Heterocycles in Life and Society

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

Heterocycles in Life and Society

An Introduction to Heterocyclic Chemistry, Biochemistry and Applications
Second Edition

by

ALEXANDER F. POZHARSKII

Soros Professor of Chemistry, Southern Federal University, Russia

ANATOLY T. SOLDATENKOV

Professor of Chemistry, Russian People's Friendship University, Russia

ALAN R. KATRITZKY

Kenan Professor of Chemistry, University of Florida, Gainesville, USA

A John Wiley & Sons, Ltd., Publication

This edition first published 2011

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Pozharskii, A. F. (Aleksandr Fedorovich)
Heterocycles in life and society / Alexander F. Pozharskii, Alan R. Katritzky, Anatoly Soldatenkov. – 2nd ed.
p. cm.
ISBN 978-0-470-71411-9 (hardback) – ISBN 978-0-470-71410-2 (paper)
1. Heterocyclic chemistry. I. Katritzky, Alan R. II. Soldatenkov, A. T. (Anatoly Timofeevich) III. Title.
QD400.P6713 2011
547′.59 – dc22

2010054024

A catalogue record for this book is available from the British Library.

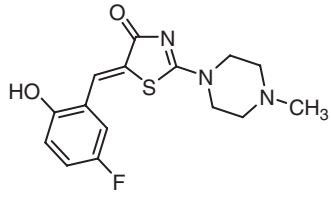
Print ISBN: 978-0-470-71411-9 (H/B) 978-0-470-71410-2 (P/B)
ePDF ISBN: 978-1-119-99838-9
oBook ISBN: 978-1-119-99837-2
ePub ISBN: 978-1-119-97013-2
eMobi ISBN: 978-1-119-97014-9

Typeset in 9/11pt Times Roman by Laserwords Private Limited, Chennai, India
Printed in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

Contents

Preface to Second English Edition

On 7 September 2009, Chemical Abstracts Service registered its 50-millionth chemical substance – a heterocyclic compound of the following structure:



Hardly a casual coincidence: heterocyclic compounds form the largest and one of the most important classes of organic compounds and some 55% of organic chemistry publications include the field. They include not only the many thousands of original articles and conference materials published annually but a great number of scientific monographs such as the multivolume *Comprehensive Heterocyclic Chemistry*, covering all fields of heterocyclic chemistry. Heterocyclic chemistry is taught worldwide at most universities and its scope is reflected in many fine text compendia and reference sources. It is therefore very strange that many general chemistry (and even organic chemistry) texts fail to include heterocycles and discuss the significance of their chemistry, or at most only in a nonsystematic manner. Furthermore, time constraints often prevent teachers of chemistry from elaborating on the manifold applications of heterocycles. This is why from the very beginning the main goal of the present book and its predecessor was to bridge this gap and to emphasize not so much the innumerable reactions of the different classes of heterocycles as their practical importance in life and society, especially their scientific applications in various branches of technology, medicine and agriculture. Our hope was, and is, that this approach will inspire the student to become involved in an immensely important and exciting field of modern chemical science and technology. The 14 years that have passed since the first edition have justified this approach. Indeed, human society, in addition to chronic old problems, now faces acute, newly recognized dangers such as climate change and ecology degradation, energy shortages, depletion of mineral resources, population growth, pandemic illnesses and so on. These challenges have forced science to become more applied and expensive but at the same time more productive and useful. This productivity results from the appearance of new powerful physical methods, apparatus as well as fundamental developments in computational techniques.

The past 10 years have been marked in biochemistry by such milestone achievements as genome decoding, clarification of ribosome structure and its activity mechanism, and wide applications of imaging techniques. Further progress has been made in medicinal chemistry where new methods of biological screening, drug delivery and drug targeting in combination with

innovative chemotherapy have been elaborated. An epochal event in science is the creation of nanotechnology which, via new materials and electronic devices, is leading to revolutionary changes in our future life. In the energy sector the growing production of biofuels, progress in development of hydrogen as a fuel, artificial photosynthesis and dye-sensitized solar cells all look very encouraging. These and other lines of development would be impossible without organic chemistry and often without heterocyclic compounds. The discussion of these themes lies at the focus of this second edition: most chapters have been substantially revised and updated, and chapter 11 is completely new.

While this book is intended for university level chemistry and biochemistry students and their instructors, it should be of interest to researchers over the whole of the chemical, biological, medical and agricultural sciences as well as in adjacent branches of science and technology. These assertions are well founded because the majority of known pharmaceutical preparations (antibiotic, neurotropic, cardiovascular, anticarcinogenic) are heterocyclic in nature; because the agricultural use of new plant development regulators and pesticides based on heterocyclic structures becomes more widespread each year; and because great attention is being paid to the synthesis and production of new kinds of thermostable polymers, highly durable fibers, fast pigments, colorants and functional dyes and of organic conductors containing heterocyclic fragments.

This book consists of 12 chapters. First, chapters (1) and (2) present the elements of the structure and properties of heterocycles and are a useful introduction to the fundamentals of their chemistry. Next, four chapters deal in a general way with the key role of heterocyclic molecules in life processes, including the transfer of hereditary information (3), the manner in which enzymes function (4), the storage and transfer of bioenergy (5) and photosynthesis (6). Chapters (7)–(9) consider the applications of heterocycles in medicine, agriculture, and industry, respectively. We have now dedicated chapter (10) to supramolecular chemistry in view of its significance. Finally, chapter (11) considers the future contribution of heterocyclic chemistry to modern trends of applied science, the latest discoveries and the prospects of finding new spheres of use for heterocycles. Chapter (12) deals with the past: specifically the emergence of heterocyclic molecules on primordial Earth, which is tightly connected with the far-reaching achievements of astrophysics. Due to modern orbital telescopes and space stations our knowledge about the origin of the Universe and its evolution has been significantly widened and deepened. On this basis new scientific disciplines are arising and strongly developing. In two of these, perhaps the most fascinating (prebiotic chemistry, synthetic biology), the role of heterocyclic compounds is especially important. In fact, a test-tube recreation of the process of molecular evolution up to synthesis of biological cells and live organisms is put forward as a not so distant perspective. It is not necessary to possess a rich imagination to foresee that the consequences of such a development of events could be even more dramatic then that of nanotechnology.

Throughout this text the student will learn to apply the knowledge gained by working on problems related to the topics covered in each chapter. Many of the 100 problems have been chosen from scientific journals and represent areas of recent significant interest. The scientists who solved these mysteries were yesterday's students. Thus, the approach to the problems will give today's students further insight into nature and a preview of what is scientifically possible. Each chapter also contains suggested further reading.

The authors have tried to organize this book in as simplified a form as possible, in as far as the scientific language is concerned. Each chapter is preceded by a piece written by a Russian poet (translated into English by E. N. Sokolyuk) or (in one case) an American poet. The selected verses may suggest subtle links with the concepts and contents of each chapter and were introduced with the hope of fruitful cross-pollination between the natural sciences and humanities, so much needed in our modern world.

In conclusion, we would like to express our warm acknowledgements to many people who helped us during the preparation of the second edition of this book. We are most grateful for helpful discussion and technical assistance from Dr Anna Gulevskaya, Dr Valery Ozeryanskii (for reading Chapter 11), Dr Vladimir Sorokin (who kindly supplied us with some fresh literature sources) and Dr John Zoltewicz.

A. F. Pozharskii
A. T. Soldatenkov
A. R. Katritzky

Preface to First English Edition

The book presents an updated translation of the Russian original '*МоЛекуЛЬЯ-ПерсТНИ*' by A. F. Pozharskii and A. T. Soldatenkov, published in 1993 by Khimiya. It has been a great pleasure to accept the invitation of my long-standing friend Sasha Pozharskii to join him and Professor Soldatenkov in producing the present English version, which follows closely the concepts and objectives of the original. We hope that this book may ignite for its readers some of the passion for heterocyclic chemistry which we the authors possess and help to repair the neglect of heterocyclic chemistry on the US academic scene. This neglect contrasts with the high importance awarded to heterocyclic chemistry and biochemistry by American industry, as well as by academic and industrial chemists alike in Europe, Japan and all over the world.

This volume could not have been produced without the help of many people. Dr Daniel Brown (Cambridge) read the whole text and made very helpful suggestions. Among many other colleagues who read parts of the work, I would like to acknowledge particularly Dr Phil Cote, Dr Alastair Monro, Dr Emil Pop, Dr Nigel Richards, Dr Eric Scriven and Dr John Zoltewicz. It is a pleasure to thank also Ms Jacqui Wells, Dr Olga Denisko and Ms Cynthia Lee for all the help they gave me in producing and finalizing the manuscript.

Alan R. Katritzky
Gainesville, Florida
April 1996

1

Molecular Rings Studded With Jewels

Fortune Goddess, in your glory, in your honor, stern Kama,
Bangles, finger-rings and bracelets I will lay before your Temple.
V. Bryusov

Readers of this book, whether or not they are students of organic chemistry, will all be aware of the vital role of proteins, fats and carbohydrates in life processes. Experience has shown that considerably less is usually known about another class of compounds which have a similar importance in the chemistry of life, namely the heterocyclic compounds or, in short, heterocycles. What are heterocycles?

1.1 From Homocycle to Heterocycle

It is rumored that the Russian scientist Beketov once compared heterocyclic molecules to jewelry rings studded with precious stones. Several carbon atoms thus make up the setting of the molecular ring, while the role of the jewel is played by an atom of another element, a heteroatom. In general, it is the heteroatom which imparts to a heterocycle its distinctive and sometimes striking properties. For example, if we change one carbon atom in cyclohexane for one nitrogen atom, we obtain a heterocyclic ring, piperidine, from a homocyclic molecule. In the same way, we can derive pyridine from benzene, or 1,2,5,6-tetrahydropyridine from cyclohexene (Figure 1.1).

A great many heterocyclic compounds are known. They differ in the size and number of their rings, in the type and number of heteroatoms, in the positions of the heteroatoms and so on. The rules of their classification help to orient us in this area.

Cyclic hydrocarbons are divided into cycloalkanes (cyclopentane, cyclohexane, etc.), cycloalkenes (e.g., cyclohexene) and aromatic hydrocarbons (with benzene as the main representative). The most basic general classification of heterocycles is similarly divided into heterocycloalkanes (e.g., piperidine), heterocycloalkenes (e.g., 1,2,5,6-tetrahydropyridine) and heteroaromatic systems (e.g., pyridine, etc.). Subsequent classification is based on the type of heteroatom. On the whole, heterocycloalkanes and heterocycloalkenes show comparatively small differences when compared with related noncyclic compounds. Thus, piperidine possesses chemical

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

properties very similar to those of aliphatic secondary amines, such as diethylamine, and 1,2,5,6-tetrahydropyridine resembles both a secondary amine and an alkene.

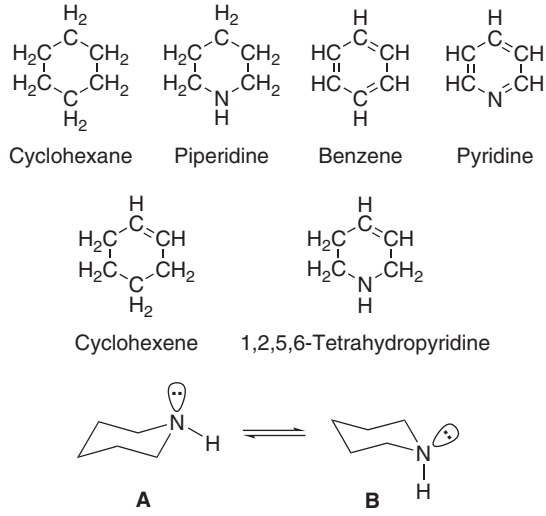


Figure 1.1 *The relationship between cyclic hydrocarbons and heterocycles and the two chair conformations of piperidine.*

An interesting feature of heterocycloalkanes and heterocycloalkenes is the possibility of their existence in several geometrically distinct nonplanar forms which can quite easily (without bond cleavage) equilibrate with each other. Such forms are called conformations. For instance, piperidine exists mainly in a pair of chair conformations in which the internal angle between any pair of bonds is close to tetrahedral (109° 28′) to minimize steric strain. In these two chair conformations (Figure 1.1), the N—H proton is in either the equatorial (A) or axial (B) position, the first being slightly preferred.

By contrast, the heteroaromatic compounds, as the most important group of heterocycles, possess highly specific features. Historically, the name ‘aromatic’ for derivatives of benzene, naphthalene and their numerous analogues came from their characteristic physical and chemical properties. Aromatic compounds differ from other groups in possessing thermodynamic stability. Thus, they are resistant to heating and tend to be oxidized and reduced with difficulty. On treatment with electrophilic, nucleophilic and radical agents, they mainly undergo substitution of hydrogen atoms rather than the addition reactions to multiple bonds which are typical for ethylene and other alkenes. Such behavior results from the peculiar electronic configuration of the aromatic ring. We consider in the next section the structure of benzene and some parent heteroaromatic molecules.

1.2 Building Heterocycles From Benzene

Each carbon atom in the benzene molecule formally participates in bond formation with its four atomic orbitals, each occupied by one electron. Three of these orbitals are hybridized and are called sp^2-orbitals. Their axes lie in the same plane and are directed from each other at an angle of 120°. These atomic orbitals overlap similar orbitals of adjacent carbon atoms or the *s*-orbitals of hydrogen

atoms, thereby forming the ring framework of six carbon–carbon bonds and six carbon–hydrogen bonds (Figure 1.2a). The molecular orbitals and bonds thus formed are called σ-orbitals and σ-bonds, respectively. The fourth electron of the carbon atom is located in an atomic *p*-orbital, which is dumbbell shaped and has an axis perpendicular to the ring plane (Figure 1.2b). If the *p*-orbitals merely overlapped in pairs, the benzene molecule would possess the cyclohexatriene structure with three single and three conjugated double bonds, as reflected in the classic representation of benzene – the Kekulé structure (Figure 1.2c). However, in reality, the benzene ring is a regular hexagon, which indicates equal overlap of each *p*-orbital with its two neighboring *p*-orbitals, resulting in the formation of a completely delocalized π-electron cloud (Figure 1.2d, e).

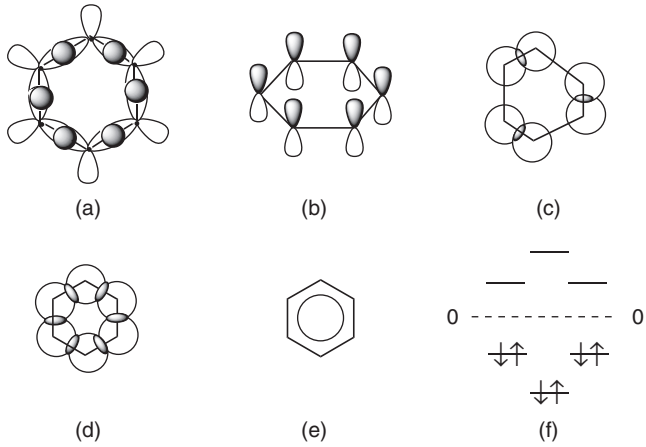


Figure 1.2 *The electronic structure of the benzene molecule: (a) framework of σ-bonds, (b) p-orbital orientation, (c) overlap of p-orbitals forming localized π-bonds (view from above), (d) overlap of p-orbitals forming delocalized π-bonds, (e) representation of the benzene ring reflecting the equivalence of all carbon-carbon bonds and the equal distribution of π-electrons, (f) energy levels of molecular π-orbitals showing electron occupation of the three orbitals of lower energy.*

Thus, in the benzene molecule as well as in the molecules of other aromatic compounds, we observe a new type of carbon–carbon bond called 'aromatic', which is intermediate in length between a single and a double bond. Standard aromatic C—C bond lengths are close to 1.40 Å, whereas the C—C distance is 1.54 Å in ethane and 1.34 Å in ethylene.

The high stability of the benzene molecule is explained by the energetic picture available from quantum mechanics. Benzene has six molecular π-orbitals. Three of these π-orbitals (bonding orbitals) lie below the nonbonding energy level and are occupied by six electrons with a large energy stabilization. The remaining three are above the nonbonding level (antibonding orbitals). Occupation of the bonding orbitals leads to the formation of strong bonds and stabilizes the molecule as a whole. Incomplete occupation of bonding orbitals, and especially the occupation of antibonding orbitals, results in considerable destabilization. Figure 1.2f shows that all three bonding orbitals in benzene are completely occupied. Hence, it is often said that benzene has a stable aromatic π-electron sextet, a concept that can be compared in its importance to the inert octet cloud of neon or the F^- anion.

In addition to the π-electron sextet, stable aromatic arrangements can also be formed by 2, 10, 14, 18 or 22 π-electrons. Such molecules contain cyclic sets of delocalized π-electrons. For example, the aromatic molecule naphthalene possesses 10 π-electrons. The number of electrons

required for a stable aromatic configuration can be calculated by the $4n + 2$ 'Hückel rule', where $n = 0, 1, 2, 3$ and so on, which was suggested by the German scientist Hückel in the early 1930s.[1]

The electronic configuration of the pyridine molecule is very similar to that of benzene (Figure 1.3a). Both compounds contain an aromatic π-electron sextet. However, the presence of the nitrogen heteroatom in the case of pyridine results in significant changes in the cyclic molecular structure. First, the nitrogen atom has five valence electrons in the outer shell, in contrast with the carbon atom which has only four. Two take part in the formation of the skeletal carbon–nitrogen σ-bonds, and a third electron is utilized in the aromatic π-cloud. The two remaining electrons are unshared, their sp^2-orbitals lying in the plane of the ring. Owing to the availability of this unshared pair of electrons, the pyridine molecule undergoes many additional reactions over and above those which are characteristic of benzene or other aromatic hydrocarbons. Second, nitrogen is a more electronegative element than carbon and therefore attracts electron density. The distribution of the π-electron cloud in the pyridine ring is thus distorted (see Chapter 2).

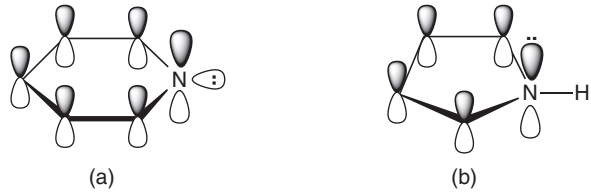


Figure 1.3 *The orientation of π-electron orbitals and unshared electron pairs in (a) pyridine and (b) pyrrole (C—H bonds are omitted).*

Heterocyclic compounds include examples containing many other heteroatoms such as phosphorus, oxygen, sulfur and so on. By substitution of a ring carbon atom we may formally transform benzene into phosphabenzene or pyrylium and thiapyrylium cations (Figure 1.4). Note that a six-membered ring which includes oxygen or another group VI element can only be aromatic if the heteroatom bears a formal positive charge (+1). Such cationic rings exist only in association with counterions like ClO_4^- or BF_4^-. Just like the nitrogen atom in pyridine, the phosphorus, oxygen and sulfur atoms donate one π-electron to the aromatic electron cloud. Such heteroatoms are often called 'pyridine-like'.

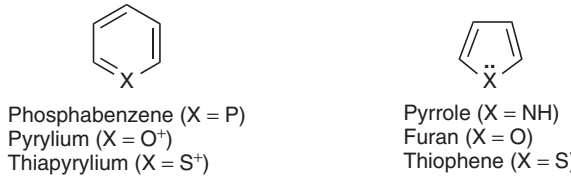


Figure 1.4 *Examples of heterocycles with pyridine-like and pyrrole-like heteroatoms.*

Formally, pentagonal aromatic heterocycles can also be derived from benzene by a heteroatom taking the place of one complete CH═CH group. Two electrons of the heteroatom *p*-orbital must

[1] For monocyclic fully conjugated compounds, the Hückel rule stops working with 26 and larger π-electron systems ($n \geq 6$). This is explained by a strong increase of inter-electron repulsion that outweighs the gain of aromatic stabilization.

now be involved in the π-system in order to obtain an aromatic sextet (Figure 1.3b). This type of heteroatom is called 'pyrrole-like' in contrast to the 'pyridine-like' nitrogen which donates only one electron to the sextet. The corresponding five-membered heterocycles containing nitrogen, oxygen or sulfur atoms are named pyrrole, furan and thiophene, respectively (Figure 1.4). One more difference between a pyridine-like heteroatom and a pyrrole-like heteroatom is obvious: the first participates with one double bond in the Kekulé structure, while the second is involved with single bonds only.

A heterocycle can contain several heteroatoms. Pyridazine, pyrimidine, pyrazine and 1,3,5-triazine are heterocyclic compounds with a single ring but two or three identical heteroatoms (Figure 1.5a). Together with pyridine and many other analogues they form the family of *azines*.

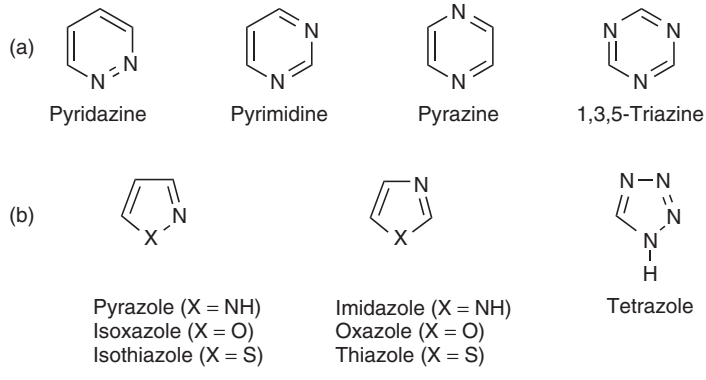


Figure 1.5 *Heterocycles of (a) the azine class and (b) the azole class.*

Five-membered heterocyclic compounds containing both pyridine-like and pyrrole-like nitrogen or other heteroatoms are called *azoles*. Pyrazole, imidazole and their oxygen and sulfur analogues belong to the azole series (Figure 1.5b).

Two or more rings are encountered in many heterocyclic compounds. The rings may be connected to each other by a single bond (as in the case of 2,2′-bipyridyl) or may be fused as shown in Figure 1.6 to form condensed systems. For example, two fused rings exist in quinoline, pteridine, indole and benzimidazole and three fused rings in acridine. In some cases a heteroatom may belong simultaneously to two (e.g., indolizine) or even three rings. Such a heteroatom is denoted a 'bridgehead' atom.

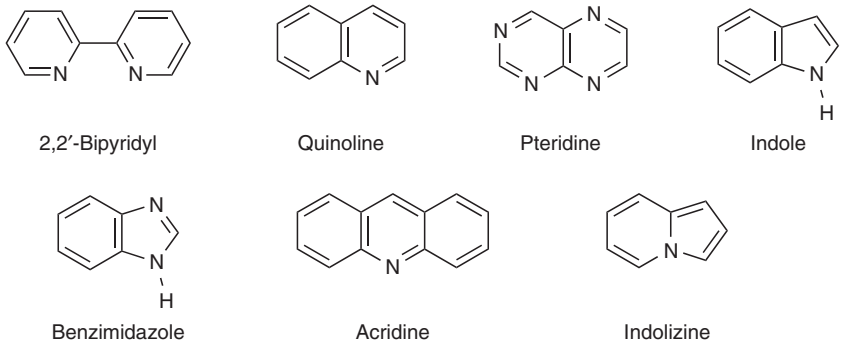


Figure 1.6 *Examples of bi- and polycyclic heterocycles.*

1.3 Some More Kinds of Heterocycles

The comparison of heterocycles with jewel-studded rings is most appropriate for five- and six-membered systems which are frequently natural products and which have become commonplace in many research laboratories. However, polymembered cycles or macrocycles have recently drawn much attention. They resemble not so much finger-rings but rather molecular bracelets or bangles. For example, aza[18]annulene is an 18-membered analogue of pyridine, and aza[17]annulene is a 17-membered analogue of pyrrole (Figure 1.7a). We focus our attention on macrocycles in subsequent chapters, especially Chapter 10.

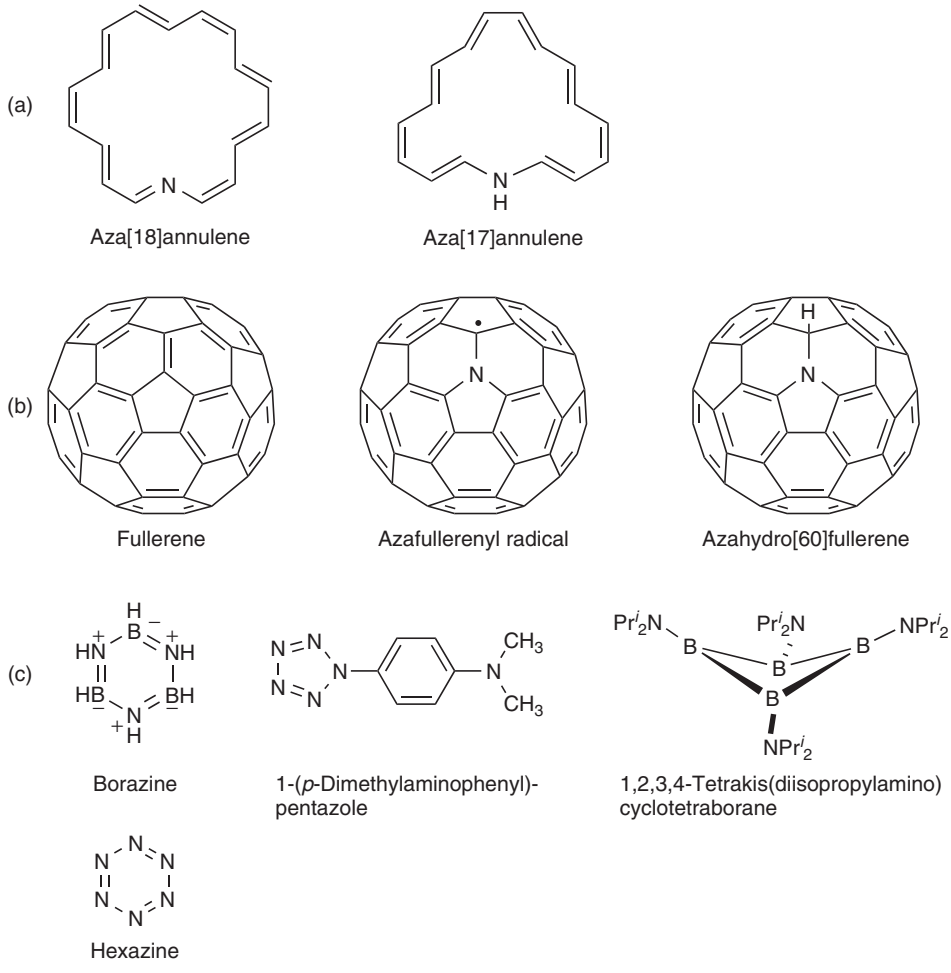


Figure 1.7 *Examples of (a) macroheterocycles, (b) azafullerenes and (c) rings without cyclic carbon atoms.*

Another recently arisen area is the chemistry of heterofullerenes – compounds in which one or more cage carbon atoms are substituted by heteroatoms. The most stable among them are

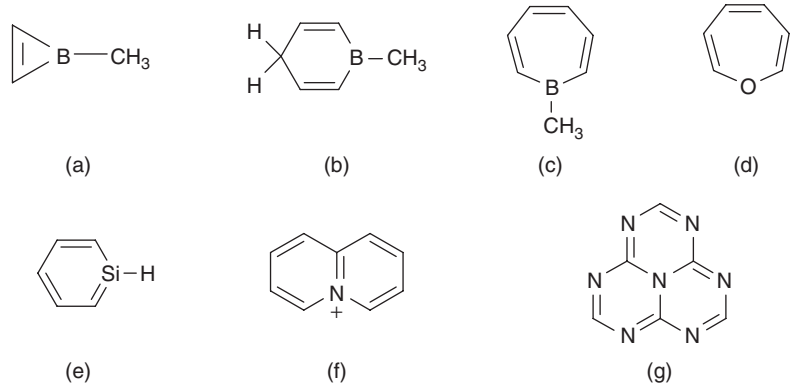
azafullerenes. The valence rules determine that, at the introduction of one nitrogen atom into the fullerene molecule C_{60}, the free radical specie $C_{59}N^{\bullet}$ should be produced. Its stabilization can be achieved either via dimerization into 2,2′-biaza[60]fullerene $(C_{59}N)_2$ or by means of hydrogen atom addition leading to green azahydro[60]fullerene $C_{59}NH$ (Figure 1.7b). Carbon nanotubes containing nitrogen or boron heteroaatoms are also known.

How many heteroatoms may be included in one ring? As many as one can imagine. A ring may, in principle, be completely constructed from noncarbon atoms (Figure 1.7c). Borazine, a well known example of such a compound, was designated 'inorganic benzene' because of its high stability. 1-(*p*-Dimethylaminophenyl)pentazole and blue-colored 1,2,3,4-tetrakis (diisopropylamino)cyclotetraborane contain five- and four-membered heterocycles composed only of nitrogen or boron atoms. The curiosity of many chemists has long been excited by a theoretical substance named 'hexazabenzene' or 'hexazine'. Numerous attempts to prepare this compound have so far ended in failure, supposedly because of its great instability and tendency to decompose to give nitrogen: $N_6 \rightarrow 3N_2$.

Of course, the examples given above by far do not cover all of the heterocyclic systems possible. In the following chapters we will become acquainted with many new ones.

1.4 Problems

1. How many chair conformations are possible for unsubstituted piperidine? How many for a 1,4-disubstituted piperidine? Draw their structures.
2. The boat conformation for saturated six-membered rings is energetically unfavorable. Account for this fact. Design the structure of a substituted piperidine in which the boat conformation is fixed.
3. Phosphacyclohexane (phosphorinane) exists almost completely in a chair conformation with the P—H bond axial. Discuss possible reasons for the stabilization of this conformation compared with the analogous piperidine conformation.
4. Indicate which of the heterocycles listed below can be formally regarded as aromatic. Explain your choices.



5. Historically, the first synthetic homocyclic aromatic system not containing carbon atoms was the golden-orange salt P_5^- Na^+. Draw its structure and explain the following facts: (i) the salt is stable only in tetrahydrofuran solution in the presence of 18-crown-6 (see Section 10.1.1), (ii) all phosphorus atoms in the anion P_5^- in solution are equivalent.
6. Draw all of the possible isomeric imidazopyridines, that is, the heterocycles which consist of fused pyridine and imidazole nuclei.
7. What is the orientation of the nitrogen lone pair of electrons in aza[18]annulene (Figure 1.7)? Is any alternative orientation possible? Discuss the orientation of the N—H bond in aza[17]annulene.
8. The relative stability (aromaticity) of five-membered heterocycles is changed in the following sequence: thiophene> pyrrole> furan. How this can be explained?
9. To avoid the formation of a free radical by placing one nitrogen atom into fullerene, one can simultaneously introduce into the molecule two heteroatoms. Draw the simplest structures of such a type.

1.5 Suggested Reading

1. Joule, J. A., *Heterocyclic Chemistry*, 4th edn, Wiley-Blackwell, Chichester, 2000.
2. Gilchrist, T. L., *Heterocyclic Chemistry*, 3rd edn, Pearson Education Press, London, 1997.

3. Katritzky, A. R., Ramsden, C. A., Joule, J. A. and Zhdankin, V. V., *Handbook of Heterocyclic Chemistry*, 3rd edn, Elsevier, 2010.
4. Katritzky, A. R., Ramsden C., Scriven E. and Taylor, R. (eds), *Comprehensive Heterocyclic Chemistry III*, vols 1–15, Elsevier, New York, 2008.
5. Katritzky, A. R., Rees, C. W. and Scriven E. (eds), *Comprehensive Heterocyclic Chemistry II*, vols 1–12, Pergamon Press, Oxford, 1995.
6. Katritzky, A. R. and Rees, C. W. (eds), *Comprehensive Heterocyclic Chemistry*, vols 1–8, Pergamon Press, Oxford, 1984.
7. Elguero, J., Marzin, C., Katritzky, A. R. and Linda, P. The tautomerism of heterocycles, in *Advances in Heterocyclic Chemistry*, *Suppl. 1*, Academic Press, New York, 1976.
8. Pozharskii, A. F., *Theoretical Basis of Heterocyclic Chemistry* (in Russian), Khimia, Moscow, 1985.
9. Katritzky, A. R. (ed.) Special issue on heterocyclic chemistry, *Chem. Rev*., 2004, **104** (5).

2

Why Nature Prefers Heterocycles

Ties subtle, full of power exist
Between the shape and flavor of a flower.
So is a brilliant unseen, until comes hour
To facet it from diamond mist.

V. Bryusov

All biological processes are chemical in nature. Such fundamental manifestations of life as the provision of energy, transmission of nerve impulses, sight, metabolism and the transfer of hereditary information are all based on chemical reactions involving the participation of many heterocyclic compounds. Why does nature utilize heterocycles? To answer this question we first describe the basic physical and physicochemical properties of the fundamental heterocyclic types.

2.1 Reactions for all Tastes

Heterocycles are involved in an extraordinarily wide range of reaction types. Depending on the pH of the medium, they may form anions or cations. Some interact readily with electrophilic reagents, others with nucleophiles and yet others with both. Some are easily oxidized, but resist reduction, while others can be readily hydrogenated but are stable toward the action of oxidizing agents. Certain amphoteric heterocyclic systems simultaneously demonstrate all of the above-mentioned properties. The ability of many heterocycles to produce stable complexes with metal ions has great biochemical significance. Such versatile reactivity is linked to the electronic distributions in heterocyclic molecules. Let us consider pyridine.

We have already seen that the nitrogen atom in pyridine induces π-electron withdrawal from the carbon atoms. As a result of this electronic shift, the carbon atoms in the *ortho* and *para* positions (relative to the nitrogen atom) acquire a partial positive charge (Figure 2.1). Thus, a π-electron deficit on the carbon skeleton is characteristic of all heterocycles containing pyridine-like heteroatoms. Such heterocycles are called π-deficient.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

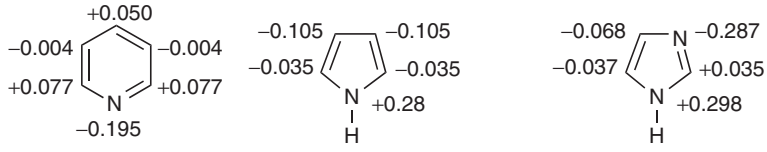
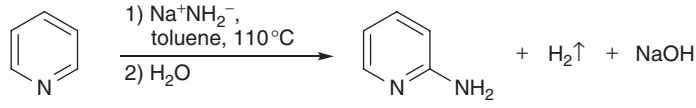
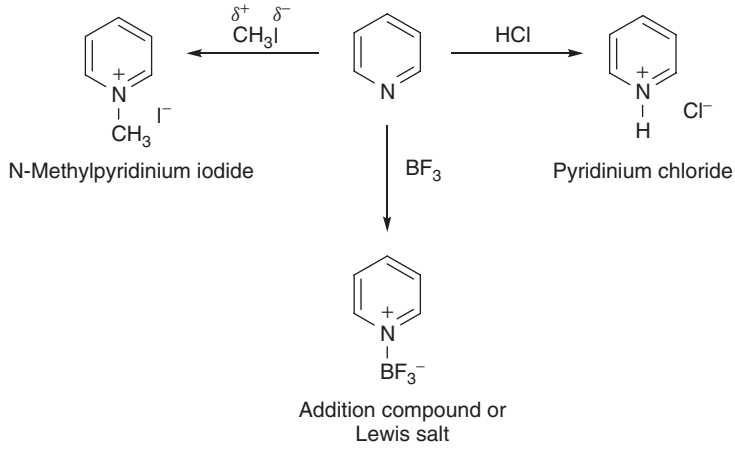


Figure 2.1 *The π-electron charges in pyridine, pyrrole and imidazole.*

A unique feature of π-deficient heterocycles is their facile interaction with negatively charged nucleophilic reagents. As a typical example, the reaction of pyridine with sodamide gives 2-aminopyridine in good yield:



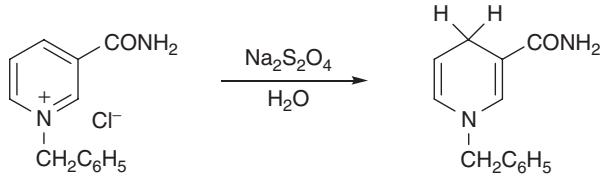
Substitution of the hydrogen atom under the action of positively charged (electrophilic) agents proceeds with difficulty or does not occur at all in π-deficient heterocycles. However, electrophiles add readily to the pyridine nitrogen owing to its unshared pair of electrons. Pyridine thus forms pyridinium and *N*-alkylpyridinium salts with acids and alkyl halides, respectively, and a zwitterionic addition compound or Lewis salt with BF_3:



Pyridine and other heterocycles containing a pyridine-like nitrogen atom behave as bases in these and similar reactions (see Section 2.2).

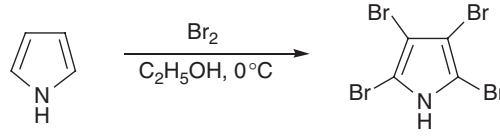
The introduction of electron-accepting groups into an organic compound lowers the energy of all molecular orbitals. Hence, such compounds donate electrons with difficulty and are thus poorly oxidized. By contrast, their ability to accept additional electrons enables such compounds to be readily

reduced. Pyridine-like heteroatoms are electron acceptors, and hence π-deficient heterocycles are reduced with ease. This is found to be the case, especially in relation to compounds which have a positively charged heteroatom, like salts of pyrylium, pyridinium and so on. For example, 1-benzyl-3-carbamoylpyridinium chloride is reduced by sodium dithionite to the corresponding 1,4-dihydropyridine derivative:

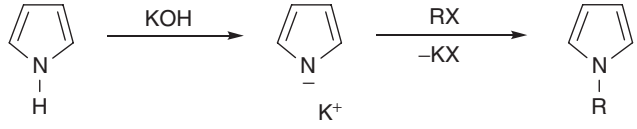


We shall see elsewhere (Sections 4.2.1 and 5.2) that nature uses this apparently simple reaction to drive a great many biologically important processes.

Quite a different situation is encountered in the case of pyrrole, furan and thiophene. Since the heteroatoms of these compounds each contribute two electrons to the π-aromatic ensemble, the cyclic system of five atoms formally has six π-electrons. As a result, in spite of the higher intrinsic electronegativity of the heteroatom, all of the carbon atoms possess excess negative charge (Figure 2.1). Such compounds are named π-excessive heterocycles. Reactions with nucleophiles agents are not common but they readily interact with electrophiles. Thus, pyrrole is almost instantly halogenated even under very mild conditions to give the tetrahalogenopyrrole, and these reactions cannot be stopped at the monosubstitution stage:



Two-electron donation to the aromatic system by the pyrrole-like heteroatom imparts a partial positive charge to the heteroatom (Figure 2.1). In the case of pyrrole and related NH-heterocycles, the N—H bond reactivity increases. N-Anions, which are readily alkylated, acylated and arylated, are thus formed under the action of bases. Such reactions are commonly used for the synthesis of various N-derivatives (note that a nonionized NH group does not, as a rule, undergo these conversions):



The molecular orbitals in π-excessive heterocycles are of high energy, and consequently these compounds are reduced with difficulty but are readily oxidized. Compounds with both pyridine-like and pyrrole-like heteroatoms, as expected, can show both π-deficient and π-excessive properties with one or the other dominant. Thus, imidazole contains two carbon atoms with a partial negative and a third with a partial positive π-charge (Figure 2.1). Its high reactivity towards halogenation is attributed to the dominant π-excessive character of the neutral molecule and especially of the imidazole anion.

2.2 Heterocycles as Acids and Bases

In the preceding section we noted the capability of nitrogen heterocycles to behave as acids or bases, the acidic properties being inherent to heterocyclic compounds containing a pyrrole-like NH group, whereas the basic properties are characteristic for those with pyridine-like nitrogen. We describe this in more detail because acid–base properties play a vital role not only in general reactivity but in many biochemical processes as well.

The acid dissociation constant (K_a) is universally used as the quantitative measure of acidity. Dissociation constants are obtained by application of the law of mass action to the acid–base equilibrium:

$$H{-}A \leftrightarrows H^+ + A^-$$

The dissociation constant K_a is equal to the anion concentration multiplied by the proton concentration, divided by the concentration of the nondissociated acid:

$$K_a = [A^-][H^+]/[HA] \tag{2.1}$$

In practice, following the analogous use of pH, it is more convenient to use the negative logarithm of K_a, the so-called acidity index pK_a:

$$pK_a = -\log K_a = -\log[A^-] - \log[H^+] + \log[HA]$$

as the value of $-\log[H^+] = pH$, then:

$$pK_a = \log\{[HA]/[A^-]\} + pH \tag{2.2}$$

It is clear from Equation (2.2) that the value of the pK_a is equal to the value of the pH when the nondissociated acid (HA) content and the anion (A^-) content are equal, that is, when the degree of dissociation is 50%.

We see that the stronger the acid, the greater the numerator and, consequently, the larger the K_a value; a larger K_a value corresponds to a smaller pK_a. Vice versa, in a series of compounds, the pK_a increases as the acidity decreases. It should be emphasized that pK_a values, which are essentially acid ionization constants, are also employed for the measurement of basicity. As a consequence of the reversibility of the dissociation process, any acid which donates its proton is thus converted to the conjugate base; similarly, a base which accepts a proton becomes the conjugate acid. Stronger acids obviously correspond to weaker conjugate bases and vice versa. Thus, for bases, the order of the pK_a changes in the opposite sense: the larger the pK_a of the conjugate acid, the stronger the base, and the weaker bases have correspondingly lower pK_a values.

The acid dissociations of pyrrole and imidazole (Figure 2.2a) are used as an example. The corresponding pK_a values are 17.5 and 14.2, respectively.[1] As pK_a is a logarithmic scale, pyrrole is a weaker acid than imidazole by a factor of $10^{3.3}$ (i.e., by a factor of 2000). This also indicates that the pyrrole anion is a stronger base than the imidazole anion by the same factor.

Whereas both pyrrole and imidazole are very weak acids, some heterocycles have pK_a values close to those of conventional acids. Tetrazole (Figure 1.5) has a pK_a of 4.89, almost equal to that of acetic acid (pK_a 4.76).

Under ordinary conditions a neutral pyrrole-like nitrogen is unlikely to add a proton because of the tendency to preserve the aromaticity of the heterocycle. In contrast, the lone electron pair of

[1] Standardized conditions must be used for the determination of ionization constants as the latter depend on solvent and temperature. The pK_a values given here were determined in aqueous solutions at 20 °C.

a pyridine nitrogen does not participate in the formation of the aromatic sextet and readily adds a proton to form a heteroaromatic cation. Thus, pyridine has a pK_a of 5.23. This value formally reflects the acidity of the pyridinium ion (Figure 2.2b), but is more often used to assess the basicity of pyridine. It can be seen that the proton of the pyridinium cation is 12 orders of magnitude more acidic than the NH of pyrrole. This is readily explained by the facile loss of a proton from the positively charged nitrogen atom in the pyridinium cation.

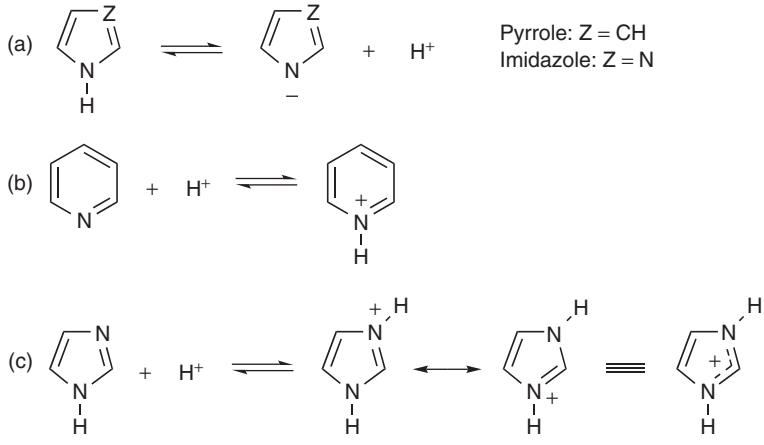


Figure 2.2 *Acid–base equilibria for pyrrole (a), imidazole (a, c) and pyridine (b).*

Obviously, heterocycles such as imidazole have amphoteric properties: imidazole is both an NH acid and a strong neutral base with a pK_a of 6.95. The imidazole ring system is frequently encountered in proteins (see Section 4.1) and is one of the strongest of all bases found in biological systems. The imidazole unit, therefore, plays an active role in proton transfer processes and the various catalytic events accompanying them. The enhanced basicity of imidazole is due to electron donation from the pyrrole nitrogen, thus favoring proton addition. The stabilized imidazolium ion can be represented by two equivalent resonance structures in which the positive charge is isolated on one nitrogen atom in the first representation and on the other in the second, or by an average structure with delocalized charge (Figure 2.2c). Section 10.1.5 contains some additional information on the basicity of nitrogen heterocycles.

2.3 Heterocycles and Metals

It is well known that minute quantities of different metals are necessary for the normal development of all living organisms. In addition to the widespread sodium, potassium, magnesium, calcium, iron and zinc, the group of 'essential metals' also includes more exotic members such as molybdenum, cobalt, chromium and others. Metals exist in organisms in the form of cations linked with various basic ligands by coordination bonds. The basic functionality may involve the amino, hydroxy or thiol groups of amino acids as well as nitrogen heterocycles (azines, azoles). The ability to form stable metallic complexes seems to be 'preprogrammed' into the structure of the heterocycles.

The fixed and outwardly directed unshared pair of electrons of a pyridine nitrogen atom is available for coordination with practically all metal ions. Thus, pyridine gives complexes of various

types: a linear arrangement with the silver ion, a tetrahedral structure with aluminum chloride, a square planar coordination compound with copper(II) chloride and a dianionic complex with cobalt(II) chloride, as shown in Figure 2.3.

The formation of coordination compounds is very similar to the production of pyridinium salts by protonation or alkylation (see Section 2.1), although some peculiarities in the electron shell configurations of some metals diversify the spatial structures of the complexes. It should be noted that the oxidation–reduction potentials and other properties of metal ions can be markedly changed by coordination. Such changes can, in turn, significantly affect the functioning of biological systems.

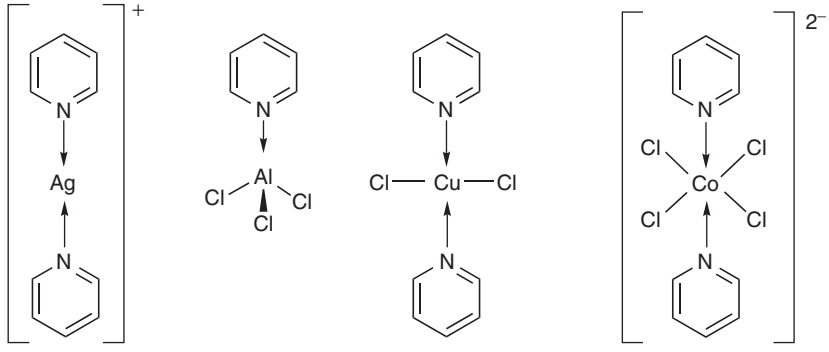


Figure 2.3 *Pyridine as a ligand in complexes.*

The number of ligands coordinated to a metal ion depends on the type and number of unfilled orbitals in the outer electron shell. For example, in order to complete its outer shell of eight electrons, aluminium(III) requires one additional pair of electrons. This electron pair can be donated by many bases, such as pyridine. The four valence bonds formed by aluminium in the $C_5H_5N{:}AlCl_3$ ensemble are formally built utilizing one 3*s*-orbital and three 3*p*-orbitals. Quantum mechanics ascertains that such bonding is achieved from the more energetically favorable mixed (hybrid) sp^3-orbitals. The best arrangement of four sp^3-orbitals in space is achieved when their axes are directed toward the corners of a regular tetrahedron. This provides for minimal interelectronic repulsion and results in a tetrahedral configuration of the complex.

A pyridine molecule can donate only one electron pair for coordination with a metal ion. Such a ligand is described as monodentate. Imidazole seems to be the most important heterocyclic monodentate ligand in biochemical processes (see Section 4.2). Polydentate ligands are able to provide several electron pairs and are highly effective. 2,2′-Bipyridyl, an example of a bidentate heterocyclic ligand (Figure 1.6), forms stable complexes with many metals, in particular with iron(II) (see Section 9.8). Strong binding here results from the so-called *chelate effect*. However, even more important is the *macrocyclic effect* found when donor centers are included in a favorable arrangement into the macrocycle with axes of their electron pairs directed into the center of the cavity. Tetra- and hexadentate ligands of such type are especially widespread. Commonly, stability constants for metal complexes of heteromacrocycles are enhanced by 4–5 powers of ten in comparison with related acyclic ligands (Figures 2.4, 2.5).[2]

The porphyrin system is a very important natural tetradentate macrocyclic ligand composed of four pyrrole rings which are linked to each other via carbon bridges (Figure 2.5). Two of the

[2] See Section 10.1 for more details on the macrocyclic effect and stability constants of metal complexes.

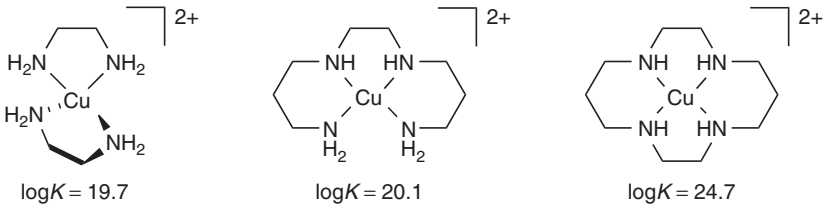


Figure 2.4 *Relative stabilities of copper complexes of polydentate nitrogen ligands.*

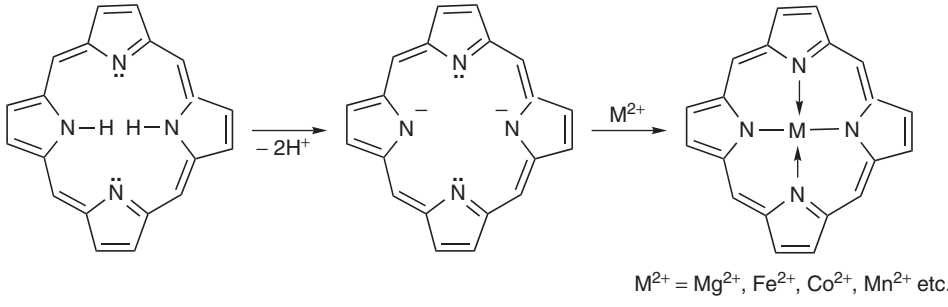


Figure 2.5 *The porphyrin molecule, its dianion and complexes with metals.*

pyrrole rings in porphyrin are in the oxidized state: their nitrogen atoms are of the pyridine type, with the unshared electron pairs oriented toward the inside of the macrocycle. If both the N—H bonds in the porphyrin molecule are ionized, a highly symmetrical dianion is formed in which all four nitrogens become equal because of delocalization of the negative charges. In this dianion all four unshared pairs of electrons are directed toward the inside of the macrocyclic cavity. The ionic radii of many metals allow them to fit within this cavity and the metal ions can be fixed in space by coordination bonds with the four porphyrin nitrogen atoms. Such complexes have considerable stability and are deeply colored. A porphyrin system which includes magnesium is part of the green plant pigment chlorophyll (see Section 6.1). A porphyrin system containing an entrapped iron(II) ion is of primary importance in respiratory and metabolic processes as it is a constituent of the red pigment of blood, hemoglobin (see Section 4.2.2). Another kind of biologically important tetrapyrrole ligand that is closely related to porphyrins is the corrin system. Its complex with the cobalt(II) ion is a structural fragment of vitamin B_{12} (see Section 4.3).

The discovery in the early 1950s of ferrocene (Figure 2.6a), the first aromatic metallic π-complex with a 'sandwich' structure, prompted chemists to investigate heterocyclic analogues. For a long time, studies were unsuccessful because the metal ions tended to coordinate more readily with the heteroatom than with the π-system. Later, ferrocene heteroanalogues such as the tetramethylpyrrole anion complex of iron(II), dipyridinechromium, pyridine–benzenechromium (Figure 2.6b–d) and a number of other sandwich-like compounds were prepared.

2.4 'There are Subtle Ties of Power . . .'

The reactions described above are accompanied by the cleavage and formation of covalent bonds of different polarity. These bonds are known to be particularly strong, their energies being between

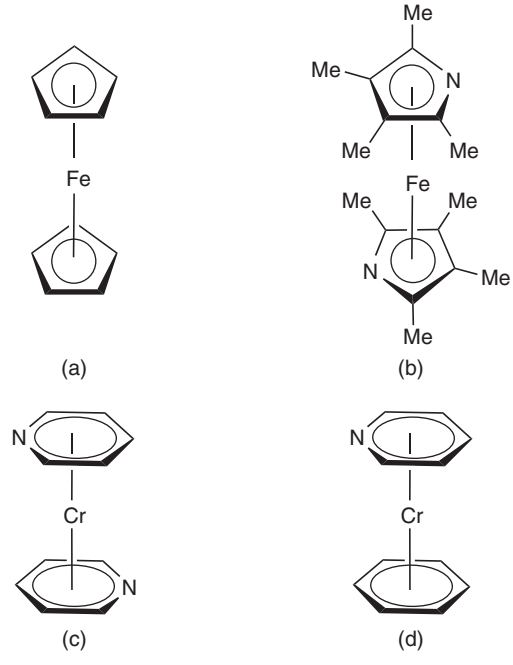


Figure 2.6 *Ferrocene (a) and some heterocyclic analogues (b–d).*

60 and 100 kcal mol^{-1}. By the sharing of electrons and the formation of covalent bonds, stable molecules are produced from which the living organism constructs numerous structures (e.g., the cell membrane). Moreover, in living systems the covalent bond serves as a reservoir of energy to be released when needed (see Chapter 5). From a biological point of view, stable covalent bonding has both advantages and disadvantages. Since such bonds are difficult to break, they cannot always provide the necessary flexibility and mobility required by living systems. For example, many biochemical reactions are reversible, the same molecule being able to react thousands of times as a result of regeneration. Thus, enzymes act as biological catalysts and hemoglobin as an oxygen carrier. These (and certain other compounds such as metal ion delivery systems) are able to change their spatial structures in a rapid and reversible manner. This is only possible when the bonds involved are much weaker than covalent linkages – such are the van der Waals–London forces, hydrogen bonding, ion pairing, dipole–dipole and other kinds of electrostatic interactions, hydrophobic effect, π-donor-acceptor interactions and so on. All are generally classified as *noncovalent interactions*.

The energies of noncovalent interactions are normally of 0.5–10 kcal mol^{-1}, only rarely overstepping these limits. It should be noted that noncovalent interactions are ubiquitous and occur not only in biologically significant molecules. They emerged in the universe at the same time as the appearance of atoms and molecules. For a beginning, let us first consider the van der Waals–London forces.

2.4.1 The van der Waals-London Interactions

It is well known that two atoms of helium cannot share their electrons to form a covalently bonded molecule, However, at very low temperatures (−269 °C) gaseous helium becomes liquefiable

evidencing some kind of weak attractions between the atoms. The nature of the attractive forces was disclosed by London at the end of 1920s. He found that movement of electrons in neighboring atoms is effectively synchronized (Figure 2.7), If at any time, say τ_1, the electron pair of one helium atom moves on the right, the electron pair of the other to minimize electron repulsion does the same and also moves to the right. Similarly, in the next moment, τ_2, electron pairs of both atoms synchronically change their positions moving to the left. Thus, each helium atom can be considered as a permanently fluctuating microdipole. The attraction of such microdipoles actually represents induced dipole–induced dipole interactions. This mechanism is also called London's or dispersion forces.

The strengths of such attractions are very small, from 0.2 to 0.5 kcal mol^{-1}, but when there are many such interactions they can substantially affect the properties of a compound. This is the reason why methane, being molecular, liquefies at −161 °C and *n*-hexane is a liquid under ordinary conditions (boiling point 69 °C), while *n*-octadecane is a solid with a relatively low melting point of 28 °C. Note that nonpolar alkane molecules have no other possible attractions except such dispersion forces.

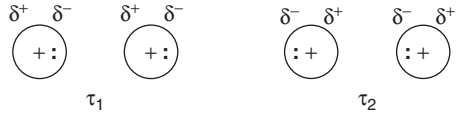


Figure 2.7 *Attraction between two helium atoms as mutually induced microdipoles.*

The most remarkable property of dispersion forces is an absence of saturation. Each atom can interact in this way with many other neighboring atoms. Since the energy of London's forces is very small and follows the law $1/r^6$, where r is the distance between the species, they should be as near to each other as possible to provide substantial attraction. However, there is a factor which resists such extra-proximity. This is the *van der Waals radius* of atoms, which outlines an area around each nucleus where practically all the atomic electron cloud can be found. When the distance between two atoms is small but larger than the sum of their *van der Waals radii*, they attract each other but when it becomes less the strong repulsive interaction between their filled electron shells rapidly arises. Such repulsion is usually called the exchange interaction or the van der Waals forces. The dispersion forces and the van der Waals forces are actually two sides of the same medal.

Typically, the van der Waals radii of atoms are larger than their covalent radii by 0.8 Å (Table 2.1). Due to the existence of the van der Waals radii, each type of atom and therefore each molecule has its specific size and shape. These parameters are essential to understand many of the properties of organic compounds, including molecular structure, steric interactions, biological properties and even reactivity. All types of molecular models which are widely used for scientific and teaching purposes are also based on van der Waals radii.

It is clear that the van der Waals–London forces are inherent to all atoms and molecules. However, there exist certain types of nonbonding interactions which are found only in compounds with certain structural characteristics. It is these specific interactions which make an essential contribution to the chemistry of living systems.

2.4.2 Hydrogen Bonding

Hydrogen atoms bound to electronegative elements such as nitrogen, oxygen or fluorine acquire a positive charge because of the high polarity of the corresponding bond. As a result of this

Table 2.1 *Selected van der Waals radii of some atoms and groups (McClinton, M. A. and McClinton, D. A.,* Tetrahedron, *1992,* ***48****, 6555)*

Atom	Radius (Å)	Atom	Radius (Å)	Group	Radius (Å)
H	1.2	F	1.4	CH_3	2.0
N	1.6	Cl	1.8	CF_3	2.7
O	1.5	Br	1.9	CCl_3	3.5
S	1.9	I	2.1	C_6H_6	1.70[a]

[a]Half-thickness of a benzene ring.

charge, and because the hydrogen atom is small, the hydrogen atom can approach other atoms and interact electrostatically with their unshared electron pairs. If the attractive force is high enough, a hydrogen atom nucleus (i.e., as a proton) can reversibly transfer to form a covalent bond with another atom. This exchange is the foundation of acid–base interactions:

$$AH + :B \rightleftharpoons A^- + HB^+$$

If, however, the attraction between the proton-donating AH group and the proton-accepting base :B is not strong enough to enable proton transfer, then a hydrogen bridge AH···B can still arise.

A classic example of hydrogen bonding is found in water, the molecules of which are aggregated in linear and three-dimensional structures. Although the strength of one hydrogen bond is not great (2–8 kcal mol^{-1}), the overall consequence of a multitude of such bonds is significant. The numerous unique properties of water (low volatility, moderate viscosity and density, specific density of ice, etc.) which allowed life on Earth to become possible are the result of hydrogen bonding.

Hydrogen bonding is certainly the most important type of noncovalent interaction between biomolecules. The polypeptide helical chains of proteins and the double-helical structure of DNA are stabilized by such interactions (see Chapter 3). The ability to form hydrogen bonds is inherent in practically all nitrogen heterocycles. Some (pyridine, other azines) are proton acceptors, others (pyrrole, indole) are proton donors, and a third group of compounds includes both proton-donating and proton-accepting functionalities (imidazole, pyrazole). Imidazole, for example, forms rather stable linear associations, whereas pyrazole is inclined to give dimers because of the specific orientation of its NH group and pyridine-like nitrogen, as can be seen in Figure 2.8a, b.

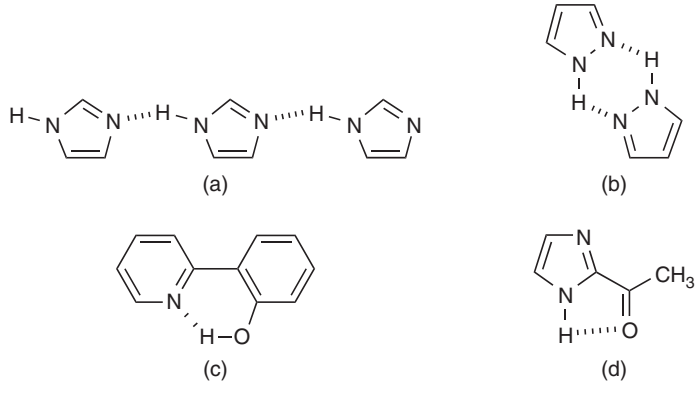


Figure 2.8 *Intermolecular and intramolecular hydrogen bonding: (a) imidazole association, (b) dimer of pyrazole, (c, d) hydrogen bonding in 2-(o-hydroxyphenyl)pyridine and 2-acetylimidazole.*

Certain heterocyclic compounds with suitably oriented functional groups can form intramolecular bonds. Such is the case for 2-(*o*-hydroxyphenyl)pyridine and 2-acetylimidazole shown in Figure 2.8c, d. Intramolecular hydrogen bonding is much stronger when it involves the construction of a six-membered ring.

2.4.3 Electrostatic Interactions

The electrostatic attraction or repulsion of charged particles is a type of noncovalent interaction as widespread as hydrogen bonding. At one time, it was thought that electrostatic interactions were characteristic only of ions. In reality, many neutral molecules engage in similar interactions especially when their electron clouds are polarized. Such molecules behave as if composed of two oppositely charged poles. These dipolar molecules can attract each other or ions (Figure 2.9).

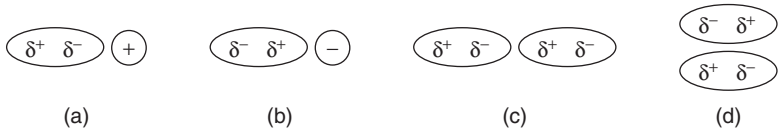


Figure 2.9 *Noncovalent interactions: (a, b) ion–dipole, (c, d) dipole–dipole.*

Almost all heterocyclic molecules are dipoles in addition to being capable of ion formation. Electrostatic interactions exert a marked influence upon heterocyclic behavior. For example, pyridine, pyrrole and 1-methylimidazole have molecular weights close to that of benzene. However, benzene demonstrates much greater volatility: it boils at 80 °C, while the heterocycles mentioned have boiling points of 115, 130 and 196 °C, respectively. Heterocyclic molecules are, to a significant degree, polar[3] and are subject to dipole–dipole associations (Figure 2.10a–c). In order to transform these associates into the vapor state, considerable energy is obviously needed, causing a decrease in the volatility of the compound. Electrostatic interactions play an essential role in biology where they participate, in particular, in optimizing the spatial arrangements of complex biomolecules. Such three-dimensional configurations can endow highly specialized biological activity. For example, the imidazole ring, found in many enzymes, exists at physiological pH to the extent of about 50% as the positively charged imidazolium ion. It is clear that the negatively charged ionized carboxylate group at the end of the protein chain will be attracted to the imidazolium cation and thus induce the relevant portion of the macromolecule to form a coil, as shown in Figure 2.10d. By contrast, if an ammonium ion is present in the chain, the repulsive force between this ion and the imidazolium ion prevents the protein chain from coiling (Figure 2.10e).

2.4.4 Molecular Complexes

In many chemical reactions, the cleavage of an existing bond and the formation of a new one are preceded by electron transfer between the molecules of the reacting compounds. As a rule, an electron is transferred from the highest occupied molecular orbital (HOMO) of the donor to the lowest unoccupied molecular orbital (LUMO) of the acceptor. As a result of the transfer the donor becomes a cation-radical, and the acceptor is converted into an anion-radical. Both particles are

[3] The polarity of a molecule may be estimated from the electronic charges of the different atoms (Figure 2.1). More definitive proof of polarity can be obtained from dipole moment values calculated as the magnitude of the distance between the centers of positive and negative charges, multiplied by the average charge. Numerical values of dipole moment are expressed in debye units (D). The higher the polarity of a molecule, the higher its dipole moment. The dipole moments of pyridine, pyrrole and 1-methylimidazole are 2.2, 1.8 and 3.8 D, respectively.

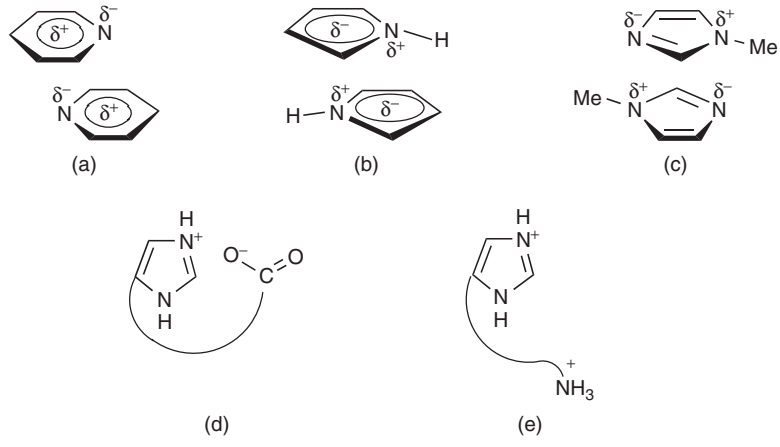


Figure 2.10 *Electrostatic interactions of heterocyclic molecules: (a–c) dipole–dipole associations, (d) attraction, (e) repulsion between charged groups of a protein chain.*

ions as they acquire charge by contributing or accepting an electron. At the same time they can be considered as radicals, for they have an odd number of electrons. Like all oppositely charged ions these cation-radicals and anion-radicals are attracted to each other and form ion-radical pairs:

$$D: + A \rightarrow D^{+\bullet}A^{-\bullet}$$

In principle, any molecule can act as donor or acceptor because it has both HOMO and LUMO, which are also called frontier orbitals. In practice, however, most compounds display a tendency toward either donating or accepting electrons. Polycyclic aromatic hydrocarbons (anthracene, benzpyrene, etc.), aromatic amines, phenols, thiophenols and other compounds with accessible unshared pairs of electrons (alcohols, esters, ketones, tertiary amines, sulfides, etc.) are typical donors. Almost all π-excessive heterocycles, particularly polynuclear compounds such as indole, carbazole, phenothiazine and others (Figure 2.11), possess good donor properties.

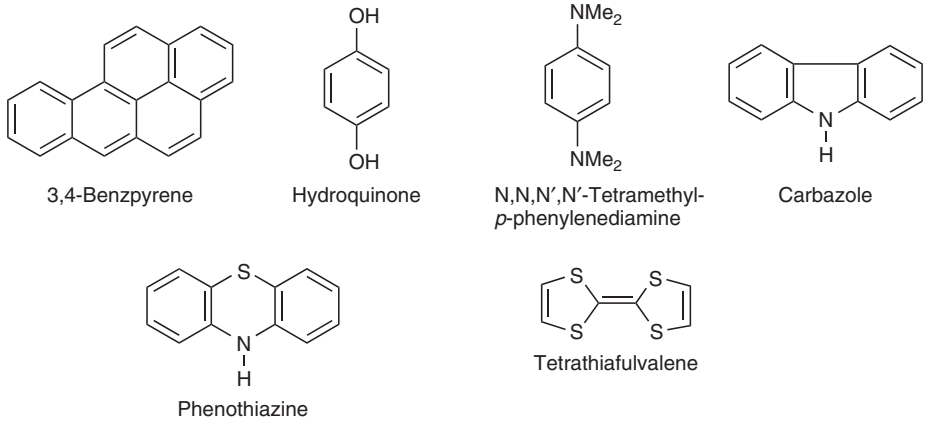


Figure 2.11 *Examples of electron donors.*

Aromatic polynitro compounds, quinones, conjugated polycyanides, inorganic substances such as molecular iodine, bromine, interhalogen compounds (ICl, IBr), heavy metal ions and so on are a few of the many types of electron acceptors commonly encountered. The acceptor class also includes all π-deficient heteroaromatic compounds and heteroaromatic cations in particular. Some examples of acceptors are shown in Figure 2.12.

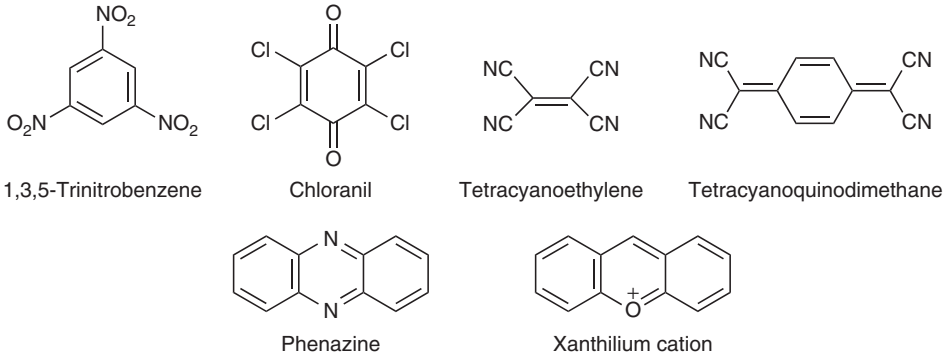


Figure 2.12 *Examples of electron acceptors.*

Both a strong donor and a strong acceptor are needed for the formation of ion-radical salts. This ensures a small energy gap between the donor's HOMO and the acceptor's LUMO which allows an electron to leap from one orbital to another, as in the case of ion-radical salt formation from tetrathiafulvalene and tetracyanoquinodimethane (Figures 2.13, 2.14a).

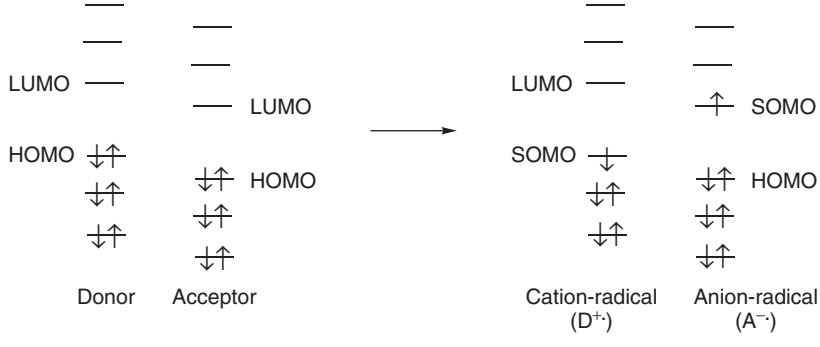


Figure 2.13 *Electron transfer in the formation of an ion-radical salt: the energy gap between the HOMO of the donor and the LUMO of the acceptor is smaller than that between the HOMO and LUMO of either the donor or the acceptor itself (SOMO = singly occupied MO).*

If the donor and acceptor are not sufficiently powerful, the energy difference between their frontier orbitals increases and the transfer of an electron becomes difficult. The possibility of these orbitals overlapping still remains and partial transfer of electron density or charge transfer may occur. The charges emerging as a result of this transfer (positive on the donor, negative on the acceptor) weakly bind the two molecules to form an association named a molecular complex

or a charge-transfer complex (CTC). Complexation of indole with chloranil generates a typical molecular complex (Figure 2.14b).

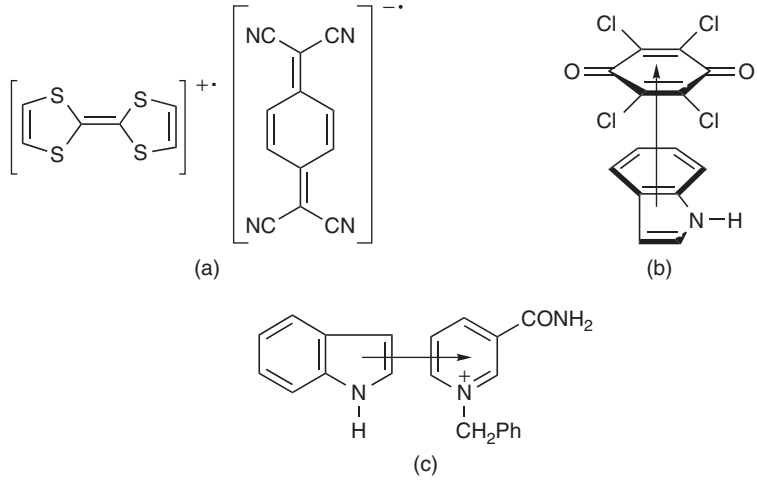


Figure 2.14 *Molecular complexes: (a) an ion-radical salt from tetrathiafulvalene and tetracyanoquinodimethane, (b) a charge transfer complex between indole and chloranil, (c) a CTC between the 1-benzyl-3-carbamoylpyridinium cation and indole.*

The CTC composition is not always in a simple 1:1 stoichiometric ratio. Two molecules of a donor may be linked with one molecule of an acceptor or vice versa. Binding energies in molecular complexes are normally below 6 kcal mol^{-1} and facile cleavage occurs in solution so that a rapid equilibrium exists with partial dissociation of the complexes into the donor and acceptor molecules.

The linkage between the components of a complex is symbolized by either a point or an arrow directed from the donor toward the acceptor:

$$D + A \leftrightarrows D^{\bullet}A \text{ or } D \rightarrow A$$

Donor and acceptor molecules tend to configure themselves into oriented layers such that maximum overlap occurs between their π-orbitals in the complex. The most indicative feature of CTC or ion-radical salt formation is the appearance of color in the reaction mixture. For example, tetrathiafulvalene, tetracyanoquinodimethane, indole and chloranil are all practically colorless compounds, but their ion-radical salts and molecular complexes shown in Figure 2.14 are greenish-black and red, respectively.

Although we have distinguished between ion-radical salts and molecular complexes, we should emphasize that there is no fundamental difference between them. In ion-radical salts the electron transfer is never complete and rarely exceeds 60%.[4] In other words, an ion-radical salt is a molecular complex in which electron transfer is quite pronounced.

Many biologically important heterocyclic compounds possess significant electron-donor or electron-acceptor ability. For example, metalloporphyrins, indoles and nucleic acid purine bases

[4] Here we draw attention to an analogy with ionic and covalent bonds. It is known that purely ionic bonds do not exist in condensed phases and that even in such a typical ionic compound as NaCl the electron transfer from sodium to chlorine does not exceed about 80%. Thus, a bond is formally referred to as ionic if the electron transfer exceeds 50%.

are all good donors. Electron-accepting properties are inherent in isoalloxazine, the main component of flavin systems (Figure 4.6), and 1-benzyl-3-carbamoylpyridinium chloride, which is used as a model for the respiratory coenzyme NAD^+. In test-tube (*in vitro*) experiments, these compounds react with various acceptors and donors to give molecular complexes. So, if indole is mixed with 1-benzyl-3-carbamoylpyridinium chloride, a yellow 1:1 molecular complex is afforded (Figure 2.14c). Such results suggest that molecular complexes also occur in living tissues (*in vivo*). Indeed, conclusive evidence has been obtained for the participation of ion-radical salts and molecular complexes in photosynthetic and respiratory processes. Electron transport may also play an important role in the action of some drugs, especially neurotropics.

2.4.5 Hydrophobic Forces

This type of noncovalent bonding interaction is not generally intrinsic to heterocyclic compounds. However, hydrophobic interactions do influence the behavior of heterocycles, especially in various life processes. If water is shaken with a nonpolar liquid, for example octane, a dispersion of tiny droplets of octane in water (i.e., an emulsion) is formed. When the agitation is stopped the octane droplets coalesce rapidly and the emulsion is converted into two liquid layers. This clearly demonstrates the presence of certain repulsive forces between the apolar octane molecules and water. These forces are named hydrophobic (Greek: *hydros* = water, *phobos* = fear).

As described above, hydrogen bonding is responsible for many of the specific properties of water. In order to dissolve a substance in water we need to break a large number of the hydrogen bonds which create the association between the water molecules. A substance can be dissolved in water only if it supplies the necessary energy for these processes to occur. Various salts, for example sodium chloride (or cooking salt), dissolve in water because the energetic expenditure is compensated by the energy released from the interactions between the Na^+ and Cl^- ions with the water dipole (solvation). Thus, 'supply and demand' is also evident in nature.

Nonionic compounds such as sugars, lower alcohols, ketones, carboxylic acids and pyridine are soluble in water because of the formation of new hydrogen bonds between these molecules and the water molecules. In contrast, the insolubility of octane, benzene and other nonpolar organic substances in water is caused by the fact that the attractive forces between the organic molecules and water are considerably weaker than those between the water molecules.

Surfactants such as trimethyloctylammonium chloride (Figure 2.15a) consist of a long hydrocarbon chain ('tail') with an ionic group at one end ('head') and display some very curious properties. When these substances dissolve in water, the water molecules repel the apolar hydrocarbon tails but are simultaneously attracted to the ionic head, thus solvating it. As a result of these contradictory tendencies the molecules of amphoteric compounds hide their 'tails' from the water dipoles by exposing only their 'heads' to the water. This curious situation results in the formation of spherical particles, namely micelles (Figure 2.15b). The formation of micelles aids the dissolution of amphoteric compounds and the solutions thus formed are somewhat turbid and opalescent because micelles are much larger in size than normal molecules.

Hydrophobic interactions therefore describe the fact that water molecules are more attracted to polar than to nonpolar compounds. The repulsive interactions with nonpolar compounds force them to gather together in specific aggregations. The useful rule of thumb 'like dissolves like' embraces this phenomenon.

How much energy is associated with hydrophobic forces? The association of two methyl groups has been calculated as a gain of 0.3 kcal mol^{-1} and that of two isobutyl groups is as much as 1.5 kcal mol^{-1} The same quantity of energy is released during the association of two phenyl groups (Figure 2.16) with a coplanar disposition of their rings ('stacking'). Stacking is a typical phenomenon for all planar rings including heterocycles. The specific role of stacking is evident in the stabilization of the DNA helical structure (see Section 3.2).

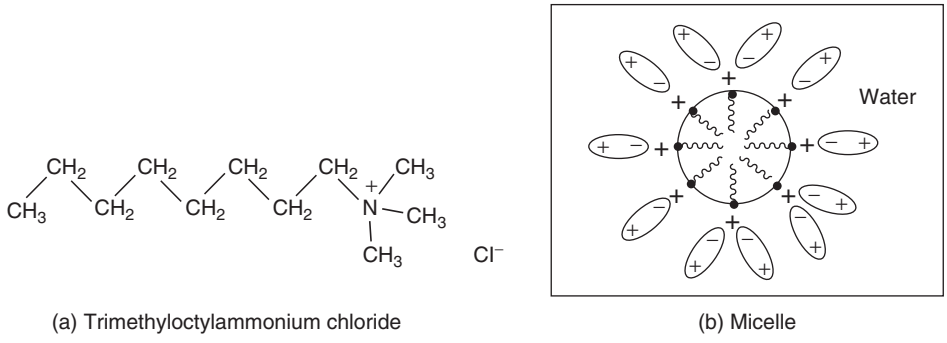


Figure 2.15 *A surfactant (a) and a representation of micelle formation (b).*

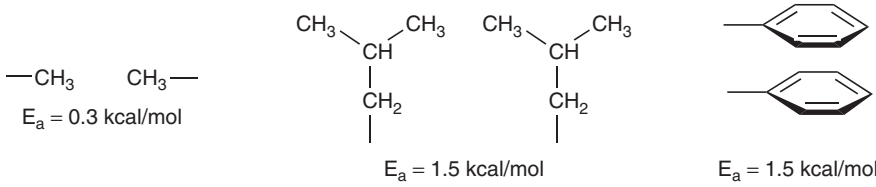


Figure 2.16 *Energies of association (E_a) of some hydrophobic groups.*

At first glance, hydrophobic forces appear to be very weak. But this is true only for the association of several small molecules. A many-fold increase in hydrophobic interaction energy is observed when hundreds of hydrocarbon groups of large biomolecules are involved, such as proteins having molecular weights up to hundreds of thousands of daltons. Associations of the hydrocarbon moieties of amino acid residues invoke the formation of hydrophobic clefts, pockets and cavities in the three-dimensional structures. Small molecules of other compounds, attracted by the same hydrophobic forces, can enter such structural clefts and sometimes fit together like a 'lock and key.' These intriguing properties of proteins determine their specific type of biological activity, whether as an enzyme, hormone or antibody.

A further important consideration connected with hydrophobic interactions is that water is our omnipresent natural solvent and is the medium for all biochemical reactions. If all biologically important compounds were water-soluble, life would be represented by a broth-like structure, as it presumably was during the early stages of chemical evolution. However, today, we see many thousands of highly organized forms of living matter. A vital prerequisite to such development is the hydrophobicity of many biomolecules: fats, proteins, polysaccharides, steroids and so on.

Hydrophobicity led to the initial structural organization of living matter at the cellular level. Cells became protected from the environment by a semipermeable membrane which vital nutrients could cross to enter the cell and metabolic wastes could pass to exit. The membrane structure (Figure 2.17a) and its functions are dependent, to a great extent, on hydrophobic forces. The membrane has a trilayer structure and consists of fats (about 40%, mainly phospholipids) and proteins (about 60%).

The external and internal walls of a cell membrane are composed of proteins, whereas the middle section is formed from both proteins and polar phospholipids having an amphoteric character (Figure 2.17b). Some protein molecules span the whole width of the membrane, forming highly

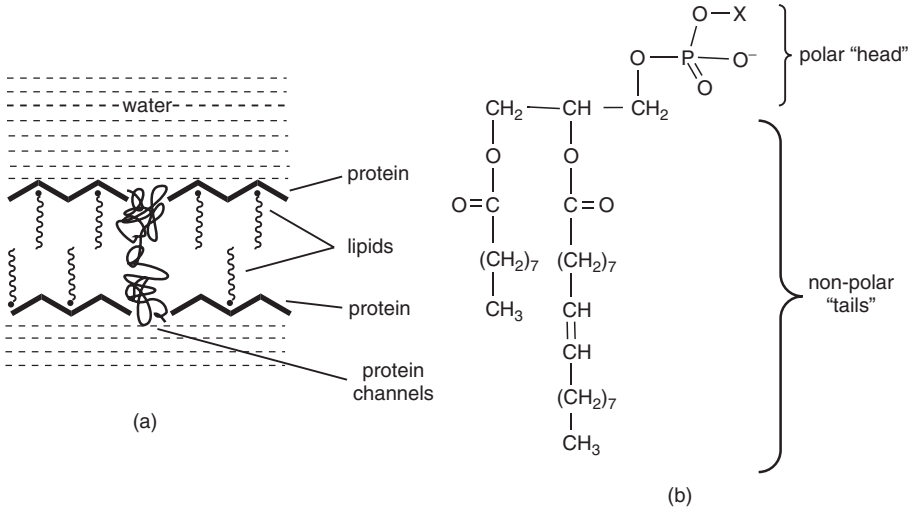


Figure 2.17 *(a) Model of the cell membrane. (b) Structure of phosphoglycerides, one of the groups of lipids which constitute cell membranes. X = different low molecular weight polar groups.*

specific channels. Through these channels polar and ionic substances can pass into a cell.[5] The phospholipids in turn are organized in two sublayers resulting from the orientation of the lipids with their nonpolar tails toward each other as in micelles. The wall proteins are associated with the polar heads of the lipids through hydrogen bonding and electrostatic attractions.

The presence of the extended lipid layer in the membrane (with a thickness of 60–70 Å, while the total thickness of the membrane is about 90 Å) allows various nonpolar molecules necessary for the functioning of the cell to penetrate the membrane.

2.5 Tautomerism: Heterocycles and Their 'Masks'

Like other classes of organic compounds, heterocycles contain compounds with the same elemental composition and molecular weight but different spatial arrangements. Such substances, called isomers, are able to exist independently and often have quite different physicochemical properties. Imidazole and pyrazole form a good example: structurally, these two isomers differ only in the arrangement of their two nitrogen atoms, but imidazole is a stronger base than pyrazole by a factor of 40 000.

When we refer to isomers, we most frequently mean two or more compounds which do not interconvert, or do so only with great difficulty. However, in chemistry there are a variety of reversible transformations in which a compound can exist in several isomeric forms in equilibrium. This type of isomerism is called tautomerism. The equation showing the equilibrium caused by the interconversion between cyanic and isocyanic acids is a simple example of tautomerism:

$$H{-}O{-}C{\equiv}N \rightleftharpoons O{=}C{=}N{-}H$$

[5] Polar moieties such as K^+ or Ca^{2+} ions are normally delivered into the cell by specialized molecular transporters that possess affinity both for the ions and for the nonpolar lipid phase (see Section 10.1.1). The lipid layer of the cell membrane also serves as a matrix into which certain key compounds can be incorporated.

As can be seen, the two isomers (or tautomers) differ in the position of proton attachment and also in the multiplicity of the bonds between oxygen, carbon and nitrogen. The migration of a proton from one heteroatom to another (sometimes to a carbon atom) often results in tautomeric interconversion. The ease of such interconversion results from the rather high acidity of the protons attached to heteroatoms. It is not surprising that the presence of different heteroatoms makes tautomerism ubiquitous in the heterocyclic series. Thus, imidazole normally exists with the proton interconverting between the nitrogen atoms at great speed (Figure 2.18a). In this case the two tautomers are indistinguishable because of the symmetry of the imidazole ring. But if we distort the symmetry by the introduction, for example, of a nitro group at position 4 or 5, the tautomers become nonequivalent and their equilibrium concentrations will be different. Figure 2.18b illustrates that the equilibrium shifts towards 1*H*-4-nitroimidazole, as indicated by the longer arrow (the ring atoms are numbered starting from the pyrrole-like nitrogen). This shift is explained by the fact that the strong acceptor nitro group decreases the electron density on all the ring atoms, but more strongly on the nitrogen atom nearest to the nitro group. The proton, therefore, is held more tightly at the nitrogen atom remote from the nitro group.

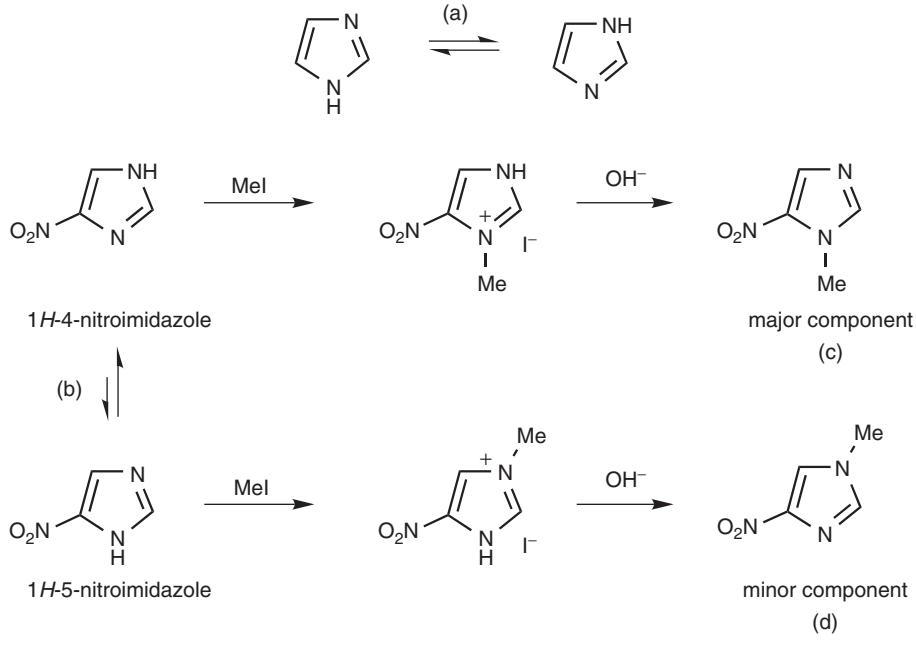


Figure 2.18 *Tautomerism of: (a) imidazole, (b) 4(5)-nitroimidazole. (c, d) N-Methylnitroimidazoles, fixed forms of tautomers.*

The study of tautomerism is very important since the structures of reaction products depend on their tautomeric equilibrium. For example, the methylation of 4(5)-nitroimidazole with methyl iodide in neutral conditions affords a mixture of 1-methyl-4-nitro- and 1-methyl-5-nitroimidazole, the latter being produced in markedly greater quantity (Figure 2.18c, d). Experimental evidence suggests that the 1*H*-4-nitro tautomer prevails in the initial mixture and that alkylation proceeds only at the pyridine-like nitrogen in neutral media. The heterocyclic bases are liberated by the action of aqueous hydroxide on the initially produced mixture of the two quaternary salts. The alkylated

products are not tautomers but as fixed forms provide models for the study of tautomerism as they have some properties very similar to those of the original individual tautomers. Individual tautomers, as a rule, cannot be isolated owing to their rapid interconversions.

To account for the behavior of tautomeric compounds, we need to realize that tautomers are, in effect, masks under which the same compounds are hidden. The name 'tautomerism' was proposed by Laar more than 100 years ago, is derived from the Greek meaning 'part of the same' (*tauto* = same, *meros* = part). One tautomeric compound can have many similar masks, which can often change depending on the media. For example, the biologically important compound purine can be written in four reasonable tautomeric forms (Figure 2.19). The fixed *N*-methyl models of all four tautomers have been prepared. In practice, however, only two tautomeric forms occur in measurable amounts: the 7*H*-tautomer and 9*H*-tautomer. Their ratio in water is near 1:1. Purine crystallizes exclusively in the 7*H*-form, whereas in a dimethyl sulfoxide solution the 9*H*-tautomer dominates.

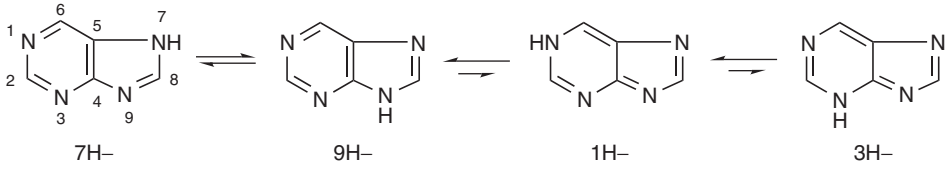
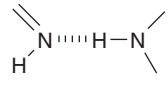


Figure 2.19 *Purine tautomers.*

In purine and imidazole the proton migrates between the ring nitrogen atoms. Tautomeric conversions in which ring heteroatoms and functional groups participate are no less widespread. 2-Hydroxypyridine, seen in Figure 2.20a, may exist in the hydroxy as well as in the oxo form. In the vapor phase and in highly dilute hexane solutions both exist, with the hydroxy form predominating by a small factor. From purely bond energy considerations, the amide structure of the 2-pyridone form would be expected to be more stable than the imidol structure of 2-hydroxypyridine. However, in low polarity media the higher aromaticity of the latter is decisive. Nevertheless, the second form, 2-pyridone, exists in the crystalline phase and dominates in all solvents more polar than hexane. The vapor phase equilibrium assesses the stability of tautomers in the absence of extraneous effects. The preponderance of the hydroxy tautomer in the gas phase is explained by its aromatic character, while the aromaticity of the 2-pyridone structure obviously depends on charge transfer which is less favored in media of low dielectric constant. The driving force for the shift in the equilibrium toward the oxo tautomer in the crystalline state and in aqueous solution lies in the stabilization of the highly polar pyridone form. A particular stabilization occurs by dimeric association in which two molecules are attached to each other by hydrogen bonding and by dipole–dipole interactions (Figure 2.20b). The dimers are further stabilized by the polar environment (crystalline lattice or polar solvents). These interactions outweigh the energy losses induced by the lower aromaticity of the pyridone form.

In 2-aminopyridine an equilibrium between the amino and imino forms is also possible (Figure 2.20c). However, only the amino structure is detectable in all phases (vapor, liquid, solid). The transition to the imino form offers no advantage in terms of bond energies and the imino form is of lower aromaticity with no offsetting stability of a dimer. The last factor is a consequence of the relatively low polarity of the imine and the decreased strength of hydrogen bonding in the



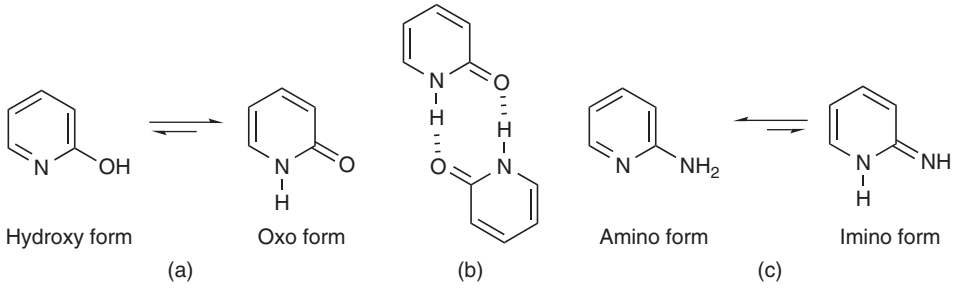
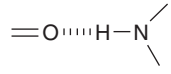


Figure 2.20 *(a) Tautomerism of 2-hydroxypyridine with 2-pyridone. (b) The 2-pyridone dimer. (c) Tautomerism of 2-aminopyridine with 2-pyridonimine.*

fragment as compared to that in the

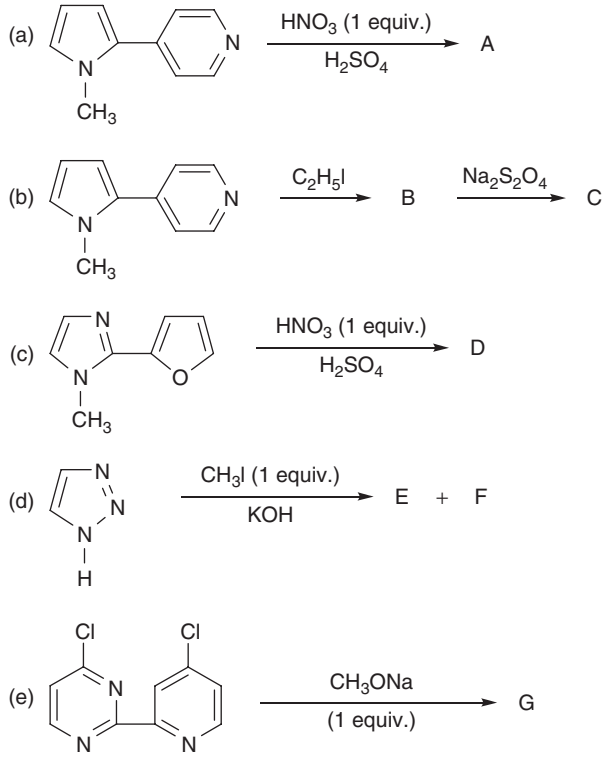


fragment, as well as geometrical considerations.

In closing this chapter, we hope that our readers will have realized the unique role of heterocycles in nature. This role is explained by the pervading influence of the heteroatoms on the reactivity, nonbonding interactions and structural modifications of heterocyclic compounds. Heterocycles have both multipurpose and specific properties which are implicit in many important chemical, biochemical and technical applications, as discussed in the following chapters.

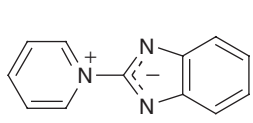
2.6 Problems

1. Piperidine (pK_a 11.22), unlike pyrrole, is a strong base. Account for this fact. Also explain the significantly higher basicity of piperidine in comparison to pyridine.
2. Benzimidazole has two ionization constants with pK_a values of 12.9 and 5.3. To which acid–base equilibria do these correspond?
3. Purine has two ionization constants: $K_a = 4.07 \times 10^{-3}$ and $K_a = 1.17 \times 10^{-9}$. Calculate the p$K_a$ values and write the equations describing the corresponding acid–base equilibria.
4. Arrange the anions of imidazole, tetrazole, purine and benzimidazole in decreasing order of basicity (see Section 2.2 and Problems 2 and 3).
5. Imidazole is nitrated by a mixture of concentrated nitric and sulfuric acids at 100 °C to give 4(5)-nitroimidazole in 75% yield. By contrast, it is difficult to halt the bromination of imidazole at the monosubstitution stage even at room temperature (reaction of bromine in chloroform leads directly to 2,4,5-tribromoimidazole). Account for the difference in reactivity.
6. Along with the dimers shown in Figure 2.8b, pyrazole also forms hydrogen-bonded structures of a higher order. Write three- and four-membered macrocycles of this type.
7. Provide structures for compounds A–G in the following reactions:

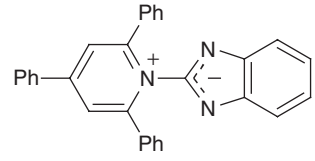


8. Assess the polarities of the three isomeric diazines pyridazine, pyrimidine and pyrazine on the basis of the dipole moment for pyridine being equal to 2.2 D.

9. The dipole moment of compound H decreases as its concentration in dioxane solution is increased. At the same time the dipole moment of compound I, being equal to 18.7 D, remains unaffected with increasing concentration. Explain.

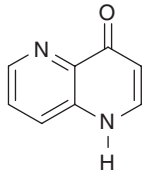


H

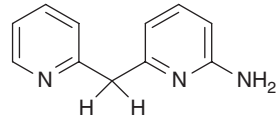


I

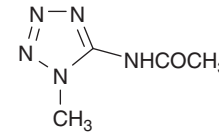
10. One part of benzene dissolves at 20 °C in 600 parts of water. However, pyridine is miscible with water in any proportion. Account for this fact.
11. A surprising property of nitrogen heterocycles is the decreased solubility of many derivatives when a CH is replaced by an OH group (the reverse of what is observed in the aliphatic series). Remember that such compounds exist mainly in their oxo forms. For example, one part of purine and one of each of its 2-oxo, 6-oxo, 8-oxo, 2,6-dioxo (xanthine) and 2,6,8-trioxo (uric acid) derivatives will dissolve in two, 380, 1400, 240, 2000 and 39 500 parts of water, respectively, at 20 °C. For comparison, one part of 1,3,7-trimethylxanthine (caffeine) dissolves in 70 parts of water. Explain.
12. Quinoline and naphthalene form 1:1 molecular complexes with bromine. The stability constants (K_c) of these complexes are equal to 115 and 0.231 mol^{-1} (CCl_4, 20 °C), respectively, and their long wavelength absorption maxima occur at 290 and 346 nm, respectively. Discuss the reasons for these differences (hint: the K_c values of the analogous 8-bromo-, 8-methyl- and 3-bromoquinoline complexes are 1.1, 4.8 and 12, respectively).
13. 1-Acetylimidazole has a half-life in aqueous solution at pH 7.0 and 25 °C of 41 min. By contrast, aliphatic amides (e.g., acetamide) are not hydrolyzed under these conditions. Account for this fact.
14. Indicate which of the following compounds can theoretically exist in alternative tautomeric forms containing intramolecular hydrogen bonding:



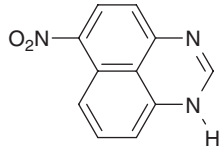
J



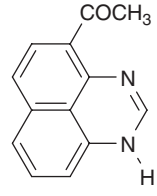
K



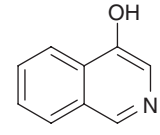
L



M

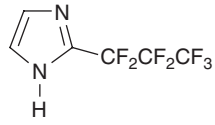


N

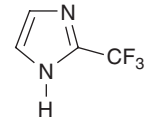


O

15. In the NMR ^{1}H spectrum of 2-perfluoropropylimidazole P (in acetone-d_6 at 25 °C and 1–2% concentration) protons H-4 and H-5 display two different signals whereas in the spectra of 2-trifluoromethylimidazole R and compound P, but at higher concentrations, they are equivalent showing a two-proton singlet. Explain these observations.



P



R

16. The two rings in sandwich structures (b) and (c) in Figure 2.6 do not rotate and therefore are held rigidly in opposite orientations. In complex (d) the rings slowly rotate. Explain.

2.7 Suggested Reading

1. Gilchrist, T. L., *Heterocyclic Chemistry*, 3rd edn, Pearson Education Press, New York, 1997.
2. Katritzky, A. R. and Pozharskii, A. F., *Handbook of Heterocyclic Chemistry*, 2nd edn, Pergamon Press, Oxford, 2000.
3. Katritzky, A. R. and Rees, C. W., (eds), *Comprehensive Heterocyclic Chemistry*, vols 1–8, Pergamon Press, Oxford, 1984.
4. Katritzky, A. R., Ramsden C., Scriven E. and Taylor, R. (eds), *Comprehensive Heterocyclic Chemistry III*, vols 1–15, Elsevier, New York, 2008.
5. Jencks, W. P., *Catalysis in Chemistry and Enzymology*, Dover, New York, 1987.
6. Matuszak, C. A. and Matuszak, A. J. Imidazole – versatile today, prominent tomorrow, *J. Chem. Educ*., 1976, **53**, 280.
7. Elguero, J., Katritzky, A. R. and Denisco O. V. Prototropic tautomerism of heterocycles: heteroaromatic tautomerism – general overview and methodology, in *Advances in Heterocyclic Chemistry*, vol. 76, Academic Press, San Diego, 2000.
8. Elguero, J., Marzin, C., Katritzky, A. R. and Linda, P. The tautomerism of heterocycles, in *Advances in Heterocyclic Chemistry*, *Suppl. 1*, Academic Press, New York, 1976.
9. Frieden, E. Non-covalent interactions: key to biological flexibility and specificity, *J. Chem. Educ*., 1975, **52**, 754.
10. Slifkin, M. A., *Charge Transfer Interactions of Biomolecules*, Academic Press, London, 1971.
11. Albert, A., *Heterocyclic Chemistry: an Introduction*, 2nd edn, Athlone Press, London, 1968.
12. Pozharskii, A. F., *Theoretical Basis of Heterocyclic Chemistry* (in Russian), Khimia, Moscow, 1985.
13. Anslyn, E. V. and Dougherty D. A., *Modern Physical Organic Chemistry*, University Science Books, Sausalito, 2006, Chap. 3.
14. Jeffrey, G. A., *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, 1997.
15. Hancock, R. D. Chelate ring size and metal ion selection. The basis of selectivity for metal ions in open-chain ligands and macrocycles, *J. Chem. Educ*., 1992, **69**, 615.

3

Heterocycles and Hereditary Information

I do not care, be what there will.
The weird sisters, with timeless skill,
Keep their wheels spinning to generate
The tangled threads of our fate.
D. Merezhkovsky

3.1 Nucleic Acids

For centuries the question of heredity and the related theme of destiny were surrounded with a mystic secrecy which provided imaginative scope for astrologers, prophets and fortune tellers. Ancient Greco-Roman mythology offers numerous tales of the Fates or Parcae, fortune goddesses who were portrayed as very old women spinning and severing the threads of human life at will. Many poets have paid them tribute; perhaps, the Russian symbolist Merezhkovsky did it especially expressively in the above quatrain written more than 100 years ago.

The significant associations made by ancient mythologists and poets, surprisingly enough, anticipated the scientific fact that the threads of life are represented in reality by very long macromolecules. These macromolecules, not tangled but exquisitely organized, indeed store all the hereditary information which predetermines the development of all life forms, be they plants, animals or mankind.

The mechanism of the transfer of hereditary information from one generation to the next was discovered in 1953 by Watson and Crick in one of the most impressive achievements of science in the twentieth century, appropriately earning a Nobel prize. Avery in 1944 established that hereditary information, in other words, the genetic code, is encrypted in the huge deoxyribonucleic acid (DNA) molecules (with molecular weights of several millions) contained in each living cell. It was Watson and Crick who demonstrated that cytosine, thymine, uracil, adenine and guanine (Figure 3.1), derivatives of the well known nitrogen heterocycles pyrimidine and purine, participate directly in the encoding of all genetic information and proposed how the information is transferred from generation to generation.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

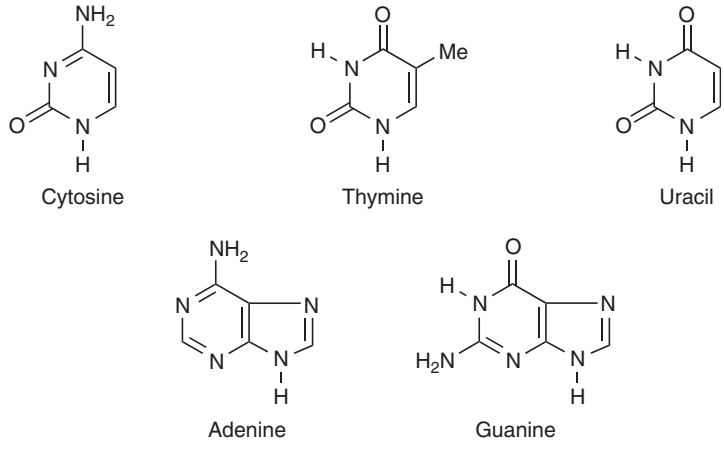


Figure 3.1 *Nucleic acid pyrimidine and purine bases.*

Cells also contain ribonucleic acid (RNA). The chief function of RNA is to control the synthesis of proteins. The DNA code is transferred to the RNA and is thus used to determine the amino acid sequence of proteins. Thus, in one sense RNA plays a secondary role relative to DNA.

We first examine the primary structures of DNA and RNA. By 'primary structure' is meant the nature of the components and their arrangement in the long molecular sequence. Complete hydrolysis of nucleic acids leads to the isolation of three types of subunits: (i) phosphoric acid, (ii) a pentose sugar and (iii) the heterocyclic bases enumerated above. At this level there are two substantial differences between DNA and RNA. Firstly, whereas both DNA and RNA contain cytosine, guanine and adenine, DNA contains thymine and no uracil, while RNA contains uracil and no thymine. Secondly, the sugar component of RNA is D-ribose, whereas the comparable subunit in DNA is D-2-deoxyribose. The names of RNA and DNA are in fact derived from the names of the corresponding sugars.

As is the case for most monosaccharides, ribose exists not in the linear form but as a mixture of cyclic hemiacetals. Intramolecular nucleophilic addition of the C-4 or the C-5 hydroxy group to the carbonyl group affords, respectively, the five-membered furanose or six-membered pyranose ring (Figure 3.2). In aqueous solution D-ribose comprises 76% of the two pyranose forms and 24% of the two furanose forms. The latter are less stable because of the angle strain inherent in five-membered rings. In both furanose and pyranose forms, the hydroxy groups at C-1 are those of hemiacetals; compounds of type $R_2C(OH)OR$ are denoted hemiacetals to differentiate them from acetals of the general formula $R_2C(OR)_2$.

As implicit in Figure 3.2 and the discussion above, the hemiacetal hydroxy group of cyclic sugars occupies one or the other of two different positions relative to the average ring plane and the neighbouring hydroxyl group at the C-2 atom. A stereoisomeric form is designated as α when these two OH groups are on same side of this ring plane and β when they are on the opposite sides. Accordingly, for ribose we have the α- and β-pyranose and the α- and β-furanose forms. In spite of the lower strain energy of pyranose forms, the ribose moieties in nucleic acids are exclusively β-D-ribofuranoses. This configuration apparently facilitates construction of long and stable (from the point of view of electronic and geometric factors) polymer chains.

The way in which the units are joined together to form nucleic acids was solved by enzymatic hydrolysis. Four crystalline acids were isolated from the hydrolysis of DNA: deoxyadenylic, deoxyguanylic, deoxycytidylic and deoxythymidylic acids (Figure 3.3). Each acid was composed

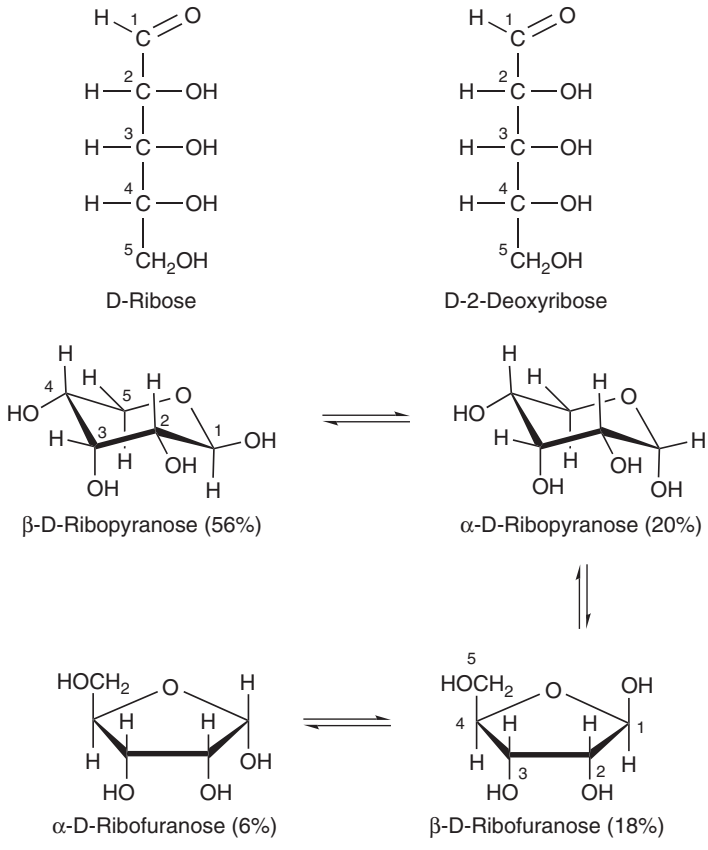


Figure 3.2 *The linear forms of ribose and deoxyribose and the cyclic forms of ribose.*

of one molecule each of a heterocyclic base, a sugar and phosphoric acid. One of the nitrogen atoms of the base was attached to the furanose ring, thus forming a nucleoside. The phosphoric ester units derived from a nucleoside and phosphoric acid are called nucleotides (or deoxyribonucleotides). The analogous enzymatic hydrolysis of RNA generates four different acids (ribonucleotides): adenylic, guanylic, cytidylic and uridylic acids (Figure 3.3).

The nucleotides of DNA and RNA have the following structural features: (i) the pentose residue is attached to the N-9 atom of purines and N-1 in pyrimidines (these positions are sterically the least hindered), (ii) the base is condensed with the sugar by displacement of the hemiacetal hydroxy group at the pentose C-1 atom, (iii) the pentose CH_2OH group is esterified by phosphoric acid. Since the other hydroxy groups of pentose can also be esterified, the names of the nucleotides shown in Figure 3.3 need further clarification. For example, a more precise name for adenylic acid is adenosine-5′-phosphate, 5′ designating the number of the carbon atom in the sugar unit which carries the phosphoric acid residue (see also Figure 3.2).

Further chemical investigations have established that nucleotides are linked one to another by a second phosphate ester bond. The bond is formed between the phosphoric acid residue of one nucleotide and the 3′-pentose hydroxy group of the other. Thus, polynucleotide chains contain 3′–5′ phosphodiester linkages. A section of the sequence is represented in Figure 3.4.

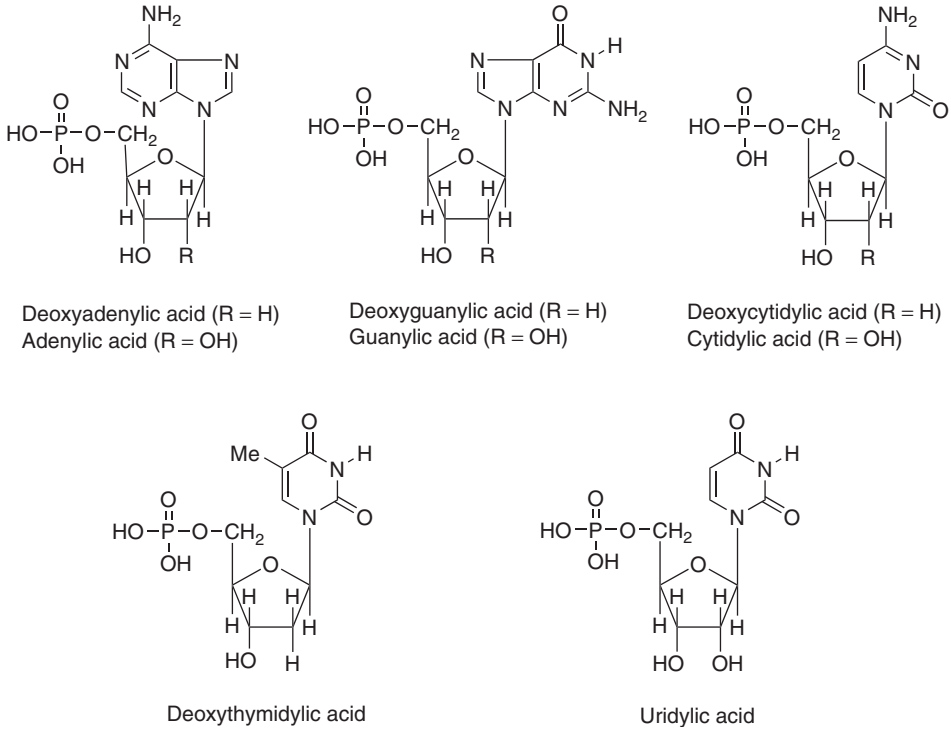


Figure 3.3 *Nucleotides formed from the enzymatic hydrolysis of DNA and RNA.*

Apart from considerations of length, the general structures of the polynucleotide chains of all nucleic acids show considerable similarity. However, they are differentiated by the variable sequences of purine and pyrimidine residues on the polyester backbone, which are of great significance. Indeed, these sequences express the genetic code. Therefore, elucidation of the primary structure of DNA and RNA above all necessitates elucidation of the precise sequence of the heterocyclic bases. Reliable and effective methods for such determination now exist. Moreover, methods for the synthesis of sequence-specific polynucleotides have been elaborated, and today one can use an automated synthesizer. Such computer-controlled instruments assemble high molecular weight nucleic acids from nucleotide units, in one sense imitating the natural assembly lines which exist in cellular organisms.

3.2 The Double Helix

The fascinating history of the discovery of the structure and mode of action of DNA was described in 1968 by Watson in his book *The Double Helix*. His account reveals that the most difficult aspect of the characterization was the elucidation of the spatial, the so-called secondary, structure of DNA. The now well known double-helix backbone structural model of DNA was consistent with all the experimental data then available, including X-ray analysis results. As demonstrated by Watson and Crick, the main features of the secondary DNA structure are as follows.

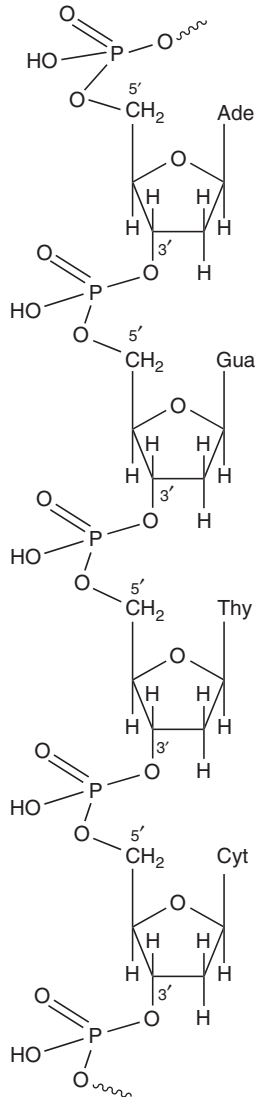


Figure 3.4 *3′–5′ Phosphodiester bonding between nucleotides in a DNA chain (Ade = adenine, Gua = guanine, Thy = thymine, Cyt = cytosine).*

1. The DNA molecule consists of two right-handed polynucleotide chains which are intertwined to form a double helix (Figure 3.5). The helices run antiparallel to each other; that is, if one moves along them in the same direction, the sequence of bonding in one sugar-phosphate chain will be $-5',3'-5',3'-5',3'-$, whereas in the other it will be the reverse configuration $-3',5'-3',5'-3',5'-$.
2. The purine and pyrimidine base units lie in the center of the double helix with their planes nearly perpendicular to the main axis of the helix. By contrast, the planes of the pentose rings are almost parallel to the main axis of the helix.

3. One coil of each helix has a length of 34 Å and includes 10 nucleotide pairs. The distance between the planes of each pair of neighboring bases in one chain is 3.4 Å.
4. Significant contributions to the secondary structure of DNA are made by several types of noncovalent bonding interactions which serve to attach the two helices together. Heterocyclic bases of one chain are bound to the appropriately situated bases of the other chain by hydrogen bonding. It is only certain specific pairs of bases which are thus attached to one another, the so-called complementary pairs [e.g., adenine–thymine (two hydrogen bonds) and guanine–cytosine (three hydrogen bonds)], as depicted in Figure 3.6. It follows that the number of purine residues in a DNA molecule exactly equals the number of pyrimidine moieties. In all types of DNA this ratio is experimentally found to be 1:1, although the ratio of each type of base within the purine and pyrimidine classes can change depending on the biological species. So, in double-stranded helical DNA, the two chains are complementary to each other. For example, human sperm contains 31% adenine, 19% guanine, 31% thymine and 19% cytosine, while the DNA of tuberculosis bacilli contains 36% each of guanine and cytosine, adenine and thymine at only 14% each.

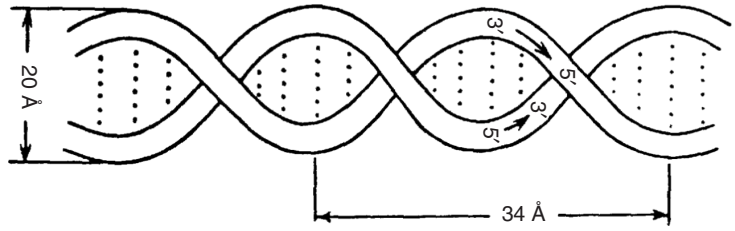


Figure 3.5 *Schematic representation of the double-helix structure of DNA (hydrogen bonding of purine-pyrimidine pairs is shown by dotted lines). Reproduced with permission from Terney,* Contemporary Organic Chemistry, *Saunders, Philadelphia, 1979. © 1967 Recthed,* Concepts of Biochemistry, *McGraw–Hill Book Co.*

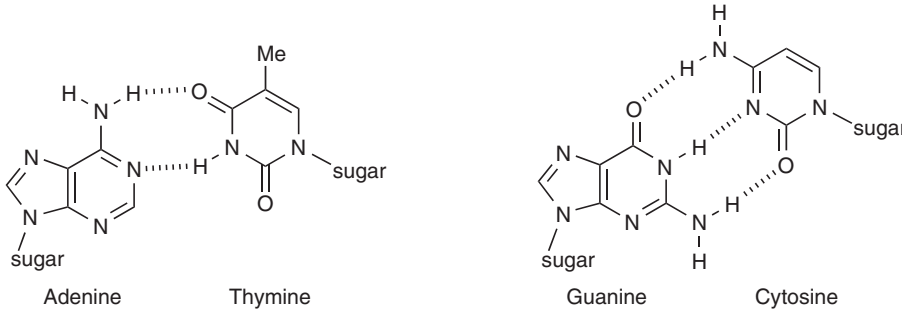


Figure 3.6 *Hydrogen bonding between complementary pairs of heterocyclic bases in DNA.*

The second force which stabilizes each individual DNA chain in the helical form as well as both chains together in the double-helix arrangement is the stacking interaction (see Section 2.4.5). Hydrophobic interactions exist between heterocyclic bases which are vertically stacked. The role of these forces in the formation of the double helix is believed to be even more important than that of hydrogen bonding.

DNA molecules also have a tertiary structure. The fact is that linear DNA could not be accommodated within a cellular structure: the length of all the DNA molecules in a single cell can reach a total of 2 m. To adopt more a compact configuration DNA is wound onto proteins called histones and these large bundles (chromatins)[1] are in turn compacted into superhelices. This allows the very long DNA molecules, containing thousands of genes and regulatory elements, to fit into the cell nucleus. The highly organized structure of each DNA molecule forms a chromosome. The number of chromosomes varies in the cells of different living organisms. For example, a single human cell contains a total of 46 chromosomes, two each of 23 types. Under a microscope, chromosomes appear as small, slightly bent rods (Figure 3.7).

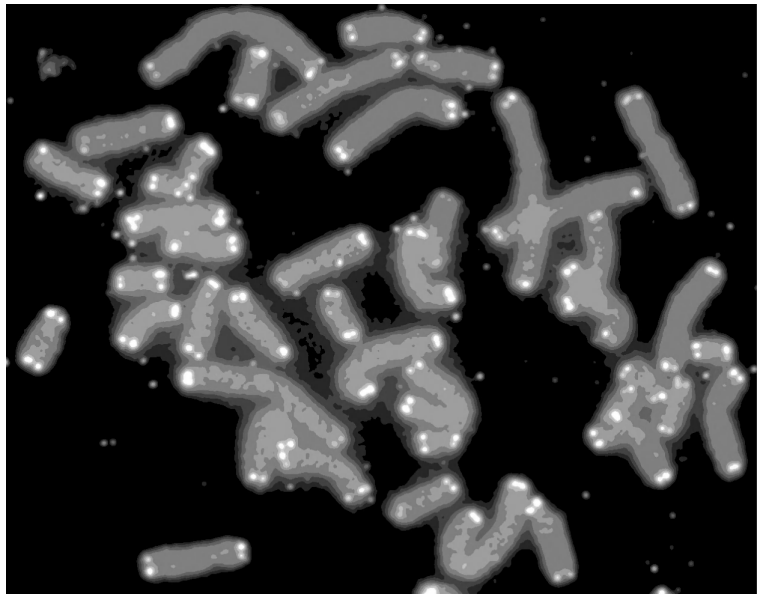


Figure 3.7 *Human chromosomes (gray) capped by telomeres (white). Retrieved from en.wikipedia; transfer was made by User: gustavocarra; original text: http://science.nasa.gov/headlines/y2006/images/telomeres/caps_med.jpg. Author: United States Department of Energy Human Genome Program. Original uploader was SierraSciSPA at en.wikipedia, on 28-01-2008. http://en.wikipedia.org/wiki/File:Telomere_caps.gif (accessed 29 October 2010)).*

Chromosomal DNA molecules are used in many biochemical experiments. Because a chromosomal DNA double helix is only 2 nm in diameter while its length ranges from millimeters to centimeters, hydrodynamic shear can easily break a DNA strand. To prevent this breakage, Japanese researchers designed microhooks and microbobbins to form a 'sewing machine' to manipulate DNA (K. Terao, M. Washiru and H. Oana, *Lab. Chip*., 2008, **8**, 1280). The microscopic tools created via photolithography, consist of a Z-shaped microhook to catch the DNA and a microbobbin, around which the DNA is wrapped. Using laser methods to manipulate the microtools, the pairs of DNA helices were captured and separated with a microhook. These scientists also wrapped and

[1] There exist many other types of DNA–protein complexes. They are commonly named nucleoproteins. Most of them fulfill regulatory or signalling functions. Proteins interact with the DNA in grooves formed on the external surface of the double helix.

unwrapped the DNA from two microbobbins. Amusingly, such experiments once again turn us to the allegory of poet Merezhkovsii as cited at the beginning of this chapter.

We next turn from DNA to consider the secondary structure of RNA. In the overwhelming majority of cases, the RNA encountered in nature is composed of a single rather than a double strand (see Figure 3.21 for an example). Nevertheless, its secondary structure is determined by the same types of interactions as are encountered in DNA. Individual portions of an RNA molecule can become associated under the influence of hydrogen bonding and stacking. When a sufficiently lengthy section of the molecule has such interactions, this section can become partly helical, resembling the gross DNA structure. For instance the secondary structure of the RNA, which controls the transfer of phenylalanine, is a typical example of a partially helical molecule (Figure 3.8), with a molecular form which, if depicted in a single plane, resembles a clover leaf. The bases in the short stems are complementary to each other, just as in DNA, forming hairpin loops. The three-dimensional structure is still more complex (Figure 3.9).

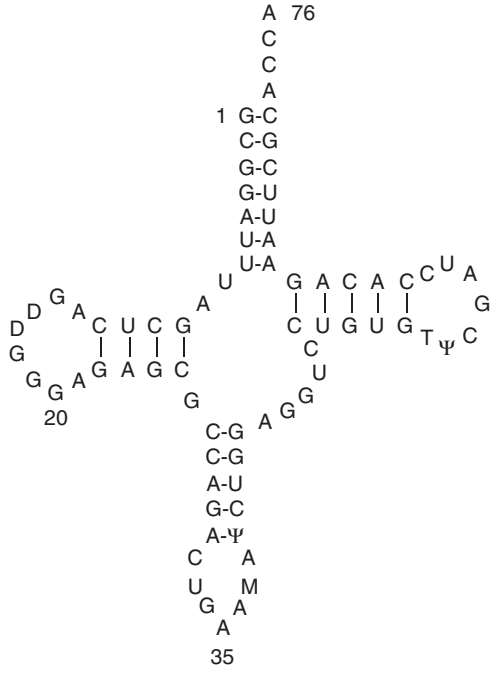


Figure 3.8 *Heterocyclic base sequence in yeast phenylalanine transfer RNA (tRNA). A single line indicates a pair of bases attached by hydrogen bonding (D = 4,5-dihydrouracil, Ψ = pseudouracil, M = dimethylguanine). Adapted from Pauling, L. and Pauling, P.,* Chemistry*, Freeman, San Francisco, 1975, with permission.*

3.3 How One DNA Doubles Itself

By the process of self-duplication, a living cell produces two new cells, each identical to the parent. The manner by which such a strikingly accurate reproduction of all of the complex biological material takes place was first suggested by Watson and Crick. They proposed that at a given time, specialized proteins direct the uncoiling of the double helix. Each of the two individual DNA

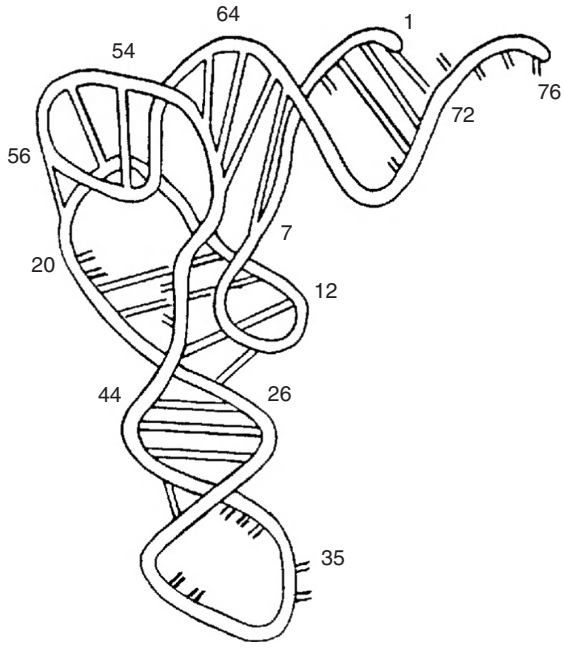


Figure 3.9 *Schematic representation of the three-dimensional structure of phenylalanine tRNA. Double lines designate hydrogen bonding between heterocyclic bases (base numbers correspond to those in Figure 3.8). Reproduced with permission from Pauling, L. and Pauling P.,* Chemistry*, Freeman, San Francisco, 1975.*

chains then begins to make its complement, each effecting the biosynthesis of a new double helix identical to the original. As a result, all the initial genetic material of the cell is doubled: the 46 initial chromosomes are converted into 92 chromosomes and the swollen parent cell then divides.

It should be noted that the entire double helix of DNA cannot be completely uncoiled simultaneously, even theoretically. The helix begins to unwind at several points like a broken zipper. Because the chains in DNA are antiparallel, one strand can be made continuously, while the other strand is made in shorter bits which are joined together by the enzyme DNA ligase. DNA synthesis is catalyzed by very complex enzymes called DNA polymerases, first studied by Kornberg. This process, called replication, is shown schematically in Figure 3.10. Scientists have been able to study replication with the aid of electron microscopes and have demonstrated the accuracy of this hypothesis.

The mechanism of replication relies on the timely delivery to the assembly line of each necessary nucleoside-5′-triphosphate and its subsequent attachment to the growing DNA chain by DNA polymerase (Figure 3.11). This reaction involves nucleophilic displacement of a pyrophosphate group ($P_2O_7^{4-}$) resulting from attack by the terminal nucleotide 3′-hydroxy substituent at the α-phosphate group of the nucleoside-5′-triphosphate. Thus, synthesis of the chain is in the 5′–3′ direction, one chain being extended continuously, the other in short sections in the opposite direction.

The new DNA molecules are constructed with the complementary heterocyclic bases. For instance, if a 5′-deoxyadenosine residue appears on a DNA strand, it is absolutely necessary for a 5′-deoxythymidine unit to be provided at the corresponding site of the new but growing DNA molecule. Thus, one double-helical DNA structure affords two new identical double-chained DNA aggregates, each composed of one original and one newly synthesized strand of DNA. The

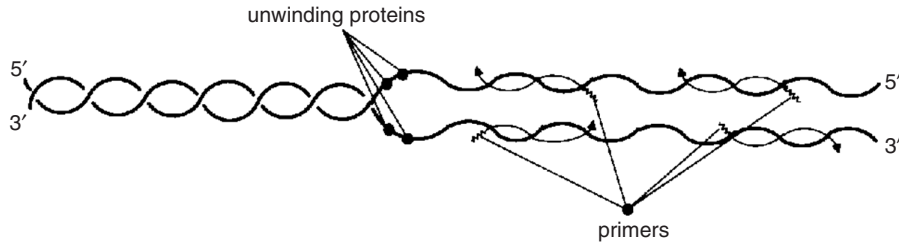


Figure 3.10 *DNA replication. Adapted from Barton, D. and Ollis, W. D. (eds),* Comprehensive Organic Chemistry*, vol. 5, Pergamon Press, Oxford, 1979, with permission.*

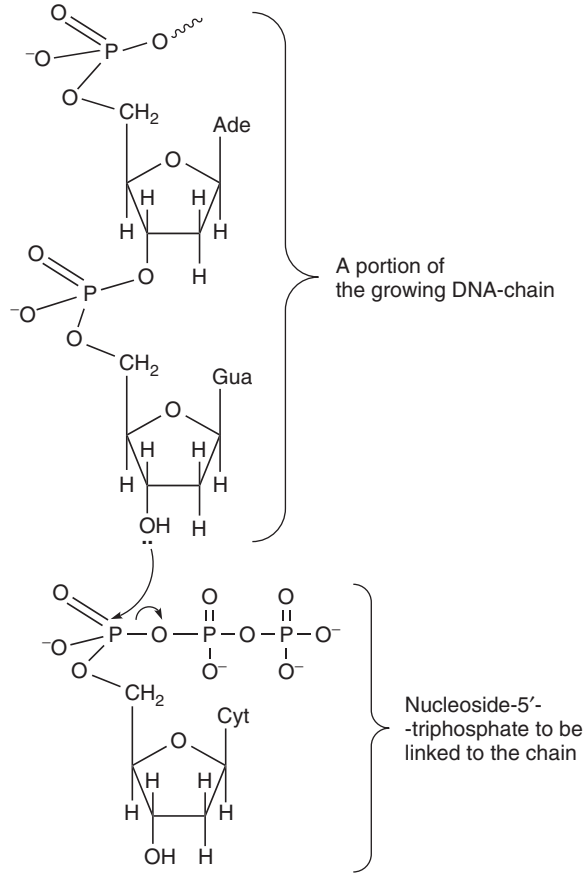
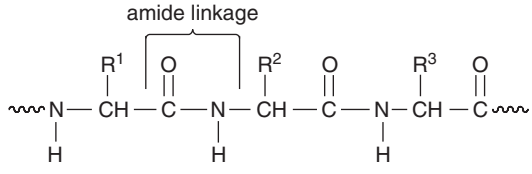


Figure 3.11 *DNA chain lengthening with the assistance of the DNA polymerase delivery system.*

high degree of fidelity is responsible for genetic stability, together with many enzymes that repair mistakes and DNA damage.

3.4 Protein Synthesis, Genetic Code and the Genome

Overall, proteins make up the most important class of natural compounds, because proteins perform an enormous variety of functions in the living organism. Cellular junctions, supporting and protective tissues, and muscle fibers are but a few examples of living structural materials constructed from proteins. Moreover, a key role of many proteins is to regulate metabolic processes (see Chapter 4). Chemically speaking, proteins are biopolymers with molecular weights ranging up to hundreds of thousands in which the polymer chains are constructed from α-amino acid residues joined by amide linkages:



When such a chain includes a relatively small number of amino acids (several dozen), the compound is named a polypeptide, or simply a peptide. A peptide with many amino acid residues is called a protein. As mentioned above, the diversity of nucleic acids stems from inclusion of only four heterocyclic bases. By contrast, proteins are derived from as many as 20 or so amino acids, giving rise to an almost infinite number of possible sequences within the polypeptide chain. These chains have different side groups (R) which include heterocyclic residues as in histidine and tryptophan (Figure 3.12). Two derivatives of pyrrolidine, namely proline and hydroxyproline, are unusual α-amino acids in that their amino groups participate in the formation of the heterocycle; they are thus secondary amines.

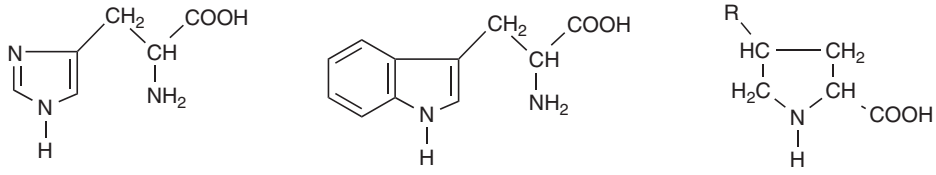


Figure 3.12 *Heterocyclic α-amino acids.*

There are obviously infinite possibilities for changing the primary structure of proteins, that is, the sequence of amino acid residues in the chains. In human organisms, thousands of proteins have been identified. Proteins are self-organizing molecules. Under the influence of noncovalent interactions between different portions of the polypeptide chain, the protein is arranged into a definite spatial form (tertiary structure), which determines its functions. Moreover, the ability of proteins to complex with metal ions and biomolecules such as nucleic acids, sugars, lipids, pigments and so on extends their chemical and biological functions.

Unlike DNA, proteins are not capable of self-reproduction. The genetic program for their synthesis is contained in the organism's genome. The genome is a complete set of DNA where all instructions on body construction and organism activity are stored. The fully decoded sequence of the human genome was announced on April 14, 2003 (*Nature*, **431**, p. 931–945. doi: 10.1038/nature0300/) and has become one of the greatest scientific achievements in human history. The human genome contains nearly three billion base pairs that reside in the 23 pairs of chromosomes (Figure 3.9). Relatively large DNA pieces where coded information is stored to make certain types of proteins are called genes. Some portions of different genes often overlap each other, particularly in microorganisms. A human organism contains about 25 000 genes, each of which produces an average of three proteins. Currently, scientists are focusing on finding the precise location of all of the protein-coding genes and nonprotein-coding sequences within the genome. This should speed up the discovery of genes responsible for such common illnesses as cancer, diabetes, heart and nerve diseases.

Significantly, protein-coding genes compose only 5% of the complete human genome. The rest consists mainly of transposons, moving sites of DNA that form and change the genome structure, and 'silent' genes, or multirepeated sequences of nucleotides. These two last types of genetic sites are said to have been included in the human genome during evolution by bacteria and viruses. The chimpanzee genome contains the same number of genes (25 DNA chromosomes, composed of 3.1 billion nucleotides), and 96.0–98.8% of the nucleotide sequence coincides with that of the human genome. Curiously, the chimpanzee genome contains a gene which protects these primates from Alzheimer's disease, possibly allowing the design of new and effective drugs against this disease.

The synthesis of a protein is an exceedingly complex biochemical process in which three types of ribonucleic acids participate: messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA). The genetic information for protein synthesis is contained in mRNA. The synthesis of mRNA itself is directed by one of the DNA strands. The chemistry of this synthesis, called transcription, is similar to the replication of DNA, but a significant difference is the participation of a different enzyme, RNA polymerase, which introduces uracil instead of thymine into the RNA chain. The molecular weights of RNA are very varied and in some cases can be in the vicinity of 10 000 or more bases.

As proteins contain 22 different amino acids and DNA includes only four bases, the incorporation of one amino acid into a protein chain cannot be encoded by one or even two bases. It has been established that a particular amino acid is encoded by a unique sequence of three consecutive bases. Such a three-letter sequence is known as a codon. Some 64 codons are theoretically possible. The relationship between codons and α-amino acids is referred to as the genetic code (listed in Table 3.1), as it determines and guides the chemistry and biology of all organisms.

The genetic code was deciphered in the following way (M. Nirenburg and H.G. Korana, Nobel prize in Physiology and Medicine, 1968). Synthetic RNAs consisting of one type of codon only were prepared and introduced into a solution containing all 20 amino acids together with enzymes prepared from bacterial cells. It was found that the synthesis of a polypeptide chain consisting of units derived from a single amino acid then took place. For example, when synthetic RNA made only from uridylic acid (UUU sequence) was used, the polypeptide isolated contained exclusively phenylalanine residues. Thus, it was clear that the UUU sequence serves as a codon for phenylalanine.

Note that the genetic code is degenerate. This means that several codons can produce the same amino acid. However, in many cases the variations in the codons for any one amino acid involve only the third codon base. Moreover, three of the codons do not correspond to any amino acid. Their function is to encode for the termination of protein synthesis. When the protein molecule under construction on messenger RNA reaches a termination codon, synthesis is halted.

Table 3.1 *Genetic code[a] (Adapted from Pauling, L. and Pauling, P.,* Chemistry, *Freeman, San Francisco, 1975, with permission)*

First Letter	Second Letter: U	Second Letter: C	Second Letter: A	Second Letter: G	Third Letter
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA *	UGA *	A
	UUG Leu	UCG Ser	UAG *	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

[a]Phe = phenylalanine; Leu = leucine; Ser = serine; Tyr = tyrosine; Cys = cysteine; Trp = tryptophan; Pro = proline; His = histidine; Gln = glutamine; Arg = arginine; Ile = isoleucine; Met = methionine; Thr = threonine; Asp = aspartic acid; Lys = lysine; Val = valine; Ala = alanine; Asn = asparagine; Glu = glutamic acid; Gly = glycine. Asterisks denote codons for protein synthesis termination.

Amino acids locate the correct position on the messenger RNA with the help of transfer RNA. In a given organism there are as many tRNAs as there are codons. tRNAs are soluble and have relatively low molecular weights (about 20 000 Da). Their molecules are arranged in 'clover leaf' configurations (see Figure 3.8). Specific enzymes 'charge' the adenosine residue situated at the end of the tRNA with the appropriate amino acids (Figure 3.8, base 76). The ribose fragment of the terminal adenosine residue has two free hydroxy groups at the 2′ and 3′ positions. The latter hydroxyl group is believed to react with an amino acid carboxy group to form an energy-rich ester bond (Figure 3.13).

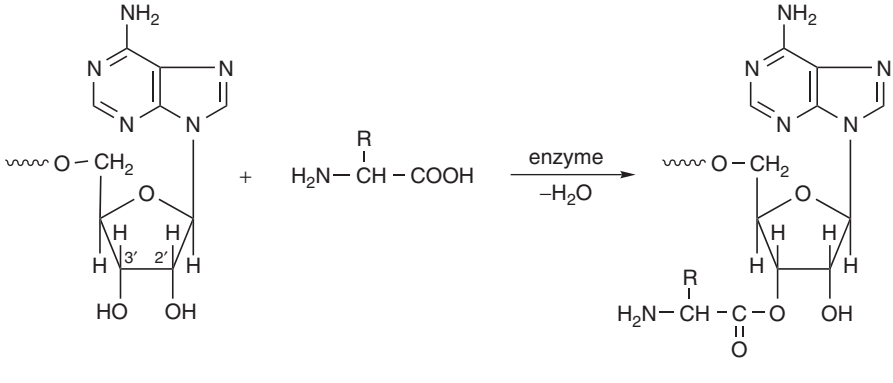


Figure 3.13 *Attachment of amino acids to transfer RNA.*

The amino acid laden tRNA is able to attach itself at the mRNA codon which corresponds to the amino acid through hydrogen bonding between the three heterocyclic bases of the codon and the complementary tRNA bases, named anticodons. For example, if the tryptophan codon is UGG, its tRNA anticodon will be CCA. The anticodons of the tRNA are situated in the central loop of the 'clover leaf.'

To commence protein synthesis, besides tRNAs and an mRNA, a ribosome is needed. A ribosome is a huge complex ($2.5–4.0 \times 10^6$ Da) composed of four ribosomal RNAs (65% of overall molecular mass) and nearly 80 proteins (35%). They form two subunits, one small and one large, which combine immediately prior to protein assembly where they visually resembles an inverted mushroom (Figure 3.14). Ribosomes play catalytic and supporting roles in protein synthesis and provide a physical site for the process. They are therefore sometimes referred to as 'protein assembly shops.'

Though both subunits act jointly, there exists a clear 'division of labor' between them. The small subunit is needed to support the mRNA in an unfolded state. The ribosome attaches to at least two mRNA codons while both the free amino acid and the growing protein chain are making the peptide bond. Two tRNAs, one charged with a single amino acid and the other with the growing polypeptide chain, are simultaneously attached to the mRNA and to the large ribosome subunit. The latter, mainly composed of RNAs, represents a surface where key stages of protein synthesis occur. To provide these functions the large subunit has two binding sites: A (aminoacyl) and P (peptidyl). One more binding site designated E (exit) is a place where tRNA or protein leave the ribosome. Protein assembly can be schematically illustrated using the β-galactosidase enzyme as an example (Figure 3.14).

At the initial step, a ribosome attaches to the mRNA to open the initiating codon AUG. This codon is complementary in sequence to anticodon UAC on tRNA carrying methionine. The methionine is delivered to the A-site where it immediately undergoes formylation by a special enzyme. As soon as the amino group of methionine is acylated the *N*-formylmethionine tRNA moves to the P-site. At this time the messenger RNA also shifts and the next site A codon thus becomes ready to receive the next tRNA with a new amino acid. Figure 3.14a depicts the stage at which the formylated methionine has just moved to site P, and the threonine-loaded tRNA has arrived at site A (codon ACC, anticodon UGG). In the following sequence, the true first step of the protein synthesis, the threonine amino group nucleophilically attacks the adjacent methionine carbonyl group causing cleavage of the CO—O ester bond between the methionine residue and its tRNA. The latter leaves the ribosome and the newly generated dipeptide moves, together with the threonine tRNA, to site P, simultaneously inducing a shift of the mRNA. Figure 3.14b demonstrates an example in which the tri-peptide OHC-NH-Met-Thr-Gly is formed, threonine-tRNA leaves the E-site and next the amino acid (designated N) enters site A. Thus, protein synthesis continues until the appearance of a termination codon at site A (Figure 3.14c shows the stop codon UAA). There is no tRNA for such a codon, but it is recognized by a special enzyme which hydrolytically cleaves the terminal ester bond. The completed protein molecule is then liberated and the ribosome dislodges the mRNA.

Each molecule of mRNA can serve as a template for the synthesis of hundreds of protein molecules. However, nature has adapted to prevent overproduction. To this end, one terminus of the mRNA strand has a site consisting of several hundred adenine residues (poly-A tail; see also Figure 3.21). After termination of the synthesis of each protein molecule, some portion of the site (containing probably two or three nucleotides) is severed. When the site is exhausted in this way, the mRNA molecule is no longer active and is rapidly degraded.

The process of formation of a peptide bond also deserves a short comment. It is assumed that after common nucleophilic addition of an amino group to a carbonyl group (Figure 3.15a), cleavage of the terminal ester proceeds with participation of a ribosyl 2′-OH group (Figure 3.15b) where

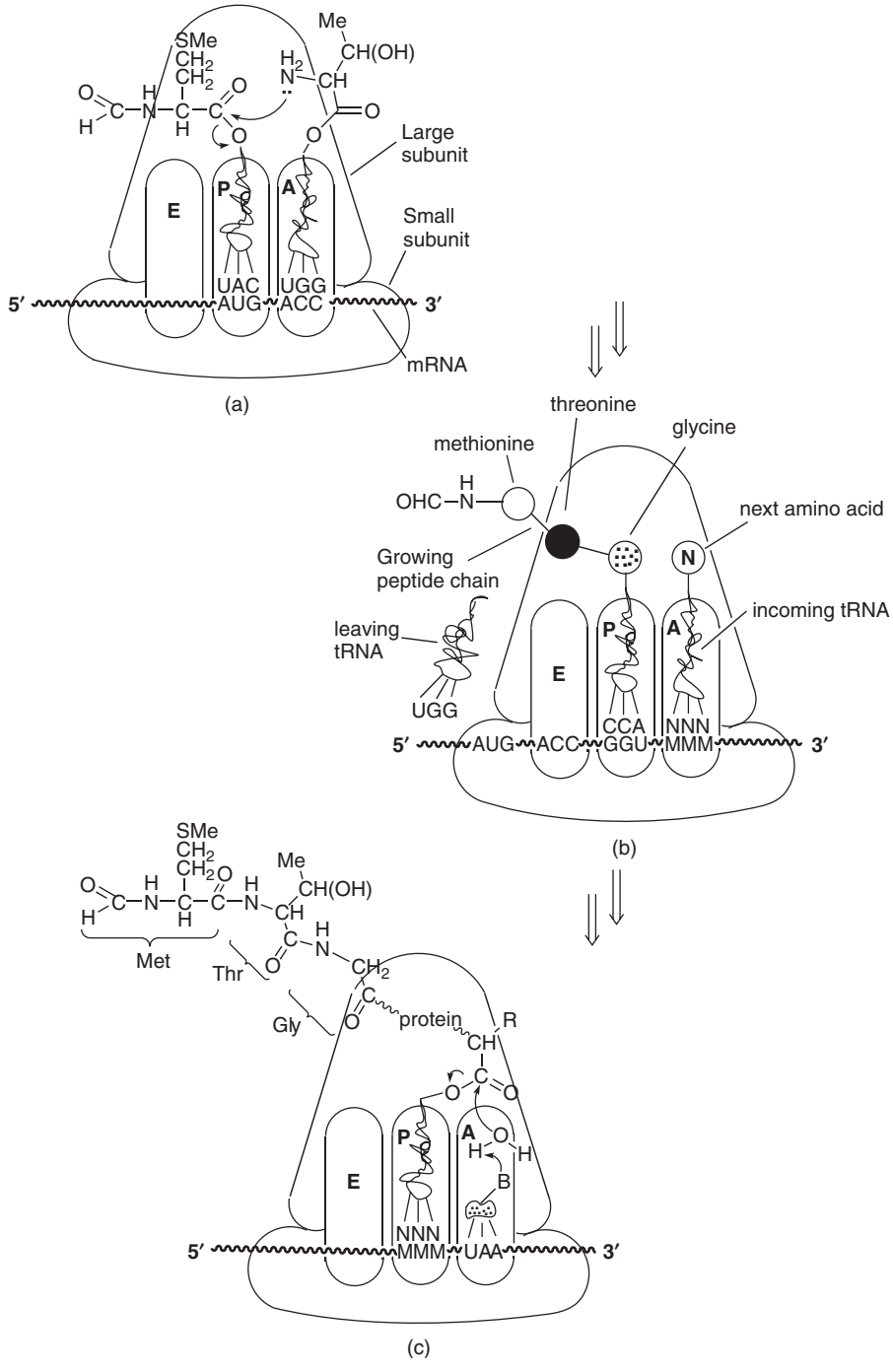


Figure 3.14 *Protein synthesis on a ribosome: (a) nucleophilic attack of threonine [Thr = MeCH(OH)] on N-formylmethionine (Met = CH_2CH_2SMe), (b) an intermediate stage with tri-peptide Met-Thr-Gly-tRNA attached to the P-site, (c) termination.*

the latter acts as a proton shuttle. The final step involves releasing transfer RNA[1] and attaching the growing polypeptide chain to the next transfer RNA[2] (Figure 3.15c).

In summary, it should be emphasized that many key details of ribosomal structure and protein synthesis were obtained by using the latest developments in electron microscopy and X-ray analysis.

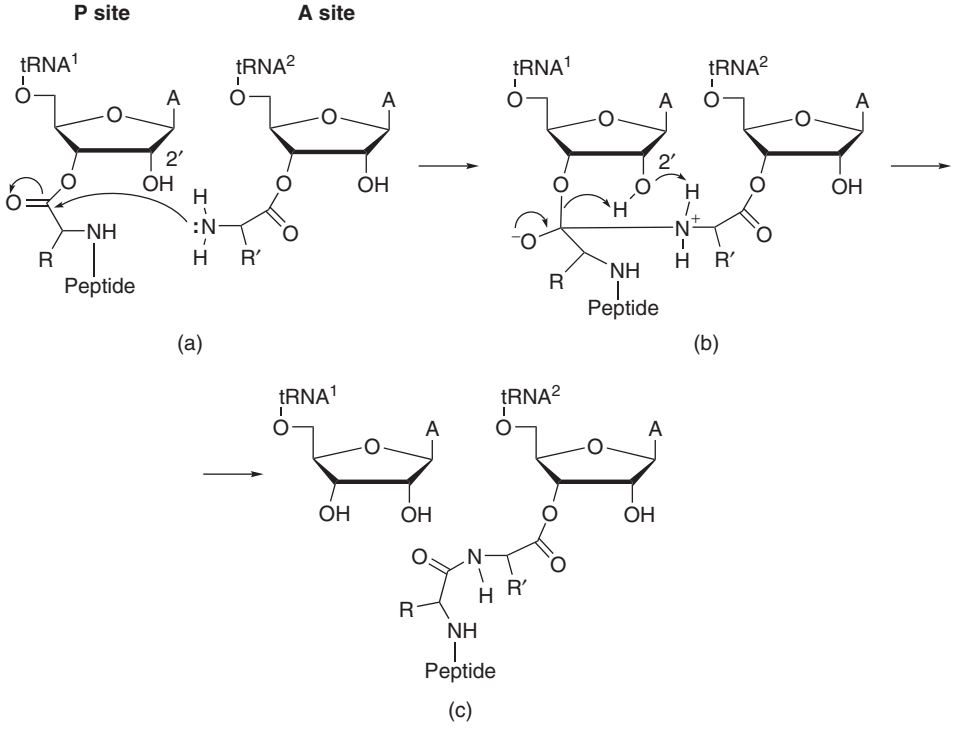


Figure 3.15 *The peptide transfer reaction catalyzed by the ribosome: (a) nucleophilic addition of an amino group to a carbonyl group, (b, c) proton-assisted cleavage of an ester group along with releasing of a tRNA[1] molecule. Reproduced with modifications from S. Borman,* Chem. Eng. News, *2007, Feb. 19, p. 14.*

3.5 What are Mutations?

In spite of the stability of biological species, individual living organisms are rather vulnerable as they are subject to continual unfavorable changes induced by various factors. From a chemical point of view, nucleic acids are the most fragile components of an organism. Their long strands are easily torn, even by simple agitation of a solution with a glass rod (see Section 3.2 for DNA 'sewing machine'). They are also susceptible to damage by many chemical reagents.

More serious chemical disorders are caused by changes in the normal sequence of the heterocyclic bases in the DNA chains. Such changes, called mutations, lead to the synthesis of abnormal proteins and, hence, often to metabolic disorders. What are the origins of mutations? First, there is a certain, purely statistical probability of an incorrect heterocyclic base being inserted into the DNA or mRNA chain undergoing synthetic assembly. We should note that this probability

is extremely low – not more than 10^{-7}. Moreover, there exists a special enzyme, part of the DNA polymerase, which 'oversees' chain construction and can eliminate erroneously incorporated nucleotides and replace them with the correct bases. This is known as proofreading. Nevertheless, because of the enormous scale of nucleic acid synthesis, a certain number of incorrect bases will remain unchanged in the chain.

The appearance of some spontaneous mutations may depend on tautomerism of the heterocyclic bases. Although these bases exist very predominantly (10^4–10^5:1) in the amino or oxo form (Figure 3.1), the less stable tautomer cannot be completely disregarded. Cytosine exists in physiological solution (water, pH 7) in equilibrium with a tiny proportion of the imino form (Figure 3.16a).

A similar, minuscule amount of the hydroxy form is present for guanine (Figure 3.16b). If cytosine and guanine exist in these minor forms at the time of DNA construction, the formation of abnormal pairs of bases becomes theoretically possible (Figure 3.16c, d). As in the normal pairs (Figure 3.6), these abnormal base pairs can be stabilized by hydrogen bonding and have the same dimensions as the normal Watson–Crick base pairs. But this subject is still open to investigation.

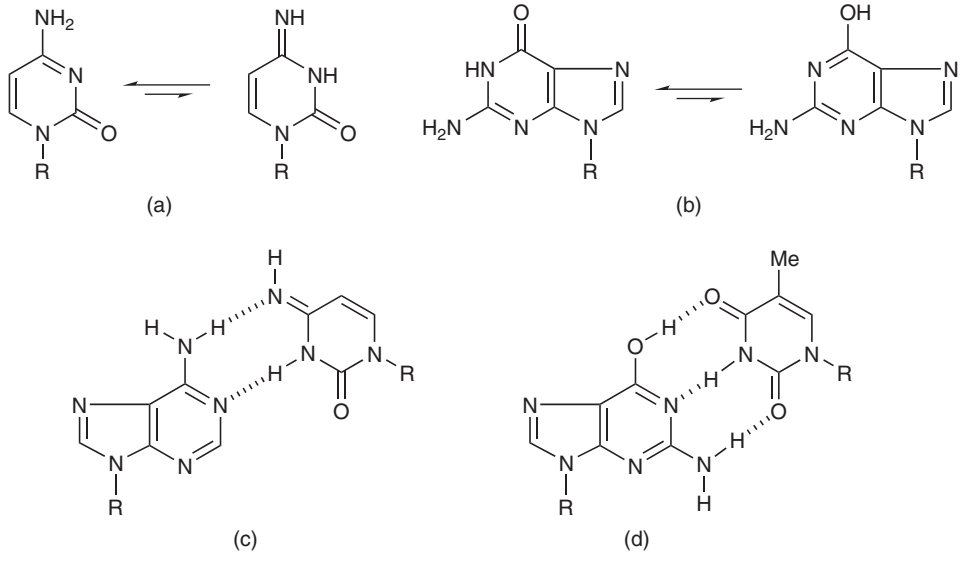


Figure 3.16 *Tautomerism in: (a) cytosine, (b) guanine. Formation of abnormal pairs: (c) adenine with cytosine in the imino form, (d) thymine with guanine in the hydroxy form. R = H or sugar.*

The most important causes of mutations are environmental effects. Penetrating radiation, for instance, brings about deep destruction of nucleic acids by numerous bond transformations. A typical example is the [2+2] cycloaddition of adjacent thymine residues leading to the formation of the thymine dimers (Figure 3.17a) and severe DNA damage. This can cause cell death (remember how our skin peels after sunburn) and even conversion of a healthy cell into a cancer cell.

Mutations are also the result of certain chemical compounds. Insufficiently tested medicines, environmental pollutants and food contaminants are among this class of mutagens. For example, nitrites (salts of nitrous acid) are added in very small quantities to meat products to impart a fresh appearance. Laboratory experiments have demonstrated that nitrous acid readily converts

adenine into hypoxanthine and the latter forms a base pair not with thymine but with cytosine (Figure 3.17b, c).

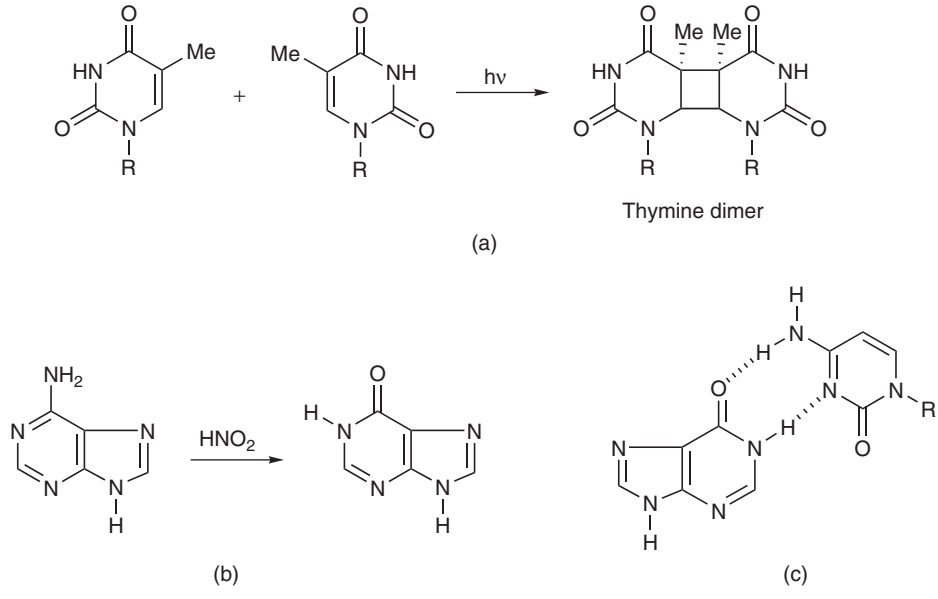


Figure 3.17 *(a) Light-induced cyclodimerization of thymine molecules. (b) Conversion of adenine into hypoxanthine. (c) Formation of the abnormal base pair hypoxanthine–cytosine. R = H or sugar.*

Organic chemists know well the high toxicity of such active alkylating agents as methyl iodide, dimethyl sulfate, methyl fluorosulfonate and diazomethane. Once introduced into an organism, these compounds alkylate N—H, O—H and S—H bonds and can also form quaternary salts with the biologically important cyclic bases (Figure 3.18). In such salts the bond between the nitrogen of the base and ribose (or deoxyribose) is readily cleaved, even by water, and the base can be lost from the RNA (or DNA) chain. Therefore, a mutagen acts like one of the mythical Parcae, cutting the thread of life.

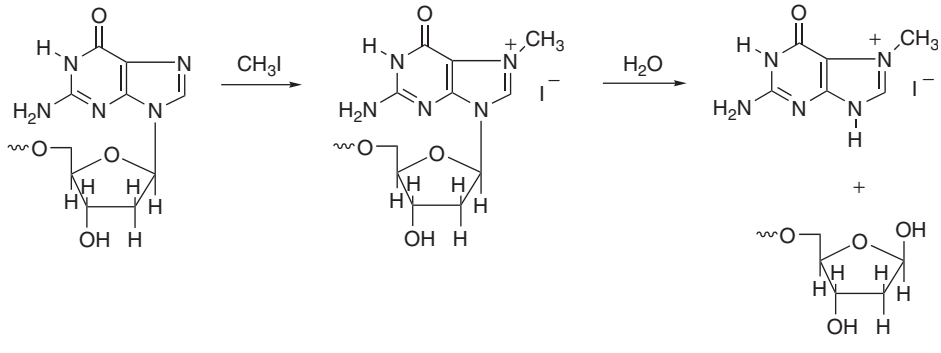


Figure 3.18 *Alkylation of the imidazole ring of deoxyguanosine and the hydrolytic decomposition of the quaternary salt.*

A further class of mutagenic compounds called 'intercalators' has also commanded the attention of scientists. These molecules have a planar polycyclic structure which gives them the possibility of penetrating the gap which exists between adjacent pairs of complementary bases in DNA. The guest molecules thus intercalated in the DNA tend to arrange themselves parallel to the plane of the bases. All polynuclear aromatic hydrocarbons and many heterocycles are potential intercalators. Under the effect of intercalation, the polynucleotide chain is strained like a spring and can become unwound in the vicinity of the defective site. The intercalator itself often serves as a matrix for the incorporation of a superfluous nucleotide into the DNA chain during its replication. Sometimes the effect of the intercalator can be rather specific. For example, the carcinogenic activity of 3,4-benzpyrene in living organisms is connected with its enzymatic oxidation to an epoxide which then intercalates into the DNA chain. This epoxide, being a strong alkylating agent, forms covalent bonds with amino groups in the nucleic acid (Figure 3.19).

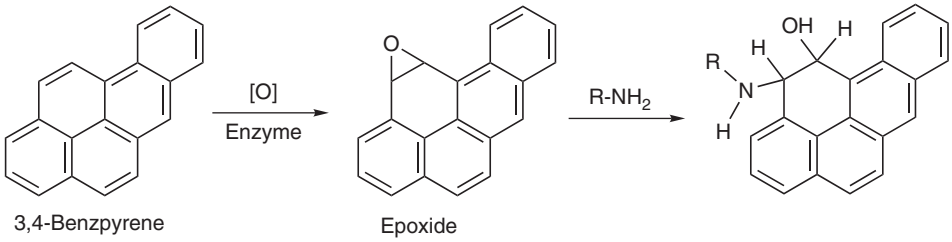


Figure 3.19 *The chemistry of the mutagenic action of 3,4-benzpyrene. R = residue of a heterocyclic base of a DNA chain.*

Both incorporation of additional bases into the DNA chain, and the loss of existing bases, can cause especially severe mutations. The net result is a frame shift, so that the wrong codon triplets are read by the protein-synthesizing machinery. The protein formed on the faulty messenger becomes unequipped to fulfill its natural biological functions or the chain is terminated early. However, under certain conditions, intercalators can be turned into life-saving drugs (see Section 7.4.3).

Milder effects are caused by point mutations when one pair of complementary bases in the DNA chain is replaced with another pair. In this case, only one different amino acid is included in the protein. If such an inclusion takes place far from the active site of the enzymatic protein, the mutation can be innocuous (a silent mutation). However, if the changed amino acid plays an important role in the protein's function, the mutation will induce functional disorders and consequently, illness. Fortunately, all organisms have a battery of 'repair' enzymes which monitor the DNA continually.

Gradual accumulation of mutations is thought to be one of the major causes of aging in living organisms. Unfavorable mutations can also be transmitted from parents to their young. Such a transfer may give rise to an inherited or genetic disease. At the present time, the chemical causes of hundreds of such molecular maladies have become clear, although most of these diseases still remain incurable. It is one thing to understand the cause; it is quite another to devise a cure. However, we can hope that genetic engineering will steadily bring us closer to the desirable goal of being able to repair genetic disorders.

One should draw attention to the recently found fact of the possibility of transducing some hereditary characteristics (tokens) without DNA participation, that is, without classic genes. So far this has been observed only in maize with violet corns and in grey mice which have a white tail tip. This phenomenon is named paramutation, and it has been shown that it might be controlled by RNA.

3.6 Mysterious Telomeres

DNA sequences are needed not only for genetic information storage. In living cells, DNA also accomplishes other functions. The most spectacular example is telomeres – the end regions of chromosomes consisting of many short repetitive DNA sequences or blocks. Each block includes not more than 6–8 nucleotides and carries no genetic information. For example, human telomeric DNA is comprised of TTAGGG blocks. One of the main functions of telomeres is capping chromosome ends, as demonstrated in Figure 3.7. Without such protection, chromosomes would not exist for a reasonable time due to slippage and resulting rearrangements. The existence of such telomeric caps is caused by the high content of guanine residues and their enhanced ability to participate in strong hydrogen bonding. Thus, four guanine units form approximately a flat tetramer called a G-quadruplex or G-tetrad. Each guanine fragment in the G-tetrad is linked with two of its neighbors by four hydrogen bonds that serve simultaneously as proton acceptors and proton donors. Additionally, G-quadruplexes are stabilized by a sodium or potassium ion that is chelated by highly polarized amide carbonyl groups (Figure 3.20a).

Flexible telomere strings can easily bend and make loops that bring the guanine fragments closer to each other, leading to the formation of a whole cascade of the G-quadruplexes resembling a book-stand (Figure 3.20b). Apart from their protective function, telomeres regulate a number of cellular processes that possibly influence aging and accompaning diseases such as cancer. The history of telomeres is too short to disclose all its secrets.

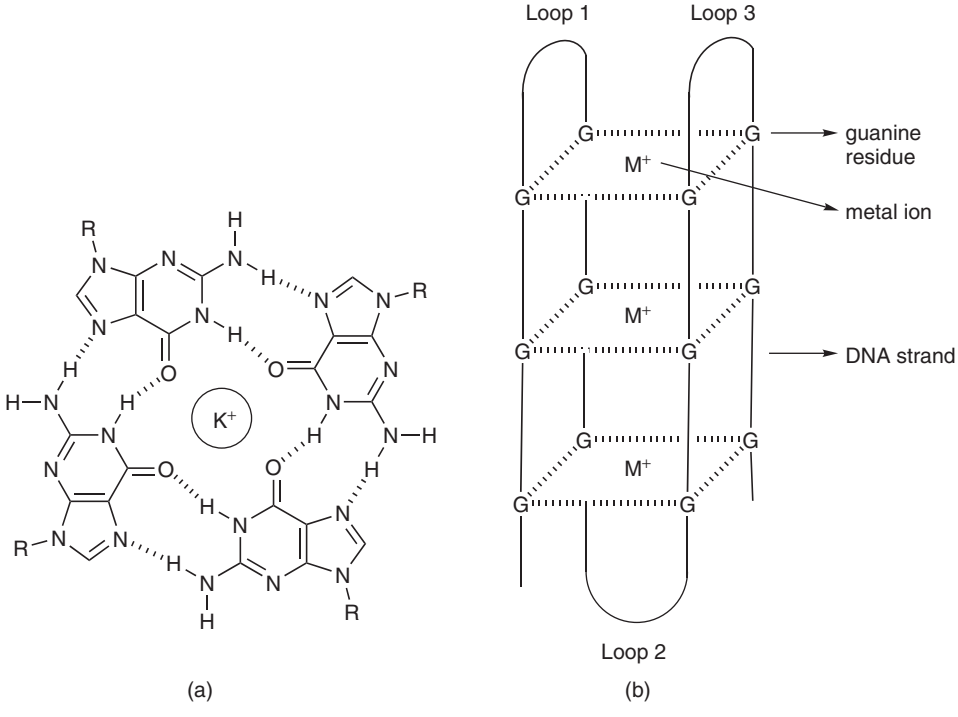


Figure 3.20 *Schematic representation of a G-quadruplex structure: (a) noncovalent bonding of four guanine residues, (b) cascade of G-tetrads formed by looping telomere DNA strands. Retrieved with modifications from en.wikipedia. Author: original uploader was Harold f at en.wikipedia. http://en.wikipedia.org/wiki/File:G-quadruplex.gif (accessed 29 October 2010).*

3.7 Gene Expression

Let us imagine orchestra instruments resting in their cases. To make them play, musicians who take out the instruments, tune them and then extract sounds using bows, or their fingers and lips should appear. The musicians also use notes, follow the movements of a conductor and sense the audience's mood. Something similar happens with DNA molecules buried deep inside chromosomes, but of course in a much more complex manner. The 'sounds' of DNA molecules, or more precisely the 'sounds' of their genes can be considered as a synthesis of functional RNAs (transcription), protein synthesis (translation) and post-translation modification of proteins. All together this is called 'gene expression' (GE) – apparently, the most crucial and sophisticated of biochemical phenomenon. Many of the latest achievements in biochemistry are connected with gene expression. Needless to say, heterocyclic nucleobases play an enormous role in GE.

Before transcription starts, the piece of DNA corresponding to the gene must be uncoiled, activated and programmed for the synthesis of the exact amount of the particular protein that the living organism needs at the moment. At the same time, the activity of many other genes may be repressed or even temporarily switched off. The corresponding regulating and controlling functions are performed by numerous proteins, RNAs and metabolites via interaction with specialized DNA-binding sites. Quite recently, it has been realized that GE is considerably influenced by environmental factors, which can cause structural and chemical changes in a genome. These last concepts are at the center of epigenetics – now an extensively developing field of science. The opinion exists that epigenetics is even more important for an organism's development and survival than purely genetic factors (nucleobase sequence).

In general, there are two ways for regulating genome activity: inducing GE and repressing GE. Permanently arising waves of gene activation–deactivation involve various signal molecules and special DNA and RNA sites, called promoters and repressors. These sites are commonly situated near the coding gene sequence. Quite often, signal molecules influence GE indirectly. For instance, they can first interact with mRNA, which then transmits the action back to the DNA and thus initiates the synthesis of necessary compounds. Notably, DNA–RNA transcription does not lead directly to mature mRNA. Primarily, so-called pre-mRNA is formed which contains, along with useful polynucleotide sequences (exons), non-useful sequences (introns; Figure 3.21a). The introns arise when noncoding parts of DNA are transcribed. Living cells have developed special mechanisms, called splicing, to remove introns. The splicing allows several exons to join with the formation of the uninterrupted coding region. Thus is formed mature mRNA ready for translation.

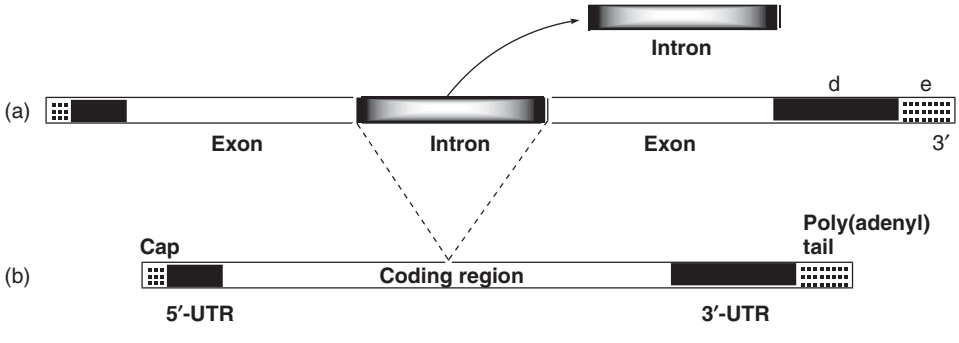


Figure 3.21 *Schematic representation of sections in: (a) pre-mRNA, (b) mature mRNA. Retrieved with modifications from: http://en.wikipedia.org/wiki/File:Pre-mRNA_to_mRNA.svg (accessed 29 October 2010).*

The structure of mRNA is rather complex and includes, in addition to the coding region, several other sections (Figure 3.21b). These are needed for protection of the mRNA, for its binding with ribosome and target metabolites and so on. The 5′ and 3′ ends of mRNA are protected with a cap and poly(A) tail, respectively. The former consists exclusively of guanosine residues; the latter represents a stretch of RNA, composed only of adenosine monophosphates. Both the end protectors are separated from the coding regions by relatively short untranslated regions, designated as 5′-UTR and 3′-UTR which fulfill mostly binding and signaling functions. For example, mRNA binds to ribosome via 5′-UTR. A landmark discovery was a recent finding in 5′-UTR of the so-called riboswitches. Riboswitch is a part of mRNA that can bind a small target molecule and this binding then affects the gene's activity. Through riboswitch the mRNA senses small-molecule metabolite concentrations and performs its own regulation. There are many types of riboswitches. Thus, different forms of purine riboswitches bind specifically either guanine or adenine. This specificity depends completely upon Watson–Crick nucleobase pairing. Such interactions activate genes responsible for producing enzymes and proteins on which purine biosynthesis and transport depend.

A fundamental type of chemical regulation of GE is methylation of cytosine residues in the DNA-coding sequence. This methylation occurs exclusively at the 5-position of cytosine nuclei in repeating sequences of cytosine–guanine (CpG). Several types of methyltransferase enzyme carry out the methylation. All of them use a reactive methyl group bound to sulfur in S-adenosyl methionine (Figure 3.22, structure **1**).

Two genome DNA regions can be subjected to such cytosine methylation: the low-density CpG region and the so-called CpG islands with a much higher concentration of CpG dinucleotides. While the former is connected with ordinary DNA–RNA transcription, CpG islands are associated with the promoter region of 76% of human genes. Evolutionary nature created cytosine methylation process to cause gene silencing in response to changes in the internal and external environment. It is needed, for example, to modulate stem cells differentiation, to avoid transcription of transposons and 'junk DNA' or to provide the organism's survival under unfavorable conditions. Moreover, methylation protects DNA from enzymatic cleavage, since restriction enzymes are unable to recognize modified cytosine residue.

There are genetic consequences of cytosine methylation. 5-Methylcytosine performs much like cytosine, pairing up with a guanine, so in the next round of cell division it is replaced with a regular cytosine. However, in biochemistry hydrolytic deamination reactions are widespread. When cytosine enters such a reaction to produce uracil (not a DNA nucleobase), this mutation is recognized by the special repairing enzyme, which replaces incorrect uracil by another cytosine. This is not so for 5-methylcytosine, whose deamination results in thymine (Figure 3.22). Unlike cytosine, thymine pairs with adenine. Since thymine is a regular DNA base, no mechanism exists to recognize and repair this error and thus the 5-methylcytosine-thymine replacement represents the most common single nucleotide mutation.

Normally, cytosine methylation can be useful only when the low-density CpG dinucleotides are extensively methylated (Figure 3.23a). However, hypomethylation can result in overexpression of growth factors and development of oncogenes (Figure 3.23b). The most dangerous consequences arise at hypermethylation of CpG islands since these lead to the complete silencing of tumor suppressor genes (Figure 3.23c).

The mechanism by which cytosine methylation regulates the transcriptional activity of genes is not yet understood in detail. Presumably, it is somehow connected with the remodeling of chromatin – complexes of DNA and the histone proteins. The methylation can influence DNA conformation and therefore the way that DNA is wrapped around histones. This result can also

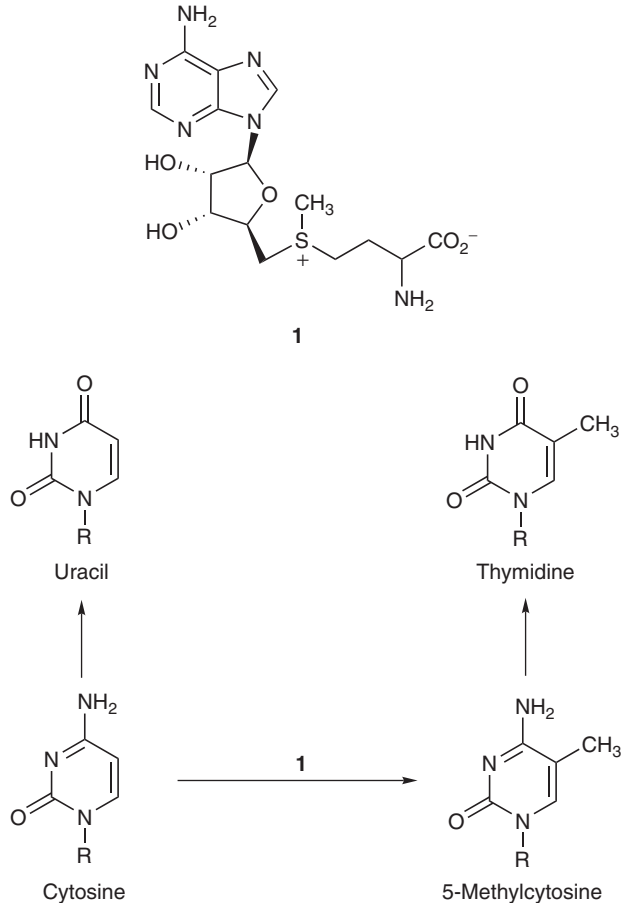


Figure 3.22 *Genome methylation of cytosine with methyltransferases (coenzyme **1** is shown) and deamination of cytosines into uracil and thymidine.*

change gene expression. Post-translational modification[2] of histone proteins is another way of gene regulation having similar consequences on GE.

One more general mechanism of regulation of GE was discovered in the late 1990s (A. Fire and C.C. Mello, Nobel Prize in Physiology and Medicine, 2006). It was termed as RNA interference (RNAi) since several types of RNA molecules cooperatively participate in the process. So-called small interfering RNA (siRNA) and microRNA (miRNA) which can bind specifically to other RNAs (e.g., mRNA, and either increase or decrease their activity by fully preventing protein production) are of central importance. The RNAi was developed in nature by evolution as an effective way to defend host cells against transposons and viral infection. Schematically, the mechanism of RNA interference is shown in Figure 3.24. In response to the appearance in a

[2] Post-translational modifications include relatively simple, sequence-independent chemical functionalizations of proteins such as acetylation, methylation, phosphorylation and so on. They commonly occur at N-termini of histones. This kind of gene regulation is considered as almost purely epigenetic.

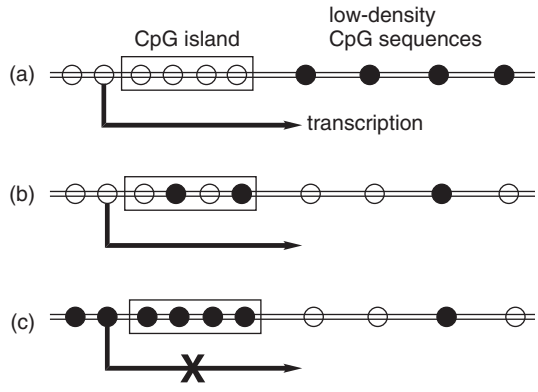


Figure 3.23 *Methylation pattern and replication mode of genome DNA: (a) normal gene, (b) hypomethylated, normally imprinted gene, (c) hypermethylated gene (white and black circles refer to unmethylated and methylated CpG dinucleotides). Adapted from Bryan, J.N., Taylor, K.H., Henry, C.J., Selting K.A., Rahmatpanah, F., Lewis, M.R. and Caldwell, C.W.,* Cancer Therapy, *2008,* ***6****, 137.*

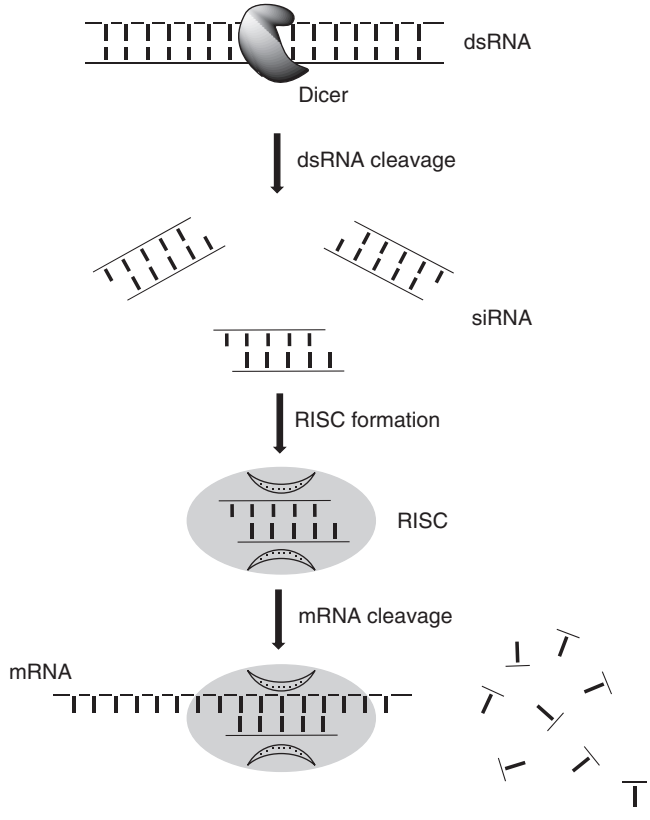


Figure 3.24 *Simplified scheme of RNA interference. Adapted with permission from Mocellin, S. and Provenzano, M.,* J. Transl. Med.*, 2004,* ***2****, 39. © 2004 BioMed Central.*

living cell of an exogenic double strand RNA (dsRNA), the cellular enzyme Dicer (which belongs to ribonuclease family) binds to dsRNA and cleaves it into short pieces of siRNA of some 20 nucleotide pairs each. These pieces of RNA pairs bind to a cellular enzyme known as RNA-induced silencing complex (RISC). The RISC uses one strand of the siRNA to bind to a complementary sequence of mRNA or to another single-stranded RNA molecule. Due to its nuclease activity, the RISC then destroys the mRNA, thus silencing expression of the foreign gene.

The miRNAs provide somewhat different level of RNA-controlled gene expression. Unlike siRNA, the miRNAs are endogenic products, which are formed by DNA transcription of the same organism. Another difference between them is that whereas siRNAs have perfect base pairing, miRNAs base pair incompletely. As a result, siRNA induces mRNA cleavage only in a single, specific place. In contrast, miRNA can bind to many different mRNA with complementary sequences. This helps to regulate gene expression broadly, particularly during the development of an organism.

Further topics dealing with the chemistry, technology and origin of nucleic acids are discussed in Sections 4.4, 11.3.2 and 12.6.8.

3.8 Problems

1. (a) Draw the structure of the polypeptide which corresponds to the sequence poly(ACUGU). (b) What complements poly(ACUUG)?
2. One chain of double-helical DNA which codes for the amino acid sequence in silk protein contains 1685 nucleotides (gene E2) including 588 adenines and 484 guanines. (a) How many codons does the DNA chain contain? How many amino acid residues does the silk protein chain incorporate? (b) What is the relative molar ratio of nucleotides in this DNA chain? In the complementary chain?
3. How long must a DNA section (called a cistron) be to code for the formation of a protein containing 200 amino acid residues?
4. A single DNA chain contains this $5'-3'$ sequence:

   ATCGTCGACGATGATCATCGGCTACTCGA.

   Determine: (a) the base sequence in the complementary DNA chain, (b) the base sequence in the mRNA translated from the first chain of DNA and (c) the sequence of amino acid units in the corresponding polypeptide.
5. Formaldehyde is a toxic compound which causes denaturation of DNA. Which DNA functional groups are most vulnerable to attack? Propose a mechanism for the chemical interaction.
6. What is the minimum number of nucleotides that must be changed or incorrectly translated in order to form the defective hemoglobin responsible for sickle-cell anemia? (As a clue, the β-chain in normal hemoglobin contains the following sequence of *N*-terminal amino acids: ValHisLeuThrProGluGluLys. In patients suffering from sickle-cell anemia the corresponding sequence is ValHisLeuThrProValGluLys.)
7. Certain proteins contain amino acid residues which have no corresponding codon in the genetic code. For example, collagen contains hydroxyproline and hydroxylysine residues. Suggest a possible mechanism for insertion of the 'noncoded' amino acids into a polypeptide chain.
8. The first totally synthetic triplex DNA (triple-stranded form), synthesized in 1957, was of no practical importance. Its potential use as 'molecular scissors' to cut DNA, which could be useful in mapping and sequencing the human genome or in blocking viral reproduction by disrupting gene transcription, is now being explored. In 1987 it was established that a short third synthetic strand (11–15 nucleotides) could be attached to the natural double DNA helix. The binding is very selective and a short strand of synthetic DNA can identify its target from millions of base pairs. Draw the triple-helical complex formed when the synthetic third strand of DNA (comprising 20 nucleotides and consisting of only two pyrimidine bases), TTTTTCTTTCTTCTTTTCTT, is attached to the following portion of the human gene only by one of its two strands: $5'\ldots$-ACGGATCCTTTTTCTTTCTTCTTTTCTTC-$\ldots 3'$. (Note: the binding follows normal chemical rules.)
9. Spider dragline silk has a tensile strength and elasticity much greater than that of steel. The complete sequence of 655 amino acid residues has been determined (A = alanine, G = glycine, L = leucine, etc.; see Xu, M. and Lewis, R. V., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 7120) as:

   QGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQGAG
   QGAGAAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAA
   AGGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAAAGGAGQ

GGYGGLGNQGAGRGGQGAAAAAAGGAGQGGYGGLGSQGAGR
GGLGGQAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQ
GAGRGGLGGQGAGAAAAAAAGGAGQGGLGGQGAGQGAGAS
AAAAGGAGQGGYGGLGSQGAGRGGEGAGAAAAAAGGAGQG
GYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAGGA
GQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRG
GLGGQGAGAVAAAAAGGAGQGGYGGLGSQGAGRGGQGAGA
AAAAAGGAGQRGYGGLGNQGAGRGGLGGQGAGAAAAAAAG
GAGQGGYGGLGNQGAGRGGQGAAAAAAGGAGQGGYGGLGSQG
AGRGGQGAGAAAAAAVGAGQEGIRGQGAGQGGYGGLGSQGS
GRGGLGGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAA
AGGVRQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYG
GLGGQGVGRGGLGGQGAGAAAAGGAGQGGYGGVGSGASAAS
AAAASRLSS.

(a) Find the two amino acids which occur most often in the sequence and calculate their overall percentage.
(b) Find similar or repeating units in the sequence and arrange the sequence in a scheme more suitable for analysis.
(c) Determine the first and second bases of the codons for the two most abundant residues. Calculate their overall percentage.

10. How you would explain the fact that biological organisms living at very high attitudes have a lower proportion of thymine residues in their genome?
11. Which base pairing is stronger, AT or GC? Correlate your answer with the fact that genome DNA promoter regions tend to have a high AT content.
12. Express an opinion about a possible reaction mechanism of (a) cytosine methylation and (b) cytosine deamination (Figure 3.22). Whose reactivity in a deamination reaction should be higher, cytosine or 5-methylcytosine?
13. A classical method to distinguish between α- and β-forms of glucose or ribose (see Figure 3.2) is their reaction with a boronic acid: α-forms produce a heterocyclic compound while β-forms remain unreacted. Explain and write down a structure of the reaction product.

3.9 Suggested Reading

1. Lehninger, A. L., Nelson, D. L. and Cox, M. M., *Principles of Biochemistry*, 4th edn, Worth, New York, 2004.
2. Wen-Hsing, L. and Graur, D., *Fundamentals of Molecular Evolution*, Sinauer, Sunderland, Mass., 1991.
3. Hawkins, J. D., *Gene Structure and Expression*, 4nd edn, Cambridge University Press, Cambridge, 1996.
4. Townsend, L. B. and Tipson, R. S. (eds), *Nucleic Acid Chemistry. Improved and New Synthetic Procedures, Methods and Techniques*, Wiley, New York, 1991.
5. Townsend, L. B. (ed.), *Chemistry of Nucleosides and Nucleotides*, vol. 2, Plenum Press, New York, 1991.
6. Neurath, H. (ed.), *Perspectives in Biochemistry*, vol. 2, American Chemical Society, Washington, D.C., 1991.
7. Wyman, J. and Gill, S. J., *Binding and Linkage: Functional Chemistry of Biological Macromolecules*, University Science, Mill Valley, Calif., 1990.

8. Blackburn, G. M. and Gait, M. J. (eds), *Nucleic Acids in Chemistry and Biology*, 2nd edn, Oxford University Press, Oxford, 1996.
9. Stryer, L., *Biochemistry*, 6th edn, Freeman, New York, 2006.
10. Watson, J. D., *The Double Helix; A Personal Account of the Discovery of the Structure of DNA*, Norton, New York, 1980.
11. Neidle, S. and Balasubramanian, S. (eds), *Quadruplex Nucleic Acids*, Royal Society of Chemistry, Cambridge, 2007.
12. Glick, B. R. and Pasternak, J. J., *Molecular Biotechnology*, 2nd edn, ASM Press, Washington, D.C., 1998.
13. Ribas de Pouplan, L. (ed.), *The Genetic Code and the Origin of Life*, Kluwer, Amsterdam, 2004.

4

Enzymes, Coenzymes and Vitamins

Can't you see, good friend of mine,
All that's sensed with eye and ear
Is but shadow, reflection
Of th'unseen by us, my dear?
V. Solov'ev

A multitude of different chemical processes occuring constantly in every living cell involve the synthesis and breakdown of millions of complex organic molecules. Scientists are fascinated by the cell as a biosynthetic machine because of the variety of functions carried out and the amazing level of coordination between all the different components. The rapid rates of reaction and the mild conditions under which these conversions take place in the organism (neutral aqueous media, temperatures from about 36 to 40 °C, normal pressure) are of particular interest. A further remarkable characteristic is the stereoselectivity of all these biochemical reactions which leads specifically to the formation of molecules with a single spatial configuration.[1] To attain the synthetic sophistication of nature has long been the ultimate goal, one toward which we are moving constantly and steadily, if slowly. The enzymes which are specific biocatalysts are of major importance in these processes. We now consider the nature of enzymes and their relationship with heterocycles.

4.1 Molecular Robots

Readers will be familiar with the manner in which robots work from real life, the cinema or TV. A robot takes an object and manipulates it according to a prearranged program. Enzymes function likewise as natural micromanipulators and therefore can be considered as molecular robots. An enzyme captures a molecule of reactant, conveys it to the reaction center, appropriately orients the

[1] The majority of biomolecules are chiral, that is, their structures possess neither an axis nor a plane of symmetry. Such molecules exist in the form of stereoisomers (optical isomers or mirror image isomers) which relate to each other as the left hand does to the right. Almost all chemical processes in organisms take place with the participation of single, strictly determined stereoisomers, the 'left-handed' or 'right-handed' molecules.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

molecule in space and if necessary activates it. The enzyme then ejects the newly formed product from the reaction zone, liberating the site for another incoming molecule of starting material.

As a rule, enzymes are proteins[2] with molecular weights up to hundreds of thousands or even millions. Enzymes can be composed of one or several polypeptide chains attached to and wound around each other in a three-dimensional manner which is determined by noncovalent interactions. This tertiary structure of enzymes means that their molecules contain clefts, pockets and/or trenches on the surface. The active site of an enzyme is one such cleft into which a specific reactant may enter 'as a key enters a lock'. Enzymes operate only with their particular 'key molecule'; that is, the enzyme can function only with a rigidly defined type of molecule (substrate). Each chemical reaction in an organism demands a specific enzyme. It is now clear why so many types of enzymes exist; at present, several thousand have been characterized.

The histidine residue is a constituent of the active site in many enzymes (Figure 3.12). The imidazole ring of histidine has a series of unique properties which enable it to show catalytic activity. First, the rather high basicity enables histidine both to form strong hydrogen bonds and also to abstract a proton from acidic groups, such as the OH group of water and alcohols. As an RO^- anion is a much stronger nucleophile than a neutral ROH molecule, the imidazole ring can catalyze a nucleophilic addition to a carbonyl group (Figure 4.1). This variety of catalysis is called general base catalysis.

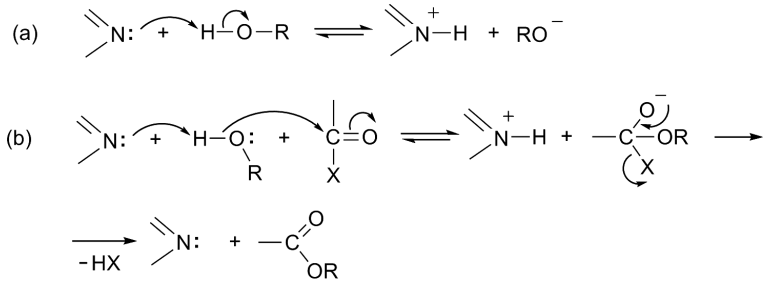


Figure 4.1 *(a) Ionization of the alcoholic hydroxy group. (b) General base catalysis with participation of the imidazole ring (the pyridine-like nitrogen atom is designated).*

In an organism, this type of catalysis is represented by the hydrolytic cleavage of protein amide bonds. The participating enzymes are called proteases. Histidine and serine residues are constituents of the protease chymotrypsin. A proposed mechanism for the action of chymotrypsin is shown in Figure 4.2. Within the enzyme, the imidazole ring of a histidine and the hydroxy group of a serine are bound together by hydrogen bonding (Figure 4.2a). When a protein molecule approaches, the imidazole nitrogen abstracts a proton from the OH group, thus activating the serine oxygen atom toward attack at the carbonyl carbon atom in the polypeptide (Figure 4.2b). The unstable activated complex, shown in Figure 4.2b, is called an enzyme–substrate complex. Further conversion, very similar to that represented in Figure 4.1, brings about cleavage of the amide bond and acylation of the enzyme at its hydroxy group (Figure 4.2c). By an analogous

[2] One of the most important biochemical discoveries of the last quarter of a century is that enzymatic activity is also inherent to polyribonucleotides – the so-called ribozymes (for details see Section 4.4).

catalytic mechanism, subsequent hydrolysis of the ester bond occurs (Figure 4.2d) with elimination of acid (RCO_2H) and regeneration of the enzyme (Figure 4.2e).

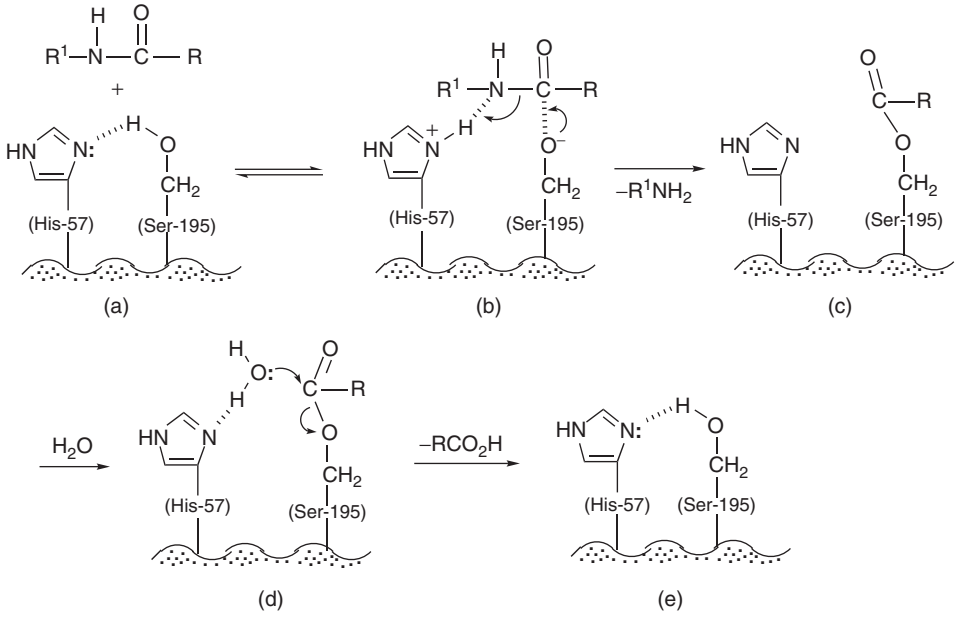
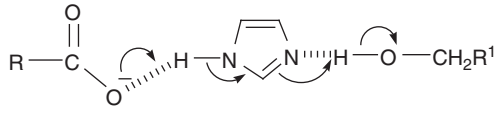


Figure 4.2 *Simplified scheme of amide bond hydrolysis catalyzed by chymotrypsin. Histidine (His-57) and serine (Ser-195) take part in the enzyme active site (numbers designate the positions of the amino acids in the protein chain). Adapted from Lehninger, A. L.,* Biochemistry*, Worth, New York, 1970, Chap. 9, p. 173, Figure 9.2, with permission.*

The scheme in Figure 4.2 is oversimplified. In reality, a third amino acid residue is involved in the active site of chymotrypsin. This is aspartic acid (or α-aminosuccinic acid). In the normal state, its ionized carboxy group and the NH proton of a histidine imidazole ring are connected by a hydrogen bond. However, in hydrolytic reactions the NH proton is donated to aspartic acid in order to increase the negative charge on the imidazole ring. This transfer facilitates the abstraction of a proton from the serine OH group. All the proton transfers are believed to be carried out synchronously through transition states which might be illustrated as:



The imidazole ring in chymotrypsin functions in an amphoteric manner, that is, both as an NH acid and as a base. Significantly, the basicity of histidine is such that it exists 50% as the neutral form and 50% as the imidazolium cation at the physiological pH of 7.4.

This peculiarity is utilized by the enzyme ribonuclease, which catalyzes P—O phosphate bond cleavage at the 3′ position of ribose (Figure 4.3). The specificity of this enzyme is such that it only severs those ribose residues which are attached to pyrimidine bases. Two histidine residues are found at the active site of ribonuclease, one in the neutral form (His-12) and the other (His-119) in the imidazolium cation form (Figure 4.3a). The neutral histidine abstracts a proton from the 2′ OH group of ribose in the enzyme–substrate complex (Figure 4.3b). The oxygen atom is thus activated and can perform intramolecular attack at the electrophilic phosphorus atom to form a cyclic phosphate diester. The P—O bond cleavage occurs simultaneously with elimination of an RNA^1—OH fragment. This process is catalyzed by the imidazolium ion of His-119 which donates its proton to the departing residue of RNA^1. The final step in the overall process involves hydrolysis of the cyclic phosphate diester (Figure 4.3c). By contrast with the initial reaction sequence (Figure 4.3b), this time the imidazolium ion of His-12 plays the role of proton donor in P—O bond cleavage, and the neutral imidazole of His-119 serves as a proton acceptor which activates a water molecule for nucleophilic attack at phosphorus. At the completion of the sequence, the enzyme molecule is regenerated to its initial state and the RNA chain is one unit smaller (Figure 4.3d).

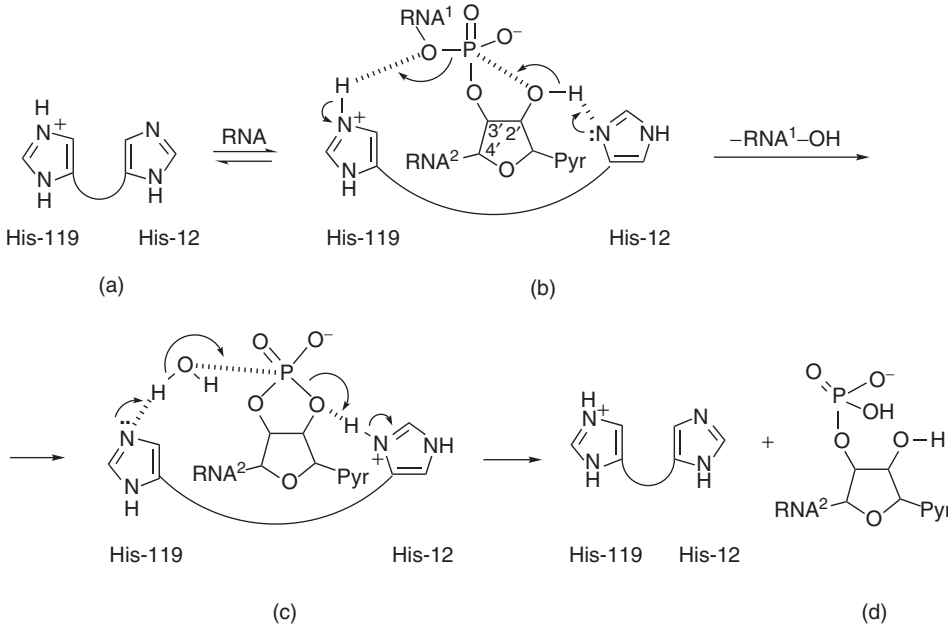


Figure 4.3 *Catalysis of the hydrolytic cleavage of a P—O bond in RNA assisted by the enzyme ribonuclease: (a) a fragment of the enzymatic active site, (b) the enzyme–substrate complex, (c) hydrolysis of the intermediate cyclic phosphate diester, (d) regeneration of the enzyme and liberation of an RNA fragment. RNA^1 and RNA^2 are different portions of the RNA chain; Pyr is a pyrimidine base.*

4.2 Coenzymes and Enzymes as 'Joint Molecular Ventures'

In addition to enzymes which possess purely protein structures (chymotrypsin, ribonuclease, etc.), there are a variety of enzymes which incorporate alongside the protein structure non-amino acid

fragments, called coenzymes or cofactors. In these cases it is the coenzyme that facilitates the necessary chemical reaction by interacting with the substrate. The role of the protein moiety (called the apoenzyme) is limited to the spatial organization of the overall process. The coenzyme is positioned in a superficial cavity of the apoenzyme and is held in place by noncovalent interactions. Sometimes, in addition, a covalent sulfide (C—S—C) or disulfide (C—S—S—C) bond is formed between the coenzyme and apoenzyme.

All reactions of coenzymes can be divided into two groups: oxidation–reduction and transfer reactions. The majority of coenzymes are derivatives of nitrogen heterocycles: we now consider some of their reactions.

4.2.1 Oxidative–Reductive Coenzymes

Many oxidative–reductive conversions occur in organic chemistry. Hydrogenation–dehydrogenation reactions form an important subclass among these (Figure 4.4a). A molecule is oxidized if it loses a hydrogen atom together with a pair of electrons, which can be in the form of an H_2 molecule or a hydride ion H^-, but not a proton or hydrogen atom. If a molecule gains a hydrogen, it is reduced. In principle, hydrogen atoms can be eliminated in pairs as H_2, for example, by heating or under the action of a porous metallic catalyst like palladium, platinum or nickel. However, more frequently the hydrogen undergoes transfer to a hydrogen acceptor molecule known as an oxidant (A) or is donated by a hydrogen donor molecule known as a reductant (AH_2). The most abundant natural oxidant is molecular oxygen, which converts hydrogen into water or, less often, into hydrogen peroxide. Alternatively, incorporation of oxygen into a molecule may occur without any change in the original number of hydrogen atoms. For example, a C—H bond may be transformed into a C—OH group (Figure 4.4b). Such reactions are also considered as oxidations on the grounds that the electron pair of a C—O bond is shifted toward oxygen (in comparison with that of the C—H bond), and thus the carbon atom loses some electron density.

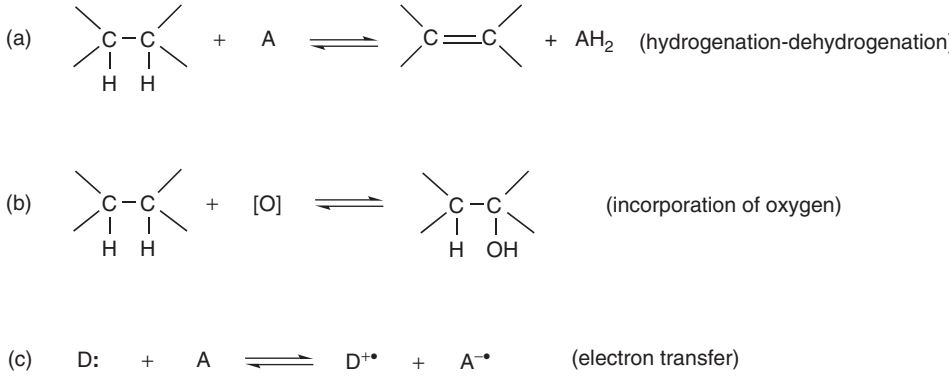


Figure 4.4 *Types of oxidation–reduction reactions.*

One further variant of oxidative–reductive conversions involves the gain or loss of electrons by a molecule without any change in the connectivity of the atoms (Figure 4.4c). This type of interaction occurs between an electron donor (D) and an electron acceptor (A). The products of such an electron transfer are a cation-radical $D^{+\bullet}$ and an anion-radical $A^{-\bullet}$ which may be bound in a molecular complex D→A. All of the above types of reactions are widely encountered in the chemistry of living organisms.

Dehydrogenases as Hydrogen Transporters

Two coenzymes, nicotinamide adenine dinucleotide (NAD) and the analogue phosphorylated at one of the hydroxy groups (NADP), are the most biologically important hydrogen transfer agents (Figure 4.5a, b).

The functional portion of both coenzymes, their 'chemical motor', is the nicotinamide residue which can exist as either: (i) the oxidized pyridinium form (NAD^+ or $NADP^+$; as shown in Figure 4.5a, b) or (ii) the reduced dihydropyridine form (NAD-H or NADP-H; as shown in Figure 4.5c). The function of the oxidized form of the enzyme is to receive a hydride ion from a hydrogen-donor substrate (Sub-H_2). In the process, the pyridinium cation is converted into a 1,4-dihydropyridine (Figure 4.5c). The second hydrogen atom of the substrate is released as a proton into solution. As with all enzymatic reactions, this process is reversible and the equilibrium can be shifted depending on the substrate type and the conditions within the living cell.

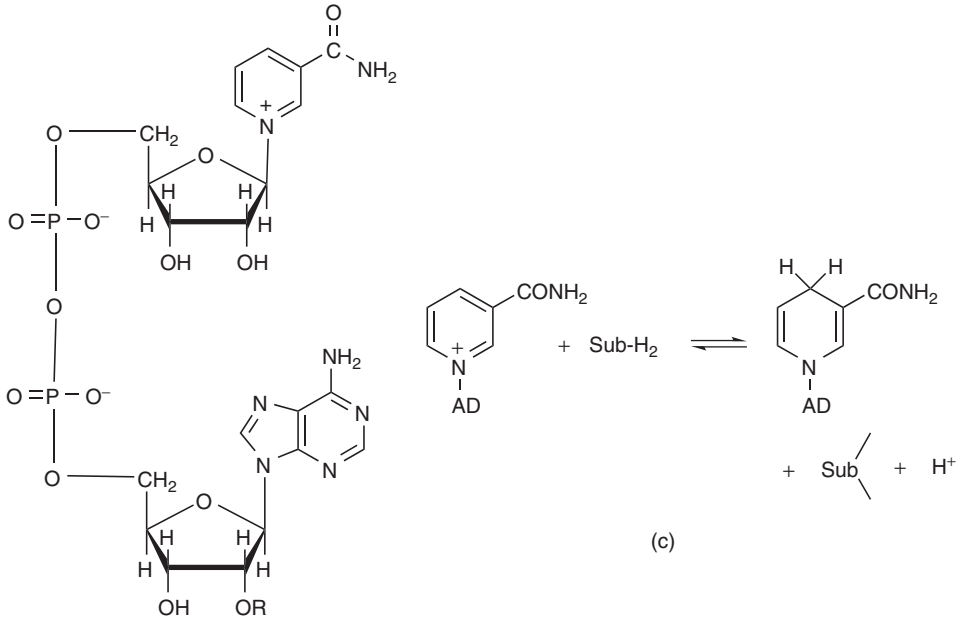


Figure 4.5 *Structure of two dehydrogenating coenzymes: (a) nicotinamide adenine dinucleotide (R = H), (b) nicotinamide adenine dinucleotide phosphate (R = PO_3H_2). (c) An oxidation–reduction process with their participation (Sub-H_2 = substrate).*

Enzymes containing these groups (NAD or NADP) are named dehydrogenases or, more precisely, pyridine-dependent dehydrogenases.[3] More than 150 dehydrogenases are known at present, and no other coenzyme class controls so many different cellular reactions. This variety is readily

[3] Dehydrogenating or hydrogenating coenzyme pairs such as NAD^+ – NAD-H are called oxidoreductases.

explained. We have already mentioned (Section 2.1) that azines, and especially azinium cations, are readily reduced. In the case of the coenzymes NAD^+ and $NADP^+$, the ease of reduction is increased by the electron-accepting carbamoyl substituent $CONH_2$. In contrast, the aromaticity of the pyridine ring is distorted in the partially reduced NAD-H. Therefore, in case of need, reverse donation by NAD-H of a hydride ion can occur readily. These diametrically opposed tendencies allow the formation of optimal redox balances whose flexibility is utilized by organisms. The activity of the pyridine-dependent dehydrogenases is significantly enhanced by several divalent metal ions (Mg^{2+}, Zn^{2+} and others). These ions complex with the coenzyme which thus becomes connected to the apoenzyme. Additional information about the NAD-H coenzyme is presented in Chapters 5, 6 and 11.

Flavin-dependent dehydrogenases make up another important class of dehydrogenases. The biological action of flavin coenzymes results from a vital heterocyclic constituent called alloxazine. Alloxazine is a tricyclic heterocycle and can be considered as being composed of two condensed fragments, pteridine and benzene rings (Figure 4.6; see also Figure 1.6). Several tautomeric forms are possible, and the two most important are called alloxazine and isoalloxazine, differing from one another by the position of a proton at N-1 or N-10.

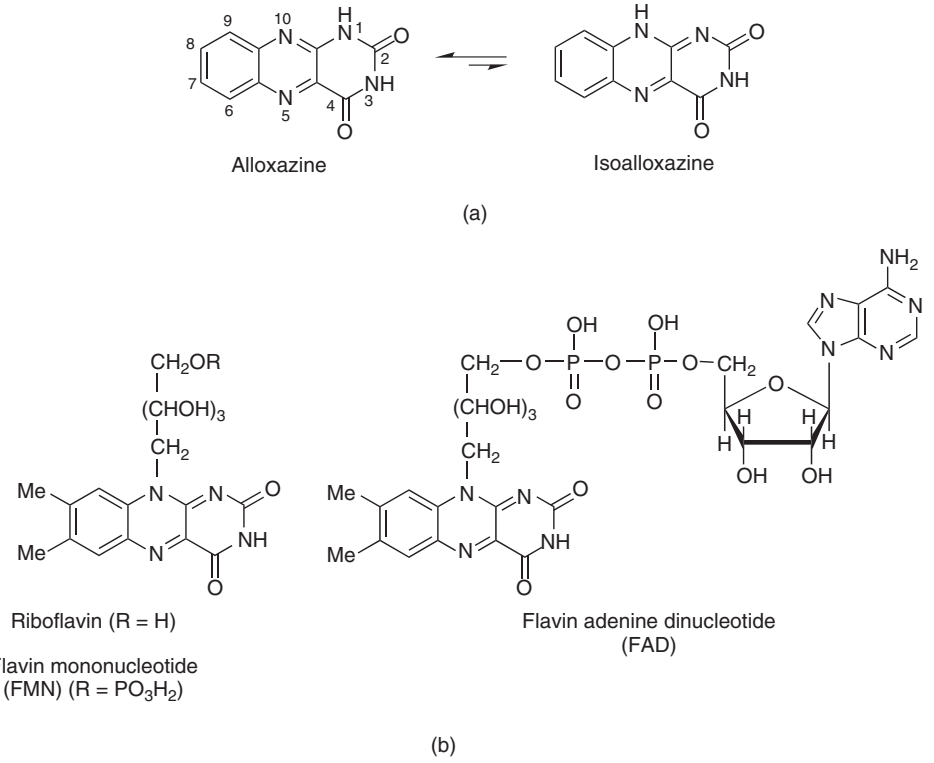


Figure 4.6 *(a) Two important alloxazine tautomers. (b) The flavin coenzymes flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD).*

Under ordinary conditions the alloxazine $\rightleftarrows$ isoalloxazine equilibrium of Figure 4.6a is almost completely shifted to the left. However, Nature has chosen to use the isoalloxazine structure

in flavin-dependent dehydrogenases. We have already described (see Section 2.5) the possibility of fixing unstable tautomers by substitution at the position to which the labile hydrogen atom is attached. Flavins are examples of such fixed tautomers. Thus, riboflavin is a 10-substituted isoalloxazine which has two methyl groups at the C-7 and C-8 positions and one ribityl residue (partly reduced acyclic ribose) at N-10. The monophosphate derivative, flavin mononucleotide (FMN), is one of the two biologically important flavin coenzymes. The second, flavin adenine dinucleotide (FAD), has a more complex structure (Figure 4.6b).

What advantages does the isoalloxazine form possess over the alloxazine tautomer from a biological viewpoint? It is well known that quinones, such as *p*-benzoquinone, are readily reduced by various reagents to hydroquinones (Figure 4.7a). The driving force for this reaction is a combination of the presence of the electron-accepting carbonyl groups and the increased aromaticity of the reduced product.

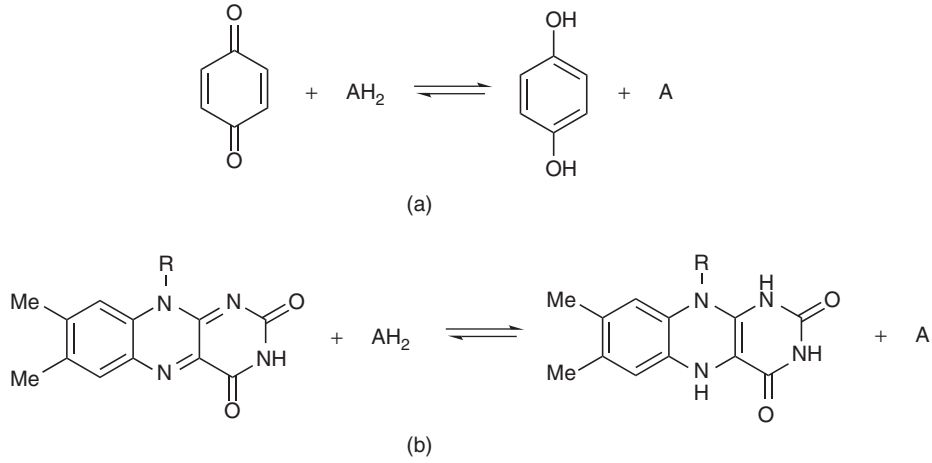


Figure 4.7 *Reduction (a) of p-benzoquinone to hydroquinone and (b) of the flavin coenzyme FMN or FAD to its H-form.*

The same forces are at play in the reduction of the FMN and FAD coenzymes whose structures incorporate the quinonoid bond systems (Figure 4.7b). Hydrogen uptake occurs at the N-1 and N-5 positions, that is, at the termini of the quinonoid chain. As a result, the uracil ring acquires a more stable π-electron structure. The reduced flavin coenzymes are designated as FMN-H and FAD-H. It is interesting to compare the chemical actions of pyridine-dependent and flavin-dependent dehydrogenases. While both transfer two electrons, they differ in the number of hydrogen atoms being transported. Coenzymes with the nicotinamide moiety accept or donate one hydrogen atom, whereas with flavins two hydrogen atoms are involved.

Recently, new coenzymes have been found which participate in hydrogenation–dehydrogenation processes. These new compounds were also found to be nitrogen heterocycles. One enzyme having the 5-deazaflavin structure was isolated from anaerobic methane-discharging bacteria (Figure 4.8a). As in the case of flavins, hydrogenation occurs at positions 1 and 5.

A further coenzyme (PQQ), a derivative of a pyrrolo[2,3-*f*]quinoline quinone (Figure 4.8b), is encountered in many important oxidoreductases such as alcohol dehydrogenase, aldehyde dehydrogenase, D-glucose dehydrogenase and others. The central *o*-quinone ring is reduced to the corresponding catechol derivative. The pyrrole and pyridine rings do not directly take part in this

reaction. However, the first ring is a π-donor and the second a π-acceptor. Their combined effect probably creates the required optimal oxidative–reductive potential.

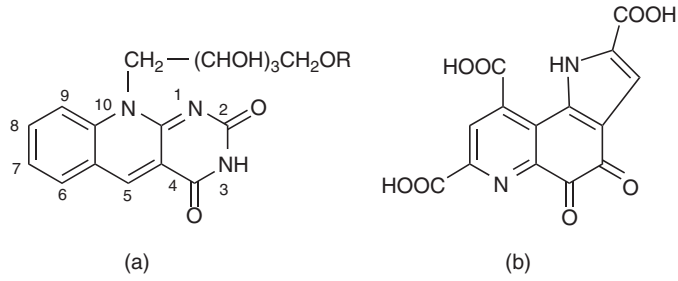


Figure 4.8 *Coenzymes transporting hydrogen with (a) 5-deazaflavin and (b) pyrrolo[2,3-f]quinoline moieties.*

Cytochromes as Electronic 'Postmen'

Oxidative–reductive enzymes whose only function is to transfer electrons from one molecule to another occur in all living cells. These electronic 'postmen' are named cytochromes.

Hemes, complexes of porphyrin with iron(II), serve as the coenzymes for numerous cytochromes. Hemes are deeply colored (usually red, brown or orange) and impart a corresponding color to the cellular structures containing them. The name cytochrome is a derivative of two Greek words (*kytos*, cell; *chroma*, color). The classification of cytochromes, designated by letters *a, b, c* and so on, is based on their color, or the position of their long wavelength absorption band. For example, cytochromes *a, b* and *c* have absorption maxima at 600, 563 and 550 nm, respectively.

We now examine in detail the porphyrins, a compound class which we will encounter many times. Porphyrins are substituted derivatives of porphin, a tetrapyrrolic macrocycle (see Figure 2.5). Depending on the type and the position of their substituents, porphyrins are classified as etioporphyrins, mesoporphyrins, coproporphyrins or protoporphyrins. The last are the most frequently encountered and contain four methyl groups, two vinyl groups and two propionic acid residues as substituents. These substituents can be attached to the four pyrrole rings in 15 different orientations, giving rise to 15 isomeric protoporphyrins. Protoporphyrin IX (Figure 4.9a) is a constituent of the cytochromes, in particular the widely studied cytochrome *c*. The coenzyme in cytochrome *c* is bound to the protein of the enzyme by two sulfide bridges which are formed by addition of the SH groups of cysteine to the vinyl groups of protoporphyrin IX (Markovnikov addition; Figure 4.9b).

It is interesting that in other cytochromes there are no covalent bonds linking the coenzyme and apoenzyme, the stability of the complex being purely the result of noncovalent interactions. Additional noncovalent interactions evidently also exist in cytochrome *c*. For example, there are two coordinate bonds between the Fe^{2+} ion and two amino acid ligands situated on either side of the plane of the porphyrin ring system and almost perpendicular to this plane. One such ligand is almost always the imidazole ring of a histidine residue. The other may also be a histidine residue or another amino acid residue such as methionine.

It is thus clear that the coordination number of iron(II) in the cytochromes is six, and therefore the electron shell is that of the nearest inert gas, krypton. Owing to complete saturation with ligands, the Fe^{2+} ion is incapable of binding with other molecules. However, the metal ion is able to lose one electron thereby becoming Fe^{3+}. It is this reversible $Fe^{2+} \rightleftarrows Fe^{3+}$ exchange which

is the basis of the activity of the cytochromes (Figure 4.9c). In the next chapter we describe the interactions of cytochromes with electron donors and acceptors in a cell.

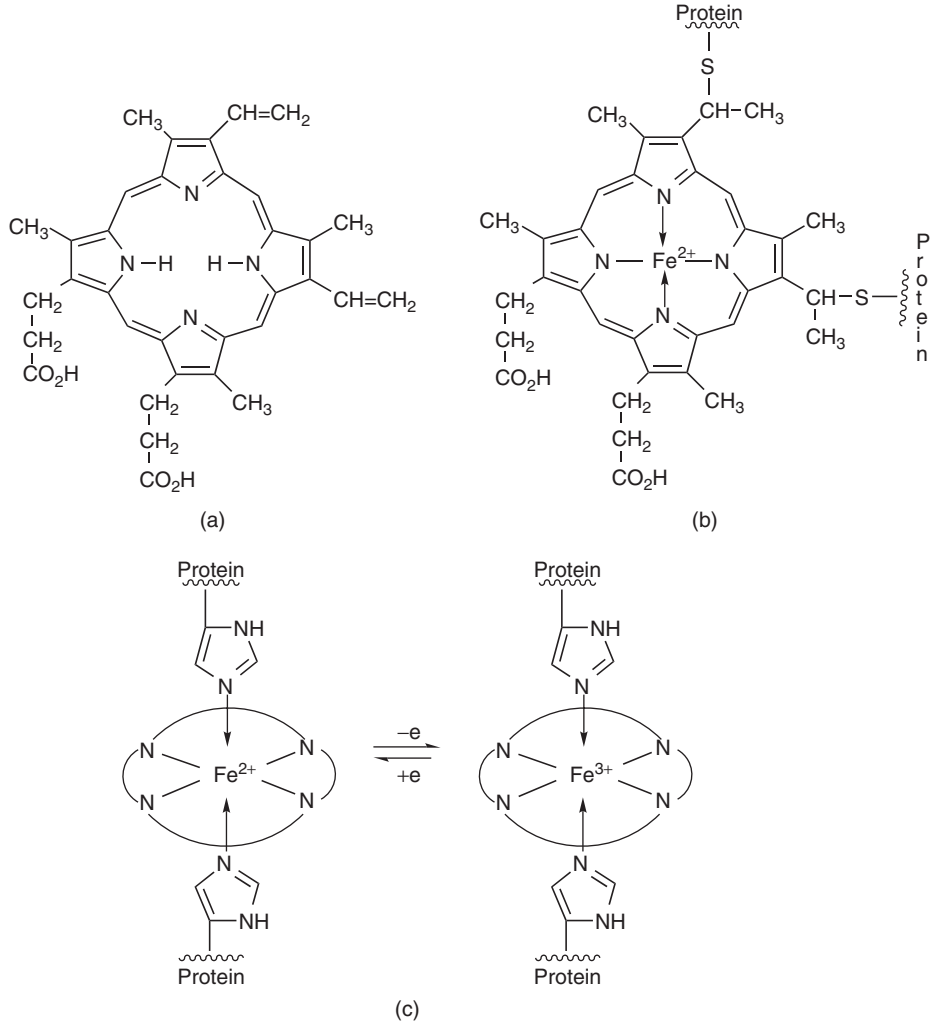
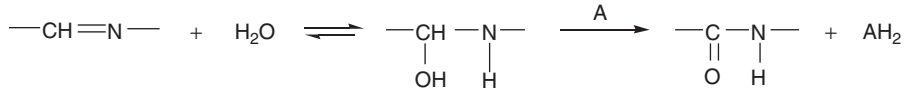


Figure 4.9 *(a) Protoporphyrin IX. (b) The coenzyme in cytochrome c. (c) Oxidation–reduction equilibrium in cytochromes.*

Oxygenases

With the exception of anaerobic bacteria, all living organisms require molecular oxygen. Oxygen is utilized both as an electron acceptor in the respiratory chain (Section 5.2.3) and for the biosynthesis of various substances, such as catecholamines, porphyrins and so on. Oxygen can also be supplied from water as occurs, for example, in the oxidative deamination of amino acids (Figure 4.31) or during the formation of oxo derivatives of purine. In the latter case, the water covalently adds

to the electron-deficient CH═N bond with subsequent dehydrogenation of the adduct formed catalyzed by a dehydrogenase:



Enzymes which promote the incorporation of molecular oxygen into organic compounds are called oxygenases. There are two types: monooxygenases, which control the introduction of one oxygen atom, and dioxygenases, which monitor the incorporation of two oxygen atoms into a substrate:

$$\text{substrate} + O_2 \longrightarrow \text{substrate-O} + (O) \quad \text{(monooxygenase)}$$

$$\text{substrate} + O_2 \longrightarrow \text{substrate-}O_2 \quad \text{(dioxygenase)}$$

In the case of monooxygenases the second atom of oxygen clearly has to be converted into a water molecule. Therefore, monooxygenase activity is accompanied by the action of a further hydrogenase enzyme which donates two hydrogen atoms to bind with this oxygen atom.

An important class of reaction catalyzed by oxygenases is the cleavage of aromatic C═C bonds near a hydroxy group, between two hydroxy groups or in an indole system. Monooxygenases which hydroxylate aromatic and aliphatic compounds without C—C bond cleavage are called hydroxylases. They can also cause *N*-dealkylation and the formation of epoxides, *N*-oxides and *S*-oxides. Alternative oxidations of tryptophan (Figure 4.10) provide typical examples of the action of mono- and dioxygenases. A monooxygenase converts tryptophan into 5-hydroxytryptophan which is then decarboxylated to give 5-hydroxytryptamine (serotonin), a central nervous system transmitter (see Section 7.4.1). In contrast, tryptophan-2,3-dioxygenase causes cleavage of the double bond between C-2 and C-3 in the pyrrole ring and thus effects formation of *N*-formylkynurenine.

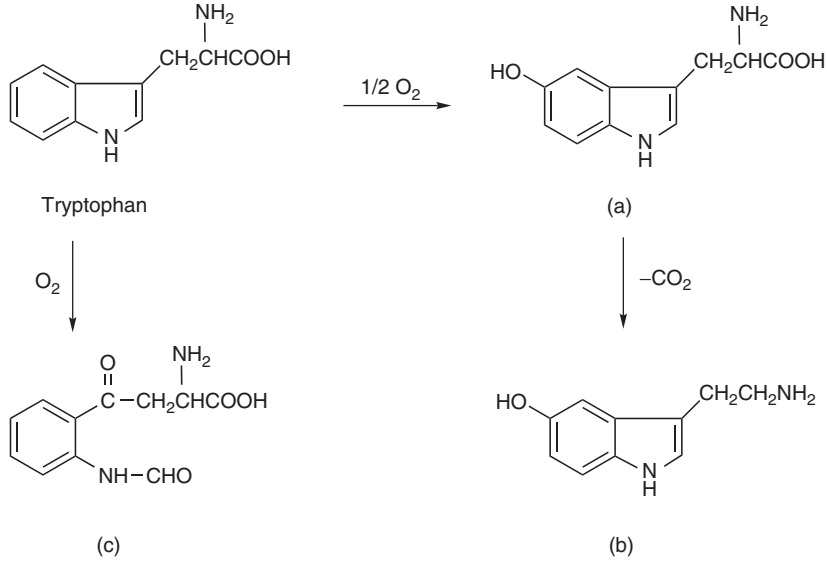


Figure 4.10 *Tryptophan conversions under the action of mono- and dioxygenases into: (a) 5-hydroxytryptophan, (b) serotonin, (c) N-formylkynurenine.*

We now consider the mechanism of action of oxygenases. The main function is the activation of an oxygen molecule, which starts from the attachment of oxygen to the coenzyme in a defined manner. Further events are determined by the type of oxidation reaction and coenzyme nature. Many oxygenases have transition metal ions as cofactors, commonly Fe^{2+}/Fe^{3+} or more rarely Cu^{+}/Cu^{2+}. They act as electron shuttles on activation of oxygen. A key problem is that ordinary (atmospheric) oxygen is chemically rather unreactive since its ground state ($^3O_2$) has two unpaired electrons in the highest occupied π* orbitals, and is spin-forbidden to react with spin-paired singlet species (Figure 4.11a). However, the triplet oxygen can react with substrate radicals. This type of reactivity typically occurs in case of dioxygenases. A characteristic example is oxidative degradation of flavonoids – heterocyclic compounds, which are derivatives of 2-phenyl-1,4-benzopyrone (Figure 4.12a; see also Figure 9.1). They are wide-spread as plant secondary metabolites. One of the most known among them is quercetin (Figure 4.12b) which is degraded under the action of the copper-dependent enzyme quercetin 2,3-dioxygenase. Its copper cation is coordinated via an oxygen atom with a glutamine residue and via nitrogen atoms with three hystidine fragments.

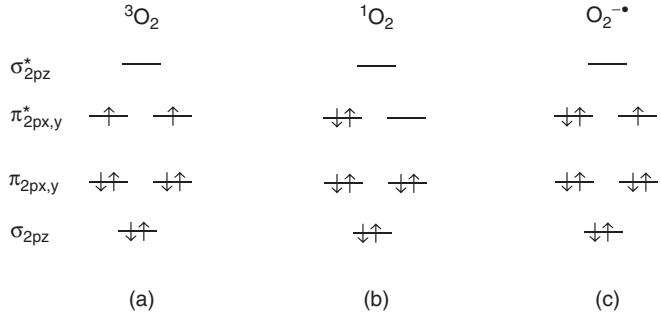


Figure 4.11 *Molecular orbitals diagrams for: (a) triplet oxygen, (b) singlet oxygen, (c) superoxide anion.*

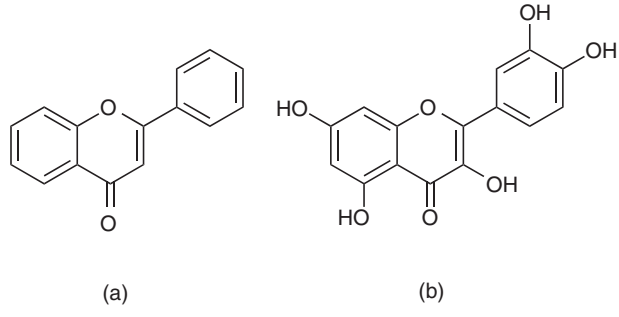


Figure 4.12 *Examples of flavonoids: (a) 2-phenyl-1,4-benzopyrone, (b) 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-1,4-benzopyrone (quercetin).*

Due to the acidic nature of the 3-OH group in quercetin on formation of an enzyme-substrate complex (Figure 4.13a) the Cu^{2+} ion is attached to this phenolate oxygen. The subsequent one electron transfer converts a Cu^{2+} ion into Cu^{+} and produces a phenoxy radical (Figure 4.13b). The latter is a resonance-stabilized specie in which spin density is distributed between the oxygen

atom at C-3 and the atom C-2. This allows triplet oxygen to attack the ring carbon atom with the formation of a peroxy radical (Figure 4.13c) and after a reversed one electron transfer – peroxide anion (Figure 4.13d). Two final stages are intramolecular nucleophlic attack on a carbonyl group followed by ring-opening and extrusion of a CO molecule (Figure 4.13e, f). Both oxygen atoms of the O_2 molecule are incorporated into the finally formed O-benzoylsalicylate.

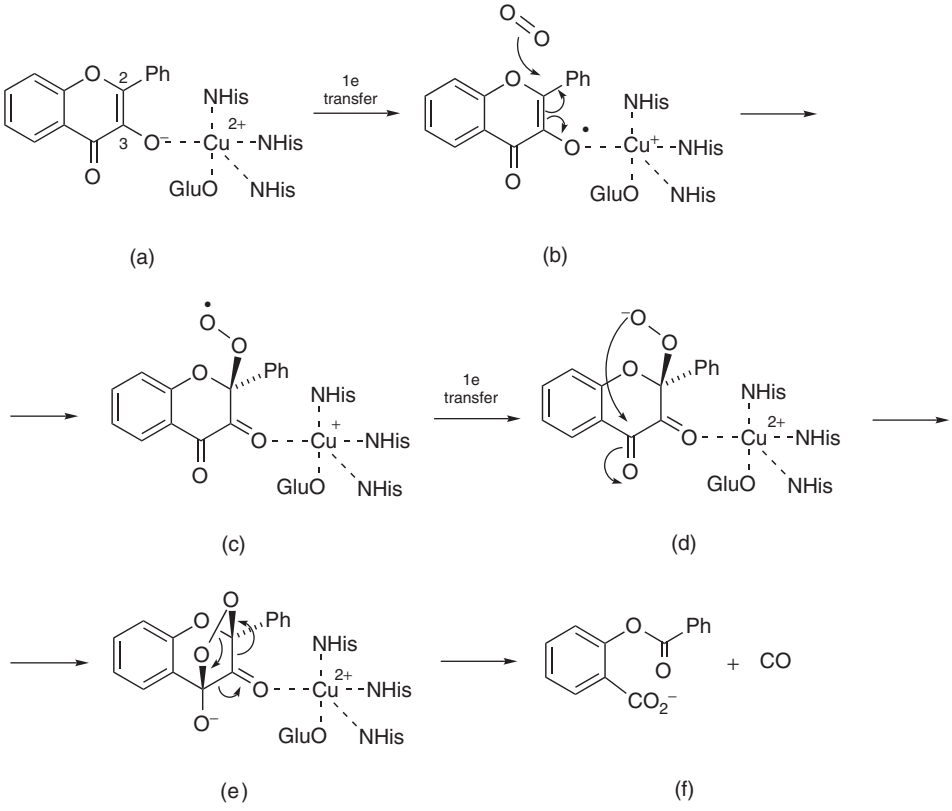


Figure 4.13 *Enzymatic destruction of quercetin (OH groups in position 5-, 7- and in 2-phenyl ring are omitted for clarity).*

In those cases when a substrate enters an oxidation reaction a nonradical (spin-paired) form of special activation of an oxygen molecule is needed. In principle, the $^3O_2$ can be activated either by its transition into a singlet excited state ($^1O_2$; Figure 4.11b), or via a one electron reduction to produce the superoxide radical-anion $O_2^{-\bullet}$ (Figure 4.11c). The superoxide specie can react easily both with electrophiles and radical species. The first type of activation can not be achieved in a living cell since $^1O_2$ is 22 kcal mol^{-1} higher in energy than $^3O_2$. Thus, the second mode is realized, mostly in monooxygenase.

Monooxygenases are of great interest as many contain heterocyclic coenzymatic fragments. Two types of heterocycles are found in monooxidases: dihydroflavins (Figure 4.7) and pterines. The latter (e.g., tetrahydrobiopterine; Figure 4.14), are derivatives of 2-aminopteridin-4-one which are widespread in naturally occurring systems (see also Figures 4.34 and 9.2). Both these coenzymes

are capable of adding an O_2 molecule, in each case at the bridgehead 4a position. Thus, adducts **1** and **2** are produced in which the oxygen is in the form of an active peroxide. The formation of a peroxide is belived to start with one electron oxidation of a coenzyme that gives superoxide anion and radical-cation **3** with a set of canonical structures differing in spin and positive charge distribution, such as **3a** and **3b**. A subsequent recombination of **3b** and $O_2^{-\bullet}$ accompaning by proton transfer produces peroxide **2**.

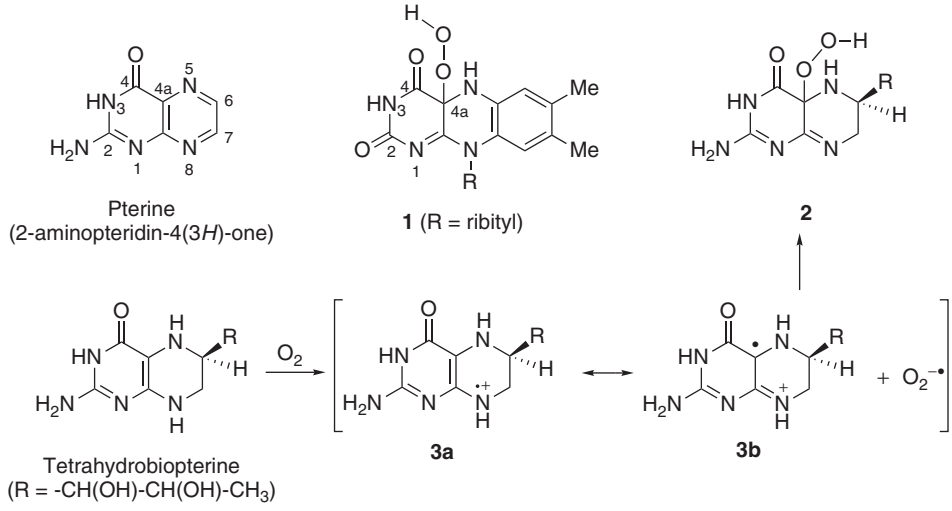


Figure 4.14 *Reduced flavins and pterines as monooxidase coenzymes and the formation of an active peroxide.*

The conversion of phenylalanine into tyrosine is a typical example of the action of such a monooxygenase (Figure 4.15). Phenylalanine oxidase, which has a tetrahydrobiopterine fragment, takes part in this reaction. Adduct **2** with oxygen oxidizes phenylalanine to an epoxide, which is then aromatized by migration of hydrogen atom H_A from the *para* position of phenylalanine to the *meta* position of tyrosine. Tetrahydrobiopterine is thus transformed into the oxidized form through a recyclization pathway via a putative acyclic intermediate. In the final step the oxidized form is reduced by an NAD-H-dependent hydrogenase, thus regenerating the original enzyme.

In some cases oxygen activation can be provided by one electron transfer from the Fe^{2+} cation to a molecule of oxygen: $Fe^{2+} \cdot O_2 \rightleftarrows Fe^{3+} \cdot O_2^{-\bullet}$. In reality, however, the substrate also participates in the activation. The initial step is the formation of an enzyme–substrate complex in which the enzyme protein is reconstructed in such a manner as to make the Fe^{2+} ion more available to the oxygen molecule. The rearrangement thus facilitates the subsequent formation of a triple complex (enzyme–substrate–O_2), inside which the oxidation takes place (Figure 4.16). Obviously, that situation in this case looks like an intermediate between mechanisms represented in Schemes 4.13 and 4.15.

This can be best exemplified by physiologically important conversion of heme into biliverdin (Figure 4.17). The highly regiospecific process is catalyzed by an enzyme called heme oxygenase. The reaction is unique because all its three steps are accompanied by O_2 activations in which the substrate itself participates. At the first step self-hydroxylation of the porphyrin meso-carbon atom occurs. The resulting hydroxygeme reacts in the second step with another O_2 to yield verdoheme

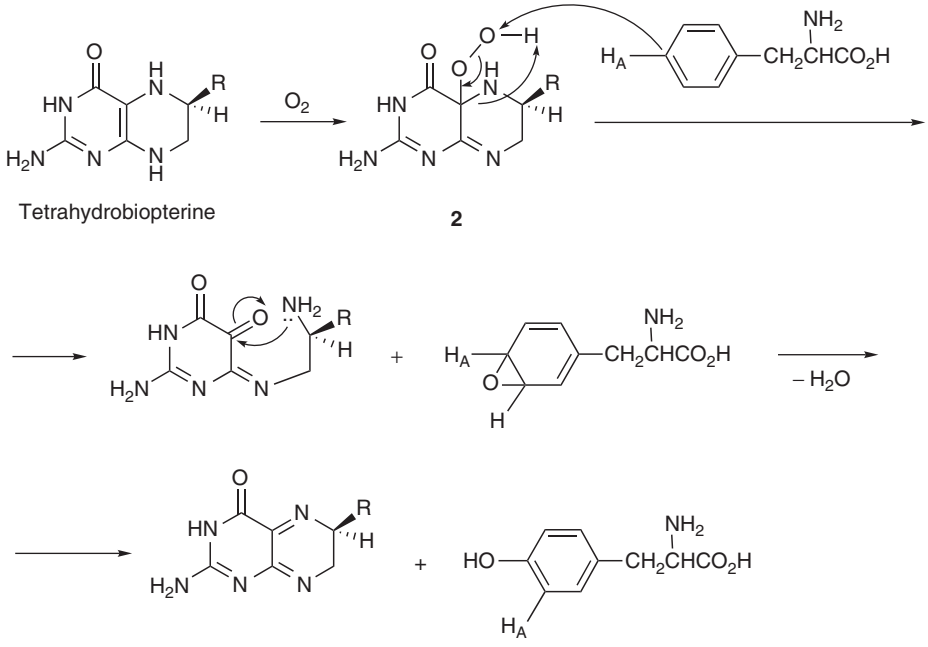


Figure 4.15 *Mechanism of the oxidation of phenylalanine to tyrosine with the coenzyme tetrahydrobiopterine.*

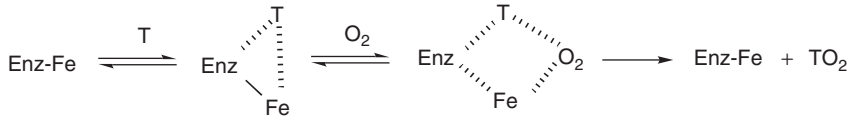


Figure 4.16 *Oxidation of a substrate (T) by the action of a dioxygenase.*

(a rare natural macrocycle with an oxonia heteroatom) and CO molecule. The third O_2 activation, by verdoheme, cleaves the porphyrin ring to form biliverdin and free Fe^{2+} ion. Conclusive evidence has been recently obtained that the Fe-OOH intermediate in its hydrated form (shown at the bottom of Figure 4.17) is a key specie of the self-hydroxylation as well as the ring-opening reaction.

An activity of cytochromes P450, that constitute a special group of oxydoreductases, somewhat recalls that of heme oxygenase. Cytochromes P450 participate in the biosynthesis of important metabolites and play a role (not yet completely investigated) in the oxidative neutralization of toxic substances received by human organisms with food, water or air. The structure of cytochromes P450 is composed of hemin and flavin cofactors; they thus belong to the class of hemoflavoproteides that differentiates them from the other cytochromes. Their wonderful ability to recognize the multitude of structurally very diverse toxicants can be explained by the remarkable conformational flexibility of their single chain protein entity (apoenzyme fragment) which is present in the cytochrome complex. The apoenzyme interacts with any chemical substrate (R—H) in such a way that the most sensitive part of the latter towards oxidation (mainly carbon atoms having the sp^3 configuration) appears in the vicinity of the active oxygenase site of the cytochrome. As a result, the aliphatic hydrophobic groups of the toxic molecule are oxidized into hydrophilic hydroxyalkyl

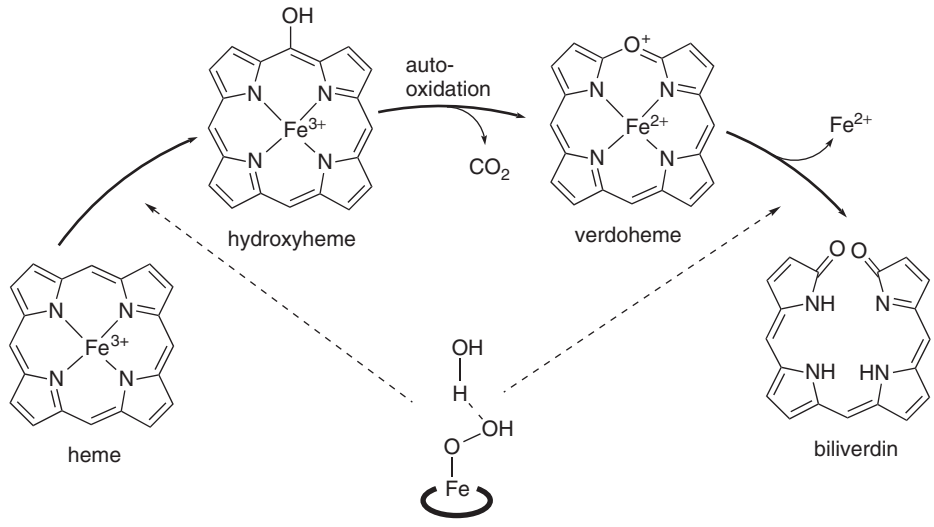


Figure 4.17 *Three-step conversion of heme into biliverdin catalyzed by heme oxygenase (substituents in porphyrin system are all omitted for clarity, see Figure 9.3 for details; Matsui, T., Unno, M. and Ikeda-Saito, M.,* Acc. Chem. Res.*, 2010,* ***43****, 240).*

groups. This process diminishes the toxicity of the metabolite (R—OH), strongly enhances its solubility in the aqueous phase which facilitates its elimination from the organism.

As shown in Figure 4.18, the first step in this sequence of biochemical neutralization of a toxic molecule in the liver comprises its bonding into an aquatic complex with cytochromes P450. Coenzyme NADP-H then helps to reduce Fe^{3+} into Fe^{2+} in the hemin part of this complex and increases its affinity towards oxygen and thus facilitates the formation of a peroxocomplex in which a molecule of oxygen is activated. Afterwards the coordination compound is reduced by one additional electron to produce a peroxocomplex with the Fe^{1+} cation. Under the action of two protons this reduced peroxocomplex is further oxygenated into another complex ($Fe^{1+} \rightarrow Fe^{3+}$) in which the oxygen hydroxylates the organic ligand (R—H) converting it into alcohol R—OH. The alcohol is then rapidly metabolized into nontoxic substances that are excreted from the organism via its kidneys.

Moreover, cytochromes P450 are also found in the membranes of skin, brain, kidneys and other organs and can catalyze dealkylation of N-, O- and S-alkyl groups, reduce azo and nitro groups into amino groups and various other reactions which neutralize toxic chemical xenobiotics.

Thus, one can conclude that heterocycles in oxygenases act as: (i) substrates for oxidation, (ii) coenzymes for oxygen activation and transfer, (iii) protein constituents, such as hystidine residues, coordinating metal ions in coenzymes, (iv) intermediates, for example, dioxetanes, in oxidation processes.

4.2.2 Coenzymes as Carriers of Molecular Species

Numerous enzymes are engaged in the transfer of various functional groups, molecules and ions. In some cases, the enzyme acts to transport a molecule; in others, enzymes deliver a group into a receiving molecule during the biosynthesis of a complex natural product. Many transferring enzymes include in their structures a heterocyclic coenzyme which is directly involved in the transfer process.

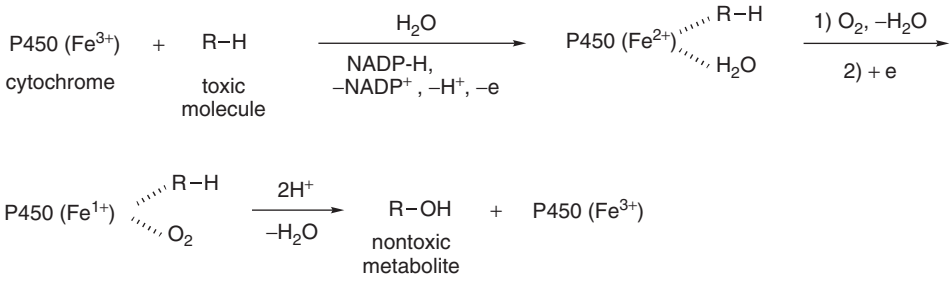


Figure 4.18 *The sequence of biochemical neutralization of a toxic molecule by cytochrome P450.*

Hemoglobin and Myoglobin as Oxygen Transporters

Hemoglobin and myoglobin are the most widespread and best known molecular transportation vehicles, and their function is to supply tissues with oxygen. Strictly speaking, they are not classified as enzymes, but they have much in common with enzymes as far as their structures and mechanism of action are concerned. Hemoglobin and myoglobin contain heme as the coenzyme. The heme is held within one of the clefts of the polypeptide globule by hydrophobic interactions and by coordination linkages between the Fe^{2+} ion and a histidine residue. Myoglobin molecules contain a polypeptide chain of 153 amino acid residues. Myoglobin is the oxygen-binding protein responsible for the transport of oxygen and also for its storage in muscle tissues. Hemoglobin is more complex than myoglobin. It is a molecular aggregate composed of four intertwined protein chains, of which two are identical α-type protein chains, and two are identical β-type chains. The α-type chains contain 140 residues each and the β-type chains contain 146 amino acid residues each. Each chain binds one heme coenzyme. Consequently, hemoglobin as a whole contains four heme residues.

The natural occurrence of hemoglobin is more extensive than that of myoglobin, the latter being encountered only in the skeletal muscles of whales, sharks and other sea creatures. Hemoglobin is believed to have been the more reliable oxygen-bearing system during biological evolution: the presence of several polypeptide chains and several hemes allows the destruction of any one unit to be compensated by the normal functioning of the others. Genetic disorders and blood diseases induced by such degradation are not uncommon in man and animals, but thanks to natural selection they do not always prove to be fatal.

A crucial difference between the heme of cytochromes and the heme of globins lies in the fact that in the latter the Fe^{2+} ion is linked with only five ligands: four nitrogens of the porphyrin system and one imidazole nitrogen of a histidine residue in the protein chain. As a result, the Fe^{2+} ion in globins can utilize its vacant sixth coordination site to bind an oxygen molecule (Figure 4.19).[4] This binding must, of course, be reversible and therefore not too strong. Oxygen becomes linked to heme under the normal partial pressure inherent in lung alveoli and is then released in tissues with an oxygen shortage; that is, where the partial pressure of oxygen is low.

It is instructive to consider what prevents the globin iron from spontaneous oxidation to iron(III), since, for example, the rapid oxidation of a cut apple exposed to air is well known. X-Ray crystallography reveals that hemoglobin contains a further histidine residue (in addition to that coordinated to iron) in close proximity to the heme and that the imidazole ring of this second

[4] The sixth valence of iron(II) in globin hemes is considered by some authors to be occupied by a water molecule. Under appropriate conditions this molecule of water is readily displaced by oxygen.

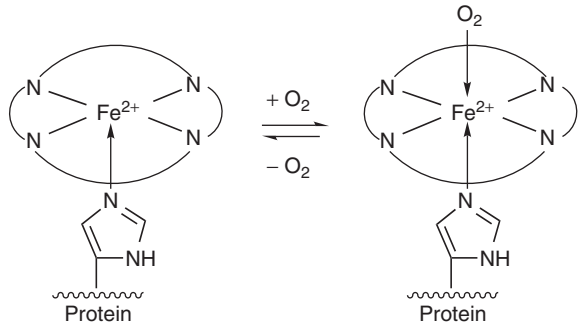


Figure 4.19 *The reversible binding of oxygen to the globin heme.*

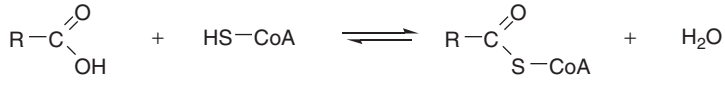
histidine is protonated. The presence of the positive charge near the Fe^{2+} cation strongly disfavors the loss of an electron, as the electrostatic repulsion between the resulting Fe^{3+} and the imidazolium cation would be great. The blood disease called ferrihemoglobinemia is an example of what can occur when this system does not function correctly. This disorder is characterized by substitution of the second histidine residue (the one not coordinated with iron) by another amino acid. This replacement in the protein chain leads to the facile oxidation of two iron atoms in such a heme to Fe^{3+}. In this state, the iron cations cannot coordinate with oxygen. As a result, the oxygen absorption capacity of the blood in such patients is diminished by a factor of two.

The chemical reactions of hemes involve not only the iron cation, but also they require participation of imidazole rings of the polypeptide histidine units and of the porphyrin system as a whole. The role of the highly conjugated porphyrin system is to delocalize and therefore stabilize charge. Such stabilization is especially important for the hemes, as the lifetime of the intermediate oxidized and reduced states must be long enough for it to exercise the corresponding biological effects.

Coenzyme A as an Acyl Group Carrier

Coenzyme A is an extended chain molecule containing the following sequence of units: adenine, ribose, pyrophosphate, pantothenic acid and 2-aminoethanethiol (Figure 4.20). Although the heterocyclic moieties of this coenzyme do not participate directly in the enzymatic reactions, consideration of its specific biochemical properties helps in a better understanding of the functions of many other coenzymes, including those based on heterocycles.

Coenzyme A (designated also as CoA or CoA-SH) transports acyl groups (RC═O) and also activates fatty acids in various bioreactions. The critical functionality in CoA is the terminal thiol which converts a carboxyl function into the corresponding thiolester:



Thiolesters of carboxylic acids have two characteristics important for the properties of coenzyme A. The first is that the energy-rich C—S bond in a thiolester is labile. Under the influence of various nucleophilic agents (water, alcohols, amines, etc.) this C—S bond is readily broken to liberate CoA with simultaneous acylation of the nucleophilic agent (Figure 4.21a). The lability of the C—S bond is controlled by several factors, the main being the high polarization of the sulfur valence electrons. A second characteristic property of thiolesters important for the biological activity of coenzyme A

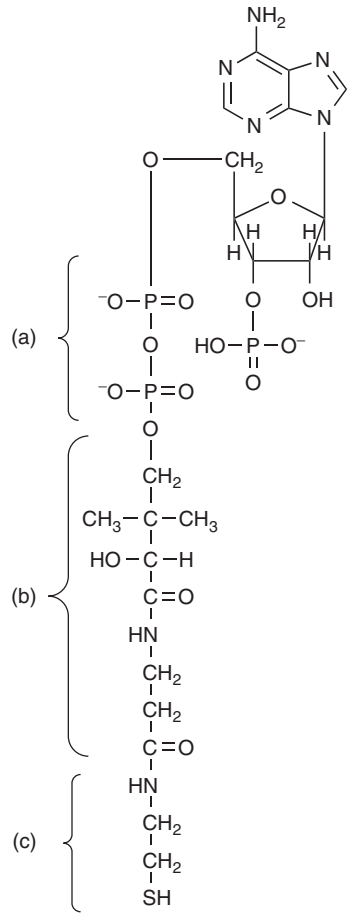


Figure 4.20 *Coenzyme A: (a) the pyrophosphate moiety, (b) the pantothenic acid residue, (c) the 2-aminoethanethiol moiety.*

is that the C—H bonds α to the carbonyl group are strongly activated. Even comparatively weak bases accept a proton from such C—H bonds to form a carbanion, as shown in Figure 4.21b.

In organic chemistry, the facile generation of carbanions by proton loss from the α-position of aldehydes, ketones and other carbonyl-containing compounds is well known. The negative charge is stabilized by conjugation with the C═O group (indicated by arrows in Figure 4.21b). In thiolesters the C—H bond activation is significantly greater than in ordinary esters, owing to the unfilled sulfur 3*d*-orbitals. Their relatively large size allows these 3*d*-orbitals partially to overlap the unshared electron pair of the α-carbon atom (a dashed line in Figure 4.21b), which imparts additional stabilization to the carbanion and facilitates its generation.

A combination of these factors is involved in the oxidative decomposition of fatty acids in organisms (Figure 4.22). This multistep process requires the participation of a large number of enzymes including the flavin- and pyridine-dependent dehydrogenases FAD and NAD^+. In the first stage (Figure 4.22a) coenzyme A is 'loaded' with an acid to form a thiolester in which the activation of the α-C—H bond enables dehydrogenation of the acid by FAD in the second stage

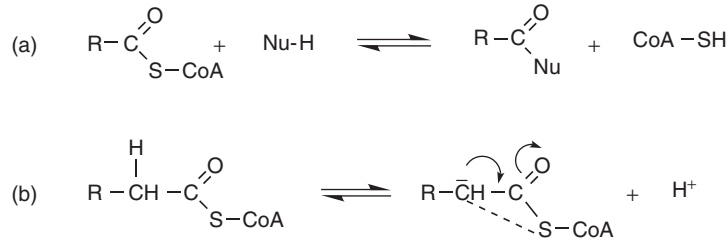


Figure 4.21 *Chemical reactions of acylated coenzyme A: (a) acyl group transfer to a nucleophilic agent. (b) ionization of the C—H bond with the formation of a carbanion.*

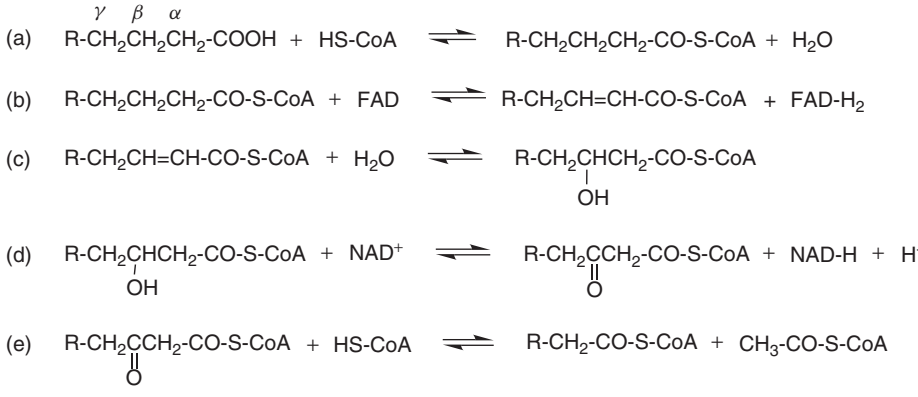
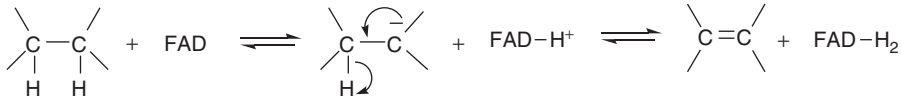
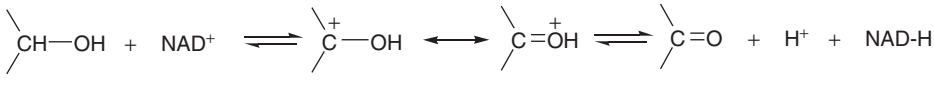


Figure 4.22 *Oxidative breakdown of fatty acids.*

of the process (Figure 4.22b). This conversion probably proceeds stepwise in two fast successive steps entailing the loss of a proton to generate a carbanion which then undergoes hydride ion loss:



The third stage (Figure 4.22c) is the hydration of the C═C bond in the unsaturated thioester thus formed. The orientation of this addition is consistent with electron delocalization theory: a hydroxide anion is added at the β-carbon atom (which has a partial positive charge owing to conjugation with the C═O group) followed by a proton at the α-carbon atom. In the fourth stage (Figure 4.22d) the hydroxy group is oxidized to a carbonyl group by an NAD^+-containing dehydrogenase, thus forming a β-keto acid (in the thiolester form):



The C—C bond situated between the keto group and the central carbon of β-keto esters are susceptible to facile cleavage by various reagents. Such a cleavage occurs during the final stage

(Figure 4.22e) of the overall process when another enzyme containing CoA converts the β-keto ester into an acetyl CoA and an acyl CoA. The original fatty acid carbon chain is thus shortened by two carbon atoms. The same process is repeated with the newly produced acyl CoA until the initial acid is completely converted into two-carbon acetyl fragments.

The acetyl products of degradation are then subjected to a versatile range of subsequent conversions within the organism. It should be noted that the enzymes containing FAD, NAD and coenzyme A participate together in the form of a single cellular ensemble in the degradation of fatty acids.

Thiamine Pyrophosphate as an Acetaldehyde Activator

Many C—H bonds in organic compounds possess low acidity and the substitution of a hydrogen atom with a metal frequently requires a strong base such as *n*-butyllithium. By such treatment, thiazole is converted into 2-thiazolyllithium, while the *n*-butyllithium is transformed into *n*-butane (Figure 4.23a).

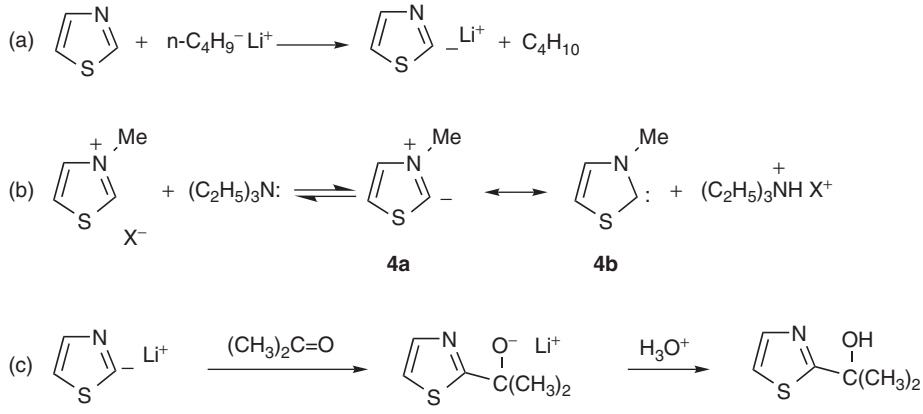


Figure 4.23 *The C-2—H bond ionization in (a) thiazole and (b) the 3-methylthiazolium cation. (c) Reaction of 2-lithiothiazole with acetone.*

Various structural changes can induce a pronounced increase in CH acidity. One, already discussed, is the introduction of a sulfur atom into the α-position relative to the C—H bond, as in the case of coenzyme A. A large increase in acidity can be induced by cation formation in heterocyclic compounds. Both factors are valid in the case of 3-alkylthiazolium salts which are able to ionize the C-2—H bond even under the action of such a comparatively weak base as triethylamine (Figure 4.23b). The resultant specie **4a** possesses carbanionic and cationic centers in close proximity and is called an ylide. Ylides are often represented in an uncharged carbenoid form **4b**.[5] Carbanions are potent nucleophiles widely used in synthesis. For instance, 2-lithiothiazole reacts with acetone to yield the corresponding alcohol (Figure 4.23c).

At the end of the 1950s it was found that similar chemical reactions occur naturally, though in a more specific manner. One of the most important coenzymes, thiamine pyrophosphate, participates in such transformations (Figure 4.24). Not surprisingly, this compound is a derivative of the thiazole cation which thus plays a leading role in biochemical reactions.

[5] Carbenes are compounds with a divalent carbon of the general formula $R^1R^2\ddot{C}$. The practical application of stable heterocyclic carbenes is discussed in Section 11.3.1.

Thiamine pyrophosphate mediates three main types of enzymatic reactions: (i) it decarboxylates α-keto acids; pyruvic acid, a key intermediate in carbohydrate and amino acid metabolism, is thus converted into acetaldehyde; (ii) thiamine pyrophosphate catalyzes the formation of acyloins (α-hydroxyalkyl ketones); (iii) the oxidative decarboxylation of pyruvic acid to form acetic acid also occurs with the help of this coenzyme. Consider the general characteristics of these three reactions.

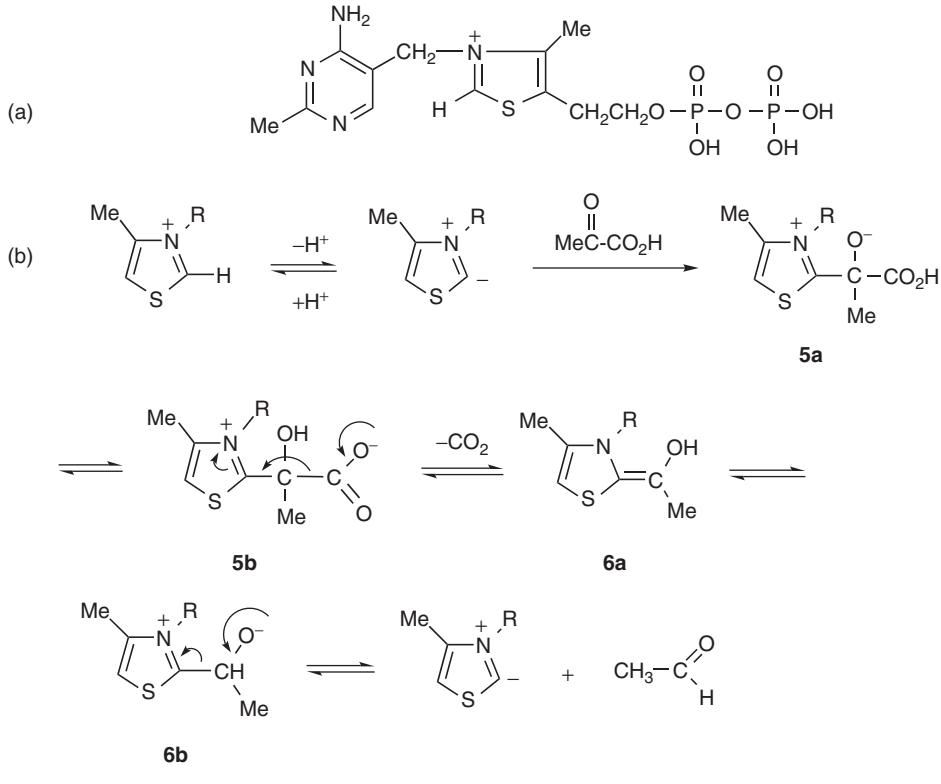


Figure 4.24 *(a) Thiamine pyrophosphate and (b) the assisted decarboxylation of pyruvic acid (only the thiazole portion of the coenzyme is shown).*

The C-2—H bond in the thiamine pyrophosphate thiazole ring is partially dissociated at physiological pH. The keto group of a pyruvic acid molecule adds to the thiamine pyrophosphate ylide (Figure 4.24b). The addition product **5a** undergoes tautomerism into **5b**, and then decarboxylation to give compound **6a** which may be considered as an 'activated acetaldehyde'. The intermediate is tautomeric with **6b**, and can give rise to two alternative further transformations: (i) decomposition into acetaldehyde and the original ylide completes the decarboxylation reaction of pyruvic acid (Figure 4.24b); (ii) the 'activated acetaldehyde' in tautomeric form **6a** can attack a second molecule of acetaldehyde (Figure 4.25). The addition product **7** thus formed is cleaved with the formation of acyloin and the ylide, enabling a second type of enzymatic reaction, which is effectively an acyloin condensation.

The oxidative decarboxylation of pyruvic acid is a more complex process involving other enzymes apart from thiamine pyrophosphate. One of these other enzymes contains as a coenzyme

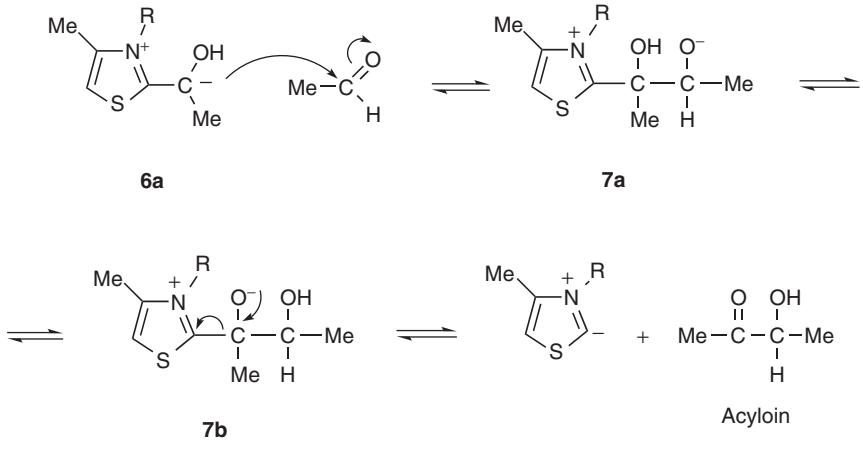


Figure 4.25 *Acyloin condensation catalyzed by thiamine pyrophosphate.*

lipoic acid (Figure 4.26), which can exist in two forms: (i) an oxidized state containing a cyclic disulfide bridge or (ii) a reduced noncyclic form with two thiol groups. The two forms are readily interconvertible by oxidation–reduction processes.

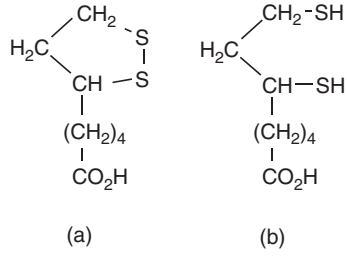


Figure 4.26 *Lipoic acid: (a) oxidized form, (b) reduced form.*

The S—S disulfide bridge is known to cleave even when treated with rather mild reductants. The 'activated acetaldehyde' serves as such a reducing agent and is represented by form **6b** shown in Figure 4.27a. Tautomer **6b** donates a hydride ion to lipoic acid and is oxidized to acetic acid. In the course of the reaction the S—S bond is ruptured, and a hydride ion is added to one sulfur atom while an acetyl group is added to the other.

In metabolic processes, acetyl groups are transported via coenzyme A. In the present case, *S*-acetyllipoic acid reacts with one of the CoA-containing enzymes with subsequent acyl group transfer to the coenzyme (Figure 4.27b). In the process lipoic acid is converted into the dithiol form. The reaction sequence ends with the oxidation of the reduced form of lipoic acid to the cyclic form by means of the FAD coenzyme (Figure 4.27c).

Pyridoxal Phosphate as an Amino Group Carrier

Normal metabolism in man and other animals requires 20 amino acids. Vertebrates are able to synthesize half this number, while the remaining essential α-amino acids must be supplied by

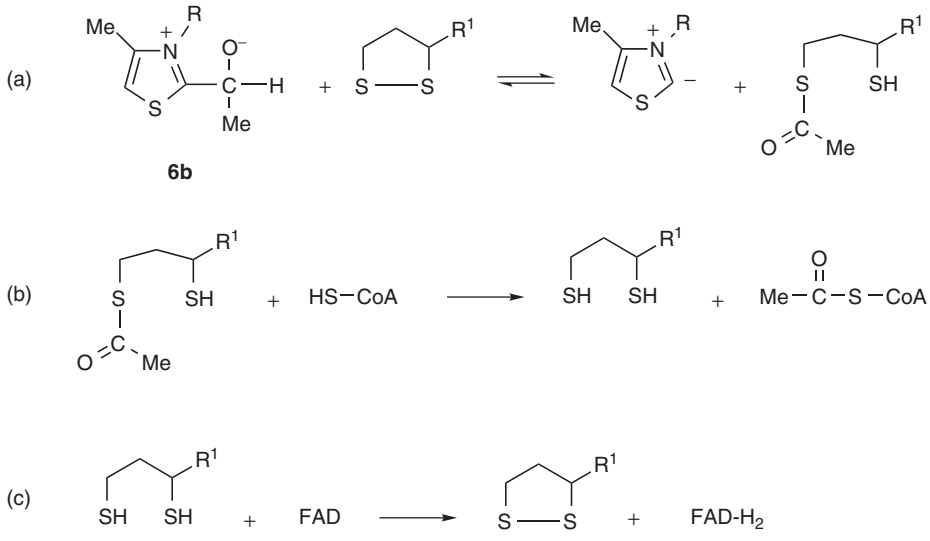
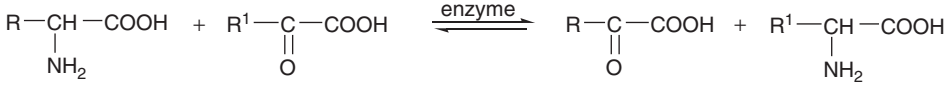


Figure 4.27 *Oxidative decarboxylation.*

their diet. A protein diet is the main source of these essential amino acids. Proteins are hydrolyzed by the digestive system of the organism to amino acids which then undergo complex chemical transformations monitored by dozens of different enzymes. Transaminases are the most important of these enzymes; they function as delivery systems by transferring amino groups and thus interconverting α-amino acids and α-keto acids:



Pyridoxal phosphate, a derivative of pyridine-4-aldehyde, serves as a coenzyme for all of transaminases (Figure 4.28). Enzymatic reactions occur with the direct participation of the aldehyde group which is transformed into an azomethine group (CH═N) by interaction with α-amino acids. The ease of this reaction is due to the interaction between the aldehyde group and the pyridine nitrogen. The electron-withdrawing effect of the nitrogen generates a partial positive charge at the 4-position of the pyridine ring (Figure 2.1), which in turn increases the positive charge on the carbon atom of the aldehyde group. This transfer of electron density substantially facilitates nucleophilic attack by the α-amino acid nitrogen atom at the aldehyde carbon.

A characteristic of the azomethine **8a** formed in the first step is the dramatic increase in C—H bond acidity at the α-position of the amino acid residue over the corresponding free amino acid. The reason for this is that the carbanion produced by C—H bond ionization is greatly stabilized by the delocalization of negative charge over the whole π-system of the molecule.

The carbanionic structure may be represented as a set of resonance structures (Figure 4.29). The pyridine ring, azomethine group and amino acid residue all lie in the same plane to maximize conjugation. This arrangement is achieved by the formation of a six-membered ring stabilized by intramolecular hydrogen bonding between the 3-position hydroxy group and the azomethine nitrogen. Pyridoxal phosphate is able to catalyze transamination only with the participation of this hydroxy group.

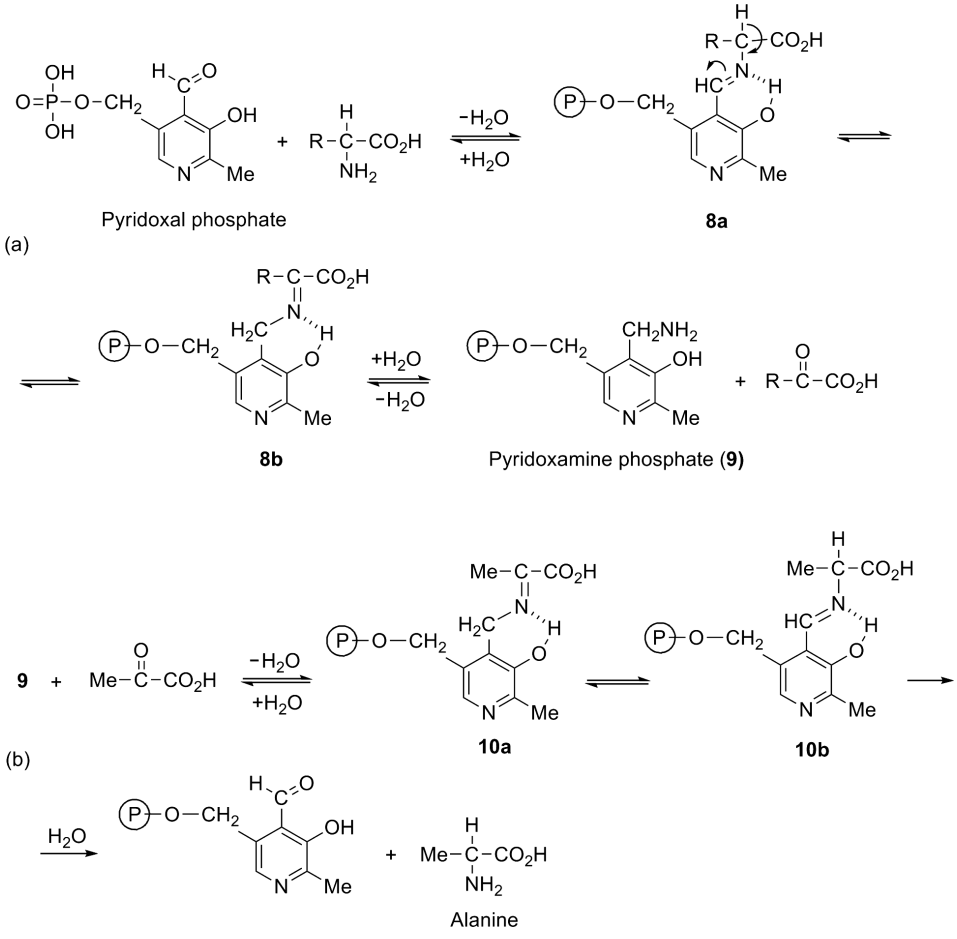


Figure 4.28 *(a) Transamination with pyridoxal phosphate participation (P = H_2PO_3). (b) Amino group transfer from α-amino acid to pyruvic acid with the formation of alanine.*

The higher acidity of the C—H bond in compound **8a** favors proton migration to the azomethine carbon and the consequent formation of tautomer **8b** in which the azomethine double bond occupies a new position between the nitrogen and carbon of the amino acid residue. Subsequent hydrolysis of structure **8b** yields pyridoxamine phosphate **9** and a keto acid, the product of oxidative deamination of the parent α-amino acid (Figure 4.28a). Pyridoxamine phosphate **9** further reacts with pyruvic acid to produce a new azomethine **10a** which tautomerizes to **10b**. Hydrolysis of **10b** leads to the regeneration of pyridoxal phosphate and the formation of alanine (Figure 4.28b).

The amino groups of all the essential α-amino acids are likewise transformed, with simultaneous conversion of pyruvic acid into alanine, each amino acid being controlled by its own transaminase. The alanine itself also undergoes deamination to re-form pyruvic acid; in this case, α-ketoglutaric acid serves as the amino group acceptor (Figure 4.30a) and is converted into glutamic acid. Glutamic acid is then oxidatively deaminated in a process that regenerates α-ketoglutaric acid, this process being mediated by glutamate dehydrogenase and its coenzyme NAD^+ (Figure 4.30b).

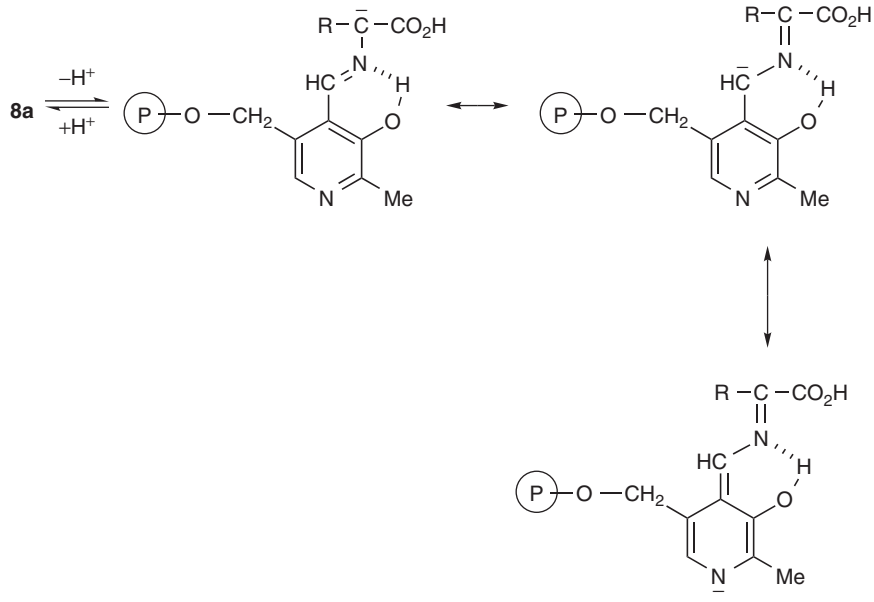


Figure 4.29 *Delocalization of the negative charge in the carbanion produced from compound* ***8a****.*

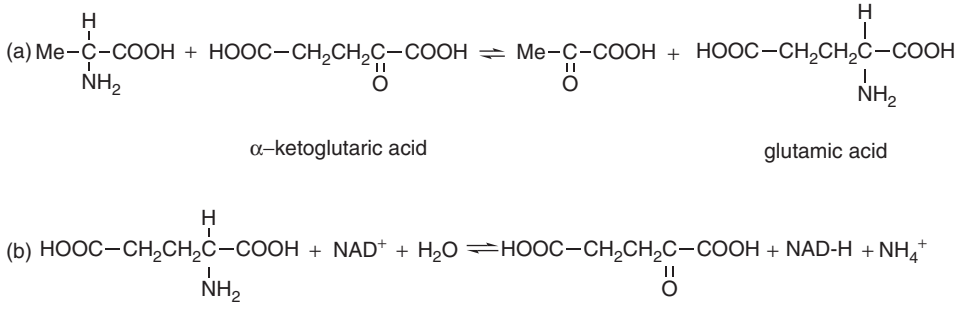


Figure 4.30 *(a) Alanine deamination. (b) Oxidative deamination of glutamic acid.*

NAD^+ abstracts a hydride ion from the α-carbon of glutamic acid, thus converting it into the corresponding immonium salt which, like all immonium salts, is readily hydrolyzed to liberate α-ketoglutaric acid and ammonia (Figure 4.31). Therefore, the metabolism of all amino acids involves the liberation of nitrogen as ammonia.

It follows that the oxidation pathways of all α-amino acids include a 'narrow gate' formed by the alanine–glutamic acid pair. This 'narrow gate' is due to the presence within the organism of an active and highly specific enzyme which targets glutamic acid only.

To complete our discussion of pyridoxal phosphate, we note that pyridoxal phosphate also catalyzes other transformations of α-amino acids, specifically decarboxylations, racemizations and aldol condensations. All of these conversions commence with azomethine formation.

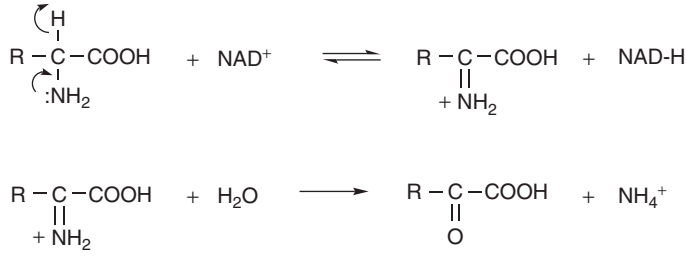


Figure 4.31 *The two-step sequence of α-amino acid oxidative deamination.*

A Coenzyme from Spinach Leaves

Heterocyclization reactions in which heterocycles are formed from acyclic compounds, and ring interconversions in which one heterocyclic ring is formed from another or from a carbocycle, are some of the most interesting and specific reactions of heterocyclic chemistry. So far, we have not mentioned these and have paid attention mainly to reactions which do not involve the heterocyclic skeleton.

There are numerous methods for heterocycle formation from acyclic compounds. The preparations of benzimidazoline (2,3-dihydrobenzimidazole) and benzimidazole from *o*-phenylenediamine on treatment with formaldehyde and formic acid, respectively, are simple examples of heterocyclization (Figure 4.32). Analogous processes are usually accompanied by the formation of intermediate compounds which are often not isolable (generally depicted in square brackets in reaction schemes). Thus, in the benzimidazoline reaction the imine base **11** is an intermediate, and in the benzimidazole reaction *N*-formyl-*o*-phenylenediamine **12** is an intermediate.

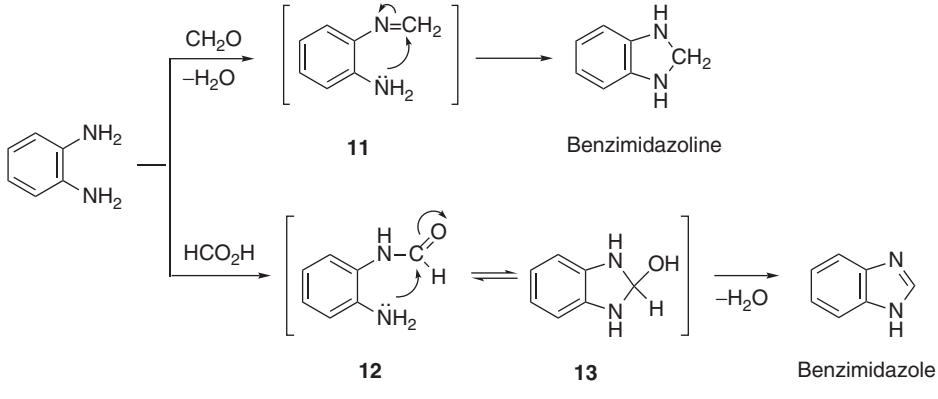


Figure 4.32 *Typical cyclization reactions of o-phenylenediamine.*

Subsequent nucleophilic addition of the amino group to the polarized C═N bond (in the case of compound **11**) or to the C═O bond (in compound **12**) takes place. Carbinol **13** appears to be the immediate precursor of benzimidazole. However, compound **13** is not isolated; similar compounds are almost instantly aromatized by the elimination of water.

In contrast to heterocyclization, ring interconversion reactions are accompanied by drastic alterations of the heterorings. Such reactions may occur with enlargement or contraction of the initial ring. Alternatively, the ring size can be preserved but its nature changed. Most ring interconversions are complex multistage processes which involve heteroring opening and subsequent cyclization. A familiar case is the conversion of pyrylium salts into pyridines upon treatment with aqueous ammonia (Figure 4.33a). In the first step one molecule of ammonia attacks the highly electron-deficient α-carbon atom of the pyrylium cation resulting in adduct **14**. Ring opening with the formation of acyclic product **15** is characteristic of such covalent adduct. Intermediate **16** then undergoes recyclization with participation of the amino and carbonyl groups.

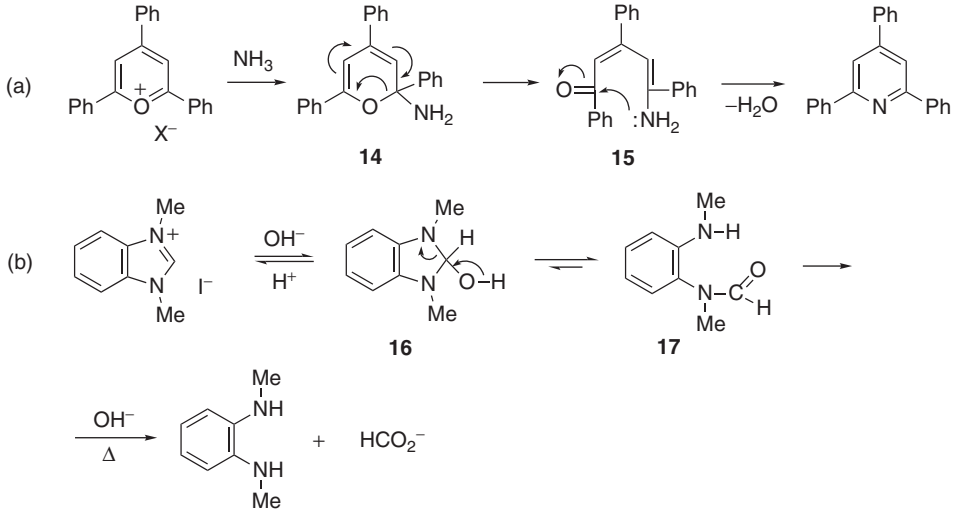
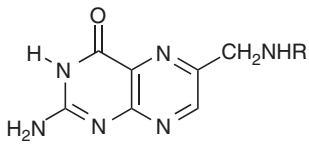


Figure 4.33 *Examples of heterocyclic cation transformations: (a) recyclization of the 2,4,6-triphenylpyrylium cation into triphenylpyridine, (b) elimination of the —CH═ group from the imidazole ring in the 1,3-dimethylbenzimidazolium cation.*

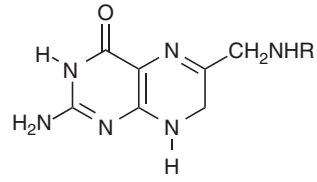
The majority of heteroring interconversions are initiated by nucleophilic attack. In certain circumstances a nucleophile may even eliminate a portion of the heteroring, and this frequently occurs when quaternary salts are treated with alkali. Even at ambient temperature, the 1,3-dimethylbenzimidazolium cation is attacked at the 2-position by hydroxide ion to form carbinol **16**, also called a pseudobase (Figure 4.33b). The pseudobase, like adduct **14**, exists predominantly in the acyclic tautomeric form **17**, which can be isolated. However, on heating, alkali deformylates **17** to produce *N,N′*-dimethyl-*o*-phenylenediamine and formic acid. The reaction as a whole represents the formal elimination of a —CH═ group from the imidazole ring. Examples of fragment eliminations from neutral heterocycles are also encountered, though they are not so common.

Cyclization and ring transformation reactions are widespread in heterocyclic chemistry and are utilized extensively in synthesis. It is not surprising that transformations of these types are found in a whole series of biochemical reactions in which insertions of one-carbon units such as Me, CH_2OH, CHO or —CH═ occur into a host of different molecules. Enzymes which transfer these groups use folic acid derivatives as coenzymes. The acid was named after the Latin word *folium* (leaf), for it was originally isolated from spinach leaves. As can be seen from Figure 4.34, folic

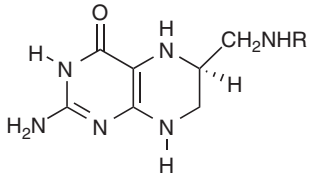
acid is a derivative of pteridine (Figures 1.6, 4.14) which has a complex substituent combining *p*-aminobenzoic acid and glutamic acid residues at the 6-position. Folic acid itself does not participate in biochemical reactions; the tetrahydro and, less often, the dihydro derivatives are the biologically active forms.



Folic acid



Dihydrofolic acid (**18**)



Tetrahydrofolic acid (**19**)

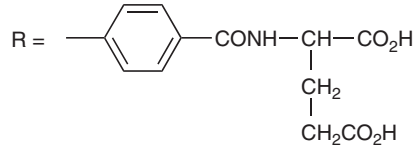
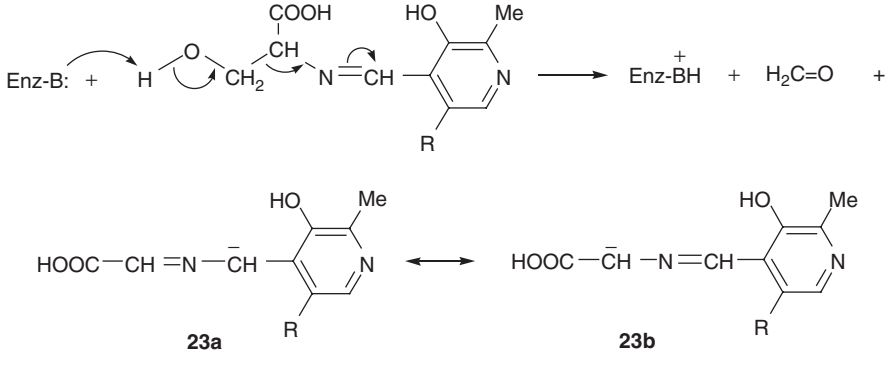


Figure 4.34 *Folic acid and its hydro derivatives.*

Tetrahydrofolic acid is first 'loaded' with a one-carbon fragment donated by the amino acid serine. In the course of this reaction, enzymatically catalyzed by serine transhydroxymethylase, tetrahydrofolic acid is converted into 5,10-methylenetetrahydrofolic acid (**20**) and serine is transformed into glycine (Figure 4.35) with the formal loss of a formaldehyde molecule.[6]

[6] Pyridoxal phosphate is a coenzyme of serine transhydroxymethylase. In general terms the formaldehyde formation from serine can be represented by the following scheme (Enz = enzyme, B: = a basic site in the enzyme, e.g., the imidazole ring of a histidine).



Anion **23b** is then protonated and hydrolyzed to produce glycine and pyridoxal phosphate.

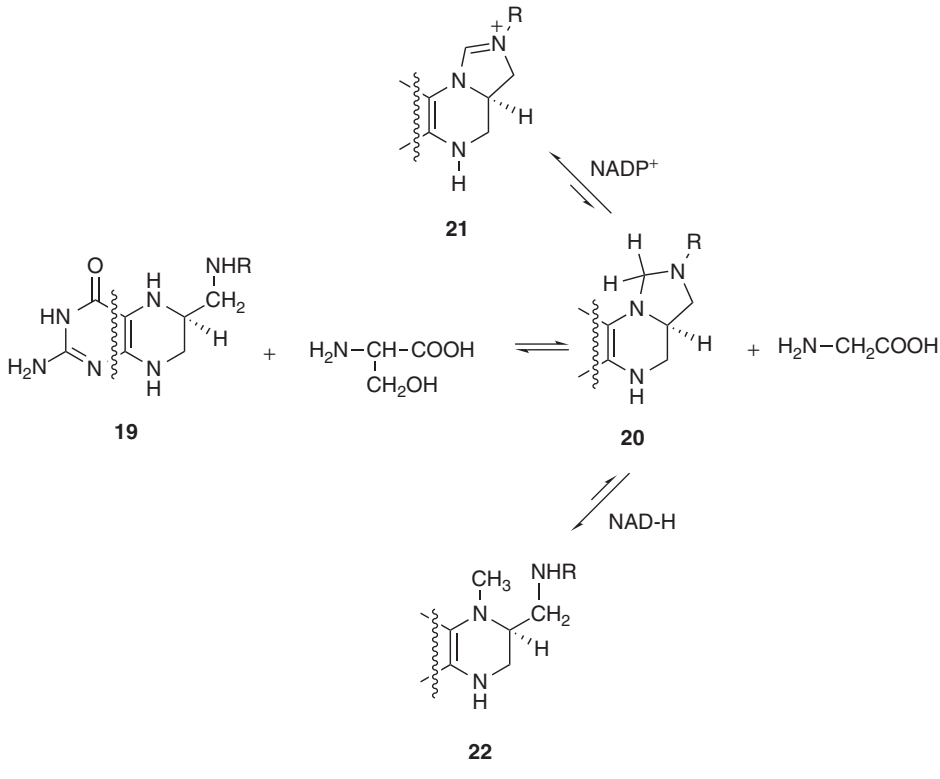


Figure 4.35 *Correlation between the tetrahydrofolic acid coenzymes containing the one-carbon fragment in different oxidation levels.*

The cyclization itself resembles the formation of benzimidazoline from *o*-phenylenediamine and formaldehyde (Figure 4.32). Similar to *o*-phenylenediamine, tetrahydrofolic acid contains the NHCCNH fragment of a vicinal diamine.

5,10-Methylenetetrahydrofolic acid (**20**) undergoes two important transformations within the organism. It can be oxidized to become 5,10-methenyltetrahydrofolic acid (**21**) by loss of a hydride ion to a dehydrogenase which contains $NADP^+$ as the coenzyme (Figure 4.35). Alternatively, by reaction with a reductase utilizing the NAD-H coenzyme, the acid is reduced with cleavage of the imidazole ring to form 5-methyltetrahydrofolic acid, that is, (**22**). All three of these derivatives of tetrahydrofolic acid, **20**–**22**, exhibit coenzymatic action. They differ from each other in the degree of oxidation of the one-carbon unit: in coenzyme **20** it is at the formaldehyde level, whereas in **21** the oxidation level is that of formic acid and in **22** that of methanol.

We now consider several instances of enzyme-mediated reactions which proceed with the participation of coenzymes **20**–**22**. One of the most important of those promoted by 5,10-methylenetetrahydrofolic acid is the conversion of 2′-deoxyuridylic acid (**24**) by *C*-methylation of the uracil ring into the 2′-deoxythymidylic acid (**25**) (Figure 4.36a). In the course of this transformation, coenzyme **20** is converted back into dihydrofolic acid (**18**), and the regeneration of tetrahydrofolic acid **19** necessitates the participation of an additional enzyme, dihydrofolate reductase, together with its coenzyme NADP-H (Figure 4.36b).

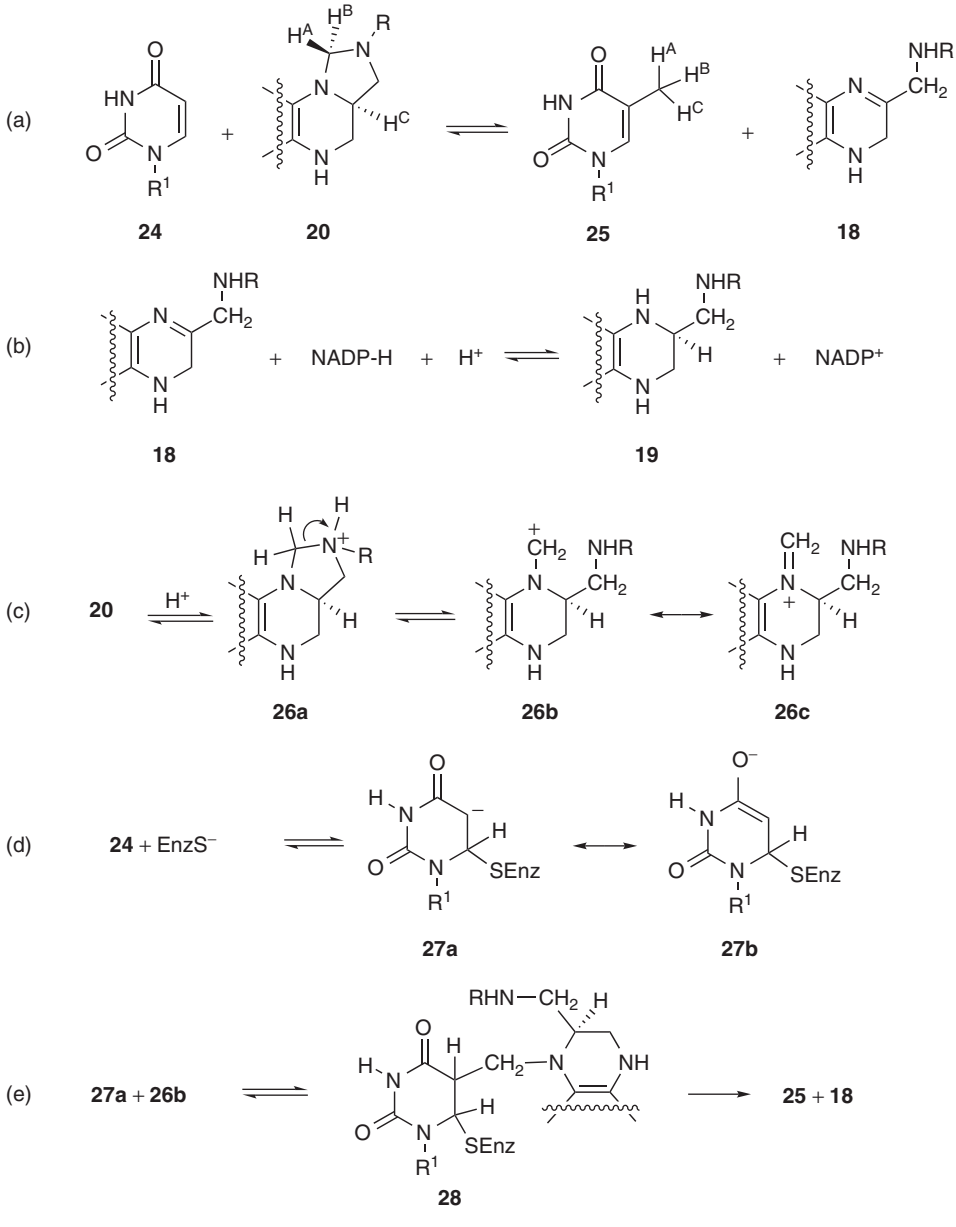


Figure 4.36 *Transformation of 2′-deoxyuridylic acid into 2′-deoxythymidylic acid (R^1 = the phosphorylated deoxyribose residue; see Figure 3.3): (a) general overview, (b) regeneration of tetrahydrofolic acid, (c) transformation of coenzyme* ***19*** *into its active methylating form* ***26b****, (d) activation of the uracil ring by enzyme addition, (e) proposed mechanism of C-methylation of the uracil nucleus.*

At first sight, coenzyme **20** appears to behave as though it introduces the carbon atom at the methanol oxidation level rather than at that of formaldehyde. This is not the case because the methyl group is created in the uracil ring in stages. Initially, the coenzyme fashions a methylene link containing H^A and H^B atoms (formaldehyde level of oxidation), the H^C atom, first connected to the enzyme C-6 atom, is transferred only at the next step. The details of the process are believed to be as follows. Coenzyme **20** functions in this reaction as the open chain analogue with a positively charged methyleneimmonium group **26** (in which the carbenium resonance structure **26b** participates) rather than as the cyclic imidazolidine **20** or its protonated form **26a** (Figure 4.36c). 2′-Deoxyuridylic acid in its turn is activated at the C-6 atom by a cysteine residue in an ancillary enzyme (Figure 4.36d). The nucleophilicity of the C-5 center in the corresponding intermediate, which can be represented as a resonance hybrid of structures **27a** and **27b**, is thus greatly enhanced. The methylene group of the coenzyme **26b** attacks the uracil ring at position C-5, and the addition product **28** then undergoes stepwise proton and enzyme elimination, C—N bond cleavage and hydride ion transfer from the C-6 atom of the coenzyme to give finally 2′-deoxythymidylic (**25**) and dihydrofolic (**18**) acids (Figure 4.36e). This formulation is, of course, simplified as the real processes are synchronous.

A major function of coenzyme **21** is its involvement in purine biosynthesis (Figure 4.37). The imidazole and pyrimidine ring closures require the incorporation of the purine C-8 and C-2 atoms.[7] In the course of these transformations coenzyme **21** undergoes certain preliminary changes. As

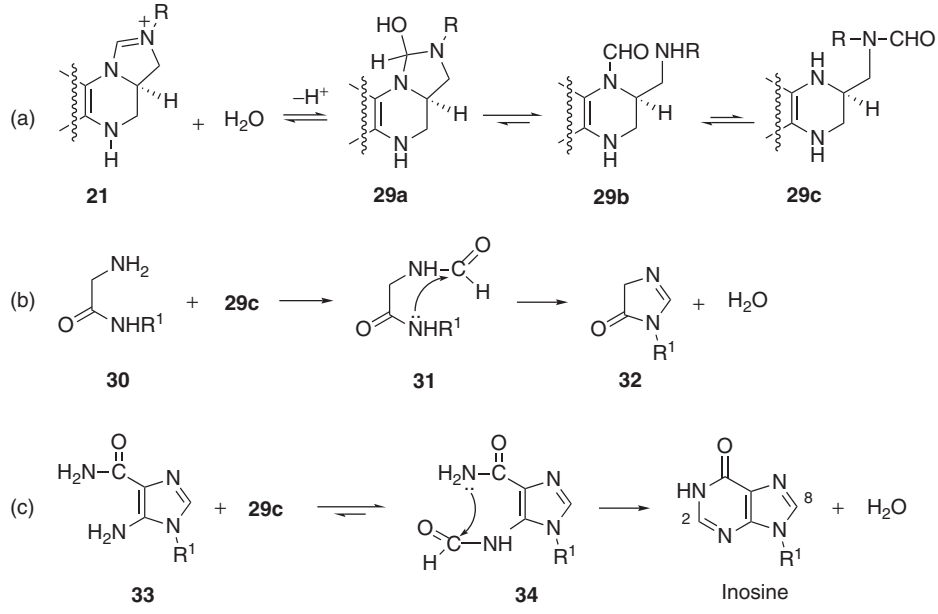


Figure 4.37 *Introduction of one-carbon fragments with the assistance of coenzyme **21** in the course of imidazole and pyrimidine ring closures during the biosynthesis of purines (R^1 = residue of phosphorylated ribose or deoxyribose; the scheme does not include the biochemical reaction sequence leading to compound **33**).*

[7] The numbering of the atoms in the purine system does not coincide with that in imidazole (Figure 2.19). In particular, the C-8 atom of purine corresponds to the C-2 atom of imidazole.

previously mentioned, treatment of imidazolium salts with alkali produces pseudobases which exist primarily in the acyclic form containing an *N*-formyl group (Figure 4.33, structure **17**). A similar situation occurs with 5,10-methenyltetrahydrofolic acid (**21**), which first adds hydroxide ion to yield a carbinol **29a** which is then stabilized by ring opening to one of the acyclic forms **29b** or **29c**. The 10-formyl derivative **29c** is the less stable but more active of the pair and is responsible for the transfer of formyl groups to *N*-ribonucleotides of glycineamide (**30**) and 5-amino-4-carbamoylimidazole (**33**). The intermediate *N*-formyl derivatives **31** and **34** undergo cyclization by dehydration (Figure 4.37b, c). The cyclized product from **34** is inosine which, in the course of subsequent reactions, provides the organism with all the necessary purine bases.

The integration of a methyl group into bioorganic molecules is carried out by means of 5-methyltetrahydrofolic acid (**22**), a coenzyme at the methanol oxidation level. A typical example is the synthesis of methionine from homocysteine (Figure 4.38). The enzyme homocysteine methyltransferase participates in the reaction alongside the coenzyme methylcobalamin which has a porphyrin-like structure incorporating a trivalent cobalt atom (see Section 4.3). The cobalt coordinates with the four pyrrole ring nitrogens and one of the nitrogen atoms of a 5,6-dimethylbenzimidazole unit. The sixth coordination site is essentially available to other ligands. In methionine formation, the methyl group of 5-methyltetrahydrofolic acid is first transferred to the methylcobalamin molecule to give a Co—Me linkage with displacement of the sixth (mobile) ligand. This Me group is later transferred to the SH function of homocysteine. Many details of this process are still not clear. For example, it is not yet known what promotes N—Me bond cleavage in coenzyme **22**, or in what form the methyl group is transferred (whether as cation, anion or radical).

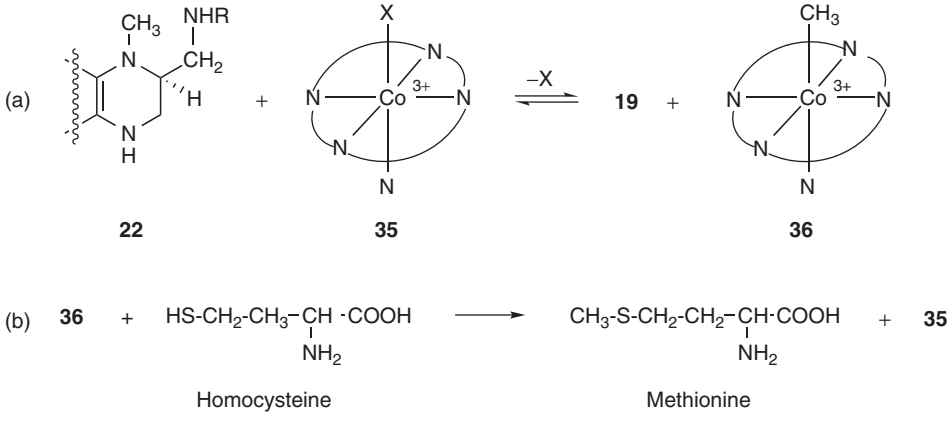


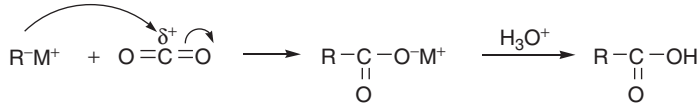
Figure 4.38 *Formation of methionine from homocysteine involving methyl group transfer by coenzyme **22** and methylcobalamin **36**.*

Biotin as a Carboxylic Group Carrier

Carbon–carbon bond formations are the cornerstones of organic chemistry, and the planning of any synthesis of an organic compound, whether it be simple or complex, must consider carefully the various possibilities of carbon skeleton construction. In living organisms carbon–carbon bonds are formed biosynthetically by a great diversity of constructions in an immense variety of bioorganic substances. As a rule, fragments containing one or two carbon atoms serve as the chemical building units used by a host of enzymes to form new C—C bonds. In this way coenzyme A combines

acetate units into chains while synthesizing fatty acids. The mechanisms of these reactions are similar to those already described for the cleavage of fatty acids, but the sequence is now reversed (refer back to Figure 4.22). Thiamine and pyridoxal phosphates mediate C—C bond formation by means of acyloin (Figure 4.25) and aldol condensations, respectively.

In the preceding section we considered the effect of the folic acid coenzymes, catalyzing the extension of a carbon skeleton by one atom in the synthesis of thymine from uracil (Figure 4.36). Carboxylation of organometallic compounds, to give the corresponding carboxylic acids, is one of the simplest and most widespread of reactions, both in nature and in the laboratory.



Just as for the majority of other reactions which produce new C—C bonds, carboxylation can be interpreted as a nucleophilic addition of a carbanion to a carbonyl function.

In an organism, the CO_2 molecule is transported and simultaneously activated by biotin. This coenzyme contains a bicyclic system of two condensed heterocyclic moieties, the completely saturated nuclei of thiophene and imidazole (Figure 4.39). The thiophene ring is substituted with a valeric acid residue which allows the biotin molecule to bind to an apoenzyme via an amide link. The influence of biotin is crucial at two stages. In the first stage a bicarbonate anion (equivalent to a CO_2 molecule) binds to the heterocyclic nitrogen. In the second stage this loosely held and activated carboxy group is transferred to the carbanion. In this fashion, the biosynthesis of oxaloacetic acid is achieved from pyruvic acid (for details, see Section 5.2.2). Another enzyme, working jointly with pyridoxal phosphate as the coenzyme, facilitates the generation of the pyruvic acid carbanion. As regards the addition of the bicarbonate anion to the cyclic amide nitrogen, it should be noted that carbonic acid, as well as the bicarbonate anion, is a hydrated form of carbon dioxide, and

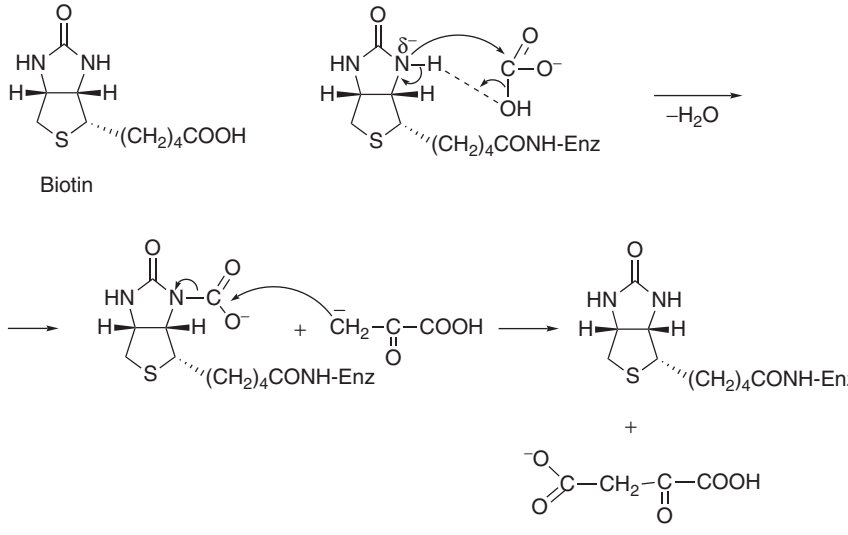


Figure 4.39 *The transformation of pyruvic acid into oxaloacetic acid as effected by biotin (Enz = apoenzyme residue).*

the three species exist in equilibrium. Therefore, in the reaction scheme the HCO_3^- ion may be interchanged with CO_2. Furthermore, all of the reaction centers of an enzyme–substrate complex are usually activated simultaneously; hence, in the transition state, the N—H bond in the amide and the C—OH bond in the HCO_3^- anion are polarized synchronously. This leads to a significant increase in both nitrogen nucleophilicity and carbon electrophilicity and, as a consequence, in the rate of the addition reaction.

4.3 Vitamins, the 'Molecules of Health'

We have seen that coenzymes are ubiquitous in their initiation and coordination of the chemistry of living organisms. We now consider the origin of the coenzymes. Major portions of coenzyme structures cannot be synthesized by human or other vertebrate organisms and must therefore be obtained from the diet. It has long been noticed that many of the vital precursors from which organisms are able to synthesize the required coenzymes contain nitrogen and may be classified as amines, more or less. The Polish scientist Funk named these substances vitamins (Latin: *vita*, life). For example, nicotinic acid (niacin) and its amide together with several related amides and esters constitutes the vitamin B_3 family. They are precursors of the pyridine-dependent dehydrogenases NAD^+ and $NADP^+$. Similarly, pyridoxal-5′-phosphate is an active form of vitamin B_6. It is biosynthesized from nutritionally supplied pyridoxine, pyridoxal and pyridoxamine which are considered as other members of the vitamin B_6 family (Figure 4.40). In some cases, vitamins are the actual coenzymes rather than their precursors: the coenzymes thiamine, riboflavin, biotin and folic acid are indeed vitamins B_1, B_2, B_7 and B_9, respectively.

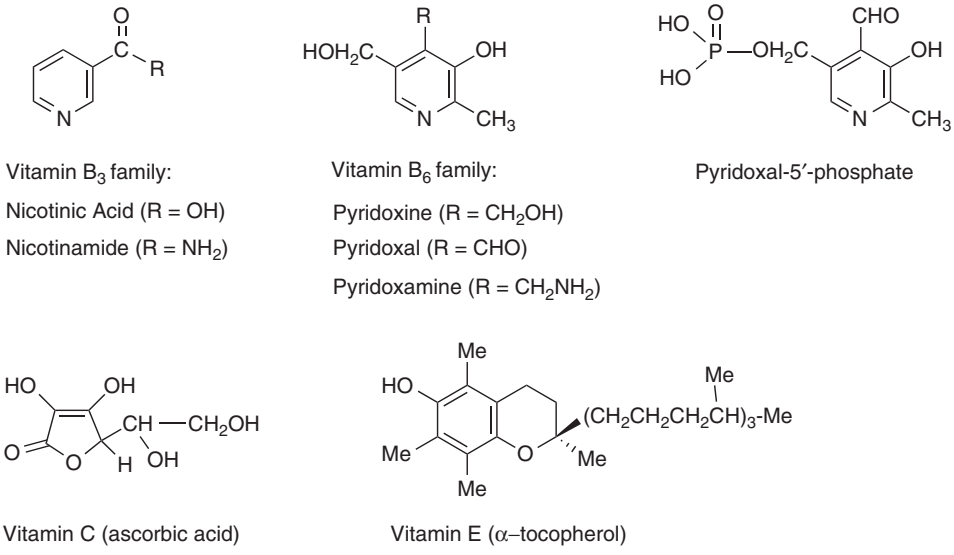


Figure 4.40 *Some heterocyclic vitamins.*

The biological activity of vitamins is so high that only a few milligrams of each are required daily. Thus, the daily requirements of nicotinic acid, riboflavin and pyridoxine are 15–20, 1–3

and 1–2 mg, respectively. Vitamin B_{12} shows even higher activity and the daily dose necessary for an adult is as low as 0.1–0.2 mg. This vitamin deserves special attention not only for its high activity but also because of its extraordinarily complicated chemical structure.

The distinguishing feature of vitamin B_{12} (cobalamin) is the ability of the cobalt(III) ion to coordinate with six donor ligands (Figure 4.41). The so-called corrin system serves as the main tetradentate ligand in this vitamin. Corrin resembles the porphyrins but differs in that two of the four pyrrole rings in corrin are directly connected (A and D rings), not through a —CH═ bridge as in the case of porphyrins. Moreover, the pyrrole rings in the corrin system are partially reduced. Vitamin B_{12} is produced by some fungi, algae and anaerobic bacteria, especially located in the alimentary canal of ruminants; therefore, corrin biosynthesis does not include steps which would require oxygen to produce a more highly oxidized aromatic porphyrin-type structure.

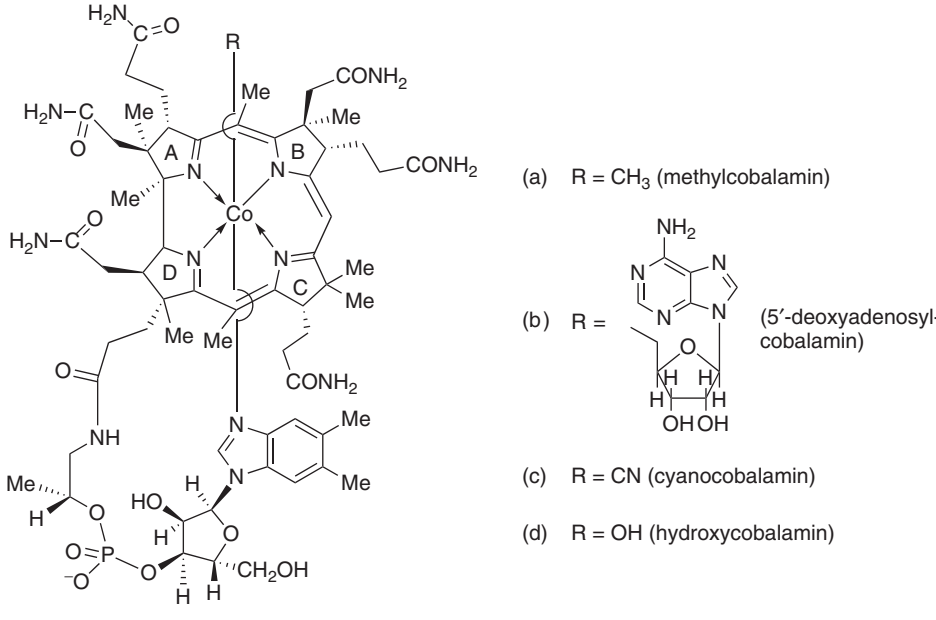


Figure 4.41 *Vitamin B_{12} (cobalamin) and its modifications.*

The fifth coordination site, oriented almost normal to the corrin plane, binds the cobalt atom with the N-3 atom of a 5,6-dimethylbenzimidazole unit. This unit is attached by a long chain formed by ribophosphate and amino acid units. The sixth ligand (R substituent in Figure 4.41) can vary, giving rise to several vitamin B_{12} modifications including methylcobalamin, adenosylcobalamin, cyanocobalamin and hydroxycobalamin. All of these analogues are present in organisms, the major component (about 50%) being the 5′-deoxyadenosyl derivative. A remarkable feature of methylcobalamin and adenosylcobalamin is the carbon–cobalt linkage. These coenzymes are the only known examples of naturally occurring organometallic compounds.

What are the functions of vitamin B_{12}? Its methyl group carrier properties have already been discussed in the synthesis of methionine (Figure 4.38). The 5′-deoxyadenosylcobalamin is a source

of the adenosyl radical that catalyzes various reactions that are triggered by the removal of a hydrogen atom from a compound. The vitamin is believed to have many other functions, among them rearrangement reactions, some of which are not yet clearly understood. However, the most important role involves the generation and management of blood. A lack of vitamin B_{12} causes pernicious anemia, one of the most serious of blood diseases. The main dietary source of this vitamin is meat, especially liver. At present vitamins B_{12} are used as food and cosmetic biologically active additives. They can diminish the risk of heart diseases and are useful as prophylactics against the brain problems of aging people. This vitamin is produced biotechnologically on an industrial scale by acetone-butanol conversion of sugars by agitating with the mixed anaerobic methane forming bacteria cultures at pH 6.4 in the presence of cobalt chloride and methanol, which increases the production of the vitamin.

Four Nobel prizes have been awarded in the areas of Chemistry, Physiology and Medicine for studies of the structure and biochemistry of vitamin B_{12}. Nevertheless, new discoveries in this field continue to appear. Thus, the biosynthesis of the 5,6-dimethylbenzimidazole ligand remained unclear for a long time. Only recently, was the enzyme BluB that is responsible for this process isolated and studied. The 5,6-dimethylbenzimidazole unit is produced from the adduct of flavin mononucleotide with an oxygen molecule (Figure 4.14, structure **1**) and results in a complex merging of hydrolytic and redox reactions still not well understood (Ealick, S.E. and Begley, T.P., *Nature*, 2007, **446**, 387). In any case this is one more example of heterocyclic ring transformation in biochemistry.

Other vitamins with a heterocyclic structure are known. Some of the most familiar are vitamin C (ascorbic acid) and vitamin E (α-tocopherol), which contain tetrahydrofuran and tetrahydrobenzopyran rings, respectively (Figure 4.40).

To sum up this section we would like to add to the ancient wisdom '*memento de mortis*' a relatively new biochemical maxim '*memento de vitaminum*', or even '*memento de heterocyclus*'.

4.4 Ribozymes: Vestiges of an Ancient World

For decades it seemed axiomatic that proteins are the only biological catalysts. However, in 1967 Francis Crick, Leslie Orgel and Carl Woese hypothesized that RNAs due to the high flexibility of their tertiary structure could also fulfill this function. Somewhat later this was confirmed experimentally by Tom Cech and Sidney Altman and in 1989 these two scientists were honored with the Nobel Prize in Chemistry. This discovery opened a new chapter in molecular biology and since then catalytic RNAs, called ribozymes, have attracted great interest. Ribozymes are found in the nucleus, mitochondria and chloroplasts of eukaryotic organisms and in some viruses. Many artificial ribozymes are now prepared synthetically. Ribozyme's main substrates are other RNA molecules or the ribozyme itself. Ribozymes can act as molecular scissors to cleave the phosphodiester bonds or as 'molecular staplers' to ligate RNA pieces. It is also believed that the ribosome which displays aminotransferase activity is actually a special kind of ribozyme (Section 3.4, Figure 3.14).

Ribozymes are relatively small molecules that encounter not more than several tens of ribonucleotide residues. Outwardly, many of them resemble a transfer RNA structure (Figure 3.8). Some of their sequences are paired with others in accordance with Watson–Crick rules and thus form a stem. Others remain unpaired and form loops (Figure 4.42a, b). Both stems and loops tend to adopt a helix-like tertiary structure. Most ribozymes have divalent metal ions, normally Mg^{2+}, as cofactors. Among the five classes of ribozymes the hammerhead ribozymes are the best known. They are named for their molecular shape when drawn flat on a piece of paper. Hammerhead

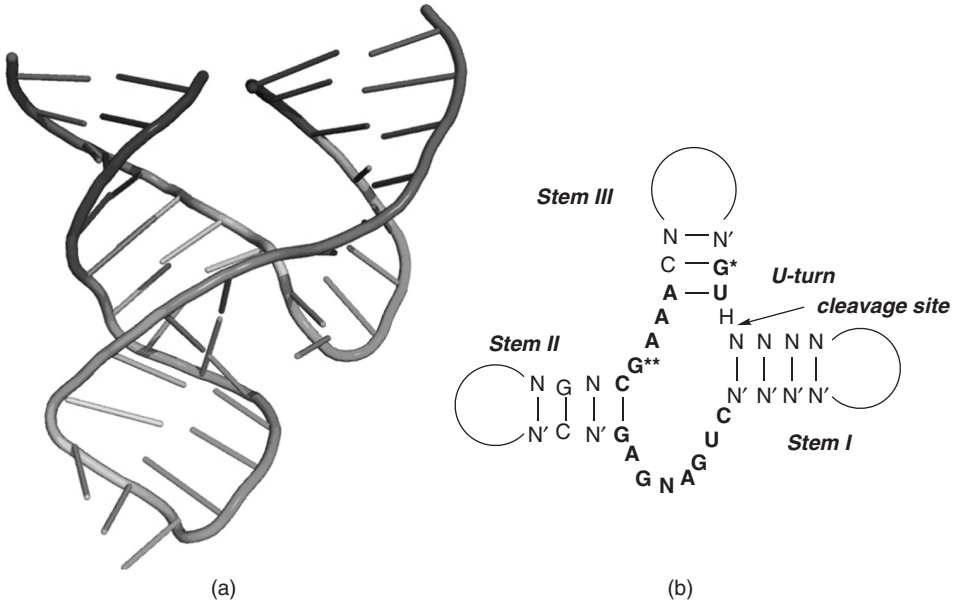


Figure 4.42 *Typical structures of hammerhead ribozyme: (a) tertiary structure (author: William G. Scott, 17 June 2007; retrieved from: http://commons.wikimedia.org/wiki/File:Minimal_hammerhead_ribozyme_structure.png; accessed 29 October 2010), (b) secondary structure (retrieved from: http://www.tulane.edu/~biochem/nolan/lectures/rna/ham.htm; accessed 29 October 2010).*

ribozymes catalyze the cleavage of an RNA backbone and their minimal consensus structure is shown in Figure 4.42b. The sequences depicted by bold letters are necessary for efficient cleavage. Ribonucleotides indicated by N and N′ are any complimentary purine–pyrimidine bases. Half-circles indicate that any number of bases may connect the ends of each stem, or they need not be connected at all. The self-cleavage site is placed always at the so-called U-turn and follows the sequence G–U–H, where G and U are conserved and crucial for catalysis and H is any base but G.

As in the case of protein enzymes, the mechanism of ribozyme catalysis follows the same general principles. Thus, general acid–base catalysis and stabilization of a transition state with various kinds of noncovalent interactions are especially important. Figure 4.43 illustrates this schematically for the self-cleavage of the phosphodiester bond between two cytosine-containing nucleotides. Stem to stem interactions that are crucial for ribozyme activation bring together in a reaction complex the cleavage site and the guanosine residues G* and G** (Figure 4.43a). Being rather acidic (pKa $\sim$ 9.3) the guanine fragment G** generates an anion that acts as a general base and accepts a proton from the ribose 2′—OH group. Negative charge thus arising on an oxygen atom facilitates its subsequent nucleophilic attack on a phosphorus atom. Another P—O bond is broken, in concert with the participation of the G* ribose residue providing acid catalysis. The transition state (Figure 4.43b) is stabilized by multiple hydrogen bonds and Mg^{2+} ions that are attracted by the negatively charged oxygen atoms. Finally, all participants in the process are released (Figure 4.43c). The 2′,3′-cyclic phosphate is not hydrolyzed by the ribozyme.

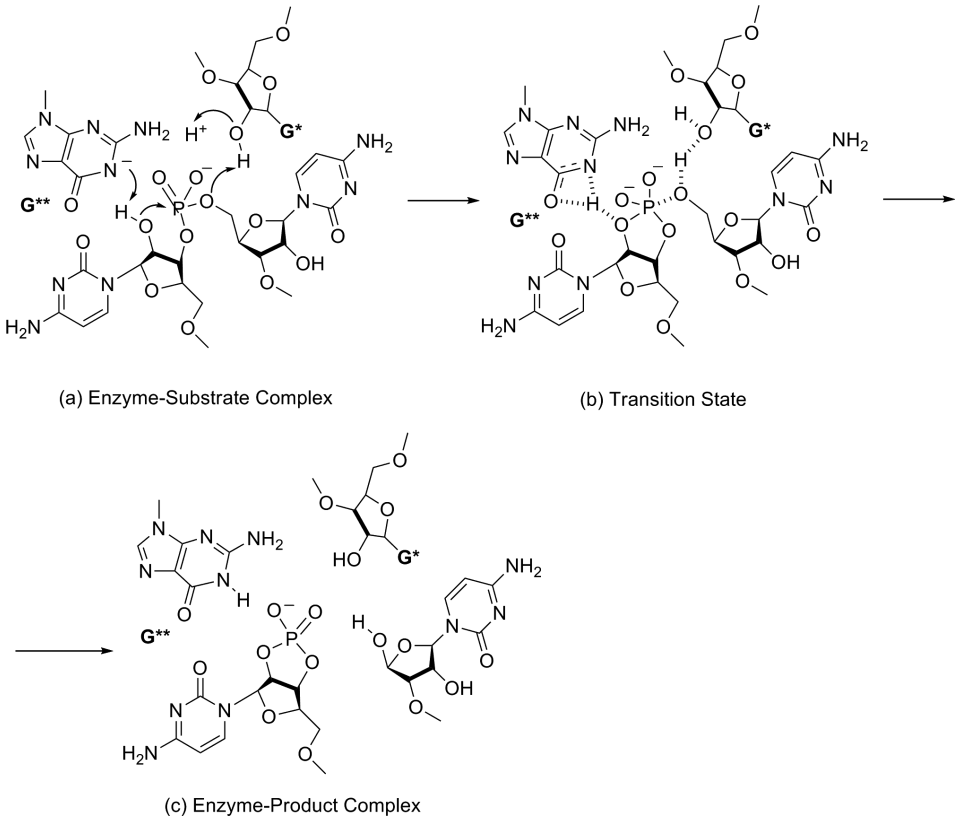


Figure 4.43 *Proposed catalytic mechanism for a hammerhead ribozyme self-cleavage reaction (Reprinted with permission from Chi, Y-I., Martick, M., Lares, M., Kim, R., Scott, W.G., et al. (2008)* Capturing Hammerhead Ribozyme Structures in Action by Modulating General Base Catalysis. *PLoS Biol 6(9): e234. doi:10.1371/journal.pbio.0060234).*

Cleavage of other RNA molecules involves a primary binding between a ribozyme and a substrate via Watson–Crick pairing (Figure 4.44a, b). Further bond breaking and product separation are shown in Figure 4.44c, d. From a practical point of view, the design of artificial ribozymes capable of reacting with a high regiospecificity to destroy any RNA molecule is especially important. For example, such a ribozyme designed to cleave the RNA of HIV and inserted into a cell would cleave all the incoming virus particles and prevent infection. The main problems in this field are to make medicinal ribozymes stable chemically and to develop the effectiveness of their transportation to the biotargets. Many other similar possibilities are opened for application in human therapy, gene engineering, biosensor devices and so on.

The discovery of ribozymes led to an exciting idea known as the ‘RNA world’. This postulates that, during the early stages of life on the Earth, small RNA molecules served both as the first biological catalysts and as the keepers of hereditary information. Later, during the process of

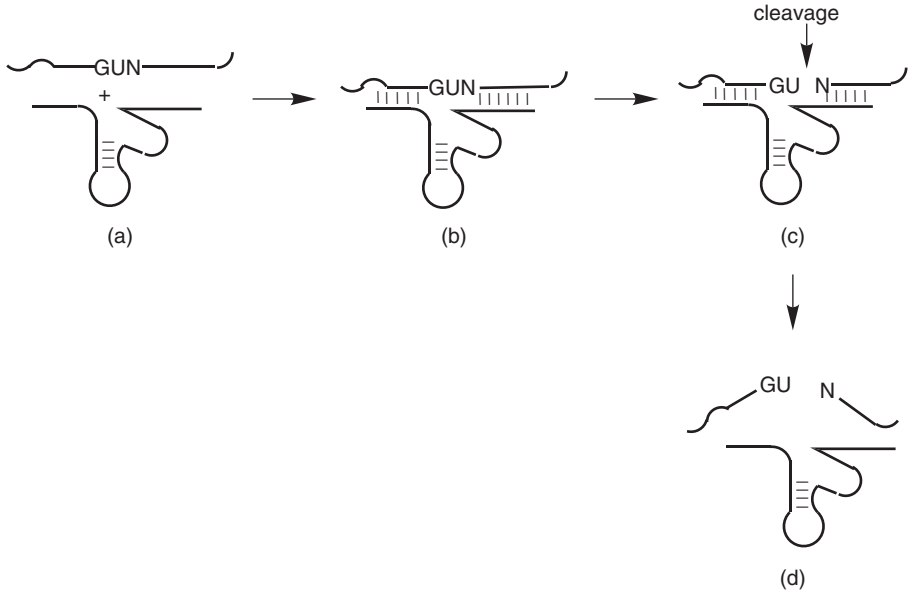


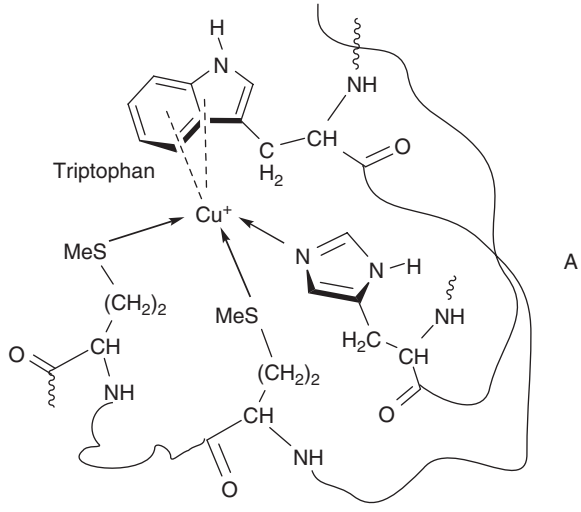
Figure 4.44 *Schematic representation of an RNA cleavage reaction: (a, b) binding of ribozyme and substrate, (c) cleavage reaction, (d) two product strands dissociate. Reprinted from http://academic.brooklyn.cuny.edu/chem/zhuang/QD/toppage1.htm.*

chemical evolution, ribozymes were replaced by chemically more stable and versatile protein enzymes and DNAs.

Diverse applications of heterocyclic coenzyme in biotechnology and science are discussed in Section 11.3.1.

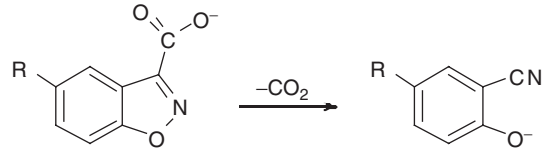
4.5 Problems

1. 2,6-Dimethyl-3,5-diethoxycarbonyl-1,4-dihydropyridine does not reduce benzaldehyde or *m*-nitrobenzaldehyde under normal conditions. In contrast, salicylic and especially 5-nitrosalicylic aldehydes are readily reduced. Account for this fact.
2. The reduction of flavins by NAD-H is catalyzed by magnesium(II) and zinc(II) ions. Suggest a mechanism for the catalysis.
3. Xanthine oxidase, a key enzyme which controls the levels of all purines in living organisms, assists the oxidation of hypoxanthine, xanthine and their precursors (adenine and guanine) to uric acid. Write the stepwise mechanism of the hypoxanthine → xanthine reaction. (Hint: the first step involves hydration of the C═N bond; FAD serves as the cofactor.)
4. Proteins usually bind metals through covalent interactions. However, the bacterial copper-trafficking protein CusF binds Cu(I) ion in a three-coordinate structure with two methionines and a hystidine. Copper also interacts via a cation-π interaction with the indole ring of a nearby tryptophan as shown in structure (A). When the tryptophan is replaced with methionine, the complex holds Cu more tightly. Suggest a possible role of metal cation-π interaction in activity of CusF.



5. What role does the 3-hydroxy substituent play in pyridoxal phosphate?
6. Draw a scheme showing the decarboxylation of tryptophan to form tryptamine catalyzed by pyridoxal phosphate.
7. Which heterocyclic coenzymes participate in the formation of new carbon–carbon bonds during biochemical reactions?
8. Which widespread heterocyclic fragment is found in the structures of many coenzymes, vitamins and energy-rich compounds (see also Chapter 5)?
9. Coenzyme PQQ (Figure 4.8b) was first isolated as the product of nucleophilic addition of acetone to one of its carbonyl groups. Similar products are formed by PQQ with many other nucleophiles such as methanol. Which one of the two carbonyl groups in PQQ is more active? Draw the structures of the PQQ adducts with acetone and methanol.

10. A biologically important decarboxylation reaction has been extensively studied using the conversion of benzisoxazole-3-carboxylates (B) into *o*-cyanophenolate anions (C) as a model. These species are convenient model compounds as they are not susceptible to general acid-base catalysis but are sensitive to other media conditions. Predict how the rate of the enzyme-catalyzed decomposition of carboxylates (B) would change if the protic solvent was replaced by an aprotic one?

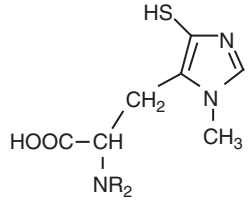


B C

11. The biochemical pathways of fertilization include protection of the embryo via oxidative crosslinking of its protein surface, as demonstrated in sea urchin eggs. The mechanism requires the bioproduction of hydrogen peroxide (which has a deleterious effect on the embryo) as regulated by the enzyme oxidase. H_2O_2 evolution results from an increased oxygen uptake by the fertilized egg (called the 'respiratory burst of fertilization'):

$$NADP - H + H^+ + O_2 \xrightarrow{\text{oxidase}} NADP^+ + H_2O_2$$

The enzyme peroxidase catalyzes the consumption of hydrogen peroxide in conjunction with the intracellular family of antioxidants known as ovothiols (D). (a) To which widespread, naturally occurring compound are the ovothiols related? (b) How can H_2O_2 levels be maintained by the use of ovothiols? Indicate schematically the ovothiol conversion. (c) Provide a mechanism for regeneration of the consumed ovothiols.



D

$NR_2 = N(CH_3)_2$ (ovothiol A)
$NR_2 = NHCH_3$ (ovothiol B)

12. Hammerhead ribozyme is rather unstable to cleavage. To prevent its decomposition on growing of crystals for X-ray measurements the guanine residue G** (Figure 4.42b) was replaced by adenine. As a result, the rate of cleavage decreases about a million times without significant structural changes. Account for this fact taking into consideration that the most basic site in adenine is the N-1 atom ($pK_a = 3.45$).

4.6 Suggested Reading

1. Lehninger, A. L., Nelson, D. L. and Cox, M. M., *Principles of Biochemistry*, 4th edn, Worth, New York, 2004.

2. Branden, C. and Tooze, J., *Introduction to Protein Structure*, 2nd edn, Garland, New York, 1999.
3. Wyman, J. and Gill, S. J., *Binding and Linkage. Functional Chemistry of Biological Macromolecules*, University Science, Mill Valley, Calif., 1990.
4. Stryer, L., *Biochemistry*, 6th edn, Freeman, New York, 2006.
5. Ferdinand, W., *The Enzyme Molecule*, Wiley, Chichester, 1987.
6. Jencks, W. P., *Catalysis in Chemistry and Enzymology*, Dover, New York, 1987.
7. Prentis, S., *Biotechnology: a New Industrial Revolution*, Orbis, London, 1984.
8. Kirby, A. J. and Hollfelder, F., *From Enzyme Models to Model Enzymes*, RSC Publishing, London, 2009.
9. Mansoorabadi, S. O., Thibodeaux, C. J. and Liu, H. The diverse roles of flavin coenzymes – nature's most versatile thespians, *J. Org. Chem*., 2007, **72**, 6329.

5

Heterocycles and Bioenergetics

I flow like a river, as lilac I bloom,
Silent as a shadow, I flame as a bonfire,
Blaze as the Sun, and shine as the Moon
And flutter as a colored butterfly.

I. Severyanin

All living organisms need energy for such diverse purposes as nerve impulses, movement, nutrient transport and maintenance of body temperature. Energy is also consumed in the biosynthesis of complex molecules such as nucleic acids, proteins, polysaccharides, phospholipids, steroids and so on. In fact, all living matter on Earth owes its existence to solar energy. However, only plants and certain bacteria can directly assimilate solar energy. The physicochemical processes by which solar radiation is absorbed by plants are collectively called photosynthesis. This phenomenon will be dealt with in Chapter 6. The present chapter is devoted to the problems of energy supply in those living organisms without a photosynthetic system, that is, the animal kingdom, including mankind.

Animals receive all of their required energy from food. In proteins, fats and carbohydrates – the main constituents of the diet – energy is stored in chemical bonds formed during photosynthesis. The process used by organisms for the release of energy from nutrient molecules is similar to that utilized by man to obtain heat from wood, coal or natural gas. This process is oxidation. The oxidizing agent used by the body is molecular oxygen inhaled during respiration. At some of the intermediate stages of biological oxidation, the role of oxidant may also be performed by organic molecules such as pyruvic acid.

Large quantities of energy are evolved in the oxidation of proteins and especially of fats and carbohydrates. For example, the combustion of one mole of palmitic acid to carbon dioxide and water liberates 2338 kcal, while one mole of glucose liberates 686 kcal:[1]

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + 686 \text{ kcal}$$

[1] Here and elsewhere, the energies quoted are free energy ΔG^{o} values. In the cell, of course, the true ΔG will be affected by the relative concentrations of the products and starting materials, according to the Nernst equation. In living cells products and starting materials are kept near equilibrium values, and therefore the true ΔG is often significantly larger than the $\Delta G^{o\prime}$.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

It is clear that if such quantities of energy were to be liberated instantly, the organism would neither be able to utilize them completely nor to conserve the energy for the future. Therefore, in the course of evolution, living organisms have developed intricate oxidative–reductive mechanisms which enable slow, adjustable rates of oxidation of food. The basis of the oxidation–reduction system is a set of diverse enzymes, chief among them being the pyridine- and flavin-dependent dehydrogenases and cytochromes. The heterocyclic compound adenosine triphosphate (ATP) is the main substance responsible for the accumulation and supply of energy for all cellular processes.

5.1 ATP as the Universal Currency of Energy

When we buy something, we usually pay with small change and bills of different values. Similarly, organisms pay for their expenditures of energy, small and large, with a type of 'molecular currency'. Such a system is required because physiological processes are based on chemical reactions with activation energies usually in the range of 7–15 kcal mol^{-1}. Thus, units of energy of the same order of magnitude need to be available.

The main molecular carrier of energy, the unit of the biological 'energy coinage', is the tetraanion of adenosine triphosphoric acid, namely adenosine triphosphate (ATP; at physiological pH, the acidic protons of adenosine triphosphoric acid are almost completely ionized). The mode of action of ATP is based on the hydrolytic elimination of one of the two terminal phosphate groups to form adenosine diphosphate (ADP) or both to form adenosine monophosphate (AMP; Figure 5.1). These processes each release approximately 9 kcal mol^{-1} of energy.

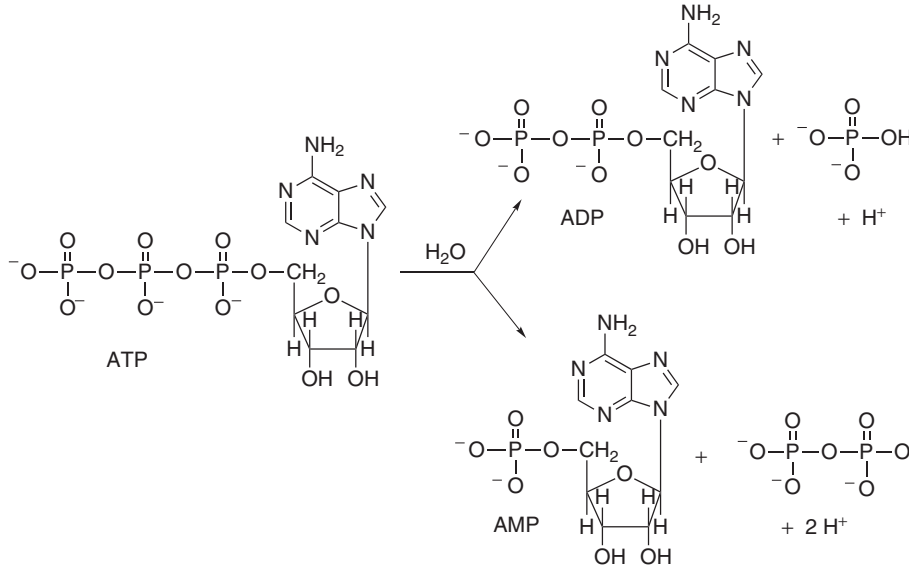


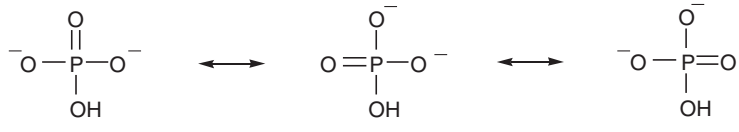
Figure 5.1 *Hydrolytic cleavage of the phosphate groups of ATP with the formation of ADP and AMP.*

In some biochemical reactions, energy transfer is carried out by guanosine triphosphate (GTP), uridine triphosphate (UTP) or cytidine triphosphate (CTP). Though there is no evidence yet for the

direct participation of the heterocyclic moieties of any of these compounds in the energy transfer, the presence of a purine unit in ATP suggests some function: the heterocyclic bases probably bind the ATP, GTP, UTP and CTP molecules (by noncovalent interactions) to the enzymes which transport them.

A number of effects facilitate the liberation of energy by hydrolysis of the pyrophosphate bonds in ATP. First, hydrolysis provides additional negatively charged oxygen atoms. As a result, the products forming on cleavage of ATP are more effectively solvated than ATP itself.[2] In other words, the strength of the bonds between phosphate residues in ATP together with the 'hydration bonds' between ATP and water are less than the strength of the 'hydration bonds' between ADP + phosphate and water.

Second, the negatively charged oxygen atoms in ATP mutually repel each other; and when the length of the pyrophosphate chain decreases, the repulsion between the remaining oxygen atoms becomes less pronounced. Moreover, ATP hydrolysis increases the possibility of resonance stabilization in the resulting adenosine diphosphate and phosphate anion compared with the initial molecule of ATP. Overlapping of the unshared electron pairs of such oxygen atoms with the vacant 3*d*-orbitals of the phosphorus atom gives rise to a substantial gain in energy:



All the above clarifies why an equilibrium ATP $\rightleftarrows$ ADP + phosphate in water is strongly shifted to the right. Nevertheless, despite the hydrolytic instability of ATP in living cells, the ATP/ADP ratio is maintained at ATP concentrations 1000-fold higher than the concentration of ADP. This displacement from equilibrium is caused by the ability of ATP to chelate metal ions very strongly. Due to the strength of such interactions, ATP exists in the cell mostly complexed with Mg^{2+} with a binding constant about 9500. This strong disturbance of equilibrium in favor of ATP provides for the evolution of a large amount of energy on hydrolysis.

In biochemistry the energy balance of a reaction is preferably expressed not in terms of the heat evolved ΔH but as the free energy change ΔG^{o}. The ΔG^{o} is connected to the equilibrium constant K by the expression:

$$\Delta G^{o} = -RT\ln \mathrm{K}$$

where R is the gas constant and T is the absolute temperature. When ΔG^{o} is negative, the products of the reaction predominate in the equilibrium mixture. By contrast, the equilibrium is shifted in favor of the reactants when ΔG^{o} is positive. Obviously, if $\Delta G^{o} = 0$, then $K = 1$; that is, the concentrations of the products and the reactants are equal in the equilibrium mixture. Many biochemical reactions are unfavorable from the thermodynamic point of view. This means that under standard conditions their ΔG^{o} values are positive. To shift the equilibrium toward product formation requires energy. The key problem becomes the mechanism of energy transfer from ATP to the reactants.

A constant temperature must be maintained for the normal functioning of the living cells of many organisms. Therefore, the heat from the hydrolysis of ATP for activation of the biochemical reactions in the cell can not be directly utilized. The biochemical procedure used to transfer energy is the so-called coupling of separate reactions by means of a common intermediate. Such

[2] The actual amount of energy liberated by hydrolysis of the pyrophosphate bonds in ATP is considerably influenced by solvation energies. Because of the strong solvation of the phosphate anion, the energy liberated is higher in a phase of low dielectric constant.

is the case, for example, in the esterification of carboxylic acids by alcohols. Under conventional conditions the equilibrium for this reaction is largely shifted toward the reactants. However, the introduction of ATP molecules into the reaction sequence allows the process to be modified to give more of the ester. The process may be represented as shown in Figure 5.2.

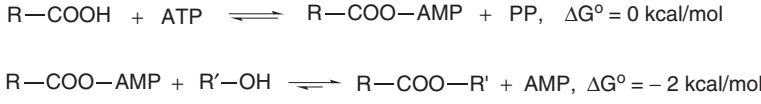


Figure 5.2 *Coupling of two reactions during the esterification of carboxylic acids in the presence of ATP.*

The initial step represents an activation of the carboxylic acid. In the course of this activation a pyrophosphate (PP) unit separates from the ATP molecule and a mixed anhydride of the carboxylic acid and adenosine monophosphoric acid is formed. The energy previously stored in the ATP molecule is now conserved in a new P—O bond between the acyl group of the carboxylic acid and AMP. This bond is called an energy-rich bond and can be compared with a compressed spring ready to be released at any moment. The release of this 'spring' takes place in the second step when the mixed anhydride reacts with an alcohol to produce the ester and AMP. This step has a ΔG^{o} value of – 2 kcal mol^{-1} which shifts the equilibrium to the right side of the equation, toward esterification. The following analogy may help. If you push your car (with a dead battery) up a steep incline by the shortest route, you must expend some effort. However, you will be compensated even if you choose to go downhill by a longer route. Such is the general principle of ATP action as the energy source.

To conserve energy in ATP, ATP is synthesized by the addition of orthophosphate to ADP or pyrophosphate to AMP. Clearly, these reactions necessitate energy expenditures. All the organic derivatives of phosphoric acid such as phosphoenolpyruvate, 1,3-bisphosphoglycerate, creatine phosphate and acetyl phosphate are energy-rich compounds. Indeed, they all are richer in energy than ATP, and can therefore operate as phosphate transfer agents on AMP and ADP (Figure 5.3

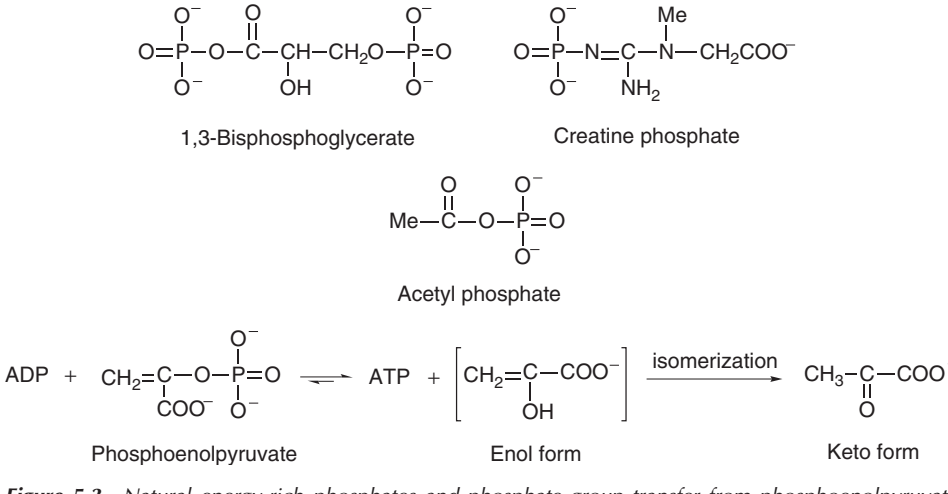


Figure 5.3 *Natural energy-rich phosphates and phosphate group transfer from phosphoenolpyruvate to ADP.*

and Table 5.1). In muscle tissues, where energy consumption is very high, ATP molecules are replenished by means of phosphate group transfers from creatine phosphate to ADP.[3] Table 5.1 reveals that phosphoenolpyruvate (an ester of phosphoric acid and pyruvic acid in the enol form) possesses the most energy-rich phosphate group of all the carriers. The additional energy capacity results from the lability of the enol form, which instantly isomerizes to the more stable keto form following donation of the phosphate anion.

Table 5.1 *Standard free energies of hydrolysis (ΔG^o) for some biologically important phosphates (adapted from Lehninger, A. L.,* Biochemistry, *Worth, New York, 1970, Chap. 14, p. 302, with permission)*

Compound	ΔG^o (kcal mol^{-1})	Compound	ΔG^o (kcal mol^{-1})
Phosphoenolpyruvate	−14.8	Glucose 1-phosphate	−5.0
1,3-Bisphosphoglycerate	−11.8	Fructose 6-phosphate	−3.8
Creatine phosphate	−10.3	Glucose 6-phosphate	−3.3
Acetyl phosphate	−10.1	Glycerol 1-phosphate	−2.2
ATP (to ADP)	−7.3		

Thus, the role of the ATP–ADP system in an organism is to transfer phosphate groups from the high energy phosphates to the low energy acceptors, such as glucose. By inspecting Table 5.1 one can conclude that this role of ATP is facilitated by its median position in the thermodynamic scale, because it readily receives phosphate groups from the high energy phosphates and also can easily transfer them to the low energy acceptors. As we shall see later, in Section 5.2.3, there exists a further method of chemical energy transfer in living cells, namely electron transfer.

5.2 Breathing

We now consider the main energetic principles of the chemical processes of breathing. Breathing is a method of extracting energy from fuel in the cell in a regulated manner by oxidation with oxygen. Respiration is a very complex chemical process comprising numerous reactions in which dozens of enzymes take part. To achieve a better understanding of the essence of respiration, we divide the whole process into three stages, as illustrated in Figure 5.4.

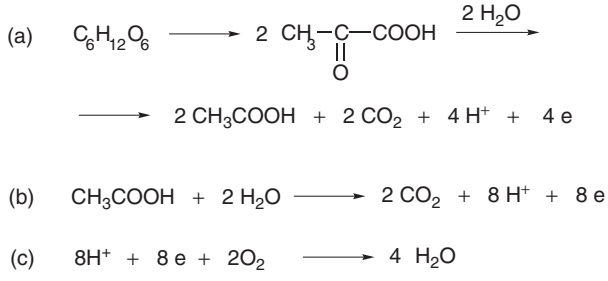


Figure 5.4 *Stoichiometric outcomes of the three principal stages of the respiratory process: (a) glycolysis and the link reaction, (b) the tricarboxylic acid (Krebs) cycle, (c) the respiratory chain.*

[3] Creatine phosphate is particularly suitable for the transfer of a phosphate group from itself to ADP and overall the process is almost energy neutral.

For example, in the oxidation of the carbohydrate glucose, the first stage, called glycolysis, involves breakdown of the sugar molecule into two three-carbon pyruvic acid residues which then undergo oxidative decarboxylation to give acetic acid in the form of acetyl CoA, MeCOSCoA. In the second stage, called the tricarboxylic acid cycle or Krebs cycle, the molecules of acetic acid are degraded with the participation of water and special electron acceptors. This degradation liberates hydrogen ions, electrons and CO_2. The third stage of the respiratory process is named the respiratory chain and concerns the transfer of the electrons, produced in the second stage, to oxygen by means of a complex series of electron carriers. Each stage of respiration produces energy for the organism from the processed raw materials, energy which is then stored in ATP molecules.

Although the first stage in the biochemical processing of fats and proteins differs slightly from glycolysis, acetic acid (in the form of MeCOSCoA) is, nevertheless, always the crucial intermediate at the end of the first stage. The second and third stages are identical regardless of the source of the cellular fuel. We now consider each stage of the respiratory process in detail.

5.2.1 Glycolysis

Glycolysis commences with the phosphorylation of a glucose molecule by ATP to produce D-glucose 6-phosphate. The glucose molecule is thus activated for a subsequent conversion; moreover, in the phosphate form it can be more readily and more strongly linked to the appropriate enzyme. D-Glucose 6-phosphate is then isomerized into D-fructose 6-phosphate which, in turn, is phosphorylated by another ATP molecule to yield D-fructose 1,6-bisphosphate (Figure 5.5).

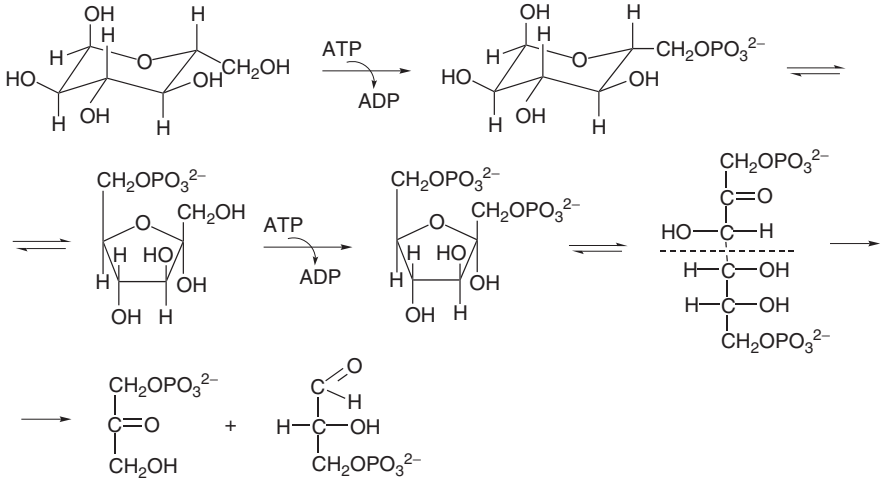


Figure 5.5 *Cleavage of a glucose molecule into two three-carbon fragments in the first steps of glycolysis.*

As for all monosaccharides, D-fructose 1,6-bisphosphate exists as an equilibrium mixture of cyclic and acyclic forms. The next step in glycolysis involves cleavage of the open chain form of D-fructose 1,6-bisphosphate into two three-carbon fragments: dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate (Figure 5.5). This transformation, a retroaldol condensation, demands considerable energy (equilibrium $\Delta G^{o} = 5.7$ kcal mol^{-1}) and is normally the rate-determining step in glycolysis. Dihydroxyacetone phosphate is isomerized into D-glyceraldehyde 3-phosphate by the action of an enzyme. As a result, one molecule of glucose gives two molecules of glyceraldehyde 3-phosphate.

From the point of view of heterocyclic chemistry, the next stage is of special significance. Oxidation of the aldehyde group in D-glyceraldehyde 3-phosphate to a carboxylic acid is followed by phosphorylation with phosphoric acid to form 1,3-bisphosphoglycerate (Figure 5.6a). The oxidation is mediated by the enzyme glyceraldehyde 3-phosphate dehydrogenase with NAD^+ as the coenzyme. A thiol (SH) group linked to the protein chain exercises an important catalytic function in the action of this enzyme. In the enzyme–substrate complex, the thiol group is believed to attach itself to the C═O group of the aldehyde to produce a hemithioacetal structure, which is able to donate a hydride ion. The hydride ion is accepted by the adjacent coenzyme, NAD^+, which is thus transformed into the corresponding reduced form NAD-H. The aldehyde group is simultaneously converted to the *S*-acylated form, that is, the thiol ester. The C—S bond in such thiol esters increases their reactivity toward nucleophiles. In the given case this bond is cleaved under the action of a very weak nucleophilic agent, inorganic phosphate, which thus gives a phosphorylated carboxy group in the product (Figure 5.6b).

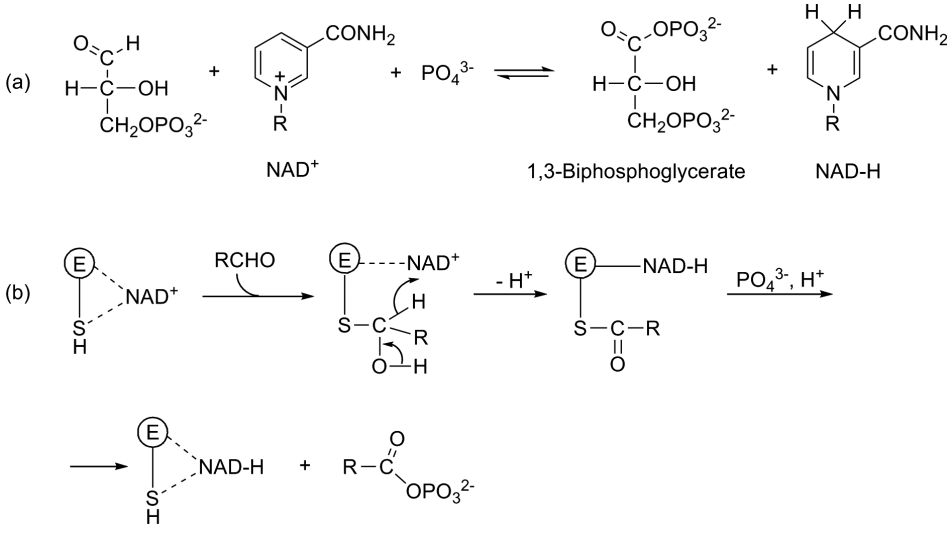


Figure 5.6 *D-Glyceraldehyde 3-phosphate conversion into 1,3-bisphosphoglycerate: (a) stoichiometry of the reaction, (b) proposed mechanism.*

The energetics of 1,3-bisphosphoglycerate are of interest. Phosphorylation of a carboxylic group requires a great deal of energy ($\Delta G^o = 11.8$ kcal mol^{-1}), and it is not immediately clear how such a reaction might occur. Again we encounter here conjugative reactions which possess a common intermediate. In the preceding stage of aldehyde oxidation, the quantity of energy ($\Delta G^o = -10.3$ kcal mol^{-1}) made available is typical of that required for the majority of oxidative processes.

In the next stage of glycolysis, 1,3-bisphosphoglycerate phosphorylates adenosine diphosphate (ADP) to form ATP and 3-phosphoglycerate (Figure 5.7a). The direction of this step is determined by the greater free energy of hydrolysis of a phosphate bond in 1,3-bisphosphoglycerate than in ATP (see Table 5.1). Some free energy is evolved in the reaction owing to the difference in the ΔG^o values of the two compounds, and this enables completion of the previous stage of 1,3-bisphosphoglycerate formation.

The 3-phosphoglycerate produced is isomerized by 3-phosphoglycerate mutase to 2-phosphoglycerate, which is then dehydrogenated to give phosphoenolpyruvate (Figure 5.7b).

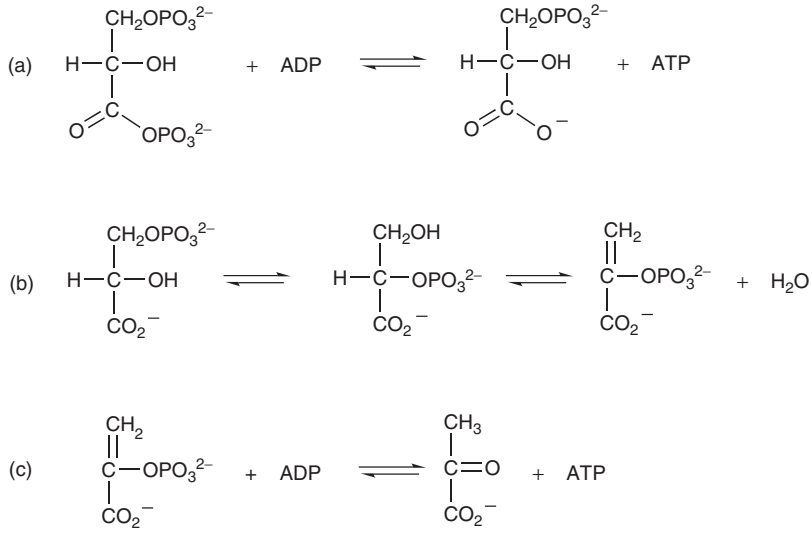


Figure 5.7 *Formation of ATP: (a) 1,3-bisphosphoglycerate conversion into 3-phosphoglycerate and ATP, (b) isomerization and dehydration of 3-phosphoglycerate, (c) transformation of phosphoenolpyruvate into pyruvate and ATP.*

The function of phosphoenolpyruvate is (as discussed in Section 5.1) to phosphorylate ADP to produce ATP and the anion of pyruvic acid (Figure 5.7c).

Though pyruvate formation is not the final step of glycolysis, the overall energetics can now be calculated. The arithmetic of the energy balance is straightforward. Two ATP molecules are lost in the initial stage (in the phosphorylation of glucose and D-fructose 6-phosphate). Bearing in mind that the cleavage of one glucose molecule releases two molecules of 1,3-bisphosphoglycerate, four new ATP molecules are formed in the final stage. Thus, glycolysis results in the liberation of about 18 kcal mol^{-1}, this energy being stored in two molecules of ATP.

What happens next to the pyruvate, varies according to the type of living organism and the conditions inside the cell. Under anaerobic conditions (the absence of oxygen) pyruvate is reduced to lactate (the anion of lactic acid) (Figure 5.8a, b) by the NAD-H which was formed during the oxidation of glyceraldehyde 3-phosphate (see also Figure 5.6a). We conclude from the above that the main purpose of the coenzyme pair NAD^+-NAD-H in glycolysis is to transfer two electrons and a hydrogen ion from glyceraldehyde 3-phosphate. The lactate formation is accompanied by an appreciable decrease in free energy ($\Delta G^o = -6.0$ kcal mol^{-1}) which represents a significant shift of the equilibrium toward the product. When a lack of oxygen occurs in the cells during intense muscular activity, lactic acid is also produced.

During the fermentation of alcohol (e.g., by beer yeast), pyruvate is decarboxylated to form acetaldehyde which is subsequently reduced to ethanol (Figure 5.8c, d). The first reaction is already familiar to us owing to the participation of the coenzyme thiamine pyrophosphate (Section 4.2.2). The reduction of acetaldehyde is carried out by another coenzyme, alcohol dehydrogenase, containing the coenzyme NAD-H.[4]

[4] In humans and other animals, alcohol dehydrogenase fulfills the opposite role to dehydrogenate ethanol to acetaldehyde. The human organism obtains ethanol not only from beer and wine but also from numerous food products, such as yogurt, fruits preserved in sugar, jams and so on.

(a) $H_3C-\underset{\underset{O}{\|}}{C}-COO^- + NAD\text{-}H + H^+ \rightleftharpoons H_3C-\underset{\underset{OH}{|}}{CH}-COO^- + NAD^+$

(b) $C_6H_{12}O_6 + 2\,PO_4^{3-} + 2H^+ + 2\,ADP \rightleftharpoons$

$\rightleftharpoons 2\,CH_3CH(OH)COO^- + 2\,H_2O + 2\,ATP$

(c) $H_3C-\underset{\underset{O}{\|}}{C}-COO^- + H^+ \overset{-CO_2}{\rightleftharpoons} CH_3\text{-}CHO \overset{NAD\text{-}H + H^+}{\rightleftharpoons} CH_3CH_2OH + NAD^+$

(d) $C_6H_{12}O_6 + 2\,PO_4^{3-} + 2H^+ + 2\,ADP \rightleftharpoons$

$\rightleftharpoons 2\,CH_3CH_2OH + 2\,CO_2 + 2\,H_2O + 2\,ATP$

Figure 5.8 *Glycolysis under anaerobic conditions: (a) pyruvate conversion into lactate on fermentation, (b) stoichiometric equation of glycolysis under lactic acid fermentation, (c) pyruvate transformation into ethanol by alcohol fermentation, (d) stoichiometric equation of glycolysis in alcohol fermentation.*

Lactic acid and ethanol are the end products of anaerobic glycolysis; in other words, they are the waste materials of these bioprocesses. Such a route for the supply of energy is extremely inefficient: the complete oxidation of one mole of glucose liberates 686 kcal of energy, but the yield during anaerobic glycolysis is but 6.8% of this total. The greater portion of the energy remains conserved in the chemical bonds of lactic acid and ethanol, which are discarded by the anaerobic organism as waste. Aerobic cells, by contrast, are able fully to extract energy from the cellular fuel in the presence of oxygen. Under aerobic conditions, the process begins with oxidative decarboxylation of pyruvic acid (Section 4.2.2), resulting in acetyl coenzyme A formation:

$$Me-\underset{\underset{O}{\|}}{C}-COO^- + NAD^+ + HS\text{-}CoA \rightleftharpoons Me-\underset{\underset{O}{\|}}{C}-S-CoA + CO_2 + NAD\text{-}H$$

This reaction proceeds with a dramatic decrease in free energy ($\Delta G^o = -8.0$ kcal mol^{-1}), causing a massive shift of the equilibrium to the right. Lipids and proteins, in addition to carbohydrates, also generate acetyl CoA under biological oxidation. Acetyl CoA is a raw material of primary importance for the next stage of the respiratory process, the tricarboxylic acid cycle.

5.2.2 The Krebs Cycle, or the 'Molecular Merry-Go-Round'

The generalized transformations depicted in Figure 5.9 detail the basic chemistry of the Krebs cycle, or tricarboxylic acid cycle. Readers with imagination may recognize it as an analogue of a merry-go-round (carousel). Indeed, as in the well known attraction, some molecules in this biochemical merry-go-round are in a permanent circular gallop to produce compounds and chemical energy at a great speed. Obviously, it is not the mechanical energy of movement, but the energy stored in different chemical reactions. The only organic chemical to continuously enter the cycle is acetic acid, in the form of the acetyl group, with the assistance of coenzyme A. Coenzyme A, in its *S*-acetyl form, transfers the acetyl group to oxaloacetic acid. This operation results in the formation of citric acid (citrate) and the regeneration of the coenzyme. The chemistry of this process, similar to an aldol condensation, is schematically depicted in Figure 5.10.

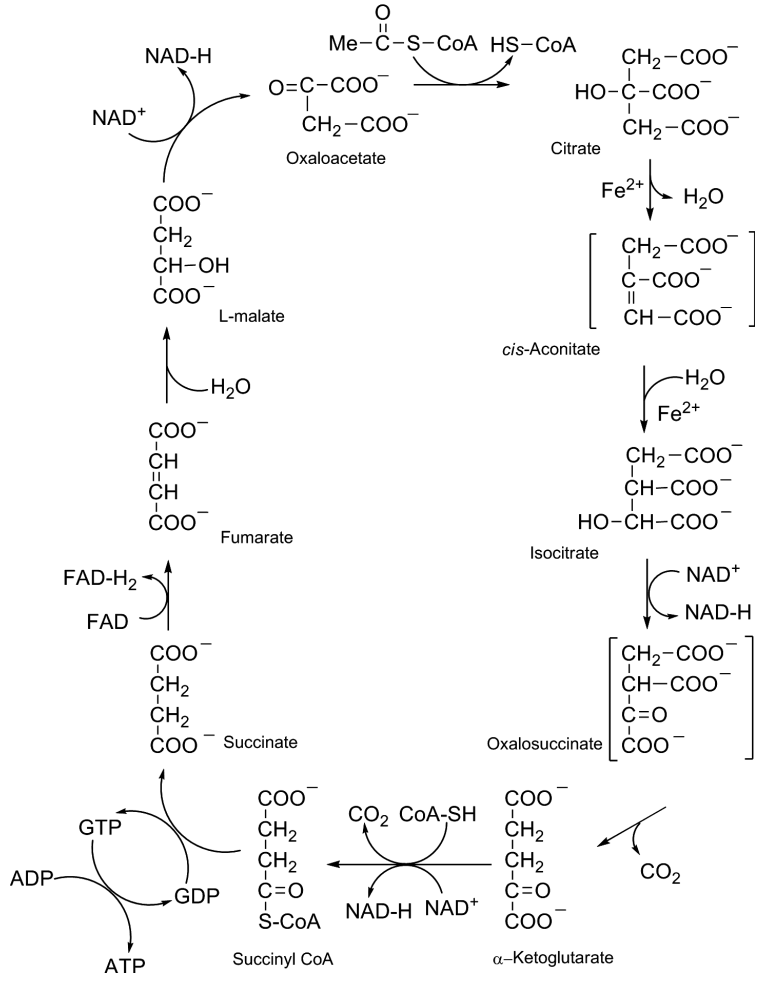


Figure 5.9 *The tricarboxylic acid or Krebs cycle (acids are represented by their anions).*

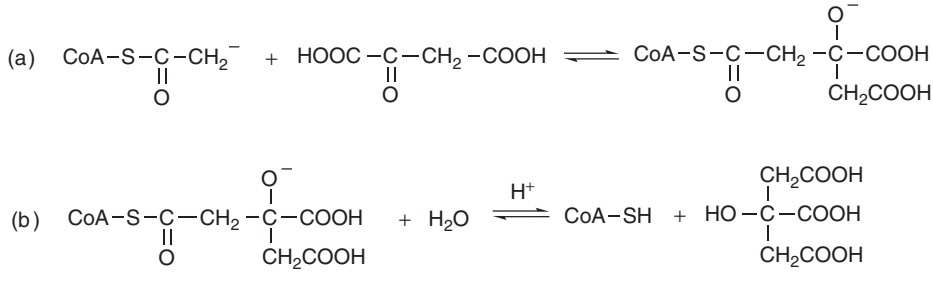


Figure 5.10 *Citric acid (citrate) formation in the Krebs cycle: (a) aldol condensation of acetyl CoA in the anionic form (see also Figure 4.21b) with oxaloacetic acid, (b) citryl CoA hydrolysis.*

According to the Krebs cycle, citric acid isomerizes into the isocitric form (isocitrate). The process is presumed to proceed via a dehydration–hydration sequence involving the intermediate *cis*-aconitic acid. The enzymatic mediator requires the presence of iron(II) ions for operation. Isocitrate is further oxidized and decarboxylated to produce α-ketoglutaric acid. The oxidation is carried out by the NAD-dependent enzyme isocitrate dehydrogenase with the participation of magnesium(II) or manganese(II) ions. Special attention should be paid to the transformation of isocitric acid into α-ketoglutaric acid. First, this is the overall rate-determining step of the Krebs cycle. Second, isocitrate dehydrogenase is inhibited by ATP and NAD-H, but conversely activated by ADP. This situation has the following consequences. When energy expenditures are intense in a cell, the amount of ATP decreases and that of ADP accordingly increases. As a result, the oxidation of isocitric acid and, consequently, the cyclic process as a whole are automatically accelerated. The intensity of electron transfer along the respiratory chain is therefore itself enhanced, causing in turn an increase in the rates both of oxidative phosphorylation and of conversion of ADP into ATP (see also Section 5.2.3). As the cellular stockpile of ATP is regenerated, isocitrate dehydrogenase inhibition begins to occur and the metabolic processes once again slow down to reach their optimal levels.

Figure 5.9 illustrates the two-step conversion of α-ketoglutaric acid into succinic acid (succinate). The first step leads to the formation of succinyl CoA and is, in principle, analogous to the oxidative decarboxylation of pyruvic acid (Section 4.2.2). A well known group of heterocyclic coenzymes, namely thiamine pyrophosphate, CoA and NAD^+, all participate. Succinyl CoA reacts subsequently with inorganic phosphate under the control of guanosine diphosphate (GDP). This pathway is a further example of energy transfer by means of reactions coupled via a common intermediate. Here, the critical intermediate is apparently the mixed anhydride of succinic and phosphoric acids. This anhydride possesses great phosphorylating power and serves as a means of transporting a phosphate group to GDP to form guanosine triphosphate (GTP). The latter phosphorylates ADP, thus connecting the α-ketoglutaric acid conversion into succinic acid with the formation of one ATP molecule (Figure 5.11).

In a subsequent stage of the Krebs cycle, succinic acid is oxidized to give fumaric acid with participation of the enzyme succinate dehydrogenase and the coenzyme flavin adenine dinucleotide (FAD; Figure 4.6). The coenzyme accepts two hydrogen atoms and is thus transformed into the reduced compound $FAD{\cdot}H_2$. Fumaric acid then undergoes hydration, initiated enzymatically by fumarase, to yield the L-stereoisomer of malic acid (L-malate). The final stage in the tricarboxylic acid cycle involves the oxidation of malic acid, monitored by the NAD-dependent enzyme malate dehydrogenase, to oxaloacetic acid. This returns us to the starting point of the overall process. Thus, the merry-go-round has come full circle and is ready to continue in a nonstop fashion.

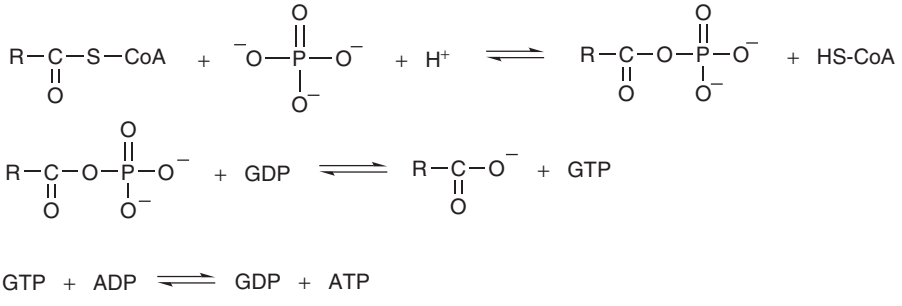


Figure 5.11 *Coupled reactions during the conversion of succinyl CoA to succinic acid ($R = HOOCCH_2CH_2$).*

Two carbon atoms per turnover are exported from the Krebs cycle in the form of CO_2 and two carbon atoms are imported as an acetyl group attached to CoA. The scheme also demonstrates that the carbon atoms entering and leaving the cycle are in different forms. A most important point is that eight hydrogen atoms are liberated as four pairs per full turn. Three pairs are used in the reduction of NAD^+ and one pair in the hydrogenation of FAD. The stoichiometric chemical equation embracing the overall process of the Krebs cycle was described earlier (Figure 5.4). The bioenergetics of the Krebs cycle will be calculated later when we consider the overall energy balance of the respiratory process.

5.2.3 The Respiratory Chain

As we now know, all of the electrons and hydrogen atoms released in glycolysis and in the tricarboxylic acid cycle are retrieved by the coenzymes NAD^+ and FAD which are thus reduced to NAD-H and $FAD{\cdot}H_2$, respectively. The subsequent path taken by these electrons to their terminal destination (i.e., to the oxygen which finally accepts them) is called the respiratory chain. A long series of electron carriers constitutes this chain (Figure 5.12).

The first reaction in the respiratory chain is the oxidation of an NAD-H molecule by a flavin mononucleotide (FMN) unit of the enzyme NAD-H dehydrogenase. The coenzyme $FMN{\cdot}H_2$ formed is then oxidized during the next stage by an iron(III) ion of a protein enzyme. Such iron is called nonheme iron to distinguish it from the heme iron which plays a decisive role in later stages of the respiratory chain. The reduced enzyme with the nonheme iron(II) component subsequently reduces coenzyme Q (CoQ), a *p*-benzoquinone derivative with a long isoprenoid side chain (Figure 5.13), in the presence of protons.

Oxidation of the reduced form of CoQ, and all of the subsequent reactions, proceeds under the control of cytochrome proteins containing heme iron (see Section 4.2.1). It is clear that electrons can be transferred from one carrier to another only when they are in the correct location and relative

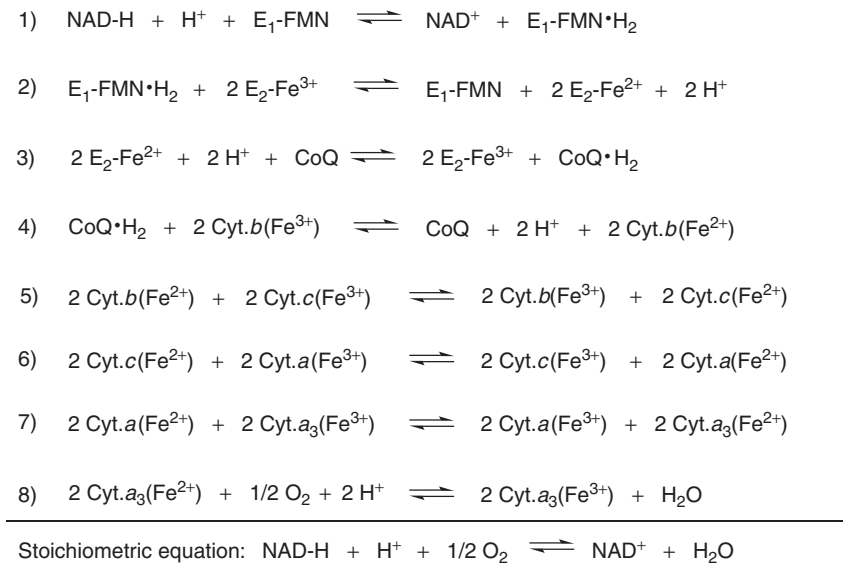


Figure 5.12 *Respiratory chain reactions (by convention, E_1 and E_2 designate the corresponding apoenzymes). Adapted from Lehninger, A. L.,* Biochemistry, *Worth, New York, 1970, Chap. 17, p. 380, with permission.*

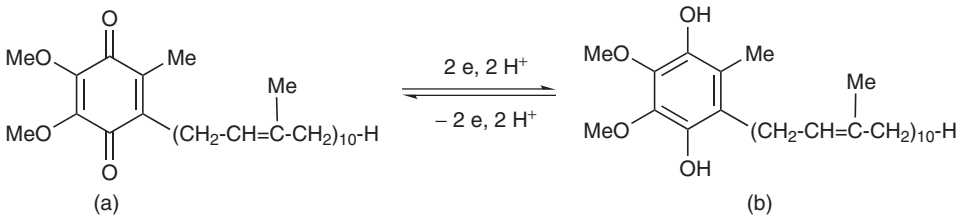


Figure 5.13 *Coenzyme Q: (a) oxidized form, (b) reduced form.*

orientation. Attempts to gain insight into the organization of the respiratory chain have indeed shown that the different components are very close to one another and form respiratory ensembles integrated into the lipid matrix, located on the inner side of the mitochondrial membrane. The components are therefore difficult to isolate and investigate. Cytochromes *a* and a_3, for instance, are aggregated in one complex called cytochrome oxidase which transports the electrons directly to oxygen. The migration of electrons along the respiratory chain is accompanied by a loss of energy, and at the lowest energy level the electrons combine with oxygen. The energy evolved in this process might be compared with that of river rapids or waterfalls. The energy of electrons proceeding to a lower level is determined by the energy gap between the initial and final levels, similar to the dependence of the energy of water flow on the difference in height between the higher and lower levels. In the case of oxidation-reduction processes the energy liberated may be calculated from the redox potentials of the substances engaged in the electron transfer (Table 5.2).

Each reductant on loss of its electrons becomes an oxidant. Similarly, an oxidant gains electrons and is thus transformed into a reductant. Together they form a coupled redox pair. The more negative the value of the redox potential (E'_0), the stronger the reductant and, correspondingly, the weaker the oxidant. On the contrary, the more positive the potential, the stronger the oxidant and the weaker the reductant. The strongest reductant listed in Table 5.2 is acetaldehyde, and the strongest oxidant is oxygen. Electron transfer in the respiratory chain occurs between agents from top to bottom of Table 5.2 in strict conformity with their redox potentials.

The free energy change for the interaction of two redox pairs can be calculated from:

$$\Delta G^{o\prime} = -nF\,\Delta E'_0 \tag{5.1}$$

Table 5.2 *Standard redox potentials (E'_0) for some pairs of compounds (two-electron transfer at pH 7.0 and 25–37 °C; adapted from Lehninger, A. L.,* Biochemistry, *Worth, New York, 1970, Chap. 17, p. 366, with permission)*

Reductant	Oxidant	E'_0 (V)
Acetaldehyde	Acetate	−0.60
H_2	$2H^+$	−0.42
Isocitrate	α-Ketoglutarate + CO_2	−0.38
NAD-H	$NAD^+ + H^+$	−0.32
$FAD{\cdot}H_2$[a]	FAD[a]	−0.11
Coenzyme Q (reduced form)	Coenzyme Q (oxidized form)	−0.05
Cytochrome *b* (Fe^{2+})	Cytochrome *b* (Fe^{3+})	0.00
Cytochrome *c* (Fe^{2+})	Cytochrome *c* (Fe^{3+})	0.26
Cytochromes *a* and a_3 (Fe^{2+})	Cytochromes *a* and a_3 (Fe^{3+})	0.285
H_2O	$^1/_2\ O_2$	0.82

[a]In association with NAD-H dehydrogenase.

where n is the number of electrons transferred, F is Faraday's constant (23.06 kcal) and $\Delta E'_0$ is the difference between the redox potentials of the electron donor and acceptor. As an example, for the redox pair NAD-H–NAD^+ and H_2O–$\frac{1}{2}O_2$, E'_0 is equal to 1.14 V. Hence, the free energy $\Delta G'_0$ is lowered by -52.6 kcal mol^{-1} in the transfer of two electrons from NAD-H to oxygen. Theoretically, this is enough to form approximately six molecules of ATP. However, experimentally the oxidation of each mole of NAD-H was found to produce three moles of ATP, and only two moles of ATP in the case of FAD·H_2. Since the energy of a phosphate bond in ATP is assumed to be -9.0 kcal mol^{-1}, the efficiency of the energy conservation in NAD-H oxidation seems to be around 50%.

In a similar way, the energetics of each separate reaction can be estimated so that the probable stage of synthesis of ATP in the respiratory chain might be elucidated. The likelihood of synthesis can be estimated at three sites, namely the locations of the following transformations:

$$1)\ \text{NAD-H} \rightleftarrows \text{FAD-H}_2 \qquad (\Delta G^{o\prime} = -9.7\ \text{kcal mol}^{-1})$$

$$2)\ \text{cytochrome } b \rightleftarrows \text{cytochrome } c \qquad (\Delta G^{o\prime} = -12.0\ \text{kcal mol}^{-1})$$

$$3)\ \text{cytochrome } (a + a_3) \rightleftarrows \text{oxygen} \qquad (\Delta G^{o\prime} = -24.7\ \text{kcal mol}^{-1})$$

The energy liberated is sufficient to combine ADP with inorganic phosphate (P_i) to form ATP. The mechanism of energy conservation in the respiratory chain, called oxidative phosphorylation, is very complicated and will not be discussed here.

We now examine the overall energy balance of glucose oxidation in the respiratory process. Calculations can be made on the basis of the formation of three moles of ATP from each mole of NAD-H oxidized by oxygen:

$$\text{NAD-H} + H^+ + 3\text{ADP} + 3P_i + \tfrac{1}{2}O_2 \rightarrow NAD^+ + H_2O + 3\text{ATP}$$

The data in Table 5.3 imply that in glycolysis and in the Krebs cycle the overall yield of NAD-H is 10 molecules and that of FAD·H_2 is two molecules. Their oxidation in the respiratory chain thus

Table 5.3 *Biosynthesis of NAD-H, FAD·H_2 and ATP during the oxidation of one mole of glucose*

Reaction	Number of moles formed		
	NAD-H	FAD·H_2	ATP[a]
Glycolysis			
1,3-Bisphosphoglycerate formation from D-glyceraldehyde 3-phosphate	2	0	0
Phosphoenolpyruvate reaction with ADP	0	0	2
Oxidative decarboxylation of pyruvate	2	0	0
Krebs cycle			
α-Ketoglutarate formation from isocitrate	2	0	0
Succinyl CoA formation from α-ketoglutarate	2	0	0
Succinic acid (succinate) formation from succinyl CoA	0	0	2
Fumaric acid formation from succinate	0	2	0
Oxaloacetic acid (oxaloacetate) formation from L-malate	2	0	0
Respiratory chain			
Oxidation of 10 moles of NAD-H by oxygen	0	0	30
Oxidation of 2 moles of FAD·H_2 by oxygen	0	0	4

[a]Two moles of ATP, produced by the phosphorylation of ADP with the assistance of 1,3-bisphosphoglycerate, are not taken into consideration since they are compensated by the expenditure of 2 moles of ATP at the start of glycolysis.

provides 34 molecules of ATP. Four more molecules of ATP are formed in the phosphorylation of ADP by phosphoenolpyruvate (during glycolysis) and by guanosine triphosphate (in the Krebs cycle). Overall, the oxidation of one mole of glucose gives 38 moles of ATP:

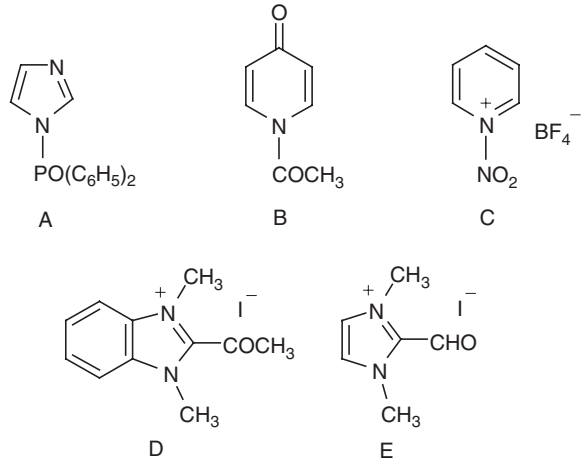
$$C_6H_{12}O_6 + 6O_2 + 38P_i + 38ADP \rightarrow 6CO_2 + 38ATP + 44H_2O$$

Since the free energy of hydrolysis of ATP to ADP is assumed to be – 9 kcal mol^{-1}, we can estimate that the oxidation of one mole of glucose liberates only approximately 342 kcal. Therefore, about half of the energy which could have theoretically been produced by the combustion of one mole of glucose is utilized.

In summary, heterocyclic molecules play an important role in living organisms as the main carriers of electrons and hydrogens during the extraction of energy from cellular fuel. In some cases (coenzymes NAD^+, FAD and their reduced forms) the structures of the heterocyclic nuclei change during the course of the reactions, whereas in others (ATP, cytochromes) the heterocyclic functions play a more passive role.

5.3 Problems

1. What are energy-rich bonds? Indicate which bonds are energy-rich in the following compounds. Provide equations showing the products of hydrolysis.



2. Explain the expression 'coupling of reactions'. Indicate its significance in biochemical processes and give two examples.
3. Aside from ATP, indicate which other compounds are utilized by living organisms in energy transfer. Draw their structures.
4. The total amount of ATP in the human body is about 0.1 mole ($\sim$ 50 g). To maintain energetic balance cells should hydrolyze 100–150 moles of ATP daily. This means that normally each human needs 50–75 kg of ATP each day, which is approximately equal to body weight. Provide a rational explanation for the source from which this amount of ATP can be derived.
5. Discuss the circumstances in which energy is released from the compounds in Problem 3.
6. Name the three main biochemical stages of respiration and give their overall chemical equations.
7. Which of the compounds listed in Table 5.1 are capable of phosphorylating ADP? Which will be phosphorylated by ATP?
8. Calculate the standard free energy change for the following reaction if the equilibrium constant K is 7.2×10^{-9} at 25 °C and pH 7:

$$(CH_3)_2CHOH + NAD^+ \rightleftharpoons (CH_3)_2C{=}O + NAD\text{-}H + H^+$$

9. How many molecules of NAD-H are formed in one turn of the Krebs cycle? How many ATP molecules does this provide for the respiratory chain?
10. Ethanol is not only a poisonous narcotic but also a source of energy. Estimate the amount of energy stored in a human organism after an intake of 100 ml of vodka.

5.4 Suggested Reading

1. Lehninger, A. L., Nelson, D. L. and Cox, M. M., *Principles of Biochemistry*, 4th edn, Worth, New York, 2004.
2. Moore, G. R. and Pettigrew, G. W., *Cytochromes* c. *Evolutionary, Structural and Physicochemical Aspects*, Springer, Berlin, 1990.
3. Stryer, L., *Biochemistry*, 6th edn, Freeman, New York, 2006.
4. Nicholls, D. G. and Ferguson, S. J., *Bioenergetics*, 3rd edn, Academic Press, New York, 2002.
5. Racker, E., *A New Look at Mechanisms in Bioenergetics*, Academic Press, New York, 1976.
6. Lynen, F., in *Nobel Lectures: Physiology and Medicine (1963–1970)*, Elsevier, New York, 1973, p. 103.
7. Newsholme, E. A. and Start, C., *Regulation in Metabolism*, John Wiley & Sons, Ltd, London, 1973.
8. Krebs, H. A. The history of the tricarboxylic acid cycle, *Perspect. Biol. Med*., 1970, **14**, 154.

6

Heterocycles and Photosynthesis

Like hymns to the luminous Sun
Breathe the lotus' delightsome
O'er the mirror-like space of the lake.
K. Bal'mont

Photosynthesis is that combination of chemical reactions and physical processes which occurs in the green parts of plants and cyanobacteria, utilizes solar energy and results in the reduction of carbon dioxide to carbohydrates. Water acts as the reducing agent; it gives up its hydrogen atoms and is consequently transformed into oxygen. Oxygen, the byproduct of photosynthesis, has promoted the development of life on Earth. The overall equation of photosynthesis is well known:

$$6CO_2 + 6H_2O + \text{light} \rightarrow C_6H_{12}O_6 + 6O_2$$

Photosynthesis also occurs in algae, phytoplankton and certain bacteria. It is of interest that some bacteria use reductants other than water. For example, green and purple sulfur bacteria reduce carbon dioxide with hydrogen sulfide. Elemental sulfur is a byproduct of this reaction:

$$6CO_2 + 12\ H_2S + \text{light} \rightarrow C_6H_{12}O_6 + 6H_2O + 12S$$

During photosynthesis the energy of sunlight is transformed into energy stored in the chemical bonds of the reduced compounds. The primary products of photosynthesis are sugars. Biosynthesis of fats and proteins can also take place in plants. The process of photosynthesis can be divided into two stages. Reactions in the first stage proceed only in the presence of light. The subsequent conversions in the second stage do not directly require energy from light. The main result of the first stage is the formation of the heterocyclic carriers of energy: (i) ATP and (ii) nicotinamide adenine dinucleotide phosphate (NADP-H) in its reduced form. In the second stage, CO_2 is reduced by NADP-H in association with various enzymes. A subsequent sequence of complex chemical reactions leads to the formation of carbohydrates.

It is evident that, in contrast to animals, 'the lotuses delightsome' of the vegetable kingdom breathe during the day not with oxygen but with carbon dioxide. From a chemical point of view, photosynthesis is the reverse of the process of breathing. Such interdependence maintains a

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

biological equilibrium and recycles carbon in nature (Figure 6.1). Plants (as phototrophs) synthesize reduced organic substances and oxygen utilizing the sun's energy, with water and carbon dioxide as their raw materials. Animals (as heterotrophs) consume plants and oxygen as raw materials to obtain energy, metabolizing them into water and carbon dioxide.

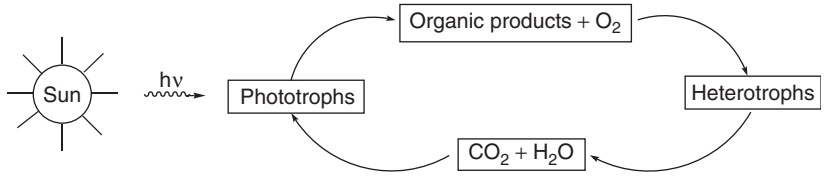


Figure 6.1 *The carbon cycle in the biosphere (Reprinted with permission from Bering, C. L.,* J. Chem. Educ.*, 1985,* ***62****, 659. © 1985 American Chemical Society).*

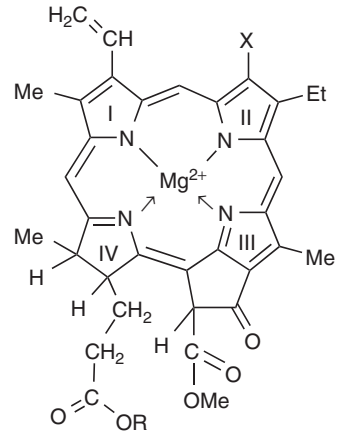
6.1 Chlorophyll: Sunlight-Receiving Antenna and Energy Carrier

Chlorophyll is the green pigment of plants which is the basis of the photosynthetic complex. Like heme in the hemoglobin of blood, the structure of chlorophyll has a porphyrin core which contains a coordinated magnesium(II) ion in its central cavity instead of iron(II) as in heme. A peculiar feature of the chlorophyll structure is the presence of a long hydrocarbon side chain formed by the alcohol phytol. It is by means of this phytyl group that the chlorophyll molecule associates with the nonpolar lipids of cellular membranes through hydrophobic interactions. Various modifications of chlorophyll are known (Figure 6.2). The most abundant is chlorophyll *a*. Chlorophyll *b* differs from chlorophyll *a* by the presence of an aldehyde group instead of a methyl group in the pyrrole ring II. The chlorophyll found in photosynthetic bacteria has a somewhat different structure and is called bacteriochlorophyll; a C—C bond in the pyrrole ring II is now saturated.

The role of chlorophyll is to capture the energy of visible light with the subsequent production of excited state electrons. A second function is to transfer these excited electrons to the corresponding electron acceptors. These functions are performed by different types of chlorophyll. The chlorophyll which transfers an electron directly to the acceptor is called the reaction center. Another type of chlorophyll acts solely to trap light energy. There is only one such reaction center per 400 molecules of light-fixing chlorophyll. 'Antenna chlorophyll', composed of 75% chlorophyll *a* and 25% chlorophyll *b*, are the molecules responsible for light fixation. To understand better the coordination between the functions of the antenna chlorophyll and the reaction center, we consider the physical fundamentals of energy absorption by these molecules.

It is well established that the electrons in conjugated molecules are distributed in discrete energy levels called molecular orbitals (MOs). In the ground state the paired electrons occupy the lowest energy levels called bonding MOs. Absorption of energy allows an electron to move from an occupied bonding MO to one of the unoccupied (nonbonding) orbitals. This migration results in excitation of the molecule. The first excited state is readily reached by transfer of an electron from a high-energy occupied MO to a low-energy unoccupied MO. When sufficient energy is supplied, the migration of electrons to higher orbitals becomes possible. Such energy is found in X-rays and ultraviolet radiation.[1] However, the energy of the radiation reaching the Earth's surface as visible

[1] X-rays, γ-rays and near-ultraviolet radiation ($\lambda <$ 300 nm) are destructive to living cells because they remove electrons from purine and pyrimidine bases, which have long wavelength absorption maxima around 270 nm. Radiation of $\lambda <$ 300 nm destroys nucleic acids and all enzyme systems. Thankfully, most of the damaging ultraviolet solar radiation is absorbed by the ozone layer of our planet. We are shielded from X-rays and γ-rays by the ionosphere and magnetosphere.



(a) X = Me, R = phytyl (see below)

(b) X = CHO, R = phytyl

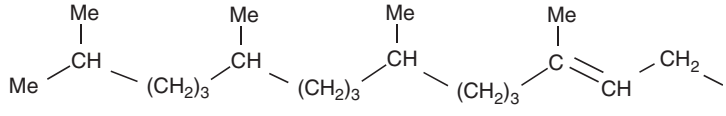


Figure 6.2 *Chlorophyll structures: (a) chlorophyll a, (b) chlorophyll b.*

light is considerably lower. Such radiation induces electron migration to the first excited state only in those molecules with rather narrow energy gaps between the highest occupied MO (HOMO) and the lowest unoccupied MO (LUMO). Small HOMO–LUMO gaps are characteristic of substances with an extensive system of conjugated bonds, which therefore have many π-electrons. Practically all organic pigments, including chlorophyll, belong to this category. The absorption spectrum of chlorophyll in the visible range (Figure 6.3) helps us to understand the origin of its green color. Chlorophyll *a* has two intense absorption bands with maxima at 450 and 660 nm. Therefore, this compound readily absorbs blue and red light but is transparent to green light (no absorption is observed in the 450–650 nm range).

Since chlorophyll absorbs only a narrow range of visible light, its effectiveness as the antenna pigment could be questionable. However, three factors increase its efficiency. First, the antenna complexes of green plants have additional pigments which readily absorb light in the very range where the chlorophyll molecules are transparent. These compounds are the carotenoids and phycobilins (Figure 6.4). The former are isoprenoid hydrocarbons containing long chains of conjugated bonds, whereas the latter are linear tetrapyrrole analogues of chlorophyll devoid of the Mg^{2+} ion. The carotenoids absorb light below 550 nm, and the phycobilins absorb in the range 570–650 nm. Incidentally, the beautiful yellow and orange colors of autumn leaves are largely due to the carotenoids, which are unmasked after the chlorophyll is destroyed.

Second, chlorophyll molecules in a cell are closely surrounded by proteins, lipids and pigments. These pigments generally extend the range of absorption by the chlorophyll molecule. For example, purified bacteriochlorophyll *a* has a long wavelength absorption maximum at 770 nm,

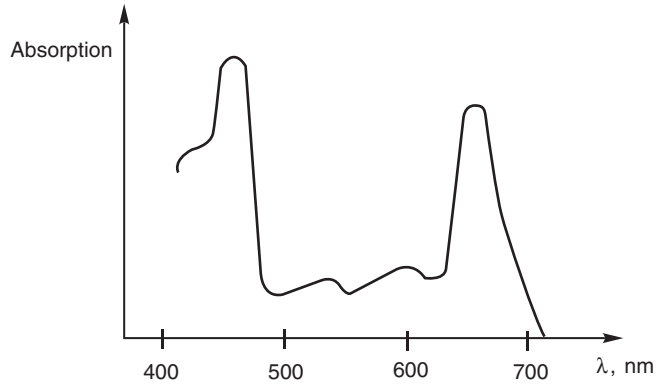
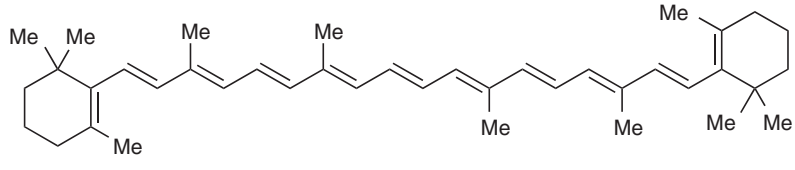
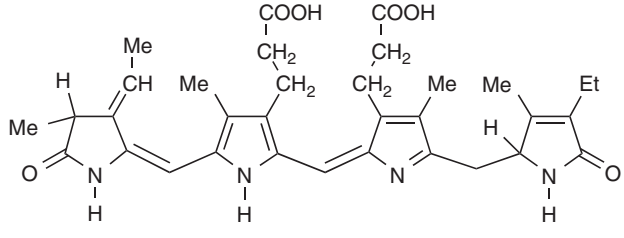


Figure 6.3 *Visible spectrum of chlorophyll a (in diethyl ether solution; Reprinted with permission from Bering, C. L.,* J. Chem. Educ.*, 1985,* ***62****, 659. © 1985 American Chemical Society).*



β–Carotene

(a)



Phycocyanin

(b)

Figure 6.4 *Structures of (a) a typical carotenoid: β-carotene and (b) a typical phycobilin: phycocyanin.*

while the same compound absorbs anywhere between 800–890 nm when involved in noncovalent interactions with various cellular components.

The third reason for the increased efficiency of chlorophyll as an antenna pigment is connected with the special properties of the molecules which distinguish them from a mere assembly of atoms. An electron transfer from one energy level (orbital) of an atom to another level takes place only with the absorption of a strictly determined quantum of energy which must correspond to the energy gap between the levels. In cases where the energy of the photon is greater or smaller, the photon is simply not absorbed. An atom in the excited state has a very short lifetime (about 10^{-12} s), after which it returns to the ground state and a quantum of energy of the same frequency

is released (Figure 6.5a, b). By contrast, the restrictions on the quantum energy are not so rigorous when a molecule is excited. Each electronic state of a molecule has numerous substates related to its vibrational and rotational motions. The laws of quantum mechanics allow the transfer between many sublevels of the ground and excited states. Therefore, light is absorbed by molecules not at a single wavelength only but over a certain range (Figure 6.5c).

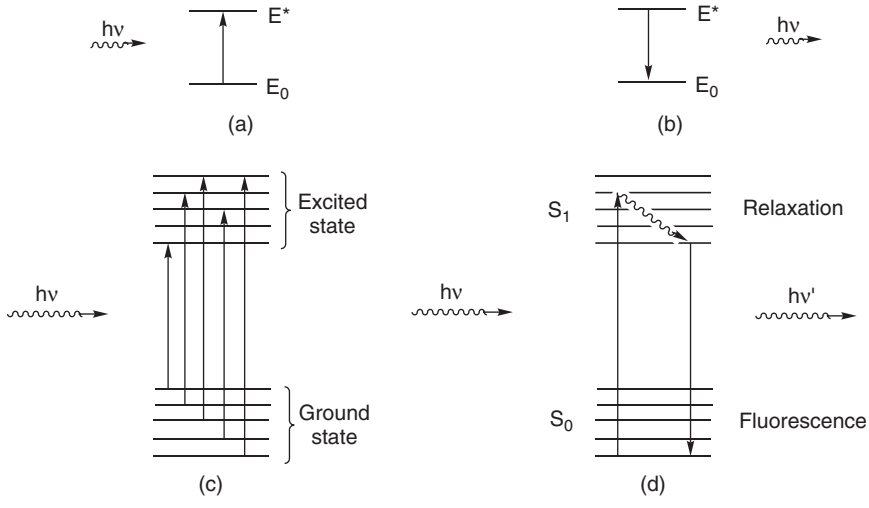


Figure 6.5 *Absorption and emission of light by an atom (a, b) and by a molecule (c, d).*

Deactivation of excited molecules has its own peculiarities. As a rule, light absorption drives an electron to one of the central excited substates. During its short lifetime (10^{-12} s) the excited molecule succeeds in transferring part of its energy to vibration or rotation or to the environment as a result of thermal equilibration. Thus, the molecule is transformed into one of the lower excited substates by a process called relaxation. When the electron returns from this lower excited state to the ground state, the molecule will emit light of a lower frequency (i.e., with a longer wavelength than that of the photon originally absorbed to excite the molecule). Such emission is known as fluorescence (Figure 6.5d).[2] Fluorescence is an undesirable phenomenon in photosynthesis because of the superfluous loss of absorbed energy. Evidently, it is not without reason that chemical evolution has selected chlorophyll as the main pigment and energy-carrying system in photosynthesis. In contrast with other pigments, such as heme, chlorophyll displays extremely feeble fluorescence.

The second method of deactivation from the photoexcited state involves energy transfer to another molecule. The excited molecule E* transfers its energy to a second molecule M, resulting in deactivation of the former (E) and excitation of the latter (M*). Such transfer does not mean that E* emits energy initially and M then captures it. The process occurs in a more specific way, resembling the mechanical resonance of two linked pendulums in which the energy of motion is directly transferred from one pendulum to the other. In the case of molecules, similar resonance energy transfer is possible only when the absorption of one molecule overlaps that of the other. This requirement is met in the case of different pigments such as carotenoids and chlorophyll and also in the case of two identical molecules of chlorophyll. In association with other pigments,

[2] Additional information on fluorescence is given in Section 9.3.1.

antenna chlorophyll absorbs quanta of incident light energy and transfers them from one molecule to another until they reach the reaction site. The energy transfer is fast and efficient only when there is very tight packing of the chlorophyll molecules.[3] The antenna chlorophyll molecules form gigantic supramolecular associations (see Chapter 10 and Section 11.1.3), and delocalization of an excited electron between the fragments results in the transfer of the energy from light excitation.

The third path of deactivation results from the ability of the molecule, once it is excited, to undergo a photochemical reaction, for example, a redox conversion. Thus, if a suitable acceptor A is in the vicinity of the excited molecule E*, an electron from the latter may migrate to the acceptor to give anion A^- and cation E^+. As E^+ has now become an oxidant, it is capable of accepting an electron from a donor D. This possibility may lead to the formation of the triad D^+EA^- with separated charges:

$$DEA \xrightarrow{h\nu} D\overset{*}{E}A \longrightarrow DE^+A^- \longrightarrow D^+EA^-$$

The molecule E itself remains apparently unchanged. Photochemists say that this substance serves as an energy carrier or a photosensitizer, making the reaction sensitive to light. Chlorophyll functions in plants in the same manner: water molecules are the initial donors of electrons and CO_2 molecules are the final acceptors. The intermediate processes are discussed in detail in the following section.

6.2 What Daylight can Achieve

At the end of the 1950s it was established that green plants and algae possess two different reaction sites where electron transfers occur from photoexcited chlorophyll to acceptor molecules. These reaction centers differ from each other in their light absorption characteristics. One center is designated as P680 since its chlorophyll absorbs at 680 nm, the other is P700 with an absorption maximum at 700 nm. The difference in absorption is caused by different environments of the chlorophyll *a* molecules in the two centers, both of which are constructed from protein–pigment complexes incorporated into a lipid matrix.[4]

Each reaction site performs its specific functions in photosynthesis but their actions are coordinated in a very efficient manner. The photosynthetic complex composed of pigment P680, antenna pigments, electron carriers, protein molecules and a number of other components all working in cooperation is called photosystem II. The analogous complex based on pigment P700 is named photosystem I.

The discovery of the two types of photosystems led to the elaboration, in 1960, of the presently accepted scheme of photochemical reactions known as scheme Z because of its resemblance to the letter Z (Figure 6.6). Incident light excites both reaction centers of a plant. Each photoexcited pigment thus becomes capable of reducing the acceptors in its photosystem. Photoexcited pigment P700 in photosystem I transfers its electron to the oxidized form of the iron-containing protein ferredoxin. The electron is then abstracted by a flavin-containing protein and is transferred to coenzyme $NADP^+$, the terminal acceptor during the light-dependent photosynthetic reactions; the $NADP^+$ is thus converted to NADP-H.

[3] The efficiency of the excitational energy transfer is inversely proportional to r^6, r being the distance between E* and M (i.e., if the distance increases by a factor of two, the efficiency is decreased 64-fold).

[4] In plants, the entire photosynthetic complex (which includes two photosystems and the associated enzymes) is located either on the surface of or inside a lipid–protein membrane. These membranes are components of the cellular organelles called chloroplasts.

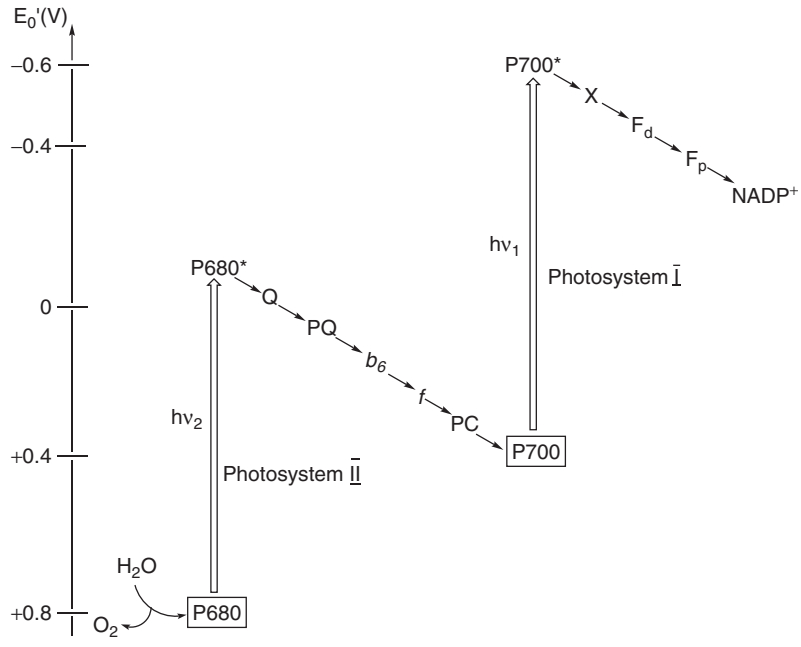


Figure 6.6 *Scheme Z of electron transfer during photosynthetic reactions. Q is the initial acceptor composed of plastoquinone and nonheme iron; PQ is the pool of plastoquinones; b_6 and f are cytochromes; PC is plastocyanin; X is tightly bound ferredoxin; Fd is ferredoxin; and F_p is a flavoprotein (Reprinted with permission from Bering, C. L.,* J. Chem. Educ.*, 1985, **62**, 659. © 1985 American Chemical Society).*

Photosystem II has its own chain of electron transporters. The excited pigment P680 first reduces plastoquinone complexed with nonheme iron. The electron is further transferred to an extensive pool of other unbound plastoquinones which serves as a reservoir of electrons during the oxidation of water. Cytochromes *b* and *f* and the copper-containing protein plastocyanin are the subsequent electron carriers in photosystem II. The oxidized pigment $P700^+$ is located at the end of this chain. Pigment $P700^+$ accepts an electron from photosystem II and reverts to the original state (P700) ready to absorb the next quantum of light.

The oxidized pigment $P680^+$ also obviously has to be converted back into the initial state. Formally, water is the ultimate reductant, but one further set of electron carriers exists between water and $P680^+$. This set includes a redox-active tyrosine amino acid (Tyr_z, Y_{161} in a matrix polypeptide chain), and the so-called oxygen-evolving complex, a cluster containing four Mn and one Ca atom (Mn_4Ca). $P_{680}{}^+$ is rapidly reduced by Tyr_z, which is hydrogen-bonded to a nearby histidine residue (Figure 6.7a). Such a hydrogen bond facilitates oxidation of Tyr_z, due to partial deprotonation of the hydroxyl group. Such catalytic motives with participation of histidine are widely spread in biochemical conversions (see Section 4.1). It is assumed that the Tyr_z radical thus formed is then reduced by Mn_4Ca with subsequent capture of electrons from water molecules (Figure 6.7b). Since to produce one oxygen molecule two water molecules should lose four electrons, the overall process likely proceeds in a step-wise manner; but the details remain unclear. Water is the original and inexhaustible source of all of the electrons involved in photosynthesis; these electrons are then employed in the reduction of plastoquinones, flavins and $NADP^+$ (Figures 4.5, 4.7, 5.13). Oxygen and protons are also liberated simultaneously.

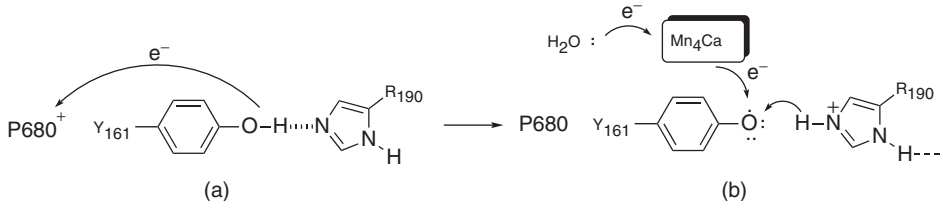


Figure 6.7 *Schematic representation of electron transfer at recovery of pigment P680.*

To understand the necessity for plants to have two discrete photosystems, we examine the energetics of electron transport. We use two formulas in the calculations: Equation (5.1) that we encountered while investigating the dependence of redox potentials on the free energy changes of reactants, and the Planck Equation (6.1) for calculating the energy of photons

$$E = h\nu = hc/\lambda \tag{6.1}$$

where ν is the frequency, λ is the light wavelength, h is the Planck constant (1.58×10^{-34} cal s) and c is the speed of light (3×10^{8} m s^{-1}).

Figure 6.6 shows the redox potentials of the main participants of electron transfer (see also Table 5.2). Water is situated at the bottom of the redox scale ($E'_0 = 0.82$ V) in accordance with the high stability of water molecules and the observation that removing an electron from water is difficult. The fact that the oxidation of water occurs under mild circumstances is a very remarkable feature of photosynthesis. As water is oxidized, the photosynthetic complex of plants must contain an oxidant with a greater positive potential than oxygen: in photosystem II this is the oxidized pigment $P680^+$. The estimated value of E'_0 for the redox couple $P680–P680^+$ falls in the range 0.8–0.9 V. Such a high value is rationalized by the tight packing of pigment P680 in the hydrophobic matrix surrounded by nonaqueous media. Water has no direct contact with $P680^+$ and therefore transfers electrons to this complex via other hydrophobic carriers, such as the manganese-containing protein mentioned above.

We emphasize once again that the photochemical reactions of photosynthesis generally culminate in the reduction of coenzyme $NADP^+$ by water to form NADP-H. The difference between the redox potentials of the two pairs $H_2O–{}^1/_2O_2$ and $NADP^+–$NADP-H ($E'_0 = -0.32$ V) amounts to -1.14 V, thus the free energy increase calculated by Equation (5.1) is approximately 52 kcal mol^{-1}, taking into account that two electrons are involved in the reduction process. A single quantum of light energy with a wavelength of 680 nm is insufficient to meet this substantial energy requirement as we can see by using Equation (6.1):

$$E = \frac{(1.58 \times 10^{-34} \text{ cal s})(3 \times 10^{8} \text{ m s}^{-1})}{680 \times 10^{-9} \text{ m}} = 6.97 \times 10^{-20} \text{ cal}$$

The value obtained is the quantum of energy taken up by one molecule. Since all of the calculations are given for one mole of compound, this value must be multiplied by the Avogadro constant:

$$E = (6.97 \times 10^{-20} \text{ cal}) \times (6.02 \times 10^{23} \text{ mol}^{-1}) = 42.2 \text{ kcal mol}^{-1}$$

This value is well below 52 kcal mol^{-1}, and therefore one quantum of light with a wavelength of 680 nm is not sufficient to reduce $NADP^+$ to NADP-H. The necessary energy can be supplied by two quanta, but the problem is how to incorporate the energy simultaneously since the reaction center of chlorophyll transfers electrons one by one. In other words, the reductive energy of the

photoexcited pigment $P680^+$ is insufficient to transport an electron to the lowest unoccupied MO of $NADP^+$. This is clearly shown in Figure 6.6, where the higher reductive potential of NADP-H can be compared with that of P680* and reduced plastoquinone Q. These considerations demonstrate the need for a second photosystem, of which the reaction center P700 is characterized by a lower energy of excitation and a less positive redox potential ($E'_0 = 0.4–0.5$ V). This potential is not sufficiently high for the oxidized form of pigment P700* to remove an electron from water. The unoxidized excited pigment P700*, however, can reduce the $NADP^+$ molecule, which implies that it is a stronger reductant than P680*. The calculations involving Equations (5.1) and (6.1) indicate that the energy of one quantum of light is more than enough to transfer an electron from P700* to $NADP^+$ (Table 6.1).

Table 6.1 *Energy characteristics of some light processes in photosynthesis*

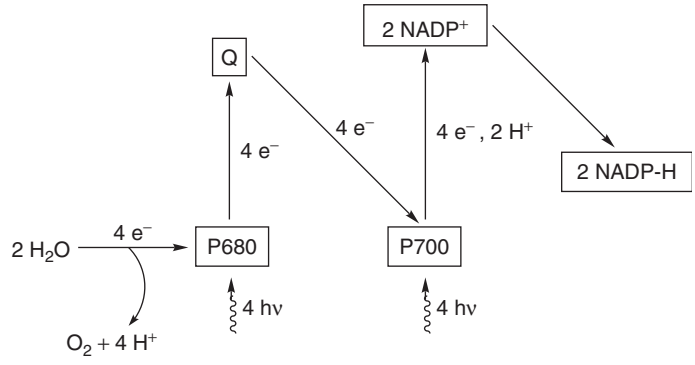
Process	Energy (kcal mol^{-1})
Free energy change during the two-electron reduction of $NADP^+$ by water ($\Delta E'_0 = -1.14$ V)	52.6
Free energy change during the one-electron reduction of plastoquinone ($E'_0 = -0.1$ V) by pigment P680 ($E'_0 = 0.83$ V;[a] $\Delta E'_0 = -0.93$ V)	21.4
Free energy change during the two-electron reduction of $NADP^+$ coenzyme by pigment P700 ($\Delta E'_0 = -0.72$ V)	33.2
Free energy change during the one-electron reduction of bound ferredoxin ($E'_0 = -0.55$ V) by pigment P700 ($E'_0 = 0.4$ V;[a] $\Delta E'_0 = -0.95$ V)	21.9
Quantity of energy in one quantum of light with a wavelength of 680 nm	42.2
Quantity of energy in one quantum of light with a wavelength of 700 nm	40.8

[a]Approximate values.

A general picture of the energetic processes occurring during the photochemical reactions of photosynthesis may be gained from the following considerations. Initially, two independent quanta of light are absorbed by pigments P680 and P700. The excited form P700* carries an electron to coenzyme $NADP^+$. The oxidized form $P700^+$ thus generated is immediately reduced by the excited pigment P680* and is thus available to absorb the next quantum of light. The oxidized form $P680^+$ is, in turn, reduced by water via an electron transfer chain shown in Figure 6.7. The next two quanta of light trapped by both photosystems promote a second electron transfer to $NADP^+$ (to be precise, to radical NADP), resulting in the formation of an NADP-H molecule by the simultaneous addition of a proton.

Thus, pigment P680 is a powerful oxidant in the oxidized form and a relatively weak reductant when photoexcited. On the contrary, pigment P700 is a weak oxidant in the oxidized state but a strong reductant when excited. Since photosynthesis requires the simultaneous availability of both a strong oxidant and a powerful reductant in the same cell, no single pigment can carry out photosynthesis. However, the joint action of both photosystems achieves the required result.

Between eight and ten quanta seem experimentally necessary to form one molecule of oxygen. This is consistent with the following simple considerations. Water is the only source of molecular oxygen in photosynthesis, one molecule of oxygen being derived from two molecules of water. This process involves the release of four protons and four electrons from two molecules of water. Since the oxidized pigment $P680^+$ functions as an electron acceptor, the transfer must involve four steps, and therefore necessitates four quanta of light. Four further quanta will, in parallel, be absorbed by pigment P700 in photosystem I in preparation for these four electrons. So, in total, eight quanta of light should be absorbed. The numbers of electrons and protons evolved are then sufficient for the formation of two molecules of NADP-H. This process is illustrated in Figure 6.8.



$$2\,H_2O + 8\,h\nu \longrightarrow O_2 + 4\,e^- + 4\,H^+$$

$$2\,NADP^+ + 4\,e^- + 2\,H^+ \longrightarrow 2\,NADP\text{-}H$$

Figure 6.8 *Light quanta and electron balance during the formation of one molecule of oxygen in the course of photosynthesis.*

NADP-H is not the only molecule in which the energy of the light quanta is stored. ATP plays a similar role. A mere glance at the Z scheme reveals two sites where a stream of electrons can flow spontaneously like water down a waterfall. Nature does not waste the energy released by this process. However, one of the sites, namely that situated between the reduced form of bound ferredoxin and $NADP^+$, provides a relatively small gain in energy ($\Delta G^o = -5.3$ kcal mol^{-1}) which is almost impossible to utilize. By contrast, the quantity of energy liberated at the second site, located between the reduced plastoquinone and the oxidized pigment $P700^+$, during the passage of even one electron ($\Delta G^o = -11.5$ kcal mol^{-1}) is quite sufficient for the formation of one molecule of ATP.[5] Experimental evidence shows that three molecules of ATP are produced from organic phosphate (P_i) during the transfer of four electrons through the site. The summary equation of the light reactions is:

$$2H_2O + 8h\nu + 2NADP^+ + 3ADP + 3P_i \rightarrow O_2 + 2NADP - H + 3ATP + 2H^+$$

The present discussion would not be complete without mention of the specificity of photochemical reactions in photosynthetic bacteria. The reaction center in such species contains bacteriochlorophyll *a* which absorbs light in the range 800–890 nm (the near-infrared region). Less energy, therefore, is needed to photoexcite bacteriochlorophyll *a* compared with chlorophyll *a*. The energy requirement equals 35.9 kcal mol^{-1} in the case of light with a wavelength of 850 nm. The energy of two such quanta is adequate, in principle, to transfer one electron from water to NAD^+.[6] However, the oxidized form of bacteriochlorophyll *a* has a positive potential ($E'_0 = 0.4$ V) which renders it unable to extract electrons from water. This is why none of the known photosynthetic bacteria can utilize water as a substrate for photosynthesis, and therefore why they do not produce oxygen. Alternatively, more easily oxidized materials such as hydrogen sulfide, alcohols, organic acids and so on are selected as substrates. This fact reflects one more remarkable feature inherent

[5] Discussion of the mechanism of energy conversion from the free electrons to the chemical bonds of ATP is beyond the scope of this book. Interested readers may find further information in the publications listed in Section 6.5.

[6] In contrast to green plants, the coenzyme NAD^+ is the final acceptor of electrons in bacteria.

to photosynthetic bacteria: they possess a single photosystem with electron transport (Figure 6.9) occurring in a cyclic fashion.[7]

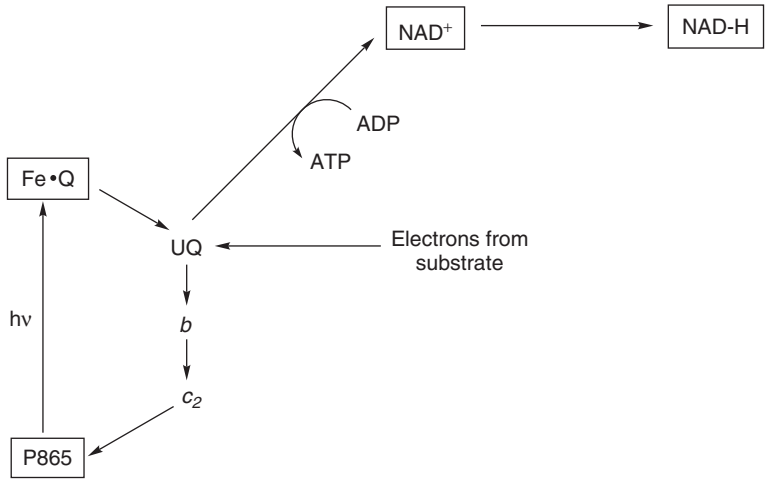


Figure 6.9 *Scheme of electron transport in photosynthetic nonsulfur purple bacteria. UQ is the ubiquinone pool; b and c_2 are cytochromes.*

In these bacteria, an electron is transferred from the photoexcited pigment P865 to the initial acceptor, the ubiquinone[8] complex containing nonheme iron (FeQ). The electron is further passed to the oxidized pigment $P865^+$ via a series of carriers including the ubiquinone pool and cytochromes *b* and c_2. The energy liberated during this process is utilized in ATP synthesis. A portion of the ubiquinone pool is reduced by ethanol or some other substrate. The electrons donated by the substrate are finally accepted by NAD^+ which thus converts to NAD-H. The specificity of this process results from the redox potential of NAD-H being substantially more negative than that of the reduced ubiquinone. The process, consequently, cannot occur spontaneously. The energy required is provided by ATP molecules which are formed in the course of the cyclic transport of electrons.

It should be mentioned that cyclic electron transport also occurs in plants. Photosystem II uses this type of operation as an auxiliary or reserve. When insufficient CO_2 is available excess NADP-H accumulates in plant cells. In such cases the further transfer of electrons to $NADP^+$ is blocked via regulatory mechanisms. The electrons then begin to flow from reduced ferredoxin back to the plastoquinone pool. This process is accompanied by the storage of energy in ATP molecules to be used for biosynthesis.

6.3 Photosynthesis Without Light

The photochemical reactions provide the photosynthetic cells with the energy to carry out their main function, the subsequent reduction of CO_2. This energy, as already mentioned, is stored in

[7] We remind readers that green plants have two photosystems complexed in a linear manner (see Figures 6.6, 6.8).

[8] Ubiquinones are structurally similar to coenzyme Q (Figure 5.13), the main difference being the length of the alkyl side chain. Plastoquinones differ from ubiquinones by the presence of methyl groups in the benzene ring in place of methoxy substituents.

two forms: as the energy of the phosphate bond in ATP and in the form of the 'reductive potential' of NADP-H coenzyme molecules. Thus, carbon dioxide fixation does not need light initiation and may even occur in darkness[9] according to the following general equation:

$$6CO_2 + 12NADP - H + 12H^+ + 18ATP \rightarrow$$
$$C_6H_{12}O_6 + 12NADP^+ + 6H_2O + 18ADP + 18P_i$$

Adding this equation to the stoichiometric equation of the light-dependent reactions (see Section 6.2) and balancing all of the coefficients, we get the simple stoichiometric equation of photosynthesis presented at the beginning of this chapter. In addition, the exact number of light quanta necessary to produce one mole of glucose can now be calculated (48).

The dark reactions of glucose photosynthesis are not of special interest from a heterocyclic viewpoint. These reactions involve a long and complicated chain of enzymatic conversions in which a number of C_3–C_7 sugars and glycols participate. Nevertheless, we do wish to 'shed some light' on the dark reactions to understand the overall picture of photosynthetic chemistry.

The assimilation of a CO_2 molecule commences with its integration into ribulose 1,5-bisphosphate (RUBP; Figure 6.10). The carboxylation is accompanied by cleavage of an intermediate C_6 compound, resulting in the formation of two molecules of 3-phosphoglycerate (PG). This RUBP carboxylase-catalyzed reaction has no simple analogues in classical organic chemistry.

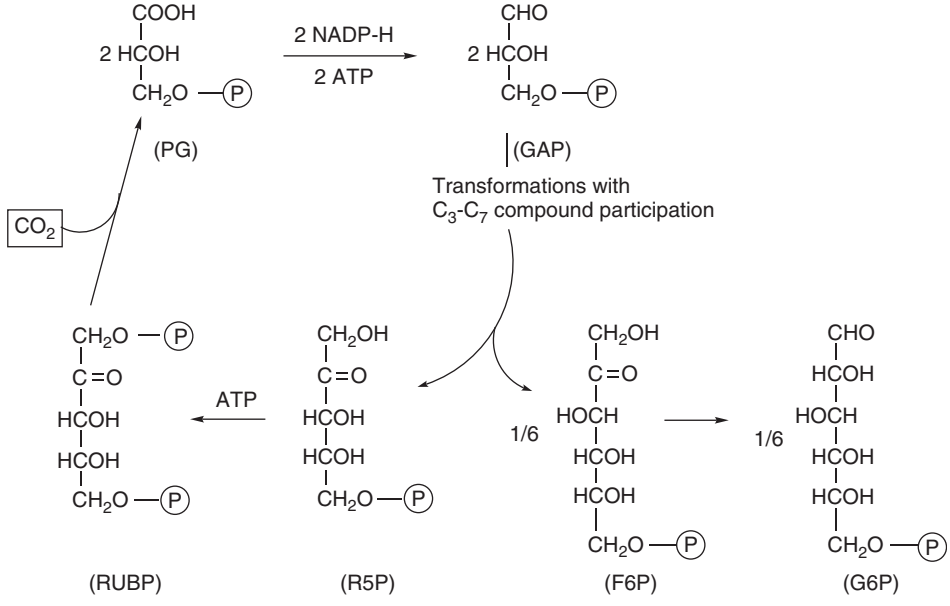


Figure 6.10 *Simplified scheme of CO_2 fixation in the reductive pentose phosphate cycle (Calvin cycle).*

Both molecules of 3-phosphoglycerate are further reduced by the coenzyme NADP-H to give glyceraldehyde 3-phosphate (GAP). As it may be recalled, the same reaction occurs during

[9] As a matter of course, periods of darkness (night) alternate with light (day). Plants normally respire at night. However, a long sojourn in the dark prevents a plant from receiving light energy. In this case it has to switch on a photorespiratory mechanism to support its living processes. As a result, carbohydrates amassed by photosynthesis become reoxidized by air to give carbon dioxide and water. An extended period of darkness eventually brings about plant exhaustion and death.

respiration, but in the opposite direction (Figure 5.6). The reversibility of this reaction allows the equilibrium to be shifted toward aldehyde formation owing to the participation of ATP and other regulatory mechanisms. The GAP molecules are subsequently subjected to complex conversions including a number of isomerizations and condensations. In this transformation pathway, two GAP molecules combine five of their carbon atoms to assemble one molecule of ribulose 5-phosphate (R5P), to which one additional carbon atom is added to yield fructose 6-phosphate (F6P). The fructose derivative is finally transformed into glucose 6-phosphate (G6P). It is important to note (Figure 6.10) that one of the carbon atoms incorporated into glucose is taken out of the transformation cycle, whereas R5P is converted to ribulose 1,5-bisphosphate (RUBP) and is recycled in the CO_2 assimilation scheme. This process, known as the Calvin cycle, was discovered in the 1950s. The assimilation of one molecule of CO_2 thus requires two molecules of NADP-H and three molecules of ATP. In total, the formation of one glucose molecule necessitates the uptake of six molecules of CO_2 in the Calvin cycle.

We have seen that almost identical heterocyclic structures are engaged in the chemical processes of both photosynthesis and respiration: the pyridine and flavin coenzymes, adenosine triphosphate, the cytochromes and a series of other tetrapyrrole compounds. However, the central function in photosynthesis is carried out by chlorophyll, which is not found in animals. In the evolution of plants there seem to be several reasons why nature chose this class of compound as the photocatalyst. First, the porphyrin system is highly aromatic and, hence, very stable chemically and thermodynamically. This is a crucial factor since each molecule of chlorophyll must undergo several thousand electron transfers without degradation. Second, the elongated π-electronic system of chlorophyll is well adapted to absorb light not only in the blue region of the spectrum but in the red and near-infrared regions also. Third, the π-electronic system is very flexible regarding the matrix, making chlorophyll a functionally versatile catalyst. Finally, chlorophyll does not lose absorbed energy by non-irradiative deactivation processes, in contrast to heme and other tetrapyrrole structures.

For several years the intriguing problems of creating artificial photosynthesis and, therefore, chlorophyll-like compounds have been the focus of attention. Section 11.1.3 covers the latest developments in this field.

6.4 Problems

1. List some of the naturally occurring modifications of chlorophyll. How do they differ from each other? What are their functions?
2. Which pigments constitute the antenna complex of green plants? Outline their specific functions.
3. How are the photoexcited molecules of chlorophyll deactivated?
4. Which compound is the terminal acceptor of electrons in the light reactions of photosynthesis? Write an equation detailing its formation. How many light quanta are required to form one molecule of this compound?
5. Why does the photosynthetic apparatus of green plants utilize two reaction centers and two photosystems? Draw the Z scheme of electron transfer and characterize in general terms the principles of its operation.
6. What is the specificity of light reactions in photosynthetic bacteria?
7. Calculate the maximum numbers of ATP molecules which can be synthesized from ADP and inorganic phosphate under standard conditions by absorption of a quantum of light of wavelength: (a) 400, (b) 550 and (c) 700 nm, respectively (assume the existence of the appropriate mechanisms).
8. Determine the standard free energy change that occurs during the transfer of one pair of electrons in photosystem I from bound ferredoxin ($E'_0 = -0.55$ V) to $NADP^+$ ($E'_0 = -0.32$ V).
9. No solar light penetrates 1500 m below see level, nevertheless, even in such completely dark places green sulfur photosynthetic bacteria (GSPB) of ellipsoidal shape have been found. They differ from other antenna complexes by their large size (about $150\times\ 75 \times 25$ nm) and a lack of protein matrix supporting the photosynthetic pigments. The bacteria contain chlorosome structures with up to 250 000 chlorophyl molecules. Suggest answers to the following questions: (a) what can be a possible source of photons at these depths? (b) what is the reason for the above structural peculiarities of GSPB?

6.5 Suggested Reading

1. Lehninger, A. L., Nelson, D. L. and Cox, M. M., *Principles of Biochemistry*, 4th edn, Worth, New York, 2004.
2. Gregory, R. P. F., *Biochemistry of Photosynthesis*, 3rd edn, Wiley, New York, 1989.
3. Clayton, R. K., *Photosynthesis: Physical Mechanisms and Chemical Patterns*, Cambridge University Press, Cambridge, 1980.
4. Blankenship R.E., *Molecular Mechanisms of Photosynthesis*, 2nd edn, John Wiley & Sons, Inc., New York, 2008.

7

Heterocycles and Health

There're four hundred and four disease'.
Some hundred and one of these
by drugs are treated with ease.
Another one hundred and one are cured by doctor and charm.
The same number of illnesses pass harmlessly by themselves.
Yet I'm with anxiety ill that the rest are incurable still.
L. Martynov

7.1 Medicines From a Natural Storehouse

For thousands of years the sick were cured worldwide by remedies obtained from 'nature's own drugstore', comprising such sources as leaves, fruits, barks and herbs. Until recently, it was not realized that successful traditional treatments of this type were frequently triggered by the presence of various heterocyclic compounds in extracts derived from plants, animals and insects.

Perhaps no other naturally occurring compound has received as much attention as quinine, both in the scientific literature and in fiction. As children reading Jules Verne's novel *The Mysterious Island*, we were all concerned for the life of Herbert when he became stricken with malaria. The youth seemed to be doomed, but a miracle occurred. On the table appeared a small box bearing the inscription 'Quinine Sulfate'. The box had been secretly left by Captain Nemo. It is further written in the novel:

> 'It [the box] contained nearly two hundred grains of a white powder, a few particles of which he carried to his lips. The extreme bitterness of the substance precluded all doubt; it was certainly the precious extract of quinine, that preeminent antifebrile.'

Within a few days Herbert was on his way to recovery.

Quinine is a representative of the alkaloids. The alkaloids comprise numerous families of nitrogen-containing organic compounds which occur widely in the plant kingdom. Alkaloids are often considered to be 'waste products' of the vital processes in plants because they are accumulated in the easily detachable parts: the bark, leaves and fruits. Cinchona trees, whose bark

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

contains quinine and approximately 20 other alkaloids, grow in the jungles of South America. The legend tells us that once upon a time an Indian sick with malaria was so thirsty in the jungle that he drank even very bitter water from a stagnant pool in which there was a fallen tree called the quina-quina. His fever abated soon and after some time he was absolutely healthy. This case became known to the natives and then to the spanish missionaries, with whom the tree 'Jesuit bark' traveled to Europe where the tree had received the name 'Cinchona'.

Almost all alkaloids are structurally derived from either aromatic or hydrogenated heterocycles. This relationship has been used as the basis for a classification system which includes, for example, the isoquinoline, pyridine, purine and quinazoline series of alkaloids. Quinine is a rather complex derivative of quinoline (Figure 7.1).

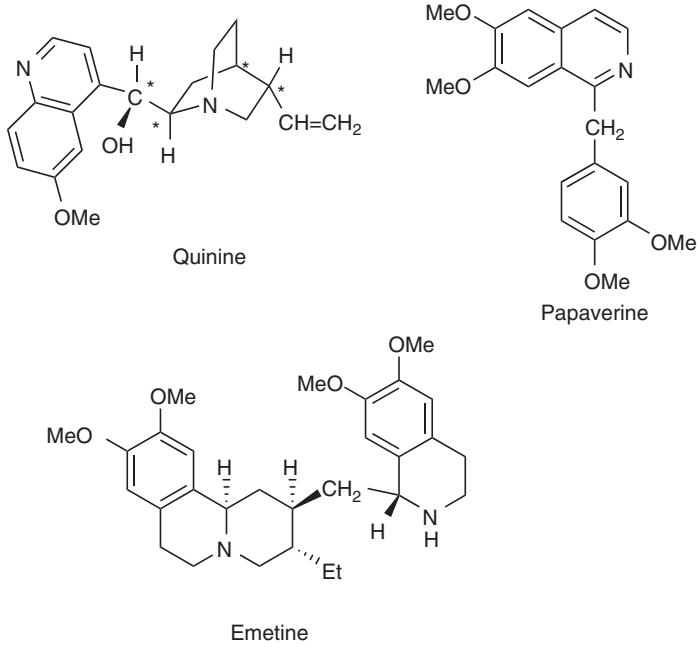
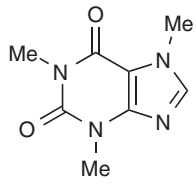


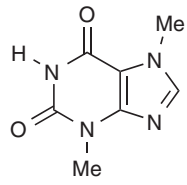
Figure 7.1 *Representative alkaloids of the quinoline and isoquinoline series.*

The quinine molecule has several asymmetric carbon (C*) atoms and can exist as a number of different stereoisomers. A dextrorotatory diastereomer of quinine, called quinidine, is used as a powerful antiarrhythmic agent in the treatment of tachycardia and ciliary arrhythmia. The isoquinoline alkaloid papaverine has found applications in therapy as a spasmolytic and vasodilator. Another isoquinoline alkaloid called emetine, extracted from the roots of the ipecacuanha plant, was for a long time an efficient remedy against amoebic dysentery until resistant strains emerged.

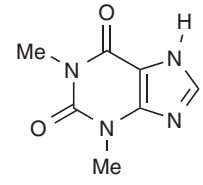
Few of us do not indulge in at least one cup of tea, coffee or cocoa daily. These drinks, having a pleasant taste and aroma, have served from time immemorial as mild tonics. Their effects are derived from caffeine, theobromine and theophylline, alkaloids of the purine group which are present in tea leaves and the beans of coffee and cacao (Figure 7.2). All of these substances are stimulants of the central nervous system (CNS) and can exert a diversity of effects on other organ systems.



Caffeine



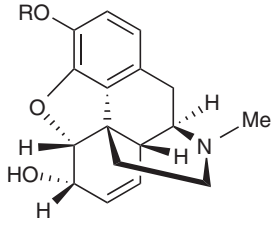
Theobromine



Theophylline

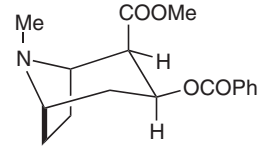
Figure 7.2 *Purine alkaloids.*

Theobromine and theophylline possess vasodilator and diuretic properties. An inherent property of many alkaloids, which makes them medically useful, is their ability to act on the central or peripheral nervous system. For example, the piperidine alkaloid morphine is well known as a pain killer and is sometimes used as an adjunct in cancer chemotherapy (Figure 7.3). These substances suppress the sensitivity of nerve endings, inhibit the conduction of impulses through nerve fibers and induce a weak hypnotic effect. *O*-Methyl morphinate, otherwise known as codeine, is used as an antispasmodic and expectorant in cough treatments. Morphine and codeine are components of opium which is obtained as a milk from unripe somniferous poppy seeds. Cocaine is present in the leaves of the coca shrub which grows in South America, South Eastern Asia and elsewhere.

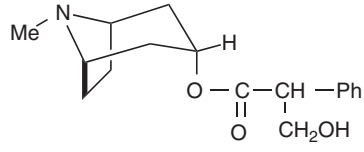


Morphine (R = H)

Codeine (R = CH_3)



Cocaine



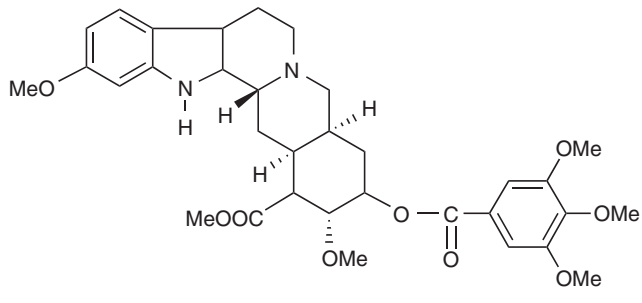
Atropine

Figure 7.3 *Piperidine alkaloids.*

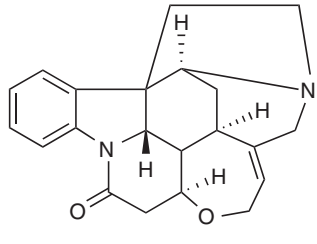
Atropine is another alkaloid of the piperidine series found in belladonna, henbane, Jamestown weed and other plants of the nightshade family and was at one time widely used by doctors in ophthalmic practice for diagnosis and treatment. Atropine is an antispasmodic muscle relaxant. When applied to the eyeball, it induces pupil dilation. The use of atropine as an antidote in cases of intoxication with narcotics, hypnotics or strong toxins such as muscarine and pilocarpine is

of great interest. Atropine is assumed to replace these poisons at the corresponding biological receptor sites, thus eliminating their toxic effect.

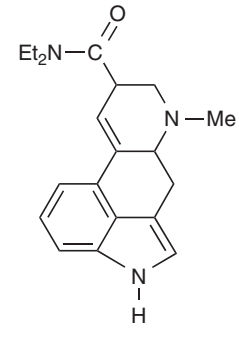
The shrub *Rauvolfia serpentina*, native to South and South Eastern Asia, contains the indole alkaloid reserpine (Figure 7.4) which is used as a tranquilizing agent. Moreover, reserpine decreases arterial blood pressure and is useful in the treatment of hypertension.



Reserpine



Strychnine



LSD

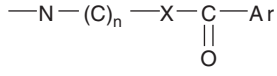
Figure 7.4 *Indole alkaloids.*

We should emphasize that the biological activity of many alkaloids depends dramatically on the dose level, as can occur with all biologically active compounds. Natural substances can be 'positive' or curing ('angelic'), but at a different dose the same compounds can also display a 'negative' or toxic ('evil') face. For example, the alkaloid strychnine in small doses acts as a cardiac stimulant. In larger doses, however, the same strychnine acts as a convulsive poison and can cause respiratory paralysis and rapid death. The powerful narcotic effects of morphine and cocaine are well known. Repeated infusion of these alkaloids causes drug addiction (morphinism, cocainism). The diethylamide of lysergic acid (the indole alkaloid with the abbreviation LSD) is a notorious drug which may induce hallucinations as a side effect. Figuratively, the ancient warning by Homer seems timely:

'Trojans, be cautious of gifts offered you by the Greeks!'

Many attempts have been made to eliminate the unfavorable physiological properties of the above-mentioned alkaloids. An efficient way to achieve this aim seemed to be to test various

substituted derivatives. However, random variation of structure is at best expensive and may lead to the synthesis of even more dangerous compounds, a notorious example of which is synthetic morphine *O,O′*-diacetate, the narcotic heroin. A more effective strategy would entail determining which fragments of the molecular structure are responsible for the useful biological activity. Thus, the analgesic effects of cocaine and a series of other local anesthetics have been shown to be induced by so-called anesthesiophorous groups, such as:



In the case of cocaine X = O and $n = 3$. A number of synthetic local anesthetics have been developed using this fragment as a basis, and the necessity for cocaine application has been reduced. Almost all alkaloids have a three-dimensional structure with an intricately developed substitution pattern. Such molecules generally include asymmetric centers and a variety of functional groups capable of forming hydrogen bonds. Pharmacologists and organic chemists describe many of the synthetic compounds as 'alkaloid-like structures'. This indicates that the substance is likely to display prominent biological activity. The specific molecular structures of alkaloids are assumed to consolidate their attachment to the biological receptors responsible for the appearance of a certain function.

7.2 Heterocycles Versus Infectious Microbes

7.2.1 In Search of 'Magic Bullets'

The twentieth century can be considered as the age of the great drug revolution. Medicinal preparations synthesized in the past 100 years have brought about a decrease in the mortality rate of numerous diseases and provide relief for many ailments. Widespread success was achieved first and foremost with infectious diseases such as pulmonary inflammation, tuberculosis, cholera and various purulent infections. For thousands of years, these afflictions were a scourge of mankind.

The drug revolution was preceded by a number of earlier discoveries. The most important of these was made by Louis Pasteur in the 1860s. He established that the source of infectious diseases was invisible pathogenic microbes. Later, the German scientist Koch elaborated procedures for growing pure bacterial cultures. It thus became possible to study the action of chemical substances on different species of bacteria.[1] For the initiator of chemotherapy, Erlich, the objective was to find a 'magic bullet' among the compounds tested which would be specific for the target pathogenic microorganism and would therefore be absolutely harmless to the patient. This approach is the core of the selective toxicity principle. At present many heterocyclic compounds have been found to possess 'magic bullet'-like properties. At the beginning of the twentieth century antibacterial activity was discovered in some heterocyclic cationic dyes. In particular, acridinium salts such as proflavine and ethacridine (Figure 7.5) were used with great success during World War I as antiseptics for the disinfection of wounds. Another well known antiseptic is the phenothiazine dye methylene blue (Figure 7.5).

The hydrazide of isonicotinic acid (isoniazid) since 1951 played a principal role in the treatment of tuberculosis, and around 1980 gave rise to hope that this disease had been conquered.

[1] Viruses, pathogenic fungi and protozoa (the simplest one-cell animals such as the cholera germ, *Plasmodium malariae*, etc.) can also serve as infectious disease inducing pathogens. Information concerning antiviral treatments can be found in Section 7.3.

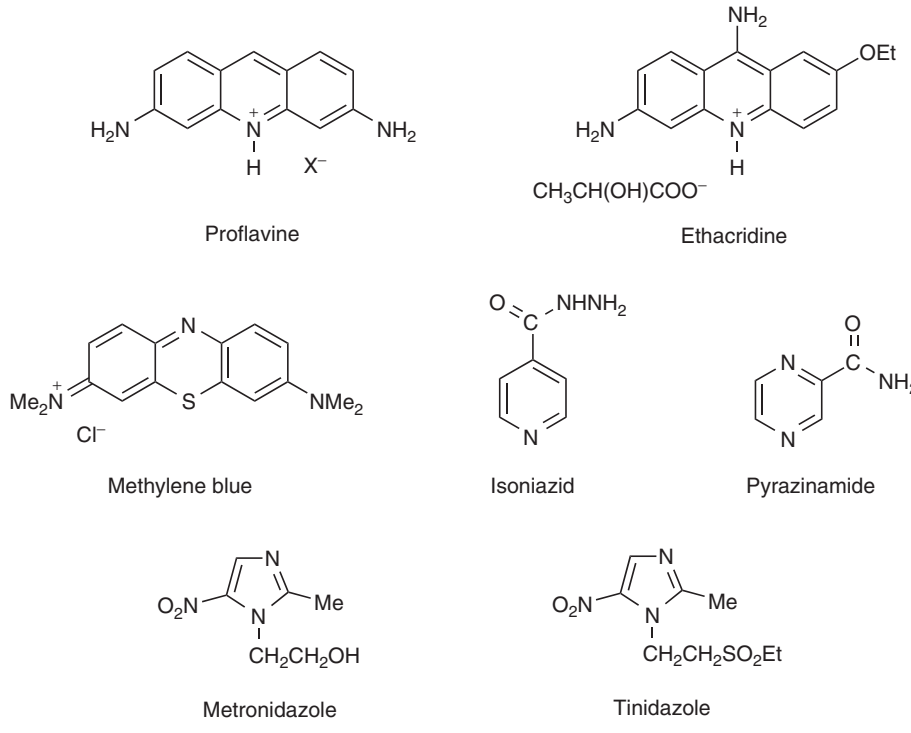


Figure 7.5 *Typical heterocyclic pharmaceuticals used in the treatment of infectious diseases.*

Unfortunately, resistant forms have since emerged. Izoniazid is now used in complex treating tuberculosis together with 2-amidopyrazine (Figure 7.5), but new developments in this area are strongly needed. Just in 2005 tuberculosis killed approximately 1.6 million people worldwide and became the second killer by infections after AIDS. The recent complete decoding of the mycobacterium H37Rv genome is expected to open the ways to the most selective synthesis of antituberculosis remedies.

From the mid-1960s nitroimidazole drugs such as metronidazole and tinidazole became available (Figure 7.5). Such preparations radically changed the treatment of *Trichomonas* infections. These heterocycles proved to be 'magic bullets' because of their high potency in relation to the parasites and surprisingly low toxicity toward humans. It is of interest that metronidazole also assists in the cure of alcoholism.

A true revolution in the struggle against infectious diseases was brought about by two classes of medical preparations: the sulfanilamides and antibiotics. Their discovery and widespread use almost coincided with one of the most dreadful events in human history, the beginning of World War II. As a result of these treatments millions of lives were saved.

7.2.2 Sulfanilamides and Heterocycles

The first commercially available sulfa drug was prontosil, widely known as red sulfanilamide (Figure 7.6a). Its strong bacteriostatic activity was discovered by German scientist G.Domagk who received the 1939 Nobel prize in Medicine. Quite soon it was found that prontosil is actually a prodrug since it is reduced in organism into colorless and also biologically active

4-aminophenylsulfonamide (white sulfanilamide, Figure 7.6b). Due to much less toxicity, white sulfanilamide after the 1950s fully substituted prontosil on the market.

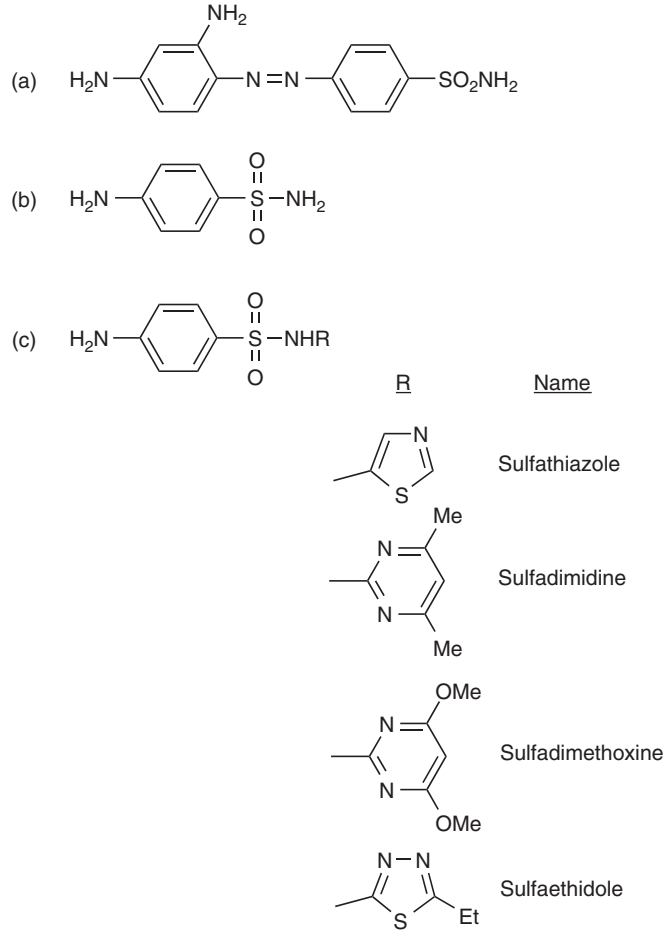


Figure 7.6 *Typical sulfa drugs: (a) red sulfanilamide, (b) white sulfanilamide, (c) heterocyclic derivatives.*

Both red and white sulfanilamides contain no heterocyclic fragments. However, the intensive research work that followed their discovery demonstrated that modification of the *p*-aminobenzenesulfonamide structure by the introduction of heterocyclic substituents into the amide markedly enhanced their biological activity. Several tens of derivatives of this type, including the well known sulfathiazole, sulfadimidine, sulfadimethoxine, sulfaethidole and others, were gradually introduced into clinical treatment (Figure 7.6c).

Sulfa drugs are highly efficient against many bacterial species and against some protozoa. Catarrhal illnesses, gastrointestinal infections, meningitis, scarlet fever, tuberculosis and bubonic plague have been successfully treated by such preparations. Simple changes in the heterocycle substitution pattern enable the formation of drugs with either short-lived or prolonged action. With

the passage of time, however, the increasing evidence of clinical toxicity of these drugs has led to a diminution in their use, and they have been replaced to a great extent by penicillins, cephalosporins and, more recently, quinolone drugs (see Sections 7.2.3 and 7.2.4).

For sulfanilamides, the biological target and mechanism of action are well established. The normal growth and development of bacteria requires *p*-aminobenzoic acid for the synthesis of folic acid, which in turn regulates the production of purines and pyrimidines (see Section 4.2.2). The basic fragment of all sulfanilamide drugs is structurally very similar to *p*-aminobenzoic acid. Therefore, the enzyme which controls the attachment of *p*-aminobenzoic acid to the pteridine ring mistakenly binds *p*-aminobenzenesulfonamide. This brings about an abrupt deceleration of folic acid biosynthesis, and consequently delays the biosynthesis of nucleic acid and bacterial cell proteins. The final result of the process is disastrous for the bacteria.[2] The reason sulfanilamides do not exert a similar negative effect on human cells is that human beings do not produce folic acid, and therefore have no need of enzymes which manipulate its biosynthesis. Man and other animals must obtain vitally important vitamin B9 from their diet.

The role of the heterocyclic radicals in sulfanilamides is not yet precisely known. However, all the heterocyclic sulfa drugs contain a pyridine-like nitrogen and the heterocycles are rather strong electron-accepting moieties. This supports the assumption that the heterocyclic fragment increases the acidity of the sulfamide N—H linkage making it close to that of *p*-aminobenzoic acid. The anion formed by dissociation of the N—H bond (Figure 7.7a) is likely to be delivered to the target infection more quickly owing to its increased solubility in blood compared with the neutral molecule. Moreover, the anion might be more readily attached to the active site of the enzyme. There may, of course, be more than one explanation. The presence of a heterocyclic substituent, together with the enhanced NH acidity, favors the conversion of sulfanilamide into its tautomeric imino form (Figure 7.7b). This tautomer is stabilized by intramolecular hydrogen bonding and may succeed in binding to and thus inhibiting the enzyme.

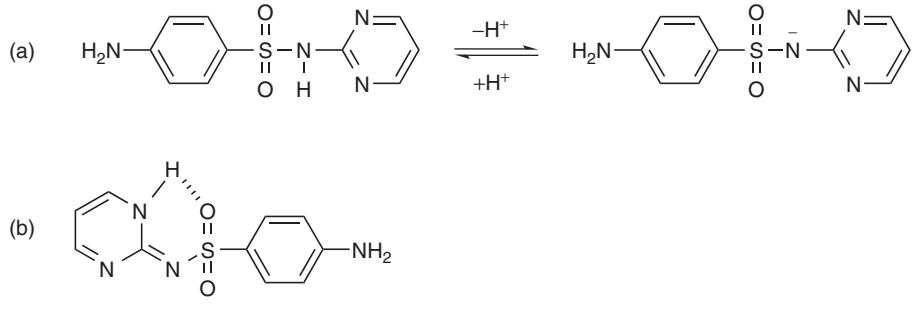


Figure 7.7 *(a) Acidic ionization of sulfadimidine and (b) tautomeric form of sulfadimidine.*

7.2.3 Antibiotics

In a broad sense, antibiotic is any chemical substance that kills bacteria or inhibits their growth. However, this definition may seem too diffuse since gastric juice, acetic acid, hydrogen peroxide and many other common substances also possess antimicrobial properties. A more accurate definition was offered by Waksman who considered antibiotics as naturally occuring compounds

[2] The creation of medicines which act as antagonists of the natural substrates (metabolites) of specific enzymes is the essence of the antimetabolite concept. We deal later with further applications of this concept.

produced by certain species of microorganisms supposedly as 'chemical weapons' in the battle against hostile microbes. Traditionally, diverse synthetic derivatives of natural antibiotics are also included in this group.

Many original antibiotics were isolated from soil samples which are laden with actinomycetes, a special kind of bacteria. Actinomycetes produce numerous secondary metabolites, including antibiotics, as a result of severe intraspecific struggle. Man has managed to 'domesticate' some species of microorganisms and use their antibiotics for protection against pathogenic bacteria, fungi and even some viruses.

The history of antibiotics began in 1928 when the Scottish scientist Fleming discovered that a *Staphylococcus* culture, contaminated with green mold, had been killed. In 1940 Florey and Chain separated the active agent from the mold of *Penicillium notatum* and named it penicillin. Three year later, the use of penicillin had spread far and wide, scoring resounding triumphs over the omnipresent microenemies of the human race. Interestingly, the structure of this efficient and low toxicity antibacterial drug, named penicillin G (Figure 7.8), was elucidated only in 1945 with the assistance of X-ray structural analysis. At the same year A. Fleming along with H. W. Florey and E. B. Chain were honored with the Nobel prize in Medicine.

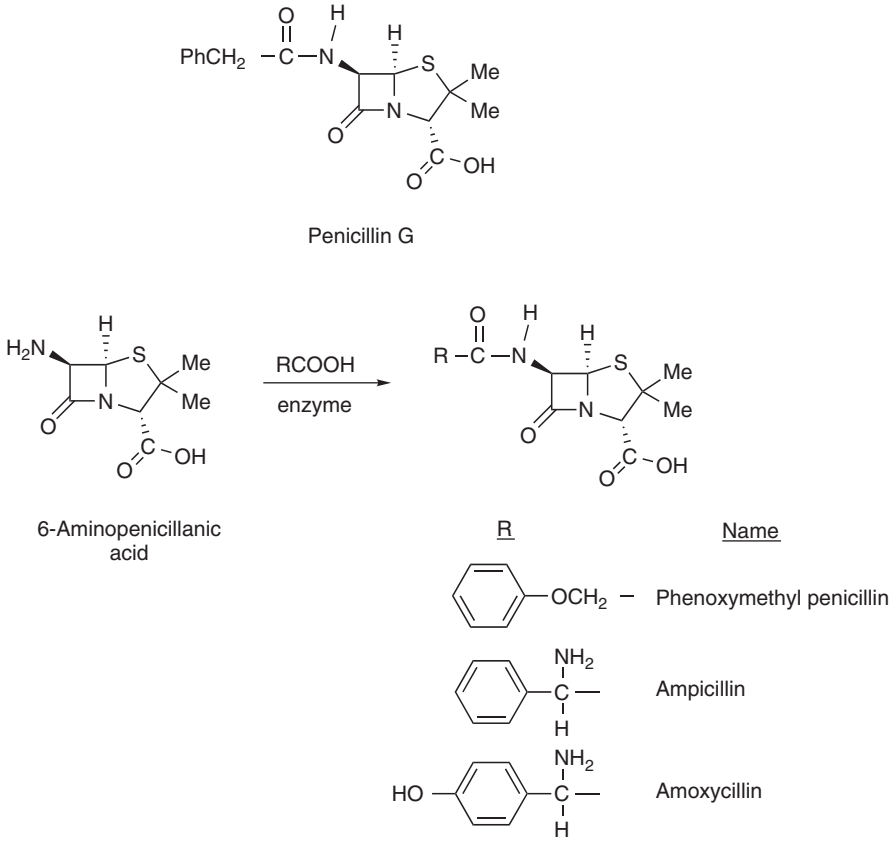


Figure 7.8 *Penicillin G and fermentative preparations of diverse penicillin antibiotics.*

The penicillin core responsible for antibiotic activity is called 6-aminopenicillanic acid. It represents a condensed heterocyclic system composed of a five-membered thiazolidine ring and a four-membered azetidine nucleus in the form of a β-lactam, the single cyclic nitrogen atom being common to both rings (Figure 7.8).

The structure of penicillin was surprising to many chemists because of the well known instability of β-lactam rings, especially toward hydrolysis. Therefore, the natural occurrence of similar structures seemed unlikely at the time. The β-lactam fragment is indeed a critical feature of the penicillin molecule, responsible for both its frailty and for its bioactivity. Once the heterocycle is ring opened (e.g., by acidic hydrolysis), the antibacterial activity is destroyed because the penicilloic acid formed is not active (Figure 7.9).

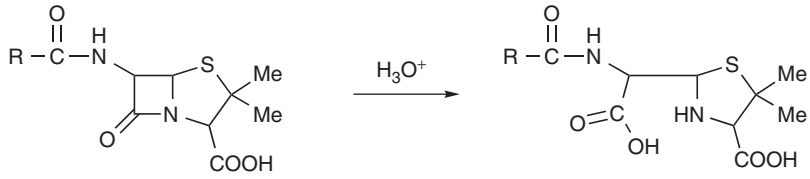


Figure 7.9 *Hydrolytic ring opening of penicillin.*

The first medicinally important penicillin G contained an amino group acylated by phenylacetic acid at the 6-position. This compound, however, had the significant disadvantage of exhibiting activity against Gram-positive bacteria only. Gram-negative species including *Escherichia coli* remained completely unaffected. Gram-negative bacteria possess a highly effective protective mechanism involving a β-lactamase enzyme (penicillinase) which cleaves the lactam OC—N bond with great selectivity, thus deactivating the antibiotic.

Attempts were made to change the structure of the original penicillin G to render it resistant toward hydrolysis. However, the only type of modification to be successful without the loss of biological activity involved changing the substitution pattern at the aminocarbonyl group in penicillin G. This involves changing the acid residue attached to the amino nitrogen. Many new penicillins are prepared biosynthetically from mixtures of water, carboxylic acid and other components using a diversity of mold fungi. In the penultimate stage of biosynthesis 6-aminopenicillanic acid is formed. Subsequent acylation by the chosen carboxylic acid gives the desired penicillin derivative (Figure 7.8).

Among the thousands of semisynthetic penicillins prepared, a number exceed penicillin G in terms of hydrolytic stability and demonstrate an appreciable effect on Gram-negative bacteria. Such was the case for phenoxymethylpenicillin, ampicillin and amoxicillin (Figure 7.8). Amoxicillin, discovered in 1972, is a broad-spectrum antibiotic which can be taken orally. However, new generations of penicillins also did not escape an appearance of antibiotic-resistant strains of bacteria capable of destroying the β-lactam ring. Several approaches were put forward to solve this problem. One of them is connected with clavulanic acid – also a β-lactam differing from penicillins by the presence in the five-membered ring of an oxygen atom instead of sulfur (Figure 7.10). Clavulanic acid has low antibacterial activity but by chance it effectively inhibits β-lactamases. Thus, an idea was born to combine in one remedy an antibiotic and β-lactamase inhibitor. A good example of its successful realization is Amoxiclav, a combination of amoxicillin and clavulanic acid.

On the way of further modification of the penicillin nucleus carbapenems and monobactams were developed and put into medicinal practice (Figure 7.10). The former are close analogues of penicillins in which a sulfur atom in the thiazolidine ring is replaced by carbon. Carbapenems

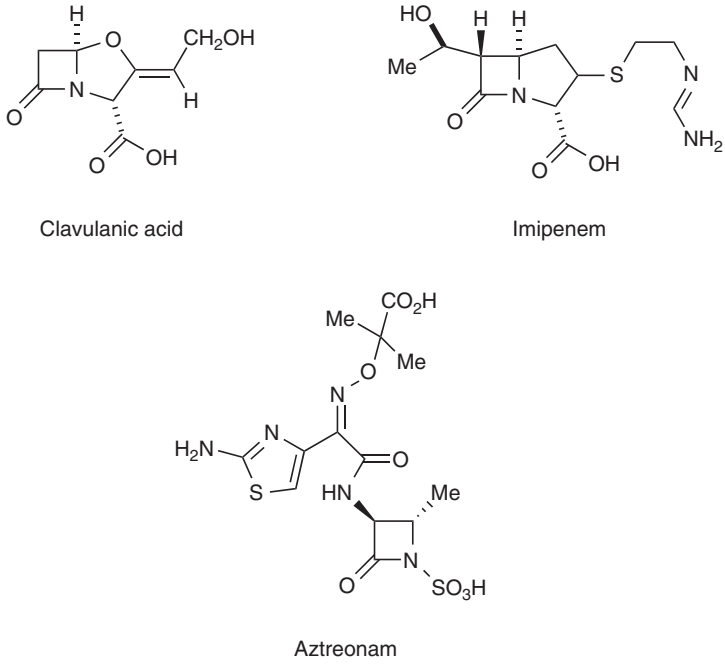


Figure 7.10 *Clavulanic acid and two representatives of carbapenems and monobactams.*

display a broad spectrum of activity and are resistant to several classes of β-lactamases. One of the most successful drugs of this group is imipenem. Monobactams, for example, aztreonam, are monocyclic β-lactam antibiotics. They are rather stable to β-lactamase and are active especially towards Gram-negative bacteria.

In 1948, the Italian scientist G. Brotzu isolated from a sewer in Sardinia cultures of *Cephalosporinium acremonium*. Soon, it was found that these cultures produced an antimicrobial substance called cephalosporin C (Figure 7.11). Because of its only moderate activity cephalosporin C is not used clinically. However, it could be hydrolytically converted into 7-aminocephalosporanic acid which served as a starting material for the preparation of semisynthetic cephalosporin antibiotics (Cephs). As penicillins, cephalosporins belong to β-lactams but differ from the former by the presence of a 1,3-thiazine ring instead of a thiazolidine. Due to the greater stability of the lactam ring, Cephs show greater effectiveness against Gram-negative bacteria than the penicillins. Moreover, the structure of cephalosporins may be modified not only at the amido nitrogen (7-position) but also at the 3-position. Numerous semisynthetic first-, second-, third- and fourth-generation Cephs have been obtained, including cefatrizine, cefuroxime and cefotaxime (Figure 7.11). Each new generation of Cephs is characterized by an increase in activity. Thus, cefotaxime (a third-generation antibiotic) is highly effective against resistant forms of Gram-negative bacteria.

All β-lactam antibiotics have the same mode of antimicrobial activity: they disrupt the synthesis of the peptidoglycan layer of bacterial cell walls. Bacteria need a strong cell wall to resist inner osmotic pressure. This strength is provided by peptidoglycan – a lattice-like polymer composed of repeated disaccharide units. Each such unit is built from N-acetylglucosamine and N-acetylmuramic acid, as shown in Figure 7.12a. Additionally, a short peptide chain, normally including L-alanine,

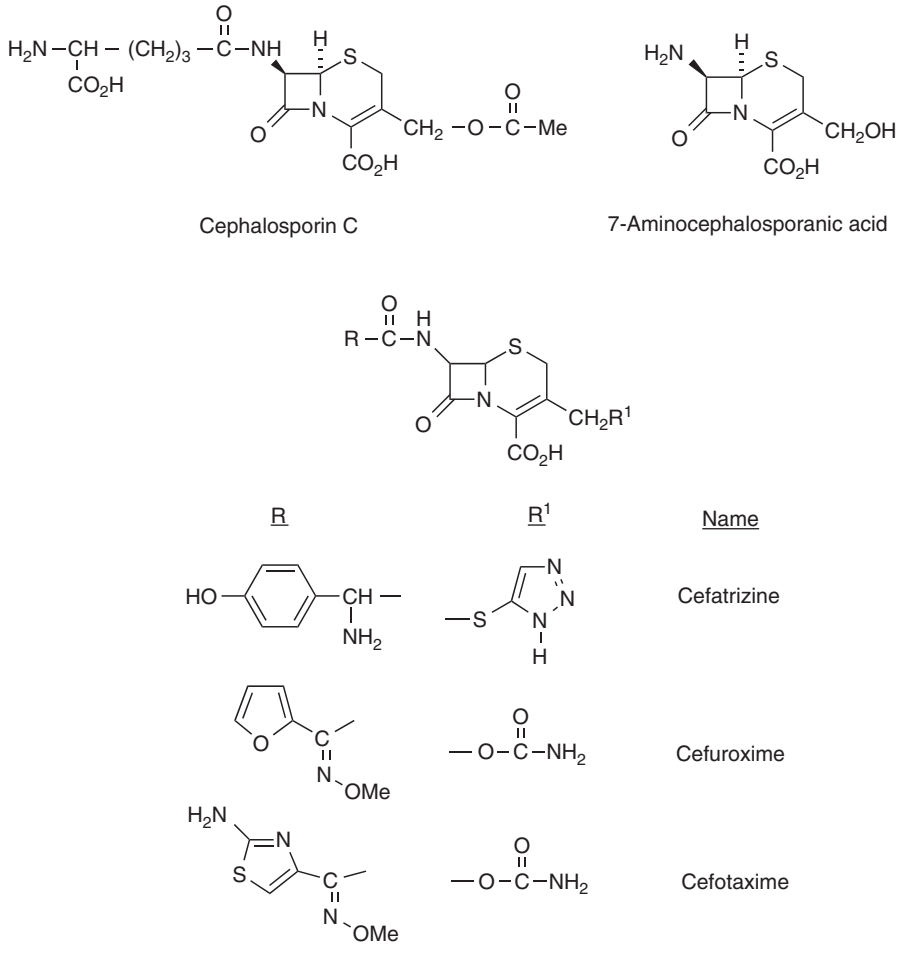


Figure 7.11 *Selected cephalosporins.*

D-glutamic acid, L-lysine and D-alanine residues, is attached to N-acetylmuramic acid. The main function of this peptide chain is to provide cross-linking between the parallel peptidoglycan strands (Figure 7.12b). The cross-linking reaction is catalyzed by the bacterial enzyme transpeptidase and proceeds with the participation of three components: the D-alanine carboxy group of one strand, the L-lysine amino group of another strand and the pentaglycine olygopeptide, $(Gly)_5$. In the beginning, transpeptidase activates the CO_2H group of D-alanine by means of its esterification with the enzymic Ser62 active site (Figure 7.13a). The enzyme-substrate complex thus formed then reacts with the end NH_2 group of pentaglycine that leads to peptide bond formation (Figure 7.13b). The next (and final) peptide bond formation between the end CO_2H group of pentaglycine and the NH_2 group of L-lysine proceeds similarly producing the necessary cross-bridge (Figure 7.13c).

The high activity of β-lactam antibiotics is caused by their ability to mimic carboxylic groups of D-alanine and pentaglycine. Interaction between Ser62 and the β-lactam fragment results in ring-opening of the latter (Figure 7.13d), blocking the activity of the transpeptidase which is responsible

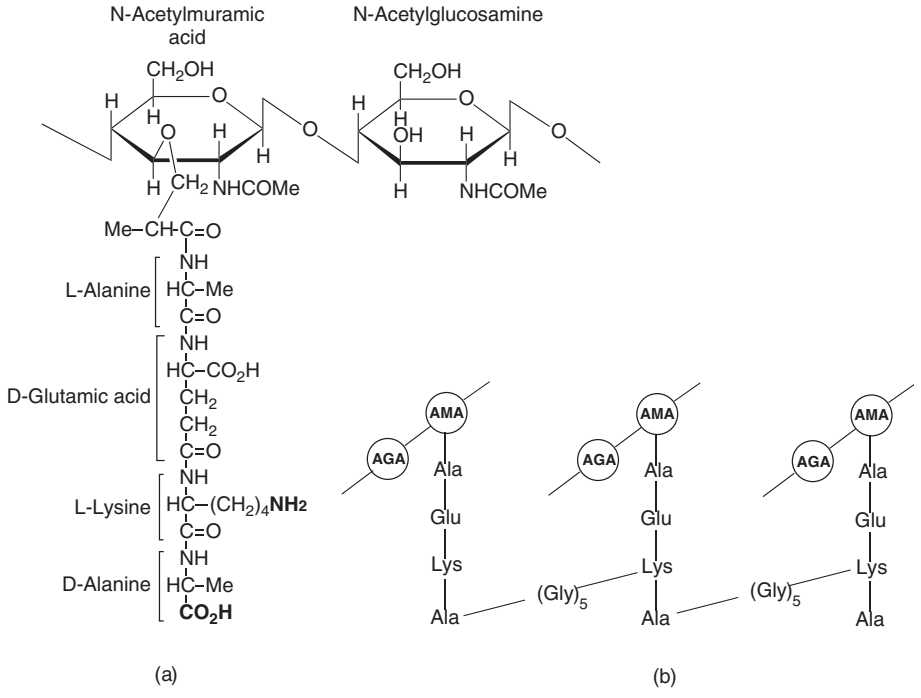


Figure 7.12 *Constituents of bacteria cell wall: (a) disaccharide unit composed of N-acetylmuramic acid (AMA) with attached oligopeptide chain and N-acetylglucosamine (AGA), COOH and NH_2 functions engaged in cross-linking reaction are indicated by bold, (b) schematic representation of cross-linking between neighboring polysaccharide strands.*

for the cross-linking. Thus, in the presence of antibiotics, the bacterial cell walls become weaker and are ruptured by osmotic pressure, ultimately killing the bacteria.

A great peculiarity of bacteria is that their reproduction cycle does not exceed 2 h, and often it is much shorter. In other words, bacteria have enough time to produce as a result of random mutations the antibiotic-resistant strains. The character of these mutations depends on the antibiotic's nature. Thus, in the case of β-lactams bacteria produce various kinds of β-lactamases. Usually the effectiveness of the antibiotic decreases considerably after 3–5 years following introduction in clinics. Consequently, there is a constant need for the development of new antibiotics. Scientists continually expand searches for new sensitive bacterial targets. Inhibiting protein synthesis or a genetic system of bacteria is considered to be especially promising. Since the start of the 1950s the antibiotic erythromycin (Figure 7.14), primarily isolated from Philippine soil samples, has found wide application. It was the first representative of the so-called macrolide antibiotics by having a large lactone ring. In the erythromycin molecule the 14-membered ring includes an ester oxygen as an heteroatom. As penicillins, erythromycin is active against Gram-positive infections but does not cause the allergic reactions that some individuals suffer after use of penicillins. A newer and much more potent macrolide drug is semisynthetic azitromycin (Figure 7.14). Unlike erythromycin, the azitromycin has an additional (nitrogen) heteroatom in its 15-membered macroring. Macrolide antibiotics have been shown to bind to the large subunit of a bacterial ribosome, blocking its exit site (Figure 3.14) and thus interupting protein synthesis. One can assume that the large surface of macrolide molecules strongly enhances their ability to such binding.

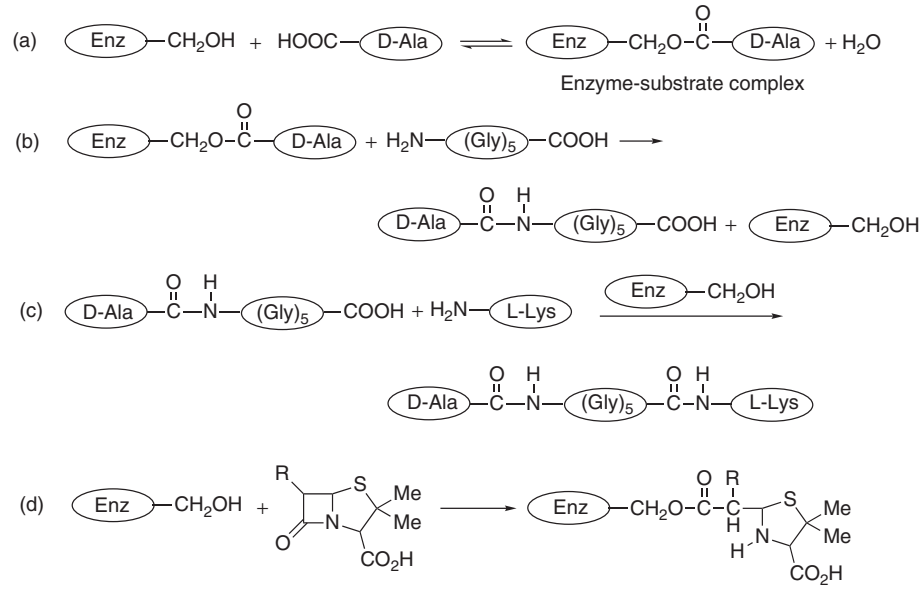
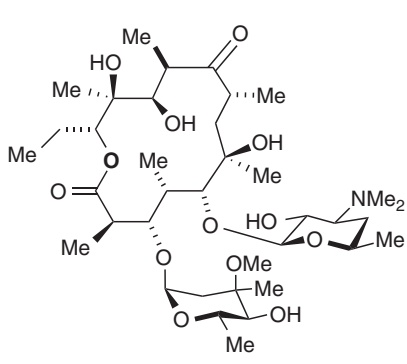


Figure 7.13 *(a) Enzyme activation of a D-alanine carboxylic group. (b, c) Cross-linking of D-alanine and L-lysine moieties via pentaglycine peptide bridge. (d) β-Lactam ring mimics the D-alanine carboxylic group deactivating enzyme.*

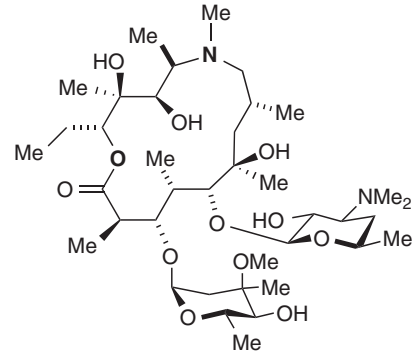
Another kind of macrocyclic drug is daptomycin – the first antibiotic of the lipopeptide type, which had been uncovered at Turkey's Mount Ararat. The daptomycin molecule contains a 31-membered heterocyclic ring with nine nitrogen and one oxygen heteroatoms (Figure 7.14). Since 2005 it has been used in the United States under the trade name Cubicin against Gram-positive skin infections. The mechanism of biological activity of daptomycin is rather complex. The antibiotic binds to a bacterial cell membrane and causes its rapid depolarization, resulting in a loss of membrane electrical potential with subsequent inhibition of protein, DNA and RNA synthesis.

7.2.4 Antibiotics From the Ocean's Depths

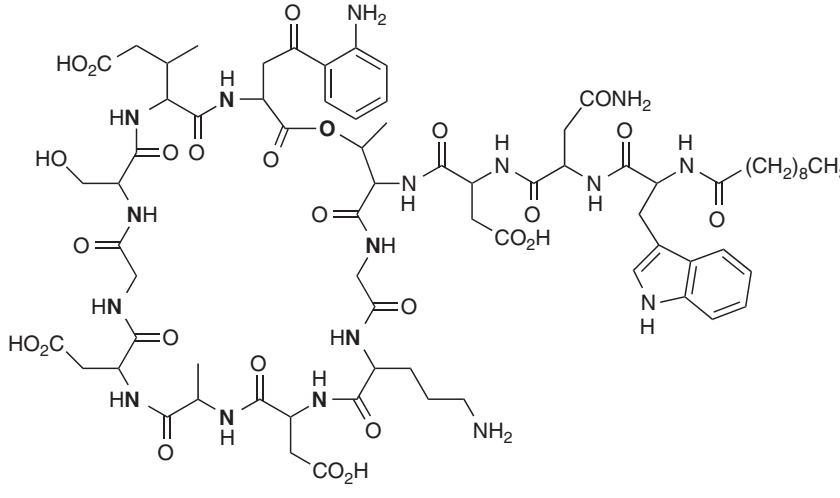
Development of novel antibiotics is both an expensive and a time-consuming business. First of all this is caused by the resourcefulness of bacteria, their unique ability to biosynthesize antibiotic-resistant strains. However, there is another reason. Apparently, the number of antibiotic structures in nature is limited. Indeed, the discovery of a new class of antibiotics happens quite rarely. Table 7.1 lists all the memorable antibiotic drugs in chronological order. One can see that for their 75-year history, starting from sulfonylamides, there were discovered only about ten types of natural antibiotics and half of these are purely synthetic. Heterocyclic compounds make up the majority of successes in this endeavor. Notably, most of these discoveries fall in the first quarter of the antibiotics era. In the next 50 years only one new type of natural antibiotic was found, namely daptomycine. Does this mean that all the low-hanging fruit have already been picked? Some scientists believe that only the tip of the iceberg has been explored up to now and many novel antibiotic structures are still hidden in less available parts of our planet, such as ocean depths or the core of snowflakes. In confirmation of such a view an unusual antibiotic, dubbed marinopyrrole, was recently isolated from ocean sediments (Figure 7.15a). Here two pyrrole rings



Erythromycin



Azitromycin



Daptomycin

Figure 7.14 *Examples of macrocyclic antibiotics.*

are connected via N and C-2 atoms. The marinopyrrole displays rather high antimicrobial activity against penicillin-resistant bacteria.

One cannot totally ignore synthetic antimicrobial preparations. As far back as the 1950s, nitrofuran drugs possessing high potency toward both Gram-positive and Gram-negative bacteria were introduced into clinical practice. Furazolidone, for example (Figure 7.15b), was used for a long time in the treatment of dysentery and enteric and paratyphoid fevers. Nitrofurans often moderate the development of microorganisms that are resistant to sulfa drugs and antibiotics. However, in recent years the use of nitrofurans has been phased out owing to their tendency to show mutagenic activity. To exclude the toxic effect the nitrofuran moiety was eliminated and the first representative of the new class of oxazolidine derivatives which show high antibacterial activity was recently found and introduced in clinical practice under the name linezolid (Figure 7.15c). This antibiotic agent binds specifically to RNA in bacterial ribosomes and inhibit protein synthesis.

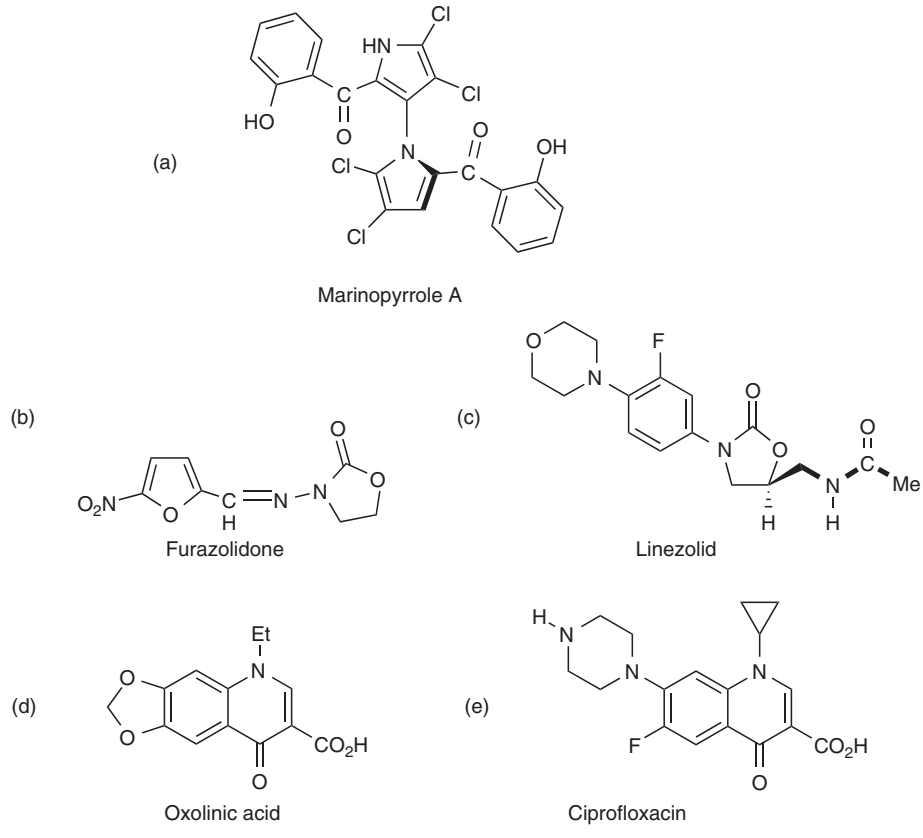


Figure 7.15 *(a) Marinopyrrole A, isolated from ocean sediments, and some purely synthetic antibacterial drugs of these series: (b) nitrofuran, (c) oxazolidinone, (d, e) 4-quinolone-3-carboxylic acid.*

Table 7.1 *Historical survey of the discovery of the main classes of antimicrobial drugs*

Year introduced	Class of drug	Origin	Presence of heterocyclic fragment
1935	Sulfonamides	Synthetic	Frequent
1942	Penicillins	Natural	Yes
1944	Aminoglycosides	Natural	Yes
1949	Chloramphenicol	Natural	No
1950	Tetracyclines	Natural	No
1952	Cephalosporins	Natural	Yes
1952	Macrolides	Natural	Yes
1950s	Nitrofurans	Synthetic	Yes
1958	Glycopeptides	Natural	Yes
1958	Rifamycins	Natural	Yes
1960s	Nitroimidazoles	Synthetic	Yes
1962	Fluoroquinolones	Synthetic	Yes
2000	Oxazolidones	Synthetic	Yes
2005	Lipopeptides	Natural	Yes

Synthetic derivatives of 4-quinolone-3-carboxylic acid has occupied an important niche among the antibacterial 'weapons' since the early 1970s. A typical representative is oxolinic acid (Figure 7.15d). Like other compounds of the series, oxolinic acid displays a wide spectrum of action, especially towards Gram-negative bacteria. These quinolone derivatives are active against bacterial strains which are resistant to sulfanilamides and antibiotics. Ciprofloxacin, containing a fluorine atom at the 6-position and an amine function at position 7, belongs to the most recent generation of the quinolone series (Figure 7.15e). The quinolone drugs target the protein enzyme DNA gyrase which unwinds the tightly packed double-helical structure of bacterial DNA. Such action prevents bacterial replication and transcription.

7.2.5 Heterocyclic Antifungal Agents

The widespread use of current antibiotics, antifungal and immunosuppressive drugs caused increasingly resistant fungal pathogens, both superficial and systemic, which led to a significant increase in mortality among AIDS, cancer and transplant patients. Due to the urgent need to cope with this problem, the past two decades of research has given rise to several new efficacious heterocyclic compounds with a broad spectrum of antifungal action. The most representative groups include derivatives of imidazole and 1,2,4-triazole.

Imidazoles such as miconazole, ketonazole, butaconazole and lanoconazole (Figure 7.16) were developed to treat not only superficial skin, nail and hair fungal infections but also systemic ones. Interestingly, only the *R*-enantiomer of lanoconazole is bioactive. The majority of triazole derivatives act successfully on systemic fungal infections, deep-seated mycoses caused primarily by *Candida* and *Aspergillus* species. Among these triazoles are fluconazole, voriconazole, terconazole and itraconazole (Figure 7.16).

Most azole derivatives interact with fungal cytochrome P-450 and inhibit biosynthesis of ergosterol and its regulatory functions in cell proliferation and membranes integrity. For example, the main target of the itraconazole and lanoconazole toxic action is the cytochrome iron atom, which is chelated by them. This results in deactivating the enzyme lanosterol-14-α-demethylase, which controls the synthesis and metabolism of the lanosterol intermediate in the ergosterol formation pathway.

7.2.6 Heterocycles Against Parasitic Diseases

Parasitic diseases such as malaria, trypanosomiasis, leishmaniasis and chronic diarrhea caused by plasmodium, trypanosome, leishmania and intestinal protozoa, respectively, affect at present many millions of people, with a high mortality rate. Malaria remains the most serious health problem: according to World Health Organization data, the annual number of cases reaches 500×10^6, of which nearly 2.5 million are fatal (mainly children). Despite its rather high effectiveness, quinine (Figure 7.1) is not an ideal antimalarial drug because of its marked toxicity. A further drawback is that cinchona trees still cannot be cultivated worldwide. The first approaches to obtaining analogues of quinine were directed to modify the structure of quinine, which was established only in 1944. In Germany and Russia this approach led to the discovery of some new efficacious preparations, namely chloroquine, plasmocid and others (Figure 7.17). Similar programs involving dozens of universities and companies were developed in the United States during World War II and the Korean and Vietnam wars. More than 30 000 compounds, mainly quinoline derivatives, were synthesized and tested for antimalarial activity. Up to the present time, chloroquine remains the most important agent used as such or in combination with quinine and other agents to cope with malaria. Chloroquine is considered to be the third largest drug by tonnage produced and consumed in the world.

The constant search for drugs for prophylaxis and curing malaria has recently given two natural leads which have a marked level of bioaction. These are artemisinin and yingzhaosu A

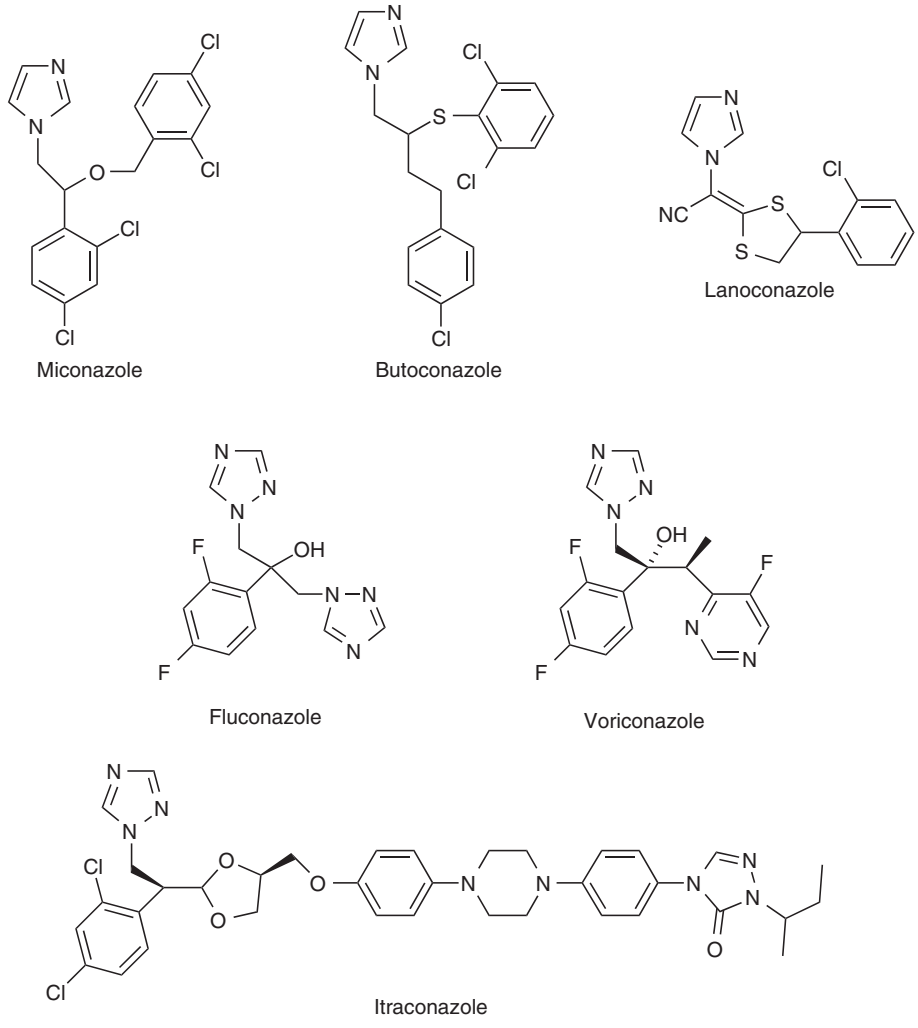


Figure 7.16 *Structures of some clinically used antifungal agents based on imidazole and 1,2,4-triazole.*

(Figure 7.17), isolated from a Chinese medicinal herb. From these a number of potent antimalarial derivatives have been synthesized, for example, arteether or recently reported compound **1**, curing the disease with just one to three doses (only tested in mice). Though the exact mechanism of biological activity of artemisinins is still unproved, it can be attributed with great probability to the well known weakness of the peroxide bond O—O. This bond is easily reduced by electron transfer from Fe^{2+}, producing a variety of radical species which can damage the parasite.

About 20 million people in the world are infected with different species of trypanosome protozoa, with many deaths every year. The disease caused by these parasites is called sleeping sickness

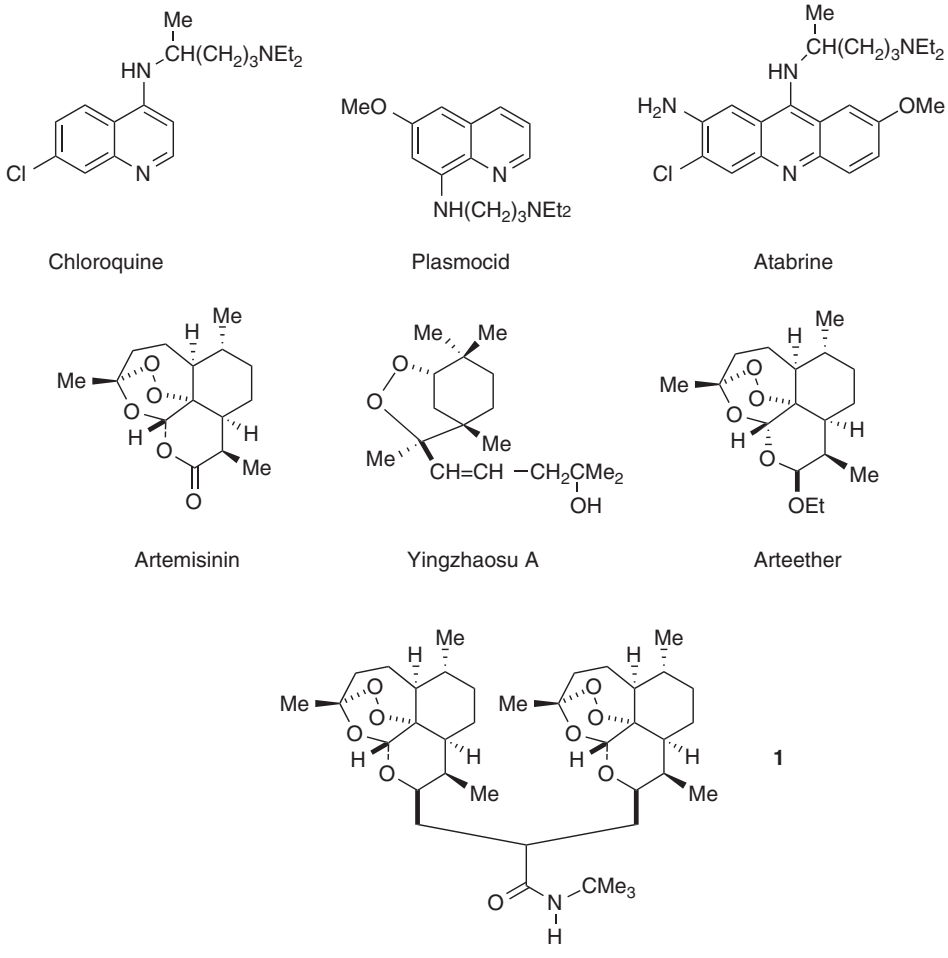


Figure 7.17 *Some heterocyclic derivatives with antimalarial activity.*

in Africa and Chaga's disease in Latin America. Only a few trypanocides are known to fight these diseases, two of which, nifurtimox and benznidazole, belong to heterocyclic derivatives and are the most effective (Figure 7.18).

In tropical and subtropical countries more than 15 million people are infected with cutaneous and visceral leishmaniasis, which is transmitted mainly by bites of sandfly females infected with leishmania. The disease emerges among AIDS patients, aggravating their treatment. Metronidazole (Figure 7.5) along with some other drugs is used to cure leishmaniasis. Thus, nitroheterocyclic preparations, despite the risk of mutagenic side effects, in many cases remain irreplaceable. One more example of this kind is nitazoxanide (Figure 7.18), a derivative of 5-nitrothiazole, which was introduced in clinical practice (in Latin America and the USA) after 1996 for the treatment of intestinal parasitic infections.

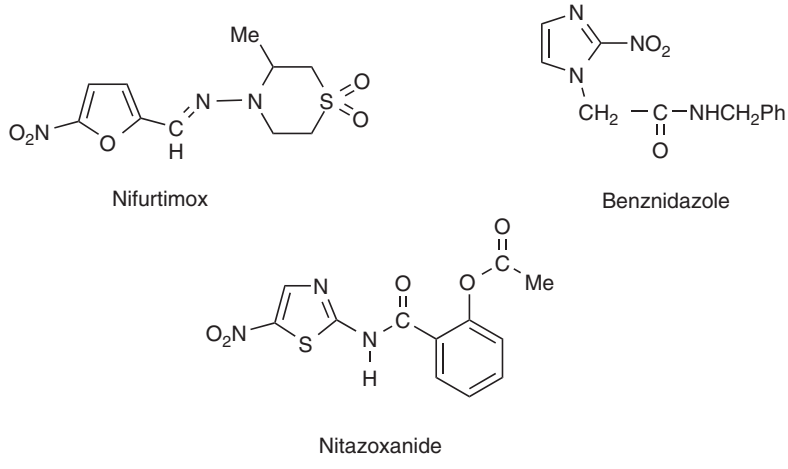


Figure 7.18 *Some heterocyclic drugs effective in the treatment of trypanosomiasis and intestinal infections.*

7.3 Heterocycles and Viral Infections

In our generation, enormous scientific efforts and financial input have been devoted to finding a cure for various viral diseases, including influenza, hepatitis, encephalitis, poliomyelitis and, of course, HIV. As a result, several rather potent antiviral drugs are currently available and their numbers grow.

What is the main barrier to the effective drug therapy of viral diseases? Surprising as it may seem, it is the relative simplicity of the structure and biochemistry of the replication of viruses that is paramount. Bacteria possess rather complicated organizations with significant differences between their cells and those of animals. We have already discussed the dependence of bacterial growth on *p*-aminobenzoic acid and some of the structural peculiarities of bacterial cell walls. However, there are further significant differences. Bacterial cells can multiply independently without penetrating into a human cell. This is not the case for viruses. Viruses are living organisms that are thousands of times smaller than bacteria. They are usually composed of a genome (DNA or RNA molecule) and sometimes a few enzymes which are stored in a protein shell called a capsid. The capsid, in turn, is enveloped by a glycoprotein capsule that plays a great role for binding and viral penetration into a host cell.

A virus can only multiply inside a host cell using its machinery. Once inside, the virus inserts its single DNA molecule, which contains a coded genome, into the DNA of the host cell. Now the invader commences the process of viral DNA replication; and the rapid assembly of protein envelopes the newly formed viral DNA using building blocks within the host cell. In these processes the parasitic guests destroy the host cell. A multitude of viruses are thus formed which are then released to invade further cells of the infected organism.

If the virus belongs to an RNA type (e.g., influenza, ebola, hepatitis A and C), it first produces its own DNA molecule inside a host cell. This process, catalyzed by the enzyme reverse transcriptase, is targeted by many antiviral drugs. In RNA-containing viruses, the RNA molecule is bound to the ribosome rather than to the mRNA of the host cell. This leads to the synthesis of virus proteins.

The biological action of viruses makes it very difficult to create an effective drug because the therapeutic agent would need to be able to differentiate between the biological target, the

virus and the host cell. Generally, an antiviral targeting technique is based on the following approaches: (i) to hinder uncoating of viruses or their penetration into the host cell, (ii) to interfere with the ability of a virus to self-reproduction once it infiltrates a human cell and (iii) to prevent completed viral particles from releasing from the host cell. There are very few agents acting via the penetration–uncoating mechanism. The best known examples are amantadine and rimantadine – derivatives of the cage hydrocarbon adamantane (Figure 7.19). They are used against rhinoviruses which cause the common cold. An equally important heterocyclic drug of this type is maraviroc, containing 1,2,4-triazole and piperidine moieties. Maraviroc was introduced in clinical practice as an anti-HIV drug in the early 2000s. It blocks the attachment of HIV to surface receptors of T-cells of the human immune system.

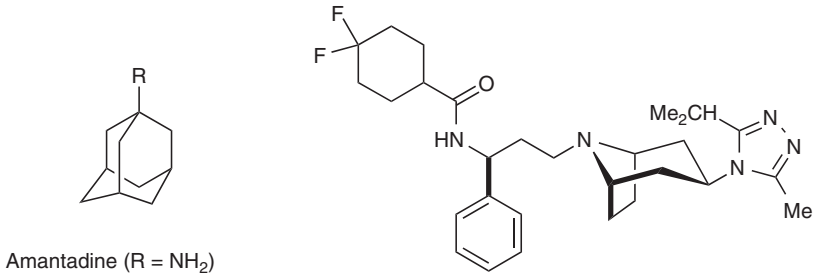


Figure 7.19 *Examples of antiviral drugs preventing virus penetration into host cells.*

The second approach turned out to be even more fruitful, since it led to the discovery of dozens of rather effective antiviral agents. Practically all of them belong to nucleoside or nucleoside-like compounds and contain natural or (rarely) purely synthetic heterocyclic bases. Now they are widely used alone or in combination with other drugs to treat herpes infections, HIV, hepatitis B and C and other dangerous diseases. The action of these drugs is normally directed against the synthesis of viral DNA, for example, via inhibition of the enzyme reverse transcriptase (RNA → DNA). For the synthesis of the highly specific antiviral drug acyclovir (Figure 7.20), the first of this kind, the Americans Elliot and Hitchings were awarded a Nobel prize in 1988. Acyclovir is highly effective against herpes infections. In the United States only, approximately 60×10^6 people are infected with this virus. Zoster encephalitis is especially dangerous, and only 25% of patients with brain inflammation previously survived. The survival rate increased to 75% with the aid of acyclovir. Acyclovir's activity results from its structural similarity to deoxyguanosine, a nucleoside necessary for DNA assembly (see Section 3.1). Once inside an infected cell, acyclovir is phosphorylated at its hydroxy group to produce the triphosphate (Figure 7.21). The first phosphorylation proceeds under the control of the enzyme thymidine kinase formed by the virus itself, while addition of the remaining two phosphate residues is mediated by cellular kinases. During the replication of viral DNA the viral enzyme DNA polymerase mistakenly inserts acyclovir triphosphate into the growing viral DNA chain instead of deoxyguanosine triphosphate. In contrast to deoxyguanosine, the drug contains no 3′—OH group to which the following nucleotide can be linked. Thus, the chain is prevented from growing further.

The reasons why acyclovir is not incorporated into the host cell DNA are not yet completely clear. We do know that human cellular kinases are unable to catalyze effectively the first phosphorylation and, furthermore, that the DNA polymerase of herpes viruses has a much greater affinity

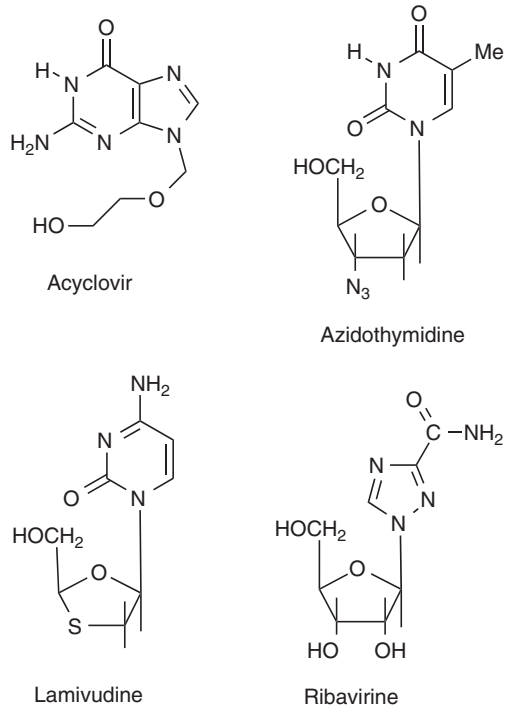


Figure 7.20 *Some antiviral drugs inhibiting synthesis of viral DNA.*

toward acyclovir triphosphate than the host cell DNA polymerase. Thus, the possibility of creating an antiviral drug with high specificity has been realized in principle.

Elliot and Hitchings introduced another antiviral drug, azidothymidine or AZT (Figure 7.20) into medical practice. AZT, now officially named zidovudine, is useful in the treatment of HIV. It impedes the human immunodeficiency virus replication process but does not cure the disease completely. The drug is an analogue of the naturally occurring nucleoside thymidine and its mechanism of action is similar to that of acyclovir, that is, consistent with the metabolite concept.

Lamivudine and ribavirine are two more representatives of antiviral drugs with nucleoside structures (Figure 7.20). Lamivudine is well tolerated and widely used for the treatment of hepatitis and HIV, in the latter case often in combination with AZT. The new standard antiviral therapy for more than 150 million worldwide patients with chronic hepatitis C is based on using ribavirin in combination with protein interferon alpha added to a polyethylene glycol (so called pegylated interferon). In such a combination approved 10 years ago, the drugs are better protected against rapid enzyme degradation and have improved physical and thermal stability and a clinically approved synergistic effect. The combination eradicates the hepatitis C virus (HCV; discovered in 1989) in more than 50% of patients and prevents liver cirrhosis and, thus, hepatocellular carcinoma. The mechanism of action of the ribavirin is very complex and consists in both direct inhibition of HCV through its RNA polymerase and an indirect pathway by enhancement of the immune system of the host to combat the viral infection through activation of defensive T-cells.

At present there seem to exist only two antiviral drugs, approved after 1999, that inhibit releasing newly formed virus particles from infected host cells. These are zanamivir (trade name Relenza)

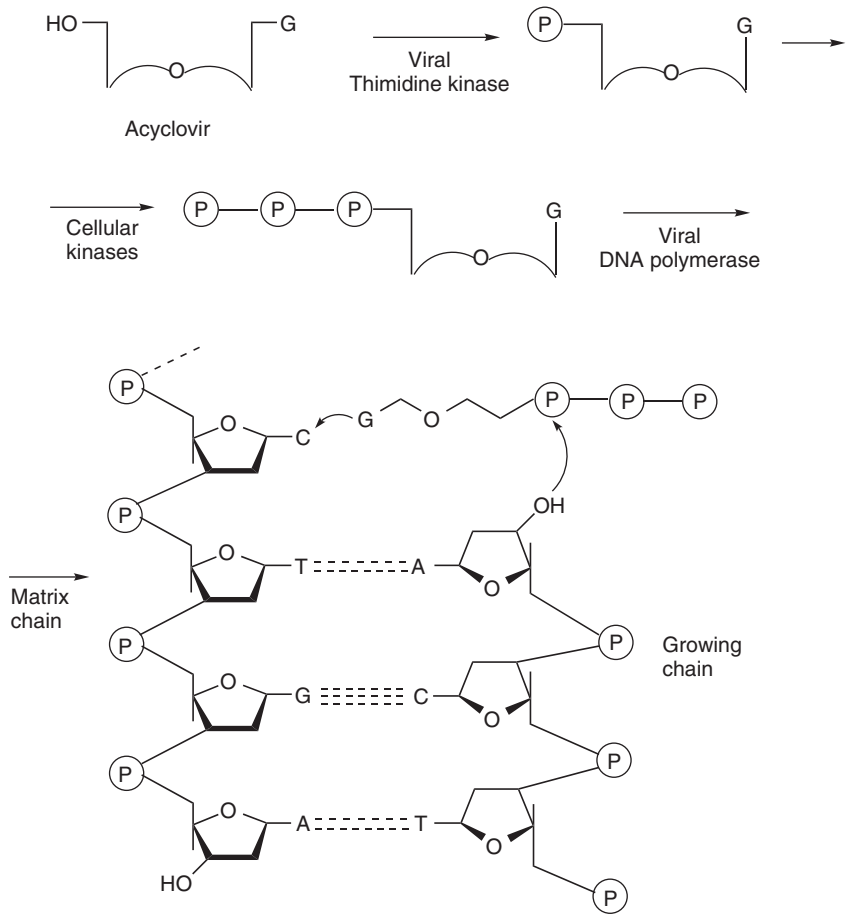
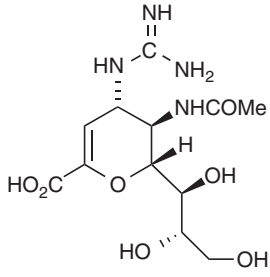


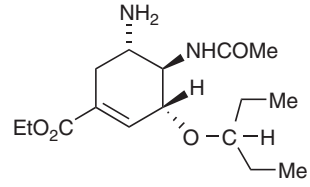
Figure 7.21 *Mechanism of acyclovir action (adapted from Hirsh, M. S. and Kaplan, J. C.,* Sci. Am.*, 1987,* ***256****, 76, with permission. © 1987 Scientific American Inc., All rights reserved).*

and oseltamivir (Tamiflu), very effective in the struggle against influenza A and B, including avian and swine modifications (Figure 7.22). Both drugs inhibit the influenza neuraminidase enzyme that cleaves sialic acid residues that connect the newly formed viruses with the surface of the infected host cell. The enzyme binds a drug and is blocked due to close similarities in structure between sialic acid and antiviral agent. In principle, the presence of a heterocyclic ring here is not necessary for antiviral activity. However, one can suggest that the oxygen heteroatom in the 2H-pyran fragment of zanamavir, as in sialic acid itself, should bring some specificity. Thus, it has been established that the carboxylic group in sialic acid exists in an ionized form that provides better binding with neuraminidase receptors. The enhanced acidity of the CO_2H group in sialic acid may be caused by a negative inductive effect of the α-hydroxyl group and pyran oxygen atom. In the zanamavir molecule the ionization of the carboxylic function, in addition to the influence of the heteroatom, is facilitated by a strongly basic guanidine moiety. From this it is understandable that tamiflu is actually a pro-drug, the ester group of which should be hydrolyzed in the human body before displaying antiviral activity. Which drug is more effective, zanamavir or

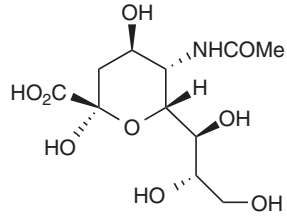
oseltamivir (Tamiflu)? The former is at least as effective as oseltamivir and has fewer side effects. However, a serious limit to zanamavir is its rather low bioavailability. Therefore, it is administered by inhalation whereas Tamiflu (oseltamivir) is taken orally, a much more popular method.



Zanamavir



Oseltamivir (Tamiflu^{TM})



Sialic acid

Figure 7.22 *Two of the most potent anti-influenza drugs which mimic sialic acid – the natural substrate for the influenza neuraminidase enzyme.*

In closing this section we remind readers that the first compounds to be utilized in the struggle against viral infections were heterocycles.

7.4 Heterocycles and the Diseases of Our Century

Since the end of the twentieth century, the two major causes of death in industrially developed countries have become cardiovascular disease and cancer. These diseases were promoted to their present dominant positions owing to the great successes in the control of infectious disease on the one hand, and to the increasingly poor dietary habits combined with increased stress and abnormal environmental factors on the other. Heart disease and cancer, together with nervous system disorders, are therefore often referred to as the 'diseases of the twenty-first century'. Some scientists consider them to be 'specific', that is, programmed into the genetic code to limit the lifespan of an individual organism. Nevertheless, it is generally held that the human race has not yet reached the upper limit of its potential, which is suggested to be 120 or even 150 years. Improvements in social and economic conditions and achievements in medicine, biology and chemistry have already increased the average life expectancy to 80 years in some countries. A central role in this progress is played by modern medicinal chemistry. The treatment of the majority of known diseases can today be carried out reasonably successfully. As regards the 'diseases of the twentieth

century', our achievements are still modest in the case of malignant tumors, but are improving in the therapy of cardiovascular illnesses, and especially in the case of nervous system diseases.

7.4.1 Heterocycles to Cure Stress, Brain Disorders and Pain

The mechanisms of nervous processes in humans are extremely complicated and cannot be discussed here. We merely note that the transmission of nervous impulses is always accompanied by a release of neurotransmitters at the nerve fiber endings. These mediators are chemical substances which affect receptors and induce a particular biochemical response, such as gastric juice secretion or elevation of arterial pressure. As a rule, each mediator interacts with several (usually two to four) specific receptors. Acetylcholine, noradrenaline, adrenaline, dopamine and serotonin are the main neurotransmitters (Figure 7.23). Serotonin functions mainly in the central nervous system and the other four mediators act on the peripheral nerve endings. Drugs generally act on the nervous system by the inhibition of neurotransmitter release or by interaction with specific receptors owing to the similarity in the structure of the drug to that of the mediator. Where a drug mimics a neurotransmitter, the drug can be either an antagonist or an agonist of the mediator. Many neurotransmitters and receptors exist. Therefore, a drug can act on a number of biotargets.

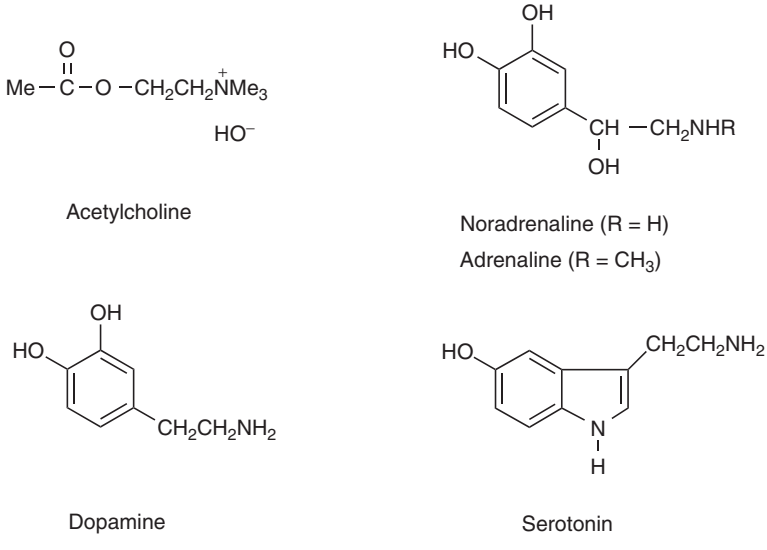


Figure 7.23 *Some neurotransmitters.*

Among the many known nervous system drugs, many different structures are encountered, ranging from aliphatic to heterocyclic, but it is the heterocycles and aromatics which predominate. As examples we mention fluoxetine and citalopram (Figure 7.24), which are recently among the most frequently prescribed antidepressant drugs. They belong to the class of the selective serotonin reuptake inhibitors (SSRI) and are used for the treatment of major depression, anxiety disorders and some related illnesses. It is well known that serotonin positively influences mood. So, it turned out that fluoxetine and citalopram specifically inhibit the mediator reuptake pump in the brain, thus increasing the extracellular level of serotonin available to bind to the postsynaptic receptor. Moreover, both the drugs and serotonin are bioisosterics and interact with the same bioreceptors and certain brain proteins. They have only weak affinity for the noradrenaline and dopamine transporter.

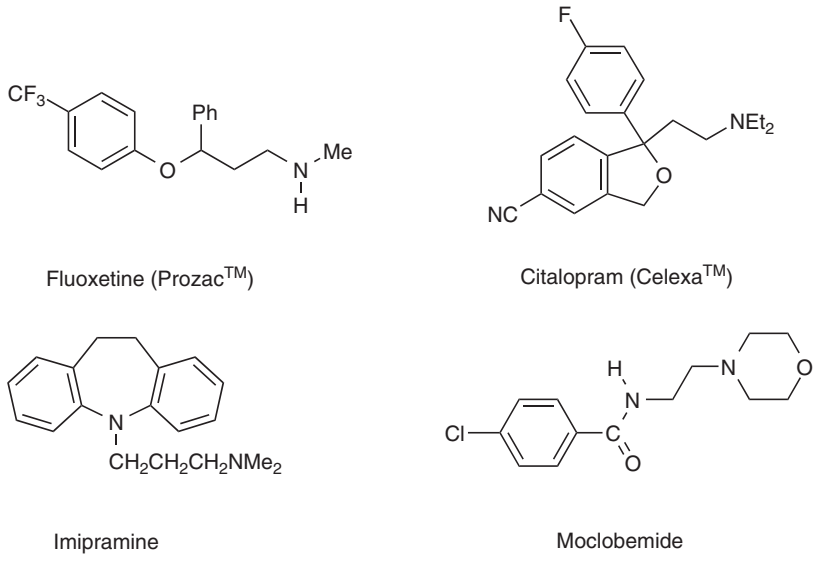


Figure 7.24 *Some antidepressant drugs.*

The era of antidepressant drugs opened at the end of the 1950s when derivatives of dibenzoazepine and some related tricyclic heterocycles were introduced in clinical practice. They are still used, especially in the case of severe major depression. Their typical representative is imipramine (Figure 7.24). The mechanism of biological activity of tricyclic depressants is similar to that of SSRI.

In an organism, serotonin, dopamine, adrenaline and noradrenaline are destroyed by the flavin-containing enzyme monoamine oxidase (MAO). To prevent this process another kind of antidepressant, the so-called oxidase inhibitors (MAOIs), has been developed. MAOIs are as active as tricyclic antidepressants but suffer from some dangerous side effects. Fortunately, a new generation of MAOIs has recently appeared. These preparations are called the reversed inhibitors of monoamine oxidase and are milder and demonstrate much higher selectivity. The morpholine derivative moclobemide (Figure 7.24) is perhaps the most successful among them.

As mentioned, neurotropic activity is an inherent property of many alkaloids, and the first drugs of this type were indeed alkaloids. The mechanism of their neurotropic action and their biological targets are well known. For example, atropine (Figure 7.3) inhibits choline receptors, thus interrupting their interaction with acetylcholine and causing muscle relaxation. Since acetylcholine is a peripheral neuromediator, atropine exerts a powerful local action and was formerly used clinically for treating intestinal, urinary and bronchial spasms. The narcotic effect of the alkaloid morphine (Figure 7.3) is accounted for by its blockage of serotonin receptors in the CNS.

The nature of the stimulatory action of caffeine and other purine alkaloids (Figure 7.2) is intriguing and is related in a rather complex way to noradrenaline and adrenaline function. During emotional excitation the adrenal cortex is known to enhance sharply the release of these two mediators, resulting in an increased flow of blood to all organs. Adrenaline and noradrenaline interact with the enzyme adenylate cyclase, which seems to be one of the receptors. Activated by the neuromediators, adenylate cyclase thus transforms adenosine triphosphate first into the monophosphate and then into cycloadenylic acid (Figure 7.25). Cycloadenylic acid plays the role of a secondary mediator (hormone) by activating phosphorylase, an enzyme which stimulates

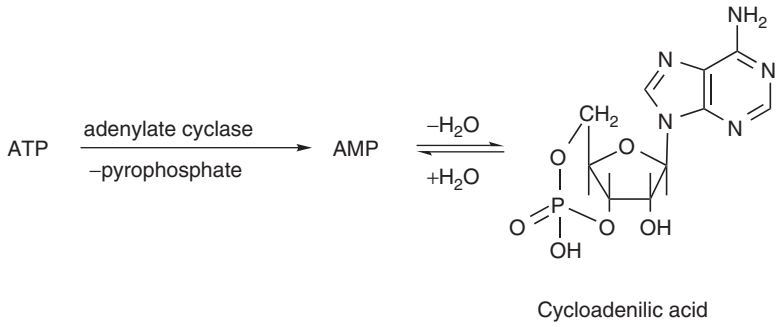
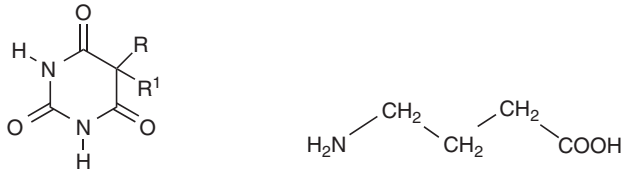


Figure 7.25 *Conversion of ATP into cycloadenylic acid.*

physiological processes such as cardiac activity and glycogenosis in the liver. As a cyclic diester, cycloadenylic acid can be hydrolyzed to AMP. This hydrolysis is catalyzed by a widely distributed natural enzyme. Caffeine and other purine alkaloids are thought to bind the enzyme concerned, resulting in an increased cycloadenylic acid concentration and thus in the onset of a stimulating effect. There is little doubt that purine alkaloids serve as antimetabolites of adenine, which probably interact with the enzyme in question under normal conditions.

Barbiturates, which are 5,5-disubstituted derivatives of barbituric acid (Figure 7.26), were the first synthetic drugs found to exert significant action on the central nervous system (CNS). Barbiturates have a hypnotic effect and suppress the CNS. Consequently, the main uses of such compounds are as tranquilizing and soporific agents. Apart from this, they are widely used as anesthetics in surgery. Figure 7.26 shows two examples of typical barbiturates, barbital[3] and Luminal (phenobarbital). The parent of these compounds, barbituric acid, was synthesized by the famous German chemist von Baeyer in 1864, but it does not markedly affect the CNS.



Barbituric acid (R = R^1 = H)

Veronal (Barbital) (R = R^1 = Et)

Luminal (R = Et, R^1 = Ph)

γ-Aminobutyric acid (GABA)

Figure 7.26 *Examples of barbiturates and their putative target: γ-aminobutyric acid (GABA).*

[3] Veronal (barbital) was the first sleep inducer to be introduced into clinical practice by Mering in 1903. Mering named the drug after the Italian town Verona. The origin of the name 'barbituric' acid is also curious. Willstaetter, who had been a student of A. von Baeyer, revealed that his then young teacher von Baeyer was attracted to a girl named Barbara, and that the first part of the acid's name was derived from her pet name Barbi. The second part of the name is much more prosaic and originates from one of the two raw materials (urea, ethyl malonate) utilized by von Baeyer to synthesize the heterocyclic ureide, malonylurea.

The mechanism of barbiturate action on the CNS is believed to enhance the activity of γ-aminobutyric acid (GABA), the main natural inhibitor of nervous processes in the mammalian brain.

Although barbiturates continued to be used in clinical practice, derivatives of 1,4-benzodiazepine such as diazepam, nitrazepam, phenazepam and others (Figure 7.27) became increasingly important from the beginning of the 1960s. Within a short period of time these derivatives gained worldwide acceptance, judging from their per capita consumption. 1,4-Benzodiazepine tranquilizers reduce or suppress fear, stress and anxiety and have also been used in surgical, pediatric and obstetric applications. Such drugs are routinely prescribed by the military for the treatment of emotional stress caused by the extreme circumstances in which soldiers live and work. 1,4-Benzodiazepine derivatives, like barbiturates, probably increase the inhibitory action of GABA in the cerebral cortex via their own specific receptors.

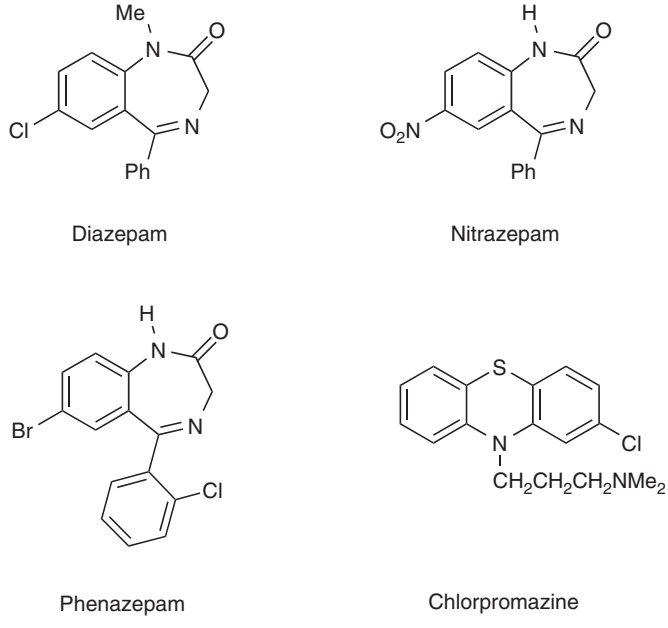


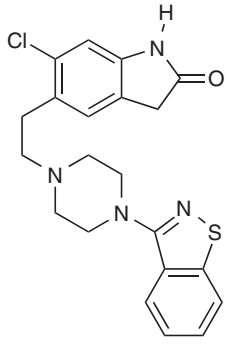
Figure 7.27 *Tranquilizers and neuroleptics.*

Our historical account skipped a few pages when we first mentioned the 1,4-benzodiazepine tranquilizers because the revolution in psycho-pharmacology was initiated with the derivatives of another heterocyclic system, namely phenothiazine (Figure 2.11). Phenothiazine derivatives were first introduced into clinical practice in the early 1950s. Chlorpromazine (aminazine) is the outstanding representative of the class (Figure 7.27) and has been widely used for the treatment of various mental disorders including schizophrenia. Phenothiazine-based drugs can reduce aggressiveness, phobias and reactions to external stimuli. Unlike the 1,4-benzodiazepines, phenothiazine derivatives are able to halt episodes of delirium and hallucinations, and phenothiazines are also devoid of pronounced somniferous (sleep-inducing) side effects. Compounds exhibiting such activity are called neuroleptics. They are believed to cancel the excitatory effects of adrenaline and,

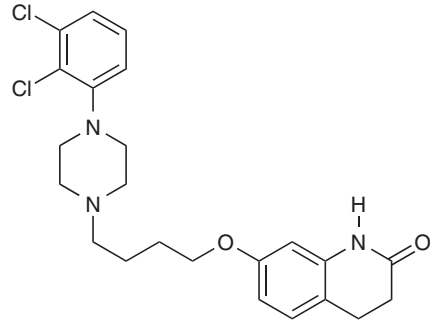
in particular, serotonin when the levels of these amines in the cerebral cortex are raised by a metabolic disorder.

The discoverer of ascorbic acid, Nobel prize winner Szent-Györgyi observed that phenothiazines, like adrenaline and serotonin, are strong electron donors (see Section 2.4.4). He therefore came to the conclusion that their biological effect is due not only to their shape but also to their electronic features. The phenothiazine system apparently forms a type of donor–acceptor molecular complex with a site on the biological target.

A few preparations of a novel class of compounds to combat schizophrenia recently appeared in the pharmaceutical market. They may be exemplified by ziprasidone and aripiprazole (Figure 7.28). These antipsychotics are antagonists not only for dopamine receptors as is the case of usual drugs (Figure 7.27), but also for serotonin ones. The derivatives of piperazine are considered as the atypical drugs and turned out to be more efficacious compared to conventional antipsychotics.



Ziprasidone



Aripiprazole

Figure 7.28 *Synthetic atypical antipsychotics.*

Drugs which have a stimulating action are used in the treatment of narcotic or hypnotic intoxications, or to increase cardiac activity when this has been slowed down during the course of an operation. Such compounds are named analeptics (Latin: *analeptica*, reanimating) and are of great importance in medicine. Analeptics influence the CNS by exciting respiratory and vasomotor centers in the cerebral cortex. Among the chief analeptics currently used are two heterocyclic derivatives, corasole and nikethamide, which are derivatives of tetrazole and pyridine, respectively (Figure 7.29).

From time to time everyone feels a rush of inspiration – a kind of inner energy, emotional excitement. This state of mind is especially appreciated by creative workers – artists, poets, musicians and scientists. Young Delvig, a friend of the famous Russian poet Alexander Pushkin, regretted: 'Not often does inspiration come to us'. There is no doubt that the basis of inspiration is of a biochemical nature, most likely connected with vasodilatation (the expanding of brain blood vessels), blood circulation improvement, concerted work of neuromediators and nervous system as a whole. As the years pass, the minutes of inspiration become more seldom and shorter, and the creative ability weakens. However, an idea to discover or invent a harmless elixir of creativity is getting to be a reality due to the appearance of a new class of medical preparations named nootrops. Nootropic preparations became very a popular means for ameliorating the integrative functions of brain, memory, attention and educational capabilities. All of them, for instance, piracetam (normabrain), etiracetam and pramiracetam, are derivatives of α-pyrrolidone which is an

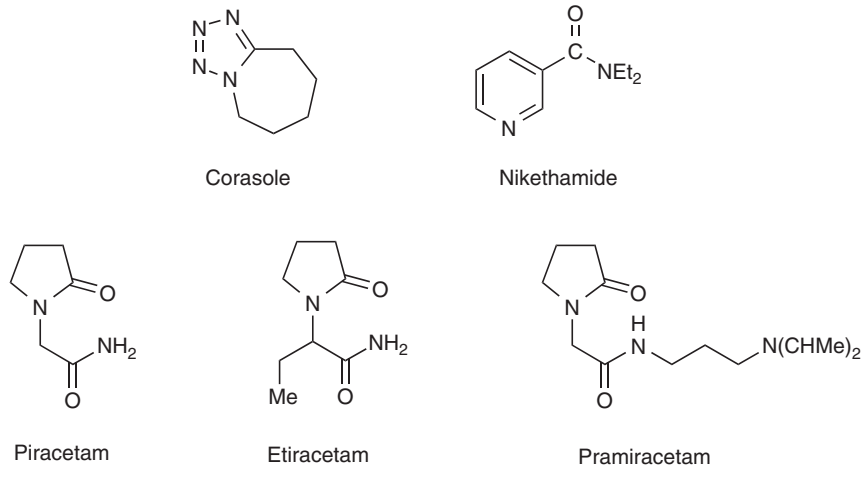


Figure 7.29 *Synthetic analeptics and nootrops.*

internal amide (lactam) of γ-aminobutyric acid (Figure 7.29). Nootrops are believed to penetrate through a hemato-encephalic barrier more easily than GABA itself and normalize the functions of this neurotransmitter. In clinical practice these acetams are used to cure patients with skull and brain traumas, stroke and atherosclerotic disturbances of brain vessels in old people and mentally retarded children. Some nootrops are even sold as a dietary supplement. The main deficiency of the known nootrops is their short-term action due to rapid metabolism by enzymes and acids in the stomach and blood. This problem demands the design of new structural analogues of this and other classes. So, recently, levetiracetam (pure S-enantiomer of etiracetam) appeared in the market. It has both significant efficacy and high tolerability and is recommended as an add-on treatment for seizures in adults. Its antiepileptic action is uniquely exerted via the brain-specific protein SV2A which serves as a binding site to this drug.

Everybody suffers at some time from physical pain caused by headache, toothache, traumas, injuries or surgical intervention. Acute pain accompanies many serious chronic illnesses. By means of pain our central nerve system signals about misbalances in one or another body organ. Since olden days, to relieve pain people used remedies from nature such as morphine, codeine and other opioids. However, due to their well known narcotic consequences, many synthetic drugs, called analgesics and anesthetics, have been developed. They are divided into opioid and non-opioid classes. The action of the former is manifested through the opioid receptors located in the CNS. They block the transfer of pain impulses and suppress the centers of pain perception in the cerebral cortex. These anesthetics include many derivatives of 4-piperidine carboxylic acid, 4-piperidol and 4-aminopiperidine (Figure 7.30a). Their first representatives (meperidine, trimeperidine, prodine) appeared as far back as the 1950s. The modern group of opioids known as fentanyls demonstrates especially high potency. Thus, parent fentanyl, which is 81 times more potent than morphine, is currently the most widely used preparation for anesthesia and analgesia. Its close analogue carfentanyl is so highly potent (10 000 times that of morphine) that it is intended for large animal use only. Episodes of such use can be seen at times on the Animal Planet Channel.

The non-opioid heterocyclic analgesics amidopyrine, antipyrine, analgin (metamizole sodium) and others (Figure 7.30b) have a similar but milder effect. These pyrazolone preparations have long been used in the treatment of headaches and neurogenic, muscular and articular pain. A rare, if not the only, example of a thiophene-based anestethic is articane, which was first introduced in

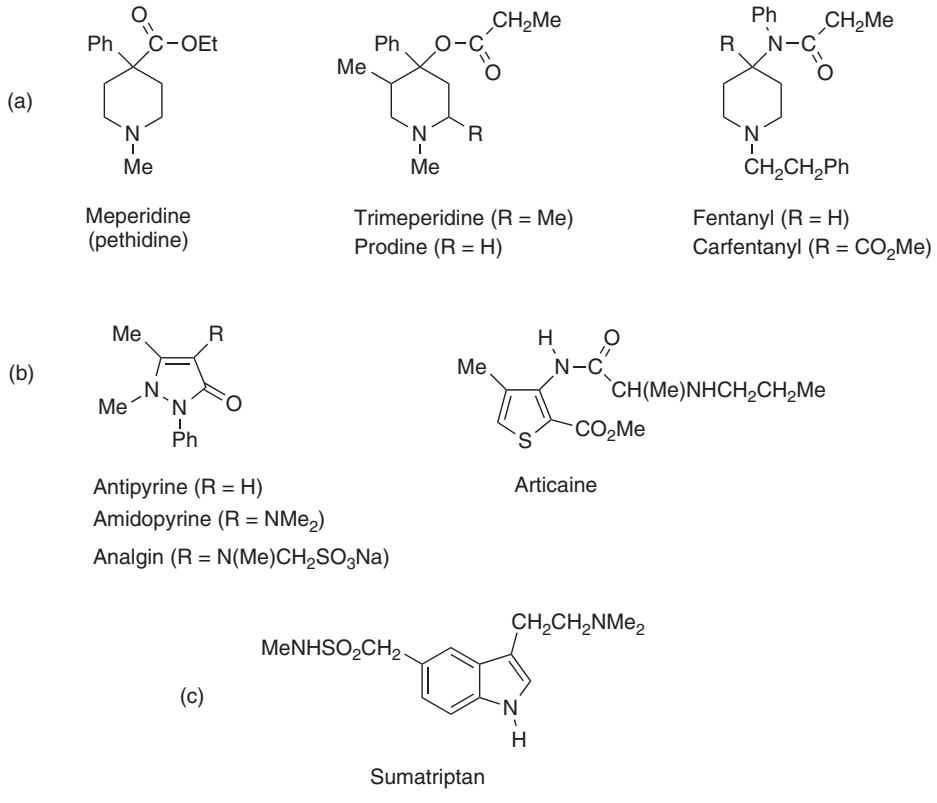


Figure 7.30 *Synthetic opioid (a) and non-opioid (b) analgesics and (c) antimigraine drug.*

clinical practice in the 1990s. Articane is widely used as a dental local anesthetic and, due to its fast action, is preferable over novocaine and lidocaine.

It is worth mentioning one more group of substances which is used to combat migraine headache. This disease is widespread and affects approximately 15% of the world's population. A majority of the recently introduced antimigraine compounds belongs to derivatives of triptamine and sumatriptan and seem to be the most popular (Figure 7.30c). It is believed that triptans selectively activate serotonin receptors (as agonists) and vasoconstrict the excessively dilated intracranial, extracerebral blood arteries under migraine. This mechanism shows that the preparations are unlikely to act as direct analgesics.

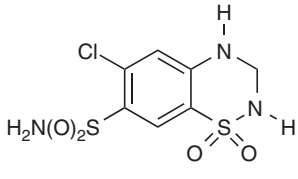
7.4.2 Heterocycles and Cardiovascular Diseases

High arterial blood pressure (hypertension), coronary spasms (stenocardia) and cardiac arrhythmia are typical manifestations of cardiovascular disorders. The drugs employed for the treatment of these diseases are called antihypertensive, antianginal (vasodilatory, antispasmodic) and antiarrhythmic agents, respectively. To begin, we highlight three features of cardiovascular agents. First, their action is often directed toward the peripheral or central nervous system. Such nervous stimulation in turn signals the appropriate regulatory mechanism to correct the abnormal deviation. Thus, the administration of hypnotic and tranquilizing drugs, such as the alkaloid

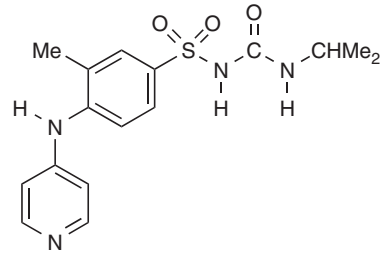
reserpine (Figure 7.4), also helps to normalize arterial blood pressure. Second, cardiovascular agents often produce more than one biological effect simultaneously (e.g., one preparation may halt cardiac spasms and concurrently cause a reduction in blood pressure). Third, the numerous cardiovascular drugs available belong to many different classes of organic compounds.

Taking into consideration the aims of the present book, we focus primarily on the heterocyclic cardiovascular drugs. Many heterocyclic classes are represented: azines and azoles, heteroaromatic and partially hydrogenated compounds, monocyclic and polycyclic and so on.

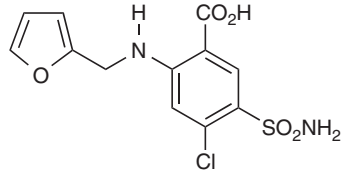
Normally, treating hypertension starts with diuretics. These decrease the amount of liquid and salts in blood vessels and thus remove the extra pressure on their walls. There are many heterocyclic derivatives among diuretics; a few are hydrochlorothiazide, torsemidet, furosemide (Figure 7.31).



Hydrochlorothiazide



Torsemidet



Furosemide (lasix)

Figure 7.31 *Some diuretics as the first-line drugs in treating hypertension.*

Another defense line against hypertension includes the so-called adrenergic receptor blockers. Alpha- and beta-adrenergic receptors are a special kind of protein which serve as a target for adrenaline, noradrenaline and to a lesser extent dopamine. Binding the catecholamines to these receptors causes a narrowing of blood vessels, increasing heart rate and as a consequence a growing blood pressure. Alpha- and beta-blockers concurrent with the natural ligands for these receptors prevent hypertension. Doxazosin and prazosin, both derivatives of 4-aminoquinazoline, are typical representatives of alpha-blockers (Figure 7.32). Interestingly, prazosin is also known as the best medication for managing severe scorpion stings. A more complex but somewhat related mechanism of activity is displayed by another heterocyclic amine, clonidine.

1-Hydrazinophthalazine (hydralazine; Figure 7.32) which exhibits a marked antihypertensive activity is a direct-acting vasodilator. This category of drugs relaxes muscles surrounding blood vessels, which results in their widening and thus lowering the blood pressure. The vasodilatory effect of alkaloids such as papaverine, theophylline, theobromine and caffeine (Figures 7.1, 7.2)

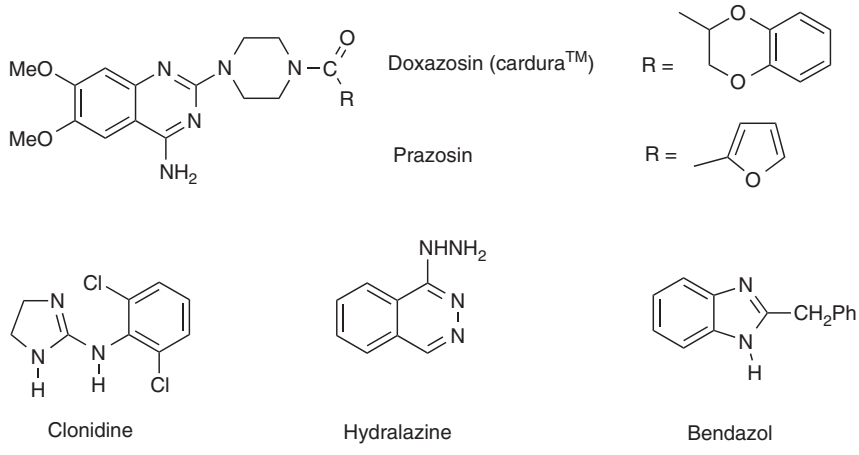


Figure 7.32 *Some adrenergic blocking agents and vasodilators.*

is also well known. A mild hypotensive effect[4] is observed in the case of 2-benzylbenzimidazole (bendazol), which is often used in combination with papaverine and other similar compounds. Interestingly, bendazol also has adaptogenic activity, that is, the ability to enhance the resistance of an organism toward unfavorable influences such as catarrhal infections.

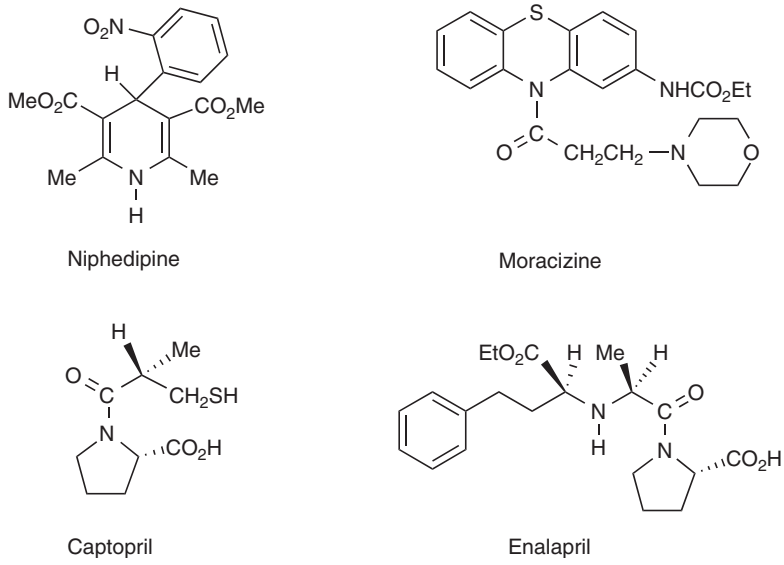


Figure 7.33 *Examples of calcium channel blockers and ACE inhibitors.*

[4] One should distinguish between 'hypertensive' and 'hypotensive' action. The former signifies an increase in blood pressure, whereas the latter indicates a reduction of either high or normal blood pressure.

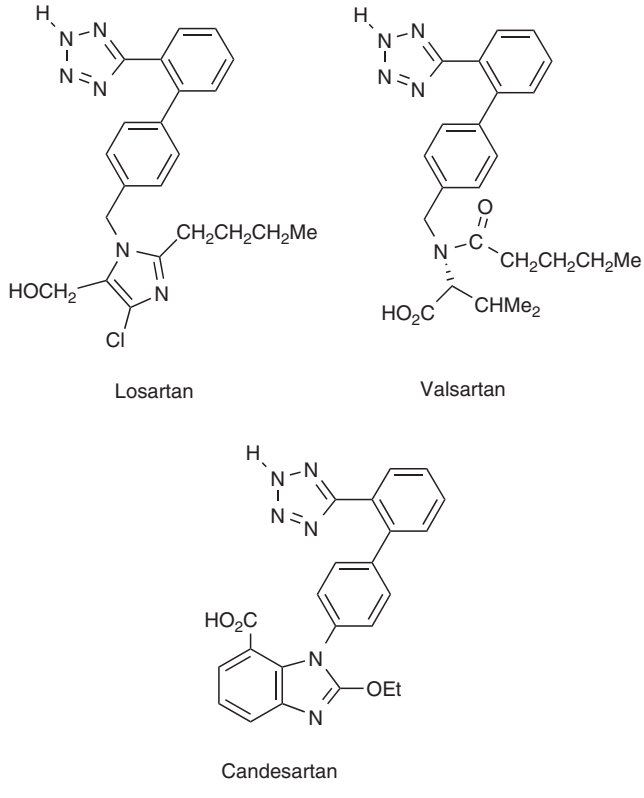


Figure 7.34 *Antihypertensive drugs of the tetrazole series.*

In the past 30 years, derivatives of 1,4-dihydropyridine have been intensively investigated because of their effectiveness in the therapy of hypertension and stenocardial seizures. One of the best known preparations of the series is nifedipine (Figure 7.33). The biological action of 1,4-dihydropyridines is via the inhibition of calcium channels. Calcium ions control a multitude of intracellular processes. In particular, calcium stimulates activity of the cardiac muscle (myocardium). However, under myocardial ischemia or infarction, cardiac function needs to be facilitated and the demand of the heart muscles for oxygen must be reduced to limit metabolism and the destruction of cell walls. These demands are met by nifedipine and its analogues, which block the channels that enable calcium transport into the cell.

Some heterocyclic compounds are employed as strong antiarrhythmic agents. The alkaloid quinidine (see Section 7.1) and the phenothiazine derivative moracizine are two examples. Captopril and enalapril (Figure 7.33) act via inhibition of angiotensin-converting enzyme (ACE) which catalyzes the formation of angiotensin II, a powerful endogenous vasoconstrictor. These drugs are thus used as antihypertensives and as adjuncts in the therapy of heart failure. Another relatively novel group of the angiotensin II receptor blockers consists of 2-diaryltetrazole derivatives – losartan, valsartan and candesartan (Figure 7.34). Currently, they are widely used in the treatment of hypertension and chronic cordial disorders.

Today, medications that bring down cholesterol and triglycerides levels in blood have become very popular. Practically all of them belong to the so-called statins and zocor, atorvastatin

(lipitor) and rosuvastatin are their especially successful representatives (Figure 7.35a). They decrease significantly the risk of myocard infarctions and the development of atherosclerosis. The drugs of this class are widely used in prophylaxis and curing of ischemic diseases of the heart, hypolipidemia and atherosclerosis. Reports have appeared as well on reducing the risk of insult in patients with sugar diabetes treated with lipitor. The bioaction of statins is connected with the effective inhibition of HMG-CoA reductase, which controls cholesterol and triglycerides biosynthesis. This NADP-dependent enzyme catalyzes reduction of β-hydroxy-β-methylglutaryl-CoA into mevalonic acid (Figure 7.35b) which is a key precursor in the biosynthesis of terpenes and steroids including cholesterol. The worldwide sale of statins has reached more than 20 billion dollars; lipitor brings about 60% of this and is currently blockbuster drug number one.[5] It is of interest that the total annual market for agents for the treatment of heart diseases has reached 70 billion dollars and that for anticancer drugs (see next section) is 55 billion dollars.

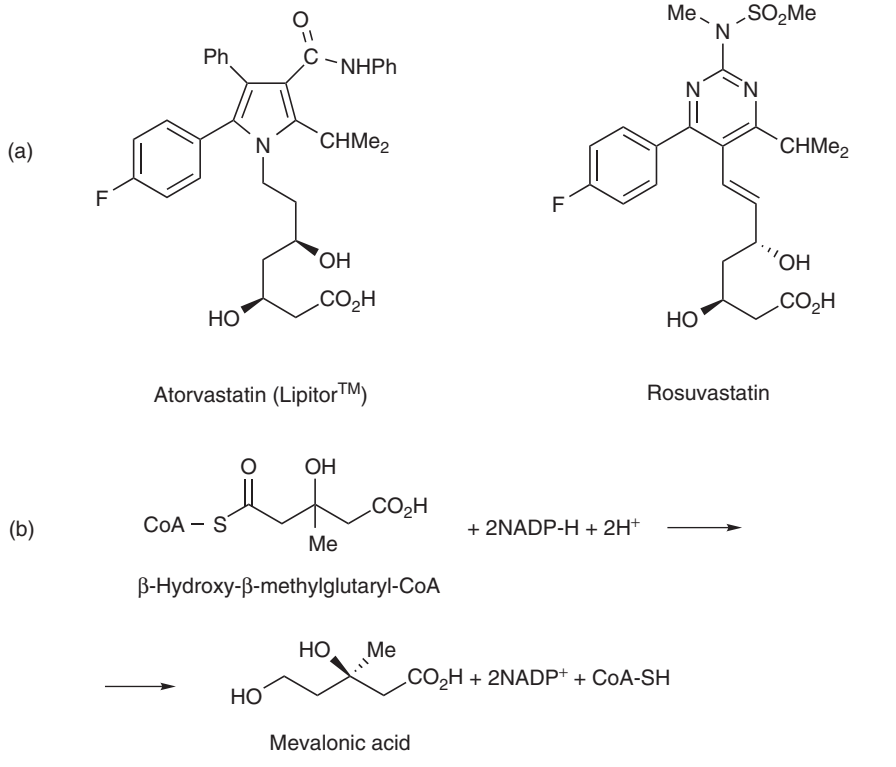


Figure 7.35 *Representatives of statin drugs (a) and the reaction (b) they target in cholesterol biosynthesis.*

7.4.3 Heterocycles and Malignant Tumors

The fight against cancer has been a principal focus for several decades. The search for a cure for cancer is of the utmost social and economic importance. Clinically effective therapeutic procedures

[5] The name 'blockbuster drug' is given to medicines whose worldwide annual sales officially exceed one billion dollars.

are being intensively sought worldwide. Regretfully, the announcement of a medicinal revolution in this field of chemotherapy would be premature. Nevertheless, the chemotherapy of tumors and especially the use of chemicals in conjunction with other methods of treatment already can significantly prolong the lifespans of many patients. In some cases, chemotherapy is able to cure the patient completely, restoring them to health and to at least their natural 'three score years and ten'.

The principal challenge in the chemotherapy of cancer lies in discovering a means to discriminate between cancerous and healthy (normal) cells. The main differences between these types of cells lie in the rate of DNA synthesis and replication and also in the rate of cell division: cancerous cells can divide 10^6 times faster than normal cells. These differences have been the focus of strategies aimed at developing a cure for the disease. Current approaches rely on the following considerations. Since tumorous cells divide more rapidly, they have an increased requirement for purine and pyrimidine nucleotides (for DNA synthesis) compared with normal cells. If we could prevent access of nucleotides to the diseased tissues, then the malignant cells would be prevented from reproducing.

This concept has led to the search for anticancer agents among the derivatives of purine and pyrimidine and to some success. Four synthetic anticancer drugs which are currently among the most clinically effective have so far resulted from this quest: 6-mercaptopurine, methotrexate, 5-fluorouracil and ftorafurum (Figure 7.36).

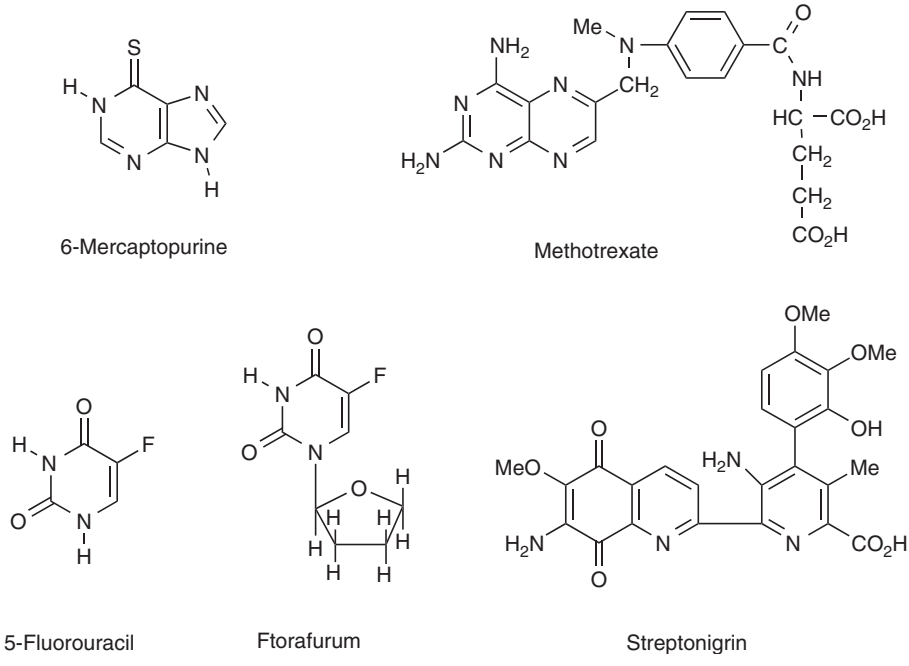


Figure 7.36 *Anticancer drugs interfering with DNA synthesis.*

6-Mercaptopurine (shown in Figure 7.36 as the dominant thione tautomer) and methotrexate are effective in the treatment of leukemia, while 5-fluorouracil and ftorafurum are used in cases of ventricular, intestinal, ovarian, pancreatic and mammary glandular tumors. The antimetabolite concept accounts for the bioactivity of all four substances. Mercaptopurine actively interferes

with DNA processes by its structural similarity to adenine and hypoxanthine. Methotrexate and fluorouracil inhibit the biosynthesis of thymidine and thereby reduce its content within cells. Thymidine is an essential nucleoside for the construction of DNA. The critical stage of thymidine synthesis appears to be the methylation of deoxyuridine monophosphate at the 5-position with the assistance of the enzyme thymidylate synthetase and a methylene adduct of tetrahydrofolic acid as coenzyme (Figure 4.35). In the course of this reaction a molecule of dihydrofolic acid is formed. For the process to be reversible, dihydrofolic acid must be hydrogenated to the tetrahydrofolic form. The enzyme dihydrofolate reductase catalyzes this transformation. It is at this point that methotrexate becomes involved. Because of the similarity to dihydrofolic acid, methotrexate inhibits dihydrofolate reductase, thus inhibiting the biosynthesis of thymidine.

Fluorouracil and ftorafurum, once delivered to the malignant cells, are converted into the 5-fluoro derivative of deoxyuridine monophosphate. Because of their resemblance to deoxyuridine monophosphate (Figure 3.3), they block the action of thymidylate synthetase and therefore also delay thymidine biosynthesis.

There are many naturally occurring heterocyclic compounds, especially among the alkaloids and antibiotics, which display anticancer activity. For instance, the antibiotic streptonigrin, containing both a pyridine and quinoline nucleus with a rather intricate substitution pattern (Figure 7.36), inhibits DNA formation and is used in the clinical therapy of lymphogranulomatosis and lymphoid leukosis.

However, a serious drawback of most existing anticancer drugs is their lack of specificity of action. As a consequence, most chemotherapeutic agents are also highly toxic toward normal cells. Thus, a more effective future treatment of malignant tumors with fewer side effects necessitates a new approach to circumvent the problem of nonspecificity. One such promising method called photodynamic therapy (PDT) was reported in the late 1980s. Primarily, its strategy was based on the use of the biologically innocuous compound 8-methoxypsoralen (or simply psoralen; Figure 7.37), which is injected into the bloodstream of patients suffering from such lethal forms of cancer as T-cell lymphoma. After a period of time, 500 ml of blood is taken and the malignant T-lymphocytes are separated from the other components of the blood by centrifugation. The separated T-lymphocytes are then suspended in solution under physiological conditions and subjected to ultraviolet (UV) radiation. The treated lymphocytes are then resuspended in the original plasma and the mixture is injected back into the patient. The process is repeated six or seven times over a certain period and can result in a dramatic amelioration in the health of the patient.

To understand what occurs at the molecular level during such treatment, we first consider the structure of methoxypsoralen. Owing to its planar geometry, this heterocyclic molecule can intercalate between neighboring pairs of DNA bases in a parallel orientation with respect to the rings (see Figure 7.37a and Section 3.5). The ability to be activated by UV light is the second characteristic of the psoralen molecule. In the UV-activated state, the carbon–carbon double bonds of the furan and α-pyran rings are highly reactive, and induce the C—5–C—6 bonds of the thymine residues to add to the double bonds of the intercalated psoralen molecules. As a result of the photochemical reaction both DNA chains become strongly crosslinked (Figure 7.37b). These crosslinks prevent DNA replication and, consequently, malignant cell growth and division. This procedure, called photophoresis, avoids any contact of the normal cells with the photoexcited drug molecules. In the non-excited state, psoralens are perfectly harmless and are in fact found in fruits and vegetables such as figs, limes and parsnips.

In a further development of the PDT, another class of organic compounds, called dye-photosensitizers (see Section 11.1.3) has been suggested together with a particular type of light to kill tumor cells. The treatment works as follows. The dye is injected either directly into tumor tissue or into the bloodstream. In the latter case the agent after some time (2–3 days) is selectively accumulated in cancer cells. Then, the tumor is irradiated with visible low energetic

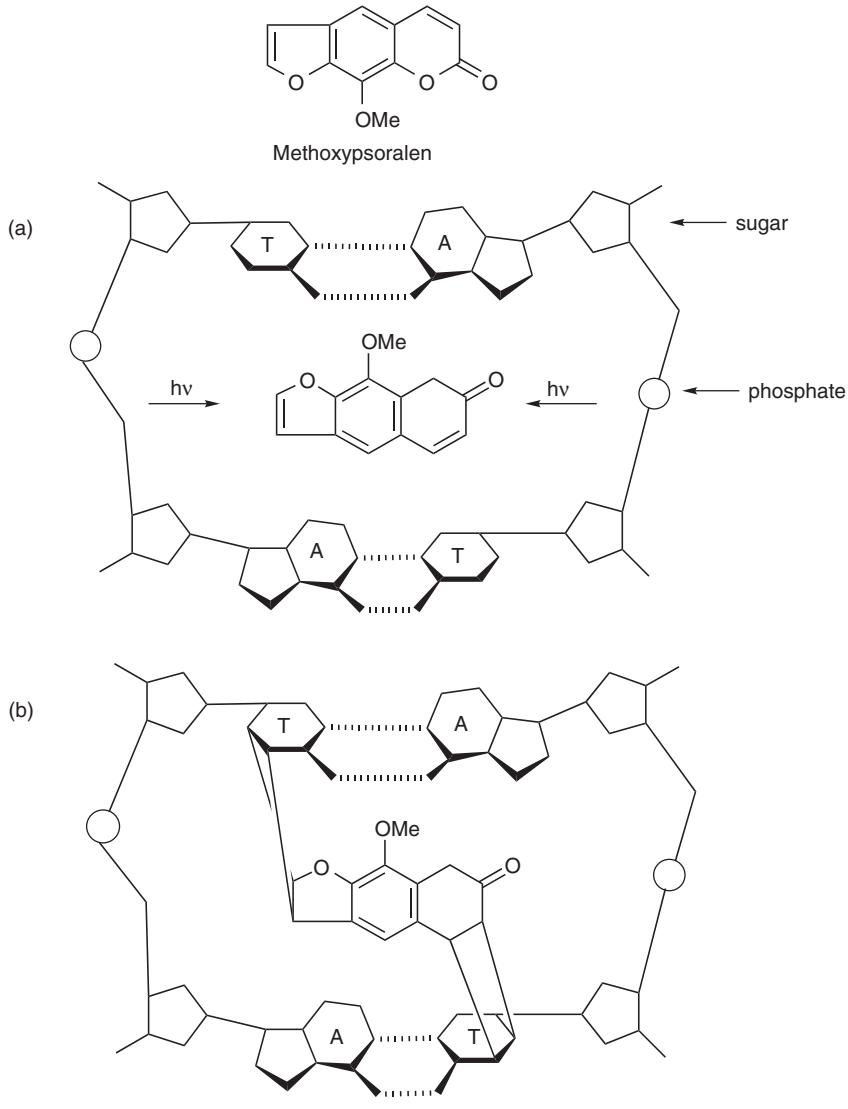
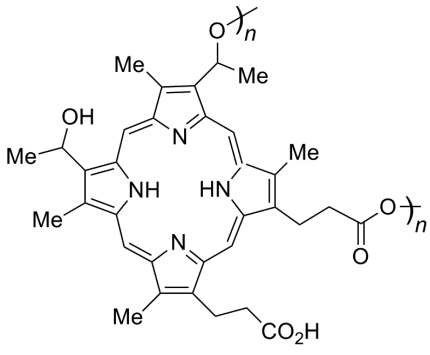


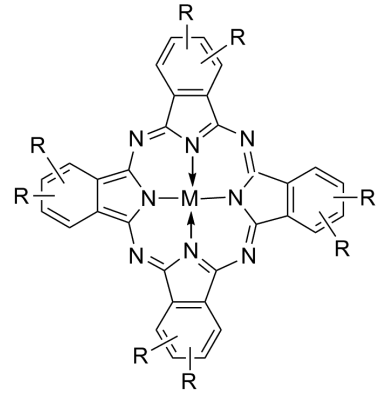
Figure 7.37 *Photointeraction of the intercalated 8-methoxypsoralen molecule with thymine residues in a DNA chain: (a) before irradiation and (b) after irradiation (A = adenine, T = thymine). Adapted from Edelson, R. L.,* Sci. Am.*, 1988,* ***259****, 68, with permission. © 1988 Scientific American Inc., George V. Kelvin.*

light (630–800 nm) in the presence of oxygen. Under these conditions the photosensitizer produces toxic singlet molecular oxygen which destroys tumor cells (see Section 4.2 and Figure 4.11). The first generation of photosensitizers for PDT is composed of oligomeric porphyrins, for example, hematoporphyrins (Figure 7.38a). The second generation consists of chlorines, bacteriochlorins (dihydro and tetrahydro derivatives of hematoporphyrins) and phthalocyanines, such as photosense and teraphthal (Figure 7.38b).



Hematoporphirins (n = 3-6)

(a)



Aluminium sulfophthalocyanine Photosense
(M = Al, R = SO_3Na)

Teraphthal (M = Co^{2+}, R = CO_2Na)

(b)

Figure 7.38 *Some photosensitizers in the photodynamic therapy of tumors.*

PDT becomes more effective under the combined administration of a photosensitizer with ascorbic acid. For example, teraphthal enhances the catalytic oxidation of vitamin C in a cell into ascorbat radicals, dehydroascorbic acid, hydrogen peroxide, hydroxyl radical and superoxide radical of oxygen, which brings about an apoptosis-driven destruction of the tumor cell.

PDT has many advantages. Thus, unlike chemotherapy, PDT can be localized; it takes very little time and it has no long-term side effects. As a whole, PDT is much cheaper and less harmful than surgery or radiotherapy. The limitation of PDT is that the light needed for activation of a photosensitizer cannot pass through more than 1 cm of tissue. Hence, PDT is mostly used for the treatment of skin cancer. However, a great deal of research and clinical studies is now underway to cancel this limitation (e.g., see Section 11.3.3).

The approach utilizing intercalating drugs for highly specific biological applications has been pursued by scientists with such success in the past few years that it has become an avenue for the creation of novel and versatile medicines. Polynuclear dyes of the acridinium salt type and methylene blue, which are used as antibacterial agents (Figure 7.5), are classic examples of intercalators. However, to cause a disruption of nucleic acid synthesis, an intercalator must become covalently linked with the DNA chain (refer back to Section 3.5).

Since anticancer drugs inhibiting DNA synthesis suffer from considerable toxicity towards normal cells, scientists are intensively searching for other biological targets. Most promising among them are various signaling and regulating cell proteins. Thus, a special class of proteins should be heavily phosphorylated before cell division starts. ATP displays itself in these reactions as a phosphorylated agent for the OH groups of serine and tyrosine residues. The process is catalyzed by the enzyme tyrosine kinase and other related kinases. Since 2000 a series of highly effective and low-toxity drugs were introduced into clinical practice, which selectively enter the catalytic cleft of tyrosine kinases and thus prevent cell division, mainly that of malignant cells. All these medications, for example, imatinib, sunitinib and bortezomib (Figure 7.39) are derivatives of nitrogen heterocycles, both aromatic and non-aromatic, π-deficient and π-excessive. They are used to treat leukemia, gastrointestinal stromal tumor, renal cell carcinoma, lymphoma and lung cancer.

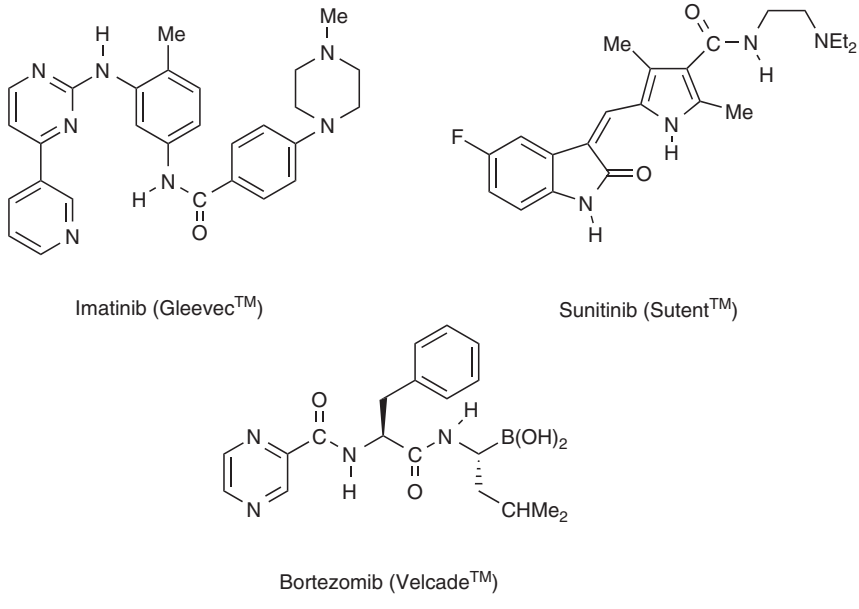


Figure 7.39 *Selected protein-targeted anticancer drugs.*

At the present time several innovative methods of anticancer therapy are being developed. To a considerable degree they are also connected with the use of heterocyclic compounds. They include methods of gene therapy, delivery of drugs by means of nanoparticles, use of short RNAs, manipulations with telomeres to restrict the number of cell divisions and so on (see Section 11.3.3).

7.5 Heterocyclic Molecules in Combat with Ulcers and Sexual Dysfunctions

Gastric ulcers are mainly caused by a misbalance in the production of gastric hydrochloric acid that leads to erosion of the stomach wall. This is a widespread illness and unsurprisingly a great need for anti-ulcer drugs exists. Notably, the first blockbuster drug in history, as far back as the 1980s, was cimetidine – an anti-ulcer preparation of an imidazole (Figure 7.40b). There are two biochemical approaches to decrease gastric juice secretion and thus to prevent ulcer formation. The first one blocks the so-called histamine H_2 receptors. Histamine or 4(5)-(2-aminoethyl)imidazole is an important neurotransmitter which is formed via decarboxylation of histidine (Figure 7.40a). Histamine can be found in each body cell where it fulfills a number of key biological functions. One of them is binding to the above-mentioned H_2 receptors that initiate a complex process of hydrochloric acid release into the stomach. Cimetidine and the even more effective furan preparation ranitidine display the property of histidine antagonists and concurrently bind to H_2 receptors.

The second approach to cure an ulcer is based on inhibition of the enzyme H^+/K^+ ATPase, which is responsible for the exit of protons into the gastrointestinal tract (proton pump function). Omeprazole (Figure 7.40c) with annual sales of six billion dollars remains the most popular and

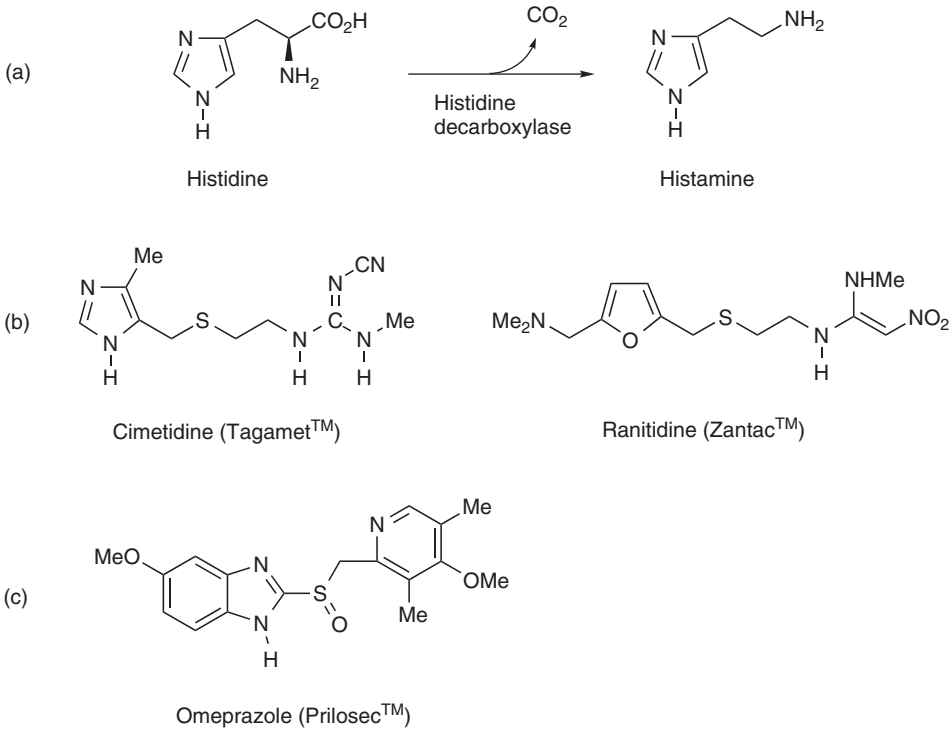
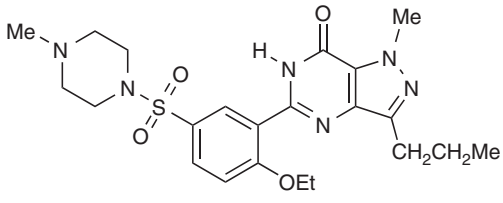


Figure 7.40 *The formation of neuromediator histamine (a) and examples of anti-ulcer drugs (b, c).*

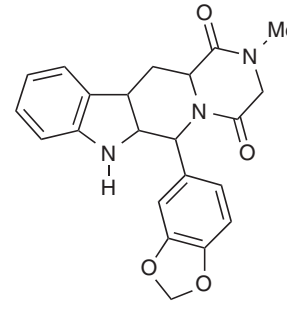
effective among this class of drugs. At present both the racemic omeprazole and its more active (S)-enantiomer (esomeprazole or nexium) are being marketed.

At the end of the twentieth and the beginning of the twentyfirst century the very first drugs efficacious in limiting sexual dysfunctions (mostly men's erectile disorders) were launched in the pharmaceutical market: Viagra and Cialis (Figure 7.41).

Viagra rapidly became very popular and in 2002 sales in the world market reached 1.7 billion dollars. Viagra enhances the relaxing action of nitrogen monoxide on guanilate cyclase and, consequently, on the smooth muscles of penis vessels. This sharply intensifies blood circulation and, hence, erection. By normalizing the erectile function, drugs of this type inhibit phosphodiesterases and thus increase the cyclic guanosine monophosphate content which controls the muscles relaxation mechanism. A patent has appeared for the production of a chewing gum which contains Viagra as a biologically active additive (as 5–100 mg of its citrate). Such a gum is recommended to be chewed for at least 2 min and no more than 0.5 h before the sexual act. It is hoped that such a manner of administration of the drug would lessen the probability of gastrointestinal disorders. Cialis is ten times more potent than Viagra despite having the same bioaction mechanism. This preparation causes a stable erection within 15 min after administration and prolongs it up to 4 h with a high degree of probability (86%). But tadalafil can have side effects, such as headache, runny nose and dyspepsia.



Sildenafil (Viagra™)

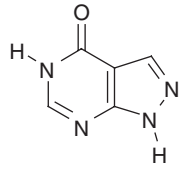


Tadalafil (Cialis™)

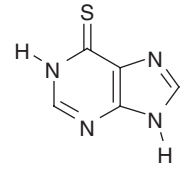
Figure 7.41 *Synthetic drugs for the treatment of erectile dysfunction.*

7.6 Problems

1. Acetanilide ($pK_a = 0.3$) and caffeine ($pK_a = 0.5$) are absorbed in the stomach at a rate of 30% per hour. In contrast, quinine ($pK_a = 8.4$) is absorbed slowly. Account for this difference.
2. What is the mechanism of the antibacterial effect of sulfa drugs? Of acridinium salts?
3. 6-Aminopenicillanic acid can presumably be formed from two naturally occurring amino acids. Name them and draw their structures, making the genetic linkage obvious.
4. Acute intoxication with phenobarbital ($pK_a = 7.2$) can be diminished by a factor of 15 by intravenous introduction of an aqueous bicarbonate solution, which increases blood and urine pH from the normal values of 7.4 and 6.0, respectively, to 8.0. Describe the chemical mechanism for the removal of this barbiturate.
5. Uric acid is usually removed from the body in urine. However, when high levels are present, it cannot be fully removed in this manner owing to its poor solubility (see Problem 11, Chapter 2). As a result, the acid begins to crystallize in joints, causing gout. Pyrazolo[3,4-*d*]pyrimidine-4-one, or allopurinol (A), is one of the most effective treatments for gout. Suggest a mechanism for the therapeutic action of allopurinol, taking into account the fact that it can also prolong the lifetime of 6-mercaptopurine (B), which is used in the treatment of leukemia (also consider the conditions described in Problem 3, Chapter 4).

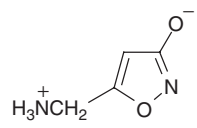


A

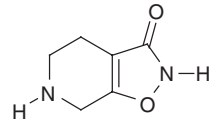


B

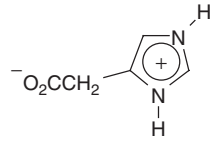
6. 5-Bromouracil and 5-fluorouracil, when introduced into organisms as markers, are incorporated differently: one into DNA, the other into tRNA. Which of the two uracils will bind to DNA? To tRNA? Explain your reasoning.
7. The five substances C–G promote or inhibit the bioaction of the naturally occurring mammalian neurotransmitter X.
   Muscimol (C) is found in mushrooms of the genus *Amanita muscaria*. It is a functional analogue (agonist) of the neurotransmitter in question. Compound D is a competitive inhibitor of the neurotransmitter in rat brain preparations. 4(5)-Imidazolylacetic acid (E) is an agonist of X. Compound F is a drug called pyracetam which exerts a nootropic (cognitive-enhancing) effect and mimics the effects of X. The alkaloid bicuculline (G) selectively antagonizes the inhibitory action of X.
   (a) Using certain structural similarities between compounds C–G and the unknown neurotransmitter, suggest a structure for X.
   (b) Determine, on the basis of Dreiding stereomodels, the distance between the ends of the pharmacophore group.



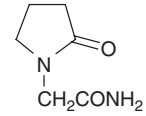
C



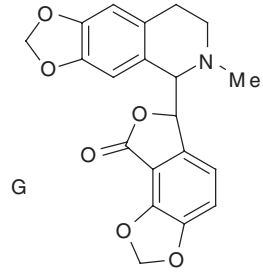
D



E



F



G

8. Macrolide antibiotics, such as azitromycin (Figure 7.14), are rather unstable to gastric acid. This is why they are normally administered as enteric coated tablets. Explain this fact.
9. The preservation of the analgesic action among a great number of 4-substituted piperidines (Figure 7.30a), synthesized on the basis of simplifying the structure of the natural alkaloid morphine (Figure 7.3), opens the possibility of finding a rule which could help design many other structures with a high probability of manifestating opioid type analgesia.
   (a) Determine a structural fragment (pharmacophore) responsible for the analgesic activity of morphine.
   (b) Formulate a rule with at least three main structural features which together could provide a designated structure explaining the analgetic properties of the opioid. Suggest one or two generalized structures.

7.7 Suggested Reading

1. Hoeprich, P. D., Jordan, M. C. and Ronald, A. R. (eds), *Infectious Diseases*, 5th edn, Lippincott, Philadelphia, 1994.
2. Lancini, G. and Parenti, F., *Antibiotics, an Integrated View*, Springer, New York, 1982.
3. Franklin, T. J. and Snow, G. A., *Biochemistry of Antimicrobial Action*, Chapman and Hall, London, 1981.
4. Moberg, C. L. and Cohn, Z. A. (eds), *Launching the Antibiotic Era. Personal Accounts of the Discovery and Use of the First Antibiotics*, Rockefeller University Press, New York, 1990.
5. Silverman, R. B., *The Organic Chemistry of Drug Design and Drug Action*, 2nd edn, Elsevier, Amsterdam, 2004.

6. Wermuth, C. G., Koga, N., Konig, H. and Metcalf, B. W. (eds), *Medicinal Chemistry for the 21st Century*, Blackwell, Oxford, 1992.
7. Thomas, G., *Fundamentals of Medicinal Chemistry*, John Wiley & Sons, Ltd, Chichester, 2004.
8. Gilman, A. G., Rall, R. W., Nies, A. S. and Taylor, P. (eds), *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 8th edn, Pergamon, New York, 1990.
9. Maxwell, R. A. and Eckhardt, S. B., *Drug Discovery: A Casebook and Analysis*, Humana, Clifton, N.J., 1990.
10. Negwer, M., *Organic-chemical Drugs and Their Synonyms (An International Survey)*, vols 1–3, VCH, New York, 1987.
11. Albert, A., *Selective Toxicity: the Physico-chemical Basis of Therapy*, 7th edn, Chapman and Hall, London, 1985.
12. Corey, E. J., Czakó, B. and Kürti, L., *Molecules and Medicine*, John Wiley & Sons, Inc., Hoboken, N.J., 2007.
13. Nicolaou, K. C. and Montagnon, T., *Molecules that Changed the World*, Wiley-VCH, Weiheim, 2007.
14. Li, J. J., Johnson, D. S., Sliskovic, D. R. and Roth, B. D., *Contemporary Drug Synthesis*, John Wiley & Sons, Inc., Hoboken, N.J., 2004.

8

Heterocycles in Agriculture

Do we need much?
No: of bread – only one slice,
With it a drop of milk.
The skies provide the salt
And clouds of white silk.
V. Khlebnikov

According to his contemporaries, the Russian poet Velemir Khlebnikov was an ascetic writer who valued the spiritual aspects of existence. Modern medicine, which warns that overnutrition is harmful to our health, is thus consistent with his poem. However, malnutrition is an equally dangerous threat and today some 1.5×10^9 people worldwide suffer from diseases caused by shortages of food and parasitic infestations. The world population growth rate has reached approximately 1.6×10^6 per week with a total of 6.5×10^9 people exerting ever-increasing demands for food. Meeting these needs will require dramatic improvements in the productivity of modern agriculture.

Exacerbating these demographic problems are the losses of agricultural production due to pests such as rodents, insects, microorganisms and weeds. Almost half of all food intended for humans is consumed or spoiled by pests (30% prior to harvest, 20% during crop transport and storage). Biological methods of plant protection are of great potential importance, but chemical control is currently still the main approach utilized. Chemical pest control is carried out by compounds known as pesticides, which include herbicides (substances used to kill weeds), insecticides, fungicides (agents used against fungal pathogens) and rodenticides. Stimulators and regulators of plant growth and development are also of great importance.

The driving force for continued research in this field is the need to meet severe ecological demands necessitating the creation of more effective pesticides. The mechanism of pesticide action is closely associated with pest biology and involves the disruption of vital biological functions. We have seen how heterocyclic compounds participate in many biological systems, and therefore it is not surprising that many pesticides are heterocyclic derivatives.

Worldwide, approximately 4–5 million tonnes of pesticides are currently applied each year at an estimated cost of 30–35 billion dollars. Herbicides comprise 60% of the total usage of pesticides in plant protection.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

8.1 A Century of Chemical Warfare Against Weeds

Though the first herbicides were applied about 100 years ago, their widespread use commenced only after World War II. Today, more than 400 different types of herbicides are utilized in agriculture, one-third of which have a heterocyclic structure. Progress in this field is evident from the decrease in the effective doses of modern herbicides [only tens of grams per hectare (ha), where 1 ha ≡ 2.477 acres] compared with their earlier counterparts (tens of kilograms per hectare). Such progress would not have been possible without an understanding of the mode of herbicide action.

The modern classification of herbicides is based on the nature of the biological target of the agent and comprises three main groups: (i) compounds which prevent photosynthesis (antiphotosynthetics), (ii) chemicals which disturb the biosynthesis of, or destroy, chlorophyll and other photosensitive pigments within the cells and (iii) materials which inhibit the biosynthesis of essential amino acids or inhibit 'dark' metabolic reactions.

Heterocyclic compounds, especially azines and azoles, are frequently employed as herbicides. Derivatives of the triazinone group, for example, ethiozin and amethidione (Figure 8.1) are important representatives of the antiphoto-synthetics. These herbicides are used to combat weeds in winter wheat crops and are characterized by rather low rates of application (0.8–1.6 kg ha^{-1}). Structurally related to the previous group are pesticides containing a triazine ring (e.g., prometryn, propazine, atrazine) or a diazine ring (e.g., terbacil and chloridazone, derivatives of uracil and pyridazinone, respectively; Figure 8.1). These herbicides are applied in slightly higher doses (1.5–1.6 kg ha^{-1}), but their action is wide-ranging. Typical applications include the destruction of weeds in cotton, sugar beet, turnips, soya, peas and sunflower crops, and also in vineyards, berry plantations and orchards.

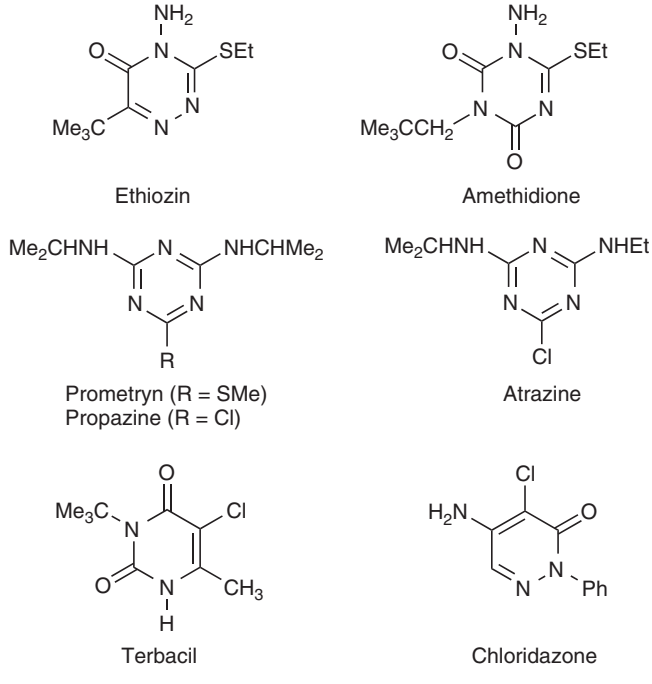


Figure 8.1 *Antiphotosynthetic herbicides of the triazine and diazine series.*

3-Amino-1,2,4-triazole, a herbicide of the azole series, has been in use for a relatively long time. Other examples from this same class include 2-trifluoromethyl-4,5-dichlorobenzimidazole (chlorflurazole) and a derivative of 1,3,4-thiadiazole named ethidimuron (Figure 8.2). Ethidimuron kills all plants and is thus used to clear roads, aerodromes, construction sites and areas surrounding high voltage transmission lines.

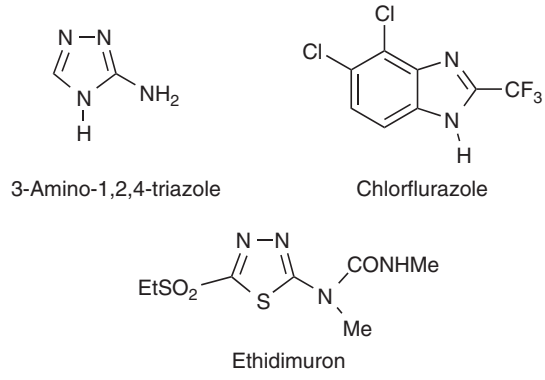


Figure 8.2 *Antiphotosynthetic herbicides of the azole type.*

All of the above-mentioned herbicides are antiphotosynthetics which interact with photosystem II (see Section 6.2). These agents bind to the membrane protein receptors of the photosynthetic complex which are in close proximity to the plastoquinone pool. As a result, electron transfer from the primary quinone acceptor to plastoquinone is interrupted. Some herbicides may disrupt the electron transport elsewhere in the electron transfer chain, but the result to the plant in both cases is the same: death.

All antiphotosynthetics contain several heteroatoms and acceptor groups. The acceptors diminish the reductive potential of the compound, and thus facilitate electron interception. Another characteristic feature of the heteroatoms and functional groups (especially C═O and NH) is their ability to bind with the amide and carbonyl functionalities of the protein receptor via hydrogen bonding and ion–dipole interactions. Herbicides which intervene with photosystem I also exist. Diquat dibromide and paraquat (methylviologen), which are quaternary salts of 2,2′-bipyridyl and 4,4′-bipyridyl, respectively, are typical representatives (Figure 8.3). They are believed to intercept electrons transmitted along photosystem I by forming stable cation-radicals. The cation-radicals reduce oxygen to the superoxide anion $O_2^{-\bullet}$ which induces the formation of hydrogen peroxide, hydroxyl radicals and singlet oxygen within the plant tissues. These species are all highly toxic and rapidly destroy plant pigments and other cellular structures. By transferring the electrons to oxygen, diquat dibromide and paraquat are regenerated. Thus, their mode of action involves electron transport.

Extensive use of antiphotosynthetic herbicides has triggered the development of mutant weeds which are resistant to these chemical agents. The gene that codes for the receptor protein to which the herbicide attaches undergoes mutations. A mutation in which only one heterocyclic base in the gene DNA is changed, leading to the substitution of a single amino acid in the polypeptide chain (e.g., serine by glycine), is sufficient to prevent adhesion of the herbicide molecule to the protein and thus to nullify its bioeffect. Thankfully, the choice of herbicides available today is sufficiently wide to allow effective control of the majority of weeds.

A serious problem remaining is that herbicides which disturb photosynthesis are still applied in comparatively high doses and can become dispersed in the environment, affecting nontargeted

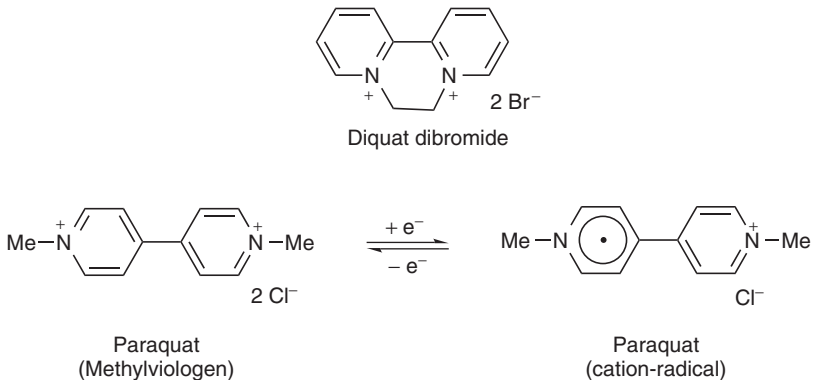


Figure 8.3 *Herbicides which interact with photosystem I.*

sectors with undersirable consequences. Therefore, since the late 1970s, interest in antiphotosynthetic herbicides has somewhat decreased. Scientists have now turned their efforts toward herbicides with a regulatory type of action. Such agents either interfere with the biosynthesis of chlorophyll, carotenoids or amino acids or change the phytohormone balance in the plant.

The biosyntheses of chlorophyll and the carotenoids in particular involve many steps. In general, disruption of any stage is detrimental to the plant. Inhibitors of photopigment biosynthesis include both heterocyclic and nonheterocyclic compounds. Diflufenican (Figure 8.4) is an example of a herbicide with such bioactivity which is used to eliminate weeds in many cereal crops. Diflufenican is valued for its low effective dose (65–250 g ha^{-1}), high selectivity and prolonged action

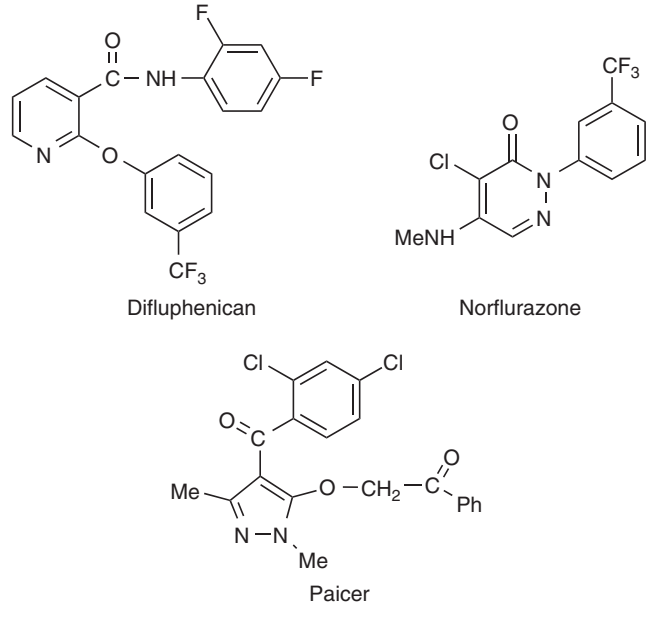
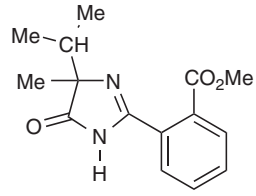


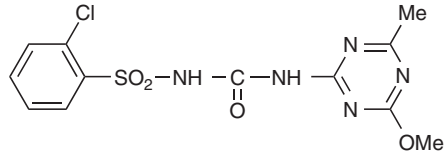
Figure 8.4 *Chlorophyll and carotenoid synthesis inhibitors.*

(a single treatment is effective for a whole season). The pyridazine herbicide norflurazone is a specific inhibitor of carotenoid biosynthesis. The pyrazole derivative Paicer (Figure 8.4) disrupts chlorophyll synthesis.

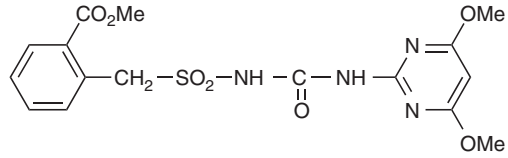
High herbicidal activity is observed for some imidazolin-5-one derivatives, such as imazamethabenz (Figure 8.5). Their structures are characterized by the presence of two alkyl groups at the 4-position of the imidazoline ring, with one of these groups being branched. Herbicides of this class act as antimetabolites by deactivating acetolactate synthetase, the main enzyme involved in the biosynthesis of amino acids with branched alkyl chains (valine, leucine, isoleucine).



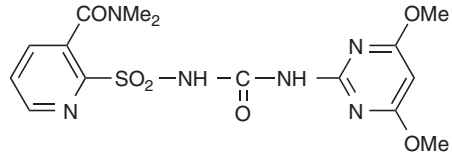
Imazamethabenz



Chlorsulfuron



Bensulfuron-methyl



Nicosulfuron

Figure 8.5 *Herbicides exerting phytohormonal effects and inhibiting amino acid synthesis.*

A recent generation of herbicides is the so-called sulfonylureas. Though their name does not reflect the presence of any heterocyclic moiety, at least one 1,3,5-triazine, pyrimidine or pyridine ring is usually present (Figure 8.5). Sulfonylureas are highly selective and are characterized by very low dose rates, which appear to be close to the theoretical limit. The optimal application rate of the most important herbicide of the series, chlorsulfuron, is 20 g ha^{-1}, but a dose as low as 5 g ha^{-1} is often effective. Chlorsulfuron kills all broad-leaved weeds in grain crops. Other sulfonylureas are applied to combat weeds in soya, cotton, sunflower and corn plantations. The pyrimidine bensulfuron-methyl (Figure 8.5) is utilized for the selective destruction of weeds in rice crops.

Sulfonylureas are unique in their high activity at such low dosages, and this suggests a phytohormonal mode of action. It has been established that these herbicides affect neither photosynthesis nor DNA synthesis, but suppress the growth and division of plant cells. It is of interest that sulfonylureas, when applied in very low concentrations, stimulate seed germination in some plants, and can also retard leaf aging and enhance biomass increase. These results also testify to their effect on plant hormonal systems.

It is worth noting that a massive usage of many herbicides can cause a toxic action on cultural plants especially in the case of rather persistent substances. To fight with this kind of nondesirable effects which could lead to losses of crops, there were synthesized special agrochemicals called safeners or antidotes. Safeners selectively enhance the activity of herbicide-detoxifying enzyme called glutathione-S-transferase in crop plants. One of them is fenclorim, a rather simple pyrimidine derivative (Figure 8.6) which is used in specific combinations with herbicides to protect crops such as corn, wheat and rice.

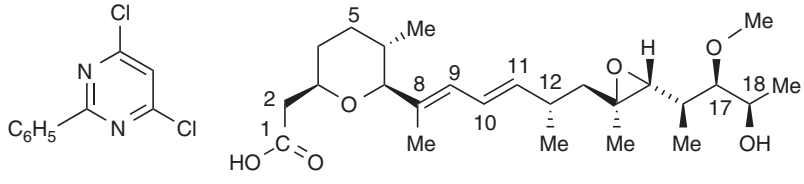


Figure 8.6 *The structures of a synthetic safener (a) and a natural herbicide herboxidiene (b).*

A natural polyketide named herboxidiene (Figure 8.6) was recently isolated as a product of actinomicet metabolism. It possesses a high herbicidal activity. Intensive structure–activity investigations (SAR) have been started to achieve its derivatives most suitable to industrial preparations.

8.2 Regulators of Plant Growth

The sulfonylureas are in fact plant growth and development regulators. The agricultural application of compounds with regulatory activity began about 60 years ago, and since that time their scope has increased. Now the chemical industry produces more than 70 growth regulators, one third of them being heterocyclic compounds. Their small percentage (5% by weight) in the overall production of pesticides is largely accounted for by their effectiveness at low concentrations.

For a long time plant growth regulators were used not so much to combat weeds but to create favorable conditions for the development of cultivated plants. The range of effects of such preparations is very wide. Plant regulators can accelerate or slow the growth, flowering or ripening of plants, and make crops more drought or frost resistant. Thus, consistently high harvests can be attained. It is now well established that plants themselves produce regulators which guarantee the punctual appearance 'on the seventh day of a bright, rounded and refined leaf decorating a sprout of an old black twig', as the old saying has it. Such naturally occurring compounds, called phytohormones, cause dramatic effects when present in minute quantities. Plant hormones are subdivided by their chemical structure and mode of action into three main groups: auxins, kinins and gibberellins, together with the single compound ethylene (Figure 8.7).

Indol-3-ylacetic acid (heteroauxin) is one of the chief growth hormones. This auxin diffuses readily along the plant stem, and its hormonal action results in the lengthening of plant cells

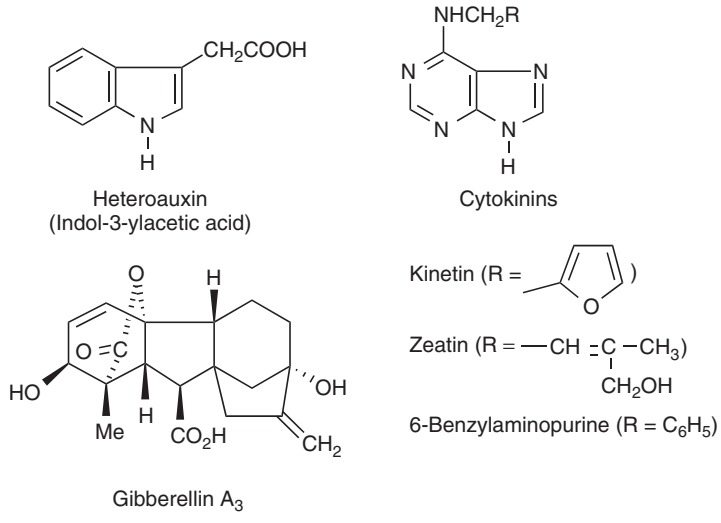


Figure 8.7 *Phytohormones.*

and stimulation of their division by increasing the rate of DNA replication. The acid induces the formation of side roots, stem sprouts and leaves, but inhibits plant growth when used in high concentrations. This acid also regulates and coordinates the growth of all plant components including the roots, stems and buds. Indol-3-ylacetic acid usually accumulates in the growing tissues of the plant: in the tips of buds and shoots, and in the young leaves and fruits. In aging tissues, synthesis of the hormone sharply decreases, and quite possibly it is this which triggers the loss of flowers, fruits and leaves. The same hormone is also found in mushrooms and some symbiotic plants. The practical uses of heteroauxin are extensive: the implantation of grafts during the vegetative propagation of fruit and berry cultures, the promotion of an increased number of side shoots in vegetables, the acceleration of ripening in fruits and the formation of seedless fruits. The main disadvantage of heteroauxin is its rapid degradation upon exposure to light. Interestingly, the synthetic homologue 4-(indol-3-yl)butyric acid is reasonably resistant toward the action of light. Recently the molecular mechanism of auxin bioactivity has been discovered (Tan, X. *et al*., *Nature*, 2007, **446**, 640).

Cytokinins are adenine derivatives substituted at the amino group. Kinetin and zeatin are their most important representatives (Figure 8.7). They enhance plant RNA synthesis and, consequently, the production of proteins in cells. Cytokinins also stimulate cell division and increase cell dimensions. In addition, cytokinins control the relationship between the roots and other parts of the plant, regulating hydration, adjusting to temperature changes and combating infections. Root tips and developing fruits are especially rich in cytokinins. 6-Benzylaminopurine is an example of a synthetic compound which shows a hormonal cytokine activity (Figure 8.7). It is not expensive, and is widely used as a growth regulator to enhance the stability of wheat and beets towards weather conditions, to stimulate branching of apple trees and roses, and to raise crop yields of melons, watermelons and pumpkins.

Some plant parasites have determined that the influx of nutrients increases at the sites of cytokinin action, and have then adapted to produce these hormones and introduce them into the plant to secure themselves an adequate food supply. The mechanism by which hormones of the cytokinin series stimulate the fission and growth of cells is unknown. Chlorsulfuron and other

sulfonylureas (Figure 8.5) possess cytokinin-like activity. However, when these compounds are applied to weeds in optimal concentrations, plant growth is suppressed by the hypercytokinase mechanism.

The gibberellins, the third class of phytohormones, are not always categorized as heterocyclic compounds despite the γ-lactone ring in their structure. Gibberellins initiate seed germination and stem growth, and promote an increase in fruit size. These compounds are secreted by cell nuclei and promote the formation of enzymes which degrade starch and seed membranes. Gibberellins also promote the synthesis of tryptophan which is converted into indolylacetic acid in sprout tips. Vines of seedless grapes are sprayed with gibberellins to increase the size both of the grapes and the overall bunches. Recently the biological target of gibberellin and the mechanism of its activity have been discovered (Murase, K., *et al., Nature*, 2008, **456**, 459).

The modern agricultural arsenal contains pesticides which allow plants to be 'immunized' against almost every imaginable pest and weed. However, compounds also exist which are capable of causing disorders in gibberellin biosynthesis, consequently stunting plant growth. They do not alter the ripening schedule and can help to achieve greater harvests by reducing the height of plants and therefore their tendency to collapse, especially in heavy rain, when harvesting of the crops can be very difficult, leading to heavy losses, if plants are blown down. Pesticides with such stunting action are known as retardants. Their typical examples are atrinal and uniconazole, a member of the 1,2,4-triazole series (Figure 8.8). Some retardants of this class are useful in regulating decorative plant cultures and lawn grass. All these retardants and analogues also have fungicidal properties and as such are used in agriculture.

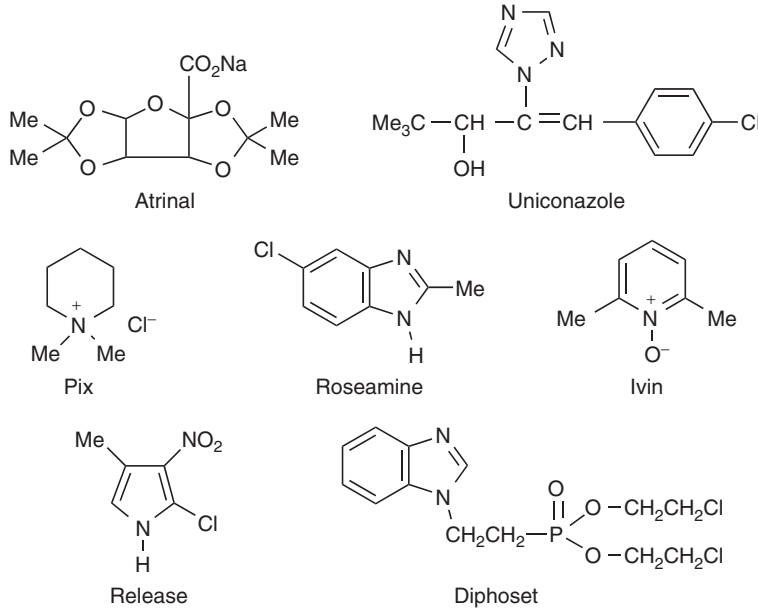


Figure 8.8 *Synthetic plant growth regulators.*

Many other synthetic regulators of plant growth and development have been prepared, each designed specifically to influence a particular biological process, such as activating or regulating photosynthesis, interfering with chlorophyll biosynthesis, stimulating plant respiration and so on.

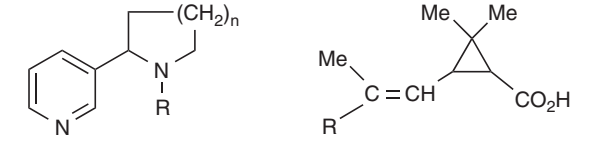
Thus, *N*,*N*-dimethylpiperidinium chloride (Pix; Figure 8.8) is employed to enhance the rate of cotton boll ripening. Roseamine, or 2-methyl-5(6)-chlorobenzimidazole, is useful in preventing cotton from becoming detached from the plant. 2,6-Dimethylpyridine 1-oxide (Ivin) stimulates the growth of tomatoes and cucumbers. A preparation named Release is successfully used to facilitate the harvesting of citrous plant fruits. The treatment of citrous plants with Release raises the level of such an endogenic phytohormone as ethylene, enhances the activity of cellulase and highly decreases the strength of attachment of a fruit to its graft. This allows an increase in the effectiveness of machine collection of fruits. It is of great importance that the flowers and unripened fruit are not affected by the synthetic regulator and that the latter demonstrates high activity when applied in very low concentrations. Diphoset (Figure 8.8) owes its plant regulating properties to 2-chloroethoxy and phosphonethyl groups which generate the phytohormone ethylene. This compound increases stability against stressing conditions in harvests of carrots, beets and cucumbers and thus ameliorates their nutrient qualities.

8.3 The Struggle Against Voracious Insects

More than 3×10^6 insect species inhabit the Earth. Of these, about 70 000 may be considered phytoparasites (or plant parasites). Such insects exact a high toll on agriculture as they devour and spoil up to 30×10^6 t year^{-1} of grain. The Colorado beetle devastates potato fields, the codling moth and other parasites damage orchards, and the barn weevil infests granaries. Locust swarms, sometimes occupying hundreds of square miles, are capable of destroying all crops. The arrival of these winged pests in huge numbers is, in some countries, equated with a national disaster. In ancient times these voracious creatures were referred to as 'hungerphorous' because their appearance *en masse* frequently caused human starvation.

The history of insecticides thus began long ago, and today more than 10^6 t (about 25% of the total production of pesticides) are applied annually in the treatment of agricultural and other terrestrial and domestic ecosystems. The famous 'insect powder' (known in Russia as 'Persian powder') can still be found in drugstores nowadays under the name 'pyrethrum'. This preparation is recommended for combating domestic insects such as ants, bugs, and fleas. Some success against agricultural pests has also been achieved.

The use of tobacco dust as an insecticide, especially against the plant louse and other small insects, has a long history. The pyridine alkaloids anabasine and nicotine, which affect the insect's CNS, are the active constituents (Figure 8.9a).



Anabasine (R = H, n = 2) Chrysanthemic acid (R = Me)

Nicotine (R = Me, n = 1) Pyrethric acid (R = CO_2H)

(a) (b)

Figure 8.9 *Components of naturally occurring insecticides from tobacco (a) and pyrethrum (b).*

Though nicotine and anabasine were the first heterocyclic compounds used in the chemical fight against pests, the importance of heterocycles in insecticide preparations is now not very

great. The current assortment of insecticides includes about 300 compounds and only 10% of them have heterocyclic structures. Modern insecticides are largely composed of aliphatic, alicyclic and homoaromatic compounds. In particular, rapid development in the chemistry of synthetic pyrethroid insecticides was triggered after the active components of pyrethrum were determined. The esters of two cyclopropane carboxylic acids, chrysanthemic and pyrethric (Figure 8.9b), were found to be responsible for the insecticidal activity. On the basis of the former, two very powerful insecticides (Figure 8.10) are chemically produced – neopinamine (80% *trans* isomer) and resmetrin [both as a mixture of *cis* and *trans* isomers and under the name cismetrin as a pure (1R) *cis* isomer]. These synthetic compounds are effective against domestic parasites, in the struggle with agricultural pests and in veterinary medicine, even on lactating cattle. The pyrethroids act by contact with insects or orally interfering with their nervous system and quickly causing paralysis ('knock-down' effect) and death of the pest. Piperonylbutoxide (Figure 8.10) enhances very much their effectiveness and therefore is widely used in mixtures with the pyrethroids.

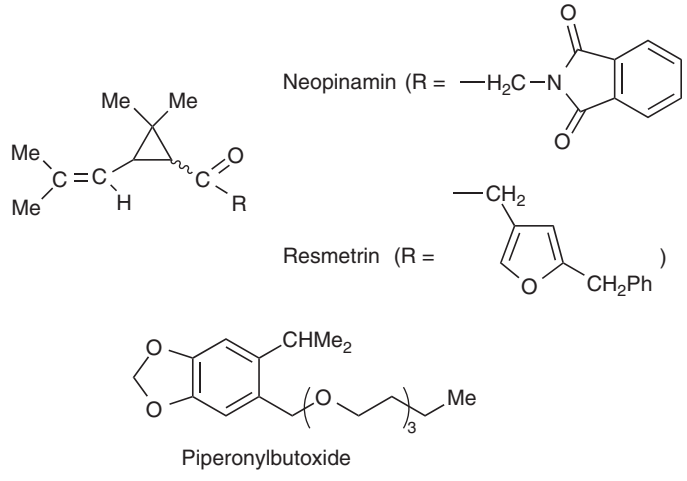


Figure 8.10 *Synthetic heterocyclic insecticides of the pyrethroid series and their synergist piperonylbutoxide.*

Many insecticides belong to the chlororganic series. A notorious example of this class is 2,2-di(4-chlorophenyl)-1,1,1-trichloroethane (DDT; Figure 8.11a), which is now banned in most countries. Recently a natural compound called epibatidine (Figure 8.11b) with insecticide effect has been separated from the skin of the frog *Epipedobatis tricolor*, which inhibits acetylcholinesterase in plant lice (aphis) and web tick organisms.[1] The aphicides such as acetamiprid and nitempyram and a systemic insecticide of another type, imidacloprid and several its analogues, were synthesized (Figure 8.11d, e). They also suppress that same enzyme but do not act on the nervous system of vertebrates. These are environmentally safe insecticides from a new group of antagonists of nicotineacetylcholine receptors causing hyperpolarization of membranes of neuronal fibers. They have two toxoforic groups, one of which includes the 2-chloro-3-aminomethylpyridine moiety. The

[1] Long ago, Colombian natives used extracts from frog skin to coat blow-darts for hunting because the frog skin is a ready source of many (about 500) poisonous alkaloids, which are needed for defense from not only insects but also carnivorous animals. One of the latest scientific achievements in this field is the isolation (from the skin of Amazonian frogs) of *Dendrobatidae*, a family previously unknown decahydroquinoline alkaloids (Figure 8.11c).

production of the neonicotineoids of the new generation has already surpassed that of pyrethroids (several thousand tonnes per year) owing to a broad spectrum of usage in agriculture. For example, they are useful as a seed dressing, for soil incorporation and application to the foliage of rice, cereals, corn, potatoes and sugar beets due to the absence of insects resistant to this class of insecticides.

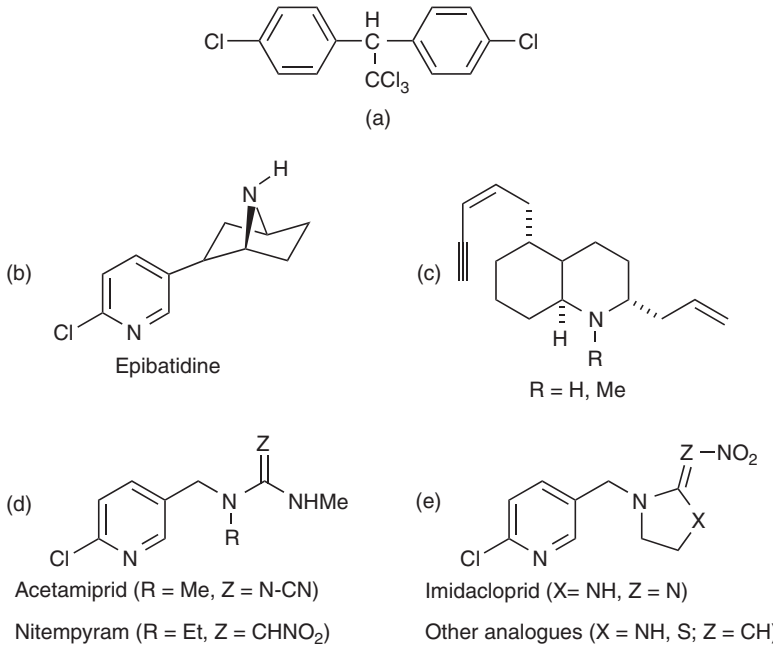


Figure 8.11 *Some synthetic chlororganic insecticides (a,d,e) and poisonous alkaloids isolated from frog skin (b, c).*

Contemporary chemical control of pests is largely achieved by organophosphorus compounds, such as the esters of phosphoric and thiophosphoric acids. Chlorophos and phtalophos are typical representatives (Figure 8.12). A number of compounds in this series, which is estimated to constitute about 43% of the total production of insecticides, contain heterocyclic fragments, mainly of the azine and azole types. As examples, the structures of diazinon, pirimiphos-methyl, menazon and phosalone are shown in Figure 8.12. Phosalone is widely used in Russia as an effective replacement for DDT against plant lice, ticks and other insects. Similar spectra of action are shown by diazinon and pirimiphos-methyl, which are recommended for application in fields and orchards. Pirimiphos-methyl is also used for disinfecting storage facilities before agricultural products are introduced. A homologue (R = Et) is especially potent against soil pests, while menazon is effective against plant lice.

Organophosphorus insecticides, as antagonists of the natural neuromediator acetylcholine (Figure 7.23), paralyze the nervous system of insects. After nerve impulse transmission, the acetylcholine is degraded in the synaptic junction to make room for the next neuromediator molecule. Hydrolytic cleavage of the acetylcholine ester group is carried out under the enzymatic control of acetylcholinesterase. Acetylcholine is thought to bind electrostatically to the enzyme at two positions through its trimethylammonium group and carbonyl carbon, the former being bound to the carboxylate anion of an aspartic acid residue of the enzyme and the latter to a

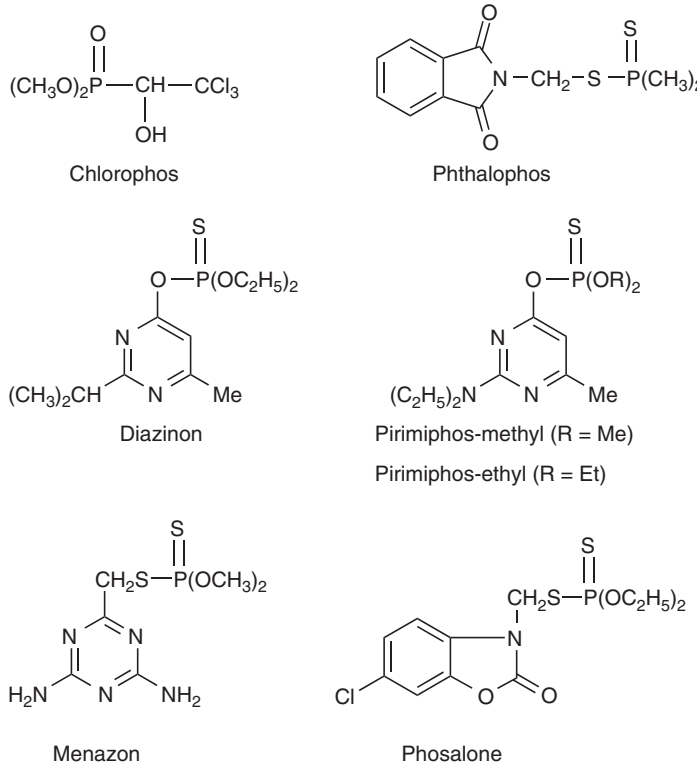


Figure 8.12 *Some organophosphorus insecticides.*

serine hydroxy moiety (Figure 8.13a). In the following step the acetyl group is cleaved from the acetylcholine molecule thus transforming it into choline, a derivative of ethanolamine containing a trimethylammonium group at the β-carbon atom (Figure 8.13b). The acetyl group linked to the serine residue in the acetylcholinesterase polypeptide chain is then readily hydrolyzed and the enzyme is regenerated.

All organophosphorus insecticides have in common their structural resemblance to the acetylcholine molecule which enables them to bind to acetylcholinesterase. An electrophilic phosphorus atom is attracted to the serine hydroxy group, while the R group, which extends some distance from the phosphorus atom, is attracted to the carboxylate anion (Figure 8.13c). For this association to occur, the R group must have a low electron density. Therefore, almost all insecticides contain electron-deficient aromatic (e.g., nitrophenyl) or heteroaromatic (e.g., pyrimidine, triazine) rings as the R group. By blocking acetylcholinesterase in this manner, organophosphorus insecticides prevent the acetylcholine molecule from approaching the enzyme. Therefore, nerve impulse transmission is interrupted and the insect's organs which rely upon cholinergic innervation are rendered ineffective. Unfortunately, organophosphorus preparations are also toxic to warm-blooded animals, including humans. Thus, we still need to create new, less dangerous and more highly selective insecticides.

Organophosphorus poisoning can occur, for example, among both agricultural workers and amateur gardeners. Moreover, many war gases are organophosphorus compounds with a similar mechanism of action. The availability of a suitable treatment in cases of organophosphorus

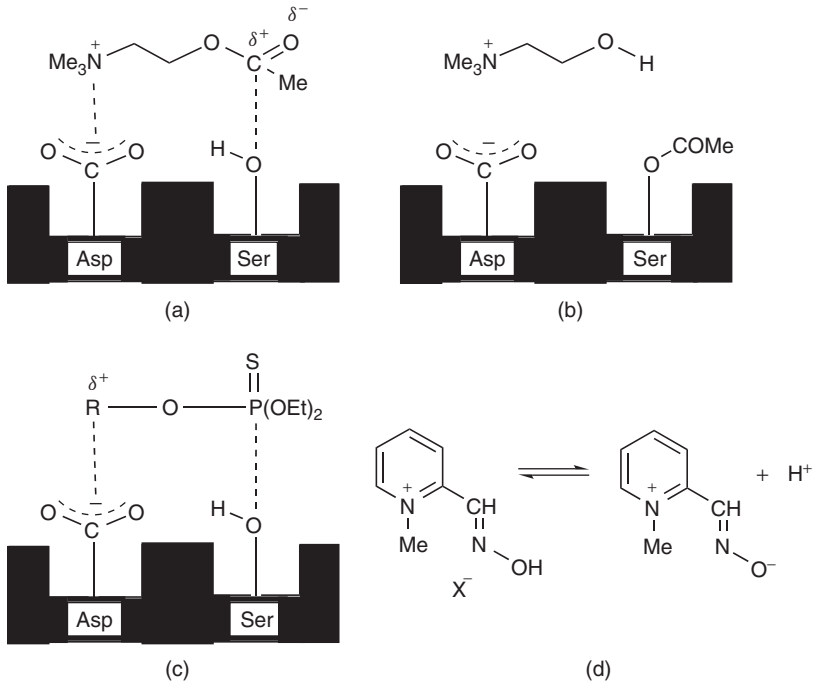


Figure 8.13 *Acetylcholinesterase interaction with acetylcholine and organophosphorus insecticides: (a) two-site binding of the enzyme with acetylcholine, (b) cleavage of the choline portion from an enzyme–substrate complex and formation of acetylated enzyme, (c) two-position binding of the enzyme to an insecticide, (d) 1-methylpyridinium-2-aldoxime as a reactivator of acetylcholinesterase. Adapted from Musil, J., Novakova, O. and Kunz, K.,* Biochemistry in Schematic Perspectives*, Avicenum, Prague, 1977, p. 45, with permission.*

poisoning is therefore important. Antidotes capable of reactivating acetylcholinesterase that has been blocked are generally heterocyclic in nature and have two common structural features, viz. an available quaternary nitrogen and an adjacent aldoxime (CH═NOH) group. 1-Methylpyridinium-2-aldoxime (Figure 8.13d) can be viewed as a typical example. The acidity of the oxime group is sufficient for ionization to the —CH═NO$^-$ anion to occur at physiological pH. The negatively charged oxygen in the anion is more strongly attracted to the phosphorus atom than is the oxygen of the serine hydroxy group, and thus the antidote displaces the organophosphorus poison from the blocked enzyme. Interestingly, if a neutral pyridine-2-aldoxime is used instead of a 1-methylpyridinium salt, the effectiveness of the reactivator decreases 1000-fold. This may imply that the reactivator, like the enzyme, acts by means of two-point contact. One hypothesis suggests that the positively charged nitrogen is necessary for the reactivator to attach to the enzyme anionic center during the expulsion of the poison. This is consistent with the fact that the distance between the positively charged nitrogen and the oxime oxygen in each antidote is close to that between the N$^+$ and the carbonyl carbon in acetylcholine.

A relatively new mode of biotoxicity against insects and ticks was discovered in some pyrrole and pyrazole derivatives (Figure 8.14). They inhibit electron transport in the chain of oxidative phosphorylation in the biosynthesis of ATP. This mechanism was proved for the natural antibiotic dioxapyrrolomycin and the synthetic insecticide chlorfenapyr. Fenpyroximate possesses uvenoid activity, suppressing larva moult in ticks by the inhibition of the NAD-H cofactor in the

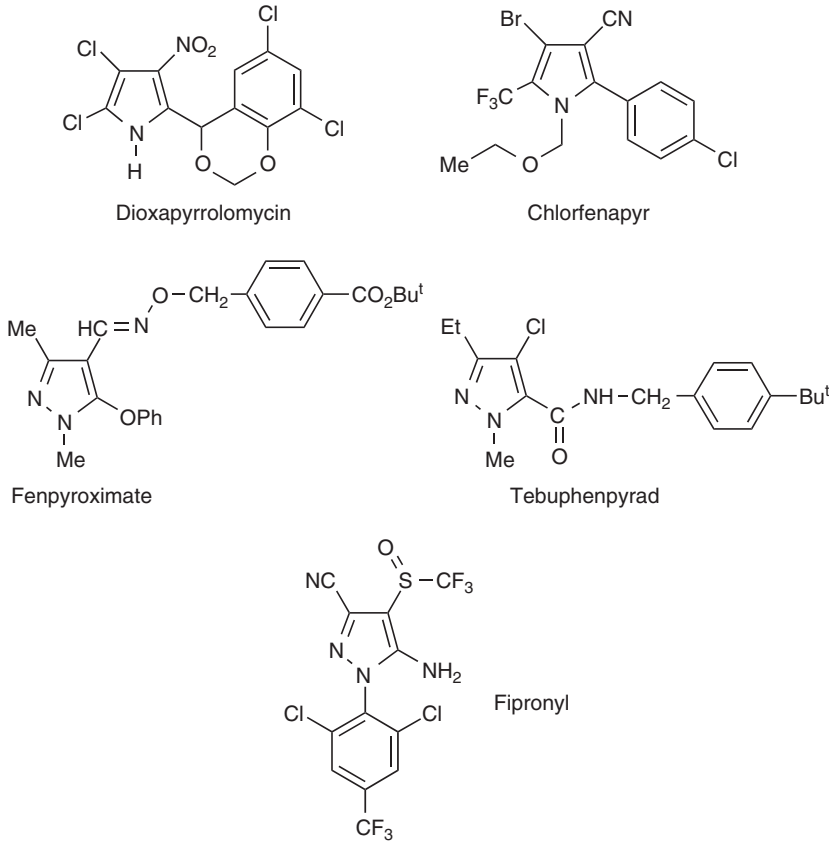


Figure 8.14 *Insecticides of pyrrole and pyrazole series interfering with electron and anion transport.*

mitochondrial oxidative phosphorylation process. Tebuphenpyrad shows systemic toxicity by a similar mechanism against a series of plant ticks.

Quantum calculations in combination with correlation analysis helped to design soil insecticides and seed protectors of a new generation, for example, fipronyl (Figure 8.14). However, this pesticide has a different mechanism of bioactivity, namely it blocks membrane conducting channels for chloride anions in GABA-dependent neurons.

Some simple 2,4,6-trihaloimidazoles (Figure 8.15) have powerful insecticidal properties with a special mechanism of toxicity targeted on the system of membrane sodium cation channels of neurons. This allows using them effectively against pests resistant to pyrethroids and organophosphorus insecticides. These biocides are transformed under the action of sun ultraviolet irradiation, air and water into derivatives of 1,2-oxet and finely into parabanic acid which has an antiseptic property.

It is of interest to note that two macrolide antibiotics – natural abamectin and its semisynthetic derivative emamectin benzoate (Figure 8.16) are presently used worldwide as insecticides and acaricides against, for example, lepidopteran pests on a variety of crops. Abamectin was isolated from the soil microorganism *Streptomyces avermitilis* and now is of much current interest for chemical modifications with the purpose of improvement and diversification of activities of its derivatives.

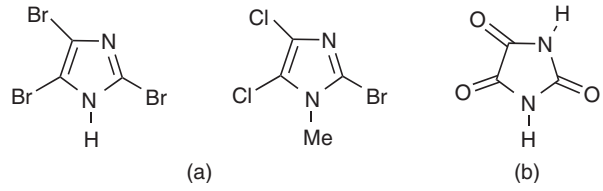


Figure 8.15 *2,4,6-Trihaloimidazoles with insecticidal activity (a) and parabanic acid (b).*

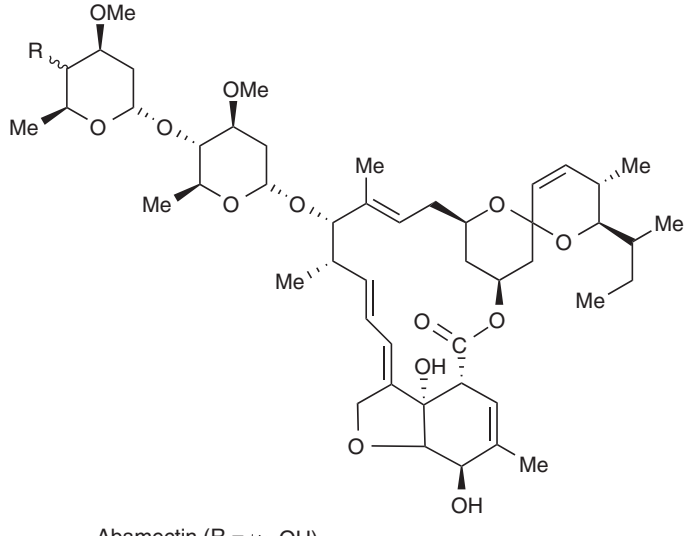


Figure 8.16 *Insecticidal macrolide antibiotics.*

In the struggle against insects a relatively new group of natural compounds named pheromones has received much attention. In nature the pheromones are used as a chemical means of communication between insects which affects their behavior. Some serve as sexual attractants, others as signals of alarm or availability of food, while others are trail markers. Pheromones are produced by the exocrine glands and are then released from the organism. Once expelled, they act as chemical signals for the same biological species. Pheromone molecules, trapped by insect receptor sites usually located in the olfactory or gustatory organs, cause a specific reaction in the receiving insect. This peculiar 'pigeon post' works with amazing efficiency and accuracy, as insects are sometimes able to receive a 'postcard' from a distance of 10 or more kilometers from the transmitting insect. The concentration of pheromone molecules may be as low as 10^{-17} mol m^{-3}.

Pheromones have been used in agriculture for some years, and quite recently have come into household use. Pheromones are employed to entice insects into traps where they are exterminated by powerful insecticides which can thus be used in a restricted site without endangering mammals. Such a method allows pesticides to be applied in very low doses economically, with high efficiency, and without polluting the environment. However, despite all of the advantages of pheromones (high

selectivity, low toxicity toward animals and man, low consumption doses) they do have one serious shortcoming, namely a high cost of production. The structures of many pheromones have been elucidated and a number have been prepared by total synthesis. The overwhelming majority are acyclic structures containing long, sometimes unsaturated, alkyl chains with a terminal functional group such as an aldehyde or ester. However, some pheromones contain heterocyclic moieties and in a number of cases the basic molecular frame is heterocyclic. The pheromone of the female bombyx (*Porthetria dispar*) contains an epoxide ring. This pheromone, disparlure (Figure 8.17), is presently manufactured on an industrial scale and is used to combat forestry pests. The pheromone of the male butterfly *Licorea ceres* and that of the 'leaf cutter' ant (*Atta texana*) are structurally simple derivatives of pyrrole (Figure 8.17). The former suppresses the motion reflexes of the female butterfly, and the latter is used by ants as a trail marker.

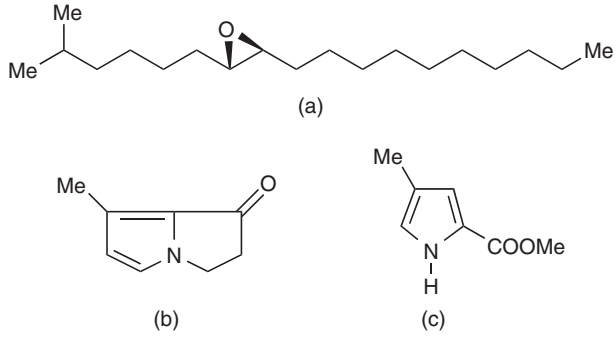


Figure 8.17 *Pheromones containing heterocyclic moieties: (a) disparlure, (b) the pheromone of the* Licorea ceres *butterfly, (c) the trail pheromone of the 'leaf-cutter' ant* Atta texana.

One of the goals in the struggle with pests is to work out a wide circle of chemical sterilizers. Their usage could cause substantial reduction in insects' number in their next generation. 5-Fluorouracyl and methotrexate (Figure 7.36) are two examples of such chemosterilizers which make some female insects barren. This kind of bioaction interfering with nucleic acid synthesis is based on the principle of antimetabolites (see Sections 7.2.2 and 7.4.3). In this case the xenobiotic is incorporated into a growing chain of nucleic acid and thus interrupts the replication process in ovarial cells and the female stops producing ova (eggs). Methotrexate begins to act at the level of inhibiting folatreductase which transforms folic acid into tetrahydrofolic acid (Figures 4.34 and 4.35).

8.4 Resisting the Kingdoms of Mustiness and Rot

Plant diseases caused by fungi are no less devastating than the aggressive invasion of insects, and often result in the total loss of the crop or harvest, even including the next season's seed supply. Therefore, efforts are ongoing to produce novel fungicides to control pathogenic fungal organisms successfully. Fungicides occupy third place after herbicides and insecticides in terms of volume of production (around 19% of the total). Fungicides are subdivided into contact and systemic agents based on the nature of their action. The latter are considered to be more valuable since they can move through the plant's vascular system. The majority of systemic fungicides in current use are derivatives of nitrogen heterocycles such as pyrazole, imidazole, triazole, pyrimidine, pyridine and so on. Triazole and benzimidazole preparations are the most effective of the class (Figure 8.18).

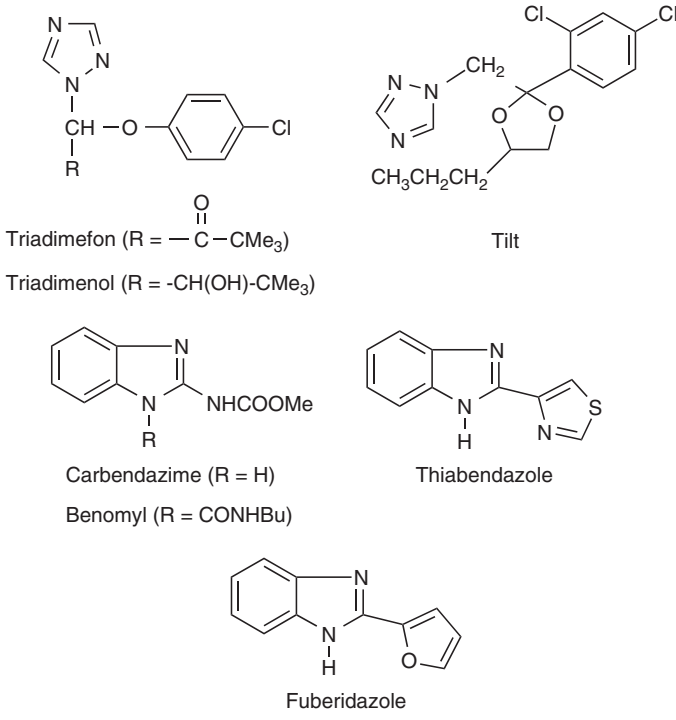


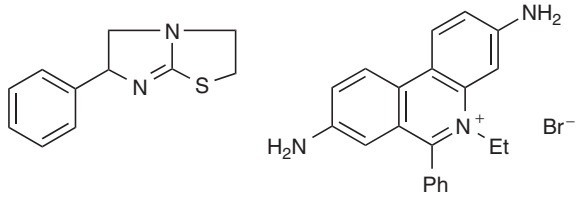
Figure 8.18 *Systemic fungicides.*

Triadimefon is the triazole-based fungicide most widely used; it has a very wide range of action and is applied during the vegetative period to combat fungal diseases of cereals, apple trees, vines, tomatoes and other plants. Triadimenol, an analogue of triadimefon containing a hydroxy instead of a carbonyl group, is utilized for seed protection and to combat mildew in cereals. Tilt also exhibits high activity and a broad spectrum of effects. Certain 1,2,4-triazole fungicides, such as triadimefon, triadimenol, uniconazole, diniconazole and other their analogues which have plant growth activity (Figure 8.8), show excellent activity against human pathogenic fungi. They belong to the large group of ergosterol biosynthesis inhibitors and their mode of action includes the inhibition of cytochrome P-450-dependent oxidative demethylation of 24-methylenedihydrolanosterol, which is a pathogen-specific precursor of ergosterol, the main sterol of pathogenic fungi. All triazole-based fungicides are of low mammalian toxicity and their application rates are low (0.4–1.0 kg ha^{-1}). These same characteristics are inherent in numerous systemic preparations.

Benzimidazole-derived fungicides were introduced into agriculture in the early 1960s and have retained their importance despite larger dose rates per cultivated hectare and the appearance of resistant forms of parasitic fungi. The chief bioactive compound of the series, benomyl, has a formidable range of action rarely encountered in other substances. Benomyl is used against pathogens of practically all cereals as well as sugar beets, vegetables, berries, fruit trees and vines and in the treatment of cotton plantations during soil tillage. Carbendazime, thiabendazole and fuberidazole are also very effective fungicides. Fuberidazole is utilized mainly in seed protection applications. Benzimidazole fungicides are assumed to disturb biofission of the fungal cell nuclei. Their highly selective toxicity is possibly due to their heightened effect on fungal rather than mammalian microtubules.

8.5 Heterocycles in Animal Husbandry

Heterocyclic compounds have found spectacular applications in animal husbandry and veterinary medicine. Meat production from animals has been significantly increased by the use of vitamin supplements (especially thiamine) in forage and by the addition of antibiotics and tranquilizers to feed. Tranquilizers alleviate stresses in the animals, which are usually inevitable under the conditions of factory farming. The struggle against various animal parasites, especially helminthic invasions, is of tremendous importance. For many years these parasitic infections were treated with phenothiazine (Figure 2.11). Today, more efficient medications, such as the highly active thiabendazole (Figure 8.18) and tetramizole (Figure 8.19), have superseded phenothiazine.



Tetramizole Ethidium bromide

Figure 8.19 *Some heterocyclic medications used against animal parasites.*

Surra is an extremely dangerous illness of cattle caused by a single-cell microorganisms called *Trypanosoma* which invade the animal's blood and tissues. The analogous human ailment is known as trypanosomiasis (see Section 7.2.6). Surra is a particular problem in countries with hot climates; for a long time the breeding of cattle in central Africa was virtually impossible owing to this parasitic scourge. Compounds based on aminophenanthridinium salts were used effectively in these cases. For instance, a single dose of ethidium bromide (Figure 8.19) was sometimes sufficient for the complete recovery of an animal afflicted with surra.

Coccidiosis infections, which usually result in severe fowl losses, can make a massive negative impact on the poultry industry. To combat this parasitic disease, veterinary medicines such as nitrofuran derivatives, sulfa drugs and antibiotics are added to the chicken's drinking water. One further heterocyclic treatment that has become useful in veterinary medicine as a cure for a number of parasitic skin diseases is nicotine sulfate.

8.6 Combinatorial Chemistry and Functional Genomics in the Synthesis of Biologically Active Heterocyclic Compounds

Despite considerable quantities of pesticides applied each year, the loss of crops in the agriculture industry is not diminishing, and in absolute tonnage it is even increasing. This is caused by two main factors. First, only 35% of all agriculturally cultivated lands in the world (mainly plantations of rice, corn, cotton, vegetables and fruit trees) are treated with pesticides. Second, over the 50-year period of wide usage of pesticides, pests have produced generations which are very resistant towards many old toxic compounds, leading them to become almost useless. To solve these problems necessitates continuing the synthesis of an enormous number of new organic compounds and choosing from them those which have the needed biological activity. The creation of combinatorial chemistry at the end of twentieth century has substantially helped to increase the

productivity of organic synthesis. For comparison the nineteenth and twentieth centuries both produced approximately 10^7 organic compounds, while the past the 20 years (1990–2010) produced approximately the same quantity owing to combinatorial chemistry.

This new principle of creation of medicinal and pesticidal compounds is based on robotic techniques of parallel micro syntheses carried out in hundreds or even thousands of miniature reactors. In such simultaneously conducted reactions in solution or on hard templates it becomes possible to obtain in one day hundreds or even thousands of new compounds from 5 mg to 1 g. Such a family of compounds is named a 'library'. In an analogous manner a whole library of compounds can be automatically tested for bioactivity to find so-called 'hits', that is, compounds with a high level of bioactivity. Such a block of hits forms a so-called focal library which allows one to choose the lead compounds. The next step of biotesting consists in choosing lead compounds possessing a maximal level of useful bioactivity. This group of leads then gives candidates likely to be drugs or pesticides. From among them one substance is selected to become a real pesticide or a drug to be produced in mass quantities and directed to the market.

Figure 8.20 shows a typical example of combinatorial synthesis of tetra and penta derivatives of 1,3-oxazolidines as potential plant growth regulators and herbicidal antidotes.

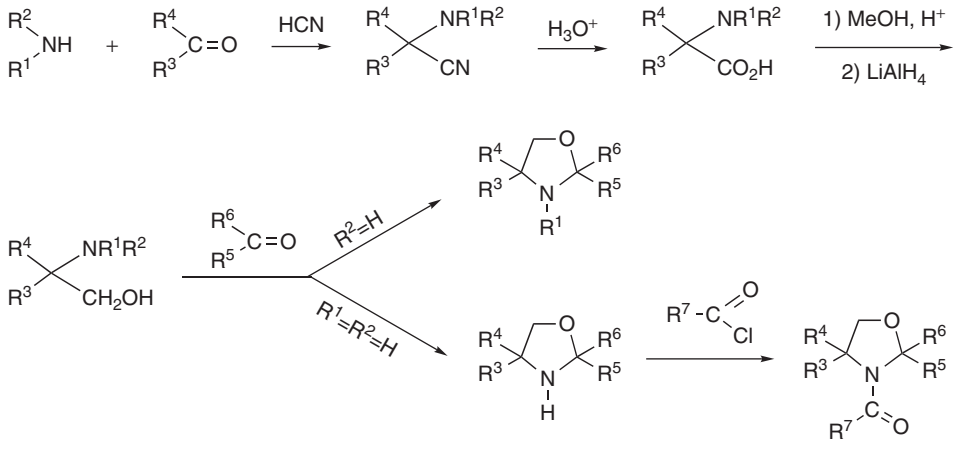


Figure 8.20 *Example of a combinatorial preparation of the 1,3-oxazolidine 'library'.*

One more novel methodology for the synthesis of biologically active compounds, which evolved in the twentyfirst century after the decoding of the human genome, is called functional genomics and proteomics because this principle is based on the knowledge of the structure of genome and the functions of its genes and proteins coded by them. Systematic work on decoding the genomes of different organisms began about 25 years ago, and by 2011 the genomic structures of more than 200 bacteria, viruses, fungi, plants, insects and vertebrates have become known. The first decoded genome was that of the weed *Arabidopsis* which contained 25 000 genes composed of 120 million pairs of bases. Since then the genomes of the soil worm nematode (about 20 000 genes), fruit fly drosophila (14 000 genes), malarial mosquito (13 600 genes) and field mouse (about 30 000 genes) have been determined. More precise data obtained in 2003 on the human genome give 24 847 genes.

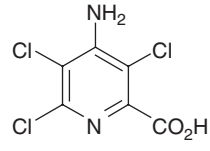
The determination of the structures and functions of genes in pathogenic microorganisms and pests has thus become a major task, as their understanding is a key to the synthesis of useful biologically active compounds as pesticides. Such compounds, new and old, will increasingly be used

to act more selectively and to control the biosynthesis of signaling endomolecules, bioreceptors and other biopeptides both in pests and in cultivated organisms.

A strategy arose after the discovery of special proteins and their genes which control apoptosis in the nematode *Caenorhabditis elegans* (Nobel Prize in 2001). This approach now allows the synthesis of organic molecules which can mimic these native proteins and thus switch off life-important cell functions, causing cell apoptosis and even phenoptosis of pathogens and pests. Another recent discovery is gene DDT-12 in the genome of the fruit fly; this gene generates resistance towards the chlororganic insecticide DDT. It opens a new perspective for pesticide and growth regulator synthesis. This gene controls the level of the most important hemoproteide cytochrome P-450 that destructs DDT in the fly organism. Thus the development of functional genomics and proteomics may substantially stimulate the synthesis of compounds acting on the genes of pests with acquired resistance and on peptides controlled by these genes.

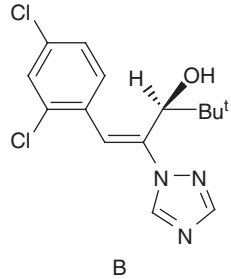
8.7 Problems

1. What physicochemical properties give rise to the herbicidal activity of paraquat and diquat dibromide? Discuss the reasons for their rather rapid inactivation in the field.
2. Picloram (A) is widely used for the control of perennial weeds and shrubs. The moderate solubility of picloram in water is dramatically increased when the herbicide is used in the potassium salt form. However, this increased solubility raises the environmental concern of potential groundwater contamination. This problem could be effectively circumvented by the formation of insoluble complexes of picloram with the metal ions present in soil and groundwater. One such nonlabile complex is readily formed by treatment of the herbicide with iron(II) ions. Indicate the structure of this complex.



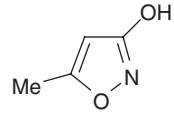
A

3. Discuss the mechanism of the biological activity of heterocyclic plant hormones. Give examples of synthetic growth regulators.
4. Nicotine is used in agriculture mainly in the monosulfate form. Give the structural formula of this salt.
5. Discuss the reasoning which led to the use of 1-methylpyridinium-2-aldoxime as an antidote for organophosphorus reagent poisoning. Why is the neutral pyridine-2-aldoxime 1000 times less active?
6. What are systemic fungicides? Give examples from the benzimidazole series.
7. Diniconazole (B) is a broad action systemic fungicide which was shown to bind stoichiometrically to cytochrome P-450 enzymes via a lone pair of electrons on one of the nitrogen atoms. Diniconazole thus disrupts ergosterol biosynthesis in fungi by mimicking the conformation of one of the intermediates, lanosterol (the cytochrome P-450 enzymes are responsible for oxidation of the lanosterol methyl group at the C-14 position). Structure B is the most active of the four possible stereoisomers. (a) Draw the structures of the other three isomers. (b) Is there any intramolecular stabilizing interaction in these isomers? (c) Which of the three nitrogen atoms of the fungicide binds to the cytochrome enzymes?



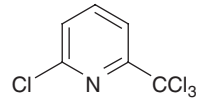
B

8. The toxoforic groups of 1,3,5-triazine-based herbicides (Figure 8.1) may inhibit the enzymes required in the photochemical transformation of water into oxygen and block the system of oxidative phosphorilation that stops the supply of bioenergetic ATFP molecules, and the plant dies. If the R substituents (Cl or SMe groups) in these herbicides are exchanged by a hydroxyl group, then the herbicidal activity is completely eliminated. Discuss this fact. Propose the other R groups that may extinguish or preserve this bioactivity.
9. Tachigaren (C) is an industrial highly active regulator of rice and beet root growth. At the same time it possesses fungicidal activity on some fungal organisms pathogenic towards these plants. It was proved that the compound itself is not really bioactive but instead one of its two isomeric glycoside derivatives does produce the regulatory effect and another shows the property of being the true fungicide. Design the structures of these bioactive xenometabolites 'packed in one nonactive flacon'.



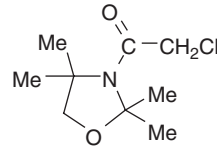
C

10. 2,4,5-Trihaloimidazoles (Figure 8.16) not only possess true insecticidal properties, they may also be considered as probactericides. Explain this fact and write chemical and photochemical transformations of such insecticides into true bactericide.
11. A series of soil bacteria utilize intensively plants nitrogen fertilizers (urea, ammonium nitrate). This process of bacterial nitrification brings about considerable loss of the fertilizers. In order to inhibit the activity of the nitrifying bacteria some heterocyclic compounds were designed and synthesized. The need for such inhibitors is very great – more than 100 000 t $year^{-1}$. One of them is nitropyrine or N-serve (D), which is in worldwide use to date, but it has some shortcomings – too volatile to be applied in fertilizers compositions and easily hydrolyzed in soil. Propose chemical and physical measures to circumvent these disadvantages.



D

12. Rewrite Figure 8.20 uncoding all of the substituents (R^1–R^7) needed to synthesize oxazoline derivative E, which is used as a safener (herbicidal antidote).



E

8.8 Suggested Reading

1. Baker, D. R., Fenyes, J. G. and Basarab, G. S. (eds), *Synthesis and Chemistry of Agrochemicals IV*, 4th edn, American Chemical Society, Washington, D.C., 1995.
2. Tomlin, C. (ed.), *The Pesticide Manual: a World Compendium: Incorporating the Agrochemicals Handbook*, 13th edn, Royal Society of Chemistry, Cambridge, 2003.
3. Bunce, N. J., *Environmental Chemistry*, Wuerz, Winnipeg, 1990.
4. Hutson, D. H. and Roberts, T. R. (eds), *Herbicides*, John Wiley & Sons, Inc., New York, 1987.
5. Ware, G. W., *Pesticides, Theory and Application*, Freeman, San Francisco, 1983.
6. Sitting, M. (ed.), *Pesticide Manufacturing and Toxic Material Control Encyclopedia*, Noyes Data, Park Ridge, 1980.
7. Tedder, J. M., Nechvatal, A. and Jubb, A. H., *Basic Organic Chemistry. Part 5: Industrial Products*, John Wiley & Sons, Ltd, London, 1975.

9

Heterocycles in Industry and Technology

I throw the garland of roses
To the world of mysterious themes,
And by it I plunge into the chaos
of unknown creative day-dreams.
V. Bryusov

Heterocycles have been indispensable in the recent far-reaching developments in science and technology for numerous applications, including the biomedical sciences, electronics, communications and aerospace technology. At the same time they remain enormously important in more traditional branches of industry such as the dye industry. The manufacture of synthetic dyes began in the second half of the nineteenth century and heterocycles immediately achieved preeminence. However, from time immemorial man has learned how to use natural pigments, of which a considerable number are heterocyclic. We begin this chapter by discussing the contribution of heterocycles to the polychromism of our world.

9.1 Heterocycles and Natural Colors

The human eye perceives the surrounding world as a multicolored picture. The existing natural colors from green grass to the palette of a butterfly wing give us immense aesthetic pleasure. We do not know whether animals share these feelings with us, but we do know that they require different colors to fulfill certain biological functions and in some cases for survival itself. Animals use their own colorings to disguise themselves, to threaten enemies, to search for mates and so on. Animals also rely on color to judge whether fruit is suitable to eat.

Colored substances usually contain an extended carbon chain composed of alternating single and double bonds. Such a chain is called a chromophore. β-Carotene (Figure 6.4), the orange pigment found in carrots and apricots, is a typical example. Carotenoids are the most widespread

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

natural pigments. We have already seen in Chapter 6 that carotenoids are the main constituents of chloroplasts found in every green plant. Another, less widely distributed, class of pigments is the quinones. 5-Hydroxy-1,4-naphthoquinone (juglone), the yellow-orange component of unripe walnut shells (Figure 9.1), is a typical representative.[1] In contrast to the carotenoids, the functional group electrons participate in the conjugation in this pigment along with electrons of the C═C bonds. When a molecule is excited by light, electrons are transported along the conjugated chain from the donor (hydroxy) to the acceptor (carbonyl) group. This process causes significant stabilization of the excited state and results in deeper coloration. Such functional groups are called auxochromes. Pyrrole-like and pyridine-like heteroatoms are efficient auxochromes in many recent far-reaching developments in science and technology. Consequently, many heterocyclic compounds are encountered among naturally occurring pigments. Green chlorophyll (Figure 6.2) and red hemoglobin (Section 4.2.2) are classic examples.

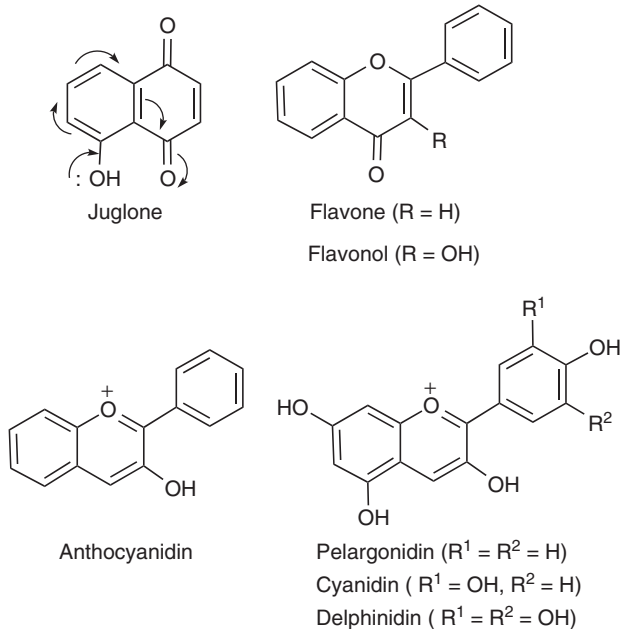


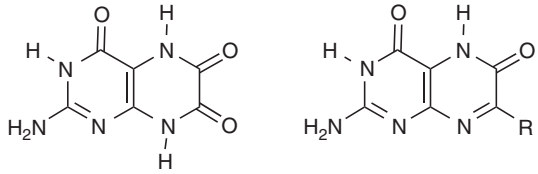
Figure 9.1 *Juglone and some flavonoid pigments.*

Another class of heterocyclic pigments includes substances with the flavonoid structure, comprising derivatives of flavone, flavonol and anthocyanidin (Figure 9.1, see also Figure 4.12). A benzopyran skeleton forms the basis of the flavonoid structure. Substituents typically include one or more phenolic hydroxy groups. Flavonoids usually exist in a methylated or glycosidated form in the plant. The most important flavonoids are the anthocyanidin glycosides, such as orange pelargonidin, red cyanidin and purple delphinidin. These pigments are responsible for the orange, red and purple/blue coloration of flowers. The colors of ripe cherries, strawberries, raspberries, plums and red apples are also the result of anthocyanidins.

[1] Juglone exists in a reduced, uncolored form in the walnut shell. When the shell is broken, the reduced compound is released with the juice and is oxidized by air to form the pigment.

Flavones and flavonols differ from anthocyanidins by an almost complete absence of light absorption in the visible region. Nevertheless, they produce vivid white and cream colors in a number of flowers such as tea rose, cherry, apple, plum, apricot and others. 'Just look! So many daisies are here and there,' wrote the Russian poet Severyanin, 'and many are in blossom; even too many; in full bloom. Their petals are triple-edged as wings, and white as silk.'

Flavonoid pigments are very seldom encountered in animals. The animal kingdom is rich in another family of heterocyclic pigments, namely the pterins (Figure 9.2). Pterins are frequently found in butterflies and other insects. Thus, the white color of the cabbage butterfly's wings arises from the presence of leucopterin. Chrysopterin imparts a yellow hue to the lemon butterfly, erythropterin produces the bright red color of the *Zegris f. Chr*. butterfly and xanthopterin is partly responsible for the yellow color of wasps.



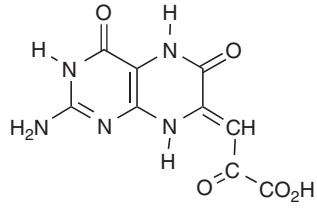


Figure 9.2 *Pterin pigments.*

Linear tetrapyrrolic compounds such as biliverdin and bilirubin are another more widespread class of natural heterocyclic pigments (Figure 9.3). They result from the oxidative breakdown of heme (see Figure 4.17). In animals primarily formed biliverdin is rapidly reduced to bilirubin, which is responsible for the yellow color of bile, urea, bruises and jaundice. Plant biliverdin goes on to form phytochromobilins, accessory light-antenna chromophores, and to chlorophyll (Figure 6.4b). Recently, bilirubin has been found for the first time in the plant kingdom: the pigment is responsible for the brilliant orange seed arils of the bird of paradise tree.

For information concerning melanin pigments, also having a heterocyclic nature and widely occurring in animal kingdom, see Section 11.1.4.

9.2 Dyes

9.2.1 From Imperial Cloaks to Jeans

Since ancient times people have used dyes, obtained from various natural sources, to adorn their dwellings and clothes and to prepare cosmetics. In the middle ages, the yellow flavonoid pigment

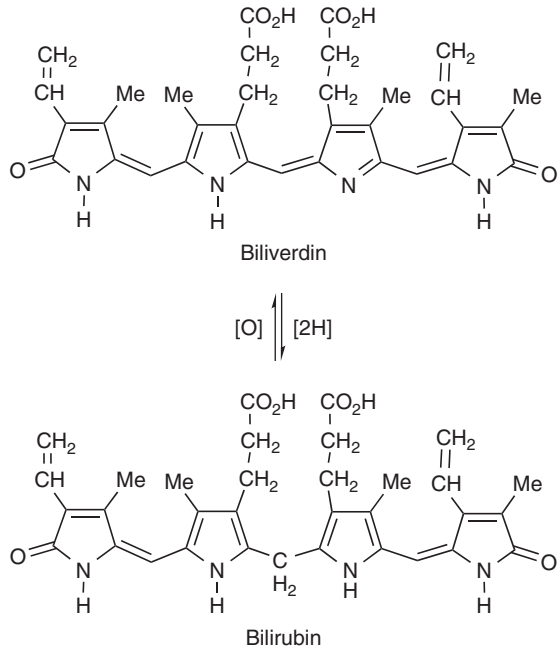


Figure 9.3 *Tetrapyrrolic pigments.*

luteolin (Figure 9.4a) was extremely popular. Luteolin was prepared from the stems, leaves and seeds of a plant called *Reseda luteola*. The blue dye indigo was separated from the leaves of another plant, *Indigofera tinctoria*. The well-known purple dye 6,6′-dibromoindigo, known as Tyrian Purple or Royal Purple, was obtained from the Mediterranean sea mollusk *Murex brabdaris*. In ancient times, Tyrian Purple was manufactured in the Phoenician town of Tyre and was originally used to color the cloaks of pharaohs, emperors and high priests.

It should be noted that indigo and dibromoindigo do not occur in nature, but are formed during the processing from natural precursors. Thus, the leaves of the indigo plant contain indican (3-hydroxyindole *O*-glycoside). When the leaves are destroyed, indican is first transformed into 3-hydroxyindole (indoxyl, which exists predominantly in the tautomeric oxoform) and then into indigo (Figure 9.4b) by the action of enzymes and oxygen.

Unfortunately, almost all natural organic dyes have shortcomings. Flavonoids, for example, suffer from poor stability toward light and chemicals. Other dyes are rather expensive to produce owing to the difficulties of their separation and purification, and their low content in the raw materials; for example, approximately 10 000 mollusks had to be processed to isolate just 1 g of Tyrian Purple! It is clear why this dye was used in ancient and medieval times only by the rich.

A new era in the dye industry began in the middle of the nineteenth century when rapid advancements in organic chemistry allowed the creation of synthetic dyes. Aromatic and heterocyclic compounds ‘stole the limelight’. The first entirely synthetic dye was mauveine, an ionic derivative of phenazine (Figure 9.5) produced by the English chemist Perkin via oxidation of a mixture of aniline and toluidines with potassium dichromate in sulfuric acid. Mauveine is red in color and is characterized by high stability to light, washing and mechanical agitation. In the past, the dye was widely used to color silk and wool.

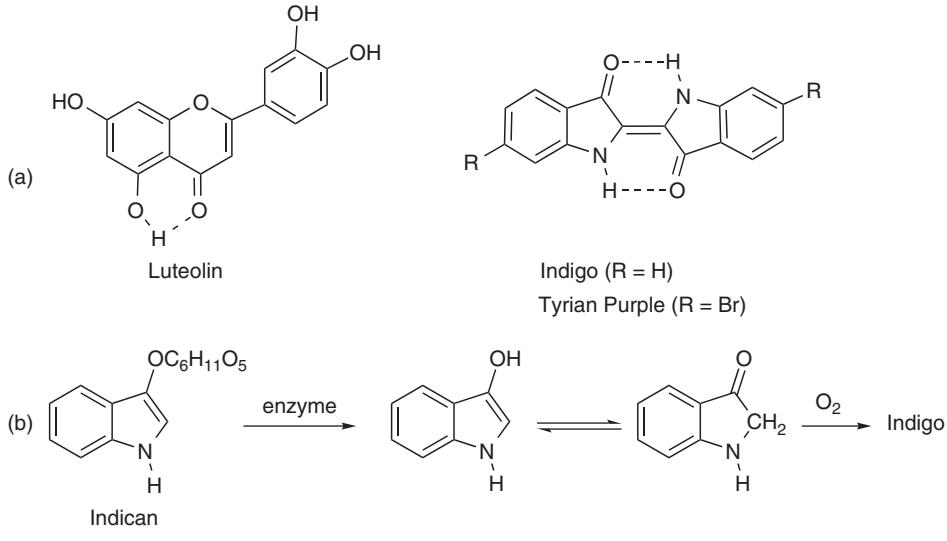


Figure 9.4 *(a) Heterocyclic dyes from natural sources. (b) Formation of indigo from indican.*

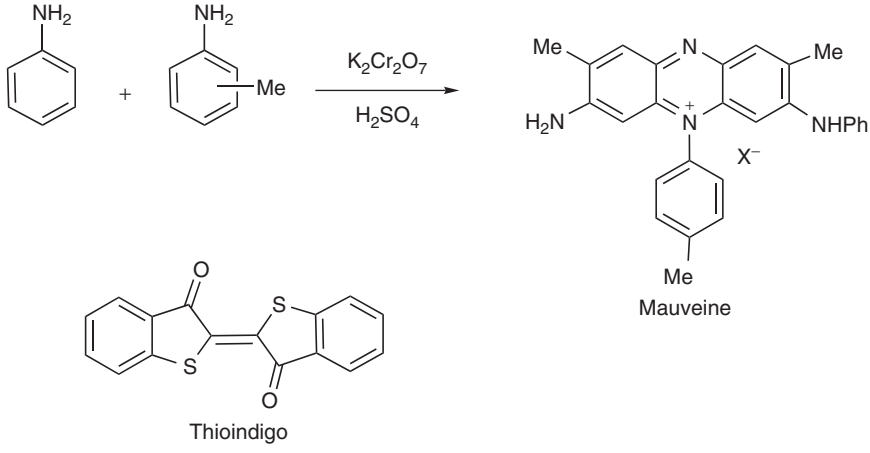


Figure 9.5 *The synthetic dyes mauveine and thioindigo.*

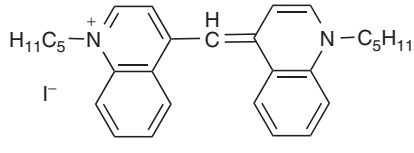
Another significant event in the history of chemical dyes occurred during 1869–1883 when von Baeyer succeeded in elucidating the structure of indigo, thereby enabling an industrial synthesis of indoxyl and subsequently indigo; for this he received the Nobel prize in 1905.[2] Since this time inexpensive synthetic indigo has become a common dye with continuing widespread use in the textile industry, mainly for coloring jeans. The success of the indigo production (more than

[2] Incidentally, von Baeyer was incorrect in ascribing the *cis* orientation to the two C═O groups in the indigo molecule. It was not until 1926 that the error was corrected.

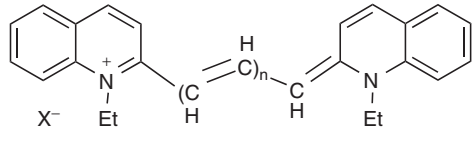
10 methods of synthesis were developed) gave rise to the production of numerous derivatives and analogues known as indigoid dyes. The replacement of nitrogen atoms by other heteroatoms was often used in the modification of the indigo structure. The red dye thioindigo (Figure 9.5) serves as a typical example.

9.2.2 'Cyanine' Means Azure

An important feature of heterocyclic dyes is the ready modification of their precise hue by structural changes. Other parameters (solubility, affinity to fabrics) are altered by conversion of the heteroatom to the cationic form. Discovery of this fact spurred the synthesis of a large family of cationic dyes. The first representative, a blue dye named cyanine (Greek: *kyanos*, azure), was synthesized in 1856 by G. Williams by heating a mixture of quinoline and 4-methylquinoline (lepidine) with isopentyl iodide in alkali. Although cyanine (Figure 9.6) did not find practical applications, it served as a prototype for the large group of cationic 'cyanine' dyes.

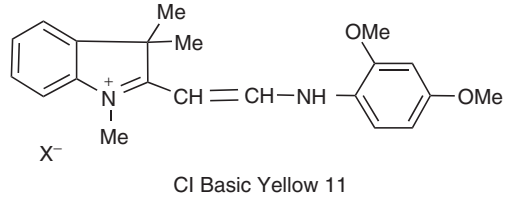


Cyanine

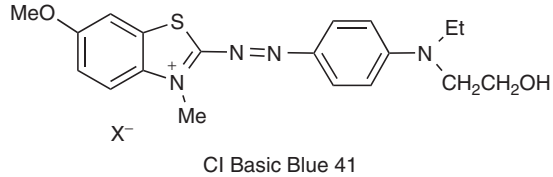


Pseudocyanine (n = 0)

Pinacyanol (n = 1)



CI Basic Yellow 11



CI Basic Blue 41

Figure 9.6 *Examples of cyanine dyes.*

Cyanine dyes generally consist of two heterocyclic ring systems connected by a bridge of conjugated carbon bonds which can vary in length. At one end, a (usually heteroaromatic) heteroatom is in a cationic state and serves as an electron acceptor; in the other, partially saturated nucleus,

the heteroatom is formally electronically neutral (pyrrole-like) and functions as an electron donor. Figure 9.6 shows three such dyes: cyanine, pseudocyanine and pinacyanol.

It is to be stressed that the difference between the heteroatoms mentioned above is grossly oversimplified as it takes into consideration only one possible canonical structure (resonance structure). In reality, the positive charge is distributed more or less equally between the nuclei because of the symmetry and conjugation of the molecule. It is this delocalization of the electrons and spreading of the charge that imparts the characteristic deep color to the cyanine dyes.

Nonsymmetrical cyanine dyes are also known. In some, familiar functional groups such as NH_2, OH, OR and so on play the role of electron donor. A typical example, CI Basic Yellow 11, is widely manufactured. The cationic azo dyes (e.g., CI Basic Blue 41; Figure 9.6) are considered to be variants of the cyanine dyes. Here, the azo group functions as the electron-transferring bridge in place of the polymethine structure previously discussed. Figure 9.7 details the synthesis of a typical example, CI Basic Red 22. In the first step, a heteroaromatic amine is diazotized and coupled with dimethylaniline or another coupling component. The azo product thus formed is transformed into a quaternary salt by heating with an alkylating agent. Owing to the nonequivalence of the three triazole ring nitrogens, two isomeric quaternary salts, in a molar ratio of 1:6, are formed. In one isomer, the two methyl groups are located at the 1- and 4-positions; in the other, they occupy the 2- and 4-positions. The commercially available dyestuff is such a mixture.

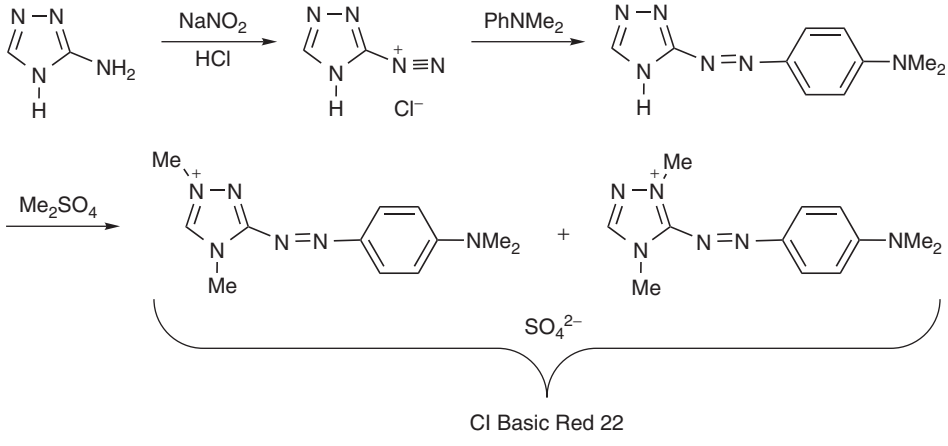


Figure 9.7 *Synthesis of the dye CI basic red 22.*

The value of cationic dyes lies in their rather high light resistance, their intense color and the wide palette available, ranging from red to violet. Moreover, they are one of the very few classes of dyes suitable for the coloration of polyacrylics, which are synthetic fibers in widespread use.

9.2.3 Phthalocyanines: Sometimes Better than Porphyrins

Despite their unique role in nature, porphyrin pigments for a long time had no industrial applications. Porphyrins and metalloporphyrins are rather weak absorbers of visible light, and their colors are insufficiently intense for these compounds to be useful as dyes. The stability of porphyrins toward the action of light and chemicals also leaves much to be desired. However, in the late 1920s and early 1930s a new class of dye was discovered with great practical importance. This porphyrin-based substance was named phthalocyanine because of its blue color and the fact that it was derived

from phthalic acid. The parent compound, unsubstituted phthalocyanine (Figure 9.8), is prepared by heating the nitrile of phthalic acid with the alcoholate of a higher alcohol. The disodium salt of phthalocyanine is formed by tetramerization of phthalodinitrile. Acidification liberates free phthalocyanine. As Figure 9.8 shows, phthalocyanine has a tetrabenzotetraazaporphyrin structure. The introduction of four additional nitrogen atoms to the meso positions of the porphyrin system, together with the annulation of each of the pyrrole rings with a benzene ring, dramatically enhances light absorption in the visible region. The stability toward light, chemicals and temperature is also enormously increased. The phthalocyanines are all very strongly colored pigments which produce bright colors that are highly resistant to acids, alkalis and high temperatures. Phthalocyanines are used to color many varied materials including fibers, paper, polymers, artificial leather and so on, and to produce varnishes, inks and dyes in solution form. Phthalocyanine itself, its complex with the copper(II) ion (copper phthalocyanine) and a derivative containing fully chlorinated benzene

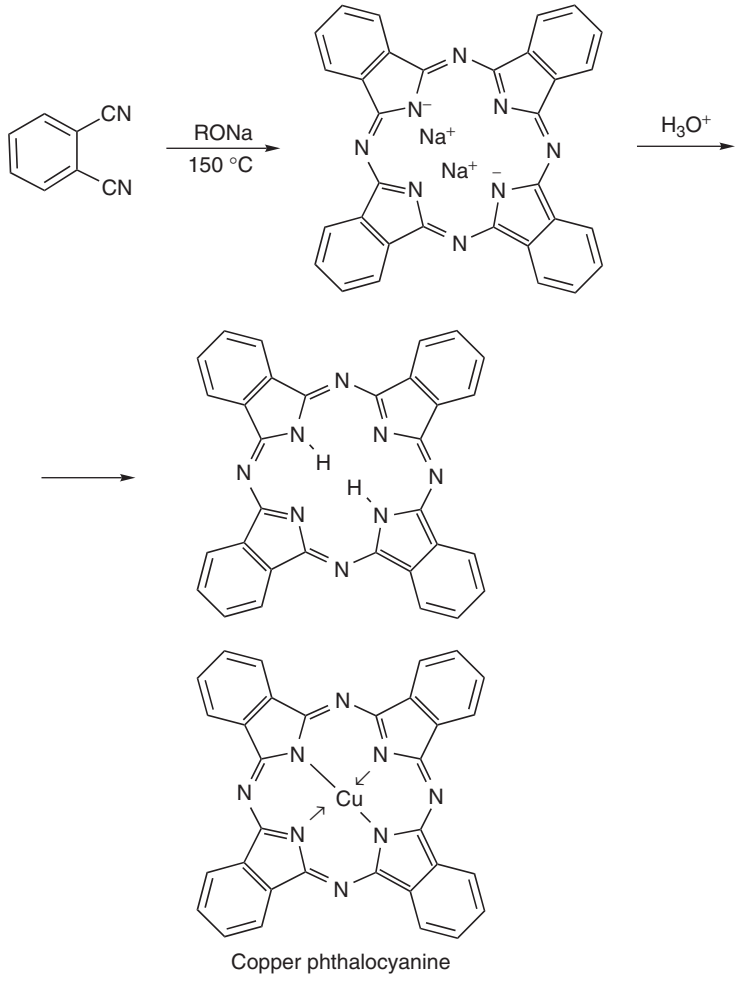


Figure 9.8 *Phthalocyanine pigments.*

rings are now the most widely used members of the class. They generate green-blue, bright blue and green colors, respectively, and are chiefly utilized as pigments (i.e., dyes which are insoluble in the medium in which the dyeing process is carried out). The worldwide production of phthalocyanines amounts to many thousands of tonnes. The discovery of phthalocyanines was a truly outstanding achievement in dye chemistry of the twentieth century.

A number of other classes of heterocyclic dyes including the phthaloperinone, quinacridone, triphenodioxazine and 3,6-diarylpyrrolo[3,4-c]pyrrole-1,4-dione systems (Figure 9.9) are also used for the preparation of pigments.

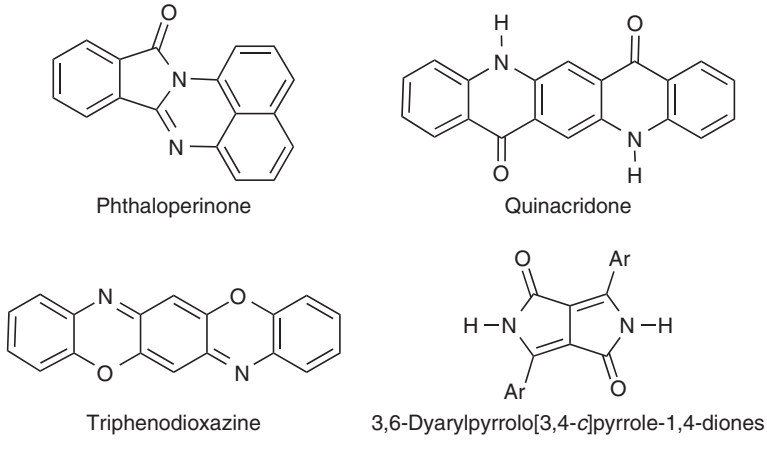


Figure 9.9 *Polynuclear chromophores used in pigment manufacture.*

9.2.4 The Anchoring of Dyes

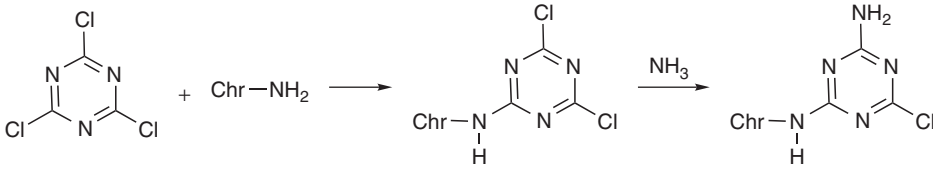
As chromophores, heterocycles can, in addition to providing color for a dye, help to fasten this color to the fiber. This was clearly demonstrated by the so-called 'reactive dyes' first synthesized in the mid-1950s. These dyes bind to a fiber by forming covalent bonds rather than simply via absorptive forces and nonbonding interactions. The effect of the binding may be compared with the effect of a strong anchor that secures a ship during a violent storm.

The structure of a reactive dye may be represented in a general way as Chr—M—X, where Chr is the molecular chromophore, X is the reactive functional group which interacts with the fiber and M is an intermediate link whose main function is to attach the X substituent. Reactive dyes are effective for coloring wool and synthetic polyamide fibers, but are most often used on cellulose fibers such as cotton. The mode of action of reactive dyes is based on the following sequence in which they combine with the hydroxy groups of the cellulose (Cel—OH) or the amido groups of the amide fibers (Am—NH):

$$\mathrm{Chr{-}M{-}X + Cel{-}OH \rightarrow Chr{-}M{-}O{-}Cel + HX}$$
$$\mathrm{Chr{-}M{-}X + Am{-}NH \rightarrow Chr{-}M{-}N{-}Am + HX}$$

Any known chromophore (e.g., a phthalocyanine, anthraquinone or, most often, an azo dye) may function as the chromophore group, X is generally Cl or F, and the M fragment is usually a heterocyclic system. Since the interaction of a reactive dye with a fiber essentially involves nucleophilic displacement of the halide, the function of the M group is to activate this displacement

(i.e., M should be an electron acceptor). Highly π-deficient azines such as 1,3,5-triazine, pyrimidine, quinoxaline and so on meet this requirement. Many reactive dyes include a triazine nucleus which is typically introduced into the dye molecule via 2,4,6-trichloro-1,3,5-triazine, also known as cyanuric chloride (Figure 9.10). Treatment of dyes containing an amino group with cyanuric chloride generates a reactive dye. However, replacement of one further chlorine by an amino group is sometimes carried out. Anchoring of the dye to the fabric occurs via the remaining chlorine atom. Figure 9.10 depicts the synthesis of a deep red reactive dye. An alternative method of anchoring dyes relies on vinyl sulfone intermediates like Chr—M—$SO_2CH{=}CH_2$, which can add NH or OH by Michael machin.



Cyanuric chloride

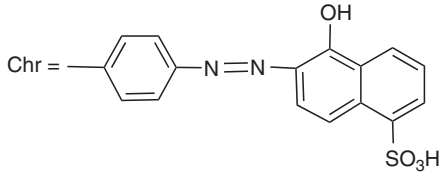


Figure 9.10 *Synthesis of reactive dyes.*

9.3 Fluorescent Agents

Many organic compounds, including some heterocycles, possess a property which causes them to glow under the action of ultraviolet or visible light. Such compounds are named luminophores, and the process by which light is radiated is called photoluminescence, or simply luminescence.[3] Luminophores, or fluorescent agents, have found diverse applications in industry, technology and science. To understand better the principles of their use, we first examine a number of specific features of luminophores.

9.3.1 Why They Shine

Almost all organic luminophores, including the aromatic hydrocarbons stilbene, terphenyl and anthracene (Figure 9.11), contain an extensive conjugated system. The greater the number of π-electrons participating in the conjugation, the narrower the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). As a result, in a compound containing extended conjugation the energy of near-ultraviolet or even visible light becomes sufficient for an electron to migrate from the HOMO to the LUMO. We have already discussed the dissipation of energy by an excited molecule in Section 6.1 (see Figure 6.5)

[3] In certain cases luminescence can be caused by sources of excitation other than light, such as radioactivity (radioluminescence), electric fields (electroluminescence), mechanical forces (triboluminescence) and so on.

and are aware that some of the absorbed energy may be lost in various nonradiative processes (e.g., vibration, rotation). Therefore, when such molecules emit a photon (by spontaneous emission; see Figure 6.5d), the energy of the photon emitted is lower than the energy of the photon which caused the excitation (this phenomenon is known as the Stokes shift). Consequently, the light emitted by the molecule will have a longer wavelength than the light absorbed. Thus, anthracene absorbs light of wavelength 380 nm and emits at 434 nm; that is, upon irradiation in the near-ultraviolet, the otherwise colorless anthracene emits blue light.

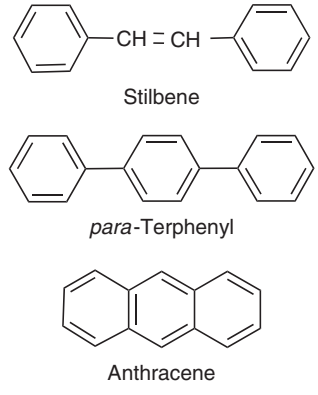


Figure 9.11 *Examples of aromatic luminophores.*

There are two kinds of luminescence: fluorescence and phosphorescence. The spin of the excited electron does not change during fluorescence (Figure 6.5d). In other words, the molecule is in the singlet excited state S_1 (the ground states S_0 of molecules other than radicals are also singlet states). In the case of phosphorescence, the spin of the excited electron changes and therefore two electrons with unpaired spins exist in the excited state of the molecule. Such a condition is called a triplet state and is designated T_1. A triplet state has a much longer lifetime than a singlet state because $T_1 \rightarrow S_0$ is a forbidden transition by the so-called selection rules, in contrast to the $S_1 \rightarrow S_0$ allowed transition. This leads to an increase in the duration of phosphorescence (up to tens of seconds), whereas fluorescence usually lasts less than 10^{-6} s. The majority of organic luminophores irradiate mainly via fluorescence, although there may be some concurrent phosphorescence.

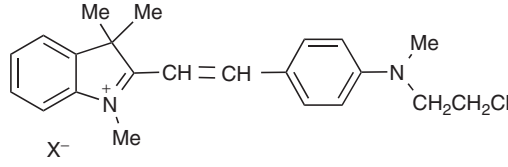
Structural rigidity in a molecule, coupled with an extensive delocalized π-electron system, favors luminescence. The inclusion of electron-donating and electron-withdrawing groups in the conjugated system also has a favorable effect. Structural rigidity is necessary to minimize loss of excitation energy into vibrational modes of the various molecular fragments. The role of the donor–acceptor groups involved in the conjugation is evident: the energy difference between the frontier orbitals is lowered and the intensity of the light absorption and emission increases. Heterocyclic rings contribute to the rigidity of the molecular framework, while the heteroatoms actively participate in the conjugation. This explains the fact that among heterocycles there are many luminophores with practical applications. We now consider a number of these in detail.

9.3.2 Safety and Aesthetics

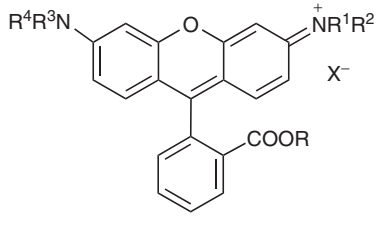
Bright dyes that stand out on illumination (e.g., by headlights) at night as well as in daylight are extremely useful. Such compounds are in demand for modern advertisements, decorative art, printing, textiles, road markings, aerodrome signs and navigation. Ordinary dyes are unsuitable

because their brightness is due to reflected light. By contrast, luminescent dyes not only reflect light but also transform a portion of the absorbed light into luminescent radiation. Fluorescence, in addition to reflected light, greatly improves the brightness and intensity of the radiation.

Cyanine dyes based on 3,3-dimethylindolinium (e.g., CI Cationic Rose 2C), rhodamines B and 6G (nitrogen analogues of fluorescein), and 1-alkylaminoanthrapyridones (Figure 9.12) are among the compounds frequently utilized as luminophores for the preparation of fluorescent dyes. Such luminophores are used in conjunction with a polymeric substrate, special adhesives and very often in combination with other dyes and luminophores. Thus, any dye color desired can be achieved with increased brightness.

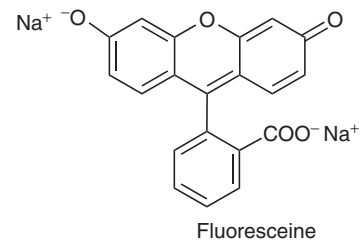


CI Cationic Rose 2C

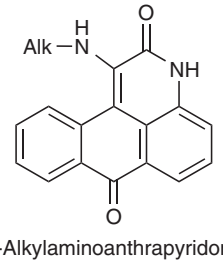


Rhodamine B (R = H, $R^1 = R^2 = R^3 = R^4$ =Et)

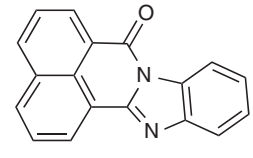
Rhodamine 6G ($R^1 = R^3$ = H, $R = R^2 = R^4$ = Et)



Fluoresceine
(disodium salt)



1-Alkylaminoanthrapyridone



1,8-Naphthoylene-1,2-benzimidazole

Figure 9.12 *Luminophores used for the preparation of fluorescent paints.*

Luminescent dyes are extensively utilized to color plastic materials and synthetic fibers. Such compounds are applied to clothes or insignia worn by road workers, mine workers, air force pilots and so on as a safety measure. Among the dyes used for this purpose, many derivatives of 1- and 6-aminoanthrapyridone and 1,8-naphthoylene-1,2-benzimidazole (Figure 9.12) can be found.

9.3.3 How to Convert White into Snow White

Perfect whiteness can rarely be achieved in the manufacture of linen, paper and plastic coatings as the starting materials often have a yellowish tint which frequently intensifies during use. This tint is

caused by the absorption of some long wavelength blue light because blue is the complementary color to yellow. To increase the whiteness of a material, one must manipulate it to reflect or radiate blue rays.[4] The procedure based on blue light reflection has a long history. In earlier times, ultramarine ('blue') or a small quantity of indigo carmine was added to the water during the washing of linen. The material thus treated ('blued') then reflected a small excess of blue radiation and appeared to be whiter. However, the whiteness thus achieved was far from ideal as the textile took on a grayish hue.

In modern times the application of so-called 'optical bleachers' has achieved the desired goal. Their effectiveness is a result of the emission of blue light by luminescence. Optical bleachers are in fact colorless fluorescent dyes. They absorb light in the near-ultraviolet range and re-emit it by fluorescence in the blue region of the visible spectrum. The application of such compounds creates the appearance of intense whiteness. It is familiar to us that white clothing appears luminescent in restaurants and clubs where soft blue light is used; this is a result of optical bleachers.

The majority of optical bleachers are derivatives of heterocycles, although their heterocyclic nuclei are not always responsible for the luminescence. Whiteners of the 4,4′-diaminostilbene-2,2′-disulfonic acid series are employed widely. The most useful of these contain triazine substituents at the amino groups (Figure 9.13, structure **1**). If chlorine atoms are retained in the triazine rings, the bleacher becomes tightly fastened to the fiber in a manner similar to that of the reactive dyes.

Optical bleachers are added to plastic materials or synthetic fibers at the temperature of the melt. Therefore, bleachers must be thermostable, as are, for example, whiteners constructed from the heterocyclic derivatives of stilbene (Figure 9.13, structures **1**–**3**). Coumarin and 1,3-diarylpyrazoline analogues (Figure 9.13, structures **4, 5**) are also useful whitening agents.

A familiar application of optical bleachers is their use in washing powders and other cleaning preparations.

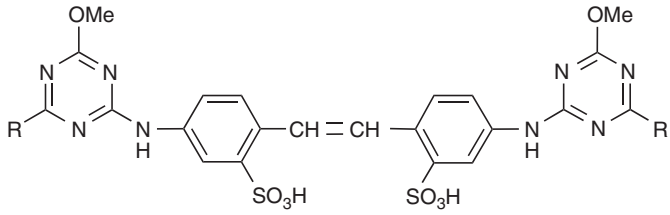
9.3.4 Markers and Tracers

Luminophores have proven to be irreplaceable as markers and tracers for a multitude of applications.[5] For instance, fluorescein is used in geological and hydrological studies to determine the directions of underground water flows and their connections to points of emergence above ground. Fluorescein, rhodamine and 1,8-naphthoylene-1,2-benzimidazole are used in luminescent flow detectors to locate microscopic fissures and other superficial damage in industrially manufactured metallic, ceramic and concrete materials. The article being tested is immersed in a luminophore solution for a period, then washed and dried. While the surface of the article may appear to be free of luminophore, a sensitive fluorescence detector allows ready detection of the compound in microcracks. A similar approach is employed in medical applications. Thus, a minute quantity of fluorescein is injected into the blood to check the permeability of blood vessels.

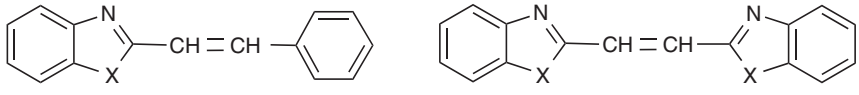
Organic luminophores have also been used in devices that register ionizing particle flux: α-, β- and γ-rays, neutrons and even neutrinos can be detected. In such cases luminophores serve as scintillators, that is, substances which produce short-lived flashes (scintillations) when struck by the ionizing particles. The number of scintillations produced is recorded by a photomultiplier. Scintillators are used in diverse branches of science and technology including nuclear and space

[4] A chemical method of eliminating yellowish tints in textiles involves treatment with an oxidizing agent. Unfortunately, the oxidation is not restricted to the yellow-colored bodies and thus some fiber damage is unavoidable.

[5] In recent decades organic dyes find increasing applications which are not directly connected with coloring textiles and other materials. Under the name *functional dyes* they are used in various fields of medicine, biotechnology and techniques ranging from photodynamic therapy (see Section 7.4.3) to lasers, solar cells, biomedical sensors, optical data storage and so on. Many such applications are discussed in this chapter and in Chapter 11.

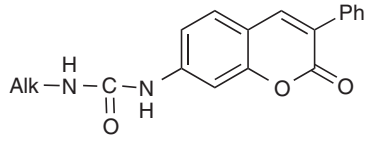


1 (R = Cl, C_6H_5NH)

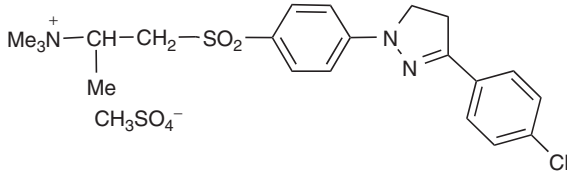


2 (X = NH, O, S)

3 (X = NH, O, S)



4



5

Figure 9.13 *Examples of optical bleachers.*

research. Among heterocyclic scintillation agents, 2,5-diphenyloxazole, 2-phenyl-5-(4-biphenylyl)-1,3,4-oxadiazole and 1,3,5-triphenyl-Δ^2-pyrazoline (Figure 9.14) dissolved in organic solvents are the most often used as liquid scintillators. As such, these solutions can be prepared in virtually unlimited volume, and so luminophores are especially useful in detecting the presence of particles at very low flow densities.[6]

The use of heterocyclic luminophores as powerful imaging agents in biomedical studies deserves separate mention and is covered in the next section.

9.3.5 Imaging and Diagnostic Agents

For many decades scientists could only dream of directly watching how processes in living organisms proceed on a molecular level. Nowadays, this is rapidly becoming a reality and heterocyclic luminophores provide here a powerful impetus. Numerous fluorescent markers capable of selective attachment to nucleic acids, lipids, polysaccharides, antibodies, cellular membranes, damaged cells

[6] Solid scintillators are also made in the form of single crystals of certain organic luminophores such as anthracene, stilbene, diphenylacetylene and others, their solid solutions in polymers can also be used. These types of scintillators, in addition to their inorganic counterparts, are often used to measure flows of high energy radiation.

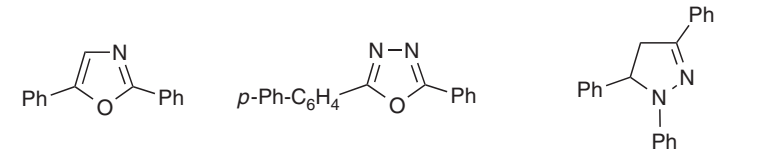


Figure 9.14 *Heterocyclic luminophores used as scintillators.*

and so on have been synthesized. Modern techniques allow selective coloration (labeling) of the corresponding biological targets with bright fluorescent dyes and their subsequent study, mainly by means of fluorescent microscopy. Cell imaging, as this approach is called, allows watching cellular events on a real time scale: gene and protein expression, protein–protein interactions, cell membrane forming and many others. It also helps to clarify drug targeting and the mechanism of drug bioactivity that is useful for further optimization of drug structure. A classical example of the application of heterocyclic luminophores in medical diagnostics is the luminescent dye acridine orange (Figure 9.15a). It 'marks' healthy and cancerous cells differently and therefore has been used in the diagnosis of malignant tumors.

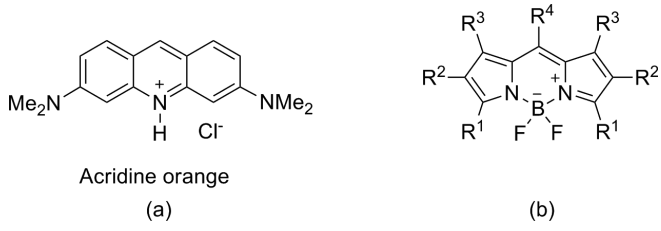


Figure 9.15 *Some heterocyclic luminophores used for imaging in cell biology.*

The great majority of imaging agents used in biomedical studies are highly polar polynuclear heterocyclic compounds. Apart from the just mentioned acridine orange dye, they include fluoresceines, rodamines (Figure 9.12) and derivatives of the dipyromethene-BF_2 system (Figure 9.15b; commonly abbreviated to BODIPY) and others. The direct coloration of biological tissues with simple dyes, as such, is rarely selective. To achieve higher selectivity a fluorophore is commonly supplied with a reactive functionality by means of which it can be attached covalently to a biological target. Hundreds of such functionalized fluorophores are now commercially available. To exclude any cell damage, a reaction leading to fixation of a fluorescent probe should not only be selective but also proceed rapidly, in high yield and under mild conditions. For this kind of chemistry the term 'click chemistry' has been introduced. Few click chemistry reactions satisfy all the above demands. One, purely heterocyclic, is the 1,3-dipolar cycloaddition of acetylenes to azides resulting in the formation of 1,2,3-triazoles (Figure 9.16a). Normally, this process goes with a moderate rate and requires the use of a cytotoxic organic solvent and a copper catalyst. However, the reactivity of the triple bond is known to increase strongly when it becomes angle-strained and activated by electron-withdrawing groups such as fluorine atoms in 3,3-difluorocyclooctyne (DIFO) molecule.

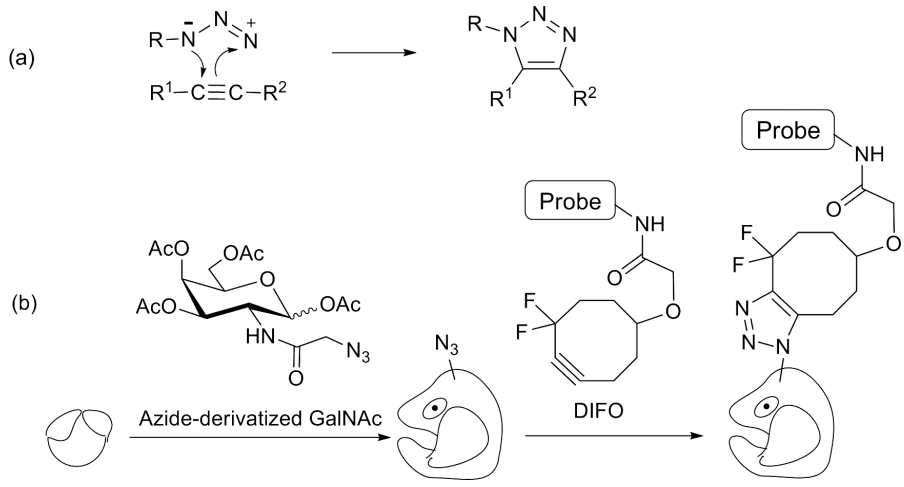


Figure 9.16 *1,3-Dipolar cycloaddition reaction of acetylenes to azides: (a) general scheme, (b) its use for fluorescent cell surface labeling in zebrafish embryos (From Laughlin, S. T., Baskin, J. M., Amacher, S. L. and Bertozzi, C. R.,* Science*, 2008,* ***320****, 664. Reprinted with permission from AAAS).*

Unlike ordinary alkynes, reactions of DIFO with azides in the absence of copper still proceed within minutes. For example, recently this reagent has been successfully employed to visualize the development of glycan cell membranes in zebrafish embryos (Figure 9.16b). Researchers first introduced azide-derivatized N-acetylgalactosamine (GAINAc) metabolically into developing embryos. Then the embryos were treated with a DIFO-containing fluorescent reagent. Finally, fluorescent microscopy was used to watch the distribution of sugars on cell surfaces and the accompaning phenomena: cell adhesion, migration, and recognition processes.

A revolutionary achievement in imaging technique and generally in biotechnology was the discovery by O. Shimomura in 1962 of the so-called green fluorescent protein (GFP). This protein composed of 238 amino acids was isolated from ocean deep-water *Aequorea victoria* jellyfish. A critical section of GFP structure is the tripeptide Ser65-Tyr66-Gly67 sequence. Three amino acid residues arranged as shown in Figure 9.17a spontaneously cyclize into 4-(*p*-hydroxybenzylidene)-imidazoline-5-one chromophore (Figure 9.17c). The first stage of the conversion represents nucleophilic addition of a glycine NH group to a serine carbonyl group accompanied by the loss of a water molecule. The thus formed 4-(*p*-hydroxybenzyl)imidazoline-5-one intermediate (Figure 9.17b) on the second stage undergoes oxidation that puts into conjugation the imidazoline C = N bond and a phenolic OH group. Due to enhanced acidity of the OH group the chromophore exists in two forms: neutral and anionic (Figure 9.17c, d). They are responsible for two excitation bands in absorption spectrum at 397 and 475 nm, respectively. The fluorescence emission providing green glow originates from the phenolate species and occurs at $\lambda_{max} = 504$ nm.

The source of light excitation for a luminophore in dark ocean depths is an intriguing question. The answer is that a second photoprotein called aequorin is present in *A. victoria* jellyfish. Aequorin also emits light due to a chemical reaction, rather than photoexcitation. This phenomenon, widespread among living creatures, is called bioluminescence (chemoluminescence).

Aequorin is composed of two distinct units, the apoprotein and the prosthetic group coelenterazine. Coelenterazine is a simple derivative of imidazo[1,2-a]pyrazine (Figure 9.18a). The aequorin apoprotein is able selectively to bind Ca^{2+} ions that triggers conformational changes and initiates

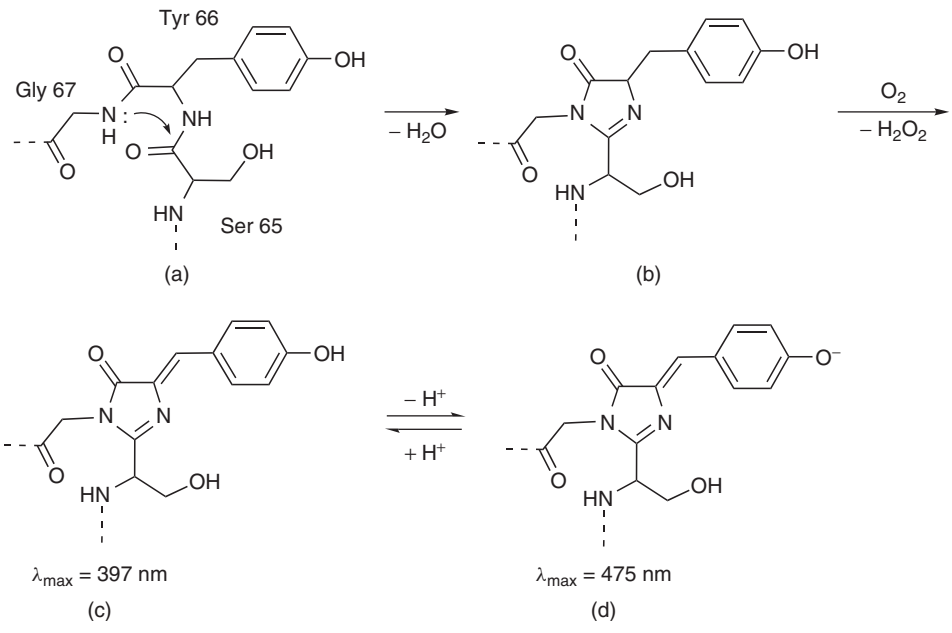


Figure 9.17 *A proposed mechanism for the formation of imidazole-5-one fluorophore in GFP (adapted with permission from Tsien, R. Y.,* Annu. Rev. Biochem.*, **67**, 509. © 1998 Annual Reviews).*

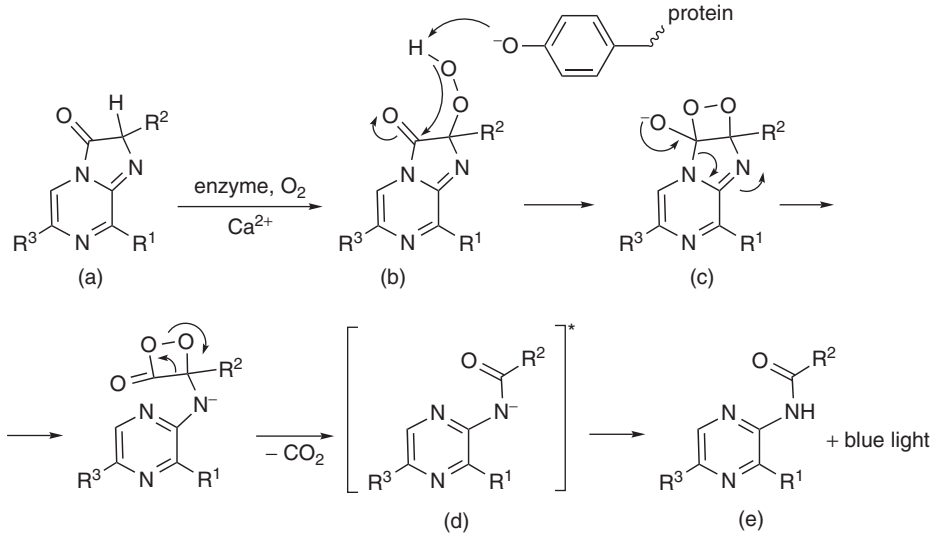


Figure 9.18 *A proposed mechanism for the bioluminescence of aequorin protein ($R^1 = C_6H_5CH_2$, $R^2 = 4\text{-}HOC_6H_4CH_2$, $R^3 = 4\text{-}HOC_6H_4$).*

oxidation of the coelenterazine. The process is catalyzed by the specific enzyme luciferase. First, an O_2 molecule is inserted into the C3-H bond of the imidazole ring to afford a hydroperoxide (Figure 9.18b). By means of basic catalysis with participation of a protein phenolate site, the hydroperoxide is then cyclized into dioxetane (Figure 9.18c). Dioxetane intermediates are participants in practically all bioluminescence reactions in living organisms. Dioxetanes are energy rich compounds due to the instability of the four-membered ring and weakness of the O–O bond. In the case of coelenterazine the oxetane is stabilized via loss of a CO_2 molecule with the formation of an excited coelenteramide (Figure 9.18d). Transition from this excited state into the ground electronic state (Figure 9.18e) is accompanied by the emission of blue light with $\lambda_{max} = 469$ nm. Since in jellyfish tissues aequorin and GFP form a tight complex, this blue light is intercepted by GFP causing the green luminescence.

A turning point in GFP studies was 1992, when the gene responsible for its formation was sequenced. Soon after biologists could introduce the GFP gene into cells of many live organisms and observe their green fluorescence. Thus, this ability of the GFP gene to serve as a reporter of gene and protein expression was established. Simultaneously, monitoring tagged proteins with fluorescence microscopy provided scientists with a powerful instrument for the study of the dynamics of RNA–protein and protein–protein interactions, protein folding and so on. It is noteworthy that, while most small fluorescent molecules such as fluoresceines are rather phototoxic for live cells, GFP is much less harmful.

Beginning at the end of the 1990s, several other GFP-like natural (wild) fluorescent proteins with yellow, red and cyan color (Figure 9.19a, b) have been discovered. They were also isolated from various marine creatures, mostly reef-building corals. At the same time, numerous color mutant fluorescent proteins have been engineered in the laboratory (Figure 9.19c–e). While all known wild fluorescent proteins differ only by a substituent in position 2 of the imidazoline ring, in mutant proteins both 2- and 4-substituents are variable. In 2008 three American scientists, M. Chalfie, O. Shimomura and R. Tsien, were awarded the Nobel Prize in chemistry for their discovery and development of the green fluorescent protein.

The diversity of fluorescent proteins has opened a door to impressive innovations in imaging technique. Perhaps, the most notable is the so-called ‘brainbow’ process in which individual neurons of an animal’s brain are mapped with fluorescent proteins, sometimes up to 100 different colors. As a result, it becomes possible to observe direct brain circuitry. There are many other uses of fluorescent proteins, some of them rather curious. Thus, breeding and marketing of fluorescent aquarium fishes, pets and various animals (mices, pigs, rabbits, etc.) is now a growing business.

In summary, chemical evolution has led to the appearance in some marine organisms of a special kind of tripeptide sequence capable of producing a unique imidazoline-4-one chromophore. The latter served as a source for a light signal with no fully understandable purposes. We are not grossly exaggerating by saying that man’s use of fluorescent proteins with completely different aims is just another whimsical turn of the evolution spiral.

9.3.6 Lasers Containing Heterocyclic Luminophores

When we speak of ‘lasers’, our imagination conjures up pictures of bright, highly focused beams of light capable of cutting metal, initiating thermonuclear synthesis, rapidly reading stored information (compact disks, magnetic cards, etc.), transferring information along glass filaments, determining distances between distant objects (e.g., the Earth and moon), surgically incising live tissue and so on. These and other uses of lasers have revolutionized modern industry, electronics and information handling, as well as medical and scientific research.

A laser is a source of light, or indeed of any form of electromagnetic radiation. Candles and electric lamps are also sources of electromagnetic radiation, but the difference is that they produce

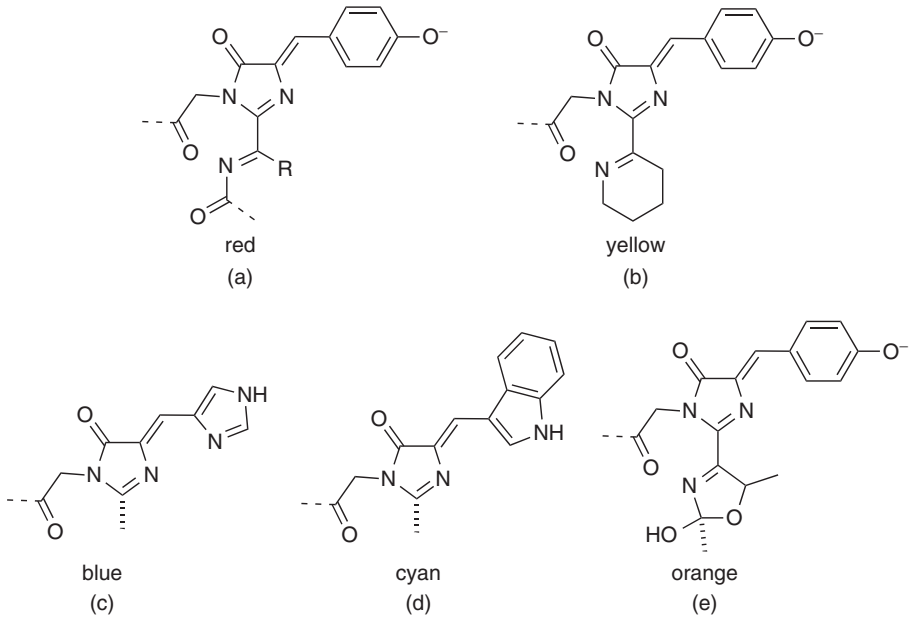


Figure 9.19 *Selected examples of natural (a, b) and engineered (c–e) fluorescent protein chromophores.*

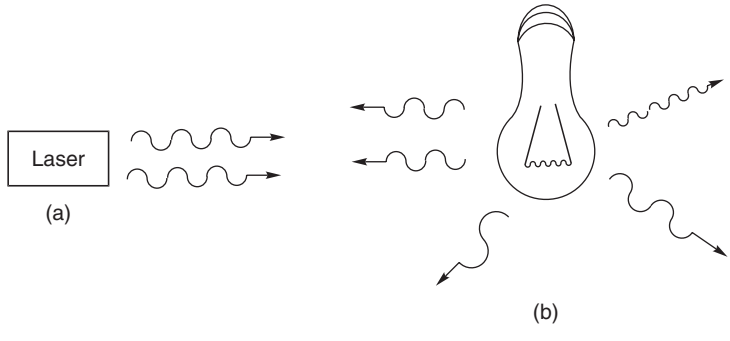


Figure 9.20 *(a) Coherent and (b) noncoherent light sources (Reprinted with permission from Kovalenko, L. J., and Leone, S. R.,* J. Chem. Educ.*, 1988,* ***65****, 681. © 1988 American Chemical Society).*

noncoherent light, that is, radiation composed of photons of different frequencies and direction or phase. Laser beams, on the contrary, are characterized by perfect coherence (Figure 9.20).

In a laser, some other form of energy is converted into coherent light. The substance which transforms the energy is called the laser active medium. Lasers are classified as gas, solid or liquid lasers, according to the aggregation state of their active medium. In gas lasers the active medium consists of atoms, ions or molecules of various compounds. In solid lasers, rare-earth metals or chromium(III) (ruby laser) salts in the crystalline or glassy form are used, and the excited ions serve as the source of radiation. The active medium in liquid lasers is composed of a solution of an organic compound. Complex compounds of rare-earth ions with various organic ligands have long been used as active media. However, in the late 1960s it was found that laser effects could

be achieved by organic luminophores alone, especially heterocyclic derivatives. We first examine the mechanism of laser action and then discuss structures.

For a laser to function, the active medium has first to be excited. This process, called laser pumping, is carried out by an additional energy source, such as an impulse lamp. There are three types of interactions between matter and light. We have already discussed two of these: photon absorption and spontaneous emission (see Sections 6.1 and 9.3.1). The third, called stimulated emission, is the type of interaction on which laser action is based. When a photon, emitted during fluorescence with a frequency $h\nu'$, acts upon another excited molecule, the emission of a second photon of the same frequency, direction and phase occurs (Figure 9.21). The process is propagated in a manner similar to a chain reaction. Two emitted photons now stimulate the emission of four photons and so on; thus, the emission of coherent radiation increases many-fold. This phenomenon is summarized by the acronym 'laser': light amplification by stimulated emission of radiation.

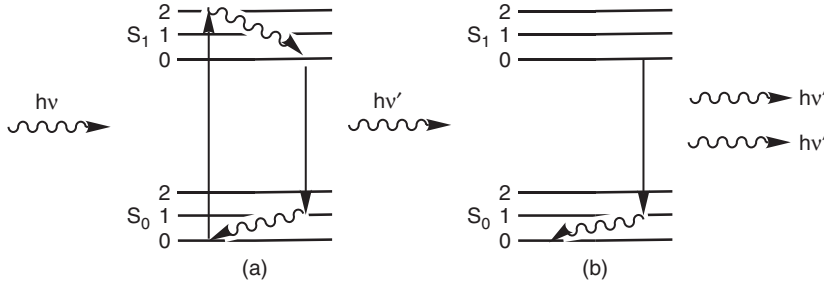


Figure 9.21 *Stimulated emission of radiation: (a) molecular absorption of a photon with frequency hν and spontaneous emission of a photon with frequency hν′ by fluorescence, (b) emission of second photon with frequency hν′ stimulated by the initial photon's effect on a second excited molecule.*

Under stimulated emission, the excited molecules radiate photons and are converted to the nonexcited state S_0. Of course, the process terminates when no excited molecules remain and the radiation ceases. To prevent this, the population of the excited state $S_1(0)$ has to be maintained at a sufficiently high and stable level. Such a state, an inverted population (Figure 9.22), is achieved by: (i) systematic pumping of the active medium, together with (ii) the construction of a system of mirrors which directs the majority of the laser radiation to reconvert molecules to the excited state.[7]

Rather stringent demands are made on the compound used in the laser. The luminophore must achieve a high quantum yield of fluorescence and must be photostable. Moreover, the absorption band should not overlap with the fluorescence band appreciably. In rhodamine 6G (Figure 9.23) only minor overlap occurs, and rhodamine 6G is therefore suitable as a laser dye.

Figure 9.22 shows that a substance capable of generating laser radiation should possess at least four appropriate energy levels between which the necessary transitions can take place. Additionally, relaxation of the molecules from the S_1 to S_0 state should be relatively slow, whereas the passages $S_1(2) \rightarrow S_1(0)$ and $S_0(1) \rightarrow S_0(0)$ should occur rapidly.

These requirements are met by a number of heterocyclic luminophores which are usually classified by the wavelengths at which they generate laser radiation (Table 9.1). The greatest generation of energy (radiation in the near-ultraviolet and blue ranges of the spectrum, i.e., 300–400 nm) is achieved by 2,5-diaryloxazoles, 2,5-diaryl-1,3,4-oxadiazoles and 2-arylbenzoxazoles.

[7] The chamber in which the laser radiation is generated contains an outlet for a portion of the radiation for the required purpose.

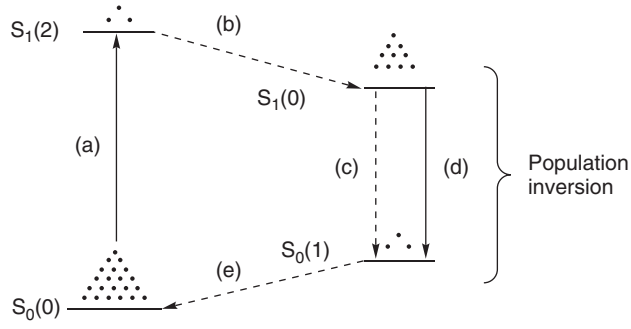


Figure 9.22 *A four-level laser (the marking of energy levels is conventional and correlates to that of Figure 9.21; each point in the figure represents one molecule): (a) excitation, (b, e) rapid relaxation, (c) slow relaxation, (d) laser radiation (Reprinted with permission from Kovalenko, L. J. and Leone, S. R.,* J. Chem. Educ.*, 1988, **65**, 681. © 1988 American Chemical Society).*

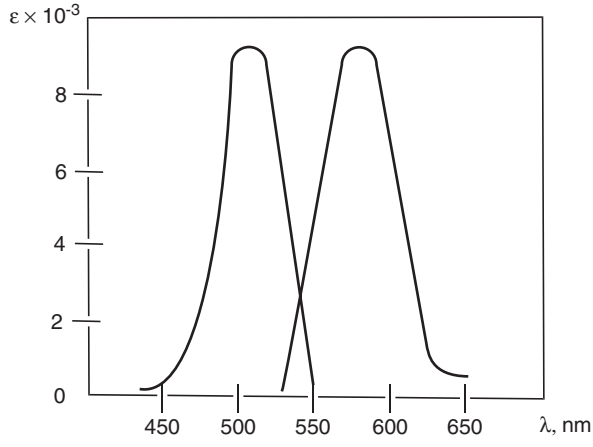
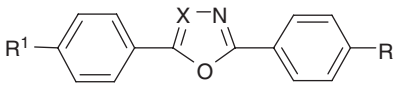
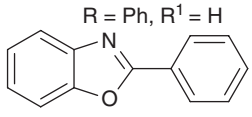
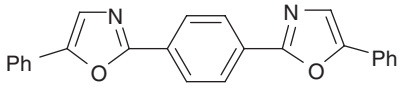
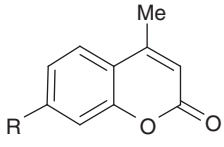
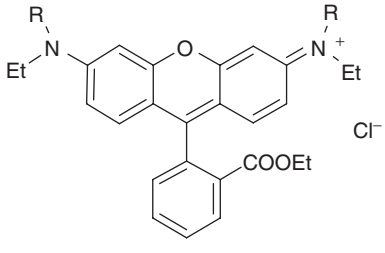
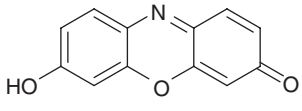
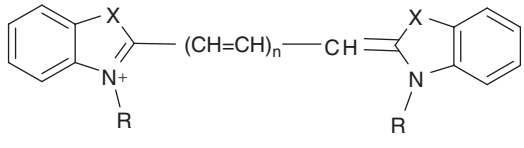


Figure 9.23 *Spectra of absorption (left) and fluorescence (right) for rhodamine 6G (in ethanol, $\lambda_{exc} = 365$ nm). Adapted from Krasovitskii, B. M., and Bolotin, B. M.,* Organic Luminescent Materials*, VCH, N.Y., p. 287. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.*

Compounds of the 1,4-bis(5-phenyloxazol-2-yl)benzene type generate radiation at somewhat greater wavelengths (400–440 nm), still within the blue region. Coumarins, especially their 7-hydroxy and 7-dialkylamino derivatives, are extensively used as luminophores in the blue and green regions of the electromagnetic spectrum.

Xanthene dyes such as rhodamines 6G and 3V and 7-hydroxy-3*H*-phenoxazin-3-one are almost ideal for generating laser radiation in the red region of the spectrum (520–650 nm). Organic compounds which lase in the infrared region (800–1200 nm) are also known. Such compounds are usually cyanine dyes with long conjugated polymethine chains. The opposing termini of the chains contain strong electron-donating and electron-withdrawing groups, usually heteroaromatic cations and their partially reduced π-excessive counterparts. Indoline, benzoxazoline, benzthiazoline, pyrylium and thiapyrylium nuclei are the heterocycles most frequently incorporated into the polymethine chain. Dipyrrometheneboron difluoride derivatives (Figure 9.15b) were also suggested as dye lasers for the red and infrared regions.

Table 9.1 *Some heterocyclic compounds used as laser active media*

Structural formula	Name	Wavelength of radiation generated (region)
 X = CH, N; R = R^1 = Ph	2,5-Diaryloxazoles 2,5-diaryl-1,3,4-oxadiazoles	300–400 nm (near-ultraviolet and blue)
 R = Ph, R^1 = H	2-Arylbenzoxazoles	300–400 nm (near-ultraviolet and blue)
	1,4-Bis(5-phenyloxazol-2-yl) benzene	400–440 nm (blue)
 R = OH, NEt_2, etc.	Substituted coumarins	440–500 nm (blue and green)
 R = H (Rhodamine 6G) R = Et (Rhodamine 3V)	Xanthene dyes	520–650 nm (red)
	7-Hydroxy-3*H*-phenoxazin-3-one and other oxazine dyes	500–800 nm (red)
 X = NPh, O, S, Se, CMe_2	Polymethine dyes	800–1200 nm (infrared)

Lasers based on organic dyes are very versatile. The width of the luminescence band in conjugated organic compounds, which may approach 200 nm (Figure 9.23) allows the frequency of the laser radiation generated to be tuned smoothly over the operating range, while gases and most solid state lasing media lack this ability. Additionally, dye lasers can offer very large pulsed energies and high average power. Their main disadvantage is their relatively low photostability that demands more complicated instrument construction.

9.4 Color Change Compounds

In the 1970s the first self-darkening eyeglasses appeared on the market. Under the name 'chameleons' they quite soon became very popular. In the shade or indoors their lenses remained clear but when exposed to sunlight they rapidly darkened, thus protecting eyes. Primarily, the 'chameleons' were produced from glass which contained embedded silver halide crystals. With light having wavelengths of 320–400 nm, the electrons of the glass combine with the colorless silver cations with the formation of elemental silver. Since the latter is visible, the lenses appear darker. The process is reversible: in the shade the silver gives electrons back and the lenses return to a clear state.

Glass lenses, however, have some disadvantages. Thus, they are rather heavy and on long use, as in the case of car drivers, create inconveniences. For this reason, producers of 'chameleons' turned to plastic lenses. Silver halides do not mix with organics and therefore cannot be employed in plasic lenses. This is why photoactive organic compounds (photochromes) came to the foreground. Remarkably, practically all photochromes used for this purpose were heterocyclic compounds. For many years spironaphthoxazines (Figure 9.24a) held the leading position. In sunlight they are reversibly converted from the colorless spirocyclic form **6** into acyclic merocyanine form **7** possessing a cyan or a dark-blue color. Another widely used class of photochromes is

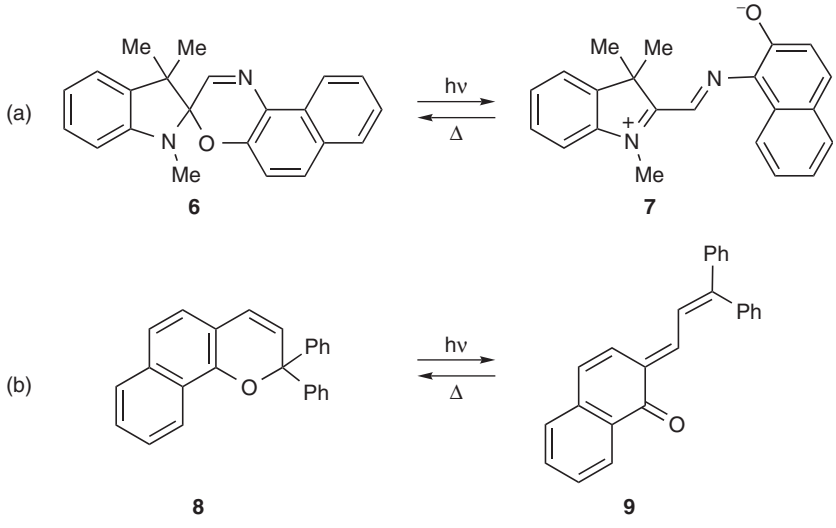


Figure 9.24 *Heterocyclic photochromes employed in the production of self-darkening lenses (Reprinted with permission from Crano, J. C., Flood, T., Kumar, A. and Van Germert, B.,* Pure Appl. Chem.*, 1996,* ***68****, 1395. © IUPAC).*

diarylnaphthopyrans which interconvert between colorless pyran **8** and colored ortho-quinoid **9** forms (Figure 9.24b).

Along with light-sensitive heterocyclic photochromes, there exist electrochromes, mechanochromes and thermochromes. They change their color in response to an electric impulse, mechanical force or temperature variations. Such compounds also have numerous practical applications. Thus, among electrochromes the already-mentioned viologens, 1,1′-dialkyl-4,4′-bipyridinium salts, are the most widespread (Figure 8.3). Under ordinary conditions they are colorless, but on action of a certain electric potential they can be reversibly reduced into stable blue-violet cation-radicals. Viologens and some other electrochromes are employed in compact displays, rear-view car mirrors and so-called smart windows regulating the amount of heat and light coming inside.

In contrast, indoline nitrospiropyrans display the property of mechanophores. Being attached to a solid polymer (Figure 9.25) they remain colorless when the polymer structure is intact. However, when structural deformations arise, the pyran ring undergoes ring-opening into the merocyanine form that is accompanied by red coloration. Actually, we have here a damage sensor signaling failures in polymer construction materials.

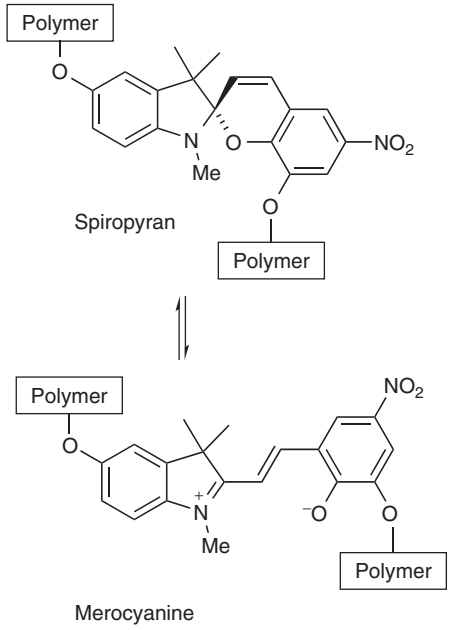


Figure 9.25 *Heterocyclic mechanophors (Reprinted by permission from Macmillan Publishers Ltd. Davis, D. A., Hamilton, A., Yang, J., Cremar, L. D., Van Gough, D., et al.,* Nature, ***459****, 68. © 2009).*

Quite a number of interesting devices are based on thermochromism. One of the very recent inventions in this field is thermoprinters for providing instant copies of digital photos. Such printers require no replaceable cartridge. Photographic paper is covered by separate layers of colorless microcrystals of specially chosen thermochromes or their mixtures and all are covered with a clear finish. Each of these layers and components melts at its own temperature, causing the formation of a colored form. Most appropriate for this purpose are the above-mentioned rhodamines and rodoles that are colorless in the ring-closed lactone form and colored in the ring-opened carboxylic acid form (Figure 9.26).

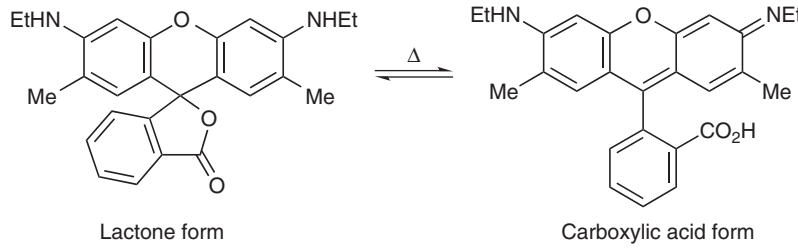


Figure 9.26 *Thermochromic conversion of rhodamine dye (Reprinted with permission from Halford, B.,* Chem. Eng. News*, 2007,* ***85*** *(37), 34. © 2007).*

Such printers create within fractions of a millisecond a full pallete of colors via a rapid succession of precisely tuned thermal pulses from its tiny heaters, each of which melts just enough of each crystalline layer to produce the proper shade. Because the crystals soften only at a high temperature – somewhere around 80 °C for the lowest-melting crystals – the images have high color stability.

The above interconversions illustrate how typical for heterocyclic compounds is the phenomenon of ring-chain tautomerism. In the next sections, we turn more attention to color change in compounds, particularly in connection with their use for writing and storage of information.

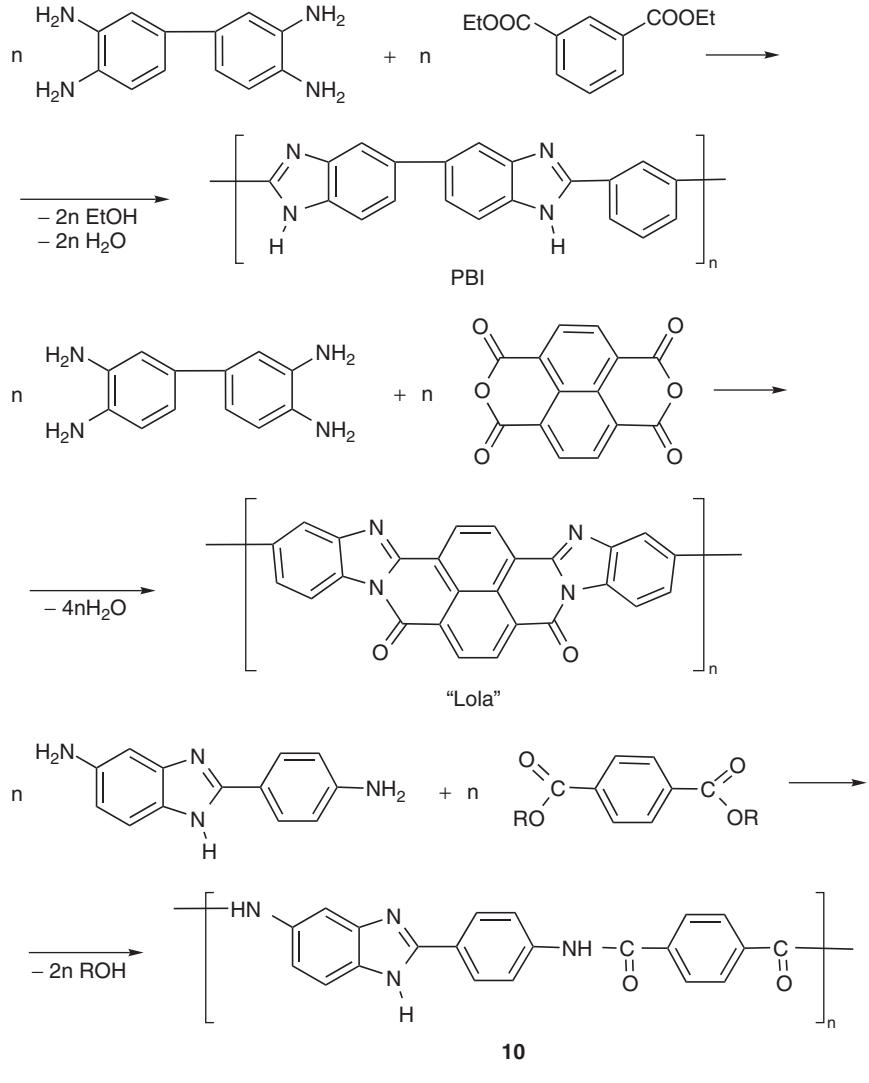
9.5 Fire Retardancy

We cannot guarantee the authenticity of the story which follows, but it seems to be an appropriate introduction to this section. At a conference, the general manager of a chemical fibers factory, after lighting a cigarette, casually used his burning lighter to ignite the necktie of an old friend. When the horrified friend recoiled, the manager said with feigned surprise, 'Mine doesn't burn. It's nonflammable,' and immediately proved his claim. In this way he succeeded in advertising his company's new product. The case was particularly impressive as the 400 °C temperature of the cigarette lighter flame instantly scorched silk, wool and cotton, and melted nylon and many other fibers. The manager's tie fabric was woven from a special type of chemical fiber distinguished by its exceptional thermostability. Obviously, the manufacture of neckties is not the major application intended for such fibers; the necktie was selected only to distinguish the manufacturer's fiber from others. However, heat resistant fibers have become essential in the modern world for the production of fireproof clothing for firemen, welders, foundrymen, pilots and astronauts. Such fibers are also utilized in the manufacture of parachutes, conveyer belts, heat-insulating material and asbestos replacements. The development of heat resistant fibers has significantly influenced progress in aerospace and aeronautics technology.

We now discuss the composition of such thermostable compounds. Chemically, these substances are polymers composed of aromatic and heteroaromatic residues. Polybenzimidazoles, and to some extent polyquinoxalines, have become especially useful. Benzimidazole and quinoxaline are highly stable molecules which do not decompose at temperatures up to 600 °C (at which temperature they exist as gases). As constituents of polymer chains, these two nuclei render the macromolecule inherently highly heat resistant and stable. The presence of N—H bonds and amide (NHC═O) groups in such polymers imparts additional durability.

In the United States, the Celanese company produces several thousand tonnes each year of the polybenzimidazole known as PBI by the condensation of 3,3′, 4, 4′-tetraaminobiphenyl with diethyl isophthalate. Another heat resistant polybenzimidazole, used in Russia to manufacture

a fiber named 'Lola', is synthesized by the condensation of tetraaminobiphenyl with the dianhydride of naphthalene-1,4,5,8-tetracarboxylic acid. The preparation of polymers from readily available benzimidazole-based monomers is also possible. In particular, very strong fibers can be obtained from polymer (**10**), which is itself formed from the polycondensation of 5-amino-2-(*p*-aminophenyl) benzimidazole with derivatives of aromatic dicarboxylic acids, such as terephthalic esters (Figure 9.27). Articles made from polybenzimidazole fibers retain their properties during prolonged periods at 200–300 °C in the presence of air. At temperatures from 300 to 350 °C their heat resistant properties are maintained for 24 h, while at 400–450 °C several hours is the limit. Such fibers exceed the heat resistance of sodium and potassium silicate glasses.



10

Figure 9.27 *Thermostable polybenzimidazoles.*

Thermostable polyquinoxalines are formed from the reaction of 1,2,4,5-tetraaminobenzene with aromatic bis-α-diketones (Figure 9.28). The resulting fibers exhibit high resistance toward corrosive agents and heating at temperatures up to 300 °C. Further thermostable fibers have been synthesized which contain other heterocyclic systems including benzoxazole units.

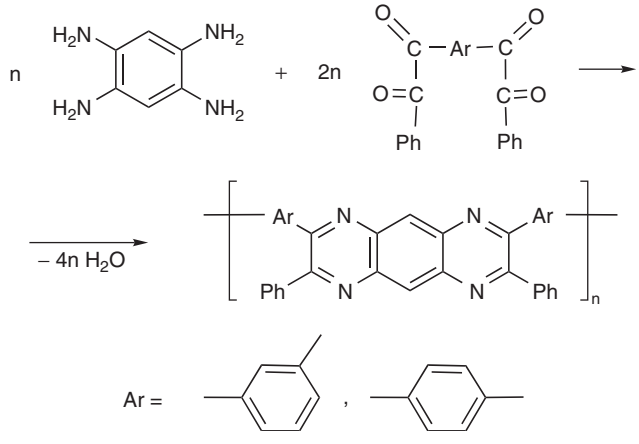


Figure 9.28 *Synthesis of thermostable polyquinoxalines.*

9.6 Photographic Materials and Recorders of Information

Recently, traditional silver halide photography has been seriously challenged by its digital counterpart. Accordingly, the need for photodyes and auxiliary materials for photography has decreased. Nevertheless, film photography is still in use, since it has some advantages in relation to wavelength range, color and exposition. The versatile array of chemicals available for use in photographic processes is largely based on heterocyclic compounds. Silver bromide crystals, which are usually used as the photosensitive material, frequently need to have their sensitivity to light enhanced by the addition of silver sulfide. As a result of such activation the unexposed sites of the film can become slightly exposed causing haziness in the photograph. To avoid this, one must add antifogging substances to the standard photographic regimen of chemicals employed nowadays. Salts and complexes are formed with the silver particles which appear in the unexposed regions of the film, thus eliminating fogging and restoring clarity and purity to the picture. Most contemporary antifoggants and photoemulsion stabilizers are heterocyclic in nature. Examples include benzotriazole, 5-nitrobenzimidazole, 1-phenyltetrazole-5-thione and 5-hydroxy-7-methyl-1,2,4-triazolo[2,3-*a*]pyrimidine (Figure 9.29).

A further disadvantage of silver bromide is that it is sensitive only toward the blue and violet region, i.e. the higher energy radiation of visible light. To widen the range of sensitivity to include the entire visible region of the spectrum, one must add so-called optical sensitizers photoemulsions. The photosensitizers are usually cyanine dyes, such as pseudocyanine or pinacyanol discussed earlier (Figure 9.6). The former enhances the sensitivity of the photoemulsion towards blue and green light, while the latter has a similar effect with respect to red light.

Heterocycles are also utilized as auxiliary compounds during other steps in the photographic process. Thus, phenidone, a derivative of pyrazole, is used in combination with hydroquinone as

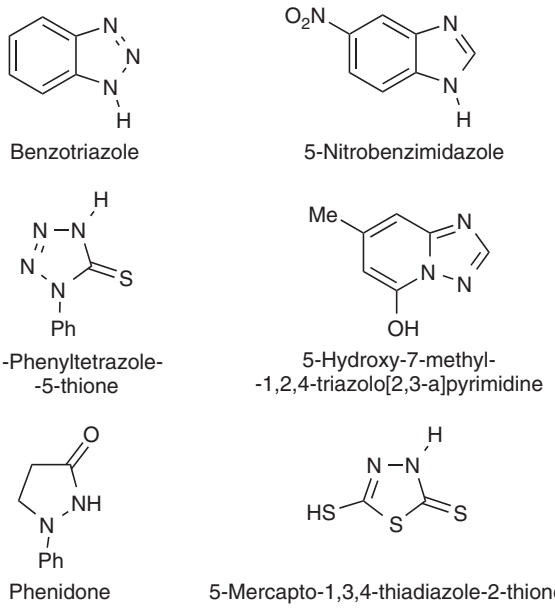


Figure 9.29 *Heterocyclic compounds used in photography (predominant tautomeric forms are shown).*

a developer, and 5-mercapto-1,3,4-thiadiazole-2-thione (Figure 9.29) is an effective toner. Toners form colored compounds with the silver particles; therefore, by careful selection from a range of compounds, black, brown and other tints can be emphasized.

New silver-free reprographic materials are eagerly sought, in part because of a shortage and increased price of silver. A classic silver-free reprographic printing process is based on the application of light sensitive diazo compounds. When exposed to light, the diazo species decompose with the liberation of nitrogen. If the exposed print is further processed with a phenol, the unexposed sites undergo azo coupling of the phenol and the unaffected diazonium salt. The final result of this conversion is the formation of an azo dye which provides color to the picture. Correspondingly, the exposed sites remain colorless. The light sensitive diazo compounds are generally homoaromatic, and therefore we will not discuss them further.

Heterocyclic compounds play a leading role in the creation of another type of silver-free reprographic material which has attracted the close attention of scientists over the past 30 years. This material is based on photochromic substances which reversibly change color by the action of light (Section 9.4). Indolylspiropyrans, one example of which has been already discussed (Figure 9.25), are distinguished by their quality and effectiveness among the various classes of organic photochromes. A closely related example is compound **12** shown in Figure 9.30, which is produced from the condensation of 1,2,3,3-tetramethylindolium iodide with 5-nitrosalicylic aldehyde. On irradiation with light of the appropriate wavelength, spiropyran **11** is transformed into the intensely purple valence isomer **12**.[8] Compound **12** is sufficiently stable for the color of a picture to be preserved over long periods. The reverse reaction (i.e., closure of the pyran ring to give the colorless

[8] Valence isomers are substances which interconvert as a result of electron and bond shifts (in contrast to other types of isomerism and tautomerism in which atoms or groups migrate).

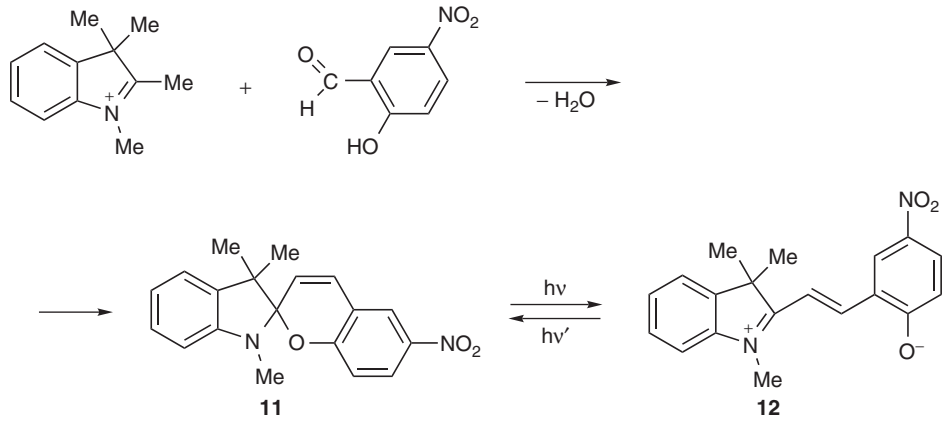


Figure 9.30 *Indolinospiropyran **11** and photoisomerization into the colored form 12.*

form **11**) is possible only upon repeated irradiation of isomer **12** by a powerful light source of a different wavelength.

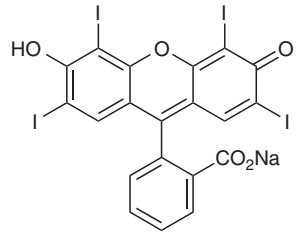
Despite their comparatively low light sensitivity, photochromes possess some remarkable advantages over silver halide-based materials. Firstly, photochromic transformations have a molecular nature that provides exceptionally high clarity and large information storage capacity to the image. Thus, the use of micro-imaging enables the information contained in a large library to be stored in compact format. For example, the 1245 pages of a Bible were reduced to about 6 cm^2. Another advantage is that as a clear color image of the subject is received immediately following exposure, the need for the traditional operations of developing and fixing is eliminated. Last but not least, photochromic materials can be highly economical as in some cases they can be utilized repeatedly after deletion of the original image.

9.7 Heterocycles as Food Additives

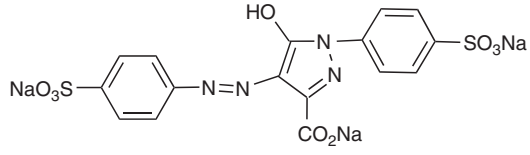
Various heterocyclic compounds are now used in the food industry as dyes, preservatives, bioadditives, aromas and flavorings. For example, of just about ten dyes cleared for use in the United States in drugs, food products and cosmetics, the following are heterocycles: red erythrosine, yellow tartrazine and blue indigo carmine (Figure 9.31).

Numerous additives for food products have a sweet taste. Most of them are sugars. The roman writer Plinium of the first century AD noted in his encyclopedia of natural sciences a sweet compound which was pressed out of sugar cane and used as a medicine. Cane cultivation (*Saccharum officinarum*) started in the twelfth century in the island of Sicily and in the sixteenth century in the Carribbean (mainly in Cuba). Saccharose or sucrose (Figure 9.32) from beet root (*Beta vulgaris*) was first obtained in the eighteenth century. The production of sucrose in the twentyfirst century rose to over 100×10^6 t $year^{-1}$. Such sweeteners as glucose (25% less sweet than saccharose) and fructose (180% the sweetness of saccharose) are much less deleterious to health in comparison with saccharose. At present more than 10×10^6 t $year^{-1}$ of glucose–fructose syrups and pure crystalline glucose and fructose are produced worldwide from starch chemically and biochemically (using the immobilized enzyme glucoamilase).

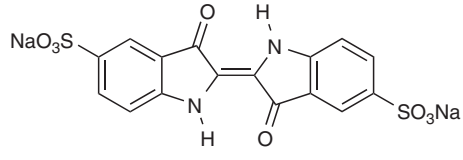
By partial substitution of hydroxyl groups by chlorine atoms in galactosaccharose a sweetener of a new generation – sucralose – was obtained (Figure 9.32), which is 600 times sweeter than



Erythrosine

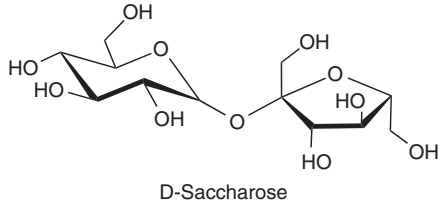


Tartrazine

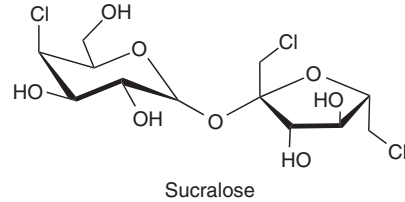


Indigo carmine

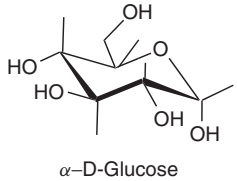
Figure 9.31 *Heterocyclic compounds used as food dyes.*



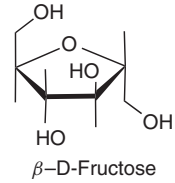
D-Saccharose



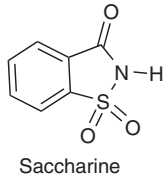
Sucralose



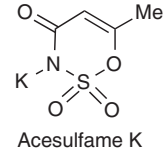
α–D-Glucose



β–D-Fructose



Saccharine



Acesulfame K

Figure 9.32 *Heterocyclic compounds used as sweeteners.*

saccharose. Sucralose is a zero-calorie sugar substitute, which is harmless towards teeth and stomach. Figure 9.32 shows two more very well known sweet heterocyclic substitutes for saccharose: saccharine (400 times sweeter than saccharose) and acesulfame K (200 times sweeter). Neither influences the insulin level in diabetic patients.

Heterocyclic vitamins C and E (Figure 4.38) and bioflavonoids such as quercetin (Figure 4.12b), dihydroquercetin, naringenin and citrus glycoside rutin (Figure 9.33) are now widely used as biologically active additives enriching various food products with vitamins and preserving them from oxidation. Flavonoids are added to such food products as vegetable and animal oils, butter, chocolate, cheese, dried milk and other products to stop the oxidative deterioration of lipids. Flavanoids chelate many metallic ions inhibiting their ability to catalyze oxidation. Moreover, bioflavonoids lessen the permeability and breakability of blood capillaries, possess antithrombotic action and reduce the risk of osteoporosis development.

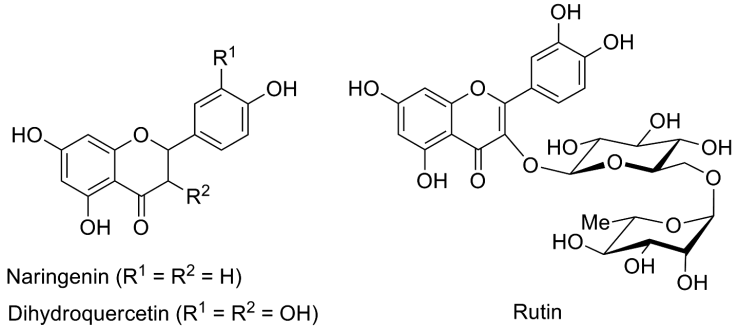
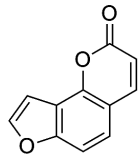


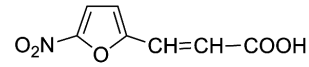
Figure 9.33 *Bioflavonoids used as antioxidants in the food industry.*

As mentioned earlier, an important industrial use of heterocycles is for the preservation of food products. Thus, thiabendazol (Figure 8.18) is extremely effective for the conservation of beets, potatoes, bananas and citrus fruits. Isopsoralene has been added to liqueurs and cosmetics, and 3-(5-nitrofuryl-2)-acrylic acid (Figure 9.34) was suggested as an effective preservative for wine. Macroheterocyclic antibiotic erythromycin A (Figure 7.14) is used as a food additive in animal husbandry. Its protective effect is similar to those of penicillins but has no allergic side-actions. It is produced by biotechnological methods (from *Streptomyces erythreus*). Semisynthetic ways of preparation of it and other macrolide antibiotics are based on intramolecular lactonisation of high carbon acids. Natamycin (or pimarycin) is another macroheterocycic antibiotic used in the protection and conservation of lard and soft cheeses and sausages. Natamycin is fabricated by biotechnology on the bacterium *S. natalensis* and it inhibits the growth and development of the yeast *Candida* and several fungi. Interestingly, note that Natamycin contains a small oxirane ring. Poly(N-vinylpyrrolidone), Figure 9.34, is produced industrially as an effective stabilizer of homogeneity of many food preparations and as a good complexing ligand for the purification of wine, beer and fruit juices.

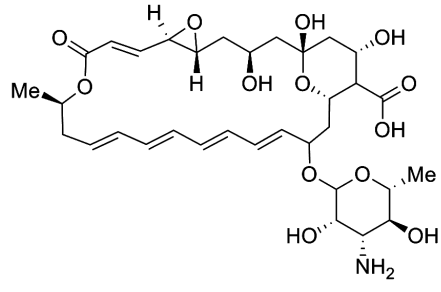
The pleasant odors of many foods are generally not attributable to a single compound. Such aromas are bouquets – a mixture embracing up to 100 volatile components as in coffee, wine and smoked foods. However, the aroma 'profile' of such a mixture is determined by a relatively small group of substances, and the components have been established in many cases. Thus, it



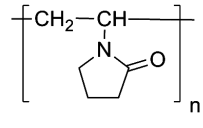
Isopsoralene



3-(5-Nitrofuryl-2)-acrylic acid



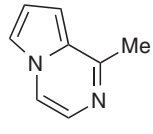
Natamycin



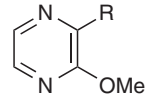
Poly-N-vinylpyrrolidone

Figure 9.34 *Examples of heterocyclic compounds used as food preservatives and stabilizators.*

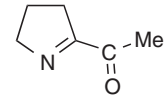
has been determined that 8-methylpyrrolo[1,2-*a*]pyrazine is the major constituent of the odor of roasted meat; 2-methoxy-3-methylpyrazine imparts the fragrance of roasted ground nuts, coffee and cocoa beans; 2-methoxy-3-*n*-hexylpyrazine simulates the aroma of pepper; and 2-acetylpyrroline is responsible for the smell of boiled rice (Figure 9.35).



Odor of roasted meat



Aroma of roasted ground nuts and cofee and cacao beans (R = Me); pepper smell (R = n-C_6H_{13})



Odor of boiled rice

Figure 9.35 *Some heterocyclic compounds with aroma properties.*

9.8 Heterocycles as Cosmetics and Perfumery Ingredients

The modern market of cosmetics and perfumery products now is over 200 billion dollars annually. Among them are heterocyclic compounds having a pleasant odor, color and other esthetic properties, biological activity useful for skin prophylactics, or properties important in manufacturing the products of high quality and preserving their fidelity. Various derivatives of tetrahydrofuran named twins and spans compose a big group of modern emulsifiers and stabilizators of cosmetic disperse systems (Figure 9.36). Spans are used as nonionogenic lipophilic foam formers, solubilisers and structure-forming agents. Twins are added as nonionic surfactants to cosmetic creams, ointments, shampoos and lotions intended for skin care.

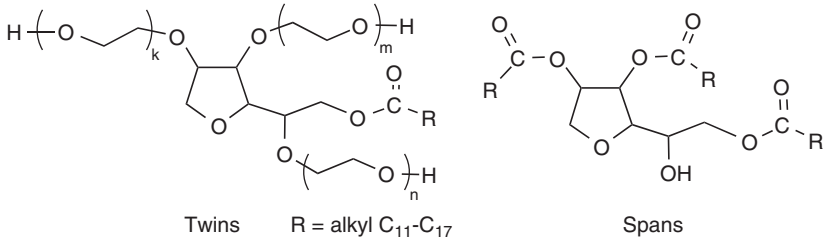


Figure 9.36 *Tetrahydrofuran derivatives used in cosmetics as emulsifiers, surfactants and stabilizing agents.*

Ambroxid is a tetramethyl substituted perhydronaphtho[2,1-b]furan which possesses a strong and stable ambergris fragrance (Figure 9.37). It is an excellent synthetic odor fixator and is used in perfumery instead of very expensive natural ambergris which contains a small quantity of ambroxid.

A series of macrocyclic compounds, containing one or two heterocyclic atoms of oxygen, have a valuable musk odor (Figure 9.37). In addition to the fragrances exaltolid and ambrettolid, they possess amber tones. Moreover, these macrolides have a synergetic property, enhancing the odour of other aromas present in perfumery and cosmetic compositions. Ambrettolid is considered to be the most expensive ingredient in the most costly perfumes. Other types of macrolid derivatives – oxalactons (2-11, cervolid) and cyclic diesters (sabinat, musconat) – have a musk odor and these are used in the fabrication of perfumes, eau-de-Colognes, toilet soaps and cosmetic products.

Nitrogen-containing heterocyclic compounds are also utilized in cosmetic products. Poly(N-vinylpyrrolidone) mentioned above (Figure 9.34) is widely used as an ingredient in hairsprays, in indian ink for eyelashes and in skin creams, owing to its property to form very thin films. In other cosmetic products this polymer plays an important role as a foam stabilizer and jelly-forming agent. The complexes of 2-mercaptopyridine N-oxide with diacetates of zinc (zinc pyrithione) or magnesium (omadin; Figure 9.38) are the active ingredients of shampoos and rinsing compositions for hair to fight dandruff formation. They also have antimicrobial and cytostatic effects.

Quinine (Figure 7.1) when used in cosmetic compositions in the form of its sulfate or arsenate is very effective in curing, tonization and strengthening hair. Orotic acid (Figure 9.38) is frequently included in gerontological cosmetics – for example in creams fighting against aging, mainly

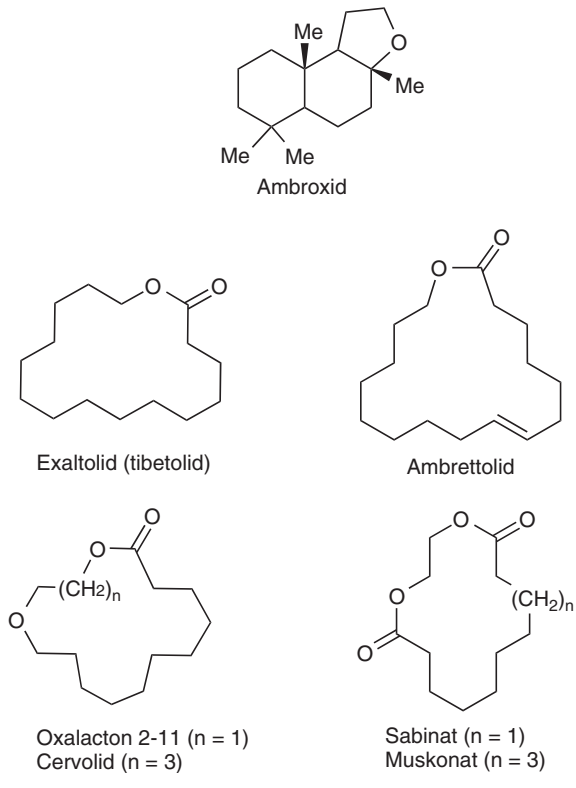


Figure 9.37 *Heterocyclic compounds as amber and musk fragrants and smell fixators in perfumery.*

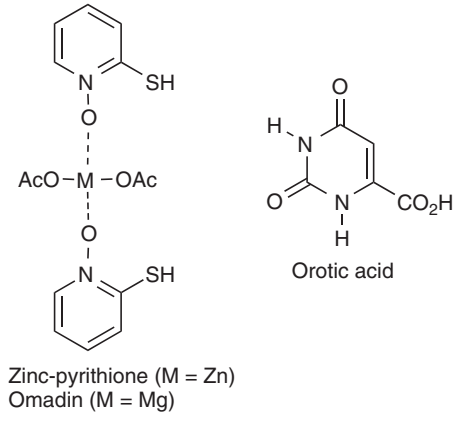
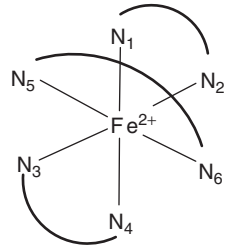


Figure 9.38 *Examples of pyridine and pyrimidine derivatives used in cosmetics.*

in face skin. Orotic acid supports all the functions of skin (its elasticity, humidity, tension, natural color), normalizing the synthesis of proteins and nucleic acids in aging skin. The purine base guanine (Figure 3.1) has now become a popular ingredient in lip pomades (lipsticks) as a mother of pearl (nacre) pigment.

9.9 Other Applications

Heterocycles occupy an important place in analytical chemistry. They are used in the determination of numerous metal ions and of inorganic and organic compounds, as well as in the extraction of metal ores. Analytical reagents for the determination of metal ions usually contain a chelating agent composed of several heteroatoms or a heteroatom and a functional group. Three classical reagents of this type are 2,2′-bipyridyl (Figure 1.6), *o*-phenanthroline and 8-hydroxyquinoline (Figure 9.39). The last is widely used for the estimation of cobalt(II), chromium(III), iron(II), vanadium(V) and other ions. The procedure is based on the formation of chelated complexes with the participation of the pyridine nitrogen and hydroxy oxygen. These complexes are colored. Therefore, the concentration of the extracted metal ion can be determined quantitatively by measuring the intensity of the colored solution spectrophotometrically following extraction with a suitable solvent. 2,2′-Bipyridyl and *o*-phenanthroline are especially useful in the analysis of the iron(II) ion. In its complexes with these compounds, iron(II) is coordinated to three ligands, causing its external electronic shell to resemble that of krypton. The complexes have an octahedral structure as depicted by:



Many heterocyclic analytical reagents contain potential thiol (SH) groups, for example, 2-mercaptobenzimidazole, 2-mercaptobenzoxazole and 2-mercaptobenzothiazole (Figure 9.39).[9] Owing to ionization of the S—H (or N—H) bonds, sparingly soluble salt-like complexes are formed (with participation of the sulfur atoms) with the ions of heavy metals such as cadmium, lead, copper and gold. Determination of the metal concentration is carried out gravimetrically or spectrophotometrically.

A novel branch of analytical chemistry where heterocyclic compounds play an important role is now rapidly developing. This is supramolecular analytical chemistry and chemosensorics. The information concerning this topic is given in Chapters 10 and 11.

Heterocycles are utilized in many other industrial and technological spheres: liquid crystals, polymeric materials, rubber stabilizers, vulcanization accelerators, energetic compounds (explosives) and so on. For example, copolymers of butadiene and 2-methyl-5-vinylpyridine are used in the production of rubbers resistant to the action of heat, oils, lubricants and gasoline. Captax

[9] All of these derivatives, although commonly named as thiols, in fact exist preferentially in the thione tautomeric form containing C═S and NH groups (see Figure 9.39).

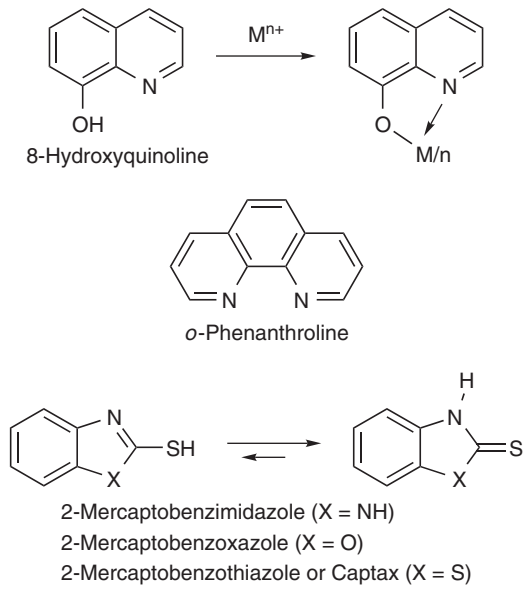
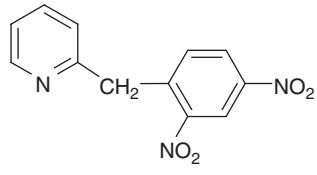


Figure 9.39 *Several heterocyclic analytical reagents.*

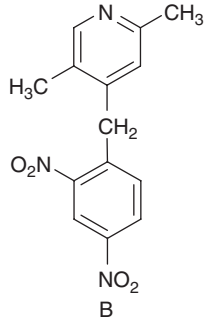
(Figure 9.39) is not only an analytical reagent but also an effective accelerator of rubber vulcanization, for which its current industrial manufacture is impressive. We should remember that the huge-scale drug and vitamin industries are to a large extent also dependent on heterocyclic compounds. The tremendous progress in this field is based on the growing use of chiral and enzyme catalysts, supercritical solvents, ionic liquids and many other recent technological achievements. A number of probable future applications of heterocycles are discussed in Chapter 11.

9.10 Problems

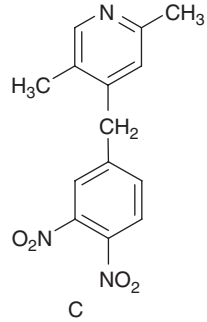
1. What are the main industrial and technological applications of heterocyclic compounds?
2. *Trans*-thioindigo and *trans-N,N′*-dimethylindigo are easily converted into the corresponding *cis* isomers. In contrast, indigo itself undergoes the analogous conversion with difficulty. Explain.
3. Dilute acids and alkalis have little influence on the color of indigoid dyes, in contrast to concentrated solutions of acids and alkalis. Thus, in concentrated sulfuric acid thioindigo changes from red (λ_{max} 546 nm) to blue (λ_{max} 641 nm). Indigo becomes green in sodium *tert*-butoxide solution. Account for these observations.
4. Hydrogenation of one, two or even three of the outer double bonds of the pyrrole nuclei of porphyrins does not cause a change in color. Suggest an explanation for this fact.
5. How could one explain the different color of biliverdin (green) and bilirubin (yellow)? Suggest the possible nature of a coenzyme, which might be responsible for the interconversion of two these pigments.
6. Interaction of a reactive dye with water leads to a decreased ability to adhere to the fiber being dyed. What is the mechanism of this undesirable reaction? To illustrate your answer, use the formulas shown in Figure 9.10.
7. What is the basic principle of optical bleacher use? Give some examples of heterocyclic optical bleachers.
8. What properties are required by laser dyes?
9. What is a photochrome? What is the principle behind the use of photochromic substances for the recording of optical information? Include a heterocyclic photochrome in your answer.
10. Dinitrobenzylpyridines A and B are photochromic substances, whereas C is not. Explain this observation, giving consideration to the structures of the photoexcited species.



A

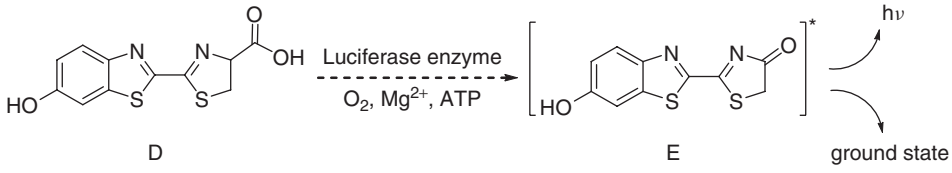


B



C

11. What physical stimuli can cause a change of color of some compounds? Name their spheres of practical application.
12. The characteristic yellow-green glowing of fireflies is caused by 2-(thiazolinyl-2)benzothiazole luminophore D. Its oxidation demands participation of Mg^{2+} ions and an ATP molecule. A final oxidation product which emits light being in an excited state is thiazoline-4-one derivative E. Suggest a scheme for its formation taking into account the role of ATP and information given in Figure 9.18.



D E

13. What is the meaning of the term 'imaging agent'? Provide some examples and explain how they are used in biotechnology.
14. Currently, in biomedical studies infrared dyes with absorption above 700 nm are especially demanded as imaging agents. What is the reason for this?
15. What does the term 'click chemistry' mean? Give an example of a 'click chemistry' reaction and indicate its practical importance.
16. Draw the structure of a luminophore that causes glowing of green fluorescent protein. Explain its natural origin and illustrate its significance for biotechnology. What is the source of photoexitation of GFP in marine organisms?

9.11 Suggested Reading

1. Zollinger, H., *Color Chemistry: Syntheses, Properies and Applications of Organic Dyes and Pygments*, 3rd edn, Wiley-VHZ, Weinheim, 2003.
2. Bamfield P., *Chromic Phenomena. Technological Application of Colour Chemistry*, The Royal Society of Chemistry, Cambridge, 2001.
3. Sung-Hoon Kim, (ed.), *Functional Dyes*, Elsevier, Amsterdam, 2006.
4. Leznoff, C. C. and Lever, A. B. P., (eds), *Phthalocyanines: Properties and Applications*, VCH, New York, 1993.
5. Bouas-Laurent, H. and Dürr, H., Organic photochromism, *Pure Appl. Chem*., 2001, **73**, 639.
6. Krasovitskii, B. M. and Bolotin, B. M., *Organic Luminescent Materials* (English translation by Vopian, V. G.), VCH, Weinheim, New York, 1988.
7. Zimmer, M., *Glowing Genes: a Revolution in Biotechnology*, Prometheus, Buffalo, 2005.
8. Pieribone, V. and Gruber, D., *Aglow in the Dark: the Revolutionary Science of Biofluorescence*, Belknap, Cambridge, 2006.
9. Tedder, J. M., Nechvatal, A. and Jubb, A. H., *Basic Organic Chemistry. Part 5: Industrial Products*, John Wiley & Sons, Ltd, London, 1975.

10

Heterocycles and Supramolecular Chemistry

At the bottom of a pitcher,
Dwarfy's met a giant teacher.
'How's a dwarf so small as thee
Got into this, just say to me?'
G. Sapghir

It took chemists a very long time to penetrate one fundamental frontier that separates the substances synthesized in their laboratories from those which provide and control the activity of living organisms. The differences between the synthetic molecules and most key biomolecules encompass not only the structural complexity of the latter but also their wonderful capability to self-organize and recognize other molecules. Classical examples of such self-organization include the formation of DNA duplexes (Section 3.2), assembly of ribosomes (Section 3.4) and specific protein structures such as chymotrypsin or hemoglobin (Sections 4.1 and 4.2). Immune reactions provide amazing examples of molecular recognition which arise as an organism's response to the xenobiotics (antigenes) that invaded it. This response switches on the synthesis of special protein antibodies able to selectively bind and neutralize substances alien to the organism. Up to the end of the 1960s, chemists could only dream of creating such synthetic molecules but this field then began to change, slowly at first and then increasingly rapidly. Artificial molecules capable of recognizing other chemical particles were discovered, then compounds capable of self-assembling were introduced and their synthesis vigorously developed. Thus, a new branch of knowledge arose in the 1980s–1990s, called 'supramolecular chemistry'. It originated in the work of three Nobel Prize laureates (1987) – two Americans, C. Pedersen and D. Cram, and one Frenchman, J.-M. Lehn.

What is supramolecular chemistry? Lehn defined it as the chemistry of intermolecular bonds that studies the associations of two or more chemical particles and the properties of such associations. Classic chemistry deals mainly with the structures of separate molecules and their reactions by which valent bonds are formed or ruptured. Unlike this, supramolecular chemistry relies mainly on noncovalent interactions such as hydrogen bonding, electrostatic and hydrophobic forces, π-π

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

stacking, metal complexation and so on. As a rule, supramolecular interactions are reversible. The associates formed under such interactions (named 'supramolecules') are able to break apart or flexibly change their structures under the influences of different stimuli. Supramolecular chemistry has strongly influenced the development of modern chemistry and science as a whole. It can be considered the precursor of nanochemistry. Current progress in such fields as medicinal chemistry, biotechnology, electronic engineering, sensors and others is greatly dependent on the methods of supramolecular chemistry, as discussed in more detail in Chapter 11.

As mentioned in Chapter 2, heterocyclic compounds have the unique ability to participate in noncovalent interactions. So it is not surprising that heterocycles are at the center of supramolecular investigations. As in natural product chemistry, the majority of synthetic molecules entering supramolecular interactions possess large or even giant molecular weights.

10.1 Molecular Recognition and Host–Guest Interactions

Molecular recognition is the selective interaction of two (sometimes more) molecules in which a larger complex plays the role of host, reversibly attracting in its 'embrace' a smaller molecule called a guest. In principle, host–guest interactions are quintessential in supramolecular chemistry. To provide better binding, the host molecule needs to have in its structure polar fragments, proton donor and proton acceptor centers, π-conjugated sites (such as aromatic rings) and so on. Along with these, it is greatly important that the shape and preorganization of the host molecule be appropriate to accept the guest. In one sense, host molecules can be viewed as molecular containers and for this reason host–guest interactions are sometimes referred to as container or receptor chemistry. We consider below the most widespread types of host–guest interactions, classifying them in accordance with the nature of the guest.

10.1.1 Cation Receptors

In 1967 Pedersen published a series of articles in the *Journal of the American Chemical Society*. The then 63 year old chemist from Du Pont described the synthesis of a new type of heterocyclic compounds which could be classified as cyclic polyethers (Figure 10.1a). In total, Pedersen synthesized more than 60 cyclic polyethers which contained 4–20 oxygen atoms, connected by CH_2CH_2 bridges. The rings ranged in size from 12 to 60 members.

The most intriguing feature of these macrocyclic polyethers was their ability to form unusually stable crystalline complexes with alkali metal ions. The practical importance of this discovery was immediately obvious as scientists did not previously have at their disposal efficient reagents for the extraction and separation of these ions.

The strength of the complex, with a preferred 1:1 polyether : ion ratio, is determined by the electrostatic attractions between the metal ion contained within the macrocyclic cavity and the ether oxygen atoms (Figure 10.1b, structure **1**). Pedersen noticed the similarity between a royal crown and the macrocyclic polyether 'crowning' an ion. To simplify the rather complicated nomenclature of these macrocycles, Pedersen proposed the new name crown ethers, which became universally accepted. The names of crown ethers include two numerals: the first designates the size of the ring, and the second the number of oxygen atoms it contains.

It was further found that the stability of each complex largely depended on the correlation between the internal cavity size of the crown ether and the ionic radius of the cation. The data in Table 10.1 indicate that 12-crown-4 forms a stable complex with the Li^+ ion, 15-crown-5 is the most suitable ether for complexation with Na^+ and 18-crown-6 is suitable for coordination with K^+, NH_4^+ and Rb^+.

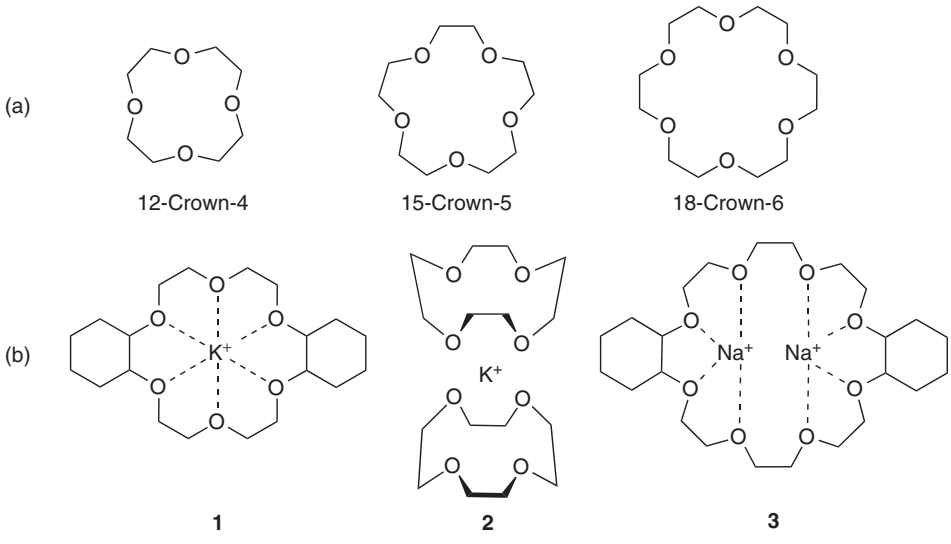


Figure 10.1 *(a) Crown ethers and (b) their complexes:* ***1*** *is a 1:1 complex of dicyclohexano-18-crown-6 with* K^+*,* ***2*** *is a 2:1 complex of 12-crown-4 with* K^+*,* ***3*** *is a 1:2 complex of dicyclohexano-24-crown-8 with* Na^+*.*

Table 10.1 *Cation diameters and cavity sizes of optimum crown ethers [adapted from Vögtle, F. and Weber, E. (eds),* Host Guest Complex Chemistry: Macrocycles: Synthesis, Structures, Applications*, Springer, Berlin, 1985, Chap. 1, p. 18, with permission]*

Cation	Cation diameter (Å)	Crown ether	Cavity diameter (Å)
Li^+	1.36	12-Crown-4	1.2–1.5
Na^+	1.90	15-Crown-5	1.7–2.2
K^+	2.66	18-Crown-6	2.6–3.2
NH_4^+	2.86	18-Crown-6	2.6–3.2
Rb^+	2.94	18-Crown-6	2.6–3.2
Cs^+	3.38	21-crown-7	3.4–4.3

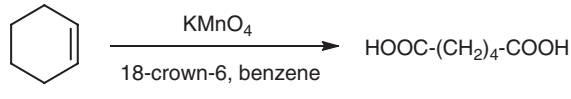
A mismatch in cation diameter and cavity size, however, does not preclude complex formation. If the cavity is too small for a cation (e.g., K^+ for 12-crown-4), complexation may still occur. In some such complexes a 2:1 ratio of constituents is observed in which one cation simultaneously coordinates with two molecules of crown ether (Figure 10.1b, 'sandwich'-like structure **2**). If the macrocycle is too large, the cavity may be occupied by two cations at the same time, forming a 1:2 molar ratio complex (Figure 10.1b, structure **3**). The cation may be enveloped by the macrocycle like a pearl in a half-open oyster shell. In each of these two cases just mentioned, the stability of the complex is significantly lower than when there is optimal correlation between the cation and crown ether dimension.

The stability constant, K_s, is a measure of complex stability. K_s is obtained by application of the law of mass action to the corresponding equilibrium:

$$K_s = \frac{[L \cdot M^+]}{[L][M^+]}$$

In this equation, $[L{\cdot}M^+]$, [L] and $[M^+]$ are the concentrations of the complex, free ligand and cation, respectively. For instance, the log K_s values for the sodium and potassium complexes of dicyclohexano-18-crown-6 are 6.4 and 8.3, respectively. Their ratio, $\sim 10^2$, is approximately an exponential measure of the selectivity of 18-crown-6 toward Na^+ and K^+ ions. In other words, a solution of the crown ether to which equal concentrations of Na^+ and K^+ ions are added will contain only one bound sodium ion for every 100 ions of potassium involved in complexation.

Crown ethers have found applications in many fields of science and technology. They are effective for the separation of alkali and rare-earth metal ions which are required for their analysis, extraction and purification. Owing to the solubility of crown ethers in nonaqueous media, alkali metal salts can be solubilized, that is, transferred from an aqueous to an organic phase. For example, potassium permanganate is itself moderately soluble in water but practically insoluble in organic solvents. However, in the presence of 18-crown-6, $KMnO_4$ becomes readily soluble in benzene, forming so-called 'purple benzene', which has exceptionally high oxidizing properties greatly surpassing those of permanganate in aqueous solution. This may be explained by the fact that in water the MnO_4^- ions are surrounded by a dense solvation shell, whereas in the nonpolar solvent, benzene, they are unsolvated or 'naked'. Purple benzene, in contrast to aqueous $KMnO_4$, readily oxidizes alkenes, alcohols, aldehydes and even alkyl groups in alkylarenes at room temperature to form carboxylic acids in almost quantitative yield. The oxidation of cyclohexene under such conditions provides adipic acid in 100% yield:



The development of crown ethers has been closely monitored by the medical profession. Substances capable of selectively binding cesium ions in the presence of Na^+ and K^+ would be potential treatments for human exposure to the very dangerous $^{137}Cs^+$ radioactive ion, a widespread radionuclide. Similar dangerous effects of $^{90}Sr^{2+}$ and other ions have been observed. Reports from many countries regarding successful research in radionuclide entrapment have recently become available.

The crown ethers proved to be the first synthetic analogues of naturally occurring substances which could transport alkali metal ions (Na^+, K^+) through cellular membranes. Natural ion transporters, ionophores, act according to the same principle as the crown ethers, but ionophores have substantially more complicated structures. Thus, one of the best known, the antibiotic valinomycin, has a macroheterocyclic depsipeptide structure composed of twelve alternating α-amino acid and α-hydroxy acid residues: D-valine, L-valine, D-hydroxyisovaleric and L-lactic acids (Figure 10.2a). Valinomycin is a specific potassium ion carrier having a K^+/Na^+ selectivity of about 10^4. In contrast to the crown ether complexes, the K^+ ion coordinates with valinomycin through the carbonyl oxygen atoms of the ester groups to form an octahedral structure of the type shown in Figure 10.2b.[1]

[1] Another mode of ion penetration through the membranes involves tubular channels. The alkaline-earth ions, particularly Ca^{2+}, are preferentially transported in this manner.

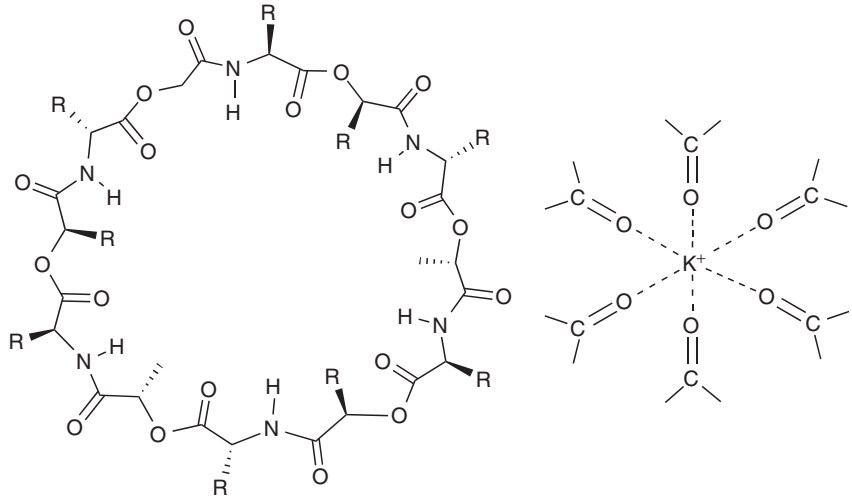


Figure 10.2 *(a) Antibiotic valinomycin (R = i-Pr) and (b) coordination fragment of its potassium complex.*

Valinomycin, like other natural cation carriers, is a reversible ionophore. Such ionophores, having penetrated a cell, liberate the cation under the effect of certain interactions and are then rapidly returned to the outside of the cell to bind another cation. The rate of these trans-membrane migrations can reach several thousand per second, and they may even operate against a concentration gradient.

The forces driving the movement of the ionophore include changes in the pH of the medium, redox potential, irradiation and other factors. Many types of reversible crown ethers which mimic natural ionophores have been prepared. Thus, the 18-crown-6 derivative **4** containing a long side chain with a terminal NH_2 group (Figure 10.3) responds well to pH changes. This crown ether forms complex **5** with K^+ in neutral or slightly alkaline media. However, upon acidification the amino group is protonated and the ammonium ion formed expels the K^+ cation from the ether cavity using its 'arm'-like chain with an ammonium 'hand' (see Figure 10.3, intramolecular complex **6**). Figuratively speaking, the crown ether 'bites its tail'. Since K^+ and NH_4^+ ions are similar in size (Table 10.1), displacement of K^+ by $NH_4{}^+$ in acidic media must be the result of entropy factors which favor intramolecular complexation of the alkylamino group. If the ammonium complex is returned to a neutral medium containing excess K^+, the potassium ions enter the cavity owing to deprotonation of the ammonium group. Thus, crown ether **4** can be compared to a shuttle moving back and forth between the external and internal walls of the membrane. By this mechanism K^+ ions are transported into the cell and H^+ ions out of the cell (Figure 10.3, structure **7**).

The functioning of natural ionophores is, of course, much more complex and efficient than that of the best synthetic reversible-type crown ether presently known. Moreover, the action of synthetic ionophores has been tested only on artificial polymer membranes. Nevertheless, progress has been substantial.

The discovery of crown ethers stimulated an army of chemists, in the literal sense of the word, to attempt the synthesis of new, more effective and more selective complexation reagents for the alkali metals. Much attention to detail was required, in particular to create internal cavities that were better organized and more capacious. While an ether macromolecule could be likened to a

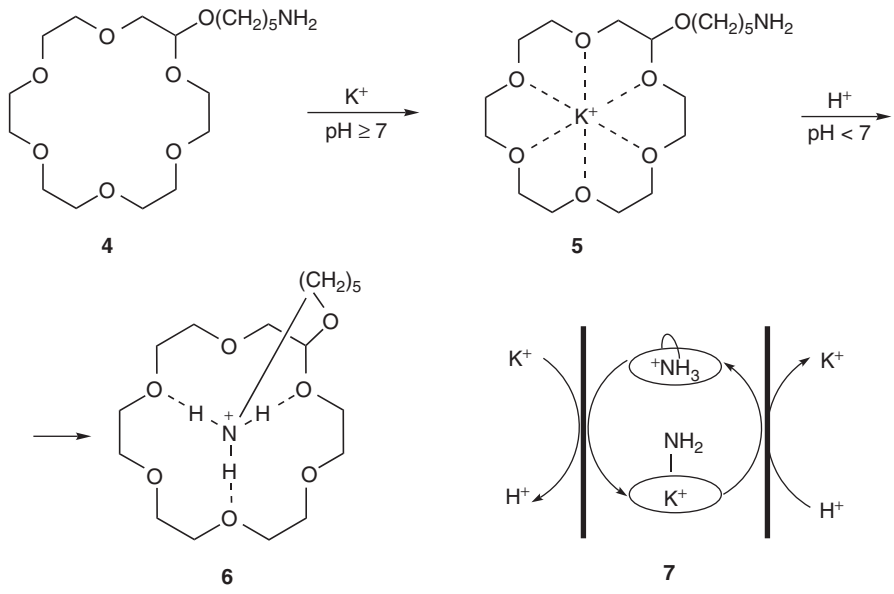


Figure 10.3 *pH-dependent crown ether properties (Nakatsuji, Y., Kobayashi, H. and Okahara, M.,* J. Chem. Soc., Chem. Commun.*, 1983, 800. Reproduced by permission of The Royal Society of Chemistry).*

crown or a hat, it then seemed reasonable that other three-dimensional molecular containers for metal ions could be prepared, such as molecular cups, jugs, pots, saucepans, barrels and so on. Major contributions to this field were made by J. Lehn and D. Cram.

Lehn began his investigations in 1968 by synthesizing three-dimensional aliphatic amino ethers **8** (Figure 10.4) which he named cryptands (Greek: *krypte*, cave or cavern). The association with caves resulted from the presence in such compounds of an internal cavity limited by three ether chains which met at two bridged nitrogen atoms. Structure **9**, [2.2.2]cryptate, is formed by complexation of a metal ion to cryptand **8** ($m = n = 1$). Complex **9** is the most intensively studied cryptate because its internal cavity size is suitable for many cations including Na^+, K^+ and Rb^+. The cations are retained within the cavity not only by the walls of this 'cavern' but also by electrostatic attraction between the cation and the electron pairs of the six oxygen and two nitrogen atoms. It is not surprising that the stabilities of [2.2.2]cryptand complexes with Na^+ and K^+ ions (picrate counter ion: log $K_s = 10.6$ and 13.2, respectively) are four or five orders of magnitude greater than those of the analogous 18-crown-6 complexes.

Compound **10** is another type of cryptand which resembles the shape of a football. The size of the cavity best fits the cesium or ammonium ion. Indeed, its complex with Cs^+ is the most stable of all known complexes of this cation. The stability of the complex of **10** with an ammonium ion (see Figure 10.4, schematic structure **11**) is largely a result of the fact that the four hydrogens of the tetrahedral NH_4^+ ion are directed toward the four cryptand nitrogens, thus allowing the formation of hydrogen bonds. Interestingly, the hydrogen bonding, together with steric shielding and electrostatic effects, results in a decrease in acidity of the NH_4^+ ion in complex **11** by six orders of magnitude compared with uncomplexed aqueous NH_4^+.

Cram commenced his study of molecular cavities in the mid-1970s. His attention was attracted by one seemingly negligible deficiency of the crown ethers and cryptands. X-Ray analysis demonstrated that both groups of compounds were not organized well enough to accept the desired

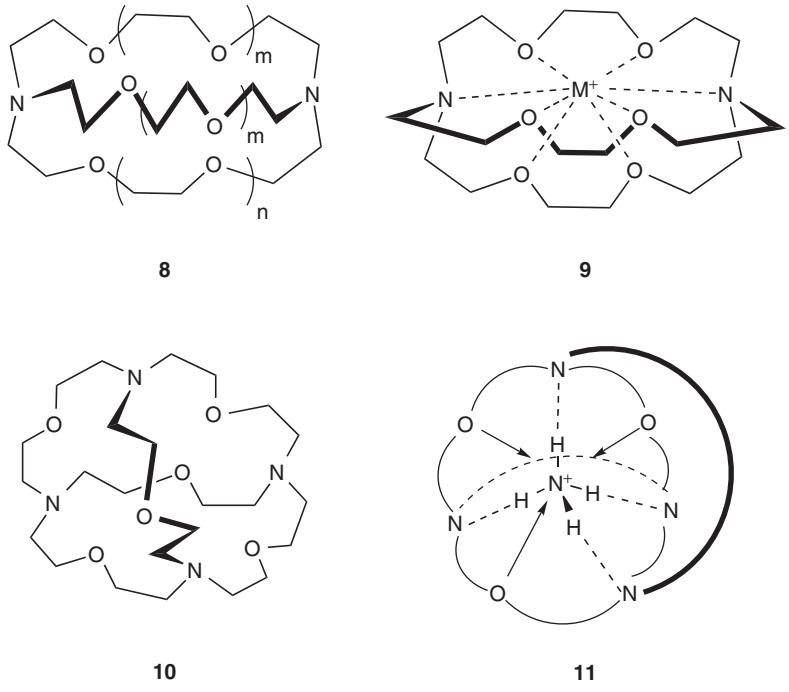


Figure 10.4 *Cryptands and their complexes.*

guest ions because their structures were somewhat 'crumpled' (Figure 10.5, structures **12** and **13**). Therefore, entry of a cation into the cavity would necessitate additional energy expenditure to reorganize and smooth out the structure, which would be reflected in the stability of the complex. Cram and his coworkers were successful in their engineering of molecular containers void of this shortcoming. In a complicated series of steps, aromatic ethers **14** and **15** with structures preorganized for complexation were prepared. The new compounds were christened spherands and cavitands. To illustrate more fully this chemical class, we include nonheterocyclic spherands (in particular **14**).

Spherands and cavitands are a type of molecular container in which the walls are lined with aromatic nuclei, the complexation being achieved by oxygen atoms. Ionic molecular or atomic guests are attracted and bound by these atoms. The containers may even be thought of as having short 'legs' if one attributes this role to the external methyl groups (structures **14** and **15**). The synthetic strategy employed by Cram proved to be highly successful. Thus, spherand **14** complexes the Na^+ cation with log $K_s = 14.1$ (Table 10.2), which is much more stable than the complexes of Na^+ with 18-crown-6 and [2.2.2]cryptand. However, the most remarkable property of **14** is its unprecedented Na^+/K^+ selectivity (10^{10}), which even surpasses all natural ionophores. Once inside spherands and cavitands, alkali metal ions are shielded to a great extent by the methyl or methylene groups linked to the oxygen atoms. These groups prevent solvation of the cation. Cavitand **15** can trap and hold prisoner small neutral molecules such as CH_2Cl_2, MeCN, SO_2 and others (for more on capturing neutral molecule, see Section 10.1.3).

Many other molecular containers have been synthesized. Of these, cryptaspherand **16**, a cryptand–spherand hybrid, is also characterized by a high K^+/Na^+ selectivity (Table 10.2) which

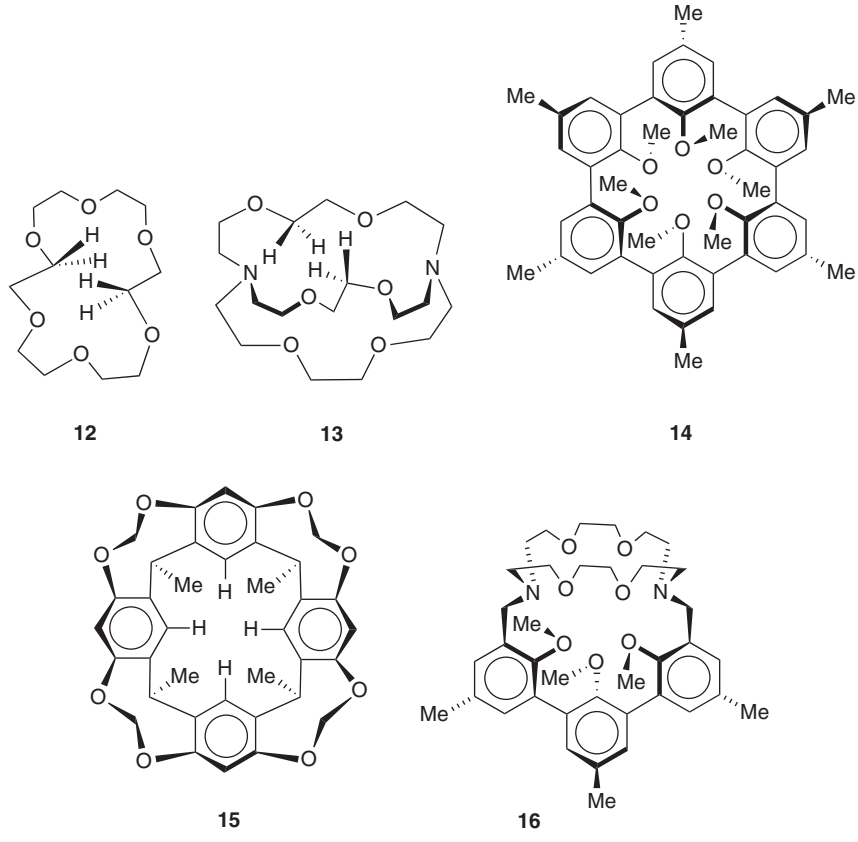


Figure 10.5 *Unorganized structures of crown ether* ***12*** *and cryptand* ***13****. Preorganized structures of spherand* ***14****, cavitand* ***15*** *and cryptaspherand* ***16****.*

Table 10.2 *Stability constants (log K_s) of macrocyclic complexes with Na^+ and K^+ (picrate counterion; Bell, T. W., Firestone, A. and Ludwig, R.,* J. Chem. Soc., Chem. Commun.*, 1989,* ***1989****, 1902. Reproduced by permission of The Royal Society of Chemistry)*

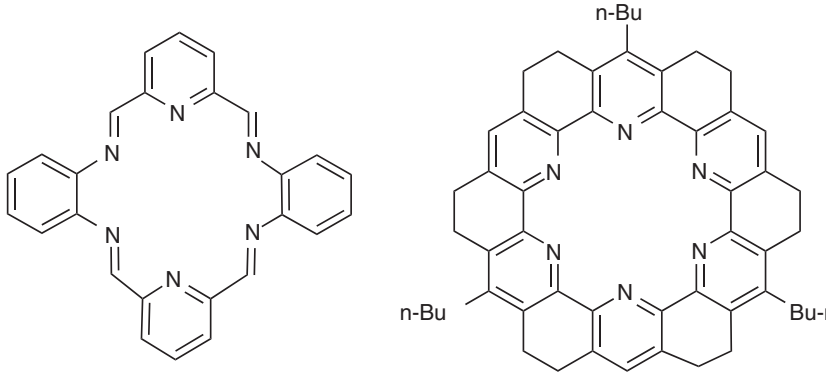
Ion	Structure number[a]				
	1	**9**	**14**	**16**	**18**
Na^+	6.4	10.6	14.1	9.9	14.7
K^+	8.3	13.2	4.4	13.9	14.3

[a]See Figures 10.1, 10.4, 10.5, 10.6.

is similar to that of the natural antibiotic valinomycin and exceeds that of many other synthetic ionophores.

In the mid-1980s interest began to shift toward aza analogues of crown ethers and cryptands. The aza group (−N =), with a specifically oriented unshared pair of electrons, has a more rigid configuration than an amine nitrogen or ether oxygen. Azamacrocycles should therefore be capable

of more precise design for the reception of guest ions. Indeed, azacrown **17** shown in Figure 10.6 reacts with the K^+ cation in a 1:1 molar ratio to give a complex which is measurably more stable than the similar complex with 18-crown-6. However, record-breaking stability was attained by the Na^+ and K^+ complexes of azacrown **18** (also called torand), which have stability constants that are even higher than for the cavitands (Table 10.2). Unfortunately, the flat and somewhat simple structure of the azacrowns can be a disadvantage because it results in low Na^+/K^+ selectivity.



17 **18**

Figure 10.6 *Azacrown ethers.*

Azacrown **19** (Figure 10.7) possesses some interesting properties which are useful in the determination of Li^+. In nonpolar solvents such as methylene chloride **19** exists in the red tautomeric form **19b**. The compound loses its color in polar media (e.g., methanol) owing to conversion to the fully aromatic tautomer **19a**. When lithium salts are added to a solution of azacrown **19** in CH_2Cl_2, Li^+ cations begin to displace protons from the internal cavity to form complex **20**. Since **20** is also colorless, the process can be monitored spectrophotometrically.

Research involving the construction of cascade or coreceptive macrocyclic ligands is being vigorously pursued. Such compounds are designed to possess two different complexation sites: one is soft (polarizable), the other is rigid (almost nonpolarizable) and acts predominantly electrostatically. A typical example is **21** shown in Figure 10.8. Macrocycle **21** has a soft coordinative center composed of two sulfur atoms and an adjacent pyridine nitrogen atom. The rigid center is formed by the cryptand portion. On treatment with a mild Lewis acid such as rhodium carbonyl, **21** forms complex **22** in which the metal ion is coordinated to the soft center with the carbonyl oriented toward the inside of the macrocycle. Further addition of a copper(II) salt effects coordination of the copper(II) ion with the cryptand fragment, with the ion being closely located to the carbonyl oxygen **23**. The copper(II) ion also coordinates with the $C{\equiv}O^+$ group, thus activating it toward nucleophilic addition. Such chemical activation imitates the function of some metal-containing enzymes.

The binding of two identical ions (e.g., two copper ions) to both centers of a cascade ligand is also possible. In this case, owing to the different chemical environments the two cations vary substantially in the final complex and, for example, the reductive potentials of the two ions are now significantly different.

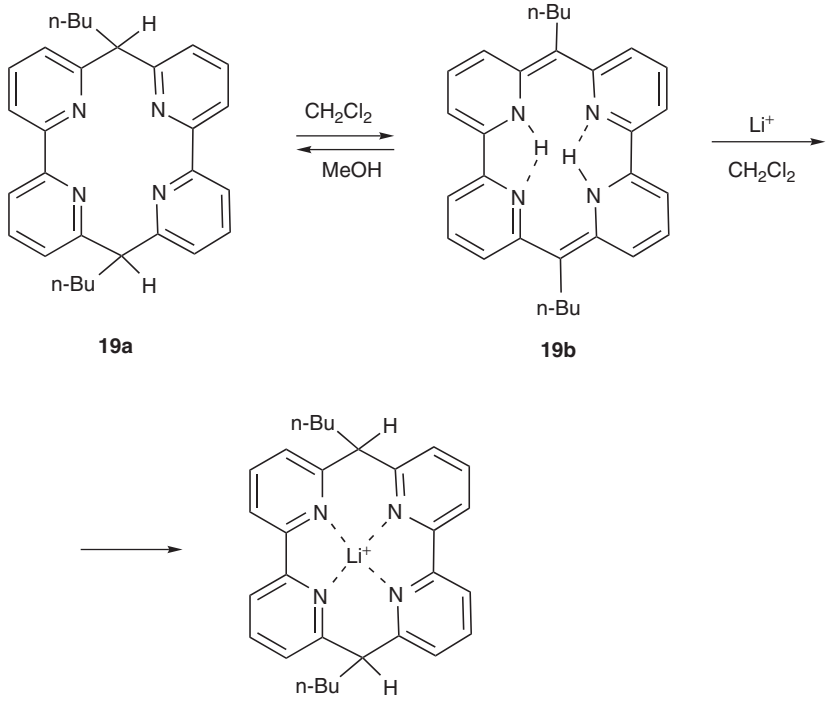


Figure 10.7 *Azacrown ethers used for Li^+ ion detection (Reprinted with permission from Ogawa S., Narushima, R. and Arai, Y.,* J. Am. Chem. Soc.*, 1984,* ***106****, 5760. © 1984 American Chemical Society).*

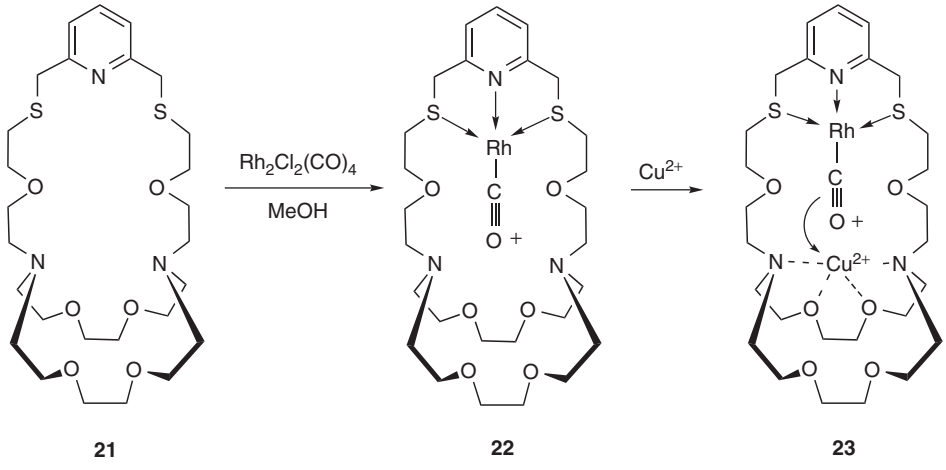
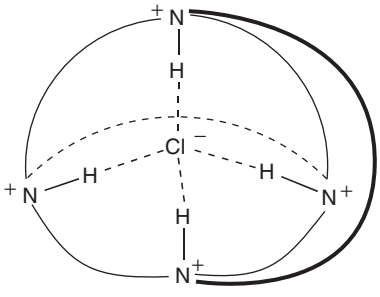


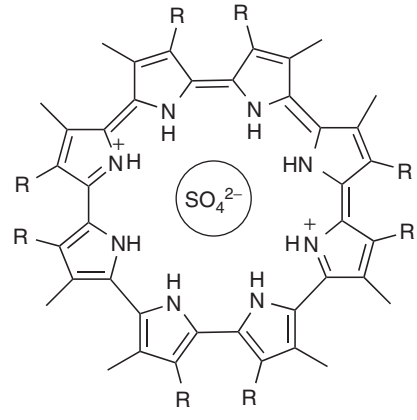
Figure 10.8 *Schematic action of a macrocyclic cascade ligand (Carroy, A. and Lehn, J.-M.,* J. Chem. Soc., Chem. Commun.*, 1986, 1234. Reproduced by permission of The Royal Society of Chemistry).*

10.1.2 Anion-, Betaine- and Ionic Associated Receptors

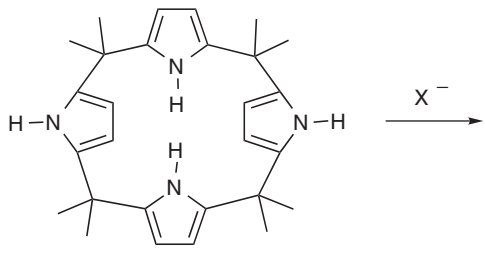
Anions play an excedingly important role in biology. About 70% of chemical processes occuring in living organisms proceed with participation of anions such as chloride, phosphate, carboxylates, polyanions of nucleic acid and so on. Their mimicking can be useful for medicinal chemistry and the creation of various molecular devices. Effective anion sensors and extractants are much required for technological and enviromental purposes. Almost all receptors for capturing anions are macroheterocyclic compounds. As a rule, they bind anions via multiple hydrogen bonds. Thus, cryptands may be adapted for anion complexation if the cavity is surrounded by positively charged centers. For instance, the chloride anion is well suited to the internal cavity size of protonated cryptand **10**, resulting in a complex represented schematically by structure **24** in Figure 10.9. Since a bromide anion could hardly occupy the same cavity because of its increased size, **10** is a good reagent for the separation of Cl^- and Br^- ions.



24



25 ($R = C_{11}H_{23}$)



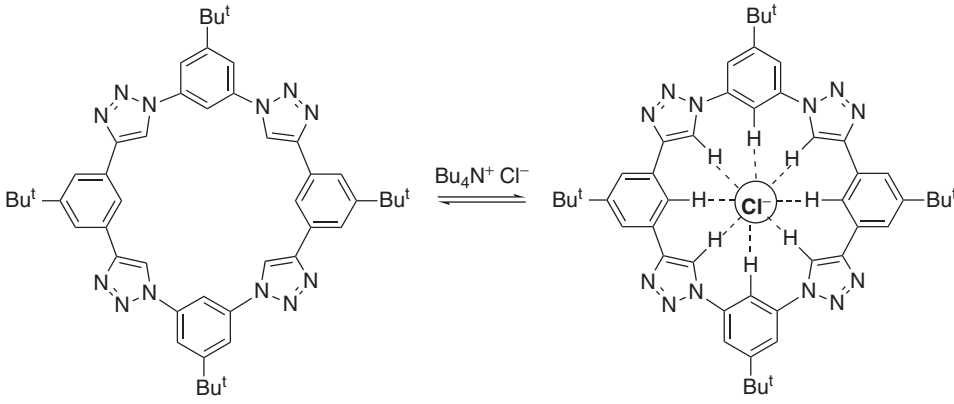
26 **27**

Figure 10.9 *Macrocyclic anion receptors acting via N–H. . .X⁻ binding (structure **25**: Reprinted with permission from Eller, L. R., Stepień, M., Fowler, C. J., Lee, J. T., Sessler, J. L. and Moyer, B. A.,* J. Am. Chem. Soc.*, 2007, **129**, 11020. © 2007 American Chemical Society; structures **26, 27**: Reprinted with permission from Gale, P. A., Sessler, J. L., Král, V. and Lynch, V.,* J. Am. Chem. Soc.*, 1996, **118**, 5140. © 1996 American Chemical Society).*

One of the most serious ecological problem is the long-term storage of nuclear waste. A proposal has been made to incorporate it into borosilicate glass 'logs' that would be stored for millennia in a geological repository. However, a small amount of corrosion-inducing sulfate anion in nuclear waste can make long-term storage difficult. Recently, the first promising candidate for a sulfate anion extractant was developed: it is a derivative of a cyclo[8]pyrrole carrying eight undecyl tails to increase molecule solubility in organic solvents. The compound has a central cavity that is just the right size for binding the SO_4^{2-} ion with the formation of complex **25** (Figure 10.9). A toluene solution of the cyclo[8]pyrrole can extract sulfate anion out of aqueous solutions that are much richer in nitrate anion – the situation typical for radioactive waste.

Anion-coordinating macrocyclic ligands need not be positively charged. Thus, in the 1990s it was found that calix[4]pyrroles, obtained by A. Baeyer over a century ago, form complexes of type **27** with halide and some other small anions, most strongly with the F^- anion ($K_s = 1.7 \times 10^4\ M^{-1}$). Notably, ligand **26** is not preorganized for such complexation and adopts a strongly nonplanar 1,3-alternate conformation in which adjacent rings are oriented in opposite directions. When an anion approaches the molecule its conformation changes, which allows participation of all four N-H bonds in anion binding, though the complex still remains puckered with the anion residing in an apical position to the root mean square plane of the calixpyrrole.

An essentially different type of anion receptor is represented by the recently prepared 1,2,3-triazole-based macrocycle **28** (Figure 10.10) with a capacious cavity ideally suited to accept a chloride anion. Most remarkable is the observation that the anion is captured via 8 hydrogen bonds of the C—H. . .X^- type, in which both triazole and benzene nuclei participate. The stability constant of resulting complex **29** is about $1.3 \times 10^5\ M^{-1}$ (CH_2Cl_2, 298 K).

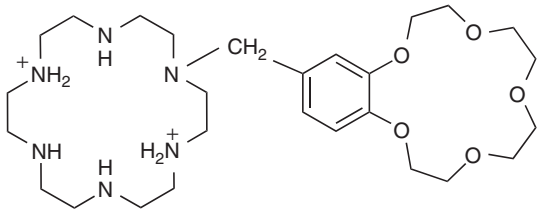


28 **29**

Figure 10.10 *Macrocyclic anion receptors acting via C–H. . .X^- binding (Li, Y. and Flood, A. H.,* Angew. Chem., Int. Ed.*, 2008, **47**, (14), 2649. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

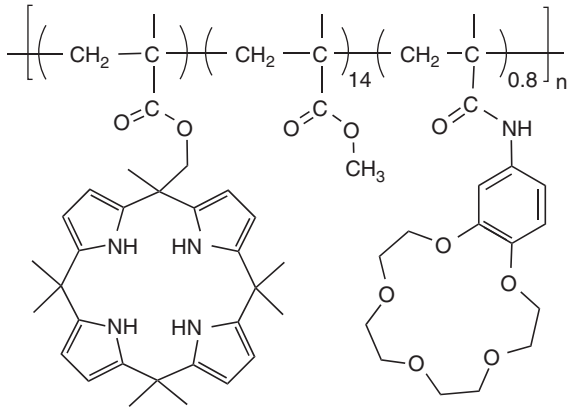
After the creation of separate receptors for cations and anions the next logical step was their joining in one single molecule to bind both types of ions at the same time. One of the first examples of such a kind was the protonated bicyclic system **30** (Figure 10.11) which forms rather stable complexes with α-amino acids and other betaines in neutral aqueous solution. The crown ether portion becomes linked to the ammonium group of the acid (see Figure 10.3, structure **6**), while

(a)



30

(b)



31

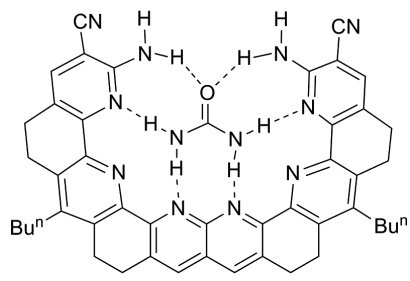
Figure 10.11 *Bifunctional receptors for binding betains (a) and salts (b) (structure **30**: Kimura, E., Fujioka, H. and Kodama, M.,* J. Chem. Soc., Chem. Commun.*, 1986, 1158. Reproduced by permission of The Royal Society of Chemistry; structure **31**: Aydogan, A., Coady, D. J., Kim, S. K., Akar, A., Bielawski, C. W., Marquez, M. and Sessler, J. L.,* Angew. Chem., Int. Ed.*, 2008, **47**, 9648. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

the positively charged polyamine ring coordinates with the carboxylate anion. In the resulting complex, ligand **30** has the conformation of a half-open book. Organic molecules have a much more complicated structure than metal cations. Therefore, figuratively speaking, on complexation with a crown ether, the latter are satisfied by a molecular 'sack for a sphere', whereas the former are more demanding and require the crown to be an exact fit such as, for instance, 'a violin for its case'.

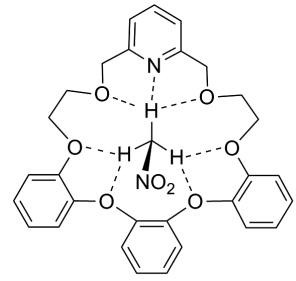
Other examples of bifunctional cation-anion receptors are the recently prepared poly(methyl methacrylate)s **31** with pendant crown-ethers and calixpyrroles (Figure 10.11). They are capable of extracting effectively from aqueous solutions potassium halide salts, such as KCl and KF. This opens a way to solve many technological problems including water purification and the separation of useful products, such as bromine and potassium, from high-salt water sources.

10.1.3 Receptors for Neutral Molecules

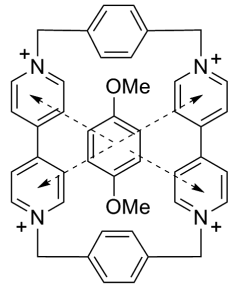
The synthesis of heterocyclic receptors capable of forming complexes of the 'guest–host' type with neutral molecules has become the focus of intense investigation in the last two decades. Such receptors are necessary for the separation of organic compounds, for the creation of new generations of drugs, for detection and analysis (chemosensors) and for resolving many other chemical and



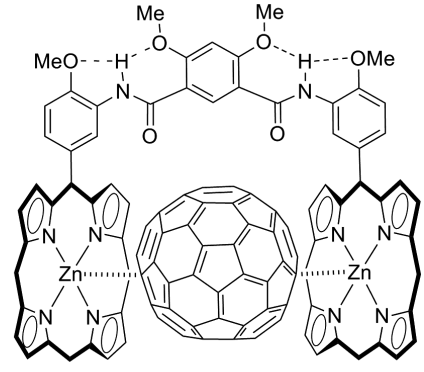
32



33



34



35

Figure 10.12 *Host–guest interaction between heterocyclic receptors and molecules of urea, nitromethane, 1,4-dimethoxybenzene and fullerene (structure* ***32****: Bell, T. W. and Hou, Z.,* Angew. Chem., Int. Ed. Engl.*, 1997,* ***36****, 1536. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission; structure* ***33****: Weber, E., Franken, S., Puff, H. and Ahrendt, J.,* J. Chem. Soc., Chem. Comm.*, 1986, 467. Reproduced by permission of The Royal Society of Chemistry; structure* ***34****: Reprinted with permission from Anelli, P. L., Ashton, P. R., Ballardini, R., Balzani, V. et al.,* J. Am. Chem. Soc.*, 1992,* ***114****, 193. © 1992 American Chemical Society; structure* ***35****: Reprinted with permission from Wu, Z.-Q., Shao, X.-B., Li, C., Hou, J.-L., Wang, K., Jiang, X.-K. and Li, Z.-T.,* J. Am. Chem. Soc.*, 2005,* ***127****, 17460. © 2005 American Chemical Society).*

technological problems. These artificial molecular receptors must mimic enzymes and biological receptors whose structures are adapted to recognize and bind such highly specific substrates.

Since even small organic molecules are rather bulky, their hosts must have a sufficiently large cavity well preorganized for multiple interactions. Most often this is a macrocycle or a half-circle molecule having a cleft or tweezer-like shape. Thus, a whole series of selective hydrogen-bonding receptors for urea have been reported. Commonly, their complexes (e.g., **32**, Figure 10.12) have 1:1 composition and are so stable that they can be recrystallized without any decomposition. Such urea receptors are acutely needed for the efficient purification of blood of kidney patients.

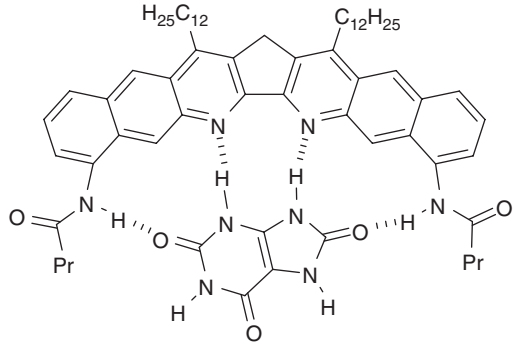
Azacrown ether **33** forms a rather stable 1:1 complex with a nitromethane molecule in which the guest is oriented almost perpendicular to the average plane of the cyclic polyether. The nitromethane methyl group is oriented toward the macrocycle and forms hydrogen bonds with the pyridine nitrogen and ether oxygen atoms. The nitro group is encircled by the three phenyl rings. As a result, the $MeNO_2$ molecule is held captive in a semispherical cavity with a radius of

about 350 nm. Nitroethane, dimethylformamide, dimethyl sulfoxide and many other solvents do not form complexes with this crown ether; the selectivity toward nitromethane is very high.

The cavity in bipyridinium cyclophane **34** is well suited for the acceptance of 1,4-disubstituted benzenes with electron-donor substituents such as 1,4-dimethoxybenzene. The host–guest interaction in this case is caused by partial transfer of π-electron density from the guest to the host molecule, as evidenced by a strictly parallel orientation of all the aromatic rings in complex **34**.

Molecular tweezer **35** with two zinc porphyrin units exhibits remarkably high binding affinity for fullerenes C_{60} and C_{70}. It is suggested that the driving force for encapsulation here is π-π stacking between the zinc porphyrin and fullerene units. Such fullerene assemblies attract great attention due to their interesting photophysical, photochemical and electrochemical properties.

Heterocyclic molecules may function not only as receptors but also as guests. Complex **36** of uric acid is an example (Kelly, T. R. and Maguire, M. P., *J. Am. Chem. Soc*., 1987, **109**, 6549):



36

$K_s = 9.1 \times 10^5\ M^{-1}$ (CH_2Cl_2–$C_6H_5CH_3$, 1:1)

10.1.4 Molecular Carcerands

In the middle of the 1980s, Cram's group synthesized molecules with structures resembling two closely associated hemispheres. These compounds were obtained from two different cavitands, one containing CH_2SH (or CH_2OH) groups on opposite peripheries, the other containing CH_2Cl groups. As expected, nucleophilic displacement of chloride occurred when equimolar quantities of these two cavitands were mixed in dimethylformamide in the presence of cesium carbonate (the latter being required for CH_2XH group ionization). This reaction joined the two hemispheric molecules together on both sides (Figure 10.13). Of course, these two hemispheres or 'cups' do not form a tight seal and a small clearance between them still remains. However, the gap is too narrow for even the smallest molecules trapped in the cups during the reaction to escape. The analogy with a prison or cage is rather appropriate and is reflected in the name given to such macromolecules – carcerands (Latin: *carcer*, a prison cell). Carcerands were the first organic compounds synthesized with capacious cavities entirely isolated from the exterior. If small molecules are 'accidentally' trapped inside the carcerand at the moment of closure, they will remain inside as 'jailbirds' and neither crystallization, chromatography nor any other conventional method can release them. Thus, during carcerand synthesis Cram and his coworkers found that diverse types

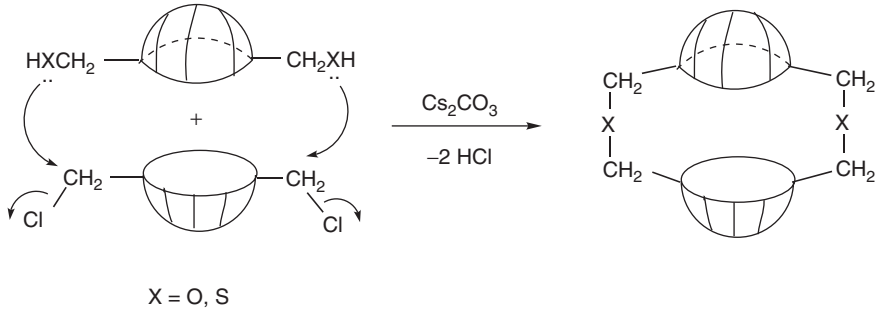


Figure 10.13 *The principle of joining two cavitand molecules to form a carcerand.*

of small particles present in the reaction mixture (dimethylformamide, cesium cations, chloride anions) were captured inside the carcerand.[2]

Subsequently, related molecules called hemicarcerands were created. They differ from carcerands by having a lower number and greater length of bridges between the cavitand rims. As a result of relatively free access to the equatorial zone, a guest molecule can easily enter into the cavity or escape from it under varying conditions.

Carcerands and hemicarcerands have found unexpected and rather specific applications in chemistry. Thus, it is possible to use the inner phase of hemicarcerands as a reactor for the synthesis of some extremely labile organic compounds and the latter, being isolated from outside space, can survive inside long enough for thorough study. A very successful example of this approach is the preparation of cyclobutadiene (**39**, Figure 10.14), an elusive hydrocarbon which had been prepared for the first time in the1960s with a lifetime of less than 5 s by irradiation of α-pyrone **38** in an argon matrix at 8 K; the reaction is accompanied by the loss of a CO_2 molecule.

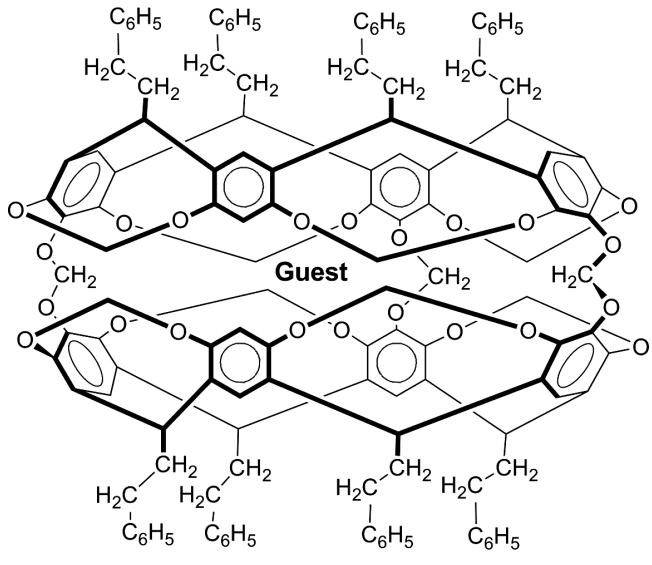
Twenty years later, Cram *et al*. conducted the same reaction inside the cavity of a hemicarcerand **37**, which is capable of accommodating small guest molecules such as DMF, benzene or α-pyrone (Figure 10.14). At first, **37** was refluxed with **38** in chlorobenzene, which led to the formation of the desired inner complex (hemicarceplex) **37**·**38**. Irradiation of this complex in $CDCl_3$ resulted in the formation of cyclobutadiene and CO_2. While the CO_2 molecule easily escaped from the cavity, **39** remained inside unchanged for hours. The main pathways of cyclobutadiene destruction are its easy cyclodimerization, cycloaddition to alkenes and autoxidation (cyclobutadiene conversion into two acetylene molecules is also possible but only proceeds at elevated temeperature). Since no other molecule can be placed into the inner space of hemicarcerand **37**, the cyclobutadiene is protected from fast decay. Similarly, another intriguing unstable molecule, 1,2-dehydrobenzene (**41**) was also generated photochemically from α-diketone **40** (Figure 10.14) and captured. Interestingly, the inner cavity of carcerands and hemicarcerands has been suggested for consideration as a new phase of matter where a host molecule provides a discreet molecular inner space as is true not only for solids (as for inclusion compounds and zeolites) but also for solutions and even for gaseous phases.

In the next section we examine molecular carcerands of a somewhat different type.

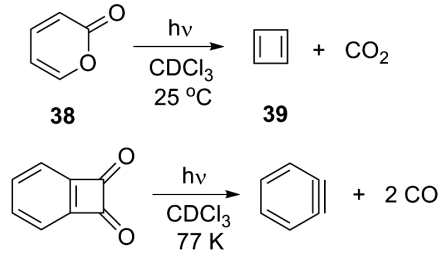
10.1.5 Molecular Containers for the Proton

The proton is one of the simplest, most abundant and important particles in nature. It is engaged in such processes as molecular and enzymatic catalysis, ionic conduction, membrane potentials and,

[2] Small linear molecules such as CO_2 or O_2 are the only species capable of thermal insertion into a carcerand by heating.



37



38 **39**

40 **41**

Figure 10.14 *Hemicarcerand **37** preventing unstable guest molecules such as cyclobutadiene (**39**) and 1,2-dehydrobenzene (**41**) from fast decomposition (Cram, D. J., Tanner, M. E. and Thomas, R,* Angew. Chem., Int. Ed.*, 1991, **30**, 1024. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission; Warmuth, R.,* Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 1347. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

of course, all acid–base equilibria. The proton is the only stable ion having no electronic shell and this explains why it is so small. One cannot compare protons with typical molecules; any tennis ball against basketball analysis is wrong. The situation is more impressive. The radius of a proton is 10^{-13} cm and is considerably less than that of other molecules and ions ($\sim 10^{-8}$ cm). Imagine a football field: if the size of a football field (more precisely two of them) corresponds to the other chemical species we commonly encounter, then a poppy seed in the center would correspond to the size of a proton.

Due to its tiny size and the absence of an electronic shell, protons should not exhibit ordinary chemical phenomena such as steric hindrance and inter-electron repulsion. As a result, protons literally stick to any molecule interacting both with free electrons and also with electrons engaged in chemical bonds. Such ions as hydroxonium, H_3O^+, and methonium, CH_5^+, well exemplify this

point. The proton affinity for electrons is so high that protons exist in the free state only in a high vacuum. Under these circumstances two questions arise: (i) how realistic is it to expect that a single proton or, say, a couple of them can be properly hidden in a certain molecular cavity and (ii) what benefit can be derived from such proton encapsulation.

In 1968 Alder *et al*. reported the abnormally high basicity of 1,8-bis(dimethylamino)naphthalene (**42**, Figure 10.15a) reaching $pK_a = 12.1$ in water or 18.2 in acetonitrile.[3] This value is substantially higher than the basicity not only of ordinary arylamines (pKa ~ 5–6) but also alkylamines (pKa ~ 10). This is caused by two main factors: (i) destabilization of the base **42** due to severe repulsion between the nitrogen lone electron pairs and (ii) the formation of a strong intramolecular hydrogen bond (IHB) in protonated form **43**. As an ordinary sponge behaves in water, molecule **42** in a proton medium slowly adds a proton and very strongly holds it. In other words, diamine **42** displays rather low kinetic but high thermodynamic activity. For this reason it was called a 'proton sponge' and this term has now become general for all such strong organic bases with similar properties. The low kinetic activity of **42** is caused by the inability of the proton to leave the cleft between the two NMe_2 groups as such. In fact, it must be extracted from the internitrogen space by another base and this process is hampered by the four bulky and hydrophobic N-methyl groups.

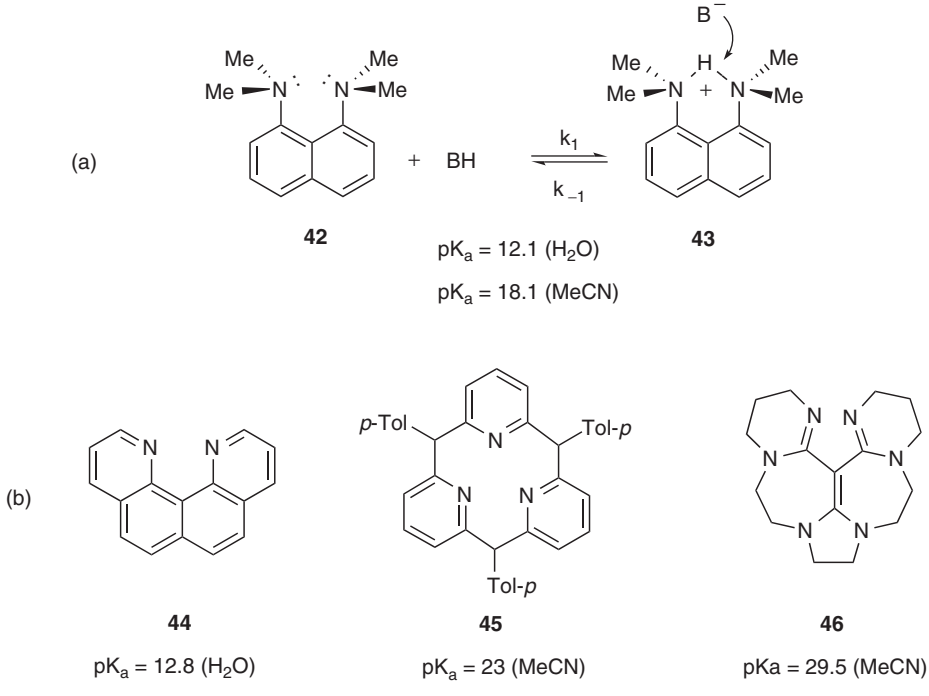


Figure 10.15 *Naphthalene proton sponge* ***42*** *and some proton sponge-like heterocyclic compounds (Pozharskii, A. F. and Ozeryanskii, V. A., in* The Chemistry of Anilines*, ed. by Rappoport, Z., John Wiley & Sons, Ltd, Chichester, 2007, Part 2, Chapter 17, p. 931–1026. © 2007 John Wiley & Sons, Ltd).*

[3] In those cases when the indication of a solvent is omitted, water is present.

Another important property of proton sponges is their very low nucleophilicity. This means that no other electrophiles, except protons, can attach to the nitrogen atoms. This observation has found a fortunate application in organic synthesis. Proton sponges are often used when a proton should be removed from a substrate having other base-sensitive functionalities.

The discovery of proton sponges strongly stimulated development of acid-base theory as well as container chemistry and some related problems. First of all, scientists have been hunting for stronger and stronger neutral organic bases. This goal was achieved using two main approaches: (i) by further destabilization of a base at the expense of better directionality of their nitrogen electron pairs, alignment of their basic centers and increasing their number and (ii) by strengthening an intermolecular hydrogen bridge in the protonated form using the heterocyclic motif. Typical examples are polynuclear heterocycles **44**–**46**. All have several pyridine nitrogen atoms whose free electron pairs are directly pointed toward each other more so than those of amine nitrogens. As expected these compounds, especially **45** and **46**, are more basic than the parent sponge **42**. However, in spite of their outward resemblance with proton sponges, **44**–**46** are kinetically active and therefore should be treated rather as proton sponge-like compounds and not as true proton sponges.

Bicyclic diamines of the cryptand type (Figure 10.4) behave towards protons completely differently. They display two kinds of protonation: inside and outside (Figure 10.16). Thus, [1.1.1]cryptand **47** first forms *out*-cations **48** and **49** with $pK_a^1 = 7.1$ and $pK_a^2 = 1.0$. Ions **48** then slowly rearrange into chelated *in*-form **50** with $pK_a \geq 17.5$ (indirect estimate). Further acidic treatment of **50** results in capturing the second proton with the formation of *in,in*-dication **51**. Dication **51** is deprotonated with a great difficulty: on being refluxed with a concentrated aqueous solution of KOH, it is partially converted into *in*-monocation **50** which cannot be further deprotonated. Thus, cryptand **47** is actually a "proton carcer".

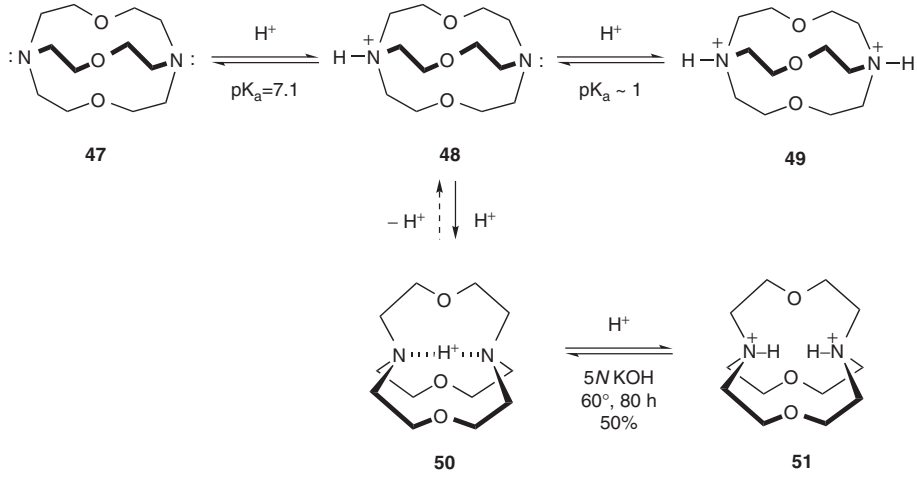
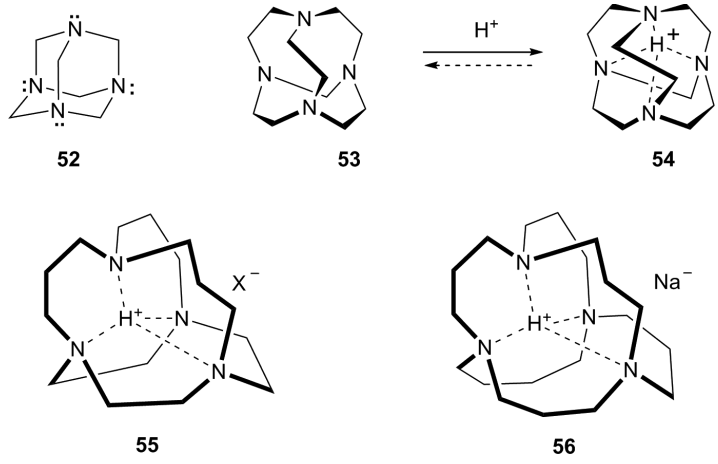


Figure 10.16 *Outside/inside protonation of [1.1.1]cryptand **47** (Reprinted with permission from Smith, P. B., Dye, J. L., Cheney, J. and Lehn, J.-M.,* J. Am. Chem. Soc.*, 1981, **103**, 6044. © 1981 American Chemical Society).*

An interesting class of cage tricyclic amines is adamanzanes (Figure 10.17). Their molecules consist of four nitrogen atoms with six alkylene bridges between them. The parent compound in

the series is well known [1^6]adamanzane or hexamethylenetetramine (**52**). [1^6]Adamanzane **52** is weakly basic since all its nitrogen electron pairs are directed outward and cannot be inverted. However, [2^6]adamanzane (hexaethylenetetramine, **53**) traps a proton so tightly that up to now this tetramine has been isolated only as cation **54**. The same is true for [$2^3.3^3$]adamanzane forming salts **55**. All attempts to remove the proton from these cations were unsuccessful. Such 'proton prisons' have found an unexpected application for the stabilization of alkalides – ionic compounds with an alkali metal anion. For example, hydrogen sodide **56** has been obtained using [3^6]adamanzane. Compound **56**, considered as a reversed sodium hydride, exists as golden needles stable below −25 °C.



52 **53** **54**

55 **56**

Figure 10.17 *Some adamanzanes and their proton complexes (Miyahara, Y., Tanaka, Y., Amimoto, K., Akazawa, T., Sakuragi, T. et al.,* Angew. Chem., Int. Ed.*, 1999, **38**, 956. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reprinted with permission from Springborg, J., Nielsen, B., Olsen, C. E. and Søtofte, I.,* J. Am. Chem. Soc.*, **124**, 5084. © 2002 American Chemical Society).*

Examples of such a kind with an encapsulated metal cation instead of a proton are also known. Metallic sodium dissolves in liquid ammonia to form a dark blue equilibrium solution of sodium atoms, ions and solvated electrons. However, sodium is much less soluble in alkylamines, for example, ethylamine. But when [2.2.2]cryptand (Figure 10.4, structure **8**, $m = n = 1$) is added to ethylamine, sodium dissolution occurs much more readily, and over time precipitation of golden crystals of [2.2.2]cryptand·2Na occurs. X-Ray analysis has indicated that one sodium atom (as a cation) is situated inside the cryptand cavity, whereas the second is located outside the cavity as an Na^- anion. Thus, the [2.2.2]cryptand stabilizes the separation of charge in the Na^+ Na^- ionic pair, which is of considerable significance.

Unlike adamanzanes and cryptands, macrocyclic tetramine **57** with a relatively large inner cavity and greater basicity behaves as a typical proton sponge (Figure 10.18). Thus, it easily abstracts a proton from chloroform with the formation of a monocation which rapidly interconverts under ambient conditions between two equivalent forms **58a** and **58b**. The resultant anion CCl_3^- breaks down to chloride and dichlorocarbene, which can be trapped.

In summary, highly basic nitrogen heterocyclic compounds can serve as proton containers. According to their kinetic activity, they can be divided into the three following categories: proton sponges, proton sponge-like compounds and proton prisons.

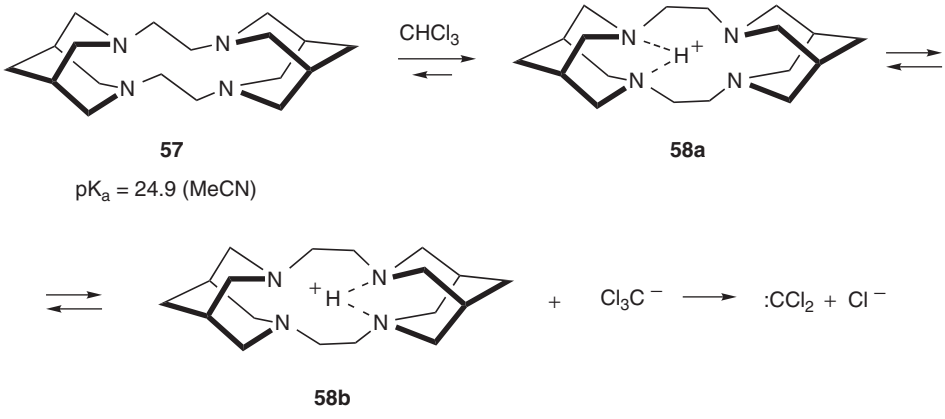


Figure 10.18 *Proton sponge behavior of tetramine* ***57*** *(Reprinted from Miyahara, Y., Goto, K. and Inazu, T.,* Tetra. Lett.*,* ***42****, 3097. With permission from Elsevier © 2001).*

10.2 Self-Assembling Molecular Systems

In addition to container chemistry, there now exists another rapidly growing main division of supramolecular chemistry that is even more exciting. This is the creation of molecules and molecular assemblies capable of self-organization. As in the case of 'host–guest' chemistry, heterocyclic compounds have occupied leading positions in the design of artificial self-assembling molecular systems. Ligands of type **59**, consisting of a number of 2,2′-bipyridyl units that are linked by relatively flexible ether groups, were an early example of such compounds (Figure 10.19). In the presence of copper(I) or silver(I) salts, two threads of such a ligand wind around the metal ions and around each other to form a double-stranded helicate **60**. The driving force for this process is the well known tendency of copper(I) and silver(I) ions to realize tetrahedral coordination geometry. As a result, each ion binds to a bidentate fragment consisting of two molecules of **59**.

In addition to coordination bonds, self-assembly can also be induced by hydrogen bonding or electrostatic interactions (Figure 10.20). Thus, macrocyclic tetracationic salt **61** and polyether **62** cocrystallize as a very stable self-organized complex ($K_a = 11150\ M^{-1}$) of the pseudorotaxane type,[4] existing in two conformations, **63a** and **63b**. During the association the linear molecule **62** passes through the macrocycle in a manner which recalls thread passing through the eye of a needle: the linear molecule is held inside the macrocycle by charge-transfer interactions between the π-deficient 4,4′-bipyridyl residues and the electron-rich 1,5-dioxynaphthalene fragments.

A powerful instrument of modern organic synthesis is the combination of the self-assembly phenomenon with classical reactions resulting in covalent bond formation. This approach is best illustrated by catenane synthesis. Catenanes (Latin: *catena*, chain) contain two or more interlocked rings similar to the links in a chain. Such exotic-looking molecules were first obtained in the beginning of the 1960s. However, laborious synthetic procedures and disappointing yields strongly hampered further progress. At that time the idea of statistical synthesis seemed most simple and

[4] A rotoxane is a mechanically interlocked supramolecule consisting of a 'dumbbell' shaped unit which is threaded through a macrocycle. Its kinetic stability is provided by bulky groups (stoppers) at the ends of the dumbbell. When the stoppers are small or absent, the rotoxane structure can be realized only through noncovalent interactions as in **63**; due to the lowered kinetic stability of **63**, such complexes are named pseudorotoxanes.

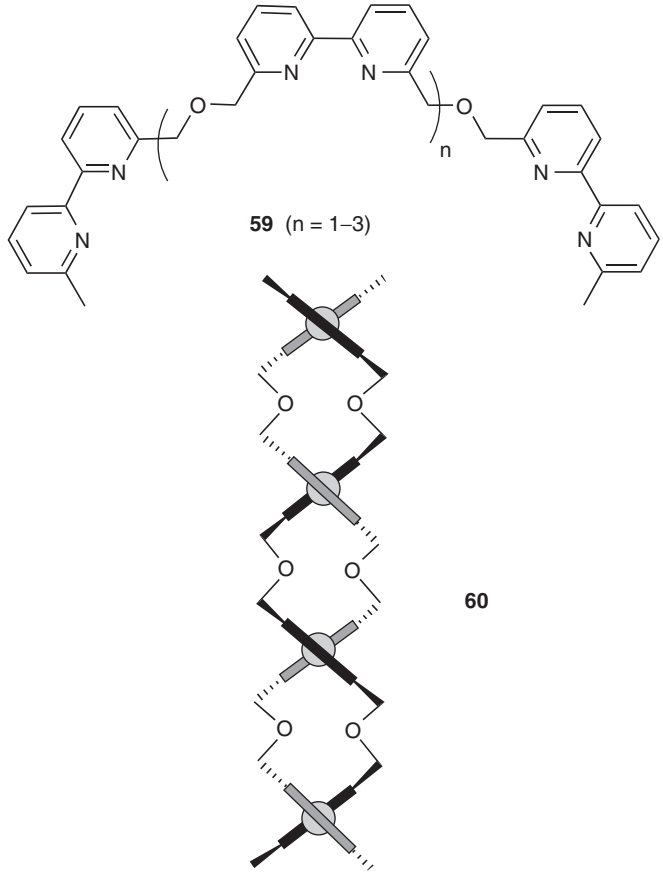


Figure 10.19 *Formation of a double-stranded helicate (shaded circles are Ag^+ or Cu^+ ions; grey and black rectangles are 2,2′-bipyridyl fragments; Lehn, J.-M. and Rigault, A.,* Angew. Chem., Int. Ed. Engl.*, 1988,* ***27****, 1095. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

attractive (Figure 10.21). This method envisages the cyclization of linear molecules containing reactive groups X and Y at their termini (one nucleophilic, the other electrophilic) to be carried out in the presence of preformed macrocycles of a definite size. At the exact moment when the new ring is being formed by an intramolecular reaction of the linear molecules, a statistically predetermined number will have 'threaded the needle' to form a [2]catenane assembly (or a [3]catenane when the [2]catenane is joined by a newly formed third ring). The probability of such an event is very low. Moreover, molecular models indicate that the would be catenane ring has to be constructed of no less than 21 atoms for it to be sterically possible for the rings to become linked.

A breakthrough in this field came from the group of the French chemist Sauvage in the early 1980s. They revived the idea of a statistical synthesis on the basis of a new principle which can be termed 'supramolecular assistance to molecular synthesis'. A manyfold increase in the probability

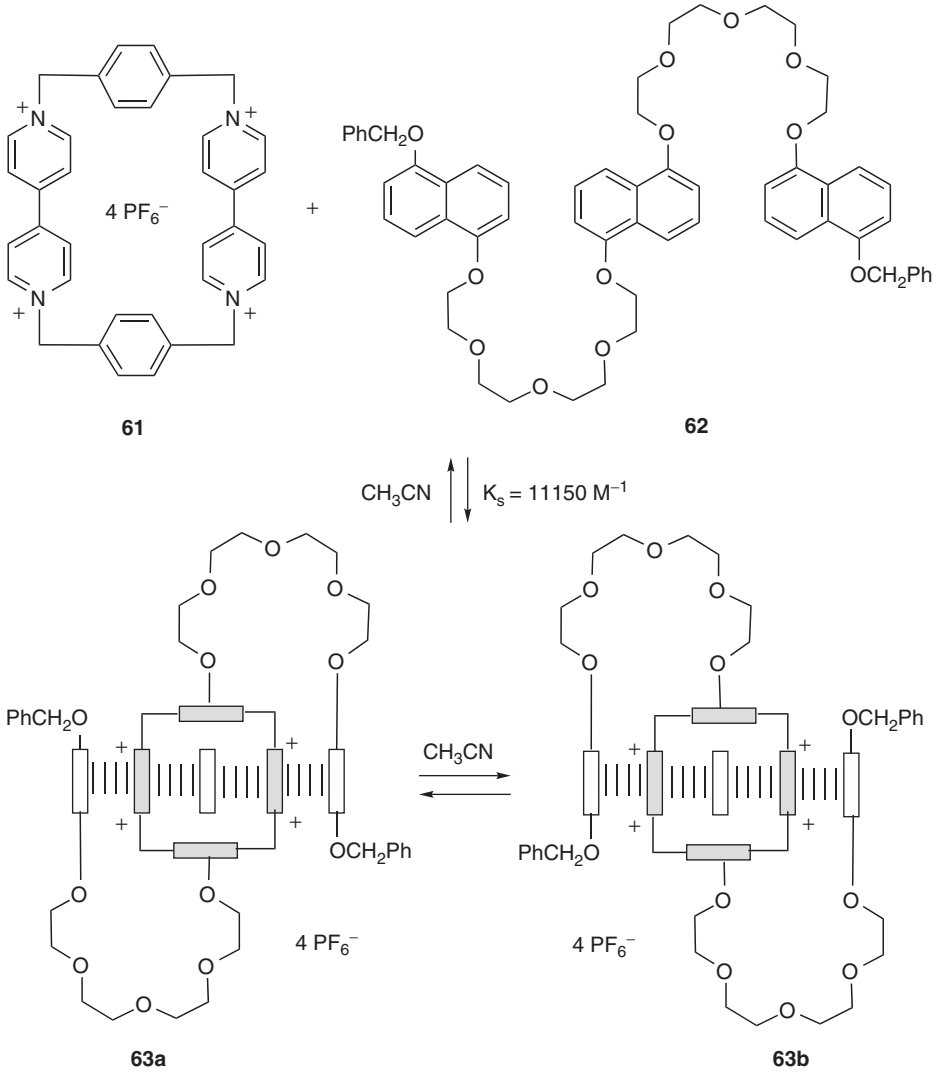


Figure 10.20 *Self-assembly induced by electrostatic interactions (shaded and white rectangles are 4,4′-bipyridyl and 1,5-dioxynaphthalene fragments, respectively; Ashton, P. R., Philp, D., Spencer, N., Stoddart, J. F. and Williams, D. J.,* J. Chem. Soc., Chem. Commun.*, 1994, 181. Reproduced by permission of The Royal Society of Chemistry).*

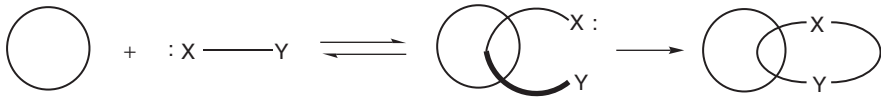


Figure 10.21 *Statistical catenane synthesis.*

of closure of the interlocking rings was achieved during the alkylation of the phenanthroline bisphenol **64** with the polyether dichloride **65** in the presence of azacrown ether **66** and a copper(I) salt (Figure 10.22). Owing to the stable 2:1 complex formed by *ortho*-phenanthroline with the Cu^+ ion (see Section 9.8), both reagents containing phenanthroline residues coordinate to form a complex of type **67** prior to cyclization. This accomplishes the formation of catenane **68** with yields up to 60%.[5]

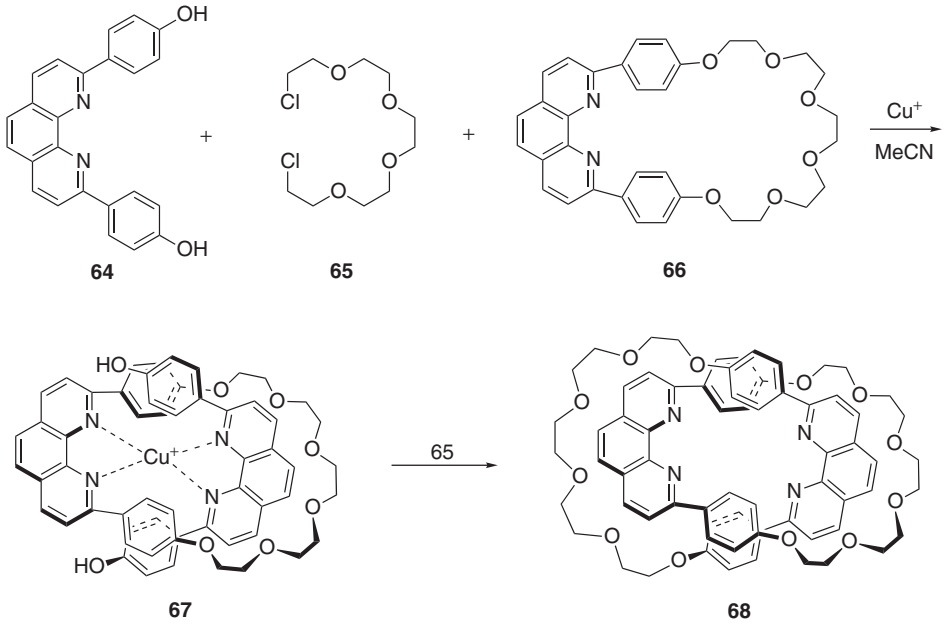


Figure 10.22 *Metal complexation assistant catenane synthesis (Albrecht-Gary, A.-M., Dietrich-Buchecker, C., Saad, Z., Sauvage, J.-P. and Weiss, J.,* J. Chem. Comm.*, 1986, 1325. Reproduced by permission of The Royal Society of Chemistry).*

A similar approach, but based on π,π-stacking interactions as template motif, was successfully realized by the J.F. Stoddart group (Figure 10.23). They managed to obtain [2]catenane **73** in 70% yield by reacting dication **69**, macrocyclic polyether **70** and *p*-xylilenedibromide **71**. Apparently, stepwise quaternization of the free nitrogen atoms is strongly facilitated by the formation of intermediate pseudorotoxane **72** which is stabilized by charge-transfer interactions as in the case of **63**. Notably, the rings in catenane **73** rotate relative to each other with some activation barrier. As a result of the rotation the phenyl rings **A** and **B** can interchange their positions forming two isoenergetic translation isomers (see Problem 10 in Section 10.3). Such molecular dynamics, possible also for rotaxanes, offer very promising applications in engineering various electronic devices, for example molecular switches. Many polycatenanes, including [5]catenane ('Olympiadane') have been synthesized in a similar manner.

More aspects of supramolecular chemistry are discussed in Chapter 11.

[5] Preparative procedures in which reagents are localized spatially by a metal ion or another particle are referred to as template syntheses.

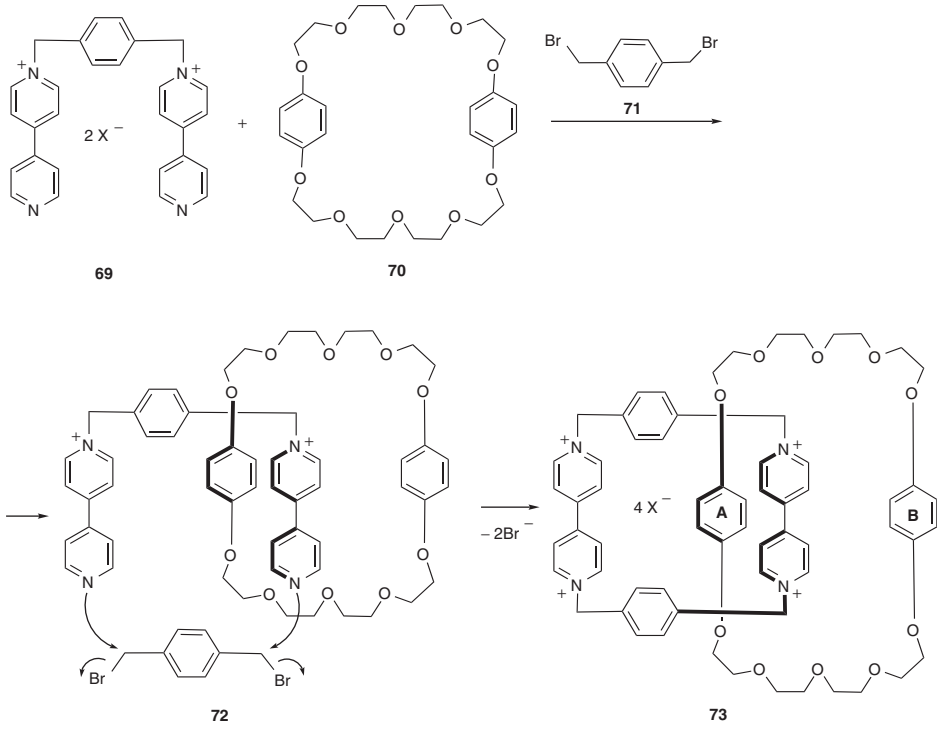
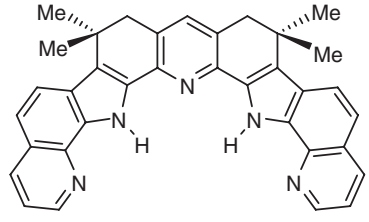


Figure 10.23 *π,π-Stacking assistant catenane synthesis (Ashton, P. R., Goodnow, T. T., Kaifer, A. E., Reddington, M. V., Slawin, A. M. Z. et al.,* Angew. Chem., Int. Ed.*, 1989, **28**, 1396. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

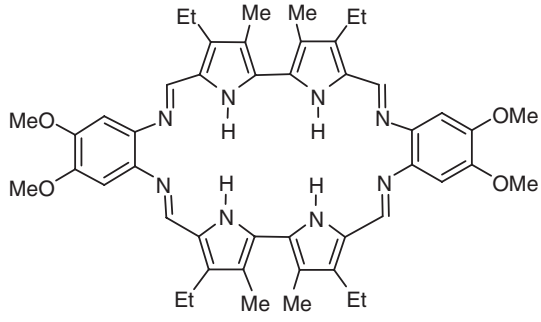
10.3 Problems

1. Pyrido[3,2-*g*]indoles are good receptors for various derivatives of urea. Suggest a structure for the 1:1 molar ratio complex between compound A and urea. Account for the significant differences between the values of the association constant K_a (l mol^{-1}, 18 °C) for the complexes of A with (a) *N*,*N*′-dimethylurea ($K_a = 118$), (b) 2-imidazolidone ($K_a = 13\,000$) and (c) barbituric acid ($K_a = 74\,200$).



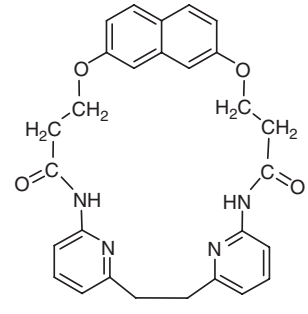
A

2. The so-called 'expanded porphyrin' B can form complexes with one or two molecules of methanol ($K_1 = 120$ and $K_2 = 30$ l mol^{-1}). Suggest structures for these complexes.



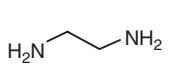
B

3. Macrocyclic receptor C forms a strong 1:1 complex with 9-methyladenine. What is the structure of the complex? (Hint: hydrogen bonding and stacking interactions stabilize the complex.)

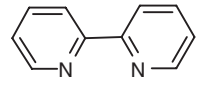


C

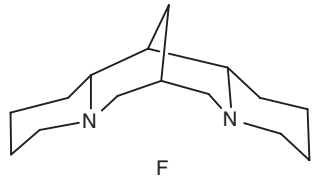
4. Crown ethers of suitable sizes (e.g., 18-crown-6) solubilize arenediazonium salts in non-polar media (e.g., chloroform). At the same time they decrease the rate of diazonium salt decomposition. Explain.
5. Complexes of the bidentate ligands D–F with the copper(II) ion are decomposed by acid in 0.006 s, 0.02 s and 295 min, respectively. Account for the differences in stability.



D

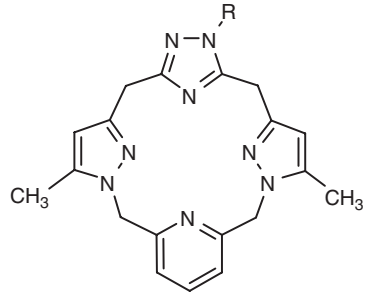


E



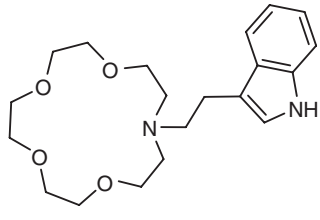
F

6. A CH_2Cl_2 solution of compound G cannot extract sodium or potassium picrates from their aqueous solutions. By contrast, a CH_2Cl_2 solution of compound H can remove these species. Explain.

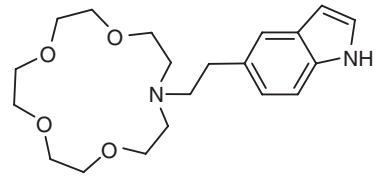


G (R = H)
H (R = (CH_3))

7. The indolyl residues in azacrown ethers I and J make their complexes with K^+ ion more stable. Suggest an explanation and write possible structures of the complexes.

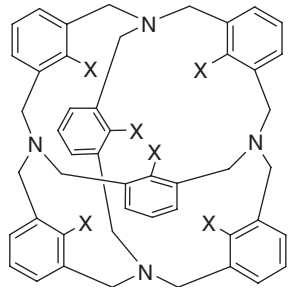


I



J

8. The cage compound K does not show any cation affinity. At the same time its hexafluoro derivative L binds potassium picrate forming the 1:1 complex with stability constant $\log K_s = 5.58$. Explain stressing the nature of the ligand–cation interaction and the possible structure of the complex.



K (X = H)
L (X = F)

9. Free tetramine **57** has little propensity to pick up a proton even in rather wet C_6D_6 or DMSO-d_6. However, once it is converted to monocation **58**, the second proton from residual water is entrained. The same phenomenon occurs when **57** is preliminarily converted into an Li^+ complex. Interpret these observations.
10. Write down a scheme for the preparation of [2]catenane similar to **68** but differing by the presence in the crown-containing ring of one tetrafluorinated phenyl ring. Suggest which of two translation isomers is preferable in this case?

10.4 Suggested Reading

1. Cram, D. J., The design of molecular hosts, guests, and their complexes (Nobel lecture), *Angew. Chem., Int. Ed. Engl*., 1988, **27**, 1009.
2. Lehn, J.-M., Supramolecular chemistry – scope and perspectives molecules, supermolecules, and molecular devices (Nobel lecture), *Angew. Chem., Int. Ed. Engl*., 1988, **27**, 89.
3. Pedersen, C. J., The discovery of crown ethers (Nobel lecture), *Angew. Chem., Int. Ed. Engl*., 1988, **27**, 1021.
4. Cram, D. J. and Cram, J.M., *Container Molecules and Their Guests*, Royal Society of Chemistry, Cambridge, 1994.
5. Gokel, G. W., *Crown Ethers and Cryptands*, Royal Society of Chemistry, Cambridge, 1991.
6. Lehn, J.-M. (ed.), *Comprehensive Supramolecular Chemistry*, Elsevier, New York, 1996.
7. Steed, J. W. and Atwood, J. L., *Supramolecular Chemistry*, John Wiley & Sons, Ltd, Chichester, 2000.
8. Special issue on supramolecular chemistry and self-assembly, *Science*, 2002, **295**, 2313–2556.
9. Special issue on supramolecular approaches to organic electronics and nanotechnology, *Adv. Mater*., 2006, **18**, 1227–1329.
10. Matsumoto, K. (ed.), *Heterocyclic Supramolecules I*, vol. 17 in series *Topics in Heterocyclic Chemistry*, Gupta, R. R. (ed.), Springer, Berlin, 2008.

11

Heterocycles and Twenty-First Century Challenges

She caught the white goose by the leg,
A goose – 'twas no great matter.
The goose let fall a golden egg
With cackle and with clatter.
A. Tennyson

In this chapter we discuss possible contributions of heterocyclic chemistry to help solve global problems of human society in the twenty-first century. In some cases this looks realistic only from a long-term perspective and the current achievements so far are modest. At the same time there are many fields where progress is appreciable and an appearance of innovative technologies based on heterocyclic compounds already occurs or is expected in the very near future.

11.1 Energy Problem

11.1.1 Biofuels

Diminishing hydrocarbon fuel reserves dictate an urgent search for alternative and renewable sources of energy. One of the rapidly developing fields is biofuels – organics obtained in a convenient form from widely abundant plant materials, such as sugar cane, algae, agricultural wastes, wood chips, etc. Currently, the two main biofuels are bioalcohol and biodiesel. Large-scale ethanol production has been successfully developed. The leading producers of ethanol as transport fuel are the United States and Brazil (70% of world production). In 2007, each of these countries manufactured more than five billion gallons of bioethanol. In Brazil, bioethanol is made from cane. The cane is shredded and squeezed. The cane juice from the crushers is diluted to a 20% sugar solution and pumped into large fermentation vats. Over 8 h, yeasts ferment the sugar to a 6–10% ethanol soup (see the chemical equations in Figures 5.4a, 5.8c). The yeasts are removed and either recycled to ferment more sugar or dried and sold as animal feed. The ethanol is then distilled

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

and dehydrated in stages to greater than 99% purity, denaturated, loaded into tanker trucks and delivered to fuel distributors. In the United States, bioethanol is made from cornstarch. Enzymes are needed to convert the starch to glucose, which is then fermented. This extra step means the Brazilian procedure is more efficient and less expensive. Bioethanol currently makes up about 13% of transportation fuel in Brazil and 4% in the United States.

Ethanol as a fuel suffers from the drawbacks of high volatilility, hygroscopicity and rather low energy density. In view of this, furan derivatives, especially 2,5-dimethylfuran (DMF), recently have been suggested as a possible alternative to bioethanol. Compared to ethanol, DMF has a higher energy density (by 40%), a higher boiling point (by 20 °C) and very low water solubility. Fructose is an effective starting material for the two-step synthesis of DMF since it can be prepared from glucose or produced directly from biomass. Fructose is first subjected to selective acid-catalyzed dehydration to produce 5-hydroxymethylfurfural (HMF). Next, HMF is converted to DMF via hydrogenation over a copper catalyst (Figure 11.1). HMF itself cannot serve as biofuel due to its high boiling point, but its derivatives are multipurpose intermediates that can be used as fuel precursors or replace petroleum-based building blocks for plastics, pharmaceuticals and fine chemicals. Of course, the commercial use of heterocyclic compounds like DMF as liquid transport fuel is a long way off. Nevertheless, rather successful initial steps have been undertaken.

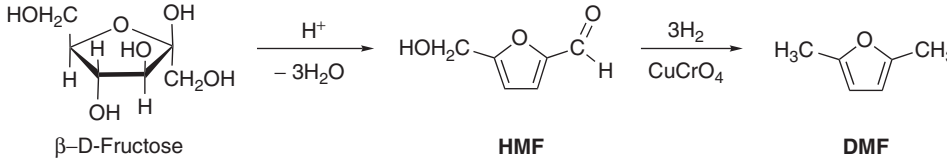


Figure 11.1 *Conversion of fructose into 2,5-dimethylfuran Reprinted by permission from Macmillan Publishers Ltd. (Román-Leshkov, Y., Barrett, J., Liu, Z. Y. and Dumesic, J. A.,* Nature*, 2007,* ***447****, 982. © 2007).*

11.1.2 Hydrogen as a Fuel

Great expectations are currently held for the hydrogen economy – wide distribution of molecular hydrogen as an ecologically pure and unlimited source of energy. Indeed, hydrogen is already being used as a fuel in hybrid cars and in rockets for spaceship launches.

Hydrogen releases energy when it is oxidized by oxygen to produce water. However, almost all the hydrogen on our planet exists in a bound state: as water, cellulose or hydrocarbons. Since O—H and C—H bonds are thermodynamically very stable, production of hydrogen from these materials reguires more energy than can be released from H_2 as fuel. Another drawback is the difficult handling of hydrogen because of its explosive character and extremely low boiling point (−252.9 °C). In principle, all these difficulties need to be overcome in order to employ hydrogen for production as a renewable raw material and energy source and to develop convenient and effective methods for hydrogen storage. Over the past decade considerable progress has been achieved in both these directions. Thus, there is the first full-scale hydrogen energy plant in Denmark, started in 2007, where hydrogen is produced by using excess wind power to electrolyze water to give oxygen and hydrogen. Molecular hydrogen can also be obtained biologically, such as by anaerobic fermentation of carbohydrates with diverse enzymes. Recently, it was discovered that, when algae is deprived of sulfur, its photosynthesis can be switched to produce hydrogen. An algae bioreactor based on this phenomenon has already surpassed 7–10% energy efficiency (the conversion of sunlight into hydrogen).

Molecular hydrogen used in modern vehicles is stored in on-board tanks in liquid or compressed form. This technology has many problems of safety, efficiency and density. Hence efficient

hydrogen storage has become one of the central limiting points of a hydrogen economy. A promising alternative is chemical storage by which a catalytic chemical reaction or just heating allows gaseous H_2 to be released from hydrogen atoms covalently bound in a molecule. Many suggestions for this purpose include metal hydrides, ammonia, carbohydrates, hydrocarbons, formic acid and others. A practical storage material must be of low molecular weight and have a high percent of available hydrogen. A leading candidate satisfying these demands is the ammonia-borane complex H_3NBH_3. Its thermal dehydrogenation in solution or in the solid state can be achieved via a transition metal-catalyzed reaction. The mixed nickel complex **2** of 1,5-cyclooctadiene and the stable heterocyclic carbene, 1,3,4-triphenyl-4,5-dihydro-1H-1,2,4-triazole-5-ylidene (**1**), is especially effective (Figure 11.2a, b). Along with its high stability, **2** provides almost quantitative hydrogen release: 18 wt% H versus the theoretical 19.6 wt%.[1] A proposed mechanism of dehydrogenation includes metal-assisted activation of a B—H bond, the formation of a σ-complex and subsequent β-elimination of H_2 (Figure 11.2c).

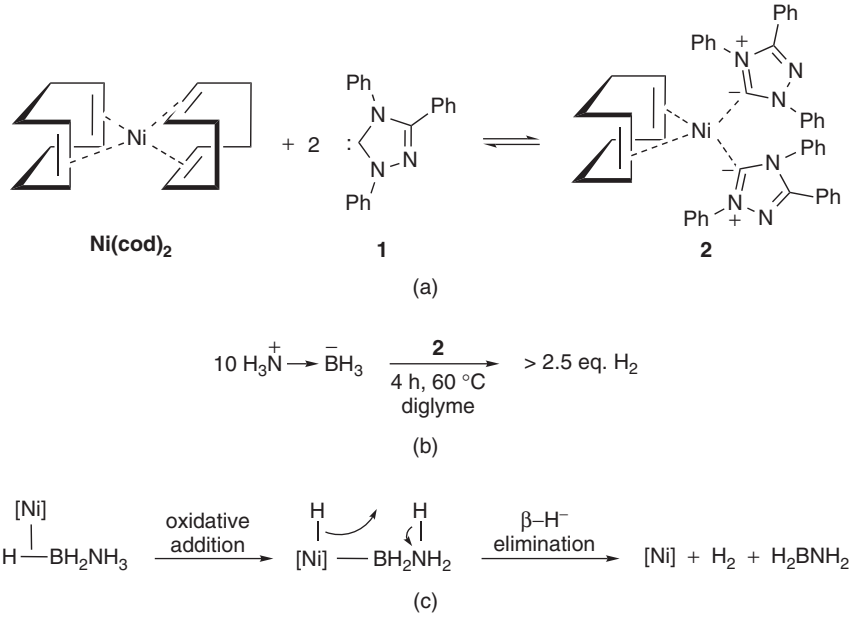


Figure 11.2 *Catalytic dehydrogenation of ammonia-borane complex: (a) preparation of Ni-carbene catalyst, (b) overall equation for H_2 formation, (c) schematic mechanism of dehydrogenation (Reprinted with permission from Keaton, R. J., Blacquiere, J. M. and Baker, R. T.,* J. Am. Chem. Soc.*, 2007,* ***129****, 1844. © 2007 American Chemical Society).*

Another attractive idea is to store molecular hydrogen inside light and highly porous materials in which it might be held by adsorption forces. The most promising among them are recently discovered metal–organic frameworks (MOF) – crystalline compounds consisting of metal ions or clusters coordinated to polyfunctional organic ligands such as dicarboxylic and polycarboxylic acids, triazoles or tetrazoles. For example, 1,3,5-tris(tetrazolyl)benzene (Figure 11.3a) on treatment with $MnCl_2$ produces a MOF, whose structure is shown schematically in Figure 11.3b. This

[1] In 2003 the United States Department of Energy set a goal to reach 9 wt% H for overall storage system weight by 2015.

material demonstrates total H_2 uptake of 6.9 wt% at 77 K and 90 bar, which at 60 g H_2 l^{-1} corresponds to a 85% storage density of that of liquid hydrogen. Even more significant, the H_2 binding energy for this MOF is about 10 kJ mol^{-1}. This is much higher than that for most other MOFs (4–7 kJ mol^{-1}) though still lower than the theoretically predicted 15 kJ mol^{-1} needed to maximize the amount of adsorbed H_2 at ambient temperature and pressure limits of 1.5–20.0 bar.

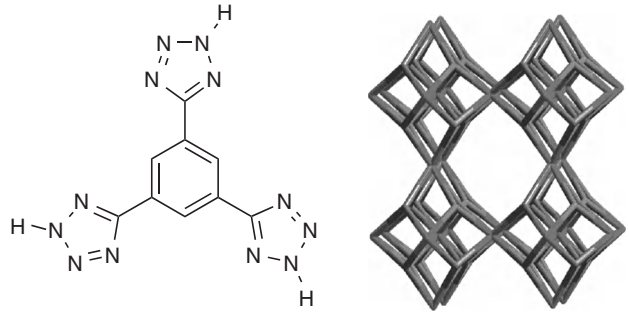


Figure 11.3 *1,3,5-Tristetrazolylbenzene (a) and schematic representation of MOF (b) forming at its interaction with $MnCl_2$ (Reprinted with permission from Dinca, M., Dailly, A., Liu, Y. et al.,* J. Am. Chem. Soc.*, 2006,* ***128****, 16876. © 2006 American Chemical Society).*

Thus, heterocyclic compounds can serve in hydrogen technologies as: (i) a feedstock (carbohydrates) for hydrogen production, (ii) hydrogen storage materials and (iii) catalytic systems for hydrogen release.

11.1.3 Direct Use of Solar Energy

Solar radiation is an inexhaustible source of clean energy. Each hour our planet receives from the Sun an amount of energy that is equal to the annual demand of all the Earth's population. Solar energy heats the Earth's surface and atmosphere, is transformed into air and water streams and is stored in chemical bonds as a result of photosynthesis. Solar energy is already used indirectly in an ecologically clean form, for example, via hydro-electric power stations, windmills, watermills or geothermal springs. However, other convenient and efficient methods of utilization and storage of solar energy are now on the agenda. This is one of the most challenging goals of modern science.

There are two main approaches to this problem: (i) direct conversion of solar energy into electricity and (ii) development of various photosyntheic devices, mimicking natural photosynthesis and often referred to as 'artificial photosynthesis'. The principal difference between these two approaches is that while the former is based almost entirely on physical processes, the basis of artificial photosynthesis is chemical transformations. We note Chapter 6 of this book first considers the topic of 'artificial photosynthesis'.

Artificial Photosynthesis

In the primary step of natural photosynthesis, plants capture the Sun's photons and utilize their energy to transform water. The free electrons and protons thus formed are fixed in the form of NADPH-coenzyme to reduce CO_2 into carbohydrates. A unique feature of natural photosynthesis is that it envolves many dozens, if not hundreds, of inorganic and organic components that are highly organized into supramolecular structures, operating with a great efficiency. Scientists are

still unable to compete successfully here with Nature. Fortunately, photosynthetic pathways can be artificially modified and even simplified, most easily by dividing them into two fundamental photosynthetic stages. One of the leading ideas is to interrupt photosynthesis at the stage of water splitting to obtain molecular hydrogen as an effective fuel and chemical reagent. Schematically, a simple device is shown in Figure 11.4. Its principle elements are: (i) an antenna complex for photon capture connected with the reaction center for charge separation, (ii) catalysts for two dark reactions – oxidation of water molecules and reduction of protons into H_2 – and (iii) a membrane to separate sections where oxygen and hydrogen are formed.

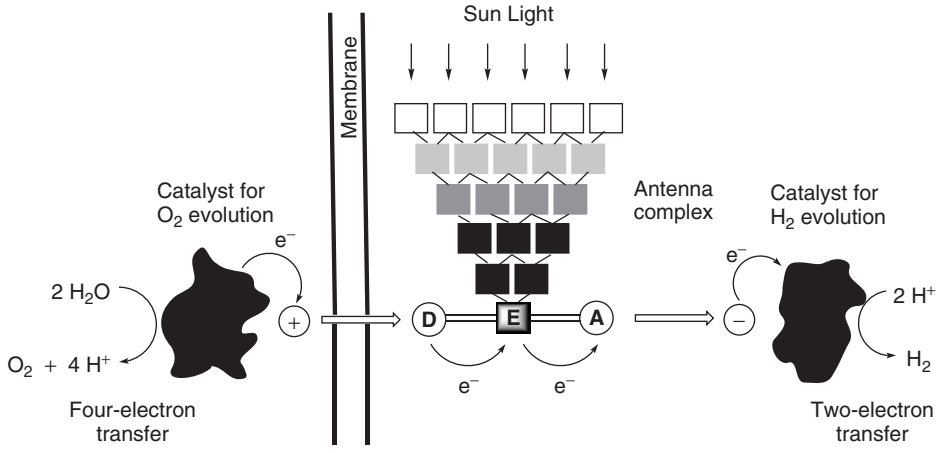


Figure 11.4 *Schematic representation of an artificial photosynthetic device for water splitting (adapted from Balzani, V., Credi, A. and Venturi, M.* Molecular devices and Machines*, Wiley-VCH, Weinheim, 2008. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

Though fully integrated devices of such type have not yet been constructed, extensive studies of the separate stages are in progress. Thus, heterocyclic compounds such as porphyrins, phthalocyanins or polypyridine metal complexes, often in conjunction with π-extended aromatics, are widely employed as light-harvesting and charge-separating units. Principally, to absorb light with a maximum of efficiency an artificial antenna system, similar to chlorophyll molecules tightly packed in the protein matrix of a green leaf, must include a large number of interconnected chromophore units. They are depicted in Figure 11.4 as small squares. In practice, a dendrimer architecture is commonly used. A typical example is shown in Figure 11.5. Light photons are captured by five 9,10-diethynylanthracene residues attached to hexaphenylbenzene. Energy from their photoexcited state is then transferred via the so-called Förster mechanism to a nearby porphyrin fragment (P) that donates one electron from its singlet excited state to a fullerene unit. All electron-transfer steps proceed with high quantum yield (80–100%) within an optimal picosecond time scale and result in the formation of charge-separated state P^+-C_{60}^-.

Unlike the antenna complex, where a single electron transfer mechanism is realized, oxidation of water and reduction of protons are multielectron processes (Figure 11.4) which demand special catalysts providing a charge pool. A search for such catalysts as well as synchronization of all elements of such a photosynthetic device is a substantial challenge. Until recently, mostly precious metals such as platinum and their oxides were tested for this purpose sometimes in combination with an organic metal complex. However, due to the tremendous scale needed for fuel production, artificial photosynthetic catalysts should be made from cheap and abundant materials.

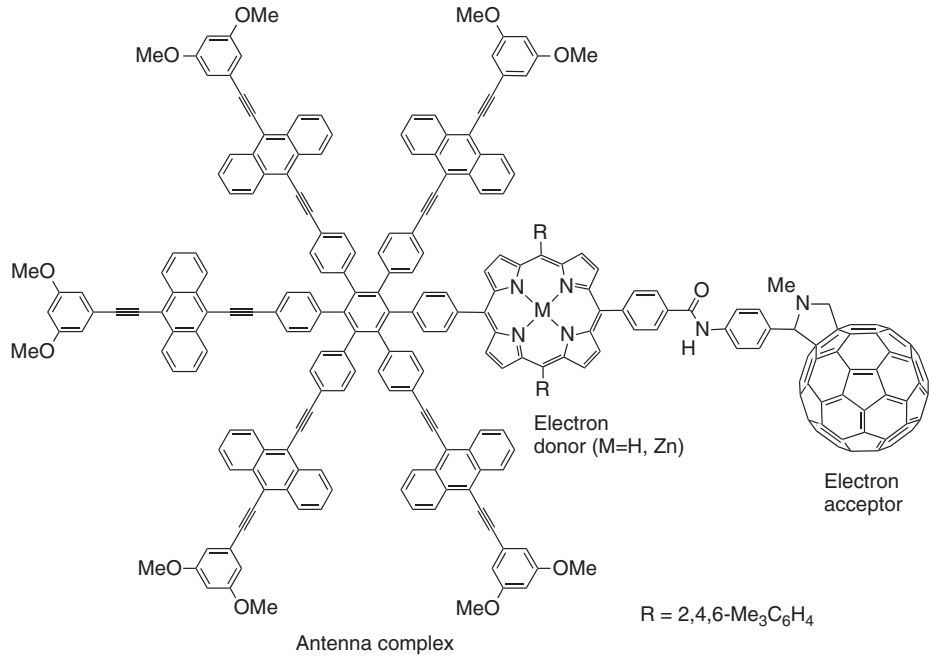


Figure 11.5 *Antenna – reaction center molecular system for artificial photosynthesis (Reprinted with permission from Kodis, G., Terazono, Y., Liddell, P. A., Andréasson, J., Garg, V., Hambourger, M., Moore, T. A., Moore, A. L. and Gust, D.,* J. Am. Chem. Soc.*, 2006,* ***128****, 1818. © 2006 American Chemical Society).*

A remarkable achievement recently discovered is an oxygen-evolving catalyst composed of Co^{2+} and HPO_4^{2-} ions (M. W. Kanan and D. G. Nocera, *Science*, 2008, **321**, 1072). Cobalt ions act as direct acceptors of electrons from water whereas phosphate anions serve as proton transfer species. The water splitting proceeds at neutral pH, 1 atm, at room temperature; recovery of precipitated cobalt species is also possible.

Another development of artificial photosynthesis consists in the employment of a 'reducing potential' arising on photoexcitation of a light antenna for the direct conversion of CO_2 into either CH_3OH ('methanol economy') or CO, which are valuable both as fuel and multipurpose raw chemicals. Ruthenium(II) polypyridine carbonyl complex, cobalt(II) trisbipyridine and cobalt(III) macrocycles catalyst with a photosensitizer reduce CO_2 into CO with good quantum yield and selectivity. A mechanistic pattern is exemplified with rhenium(I) bipyridine complexes **3** (Figure 11.6). The photoexcited state **3*** is quenched by triethanolamine (TEA) acting as one-electron donor. Since the 18-electron valence shell of a rhenium atom in **3** is filled, the extra electron from TEA occupies the lowest empty π-orbital of the bipyridine producing radical-anion $\mathbf{3}^{-\bullet}$. This triggers a loss of ligand L^- and the formation of **4** in which the outer sphere of the Re atom becomes unfilled. Therefore, **4** can now react with CO_2 to give adduct **5**.[2] Within this adduct the CO_2 molecule undergoes a two-electron reduction, adding one electron intramolecularly as depicted by the curved arrow in **5** and another electron from the radical-anion $\mathbf{3}^{-\bullet}$. As a result, a

[2] Structure **5** is conditional since the exact mode of CO_2 binding remains unknown.

CO molecule evolves and initial complex **3** is restored via addition of L^- to **6**. Thus, a bypyridine residue here participates not only as a bidentate ligand but also as an electron acceptor stabilizing the excited state and as an effective electron-carrier in the reduction process.

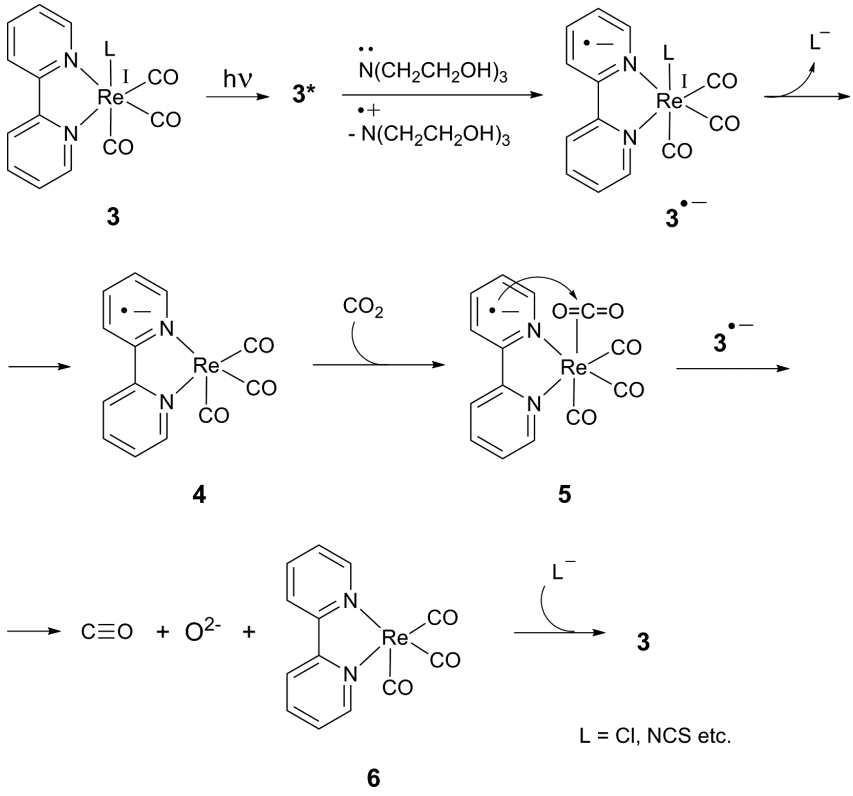


Figure 11.6 *Photocatalytic reduction of carbon dioxide into CO (Reprinted with permission from Takeda, H., Koike, K. and Ishitani, O.,* J. Am. Chem. Soc.*, 2008,* ***130****, 2023. © 2008 American Chemical Society).*

Solar Cells

From a practical point of view, electricity is the most convenient form of energy. Therefore, direct conversion of sun radiation into electric current seems especially attractive. The first 'sun-electricity' devices known as solar cells (or 'photovoltaic' cells) appeared as far back as the 1950s. Their further development was strongly accelerated by the space industry, growing energy shortages and climate change.

Classical solid-state solar cells are made from silicon or rarely from other semiconductor materials (CdTe, Ge, etc.). These semiconductors conduct electricity but have a much higher resistance than metals. The resistance of a semiconductor is very sensitive to impurities. One can precisely improve the properties of a semiconductor by *doping*, that is, adding different kinds and amounts of impurity atoms. In terms of electrical properties, doped silicon is divided into n-type and p-type. The n-type is enriched by free electrons, while the p-type is produced by atoms missing an electron (called a hole; Figure 11.7a). When p-type and n-type silicon are joined to form the

so-called p-n junction (Figure 11.7b), some of the free electrons in the n-region diffuse across the junction and combine with holes to form negative ions. In so doing they leave behind positive ions at the donor impurity sites. The combining of electrons and holes depletes the holes in the p-region and the electrons in the n-region near the junction. Thus, a depletion zone (Figure 11.7c) arises near the junction that interrupts any further electron transfer due to coulombic forces. Other electrons from the n-region cannot migrate because they are repelled by the negative ions in the P region and attracted by the positive ions in the N region. However, if one applies a voltage with the polarity indicated in Figure 11.7d electrons will flow with very small resistance to the positive pole providing conductivity. Contrary to this, a reverse voltage (Figure 11.7e) impedes the flow of electrons across the junction since now they should be driven away from the junction, preventing conduction. Such conductivity is the basis of diodes, widely used in electronics, and of solar cell action.

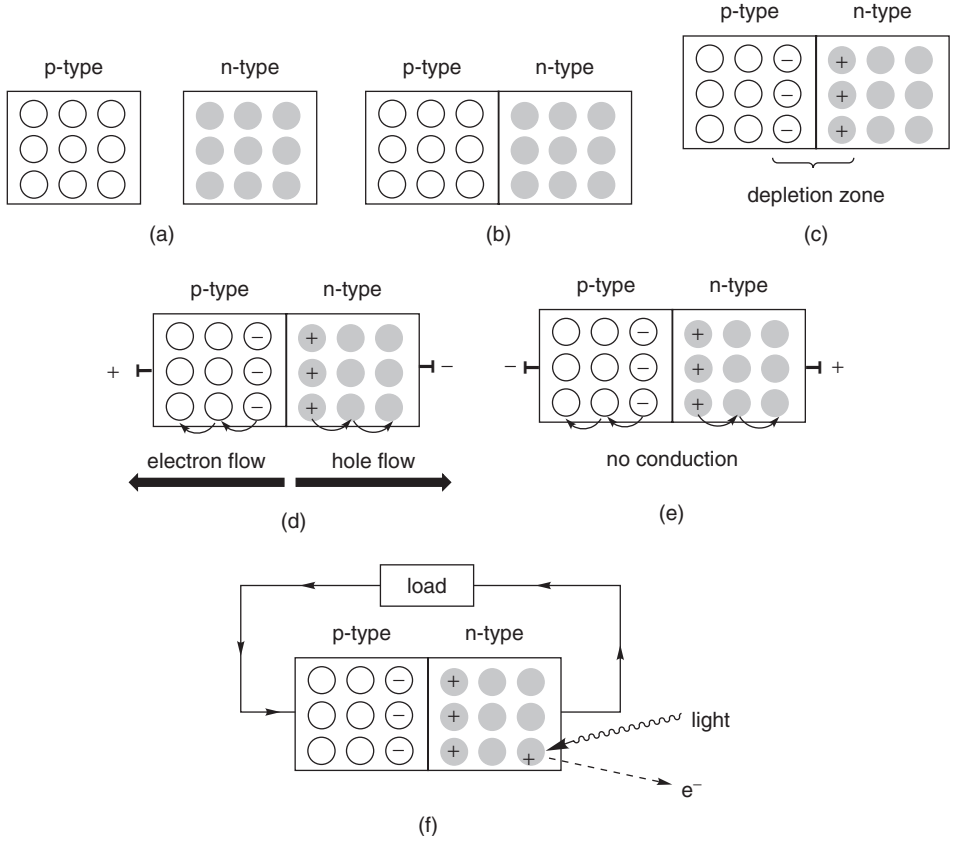


Figure 11.7 *Schematic representation of different states of a semiconductor p,n-junction (a–e) and the principle of operating of a solar cell (f). White circles designate electron-acceptor atoms (holes) and gray circles electron-excessive centres.*

Schematically, how a solar cell operates is shown in Figure 11.7f. When light photons strike the p-n junction, some electrons receive enough energy to leave the silicon atoms and become mobile

electrons. When an electron leaves the atom, a hole is created at the atom's electron site. The intrinsic electric field at the p-n junction forces the mobile electron to move towards the n-side. The intrinsic electric field also causes the electrons from the neighboring atoms to fill the vacant sites successively, so the hole appears to move towards the p-side. The mobile electrons and the holes move in opposite directions, separating the positive and negative charges. This builds up an electric potential and when the p-n junction is connected to an external circuit through the front and back contacts, the circuit provides a pathway for the separated electrons and holes to recombine, resulting in an electric current flowing through the circuit. This process can be used to recharge batteries or to power various instruments.

Despite increasing efficiency (now 15% on average) silicon solar cells have disadvantages, in particular, complicated manufacture and sensitivity to environmental conditions, causing other types of solar cells to be developed. Apparently, the most successful among them is the so-called dye-sensitized solar cell (DSC). A typical DSC (Figure 11.8a) comprises a light-transparent glass carrying a thin layer of highly porous nanocrystalline TiO_2 as a semiconductor. It is soaked with a dye that provides effective light absorption. Between this part of the cell (called photoanode) and a cathode, an electrolyte is placed which contains a redox-system (commonly an iodide/triiodide couple) capable of fast regeneration of the oxidized dye. Numerous functional dyes or organic conjugated polymers are employed as photosensitizers and conducting materials. Presently the most efficient sensitizers for DSC appear to be ruthenium(II) dipyridine and tripyridine complexes such as those depicted in Figure 11.8b.

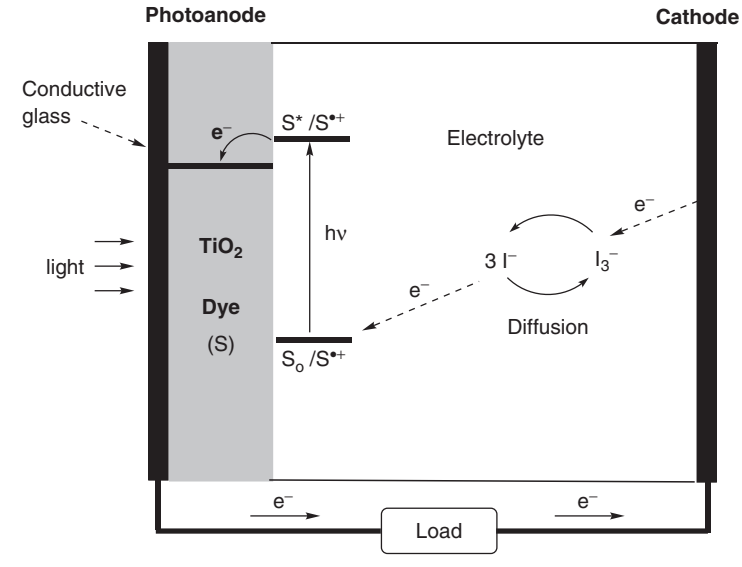
Sunlight enters the cell through the transparent left contact, striking the dye (S) on the surface of the TiO_2. An excited state of the dye (S*) is formed, from which an electron can be 'injected' directly into the conduction band of the TiO_2, and from there it moves by diffusion to the clear anode on the left. The dye molecule missing an electron ($S^{\bullet+}$) will decompose if another electron is not provided. The dye strips one electron from iodide in the electrolyte next to the TiO_2, oxidizing it into triiodide. This reaction must occur quickly compared to the time that it takes for the injected electron to recombine with the oxidized dye molecule. The triiodide then recovers its missing electron by diffusing to the cathode which re-introduces the electrons after flowing through the external circuit.

DSCs have reached conversion efficiencies exceeding 11%. With a very low cost of fabrication, compatibility with flexible substrates and mechanical robustness, DSCs are currently considered to be the most promising future solar technology. Wide commercial applications by 2020 are forecast.

Another similar device where heterocyclic substrates play an important role is the hybrid biofuel cell (Figure 11.9a). In such a cell, a dye with a photooxidative potential, for example, an appropriate porphyrin derivative (Figure 11.9b), is employed for hydrogen production from a natural source such as glucose or ethanol. Organic fuel placed between two electrodes is primarily dehydrogenated by soluble NAD^+ enzyme producing its NAD-H form, oxidized components and protons. NAD-H immediately ejects an electron in radical cation $S^{\bullet+}$ regenerating both photosensitizer and NAD^+ for the next cycle. Protons move through a special membrane to the cathode and are reduced to H_2. The electron circuit is completed by a connecting wire that passes the electrons obtained at the anode to a counter-electrode.

Solar Energy Conservation

Investigations concerning solar energy conservation are of great importance. Certain organic compounds, including heterocycles, can absorb sunlight and to store the solar energy by transforming it into an energy-rich product. In practice, all such transformations are represented by valence or *cis–trans* isomerizations. A classical example of such compounds is *trans*-azobenzene which is converted photochemically into its energy-rich *cis*-isomer (Figure 11.10a). Indigo *N,N*-diacyl derivatives similarly undergo *trans–cis* photoisomerization



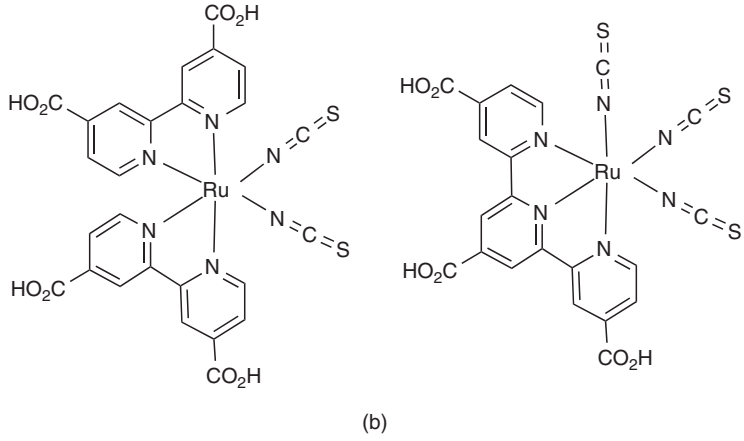
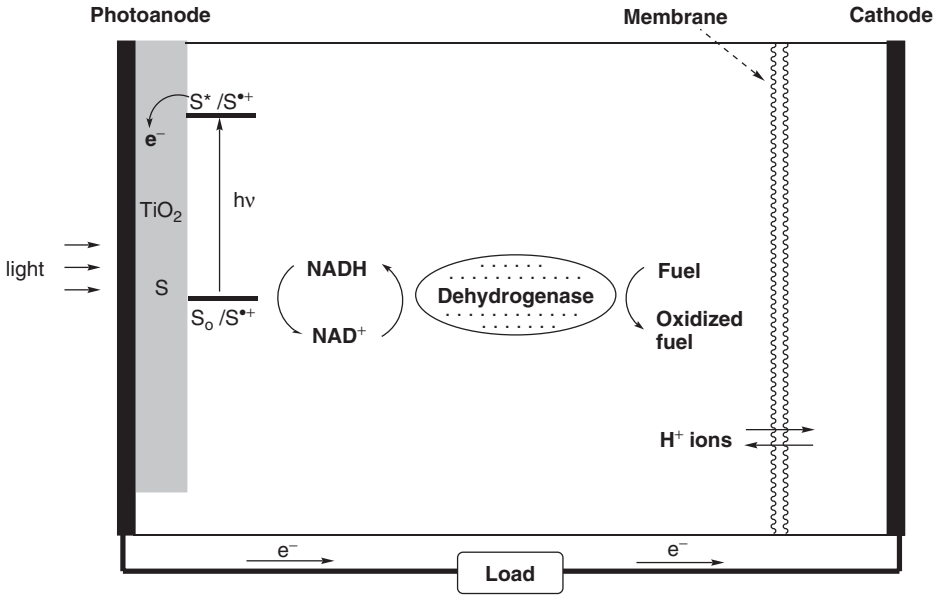


Figure 11.8 *(a) Schematic representation of dye-sensitized solar cell. (b) Examples of the ruthenium sensitizers (adapted with permission from from Grätzel, M.,* Inorg. Chem.*, 2005,* ***44****, 6841. © 2005 American Chemical Society).*

(Figure 11.10b). 1-Methyl-5-phenyl-Δ^2-pyrazoline undergoes valence isomerization to form energy-rich 2-phenylcyclopropylazomethane (Figure 11.10c). The stored energy resulting from these conversions amounts to around 10 kcal mol^{-1}. Nonheterocyclic norbornadiene containing a quadricyclane system (Figure 11.11d) holds the record for the largest energy conservation capacity (26 kcal mol^{-1}). The isomerized energy-rich products need to be reasonably stable under ordinary conditions and, for the compounds to be of practical use, an additional requirement is that the stored energy should be liberated on demand, preferably in the form of heat. The energy-rich product must usually be stimulated, e.g., by catalysts or mild heating, to effect energy



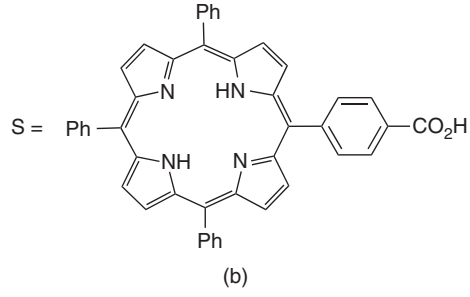


Figure 11.9 *Schematic representations of: (a) hybrid photobiofuel cell, (b) porphyrin photosensitizer (Reprinted from Moore, A. L., Gust, D. and Moore, T. A.,* L'Actualité Chimique*, 2007,* ***308/309****, 50).*

release. Once the energy is discharged in the form of heat, the molecule reverts to the original low energy state.

Reversible phototransformations could potentially be employed in small heating systems such as domestic heaters. On bright sunny days a substance placed in a specialized tank on the roof would accumulate solar energy which could then be utilized to heat the dwelling at night or in cool weather. Figure 11.11 is an energy diagram of the processes which occur during the storage of solar energy. An initial product R, upon absorption of a quantum of light, is transferred to an excited triplet state (R*). Further stabilization is achieved by conversion to the high energy product P which stores part of the absorbed energy (ΔE) as chemical bond energy.

Such transitions can be of practical importance provided that the substance can be cycled repeatedly. However, experience has demonstrated that in due course, most substances are gradually destroyed. Thus, the search for materials capable of withstanding many energy conversion cycles has begun in earnest.

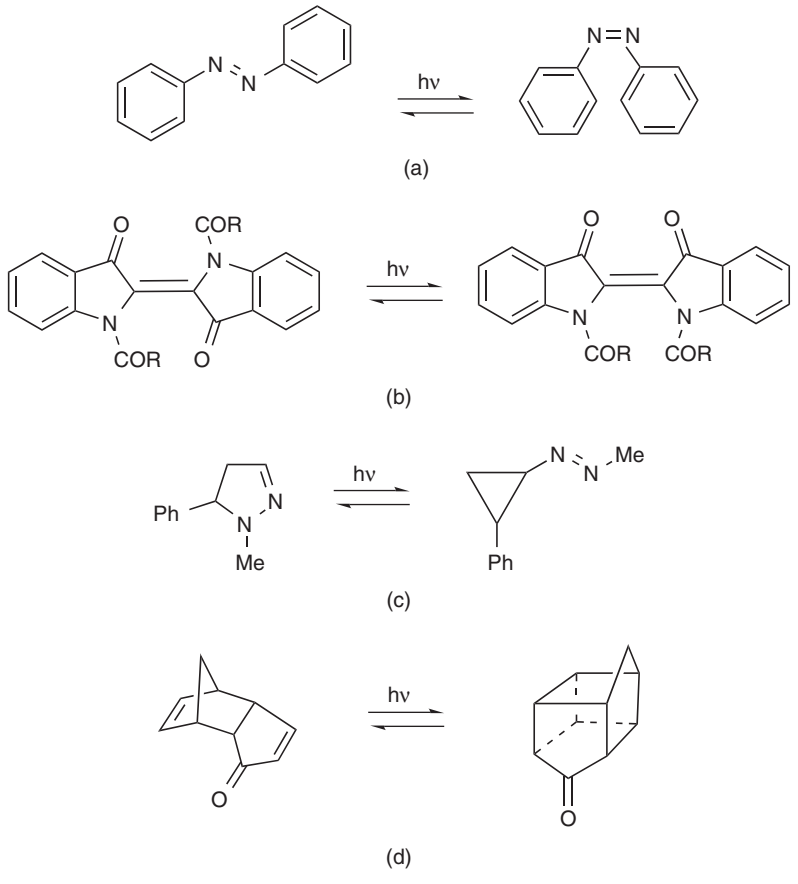


Figure 11.10 *Examples of photoisomerizations utilized for solar energy conservation (Scharf, H.-D., Fleischhauer, J., Leismann, H., Ressler, I., Schleker, W. and Weitz, R.,* Angew. Chem., Int. Ed. Engl.*, 1979,* ***18****, 652. © 1979 by Verlag Chemie, GmbH, Germany).*

11.1.4 Conducting Materials

Just as for energy storage, the development of effective conducting materials is also of great importance. The synthesis of such materials has been a major achievement of heterocyclic chemistry in the past three decades, as we now introduce with some background.

An electric current is caused by movement of free electrons or other charged particles through a conductor. Normally, conductive material resist this movement. Such resistance, R, is directly proportional to conductor length and inversely proportional to its cross-sectional area, A (equation 11.1). Parameter ρ, called *resistivity*, reflects the intristic ability of the material to conduct at l and A equal to 1. The inverse of *resistivity* is *conductivity,* σ (equation 11.2), expressed in units S cm^{-1} or Ω^{-1} cm^{-1} (S – Siemens, Ω – ohms).

$$R = \frac{\rho l}{A} \tag{11.1}$$

$$\sigma = \frac{1}{\rho} \tag{11.2}$$

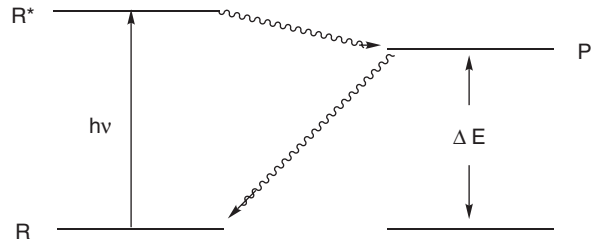


Figure 11.11 *Energy diagram for the conservation of light energy.*

Commonly, all materials are graded as conductors on the basis of their ρ or σ values. Thus, conventional metals such as Fe, Cu or Ag exhibit the highest conductivity ($\sigma = 10^6 - 10^8$ S cm^{-1}). The opposite side of the conductivity scale is occupied by insulators whose conductivity is negligible: 10^{-18} S cm^{-1} for quartz or 10^{-10} S cm^{-1} for glass. Classical semiconductors are in the middle: 10^{-2} S cm^{-1} for Ge and 10^{-5} S cm^{-1} for Si. Unlike semiconductors whose conductivity rises with increasing temperature, many conducting materials demonstrate different behavior. Their conductivity gradually increases on lowering the temperature and at a certain point (the critical temperature, T_c), close to absolute zero, the resistance completely disappears. This special state is called superconductivity. An electric current flowing in a circle of superconducting wire can persist indefinitely with no power source. Though superconducting materials are already used in particle accelerators, space apparatus, magnetic resonance imaging etc., their wider application is hindered by the necessity of their strong cooling with expensive liquid helium. Therefore the search for so-called high-temperature superconductors is now a prime focus of scientific studies.

The majority of organic substances are insulators. However, if a compound has an extended conjugated chain, it can behave as a semiconductor or even a conductor, due to the well known ability of π-electrons to delocalize. The simplest examples are graphite, graphene or nanotubes having metallic conductivity ($\sigma \sim 10^6$ S cm^{-1}). Another type of organic conductor includes conjugated polymers with polyacetylene as their parent representative (Figure 11.12a). However, polyacetylene itself exhibits only semiconductor properties because, as in other polyenes, for example in 1,3,5-hexatriene, π-electron delocalization is not as effective as in benzene or graphite, and therefore their C—C bonds alternate with lengths about 1.35 and 1.45 Å. However, in polyene-based cations, radicals or radical-cations, the delocalization strongly increases and bond lengths become perfectly equalized (Figure 11.12b–d). This gave scientists the idea of doping polyenes to enhance conductivity. Indeed, chemical or electrochemical oxidation of polyacetylene leads to the appearance of a great number of free electrons and holes in its structure that dramatically increases conductivity. Thus, exposure of polyacetylene to I_2 produces material with $\sigma = 10^3 - 10^5$ S cm^{-1}. The doping results in the formation of partially delocalized radical ions called polarons (Figure 11.13a) whose migration along the polymer backbone creates a high conductivity. In 2000, Shirakawa, MacDiarmid and Heeger shared the Nobel Prize in Chemistry for their achievements in the field of conducting polyacetylenes.

This success has spurred the creation of many other polyacetylene-like organic polymers including heterocyclic ones, such as *para*-polyphenylenes, polythiophenes, polypyrroles and polyindoles (Figure 11.13b–e). Poly(3-hexyl)thiophene, the first organic polymer, showed superconductivity below 2.5 K. The main advantages of organic superconductors over their ceramic and metallic counterparts are low cost, lower density and the possibility of fine-tuning their electrical properties. These compounds are prepared by chemical or electrochemical oxidative polymerization of the parent heterocycles. Films formed from the heterocyclic polymers are good conductors and have

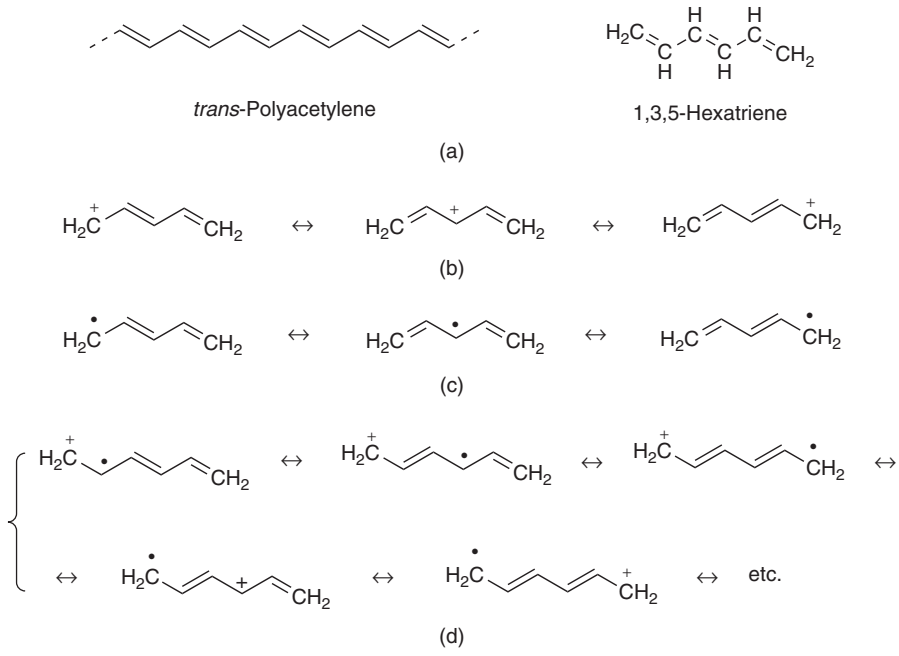


Figure 11.12 *(a) Localized double bonds in* trans*-polyacetylene and 1,3,5-hexatriene and bond-, charge- and electron delocalization in: (b) pentadienyl cation, (c) pentadienyl radical, (d) radical cation of 1,3,5-hexatriene.*

found extensive applications in the manufacture of chemically modified electrodes and sensors, solid batteries and diverse composite materials. Synthetic electroconductive nylon or polyester fibers covered with a thin layer of polypyrrole have been used as a coating for airplane fuselages to prevent the reflection of radar signals. The same textile has also been used to shield electronic equipment including computers and control systems.

Another valuable property of conjugated polymers is their deep color resulting from a very small energy gap between the highest occupied and the lowest empty molecular orbitals. Thus, all polymers shown in Figure 11.13 with large enough monomer units are normally black. However, their optical properties can be modified by external stimuli (solvent, temperature or applied potential) or structural changes. This response is caused by twisting of the polymer backbone that disrupts conjugation. Due to their interesting optical properties, conjugated polymers are especially attractive as both chemosensors and as sensitizers in dye-sensitized solar cells.

Excitingly, the polyacetylene motif is found in living nature. Thus, the brown-black pigment melanin found in animal skin, hair, fur and some internal organs (brain, ear, eyes, etc.) has in its structure a polyindolequinone chromophore (Figure 11.13f). Actually, melanin consists of a mixture of oligomers forming amorphous black particles of different sizes and shapes. Apparently, the main biological function of melanin is protection of the organism from harmful ultraviolet (UV) radiation. Melanin dissipates more than 99.9% of the absorbed UV radiation as heat, preventing the generation of free radicals. Common sunburn, as well as the natural dark-colored skin of people living near the equator is nothing more than an organism's self-protection against the sun's radiation via extensive melanin biosynthesis.

Apart of polyindole-based melanin (eumelanin) there is another type of melanin called pheomelanin. It differs from eumelanin with its brown-red color caused by the presence in its structure of

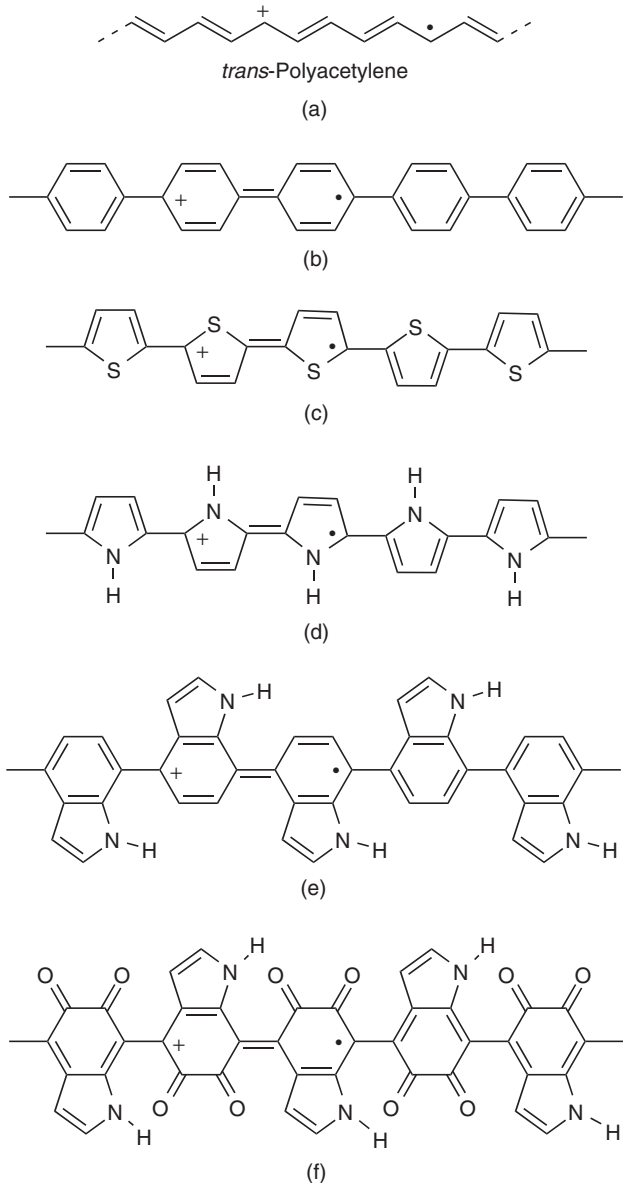


Figure 11.13 *Structures of doped conducting polymers: (a) polyacetylene, (b) p-polyphenylene, (c) polythiophene, (d) polypyrrole, (e) polyindole, (f) fragment of eumelanin pigment.*

a benzothiazine or benzothiazinone chromophore (Figure 11.14a, b). Pheomelanin is found in hair and skin where it can coexist with eumelanin. Large amounts of pheomelanin are concentrated in the lips, nipples or glands and especially in red hair. Both eumelanin and pheomelanin are biosynthesized from the amino acid tyrosine, which is primarily converted into dopa (Figure 11.14c). The heterocyclization in the case of pheomelanin proceeds with the participation of L-cystein and results in the formation of cysteinyldopa as an intermediate (Figure 11.14d).

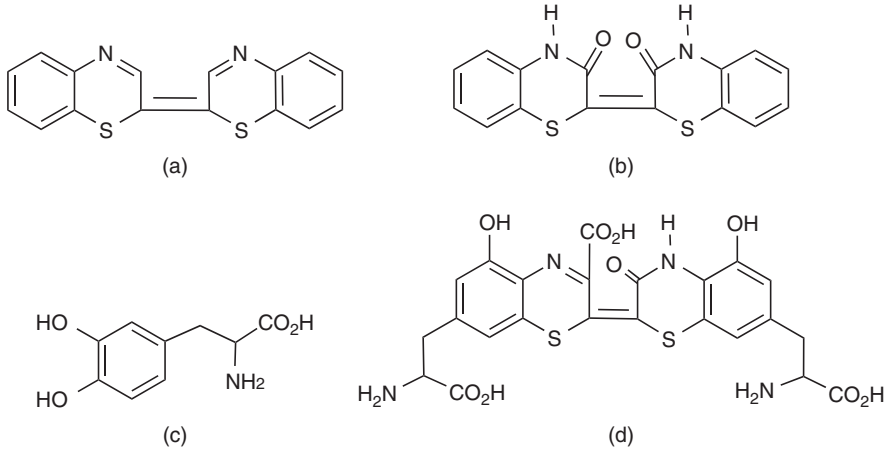


Figure 11.14 *(a, b) Benzothiazine chromophores occuring in pheomelanins. (c, d) Dopa and cysteinyldopa – biosynthetic precursors of eumelanins and pheomelanins.*

After doping, the black particles of melanin demonstrate high conductivity, $\sigma \sim 1$ S cm^{-1}. Melanin's conductivity is thought to be connected with its biological functions especially in brain neurons, eyes and inner ear. Charles Darwin first noticed that white cats with blue eyes, suffering a lack of melanin, are almost always deaf. Today, it is established that many human diseases such as melanoma, Parkinson's disease, albinism or deafness are influenced by melanin deficiency.

As a good semiconductor melanin has been successfully tested in various electronic devices as an innovative sensor, energy generator or switching agent. Obviously, further mimicking of melanin's functions could be productive technologically.

Ion-radical salts or charge transfer complexes compose a second important class of organic conductors. Their preparation requires a strong electron donor, which can be readily, polarized and a powerful electron-withdrawing group. Well known tetrathiafulvalene (TTF; Figure 2.11) was the first electron-donating component widely used for such materials. Its ion-radical salt with tetracyanoquinodimethane (TCNQ; see Figures 2.12 and 2.14a), first obtained in 1972, was one of the earliest organic compounds to display conductivity of the metallic type ($\sigma = 1.7 \times 10^4$ S cm^{-1}), initiating the search for more effective electron donors and acceptors.

The search for electron acceptors initially focused on π-deficient heterocycles with strong electron-withdrawing groups. 2,4,6-Tricyano-1,3,5-triazine (**7**), 3,6-dicyano-1,2,4,5-tetrazine (**8**) and the bisthiadiazole system (**9**) based on TCNQ serve as examples (Figure 11.15). However, donor systems have proved to be more important: their structural features usually involve substitution of the sulfur atoms in TTF by selenium and extension of their conjugated system by incorporating multiple bonds and additional electron donor groups. The investigation of tetramethyltetraselenofulvalene (TMTSF, **10**) and bis-(ethylenedithiolene)tetrathiafulvalene (ET, **11**; Figure 11.15) provided particularly valuable results. Dozens of superconductors were

discovered among the ion-radical salts of **10** and **11** with general formula $(D_2)^+ X^-$ (they are called Bechgaard salts.) The counteranion X^- was not necessarily an organic anion-radical of the $TCNQ^-$ type; indeed, enhanced properties were shown by the inorganic anions ClO_4^- or PF_6^- for the TMTSF-based salts and by the linear anions I_3^-, IBr_2^-, AuI_2^- and so on for the ET salts. Thus, the ion-radical salt $(TMTSF)_2^+ \cdot PF_6^-$ becomes a superconductor at 2 K under a pressure of 4 kbar. The salt $(ET)_2^+$ $[Cu(SCN^-)_2]Cl$ holds the record for superconductivity among heterocyclic compounds.[3] The critical temperature (T_c) for transition into the superconductive state is 12.8 K (at 0.3 kbar). Such salts are usually prepared by controlled electrochemical oxidation of the organic donors in the presence of the corresponding inorganic salts.

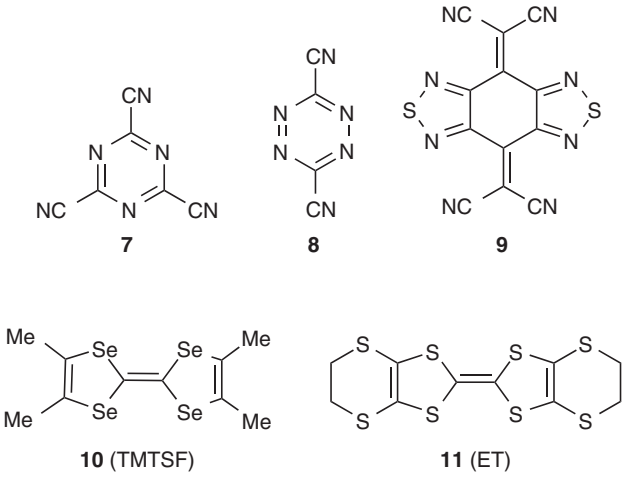


Figure 11.15 *Electron acceptors (**7–9**) and donors (**10** and **11**) used for the preparation of ion-radical 'organic metals'.*

X-Ray analysis has provided much insight into the phenomenon of high conductivity of ion-radical salts. In organic metals the ionic moieties are assembled in a parallel stacking arrangement with the cations and anions constituting different piles (Figure 11.16a). Conductivity occurs through each stack which has a quasimonomeric character, reflecting the nature of crystalline 'organic metals'. A possible mechanism for the conductivity of cation-radical piles is depicted in Figure 11.16b. The electron transfer from donor to acceptor is always incomplete. Even in the salt $TTF^{+\bullet}$ $TCQDM^{-\bullet}$, electrons are only about 60% transferred, meaning that the stacks include cation-radicals together with neutral donor molecules. Such partial filling of molecular orbitals in the stack's units favors electron transfer and therefore provides conductivity, equally true for the Bechgaard salts in which the number of anions is about half that of the π-donor component (Figure 11.16c).

Since metallic conductivity along the chains of organic ion-radicals is provided by π-electrons, the molecular π-orbitals of adjacent particles should partially overlap and thus preferably be parallel to each other. Indeed, both these conditions are met: (i) all effective organic donors and acceptors have planar structures and hence a favorable orientation, (ii) the ionic particles

[3] The highest recorded T_c for an organic superconductor at standard pressure is 33 K observed for the alkali-doped fullerene $RbCs_2C_{60}$. This can be in part attributed to the three-dimensional (metal-like) character of fullerene conductivity, differing from one- or two-dimensional conductivity in ion-radical salts, conjugated polymers and graphite.

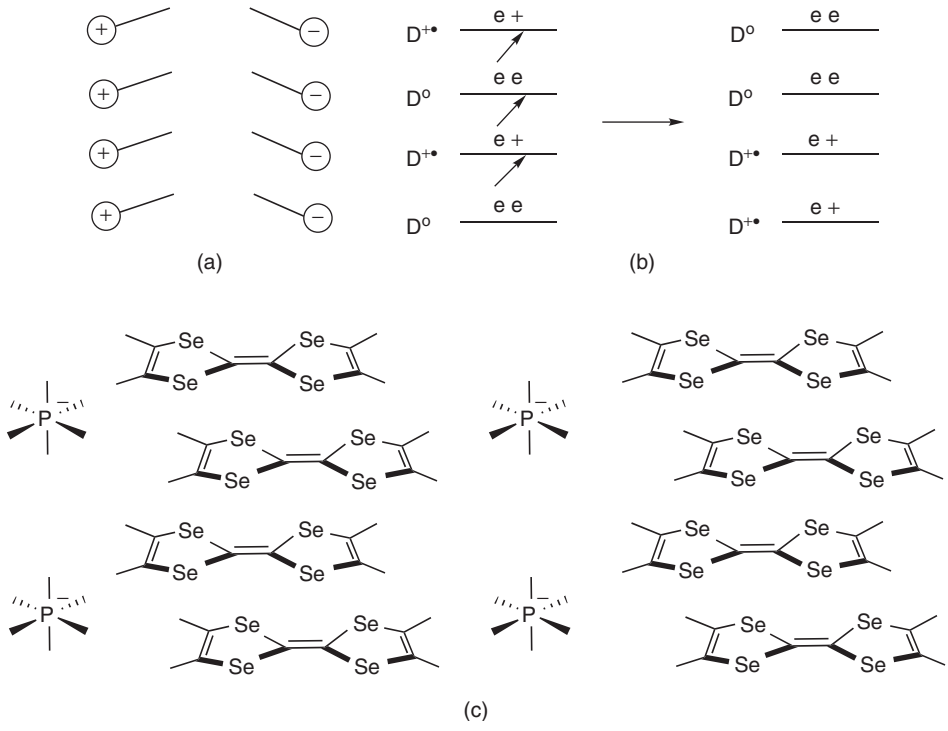


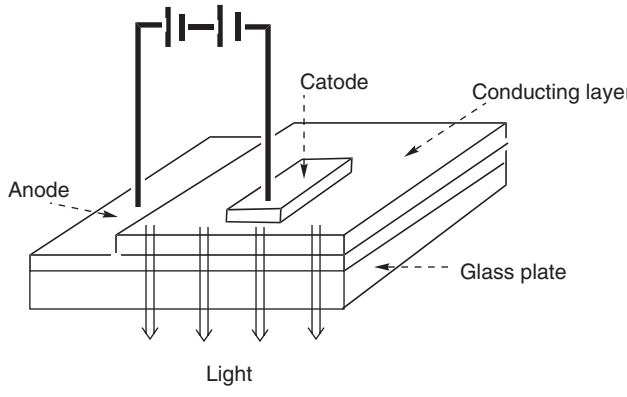
Figure 11.16 *Orientation of ion-radical complexes in the solid 'organic metals': (a) general view for complexes of TTF·TCNQ type, (b) possible mechanism of conductivity in such complexes, (c) arrangement of* $(TMTSF)_2^{+\bullet}$ PF_6^- *from X-ray data.*

in the stacks are separated by distances which are much smaller than the sum of their van der Waals radii.

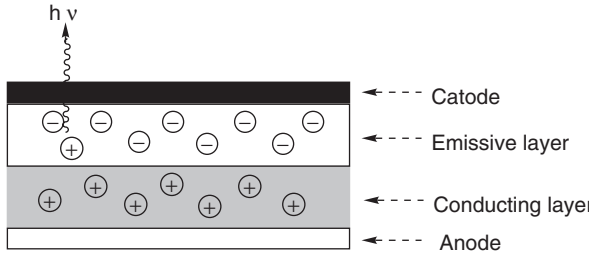
Although existing 'organic metals' cannot presently compete with nonorganic conductors, in particular the recently discovered high temperature superconductors, we can expect further developments with these 'organic metals.'

Since 1990, organic semiconductors have received much attention as electroluminescent materials for organic light-emitting diodes (OLEDs). In the simplest case an OLED is a device in which a film of organic semiconductor is placed between two electrodes where at least one of the sides is transparent (Figure 11.17a). When a voltage is applied across the OLED the conductive layer is divided into two zones. The zone that is closer to the anode becomes positively charged and the one nearer to the cathode gets extra electrons and becomes negatively charged. Under the influence of electrostatic forces electrons and holes in the conductive layer move toward each other and recombine causing emission of radiation in the visible region (most often red, green or blue; Figure 11.17b). As conductive material, organic polymers are commonly used, for example, *para*-polyphenylene or polythiophenes (e.g., **12**; Figure 11.17). OLEDs based on small conducting molecules such as tris(8-hydroxyquinalinato)aluminium **13** (Figure 11.17) are also known.

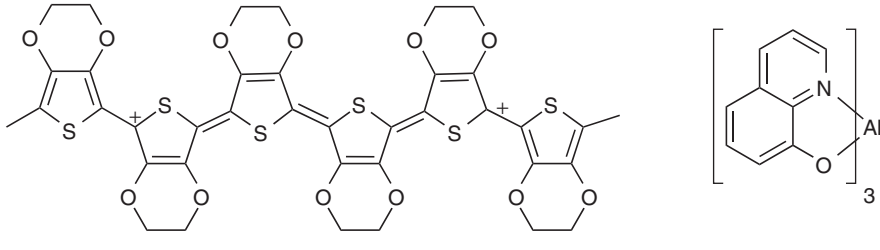
OLEDs represent extremely useful technology. Their commercial application as light sources, in advertising, computer monitors, TV screens and in numerous portable systems grows constantly.



(a)



(b)



12 **13**

(c)

Figure 11.17 *(a) Schematic representation of monolayered OLED. (b) Recombination of electrons and holes arising in a conducting layer. (c) Structures of some heterocyclic semiconductors used in OLEDs.*

11.2 Ecology and Green Chemistry

The burning of vast amounts of fossil fuels, depleting of forests and using ecologically hostile technologies, leading to increasing carbon dioxide concentration in the Earth's atmosphere largely contribute to a progressive worsening of our environment and to global warming. A steady transition to renewable energy sources as well as a search for alternative ways to decrease carbon

dioxide in the atmosphere could improve the ecological situation. Thus, considerable amounts of CO_2 emitted by coal-fired power plants are already captured by various amine reagents or sequestered deeply underground to be trapped by carbonate minerals or water. Especially attractive technologies are those in which carbon dioxide is used as an unlimited feedstock for large-scale production of valuable chemicals and fuel. One such innovation is connected with photocatalytic reduction of CO_2 into carbon monoxide, as discussed above (Figure 11.6). However, presently, more realistic is the catalytic reduction of carbon dioxide into methanol by hydrogen. Several pilot plants based on this approach are already operating in Japan. A composition of metal and metal oxides such as $Cu/ZnO/Al_2O_3$ is commonly employed in this technology as catalyst. The present energy-consuming and slow process demands further search for more efficient catalysts.

Recently, stable heterocyclic carbenes, for example, 1,3-dimesitylimidazole-2-ylidene (Figure 11.18, structure **15b**) have been suggested to substitute for the metal – metal oxide catalysts. Such carbenes are easily obtained on treatment of imidazolium salts **14** with a suitable strong base like sodium hydride. N-Heterocyclic carbenes are commonly represented as resonance hybrid of carbenoid **15b** and ylid **15a** structures. As ylids they are highly nucleophilic and can

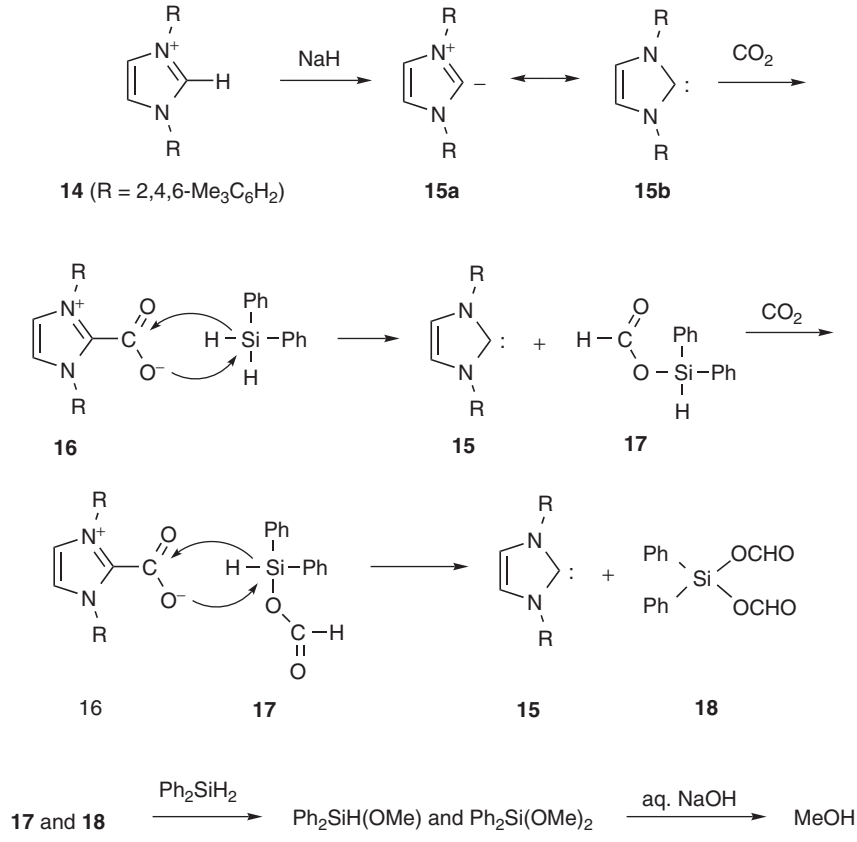


Figure 11.18 *Reduction of carbon dioxide into methanol catalyzed by 1,3-dimesitylimidazole-2-ylidene (**15**, R = 2,4,6-$Me_3C_6H_2$; Riduan, S. N., Zhang, Y. and Ying, J. Y.,* Angew. Chem., Int. Ed.*, 2009, **48**, 3322. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission).*

add a carbonyl compound to activate it for further conversions (see also Sections 4.2.2 and 11.3.1 for more details). Thus, **15** forms with CO_2 imidazolium carboxylate **16**, easily reducible by diphenylsilane at room temperature into formoxy- **17** and diformoxysilanes **18**. A subsequent two-staged reduction and alkaline hydrolysis of **17** and **18** furnishes methanol in over 90% yield.

The chemical industry is a cornerstone of modern civilization, but its wastes are very hazardous for human health and our environment. For example, about 20×10^6 t of volatile organic compounds are emitted annually into the Earth's atmosphere. Therefore, for the past two decades great attention has been directed to so-called 'green chemistry'. This term refers to dozens of principles that should be applied to modern chemical technologies. Some of them are: (i) whenever possible, use and generate substances causing little or no toxicity to human health and the environment, (ii) avoid any auxiliary substances such as solvents, separation agents and so on, (iii) minimize energy requirements, (iv) use renewable starting materials, (v) use catalytic rather than stoichiometric reagents and be as selective as possible. The above can be illustrated by two examples in which heterocyclic compounds play an essential role.

The first example concerns manufacturing polylactic acid (PLA), a biodegradable thermoplastic polyester which has multiple applications, ranging from textiles to food packaging. PLA is made by a three-stage procedure from renewable resources, such as corn starch or cane sugar. First, the plant sugars undergo bacterial fermentation producing lactic acid. The latter cannot be directly polymerized since the water molecule evolved at each esterification step reverses the reaction, thus preventing the formation of a high molecular weight polymer. It is more advantageous to transform catalytically lactic acid into its dimeric lactide which then easily polymerizes in the melt without water production (Figure 11.19a). Although dimerization also generates water, it can be separated by distillation prior to polymerization. Thus, the heterocyclic structure of the lactide provides an advantage just by storing polymerization units in an anhydrous state.

The second example concerns [2+2]-photodimerization of 1,2-bis(4-pyridyl)ethylene (4,4′-bpe) into the corresponding cyclobutane derivative. To dimerize, two olefin fragments should align in a parallel manner and be close to each other (<3.8–4.2 Å). Such orientation allows their π-orbitals to overlap, a key factor for cyclodimerization. However, in the solid state 4,4′-bpe is photostable because it forms a layered structure in which olefins of neighboring layers lie orthogonal and separate by 6.52 Å. In a liquid phase the favorable orientation of molecules is hampered by dilution and diffusion and the process is not selective and effective. The problem has been solved by fixing two 4,4′-bpe molecules in a template matrix with two molecules of resorcinol, shown in Figure 11.19b, structure **19**. In this arrangement, the olefin units are strictly parallel and are separated by 3.65 Å. Unsurprisingly, irradiation of solid **19** leads to the formation of the cyclobutane complex **20** with 100% yield. Such supramolecule-based strategy excludes using solvents required in many other similar reactions. Moreover, here resorcinol acts as catalyst since it can be recycled.

Many recent developments of green chemistry involve ionic liquids (ILs). By definition, ILs are organic salts melting below 100 °C; many even exist as liquids at room temperature. A great majority of ILs are quaternary salts of nitrogen heterocycles: pyrrolidinium, pyridinium, isoquinolinium, imidazolium and so on (Figure 11.20). To provide moisture stability ILs often possess low nucleophilic anions such as BF_4^-, PF_6^-, TfO^- and so on. The low melting point of ILs is attributed to their molecular asymmetry, making difficult for their ions the regular packing in crystals which occurs for conventional inorganic salts like NaCl.

The main ecological benefits of ILs consist in an extremely low vapor pressure and their ability to be readily recycled. ILs also show excellent thermal stability, low combustibility and wide liquid regions. ILs are completely ionized, but show low coordinating ability, which is especially favorable for catalytic reactions. Almost all the properties of ILs can be tuned easily by changing the heterocyclic system, substituents and counterions. Thus, they can be made immisible with

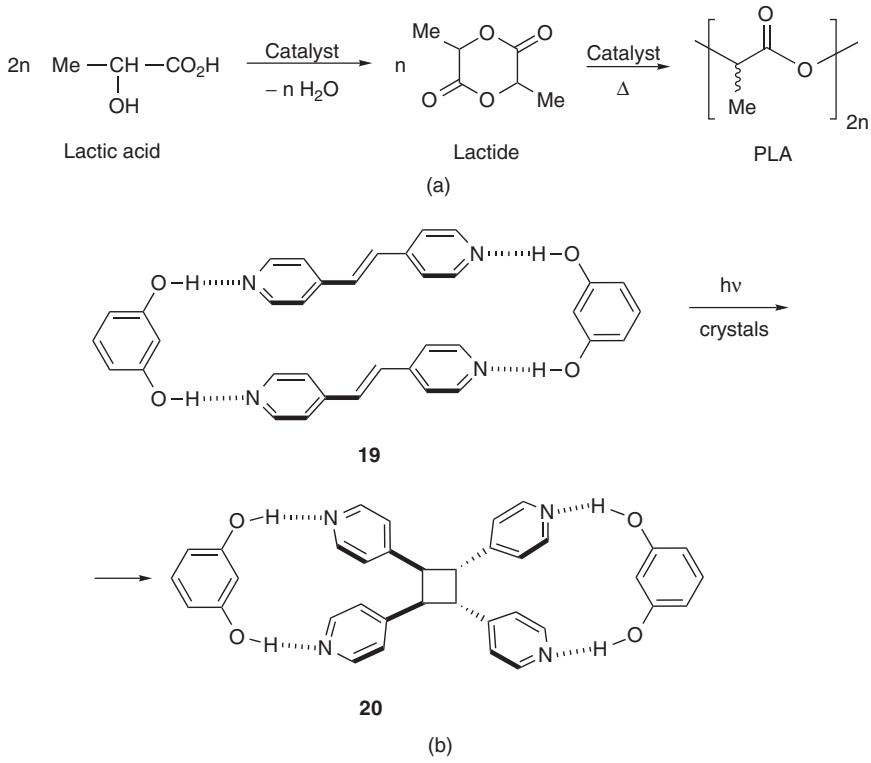
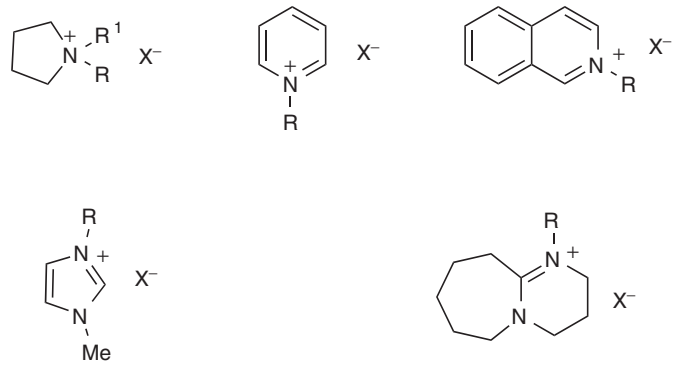


Figure 11.19 *Selected examples of 'green chemistry' reactions, without the use of solvents: (a) preparation of polylactic acid, (b) hydrogen bonding-assisted photodimerization of 1,2-bis(4-pyridyl)ethylene (Reprinted with permission from MacGillivray, L. R., Reid, J. L. and Ripmeester, J. A.,* J. Am. Chem. Soc.*, 2000,* ***122****, 7817. © 2008 American Chemical Society).*

water or organic solvents, providing flexibility for various reactions and separation schemes. What is especially impressive, some ILs can dissolve materials such as cellulose, coal, metals, metal oxides and even rocks.

ILs have found increasing applications as reaction media for various kinds of organic reactions both in the laboratory and in industry. A simple example is the Friedel–Crafts acylation of naphthalene with acetyl chloride. The classical version of the reaction is carried out in nitrobenzene or nitromethane and gives the 2-isomer as the major product and the acylating agent is thought to be an $AcCl–AlCl_3$–nitrobenzene complex. However, when the ionic liquid 1-ethyl-3-methylimidazolium chloride is used as solvent, the ratio of 1- and 2-acetylnaphthalenes is completely reversed to give the thermodynamically unfavored 1-isomer as the dominant product (Figure 11.21). Obviously, the position of attack on naphthalene is determined to a large extent by steric factors. In the ionic liquid, the acylating agent is thought to be the free acylium ion, which is much smaller than the $AcCl–AlCl_3–PhNO_2$ complex, and attacks at the sterically hindered but kinetically more favorable 1-position. Here the ionic liquid displays its strong ionizing ability.

The PetroChina company now employs an aluminium chloride-based ionic liquid in place of sulfuric or hydrofluoric acid catalysts for the alkylation of isobutene, a key refinery intermediate to make gasoline. The annual production of gasoline has reached 65 000 t by this method and this is presently by far the largest commercial process utilizing an ionic liquid.



R = Et, $X^- = SO_4^{2-}$, m.p. = 70 °C

R = Et, $X^- = NO_3^-$, m.p. = 38 °C

R = Et, $X^- = TfO^-$, m.p. = –9 °C

R = Et, $X^- = MeCO_2^-$, m.p. = –45 °C

R = Bu, $X^- = BF_4^-$, m.p. = –80 °C

Figure 11.20 *Some types of heterocyclic quaternary salts used as ionic liquids.*

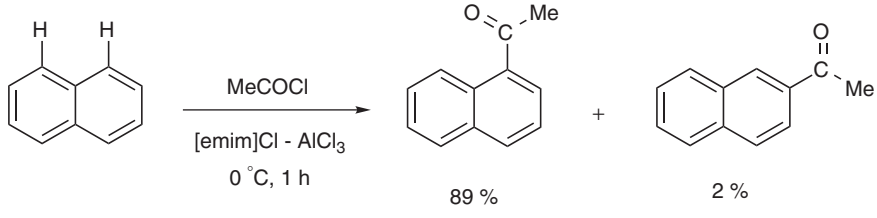


Figure 11.21 *Friedel-Crafts acylation of naphthalene in 1-ethyl-3-methylimidazolium chloride, [emim]Cl, as ionic liquid (Adams, C. J., Earle, M. J., Roberts, G. and Seddon, K. R.,* Chem. Comm.*, 1998,* ***1998****, 2097. Reproduced by permission of The Royal Society of Chemistry).*

Historically, the first major industrial application of ILs was the *b*iphasic *a*cid *s*cavenging utilizing *i*onic *l*iquids (BASIL) process of BASF. It dramatically improved the synthesis of alkoxyphenylphosphines, which serve as precursors to make the photoinitiators used in UV-curable coatings and other applications. The alkoxy phenyl phosphines were previously obtained by the interaction of chloro phenyl phosphines with an appropriate alcohol in the presence of triethylamine as an acid scavenger. However, this had the serious drawback of the formation of insoluble triethylamonium chloride as a waste byproduct that is difficult to handle, requiring still infiltration etc. In the BASIL process triethylamine is replaced by 1-methylimidazole. The imidazole scavenges HCl to form 1-methylimidazolium chloride as an ionic liquid. As a result the formation of two liquid phases occurs, the upper one containing alkoxyphenylphosphine and the lower the liquid imidazolium salt (Figure 11.22). The latter can be easily separated and (after preliminary basification) used again. The process uses a much smaller reactor, speeds up the

reaction and significantly increases the yield. Overall, the productivity of the BASIL process exceeds that of traditional technology by a factor of 80 000.

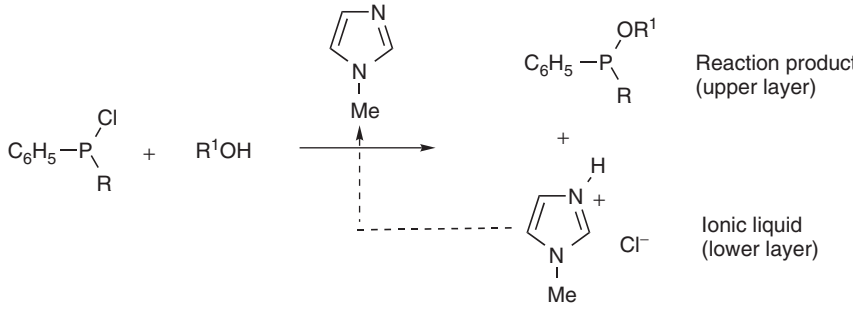


Figure 11.22 *Preparation of alkoxyphenylphosphines via the BASIL process.*

Another IL-based innovative technology is the Difasol process developed by Institut Français du Pétrole, used for the production of low-branched hexenes and octenes by catalytic dimerization of propene and butenes (Figure 11.23a). In the Difasol process a precursor of the nickel catalyst $NiCl_2L_2$ is first dissolved in the ionic liquid [bmim]Cl-$AlCl_3$-$EtAlCl_2$. This system serves both as solvent and Lewis acid, producing an active form of catalyst (Figure 11.23b) and, equally importantly, stabilizes the active nickel particles. 1-Butene is then continuously passed through the ionic liquid, yielding octenes that are separated in the upper phase, immisible with the ionic liquid layer. The catalyst is not dissolved in the octene phase and therefore can be reused together with the ionic liquid. The Difasol process has many advantages: only 1 g of catalyst is needed to dimerise 250 kg of butene, the process proceeds under mild conditions (atmospheric pressure, −15 to −5 °C) and

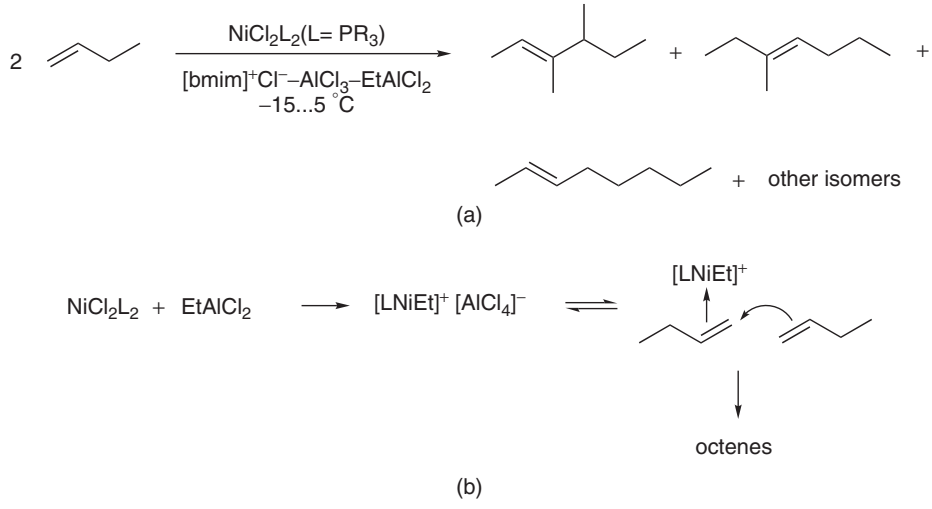


Figure 11.23 *Difasol process for conversion of butene-1 into isomeric octenes: (a) general scheme, (b) suggested mechanism for catalyst activation, [bmim]Cl – 1-butyl-3-methylimidazolium chloride.*

provides higher selectivity and yield of dimers together with a considerable reduction of chemical waste and energy use. C_8 olefins are then converted through hydroformylation-reduction into C_9 alcohols used in production of dialkyl phthalates for polyvinylchloride plasticizers.

Cellulose is the earth's most widespread natural organic chemical and is highly important as a biorenewable resource. Every year, chemical industry uses about 200×10^6 t of cellulose for the production of paper and many other materials. However, more intensive exploitation of cellulose is prevented by the lack of a suitable solvent. Making cellulosic fibers from so-called dissolving pulp currently involves the use, and subsequent disposal, of great volumes of various chemical auxiliaries, such as carbon disulfide. Major volumes of waste water are also produced. Recently, the process was greatly simplified by the use of ILs, for instance, 1-butyl-3-methylimidazolium chloride, which serve as solvents and are recycled.

Ionic liquids have several properties that make them useful in gas storage and handling. Thus, ILs can be used instead of pressurized cylinders as a transport medium for poisonous gases such as trifluoroborane, phosphine or arsine, which are needed for selective doping silicon semiconductors by B, P or As atoms. The gases are dissolved in the ionic liquids at or below atmospheric pressure and easily withdrawn from the containers by applying a vacuum.

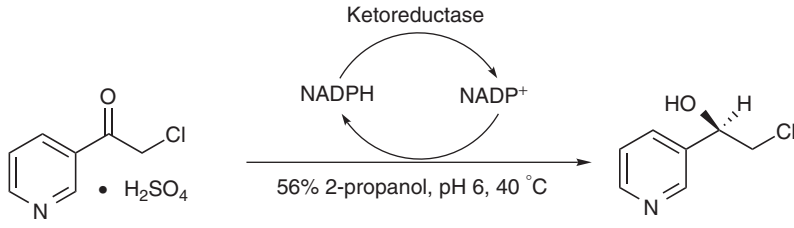
Other uses of ILs include: (i) the production of nanoparticles, (ii) electrolytes in sensors, solar cells and lithium-ion and other types of batteries, (iii) lubrication, (iv) performance additives for paints and coatings, (v) pharmaceutical and drug delivery systems and so on. Ionic liquids will probably find many further commercial applications.

11.3 Biotechnology and Related Problems

Methods now attributed to biotechnology have long been used in everyday life: plant selection, milk and fruit juice fermentation, brewing, isolation of dyes and remedies from natural sources and so on. The term 'biotechnology' first came into common use only in the beginning of the 1980s, when the first genetically modified microorganism was produced and applied for treating oil spills. Recently the decoding of the human and of many other living organisms' genomes has opened new horizons for biotechnology, making it one of the most rapidly developing fields of human activity. Biotechnology is now defined as 'any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use' [*The Convention on Biological Diversity (Article 2. Use of Terms)*, United Nations, 1992]. Biotechnology is applied presently in five major areas: (i) health care and medicine, (ii) agriculture and food production, (iii) industrial production of nonfood materials (e.g., biofuels or biodegradable plastics), (iv) environmental uses and (v) analytical and diagnostic uses. Heterocyclic compounds play a significant role in many of these applications, for instance, through heterocyclic coenzymes, nucleic acid sequences or medicines. Accordingly, we next focus on current developments of enzyme and nucleic acid technologies as well as therapeutics.

11.3.1 Enzyme Technologies

Enzyme technology, being a subfield of biotechnology, deals with the manufacturing of both bulk and high added-value products (foods, fine chemicals, pharmaceuticals, etc.) utilizing enzymes as biocatalysts. Enzymes are also used to improve services such as washing clothes and environmental processes. Two typical examples comprise the industrial production of antibiotics (Section 7.2.3) and of drug intermediates. Thus Merck used the NADP-H-dependent ketoreductase enzyme to prepare **21** (Figure 11.24). Biocatalytic syntheses are appealing because they run under mild conditions with less solvent while giving high enantioselectivity and yield, a process more environmentally friendly and economically attractive.



21

Figure 11.24 *Use of ketoreductase enzyme for industrial synthesis of pyridine alcohol **21**.*

However, as proteins, enzymes bring certain practical disadvantages. Thus, the typical enzyme usually works on just one specific substrate. Even more serious is the occurrence of fast enzyme denaturation on heating, pH changes or under the action of various chemicals, for example, surfactants. Therefore, scientists are interested in the development of so-called artificial enzymes. These are relatively small molecules, possessing catalytic activity similar to natural enzymes but having higher stability and more general applicability to a particular class of reaction, rather than just to a specific substrate.

A significant development in this area since 1990 is the creation of diverse catalytic systems based on nitrogen heterocyclic carbenes (NHC). This originates from the ability of the enzyme thiamine pyrophosphate, which is responsible for the metabolism of α-keto acids. In this process thiamine pyrophosphate generates thiazol-2-ylidene (Figure 4.23, structure **4b**) which instantly undergoes nucleophilic addition to a C=O group and triggers the indicated conversions as exemplified in Figures 4.24 and 4.25. Thiazol-2-ylidenes and other carbenes were long considered as illusive short-lived species. However, in the 1990s the first stable carbenes were desribed: almost all were derivatives of imidazole, 1,2,4-triazole or other nitrogen heterocycles (see, e.g., structures **1** and **15** in Figures 11.2 and 11.18). Their stability is believed to be due to three main factors: (i) the electron-donor effect of pyrrole nitrogen heteroatoms (cf. resonance structures **4a** and **4b** in Figure 4.23), (ii) the protective action of bulky N-substituents (adamantyl, 2,6-diisipropylphenyl or mesityl as in **15** (Figure 11.25)) and (iii) the high aromaticity of azole rings.

As mentioned previously (Section 4.2.2), a key property of NHCs is their high nucleophilicity, very helpful for both classical reactions and unconventional syntheses which are otherwise difficult. Thus, cyclic macromolecules were unavailable for decades, intriguing chemists and material scientists. Recently discovered is the zwitterionic ring-opening polymerization of lactide (L) initiated by 1,3-dimesitylimidazole-2-ylidene **15** and leading to the generation of **22** and other cyclic polylactides with well defined molecular weights between $M = 5000$ and 30 000 g mol^{-1} with narrow polydispersity. The process is extremely rapid; its proposed mechanism is in Figure 11.25.

Stable NHCs are strong bases (pKa $\sim$ 24) that are often used in deprotonation procedures. NHCs are also excellent carbon ligands for practically all metal cations, from alkali and transition metals to lanthanides and actinides. Many such complexes, for example, **23**–**26** (Figure 11.26), as well as the NHCs themselves, are commercially available.

The gold carbene complex **23** in conjunction with $AgSbF_6$ is a highly efficient catalytic system for the hydration of a wide array of alkynes to ketones. The catalyst beneficially replaces toxic mercury(II) salts used in the classical Kucherov reaction and operates under acid-free conditions at very low catalyst loadings (Marion, N., Ramon, R.S., and Nolan, S.P., *J. Am. Chem. Soc*., 2009, **131**, 448):

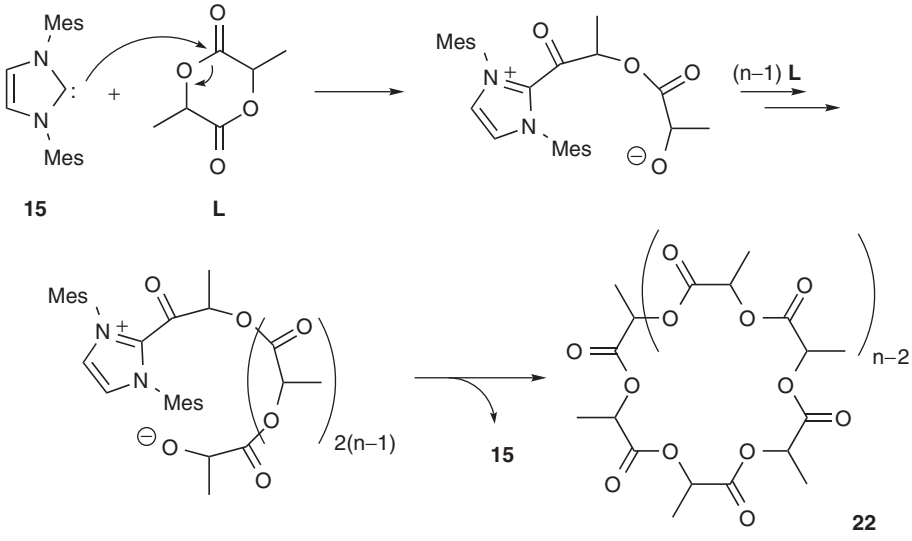


Figure 11.25 *Ring-opening cyclopolymerization of lactide initiated by carbene* ***15*** *(Reprinted with permission from Jeong, W., Shin, E. J., Culkin, D. A., Hedrick, J. L. and Waymouth, R. M.,* J. Am. Chem. Soc.*, 2009,* ***131****, 4884. © 2009 American Chemical Society).*

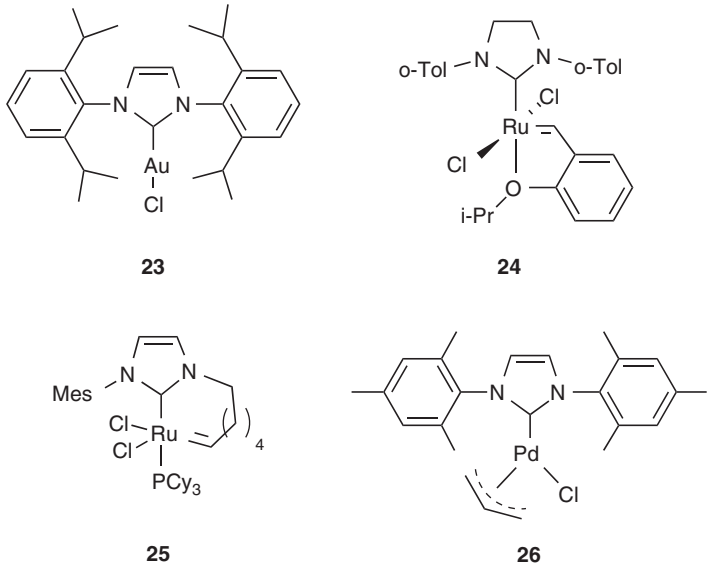
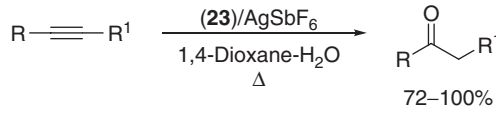
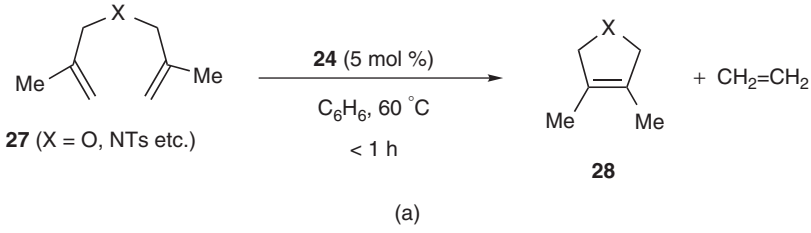


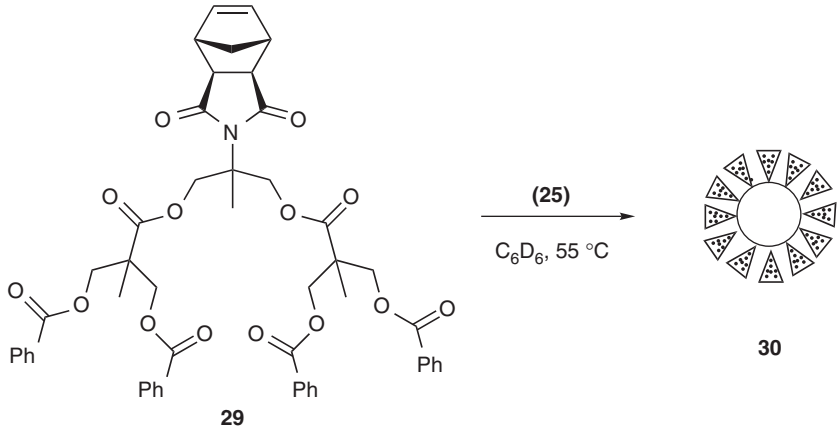
Figure 11.26 *Selected metal complexes of imidazole-based stable carbenes.*



Ruthenium carbene complexes such as **24** and **25** are widely used in petrochemistry, medicinal chemistry and discovery work for olefin metathesis. As the reaction proceeds, C=C bonds are broken (producing metalcarbenes RCH = [M] intermediates) and then re-formed with an exchange of substituents, ring closing, ring-opening, or polymerization. For example, terminal dienes **27** with Ru imidazoline carbene catalyst **24** cyclize into 2,5-dihydroderivatives of furan and pyrroles with >95% conversion; ethylene is other product (Figure 11.27a).



(a)



(b)

Figure 11.27 *(a) Ring-closing metathesis (Reprinted with permission from Stewart, I. C. et al.,* Org. Lett.*, 2007,* ***9****, 1589–1592. © 2007 American Chemical Society). (b) Ring-opening metathesis polymerization (Boydston, A. J., Holcombe, T. W., Unruh, D. A., Frechet, M. J. and Grubbs, R. H.,* J. Am. Chem. Soc.*, 2009,* ***131****, 5388. © 2009 American Chemical Society).*

In the beautiful example of ring-expansion metathesis polymerization shown in Figure 11.27b, a dendronized alkene monomer **29** in the presence of a catalytic amount of **25** forms cyclic dendronized polymer, schematically shown as **30**. AFM imaging has confirmed toroidal features of the macromolecules with diameters of 35–40 nm, heights ranging from 5 to 9 Å, and internal diameters ranging from ca. 5 to 7 nm. Such nanoscale molecular architectures with well defined shapes and dimensions should find application in drug delivery and nanotechnology.

Palladium carbene complexes of type **26** are widely employed in various cross-coupling reactions permitting C-C bond formation (Suzuki-Miyaura, Heck, Sonogashira and some others; see Marion, N. and Nolan, S. P., *Acc. Chem. Res*., 2008, **41**, 1440).

Being effective carbon ligands, NHCs can stabilize unstable, sometimes intriguing, electron-deficient molecules, for example, diborene, HB=BH, or bis-silylene, Si_2 (Figure 11.28), a fundamental contribution into the chemistry of the main-group elements. Previously, no one had been able to get two boron atoms to form a double bond in a stable, neutral molecule. By analogy, the naked Si_2 molecule without ligands could only be examined spectroscopically in the gas phase or in an argon matrix at very low temperature.

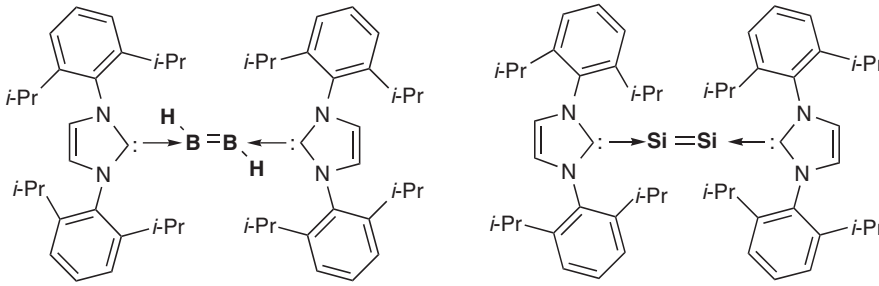
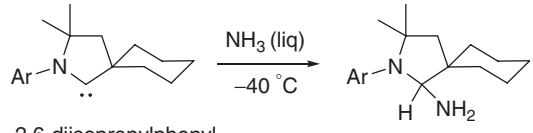


Figure 11.28 *Carbene stabilization of B_2H_2 and Si_2 molecules (Reprinted with permission from Wang, Y., Quillian, B., Wei, P., Wannere, C. S., Xie, Y., King, R. B., Schaefer, III, H. F., Schleyer von, P. R. and Robinson, G. H.,* J. Am. Chem. Soc., *2007,* ***129****, 12412. © 2007 American Chemical Society; Wang, Y., Xie, Y., Wei, P., King, R.B., Schaefer, III, H. F., Schleyer von, P. R. and Robinson, G.H.,* Science, *2008,* ***321****, 1069).*

Hydrogen gas and ammonia are among the most important raw chemicals in modern industry. Traditionally, transition-metal complexes are used to chemically activate or cleave the H_2 molecule and very few examples of NH_3 scission are known. However, carbenes, in which the carbene center is sandwiched between an amino group and an alkyl group, have enough nucleophilicity to cleave H_2 and NH_3 molecules. The resulting fragments (H and/or NH_2) become attached to the carbene center. With liquid ammonia, the reaction occurs in high yield at around −40 °C (Frey, G. D., Lavallo, V., Donnadieu, B., Schoeller, W. W. and Bertrand, G., *Science*, 2007, **316**, 439):



Ar = 2,6-diisopropylphenyl

This observation should find many applications if a method is developed to use the carbene as a catalyst transferring H and NH_2 species to appropriate substrates.

Currently, there is considerable interest in the enzyme methane monooxygenase (MMO). Some bacteria (methanotrophs) produce energy for their metabolism with its help, consuming methane (or other alkanes) by oxidizing it into methanol. The process demands the participation of NAD(P)-H and ultimately leads to the reductive scission of O_2 and the insertion of an oxygen atom into a C—H bond (Figure 11.29a). MMO has a complex structure, not yet clarified. Commonly, two modifications are distinguished: the soluble and the so-called particulate form. The soluble form

has a nonheme diiron active site. In the reduced state each Fe^{II} ion is surrounded by six ligands, one of which is a histidine residue (Figure 11.29b).

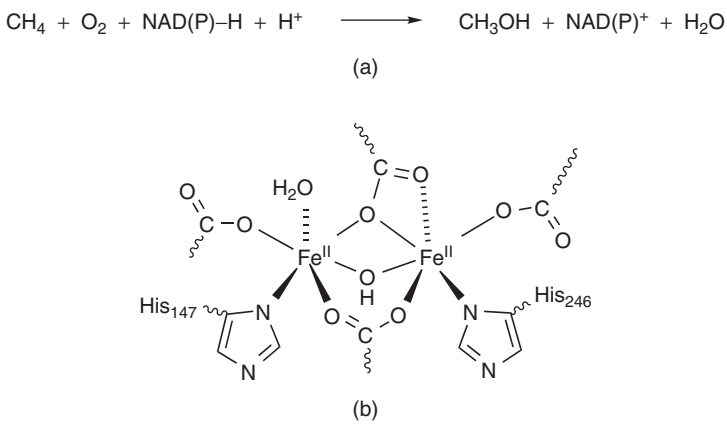


Figure 11.29 *Enzyme methane monooxygenase activity: (a) summary equation of MMO-catalyzed oxidation of methane into methanol, (b) suggested structure of MMO's active site.*

Because of the high energy of the C—H bond in CH_4 (99 kcal mol^{-1}), a large activation barrier needs to be overcome which can be significantly lowered by MMO, giving the process practical significance. Thus, modelling MMO could be important for more effective use of natural gas, for environmental purposes (e.g., treating oil spills) as well as for better understanding methane (a green house gas) recycling in nature.

11.3.2 DNA Technologies

Soon after the discovery of the double helix structure, scientists realized that the unique self-organization, recognition and information properties of DNA molecules can be applied for diverse purposes, far beyond biochemistry. Thus, the development of DNA technologies began as a specific branch of biotechnology. DNA technologies have already revolutionized fields as varied as forensic sciences, history and anthropology, informatics and medicinal therapy. Below we summarize some of these selective applications of DNA technologies.

Polymerase Chain Reaction

The elucidation of the DNA replication mechanism induced scientists to develop a simple and universal biosynthetic path for amplification of a great variety of small (~2 kilobase pairs on average) DNA fragments, including artificial polynucleotide pools. This technique, based on using DNA polymerase, is named the polymerase chain reaction (PCR) and has now become one of the most powerful and fruitful among the DNA technologies. It allows amplification of even minute quantities of DNA. In 1993 American scientist K.B. Mullis who invented the PCR method was awarded the Nobel Prize in Chemistry.

The rapid preparation of millions of copies of a polynucleotide by PCR relies on thermal cycling, consisting of repeated heating and cooling of the mixture of special components. The latter include: (i) a double-strand DNA sample to be amplified, (ii) two short synthetic oligonucleotides called primers, (iii) four deoxynucleotide triphosphates (dATP, dGTP, dCTP, dTTP), (iv) a thermostable

DNA polymerase, (v) metal ions (Mg^{2+}, K^{+}) and (vi) buffer solution. Each cycle includes three basic stages (Figure 11.30). In the first stage, called denaturation or 'DNA melting', a DNA duplex unwinds and separates into two single strands on heating for 1–5 min at 95 °C. On subsequent cooling to 54 °C, the two primers anneal to the DNA chain specifically at their complementary sequences. In this artificial assembly (chimeric DNAs), one primer is complementary to one target site of the DNA molecule, while the other primer is complementary to the opposite strand of the DNA at another distant site. The sequences of the two primers consisting of 18–25 nucleotides are designed so they are oriented with their 3′-hydroxyl ends towards each other in the chimeric DNAs. At the third stage, the extension of polynucleotide chains is carried out for 1–2 min at 72 °C, and the primers are elongated into polynucleotides by incorporation of the deoxynucleotides present in the initial mixture. This polymerization is catalyzed by a thermostable DNA polymerase enzyme.[4] The enzymatic synthesis is initiated at the 3′-terminals of the primers and the lengthening of both new polynucleotide chains proceeds towards each other along the uncoiling DNA template. The whole process is then repeated over and over during subsequent cycles. In the second (and succeeding) cycle, two types of template participate: (i) the original DNA strands and (ii) the newly formed DNA strands. However, after a few additional cycles, newly synthesized DNA fragments quickly predominate. Since at each elongation step, the amount of DNA target is doubled and since the DNA polymerase joins a thousand bases per minute, the whole PCR procedure demands not more than 30–40 cycles and takes a few tens of minutes.

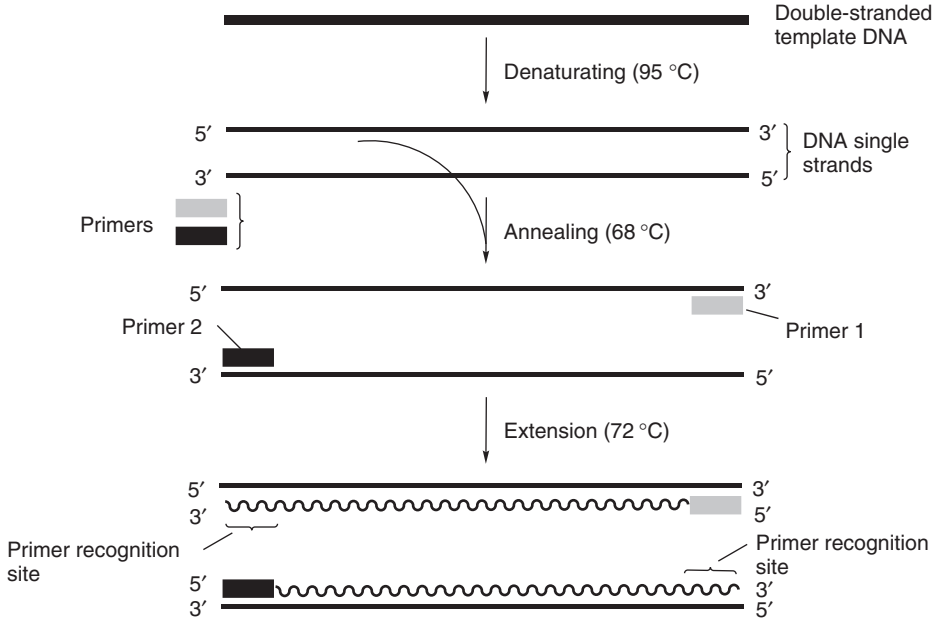


Figure 11.30 *The first cycle of a polymerase chain reaction (Copyright: IPGRI and Cornell University, 2003. Retrieved with small modifications from: www.bioversityinternational.org.; file name: MolMarkers Vol1 III AFPLs.pdf).*

[4] Since heating is required for the separation of the DNA-strand, DNA polymerases used in PCR should withstand relatively high temperatures. Fortunately, such enzymes are produced by certain hot-spring bacteria and can be isolated.

Typically amplification of only a specific region of the DNA is the aim. Both ends of such a DNA target should be restricted by primers. However, in the first cycle each single DNA strand adds only one primer. Because of this two new polynucleotide chains, much longer than needed, are formed in the first cycle. In the second cycle, their shortening is ensured by the annealing of the second primer to its complementary site on the DNA that already contains the first primer. The short chains thus obtained in the second and the third cycles are used as the templates for creating millions of copies of the desired length.

The PCR method has many applications and variations. For example, it is widely used for the selective amplification of a specific region of genome DNA that allows functional analysis of genes, estimation of gene expression, diagnosis of hereditary and infectious diseases and so on. Since the PCR technique in a short time can supply an investigator with a large amount of pure DNA, it has found major application in forensic analysis under the name DNA profiling or genetic fingerprinting.

DNA Profiling

As already mentioned, genomic DNA can be divided roughly into coding and noncoding areas. The coding area determines protein production and is very similar for different individuals in the sense of the nucleobase sequences. Noncoding (so-called junk) DNA has large sections very specific for each living individual. This specificity can be qualified by the number of tandem repeats in such sections. The terms 'tandem repeats' or 'short tandem repeats' (STR) refer to a sequence of several nucleotides (usually between 2 and 20) which is repeated with the repeats adjacent to each other. The following tandem repeat in which the sequence ATCG is repeated four times serves as a typical example:

A – T – C – G – A – T – C – G – A – T – C – G – A – T – C – G

Laboratory analysis of the number and length of such sequences (in fact short DNA fragments; generally fewer than 400 bases) provide a DNA profile for an individual. At the beginning of the analysis, these DNA fragments are separated from the entire DNA strand and amplified. This is done by means of the PCR technique using primers corresponding to the ends of each repeating sequence. Next, the length of each fragment is determined by gel electrophoresis – a technique in which molecules are separated by the difference in their net charge in the presence of an externally applied electric field. To provide a good estimate, DNA fragments with known lengths are run alongside the test speciments and the distances migrated are compared. To visualize an experimental picture, the chromatogram is treated with a special fluorescent dye and then is photographed under UV light. Quaternary heteroaromatic salts such as ethidium bromide (Figure 11.31a) or SYBR Green I (Figure 11.31b) are the dyes most often used.

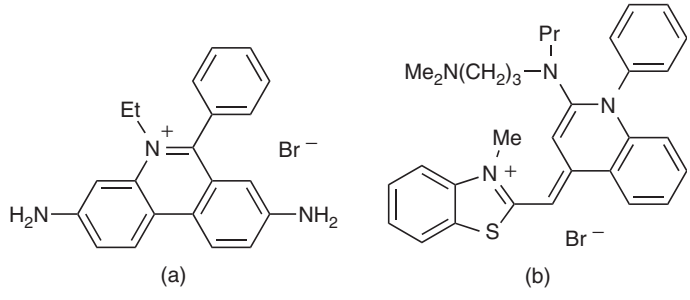


Figure 11.31 *Fluorescent dyes for staining nucleic acids in gel electrophoresis.*

Generating a DNA profile usually involves analysis of an individual's DNA for ten different STRs on different chromosomes. Statistically, no two people (except identical twins) are likely to have the same number of repeats in all of these STRs. Figure 11.32 illustrates this point. The child has inherited only five of his/her repeats from each parent, thus each child has his/her own unique DNA profile.

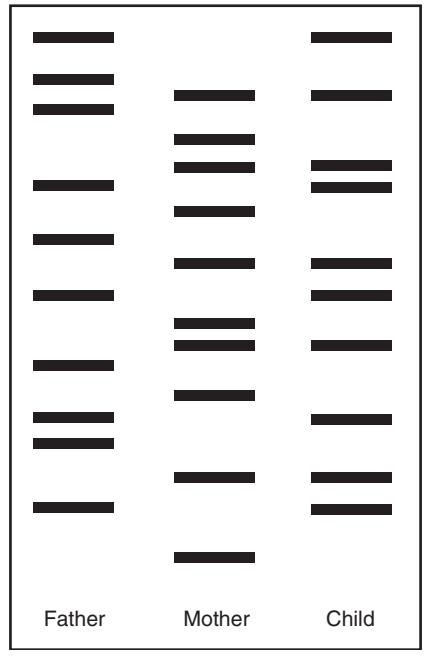


Figure 11.32 *DNA profiles of two parents and their child (each band corresponds to STR of definite length).*

Today DNA fingerprinting has become a highly reliable method of conclusive personal identification. Several large DNA database exist in the United States, England and other countries. They are very helpful in criminal investigations and clarification of genetic family relationships.

DNA Nanotechnologies

Molecules of single-strand DNA, just as for any polynucleotide, are sticky molecules since they spontaneously bind other similar molecules or even their own through classical Watson–Crick pairing. American crystallographer N.C. Seeman was the first who made use of this property for the controllable construction of nanoscale structures of different shapes from pieces of DNA. Not only single-strand but also two-strand DNAs can be employed for such architecture. In the latter case a double helix has a single-stranded overhang (often called a 'sticky end') to bind another overhang DNA with antisense sequence as shown in Figure 11.33a. In this way, four single-strand DNAs with strictly programmed sequences can form a 4-arm junction (Figure 11.33b) as well as many other two- and three-dimensional constructions: faces, stars, road maps, polyhedrons and so on.

Recently, DNA nanotechnology has expanded considerably due to the invention of programmable folding of long, single-stranded DNA into various arbitrary shapes. This technique developed by American scientist P.W.K. Rothemund (*Nature*, 2006, **440**, 297) is now known

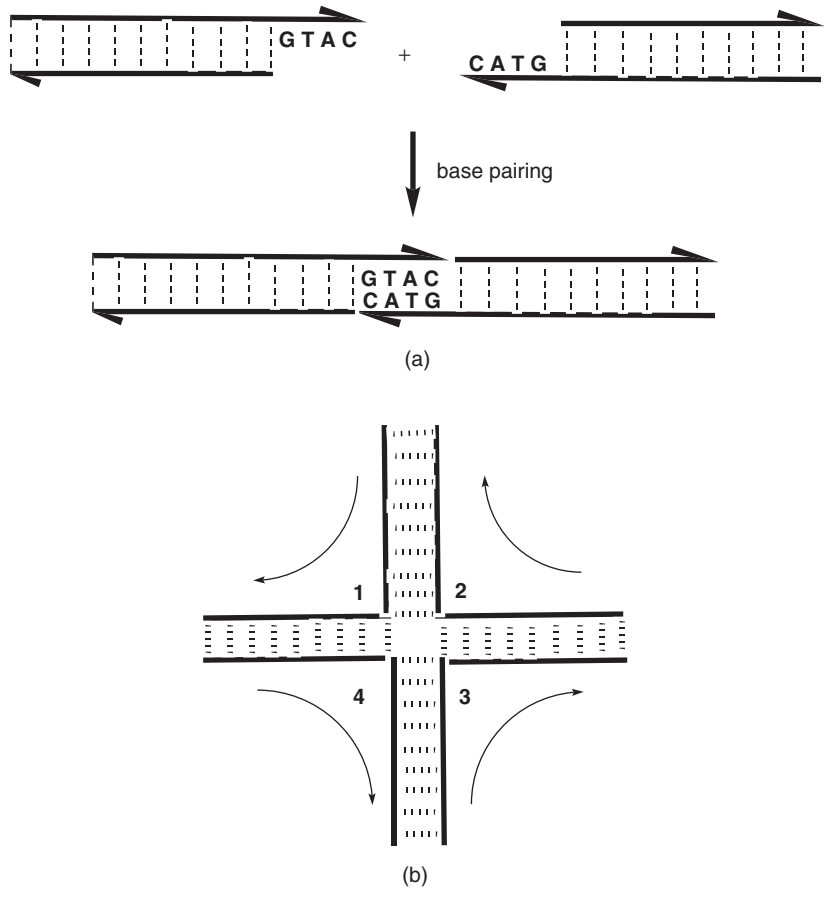


Figure 11.33 *Assembly of artificial DNA molecules using DNA nanotechnology: (a) association of two sticky-ended DNA molecules, (b) schematic structure of a 4-arm junction (**1–4** are four separate DNA single strands; the curved arrows indicate 5′ → 3′ direction of the backbones).*

as 'DNA origami'. One of its most impressive achievements is the building of a nanoscale DNA box (Figure 11.34) fabricated from the circular, single-stranded DNA genome of the M13 bacteriophage. In the first stage the DNA was heated with 220 staple strands – short polynucleotides with specially designed nucleobase sequences. The latter annealed to target sites of DNA providing its pre-programmed folding. This stage resulted in the formation of six interconnected DNA sheets arranged in two parallel rows with three sheets in each. To connect the sheets into a cubic structure in the second stage they were assembled with 59 further staple strands. The DNA box thus obtained was 42 nm high 36 nm wide and 36 nm deep. The box was equipped with a dual lock and key system (two short strands of DNA) to hold the lid opened or closed. To detect the opening process two fluorescent dyes C3 and C5 were inserted in the lid and the upper edge of the front sheet. Efficient lighting occurred only in closed state due to the close proximity of the dyes and the FRET effect (fluorescent resonance energy transfer). Dyes Cy3 and C5 belong to the indoline family of cyanines:

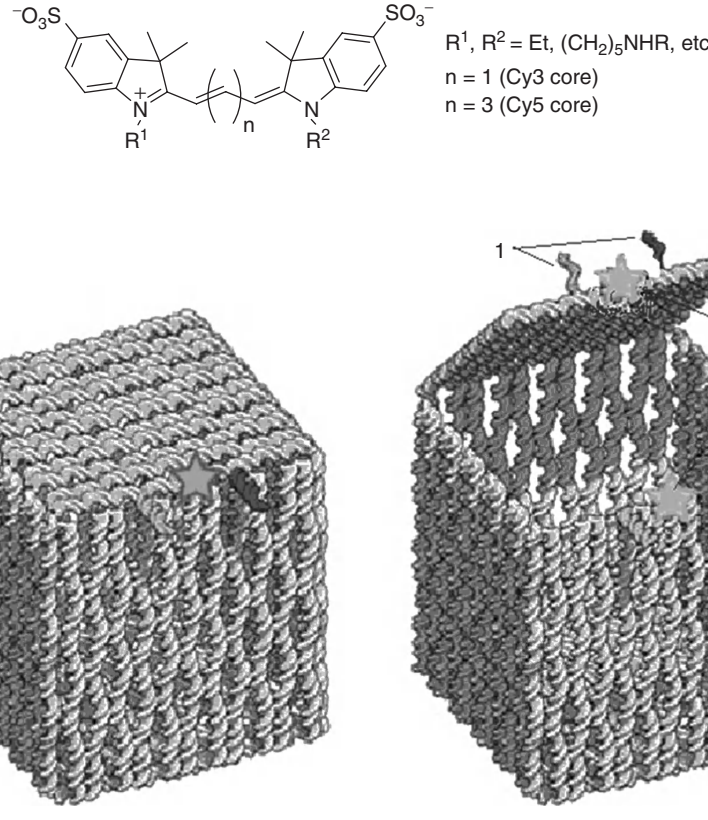


Figure 11.34 *A nanoscale box built by DNA origami:* 1 *– short polynucleotides that fulfill a lock/key function,* 2 *– fluorescent dyes signaling on opening-closing the box (Reprinted by permission from Macmillan Publishers Ltd: Andersen, E. S., Dong, M., Nielsen, M. M. et al.,* Nature*, 2009,* ***459****, 73 © 2009).*

Nanoscale DNA structures, including DNA nanocontainers, can find various applications, for example as drug delivery systems or logic sensors for multiple-sequence signals. Thus, DNA computing, which uses DNA instead of the classical silicon-based computer technologies, is currently progressing rapidly. Due to the information-encoding and recognition capability of DNA together with the enzymatic machinery available for DNA manipulation, DNA molecules can serve as input, output and software. DNA computers will be faster and smaller than any other computers built so far and will demand extremely low power consumption.

11.3.3 New Trends in Health Care

Decoding the human genome raised great hopes for a rapid development of gene therapy. Gene therapy is very promising for the treatment of both inherited and acquired diseases including cancer, diabetes, and cardiovascular and nervous system disturbances. Moreover, in the not so distant future, gene therapy shows promise for transformation into personalized medicine taking into account the specific genetic profile of patients. Commonly, the concept of gene therapy

involves the transfer of genetic material into a cell or tissue with the goal of curing a desease or even its prevention. However, due to the discovery of multiple mechanisms for the regulation of gene expression, the term 'gene therapy' now also includes therapeutic intervention with short RNAs and other small molecules to influence gene expression.

The strategy for development of gene therapy involves the elaboration of methods of genetic testing, gene engineering, transfer of genetic material into a cell with subsequent regulation of its activity and selectivity. Among other tasks, we mention a search for so-called biomarkers – cell compounds accompaning each genetic illness. Apart from significant achievements in gene diagnostics, no serious breakthrough in gene therapy has yet been achieved. One of the most challenging problems is the delivery of genetic material into a cell. Difficulties arise particularly from the large size of polynucleotides and their high negative charge; for example, siRNA molecules on average represent ~40 charged polyanions. Such a molecule cannot penetrate the negatively charged cell membrane. To overcome these difficulties, multiple delivery systems (vectors) are being intensively tested. Lipid- and polymer-based nanoparticles are being developed for the delivery of RNAi. The first genetically targeted therapeutic may be created in the next five years; most likely, they will involve small interfering RNA technologies, siRNA.

Apart from gene therapy, several other innovative approaches to drug medicine have been suggested, including the so-called molecular logic gate concept, by which a molecule must perform some kind of logic computation before it can help defective cells. A prototypical example of such a drug is BODIPY dye **33** (Figure 11.35, cf. Figure 9.15b). Compound **33** recognizes two characteristics inherent to cancer cells: high concentrations of H^+ (pH ~4) and Na^+ (up to three times that in normal tissues). When both conditions are satisfied, which in computing terminology is referred to as 'AND' logic, the molecule, in the presence of light, generates singlet oxygen and other reactive oxygen species that destroy the infected cells. This approach can be considered as an advanced modification of photodynamic therapy (Section 7.4.3). Many other molecular logic gates based on heterocyclic systems have since been suggested (Szacilowski, K., *Chem. Rev*., 2008, **108**, 3481).

As explained above (Section 9.4), light induces in some molecular systems (photoswitches) reversible structural changes which are used in materials science technologies. It is now evident that similar photoswitches can be applied in medicine. Thus, bis(pyridinium)dithienylethene **34** exists in two forms, depending on the wavelength of the irradiation light (Figure 11.35). Ultraviolet light switches it from a colorless, ring-open form **34a** to a blue, ring-closed isomer **34b**, and visible light causes the 1,3-cyclohexadiene ring to reopen. The molecule even maintains its photoswitching ability inside living organisms such as the worm *Caenorhabditis elegans*, in which it can be cycled multiple times. When *C. elegans* worms are fed with the ring-open form **34a**, they behave normally. However, if such worms are exposed to 365 nm UV light, **34a** is converted into the ring-closed form **34b** and paralyzes the worms. Shining visible light longer than 490 nm on the worms reverses the paralysis, which is likely caused by the photoisomer **34b** disrupting a metabolic electronic pathway involved in energy production.

11.3.4 Heterocycles as Molecular Sensors

Broadly, a sensor can be defined as any device allowing detection of a distinct kind of matter or energy. Thus, radars, lasers, sonars and infrared optics, which are used for the observation of distant or hidden objects, can be considered as sensors. Quite often, people borrow ideas for the construction of such sensor tools from living nature. Indeed, Nature generously endowed animals with multiple biodevices helping them in navigation, communication, searching for food, discerning danger and so on. Until recently, scientists engineered sensors mainly for the detection of bulk items. But a large majority of biosensors act at a molecular level. They can distinguish the concentrations of different molecules that have key significance for the regulation of numerous

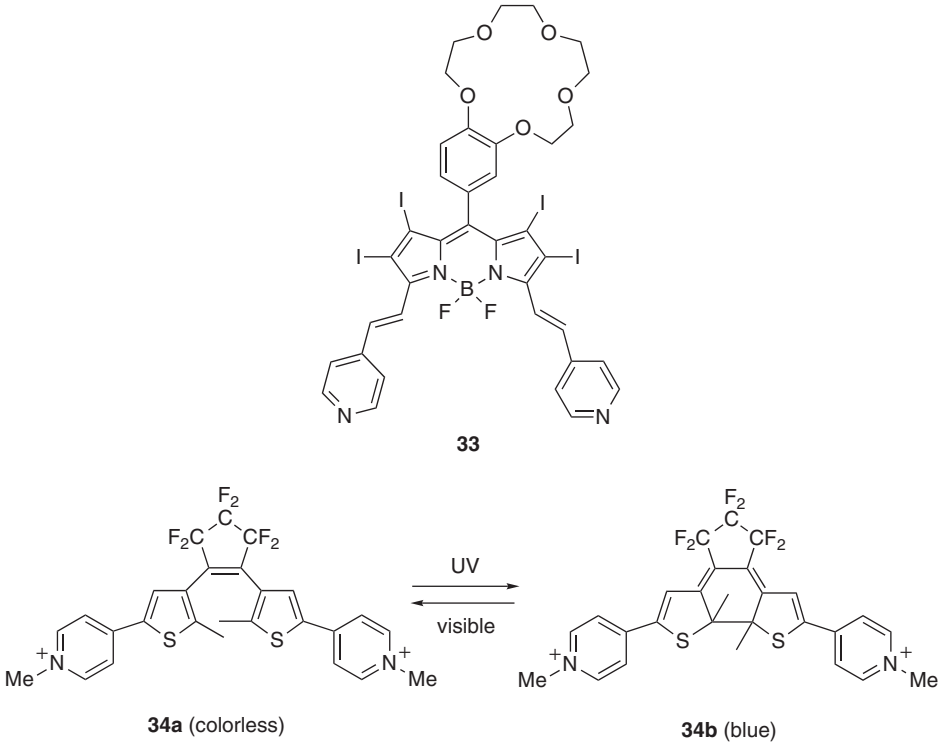


Figure 11.35 *Examples of heterocyclic derivatives switching their biological activity on chemical contact (structure **33**: Reprinted with permission from Ozlem, S. and Akkaya, U.,* J. Am. Chem. Soc.*, 2009, **131**, 48. © 2009 American Chemical Society) or light input (structure **34**: Reprinted with permission from Al-Atar, U., Fernandes., R., Johnsen, B., Baillie, D. and Branda, N. R.,* J. Am. Chem. Soc.*, 2009, **131**, 15966. © 2009 American Chemical Society).*

biochemical processes. A striking example is butterfly antenna which detect in air tiny amounts of pheromones (Section 8.3) or a dog nose detecting a single odor among many dozens of others. Therefore, the creation of an 'electron nose' and an 'electron tongue', as well as molecular sensors in general, is a major aim of modern science. Molecular sensors are also in high demand for environmental monitoring, detecting explosive and poisonous compounds in transport baggage, protecting money and other valuable documents and so on. However, the most important applications of molecular sensors lie in the biomedical sciences, especially in medicinal diagnostics.

Commonly, an artificial molecular sensor is a specially designed individual organic compound or a combination of several components differing by high selectivity and sensitivity. In many cases the action of the molecular sensor is based on molecular recognition and is reversible. Such a host–guest interaction with an analyte is reflected in the recently introduced term 'supramolecular analytical chemistry'. An important process in the use of molecular sensors, especially in live tissues, is the detection of one singular interaction with a target molecule. For this purpose, fluorescent methods and compounds which undergo color changes are often employed (see Sections 9.3.5 and 9.4). Many examples of molecular sensors, including a damage sensor (Figure 9.25), crown-ethers, receptors for anions and neutral molecules were provided in Section 10.1. Some other recent developments are given below, emphasizing the penetration of nanotechnologies into this field.

Every moment hundreds of different compounds are involved in a cell's biochemical reactions. Monitoring these processes provides timely disclosure of any misbalances and allows tracing a course for their correction. Imaging agents are especially effective for such diagnostics (see Section 9.3). However, their employment is difficult in many cases. A typical example of such an analyte is hydrogen peroxide. H_2O_2 is a normal metabolite of many oxidation reactions but its overproduction is usually observed on development of numerous inflammatory diseases. Imaging hydrogen peroxide *in vivo* is hampered by its low concentration and low reactivity in comparison with other reactive oxygen species ($O_2^{\bullet -}$, $OH^{\bullet}$, NO). Recently, the so-called peroxalate nanoparticles were developed for this purpose. They are formulated from hydrophobic polymer **35** that contains peroxalate esters in its backbone and a fluorescent dye (e.g., pentacene **37**). The peroxalate nanoparticles image H_2O_2 through a two-step process (Figure 11.36). First, hydrogen peroxide diffuses into the nanoparticles and reacts with the oxalate ester groups, generating dioxetanedione **36** – a high-energy heterocyclic compound. The dioxetanedione then quickly decomposes, producing carbon dioxide and photons of light that excite chemically the fluorescent dye **37**. A key advantage of using peroxalate nanoparticles is their ability to generate chemiluminescence at high emission wavelengths (>600 nm). This is ideally suitable for deep-tissue imaging (>1 cm) owing to minimal absorption by hemoglobin, water and lipids at these wavelengths. The method is tunable (by changing the fluorescent dye), highly sensitive (detects H_2O_2 at concentrations as low as 250 nM) and selective. Other reactive oxygen species, such as superoxide or nitric oxide, are unable to generate dioxetanedione under the same conditions.

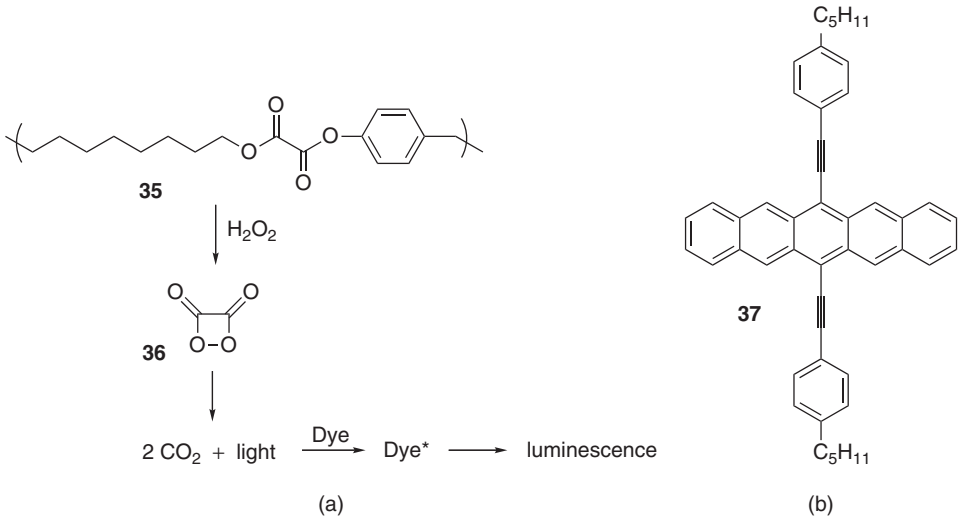


Figure 11.36 *A principle for imaging hydrogen peroxide* in vivo*: (a) hydrogen peroxide oxidation of peroxalate ester and generation of oxetanedione and light quanta, (b) pentacene derivative as fluorescent dye (Adapted by permission from Macmillan Publishers Ltd: Lee, D., Khaja, S., Velasquez-Castano, J.C., et al.,* Nat. Mater.*, 2007,* ***6****, 765, © 2007).*

The second example of a nanoscale molecular sensor deals with the routine detection of mercuric ion (Hg^{2+}) in water reservoirs, a great problem since mercury is a widespread and dangerous environmental pollutant. Numerous current methods for Hg^{2+} detection suffer from limited sensitivity and selectivity and the sensors used are kinetically unstable and incompatible with aqueous

environments. A new technique is based on using DNA-functionalized gold nanoparticles (DNA-Au NPs) that are about 100 nm in size and form a stable colloid of an intense red color in water or another fluid. Gold nanoparticles have a large affinity for sulfur atoms that are widely used for their functionalization. For example, the DNA-Au NPs are obtained by treatment of gold nanoparticles with HS-functionalized DNA. DNA-Au NPs have very high extinction coefficients (three to five orders of magnitude higher than those of common organic dyes). They are easily hybridized to complementary particles which exhibit extremely sharp melting transitions (for DNA melting, see Section 11.3.2). Due to their unique properties, DNA-Au NPs have been suggested as sensors for proteins, oligonucleotides, metal ions and various small molecules. Their use for detecting Hg^{2+} ions relies on the following principles (Figure 11.37).

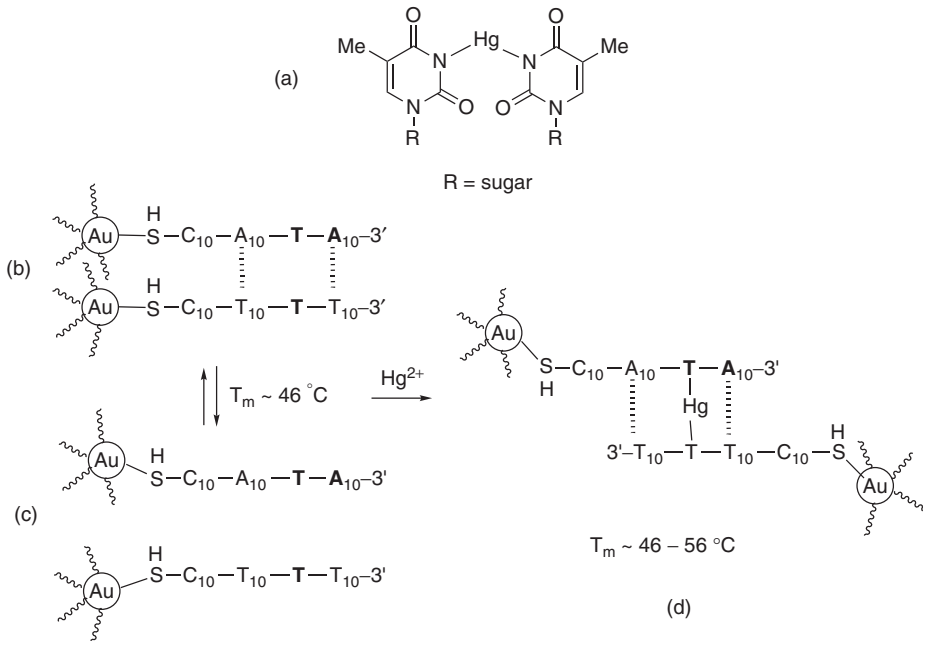
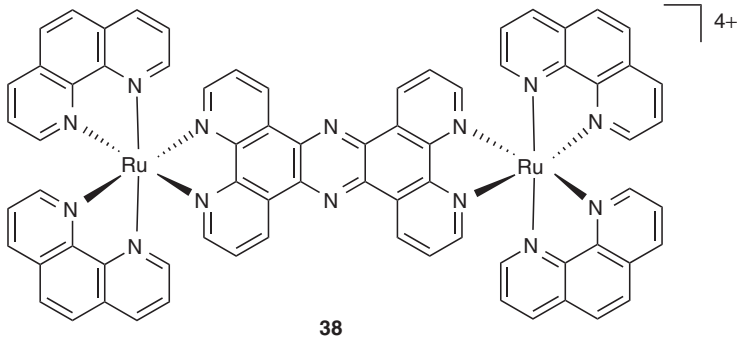


Figure 11.37 *Basic elements of colorimetric detection of mercuric ion with DNA-functionalized gold nanoparticles: (a) thymine-Hg^{2+}-thymine base pair, (b) binding of two DNA-AuNPs via complementary base pairing (in the absence of Hg^{2+} ions), (c) melting of the aggregates, (d) binding of the nanoparticles in the presence of Hg^{2+} ions (wave lines designate polynucleotide chains identical with the ones shown). Adapted from Lee, J.-S., Han, M. S. and Mirkin, C. A.,* Angew. Chem., Int. Ed.*, 2007, **46**, 4093. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.*

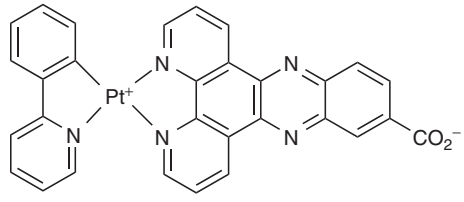
First, the well known ability of thymine derivatives to form mercuric salts with the participation of its acidic N3-H bond (Figure 11.37a) is used. Second, two types of DNA-Au NPs are made, each functionalized with different thiolated-DNA sequences: 5′-HS-C_{10}-A_{10}-**T**-A_{10}-3′ (probe A) and 5′-HS-C_{10}-T_{10}-**T**-T_{10}-3′ (probe B). These sequences are complementary except for a single thymidine-thymidine mismatch (shown in bold as **T**). When mixed, both types of particles form stable aggregates due to A-T base pairing (Figure 11.37b). Without Hg^{2+} the aggregates melt with dramatic purple to red color change at about 46 °C thereby converting into free particles (Figure 11.37c). In the presence of Hg^{2+}, however, the aggregates melt at temperatures higher than 46 °C because of the coordination of Hg^{2+} to the two thymidines that make up the T-T

mismatch (Figure 11.37d). To measure precisely the T_m value, the solution is slowly heated while its extinction is monitored at 525 nm, where the Au NP probes exhibit the maximum intensity. The concentration of Hg^{2+} is obtained from its linear dependence on the melting temperature. The present limit of detection is approximately 20 ppb Hg^{2+}, which is the lowest ever reported for a colorimetric Hg^{2+} sensing system. In addition, the chelating ability of the thymidines that form the mismatch in the oligonucleotide duplex is extremely selective for Hg^{2+}.

Imaging agents help to visualize distinct types of molecules in living cells and to detect their concentration. In principle, they are also able to provide structural information that is especially important for such complex compounds as proteins and DNA. However, many currently available fluorescent dyes suffer from low water solubility, photobleaching, high toxicity and low membrane permeability. In addition, they require ultraviolet light for illumination, which can induce extensive DNA photodamage. In this context, a new generation of DNA imaging agents has recently appeared. This includes complexes of chelating azaheterocycles (2,2′-dipyridyls, o-phenantroline, etc.) with such metal ions as Ru^{2+}, Pt^{2+}, Rh^{3+}. Two representatives, **38** and **39**, are depicted in Figure 11.38. They almost completely lack the above shortcomings and because of easy metal to ligand charge transfer excitation are activated in the safer visible region. The mechanism of action of such DNA sensors is due to their flat geometry that is based on their intercalation effect (see Sections 3.5 and 7.4.3). For example, **39** binds specifically and very effectively in this way to telomere G-quadruplexes. This binding is accompanied by about a 290-fold increase in the intensity of photoluminescence (λ_{max} 513 nm) that allows the visualization of special conformations of DNA such as G-quadruplexes. Interestingly, such interactions that also stabilize the telomere quadruplex structure are one of the most important strategies in preventing cancer deseases.



38



39

Figure 11.38 *Selected metal complexes for direct imaging of DNA structures in living cells: structure **38** (Reprinted by permission from Macmillan Publishers Ltd: Gill, M. R., Garcia-Lara, J., Foster, S. J., Smythe, C., Battaglia, G. and Thomas, J. A.,* Nat. Chem.*, 2009, **1**, 662 © 2009), structure **39** (Reprinted with permission from Ma, D-L., Che, C-M. and Yan, S-C.,* J. Am. Chem. Soc.*, 2009, **131**, 1837. © 2009 American Chemical Society).*

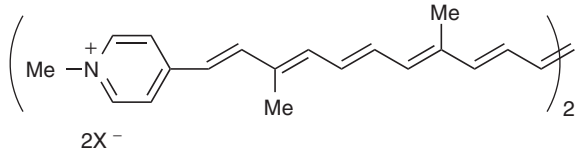
11.4 From Molecular Devices to Molecular Computer

Throughout human history people have invented countless instruments and devices helping in their work, movement, communication or providing better living conditions. Tools, scissors, sewing needles, yachts, bicycles, airplanes, telephones and computers are just a few examples. These are all commonly designated as macroscopic devices. An innovation of the present is that a new generation of macroscopic devices such as the scanning tunneling microscope or a modern X-ray diffractometer allows us to see the world at the molecular level. This brings many advantages, for example, in the creation of high quality functional materials and, what is especially significant, the design of molecules capable to perform useful tasks which cannot be accomplished by common macrodevices: such include targeted drug delivery, cleaning of blood vessels, corrections of genetic distortions, cell diagnostics and so on. Scientists hope to achieve these and many other goals by means of artificial molecular machines – tiny synthetic devices (nanodevices) made of specially constructed single molecules or a combination of such molecules.

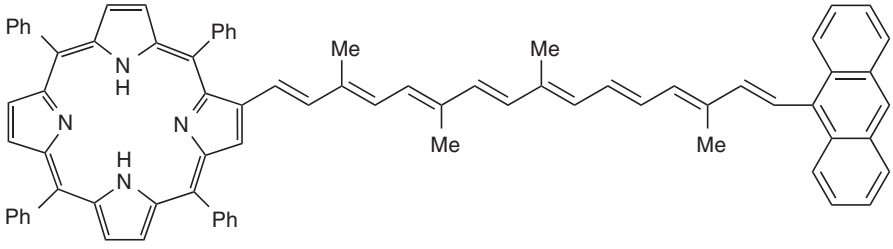
In fact, numerous molecular machines have functioned in living organisms for millions of years: ribosomes, enzymes, ionophores, proton pumping systems, motor proteins and so on. Though we are still very far from competing with Nature in such operations, many successful steps in this direction have been already done. Most are connected with microdevices for the so-called molecular electronics: such include molecular switches, motors, shuttles, tweezers, wires and so on. This completely new area of science is strongly spurred by the development of supramolecular chemistry and is rapidly merging with nanotechnology.

Unique structural and electronic properties together with diverse reactivity (Chapter 2) make heterocycles excellent materials for molecular devices. This book has already provided examples of molecular switches (Figures 9.24, 9.25, 10.23, 11.35), shuttles (Figure 10.3), tweezers (Figure 10.12), devices for artificial photosynthesis (Figures 11.4, 11.5) and drug delivery (Figure 11.34). Some more sophisticated examples are given below.

The idea of creating so-called molecular wires, the conductors necessary for molecular electronics, is of great interest (Section 11.1.4). Bispyridinium salt **40** (Figure 11.39), a vinyl analogue of methylviologen (Figure 8.3), can be regarded as a prototype. In this salt the conjugated polyene



40



41

Figure 11.39 *Molecular wires to convey electric current (**40**) and light energy (**41**).*

chain plays the role of a wire and the heterocyclic cations constitute the specialized molecular contacts. A potential difference appearing between the two ends of the chain causes one of the pyridine rings to be reduced and electrons to enter this molecular system and migrate to the opposite end. Thus, a current arises in the molecular 'wire'. When placed across a cellular membrane, such wires can function as electron channels.

This idea has been extended to the development of molecular light transformers. One such compound **41** is also a polyene connected to an anthracene molecule at one end and a tetraphenylporphyrin system at the other. Light of wavelength 256 nm first excites the anthracene unit. The energy absorbed is carried along the conjugated chain to the tetraphenylporphyrin ring, resulting in dramatically intense carmine red (656 nm) light emission. Similar systems can be used to transmit various types of signal, separating charges, locating biological objects by luminescence and so on.

A more complex artificial nanodevice capable of performing signal processing in solution is shown in Figure 11.40. The molecule in Figure 11.40 mimics the function played by a macroscopic electron extension cable. The complex heterocyclic system comprises three main components: (i) tris(2,2′-bipyridine)ruthenium unit **A**, which functions as a light-powered electron source, (ii) 1,1′-dioctyl-4,4′-bipyridinium dication **E**, which is an electron sink and (iii) bridging molecule **C**, which provides an extension cable. In CH_2Cl_2 solution all three units (**A, E, C**) can reversibly self-assemble by means of two distinct plug/socket junctions. The sockets are represented by two crown-ether moieties, dibenzo[24]crown-8 **B** attached to the **A** unit, and benzonaphtho[36]crown-10 **D** connected to the end of the conducting cable **C**. The other end of the cable **C** contains a dialkylammonium group which plays a role of a plug entering the socket **B** due to hydrogen bonding with ether oxygens. In its turn, 4,4′-bipyridinium dication **E** plugs into the socket **D** by π-electron donor–acceptor bonding. When the whole system is self-assembled, electrons generated by visible photoexcitation of the unit **A** are transmitted to the component **E** on a nanosecond timescale.

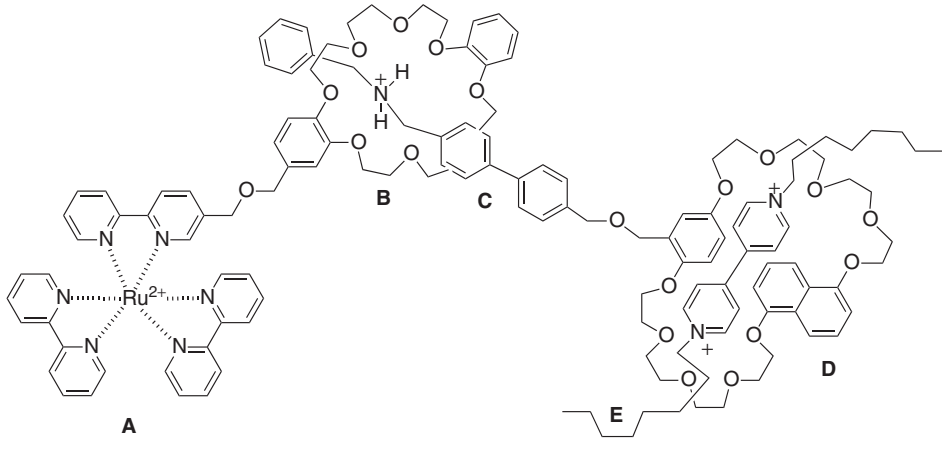


Figure 11.40 *Self-assembling supramolecular extension cable:* ***A*** *electron source component,* ***B*** *hydrogen-bonding socket,* ***C*** *dialkylammonium center as plug 1,* ***D*** *benzonaphtho[36]crown-10 as π-electron-rich socket 2,* ***E*** *1,1′-dioctyl-4,4′-bipyridinium dication as plug 2 (adapted from Ferrer, B., Rogez, G., Credi, A., Ballardini, R., Gandolfi, M. T., Balzani, V., Liu, Yi, Tseng, H.-R. and Stoddart, F.,* Proc. Natl Acad. Sci. USA*, 2006,* ***103****, 18411. Copyright 2006 National Academy of Sciences, U.S.A).*

A highly challenging goal for scientists and engineers is the development of a molecular computer. Many future projections on this subject include a great deal of science fiction. The challenge

arises because of constant efforts to reduce computer size while greatly increasing power and efficiency. Moore's law describes the long-term trend in the efficiency of computing hardware: the number of transistors that can be placed on an integrated circuit has doubled approximately every 18 months. Today's personal computers are already several times more powerful than the supercomputers used around 1970 to place men on the moon. However, this increase in efficiency is expected to stop around 2015, for the reasons described below.

A computer is a programmable machine that receives input, stores and manipulates information and provides output in a useful format. A computer has four main components: the arithmetic logic unit (ALU), the control unit, the memory and the input and output (I/O) devices. These parts are connected by groups of wires called busses. Inside each of these parts are 10^3–10^{12} small electrical circuits which can be turned 'off' or 'on' by means of electronic switches. The circuits are arranged in logic gates so that one or more of the circuits may control the state of one or more of the other circuits. The control unit, ALU, registers, and basic I/O are typically constructed on a single integrated circuit called a central processing unit (CPU) commonly known as a microprocessor – which comprises the heart of a computer. For our discussion *computer memory* is highly important.

Colloquially, computer memory refers to the physical devices used to store data or programs. There are two main types of memory: temporary or volatile and permanent or nonvolatile. The first operates as random access memory (RAM) and requires power to maintain the stored information; when this power is switched off the information is lost. Volatile memory makes use of integrated circuits consisting of silicon-based transistors. By contrast, nonvolatile memory, which operates more slowly, retains the stored information even when not powered. Nonvolatile memory is read-only memory (ROM) commonly based on the use of magnetic materials. Hard disks and floppy discs are the devices commonly used for the storage of ROM.

In computing, the basic unit of information is a *bit* (a contraction from the words *binary digit*). A bit is the amount of information that can be stored by any physical system switching between two stable states, to represent 0 and 1. Such a system is called bistable. Bits can be implemented in many forms by almost any on/off switch: a silicon transistor, an electrical capacitor that can store or lose a charge, a magnet with its polarity up or down, a surface that can have a pit or not. The two values can also be interpreted as logical values (*true/false, yes/no*), algebraic signs (+/−), activation states (*on/off*) and so on.

Computers represent information in binary code, written as sequences of 0s and 1s. A binary string of eight digits (bits), for example, can represent any of $2^8 = 256$ possible values and can therefore correspond to a variety of different symbols, letters or instructions. Along with eight-bit code, octal, decimal or hexadecimal notation has been also developed. In the eight-bit ASCII code, a lowercase 'h' is represented by the bit string 01101000 and the word 'heterocycle' is written as follows.

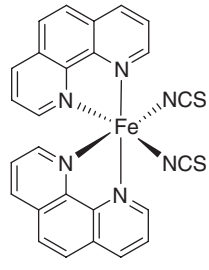
h	e	t	e	r	o	c	y	c	l	e
01101000	01100101	01110100	01100101	01110010	01101111	01100011	01111001	01100011	01101100	01100101

Thus, to designate a single 11-letter word demands 88 transistors. Not surprisingly, modern microprocessors contain more than 3×10^9 transistors and further increases in computer memory are urgently required. So far, this problem has being alleviated by continuous miniaturization of silicon integrated microcircuits. However, further progress in this direction is now approaching technical and fundamental limits including those caused by electron tunneling, difficulties arising at lithographic fabrication of nanoscaled (<50 nm) SiO_2 layer. Among various proposals to overcome this obstacle an idea of using molecular switching units looks especially promising.

Ideally, each such molecule could store one bit of information that together with a small molecular size (1–100 nm) should provide very high informational density and computational capability. Molecular structures possess a wide range of optical, electric, magnetic and mechanical properties; they can be tuned via synthetic procedures; phenomena of self-assembly and molecular recognition can be employed for creation of such molecular devices. In recent years, a series of different molecular switching units have been tested based on change in color, spin states, redox states, conformation and shape, which can be influenced by various external stimuli.

The creation in 2007 of a 160 000-bit molecular electronic memory circuit based on bistable [2]rotaxane was a significant achievement in this area (Figure 11.41). The molecule in Figure 11.41 consists of long polyether pivot, threaded through bis-4,4′-bipyridinium macrocycle (see also Figure 10.20, structure **61**). In the normal state, the macrocycle resides on a tetratiofulvalene station that corresponds to 0. When a positive volatage (+1.5 V) is applied, the TTF fragment is oxidized into TTF^{1+} or TTF^{2+} and a fast translation of the macrocycle from the TTF^{+} site to dioxynaphthalene site occurs that corresponds to '1'. Scientists had engineered the circuit with a cross-bar architecture that consisted of 400 Si bottom nanowire electrodes crossed by 400 Ti top nanowire electrodes. A monolayer of the [2]rotoxanes was sandwiched between these electrodes. Due to the presence at the ends of the rotaxane of bulky hydrophobic and hydrophilic stoppers each molecule was oriented perpendicular to the electrodes: with the hydrophobic stoppers closer to the Si nanowires and the hydrophilic stoppers closer to Ti nanowires. Each bit corresponds to an individual junction defined by a Si bottom and Ti top nanowire forming in total $400 \times 400 = 160\,000$ bits. Appling voltage to one horizontal and one vertical wire one can switch [2]rotaxane molecule between two states thus writing or reading one bit of information. Molecular electronic memory density reached in this work is 10^{11} bits cm^{-2} that is roughly analogous to the dimensions of RAM circuit projected to be available by 2020.

The development of molecular switching units for hard disk memory (ROM) is also in progress. Currently, computer hard disks store data by defining the magnetic anisotropy orientation of a spinning disk with chrome oxide carrier. So far, steady increase of memory density was achieved by shrinking the size of the effective magnetic storage regions (domains). However, this trend will soon reach the so-called superparamagnetic limit at which ambient heat can cause de-orientation of magnetic domains. To escape this problem molecular spin-transition compounds have been suggested. Their typical and perhaps the most successful example is the iron(II) complex with 1,10-phenanthroline system **42**. The application of this compound is based on the ability of Fe^{2+} ion in an octahedral ligand field to switch its $3d^6$ valence shell between paramagnetic (S = 2) high-spin and diamagnetic (S = 0) low-spin states. This transition can be triggered by such external stimuli as temperature, pressure or electromagnetic radiation.



42

Scientists have long compared a computer with the human brain, which also receives and stores information, processes and provides output. Moreover, even single-celled animals and plants

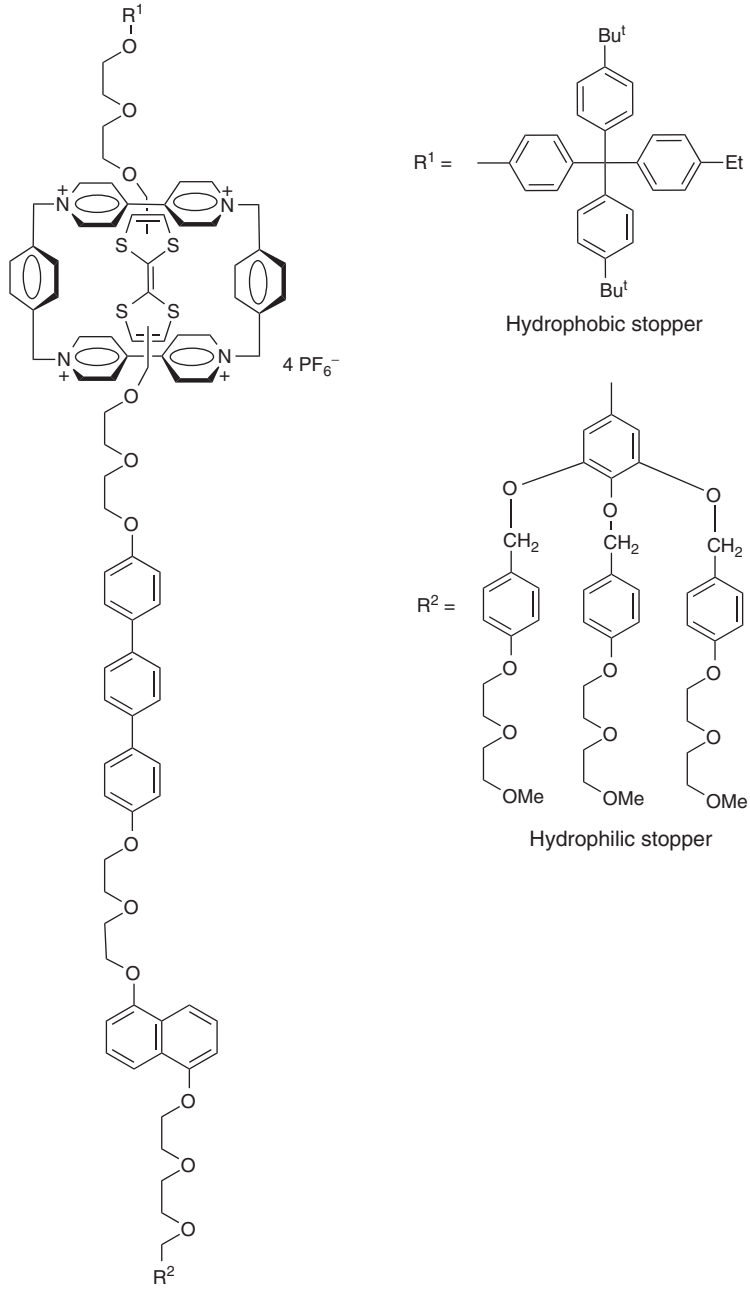


Figure 11.41 *Bistable [2]rotoxane molecule used in molecular electronic memory circuit; the non-oxidized, low-conductance form is shown which corresponds to the '0' state (Reprinted by permission from Macmillan Publishers Ltd: Green, J. E., Choi, J. W., Boukai, A., Bunimovich, Y., Johnston-Halperin, E., DeIonno, E., Luo, Y., Sheriff, B. A., Xu, K., Shin, Y. S., Tseng, H.-R., Stoddart, J. F. and Heath, J. R.,* Nature*, 2007,* ***445****, 414. © 2007).*

constantly switch ‘off’ and ‘on’ metabolic processes by enzyme stimulation. These and many other observations had initiated a development of a biological type of molecular computer. In biocomputers, biologically derived materials perform computational functions. One can distinguish three types of biocomputers: biochemical computers, biomechanical computers and bioelectronic computers.

Biochemical computers achieve computational functionality by using various biochemical reactions as input, and provide chemical output in forms such as changing concentration of reactants or reaction products, the presence of the particular product and so on. Biomechanical computers rely on the nature of biomolecules to adopt certain structural configurations under certain chemical conditions. In this case, the mechanical, three-dimensional structure of the product is detected and interpreted appropriately as a calculated output. Bioelectronic computers comprise specifically designed biomolecules that conduct electricity in highly specific manners based upon the initial conditions that serve as the input of the bioelectronic system.

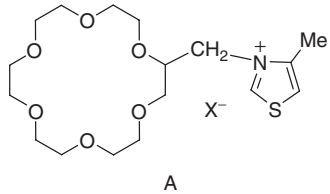
At the present time, the most advanced biomachines for computing are based on DNA molecules including specially engineered DNA nucleotides. Such biocomputers can perform both logic and mathematical calculations. There are many forms of binary signals used in DNA computing: an ability of DNA strands to be opened or closed in response to external input, DNA restriction-ligation, encoding defined types of mRNA and the synthesis of special proteins whose concentration is then registered and so on. Commonly, the corresponding enzymes serve as hardware in such computing biodevices. Recently, biocomputing systems were coupled with standard silicon based chips for the first time: an event marking the integration of biological and electro-mechanical systems at the subcellular level.

Developing the technology of biocomputers is a popular, rapidly growing research field, likely to see rapid future progress. Scientists believe that a DNA computer coupled with an input/output module will be capable of diagnosing cancerous activity within a cell and then releasing an anticancer drug upon diagnosis. The creation of a successfully working molecular computer will open a door for the use of various types of microscopic nanomachines able to build and repair things at a fundamental level. It will mean the appearance of an era of new technology that will change everything from medicine to space exploration.

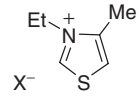
The rapid development of science and technology does not allow us to cover all of the current trends concerning heterocyclic chemistry in one chapter. The authors are aware of this and encourage readers to follow new achievements in this field.

11.5 Problems

1. Suggest practically useful products which can be prepared from 5-hydroxymethylfurfural (Figure 11.1).
2. The surface of a rotated liquid is parabolic. Using this observation, I. Newton in 1670 suggested a reflecting telescope with a mercury-based parabolic mirror. Recently, the use of an ionic liquid has been suggested instead of mercury in such a telescope. When placed on the moon it could work in the infrared region and, with a diameter of 100 m, it could allow a much better view of the early Universe than modern instruments. Since ionic liquids have no reflecting ability, their surface in such a telescope should be covered with silver nanoparticles. What specific properties must be characteristic for an ionic liquid used in the moon telescope? What prevents placing a mercury-based telescope on the moon? (Hint: the temperature of moon's surface can decrease to $-147\,^{\circ}C$.)
3. Due to the very low Lewis acidity of its counter-anion, the ionic liquid [bmim][Cu_2Cl_3] can effectively and reversibly absorb and store gases including such poisons as PH_3 and AsH_3. As a result, a nontoxic complex compound is formed with a molar ratio of components IL : PH_3 = 1:2. How many liters of phosphine gas can be stored in 10 l of [bmim][Cu_2Cl_3], if the density of this ionic liquid is 1.5 g cm^{-3}?
4. A crown containing the thiazolium salt A in the presence of K^+ or Na^+ (but not Li^+) enhances the rate of pyruvic acid decarboxylation (see Figure 4.24) by one order of magnitude compared with the 3-ethyl-4-methylthiazolium cation B. Suggest an explanation.

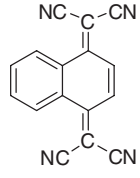


A

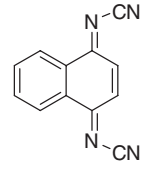


B

5. Suggest a mechanism for the interaction of the fluorescent dyes shown in Figure 11.31 with DNA double strands. Explain why these dyes also stain single-stranded DNA and RNA but to a lower degree.
6. Suggest a mechanism for the influence of H^+ and Na^+ concentrations on the ability of compound **33** to generate singlet oxygen. Note that **33** absorbs light in MeCN solution at ~625 nm and this band moves to ~655 nm in acidic medium.
7. *N*,*N*′-Dicyano-1,4-naphthoquinodiimine (D) forms a 1:1 charge transfer complex with tetrathiafulvalene which possesses high electrical conductivity. However, tetracyano-1, 4-naphthoquinodimethane (C) does not form a complex with TTF. Explain.

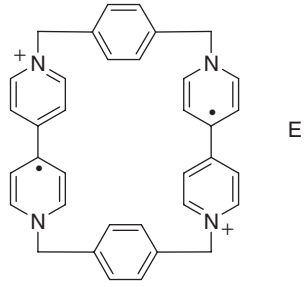


C

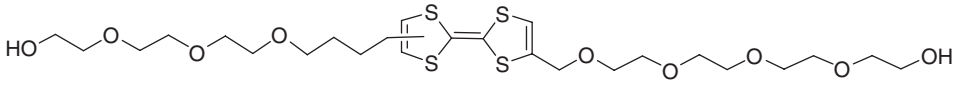


D

8. The diradical dication E cannot be oxidized to the corresponding tetracation in MeCN solution at – 0.33 V (vs SCE). However, when E is threaded by electron-donor wire, F, it is oxidized under the same conditions. Explain.



E



F

9. Photosensitizers used in DSCs and hybrid biofuel cells usually have carboxylic acid groups in their molecules (see Figures 11.8b and 11.9b). Give a reasonable explanation.
10. Show the three main stages in the biosynthesis of a polynucleotide pool from a DNA duplex by the PCR method. What weak interactions and chemical reactions occur during this process? (First analyze Figure 3.10 and the text of Section 3.3).
11. The supramolecule depicted in Figure 11.40 can be switched on/off by two distinct, reversible assembling/disassembling processes that are governed by chemical inputs. Suggest molecular systems that could form these inputs.

11.6 Suggested Reading

1. Rifkin, J., *The Hydrogen Economy*, Tarcher/Putnam, New York, 2002.
2. Olah, G. A., Goeppert, A. and Prakash, G. K. Chemical recycling of carbon dioxide to methanol and dimethyl ether: from greenhouse gas to renewable, environmentally carbon neutral fuels and synthetic hydrocarbons, *J. Org. Chem*., 2009, **74**, 487.
3. Gratzel, M. Solar energy conversion by dye-sensitized photovoltaic cells, *Inorg. Chem*., 2005, **44**, 6841.
4. Williams, J. M., Ferraro, J. R., Thorn, R. J., Carlson, K. D., Geiser, U., Wang, H. H., Kini, A. M. and Whanglo, M.-H., *Organic Superconductors (Including Fullerenes). Synthesis, Structure, Properties, and Theory*, Prentice-Hall, Englewood Cliffs, N.J., 1992.
5. Jerome, D. The development of organic conductors: organic superconductors, *Solid State Sciences*, 2008, **10**, 1692.
6. Martins, M. A. P., Frizzo, C. P., Moreira, D. N., Zanatta, N. and Bonacorso, H.G. Ionic liquids in heterocyclic chemistry, *Chem. Rev*., 2008, **108**, 2015.
7. Deetlefs, M. and Seddon, K. R. Assessing the greenness of some typical laboratory ionic liquid preparations, *Green Chem*., 2010, **12**, 17.
8. Ritter, S. R. Ionic liquids go to market, *Chem. Eng. News*, 2008, **86** (39), 36.

9. Prentis, S., *Biotechnology: A New Industrial Revolution*, Orbis, London, 1984.
10. Aehle, W. (ed.), *Enzymes in Industry*, 3rd edn, Wiley-VCH, Weinheim, 2007.
11. Kirby, A. J. and Hollfelder, F., *From Enzyme Models to Model Enzymes*, RSC Publishing, Oxford, 2009.
12. Hahn, F. E. and Jahnke, M. C. Heterocyclic carbenes: synthesis and coordination chemistry, *Angew. Chem., Int. Ed*., 2008, **47**, 3122.
13. Matsumoto, K. (ed.), *Heterocyclic Supramolecules I*, vol. 17 in series *Topics in Heterocyclic Chemistry*, Gupta, R. R. (ed.), Springer, Berlin, 2008.
14. Balzani, V., Credy, A. and Venturi, M., *Molecular Devices and Machines*, 2nd edn, Wiley-VCH, Weinheim, 2008.
15. Feringa, B. L. The art of building of small: from molecular switches to molecular motors, *J. Org. Chem*., 2007, **72**, 6635.

12

The Origin of Heterocycles

The whisper might be born before the lips.
In the woodlessness the leaves were falling down,
And those to whom we dedicate our wisdom,
Before this wisdom had their features crowned.
O. Mandelstam

In this final chapter we try to answer the question posed in Pushkin's poem *Water-nymph:*

'Where art thou from, sweet child of beauty?'

How did the heterocycles essential for life such as the purine and pyrimidine bases, porphyrins, 1,4-dihydronicotinamide, indoles and amino acids first appear on Earth? The philosopher Heraclites of Ephesus wrote 2500 years ago:

'Only those who know the origin and development of matter understand its nature.'

The same idea was offered more simply by the popular Russian writer Koz'ma Proutkov:

'Find the roots of everything and thou willst understand much.'

We know today that many of the above-mentioned heterocyclic compounds can be synthesized by living organisms, and the mechanisms of their biosynthesis are well established. However, to date, the nonbiological origin of heterocycles on Earth remains controversial. It seems obvious that, prior to the appearance of the first primitive life forms on Earth, a stockpile of versatile raw materials must have developed. Thus, the question of the origin of heterocycles is a vital part of the overall mystery concerning the origins of life. While we are far from knowing the whole story, scientific comprehension of the issue has advanced significantly toward the original source of living matter. The origin of life on primitive Earth (protoearth) is presumed to be a result of

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

chemical evolution. We now trace events back to the very beginning, relying on the data gained from cosmochemistry, astrophysics, geology and biology.

12.1 The Origin of the Universe and the Appearance of Chemical Elements

Heterocycles, like all other organic compounds, are formed predominantly from six elements (C, H, N, O, P, S), which are sometimes called 'organogens' as all living species are constructed from them. Of these, hydrogen, carbon, nitrogen and oxygen are especially widespread in the galaxy. It is of interest to consider how and why these elements originated and what their purpose was. Table 12.1 summarizes the main phases in the evolution of the universe.

In accordance with modern theory based on theoretical calculations, computer modeling and astrophysical observations the universe appeared about 14×10^9 years ago as the result of a

Table 12.1 *Time scale of chemical evolution (adapted with modification from Silk, J.,* The Big Bang, *Freeman, New York, 1989, Chap. 4, p. 72, Table 4-1 with permission)*

Evolutionary stage	Time after 'big bang'	Characteristic processes
Big bang	0 to 10^{-36} s	Explosion of a small volume of matter with incredibly high density, temperature and pressure
	10^{-36} to 10^{-11} s	Cosmic inflation – exponential growth of the earliest universe, formation of quark–gluon plasma, electrons and other elementary particles
	10^{-11} to 10^{-6} s	Combination of quarks and gluons with the formation of protons and neutrons
	3–10 min	Big bang nucleosynthesis – combination of protons and neutrons with the formation of deuterium, helium and nuclei of minor light elements (Li, B, Be); most protons remained uncombined as hydrogen nuclei
Formation of the first light elements	380×10^3 years	Combining nuclei of light elements with electrons; formation of hydrogen and helium
Formation and evolution of the first stars	$0.5 - 1.0 \times 10^9$ years	Start of nuclear synthesis of carbon, nitrogen, oxygen and other elements up to iron and nickel
Explosions of supernovae; formation of gas and dust nebulas	5×10^9 years	Nuclear synthesis of heavy elements (radioactive 'clock' commences) and beginnings of chemical evolution in the cosmos
Evolution of gas and dust nebulas; formation of the second generation of stars: appearance of the solar and other planetary systems	$5 - 14 \times 10^9$ years	Synthesis of interstellar particles and molecules
Formation and development of the Earth	9×10^9 years	Molecular evolution on Earth; synthesis of heterocycles and other biologically important substances; biogenesis and biological evolution

gigantic event known as the 'Big Bang'. This involved the explosion of an extremely dense and hot nucleus of matter. The following very short interval (10^{-36} to 10^{-11} s) of exponential growing of the fireball thus arising is called cosmic inflation. It resulted in tremendous expansion and cooling (to 10^{28} K) of the earliest universe and the appearance of quark–gluon plasma and many elementary particles, including neutrinos, electrons and photons. At about 10^{-6} s, quarks and gluons combined to produce neutrons and protons. A few minutes later, when the temperature decreased to 10^9 K, neutrons combined with protons, leading to the formation of atomic nuclei of some light chemical elements, isotopes such as deuterium and helium and much lesser amounts of lithium and beryllium. Most protons however remained uncombined as hydrogen nuclei. This last stage of the Big Bang is known as the Big Bang nucleosynthesis.

After 4×10^5 years the universe had cooled to about 3000 K as a result of its further expansion. Electrons began to combine with protons and other light atomic nuclei to form the first atoms, mostly hydrogen and helium, and these new substances became separated from the cosmic radiation (discovery of the cosmic microwave background radiation is one of the most conclusive confirmations of this theory). The resulting matter, under the action of gravitational forces, began to form into stars and galaxies. The first stars were composed of hydrogen (75%) and helium (25%) and they served as factories where the synthesis of most of the other elements occurred. Due to gravitational compression in a star, a high temperature develops and at 5×10^6 K nuclear fusion of hydrogen begins, resulting in the formation of helium. The helium, as the heavier substance, condenses, increasing star density and temperature even more. When the temperature approaches 10^8 K, the nuclear fusion of helium is triggered, as a result of which three helium atoms form carbon. By a similar scenario, carbon is further converted into oxygen, oxygen into silicon and silicon into iron and nickel. These last two elements are the final products of nuclear reactions inside stars, since the formation of heavier elements is thermodynamically unfavorable. On the whole, during star nucleosynthesis, lighter elements are transformed into heavier by the capture of protons, neutrons and α-particles. The following is a scheme depicting the synthesis of some of the organogenic elements:

$$^1H + {}^1H \rightarrow {}^2D \qquad {}^{12}C + {}^1n_0 \rightarrow {}^{13}C$$

$$^2D + {}^1H \rightarrow {}^3He \qquad {}^{13}N + {}^1H \rightarrow {}^{14}N$$

$$2\,{}^3He \rightarrow {}^4He + 2\,{}^1H \qquad {}^{12}C + {}^1H \rightarrow {}^{13}N$$

$$3\,{}^4He \rightarrow {}^{12}C \qquad {}^{12}C + {}^4He \rightarrow {}^{16}O$$

Due to convection carbon, oxygen, nitrogen and other elements are continuously brought to the stellar surface and fly away as star wind, enriching cosmic space. The fate of a large star (over 10× the mass of the Sun) when nuclear reactions cease is extremely spectacular. When it can no longer maintain a dynamic equilibrium, a huge mass of iron suddenly undergoes gravitational collapse, causing a shock wave of incredible force. Instantly the star throws off its outer shell and the chemical elements accumulated inside it fly away in all directions. The remaining stellar nucleus after such an explosion becomes so dense that, if a star is large enough, a 'neutron star' or even a 'black hole' can arise. The tremendous temperature and pressure developing in a star explosion triggers the nuclear synthesis of all the other elements heavier than iron, including a natural set of radioactive elements.

Exploding stars are astronomically referred to as 'supernovae'. Their explosion is accompanied by extraordinarily bright flashes of light. A supernova flash is a rather rare event. It is estimated that, in an average galaxy, supernovae occur about once in 50 years. The first supernova to be recorded was observed in the middle of the eleventh century. This star could be seen, even during the day, for an entire year. Nowadays, due to space telescopes such as the Hubble instrument, a number of other more distant supernovae have been registered and studied.

Nebulas in the skies represent the characteristic remains of extinct supernovae. These gas and dust residues, which gradually diffused from the site of the catastrophe, are evidence of the ancient explosions. However, gravitational forces then gathered many of these residues together and agglomerated them into dense star-like formations (suns) of smaller sizes. The evolution of gas and dust nebulas enriched with heavy elements led to the appearance of planets. Our galaxy is thought to have about 10^9 stars with planetary systems similar to our solar system.

After their hydrogen burning ceases, stars with a mass equal or less than the Sun have a potential for only one more nuclear cycle, namely that based on helium. The carbon thus formed raises up and scatters in cosmic space as graphite or presumably, tiny diamond particles (searching for the latter is a recent scientific aim). Ultimately, small stars are converted into so-called 'white dwarfs'. Evidence has been recently obtained that, among the 'white dwarfs', purely carbon stars exist.

12.2 Interstellar Molecules

In fact, real chemical evolution begins when the first chemical reactions leading to the formation of molecules becomes possible. This process demands much milder conditions than those existing inside stars. Apparently, they are already provided in upper stellar atmospheres where multiple interatomic collisions produce diatomic molecules such as CH, C_2, CN, CO, NH or OH which are then ejected into interstellar space in pulses which derive from pressure variations between the surface of the star and space. Low-temperature synthesis of more complicated molecules takes place in cosmic gas and dust nebulas. For example, due to the presence of a large amount of hydrogen in space, hydrogenation of the above diatomic species provides multiatomic molecules such as methane, water, hydrogen cyanide, ammonia and others (most of the ammonia is then converted via oxidation or dehydrogenation into much more stable molecular nitrogen). These hydrides are seen everywhere in the cosmos, including in our solar system. For example, numerous lakes of liquid methane were recently photographed by the 'Cassini' spacecraft on the surface of Titan, Saturn's satelite.

When the sky was observed only by optical spectroscopic methods, we could not detect organic molecules in space. However, in the late 1960s radiotelescopes were introduced into practical astronomy. Immediately, in place of the previous silent 'darkness of the night', a multivoiced chorus of interstellar molecules was heard. Formaldehyde was the first organic molecule to be registered in far-off space by radioastronomers. The first signal was detected in 1969 at a wavelength of 6.2 cm. Formaldehyde was later shown to be one of the most abundant compounds in space, and was thus dubbed the 'universal molecule'. About 150 cosmic substances have since been discovered, and the most important (as likely precursors of heterocyclic compounds) are listed in Table 12.2. The majority of these molecules were found in the center of our galaxy and in the Orion nebula.

Along with radio frequencies, microwave, ultraviolet, infrared and mass spectroscopy are now widely used to detect interstellar molecules. Though obtaining conclusive proof in support of one structure is strongly complicated by overlapping absorption bands of different substances, great progress in this field has been achieved. Perhaps, the most remarkable finding is the recent discovery in space of large quantities of benzene and polycyclic aromatic hydrocarbons (PAH): naphthalene, anthracene, chrysene and others. In particular, they were registered by the 'Spitzer' infrared space telescope (launched by NASA in 2007) in the Orion nebula and the 'Cassini' mass/charge spectrometer in Titan's atmosphere. An intriguing point in the origin of PAH is the nature of the building blocks participating in the formation of multiple carbon–carbon bonds. There is weighty evidence that acetylenic compounds (Table 12.2) abundant in space play such a role. Their precursors, in turn, can be methane, ethane, ethylene as well as other C_2 molecules.

Table 12.2 *Selected molecules detected in interstellar medium (retrieved from http://en.wikipedia.org/wiki/List of molecules in interstellar space; accessed 9 November 2010)*

Formula	Name	Formula	Name
NH_3	Ammonia	$H_2NCH_2C{\equiv}N$	Aminoacetonitrile
H_2O	Water	MeOH	Methanol
CH_4	Methane	EtOH	Ethanol
$CH_2{=}CH_2$	Ethylene	MeOMe	Dimethyl ether
C_6H_6	Benzene	$(CH_2)_2O$	Ethylene oxide
$HC{\equiv}CH$	Acetylene	H_2CO	Formaldehyde
$HC{\equiv}C\cdot$	Ethynyl radical	MeCHO	Acetaldehyde
$MeNH_2$	Methylamine	$HOCH_2CHO$	Glycolaldehyde
$CH_2{=}NH$	Methyleneimine	$Me_2C{=}O$	Acetone
$HC{\equiv}N$	Hydrogen cyanide	HCO_2H	Formic acid
$\cdot C{\equiv}N$	Cyano radical	HCO^+	Formyl cation
$MeC{\equiv}N$	Acetonitrile	HCO_2Me	Methyl formate
$CH_2{=}CHC{\equiv}N$	Acrylonitrile	$HCONH_2$	Formamide
$HC{\equiv}CC{\equiv}N$	Cyanoacetylene	H_2NCONH_2	Urea
$HC{\equiv}CC{\equiv}CC{\equiv}N$	Cyanodiacetylene	$H_2NCH_2CO_2H$	Glycine
$H_2NC{\equiv}N$	Cyanamide	$H_2C{=}S$	Thioformaldehyde

Measurements by Cassini's apparatus have established that Titan's upper atmosphere consists mainly of methane and nitrogen. Under solar ultraviolet radiation and Saturn's energetic particles, CH_4 and N_2 molecules undergo dissociation and ionization followed by recombination of the thus formed ion and radical species. This produces a considerable amount of C_2H_6, C_2H_4, C_2H_2, C_4H_2, HCN and a variety of other compounds (Figure 12.1a). In laboratory experiments, acetylenes can easily be cyclized into benzene and other aromatic compounds including PAH. Apparently, such reactions could also proceed in the cosmos, especially in interstellar clouds of high density. However, for most space, including the atmospheres of small planets, a very low concentration of particles is typical, in which random collisions of neutral molecules occur extremely rarely. Under these circumstances the reactions between positively charged ions and neutral molecules are much more effective due to the electrostatic attraction of molecular electron clouds to cationic particles. This type of ion-neutral chemistry has been suggested for the formation of benzene in Titan's atmosphere (Figure 12.1b). Attention should be drawn to the loss of hydrogen in each stage of the process. It supposedly results from the continuous escape of hydrogen from the planet exosphere that shifts an equilibrium towards more unsaturated compounds. Benzene in its ionized forms serves further as a precursor of PAH via similar conversions. Notably, at altitudes of 200–500 km benzene is a main constituent of Titan's atmosphere.

Interstellar molecules are thought to result not only from homogeneous synthesis during the collision of ions and simple molecules in the vapor phase. Also of importance are heterogeneous reactions which take place on the surface of dust particles and granules that make up the interstellar dark clouds. Graphite or silicates covered with dirty ice formed from frozen methane, ammonia, formaldehyde and so on are the constituents of these granules. In laboratory experiments designed to reproduce space conditions (high vacuum, 10 K temperature) many of the known interstellar molecules including formaldehyde, formamide, formic acid and others have been synthesized by UV irradiation of mixtures of NH_3, H_2O, CO and CH_4. It has also been established that during solid phase polymerization induced by radiation, formaldehyde produces polyoxymethylene and polysaccharides at temperatures from 4 to 140 K. Radioastronomers have also discovered

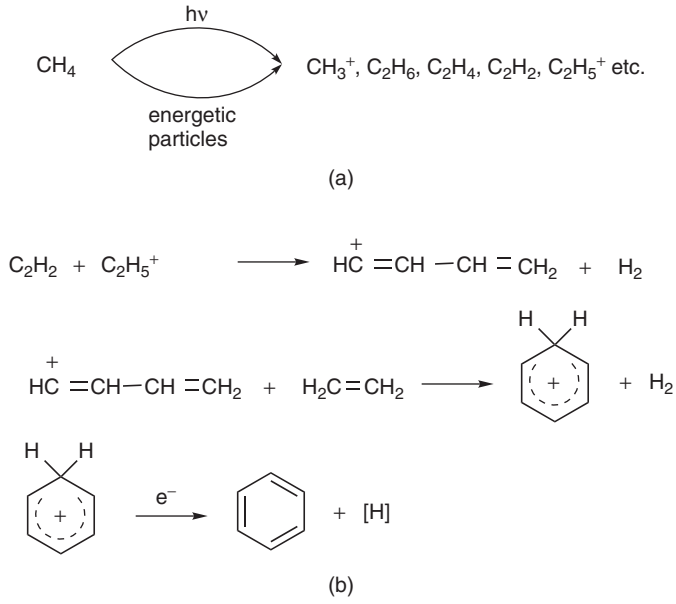


Figure 12.1 *Possible mechanism for the formation of benzene from methane in Titan's upper atmosphere (From Waite, J. H. Jr, Young, D. T., Cravens, T. E., Coates, A. J., Crary, F. J., Magee, B. and Westlake, J.,* Science, *2007,* ***316****, 870. Reprinted with permission from AAAS).*

formaldehyde polymers in interstellar dust clouds. This observation became the foundation of the 'cold prehistory of life' theories.

There are not too many reports on heterocyclic compounds found in the cosmos. The simplest such molecule is ethylene oxide detected in interstellar space in 1997. It is thought to form either via reaction of ethyl radical with atomic oxygen (Figure 12.2a; this conversion proceeds effectively in laboratory experiments) or by a two-step process starting from the attack of a methyl cation on an ethanol molecule (Figure 12.2b).

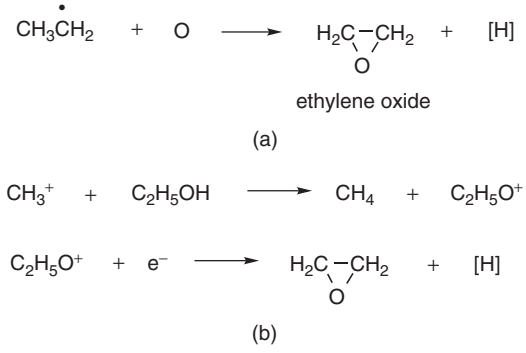


Figure 12.2 *Two possible mechanisms for the formation of ethylene oxide in interstellar space (Reprinted with permission from Dickens, J. E., Irvine, W. M., Ohishi, M., Ikeda, M., Ishikawa, S., Nummelin, A. and Hjalmarson, A.,* The Astrophysical Journal, *1997,* ***489****, 753. © IOP Publishing).*

For more complex heterocyclic molecules, indirect evidence was obtained for the existence of porphyrin-like compounds. A large number of compounds with a triple C≡N bond are found in interstellar space (Table 12.2). In the laboratory, nitriles are widely used to perform heterocyclization reactions. For example, metal-catalyzed coupling of acetylenes and nitriles allows the preparation of various pyridine derivatives (Figure 12.3). Scientists believe that similar processes are possible in space, especially on the surface of granules. In this connection a search for pyridine, quinoline and other azaarenes in interstellar space has been intensified though so far without success.

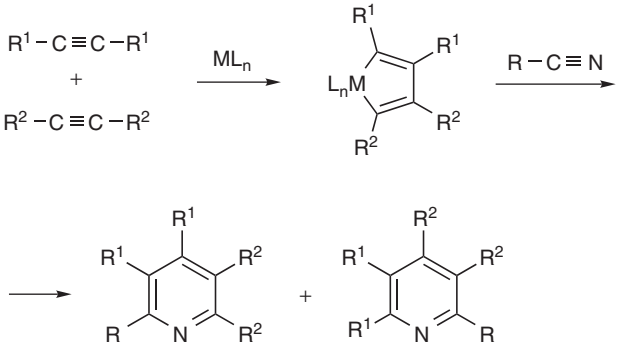


Figure 12.3 *Transition metal-catalyzed laboratory synthesis of pyridines from an acetylene–nitrile mixture (ML_n – metal–ligand complex; Reprinted with permission from Takahashi, T., Tsai, F. Y. and Kotora, M.,* J. Am. Chem. Soc.*, 2000,* ***122****, 4994. © 2000 American Chemical Society).*

Undoubtedly, the most interesting heterocyclic compound so far observed in asteroids and comets is adenine. Adenine is formally a hydrogen cyanide pentamer and numerous experiments have indeed confirmed that it is formed from HCN in vapor, liquid and condensed phases. Hydrogen cyanide is widely distributed in space where it can be formed via collision of the $^{\bullet}$CN radical with various hydrogen atom donors (H, H_2, H_3^+, etc.). It is assumed that another mechanism for the formation of HCN is also possible. Thus, significant quantities of HCN are produced in almost all the model experiments irrespective of the source of the nitrogen (ammonia or molecular nitrogen) and carbon (methane or CO_2). The general reactions can be represented as follows:

$$CH_4 + NH_3 \rightarrow HCN + 3H_2$$
$$3H_2 + N_2 + 2CO \rightarrow 2HCN + 2H_2O$$

Methylamine is believed to be an intermediate which is subsequently dehydrogenated to HCN:

$$CH_3NH_2 \rightarrow CH_2 = NH \rightarrow HCN$$

Hydrogen cyanide is a unique species in that the high polarization of the neutral molecule results in a strong electrophilic character while the CN^- anion possesses high nucleophilicity. This doubly reactive compound thus readily polymerizes to produce dimer, trimer and oligomers (Figure 12.4). Carbon–carbon chains bearing highly reactive functional groups such as C≡N, NH_2 and then =NH are thus constructed. Among these oligomers most important as adenine precursors are tetrameric diaminomalononitrile (DAMN), diaminofumaronitrile (DAFN) and especially 5-amino-1H-imidazole-4-carbonitrile (AICN). The former is a primary product of addition of the fourth

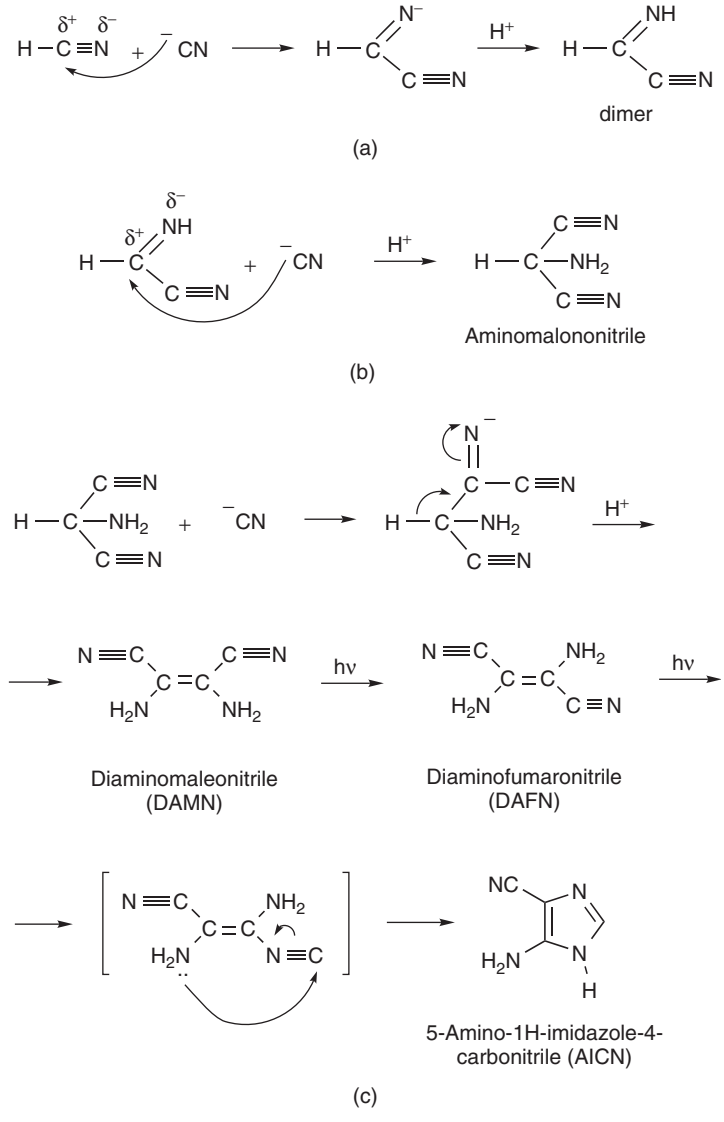


Figure 12.4 *Formation of hydrogen cyanide dimer (a), trimer (b) and tetramers (c).*

HCN to aminomalononitrile. It further isomerizes under photolytic conditions into DAFN, which is converted (also photolytically) into AICN. Note that the last reaction should be accompanied by the formation of an isonitrile intermediate in the isomerization of DAFN.

AICN has been shown to easily add the fifth HCN molecule producing amidine **1a**, existing also in the tautomeric forms **1b** and **1c** (Figure 12.5). The latter are both capable of thermal or photolytic transformation into adenine without any sizable activation barrier. All of the above follows from numerous model experiments and sophisticated quantum-chemical calculations. This

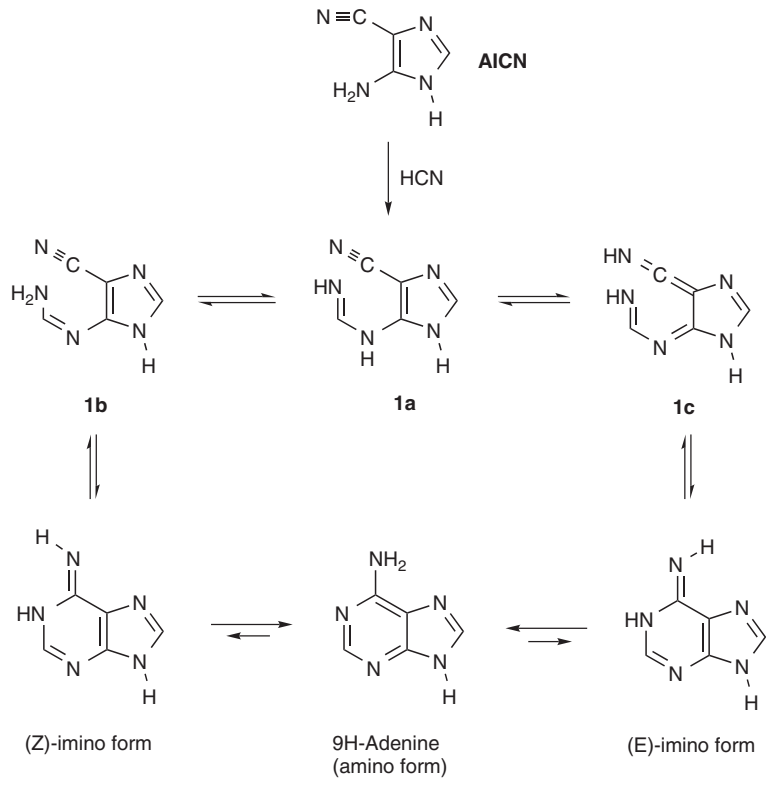


Figure 12.5 *Formation of adenine from 5-amino-1H-imidazole-4-carbonitrile (Reprinted with permission from Glaizer, R., Hodgen, B., Farrelly, D. and McKee, E.,* Astrobiology, *2007, 7, 455. © Mary Ann Liebert, Inc. Publishers).*

means that the appearance of hydrogen cyanide in cosmic space predetermines the inevitable formation of adenine.

12.3 Organic Compounds in Comets and Meteorites

The abundance of organic molecules in interstellar space suggests that they could have been transported to the Earth's surface during the passage of the solar system through the harsh environment of gas and dust nebulas. Another potential source of simple precursors for more complicated molecules is the carbon-based material from which comets and meteorites are composed. We underline in this context the similarity between the isotopic compositions of comet and terrestrial carbon. The Earth is thought to have collided with comets more than one hundred times in its history. The data in Table 12.3 illustrate the role of cosmic organic carriers in 'seeding' our planet with carbon-based matter. We draw attention to the footnote of Table 12.3 which contains some 'food for thought', including a comparison of the carbon content in the Earth's sedimentary rocks with the quantity of carbon derived from the cosmos.

A number of nitrogen heterocyclic compounds were registered in cometary dust by mass spectroscopy during the encounter of the Vega 1 spacecraft with Halley's comet in 1986. Their list,

Table 12.3 *Cosmic contributions in 'seeding' the Earth with carbonaceous matter during the first* 2×10^9 *years*[a] *(adapted from Lazcano-Araujo, A. and Oro, J., in* Comets and the Origin of Life, *ed. C. Ponnamperuma, Reidel, Dordrecht, 1981, with permission)*

Delivery system	Weight of carbon (t)
Comets	10^{16}
Interplanetary dust (particle size less than 1 mm)	2×10^{13}
Meteorites	$10^{12} - 10^{18}$
Interstellar clouds	10^9
Solar wind	1.5×10^8

[a]In the sedimentary layer of modern Earth the carbon content is assessed to be between 1.2×10^{16} and 1.9×10^{16} t. All living matter on Earth contains about 10^{12} t of carbon.

along with the already mentioned adenine, includes purine, pyrimidine, pyridine, imidazole, pyrrole and Δ^3-pyrroline. Halley's comet is also composed of many other carbon-containing neutral molecules, radicals and ions, such as C_2, C_3, CO, CO_2, CS, HCN, ·CN, $\cdot CH_2$, ·CH, C^+, CO^+ and CN^+. The nucleus of the comet contains frozen methane under a black asphalt-like layer. This shell may itself be the product of radiation-induced methane polymerization, as has been demonstrated experimentally.

Exciting results have been obtained in the search for abiogenic organic compounds in meteorites. In the early morning of 28 September 1969, a large meteorite fell to Earth near the Australian town of Merchison. This stone meteorite, thus named 'Merchison', was of the carbonaceous chondrite type,[1] and its extractable matter was found to contain an array of organic substances. Among the heterocyclic compounds observed were three amino acids (pipecolinic acid, histidine, proline), porphyrins, pyrimidines, triazines and purines. The pyrimidine derivatives detected, shown in Figure 12.6, were found to be quite different from the known 'terrestrial' pyrimidines of biological origin. Although the purine bases found in meteorites have familiar structures (e.g., adenine and guanine were separated from the carbonaceous matter of the 'Orguey' meteorite), scientists have no doubt that all of this carbonaceous matter has been formed abiogenically.

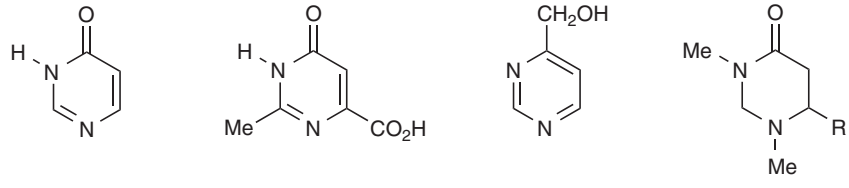


Figure 12.6 *Pyrimidine bases found in the 'Merchison' meteorite: (a) pyrimidin-4(3H)-one, (b) 2-methyl-6-oxo-1,6-dihydropyrimidine--6-carboxylic acid, (c) 4-hydroxymethylpyrimidine, (d) 6-alkyl-1,3-dimethyltetrahydropyrimidin-4(1H)-ones (R = Et, Pr*^{n}*).*

[1] Such meteorites are composed of spherical silicate formations called chondrules. Carbonaceous chondrites contain much greater quantities of carbon-based particles (up to 7%) than other types of meteorites (e.g., iron analogues). Chondrites are readily broken up. About 30% of the carbonaceous matter can be extracted by solvents; the remaining unextractable matter is polymeric in nature.

12.4 Do Heterocycles Exist on the Moon and Mars?

Specimens of lunar rocks and dust retrieved from the moon between 1969 and 1972 by the American spaceship 'Apollo' and the Russian robotic station 'Luna' contained only tiny quantities of carbon (2–200 ppm). The rocks contained small amounts of CO_2, methane, ethane, propane, acetylene, benzene and toluene and trace levels of amino acids. Heterocyclic compounds were not detected, except for the extraction of an insignificant quantity of porphyrin-like substances ($0.005–0.1\ mg\ g^{-1}$) in one instance. There has been no satisfactory explanation to date for the low content of organic matter in the lunar materials.

The chemical probing of Martian soil by the American spacecrafts 'Viking' in 1976 and 'Phoenix' in 2008 demonstrated the complete absence of organic matter although the presence of water and CO_2 on Mars is well established. The wet chemistry laboratory of 'Phoenix' detected in Martian soil the presence of ions Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^- and ClO_4^-, the last being rather unexpected. Since 2000, great interest has been attracted to Mars meteorites found on Earth (overall number is 34). Trace amount of organic matter, identified as polycyclic aromatic hydrocarbons, were detected only in the one that landed in Antarctica. While all lunar and Martian studies so far revealed 'lifeless' deserts, hope still exists that organic oases will one day be found.

12.5 The Atmosphere of Earth and Other Planets

To understand the modes of origin of biologically important organic compounds, we need first to be aware of the significant changes over geological time in the chemical compositions of the atmospheres of the Earth and the other planets within the solar system. Since all of the planets were formed by the agglomeration of gas and dust clouds consisting mainly of hydrogen and helium, there is no doubt that these gases provided the first atmospheres. Each of the heavy planets (Jupiter, Saturn, Uranus, Neptune) retains such a primeval atmosphere. Their gravitational fields are sufficient to hold these light gases within their atmospheres. The smaller planets (Mercury, Venus, Earth, Mars), on the contrary, lost their primeval atmospheres to the cosmic vacuum. As a result of volcanic eruptions and other gaseous discharges, their atmospheres have gradually become heavier because of the presence of carbon dioxide, water vapor, ammonia, nitrogen and the heavier inert gases such as argon. Such atmospheres are referred to as 'secondary'.

The evolution of the gas envelope progressed further on Earth. Scientists believe that the secondary atmosphere on Earth consisted mainly of water vapor, the condensation of which initiated the formation of oceans. Water vapor in the upper layers of the atmosphere, under the influence of lightning discharges and ultraviolet radiation, dissociated to give oxygen and a thin layer of ozone. Up to this time, although the chemical precursors essential for the origin of life already existed on the Earth's surface (see Section 12.6), intense ultraviolet radiation had prevented further progress. It is well known that ozone absorbs ultraviolet rays, and therefore the formation of a protective ozone blanket allowed the synthesis of complex biomolecules and the development of primitive living organisms such as photosynthetic bacteria, simple algae, mosses, lichens and so on.[2] The development of these early life forms and the evolution of more complex plants led to a 'refueling' of the Earth's atmosphere with oxygen through photosynthesis, which gradually resulted in the atmospheric composition we know today. The atmospheric compositions of the planets in our solar system are listed in Table 12.4.

[2] Some scientists believe that the first living organisms on Earth appeared prior to the introduction of oxygen to the atmosphere. Anaerobic bacteria may have been able to develop in a layer of water thick enough to protect them from the ultraviolet radiation prior to the formation of the ozone shield.

Table 12.4 *Compositions of the atmospheres of solar system planets [adapted from Marov, M. Ya.,* Planets of the Solar System *(in Russian), 2nd edn, Nauka, Moscow, 1986, Chap 5, p. 240, Table 4, with permission]*

Planet	Surface temperature (°C)	Main components of atmosphere (vol%)
Mercury	−185 to 510	H_2 (<18), He (<20), Ne (40–60), Ar (<2), CO_2 (<2); atmosphere is highly rarefied
Venus	500	CO_2 (95), N_2 (<3), H_2O (0.1–1.0), O_2 (<0.01), NH_3 (<0.01)
Earth	−80 to 50	N_2 (78), O_2 (21), Ar (0.9), H_2O (0.1–1.0), CO_2 (0.03)
Mars	−80 to 16	CO_2 (95), N_2 (2.7), Ar (1.6), O_2 (0.13), CO (0.077), H_2O (0.03)
Jupiter	−140	H_2 (87), He (12.8), NH_3 (0.01), CH_4 (0.01), PH_3
Saturn	−180	H_2, traces of CH_4, NH_3, N_2, He
Uranus	−170	H_2, CH_4
Neptune	−220 to −160	H_2, CH_4

12.6 Heterocycles and the Origin of the Biosphere

Until the middle of the twentieth century the question of chemical evolution on Earth attracted the attention chiefly of biologists, geologists, paleontologists, poets and philosophers. The French writer de Saint-Exupery once wrote:

> '. . .from a boiling lava-flow, from a stellar substance our life is born . . . a noble cosmic flower.'

Poetic intuition is deep and a philosopher's mind is inquisitive and perceptive. However, a researcher in the natural sciences must probe ideas by experimentation. Chemists have succeeded in applying this concept by successfully reproducing in the laboratory the conditions which existed on primitive Earth. Under these circumstances, practically all of the biochemically important complex organic molecules (including heterocycles) that provided the basis for life on Earth could be prepared from simple precursors.[3]

12.6.1 Simple Precursors of Heterocycles

The data in Table 12.3 indicate that the Earth contains an enormous quantity of organic matter, but early in the history of the universe the variety of its components was not sufficient for the formation of all of the biological compounds. Scientists are now certain that almost all of the compounds necessary for the generation of life could have arisen on the protoearth both in the surrounding atmosphere and on its surface, especially in the ocean. Abiotic syntheses could have taken place on primitive Earth utilizing one or more of the following energy sources: (i) solar ultraviolet radiation, (ii) electrical lightning discharges, (iii) ionizing radiation emanating from the upper layers of the terrestrial crust (at depths reaching 1 km; calculations show that the greatest

[3] This branch of chemistry is called prebiotic (or abiotic) chemistry. Prebiotic chemistry has been defined as 'a search for processes in which organic matter may be brought to life through self-organization' (Sutherland, J. D., *Angew. Chem. Int. Ed*., 2007, **46**, 2354). Prebiotic synthesis is plausible if it deals with starting materials which could be abundant at the site of synthesis; reactions must be conducted in water or without solvent and yields of the products should be reasonable.

source of energy was fission of the potassium-40 isotope), (iv) volcanic heat and (v) shock waves, particularly those resulting from meteorite showers.

It is now generally accepted that the primeval Earth possessed a reducing atmosphere, the chief components of which were methane, ammonia, water and hydrogen. In the 1950s the American scientist Miller demonstrated at the University of Chicago that, under the action of electrical discharges, a mixture of these gases contained in a sealed sterile apparatus could be converted into an array of organic products. Formaldehyde, acetaldehyde, formic acid, acetic acid, succinic acid, lactic acid, fumaric acid, urea, a number of amino acids and HCN were among those found. Variations of the primary mixture composition (e.g., increasing the oxidizing character by the addition of CO_2 or CO, replacing nitrogen by ammonia, etc.) or the use of a different energy source (ultraviolet radiation, bombardment with high-speed α-particles or electrons, even imitating a volcanic eruption) did not significantly affect the experimental results.[4]

Some of these gas phase reactions must be radical in nature. The initial molecules are broken into highly reactive fragments: atomic hydrogen and $Me^{\bullet}$, $NH_2^{\bullet}$ and $OH^{\bullet}$ radicals. These species recombine, dimerize or dissociate further into still smaller particles. Thus, a methyl radical can successively lose all of its hydrogen atoms and ultimately become atomic carbon, as has been proven by the formation of graphite grains in model experiments. Combination of $Me^{\bullet}$ with $OH^{\bullet}$, or of $Me^{\bullet}$ with $NH_2^{\bullet}$, would give methanol or methylamine, respectively. Methanol can be transformed by a similar radical mechanism into formaldehyde, which in turn can be converted to acetaldehyde, formic acid, formamide, glyoxal and other products. In these transformations, the formyl radical $^{\bullet}CHO$ is an intermediate (Figure 12.7).

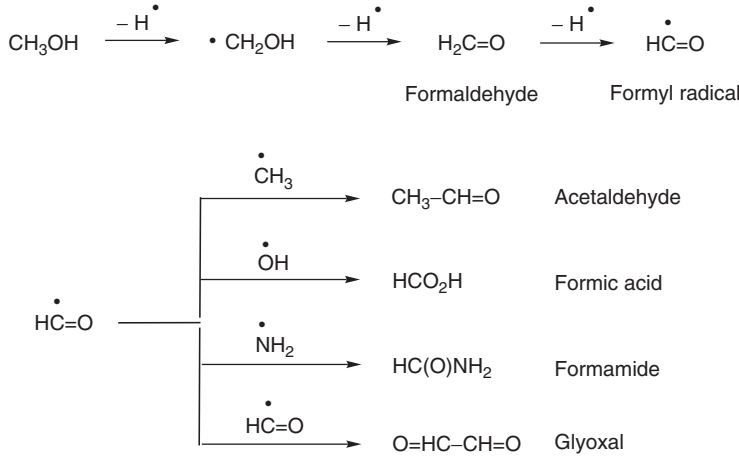
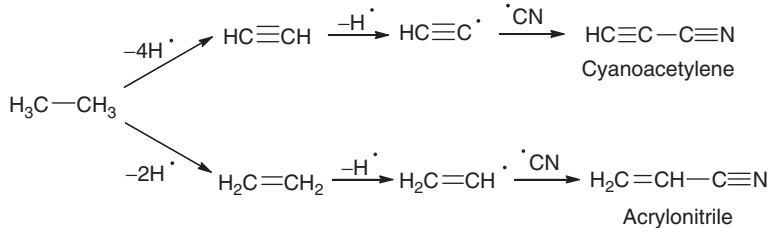


Figure 12.7 *Probable origin of formaldehyde and some of its transformations in the model experiments carried out by Miller.*

We have already mentioned the high significance of hydrogen cyanide in interstellar chemistry (see Section 12.5). There is striking evidence that HCN also played a key role in the formation of many important biomolecules including amino acids, purines, pyrimidines and imidazoles on

[4] Recently Miller's samples that were saved in glass vials have been reanalyzed with modern liquid chromatography and mass spectrometry. Miller's main conclusions were largely confirmed but the reinvestigation has disclosed a considerably larger number of amino acids and alkyl amines formed in his experiments.

Earth. Cyanoacetylene and acrylonitrile (cyanoethylene) are two further compounds of importance in chemical evolution. Their formation can be envisioned as the result of the combination of a $CN^{\bullet}$ radical with an ethynyl or vinyl radical, respectively:



All the above-mentioned radicals, in addition to the simple compounds formed in model experiments, have been detected in interstellar space, in the heads of comets and in the atmospheres of several planets. This provides support for the belief that the model experiments accurately depict chemical evolution on Earth, and reflects the universality of the chemical reactions which take place in various regions of the universe.

12.6.2 Heterocyclic Amino Acids

We have previously mentioned that α-amino acids of nonbiological origin are constantly transported to the Earth via cosmic carriers such as meteorites. However, all the conditions necessary for their spontaneous appearance also appear to have existed on the primitive Earth. In fact, many α-amino acids, including those essential for living organisms, have been prepared in laboratory experiments in which the gaseous mixtures imitate the primeval Earth's atmosphere. The major pathway for their formation seems to be the Strecker synthesis; as shown in Figure 12.8, an aldehyde reacts with a mixture of hydrogen cyanide, ammonia and water. The aldehyde and ammonia first produce an aldimine which subsequently undergoes addition of hydrogen cyanide to its C=N bond. This is followed by hydrolysis of the aminonitrile thus formed to give the corresponding α-amino acid.

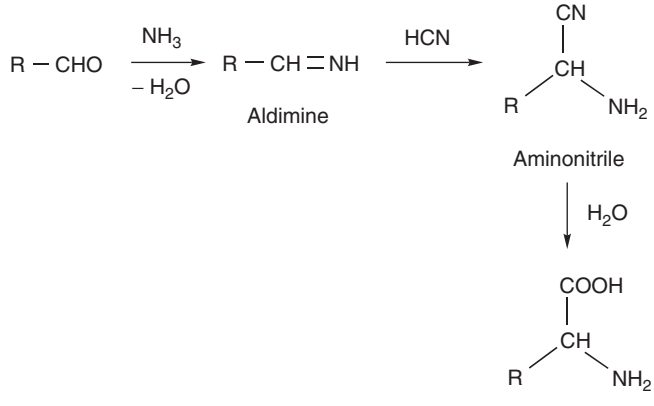


Figure 12.8 *The Strecker synthesis of α-amino acids.*

Unique chemical mechanisms for the abiogenic formation of the heterocyclic amino acids histidine, tryptophan, proline and hydroxyproline (Figure 3.12) are not as yet generally accepted. Histidine has been found in mixtures formed by heating an aqueous solution of formaldehyde with ammonia at 185 °C. The formation of methyleneimine, hydrogen cyanide, acetaldehyde, glyoxal and formic acid is highly probable under these conditions. The subsequent reaction of glyoxal with ammonia and formaldehyde, well known to preparative chemists, can produce imidazole (Figure 12.9a).[5] Imidazole contains two carbon atoms (at positions 4 and 5) with partial negative charges (Figure 2.1) which can be readily attacked by various electrophilic agents. Thus, the addition of glyoxal to imidazole can produce the hydroxy aldehyde **2**, which can be converted (e.g., by the formate anion) to imidazolylacetaldehyde (**3**). The Strecker synthetic pathway then converts **3** to histidine (Figure 12.9b).

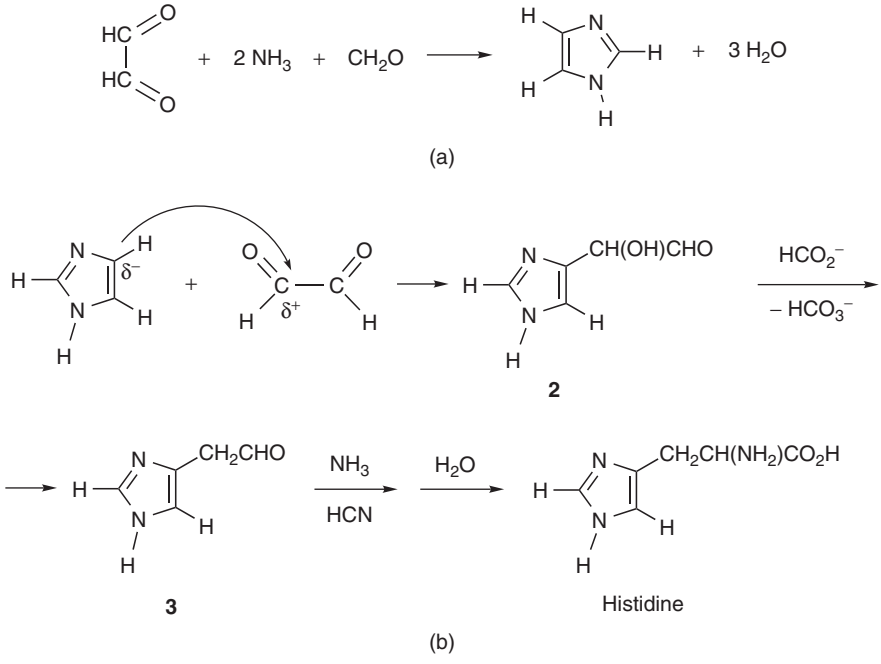
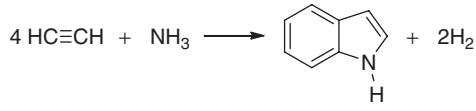


Figure 12.9 *Possible scheme for the abiotic synthesis of histidine.*

A fundamentally similar scheme can be proposed for the synthesis of tryptophan from indole. The reaction with glyoxal occurs at the C-3 position of indole, the position of highest electron density. Evidence for the abiotic synthesis of indole was obtained by the irradiation of aqueous solutions of formaldehyde and ammonium nitrate. Moreover, appreciable quantities of indole are produced during the pyrolysis of mixture of methane (or other low alkanes) and ammonia, a pathway involving the production of acetylene that reacts further with ammonia in a number of steps to form various substances including indole.

[5] The formation of imidazole (often along with indole, tryptamine, urea and other nitrogen bases) was also demonstrated in model syntheses. Imidazole formed when mixtures of CH_4, NH_3, H_2 and H_2O were bombarded with a stream of electrons. Irradiation of an aqueous solution of ammonium nitrate and formaldehyde similarly produced imidazole.



Tryptophan can also be produced from serine. Serine can readily be prepared from glycine and formaldehyde and is a frequent byproduct of abiotic synthetic procedures aimed at providing amino acids. In the formation of tryptophan, serine is believed to be first converted into unstable dehydroalanine which then rapidly undergoes a Michael reaction (addition to an activated double bond) with indole (Figure 12.10).

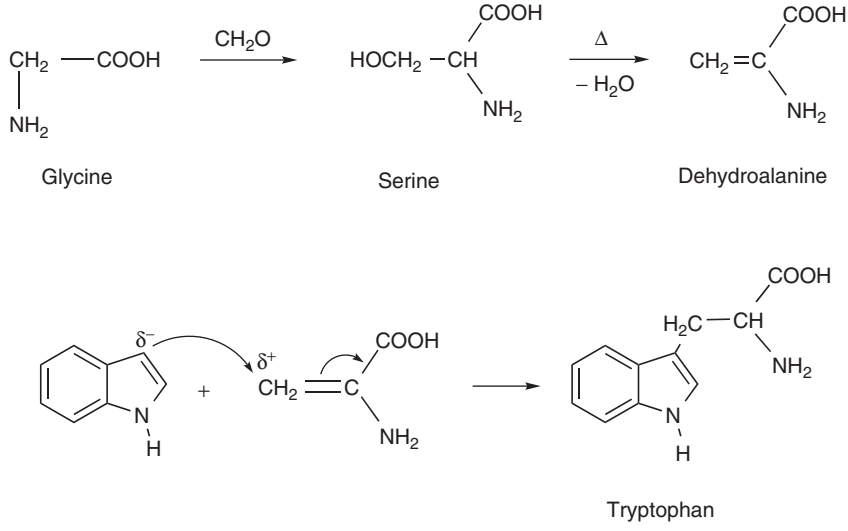


Figure 12.10 *Formation of tryptophan from indole and serine.*

Proline (2-pyrrolidinecarboxylic acid) was also found among the abiotic reaction products. Proline is formed with other amino acids by the application of electrical discharges to mixtures of CH_4, N_2 and H_2O or by heating mixtures of CH_4, NH_3 and H_2O at 900 °C. Interestingly, proline was separated from the aqueous extracts of lava samples taken during the eruption of Maunu Ulu in the Hawaiian islands.

12.6.3 Pyrroles and Porphyrins

Definitive evidence for the theory of the abiogenic origin of porphyrins, pyrimidine bases and a multitude of other heterocyclic compounds was provided when they were detected in ashes and rocks brought up from below the Earth's crust during volcanic eruptions. Traces of porphyrins were also identified among the products in model CH_4—NH_3—H_2O mixtures treated with electrical discharges. The greatest variety of laboratory-synthesized porphyrins was achieved using pyrrole as the precursor. Mixtures containing formaldehyde (or other aldehydes) and pyrrole were treated with ultraviolet light, γ-rays, electrical discharges and heat (100–180 °C). The formation of porphyrins was observed in all cases, but the yields were increased by the presence of oxygen. Oxygen was demonstrated to promote oxidation of the originally produced, partially hydrogenated porphyrins (porphyrinogens) to the final products. Various polypyrrolylmethanes,

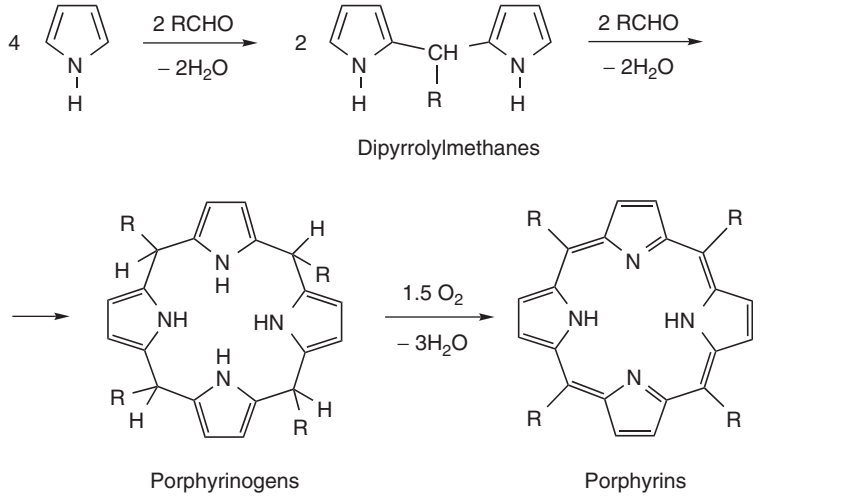


Figure 12.11 *Formation of porphyrins from pyrrole.*

particularly, dipyrrolylmethanes (Figure 12.11), were also found to be precursors of the porphyrinogens.

Pyrrole itself is produced in many different abiogenic conversions such as the interaction of acetylene with ammonia (the Chichibabin reaction in the presence of natural clays), of acetylene with hydrogen cyanide, of glucose with ammonia, and in the pyrolysis of the diammonium salt of mucic acid. However, porphobilinogen (Figure 12.12) is the most important pyrrole system from the point of view of the beginning of life on Earth. This compound is the biosynthetic precursor of practically all the naturally occurring porphyrins such as chlorophyll, hemoglobin, and vitamin B_{12}. The abiotic synthesis of porphobilinogen occurs readily by reaction of succinic acid with glycine. δ-Aminolevulinic acid is an intermediate product in this reaction, as in the biochemical synthesis carried out by living organisms.

12.6.4 Furanose Sugars

Monosaccharides are essential biological building blocks, as are the amino acids and the purine and pyrimidine bases. Since saccharide biomolecules (furanoses and pyranoses) exist preferentially in the cyclic or, more precisely, the heterocyclic form, it is pertinent to investigate the possibility of their arising abiotically. Primarily, the research in this field has focused on an observation by the Russian chemist Butlerov who noted in 1861 that heating aqueous solutions of formaldehyde with barium or calcium hydroxides resulted in the formation of mixtures of carbohydrates (formose reaction). Almost 30 compounds (including trioses, tetroses, pentoses, hexoses) were later identified among the products of this condensation, including such important molecules as ribose, fructose and glucose. The base-catalyzed transformation shown in Figure 12.13 involves a sequence of consecutive aldol condensations of formaldehyde and intermediate glycolaldehyde and glyceraldehyde. Both formaldehyde and glycolaldehyde have been discovered in interstellar molecular clouds (Table 12.2), providing substantial evidence for their existance on the early Earth.

What is the probability of such a pathway ‘seeding’ the Earth with carbohydrates during the prebiological period? At present, serious doubts exist that sugars could accumulate on primordial Earth as a result of the formose reaction. Indeed, the highly basic medium required was

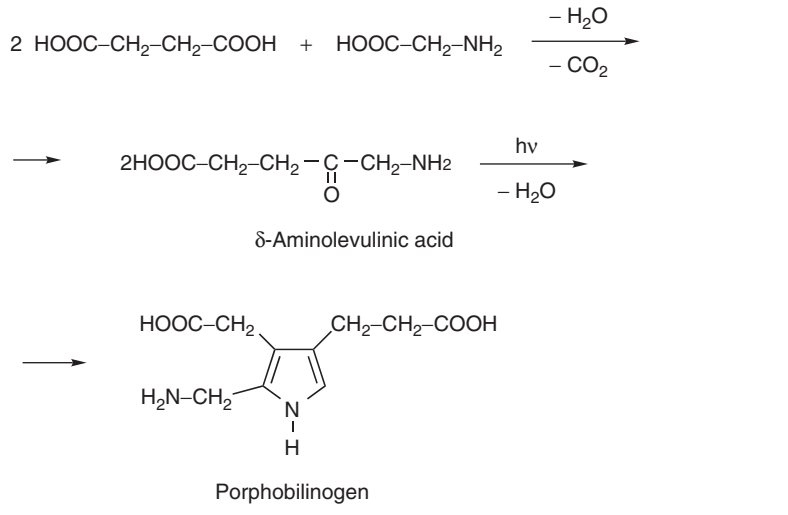


Figure 12.12 *Abiotic synthesis of porphobilinogen.*

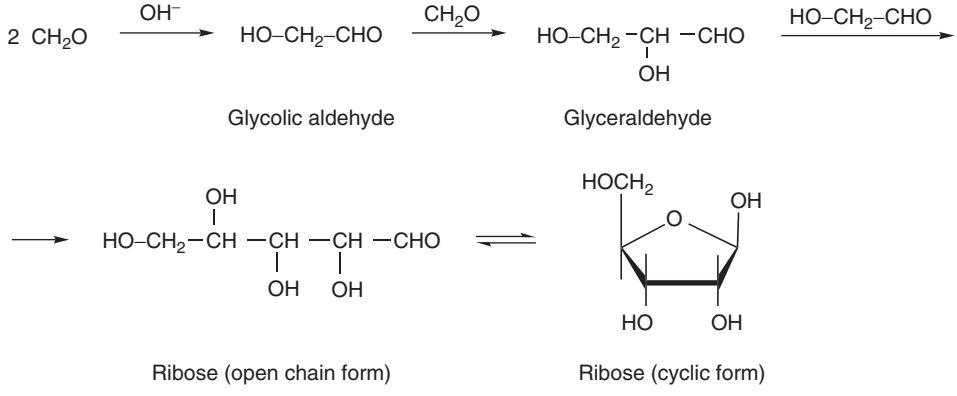
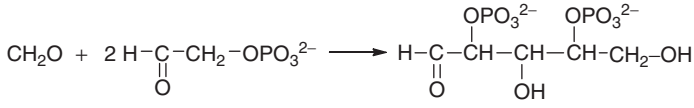


Figure 12.13 *Formation of monosaccharides by the base-catalyzed condensation of formaldehyde.*

generally believed never to have existed in the Earth's oceans. Furthemore, ribose and other sugars are unstable under strongly alkaline conditions. In addition, the yield of ribose, one of the most prominent monosaccharide in the formose reaction, is very small (~1%). A current opinion, based on a number of successful model experiments, is that the prebiotic formation of carbohydrates occurred as a result of the Lewis acid-catalyzed cross-aldolization of formaldehyde, glycolic aldehyde and glyceraldehyde under conditions close to neutral. Thus, it was demonstrated that solutions of formaldehyde, on being heated in the presence of apatite and kaolinite clays, produce a wide palette of carbohydrates, including ribose and deoxyribose. In other experiments pentoses were obtained either on the incubation of formaldehyde and glycolaldehyde in the presence of various borate minerals or on the interaction of formaldehyde with glycolaldehyde phosphate. The latter reaction leads to the direct formation of phosphorylated sugars. The reaction is diastereoselective,

the main product being ribose 2,4-diphosphate with the same configuration of substituents as in natural RNA ribose.



The results of aldolization of an equimolar mixture of glycolic and glyceric aldehydes in the presence of a Zn-proline catalyst in water at ambient temperature are also convincing (Figure 12.14a). They have demonstrated that a self-condensation of each aldehyde is less favored than their cross-aldolization and the total yield of pentoses substantially exceeds that of tetroses and hexoses. Ribose alone accounts for 20% of the total product and displays high stability (no epimerization for two weeks). It is noteworthy that the conditions employed are comparable with a prebiotic environment. Moreover, zinc is an abundant transition metal and proline could form along with other aminoacids on early Earth (Section 12.6.2) or come from extraterrestrial sources. Some key stages of the aldolization mechanism are shown in Figure 12.14b. They reflect an acidification of an α-CH_2 group via coordination of a carbonyl oxygen with Zn^{2+} ion followed by enolization and nucleophilic addition of the enol thus formed to the carbonyl group of the second reagent.

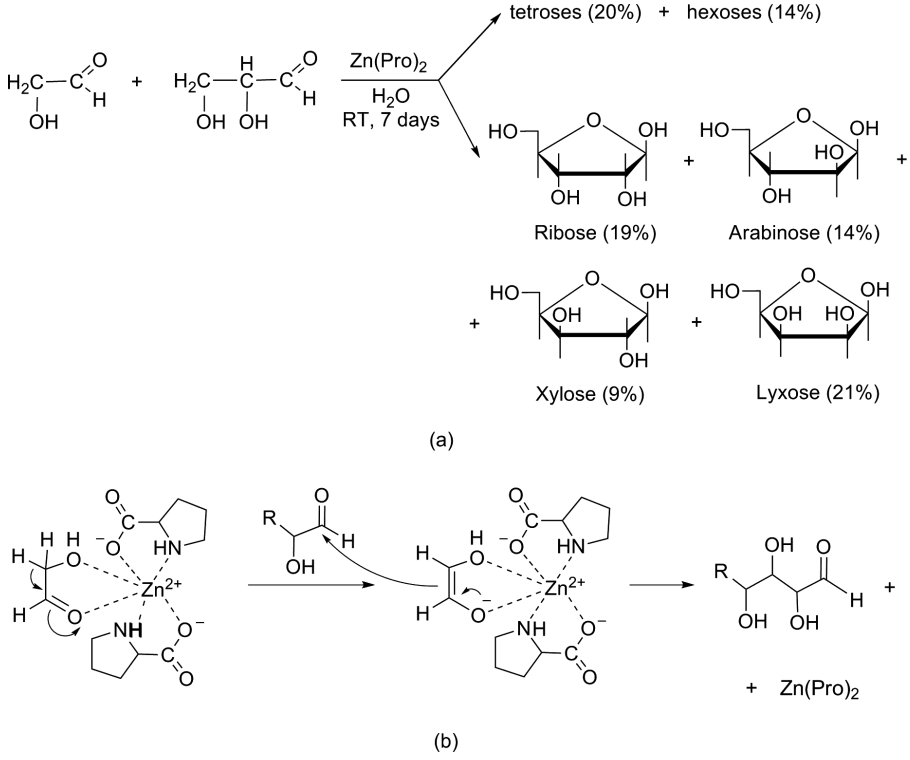


Figure 12.14 *Zn-proline catalyzed cross-aldolization of glycolaldehyde and rac-glyceraldehyde: (a) general results, (b) proposed mechanism (Kofoed, J., Reymond, J.-L. and Darbre, T.,* Org. Biomol. Chem.*, 2005,* ***3****, 1850. Reproduced by permission of The Royal Society of Chemistry).*

12.6.5 Nicotinamide

Life on Earth in the primary 'broth' is unlikely to have commenced in the absence of vitamins (coenzymes) and their precursors. We first consider nicotinamide, the simplest and most important coenzyme. Nicotinonitrile, whose formation was noted in a number of model experiments, was very probably the direct precursor. The action of electrical discharges on CH_4—N_2—H_2 mixtures yields 1% of nicotinonitrile. It is believed that acrylonitrile may be an intermediate in the synthesis of nicotinonitrile (see Section 12.6.1): two molecules combine by a Diels–Alder interaction to form a pyridine ring. Hydrolysis of the nitrile group leads successively to nicotinamide and nicotinic acid (Figure 12.15). Nicotinonitrile can also be prepared by condensation of cyanoacetylene, propenal and ammonia, which themselves could have readily been formed in the Earth's protoatmosphere. The existence of nicotinamide in the primitive ocean suggests that the coenzyme NAD^+ and its reduced form NAD-H may have arisen during the early stages of chemical evolution and controlled some biochemically important redox reactions.

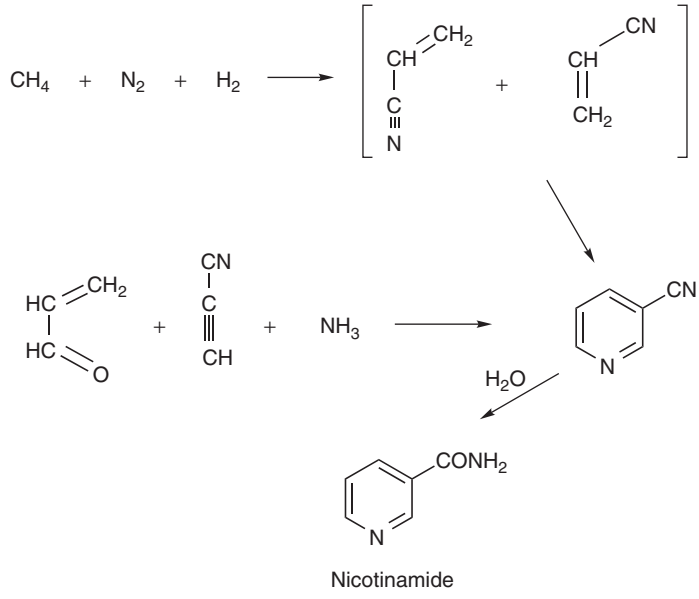


Figure 12.15 *Abiotic pathways for nicotinamide formation.*

12.6.6 Purines and Pyrimidines

As mentioned (Section 12.2) adenine is a very abundant molecule in space due to its ease of formation from hydrogen cyanide. Therefore, one can believe that large amounts of adenine were delivered to Earth with meteorites and comets. However, it also follows from Miller's experiments (Section 12.6.1) that the conditions existing on protoearth could favor the abiotic formation of adenine and guanine – another purine nucleobase. The Spanish researcher Oro obtained experimental evidence for this hypothesis as far back as 1963. He observed the formation of adenine while heating concentrated aqueous solutions of ammonium cyanide. Latter scientists demonstrated that low temperatures could favor purine base formation even to a greater extent. Under freezing

conditions, growing ice crystals trapped compounds, such as cyanides, thus enhancing their local concentration in microscopic pockets of liquid where they could collide more often. Importantly, low temperatures prevent the fast decomposition of nucleobases, allowing for their accumulation prior to the next stage of chemical evolution – the formation of nucleosides and nucleotides.

5-Aminoimidazole-4-carbonitrile (AICN), a tetramer of HCN (Figure 12.4), is generally considered a key precursor of adenine and guanine. Hydrolysis of the cyano group converts AICN into 5-aminoimidazole-4-carboxamide and the following addition to dicyanogen with a subsequent cyclocondensation yields guanine (Figure 12.16).

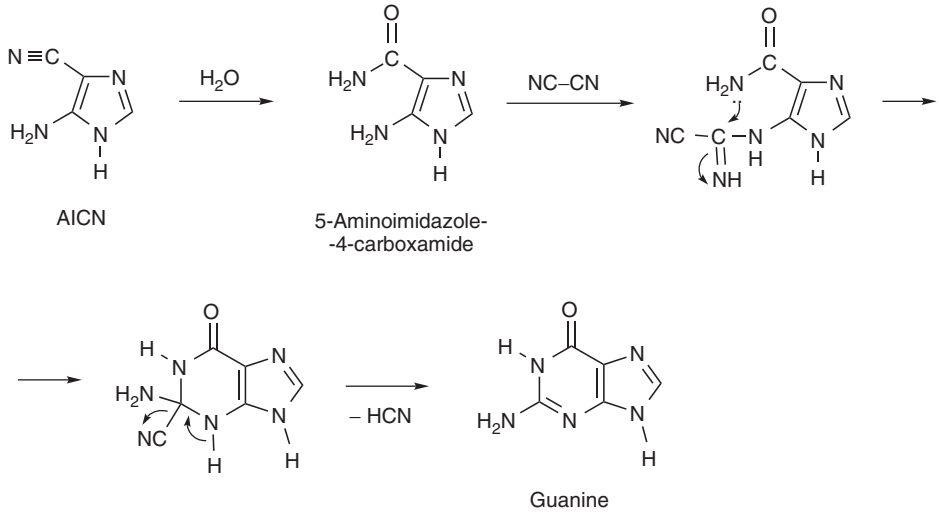


Figure 12.16 *Putative scheme for the abiotic synthesis of guanine.*

Adenine and guanine are also formed in 1.0 and 0.5% yield, respectively, by ultraviolet irradiation of solutions containing HCN or NaCN. It was further established that they can result from the application of either high voltage electrical discharges or cold plasma to gaseous mixtures of CH_4, NH_3 and H_2 in the presence of apatite; in these conversions hydrogen cyanide appears to be an important intermediate. Glycine (a common product of abiogenic reactions) is transformed into purines upon heating.

Many of the details regarding the abiotic origin of pyrimidines are still unclear. So far, the only established synthetic precursor is urea. Originally, all model experiments aimed to obtain pyrimidine nucleobases from urea resulted in low yields. For example, heating urea with cyanoacetylene (100 °C, 5 h) leads to only 5% of cytosine. However, in 1995 it was demonstrated that cyanoacetaldehyde (formed at hydration of cyanoacetylene) is a more successful reagent in this reaction. Thus, heating cyanoacetaldehyde with a concentrated urea solution – such as might exist in an evaporating lagoon on early Earth – produces cytosine with a yield up to 53% (Figure 12.17a). Uracil is then formed from cytosine at hydrolysis and thymine can be produced via methylation of uracil with formaldehyde and formic acid (Figures 12.17b, 3.22).

12.6.7 Nucleosides and Nucleotides

We now have a general concept of how the heterocyclic compounds necessary for life appeared on Earth. The origin of the first biological molecules capable of autoreproduction, is associated with

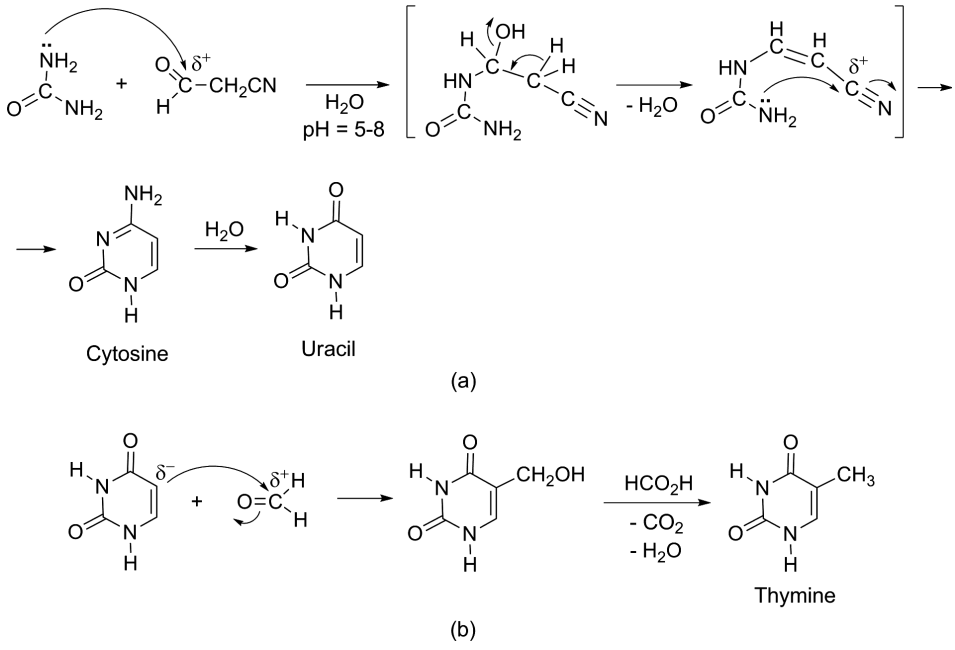


Figure 12.17 *Prebiotic synthesis of pyrimidine nucleobases.*

the subsequent stages of chemical development and especially with the appearance of polymeric compounds such as polynucleotides, polypeptides and polysaccharides. Our description of this epoch in molecular evolution is, to a large extent, speculative. We commence with an examination of the syntheses of nucleosides and nucleotides from purine and pyrimidine bases (see Section 3.1).

If the nucleobases coexisted in primordial 'soup' with monosaccharides and many other substances, including inorganic phosphates, their assembly must proceeded with reasonable selectivity both between sugar OH groups and nitrogen heteroatoms (e.g., N-6, N-7 or N-9 for adenine). An activation of these functionalities was also essential since hydroxyl groups can not be directly replaced by nucleophiles at neutral pH and nitrogen heteroatoms are not sufficiently nucleophilic. Chemists still cannot achieve such remarkable selectivity and so they have to protect any OH or NH_2 group that is not to react. This can be illustrated by a classical laboratory method for the preparation of adenosine and other purine nucleosides (Figure 12.18). The multistaged synthesis involves the preliminary preparation (not shown here) of protected 9-chloromercuriadenine (**4**) and ribofuranosyl chloride (**5**), their interacton in boiling xylene and the final elimination of acetyl groups. Further preparation of adenosine-5′-monophosphate demands a similar set of protection–deprotection procedures. Obviously, such chemistry could not be realised on early Earth because of its complexity, namely, the use of a nonaqueous solvent and nonprebiotic reagents and a lack of self-organization.

Many of the experiments aiming to develop a plausible prebiotic synthesis of purine and pyrimidine nucleosides have been inconclusive and this is the weakest place in the chain of prebiotic reactions leading to polymers. Thus, heating in a dry state adenine or guanine with ribose and sodium trimetaphosphate in the presence of $MgCl_2$ resulted in the formation of adenosine (Figure 12.19) or guanosine with only 4 or 9% yield, respectively. This process with its unclear

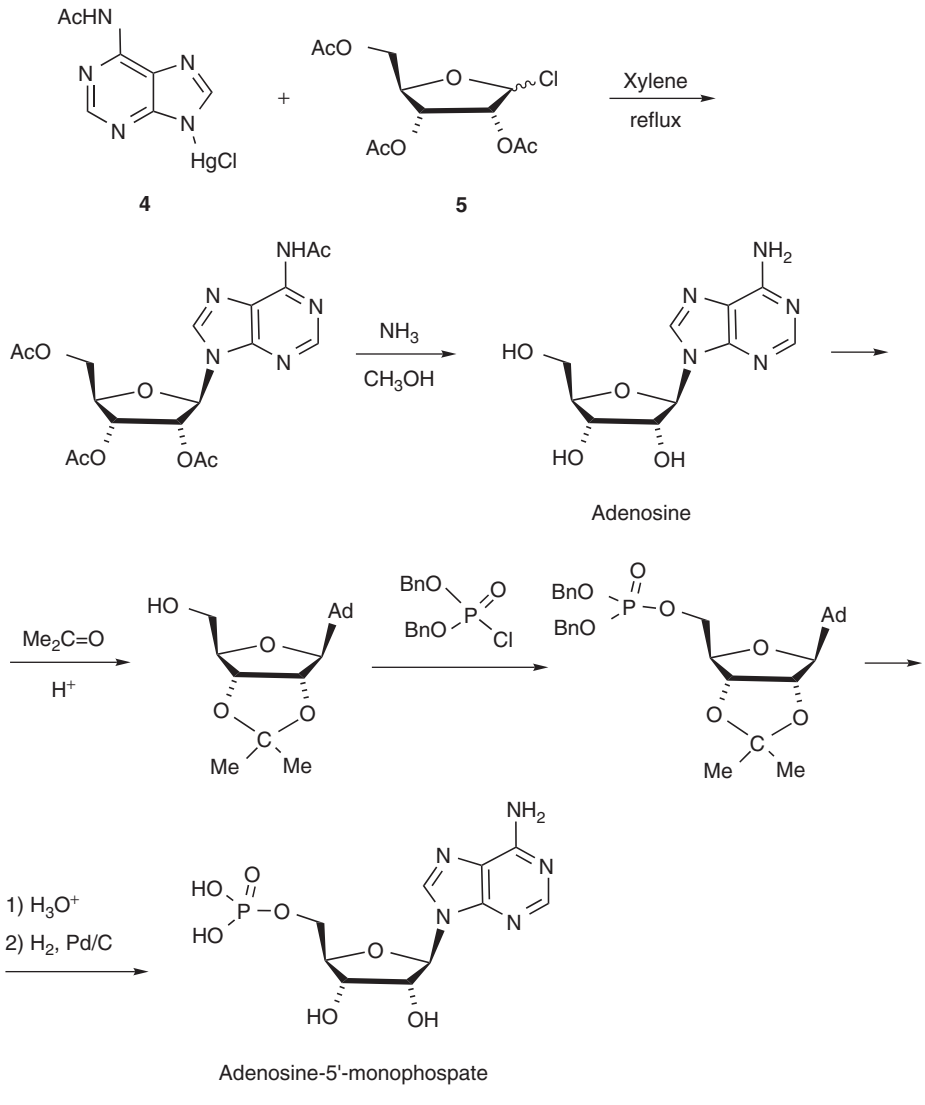


Figure 12.18 *Laboratory synthesis of adenosine and adenosine-5′-phosphate (Bn = $C_6H_5CH_2$).*

mechanism was non-selective and considerable amounts of 6-ribosylamino- and 6-ribosylamino-9-ribosylpurines were also formed in the case of adenine. No direct union of cytosine or uracil with ribose to prepare pyrimidine nucleosides abiotically has been reported, although several indirect approaches were suggested.

Imitating the prebiotic syntheses of nucleotides is also a complex and still unresolved problem. The main difficulties consist in identifying an ancient phosphorylation agent and excluding water as the solvent. The phosphorylation of alcohols is energetically unfavorable as water hydrolyzes an ester bond that shifts the equilibrium towards the starting compounds. Preparatively, this is commonly overcome by special activation of the hydroxy group and the use of condensing reagents

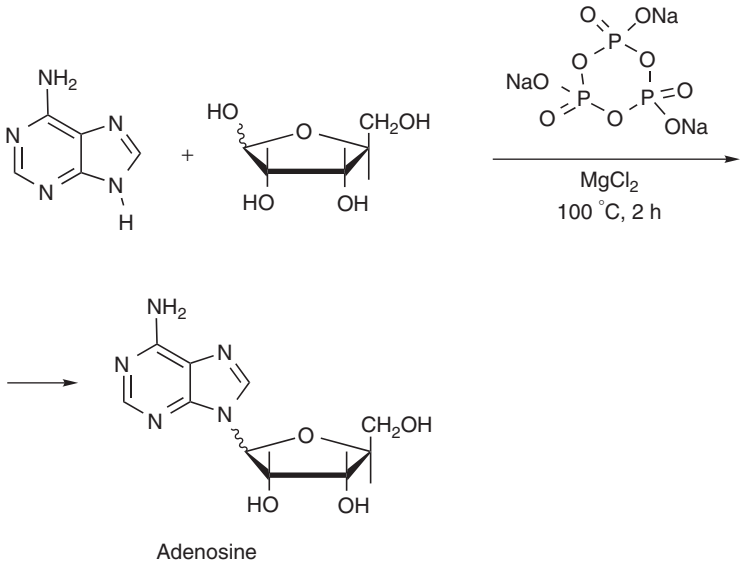


Figure 12.19 *Prebiotic synthesis of adenosine.*

that bind water. Carbodiimides easily react with water to form urea derivatives, serving as a typical example. Perhaps under the conditions of protoearth, the function of activating and condensing agents could be fulfilled by cyanate, cyanamide or urea. Almost all the phosphorus on Earth exists in rocks and minerals as insoluble calcium orthophosphate and polyphosphates. Therefore, these materials are the most plausible source of phosphate for the solid-phase prebiotic nucleotide synthesis. For example, uridine very slowly reacts with a large excess of urea and $Ca_3(PO_4)_2$ at 100 °C, giving a complex mixture of O-phosphorylated products. The process is strongly accelerated when calcium phosphate is replaced by ammonium phosphate; possibly such accelaration is caused by proton catalysis.

The absence of any demonstrated prebiotic synthesis for nucleosides and nucleotides stimulated alternative ideas. In particular, perhaps these compounds were formed on protoEarth not from free nucleobases and ribose but indirectly from some other precursors. Originally, scientists believed their formation should be preceded by the biosynthesis of nucleobases which then couple with a monosaccharide and a phosphate. However, this not the case. In reality, as shown by adenosine biosynthesis, a long chain of complex enzymatic reactions results first in a step-wise attachment of the imidazole ring to a phosphorylated ribose that is followed by pyrimidine ring construction (Figure 12.20). Note that all four nitrogen heteroatoms of the purine system are derived from three amino acid precursors – glycine, glutamine and aspartic acid.

The fruitfulness of a nonstandard approach was recently demonstrated by J. D. Sutherland and his coworkers (*Nature*, 2009, **459**, 239), who discovered an effective route of prebiotic synthesis for activated pyrimidine ribonucleotides which bypasses free pyrimidine bases and ribose (Figure 12.21). As starting compounds, two- and three-carbon building blocks abundant in the Universe and on early Earth were used: cyanamide, cyanoacetylene, glycolaldehyde, glyceraldehyde and inorganic phosphate (see Table 12.2 and Section 12.6.1). In the first stage 2-aminooxazole **6** is

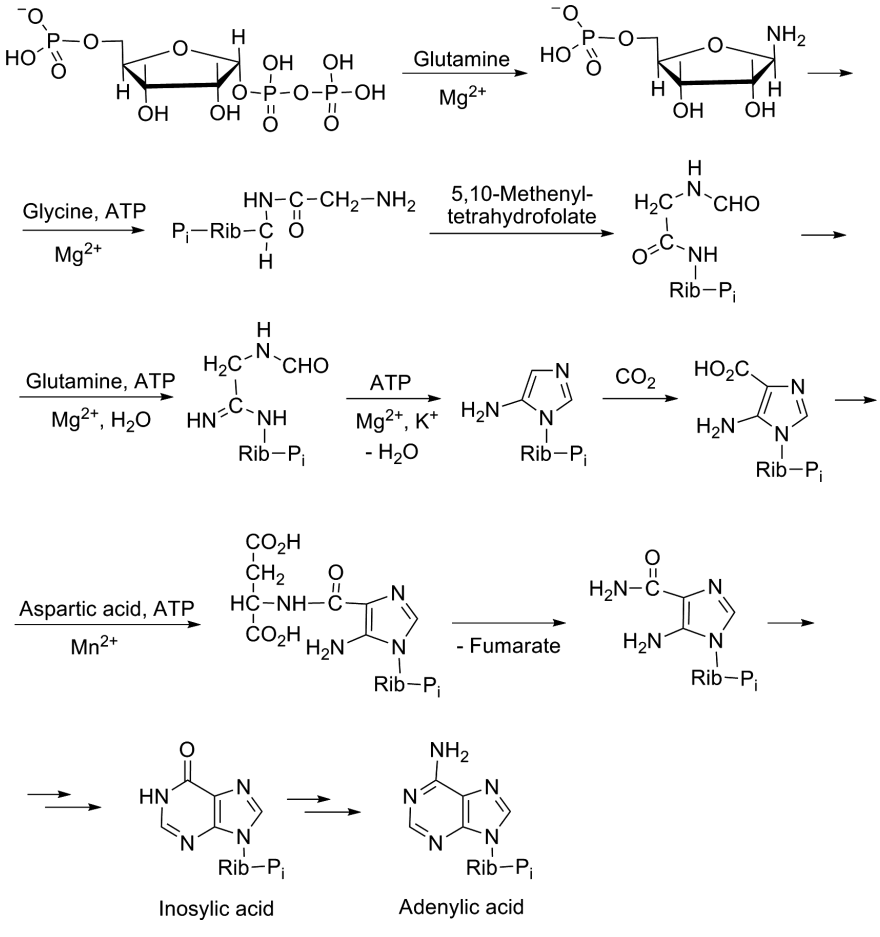


Figure 12.20 *Biosynthesis of purine nucleotides (Rib-P_i – 5-phosphorylribosyl-1).*

obtained by reaction of NH_2CN with glycolaldehyde. This amine enters a cycloaddition reaction with glyceraldehyde producing arabinose derivative **7** together with minor quantities of some other pentose amino-oxazolines. Interaction of **7** with cyanoacetylene gives anhydrobase **8**, then reacting with phosphate to yield the β-ribocytidine-2′,3′-cyclic phosphate **10**, possibly via intermediate **9**. It is noteworthy that the intramolecular S_N2 substitution of the oxazoline oxygen atom by the phosphate oxygen results in the inversion of the stereochemical configuration with the transformation of arabinose into ribose. The whole cycle is completed by a final photochemical step in which the cytidine fragment is hydrolytically converted into uracil derivative **12** through photohydrate **11** (cf. Figure 12.17). All steps proceed under mild conditions (water, pH 6.5–7.0, room temperature or moderate heating), with excellent to moderate yields and good regio- and stereoselectivity. The authors of this discovery view the above prebiotic synthesis as predisposed; if true, Nature created beautiful examples of heterocyclic chemistry as far back as in the beginning of its existence.

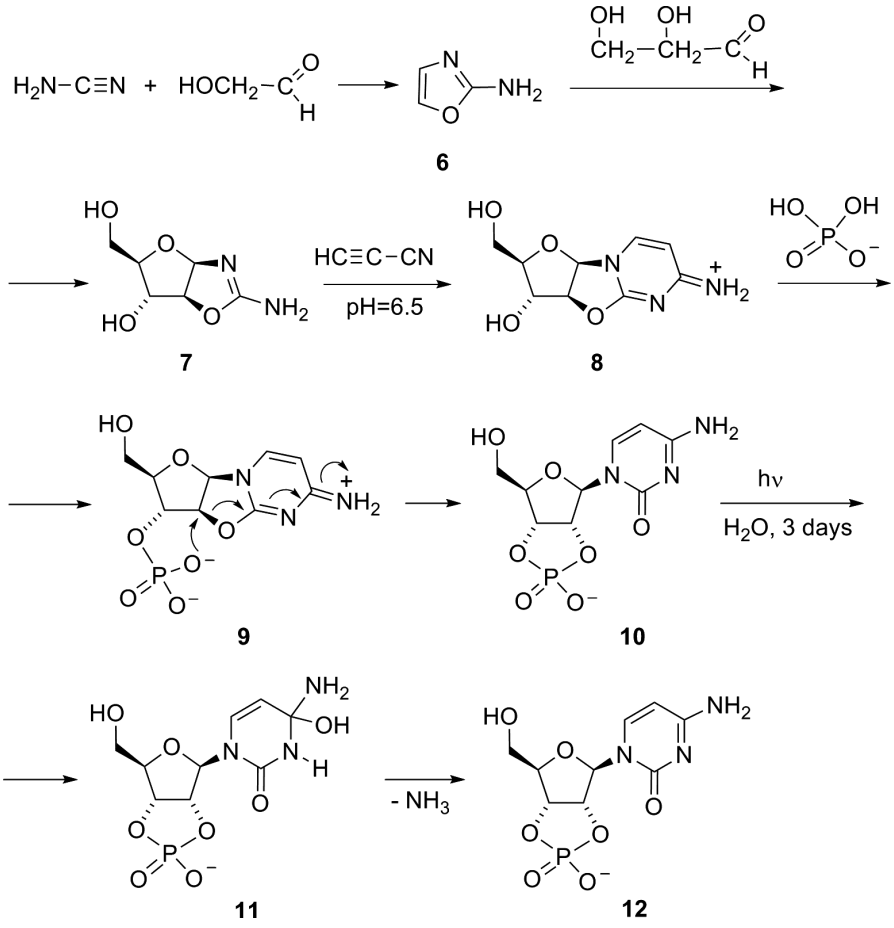


Figure 12.21 *Pyrimidine ribonucleotide assembly bypassing ribose and free pyrimidine nucleobases (Reprinted with permission from Macmillan Publishers Ltd. Powner, M.W., Gerland, B. and Sutherland, J.D.* Nature*, 2009,* ***459****, 239. © 2009).*

12.6.8 Polynucleotides and the Birth of 'Animated' Organic Molecules

The prebiotic synthesis of polynucleotides and the origin of nucleic acids with a defined sequence of bases are the least understood areas of chemical evolution. Modern theory assumes that a genetic system can be viable only if the length of the informational polymer is in the range of 30–60 monomers. This Rubicon was crossed when nonliving matter was animated and acquired the primitive capacity to reproduce, that is, to record and transfer chemical information, to possess enzyme-like catalytic properties and to have the ability to store and transmit energy – in other words, to show the characteristics of primary metabolism. In the course of molecular evolution these qualities were further adapted and refined.

The transition from nucleotides to polynucleotides in aqueous solution is thermodynamically unprofitable and cannot proceed spontaneously to a significant extent. All the essential nucleotides based on uridylic, cytidylic and adenylic acids (Figure 3.3) on evaporation of their acidic solutions

and subsequent heating at 65–150 °C give complex mixtures of short oligonucleotides with random 2′-5′ and 3′-5′ phosphodiester linkages. Such polynucleotides could not serve as the basis for a genetic code, not only because of their short length but also because they contain only one purine or pyrimidine base.

Polymerization of nucleotides in water requires an external catalyst. Originally, the best method of activation was probably by conversion of nucleotides into phosphoramidates such as a nucleoside-5′-phosphorimidazolide (**13**). Various minerals and metal cations, especially Pb^{2+} and uranyl $[UO_2]^{2+}$, could have been successful as catalysts. For example, incubation of activated nucleotides of type **13** with montmorillonite (alumosilicate mineral clay) provides, via dimers **14**, polynucleotide chains up to 30–60 monomers long and mostly with the 3′-5′ sequence (Figure 12.22).

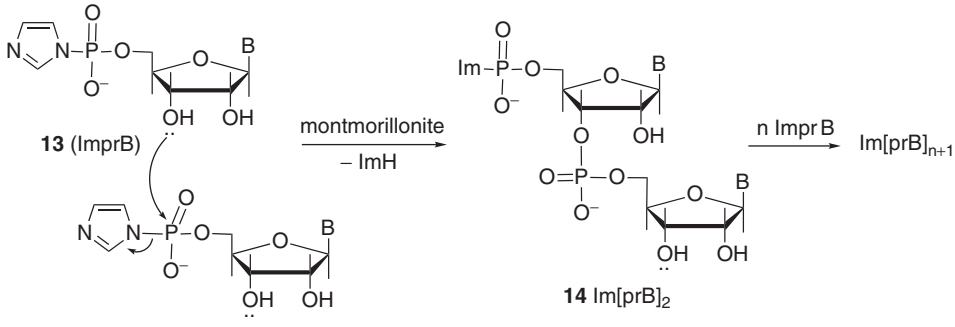


Figure 12.22 *Prebiotic synthesis of oligomeric polynucleotides from preactivated nucleotides (Im – imidazolyl-1, B – any nucleobase, pr – phosphorylribose).*

Under the conditions which existed on the protoearth, the inclusion of various bases into the polynucleotide chain could have been caused by their high concentration in certain reservoirs, or on solid surfaces. Various metal cations such as Mg^{2+}, Ca^{2+}, Zn^{2+}, Fe^{3+} and so on are believed to have functioned as adsorptive centers in the clays. As the cations differ in their affinities toward each of the bases, they became a means of regulating the selection of nucleotides to be attached to the growing chain. It is thought that the first polynucleotides were probably of relatively low molecular weights and consisted of a limited number of bases. Adenine–thymine oligonucleotides may have prevailed and played a role in the primitive transport of RNA.

Undoubtedly, the most critical event in prebiotic evolution was the appearance of the first self-replicating molecules, commonly thought to be single-strand polynucleotides that included both ribose and deoxyribose with a variable number of heterocyclic bases. The exact sequence and nature of all the steps that initiate such replications remain unknown. However, the opinion is wide spread that it includes the synthesis of complementary RNA on a preformed RNA template, the mechanism involves free-floating nucleotides and short oligonucleotides that approach the matrix and stick to it through Watson–Crick base pairing, followed by ligation (Figure 12.23). Thus, the matrix sequence of bases catalyzed the reproduction of similar polynucleotides.

Numerous model experiments of template-directed synthesis suggest that for effective ligation (optimal rate, 3′-5′ linkage, etc.), preliminary activation of the nucleotides and the presence of divalent metal cations are needed. On early Earth the efficiency of the syntheses also would have been enhanced after the transition from open aqueous media (oceans, lakes, lagoons) to closed, phase-limited systems such as bubbles, microspheres and droplets, which are considered to be protocytes – the first models of living cells. Owing to the polymerization processes, concentration

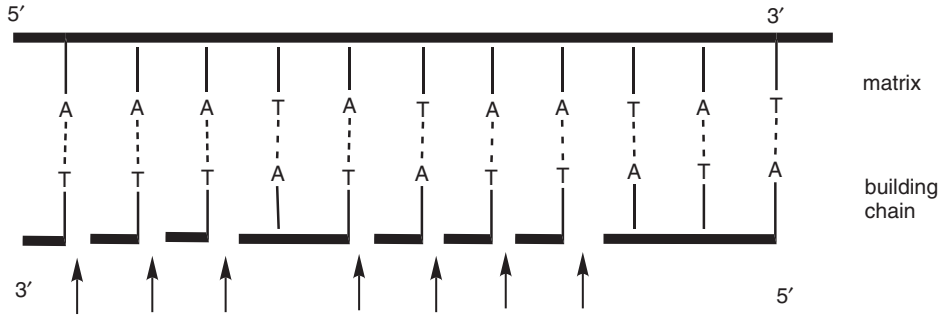


Figure 12.23 *Schematic representation of template-directed RNA synthesis (arrows indicate ligation sites).*

gradients of nucleotides and amino acids could appear in protocytes isolated from the environment by a pellicle, a membrane. As a consequence of the osmotic pressure which developed as a result of the concentration gradient, these substances were pumped from the external aqueous solution (the so-called 'primary broth') to the inside of the microspheres. In this manner a primitive metabolism accompanied by promatrix synthesis of biologically significant polymers may have originated.

In modern living cells the polymerization of nucleotides is catalyzed by the protein enzyme polymerase and by nucleic acids mediated protein synthesis. From this the intriguing 'chicken and egg' dilemma arises: which came first, the proteins themselves or the nucleic acids. We have already seen that various experimental abiotic conditions yield complex arrays of amino acids. Such conditions also lead to the formation of peptides. The evidence gathered thus far suggests that during the early stages of molecular evolution, polypeptides and polynucleotides came into existence independently of each other and in the case of their appearance in the same phase-isolated system, evolutionary selection factors came into play. Macromolecules with the inherent ability to store and transfer information were more stable and adapted to new environmental situations and were thus naturally selected. Polypeptides and polynucleotides seem to comply with the rule of cross-stereocomplementarity. This means that a mutually specific recognition should exist between them, based on hydrophobic and other nonbonding interactions and also on steric relationships. Peptides initially may have played the dominant role but control gradually shifted to the nucleic acids which carry more precise information and are more stable.

Natural selection eventually retained those structures and fragments which ensured the greatest thermodynamic stability of the macromolecules, the highest stability of complementary base pairs, the best selectivity and the optimum rate of matrix replication. Eventually, separation of macromolecules into DNAs and RNAs also occurred. The former, being more stable, were useful for the storage of information; the latter, less stable owing to the interaction of the two hydroxyl groups at the C-2 and C-3 positions of the ribose residue, were nevertheless more efficient in the transport of amino acids and in the synthesis of proteins. Thus, nature separated the 'chaff from the wheat' and the 'sheep from the goats' at the molecular level. The principle of nitrogen base complementarity was achieved only when the appropriate asymmetry of the sugars in the nucleic acids had developed. Over time it became possible for the D-form of ribose to be separated from the L-enantiomer. The development of the double-helix structure endowed a higher stability towards random modifications and mutations than the single-chain macromolecule. The structural peculiarities of the polynucleotides were enhanced by the helical conformation, thus making these molecules hereditary. The evolutionary process led to a remarkable result – the genetic code of the modern bioworld.

In Section 4.4 we saw that in addition to their role as carriers of genetic information, RNAs also possess some of the catalytic properties of enzymes. Hence, they have been called 'ribozymes'. The most important ribozyme in nature is the ribosome which catalyzes protein synthesis (Section 3.4). Ribozymes can also hydrolyze phosphodiester bonds, causing the enzymatic cleavage of another RNA and even self-splicing. In the early 2000s it had been found that specially tailored ribozymes possess properties of RNA polymerase. The latter gave strong support to the idea of an early 'RNA world' (Section 4.4).

However, neither the prebiotic synthesis of ribonucleotides nor RNA replication would have been easy under the conditions of the primitive earth. Indeed, the RNA world itself may not be the first coding system to have arisen on the protoearth. A whole number of simpler genetic systems which might have preceded RNA were therefore brought into consideration: pyranosyl RNA (*p*RNA), threosyl RNA (TNA), peptide RNA (PNA) and some others (Figure 12.24a–c). As RNAs, these artificial nucleic acids are able to form a double helix and in some cases (PNA) chimeric complexes with RNAs. It is noteworthy that, despite differences in their backbones, all these nucleic acids have the same heterocyclic bases for coding genetic information.

In 1995 Orgel and coworkers had shown that an achiral decameric oligopeptide having attached cytidine bases (PNA-C_{10}) and structure shown in Figure 12.24c served as a template for the

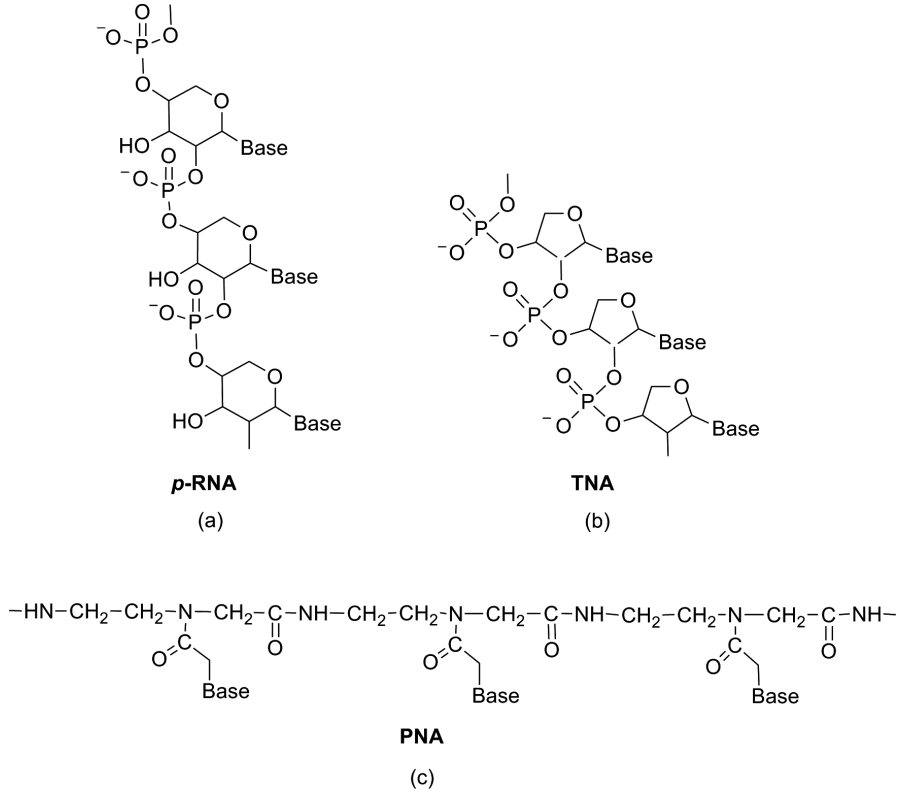


Figure 12.24 *Some artificial nucleic acids as possible candidates for a pre-RNA world.*

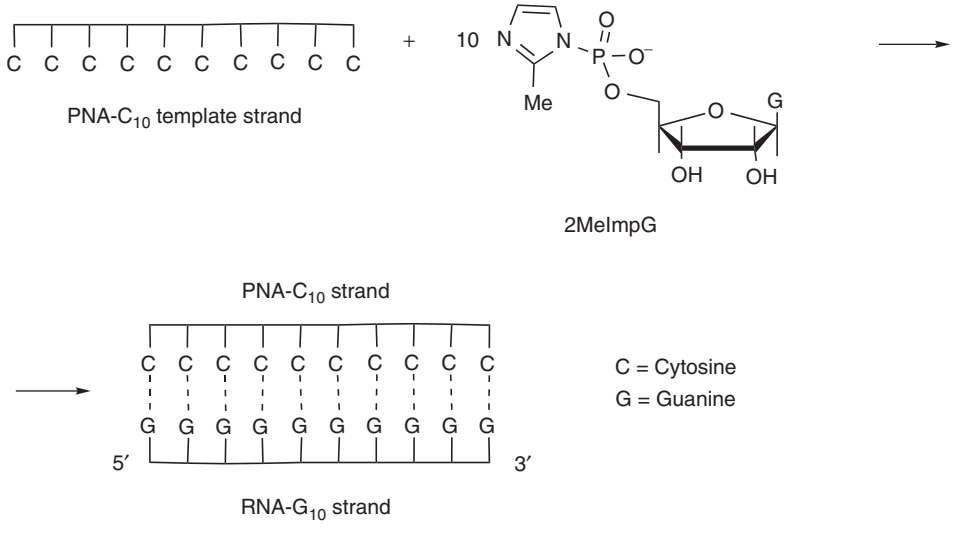


Figure 12.25 *Template-directed synthesis of an RNA-G_{10} oligomer (PNA-C_{10} as template, 2MeImpG as substrate).*

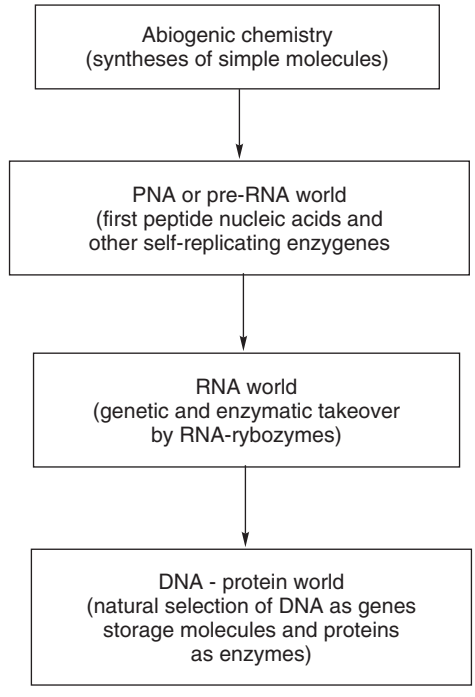


Figure 12.26 *General scheme of evolution of information-carrying polymers from abiotic molecules.*

synthesis of a complementary RNA decamer by successive ligation of guanosine mononucleotide units, each activated by a 2-methylimidazolyl group (2MeImpG; Figure 12.25).

The PNA template accelerates the intermolecular condensation of the guanosine mononucleotide substrate, probably through a complex with Watson–Crick base pairing. In the course of this transformation all the possible *n*-mers are produced. The di- and trimers initially formed possess predominantly the nonnatural 2′-5′ phosphodiester linkage. However, the steps of further elongation to tetramers, pentamers and so on are highly specific and give the natural 3′-5′ linkages. This means that a generic transfer has been effected from PNA ('enzygenes') to RNA ('ribozymes') without

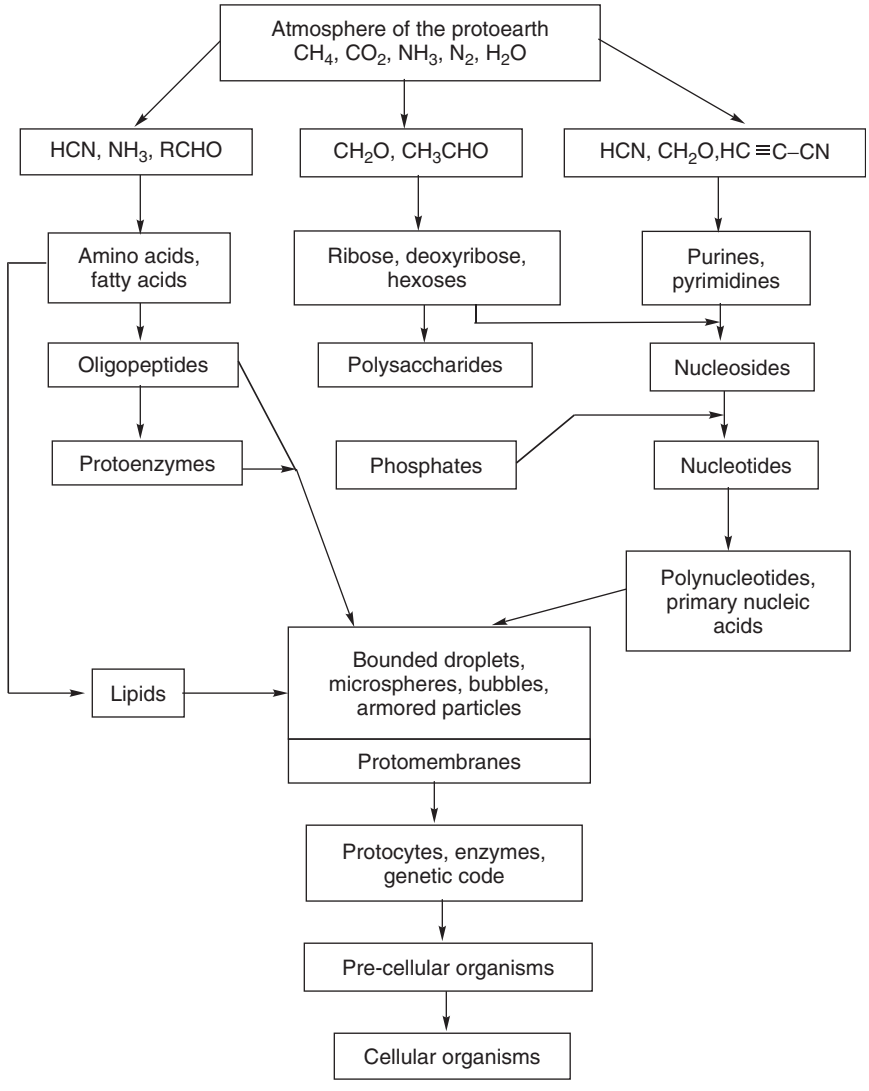


Figure 12.27 *General scheme of chemical evolution (adapted from Lazcano-Araujo, A. and Oro, J., in* Comets and the Origin of Life*, (ed. C. Ponnamperuma), Reidel, Dordrecht, 1981, with permission).*

loss of information. Significantly, a cytidine DNA decamer can direct the ligation of guanosine PNA dimers in this experiment, which suggests the possibility of reverse transcription and the probability of the partial evolution of more than one self-replicating protosystem. This could fill the gap between contemporary information-carrying polymers and the prebiotic chemistry of organic molecules (Figure 12.26).

Natural selection at the nucleic acid level continues to function at the present time. Without this line of reasoning it is difficult to understand the origin of bacteria that live in the cooling water of atomic piles in which the radioactivity reaches 1000 roentgens. Without natural selection we could not account for mutant viruses, bacteria and lower fungi becoming resistant to once highly effective antibiotics.

The most enigmatic chapter in the history of molecular evolution involves the origin of the simplest single-cell structures and the subsequent formation of single-cell organisms. In order for life to commence, all the necessary components (proteins, carbohydrates, nucleic acids, fats, etc.) had to be accumulated in a confined phase of limited area. Scientists have devoted great efforts in modeling a variety of procellular structures based on this requirement: bounded droplets (automatically formed from polypeptide and polynucleotide solutions), proteinoid microspheres, foam bubbles, armored microspheres (grains of alumoferrosilicate coated with a lipoid layer) and so on. It has been demonstrated that such systems possess a number of features characteristic of primary cells. In particular, they are distinguished by their high stability, permeability, specific catalytic properties and, most importantly, their ability to grow (owing to their exchange of low

Table 12.5 *Geochronological scale of chemical evolution [adapted from Oparin, A. I.,* Problems of Appearance and Essence of Life *(in Russian), Nauka, Moscow, 1973, p. 171, Figure 6, with permission]*

Period of evolution	Age[a] ($\times 10^9$ years)	Geological era (duration $\times 10^9$ years)	Characteristics of stage (duration $\times 10^9$ years)
Origin of the solar system	5.0–6.0		
Formation of Earth	4.5–5.0		Beginning of molecular evolution on Earth
Formation of Earth's crust	4.5–4.0		Abiogenic synthesis of organic compounds (5.0–3.0)
Formation of secondary reductive atmosphere	4.0–3.5	Archaean (4.5–2.8)	Formation of ocean and primary 'broth' (4.0–3.0)
End of reductive and beginning of transitional atmosphere	3.5–2.0		Formation of protocytes and protoenzymes; development of primary nucleic code (3.0–2.0); anaerobic epoch begins
Formation of oxygen atmosphere	2.0–1.0	Proterozoic (2.7–0.7)	Start of photosynthesis; appearance of aerobes
Beginning of Cambrian era	1.0 to present time	Paleozoic (0.7–0.23)	Biogenesis, biological evolution
		Mesozoic (0.23–0.067)	Appearance of mammals
		Cenozoic (0.067 to present)	Appearance of anthropoids

[a]Age before present as estimated by radio-dating of uranium lead (scale up to 5×10^9 years).

molecular weight substances with the surrounding media), divide and degrade. The branch of modern science that aims on making a fully synthetic cell and ultimately engineering various artificial organisms is called synthetic biology. It is likely that the consequences of synthetic biology for the future of mankind could by far exceed those of nanotechnology.

In this chapter, we have traced the course of chemical evolution from the primeval universe of atoms through simple molecules to highly complex polymers. We have observed how molecular structures spontaneously, but with natural regularity, became more and more complicated and the manner in which the natural selection of structures occurred. We have focused on the natural selection of heterocyclic compounds which were the most feasible to perform biologically important roles. This self-improvement is the essence of chemical evolution, the inescapable, preprogrammed drive for life. As Albert Einstein wrote:

> 'Life is predetermined by the existence of atoms, and the mystery of all existence is contained in the very lowest step.'

The general scheme of chemical evolution and the corresponding time scale are depicted in Figure 12.27 and Table 12.5. The data indicate that Nature took $0.5 - 1.0 \times 10^9$ years to play the game of chemical evolution and to evolve to the simplest forms of life. This is not considered long-term on the geological scale of time. Indeed, imprints of algae biofossils 3.4×10^9 years old have been found on rock samples extracted in Swaziland (Southern Africa). Analogous natural antiquities in the schistose quartzites of Western Australia are dated at 3.5×10^9 years.

12.7 Problems

1. An alternative hypothesis for the origin of life on Earth is based on the assumption that chemical evolution (up to the development of peptides and oligonucleotides) occurred not in the primordial ocean but in the atmospheric clouds of the protoearth. Provide ideas to support this cloud droplet chemistry theory. (Hint: consider the availability of initial reactants, energy sources and conditions required to carry out the chemical transformations leading to the macromolecules, and see Figure 12.25.)
2. What inherent structural properties of hydrogen cyanide determined its key role in the abiotic origin of the biologically important heterocyclic systems?
3. Why freezing conditions could favor participation of hydrogen cyanide in prebiotic reactions?
4. In early experiments in which mixtures of H_2 (20%), CH_4 (40%), NH_3 (40%) and water vapor were subjected to electrical discharges, a number of organic products including HCN and amino acids were obtained. However, purines and pyrimidines were not identified in the final mixtures. Suggest an explanation.
5. Suggest a possible scheme for the abiotic formation of tryptophan from indole and xanthine from available precursors.
6. Recently, the dipeptide histidylhistidine (His-His) was prepared by a prebiotic synthetic method. Small His-containing peptides, and histidine itself, are assumed to promote some prebiotic reactions. Suggest reaction products and postulate a mechanism involving the catalytic action of the dipeptide His-His when an aqueous mixture (pH 8.0) of 0.1 M 2′-deoxyadenosine-5′-monophosphate (dAMP) and 0.02 M His-His is heated at 90 °C for 64 h in 10 h cycles to evaporation (conditions consistent with a primeval evaporating pond under tidal influences).
7. Suggest a scheme for the abiotic formation of porphobilinogen. What role may this substance have played in the emergence of life on Earth?
8. What factors may cause an elevated content of adenine in DNA and other important biomolecules?
9. In transfer RNAs the modified nucleobases, 2-thiocytosine and 2-thiouracil, are sometimes met. Suggest a possible prebiotic scheme for their formation.
10. What could be plausible phosphorylation agents on primitive Earth and how did they influence conditions of phosphorylation of nucleosides?
11. Recently, the coenzymes adenosine diphosphate glucose, guanosine diphosphate glucose and cytidine diphosphoethanolamine were synthesized by a nonenzymatic prebiological method designed to model the primitive metabolism in the first Archaean cells. Nucleotides of this type may be regarded as metabolic RNA fossils. What role might these coenzymatic molecules have played in the first living cells? (For hints see Chapters 4 and 5.)
12. What causes the activation of nucleotides on their conversion into phosphorimidazolides (Figure 12.22)?

12.8 Suggested Reading

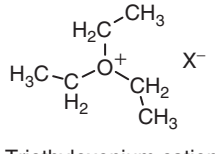
1. Zaikowski, L. and Friedrich, J. (eds), *Chemical Evolution across Space and Time. From the Big Bang to Prebiotic Chemistry*, Oxford University Press, Oxford, 2008.
2. Mason, S. F., *Chemical Evolution. Origin of the Elements, Molecules, and Living Systems*, Clarendon Press, Oxford, 1991.

3. Eschenmoser, A. and Loewenthal, E., Chemistry of potentially prebiological natural products, *Chem. Soc. Rev*., 1992, **21**, 1.
4. Altman, S., Enzymatic cleavage of RNA by RNA (Nobel lecture), *Angew. Chem., Int. Ed. Engl*., 1990, **29**, 749.
5. Cech, T. R., Self-splicing and enzymatic activity of an intervening sequence RNA from *Tetrahymena* (Nobel lecture), *Angew. Chem., Int. Ed. Engl*., 1990, **29**, 759.
6. Orgel, L.E., Prebiotic chemistry and the origin of the RNA world, *Crit. Rev. Biochem. Mol. Biol*., 2004, **39**, 99.
7. Fox, S. W., *The Emergence of Life: Darwinian Evolution from the Inside*, Basic Books, New York, 1988.
8. Dyson, F., *Origins of Life*, Cambridge University Press, Cambridge, 1999.
9. Maynard, S. J., and Szathmary, E., *The Origins of Life*, Oxford University Press, New York, 1999.
10. Zubay, G., *Origins of Life on the Earth and in the Cosmos*, Academic Press, New York, 2000.
11. Luisi, P. L., *The Emergence of Life: From Chemical Origins to Synthetic Biology*, Cambridge University Press, Cambridge, 2006.
12. Wagnière, G. H., *On Chirality and the Universal Asymmetry*, Wiley-VCH, Zürich, 2007.
13. Emeline, A. V., Otroshchenko, V. A., Ryabchuk, V. K. and Serpone, N., Abiogenesis and photostimulated heterogeneous reactions in the interstellar medium and on primitive Earth: relevance to the genesis of life, *J. Photochem. Photobiol. C: Photochem. Rev*., 2003, **3**, 203.
14. Calvin, M., *Chemical Evolution: Molecular Evolution Towards the Origin of Living Systems on the Earth and Elsewhere*, Oxford University Press, New York, 1969.
15. Bernal, J. D., *The Origin of Life*, World Publishing, Cleveland, OH, 1967.
16. Oparin, A. I., *The Origin of Life on the Earth*, 3rd edn, Academic Press, New York, 1957.

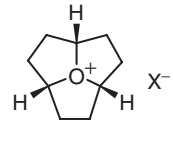
Conclusion

Long before the first living creatures, heterocycles arose on our planet, their appearance predetermined by the fundamental laws of chemical evolution. Together with other classes of organic compounds, heterocycles promoted the formation of life on Earth. We can now synthesize such life-based materials and many others not occurring in nature, and science has taught us how to use heterocycles to improve the quality of human life and to explore the secrets of nature.

Why is it appropriate to emphasize specifically the role of heterocycles? Analogies to the roles of other classes of organic compounds are easily found. In fact, dyes, luminophores, pesticides and drugs do not have to be heterocyclic in structure. In a similar fashion there are many common features in chemistry and physics between such related compounds as pyrrole and aniline, or between pyridine and nitrobenzene. However, the introduction of a heteroatom into a carbocyclic ring imparts fundamentally new properties. A typical example is the dramatic differences between trialkyloxonium (Meerwein) salts and their oxatriquinane heterocyclic counterpart. While the former are exceedingly sensitive to moisture, oxatriquinane cations resist water and can be even recrystallized from it. Heterocycles are chemically more flexible and structurally more rigid to respond to the many demands of biochemical systems. This is why nature selected the heterocycles pyrrole and pyridine, and not the homocycles aniline and nitrobenzene, as the basis of so many essential biological systems.



Triethyloxonium cation
(water sensitive)



Oxatriquinane
(water resistant)

This book was written with the prime intention of drawing attention to this specific nature of heterocycles. New types of heterocyclic structures and methods of their synthesis are continuously being discovered and without doubt the chemistry of heterocycles will continue to progress. Heterocyclic compounds are now deeply involved in many modern technologies: biotechnology, medicine, nanotechnology, sensorics, informatics, new materials as well as energetics and ecology.

At the conclusion of this book we wish to look back to the beginning and to unite two different themes: one mentioned in the preamble to the first chapter, and the other here, in an allegory of the Russian poet-symbolist Zinaida Hippius:

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

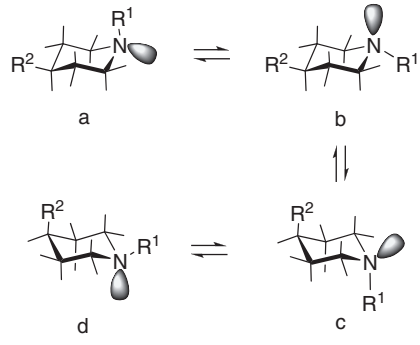
‘For a long time, of her, we were singing praises,
Of the pretender queen.
The whisps of fragrant fumes are still being felt in hazes,
The twinkling shrine still seen.’

The authors emphasize that in the creation of this book they have learned much about the role of heterocyclic chemistry in life and society. They hope that the reader has been likewise informed. Undergraduate, graduate and postdoctoral students are invited to consolidate and expand their knowledge of organic chemistry and the applications of heterocycles in the Problem sections at the end of each chapter.

Answers and References to Selected Problems

Chapter 1

1. In a 1,4-disubstituted piperidine, four equilibrating conformers, that is, (a–d), are possible (hydrogen atoms are omitted):



   The equilibrium proportions of such isomers depend on the nature of the substituents and their positions in relation to the nitrogen atom (Eliel, E. L., Kandasamy, D., Yen, C.-Y. and Hargrave, K. D., *J. Am. Chem. Soc.*, 1980, **102**, 3698).
2. 4-Hydroxy-1-methylpiperidine can exist in a boat conformation owing to intramolecular hydrogen bonding. The boat conformation is also fixed by the existence of a bridge (e.g., CH_2CH_2) between the 1-position and 4-position.
3. See: Lambert, J. B., Oliver, W. L. and Jackson, G. F., *Tetrahedron Lett.*, 1969, 2027.
4. A, C, E and F.
5. See: Baudler, M., Akpapoglou, S., Ouzounis, D., *et al., Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 280.
6. Excluding different tautomeric forms, there are four isomers. Two contain nitrogen at the bridgehead.
7. Owing to the flexibility of their rings, azaannulenes can exist in various conformations with the nitrogen lone electron pair or N—H bond oriented either inside or outside the cycle depending on the 'internal' or 'external' disposition of the nitrogen atom.

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications,
Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.
 Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9

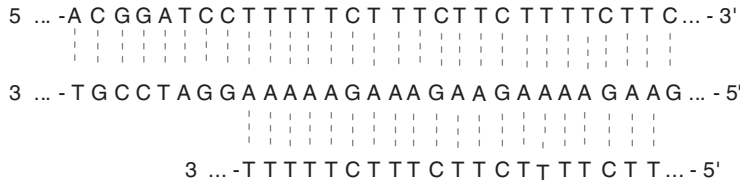
8. This sequence results from the relative electronegativity of heteroatoms. Oxygen has the largest electron affinity and donates its free electron pair into aromatic π-system with greater difficulty than nitrogen and sulfur atoms.
9. A nitrogen or boron atom can be placed into an azafullerenyl radical (Figure 1.7) instead of a carbon atom carrying an unpaired electron.

Chapter 2

1. Piperidine is more basic than pyrrole owing to participation of the lone electron pair of the pyrrole nitrogen in the formation of the pyrrole aromatic π-sextet. The higher basicity of piperidine compared to pyridine is accounted for by the different hybridizations of their nitrogen atoms: sp^3-hybridization in piperidine and sp^2 in pyridine. Orbitals of greater *s*-character display higher electronegativity and therefore lower affinity toward a proton.
2. The pK_a of 12.9 corresponds to ionization of the N—H bond in neutral benzimidazole to give the benzimidazole anion. The pK_a of 5.3 is for ionization of the benzimidazolium cation to give neutral benzimidazole.
3. p$K_a' = 2.39$ (ionization of the N—H bond in the purine cation), p$K_a'' = 8.93$ (ionization of the N—H bond in neutral purine).
4. Basicities of anions: imidazole> benzimidazole> purine> tetrazole.
5. The nitration of imidazole in a strongly acidic medium occurs on the relatively inert imidazolium cation, whereas bromination in organic solvents occurs via the much more active neutral molecule.
7. (a) Nitration occurs at the free α-position of the pyrrole ring. (b) The *N*-ethylpyridinium salt is formed, which upon reduction gives the corresponding 1,4-dihydropyridine derivative. (c) The α-position of the furan cycle is exclusively nitrated. (d) A mixture of 1-methyl- and 2-methyl-1,2,3-triazole is formed. (e) The chlorine atom in the pyrimidine nucleus is more active toward substitution and is replaced by a methoxy group.
8. The dipole moment is a vector value. The vector is considered to be oriented toward the negative pole. In pyridine, the vector sum is obviously directed from the center of the ring toward the nitrogen atom. By adding such vectors oriented at 60° (in pyridazine), 120° (in pyrimidine) and 180° (in pyrazine) one can assess quite accurately the relative dipole moments for these diazines: pyridazine> pyrimidine> pyrazine.
9. See: Alcalde, E., Dinares, I., Fayet, J.-P., Vertut, M.-C. and Elguero, J., *J. Chem. Soc., Chem. Commun.*, 1986, 734.
11. See: Pfleiderer, W., in *Physical Methods in Heterocyclic Chemisty*, (ed. A. R. Katritzky), Academic Press, New York, 1963, Chap. 4.
12. See: Eisch, J. J. and Jaselskis, B., *J. Org. Chem.*, 1963, **28**, 2865.
13. See: Staab, H. A., *Angew. Chem., Int. Ed. Engl.*, 1962, **1**, 351.
14. J, K, L and N.
15. See: Kimoto, H., Fojii, S. and Cohen, L.A., *J. Org. Chem*., 1982, **47**, 2867.
16. One can suggest that at rotation of rings in structures (b) and (c) a substantial barrier exists caused by repulsion of negatively charged nitrogen atoms when they are in a *cis*-orientation relative to each other.

Chapter 3

1. (a) Poly(ThrValLeuTyrCys). (b) The dipeptide LeuAsp.

2. (a) There are 561 codons and 561 residues. (b) In the first chain [A] = 0.35, [G] = 0.29 and [T + C] = 0.36. In the complementary chain [T] = [0.35], [C] = 0.29 and [A + G] = 0.36. See: Regier, J. C. and Pacholski, P., *Proc. Natl Acad. Sci. USA*, 1985, **82**, 6035.
3. Campbell, J. A., *J. Chem. Educ.*, 1976, **53**, 447, Problem Q248.
4. (a) 5′-TCGAGTAGCCGATGATCATCGTCGACGAT-3′; (b) 5′-UCGAGUAGCCGAUGAUCAUCGUCGACGAU-3′; and (c) Ser-Ser-Ser-Arg, Ser-Ser-Ser-Thr (adapted from Lehninger, A. L., *Biochemistry*, Worth, New York, 1970, Chap. 31, p. 727, Problem 6, with permission).
5. See: Huang, H., Solomon, M. S. and Hopkins, P. B., *J. Am. Chem. Soc.*, 1992, **114**, 9240.
6. Two (adapted from Ternay, A. L. Jr, *Contemporary Organic Chemistry*, 2nd edn, Saunders, Philadelphia, 1979, with permission).
7. Hydroxyproline and hydroxylysine are derivatives of the coded proline and lysine; it is known that these two coded amino acids are first incorporated into the polypeptide chain and are subsequently enzymatically hydroxylated.
8. 

9. (a) Glycine and alanine; total 68%. (b) AGR, GGL, GGQ, GAG, AAAAAA and GGAGQG-GYGGXQG. (c) Guanine and cytosine; 82%.
10. As illustrated in Figure 3.17a, thymine easily undergoes photodimerization, which causes severe DNA mutations. Due to strong ultraviolet radition in high attitudes living organisms with an ordinary content of thymine could not survive.
11. The GC aggregate is stronger because it involves three hydrogen bonds unlike two in the AT pair. Since the DNA promoter region fulfills a catalytic function, its structure should be relatively flexible and that is possible when the AT content is higher.

Chapter 4

1. See: Pandit, U. K. and Mas Cabre, F. R., *J. Chem. Soc., Chem. Commun.*, 1971, 552.
2. See: Fukuzumi, S., Kuroda, S., Goto, T., Ishikawa, K. and Tanaka, T., *J. Chem. Soc., Perkin Trans. 2*, 1989, 1047.
4. The lower affinity afforded by the tryptophan cation-π interaction may help the copper trafficking protein to release its metal cargo more easily.
8. The adenosyl moiety is found in NAD^+, $NADP^+$, FAD, CoA, ATP, ADP and so on.
9. See: Itoh, S., Ogino, M., Fukui, Y., Murao, H., Komatsu, M., Ohshiro, Y., Inoue, T., Kai, Y. and Kasai, N., *J. Am. Chem. Soc.*, 1993, **115**, 9960.
10. The rate would increase approximately 10-fold (Lewis, C., Kramer, T., Robinson, S. and Hilvert, D., *Science*, 1991, **253**, 1019). In protic (aqueous) solvents the substrate is highly stabilized through hydrogen bonding. Aprotic solvents force charge delocalization, stabilize the transition state by dispersion interactions, facilitate passage of the substrate from solution into the predominantly hydrophobic active site of the decarboxylase and thus accelerate decarboxylation of the carboxylate anion. X-Ray analysis has shown that the carboxylate-binding

site is highly hydrophobic in the case of histidine decarboxylase (Gallagher, T., Snell, E. E. and Hackert, M. L., *J. Biol. Chem.*, 1989, **264**, 12 737).

11. (a) Histidine. (b, c) Ovothiols act as antioxidants by enzymatic consumption of H_2O_2. During this process the ovothiols are oxidized to the disulfides. Enzymatic reduction of the disulfides by NADP-H regenerates the ovothiols (Shapiro, B. M., *Science*, 1991, **252**, 533).
12. A key role is played by the relative basicity of guanine anion and neutral adenine molecule differing by about million times in favor of the former (pKa ~9.3 and 3.45, respectively). Due to this an adenine residue placed in ribozyme instead of G** can not perform basic catalysis as depicted in Figure 4.43a.

Chapter 5

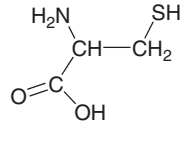
1. N—P bond in A, N—COMe bond in B, N—NO_2 bond in C, C—COMe bond in D, C—CHO bond in E.
4. An explanation of this illusive paradox lies in multiple recycling of ATP. The total amount of ATP + ADP in the human body remains nearly constant. Therefore, one can easily calculate that each ATP molecule is recycled 1000–1500 times during a single day.
8. 11.1 kcal mol^{-1} (adapted from Lehninger, A. L., *Biochemistry*, Worth, New York, 1970, Chap. 14, p. 311, Problem 4, with permission).
10. There are about 40 ml of ethanol in 100 ml of vodka. Oxidation of this quantity of ethanol in the Krebs cycle and in the respiratory chain provides 8.4 moles of ATP, corresponding to around 76 kcal mol^{-1}.

Chapter 6

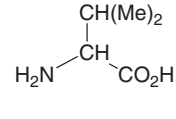
7. (a) Nine, (b) seven, (c) five molecules of ATP, respectively (adapted from Lehninger, A. L., *Biochemistry*, Worth, New York, 1970, Chap. 21, p. 480, Problem 1, with permission).
8. 10.6 kcal mol^{-1}.
9. (a) GSPB normally inhabit the zones near hot hydrothermal vents. One can suggest that the latter emit geothermal light at the very edge of the red spectrum. (b) Obviously, a great lack of photons at these depths is the main factors determing the large size of GSPB and the absence of protein matrix in their structure.

Chapter 7

1. See: Campbell, J. A., *J. Chem. Educ.*, 1977, **54**, 309, Problem Q302.
3.



Cysteine

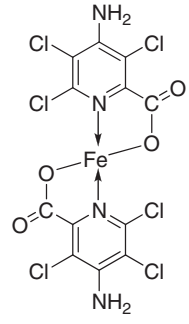


Valine

4. See: Campbell, J. A., *J. Chem. Educ.*, 1977, **54**, 369, Problem Q306.
5. The structural similarity of allopurinol to the purines suggests that it functions as an antimetabolite. Allopurinol supposedly inhibits the enzyme xanthine oxidase which is involved in the metabolism of purines into uric acid.
6. See: Campbell, J. A., *J. Chem. Educ.*, 1977, **54**, 247, Problem Q294.
7. (a) γ-Aminobutyric acid (Figure 7.26). (b) Between 0.6 nm (in the extended conformation) and 0.45 nm (in the folded conformation) in the zwitterionic form of γ-aminobutyric acid; in C and E between 0.55 and 0.59 nm; in D 0.44 nm; in G 0.56 nm (Sytinsky, I. A., Soldatenkov, A. T. and Lajtha, A., *Prog. Neurobiol.*, 1987, **10**, 89).

Chapter 8

2.

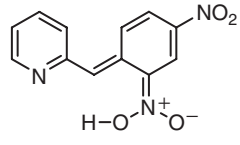


7. (a) *Cis–trans* isomerism around the external C—C bond and optical isomerism due to the presence of an asymmetric carbon atom give rise to a total of four isomers. (b) Intramolecular hydrogen bonding between the OH group and the N-2 atom of the triazole ring is possible. (c) The N-4 atom owing to its sterically uncrowded position and highest nucleophilicity (Katagi, T., *J. Agric. Food Chem.*, 1988, **36**, 344).

Chapter 9

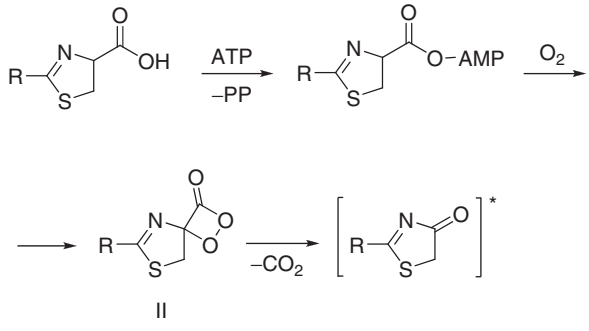
2. Conjugation in indigoid dyes decreases the order of the C—C bond existing between the two nuclei, thereby facilitating rotation about the bond and, consequently, interconversion between the *cis* and *trans* forms. However, in indigo itself the *trans* form is additionally stabilized by intramolecular hydrogen bonds, making conversion to the *cis* form significantly more difficult.
3. Indigoid dyes are rather weak bases and NH acids and are ionized only under the action of strong acids and alkalis. Thus, concentrated sulfuric acid causes protonation at the C═O group, enhancing its electron-accepting properties and making conjugation more effective. The degree of conjugation is also increased when the *N*-anion is formed under the action of sodium *tert*-butoxide.
4. The conjugated 18π-electron system, including the four nitrogen atoms and internal C—C and C═N bonds, is believed to be the main chromophore in the porphyrin system which provides annular conjugation and coloration. The chromophore does not appear to be affected by the hydrogenation of the external C═C bonds.

5. Biliverdin molecule has a more extended π-conjugated system including four pyrrole rings. In contrast, in bilirubin molecule the conjugated chain is interrupted by a CH_2 group which connects two internal pyrrole rings.
10. Photochromism in compounds A and B is due to migration of a proton from the CH_2 group to the *ortho*-nitro group with formation of a colored photoexcited isomer (e.g., I, which is blue in contrast to the tan isomer A). This proton transfer probably occurs as an intramolecular process, which is impossible in the case of compound C. However, the presence of the *para*-nitro group is also important since it enhances the acidity of the CH_2 protons (Gilfillan, E. D. and Pelter, M. W., *J. Chem. Educ.*, 1994, **71**, A4; and Prostakov, N. S., Krapivko, A. P., Soldatenkov, A. T., Furnaris, K., Savina, A. A. and Zvolinskii, V. P., *Chem. Heterocycl. Compd*, 1976, **16**, 312).



I

12. It is assumed that ATP first activates the carboxylic group in the firefly luminophore (*cf*. Figure 5.2) that facilitates the formation of dioxetanone derivative II loosing CO_2 molecule and emitting light:



II

R = 6-hydroxybenzothiazolyl-2

14. Luminophores emitting light in the infrared region are less damaging for living tissues than dyes emitting in the visible part of electromagnetic spectrum.

Chapter 10

1. Hedge, V., Hung, C.-Y., Madhukar, P., *et al., J. Am. Chem. Soc.*, 1993, **115**, 872.
2. Sessler, J. L., Mody, T. D. and Lynch, V., *J. Am. Chem. Soc.*, 1993, **115**, 3346.
3. Hamilton, A. D., *J. Chem. Educ.*, 1990, **67**, 821.
4. Gokel, G. W. and Cram, D. J., *J. Chem. Soc., Chem. Commun.*, 1973, **481**.
5. Busch, D. H., *Chem. Rev.*, 1993, **93**, 847.
6. Tarrago, G., Marzin, C., Najimi, O. and Pellegrin, V., *J. Org. Chem.*, 1990, **55**, 420.

7. Hu, J., Barbour, L. J. and Gokel, G. W., *J. Am. Chem. Soc.*, 2002, **124**, 10 940.
8. Takemura, H., Kon, N., Yasutake, M., Kariyazono, H., Shinmyozu, T. and Inazu, T., *Angew. Chem., Int. Ed*., 1999, **38**, 959.
9. Miyahara, Y., Goto, K. and Inazu T., *Tetrahedron Lett*., 2001, **42**, 3097.
10. Gilard, R. E., Stoddart, J. F., White, A. J. P., Williams, B. J. and Williams, D. J., *J. Org. Chem*., 1996, **61**, 4504.

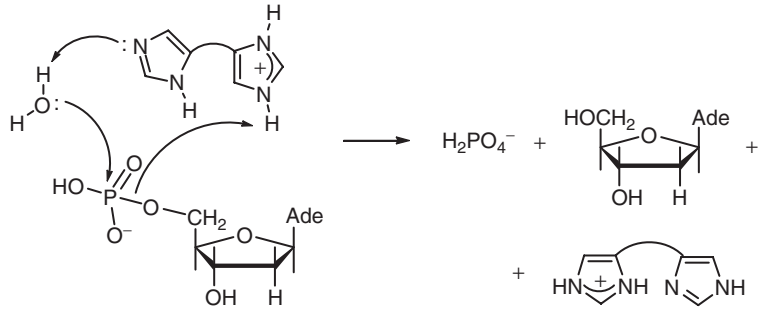
Chapter 11

1. By selective oxidation 5-hydroxymethylfurfural can be converted into 2,5-furandicarboxylic acid, which can be used as a replacement for terephthalic acid in the production of polyethyleneteraphthalate. The reduction of HMF can lead to 2,5-dihydroxymethylfuran or 2,5-bis(hydroxymethyl)tetrahydrofuran, which can serve as the alcohol component in the production of polyesters.
2. Borra, E. F., Seddiki, O., Angel, R., *et al., Nature*, 2007, **447**, 979.
3. 1792 liters.
4. Yano, Y., Kimura, M., Shimaoka, K., *et al., J. Chem. Soc., Chem. Commun*., 1986, 160.
5. It is believed that stacking of the dyes between nucleic acid base pairs and their hydrophobic surrounding therein are responsible for this phenomenon. By moving into an hydrophobic environment and away from water, the dyes are forced to shed any water molecules that were associated with it. As water is an efficient fluorescent quencher, the removal of these water molecules allows the dyes to fluoresce.
6. Ozlem, S. and Akkaya, U., *J. Am. Chem. Soc*., 2009, **131**, 48.
7. Aumüller, A., Hädicke, E., Hünig, S., *et al., Angew. Chem., Int. Ed. Engl.*, 1984, **23**, 449.
8. Balzani, V., Credi, A., Mattersteig, G., *et al., J. Org. Chem*., 2000, **65**, 1924.
9. Due to the presence of carboxylic functions dye-sensitizer molecules are more strongly retained by metal oxide nanoparticles.
11. Connections between B and C units and between D and E units are controlled by acid/base and redox inputs, respectively.

Chapter 12

1. Prebiotic reactants such as CH_4, H_2O, NH_3 and N_2 were present in the atmosphere and simple, open chain and heterocyclic organic molecules might arise from such precursors under the action of UV light, corona discharge, lightning or heating from asteroid impacts. These molecules may have been prevented from falling to the Earth's surface by atmospheric moisture adhesion. Being entrapped by cloud droplets which contain catalytic clay particles, these molecules could yield biologically important macromolecules by polymerization and multiple evaporation–condensation processes. Oligomerization and polymerization of amino acids, purine bases and pyrimidine bases could have occurred via dehydration which would have been more difficult in the ocean than in the atmospheric droplets. Complex molecules might also be supplied to the Earth by comets and meteorites.
3. There are two possible reasons: (i) HCN is a volatile compound and a lower temperature caused its accumulation on the Earth's surface, (ii) crystallization of water ejected cyanides from growing crystals and they separated in microscopic liquid pockets.
4. The likely cause is the tremendous excess of hydrogen in the initial mixtures. Thermodynamic calculations show that the formation of heterocyclic molecules is unlikely in highly reductive mixtures mimicking the protoearth's atmosphere.

6. Here dAMP is hydrolytically dephosphorylated. The reaction proceeds via a proton transfer mechanism (concerted general acid–base catalysis by the dipeptide):



9. Interaction of cyanoacetaldehyde and thiourea at heating is known to give thiocytosine from which thiouracil is formed by hydrolysis. Thiourea is a potentially prebiotic compound since it can be produced from NH_4CNS or by reaction of H_2S with cyanamide.
11. The first two compounds might be important in: (i) the activation of such nutrients as carbohydrates in the primordial 'broth' by means of phosphorylation, (ii) the extraction of energy from fuel molecules and the simultaneous storage in the energy-rich bonds of these high energy phosphorylated compounds and (iii) the synthesis of biochemical building blocks. The third cofactor might be useful in the transfer of ethanolamine to diacylglycerol in the biosynthesis of phospholipids necessary for simple membrane formation and in providing the primeval cells with individual identities (Mar, A. and Oro, J., *J. Mol. Evol.*, 1991, **32**, 201).
12. Both the pyrrolic nitrogen atom and the phosphorus atom in phosphorimidazolides carry partial positive charge and therefore the N—P bond is energy-rich (Section 5.1). Its hydrolysis or alcoholysis eliminates electrostatic repulsion, thus releasing energy.

Index

Heterocycles in Life and Society: An Introduction to Heterocyclic Chemistry, Biochemistry and Applications, Second Edition. Alexander F. Pozharskii, Anatoly T. Soldatenkov and Alan R. Katritzky.

Published 2011 by John Wiley & Sons, Ltd. ISBN: 978-0-470-71411-9