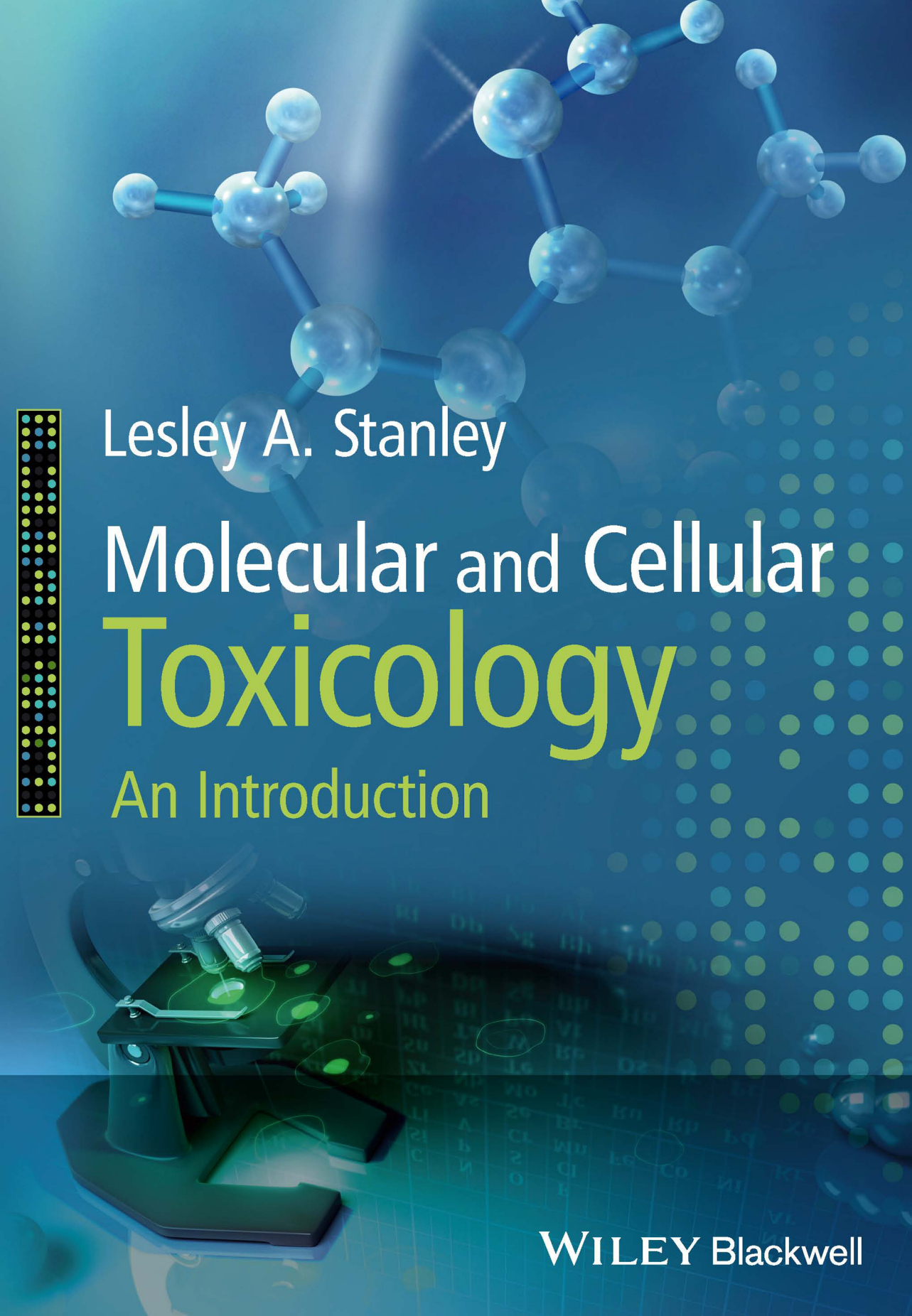

Lesley A. Stanley
Molecular and Cellular
Toxicology
An Introduction
WILEY Blackwell

Molecular and Cellular Toxicology

Molecular and Cellular Toxicology

An Introduction

Dr Lesley A. Stanley
Consultant in Investigative Toxicology, UK

WILEY Blackwell

This edition first published 2014 © 2014 by John Wiley & Sons, Ltd

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

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Stanley, Lesley A., author.
Molecular and cellular toxicology : an introduction / Dr. Lesley A. Stanley.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-119-95207-7 (cloth) – ISBN 978-1-119-95206-0 (paper)
I. Title.
[DNLM: 1. Toxicity Tests – methods. 2. Molecular Biology. 3. Xenobiotics – toxicity. QV 602]
RA1199.4.A38
615.9′07 – dc23
2013048370

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

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Set in 10/12pt TimesTenRoman by Laserwords Private Limited, Chennai, India

1 2014

Dedication

Dedicated to the memory of Elizabeth Stanley and Margaret Orr,
two wonderful mothers

Contents

Foreword

by Dr Marilyn J. Aardema
Marilyn Aardema Consulting, LLC

Technical advances and bold initiatives like the National Academy of Science's *Toxicology in the 21st Century*, along with socio-political pressures such as the 3Rs (replace, reduce and refine the use of animals in experiments) have led to remarkable changes in the field of toxicology over the past several years. Toxicology is undergoing a major shift towards assessing and understanding damage at the cellular and tissue level. This comprehensive, well-written book focusses on the timely topic of current advances in the field of molecular and cellular toxicology. The book starts with a review of how cells and tissues respond to damage and the consequences of damage that overwhelm normal cellular protective mechanisms. With this background, new technologies for evaluating cellular and tissue damage along with investigating the toxicological outcomes are described. This includes the use of 'omics technologies (transcriptomics (changes in RNA), proteomics (changes in proteins)), metabolomics (changes in products of metabolism; Chapter 3), the use of 3D tissue models to obtain a more biologically relevant assessment of toxicity (Chapters 4, 8) and the use of *in silico* approaches for predicting toxicological effects (Chapter 6). The final Chapter 10 provides a glimpse forward at emerging technologies that are sure to impact the field of molecular and cellular toxicology further in the years ahead. These and the other chapters in the book not only provide essential reading on recent technology developments, but also provide up-to-date information on the drivers behind these advances, and the global efforts towards validation and incorporation of new approaches into the toxicology paradigm.

This book will be invaluable to all those interested in the latest advances in toxicology including postgraduate/graduate life science students interested in toxicology as well as individuals starting out in the field of cosmetics, consumer products, pharmaceutical and testing industries who need knowledge of current approaches in toxicology. I commend the editors and author, Dr Lesley Stanley, on this valuable contribution to the field of Toxicology.

Preface

Over the past 10 years the subject of toxicology has changed dramatically, moving from a discipline which was once firmly wedded to traditional (some might say old-fashioned) methods to one which is keen to embrace the innovative techniques emerging from the developing fields of cell culture and molecular biology. Over the same period our ability to predict outcomes using computer models has also progressed to an astonishing degree.

The availability of novel methods has had a great deal of influence on the development of new approaches in toxicology; another key impetus has been the need to reduce the use of animals in experimentation and testing. This was originally driven by public distaste (even revulsion) for the practice. The scientific approach has been codified in terms of the so-called '3Rs' (Reduction, Refinement and Replacement) and is now enshrined in legislation, particularly the 7th Amendment to the European Union Cosmetics Directive and the REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) regulations.

The 7th Amendment prohibits the marketing within the EU of cosmetics containing ingredients which have been tested on animals. This has made the development of 'alternative' (non-animal and computer-based) methods a key priority, particularly for the cosmetics industry. At the same time, the pharmaceutical industry has been pushed in the same direction by the need for cost-effective ways of screening large numbers of potential drug molecules to identify those with the elusive property of 'druggability' (the rare combination of efficacy and pharmacological properties which allows oral dosing one or a few times per day without accumulation to toxic levels). With these factors in mind, I hope that this book will be useful to individuals at the start of their careers in these industries.

The REACH initiative was first suggested in an EC White Paper released in 2001. A formal legislative proposal was issued in 2003 and the regulations came into force on 1 June 2007. REACH requires companies which manufacture or import chemical substances in quantities greater than 1 tonne to assess their adverse effects in terms of toxicology and consequences of release into the environment. It differs from previous legislation in that it makes the industry responsible for managing the risks associated with the chemicals it uses. Another

important feature of REACH is that it explicitly states that, where possible, animal tests should be replaced by alternatives such as *in vitro* tests or computer predictions.

The aim of this book is to introduce recent developments in the fields of molecular and cellular toxicology to an audience of life scientists at the final year undergraduate and early postgraduate levels. The wide ranging nature of the subject and the rapid rate of progress, however, mean that it cannot be comprehensive; instead, what has been attempted is a broad-brush sketch of the landscape illustrated with examples which highlight the key points and interesting developments.

It will be evident from the wide range of topics mentioned even in this brief preface that, in order to keep the book concise and readable (as well as affordable), a certain amount of background knowledge has had to be assumed. In particular, the reader will need to have a solid foundation in the following:

- Basic biochemistry, molecular and cellular biology
- Conventional laboratory methods using DNA, RNA and protein
- Tissue structure and histology
- Pathological processes and their histopathological consequences

An elementary understanding of the metabolism and toxic effects of chemicals would be helpful but is not essential, since the book aims to cover the relevant processes.

For those who do need to look up some background information on these topics, I recommend the following textbooks:

References

Alberts, B. *et al*. (2008) *Molecular Biology of the Cell*, 5th edn, Garland Science, Taylor and Francis Group, LLC, New York. ISBN: 978-0-8153-4105-5

Berg, J.M. *et al*. (2006) *Biochemistry*, 6th edn, W.H. Freeman and Co. Ltd. ISBN: 978-0-7167-8724-2

Bolsover, S.R. *et al*. (2011) *Cell Biology*: *A Short Course*, 3rd edn, Wiley-Blackwell. ISBN: 978-0-470-52699-6

Male, D. *et al*. (2012) *Immunology*, 8th edn, Elsevier-Saunders, Philadelphia. ISBN: 0323080588

Peckham M. (2011) *Histology at a Glance* Wiley-Blackwell. ISBN: 978-1-4443-3332-9.

Stevens, A. *et al*. (2008) *Core Pathology*, 3rd edn, Mosby-Elsevier. ISBN: 978-0-7234-3444-3

Acknowledgements

My Editor, Nicky McGirr, has been a source of inspiration throughout the course of this project. I am immensely grateful to her for her excellent advice, and particularly for her unflagging enthusiasm which kept me going through all the difficult bits.

I would like to thank a number of people who were kind enough to let me use some of their material in this book. Dr Elaine Johnstone (Department of Oncology, University of Oxford) gave me permission to use material from one of her lectures on pharmacogenetics and genome wide association studies in Chapter 3; part of the text upon which Chapters 4 and 6 are based was provided by Dr Paul Brantom (Brantom Risk Assessment Ltd) with permission from the cosmetics industry association Colipa; and Dr Gill Clare (Independent Consultant on Genetic Toxicology) provided invaluable material and advice for Chapter 8. In addition, many people (some of whom do not even know me) generously allowed me to use their illustrations; they are too numerous to list individually, but would like to record my thanks to all of them.

I am also grateful to Mrs Roberta Logan and Drs Eian Massey, Robin Whelpton and Gary Hutchison, all of whom provided constructive comments on various versions of the manuscript during its preparation.

This project could not have been completed without the encouragement of many of my friends. In particular, I would like to thank Sarah Young for letting me work in her house while the builders were in mine and Julie McDowell for being my gym buddy.

Finally, it is my duty and pleasure to thank my husband, Nigel Orr, for his unfailing support, his infinite tolerance and for never being without a secret supply of chocolate.

Abbreviations

3Rs	replacement, refinement and reduction (of the use of animals in research)
4-ABP	4-aminobiphenyl
AAF	acetylaminofluorene
ADME	Absorption, Distribution, Metabolism and Excretion
AFB1	aflatoxin B1
AhR	arylhydrocarbon receptor
ALT	alanine aminotransferase
ASO	allele-specific oligonucleotide
ASPCR	allele-specific polymerase chain reaction
AST	aspartate aminotransferase
ATP	adenosine triphosphate
AUC	area under the plasma concentration–time curve
BAC	bacterial artificial chromosome
BBB	blood-brain barrier
BMD	benchmark dose
bp	base pair
BrdU	bromodeoxyuridine
CAR	constitutive androstane receptor
cdk	cyclin-dependent kinase
cDNA	copy DNA
CEBS	Chemical Effects on Biological Systems
ChIP	chromatin immunoprecipitation
CHO	Chinese hamster ovary
CIN	cervical intraepithelial neoplasia
CITCO	6-(4-chlorophenyl)-imidazo[2,1-b]thiazole-5-carbaldehyde
CL_{INT}	intrinsic clearance
C_{MAX}	maximum (plasma) concentration
CNS	central nervous system
COMET	Consortium for Metabonomic Toxicology
CPMP	European Committee for Proprietary Medical Products
CYP	cytochrome P450

DDI	drug–drug interaction
DEHP	diethylhexylphthalate
DEN	diethylnitrosamine
DILI	drug-induced liver injury
DMBA	7,12-dimethylbenz(a)anthrancene
DMN	dimethylnitrosamine
DMSO	dimethyl sulphoxide
EC_{50}	concentration giving 50% of maximal effect
ECHA	European Chemical Agency
ECVAM	European Centre for the Validation of Alternative Methods
EFSA	European Food Safety Authority
EMA	European Medicines Agency
ENU	ethylnitrosourea
EPA	US Environmental Protection Agency
ESC	embryonic stem cell
EST	embryonic stem cell test
EU	European Union
FABP	fatty acid binding protein
FDA	US Food and Drug Administration
floxed	flanked with loxP sites
GC	gas chromatography
GFP	green fluorescent protein
γGT	gamma glutamyl transpeptidase
GI	gastrointestinal
GLP	Good Laboratory Practice
GSH	glutathione
GST	glutathione *S*-transferase
GTP	guanosine triphosphate
GWAS	genome-wide association study
HCC	hepatocellular carcinoma
hERG	human ether-a-go-go related gene
HO-1	haem oxygenase 1
HPRT	hypoxanthine phosphoribosyltransferase
HRN™	Hepatic Reductase Null™
HTS	high throughput screening
i.p.	intraperitoneal
i.v.	intravenous
IC_{50}	concentration giving 50% inhibition
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
ILSI	International Life Sciences Institute
iPSC	induced pluripotent stem cell
IVIVE	*in vitro–in vivo* extrapolation
kb	kilobase
kD	kilo Dalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
K_M	Michaelis constant

K_{OW}	octanol-water partition coefficient
LC	liquid chromatography
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LD_{50}	dose giving 50% lethality
LDH	lactate dehydrogenase
LOAEL	lowest observed adverse effect level
LPS	lipopolysaccharide
MALDI	matrix-assisted laser desorption ionisation
MAPK	mitogen-activated protein kinase
MDCK	Madin-Darby canine kidney
MDR	multidrug resistance protein
MHLW	Japanese Ministry of Health, Labour and Welfare
MHRA	Medicines and Healthcare Products Regulatory Agency
MIAME	Minimum Information About a Microarray Experiment
MIAPE	Minimum Information About a Proteomics Experiment
MNU	methylnitrosourea
MOE	margin of exposure
mRNA	messenger RNA
MRP	multi-drug resistance-associated protein
MS	mass spectrometry
MTD	maximum tolerated dose
NADPH	nicotinamide adenine dinucleotide phosphate
NAT	*N*-acetyltransferase
NHS	UK National Health Service
NIEHS	US National Institute of Environmental Health Sciences
NIH	US National Institutes of Health
NMR	nuclear magnetic resonance
NOAEL	no observed adverse effect level
NTP	US National Toxicology Programme
OECD	Organisation for Economic Co-operation and Development
p.o.	perioral
PAH	polycyclic aromatic hydrocarbon
PAMPA	Passive Artificial Membrane Permeability Assay
P_{app}	apparent permeability
PB	phenobarbital
PBBK	physiologically based biokinetic
PBPK	physiologically based pharmacokinetic
PBTK	physiologically based toxicokinetic
PCN	pregnenlonone 16α-carbonitrile
PCR	polymerase chain reaction
PhIP	2-amino-1-methyl-6-phenylimidazo-[4,5-b]pyridine
pK_a	acid dissociation constant
PPARα	peroxisome proliferater activated receptor α
PPD	*p*-phenylene diamine
PXR	pregnane X-receptor
QA	quality assurance
QC	quality control
QSAR	quantitative structure–activity relationship

QSPR	quantitative structure–permeability relationship
RFLP	restriction fragment length polymorphism
RHE	reconstructed human epidermis
RIVM	Netherlands National Institute for Public Health and the Environment
RSMN	reconstructed skin micronucleaus
RT-PCR	reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulphate
SELDI	surface-enhanced laser desorption/ionization
SHE	Syrian hamster embryo
SNP	single nucleotide polymorphism
SOD	superoxide dismutase
SULT	sulphotransferase
SXR	steroid X receptor
$t_{1/2}$	half life
TCDD	2,3,7,8-tetrachlorodibenzo-*p*-dioxin
TCPOBOP	1,4-bis[2-(3,5-dichloropyridyloxy)]benzene
TGP	Toxicogenomics Project in Japan
T_M	melting temperature
TNF	tumour necrosis factor
TOF	time-of-flight
TPA	12-*O*-tetradecanoyl phorbol 13-acetate
TTC	threshold of toxicological concern
UDS	unscheduled DNA synthesis
UGT	UDP-glucuronyl transferase
ULN	upper limit of normal
UV	ultraviolet
V_{max}	maximum velocity

About the Companion Website

This book is accompanied by a companion website:

www.wiley.com/go/stanley/moleculararcellulartoxicology

The website includes:

- Powerpoints of all figures from the book for downloading
- PDFs of all tables from the book for downloading

1

Background to Molecular and Cellular Toxicology

1.1 What do we mean by molecular and cellular toxicology?

The Society of Toxicology[1] defines Toxicology as 'the study of the adverse effects of chemical, physical or biological agents on people, animals and the environment' and toxicologists as 'scientists trained to investigate, interpret and communicate the nature of those effects'. In the disciplines of molecular and cellular toxicology, toxicologists make use of the many new techniques which are becoming available in the molecular life sciences to understand the underlying mechanisms by which these agents damage cells, tissues and entire organisms. The main aims of toxicity testing, whether during pre-clinical drug development, in the course of safety assessment of cosmetic ingredients and consumer products or while evaluating the potential consequences of exposure to industrial and environmental chemicals, are to construct a toxicological profile of the chemical and to identify a threshold dose (if any).

The topic of this book is how molecular and cellular techniques can be used to study the toxicity of exogenous chemicals, referred to in the trade as *xenobiotics*. The primary target organs for xenobiotic toxicity are usually those which are exposed to xenobiotics and their metabolites because of the roles they play as portals of entry, sites of metabolism and/or organs of excretion. The molecular and cellular consequences of exposure are summarised in Figure 1.1.

Despite the many scientific advances made in the life sciences over the last couple of decades, which include spectacular advances in the fields of molecular biology, biotechnology and bioinformatics, the basic concepts of regulatory

[1] http://www.toxicology.org/

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

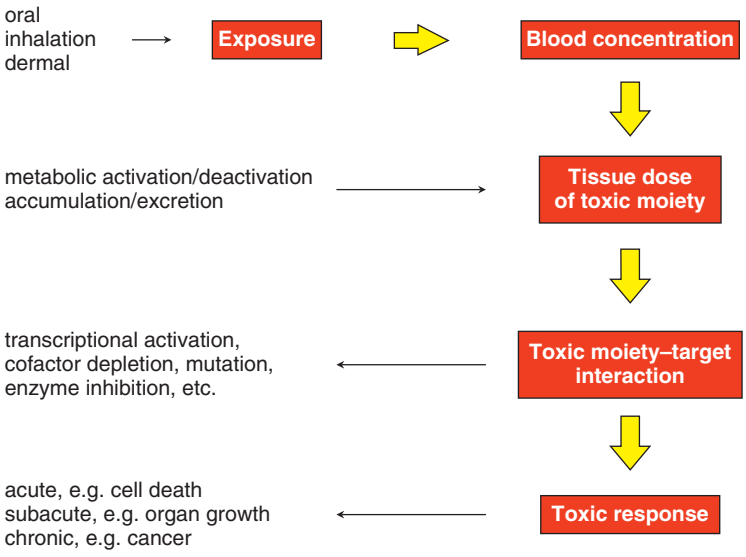


Figure 1.1 Consequences of exposure to a toxic insult (source: Dr Cliff Elcombe, CXR Biosciences Ltd. Reproduced with permission of Dr Cliff Elcombe)

toxicology have hardly changed over the same period. For example, although the classical LD_{50} (dose giving 50% lethality) test for oral toxicity and the Draize tests for eye or skin irritancy are widely considered to cause unacceptable suffering to laboratory animals, they are still widely used and the development of non-animal alternatives has been slow, to say the least. However, the implementation of both the 7th Amendment to the European Union (EU) Cosmetic Ingredient Directive and the Registration, Evaluation, Authorisation and restriction of Chemicals (REACH) regulations during the early 2000s has provided a strong stimulus for further developments.

There is an acute need for this to be reflected in a paradigm shift in the field of toxicology to take advantage of the new opportunities offered by modern developments in the life sciences, including new *in vitro* models, alternative whole organism (non-mammalian) models and the exploitation of 'omics methods, high throughput screening (HTS) technologies and molecular imaging technologies.[2]

1.2 Tissues and their maintenance

Tissues are made up of cells of various types plus the extracellular space which surrounds them. The extracellular space is filled with extracellular matrix, the proportion and structure of which depends on the tissue type. Epithelia, for example, consist mainly of sheets of epithelial cells with very little extracellular matrix whereas connective tissue contains few cells and a lot of extracellular

[2] Reviewed from a 3Rs perspective by Van Vliet (2011).

matrix. The proteins of the extracellular matrix are linked to cytoskeletal proteins through the plasma membrane and are able to influence cell development, migration, proliferation, shape and function.

All tissues have certain basic requirements including mechanical strength, access to nutrients and removal of waste, connection to the nervous system, removal of debris and protection against infection. Specialised (differentiated) cells provide these and other functions.

During the process of embryonic development, the fertilised ovum proliferates and the resulting daughter cells differentiate to form three germ layers:

- The endoderm gives rise to the epithelia of the gut and its associated organs (lung, liver and pancreas).
- The ectoderm gives rise to the outer surface epithelia (epidermis, buccal epithelium and outer cervical epithelium) and neuroectodermal tissues.
- The mesoderm gives rise to the embryonic mesenchyme and thence to the connective tissue and supporting tissues including bone, cartilage, muscle, vascular tissue and haematopoietic system.

The products of this process are the various differentiated tissues of the body. Even when removed from their normal environment, differentiated cells retain their specialised characteristics; for example, glandular cells still secrete mucin, fibroblasts still make extracellular matrix and macrophages still carry out phagocytosis. Differentiated cells can still respond to the environment and some cell types can adapt quite dramatically: for example, fibroblasts can convert into cartilage cells, liver cells can express different enzymes and mammary cells can switch milk proteins on or off. Some cells, however, are terminally differentiated, having become so specialised that they have lost the ability to divide.

1.2.1 Stem cells

Terminally differentiated tissues are maintained by stem cells, precursors which are not themselves differentiated but are committed to produce a particular type of terminally differentiated cell. A stem cell can be defined as ‘a cell which can proliferate either symmetrically or asymmetrically in response to an appropriate external signal’;[3] in other words, under one set of circumstances it will divide to produce two stem cells and under other circumstances it will divide to generate one stem cell and one progenitor cell which can give rise to a differentiated cell lineage. The signals to which stem cells can respond include growth factors, levels of oxygen and antioxidants and growth substrates (e.g. feeder layers, extracellular matrix).

Stem cells can divide without limit and on division the daughter cells have a choice either to remain as a stem cell or embark on terminal differentiation. The final differentiated state of the majority of stem cells is pre-determined (e.g. muscle satellite cell, spermatogonium), although some stem cells are pluripotent (can differentiate into many cell types). Organ-specific stem cells have two defining properties, the ability to self-renew and the potential to differentiate into

[3] For a review, see Kang and Trosko (2011).

organ-specific cell types. The various types of stem cells have different potencies (i.e. abilities to generate different classes of progeny):

- **A totipotent stem cell** can generate an entire new organism. The definitive totipotent stem cell is the fertilised egg; following implantation, the totipotent fertilised egg becomes committed to form an embryonic pluripotent stem cell.
- **A pluripotent stem cell** can give rise to any other type of cell but not to an entire new organism. Pluripotent cells give rise to committed progenitor cells which can only mature into one type of cell (i.e. each one is unipotent) and this maturation process involves differentiation, which is controlled by growth factors and the surrounding environment.
- **Multi-potent stem cells** can produce a limited number of cell types and are committed to become part of a particular organ. They give rise to lineages of progenitor cells.
- **Progenitor cells** are committed to a particular lineage (e.g. the haematopoietic system) and give rise to terminally differentiated cells, which do not divide further.

1.3 Tissue damage

Living tissues are constantly exposed to environmental changes to which they respond with modifications of metabolism and growth.

- **Primary (direct) injury** involves an interaction between the chemicals and the components of the cell. Toxic cell injury requires high concentrations of toxic compounds and, in some cases, metabolic activation. It may involve membrane damage (e.g. lipid peroxidation induced by carbon tetrachloride in the liver).
- **Secondary (indirect) injury** involves changes in the cellular environment (e.g. oxygen tension, nutrient supply, hormone levels).

1.3.1 Consequences of tissue injury

The primary responses following tissue damage due to an injury are cell death and acute inflammation. The pathological stimuli responsible may be endogenous (e.g. hormones, autoimmunity, anoxia) or exogenous (e.g. radiation, drugs/chemicals, infections, mechanical trauma, heat or nutritional imbalances). The pathological changes observed following a toxic insult give an indication of the vulnerability of certain organ systems, and their nature and severity may give an insight into the toxicity of the compound. However, pathological changes, as revealed by microscopy, do not necessarily provide information about the sub-cellular and molecular processes involved.

The long-term consequences of injury depend on the ability of the tissue to regenerate and on whether the damaging agent persists. They include regeneration, healing by repair and chronic inflammation. The final outcome

may be restoration (complete healing with full functionality) or fibrosis/scar formation.[4]

Following an episode of tissue damage, the following may occur:

- Full regeneration: an optimal response, but only occurs in the liver in higher organisms.
- Removal/repair of necrotic tissue leading to restitution or fibrosis (scar formation).
- Alterations to necrotic tissue (e.g. calcification).

Acute and chronic inflammation Acute inflammation is the commonest early response to tissue damage and destruction. The classical clinical indications of acute inflammation are *rubor* (redness), *calor* (heat), *dolor* (pain), *tumor* (swelling) and loss of function. If the injury is not too severe and the damaging agent has been removed, this will rapidly subside and the tissue will start to heal itself, either by restoration or by scar formation. Restoration occurs when there is minimal damage to the tissue architecture and comprises restoration of the normal structure and function of the tissue without forming a scar. It requires the supporting stroma to be intact and the damaged cells must be able to regenerate as in, for example, liver regeneration following acute liver damage. Regeneration depends on the ability of cells to divide, which means that it usually involves stem cells such as those in the gastrointestinal (GI) tract, urinary tract, skin, lymphoid tissue and the haemopoietic system. Cells such as hepatocytes which can come out of quiescence are also able to regenerate, but terminally differentiated cells such as cardiac myocytes and neurons cannot regenerate.

If, however, the damaging agent persists, the tissue will become chronically inflamed. In chronic inflammation, the processes of necrosis, organisation and repair all occur simultaneously. Chronic inflammation occurs in situations such as long-term alcohol abuse, where the ability of the liver to restore itself is overwhelmed by continual exposure to alcohol and progresses through chronic inflammation to fibrosis, cirrhosis and ultimately liver failure. The macrophage, which arises as a result of monocyte differentiation in response to interferon γ, is the main effector cell in chronic inflammation. Activated macrophages are also called epithelioid cells and can fuse to form multi-nucleate histiocyte giant cells, which have both phagocytic and secretory roles in chronic inflammation.

Restoration Restoration occurs when there has been minimal damage to the tissue architecture and cells can re-grow. The end result of this process is restoration of normal tissue structure and function without scarring. In order for this to occur the acute inflammatory response must be terminated appropriately. The support stroma must remain intact and the damaged cells must be able to regenerate. This process is most clearly seen in response to acute liver damage (e.g. two-thirds partial hepatectomy), when the full mass and function of the liver is restored within a few days. The ability of a tissue to regenerate depends on the ability of its cells

[4] For a beautifully illustrated review of cellular processes involved in healing, see Abreu-Blanco *et al.* (2012).

to divide, so this process is dependent upon the presence of stem cells (e.g. in the GI tract, urinary tract, skin, lymphoid tissue and haematopoietic system) or cells which can come out of quiescence (e.g. liver, kidney). Terminally differentiated cells (e.g. cardiac myocytes, neurones) cannot regenerate, so tissues made up of these cell types are particularly vulnerable to injury.

Scar formation If a tissue is too severely damaged for restoration to be possible, healing can occur by means of organisation and repair, leading to the formation of a scar. Macrophages phagocytose dead tissue and inflammatory exudate and existing capillaries bud into the damaged area leading to the formation of vascular granulation tissue (organisation). Proliferation of fibroblasts within this tissue causes it to develop into fibrovascular granulation tissue, which gradually fills with collagen secreted by the fibroblasts to form a collagenous scar (repair). This process can be impaired by inadequate nutrition, ischaemia , infection, disease (e.g. diabetes) and the presence of foreign material.

1.3.2 Reversible changes in cells and tissues

Each tissue is an intricate mixture of different cell types and this organisation is maintained even though individual cells are constantly dying and being replaced. An appropriate balance between cell growth and cell death is therefore essential for the maintenance of homeostasis. An excess of cell growth over cell death leads to disorders of cell accumulation (e.g. cancer) and insufficient cell growth combined with excessive cell death leads to disorders of cell loss (toxicity/atrophy). Changes in cellular growth patterns may involve changes in the size of cells or in their number. Alterations in the differentiation state of cells (dedifferentiation or metaplasia) may also occur.

Xenobiotics may induce the following reversible changes in the pattern of cellular growth:

- **Hypertrophy** is an increase in the size of the individual cells within a tissue. It may be a physiological adaptation (e.g. hypertrophy of skeletal and cardiac muscle in athletes), an adaptive response to stress (e.g. hepatocyte hypertrophy due to enzyme induction and proliferation of the smooth endoplasmic reticulum) or a pathological effect (e.g. heart muscle in hypertension). The opposite of hypertrophy (cell shrinkage) is called *atrophy*.
- **Hyperplasia** is an increase in cell number leading to an increase in the volume of an organ. By definition, it can only occur in cell types which have retained the ability to divide, and is therefore not seen in terminally differentiated tissues (brain, skeletal muscle). Hyperplasia may be a physiological process (e.g. in the lactating breast) or a repair process (e.g. wound healing). It can also have pathological consequences because it is necessary for fixing (i.e. making irreversible) DNA damage, which increases the risk of neoplasia. Hyperplasia can occur in response to toxic stimuli in epithelial cells (e.g. renal tubule, pulmonary alveolar epithelium, intestinal epithelium and epidermis), blood cells, thyroid cells and bone tissue. The liver, despite having a very low rate of cell proliferation under normal circumstances, can come out

of quiescence and respond with a spectacular proliferative response following chemical or physical damage. The opposite of hyperplasia (reduced cell proliferation) is called *hypoplasia*.

- **Metaplasia** refers to the reversible replacement of one type of adult cell by a simpler mature cell type due to abnormal differentiation of a stem cell. This is commonly an adaptation to stress (e.g. as a result of chronic inflammation). A classical example of metaplasia is the replacement of the columnar epithelium of respiratory tract (the bronchial lining) with squamous epithelium in smokers (squamous metaplasia). This is a reversible event and therefore not classified as part of the neoplastic process, but it can lead to dysplasia, which is irreversible.

1.3.3 Irreversible changes in cells and tissues

When a tissue is exposed to a sub-lethal dose of a toxin it may undergo reversible adaptive changes in order to cope with the insult. Such changes may be accompanied by morphological alterations, but these will regress if the insult is removed. However, if adaptive changes are insufficient to overcome the insult, the cell will progress to irreversible damage. The likelihood of this happening depends upon the cell type and its metabolic state at the time of injury; for example, ischaemia will cause irreversible damage after a few minutes in neurons, after 1–20 min in cardiac myocytes and after 1–2 h in epithelial cells of the renal proximal tubule. Severe or chronic toxic insults can therefore lead to irreversible changes in cell growth and differentiation.

- **Dysplasia** is an abnormal change in the arrangement and size of cells in a tissue. It can sometimes be reversed, but is generally considered to represent a point of commitment to the carcinogenic process.
- **Neoplasia** literally means new growth: a group of cells which are growing in an uncontrolled manner.
- **Anaplasia** refers to the regression of the physical characteristics of a cell towards a more primitive or undifferentiated type and is a common feature of malignant tumours.

1.4 Tissue responses to injury

The main targets for damage within the cell are the cell membrane, mitochondria, cytoskeleton and DNA. The biochemical mechanisms involved include ATP loss, release of calcium into the cytoplasm, reactive oxygen metabolites, structural damage to membranes and cytoskeleton and DNA damage which can be lethal or lead to mutations.

1.4.1 Oxidative stress

Oxidative stress underlies a vast number of human diseases as well as mechanisms of toxicity of drugs and chemicals; in addition, exposure to environmental

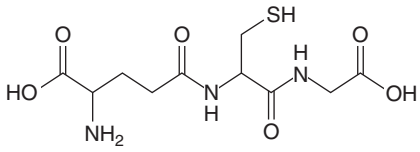


Figure 1.2 Structure of glutathione

chemicals can cause a variety of human diseases by mechanisms which involve oxidative stress.

Oxidative stress has been implicated in the toxicity of a plethora of drugs and environmental chemicals. One of the key intracellular molecules involved in this process is the tripeptide glutathione (γ-glutamyl-cysteinyl-glycine (GSH) (Figure 1.2). Glutathione is the most abundant low-molecular-weight thiol found within cells; by cycling between its reduced state (GSH) and its oxidised state (glutathione disulfide) it helps to maintain the appropriate redox status within the cell. Many inducers of oxidative stress exert their toxic effects by causing GSH depletion as a consequence of generation of reactive oxygen species, which arise from a variety of sources, being either endogenously generated or produced by environmental agents such as xenobiotics, UV irradiation and infectious organisms.[5] In the event that a cell produces more reactive oxygen species than can be detoxified the result is DNA damage, lipid peroxidation and cell death. The consequence of this is an acute or a chronic disease.

The three major types of reactive oxygen species are as follows:

- **Superoxide anion radical (O_2^-)**, which is present constitutively in cells because of leakage from the mitochondrial respiratory chain.
- **Hydrogen peroxide (H_2O_2)**, resulting from the dismutation of $O_2^{-\cdot}$ or directly from the action of oxidase enzymes.
- **Hydroxyl radical (•OH)**, a highly reactive species that can modify purine and pyrimidine bases and cause DNA strand breaks resulting in DNA damage.

Reactive oxygen species are natural by-products of cellular metabolism, and oxidative stress is tightly regulated by the balance between their production and removal. Stress response pathways have evolved to protect cells against oxidative stress and environmental challenge, as well as to repair damage. Indeed, all organisms have enzymes which can scavenge superoxide and H_2O_2 (Imlay, 2008). Oxidative stress is a very complex problem because of the number of different pathways which exist in order to protect against different oxidants. In mammalian cells, for example, the enzymes involved include superoxide dismutases (SODs), catalase, glutathione *S*-transferases (GSTs), glutathione peroxidase, reduced nicotinamide adenine dinucleotide phosphate (NADPH)-quinone oxidoreductase, haem oxygenase-1 (HO-1), dual-specificity phosphatases, thioredoxin and peroxiredoxins.

The main intracellular site where reactive oxygen species are generated is the mitochondrion, and mitochondrial energy metabolism is quantitatively the most important source of reactive oxygen species in the majority of eukaryotic cell types (Kowaltowski *et al*., 2009; Khansari *et al*., 2009; Murphy, 2009; Kagan *et al*.,

[5] For a review, see Auten and Davis (2009).

2009). The production of reactive oxygen species in mitochondria is a normal consequence of respiration, but the different species generated during this process can differ markedly in their reactivity and lifetime; for example, hydroxyl radical reacts almost instantaneously with adjacent molecules whereas semiquinones may be stable for days, weeks or months (Pryor *et al*., 2006). In fact, mitochondria can produce reactive oxygen species even under conditions of hypoxia and this may have implications for mitochondrial redox signalling (Murphy, 2009).

The primary reactive oxygen species generated by mitochondria is superoxide, which is produced by one-electron reduction of O_2 and metabolised by SOD in the inter-membrane space. The resulting H_2O_2 is relatively unreactive; however, if it is not metabolised, it may go on to form hydroxyl radical via a Fenton reaction (Kowaltowski *et al*., 2009; Pryor *et al*., 2006). If there is a lack of balance between reactive oxygen species generation and antioxidant defence mechanisms, reactive oxygen species can leak from the mitochondrion causing damage to cellular targets, including cell membrane fatty acids (forming lipid peroxides), cellular proteins (damaged proteins may accumulate up to toxic levels causing cell death) and DNA (causing DNA strand breaks and deletions). The consequent damage has a number of sequelae including modulation of survival signalling molecules, triggering of cell death pathways and production of proinflammatory cytokines and chemokines (Khansari *et al*., 2009; Roberts *et al*., 2009; Pan *et al*., 2009).

Oxidative stress and chronic inflammation are characteristic of a wide variety of human diseases (Brenneisen *et al*., 2005), but it is often difficult to determine whether oxidative stress is a primary cause of cell death or a physiological consequence of the induction of cell death pathways. This has led to the identification of a 'growing need for simple, convenient, and reliable markers for the assessment both *in vitro* and *in vivo* of the metabolic/oxidative distress and of its modulation … [by] … pharmaceutical products' (D'Alessandro *et al*., 2011).

Oxidative stress also plays a role in cellular senescence and cancer. Unbalanced regulation of the production of reactive oxygen species appears to initiate cellular senescence programmes via multi-faceted mechanisms including the direct induction of mutations (Pan *et al*., 2009), and the altered metabolic state of cancer cells (associated with aerobic glycolysis) makes them particularly susceptible to reactive oxygen species damage linked to the accumulation of mutations (D'Alessandro *et al*., 2011). Indeed, the interaction between reactive oxygen species and cellular senescence has been suggested as a target for cancer therapy (Pan *et al*., 2009).

The transcription factor Nrf2 plays a key role in cellular responses to oxidative stress.[6] Under normal conditions, Nrf2 is located in the cytoplasm and is bound to the accessory protein Keap1, which targets Nrf2 for proteasome-mediated degradation. Under conditions of oxidative stress, however, Keap1 loses its ability to bind Nrf2, which is then able to translocate to the nucleus and bind to the antioxidant response elements in the 5′ regulatory regions of target genes.

Example: Oxidative stress and cardiovascular disease

Oxidative stress is known to play a role in the aetiology of cardiovascular diseases (atherosclerosis, coronary artery disease and myocardial infarction) as

[6] For a review, see Limon-Pacheco and Gonsebatt (2009).

well as stroke, dementia, Parkinson's disease and cancer. Excessive oxidative stress and chronic inflammation are characteristic features of cardiovascular disease, which features increased production of reactive oxygen species, compromised antioxidant defences (e.g. GSH depletion) and increased circulating levels of proinflammatory cytokines (Lee *et al*., 2011). Oxidative stress may be involved in the pathogenesis of cardiovascular disease, and the Nrf2 pathway has been implicated in this process, especially in the sedentary elderly. Proposed strategies for prevention range from lifestyle changes including dietary supplementation (increased consumption of broccoli, curcumin) and increased exercise to pharmaceutical interventions such as prophylactic treatment with the ubiquitin-proteasome inhibitor MG132, which may protect cardiomyocytes against oxidative stress via Nrf2-mediated up-regulation of antioxidant genes.

1.4.2 Necrosis and apoptosis

The two main mechanisms by which cells can die are necrosis and apoptosis (also known as programmed cell death).[7] If repair mechanisms and changes in gene expression are insufficient to allow the cell to cope with a toxic insult, the consequence is cell death by one of these two mechanisms. Many of the key players in apoptosis have now been identified and it turns out that many of the genes which control cell growth are also involved in apoptosis, allowing the processes of cell division and death to be tightly coregulated. This is crucial for the maintenance of homeostasis; indeed, evasion of apoptosis signals is a hallmark of human cancer, and the genes involved in regulation of the cell cycle and apoptosis represent potential molecular targets for the control of human diseases where inappropriate apoptosis is prominent (e.g. cancer and degenerative disorders).

The occurrence of necrosis is not determined by factors that are intrinsic to the cell but by changes in the cellular environment. In contrast, apoptosis is a physiological process and plays a role in the maintenance of tissue homeostasis. Apoptosis is an active process leading to cell death via an ordered sequence of events. Its morphologic appearance is very different from that of necrosis (Figure 1.3). Necrosis involves the swelling and rupture of the injured cells, whereas apoptosis involves a specific series of events that leads to the dismantling of the internal contents of the cell. Apoptosis is a normal part of development, being involved in processes including

- resorption of the tail during metamorphosis of a tadpole;
- removal of the webbing between the digits of hands and feet during mammalian embryonic development.

Necrosis is an unregulated mechanism involving dilation of the endoplasmic reticulum, dissolution of lysosomes and ribosomes, mitochondrial swelling and an increase in cell volume. Necrosis does not, however, involve gross changes in chromatin structure. In contrast, during apoptosis, the cytoplasmic organelles are well preserved and there is actual shrinkage of the cell and nucleus. The biochemical changes which occur during apoptosis include a moderate increase in

[7] For a general review, see Han, Kim and Kim (2008).

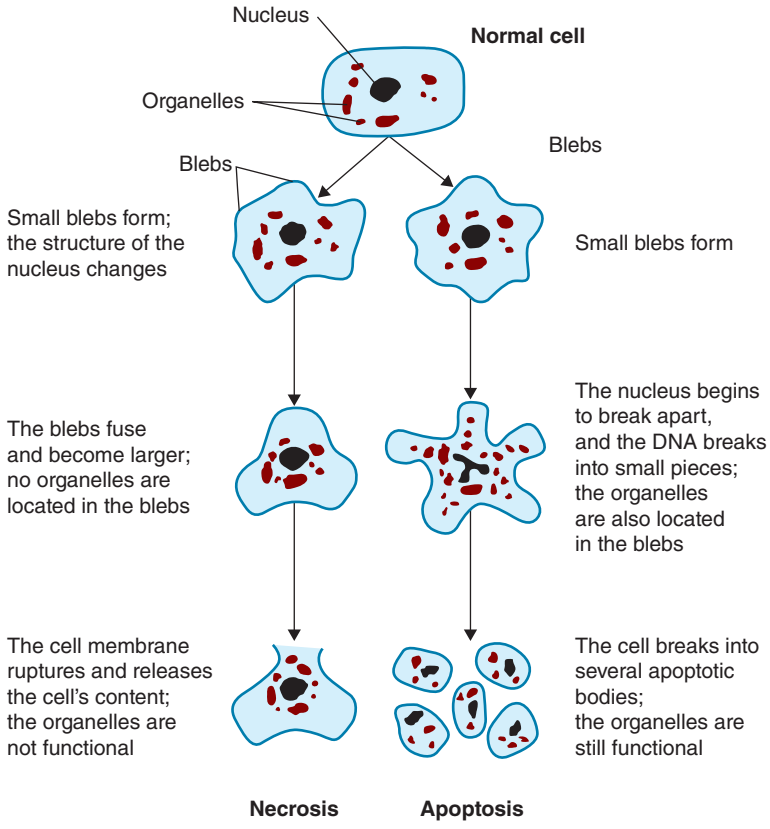


Figure 1.3 Structural changes of cells in necrosis and apoptosis (Source: Goodlett and Horn (2001); figure 1. Reproduced with permission of Charles Goodlett)

intracellular [Ca^{2+}] and total shutdown of protein and RNA synthesis. Following condensation of the nuclear chromatin, activation of Ca^{2+}/Mg^{2+} endonuclease produces distinctive chromatin fragments which may be viewed as a ladder on an agarose gel (Hooker *et al*., 2012).

Apoptosis is characterised by distinctive morphological changes including decreased cell volume, increased cell density, compaction of cytoplasmic organelles (except for the mitochondria, which remain morphologically normal) and dilation of the endoplasmic reticulum. Nucleolar disintegration, budding and separation of nucleus and cytoplasm into multiple small membrane-bound apoptotic bodies occurs, followed by progressive degeneration of residual nuclear and cytoplasmic structures and condensation of chromatin into crescent-shaped caps at the cell periphery.

A key process in apoptosis is the activation of a series of highly conserved cysteine proteases called *caspases* which act as common effector molecules in various forms of cell death. Caspases are produced as inactive precursors called *procaspases* which can be activated either by oligomerisation (initiator caspases e.g. caspases 8 and 9) or by proteolytic cleavage to create active enzymes in a proteolytic

cascade (effector caspases e.g. caspase 3). Once they are activated, caspases cleave other proteins within cells resulting in efficient and precise killing of the cell in which they are activated.

Apoptosis can be initiated via two different pathways, the extrinsic and intrinsic pathways, each of which involves the activation of specific caspases:[8]

- **The extrinsic pathway** is triggered by the binding of so-called death signals such as tumour necrosis factor (TNF), Fas ligand or TNF-related apoptosis-inducing ligand to the corresponding receptors (TNFR, Fas receptor, DR4/5). This triggers the recruitment of adaptor proteins and activation of the initiator caspases 8 and 9. Interestingly, caspase 8 mutations have been detected in some cancers and can act as dominant negative mutations, blocking apoptotic cell death and having a profound impact on the cancer cell's ability to undergo apoptosis (Fulda, 2009).
- **The intrinsic pathway** is mediated by mitochondrial damage, which leads to cytochrome c release. Cytochrome c interacts with an adaptor molecule (Apaf-1) and procaspase 9 to form a complex called an *apoptasome* leading to the activation of caspase 9. Inhibitor of Apoptosis Proteins (IAPs) inhibit caspases and procaspases and are themselves controlled by other mitochondrial proteins (Smac/DIABLO and Omi/HtrA2). This process is regulated by various factors including members of the Bcl-2 family, Bax and Bak.

Both pathways lead to the activation of so-called executioner caspases (caspase 3 and caspase 7) by caspases 8 and 9.

Apoptosis appears to be the major pathway of cell death triggered by DNA damage. It is thought to eliminate genetically damaged cells and, therefore, counteracts carcinogenesis. The genes involved in regulating this process represent potential molecular targets for the control of human diseases which feature inappropriate apoptosis (e.g. cancer and degenerative disorders).

Role of calcium in cell death Calcium signals are responsible for the regulation of many vital cell functions; cellular Ca^{2+} overload or perturbation of intracellular Ca^{2+} compartmentalisation can cause cytotoxicity. The point of no return is thought to be when the sarcolemma can no longer bind Ca^{2+} and the mitochondria start to take up the excess calcium (Rasola and Bernardi, 2011). Cell death can be brought about by a loss of Ca^{2+} homeostatic control, but can also be triggered by more subtle changes in Ca^{2+} distribution within intracellular compartments.[9]

Normal intracellular calcium levels range between 10^{-7} and 10^{-6} M, approximately four orders of magnitude lower than in the extracellular fluid, thanks to the activity of plasma membrane Ca^{2+} ATPases and the Na^{+}/Ca^{2+} exchanger which remove Ca^{2+} from the cell. This concentration gradient allows rapid influx of Ca^{2+} into the cell if the plasma membrane channels open, for example in response to a toxic insult. An unphysiological increase in cytosolic calcium concentration (to $>10^{-5}$ M) can either cause necrosis or contribute to apoptosis by activating

[8] For a review from the perspective of renal toxicity, see Servais *et al.* (2008).
[9] For a detailed review of this topic, see Zhivotovsky and Orrenius (2011).

physiological calcium-dependent processes and cellular responses that are not normally affected by calcium. These include changes in cellular shape, blebbing (a bleb is an irregular bulge in the plasma membrane of a cell caused by localised decoupling of the cytoskeleton from the plasma membrane), changes in ionic conduction, excessive contraction/transmitted release, enzyme activation (proteases, nucleases, phospholipases, transglutaminases) and inhibition (adenylate cyclase). This is associated with a cycle of decreased ATP levels due to activation of the Ca^{2+} pump, subsequently causing reduced functioning of the pump itself due to insufficiency of ATP.

1.4.3 Neoplasia

The term *neoplasia* refers to the development of a collection of cells which grow in an uncontrolled manner, usually resulting in the formation of a tumour. Neoplasia may be benign or malignant and its causes include foreign chemicals (xenobiotics), hormones, viruses, inherited genes, radiation and physical agents (e.g. inert implanted plastic or metal film).

The cellular changes observed during neoplasia include hypertrophy, hyperplasia and alterations in differentiation which may be associated with changes in the histological appearance of cells. It is important to note that the terms cancer and tumour are not synonymous: the word tumour applies to any readily defined mass of tissue distinct from normal tissue but not necessarily made up of abnormal cells; in other words, the term tumour just means a lump. The definition of a benign tumour is that it is restricted to its site of origin; a malignant tumour is one which is metastatic.

Cancer is defined as a heritably altered, relatively autonomous growth of tissue which occurs when a cell or group of cells begins to multiply more rapidly than normal leading to the development of a malignant tumour: it is, by definition, a disease of aberrant cells and is a consequence of both uncontrolled cell division and the loss of normal patterns of differentiation leading to the autonomous growth of these abnormal cells. A true cancer is a malignant neoplasm made up of morphologically transformed, autonomously replicating (malignant) cells. The characteristics of morphologically transformed cells include uncontrolled cell division, loss of contact inhibition, defective cell cycle control, lack of balance between cell division and cell death, breakdowns in cell–cell communication and dedifferentiation.

Metastasis occurs when individual cells or small groups of cells break away from the primary tumour and migrate to other sites within the body, where they begin to grow into secondary tumours (metastases). Malignant tumours tend to be locally invasive, fast-growing and anaplastic whereas benign tumours are usually encapsulated and are generally slow growing.

1.4.4 The initiation–promotion paradigm

Carcinogenesis is defined as the process by which normal cells are transformed into cancer cells.[10] A carcinogen, therefore, is any substance that causes cancer.

[10] For an overview of the history of ideas about carcinogenesis, see Mitrus *et al.* (2012).

The link between cancer and exposure to specific exogenous substances was first identified in the 18th century, when in 1761 John Hill noticed an increase in nasal cancer associated with long-term use of snuff and in 1775 Percival Pott observed that chimney sweeps often suffered from scrotal cancer. The role of industrial exposure to aromatic amines in the induction of bladder cancer was demonstrated in 1895, when Rehn reported an increase in bladder cell tumours in workers in the dye and rubber industries.[11]

Some cancers arise as a result of genetic susceptibility or alterations in homeostasis (e.g. hormonal changes) while others are caused by exposure to carcinogenic substances. Experimentally, a compound is considered to be carcinogenic if its administration to laboratory animals induces a statistically significant increase in the incidence of one or more histological types of neoplasia compared with animals in the control group, which are not exposed to the substance.

According to the European Chemicals Agency (ECHA)[12]:

> Chemicals are defined as carcinogenic if they induce tumors, increase tumor incidence and/or malignancy, or shorten the time to tumor occurrence. Benign tumors that are considered to have the potential to progress to malignant tumors are generally considered along with malignant tumors. Chemicals can induce cancer by any route of exposure (e.g. when inhaled, ingested, applied to the skin, or injected), but carcinogenic potential and potency may depend on the conditions of exposure (e.g., route, level, pattern, and duration of exposure).

The concept of cancer as a multi-step process was first proposed by Armitage and Doll in 1954 and extended by Moolgavkar and Knudsen in 1981 (Armitage and Doll, 1954, Moolgavkar and Knudson, 1981). The idea that the process of carcinogenesis could be divided into discrete stages arose when it was observed that tumours could be induced by painting mouse skin with polycyclic aromatic hydrocarbons (PAHs) then with croton oil, but not if these substances were applied in the reverse order. This is now known as the initiation–promotion paradigm (Figure 1.4) and reflects the fact that the carcinogenic process involves both genetic damage and changes in the growth of cells and tissues. Cell replication is necessary for fixing DNA damage (i.e. making it irreversible), therefore increasing the risk of neoplasia, and hyperplasia is central to the promotion phase during which tumour growth occurs.

Initiation A cancer originates from an initiated cell which multiplies clonally, escapes apoptosis and accumulates a collection of genetic and/or epigenetic alterations which allow it to escape from normal control mechanisms. Mutagenesis, the process by which mutations occur, is a key aspect of carcinogenesis. A mutation is a change in the sequence of bases in a DNA molecule, and any insult which causes a mutation is known as a *mutagen*. Carcinogenesis almost always involves mutagenesis, and many carcinogens are also mutagens, but it is a complex process involving defects in many different biological control mechanisms so mutagenesis is not, in and of itself, sufficient to generate cancer.

[11] For interesting reviews on this topic, see Oliveira *et al.* (2007).

[12] ECHA *Guidance on information requirements and chemical safety assessment* Chapter R.7a: Endpoint specific guidance, Version 2.0, November 2012, Section R.7.7.8.1.
http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_r7a_en.pdf

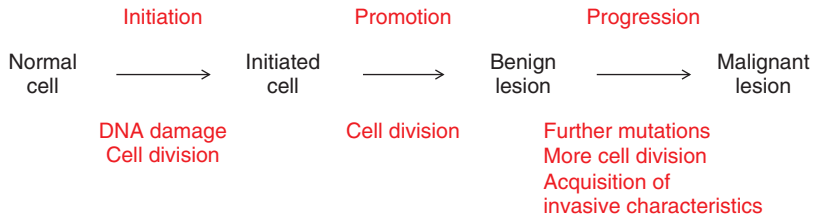


Figure 1.4 The initiation–promotion–progression paradigm

Initiation involves the acquisition of irreversible genetic alterations as a consequence of mutation of one or more key genes. This may occur spontaneously or as a result of DNA damage caused by chemical or physical damage (e.g. due to ionising radiation). Initiation gives rise to a single initiated cell which is phenotypically indistinguishable from the surrounding normal cells but is predisposed to give rise to a neoplastic lesion. It is a rapid and irreversible process and the characteristics of initiation are transmitted to daughter cells when the initiated cell divides. The ability of chemicals which induce DNA damage (genotoxicity) to cause initiation and ultimately tumour development has led to an intense focus on the identification and regulation of genotoxic chemicals.

Promotion Promotion involves an increase in the number of initiated cells to produce a benign tumour. Increases in cell number can result from either increased cell proliferation or reduced cell death, so increased cell division and evasion of apoptosis are both considered to be critical in the promotion process. Promotion is believed not to require further DNA damage.

The importance of cell proliferation in the promotion process is twofold: (i) prior to DNA replication, DNA damage can be repaired, but proliferating cells have less time to repair DNA damage, and DNA replication during cell division can fix this damage, making it permanent and irreversible. This can be particularly important if the DNA damage occurs in a stem cell which can give rise to multiple progeny, and pluripotent stem cells are now believed to be the key target cells in carcinogenesis. Indeed, cells which are fully differentiated or committed to differentiation are unlikely to give rise to tumours as they are already programmed to die. (ii) Cell proliferation converts an individual initiated cell to a detectable lesion (a preneoplastic focus and subsequently a benign tumour).

By definition, the promotion process involves hyperplasia which may be augmentative (proliferation over and above that which is required for tissue maintenance) or regenerative (proliferation which is necessary for tissue repair) (Figure 1.5). Tumour promoters contribute towards the fixation of DNA damage in the form of mutations, enhance epigenetic processes which alter gene expression and cause changes in cellular growth control. Some promoters are tissue specific whereas others can act on more than one tissue type.

Long-term treatment with certain promoting agents can induce neoplasia, apparently without the need for initiation, the process known as *non-genotoxic carcinogenesis*. Some investigators believe that this indicates a genotoxic effect which has been missed by conventional genotoxicity assays, but it is equally

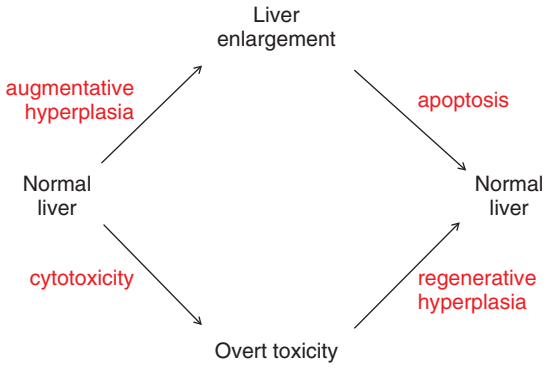


Figure 1.5 Augmentative and regenerative hyperplasia

possible that non-genotoxic carcinogenesis occurs as the result of promotion of pre-existing, spontaneously initiated cells.

The cell cycle During the cell cycle, a cell must replicate its DNA and duplicate its contents, then divide in two. In the case of unicellular organisms each cell division cycle makes a new organism whereas in the case of multi-cellular organisms many rounds of cell division are required to make a new organism. Cell division is also needed in the adult body in order to replace cells which have been lost as a consequence of apoptosis or necrosis.

Eukaryotic cells divide and grow at different rates. The eukaryotic cell cycle can be resolved into four phases (Figure 1.6):

- G1 phase (growth of the cell)
- S-phase (DNA replication)
- G2 phase (growth and error checking)
- M phase (cell divides)

Cells can leave the cell cycle during and go into quiescence (G0 phase); indeed, some cell types (e.g. hepatocytes) spend most of their time in G0. During M phase the cell divides (cytokinesis) and its DNA is distributed evenly between the daughter cells.

The cell cycle must be controlled to allow time for synthesis of new proteins, replication of DNA and checking for DNA damage. Checkpoints are required in order to ensure that each process is complete before the next one starts. If this does not happen, delays and interruptions occur; this is what happens in cancer. At a cell cycle checkpoint the cell cycle pauses until the appropriate feedback signals have been received. The key checkpoints in the eukaryotic cell cycle are

- start (in G1, just before the beginning of S-phase), when the cell becomes committed to the cell cycle
- entry to M phase (at the end of G2)
- exit from M phase

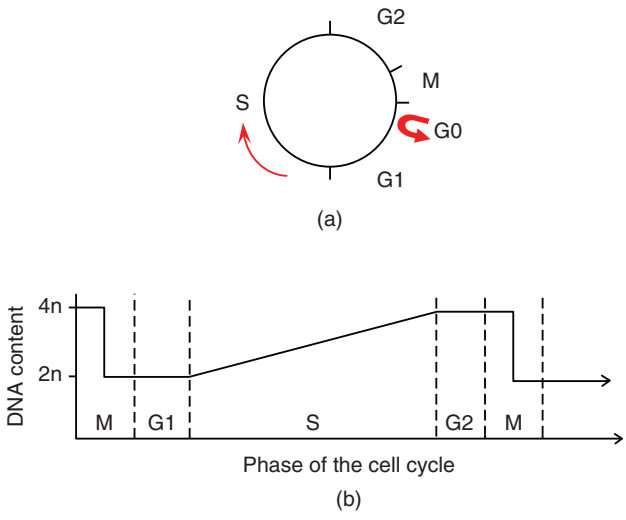


Figure 1.6 The cell cycle and cell division. (a) The phases of the cell cycle (b) DNA content of cells at different stages of the cell cycle (source: Dr Jerry Styles. Reproduced with permission of Dr Jerry Styles)

Cell cycle checkpoints are tightly controlled and monitored by proteins which include the cyclins and cyclin-dependent kinases (CDKs). The mitotic cyclins bind to CDK molecules during G2, permitting entry to mitosis, while the G1 cyclins bind to CDK molecules during G1 and are required for entry into S-phase. The CDKs act by phosphorylating specific proteins on serine or threonine residues.

As a consequence of the importance of cell cycle control, increases in cell replication are indicative of tumour promotion. Changes in cell replication may be detected by a variety of methods including:

- Incorporation of ^{3}H- or ^{14}C-labelled thymidine into cellular DNA (detected by autoradiography and scintillation counting, respectively)
- Incorporation of bromodeoxyuridine (BrdU) into cellular DNA
- Up-regulation of proliferating cell nuclear antigen, an endogenous protein whose expression increases during S-phase

These techniques can, in theory, be applied to any tissue but it is easier to detect increased cell proliferation over background in tissues with a low background rate of replication, such as the liver.

Under normal circumstances adjacent cells communicate with each other to ensure that proliferation is properly controlled. Cell–cell contact is a key aspect of this process. When non-transformed adherent cells are grown in culture they become arrested in G0 when they are touching each other on all sides, thus forming a confluent monolayer. This process, which is known as *contact inhibition*, is controlled by cyclin-dependent kinases and mediated via cell membrane proteins such as N-cadherin.

Cell–cell adhesion in epithelial cells is mediated by integrins, transmembrane proteoglycans and calcium-dependent cell–cell adhesion molecules

called *cadherins*. In particular, the calcium-dependent homophilic interactions of E-cadherin, which induce contact inhibition, maintain the epithelial cell phenotype and prevent migration. E-cadherin communicates with the cellular interior by catenin-mediated interactions with the actin cytoskeleton. One of the functions of E-cadherin is to sequester β-catenin, reducing the levels observed in the cytoplasm. Cytoplasmic levels of β-catenin are also regulated by proteolysis. Accumulation of β-catenin occurs physiologically during embryonic development and pathologically during tumorigenesis. In contrast, α-catenin, which associates with desmosomal cadherins as well as E-cadherin, is believed to have an inhibitory effect on processes associated with tumour development.

Disruption of cell–cell communication facilitates the dedifferentiation of cells to a more mesenchymal phenotype. This is a normal process during embryonic development but occurs pathologically as benign tumours progress towards a more invasive/metastatic phenotype. Loss of expression of E-cadherin and β-catenin is a key aspect of this process, which is known as the *epithelial–mesenchymal transition* and is associated with the induction of proliferative, mesenchymal and invasive genes leading towards a more malignant phenotype.

Preneoplastic lesions A preneoplastic lesion is a recognisable group of cells which has undergone the early stages of the carcinogenic process but is not yet fully committed to forming a tumour and can, under certain circumstances, regress to form apparently normal tissue. For example, a well-characterised series of phenotypic changes occurs during the early stages of rodent liver carcinogenesis *in vivo*. The methods used to detect these changes often involve the use of so-called initiation-promotion protocols (Pitot, 2007),[13] in which animals are treated with a single dose of a potent genotoxic compound such as 2-acetylaminofluorene (2-AAF) (initiation) followed by regular treatment with a compound such as phenobarbital (PB) which induces cell proliferation (promotion). In some cases two-thirds partial hepatectomy is used to stimulate cell proliferation because the rodent liver can survive removal of two-thirds of the liver. Following this surgery it undergoes a period of rapid growth as a result of which the full weight of the liver is restored in less than a week.

Other early markers of tumour development include preneoplastic enzyme changes (e.g. up-regulation of γ-glutamyl transpeptidase (γGT) and glutathione-*S*-transferases (GSTs) in rodent liver), development of morphologically altered preneoplastic foci, early changes in ploidy and nuclearity, oncogene activation, activation of growth factors and altered cell–cell communication. All these can be seen to be logically related to the acquisition of cancerous properties by the cells: up-regulation of detoxifying enzymes may confer resistance to the cytotoxic effects of the carcinogen, morphological changes are diagnostic of the transformed phenotype, oncogene activation and growth factor up-regulation enhance proliferation and loss of cell–cell communication allows the transformed cell to ignore growth retarding signals from surrounding cells.

Examples of types of preneoplastic lesions which can arise in animals during chemical carcinogenesis include altered hepatic foci, prepapillomas of the skin and aberrant crypt foci in the colon. Among the best-characterised preneoplastic

[13] For a fascinating personal account of the background to these studies, see Pitot (2007).

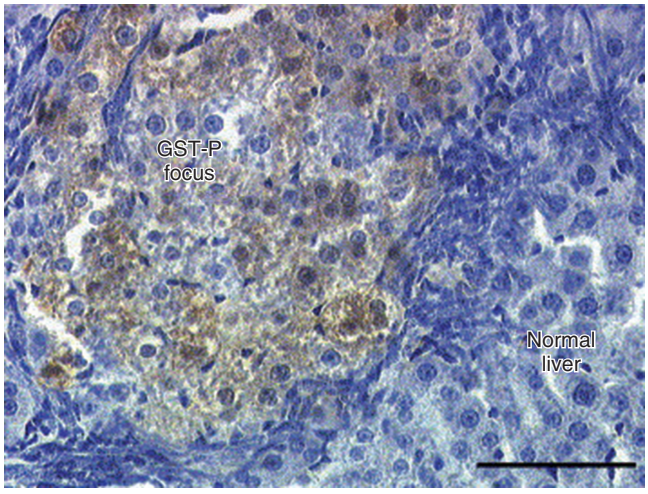


Figure 1.7 Photomicrograph of a GST-P positive preneoplastic nodule. This example is taken from a male Wister rat subjected to an initiation-promotion regime as follows: injected intraperitoneally with a single dose of diethylnitrosamine (DEN) dissolved in water (100 mg kg body weight) on day 1, then dosed with 2-acetylaminofluorene (2-AAF) (20 mg kg body weight p.o.) in dimethyl sulfoxide/carboxymethylcellulose on days 7, 8 and 9, and 70% partial hepatectomy on day 10. Transverse histological sections (6 μm thick) were stained using a rabbit primary antibody against the placental form of glutathione *S*-transferase (GST-P, 1:250). The GST-P positive hepatocytes are stained brown while normal hepatocytes are blue. (Bar = 100 μm, haematoxylin counterstain). (source: Gonzalez de Mejia *et al.* (2004); figure 6. Reproduced with permission of Elsevier)

lesions of toxicological relevance are hepatic preneoplastic foci (Figure 1.7), which develop in rodent liver in response to treatment with hepatic carcinogens.[14] They are easy to identify, being characterised by an altered pattern of expression of various markers which can be detected by means of immunohistochemistry. The most commonly used markers are GST-P and gamma glutamyl transpeptidase (γGT); others include the diubiquitin-like molecule FAT10, α_2-macroglobulin, fatty acid synthase and α-fetoprotein. One of the most useful of these markers is FAT10, which is also overexpressed in 70–90% of human hepatocellular carcinomas (HCCs). Single FAT10-positive cells appear early in the carcinogenic process, possibly as a consequence of epigenetic alterations. They have a growth advantage compared with normal hepatocytes, and are thought to represent a subpopulation of initiated cells which are resistant to cytotoxicity in the presence of a strong growth stimulus (French, 2010).

Studies on preneoplastic stem cells suggest that, at least in rats, susceptibility to the early stages of hepatocarcinogenesis is a consequence of the activation of several low penetrance genes and a single predominant susceptibility gene (French, 2010). Epigenetic phenomena also play a role; for example, hypomethylation is observed in the hepatocytes of susceptible F344 rats.

[14] For an atlas of the difference types of liver lesions found in mice, see http://www.niehs.nih.gov/research/resources/liver.

Preneoplastic lesions in humans If clearly defined and easy to identify, preneoplastic lesions can be of great value both to the toxicologist and to the oncologist. Because they occur early in the carcinogenic process, the toxicologist can use them as an early indicator of potential future tumour development, while the oncologist can use them for early diagnosis and treatment. For example, dysplastic foci (<1 mm diameter) and nodules (1 mm – 1 cm diameter) have been identified as preneoplastic lesions in human liver. These may have either a large cell or small cell morphology; large cell dysplasia is not a precursor for HCC whereas small cell dysplasia is considered to be a preneoplastic lesion (French, 2010).

Identification of preneoplastic lesions can also play a key role in cancer screening. Cancer causes approximately 840 000 deaths annually in the EU and over 1 300 000 new cases are reported each year. In many cases, by the time a tumour causes symptoms, it is too advanced to treat, but if tumours can be detected when they are very small, or better still, people who have not yet developed a tumour but have a high risk of doing so can be identified; the chances of successful treatment are much improved. As well as saving lives, a well-organised screening programme can save the health services an immense amount of money and time. In order to establish a screening programme, two key requirements must be met: There must be a test or procedure which will detect the cancer before symptoms develop (preferably at the preneoplastic stage) and there must be evidence that treatment at this earlier stage of the disease will result in an improved outcome. The most successful screening programme to date is the one for cervical cancer.

Example: Early diagnosis and treatment: the cervical smear test

The cervical cancer screening programme depends upon the identification of preneoplastic cells in otherwise healthy women. In this programme, precancerous changes in the cervix are detected by looking for abnormal cells which arise in its lining.

Precancerous change in the cervix is called *cervical intraepithelial neoplasia* (CIN). In this condition, abnormal cells with large, oddly shaped nuclei are seen in the lining of the cervix. The process starts with CIN I and progresses through CIN II to CIN III. The majority of these changes will eventually revert to normal, but in a few cases they precede the development of a tumour.

When a woman goes for a smear test, cells are scraped from the cervix using a spatula, spread on a slide, stained using a method developed by Papanicolau in 1942 (the Pap test), and examined to identify any which are a strange colour or have an enlarged or bizarrely shaped nucleus. If a mild abnormality (dyskaryosis) is detected, the woman will be asked to return in 3–6 months. If the changes persist, she will be referred for colposcopy. In the colposcopy procedure a gynaecologist examines the cervix directly and removes any abnormal tissue. This procedure cures the problem at the same time as confirming the diagnosis.

The problem of cervical cancer screening is that the current screening programme depends upon the detection of dyskaryotic cells by trained cytology screeners. The process is labour-intensive, expensive, subjective and prone to a significant incidence of errors. Any method giving a simple yes/no answer would be a significant improvement, especially if it lent itself to automatic sample processing and analysis.

Could CYP1B1 be the answer? Immunohistochemical analysis of the expression of a protein called CYP1B1 in human tumours and corresponding histologically normal tissues (Murray *et al*., 1997), suggested that this cytochrome P450 (CYP) isozyme might have a future as a tumour marker since CYP1B1 expression was detected in 122/127 tumours and 0/130 normal tissue samples.

The question that arose from this was: At what stage of tumour development is CYP1B1 expression initiated? Cervical cancer was selected for further study for the reasons outlined earlier, and the results of this study indicated that expression of CYP1B1 was detectable in cervical lesions and individual exfoliated cervical epithelial cells from a subject with CIN. The location of the staining corresponded with the area occupied by abnormal cells. In contrast with the difficulty of picking out abnormal cells in a Pap smear, individual CYP1B1 positive cells could readily be detected in smear samples from patients with CIN (Figure 1.8). Thus, CYP1B1 appeared to be a good marker for precancerous changes in the cervix. The potential advantage of this test is that it gives an unequivocal result: normal cells are blue whereas abnormal cells are brown. This test would be much easier to interpret than the Pap test, and would lend itself to automation, making screening cheaper and less labour-intensive.

The result of initiation followed by promotion is a benign tumour. The definition of a benign tumour is that it does not metastasise. Benign tumours do undergo excessive growth and immortalisation, but are unable to spread metastatically and are generally slow growing. The promotion process is reversible and subject to physiological modification; if the promoting stimulus is removed regression can occur as a result of the induction of apoptosis.

Progression The final stage of carcinogenesis, leading to a malignant tumour, is progression. Progression involves the acquisition of metastatic capability by the tumour, converting it from a benign tumour to a malignant one. A malignant tumour is one which is metastatic and a true cancer is, by definition, a malignant neoplasm. Malignant tumours tend to be locally invasive, fast-growing and anaplastic. Metastasis occurs when individual cells or small groups of cells break away from the primary tumour and migrate to other sites within the body, where they begin to grow into secondary tumours (metastases). The processes involved in progression are less clearly defined than those involved in initiation and promotion, but are believed to include loss of cell–cell communication within the tumour and the acquisition by individual cells of the ability to detach from the main tumour, invade adjacent tissues, migrate via the blood stream and establish new colonies in distant tissues (metastasis).

Progression is the conversion of a preneoplastic/benign lesion to the so-called malignant phenotype. The characteristics of this phenotype include neoplastic differentiation, unlimited replicative potential, ability to evade apoptosis, sustained angiogenesis, telomerase overexpression, self-sufficiency in growth signals and insensitivity to anti-growth signals. Collectively, these changes confer on the tumour cells the capacity for tissue invasion and metastasis.

Progression to malignancy involves both genetic and epigenetic mechanisms leading to the ability to proliferate indefinitely, independent of stimulation. This process is irreversible and involves changes in the biochemical, metabolic and morphological characteristics of the cells within the lesion. Malignant tumours

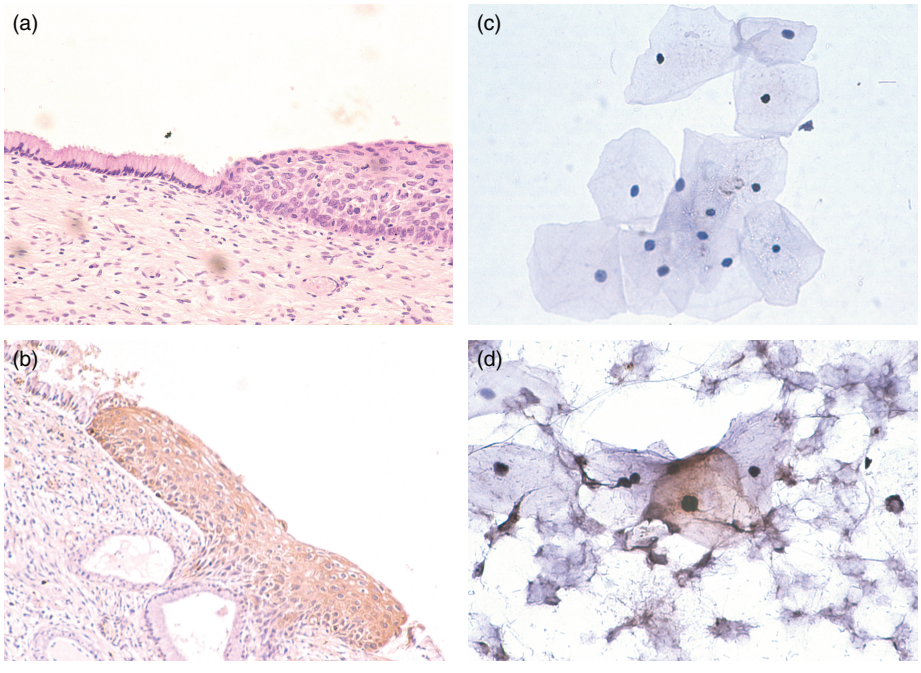


Figure 1.8 CYP1B1 as a preneoplastic marker in cervical intraepithelial neoplasia. (a) Normal structure of the endocervical junction stained with haematoxylin and eosin; (b) Endocervical junction exhibiting cervical intraepithelial neoplasia (CIN), stained for CYP1B1 (brown) and counterstained with haematoxylin (blue); (c) Cells taken at colposcopy and immunostained with preimmune (negative control) serum; (d) Cells taken at colposcopy and immunostained with a monoclonal antibody raised against CYP1B1. The immunostain appears brown; haematoxylin (blue) counterstain was used to reveal the cells. Immunostaining undertaken by Dr Thoung Hoang; photography by Mr Mike T. Ball

are characterised by excessive growth, immortalisation and metastatic spread.[15] Malignant cells are able to separate from the parent tumour, migrate to distant locations and establish new tumours. The characteristics which make this possible include rapid growth, invasiveness and the ability to stimulate angiogenesis, with invasiveness being the earliest manifestation. Metastasis entails a strong component of tumour–host interactions.

Proteins which play key roles in the progression phase of carcinogenesis include the proteins CD44 and osteopontin, stress response genes that contribute to host defences. Osteopontin is an early mediator of host defences and provides protection against intracellular pathogens while CD44 is a transmembrane glycoprotein which acts as a homing receptor and is normally expressed in lymphocytes and macrophages. Both genes are dysregulated in autoimmune diseases.

One of the key aspects of the progression phase of carcinogenesis is acquisition of the so-called mutator phenotype (Loeb, 2011). This term was coined to convey the fact that cancer cells have an increased error rate during DNA synthesis, possibly arising as a consequence of mutations in genes which govern genetic stability (e.g. those responsible for DNA repair, DNA replication, chromosomal segregation and cell cycle checkpoint control).

[15] For a review see Weber (2008).

Cells undergoing the processes of promotion and progression exhibit extremely diverse phenotypes, reflecting the different combinations of genetic and epigenetic alterations which occur during the neoplastic process. Selection of the cell population having the most favourable genotype/phenotype for survival occurs; indeed, only 1/100 individual GST-P$^+$ hepatocytes goes on to form an altered hepatic focus, let alone a full-blown tumour. This phenotypic diversity and its molecular basis remain to be fully characterised.

Epigenetic changes Cancer cells have an altered epigenotype compared with the tissues from which they arise, and this epigenetic switch is characterised by changes in the level and placement of DNA methylation and histone modification, but much work remains to be done to resolve the role of epigenetic changes such as DNA methylation in the maintenance and reversibility of promotion and progression.[16]

It is, for example, known that DNA methyltransferase 1 is overexpressed during tumour development. Hypomethylation of long interspersed nuclear elements (LINE-1 repeats), satellite DNA and moderately repeated DNA sequences is accompanied by hypermethylation of localised promoter-associated CpG islands which are usually unmethylated in normal cells, rendering them transcriptionally silent. In addition, deamination of methylated cytosine residues forms thymine, an error which is difficult to correct because DNA repair enzymes are unable to determine which base in the resulting mismatched pair is the correct one. Thus genome-wide hypomethylation can increase mutation rates leading to genomic instability in developing tumours, demonstrating the link between organisation of the genome and replication/repair. Changes in histone modifications are also observed in cancer; loss of lysine 16 acetylation and lysine 20 methylation in histone H4 is a hallmark of human cancer cells and is associated with hypermethylation at repetitive sequences.

Epigenetic alterations contribute to all phases of cancer development, including initiation, promotion, invasion, metastasis and resistance to chemotherapy; indeed, epigenetic biomarkers such as methylated DNA sequences, modified histones and microRNAs may be of value in diagnostic and prognostic monitoring. Furthermore, epigenetic processes are reversible: DNA methyltransferases and histone deacetylases therefore represent targets for chemotherapeutic inhibition. The antiretroviral drug azidothymidine is an example of a DNA methyltransferase inhibitor.

1.5 Key concepts in toxicology

1.5.1 Risk and hazard[17]

Evaluation of the harmful effects of substances in the environment may be based either upon hazard alone or upon an evaluation of risk. It always starts

[16] For an interesting, rather speculative discussion, see Szyf (2011).

[17] This section has been extracted from the Hazardous Substances Advisory Committee *Approach on Hazard and Risk Assessment of Substances* (2013) with the kind permission of the Chair and Secretariat. For the full statement, see https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/209709/hsac-hazard-risk-assessment-substances.pdf.

by identifying potential targets and routes of exposure. This is followed by hazard identification/characterisation with the option of proceeding to exposure assessment and risk identification/characterisation. A substance's potentially hazardous properties can include toxicity, persistence and ability to bio-accumulate. Risk is assessed by evaluating these properties against the predicted concentrations to which organisms are exposed. The hazard associated with a particular substance is its intrinsic ability to cause harm, while risk is the probability that such harm will occur in practice; it depends upon exposure, and the probability of risk from a particular hazard is almost always <100%.

The process involves three stages (i) hazard identification and characterization, (ii) exposure assessment and (iii) risk identification and characterisation, each of which is described in the following text:

- **Hazard identification and characterisation:** Screening tests must be amenable to testing many substances for
  - physicochemical properties (e.g. flammability, boiling point, water solubility and relative density); these determine how a substance behaves in different environmental compartments and influence its degradability and persistence;
  - toxicological properties for humans such as irritancy, sensitisation, acute and chronic effects including carcinogenicity, mutagenicity and reproductive effects; and
  - ecotoxicological properties including acute and chronic toxic effects on different species as well as processes such as bio-accumulation and bio-magnification.

  For ethical, time and cost-saving reasons hazards may initially be predicted from molecular structure (*in silico*), available information on structurally similar chemicals (read-across) and short-term tests. These are conducted *in vitro* wherever possible in order to avoid the unnecessary use of animals, and recent improvements in miniaturisation and robotics have allowed the development of a range of HTS approaches. As the objective is to identify potential hazards, these initial screening tests tend to be biased towards increasing sensitivity at the expense of specificity.
- **Exposure assessment:** It is exposure which determines whether a biological target, which may range from a single species to an entire ecosystem, is likely to experience adverse effects. Only if the target species is exposed to a harmful dose does a substance actually represent a risk.
- **Risk identification and characterisation:** The objective of risk identification and characterisation is to assess the probability and extent of occurrence of adverse effects of chemicals in the environment realistically, taking into account predicted exposure conditions over time. The potential for induction of adverse effects at the most sensitive stages of development (e.g. the embryo or foetus) is particularly important.

The validity and usefulness of any risk assessment are based on many factors. A major consideration is the quality and comprehensiveness of the underpinning scientific evidence. Good regulatory decisions are most likely to result from high quality evidence, taking into account all aspects of the risk assessment process.

1.5.2 Variability and uncertainty

If one measures a biological parameter, including toxicity, in a group of individuals the individual values obtained will differ because of two factors, variability and uncertainty.[18] In the context of toxicology these may be defined as follows:

- **Variability** is defined as observable diversity in biological sensitivity or response, and in exposure parameters. It is caused by inherent biological differences between species, strains, sub-strains and individuals and cannot be reduced. Variability in response to toxic chemicals is determined by the fate of the chemical within the body (toxicokinetics) and the toxicity of the chemical and its metabolites (toxicodynamics). Variability in both toxicokinetics and toxicodynamics relates to a combination of factors that are inherent to the organism, and other factors relating to the physiology and environment of the individual, which change over time. The inherent characteristics include species, sex and genotype. The modulating factors include the physiology of the individual (e.g. age, stage of development, disease or nutritional deficiency, environment and lifestyle factors) and other chemical exposures originating from diet and lifestyle. In principle, variability is measurable, and lack of knowledge of variability is a source of uncertainty (i.e. uncertainty about the variability).
- **Uncertainty** is defined as imperfect knowledge concerning the present or future state of an organism, system or (sub) population under consideration. Uncertainty refers to lack of knowledge, which can often be reduced by undertaking appropriate studies or by increasing the sophistication or power of studies.

Variability is also sometimes called type A uncertainty, aleatory uncertainty, inherent uncertainty or irreducible uncertainty; however, the term variability is preferable because variability can be reproducibly quantified and therefore is not actually uncertain. The variability of a parameter represents true heterogeneity amongst individuals which cannot be reduced and has to be taken into consideration during risk assessment. Examples of this heterogeneity include physiological conditions and states, for example, pregnancy, the functional maturation of organs, body weight and composition, respiratory rate and food consumption.

The sources of variability in human susceptibility to toxic chemicals include both toxicokinetic and toxicodynamic factors. Toxicokinetics includes the processes of intestinal uptake and transfer (i.e. absorption) of substances by the body, the biotransformation these substances undergo, the distribution of the substances and their metabolites between tissues, and the elimination of the substances and metabolites from the body, while toxicodynamics is the process of interaction of chemical substances with target sites and the subsequent events which lead to adverse effects. Each of the stages in toxicokinetics and toxicodynamics is a potential source of variability and uncertainty.

[18] For a comprehensive discussion of this topic, see the Committee on Toxicity *Report on Variability and Uncertainty in Toxicology of Chemicals in Food, Consumer Products and the Environment* on http://cot.food.gov.uk/cotreports/cotwgreports/cotwgvut.

Other sources of variability in response are more generic; they include physiological factors (e.g. age, nutritional status, obesity and exercise) and disease states (e.g. diabetes mellitus). Age can be a particularly important factor, and risk assessments should always consider the potentially increased susceptibility to toxic chemicals of the developing child (embryo, foetus and child). Infants may be sensitive or resistant because of immaturity of enzymes and other processes involved in the elimination of chemicals (toxicokinetics), whilst infants and young children may show increased sensitivity at receptors or other macromolecules where toxins exert their effects (toxicodynamics). Growth, changing body composition and differentiation of tissues and organs in early life, childhood and adolescence may lead to increased susceptibility to environmental agents compared with adults; on the other hand, young animals, including humans, may be able to metabolise some agents better than do adults, and have a better ability to recuperate or compensate for adverse events than adults, in whom the capacity for architectural and functional regeneration of organs and tissues is reduced. At the other end of life, the elderly may have increased susceptibility to some types of toxic insult because of declining liver and kidney function, and this may be exacerbated by the fact that the majority of elderly people are being treated with multiple therapeutic agents (so-called polypharmacy).

1.5.3 Threshold and non-threshold dose responses

The conventional approach to risk assessment involves the identification of a threshold dose (i.e. one below which there is no toxic response) in laboratory animals and then the application of a so-called uncertainty factor or "safety factor" to allow for the possibility that humans are more sensitive to the observed effect than the animals in which the tests were conducted. The identification of threshold doses is particularly important in chemical risk assessment because it is considered to be possible to identify a safe level of exposure for a chemical which exhibits a threshold response, whereas for a chemical which exhibits no threshold it is assumed that no amount of exposure can be considered safe.[19]

The parameters used in these calculations include:

- The no observed adverse effect level (NOAEL)
- Various benchmark doses (BMD)s

Under the REACH regulations, two new parameters, the Derived No Effect Level and Derived Minimal Effect Level, have been defined. The Derived No Effect Level is calculated from the NOAEL or an appropriate BMD and applies to effects which display a threshold, while the Derived Minimal Effect Level is used where no threshold can be defined and is considered to represent a low, possibly theoretical and hopefully tolerable risk.

[19] For more on threshold effects, see http://www.popstoolkit.com/riskassessment/module/exposure+and+toxicity+analysis/toxicity/ threshold+contaminants.aspx and http://www.toxicology.org/isot/RC/northland/blackburn.pdf.

Most chemically induced toxic effects exhibit a threshold, and it is therefore assumed to be possible to set a safe dose (e.g. an acceptable daily intake) based on a NOAEL with the application of appropriate safety and uncertainty factors. However, according to the initiation-promotion-progression theory of carcinogenesis, a single mutation could theoretically initiate a tumour; in other words, there may be no threshold for tumour induction. In the absence of a threshold, conventional threshold-based risk assessments cannot be applied; one solution to this problem is to adopt the margin of exposure (MoE) approach developed as a result of work sponsored by the European Food Safety Authority (EFSA) (Benford *et al*., 2010). The MoE is the ratio of a risk assessment parameter obtained from animal studies (e.g. a dose giving a particular incidence of tumours) to the anticipated exposure level. In the United States, an MoE of >100,000 is considered to indicate low risk, MoEs between 10 000 and 100 000 indicate moderate risk and a MoE <10,000 indicates high risk. The European approach is less rigid, and chemicals are considered on a case-by-case basis.

The vast number of compounds to which humans are exposed to and the cost and resources needed in order to assess all their hazards make it impossible to conduct a complete risk assessment for every chemical. A pragmatic approach to this problem is to attempt to identify a generic level of exposure which is considered acceptable for any chemical, the threshold of toxicological concern (TTC) (Munro *et al*., 2008). This approach is based on the concept that a reasonable assurance of safety can be given, even in the absence of chemical-specific toxicity data, provided that the exposure is sufficiently low: i.e. one can identify an exposure below which there is no significant risk to human health.

Proposed TTCs have been based on the knowledge gained from 50 to 60 years' worth of risk assessments and balance the uncertainties inherent in extrapolation from existing toxicology data against a predicted or known low level of exposure. The approach was initially used for food packaging materials and subsequently extended to a wider range of chemicals. The original TTC for food packaging materials and additives was set at 0.5 ppb (equivalent to 1.5 μg/person/day) by the US Food and Drug Administration (FDA).

The development of TTCs for other compounds has attempted to take into account the molecular structures of the compounds under consideration. Compounds are categorised into three classes (Munro *et al*., 2008):

- **Class I:** Simple structures which are efficiently metabolised and have low potential toxicity.
- **Class II:** Less clearly innocuous than Class I, but no positive indications of toxicity or significant uncertainties.
- **Class III:** Have structural features which preclude the assumption of safety or suggest significant toxicity.

Statistical analysis of NOAELs indicated that TTCs could be set at 1800 μg/person/day for Class I, 540 μg/person/day for Class II and 90 μg/person/day for Class III (based on a body weight of 60 kg). Additional TTCs of 0.15 μg/person/day and 18 μg/person/day were defined for potential genotoxins and organophosphates, respectively. These values are now commonly used to evaluate food flavourings and are in the process of being extended to other categories of compounds,

although this can be problematic in some cases: for example, the application of the TTC of 0.15 μg/person/day to genotoxic impurities in medicinal products would result in the rejection of most drugs because of the use of highly reactive reagents in their synthesis. Overall, the TTC approach is considered to be a pragmatic solution to the problem of assessing the safety of large numbers of compounds for which limited data are available. It reduces the number of unnecessary animal tests and is conceptually simple and easy to apply. This approach is now accepted as part of the risk assessment of food flavourings and packaging materials and is in the process of being extended to other substances.

A concept which goes beyond thresholds is that of hormesis. This is based on the observation that the dose responses for some effects (including some genotoxic processes) can be J-shaped, that is, higher values observed at extremely low/zero doses than at low conventional doses. This may be explicable in terms of, for example, the induction of DNA repair enzymes by low doses of genotoxic agents leading to repair of endogenous lesions which would be left unrepaired in untreated animals/cells. It can be difficult to determine the shape of dose–response curves at very low doses in hazard assessment studies because the number of doses used is often inadequate, the dose spacing may be inappropriate and few doses below the NOAEL are considered, so it is difficult to prove or disprove the existence of hormesis. Some commentators believe that hormesis should be given much more credence during risk assessment, and it has, indeed, been suggested that hormesis should be the default assumption (Calabrese, 2005). However, this is still the subject of heated toxicological debate (Calabrese *et al.*, 2011, 2012; Zeiger and Hoffmann, 2012).

1.5.4 The regulatory context

The role of the toxicologist is to ensure the safety of people who may be exposed to chemicals via the environment, in the workplace or voluntarily through using cosmetic ingredient, consumer products or pharmaceuticals. The data generated and reported by toxicologists must therefore be entirely reliable, and this requires total integrity on the part of everybody involved. While one would hope that all toxicologists work to the highest professional and ethical standards, this has not always been the case, and it has been necessary to implement regulations to ensure the maintenance of appropriate standards in the industry. Furthermore, in order to compare the potential toxic effects of different substances, data must be generated in a standardised manner. This is facilitated by the provision of test guidelines by a variety of international agencies.

The main regulations addressing the issue of data generation and reporting are the Good Laboratory Practice (GLP) regulations. These were first drafted in response to scandals such as the one relating to the initial safety assessment of the artificial sweetener aspartame in the early 1970s, when the manufacturer Searle was accused of bad practice in the conduct of toxicity tests and misrepresentation of the results obtained. The GLP regulations do not provide step-by-step instructions for toxicology testing; rather, they specify various generic aspects of good practice including the use of Standard Operating Procedures, the need for verification at all stages of a laboratory investigation, and criteria

for record keeping, reporting and archiving. Since the implementation of the GLP regulations, additional regulations have been drawn up: these include Good Manufacturing Practice for the manufacture of medicinal products, Good Distribution Practice regulations covering transportation of medicinal products, Good Clinical Practice regulations for clinical trials and Good Pharmacovigilance Practice relating to the monitoring of the safety of medicines. In the United Kingdom, the agency responsible for invigilating these regulations is the Medicines and Healthcare Products Regulatory Authority (MHRA).[20]

International bodies including the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) and the Organisation for Economic Co-operation and Development (OECD) provide formal regulatory guidelines specifying procedures for various aspects of pharmaceutical manufacturing and *in vivo*/*in vitro* testing. The ICH Guidelines are categorised under the headings of Quality Guidelines, Safety Guidelines, Efficacy Guidelines and Multi-disciplinary Guidelines,[21] while OECD Guidelines cover tests for the physical–chemical properties of chemicals, environmental effects, degradation and accumulation in the environment, human health effects and other areas.[22]

When conducting tests for regulatory purposes it is essential to ensure that they are conducted according to the appropriate guidelines and are, where appropriate, compliant with the GLP regulations; regulatory authorities such as the MHRA in the United Kingdom, the European Medicines Agency (EMA) in Europe and the US FDA specify the circumstances under which GLP-compliant data generated according to international guidelines are required.

1.5.5 Limitations of whole animal studies

In vivo studies have a number of limitations which have led to ever-increasing efforts to develop better and more informative alternatives. Alternative approaches include endpoint assays, tissue slices, toxicokinetic modelling, structure–activity relationships, database interrogation and tissue culture (Kniewald *et al*., 2005). The available *in vitro* systems for studying target organ toxicity include perfused organ systems, precision-cut tissue slices, established cell lines, isolated cells in suspension and primary cell cultures. The most important factors in selecting an *in vitro* system are retention of differentiated functions and demonstration of good *in vivo–in vitro* correlations for the phenomena of interest. The greatest progress in the use of *in vitro* test systems has been in the areas of target organ toxicity (particularly hepatic, renal and neural toxicity) and local ocular–dermal toxicity.

The drive to reduce animal usage in toxicity testing has led to the establishment of international bodies to oversee the move towards the use of alternatives. These include the European Centre for the Validation of Alternative Methods (ECVAM),[23] which was established in 1991 in response to an EU Directive

[20] http://www.mhra.gov.uk/Howweregulate/Medicines/Inspectionandstandards/index.htm
[21] http://www.ich.org/products
[22] http://www.oecd.org/env/ehs/testing/oecdguidelinesforthetestingofchemicals.htm
[23] http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam

mandating active support for the development, validation and acceptance of methods to replace, reduce or refine the use of animals in laboratories and its US equivalent, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM).[24] Both these organisations support progress towards the reduced use of animals in toxicity testing by monitoring the evaluation of new methods and organising collaborative validation programmes.

3Rs and the place of *in vitro* tests in safety evaluation The 3Rs principle mandates the Replacement, Reduction and Refinement of the use of animals in scientific experimentation. The need for good *in vitro* toxicity tests has become even more urgent following the implementation of the REACH regulations in 2007. The need to test the toxicity of an estimated 30,000 existing chemicals that are currently marketed in volumes greater than 1 tonne per year has led to predictions that as many as four million animals will be used in testing unless suitable alternative methods become available (Clemedson, 2008).

The use of *in vitro* methods in regulatory toxicity testing has conventionally been limited to defining the potential for *in vivo* toxicity and verifying negative results. The effort to further develop and validate better *in vitro* alternatives has become increasingly urgent following the enactment of the REACH regulations in June 2007 and the 7th Amendment in March 2009. In particular, the 7th Amendment has prohibited animal testing of cosmetic ingredients, including the use of animals for skin irritation, corrosion, and genotoxicity testing, since March 2009. Furthermore, the use of animals for absorption, distribution, metabolism and excretion (ADME) testing of cosmetic ingredients has been banned since March 2013.

The conventional view is that *in vitro* tests are suitable for hazard identification but not for the other aspects of risk assessment, and the use of *in vitro* tests has tended to comply with these perceived limits. In the pharmaceutical industry, however, strenuous efforts have been made to overcome these limitations and predictions of bioavailability, pharmacokinetics and pharmacodynamics are now integrated into the early stages of drug development, being conducted concomitantly with efficacy tests. The results obtained are used to refine the process of selection of drug candidates and this approach has been found to save time, money and resources in the early stages of drug development.

The development of integrated testing strategies is key to the implementation of the 3Rs in risk assessment. This entails using a range of different types of information (of the kinds described in the remainder of this book) and combining such different inputs rationally in regulatory decision-making. Animal testing should be a final resort, carried out only in cases where insufficient pre-existing information is available. Before doing so, a number of questions should be addressed:

- Can a TTC approach be justified?
- Is there scope for the use of an integrated testing strategy or an *in silico* approach?

[24] http://iccvam.niehs.nih.gov/

- Is it possible to extrapolate from pre-existing information, including data from the literature or other information which is in the public domain? While this may not have been generated according to GLP standards, it can be used in some settings; for example, the REACH regulations do not require all the information in a dossier to meet GLP criteria.

It is important to optimise *in vivo* assays in order to obtain the best possible data without causing undue pain and suffering.[25] Annexes VII to X in the REACH regulations provide specific examples of situations in which animal testing can be avoided by taking advantage of pre-existing information. The sequence in which tests are performed is critical: it is important to give thought to planning a sequence of tests which will generate the maximum amount of information with the lowest severity and least suffering to animals.

In order to identify the chronic effects of a test compound, suitably long-lived *in vitro* systems, representing all cell types of each organ/tissue are needed. These must be capable of providing an estimate of the potential impact of repeated exposure on those organs/tissues. Any effects in these *in vitro* systems will have to be assessed by reference to kinetic and modelling data which allow the prediction of potential tissue and organ interactions which may be important in risk assessment. The final stage of risk assessment is the evaluation of risk to the consumer from exposure which may arise from the intended use of the chemical. This requires establishment of a quantitative link between the results of the *in vitro* assays and human exposure.

1.5.6 Use of human tissues in toxicology

The use of human tissue for *in vitro* toxicology is becoming increasingly popular because it is considered to provide more relevant data than do animal models. Human tissue-based methods also avoid the animal welfare concerns inherent in conventional animal-based toxicology studies, although they have their own problems relating to the legal, ethical and informed consent issues surrounding human tissue acquisition.

Three main sources of human tissue are typically used in drug development:

- **Tissues residual to surgery:** tissues not required for diagnosis or which are generated by cosmetic procedures can be accessed rapidly and stored as fresh, fixed or frozen tissues.
- **Tissues and organs from transplant procedures:** organ donation rightly takes precedence over research; however, many organs cannot be used in a transplant procedure and may be consented for use in medical research.
- **Tissues retrieved post-mortem:** these tissues are most often frozen or fixed and used in target discovery or identification.

The regulatory environment surrounding the use of human tissues in UK research has become much more demanding following the implementation of the Human Tissue Act (2004) and the Human Tissue (Scotland) Act (2006). In particular there is now a requirement for all end users to register and have a designated

[25] For a detailed discussion of this issue, see Madden *et al.* (2012).

person to deal with approved procedures and record keeping. Key facts relating to the Human tissue Act (2004) are summarised in Box 1.1.

Box 1.1 Human Tissue Act 2004[26]

The Human Tissue Act 2004 replaced the Human Tissue Act 1961, the Anatomy Act 1984 and the Human Organ Transplants Act 1989 as they relate to England and Wales, and the corresponding Orders in Northern Ireland.

The Human Tissue Act 2004 covers England, Wales and Northern Ireland. It established the HTA to regulate activities concerning the removal, storage, use and disposal of human tissue. Consent is the fundamental principle of the legislation and underpins the lawful removal, storage and use of body parts, organs and tissue. Different consent requirements apply when dealing with tissue from the deceased and the living. The Human Tissue Act 2004 lists the purposes for which consent is required (these are called *Scheduled Purposes*).

There is separate legislation in Scotland - the Human Tissue (Scotland) Act 2006. While provisions of the Human Tissue (Scotland) Act 2006 are based on authorisation rather than consent, these are essentially both expressions of the same principle.

Further information about the Human Tissue (Scotland) Act 2006[27]

The key points of the Human Tissue Act 2004

The Human Tissue Act 2004 regulates the removal, storage and use of human tissue. This is defined as material that has come from a human body and consists of, or includes, human cells.

The Human Tissue Act 2004 creates a new offence of DNA 'theft'. It is unlawful to have human tissue with the intention of its DNA being analysed, without the consent of the person from whom the tissue came.

The Human Tissue Act 2004 makes it lawful to take minimum steps to preserve the organs of a deceased person for use in transplantation while steps are taken to determine the wishes of the deceased, or, in the absence of their known wishes, obtaining consent from someone in a qualifying relationship.

Offences under the Human Tissue Act 2004

Removing, storing or using human tissue for Scheduled Purposes without appropriate consent.

Storing or using human tissue donated for a Scheduled Purpose for another purpose.

Trafficking in human tissue for transplantation purposes.

Carrying out licensable activities without holding a licence from the HTA (with lower penalties for related lesser offences such as failing to produce records or obstructing the HTA in carrying out its power or responsibilities).

Having human tissue, including hair, nail and gametes (i.e. cells connected with sexual reproduction), with the intention of its DNA being analysed

[26] http://www.hta.gov.uk/legislationpoliciesandcodesofpractice/legislation/humantissueact.cfm, Downloaded with the permission of the Human Tissue Authority.
[27] http://www.hta.gov.uk/_db/_documents/Information_about_HT_%28Scotland%29_Act.pdf

without the consent of the person from whom the tissue came or of those close to them if they have died. (Medical diagnosis and treatment, criminal investigations, etc. are excluded).

The first four offences only apply in England, Wales and Northern Ireland, although the Human Tissue (Scotland) Act 2006 has similar offences and penalties. The offence of DNA theft applies UK-wide. To find out more about the offences you can download the Human Tissue Act 2004[28] and Human Tissue (Scotland) Act 2006.[29]

(Source: Reproduced with permission of the Human Tissue Authority)

The remaining issues surrounding the routine use of human tissues in toxicology include:

- **The need to ensure a regular, reliable supply:** In particular, suitable infrastructure must be in place if using fresh tissues and it may be necessary to have an on-call rota of individuals trained to receive and process tissue at short notice.
- **Inter-individual variability:** This may be perceived to be both a limitation of the use of human tissues (since it will increase the variability of the results of *in vitro* studies) and an advantage (since this represents the true range of responses within the general population).
- **Regulatory acceptability:** Regulatory bodies are seeing more and more data from *in vitro* studies with human tissue; this is therefore becoming less of an issue as time moves on.

The expert view is that *in vitro* human tissue has great potential as an experimental system in various aspects of drug development as well as in other aspects of toxicology, but that there is still room for improvement and expansion of its application (Clotworthy, 2012).

1.6 Summing up

This chapter should have provided sufficient background to allow the readers to put into context the molecular and cellular processes described in subsequent chapters. The further reading recommended here is intended to help readers who have found any of the concepts unfamiliar, and therefore focuses on key pathology and toxicology textbooks containing background information relevant to the remaining chapters of this book.

Self-assessment questions

- It could be said that 'Apoptosis is a double-edged sword'. Do you agree with this statement?
- What are the ethical issues involved in the use of human tissues for research purposes?

[28] http://www.legislation.gov.uk/ukpga/2004/30/contents
[29] http://www.legislation.gov.uk/asp/2006/4/contents

Background reading

Greim, H. and Snyder, R. (2008) Toxicology and Risk Assessment: A Comprehensive Introduction, John Wiley and Sons, Ltd: Chichester, UK; ISBN: 978-0-470-86893-5.

Niesink, R.J.M. *et al*. (1995) Toxicology: Principles and Applications, CRC Press ISBN: 0-8493-9232-2.

Stevens, A. *et al*. (2009) Core Pathology, 3rd Edition Mosby/Elsevier Ltd ISBN: 978-0-7234-3444-3.

References

Abreu-Blanco, M.T., Watts, J.J., Verboon, J.M. and Parkhurst, S.M. (2012) Cytoskeleton responses in wound repair. *Cellular and Molecular Life Sciences*, **69**, 2469–2483.

Armitage, P. and Doll, R. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *British Journal of Cancer*, **8**, 1–12.

Auten, R.L. and Davis, J.M. (2009) Oxygen toxicity and reactive oxygen species: the devil is in the details. *Pediatric Research*, **66**, 121–127.

Benford, D., Bolger, P.M., Carthew, P. *et al*. (2010) Application of the margin of exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. *Food and Chemical Toxicology*, **48** (Suppl 1), S2–S24.

Brenneisen, P., Steinbrenner, H. and Sies, H. (2005) Selenium, oxidative stress, and health aspects. *Molecular Aspects of Medicine*, **26**, 256–267.

Calabrese, E.J. (2005) Should hormesis be the default model in risk assessment? *Human and Experimental Toxicology*, **24**, 243.

Calabrese, E.J., Stanek, E.J. III, and Nascarella, M.A. (2011) Evidence for hormesis in mutagenicity dose-response relationships. *Mutation Research*, **726**, 91–97.

Calabrese, E.J., Stanek, E.J. III, and Nascarella, M.A. (2012) A detailed re-assessment supports the conclusion of the paper that hormesis is commonly observed in the Ames assay. *Mutation Research*, **747**, 157.

Clemedson, C. (2008) The European ACuteTox project: a modern integrative *in vitro* approach to better prediction of acute toxicity. *Clinical Pharmacology and Therapeutics*, **84**, 200–202.

Clotworthy, M. (2012) The application of human tissue for drug discovery and development. *Expert Opinion on Drug Discovery*, **7**, 543–547.

Cohen, S.M. and Arnold, L.L. (2011) Chemical carcinogenesis. *Toxicological Sciences*, **120** (Suppl 1), S76–S92.

D'Alessandro, A., Rinalducci, S. and Zolla, L. (2011) Redox proteomics and drug development. *Journal of Proteomics*, **74**, 2575–2595.

French, S.W. (2010) Molecular events in hepatic preneoplasia: a review. *Experimental and Molecular Pathology*, **88**, 219–224.

Fulda, S. (2009) Caspase-8 in cancer biology and therapy. *Cancer Letters*, **281**, 128–133.

Gonzalez De Mejia, E., Ramirez-Mares, M. V., Arce-Popoca, E., Wallig, M., *et al.* 2004. Inhibition of liver carcinogenesis in Wistar rats by consumption of an aqueous extract from leaves of Ardisia compressa. *Food and Chemical Toxicology*, **42**, 509–516.

Goodlett, C. R. & Horn, K. H. 2001. Mechanisms of alcohol-induced damage to the developing nervous system. *Alcohol Res Health*, **25**, 175–84.

Han, S.I., Kim, Y.S. and Kim, T.H. (2008) Role of apoptotic and necrotic cell death under physiologic conditions. *BMB Reports*, **41**, 1–10.

Hooker, D.J., Mobarok, M., Anderson, J.L. *et al.* (2012) A new way of measuring apoptosis by absolute quantitation of inter-nucleosomally fragmented genomic DNA. *Nucleic Acids Research*, **40**, e113.

Imlay, J.A. (2008) Cellular defenses against superoxide and hydrogen peroxide. *Annual Review of Biochemistry*, **77**, 755–776.

Kagan, V.E., Wipf, P., Stoyanovsky, D. *et al.* (2009) Mitochondrial targeting of electron scavenging antioxidants: regulation of selective oxidation vs random chain reactions. *Advanced Drug Delivery Reviews*, **61**, 1375–1385.

Kang, K.S. and Trosko, J.E. (2011) Stem cells in toxicology: fundamental biology and practical considerations. *Toxicological Sciences*, **120** (Suppl 1), S269–S289.

Khansari, N., Shakiba, Y. and Mahmoudi, M. (2009) Chronic inflammation and oxidative stress as a major cause of age-related diseases and cancer. *Recent Patents on Inflammation & Allergy Drug Discovery*, **3**, 73–80.

Kniewald, J., Kmetic, I., Gaurina-Srcek, V. and Kniewald, Z. (2005) Alternative models for toxicity testing of xenobiotics. *Arhiv za Higijenu Rada i Toksikologiju*, **56**, 195–204.

Kowaltowski, A.J., De Souza-Pinto, N.C., Castilho, R.F. and Vercesi, A.E. (2009) Mitochondria and reactive oxygen species. *Free Radical Biology and Medicine*, **47**, 333–343.

Lee, S., Park, Y., Zuidema, M.Y. *et al.* (2011) Effects of interventions on oxidative stress and inflammation of cardiovascular diseases. *World Journal of Cardiology*, **3**, 18–24.

Limon-Pacheco, J. and Gonsebatt, M.E. (2009) The role of antioxidants and antioxidant-related enzymes in protective responses to environmentally induced oxidative stress. *Mutation Research*, **674**, 137–147.

Loeb, L.A. (2011) Human cancers express mutator phenotypes: origin, consequences and targeting. *Nature Reviews Cancer*, **11**, 450–457.

Madden, J.C., Hewitt, M., Przybylak, K. *et al.* (2012) Strategies for the optimisation of *in vivo* experiments in accordance with the 3Rs philosophy. *Regulatory Toxicology and Pharmacology*, **63**, 140–154.

Mitrus, I., Bryndza, E., Sochanik, A. and Szala, S. (2012) Evolving models of tumor origin and progression. *Tumour Biology*, **33**, 911–917.

Moolgavkar, S.H. and Knudson, A.G. Jr. (1981) Mutation and cancer: a model for human carcinogenesis. *Journal of the National Cancer Institute*, **66**, 1037–1052.

Munro, I.C., Renwick, A.G. and Danielewska-Nikiel, B. (2008) The threshold of toxicological concern (TTC) in risk assessment. *Toxicology Letters*, **180**, 151–156.

Murphy, M.P. (2009) How mitochondria produce reactive oxygen species. *Biochemistry Journal*, **417**, 1–13.

Murray, G.I., Taylor, M.C., McFadyen, M.C. *et al.* (1997) Tumor-specific expression of cytochrome P450 CYP1B1. *Cancer Research*, **57**, 3026–3031.

Oliveira, P.A., Colaco, A., Chaves, R. *et al.* (2007) Chemical carcinogenesis. *Anais da Academia Brasileira de Ciências*, **79**, 593–616.

Pan, J.S., Hong, M.Z. and Ren, J.L. (2009) Reactive oxygen species: a double-edged sword in oncogenesis. *World Journal of Gastroenterology*, **15**, 1702–1707.

Pitot, H.C. (2007) Adventures in hepatocarcinogenesis. *Annual Review of Pathology*, **2**, 1–29.

Pryor, W.A., Houk, K.N., Foote, C.S. *et al.* (2006) Free radical biology and medicine: it's a gas, man! *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, **291**, R491–R511.

Rasola, A. and Bernardi, P. (2011) Mitochondrial permeability transition in Ca(2+)-dependent apoptosis and necrosis. *Cell Calcium*, **50**, 222–233.

Roberts, R.A., Laskin, D.L., Smith, C.V. *et al.* (2009) Nitrative and oxidative stress in toxicology and disease. *Toxicological Sciences*, **112**, 4–16.

Servais, H., Ortiz, A., Devuyst, O. *et al.* (2008) Renal cell apoptosis induced by nephrotoxic drugs: cellular and molecular mechanisms and potential approaches to modulation. *Apoptosis*, **13**, 11–32.

Szyf, M. (2011) The implications of DNA methylation for toxicology: toward toxicomethylomics, the toxicology of DNA methylation. *Toxicological Sciences*, **120**, 235–255.

Van Vliet, E. (2011) Current standing and future prospects for the technologies proposed to transform toxicity testing in the 21st century. *Altex*, **28**, 17–44.

Weber, G.F. (2008) Molecular mechanisms of metastasis. *Cancer Letters*, **270**, 181–190.

Zeiger, E. and Hoffmann, G.R. (2012) An illusion of hormesis in the Ames test: statistical significance is not equivalent to biological significance. *Mutation Research*, **746**, 89–93.

Zhivotovsky, B. and Orrenius, S. (2011) Calcium and cell death mechanisms: a perspective from the cell death community. *Cell Calcium*, **50**, 211–221.

2

Individual Susceptibility to Toxic Chemicals

2.1 Introduction

The fact that individuals differ in their response to toxic agents has been known for a very long time.[1] The phenomenon of favism, for example, was first described by the Ancient Greeks. Favism is the most common human enzyme deficiency and describes the phenomenon whereby haemolysis occurs in certain individuals on ingestion of fava beans due to a deficiency of the enzyme glucose-6-phophate dehydrogenase. The defect is sex-linked, being transmitted from mother (usually a healthy carrier) to son (or daughter, who would be a healthy carrier too). It is the commonest human enzyme deficiency, affecting some six million people, and exists in more than 400 genetic variants. Agents that can cause haemolysis in glucose-6-phophate dehydrogenase-deficient individuals include antimalarial drugs (primaquine, pamaquine, pentaquine and chlorophine), sulfonamides, nitrofurans (e.g. furadantin), acetaminophen, naphthalene, certain vitamin K derivatives, acetylsalicylic acid and probenecid (Benemid). The favism-inducing toxins in beans are believed to be divicine and isouramil, the aglycone moieties of vicine and convicine.

Modern interest in the role of xenobiotic metabolising enzymes in differences between individuals'susceptibility to the adverse effects of drugs and chemicals began with several serendipitous observations, including the fact that so-called slow acetylators had increased susceptibility to isoniazid-induced neuropathy and hydralazine-induced systemic lupus erythematosus and the dramatic effects of the antihypertensive drug debrisoquine in an individual who subsequently turned out to be what is now called a CYP2D6 poor metaboliser (a laboratory worker studying the metabolism of the drug took a normal therapeutic dose and collapsed due

[1] For a review of pharmacogenetics in history and today, see Pirmohamed (2011).

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

to a catastrophic drop in blood pressure). These, and other, observations led to the initiation of the discipline now known as toxico- (or pharmaco-) genetics.

This chapter will introduce the technologies used in toxicogenetics and exemplify their impact on susceptibility to chronic and acute toxic effects. The current status of the field will be evaluated and possible future developments will be identified.

2.2 Toxicogenetics and toxicogenomics

The disciplines of toxico/pharmacogenetics and toxico/pharmacogenomics address genetic variability by considering the role of an individual's genetic makeup in determining the effects of xenobiotics to which they may be exposed.[2] Toxicogenetics is the study of genetic variation in specific candidate genes relevant to a particular toxic response and is based on an existing paradigm such as the proposed mechanism of carcinogenesis. It is relatively inexpensive and easy to perform with current technologies, and because it arises from a biological hypothesis the biological consequences of the results obtained are usually relatively easy to determine. This is how the study of genetic influences on individual susceptibility to toxic agents started back in the 1960s. Toxicogenomics is the study of genetic variation across the whole genome in an unbiased manner. This involves the simultaneous study of many genetic variants (currently 300 000 to 1.2 million markers have been identified in the human genome), and is by definition an unbiased approach, but its disadvantage is that the functional consequences of the observations made are usually unknown.

The study of genetic variation in drug responsiveness has evolved from candidate gene studies to studies of variation across the whole human genome (Madian *et al*., 2012). Initial successes included the identification of key variants in xenobiotic metabolising enzymes which affect responses to drugs and carcinogens. The field subsequently expanded to include the analysis of regulatory regions and to be able to look at more subtle effects. In recent years it is considered to have progressed from hypothesis-driven approaches to open-ended approaches which do not involve any preconceptions, although the philosophical question of whether this actually represents progress is beyond the scope of this book.

The causes of genetic variability in response to therapeutic (or toxic) agents can include interindividual differences in drug targets, which may have a direct effect on responsiveness, and in drug metabolism, which may lead to differences in the rate of clearance of a drug. Such individual variation is often initially observed serendipitously when individuals suffer so-called idiosyncratic responses to drugs and xenobiotics. Differences in therapeutic responses are also sometimes identified during clinical trials but in many cases these are insufficiently large to provide a statistical likelihood of identifying very rare events and adverse side effects are often not reported until a drug has been on the market for some time.

[2] The pharmaco- and toxico- versions of these terms are essentially synonymous except that pharmacogenetics/genomics addresses both beneficial (therapeutic) and adverse effects whereas toxicogenetics/genomics focuses on adverse effects, which may be either short- or long-term. The terms *toxicogenetics* and *toxicogenomics* will be used in this chapter unless specifically referring to a therapeutic effect, but it should be taken that the general aspects of the discussion are equally relevant to pharmacogenetics/genomics.

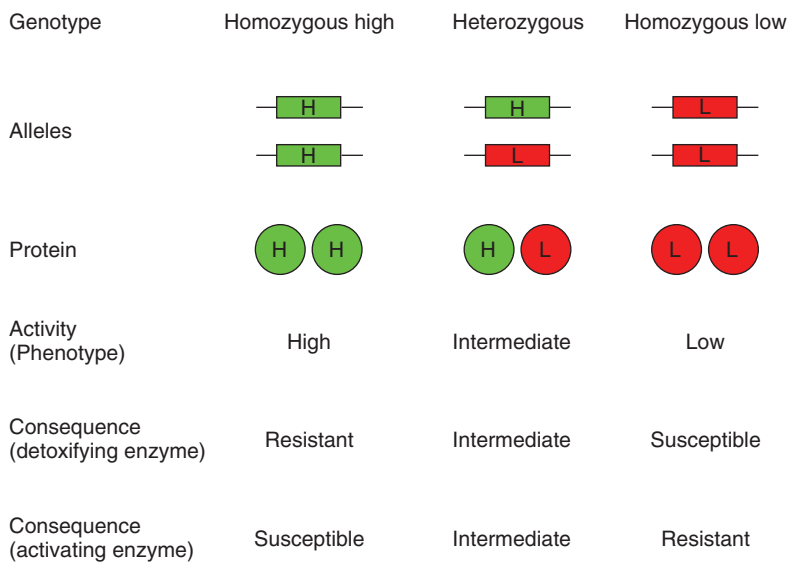


Figure 2.1 Genotype and phenotype and their consequences. This figure illustrates the possible consequences of polymorphism in a hypothetical xenobiotic metabolising enzyme that has both a high activity and a low activity variant. In the case of an individual who is homozygous for the high activity variant, two high activity alleles will be present in the genome (genotype) and these will be transcribed and translated to form entirely high activity enzyme molecules. Thus the individual's liver will have high enzyme activity (phenotype). If the enzyme in question is a detoxifying enzyme, the individual is predicted to have reduced susceptibility to toxicity, whereas if it is a metabolic activating enzyme the individual may have increased susceptibility. The reverse is the case for an individual who is homozygous for the low activity variant while a heterozygous individual has one high and one low activity allele, a mixture of high and low activity enzyme molecules in the liver, and hence an intermediate phenotype (source: Stanley (2008); figure 2.39. Reproduced with permission of John Wiley & Sons, Ltd)

Toxicogenetic variation is determined by the genes which encode proteins responsible for the metabolic processing of, and response to, xenobiotics. Each gene is found at a specific locus. At each locus there may be one or more possible sequence variants, or alleles. The existence of more than one possible allele at a particular locus is called a *polymorphism*, and loci where multiple alleles exist are described as polymorphic. A polymorphism is defined as a sequence variant which is present in at least 1% of the population. Humans, being diploid, have two copies of each locus (i.e. they carry two alleles of each gene), and the combination of alleles carried by an individual defines his or her genotype at that locus. If both copies are the same, the individual is said to be homozygous at this locus; if two different alleles are present, the individual is said to be heterozygous.

The consequences of polymorphic variation at a hypothetical allele are illustrated in Figure 2.1.

2.3 Genotyping and phenotyping

Many polymorphisms affect the amino acid sequence of a protein, but some affect the non-coding regulatory regions of genes. Polymorphisms which affect

the amino acid sequence may be detected either at the level of nucleotide sequence or amino acid sequence. Alternatively, those which affect function may be detected directly by measuring the function of the protein (e.g. its enzyme activity). This is the phenotype; it is a function of the underlying DNA sequence but may also be influenced by regulatory, physiological and environmental factors as well as by pathological conditions. Polymorphisms in non-coding regions may affect the level of expression of a particular protein and can be detected at the nucleotide sequence level or by looking at mRNA or protein expression. The complete absence of a gene may also be classified as a polymorphism.

2.3.1 Genotyping

Techniques which look at the nucleotide sequence directly are called *genotyping methods*, whereas those which address expression and function are known as *phenotyping approaches* (Daly, 2004). Genotyping involves examining the altered nucleotide sequences of polymorphic variants in the DNA itself. This approach traditionally involved cloning and sequencing the gene of interest from different individuals, but recent technological developments mean that genotyping is now usually undertaken using polymerase chain reaction (PCR)-based methods.

SNPs The vast majority (99.5%) of the human genome is common to everybody; the differences between individuals (three million bases) are mainly single nucleotide polymorphisms (SNPs), which are estimated to account for as much of 90% of human genetic variations.[3] On an average, there is about one SNP for every 300 bases, meaning that an average-sized gene of 25 kb may contain up to 80 SNPs, but they are not evenly spread through the genome: there are SNP hot spots where the density of SNPs is much higher than in other regions, and 99% of SNPs are found in regions which do not contain genes. The most common types of SNP are transitions (purine-purine (C-T) and pyrimidine-pyrimidine (A-G)). Since 5-methylation of cytosine is the commonest base modification in DNA, and 5-methylcytosine can be deaminated to form thymine, C to T SNPs are the most common type in the human genome.

The changes caused by SNPs may be silent, harmless or latent. Most SNPs are located in non-coding/non-regulatory regions of the genome and are functionally silent; it is currently estimated that only about 2000 cause an actual amino acid change. These missense or non-synonymous SNPs alter the amino acid sequence of the cognate protein and may affect its function, whereas synonymous SNPs do not have any effect on coding information, although if located within a regulatory region they can affect responses to drugs and toxic chemicals.

Most SNPs are biallelic, that is, there are only two possible variants (usually a major and a minor variant), so they are relatively easy to type using automated methods. In addition, they exhibit a fairly low rate of recurrent mutation, which means that they act as stable markers of human evolutionary history. They are usually detected by PCR-based methods, which have conventionally included allele-specific oligonucleotide (ASO) hybridisation, restriction fragment length polymorphism (RFLP), allele-specific PCR (ASPCR) and PCR followed by

[3] For a concise review in the context of cancer chemoprevention, see Velasquez and Lipkin (2005).

direct sequencing (Box 2.1). The fact that PCR products are usually quite short (a few hundred nucleotides) means that they are easy to sequence, and recent advances in sequencing technology and automation have made it possible to sequence the products of hundreds of PCR reactions simultaneously. As a consequence, PCR followed by direct sequencing is now the commonest method used to detect SNPs.

Box 2.1 PCR-based methods for detecting SNPs

Allele-specific oligonucleotide hybridisation

Allele-specific oligonucleotide hybridization depends on the fact that, at the melting temperature (T_m) of a probe-target DNA duplex, any mismatched probes can be washed off. It uses a short oligonucleotide probe, ideally about 19 nucleotides long, which will specifically recognise either the normal (wild-type) sequence or a specific sequence variant. The best probes to use for allele-specific oligonucleotide hybridisation are antisense probes; in other words, they bind to the sense strand. The method used is as follows:

1. Amplify a region of DNA which has the sequence variant near the middle.
2. Transfer the PCR products to nylon membranes by slot/dot blotting (similar to a Southern blot).
3. Incubate the membranes with radioactively labelled probes: a separate membrane for each probe. One probe matches the normal (wild-type) sequence; there is a probe to match each possible sequence variant.
4. Wash the membranes at exactly T_m. Perfectly matched sequences will remain bound whereas mismatched sequences will wash off.
5. Expose the membranes to autoradiography film to visualise the results.

As an example, consider this sequence:

CGAGTAACGTGCATGC**T**AGCTAGTCAGTCA

which can be mutated to:

CGAGTAACGTGCATGC**G**AGCTAGTCAGTCA

Two probes are required, for example, GCACGTACG**A**TCGATCAGT and GCACGTACG**C**TCGATCAGT

To work out the melting temperature of the duplex:

$$T_m(\text{in}^\circ\text{C}) = (2 \times \text{AT}) + (4 \times \text{GC})$$

$$\text{For probe 1}: T_m = (2 \times 8) + (4 \times 10) = 56^\circ\text{C}$$

$$\text{For probe 2}: T_m = (2 \times 7) + (4 \times 11) = 58^\circ\text{C}$$

Restriction fragment length polymorphism
The principle of RFLP depends on the recognition of specific sequences by restriction enzymes. Restriction enzymes recognise palindromic sequences which are called *restriction sites*. For example,

AG**G**TCCCACCT
TC**C**AGGGTGGA

If a restriction site is mutated, the enzyme will no longer cleave the DNA. For example,

AG**T**TCCCACCT
TC**A**AGGGTGGA

In RFLP analysis, the region of interest is amplified, as for allele-specific oligonucleotide hybridisation. The PCR products are then incubated with a restriction enzyme and run on a gel. The wild-type sequence will be cleaved (generating two bands) whereas the variant sequence is not (so only one band is seen). This method depends on being able to find a restriction enzyme which recognises the sequence variant in which you are interested.

Allele-specific PCR
A PCR reaction will only work if the primers are perfectly matched to the DNA sequence. It is possible to take advantage of this to screen for sequence variants. This time the primers are designed so that one of them ends just at the point of the sequence variant.

For example, using the sequence we looked at before (CGAGTAACGT GCATGC**T**AGCTAGTCAGTCA), which can be mutated to CGAGTAACG TGCATGC**G**AGCTAGTCAGTCA, two different 5′ primers may be made:

- Primer 1: CGAGTAACGTGCATGC**T**
- Primer 2: CGAGTAACGTGCATGC**G**

The same 3′ primer is used as before and the annealing temperature is set to be close to T_m. In this case, Primer 1 will only amplify the wild-type sequence and Primer 2 will only amplify the variant sequence.

Currently, the favoured high throughput sequencing (HTS) method for SNP detection is direct DNA sequencing, which has been made feasible by the spectacular improvements in sequencing technology over the last 20 years. However, other high throughput genotyping methods (loosely defined as techniques involving the use of automation to increase the throughput of an experimental procedure) such as array-based methods are becoming increasingly popular.[4] Improvements in technologies initially made it quicker to run and detect single SNPs in 96-sample format using methods such as Taqman® real-time PCR (RT-PCR) than in individual Eppendorf-type tubes; now it is possible to detect

[4] Reviewed by Bhasker and Hardiman (2010).

multiple SNPs on one sample, with multiplexing or array-format genotyping. This is becoming much more cost effective (less than 10 p per genotype), which is essential to make it a feasible approach for large studies. For example, the Wellcome Trust Case-Control Consortium[5] study covering 13 disease areas has over 60 000 samples to analyse, so minimal cost per sample is a key consideration.

In DNA microarray technology, 20 000 to 100 000 unique DNA molecules are robotically applied to the surface of silicon wafers which are approximately the size of a microscope slide. Microarray tools such as the TaqMan OpenArray®[6] are now in regular use for DNA sequence analysis, genotyping, and molecular diagnostics as well as for monitoring gene expression. The TaqMan OpenArray® genotyping plate is a 63-mm × 19-mm mid-density reaction plate containing 3072 reaction through-holes, each of which takes a 33-nl reaction volume. Individual through-holes are preloaded with a TaqMan® assay so that all that has to be done is to add DNA and reaction mix and run the PCR. Multiple samples can be run on a single array, and a variety of arrays is available; these can be specific for specific classes of genes (e.g. drug metabolism arrays) or can be custom made.

2.3.2 Phenotyping

Genotyping is becoming increasingly popular as a means of evaluating toxicogenetic variation because of the advent of high capacity automated PCR methods. However, it is important to remember that it is the phenotype which determines function and will mediate any observed effects on susceptibility. Genotype is determined by the DNA sequence and is unchanging, whereas phenotype may change during life due to developmental changes in gene expression and, potentially, xenobiotic-induced effects such as induction or enzyme inhibition. Traditional methods for phenotyping xenobiotic metabolising enzymes involve measuring enzymatic activity using diagnostic substrates. The advantage of this approach is that it directly addresses the actual function of the enzyme and does not require assumptions about genotype–phenotype correlations, but it has become less popular in recent years because it is more labour-intensive than genotyping.

The genotype can be determined by looking at any cell type because all the cells of an individual, except for red blood cells and germ cells, contain the same genomic DNA sequence. This means that it is not necessary to examine genotype in the target tissue in which toxicity is expressed, and that more accessible surrogate tissues can be used. However, in order to look at levels of expression and function, it is important to identify and access the tissue in which the phenotype is actually expressed. In the case of hepatic enzymes, this would entail obtaining a liver biopsy sample, which is clearly impracticable when examining healthy populations. Instead, white blood cells are usually used for this purpose since they may be obtained by relatively non-invasive methods; however, it must be recognised that this might not reflect the situation in the liver or the target tissue.

Many investigators prefer to use mRNA-based techniques such as TaqMan® RT-PCR, which is a simple and rapid method to quantify the expression of a

[5] http://www.wtccc.org.uk/

[6] http://tools.invitrogen.com/content/sfs/brochures/cms_057219.pdf

particular mRNA rather than using labour-intensive protein-based approaches. However, the results of these studies must be examined critically because differences in mRNA levels may not be reflected at the protein, and hence functional, level.

2.3.3 Correlating genotype and phenotype

The drawback of genotyping methods in toxicogenetics is that they provide only an indirect measure of the actual activity of the enzyme or protein of interest. To interpret the results it is necessary to make the assumption that in a given individual the phenotype is a direct function of the genotype. In order to justify this assumption it is necessary to demonstrate a clear correlation. In some cases, such as the aromatic amine metabolising enzyme *N*-acetyltransferase Type 2 (NAT2), a clear correlation has been established. However, in other cases the relationship between genotype and phenotype is more complex.

Example: The difficulty of relating genotype to *in vitro* and *in vivo* phenotype

There is a large variation in activity levels of serum paraoxonase 1, an enzyme that is involved in the metabolism of organophosphate pesticides such as diazoxon, between individuals. This is important because people with very low paraoxonase 1 activity have increased susceptibility to organophosphate toxicity; in addition, polymorphic variation in paraoxonase 1 has been implicated in susceptibility to a number of diseases including (controversially) Parkinson's disease (Furlong *et al.*, 2010). Two polymorphic sites, at codons 55 (Leu^{55} to Met^{55}) and 192 (Gln^{192} to Arg^{192}) have been identified within the human paraoxonase 1 gene. These are relatively common and lead to structural changes in the protein. The literature on the consequences of these polymorphisms is contradictory due to an artefact generated by the method used to assay the activity. When assayed under published conditions, that is, in the presence of 2M NaCl, the Gln^{192} isoform was more active towards diazoxon than the Arg^{192} isoform, but when injected into paraoxonase 1 null mice both isoforms gave similar levels of protection against a diazoxon challenge. Further experimentation, using more physiological salt concentrations, revealed that high concentrations of NaCl differentially inhibit the Arg^{192} isoform of paraoxonase 1 and that the concentration of substrate was also critical (O'Leary *et al.*, 2005).

Another point arising from the example of paraoxonase 1 is the importance of considering actual function, as well as genetics, when evaluating the relevance of metabolic polymorphisms as risk factors in susceptibility to the adverse consequences of exposure to xenobiotics. In the case of paraoxonase 1, there was more variation in activity between individuals having the same genotype than there was between the activities associated with different genotypes, meaning that genotyping alone is insufficient to determine the role of paraoxonase 1 in risks associated with exposure to organophosphates.

To summarise, if genotyping methods are to be used as a way of evaluating the role of polymorphisms in toxic processes, it is very important that a good correlation has to be established between genotype and phenotype. In order to understand the effects of polymorphisms on processes in toxicity, it is essential to

generate the highest quality data using the best possible study design and methodology and the use of validated assays is critical. Furthermore, the interpretation of the data generated depends upon a sound knowledge of the range of polymorphisms which exist at any given locus. Ideally, one should be confident that all the possible allelic variants of the gene in question have been identified, and considerable work is still required in order to complete this database.

2.4 Polymorphic xenobiotic metabolism

The evaluation of tissue-specific detoxification and metabolic activation is highly challenging because of the complexity of xenobiotic metabolising systems, with multiple pathways of metabolism occurring in a single tissue, different patterns of expression in different tissues and distinct but related enzymes in different species. The existence of multiple genes, the broad specificity of individual enzymes, generation of diversity by differential splicing and the formation of heterodimers adds to these challenges. Furthermore, interindividual variability can arise as a result of polymorphisms which are often linked to disease susceptibility.

The primary function of xenobiotic metabolism is to render hydrophobic chemicals more hydrophilic so that they can be excreted (Figure 2.2); however, this often involves the generation of highly reactive intermediates and a single enzyme can mediate both activation and detoxification. Xenobiotic metabolism is therefore a balancing act between detoxification and metabolic activation (Figure 2.3).

When a cell is exposed to a xenobiotic, four metabolic processes take place:

- **Phase 0:** passage of a chemical through the cell membrane into the cytosol
- **Phase I:** metabolism introducing a reactive group into the molecule
- **Phase II:** transfer of polar groups onto the products of Phase I metabolism
- **Phase III:** transport of water-soluble metabolites of Phase I/II metabolism out of the cell

Phase I and II xenobiotic metabolism is mediated by a diverse group of enzymes. The key enzyme families involved are summarised in Figure 2.4. In Phase I metabolism, a highly lipophilic compound such as a PAH is oxidised by addition of a hydroxyl or epoxide group; in Phase II, the resulting highly reactive electrophile is conjugated to a hydrophilic group such as glutathione, making it sufficiently water-soluble to be excreted.

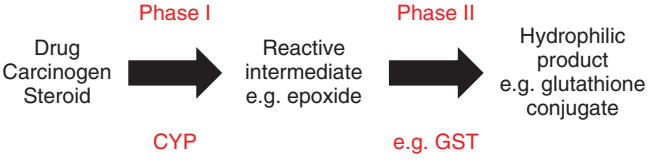


Figure 2.2 The role of xenobiotic metabolising enzymes in metabolic activation and detoxification

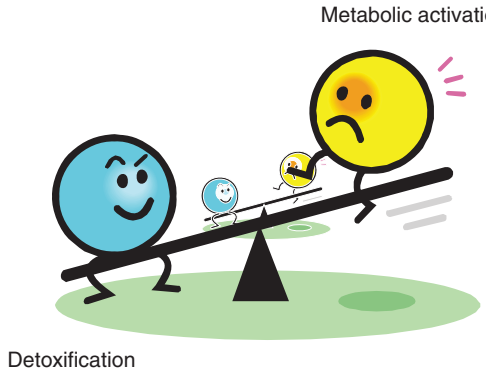


Figure 2.3 Xenobiotic metabolism is a balancing act! If detoxification outweighs metabolic activation, the xenobiotic will be removed safely from the body. If metabolic activation outweighs detoxification, the consequence may be protein binding or DNA binding leading to cytotoxicity, DNA damage or other toxic effects

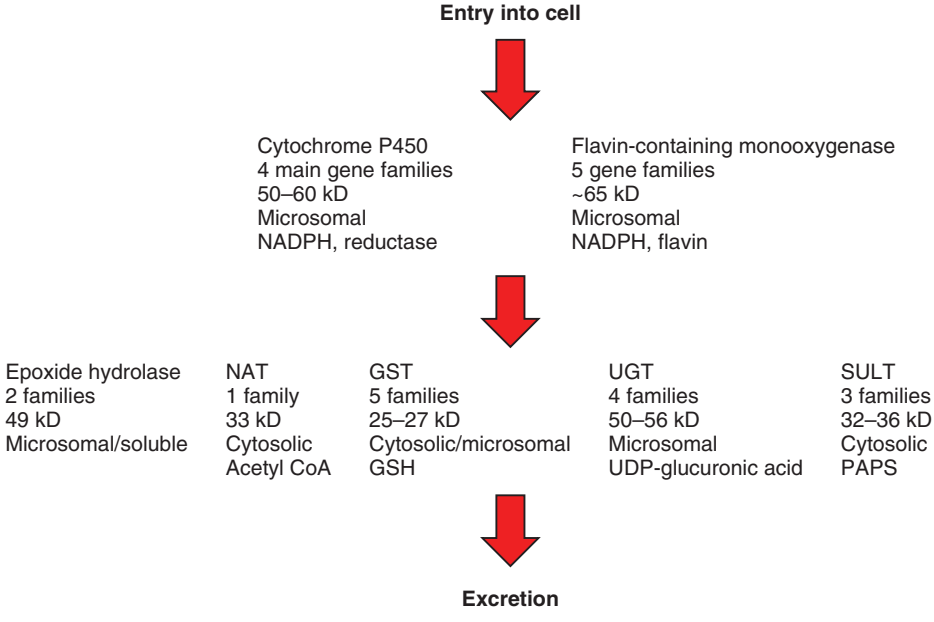


Figure 2.4 Key xenobiotic metabolising enzyme families

Phase I metabolism is dominated by CYP enzymes, which incorporate one atom from molecular oxygen into a substrate leading to the generation of reactive intermediates such as epoxides and are involved in both detoxification and activation of xenobiotics. Box 2.2 provides a brief introduction to CYP enzymology and nomenclature. Members of the CYP1, CYP2 and CYP3 gene families are generally involved in the metabolism of xenobiotics; they account for 75–80% of Phase I xenobiotic metabolism and 65–70% of the clearance of clinically used

drugs. Polymorphism in this gene family plays a key role in a variety of adverse drug reactions as well as in responsiveness to cancer chemotherapy.[7]

Box 2.2 Cytochrome P450 enzymology and nomenclature

The cytochrome P450 enzyme family, a group of nicotinamide adenine dinucleotide phosphate (NADPH)-dependent microsomal enzymes involved in the biotransformation of azo dyes, was identified by James and Elizabeth Miller and their associates in the late 1940s. The reactions catalysed by these enzymes also proved to be important in the processing of other carcinogens as well as drugs and steroid hormones, and subsequent work identified the agent responsible as a liver pigment which bound carbon monoxide. This pigment, a b-type cytochrome containing iron protoporphyrin IX as a prosthetic group, was named 'cytochrome P450' because the complex between the reduced pigment and carbon monoxide had an absorption maximum of 450 nm.

Cytochrome P450 is not a single entity but comprises a superfamily of related enzymes which are widely distributed in animals, plants and protists. The human genome contains 57 active CYP genes which fall into 18 families. Mammalian CYPs are integral membrane proteins of molecular weight 50–60 kD which are found in high concentrations in the endoplasmic reticulum and in lower concentrations in the mitochondrial, nuclear and plasma membranes. The organ containing the highest concentration of CYPs is the liver; however, organs which are accessible to compounds from the environment, including skin, GI tract and lung, contain significant levels of some CYPs.

The members of the CYP family have broad, overlapping specificities with regard to both the substrates metabolised and the sites of attachment within the substrate molecule. Many biotransformation reactions involving both xenobiotic and endogenous substrates are dependent on CYP catalysis. The most common reaction catalysed is mono-oxygenation, but CYPs can also mediate epoxidation, peroxygenation, N-, S- and O-dealkylation, N- and S-oxidation, dehalogenation, desulfurisation, reduction of nitro-, azo- and N-oxide groups, peroxides, epoxides and deamination. The xenobiotic substrates metabolised include many drugs, carcinogens and probe substrates; endogenous CYP substrates include fatty acids, prostaglandins, leukotrienes, thromboxanes, steroid hormones and bile acids.

The CYP superfamily is thought to be derived from a single ancestral gene which may have existed early in evolution: CYPs have been characterised in some very primitive organisms, including the bacterium *Pseudomonas putida*, whose camphor-metabolising enzyme was the first CYP to have its three-dimensional structure determined by X-ray crystallography. The CYP superfamily is thought to have arisen as a result of a series of duplication events; rapid evolution of this family occurred, especially after the divergence of the mammals.

In spite of the rapid evolution of CYPs, all members of this superfamily retain a number of features in common and exhibit a high level of sequence

[7] Reviewed by Johansson and Ingelman-Sundberg (2011).

homology. They are therefore classified into gene families on this basis, each gene family being allocated a number while members of a gene family are classified using letters (for subfamily) and numbers (for each individual member) (Nelson, 2009).[8] CYPs are named by a nomenclature committee which was formed in 1999 and is responsible for allocating allele designations; it maintains a web page for each of the genes for which it is responsible (currently 29 CYPs plus cytochrome P450 reductase). Investigators who discover a new CYP variant are required to provide the committee with sequence information, at least covering exons, intron–exon junctions and linkages to other variants, preferably prior to submission of a manuscript. This information is held in confidence until publication or authorisation of release.

The nomenclatures allocated by the committee include family and subfamily designations plus a number indicating the specific member. Thus, for example, the major ethanol-metabolising CYP enzyme in humans is designated CYP2E1 for Family 2, subfamily E, member 1. The corresponding gene is *CYP2E1* (since gene names are conventionally italicised). Mouse genes are given lower case names, so the mouse homologue of CYP2E1 is called Cyp2e1 (gene: *Cyp2e1*).

Although CYP-mediated metabolism is essential to the detoxification of many compounds, the highly reactive intermediates formed during this process are often more toxic than the parent compounds and have the capacity to bind to cellular macromolecules such as proteins and DNA, causing severe cellular damage. This process often results in mutation and can lead to carcinogenesis. Thus, metabolic activation can occur as a result of Phase I metabolism because CYP enzymes can generate reactive metabolites that are more toxic than the original substrate.

Phase II metabolism is responsible for the conjugation of xenobiotics, which typically leads to more water-soluble and therefore more easily excretable compounds.

- **Glutathione *S*-transferases (GSTs):** Many reactive electrophilic compounds form conjugates with glutathione. These conjugation reactions are catalysed by GSTs, although strong electrophiles may also react non-enzymatically with GSH. All cells contain GSH in concentrations which may reach 10 mM in some tissues. Glutathione conjugates of xenobiotics are degraded to *N*-acetylated cysteine thioethers (mercapturic acids), which are subsequently excreted.
- **UDP-glucuronyl transferases (UGTs)** catalyse the transfer of D-glucuronic acid from UDP-glucuronic acid to functional groups of aliphatic and aromatic alcohols, carboxylic acids, amines, hydroxylamines, amides and thiols and have broad-substrate specificities. These reactions give rise to *O*-, *N*-, *S*- and *C*-glucuronides. In this reaction a β-glycosidic bond is formed between the aglycone (xenobiotic) and the glucuronic acid. UGTs are also essential for the glucuronidation, and subsequent excretion, of bilirubin.
- **Sulphotransferases (SULTs)** use 3′-phosphoadenosine-5′-phosphosulphate as sulphate donor and are involved in the metabolism of endogenous substrates such as biogenic amines and steroid hormones. The sulphonates produced

[8] For the latest updates on CYP nomenclature, see the Cytochrome P450 Homepage (http://drnelson.uthsc.edu/CytochromeP450.html).

exist preferentially in the ionised form at physiological form at physiological pH, which markedly increases the water solubility of the xenobiotic; however, sulphonation may also give rise to toxic reactive metabolites.

- ***N*-acetyltransferases (NATs)** are cytosolic enzymes which mediate the detoxification of aromatic amines and hydrazines via *N*-acetylation. Their substrates include aromatic amines, hydrazines, hydrazides, sulfonamides, some aliphatic primary amines and hydroxylamines. *N*-acetylation is an important detoxifying process for most of these (especially aromatic amines), but NATs may also mediate the metabolic activation of carcinogenic heterocyclic amines by catalysing their *O*-acetylation. Humans have two functional NAT genes *(NAT1* and *NAT2*) found, along with the pseudogene *NAT*p, on chromosome *8p*22.

Epoxides are detoxified to diols by epoxide hydrolases, which protect cells from the formation of both DNA and protein adducts by these very reactive species which are found in the diet or formed during Phase I metabolism. In the course of the oxidative metabolism of olefins and aromatic compounds, the reactive epoxides that are formed are substrates of epoxide hydrolase. Generally, the hydrolysis of epoxides represents a detoxification reaction in xenobiotic metabolism, but in the case of certain PAHs, epoxidation followed by epoxide hydrolase-mediated hydrolysis comprises a metabolic activation process.

In summary, therefore,

- Phase I metabolism generates reactive intermediates.
- In the absence of adequate Phase II metabolism these may:
  - bind to DNA forming nucleotide adducts and possibly acting as carcinogens
  - bind to proteins forming peptide adducts and possibly acting as immunogens
- In the presence of active Phase II enzymes the consequences may be:
  - further activation, e.g. formation of diol epoxides
  - formation of hydrophilic products ready to be excreted by efflux transporters

2.4.1 Polymorphic xenobiotic metabolising enzymes

Polymorphisms exist among both Phase I and Phase II xenobiotic metabolising enzymes. The CYP gene family, for example, is very diverse, comprising 57 genes in 18 gene families (Sim and Ingelman-Sundberg, 2010). Where polymorphisms do exist, a star number is allocated to each polymorphic variant. Thus, for example, the designation *CYP2C9*2* represents the following:

- Cytochrome P450
- Family 2
- Subfamily C
- Member 9
- Polymorphic variant 2

The basis of naming of other xenobiotic metabolising enzymes is similar, except that it is simpler where subfamilies do not exist; thus, for example, *NAT2*4* represents the following:

- *N*-acetyltransferase
- Family 2
- Polymorphic variant 4

Humans have traditionally been classified as either rapid or slow acetylators depending upon their NAT2 phenotype; in most Caucasian populations approximately 60% of individuals are slow acetylators. This variant is the main rapid acetylator variant of NAT2.

2.4.2 The role of xenobiotic metabolising polymorphisms in susceptibility to toxic agents

Various metabolic polymorphisms are known to have a significant impact on individual susceptibility to toxic and carcinogenic compounds, although the literature in this area can often be contradictory. In this chapter, rather than presenting a long list of polymorphisms which may or may not be involved in susceptibility to chemically induced disease processes and toxicity, a few of the better characterised examples will be presented.

Example: Interindividual variability in response to the anticoagulant warfarin

Individuals' responses to medication vary both in terms of beneficial effects and toxicity, leading to a range of possible responses:

- Favourable response and acceptable tolerability (the desired outcome).
- Little or no response (drug does not work).
- Serious adverse effects (to be avoided at all costs).

Personalised drug therapy is especially desirable for drugs which have a narrow therapeutic index, particularly when the consequences of toxicity are life-threatening. One such drug is warfarin, a coumarin anticoagulant which is one of the most widely used treatments for thromboembolic (blood clotting) diseases worldwide.[9] Indeed, it is the most widely used drug worldwide for the treatment of thromboembolic events including atrial fibrillation and deep vein thrombosis.

Warfarin is a good example of a drug where toxicogenetic factors affecting both toxicokinetics (metabolism) and toxicodynamics (mechanism of action) influence responses to therapy.[10] There is enormous variation in individuals' responses to treatment with warfarin and it can take a very long time (sometimes months) to identify a dose which will adequately control blood clotting. There is a risk of excessive bleeding if the initiating dose is too high, while control of clotting is ineffective if the dose is too low. Warfarin acts by antagonizing the vitamin K cycle, an essential part of the process of blood clotting, inhibiting the regeneration of reduced vitamin K which is an essential cofactor for the clotting cascade.

[9] Reviewed by Rettie and Tai (2006).
[10] For a concise review, see Cavallari (2012).

The dose of warfarin required for stable clot control is dependent on both genetic and non-genetic factors. The non-genetic factors involved include age, gender, body mass index and previous drug history while the key genetic factors are the CYP isoform CYP2C9 and the enzyme VKORc02, which catalyses the rate determining step of the vitamin K cycle, controlling the clotting process.

Warfarin is a natural product and is administered as racaemic mixture of the R- and S-stereoisomers of the drug. S-warfarin is three to five times more potent as an inhibitor than R-warfarin. There are several metabolic clearance pathways for R-warfarin, but the main route by which S-warfarin is metabolised in the human liver is via CYP2C9, generating harmless 6- and 7-hydroxy metabolites which are excreted in the urine. CYP2C9 maps to chromosome 10q24.2 and contains nine exons encoding a typical 60 kD CYP protein. The reference sequence for CYP2C9 is CYP2C9*1, while the two most important CYP variants for warfarin dosing and prevention of adverse events are CYP2C9*2 and CYP2C9*3. These alleles encode Arg144Cys and Ile359Leu variants of the CYP2C9 protein, both of which are functionally defective (having 70 and 5% of the catalytic efficiency of CYP2C9*1, respectively). The allele frequencies of CYP2C9*2 and CYP2C9*3 are about 12 and 8%, respectively, meaning that approximately 40% of the Caucasian population will carry at least one of these alleles (they are much rarer in populations of African and Asian origin). Individuals with the CYP2C9*2 and CYP2C9*3 variants exhibit impaired warfarin clearance; they are likely to need lower doses of warfarin and have an increased risk of bleeding complications during therapy.

VKORc02 is the vitamin K epoxide reductase complex 1 gene, which contains three exons mapping to the short arm of chromosome 16 and encodes an 18-kD integral membrane protein, the enzyme vitamin K epoxide reductase. Homozygous coding region mutations in VKORc02 cause a severe condition called *multiple coagulation factor deficiency type II*, while individuals who are heterozygous for coding region mutations tend to be warfarin resistant (i.e. their blood still clots even after treatment with warfarin). However, these missense mutations are rare within the population (much less than 0.1%), so this form of genetic variation accounts for relatively little of the population variability in warfarin responsiveness. Polymorphisms in the regulatory region are a much more common source of variability, correlating strongly with warfarin responsiveness, and the estimated dose of warfarin is now tailored to a promoter SNP G3673A (rs9923231), which is located −1639 nucleotide base pairs upstream from the ATG start codon in the *VKORc02* gene. This polymorphism alters a VKORc02 transcription factor binding site and luciferase assays show that the activity of the G allele was increased by 44% over the activity of the A allele. Additionally, analysis of VKORc02 mRNA isolated from human liver samples showed that carriers of the A allele at position 3673 had reduced amounts of VKORc02 mRNA. Patients who have the AA genotype (or AA haplotype) are the most warfarin sensitive and therefore often require lower warfarin doses.

This is a good example of the potential benefit of personalised drug therapy. Polymorphisms in CYP2C9 account for about 10% and those in VKORc02 for about 15–30% of population variability in warfarin dose requirements in Caucasian and Asian populations. Overall, the combined effects of genetic variation in CYP2C9 and VKORc02 can explain up to 60% of variability in warfarin dose requirements, and it has been estimated that genetic testing for CYP2C9 and

VKORc02 could reduce the risk of hospitalization with warfarin toxicity by 30% (Madian *et al*., 2012). The International Warfarin Pharmacogenetics Consortium has developed an algorithm for estimating appropriate warfarin doses based on clinical and genetic data from a large population of patients (Klein *et al*., 2009). The factors taken into account were demographic characteristics, CYP2C9 and VKORc02 genotype, use of concomitant medications, and measured clotting parameters (target vs treatment international normalised ratio, a parameter based on prothrombin time) in relation to the stable therapeutic dose of warfarin required. Use of the algorithm predicted the required dose of warfarin better than a fixed dose approach or an algorithm based only on clinical parameters, especially in individuals requiring daily doses lower than 3 mg or higher than 7 mg (the normal daily dose usually falls between 2 and 5 mg kg^{-1}). The FDA now recommends that reduced doses of warfarin are considered in individuals known to have a variant CYP2C9 or VKORc02 allele, although it does not actually recommend genetic testing at the moment.

Example: Interindividual variability in susceptibility to industrial chemicals

Relatively few well characterised examples of acute toxicity affected by polymorphic xenobiotic metabolising enzyme currently exist. One of the best examples to date is acrylonitrile, an industrial solvent which is, in addition to being carcinogenic, acutely toxic (Thier *et al*., 2002; Hamdy *et al*., 2012). Acrylonitrile is metabolised via oxidative and reductive routes. The GSH-dependent detoxification pathway leads via the primary metabolites *S*-cyanoethylglutathione and *S*-methylglutathione to the mercapturic acids *N*-acetyl-*S*-cyanoethylcysteine and *N*-acetyl-*S*-cyanomethylcysteine, which are the final urinary excretion products. The alternative oxidative pathway of acrylonitrile metabolism, mediated by the isozyme CYP2E1, generates a reactive epoxide metabolite, cyanoethylene oxide (CEO), which is further metabolised to cyanide (Figure 2.5) (Wang *et al*., 2002). Both acrylonitrile and CEO also react with tissue thiols leading to GSH depletion. The acute toxicity of high doses of acrylonitrile is thought to be gated by the depletion of hepatic GSH, and this has led to the recommended use of *N*-acetylcysteine as an antidote to acrylonitrile poisoning.

There is significant inter- and intra-species variability in the acute toxicity of acrylonitrile. Humans are known to have a higher level of CYP2E1-mediated oxidative metabolism of acrylonitrile than rodents, although inducing agents such as acetone can be used to upregulate CYP2E1 activity in rats for use in experimental studies (Suhua *et al*., 2010), and this makes it difficult to extrapolate directly from the results of animal experiments to the effects of acrylonitrile in humans. In addition, humans have an active epoxide hydrolase pathway which leads to the generation of cyanide from CEO whereas this is not the case in rodents. Oxidative metabolism, which leads to the formation of cyanide, seems to be much less important in animals than in humans. The acute toxicity of acrylonitrile in animals seems to be mediated by GSH depletion whereas, in humans, acute toxicity is largely determined by the metabolic formation of cyanide.

The formation of an *N*-(cyanoethyl)valine adduct at the *N*-terminus of haemoglobin can be used as a biomarker which is indicative of blood levels of acrylamide (Thier *et al*., 1999). By measuring levels of *N*-(cyanoethyl)valine in relation to acrylonitrile exposure, higher adduct levels have been demonstrated

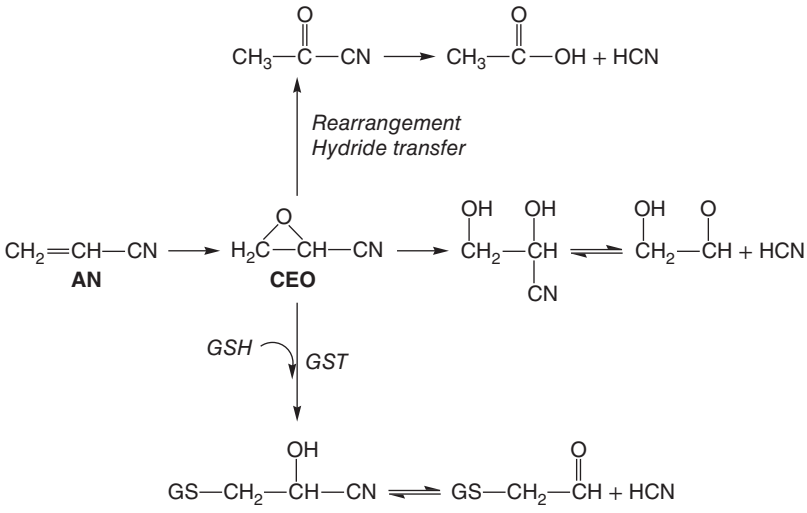


Figure 2.5 A proposed scheme showing the role of cytochrome P450 and epoxide hydrolase enzymes in the metabolism of acrylonitrile to cyanide (source: Wang *et al.* (2002); figure 1. Reproduced with permission of the American Society for Pharmacology and Experimental Therapeutics)

in individuals with at least one copy of a CYP2E1 promoter variant ($A_{316}G$) (Thier *et al.*, 2002). This may be associated with individual variation in the inducibility of CYP2E1. Slower CYP2E1-mediated metabolism of acrylonitrile in some individuals could lead to lower levels of metabolic activation (to CEO and cyanide) and increased blood levels of the parent compound, reflected in higher levels of the biomarker.

Initially, studies on individual cases of acrylonitrile intoxication provided clues about the role of metabolic polymorphisms in the detoxification of acrylonitrile in humans (Thier *et al.*, 2000). When the effects of acute acrylonitrile exposure were compared in two individuals, one with low and one with high GST activity, the individual with low activity experienced headache, nausea and vomiting. Furthermore, the level of hydrocyanic acid in his blood was within the lethal range, although fortunately he recovered following treatment with the antidote, *N*-acetylcysteine. This was consistent with the hypothesis that, particularly in individuals with low GST activity, toxicity is gated by GSH depletion. If insufficient GSH is available for conjugation (or the activity of GST is too low), free acrylonitrile may become available and gain entry into the CYP2E1-mediated oxidative pathway leading to toxicity.

Biomonitoring studies have suggested a role for the GST isozyme GSTP1 in the human metabolism of acrylonitrile. Individuals with the wild-type form of GSTP1 (GSTP1*A) appear to have a lower level of *N*(cyanoethyl)valine adducts (i.e. they detoxify acrylamide more efficiently) compared with those carrying a polymorphism at codon 104 of this gene (GSTP1*B and GSTP1*C) (Thier *et al.*, 2001). This is thought to be due to altered affinity of the enzyme for electrophilic substrates: the wild-type GSTP1*A enzyme efficiently conjugates GSH to electrophilic substrates (i.e. it has a low K_m and a high K_{cat}/K_m ratio) whereas the

GSTP1*B and GSTP1*C variants have higher K_m and lower K_{cat}/K_m values. This would be expected to lead to less efficient detoxification in the individuals who carry variant forms of GSTP1 and potentially to increased susceptibility to acrylonitrile toxicity. Furthermore, multiplex analysis has implicated the combined genotypes of CYP2E1, GSTM1 and mEH4 in urinary excretion and the formation of acrylamide and glycidamide adducts in individuals exposed to acrylamide in the workplace (Huang *et al*., 2011, Huang *et al*., 2012). The prediction of blood concentrations by means of pharmacokinetic modelling also has an important role to play (Takano *et al*., 2010).

These results suggest that the biomonitoring of industrial exposure to acrylonitrile should be supported by genotyping and illustrate the way in which the combination of toxicogenetic data with pharmacokinetic modelling and biomonitoring of actual exposures could, in the future, be used to identify and protect susceptible subgroups within the exposed subpopulation.

Example: Interindividual variability in susceptibility to bladder carcinogens

Aromatic amines such as 4-aminobiphenyl (4-ABP) are used in a number of industrial processes and are also found in cigarette smoke. Bladder cancer is associated with occupational exposure to aromatic amines in the rubber, textile, dye and chemical industries. Smoking is also a major risk factor (~66% of bladder cancers in Western countries are attributable to cigarette smoking) and this is thought to be due to the presence of aromatic amines as well as PAHs and nitrosamines in cigarette smoke.

Carcinogenic aromatic amines are metabolised by *N*-acetylation. Individuals may be classified as rapid or slow acetylators depending upon the variant of NAT2 they carry.[11] Acetylation by NAT2 was an early example of pharmacogenetic variation in that the half-life of the anti-tuberculosis drug isoniazid has been known for many years to be shorter in rapid acetylators than slow acetylators, leading to reduced therapeutic efficacy, whereas isoniazid-induced polyneuropathy is associated with the slow acetylator phenotype of NAT2 (Kinzig-Schippers *et al*., 2005). The slow acetylator phenotype is also associated with susceptibility to hydralazine-induced systemic lupus erythematosus, and vigilance is required when prescribing hydralazine for the chronic treatment of hypertension (Finks *et al*., 2006).

One of the reasons that the role of the NAT2 polymorphism in mediating the effects of xenobiotics is relatively well understood is that convenient methods exist for measuring the activity of this enzyme both *in vitro* and *in vivo*. The most commonly used *in vitro* probe substrate for NAT2 is sulfamethazine, but for *in vivo* phenotyping it is convenient to use caffeine since volunteers may be administered this compound without inciting concerns regarding toxicity.

In addition to the availability of convenient and reliable phenotyping methods, various methods are available for the determination of NAT2 genotype. From the early 1990s, methods based on PCR using combinations of allele-specific PCR and RFLP were developed (see Hickman *et al*., 1992). At the time these were considered to be a great advance but nowadays they seem very slow and labour-intensive, and given that there are now known to be at least 65 genetic

[11] For a detailed discussion of polymorphic variants of NATs and their consequences, see Hein, 2009.

Table 2.1 Biochemical consequences of some amino acid substitutions found in NAT2

NAT2 variant	Amino acid change	Consequence
G191A	$Arg^{64}Gln$	Reduced activity and protein level due to protein instability; interactions at Arg64 are required for structural stability
T341C	$Ile^{114}Thr$	Reduction in active enzyme level due to enhanced protein degradation; no change in intrinsic stability or kinetic parameters
G364A	$Asp^{122}Asn$	Asp^{122} is part of the catalytic triad so this polymorphism has a direct effect on activity; this variant may also have increased susceptibility to enzyme acetylation, increasing proteasomal degradation
A434C	$Gln^{145}Pro$	Reduced protein level with no change in intrinsic stability, possibly due to a change in secondary structure leading to proteasomal degradation
G590A	$Arg^{197}Gln$	Reduced activity and protein levels due to reduced protein thermostability
G857A	$Gly^{286}Gln$	Reduced activity for some substrates but not others. Reduced protein and thermal stability

(source: Data from Hein (2009))

variants of NAT2 it is clearly not possible to use this method to look at all of these. However, seven key SNPs have been identified and found to identify most of the variant alleles. Indeed, studies using panels of two, three, four or all seven SNPs, in combination with a proposed tag SNP which has been proposed as an activity biomarker, have demonstrated that genotyping for four SNPs (191 G>A, 341 T>C, 590 G>A and 857 G>A) predicts phenotype (as measured using sulfamethazine) with at least 98.4% accuracy in a range of ethnic populations (Hein and Doll, 2012, Suarez-Kurtz *et al*., 2012). This was as good as using the seven SNP panel and better than any of the other panels or the tag SNP.

The availability of convenient genotyping methods and the fact that a clear correlation between genotype and phenotype was quickly established for this enzyme have facilitated epidemiological analysis of the NAT2 polymorphism. The deduction of NAT2 phenotypes from genotypes is based upon the assumption that rapid and slow alleles are codominant; this means that as well as rapid acetylators (with two rapid alleles) and slow acetylators (with two slow alleles), it is also possible to identify heterozygotes (i.e. people with one rapid and one slow allele) as an intermediate group. There is further heterogeneity in the slow acetylator phenotype because each SNP has a distinct effect on the activity and stability of the corresponding protein (Table 2.1).

In an influential study published in the 1980s, it was found that 22/23 (95.7%) dye factory employees (or ex-employees) with bladder cancer were slow acetylators, whereas only approximately 60% of controls were slow acetylators (Cartwright *et al*., 1982). Although subsequent studies found somewhat lower slow acetylator frequencies in bladder cancer patients (Risch *et al*., 1995) (Table 2.2), the observation that slow acetylators have an increased risk of bladder cancer is now well established, and has been confirmed by meta-analysis (Garcia-Closas *et al*.,

Table 2.2 Epidemiological analysis of the NAT2 polymorphism in bladder cancer

	No. of slow acetylators/total in study	
Exposure	Patients	Controls
Dye factories	22/23 (95.7%)	–
Urban	74/111 (66.6%)	118/207 (57%)
Industrial	44/62 (71%)	–
No exposure	83/127 (65.4%)	26/59 (44.1%)

(source: Data from Cartwright *et al.* (1982) and Risch *et al.* (1995))

2005). This effect seems to be specific to smokers: NAT2 slow acetylators are especially susceptible to the adverse effects of smoking on bladder cancer risk whereas slow acetylation does not increase bladder cancer risk among subjects who have never smoked.

The role of NAT2 should not be considered in isolation since susceptibility to bladder cancer is also modulated by interaction with other xenobiotic metabolising enzymes. The metabolic activation of aromatic amines is illustrated in Figure 2.6. The first step in this process is hepatic *N*-hydroxylation by the CYP isozyme CYP1A2 (Kadlubar and Badawi, 1995). The resulting *N*-hydroxyl aromatic amines form GSH, glucuronide or sulphate conjugates (catalysed by GSTs, UGTs and SULTs, respectively). These soluble metabolites are transported to the bladder via the systemic circulation and may then be taken up by transitional epithelial cells. The conjugates may be hydrolysed, either under acidic conditions in the urine itself or as a result of intracellular hydroxylase activity, regenerating the original *N*-hydroxylamine.

The other NAT isozyme (NAT1) metabolises *p*-aminobenzoic acid as well as endogenous substrates such as *p*-aminobenzoyl glutamate, is expressed in a number of tissues including the urothelium (Stanley *et al.*, 1996), and is thought to play a key role in cellular homeostasis. Like NAT2, it can metabolise aromatic amines, but it is as yet unclear what the consequences of this polymorphism are for bladder cancer susceptibility. It is currently believed that NAT1 *O*-acetylates *N*-hydroxy aromatic amines *in situ* in the transitional epithelium, forming highly reactive *N*-acetyoxy esters which are able to bind directly to DNA, potentially leading to the initiation of carcinogenesis. This illustrates the way in which the same enzyme may mediate detoxification in some circumstances and metabolic activation in others: NAT1-mediated *N*-acetylation is a first-pass detoxification route for aromatic amine carcinogens in the skin whereas in the colon NAT1 mediates metabolic activation of the *N*-hydroxy aromatic amine metabolites of the same aromatic amines.

While individuals can readily be classified as rapid or slow acetylators according to their *NAT2* genotype/phenotype, the genotype/phenotype correlation for *NAT1* is less clear. Most of the polymorphic variants of NAT1 result from one or more SNPs although in some cases small deletions are observed. Interestingly, cell biology studies have shown that some of these variant alleles fail to fold correctly and accumulate in aggresomes where they are targeted for ubiquitinylation and proteolysis.

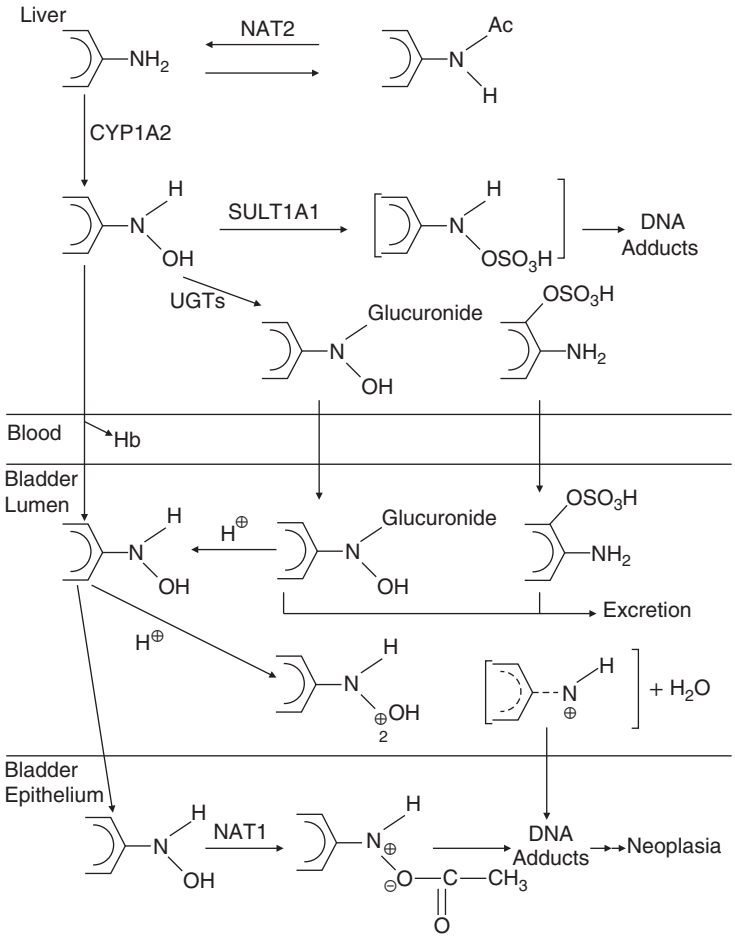


Figure 2.6 Pathways of aromatic amine metabolism in the liver and bladder (source: Stanley (2008); figure 2.40. Reproduced with permission of John Wiley & Sons, Ltd. Original figure from Kadlubar and Badawi (1995). Reproduced with permission of Elsevier)

The difficulty in relating genotype to phenotype in the case of NAT1 arises because NAT1 phenotype is influenced by factors other than SNPs in the coding region (there is, for example, evidence that NAT1 is subject to transcriptional regulation via the glucocorticoid receptor (Bonamassa *et al.*, 2012)), and because the distribution of NAT1 activity (at least in red blood cells) is unimodal rather than bi/tri modal as for NAT2. The current recommendation is that individuals who are homozygous or heterozygous for a slow allele of NAT1 are classified as slow NAT1 acetylators. The alleles to which this applies are NAT1*14A, NAT1*14B, NAT1*17, NAT1*19 and NAT1*22. There is some confusion regarding NAT1*11, which some people classify as a rapid acetylation allele while other classify it as slow, and NAT1*10, which carries SNPs in the 3′ untranslated region but not in the coding region.

The GST isozyme GSTM1 may also contribute to the risk of bladder cancer. The GSTM1 polymorphism takes the form of a null allele, that is, a missing gene,

designated GSTM0, and individuals may carry two, one, or no copies of the GSTM1 gene. Case-control studies and meta-analyses indicate that the risk of bladder cancer is increased by approximately 50% in homozygous GSTM1 null individuals (Engel *et al*., 2002). This effect, unlike that observed in relation to NAT2, is similar in both smokers and non-smokers (Moore *et al*., 2011).

Thus, with respect to bladder cancer, the toxicogenetic situation is relatively clear-cut. Activating and detoxifying enzymes in the liver compete for the aromatic amine substrate, leading to the generation of metabolites, which are subsequently delivered to the bladder. In this scenario, low levels of hepatic *N*-acetylation (NAT2) allow *N*-hydroxylation by CYP1A2 to predominate. Following further by GSTs and transport to the bladder, hydrolysis regenerates a chemically reactive *N*-hydroxyl aromatic amine which is susceptible to *O*-acetylation by NAT1 leading to the initiation of carcinogenesis. The hypothesis based upon this scheme is that individuals who are slow for NAT2, rapid for CYP1A2, GSTM1 null and rapid for NAT1 have the highest risk of developing bladder cancer. Certainly, slow acetylators (at least those who smoke) can be shown to have an increased risk, as are GSTM1 null individuals. There is also evidence that individuals with the slow NAT2 in association with a rapid CYP1A2 phenotype have a further increase in their risk of bladder cancer if they smoke (Tao *et al*., 2012).

It is unusual for studies on toxicogenetic factors in cancer to yield the kind of relatively clear-cut results seen in the case of bladder cancer. Attempts to elucidate the role of the metabolic polymorphisms in colorectal cancer illustrate some of the difficulties in obtaining definitive answers.

Example: Individual susceptibility to colorectal cancer

An individual's risk of colorectal cancer is modified by both genetic and environmental/lifestyle factors, including a first degree family history of colorectal cancer, overeating, physical inactivity, a high intake of red meat, alcohol use, smoking and a low intake of vegetables. Heterocyclic amines are generated during chemical reactions which occur during the cooking of red meat. Thus, people who frequently consume cooked meat are exposed to heterocyclic amines on a regular basis, and these compounds are thought to be the central carcinogens which mediate the carcinogenic process in the colon; haem iron and nitrate/nitrite intake may also play a role (Cross *et al*., 2010). The heterocyclic amine 2-amino-1-methyl-6-phenylimidazo-[4,5-b]pyridine (PhIP) is considered to be of particular significance with respect to colorectal cancer because it is usually the predominant heterocyclic amine found in cooked meat, it is mutagenic in bacterial and mammalian cell-based assays, and it is a colon carcinogen in rats (Gooderham *et al*., 2001) (Figure 2.7).

Heterocyclic amines such as PhIP are subject to both ring- and *N*-oxidation mediated by CYPs; the *N*-hydroxy metabolites are direct acting mutagens whereas the ring hydroxylated metabolites are not mutagenic. The main CYP isozyme involved in the Phase I metabolism of heterocyclic amines is CYP1A2; indeed small genotyping studies have suggested that the CYP1A2*1F variant of CYP1A2 could be a stand-alone risk factor for colorectal cancer (Moonen *et al*., 2005; Saebo *et al*., 2008). Both the parent heterocyclic amines and their Phase I

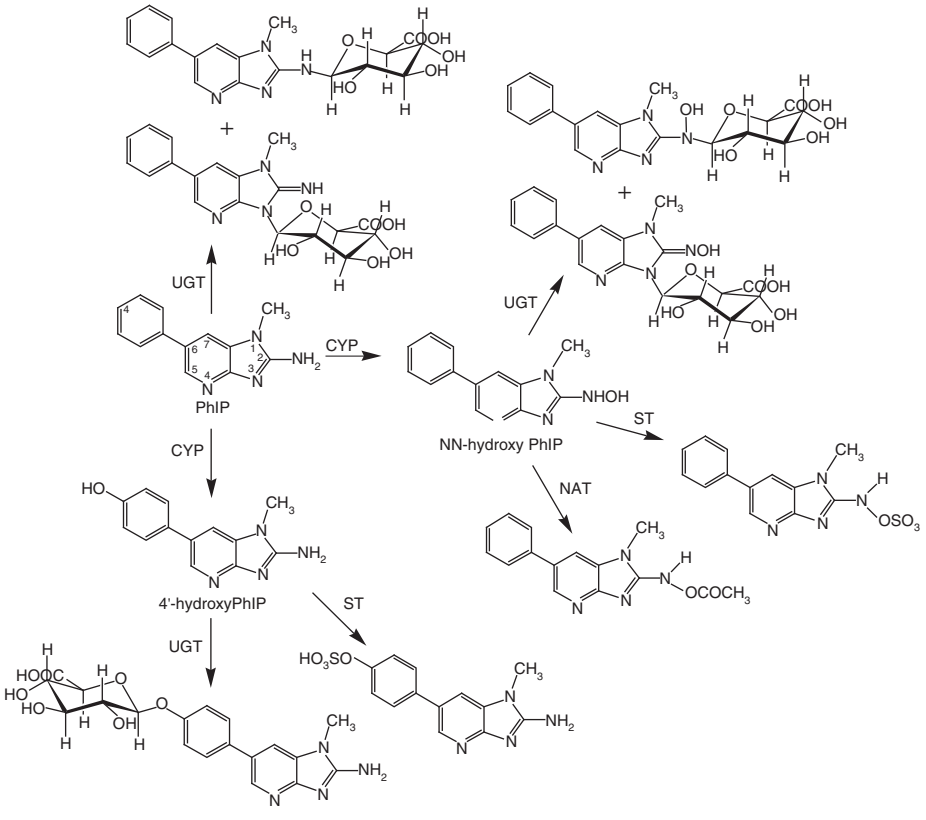


Figure 2.7 Pathways of heterocyclic amine metabolism (source: Stanley, 2008; figure 2.41. Reproduced with permission of John Wiley & Sons, Ltd. Original figure from Gooderham *et al.* (2001). Reproduced with permission of the American Society for Pharmacology and Experimental Therapeutics)

metabolites also undergo Phase II metabolism catalysed by GSTs, UGTs, SULTs and NATs. The key metabolites implicated in colorectal cancer induction are the *N*-acetoxy derivatives, which may be generated by the action of NAT1 and NAT2. The Phase I metabolism of heterocyclic amines is primarily hepatic, but *O*-acetylation can occur in either the liver or in the colon.

In humans, the generation of reactive genotoxic metabolites via *N*-oxidation is the primary route of oxidative metabolism for heterocyclic amines. In experimental animals, on the other hand, both activation and detoxification of these compounds occur, and this may lead to differences in susceptibility to, and the tissue specificity of, their carcinogenic effects. The species differences observed in the oxidative metabolism of heterocyclic amines illustrate the difficulty of extrapolating from animal to human, especially where polymorphic enzymes are involved.

The pathway of metabolism of heterocyclic amines suggests that the rapid NAT2 phenotype would be likely to confer an increased risk of colorectal cancer, especially when combined with high CYP1A2 activity and regular consumption of

well-done red meat. Phenotyping studies in the early 2000s suggested that this was the case, at least in smokers (Le Marchand *et al*., 2001). It has also been suggested that the NAT1*10 variant of NAT1, which is associated with high acetylation activity in colon tissue, may be a risk factor (Yeh *et al*., 2009). However, the results of studies using genotyping alone have been contradictory (Brockton *et al*., 2000, Cleary *et al*., 2010), especially when the studies in question are small (da Silva *et al*., 2011), and meta-analysis of 40 studies has not indicated any clear link between NAT2 genotype and colorectal cancer susceptibility (Zhang *et al*., 2012).

The difference between aromatic amine carcinogenesis, where the rapid acetylator phenotype confers protection, and heterocyclic amine carcinogenesis, where the rapid acetylator phenotype is, if anything, a risk factor, is a function of the chemistry of the two classes of amines. Heterocyclic amine carcinogens exhibit steric hindrance to the exocyclic amine moiety, which makes them resistant to *N*-acetylation, and this means that they are more likely to undergo *O*- but not *N*-acetylation, whereas aromatic amine carcinogens can undergo both reactions.

It is evident that the determination of colorectal cancer susceptibility is multi-factorial, involving both genetic and environmental factors. There is some evidence that the risk of colorectal cancer is increased in individuals who are current or ex-smokers, in individuals who prefer well-done red meat, and have the rapid phenotypes of both CYP1A2 and NAT2. There is also evidence that GSTM1 and GSTT1 null genotypes confer additional risk for colorectal cancer, at least in Caucasian populations (Economopoulos and Sergentanis, 2010), although this does not emerge even from all large studies (Smits *et al*., 2003). The confusion over the relative roles of different polymorphic xenobiotic-metabolising enzymes in modulating susceptibility to colorectal cancer illustrates the importance of adequate study size in toxicogenetic studies.

2.5 Study numbers and effect size

The majority of studies on polymorphic variation and risk of toxicity or disease have, for scientific and resource reasons, been quite small (a few hundred subjects, at most) and have concentrated on one or two key genes. However, redundancy within xenobiotic metabolising pathways means that, if the activity of one enzyme is reduced, potential toxic agents may be directed down other pathways. For example, in smokers, various enzymes, including GSTM, GSTT, GSTP, NAT1 and NAT2, may exert combined effects on DNA adduct formation and cancer susceptibility. In addition, it is now clear that metabolism is not the only phase of drug disposition which is subject to polymorphic variation: polymorphisms in nuclear receptors and drug transporters probably also exert toxicogenetic effects, and our understanding of how these genes determine susceptibility to toxic agents is much less advanced than in the case of xenobiotic metabolising enzymes, partly because the size of the epidemiological studies required in order to understand the roles of all the possible polymorphic variants of these proteins and variation in drug response was considered prohibitive before the year around 2000.

Unlike so-called Mendelian diseases such as sickle cell anaemia and cystic fibrosis, in which alterations in a single gene explain the vast majority of occurrences,

common diseases are likely to be influenced by multiple genes, each with a relatively small effect, whose actions in concert with each other and with environmental influences cause clinical disease. The effect size of an association is the likelihood of response to a drug in individuals with the susceptibility allele as compared with those without that allele. In case-control studies the magnitude of the effect is usually expressed as an odds ratio (OR). If the OR deviates from 1.0, the risk is either increased (OR >1) or decreased (OR <1). The multi-factorial nature of most common diseases means that relative risk estimates are usually low, in the range of 1.5 to 3.0; in fact, pharmacogenomic markers often confer only about a doubling of risk (i.e. an OR of 2.0)

Many preliminary studies aimed at the identification of risk factors are relatively small, with only 100–300 cases and controls. This number is sufficient to detect common polymorphisms which double risk, but will not detect rare polymorphisms or those which cause less than a doubling in risk. Much larger studies and/or meta-analysis are required to reveal small increases in risk. It is also important to note that polymorphic variants may be present at different frequencies within ethnic groups; for example, 45% of Caucasians but only 10% of Japanese individuals are NAT2 slow acetylators. This means that the implications of the identification of a particular variant may differ depending on the ethnic makeup of the exposed population. It may not always be possible to extrapolate risk assessments from one ethnic group to another.

The best approach is now considered to be to concentrate effort on large studies using thousands rather than hundreds of cases and controls, such as the international project on Genetic Susceptibility to Environmental Carcinogens (Gaspari *et al*., 2001), which was initiated in 1996 with the aim of undertaking pooled data analysis on polymorphisms in genes known to affect metabolic susceptibility. Its aims were to quantify gene frequencies in healthy populations by ethnicity and geographical distribution and then to study the association between each gene frequency and cancer. By working together, the investigators hoped to be able to pool a large enough number of subjects to study the effects of multiple gene polymorphisms on cancer risk. Contributions were invited from investigators who had published case-control studies on Phase I and II polymorphisms up to June 1999. As of February 2008, it involved 185 investigators, 304 studies and 124 456 subjects (53 072 cancer cases and 71 384 controls) who had been were genotyped for one or more of CYP1A1, CYP1B1, CYP2E1, CYP2D6, GSTM1, GSTM3, GSTP1, GSTT1, NAT1, NAT2, EH, ADH3, NQO1, MPO and MTHFR. As well as providing information about the geographic and ethnic distribution of known polymorphisms in these genes (Garte and GSEC, 2001), this provided a resource which could be used to evaluate the roles of multiple metabolic polymorphisms in disease susceptibility. The data are now being used to further our understanding of susceptibility to a range of neoplastic conditions, including colorectal cancer (Taioli *et al*., 2009), lung cancer (Cote *et al*., 2009, Raimondi *et al*., 2005, Raimondi *et al*., 2006), oral and pharyngeal cancers (Varela-Lema *et al*., 2008), gastric cancer (Boccia *et al*., 2008), bladder cancer (Kellen *et al*., 2007) and breast cancer (Paracchini *et al*., 2007). These studies have highlighted the difficulties involved in assessing the true consequences of variation in genes whose individual effects are of borderline significance, leading to the establishment of guidelines for the evaluation of cumulative epidemiological evidence (Ioannidis *et al*., 2008).

2.6 Recent developments

The limited methodology available meant that the epidemiology studies of the 1960s–1980s tended to focus on the postulated effects of polymorphisms in individual xenobiotic metabolising enzymes on susceptibility to adverse effects such as cancer. Candidate gene approaches such as those which led to the identification of the roles of NAT2 in bladder cancer, GSTP1 in acrylonitrile toxicity and CYP2C9 in warfarin haemorrhage are limited in that they only look at selected genes and so may miss other important toxicogenetic factors. However, the effects of individual polymorphisms are often subtle because susceptibility to environmentally induced disease is multifactorial, being determined by numerous genetic factors as well as environmental exposures. With the advent of HTS methods for SNP analysis it has become possible to look at thousands of markers across the human genome in so-called genome-wide association studies (GWAS) without making prior assumptions as to which markers might confer susceptibility.

2.6.1 Genome-wide association studies

A GWAS, also known as a whole genome association study, is an examination of genetic variation across the genome, designed to identify genetic associations with observable biological phenomena such as health conditions. In human studies, this might include traits such as blood pressure or weight, or susceptibility to specific diseases or conditions such as serious adverse drug reactions (Daly, 2012). The goal of a GWAS is to test the links between subtle changes in the DNA sequences of individuals within a population and a trait (or phenotype) by assaying the majority of common SNPs across the entire genome. Its advantage over the candidate gene approach is that all genes are considered equal, with no influence from preconceived ideas. This falls under the heading of functional genomics, which is defined as the study of relationships between particular genotypes and specific phenotypes (essentially genotype–phenotype correlation at the genomic level). One early contribution of the GWAS approach was in identifying the role of VKORc02 (as well as CYP2C9) in the pharmacogenetics of warfarin responsiveness; however, the limitation of this approach in toxicogenomics lies in the attempt to associate particular SNPs with exposures, as prior exposure to toxic chemicals may not have been recorded, and there is no opportunity to conduct family studies because it is unethical to expose other family members to a toxic chemical in order to test an association.

High throughput genotyping platforms are good at genotyping SNPs but these only explain a subset of pharmacogenetic variation. They are not so good at identifying structural variants such as insertions, deletions, inversions and copy number variants, leading to an inherent tendency to focus on a subset of possible types of genetic variation.

Copy number variants are defined as DNA segments ranging in size from 1 kb to several megabytes which vary among individuals due to deletion, insertion, duplication or complex recombination: they are less common than SNPs but there is evidence that they can play a key role in drug responsiveness and disease processes (He *et al.*, 2011); for example, some individuals are classified as CYP2D6 ultrarapid metabolisers because they carry multiple copies of the

CYP2D6 gene leading to very high levels of activity and consequent resistance to drugs which undergo CYP2D6-mediated clearance (de Leon, 2007). This issue is being addressed in the Copy Number Variation Project, co-ordinated via the Wellcome Trust Sanger Institute.[12]

The advantage of the GWAS approach is simultaneous, unbiased testing of millions of SNPs while its disadvantage is that it does not provide functional information for the implicated loci, so when interpreting a GWAS study it is important to ask two questions:

- Is any association detected real?
- Are the results actually of any use?

Haplotypes A haplotype is a combination of alleles which are physically linked (in linkage disequilibrium) along a chromosome, in other words multiple loci that are inherited as a unit (Figure 2.8).[13] In contrast, SNPs that are separated by a large distance are typically not very well correlated, because recombination occurs in each generation, mixing the allele sequences of the two chromosomes. Haplotypes can be assigned to both the chromosomes in a diploid cell, so for any given region (apart from on the sex chromosomes in males) there is one haplotype from the mother and one from the father.

Two main approaches are adopted for the analysis of haplotypes:

- **Use of tag SNPs:** The alleles of nearby SNPs on a single chromosome are correlated. This means that if the allele of one SNP for a given individual is known, the alleles of nearby SNPs can often be predicted because each SNP arose in evolutionary history as a single mutation, and was then passed down to descendants surrounded by other, earlier, mutations.
  In this approach genomic patterns of linkage disequilibrium are used to select a set of markers that are statistically associated with many other variants in the genome. Linkage disequilibrium ranges from 0 (no linkage disequilibrium) to 1 (complete linkage disequilibrium) and declines with distance over the range from 10 to 100 kb (in other words, loci separated by 100 kb can usually recombine freely); over shorter distances, linkage disequilibrium is very variable and there is sometimes extensive recombination even around a single gene.

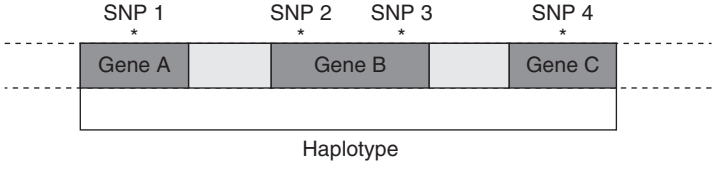


Figure 2.8 Relationship between SNPs, genes and haplotypes. A haplotype is a section of chromosome which is inherited as a block. It may contain one or more genes and/or a variety of SNPs which are co-inherited because of linkage diequilibrium

[12] http://www.sanger.ac.uk/research/areas/humangenetics/cnv/

[13] For a clear explanation of haplotypes and linkage disequilibrium, see Manolio, Brooks and Collins (2008).

- **Sequence-based analysis:** In this approach, potentially functional SNPs in or near exons are assayed directly. This approach is considered to be better for rare variants and complex diseases but may miss regulatory variants which are thought to be important in determining drug response.

In practice, the study of haplotypes provides increased statistical power to detect associations and map disease mutations compared with single marker approaches. This approach means that less genotyping effort is required to evaluate the role of common genetic variants. However, it is important to note that a haplotype is not a permanent entity, and there has been considerable discussion about the existence or otherwise of haplotype blocks. A haplotype block is defined as a block of haplotypes that shows limited genetic diversity within a specific population; for example four or five common haplotypes may be seen in a given population. Early descriptions of genome-wide haplotype structure tended to describe the genome in terms of a fixed arrangement of haplotype blocks, but this way of thinking fails to allow for the underlying recombination rate, which means that a particular haplotype block is not necessarily inherited as a unit. A more accurate way to describe genome-wide haplotype structure would be as block-like regions of reduced recombination showing limited haplotype diversity, punctuated by recombination hotspots (Barnes, 2006).

2.6.2 Collaborative programmes

GWAS require large populations to retain adequate power to detect any significant effect. Over the last couple of decades, several international collaborations have been established in order to address this issue.

The SNP Consortium: The SNP Consortium, which emerged from a Glaxo-Wellcome initiative in 1998, was established in the spring of 1999 (Holden, 2002). The resulting not-for-profit organisation based in the United States was supported by the Wellcome Trust, leading academic centres and 13 pharmaceutical and technological companies. The primary aim of the SNP Consortium was to create a high density SNP map of the human genome in order to facilitate:

- Future development of personalised medication, to minimise side effects and maximise therapeutic benefit.
- Identification of individuals who would benefit from preventative medication.
- Better design of clinical trials now to allow information to be collected for future analyses.

It set out to identify SNPs without biasing the analysis towards coding regions within the genome, although uniform coverage did not prove to be possible because some regions were refractory to cloning. Data release from the SNP Consortium was completed in 2001; 1.7 million SNPs were identified and 1.5 million of these were mapped (1.3 million to a single location and the rest to two or more locations).

The International HapMap project: The: HapMap project[14] was a multi-country effort to identify and catalogue genetic similarities and differences in human beings. It involved researchers in Canada, China, Japan, Nigeria, the United Kingdom and the United States. Whereas the SNP Consortium focused on individual SNPs, the HapMap project looked at haplotypes with the aim of characterising the most common haplotypes in major ethnic groups. It was designed to create a public, genome-wide database of patterns of common sequence variation to guide future genetic studies of human health and disease. The hypothesis underlying the HapMap approach was the 'common disease – common variant' hypothesis, which makes the assumption that most genetic influences on common diseases are attributable to a limited number of allelic variants which are present in 1–5% of the population. Beyond this, however, the HapMap approach was free of reliance on preconceived ideas and has been described as 'a comprehensive, systematic, even agnostic approach' (Manolio *et al*., 2008).

The DNA samples for the HapMap came from a total of 270 people grouped, where possible, in trios (two parents and an adult child). The groups sizes analysed were chosen to be large enough to characterise linkage disequilibrium and haplotypes within the populations they represent, but they are recognised not to be large enough to reflect an entire population or ethnic group (Barnes, 2006).

The samples studied were from

- US residents with northern and western European ancestry (30 trios)
- The Yoruba people of Ibadan, Nigeria (30 trios)
- Unrelated Japanese individuals from the Tokyo area (45 individuals)
- Unrelated Chinese individuals from Beijing (45 individuals)

When analysing HapMap data the Japanese and Chinese subpopulations are often combined because they are genetically very similar.

The HapMap project was designed to allow genome-wide discovery and characterisation of SNPs from major worldwide populations. Genotyping of a million SNPs was carried out in Phase I of the project and published by the International HapMap Consortium in 2005 (IHC, 2005); 3.1 million more were added in 2007 (IHC, 2007). In Phase III of the project, a further seven populations were added. All the data, from these 1301 extended HapMap samples and those analysed previously, are freely available from the HapMap data co-ordination centre; the HapMap genome browser[15] can be used to view linkage disequilibrium results and haplotypes across a gene or region of interest, select tagging SNPs and export genotype/linkage disequilibrium data in single or multiple populations. In addition, actual samples from the HapMap project are held at the Coriell Institute for Medical Research, USA.[16]

In addition to DNA samples, lymphoblastoid cell lines have been generated from the HapMap participants' peripheral blood mononuclear cells. These are also held at the Coriell Institute. They are being offered as a resource for testing the biological consequences of the genotypes identified in the HapMap and have

[14] http://hapmap.ncbi.nlm.nih.gov/
[15] http://hapmap.ncbi.nlm.nih.gov/cgi-perl/gbrowse/hapmap28_B36/
[16] http://www.coriell.org/

been used in a proof-of-principle study to demonstrate that it is possible to use human lymphoblastoid cells as an *in vitro* model system to identify genetic factors affecting responses to xenobiotics (O'Shea *et al*., 2011), but they are, of course, subject to the caveats about use of surrogate tissues mentioned at the beginning of the chapter. They are likely to be a useful resource for studying the haematotoxic side effects of chemotherapy (Zhang and Dolan, 2010), but apart from that their usefulness as a model for target tissue effects may be quite limited.

Manolio *et al*. (2008) characterise the HapMap as an extensive international collaborative effort in which common objectives were agreed upon and pursued in a highly focused, cooperative approach; they consider this to have been key to its success. Their evaluation is that successful GWAS are the most visible and exciting outcomes of HapMap to date, providing novel and unexpected insights into the pathophysiology of disease. The HapMap project has also contributed to the development of high throughput genotyping and analytic methods and provided samples for validation and standardization of variation detection methods. The way in which it has expanded our understanding of evolutionary pressures and natural selection, defining genetic relationships across populations, is expected to produce new advances in the prevention and treatment of common diseases.

The International HapMap project succeeded in identifying variants and/or genomic regions associated with some 40 or so complex diseases and traits in a diverse range of populations. These include conditions such as cancer, inflammatory bowel disease and cardiovascular disease which have a known link with the environment and/or exposure to xenobiotics. Interestingly, the evidence suggests that disease susceptibility is more often found to be associated with non-coding SNPs than those in the coding regions, and this has been taken to suggest that the regulation of protein expression is more important than changes in the structure and/or function of the proteins themselves. This may be because coding sequence changes which significantly affect protein function are strongly selected against and are therefore too rare to be detected as SNPs under the conditions used in the HapMap study.

The HapMap has introduced a new paradigm into genomic research by making possible the cost-efficient assessment of genomic variation within an individual; when the project began, the cost per sample was about 40¢, by 2005 this had dropped to 1¢ and by 2008 it had dropped a further 10-fold to about 0.1¢. Its results are, at the time of writing, still being evaluated but it is clear that this project has provided insights into evolutionary pressures on the human genome and it has facilitated functional investigation and cross-population comparisons of candidate disease genes. It offers the possibility of developing genetic methods to screen for disease susceptibility and identifying new drug targets, and its potential contribution to toxicogenomics is being followed up as part of the 1000 Genomes and Environmental Genome Projects.

The 1000 Genomes project: : The availability of HTS methods has made it possible to sequence multiple human genomes relatively cheaply, leading to the initiation of the 1000 Genomes project[17] in 2007. This project is run

[17] http://www.1000genomes.org/home

by a consortium involving participants from the United Kingdom, United States, Germany and China and its aim is 'to provide a deep characterization of human genome sequence variation as a foundation for investigating the relationship between genotype and phenotype' (Abecasis *et al.*, 2010). It uses HTS technologies for the in-depth characterization of sequence variation in the human genome, and aims to characterise at least 95% of variants found with a frequency of ≥1% (i.e. meeting the criteria for a polymorphism) in each of five major human population groups (Europe, East Asia, South Asia, West Africa and the Americas). In addition, it set out to characterise low frequency alleles (with an incidence greater than about 0.1%) in coding regions because, as mentioned earlier, functional allele changes in coding regions tend to occur with low frequency. Unlike the SNP Consortium and the HapMap, the 1000 Genomes project aims to discover as many variants as possible present in more than 1% in the human population, including copy number variants as well as SNPs. Many of the samples used in the 1000 Genomes project are lymphoblastoid cell lines from the International HapMap project, allowing the relationships between genotype and cellular phenotype to be examined in detail (bearing in mind the caveats about use of lymphoid cells as a surrogate tissue).

The pilot phase of the 1000 Genomes project comprised three exercises: low coverage whole genome sequencing of 179 individuals from the four HapMap Phase I/II population groups; high coverage whole genome sequencing of two HapMap mother-father-child trios (one European and one African) and exon targeted sequencing of 697 individuals from the HapMap Phase III population groups (Abecasis *et al.*, 2010). The locations, allele frequencies and local haplotype structures of 15 million SNPs, a million short insertions/deletions and 20 000 other structural variants were defined in this phase of the project. This revealed that, on an average, each individual carries about 250–300 loss of function mutations and, surprisingly, as many as 50–100 variants which have been implicated in the aetiology of genetic disorders.

The next stage of the 1000 Genomes project will focus on whole genome sequencing. Whole genome sequencing is the gold standard for individualised full sequence information. Until now it has been cost-prohibitive but it is becoming cheaper; the '$1000 genome challenge' has been set as the target for sequencing an entire genome. If this is met it should enable full prospective personalised therapy once associations become more certain. The aim of the complete 1000 Genomes project has actually been extended to genetic characterisation of 2500 individuals from 27 populations by means of low coverage whole genome sequencing, array-based genotyping and deep targeted sequencing of all coding regions, thus representing a big step forward towards a complete description of human DNA polymorphism (Abecasis *et al.*, 2010). The array-based genotyping methods developed for use in the 1000 Genomes project permit simultaneous testing of five million SNPs.

By March 2012, the 1000 Genomes project had generated more than 260 Tb of data which are publicly accessible via the 1000 Genomes web-based browser.[18]

[18] http://www.1000genomes.org/ensembl-browser

It is hoped (Zhang and Dolan, 2010) that the integration of 1000 Genomes data with systematic phenotyping information (e.g. about drug responses) will make it possible to identify new candidate susceptibility loci and pinpoint causal variants for disease, toxicity and adverse side effects of drugs. A step towards this aim has been taken by incorporating data from the 1000 Genomes project into the database maintained by the Very Important Pharmacogenes component of PharmGKB[19] (Whirl-Carrillo *et al*., 2012), an initiative aimed at providing annotated information regarding genes, variants, haplotypes and splice variants which are relevant to pharmacogenetics and pharmacogenomics (Thorn *et al*., 2010). The Very Important Pharmacogenes database contains information, including that from the 1000 Genomes project, about 39 key polymorphic genes involved in pharmacological responses and PharmGKB issues regular updates on Very Important Pharmacogenes (see Lamba *et al*., 2012; McDonagh *et al*., 2012; Thorn *et al*., 2012a,2012b.[20]

The Environmental Genome Project: The Environmental Genome Project is an extension of the Human Genome Project initiated by the US National Institute of Environmental Health Sciences (NIEHS)[21] in 1997 (Olden and Wilson, 2000). Its aim is to achieve a comprehensive characterisation of SNPs in key biological pathways involved in toxic responses, including DNA repair, cell cycle control and drug metabolism. To obtain comprehensive variation discovery in this important set of candidate genes, a PCR-based approach was implemented by the Environmental Genome Project. Overlapping PCR amplicons covering the entirety of each gene, including the 5′ upstream promoter (~2 kb), all intronic and exonic regions, and 3′ downstream flanking regions were generated. Each amplicon was fully sequenced to generate a composite (diploid) chromatogram for each individual in the DNA panel, and genotypes were assigned at each polymorphic position. The project initially focussed on a so-called polymorphism discovery resource of 90 individuals selected to represent the US population, but the focus later moved to a group of 95 individuals who had already been examined as part of the International HapMap project. In this way, the Environmental Genome Project was complementary to the International HapMap project.

The presence of SNPs at a frequency of about one every 300 bases in the human genome meant that it was too expensive to sequence all of them in routine studies. This led Environmental Genome Project investigators to attempt to identify as subset of SNPs which can act as tags for specific haplotypes. This analysis did, however, raise some questions about the value of the HapMap SNP portfolio. The results of two studies which set out to answer the question whether HapMap SNPs are sufficient for the selection of tag SNPs for the most common haplotypes indicated that, while the HapMap is a rich resource for haplotype identification, the HapMap data are not robust enough to identify all the common haplotypes in key genes (Tantoso *et al*., 2006, Taylor *et al*., 2006), indicating that yet more detailed

[19] http://www.pharmgkb.org/

[20] For more on PharmGKB and other key databases in the area of pharmacogenetics, see Sim, Altman, and Ingelman-Sundberg (2011).

[21] http://www.niehs.nih.gov/.

analysis of human SNP profiles is still required for a complete understanding of variation across the human genome.

The Environmental Genome Project includes initiatives in several major research areas:

- **SNP Discovery and GeneSNPs database:** Sequencing of over 600 prioritised environmentally relevant genes thought to play a role in susceptibility to environmental exposures in a panel of 95 individuals representing the ethnic diversity found in the United States. The project has given particular consideration to genes, such as those encoding CYPs, which are known to play a role in susceptibility to toxic chemicals.[22]
- **Comparative Mouse Genomics Consortium:** Development of important mouse models to study the functional significance of human DNA polymorphisms, primarily within the prioritised DNA repair and cell cycle control pathways (Wiley *et al*., 2006).
- **Ethical, legal and social implications of genomic research:** The Environmental Genome Project raises a number of ethical questions around the issues of genetic testing, health insurance and employment law. If you are interested in these, they are described very well by Robert and Smith (Robert and Smith, 2004).

The information gathered in these areas will help build an understanding of the complex interrelationships between environmental exposure, genetic susceptibility, and human disease.

2.7 The UK Biobank

The UK Biobank project[23] represents an important step forward in that it is an example of a large, centralised prospective population study which will provide data and specimens for analysis for decades to come. It originated from a meeting jointly hosted in 1999 by the Medical Research Council and the Wellcome Trust at which the urgent need for a large-scale prospective population study to address interactions between genetic and environmental causes of disease was identified (Ollier *et al*., 2005). The UK Biobank was established as a charitable company in 2003. Its co-ordination centre was based at the University of Manchester with oversight by an independent Board of Directors.

The aim of the UK Biobank study is to characterise the role of genetic factors, lifestyle and environmental exposures in the major diseases of middle and late age in 500 000 people from all around the United Kingdom who were aged 40–69 at the start of the study. This age range was selected to represent the subset of the population in which the most common disease phenotypes occur at a frequency high enough to generate sufficient statistical power for meaningful analysis within 5–10 years of study commencement (Ollier *et al*., 2005). Prospective studies of this

[22] For a detailed discussion of coverage of CYP polymorphisms by the Environmental Genome Project, see Rieder *et al*. (2008).
[23] http://www.ukbiobank.ac.uk/

kind have to be very large because only a minority of participants will develop any given condition during the duration of the study (Collins, 2012).

Over the next 20–30 years UK Biobank will allow approved researchers to use these unique resources to study the progression of chronic illnesses; it will also make it possible to investigate the determinants of good health in the middle aged and elderly. This prospective approach allows exposures to be assessed before they are affected by the presence of a disease or its treatment, and before behaviours change in response to developing disease. It also means that conditions which cannot be studied retrospectively, such as fatal diseases and severe degenerative conditions, can be examined. The overall aim is to help researchers develop new and better ways to prevent, diagnose and treat many common diseases of middle and old age.

The key to the UK Biobank plan was to collect specimens and exposure data prior to the onset of disease, making it possible to examine exposures without potential bias due to outcomes as well as undertaking before and after comparisons of biomarkers. The study design adopted was what is called a community-based prospective cohort study. Postal invitations were issued to nine million individuals aged between 40 and 69 years. Of these, 500 000 agreed to participate and were invited to a local assessment centre for an initial examination between 2007 and 2010. This low response rate, with the associated need to contact many more people than eventually participated, was accepted because of the desire to achieve a sufficiently diverse study population. The aim was to construct a representative sample population in terms of ethnicity and sociodemographics, with equal numbers of men and women and containing balanced numbers in 5-year age bands from 40 to 69 years. The baseline assessment involved an interview, health and lifestyle questionnaire, and measurement of various clinical parameters (Manolio *et al*., 2012). In addition, a key part of this project was that samples of blood and urine were taken at this initial interview (saliva collection was added to the protocol in the later stages (Pramanik *et al*., 2012)), providing an invaluable resource for future analysis. Informed consent was also obtained for health outcomes to be followed via NHS medical records and the aim was to collect ongoing data on every participant who did not withdraw from the study.

The approach to sample collection, processing and archiving was an enormous challenge for the UK Biobank team because of the large number of samples to be collected, the need to conduct multiple analyses on each participant and the intention to archive samples for future analyses which could not be specified at the outset (Elliott and Peakman, 2008). In order to meet these requirements, best practice was adopted from industry and standardised procedures with strict quality assurance (QA) and quality control (QC) were implemented. The overriding consideration was the need to generate the highest quality data; to this end, minimal local processing was undertaken and the samples were shipped promptly from the assessment centres to a central facility for storage at −80°C (working archive) and in liquid nitrogen (backup archive). Haematology samples were analysed immediately, and other samples from each participant were stored at two geographically separate locations where sample movements were carried out by automated robotic systems.

In order to conserve the precious biological specimens in the Biobank, the decision was taken not to distribute samples for analysis (Ollier *et al*., 2005); instead,

any required analyses are commissioned directly by the Biobank management from approved laboratories. This process was agreed in order to ensure optimal QC, maximal efficiency of sample use and timely acquisition of data.

Data management is also a key aspect of the UK Biobank project (Ollier *et al.*, 2005). In order to ensure the protection of participants' confidential information while allowing appropriate dissemination of results, the data are held in three repositories: a protected resource (health records), a managed resource (non-sensitive data held so as to protect users and ensure the scientific/ethical integrity of the study) and an open resource (containing results and non-confidential research data).

The UK Biobank is considered to be a model for the good conduct of large prospective epidemiology studies (Manolio *et al.*, 2012). In particular, the recruitment process (with its focus on maintenance of diversity within the study population, use of local assessment centres and provision of a helpline for potential participants) is viewed as an example of good practice, as is the direct link to National Health Service (NHS) records which simplifies follow-up. In addition, the establishment of an automated, industrial-scale specimen processing and storage system is recognised as a major step forward in the conduct of prospective epidemiology studies. Furthermore, the UK Biobank has been recognised for its high standards of ethics and governance and it has been identified as an example of altruism since the participants are donating their time, information and samples for no personal benefit (Thornton, 2009)

The availability of prospective samples from hundreds of thousands of individuals will allow the UK Biobank to be used as a resource to assess the relevance of different exposures to disease incidence through retrospective case-control studies, and with the launch of the UK Biobank resource for use by all researchers on 30 March 2012 (Collins, 2012) this phase of the project is now well under way. The UK Biobank resource is available for any health-related research that is in the public interest, without any exclusivity or preferential access.

2.8 Conclusions

Individual variation in responses to toxic chemicals is often caused by underlying genetic variation, and the emerging field of toxicogenetics promises to enhance our understanding of susceptibility to toxic agents in ways which could not have been imagined even a few years ago. Genetic testing is now becoming more widely accepted but attitudes to it are often based on a simplistic view of what toxicogenetics can actually achieve. For example, no current test can predict outcomes with 100% accuracy and rare variants have large effects, but explain only a small proportion of variability within the population.

The availability of high throughput PCR-based methods has made it possible to examine large numbers of samples for many different polymorphisms simultaneously, generating huge volumes of data, which should be of great value in informing risk assessment as long as the limitations of the available methods are taken into account (Hong *et al.*, 2010a, Hong *et al.*, 2010b, Gamazon *et al.*, 2012). Large population studies with multiple markers have the potential to explain a large proportion of variability, but smaller combined effects are harder

to detect consistently. Whole genome sequencing for each individual may allow susceptibility to be predicted, at least to a limited extent, as well as making personalised therapy a realistic prospect, but it poses practical problems of cost effectiveness and data handling for the health care provider as well as ethical and legal issues.

The strength of the toxicogenetic approach is that it makes it possible to examine susceptibility to xenobiotics in human populations, thus answering concerns relating to differences in susceptibility between humans and animals. However, the nature of toxicogenetic analysis, including the fact that, for ethical reasons, chemicals may not be administered to humans at known toxic doses and invasive methods may not be used to obtain relevant tissue samples, means that a degree of variability and uncertainty will always be present.

In conclusion, therefore, polymorphisms in xenobiotic metabolism and its regulation clearly contribute to susceptibility to environmental toxins and carcinogens. The effects of individual genes are often subtle but specific combinations of multiple genotypes, found in specific subpopulations, may have a marked effect on susceptibility. The current status of research is that the characterisation of SNPs, haplotypes and their distribution is currently in progress. We do not have sufficient data yet to allow this information to be used in the risk assessment process, but significant developments are anticipated once the baseline data are all in.

Self-assessment questions

- What techniques would you use to identify a novel SNP and determine its role in disease susceptibility?
- Is it ever worth trying to delineate the contribution of an individual xenobiotic metabolising enzyme to the risk of cancer development in an exposed population?
- What are the ethical issues involved in predicting cancer susceptibility on the basis of genetic predisposition?
- Does the move from hypothesis-driven approaches to open-ended approaches represent progress?

Background reading

Xenobiotic metabolism:

Coleman, M.D. (2010) Human Drug Metabolism: An Introduction, 2nd edn, Wiley-Blackwell, Chichester, UK. ISBN: 978-0-470-74216-7

Gibson, G.G. and Skett, P. (2001) Introduction to Drug Metabolism, 3rd edn, Nelson-Thornes, Cheltenham, UK. ISBN: 978-0-748-76011-4

Molecular techniques

Dale, J.W. and von Schantz, M. (2007) From Genes to Genomes: Concepts and Applications of DNA Technology. John Wiley and Sons, Ltd Chichester, UK. ISBN: 978-0-470-01734-0

References

Abecasis, G.R., Altshuler, D., Auton, A. *et al.* (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.

Barnes, M.R. (2006) Navigating the HapMap. *Briefings in Bioinformatics*, **7**, 211–224.

Bhasker, C.R. and Hardiman, G. (2010) Advances in pharmacogenomics technologies. *Pharmacogenomics*, **11**, 481–485.

Boccia, S., Hung, R., Ricciardi, G. *et al.* (2008) Meta- and pooled analyses of the methylenetetrahydrofolate reductase C677T and A1298C polymorphisms and gastric cancer risk: a huge-GSEC review. *American Journal of Epidemiology*, **167**, 505–516.

Bonamassa, B., Ma, Y. and Liu, D. (2012) Glucocorticoid receptor-mediated transcriptional regulation of *N*-acetyltransferase 1 gene through distal promoter. *The AAPS Journal*, **14**, 581–590.

Brockton, N., Little, J., Sharp, L. and Cotton, S.C. (2000) *N*-acetyltransferase polymorphisms and colorectal cancer: a HuGE review. *American Journal of Epidemiology*, **151**, 846–861.

Cartwright, R.A., Glashan, R.W., Rogers, H.J. *et al.* (1982) Role of *N*-acetyltransferase phenotypes in bladder carcinogenesis: a pharmacogenetic epidemiological approach to bladder cancer. *Lancet*, **2**, 842–845.

Cavallari, L.H. (2012) Tailoring drug therapy based on genotype. *Journal of Pharmacy Practice*, **25**, 413–416.

Cleary, S.P., Cotterchio, M., Shi, E. *et al.* (2010) Cigarette smoking, genetic variants in carcinogen-metabolizing enzymes, and colorectal cancer risk. *American Journal of Epidemiology*, **172**, 1000–1014.

Collins, R. (2012) What makes UK Biobank special? *Lancet*, **379**, 1173–1174.

Cote, M.L., Chen, W., Smith, D.W. *et al.* (2009) Meta- and pooled analysis of GSTP1 polymorphism and lung cancer: a HuGE-GSEC review. *American Journal of Epidemiology*, **169**, 802–814.

Cross, A.J., Ferrucci, L.M., Risch, A. *et al.* (2010) A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. *Cancer Research*, **70**, 2406–2414.

Da Silva, T.D., Felipe, A.V., De Lima, J.M. *et al.* (2011) *N*-Acetyltransferase 2 genetic polymorphisms and risk of colorectal cancer. *World Journal of Gastroenterology*, **17**, 760–765.

Daly, A.K. (2004) Development of analytical technology in pharmacogenetic research. *Naunyn-Schmiedeberg's Archives of Pharmacology*, **369**, 133–140.

Daly, A.K. (2012) Using genome-wide association studies to identify genes important in serious adverse drug reactions. *Annual Review of Pharmacology and Toxicology*, **52**, 21–35.

De Leon, J. (2007) The crucial role of the therapeutic window in understanding the clinical relevance of the poor versus the ultrarapid metabolizer phenotypes in subjects taking drugs metabolized by CYP2D6 or CYP2C19. *Journal of Clinical Psychopharmacology*, **27**, 241–245.

Economopoulos, K.P. and Sergentanis, T.N. (2010) GSTM1, GSTT1, GSTP1, GSTA1 and colorectal cancer risk: a comprehensive meta-analysis. *European Journal of Cancer*, **46**, 1617–1631.

Elliott, P. and Peakman, T.C. (2008) The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *International Journal of Epidemiology*, **37**, 234–244.

Engel, L.S., Taioli, E., Pfeiffer, R. *et al.* (2002) Pooled analysis and meta-analysis of glutathione *S*-transferase M1 and bladder cancer: a HuGE review. *American Journal of Epidemiology*, **156**, 95–109.

Finks, S.W., Finks, A.L. and Self, T.H. (2006) Hydralazine-induced lupus: maintaining vigilance with increased use in patients with heart failure. *Southern Medical Journal*, **99**, 18–22.

Furlong, C.E., Suzuki, S.M., Stevens, R.C. *et al.* (2010) Human PON1, a biomarker of risk of disease and exposure. *Chemico-Biological Interactions*, **187**, 355–361.

Gamazon, E.R., Skol, A.D. and Perera, M.A. (2012) The limits of genome-wide methods for pharmacogenomic testing. *Pharmacogenetics and Genomics*, **22**, 261–272.

Garcia-Closas, M., Malats, N., Silverman, D. *et al.* (2005) NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet*, **366**, 649–659.

Garte, S. and GSEC (2001) Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiology, Biomarkers and Prevention*, **10**, 1239–1248.

Gaspari, L., Marinelli, D. and Taioli, E. (2001) International collaborative study on genetic susceptibility to environmental carcinogens (GSEC): an update. *International Journal of Hygiene and Environmental Health*, **204**, 39–42.

Gooderham, N.J., Murray, S., Lynch, A.M. *et al.* (2001) Food-derived heterocyclic amine mutagens: variable metabolism and significance to humans. *Drug Metabolism and Disposition*, **29**, 529–534.

Hamdy, N.M., Al-Abbasi, F.A., Alghamdi, H.A. *et al.* (2012) Role of neutrophils in acrylonitrile-induced gastric mucosal damage. *Toxicology Letters*, **208**, 108–114.

He, Y., Hoskins, J.M. and McLeod, H.L. (2011) Copy number variants in pharmacogenetic genes. *Trends in Molecular Medicine*, **17**, 244–251.

Hein, D.W. (2009) *N*-acetyltransferase SNPs: emerging concepts serve as a paradigm for understanding complexities of personalized medicine. *Expert Opinion on Drug Metabolism & Toxicology*, **5**, 353–366.

Hein, D.W. and Doll, M.A. (2012) Accuracy of various human NAT2 SNP genotyping panels to infer rapid, intermediate and slow acetylator phenotypes. *Pharmacogenomics*, **13**, 31–41.

Hickman, D., Risch, A., Camilleri, J.P. and Sim, E. (1992) Genotyping human polymorphic arylamine *N*-acetyltransferase: identification of new slow allotypic variants. *Pharmacogenetics*, **2**, 217–226.

Holden, A.L. (2002) The SNP consortium: summary of a private consortium effort to develop an applied map of the human genome. *Biotechniques, Suppl*, **22-4**, 26.

Hong, H., Shi, L., Su, Z. *et al.* (2010a) Assessing sources of inconsistencies in genotypes and their effects on genome-wide association studies with HapMap samples. *Pharmacogenomics Journal*, **10**, 364–374.

Hong, H., Su, Z., Ge, W. *et al.* (2010b) Evaluating variations of genotype calling: a potential source of spurious associations in genome-wide association studies. *Journal of Genetics*, **89**, 55–64.

Huang, Y.-F., Chen, M.-L., Liou, S.-H. *et al.* (2011) Association of CYP2E1, GST and mEH genetic polymorphisms with urinary acrylamide metabolites in workers exposed to acrylamide. *Toxicology Letters*, **203**, 118–126.

Huang, Y.F., Chiang, S.Y., Liou, S.H. *et al.* (2012) The modifying effect of CYP2E1, GST, and mEH genotypes on the formation of hemoglobin adducts of acrylamide and glycidamide in workers exposed to acrylamide. *Toxicology Letters*, **215**, 92–99.

IHC (2005) A haplotype map of the human genome. *Nature*, **437**, 1299–1320.

IHC (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.

Ioannidis, J.P., Boffetta, P., Little, J. *et al.* (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. *International Journal of Epidemiology*, **37**, 120–132.

Johansson, I. and Ingelman-Sundberg, M. (2011) Genetic polymorphism and toxicology – with emphasis on cytochrome P450. *Toxicological Sciences*, **120**, 1–13.

Kadlubar, F.F. and Badawi, A.F. (1995) Genetic susceptibility and carcinogen-DNA adduct formation in human urinary bladder carcinogenesis. *Toxicology Letters*, **82-83**, 627–632.

Kellen, E., Hemelt, M., Broberg, K. *et al.* (2007) Pooled analysis and meta-analysis of the glutathione *S*-transferase P1 Ile105Val polymorphism and bladder cancer: a HuGE-GSEC review. *American Journal of Epidemiology*, **165**, 1221–1230.

Kinzig-Schippers, M., Tomalik-Scharte, D., Jetter, A. *et al.* (2005) Should we use *N*-acetyltransferase type 2 genotyping to personalize isoniazid doses? *Antimicrobial Agents and Chemotherapy*, **49**, 1733–1738.

Klein, T.E., Altman, R.B., Eriksson, N. *et al.* (2009) Estimation of the warfarin dose with clinical and pharmacogenetic data. *The New England Journal of Medicine*, **360**, 753–764.

Lamba, J., Hebert, J.M., Schuetz, E.G. *et al.* (2012) PharmGKB summary: very important pharmacogene information for CYP3A5. *Pharmacogenetics and Genomics*, **22**, 555–558.

Le Marchand, L., Hankin, J.H., Wilkens, L.R. *et al.* (2001) Combined effects of well-done red meat, smoking, and rapid *N*-acetyltransferase 2 and CYP1A2 phenotypes in increasing colorectal cancer risk. *Cancer Epidemiology, Biomarkers and Prevention*, **10**, 1259–1266.

Madian, A.G., Wheeler, H.E., Jones, R.B. and Dolan, M.E. (2012) Relating human genetic variation to variation in drug responses. *Trends in Genetics*, **28**, 487–495.

Manolio, T.A., Brooks, L.D. and Collins, F.S. (2008) A HapMap harvest of insights into the genetics of common disease. *Journal of Clinical Investigation*, **118**, 1590–1605.

Manolio, T.A., Weis, B.K., Cowie, C.C. *et al.* (2012) New models for large prospective studies: is there a better way? *American Journal of Epidemiology*, **175**, 859–866.

McDonagh, E.M., Wassenaar, C., David, S.P. *et al.* (2012) PharmGKB summary: very important pharmacogene information for cytochrome P-450, family 2, subfamily A, polypeptide 6. *Pharmacogenetics and Genomics*, **22**, 695–708.

Moonen, H., Engels, L., Kleinjans, J. and Kok, T. (2005) The CYP1A2-164A-->C polymorphism (CYP1A2*1F) is associated with the risk for colorectal adenomas in humans. *Cancer Letters*, **229**, 25–31.

Moore, L.E., Baris, D.R., Figueroa, J.D. *et al.* (2011) GSTM1 null and NAT2 slow acetylation genotypes, smoking intensity and bladder cancer risk: results from the New England bladder cancer study and NAT2 meta-analysis. *Carcinogenesis*, **32**, 182–189.

Nelson, D.R. (2009) The cytochrome P450 homepage. *Human Genomics*, **4**, 59–65.

O'Leary, K.A., Edwards, R.J., Town, M.M. and Boobis, A.R. (2005) Genetic and other sources of variation in the activity of serum paraoxonase/diazoxonase in humans: consequences for risk from exposure to diazinon. *Pharmacogenetics and Genomics*, **15**, 51–60.

O'Shea, S.H., Schwarz, J., Kosyk, O. *et al.* (2011) *In vitro* screening for population variability in chemical toxicity. *Toxicological Sciences*, **119**, 398–407.

Olden, K. and Wilson, S. (2000) Environmental health and genomics: visions and implications. *Nature Reviews Genetics*, **1**, 149–153.

Ollier, W., Sprosen, T. and Peakman, T. (2005) UK Biobank: from concept to reality. *Pharmacogenomics*, **6**, 639–646.

Paracchini, V., Raimondi, S., Gram, I.T. *et al.* (2007) Meta- and pooled analyses of the cytochrome P-450 1B1 Val432Leu polymorphism and breast cancer: a HuGE-GSEC review. *American Journal of Epidemiology*, **165**, 115–125.

Pirmohamed, M. (2011) Pharmacogenetics: past, present and future. *Drug Discovery Today*, **16**, 852–861.

Pramanik, R., Thompson, H., Kistler, J.O. *et al.* (2012) Effects of the UK Biobank collection protocol on potential biomarkers in saliva. *International Journal of Epidemiology*, **41**, 1786–1797.

Raimondi, S., Boffetta, P., Anttila, S. *et al.* (2005) Metabolic gene polymorphisms and lung cancer risk in non-smokers. An update of the GSEC study. *Mutation Research*, **592**, 45–57.

Raimondi, S., Paracchini, V., Autrup, H. *et al.* (2006) Meta- and pooled analysis of GSTT1 and lung cancer: a HuGE-GSEC review. *American Journal of Epidemiology*, **164**, 1027–1042.

Rettie, A.E. and Tai, G. (2006) The pharmocogenomics of warfarin: closing in on personalized medicine. *Molecular Interventions*, **6**, 223–227.

Rieder, M.J., Livingston, R.J., Stanaway, I.B. and Nickerson, D.A. (2008) The environmental genome project: reference polymorphisms for drug metabolism genes and genome-wide association studies. *Drug Metabolism Reviews*, **40**, 241–261.

Risch, A., Wallace, D.M., Bathers, S. and Sim, E. (1995) Slow *N*-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer. *Human Molecular Genetics*, **4**, 231–236.

Robert, J.S. and Smith, A. (2004) Toxic ethics: environmental genomics and the health of populations. *Bioethics*, **18**, 493–514.

Saebo, M., Skjelbred, C.F., Brekke Li, K. *et al.* (2008) CYP1A2 164 A-->C polymorphism, cigarette smoking, consumption of well-done red meat and risk of developing colorectal adenomas and carcinomas. *Anticancer Research*, **28**, 2289–2295.

Sim, S.C., Altman, R.B. and Ingelman-Sundberg, M. (2011) Databases in the area of pharmacogenetics. *Human Mutation*, **32**, 526–531.

Sim, S.C. and Ingelman-Sundberg, M. (2010) The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Human Genomics*, **4**, 278–281.

Smits, K.M., Gaspari, L., Weijenberg, M.P. *et al.* (2003) Interaction between smoking, GSTM1 deletion and colorectal cancer: results from the GSEC study. *Biomarkers*, **8**, 299–310.

Stanley, L.A., Coroneos, E., Cuff, R. *et al.* (1996) Immunochemical detection of arylamine *N*-acetyltransferase in normal and neoplastic bladder. *Journal of Histochemistry and Cytochemistry*, **44**, 1059–1067.

Suarez-Kurtz, G., Vargens, D.D., Sortica, V.A. and Hutz, M.H. (2012) Accuracy of NAT2 SNP genotyping panels to infer acetylator phenotypes in African, Asian, Amerindian and admixed populations. *Pharmacogenomics*, **13**, 851–854 author reply 855.

Suhua, W., Rongzhu, L., Wenrong, X. *et al.* (2010) Induction or inhibition of cytochrome P450 2E1 modifies the acute toxicity of acrylonitrile in rats: biochemical evidence. *Archives of Toxicology*, **84**, 461–469.

Taioli, E., Garza, M.A., Ahn, Y.O. *et al.* (2009) Meta- and pooled analyses of the methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism and colorectal cancer: a HuGE-GSEC review. *American Journal of Epidemiology*, **170**, 1207–1221.

Takano, R., Murayama, N., Horiuchi, K. *et al.* (2010) Blood concentrations of acrylonitrile in humans after oral administration extrapolated from *in vivo* rat pharmacokinetics, *in vitro* human metabolism, and physiologically based pharmacokinetic modeling. *Regulatory Toxicology and Pharmacology*, **58**, 252–258.

Tantoso, E., Yang, Y. and Li, K.B. (2006) How well do HapMap SNPs capture the untyped SNPs? *BMC Genomics*, **7**, 238.

Tao, L., Xiang, Y.B., Chan, K.K. *et al.* (2012) Cytochrome P4501A2 phenotype and bladder cancer risk: The Shanghai bladder cancer study. *International Journal of Cancer*, **130**, 1174–1183.

Taylor, J.A., Xu, Z.L., Kaplan, N.L. and Morris, R.W. (2006) How well do HapMap haplotypes identify common haplotypes of genes? A comparison with haplotypes of 334 genes resequenced in the environmental genome project. *Cancer Epidemiology, Biomarkers and Prevention*, **15**, 133–137.

Thier, R., Balkenhol, H., Lewalter, J. *et al.* (2001) Influence of polymorphisms of the human glutathione transferases and cytochrome P450 2E1 enzyme on the metabolism and toxicity of ethylene oxide and acrylonitrile. *Mutation Research*, **482**, 41–46.

Thier, R., Lewalter, J. and Bolt, H.M. (2000) Species differences in acrylonitrile metabolism and toxicity between experimental animals and humans based on observations in human accidental poisonings. *Archives of Toxicology*, **74**, 184–189.

Thier, R., Lewalter, J., Kempkes, M. *et al.* (1999) Haemoglobin adducts of acrylonitrile and ethylene oxide in acrylonitrile workers, dependent on polymorphisms of the glutathione transferases GSTT1 and GSTM1. *Archives of Toxicology*, **73**, 197–202.

Thier, R., Lewalter, J., Selinski, S. and Bolt, H.M. (2002) Possible impact of human CYP2E1 polymorphisms on the metabolism of acrylonitrile. *Toxicology Letters*, **128**, 249–255.

Thorn, C.F., Aklillu, E., Klein, T.E. and Altman, R.B. (2012a) PharmGKB summary: very important pharmacogene information for CYP1A2. *Pharmacogenetics and Genomics*, **22**, 73–77.

Thorn, C.F., Ji, Y., Weinshilboum, R.M. *et al.* (2012b) PharmGKB summary: very important pharmacogene information for GSTT1. *Pharmacogenetics and Genomics*, **22**, 646–651.

Thorn, C.F., Klein, T.E. and Altman, R.B. (2010) Pharmacogenomics and bioinformatics: PharmGKB. *Pharmacogenomics*, **11**, 501–505.

Thornton, H. (2009) The UK Biobank project: trust and altruism are alive and well: a model for achieving public support for research using personal data. *International Journal of Surgery*, **7**, 501–502.

Varela-Lema, L., Taioli, E., Ruano-Ravina, A. *et al.* (2008) Meta-analysis and pooled analysis of GSTM1 and CYP1A1 polymorphisms and oral and pharyngeal cancers: a HuGE-GSEC review. *Genetics in Medicine*, **10**, 369–384.

Velasquez, J.L. and Lipkin, S.M. (2005) What are SNPs and haplotypes and how will they help us manage the prevention of adult cancer? *Current Oncology Reports*, **7**, 475–479.

Wang, H., Chanas, B. and Ghanayem, B.I. (2002) Cytochrome P450 2E1 (CYP2E1) is essential for acrylonitrile metabolism to cyanide: comparative studies using CYP2E1-null and wild-type mice. *Drug Metabolism and Disposition*, **30**, 911–917.

Whirl-Carrillo, M., McDonagh, E.M., Hebert, J.M. *et al.* (2012) Pharmacogenomics knowledge for personalized medicine. *Clinical Pharmacology and Therapeutics*, **92**, 414–417.

Wiley, J.C., Prattipati, M., Lin, C.P. and Ladiges, W. (2006) Comparative Mouse Genomics Centers Consortium: the Mouse Genotype Database. *Mutation Research*, **595**, 137–144.

Yeh, C.C., Sung, F.C., Tang, R. *et al.* (2009) Polymorphisms of cytochrome P450 1A2 and *N*-acetyltransferase genes, meat consumption, and risk of colorectal cancer. *Diseases of the Colon & Rectum*, **52**, 104–111.

Zhang, L., Zhou, J., Wang, J. *et al.* (2012) Absence of association between *N*-acetyltransferase 2 acetylator status and colorectal cancer susceptibility: based on evidence from 40 studies. *PLoS One*, **7**, e32425.

Zhang, W. and Dolan, M.E. (2010) Impact of the 1000 genomes project on the next wave of pharmacogenomic discovery. *Pharmacogenomics*, **11**, 249–256.

3
‘Omics Techniques

3.1 ‘Omics and bioinformatics

The term ‘*omics* is a convenient, colloquial way of referring to high-content technologies such as transcriptomics (which looks at RNA), proteomics (which looks at proteins) and metabolomics (which looks at the products of metabolism). ‘Omics techniques use HTS techniques to generate the huge amounts of data that are required in order to achieve a comprehensive understanding of biological processes such as pathways of toxicity, while bioinformatics is the term for the computational methodology used to store, analyse and interpret the vast quantity of data thus generated.[1] The recent rapid growth in the use of ‘omics throughout the life sciences has been a function of concurrent improvements in both types of technology.

‘Omics analysis requires the use of technologies which are able to work with complex biological samples robustly, flexibly and cost-effectively in order to generate results with high sensitivity and specificity. The data generated, however, make sense only if the experimental effort is supported by analytical resources and tools which make it possible to understand and interpret them, and this is where bioinformatics fits in. Bioinformatics, in turn, involves at least two types of activity: data mining (searching biological and literature databases to discover what is already known) and data analysis (comparing complex datasets from samples under different conditions, e.g. before and after treatment, diseased and healthy). It includes data processing and molecular identification, statistical data analysis, pathway construction and data modelling in a system-wide context.

[1] For a clear summary of ‘omics, bioinformatics and the relationships between them, see Schneider and Orchard (2011).

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/molecularcellulartoxicology

3.2 Transcriptomics

Transcriptomics is also known as transcriptional profiling, global analysis of gene expression or genome-wide expression profiling. This approach can be used to obtain information about the global pattern of RNA expression and the way in which it changes in response to a stimulus. It involves the analysis of large numbers of RNA transcripts using microarrays loaded with either short, single-stranded oligonucleotides (~20–60 nucleotides) synthesised *in situ* or cDNA applied by spotting or printing.[2]

3.2.1 Methodology

The most popular microarray platforms include those provided by Affymetrix[3] and Agilent.[4] The principle of transcriptional profiling using microarray methods is illustrated in Figure 3.1.

Experimental techniques Microarray analysis always involves the comparison of two samples (e.g. control and treated, or two different tissues) so the information it provides is relative rather than absolute. Figure 3.2 shows the workflow for a typical two-colour microarray experiment using oligonucleotide arrays where the aim is to investigate the effects of drug treatment; in this example RNA is isolated from treated and control tissues.

A typical transcriptional profiling experiment requires about 0.5 μg of mRNA, equivalent to about 20 μg of total RNA. If necessary, the mRNA component can be amplified by RT-PCR. The integrity and purity of RNA is checked (this is very important because the quality of the RNA samples used has a critical effect on the results obtained during transcriptional profiling) and the RNAs are labelled with fluorescent nucleotides (usually green cy3 for the control and red cy5 for the test). They are then mixed together and incubated with the microarray (usually overnight at 42–50 °C), preferably in a purpose-made hybridisation chamber which maintains constant temperature and humidity. After thorough washing, fluorescence data is collected using a laser scanner which quantifies fluorescence in the two-colour channels for each spot or gene on the array, generating a digital image. Following further processing to generate a numerical readout and remove background, with human intervention to remove artefacts if necessary, and normalisation to correct for errors in labelling, hybridisation and scanning, the reading is transformed to a ratio and represented as a false colour image. The outputs from transcriptional profiling may be presented in various formats, one of the most popular among them being the heat map (Figure 3.3).

The aspects of experimental design which must be considered carefully when planning a transcriptional profiling study are essentially those which apply to any biological experiment, but they merit reiteration in this context. They include the appropriate choice of sampling times (there is no point in undertaking a sophisticated analysis of gene expression at a time point before or after the response has

[2] For a review, see Trevino, Falciani and Barrera-Saldana (2007).
[3] http://www.affymetrix.com/estore/index.jsp
[4] http://www.genomics.agilent.com/en/home.jsp

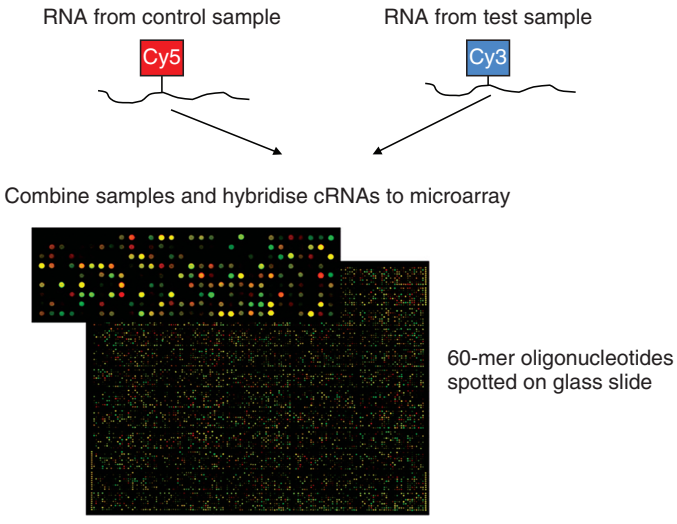


Figure 3.1 How two-colour transcriptional profiling works. The technique of transcriptional profiling using spotted oligonucleotide arrays is based on labelling RNA from your test and control samples with two different colours. RT primed from an oligo dT primer is used to incorporate fluorescent nucleotides in the first strand of cDNA synthesis. Once the RNA is labelled the two samples are combined and hybridised to the array which consists of oligonucleotides for each gene spotted on a glass slide. Two-colour labelling of the samples facilitates assessment of differences in the abundance of RNA species in the two samples. Where test RNA is labelled with green (Cy3) and control red (Cy5) if there is an abundance of RNA for a particular gene in the test sample the spot containing the probe for this gene on the array will turn green. No change is indicated by yellow, because there is an equal amount of red and green in the samples, and lower RNA for the gene in the test samples is indicated by a red spot. Hence, it is the differences in intensities of the two colours that enables you to determine if a gene is up- or down-regulated (source: Adapted from Dr Simon Plummer, MicroMatrices® Associates Ltd. Reproduced with permission of Dr Simon Plummer)

taken place!), dose and the endpoints to be measured. More general considerations include the following:

- **Pooling**: Pooling of samples can be legitimate if variability is small and the aim is to look at tissue effects rather than at single cells, but only if the dynamic response is likely to be reasonably consistent across the tissue under consideration. It should be noted that pooling can, in itself, create artefacts (e.g. by inducing cellular stress responses).
- **Replication**: This minimises unintended confounding and allows the true extent of variability to be estimated. It can improve both the precision and power of the analysis. Both biological and technical replicates are required (i.e. analysis of multiple independent samples and, where possible, replicate measurements on each sample).[5]

[5] For a detailed discussion of different experimental designs for microarray experiments, see http://discover.nci.nih.gov/microarrayAnalysis/Experimental.Design.jsp.

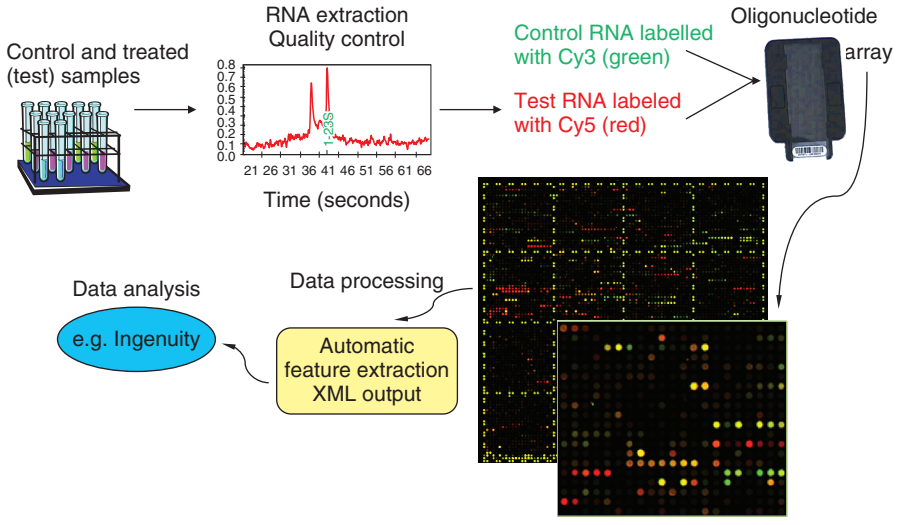


Figure 3.2 Transcriptional profiling workflow. This shows the workflow for a typical two-colour microarray experiment using oligonucleotide arrays where we are interested in examining the effects of a compound. RNA is isolated from drug treated and control tissues. Integrity and purity of RNA is checked. The RNAs are then labelled with fluorescent nucleotides, mixed and hybridised on the array. The data for each gene on the array is collected using a scanner which quantifies fluorescence in the two-colour channels for each spot or gene on the array. Data processing is an important step because it enables calculation of the statistical significance of the changes that have taken place. Lastly, we have to make sense of what these changes mean which involves the use of other software for data analysis (source: Adapted from Dr Simon Plummer, MicroMatrices® Associates Ltd. Reproduced with permission of Dr Simon Plummer)

Quality assurance The results of a transcriptional profiling experiment can only provide insights into underlying biological processes if all the sequences on the arrays used are correctly identified and annotated and the data are stored appropriately. In addition, the arrays must be manufactured to a consistent standard in order to yield meaningful results. This was a major issue in the early days of development of the technology but has been overcome to some extent by the increased use of well-characterised microarrays and reagents from reputable manufacturers.

The technical limitations of microarray-based methods include the following:

- They rely upon sequence-specific probe hybridisation and are therefore very dependent upon the range and quality of probes available.
- They can undergo background and cross-hybridisation problems.
- There can be issues with the dye-based detection methods used.
- Experimental design can be constrained by the types of chip or microarray available.

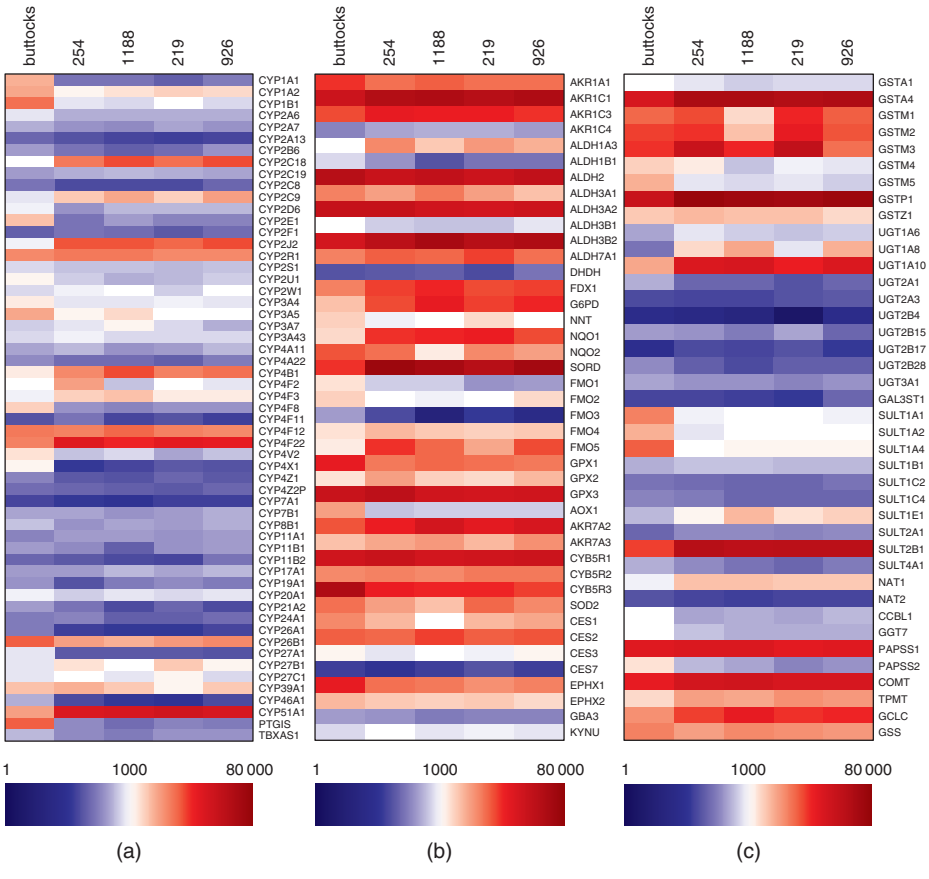


Figure 3.3 Heatmap of xenobiotic metabolising gene expression in human skin samples. The signal intensity values are shown on a logarithmic scale for each half of the range of signals covered. Lower signal intensity genes (<1000) are expressed in blue to white colour, higher signal intensity genes (>1000) in white to red colour. The signal intensity value of 1000 was chosen as the middle point of the scale because it was the closest in magnitude order to the median of the distribution of all signal values. (a) CYP genes (b) non-CYP Phase I genes and (c) Phase II genes (source: Hu *et al.* (2010); figure 2. Reproduced with permission of Elsevier)

The international standard for reporting microarray experiments is defined by Minimum Information About a Microarray Experiment (MIAME). According to the MIAME homepage,[6] MIAME describes the minimum information that is needed to enable the interpretation of the results of the experiment unambiguously and potentially to reproduce the experiment. The six most critical elements contributing towards MIAME are summarised in Box 3.1.

[6] http://www.mged.org/Workgroups/MIAME/miame.html

Box 3.1 Minimum Information About a Microarray Experiment – MIAME 2.0

The following six elements must be provided to support microarray-based publications.

1. **The raw data for each hybridisation**

   The raw data are defined as data files produced by the microarray image analysis software, such as CEL files for Affymetrix or GPR files for GenePix. These files should be provided in the native formats and should match their respective array designs.

2. **The final processed data for the set of hybridisations in the experiment (study)**

   The processed data is defined as the normalised and/or summarised data on which the conclusions in the related publication are based. For instance, these can be MAS5 or RMA normalised data matrices for Affymetrix data. In gene expression experiments the final processed data is typically a matrix of genes and experimental conditions characterising the expression of each gene under each condition. The identifiers used in these processed data files should match the array annotation or locations on the arrays.

3. **The essential sample annotation, including experimental factors and their values**

   Experimental factors (conditions) and their values are the most essential information about the samples used in the experiment. The experimental factors are the key experimental variables in the experiment, for instance 'time' in time series experiments, 'dose' in dose–response experiments, 'compound' in compound treatment experiments, or 'disease state' (normal or otherwise) in disease studies. The same experiment may have several experimental factors, for example, compound, dose and time may all be experimental factors in a dose response experiment in which several compounds are used to treat samples over a time course. In addition to experimental factor values, additional sample information that is required to interpret the experiment must be given, for instance, the organism and organism part from which the sample has been taken.

4. **The experiment design including sample data relationships**

   The purpose of the experimental design description is simply to specify the essential relationships between different biomaterials, such as samples and arrays, and the data files which are produced in each hybridisation. In a simple one channel-one sample-one array experiment, this may be a table listing all samples and the respective raw data files. If relevant, it is important to show which hybridisations in the experiment are replicates, and which are technical and which are biological replicates. More generally, the experimental design can be described as a graph where nodes represent biomaterials (e.g. samples or their sources) and data objects (e.g. files), and

edges or arrows show their relationships. MAGE-TAB provides a simple format to encode such graphs.

5. **Sufficient annotation of the array design**

   Essential array design information is the reporter (probe) sequence information and/or the database accession numbers that characterise a sequence. For synthetic oligonucleotides the precise DNA sequence must be given. For commercial or other standard array platforms this information is typically provided by the array vendors or manufacturers.

6. **Essential experimental and data processing protocols**

   The essential laboratory and data processing protocols are usually described in the journal methods section. It is sufficient to simply refer the standard experimental or data processing protocols, such as MAS5 or RMA. However, if a protocol depends on parameters that can be varied, their values should be provided. If novel or non-standard data processing protocols are used, these should be described in sufficient detail to allow the user to understand what exactly has been done in the experiment and how the data has been analysed to reach the conclusions of the study.

(source: Minimum Information About a Microarray Experiment – MIAME 2.0. Reproduced with permission of the Functional Genomics Data Society)

Data processing Data processing is a critical part of transcriptional profiling because of the large number of genes being analysed in each experiment. This can be a complicated and time-consuming process, but essentially it involves determination of the statistical significance of the changes that have taken place and determination of the meaning of the changes observed using pathway analysis software such as Ingenuity Pathway Analysis (Ingenuity® Systems).[7] The proper statistical analysis of transcriptional profiling data is critical because of the number of simultaneous analyses being conducted (think about it: for an experiment considering 10 000 genes, the application of the conventional test for statistical significance ($p \leq 0.05$) would generate no less than 500 false positives).

Databases The process of converting raw data into knowledge and understanding starts with collecting, selecting, cleaning and storing the data in order to populate a database. The desired output of this is an archive of data stored in a uniform and efficient manner. This primary (archived) database contains information directly deposited by contributors and can include published data, DNA sequences, protein structures and expression profiles, depending on the nature of the study being conducted, while secondary (derived) databases are so called because they contain the results of further analysis of primary database information.

Transcriptomics shares with all the 'omics approaches the challenge of handling ever-increasing amounts of complex information effectively and flexibly. In order to meet this challenge, formal methods have been developed to allow information to be managed consistently and transparently. The methods developed for

[7] http://www.ingenuity.com/

this involve the use of controlled vocabularies and ontologies to organise large quantities of information.

A controlled vocabulary is a standardised dictionary of terms used to represent and manage information. This allows the database to be subjected to keyword searching with a degree of confidence that all relevant information within it will be located. One such controlled vocabulary, used to capture disease data, is called MEDIC (*MErged DIsease voCabulary*) (Davis *et al*., 2012). MEDIC, which is freely available for downloading, takes advantage of existing disease vocabularies such as MeSH (as used in MedLine) and is updated monthly to ensure currency. Using MEDIC, chemicals and genes can be associated with a disease process via either of two types of interaction: M indicates a biomarker or part of a molecular mechanism and T represents a putative therapy or therapeutic target.

Ontologies such as Gene Ontology use species-independent terminology to describe gene products in terms of the biological processes in which they are involved, the cellular components with which they interact and their molecular functions.

As transcriptional profiling and other 'omics techniques have become more and more central to toxicology numerous publicly accessible databases[8] have been established with the aim of making 'omics findings available to the wider scientific community:

- **Chemical Effects in Biological Systems (CEBS):** The CEBS database[9] is a public repository for toxicogenomics data developed by the US National Centre for Toxicological Research, part of the NIEHS. CEBS was the first public repository to be able to integrate biological, toxicological and histopathological observations with the outcomes of 'omics studies. Development of the CEBS database, which can store data from studies in experimental animals, cultured cells or humans, commenced in 2002 and initially focused on the capture of microarray and proteomics data. In the second phase of development, which started in 2003 and was completed in 2006, information about study design and toxicological outcomes was incorporated (Waters *et al*., 2008).
  During the development of the CEBS database, careful consideration was given to the need to ensure a common minimum standard for the data deposited, capture all relevant data and present these data in such a way that they can be interrogated effectively by users. The data fields used were initially defined based upon terminology commonly encountered in a standard acute toxicity study, but the database was designed to be extendable so as to capture relevant information from other types of study (e.g. carcinogenicity, reproductive toxicity, neurotoxicity) (Fostel *et al*., 2005). This approach allows 'omics data to be integrated with pathophysiological observations, thus facilitating phenotypic anchoring (the process by which gene expression changes are linked to changes in phenotype).
  The CEBS data dictionary (CEBS-DD[10]) was developed to allow archiving, retrieval and exchange of high-content data sets including transcriptional profiles. This describes incoming data on the basis of alignment with public

[8] For a complete list of current 'omics databases, see http://www.ebi.ac.uk/services
[9] http://www.niehs.nih.gov/research/resources/databases/cebs/index.cfm
[10] http://www.niehs.nih.gov/research/resources/databases/cebs/

standards and proprietary data formats and feeds in to the Biomedical Investigation Database,[11] which uses a workflow design to capture information about study design and histopathology.

- **Comparative Toxicogenomics Database** is a community-supported public resource based at the Mount Desert Island Biological Laboratory with support from NIEHS and the National Centre for Research Resources of the National Institutes of Health (NIH). Its aim is to help researchers understand the connection between environmental chemicals and gene products, and their effects on human health focussing on the effects of environmental chemicals and functioning as a tool for hypothesis generation as well as a data repository (Davis *et al*., 2011). Text mining is used to prioritise data for inclusion, but the information in the database is curated manually to ensure its validity. As of July 2010, data from 21 600 publications had been curated, making a total of approximately 1.4×10^6 chemical-gene-disease connections available for exploration and analysis. The database also contains Kyoto Encyclopedia of Genes and Genomes (KEGG)[12] and reactome pathway data, links to the gene pages at PharmGKB and the chemical pages at DrugBank and enhanced data features including enriched gene ontology terms which can be used to understand chemical interactions as well as gene–disease associations.
- **DrugMatrix®:** The DrugMatrix® database[13] was originally generated commercially by Iconix Pharmaceuticals, but was subsequently purchased and released into the public domain by the NIEHS (Ganter *et al*., 2005). It represents the results of experiments in which 600 compounds (including 400 FDA-approved drugs) were administered to rats and the effects were profiled in up to seven tissues using Uniset Rat I and Human I Expression Bioarrays representing some 10 000 genes. Two dose levels, the maximum tolerated dose (MTD) and a 'fully effective dose' (i.e. the therapeutically active dose, or 10-fold below the high dose for non-pharmacologically active compounds), were used in order to capture both toxic and pharmacological gene expression profiles.
- **Toxicogenomics Project – Genomics Assisted Toxicity Evaluation System (TG-GATES):** The Toxicogenomics Project (TGP) in Japan[14] was coordinated by the Japanese National Institute of Health Sciences and NI Biomedical Innovation Institute. It ran from 2002 to 2007, involved government investigators and 16 Japanese pharmaceutical companies and studied the effects of toxic chemicals in the rat, focussing on effects in the liver and kidney (Uehara *et al*., 2010). Bridging studies were also carried out in human and rat hepatocytes. The aim of this programme was to investigate the value of single dose and 28-day toxicogenomics studies in predicting the outcomes of lifetime bioassays. Single dose studies were conducted at three doses (usually a basal dose, 3 times and 10 times basal dose) with termination after 3, 6, 9 and 24 h; repeated dose studies were terminated after 3, 7, 14 and 28 days with termination at 24 h after the final dose.

[11] https://dir-apps.niehs.nih.gov/arc/
[12] http://www.genome.jp/KEGG
[13] https://ntp.niehs.nih.gov/drugmatrix/index.html
[14] http://www.tgp.nibio.go.jp/index-e.html

The resulting data (Affymetrix Rat RG230-2.0 array data and conventional toxicological parameters for prototypical toxicants in rat liver and kidney and in cultured rat and human hepatocytes) were deposited in the TG-GATES database,[15] as a single-project, open access database. The database can be searched by compound name or pathological finding, and gene expression data associated with specific phenotyping information can be downloaded. A database of pathological images is also available.[16] The usefulness of the TG-GATES resource has been recognised internationally, including within the EU (Gocht and Schwarz, 2012).

Example: Identifying biomarkers of non-genotoxic carcinogenesis

The TG-GATES database has proved useful in identifying biomarkers of effect for non-genotoxic carcinogens whose mode of action is thought to involve oxidative stress. In a comparison of methapyrilene and thioacetamide, for example, 112 genes whose expression profile can be used to predict oxidative stress with high precision were identified in 24 h and 28 day studies in 5-week-old Sprague-Dawley rats. The use of this panel of biomarkers has been claimed to be able to predict oxidative stress-related hepatocarcinogenicity as early as 24 h after a single dose (Uehara *et al*., 2008). In a follow-up study (Uehara *et al*., 2011), data from 28 day studies was used to refine the analysis with the aim of reducing the number of false positives generated. Using data from all 150 compounds in the TG-GATES database it was possible to refine the predictions to give 99% sensitivity and 97% specificity, with almost complete elimination of false positives. This method was found to be predictive for both genotoxic and non-genotoxic carcinogens. The authors believe that it will have utility both in prospective screening for hepatocarcinogenicity and in the prioritisation of substances for carcinogenicity testing.

Example: Liver toxicity related to defects in blood clotting

Abnormal haemostasis is often associated with hepatotoxicity but the precise mechanisms involved are unclear. They may include reduced synthesis, increased consumption and/or reduced clearance of clotting factors. In another application of the TG-GATES approach, Hirode *et al*. (2009) attempted to identify mechanism-based biomarkers for the diagnosis/prognosis of hepatotoxicity-related coagulopathies.

The known mechanisms of the action of aspirin and propylthiourea were confirmed by principal component analysis and shown to be independent of liver enzyme release. For the other compounds tested, the dominant changes were in lipid metabolism and blood clotting. There was a decrease in the mRNA levels of some clotting factors, but this was not associated with liver enzyme release, suggesting that toxicity was due to a metabolic defect rather than necrosis. There was also evidence for a possible compensatory mechanism involving vitamin K. The gene Vkorc03, which synthesises vitamin K from vitamin K epoxide, was generally down-regulated by hepatotoxins, but the paralogous gene Vkorc03L1 was

[15] http://dbarchive.biosciencedbc.jp/en/open-tggates/desc.html
[16] http://dbarchive.biosciencedbc.jp/en/open-tggates-pathological-images/desc.html

up-regulated suggesting that the system was trying to maintain vitamin K homeostasis in the face of disruption due to hepatotoxin exposure.

The TGP developed a new scoring method called the *differentially expressed gene score* (D-score), a statistical approach to data analysis which helps to alleviate the effects of poor data quality on the analysis of transcriptomic data. This parameter, first reported in 2009 (Kiyosawa *et al*., 2009), captures the overall direction of expression changes as well as their magnitude and is well suited to HTS. Further analysis of the database using network inference approaches based on gene set-level data rather than individual gene data has made the analysis more robust, reduced the number of individual inputs required and improved comparability between studies, thus enhancing the gene set and phenotype-level relational capability of the database (Kiyosawa *et al*., 2010).

Example: Use of D-scores to analyse the effects of bromobenzene on rat liver

An example of the way that D-scores can be used to explain the transcriptional effects of toxins is illustrated in Figure 3.4. Rats were treated with 300 mg kg^{-1} bromobenzene, and their livers were collected at 2, 6, 12 or 24 h after the treatment, and a GeneChip® analysis was conducted. Red and blue colours indicate high and low D-scores for the gene sets, respectively. After 2 h, the effects seen were confined to DNA damage. As time went on the effects spread to include oxidative stress and GSH depletion; subsequently pathways related to GSH homeostasis, endoplasmic reticulum (ER) stress, inflammation and carcinogenesis were activated while cholesterol synthesis and steroid hormone metabolism were activated.

Toxicogenomic data may also be deposited in generic repositories such as ArrayExpress, ArrayTrack, and Gene expression Omnibus; however, specialist databases such as TG-GATES and DrugMatrix® have a number of advantages over these. Between them, TG-GATES and DrugMatrix® represent a repository of data from approximately 30 000 liver gene expression profiles associated with some 500 drugs having a wide range of indications, including data from both *in vitro* and *in vivo* studies generated using uniform experimental designs (Chen *et al*., 2012).

3.2.2 Proof of principle

A relatively simplistic approach to the problem of analysing changes in hepatic gene expression in response to hepatotoxicants is to use customised or focused DNA microarrays. This approach, which was particularly popular in the early days of transcriptional profiling when commercial systems were prohibitively expensive, is relatively cost-efficient and requires less time than a whole-genome approach.

Example: Use of a home-made microarray to characterise hepatotoxicity

In a method-development exercise, Inadera *et al*. () developed a home-made cDNA microarray, the Mouse Liver Stress Array Version 1.0, containing 355 unique genes involved in drug metabolism, cellular stress, inflammation and liver-related proteins. These genes were selected as being representative of

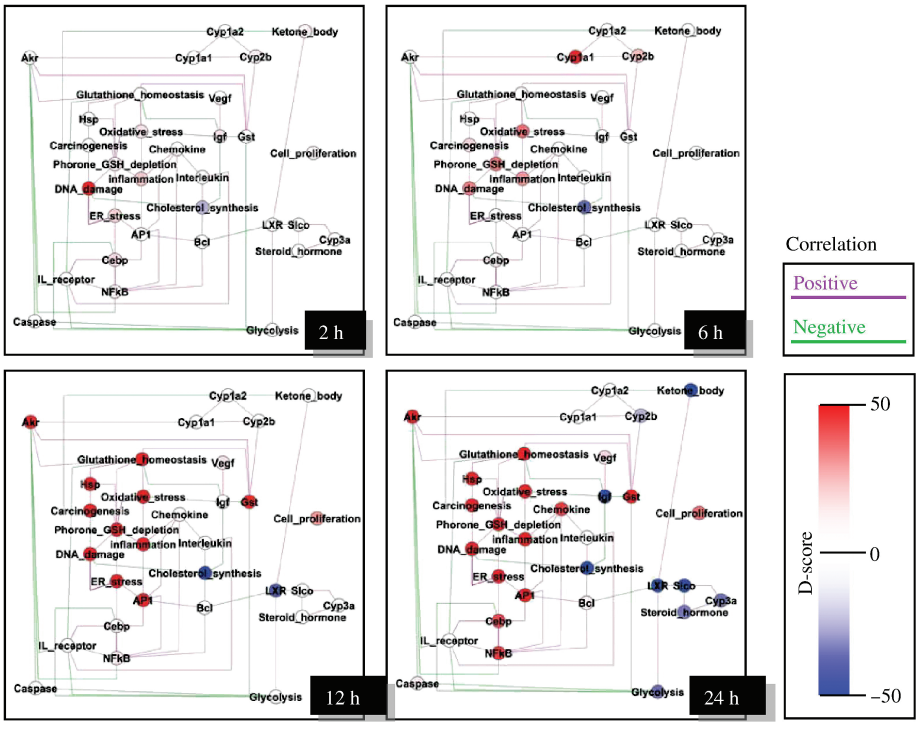


Figure 3.4 Time course of D-score profile for bromobenzene-treated rat liver. Rats were treated with 300 mg kg^{-1} bromobenzene, the livers were collected at 2, 6, 12 or 24 h after the treatment, and GeneChip® analysis was conducted. Red and blue colours indicate high and low D-scores for the gene sets, respectively. At 2 h, the D-score for DNA damage genes was increased, indicating that the genes associated with DNA damage were overall up-regulated at this time point. At 6 h, D-scores for glutathione depletion and oxidative stress were increased, and those for *Cyp1a1* and *Cyp2b* genes were increased as well. At 12 h, the increased D-scores for HSP, carcinogenesis, glutathione homeostasis, ER stress and AP1 genes were observed, suggesting that the increased D-score for DNA damage genes triggered at 2 h affected its neighboured nodes and the area of activated nodes was propagated. In addition, the increase in the D-scores for *Cyp1a1* and *Cyp2b* genes disappeared at 12 h, and instead those for *Akr* and *Gst* genes, which are first neighbours of *Cyp1a1* and *Cyp2b* genes, respectively, were increased at 12 h. Furthermore, the D-score for cholesterol synthesis genes was decreased, which has a negative correlation with DNA damage genes. At 24 h, the D-scores for *Cebp* and *NfkB* genes were increased and those for *LXR* and *Slco* genes were decreased. Overall, the established network inferred based on the TG-GATEs data set was well applicable to microarray data on bromobenzene-treated rat liver which was obtained from outside TGP (source: Kiyosawa *et al.* (2010); figure 6. Reproduced with permission of Elsevier)

homeostatic control and detoxification processes in the liver. The transcripts represented on the Mouse Liver Stress Array included 197 genes related to drug metabolism (80 transferases, 75 oxidoreductases, 18 hydrolases, 13 isomerases and 11 lyases), 117 inflammation-related genes (cytokines/chemokines and their receptors), 35 genes known to be highly expressed in the liver (e.g. α2-HS-glycoprotein, apoE, MUP-1) and 6 positive/negative control genes.

Using this customised microarray, they analysed gene expression changes in the mouse liver treated by 11 known hepatotoxicants which are believed to act via a range of mechanisms:

- The oestrogen 17β-estradiol (0.2 mg kg^{-1})
- The putative endocrine disrupting chemical bisphenol A (100 mg kg^{-1})
- The Gram-negative bacterial cell wall component lipopolysaccharide (LPS: 4 mg kg^{-1})
- The solvent trichloroethylene (500 μl kg^{-1})
- The nuclear receptor ligand pregnenolone 16α-carbonitrile (PCN: 100 mg kg^{-1})
- The tumour promoter phenobarbital (PB) (10 or 100 mg kg^{-1} /day^{-1} for 3 days)
- The plasticiser and nuclear receptor ligand diethylhexylphthalate (DEHP: 100 mg kg)
- The biocide triphenyltin (10 mg kg^{-1})
- The solvent and dry cleaning agent carbon tetrachloride (5 or 50 mg kg^{-1})
- The nephrotoxin cadmium chloride (1 or 10 mg kg^{-1})
- The nuclear receptor ligand 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP; 0.3 mg kg^{-1})

The data were analysed by means of heat mapping and Ingenuity Pathway Analysis. As this was primarily a method-development exercise, no novel findings were reported; rather, the analysis focussed on establishing that the changes in gene expression which occurred in response to each compound were those expected from previous studies. However, the paper does report one or two interesting findings; for example, the pattern of gene expression changes seen in response to bisphenol A (a compound used to make plastic clear and shatterproof) were similar to those seen with 17β-estradiol, consistent with its proposed endocrine disrupting effects, whereas a completely different profile was observed following treatment with DEHP (which is also found in plastics, being used as a plasticiser). Overall, this study may be viewed as proof-of-principle for the use of transcriptional profiling to characterise the molecular pathways involved in hepatotoxicity.

3.2.3 Hepatotoxicity

Transcriptomic approaches are becoming increasingly popular in the study of hepatotoxicity, both *in vitro* and *in vivo*. The following are a few examples, but new ones are being published all the time.

Example: Hepatotoxicity of acetaminophen

One powerful way to use microarray analysis in target organ toxicity is to compare toxic and non-toxic compounds with similar structures, for example, acetaminophen (*N*-acetyl-*p*-aminophenol) and its non-toxic regioisomer *N*-acetyl-*m*-aminophenol. In a study comparing these compounds (Priyadarsiny *et al*., 2008), C57BL/6J mice were treated with each compound at 400 mg kg^{-1} intra-peritoneally (i.p.). The major pathways affected by acetaminophen were

energy homeostasis and tissue repair machinery. Both compounds caused liver damage within 6 h of dosing but the effect of acetaminophen was approximately twice that of *N*-acetyl-*m*-aminophenol. Centrilobular necrosis was observed with acetaminophen but not with *N*-acetyl-*m*-aminophenol.

The transcriptional changes associated with these responses were examined using the CodeLink Uniset I mouse 20K oligonucleotide array. Differential regulation of 896 genes was observed; 648 of these were up-regulated and 248 were down-regulated. Acetaminophen up-regulated 62 genes and down-regulated 70 genes to a greater extent than did an equivalent dose of *N*-acetyl-*m*-aminophenol. The genes up-regulated by acetaminophen play a role in apoptosis, stress responses, cell growth and proliferation, cell signalling, inflammatory responses, adhesion and cell structure. Those which were down-regulated are involved in lipid metabolism, development, signal transduction and cell growth. CYP expression was also down-regulated in response to acetaminophen. The main family of genes up-regulated by acetaminophen were those of the mitogen-activated protein kinase (MAPK) and cytokine signalling pathways. This seemed to be associated with the induction of inflammation and fibrosis.

Example: Strain and species differences in hepatotoxicity

One disadvantage of studies which use inbred rodent strains to analyse transcriptomic changes in target organ toxicity is that they do not adequately reflect the variability of human population. Attempts have been made to address this problem by using a laboratory mouse strain diversity panel comprising 36 inbred strains derived from *Mus mus musculus*, *Mus mus castaneous* and *Mus mus domesticus* (Harrill *et al*., 2009). Acetaminophen (300 mg kg^{-1}) induced centrilobular necrosis in the liver 24 h after dosing, consistent with effects reported previously. This was associated with minor inflammatory infiltration into the hepatic parenchyma as well as varying degrees of intrahepatic haemorrhage. A wide range of toxic responses occurred among the mouse strains tested: 30 strains exhibited mild necrosis (<40% damage) while 6 exhibited severe necrosis (40–100% damage). The most sensitive strains were CBA/J and B6C3F1/J.

Transcriptional profiles associated with strain, treatment and liver necrosis were identified. It was possible to identify 26 population-wide biomarkers of acetaminophen hepatotoxicity (i.e. transcripts affected significantly by treatment and toxicity outcome but not by genotype), 17 of which were up-regulated while 9 were down-regulated. They represented a closely linked molecular network of intracellular pathways associated with hepatocyte death as well as genes which had not been associated with the effects of acetaminophen previously. The authors propose that these represent mechanism-relevant biomarkers of liver necrosis in the mouse, although they note that they have not yet been linked to human responses.

Attempts to use transcriptional profiling to identify susceptible and resistant genotypes have also been made in rats (Yun *et al*., 2009). Male Sprague-Dawley rats were subjected to a pre-treatment liver biopsy (removing ~10 mg of tissue), allowed to recover for 3 weeks, and then treated twice with carbon tetrachloride (2 ml kg^{-1} periorally (p.o.)). Blood was collected 24 h after the second dose and liver enzyme release was measured. Ten rats, five with high liver enzyme

levels ("susceptible") and five with low levels ("resistant"), were selected for transcriptional profiling using the Affymetrix GeneChip® Rat Gene 1.0 ST array system. Note that this was conducted using the biopsy tissue taken prior to treatment, not after carbon tetrachloride exposure. The results indicated wide inter-individual variation in hepatotoxicity associated with hepatocellular steatosis and inflammatory cell infiltration. The genes identified by transcriptional profiling were subjected to functional classification using the PANTHER system (Protein ANalysis THrough Evolutionary Relationships).[17]

The susceptible group of rats had innately lower expression of genes associated with immunity and defence, lipid, fatty acid and steroid metabolism, transport systems (including anion transport) and complement-mediated immunity. The conclusion drawn was that susceptibility to fatty liver induced by toxicants such as carbon tetrachloride is a function of imbalances in hepatic fatty acid flow and triacylglyceride synthesis and excretion.

Example: Species-specific responses to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD)

The responses of orthologous mouse and rat genes to TCDD were initially analysed in Sprague-Dawley rats and C57BL/6 mice, indicating surprisingly limited species overlap: only 33/3087 orthologous genes showed the same response in both species (Boverhof *et al.*, 2006). This represented 15.1% of the rat genes which responded to TCDD and 12.8% of the mouse genes, so approximately 85% of hepatic responses to TCDD were species-specific. The pattern of gene expression observed was consistent with a response mediated by Nrf2, and the results suggested that a small core of genes is responsible for mediating common processes of TCDD hepatotoxicity in rats and mice.

In a follow-up study, the responses of C57BL/6 mice and Long Evans rats were compared (Boutros *et al.*, 2008) using the MOE430-2 array of 45 101 probes for mouse and the RAE230A array of 15 293 probes for rat. Again, the commonality between species was very limited (33/8125 orthologous genes, three of which changed in opposite directions). Interestingly, genes which responded to TCDD in one species were less likely to respond in the other species than would be expected by chance alone.

This is surprising because the hepatotoxic outcomes observed in mice and rats following TCDD treatment are very similar. It seems likely that aspects of hepatotoxicity which are common to both species are mediated by members of the core group of 30 genes which exhibit species-independent regulation by TCDD. This includes *Cyp1a1*, *Cyp1a2*, *Cyp1b1*, *Nqo1* and *Tiparp*. The identification of these genes is reassuring in that the *Cyp1* gene family is representative of a common core set of genes which were previously known to be subject to modulation by TCDD in both rats and mice.

It should be noted that the time point chosen for this study was 19 h, prior to the onset of overt toxicity. It is possible that the results of analysis during the fulminant phase of hepatic toxicity would provide a more clear-cut answer. Overall, this study indicated that mRNA changes in the hepatocytes themselves do not

[17] http://www.pantherdb.org/

necessarily relate directly to hepatotoxicity. One possibility is that changes outside the liver (e.g. in the haematopoietic system) play a key role in determining hepatic target organ toxicity; another is that changes in a minor cell population such as the Kupffer cell play the central role in toxicity. The third possibility, that common pathways are perturbed through different genes in each species, was not supported by the results of this study.

The suggestion of profiling additional strains of rat has been acted upon by Moffat *et al*. (2010), who looked at TCDD responsiveness in the TCDD-resistant Han Wistar rat and the sensitive Long Evans strain. The LD_{50} of TCDD in Han Wistar rats is 1000 times higher than that in the Long Evans strain due to a point mutation in the Ah-receptor (AhR) gene which leads to the expression of a truncated mRNA with 38 or 43 amino acids missing from the transactivation domain.

Two additional strains have been derived from Han Wistar and Long Evans by breeding:

- LnA carries the variant AhR and is dioxin-resistant.
- LnC is wild type for AhR and is dioxin-sensitive.

In the study of Moffat *et al*. (2010), rats were dosed with TCDD ($100\,\mu g\,kg^{-1}$) and hepatic mRNA abundance was analysed 19 h after TCDD treatment using the Affymetrix RAE230A gene array. The resistant strains exhibited fewer transcriptional changes in response to TCDD than did the sensitive strains, but *Cyp1a1*, *Cyp1a2* and *Cyp1b1* were responsive in all strains. Responses to TCDD were classified as Type I (similar in sensitive and resistant strains; 25 genes in this study) or Type II (differ between sensitive and resistant strains; 46 genes in this study). Type II genes appear to be those which are integral to the mechanisms of hepatotoxicity, wasting and lethality. This class of genes, which accounted for the main differences between the strains, were associated with pathways relating to lipid metabolism, cellular membrane function and energy metabolism. In the resistant strains there was no evidence for derailment of energy metabolism via effects of lipid homeostasis, protein metabolism or ATP production/utilisation.

Example: Transcriptional profiling in precision-cut liver slices

The cellular pathways leading to organ-specific toxicity are often multifaceted, involving interactions between several cell types and biochemical networks of cell–cell and cell–matrix interactions. Precision-cut tissue slices are relatively easy to prepare and retain normal tissue architecture. They can be prepared from a variety of tissues, although they are most commonly prepared from liver and kidney using a purpose-designed tissue slicer such as the Brendel-Vitron model. Livers and kidneys are easily cored and sliced by this methodology and mouse liver slices made in this way, for example, retain their viability for at least 24 hrs. The advantage of using precision-cut tissue slices for transcriptional profiling studies, as in other contexts, is that all cell types (including Kupffer cells) are represented and the normal cell–cell interactions of the liver remain intact in the

interior of the slice. In addition, the fact that multiple slices can be made from a single organ sample means that replication is possible and allows statistical methods to be applied.

Elferink *et al*. (2008) postulated that transcriptomic analysis of precision-cut liver slices could be used to identify a small set of genes which would be indicative of a specific mechanism of hepatotoxicity and could be used in the screening of new drugs. The following four model hepatotoxins were used:

- LPS (1 mg kg^{-1} i.p.) causes toxicity via an inflammatory mechanism.
- Acetaminophen (1800 mg kg^{-1} p.o.) causes necrosis.
- Carbon tetrachloride (8 g kg^{-1} p.o.) causes both fibrosis and necrosis.
- Gliotoxin (3 mg kg^{-1} i.p.) causes apoptosis.

Transcriptomic analysis was carried out using the Affymetrix RGU34A array and the data were analysed using ToxShield™ software,[18] which allows prediction of human hepatotoxicity from rat gene expression data. The ToxShield™ model is based on 100 compounds with diverse mechanisms of toxicity (necrosis, cholestasis, micro/macrovesicular steatosis, hepatitis, genotoxicity, non-genotoxic carcinogenicity and nuclear receptor agonism). It also includes data on nephro-and cardiotoxicity as well as data from non-toxic compounds.

The results indicated that acetaminophen, LPS and carbon tetrachloride induced time-dependent increases in the number of genes whose expression was altered whereas in the case of gliotoxin the number of genes altered reduced with time. This suggested that gliotoxin induced a transient effect whereas the effects of the other three toxins were progressive. Only a very small number of genes were common to all the four toxins. The ToxShield model made predictions which corresponded with the known mechanisms of action of three of the four test compounds; the exception was gliotoxin, which was predicted to be non-toxic, possibly as a function of the concentration used in this particular experiment.

In a follow-up study, the same team extended their analysis to include precision-cut human liver slices (Elferink *et al*., 2011). They examined the stability of gene expression in precision-cut liver slices and measured inter-individual variability due to donor age and sex as well as the source from which the liver was obtained. In order to evaluate toxic responses in this system, slices were treated with acetaminophen (1 mM or 5 mM for 3 or 24 h). Analysis was conducted using the Affymetrix Human Genome U133 Plus 2.0 array, which carries 54 675 transcripts representing some 25 000 genes. The results indicated that gene expression was reasonably stable, the main changes being associated with remodelling of the cytoskeleton and the cell adhesion, while there were hardly any changes in the expression of genes associated with xenobiotic metabolism, cellular stress and toxicity. Towards the end of the 24-h incubation period, changes in gene expression consistent with hepatic stellate cell activation, proliferation of myofibroblast-like cells and deposition of collagen began, suggesting that fibrogenic pathways had begun to function. There was little evidence of inter-individual variability due to the age or sex of the donor, nor to pathological

[18] http://www.genelogic.com/knowledge-suites/toxexpress-program/toxshield-suite

condition. The effects of acetaminophen were consistent with its known mechanism of toxicity (impairment in energy metabolism and mitochondrial function, oxidative stress, signalling, apoptosis and inflammation).

3.2.4 Extrahepatic toxicity

Many toxic compounds exert effects in more than one tissue, and the effort involved in analysing such effects on a one-by-one basis has, in many cases, already taken many years without leading to a full understanding. 'Omics techniques such as transcriptional profiling make it possible to obtain a global overview of molecular effects in different tissues, while bioinformatic approaches allow the similarities and differences between the responses of different tissues to be identified.

Example: Nephrotoxicity of TCDD

The toxic effects of TCDD include thymic atrophy, hepatotoxicity, teratogenesis, a wasting syndrome and ultimately death. The kidney is also a target for TCDD toxicity during murine foetal development (Boutros *et al.*, 2009). The vast majority of TCDD-induced toxicities are mediated via the AhR, but the pathways that lie between AhR activation and specific toxicities are mostly unclear. In the liver it is known that:

- The response of adult liver to TCDD requires the presence of a functional AhR protein.
- Large numbers of genes are perturbed by TCDD in wild-type liver.
- Large numbers of genes are perturbed by genetic ablation of the AhR.
- These genes lie on specific pathways which are associated with specific transcription factor binding sites which also have a role in basal liver function.

In a study where the aim was to establish similarities and differences between hepatic and renal responses to TCDD, adult male C57BL/6 mice and mice lacking the Ah-receptor (AhR null) were dosed with TCDD (1 mg kg^{-1} p.o.) and sacrificed 19 h later (adult mice were used for this microarray study for practical reasons). Hepatic and renal gene expression were analysed using Affymetrix MOE430-2 microarrays.

Only five genes were altered by TCDD in the kidneys of AhR null mice, and only seven genes in the liver, suggesting that a functional AhR is required for almost all transcriptional responses to TCDD in murine liver and kidney. Even in wild-type C57BL/6 mice, only 17 kidney genes exhibited AhR-dependent changes in response to TCDD whereas 297 liver genes were regulated in this manner. This suggests that the adult kidney is minimally affected by TCDD, at least at this time point and dose. The basal expression of 44 genes was altered in both the kidney and liver of untreated AhR null mice; five of these responded to TCDD in both tissues.

In this study the response of the kidney clearly differed from that of the liver in that the adult kidney was relatively refractory to TCDD toxicity; however,

further work is required in order to understand the renal effects of TCDD at the transcriptional level because renal transcriptional responses differ markedly between adult and foetal mice.

Example: Metabolic acidosis

Transcriptional profiling has also been used to address adaptive responses in renal acid excretion which occur in response to disturbances in acid-base balance.

Microarray gene expression analysis has been used to study the molecular consequences of acute (2 days) and subchronic (7 days) metabolic acidosis in C57BL/6J mouse kidney (Nowik *et al*., 2008). Metabolic acidosis was induced by administering ammonium chloride/2% sucrose via the drinking water, the control being sucrose only. Transcriptional changes were analysed using Applied Biosystems Mouse Genome Survey microarrays containing 32 996 curated genes and 44 498 possible mRNAs.

Approximately a third of all the genes on the array were expressed at detectable levels in the kidney and more than 4000 genes were affected by metabolic acidosis. The pathway most affected by metabolic acidosis was oxidative phosphorylation; genes encoding enzymes and transporters associated with ammonia synthesis were also affected. The largest group of transcripts affected was the encoding solute carrier transporters of the Slc036 (monocarboxylic acid) and Slc25 (mitochondrial) gene families.

Overall, the kidney was strongly affected as a whole by metabolic acidosis at many different levels, the main effects being on energy homeostasis, acid excretion and water/electrolyte balance. Genes involved in apoptosis and cell proliferation were also markedly affected, suggesting that an adaptive remodelling process was being induced.

3.3 Proteomics

Proteomic tools produce a global view of complex biological systems by examining complex protein mixtures using large-scale HTS technologies (Barrier and Mirkes, 2005). Proteomic methods speed up the processes of protein separation, quantification and identification and, depending upon the methodology used, can address post-translational modifications and protein–protein interactions as well as accessing 'the parts other [techniques] cannot reach' (e.g. biofluids and subcellular compartments (Merrick and Madenspacher, 2005). Toxicoproteomics is the subdiscipline of proteomics that integrates traditional toxicology and pathology with differential protein and gene expression data to generate information about the effects of toxic compounds at the protein level.[19]

The methodology for proteomic analysis has developed rapidly during the first decade of the 21st century. Around the year 2000, it was perfectly acceptable to carry out a so-called proteomic analysis by hand using conventional 2D-gel electrophoresis whereas by 2010 many proteomic studies depended upon the availability of highly developed analytical methods, usually based on mass spectrometry (MS). There was a corresponding increase in the amount and quality of data which could be generated but also in the costs incurred.

[19] Reviewed by Wetmore and Merrick (2004).

3.3.1 Methodology

Experimental techniques Over the last few years proteomic methodology has evolved from conventional biochemical techniques such as 2D-gel electrophoresis to HTS technologies including various types of MS and microarray methods (Merrick and Madenspacher, 2005).

As the name suggests, 2D-gel electrophoresis involves separating proteins in two dimensions, usually by isoelectric point in one dimension and molecular weight in the other. The proteins on the gel are visualised (e.g. by silver staining), digested *in situ* (e.g. using trypsin), extracted and identified. The advantages of this approach are that it is conceptually simple and relatively cheap to set up, but it is time consuming and labour-intensive and has poor reproducibility. Unusually large, small, hydrophobic and basic proteins are difficult to analyse by this method and its resolution is limited, especially for proteins of low abundance.

A modification of this is 2D-fluorescence difference gel electrophoresis, which allows relative quantification of protein expression in more than one sample. Two samples are labelled with different fluorescent dyes and run together on a 2D-gel electrophoresis system. Relative expression in the two samples is quantitated by comparing the intensities of the two dyes in each spot on the gel. The spots are then excised and the proteins identified by MS. A third dye may be used for normalisation of the fluorescence intensities.

MS methods allow the rapid, accurate identification of proteins in complex mixtures and were initially used to identify protein spots on 2D-gels. More recently, techniques have been developed for the use of MS techniques in proteomic analysis without the need for prior electrophoretic separation.

- **Single stage MS:** This method can be used to identify the mass:charge (m/z) ratio for each peptide within a sample and is suitable for proteomic analysis in species with relatively small, completely sequenced genomes.
- **Tandem mass spectrometry (MS/MS):** This method is preferable when attempting proteomic analysis of species with large genomes, although there is still a requirement for complete sequence availability. The equipment used includes triple quadrupole MS, ion-trap and time-of-flight (TOF) instruments, usually used in electrospray ionisation mode and linked to a prefractionation system (e.g. liquid chromatography (LC) or capillary electrophoresis). For example, multi-dimensional protein identification technology allows the identification of a greater range of proteins than 2D-gel electrophoresis. In this method, a complex mixture of proteins is digested (e.g. using trypsin) then subjected to multi-dimensional chromatographic separation followed by electrospray ionisation MS/MS. This method requires large-scale, expensive equipment but provides very high-quality protein identification and can be combined with stable isotope labelling to allow relative quantitation (e.g. using the isotope-coded affinity tag method).
- **Surface-enhanced laser desorption ionisation (SELDI):** This HTS chip-based MS technology, a variant of matrix-assisted laser desorption ionisation (MALDI), is used to produce proteomic fingerprints from small biological samples. The sample is applied to a chip, washed and overlaid with a MALDI matrix. Proteins are then released by laser desorption and analysed by TOF-MS. This approach does not allow the identification of individual proteins but generates a diagnostic spectrum of m/z peaks which can be used.

Array-based proteomic methods are also becoming popular for the HTS analysis of protein expression. These work on a similar principle to DNA arrays but are only available in relatively low-density formats, and the problem of maintaining correct protein folding and functionality remains an issue with these techniques. The types of arrays which are currently available include the following:

- **Protein arrays:** These can be used to address protein–protein interactions.
- **Antibody arrays:** As the name suggests, these comprise an array of antibodies and usually operate on a similar principle to enzyme-linked immunoassay methods. They usually focus on a specific group of proteins such as a particular signalling pathway or a protein family.
- **Reverse phase assays:** In this type of array, sample proteins are applied to the array in serial dilutions and probed using antibodies. The main use of reverse phase arrays is for the quantitation of expression of a particular subset of proteins.
- **ChIP-chip assays:** Chromatin immunoprecipitation (ChIP) arrays are used to identify protein binding sites across the whole genome. The ChIP method is used to determine the location of DNA-binding sites on the genome for a particular protein of interest. This technique gives a picture of nuclear protein-DNA interactions in living cells or tissues. The principle of the assay is that DNA-binding proteins (including transcription factors and histones) in living cells can be cross-linked to the DNA that they are binding. The crosslinking is often accomplished by applying formaldehyde to the cells (or tissue). Following crosslinking, the cells are lysed and the DNA is fragmented by sonication. By using an antibody that is specific to a putative DNA-binding protein, the protein-DNA complex can then be immunoprecipitated. The immunopurified protein-DNA complexes are then heated to reverse the formaldehyde crosslinking of the protein and DNA complexes, allowing the DNA to be separated from the proteins. The identity and quantity of the DNA fragments isolated can then be determined by PCR.

In order to determine where a protein binds on a genome-wide scale, a DNA microarray can be used (ChIP-on-chip or ChIP-chip) allowing for the characterisation of the cistrome, which is defined as the set of *cis*-acting targets (DNA-binding sites) of a trans-acting factor (transcription factor, pioneer factor, restriction enzyme, etc.) on a genome scale. ChIP-sequencing has also recently emerged as a new technology that can localise protein binding sites in a high throughput, cost-effective manner.

The results of ChIP-chip analysis may be extended and confirmed by direct sequencing of the putative protein binding sequences identified. However, it is still necessary to verify that the binding of a particular protein (e.g. a transcription factor) adjacent to a particular gene is actually a part of its regulatory mechanism.

'Omics techniques have been criticised on the grounds that they do not take into account the consequences of post-translational modifications such as phosphorylation, acetylation, glycosylation, nitrosation, ubiquitination, farnesylation and sumoylation. However, methods are now available to address at least some of these processes, and the discipline of proteomics is now considered to include proteome mapping, functional proteomics, protein profiling, structural proteomics and the study of post-translational modifications and protein–protein interactions

Quality assurance By analogy with MIAME, the Human Proteome Organisation has defined reporting guidelines for proteomics experiments in the Minimum Information About a Proteomics Experiment (MIAPE).[20]

Data processing One of the challenges of proteomics is the identification of the many proteins detected by MS-based approaches, which provide only molecular weight data (and in some cases the molecular weights of fragments). The spectra generated by MS-based analysis of protein mixtures are interpreted with the aid of protein databases such as SEQUEST, Mascot, PROWL, COMET and ProteinProspector.[21]

Databases The issues affecting data management in proteomics are similar to those discussed in the context of transcriptional profiling, with the additional complication that it is necessary to be able to store information about protein–protein interactions as well as that relating to individual proteins. In order to deal with these issues, the Human Proteome Organisation has developed the Proteomics Standards Initiative,[22] which defines community standards for data representation in proteomics to facilitate data comparison, exchange and verification.

The key data storage facilities in proteomics include the following:

- **UniProtKB**,[23] originally established by the Swiss Institute of Bioinformatics and the European Bioinformatics Institute with the aim 'to provide the scientific community with a comprehensive, high quality and freely accessible resource of protein sequence and functional information', and manages a protein knowledgebase consisting of two sections:
  - Swiss-Prot, which is manually annotated and reviewed.
  - TrEMBL, which is automatically annotated and is not reviewed.

  UniProtKB/TrEMBL contains high quality computationally analysed records that are enriched with automatic annotation and classification. These unreviewed entries are stored separately from the manually reviewed UniProtKB/Swiss-Prot entries so that the high- quality data of the latter is not diluted.
- **The PRIDE Proteomics IDEntifications database**[24] is a centralised, standards compliant, public data repository for proteomics data, including protein and peptide identifications, post-translational modifications and supporting spectral evidence. It provides a standards compliant public repository for MS-based proteomics data and allows access to the experimental evidence for the existence of hypothetical proteins (e.g. identified by sequence analysis) as well as information about their tissue-specific expression. PRIDE is linked to the protein annotation database InterPro,[25] a resource that provides

[20] http://www.psidev.info/node/91
[21] http://fields.scripps.edu/sequest/; http://www.matrixscience.com/integra.html; http://prowl.rockefeller.edu/; http://comet-ms.sourceforge.net/; http://prospector.ucsf.edu/prospector/mshome.htm
[22] http://www.psidev.info/
[23] http://www.uniprot.org/
[24] http://www.ebi.ac.uk/pride/

functional analysis of protein sequences by classifying them into families and predicting the presence of domains and important sites using predictive models known as signatures from the member databases of the InterPro consortium.

3.4 Metabolomics/metabonomics

Transcriptional profiling and proteomics are powerful approaches for the identification of molecular pathways involved in target organ toxicity, but they have the limitation that they examine systems at the level of RNA and protein expression rather than addressing actual metabolic outcomes. However, advances in metabolic profiling technology are opening up the possibility of identifying unique metabolic profiles associated with specific mechanisms of toxicity.

Metabolic profiling, now known as metabolomics or metabonomics, has been defined as 'the quantitative measurement of the time-related multi-parametric metabolic response of living systems to pathophysiological stimuli or genetic modification' (Nicholson *et al*., 1999). It involves the analysis of low-molecular weight compounds in biological fluids in order to obtain a complete profile of the small molecule content of the sample. The types of analytes considered are small endogenous compounds, including sugars, amino acids, organic acids and creatinine; thus, this approach provides a direct functional readout of the physiological state of the organism. The advantages of this approach include the fact that it addresses systemic effects and the non-invasive method of sample collection, which reduces distress to experimental animals, allows for repeated sampling from a single animal over time and means that human volunteer studies can be conducted using the same methodology as in experimental animals, thus facilitating interspecies extrapolation.

The terms *metabolomics* and *metabonomics* are often used interchangeably, although some researchers define the two terms differently; some investigators define metabonomics as the analysis of low-molecular weight compounds in biological fluids by means of nuclear magnetic resonance (NMR) and metabolomics as the same type of analysis using LC-MS or gas chromatography (GC)-MS while others use metabolomics for cell-based metabolite profiling and metabonomics to refer to a system with a biology-based approach. For the remainder of this chapter, the term metabolomics will be used and should be taken to include metabonomics where appropriate.

The metabolomic approach is complementary to the genomic, transcriptomic and proteomic approaches described earlier and in Chapter 2 (van Ravenzwaay *et al*., 2007). Although, like these, it is a high-content approach, it differs from the other methods in that it:

- addresses the biochemical products of metabolism rather than the genes, RNAs and proteins which lead to their generation;
- allows detection of actual changes in levels of metabolites as a consequence of transcriptional/translational changes;

[25] http://www.ebi.ac.uk/interpro/

- focuses on metabolites of known structure and function, making it easier to determine the biological significance of the changes observed;
- may actually be simpler to interpret than other 'omics methods, as the number of metabolites detected is much smaller than the total number of genes/proteins in the genome; and
- is based on biofluids which carry information about systemic, rather than tissue-specific effects, and can be sampled repeatedly in order to generate a time course in a single animal or human.

Metabolomics combines strategies for the identification and quantification of cellular metabolites using sophisticated analytical techniques with the application of statistical and multi-variant methods for information extraction and data interpretation.[26] The term metabolomics covers several approaches (Roessner and Bowne, 2009):

- **Targeted metabolomics** involves the determination and quantification of a predetermined set of known metabolites (usually <50 in number) using analytical techniques which have been optimised for the compounds of interest, and is really just an extension of conventional multiplex analysis. This technique does not require the use of high resolution, accurate mass detection.
- **Metabolic fingerprinting** is the identification of metabolic signatures which may be used, for example, in classifying compounds according to the modes of action. It is typically conducted on minimally prepared peripheral biofluids (i.e. urine or blood) and the peaks identified may or may not be assigned to molecular origin, which is often the rate-limiting step in data analysis.
- **Metabolomics/metabolite profiling** is a non-targeted approach involving the application of a range of complementary analytical methods in order to identify and quantify as many metabolites as possible. This method involves the simultaneous analysis of >50 analytes and is best conducted using high resolution, accurate mass detection.

3.4.1 MS-based metabolomics

A number of practical issues need to be taken into account when planning and conducting a metabolomics experiment, the workflow for a typical example of which is illustrated in Figure 3.5. Some of these are common to all types of toxicological experiment (e.g. selection of appropriate test and control groups), while others are specific to the method and matrix being used (normalisation of results, especially for urine, and preventing clogging of the electrospray source of the MS). Urine is the matrix of choice for metabolomic analysis since it is easily sampled, simple to analyse and provides information about polar metabolites; the use of urine as a matrix for analysis has the advantages that it is non-invasive and circulating levels of metabolites are concentrated by renal filtration. However, the

[26] For a very detailed review of the analytical methods used and methods of statistical analysis, see Mishur and Rea (2012).

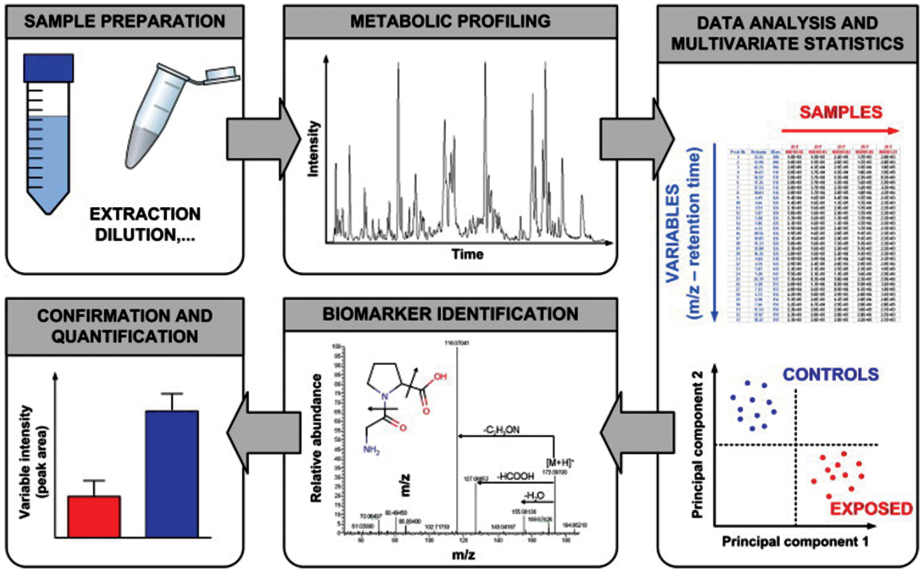


Figure 3.5 The MS-based metabolomics flow chart (source: Roux *et al.* (2011); figure 3. Reproduced with permission of the Canadian Society of Clinical Chemists)

amount of each metabolite excreted via the urine is highly variable and this has to be taken into account during analysis. Other matrices such as plasma can provide information about other classes of metabolites (e.g. lipids). Appropriate sample preparation is a key to obtaining interpretable results: urine samples can simply be diluted with water, whereas matrices with high protein content (plasma, cerebrospinal fluid and cell samples) need to be precipitated using organic solvents such as methanol, ethanol or acetonitrile, with or without sonication or mechanical agitation to break up the cells. Such approaches allow the coupling of classical histological and biochemical tools with 'omics and imaging approaches in a combined approach to the elucidation of mechanisms.[27]

Data evaluation and interpretation in metabolomics, as in other 'omics approaches, generally requires the application of multi-variate statistical analysis. In order to generate a quantitative metabolite table, it is necessary to annotate the peaks and then quantitate each analyte. In the case of NMR, the signals obtained are directly proportional to the concentration of the corresponding entity in the analytical matrix, so that quantitation is relatively straightforward, although overlapping peaks can create problems. The quantitation of MS data is more problematic due to the generation of multiple product ions during molecular fragmentation and the fact that only relative quantitation (against multiple concentrations of authentic standards) is possible.

[27] For more on the applications of LC-MS/MS-based metabolomics in clinical chemistry and toxicology, see Roux *et. al*. (2011) and Robertson, Watkins and Reily (2011).

The metabolomic approach has a number of advantages over genomic, transcriptomic and proteomic approaches:

- It allows the detection of actual changes in levels of metabolites as a consequence of transcriptional/translational changes, that is, it measures changes in metabolism directly rather than indirectly.
- It addresses small molecule metabolites whose structure can be defined and whose biochemical function is, in many cases, known. This should, in theory at least, make it easier to predict the consequences of the changes observed.
- It provides high level, integrated information.
- It can be performed using blood or urine samples. These are more representative of processes in the system as a whole than are the single tissue samples used in transcriptomic or proteomic analysis.
- The procedure for obtaining samples is relatively non-invasive and the sensitivity of the techniques used means that the samples used can be very small. This means that it is possible to take serial samples from a single individual (human or animal; even mouse) and study the time course of a response without the need for sample pooling and creation of composite time courses.

The results obtained in metabolomic studies are often simpler than those obtained in, for example, proteomic studies because the number of small molecule metabolites present in biological fluids is smaller than the total number of proteins expressed. This means that metabolomics has a better chance of identifying meaningful changes because they are expressed against a lower level of background noise than in proteomics.

Metabolomics is, however, not without its problems because no single analytical technique is able to characterise the entire metabolome, data interpretation and information extraction remain challenging and the data obtained in metabolomic analysis do not always correspond to preconceptions based upon conventional biological/biochemical experimentation.

In an effort to deal with these issues, the Metabolomics Society has appointed an Oversight Committee as part of the Metabolomics Standards Initiative (MSI)[28] to monitor, coordinate and review the efforts of working groups in specialist areas that will examine standardisation and make recommendations with a view to specifying the minimal guidelines on Core Information for Metabolomics Reporting, analogous to the MIAME and MIAPE guidelines.

Metabolic profiling of blood and urine has potential applications in both medicine and toxicology. In the medical setting, it has the potential to identify metabolic signatures for use in diagnosis, determination of prognosis and monitoring of treatment as well as in mechanistic research. In the context of animal experimentation it can be seen as a 3Rs approach since it can be used to reduce the number of animals used by taking serial samples and analysing many metabolites simultaneously. It also contributes to systems biology by providing data which can be used in the development of computational models to predict the behaviour of a biological system following perturbation by, for example, a toxic insult.

[28] http://msi-workgroups.sourceforge.net/

Endogenous metabolic profiling is problematic because important biomarkers tend to be small, polar molecules which are difficult to analyse by conventional methods and are sensitive to temporal and inter-individual variation. This necessitates the acquisition of reproducible high-quality data using, for example, integrated LC-MS methodology.

Example: Metabolomic changes in rat liver due to carbon tetrachloride

The potential value of the metabolomic approach to target organ toxicity is illustrated by the work of Lin *et al*. (2009), who used the model compound carbon tetrachloride (which induces hepatic steatosis, necrosis, fibrosis, and ultimately hepatocarcinogenesis) to characterise metabolomic changes in rat liver. Wistar rats were treated with carbon tetrachloride (1 ml kg^{-1}) and urine collected for 7 days. Samples were also taken for blood biochemistry and histology. The response to carbon tetrachloride was consistent with that observed previously: there was a marked increase in liver enzyme release which peaked on day 3 and returned to normal by day 6, and this was associated with acute hepatocellular damage (formation of lipid droplets, vacuolation and fibrosis).

Metabolomic analysis (based on LC-MS) was used to identify endogenous metabolites which were considered to be biomarkers of toxicity. Principal component analysis was used to identify 23 compounds associated with the metabolic response following carbon tetrachloride treatment. Biomarker elucidation with the aim of detecting variations in chromatographic fingerprints and identifying the compounds responsible allowed 11 biomarkers to be identified unambiguously. These were citric acid, hippuric acid, taurine, L-isoleucine, creatinine, L-threonine, L-serine, β-alanine, L-histidine, L-arginine and L-lysine. Some of these (citric acid, hippuric acid, creatinine and taurine) are considered to be generic markers of toxicity while others may be identified as specific biomarkers with further mechanistic understanding.

Example: Use of metabolomic data for biological grouping of compounds

A statistical pairwise comparison of metabolomic changes caused by data-rich compounds in the MetaMap®Tox database has been undertaken with the aim of evaluating their biological and toxicological relevance (van Ravenzwaay *et al*., 2012). Groups of at least three reference compounds with a common mode of action were selected from the MetaMap®Tox database.[29] Common effects on metabolites were identified (by using GC-MS and LC-MS/MS approaches the investigators were able to detect and quantify 238 metabolites with molecular weights between 80 and 1500 Da) and evaluated on the time and dose response, and then at least one additional chemical which matched this metabolite pattern was identified. The idea was that the compounds thus identified would have the same mode of action as the reference compounds; compounds which do not share this mode of action should not exhibit the same metabolic profile.

The results distinguished between compounds which are structurally similar but toxicologically dissimilar (e.g. 2-AAF and 4-acetylaminofluorene (4-AAF)), and

[29] http://www.metanomics-health.de/metamaptox-studies-1107.html

were interpreted as providing a biological underpinning for chemical groupings. This may be of value in the context of REACH, where read-across is a central plank of the evaluation of novel substances, and the investigators suggest the development of a new concept, the Quantitative Biological Activity Relationship, to underpin this process.

3.4.2 NMR-based metabolomics

The Consortium for Metabonomic Toxicology (COMET),[30] led by Imperial College and involving six major pharmaceutical companies, was established in order to investigate the utility of NMR-based metabolomic approaches in the toxicological assessment of drug candidates. The COMET project started in 2001 and its first phase ran for 3 years (Lindon *et al.*, 2003). Participants identified the following strengths of the metabolomic approach:

- Analysis is based on the use of biofluids (e.g. urine) therefore the approach is non-invasive.
- Early detection of metabolomic changes permits the identification of biomarkers which precede overt histopathological changes.
- Metabolomic analysis gives a global overview of systemic biochemical changes, providing additional mechanistic insights.

The project applied metabolomic techniques to generate a large database of model toxins. In the COMET1 project, 147 toxin treatments were evaluated at three doses in Sprague-Dawley rats and B6C3F1 mice, with termination 48 h and 7 days after dosing. Blood samples were collected at $t = 24$ h, 48 h and 7 days while urine was collected over 8 days, including a 24-h baseline collection. The focus was on hepatotoxins and nephrotoxins, leading to the creation of a metabolomic database of ^{1}H-NMR spectra of rodent urine and serum (Lindon *et al.*, 2005, Ebbels *et al.*, 2007). The choice of ^{1}H-NMR as the analytical method was based on the belief that it is best suited to the study of toxic events because it can be used to carry out simultaneous multi-component analyses on biological materials without bias caused by expectations regarding the types of changes to be observed. This method was also shown to have good inter-laboratory consistency; the differences between laboratories were certainly smaller than the toxin-induced differences in biochemistry.

The COMET2 project extended this to a consideration of mechanisms of toxicity, combining NMR and LC-MS-based metabolomic analysis with transcriptomic analysis. The system was used to characterise metabolomic patterns involved in galactosamine-induced hepatotoxicity and bromoethanamine-induced renal papillary necrosis (Coen, 2010).

Data from the COMET project were used to develop an expert system for the prediction of toxic effects based on chemometric analysis of the observed ^{1}H-NMR changes, adopting a neural networking approach. Once trained, the expert system was able to predict the potential effects of novel compounds rapidly and automatically; however, it was difficult to ascertain from the networking

[30] http://www1.imperial.ac.uk/surgeryandcancer/divisionofsurgery/biomol_med/key_projects/comet/

model which of the original sample descriptors actually informed the classification (Lindon *et al*., 2003).

Data from the COMET project study on hydrazine have also been used to develop a mathematical model to describe longitudinal metabolomic time course data (Berk *et al*., 2011). This is a particularly challenging issue because of the need to represent time courses from hours to months and the possibility of disease progression and physiological changes overlaying toxic processes. By identifying peaks associated with metabolites which had previously been shown to be differentially regulated by hydrazine, it was possible to construct a mathematical model of the time course of its effects. Interestingly, the response in the majority of animals followed the modelled curve closely, but a small number responded differently to the stimulus and by modelling the responses of individual animals separately it was possible to classify them as strong or weak and fast or slow responders.

3.5 Integrating different types of 'omics data

'Omics data addresses the genetic regulation and behaviour of cells at a number of different levels which interact through a variety of feedback mechanisms. The four main levels usually considered under the heading of 'omics are the genome, the transcriptome, the proteome and the metabolome. These are in turn modulated via dynamic mechanisms involving DNA methylation (the epigenome), RNA stability (the microRNAome), protein phosphorylation (the phosphoproteome) and metabolic flux (the fluxome) (Figure 3.6). The technology for examining the four main 'omic levels is now well established, but we are still in the early stages of trying to decipher how they interact via the epigenome, the microRNAome, the phosphoproteome and the fluxome.

One way of handling the complexity of 'omics information is to adopt a clustering approach. This involves identifying groups of genes which are likely to be functionally related or influenced by one or more common upstream factors. As well as simplifying the analysis of genes of known function, this approach can be used to assign functions to previously uncharacterised genes. However, it should be borne in mind that clustering methods only indicate that the genes in a cluster are co expressed. This does not necessarily imply that they are coregulated.

These methods cannot completely overcome the problem of how to integrate diverse data sets such as transcriptional profiles, transcription factor binding information and sequence data into a common modelling framework. The temptation is to create the simplest possible model, but it is essential to recognise that each interaction measured is, itself, generated by multiple biochemical processes. Sophisticated mathematical approaches such as the construction of so-called "infinite Gaussian mixture models" using non-parametric Bayesian analysis allow this issue to be addressed. In this approach each gene expression profile is assumed to represent a multi-dimensional vector of measurements and each cluster is assumed to be described by a multivariate Gaussian probability distribution.[31]

[31] The mathematics behind this approach are beyond the scope of this book but further information may be found in Cooke, Savage, and Wild (2009).

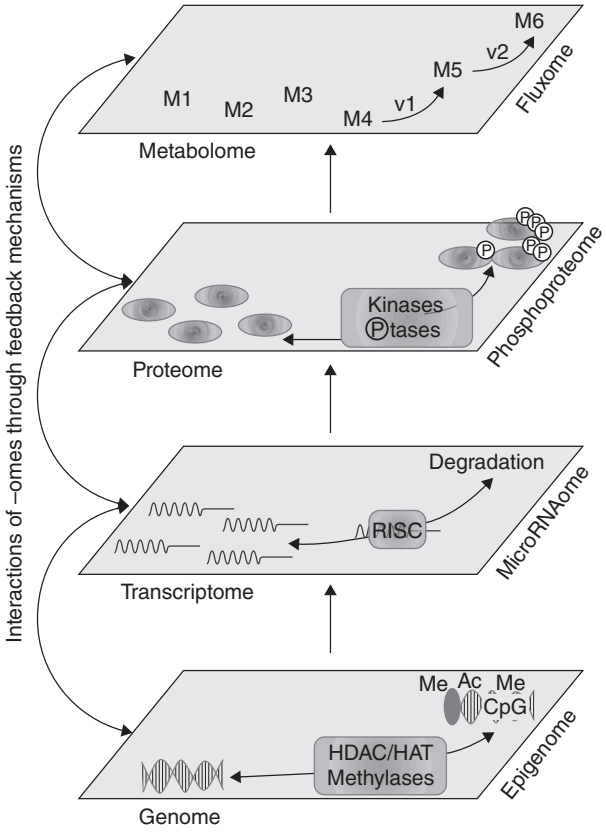


Figure 3.6 The 'omic levels and their interactions. The 'omes can be separated into biological levels, comprising the genome, transcriptome, proteome and metabolome. The action of these core omes can be influenced by modifier omes that regulate a single biological level; such modifier omes include the epigenome, miRNAome, phosphoproteome and fluxome. Finally, interactions may occur vertically between each of the biological levels, with feed-back and feed-forward systems co-ordinating both the core- and modifier-omic outputs from each biological levels (source: Plant (2008). Reproduced with permission of Elsevier)

3.5.1 'Omics in drug discovery

'Advances in molecular profiling approaches, such as transcriptomics, proteomics and metabolomics, offer the potential to provide a more comprehensive insight into toxicological effects than hitherto possible' (Gallagher *et al.*, 2009). 'Omics technologies are now pervasive within all aspects of the drug development cycle and particularly in the preclinical phase, although conventional methods such as histopathology assessment still occupy a central position in safety assessment.

The 'omics technologies have two principal applications in drug development:

- Mechanistic studies to characterise mechanisms of action.
- Predictive studies seeking markers for the prediction of whether a certain compound will be toxic to humans.

Before relying on 'omics approaches to address these issues, however, it is essential to ensure that the results obtained are of at least the same quality as those generated by conventional approaches and that the novel approaches confer pragmatic benefits in terms of reduced cost and/or lower error rates. Currently, the best chances of achieving these requirements is in the early prediction of long term adverse effects such as carcinogenicity and reproductive toxicity, where the conventional approaches require an entire (rodent) lifetime.

In addition, the issue of QA must always be borne in mind. Transcriptional profiling is particularly prone to the generation of artefacts, and in recent years considerable effort has been devoted to the establishment of reliable validated methods for RNA extraction, microarray analysis and data management. In particular, the MicroArray Quality Control (MAQC) project[32] has helped to improve microarray and next-generation sequencing technologies and foster their proper applications in discovery, development and review of FDA regulated products (Fan *et al*., 2010).

Commercial offerings in the use of gene arrays for prediction of toxicity are supported by very little peer-reviewed work specific to those services. However it is clear that such commercial technologies may find use in certain circumstances where there is a need for comparison of a range of molecules in respect of a potential biomarker for a specific toxic endpoint. Such markers may not need to be 100% predictive in order to be useful in drug development. An example of such a service is ToxExpress[33] from Gene Logic, described on the company's web site as 'a flexible and enabling program that brings the power of toxicity-based gene expression coupled with classical toxicology endpoints to lead optimisation and drug safety studies'.

3.5.2 'Omics profiles as biomarkers of toxicity

Biomarkers are measurable and quantifiable biological parameters which serve as indices for health and physiology-related outcomes. In 2001, the NIH defined a biomarker as 'a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes or pharmacologic responses to a therapeutic intervention' (BDWG, 2001), while the NIEHS definition[34] focuses more on their mechanistic role, stating that 'Biomarkers are key molecular or cellular events that link a specific environmental exposure to a health outcome'. Biomarkers are related to, but distinct from, risk factors and surrogate endpoints, although in some cases a biomarker can also act as a surrogate endpoint.

Transcriptional profiling is a powerful approach for the identification of molecular pathways involved in target organ toxicity, but it has the limitation that it examines systems at the level of RNA expression rather than addressing actual biochemical outcomes. However, advances in metabolomic technology have made it possible to interrogate biological samples with a view to identifying unique metabolic fingerprints which provide 'a systems level window into cellular

[32] http://www.FDA; for additional information, see Currie (2012).
[33] http://www.genelogic.com/knowledge-suites/toxexpress-program
[34] http://www.niehs.nih.gov/health/topics/science/biomarker

homeostasis as the concentrations and fluxes of key cellular metabolites alter in response to interventions and external stressors such as administration of a toxin or onset of a disease' (Coen, 2010). This approach involves the use of NMR spectroscopy and MS to generate non-targeted, information-rich metabolic profiles followed by the use of multi-variate statistical methods to extract biomarker information from the resulting complex data sets. It has already proved to be powerful in the assessment of temporal metabolic fluctuation during the onset, progression and resolution of a toxic response. One advantage of this approach is that it simultaneously provides information on drug or toxin metabolite profiles and endogenous metabolic processes which occur in response to an insult.

The state of the art in toxicity biomarkers is evolving but while companies are attempting to relate target organ effects to changes in gene expression profiles, particularly in liver, a set of genes which can be considered to be predictive of toxicity has yet to be identified. Rather, expression profiling is being used on a case by case basis to provide a mechanistic understanding of the target and implications of its agonism or antagonism. Such information can be identified and used to define a customised set of parameters which can be used to evaluate future drug safety studies.

However, problems still remain, particularly with respect to the possibility of applying 'omics biomarkers in a clinical context. There are issues of technology standardisation (it has been said (Molitoris *et al*., 2008) that proteomics techniques such as SELDI just reveal "unidentified flying particles"). Data reproducibility continues to be an issue and it can be difficult to discriminate between experimental error, genuine inter-individual variability and uncertainty in disease processes.

The number of genes in a transcriptomic fingerprint may affect its predictive value: a fingerprint with too many genes (e.g. 5000) may lack discriminatory power due to the presence of background noise and chemical-specific effects. The optimal number of genes for a diagnostic fingerprint is thought to be about 100 (see Uehara *et al*. (2008)). Some investigators believe that a fingerprint of about 10 genes is sufficient to predict effects; however, this is arguably not an 'omics strategy at all since such a small number of genes may readily be addressed using conventional molecular methods such as RT-PCR.

It is not likely that genomics or any of the other pattern-based technologies will provide usable alternatives to conventional animal studies in the short-term but as part of an integrated strategy they are ultimately certain to play a part. In a 2006 review of the state of the art (Maggioli *et al*., 2006) it was concluded that some of the challenges to be faced by this technology before it could fulfil the promise of replacing toxicological studies were

- building databases of gene-expression profiles for known toxins;
- ensuring comparability of data generated using different technologies;
- improving prediction of toxicity across species;
- improving prediction of toxicity from *in vitro* to humans.

A second review (Collings and Vaidya, 2008) details some of the significant developments of subsequent years paying particular attention to the quantitative aspects and the integration of data from studies of genomics, proteomics

and metabolomics. These techniques have been used to predict carcinogenic potential *in vivo* on the basis of genomics analysis of short-term rat studies (Ellinger-Ziegelbauer *et al*., 2008, 2009). Even in this well-studied area, the predictive power was less than 90%, demonstrating a need for further development. The lack of current data on the transfer of this technology to *in vitro* responses is one of the primary weaknesses in its use as a read-across for animals in risk assessment, but it is likely that as this methodology establishes reliably predictive biomarkers for specific toxicities then its application to *in vitro* studies may become possible.

3.6 Remaining issues with 'omics approaches

When molecular biologists and toxicologists first got together and started to apply 'omics technologies to questions in toxicology, studies were largely undertaken by individual groups in academia, and this led to number of issues in relation to inconsistencies in experimental design, lack of standardisation of protocols, a multiplicity of normalisation and analysis procedures, inadequate QC measures and standards and consequent problems in terms of the comparability and reproducibility of assays. These, together with a lack of effective data sharing and the absence of formal reporting standards, meant that after an initial burst of enthusiasm leading to a rapid increase in publication rate from 2000 to 2006, there was a slowing down in the later 2000s. However, in recent years the development of high-quality microarrays, standardised protocols, accurate scanning technologies and robust computational methods have helped these technologies to mature into useful, and easy to use, investigational tools.

It is, however, still the case that 'omics approaches can be subject to a very high false positive and false negative rate, meaning that while they may have reasonable accuracy they can lack precision. Reliable quantitation is required to allow quantitative dose–response relationships, necessary for read-across, to be derived, although collaborative efforts to collect and collate control data and identify sources of variation in baseline gene expression levels have helped to define what is normal in profiling terms (Boedigheimer *et al*., 2008). In addition, the use of different bioinformatic approaches and the statistical analysis applied can have as much, if not more, effect on the outcome of an 'omics study as the platform used, an issue which is often not addressed explicitly in academic publications. Data management and storage is also an issue because all the 'omics technologies generate large quantities of data which require the highest standards of storage and reporting. Based on current technology, DNA microarrays have much greater information capacity than proteomic methods, although this is being addressed by the development of HTS MS-based methods.

Various other questions also arise when trying to interpret 'omics results in terms of mechanistic toxicology. In terms of the available methodologies, it has been stated that 'the technology for transcriptome analysis is highly advanced, allowing the complete transcriptome to be analysed in a single experiment; in comparison, the technological difficulties associated with the analysis of phosphoproteins means that coverage of a phosphoproteome is much more sparse' (Plant, 2008). The current predominance of transcriptomics studies is mainly due to the

current state of development of the technology rather than any consensus that this is the best level to address effects. There is not yet a universal agreement on the meaning and validity of 'omic markers of toxicity, especially in relation to transcriptional profiling, since changes in mRNA expression generally precede the corresponding changes in protein expression and the level of mRNA expressed does not always agree with that of the corresponding protein and consequent biological effects. Indeed, it is self-evident that the effects of toxic compounds are not a direct consequence of changes in mRNA expression but are mediated via the protein products of these mRNAs. Thus, the measurement of mRNA changes is a surrogate endpoint. It is usually an appropriate one, but there is a considerable amount of evidence that changes at the transcript level do not necessarily reflect protein expression or activity. This may be the case for a number of reasons, some of which may be resolved experimentally while others are intrinsic to the relationship between mRNA and protein expression.

Furthermore, the need to extract mRNA of protein from tissues for transcriptomic/proteomic analysis means that structural information is lost, since all cells within the treated tissue (both target and non-target cells) are represented in the samples analysed. This may lead to a dilution effect so that changes taking place in a small target population are drowned out by negative data from a larger, non-responsive population (Merrick and Madenspacher, 2005).

If these issues are borne in mind, a productive relationship may be developed between transcriptomic and proteomic approaches. For example, preliminary transcriptomic studies can be used to guide and inform the design of toxicoproteomic studies. The further challenge lies in conducting the appropriate hypothesis-driven follow-up studies which will bring biological meaning to the data contained in lists of altered transcripts, proteins and metabolites identified in transcriptomic, proteomic and metabolomic studies.

MS-based proteomic techniques are also limited by the fact that MS is intrinsically a non-quantitative method (Maier *et al*., 2009). This can, to some extent, be overcome by stable isotope labelling and label-free quantitation methods but these methods still tend to provide only relative rather than absolute quantitation results.

When trying to interpret transcriptomics data in terms of predicted protein levels, or to correlate transcriptomics and proteomic data, it is important to remember that relationships between protein and mRNA abundances are not normally distributed. This is a consequence of a variety of factors including post-transcriptional effects, post-translational phenomena (e.g. protein half life), noise and experimental error. As a general rule, therefore, the use of a ranking method such as the Spearman rank correlation coefficient is preferable to attempts at absolute quantitation.

3.7 Conclusions

It has been concluded that a concerted effort is required on the part of regulatory, academic and industry participants in order to 'move 'omic analysis from an expensive "toy" that gives mechanistic hints to a tool that provides regulatory

standard data' (Plant, 2008). Once this has been achieved it should become possible to meet the challenge of determining which of the molecular events occurring following exposure to low doses of chemicals lead to pathological outcomes versus those which may be adaptive, beneficial or irrelevant (Aardema and MacGregor, 2002). This remains challenging because the aim of regulatory testing is to classify compounds as either toxic or non-toxic, whereas the reality is that the effects of most compounds are dependent on dose and time (i.e. they occur at a particular dose and after a certain period of time). Initial hopes that microarray analysis could be used to define gene expression signatures which are diagnostic for particular types of toxicity have not been fulfilled; however, with careful analysis, 'omics data can already be used to help generate or support a weight-of-evidence case for a particular mode of action or human risk assessment (Kienhuis *et al*., 2011, Currie, 2012). As they mature they will no doubt contribute to the elucidation of mechanisms of action and help us to achieve a detailed understanding of the effects of xenobiotics on living organisms.

Self-assessment questions

- What are the relative merits of transcriptomics, proteomics and metabolomics in the context of toxicology?
- Will the application of 'omics approaches make it possible to define molecular fingerprints for specific mechanisms of toxicity?
- To what extent do toxicoproteomic methods go beyond the detection of 'unidentified flying particles'?

Background reading

Dale, J.W. and von Schantz, M. (2007) From genes to genomes: Concepts and applications of DNA technology, John Wiley and Sons, Ltd, Chichester, UK. ISBN: 978-0-470-01734-0

References

Aardema, M.J. and MacGregor, J.T. (2002) Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. *Mutation Research*, **499**, 13–25.

Barrier, M. and Mirkes, P.E. (2005) Proteomics in developmental toxicology. *Reproductive Toxicology*, **19**, 291–304.

BDWG (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, **69**, 89–95.

Berk, M., Ebbels, T. and Montana, G. (2011) A statistical framework for biomarker discovery in metabolomic time course data. *Bioinformatics*, **27**, 1979–1985.

Boedigheimer, M.J., Wolfinger, R.D., Bass, M.B. *et al*. (2008) Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genomics*, **9**, 285.

Boutros, P.C., Yan, R., Moffat, I.D. *et al.* (2008) Transcriptomic responses to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) in liver: comparison of rat and mouse. *BMC Genomics*, **9**, 419.

Boutros, P.C., Bielefeld, K.A., Pohjanvirta, R. and Harper, P.A. (2009) Dioxin-dependent and dioxin-independent gene batteries: comparison of liver and kidney in AHR-null mice. *Toxicological Sciences*, **112**, 245–256.

Boverhof, D.R., Burgoon, L.D., Tashiro, C. *et al.* (2006) Comparative toxicogenomic analysis of the hepatotoxic effects of TCDD in Sprague Dawley rats and C57BL/6 mice. *Toxicological Sciences*, **94**, 398–416.

Chen, M., Zhang, M., Borlak, J. and Tong, W. (2012) A decade of toxicogenomic research and its contribution to toxicological science. *Toxicological Sciences*, **130**, 217–228.

Coen, M. 2010. Metabonomic approaches for the pre-clinical study of model renal and hepatic toxins. *Proceedings of the BTS Annual Congress*, *Heriot-Watt University*, *28–31 March*, S025.

Collings, F.B. and Vaidya, V.S. (2008) Novel technologies for the discovery and quantitation of biomarkers of toxicity. *Toxicology*, **245**, 167–174.

Cooke, E.J., Savage, R.S. and Wild, D.L. (2009) Computational approaches to the integration of gene expression, ChIP-chip and sequence data in the inference of gene regulatory networks. *Seminars in Cell and Developmental Biology*, **20**, 863–868.

Currie, R.A. (2012) Toxicogenomics: the challenges and opportunities to identify biomarkers, signatures and thresholds to support mode-of-action. *Mutation Research*, **746**, 97–103.

Davis, A.P., King, B.L., Mockus, S. *et al.* (2011) The Comparative Toxicogenomics Database: update 2011. *Nucleic Acids Research*, **39**, D1067–D1072.

Davis, A.P., Wiegers, T.C., Rosenstein, M.C. and Mattingly, C.J. (2012) MEDIC: a practical disease vocabulary used at the Comparative Toxicogenomics Database. *Database (Oxford)*, **2012**, bar065.

Ebbels, T.M., Keun, H.C., Beckonert, O.P. *et al.* (2007) Prediction and classification of drug toxicity using probabilistic modeling of temporal metabolic data: the Consortium on Metabonomic Toxicology screening approach. *Journal of Proteome Research*, **6**, 4407–4422.

Elferink, M.G., Olinga, P., Draaisma, A.L. *et al.* (2008) Microarray analysis in rat liver slices correctly predicts *in vivo* hepatotoxicity. *Toxicology and Applied Pharmacology*, **229**, 300–309.

Elferink, M.G., Olinga, P., Van Leeuwen, E.M. *et al.* (2011) Gene expression analysis of precision-cut human liver slices indicates stable expression of ADME-Tox related genes. *Toxicology and Applied Pharmacology*, **253**, 57–69.

Ellinger-Ziegelbauer, H., Aubrecht, J., Kleinjans, J.C. and Ahr, H.J. (2009) Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. *Toxicology Letters*, **186**, 36–44.

Ellinger-Ziegelbauer, H., Gmuender, H., Bandenburg, A. and Ahr, H.J. (2008) Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term *in vivo* studies. *Mutation Research*, **637**, 23–39.

Fan, X., Lobenhofer, E.K., Chen, M. *et al.* (2010) Consistency of predictive signature genes and classifiers generated using different microarray platforms. *Pharmacogenomics Journal*, **10**, 247–257.

Fostel, J., Choi, D., Zwickl, C. *et al.* (2005) Chemical Effects in Biological systems--Data Dictionary (CEBS-DD): a compendium of terms for the capture and integration of biological study design description, conventional phenotypes, and 'omics data. *Toxicological Sciences*, **88**, 585–601.

Gallagher, W.M., Tweats, D. and Koenig, J. (2009) Omic profiling for drug safety assessment: current trends and public-private partnerships. *Drug Discovery Today*, **14**, 337–342.

Ganter, B., Tugendreich, S., Pearson, C.I. *et al.* (2005) Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *Journal of Biotechnology*, **119**, 219–244.

Gocht, T. and Schwarz, M. (2012) Towards the replacement of *in vivo* repeated dose systemic toxicity testing, European Commission and Cosmetics Europe. http://www.seurat-1.eu/pages/library/seurat-1-annual-report.php

Harrill, A.H., Ross, P.K., Gatti, D.M. *et al.* (2009) Population-based discovery of toxicogenomics biomarkers for hepatotoxicity using a laboratory strain diversity panel. *Toxicological Sciences*, **110**, 235–243.

Hirode, M., Omura, K., Kiyosawa, N. *et al.* (2009) Gene expression profiling in rat liver treated with various hepatotoxic compounds inducing coagulopathy. *Journal of Toxicological Sciences*, **34**, 281–293.

Hu, T., Khambatta, Z. S., Hayden, P. J., Bolmarcich, J., *et al.* 2010. Xenobiotic metabolism gene expression in the EpiDerm® *in vitro* 3D human epidermis model compared to human skin. *Toxicology In Vitro*, **24**, 1450–63.

Inadera, H., Tachibana, S., Takasaki, I. *et al.* (2008) Expression profile of liver genes in response to hepatotoxicants identified using a SAGE-based customized DNA microarray system. *Toxicology Letters*, **177**, 20–30.

Kienhuis, A.S., Bessems, J.G., Pennings, J.L. *et al.* (2011) Application of toxicogenomics in hepatic systems toxicology for risk assessment: acetaminophen as a case study. *Toxicology and Applied Pharmacology*, **250**, 96–107.

Kiyosawa, N., Ando, Y., Watanabe, K. *et al.* (2009) Scoring multiple toxicological endpoints using a toxicogenomic database. *Toxicology Letters*, **188**, 91–97.

Kiyosawa, N., Manabe, S., Sanbuissho, A. and Yamoto, T. (2010) Gene set-level network analysis using a toxicogenomics database. *Genomics*, **96**, 39–49.

Lin, Y., Si, D., Zhang, Z. and Liu, C. (2009) An integrated metabonomic method for profiling of metabolic changes in carbon tetrachloride induced rat urine. *Toxicology*, **256**, 191–200.

Lindon, J.C., Keun, H.C., Ebbels, T.M. *et al.* (2005) The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. *Pharmacogenomics*, **6**, 691–699.

Lindon, J.C., Nicholson, J.K., Holmes, E. *et al.* (2003) Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. *Toxicology and Applied Pharmacology*, **187**, 137–146.

Maggioli, J., Hoover, A. and Weng, L. (2006) Toxicogenomic analysis methods for predictive toxicology. *Journal of Pharmacological and Toxicological Methods*, **53**, 31–37.

Maier, T., Guell, M. and Serrano, L. (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Letters*, **583**, 3966–3973.

Merrick, B.A. and Madenspacher, J.H. (2005) Complementary gene and protein expression studies and integrative approaches in toxicogenomics. *Toxicology and Applied Pharmacology*, **207**, 189–194.

Mishur, R.J. and Rea, S.L. (2012) Applications of mass spectrometry to metabolomics and metabonomics: detection of biomarkers of aging and of age-related diseases. *Mass Spectrometry Reviews*, **31**, 70–95.

Moffat, I.D., Boutros, P.C., Chen, H. *et al.* (2010) Aryl hydrocarbon receptor (AHR)-regulated transcriptomic changes in rats sensitive or resistant to major dioxin toxicities. *BMC Genomics*, **11**, 263.

Molitoris, B.A., Melnikov, V.Y., Okusa, M.D. and Himmelfarb, J. (2008) Technology Insight: Biomarker development in acute kidney injury--what can we anticipate? *Nature Clinical Practice Nephrology*, **4**, 154–165.

Nicholson, J.K., Lindon, J.C. and Holmes, E. (1999) 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, **29**, 1181–1189.

Nowik, M., Lecca, M.R., Velic, A. *et al.* (2008) Genome-wide gene expression profiling reveals renal genes regulated during metabolic acidosis. *Physiological Genomics*, **32**, 322–334.

Plant, N. (2008) Can systems toxicology identify common biomarkers of non-genotoxic carcinogenesis? *Toxicology*, **254**, 164–169.

Priyadarsiny, P., Khattar, S.K., Malik, R. *et al.* (2008) Differential gene expression analysis of a known hepatotoxin, *N*-acetyl-*p*-amino-phenol (APAP) as compared to its non-toxic analog, *N*-acetyl-*m*-amino-phenol (AMAP) in mouse liver. *Journal of Toxicological Sciences*, **33**, 163–173.

Robertson, D.G., Watkins, P.B. and Reily, M.D. (2011) Metabolomics in toxicology: preclinical and clinical applications. *Toxicological Sciences*, **120** (Suppl 1), S146–S170.

Roessner, U. and Bowne, J. (2009) What is metabolomics all about? *Biotechniques*, **46**, 363–365.

Roux, A., Lison, D., Junot, C. and Heilier, J.F. (2011) Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review. *Clinical Biochemistry*, **44**, 119–135.

Schneider, M.V. and Orchard, S. (2011) Omics technologies, data and bioinformatics principles. *Methods in Molecular Biology*, **719**, 3–30.

Trevino, V., Falciani, F. and Barrera-Saldana, H.A. (2007) DNA microarrays: a powerful genomic tool for biomedical and clinical research. *Molecular Medicine*, **13**, 527–541.

Uehara, T., Hirode, M., Ono, A. *et al.* (2008) A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology*, **250**, 15–26.

Uehara, T., Ono, A., Maruyama, T. *et al.* (2010) The Japanese toxicogenomics project: application of toxicogenomics. *Molecular Nutrition & Food Research*, **54**, 218–227.

Uehara, T., Minowa, Y., Morikawa, Y. *et al.* (2011) Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database. *Toxicology and Applied Pharmacology*, **255**, 297–306.

Van Ravenzwaay, B., Cunha, G.C., Leibold, E. *et al.* (2007) The use of metabolomics for the discovery of new biomarkers of effect. *Toxicology Letters*, **172**, 21–28.

Van Ravenzwaay, B., Herold, M., Kamp, H. *et al.* (2012) Metabolomics: a tool for early detection of toxicological effects and an opportunity for biology based grouping of chemicals-from QSAR to QBAR. *Mutation Research*, **746**, 144–150.

Waters, M., Stasiewicz, S., Merrick, B.A. *et al.* (2008) CEBS–Chemical Effects in Biological Systems: a public data repository integrating study design and toxicity data with microarray and proteomics data. *Nucleic Acids Research*, **36**, D892–D900.

Wetmore, B.A. and Merrick, B.A. (2004) Toxicoproteomics: proteomics applied to toxicology and pathology. *Toxicologic Pathology*, **32**, 619–642.

Yun, J.W., Kim, C.W., Bae, I.H. *et al.* (2009) Determination of the key innate genes related to individual variation in carbon tetrachloride-induced hepatotoxicity using a pre-biopsy procedure. *Toxicology and Applied Pharmacology*, **239**, 55–63.

4

In Vitro Methods for Predicting *In Vivo* Toxicity

4.1 *In vitro* toxicology

Toxicology has traditionally been studied using animal models; however, in recent years there has been a strong impetus to develop and validate non-animal (*in vitro* and *in silico*) models. The drivers for this have included the following:

- Public concern about the use of animals in experimentation.
- The implementation of the 7th Amendment, which dictates that *in vivo* methods can no longer be used to characterise the disposition properties of cosmetic ingredients.
- The enactment in 2007 of the REACH regulations, which has created the need for a huge quantity of toxicity and disposition data which would be prohibitively costly to generate entirely using *in vivo* models (REACH also includes a requirement to minimise the use of experimental animals).
- The need for HTS methods for the pharmaceutical industry to use in drug discovery and development.

These have promoted an unprecedented expansion of *in vitro* testing. It has proved reasonably easy to test for acute toxicity *in vitro*, but the development of *in vitro* methods for repeated dose toxicity testing has proved much more challenging and strenuous. Method development and validation efforts are still in progress.

4.2 Tissue culture

Many of the techniques of *in vitro* toxicology depend upon the ability of mammalian cells to grow in culture. The discipline of tissue culture originated with

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

the discovery that if small pieces of tissue were placed in nutrient medium, cells would start to migrate out of these explants and could remain viable for days to weeks. Subsequently, methods for disaggregating tissues using either physical or enzymatic methods and plating out the resulting cell suspensions made it possible to prepare homogeneous cultures. Both these types of culture are referred to as primary cultures. A further significant development came when the first human tumour-derived cell line, HeLa, was created in 1952.

The term tissue culture covers a variety of types of culture:

- **Organ culture** refers to a three-dimensional (3D) culture of intact tissue which is grown at a gas–liquid interface and retains some or all of the histological features the tissue had *in vivo*. A histiotypic culture is similar except that cells of a single type have been dispersed and subsequently re-aggregated to create a three-dimensional structure with tissue-like cell density, while an organotypic culture is created using the same procedure but recombining cells of different lineages to form a so-called tissue equivalent.
- **Primary explant culture** describes the situation in which cells migrate out of a fragment of tissue placed at a solid–liquid interface. The tissue sample is disaggregated by chopping finely with a scalpel or by enzymatic digestion. The resulting fragments are placed in a Petri dish with a thin layer of cell culture medium, after which cells gradually migrate out from the tissue fragments. When enough cells have migrated, the remaining fragments can be washed away and the cells subcultured carefully into a flask.
- A **primary culture** is one that is derived from intact or dissociated tissues or from organ fragments. Such a culture is considered to be a primary culture until the first time it is passaged or subcultured (i.e. transferred into a new culture vessel in order to propagate it as a cell line or set up replicate cultures). At this point it becomes a cell line, but not yet an established line. Primary cell culture models are biochemically and physiologically similar to the corresponding cells *in vivo*. They have increased longevity compared with perfused organs, precision-cut tissue slices and isolated cells and time in culture allows them to recover from the trauma of preparation prior to use. Differentiated functions are often maintained and primary cultured cells can be used for acute and subchronic toxicity studies.
- A **cell line** comes into existence the first time a culture is passaged. Cell lines have both advantages and disadvantages compared with primary cultures. The advantages include the scope to grow a large (in theory an infinite) number of cells, the homogeneity of the cell population and the fact that long-term culture allows thorough characterization of the cell population. The disadvantages, which are discussed further elsewhere in this chapter, include loss of differentiated properties and genetic instability; in addition, the reagents required are expensive and the maintenance of cell lines is a labour-intensive task requiring expertise in sterile technique.

The attributes of 2D-and 3D-tissue culture, as compared with *in vivo* methods, are summarised in Table 4.1. Tissue culture is also, arguably, economical compared with the cost of setting up a large animal study (which is becoming increasingly expensive, especially as welfare requirements become more stringent), although

Table 4.1 Relative merits of *in vivo* and cell culture models for toxicity testing

2D-culture	3D-culture	*In vivo*
Easy to manipulate	Moderately easy to manipulate	Limited scope for manipulation
Does not reflect the *in vivo* situation	Partially reflects the *in vivo* situation	By definition, fully reflective of the *in vivo* situation
In vitro studies using human cells may be more relevant than *in vivo* studies in animals	*In vitro* studies using human cells may be more relevant than *in vivo* studies in animals	Does not necessarily reflect human responses
Uniform cell population, so cellular heterogeneity is not an issue, but conversely true diversity of cells within any given tissue is not taken into account	Extent of heterogeneity within the spheroid is difficult to evaluate	Takes into account the diversity of cell populations, but this makes analysis very complex
Cells of the target organ/tissue can be studied under tightly controlled conditions	Less tightly controlled conditions: it is difficult to know what is happening in the centre of a spheroid	Conditions within the living organism cannot be defined/controlled
Uniform physical and chemical environment	Culture environment is uniform, but conditions within the spheroid are ill-defined	Variable physical and chemical environment
Nutritional and hormonal status, oxygen supply and exposure to test compound are tightly controlled	Culture conditions can be defined, but conditions within the spheroid are difficult to determine	Difficult to control nutritional and hormonal status of target cells, oxygen supply and exposure to test compound
Fewer variables to consider (e.g. no input from blood flow, hormonal effects, nerve inputs), but does not take all relevant variables into account	More variables require consideration, but not as many as *in vivo*	Many variables to consider (e.g. input from blood flow, hormonal effects, nerve inputs), but it is in theory possible to take all relevant variables into account
Assay duration depends on the lifespan of cells in culture	3D-culture may enhance viability and lifespan of cultured cells	Duration depends on the lifespan of the species being used
Test compounds have direct access to target cells	Test compounds have access to the outside of the spheroid but the extent of access to the interior is difficult to verify	Impossible to be sure that compound is reaching the target tissue

Toxic metabolites pass into the medium and are diluted into the large volume of medium compared with that of the cell monolayer	Possible accumulation of toxic metabolites leading to central necrosis within the spheroid	Toxic metabolites may be removed by the normal processes of disposition
Potentially compromised cell function and behaviour, including cell–cell interactions, signal transduction and appropriate patterns of gene expression	Improved cell function and behaviour, including cell–cell interactions, signal transduction and appropriate patterns of gene expression	Fully reflects cell function and behaviour, including cell–cell interactions, signal transduction and appropriate patterns of gene expression
Individual cells can be examined morphologically	Spheroids can be examined individually, but not individual cells	Not possible to examine target cells prior to termination (except for blood cells and sperm)
Organ-specific functions cannot be replicated using individual cell lines	Better reproduction of and maintenance of organ-specific functions	Organ-specific functions remain intact
Loss of structural information as a result of propagation on a two-dimensional surface rather than in three dimensions	Improved structural information as a result of propagation in three dimensions	Structure is intact, but inaccessible during the course of the test
May lack systemic factors required for homeostatic regulation *in vivo*	May lack systemic factors required for homeostatic regulation *in vivo*	Includes all systemic factors required for homeostatic regulation *in vivo*
Small quantities of test materials can be used (particularly important in the early stages of preclinical development of pharmaceuticals)	Small quantities of test materials can be used (particularly important in the early stages of preclinical development of pharmaceuticals)	Relatively large quantities of test materials are required
Difficulty in setting appropriate concentrations of test items for use in cell culture studies	Difficulty in setting appropriate concentrations of test items for use in cell culture studies	Dose setting can be a problem *in vivo* too
Limited scope of cell culture systems to account for kinetics and biotransformation	Some evidence that disposition processes may be maintained better in 3D-culture	All disposition processes are intact
Derivation of NOAELs from *in vitro* test systems is difficult	Derivation of NOAELs from *in vitro* test systems is difficult	Derivation of NOAELs from *in vitro* test systems is a routine procedure (but dependent on appropriate dose setting)
Most commercial *in vitro* toxicity endpoints are still geared to monolayer cultures in multi-well plates	Most commercial *in vitro* toxicity endpoints are still geared to monolayer cultures in multi-well plates and may not be easy to adapt to 3D-cultures	Guidelines are available for the assessment of toxicity endpoints which are relevant *in vivo*

the amount of equipment and reagents needed for cell culture mean that initial set-up costs can be high; indeed, the cost of preparation of cells grown in culture is estimated to be approximately ten times that of the equivalent wet weight of fresh animal tissue. In addition, small quantities of test materials can be used, and this is particularly important in the early stages of pre-clinical development of pharmaceuticals.

However, the use of cultured cells has a number of drawbacks which should always be taken into account when deciding on the approach to be used to address a particular problem. Scientific issues are listed in Table 4.1; in addition, human factors which should be taken into account include the need for technical expertise, which means that cell culture cannot be undertaken on an *ad hoc* basis, as well as the fact that only small quantities of cells can be grown using conventional culture methods. The maximum batch size for a small laboratory staffed by 2–3 people is about 1–10 g wet weight of cells and any amount above about 100 g wet weight would constitute an industrial pilot-scale exercise.

Cultured cells may be anchorage-dependent (requiring a surface to attach to in order to be able to survive, multiply and/or express differentiated functions) or anchorage independent (able to grow as a suspension in liquid or semi-solid medium). Anchorage-dependent cells usually grow as monolayers; important requirements for the growth of monolayers include a suitable substrate for attachment and the existence of cell–cell contacts (especially for epithelial cells). For some cell lines the plating density is also critical. In order to generate a culture which reflects, as completely as possible, the behaviour of the corresponding cell type *in vivo* it is necessary to consider a variety of characteristics of the environment in which the cells are cultured, including the following:

- **The nature of the underlying substrate:** Most cell lines can produce their own extracellular matrix, but it may be necessary to provide an exogenous source of extracellular matrix to allow the propagation of some specialised cell types and for primary culture. This helps to promote cell adhesion and enhance cell spreading, proliferation and coverage of the underlying substrate. If exogenous extracellular matrix is required it can be prepared by mixing the appropriate purified components (e.g. basic polymers such as poly-L-lysine or purified extracellular matrix components such as fibronectin, laminin and collagen Types I and IV) or using cultured cells to generate it and then washing off the producer cells. Alternatively, a commercial matrix such as Matrigel® can be used. Matrigel® is a solubilised basement membrane preparation extracted from the Englebreth-Holm-Swarm tumour line and is provided as a sterile liquid which can be used to form a thin surface layer over an underlying substrate or as a 3D-gel.[1]
- **The degree of cell–cell contacts:** Contact-mediated (cell–cell and cell–matrix) and diffusible (soluble nutritional/hormonal) signals play a key role in determining how well cell plate out and survive in culture. Under normal *in vitro* conditions, where only one cell type is present, only autocrine and homocrine interactions are possible, but the use of conditioned medium and/or feeder layers to provide factors synthesised by other cell types may improve the situation.

[1] http://www.corning.com/lifesciences/surfaces/en/matrigel.aspx

- **The characteristics of the nutrient medium used:** The medium has to provide the cells with essential nutrients, vitamins, cofactors, metabolic substrates, amino acids, inorganic ions and trace elements. It should also be buffered in order to keep the cells within the appropriate pH range and to limit the effects of acidic products of metabolism (such as CO_2 and lactic acid). Indicators such as Phenol Red are often added to cell culture media in order to allow pH to be visualised easily, although caution should be exercised since Phenol Red has been shown to have weak oestrogenic activity. Finally, cell culture media are usually supplemented with serum (e.g. foetal bovine serum), which provides a range of unspecified growth factors and hormones which help the cells to grow optimally. However, caution should again be exercised since this adds an element of uncertainty to the composition of the medium and can be a source of contamination by mycoplasma, viruses and, in theory, prion diseases such as bovine spongiform encephalitis.
- **The incubation conditions:** Physical factors which are important to the successful growth of cells include the incubation temperature, composition of the gas phase and relative humidity. In particular, mammalian cells are very susceptible to heat shock so it is advisable to set cell culture incubators slightly below the notional optimum growth temperature of 37 °C (say about 36.5 °C).

4.2.1 Primary cell cultures

The ease of culture of a particular cell type is a function of its embryonic origin. Cells of mesenchymal (osteogenic, chondrogenic, adipogenic or myogenic) origin grow readily in culture; this is why fibroblasts tend to overgrow other cell types in the primary cultures of most tissues. Mesenchymal cells are usually anchorage-dependent with a spindle-shaped morphology. Primary cultures of mesenchymal cells are robust and easy to prepare by either explant or enzymatic methods. Cell populations of this embryonic origin also tend to form cell lines readily, possibly because they contain a larger proportion of stem and precursor cells with greater capacity for self-renewal than adult cells. Similarly, tissues with the capacity for renewal *in vivo* (e.g. those of epidermal, intestinal epithelial and haematopoietic origin) lend themselves to the preparation of cell lines whereas tissues which reproduce only under stress often contain few precursor cells and tend to have a limited life span *in vitro*. Tumour cells adapt to culture more readily than normal cells, probably because they have greater proliferative capacity than normal cells.

Endodermal cell types (those derived from the liver, pancreas and respiratory tract) are difficult to culture, but are commonly used in cellular toxicology studies because they represent key organs of xenobiotic exposure and excretion. The type of cultured cell most commonly used in toxicity testing is the primary cultured hepatocyte. This model is considered to be the gold standard for the *in vitro* study of metabolism-dependent toxicity. In order to make a primary culture of hepatocytes, the liver must be digested gently to release the hepatocytes without damaging them. This is done by perfusing the liver with digestive enzymes. In the case of a rodent liver the whole liver is commonly used; when preparing hepatocytes from human liver one lobe will generally be perfused. The tissue should be as fresh as

possible. In the case of human liver this means obtaining it as soon as possible after surgery or after brain death has been established. In order to ensure that the cells retain their differentiated characteristics it is essential that cultures are used quickly (within 4 days) and provided with appropriate signals for differentiation. This can be achieved by coculture with epithelial cells or growth in 3D culture.

The use of primary cultured hepatocytes has been facilitated by the development of cryopreservation techniques which allow viable cells to be stored in liquid nitrogen and taken out for culturing as required. Recent advances in the use of primary human hepatocytes have included the development of new assays using these plateable, cryopreserved hepatocytes.

Example: The Metabolic Comparative Cytotoxicity Assay

The Metabolic Comparative Cytotoxicity Assay (Li, 2009) evaluates toxicity in Chinese hamster ovary (CHO) cells and human hepatocytes in the presence and absence of the CYP inhibitor 1-aminobenzotriazole (a mechanism-based inhibitor of CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4). Enhanced cytotoxicity in human hepatocytes compared with CHO cells suggests that CYP-mediated metabolism plays a role in the observed effects. The Cytotoxic Metabolic Pathway Identification Assay is recommended as a follow up in cases where CYP-mediated metabolism is implicated in the mechanism of toxicity. In this assay, cytotoxicity is evaluated in the presence of isoform-specific CYP inhibitors.

The Metabolic Comparative Cytotoxicity Assay helps to identify cases in which hepatic Phase I metabolism plays a role in toxicity while the Cytotoxic Metabolic Pathway Identification Assay allows the involvement of individual isoforms to be characterised. For example, in a proof of principle study, plateable cryopreserved human hepatocytes from two donors were used to evaluate the role of CYPs in the cytotoxic effects of aflatoxin B1 (AFB1). As expected, AFB1 was more cytotoxic in human hepatocytes than in CHO cells. In the Cytotoxic Metabolic Pathway Identification Assay its cytotoxicity was inhibited by ketoconazole (a CYP3A4 inhibitor) and ditheylthiocarbamate (CYP2A6) but not by the other CYP inhibitors tested (furafylline, sulfaphenazole and quinidine), thus implicating CYP3A4 and CYP2A6 in metabolic activation.

Example: Microarray screening in primary hepatocyte cultures

The feasibility of microarray screening for hepatotoxicity *in vitro* using rat primary hepatocytes has also been investigated using Affymetrix Rat Toxicology U34 arrays (Suzuki *et al*., 2008a). Primary cultures grown on collagen I coated plates were exposed to test compounds at 1/3 of the cytotoxic concentration for 6 or 24 h. The compounds tested were acetaminophen, cyclophosphamide, clofibrate, chlorpromazine, lithocholic acid, cisplatin, diclofenac and disulfiram. It was found that this system could be used to classify compounds as cholestatic hepatotoxicants, non-cholestatic hepatotoxicants and non-hepatotoxicants based on modes of toxicity.

Renal tubular cells may also be grown in primary culture and used to assess nephrotoxicity. They may be used as mixed cultures, or the individual populations

(proximal tubular, distal tubular and collecting duct lining cells) can be separated to generate pure cultures. Such cultures are considered to be a good model for the intact nephron because they

- express tubular segment-specific markers and transporters;
- form tight junctions and microvilli;
- respond appropriately to parathyroid hormone and vasopressin;
- exhibit proximal tubule endocytic capacity;
- produce interleukin and osteopontin;
- exhibit crystal retention capacity.

Primary cultures of renal tubular cells currently represent the closest approximation to the *in vitro* situation in the kidney, and gene expression analysis of the effects of nephrotoxic drugs in rat primary renal cortical tubular cells (Suzuki *et al*., 2008b) suggests that an *in vitro* gene expression analysis approach using transcriptional profiling is feasible for screening for direct tubular toxicity of drugs with the potential to help to clarify the underlying mechanisms of tubular toxicity.

Example: Use of primary cultures of proximal tubular cells to evaluate renal toxicity

Primary cultures of renal proximal tubular cells have the potential to be useful in the evaluation of renal toxicity because they reflect the *in vivo* situation well in terms of the expression of xenobiotic metabolising enzymes and transporters while having the scope for increased throughput compared with *in vivo* studies. This was demonstrated in a proof-of principle study (Cai *et al*., 2009) in which the uptake and cytotoxicity of three candidate anti-inflammatories was compared in human and monkey proximal tubule cell cultures. The results indicated that human proximal tubule cells were equally or less sensitive to the adverse effects of these compounds, correlating with *in vivo* observations and leading to the conclusion that *in vitro* primary human and monkey proximal tubule cells are suitable for use in the screening of compounds for the quantitative prediction of *in vivo* outcomes. In terms of the 3Rs this represents a useful Refinement technique, and while it does not avoid the use of animals altogether, the fact that multiple cultures can be initiated from the kidneys of one monkey does mean that it falls into the category of Reduction compared with *in vivo* methods.

Example: Use of human renal tubular cells *in vitro*

Human tubule cells grown on Transwell® permeable filter supports for 12 days have been validated as a robust, polarised primary cell culture model of the renal tubule (Brown *et al*., 2008). The use of human renal cells in this context was considered important because of species differences in the properties and expression pattern of renal transporters. Net transepithelial fluxes of *p*-aminohippuric acid were demonstrated in proximal tubular cells but not distal tubular cells, as expected, and creatinine transport was also observed. Functional expression of drug transporters in the apical membrane was demonstrated. Mixed cultures were

prepared by collagenase digestion and Percoll density gradient centrifugation yielding 3.5×10^6 cells/g wet weight (enough for 24–36 × 6.5 mm Transwell® cultures) and comprised mostly proximal tubular cells. Pure cultures were made by trypsinising mixed cultures and subjecting the cells to fluorescence-activated cell sorting using leucine aminopeptidase as a marker for proximal tubular cells. The yield of purified proximal tubular cells per sample was sufficient for about six wells at 50 000 cells/well, but mixed cultures were found to be preferable for most experiments.

4.2.2 Established cell lines

The process by which a primary culture may be generated from tissue or organ fragments and then subcultured to form an established cell line which may subsequently evolve, becoming immortalised and finally transformed, is summarised in Figure 4.1. Once the cells have been passaged several times and seem to have formed a stable cell line (usually after about three passages), stocks can be frozen in liquid nitrogen and the cell line can be considered to be established.

In order for a cell line to become established it is essential for the number of cells to increase, implying that the cell type which eventually predominates will have a high proliferative capacity, and it should be noted that the process of passaging, which involves separating the cells from their underlying substrate by

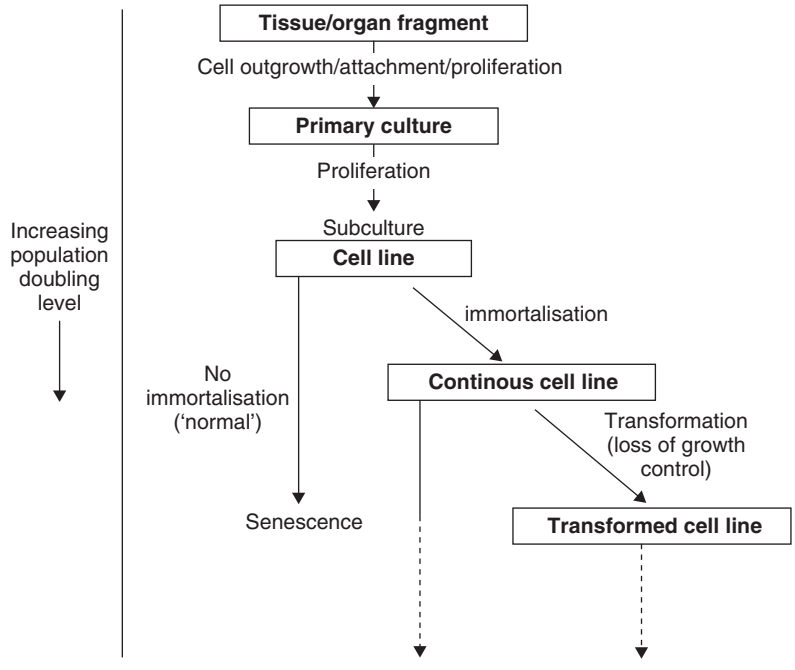


Figure 4.1 Scheme depicting the origin and progression of cell lines with population doubling level. Note that in some cases transformation can occur in the absence of immortalisation (source: Davis (2001); figure 4.1. Reproduced with permission of John Wiley & Sons, Ltd)

scraping or enzymatic digestion (e.g. with a mixture of trypsin and a chelating agent such as ethylenediaminetetraacetic acid (EDTA)), is intrinsically stressful. It involves breaking both cell–cell and cell–substrate interactions, causing cellular injury from which a proportion of cells will not recover. This means that passaging is effectively a selection procedure which may cause a cell line to evolve and change its phenotype over time.

Cell lines which have been established for long-term *in vitro* culture are, by definition, viable in the long term and are relatively easy to maintain, but it is a characteristic of continuously cultured cell lines that they undergo dedifferentiation. Furthermore, the avoidance of cellular senescence, which is essential for the formation of a continuous cell line, is characteristic of the process of transformation. This process is probably associated with the capacity for genetic variation leading to a heterogeneous population from which the most proliferative cell type is effectively selected. Indeed, most continuous cell lines are both aneuploid and heteroploid. Many of them also have an effectively infinite lifespan (immortalisation) and may be morphologically transformed. It is therefore important to work within a very narrow range of passage numbers and to have highly effective procedures in place for stock control, prevention of cross-contamination and authentication of the cell lines used.[2]

Historically, loss of differentiated cellular characteristics has been blamed on dedifferentiation associated with the ability to proliferate; in general, cells which proliferate rapidly tend to exhibit dedifferentiation. However, a number of other explanations are also possible. The wrong lineage of cells may have been selected *in vitro*; the desired cell type may have undergone terminal differentiation and been overgrown by undifferentiated cells of the same lineage; or the cells may have reverted to a more primitive phenotype, or possibly even to a stem cell. This is similar to the process which occurs in the liver during regeneration, when rapid proliferation of a stem-cell-like cell type to replace the lost tissue is followed by differentiation to form hepatocytes. An alternative possibility is that the appropriate factors required for differentiation have been omitted from the culture environment/medium; some cell lines can be encouraged to differentiate by manipulating the culture conditions (e.g. by slowing down the growth of the cells or manipulating the concentrations of hormones and growth factors in the culture medium).

Numerous different cell lines, of human and animal origin, are now available commercially or via research collaborations. Examples of some commonly used liver, lung and colon cell lines are shown in Figure 4.2. When using commercially available cell lines it is strongly recommended that cells are initially obtained from a recognised cell culture collection such as the European Collection of Animal Cell Cultures (ECACC), the American Type Culture Collection (ATCC) or the German Collection of Microorganisms and Cell Cultures, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). The reason for this is that cells from these sources undergo detailed QC to ensure that they are healthy and are free of contamination by microorganisms and other cell lines. It is particularly important to use cell lines from reputable sources because there is evidence that many stocks of human cell lines have been contaminated with cells

[2] For further information on these procedures, see the EC Good Cell Culture Guidelines; Coecke *et al*. (2005).

from other lines and even other species. In the 1960s it was shown that HeLa cells are a common contaminant of human tumour-derived cell lines (Gartler, 1967, 1968), although with the development of techniques for analysing marker chromosomes and enzyme polymorphisms, the authenticity of certain cell lines was confirmed (Fogh *et al*., 1977, Wright *et al*., 1981). These methods, together with DNA-based techniques such as DNA fingerprinting and short tandem repeat analysis (Nims *et al*., 2010), mean that there is now no excuse for using misidentified or cross-contaminated cell lines. Unfortunately, however, the continuing prevalence of poor tissue culture practices such as inadequate segregation of cell lines and reagents means that many published cell culture studies (possibly as many as 15–20%) are still found to be compromised (Nardone, 2007, Lucey *et al*., 2009, Capes-Davis *et al*., 2010).

4.3 Acute toxicity *in vitro*

The first step towards replacing the use of animals in toxicity testing is to identify *in vitro* methods which can be used to predict acute toxicity. This involves the selection of appropriate methods of cytotoxicity testing and the choice of a suitable cell line model.

4.3.1 Cytotoxicity testing

Assays for cytotoxicity have conventionally included dye exclusion and dye uptake assays (e.g. Trypan Blue exclusion, Neutral Red uptake) and colony forming assays; however, these are becoming increasingly unpopular because they tend to be laborious, time consuming and insensitive. Most importantly, from the perspective of the modern *in vitro* laboratory, they do not adapt well to HTS approaches. Modern cytotoxicity testing methods focus on parameters which are better suited to HTS, including biochemical events associated with loss of viability and changes in membrane integrity leading to cellular disintegration.[3]

Viability assays based on dye reduction have now been popular for several decades. The first of these was the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay, in which the parameter measured was reduction of the yellow formazan salt MTT to an insoluble purple crystalline product. The MTT assay was the first cell viability assay which could be carried out using 96-well plates, but it was limited by the fact that the product of the reaction was insoluble and had to be solubilised using a solvent such as dimethyl sulfoxide (DMSO) before reading the results. This meant that the MTT assay could only be used as a single time-point assay. The next tetrazolium assay to be developed was the XTT (2,3-bis (2-methoxy-4-nitro-5-sulfophenyl)-5-[(phenylamino) carbonyl]-2H-tetrazolium hydroxide) assay, in which the product of the reaction was soluble; however, unfortunately, in this case it was the substrate which had solubility issues.

Third generation dye reduction assays include the MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) and WST

[3] Reviewed by Niles, Moravec and Riss (2008).

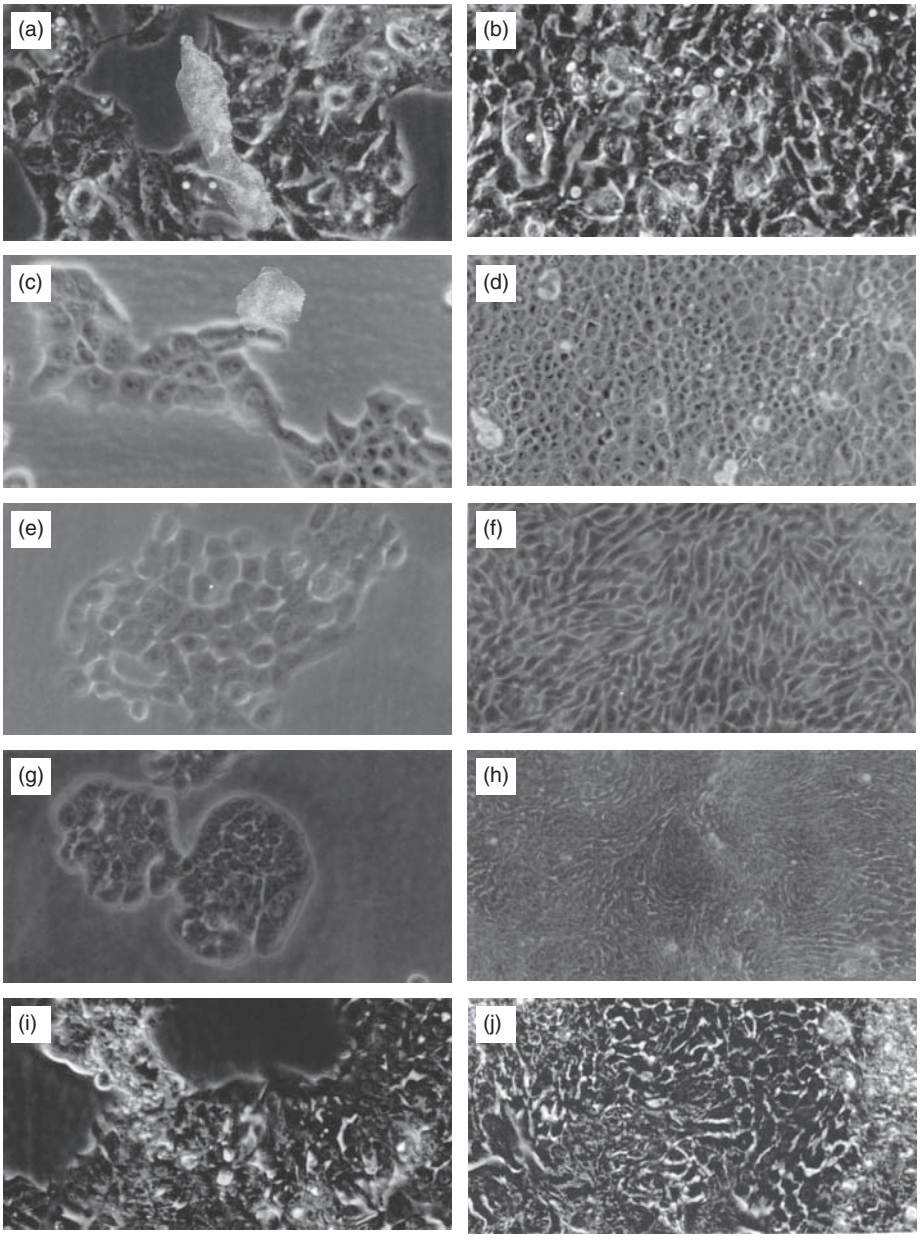


Figure 4.2 Morphology of human liver, lung and colon tumour-derived cell lines. Logarithmic and confluent cultures were photographed at a magnification of 100–150× using an Olympus C-35DA-2 automatic camera attached to an Olympus TO41 phase contrast microscope. White light (setting 12) was passed through a green filter and the exposure set at 0.1 s. (a) HepG2 (liver) – Logarithmic phase; (b) HepG2 (liver) – Confluent phase; (c) NCI H322 (lung) – Logarithmic phase; (d) NCI H322 (lung) – Confluent phase; (e) NCI H358 (lung) – Logarithmic phase; (f) NCI H358 (lung) – Confluent phase; (g) HT29 (colon) – Logarithmic phase; (h) HT29 (colon) – Confluent phase; (i) LS174T (colon) – Logarithmic phase; (j) LS174T (colon) – Confluent phase

(water soluble tetrazolium) assays, in which the substrates are readily soluble, easy to prepare and stable in solution. These assays can be enhanced by the use of electron coupling reagents to amplify the signal.

The Resazurin assay, which is based on the reduction of the blue dye Alamar Blue to a pink fluorescent resorufin product by intracellular reductase and diaphorase enzymes, has also become increasingly popular in recent years. This assay can be run in real time mode or as an endpoint assay for HTS, in which case the dye is added to the cultures 1–3 h before reading the plates.

Homeostatic control ensures that ATP levels are relatively constant in viable cells, but rapid ATP depletion occurs as a result of cytotoxicity, and bioluminescent ATP-based assays, based on the use of luciferase enzymes which have been modified to generate a steady luminescent signal, are a popular alternative to dye reduction assays. These are now the method of choice in HTS laboratories because they are simple, rapid and are not affected by the interference from test compounds, a common problem with fluorescent assays. They do need to be used judiciously, however: the procedure kills the cells, so they cannot be combined with measurement of other parameters, and it is important to check that the test compounds being used are not luciferase inhibitors.

Enzyme release-based assays represent an alternative to assays based on changes in cellular biochemistry. These involve the measurement of constitutive, conserved, stable enzymes which are released from dying cells as their membranes break down. This approach is well suited to the toxicology and drug discovery setting because it is amenable to adaption for HTS and represents a *bona fide* toxic response. The most popular version of the type of assay, which has been used both *in vitro* and *in vivo* for many years, is based on the release of lactate dehydrogenase (LDH). However, this cannot be conducted on cells grown in serum-supplemented media, and problems can arise as a result of signal quenching, autofluorescence and enzyme inhibition, so additional assays – for example – those based on the release of adenylate kinase and glyceraldehydes-3-phosphate dehydrogenase) have been developed more recently. Modern enzyme release assays are available in HTS format (384/1536 wells) with short incubation time (5–30 min) and in either fluorescent or luminescent formats.

4.3.2 Choice of cell line

Cell lines from different tissues can be used to compare the target organ toxicity of xenobiotics.

Example: Use of human hepatocellular and renal tubular cell lines for *in vitro* toxicity testing

The human hepatoma cell line Bel-7402 and the renal tubular cell line HK-2 have been used to compare the hepatotoxic and nephrotoxic effects of a range of xenobiotics (Zhang *et al.*, 2007). Acetaminophen, mitomycin C and copper chloride were more toxic in liver cells than kidney cells (based on concentrations giving 50% inhibition (IC_{50} values) in the two cell lines), whereas cadmium chloride was more toxic to kidney cells. These results correlated with the effects of these chemicals *in vivo*. However, phenacetin and aristolochic acid were more toxic to the liver

cell line than the kidney cell line *in vitro* despite being kidney toxicants *in vivo*. This observation, which may have been due to the fact that cell culture systems do not take into account the toxicokinetic processes which help to determine target organ toxicity *in vivo*, illustrates the need for caution in extrapolating from *in vitro* to *in vivo* systems.

The replacement of the LD_{50} test (the test which identifies a dose of chemical which kills 50% of the treated animals) is a key priority. During the first decade of this century two programmes, 'Evaluation-guided Development of New *In vitro* Test Batteries' and 'Multi-centre Evaluation of *In vitro* Cytotoxicity', set out to develop an acute test battery which could replace the LD_{50} test (Clemedson *et al*., 2002, Clemedson *et al*., 2003). Of 50 chemicals tested, the majority showed *in vitro* concentrations giving 50% of maximal activity (EC_{50}s) which correlated well with human lethal blood concentrations. Most of the chemicals for which this correlation did not apply were neurotoxicants. For these chemicals the human neuroblastoma cell line SH-SY5Y was proposed as a suitable model system although currently no single cell type possessing all the relevant receptors is available.

These programmes laid the foundations of a project funded under the EU Framework Six Programme entitled ACuteTox (Clemedson *et al*., 2007, Clemedson, 2008) which was designed to further extend the range and capabilities of *in vitro* testing for acute toxicity. The aim of the ACuteTox project was to demonstrate that the animal tests for acute systemic toxicity currently used for regulatory guideline purposes could be replaced by a combination of alternative assays. The results of the Multi-centre Evaluation of *In vitro* Cytotoxicity programme had indicated that, while correlations between *in vitro* EC_{50}s and human lethal blood concentrations were generally good, basal cytotoxicity did not always provide a satisfactory prediction of acute systemic toxicity. The main reason for this was the assumption that the concentration of a compound to which cells are exposed in an *in vitro* test is comparable to that to which they would be exposed *in vivo*. This would only be true in a situation where there is rapid absorption, slow elimination and no metabolic activation. The assumption that systemic toxicity is not driven by organ-specific effects is also implied. Thus, the reasons for poor correlations between *in vitro* cytotoxicity and *in vivo* acute toxicity include poor bioavailability, organ-specific effects and rapid metabolic elimination. In addition, effective *in vitro* concentrations may be lower than those intended due to factors such as protein binding within cell culture medium.

The primary aim of the ACuteTox project was to permit the regulatory classification of high-production volume chemicals, a secondary objective being to contribute data for use in risk assessment. It was also noted that a test battery of this type could have applications during drug development (Clemedson, 2008), perhaps in reviewing candidate compounds with a view to identifying those which are unlikely to have suitable druggability.

The outcome of the ACuteTox project was the recommendation of a test battery of five to seven *in vitro* and *in silico* tests including a HTS basal cytotoxicity test (e.g. the 3T3 cell neutral red uptake test), prediction of access of parent compound and metabolites to the target tissue, *in silico* estimation of uptake and distribution and predictions of organ-specific toxicity, particularly neurotoxicity, nephrotoxicity, hepatotoxicity and haematopoietic effects. This would allow the toxicity of

90% of compounds to be predicted; however, the project team recognised that the desired test battery would be too expensive for routine use (Clemedson, 2008) and suggested a compromise in which two to three tests could be used, with the proviso that a neurotoxicity endpoint was essential but with the note that hepatotoxicity testing was not required for the prediction of acute toxicity.

At the time of writing, the final results of the ACuteTox programme are in the process of being published. An analysis of rodent LD_{50} data for 97 reference substances for variability, reliability and predictive capacity has been completed, allowing performance criteria for *in vitro* methods to be developed (Hoffmann *et al*., 2010), and a candidate neurotoxicity assay has been identified (Gustafsson *et al*., 2010). Further publications arising from the programme were in press at the time of writing (Clothier *et al*., 2012; Kinsner-Ovaskainen *et al*., 2013; Kopp-Schneider *et al*., 2013; Prieto *et al*., 2012).

4.3.3 Liver

The hepatocyte is one of the cell types most commonly used for *in vitro* toxicology studies because of the importance of the liver in xenobiotic metabolism and clearance. Hepatocytes do not form cell lines because they very rarely divide. They can, however, be cultured as monolayers for up to about 4 days, long enough to be useful for some types of experiments.

The requirement to have fully functioning cells for the duration of study is an essential step if hepatocyte cultures are to be used for prolonged periods of exposure. These problems of long-term viability and functionality are beginning to be overcome by the use of a range of techniques including gel sandwich (using collagen to provide a structured 3D-base for the cells), liver spheroids (again providing a 3D-frame using a gel sphere as the base) and precision-cut liver slices (which retain the structural relationships of the original tissue).

The benefit of creating a 3D-structure within hepatocyte cultures has been shown in several ways, both for primary cells and for cell lines. The use of a sandwich approach with a collagen underlayer and proprietary extracellular matrix (Matrigel®) overlayer has been described as improving the differentiation and performance of primary human hepatocytes and HepG2 (human hepatoblastoma) cells in culture.

Human liver-derived cell lines Comparison studies have concluded that primary cells in culture are more like intact liver than are cell lines such as HepG2 and Huh7, but for reproducibility of response it is often considered preferable to use a stable cell line for *in vitro* testing rather than primary cells.

The optimum choice of a cell-line for testing has yet to be resolved; however, many metabolic and short-term mechanistic studies have used HepG2 cells, and a significant amount of prior knowledge is available on the performance of this line. The genomic profile of the HepG2 line has been compared with that of primary human hepatocytes following exposure to ciprofloxacin, levofloxacin, gatifloxacin or trovafloxacin revealing a similarity in the response of both cell types to the latter test substance (Liguori *et al*., 2008). Trovafloxacin had been demonstrated to have a unique toxicity in human hepatocytes and this was reflected in a specific gene

expression profile, which was replicated in the HepG2 cells, although other, less critical, end-points showed some differences between the gene expression profiles in the two cell types. The data support the possibility that HepG2 cells could find use in *in vitro* toxicity assays.

Example: Multi-endpoint testing using HepG2 cells

A multi-endpoint mechanistic approach to using HepG2 cells in acute toxicity testing has been described (Flynn and Ferguson, 2008). Different aspects of hepatotoxicity were modelled using *in vitro* assays, using:

- Total double stranded DNA as an indicator of cell death.
- Rhodamine uptake and retention as a marker for mitochondrial membrane depolarisation/transporter induction/inhibition.
- Nile Red uptake as a marker for steatosis.
- Dihydrochlorofluorescein oxidation as a marker for oxidative stress.
- Ethoxyresorufin-O-deethylation as a marker for modulation of CYP1A activity.
- Benzyloxyresorufin-O-debenzylation as a marker of CYP2B and CYP3A activity.

The data demonstrated some success in fitting the known toxicity of the hepatotoxicant set (6 compounds) and predicting potential hepatotoxicity among the unknowns (11 compounds). These procedures might form the basis of a useful approach to repeated-dose testing, particularly the statistical approaches for deriving a suitable prediction model from multiple end-points.

The HepaRG cell line, derived from the tumour of a female patient suffering from HCC and hepatitis C infection, has also become increasingly popular in recent years.[4] This cell line appears to be more representative of normal human liver than many of the other cell lines in common use, at least according to transcriptional profiling. In early stage cultures it has an epithelial phenotype, but it is thought to represent a population of bipotent progenitor cells because in longer term culture it can differentiate to form cholangiocytes and hepatocyte-like cells. After culturing in the presence of DMSO for about 2 weeks, about 50% of the cells are hepatocyte-like and form structures resembling bile canaliculi. They express hepatocyte-like levels of xenobiotic metabolising enzymes (CYPs, GSTs and UGTs) and transporters, but this is dependent upon continued exposure to DMSO. It should also be noted that genotype of the HepaRG cell line indicates that it is likely to have a poor/intermediate metaboliser phenotype for CYP2D6 and CYP2C9.

The main use of the HepaRG line so far has been for *in vitro* studies of CYP induction and associated functions, although optimisation studies have shown that it can respond to toxic insults by up-regulation of pathways of toxicity including steatosis, phospholipidosis and cholestasis (Antherieu *et al.*, 2012). In addition, proof-of-principle studies have indicated that it is a suitable model for studying toxicity, especially that of compounds requiring metabolic activation (e.g. acetaminophen and heterocyclic amines).

[4] Reviewed by Andersson, Kanebratt and Kenna (2012).

HepaRG cell populations have been described as resembling an average population of human hepatocytes, with the caveats that they seem to contain a subpopulation of progenitor or primitive biliary epithelial cells, exhibit chromosomal abnormalities and express a number of genes related to their transformed phenotype (Rogue *et al*., 2012). A comparison between HepaRG and other human hepatocellular culture models (HepG2, Hep3B, Huh7, SK-HEP-1 and primary hepatocytes), based on RT-PCR analysis of the expression of 251 xenobiotic metabolising enzymes, has indicated that the pattern observed in HepaRG cells is the most similar to that of human hepatocytes; however, none of the established cell lines fully reflected the pattern seen in primary human hepatocytes (Guo *et al*., 2011). Similarly, a whole-genome transcriptional profiling analysis of basal gene expression in HepaRG and HepG2 cells indicated that HepaRG cells bore a closer resemblance to primary hepatocytes; however, profiling after exposure to AFB1, benzo(a)pyrene (B(a)P), cyclosporin A, 17β-estradiol and TCDD suggested that HepG2 cells were better at discriminating between the effects of genotoxic and non-genotoxic carcinogens (Jennen *et al*., 2010). In combination, the results of these studies show that both cell lines have a role to play in *in vitro* toxicology; however, they also highlight the need for thorough characterisation of the cell lines being used and caution in interpreting the results obtained.

In addition to conventional 2D-monolayer cultures, both HepaRG and HepG2 cells are starting to be used in 3D-culture systems. Both have been grown as spheroids and subjected to preliminary characterisation in terms of enzyme expression and responses to representative toxic compounds. In each case the results indicated that 3D-cultures had significantly improved functions compared with the same cell lines in monolayer culture (Leite *et al*., 2012; Fey and Wrzesinski, 2012).

Overall, in the limited characterisation undertaken to date, the HepaRG cell line has shown great promise as an *in vitro* model for hepatotoxicity and xenobiotic metabolism; however, the HepG2 line has a long history of use in a variety of contexts, and it continues to be used as a representative human liver cell line in, for example, multi-cell line approaches to the determination of organ-specific toxicity (Lin and Will, 2012).

4.3.4 Skin

The skin is the largest organ of the body and has four key functions: protection against physical and chemical insults (barrier function), sensation (touch), thermoregulation and metabolism. The three main layers of the skin are the epidermis, beneath which is found the dermis and the underlying hypodermis. The epidermis is made up of stratified squamous keratinising epithelium comprising five main sublayers: from top to bottom, the stratum corneum, stratum lucidum, stratum granulosum (granule cell layer), stratum spinosum (prickle cell layer) and stratum germinativum/stratum basale (basal cell layer). The basal cell layer, which is responsible for renewing the skin, contains stem cells together with keratinocytes, melanocytes and a subpopulation of neuroendocrine cells called Merkel cells. The main cell type in the upper layers of the skin is the keratinocyte. These move up through the layers of the skin and undergo terminal differentiation to form the squames (dead, scale-like cells) of the stratum corneum.

The *in vitro* models available for the study of chemical effects on the skin include organotypic cultures of pig and guinea pig skin (which is really just a Refinement approach because it still requires the use of experimental animals) as well as 3D human skin models. *In silico* models for skin toxicity are also improving rapidly in relevance and utility.

***Ex vivo* human and animal skin** After liver, skin is the most desirable tissue for use in toxicity testing. This is generally prepared in the form of skin discs, cut to a thickness of approximately 500 μM using a dermatome and comprising the epidermis plus part of the dermis. Skin discs are used in testing for penetration, metabolism of topically applied compounds, irritation and corrosivity testing. The use of skin preparations for corrosivity and phototoxicity testing is the only human tissue based technique which has undergone a full validation process under EU regulations. A complete list of the available non-animal test methods for skin irritation/corrosion may be found on the ALTEX web site.[5]

Various tissue banks now provide *ex vivo* human skin samples on a semi-commercial basis. Viable skin is commonly obtained as surgical waste from breast reduction or tummy tuck operations. Abdominal skin from the latter source is considered preferable since it is believed to carry out more of the normal functions of the skin than does breast tissue. The main disadvantage of using *ex vivo* skin is variability between samples, due both to inter-individual differences in enzyme expression and possible variability in quality. Both interspecies and inter-laboratory variation is observed. The use of human tissue has a number of additional disadvantages, including difficulty in obtaining fresh human skin for immediate experimental use and source-dependent variability in tissue handling and tissue quality which can affect the level of xenobiotic metabolising enzymes. In addition, depending on the criteria specified for the study, the samples might also have undergone a freeze-thaw cycle.

Animal skin is easier to work with than human skin because it can usually be used fresh whereas human skin is often obtained in the frozen state. Pig and guinea pig skin are widely considered to be the best surrogates for human skin for use in organotypic culture (Barbero and Frasch, 2009). Pig skin is a particularly good model because it has a thin hair coat, a thick epidermis with well-differentiated underlying structure, a dermis with a well-differentiated papillary body and a large content of elastic tissue. However, it has been argued that the hairless guinea pig is a better model because it has:

- Smooth skin, devoid of hair.
- Sufficient skin for multiple replicates.
- Thick epidermis with distinct strata.
- Presence of both serrated and non-serrated keratinocytes.
- The outlines of corneocytes are similar to those in human skin.

Despite some issues, both pig and guinea pig skin are good surrogates for human skin. When data from the various systems are subjected to correlation analysis, the slopes are close to unity and the intercepts are near zero. In general, the animal

[5] http://alttox.org/ttrc/toxicity-tests/skin-irritation

models give excellent results in terms of calculated permeability coefficients but human skin may still be required for the measurement of lag times.

Three-dimensional human skin models The variability and limited availability of human skin samples have led various investigators to develop *in vitro* 3D-skin models to study metabolic behaviour of human skin, as well as efficacy and safety of chemical and product exposures. Reconstructed human epidermis (RHE) models are now commonly used for the *in vitro* evaluation of topically applied chemicals with respect to various characteristics, including cutaneous irritancy, percutaneous absorption and cutaneous xenobiotic metabolism. These reconstructed epidermis models have a number of strengths compared with either *ex vivo* skin or monolayer cultures:

- Their structure mimics the architecture of the normal epidermis.
- They can be used to assess biological parameters underlying various types of skin damage.
- They permit direct application of chemicals and finished products to the stratum corneum.

RHE models are derived from human keratinocyte cultures which undergo differentiation to form three-dimensional structures that closely resemble human skin.[6] The available models include the following.

- **The EpiDerm™ model** (Cannon *et al*., 1994), marketed by MatTek corporation,[7] is a three-dimensional multi-layered skin model derived from human infant foreskin keratinocyte cultures which are induced to differentiate by air-lifting the growing cultures so that the apical surface is exposed to the atmosphere while the basal layers remain in contact with the culture media. Following differentiation to form basal, spinous, granular and cornified layers, the EpiDerm™ model resembles human epidermis (Figure 4.3).
- **The EpiSkin™ model** (Tinois *et al*., 1991), marketed by SkinEthic™ laboratories, is derived from normal adult human keratinocytes removed during plastic surgery. These are grown on a specially developed dermal substrate comprising oxidised Types I and II collagen covered with a film made up of native and oxidised Type IV collagen. When exposed to the atmosphere, cells grown in this way differentiate to form a stratified epithelium with an apparently continuous basement membrane and a dense horny apical surface. Its performance in corrosivity testing has been evaluated against OECD guideline requirements and found to be satisfactory, at least for a small number of compounds (one positive control, one negative and seven test compounds) (Deshmukh *et al*., 2012).
- **SkinEthic™** is a third reconstructed epidermis model marketed by SkinEthic™ Laboratories (Rosdy and Clauss, 1990). This model, in which second passage human keratinocytes derived from surgically excised adult skin are grown on filters at the air liquid interface of a chemically defined medium,

[6] For a review comparing these three models, see Netzlaff *et al*. (2005)
[7] For more information, see http://www.mattek.com/pages/in-vitro-toxicology/.

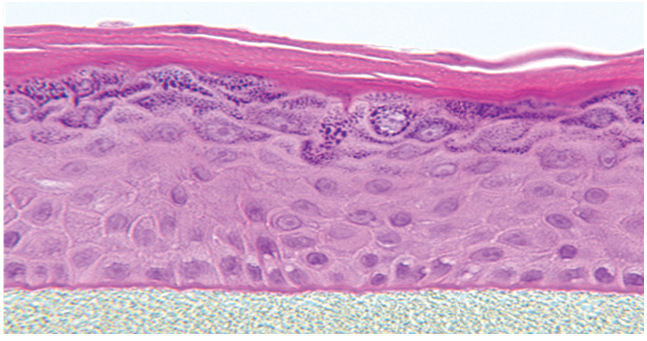


Figure 4.3 Morphology of the EpiDerm™ reconstructed human epidermis model. Haematoxylin and Eosin stained cross-section of the EpiDerm™ Tissue Model at 200× magnification. (source: MatTeck Corporation. Reproduced with permission of MatTek Corporation)

generates a multi-layered epithelium whose architecture is similar to that of human skin at the light and electron microscope levels. Layers corresponding to the stratum basale, stratum spinosum and stratum granulosum form and are topped by an anucleated stratum corneum containing at least 10 compact cell layers.

- **LabCyte™ EPI-MODEL** (Niwa *et al.*, 2009) consists of normal human epidermal keratinocytes forming a multi-layered, skin-like structure on specially prepared permeable cell culture inserts (12 mm diameter; culture area 1.1 cm^2). The cells in this model organise themselves into consecutive basal, spinous and granular layers and a multi-layered stratum corneum with a high concentration of keratohyalin granules and desmosomes. This model, which is similar to EpiSkin™ and is commercially available in Japan,[8] has been characterised in terms of changes in gene expression following exposure to the detergent sodium dodecyl sulfate (SDS) (Niwa *et al.*, 2009) and ability to predict percutaneous absorption and human pharmacokinetics of topically applied compounds (Hikima *et al.*, 2012). Evaluation against the ECVAM performance standards for skin irritancy testing and OECD Test Guidelines for *in vitro* skin corrosion tests has indicated its suitability for these test methods (Katoh *et al.*, 2009, Katoh *et al.*, 2010).

All these models have a similar lipid profile (although they do tend to contain intracellular lipid droplets which are not seen in human skin) and express a similar pattern of biochemical markers (e.g. keratins 1 and 10) to those of human skin. They do, however, express some additional markers such as keratin 6 and the protein SKALP, which is associated with psoriasis and wound healing *in vivo*. Overall, their general histology, composition and biochemistry are representative of human skin. They are consistently more permeable, but this disadvantage is considered to be overcome by their consistency and reduced variability compared with *ex vivo* skin.

[8] For details, see http://www.jpte.co.jp/english/business/LabCyte

Skin corrosion and irritation Skin irritation is defined as a local, non-immunogenic inflammatory reaction which appears shortly after exposure and usually disappears after a few days. The processes involved in skin and eye irritation have conventionally been perceived as relatively simple, and were therefore expected to be relatively easy to replicate *in vitro*. To some extent, this has proved to be the case, although the predictivity of alternative tests in terms of human risk assessment is not always satisfactory (Basketter, 2012).

The most severe form of chemically induced skin toxicity is corrosion, defined as irreparable damage to the skin which encompasses both the epidermis and dermis. Corrosion involves direct chemical damage leading to disintegration and irreversible damage at the site of contact. It requires the chemical to be able to penetrate the stratum corneum by diffusion or erosion and exert direct cytotoxicity in the underlying layers. Irritation is a less severe effect, comprising a reversible local inflammatory reaction mediated by the local innate immune system. It involves the release of inflammatory mediators which cause vasodilation and fluid leakage from capillaries, induce chemotaxis of mast cells and neutrophils and stimulate nerve endings leading to itching and stinging sensations. Irritant contact dermatitis is caused by occupational and domestic exposure of the skin to strong irritants such as detergents and alkalis, and commonly affects the hands.

The identification of primary irritants has conventionally been achieved by means of macroscopic observation of effects on the skin/eye *in vivo*. Traditional animal testing for irritants has used the rabbit because its skin is very thin, allowing subepidermal and subcutaneous inflammation to be recognised. In the Draize test, for example, the test item is applied to shaved rabbit skin which is then covered (occluded) and left for 24 h. Erythema and oedema are then assessed at time points of 24 and 72 h. Draize tests have been heavily criticised because of animal welfare concerns, and they also have a number of scientific deficiencies, including lack of mechanistic information, subjectivity in scoring, irrelevance to accidental human exposure, cost (in terms of both animal numbers and operator time) and lack of reproducibility (including variability according to strain, age and sex of the test animals used).

Draize tests are now being phased out in favour of *in vitro* alternatives; in particular, the 7th Amendment outlawed the use of animals for irritancy testing of cosmetic ingredients with effect from 2009. However, as of 2010 the OECD guidelines for acute dermal irritation and corrosivity testing of chemicals and the guideline for non-clinical dermal tolerance testing of medicinal products were still based on the use of the Draize test.

The use of *in vivo* models to study dermal penetration and irritation has a number of disadvantages, including the fact that they tend to over-predict effects, significant species differences are observed, and they require a large number of animals (and are therefore costly and cumbersome). Furthermore, implementation of the 3Rs principle, together with enactment of the 7th Amendment and the REACH regulations, has led to increased utilisation of alternative methods for the testing of potential skin toxins, particularly irritants.

Both EpiDerm™ and EpiSkin™ can be used for the *in vitro* evaluation of topically applied chemicals with respect to various characteristics, including cutaneous irritancy, percutaneous absorption, dermal xenobiotic metabolism and genotoxicity. A comparison of historical results obtained using EpiDerm™ and EpiSkin™

with those of the Draize tests and human skin irritation assessed in a volunteer study indicated that the concordance between the Draize test and actual human effects was poor (56%), whereas the two reconstructed skin models, while not perfect, did give much better predictions (69.6% for EpiSkin™ and 76.0% for EpiDerm™) (Jirova *et al*., 2010).

Example: Application of the EpiSkin™ model in microarray analysis

Skin irritation potential is defined as the reversible inflammatory response of the epidermis to a topically applied substance. A typical application is testing the skin irritation potential of new cosmetic product formulations, while in development, to identify chemicals that might induce adverse skin reactions. The EpiSkin™ model has been used for microarray analysis of skin irritation using the DualChip™ Human Aging array, a low- density array of 240 genes associated with the keratinocyte phenotype, cell senescence and stress responses (Borlon *et al*., 2007). The standard EpiSkin™ skin irritation test, in which chemicals are applied topically, left in contact with the tissue construct for 15 min and rinsed off, was used. Microarray analysis was carried out after a 42-h recovery period. The test chemicals comprised four irritants (potassium hydroxide, 1-bromopentane, 10-undecenoic acid and heptanal, used at sub-cytotoxic concentrations) and four non-irritants (3,3-dithiodipropionic acid, 4-amino-1,2,4-triazole, 4,4-methylene-*bis*-(2,6-butyl)phenol and 3-chloronitrobenzene).

This study identified 51 genes whose expression was significantly affected ($p < 0.01$), 16 of which responded to irritants but not to non-irritants. The main molecular mechanisms with which these genes were associated were the cell cycle, chromosomal processing and cell differentiation (e.g. cyclin D1, plasminogen activator inhibitor 1, cathepsin H and cystatin 6); apoptotic genes did not seem to respond to irritants. Certain responses were also associated with specific irritants, indicating that toxicogenomic technologies may be of value in identifying novel mechanisms of skin irritancy.

Reconstructed human skin models are now widely accepted for corrosivity, and several protocols have been subjected to formal validation by ECVAM.[9] Four methods, based on EpiSkin™, EpiDerm™, SkinEthic™ and EpiCS® (EST-1000), have been accepted by the EU (method description B.40 bis) and OECD (Test Guideline #431), and approved protocols are available via the ECVAM web site.

Three reconstructed human skin models have undergone formal validation by ECVAM and as a result it has recommended tests using EpiSkin™, EpiDerm™, and SkinEthic™ as non-animal methods for irritancy testing (OECD Test Guideline #439/EU Test Method B.46). The documentation relating to the ECVAM validation process is available via the ECVAM web site. Performance standards for alternative tests were developed in 2005–2007 and applied to the EpiSkin™ and EpiDerm™ test methods, and the inclusion of IL-1α release as an additional endpoint was identified as a useful adjunct to the EpiSkin™ test. The SkinEthic™ test and a modified version of the EpiDerm™ test (EpiDerm™ SIT) were endorsed by ECVAM in 2008 following a catch-up/me too validation exercise (Alepee *et al*., 2010, Tornier *et al*., 2010).

[9] For full details and links to key documents, see http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/validation-regulatory-acceptance.

Phototoxicity Phototoxicity is defined as a toxic response elicited after initial exposure of the skin to a chemical with subsequent exposure to light. It can occur in response to exposure to chemicals which absorb visible or UV light. ECVAM has approved an *in vitro* phototoxicity assay, the 3T3 NRU PT assay, which measures cytotoxicity of compounds in mouse (Balb/c) 3T3 fibroblasts in the presence and absence of simulated sunlight. This assay is covered by OECD Test Guideline #432, EU Test Method B.42 and INVITTOX Protocol No. 78. It is considered to be capable of identifying chemicals as being definitively non-phototoxic (i.e. a negative result at a high concentration is sufficient confirmation of photosafety), but has been criticised on the grounds that it is over-sensitive and over-predictive. It is also limited in that the 3T3 cell culture model lacks metabolic capability and does not replicate the barrier function of human skin. Furthermore it is unable to replicate *in vivo* exposure conditions and provides no information about the mechanism of any effects observed; for example, it cannot discriminate between direct phototoxins and photoallergens.

The 3T3 NRU PT test is now used in hazard assessment as a Tier I test for phototoxicity. Positive results in the 3T3 NRU PT test should not preclude further development of the compound but do indicate that follow-up testing is required (Ceridono *et al*., 2012). In the case of a putative positive, the compound can be subjected to Tier II testing with the aim of excluding false positives. Tier II testing often uses RHE models such as EpiDerm™ or EST-100 (H3D PT), especially for compounds such as cosmetic ingredients which are likely to be applied topically. RHE models have the advantage that they are metabolically competent and do have a barrier function, which can be modulated by adjusting the culture conditions. They are also of consistent quality and are easier to treat with chemicals which have poor solubility.

Skin sensitisation Allergic contact dermatitis may occur in predisposed individuals in response to skin contact with allergens such as nickel, cosmetics, dyes and rubber. It tends to be localised at the site of contact (e.g. the wrist or the back of the neck, where jewellery makes contact). Skin sensitisation, which can lead to allergic contact dermatitis, can be described as a delayed-type hypersensitivity reaction induced by exposure to a low-molecular weight reactive chemical (a hapten) which can penetrate the skin and bind to the proteins, possibly as a result of metabolic activation and/or chemical oxidation. Skin sensitisers are able to induce allergic contact dermatitis, which comprises two phases, induction of immunological memory following exposure to a sensitising substance and elicitation of a cell mediated allergic response by subsequent exposure to the same allergen[10] The marked clinical, pathological and immunological similarities between chemically induced skin irritation and allergic contact dermatitis make it very important to consider the mechanism when evaluating these processes (Ku *et al*., 2009).

Skin sensitisation has historically been detected by means of tests in guinea pigs (the guinea pig maximisation test and the Buehler occluded patch tests) while the murine local lymph node assay has been preferred more recently. This assay is based upon the quantification of cell proliferation in the draining auricular lymph nodes of mice after repeated topical application of the test compound and is considered in 3Rs terms to be a Refinement because it provides for better animal

[10] For further detail, see Male *et al.* (2012) Immunology (6th Edn) Chapter 26.

welfare than the guinea pig methods. The scientific advantage of this assay is that it gives an indication as to the relative potency of different skin sensitisers as well as a yes/no answer regarding skin sensitising potential.

The toxicological evaluation of potential skin sensitisers without the use of animals presents a huge challenge, particularly to the cosmetics industry, because it is essential to assess the potential of cosmetic ingredients to induce skin sensitisation under the conditions of use which could lead to allergic contact dermatitis in humans, but there is no single non-animal model which can reproduce all the processes which need to be taken into consideration. Indeed, no *in vitro* methods have as yet been accepted as partial or full replacements for existing animal tests and the possibility to recommend a single non-animal test method in the foreseeable future is unlikely. Consequently, the risk assessment of cosmetic ingredients for skin sensitising potential does not follow a standardised procedure but tends to be carried out on a case-by-case basis following a weight of evidence approach in which results from a number of alternative (*in vitro*/*in silico*) methods are combined. These can include the following.

- Characterisation of physicochemical properties and read-across based on structurally similar chemicals.
- Toxicokinetic modelling of skin exposure and bioavailability.
- Direct peptide reactivity assay (DPRA) for modification of skin proteins (alternatively, these may be predictable *in silico* on the basis of the known reactivity of the chemical with cellular nucleophiles.[11]
- Assessment of dermal metabolism using *ex vivo* human skin tissue, RHE constructs and immortalised keratinocyte models.
- Characterisation of innate immune responses in the skin, for example, the role of dermal fibroblasts and/or epidermal keratinocytes in activating Langerhans and dendritic cells.
- Measurement of sensitiser-induced dendritic cell activation, currently achieved by means of either the myeloid U937 skin sensitisation test (MUSST) or the human cell line activation test (h-CLAT). New biomarkers for dendritic cell activation are also under investigation; these include activation of p38 mitogen-activated protein kinase (MAPK) and inhibition of ERK 1/2 kinase.
- Identification of sensitiser-specific T-cell responses remains an issue. The best assay currently available is the T-cell priming assay in which sensitiser-treated, irradiated human monocytes are co-cultured with CD25-depleted T-cells, which have a reduced activation threshold, making them more sensitive to weak sensitisers, but the need for a better T-cell activation assay is still an issue in skin sensitisation.
- Where available, previous human experience and historical data from animal studies may be included. The types of human data which may be available include consumer experience and feedback, results from clinical studies, occupational exposure data, case reports, consumer tests and the results of volunteer studies.

Overall, the current guiding principles for the safety assessment of skin sensitisation start with hazard identification, including an examination of existing data to

[11] Discussed in detail by Goebel et al. (2012).

see whether the compound can be classified as a non-sensitiser. This is followed by an evaluation of physicochemical properties and the generation of *in vitro* data including assessment of metabolism. Three *in vitro* tests, the DPRA, MUSST and h-CLAT are in the process of validation by ECVAM and these are currently the preferred methods. An additional method, the Keratinosens assay, has undergone non-ECVAM trials and at the time of writing these are being peer-reviewed by ECVAM. In order to extend this hazard assessment to a full risk assessment, exposure data based on knowledge about the habits of use of cosmetics and toiletries, refined by consideration of dermal availability, may also be applied. Thus a weight of evidence approach in which available alternative tests are combined provides a pragmatic starting point for moving towards a non-animal based quantitative risk assessment.

Evaluation As a result of changes in legislation, *in vitro* methods for determining dermal toxicity are becoming essential for the testing of cosmetic ingredients. The pharmaceutical industry has also come to recognise the value of *in vitro* model systems to assess the safety and efficacy of drugs during development. However, the use of this type of test does involve a number of assumptions:

- Absorption and metabolism *in vitro* reflect the corresponding processes *in vivo*.
- The concentrations to which cells are exposed *in vitro* correspond to exposures *in vivo*.
- Additional cell types such as fibroblasts, melanocytes and Langerhans cells (which are not represented in cell culture systems) are not involved in toxicity.

In addition, the use of *in vitro* tests is difficult for some types of formulation: in particular, the testing of powders, gels, creams and other complex topical formulations *in vitro* can be challenging.

4.3.5 Eye

A variant of the Draize tests is conventionally (and notoriously) used to assess the ocular irritation potential of products such as shampoos; however, as discussed earlier, the Draize test is associated with significant animal welfare issues. It has also been criticised on scientific grounds because of its subjectivity, overpredictiveness and inter/intra-laboratory variability. One approach to this problem has been to develop *in vitro* systems to model effects on the eye. *In vitro* ocular irritation tests include cell-based cytotoxicity methods, reconstituted tissue models, organotypic culture methods, chorioallantoic membrane methods and isolated organ methods. Cell-based assays are effective for the study of water soluble materials but have limited use for other types of formulation such as insoluble materials and powders.

Organotypic models such as intact bovine lenses in organ culture can be used as a surrogate for the cornea. This type of system generates results which compare favourably with Draize test scores; it has the advantage of greater sensitivity, giving a wider range of scores than the Draize test and can be considered to be a Refinement in 3Rs terms.

Comparisons of the available organotypic methods reveal a variety of advantages and disadvantages specific to particular systems (Barile, 2010). For instance, the limitations of the bovine cornea are its dimensions and the thickness of the cornea, which creates difficulty in screening mild irritants. The same applies to pig cornea. In addition, due to the slaughtering process, obtaining eyes from processed cows and pigs, as well as the labour required to remove the eye, is an important issue. While these four *in vitro* test methods are not currently being considered as Replacements for the *in vivo* rabbit eye test, further optimization and validation of these methods, and other *in vitro* techniques, is necessary to reduce and refine animal use for ocular safety testing (Barile, 2010).

Corneal epithelial cells, the first exposed cells on the surface of the eye, have been used to investigate issues linked to the biological role of tissue and gene regulation.[12] In order to maximise the information obtained, *in vitro* test methods should involve exposure times that reflect actual accidental exposure, make the best possible use of existing Draize test data (even if it is variable) and include an effective prediction model. One such test is the Short Term Exposure (STE) test (Takahashi *et al*., 2008), which can be evaluated after exposures as short as 5 min. The end point for this simple assay is cell viability and it has been promoted on the basis of being a promising, easily standardised alternative test with great potential.

The basic assumptions implicit in the use of corneal cell culture models are similar to those for keratinocyte models. One consideration which must be borne in mind is that false positives can occur because of the absence of mechanical protective mechanisms such as tear rinsing and blinking. Primary corneal cultures are commonly made using rabbit corneal cells, probably as a result of the prior use of rabbits in the Draize test, and these cells do seem to be representative of most mammalian corneal cells. Adult rabbit corneas are excised, the endothelial layer is removed and the epithelium is separated from the stroma by dispase digestion and gentle scraping. The epithelial layer is trypsinised, and grown to confluency (taking 3–4 days) and used in testing. This method typically yields a homogeneous population of corneal epithelial cells with very few contaminating cell types. This use of rabbit corneal cell culture does represent a gesture towards the 3Rs, but is not a true Replacement approach because it still requires the use of animals.

The normal cellular repair mechanisms are present in corneal cell cultures. This may lead to problems if loss of contact inhibition allows migration and re-growth, thus masking toxicity, or (if temporary detachment occurs) leading to false test positives. The isolation procedure itself may induce cellular stress and the enzyme profile of the cells changes over time in culture so it is important to ensure that the endpoints measured are timely and relevant.

Example: Use of human corneal epithelial cells to study oxidative DNA damage

Human corneal epithelial cells can also be grown in culture. They do not normally proliferate *in vitro* but can be stimulated to do so. The proliferative capacity of human corneal epithelial cells is dependent on the age of the cornea donor and is a function of telomere length. Studies using human corneal epithelial cells from different donors indicate that the amount of oxidative DNA damage in human corneal epithelial cells increases with age. Artificial induction of oxidative DNA

[12] For a review see Castro-Munozledo (2008).

damage using H_2O_2 in human corneal epithelial cells from young donors leads to a reduction in the cells' proliferative capacity (Joyce *et al*., 2009), and this may be of importance *in vivo* because UV exposure and the high metabolic capacity of corneal cells may lead to oxidative stress and consequent DNA damage. The similarity in relative intensity and pattern of oxidative damage in *ex vivo* corneal samples and human corneal epithelial cells indicates that the relative level of oxidative damage does not change on culturing, supporting the use of this model for *in vitro* toxicology studies.

Monolayer cultures are unable to mimic the complex processes involved in ocular exposure and toxicology and today, engineered 3D *in vitro* culture systems facilitate the use of multi-layered, standardised tissues which adequately represent the human corneal epithelium (Cotovio *et al*., 2010).

- **EpiOcular™:** The EpiOcular™ model[13] is an *in vitro* human corneal model designed to replace the traditional animal Draize eye test. It is a 3D *in vitro* human corneal epithelium model comprising of normal human-derived epidermal keratinocytes cultured on a permeable polycarbonate membrane. The cells form a stratified, squamous multi-layered epithelium similar to that of the eye cornea. The tissue construct has an air–liquid interface and exhibits morphological and growth characteristics that mimic *in vivo* conditions. The model provides data that resolves degrees of cellular cytotoxicity as a reflection of ocular irritation, from moderate-to-very mild irritancy range through responses induced by high moderate-to-severe irritation. Preliminary evaluation suggests that the assay is applicable to both hydrophilic and hydrophobic test materials in either liquid or solid state, and can differentiate mild to moderate irritants while identifying severe irritants. In combination with an organotypic model, the *in vitro* battery could possibly be useful to resolve the full range of irritation potentials (Eskes *et al*., 2007).
- **Human corneal epithelium™**: Similar to the EpiOcular™ model, the SkinEthic™ Human Corneal Epithelium (HCE™) model[14] is a human corneal epithelium model which consists of immortalised human corneal epithelial cells cultured at the air–liquid interface on a polycarbonate substrate membrane. The resulting epithelial tissue lacks a stratum corneum and thus morphologically resembles the corneal mucosa of the human eye. *In vivo*/*in vitro* comparisons in formal validation studies sponsored by ECVAM currently indicate the usefulness of the model as a pre-screen for ocular irritation of test ingredients and raw materials. Although the method has not completed a complete formal validation study, the HCE™ model is also suitable to detect corneal repair and recovery *in vitro*.

Retinal pigment epithelial cells The key functional component of the eye is the neural retina. Healthy retinal pigmented epithelial cells are essential for the functioning and survival of the neural retina. Oxidative stress plays a key role in dysfunction of the retinal pigmented epithelium and the pathogenesis of age-related macular degeneration.

[13] http://www.mattek.com/pages/products/EpiOcular
[14] http://www.skinethic.com/HCE.asp

In vitro models for studying the retinal pigmented epithelium include monkey choroidal epithelial cells and rat retinal ganglion cells. Human retinal pigmented epithelial cell lines are also available; these include the RF6A, RGC-5 and ARPE-19 lines. The ARPE-19 cell line has been used to evaluate the effects of sublethal oxidative stress on retinal pigmented epithelial cell attachment and migration *in vitro*. The cells were resistant to the effects of H_2O_2 up to 400 μM and could be made resistant to higher concentrations (up to 1 mM) by 24-h pre-treatment with a low concentration of H_2O_2. Pre-conditioning conferred significant protection against subsequent H_2O_2-induced cell death and affected cell attachment in a dose-dependent manner. This finding may help in understanding the pathogenesis of diseases in which oxidative stress plays an important role and in determining the suitability of certain treatment strategies, in particular retinal pigmented epithelial cell transplantation, in the treatment of age-related macular degeneration.

Evaluation In terms of the selection of *in vitro* models for ocular toxicity, different types of models have different strengths and weaknesses. *In vitro* methods which include the stroma as well as viable cells have greater potential for identifying severe irritants while epithelial models are better for distinguishing between non-classified substances and mild irritants. Compounds which lie between these extremes are difficult to classify *in vitro*. Bottom-up testing is recommended if weak or no irritancy is predicted whereas a top-down approach is preferred where severe irritancy is expected (Scott *et al*., 2010).

4.4 Repeated dose toxicity

As of 2011, very few examples of *in vitro* repeat-dose exposure studies had been reported in the peer-reviewed literature, and the consensus of expert opinion was that suitable Replacements for animal tests would not be available until 2018–2020 (Adler *et al*., 2011). The experts were more optimistic about the likely availability of *in vitro* tests for hazardous properties, which they thought would be available by 2017, but they noted that this would not be sufficient for a full safety assessment. Overall, the experts agreed that the scientific basis for complete Replacement of animal testing was still not fully established and would require additional time beyond the 2013 deadline set by the 7th Amendment. In particular, the problems of predicting potency and dose–response relationships from the available alternative methods were highlighted, along with issues related to the definition of general systemic toxicity, reproductive toxicity and carcinogenicity. The following specific difficulties relating to the use of alternative methods were identified:

- Single cell-based assays cannot represent the complex interplay between different cell types.
- The role of inter-cellular mediators is difficult to model *in vitro*.
- The accuracy with which cells *in vitro* reflect the behaviour of cells *in vivo* is still doubtful.

In order to address these issues, a major collaborative programme, 'Safety Evaluation Ultimately Replacing Animal Testing' (SEURAT-1), funded under the

EU Framework Seven Programme was initiated in January 2011 (Hengstler *et al*., 2012; Vinken *et al*., 2012). The aim of this programme is to develop a long-term strategy for R&D work leading to pathway-based human safety assessments in the field of repeated dose systemic toxicity testing by establishing animal-free innovative toxicity testing methods. The objectives were as under:

- To develop highly innovative tools and methodology that can ultimately support regulatory safety assessment.
- To formulate and implement a research strategy based on generating and applying knowledge of mode of action.
- To demonstrate proof-of-principle at multiple levels – theoretical, systems and application.
- To provide the blueprint for expanding the applicability domains – chemical, toxicological and regulatory.

A review of EU Scientific Committee opinions on cosmetic ingredients indicated that the organs most commonly affected in repeated dose toxicity tests were the liver, kidney and spleen (Vinken *et al*., 2012). The SEURAT-1 programme was designed to focus on the liver, addressing three key issues in repeated dose toxicity (liver fibrosis, idiosyncratic drug-induced liver injury (DILI) and steatosis/steatohepatitis). The programme of research was based upon the identification of the so-called molecular initiating event which defines the beginning of a toxicological pathway or adverse outcome pathway. Such events may be unique or common to many pathways. As part of the characterisation of molecular initiating events, it is important to determine whether it is necessary to reach a threshold value in order for a certain pathway to be disturbed, and to distinguish between pathways which begin with an irreversible interaction and those which start with a reversible (e.g. receptor-mediated) interaction. It is assumed that repeated dose toxicity is not simply a consequence of prolonged activity of the same pathways that cause acute toxicity. The strategy adopted included preparation of stable human cell lines, development of organ simulating devices, computational modelling and estimation, and recognition of the potential contribution of innovative systems biology approaches. It involved six sub-projects whose inter-relationships are illustrated in Figure 4.4.[15]

- **SCR&Tox:** Stem cells for relevant efficient extended and normalised toxicology.
- **HeMiBio:** Hepatic Microfluidic Bioreactor.
- **DETECTIVE:** Detection of endpoints and biomarkers for repeated dose toxicity using *in vitro* systems.
- **COSMOS:** Integrated *in silico* models for the prediction of human repeated dose toxicity of cosmetics and to optimise safety.
- **NOTOX:** Predicting long-term toxic effects using computer models based on systems characterisation of organotypic cultures.
- **ToxBank:** Supporting integrated data analysis and servicing of alternative testing methods in toxicology.

[15] For details of the current status of each of these projects, see the most recent Annual Report; at the time of writing, Volume 2: Gocht and Schwarz (2012).

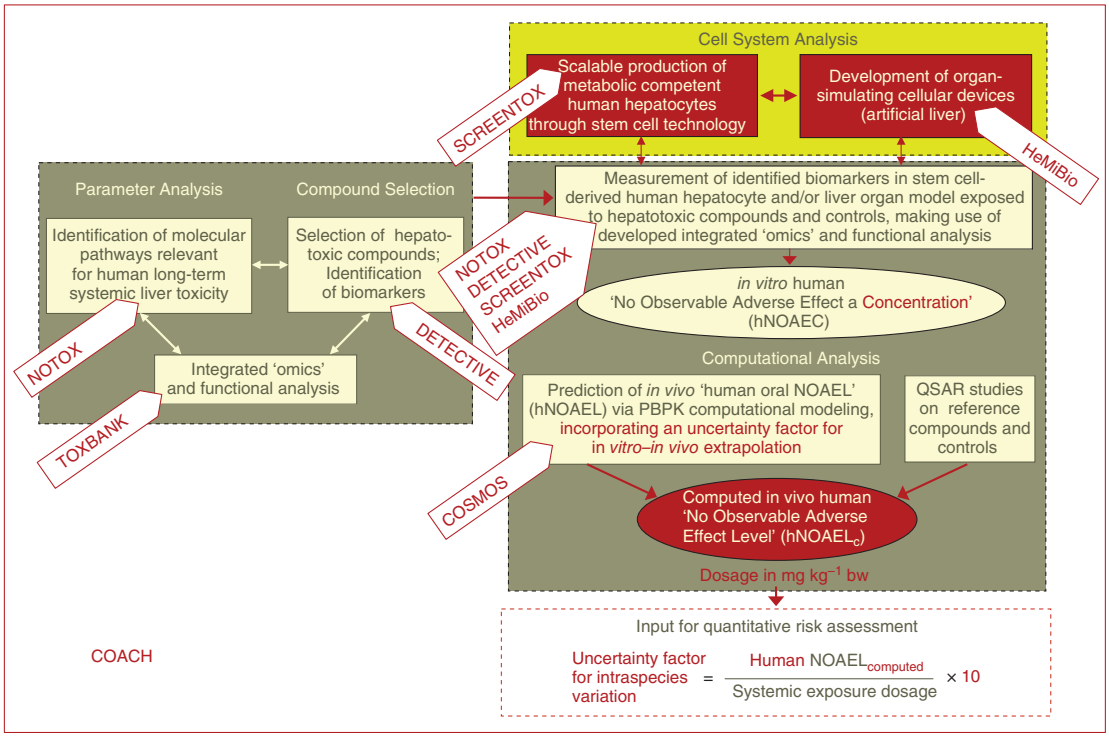


Figure 4.4 Proposal for an integrated approach for the *in vitro* assessment of repeated dose hepatotoxicity. The contributions of the different partners involved in the ongoing EU-Colipa initiative are indicated. hNOAEC, human no observable adverse effect concentration; hNOAEL, human oral NOAEL; hNOAEL$_C$, computed hNOAEL; NOAEL, no observable adverse effect level; QSAR, quantitative structure–activity relationship. (source: Vanhaecke *et al.* (2011); figure 1. With kind permission of Springer Science and Business Media)

4.5 Reproductive toxicity

The development of alternative tests for reproductive toxicity is a priority because the demands of REACH have the potential to lead to a vast number of animals being used for testing; indeed, it has been estimated that reproductive toxicity testing accounts for 60% of animal usage under REACH. In addition, reproductive toxicity testing *in vivo* is extremely costly and time consuming: the OECD guidelines specify tests covering all the critical stages of reproduction and development over two generations. The consequence is an urgent need for new and improved alternative methods. Currently the two most popular *in vitro* tests are the rat whole embryo culture assay and the embryonic stem cell test (EST) (Piersma, 2011).

The rat whole embryo culture assay makes use of rat embryos at 10–12 days postfertilisation. This is the period during which morphogenesis takes place; this includes development of the heart, neural tube closure, formation of the vertebral column and establishment of the craniofacial structures and limb buds, as well as key changes in gene expression, the timing of which is similar *in vitro* and *in vivo*. All of these processes can be monitored in culture, allowing embryotoxicity to be studied during a critical period of development and in the absence of maternal toxicity; however, it should be noted that this assay is technically demanding and involves removal of embryos from a pregnant dam so it is not strictly a non-animal method.

The EST addresses the differentiation of blastocyst-derived mouse pluripotent embryonic stem cells (ESCs) which can be induced to differentiate into a wide range of cell types including cardiomyocytes, the chosen endpoint of the assay as used at the Netherlands National Institute for Public Health and the Environment (RIVM) (Piersma, 2011). In this version of the assay (Figure 4.5) the ESCs are cultured using the hanging drop technique and assemble themselves into embryoid bodies which resemble the egg cylinder stage of a mouse embryo. These contain endodermal, ectodermal and mesodermal layers. Further differentiation is induced by cultivation in suspension culture for two days followed by plating onto tissue culture plates; on day 10 of differentiation the effects of test compounds are assessed by counting the number of contracting (i.e cardiomyocyte-like) cells. Other differentiation pathways such as those leading to neurons (Baek *et al*., 2012), osteoblasts, adipocytes and hepatocytes, can also be addressed.

The EST is unique in that it is a relatively simple, cell-based assay which incorporates every stage of differentiation. It is good at detecting strong developmental toxicants but struggles to discriminate between weak positives and negatives. The main effects of strong developmental toxicants were on cell proliferation during the first three days of culture, and one way to focus on the differentiation phase is to delay chemical exposure until after the first three days of culture.

The microscopic endpoints of the EST can be combined with transcriptional profiling, preferably 3–4 days into the protocol, or with proteomic profiling (Osman *et al*., 2010, van Dartel and Piersma, 2011). The ESC test is currently very popular although it requires further optimisation; one direction which is currently under investigation is the use of human ESCs (although there are, of course, ethical implications to this approach).

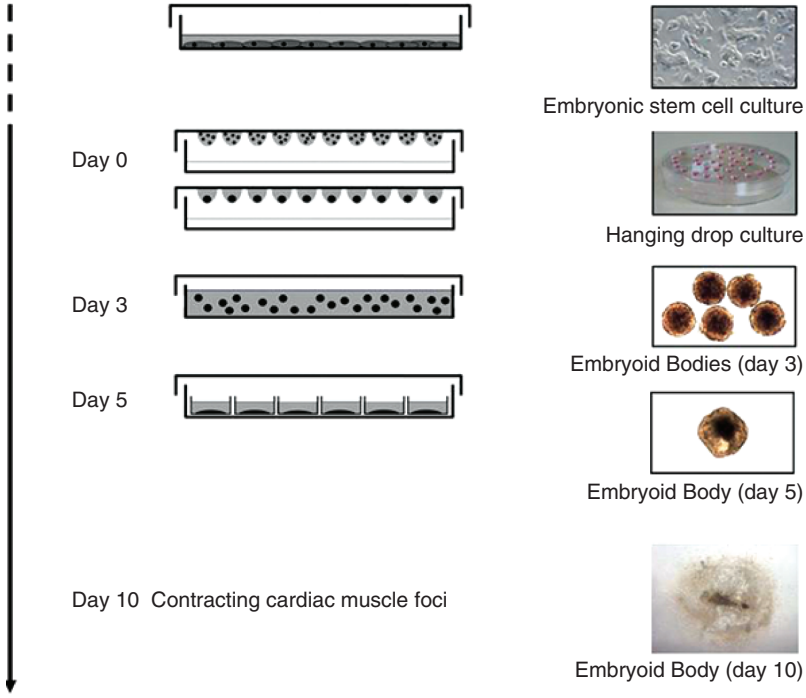


Figure 4.5 Differentiation protocol for embryonic stem cell (ESC) differentiation. Undifferentiated ESCs are cultured as hanging drops. Single-cell suspensions aggregate and form embryoid bodies (EB). EB are further cultured in suspension culture. At day 5, EB are plated on tissue culture plates. The morphology of routine culture and untreated EB at days 3, 5 and 10 is shown. The classical endpoint is microscopic evaluation of contracting foci. (source: van Dartel and Piersma (2011) figure 1. Reproduced with permission of Elsevier)

A validation study addressing three potential *in vitro* embryotoxicity tests was conducted by ECVAM between 1997 and 2000 (Genschow *et al.*, 2002). The tests were the EST, the whole-embryo culture test and the micromass test which subsequently lost popularity because of its poor predictivity and the fact that fresh embryonic material is needed each time the test is run (Piersma, 2011). Each of the three tests predicted the correct classification (non-embryotoxic, weakly embryotoxic or strongly embryotoxic) for just under 80% of chemicals and met the criteria for validation. A review of the potential for practical use of these tests (along with the newer zebrafish embryo culture assay) concluded that the EST performed at least as well as the other assays and being a pure *in vitro* test should be developed to permit a higher throughput of chemicals (van der Laan *et al.*, 2012).

The range of *in vitro* test systems available to detect effects of chemicals on mammalian fertility includes tests using follicle culture, oocyte culture, pre-implantation embryos, mature sperm, spermatagonial stem cells, Sertoli cells, seminiferous tubules, Leydig cells, pituitary cells, testis slices, co-culture of Sertoli and germ cells, endocrine–binding and transcriptional techniques (Bremer *et al.*, 2005, 2007). The ReProTect project (2004–2009) aimed to

evaluate the most promising of these methodologies. An overview of this project defined the components of reproductive toxicity which made up the different research streams (Hareng *et al*., 2005). The key areas identified were fertility, implantation and prenatal development together with a number of crosscutting technologies (*in silico*, sensor technology, array technology, biotransformation and receptor interaction). A test battery of 14 methods, focussing on endocrine disruption, fertility and embryonic development, was developed to a point where the tests could be subjected to evaluation against each other and *in vivo* results. The tests developed included a number of ligand binding and reporter assays for anti-androgenic and anti-oestrogenic effects, the mouse follicle bioassay, bovine *in vitro* maturation and *in vitro* fertilization assays, the mouse embryonic peri-implantation assay, the Ishikawa cell test, whole embryo culture, the EST and the ReProGlo assay.[16]

In the final year of the programme, a ring trial was conducted to see how well the test battery performed against historical *in vivo* results (Schenk *et al*., 2010). A panel of experts chose ten test chemicals, each of which was subjected to all 14 tests, and the results were evaluated against 30 endpoints (male and female fertility and developmental toxicity for each chemical). The test battery made the correct prediction in 23 of 30 cases and the wrong prediction in four cases. The remaining three endpoints were those for bisphenol A (male and female fertility effects and developmental toxicity); the reproductive effects of this chemical differ according to the route of administration because it undergoes rapid metabolism following oral administration and the *in vitro* tests generated ambivalent results. Overall, the conclusion of the study was that the ReProTect test battery "allowed a robust prediction of adverse effects on fertility and embryonic development of the 10 test chemicals *in vivo*" and that "The procedure used here, a nearest neighbour analysis coupled with a weight of evidence approach, may guide future activities in the field of alternative toxicity testing" (Schenk *et al*., 2010). The main problems observed were in the prediction of effects on male fertility, and the authors recognised that the lack of a reliable test for these effects was a drawback of the proposed test battery. The test battery was also unable to deal with compounds which undergo metabolic activation.

4.6 Stem cell-derived systems

The ability to provoke stem cells to differentiate into various types of adult cells *in vitro* has led to a great deal of enthusiasm for their potential use in toxicity testing. The idea is that a human stem cell line could be maintained indefinitely in culture and triggered to differentiate into target cells (e.g. hepatocytes, neurons or cardiomyocytes) as required. This would, in theory, provide an unlimited source of fresh human target cells with normal genomes and phenotypic characteristics for use in testing.

Human embryonic stem cells (hESCs) isolated from the inner cell mass of a human embryo on days 5–6 of development have, at least theoretically, the

[16] A special issue of Reproductive Toxicology detailing all these assays was published in 2010 (Vol 30 Issue 1).

potential for unlimited self-renewal while maintaining a stable genotype and phenotype and can differentiate into the majority of cell types, including hepatocyte-like cells which have been proposed as an *in vitro* model for the hazard assessment of potential chemical carcinogens (Yildirimman *et al*., 2011). However, the use of cells derived from human embryos is clearly beset by ethical problems. One way round this problem is to use so-called induced pluripotent stem cells (iPSCs), which can be derived from various non-pluripotent cell types by overexpression of key genes.[17] This approach is less ethically compromised and has the secondary advantage that a wider donor pool is available. Following transfection of key genes, including transcription factors (e.g. Oct-3/4, Sox-2) and induction enhancers (e.g. Klf4, c-myc), colonies of iPSCs develop over a period of 3-4 weeks and can be isolated by means of phenotypic and reporter gene/antibiotic selection.

Both these types of stem cell (hESCs and iPSCs) are similar in terms of growth properties, morphology, marker expression, telomerase activity and pluripotency, although the possibility of subtle differences cannot be excluded. Unfortunately both are also difficult to culture because of their tendency to grow as as microcolonies, the passaging of which requires microdissection and growth on feeder layers. In addition the critical factors for maintaining the undifferentiated state are not yet fully understood and, unless cultured very carefully, they tend to undergo undirected differentiation. Furthermore, their use in toxicity testing depends upon their ability to reflect key chronic toxicity effects such as steatosis, cholestasis and cardiomyopathy, which has not yet been convincingly demonstrated.

The same problem besets attempts to develop stem cell-derived target cells representing other tissue types; however, successfully developed models from differentiated stem cells could achieve the performance benefits of primary cell cultures from stored resources and thus eliminate dependence on scarce supplies for human tissue and/or animal sacrifice for primary animal cells. The current state of play is that cell types with the potential to be used to study neurotoxicity, cardiotoxicity, reproductive and developmental toxicity are available, but international agreement is required prior to further evaluation and validation. The expression of appropriate differentiation markers must be demonstrated, and thorough QC will be key to assuring the reliability of these systems.

The scope for this type of approach has been extended recently by the discovery of mesenchymal and neural stem cells in human umbilical cord blood, thus providing an abundant source of suitable cells for *in vitro* testing by a non-invasive method. These mesenchymal stem cells can differentiate into adipocytes, osteocytes and chondrocytes and also go on to form ectodermal and endodermal lineages. These could be used to examine how toxicants affect differentiation and to assess their effects on cellular senescence. Furthermore, the ability of neural stem cells to differentiate into astrocytes, oligodendrocytes and neurons in so-called neurosphere culture indicates the potential for developing novel methods for neurotoxicity testing and identifying compounds which can be used to treat neurodegenerative disorders (Kang and Trosko, 2011). A method for directing the differentiation of iPSCs down the retinal pigmented epithelium route has also been described (Zahabi *et al*., 2012).

Recent developments in our understanding of stem cell biology and the role of cell–cell communication have led to improved *in vitro* assays using 3D-cultures

[17] Reviewed by Drews *et al.* (2012).

of organ-specific stem cells which have the ability to create their own microenvironment. The basis of these assays is that disruption of this microenvironment, and hence the ability of the cells to control their own proliferation, differentiation and apoptosis, could give an indication of potential toxicity. Such assays are now being used experimentally to study developmental toxicity and the effects of known carcinogens. The ultimate aim of this work is to be able to use a battery of human stem cells from key organs and induce them to form 3D cultures containing most of the major cell types, along with the appropriate extracellular matrix. This approach could be extended to include co-culture systems in which stem cells from key target organs are combined with other cell types such as those of the immune system.

4.7 Conclusions

Numerous cell culture methods have been developed and used for exploring short-term effects or mechanisms of action of substances on specific target tissues and organs. Generally, however, these have not been explored as models for prediction of effects of unknown substances. There are very few examples of investigation of repeated dose exposure or modelling of chronic effects in any tissues. The development of a full set of tissue/organ specific assays, which is a key part of any integrated strategy, is considered by all reviewers to be a long-term challenge and unlikely to be realised quickly.

A non-animal risk assessment for systemic effects will require a fundamental paradigm shift, since it is generally accepted that it is impossible to simply replace a 90-day rat study with a battery of *in vitro*/*in silico* techniques. The first step will therefore be to define what a new integrated risk assessment approach would need to look like. It is envisaged that considerable investment is required in basic research to develop and validate a range of *in vitro* models capable of forming part of such an integrated approach, and the incentive provided by legislation such as the 7th Amendment is providing significant impetus in this direction.

A second important point relates to the use of cell lines as opposed to primary cultures. Since the use of primary cultures depends upon the regular sacrifice of animals they cannot be the long-term answer to Replacement. However such methods may provide indicators of the relevant markers of toxicity which are required from cell line assays and in the absence of other approaches, at this stage, may provide some useful information.

Future perspectives in *in vitro* toxicity testing include the potential development of patient-specific iPSCs for individualised testing and systematic industrial iPSC banking, permitting custom-made iPSC derived cell lines to be created on demand.

Self-assessment questions

- Under what circumstances would you use primary hepatocytes to assess a cytotoxic response, and when would a human liver tumour-derived cell line be preferable?

- What are the prospects for successful implementation of the 7th Amendment ban on animal testing of cosmetic ingredients?
- Evaluate the use of reconstructed human skin models for testing of cosmetic ingredients.
- What are the options for using alternative models for ocular toxicity testing?
- To what extent will stem cell methods replace animal tests for the routine safety assessment of chemicals?

Background reading

Davis, J.M. (ed) (2011) Animal Cell Culture. Essential Methods, Wiley-Blackwell, Chichester. ISBN: 0470666586

Freshney, R.I. (2010) Culture of Animal Cells (6th Edition) Wiley-Blackwell, Hoboken. ISBN: 0470528129.

References

Adler, S., Basketter, D., Creton, S. *et al.* (2011) Alternative (non-animal) methods for cosmetics testing: current status and future prospects 2010. *Archives of Toxicology*, **85**, 367–485.

Alepee, N., Tornier, C., Robert, C. *et al.* (2010) A catch-up validation study on reconstructed human epidermis (SkinEthic™ RHE) for full replacement of the Draize skin irritation test. *Toxicology In Vitro*, **24**, 257–266.

Andersson, T.B., Kanebratt, K.P. and Kenna, J.G. (2012) The HepaRG cell line: a unique *in vitro* tool for understanding drug metabolism and toxicology in human. *Expert Opinion on Drug Metabolism & Toxicology*, **8**, 909–920.

Antherieu, S., Chesne, C., Li, R. *et al.* (2012) Optimization of the HepaRG cell model for drug metabolism and toxicity studies. *Toxicology In Vitro*, **26**, 1278–1285.

Baek, D.H., Kim, T.G., Lim, H.K. *et al.* (2012) Embryotoxicity assessment of developmental neurotoxicants using a neuronal endpoint in the embryonic stem cell test. *Journal of Applied Toxicology*, **32**, 617–626.

Barbero, A.M. and Frasch, H.F. (2009) Pig and guinea pig skin as surrogates for human *in vitro* penetration studies: a quantitative review. *Toxicology In Vitro*, **23**, 1–13.

Barile, F.A. (2010) Validating and troubleshooting ocular *in vitro* toxicology tests. *Journal of Pharmacological and Toxicological Methods*, **61**, 136–145.

Basketter, D.A. (2012) Non-animal approaches in skin toxicology. *Archives of Toxicology*, **86**, 1159–1160.

Borlon, C., Godard, P., Eskes, C. *et al.* (2007) The usefulness of toxicogenomics for predicting acute skin irritation on *in vitro* reconstructed human epidermis. *Toxicology*, **241**, 157–166.

Bremer, S., Balduzzi, D., Cortvrindt, R. *et al.* (2005) The effects of chemicals on mammalian fertility. The report and recommendations of ECVAM Workshop 53 – the first strategic workshop of the EU ReProTect Project. *Alternatives to Laboratory Animals*, **33**, 391–416.

Bremer, S., Brittebo, E., Dencker, L. *et al.* (2007) *In vitro* tests for detecting chemicals affecting the embryo implantation process. The report and recommendations of

ECVAM workshop 62 – a strategic workshop of the EU ReProTect project. *Alternatives to Laboratory Animals*, **35**, 421–439.

Brown, C.D., Sayer, R., Windass, A.S. *et al.* (2008) Characterisation of human tubular cell monolayers as a model of proximal tubular xenobiotic handling. *Toxicology and Applied Pharmacology*, **233**, 428–438.

Cai, H., Agrawal, A.K., Putt, D.A. *et al.* (2009) Assessment of the renal toxicity of novel anti-inflammatory compounds using cynomolgus monkey and human kidney cells. *Toxicology*, **258**, 56–63.

Cannon, C.L., Neal, P.J., Southee, J.A. *et al.* (1994) New epidermal model for dermal irritancy testing. *Toxicology In Vitro*, **8**, 889–891.

Capes-Davis, A., Theodosopoulos, G., Atkin, I. *et al.* (2010) Check your cultures! A list of cross-contaminated or misidentified cell lines. *International Journal of Cancer*, **127**, 1–8.

Castro-Munozledo, F. (2008) Corneal epithelial cell cultures as a tool for research, drug screening and testing. *Experimental Eye Research*, **86**, 459–469.

Ceridono, M., Tellner, P., Bauer, D. *et al.* (2012) The 3T3 Neutral Red uptake phototoxicity test: practical experience and implications for phototoxicity testing--the report of an ECVAM-EFPIA workshop. *Regulatory Toxicology and Pharmacology*, **63**, 480–488.

Clemedson, C. (2008) The European ACuteTox project: a modern integrative *in vitro* approach to better prediction of acute toxicity. *Clinical Pharmacology and Therapeutics (St. Louis)*, **84**, 200–202.

Clemedson, C., Dierickx, P.J. and Sjostrom, M. (2003) The prediction of human acute systemic toxicity by the EDIT/MEIC *in vitro* test battery: the importance of protein binding and of partitioning into lipids. *Alternatives to Laboratory Animals*, **31**, 245–256.

Clemedson, C., Kolman, A. and Forsby, A. (2007) The integrated acute systemic toxicity project (ACuteTox) for the optimisation and validation of alternative *in vitro* tests. *Alternatives to Laboratory Animals*, **35**, 33–38.

Clemedson, C., Nordin-Andersson, M., Bjerregaard, H.F. *et al.* (2002) Development of an *in vitro* test battery for the estimation of acute human systemic toxicity: an outline of the EDIT project. Evaluation-guided Development of New *In Vitro* Test Batteries. *Alternatives to Laboratory Animals*, **30**, 313–321.

Clothier, R., Gomez-Lechon, M.J., Kinsner-Ovaskainen, A. *et al.* (2012) Comparative analysis of eight cytotoxicity assays evaluated within the ACuteTox Project. *Toxicology In Vitro* .

Coecke, S., Balls, M., Bowe, G. *et al.* (2005) Guidance on good cell culture practice. a report of the second ECVAM task force on good cell culture practice. *Alternatives to Laboratory Animals*, **33**, 261–287.

Cotovio, J., Grandidier, M.H., Lelievre, D. *et al.* (2010) *In vitro* assessment of eye irritancy using the Reconstructed Human Corneal Epithelial SkinEthic™ HCE model: application to 435 substances from consumer products industry. *Toxicology In Vitro*, **24**, 523–537.

Deshmukh, G.R., Kumar, K.H., Reddy, P.V. and Rao, B.S. (2012) *In vitro* skin corrosion: Human skin model test – a validation study. *Toxicology In Vitro*, **26**, 1072–1074.

Drews, K., Jozefczuk, J., Prigione, A. and Adjaye, J. (2012) Human induced pluripotent stem cells--from mechanisms to clinical applications. *Journal of Molecular Medicine (Berlin)*, **90**, 735–745.

Eskes, C., Cole, T., Hoffmann, S. *et al.* (2007) The ECVAM international validation study on *in vitro* tests for acute skin irritation: selection of test chemicals. *Alternatives to Laboratory Animals*, **35**, 603–619.

Fey, S.J. and Wrzesinski, K. (2012) Determination of drug toxicity using 3D spheroids constructed from an immortal human hepatocyte cell line. *Toxicological Sciences*, **127**, 403–411.

Flynn, T.J. and Ferguson, M.S. (2008) Multiendpoint mechanistic profiling of hepatotoxicants in HepG2/C3A human hepatoma cells and novel statistical approaches for development of a prediction model for acute hepatotoxicity. *Toxicology In Vitro*, **22**, 1618–1631.

Fogh, J., Wright, W.C. and Loveless, J.D. (1977) Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *Journal of the National Cancer Institute*, **58**, 209–214.

Gartler, S.M. (1967) Genetic markers as tracers in cell culture. *National Cancer Institute Monograph*, **26**, 167–195.

Gartler, S.M. (1968) Apparent Hela cell contamination of human heteroploid cell lines. *Nature*, **217**, 750–751.

Genschow, E., Spielmann, H., Scholz, G. *et al*. (2002) The ECVAM international validation study on *in vitro* embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Alternatives to Laboratory Animals*, **30**, 151–176.

Gocht, T. and Schwarz, M. (2012) Towards the Replacement of *In Vivo* Repeated Dose Systemic Toxicity Testing, European Commission and Cosmetics Europe. http://www.seurat-1.eu/pages/library/seurat-1-annual-report.php

Goebel, C., Aeby, P., Ade, N. *et al*. (2012) Guiding principles for the implementation of non-animal safety assessment approaches for cosmetics: skin sensitisation. *Regulatory Toxicology and Pharmacology*, **63**, 40–52.

Guo, L., Dial, S., Shi, L. *et al*. (2011) Similarities and differences in the expression of drug-metabolizing enzymes between human hepatic cell lines and primary human hepatocytes. *Drug Metabolism and Disposition*, **39**, 528–538.

Gustafsson, H., Runesson, J., Lundqvist, J. *et al*. (2010) Neurofunctional endpoints assessed in human neuroblastoma SH-SY5Y cells for estimation of acute systemic toxicity. *Toxicology and Applied Pharmacology*, **245**, 191–202.

Hareng, L., Pellizzer, C., Bremer, S. *et al*. (2005) The integrated project ReProTect: a novel approach in reproductive toxicity hazard assessment. *Reproductive Toxicology*, **20**, 441–452.

Hengstler, J.G., Marchan, R. and Leist, M. (2012) Highlight report: towards the replacement of *in vivo* repeated dose systemic toxicity testing. *Archives of Toxicology*, **86**, 13–15.

Hikima, T., Kaneda, N., Matsuo, K. and Tojo, K. (2012) Prediction of percutaneous absorption in human using three-dimensional human cultured epidermis LabCyte™ EPI-MODEL. *Biological and Pharmaceutical Bulletin*, **35**, 362–368.

Hoffmann, S., Kinsner-Ovaskainen, A., Prieto, P. *et al*. (2010) Acute oral toxicity: variability, reliability, relevance and interspecies comparison of rodent LD_{50} data from literature surveyed for the ACuteTox project. *Regulatory Toxicology and Pharmacology*, **58**, 395–407.

Jennen, D.G., Magkoufopoulou, C., Ketelslegers, H.B. *et al*. (2010) Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification. *Toxicological Sciences*, **115**, 66–79.

Jirova, D., Basketter, D., Liebsch, M. *et al*. (2010) Comparison of human skin irritation patch test data with *in vitro* skin irritation assays and animal data. *Contact Dermatitis*, **62**, 109–116.

Joyce, N.C., Zhu, C.C. and Harris, D.L. (2009) Relationship among oxidative stress, DNA damage, and proliferative capacity in human corneal endothelium. *Investigative Ophthalmology & Visual Science*, **50**, 2116–2122.

Kang, K.S. and Trosko, J.E. (2011) Stem cells in toxicology: fundamental biology and practical considerations. *Toxicological Sciences*, **120** (Suppl 1), S269–S289.

Katoh, M., Hamajima, F., Ogasawara, T. and Hata, K. (2009) Assessment of human epidermal model LabCyte™ EPI-MODEL for *in vitro* skin irritation testing according to European Centre for the Validation of Alternative Methods (ECVAM)-validated protocol. *Journal of Toxicological Sciences*, **34**, 327–334.

Katoh, M., Hamajima, F., Ogasawara, T. and Hata, K. (2010) Assessment of the human epidermal model LabCyte™ EPI-MODEL for *in vitro* skin corrosion testing according to the OECD Test Guideline 431. *Journal of Toxicological Sciences*, **35**, 411–417.

Kinsner-Ovaskainen, A., Prieto, P., Stanzel, S. and Kopp-Schneider, A. (2013) Selection of test methods to be included in a testing strategy to predict acute oral toxicity: an approach based on statistical analysis of data collected in phase 1 of the ACuteTox project. *Toxicology In Vitro*, **27**, 1377–94..

Kopp-Schneider, A., Prieto, P., Kinsner-Ovaskainen, A. and Stanzel, S. (2013) Design of a testing strategy using non-animal based test methods: lessons learnt from the ACuteTox project. *Toxicology In Vitro*, **27**, 1395–401.

Ku, H.O., Jeong, S.H., Kang, H.G. *et al.* (2009) Gene expression profiles and pathways in skin inflammation induced by three different sensitizers and an irritant. *Toxicology Letters*, **190**, 231–237.

Leite, S.B., Wilk-Zasadna, I., Zaldivar, J.M. *et al.* (2012) Three-dimensional HepaRG model as an attractive tool for toxicity testing. *Toxicological Sciences*, **130**, 106–116.

Li, A.P. (2009) Metabolism Comparative Cytotoxicity Assay (MCCA) and Cytotoxic Metabolic Pathway Identification Assay (CMPIA) with cryopreserved human hepatocytes for the evaluation of metabolism-based cytotoxicity *in vitro*: proof-of-concept study with aflatoxin B1. *Chemico-Biological Interactions*, **179**, 4–8.

Liguori, M.J., Blomme, E.A. and Waring, J.F. (2008) Trovafloxacin-induced gene expression changes in liver-derived *in vitro* systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metabolism and Disposition*, **36**, 223–233.

Lin, Z. and Will, Y. (2012) Evaluation of drugs with specific organ toxicities in organ-specific cell lines. *Toxicological Sciences*, **126**, 114–127.

Lucey, B.P., Nelson-Rees, W.A. and Hutchins, G.M. (2009) Henrietta Lacks, HeLa cells, and cell culture contamination. *Archives of Pathology and Laboratory Medicine*, **133**, 1463–1467.

Nardone, R.M. (2007) Eradication of cross-contaminated cell lines: a call for action. *Cell Biology and Toxicology*, **23**, 367–372.

Netzlaff, F., Lehra, C.-M., Wertzb, P.W. and Schaefer, U.F. (2005) The human epidermis models EpiSkin™, SkinEthic™ and EpiDerm™: an evaluation of morphology and their suitability for testing phototoxicity, irritancy, corrosivity, and substance transport. *European Journal of Pharmaceutics and Biopharmaceutics*, **60**, 167–178.

Niles, A.L., Moravec, R.A. and Riss, T.L. (2008) Update on *in vitro* cytotoxicity assays for drug development. *Expert Opinion Drug Discovery*, **3**, 655–669.

Nims, R.W., Sykes, G., Cottrill, K. *et al.* (2010) Short tandem repeat profiling: part of an overall strategy for reducing the frequency of cell misidentification. *In Vitro Cellular and Developmental Biology - Animal*, **46**, 811–819.

Niwa, M., Nagai, K., Oike, H. and Kobori, M. (2009) Evaluation of the skin irritation using a DNA microarray on a reconstructed human epidermal model. *Biological and Pharmaceutical Bulletin*, **32**, 203–208.

Osman, A.M., Van Dartel, D.A., Zwart, E. *et al.* (2010) Proteome profiling of mouse embryonic stem cells to define markers for cell differentiation and embryotoxicity. *Reproductive Toxicology*, **30**, 322–332.

Piersma, A. 2011. Development and implementation of alternative methods in reproductive toxicology. RIVM 340700005/2011. http://www.rivm.nl/bibliotheek/rapporten/340700005.pdf

Prieto, P., Kinsner-Ovaskainen, A., Stanzel, S. *et al.* (2012) The value of selected *in vitro* and *in silico* methods to predict acute oral toxicity in a regulatory context: results from the European Project ACuteTox. *Toxicology In Vitro*.

Rogue, A., Lambert, C., Spire, C. *et al.* (2012) Interindividual variability in gene expression profiles in human hepatocytes and comparison with HepaRG cells. *Drug Metabolism and Disposition*, **40**, 151–158.

Rosdy, M. and Clauss, L.C. (1990) Terminal epidermal differentiation of human keratinocytes grown in chemically defined medium on inert filter substrates at the air-liquid interface. *Journal of Investigative Dermatology*, **95**, 409–414.

Schenk, B., Weimer, M., Bremer, S. *et al.* (2010) The ReProTect Feasibility Study, a novel comprehensive *in vitro* approach to detect reproductive toxicants. *Reproductive Toxicology*, **30**, 200–218.

Scott, L., Eskes, C., Hoffmann, S. *et al.* (2010) A proposed eye irritation testing strategy to reduce and replace *in vivo* studies using Bottom-Up and Top-Down approaches. *Toxicology In Vitro*, **24**, 1–9.

Suzuki, H., Inoue, T., Matsushita, T. *et al.* (2008a) *In vitro* gene expression analysis of hepatotoxic drugs in rat primary hepatocytes. *Journal of Applied Toxicology*, **28**, 227–236.

Suzuki, H., Inoue, T., Matsushita, T. *et al.* (2008b) *In vitro* gene expression analysis of nephrotoxic drugs in rat primary renal cortical tubular cells. *Journal of Applied Toxicology*, **28**, 237–248.

Takahashi, Y., Koike, M., Honda, H. *et al.* (2008) Development of the short time exposure (STE) test: an *in vitro* eye irritation test using SIRC cells. *Toxicology In Vitro*, **22**, 760–770.

Tinois, E., Tiollier, J., Gaucherand, M. *et al.* (1991) *In vitro* and post-transplantation differentiation of human keratinocytes grown on the human type IV collagen film of a bilayered dermal substitute. *Experimental Cell Research*, **193**, 310–319.

Tornier, C., Amsellem, C., Fraissinette Ade, B. and Alepee, N. (2010) Assessment of the optimized SkinEthic™ Reconstructed Human Epidermis (RHE) 42 bis skin irritation protocol over 39 test substances. *Toxicology In Vitro*, **24**, 245–256.

Van Dartel, D.A. and Piersma, A.H. (2011) The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. *Reproductive Toxicology*, **32**, 235–244.

Van Der Laan, J.W., Chapin, R.E., Haenen, B. *et al.* (2012) Testing strategies for embryo-fetal toxicity of human pharmaceuticals. Animal models vs. *in vitro* approaches: a workshop report. *Regulatory Toxicology and Pharmacology*, **63**, 115–123.

Vanhaecke, T., Pauwels, M., Vinken, M., Ceelen, L., *et al.* (2011). Towards an integrated *in vitro* strategy for repeated dose toxicity testing. *Archives of Toxicology*, **85**, 365–366.

Vinken, M., Pauwels, M., Ates, G. *et al.* (2012) Screening of repeated dose toxicity data present in SCC(NF)P/SCCS safety evaluations of cosmetic ingredients. *Archives of Toxicology*, **86**, 405–412.

Wright, W.C., Daniels, W.P. and Fogh, J. (1981) Distinction of seventy-one cultured human tumor cell lines by polymorphic enzyme analysis. *Journal of the National Cancer Institute*, **66**, 239–247.

Yildirimman, R., Brolen, G., Vilardell, M. *et al*. (2011) Human embryonic stem cell derived hepatocyte-like cells as a tool for *in vitro* hazard assessment of chemical carcinogenicity. *Toxicological Sciences*, **124**, 278–290.

Zahabi, A., Shahbazi, E., Ahmadieh, H. *et al*. (2012) A new efficient protocol for directed differentiation of retinal pigmented epithelial cells from normal and retinal disease induced pluripotent stem cells. *Stem Cells and Development*, **21**, 2262–2272.

Zhang, L., Mu, X., Fu, J. and Zhou, Z. (2007) *In vitro* cytotoxicity assay with selected chemicals using human cells to predict target-organ toxicity of liver and kidney. *Toxicology In Vitro*, **21**, 734–740.

5

In Vitro Methods for Absorption, Distribution, Metabolism and Excretion

5.1 Why study ADME *in vitro*?

Four key processes determine the fate of a xenobiotic following exposure *in vivo* (Figure 5.1); these are as follows:

- Absorption (the rate and extent to which the xenobiotic gets into the systemic circulation)
- Distribution (to the organs and tissues of the body)
- Metabolism (by xenobiotic metabolising enzymes)
- Excretion (removal from the body)

The abbreviation ADME (Absorption, Distribution, Metabolism, Excretion) is used to refer to these processes as a group, and it is the balance between the different aspects of ADME which determines the pharmacokinetics or toxicokinetics of a compound *in vivo*.[1]

The impetus to develop better *in vitro* methods for ADME has arisen from the same factors which have driven developments in *in vitro* toxicology (public concern about animal experimentation, the 7th Amendment, REACH and the needs of the pharmaceutical industry). As a consequence of these developments the methodologies available for the study of ADME *in vitro* progressed dramatically during the first 10–12 years of the twenty-first century.[2]

[1] The terms *pharmacokinetics* and *toxicokinetics* are essentially interchangeable since they describe the same set of processes; it is just the final outcomes (beneficial or adverse) which differ. Since this is a toxicology textbook, the term toxicokinetics will be used for generic discussion.

[2] For a review of these developments, and the responses of different industry sectors to them, see Schroeder *et al*. (2011).

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

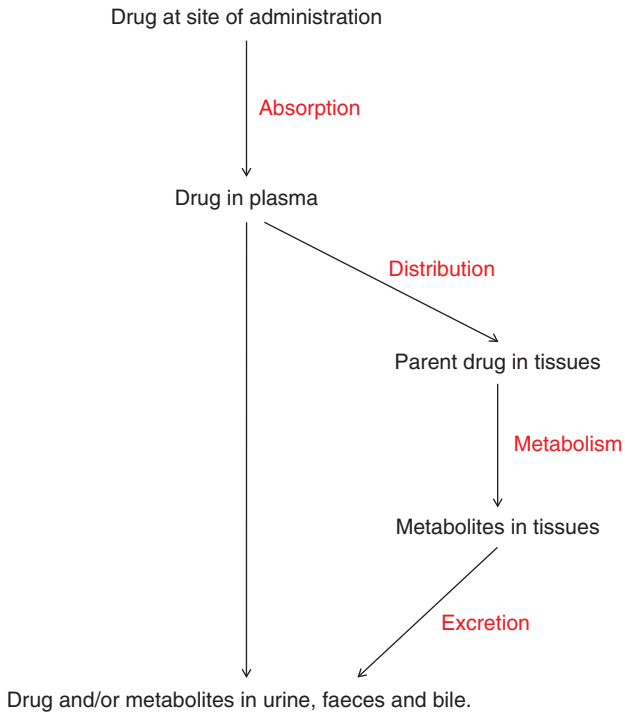


Figure 5.1 Absorption, Distribution, Metabolism and Excretion

5.2 Absorption

5.2.1 Dermal penetration

The rate limiting barrier to compound entry via the skin is the stratum corneum which can be considered to comprise an amorphous lipophilic layer and a hydrophilic layer. Since the stratum corneum is non-viable it is thought that the permeation of molecules through this layer is governed by diffusion. Following penetration across the stratum corneum, compounds diffuse across the epidermis and dermis and are carried away via the cutaneous microvasculature. The blood supply to the skin is plentiful and the large volume of blood passing through the skin acts as a sink for diffusing molecules (Yamashita and Hashida, 2003).

The systems available for experimental analysis of xenobiotic transport range from relatively simple chemical approaches to complex *in vivo* systems.[3] In this section, methods are discussed in approximate order of complexity, beginning with the chemical methods and moving on to primary and continuous *in vitro* cell culture methods. Various methods are used to evaluate the ability of compounds to cross membranes by passive diffusion (for a brief introduction to passive and active transport processes, see Box 5.1). The most popular of these for screening purposes is the Parallel Artificial Membrane Permeability Assay (PAMPA) (Avdeef and Tsinman, 2006; Di *et al*., 2003; Sugano *et al*., 2003; Li, 2005).

[3] For a detailed review, see Stanley *et al*. (2009).

Box 5.1 Transport across cell membranes

The cell membrane must allow transport, which is vital for life, including the selective transport of molecules/ions. This is problematic because many of the molecules required by the cell are very hydrophilic whereas the cell membrane is a largely hydrophobic environment. Furthermore, the access of foreign molecules to the interior of the cell must be controlled. In order to deal with this problem, cells have evolved a variety of mechanisms for facilitating and, where necessary, controlling the influx and efflux of molecules of various kinds.

Passive diffusion is the simplest form of transport across a membrane. In this process, molecules or ions move in the most thermodynamically favourable direction without the involvement of specific transport systems or molecules. In the case of uncharged species this is an entropic process and molecules always move down a concentration gradient. For charged species the process is determined by the electrochemical potential difference across the membrane and depends on the charge carried by the particle being transported.

The problem is that passive diffusion alone is too slow to sustain life. The simplest solution to this problem is facilitated diffusion. The term *facilitated diffusion* describes the process by which some ions/molecules cross membranes more rapidly than one would expect for passive diffusion. This occurs because specific proteins facilitate transport by a process which displays saturation kinetics.

Passive and facilitated diffusion processes are relatively simple in that they involve the movements of molecule down a concentration gradient. In contrast, transport against a concentration gradient requires energy input, in other words an active transport system. The most common energy input for this process is ATP hydrolysis. Active transport requires an energy coupling device in which chemical and mechanical energy are coupled. In practice, this usually involves ATP hydrolysis coupled to transport although alternative energy sources such as light or energy stored as ion gradients may also be used. Sometimes the transport of two ions or molecules is coupled: in an anti-port process two entities move in opposite directions while in a symport process two entities (e.g. oppositely charged ions) move in the same direction.

Chemical methods for measuring intestinal absorption The PAMPA assay, available commercially from pIon, Inc,[4] uses an artificial membrane made by coating highly permeable membranes with long-chain fatty acids. The system consists of a hydrophobic filter coated with a mixture of lecithin and phospholipids dissolved in an inert organic solvent such as dodecane. It is completely artificial with no pores or active transport systems. The rate of permeation across this artificial membrane barrier has been shown to correlate well with the extent of drug absorption in humans.

The PAMPA system has the advantages (less complex than biological systems, consistent, controllable, lends itself to automation and HTS) and disadvantages (oversimplifies complex processes) normally associated with such entirely

[4] http://www.pion-inc.com/

artificial systems. Its lipid composition differs from that of biological membranes (it lacks sphingomyelin) and it may not exist as a bilayer, being mounted on a sponge-like filter support.

The PAMPA assay can be varied to mimic any biological membrane of interest, including skin, and the donor and acceptor compartments can also be modified to mimic *in vivo* conditions. Overall, it is considered to capture passive transcellular diffusion across lipid barriers but does not take into account paracellular pores or active transport mechanisms mediated by influx/efflux transporters (Naik and Cucullo, 2012). In terms of the passage of compounds through the skin, the stratum corneum can be considered as a lipid layer which alternates with an aqueous medium below it, and as such it lends itself at least to preliminary investigation using non-biological systems such as PAMPA; indeed, pION has recently introduced a new version of PAMPA specifically designed to model skin permeation.[5] The transport of lipophilic chemicals through the stratum corneum is limited by the need to transfer directly into this aqueous medium, meaning that highly lipophilic compounds tend to remain in the stratum corneum. Data on skin transport indicate that there are at least two parallel pathways for diffusion through the stratum corneum, the polar or aqueous pathway and the non-polar or lipophilic pathway.

Three-dimensional human skin models Extensive histological and biochemical characterisation has demonstrated that the available RHE models (EpiSkin™, EpiDerm™ and SkinEthic™), while still requiring refinement in terms of tissue homeostasis and barrier properties, do reproduce many characteristics of normal human epidermis. Both EpiDerm™ and EpiSkin™ are suitable for skin permeation studies, providing results which agree with those obtained using *ex vivo* human skin. The method used is that described in OECD Test Guideline #428 (for percutaneous absorption by human/animal skin), but the use of RHE models for this procedure was not fully approved as an alternative method as of 2012. A multi-centre validation study reported in 2008 (Schafer-Korting *et al.*, 2008) had the aim of achieving 'a validated alternative approach, based on biotechnologically manufactured and commercially available human skin models, to be used for the development and testing of new, or not yet finally characterised, chemical entities'. The models tested were EpiDerm™ (EPI-606-X), EpiSkin™ (J13) and SkinEthic™ (RHE-L-17), which were compared with human and pig skin. A range of test materials was used, caffeine and testosterone being investigated in detail while seven other compounds with a wide range of molecular weights and lipophilicities were also evaluated. The lag times observed with the RHE models were on the order of minutes and always shorter than 2 h, whereas lag times of 0.5–2.5 h and 4.5–11 h were observed with human and pig skin, respectively. Good correlations were observed in all but two of the comparisons, and overall it was concluded that EpiDerm™, EpiSkin™ and SkinEthic™ are appropriate alternatives to human and pig skin for the assessment of skin permeation and penetration. The RHE models were able to predict rank orders of solute permeability correctly, although it should be noted that they tend to be more permeable than

[5] http://www.pion-inc.com/news/story/pion-launches-skin-pampa-test-system-modern-high-throughput-compliment-conventional-f

ex vivo human skin. Interestingly, in this study the RHE models did not generate more reproducible data than excised human skin, even though this had previously been one of the selling points for these models. The various RHE models are now beginning to be used to examine the dermal absorption of drugs and cosmetic ingredients (Hikima *et al*., 2012; Gabbanini *et al*., 2010; Christensen *et al*., 2011). Christensen *et al*. (2011) note that results obtained using these 3D models should be treated with caution when trying to evaluate topical therapeutic agents, but that the models can be valuable in making comparisons between commercial products.

With respect to percutaneous absorption, a series of *in vitro* uptake studies indicated that both EpiDerm™ and EpiSkin™ were able to predict rank orders of solute permeability correctly, but it was noted that absolute values did not necessarily agree with those obtained using *ex vivo* human skin (the RHE models had greater permeability to xenobiotics) and that small effects such as those due to changes of vehicle were less predictable using these models (Dreher *et al*., 2002a; Dreher *et al*., 2002b; Lotte *et al*., 2002). A prevalidation study comparing the uptake of the model compounds caffeine and testosterone by EpiDerm™, EpiSkin™ and SkinEthic™ (Schafer-Korting *et al*., 2006) indicated that lag times for uptake were shorter and less variable in the RHE skin models than in *ex vivo* human skin, consistent with the incomplete barrier function of these models. Similar results have been obtained using the Phenion®FT skin model (a 3D RHE model constructed from primary human foreskin keratinocytes and fibroblasts which expresses a range of dermal differentiation markers and has a functional stratum basale, as indicated by Ki67 expression[6]) in comparison with excised pig skin. Phenion®FT was found to have a barrier function weaker than that of excised pig skin but similar to those of other 3D RHE models (Ackermann *et al*., 2010). In this study, a range of lag times was observed: those with the Phenion®FT model were shorter than those with pig skin and EpiSkin™, but longer than those with EpiDerm™ and SkinEthic™.

Example: Use of EpiDerm™ to assess skin permeation by hair dye ingredients

Results obtained using the aromatic amine hair dye constituent *p*-phenylenediamine (PPD) also suggest an incomplete barrier function for the EpiDerm™ model since the extent of penetration with the EpiDerm™ tissue was higher than dermal penetration data for PPD obtained from studies in human volunteers (Hu *et al*., 2009). The total recovery of PPD-equivalents in culture media after a 30-min topical exposure in the EpiDerm™ model was 12.5% (5.32% as PPD, 1.24% as monoacetyl-PPD and 5.93% as diaectyl-PPD) when 2500 μM PPD was applied. Reported systemic exposure in humans after exposure via hair dye use is much lower. Goetz *et al*. (1988) reported that the fraction of PPD equivalents excreted in urine was 0.04–0.25% of the amount applied during hair dyeing. More recently, Nohynek *et al*. (2004b) reported that the fraction of PPD-equivalents applied in a hair dye that was excreted in the urine was 0.43%.

Evaluation Skin permeation is a multifactorial multistep process which is affected by many factors, making it very difficult to predict. In some cases, *in vitro*

[6] For further information, see http://www.phenion.com/downloads/phenion_ft-factsheet_rz_sept06.pdf.

systems can actually underpredict absorption profiles and pharmacodynamic effects following topical exposure, although most *in vitro* model systems tend to overpredict percutaneous absorption. The use of *in vitro* permeation methods can assist in the prediction of rank orders of compound uptake, but despite recent progress it is still difficult to define *in vitro–in vivo* correlations for this exposure route. This has been identified as being more challenging than the definition of *in vitro–in vivo* correlations following oral dosing. A method for achieving this objective is critically needed.

With respect to the prediction of dermal penetration, the EpiDerm™ and EpiSkin™ models are morphologically and biochemically comparable to normal human skin in terms of architecture, homeostasis and lipid composition. They are already extensively used in the cosmetics industry, being recommended by the cosmetics industry association (Cosmetics Europe)[7] for skin irritation testing. Despite some limitations they are also used for tests of skin permeation and dermal metabolism. These models represent a valid approach to the problem of characterising dermal absorption and metabolism *in vitro* and their continued use is recommended.

5.2.2 Gastrointestinal absorption

The GI tract is the main route of absorption for xenobiotics ingested with food and ingestion is second only to inhalation as the main route of exposure to toxins. Food-borne xenobiotics can be absorbed from the GI tract, entering the blood and travelling to body organs to cause toxic or beneficial effects. Oral exposure also occurs as a consequence of deliberate administration (e.g. drugs, food additives and dietary supplements). GI absorption is also relevant for toxicological studies with laboratory animals to which substances are administered in the diet or in emulsified or dissolved form by gavage.

The consequence of direct contact between the GI tract and ingested xenobiotics is that the GI tract is potentially exposed to them in high concentrations and because it has a large surface area, the GI tract is in direct contact with materials entering from the environment and can be exposed to high concentrations of xenobiotics following oral exposure. The GI tract is the major port of entry for ingested materials and is the site of absorption as well as, in some cases, metabolism of ingested chemicals. It has two major roles in the context of toxicology: to form a barrier to ingested substances and to act as a modifier of specific substances.

The route and pattern of exposure of a given xenobiotic can have a marked effect on its toxicity due to differences in pharmacokinetics and/or local effects at the site of contact.

Example: Effect of route and pattern of exposure on the toxicity of carbon tetrachloride

Carbon tetrachloride used to be widely used as a dry cleaning agent and in fire extinguishers. It was sold over the counter by retail chemists and household

[7] https://cosmeticseurope.eu/index.php

suppliers, so there has been widespread human exposure in both occupational and domestic settings over the years, primarily by inhalation but also by oral and dermal absorption.

In rodents, carbon tetrachloride is a potent hepatotoxin if administered orally, but not if administered by inhalation. The effects on male Sprague-Dawley rats of inhaled and oral doses of carbon tetrachloride have been compared in terms of differences in pharmacokinetics and hepatotoxicity (Sanzgiri *et al.*, 1995).

Absorption occurred rapidly following inhalation, high levels of carbon tetrachloride being found in the blood at the first sampling point (5 min after the start of inhalation). Near steady-state levels were reached within the first two to three sampling times and maintained over the whole 2-h exposure period. When two dose levels were compared, maximum plasma concentrations (C_{max}) and area under the plasma concentration-time curve (AUC) increased approximately proportionately (13–15-fold) with a 10-fold increase in dose.

The pattern with oral ingestion was quite different. Blood concentrations increased steadily over the 2 h of gastric infusion and then declined slowly after the end of the infusion. The increase in blood concentration with a 10-fold increase in dose was disproportionate: C_{max} and AUC increased 20–30-fold with a 10-fold increase in dose. At high doses, clearance values were similar for inhalation and ingestion routes, but at low doses AUC values were much higher for inhalation than ingestion.

When bolus gavage and infusion dosing regimens were compared, C_{max} values were much higher in the gavage groups than the infusion groups and clearance was lower in the gavage groups, although the terminal elimination of half lives in the two groups were similar. There was a corresponding difference between regimes in the target organ toxicity of carbon tetrachloride. Animals dosed by bolus gavage exhibited severe hepatic injury, probably because the sudden entry of a large bolus dose of carbon tetrachloride via the hepatic portal vein overwhelms the detoxifying capacity of the liver. Thus the liver may be susceptible to ingested carbon tetrachloride because it is the first organ exposed following absorption via the GI tract. This may not be relevant to situations where exposure occurs via inhalation (which is more likely for human exposure) and therefore the hepatotoxicity observed in rats following a large oral bolus dose may not represent the real world risks associated with human exposure to carbon tetrachloride.

Absorption is the extent to which a drug or xenobiotic is absorbed from the gut lumen into the portal circulation. Both passive and active transport processes take place in the intestine. Water soluble compounds are absorbed passively through the apical cell membrane and passage across the basolateral membrane is also mainly passive, although pore-mediated diffusion and carrier-mediated transport also occur in the basolateral membrane.

The physical factors affecting absorption include the following.

- **Blood flow to the absorption site**: The greater the blood flow, the more rapidly drugs are absorbed.
- **Total surface area available for absorption**: The intestine is covered in microvilli which increase its surface area for absorption, hence increasing the efficiency of absorption.

- **Contact time at the absorption site**: The compound has to be at the site of absorption long enough to get absorbed: this is why oral drugs are ineffective during bouts of diarrhoea.

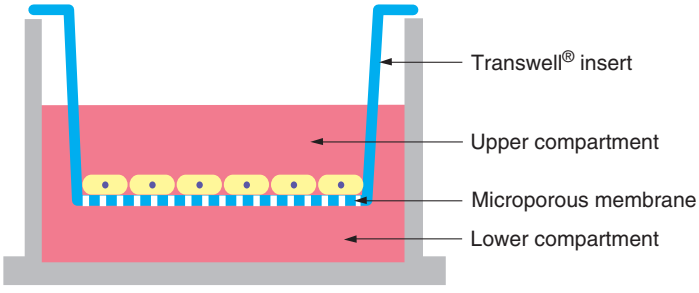
Drugs and xenobiotics may cross the intestine via both paracellular and transcellular routes (Chan *et al.*, 2004). Paracellular transport of small hydrophilic compounds (passage between adjacent cells) can occur but is limited by intercellular tight junctions. Transcellular transport (passage through cells) entails uptake across the apical membrane of the mucosal cell, transport across the cytosol and exit via the basolateral membrane.

Solubility and intestinal permeability are the most important properties that determine xenobiotic uptake following oral administration (Boobis *et al.*, 2002). Pathways for the intestinal absorption of xenobiotics include passive transcellular/paracellular absorption, facilitated diffusion, active or secondary active uptake, absorption limited by efflux transporters and intestinal metabolism followed by uptake of parent compound and/or metabolite. Passive diffusion via the paracellular and transcellular routes is the most important route for the majority of compounds, meaning that models which measure passive diffusion can be used as a preliminary screening method for the prediction of *in vivo* absorption.

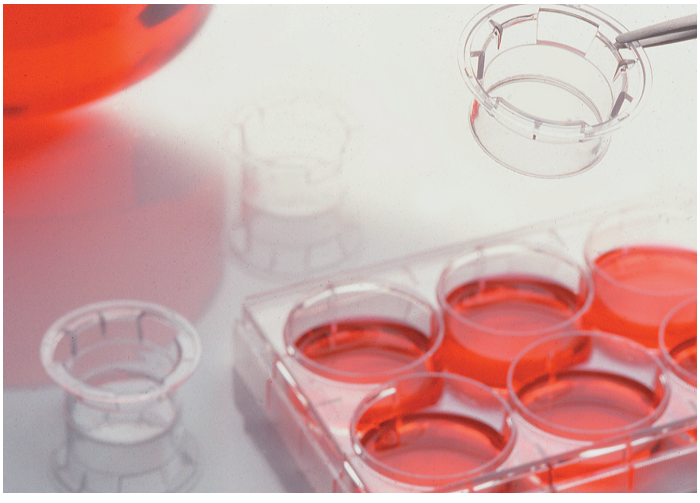
Chemical methods for measuring intestinal absorption The original PAMPA system used a proprietary lipid mixture developed for the prediction of oral/intestinal absorption and has recently been improved by using a lipid composition more similar to that of the intestinal brush border membrane, and when combined with cell-based systems for the characterisation of efflux mechanisms, this gives a sufficient understanding of the properties of drug candidates for decision making during discovery. It is resistant to a wide range of pHs and so can also be used to assess the effects of different pH conditions on permeability, from pH 1.0–2.5 (representing the pH of the stomach contents) to pH 6.6–7.0 (comparable with the pH inside the small intestine).

Cell culture models Various primary and continuous culture models have been used in attempts to simulate transport across the intestinal barrier. These have included systems comprising a single cell type as well as coculture models. In order to mimic an intact membrane, the cells have to be grown on the surface of a permeable membrane. These are usually obtained in the form of Transwell® inserts, convenient, easy-to-use permeable support devices which have been designed to produce a cell culture environment that closely resembles the *in vivo* state (Figure 5.2). The Transwell® system allows polarized cells to feed basolaterally and thereby carry out metabolic activities in a relatively natural fashion. The permeable growth supports feature a thin translucent polycarbonate membrane in a polystyrene plate treated for optimal cell attachment, thus allowing microscopic observation of the cells as they grow in order to ensure that a tight monolayer has developed prior to use in a transport assay.

The cells are grown on the surface of the membrane and once they become confluent it is necessary to verify that they have formed an intact membrane barrier. The main method used to verify barrier integrity is transepithelial electrical resistance (TEER). For the analysis of vectorial transport, the ratio of compound



(a)



(b)

Figure 5.2 The Transwell® apparatus (a) Schematic diagram of the Transwell® assembly. The porous bottom of the insert provides independent access to both sides of a cell monolayer giving researchers a versatile tool to study cell transport and other metabolic activities *in vitro*. The porous bottom of the insert provides independent access to both sides of a cell monolayer giving researchers a versatile tool to study cell transport and other metabolic activities *in vitro* (source: Images courtesy of Corning Inc.). (b) Setting up a 24-mm Transwell®-Clear Insert. The medium is added to the culture plate first; then the medium and cells are added to the Transwell® insert

flux across the membrane from the apical to basolateral (A-B) and basolateral to apical (B-A) side of the monolayer is calculated. Transport from A-B represents passive diffusion whereas transport from B-A represents active transport, for example, mediated by drug transporters.

One of the problems with the use of primary culture models is that a new culture has to be established for every experiment. This means that primary culture models tend to be labour-intensive to use and often lack the necessary reproducibility to compare results across a series of experiments. In order to meet the preference for a model which is reproducible and easy to work with, models using continuously cultured cell lines have been developed. Their advantages include

the feasibility of measuring functional transporter activity, the ability to control the experimental conditions precisely, the option of over-expressing the transporter of interest and the absence of confounding physiological processes. The most popular cell line for permeability screening is Caco-2, a human colonic adenocarcinoma cell line which undergoes spontaneous enterocytic differentiation in culture, forming polarised cells with well-established tight junctions resembling the human intestinal epithelium. Differentiated Caco-2 cells express a range of influx and efflux transporters, although at a lower level than do intestinal epithelial cells *in vivo*.

The Caco-2 assay is the most popular and extensively characterised cell-based system used to examine the uptake of xenobiotics (Balimane *et al*., 2006). It is popular for this purpose, because it is of human origin, grows as a simple monolayer with differentiated epithelial characteristics (including maintenance of microvilli) and is easy to maintain. In addition, it expresses the appropriate polarised phenotype *in vitro*. Being a single cell line, it is free of mucus and underlying submucosal and muscular tissue which could interfere with drug transport.

The permeation of drugs across Caco-2 monolayers grown on Transwell® inserts has been shown to correlate well with oral absorption in humans and this assay is considered to be the gold standard for permeability assessment during drug discovery, although it has disadvantages associated with the fact that it is cost and resource intensive, subject to experimental variability and not amenable to adaptation to HTS mode. In addition, the Caco-2 model has a number of biological limitations:

- It tends to underestimate the permeability of compounds which are absorbed via paracellular pores because the tight junctions formed by Caco-2 monolayers are tighter than those *in vivo*.
- Despite the fact that it expresses some membrane transporters, the correlation between Caco-2 permeability and *in vivo* absorption is poor for compounds which are subject to transporter-mediated uptake.
- Unmodified Caco-2 cells express relatively low levels of efflux transporters, even following treatment with inducing agents, and have little, if any drug metabolising activity (Li *et al*., 2003). Efflux transporter expression in Caco-2 cells is up-regulated during the proliferative stage of growth and down-regulated after the cells reach confluency (Goto *et al*., 2003).
- In addition, Caco-2 cells express a phenotypically variable and poorly characterised complement of other drug transporters (Taipalensuu *et al*., 2004). These additional transporters could compromise the results of studies aimed at the specific characterisation of particular transport processes.

One of the problems with Caco-2 is that this cell line is very slow growing making the Caco-2 assay time consuming and labour-intensive. In the conventional Caco-2 assay, the cells have to be cultured for 21 days prior to use in the actual assay, and this has implications in terms of cost, resource and the risk of contamination.

The Caco-2 assay has similar predictivity to the PAMPA assay but is more versatile in that it incorporates transcellular passive permeability, paracellular transport and certain aspects of transporter-mediated influx and efflux. However, the PAMPA assay is much less labour-intensive than the Caco-2 assay, being relatively

cheap and easy to set up, available in 96-well format and having the ability to be carried out in a single day. In addition, it is possible to use relatively high concentrations of solvents such as DMSO in this assay so that issues of solubility are less problematic than they would be with a cell-based assay. The PAMPA assay does, however, have a tendency to underestimate the absorption of compounds which are actively absorbed via drug transporters.

One problem with intestinal cell culture models such as Caco-2 is that they only represent the absorption phase of ADME without taking into account any of the other stages of the disposition process. A simple approach to this problem is to co-culture an intestinal cell line such as Caco-2 with a target cell type (e.g. hepatocytes). The suspended design of the Transwell® system allows for undamaged co-culturing of cells in the lower compartment so, for example, Caco-2 cells may be grown on Transwell® filters while hepatocytes are maintained in suspension in the bottom compartment. This approach has the advantage that it takes both absorption and hepatic metabolism into account, but otherwise it does not overcome the known limitations of the Caco-2 assay method. It represents a useful refinement to the Caco-2 assay but increases the complexity of the assay and is probably not necessary in a routine screening setting.

A number of other cell lines of human and animal intestinal origin have been used to study xenobiotic absorption. The human colon tumour-derived cell line HT-29 has some value in this respect since it can be induced to differentiate the forming, mucus-secreting goblet cells (which are not found in Caco-2 cultures). However, the extensive historical database and commercial availability of the Caco-2 model means that this line is still preferred for routine screening purposes. Apart from Caco-2, the cell lines most commonly used for absorption screening are the LLC-PK1 (porcine kidney) and MDCKII (Madin-Darby canine kidney epithelial cell) lines (Aszalos, 2004). These are both epithelial lines with low transporter expression and have the disadvantage that neither is of human origin.

- **LLC-PK1**: The porcine kidney cell line LLC-PK1 has been used as a model for drug transport studies for a number of years. LLC-PK1 cells have a polarised phenotype in culture and constitutively express some drug transporter activities in their own right (but these are, of course, porcine drug transporters). The presence of porcine drug transporters in LLC-PK1 cells may compromise the results of *in vitro* studies using this cell line (Yamazaki *et al*., 2001). The main disadvantage of the LLC-PK1 cell line is its non-human (and non-rodent) origin. However, this cell line is easy to work with, and, as shown by Brimer *et al*. (2000b), responds well to transduction with adenoviral vectors.
- **MDCKII**: The canine kidney cell line MDCKII has been used as a model for drug transport studies for several decades. This cell line was recently used to establish a drug transporter assay which fulfils the criteria for regulatory validation, (i.e. it is well-defined, is robust and transferable, has good dynamic range, produces interpretable data of relevance for clinicians and has reasonable throughput) (Keogh and Kunta, 2006). However, the assay was limited by toxicity and solubility and the IC_{50} values obtained were acknowledged to be system-specific. A further limitation, which is not mentioned by the authors, is that this assay uses a canine cell line. There are known to be marked species

differences between dogs and humans in both passive and active transport processes, and this could be a confounding factor in the interpretation of this assay. The MDCKII line does, however, have the advantage of forming confluent monolayers more quickly than does Caco-2 (3 days compared with 3 weeks).

Various sublines of MDCKII and LLC-PK1 have been engineered to express specific transporters for the study of molecular mechanisms of absorption, the best characterised of these derivatives being the MDCK-MDR1 line which has been engineered to express human transporters.

When considering the role of the intestine in xenobiotic disposition it is important to remember that the intestine expresses xenobiotic metabolising enzymes and can therefore contribute to the metabolism component of ADME as well as the absorption component. In an attempt to circumvent this problem, sublines of MDCKII and LLC-PK1 have been engineered for stable expression of CYP3A4. Both these cell lines gave good levels of CYP3A4 expression, the measured activities in the engineered LLC-PK1 cell line being comparable with those observed in human intestinal preparations (Brimer *et al*., 2000a). In contrast, Caco-2 cells do not express measurable levels of CYP3A4 so they are a poor model for the uptake and intestinal metabolism of CYP3A4 substrates and they do not respond well to attempts to engineer them to express CYP3A4 from an expression construct. Unfortunately, therefore, there does not seem to be scope for introducing this metabolic activation system into the Caco-2 cell line model.

Evaluation In the pharmaceutical industry, intestinal absorption is typically predicted using a combination of physicochemical and cell-based models whose strengths and limitations can be summarised as follows:

- **PAMPA**: This method is less complex than biological systems, is consistent and controllable and lends itself to automation and HTS, but it oversimplifies complex processes, quantitates only passive diffusion processes and does not address transporter function.
- **Primary culture models**: These are representative of normal cells and, if made using cell from an appropriate donor organism, high cell yields are possible. Primary cultured cells closely resemble *in vivo* cells and can either be used fresh or cryopreserved. However, this approach is time consuming and labour-intensive, considerable batch-to-batch variability is observed and it does not take normal tissue architecture or systemic processes into account.
- **Continuously cultured cell lines**: If used within an appropriate range of passage numbers and cultured with care, continuously cultured cell lines should exhibit less variability than primary cultures and can be characterised in detail because the number of cells available is theoretically unlimited. It is necessary to measure transport in both directions in order to allow for passive diffusion, which can exceed directional transport and this method does not take normal tissue architecture or systemic processes into account.
  - Caco-2 is of human origin and can be modified to express higher levels of and/or CYP3A4 but it is slow growing and expresses a phenotypically variable and poorly characterised complement of other transporters which can change due to induction/inhibition following xenobiotic treatment.

- LLC-PK1 is easy to work with and can easily be modified to express higher levels of transporter and/or CYP3A4. However, it is of non-human origin (porcine kidney) so the results obtained may be compromised by the presence of porcine transporters. In addition, it uses cells of kidney origin as a model for intestinal uptake.
- MDCKII is easy to work with and has been used to establish a validated screening assay, but again it is of non-human origin (canine kidney) so the results may be compromised by the presence of canine transporters. Like the LLC-PK1 line it uses cells of kidney origin as a model for intestinal uptake.

The PAMPA assay has been found to give a reasonable indication of intestinal absorption and generates results which are, in many cases, comparable to those obtained using the more labour-intensive Caco-2 assay. It is relatively cheap and easy to implement, can be run in HTS mode and can tolerate a wider range of pH and higher solvent concentrations than the Caco-2 assay. Thus, in order to obtain the maximum permeability and absorption information with minimal input of time and resources, a combination of the PAMPA and Caco-2 assays is becoming popular. These models are considered to offer the right balance between predictability and throughput, PAMPA being used as a HTS primary screening tool while the Caco-2 assay is used as a low throughput but more predictive secondary tool to delineate different pathways of absorption and to assess the transporter interaction potential of drug candidates (Balimane *et al.*, 2006). For passively absorbed compounds, the PAMPA and Caco-2 models show good correlation with absorption in humans. They both work well in terms of classifying compounds as having high, medium or low permeability although they exhibit wide variability in the middle of the dose response curve. Use of the two assays in combination can give an indication as to the mechanisms involved in absorption of a particular compound: for example, high PAMPA permeability but low Caco-2 permeability may indicate active efflux rather than low-intrinsic permeability, suggesting that the compound is a substrate for transporters. Such a compound would be well absorbed at high doses (as in the case of an orally administered drug) due to the saturation of transporter-mediated transport, but in the case of low exposures its uptake would be low.

It should be noted that all *in vitro* approaches to the prediction of intestinal absorption are limited by their inability to model physiological factors such as gastric emptying rate, GI transit and emptying rates, and GI pH. It should also be noted that *in vitro* absorption assays alone cannot be used to predict the time of maximal plasma concentration because these parameters are not exclusively related to absorption, being significantly affected by rates of disposition, metabolism and excretion.

5.3 Distribution

The *in vitro* prediction of tissue distribution is particularly challenging because of the involvement of systemic effects which may be determined by several different spatially distinct organs and tissues.

5.3.1 Protein binding

Plasma protein binding is a key factor in determining the pharmacokinetics of xenobiotics since the free fraction in blood is a key determinant of clearance. The methods used to characterise protein binding include micro-dialysis, equilibrium dialysis and ultra-filtration. However, it is difficult to predict biokinetic effects directly from the results of plasma protein binding assays because the relationship between plasma protein binding and biokinetics is non-linear. There is general agreement that biokinetics are not significantly affected if the proportion bound to plasma proteins is <95%, but above that value, the effects of protein binding are difficult to predict. It is difficult to discriminate experimentally between 99 and 99.9% protein binding, but this may be an important difference because this has an order of magnitude effect on the free concentration of the compound (1% vs 0.1%). In addition, most analyses assume that the unbound concentration of compound is the same in plasma/blood and inside the hepatocyte. However, the tightness of plasma protein binding may have significant consequences and it is often unclear whether plasma protein binding is so tight that processes such as permeation and hepatic metabolism are significantly restricted or whether molecules are so readily released from binding proteins that these processes are essentially unaffected. Correction for the blood/plasma unbound fraction improves *in vitro–in vivo* correlations, and in some cases it is also necessary to correct for binding to microsomal proteins.

5.3.2 Blood-brain barrier

The blood-brain barrier (BBB), which is responsible for controlling the entry of xenobiotics into the brain, comprises a monolayer of endothelial cells connected by complex tight junctions which severely limit paracellular transport and is the most important barrier between the central nervous system (CNS) and the systemic circulation. The main cellular components of the BBB are endothelial cells, pericytes and astrocytes. Astrocytes have a protective role within the BBB as well as serving as scaffolds, guiding neurons to innervate the BBB and directing blood vessels to their correct locations. It provides a fourfold defence against unwanted molecules:[8]

- A paracellular barrier formed by interendothelial tight and adherens junctions, which act as a physical barrier, forcing most molecular traffic to go by the transcellular route.
- A transcellular barrier created as a result of a low level of transcytosis and endocytosis.
- An enzymatic barrier mediated by xenobiotic metabolising and other enzymes.
- High expression of efflux and SLC transporters.

The BBB is highly specialised and differs from the intestinal epithelial barrier with respect to both passive and active absorption characteristics. In particular,

[8] For a concise, well-organised general review of *in vitro* models, see Wilhelm, Fazakas, and Krizbai (2011).

it is highly resistant to the passage of hydrophilic molecules. It also serves a carrier function, allowing essential nutrients into the brain and removing unwanted products of metabolism. In addition to these physiological characteristics, BBB endothelial cells express an array of metabolic enzyme systems and efflux transporters forming a biochemical barrier to compound uptake.

The simplest experimental models for studying transport across the BBB are artificial membrane systems such as PAMPA, which has been adapted for use as a model of the BBB by using porcine brain lipids to give a better approximation of the lipid composition of the BBB.

The key requirements for an *in vitro* cellular model of the BBB are as follows:

- Well characterised for membrane integrity with physiologically realistic cell architecture.
- Restrictive paracellular permeability, with appropriate expression of tight junctions between adjacent endothelial cells.
- *In vivo*-like asymmetrical expression of transporter molecules allowing the study of endothelial cell polarity.
- Expression of functional efflux mechanisms.
- Ability to discriminate the permeability of substances according to lipophilicity and molecular weight.
- Easy to culture, readily available and transferable.

The ability to reproduce the effects of haemodynamic and systemic/inflammatory insults on the BBB is also desirable.

In order to verify the paracellular barrier characteristics of *in vitro* models, two key methods are used: measurement of TEER and assessment of permeability to tracer substances of known molecular weight (e.g. sodium fluorescein and FITC-dextran). The TEER cut-off value for use of monolayer cultures in transport experiments is 120–130 Ohm $\times$ cm^2; above this value there is no further change in permeability.

The systems used to model the BBB include brain microvessel endothelial cell cultures of rodent, bovine, porcine and non-human primate origin.[9] Primary human brain endothelial cells have been available since the 1980s. They are marketed commercially and conserve the properties of the BBB but their preparation requires a high level of technical skill and they tend to have a very short lifespan so they are more useful for research than routine use. The cells can be passaged and cryopreserved, although their differentiated properties tend to be lost upon subculturing so cryopreservation should always be done at the earliest possible passage. Primary bovine brain endothelial cells are now commercially available, although since they need to be co-cultured with rat astrocytes they still entail significant animal use (Culot *et al.*, 2008). This model reproduces some of the complexities of the *in vivo* situation and gives results which correlate well with *in vivo* data, but it still requires scaling down (it is currently available in 24-well format) and does not really meet the needs of industry with respect to abandoning the use of animals in testing.

Immortalised endothelial cells of rodent, bovine and human origin are available. These lines, which have been transformed using viral transforming genes, are homogeneous, phenotypically stable, reproducible, relatively easy to culture

[9] Reviewed by Prieto *et al.* (2004).

and can be banked in order to overcome the problem of supply for use in HTS testing. However, they are of limited use as a model for the BBB because they lack restrictive paracellular permeability and may have undergone morphological transformation. Unfortunately relatively few continuously cultured cell lines of brain endothelial origin are available. The best characterized of these are RBE4 and GP8 (rat) and hCMEC/D3 (human). The most commonly used of these is RBE4, which expresses a range of BBB transporters and xenobiotic metabolising enzymes, but this can only be grown under static coculture conditions.

An alternative to the use of brain-derived cell lines for permeability screening is to use cell lines of other tissue origins. If these originate from tissues which express their own transporters (e.g. intestine, kidney) they may be used in their unmodified state. They may also be modified by transfecting in drug transporters of human or other species origin in order to improve their ability to mimic a particular barrier (e.g. the intestinal barrier or the BBB). Cell lines of non-cerebral origin may adopt some of the characteristics of the BBB when co-cultured with astrocytes. Cell lines which have been cultured in this manner include human umbilical vein endothelial cells (HUVEC), ECV304 (a subline of the T24 bladder carcinoma-derived cell line), LLC-PK1, MDCK II, and Caco-2. Various sublines of MDCKII and LLC-PK1 engineered to express specific transporters, such as the MDCK-MDR1 line which expresses human transporters, have also been used to model the BBB. These systems are, as yet, poorly characterised and they require validation. In particular, there is disagreement about the usefulness of Caco-2 cells cocultured with astrocytes as a model of the BBB since Caco-2 is derived from colonic epithelial tumour cells rather than from the endothelium.

Co-culturing with other cell types (e.g. glial cells, pericytes) can improve the expression of BBB marker enzymes, transporters and tight junctions by cultured endothelial cells. The addition of astrocytes leads to the formation of more stringent endothelial tight junctions and the creation of a culture which is more representative of the BBB *in situ* (Naik and Cucullo, 2012). The presence of glia and establishment of glial-endothelial cell interactions increases expression of tight junctions and transporters. It also promotes the expression of brain endothelial markers including OX-26, alkaline phosphatase, acetylcholine esterase and Na^+K^+ATPase .

If it is necessary to avoid contact between the two cell lines, this can be achieved by culturing the astrocytes on the bottom of the cell culture well and the endothelial cells on a porous support such as a Transwell® membrane. However, this is a complicated and labour-intensive procedure which can lead to increased variability in results and is not suitable for routine use. The addition of a neuronal cell line to the system can provide a target for the evaluation of neurotoxicity, but this is subject to the same caveats as for the two cell type system.

Static models of the BBB, such as those grown on Transwell® filters, offer a simplified view of the BBB and are suitable for preliminary screening experiments (Naik and Cucullo, 2012). These conventional cell culture models have their uses when trying to replicate the BBB, but they do not take into account shear stress, which can have a key influence on the behaviour of the BBB. Endothelial cells grown under static conditions tend to undergo dedifferentiation and grow in an uncontrolled manner due to the absence of growth controlling factors such as laminin. Physiological shear stress, detected by mechanosensors, is a key factor

in allowing endothelial cells to differentiate and express their barrier forming phenotype *in vitro*. Endothelial cells grown in the presence of shear stress are larger and flatter than those grown under static conditions and contain abundant endocytic vesicles, microfilaments and clathrin-coated pits.

Dynamic systems, in which the cells are cultured in the presence of flowing medium to create shear stress, can generate much higher TEER values than static systems (up to 1000 Ohm × cm^2 vs ~70 Ohm × cm^2). This may be achieved by co-culturing brain microvascular endothelial cells and astrocytes in hollow-fibre cultures and exposing them to quasi-physiological pulsatile laminar shear stress with, if desired, concomitant circulation of blood cells. The best results are obtained in models based on primary endothelial cells; the main limitation of endothelial cell lines is that they form cultures with relatively high paracellular permeability.

This type of system makes it possible to reproduce multiple functional properties and physiological responses typical of the BBB, including low permeability to K^+ and polar molecules, high TEER, negligible extravasation of proteins and expression of specialised transporters, ion channels and efflux systems. However, certain limitations do remain:

- It does not allow visualisation of the intraluminal compartment.
- It is not suitable for HTS.
- It makes it difficult to characterise the cells during and after the experiment.
- It requires a relatively large number of cells (approximately a million per experiment).

Evaluation A study organised by ECVAM compared a range of *in vitro* models for the prediction of BBB permeability in comparison with the mouse brain uptake assay. The conclusions of this study were as follows:

- The linear correlation between *in vitro* and *in vivo* assays was poor.
- The *in vitro* assays were unable to predict active influx.
- The MDCK-MDR1 cell line was best at distinguishing between active efflux and passive mechanisms.

The conclusion drawn was that factors other than passive permeability influence the distribution of compounds to the brain *in vivo*. No single *in vitro* test was recommended but the ECVAM team recommended that the possibility of using a battery of *in vitro* tests should be investigated. The need for further basic research on species differences, metabolic capacity and the characteristics of the human BBB in order to allow the BBB to be modelled adequately in a routine testing setting was noted.

Overall, *in vitro* cell-based BBB models represent a cost-effective, versatile and simple solution to the problem of predicting BBB penetration. They are, however, limited by the fact that endothelial cells tend to dedifferentiate in culture and by the need for further validation (Naik and Cucullo, 2012). Future development will be driven by the need for HTS models suitable for the pharmaceutical industry to use in screening and more realistic models for basic and translational research.

5.3.3 Other protective barriers

Blood-placenta barrier During gestation, the foetus is exposed to xenobiotics from the mother's bloodstream by placental transfer. Compounds mainly cross the placenta by simple passive diffusion although facilitated diffusion, active transfer and pinocytosis can also play a role. The factors which determine the transfer of a given molecule across the placenta are similar to those which determine transfer across other biological barriers, that is, molecular weight, acid dissociation constant (pK_a), lipophilicity and protein binding. The placental factors which affect transfer include thickness, surface area, the presence of carrier systems and the lipid and protein content of placental membranes. Two key physiological factors also affect the placental transfer of xenobiotics (Myllynen *et al.*, 2005):

- Foetal blood is more acidic than maternal blood and in the case of molecules which are subject to ionisation this can create a gradient of free unionised compound favouring transfer towards the foetus.
- Plasma protein binding is greater in the foetus than in the maternal circulation and, due to the immature state of foetal xenobiotic metabolism, the half life of xenobiotics in the foetal circulation can be longer than that in adults.

The placenta appears to have evolved a range of mechanisms in order to protect the foetus against this xenobiotic stress: for example, transporters on the maternal brush border of the syncytiotrophoblast pumps xenobiotics away from the foetus and back into the maternal circulation. However, in contrast, the placenta seems to have low biotransformation capacity compared with the liver.

Blood-testes barrier The blood-testes barrier is formed by tight intracellular junctions between adjacent Sertoli cells in the seminiferous tubule and protects the spermatocytes, spermatids and spermatozoa from blood-borne toxic substances. It also prevents the by-products of gametogenesis from entering circulation, lest they should stimulate an autoimmune reaction.

Many candidate drug compounds fail in development as a consequence of adverse effects on the testes, and there is an urgent need for suitable *in vitro* models for use in predicting potential testicular toxicity. This work is still at an early stage, but preliminary results on the development of a two chamber *in vitro* model of the blood testis barrier using 'adult' rat Sertoli cells were presented by GlaxoSmithKline at the *In vitro* Toxicology Society meeting held at the University of London in November 2012.

5.4 Metabolism

Biotransformation is a key factor in xenobiotic clearance, and still represents a poorly understood component of ADME (Boobis *et al.*, 2002). In addition, biotransformation may affect the toxicity of a xenobiotic either via conversion to a more active metabolite or by changing the pattern of biodistribution (for example, a topically applied compound may be rendered more absorbable by metabolism in the skin).

5.4.1 Skin

The ideal model for studying xenobiotic metabolism in human skin is, in theory, primary *ex vivo* human skin explants. This model has the advantage that all the necessary xenobiotic metabolising enzymes should be present; however, the main disadvantage of *ex vivo* skin is batch-to batch variability in tissue quality and handling. In addition, there is individual variability in the biological characteristics of skin, including the expression of key enzymes such as CYPs. This represents both an advantage and a disadvantage of this system: it adds to potential variability in the results obtained, but the study of a number of individual human skin samples can contribute to an understanding of individual differences in metabolism, which is an important aspect of risk assessment.

RHE models, including EpiDerm™ and EpiSkin™, have been shown to be metabolically competent (that is, they are able to metabolise xenobiotics in a way that is, at least in some respects, comparable with that of human skin). All the models tested exhibit NADPH:quinone oxido reductase activity and GST-mediated metabolism of chlorodinitrobenzene (Harris *et al*., 2002a, Harris *et al*., 2002b). In terms of CYPs, activity for ethoxyresorufin O-deethylation is below the level of detection in all the models unless induced by treatment with 3-methylcholanthrene (3-MC) (10 μM for 18 h); the EpiDerm™ model gives the most consistent results in response to induction.

Example: Biotransformation of *p*-aminophenol (PAP) and *p*-phenylenediamine

If *in vitro* RHE models are to be used to predict responses to cutaneous xenobiotic exposures they must express the multitude of xenobiotic metabolising pathways present in human skin, and for genotoxicity assays, they must appropriately detect genotoxins that require metabolism. To evaluate the dermal metabolism of aromatic amine hair dyes, the biotransformation of PAP and PPD, both primary precursors in oxidative hair dyes, has been characterised in the EpiDerm™ skin model (Hu *et al*., 2009).

The metabolism results were consistent with other *in vitro* and *in vivo* data for these compounds, which indicate that NAT1-mediated metabolism in the skin is a major route of aromatic amine detoxification following dermal exposure. Despite the known fact that RHE models such as EpiDerm™ and EpiSkin™ have higher permeability for xenobiotics than intact skin, the vast majority of a topically applied dose of PAP or PPD is excluded from skin tissue, consistent with the *in vivo* situation. Importantly, EpiDerm™ exhibited good batch-to-batch and donor-to-donor reproducibility, and biotransformation performance of the tested chemicals was comparable to the EpiSkin™ model. Additional studies to investigate the expression/activity of xenobiotic metabolising enzymes in the EpiDerm™ model compared to human skin will be important to further characterise the utility of this model.

In order to demonstrate that 3D human skin models have xenobiotic metabolism pathways representative of those in human skin, gene expression has been analysed in the EpiDerm™ *in vitro* model compared to normal full-thickness human buttock skin (Hu *et al*., 2010). This identified 139 major genes involved in Phase I and Phase II reactions for the biotransformation of xenobiotics. The analysis of the presence or absence of expression of genes encoding xenobiotic

Figure 5.3 Venn diagram comparing genes expressed in EpiDerm™ tissues vs full thickness human buttock skin. The numbers of genes identified as being expressed in at least one of the four EpiDerm™ tissues and/or in FTHBS are categorized according to the subsets of samples in which they were detected. Genes whose expression was not detected in any of the tissues tested are shown separately (source: Hu *et al*. (2010); figure 1. Reproduced with permission of Elsevier)

metabolising enzymes indicated excellent agreement between full thickness human buttock skin and EpiDerm™ tissues (Figure 5.3). Of 139 genes analysed, only five (~3.5%) genes were unique to EpiDerm™ tissue (i.e. expressed in EpiDerm™ tissue and not in full thickness human buttock skin) and only 13 genes were expressed in full thickness human buttock skin but not in EpiDerm™ tissue. Overall, 121 genes (87%) showed consistency between the two skin tissue types evaluated (83 present in both EpiDerm™ and full thickness human buttock skin and 38 present in neither tissue type).

The results of these studies indicated that the expression of xenobiotic metabolism genes in the EpiDerm™ model is very similar to that in human skin, supporting the conclusion that this is a relevant *in vitro* system for the study of cutaneous exposures, metabolism, and toxicity. The fact that the profile of xenobiotic metabolising gene expression in the EpiDerm™ model is, overall, very similar to that of human skin (Jackh *et al*., 2011; Hu *et al*., 2010; Neis *et al*., 2010; Gotz *et al*., 2012b; Gotz *et al*., 2012a) suggested that the few differences observed were unlikely to affect the outcome of generic screening assays. There may be some changes in the substrate specificity of metabolism and/or the pathways followed by specific xenobiotics that would need to be taken into consideration when undertaking detailed evaluation of the metabolism and potential cutaneous toxicity of a specific compound. For example, CYP activities in microsomes prepared from EpiDerm™ and Phenion® FT tissues seem to be very low whereas flavin-containing mono-oxygenase activities in both models are relatively high (about 10–20% of those in hepatic microsomes), suggesting that this is the primary oxidising enzyme system in these models (Jackh *et al*., 2011). Substantial cyclo-oxygenase activity has also been detected in 3D RHE and human skin microsomes (Gotz *et al*., 2012b).

These data are consistent with reports by other investigators indicating a generally high level of expression of detoxifying enzymes and low expression of enzymes involved in metabolic activation in human skin. The results provide further support for the concept that xenobiotic metabolising enzymes expressed in the skin may contribute to the role of skin as a protective barrier by detoxifying xenobiotics following topical exposure. This is an important consideration in the safety assessment of topically applied compounds.

5.4.2 Gastrointestinal tract

Chemical modification of xenobiotics in the GI tract can occur due to biotransformation, bacterial metabolism and/or acid catalysis. The GI tract is metabolically capable and expresses various xenobiotic metabolising enzymes, including monoamine oxidases, CYPs, UGTs, NATs and SULTs. The cells at the tips of the intestinal villi have 6–10 times higher CYP activities than those in the crypts. In addition, the levels at which these enzymes are expressed vary along the length of the intestine. In general, xenobiotic metabolising capacity is highest in the proximal small intestine and lower in the distal regions of the GI tract. The exception to this is CYP activity, which is present throughout the GI tract although the spectrum of isoforms expressed exhibits regional variation.

The intestine is the main route of elimination of lipophilic substances, more so than the urinary route. This process is usually an indirect one which involves uptake via the intestinal mucosa, transport to the liver, hepatic metabolism, efflux into the bile canaliculi via efflux transporters and excretion into faeces with the bile. In addition, some lipophilic compounds are taken up by intestinal bacteria and eliminated when the bacteria die and are ejected with the faeces.

Enterohepatic recirculation Enterohepatic recirculation, which involves the biliary excretion of xenobiotics followed by reabsorption in the colon, either unchanged or following bacterial deconjugation, is a retention mechanism for a variety of toxic substances.

Compounds which are subject to enterohepatic recirculation include bile salts, organometallic compounds, organochlorine compounds and numerous drugs (e.g. diazepam, testosterone and spironolactone). The effects of oral contraceptives are prolonged due to enterohepatic recirculation, and contraceptive failure may occur when antibiotics are taken concomitantly with oral contraceptives due to reduced bacterial metabolism and compromised enterohepatic recirculation. Enterohepatic recirculation may, in turn, be affected by xenobiotics either as a consequence of effects on hepatobiliary transport or via effects on intestinal reabsorption.

5.4.3 Liver

Liver is by far the most heavily used human tissue in the context of toxicology because of its important physiological functions and role in the metabolism of endogenous and exogenous compounds. The most common target organ toxicity observed during drug development is hepatotoxicity, which is usually related to drug metabolism and/or rates of hepatic clearance, and the liver is one of

the organs most commonly affected in repeated dose toxicity tests of cosmetic ingredients (Vinken *et al*., 2012). The pharmaceutical industry has embraced the *in vitro* approach to prediction of xenobiotic metabolism and clearance.

When a drug is absorbed via the GI tract it goes into the portal circulation and is delivered straight to the liver. Only after it has passed through the liver does it enter the systemic circulation, and the metabolism of the drug in the liver and intestinal wall before it even reaches the systemic circulation is called *first pass metabolism*. It is important to understand first pass metabolism because this process determines the effectiveness of orally administered drugs as well as influencing the toxicity of xenobiotics following oral exposure. Variations in hepatic metabolism can lead to variability in drug responses, and drug–drug interactions (DDIs) may happen as a result of effects on first pass metabolism. The rate and extent of first pass metabolism may be affected by a variety of factors, including genetic polymorphisms, differences in gene expression and liver disease.

For some drugs which are extensively metabolised by the liver, almost all the drug is cleared as it passes through the liver. These drugs are described as flow limited and are the ones where problems with first pass hepatic metabolism can be expected. For another group of drugs, the liver only metabolises them very inefficiently, and the hepatic extraction ratio is close to zero. This means that very little of the drug is cleared by the liver.

The amount of an oral dose which reaches the systemic circulation determines its oral bioavailability. Oral bioavailability is a function of the rates of absorption and clearance. If the same dose of drug is administered both orally and as a single intravenous (i.v.) dose, the bioavailability can be calculated from the AUCs for the two dosing routes:

$$\text{Bioavailability} = \frac{\text{AUC (oral)}}{\text{AUC (i.v.)}}$$

Assuming that the drug is absorbed efficiently, first pass clearance in the liver is a major determinant of bioavailability

$$\text{Bioavailability} = 1 - \text{hepatic extraction ratio}$$

- for flow limited drugs where the hepatic extraction ratio is approximately 1, bioavailability is approximately 0 (0%) and
- for drugs with a very low hepatic extraction ratio, bioavailability is approximately 1 (100%).

Metabolic clearance may be calculated based upon the disappearance of the parent compound or appearance of metabolites in *in vitro* incubations. It is usually expressed in terms of *in vitro* half life ($t_{1/2}$) and intrinsic clearance (CL_{INT}). CL_{INT}, which is defined as the maximum activity of the liver towards test compound in the absence of other physiological factors (hepatic blood flow, matrix binding), can be used to calculate secondary pharmacokinetic values such as hepatic clearance.[10] In order to extrapolate from *in vitro* to *in vivo* clearance it is necessary to

[10] For an explanation of the various theoretical models used, see Baranczewski, *et al*. (2006).

extrapolate from *in vitro* CL_{INT} to CL_{INT}/g of liver and hence to CL_{INT}/whole liver. Both these factors are highly variable, but generally accepted values are available in the literature.

In addition to calculating CL_{INT}, data from *in vitro* incubations can be used to determine enzyme kinetic parameters. These may give an indication of potential problems such as non-linear kinetics, which may cause difficulties in predicting dose responses *in vivo*. This may be a consequence of parallel and sequential pathways of metabolism, the effects of blood flow and plasma protein binding, extrahepatic metabolism and renal elimination or a function of actual non-linearity in enzyme kinetics: for example, the kinetics of CYP3A4 are known to deviate significantly from the Michaelis-Menten model, and this can lead to problems in extrapolating to the *in vivo* situation (Houston and Kenworthy, 2000).

CYP-mediated metabolism The use of primary human liver preparations to study drug metabolism and DDIs have become a routine part of pharmaceutical development, and the prediction of CYP-mediated metabolic pathways is particularly well advanced (MacGregor *et al*., 2001).

Human liver-derived systems

- **Precision-cut tissue slices** represent the most complex *in vitro* system which can be used routinely to assess metabolic clearance, and mimic the *in vivo* system most closely. The tissue slice model can be described as a multicellular 3D *in vitro* model which possesses the biologically relevant structural and functional features of *in vivo* tissues and in which cell–cell and cell–extracellular matrix interactions are preserved (Vickers and Fisher, 2004). If human liver is used this does represent a non-animal approach; if tissue of animal origin is used this may be considered to be a 3Rs Reduction/Refinement approach rather than a Replacement. The use of human and animal tissue slices provide a bridge between animal-derived data and the human situation.

  The key attributes of the tissue slice system are multi-cellular complexity and 3D organisation. The presence of the extracellular matrix in the tissue slice contributes to the maintenance of differentiation and provides positional information, allowing for the consideration of zonal effects. Various cell types within the slice contribute to biotransformation and the release of intercellular mediators.

  The most important factor which determines the usefulness of a tissue slice preparation is tissue quality, although the medium used for slice preparation and the culture conditions also make a contribution. In the case of the liver, tissue quality is affected by handling, fibre content and cold ischaemia time. Cold ischaemia time is defined in liver transplantation as the time from clamping of the donor aorta to the time the transplanted organ becomes incorporated into the recipient's circulatory system. It is a critical factor determining the short-term functionality of the transplant as well as the long-term survival of both the graft and its recipient (Stahl *et al*., 2008). In the experimental setting the equivalent would be the time from clamping of the donor aorta to the time a culture is established and the shorter it is, the better the functionality of the resulting preparation. Use of a dynamic organ culture system, which

alternates exposure to the medium and atmosphere, provides superior results. The slice thickness must be greater than 100 μM to ensure an adequate ratio of undamaged to damaged cells, since one or two layers of cells are always damaged during the slicing procedure. This means that even very good slice preparations do not exhibit more than 80% viability (De Graaf *et al*., 2002).

The tissue slice system can provide predictions of *in vivo* metabolism which are comparable to those obtained using isolated liver cells, but in comparison with hepatocyte-based and microsomal approaches, it is labour-intensive and not well established for the prediction of metabolic clearance or DDIs. In addition, the factors influencing the maintenance of CYP expression in slices still require elucidation and because of compound concentration gradients within the tissue slice, not all cells within the slice are able to participate in metabolism (Kedderis and Lipscomb, 2001). This limits the usefulness of the tissue slice model.

- **Freshly prepared hepatocytes** may be isolated from human or animal liver by collagenase perfusion. In the case of human liver, hepatocytes can be isolated from non-transplantable donor tissue following up to (and occasionally more than) 24-h cold preservation (Li, 2007). This method is now considered routine although it could still be further optimised in terms of consistency and yield of cells. Improved distribution mechanisms and access to information, as well as better regulation and ethical standards, will help to facilitate the use of human donor tissue, particularly liver, for routine testing (Sladowski *et al*., 2005). Isolated hepatocytes may be used in a variety of different culture/incubation systems including 3D/sandwich cultures, monolayer cultures and short-term suspension systems.
- **Hepatocyte suspensions**: It is generally agreed that isolated hepatocytes retain most of the key properties of the liver; in particular, hepatocyte suspensions predict CL_{INT} more accurately than do either tissue slices or microsomal incubations (Kedderis and Lipscomb, 2001). In comparison with primary hepatocyte cultures, suspensions give higher CL_{INT} values for high-turnover compounds, partly because uptake in the monolayer culture system is limited by surface area. However, the situation is reversed with low-turnover compounds: monolayer cultures give higher CL_{INT} values for this class of compounds, possibly due to improved enzyme stability in the cultured cells (Griffin and Houston, 2005). Cultured hepatocytes may, therefore, be a better screening system for low-turnover compounds where it is essential to obtain accurate clearance values, but in general short-term hepatocyte suspensions represent a much more convenient screening system for routine use, especially now that cryopreserved hepatocyte suspensions are readily available.
- **Primary cultured hepatocytes**: Hepatocytes can be maintained in monolayer culture for until several weeks although their use for metabolism studies is limited to a few days because dedifferentiation occurs over time and almost all their CYP activity is lost within the first 96 h in culture. The quality and differentiation status of primary hepatocyte cultures may be improved by the use of an extracellular matrix equivalent such as Matrigel® (Gross-Steinmeyer *et al*., 2005; Olsavsky *et al*., 2007; Page *et al*., 2007). Other approaches to enhance

hepatocyte differentiation *in vitro* include addition of DMSO (up to 2%) to the medium and coculture with epithelial cells derived from biliary primitive cells. The disadvantages of the use of primary cultured hepatocytes, whether in monolayer or 3D culture, are that it is costly and time-consuming and only a limited number of compounds can be tested at one time.

- **Cryopreserved hepatocytes**: Techniques for hepatocyte cryopreservation have improved significantly in recent years and this method is now considered routine as long as high-quality (transplantation grade) tissue is used (Li, 2007; Fisher *et al*., 2001). Cryopreserved human hepatocytes are commercially available from a number of companies including Invitrogen®. Cambridge Bioscience, Lonza, BD Biosciences and XenoTech; a Google search on 'human hepatocytes' will quickly reveal the identities of other companies which are currently active in this field. It is now possible to generate monolayer cultures from cryopreserved hepatocytes if damage to the cell membrane is minimised during the initial preparation, and this system can be used to evaluate CYP induction by xenobiotics as well as in short-term metabolic clearance studies (Garcia *et al*., 2003; Roymans *et al*., 2004; Kafert-Kasting *et al*., 2006; Alexandre *et al*., 2012). In comparison with freshly isolated cells, cryopreserved hepatocytes have similar CYP, UGT, GST and SULT isoform activities, and have been shown to give predictions of CL_{INT} which are comparable to those derived from sandwich cultured hepatocytes, freshly isolated cells and human liver microsomes (Griffin and Houston, 2004, Hallifax *et al*., 2005, Lau *et al*., 2002). However, cryopreserved cells exhibit more batch-to-batch variability than fresh preparations and are unsuccessful in predicting the clearance of compounds which exhibit complex kinetics.

  Overall, the cryopreservation system has both strengths and limitations; its strengths include ease of experimentation (allowing planning), the scope to repeat experiments on the same donor and the option of selecting a preferred donor, while among its limitations are limited incubation times (up to 5 h only for non-plateable cells). Furthermore, intracellular GSH levels are drastically reduced following cryopreservation and transporter activities are only preserved in a subset of batches. However, the use of cryopreserved hepatocytes has become much more routine in recent years, although opinions differ as to their value for use in the prediction of human CL_{INT} (Lau *et al*., 2002; Naritomi *et al*., 2003; Griffin and Houston, 2004). Recent developments have also allowed the cryopreservation of precision-cut liver slices using non-toxic, high-molarity vitrification solutions (de Graaf *et al*., 2007).

Human tumour-derived cell lines In general, human tumour-derived cell lines are of limited use in the characterisation of xenobiotic metabolism because they tend to have very low levels of expression of the relevant enzymes, particularly CYPs. However, an evaluation of the HepaRG cell line has suggested that it expresses levels of xenobiotic metabolising enzymes at levels comparable to those seen in pooled human hepatocyte preparations and provides similar results in studies of drug metabolism and clearance (Houston *et al*., 2012, Zanelli *et al*., 2012). It has been concluded, therefore, that this cell line may have value in the prediction of CL_{INT} as long as appropriate empirical correction factors are used.

Stem cell-derived hepatocytes Methods for inducing the differentiation of murine and human embryonic stem cells along the hepatocyte lineage have undergone marked improvement in recent years (Sartipy *et al*., 2007; Ameen *et al*., 2008; Soto-Gutierrez *et al*., 2008). It is now possible to generate cells with a strong superficial resemblance to hepatocytes, including the production of hepatocyte markers such as albumin and expression of hepatocyte-like histological appearance. However, the functionality of these cells in terms of ability to replicate the metabolic functions performed by hepatocytes *in vivo* remains questionable. In particular, the expression of higher differentiated functions such as CYP-mediated metabolism has not been adequately demonstrated. In addition, before embryonic stem cell-derived hepatocytes can be used routinely for the characterisation of xenobiotic metabolism it will be necessary to demonstrate consistent differentiation to the same endpoint on every occasion the cells are used.

There has been considerable interest in the use of hepatocytes generated by the *in vitro* differentiation of stem cells to provide a more stable system and a reliable source of non-immortalised cells with a consistent phenotype, both for regenerative medicine and for research use in hepatology and toxicology (Dalgetty *et al*., 2009). This approach has been successfully applied to mouse and human embryonic stem-cells (Kulkarni and Khanna, 2006; Hay *et al*., 2008a; Hay *et al*., 2008b; Zamule *et al*., 2011). The differentiation of human ESCs into hepatocyte-like cells can be achieved using a three-step protocol in which the cells are primed to form definitive endoderm by treatment with activin A and sodium butyrate (not required in all protocols (Cai *et al*., 2007)), induced to differentiate using DMSO and then encouraged to mature with HGF and oncostatin M (Hay *et al*., 2008b). The cells differentiate into hepatocyte-like cells and exhibit morphological and specific gene changes typical of each stage of differentiation (Hay *et al*., 2008a; Hay *et al*., 2008b). They exhibit hepatocyte-like functions such as glycogen storage and expression of immunochemically detectable fibrinogen, fibronectin, α2-macroglobulin, cytochrome P450 reductase, CYP2C9, CYP2C19, CYP2D6 and CYP3A (identified by RT-PCR as a mixture of CYP3A4 and CYP3A7). Enzymatic activities towards CYP substrates including midazolam, tolbutamide, bufuralol and phenacetin can also be detected in extracts from these cells (Hay *et al*., 2008b).

There are two problems with the use of human ESCs to generate hepatocytes *in vitro*: the number of human ESC lines available is small, and the use of cells originally derived from aborted foetuses raises serious ethical questions. Considerable effort has therefore recently been focussed on the establishment of methods for inducing adult human cells to differentiate into hepatocytes.

Human iPSCs are reprogrammed mature somatic fibroblasts which represent a pluripotent cell population able to generate all primary cell types *in vitro*, and a method has recently been reported for inducing them to differentiate along the hepatocyte lineage (Sullivan *et al*., 2010). The advantage of this approach is that the iPSCs can be obtained by relatively non-invasive methods, opening up the possibility of studying cells representing a wide range of polymorphic variants and ethnic backgrounds. The method for generating iPSCs involves infecting fibroblasts with retroviruses encoding the transcription factors Oct-4, Sox2, Klf4 and c-MYC; the iPSCs are then induced to differentiate by a stepwise process in which they are first treated with Activin A (for 3 days) and Wnt-3 (for 5 days) then

grown in the presence of DMSO and induced to mature using HGF and oncostatin M. The resulting hepatocyte-like cells express a range of hepatic markers (αFP, HNF-4, Sox-17, CXCR4 and CYP7A1), while the pluripotency marker Oct-4 is down-regulated. They also express functional CYP1A2 and CYP3A4.

In principle, an even better way to generate hepatocyte-like cells would be from unwanted human tissues such as placenta, taking advantage of the fact that adult mesenchymal stem cells can be induced to differentiate along the hepatocyte lineage. Placental mesenchymal stem cells have high proliferative capacity and the preparation of hepatocytes from this cell type would be less labour-intensive than the differentiation of iPSCs. Such a procedure would also obviate the need to use retroviruses as well as being ethically preferable to the use of hESCs. Briefly, the procedure used involves pre-culturing the cells in the presence of EGF, β-FGF and BMP-4 to inhibit cell proliferation and then exposing them to cytokines (oncostatin M, HGF) and hormones (dexamethasone, insulin) to encourage differentiation along the hepatocyte lineage (Lee *et al*., 2012). The resulting cells are hepatocyte-like in appearance and express a range of hepatocyte markers, although further characterisation and validation is required before they can be reliably used as a replacement for *bona fide* hepatocytes. This raises the point that before stem cell-derived hepatocytes can be used in toxicity assessments it is necessary to verify that appropriate differentiation has been achieved. It may be possible to persuade cultured stem cells to differentiate into something which looks very much like a hepatocyte; however, it is very difficult, if not impossible, to be sure that the resulting cell actually is a hepatocyte and will respond accordingly in every situation.

Human liver microsomes Microsomal preparations are made by centrifuging tissue homogenates at 9000–11 000 g (20 min) to remove cellular debris then at approximately 230 000 g (1 h) to precipitate the intracellular membranes. The supernatant from this step is the cytosolic fraction (containing soluble enzymes) while the pellet is the microsomal fraction. The pellet is washed by repeating the ultracentrifugation step and then re-suspended for use in assays. The microsomal fraction contains membrane bound enzymes such as CYPs and UGTs and is used in the identification of metabolites and for determining which enzyme families have the capacity to be involved in metabolism of test compounds. The supernatant from the first centrifugation step at 9000 g (S9), which contains both cytosolic and microsomal enzymes, can be used as a metabolic activation system for genotoxicity assays such as the Ames test. This system, which is the one traditionally used for measuring CYP activities, can provide results which are complementary to those from whole cell/tissue systems.

Example: Use of microsomal incubations to assess the likelihood of DDIs

One of the most common reasons for DDIs is inhibition of specific CYP isozymes. A drug may be metabolised specifically by a particular CYP, but may have the capacity to inhibit a whole range of different isozymes, thus affecting the pharmacokinetics of other therapeutic agents taken concomitantly.

DDIs often occur as a result of competitive inhibition of CYP-mediated metabolism in the liver, although they can also involve the up- or down-regulation of CYP expression. They can, in many cases, be predicted by means of simple

in vitro incubations in which the inhibitory effect of a test compound is compared with those of known inhibitors of individual CYP isoforms.

As a general rule of thumb, DDIs are considered

- possible if the IC_{50} for competitive inhibition of CYP activity is <10 μM;
- likely if the IC_{50} is <1 μM.

An alternative rule of thumb is that no DDIs are expected if the ratio of C_{max}/K_i is <0.1 (Huang and Rowland, 2012; Kosugi *et al*., 2012). This ratio is recommended as a risk index by both the FDA and EMA, although the precise parameters used by the two agencies does differ: the FDA considers K_i and C_{max} based on total concentrations whereas the EMA prefers the use of free (i.e. non-protein bound) concentrations. DDIs can also occur as a result of irreversible (suicide) inhibition, as revealed by the difference between the ICs_{50} with and without a 30-min pre-incubation step. An example of this type of analysis is illustrated in Figure 5.4.

Recombinant cells/enzymes Heterologously expressed CYPs and other enzymes are now routinely used to predict drug clearance via specific metabolic routes (Proctor *et al*., 2004) and are available from a range of commercial suppliers including Cypex and Invitrogen.[11] The systems available include mammalian, yeast, insect and bacterial cells (Fujita and Kamataki, 2002). Yeast, insect and mammalian systems have the advantage of being of eukaryotic origin but their yields tend to be low, whereas bacterial systems are cheaper, easier to use and give a high protein yield relatively quickly. One problem with these systems is that the availability of accessory proteins (e.g. cytochrome P450 reductase, cytochrome b5) may be low, although this can be overcome by co-expressing these proteins either as separate genes or as fusion proteins (Venkatakrishnan *et al*., 2000).

The use of heterologously expressed enzymes has the advantages of simplicity, the ability to evaluate the behaviour of individual enzymes in isolation and scope for the establishment of HTS assays. However, the limitation of this approach is that the contributions of different enzymes in a mixed system are not taken into account and the specific activities of heterologously expressed enzymes often differ from the equivalent activities measured in human liver microsomes. In addition, the results obtained show wide variability, partly dependent upon the choice of probe substrate (Proctor *et al*., 2004). Although the use of heterologously expressed enzymes has been recommended for the prediction of metabolism and toxicity in humans (Fujita and Kamataki, 2002), the use of these systems for direct *in vitro–in vivo* extrapolation is not well established.

Glucuronidation The majority of efforts aimed at the prediction of metabolic clearance have focussed on CYP-mediated metabolism, but it has recently become clear that glucuronidation also plays a key role in xenobiotic clearance. However, little is known regarding the relationship of *in vitro* glucuronidation to UGT-mediated xenobiotic clearance *in vivo*. Among the reasons for this are

[11] http://www.cypex.co.uk/; http://www.invitrogen.com/site/us/en/home.html.

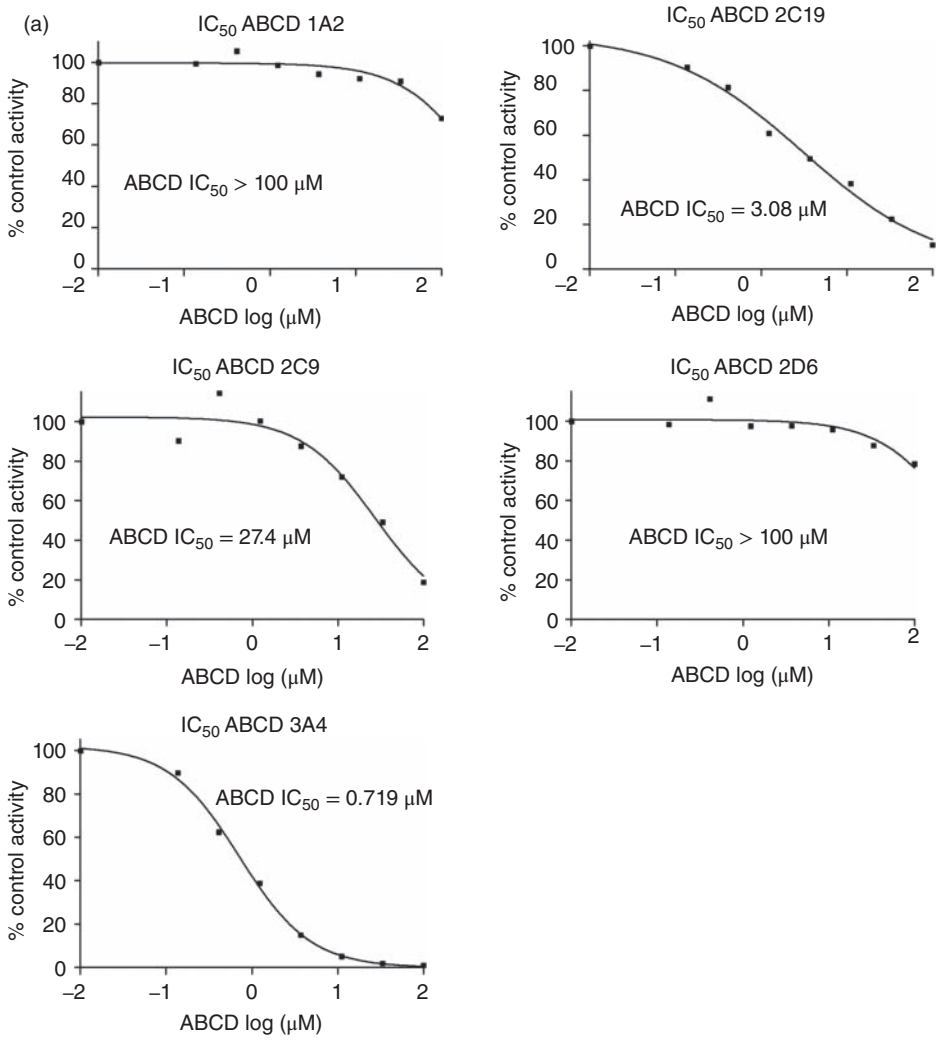


Figure 5.4 Screening for potential drug–drug interactions by *in vitro* inhibition of CYP activities. Incubation mixtures contained HEPES buffer, an NADPH regeneration system, pooled human liver microsomes (0.5 mg protein), inhibitor and substrate or substrate cocktail. Substrate cocktail final concentrations were 10 uM phenacetin, 100 uM tolbutamide, 10 uM omeprazole, 10 uM bufuralol and 10 uM midazolam. After addition of the test compound, the samples were pre-incubated for 4 min prior to addition of substrate or substrate cocktail. Following a 20-min incubation at 37 °C, the reactions were terminated with 250 ul of ice cold methanol containing 5 uM dextrorphan as an internal standard. The samples were vortexed briefly and centrifuged at approximately 13 000 × g for 5 min. The supernatant from each sample was then transferred to a deep well 96-well plate for LC-MS/MS analysis. (a) Competitive inhibition by compound ABCD: This compound is likely to cause drug–drug interactions via competitive inhibition of CYP3A4 and possibly also by inhibiting CYP2C19. (b) Suicide inhibition by compound XXXX: This figure illustrates time dependent 'suicide' (irreversible) inhibition of CYP1A2 by a test compound. Pre-incubated samples are in red, showing the shift in the IC_{50} value when the compound is pre-incubated for 30 min before the addition of the substrate when compared to a no pre-incubation step. This compound could cause drug–drug interactions as a result of suicide inhibition of CYP1A2 (source: CXR Biosciences Ltd. Reproduced with permission of CXR Biosciences Ltd)

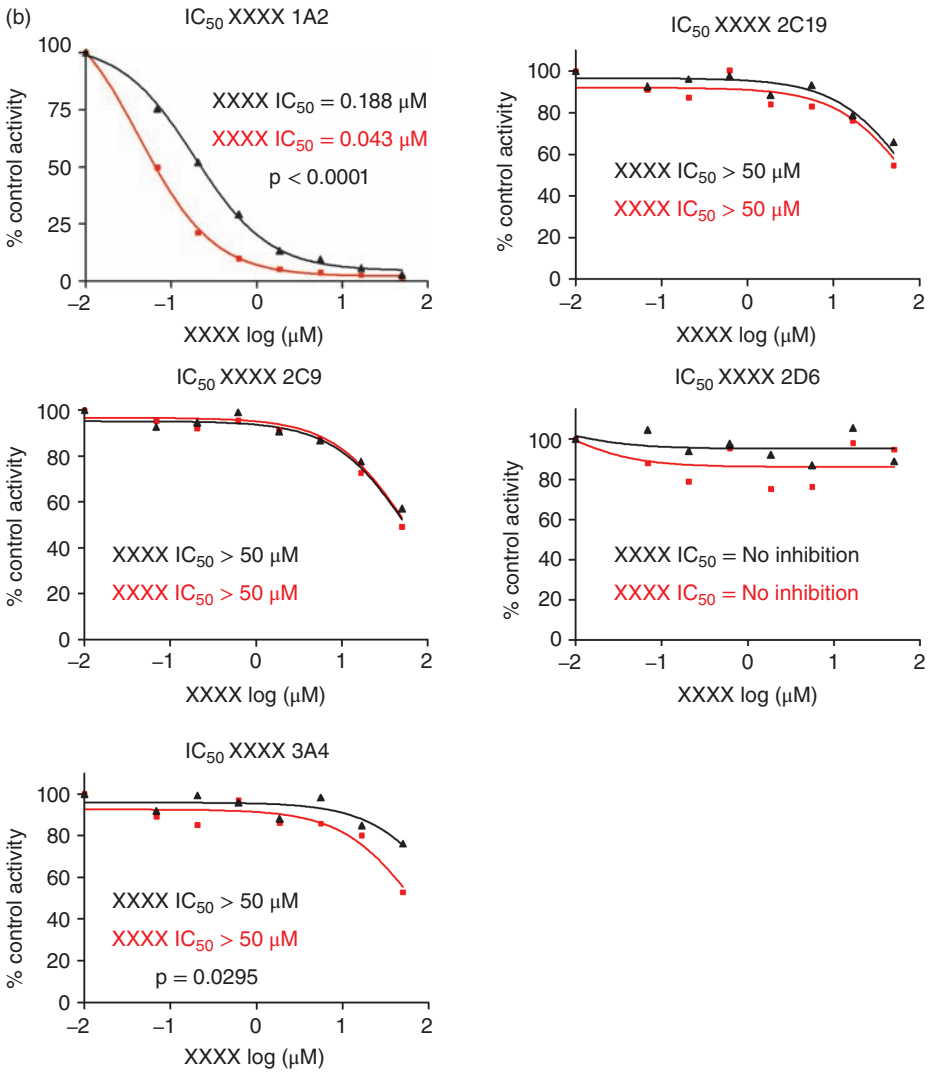


Figure 5.4 (*continued*)

the broad, overlapping substrate specificities of different UGTs, which make it difficult to identify selective probe substrates. In addition, the relative rarity of DDIs involving glucuronidation has reduced perceived urgency with regard to establishing methods for predicting glucuronidation (Miners *et al.*, 2006).

In principle, it is possible to calculate CL_{INT} via UGTs using the same techniques and assumptions as for CYPs, although additional empirical scaling factors may be required. However, the rate of glucuronidation is highly dependent upon incubation conditions (pH, cofactor concentrations, buffer type and ionic strength). Another significant problem with attempts to characterise UGT activity is that there is no universal assay to determine total UGT in a sample equivalent to the

carbon monoxide difference spectrum for CYPs. Underprediction of CL_{INT} via UGTs is a problem with both microsomal and intact hepatocyte incubation systems, although this problem is less marked in the case of intact hepatocytes. In addition, extrahepatic glucuronidation may need to be taken into account when trying to predict the clearance of UGT substrates *in vivo* (Miners *et al*., 2004).

Evaluation The pharmaceutical industry has embraced *in vitro* approaches to the prediction of xenobiotic metabolism and clearance. Such methods are now widely used for the preclinical prediction of the metabolic properties of drug candidates. However, while *in vitro* methods have been shown to be useful tools for the prediction of metabolic stability and DDI potential, there is still no rapid and secure route for extrapolating from in house or published *in vitro* data to the *in vivo* human situation. The current position is that *in vitro* methods can identify the extremes of metabolic stability and instability but approaches to quantitative *in vitro*–*in vivo* correlation still require improvement. The interpretation of *in vitro* metabolism data can be complicated and difficult because of the need to take into account many different metabolic pathways (not all of which are necessarily represented in the *in vitro* system used), the role of transporters as barriers to uptake and/or enhancers of uptake and difficulty in determining the unbound fraction of compound in plasma, blood and the *in vitro* test system used. In addition, there is evidence that the prediction of clearance using intact cells can be affected by saturable uptake into lysosomes within the hepatocyte.

Reaction phenotyping of xenobiotics metabolised by glucuronidation is currently feasible but experimentally challenging. Problems with data interpretation can arise due to the dependence of UGT activity on incubation conditions, the source of enzyme and the atypical kinetic behaviour of UGTs. *In vitro*–*in vivo* correlations may be influenced by fundamental questions relating to xenobiotic glucuronidation, including the implications of dimerisation, dependence on accessory proteins and the effects of the local membrane and cellular environment.

5.5 Excretion

Removal of drugs and other xenobiotics from the body can occur via a number of routes. The most common is via the kidney to the urine. Other routes may include bile, faeces, exhaled air, sweat, saliva and, in nursing mothers, milk. Perspiration is also an elimination route. Substances below a certain molecular size are excreted via the healthy kidney; in man, the cut-off weight for renal excretion is approximately 300. Substances which are excreted with the faeces may have entered the intestinal tract from the liver via the bile or have been excreted directly through the intestinal membrane. Volatile substances and gases can be eliminated via the lung by exhalation. The elimination of xenobiotics with breast milk can result in internal exposure of the baby.

The two major routes of drug elimination from the body are excretion as unchanged drug by the kidneys and elimination by metabolism in the liver, which plays a key role in the elimination of lipophilic compounds. The amount of drug excreted unchanged via the kidneys can be calculated by administering a dose of the drug, collecting all the urine produced over a suitable period (e.g.

24–48 h) following the dose, and measuring the amount of unchanged drug in the urine. This value can be used to calculate the fraction excreted unchanged; the amount metabolised is assumed to be (1 − fraction excreted unchanged). The fraction excreted unchanged can range from close to zero (e.g. morphine) to close to 1 (e.g. penicillin). Total body clearance of a drug is the sum of all the clearance processes in the body. Since the two main processes are renal clearance (excretion) and hepatic clearance (metabolism):

$$\text{Total clearance} = \text{Hepatic clearance} + \text{Renal clearance}$$

Renal excretion of xenobiotics is believed to be a largely passive, and therefore readily a predictable process. However, high molecular weight lipophilic molecules are often associated with significant hepatic metabolism and excretion via the biliary route. Both these processes are currently difficult to predict using *in vitro* methods.

5.5.1 Biliary excretion

The threshold molecular weight for compounds to be excreted into bile is approximately 325 in rats and 500–600 in humans. For compounds above this molecular weight, biliary excretion can significantly impact systemic exposure, biological effects and toxicity. In addition, compounds which are excreted into bile tend to undergo enterohepatic recirculation. Biliary excretion is therefore an important phenomenon, but is poorly understood in humans due to the difficulty of obtaining bile samples from healthy human subjects. The contribution of biliary excretion to the disposition of xenobiotics has often been underestimated due to a lack of methodologies to quantify the proportion of compound which undergoes enterohepatic recirculation and to poor understanding of the processes involved in hepatic excretion and reabsorption from the canalicular space, bile duct and intestine.[12]

With respect to *in vitro* methods to address biliary excretion, various approaches are available. The approach which best represents the *in vivo* situation is the use of sandwich cultured hepatocytes. These develop extensive functional bile canalicular networks, acquire and maintain normal cell polarity and allow direct access to the hepatocyte and adjacent biliary compartment. The majority of this work has been carried out using rat hepatocytes but the technique is now being applied to human hepatocytes as well (Kostrubsky *et al*., 2003). This system can be used to determine *in vitro* intrinsic biliary clearance, although it is technically demanding and labour-intensive. The use of hepatocyte suspensions is a simpler and less labour-intensive approach and has been adopted by some investigators, but its lack of 3D tissue structure means that this system is severely limited in terms of its applicability to biliary clearance studies.

An alternative approach has been to use continuously cultured cell lines such as MDCK-MDR1. This system can be used to mimic vectorial transport processes in human liver but it is limited by the fact that the cell line used is of canine kidney rather than human liver origin and by the absence of appropriate transport proteins, enzymes and cofactors. These limitations need to be taken into account

[12] Reviewed by Ghibellini *et al*., (2006).

when attempting to extrapolate to *in vivo* biliary clearance. However, this method does have the advantage of being simple to use and more readily available than primary human hepatocytes.

5.5.2 Renal clearance

Despite representing <1% of total body weight, the kidneys receive 20% of total cardiac output and are important tissues because of their key role in the excretion of xenobiotics and metabolites. The rapid rate of renal perfusion allows xenobiotics to reach high concentrations in the kidney, and this can be exacerbated by ion transport systems and renal concentration mechanisms. Precision-cut human kidney slices are becoming increasingly popular as a system for *in vitro* toxicology studies.

Passive and active tubular transport of xenobiotics from the blood to the lumen of the kidney tubule is determined by interplay between basolateral permeability, surface area of renal tubular cells, the unbound fraction in blood, blood flow and the architecture of the tubular system. Fractional tubular reabsorption is determined by brush border permeability and surface area, intra-luminal radius, flow characteristics and residence time. In addition to purely physical characteristics, permeability depends on the biochemical characteristics of the brush border membrane and the expression of drug transporters, as well as on compound concentration and pH.

The primary site of both carrier mediated transport and CYP-mediated metabolism in the kidney is in the proximal convoluted tubule. The kidney is not believed to play a major role in the total metabolism of CYP substrates as these are likely to have undergone extensive metabolism in the liver before reaching the kidney, but the kidneys do express high levels of UGTs and may participate in the metabolic clearance of glucuronidation substrates. However, the number and fraction of renal tubular cells with UGT activity and the impact of this on regional permeability within the renal tubule are unknown.

These significant data gaps mean that it is difficult to predict the renal metabolic clearance of xenobiotics (Fagerholm, 2007), and the available methods for predicting clearance via the major excretion pathways of the liver and kidney perform poorly. Given that many compounds are eliminated via these routes, especially renal elimination, this is an important omission which needs to be rectified. The recent development of methods for cryopreservation of precision-cut kidney slices using non-toxic, high-molarity vitrification solutions opens up the possibility of using this system to further characterise the role of renal metabolism and elimination in xenobiotic clearance *in vitro*. However, this method has only been developed for rat kidney slices, and still requires further optimisation even in this case, so it does not currently represent a way of avoiding the use of animals.

5.6 Conclusions

Despite considerable efforts by numerous research groups, it is still not currently possible to construct a complete picture of the ADME of a test compound based only on non-animal approaches. However, considerable progress has been made during the last 5–10 years and it is possible to identify a number of methods and

approaches which merit further evaluation. In particular, significant progress has been made in the development of methodology to assess absorption (via the GI tract and/or skin) and hepatic metabolism. There is, for example, sufficient evidence to support the view that the EpiDerm™ model is a valuable, robust *in vitro* tool to help address the need for improved non-animal tests for evaluation of efficacy, metabolism, and toxicity of cutaneous exposures. In contrast, further progress is required before the distribution and excretion processes can be predicted with confidence using *in vitro* and non-animal approaches.

Self-assessment questions

- What are the relative merits of the cell-based and artificial model systems used to predict the intestinal absorption of xenobiotics?
- Evaluate the *in vitro* systems available for characterising the hepatic metabolism of a putative CYP substrate.
- How can *in vitro* systems be used to predict the systemic clearance of a drug candidate intended for oral dosing?

References

Ackermann, K., Borgia, S.L., Korting, H.C. *et al*. (2010) The Phenion® full-thickness skin model for percutaneous absorption testing. *Skin Pharmacology and Physiology*, **23**, 105–112.

Alexandre, E., Baze, A., Parmentier, C. *et al*. (2012) Plateable cryopreserved human hepatocytes for the assessment of cytochrome P450 inducibility: experimental condition-related variables affecting their response to inducers. *Xenobiotica*, **42**, 968–979.

Ameen, C., Strehl, R., Bjorquist, P. *et al*. (2008) Human embryonic stem cells: current technologies and emerging industrial applications. *Critical Reviews in Oncology/Hematology*, **65**, 54–80.

Aszalos, A. (2004) P-glycoprotein-based drug-drug interactions: Preclinical methods and relevance to clinical observations. *Archives of Pharmacal Research*, **27**, 127–135.

Avdeef, A. and Tsinman, O. (2006) PAMPA - A drug absorption *in vitro* model: 13. Chemical selectivity due to membrane hydrogen bonding: In combo comparisons of HDM-, DOPC-, and DS-PAMPA models. *European Journal of Pharmaceutical Sciences*, **28**, 43–50.

Balimane, P.V., Han, Y.H. and Chong, S. (2006) Current industrial practices of assessing permeability and P-glycoprotein interaction. *The AAPS Journal*, **8**, E1–E13.

Baranczewski, P., Stanczak, A., Sundberg, K. *et al*. (2006) Introduction to *in vitro* estimation of metabolic stability and drug interactions of new chemical entities in drug discovery and development. *Pharmacological Reports*, **58**, 453–472.

Boobis, A., Gundert-Remy, U., Kremers, P. *et al*. (2002) *In silico* prediction of ADME and pharmacokinetics. Report of an expert meeting organised by COST B15. *European Journal of Pharmaceutical Sciences*, **17**, 183–193.

Brimer, C., Dalton, J.T., Zhu, Z. *et al*. (2000a) Creation of polarized cells coexpressing CYP3A4, NADPH cytochrome P450 reductase and MDR1/P-glycoprotein. *Pharmaceutical Research*, **17**, 803–810.

Brimer, C., Dalton, J.T., Zhu, Z.X. *et al*. (2000b) Creation of polarized cells coexpressing CYP3A4, NADPH cytochrome P450 reductase and MDR1/P-glycoprotein. *Pharmaceutical Research*, **17**, 803–810.

Cai, J., Zhao, Y., Liu, Y. *et al*. (2007) Directed differentiation of human embryonic stem cells into functional hepatic cells. *Hepatology*, **45**, 1229–1239.

Chan, L.M.S., Lowes, S. and Hirst, B.H. (2004) The ABCs of drug transport in intestine and liver: efflux proteins limiting drug absorption and bioavailability. *European Journal of Pharmaceutical Sciences*, **21**, 25–51.

Christensen, J.M., Chuong, M.C., Le, H. *et al*. (2011) Hydrocortisone diffusion through synthetic membrane, mouse skin, and EpiDerm™ cultured skin. *Archives of Drug Information*, **4**, 10–21.

Culot, M., Lundquist, S., Vanuxeem, D. *et al*. (2008) An *in vitro* blood-brain barrier model for high throughput (HTS) toxicological screening. *Toxicology In Vitro*, **22**, 799–811.

Dalgetty, D.M., Medine, C.N., Iredale, J.P. and Hay, D.C. (2009) Progress and future challenges in stem cell-derived liver technologies. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, **297**, G241–G248.

De Graaf, I.A., Draaisma, A.L., Schoeman, O. *et al*. (2007) Cryopreservation of rat precision-cut liver and kidney slices by rapid freezing and vitrification. *Cryobiology*, **54**, 1–12.

De Graaf, I.A., Van Meijeren, C.E., Pektas, F. and Koster, H.J. (2002) Comparison of *in vitro* preparations for semi-quantitative prediction of *in vivo* drug metabolism. *Drug Metabolism and Disposition*, **30**, 1129–1136.

Di, L., Kerns, E.H., Fan, K. *et al*. (2003) High throughput artificial membrane permeability assay for blood-brain barrier. *European Journal of Medicinal Chemistry*, **38**, 223–232.

Dreher, F., Fouchard, F., Patouillet, C., Andrian, M., *et al*. 2002a. Comparison of cutaneous bioavailability of cosmetic preparations containing caffeine or alpha-tocopherol applied on human skin models or human skin *ex vivo* at finite doses. *Skin Pharmacol Appl Skin Physiol*, **15** Suppl 1, 40–58.

Dreher, F., Patouillet, C., Fouchard, F., Zanini, M., *et al*. 2002b. Improvement of the experimental setup to assess cutaneous bioavailability on human skin models: dynamic protocol. *Skin Pharmacol Appl Skin Physiol*, **15** Suppl 1, 31–9.

Fagerholm, U. (2007) Prediction of human pharmacokinetics - renal metabolic and excretion clearance. *Journal of Pharmacy and Pharmacology*, **59**, 1463–1471.

Fisher, R.L., Gandolfi, A.J. and Brendel, K. (2001) Human liver quality is a dominant factor in the outcome of *in vitro* studies. *Cell Biology and Toxicology*, **17**, 179–189.

Fujita, K. and Kamataki, T. (2002) Genetically engineered bacterial cells co-expressing human cytochrome P450 with NADPH-cytochrome P450 reductase: prediction of metabolism and toxicity of drugs in humans. *Drug Metabolism and Pharmacokinetics*, **17**, 1–22.

Gabbanini, S., Matera, R., Beltramini, C. *et al*. (2010) Analysis of *in vitro* release through reconstructed human epidermis and synthetic membranes of multi-vitamins from cosmetic formulations. *Journal of Pharmaceutical and Biomedical Analysis*, **52**, 461–467.

Garcia, M., Rager, J., Wang, Q. *et al*. (2003) Cryopreserved human hepatocytes as alternative *in vitro* model for cytochrome P450 induction studies. *In Vitro Cellular and Developmental Biology - Animal*, **39**, 283–287.

Ghibellini, G., Leslie, E.M. and Brouwer, K.L. (2006) Methods to evaluate biliary excretion of drugs in humans: an updated review. *Molecular Pharmaceutics*, **3**, 198–211.

Goetz, N., Lasserre, P., Bore, P. & Kalopissis, G. 1988. Percutaneous absorption of *p*-phenylene diamine during an actual hair dyeing procedure. *Int J Cosmet Sci*, **10**, 63–73.

Goto, M., Masuda, S., Saito, H. and Inui, K. (2003) Decreased expression of P-glycoprotein during differentiation in the human intestinal cell line Caco-2. *Biochemical Pharmacology*, **66**, 163–170.

Gotz, C., Pfeiffer, R., Tigges, J. *et al*. (2012a) Xenobiotic metabolism capacities of human skin in comparison with a 3D epidermis model and keratinocyte-based cell culture as *in vitro* alternatives for chemical testing: activating enzymes (Phase I). *Experimental Dermatology*, **21**, 358–363.

Gotz, C., Pfeiffer, R., Tigges, J. *et al*. (2012b) Xenobiotic metabolism capacities of human skin in comparison with a 3D-epidermis model and keratinocyte-based cell culture as *in vitro* alternatives for chemical testing: Phase II enzymes. *Experimental Dermatology*, **21**, 364–369.

Griffin, S.J. and Houston, J.B. (2004) Comparison of fresh and cryopreserved rat hepatocyte suspensions for the prediction of *in vitro* intrinsic clearance. *Drug Metabolism and Disposition*, **32**, 552–558.

Griffin, S.J. and Houston, J.B. (2005) Prediction of *in vitro* intrinsic clearance from hepatocytes: comparison of suspensions and monolayer cultures. *Drug Metabolism and Disposition*, **33**, 115–120.

Gross-Steinmeyer, K., Stapleton, P.L., Tracy, J.H. *et al*. (2005) Influence of Matrigel®-overlay on constitutive and inducible expression of nine genes encoding drug-metabolizing enzymes in primary human hepatocytes. *Xenobiotica*, **35**, 419–438.

Hallifax, D., Rawden, H.C., Hakooz, N. and Houston, J.B. (2005) Prediction of metabolic clearance using cryopreserved human hepatocytes: kinetic characteristics for five benzodiazepines. *Drug Metabolism and Disposition*, **33**, 1852–1858.

Harris, I.R., Siefken, W., Beck-Oldach, K. *et al*. (2002a) Comparison of activities dependent on glutathione *S*-transferase and cytochrome P-450 IA1 in cultured keratinocytes and reconstructed epidermal models. *Skin Pharmacology and Applied Skin Physiology*, **15** (Suppl 1), 59–67.

Harris, I.R., Siefken, W., Beck-Oldach, K. *et al*. (2002b) NAD(P)H:quinone reductase activity in human epidermal keratinocytes and reconstructed epidermal models. *Skin Pharmacology and Applied Skin Physiology*, **15** (Suppl 1), 68–73.

Hay, D.C., Fletcher, J., Payne, C. *et al*. (2008a) Highly efficient differentiation of hESCs to functional hepatic endoderm requires ActivinA and Wnt3a signaling. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 12301–12306.

Hay, D.C., Zhao, D., Fletcher, J. *et al*. (2008b) Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development *in vivo*. *Stem Cells*, **26**, 894–902.

Hikima, T., Kaneda, N., Matsuo, K. and Tojo, K. (2012) Prediction of percutaneous absorption in human using three-dimensional human cultured epidermis LabCyte™ EPI-MODEL. *Biological and Pharmaceutical Bulletin*, **35**, 362–368.

Houston, J.B. and Kenworthy, K.E. (2000) *In vitro-in vivo* scaling of CYP kinetic data not consistent with the classical Michaelis-Menten model. *Drug Metabolism and Disposition*, **28**, 246–254.

Houston, J.P., Kohler, J., Bishop, J.R. *et al*. (2012) Pharmacogenomic associations with weight gain in olanzapine treatment of patients without schizophrenia. *Journal of Clinical Psychiatry*, **73**, 1077–1086.

Hu, T., Bailey, R.E., Morrall, S.W. *et al*. (2009) Dermal penetration and metabolism of *p*-aminophenol and *p*-phenylenediamine: application of the EpiDerm™ human reconstructed epidermis model. *Toxicology Letters*, **188**, 119–129.

Hu, T., Khambatta, Z.S., Hayden, P.J. *et al*. (2010) Xenobiotic metabolism gene expression in the EpiDerm™ *in vitro* 3D human epidermis model compared to human skin. *Toxicology In Vitro*, **24**, 1450–1463.

Huang, S.M. and Rowland, M. (2012) The role of physiologically based pharmacokinetic modeling in regulatory review. *Clinical Pharmacology and Therapeutics*, **91**, 542–549.

Jackh, C., Blatz, V., Fabian, E. *et al*. (2011) Characterization of enzyme activities of cytochrome P450 enzymes, flavin-dependent monooxygenases, *N*-acetyltransferases and UDP-glucuronyltransferases in human reconstructed epidermis and full-thickness skin models. *Toxicology In Vitro*, **25**, 1209–1214.

Kafert-Kasting, S., Alexandrova, K., Barthold, M. *et al*. (2006) Enzyme induction in cryopreserved human hepatocyte cultures. *Toxicology*, **220**, 117–125.

Kedderis, G.L. and Lipscomb, J.C. (2001) Application of *in vitro* biotransformation data and pharmacokinetic modeling to risk assessment. *Toxicology and Industrial Health*, **17**, 315–321.

Keogh, J.P. and Kunta, J.R. (2006) Development, validation and utility of an *in vitro* technique for assessment of potential clinical drug-drug interactions involving P-glycoprotein. *European Journal of Pharmaceutical Sciences*, **27**, 543–554.

Kostrubsky, V.E., Strom, S.C., Hanson, J. *et al*. (2003) Evaluation of hepatotoxic potential of drugs by inhibition of bile-acid transport in cultured primary human hepatocytes and intact rats. *Toxicological Sciences*, **76**, 220–228.

Kosugi, Y., Hirabayashi, H., Igari, T. *et al*. (2012) Evaluation of cytochrome P450-mediated drug-drug interactions based on the strategies recommended by regulatory authorities. *Xenobiotica*, **42**, 127–138.

Kulkarni, J.S. and Khanna, A. (2006) Functional hepatocyte-like cells derived from mouse embryonic stem cells: a novel *in vitro* hepatotoxicity model for drug screening. *Toxicology In Vitro*, **20**, 1014–1022.

Lau, Y.Y., Sapidou, E., Cui, X. *et al*. (2002) Development of a novel *in vitro* model to predict hepatic clearance using fresh, cryopreserved, and sandwich-cultured hepatocytes. *Drug Metabolism and Disposition*, **30**, 1446–1454.

Lee, H.J., Jung, J., Cho, K.J. *et al*. (2012) Comparison of *in vitro* hepatogenic differentiation potential between various placenta-derived stem cells and other adult stem cells as an alternative source of functional hepatocytes. *Differentiation*, **84**, 223–231.

Li, A.P. (2005) Preclinical *in vitro* screening assays for drug-like properties. *Drug Discovery Today: Technologies*, **2**, 179–185.

Li, A.P. (2007) Human hepatocytes: isolation, cryopreservation and applications in drug development. *Chemico-Biological Interactions*, **168**, 16–29.

Li, Q., Sai, Y., Kato, Y. *et al*. (2003) Influence of drugs and nutrients on transporter gene expression levels in Caco-2 and LS180 intestinal epithelial cell lines. *Pharmaceutical Research*, **20**, 1119–1124.

Lotte, C., Patouillet, C., Zanini, M., Messager, A., *et al*. 2002. Permeation and skin absorption: reproducibility of various industrial reconstructed human skin models. *Skin Pharmacol Appl Skin Physiol*, **15** Suppl 1, 18–30.

MacGregor, J.T., Collins, J.M., Sugiyama, Y. *et al*. (2001) *In vitro* human tissue models in risk assessment: report of a consensus-building workshop. *Toxicological Sciences*, **59**, 17–36.

Miners, J.O., Knights, K.M., Houston, J.B. and MacKenzie, P.I. (2006) *In vitro-in vivo* correlation for drugs and other compounds eliminated by glucuronidation in humans: pitfalls and promises. *Biochemical Pharmacology*, **71**, 1531–1539.

Miners, J.O., Smith, P.A., Sorich, M.J. *et al.* (2004) Predicting human drug glucuronidation parameters: application of *in vitro* and *in silico* modeling approaches. *Annual Review of Pharmacology and Toxicology*, **44**, 1–25.

Myllynen, P., Pienimaki, P. and Vahakangas, K. (2005) Human placental perfusion method in the assessment of transplacental passage of antiepileptic drugs. *Toxicology and Applied Pharmacology*, **207**, 489–494.

Naik, P. and Cucullo, L. (2012) *In vitro* blood-brain barrier models: current and perspective technologies. *Journal of Pharmaceutical Sciences*, **101**, 1337–1354.

Naritomi, Y., Terashita, S., Kagayama, A. and Sugiyama, Y. (2003) Utility of hepatocytes in predicting drug metabolism: comparison of hepatic intrinsic clearance in rats and humans *in vivo* and *in vitro*. *Drug Metabolism and Disposition*, **31**, 580–588.

Neis, M.M., Wendel, A., Wiederholt, T. *et al.* (2010) Expression and induction of cytochrome P450 isoenzymes in human skin equivalents. *Skin Pharmacology and Physiology*, **23**, 29–39.

Nohynek, G. J., Skare, J. A., Meuling, W. J., Hein, D. W., *et al.* 2004. Urinary acetylated metabolites and *N*-acetyltransferase-2 genotype in human subjects treated with a *para*-phenylenediamine-containing oxidative hair dye. *Food Chem Toxicol*, **42**, 1885–91.

Olsavsky, K.M., Page, J.L., Johnson, M.C. *et al.* (2007) Gene expression profiling and differentiation assessment in primary human hepatocyte cultures, established hepatoma cell lines, and human liver tissues. *Toxicology and Applied Pharmacology*, **222**, 42–56.

Page, J.L., Johnson, M.C., Olsavsky, K.M. *et al.* (2007) Gene expression profiling of extracellular matrix as an effector of human hepatocyte phenotype in primary cell culture. *Toxicological Sciences*, **97**, 384–397.

Prieto, P., Blaauboer, B.J., De Boer, A.G. *et al.* (2004) Blood-brain barrier *in vitro* models and their application in toxicology. The report and recommendations of ECVAM Workshop 49. *Alternatives to Laboratory Animals*, **32**, 37–50.

Proctor, N.J., Tucker, G.T. and Rostami-Hodjegan, A. (2004) Predicting drug clearance from recombinantly expressed CYPs: intersystem extrapolation factors. *Xenobiotica*, **34**, 151–178.

Roymans, D., Van Looveren, C., Leone, A. *et al.* (2004) Determination of cytochrome P450 1A2 and cytochrome P450 3A4 induction in cryopreserved human hepatocytes. *Biochemical Pharmacology*, **67**, 427–437.

Sanzgiri, U.Y., Kim, H.J., Muralidhara, S. *et al.* (1995) Effect of route and pattern of exposure on the pharmacokinetics and acute hepatotoxicity of carbon tetrachloride. *Toxicology and Applied Pharmacology*, **134**, 148–154.

Sartipy, P., Bjorquist, P., Strehl, R. and Hyllner, J. (2007) The application of human embryonic stem cell technologies to drug discovery. *Drug Discovery Today*, **12**, 688–699.

Schafer-Korting, M., Bock, U., Diembeck, W. *et al.* (2008) The use of reconstructed human epidermis for skin absorption testing: results of the validation study. *Alternatives to Laboratory Animals*, **36**, 161–187.

Schafer-Korting, M., Bock, U., Gamer, A., Haberland, A., *et al.* 2006. Reconstructed human epidermis for skin absorption testing: results of the German prevalidation study. *Alternatives to Laboratory Animals*, **34**, 283–94.

Schroeder, K., Bremm, K.D., Alepee, N. *et al.* (2011) Report from the EPAA workshop: *in vitro* ADME in safety testing used by EPAA industry sectors. *Toxicology In Vitro*, **25**, 589–604.

Sladowski, D., Combes, R., Van Der Valk, J. *et al.* (2005) ESTIV questionnaire on the acquisition and use of primary human cells and tissue in toxicology. *Toxicology In Vitro*, **19**, 1009–1013.

Soto-Gutierrez, A., Navarro-Alvarez, N., Caballero-Corbalan, J. *et al.* (2008) Endoderm induction for hepatic and pancreatic differentiation of ES cells. *Acta Medica Okayama*, **62**, 63–68.

Stahl, J.E., Kreke, J.E., Malek, F.A. *et al.* (2008) Consequences of cold-ischemia time on primary nonfunction and patient and graft survival in liver transplantation: a meta-analysis. *PLoS One*, **3**, e2468.

Stanley, L.A., Horsburgh, B.C., Ross, J. *et al.* (2009) Drug transporters: gatekeepers controlling access of xenobiotics to the cellular interior. *Drug Metabolism Reviews*, **41**, 27–65.

Sugano, K., Nabuchi, Y., Machida, M. and Aso, Y. (2003) Prediction of human intestinal permeability using artificial membrane permeability. *International Journal of Pharmaceutics*, **257**, 245–251.

Sullivan, G.J., Hay, D.C., Park, I.H. *et al.* (2010) Generation of functional human hepatic endoderm from human induced pluripotent stem cells. *Hepatology*, **51**, 329–335.

Taipalensuu, J., Tavelin, S., Lazorova, L. *et al.* (2004) Exploring the quantitative relationship between the level of MDR1 transcript, protein and function using digoxin as a marker of MDR1-dependent drug efflux activity. *European Journal of Pharmaceutical Sciences*, **21**, 69–75.

Venkatakrishnan, K., Von Moltke, L.L., Court, M.H. *et al.* (2000) Comparison between cytochrome P450 (CYP) content and relative activity approaches to scaling from cDNA-expressed CYPs to human liver microsomes: ratios of accessory proteins as sources of discrepancies between the approaches. *Drug Metabolism and Disposition*, **28**, 1493–1504.

Vickers, A.E. and Fisher, R.L. (2004) Organ slices for the evaluation of human drug toxicity. *Chemico-Biological Interactions*, **150**, 87–96.

Vinken, M., Pauwels, M., Ates, G. *et al.* (2012) Screening of repeated dose toxicity data present in SCC(NF)P/SCCS safety evaluations of cosmetic ingredients. *Archives of Toxicology*, **86**, 405–412.

Wilhelm, I., Fazakas, C. and Krizbai, I.A. (2011) *In vitro* models of the blood-brain barrier. *Acta Neurobiologiae Experimentalis (Warsaw)*, **71**, 113–128.

Yamashita, F. and Hashida, M. (2003) Mechanistic and empirical modeling of skin permeation of drugs. *Advanced Drug Delivery Reviews*, **55**, 1185–1199.

Yamazaki, M., Neway, W.E., Ohe, T. *et al.* (2001) *In vitro* substrate identification studies for P-glycoprotein-mediated transport: species difference and predictability of *in vivo* results. *Journal of Pharmacology and Experimental Therapeutics*, **296**, 723–735.

Zamule, S.M., Coslo, D.M., Chen, F. and Omiecinski, C.J. (2011) Differentiation of human embryonic stem cells along a hepatic lineage. *Chemico-Biological Interactions*, **190**, 62–72.

Zanelli, U., Caradonna, N.P., Hallifax, D. *et al.* (2012) Comparison of cryopreserved HepaRG cells with cryopreserved human hepatocytes for prediction of clearance for 26 drugs. *Drug Metabolism and Disposition*, **40**, 104–110.

6

In Silico Methods and Structure–Activity Relationships

6.1 Why *in silico*?

The topic of this chapter is the use of *in silico* methods to predict systemic toxicity, drug/toxin uptake and pathways of metabolism. The term *in silico* means the use of computational methods to create models which can be used to make predictions about ADME and toxicity, thus saving the time and expense involved in carrying out 'real-life' studies. *In silico* approaches can be particularly helpful in contexts such as the drug discovery stage of pharmaceutical development, where the quantity of a candidate compound available may be very small. In this setting, *in silico* predictions may be used to inform the design of key experiments thus maximizing the amount of information obtained using the precious material from early syntheses.

Many *in silico* methods for the prediction of ADME and toxicity depend upon the establishment of quantitative structure–activity relationships (QSARs). As the name suggests, these are mathematical models which link the chemical structure of a compound to some aspect of its biological activity (e.g. receptor binding, susceptibility to CYP-mediated metabolism).[1] In some cases QSARs are based purely on extrapolation from observational data (e.g. dose–response curves) but the most informative QSARs are those which are based on mechanistic or structural information (e.g. the structure of a receptor's ligand binding site).

[1] For an introductory review, see Modi *et al.* (2012).

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/molecularcellulartoxicology

6.2 Predicting the ADME characteristics of xenobiotics

The use of *in vitro* and computational approaches as early as possible in the drug discovery process, the commercial development of a new cosmetic/consumer product or the assessment of potential risks associated with workplace/environmental exposure to an industrial chemical helps to reduce the number of safety issues needing to be addressed using *in vivo* methods. The traditional approach involves QSARs and expert systems which use statistical and learning approaches, molecular descriptors and experimental data to model biological processes such as intestinal absorption, hepatic metabolism and oral bioavailability; however, these approaches are compromised by data gaps and fail to take into account the available structural information about the proteins which participate in the processes under consideration. Alternative approaches, based on 3D-protein structures, began to be applied in the early 2000s and have started to contribute to the knowledge base, although there are still significant challenges due to the broad, overlapping substrate specificities of xenobiotic metabolising enzymes and other key proteins.[2]

6.2.1 Absorption

One of the drivers of research on systemic pharmacokinetics is the need to improve plasma concentrations of potent drugs which have very poor bioavailability. Numerous *in silico* approaches have been developed by and for the pharmaceutical industry in this endeavour; these include models for the prediction of oral absorption, transporter activity, bioavailability, volume of distribution, metabolism and CYP induction/CYP inhibition (Davis and Riley, 2004). On the other hand, the cosmetics industry has not, to date, taken up many of these strategies, partly because the needs of the two industries with respect to bioavailability and metabolic stability are very different.

The aim of *in vitro* ADME screening in the earliest stages of drug discovery is, essentially, to classify compounds as 'good' (high bioavailability/metabolically stable) or 'bad' (low bioavailability/metabolically unstable) (Walum *et al.*, 2005) with a view to facilitating candidate selection. Numerous *in silico* approaches have been developed by and for the pharmaceutical industry in this endeavour; these include models for the prediction of oral absorption, transporter binding, bioavailability, volume of distribution, metabolism and CYP induction/CYP inhibition (Davis and Riley, 2004). More accurate assays are used later in drug development to generate quantitative values in order to predict actual doses required to induce a clinical effect and to establish a provisional therapeutic index. In contrast, the ideal compound for the cosmetics industry will have very low bioavailability and the small proportion which does enter the body will be cleared rapidly, preferably without the need for metabolism. If metabolism does occur it should be rapid and lead to the generation of inactive metabolites. Thus the needs of the cosmetics industry with respect to the use of *in vitro* alternatives for predicting ADME *in vivo* are the mirror image of those of the pharmaceutical industry.

[2] For a detailed review, see Moroy *et al.* (2012).

Dermal penetration The physicochemical characteristics which determine the passive uptake of a xenobiotic are its lipophilicity, H-bonding and ionisation properties; molecular size and shape; polarity and flexibility (Balimane *et al.*, 2000). These can be used in mathematical models to estimate more complex physicochemical parameters such as the octanol-water partition coefficient (K_{OW}), which is conventionally accepted as a measure of lipophilicity and is used for the prediction of permeability. Parameters such as K_{OW} can, in turn, be used to develop quantitative structure–permeability relationships (QSPRs), a class of QSAR which is specific to the study of membrane permeability. The conceptual simplicity of the PAMPA assay makes it a good candidate for this purpose.

The importance of dermal absorption, both in the safety evaluation of cosmetics and the discovery/development of topically active drug formulations, has driven efforts to derive dermal QSPRs in an attempt to relate percutaneous chemical uptake to known physicochemical parameters. Furthermore, REACH explicitly encourages the greater use of QSARs and QSPRs in order to reduce costs and use of animals in toxicity testing. This type of model is based upon the assumption that uptake via the skin is essentially a passive process; indeed, it has been known for decades that partitioning and solubility are key to dermal penetration. The key parameters required for the development of QSPRs are hydrophobicity, molecular size and H-bonding capacity (Moss *et al.*, 2002), and modern models range from simple ones which consider the stratum corneum (or entire skin) as a single compartment to those which explicitly attempt to take into account the structural complexity of the skin.[3]

The aims of QSPR models range from the screening of potential drug candidates for transdermal delivery to the assessment of potential risks following pesticide exposure. The key focus has been to define the permeability coefficient (k_p) of a compound, which is defined as the steady-state permeation rate across a membrane (J_{ss}) between a donor and a receptor compartment normalised against the concentration gradient between the two compartments (ΔC_v). Thus, $k_p = J_{ss}/\Delta C_v$ and, in fact, when the concentration in the receptor compartment is effectively zero this can be simplified to $k_p = J_{ss}/C_v$.

Most QSPR models are based on algorithms which calculate k_p for the penetration of a compound applied in an aqueous formulation, mainly because there is a large historical database of this kind of information. The most commonly used model was developed by Potts and Guy (1992) and is based on compounds with molecular weights between 18 and >750 and logP between −3 and 6. The formula derived by Potts and Guy is

$$\text{Log}k_p = -6.3 + 0.71\,\text{logP} - 0.061\,\text{MW}$$

Other models have been proposed, but these tend to be based on more limited datasets.

It is important to recognise that many QSPR models are limited by factors including the use of data from a variety of sources, formulation effects and a failure to take into account different endpoints (surface deposition, superficial skin

[3] For a comprehensive review of QSPR and other approaches to the prediction of skin permeability, see Mitragotri *et al.* (2011).

penetration, delivery to skin appendages and transdermal delivery) (Moss *et al*., 2002). A long list of potential sources of error in QSPRs has been identified by Dearden *et al*. (2009) in a very readable review of 'how not to do it'. They highlight the importance of adequate validation, including cross-validation within the training set and external validation (i.e. using test compounds which are not part of the training set). Thorough external validation is now considered to be essential, but unfortunately Dearden *et al*. are forced to conclude that many published QSAR/QSPR analyses still contain errors. Let the reader beware!

The inclusion of QSAR/QSPR data in REACH evaluations has highlighted the need for a flexible scientific validation process to allow regulatory agencies and companies to establish the reliability of QSAR predictions. To that end, the OECD has defined a set of principles for validation, for regulatory purposes, of QSAR models.[4] The requirements identified by OECD are

1. a defined endpoint;
2. an unambiguous algorithm;
3. a defined domain of applicability;
4. appropriate measures of goodness-of-fit, robustness and predictivity;
5. a mechanistic interpretation, if possible.

Unfortunately, when these principles were used to evaluate 33 QSPR models (Bouwman *et al*., 2008) it was found that only one (the model of Magnusson *et al*. (2004)) gave a relatively good correlation between predicted and experimental values, and even with this one the absolute values derived from QSPRs could be out by two to three orders of magnitude! This led Boumann *et al*. to conclude that the use of QSAR is currently limited to initial screening, although they also recognise the potential value of QSPRs as part of a weight of evidence approach or an intelligent testing strategy. Meanwhile, there is an urgent need for models which can take into account the complexity of the skin; the problem is that this tends to make the models more complicated, less transparent and therefore less suitable for use by non-experts.

The QSPR approach has some intrinsic limitations: it is not appropriate for formulations whose components modulate the barrier properties of the skin, and it struggles to deal with complex multicomponent or multiphasic formulations (such as many cosmetics and topical drug formulations). Research into improved methods for modelling dermal uptakes has therefore involved a range of approaches including, in addition to improved QSPRs based on chemical structures (e.g. fragment descriptors, neural network-based approaches), steady-state models, porous pathway models, transient models and pharmacokinetic models of the skin. An alternative approach is to recognise the non-linear nature of human skin permeability and apply non-linear methods such as Gaussian process regression (Moss *et al*., 2011; Brown *et al*., 2012).

Mathematical models provide a way of assessing the likelihood of internal exposure to compounds which come into contact with the skin, but at the time of writing there is still some resistance to using such models in risk assessment. It has been suggested that this is because the models are not accessible to risk assessors

[4] http://www.oecd.org/env/ehs/risk-assessment/37849783.pdf

and industrial hygienists, who are confused as to how to use their outputs. There is also an issue with regard to the legitimacy and applicability of models which are currently available. In addition, most of the available models are limited to single chemicals in simple matrices such as water, which seriously limits their usefulness. Overall, the need to keep the models simple enough to be used by the experimental community while still explaining complex real-world phenomena has not yet been met.

Efforts have been made to combine *in silico* approaches with the application of RHE models in order to improve the prediction of skin penetration of a range of test materials (Schafer-Korting *et al.*, 2008). Unfortunately, while the experimental performance of the RHE models was satisfactory, the results of *in silico* analysis were not. When apparent permeability (P_{app}) values derived from the permeation of excised human skin, pig skin and RHE preparations were correlated with P_{app} values calculated on the basis of molecular weight and lipophilicity, the results obtained by calculation using the DERWIN version 1.43 (US Environmental Protection Agency, 2000) and SKINPERM Version 3.2 program packages were closely related ($r = 0.947$), but poor correlation was observed between experimental and calculated data, particularly when so-called problem compounds (clotrimazole, digoxin, ivermectin and mannitol) whose physicochemical properties lay outside the range covered by the substances used in model development were included, but also when more conventional compounds were considered. This seemed to be a problem with the *in silico* methods used rather than the *in vitro* test systems themselves.

An evaluation of the QSPR approach by Godin and Touitou (2007) emphasises a number of specific uncertainties with respect to its application to percutaneous absorption, for example:

- Limitations in terms of statistical fit
- Failure under severe non-linear conditions
- Inability to extrapolate conclusions to other systems, especially when the carrier effects have to be taken into account
- Issues of validity with respect to mechanisms of permeation

More recent attempts to predict skin permeability have involved the use of artificial neural networks. This empirical approach takes into account various molecular and structural descriptors without trying to apply any mechanistic consideration. Artificial neural networking appears to provide superior predictions of skin permeability compared with mechanistic models and is considered to be a promising technique for the non-linear modelling of complex cause and effect relationships.

Ocular penetration Attempts to model direct ocular penetration have focussed on the development of a simple pharmacokinetic model which assumes that permeation can be described by a single rate constant. However, in reality xenobiotic permeation via this route is a complex phenomenon and the molecule has to permeate across several tissues and barriers in both the anterior and posterior eye. A more sophisticated model, comprising three successive barriers (the sclera, the choroid and the retinal pigmented epithelial cell barrier), has now been

developed. This model, which assumes absorption over an area of 1 cm^2 and no lateral diffusion, has generated promising results although it has a number of limitations:

- It does not take into account losses from the scleral surface by non-productive loss processes.
- The concentration of test compound is assumed to remain constant at the scleral surface to simulate steady-state concentrations and fluxes.
- The model lacks some physiological features and only considers passive diffusion across the retinal pigmented epithelium.
- The data used has been collected from several different sources.

Another limitation is that this model addresses access from the systemic circulation into the eye but not from the eye to the systemic circulation. Further development of the model may allow the disposition of xenobiotics following ocular application, but at the present time it is limited in its application and not suitable for use in a routine setting.

Intestinal absorption The two main routes via which humans absorb drugs and xenobiotics are through the skin and GI tract. For drugs, the choice of a route of administration is determined by the chemical properties of the drug and the therapeutic objective (e.g. rapid action or sustained effect). Routes of drug administration are classified as local (topical) or systemic (enteral or parenteral).

While it is highly efficient in absorbing desirable substances such as nutrients, the small intestinal mucosa forms an important barrier to the absorption of xenobiotics. The intestinal absorption of xenobiotics is affected by conditions in the GI tract and by the characteristics of the xenobiotic in question (e.g. susceptibility to intestinal first pass metabolism, ionisation state). The biological factors involved include the area and condition of mucosal surfaces, the degree of gastric emptying, GI motility and the presence of food.

The optimum route of administration for drugs is the oral route since this is convenient and usually does not require close medical supervision. However, although oral dosing is the commonest route of administration for pharmaceutical agents, it is also the most variable in terms of achieving the desired effect because the pathway from the mouth to the target tissue can be complicated, and there can be problems with degradation in the stomach. In addition, a key consideration is the extent to which the drug undergoes first pass metabolism in the GI tract and liver before it reaches the systemic circulation. Rectal administration (by means of a suppository) can help to reduce these problems because a significant percentage of the rectal blood flow bypasses the liver.

Software packages such as GastroPlus™ from Simulations Plus[5] use mathematical models to simulate oral absorption. GastroPlus™ uses a proprietary Advanced Compartmental Absorption and Transit model which models the intestinal tract as nine compartments (the stomach, seven sections of the small intestine and the colon) and can be used to predict dissolution and disintegration of formulations, passage along the GI tract at different pHs, intestinal absorption and transfer across enterocytes to reach the circulation. It yields a range of results reflecting the

[5] http://www.simulations-plus.com/Products.aspx?pID=11

rate and extent of absorption processes in different sections of the GI tract and can also be adapted to include carrier-mediated drug transport and drug metabolism.[6] The initial use of GastroPlus™ was to model oral absorption but more recently it has been extended to handle complex processes such as release from formulations, absorption/excretion and intestinal transit. GastroPlus™ can also contribute to the prediction of systemic pharmacokinetics, as discussed later in this chapter.

One way of avoiding the need to purchase highly specialised software for the prediction of permeability and absorption is to use a simple Microsoft® Windows® based package. To this end, several Excel®-based applications are available (Langenbucher, 2007; McCoy *et al*., 2012). These include PHDISSOC, which models electrolytic dissolution and its effect on solubility and partitioning; SOLUPART, which predicts absorption potential; GIABS, a mechanistic model simulating drug transport and absorption through the GI tract; and TK Modeller, which predicts blood/plasma concentrations resulting from single, multiple bolus or dietary dosing.

The majority of *in silico* methods for the prediction of xenobiotic absorption do not take into account the activity of transporters or xenobiotic metabolising capacity of the intestine. The available computational methods for predicting transporter interactions typically involve determination of the extent to which a given compound will inhibit the *in vitro* uptake of a probe substrate and their stage of development is at least a decade behind that for predicting CYP-mediated metabolism (Ekins *et al*., 2012). Further developments in this area await the construction of initial models based on a relatively small number of compounds, focussing on inhibitory interactions because they are relatively easy to measure, followed by high throughput analysis to build up the number of compounds considered. Once sufficient information is available, this can be combined with improved 3D-structural models.

In order to extend the available methods, information about the following would be required (Galetin and Houston, 2006; Fagerholm, 2007a):

- Enterocytic blood flow
- Cellular heterogeneity
- Spatial organisation of CYPs and UGTs
- Intestinal distribution and absorptive area
- The impact of efflux and reabsorption
- The number, distribution and metabolic activity of intestinal enterocytes would be required

The latter three factors have been incorporated into a simple first pass intestinal wall elimination model based on human cell *in vitro* clearance values, *in vivo* first pass data for reference compounds and CL_{INT} for test compound together with fractional absorption from the small intestine (Fagerholm, 2007a). This method works well for CYP3A4 substrates with various degrees of intestinal uptake and CL_{INT} as well as for substrates of other CYP isozymes but requires further data before it can be used in a routine setting. A better understanding of the mechanisms involved in intestinal xenobiotic uptake and metabolism is required before

[6] For a general review, see Lupfert and Reichel (2005).

it becomes possible to predict intestinal first pass metabolism from *in vitro* data on intestinal CL_{INT} (Galetin and Houston, 2006).

6.2.2 Distribution

Tissue composition-based approaches to xenobiotic distribution make use of a variety of outputs, taking into account physicochemical properties, acid dissociation constant (pK_a) and, in some cases, protein binding (Boobis *et al*., 2002). These factors are incorporated into partitioning models developed using information on tissue and blood composition and blood flow to the tissues of interest. Semi-empirical approaches, which use *in vivo* kinetic data, may also be used to predict the tissue distribution of xenobiotics. The principle of this approach is to partition the volume of distribution at steady state, as determined experimentally, to predict individual tissue-plasma ratios.

Blood-brain barrier penetration Accepted *in silico* models of the BBB focus on characteristics similar to those which determine intestinal absorption (e.g. lipophilicity, H-bonding and ionisation characteristics, pK_a and molecular size).[7] Conventional attempts to predict BBB penetration have used partition coefficients measured in various solvent systems, but this is a very simplistic approach to a complex problem. Commercially available packages which adopt this approach include ACD/logP; this is included as part of the chemical drawing package ACD/ChemSketch, which is available in both commercial and freeware versions.[8]

Other factors which have to be taken into account include free fraction in blood and brain, ability to partition into capillary membranes, blood flow, distribution into the brain parenchyma and active efflux. Simple rule-based approaches to the *in silico* prediction of BBB penetration are useful in so far as they contribute to an understanding of the underlying mechanisms involved, but they may not reflect the complexity of BBB function. More sophisticated approaches which have been adopted recently have used neural networks to model the relationship between the log blood-brain partition coefficient (LogBB, measured *in vivo*) and a set of computed property descriptors. This *in silico* approach was able to predict LogBB with reasonable accuracy but could not predict the free drug concentration in the brain, the key parameter which determines biological effects.

The available *in silico* models of the BBB are rapid and relatively inexpensive, but are phenomenologically based and rely on limited datasets. In order to improve these models further, better/larger *in vivo* datasets and improved *in vitro* models are required. Increased ease, speed and throughput may be achieved but this often occurs at the cost of decreased accuracy of prediction.

In silico models of the BBB are still limited by the accessibility, acquisition and reliability of the data sets necessary for constructing a QSAR and need to be supported by further *in vitro* and *in vivo* studies (Naik and Cucullo, 2012). This means that they cannot yet be reliably used to predict the brain distribution of test compounds.

[7] For a review, see Goodwin and Clark (2005).
[8] http://www.acdlabs.com/products/draw_nom/draw/chemsketch/ and http://www.acdlabs.com/resources/freeware/chemsketch/

6.2.3 Metabolism

Hepatic metabolism The parameters which need to be taken into account during the development of xenobiotic metabolism prediction models include the physicochemical characteristics of the substrate (molecular sites sensitive to oxidation or conjugation, 3D-structure, molecular surface properties, electronic structure, quantum mechanics, polarity, hydrophobicity, lipophilicity, H-bonding capabilities and 3D-molecular interaction fields) and a range of biological/biochemical parameters including protein structure, especially of the active site, enzyme specificity/regioselectivity, accessibility of the binding site, specific activity and reaction mechanism (Boobis *et al*., 2002).

Modelling approaches in which the three-dimensional interactions between chemical substrates and enzymes of biotransformation are predicted at the whole molecule and active site level may allow the prediction of susceptibility to metabolism by particular enzyme families (e.g. CYPs). This approach involves calculating the quantum mechanical properties of molecules in order to establish their intrinsic reactivities then predicting the interactions between metabolising enzymes and other biological receptors based on electronic and 3D considerations. It requires prior structural characterisation of the enzymes involved and has been facilitated by the availability of crystal structures for a number of CYPs (although until recently these were largely predicted based upon microbial enzyme structures). This approach has been used to predict both metabolism by specific enzymes and metabolic stability as a general property, although it should be noted that the introduction of bias depending upon the core structures used to develop an individual model may limit its applicability (Ekins, 2003).

In silico methods for the prediction of drug metabolism include rule-based, ligand-based and ligand–protein-based approaches as well as pharmacophore and 3D-QSAR-based strategies. Pharmacophore models, which describe the critical 3D arrangement of functional groups in the ligand that is responsible for generating a biological response (Ekins *et al*., 2000, Ekins and Obach, 2000), have been developed for some enzymes based on substrate/inhibitor shape, electronic properties and molecular conformation. Examples of available software for 2D and 3D-QSAR, pharmacophore modelling and metabolite prediction include MetaDrug (Reuters), CoMFA (*C*omparative *M*olecular *F*ield *A*nalysis; Tripos), Pentacle (Molecular Discovery) and GOLPE (*G*enerating *O*ptimal *L*inear *P*LS *E*stimations/; Multivariate Infometric Analysis S.r.I).[9] Until recently, proprietary packages like these were the only ones available, but two open-source software packages, 3-D QSAutogridR and Open3DQSAR,[10] have recently been released (Tosco and Balle, 2011; Ballante and Ragno, 2012). One of these, 3-D QSAutogridR, has been used to create the first open-access 3D-QSAR server[11] (Ballante and Ragno, 2012).

Among the most popular packages used for the prediction of routes of metabolism and potential toxicity are Meteor (Lhasa Limited) and MetabolExpert

[9] http://thomsonreuters.com/content/science/pdf/ls/metadrug-cfs-en.pdf; http://www.tripos.com/index.php?family=modules,SimplePage,,,&page=QSAR_CoMFA; http://www.moldiscovery.com/soft_pentacle.php; http://www.miasrl.com/golpe.htm.

[10] http://open3dqsar.sourceforge.net/?Description

[11] www.3d-qsar.com

(CompuDrug) (Boobis *et al*.; 2002, Jolivette and Ekins, 2007; Marchant *et al*., 2008).[12] These packages were developed on the basis of structure-metabolism rules which allow prediction of the metabolic fate of a compound and are widely used by the pharmaceutical industry during the discovery and early development phases of drug development.

- **Meteor** is one of a suite of programs developed by Lhasa Limited, a not-for-profit organisation devoted to the promotion of data and knowledge sharing in chemistry and the life science (Marchant *et al*., 2008). The other programs offered by Lhasa Limited are Zeneth (for predicting forced chemical degradation pathways of organic compounds), Derek Nexus (for predicting toxicity), and Vitic Nexus (a chemical database and information management system). Its aim is to deduce risk from existing knowledge using rules taken from the files of mechanistic organic chemistry. Meteor is a rule-based empirical software tool which compares the query structure against a database of known biotransformations and applies absolute reasoning rules to evaluate the likelihood of a biotransformation, classifying it as probable, plausible, equivocal, doubtful or improbable. Relative reasoning is then used to rank the comparative likelihood of possible biotransformations based on a set of relative precedences (e.g. 'N-demethylation is more likely than O-demethylation'). One of the strengths of Meteor is that it predicts actual biotransformations rather than just potential sites of metabolism on a molecule. An example of the type of output delivered by Meteor is illustrated in Figure 6.1.
- **MetabolExpert**, offered by CompuDrug, is a rule-based expert system which predicts potential metabolites by rules constructed from known metabolic transformations on the basis of >400 literature references.

Example: Use of MetabolExpert combined with Glide to identify substrates for CYP2C9

Combined approaches can increase confidence in the prediction of sites of metabolism; for example, when a three-step structure-based approach combining identification of possible CYP2C9 metabolites using a rule-based system (MetabolExpert) with docking to the active site of CYP2C9 (Glide[13] Version 5.0) was compared with approaches based on substrate docking alone, the results were comparable with those achieved by literature approaches used for site of metabolism prediction and demonstrated that metabolite docking performed better than substrate docking on the dataset used (Tarcsay *et al*., 2010). With modification by adding recently identified hydroxylation reactions, MetabolExpert correctly predicted 82% of true metabolites and docking predictions by Glide were 84% correct. By combining the results of the two methods, a correct metabolite was found among the three best-ranked options in 69% of the cases. The main advantages of the method were automatic evaluation of the results and the direct prediction of the metabolite structures instead of the site of metabolism only.

[12] https://www.lhasalimited.org/meteor/for; http://www.compudrug.com/?q=node/36.
[13] http://www.schrodinger.com/productpage/14/5/

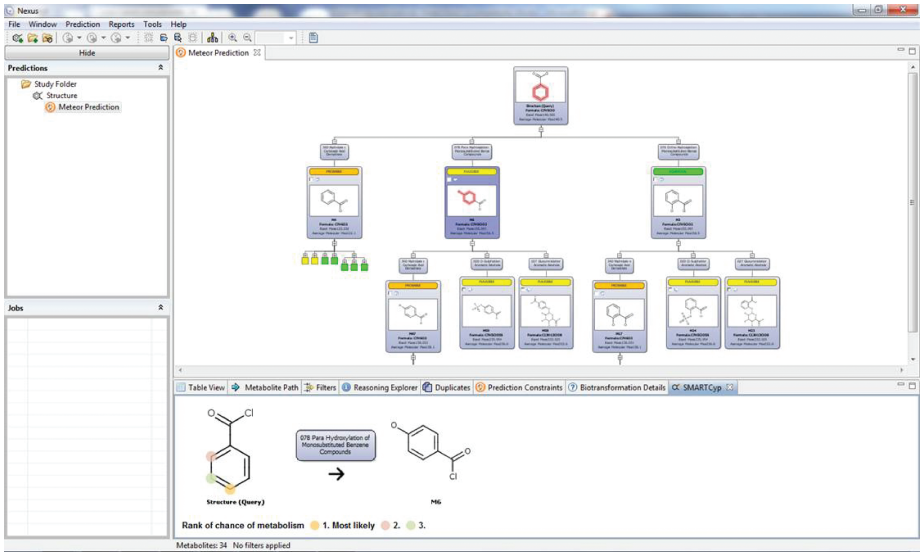


Figure 6.1 Meteor screenshot. The prediction diagram was created using Meteor Nexus v1.5.2, Meteor 2012.1.0.0 Knowledge base, prediction constraints = Maximum. Depth: 2, Default all other options (source: Lhasa Limited. Reproduced with permission of Lhasa Limited)

Both Meteor and MetabolExpert are easy to use, give rapid answers and are linked to toxicity prediction packages (Derek and HazardExpert, respectively). They allow the user to participate in the generation of a metabolic tree by distinguishing between Phase I and Phase II reactions and introducing rules, biotransformations, examples and literature references. However, one of their problems, as with all *in silico* packages designed to predict biotransformation, is that they often predict large numbers of metabolites, many of which may never be identified experimentally. It is therefore necessary to constrain the analysis in order to provide realistic predictions (Marchant *et al.*, 2008).

The relative performance and validation of different metabolism prediction packages has been controversial (Coecke *et al.*, 2006).

Example: Comparison of Meteor, MetaSite and StarDrop for predicting Phase I metabolism

Efforts have recently been made to compare different packages directly (T'jollyn *et al.*, 2011). Meteor (Version 11) was compared with MetaSite,[14] an automated docking model which predicts Phase I metabolism (Version 3.0.1) and StarDrop,[15] which uses a quantum mechanical approach to predict the involvement of CYP3A4, CYP2D6 and CYP2C9 in the metabolism of a query molecule on the basis of the energy barrier in the removal of an electron, which is the rate-determining step in product formation (Version 3.4). The latter two models identify potential metabolic hotspots on a query molecule rather than predicting

[14] http://www.moldiscovery.com/soft_metasite.php
[15] http://www.optibrium.com/stardrop/

actual biotransformations; thus Meteor might identify two or more possible metabolites where the other two just identify a metabolic hotspot. The three packages were compared in terms of sensitivity (how many of the actual *in vivo* metabolites are identified) and precision (how many of the metabolites identified by the package are actually observed *in vivo*). A test set of 22 compounds with divergent chemical structures, therapeutic indications, extent of metabolism, number of known metabolites and metabolic liabilities were used and the conclusions were as follows:

- Meteor was more sensitive than the other packages (i.e. it identified more potential metabolites) but it had a tendency for over prediction.
- StarDrop and MetaSite yielded very precise, but much less sensitive results which may be more useful for lead optimisation and 'soft-spot' prediction.
- Before choosing a package to predict metabolism, it is necessary to make a decision as to whether to prioritise sensitivity ('what might happen') or precision ('what will happen').

In particular, reliable methods for predicting major metabolites and their further effects in a complete metabolic system are needed, both for the selection of molecules for synthesis and development in the pharmaceutical industry and for safety evaluation without the use of animals in the cosmetics industry.

UGTs The atypical kinetic behaviour of UGTs, which is consistent with the existence of allosteric effector sites or simultaneous binding of two substrate molecules at a single active site, is particularly difficult to model. However, human UGTs belong to the glycosyltransferase (GT-1) family, for which a number of structures are available, and despite the fact that the overall homology between UGTs and GT-1s is only about 20% the structural fold is predicted to be highly conserved. This has allowed a number of 2D/3D-QSARs and pharmacophore models for UGTs to be developed.[16] These provide good predictions of xenobiotic metabolism by several UGT isozymes although they can be difficult to interpret because of the complexity of UGT–substrate interactions and the lack of a clear relationship between substrate binding and metabolic turnover.

Example: *In silico* modelling of UGTs using the VolSurf approach

The VolSurf approach lends itself to the *in silico* modelling of enzymes whose 3D-structures have not been determined and is of value as an alternative to 3D-QSAR and CoMFA methods. VolSurf does not require alignment of substrate molecules to a predicted active site structure; rather, it uses computational methods to generate 2D-molecular descriptors from 3D-molecular interaction energy grid maps. This approach has recently been exploited to develop models for two human UGTs, UGT1A10 and UGT2B7 (Dong and Wu, 2012; Ako *et al.*, 2012).

The polymorphic enzyme UGT1A10 has broad substrate specificity which allows it to metabolise phenols, nitrosamines, flavonoids, oestrogens and PAHs. It is of clinical relevance because it is responsible for the rapid clearance of

[16] Reviewed by Dong *et al.* (2012).

raloxifene, a selective oestrogen receptor modulator which is used to prevent and treat osteoporosis, reducing its oral bioavailablity to about 2%. Use of a diverse training set including fatty acids, simple phenols, coumarins, flavonoids, stilbenes, nitrosamines, lignans, PAHs and oestrogens in VolSurf identified a number of chemical features associated with the enzyme–substrate interaction, including size and shape, hydrophilic and hydrophobic regions, polarizability and the so-called capacity factor (the ratio between hydrophilic regions and total surface area). While it may seem self-evident that these features would be expected to affect enzyme–substrate interactions, it should be noted that features such as size and shape, as well as hydrophobic regions, had not proved to be important in *in silico* models of other UGTS (UGT1A1 and UGT1A4).

Another polymorphic UGT, UGT2B7, plays a major role in the glucuronidation of a number of widely used drugs including codeine, ketoprofen and gemfibrozil. It can metabolise both O- and N-glycones, and is able to accept a range of hydroxyl classes including phenolic and alcoholic hydroxyls as well as carboxylic acids. Its broad substrate specificity suggests that it has a general role in chemical defence. Both pharmacophore and VolSurf approaches have been used to model the substrate interactions of UGT2B7. A pharmacophore developed using a training set of 36 structurally diverse substrates identified the following key pharmacophore features:

- one glucuronidation site;
- one hydrogen bond acceptor;
- three hydrophobic regions.

The features identified by VolSurf as being associated with better activity were size and shape, hydrophobicity, hydrogen-bonding capacity, polarizability, molecular weight and LogP, while high values for capacity factor were detrimental to activity. Thus, VolSurf highlighted the importance of hydrophobic and hydrogen-bonding interactions in the functioning of UGT2B7, while hydrophilic interactions were associated with poor substrate binding. In this analysis, the pharmacophore model proved to be similar to one previously developed for UGT1A1, and thus helping to explain the overlapping substrate specificities of these UGTS. However, the authors note that one limitation of the pharmacophore approach is that it does not take the potential flexibility of the active site into account, even though such flexibility is known to exist in a number of xenobiotic metabolising enzymes (e.g. CYP1A1, SULT1A1).

Both these studies have generated models which should be of value in the early identification of drug candidates which are likely to exhibit rapid, and potentially problematic, UGT-mediated clearance in the clinic.

Evaluation As mentioned earlier, the pharmaceutical industry has embraced *in vitro* and *in silico* approaches to the prediction of xenobiotic metabolism and clearance. Such methods are now widely used for the preclinical prediction of the metabolic properties of drug candidates. However, while *in vitro* methods have been shown to be useful tools for the prediction of metabolic stability and drug-drug interaction potential, there is still no rapid and secure route for extrapolating from in house or published *in vitro* data to the *in vivo* human situation (Baranczewski *et al*., 2006; Nagilla *et al*., 2006).

An additional factor which may impact the prediction of clearance and toxicity in a repeated dose setting is the induction of CYP expression. The pharmaceutical industry takes a great interest in this phenomenon because it can lead to loss of efficacy and/or DDIs. It is likely to be less of an issue for the cosmetics industry but may still be important in that it may lead to altered (increased or reduced) toxicity in repeated dose studies compared with acute studies.

Reaction phenotyping of xenobiotics metabolised by glucuronidation is currently feasible but experimentally challenging. Problems with data interpretation can arise due to the dependence of UGT activity on incubation conditions, the source of enzyme and the atypical kinetic behaviour of UGTs. *In vitro–in vivo* correlations may be influenced by fundamental questions relating to xenobiotic glucuronidation, including the implications of dimerisation, dependence on accessory proteins and the effects of the local membrane and cellular environment.

6.2.4 Excretion

Most *in silico* clearance models assume that the dominant process involved in xenobiotic clearance is hepatic metabolism, and relatively little work has been done on the *in silico* prediction of excretion (Boobis *et al*., 2002). Passive excretion can theoretically be predicted using the same methods as for tissue distribution but, in practice, this approach has not been developed because metabolic stability *in vitro* and *in vivo* and initial animal pharmacokinetic studies give sufficient indications as to the potential significance of renal excretion (Boobis *et al*., 2002).

It is possible to include renal and biliary clearance in *in silico* models if these processes are likely to play a significant role in clearance. In principle, passive glomerular filtration, tubular reabsorption and secretion can be predicted from the physicochemical properties and plasma protein binding characteristics of a compound (Lilienblum *et al*., 2008) and renal clearance in humans can be predicted by allometric scaling using rat data corrected for species differences in the glomerular filtration rate (Davis and Riley, 2004). However, this approach is limited by the fact that the major route of excretion is not necessarily the same in humans and animals (Fagerholm, 2007b) and active secretion and reabsorption are even less predictable. In theory, therefore, *in silico* modelling of renal clearance is possible but this has not been attempted to date and more *in vitro* information is required in order to allow the *in silico* prediction of biliary clearance.

6.3 Physiologically based biokinetic modelling

Physiologically based biokinetic (PBBK[17]) models are mathematical models which describe the pharmacokinetics of compounds in the context of the known anatomy and physiology of an organism. They evolved from Monte Carlo

[17] Throughout this chapter, the term physiologically based biokinetic (PBBK) modelling will be used in preference to the (essentially synonymous) terms physiologically based pharmacokinetic (PBPK) and physiologically based toxicokinetic (PBTK) modelling since this term has universal applicability in the context of both pharmacokinetic and toxicological processes.

simulation methods which first became available in the mid-1980s and involve combining experimental results, *in vitro* data, information from the literature and computational techniques in order to simulate the *in vivo* pharmacokinetics of a test compound. This can facilitate the evaluation of existing data, help to identify and weigh the importance of data gaps, allow 'what if' questions to be posed and contribute to the design of 'real world' experiments. The resource intensive nature of this approach initially slowed its uptake in the pharmaceutical industry and until recently it was mainly an academic exercise, but improvements in the power and speed of computing are now helping to overcome this problem and over the last few years the number of regulatory submissions including PBBK data has increased sharply (Zhao *et al*., 2012); indeed, it has been estimated that by 2016 about 20% of pharmaceutical research and development budgets will be devoted to *in silico* modelling. This development has largely been driven by the increased availability of user-friendly software which has generated enthusiasm for more informative mechanistic models.

In a PBBK model, an array of drug-independent parameters such as genetic variation, age and disease state is incorporated into a multi-compartment model with the aim of predicting key biokinetic parameters such as AUC and C_{max}. The compartments considered correspond to either anatomical structures (e.g. liver, kidney) or tissue types (e.g. fat, muscle) and the interconnections correspond to the flow of body fluids (blood or lymph) (Figure 6.2); (Clewell and Clewell, 2008).[18]

The following are the key parameters required for the construction of a PBBK model:

- Biochemical parameters (CL_{INT}, Michaelis constant (K_M) and maximal velocity (V_{max})) – the use of human hepatocytes is preferred but the use of microsomes is considered adequate in many cases.
- Plasma membrane permeabilities – for xenobiotics whose membrane permeability is very low, the rapid equilibrium assumption may not apply.
- Renal and biliary excretion clearances.
- Plasma protein binding.
- Tissue-to-plasma partition coefficients – It is considered reasonable to assume that tissue-to-plasma unbound ratios and volumes of distribution are similar in humans and animals.
- Physiological parameters (tissue mass, blood flow rate).

The biochemical parameters used can be derived from *in vitro* systems or by allometric scaling from animal data.[19] *In vitro* data can be used to define the compound's mode of action, which can then be used to inform the choice of a mathematical model for PBBK modelling. For the prediction of toxicity, integration of pharmacokinetic and pharmacodynamic information with *in vitro* toxicity studies using human hepatocytes is recommended.

Population pharmacokinetics is now seen as an essential tool in drug development; however, the conventional 'top-down' approach, in which models

[18] For a review of the processes involved in PBBK modelling, see Clewell and Clewell (2008).
[19] http://www.nature.com/scitable/knowledge/library/allometry-the-study-of-biological-scaling-13228439

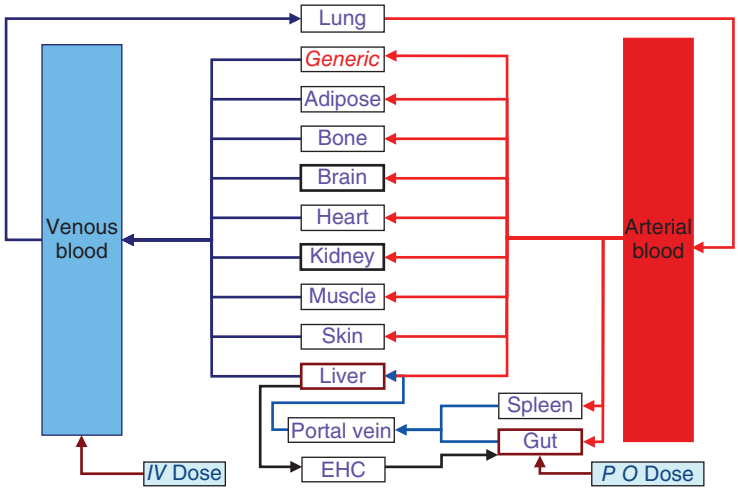


Figure 6.2 Structure of a generic PBPK model of the mammalian body. PBBK models are constructed using anatomical, biochemical and physiological data to calculate steady-state concentrations of compound in various tissues, taking into consideration blood flow and biodistribution. The model can be used to predict hepatic and renal clearance (CL_H and CL_R). The main use of this type of model is in extrapolating preclinical data to man and, since it is physiologically based, it can be used to predict what happens in the event of changes to certain biological processes (e.g. renal impairment, liver cirrhosis, diabetes) (source: Adapted from Bois *et al.* (2010); figure 1. Reproduced with permission of Elsevier)

are constructed using clinical data and the model parameters and covariance are estimated statistically, is limited in that it does not take into account all relevant biological variation. In contrast, PBBK models adopt a 'bottom-up' approach to describe whole biological systems in physiological terms that have relevance to xenobiotic distribution, mode of action and underlying biochemical processes. They include both system-dependent and drug-dependent elements. The power of PBBK modelling is that, in cases where its predictions differ from actual observations, the model can be interrogated in order to find out why.

The most basic PBBK models focus on individual biokinetic processes such as absorption, distribution and clearance whereas more complex models integrate these processes to predict whole body biokinetic effects. Classical PBBK approaches were developed to model measured complex endpoints such as *in vivo* plasma-time profiles, but PBBK models can also be used to predict the behaviour of a xenobiotic in new situations or to rank compounds in terms of their biokinetic properties. The combination of PBBK modelling with *in vitro–in vivo* extrapolation (PBBK-IVIVE) allows virtual populations to be established by building up mechanistic and physiologically based pharmacokinetic models which incorporate known variability in demographic and biological factors in a 'bottom-up' approach which yields better results than allometric scaling.[20] The factors which should be considered during this process

[20] For a well-written, widely cited review, see Rostami-Hodjegan and Tucker (2007) and for an enthusiastic summary of the current state of play (in 2012) see Rostami-Hodjegan (2012).

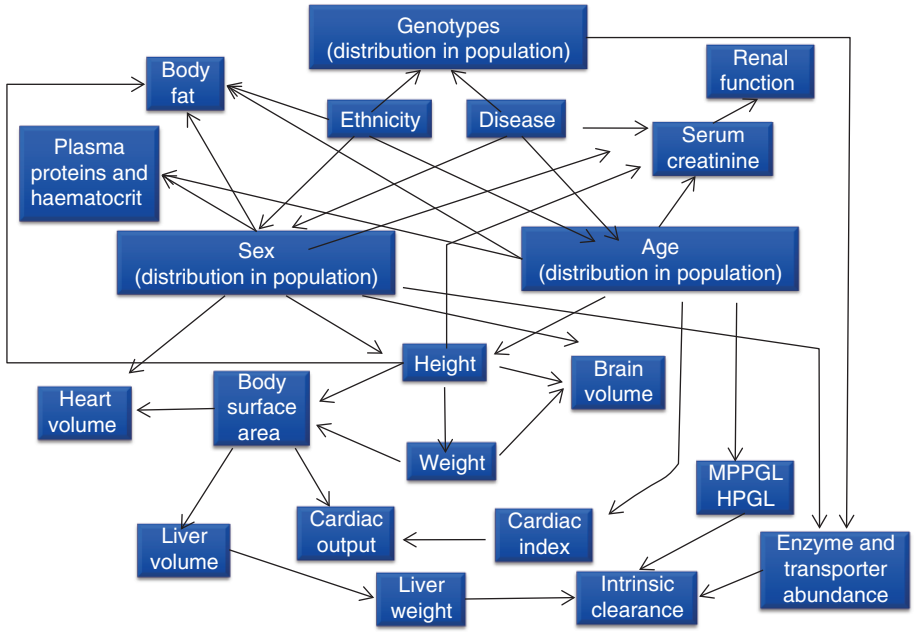


Figure 6.3 The inter-correlation of covariates affecting a chemical's clearance from the body (source: Bois *et al.* (2010); figure 2. Reproduced with permission of Elsevier)

are illustrated in Figure 6.3. These drug-independent parameters can, in turn, be linked to drug-specific physicochemical properties and ADME parameters including CL_{INT}, hepatic blood flow, serum/plasma protein binding and (where available) the activities of uptake and efflux transporters, allowing population variability in ADME to be predicted and individuals at risk to be identified before administering a new drug to humans. In addition, comparison of the results of PBBK modelling with those of *in vivo* studies can give an indication as to the accuracy of predictions made using *in vitro* methods: for example, if clearance predictions based on microsomal incubations correlate closely with actual *in vivo* values this suggests that the main route of elimination is via CYP-mediated metabolism. Furthermore, the PBBK-IVIVE approach allows potential variability in clearance to be simulated, thus reducing uncertainty when predicting the consequences of human exposure. This may be crucial in identifying individuals who might exhibit susceptibility to DDIs or toxicity.

The increasing popularity of PBBK modelling, together with developments which have made user interfaces much easier to work with, have led to a number of companies offering PBBK modelling packages on a commercial basis. The UK company Cyprotex[21] offers two such packages, Cloe PK® (for a full description, see Leahy (2006)) and Cloe HIA® (Thomas *et al.*, 2008), while the German company Bayer has PK-Sim®[22] (Willmann *et al.*, 2005; Willmann *et al.*, 2012),

[21] http://www.cyprotex.com/home/
[22] http://www.systems-biology.com/products/pk-sim.html

a PBBK package which can be integrated with Bayer's graphical modelling tool MoBi® to permit modelling of biological effects at any level from whole body to the molecular (Eissing *et al*., 2011). Perhaps the most widely known and respected offerings, however, are GastroPlus™ (from Simulations Plus) and the Simcyp[23] family of simulation modules.

GastroPlus™ The capabilities of GastroPlus™ have been expanded significantly since the days when it was just used to predict intestinal absorption of orally administered compounds. According to the Simulations Plus web site: 'Since 1997, Simulations Plus has evolved the ACAT™ Advanced Compartmental Absorption and Transit model to a high state of refinement, providing the industry's most accurate, flexible, and powerful simulation program. This smoothly integrated platform combines a user-friendly interface with sophisticated science to help you make better project decisions … faster!' GastroPlus™ Version 8.5 has a wide range of modules permitting mechanistically based simulation of intravenous, oral, oral cavity, ocular, intranasal and pulmonary absorption, pharmacokinetics and pharmacodynamics in humans and animals.[24]

The use of GastroPlus™ to predict the intestinal metabolism of CYP3A substrates and its effect on oral bioavailability has started to be explored recently (Heikkinen *et al*., 2012; Sinha *et al*., 2012). The results obtained suggest that generic PBBK modelling software packages such as GastroPlus™ can be used to predict the oral bioavailability of compounds where intestinal metabolism plays a significant role, although the precision of the predictions is limited by the precision with which input parameters such as intrinsic metabolic clearance and human jejunal effective permeability are measured (Heikkinen *et al*., 2012). However, it does seem to be possible to take advantage of known human liver data for CYP3A4-dependent metabolism (with appropriate scaling factors) rather than having to incorporate measured rates of intestinal metabolism *in vitro*, at least for compounds with good aqueous solubility.

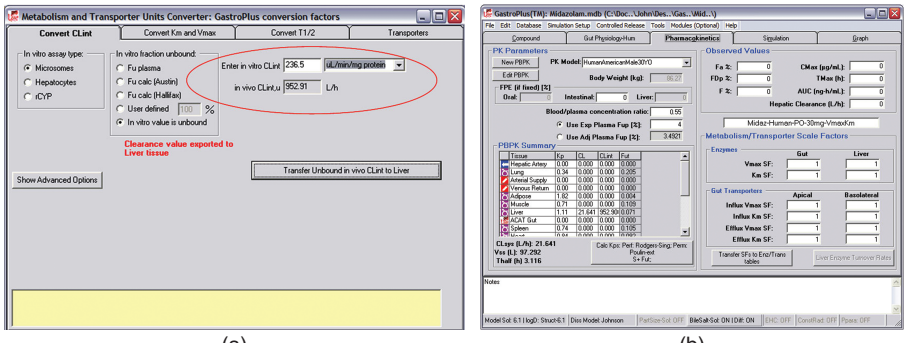
Example: Use of GastroPlus™ to predict first-in-human outcomes from *in vitro* data

Figure 6.4 illustrates the results obtained using GastroPlus™ Version 8.5 to predict the pharmacokinetics of midazolam (as a representative compound which is a CYP3A4 substrate) from *in vitro* data.

- Given: Measured *in vitro* physicochemical properties and human microsomal intrinsic clearance for midazolam.
- Predict: First-in-human outcomes (Fa%, F%, and plasma concentration-time profile) for a 30-mg oral tablet dose.
- Methods: The measured data for midazolam are entered into the program:
  - Solubility = 0.00103 mg ml^{-1} @ pH 7.8
  - logP = 2.8; pKa (base) = 6.15

[23] http://www.simcyp.com/

[24] Details may be found in the GastroPlus™ Version 8.5 brochure, available from http://www.simulations-plus.com/images/pdf/gastro8.5brochure.pdf.



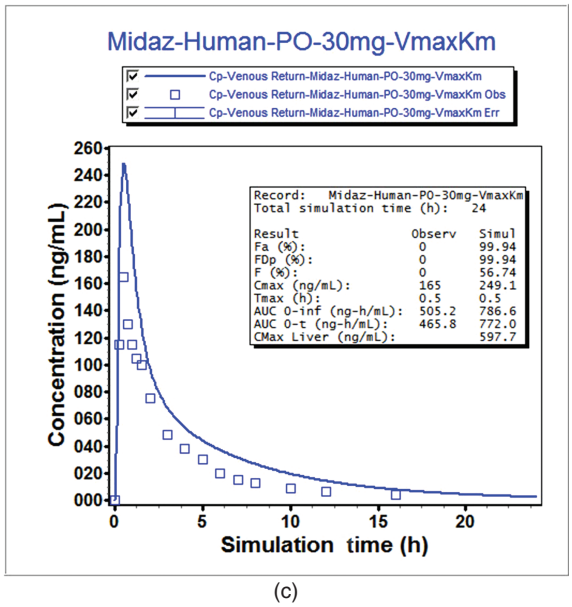


Figure 6.4 Use of GastroPlus™ to predict first-in-human outcomes from *in vitro* data (source: Figures and text were reproduced from GastroPlus™ 8.5 documentation, with permission from Simulations Plus, Inc. (www.simulations-plus.com))

- Fraction unbound in plasma (%) = 4
- Blood-plasma concentration ratio = 0.55
- P_{eff} (from chemical structure) = 6.4E-4 cm s^{-1}
- The measured *in vitro* human microsomal intrinsic clearance is 236.5 µl min^{-1} mg^{-1} protein.

Using the built-in Units Converter tool (shown on page 6 of the GastroPlus™ Version 8.5 brochure), we convert the *in vitro* microsomal intrinsic clearance to an *in vivo* CL_{INT} as shown in Figure 6.4a. This value is transferred to the liver in the default 30-year-old American male PBBK model as shown in Figure 6.4b.

- Results: Fa% = 100, F% = 56.7% (literature predicts ~ 35–40%).
  The predicted C_{max} and AUC are within twofold of the observed values, which are acceptable for a first-in-human prediction, as shown in Figure 6.4c.

Simcyp The Simcyp project, a co-operative venture involving a consortium of Pharma companies and academics, had the objective of developing a user-friendly PBBK program and database with simple visual outputs that could be used as a framework for integrating *in vitro* data and predicting human *in vivo* pharmacokinetics in virtual patient populations to permit modelling of mechanistic IVIVE and metabolic DDIs. The advantages of this type of consortium-based approach, which also involved regulatory agencies and academic centres of excellence, are that it encourages precompetitive data sharing and allows industry guidance on further development of the package. Simcyp was spun out from the University of Sheffield in 2001 and is now a fully fledged commercial operation with three main offerings:

- **The Simcyp Population-based Simulator** (Jamei *et al*., 2009; Jamei *et al*., 2013) is a modular platform and database used for model based drug development based on mechanistic modelling and simulation of oral absorption, tissue distribution, metabolism and excretion in both healthy and diseased populations. It also allows the prediction of metabolic DDIs. The unique feature of the Simcyp Population-based Simulator is that it allows the prediction of the population distribution (median values and variability) of clearance, including individuals at the extremes of the population range, rather than just characterising a single 'average' individual. The modular structure of the package comprises a data store containing population data and compound data. By integrating this information with *in vitro* data and combining data for up to four populations it allows the prediction of pharmacokinetic behaviour in 'real-world' populations, providing either a single, steady-state estimate of predicted plasma concentration (PK Parameters) or a simulated time course (PK Profiles). The data are delivered in Microsoft® Excel® format, making it easy to integrate with other applications.
  The incorporation of data on enzyme inhibition and induction data allows the Simulator to be used to predict metabolic drug–drug interactions, and it can also be used to address ethnic differences in pharmacokinetics, although the predictions made in ethnic populations may be less reliable than for Caucasian populations because of a shortage of relevant data on parameters such as hepatic protein concentration, cardiac output and intestinal CYP3A concentration in non-Caucasian populations.
  Simcyp describes the Simulator as being fast, user-friendly, up-to-date, transparent, cost-saving and time-saving. In independent assessments it was found to provide good predictions of oral bioavailability of an example of a lipophilic investigational drug (Sinha *et al*., 2012) and the magnitude of DDIs for mechanism-based CYP3A inhibitors in cases where substrate pharmacokinetics could be predicted adequately (Wang, 2010).
  A new Simulator module is released every year in response to advice from the Simcyp Consortium, which has responsibility for ensuring that its development is in line with user requirements. This process has steered the evolution of the Simulator from a simple, static model for predicting DDIs to a full PBBK model (Figure 6.5). At the time of writing, Simcyp is in the progress of extending the Simulator further to allow PBBK modelling following dermal application (Polak *et al*., 2012) and in pregnant populations

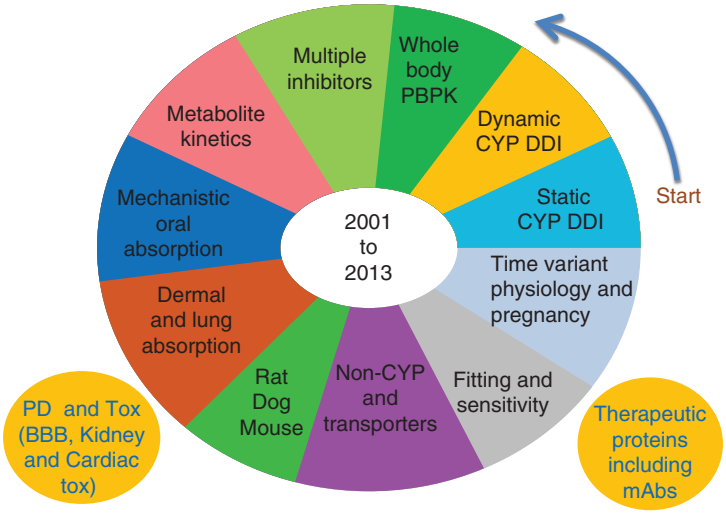


Figure 6.5 The chronology of expansion of the Simulator features from 2001 to 2013 under the Simcyp Consortium guidance. The development started with static metabolic drug–drug interaction calculations then dynamic drug–drug interaction models followed by whole body PBPK and so on (source: Jamei *et al.* (2013); Figure 1)

(Abduljalil *et al.*, 2012; Gaohua *et al.*, 2012). The company is also evaluating its use to predict changes in pharmacokinetics in morbidly obese populations following bariatric surgery (Darwich *et al.*, 2012) and working on adapting it to handle data on therapeutic proteins.

- **The Simcyp Paediatric Simulator** allows pharmacokinetic behaviour to be modelled in infants, neonates and children. This provides valuable information relevant to first-time dosing decisions and the design of clinical studies. It allows population variability in pharmacokinetics to be simulated over any age range and potential drug-drug interactions to be quantified. Predictions can be made either from *in vitro* data, or from adult *in vivo* values by retrograde modelling.
- **Simcyp Animal** is a whole-body PBBK modelling platform for rat, dog and null mouse which is available to Simcyp Consortium members and associates. This can be used to assess pharmacokinetic properties, evaluate formulation and food effects on drug absorption, predict concentration-time profiles in plasma, tissues and organs and investigate the formation and kinetics of primary metabolites. Simulations with Simcyp Animal have been used to increase confidence in PBBK-IVIVE before moving to human simulations. When used in conjunction with the Simcyp Population-based Simulator, Simcyp Animal allows comparison of human and animal data without relying upon allometric scaling. Simcyp Dog is particularly suited to the study of oral drug absorption, while Simcyp's null mouse simulator allows users to investigate how the addition or removal of specific genes controlling drug metabolising enzymes and transporters affects the ADME properties of a drug.

Modern PBBK models such as those provided by Simcyp have utility in a variety of contexts that include

- **Predicting the oral bioavailability of problem compounds:** Lipophilic compounds represent an intractable problem for the prediction of oral bioavailability.

Example: Prediction of oral bioavailability using GastroPlus™ and Simcyp in combination

A case study using a lipophilic investigational drug candidate which is a known CYP3A4 substrate (Sinha *et al*., 2012) illustrates these problems. In this study, GastroPlus™ was used to predict pharmacokinetics in the rat and Simcyp was used to predict human bioavailability; both predictions were then compared with *in vivo* measurements. When GastroPlus™ was used to predict pharmacokinetics in the rat, the nanosuspension could be modelled as a liquid but not as a suspension, but when Simcyp was used to predict human oral bioavailability it could only be modelled as an immediate release formulation, not as a liquid. This was thought to be because of the setting in Simcyp which assumes complete absorption of fully dissolved compounds; the Simcyp predictions also had to be adjusted to take into account bile micelle solubilisation, segmental solubility, colonic absorption and optimisation of the predicted free fraction of drug within enterocytes. Subject to these issues, however, the predictions of both software packages fitted the experimental data well.

- **Prediction of DDIs:** Dynamic models such as those generated by PBBK modelling are increasingly being used to evaluate the potential for drug candidates to become involved in DDIs; indeed, the FDA has recently proposed a decision tree comprising various approaches including PBBK modelling for the prediction of drug–drug interactions.[25]
- **Development of paediatric study plans:** Drugs have historically been used off-label (i.e. in an unapproved age group, dose or formulation) in paediatrics, but this is risky because simple weight adjusted dose adjustment is not enough to ensure efficacy and safe dosing in children. Historically, PBBK modelling has played a key part in paediatric environmental toxicity assessment and it is increasingly being used to support paediatric drug development projects, although a degree of scepticism still remains within the paediatric medical community.[26] Paediatric PBBK modelling cannot currently replace key clinical trials if it is feasible to perform them, especially given that adult safety profiles cannot be transferred directly to children even when the disease processes involved are the same, but the use of PBBK modelling in paediatric applications is now recognised by the FDA,[27] and improved analytical

[25] http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm080499.htm

[26] For detailed coverage of this topic, including recent progress and remaining issues, see Barrett *et al*. (2012).

[27] http://www.fda.gov/downloads/AboutFDA/%2520CentersOffices/CDER/ManualofPoliciesProcedures/ucm073007.pdf

technologies offer the potential to use micro-dosing approaches to generate actual pharmacokinetic data in support of paediatric PBBK modelling.

- **Effects of organ impairment:** Changes such as those typical of renal impairment can be incorporated into PBBK models, although the consequences of multiple impairments can be difficult to model because of insufficient availability of clinical data for use in model validation.
- **Prediction of drug disposition** in pregnant women and in patients with rare diseases, where *in vivo* studies would be unethical and/or impractical (Huang and Rowland, 2012).

In order for PBBK modelling information to be included in a regulatory submission, a certain amount of methodological detail must be provided (Zhao *et al.*, 2012). This includes, but is not limited to: workflow of model construction, verification, modification (if any) and application; information on the input parameters; software details (minimally name and version); details of the experimental design of the simulations and sensitivity analysis.

This discussion has largely focussed on applications of PBBK in drug development because many of the key developments have been driven by the needs of the pharmaceutical industry; however, PBBK modelling also has great potential in improving the assessment of risk to humans subjected to environmental exposures. By integrating chemical-specific dosimetry data with physiological parameters and constraints, PBBK modelling can be used to reduce uncertainty in extrapolation and the resulting human risk estimates can be made more reliable. In order to achieve this, the models used must be subjected to careful QA – 'being published is not enough' (McLanahan *et al.*, 2012) – but subject to thorough review, preferably involving a QA check of the model code and a dialogue between the model developer and evaluator, published PBBK models may make a significant contribution to human health risk assessment, especially under circumstances where there is no time to develop a custom model *de novo* for use in risk assessment.[28]

Evaluation Modern PBBK models provide mechanistic, dynamic, time-based, assumption-testable profiles based on both drug- and system-dependent parameters (Huang and Rowland, 2012). User-friendly commercial packages are now available, meaning that pharmacologists who are not primarily modellers can use PBBK tools to tackle drug development and regulatory issues (Zhao *et al.*, 2012), but there is still a need for more *in vivo* and *in vitro* data to support further model development and ongoing refinement will continue to be required.

6.4 Toxicity

The factors which influence target organ toxicity include toxicokinetics, biotransformation and toxicodynamic factors. The tissue accumulation of xenobiotics is a function of their solubility characteristics. Water-soluble polar compounds are eliminated via the kidneys, bile, sweat and intestinal secretions.

[28] For a detailed discussion of the criteria which should be applied, see McLanahen *et al.*, 2012.

In contrast, the excretion of lipophilic compounds is hampered because they can be sequestered in body fat, so that at equilibrium the ratio of lipophilic chemicals in blood:body fat is extremely low. In addition, although lipophilic compounds are excreted into the gut or into kidney tubules they are rapidly reabsorbed and re-enter the circulation. However, metabolism in more polar compounds can enhance their elimination from the body.

6.4.1 Exposure modelling

Adults are primarily exposed to xenobiotics via the diet, inhalation and skin exposure. High level exposures can occur in the workplace, and health and safety regulations are in place to minimize this type of exposure. Low level exposures via food, water and air are more difficult to model and control.

Mathematical exposure modelling is an indirect method of determining exposure, particularly human exposure to environmental contaminants. It is useful when direct measurement of pollutant concentration is not feasible due to the need for continuous monitoring using complex and expensive laboratory equipment. The ability to make inferences in the absence of direct measurements makes exposure modelling a powerful tool for predicting exposures by exploring hypothetical situations.

Two critical pieces of information are needed in order to calculate human exposure:

- the whereabouts of the individual or individuals being exposed;
- the concentration of the pollutants in the different locations.

These can be expressed mathematically in terms of the time spent by a person in those different locations multiplied by the time-averaged air pollutant concentrations occurring in those locations. In the absence of data obtained from direct observation, human activity pattern data may be used. These data can be used to create inhalation exposure models which can serve as useful public health tools for epidemiology, education, intervention, risk assessment, and creation of air quality guidelines.

The exposure of children to xenobiotics, whether deliberately (e.g. pharmaceuticals) or inadvertently (e.g. food additives, contaminants and pollutants) is a potential cause for concern and an area where modelling approaches may play a key role, since it is not ethically appropriate to conduct even low dose volunteer studies in small children. However, the state of understanding regarding the disposition of xenobiotics in children lags far behind that in adults for a number of reasons (Ginsberg *et al.*, 2004):

- Ethical considerations mean that it is inappropriate to test xenobiotics in children even using microdose protocols.
- Children's biokinetics differ from those in adults due to
  - smaller body size;
  - different ratios of fat:muscle:water;
  - higher breathing and metabolic rates;
  - immaturity of enzymes and clearance systems.

- It is not possible to generalise across age groups because of the wide diversity in general physiology and xenobiotic disposition across the age range from the neonate to the adolescent.
- New methodologies for the prediction of xenobiotic disposition (e.g. PBBK modelling) have yet to be applied systematically to risk assessment for children.

Predictions in the neonate are particularly difficult because of rapid changes in liver and kidney function and anatomical remodelling of the vasculature which occur immediately after birth, although efforts have been made to create general PBBK models for drug and xenobiotic disposition in infants and children.

In order to fill these data gaps, the increased use of modelling approaches using toxicokinetic data from animals and adults and/or surrogate chemicals for which data are available in children has been recommended (Ginsberg *et al.*, 2004). In order for this approach to be successful it is necessary to take into account age-dependent changes in body weight, tissue volume, blood flow and bone formation/resorption rates. When physiological parameters from animal studies are used as the basis for extrapolation to humans, additional uncertainty is incurred because of the difficulty in extrapolating from adult animals to children. Data from juvenile animals cannot readily be used for this purpose as this would increase uncertainty in the model. An additional difficulty is the fact that there is no easy way to identify the stage of development at which significant changes in pharmacokinetics occur, meaning that the points at which borderlines are drawn are arbitrary. This makes it necessary to consider very small age ranges especially when making predictions about xenobiotic disposition in early postnatal stages.

6.4.2 Prediction of toxicity

Various commercially available packages provide the basic tools for *in silico*/QSAR approaches to the prediction of toxicity.[29] These can be used for REACH classification and to meet the requirements of the 7th Amendment, and can also be useful in evaluating the potential toxicity (especially genotoxicity and carcinogenicity) of pharmaceutical impurities (Fioravanzo *et al.*, 2012). They include the following:

- **MCASE** (Multiple Computer Automated Structure Evaluation; also available as a Windows-based version, MC4PC) is a machine-learning tool based on a hierarchical analysis of molecular moieties with specific biological functionality (pharmacophores, toxicophores or so-called biophores) (Saiakhov and Klopman, 2008). It was originally developed in 1982, released for general use in 1996 and adopted by the FDA in 1998. This software is theoretically able to predict any kind of biological effect, although it was designed with toxicological endpoints in mind. It has the capacity to generate 'expert knowledge' directly from crude experimental data, with no prior toxicological expertise required. It does not provide mechanistic information as such, but the identification of biophores often suggests a mechanistic interpretation

[29] For an overview in the general context of computational toxicology, see Rusyn *et al.* (2012).

and there is a partner program, MetaPC, which can help with this. It is set up in such a way as to make validation easy; the user can either set up custom validation runs or make use of the generic validation report which is provided within the package. MCASE includes a number of modules which are directly relevant to REACH assessment including ADME, irritation, acute/subchronic/chronic toxicity, genotoxicity, carcinogenicity, developmental toxicity and environmental endpoints (Saiakhov and Klopman, 2008). Its use is also recommended by the Danish Environmental Protection Agency and US FDA (Fioravanzo *et al*., 2012).

- **Derek Nexus**, the successor to Derek for Windows, was released by Lhasa Limited in 2009.[30] Lhasa Limited has also responded to the ICH M7 guidelines addressing the use of *in silico* tools to predict potential genotoxic impurities in pharmaceuticals, which require the use of two *in silico* methods, by developing Sarah Nexus as an additional tool complementary to Derek.

  Derek is an expert system which uses a knowledge base comprising alerts, examples and rules for the prediction of toxicity, allowing users to gather information about possible modes of action and providing supporting data in terms of comments, literature references, and toxicity results. Derek provides outputs in a variety of formats (Figure 6.6; Table 6.1). One of the advantages of this approach is that it encourages the sharing of proprietary data in order to validate predictions and allows non-confidential knowledge to be derived from anonymised proprietary data without compromising commercial sensitivity. Derek also links directly to Vitic, a chemically intelligent toxicity database designed to allow efficient access to pre-existing toxicology information which covers many traditional areas of toxicological investigation such as carcinogenicity, mutagenicity, sensitisation and skin irritation and incorporates hepatotoxicity data accumulated from peer-reviewed journals. It is claimed to have the potential to incorporate in-house data on other aspects of toxicity, making it amenable to development in ways relevant to the needs of a particular user group or industry; however, as with all *in silico* systems, the usefulness of Derek depends on the issue under investigation and whether the chemicals under consideration fall within the Applicability Domain of the program.
- **HazardExpert** is another knowledge-based system which identifies toxicophores and applies toxicological knowledge, QSAR models and fuzzy logic to make predictions which are expressed in a format similar to that used by the International Agency for Research on Cancer to classify carcinogens (from 1: highly probable toxicity to 4: no toxicity). Its algorithm is accessible to the user, allowing it to be customised.
- **TOPKAT** is a linear QSAR system which was released for general use in 2007. It makes predictions using computer-assisted technology and can provide either a yes/no answer (e.g. for genotoxicity) or continuous measurements for quantitative endpoints (e.g. LD_{50}s). TOPKAT is believed to be the only commercially available software package which can predict a lowest observed adverse effect level (LOAEL) or NOAEL for a chemical from its structure (Rupp *et al*., 2010), although there is still a significant margin of error in the predictions made. In an evaluation of 807 chemicals of known structure, it was

[30] Further information about the history of Lhasa Limited can be found at http://www.lhasalimited.org/about-us.htm.

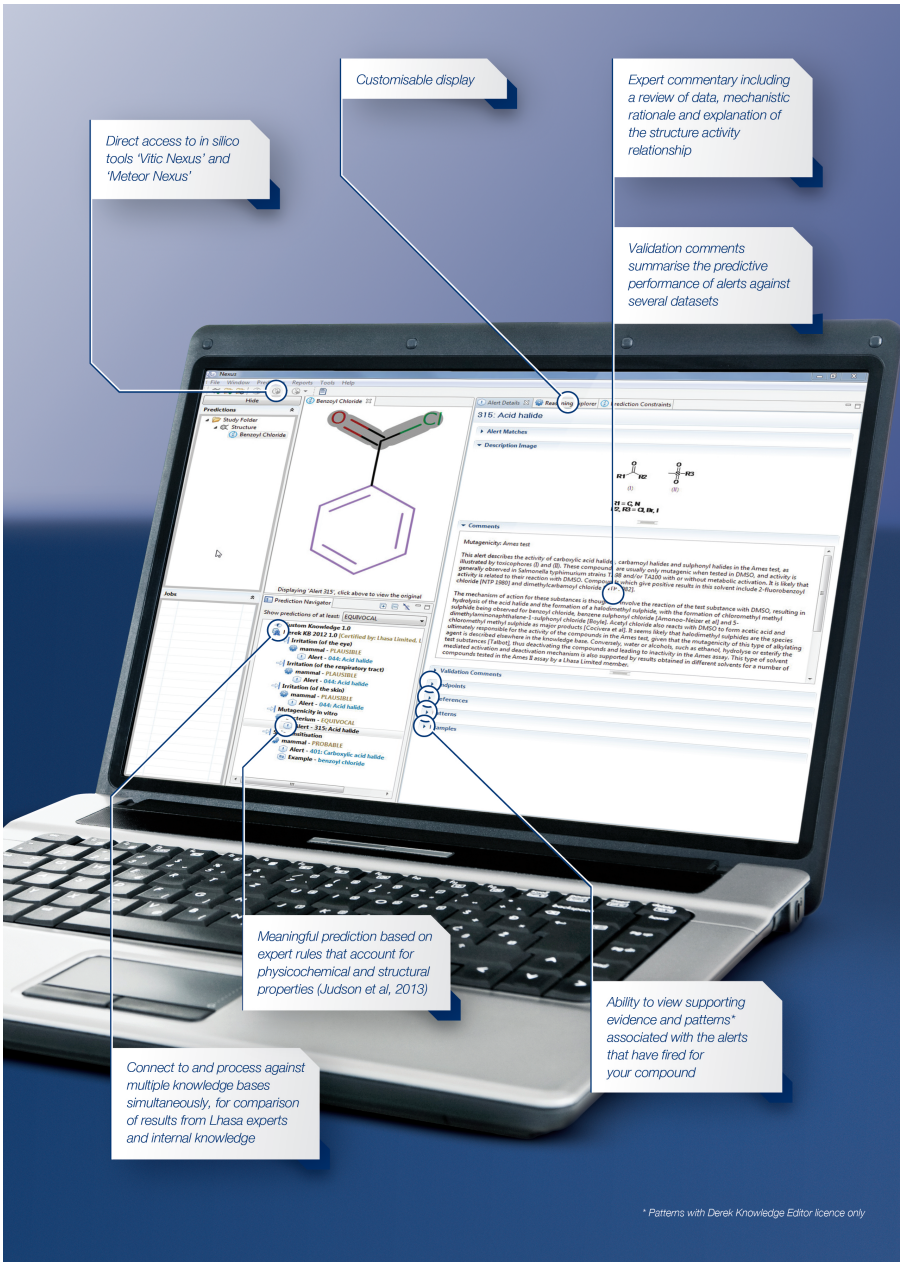


Figure 6.6 Features of Derek Nexus 3.0.1 (source: Lhasa Limited. Reproduced with permission of Lhasa Limited)

Table 6.1 A Derek Nexus summary report on benzyl chloride using the default processing constraints

Derek Nexus Report

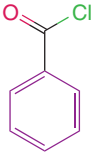
Author
Lhasa Limited

Report date
23 July 2013 13:37:14

Prediction date
23 July 2013 13:17:32

Program version
Derek Nexus: 3.0.1, Nexus: 1.5.1

Submitted compound



Predictions
Derek KB 2012 1.0
KnowledgeBase name
Derek KB 2012 1.0

KnowledgeBase version
1.0

KnowledgeBase last modified
29 November 2012 11:56:30

KnowledgeBase location
C:/Program Files/Lhasa Limited/Lhasa Knowledge Suite – Nexus 1.5/KnowledgeBases /org.lhasalimited.pluto.knowledge.derek_1.0.7

KnowledgeBase certified by
Lhasa Limited, Leeds, Yorkshire, UK

Reasoning summary and alerts found
Irritation (of the eye) in mammal is PLAUSIBLE
044 Acid halide
Irritation (of the respiratory tract) in mammal is PLAUSIBLE
044 Acid halide
Irritation (of the skin) in mammal is PLAUSIBLE
044 Acid halide
Mutagenicity *in vitro* in bacterium is EQUIVOCAL
315 Acid halide
Skin sensitisation in mammal is PROBABLE
401 Carboxylic acid halide

Table 6.1 (*Continued*)

Reasoning glossary:

Certain
There is proof that the proposition is true

Probable
There is at least one strong argument that the proposition is true and there are no arguments against it

Plausible
The weight of evidence supports the proposition.

Equivocal
There is an equal weight of evidence for and against the proposition.

Doubted
The weight of evidence opposes the proposition.

Improbable
There is at least one strong argument that the proposition is false and there are no arguments that it is true.

Impossible
There is proof that the proposition is false.

Open
There is no evidence that supports or opposes the proposition

Contradicted
There is proof that the proposition is both true and false

(source: Lhasa Limited. Adapted with permission of Lhasa Limited)

only possible to use TOPKAT to predict LOAELs for 347 and of these, only 24% of the predictions were within twofold of the true value, although 93% were within two orders of magnitude, leading to authors proposing an additional safety factor of 100-fold when deriving a guidance value such as a tolerable daily intake from a LOAEL or NOAEL predicted by *in silico* methods (Rupp *et al*., 2010). In a comparison exercise (Mombelli, 2008), TOPKAT was better than other programs at identifying true negatives for skin irritation (25/42) but also generated a number of false positives (13/42).

Various open source models such as Toxtree and Toxmatch are also available via the EC Joint Research Centre Institute for Health and Consumer Protection.[31] These can be used for REACH classification by means of read-across based on common functional groups, precursors/breakdown products and changes in properties across a chemical group. Toxmatch uses analogue approaches (based on small groups of chemicals which are too limited to allow trends to be identified) and category analysis (where an extensive range of analogues allows trends to be detected) for systematic read-across and to assist in the development and assessment of QSARs (Patlewicz *et al*., 2008), while ToxTree predicts different types of

[31] http://ihcp.jrc.ec.europa.eu/

toxicological hazard and modes of action by applying decision tree approaches to permit grouping and hazard profiling of chemicals (Pavan and Worth, 2008). An additional program, DART (Decision Analysis by Ranking Techniques) can be used for transparent multifactorial evaluation including cost-benefit analysis. In addition, the OECD provides a freely accessible QSAR Toolbox (The OECD QSAR Toolbox for Grouping Chemicals into Categories[32]) to facilitate the use of QSAR analysis for REACH and other purposes. The Toolbox serves as a platform that incorporates various modules and databases from other sources, including ToxTree and several MCASE modules (Saiakhov and Klopman, 2008, Mombelli and Devillers, 2010).

A novel approach to the prediction of toxicity is becoming possible now that the molecular structures of receptors, enzymes and target proteins are becoming available. This has permitted the development of packages which make predictions based on the binding of chemicals to these proteins. One of these, VirtualToxLab,[33] uses a combination of automated, flexible docking and multi-dimensional QSAR to estimate the toxic potential of drugs, chemicals and natural products. The underlying philosophy of the VirtualToxLab approach is to estimate the toxic potential of a compound in terms of its binding affinity towards a series of proteins which are known or suspected to trigger adverse effects. In order to do this, it makes use of validated models for 16 key proteins: the receptors for estrogens (α and β), progesterone, androgens, glucocorticoids, mineralocorticoids and thyroid hormone (α and β) plus various nuclear receptors, hERG, CYP1A2, CYP2C9, CYP2D6 and CYP3A4. The result is a value, the toxic potential, between 0 (no cause for concern) and 1 (extreme concern) which can be interpreted as a toxicity alert. This approach is attractive in that it is specific and target-oriented; however, it does have a number of limitations. The authors recognise that it ignores issues relating to bioavailability, ADME and access to the target, and that compounds with a low toxic potential as defined by VirtualToxLab may not be benign because they could exert their effects via other targets or mechanisms (Vedani *et al*., 2012). It is also important to highlight that this approach considers only a very limited number of potential molecular targets and, perhaps most importantly, that it is purely a hazard-based approach (that is, it only considers the possibility of toxicity, not whether there is any risk that this would actually occur in the real world). The VirtualToxLab program is freely available to researchers in universities, government and regulatory agencies and not-for-profit organisations (OpenVirtualToxLab[34]) and the regular VirtualToxLab program can be accessed 'at a modest fee' (Vedani *et al*., 2012).

Acute toxicity The 7th Amendment ban on the use of animals for toxicity testing of cosmetic ingredients has stimulated efforts to develop *in silico* prediction methods. The ideal model needs to consider an adequate number of molecules to allow sufficient statistical representation of all possible structures. In addition, the range of compounds included should represent a wide range of quantified toxic potencies and be amenable to mechanistic interpretation.

[32] http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm
[33] http://www.biograf.ch/index.php?id=home
[34] http://www.biograf.ch/data/projects/OpenVirtualToxLab.php

Open Source tools such as Toxmatch and ToxTree can be used to screen for the potential to induce various types of dermal toxicity including skin irritation, corrosion and sensitisation (Gallegos-Saliner *et al.*, 2008; Scheel and Keller, 2012). These can identify structural alerts for key processes such as protein binding together with indications of possible modes of action; however, because protein binding, though predictable, is only the first step in the adverse outcome pathway they can only provide yes/no answers with no indication of potency (Scheel and Keller, 2012). The same caveat also applies to commercial packages such as Derek, TOPKAT and MCASE.

Example: Prediction of skin irritation using a QSPR approach

In an effort to develop a skin irritation model for a wide range of compounds a QSPR has been developed using rabbit Draize test data for 189 chemicals of diverse classes (Golla *et al.*, 2009). The QSPR model was initially developed using exclusively structural descriptors and then expanded using literature data and functional group information to improve the predictive ability of the model. It allowed skin irritation to be predicted with an r^2 value of 0.78 and the final set of descriptors chosen exhibited a good correlation with molecular size, reactivity and skin penetration. However, because the accuracy and reliability of any *in silico* method is dependent upon the quality of the experimental data used to develop it, the authors of the study noted (Golla *et al.*, 2009) that more human-like models for dermal effects are required.

Example: Use of *in silico* packages to predict skin sensitisation

Two *in silico* tools, Derek and the OECD Toolbox, have been compared in terms of their ability to detect skin sensitisation (Goebel *et al.*, 2012). When their ability to predict the skin sensitising potential of 249 chemicals was compared, neither showed good concordance with the murine local lymph node assay: the concordance for DEREK (Version 10) was 63% while that for the OECD Toolbox was only 53%. Worryingly, both yielded numerous false positives; however, negative predictivity reached up to 70%, suggesting that it may be possible to incorporate QSAR as part of a combination of approaches to obtain information for priority setting, and using the OECD Toolbox with the data gap filling function did improve the overall concordance of the predictions with the LLNA to 77%. Overall, the conclusion was that *in silico* approaches do have potential as part of a weight of evidence approach (Goebel *et al.*, 2012).

The importance of hepatotoxicity as a reason for failure during drug development, as well as its status as a common finding during chemical risk assessment, means that there is an urgent need for reliable methods of predicting this outcome *in silico*; however, the literature to date contains relatively few convincing examples of the successful application of the approach.

Overall, QSAR programs and structural alerts are considered to perform well for compounds with low human exposure levels and have the potential to supplement the results of routine tests (e.g. for genotoxicity) as part of a weight of evidence approach (Mayer *et al.*, 2008). A collaborative program aimed at predicting organ toxicity from QSARs based on histopathological findings (part of the MetaTox consortium) succeeded in predicting specific endpoints such as liver

necrosis and liver enlargement but struggled to deal with complex outcomes such as hepato- and nephrotoxicity (Myshkin *et al*., 2012). It is also worth noting that the use of two or more programs to address the same question can improve the reliability of the predictions made; by pooling all the positive results sensitivity can be increased, or by looking for consensus predictions better confidence can be achieved (Fioravanzo *et al*., 2012; Matthews *et al*., 2008; Matthews *et al*., 2009b).

Example: Use of principal component analysis to predict hepatotoxicity

Efforts to model *in vitro* lactate dehydrogenase release using the Soft Independent Modelling of Class Analogy approach accurately predicted the hepatotoxicity of more than 90% of a set of compounds of known hepatotoxicity, although it was not tested against an unknown data set (Clark *et al*., 2004). The limited applicability and lack of independent assessment of this model provide no indication of wider potential use. Dose-dependent human hepatotoxicity has also been predicted using a decision tree algorithm based upon 25 physicochemical properties (Cheng and Dixon, 2003). The model was refined to minimise erroneous classification and allocated each chemical to either hepatotoxic or non-hepatotoxic categories. This approach predicted dose dependent human hepatotoxicity with an accuracy of around 80%.

Example: Use of neural networks and molecular structure-based methods to predict toxicity

A further approach, using neural networks and computational techniques to characterise molecular structure, has been applied to a range of end-points including mode of action, gene expression, hepatotoxicity and neurotoxicity (Piotrowski *et al*., 2007). The method attempted to capture the entire information about the structural and electronic properties of the molecules using 30 variables, although it was noted that, of these, 18 were sufficient for predicting toxicity. The mode of ecotoxicological action of test substances was predicted with 95% classification accuracy and classification of hepatotoxic chemicals into acute or secondary liver toxicity was achieved with 100% accuracy. This performance was not adversely affected by the inclusion of additional chemicals showing no hepatotoxicity. Additional studies to predict lesion types were 97% accurate for necrosis, 96% for fatty liver, 98% for cirrhosis, 94% for carcinoma and 98% for cholestasis. The predictive model for neurotoxicity was based on BBB partitioning. The results of this analysis gave a correlation coefficient of 0.86 between the observed and expected values.

The same approach has been applied to microarray data and correctly predicted gene expression levels for all but two of the test compounds. Although integration of microarray and *in silico* approaches is a logical development of these techniques another use of this approach with a number of drugs as model compounds (Ekins *et al*., 2006) yielded variable results, indicating the need for more refinement and development before this method will be of practical applicability.

Example: *In silico* prediction of hepatobiliary and urinary adverse effects

More recently, the utility of *in silico* methods to predict hepatobiliary and urinary adverse effects has been evaluated by a team at the FDA Center for Drug Evaluation and Research (Matthews *et al*., 2009a, Matthews *et al*., 2009b). Adverse events

were clustered into five types of liver injury: liver enzyme disorders, cytotoxic injury, cholestasis and jaundice, bile duct disorders, and gall bladder disorders, and six types of urinary tract injury: acute renal disorders, nephropathies, bladder disorders, kidney function tests, blood in urine, and urolithiasis. The team evaluated MC4PC in comparison with three other QSAR programs (MDL-QSAR,[35] BioEpisteme[36] and Leadscope Predictive Data Miner[37]). When used on its own, MC4PC performed well in terms of high specificity (>90%) and a low false positive rate (<10%), although its sensitivity was poor (~20%). Derek for Windows, which includes expert rules for the prediction of hepatotoxicity in animals and humans, was also used in this study (Versions 9 Knowledge 2006 and 10.0). The results of this study demonstrated that the Derek for Windows had 92% specificity, but only 20% sensitivity for predicting pharmaceuticals with hepatobiliary adverse events. The analysis was run without prior knowledge of the FDA hepatotoxicity database; the study therefore served as an external validation of Derek for Windows, demonstrating that Derek can be used to create direct links between proposed modes of action and observed hepatic adverse events, although with limited sensitivity (Matthews *et al*., 2009a).

Embryotoxicity and reproductive toxicity Reproductive toxicity is one of the most challenging endpoints for QSAR modelling because of the subtlety and complexity of the endpoints and the shortage of reliable data for model building. Many attempts to develop QSAR prediction of embryotoxicity and reproductive toxicity have been reported, but these have all had limitations either of applicability or accuracy (Matthews *et al*., 2007a, Matthews *et al*., 2007b). The problem is that reproductive and developmental toxicity are complex endpoints with diverse outcomes ranging from gross anatomical abnormalities to subtle behavioural effects. These are often described in purely qualitative terms, making them very difficult to model mathematically. It has also been suggested that the mathematical modelling community is not particularly interested in attempting to model reproductive effects *in silico* (Cronin and Worth, 2008). However, the demands of REACH mean that there is an urgent need for *in silico* methods to predict these important endpoints, especially given that reproductive toxicity testing is believed to account for 60% of animal usage under REACH (Piersma, 2011).

The problem in developing such models is that there is a perceived lack of high quality data and insufficient knowledge of modes and mechanisms of action with which to build them. The available models fall into two categories, so-called "local" models which address only a restricted set of compounds and global models which attempt to model a heterogeneous group of compounds with diverse structures and properties. Local models tend to be very specific ones, often developed in house during the investigation of a particular class of compounds whereas commercial models such as Derek, Topkat, MC4PC and HazardExpert tend to be global in scope. These are able to represent a broad range of chemistries and mechanisms, but tend to be less transparent than the

[35] http://www.akosgmbh.de/Archive/mdl_qsar.htm
[36] No website at present
[37] http://www.leadscope.com/product_info.php?products_id=41

more focussed local models. Read-across approaches such as ToxMatch can also be used to address reproductive effects.

One popular approach to the problem of modelling reproductive and developmental toxicity is to reduce it to the question of whether compounds bind to key receptors such as the oestrogen and androgen receptors, which at least has the advantage that the mechanisms involved are well understood. The CoMFA approach is particularly well suited to the assessment of receptor binding phenomena. The ReProTect project included a small QSAR effort which aimed to predict transfer of compounds across the placental and blood-testis barriers. Some success was achieved with the placental membrane, but modelling of passage across the blood-testis barrier proved more challenging.

Example: Prediction of reproductive toxicity using MC4PC

Multicase MC4PC software and a weight of evidence based approach from a wide range of end-points in different species have been compared in relation to prediction of reproductive toxicity (Matthews *et al*., 2007a, Matthews *et al*., 2007b). It was possible to construct QSARs for male and female reproductive toxicity, foetal dysmorphogenesis, functional toxicity, mortality, growth and newborn behavioural toxicity. Apart from these defined aspects other areas of reproductive toxicity proved difficult to model for various reasons; however, this approach may have potential for further development, both to cover those end-points not so far modelled and to further optimise its performance for specific end-points.

Example: Use of more than one model to enhance the prediction of reproductive toxicity

A weight of evidence approach using five approaches (the CAESAR model developed as part of the EU Caesar project, the Reproductive Toxicology super-endpoint in Derek for Windows, the oestrogen receptor binding profiler in the OECD Toolbox and read-across using ToxMatch) has been shown to generate more reliable predictions than separate use of the individual models (Hewitt *et al*., 2010). In particular, this study found that the performance of Derek in predicting reproductive effects was relatively poor (26.2% correct predictions), similar to that recorded in a previous evaluation of performance of three different approaches (Hulzebos *et al*., 2005). The authors point out, however, that this is not a criticism of Derek as a system; rather, it reflects the paucity of suitable data to drive the model's predictions (Hewitt *et al*., 2010).

In order for mathematical models to be useful in the prediction of reproductive effects, effort is needed to resolve the following issues: the models need to be more mechanistically focussed, the existing models need to be evaluated against the OECD criteria, and more/better data are needed to support improvements. Such efforts are important in order to improve the prediction of reproductive effects and reduce the vast number of animals used in testing to meet the demands of REACH. Overall, current thinking is that the application of alternative methods to reproductive toxicity testing requires the use of a tiered approach with a battery of alternative tests, which could involve mathematical models once these are

more fully developed. One current effort with this aim is the US Environmental Protection Agency (EPA) Virtual Embryo Project (v-embryo™).[38]

Chronic toxicity The majority of methods for predicting chronic mammalian toxicity by QSAR use general end-points rather than organ-specific effects. A review of the use of QSAR as an alternative to animal studies for sensitisation and irritation also summarised the application of QSAR to other areas of toxicity (Gerner *et al*., 2004). For repeat-dose toxicity the conclusion was that QSAR is limited to the prediction of LOAEL values. The origin of this comment appears to be an exercise which used specifically designed TOPKAT QSAR software, based on chemicals representing five structural classes (acyclics, alicyclics, heteroaromatics, multiple benzenes and single benzenes). The model was tested for ability to predict the chronic LOAEL of chemicals from a pesticide dataset and a second mixed dataset of chemicals. Results showed the prediction of LOAEL to be within a factor of 2 for 43% of chemicals, 5 for 66% and 10 for 80%. Although at the time of reporting it was concluded that this QSAR approach had only limited application the concept may have some potential for further refinement in relation to specific chemical groups.

6.5 Conclusions

Numerous software packages are now available for the prediction of ADME properties, particularly passive absorption and CYP-mediated metabolism. These range from generic mathematical modelling packages to highly specialised packages designed to predict the drug-like properties of novel compounds, although it is important to note that there is considerable fluidity in the software market: several packages and companies which are well reviewed in the recent literature can no longer be located by means of Google searching.

The use of QSAR models for prediction of toxicity has been attempted using a wide variety of approaches but these are usually constructed to describe a specific data set and frequently show weaknesses when a different test set is used. Some recent approaches use a more sophisticated approach to description of the molecules studied and thus potentially greater predictive power and independence of the training set. Read-across databases such as those proposed for REACH, and developments of these, which are based upon structural similarity, may have use as a first line of data gathering on any novel chemicals.

For the Replacement of animal testing in the risk assessment of repeated dose exposure a considerable investment is required in basic research to develop and validate a range of *in vitro* models capable of forming part of an integrated approach. QSAR and read across methods need to be refined and overall strategies devised incorporating kinetic and metabolic studies to permit quantitative interpretation of results in terms of consumer risk.

The field of PBBK modelling is developing rapidly and the prospect of applying this approach to the integration of *in vitro* and *in silico* data in order to develop

[38] http://epa.gov/ncct/v-Embryo/

complete ADME models is attractive in cases where the results of preliminary screening are not sufficient for decision making and risk assessment (for example, in cases where significant uptake occurs or hepatic clearance is slow). The processes of distribution and excretion are still difficult to model on the basis of non-animal tests and it is to the understanding of these processes that PBBK modelling could make a significant contribution.

When combined with knowledge about receptors and pathophysiological processes, the PBBK-IVIVE approach represents a significant move towards the coming era of systems pharmacology; however, the cost of developing the huge databases required and the need for continual updating as new information becomes available mean that future progress will be dependent upon the willingness of academic and commercial partners to share data and engage in collaborative programs to develop generic, widely accepted models on common platforms. This is a big challenge, but if achieved it will create an environment of "modelling for all" allowing non-specialist users to address specific questions while expert modelling groups are free to work on further improvements to the approach.

In conclusion, a wide variety of *in vitro* and *in silico* approaches is available for use in predicting the disposition and adverse effects of drugs and other xenobiotics. Individually, none of these can provide a complete picture of this complex process, but by integrating data from several *in vitro* studies and *in silico* models effective predictions can be made. This process is illustrated in Figure 6.7.

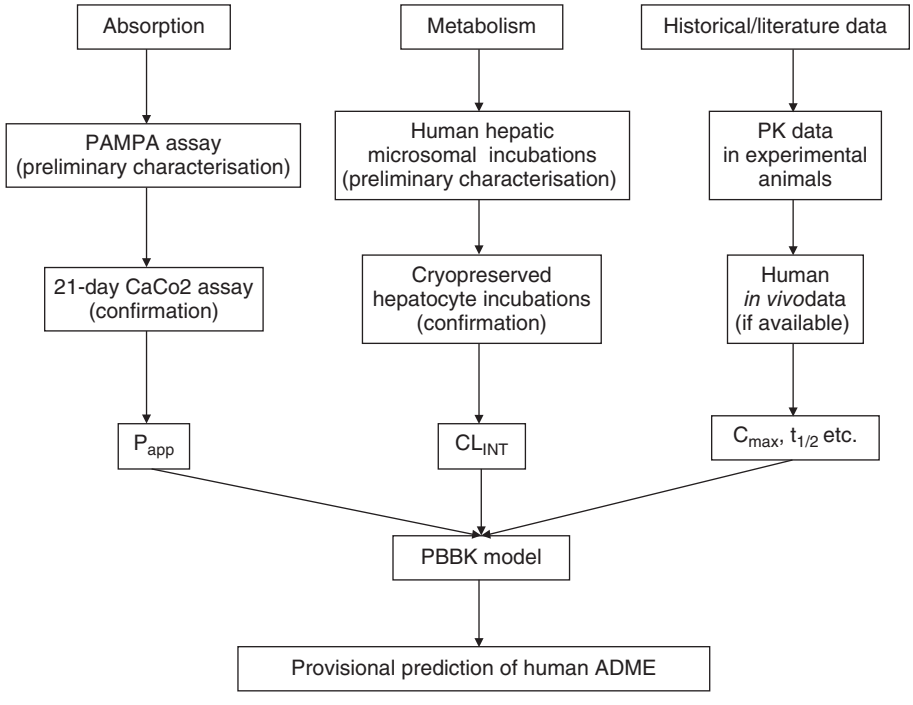


Figure 6.7 Integration of *in vitro* and PBBK approaches for the prediction of ADME

Self-assessment questions

- What factors would you consider when selecting a software package for use in the prediction of xenobiotic uptake and metabolism?
- How can PBBK modelling help to accelerate and reduce the costs of drug development?

References

Abduljalil, K., Furness, P., Johnson, T.N. *et al.* (2012) Anatomical, physiological and metabolic changes with gestational age during normal pregnancy: a database for parameters required in physiologically based pharmacokinetic modelling. *Clin Pharmacokinetics*, **51**, 365–396.

Ako, R., Dong, D. and Wu, B. (2012) 3D-QSAR studies on UDP-glucuronosyltransferase 2B7 substrates using the pharmacophore and VolSurf approaches. *Xenobiotica*, **42**, 891–900.

Balimane, P.V., Chong, S. and Morrison, R.A. (2000) Current methodologies used for evaluation of intestinal permeability and absorption. *J Pharmacol Toxicol Methods*, **44**, 301–312.

Ballante, F. and Ragno, R. (2012) 3-D QSAutogrid/R: an alternative procedure to build 3-D QSAR models: methodologies and Applications. *J Chem Inf Model*, **52**, 1674–1685.

Baranczewski, P., Stanczak, A., Sundberg, K. *et al.* (2006) Introduction to *in vitro* estimation of metabolic stability and drug interactions of new chemical entities in drug discovery and development. *Pharmacol Rep*, **58**, 453–472.

Barrett, J.S., Della Casa Alberighi, O., Laer, S. and Meibohm, B. (2012) Physiologically based pharmacokinetic (PBPK) modeling in children. *Clin Pharmacol Ther*, **92**, 40–49.

Bois, F. Y., Jamei, M. and Clewell, H. J. (2010). PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology*, **278**, 256–67.

Boobis, A., Gundert-Remy, U., Kremers, P. *et al.* (2002) *In silico* prediction of ADME and pharmacokinetics. Report of an expert meeting organised by COST B15. *Eur J Pharm Sci*, **17**, 183–193.

Bouwman, T., Cronin, M.T., Bessems, J.G. and Van De Sandt, J.J. (2008) Improving the applicability of (Q)SARs for percutaneous penetration in regulatory risk assessment. *Hum Exp Toxicol*, **27**, 269–276.

Brown, M.B., Lau, C.H., Lim, S.T. *et al.* (2012) An evaluation of the potential of linear and nonlinear skin permeation models for the prediction of experimentally measured percutaneous drug absorption. *J Pharm Pharmacol*, **64**, 566–577.

Cheng, A. and Dixon, S.L. (2003) *In silico* models for the prediction of dose-dependent human hepatotoxicity. *J Comput Aided Mol Des*, **17**, 811–823.

Clark, R.D., Wolohan, P.R., Hodgkin, E.E. *et al.* (2004) Modelling *in vitro* hepatotoxicity using molecular interaction fields and SIMCA. *J Mol Graph Model*, **22**, 487–497.

Clewell, R.A. and Clewell, H.J. III (2008) Development and specification of physiologically based pharmacokinetic models for use in risk assessment. *Regul Toxicol Pharmacol*, **50**, 129–143.

Coecke, S., Ahr, H., Blaauboer, B.J. *et al.* (2006) Metabolism: a bottleneck in *in vitro* toxicological test development. The report and recommendations of ECVAM workshop 54. *Altern Lab Anim*, **34**, 49–84.

Cronin, M.T.D. and Worth, A.P. (2008) (Q)SARs for predicting effects relating to reproductive toxicity. *QSAR Combinat Sci*, **27**, 91–100.

Darwich, A.S., Pade, D., Ammori, B.J. *et al.* (2012) A mechanistic pharmacokinetic model to assess modified oral drug bioavailability post bariatric surgery in morbidly obese patients: interplay between CYP3A gut wall metabolism, permeability and dissolution. *J Pharm Pharmacol*, **64**, 1008–1024.

Davis, A.M. and Riley, R.J. (2004) Predictive ADMET studies, the challenges and the opportunities. *Curr Opin Chem Biol*, **8**, 378–386.

Dearden, J.C., Cronin, M.T. and Kaiser, K.L. (2009) How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). *SAR QSAR Environ Res*, **20**, 241–266.

Dong, D., Ako, R., Hu, M. and Wu, B. (2012) Understanding substrate selectivity of human UDP-glucuronosyltransferases through QSAR modeling and analysis of homologous enzymes. *Xenobiotica*, **42**, 808–820.

Dong, D. and Wu, B. (2012) *In silico* modeling of UDP-glucuronosyltransferase 1A10 substrates using the VolSurf approach. *J Pharm Sci*, **101**, 3531–3539.

Eissing, T., Kuepfer, L., Becker, C. *et al.* (2011) A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. *Front Physiol*, **2**, 4.

Ekins, S. (2003) *In silico* approaches to predicting drug metabolism, toxicology and beyond. *Biochem Soc Trans*, **31**, 611–614.

Ekins, S., Andreyev, S., Ryabov, A. *et al.* (2006) A combined approach to drug metabolism and toxicity assessment. *Drug Metab Dispos*, **34**, 495–503.

Ekins, S. and Obach, R.S. (2000) Three-dimensional quantitative structure activity relationship computational approaches for prediction of human *in vitro* intrinsic clearance. *J Pharmacol Exp Ther*, **295**, 463–473.

Ekins, S., Polli, J.E., Swaan, P.W. and Wright, S.H. (2012) Computational modeling to accelerate the identification of substrates and inhibitors for transporters that affect drug disposition. *Clin Pharmacol Ther*, **92**, 661–665.

Ekins, S., Waller, C.L., Swaan, P.W. *et al.* (2000) Progress in predicting human ADME parameters *in silico*. *J Pharmacol Toxicol Methods*, **44**, 251–272.

Fagerholm, U. (2007a) Prediction of human pharmacokinetics--gut-wall metabolism. *J Pharm Pharmacol*, **59**, 1335–1343.

Fagerholm, U. (2007b) Prediction of human pharmacokinetics - renal metabolic and excretion clearance. *J Pharm Pharmacol*, **59**, 1463–1471.

Fioravanzo, E., Bassan, A., Pavan, M. *et al.* (2012) Role of *in silico* genotoxicity tools in the regulatory assessment of pharmaceutical impurities. *SAR QSAR Environ Res*, **23**, 257–277.

Galetin, A. and Houston, J.B. (2006) Intestinal and hepatic metabolic activity of five cytochrome P450 enzymes: impact on prediction of first-pass metabolism. *J Pharmacol Exp Ther*, **318**, 1220–1229.

Gallegos-Saliner, A., Poater, A., Jeliazkova, N. *et al.* (2008) Toxmatch--a chemical classification and activity prediction tool based on similarity measures. *Regul Toxicol Pharmacol*, **52**, 77–84.

Gaohua, L., Abduljalil, K., Jamei, M. *et al.* (2012) A pregnancy physiologically based pharmacokinetic (p-PBPK) model for disposition of drugs metabolized by CYP1A2, CYP2D6 and CYP3A4. *Br J Clin Pharmacol*, **74**, 873–885.

Gerner, I., Spielmann, H., Hoefer, T. *et al.* (2004) Regulatory use of (Q)SARs in toxicological hazard assessment strategies. *SAR QSAR Environ Res*, **15**, 359–366.

Ginsberg, G., Slikker, W. Jr., Bruckner, J. and Sonawane, B. (2004) Incorporating children's toxicokinetics into a risk framework. *Environ Health Perspect*, **112**, 272–283.

Godin, B. and Touitou, E. (2007) Transdermal skin delivery: predictions for humans from *in vivo*, ex vivo and animal models. *Adv Drug Deliv Rev*, **59**, 1152–1161.

Goebel, C., Aeby, P., Ade, N. *et al.* (2012) Guiding principles for the implementation of non-animal safety assessment approaches for cosmetics: skin sensitisation. *Regul Toxicol Pharmacol*, **63**, 40–52.

Golla, S., Madihally, S., Robinson, R.L. Jr. and Gasem, K.A. (2009) Quantitative structure-property relationships modeling of skin irritation. *Toxicol In Vitro*, **23**, 176–184.

Goodwin, J.T. and Clark, D.E. (2005) *In silico* predictions of blood-brain barrier penetration: considerations to "keep in mind". *J Pharmacol Exp Ther*, **315**, 477–483.

Heikkinen, A.T., Baneyx, G., Caruso, A. and Parrott, N. (2012) Application of PBPK modeling to predict human intestinal metabolism of CYP3A substrates - an evaluation and case study using GastroPlus™. *Eur J Pharm Sci*, **47**, 375–386.

Hewitt, M., Ellison, C.M., Enoch, S.J. *et al.* (2010) Integrating (Q)SAR models, expert systems and read-across approaches for the prediction of developmental toxicity. *Reprod Toxicol*, **30**, 147–160.

Huang, S.M. and Rowland, M. (2012) The role of physiologically based pharmacokinetic modeling in regulatory review. *Clin Pharmacol Ther*, **91**, 542–549.

Hulzebos, E., Sijm, D., Traas, T. *et al.* (2005) Validity and validation of expert (Q)SAR systems. *SAR QSAR Environ Res*, **16**, 385–401.

Jamei, M., Marciniak, S., Edwards, D. *et al.* (2013) The Simcyp population based simulator: architecture, implementation, and quality assurance. *In Silico Pharmacology*, **1** (9).

Jamei, M., Marciniak, S., Feng, K. *et al.* (2009) The Simcyp population-based ADME simulator. *Expert Opin Drug Metab Toxicol*, **5**, 211–223.

Jolivette, L.J. and Ekins, S. (2007) Methods for predicting human drug metabolism. *Adv Clin Chem*, **43**, 131–176.

Langenbucher, F. (2007) Handling of computational *in vitro/in vivo* correlation problems by Microsoft Excel: V. Predictive absorbability models. *Eur J Pharm Biopharm*, **67**, 293–299.

Leahy, D.E. (2006) Integrating *in vitro* ADMET data through generic physiologically based pharmacokinetic models. *Expert Opin Drug Metab Toxicol*, **2**, 619–628.

Lilienblum, W., Dekant, W., Foth, H. *et al.* (2008) Alternative methods to safety studies in experimental animals: role in the risk assessment of chemicals under the new European Chemicals Legislation (REACH). *Arch Toxicol*, **82**, 211–236.

Lupfert, C. and Reichel, A. (2005) Development and application of physiologically based pharmacokinetic-modeling tools to support drug discovery. *Chem Biodivers*, **2**, 1462–1486.

Magnusson, B.M., Anissimov, Y.G., Cross, S.E. and Roberts, M.S. (2004) Molecular size as the main determinant of solute maximum flux across the skin. *J Invest Dermatol*, **122**, 993–999.

Marchant, C.A., Briggs, K.A. and Long, A. (2008) *In silico* tools for sharing data and knowledge on toxicity and metabolism: Derek for Windows, Meteor, and Vitic. *Toxicol Mech Methods*, **18**, 177–187.

Matthews, E.J., Kruhlak, N.L., Benz, R.D. *et al.* (2009a) Identification of structure-activity relationships for adverse effects of pharmaceuticals in humans: Part C: use of QSAR and an expert system for the estimation of the mechanism of action of drug-induced hepatobiliary and urinary tract toxicities. *Regul Toxicol Pharmacol*, **54**, 43–65.

Matthews, E.J., Kruhlak, N.L., Benz, R.D. *et al.* (2008) Combined use of MC4PC, MDL-QSAR, BioEpisteme, Leadscope PDM, and Derek for Windows software to achieve high-performance, high-confidence, Mode of Action-based predictions of chemical carcinogenesis in rodents. *Toxicol Mech Methods*, **18**, 189–206.

Matthews, E.J., Kruhlak, N.L., Daniel Benz, R. and Contrera, J.F. (2007a) A comprehensive model for reproductive and developmental toxicity hazard identification: I. Development of a weight of evidence QSAR database. *Regul Toxicol Pharmacol*, **47**, 115–135.

Matthews, E.J., Kruhlak, N.L., Daniel Benz, R. *et al.* (2007b) A comprehensive model for reproductive and developmental toxicity hazard identification: II. Construction of QSAR models to predict activities of untested chemicals. *Regul Toxicol Pharmacol*, **47**, 136–155.

Matthews, E.J., Ursem, C.J., Kruhlak, N.L. *et al.* (2009b) Identification of structure-activity relationships for adverse effects of pharmaceuticals in humans: Part B. Use of (Q)SAR systems for early detection of drug-induced hepatobiliary and urinary tract toxicities. *Regul Toxicol Pharmacol*, **54**, 23–42.

Mayer, J., Cheeseman, M.A. and Twaroski, M.L. (2008) Structure-activity relationship analysis tools: validation and applicability in predicting carcinogens. *Regul Toxicol Pharmacol*, **50**, 50–58.

McCoy, A.T., Bartels, M.J., Rick, D.L. and Saghir, S.A. (2012) TK Modeler version 1.0, a Microsoft® Excel®-based modeling software for the prediction of diurnal blood/plasma concentration for toxicokinetic use. *Regul Toxicol Pharmacol*, **63**, 333–343.

McLanahan, E.D., El-Masri, H.A., Sweeney, L.M. *et al.* (2012) Physiologically based pharmacokinetic model use in risk assessment–Why being published is not enough. *Toxicol Sci*, **126**, 5–15.

Mitragotri, S., Anissimov, Y.G., Bunge, A.L. *et al.* (2011) Mathematical models of skin permeability: an overview. *Int J Pharm*, **418**, 115–129.

Modi, S., Hughes, M., Garrow, A. and White, A. (2012) The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries. *Drug Discov Today*, **17**, 135–142.

Mombelli, E. (2008) An evaluation of the predictive ability of the QSAR software packages, Derek, HazardExpert and TOPKAT, to describe chemically-induced skin irritation. *Altern Lab Anim*, **36**, 15–24.

Mombelli, E. and Devillers, J. (2010) Evaluation of the OECD (Q)SAR Application Toolbox and Toxtree for predicting and profiling the carcinogenic potential of chemicals. *SAR QSAR Environ Res*, **21**, 731–752.

Moroy, G., Martiny, V.Y., Vayer, P. *et al.* (2012) Toward *in silico* structure-based ADMET prediction in drug discovery. *Drug Discov Today*, **17**, 44–55.

Moss, G.P., Dearden, J.C., Patel, H. and Cronin, M.T. (2002) Quantitative structure-permeability relationships (QSPRs) for percutaneous absorption. *Toxicol In Vitro*, **16**, 299–317.

Moss, G.P., Sun, Y., Wilkinson, S.C. *et al.* (2011) The application and limitations of mathematical modelling in the prediction of permeability across mammalian skin and polydimethylsiloxane membranes. *J Pharm Pharmacol*, **63**, 1411–1427.

Myshkin, E., Brennan, R., Khasanova, T. *et al.* (2012) Prediction of organ toxicity endpoints by QSAR modeling based on precise chemical-histopathology annotations. *Chem Biol Drug Des*, **80**, 406–416.

Nagilla, R., Frank, K.A., Jolivette, L.J. and Ward, K.W. (2006) Investigation of the utility of published *in vitro* intrinsic clearance data for prediction of *in vivo* clearance. *J Pharmacol Toxicol Methods*, **53**, 106–116.

Naik, P. and Cucullo, L. (2012) *In vitro* blood-brain barrier models: current and perspective technologies. *J Pharm Sci*, **101**, 1337–1354.

Patlewicz, G., Jeliazkova, N., Gallegos Saliner, A. and Worth, A.P. (2008) Toxmatch-a new software tool to aid in the development and evaluation of chemically similar groups. *SAR QSAR Environ Res*, **19**, 397–412.

Pavan, M. and Worth, A.P. (2008) Publicly-accessible QSAR software tools developed by the Joint Research Centre. *SAR QSAR Environ Res*, **19**, 785–799.

Piersma, A. 2011. Development and implementation of alternative methods in reproductive toxicology. RIVM 340700005/2011.

Piotrowski, P.L., Sumpter, B.G., Malling, H.V. *et al.* (2007) A toxicity evaluation and predictive system based on neural networks and wavelets. *J Chem Inf Model*, **47**, 676–685.

Polak, S., Ghobadi, C., Mishra, H. *et al.* (2012) Prediction of concentration-time profile and its inter-individual variability following the dermal drug absorption. *J Pharm Sci*, **101**, 2584–2595.

Potts, R.O. and Guy, R.H. (1992) Predicting skin permeability. *Pharm Res*, **9**, 663–669.

Rostami-Hodjegan, A. (2012) Physiologically based pharmacokinetics joined with *in vitro-in vivo* extrapolation of ADME: a marriage under the arch of systems pharmacology. *Clin Pharmacol Ther*, **92**, 50–61.

Rostami-Hodjegan, A. and Tucker, G.T. (2007) Simulation and prediction of *in vivo* drug metabolism in human populations from *in vitro* data. *Nat Rev Drug Discov*, **6**, 140–148.

Rupp, B., Appel, K.E. and Gundert-Remy, U. (2010) Chronic oral LOAEL prediction by using a commercially available computational QSAR tool. *Arch Toxicol*, **84**, 681–688.

Rusyn, I., Sedykh, A., Low, Y. *et al.* (2012) Predictive modeling of chemical hazard by integrating numerical descriptors of chemical structures and short-term toxicity assay data. *Toxicol Sci*, **127**, 1–9.

Saiakhov, R.D. and Klopman, G. (2008) MultiCASE Expert Systems and the REACH Initiative. *Toxicol Mech Methods*, **18**, 159–175.

Schafer-Korting, M., Bock, U., Diembeck, W. *et al.* (2008) The use of reconstructed human epidermis for skin absorption testing: Results of the validation study. *Altern Lab Anim*, **36**, 161–187.

Scheel, J. and Keller, D. (2012) Investigation of the skin sensitizing properties of 5 osmolytic prodrugs in a weight-of-evidence assessment, employing *in silico*, *in vivo*, and read across analyses. *Int J Toxicol*, **31**, 358–363.

Sinha, V.K., Snoeys, J., Osselaer, N.V. *et al.* (2012) From preclinical to human--prediction of oral absorption and drug-drug interaction potential using physiologically based pharmacokinetic (PBPK) modeling approach in an industrial setting: a workflow by using case example. *Biopharm Drug Dispos*, **33**, 111–121.

T'jollyn, H., Boussery, K., Mortishire-Smith, R.J. *et al.* (2011) Evaluation of three state-of-the-art metabolite prediction software packages (Meteor, MetaSite, and StarDrop) through independent and synergistic use. *Drug Metab Dispos*, **39**, 2066–2075.

Tarcsay, Á., Kiss, R. and Keser, G. (2010) Site of metabolism prediction on cytochrome P450 2C9: a knowledge-based docking approach. *J Comput Aided Mol Des*, **24**, 399–408.

Thomas, S., Brightman, F., Gill, H. *et al.* (2008) Simulation modelling of human intestinal absorption using Caco-2 permeability and kinetic solubility data for early drug discovery. *J Pharm Sci*, **97**, 4557–4574.

Tosco, P. and Balle, T. (2011) Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields. *J Mol Model*, **17**, 201–208.

Vedani, A., Dobler, M. and Smiesko, M. (2012) VirtualToxLab - a platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicol Appl Pharmacol*, **261**, 142–153.

Walum, E., Hedander, J. and Garberg, P. (2005) Research perspectives for pre-screening alternatives to animal experimentation On the relevance of cytotoxicity measurements, barrier passage determinations and high throughput screening *in vitro* to select potentially hazardous compounds in large sets of chemicals. *Toxicol Appl Pharmacol*, **207**, 393–397.

Wang, Y.H. (2010) Confidence assessment of the Simcyp time-based approach and a static mathematical model in predicting clinical drug-drug interactions for mechanism-based CYP3A inhibitors. *Drug Metab Dispos*, **38**, 1094–1104.

Willmann, S., Lippert, J. and Schmitt, W. (2005) From physicochemistry to absorption and distribution: predictive mechanistic modelling and computational tools. *Expert Opin Drug Metab Toxicol*, **1**, 159–168.

Willmann, S., Thelen, K. and Lippert, J. (2012) Integration of dissolution into physiologically-based pharmacokinetic models III: PK-Sim®. *J Pharm Pharmacol*, **64**, 997–1007.

Zhao, P., Rowland, M. and Huang, S.M. (2012) Best practice in the use of physiologically based pharmacokinetic modeling and simulation to address clinical pharmacology regulatory questions. *Clin Pharmacol Ther*, **92**, 17–20.

7

Transgenic Animal Models for ADME and Systemic Toxicity

7.1 Transgenic models and their use in toxicology

The traditional discipline of toxicology has focused on the use of animal models to study ADME and toxicity. However, a vast amount of data indicates that animals differ from humans in important ways. They may either possess biological functions which are absent in humans or lack functions which humans have. Even where a particular protein is expressed in both humans and animal models there may be subtle (or even dramatic) differences between the two systems.

Recent advances in transgenic technology have made it possible to address these issues in a variety of ways. This chapter aims to provide sufficient background to allow the reader to understand the literature in this fast-moving field, but for the latest, up-to-date information reference to the websites of key players in the field is recommended.

In most cases the species used to generate transgenic models has been the mouse because the process of transgenesis is technically easier in this species than in others such as the rat. The simplest approach takes advantage of the natural propensity of ESCs to undergo homologous recombination to remove an endogenous gene using a targeting vector, thus generating a so-called null or knockout (KO) mouse. The same technology can be used to create knock-in mice in which a piece of exogenous DNA is introduced as part of the same process by which the endogenous gene is knocked out. In contrast, transgenesis is the process of adding a gene, which may be a full-length or truncated cDNA sequence (native or mutated), a genomic DNA fragment, antisense sequence or a microRNA coding

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/molecularcellulartoxicology

element. This transgene is typically linked to a promoter which drives its expression. Depending upon the desired characteristics of the transgenic model, this may be a strong generic promoter (thus generating a model in which the gene of interest is overexpressed) or one which drives tissue-specific expression (e.g. in the liver).[1]

In conventional transgenics, the transgene integrates randomly at a single site in the genome. The location of integration cannot be predicted in advance, nor can the number of copies, which may range from one to as many as 50 in end-to-end (concatameric) orientation. More advanced models include the following.

- Conditional knockouts, in which a particular gene is engineered to target it for deletion at a specific key point in development.
- Humanised models, in which one or more key murine genes is/are replaced with its human equivalent.
- Reporter models, knock-in mice engineered to carry a chimeric gene in which the expression of a readily detectable reporter molecule (e.g. a fluorescent protein) is driven by the regulatory elements of the gene of interest, thus allowing the regulation of expression (but not function) to be examined in detail.

This is a very fast-moving field and a complete review of the available models is beyond the scope of a single chapter. Instead, this chapter will provide a few edited highlights to illustrate the current status of the field and suggest further reading for those who are interested in pursuing the topic further.[2]

7.2 ADME models

The effects of xenobiotics are mediated by changes in the expression of key genes such as xenobiotic metabolising enzymes and stress response genes. These are conventionally studied using animal models; however, this approach can be highly problematic because of the following:

- There are marked species differences in responsiveness to xenobiotics, particularly those which act via receptor-mediated mechanisms.
- The responses observed are often subtle and complex, involving changes in the expression of genes whose expression is difficult to detect and whose function is difficult to measure.

The need of the pharmaceutical industry to be able to model drug handling in humans, combined with the enthusiasm of many people in this industry for the concept of novel models and new testing strategies, has facilitated the development of a variety of humanised models which can be used to study ADME (Figure 7.1).

[1] For an excellent general review of the considerations involved in the research use of transgenic mice, see Davis *et al.*, 2012. The specific examples in that review are taken from the field of cardiac research, but the general points made and issues raised are relevant to the use of transgenic models in any research context.

[2] For a detailed review of this topic, see Boverhof *et al.* (2011).

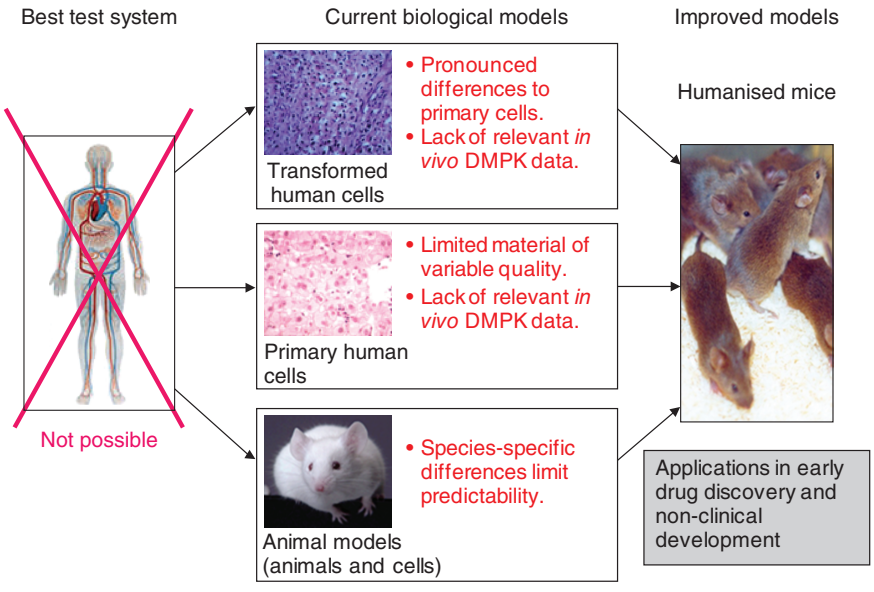


Figure 7.1 Rationale for the use of humanised models in ADME and toxicology. It is not ethical to conduct preliminary toxicity and ADME testing in humans, but current preclinical models are clearly not serving us well. Up to 70% of drug attrition during preclinical development is caused by non-clinical toxicity, poor pharmacokinetics, clinical safety issues or lack of efficacy. Furthermore, only one in nine compounds entering Phase I studies gets approved by regulatory authorities. Indeed, overall the cost for each ultimately successful drug can be as much as $1billion. Mice humanised for major xenobiotic processing systems (including nuclear receptors, drug transporters and xenobiotic metabolising enzymes) represent one promising approach to solving this problem (source: Scottish Enterprise. Reproduced with permission of Scottish Enterprise). DMPK: Drug metabolism and pharmacokinetics

Conventional humanised models tend to use a human cDNA or genomic sequence (for a nuclear receptor, CYP or Phase II enzyme) driven by a strong tissue-specific promoter (e.g. the albumin promoter, to drive expression in the liver). These models are limited in that the levels and location of expression observed usually fail to correspond either with that of the human gene (in so far as it has been characterised) or the endogenous murine gene. More advanced models for the physiologically appropriate expression of nuclear receptors and CYPs have recently been developed. These use large tracts of genomic sequence to control the expression of the human transgene in the appropriate tissues and at physiologically relevant levels. This chapter will make reference to conventional approaches and recommend further reading for further study of the models generated using these methods, but its primary focus will be the more innovative models which have been reported since about 2008. The most fully developed of these are models for nuclear receptor function, CYP-mediated metabolism (Phase I), Phase II xenobiotic metabolism and drug transport (Phase III). Figure 7.2 illustrates the position of these targets within the process of hepatic xenobiotic metabolism.

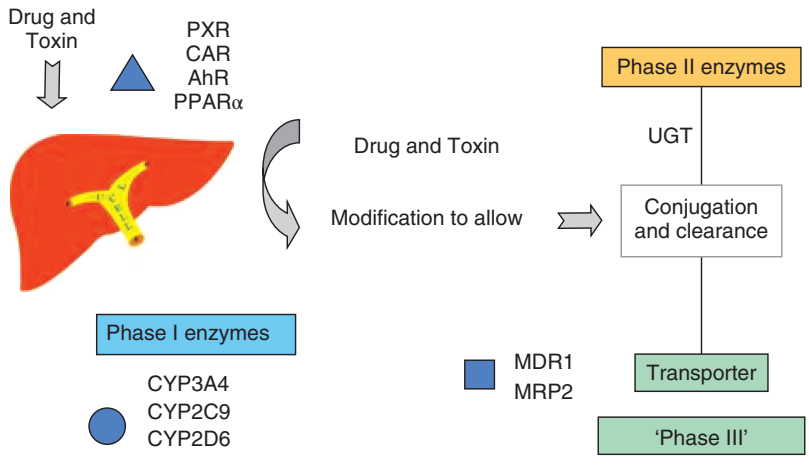


Figure 7.2 Targets for transgenesis in the process of hepatic xenobiotic metabolism (source: Scottish Enterprise. Reproduced with permission of Scottish Enterprise)

7.2.1 Nuclear receptor models

PXR and CAR The nuclear receptors pregnane X receptor (PXR) and constitutive androstane receptor (CAR) were identified in the late 1990s and subsequently shown to play key roles in the regulation of drug-metabolising enzymes, particularly the CYP2B and CYP3A families of CYPs. They are members of a family of nuclear receptors which includes the steroid, retinoid, and thyroid hormone receptors. These are characterised by the presence of a DNA-binding domain with two zinc finger motifs and a conserved ligand-binding domain which, in addition to the ligand-binding site, contains dimerisation motifs and trans-activation domains. In the case of PXR and CAR, the dimerisation partner is the retinoid X receptor.

Mouse PXR was identified in 1998 by using an expressed sequence tag to screen a mouse liver library (Kliewer *et al.*, 1998). The receptor cloned by this approach was shown to be activated by derivatives of dexamethasone and pregnenolone. At approximately the same time, the human receptor steroid X receptor (SXR), the human homologue of murine PXR, was cloned and shown to be related to CAR and the 1,25-dihydroxyvitamin D receptor (Blumberg *et al.*, 1998; Lehmann *et al.*, 1998).

The first reports referring to CAR as such were published in 1997 (Choi *et al.*, 1997), although the human analogue of this orphan receptor (MB67) had been identified previously (Baes *et al.*, 1994). The discovery of CAR made it necessary to accommodate the concept of a constitutively active receptor. CAR is subject to regulation by two classes of ligands: agonists (e.g. 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP)) and inverse agonists (e.g. androstanol) (Tzameli and Moore, 2001). One of the early difficulties in studying CAR was that high concentrations of PB are required for activation; indeed, it is still unclear whether PB actually activates CAR via a direct binding interaction. However, the discovery of two more potent ligands: TCPOBOP, a potent and long-lasting PB-like inducing agent in the mouse, and a novel imidazole

derivative, 6-(4-chlorophenyl)-imidazo[2,1-b]thiazole-5-carbaldehyde (CITCO), a specific agonist for human CAR, opened up the possibility of studying species-specific CAR responses (Maglich *et al.*, 2003).

It soon emerged that PXR mediated the non-glucocorticoid receptor-mediated mechanism of induction of the CYP3A family, while it appeared that CAR may be the long sought-after PB receptor. In addition to regulating key CYPs, these receptors are involved in the regulation of phase II drug-metabolising enzymes and transmembrane drug transporters.[3]

Both PXR and CAR act as transcription factors. The basic mechanism by which PXR and CAR are thought to activate transcription is a conventional one, involving xenobiotic binding, heterodimerisation and interactions with the 5′ regulatory sequences of target genes, as summarised in Figure 7.3. The mechanism of action of PXR is similar to that of conventional nuclear receptors such as the steroid hormone receptors except that, instead of binding a narrow range of ligands with very high affinity, it binds a wide range of ligands with relatively low affinity.

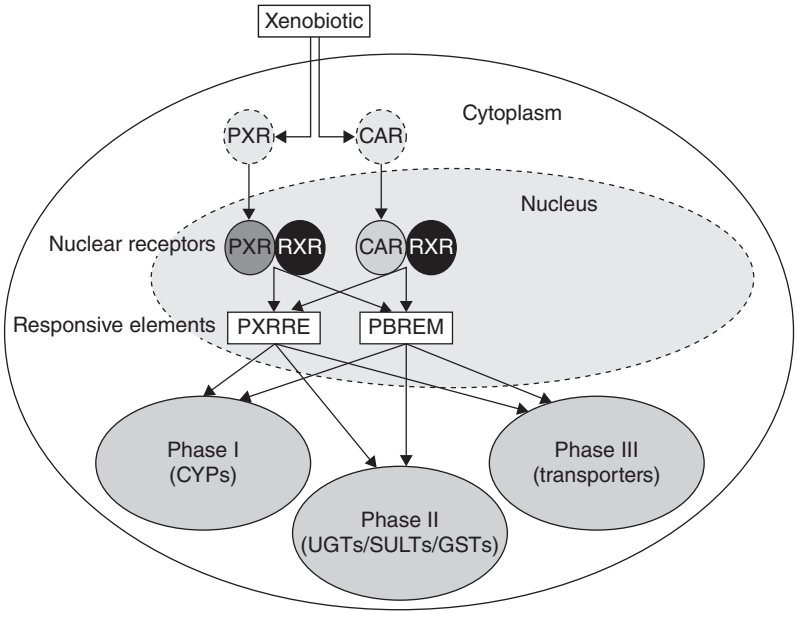


Figure 7.3 **Cross-talk between PXR and CAR.** A multitude of xenobiotics can interact with PXR or CAR or both. Ligand binding leads to the translocation of the receptor from the cytoplasm to the nucleus, where both PXR and CAR use the retinoic X receptor (RXR) as a partner for heterodimerization and bind to corresponding cis-acting elements of their target genes. PXR and CAR mediate gene regulation through certain direct repeats, known as the PXR responsive element (PXRRE) and the PB responsive element module (PBREM). Target genes that contain PXRREs and PBREMs in their promoters can be regulated both by PXR and CAR. Among those that are subject to PXR and CAR regulation are a number of genes encoding Phase I or Phase II drug-metabolising enzymes or transporters (source: Sheer *et al.* (2008), figure 1. Reproduced with permission of the American Society for Clinical Investigation)

[3] For a detailed review of this topic and the transgenic models outlined later, see Stanley *et al.* (2006).

This pattern of ligand responsiveness may have evolved to allow this receptor to deal with the vast number of exogenous chemicals to which organisms are exposed. The ligand responsiveness of CAR is more complex. This receptor is constitutively active and is able to respond to both true agonists such as TCPOBOP (which up-regulate its activity) and inverse agonists such as androstanol (which down-regulate it). However, the picture with CAR is less clear than that with PXR because of the possibility that constitutively active CAR may be excluded from the nucleus in order to prevent unwanted induction of gene expression.

Marked species differences in responsiveness to CYP-inducing agents have been known for many years, and with the identification and structural characterisation of PXR and CAR it became clear that many of these differences are due to intrinsic differences between these receptors in different species. Such differences are likely to play an important role in drug pharmacokinetics and toxicity as well as in endogenous metabolism and pathological processes, making it difficult to extrapolate directly from, for example, results in mice in order to make predictions about drug disposition in man. The characterisation of species differences in CAR activity is complicated by the existence of both agonists and inverse agonists (i.e. compounds which switch off the constitutive activity of CAR), although the key residues involved in differences between murine CAR and human CAR in responsiveness to agonists and inverse agonists have been identified (Ueda *et al*., 2002).

A wide range of technologies has recently been developed to facilitate the screening of compounds for their interactions with transcription factors such as CAR and PXR. Such assays provide a valuable tool for application in drug development. Among the *in vitro* methods commonly used to study ligand interactions with PXR and CAR, the most common approach is to use reporter constructs which deliver luminescent or biochemical readouts and are transiently transfected into either continuously cultured cell lines or primary cells.

In vitro assays have the advantage of rapidity, but clearly have significant limitations based on the reporter systems being applied, the cell line being used for the assay and the fact that they give no information about cellular effects *in vivo*. The generation of humanised models provides a novel way of understanding the functions of these genes in drug metabolism and drug safety. Examples of processes which would be predicted to exhibit mouse–human differences due to species differences in PXR and CAR include drug metabolism and transport (Phase I, II and III) as well as homeostatic control mechanisms (such as thyroid and haem homeostasis), acute toxicity, carcinogenesis, liver damage, infection and inflammation. Some of these differences have been rationalised at the molecular level, but many remain to be elucidated and the development of null and humanised mouse models represents a significant step forward in this effort. In particular, the development of novel models will help us to reach a fuller understanding of CAR activity, which is particularly complex because of the existence of ligands with different effects (agonists, partial agonists and antagonists) as well as the many factors which appear to affect CAR activity indirectly. Novel models, including PXR and CAR null and humanised mice, represent invaluable tools for the dissection of these processes.

PXR null and PXR-humanised mouse models PXR null mice, generated by engineering a murine PXR replacement targeting vector from a 129/Sv genomic

bacterial artificial chromosome (BAC) library into ESCs (Xie *et al*., 2000), have no overt phenotype and no significant changes in blood chemistry in comparison with wild-type controls, although their hepatic Cyp3a11 levels are elevated.

Mouse lines which constitutively overexpress human PXR have been generated using two transgenes, Alb-SXR and Alb-VPSXR, which drive the hepatic expression of human PXR under the control of the mouse albumin promoter/enhancer (Xie *et al*., 2000). The Alb-SXR transgene encodes normal human PXR, whereas Alb-VPSXR encodes a constitutively active form of the receptor. The Alb-SXR transgene was subsequently bred onto a PXR null background to generate mice which carried human PXR instead of murine PXR. Alb-SXR mice respond to PXR ligands is a manner more characteristic of humans than wild-type mice. Interestingly, introduction of the constitutively active human PXR construct VPSXR led to an overt phenotype including hepatomegaly, histologic liver toxicity and growth retardation.

A follow-up from this study again involved the use of the VPSXR construct, this time driven by the fatty-acid-binding protein (FABP) promoter (Gong *et al*., 2006). In this model, constitutively active human PXR is expressed in the liver and GI tract but not in the stomach or kidney. This confers increased sensitivity to toxins which act via oxidative stress pathways (specifically, in this case, paraquat), suggesting that PXR may play a role in mammalian oxidative stress responses. Similar sensitivity was observed in Alb-VPSXR mice as well as in wild-type mice treated with PCN. Increased sensitivity was associated with GSH depletion (despite transient up-regulation of GST in response to paraquat) and reduced expression of SODs and CAR, these responses being independent of the Nrf2-KEAP1 pathway.

Both these models were made using human PXR cDNAs, so they lack human regulatory sequences. A further model has been made using a BAC construct containing the complete human PXR gene including all its 5′ and 3′ flanking sequences (Ma *et al*., 2007a). Following the introduction of the human PXR construct the mice were bred onto a PXR null background, backcrossed to a C57BL/6 genetic background for at least four generations and bred to homozygosity for human PXR.

These mice expressed human PXR mRNA in an appropriate tissue-specific manner (i.e. in the liver and GI tract (except stomach) but not in the heart or lung). Unlike in humans, where the expression of PXR in the liver is much higher than in the GI tract, the levels of expression in liver and GI tract were similar. As expected, the expression of Cyp3a11 was up-regulated in response to rifampicin ($10\,mg\,kg^{-1}$) but not PCN in the humanised mice (the reverse of the pattern in wild-type animals) and the C_{max} and AUC of midazolam were correspondingly reduced (by about 60%). Interestingly, hepatic expression of Cyp3a11 took a week to decline after the last dose of rifampicin, leading the authors to remark on the possibility of DDIs in humans even after ceasing treatment with drugs which are PXR ligands.

Example: Human-specific effects of the PXR agonist rifamixin

This model has been used to explain the beneficial effects of the antibiotic rifaximin in inflammatory bowel disease (Ma *et al*., 2007b). Rifaximin is a human-specific intestinal CYP3A inducer conventionally used to treat traveller's diarrhoea; it can also be used relieve chronic GI disorders. In PXR-humanised mice, a dose of $25\,mg\,kg^{-1}$ induced about fourfold up-regulation of Cyp3a11 in

the gut. Pharmacokinetic analysis showed that the compound was not absorbed following administration in the diet, suggesting that its oral bioavailability in humans is likely to be minimal and that its effects in the gut are essentially topical. In humanised mice, rifaximin provided significant benefits in artificially induced inflammatory bowel disease whereas rifampicin, which is absorbed, actually exacerbated the condition (Cheng *et al*., 2010). The beneficial effect was not observed in wild-type or PXR null mice, suggesting that it is mediated via a PXR-specific effect, rather than being due to the antibiotic properties of the drug (Cheng, Shah, and Gonzalez, 2012b). Recently, wild-type, PXR null and human PXR mice have been used in studies aimed at establishing whether rifaximin is safe for long-term clinical use (Cheng *et al*., 2012a). Treatment with rifaximin over a 6-month period revealed induction of hepatocellular degeneration, nodular hyperplasia and mild fibrosis in human PXR mice, although this was not associated with liver enzyme release. This effect was not seen in wild-type or PXR null mice, suggesting that it was dependent on the presence of human PXR, and the authors concluded that 'despite the potential adverse effect of long-term administration of rifaximin, it is still within a safety tolerance applied in clinical therapy to treat GI diseases'. While the conclusion drawn from this study may be considered rather optimistic, it does serve to illustrate the usefulness of PXR null and humanised models in dissecting the potential beneficial and adverse effects of drugs which may be human-specific PXR ligands.

CAR null and CAR humanised mouse models Several lines of CAR null mice have been generated independently and used for different purposes. One line was generated using a construct in which the coding region of β-galactosidase replaced the murine CAR coding sequence close to the promoter, obliterating part of the DNA-binding domain (Wei *et al*., 2000). The resulting mice did not express murine CAR mRNA, had no overt phenotype and were resistant to the effects of the CAR ligands PB (100 mg kg^{-1}) and TCPOBOP (3 mg kg^{-1}). They were not only resistant to the induction of Phase I drug metabolism by PB and TCPOBOP but also had reduced basal expression of these enzymes. This contrasts with the results obtained in PXR null mice, which had elevated basal Cyp3a11 expression. Two other groups have independently generated murine CAR null mice and used them in attempts to evaluate the role of murine CAR in hepatocarcinogenesis and to identify the overlapping but distinct sets of genes regulated by murine PXR and CAR. In addition, PXR/CAR double null mice have been made by inter-breeding PXR null and CAR null mice.

Example: Role of CAR in metabolic diseases

It has also been suggested that CAR could play a role in the relief of obesity and diabetes (Gao *et al*., 2009). When the responses of wild-type mice with obesity/Type II diabetes induced by administration of a high-fat diet were compared with those of CAR null mice, it was found that treatment with TCPOBOP could prevent the development of obesity or reverse preinduced obesity in wild-type mice. It also improved insulin sensitivity in both genetically obese (*ob/ob*) and high-fat diet-induced Type II diabetic models, whereas CAR null mice had intrinsic insulin resistance which could not be relieved by TCPOBOP.

A CAR-humanised mouse line has also been described. A construct containing the coding sequence for human CAR was linked to the albumin promoter. A region from an abundantly expressed gene (rabbit β-globin) was added between the promoter and the human CAR receptor coding sequence to enhance the expression of the human CAR receptor. This combination of regulatory elements leads to dramatic overexpression of human CAR in the liver of these mice.

Evaluation Our understanding of the role of PXR and CAR in the control of endogenous metabolism and pathological changes therein is improving rapidly, although it is still far from complete. Nuclear receptor-humanised mice have already been used in a number of studies addressing acute toxicity, mechanisms of non-genotoxic carcinogenesis, thyroid homeostasis, haem homeostasis, liver pathology, the effects of bile acid accumulation, bilirubin clearance and the role of PXR as a bile acid chemosensor. The data obtained so far indicate that PXR and CAR may well play a role in the control of bile acid and cholesterol homeostasis. However, PXR and CAR seem to have a number of paradoxical effects which are, as yet, incompletely understood. Further work is required and this is a prime area for further investigation using novel null and humanised models. Existing models have already started this process; for example, a transgenic mouse system that allows conditional expression of a constitutively activated human CAR (VP-CAR) created using the TetRE-VP-CAR transgene, which allows the expression of human CAR to be controlled by treatment with doxocycline[4] has been used to investigate the role of CAR in responsiveness to lithocholic acid and regulation of SULT expression. The ongoing development of new models which allow more *bona fide* expression of PXR and CAR *in vivo*, as well as mice humanised for both receptors, will further enhance this process, providing much greater clarity to our understanding of these important receptors.

The tADMET™ PXR/CAR panel The interactions of chemicals with PXR and CAR are complex because they regulate common genes and a single chemical agent can often interact with both receptors. In order to fully understand the relative importance of the receptors in the efficacy and safety of drugs, it is critical to develop models where the role of nuclear receptors can be evaluated separately and in combination. Such an undertaking has been reported by Scheer and colleagues, who have generated and characterised a range of new humanised mouse models for PXR and CAR.

In contrast to previous approaches, Scheer *et al.* (2008, 2010) used knock-in strategies to express human PXR and CAR under the control of their corresponding mouse promoters, with concurrent interruption of the expression of the endogenous genes. The minimal genetic complexity of this system, requiring only one genetic alteration for each receptor, allowed them to establish a panel of PXR/CAR mouse models.

The humanised receptor mice exhibited species-specific differences in the interaction with known drugs. In agreement with previous studies, PXR-humanised was strongly activated by rifampicin, but only weakly by dexamethasone and

[4] For an explanation of how the Tet on/off system works, see Davis *et al.* (2012)

PCN. In the case of CAR-humanised mice, CITCO was a potent activator, while TCPOBOP barely activated the human receptor. The humanised mice therefore appear to reliably reflect the interaction of compounds with these human receptors. The activation of PXR by lower doses of rifampicin in humanised mice compared to wild-type mice was also confirmed by 7-benzyloxyquinoline activity measurements. Similarly, associated with the induction of Cyp2b10 in the CAR-humanised mice, a marked induction in pentoxyresorufin O-dealkylation activity was observed at all doses tested.

The relative genetic simplicity of their models allowed Scheer and colleagues to use a combinatorial approach to generate mice humanised for multiple pathways of drug metabolism. They generated a panel of PXR/CAR mouse models, including single null and humanised mouse lines and all possible combinations of these. This panel makes it possible to discriminate between PXR- and CAR-mediated effects by uncovering the cross-talk between these two receptors and evaluating its significance *in vivo*.

Example: Use of the PXR/CAR panel to dissect mechanisms of CYP induction

Scheer *et al*. used the PXR/CAR panel to assess whether the activation of human and (to a lesser extent) murine PXR by PB, as observed *in vitro*, would have any significance for the regulation of Cyp3a11 or Cyp2b10 by this compound *in vivo*. In wild-type animals, administration of PB at 40 mg kg^{-1} markedly induced Cyp2b10 expression and, to a lesser extent, Cyp3a11. Induction of both these proteins was observed in all mouse lines with a humanised CAR, demonstrating that the human receptor responds equally to PB. In mouse lines lacking CAR, the induction of Cyp2b10 and Cyp3a11 was strongly attenuated, whereas it was unaffected in PXR null animals. These effects were also reflected in the measurement of 7-benzyloxyquinoline and pentoxyresorufin O-dealkylation activity activities. These data indicate that CAR is the transcription factor responsible for induction of Cyp2b10 and Cyp3a11. The fact that induction of both Cyp3a11 and Cyp2b10 was also absent in the PXR-humanised/CAR null mice indicated that human PXR is not able to compensate for the loss of CAR activity.

Example: Species differences in hepatocellular hypertrophy mediated by PXR and CAR

The PXR-humanised, CAR-humanised and corresponding null models developed by Scheer and his colleagues have been used to show that CYP up-regulation and hepatocellular hypertrophy mediated by PXR and CAR can be separated from induction of cell proliferation in the liver (Ross *et al*., 2010). Treatment with PB (80 mg kg^{-1} day^{-1}) or chlordane (10 mg kg^{-1} day^{-1}) led to centrilobular hypertrophy, liver enlargement and CYP induction in PXR-humanised/CAR humanised and wild-type mice but not in the PXR null/CAR null line; in contrast, however, cell proliferation occurred only in wild-type mice, not in either PXR/CAR humanised or PXR/CAR null animals. This is consistent with the widely held view that, while these chemicals exert biochemical effects in both human and rodent liver, human hepatocytes are likely to be resistant to their capacity to induce cell proliferation. Furthermore, it implicates PXR and/or CAR in this species difference and provides reassurance regarding the question of whether these compounds can cause or exacerbate hepatocarcinogenesis in humans.

Evaluation The above-mentioned examples demonstrate that the use of a combination of CAR and PXR mice greatly facilitates the elucidation of ways in which these transcription factors can affect drug disposition and toxicity in man. Mice nulled and/or humanised for single or multiple transcription factors involved in drug disposition in man have many potential applications in toxicology and preclinical drug development. One of the ongoing goals of work on the tADMET™ programme is to further increase the value of these models by introducing human Phase I, II, and III genes onto the CAR/PXR-humanised background. The most up-to-date information on these models may be found on the tADMET™ page of the Taconic web site.[5]

7.2.2 Xenobiotic metabolism models

The Hepatic Reductase Null™ mouse The Hepatic Reductase Null (HRN™) Mouse is an *in vivo* model used to study the effects of CYP-mediated metabolism. It was developed in response to the difficulties associated with deleting individual CYP genes. If the gene in question plays an essential housekeeping role, this will usually lead to either embryonic lethality or perinatal morbidity whereas if it just has a xenobiotic-metabolising function, the effects of deletion are often subtle and other CYPs may compensate for the absent enzyme. Theoretically one might try to create a mouse which was nulled for all CYPs, but because the CYP genes are scattered throughout the mammalian genome this is not technically possible. However, cytochrome P450 reductase is an absolute requirement for activity of the CYP-dependent monooxygenase system. While complete deletion of reductase is developmentally lethal, it became apparent to Henderson and colleagues that specific postnatal deletion could be used to generate a model in which CYP function is more-or-less completely absent in one or more tissues of an otherwise normal mouse. Thus, one solution to the problem was to create a conditional deletion of reductase at defined developmental stages (Henderson *et al*., 2003) or in a tissue-specific manner (Finn *et al*., 2007). This was achieved by flanking the reductase gene with loxP sites then using the cre/loxP system[6] to delete the gene in a tissue-specific manner. The first of these models was the HRN™ mouse, generated by breeding mice carrying a floxed (flanked by LoxP) reductase gene with transgenic mice expressing Cre recombinase under the control of the rat albumin promoter to engineer a conditional deletion of the reductase gene in the adult liver. The resulting HRN™ mice lack hepatic reductase protein and activity but express normal levels of reductase in the kidney. They are viable and fertile and have no overt phenotype despite a complete lack of hepatic CYP activity, although HRN™ mice do tend to develop pale, mottled enlarged livers (characteristic of steatosis) as they mature. The absence of hepatic CYP-mediated metabolism in the HRN™ mouse is demonstrated by the fact that HRN™ mice have severely compromised hepatic testosterone metabolism, with almost complete ablation of 7α- and 16β-hydroxylation and 90% reduction in 6β-hydroxylation despite up-regulation of the corresponding CYP proteins (Henderson *et al*., 2003).

[5] http://www.taconic.com/wmspage.cfm?parm1=1792

[6] For an explanation of this technology, see Boverhof *et al*. (2011).

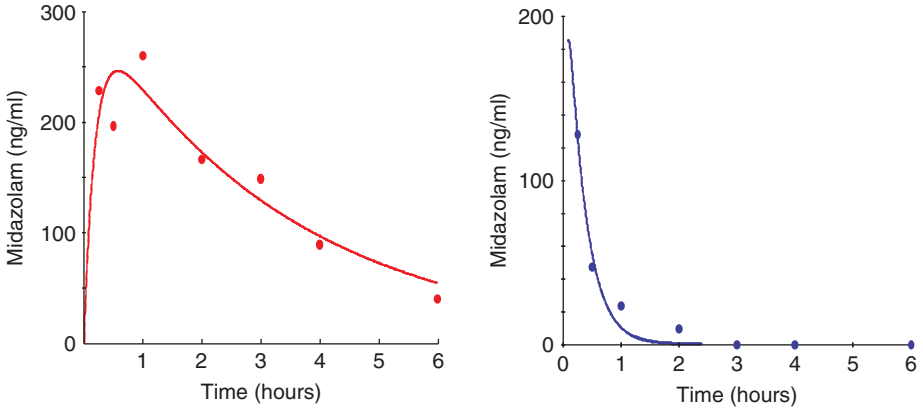


Figure 7.4 The pharmacokinetic profile of midazolam is markedly altered in HRN™ mice. Mice were dosed orally with midazolam (2.5 mg kg^{-1}). Blood samples were taken over the following 6 h and analysed for midazolam by LC-MS/MS; Left: HRN™; Right: Wild-type. (source: CXR Biosciences Ltd. Reproduced with permission of CXR Biosciences Ltd)

Table 7.1 Pharmacokinetic parameters of midazolam in HRN™ and wild-type mice

Parameter	HRN™ mouse	Wild-type mouse
$AUC_{0-6\,h}$ (h ng ml^{-1})	934	76
T_{max} (h)	0.85	0.09
C_{max} (ng ml^{-1})	244	185
CL-F (ml h^{-1})	2680	32 900

Data from CXR Biosciences Ltd.

The HRN™ mouse model can be used to study the role of CYP-mediated metabolism in drug pharmacokinetics and toxicity (Henderson, Pass, and Wolf, 2006). This has been demonstrated by showing that the pharmacokinetic profile of midazolam is markedly altered in HRN™ mice (Figure 7.4 and Table 7.1) while the toxicity of acetaminophen is significantly attenuated in HRN™ mice, this effect being associated with resistance to the GSH depletion which is normally seen following a toxic dose of acetaminophen. This model can be used to discriminate between effects mediated by Phase I and Phase II metabolism: the pharmacokinetics of diclofenac, for example, are virtually the same in HRN™ and wild-type mice (Henderson, Pass, and Wolf, 2006), while its metabolism proceeds mainly via acyl glucuronide formation and taurine amide conjugation, demonstrating the utility of this model in understanding the impact of reactive metabolite formation via non-CYP-mediated routes (Pickup *et al.*, 2012). It has also been used to discriminate between CYP-mediated metabolic activation and detoxification of genotoxic carcinogens (Arlt *et al.*, 2012; Arlt *et al.*, 2011; Arlt *et al.*, 2008).

Example: Better dosing regimens for cancer chemotherapy

One application of the HRN™ model is to inform the development of better dosing regimens to promote the beneficial effects of drugs and minimise

their adverse side effects. A good example of this is cyclophosphamide, which undergoes CYP-mediated metabolism to form 4-hydroxycyclophosphamide. This metabolite is thought to mediate both the anticancer effects of cyclophosphamide and its myelotoxicity. Studies using HRN™ mice indicated that the anticancer effects of cyclophosphamide were related to total exposure (AUC in pharmacokinetic terms) whereas myelotoxicity was a function of acute exposure (C_{max}) (Pass *et al*., 2005). This suggests that slow infusion of cyclophosphamide over a period of time, instead of bolus administration, might favour the anticancer effect of the drug over its myelotoxicity; indeed, this approach has been shown to be beneficial, at least in the paediatric setting (Yule *et al*., 2001).

The HRN™ model is, of course, liver specific. It is complemented by a more recent model in which reductase was knocked out by cross-breeding mice with loxP sites integrated in introns 2 and 5 of the reductase gene with mice expressing Cre recombinase under the control of the villin promoter, leading to deletion of the reductase gene within a month of birth (Zhang *et al*., 2009). The resulting mice are viable, fertile and of normal size with no obvious physiological abnormalities. They express no reductase in the intestine, and as in the liver of the HRN™ mouse the intestine of this line exhibits up-regulation of CYPs including Cyp1a1, Cyp2b, Cyp2c and Cyp3a. The effects of this phenotype on the pharmacokinetics of nifedipine indicated that the deletion of intestinal reductase has clear consequences in terms of oral bioavailability, C_{max} being increased by 1.8-fold and AUC by 1.6-fold; however, further studies, preferably including the HRN™ model for comparison, are required in order to determine the usefulness of this line.

CYP models The use of *in vivo* models for the study of xenobiotic metabolism makes it possible to take into account blood flow, the rate and extent of biodistribution, the tissue distribution of xenobiotic metabolising enzymes and interactions with other xenobiotic handling systems such as drug transporters. However, it is well established that xenobiotic metabolism in experimental animals, including mice, differs significantly from that of humans. In an effort to represent human xenobiotic metabolism more accurately; therefore, numerous so-called *humanised* models expressing one or more CYPs in the presence or absence of their mouse homologues and under the control of various promoter and enhancer sequences have been described over the last decade or so. A complete review of all of these is beyond the scope of this chapter; instead it will focus mainly on the CYP3A family because of the importance of this isoform in the metabolism of pharmaceuticals and in drug–drug interactions.[7]

CYP3A The importance of the CYP3A gene family, particularly CYP3A4, in the metabolism of drugs and other xenobiotics has led to great efforts in the generation of Cyp3a null and CYP3A4-humanised mice.[8] One of the key groups in this field has been the one led by Alfred Schinkel. This group has developed

[7] The review of Cheung and Gonzalez (2008) ibid. is strongly recommended for further reading.
[8] For a detailed description of these and a wide-ranging discussion of the biological implications of CYP3A and P-glycoprotein, together with a detailed theoretical discussion of their pharmacokinetic roles, see Van Waterschoot and Schinkel (2011).

a range of Cyp3a null and CYP3A-humanised mouse models for the study of CYP3A-mediated xenobiotic metabolism.[9]

The Cyp3a null model developed by this group completely lacks Cyp3a expression. It was generated by making two individual lines (one in which the seven closely linked Cyp3a genes had been knocked out by inserting loxP sites in the Cyp3a57 gene and downstream of the Cyp3a59 gene then deleting the floxed region by Pgk–Cre recombinase transfection in ESCs and one in which the distal Cyp3a13 gene was inactivated by replacing its promoter region and exons 1/2 with a Pgk–hygromycin cassette), then interbreeding them to generate a line which is completely null for Cyp3a (van Herwaarden *et al*., 2007). The resulting mice were viable and fertile. They had no physiological, haematological or clinical chemistry abnormalities, but there were compensatory changes in Cyp2c55 expression which led to unexpected findings such as the fact that Cyp2c55 was able to compensate for the absence of Cyp3a in the clearance of midazolam but not triazolam (van Waterschoot *et al*., 2009b; van Waterschoot *et al*., 2008). This illustrates the need for caution in interpreting data obtained using genetically modified mice, since the up-regulation of Cyp2c55 was unexpected; however, the effects on midazolam are unlikely to be relevant to humans because human CYP2C enzymes cannot metabolise midazolam.

The complete Cyp3a null model was subsequently humanised by interbreeding with transgenic mice which express CYP3A4 either in the liver (under the control of the apoliprotein E promoter) or in the small intestine (under the control of the villin promoter) (van Herwaarden *et al*., 2007; van Herwaarden *et al*., 2005). Tissue-specific hepatic and intestinal CYP3A4 levels in the two models were consistent with the promoter used to drive expression and similar to those in the corresponding tissues in humans and wild-type mice (van Herwaarden *et al*., 2007), the K_m and V_{max} values determined using hepatic and intestinal microsomes were comparable with the corresponding values from equivalent human tissue microsomes, and the presence of CYP3A4 normalised the expression of Cyp2c55 in the humanised lines (van Waterschoot *et al*., 2009c).

Example: Contribution of hepatic and intestinal CYP3A4 to the metabolism of docetaxel

In a demonstration of their utility, these lines were used to evaluate the relative contributions of hepatic and intestinal CYP3A4-mediated metabolism to the metabolism of the anticancer drug docetaxel. Docetaxel metabolite formation, which was absent in the combined Cyp3a null model, was reinstated in the predicted tissue-specific manner in the humanised models. This indicated that Cyp3a-mediated metabolism is a key determinant of oral docetaxel exposure; indeed, expression of CYP3A4 in the intestine (only) was sufficient to virtually block docetaxel entry from the gut following oral dosing although it had very little effect on its systemic clearance following intravenous administration. In contrast, hepatic CYP3A4 (expressed in the absence of intestinal CYP3A4) greatly reduced systemic exposure following intravenous administration but had little effect on oral absorption (van Herwaarden *et al*., 2007).

[9] For a concise summary of the available CYP3A models, see Van Herwaarden *et al*. (2009).

The species-specific regulation of CYP gene expression by nuclear receptors such as PXR means that models which are only humanised for a single CYP have limited applicability since they are, essentially, just mice with a human gene inserted. A key target for investigators in the field has, therefore, been to engineer mouse lines in which one or more human CYPs is/are regulated by one or more human receptors (e.g. CYP3A4 regulated by PXR[10]). One example of this approach is the mouse reported by Ma and colleagues. This mouse was generated by random transgenesis using BAC clones encoding human CYP3A4 and CYP3A7 (previously used to make a CY3PA4-humanised mouse on a wild-type background (Cheung *et al*., 2006)) and PXR together with their 5′ and 3′ flanking sequences. Humanised CYP3A4 mice were generated first, and then interbred with PXR null and humanised mice to generate double transgenics (TgCYP3A4/human PXR) (Ma *et al*., 2008). Treatment of these mice with PXR ligands generated a response which was more human-like than mouse-like (robust induction with rifampicin and weak induction with PCN). However, an unexpected sexual dimorphism was observed in that the adult male TgCYP3A4/human PXR mice did not express CYP3A4 in the liver whereas expression was sustained in the female mice. The authors state that it is unclear whether this reflects expression in humans because of the many CYP3A4-inducing agents to which humans are exposed to in adult life, which could mean that the high level of CYP3A4 expression observed in adult human liver in both sexes is a consequence of ongoing induction rather than constitutive expression. Nevertheless, this does represent a potential limitation of the model and serves to illustrate the pitfalls of trying to carry out multiple humanisations by repeated random transgenesis. Not only does the repeated use of random transgenesis increase the likelihood that one or more of the transgenes will be affected by positional effects (i.e. aberrant expression as a consequence of the site at which the transgene integrates), but it also means that the generation of a multiply modified line requires repeated rounds of breeding, which is very labour-intensive and leads to a great deal of wastage.

Example: Role of PXR in acute hepatotoxicity

This model has been used to examined the relationship between human PXR-regulated CYP3A4 function and the risk of toxicity in response to acetaminophen (Cheng *et al*., 2009). Treatment with rifampicin (to up-regulate human PXR-mediated responses) markedly increased the hepatotoxicity of acetaminophen in this model, being associated with increased expression of CYP3A4 and GSTA1, production of electrophilic metabolic intermediates, and depletion of GSH. This seemed to be associated with increased oxidative stress in the transgenic mice since they exhibited significantly higher cellular H_2O_2 levels than controls following treatment with acetaminophen (200 or 400 mg kg^{-1}). The production of more reactive oxygen species following acetaminophen treatment led to increased oxidative stress-induced hepatotoxicity and nephrotoxicity.

The problem with many of the humanised models described before around 2010 was that they were created by conventional transgenesis, meaning that it was very

[10] For succinct summaries of the available CYP3A null and humanised models, see Cheng, Ma, and Gonzalez (2011).

difficult to humanise multiple genes simultaneously. The knock-in approach used by Scheer *et al.* (2010, 2008) was therefore a significant step forward because it allowed a modular approach to humanisation. As well as humanising both PXR and CAR in a single mouse, this approach makes it possible to replace entire clusters of murine CYP genes with their human homologues. For example, Scheer and his colleagues have used a targeted gene replacement strategy to replace seven murine CYP3a genes on chromosome 5 with a BAC carrying human CYP3A4 and CYP3A7 (Hasegawa *et al.*, 2011). This allowed CYP3A4 to be expressed in mice under the control of its own promoter, leading to the interesting observation that expression in the liver was relatively low compared to robust intestinal expression of this gene whereas CYP3A4 is expressed at high (though variable) levels in human liver. This raises the question of whether the humanised mouse is failing to reflect human expression patterns correctly, or whether the high level of expression in humans is a consequence of exposure to inducing agents in the human diet and environment, which is perhaps more likely.

Interbreeding of huCYP3A4 mice with PXR-humanised and CAR-humanised mice made it possible to generate mice humanised for all three genes (PXR/CAR/CYP3A4/3A7- humanised) and use them to rank the potency of different PXR ligands in terms of their potential ability to up-regulate CYP3A4 expression in human liver. Expression of CYP3A4 was, as expected, up-regulated by rifampicin but not by PCN in this model.

Example: Pharmacokinetics of triazolam in triple humanised mice

When the effects of rifampicin, sulfinpyrazole and pioglitazone on the pharmacokinetics of the CYP3A4 substrate triazolam were characterised in triple humanised mice using clinically relevant doses (3–10 mg kg^{-1} for rifampicin, 2–10 mg kg^{-1} for sulfinpyrazole and ~2 mg kg^{-1} for pioglitazone), the reduction in AUC corresponded closely with the effects observed in the clinic (Table 7.2),

Table 7.2 Pharmacokinetic parameters of triazolam for different doses of rifampicin, sulfinpyrazone and pioglitazone in PXR humanised/CAR humanised/huCYP3A4/3A7 mice

Ligand	Dose (mg kg^{-1})	Reduction in AUC (%)	Effect in humans (%)
Rifampicin	1	52	73–96
	3	63	
	10	91	
Sulfinpyrazole	0.5	1	39
	2	15	
	10	37	
Pioglitazone	2	2	0–26
	10	0	
	50	−33 (increase)	

PXR/CAR/CYP3A4/3A7-humanised mice were given different oral doses of rifampicin, sulfinpyrazole or pioglitazone followed by a 5 mg kg^{-1} oral dose of triazolam. Data represent the mean ± SD. (n = 3 mice per compound and treatment). The triazolam C_{max} and AUC values from mice treated with rifampicin, sulfinpyrazole or pioglitazone were compared with those from the corresponding vehicle-treated control group.
Adapted from Hasegawa *et al.* (2011); table 2. Reproduced with permission of the American Society for Pharmacology and Experimental Therapeutics.

suggesting a possible use of mouse lines such as this to inform the design of chronic toxicities. It was interesting to note that in case of pioglitazone, which is not only a PXR activator but also a CYP3A4 inhibitor, there was actually an increase in the AUC of triazolam at the higher dose tested ($50\,mg\,kg^{-1}$), consistent with a dominant effect on CYP3A4 inhibition over induction at this dose. At a lower dose of $10\,mg\,kg^{-1}$, pioglitazone had no effect on AUC, suggesting that the induction and inhibition of CYP3A4 may have balanced each other out at this dose.

CYP2D In an even more ambitious gene replacement exercise, Scheer and his colleagues have also created a panel of lines in which they have replaced the entire murine Cyp2d cluster (containing nine functional genes) with different isoforms of CYP2D6 (CYP2D6*1, CYP2D6*2 and a novel variant provisionally called *CYP2D6*.N). The complete strategy used is too complex to describe here, but briefly what they did was to use two rounds of targeting in mouse ESCs to flank the mouse Cyp2d cluster with lox P sites, insert an expression cassette containing a human CYP2D6 gene (including 9 Kb of promoter, all the introns and exons, 700 bp of 3' flanking sequence and the poly A motif) and subsequently delete the mouse Cyp2d cluster by crossing the resulting mice with a line which expresses Cre recombinase in the germ line, leading to cre-mediated recombination of the floxed region. By also adding Flp recombinase recognition sites in the targeting vector, it was possible to delete the CYP2D6 construct to delete the entire Cyp2d cluster. This represents a novel method for creating humanised mouse lines in which large multi-gene families of linked genes have been deleted. The lines still require detailed characterisation, but preliminary results (Scheer *et al.*, 2012) suggest that, at least for certain CYP2D6 substrates such as bufuralol, the huCYP2D6, wild-type and Cyp2d null lines may be representative of ultrarapid, extensive and poor metabolisers, respectively.

These developments in targeted transgenesis have made it much easier to generate mice carrying multiple humanised transgenes which are expressed in a physiologically appropriate manner. However, not everybody considers this to be the definitive answer to the problem of how to humanise xenobiotic responses in mice. Igarashi and colleagues, for example, are worried that a fully humanised PXR would not bind effectively to the regulatory sequences of target genes in the mouse. One way round this problem is to retain the mouse flanking sequences and DNA-binding domain of PXR and just replace the ligand-binding domain with the corresponding human sequence. This has been done in creating the hSXRki mouse, which contains a cDNA encoding the human ligand-binding domain homologously recombined into the mouse gene at the 3′ end of exon 3 of PXR (Igarashi *et al.*, 2012). This line expresses PXR mRNA in a range of target tissues and responds to rifampicin with induction of Cyp3a11 and Ces6.

At the time of writing the model is too new to be able to determine whether it will actually prove to reflect human PXR-mediated responses any better than the tADMET™ models; however, the lack of effective binding to target regulatory sequences should not be a problem in the tADMET™ models such as the PXR/CAR/CYP3A4/3A7-humanised line, since these contain fully humanised target genes as well as humanised receptors, although even in these models the human receptor will have to interact with endogenous mouse dimerisation

partners and coregulators. The attraction of the tADMET™ models is that they contain entire tracts of relevant DNA, including all the necessary human flanking sequences, so their responses should, in principle be, as reflective of the human situation as is possible in a mouse. Further tADMET™ models are in development; for the latest developments see the tADMET™ page on the Taconic web site.[11]

UGT models Genetic engineering of the UGT1 gene is particularly difficult because this gene generates functional diversity via an unusual mechanism based on an alternative splicing mechanism. Briefly, the UGT1 gene has nine different exons 1 distributed over 150 kb of the genome. Each of these can be spliced to the conserved exons 2–3 to generate an mRNA encoding a functional UGT1 protein. The difficulty of carrying out genetic engineering using such large tracts of DNA meant that the UGT family lagged behind other xenobiotic metabolising enzymes in the generation of null and humanised models. However, a series of papers describing UGT models has been published since 2005 by the group of Robert Tukey. The first UGT1-humanised mouse, containing a BAC construct including all the functional exon 1 sequences (UGT1A1–UGT1A10) as well as the conserved exon 2–5 sequences, was generated by conventional transgenesis and bred onto a wild-type C57BL/6N background (Chen *et al.*, 2005). Somewhat surprisingly, this line had no obvious phenotypic abnormalities despite carrying multiple copies of such a large construct and it expressed human UGT1A1, UGT1A4 and UGT1A6 in the liver and GI tract. The basal level of expression varied depending upon the founder mouse, which is not unusual for lines generated by conventional transgenesis because the construct will have integrated at a different site in each founder. The expression of the various UGT1 proteins was inducible by TCDD, PCN and the peroxisome proliferator WY-14643, demonstrating regulation via the AhR, PXR and peroxisome proliferator-activated receptor (PPAR) α (Chen *et al.*, 2005; Senekeo-Effenberger *et al.*, 2007). These mice have been used to study the Nrf2-KEAP1-mediated regulation of UGT1 in response to oxidative stress (Yueh and Tukey, 2007) and the role of UGT-mediated metabolism in the pharmacokinetics of various drugs (Argikar *et al.*, 2009; Cai *et al.*, 2010) because UGT1A4 is a pseudogene in mice they are considered to be fully humanised for UGT1A4 without any further engineering (Argikar *et al.*, 2009), but in order to eliminate interference from murine UGT1 activity the humanised line has been crossed to generate humanised mice with a Ugt1 null background (Cai *et al.*, 2010). This had to be done in two rounds of breeding because the Ugt1 null genotype is lethal within 7 days after birth due to fatal hyperbilirubinaemia (Nguyen *et al.*, 2008), so the humanised genotype had to be first bred onto a $Ugt1^{+/-}$ background, then the offspring were interbred to generate UGT1A/$Ugt1^{-/-}$ mice.

Example: Role of UGTs in neonatal jaundice

Once humanised mice with the Ugt null genotype were available they could be used to study the main pathological process involving UGTs, neonatal hyperbilirubinaemia. In humans, this can range from a mild jaundice, which is common in otherwise healthy babies due to a delay in the neonatal up-regulation of

[11] http://www.taconic.com/wmspage.cfm?parm1=1792

UGT1A1 expression and normally resolves itself within a few days, to the severe phenotype of Criggler–Najjar syndrome, characterized by a complete inactivation of UGT1A1-dependent bilirubin glucuronidation activity. An intermediate form of hyperbilirubinaemia is observed in Gilbert's Syndrome, which is associated with a variant form of UGT1A1 (UGT1A1*28). The generation of lines of humanised mice with either UGT1A1*1 or UGT1A1*28 on a Ugt1 null background made it possible to characterise hyperbilirubinaemia under normal or Gilbert's Syndrome-like conditions in the absence of murine Ugt1 activity (Fujiwara *et al*., 2010). The accumulation of unconjugated bilirubin was slower in humanised than Ugt1 null mice, although the serum levels reached (>150 mg ml^{-1} by day 14) were high enough to cause seizures and severe neurological damage (kernicterus). If the mice survived beyond this point, serum bilirubin levels declined rapidly as intestinal UGT1A1 expression increased with age. Consistent with observations in Gilbert's syndrome, the level of expression of UGT1A1*28 was lower in humanised mice than that of UGT1A1*1. Thus humanised mice which survive beyond the age of 2 weeks can be used as a model for the study of adult jaundice as well as the accumulation of unconjugated bilirubin in brain and onset of seizures in neonatal hyperbilirubinaemia.

Example: Gestational regulation of UGT1

In recent studies, the UGT1A1*1 and UGT1A1*28 transgenes were bred onto PXR null (Xie *et al*., 2000) and CAR null genotypes with a view to studying the roles of PXR and CAR in regulating UGT1 expression during gestation (Chen *et al*., 2012). This exercise revealed that, while CAR seemed to have little role in the regulation of expression during gestation, PXR plays a key role in controlling gestational serum bilirubin levels. The inducibility of UGT1A1 during gestation was markedly reduced on a PXR null background whereas this was reversed in newborn mice. Up-regulation of UGT1A1 expression on the PXR null background prevented the neonatal hyperbilirubinaemia which was seen on a PXR wild-type background, suggesting that non-ligand-bound PXR acts as a repressor of UGT1A1 expression in newborn mice. Interestingly, neonatal UGT1A1/Ugt1a$^{-/-}$ mice given human breast milk developed a more severe hyperbilirubinaemia than those given baby formula; this did not happen on a PXR null background but did still happen in CAR null mice, and was interpreted as indicating that human breast milk contains factors which suppress the expression of UGT1A1 (Fujiwara *et al*., 2012). Further studies using UGT1 humanised and null mice in combination with various transcription factor models should help to enhance our understanding of the processes involved in neonatal jaundice and more severe forms of hyperbilirubinaemia, and hopefully lead to better preventive strategies and/or treatments.

7.2.3 Drug transporter models

Drug transporter null mouse lines Drug transporters, membrane proteins which mediate the cellular uptake and efflux of molecules, play a key role in toxicokinetics because they influence the oral bioavailability, disposition and excretion of xenobiotics. The field of drug transporter research is less advanced

than those of nuclear receptor function and CYP-mediated metabolism, but is now advancing rapidly. Until recently the generation of transgenic drug transporter models focused on the ABC transporters (the multidrug resistance (MDR) and multidrug-resistance-related protein (MRP)) gene families,[12] but over the last few years considerable effort has also been devoted to the organic anion transporters (OATPs), which mediate the sodium-independent uptake of endogenous compounds and xenobiotics (Iusuf, Van De Steeg, and Schinkel, 2012). Organic anion-transporting polypeptides are localised in the liver, kidney, intestine and capillary endothelial cells of the brain and may therefore play a role in the CNS action and toxicity of drugs and toxicants.

Hepatic efflux transporters Hepatic efflux transporters are found both in the apical (canalicular: e.g. MDR1) and basolateral (sinusoidal: e.g. MRP1 and MRP3) membranes of the hepatocyte. The apical transporters move substrates into the bile while the basolateral ones move them towards the systemic circulation.[13]

The product of the human MDR1 gene was first identified as the MDR protein, P-glycoprotein, which was responsible for drug resistance in many tumour-derived cell lines. Its expression was subsequently detected in the membranes of a number of normal polarized cell types, including the brush borders of renal proximal tubules, the biliary membranes of hepatocytes, the apical membranes of intestinal mucosal cells, the adrenal gland, and capillary endothelial cells of the brain and testis. The expression of P-glycoprotein at the tip of the intestinal villus means that it is ideally placed to limit compound absorption from the intestinal lumen. The pattern of P-glycoprotein expression, together with its known role in drug resistance, suggested that it might have a normal function in the removal of xenobiotics and other potentially toxic substances from susceptible tissues.

The effort to generate transgenic models for drug transport via P-glycoprotein is complicated by the fact that mice have two *Mdr1* genes (*Mdr1a* and *Mdr1b*) whose combined functions are believed to add up to those of P-glycoprotein in humans. This means that any attempt to delete or humanize *Mdr1* function in mice must address both the murine genes, and in recent years mice deficient for two or more transporters have been generated and characterised to look at the functional overlap and complementarity of ABC transporters.

Mdr1a null mice were first reported in 1994. They were viable, fertile and apparently normal. *Mdr1a* mRNA was detectable in null mice but no functional P-glycoprotein was produced. However, a compensatory increase in hepatic and renal *Mdr1b* expression (at mRNA and protein levels) was detected in heterozygous ($Mdr1a^{+/-}$) and homozygous ($Mdr1a^{-/-}$) mice. The possibility of confounding effects due to overexpression of *Mdr1b* must therefore be borne in mind when evaluating the results of studies in *Mdr1a* null mice.

Interest in the susceptibility of these mice to toxic chemicals arose when the mice were sprayed with ivermectin to treat a mite infestation. The null mice were 50–100 times more sensitive to the adverse effects of this pesticide, which is normally non-toxic to mice, and the concentration of ivermectin in their brains was 90× higher than in wild-type littermates following oral dosing at 0.2 mg kg^{-1}.

[12] For a detailed review of this topic, and the mouse models described later, see Stanley *et al.* (2009).
[13] For a general review, see Kock and Brouwer (2012).

The mice also exhibited slower clearance and increased tissue concentrations when dosed with the anticancer drug vinblastine. The *Mdr1a* null mouse model subsequently found wide use in the study of *in vivo* distribution of drugs and xenobiotics although the model has now been largely superseded by the Mdr1a/Mdr1b double null model.

The generation of *Mdr1b* null mice was reported in 1997. *Mdr1b* null mice had normal viability, fertility and lifespan and no obvious physiological abnormalities. Their macro- and microscopic anatomy was normal. The *Mdr1b* null phenotype had no effect on the expression of *Mdr1a* at the mRNA level and in contrast with the results obtained in *Mdr1a* null mice, the *Mdr1b* null phenotype had no effect on the pharmacokinetics of digoxin.

The possibility of erroneous results due to overexpression of *Mdr1b* in *Mdr1a* null mice made it critically important to develop a mouse line in which both genes were knocked out. The close linkage between *Mdr1a* and *Mdr1b* made it impossible to generate double null mice by interbreeding *Mdr1a* null and *Mdr1b* null mice so a new line in which both genes were deleted was generated. Like the *Mdr1a* and *Mdr1b* single null mice, the resulting animals were normal with respect to development, viability, fertility, macro/microanatomy and biochemical/haematological parameters. No changes in bile flow or biliary output of phospholipids, bile acids, GSH and cholesterol were observed, neither were there any changes in the expression of other transporter genes.

Mrp1 null mice were reported by two groups in 1997. Both lines were healthy and fertile with no histological abnormalities and Western blot analysis indicated that no Mrp1 protein was expressed by these mice. The results of experiments using erythrocytes and bone marrow-derived mast cells indicated that a key GSH conjugate transporter was absent from these mice. *Mrp1* null mice were hypersensitive to etoposide both *in vivo* and *in vitro*, *in vivo* toxicity being associated with the development of polyuria and diabetes insipidus.

Mrp2 null mice have been generated commercially by Deltagen, Inc. The absence of *Mrp2* was confirmed by Western blot analysis. Null mice had no gross abnormalities, their lifespan was normal and offsprings were born at the predicted Mendelian ratios. No liver toxicity was observed in null mice, but total circulating bilirubin levels were fivefold higher in null animals than in wild-type mice.

A second, independently generated, line of *Mrp2* null mice was reported in 2006. Again the *Mrp2* null mice were fertile, were born at predicted Mendelian ratios and had normal lifespan and body weights. They had no obvious macro- or microscopic abnormalities except for some liver enlargement (120–125% of wild type) which was associated with slight (twofold) elevations in *Mrp*3 and *Mrp*4 expression. The only change in clinical chemistry parameters in *Mrp2* null mice was a moderate increase in total bilirubin which was mainly due to conjugated bilirubin.

Various double and triple null mouse lines, including *Mdr1a/Mdr1b* double null (functionally equivalent to a P-glycoprotein null), *Mdr1a*/1b/*Mrp1*, *Mdr1a*/1b/*Mrp2*, *Mdr1a*/1b/*Bcrp* and triple null mice have also been generated, opening up the possibility of discriminating between the roles of P-glycoprotein, *Mrp1* and *Mrp2 in vivo*.[14]

[14] For an overview of the various multiple knockout models available as of 2009, see Lagas, Vlaming, and Schinkel (2009).

Combined deletion of CYPs and transporters Mice lacking the functions of both Cyp3a and P-glycoprotein have been generated by crossing the existing Cyp3a null and *Mdr1a/1b* null lines on a 99% FVB background (van Waterschoot *et al.*, 2009a). Despite lacking two important genetic regions, these mice are viable, fertile and have normal body weights and lifespans. They have no marked physiological abnormalities and their clinical chemistry and haematology profiles are more-or-less normal apart from moderate increases in serum creatinine, urea and triglycerides, which were not seen in the parent lines. However, following oral dosing with docetaxel they exhibited a 72-fold increase in AUC (17-fold following i.v. dosing) and a corresponding increase in oral bioavailability, which was about 45% in this line compared with 10% in wild-type mice. This is associated with exacerbated toxicity in the combined null model: following three daily oral doses of 10 mg kg^{-1}, combined null mice died on days 4–6 due to severe intestinal toxicity whereas no marked toxicity was observed in wild-type or *Mdr1a/1b* null mice. This effect, which is likely to be a consequence of the inability of enterocytes lacking *Mdr1* and Cyp3a activities to exclude and/or metabolise docetaxel as it enters from the gut lumen, may have implications for the use of drugs with narrow therapeutic windows, both in terms of possible DDIs and interindividual differences in susceptibility to adverse side effects.

Hepatic uptake transporters The OATPs comprise six families, among which the most important for drug uptake are the OATP1A, OATP1B and OATP2B1 families.[15] The substrates of OATP1A include bile acids, oestrogen derivatives, peptides and drugs (including fexofenadine, pravastatin, rosuvastatin, D-penicillamine and methotrexate), while OATPs 1B1 and 1B3 also transport conjugated and unconjugated bilirubin, thyroid hormones, digoxin, paclitaxel, rifampicin, enalapril and SN-38. The functions of the OATP2 family still require clarification.

Various null mouse models are available, but as there are no clear orthologous relationships between the human and mouse OATP1A and 1B families the issue of humanisation is problematic (Iusuf, Van De Steeg, and Schinkel, 2012). Furthermore, the significant overlap in tissue distribution and substrate specificity of *Oatp1a* and *Oatp1b* leads to redundancy and means that there is likely to be compensation in single null lines. To get round this problem, Van de Steeg *et al.* (2010) created a mouse model in which all five established *Oatp1a* and *Oatp1b* genes were knocked out. They achieved this by floxing both ends of the gene cluster which encodes *Oatp1a* and *Oatp1b* (620 kb) and then deleting them by transfecting Pgk–Cre recombinase into the floxed ESCs. The *Oatp1c* gene was left intact because it is believed to have a specific physiological role, possibly involving thyroid hormone uptake. The resulting mice were viable, fertile and had a normal lifespan, although they exhibited increased body weight in adulthood and gradually developed jaundice due to hyperbilirubinaemia. Apart from slight down-regulation of *Mdr1a*, there were no marked changes in the expression of xenobiotic metabolising and transporter systems in these mice. Their clinical chemistry profiles were essentially normal, but they did have >40-fold higher

[15] For a general review, see Fenner *et al.* (2012).

plasma bilirubin levels, mainly due to conjugated bilirubin, compared with both wild-type and single null lines (in which the effects seen are subtle). This is thought to reflect a cycling mechanism in the liver, involving a so-called 'hepatocyte-hopping' process whereby bilirubin is excreted from hepatocytes into the sinusoids following conjugation, taken up again further downstream by OATPs and then finally excreted into the bile (Iusuf, Van De Steeg, and Schinkel, 2012). Inhibition of this process could generate higher levels of conjugated bilirubin in venous blood leaving the liver, thus accounting for the increased circulating levels observed in double null mice. A similar effect could explain the sudden rise in plasma conjugated bilirubin that is sometimes observed in humans following drug treatment; it could be due to inhibition of OATP-mediated transport.

Other studies using *Oatp1a/1b* double null mice have implicated these transporters in bile acid homeostasis, transport of anionic oestrogen conjugates and drug clearance.

Example: Role of OATP1B1 in the disposition of methotrexate

The anionic anticancer/antirheumatic drug methotrexate is a substrate for human OATP1B1/murine *Oatp1b1 in vitro*. A marked decrease in hepatic uptake (25-fold) is observed in *Oatp1a/1b* null mice following i.v. dosing, while increased hepatic uptake and decreased plasma concentrations are observed in mice humanised for OATP1B1 (Iusuf, Van De Steeg, and Schinkel, 2012). The plasma AUC of methotrexate is increased in *Oatp1a/1b* null mice after either i.v. or oral dosing, and the observed differences are sustained for 2 h following an oral dose. This suggests a possible mechanism for the observation that patients with low-activity variants of OATP1B1 have increased plasma concentrations of methotrexate associated with reduced GI toxicity, possibly due to reduced liver uptake resulting in diminished biliary excretion of methotrexate, which thus reduced direct intestinal exposure and hence the intestinal toxicity of this drug. A similar effect is observed in wild-type mice treated with the *Oatp1a/1b* inhibitor rifampicin ($20\,mg\,kg^{-1}$ i.v. 3 min before dosing with methotrexate at $10\,mg\,kg^{-1}$ i.v.) (van de Steeg *et al.*, 2010).

Research using drug transporter null mouse lines Drug transporter null mouse models have been used to study systemic and tissue-specific toxicity, effects on pharmacokinetics and bioavailability, transport across the BBB and hepatic uptake/excretion. They may also have value as disease models and in the study of pathological processes. For example, the *Mdr1a* null mice developed by Schinkel and colleagues have a tendency to develop intestinal inflammation even when kept under specific pathogen-free conditions. The condition was similar to the human inflammatory bowel diseases ulcerative colitis and Crohn's disease. It has therefore been suggested that *Mdr1a* null mice could be used as a model for human inflammatory bowel disease as it develops colitis without exhibiting any other gross morphological abnormalities. The development of colitis, which is more severe in *Mdr1a* null mice housed under conventional than specific pathogen-free conditions, appears to involve epithelial dysfunction rather than defects in the immune system and it has been hypothesised that the disease develops because of an inability to efflux bacterial toxins to the intestinal lumen for excretion.

7.3 Reporter models

The changes in gene expression which occur in response to toxic chemicals are often difficult to measure directly because, for example, the functions of the molecules induced are difficult to measure or the molecules themselves are very labile. Developments in transgenic technologies, particularly the availability of coloured, fluorescent and luminescent reporter molecules which can be linked to the regulatory regions of key genes, have provided the opportunity to generate reporter models which can be used to develop *in vitro* screening tests (Emter, Ellis, and Natsch, 2010; Uibel *et al*., 2010; Hendriks *et al*., 2012), elucidate mechanisms of toxicity and potentially for *in vivo* screening.[16] The models developed to date have tended to focus on CYP induction (using the regulatory region of a CYP gene to drive the expression of a reporter molecule) and responses to insults such as oxidative stress or DNA damage.

According to Maggi *et al*. (2004), the ideal reporter model for studying the effects of pharmacologically active/toxic chemicals would allow the following.

- Rapid assessment of all organs where a given compound is active.
- Evaluation of the response to the compound after repeated administration.
- Measurement of the response of different tissues following administration via different routes.
- Evaluation of the minimum dose of compound necessary to obtain the toxicological response.
- Identification of active metabolites and their profiles of action.
- Detection of sites of compound accumulation during chronic administration.

In order to use reporter models as surrogates for studies on endogenous gene expression, it is essential to verify that the expression of the reporter accurately reflects that of the endogenous gene. It is often difficult to meet this requirement, and the toxicological use of reporter models has been held back because of the amount of work involved in validation. This section therefore summarises the types of reporter which are currently available, focussing on the strengths and limitations of the different types of reporter gene used, and provides examples from the small number of models for which toxicological data are currently available.

7.3.1 LacZ-*based models*

The *LacZ* gene of the bacterium *Escherichia coli* has, for a number of years, been a popular reporter for demonstrating the cellular localisation of transgene expression from a variety of promoters. It is particularly useful for studies involving regulation via *cis*-acting regulatory sequences (Takahashi *et al*., 2000). It has been extensively used in the study of developmental gene regulation as well as for the study of CYP3A4 regulation and toxic responses involving oxidative stress and DNA damage. Its strengths as a reporter include the fact that it is easy to assay and suitable for visualisation on tissue sections by microscopy.

A mouse model engineered by conventional transgenesis to contain 13 kb of the 5′ flanking region of the human CYP3A4 gene (−13 kb to +53 bp, including

[16] For general background to the different types of reporters available, see Ghim *et al.* (2010).

the XREM) linked to *LacZ* (-13CYP3A4/*LacZ*) and a sister model containing a smaller fragment (–3200 to +53 bp) of the 5′ flanking region (-3.2CYP3A4/*LacZ*) have been used to study the tissue-specific expression of CYP3A4 (Robertson *et al*., 2003). No reporter expression was detected in the -3.2CYP3A4 model, whether or not the animals were treated with inducing agents, suggesting that this proximal region of the CYP3A4 gene is not sufficient to drive expression in mouse liver. Of the four –13CYP3A4 lines examined, two did not express the reporter, one exhibited both basal and induced expression and one expressed the reporter only after induction. This is typical of the problems experienced with conventional transgenesis: the transgene integrates in different locations in each founder line and this may lead to marked differences in its regulation and expression.

The line showing the greatest extent of induction was characterised in terms of tissue-specific reporter expression and response to different inducing agents. The pattern of expression of the transgene corresponded with that of endogenous Cyp3a11, it responded as predicted to inducing agents and induction by dexamethasone (1–100 mg kg^{-1} i.p.) was dose-dependent, although this was associated with recruitment of additional cells with the capacity to express the reporter rather than up-regulation in cells which are already expressing it. Furthermore, comparison of reporter expression between embryos at 17 days of gestation and 21 days postpartum demonstrated developmental regulation. These studies, which confirm that the proximal promoter alone cannot drive expression but 5′ flanking sequences of CYP3A4 can drive expression in absence of the remainder of gene cluster, laid the foundation for the use of this model in studies of CYP3A4 regulation by xenobiotics and during preclinical screening of drug candidates to ensure the exclusion of potential CYP3A4- inducing agents from Phase I trials, although the authors acknowledge that it is dependent upon the activity of mouse receptors (PXR and CAR). These mice have been used to look at feed-forward regulation of bile acid detoxification, using bile duct ligation as a model of acute cholestasis (Stedman *et al*., 2004), elegantly demonstrating up-regulation of CYP3A4 expression in response to the accumulation of toxic bile acids.

The expression of *LacZ* in reporter mice can also be used to visualise the effects of toxic insults such as DNA damage and oxidative stress. One example of this approach is a model developed using the promoter of the murine cyclin-dependent kinase inhibitor 1 (p21/Waf1) gene, which mediates cell cycle G1 phase arrest in response to a variety of types of cellular stress. A 4.5 kb fragment covering the entire p21 promoter region was amplified from CD-1 mouse genomic DNA, ligated to the *LacZ* reporter gene and microinjected into C57BL/6 x CBA zygotes (Vasey *et al*., 2008). The resulting transgenic lines responded to insults such as etoposide (40 mg kg^{-1}, inducing DNA damage), LPS (1.5 mg kg^{-1}, inducing acute inflammation) and mercuric chloride (8.5 mg kg^{-1}, inducing oxidative stress) with up-regulation of the reporter in a number of tissues. The spatial expression profile of the reporter was compound-specific, suggesting that this line could be used to identify the type of stress induced by a particular compound as well as for studies of tissue-specific toxicity, developmental changes in p21 expression and disease processes involving p21.

Example: Developmental expression of p21 in the p21–*LacZ* mouse model

An initial study aimed at examining developmental changes in reporter expression in the p21–*LacZ* mouse model was reported in 2011 (Vasey *et al*., 2011).

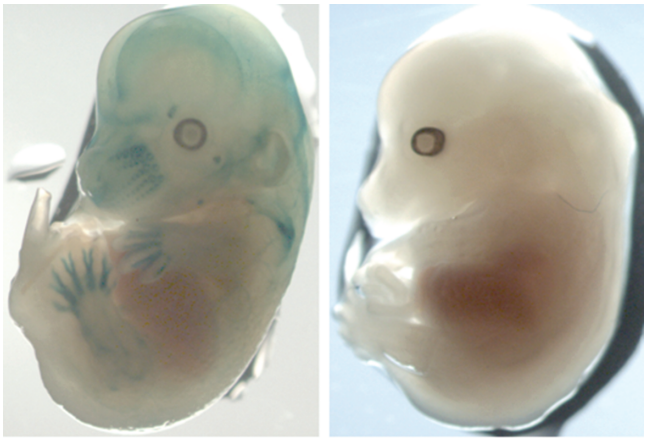


Figure 7.5 ***LacZ* staining in p21–*LacZ* and nontransgenic 14.5 dpc embryos.** *LacZ* staining was carried out on 14.5 dpc embryos of p21–*LacZ* and non-transgenic CBA/C57BL/6 genetic background. Staining was detected in various regions of the p21–*LacZ* embryo while no staining was detected in the nontransgenic embryo (source: Vasey *et al*. (2008); supplementary figure 1. Reproduced with permission of Elsevier)

Reporter expression was visualised in mouse embryos at days 11.5, 12.5 and 14.5 of gestation. Expression was extremely restricted on day 11.5, but by days 12.5–14.5 reporter expression was evident in the scapula, head and limb regions. The most intense staining was observed in areas which were undergoing remodelling, such as between the digits of the developing paws (Figure 7.5). Interesting, expression of the reporter did not occur concurrently with apoptosis (as was initially expected); instead, it seemed to precede the onset of apoptosis. Suggested explanations for this include the possibility that p21 is responsible for triggering cells to exit the cell cycle prior to apoptosis, or that p21 is responsible for setting the boundaries of remodelling so that apoptosis does not exceed the requirements of normal development. This study illustrates the way in which a reporter model such as the p21–*LacZ* mouse can be used to reveal the spatial and temporal patterns of p21 expression in areas undergoing developmental remodelling, terminal differentiation and apoptosis. By implication, this model will be also valuable in studying the ways in which toxic insults can disrupt these important processes.

Another *LacZ*-based reporter line has been made to express *LacZ* under the control of the HO-1 promoter, which mediates cellular responses to heavy metals and oxidative stress (Young *et al*., 2010). A BAC clone was engineered to contain HO-1 promoter-driven *LacZ* by replacing exon 1 and intron 1 of the *Hmox1* gene with the *LacZ* cassette. The resulting reporter construct contained most of the *Hmox1* gene together with 16.5 kb of 5′ flanking sequence and 8 kb of 3′ flanking sequence. Instead of allowing it to integrate randomly, insertion was targeted to the housekeeping hypoxanthine-guanine phosphoribosyltransferase (HPRT) locus on the X-chromosome. This locus was chosen because it is effective at supporting promoter-dependent transgene expression, although it can be problematic because, being on the X-chromosome, it is subject to the process of X-inactivation in females, which leads to mosaicism whereby only about 50% of

the cells in a given tissue will have an active X-chromosome carrying the transgene. This is not a problem for secreted markers or those which are measured by biochemical methods, but it means that *LacZ* staining is seen as a patchwork pattern in responsive tissues of female mice.

Example: Effects of toxic chemicals on HMOX–*LacZ* reporter expression

Intense expression of *LacZ* was visualised in the spleen and alveolar epithelium of transgenic mice, while low-level basal expression was also detected in the liver, thymus and brain. When HMOX–*LacZ* mice were treated with cadmium chloride (0, 1, 2, 4 or 8 mg kg^{-1} i.p.), evaluation of Haematoxylin and Eosin stained sections indicated no marked hepatic abnormalities in naive and vehicle-treated control animals, whereas the appearance of liver from the cadmium chloride treated animals indicated acute toxicity manifesting as hepatocyte vacuolation and areas of necrosis associated with deposits of erythrocytes. This was associated with *LacZ* staining whose intensity increased with dose. The *LacZ* response coincided with the induction of HO-1 mRNA but preceded and was more sensitive than liver enzyme release, demonstrating the potential of *LacZ* as a sensitive reporter gene for the early detection of heavy metal effects.

Further studies were conducted using haemin (65 mg kg^{-1}, inducing hypoxia), etoposide (40 mg kg^{-1}, inducing DNA damage) and LPS (1.5 mg kg^{-1}, inducing acute inflammation). All three compounds induced reporter expression in the liver and lung but with different patterns within and between tissues. Hepatic induction by haemin peaked about 8 h after dosing, lagging behind the up-regulation of the endogenous HO-1 mRNA (presumably due to the time taken for mRNAs to be processed and translated into protein). Staining could still be seen after 24 h, but the authors note that the *LacZ* reporter is limited in terms of characterising the time taken for a response to decline because the enzyme itself has a half-life of about 24 h. The specific spatial patterns of induction seen with different types of stress suggest that one application of this model might be to classify the nature of the cellular stresses induced by test compounds.

7.3.2 Green fluorescent protein-based models

Aequorea victoria is a brightly luminescent jellyfish which emits light from yellow tissue masses that consist of about 6000–7000 photogenic cells each. The components required for bioluminescence include a Ca^{2+}-activated photoprotein, aequorin, which emits blue-green light, and an accessory green fluorescent protein (GFP), which accepts energy from aequorin and re-emits it as green light. The gene encoding *A. victoria* GFP has been cloned and has emerged as a versatile visual reporter gene (Ikawa *et al.*, 1998). GFP is an extremely stable protein of 238 amino acids which fluoresces maximally when excited at 400 nm with a lesser peak at 475 nm, and fluorescence emission peaks at 509 nm. Unlike enzyme-based fluorescent systems, the GFP reporter does not require a substrate in order to emit fluorescence, but only needs the input of energy in order to excite the fluorophore.

The major strength of GFP as a reporter is that it allows the direct imaging of fluorescence in living cells. As compared with LacZ (Chiocchetti *et al.*, 1997), its advantages include the following.

- Sensitivity: This allows the detection of cells with low levels of reporter expression to be identified.
- Expression throughout the lifespan: GFP is expressed throughout the lifespan of transgenic mice, from the preimplantation stage to adulthood. When a promoter driving ubiquitous expression is used, GFP transgenic mice can be identified at birth because they are bright green!
- Lack of toxicity: GFP is non-toxic, does not alter the subcellular localisation of other proteins and does not interfere with cell growth or function.
- Good distribution: The reporter diffuses evenly throughout the cell and can travel long distances, for example, to the ends of neurons and glia.
- Stability: The reporter is not affected by chemical denaturants.
- Simplicity of sample handling: No chemical fixation, addition of an exogenous substrate or antibody labelling is required.

Wild-type GFP fluoresces relatively weakly, and a second generation of coloured fluorescent reporters has been developed in order to overcome this limitation. Many of these contain a mutation within the chromophore region of the protein (e.g. S65T encodes an amino acid change from serine to threonine at residue 65, which shifts the excitation peak of the protein to 489 nm, optimising its spectrum for detection using fluorescence microscopy or flow cytometry). Furthermore, the development of enhanced colour variants of GFP (e.g. enhanced cyan fluorescent protein (blue-shifted) and enhanced yellow fluorescent protein (red-shifted)) has made it possible to combine multiple fluorescent reporters in a single animal, allowing two or more responses to be evaluated simultaneously (Hadjantonakis, Macmaster and Nagy, 2002).

The potential use of GFP as a reporter for effects induced by xenobiotics is illustrated by a transgenic line expressing GFP driven by the rat CYP1A1 promoter on a mixed C57BL/6 and CBA background. A red shift variant of GFP which has brighter fluorescence and higher expression levels in mammalian cells was expressed under the control of the flanking sequence and promoter region of rat CYP1A1. Intriguingly, transgenic offsprings were easily recognisable by their green ears (!) (Figure 7.6) and it turned out that the reporter was constitutively expressed in the sebaceous gland, leading to the suggestion that CYP1A1 could play a role in acne (Rowe *et al*., 2008).

Preliminary experiments in which the CYP1A1-GFP reporter mice were treated with 3-MC (40 mg kg^{-1} in corn oil i.p.) three times at 24-h intervals and sacrificed 24 h after the final dose demonstrated up-regulation of the reporter in the liver as expected; however, no further studies using these mice have been reported.

7.3.3 Luciferase-based models

In recent years, the firefly luciferase gene has become the reporter of choice in many contexts. The advantages of using luciferase as a reporter include:

- rapid turnover/short half-life
- absence of accumulation in target cells
- predicted good signal/noise ratio in mammalian cells (because it is an insect enzyme)
- ease of quantitation in tissue extracts

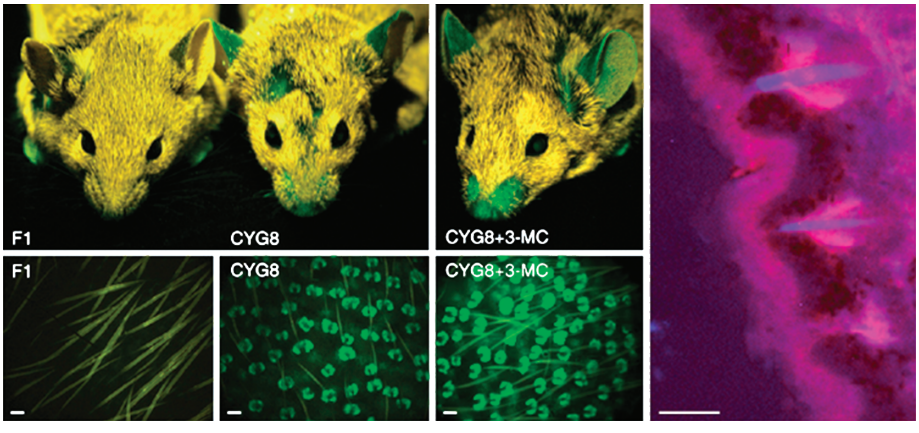


Figure 7.6 CYP1A1 gene activity in sebaceous gland of mice. (a) Wild-type and (b, c) transgenic CYP1A1–EGFP live mouse and ear punch illuminated with blue light of wavelength 488 nm and detected by monitoring light of wavelength 509 nm. Examples of basal transgene activity (b) or transgene activity induced by intraperitoneal injection of 3-methylcholanthrene suspension in corn oil (40 mg g^{-1} body weight) are shown. (d) Immunostained ear section of mouse shown in (b) confirming predominant CYP1A1 (pink) location to the sebaceous gland surrounding the hair shaft (blue). Bar = 1 mm (source: Rowe *et al.* (2008); figure 1. Reproduced with permission of Nature Publishing Group)

- availability of high-sensitivity enzymatic assays for luciferase activity
- availability of monospecific antibodies for *in situ* localisation
- suitability for optical imaging

Luciferase-based reporter systems have also been used to study the induction of human CYP genes in a mouse setting; for example, a mouse line (CYPLucR) which expresses luciferase under the control of nucleotides −1612 to +292 of the human CYP1A1 promoter provided the opportunity 'to investigate the contribution of regulatory factors that modulate gene expression as well as structural determinants of the gene that may influence expression patterns *in vivo*' (Galijatovic *et al.*, 2004). The results obtained were compared with those in a line containing the full-length human CYP1A1 gene together with 9 kb of flanking sequence (CYP1A1N). The hepatic and pulmonary expression patterns of the luciferase reporter gene and the human transgene following TCDD induction were similar; luciferase was localised to hepatocytes and corresponded with the induction of both murine Cyp1a1 and human CYP1A1. However, the inducibility of the CYPLucR and CYP1A1N transgenes differed in some ways from that of the endogenous murine Cyp1a1 gene: in contrast to the endogenous gene, neither transgene was inducible in the kidney, whereas the CYPLucR reporter gene was expressed at unexpectedly high levels in the brain. These problems may be a function of

- the use of a limited amount of flanking DNA to control the expression of the reporter, possibly omitting key regulatory regions;

- the absence of essential human transcription factors in certain mouse tissues;
- the loss of long-range human gene–gene interactions in the transgenic mouse model;
- the effects due to the location of the integrated transgene (integration site effects).

The company Xenogen (now owned by Perkin-Elmer[17]) has focussed on the development of luciferase-based reporter mice and their use in whole body bioluminescent imaging to visualise the expression of xenobiotic-responsive genes such as CYPs (Zhang *et al.*, 2003a; Zhang *et al.*, 2003b; Zhang *et al.*, 2004a; Zhang *et al.*, 2004b) as well as those which play a role in toxic responses. They argue that the generation of transgenic promoter-luciferase animals for genes regulated by specific toxic processes, coupled with real-time evaluation of site-specific gene expression will provide novel, non-invasive biomarkers which are predictive of developing toxicity *in vivo*. The procedure for bioluminescent imaging of live luciferase-transgenic rodents is summarised in Box 7.1.

Box 7.1 Bioluminescent imaging of mice

The method used for *in vivo* imaging requires a series of injections and manipulations but is otherwise non-invasive. At the appropriate time points after dosing or other procedures, the following sequence is performed.

- Luciferin (150 mg kg^{-1} i.p.) is injected 20 min before imaging.
- Mice are anaesthetised (e.g. with Nembutal (pentobarbital; 25–50 mg kg^{-1}) or isofluorane) immediately before imaging.
- Images are captured with integration times of 1–10 min. This must be done in total darkness; the results are viewed as pseudo-colour images, where low bioluminescence shows as blue and high as red. Light generated by bioluminescence can penetrate tissues at low but detectable levels. Red light penetrates better than blue because haemoglobin absorbs blue light, and some is lost due to scattering (net losses being ~10-fold per centimetre depth of tissue).
- Black and white reference images are captured under low light illumination and the bioluminescent and reference images are superimposed.
- The animals are then kept warm and allowed to recover.

For repeated imaging (e.g. over a time course), luciferin must be injected every time.

It is important to ensure that the anaesthetic used does not adversely affect the bioluminescent signal, which can occur in two ways: some anaesthetics are direct inhibitors of luciferase activity, while their cardiovascular effects could adversely affect the biodistribution of luciferin, altering both maximal signal intensity and signal kinetics. These haemodynamic effects seem to be more of a problem than direct inhibition, and it is important to standardise

[17] http://www.perkinelmer.com/Catalog/Category/ID/In%20Vivo%20Imaging

the anaesthetic procedure in terms of dose and the time between anaesthesia and injection of the substrate (Keyaerts *et al*., 2012). The best anaesthetics, in terms of minimal effects on bioluminescence, were pentobarbital and avertin whereas isofluorane, despite being easier to use, did have an adverse effect on the bioluminescent signal.

One example of this approach is the panel of bioluminescent reporter mouse and rat lines designed to report on the regulation of human CYP3A4 (Zhang *et al*., 2003a; Zhang *et al*., 2003b; Zhang *et al*., 2004b). In order to create these models a human genomic 3A4 BAC clone was used to construct a 15 kb DNA fragment in which 13 kb of the CYP3A4 promoter was linked via a chimeric intron to a firefly luciferase reporter cassette (Zhang *et al*., 2003a; Zhang *et al*., 2003b). The construct included both the distal enhancer (−7.8 kb) and the proximal XRE (−150 bp). Disconcertingly, however, the expression of the luciferase reporter did not correlate with that of the endogenous Cyp3a in mice or rats and in terms of inducibility, the greatest extent of induction was observed with clotrimazole. This marked response to clotrimazole was unexpected since the system relies on the endogenous murine PXR; it may reflect reporter activation via other receptors such as CAR, but the lack of correlation with regulation of endogenous CYP3A means that more work is required before this line can be relied upon in drug discovery/development, and the authors recognise that the dominance of host factors in inducibility via PXR means that even humanised rats and mice may not accurately reflect the human response unless the receptors are also humanised.

Efforts in relation to detection of overt toxicity have focussed on the regulation of haem oxygenase (Malstrom *et al*., 2004; Zhang *et al*., 2001). A bioluminescent model was constructed using a 15 kb genomic fragment containing the mouse HO-1 promoter upstream of the firefly luciferase reporter cassette (Zhang *et al*., 2001). The reporter was widely expressed in mouse pups, the most intense expression being seen in the cranium, corresponding with the known expression of HO-1 in the neonatal mouse brain. High-level expression was also detected in the testes of adult male mice. Injection of cadmium chloride into the left thigh (20 µl of 20 mM) led to strong local expression of the reporter, peaking at 9 h and starting to decline by 24 h after injection. This was a local effect and was not observed in the contralateral thigh.

Example: Biophotonic imaging of the renal toxicity of chloroform

As part of a programme to evaluate the potential of biophotonics for predictive toxicology, Weir *et al*. (2005) conducted a series of studies in which HO-1.luc reporter mice were dosed with chloroform (200 mg kg^{-1}, p.o., daily for 5 days) to induce renal toxicity. Bioluminescence was clearly visible over the kidneys of chloroform-treated mice (Figure 7.7). Comparison of the *in vivo* response of the luciferase reporter with markers of toxicity measured *ex vivo* (differential gene expression of adaptive antioxidant response genes, clinical chemistry and microscopic examination) confirmed the gender-specific difference in chloroform renal toxicity in HO-1.luc transgenic mice and correlated with the expression of the endogenous HO-1 gene. These studies illustrate the capacity of biophotonics for

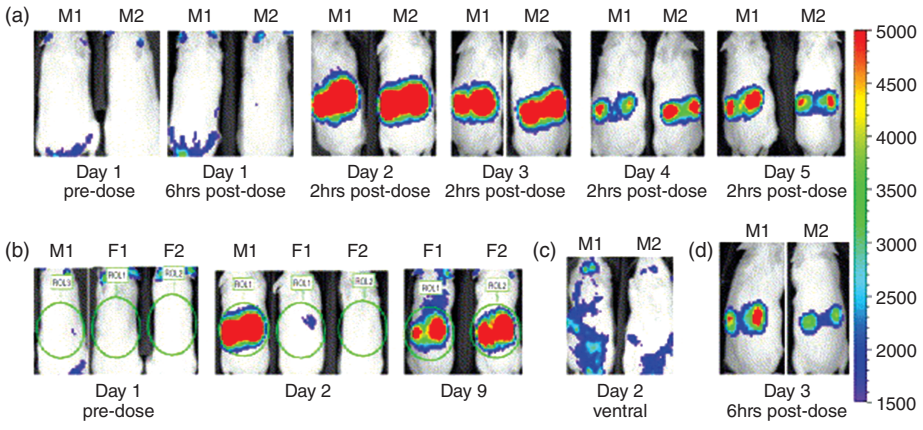


Figure 7.7 Examples of biophotonic images captured from male and female mice treated with chloroform or vehicle control. (a) Paired whole body images of chloroform-treated male animals (M1 and M2) pre-dose and following dosing on Days 1–5. (b) Paired whole body dorsal images of chloroform-treated female animals (F1 and F2) pre-dose and following dosing on Days 2 and 9. The male mouse images are included for comparison. (c) The ventral images on Day 2 are included for comparison. (d) Dorsal biophotonic images taken 6 h after dosing (source: Weir *et al.* (2005); figure 1. Reproduced with permission of Elsevier)

real-time detection of site-specific gene expression, which may be predictive of developing toxicity.

Example: Use of the HO-1.luc mouse to evaluate potential treatments for neoneatal jaundice

The HO-1.luc reporter mouse has potential applications in terms of drug discovery as well as toxicity testing. For example, it has been used to identify possible treatments to reduce HO-1-mediated bilirubin production in neonatal jaundice. The problem here is that many compounds which inhibit HO-1 also up-regulate its expression, so Zhang *et al.* (2002) compared the effects of three inhibitors (zinc bis-glycol porphyrin, tin protoporphyrin and zinc protoporphyrin) in comparison with cadmium chloride. The results showed that zinc bis-glycol porphyrin met the criteria for a therapeutic candidate since it inhibited the activity of the enzyme but did not up-regulate reporter expression *in vivo*.

Example: Ischaemia-reperfusion injury in the HO-1.luc mouse

This HO-1.luc reporter model has also been used to study the consequence of ischaemia-reperfusion injury, which is an important issue in human liver transplantation (Su *et al.*, 2006). In order to induce ischaemia, the arterial and portal venous blood supply to the left lobes of the liver was cut off using an atraumatic vascular clip. After 60 min of warm ischemia, the clip was removed, thus initiating hepatic reperfusion, and the wound was closed in layers with Prolene sutures. Bioluminescence was detectable 3 h after reperfusion and peaked at 9 h, then returned to baseline 48 h after reperfusion. The principal

source of this bioluminescence was the post-ischaemic lobe of the liver. The animal welfare benefits of using a bioluminescent approach are obvious here: instead of operating on, and subsequently killing, numerous mice the response could be measured over time in a single group of mice. In this case, 8 mice were used to examine eight time points, whereas if conventional approaches had been used this would have required 64 mice.

7.3.4 Evaluation

The bioluminescent approach has both strengths and limitations. Importantly, both in practical and 3Rs terms, it is essentially non-invasive (apart from the need to inject luciferin) and the response of an individual animal can be examined over time (rather than having to dose and sacrifice different groups of animals for every time point). This has advantages in terms of animal welfare (since fewer animals are used) and in situations where limited amounts of test compound are available (e.g. in drug discovery), although it should be noted that the time course of induction in rodents may not be the same as that in humans. Experiments in bioluminescent reporter mice can be used as a pre-screen to identify suitable doses and time points in preparation for more detailed studies. Furthermore, if animals are allowed to recover after imaging, the responses of an individual animal to different compounds can be examined; it has been shown that responses are reproducible if a 1-week interval is allowed between experiments. The use of intact, metabolically competent animals means that induction due to metabolites that are only produced *in vivo* can be detected. Finally, bioluminescence can be quantified accurately, allowing a numerical readout to be obtained.

The limitations of the bioluminescent approach include the fact that the visual readout obtained is rather vague in terms of localisation (being described as, for example, 'over the ventral upper abdomen' or 'dorsal area overlying the kidneys'), although of course it does still allow induction to be visualised in live animals, which was not previously possible. In addition, the animals have to be dosed with luciferin and anaesthetised prior to imaging.

7.4 Conclusions

The conventional view of ADME and toxicology studies in drug development is that they are part of a box-ticking exercise which has to be completed prior to clinical development. However, as better and better transgenic models become available their use is likely to extend both backwards into the process of drug discovery (studies to verify bioavailability prior to candidate selection) and forwards into the clinical phase (studies to explain clinical observations during Phase I, II and III studies). However, it should be clear from the discussion in this chapter that such models have to be characterised carefully and used judiciously. They show enormous promise, but there is no guarantee that the results of humanising an individual gene will be those predicted in advance, and even if they do correspond to expectations, the traditionalist would still consider that the problem with a humanised mouse is that it is neither a human nor a mouse!

Self-assessment questions

- Evaluate the tADMET™ PXR/CAR panel in comparison with previous transgenic models for the study of nuclear receptor function.
- What can the HRN™ mouse tell us about CYP-mediated processes in drug metabolism?
- Can reporter mouse models provide information about gene expression that could not be obtained using conventional animal models?

Background reading

Dale, J.W. and von Schantz, M. (2007) *From Genes to Genomes: Concepts and Applications of DNA Technology*. John Wiley and Sons, Ltd, Chichester, UK. ISBN: 978-0-470-01734-0

References

Argikar, U.A., Senekeo-Effenberger, K., Larson, E.E. *et al.* (2009) Studies on induction of lamotrigine metabolism in transgenic UGT1 mice. *Xenobiotica*, **39**, 826–835.

Arlt, V.M., Poirier, M.C., Sykes, S.E. *et al.* (2012) Exposure to benzo[a]pyrene of Hepatic Cytochrome P450 Reductase Null (HRN™) and P450 Reductase Conditional Null (RCN) mice: detection of benzo[a]pyrene diol epoxide-DNA adducts by immunohistochemistry and ^{32}P-postlabelling. *Toxicology Letters*, **213**, 160–166.

Arlt, V.M., Singh, R., Stiborova, M. *et al.* (2011) Effect of hepatic cytochrome P450 (P450) oxidoreductase deficiency on 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine-DNA adduct formation in P450 reductase conditional null mice. *Drug Metabolism and Disposition*, **39**, 2169–2173.

Arlt, V.M., Stiborova, M., Henderson, C.J. *et al.* (2008) Metabolic activation of benzo[a]pyrene *in vitro* by hepatic cytochrome P450 contrasts with detoxification *in vivo:* experiments with hepatic cytochrome P450 reductase null mice. *Carcinogenesis*, **29**, 656–665.

Baes, M., Gulick, T., Choi, H.S. *et al.* (1994) A new orphan member of the nuclear hormone-receptor superfamily that interacts with a subset of retinoic acid response elements. *Molecular and Cellular Biology*, **14**, 1544–1552.

Blumberg, B., Sabbagh, W., Juguilon, H. *et al.* (1998) SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Genes & Development*, **12**, 3195–3205.

Boverhof, D.R., Chamberlain, M.P., Elcombe, C.R. *et al.* (2011) Transgenic animal models in toxicology: historical perspectives and future outlook. *Toxicological Sciences*, **121**, 207–233.

Cai, H., Nguyen, N., Peterkin, V. *et al.* (2010) A humanized UGT1 mouse model expressing the UGT1A1*28 allele for assessing drug clearance by UGT1A1-dependent glucuronidation. *Drug Metabolism and Disposition*, **38**, 879–886.

Chen, S., Beaton, D., Nguyen, N. *et al.* (2005) Tissue-specific, inducible, and hormonal control of the human UDP-glucuronosyltransferase-1 (UGT1) locus. *Journal of Biological Chemistry*, **280**, 37547–37557.

Chen, S., Yueh, M.F., Evans, R.M. and Tukey, R.H. (2012) Pregnane-X-receptor controls hepatic glucuronidation during pregnancy and neonatal development in humanized UGT1 mice. *Hepatology*, **56**, 658–667.

Cheng, J., Krausz, K.W., Tanaka, N. and Gonzalez, F.J. (2012a) Chronic exposure to rifaximin causes hepatic steatosis in pregnane X receptor-humanized mice. *Toxicological Sciences*, **129**, 456–468.

Cheng, J., Ma, X. and Gonzalez, F.J. (2011) Pregnane X receptor- and CYP3A4-humanized mouse models and their applications. *British Journal of Pharmacology*, **163**, 461–468.

Cheng, J., Ma, X., Krausz, K.W. *et al*. (2009) Rifampicin-activated human pregnane X receptor and CYP3A4 induction enhance acetaminophen-induced toxicity. *Drug Metabolism and Disposition*, **37**, 1611–1621.

Cheng, J., Shah, Y.M. and Gonzalez, F.J. (2012b) Pregnane X receptor as a target for treatment of inflammatory bowel disorders. *Trends in Pharmacological Sciences*, **33**, 323–330.

Cheng, J., Shah, Y.M., Ma, X. *et al*. (2010) Therapeutic role of rifaximin in inflammatory bowel disease: clinical implication of human pregnane X receptor activation. *Journal of Pharmacology and Experimental Therapeutics*, **335**, 32–41.

Cheung, C. and Gonzalez, F.J. (2008) Humanized mouse lines and their application for prediction of human drug metabolism and toxicological risk assessment. *Journal of Pharmacology and Experimental Therapeutics*, **327**, 288–299.

Cheung, C., Yu, A.M., Chen, C.S. *et al*. (2006) Growth hormone determines sexual dimorphism of hepatic cytochrome P450 3A4 expression in transgenic mice. *Journal of Pharmacology and Experimental Therapeutics*, **316**, 1328–1334.

Chiocchetti, A., Tolosano, E., Hirsch, E. *et al*. (1997) Green fluorescent protein as a reporter of gene expression in transgenic mice. *Biochimica et Biophysica Acta*, **1352**, 193–202.

Choi, H.S., Chung, M.R., Tzameli, I. *et al*. (1997) Differential transactivation by two isoforms of the orphan nuclear hormone receptor CAR. *Journal of Biological Chemistry*, **272**, 23565–23571.

Davis, J., Maillet, M., Miano, J.M. and Molkentin, J.D. (2012) Lost in transgenesis: a user's guide for genetically manipulating the mouse in cardiac research. *Circulation Research*, **111**, 761–777.

Emter, R., Ellis, G. and Natsch, A. (2010) Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers *in vitro*. *Toxicology and Applied Pharmacology*, **245**, 281–290.

Fenner, K.S., Jones, H.M., Ullah, M. *et al*. (2012) The evolution of the OATP hepatic uptake transport protein family in DMPK sciences: from obscure liver transporters to key determinants of hepatobiliary clearance. *Xenobiotica*, **42**, 28–45.

Finn, R.D., McLaren, A.W., Carrie, D. *et al*. (2007) Conditional deletion of cytochrome P450 oxidoreductase in the liver and gastrointestinal tract: a new model for studying the functions of the P450 system. *Journal of Pharmacology and Experimental Therapeutics*, **322**, 40–47.

Fujiwara, R., Chen, S., Karin, M. and Tukey, R.H. (2012) Reduced expression of UGT1A1 in intestines of humanized UGT1 mice via inactivation of NF-kappaB leads to hyperbilirubinemia. *Gastroenterology*, **142** (109–118), e2.

Fujiwara, R., Nguyen, N., Chen, S. and Tukey, R.H. (2010) Developmental hyperbilirubinemia and CNS toxicity in mice humanized with the UDP glucuronosyltransferase 1 (UGT1) locus. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 5024–5029.

Galijatovic, A., Beaton, D., Nguyen, N. *et al*. (2004) The human CYP1A1 gene is regulated in a developmental and tissue-specific fashion in transgenic mice. *Journal of Biological Chemistry*, **279**, 23969–23976.

Gao, J., He, J., Zhai, Y. *et al.* (2009) The constitutive androstane receptor is an anti-obesity nuclear receptor that improves insulin sensitivity. *Journal of Biological Chemistry*, **284**, 25984–25992.

Ghim, C.M., Lee, S.K., Takayama, S. and Mitchell, R.J. (2010) The art of reporter proteins in science: past, present and future applications. *BMB Reports*, **43**, 451–460.

Gong, H., Singh, S.V., Singh, S.P. *et al.* (2006) Orphan nuclear receptor pregnane X receptor sensitizes oxidative stress responses in transgenic mice and cancerous cells. *Molecular Endocrinology*, **20**, 279–290.

Hadjantonakis, A.K., Macmaster, S. and Nagy, A. (2002) Embryonic stem cells and mice expressing different GFP variants for multiple non-invasive reporter usage within a single animal. *BMC Biotechnology*, **2**, 11.

Hasegawa, M., Kapelyukh, Y., Tahara, H. *et al.* (2011) Quantitative prediction of human pregnane X receptor and cytochrome P450 3A4 mediated drug-drug interaction in a novel multiple humanized mouse line. *Molecular Pharmacology*, **80**, 518–528.

Henderson, C.J., Otto, D.M.E., Carrie, D. *et al.* (2003) Inactivation of the hepatic cytochrome P450 system by conditional deletion of hepatic cytochrome P450 reductase. *Journal of Biological Chemistry*, **278**, 13480–13486.

Henderson, C.J., Pass, G.J. and Wolf, C.R. (2006) The hepatic cytochrome P450 reductase null mouse as a tool to identify a successful candidate entity. *Toxicology Letters*, **162**, 111–117.

Hendriks, G., Atallah, M., Morolli, B. *et al.* (2012) The ToxTracker assay: novel GFP reporter systems that provide mechanistic insight into the genotoxic properties of chemicals. *Toxicological Sciences*, **125**, 285–298.

Igarashi, K., Kitajima, S., Aisaki, K. *et al.* (2012) Development of humanized steroid and xenobiotic receptor mouse by homologous knock-in of the human steroid and xenobiotic receptor ligand binding domain sequence. *Journal of Toxicological Sciences*, **37**, 373–380.

Ikawa, M., Yamada, S., Nakanishi, T. and Okabe, M. (1998) 'Green mice' and their potential usage in biological research. *FEBS Letters*, **430**, 83–87.

Iusuf, D., Van De Steeg, E. and Schinkel, A.H. (2012) Functions of OATP1A and 1B transporters *in vivo:* insights from mouse models. *Trends in Pharmacological Sciences*, **33**, 100–108.

Keyaerts, M., Remory, I., Caveliers, V. *et al.* (2012) Inhibition of firefly luciferase by general anesthetics: effect on *in vitro* and *in vivo* bioluminescence imaging. *PLoS One*, **7**, e30061.

Kliewer, S.A., Moore, J.T., Wade, L. *et al.* (1998) An orphan nuclear receptor activated by pregnanes defines a novel steroid signaling pathway. *Cell*, **92**, 73–82.

Kock, K. and Brouwer, K.L. (2012) A perspective on efflux transport proteins in the liver. *Clinical Pharmacology and Therapeutics*, **92**, 599–612.

Lagas, J.S., Vlaming, M.L. and Schinkel, A.H. (2009) Pharmacokinetic assessment of multiple ATP-binding cassette transporters: the power of combination knockout mice. *Molecular Interventions*, **9**, 136–145.

Lehmann, J.M., McKee, D.D., Watson, M.A. *et al.* (1998) The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *Journal of Clinical Investigation*, **102**, 1016–1023.

Ma, X., Cheung, C., Krausz, K.W. *et al.* (2008) A double transgenic mouse model expressing human pregnane X receptor and cytochrome P450 3A4. *Drug Metabolism and Disposition*, **36**, 2506–2512.

Ma, X., Shah, Y., Cheung, C. *et al.* (2007a) The pregnane X receptor gene-humanized mouse: a model for investigating drug-drug interactions mediated by cytochromes P450 3A. *Drug Metabolism and Disposition*, **35**, 194–200.

Ma, X., Shah, Y.M., Guo, G.L. *et al.* (2007b) Rifaximin is a gut-specific human pregnane X receptor activator. *Journal of Pharmacology and Experimental Therapeutics*, **322**, 391–398.

Maggi, A., Ottobrini, L., Biserni, A. *et al.* (2004) Techniques: reporter mice - a new way to look at drug action. *Trends in Pharmacological Sciences*, **25**, 337–342.

Maglich, J.M., Parks, D.J., Moore, L.B. *et al.* (2003) Identification of a novel human constitutive androstane receptor (CAR) agonist and its use in the identification of CAR target genes. *Journal of Biological Chemistry*, **278**, 17277–17283.

Malstrom, S.E., Jekic-McMullen, D., Sambucetti, L. *et al.* (2004) *In vivo* bioluminescent monitoring of chemical toxicity using heme oxygenase-luciferase transgenic mice. *Toxicology and Applied Pharmacology*, **200**, 219–228.

Nguyen, N., Bonzo, J.A., Chen, S. *et al.* (2008) Disruption of the ugt1 locus in mice resembles human Crigler-Najjar Type I disease. *Journal of Biological Chemistry*, **283**, 7901–7911.

Pass, G.J., Carrie, D., Boylan, M. *et al.* (2005) Role of hepatic cytochrome P450s in the pharmacokinetics and toxicity of cyclophosphamide: studies with the hepatic cytochrome P450 reductase null mouse. *Cancer Research*, **65**, 4211–4217.

Pickup, K., Gavin, A., Jones, H.B. *et al.* (2012) The hepatic reductase null mouse as a model for exploring hepatic conjugation of xenobiotics: application to the metabolism of diclofenac. *Xenobiotica*, **42**, 195–205.

Robertson, G.R., Field, J., Goodwin, B. *et al.* (2003) Transgenic mouse models of human CYP3A4 gene regulation. *Molecular Pharmacology*, **64**, 42–50.

Ross, J., Plummer, S.M., Rode, A. *et al.* (2010) Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane *in vivo*. *Toxicological Sciences*, **116**, 452–466.

Rowe, J.M., Welsh, C., Pena, R.N. *et al.* (2008) Illuminating role of CYP1A1 in skin function. *Journal of Investigative Dermatology*, **128**, 1866–1868.

Scheer, N., Kapelyukh, Y., McEwan, J. *et al.* (2012) Modeling human cytochrome P450 2D6 metabolism and drug-drug interaction by a novel panel of knockout and humanized mouse lines. *Molecular Pharmacology*, **81**, 63–72.

Scheer, N., Ross, J., Kapelyukh, Y. *et al.* (2010) *In vivo* responses of the human and murine pregnane X receptor to dexamethasone in mice. *Drug Metabolism and Disposition*, **38**, 1046–1053.

Scheer, N., Ross, J., Rode, A. *et al.* (2008) A novel panel of mouse models to evaluate the role of human pregnane X receptor and constitutive androstane receptor in drug response. *Journal of Clinical Investigation*, **118**, 3228–3239.

Senekeo-Effenberger, K., Chen, S., Brace-Sinnokrak, E. *et al.* (2007) Expression of the human UGT1 locus in transgenic mice by 4-chloro-6-(2,3-xylidino)-2-pyrimidinylthioacetic acid (WY-14643) and implications on drug metabolism through peroxisome proliferator-activated receptor alpha activation. *Drug Metabolism and Disposition*, **35**, 419–427.

Stanley, L.A., Horsburgh, B.C., Ross, J. *et al.* (2006) PXR and CAR: nuclear receptors which play a pivotal role in drug disposition and chemical toxicity. *Drug Metabolism Reviews*, **38**, 515–597.

Stanley, L.A., Horsburgh, B.C., Ross, J. *et al*. (2009) Drug transporters: gatekeepers controlling access of xenobiotics to the cellular interior. *Drug Metabolism Reviews*, **41**, 27–65.

Stedman, C., Robertson, G., Coulter, S. and Liddle, C. (2004) Feed-forward regulation of bile acid detoxification by CYP3A4 - Studies in humanized transgenic mice. *Journal of Biological Chemistry*, **279**, 11336–11343.

Su, H., Van Dam, G.M., Buis, C.I. *et al*. (2006) Spatiotemporal expression of heme oxygenase-1 detected by *in vivo* bioluminescence after hepatic ischemia in HO-1/Luc mice. *Liver Transplantation*, **12**, 1634–1639.

Takahashi, E., Miyamoto, N., Kajiwara, N. *et al*. (2000) Expression analysis of *Escherichia coli lacZ* reporter gene in transgenic mice. *Brain Research. Brain Research Protocols*, **5**, 159–166.

Tzameli, I. and Moore, D.D. (2001) Role reversal: new insights from new ligands for the xenobiotic receptor CAR. *Trends in Endocrinology and Metabolism*, **12**, 7–10.

Ueda, A., Kakizaki, S., Negishi, M. and Sueyoshi, T. (2002) The residue, threonine 350 confers steroid hormone responsiveness to the mouse nuclear receptor CAR. *The FASEB Journal*, **16**, A146–A.

Uibel, F., Muhleisen, A., Kohle, C. *et al*. (2010) ReProGlo: a new stem cell-based reporter assay aimed to predict embryotoxic potential of drugs and chemicals. *Reproductive Toxicology*, **30**, 103–112.

Van De Steeg, E., Wagenaar, E., Van Der Kruijssen, C.M. *et al*. (2010) Organic anion transporting polypeptide 1a/1b-knockout mice provide insights into hepatic handling of bilirubin, bile acids, and drugs. *Journal of Clinical Investigation*, **120**, 2942–2952.

Van Herwaarden, A.E., Smit, J.W., Sparidans, R.W. *et al*. (2005) Midazolam and cyclosporin a metabolism in transgenic mice with liver-specific expression of human CYP3A4. *Drug Metabolism and Disposition*, **33**, 892–895.

Van Herwaarden, A.E., Van Waterschoot, R.A. and Schinkel, A.H. (2009) How important is intestinal cytochrome P450 3A metabolism? *Trends in Pharmacological Sciences*, **30**, 223–227.

Van Herwaarden, A.E., Wagenaar, E., Van Der Kruijssen, C.M. *et al*. (2007) Knockout of cytochrome P450 3A yields new mouse models for understanding xenobiotic metabolism. *Journal of Clinical Investigation*, **117**, 3583–3592.

Van Waterschoot, R.A., Lagas, J.S., Wagenaar, E. *et al*. (2009a) Absence of both cytochrome P450 3A and P-glycoprotein dramatically increases docetaxel oral bioavailability and risk of intestinal toxicity. *Cancer Research*, **69**, 8996–9002.

Van Waterschoot, R.A., Rooswinkel, R.W., Sparidans, R.W. *et al*. (2009b) Inhibition and stimulation of intestinal and hepatic CYP3A activity: studies in humanized CYP3A4 transgenic mice using triazolam. *Drug Metabolism and Disposition*, **37**, 2305–2313.

Van Waterschoot, R.A., Rooswinkel, R.W., Wagenaar, E. *et al*. (2009c) Intestinal cytochrome P450 3A plays an important role in the regulation of detoxifying systems in the liver. *FASEB Journal*, **23**, 224–231.

Van Waterschoot, R.A. and Schinkel, A.H. (2011) A critical analysis of the interplay between cytochrome P450 3A and P-glycoprotein: recent insights from knockout and transgenic mice. *Pharmacological Reviews*, **63**, 390–410.

Van Waterschoot, R.A., Van Herwaarden, A.E., Lagas, J.S. *et al*. (2008) Midazolam metabolism in cytochrome P450 3A knockout mice can be attributed to up-regulated CYP2C enzymes. *Molecular Pharmacology*, **73**, 1029–1036.

Vasey, D.B., Wolf, C.R., Brown, K. and Whitelaw, C.B. (2011) Spatial p21 expression profile in the mid-term mouse embryo. *Transgenic Research*, **20**, 23–28.

Vasey, D.B., Wolf, C.R., MacArtney, T. *et al*. (2008) p21-LacZ reporter mice reflect p53-dependent toxic insult. *Toxicology and Applied Pharmacology*, **227**, 440–450.

Wei, P., Zhang, J., Egan-Hafley, M. *et al*. (2000) The nuclear receptor CAR mediates specific xenobiotic induction of drug metabolism. *Nature*, **407**, 920–923.

Weir, L.R., Schenck, E., Meakin, J. *et al*. (2005) Biophotonic imaging in HO-1.luc transgenic mice: real-time demonstration of gender-specific chloroform induced renal toxicity. *Mutation Research*, **574**, 67–75.

Xie, W., Barwick, J.L., Downes, M. *et al*. (2000) Humanized xenobiotic response in mice expressing nuclear receptor SXR. *Nature*, **406**, 435–439.

Young, R., Wolf, C.R., Brown, K. *et al*. (2010) Spatial monitoring of toxicity in HMOX-LacZ transgenic mice. *Transgenic Research*, **19**, 897–902.

Yueh, M.F. and Tukey, R.H. (2007) Nrf2-Keap1 signaling pathway regulates human UGT1A1 expression *in vitro* and in transgenic UGT1 mice. *Journal of Biological Chemistry*, **282**, 8749–8758.

Yule, S.M., Price, L., Cole, M. *et al*. (2001) Cyclophosphamide metabolism in children following a 1-h and a 24-h infusion. *Cancer Chemotherapy and Pharmacology*, **47**, 222–228.

Zhang, Q.Y., Fang, C., Zhang, J. *et al*. (2009) An intestinal epithelium-specific cytochrome P450 (P450) reductase-knockout mouse model: direct evidence for a role of intestinal P450s in first-pass clearance of oral nifedipine. *Drug Metabolism and Disposition*, **37**, 651–657.

Zhang, W., Contag, P.R., Hardy, J. *et al*. (2002) Selection of potential therapeutics based on *in vivo* spatiotemporal transcription patterns of heme oxygenase-1. *Journal of Molecular Medicine (Berlin)*, **80**, 655–664.

Zhang, W., Feng, J.Q., Harris, S.E. *et al*. (2001) Rapid *in vivo* functional analysis of transgenes in mice using whole body imaging of luciferase expression. *Transgenic Research*, **10**, 423–434.

Zhang, W., Moorthy, B., Chen, M. *et al*. (2004a) A Cyp1a2-luciferase transgenic CD-1 mouse model: responses to aryl hydrocarbons similar to the humanized AhR mice. *Toxicological Sciences*, **82**, 297–307.

Zhang, W.S., Purchio, A., Chen, K. *et al*. (2003a) *In vivo* activation of the human CYP3A4 promoter in mouse liver and regulation by pregnane X receptors. *Biochemical Pharmacology*, **65**, 1889–1896.

Zhang, W.S., Purchio, A.F., Chen, K. *et al*. (2003b) A transgenic mouse model with a luciferase reporter for studying *in vivo* transcriptional regulation of the human CYP3A4 gene. *Drug Metabolism and Disposition*, **31**, 1054–1064.

Zhang, W.S., Purchio, A.F., Coffee, R. and West, D.B. (2004b) Differential regulation of the human CYP3A4 promoter in transgenic mice and rats. *Drug Metabolism and Disposition*, **32**, 163–167.

8

Genotoxicity and its Measurement

8.1 Genotoxicity testing

According to the Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment (COM) (COM, 2011):

> Genotoxicity refers to interaction with, or damage to, DNA and/or other cellular components which regulate the fidelity of the genome. It is a broad term that, as well as mutation .... includes damage to DNA such as the production of adducts, by the chemical itself or its metabolites.

A genotoxic chemical is one which interacts adversely with the genome in one or more tests (including mutagenicity assays). A mutagenic chemical is defined as 'a genotoxin that has been shown to cause a heritable change in the phenotype of a cell or organism via a stable alteration of the genome'. Genotoxicity is an endpoint in its own right; in addition, most carcinogens are genotoxic (i.e. they damage DNA directly). They are positive in mutagenicity assays *in vitro* and *in vivo*, and may cause adduct formation, induce DNA repair and act as tumour initiators.

Mutations are irreversible and there are no warning signs to warn people to avoid mutagenic compounds because the effects of exposure to relatively low concentrations may not be immediately apparent. Furthermore, because the process leading to mutation requires cell division as well as DNA damage, mutations may not be evident until many years after exposure, especially in cell types which only divide very infrequently.

Since mutations are more likely to be deleterious than beneficial, they are always considered to be associated with potential risks; indeed, they play a role in numerous pathological conditions including atherosclerotic plaques, senile cataracts and gall bladder disease as well as in aging. Furthermore, heritable

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

chromosomal alterations in somatic cells are implicated in the development of cancer and those in germinal cells may lead to inherited disorders, stillbirth and spontaneous abortion.[1]

Numerous testing systems have been developed for the identification of genotoxic chemicals. The simplest of these use microorganisms such as bacteria and yeast. Additional tests using mammalian systems include *in vitro* tests, *in vivo* tests using conventional laboratory species and more sophisticated tests using transgenic animals carrying genes which can report on the nature and frequency of mutations occurring *in vivo*.

This chapter is not intended to be a comprehensive review of genotoxicity testing methods or a step-by-step guide on how to conduct a genotoxicity testing programme; rather, the aim is to describe the standard tests briefly and then introduce some novel methods which have become possible as a result of recent developments in mammalian genetics and transgenic technology. Further information on the standard tests is provided elsewhere.[2]

8.2 Core *in vitro* tests

8.2.1 The Ames test

The gold standard method for testing the genotoxicity of chemicals is the Ames test,[3] which is used worldwide as an initial screen to determine the mutagenic potential of new chemicals and drugs. Data from this test are required for submission of data to regulatory agencies for registration or acceptance of products including drugs and biocides. The test, which was developed in the early 1970s,[4] evaluates a chemical's ability to induce mutations in *Salmonella typhimurium* by a variety of mechanisms, for example, base pair substitution, induction of frame-shifts and oxidative damage. It is the first test done in regulatory regimes unless it is considered inappropriate for specific reasons (e.g. in the case of antibacterial test substances). The Ames test is held to be the best available *in vitro* test for genotoxicity, being a relatively predictive indicator of genotoxic carcinogens (it identifies 70–90% of known genotoxic carcinogens).

The Ames test uses strains of *S. typhimurium* (and in some cases *Escherichia coli*) which carry genetic defects making them dependent upon specific nutrients in the medium on which they are grown. In the Ames test, tester strains are plated on minimal medium and exposed to the chemical under investigation. If the chemical is mutagenic, some of the cells acquire mutations which allow them to grow on minimal medium and are therefore able to develop into colonies. By using different tester strains, each of which identifies a particular type of mutation, it is possible to identify the type of adduct formed by the chemical being tested.

---

[1] For a discussion of the human health significance of mutagenicity, see paper MUT/2012/14 on the COM web site, http://www.iacom.org.uk/papers/.

[2] For comprehensive *in vitro* and *in vivo* protocols, see Parry and Parry (2012).

[3] See http://ntp.niehs.nih.gov/?objectid=16D6D969-A9F4-0B6F-2B81F2FB53BBFD3E

[4] For a review describing how the Ames test was developed and providing detailed instructions for performing and interpreting the test, see Mortelmans and Zeiger (2000).

Many genotoxic compounds require metabolic activation, because bacterial indicator cells lack many of the key enzymes involved in this process (especially CYPs), bacterial mutagenicity tests such as the Ames test are almost invariably run in the presence of hepatic S9 mix. An S9 mix is the supernatant which remains after centrifugation of a tissue homogenate at 9000 g and contains both microsomal and cytosolic xenobiotic metabolising enzymes. Uninduced S9 is of limited usefulness because it lacks many of the key enzymes involved in metabolic activation, and the most common type of S9 used is made using liver from rats treated with Aroclor 1254, a mixture of polychlorinated biphenyls which potently induces the expression of several families of CYPs and other xenobiotic metabolising enzymes. However, concerns about the environmental effects of this material have led to the adoption of alternatives such as a mixture of β-naphthoflavone and PB. The assay may also be customised by using tissue-specific S9 mixes (e.g. from lung).

8.2.2 *In vitro* gene mutation tests using mammalian cells

The Ames test is the best measure available for induction of mutations, but it does not measure aneugenic or clastogenic effects, or the effects of chemicals on mammalian cell replication. Follow-up testing using mammalian systems, either by *in vitro* or *in vivo*, is therefore usually required.[5]

One of the most popular mammalian cell assays is the so-called mouse lymphoma assay, which is recommended by the US FDA and EPA. European bodies and particularly the ICH also drove the acceptance of this test as a single screen. This uses the mouse lymphoma L5178Y ($tk^{+/-}$) cell line to detect changes at the thymidine kinase (*tk*) locus due to base changes, frameshift mutations and small deletions. The mouse lymphoma assay has the potential to detect both mutagenic and clastogenic events at the thymidine kinase locus of $tk^{+/-}$ L5178Y mouse lymphoma cells by measuring resistance to the lethal nucleoside analogue trifluorothymidine. When trifluorothymidine is added to the culture medium, the $tk^{+/-}$ cells die, but $tk^{-/-}$ cells are unable to incorporate trifluorothymidine into their DNA and so they are resistant to its genotoxic effects. The $tk^{+/-}$ assay can detect various types of genetic damage involving chromosome 11, the location of the *tk* gene, including point mutations and chromosome aberrations (Ogawa *et al.*, 2009). Depending on the type of DNA damage sustained, the resulting resistant colonies may be large (indicative of gene mutations) or small (indicative of chromosomal aberrations). The mouse lymphoma assay is usually used as part of a test battery for regulatory submissions under OECD and ICG guidelines and is offered commercially by a number of contract research organisations.[6]

Another mammalian cell mutation assay is the HPRT assay. The HPRT gene is located on the X chromosome in mammalian cells and is used as a model gene to investigate mutational effects in mammalian cell lines such as the V79 Chinese hamster ovary cell line. The methodology used is similar to that of the mouse

[5] The subject of *in vitro* genotoxicity testing was subjected to systematic review in 2009: Muller (2009).
[6] A typical test protocol may be found at http://www.genpharmtox.de/downloads/AssaySheetMOLY.pdf

lymphoma assay.[7] It is based on the principle that cells with an intact HPRT gene are susceptible to the toxic effects of 6-thioguanine, whereas mutant cells are resistant to this analogue of the naturally occurring purine bases hypoxanthine and guanine. The HPRT assay focuses on small changes such as point mutations and small deletions because large deletions in the X chromosome are lethal; this also means that the background mutation frequency detected by the assay is lower than that of the mouse lymphoma assay because the majority of spontaneous mutations are lethal and are therefore not measured in the assay. A major drawback of this test, however, is that too few cells are generally exposed and then selected for mutation frequency so it tends to lack statistical power.

In order to verify that observations made in rodent systems are relevant to humans, *in vitro* tests using human peripheral blood lymphocytes and cultured human cell lines are used. It is possible to use other cell lines, such as MCL-5, a human lymphoblastoid cell line which has been stably transfected with four human P450s (CYP2A1, CYP2E1, CYP3A4 and CYP2A6). This cell line also expresses CYP1A1 and microsomal epoxide hydrolase, allowing it to be used to identify mutagens which require metabolic activation in order to exert their effects (Crespi *et al*., 1991)); however, it is very slow growing and this limits its utility in tests such as the HPRT assay, in which achievement of a suitable rate of cell division is critical to the generation of interpretable results.

Technical considerations for mammalian cell-based tests For all the core *in vitro* mammalian tests, the following need to be stressed:

- Selection of an appropriate cell line is essential (Fowler *et al*., 2012a). In particular, it is essential to ensure that the rate of cell division is adequate in cells during the treatment and assessment periods, and that cells at all stages of the cell cycle are present during the exposure period.
- Sufficient time must be allowed between the end of exposure and measurement for fixation of incident damage, whether as mutations, chromosome aberrations or micronuclei, but the assessment should not be too long after the end of exposure and measurement or the damaged cells will disappear.
- Adequate levels of cytotoxicity must be ensured (Fowler *et al*., 2012b); insufficient and excessive levels are both undesirable. The relevant regulatory guidelines specify the levels to aim for, and suitable concurrent methods of assessing cytotoxicity should be incorporated into each assay.

The results of all the regulatory tests should be compared with historical control values to assess the validity of the test and then to assess the biological relevance of any positive result.

8.2.3 The *in vitro* chromosome aberration test

Chromosomal mutations may occur in somatic cells (leading to neoplasia) or in germ cells (leading to birth defects and developmental abnormalities). Types of chromosomal mutation include the following:

[7] A typical test protocol may be found at http://www.genpharmtox.de/downloads/AssaySheetsHPRT.pdf.

- Structural changes: Stable aberrations such as translocations and inversions and unstable aberrations including gaps, breaks, dicentric chromosomes, double minutes and rings.
- Numerical changes: Aneuploidy (loss or gain of chromosomes) and polyploidy (gain of multiples of the diploid chromosome complement).

In the mammalian chromosome aberration test, mammalian cells are cultured *in vitro*, exposed to the test compound for a known amount of time and then harvested. Microscope slides are prepared and the frequency of asymmetrical structural chromosome aberrations is determined. The cell types used include rodent cell lines as well as human peripheral blood lymphocytes, which do not normally divide *in vitro* but can be stimulated to do so by the addition of a mitogen such as phytohaemagglutinin. The key to success with this assay is careful technique, and good cell growth during the assay is critical to a successful outcome. High quality metaphase preparations, use of a trained and experienced observer who is familiar with the karyotype of the cells being used, robust acceptance criteria and careful evaluation of the data all help to generate an accurate test result. Special care is needed when using transformed cell lines as these may have unstable karyotypes, so experience is needed in order to distinguish between chemically induced damage and natural variation.

A flow chart of the process is shown in Figure 8.1. The stage of the cycle when cells are exposed affects the sensitivity to mutagens, and so exposure to the test

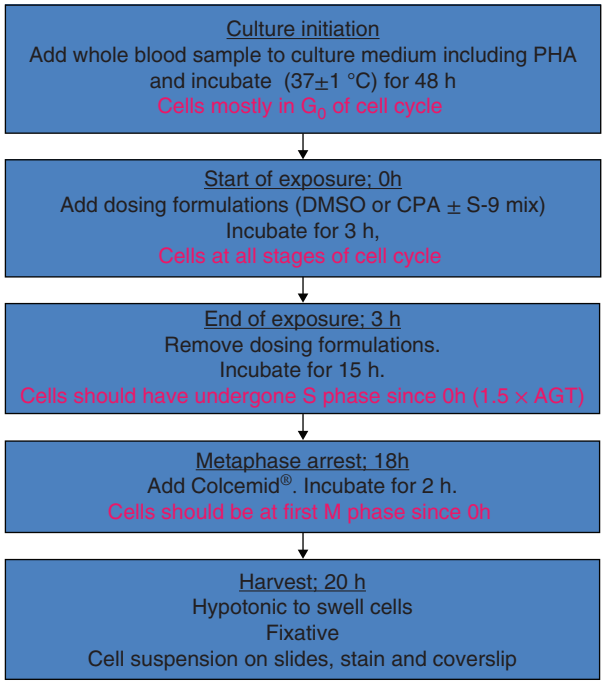


Figure 8.1 Timings of the *in vitro* human lymphocyte chromosome aberration test (source: Clare (2012); figure 1, p.73. Reproduced with kind permission from Springer Science and Business Media)

substance commences approximately 48 h after culture initiation, when cells are dividing and at all stages of the cell cycle, and the initial synchrony of the cells is less. After exposure to most mutagens, cells need to undergo S phase for chromosome damage to manifest. Since many of the asymmetrical structural chromosome aberrations prevent unlimited cell division, it is important to allow long enough for damaged cells to appear but not long enough for the cells to be lost (i.e. in the first metaphase after dosing which occurs halfway through the second cell cycle after exposure, usually 12–18 h after dosing).

A typical example of the type of cytogenetic damage seen by light microscopy following Giemsa staining is illustrated in Figure 8.2. In the regulatory chromosome aberration test, 100 metaphases from each culture are analysed for chromosome aberrations, 50 per slide, making 200 per concentration of test substance. The exception is where 10 cells with structural aberrations (excluding gaps) have been noted on a slide, when analysis may be terminated. Only cells with 44–46 chromosomes are considered acceptable for analysis; cells with more than 46 chromosomes (that is, polyploid, hyperdiploid, or endoreduplicated cells) are recorded separately and may be quantified if required.

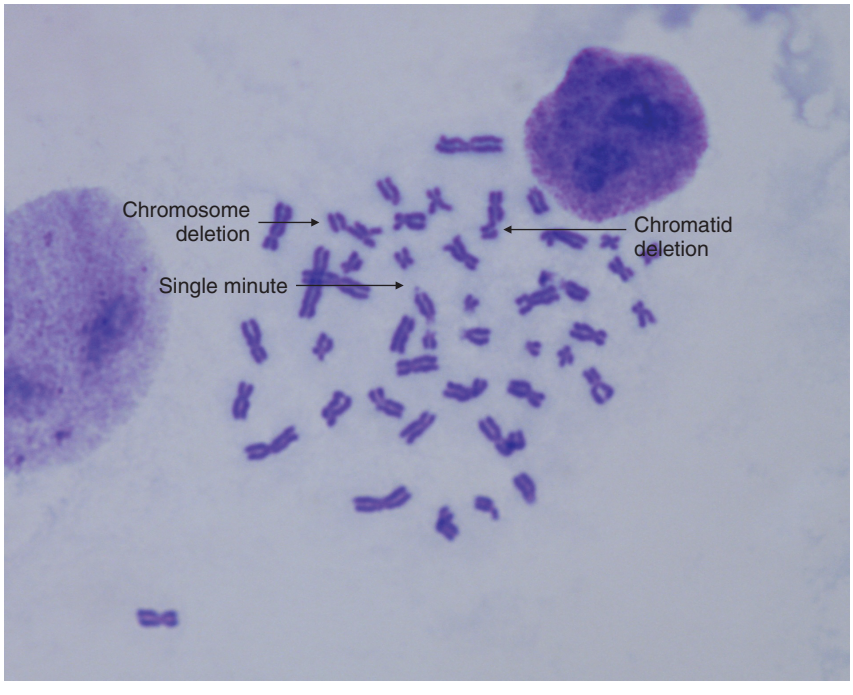


Figure 8.2 Results of the chromosome aberration test. Photograph of human lymphocyte chromosomes from cell exposed to CPA at 12.5 μg ml^{-1}, stained with Giemsa (1000× magnification) (source: Clare (2012); figure 5, p.82. Reproduced with kind permission from Springer Science and Business Media)

After completion of analysis and decoding, the proportions of aberrant cells per culture (often expressed as a percentage, since 100 cells are routinely analysed) are tabulated as:

- Cells with structural aberrations including gaps.
- Cells with structural aberrations excluding gaps. This information is used to draw a conclusion as to the clastogenic potential of a test substance as gaps may occur by non-genotoxic modes of action.
- Polyploid, endoreduplicated or hyperdiploid cells. An increase in polyploid cells may indicate that a chemical has the potential to induce numerical aberrations, when use of the *in vitro* micronucleus test should be considered.

It is important to establish that the effects observed are a genuine consequence of exposure to the test compound, rather than being due to random variation. The following evaluation criteria help in the effort to avoid ascribing small, random, statistical increases as a positive response, particularly when compared with zero values in the negative control cultures:

1. The proportion of cells with structural aberrations at one or more concentrations exceeds the concurrent vehicle control values and the historical negative control (normal) range. If no historical control range is available, it is necessary to increase the number of concurrent negative control cultures.
2. A statistically significant ($p \leq 0.05$) increase in the proportion of cells with structural aberrations (excluding gaps) is observed.

Evidence of a concentration-related trend in the proportion of cells with structural aberrations (excluding gaps) tends to support the conclusion but is not considered essential in the evaluation of a positive result.

8.2.4 The *in vitro* micronucleus assay

Micronuclei are small cytoplasmic inclusion bodies containing genomic DNA in the form of entire or partial chromosomes. A micronucleus arises when a fragment of DNA or a whole chromosome remains detached from the mitotic spindle during chromosomal separation, either because it lacks a centromere or because the centromere does not attach properly, and therefore gets separated from the rest of the chromosomes when they migrate to the cellular poles. When the daughter nuclei form, this fragment becomes surrounded by its own membrane, thus forming a micronucleus.

The actin inhibitor cytochalasin B can be included in the culture medium to block cytokinesis immediately after the nucleus has divided in order to focus the assay on cells that have divided during or since exposure to the test compound; in other words those in which any damage may reasonably be assumed to have occurred as a result of exposure. The biological process underlying the micronucleus assay is illustrated in Figure 8.3. In the presence of cytochalasin B, the cells

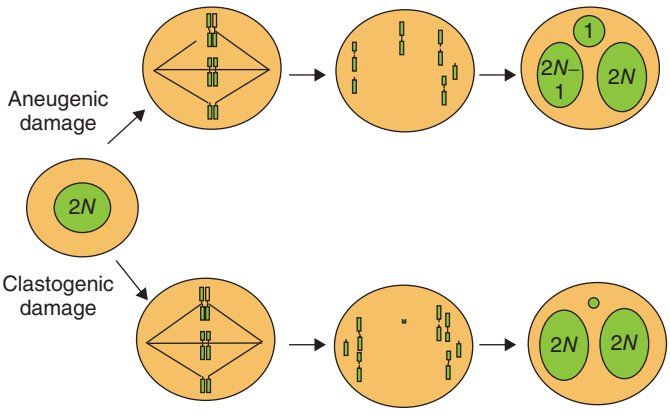


Figure 8.3 The micronucleus assay with cytochalasin B-induced cell cycle arrest (source: Dr Jerry Styles. Reproduced with permission of Dr Jerry Styles)

become arrested immediately after nuclear division resulting in the formation of a binucleated cell, and by scoring the number of binucleated cells containing micronuclei as a percentage of total binucleated cells it is possible to identify the extent of chromosomal damage which has occurred in response to chemical exposure. Micronuclei containing entire chromosomes can be distinguished from those containing acentric fragments by fluorescent *in situ* hybridisation using commercially available fluorescent probes, making it possible to distinguish between anuegenic effects (where the micronuclei contain entire chromosomes) and clastogenic effects (where they contain acentric fragments).

The *in vitro* micronucleus assay lends itself to automation by the use of flow cytometry for scoring (Avlasevich *et al.*, 2011; Dertinger *et al.*, 2011c). The reported advantages of flow cytometry include increased objectivity, leading to reduced inter- and intra-laboratory variation, scope for scoring a larger number of cells (5000 cells in 2–3 min compared with 1000–2000 cells for conventional methods) increased efficiency and faster throughput including scope for adaptation to 96-well format. The problems associated with flow cytometry for scoring micronuclei, however, include the need to lyse cells in order to detect micronuclei by flow cytometry, the fact that other types of subcellular fragments and microorganisms can be misidentified as micronuclei and the inability of flow cytometry analysis to relate micronuclei to individual cells or cell types. In addition, samples cannot be stored for verification/repeat analysis.

An alternative to flow cytometry for micronucleus scoring is the use of laser scanning cytometry, a technique which combines the strengths of flow cytometry and image analysis. In this technique (Darzynkiewicz *et al.*, 2011), cells grown on slides or multiwell plates are analysed *in situ* (suspension cells have to be applied to a slide by cytocentrifugation prior to analysis). The slides are subjected to laser scanning and can also be viewed under white light for visual examination. Images can be downloaded and the slides themselves can be stored for follow-up evaluation. The throughput of this method is intermediate between those of purely visual examination and flow cytometry (~1000 cells in 3–5 min) and it has the advantage over flow cytometry that visual characteristics of the cells (e.g. presence

of centromeres/telomeres, nuclear proteins and chromosome-specific markers) can be examined. Its limitation is that complete separation between adjacent cells is required in order for the image analysis software to detect them as separate objects.

A dual-dye sequential staining procedure has been developed by Avlasevich *et al*. (2011). In this method, ethidium bromide monoamide (which binds covalently to DNA following photoactivation) is used to identify dead and dying cells while the pan-DNA fluorescent stain SYTOX® Green is used to detect the micronuclei. This allows differential staining of healthy cells ($EMA^-/SYTOX^+$) and necrotic/apoptotic cells ($EMA^+/SYTOX^+$) so that interference from dead and dying cells can be eliminated from the analysis. This assay has the advantage that it can provide information about cytotoxicity as well as micronucleus counts, but it requires further development in order to improve its compatibility with primary cells and ability to discriminate between *bona fide* micronuclei and fragments produced during apoptosis.

The *in vitro* micronucleus assay is now generally considered to be an appropriate replacement for other chromosome aberration assays and is the subject of an OECD Test Guideline (#487). It has been subjected to retrospective validation by ECVAM (Corvi *et al*., 2008), is now accepted at EU level and has been incorporated into the requirements of the REACH regulations. It has been claimed to have numerous advantages over other assays, including

- the ability to detect both clastogens and aneugens, including compounds which cause mitotic delay, apoptosis, chromosome breakage, chromosome loss and non-disjunction;
- robustness, accuracy and good statistical power: thousands of cells can be scored relatively quickly whereas in other chromosomal aberration assays only a few hundred cells are usually scored.

Its main disadvantage is that it does not provide any information regarding the type of structural chromosomal damage which has occurred.

Variants of the core micronucleus assay

Application of the in vitro micronucleus assay to human* in vivo *exposure One of the advantages of the *in vitro* micronucleus assay is that it can be used to characterise aneugenic and clastogenic effects in human cells, and human lymphocytes have, for many years, been used for the *in vitro* micronucleus assay by exposing them to test material after culturing in the presence of PHA for 44–48 h. It can also be used for monitoring genetic damage in humans[8] by analysing various other cell types, including erythrocytes and exfoliated epithelial cells from exposed and unexposed individuals, and there is evidence that the baseline frequency of micronuclei in cancer patients, even before chemo- or radiotherapy, is significantly elevated compared with that in healthy controls. Furthermore, the frequency of micronuclei in peripheral blood mononuclear cells from healthy individuals is predictive of a risk of developing cancer and it has been suggested that micronucleus formation is associated with events in the early stages of

[8] For a review, see Speit, Zeller, and Neuss (2011).

carcinogenesis and could be used as a biomarker of risk (Bonassi *et al*., 2011b); however, it would not be appropriate to implement the use of such a biomarker unless a clear benefit can be demonstrated (i.e. appropriate management strategies would have to be in place for individuals identified as being at risk).

The International Human Micronucleus Project (HUMN and $HUMN_{XL}$, which focuses on the use of buccal mucosal cells) is a collaborative effort with the aim of co-ordinating research on the application of the micronucleus assay to humans, especially with respect to intra/interlaboratory variation and the effects of age, gender and smoking status (Fenech *et al*., 2011; Bonassi *et al*., 2011a). It has concluded that the human buccal mucosal cell assay offers minimal invasiveness with the potential for population monitoring subject to the development of a standardised protocol, and also aims to develop micronucleus assays in other accessible human cell types (erythrocytes, nasal epithelium and hair root cells).

Mouse bone marrow and peripheral blood micronucleus assays Rodent micronucleus assays detect clastogenicity by measuring micronuclei formed from acentric chromosome fragments in immature (polychromatic) erythrocytes in the bone marrow or in reticulocytes in peripheral blood. The most commonly conducted *in vivo* assay for clastogens is the rodent erythrocyte micronucleus assay, which is relatively simple and sensitive and is now the subject of an OECD Test Guideline (#474).

In the *in vivo* micronucleus assay, animals are usually given a single dose of test material and sampled 24–48 h for bone marrow and 36–72 h later for peripheral blood effects. The alternative is to administer multiple doses with sampling 18–24 h after the last dose for bone marrow and 36–48 h later for peripheral blood. It is necessary to score at least 2000 polychromatic erythrocytes per animal and to determine the proportion of polychromatic erythrocytes among total erythrocytes in order to assess cytotoxicity.

This assay detects chromosome breakage and whole chromosome loss in polychromatic erythrocytes and subsequently in normochromic erythrocytes in peripheral blood. Micronuclei are seen as small nucleus-like inclusions in erythrocytes from which the main nucleus has been expelled and are the only DNA-containing structures in immature red blood cells, making them easy to identify. Small blood samples can be used for scoring, meaning that repeated sampling from a single animal is possible and, for example, pre- and post-treatment micronucleus counts can be taken in order to compare spontaneous and induced micronucleus scores. The preferred dye for micronucleus staining is Acridine Orange, which stains both RNA (red) and DNA (green) and is simple and convenient for routine scoring.

The advantages of the mammalian peripheral blood micronucleus assay, as compared with metaphase analysis, are that the endpoint is simple and there is scope for automated scoring by image analysis and flow cytometry (Heddle *et al*., 2011). It can, in principle, detect both clastogens and aneugens and is highly sensitive because of the low spontaneous frequency of micronucleus formation and the large number of cells which are available for sampling. In addition, it can readily be incorporated into a standard repeat dose toxicity assay in place of bone marrow smears. There is a large database of historical information on this assay and it is widely accepted by regulatory authorities.

Micronucleus assays can, in theory, be carried out in any cell type which undergoes continuous cell division. Tissues with a low proliferative index (e.g. liver) can be used if replication is stimulated; in the case of liver this may be achieved by two-thirds partial hepatectomy or treatment with a growth-stimulating chemical (to induce restorative hyperplasia).

Tissues in which the micronucleus assay has been performed for research purposes include colon, skin and young rodent liver (which continue to divide until about 5–6 weeks after birth) (Morita, MacGregor, and Hayashi, 2011). Male germ cells can also be used to provide information about potential reproductive effects. These assays have not, however, gained recognition for regulatory purposes because they have not yet been fully characterised. In particular, characterisation of cell turnover rates is required in order to ensure that an adequate delay between final treatment and sampling is chosen.

8.3 Assessment of genotoxicity for regulatory purposes

Various recommendations regarding choice of mutagenicity assay have been issued by regulatory and other bodies. Most of them involve a tiered approach in which Stage I (Basic Testing) involves the use of a bacterial assay (usually the Ames test) and an *in vitro* mammalian test (preferably the *in vitro* micronucleus assay) to detect the potential to cause mutations; putative positives then go on to Stage II (Follow-up Testing), which may entail further testing *in vitro* or *in vivo* (Pfuhler *et al*., 2007).

The UK position on genotoxicity testing is defined by the COM, which has stated that 'information on the three levels of mutation, namely gene, clastogenicity (i.e. structural chromosome aberrations) and aneuploidy (i.e. numerical chromosome aberrations) is necessary to provide comprehensive coverage of the mutagenic potential of a chemical'. Its position is that no single validated assay can provide information on all three types of damage, and guidance issued in 2011 (COM, 2011) recommends a staged approach to testing in which the following take place:

- **Stage 0** involves preliminary consideration of the chemical structure in question (e.g. physicochemical properties, structure-activity relationships and *in silico* predictions). This stage can include HTS tests where appropriate.
- **Stage 1** involves *in vitro* testing using the Ames test and the *in vitro* micronucleus test. A clear positive in either test is sufficient to identify a compound as an *in vitro* mutagen, while a clear negative in both assays justifies the conclusion that the chemical lacks mutagenic activity *in vitro*. However, follow-up investigations may be required if tumours are found *in vivo*, the metabolic activation system used *in vitro* is inappropriate of human-specific metabolites which are identified. Additional information from testing in RHE models is also useful where extensive dermal exposure is expected and/or *in vivo* testing is banned (e.g. for cosmetics). It is believed that no relevant rodent carcinogens or *in vitro* genotoxins would remain undetected in an *in vitro* battery comprising the Ames test and the *in vitro* micronucleus test (Kirkland *et al*., 2011).

- **Stage 2** involves *in vivo* testing to investigate one or more of the following:
  - The endpoints identified in Stage 1
  - Genotoxicity in site of contact and/or target tissues
  - Potential for germ cell toxicity
  - *In vivo* mutagenicity of chemicals which were negative in Stage 1 but where there is high/moderate and/or prolonged exposure

In vivo tests can be used to address particular issues, for example, for potential site of contact mutagens or to investigate the results of carcinogenicity assays in particular tissues. However, the COM considers that animals should only be used when there is no suitable alternative and that a negative result in a single carefully selected and conducted *in vivo* assay is usually sufficient to address a positive result found *in vitro*. The aim should be to clarify the *in vivo* genotoxicity of compounds which are positive in Stage 1 and for which there is a need to ascertain whether genotoxicity is expressed *in vivo*. The minimum number of animals consistent with obtaining valid results should be used (with historical controls being considered acceptable in order to minimise animal usage).

8.4 Novel *in vitro* methods

8.4.1 GreenScreen HC

The proprietary GADD45a–GFP GreenScreen HC assay (provided by Gentronix[9]) monitors genotoxin-induced transcription of the GADD45a gene using a GFP-coupled reporter gene. GADD45a is a growth arrest and DNA-damage-inducible gene whose transcript levels are increased (3–14-fold) following stressful growth arrest conditions and treatment with DNA-damaging agents (Walmsley, 2008). Genetic, biochemical and genomic approaches have implicated the GADD45a gene in responses to genotoxic damage, and GADD45a null mice have increased susceptibility to tumour induction by ionising radiation and genotoxic chemicals. The nature of this assay, which measures transcriptional induction in response to genotoxicity rather than genotoxicity *per se*, means that it records cellular responses to genomic damage as an end point rather that identifying a particular type of DNA damage or failure of DNA repair.

The GFP reporter used in the GreenScreen HC assay is encoded by a stable replicating plasmid which includes all the essential *cis*-acting elements of the human GADD45a gene coupled to the GFP coding sequence. The GADD45a gene is only inducible in the presence of a functional p53 gene, so in contrast with several core *in vitro* tests which use apoptosis incompetent cell lines, the assay uses the *TK*6 p53-competent human lymphoblastoid cell line, possibly enhancing the relevance of the observations made.

This assay is reported to detect a wide range of genotoxins including point mutagens, clastogens and aneugens. The cells need to pass through a complete cell cycle in order to detect all possible classes of genotoxic damage, so the exposure time used is usually about 20 h, while DNA damage can delay the cell cycle so data

[9] http://www.gentronix.co.uk/

are usually collected at $t = 24–48$ h. Over this time period, exposure of the cells to genotoxins up-regulates reporter expression and causes the cells to fluoresce green. The resulting fluorescence is normalised to the cell density within the culture, making it possible to distinguish between cultures containing a large number of weakly positive cells and a small number which are fluorescing intensely, as well as allowing cytotoxic effects to be identified.

The GreenScreen HC assay can be compromised by interference due to strongly coloured and/or fluorescent reaction components which absorb or emit around the maxima of GFP (absorption 488 nm/emission 520 nm). This is not usually an issue with pharmaceutical test compounds, but it can be problematic when assaying compounds which require metabolic activation since S9 mix itself absorbs at these wavelengths. One way round this is to carry the assay out in the presence of S9 and then using a flow cytometry step to separate out the cells and measure their fluorescence in the absence of interference.

The assay requires only small quantities of test compound and was originally established in a 96-well format, making it a medium throughput, although a recent improved version (BlueScreen, a luminescent assays using luciferase from the marine crustacean *Gaussia princeps*) works in 384-well plates. It has not yet achieved regulatory recognition but it meets the ECVAM criteria for regulatory acceptance (a clearly defined endpoint, robustness and transferability) and is currently considered to be of potential value as part of a weight of evidence approach (Lynch *et al*., 2011b). It compares favourably with the Ames test, having sensitivity of 84% (22/26 known positives, compared with 75%) and a specificity of 98% (54/55 known negatives, compared with 60%) (Walmsley, 2008).

8.4.2 The Reconstructed Skin MicroNucleus assay

The conventional way to resolve questions raised by the high false positive rate observed *in vitro* genotoxicity tests has been to follow up putative positives with one or more of the *in vivo* tests described previously. Since the implementation of the 7th Amendment ban on genotoxicity testing of cosmetic ingredients in 2009, however, this has no longer been possible and this has caused an active debate as to the use of *in vitro* tests in conducting safety assessments of substances such as hair dye ingredients.

The issue with oxidative hair dyes is that they are designed to undergo a chemical reaction on/near the skin when the ingredients are mixed. This means that an uncharacterised mixture of potentially genotoxic products may be generated. However, the 7th Amendment bans all animal testing of cosmetic ingredients for genotoxicity despite the fact that conventional *in vitro* tests alone do not provide an appropriate assessment of hazard since it has been demonstrated that the standard initial battery of *in vitro* assays has a high level (up to 95%) of false positive results (Kirkland *et al*., 2005).

The components of a typical oxidative hair dye are a primary starting material (e.g. PPD), couplers/modifiers (e.g. *m*-phenylenediamine), oxidising agents (e.g. H_2O_2) and alkalinising agents (e.g. ammonia). Unfortunately, none of the common *in vitro* genotoxicity tests (the Ames test, chromosomal aberration tests and mouse lymphoma assay) are effective at discriminating between carcinogenic and non-carcinogenic aromatic amines since all aromatic amines tend to be positive

in *in vitro* genotoxicity tests, especially in the presence of a metabolic activation system. This problem is exacerbated by the fact that some aromatic amines are detoxified by N-acetylation whereas others undergo metabolic activation by this family of enzymes.

The 7th Amendment ban on the marketing of cosmetics containing ingredients that have been tested on animals meant that the conventional approach to determining whether *in vitro* mutagens exhibit mutagenic potential *in vivo* by means of bioassays is no longer available in the case of hair dye ingredients. This has led to a considerable degree on concern in the hair dye industry and elsewhere.[10] Indeed, the Scientific Committee on Consumer Products acknowledged in 2009[11] that

> ....after 11 March 2009, in many cases, it will not be possible to evaluate the mutagenic potential of cosmetic ingredients on a sound scientific basis. Because the potential mutagenicity of these ingredients is of major concern, an important part of the toxicological evaluation of cosmetic ingredients cannot be accomplished.

Since positive results *in vitro* are typically followed up by *in vivo* genotoxicity assays and these are no longer possible for cosmetic ingredients, there has been a focused effort to develop improved *in vitro* methods especially for evaluation of cosmetic ingredients. One such assay is the reconstructed skin micronucleus (RSMN) assay, a novel micronucleus assay using EpiDerm™ RHE constructs. The RSMN assay in EpiDerm™ tissues has been shown to discriminate correctly between direct-acting genotoxins and non-genotoxins, and the assay has been the subject of an international collaborative validation effort, supported by Cosmetics Europe. It is considered to be a relevant system for the study of cutaneous exposure to potential genotoxins and is recommended as a follow-up test following the identification of *in vitro* genotoxicity positives (Kirsch-Volders *et al*., 2011). Indeed, if validated successfully it could become a Stage I assay.

The advantages of using recombinant skin models for genotoxicity testing in this context include their relevant human metabolic capacity, the presence of a functional stratum corneum (and hence relevant target cell exposure) and the fact that they exhibit more normal DNA repair and cell cycle control as a result of having been prepared from primary human cells (Aardema *et al*., 2010). Importantly, because these tissue models are composed of normal human cells, they allow more physiologically relevant exposure conditions and more relevant metabolism versus exogenous liver S9 added to current *in vitro* assays.

8.5 Novel *in vivo* assays for gene mutations

8.5.1 The Pig-A assay

A relatively new method for determining mutation frequencies and rates uses an endogenous gene, *Pig-A*, whose structure, function and localisation of expression are highly conserved between rodents (*Pig-a*) and humans (*PIG-A*)

[10] http://forums.alttox.org/index.php?topic=297.0

[11] http://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_s_08.pdf

(Peruzzi *et al*., 2010). The background frequency of *PIG-A* mutations in humans is about 1×10^{-5} and *PIG-A* was originally identified as the gene involved in the rare blood disease paroxysmal nocturnal haemoglobinuria. It is an X-linked housekeeping gene encoding an enzyme subunit which is involved in the biosynthesis of the glycosyl phosphatidyl inositol anchor. This tethers extrinsic membrane proteins to the plasmalemma, and although other autosomal genes which encode proteins are involved in anchoring it is considered reasonable to assume that the loss of ability to anchor glycosyl phosphatidyl inositol linked proteins to the plasmalemma is primarily due to mutations in *Pig-A* (Dertinger and Heflich, 2011). The presence or absence of Pig-A protein can readily be detected by flow cytometry using antibodies which recognise glycosyl phosphatidyl inositol anchored proteins. Any mutation which causes loss of Pig-A activity (which can range from single base substitutions and frameshift mutations to occasional large deletions) can be detected in terms of the loss of these proteins from blood cells (the granulocyte population being the easiest to analyse). The Pig-A assay discriminates between erythrocytes in which the *Pig-A* gene has been mutated and wild-type cells by immunofluorescent labelling of the cell surface marker CD-59: wild-type cells bind the label while mutant cells, because they lack the anchor protein, remain unlabelled.

In rodents, the *Pig-a* mutation rate is highly responsive to treatment with strong mutagens, while the ability to carry out *in vivo* sampling before and after dosing and the accumulation of mutant erythrocytes over time meant that the Pig-A assay can be used to monitor the genotoxicity of weak mutagens with repeat dosing (e.g. during a 28-day study).

The Pig-a assay still requires substantial method development and interlaboratory validation, but it is relatively quick, inexpensive, requires only small samples for analysis and can be carried out using common laboratory strains. A proprietary kit has recently become available.[12] With further development, this assay could readily be integrated into regulatory tests and it has been identified as a high priority assay for further development (Schuler *et al*., 2011). It has been the subject of an international collaborative trial aimed at establishing its reproducibility, scope for incorporation into subchronic toxicity studies and comparability with core assays (Dertinger and Heflich, 2011; Dertinger *et al*., 2010; Dertinger *et al*., 2011a; Dertinger *et al*., 2011b; Lynch *et al*., 2011a). The assay can be used in both mice and rats, although the kinetics and persistence of the response differs between species (Bhalli *et al*., 2011). However, as currently configured it does not provide any more information than the *in vivo* micronucleus assay; its utility would be enhanced if it could be made applicable to cell types other than the erythrocyte.

8.5.2 *In vivo* assays using transgenic mouse models

Transgenic rodent-based mutation assays are based on the use of transgenic mice (and rats) which contain multiple copies of chromosomally integrated plasmid/phage shuttle vectors harbouring reporter genes for the detection of mutations.[13] Their development began with two models, the Muta™Mouse

[12] http://litronlabs.com/in_vivo_mutation.html

[13] For a detailed review of this topic, see Boverhof *et al*. (2011)

(which carries the bacterial *lacZ* gene) and the Big Blue® (carrying the bacterial *lac*I gene) (Nohmi, Suzuki, and Masumura, 2000).[14] Each of these lines carries a stably integrated reporter gene which is integrated into the mouse chromosome as part of a λ shuttle vector and is present in every cell. They allow mutations induced in a genetically neutral transgene to be used to study the mutagenic effects of chemicals in any tissue, and are particularly useful for examining the site of contact effects (e.g. in the skin following topical application, in the stomach following oral dosing and in the lung following inhalation exposure) (Dean *et al*., 1999). Transgenic models can also be used to examine potential mutagenic activation by examining effects in the liver and to address mutagenicity in target tissues, if known.

In transgenic mouse mutagenicity assays, transgenic mice carrying λ vectors are treated with test material and the resulting *in vivo* mutagenic events are scored by recovering the reporter genes from mouse genomic DNA and analysing the phenotype of the reporter gene in a bacterial host (Box 8.1). This was initially achieved by identifying the colour of the plaques produced when the recovered phage was used to infect *E. coli* cells grown in the presence of X-*gal*, as follows.

- **The Muta™Mouse model** carries a bacterial *lacZ* gene (3126 bp). This transgene encodes the enzyme β-galactosidase, which metabolises the colourless substrate X-gal to an insoluble blue product. When this gene is mutated, an inactive enzyme is produced and the plaques formed are colourless or pale blue, whereas the plaques formed by wild-type phage are dark blue.
- **The Big Blue® model** uses the bacterial *lac*I gene (1080 bp), which encodes the *lacZ* repressor, as a mutational target gene. In mutant phage, the repressor is inactive, active β-galactosidase is expressed and blue plaques are formed in the presence of the chromogenic substrate X-gal. Wild-type phage produce colourless plaques because the expression of β-galactosidase is blocked by the repressor.

Box 8.1 Procedures for the Muta™Mouse and Big Blue® assays

The procedures for the Muta™Mouse (*lacZ* gene) and Big Blue® (*lacI* gene) assays are similar (Dean *et al*., 1999). In both assays, animals carrying bacterial target genes are treated with the test compound. The transgenes are then recovered by *in vitro* packaging of mammalian DNA and the resulting phage is used to infect *E. coli* under conditions which allow selection or HTS for scoring of mutations. The first few steps of the two assays are similar.

1. Transgenic mice (for Big Blue®, rats are also available) are treated with the chemical of interest. The protocol used depends upon the sensitivity and proliferation rate of the target tissue, but daily treatment for 28 days is normally recommended.

[14] The subsequent development of this class of assay has been the subject of a very thorough review Lambert *et al*. (2005).

2. Sufficient time is allowed for any mutations to be fixed *in vivo*. The time required can range from ~1 week for rapidly dividing tissue (e.g. bone marrow, intestine) to at least a month for quiescent tissues (e.g. liver); correct timing of sacrifice is critical. Following 28 days' repeated dosing, a period of 3 days after the final dose is recommended.
3. Tissues of interest are removed and high-molecular weight DNA is extracted.
4. The target gene (*lacZ* or *lac*I) is incorporated into λ preheads using a commercial λ packaging extract to generate viable λ phage.
5. The phage is used to infect *E. coli* cultures growing on medium containing X-gal.
6. After 16 h, the plaques produced are scored.
7. If desired, DNA sequencing of the mutant phage can be used to characterise the mutation spectra. However, this is laborious, adds to the cost of the assay and is not required for regulatory purposes.

One of the issues with early efforts to apply these models was their reliance on visual colony counting, which is time consuming and tends to be subjective (Bielas, 2002). A positive selection system taking advantage of *E. coli gal*E, the gene encoding galactose epimerase which converts toxic galactose-6-phosphate to glucose-6-phosphate, was therefore introduced. *E. coli* strains which lack a functional *gal*E gene are sensitive to galactose because they accumulate toxic intermediates during galactose metabolism, and this can be used to select the *lacZ*$^-$ genotype. In the presence of phenyl-β-D-galactoside, functional β-galactosidase generates toxic galactose-6-phosphate; thus, only *lacZ*$^-$ mutants can make plaques on a lawn of *gal*E$^-$ *E. coli*.

Second-generation transgenic mutation assays which take advantage of the cII gene of λ phage have also been developed. These allow positive selection for base changes and frameshift mutations and can be combined with either the Muta™Mouse or Big Blue® system. The cII gene is derived from the λ phage vector and encodes a repressor gene which controls the λ lytic/lysogenic cycle. It can be used for positive selection in *Hfl* protease-deficient strains of *E. coli*. In *Hfl*$^-$ *E. coli* strain G1225, phages with a functional cII gene are unable to enter the lytic cycle: in other words, only phage with a mutated cII gene can make plaques on a lawn of *Hfl*$^-$ bacteria. The cII system can be used in combination with either the Muta™Mouse or Big Blue® system, and since this is a separate selection system it allows independent confirmation of results obtained by colorimetric screening.

The *gpt* delta model allows detection of point mutations using 6-thioguanine (only *gpt*$^-$ cells can survive on plates containing 6-thioguanine) and deletions using the *spi*$^-$ system. The *gpt* gene of *E. coli* encodes the enzyme guanine phosphoribosyl transferase, which is involved in the bacterial purine salvage pathway and catalyses an obligatory step in the incorporation of guanine into DNA. It also phosphorylates 6-thioguanine, which is toxic to cells when incorporated into DNA. *E.coli* cells with a functional *gpt* gene cannot survive on plates supplemented with 6-thioguanine because of this toxicity, so under these conditions only *gpt*$^-$ cells form colonies. The detection of deletions in *gpt* delta mice takes

advantage of *spi*$^-$ (sensitive to P2 interference), which preferentially selects deletion mutants of λ phage. The *spi*$^-$ selection assay is based upon the fact that the growth of wild-type λ phage is restricted in *E. coli* cells carrying λ prophage P2. Replacement of the λ *red* and *gam* genes with foreign DNA (due to chemical-induced rearrangement) generates *spi*$^-$ phage which can grow in P2 lysogenic *E. coli* cells.

The available commercial transgenic models such as Muta™Mouse and Big Blue® are limited in their responsiveness to clastogens since they are not designed to detect chromosomal rearrangements (Lynch *et al*., 2008). The *lacZ* plasmid transgenic model was developed in order to address this (Mahabir *et al*., 2008). As its name suggests, this contains a *lacZ* plasmid construct and it is able to detect small mutations as well as large deletions (≥500 bp). In the *lacZ* plasmid assay, isolated genomic DNA from treated mice is digested with *Hin*dIII to release single linearised copies of the plasmid from the tandem array. These are purified by adsorption onto magnetic beads coated with *lac* repressor protein, recircularised with T4 ligase, electroporated into *gal*E–*lacZ*$^-$ *E. coli* and selected on P-*gal*. This method allows the detection of a range of deletions arising within the tandem array as well as deletions extending beyond the *lacZ* target gene into adjacent 3′ flanking sequences. Short-term mouse embryonic fibroblast cultures from this line are useful for studying mechanisms by which chemicals can alter genome maintenance in mouse cells (Busuttil *et al*., 2006; Mahabir *et al*., 2009).

Strengths and limitations Transgenic rodent mutation assays allow phenotypically neutral gene mutations to be analysed in a range of tissues. Compared with other *in vivo* mutation assays, they are technically undemanding, give reproducible results and involve relatively simple molecular analysis. They are flexible with respect to route of administration, applicable to many tissues (although they may require separate optimisation for each tissue) and use relatively few animals compared with conventional methods (Eastmond *et al*., 2009). However, the limitations of these models, which should be taken into account when interpreting the results obtained, include the following.

- **Use of non-expressed transgenes:** The fact that the transgenes used in the Muta™Mouse, Big Blue® and *gpt* delta models are not expressed *in vivo* is presented as an advantage, since they are considered to be phenotypically neutral. However, there is evidence that DNA repair is preferentially targeted to transcriptionally active genes, so damage detected in non-expressed transgenes may not reflect that which might occur in a key functional gene.
- **High spontaneous mutation frequency:** The transgenes in rodent transgenic mutation models have mutation rates of $\sim 4 \times 10^{-5}$, much higher than those of endogenous genes in the same animals (Young *et al*., 1995; Lemieux *et al*., 2011). This may mean that a heavily methylated, transcriptionally inactive transgene has increased susceptibility to mutation, leading to an overestimate of the hazard associated with compounds tested in these transgenic systems. The absolute values observed vary with age, tissue and the animal model (*lacZ* > *lac*I > *gpt* > *spi*$^-$). This increased mutation frequency may be because, in addition to their high GC content, they have a high CpG content and

multicopy concatameric structure which favour hypermethylation and are associated with their transcriptionally inactive nature. It is also not known whether repair of damage in the transgenes differs from that in endogenous genes. In addition, *ex vivo* and *in vitro* mutations are theoretically possible in these systems, although they are thought to be extremely rare in properly conducted studies.

- **Characterisation of transgenes:** The use of transgenic mouse models for mutation testing is dependent upon adequate characterisation, but in fact verification of the sequence of the Muta™Mouse transgene was only published in 2010 (Shwed *et al*., 2010). Furthermore, detailed characterisation of the λGT10 transgene in Muta™Mouse has demonstrated the existence of at least 10 defective, rearranged copies of the vector and shown that only about 2/3 of the transgene copies are actually amenable to recovery and use in *E.coli* mutation assays (Shwed *et al*., 2010).
- **Lack of sensitivity to clastogens:** Transgenic rodent mutation assays have been criticised on the grounds that they do not respond equally to all types of DNA damage. The types of damage detected are primarily base pair substitution mutations together with a few frameshift mutations and small insertions/deletions. These assays are relatively insensitive to clastogens (compounds which cause DNA breaks leading to large chromosomal deletions).
- **Genetic background:** The Big Blue® mouse model is available on C57BL/6 and B6C3F1 backgrounds. The B6C3F1 is the strain used by the US National Toxicology Program (NTP), so results obtained using the Big Blue® system may be interpreted in the light of those of NTP bioassays. The Big Blue® transgene is also available as a transgenic rat model. However, the genetic background of the Muta™Mouse system is CD2F1. This background is not commonly used for carcinogenicity testing, leading to questions regarding the relevance of results in this model.
- **Determination of the cell type undergoing DNA damage:** The Muta™Mouse and Big Blue® assays were developed for use *in vivo*, but one limitation of the *in vivo* Big Blue® assay is that it is difficult to identify the specific cell type carrying the mutation since DNA from whole tissues is usually examined representing an average of all the cell types present. With this in mind, cell lines have been generated from some transgenic models for use *in vitro*. These include a Big Blue® mouse embryonic fibroblast line, a series of Big Blue® rat lines (the mammary gland epithelial line BBR-ME, the mammary gland fibroblast line BBR-MFib and the oral cavity epithelial line BBR-OE) (McDiarmid *et al*., 2001) and a *lac*I-containing Rat2 embryonic fibroblast line.
- **Cost!** These assays are extremely expensive compared with the core assays described previously.

As long as these limitations are taken into account, transgenic rodent mutation assays are suitable for compounds which are poorly absorbed, highly reactive and/or rapidly metabolised, especially in cases where the core assays have generated equivocal results (Eastmond *et al*., 2009). In some cases, they may be more relevant to risk assessment than conventional assays because of their ability to detect effects at the site of first contact.

Example: Site of contact effects of β-propiolactone in the stomach

The compound β-propiolactone is a site of contact mutagen which causes stomach and skin tumours following oral and topical administration, respectively. This compound is negative in the mouse bone marrow micronucleus and liver unscheduled DNA synthesis (UDS) assays following oral dosing but when the Muta™Mouse assay was used it proved to be positive in the stomach following oral dosing, consistent with its known carcinogenic effect in this tissue. The use of Muta™Mouse or Big Blue® as site of contact assays has therefore been recommended for compounds which are positive in *in vitro* mutagenicity assays but negative in conventional *in vivo* assays. As well as providing additional evidence as to whether the *in vitro* positives for such compounds are true or false results, this approach provides the opportunity to investigate their effects in the tissues which are likely to be exposed in an occupational or therapeutic setting.

8.6 DNA damage and its repair

Genotoxic damage does not have biological consequences unless it is fixed by cell division, but if lesions induced by DNA-damaging agents are not repaired before replication occurs, they can go on to form mutations in critical genes. The types of DNA damage observed in mammalian cells include single-DNA strand breaks, double-DNA strand breaks, chemical modifications to bases and sugars (adduct formation), mismatches and inter-/intra-strand cross-links (Maynard *et al*., 2009). The presence of adducts can affect replication, leading to substitutions, or breaks in the DNA chain which can lead to mutations or loss of genetic material. The genetic changes which have been linked to adduct formation include deletions, frameshift mutations and nucleotide substitutions.

All types of mutations require cell division to become evident; gene mutations need sufficient divisions to form a clone before they can be detected, whereas numerical and structural chromosome aberrations are best visualized in the division after exposure, because these types of damage usually compromise the potential for division and/or limit the potential life-span of the daughter cells.

8.6.1 DNA damage

Chemically induced damage to DNA can include adduct formation and the induction of DNA strand breaks. Methods are available for detecting each of these types of damage; these include alkaline elution and unwinding assays based upon the effects of DNA strand breaks on the behaviour of DNA molecules under alkaline conditions, which date from the early 1970s, and newer techniques such as the Comet and ^{32}P-postlabelling assays. However, methods such as the alkaline elution method and the DNA alkaline unwinding assay,[15] which detect single-DNA strand breaks in the DNA of treated cells, have fallen out of favour in recent years and have now been almost entirely superseded by the Comet assay, a quick, simple and sensitive method for detecting single- and double-stranded DNA strand breaks in individual cells (Tice *et al*., 2000; Collins *et al*., 2008).

[15] For background on these assays, see Ahnstrom (1988).

The Comet assay The Comet (single-cell gel electrophoresis) assay evaluates compounds in terms of their ability to cause DNA strand breaks and/or acid labile sites and is used to detect DNA damage in cells which have been exposed to chemicals or physical insults *in vitro* or *in vivo*. It is based upon the ability of negatively charged DNA-containing DNA strand breaks to move through an agarose gel in response to application of an electric field and is colloquially known as the Comet Assay because the appearance of the cells after electrophoresis looks like a comet, with a solid sphere (the nucleus) and a tail formed by the movement of fragmented DNA. When carrying out a Comet assay it is important to take cytotoxicity into account in order to distinguish between DNA damage due to genotoxicity and the DNA degradation that occurs during necrosis; if this is done properly the extent of migration of the DNA is a function of the frequency of DNA strand breaks due to chemical exposure.

In the Comet assay (Box 8.2), a suspension of cells which have been treated with the suspect genotoxin is mixed with low melting point agarose and mounted on a microscope slide. The cells are then lysed with alkali, electrophoresed at high pH and stained with a fluorescent dye. The intensity of the resulting comet tail can be assessed visually or by using image analysis software.

Box 8.2 The Comet assay procedure

The preferred version of the Comet assay for detecting genotoxins is the alkaline version (carried out at pH 13 or above), which can detect single-DNA strand breaks, alkali-labile sites, DNA–DNA/DNA–protein cross-links and incomplete DNA repair. It can be applied to cells following *in vivo* treatment or *in vitro* exposure (with the use of a metabolic activation system, if appropriate). Briefly, the procedure is as follows.

1. Cells are layered onto a microscope slide in low melting point agarose.
2. The cells are lysed for 1 h in a high-salt/detergent solution to release the DNA.
3. Alkali exposure is then used to denature the DNA, revealing the single-DNA strand breaks and causing the alkali-labile sites to be expressed as single-DNA strand breaks.
4. Electrophoresis is carried out at pH 13.
5. The gel is neutralised using, for example, Tris buffer at pH 7.5.
6. The comets are stained with a fluorescent dye (e.g. ethidium bromide, propidium iodide, DAPI or YoYo-1).
7. Tail length, tail intensity or so-called 'tail moment' (a function of length and intensity) is measured using an image analyser such as the Comet Assay IV system provided by Perceptive Instruments.[16] For cells exposed *in vitro*, at least 50 cells per exposure (25/slide) should be analysed and for *in vivo* treatment the corresponding number is at least 100/animal (50/slide and 6 animals/group), but no particular advantage is derived from increasing the number of cells analysed per slide above 50 or further increasing the number of animals per group (Smith *et al.*, 2008).

[16] http://www.perceptive.co.uk/cometassay/

The Comet assay is sensitive to very low levels of DNA damage (it can detect from ~100 to several thousand breaks per cell) and requires only small numbers of cells (and correspondingly small amounts of test material). It can be applied to human lymphocytes, making it a useful technique for the biomonitoring of DNA damage in human populations, and has become popular because of its flexibility, cheapness, rapidity and ease of application. It now has its own online interest group[17] and has a role in *in vivo* testing because it can be applied to a variety of target organs as well as blood.

Example: Detection of single- and double-DNA strand breaks in cervical cancer

In a nice clinical example, a group in Mexico has investigated global DNA damage in epithelial cells from women with low- and high-grade CIN lesions and tumours using the neutral Comet assay (which is specific to double-DNA strand breaks) and the alkaline Comet assay (which detects both single- and double-DNA strand breaks) (Cortes-Gutierrez *et al*., 2012). Distinguishing between these two types of damage is important because single-DNA strand breaks are readily repaired and do not represent a severe threat to genomic integrity, whereas double-DNA strand breaks are likely to lead to chromosomal abnormalities. The types of comet observed, and their classification, are illustrated in Figure 8.4. In this small study (20 patients and 10 controls) there was a clear increase in global DNA damage and double-DNA strand breaks in women with more advanced tumours, indicating that tumour progression was associated with DNA damage and possibly with the acquisition of genomic instability, consistent with the concept of the mutator phenotype. In the case of these cervical lesions this process was thought to be associated with Herpes Virus 16 infection. The results of this study suggest that, at least in principle, the Comet assay could be used as a prognostic test in cervical cancer; however, a lot more evaluation would be required before this could be implemented in the clinic.

The choice of tissue for study in the *in vivo* Comet assay is influenced by the question under consideration (e.g. target organ, effects at the site of first contact or a measure of systemic exposure), the accessibility of the tissue and the ease of separating intact cells for the assay. Cells are easily obtained from blood and bone marrow by simple centrifugation and filtration whereas mechanical disruption is required for solid tissues such as liver and stomach. During this procedure, processes such as endonuclease activation, oxidative DNA damage and nuclear disruption must be prevented. Even so, background DNA damage can still be a problem in certain tissues; for example, higher levels of DNA damage and non-detectable cell nuclei tend to be observed in stomach compared with other tissues because of endogenous damage and damage during cell preparation (Smith *et al*., 2008).

For the *in vivo* Comet assay, positive controls should be chosen on the basis that they are known to cause DNA damage in the tissue of interest. The most popular positive controls are ethyl methane sulfonate (EMS) and 2-AAF, but both of these are problematic because of their physical properties: 2-AAF induces DNA damage in the liver but is difficult to formulate; EMS induces DNA damage in

[17] http://www.cometassay.com/

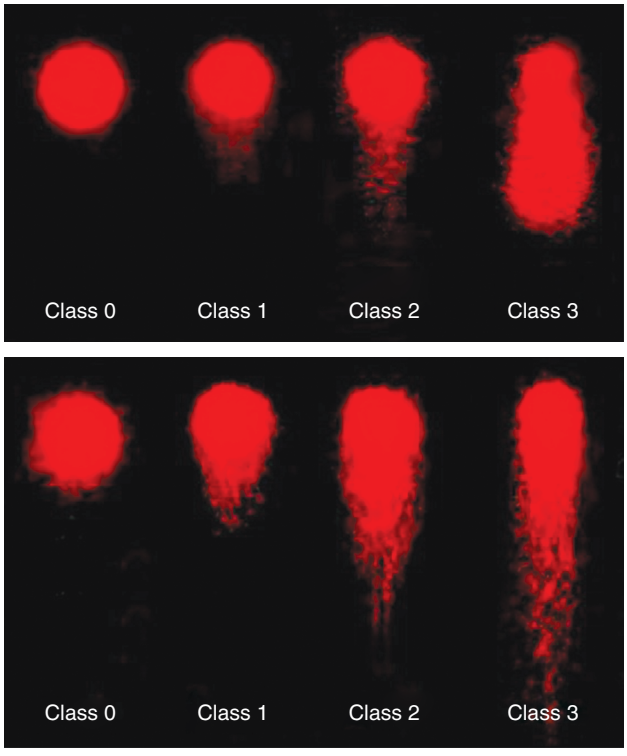


Figure 8.4 Example of a Comet assay. Global DNA damage and DNA double strand breaks were evaluated in cervical epithelial cells from Mexican women with low-grade and high-grade squamous intraepithelial lesions using alkaline and neutral comet assays to help build a better picture of the degree and type of DNA damage present in these patients. Comets in the alkali assay revealed global DNA damage and those in the neutral nondenaturing assay exhibited DNA double strand breaks. Classification of the comets; alkaline comet assay reflecting global DNA damage (upper panel), and neutral comet assay, mainly reflecting double strand breaks (lower panel). Slides were stained with propidium iodide (1 μg ml^{-1}) and examined at 400× magnification using an inverted fluorescence microscope equipped with an excitation filter at 549 nm and a barrier filter at 590 nm and attached to a video camera. Each comet was assigned a value of 0–3 according to its class: undamaged cells with normal nuclei (Class 0; no or barely detectable tails), 1–3 representing increasing relative tail intensities from a halo around the nucleus (Class 1) through gradual increases in the length of the comet tail evolving in parallel with a decrease in the nuclear DNA content (Class 2 and Class 3) (source: Cortes-Gutierrez *et al.*, (2012); figure 1)

all tissues, but its use in restricted because of its volatility. An alternative positive control is methylnitrosourea (MNU), which induces DNA damage in a range of mouse tissues and adducts in all tissues of the rat. This compound induces comet formation in Han Wistar rat liver, stomach, bone marrow and peripheral blood 3 and 24 h after dosing (Smith *et al.*, 2008).

Overall, the *in vivo* Comet assay is capable of giving information about the different types of damage present in cellular DNA and also about the cell's ability to repair the damage (Collins *et al.*, 2008); however, the utility of the *in vitro* version of the assay is considered to be more limited. With regard to the choice of

assays for Stage 2 testing, the rodent Comet assay has gained acceptance as a standard technique for evaluating DNA damage when an *in vivo* test is required for regulatory purposes (Smith *et al.*, 2008). At the time of writing it does not yet have an OECD guideline (EFSA, 2012), but one is in preparation and should be signed in 2013.

^{32}P-postlabelling A number of highly sensitive methods are available for the detection of low levels of DNA damage in a research context, although these vary greatly in terms of sensitivity, selectivity and practicality for human studies (Farmer and Singh, 2008). They include mass spectrometry methods and ^{32}P-postlabelling, both of which can detect adduct levels down to about 1 per 10^9 to 10^{12} nucleotides. The significance of such low levels of adduct formation is still the subject of debate because the background level of DNA damage can be as high as 1 per 10^5 nucleotides (Brink *et al.*, 2009).

In the ^{32}P-postlabelling method, sample DNA (which must be of the highest possible quality) is digested to nucleoside 3′-monophosphates, enriched to separate adducted and non-adducted nucleotides and 5′-phosphorylated using γ-^{32}P-ATP and polynucleotide kinase. It is important to ensure that complete hydrolysis of the DNA is achieved, giving the appropriate proportions of all four bases, since incomplete hydrolysis is associated with adduct loss and high background (Munnia *et al.*, 2007). The resulting mixture of labelled nucleotides is separated on 2D thin layer chromatography or HPLC, and adducts are identified by comparison with standards. This method typically detects a mixture of adducts which form the so-called ‘diagonal radioactive zone’ and has broad specificity for bulky adducts such as those formed by PAHs (e.g. B(a)P).

Example: Effects of fish oil on adduct formation by polycyclic aromatic hydrocarbons

In a study aimed at investigating the potential value of fish oil as a chemoprotective agent, B6C3F1 mice were given a diet supplemented with fish oil (11.5%; controls received corn oil) for 30 days then injected with a mixture of PAHs (Zhou *et al.*, 2011). Liver DNA was extracted and analysed by ^{32}P-postlabelling. The pattern of adducts detected is illustrated in Figure 8.5. The pattern of adducts detected was qualitatively similar between mice-fed fish oil and corn oil, the major adduct detected being B(a)P diol epoxide-dG, but the formation of both total adducts and B(a)P diol epoxide-dG a week after dosing was lower in mice treated with fish oil than in those which were fed a corn oil-containing diet. This was associated with up-regulation of Cyp1a1 and GstT1 in the fish-oil-treated mice and with attenuation of overt toxicity. These results suggest, firstly, that fish oil could potentially have a chemoprotective effect (at least in mice!) and secondly, that Cyp1a1 may be playing a role in detoxification rather than metabolic activation in this setting.

The advantages of the ^{32}P-postlabelling method are that it requires only a few micrograms of DNA per analysis, is highly sensitive and is applicable to a wide variety of adduct structures (although it cannot be used for adduct characterisation). Conversely, it is labour intensive, requires the use of high levels

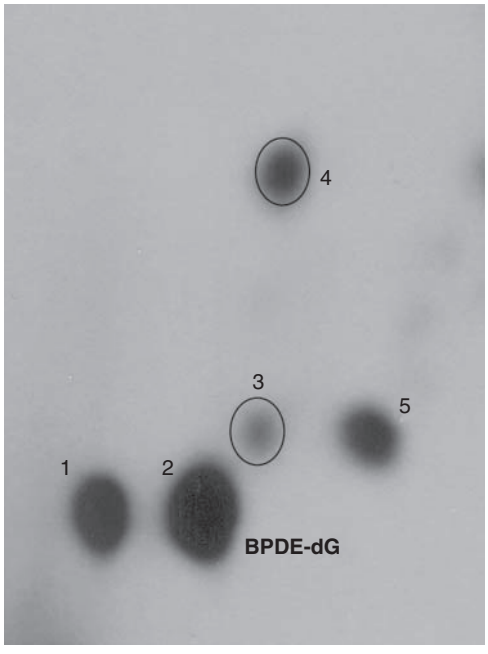


Figure 8.5 Typical representative pattern of ^{32}P-postlabeled PAH-DNA adducts in mouse liver. Mice were treated with fish oil (11.5% in the diet) for 30 days, then injected with a mixture of high-dose PAHs i.p. (total of 16.48 mg per 20 g body weight). Animals were terminated at 7 days after treatment and DNA adducts were detected by 32P-postlabelling. Spot 2 represents the deoxyguanine adduct of B(a)P diol epoxide while total PAH-DNA adducts were quantitated by adding up the intensities of spots 1–5 (source: Zhou *et al.* (2011), figure 1)

of radioactivity and has limited scope for automation (Kovacs *et al.*, 2011). Its sensitivity is limited by the specific activity of the γ-^{32}P-ATP used and the yield of the labelling process, but despite its limitations, it is still considered to be better than MS-based detection for many applications (Farmer and Singh, 2008). It is more widely applicable, especially if the amount of DNA available is limited, and can be used for global screening of adduct formation. This is an important advantage because one of the problems is trying to characterise the effects of real life exposure to mixtures of potentially genotoxic chemicals.

8.6.2 DNA repair

Damaged DNA has to be repaired in order to maintain the integrity of the genome. Protective mechanisms such as metabolic detoxification and DNA repair can act before or after DNA damage has occurred, but must take effect before the lesion is fixed and converted to a mutation. When nucleotides have been damaged, however, DNA repair enzymes can introduce incorrect nucleotides leading to fixation of the damage and, following DNA replication, to a mutation.

DNA repair is co-ordinated with the cell cycle and involves several key pathways: base excision repair, nucleotide excision repair, mismatch repair and repair of double-DNA strand breaks.[18]

- **Base excision repair** involves the recognition and repair of base modifications, abasic sites and single-DNA strand breaks, that is, small lesions where the structure of the double helix is not too severely distorted. The main source of this type of damage is attack by reactive oxygen species, which can lead to high rates of damage even in the absence of exogenous DNA-damaging agents. One example of this type of damage is the formation of 8-oxoguanine, an indicator of oxidative stress which can occur as a result of oxidation of a guanine base within DNA or incorporation of an oxidised dGTP moiety during DNA synthesis. In this case, mutagenicity arises from the fact that 8-oxoguanine can mismatch with adenine, leading to generation of a GC to TA mutation in the next round of DNA synthesis. The insertion of 8-oxoguanine into DNA is prevented by the enzyme 8-oxo-7,8-dihydroguanosine triphosphatase, which hydrolyses 8-oxo dGMP, but when an 8-oxoguanine residue is inserted or generated *in situ* it can be repaired by the DNA glycosylase enzyme 8-oxoguanine DNA glycosylase I, which cleaves the glycosidic bond leaving an abasic site.

  A related type of repair is mediated by the enzyme O^6-methylguanine DNA methyl transferase, which specifically repairs O^6-methylguanine adducts. This is an unusual enzyme (in fact according to the strictest criteria it is not an enzyme at all) because the reaction is a stochiometric so-called suicide reaction in which the enzyme molecule itself is destroyed. *De novo* synthesis of an enzyme molecule is therefore required for the repair of each adduct. The reason for this is that the enzyme acts via a sulfydryl mechanism in which the methyl group from O^6-methylguanine is irreversibly bound to the enzyme. No human diseases are known to be associated with base excision O^6-methylguanine DNA methyltransferase defects, but cell lines with such defects have increased susceptibility to the toxicity of methylating agents.
- **Nucleotide excision repair** is the major repair pathway for removal of pyrimidine dimers and bulky lesions which distort the DNA helix, such as those caused by alkylating agents, UV radiation and some types of reactive oxygen species. This type of repair is defective in the disease xeroderma pigmentosum, an autosomal recessive genetic disorder in which patients develop tumours on any area of skin which are exposed to sunlight. Patients with this disease also have strange skin pigmentation (hence the name of the disorder). Xeroderma pigmentosum patients can be classified into seven complementation groups, implying the existence of seven possible defective genes. These have now been identified and shown to be involved in the process of nucleotide excision repair. It has also recently been shown that families with a high incidence of basal cell carcinoma of the skin have reduced rates of nucleotide excision

[18] For a review of the repair of oxidative lesions, see Berquist and Wilson (2012) and another in the context of cancer, see Wallace, Murphy, and Sweasy (2012).

repair compared with the rest of the population. In addition, the nucleotide excision repair capacity of normal individuals declines by about 1% per year as they age; this could help to explain the very high incidence of basal cell carcinoma in elderly people who have worked in the sunshine for many years (e.g. building workers, shepherds).

- **Mismatch repair** is an editing function in which an erroneously incorporated nucleotide or a small insertion/deletion is removed and replaced with the correct one. This type of error usually arises as a result of replication errors (which may be spontaneous or occur as a consequence of chemical damage) or polymerase slippage. Defects in mismatch repair are found in 10–20% of colorectal tumours, and are present in the germline of patients with hereditary non-polyposis colon cancer. These patients have an elevated rate of mutation which makes them more susceptible to acquiring further mutations leading to the development of tumours.
- **Repair of double-DNA strand breaks**: This type of repair becomes necessary when the sugar phosphate backbones of both strands of the double helix have been broken, which can occur in response to insults such as ionising radiation, reactive oxygen species and chemical damage, but can also happen spontaneously during DNA replication. DNA double strand breaks are a potent form of genotoxic damage.

The induction of DNA repair can be used as a marker of DNA damage. The rationale behind this approach is that cells have evolved mechanisms to repair DNA damage which can lead to genotoxicity, and that the extent of this repair activity can be used to estimate the extent of DNA damage which has occurred in response to an insult. The most commonly used method involves measuring so-called UDS (which occurs as a consequence of DNA damage leading to the need for DNA repair).

The unscheduled DNA synthesis assay The UDS assay measures DNA excision repair (long-patch DNA repair) in order to provide indirect evidence of the DNA-damaging capacity of the test compound. It can be carried out either *in vitro* or *in vivo* and measures DNA synthesis induced for the purpose of repairing excised segments of DNA-containing chemically induced damage. The radioactive label ^{3}H-thymidine is used to identify newly synthesised DNA, usually in liver because of the low background level of replication in this tissue. The UDS assay is more sensitive to damage that induces nucleotide excision repair than that which induces base excision repair because the former involves the incorporation of ~100 new nucleotides, whereas the latter only requires insertion of one to three nucleotides; thus, more ^{3}H-thymidine is incorporated during nucleotide excision repair than during base excision repair.

The UDS assay is usually conducted in rats using at least two dose levels. The animals are sacrificed 2–4 h and 12–16 h after dosing, hepatocyte cultures are prepared and the cells are exposed to ^{3}H-thymidine for 3–8 h. Following autoradiography, cytoplasmic and nuclear grain counts are taken and the nuclear net grain count (nuclear count – cytoplasmic count) is calculated on a cell-by-cell basis in at

least 100 cells. Chemicals which induce a significant increase in net nuclear grain count (i.e. UDS) are considered to have induced DNA damage. The UDS assay can be applied to any cell type with a low enough background replication rate and is adaptable for use in spermatocytes to identify germ cell effects. It also has the advantage that it addresses the entire genome rather than specific loci.

A dual-labelling approach has been used to improve the sensitivity and productivity of the UDS assay and reduce its subjectivity (Li *et al.*, 2008). This uses a fluorescent DNA intercalating stain (DAPI, 4′,6-diamidino-2-phenylindole) to identify viable cell nuclei while those undergoing DNA repair are labelled with ^{3}H-thymidine. In this method, nuclei replicating cells appear as solid black objects, apoptotic nuclei are seen as fragments and nuclei undergoing UDS are mottled in appearance. This approach makes it possible to distinguish between real nuclei and labelling artefacts.

A purely *in vitro* version of the UDS test can be carried out using the HepG2 cell line (Valentin-Severin *et al.*, 2004). These cells do divide in culture, though relatively slowly, but it is relatively easy to distinguish between replicative DNA synthesis (which leads to intensely black stained nuclei) and UDS (which leads to deposition of individual grains in nuclei).

Evaluation Although the Comet and UDS assays can be carried out *in vitro* they add little to the value of the results of the core tests, and their main utility is *in vivo*. The rat liver UDS assay and Comet assay give similar results, but the Comet assay is better at identifying rodent carcinogens and is therefore preferred by some advisory bodies (COM, 2011). Tests for DNA strand breakage are best described as supplementary problem solving tests. The UDS assay is, as stated, specific and insensitive, but it does have a role if the target organ is the liver and the chemical class is likely to cause lesions that result in DNA excision repair only (i.e. it will only detect compounds that produce bulky adducts). In addition, the UDS assay does not identify the actual nature of the damage to DNA, nor does it provide any information about the fidelity of DNA repair.

8.7 Thresholds

Direct-acting genotoxicants can induce mutation as a result of a single reaction of the compound with DNA and are therefore conventionally considered to act by a ‘single hit, single target’ mechanism. Genotoxic processes are therefore conventionally assumed not to be subject to thresholds. However, consideration of micronucleus data for a range of aneugens and clastogens (Elhajouji *et al.*, 2011) indicates that aneugens do frequently exhibit thresholds, whereas the data for clastogens are less clear-cut.

The potentially severe consequences of exposure to genotoxins and the consequent importance of carrying out a correct risk assessment have led to intense debate around the question of whether genotoxicity does exhibit thresholds; experts struggle to agree on the level of adducts below which no adverse effects would be expected (Pottenger *et al.*, 2009). However, as our understanding of

metabolic detoxification, mechanisms of genotoxicity and DNA repair improves, it is becoming apparent that certain genotoxic processes may exhibit thresholds. The types of thresholds observed include the following:

- **Metabolic thresholds:** These are observed when a compound which undergoes metabolic elimination at low doses exceeds the capacity of detoxification enzymes within the cell leading to toxicity and/or the development of neoplasia.
- **Threshold for aneugens:** Aneugens may act on non-DNA targets rather than interacting directly with DNA. This class of compounds may need to occupy a critical number of target sites in order to exert a biological effect; it includes spindle poisons which require multiple interactions with the mitotic spindle before they disturb microtubule dynamics and affect chromosome segregation.
- **Other thresholds:** These may apply to compounds which act via indirect mechanisms such as oxidative damage, cytotoxicity and inhibition of DNA synthesis. Homeostatic mechanisms such as DNA repair, target redundancy and the maintenance of a reducing environment within cells help to prevent damage from becoming permanent until the point at which they become saturated, and this can lead to the existence of a threshold dose response.

Example: Existence of thresholds in the effects of DNA-oxidising agents

A study of the effects of a range of DNA-oxidising agents in human lymphoblastoid *TK*6 cells (Platel *et al*., 2009) has provided some evidence regarding the existence of thresholds for this type of DNA damage. Indeed, the dose response curves often had more than one plateau, suggesting the existence of complex cellular mechanisms with multiple dose-dependent interactions. Further clarification of the existence of thresholds for genotoxicity requires an improved understanding of the consequences of metabolism and the repair of DNA lesions, together with improvements in the availability and user-friendliness of mathematical models.

The COM has issued a statement on the existence of thresholds for *in vivo* mutagens (COM, 2010). This notes that threshold modes of actions have been identified for *in vivo* mutagens which act on non-DNA targets (e.g. microtubule effects) and protective mechanisms (e.g. DNA repair) or via overloading of detoxification mechanisms (e.g. GSH depletion). It emphasizes the need to consider the statistical power of studies used to define thresholds, the importance of careful dose selection, especially around the observed no effect level, and the need to justify the use of mathematical models on a case by case basis. The statement reaffirms the default assumption that, in the absence of contradictory evidence, no safe level of exposure to genotoxic chemicals can be quantified. However, it recognises the need to consider biologically meaningful thresholds and advises that an appropriate strategy should be devised for each chemical under consideration to identify thresholds and/or NOAELs for potential threshold modes of action. The exact wording of the COM conclusion is reproduced in Box 8.3.

Box 8.3 Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment Guidance Statement: Thresholds for in vivo *mutagens*

(i) The COM reaffirmed the default position that for *in vivo* mutagens, in the absence of mechanistic data to infer a threshold, it is prudent to assume that there is no threshold for mutagenicity.
(ii) If there is good reason to consider that a threshold mode of action is appropriate, then it is necessary to investigate the biologically meaningful threshold for all genotoxic effects that have been reported.
(iii) An appropriate strategy should be devised for each chemical under consideration to identify threshold dose levels or NOELs for all potential thresholded modes of action of genotoxicity, which may include either *in vitro* or *in vivo* studies.

8.8 Conclusions

Numerous methods are available for the identification of genotoxicants; the induction of point mutations and chromosomal alterations can be quantified in bacteria, mammalian cells, *in vitro* tests and transgenic models (Muta™Mouse and Big Blue® mouse lines and associated second generation models) while DNA strand breaks are detected by the Comet assay and DNA repair can be measured in terms of UDS.

Highly sensitive methods such as ^{32}P-postlabelling and MS have made it possible to detect DNA adducts at frequencies as low as $1/10^{12}$ nucleotides, allowing dose responses for genotoxicity in relation to DNA damage to be determined; however, the significance of low dose responses to genotoxic chemicals in terms of human health effects remains debatable, as does the significance of this in terms of risk assessment. Given that the background level of adducts is $1/10^5$ nucleotides the significance of additional adducts at levels as low as 1 per 10^{12} nucleotides requires explanation in terms of the identification of critical targets and corresponding biological effects. In general, the lowest level of DNA adducts found is limited only by the limit of detection of the assay used to detect them, whereas the lowest effect level for a biological effect is defined by the nature of the response (Brink *et al.*, 2007). This means that DNA adducts can already be detected at dose levels much lower than those required for a biological response, and this gap will increase as detection methods become more sensitive.

Self-assessment questions

- What additional information could you obtain by evaluating the genotoxicity of a test item using a transgenic mouse model such as Muta™Mouse rather than a simple Ames test?

- Given that it is no longer possible to test the genotoxicity of cosmetic ingredients *in vivo*, how would you go about ensuring the safety of the ingredients of a new skin cream?
- If the initiation of genotoxic carcinogenesis requires an actual mutation, how useful is it to be able to detect adducts at a frequency of less than one per genome?

References

Aardema, M.J., Barnett, B.C., Khambatta, Z. *et al.* (2010) International prevalidation studies of the EpiDerm™ 3D human reconstructed skin micronucleus (RSMN) assay: transferability and reproducibility. *Mutation Research*, **701**, 123–131.

Ahnstrom, G. (1988) Techniques to measure DNA single-strand breaks in cells: a review. *International Journal of Radiation Biology*, **54**, 695–707.

Avlasevich, S., Bryce, S., De Boeck, M. *et al.* (2011) Flow cytometric analysis of micronuclei in mammalian cell cultures: past, present and future. *Mutagenesis*, **26**, 147–152.

Berquist, B.R. and Wilson, D.M. 3RD (2012) Pathways for repairing and tolerating the spectrum of oxidative DNA lesions. *Cancer Letters*, **327**, 61–72.

Bhalli, J.A., Pearce, M.G., Dobrovolsky, V.N. and Heflich, R.H. (2011) Manifestation and persistence of *Pig-a* mutant red blood cells in C57BL/6 mice following single and split doses of N-ethyl-N-nitrosourea. *Environmental and Molecular Mutagenesis*, **52**, 766–773.

Bielas, J.H. (2002) A more efficient Big Blue® protocol improves transgene rescue and accuracy in a adduct and mutation measurement. *Mutation Research*, **518**, 107–112.

Bonassi, S., Coskun, E., Ceppi, M. *et al.* (2011a) The HUman MicroNucleus project on eXfoLiated buccal cells (HUMN(XL)): the role of life-style, host factors, occupational exposures, health status, and assay protocol. *Mutation Research*, **728**, 88–97.

Bonassi, S., El-Zein, R., Bolognesi, C. and Fenech, M. (2011b) Micronuclei frequency in peripheral blood lymphocytes and cancer risk: evidence from human studies. *Mutagenesis*, **26**, 93–100.

Boverhof, D.R., Chamberlain, M.P., Elcombe, C.R. *et al.* (2011) Transgenic animal models in toxicology: historical perspectives and future outlook. *Toxicological Sciences*, **121**, 207–233.

Brink, A., Richter, I., Lutz, U. *et al.* (2009) Biological significance of DNA adducts: comparison of increments over background for various biomarkers of genotoxicity in L5178Y *tk(+/–)* mouse lymphoma cells treated with hydrogen peroxide and cumene hydroperoxide. *Mutation Research*, **678**, 123–128.

Brink, A., Schulz, B., Stopper, H. and Lutz, W.K. (2007) Biological significance of DNA adducts investigated by simultaneous analysis of different endpoints of genotoxicity in L5178Y mouse lymphoma cells treated with methyl methanesulfonate. *Mutation Research*, **625**, 94–101.

Busuttil, R.A., Rubio, M., Dolle, M.E. *et al.* (2006) Mutant frequencies and spectra depend on growth state and passage number in cells cultured from transgenic *lacZ*-plasmid reporter mice. *DNA Repair (Amst)*, **5**, 52–60.

Collins, A.R., Oscoz, A.A., Brunborg, G. *et al.* (2008) The Comet assay: topical issues. *Mutagenesis*, **23**, 143–151.

COM (2010) Guidance statement: thresholds for *in vivo* mutagens. http://www.iacom.org.uk/guidstate/documents/Thresholdsforinternetfinal.pdf

COM (2011) Guidance on a strategy for genotoxicity testing of chemical substances. http://www.iacom.org.uk/guidstate/documents/COMGuidanceFINAL2.pdf

Cortes-Gutierrez, E.I., Hernandez-Garza, F., Garcia-Perez, J.O. *et al.* (2012) Evaluation of DNA single and double strand breaks in women with cervical neoplasia based on alkaline and neutral comet assay techniques. *Journal of Biomedicine and Biotechnology*, **2012**, 385245.

Corvi, R., Albertini, S., Hartung, T. *et al.* (2008) ECVAM retrospective validation of *in vitro* micronucleus test (MNT). *Mutagenesis*, **23**, 271–283.

Crespi, C.L., Gonzalez, F.J., Steimel, D.T. *et al.* (1991) A metabolically competent human cell line expressing five cDNAs encoding procarcinogen-activating enzymes: application to mutagenicity testing. *Chemical Research in Toxicology*, **4**, 566–572.

Darzynkiewicz, Z., Smolewski, P., Holden, E. *et al.* (2011) Laser scanning cytometry for automation of the micronucleus assay. *Mutagenesis*, **26**, 153–161.

Dean, S.W., Brooks, T.M., Burlinson, B. *et al.* (1999) Transgenic mouse mutation assay systems can play an important role in regulatory mutagenicity testing *in vivo* for the detection of site-of-contact mutagens. *Mutagenesis*, **14**, 141–151.

Dertinger, S.D. and Heflich, R.H. (2011) *In vivo* assessment of *Pig-a* gene mutation – recent developments and assay validation. *Environmental and Molecular Mutagenesis*, **52**, 681–684.

Dertinger, S.D., Phonethepswath, S., Franklin, D. *et al.* (2010) Integration of mutation and chromosomal damage endpoints into 28-day repeat dose toxicology studies. *Toxicological Sciences*, **115**, 401–411.

Dertinger, S.D., Phonethepswath, S., Weller, P. *et al.* (2011a) Interlaboratory *Pig-a* gene mutation assay trial: studies of 1,3-propane sultone with immunomagnetic enrichment of mutant erythrocytes. *Environmental and Molecular Mutagenesis*, **52**, 748–755.

Dertinger, S.D., Phonethepswath, S., Weller, P. *et al.* (2011b) International Pig-a gene mutation assay trial: evaluation of transferability across 14 laboratories. *Environmental and Molecular Mutagenesis*, **52**, 690–698.

Dertinger, S.D., Torous, D.K., Hayashi, M. and MacGregor, J.T. (2011c) Flow cytometric scoring of micronucleated erythrocytes: an efficient platform for assessing *in vivo* cytogenetic damage. *Mutagenesis*, **26**, 139–145.

Eastmond, D.A., Hartwig, A., Anderson, D. *et al.* (2009) Mutagenicity testing for chemical risk assessment: update of the WHO/IPCS Harmonized Scheme. *Mutagenesis*, **24**, 341–349.

EFSA (2012) Minimum Criteria for the acceptance of *in vivo* alkaline Comet Assay Reports. *EFSA Journal*, **10**, 2977.

Elhajouji, A., Lukamowicz, M., Cammerer, Z. and Kirsch-Volders, M. (2011) Potential thresholds for genotoxic effects by micronucleus scoring. *Mutagenesis*, **26**, 199–204.

Farmer, P.B. and Singh, R. (2008) Use of DNA adducts to identify human health risk from exposure to hazardous environmental pollutants: the increasing role of mass spectrometry in assessing biologically effective doses of genotoxic carcinogens. *Mutation Research*, **659**, 68–76.

Fenech, M., Holland, N., Zeiger, E. *et al.* (2011) The HUMN and HUMNxL international collaboration projects on human micronucleus assays in lymphocytes and buccal cells–past, present and future. *Mutagenesis*, **26**, 239–245.

Fowler, P., Smith, K., Young, J. *et al*. (2012a) Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. *Mutation Research*, **742**, 11–25.

Fowler, P., Smith, R., Smith, K. *et al*. (2012b) Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. *Mutation Research*, **747**, 104–117.

Heddle, J.A., Fenech, M., Hayashi, M. and MacGregor, J.T. (2011) Reflections on the development of micronucleus assays. *Mutagenesis*, **26**, 3–10.

Kirkland, D., Aardema, M., Henderson, L. and Muller, L. (2005) Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutation Research*, **584**, 1–256.

Kirkland, D., Reeve, L., Gatehouse, D. and Vanparys, P. (2011) A core *in vitro* genotoxicity battery comprising the Ames test plus the *in vitro* micronucleus test is sufficient to detect rodent carcinogens and *in vivo* genotoxins. *Mutation Research*, **721**, 27–73.

Kirsch-Volders, M., Decordier, I., Elhajouji, A. *et al*. (2011) In vitro genotoxicity testing using the micronucleus assay in cell lines, human lymphocytes and 3D human skin models. *Mutagenesis*, **26**, 177–184.

Kovacs, K., Anna, L., Rudnai, P. and Schoket, B. (2011) Recovery of bulky DNA adducts by the regular and a modified ^{32}P-postlabelling assay; influence of the DNA-isolation method. *Mutation Research*, **721**, 95–100.

Lambert, I.B., Singer, T.M., Boucher, S.E. and Douglas, G.R. (2005) Detailed review of transgenic rodent mutation assays. *Mutation Research*, **590**, 1–280.

Lemieux, C.L., Douglas, G.R., Gingerich, J. *et al*. (2011) Simultaneous measurement of benzo[a]pyrene-induced *Pig-a* and *lacZ* mutations, micronuclei and DNA adducts in Muta™Mouse. *Environmental and Molecular Mutagenesis*, **52**, 756–765.

Li, W., Choy, D.F., Post, J.M. and Sullivan, M.E. (2008) A dual-labeling method to quantify unscheduled DNA synthesis in primary cells. *Journal of Pharmacological and Toxicological Methods*, **57**, 220–226.

Lynch, A.M., Giddings, A., Custer, L. *et al*. (2011a) International *Pig-a* gene mutation assay trial (stage III): results with N-methyl-N-nitrosourea. *Environmental and Molecular Mutagenesis*, **52**, 699–710.

Lynch, A.M., Mahabir, A.G., Bradford, A. *et al*. (2008) Is Muta™Mouse insensitive to clastogens? *Mutation Research*, **652**, 145–150.

Lynch, A.M., Sasaki, J.C., Elespuru, R. *et al*. (2011b) New and emerging technologies for genetic toxicity testing. *Environmental and Molecular Mutagenesis*, **52**, 205–223.

Mahabir, A.G., Van Benthem, J., Korsten, H. *et al*. (2008) Detecting genotoxic effects of potential clastogens: an *in vivo* study using the transgenic *lacZ* plasmid and the Muta™Mouse model. *Mutation Research*, **652**, 151–157.

Mahabir, A.G., Zwart, E., Schaap, M. *et al*. (2009) *lacZ* mouse embryonic fibroblasts detect both clastogens and mutagens. *Mutation Research*, **666**, 50–56.

Maynard, S., Schurman, S.H., Harboe, C. *et al*. (2009) Base excision repair of oxidative DNA damage and association with cancer and aging. *Carcinogenesis*, **30**, 2–10.

McDiarmid, H.M., Douglas, G.R., Coomber, B.L. and Josephy, P.D. (2001) Epithelial and fibroblast cell lines cultured from the transgenic Big Blue® rat: an *in vitro* mutagenesis assay. *Mutation Research*, **497**, 39–47.

Morita, T., MacGregor, J.T. and Hayashi, M. (2011) Micronucleus assays in rodent tissues other than bone marrow. *Mutagenesis*, **26**, 223–230.

Mortelmans, K. and Zeiger, E. (2000) The Ames *Salmonella*/microsome mutagenicity assay. *Mutation Research*, **455**, 29–60.

Muller, L. (2009) 4.3 *In-vitro* genotoxicity tests to detect carcinogenicity: a systematic review. *Human and Experimental Toxicology*, **28**, 131–133.

Munnia, A., Saletta, F., Allione, A. *et al.* (2007) ^{32}P-Post-labelling method improvements for aromatic compound-related molecular epidemiology studies. *Mutagenesis*, **22**, 381–385.

Nohmi, T., Suzuki, T. and Masumura, K. (2000) Recent advances in the protocols of transgenic mouse mutation assays. *Mutation Research*, **455**, 191–215.

Ogawa, I., Furukawa, S., Abe, M. *et al.* (2009) Multi-endpoint genotoxic assay using L5178Y (*Tk(+/−)*-3.7.2c) cells. *Journal of Toxicological Sciences*, **34**, 547–553.

Parry, J.M. and Parry, E.M. (2012) Genetic Toxicology: Principles and Methods, *Methods in Molecular Biology* Vol. 817.

Peruzzi, B., Araten, D.J., Notaro, R. and Luzzatto, L. (2010) The use of *PIG-A* as a sentinel gene for the study of the somatic mutation rate and of mutagenic agents *in vivo*. *Mutation Research*, **705**, 3–10.

Pfuhler, S., Albertini, S., Fautz, R. *et al.* (2007) Genetic toxicity assessment: employing the best science for human safety evaluation part IV: recommendation of a working group of the Gesellschaft fuer Umwelt-Mutationsforschung (GUM) for a simple and straightforward approach to genotoxicity testing. *Toxicological Sciences*, **97**, 237–240.

Platel, A., Nesslany, F., Gervais, V. and Marzin, D. (2009) Study of oxidative DNA damage in TK6 human lymphoblastoid cells by use of the *in vitro* micronucleus test: determination of No-Observed-Effect Levels. *Mutation Research*, **678**, 30–37.

Pottenger, L.H., Carmichael, N., Banton, M.I. *et al.* (2009) ECETOC workshop on the biological significance of DNA adducts: summary of follow-up from an expert panel meeting. *Mutation Research*, **678**, 152–157.

Schuler, M., Gollapudi, B.B., Thybaud, V. and Kim, J.H. (2011) Need and potential value of the *Pig-a in vivo* mutation assay – A HESI perspective. *Environmental and Molecular Mutagenesis*, **52**, 685–689.

Shwed, P.S., Crosthwait, J., Douglas, G.R. and Seligy, V.L. (2010) Characterisation of Muta™Mouse lambda gt10-*lacZ* transgene: evidence for *in vivo* rearrangements. *Mutagenesis*, **25**, 609–616.

Smith, C.C., Adkins, D.J., Martin, E.A. and O'Donovan, M.R. (2008) Recommendations for design of the rat comet assay. *Mutagenesis*, **23**, 233–240.

Speit, G., Zeller, J. and Neuss, S. (2011) The *in vivo* or *ex vivo* origin of micronuclei measured in human biomonitoring studies. *Mutagenesis*, **26**, 107–110.

Tice, R.R., Agurell, E., Anderson, D. *et al.* (2000) Single cell gel/comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing. *Environmental and Molecular Mutagenesis*, **35**, 206–221.

Valentin-Severin, I., Thybaud, V., Le Bon, A.M. *et al.* (2004) The autoradiographic test for unscheduled DNA synthesis: a sensitive assay for the detection of DNA repair in the HepG2 cell line. *Mutation Research*, **559**, 211–217.

Wallace, S.S., Murphy, D.L. and Sweasy, J.B. (2012) Base excision repair and cancer. *Cancer Letters*, **327**, 73–89.

Walmsley, R.M. (2008) GADD45a-GFP GreenScreen HC genotoxicity screening assay. *Expert Opinion on Drug Metabolism & Toxicology*, **4**, 827–835.

Young, R.R., Rogers, B.J., Provost, G.S. *et al.* (1995) Interlaboratory comparison: liver spontaneous mutant frequency from lambda/lacI transgenic mice (Big Blue) (II). *Mutation Research*, **327**, 67–73.

Zhou, G.D., Zhu, H., Phillips, T.D. *et al.* (2011) Effects of dietary fish oil on the depletion of carcinogenic PAH-DNA adduct levels in the liver of B6C3F1 mouse. *PLoS One*, **6**, e26589.

9

Oncogenes and the Identification of Human Carcinogens

9.1 Introduction

Numerous molecular changes occur during tumourigenesis. Some of these are key mechanistic steps in tumour development, but others occur as a consequence of cellular transformation. This chapter describes some of them and discusses experimental methods for investigating the carcinogenic process in its entirety. The evolution of transgenic technology has made it possible to develop a number of transgenic models which can assist in the identification and characterisation of potential human carcinogens, including transgenic mouse lines carrying oncogenes or tumour suppressor genes which have been generated with a view to improving the sensitivity and reducing the duration of carcinogenesis assays. This chapter describes several of these models and places them in the context of conventional bioassays.

9.2 Identification of human carcinogens

9.2.1 The lifetime carcinogenicity bioassay

Genotoxicity tests are central to carcinogenic risk assessment, but although genotoxicity is a key process in carcinogenesis, it is not the only factor which leads to tumour development, and the classification of a compound as a mutagen is not sufficient to identify it as a carcinogen. Indeed, no short-term assay is capable of identifying carcinogens definitively. The only way to identify a carcinogen

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

unequivocally is by exposing an animal to the compound and observing the development of tumours over its entire lifetime. Assays of this kind, which are usually performed in rodents because of their small size and relatively short lifetime (otherwise the assay would be prohibitively time consuming and expensive) are referred to as *lifetime bioassays* or *2-year bioassays*. Rodents (rats or mice) are exposed to the compound of interest either by means of a single dose early in neonatal development (this approach is more appropriate for strong genotoxins which have been identified as potential initiating agents) or by regular dosing (by gavage, inhalation or dermal exposure, depending on the anticipated route of human exposure) throughout life. The animals' health is monitored throughout life and morbidity, mortality and tumour incidence are recorded.

9.2.2 The National Toxicology Program 2-year bioassay

The largest programme of bioassays, and the one which is considered most definitive in the identification of potential human carcinogens, is the one run by the US NTP,[1] which is based at the NIEHS. The NTP carries out lifetime bioassays in both rats and mice in order to provide carcinogenicity data for use in risk assessment. The protocol used for these bioassays is summarised in Box 9.1.

Box 9.1 The NTP 2-Year Study Protocol

The purpose of this study is to determine the toxicologic and/or carcinogenic effects of long-term exposure on rats and mice.

1. Treatment:
   After a 10- to 14-day quarantine period, animals are assigned at random to treatment groups. Rats and mice receive the test agent for 104 weeks via a defined route of exposure at three treatment concentrations plus controls. Animals have continuous *ad libitum* access to dosed-feed and dosed-water in those exposure studies. For inhalation, gavage and dermal studies, animals are treated five times per week, weekdays only. Male mice are housed individually and rats and female mice are group-housed, except for inhalation and dermal exposure studies in which all rats and mice are housed individually.

	Animals		Sexes		Species		Test Groups		Total
Treatment	50	×	2	×	2	×	3	=	600
Controls	50	×	2	×	2	×	1	=	200
Sentinel Animals	15	×	2	×	2	×	–	=	60
Total									860

[1] http://ntp.niehs.nih.gov/?objectid=72015DAF-BDB7-CEBA-F9A7F9CAA57DD7F5

2. Observations:
   Individual animal body weights for test and control group animals are recorded on day one on test and at 4-week intervals thereafter except for dosed-feed and dosed-water studies, which are recorded weekly for the first 13 weeks and monthly thereafter. If life-threatening tumours develop, a significant number of deaths occur, or a significant effect on body weight is observed, the weighing frequency may be increased to every 2 weeks.
   Animals are observed twice daily at least 6 h apart (before 10:00 a.m. and after 2:00 p.m. including holidays and weekends) for moribundity and mortality. Animals found moribund or showing clinical signs of pain or distress are humanely euthanized. Formal examinations for clinical signs of toxicity are made and recorded at 4-week intervals. For dosed-feed or dosed-water studies, food consumption/water consumption is measured and recorded weekly for the first 13 weeks and monthly thereafter.
3. Necropsy and pathology:
   - Necropsy: A complete necropsy is performed on all treated and control animals that either die or are sacrificed. All tissues required for complete histopathology are trimmed, embedded, sectioned and stained with hematoxylin and eosin for histopathologic evaluation. (http://NTP.NIEHS.nih.gov/index.cfm?objectid=070CC36D-0356-DD41-9D9BAB3216C57CE2).
   - Histopathology: All animals in all treatment groups that die (or are sacrificed in a moribund condition) and those that complete the 104-week exposure are subjected to a complete necropsy and slides of all tissues required for complete histopathologic evaluation are prepared and evaluated. (http://NTP.NIEHS.nih.gov/index.cfm?objectid=070CB534-FDD9-ACB0-1654A9869474C5B6).

The NTP bioassay programme has generated a huge amount of valuable data which has provided information on carcinogenic risk assessment for more than 30 years. The vast majority of known human carcinogens are positive in rodent bioassays and, indeed, many compounds which are now known to be carcinogenic in humans were first identified as carcinogens using the lifetime bioassay. However, there are issues with the lifetime bioassay approach and design of the assay itself. For example, it has been criticised as follows (Boobis *et al.*, 2009):

> The two-year cancer bioassay in rodents remains the primary testing strategy for in-life screening of compounds that might pose a potential cancer hazard. Yet experimental evidence shows that cancer is often secondary to a biological precursor effect, the mode of action is sometimes not relevant to humans, and key events leading to cancer in rodents from nongenotoxic agents usually occur well before tumorigenesis and at the same or lower doses than those producing tumors.

Issues with the lifetime bioassay include the problem that it has not been confirmed whether rodent models are representative of carcinogenesis in humans (many of the effects seen in animals are of little relevance to humans), the fact that concomitant exposures are not addressed and the long duration of the studies.

Two aspects of the NTP bioassay which are subject to debate and criticism are the approaches taken in dose setting and the choice of mouse and rat strains used.

- **Dose setting:** The NTP bioassay involves the use of three dose groups (high, medium and low). The top dose is usually set to be as near as possible to the MTD (i.e. the highest dose that animals can tolerate for chronic daily (or 5× weekly) exposure). This is justified on the basis that it should increase the ability of the assay to detect carcinogens of low potency. However, the use of such a high dose may be criticised on several grounds as follows.
  - It has been argued that these doses are too high because they overwhelm the protective mechanisms of the organism (e.g. metabolic detoxification, DNA repair).
  - There is good evidence that mechanisms of carcinogenesis may be different at high doses compared with lower doses. In terms of risk assessment, this could be problematic since humans are not exposed to very high doses of potential carcinogens and so a mechanism which is only observed at a very high dose may not be relevant to humans.
  - Even in the case of a carcinogen which acts by the same mechanism across the entire dose range, the use of a very high dose level may provide results which are not relevant to humans simply because humans never experience such high exposures. There is therefore the risk of generating false positives and raising concerns about a chemical which is without appreciable risk in a human setting.
- **Choice of rodent strains:** The NTP bioassay makes use of mice and rats,[2] but there are issues in extrapolating from either of these species to humans as well as in relation to the specific strains used in the assay.
  - There is a fundamental difference in the spectrum of common age-related cancers between rodents and humans; mice exhibit a preponderance of mesenchymal tumours (lymphomas and sarcomas), whereas epithelial tumours (papillomas, adenomas and carcinomas) are more prevalent in humans (Storer *et al*., 2010).
  - Data going back half a century show that inbred mouse strains have widely differing susceptibilities to hepatocarcinogenesis. Some strains, including C57BL/6J, C57BL/10Sn, DBA/2J and Balb/c, have a very low spontaneous liver tumour incidence (<4% of animals develop liver tumours) whereas others, such as C3H/HeJ, have high incidences (40–50%); similarly, strains can be classified as resistant (e.g. Balb/c, C57BL/6J) or susceptible (e.g. C3H/HeJ, CBA) to chemically induced liver tumours. The mouse strain used in the NTP bioassay is the B6C3F1/N hybrid. This hybrid, which is an F1 cross between two inbred lines, a susceptible line (C3H) and a resistant line (C57BL/6), has a spontaneous liver tumour incidence

[2] For a discussion of the use of inbred and outbred mouse and rat strains in toxicology, see Inselman *et al*. (2011).

of up to 40%. However, it has a low incidence of other types of tumour and identifies known carcinogens efficiently. Furthermore, it is hardy, disease resistant and breeds well. Having reviewed this information (King-Herbert and Thayer, 2006), the NTP decided in 2005 to continue using the B6C3F1/N mouse in its bioassays, with the recommendation that the genomic sequences of the two parental lines should be determined.

– Until recently, the rat strain used in the NTP bioassay was the F344/N. This line has a high background incidence of certain tumour types (including testicular interstitial tumours and mononuclear cell leukaemia) and can have poor fertility. Recently, however, the F344/N strain has been replaced by the more robust Harlan Sprague-Dawley strain (King-Herbert, Sills, and Bucher, 2010). This has resolved some of the strain-specific issues related to the use of the F344/N strain but issues related to rat-specific mechanisms of carcinogenicity still remain: for example, saccharin is a male rat-specific bladder carcinogen (it acts via a mechanism involving deposition of α2 microglobulin which does not occur in humans) so extrapolation from rat bladder carcinogenesis to human cancer risk is problematic.

Mammalian carcinogenicity studies are the gold standard for predicting the effects of chemical carcinogens. Embarking on a mammalian carcinogenicity study is, however, a serious commitment since these assays are of long duration (more than 2 years) and are very expensive in terms of time, money (can be >$2 million), animals (>1000 per assay) and resources (Storer *et al*., 2010). Scientific concerns remain because the assay exhibits a preponderance of species-specific responses, is difficult to extrapolate to low doses and may not contribute to the understanding of carcinogenic mechanisms (Storer *et al*., 2010). Careful consideration therefore has to be given to the choice of species and strains, dose levels and route of administration. The limitations of the assay in terms of predictive power and correlation with other species and man must always be borne in mind when planning, executing and interpreting a lifetime bioassay.

9.3 Genetic changes in cancer

Two major classes of genes, the oncogenes and the tumour suppressor genes, have been identified as playing important roles in mechanisms of carcinogenesis. Interestingly, the theoretical existence of these types of genes was predicted by Boveri in 1914, before DNA was even identified as the genetic material.[3]

Oncogenes tend to be genes which are involved in the regulation of cell proliferation and are activated by mutations or overexpression during the carcinogenic process (gain-of-function mutations). They were initially identified by their ability to induce morphological transformation in NIH3T3 mouse fibroblasts. The first human oncogene to be identified was the Harvey *ras* oncogene (H-*ras*), a mutated intracellular signalling molecule which is permanently switched on.

[3] See Cohen and Arnold (2011).

This was identified in the human bladder tumour T24, which was shown to contain a transforming gene homologous to that of the Harvey murine sarcoma virus v-Ha-*ras*. *Ras* genes encode 21 kDa proteins which mediates signal transduction between tyrosine kinase receptors and the nucleus. They are involved in second messenger signalling between growth factor receptors and the nucleus and can be activated by point mutations at codons 12, 13, 61 or 113–117. *Ras*-mediated signal transduction involves the hydrolysis of GTP; following GTP hydrolysis, guanosine diphosphate is released and the protein returns to the inactive signalling conformation. Mutated *ras* proteins lack GTPase activity and therefore remain in the active signalling conformation all the time (Figure 9.1).

Tumour suppressor genes are cell cycle control genes whose expression is lost, either as a result of deletion or mutation, during the carcinogenic process. Their existence was first predicted in the early 1970s on the basis of tumour incidence statistics. This prediction was based on genetic analysis of the eye

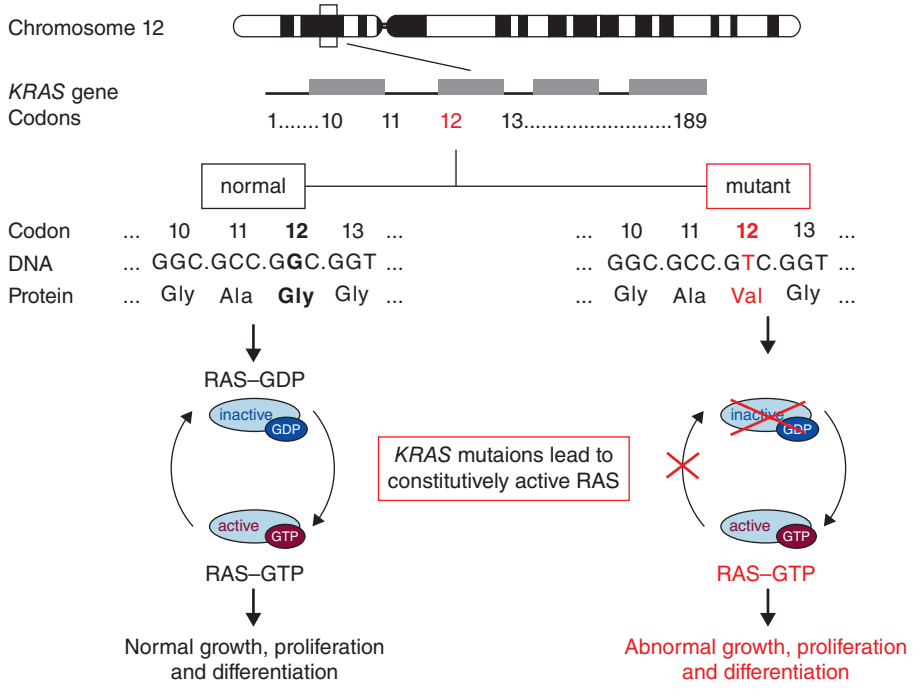


Figure 9.1 Role of *K-ras* mutations in oncogenic activation of intracellular signalling. The human *KRAS* gene, located on chromosome 12, encodes a small G-protein that functions downstream of EGFR-induced cell signalling. This G-protein belongs to the family of RAS proteins involved in signal transduction pathways that regulate cell development and function. RAS proteins normally cycle between active (RAS–GTP) and inactive (RAS–GDP) conformations. Somatic missense mutations in codon 12 of the *KRAS* gene, leading to single amino acid substitutions such as p.Gly12Val, are the most common alterations found in colorectal tumours. These *KRAS* mutations result in RAS proteins that are constitutively in the active RAS–GTP conformation. Unlike wild-type RAS proteins which are deactivated after a short time, the mutated RAS proteins cause continuous activation of RAS signalling pathways in the absence of upstream stimulation of EGFR/HER receptors. This oncogenic activation of RAS signalling pathways leads to abnormal cell growth, proliferation and differentiation (source: van Krieken *et al.* (2008); figure 2)

tumour retinoblastoma, which indicated that two genetic events were required in order for a tumour to develop. It was hypothesised that in hereditary cases of retinoblastoma one copy of the gene has a germline mutation, whereas sporadic cases are much rarer because independent mutations on both copies of the gene are required. Tumour suppressor genes were subsequently shown to be involved in the development of many human tumours. Examples of tumour suppressor genes include the gene responsible for retinoblastoma and the p53 tumour suppressor gene (Box 9.2).

Box 9.2 Tumour suppressor genes

RB: the retinoblastoma gene

The RB protein is a 105 kDa (928 amino acid) protein which represses the transcription of genes required for DNA replication and cell division as well as playing a role in apoptosis. Phosphorylation is critical for its function and is mediated by cyclin and CDK complexes. The hypophosphorylated, active form of RB is found during the G0/G1 phase of the cell cycle whereas the hyperphosphorylated, inactive form dominates in late G1, just before the start of S-phase. Mutations in RB have been associated with the development of osteosarcomas, prostate and breast tumours as well as retinoblastoma, the setting in which it was first identified; functional inactivation can also occur in the absence of actual mutations.

p53

One of the purposes of cell cycle checkpoints is to allow the DNA to be checked for errors and repaired if necessary. The p53 tumour suppressor protein plays a key role in this process, hence its nickname 'Guardian of the genome'. It is involved in the decision between cell division and apoptosis and is inactivated in the hereditary cancer susceptibility condition Li–Fraumeni syndrome. The p53 gene encodes a 393 amino acid nuclear phosphoprotein which binds to DNA as a tetramer and suppresses tumour cell proliferation by inducing apoptosis, cell cycle arrest or senescence in response to stresses such as DNA damage, oncogene activation and hypoxia (Kenzelmann Broz and Attardi, 2010). In particular, it arrests the cell cycle at the entry to S phase (the G1 checkpoint) and just before metaphase (G2 checkpoint) if DNA damage is detected, allowing time for the DNA to be repaired.

The p53 protein is subject to various post-translational modifications (phosphorylation, acetylation, sumoylation and glycosylation) and its activity is modulated via protein stability. Mutations of p53, which may lead to loss of function or expression of variant proteins, are among the most commonly observed mutations in human tumours. Some variant forms of p53 can exert dominant negative effects, and they can also have so-called *gain-of-function* properties. Additional alterations which occur in tumour cells (e.g. changes in the RB pathway) can stabilise mutant p53 and contribute to acquisition of invasive and metastatic properties.

No single gene can account for the development of cancer; individual genes can be said to contribute to, but not to cause, cancer. Genes which play a key role in

the carcinogenic process include p53, p21 and RB, all of which play crucial roles in cellular protection because of the central role they play in control of the cell cycle. Each of these has a specific role: RB controls the rate of cell proliferation and promotes terminal differentiation, p53 interrupts the cell cycle at G1 to allow DNA damage to be repaired and can induce apoptosis in order to prevent the replication of damaged DNA, and p21 acts as an inhibitor of CDKs, providing a functional link between p53 and the cell cycle.

9.3.1 Methods for detecting activated oncogenes

The NIH3T3 transfection assay was the only method for detecting activated oncogenes prior to the invention of the PCR method. In this assay, high-molecular weight DNA from a tumour is transfected into the NIH3T3 cell line by calcium phosphate precipitation or lipofection. The cells are then grown up and evaluated for transformation either entirely *in vitro* (the focus assay) or by means of an *in vivo* step to detect actual tumourigenicity (the nude mouse tumourigenicity assay). If the transfected DNA contains an oncogene, the recipient cells become transformed: that is, they can form foci of morphologically transformed cells in culture or develop into tumours *in vivo*. The transforming gene in the resulting tumour or focus is identified by Southern blotting and PCR-based methods such as ASO hybridisation, RFLP and direct sequencing (Figure 9.2). However, this assay has numerous limitations: it is extremely expensive and labour-intensive, it can only detect a small subset of potential oncogenes (those of the *ras* complementation group), and it has a very high rate of false positives. In addition, it uses animals in an ethically unacceptable way (essentially just as receptacles in which to grow tumours) so it is not surprising that the advent of PCR-based methods which could identify oncogene mutations directly made the NIH3T3 assay obsolete by the early 1990s.

9.3.2 *In vitro* transformation assays

The principle underlying the NIH3T3 cell assay has, more recently, been exploited in the development of *in vitro* cell transformation assays which directly measure the transforming potential of putative carcinogens with a view to predicting 2-year bioassay positives and thus reducing animal usage and its associated welfare issues.[4] These assays can be used as a follow-up to a positive *in vitro* genotoxicity result and may be of value in terms of their contribution to a weight of evidence case regarding genotoxic carcinogenicity or the establishment of a TTC; they are also able to detect some non-genotoxic carcinogens and can contribute to the classification of carcinogens and non-carcinogens. Their value is context-dependent: it may be rather limited in the pharmaceutical industry, which really needs a definitive *in vivo* carcinogenicity result for risk assessment, but they are particularly useful in the chemical, consumer products and cosmetics industries as these try to meet the 3Rs requirements of REACH and the 7th Amendment.

[4] The history of this approach is summarised in Schechtman (2012).

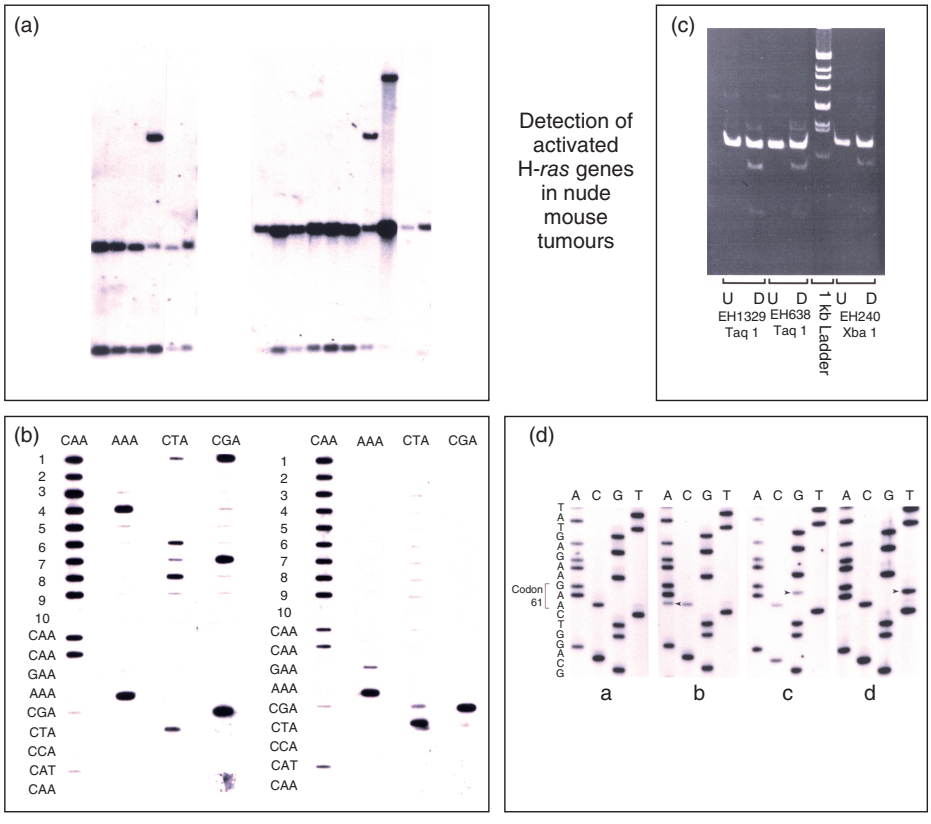


Figure 9.2 Detection of activated H-*ras* genes in nude mouse tumours. Tumours resulting from a nude mouse tumourigenicity assay using DNA from vinyl carbamate-induced mouse liver tumours were analysed for H-*ras* activation by (a) Southern blotting (b) allele-specific oligonucleotide hybridisation (c) restriction fragment length polymorphism (d) direct sequencing of PCR products(source: Stanley *et al.* (1992); figures 1, 3, 4 and 5 by permission of Oxford University Press)

The phenotypic changes evaluated in these assays reflect some stages of multistep carcinogenesis; they include immortalisation, morphologic changes (including altered morphology, basophilic staining), formation of foci (criss-cross growth patterns, multilayered growth and a generally disorganised appearance), aneuploidy, genetic instability, anchorage-independent growth and *in vivo* tumourigenicity.

In vitro rodent cell transformation assays were the subject of an OECD detailed review paper published in 2007 (#31) (Vasseur and Lasne, 2012). This evaluated three assays, the Syrian hamster embryo (SHE) cell assay (which dates back to the 1960s and uses a karyotypically normal cell type as its starting point), the BALB/c 3T3 assay and the C3H 10T$^1/_2$ assay (both of which start with immortalised aneuploid cell lines and are considered to be more relevant to the later stages of carcinogenesis). The conclusion of the detailed review was that the SHE and BALB/c 3T3 assays were suitable for the development of OECD Test Guidelines, whereas

the C3H 10T½ assay had potential utility in the elucidation of molecular mechanisms (e.g. using 'omics approaches) but was not suitable for OECD guideline purposes.

As a consequence of the OECD detailed review, ECVAM initiated a prevalidation exercise (Vanparys *et al*., 2011), the outcome of which was that the SHE and BALB/c 3T3 assays generally showed acceptable reproducibility and transferability as long as dose selection was done carefully, although the subjective nature of the scoring system was an issue. The conclusion from this exercise was that OECD Test Guidelines should be prepared for the SHE and BALB/c 3T3 assay (Corvi *et al*., 2012; Tanaka *et al*., 2012; Pant *et al*., 2012; Maire *et al*., 2012b) and recommended protocols for these assays are now available (Maire *et al*., 2012a; Sasaki *et al*., 2012). However, further work is still required in order to resolve the remaining issues concerning *in vitro* cell transformation assays (Creton *et al*., 2012). For example, the COM does not consider the SHE cell assay to be suitable for regulatory purposes because of the subjective nature of the assessment of foci, and also because the cell line used is of rodent origin. In addition, there is evidence that the number of genetic events required for the transformation of human cells is greater than for rodent cells. Furthermore, the mechanism underlying the assay is not fully understood and has not been clearly shown to be that involved in actual carcinogenesis.

9.3.3 *Ras* oncogene activation during tumour development

The oncogene family whose role in the initiation and development of both human and experimental tumours is best characterized is the *ras* family. The *ras* oncogene family comprises three genes: H-*ras* and K-*ras* are homologous to viral transforming genes, while N-*ras* was first identified in a human neuroblastoma. K-*ras* mutations are commonly observed in human tumours and mutations of the other *ras* oncogenes are also found in particular tumour types (Prior, Lewis, and Mattos, 2012).

In the mid-1980s, the idea arose that the spectrum of oncogene mutations observed in experimentally induced tumours might help to elucidate mechanisms of carcinogenesis in model systems and even make it possible to distinguish between genotoxic and non-genotoxic carcinogens *in vivo*, leading to a flurry of activity in the late 1980s and early 1990s aimed at substantiating this hypothesis.[5]

Example: Mouse skin tumours

The SENCAR mouse skin tumour system is a good model for studies on initiation, promotion and progression. Early studies in this model revealed H-*ras* activation in 7,12-dimethylbenz(a)anthracene (DMBA)-induced/12-*O*-tetradecanoyl phorbol-13-acetate (TPA)-promoted skin carcinomas. The mutation involved was a CAA–CTA mutation at codon 61. Similar results were obtained with dibenz*[c,h]*acridine and urethane but not B(a)P. In each case, the mutations observed corresponded to the known DNA-binding properties of the carcinogens.

[5] For an old, but still relevant, review of this topic see Stanley (1995).

These results suggested that mutation of the H-*ras* gene by point mutation was an initiating event in mouse skin tumourigenesis. Detection of H-*ras* mutations in benign papillomas substantiated the hypothesis that *ras* activation is an initiating event, at least in murine skin tumour development.

Initiation could also be achieved by painting a suspension of Harvey murine sarcoma virus or the T24 tumour oncogene directly onto mouse skin and was shown to be an early event, since mutations could be detected in mouse skin within a week of treatment with DMBA. In SENCAR mice, mutations can be detected in skin tumours induced by treatment with a promoter alone, suggesting that the genetic defect in this mouse strain is an initiating event.

Example: Mouse lung tumours

Strain A mice are highly susceptible to lung tumours whereas C57BL/6 mice are relatively resistant. K-*ras* mutations are seen in the majority of Strain A lung tumours. The mutations are usually those predicted from the genotoxic properties of the carcinogens used.

- methylating agents (e.g. dimethylnitrosamine (DMN)) induce tumours with mutations consistent with the formation of O^6-methylguanine.
- ethylating agents (e.g. ethylnitrosourea (ENU), diethylnitrosamine (DEN) (DEN)) induce tumours with mutations consistent with the formation of O^4-ethylthymidine.

A huge amount of information, mostly relating to the B6C3F1 mouse hybrid used by the NTP, is available regarding spectra of oncogene mutations in carcinogen-induced mouse liver tumours. Activated *ras* oncogenes, particularly H-*ras* mutants, are commonly detected in chemically induced hepatocellular tumours from B6C3F1 mice; indeed, when strong genotoxic carcinogens are used, the vast majority (up to 100%) of the resulting tumours contain carcinogen-specific H-*ras* mutations. This observation led Marshall Anderson and others to suggest that the profile of H-*ras* mutations found in B6C3F1 mouse liver tumours could be used as an indicator of the genotoxic properties of the carcinogen which induced the tumours.

Approximately 40% of B6C3F1 mice develop liver tumours spontaneously, and many of these tumours contain H-*ras* mutations. When genotoxic carcinogens are used to induce liver tumours in this strain, up to 100% of them contain H-*ras* mutations. This observation has been exploited in studies aimed at understanding the factors that determine strain differences in susceptibility to hepatocarcinogenesis.

Strain differences in susceptibility to hepatocarcinogenesis Mutations can be detected in DEN-induced B6C3F1 liver lesions 11 weeks following carcinogen treatment, suggesting that H-*ras* activation and/or propensity for clonal expansion of H-*ras*-containing cells may be involved in susceptibility. In contrast, *ras* activation is a rare event in liver tumours from resistant mice such as C57BL/10, usually occurring in only ~10% of tumours even when a potent genotoxic carcinogen is used.

Table 9.1 4-Aminobiphenyl-induced mutation fraction in sensitive and resistant mouse liver

Mouse strain	Treatment	Mutant fraction	Increase in initiated mice
B6C3F1	DMSO	1.0×10^{-5}	34.5-fold
	4-ABP	44.9×10^{-5}	
C57BL/6N	DMSO	1.4×10^{-5}	5-fold
	4-ABP	7.0×10^{-5}	

Adapted from Parsons *et al.* (2005). Reproduced with permission of John Wiley & Sons, Inc.

In a subsequent study (Parsons *et al.*, 2005), measurement of the mutant fraction of H-*ras* codon 61 CAA-AAA mutations in control and 4-aminobiphenyl (4-ABP)-initiated mouse liver from sensitive (B6C3F1) and resistant (C57BL/6N) strains (Table 9.1) suggested that initiated cells exist in sensitive and resistant strains even without carcinogen treatment. These results are consistent with the idea that the low incidence of spontaneous lesions in the resistant strain is due to resistance to initiation. The resistant strain may also be more refractory to promotion than the sensitive strain.

This contrasts with results obtained using the BigBlue® system: the strain background of the mice used in the Big Blue® assay does not seem to have any effect on the *lac*I mutant frequencies detected. When the assay was conducted in two different laboratories, each using both C57BL/6 (liver tumour resistant) and B6C3F1 (sensitive) mice, the results obtained were amazingly consistent. In each case the mutant frequency detected was about 4×10^{-5} (Young *et al.*, 1995; Lemieux *et al.*, 2011). These results suggest that strain differences in *ras* mutation frequency may reflect the fact that the H-*ras* gene has a functional role in hepatocarcinogenesis whereas the *lac*I transgene in the Big Blue mouse is phenotypically neutral.

The extent of H-*ras* activation in mouse liver tumours is affected by the dose of carcinogen as well as the strain of mouse used. For example, when a high dose of DEN is used to induce tumours in B6C3F1 mice, the spectrum of mutations seen is similar to that in spontaneous tumours. When a lower dose is used, however, a more carcinogen-specific pattern is observed. The results suggest that high doses of carcinogen may be causing cytotoxicity, thus promoting spontaneous lesions, whereas at lower doses a classical genotoxic mechanism operates.

9.3.4 *Non-ras* oncogenes

Ras genes were easy to study in the early days because they are highly active in the NIH3T3 transfection assay and are activated by point mutations which can be detected by PCR. With recent advances in detection methods, mutagenic activation of other oncogenes has become easier to identify and this has provided preliminary evidence for the possibility that oncogenes other than *ras* play a role in tumour development in resistant strains of mouse. For example, in a study analysing codon 637 mutations in the B-*raf* oncogene (another component of the *ras* pathway), activation of B-*raf* was shown to predominate in C57BL/6J mouse liver tumours induced by DEN ($20\,mg\,kg^{-1}$) or DMBA ($20\,\mu mol\,kg^{-1}$), whereas H-*ras* activation predominated in C3H/He and B6C3F1 tumours (Buchmann *et al.*, 2008). Furthermore, B-*raf* mutations were observed in 15/17 (88.2%)

preneoplastic lesions in C57BL/6J liver, suggesting that B-*raf* activation was an early (possibly initiating) event in this strain.

9.3.5 Evaluation

Approximately 40% of B6C3F1 mice develop liver tumours spontaneously, and many of these tumours contain H-*ras* mutations. When genotoxic carcinogens are used to induce liver tumours in this strain, the vast majority of them contain H-*ras* mutations; however, *ras* mutations are rarely detected in rat liver tumours. The only exception is seen in tumours induced by AFB1, which induces tumours containing K- and N-*ras* genes mutated at codons 12 and 13. The significance of this is difficult to assess since the mouse does not develop liver tumours in response to AFB1. Hepatocellular tumours in humans are rare in the Western world, although they are common in populations exposed to AFB1 (e.g. through eating contaminated peanuts), and practically never contain activated *ras* genes. This raises the question: Can data on *ras* oncogene activation in rodent liver tumours help us to understand the effects of human carcinogens? In this context it is interesting to note that there are marked similarities in global gene expression profiles between spontaneous B6C3F1 mouse liver tumours and human HCC, suggesting that findings in the NTP mouse model may have relevance to human hazard after all (Hoenerhoff *et al*., 2011).

Ras oncogenes clearly play a role in experimental tumour development, especially with regard to initiation in the mouse skin and rat mammary tumour models. In the mouse liver and lung models, a strong relationship has been demonstrated between DNA adduct formation, specific mutations and the development of tumours in response to strong genotoxic carcinogens. However, the results are less clear cut for weakly and non-genotoxic carcinogen-induced tumours.

9.4 Non-genotoxic carcinogenesis

The separation of carcinogenesis into the stages of initiation, promotion and progression has led to various schemes for distinguishing carcinogens according to which stage of the process they affect. Such distinctions include separation of complete vs incomplete carcinogens (complete carcinogens can induce tumours without any additional treatment, whereas incomplete carcinogens are only effective when co-administered with another substance). Classical tumour promoters are considered to be insufficient to induce tumours on their own, but may drive tumour development; their role in human tumours and the associated carcinogenic risk to humans is still the subject of debate.

The conventional classification of carcinogens, however, divides them into two classes, genotoxic and non-genotoxic (Figure 9.3). Non-genotoxic carcinogens and indirect genotoxicants do not damage DNA directly and may modify or interact with multiple and/or partially redundant molecular targets. They are negative in mutagenicity assays *in vitro* and *in vivo*, do not form DNA adducts and do not induce DNA repair. Non-genotoxic carcinogens can act via multiple mechanisms

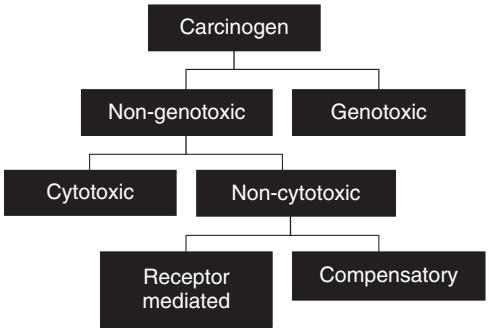


Figure 9.3 Classification of carcinogens (source: Dr Cliff Elcombe. Reproduced with permission of Dr Cliff Elcombe, CXR Biosciences Ltd)

(Hernandez *et al*., 2009): they may be cytotoxic (leading to regenerative hyperplasia) or mitogenic (inducing augmentative hyperplasia) and their effects may or may not be receptor-mediated.

9.4.1 Non-receptor-mediated mechanisms

Tissue-specific toxicity can play an important role in non-genotoxic carcinogenic processes.

Example: Saccharin-induced bladder tumours in male rats

One of the best characterised examples of this type of mechanism is the induction of bladder tumours by saccharin in rats. Lifetime dietary exposure to saccharin induces a higher level of bladder tumours in male than female rats; mice are resistant to this effect. The mechanism involved is precipitation of calcium phosphate-containing crystals from the urine, causing irritation and cytotoxicity to the urothelium and leading to regenerative hyperplasia. The increased incidence in male rats is due to the presence in their urine of the protein α_{2u}-microglobulin, which exacerbates the precipitation of these crystals. No bladder tumours are seen in mice because the concentrations of calcium and phosphate are lower in their urine, and primates (including humans) are not at risk because the osmolality of their urine is much lower than that in rodents, so that this mechanism of tumour induction is not considered to be relevant to humans.

It is interesting to note that a related mechanism of tumour induction occurs in rats treated with melamine, which produces urinary calculi and bladder tumours in rats. This has led to possible concern for the survivors of the population of children exposed to melamine via infant formula in China in 2007/2008.[6] At the time this exposure led to urinary obstruction, renal injury and hydronephrosis, and there is concern that urinary or renal tumours could also result although it is hoped that this will not be the case because of the transient nature of the exposure. The final outcome will depend upon the threshold for tumour induction, which has never been characterised in humans.

[6] http://www.telegraph.co.uk/news/worldnews/asia/china/4315627/Two-sentenced-to-death-over-China-melamine-milk-scandal.html

Chemically induced alterations in key metabolic pathways may also lead to tumour development

Example: Induction of thyroid tumours

Anti-thyroid drugs such as 6-methyl-2-thiouracil, 6-propyl-2-thiouracil and thiourea increase the incidence of thyroid tumours in rodents by disrupting the biosynthesis and/or secretion of thyroid hormones. In rats, goitrogenic agents and enzyme inducers increase the peripheral metabolism of thyroid hormones by inducing hepatic metabolising enzymes (e.g. UGTs), thus increasing thyroid-stimulating hormone levels in an attempt to compensate and causing a direct stimulus to thyroid follicular cells leading to the development of benign and malignant tumours. Levels of thyroid hormone can also be disrupted by xenobiotics that inhibit 5′-monodeiodinase, the enzyme which converts T4 to biologically inactive T3. In rats, this leads to up-regulation of thyroid-stimulating hormone expression, follicular cell hypertrophy and hyperplasia and ultimately to the development of thyroid follicular cell tumours. Humans are not subject to this effect, though, because they have a circulating thyroid-binding globulin which controls the availability of free thyroid hormones. In humans, hypothyroidism does not lead to follicular cell proliferation despite increased levels of thyroid-stimulating hormone and there is no evidence for chemical induction of thyroid tumours in humans.

9.4.2 Receptor-mediated mechanisms

Example: PPARα mediated hepatocarcinogenesis

Peroxisome proliferator receptor α (PPARα), so named because its activation induces a spectacular increase in the number of peroxisomes within hepatocytes, mediates the pleiotrophic response to a class of chemicals, the so-called peroxisome proliferators, which include phthalate plasticisers such as DEHP and fibrate lipid-lowering drugs such as clofibrate and ciprofibrate.[7] It plays a key homeostatic role in that it regulates lipid homeostasis, but there are important issues in relation to PPARα because there is evidence that it may be responsible for peroxisome proliferator-induced carcinogenesis, at least in rodent liver. The relevance of this phenomenon, and the role therein of PPARα, has been the subject of intense debate for many years because human cells appear to be resistant to the *in vitro* effects of PPARα ligands compared with their rodent equivalents. The points in the proposed carcinogenic mechanism at which this resistance may occur are highlighted in Figure 9.4.

In an attempt to resolve this issue, a humanised model which expresses human PPARα on a murine PPARα null background has been generated (Cheung *et al.*, 2004). The transgene used contained a human PPARα cDNA under the control of a tetracycline-responsive regulatory sequence. This construct allowed the expression of hPPARα to be controlled by the administration or withdrawal of the drug doxocycline. These mice exhibited decreased serum triglycerides and increased expression of genes encoding peroxisomal, mitochondrial and

[7] For more on PPARα and the related receptors PPARγ and PPARδ, see Wahli and Michalik (2012).

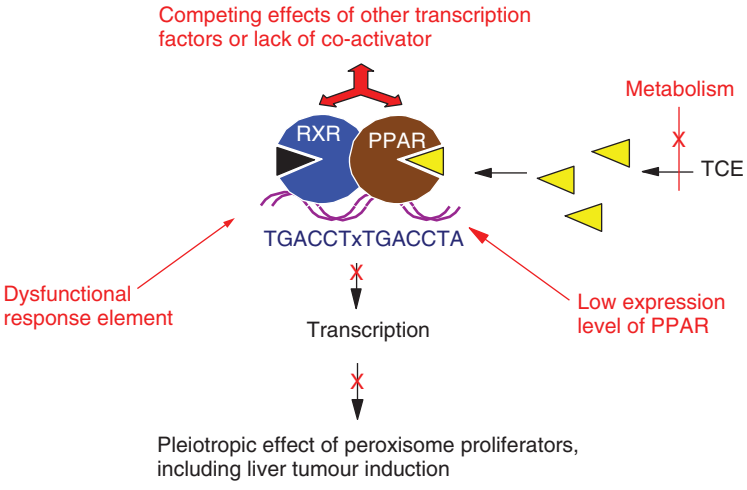


Figure 9.4 PPARα-mediated carcinogenesis. Steps which might cause humans to have reduced susceptibility compared with rodents are highlighted in red (source: Reproduced with permission of Dr Cliff Elcombe, CXR Biosciences Ltd)

microsomal fatty-acid-oxidising enzymes compared with PPARα nulls, although the effect of the human PPARα was less marked than that of murine PPARα. However, they failed to respond to WY-14643 with increases in the expression of cell cycle control genes or replicative DNA synthesis (as measured by BrdU labelling). The results obtained suggested that human PPARα is relatively ineffective in activating the genes required for hepatic cell proliferation as compared with murine PPARα. Consistent with these observations, further studies with this model demonstrated that low doses of the peroxisome proliferator ammonium perfluorooctanoate (0.1 and 0.3 $mgkg^{-1}$/day for 2 weeks) significantly activated PPARα responses and increased hepatic triglyceride and serum cholesterol levels in wild type but not humanised mice (Nakamura *et al.*, 2009). The expression of PPARα in the humanised mice was at least as high as that in wild-type animals, consistent with the idea that human PPARα has reduced functionality compared with murine PPARα. This was in line with the contention that peroxisome-proliferator-induced hepatocarcinogenesis is a rodent-specific effect and peroxisome proliferators therefore do not constitute a risk to humans.

One of the limitations of the original PPARα humanised mouse model was that it was created using a cDNA, so it lacked the human regulatory sequences. A new, improved model was subsequently created using the complete genomic human PPARα sequence (on a P1 phage artificial chromosome) on a PPARα SV129 null background (Yang *et al.*, 2008). This mouse line can be used to examine extrahepatic PPARα-mediated effects since it expresses PPARα in various tissues (including heart and kidney) as well as the liver, and it has the advantage that the expression of PPARα is controlled by its own (human) promoter. Examination of fatty acid metabolism in this model revealed that it has reduced serum triglyceride

levels without altered expression of lipoprotein lipase and apolipoproteinCIII and exhibits robust induction of genes encoding the enzymes involved in fatty acid oxidation. Use of this line allowed examination of the question of whether human PPARα is able to mediate proliferative responses to peroxisome proliferators: hPPARα activation-induced peroxisome proliferation, which was associated with slight hepatomegaly due to hypertrophy (essentially cellular enlargement to accommodate the extra peroxisomes), but not up-regulation of cell cycle control genes (cyclin D1, CDK4) was observed. This provides further support for the concept that the effects of peroxisome proliferators on peroxisome proliferation and lipid metabolism (seen in both human and rodent systems) are distinct from those involved in liver enlargement and hepatocyte proliferation. This may be solely due to differences in PPARα itself between humans and rodents, but it could also involve generic differences between human and rodent hepatocytes (e.g. in the pattern of expression of coregulatory proteins which might modulate PPARα-mediated effects).

These results are very interesting, but the hPPARα model used is self-evidently a highly artificial system in which to study species differences in carcinogenesis. More physiological models are now becoming available and will help to clarify whether the conclusions drawn from existing models have general applicability in relation to the risk assessment of peroxisome proliferating chemicals.

9.4.3 When is a genotoxic carcinogen not a genotoxic carcinogen?

The vexed question of whether non-genotoxic carcinogens represent the same risk to human health as do genotoxic carcinogens remains to be resolved and anyway, almost every chemical can be shown to be genotoxic if enough tests are conducted using high enough concentrations/doses. In addition, techniques such as ^{32}P-postlabelling can be highly subjective. Furthermore, it can be very difficult to control for/take into account concomitant exposures and the fact that the development of neoplasia is known to induce genomic instability may also lead to the detection of mutations even in tumours induced by non-genotoxic agents.

H-*ras* mutations are observed in spontaneous mouse liver tumours and those induced by non-genotoxic carcinogens as well as in genotoxic carcinogen-induced tumours. The frequency of H-*ras* codon 61 mutations in tumours induced by non-genotoxic carcinogens is usually lower than that in spontaneous or genotoxic carcinogen-induced tumours, but novel mutations (e.g. at codon 117) have been identified in B6C3F1 mouse liver tumours induced by well-characterised non-genotoxic carcinogens such as the peroxisome proliferators methylclofenapate and clofibrate, which are believed to be non-genotoxic. This leads to the question: Does this mean that these peroxisome proliferators are actually genotoxic?

Figure 9.5 illustrates an alternative mechanism by which tumours with unusual mutations such as codon 117 mutations could arise without the necessity of genotoxicity. This model is based on the evidence that initiated cells are present in mouse liver even in the absence of a genotoxic insult, and it proposes that different

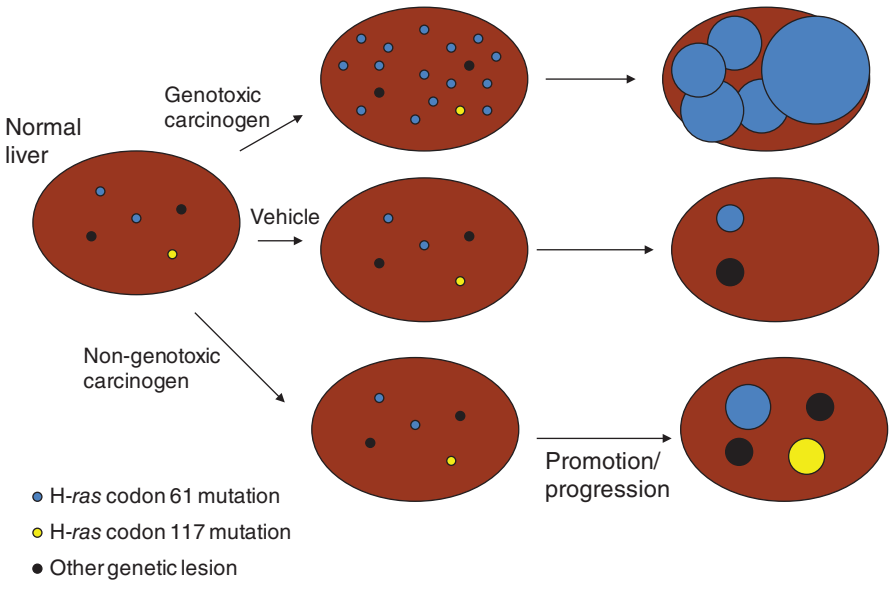


Figure 9.5 An alternative mechanism of action for some non-genotoxic carcinogens

types of initiated cells with different molecular lesions exist even prior to treatment with a carcinogen. In Figure 9.5:

- blue indicates tumours containing H-*ras* codon 61 mutation
- yellow indicates tumours carrying unusual H-*ras* mutations, for example, at codon 117
- black indicates tumours which have mutations of other types

In this mechanism it is assumed that initiated cells containing various types of mutations are present in, for example, the mouse liver even in the absence of chemical treatment. Treatment with a strong genotoxic carcinogen will initiate additional cells carrying by inducing specific mutations such as those at codon 61 (blue), and these go on to form tumours carrying corresponding mutations. Even in the absence of carcinogen treatment, a small proportion of spontaneously initiated cells will go on to form tumours containing various mutations (blue, black). Other initiated cells (yellow) remain dormant and fail to form tumours spontaneously or in genotoxic carcinogen-treated liver. Treatment with a non-genotoxic carcinogen, however, may promote out this class of initiated cells which would otherwise not go on to form tumours, thus generating tumours with novel mutations which are not seen in spontaneous tumours or those induced by genotoxic mechanisms (yellow).

This mechanism is feasible because it is known that different promoters can promote different classes of initiated lesions (Pitot, 2007), and its consequence is that tumours containing novel H-*ras* mutations can be observed following treatment with a non-genotoxic carcinogen, not as a consequence of carcinogen-induced mutational damage but as a result of preferential promotion of a rare subset of initiated cells.

9.5 Transgenic models for short-term carcinogenicity bioassays

The conventional regulatory approach to carcinogen testing, in which lifetime bioassays are conducted in two species (usually rats and mice), is unsatisfactory for a number of reasons. This approach is believed to detect all human carcinogens, but as well as being unwieldy, time consuming and expensive, it has low accuracy (69%) due to a very high incidence of false positives. In the 1990s, these issues, combined with improved understanding of the roles of oncogenes and tumour suppressor genes in tumour development, led to an intensive effort to develop novel assays using transgenic strains containing defects in key pathways of carcinogenesis.[8] The advantages of these assays are that they use fewer animals and can provide mechanistic information as well as a yes/no answer as to carcinogenicity. In addition, the tumours develop rapidly, so the assays are shorter (24–26 weeks) than the conventional lifetime bioassay. In addition, each of the assays has its own individual strengths and limitations.

9.5.1 *RasH2*

The rationale for the development of the *rasH2* model[9] (official designation: CB6F1-TgN(RasH2)) was that integration of the c-Ha-*ras* oncogene would provide additional copies of a relevant genetic target for codon-specific mutational activation. The line carries several copies of the human c-Ha-*ras* oncogene integrated in tandem array and expressed under the control of its own promoter. The parental strain has been backcrossed to a C57BL/6 background for more than 21 generations, yielding a genetically stable line. Male transgenic *rasH2* mice (Tg C57BL/6J) are crossed with inbred BALB/cByJ females to produce F1 *rasH2* mice (CB6F1-Tg *rasH2*). Individual animals are tested for the transgene by PCR at 3 weeks of age to differentiate between transgenic (Tg) and non-transgenic (non-Tg) littermates. The homozygous transgenic genotype is lethal during embryogenesis and thus these pups are eliminated from the population.

The *rasH2* transgene is expressed in normal tissues at 2-3 times the levels observed in wild-type tissues but the mice develop few tumours before 33 weeks of age. At around 8–9 months of age, the mice start to develop spontaneous tumours and ~50% of mice have such tumours by 18 months of age. The tissues affected at age 8–9 months are the lung (bronchoalveolar lung adenomas and carcinomas, incidence 7.4%), spleen (haemangiomas and haemangiosarcomas, 5.4%), forestomach (squamous cell papillomas and carcinomas, 2.4%) and skin (1.2%). The mice also exhibit some non-neoplastic pathology including a high incidence of degenerative skeletal muscle myopathy (>80%) and occasional vascular anomalies (<1%).

The strengths of this model are that it can detect both genotoxic and non-genotoxic compounds, any conventional route of administration can be used

[8] For detailed reviews of this topic, see Jacobson-Kram, Sistare, and Jacobs (2004), Boverhof *et al.* (2011) and Taneja *et al.* (2011).
[9] http://www.ciea.or.jp/rash/rash2mice.htm

Table 9.2 Comparison of the *rasH2* mouse line with conventional methods for detection of carcinogens

Compound	Ames test	Two-year bioassay	*rasH2*
Benzene	– (positive in chromosome aberration test)	+	+
Cyclophosphamide	+	+	+
Cyclosporin A	–	–/equivocal	–/equivocal
Diethylstilbestrol	–	Equivocal	+
17α-estradiol	–	+ (mouse mammary tumours)	–
Phenacetin	+	+	+

Adapted from Morton *et al.* (2002). Reproduced with permission of the Society of Toxicologic Pathology.

and the model responds robustly to positive control compounds (e.g. MNU, urethane). A comparison of results obtained using the *rasH2* model with historical results from the Ames test and the 2-year bioassay yielded the results shown in Table 9.2 (Morton *et al.*, 2002). These results indicate that the *rasH2* model is able to detect genotoxic, and some non-genotoxic, carcinogens but is limited with respect to compounds which act via other mechanisms (e.g. immunosuppressive agents and hormonal carcinogens). Unexpectedly, the majority of tumours induced by genotoxic carcinogens in the *rasH2* mouse line are found in the lung, spleen, forestomach and skin rather than in the liver.

It was initially surprising that introduction of an H-*ras* gene (which is normally associated with the development of liver tumours in mice) actually generated a model with a high incidence of lung tumours. However, Japanese workers (Umemura *et al.*, 2002) have capitalised on this observation by using the *rasH2* model to develop an *in vivo* bioassay for the detection of pulmonary carcinogens. This assay involves a single treatment with the putative carcinogen followed after a week's break by promotion using butylated hydroxytoluene (40 mg kg^{-1} orally, five doses per week for 9 weeks) and has been shown to give positive results with a number of known carcinogens: DMN (15 mg kg^{-1}), DEN (100 mg kg^{-1}), ENU (120 mg kg^{-1}), 3-MC (100 mg kg^{-1}), DMBA (5 mg kg^{-1}) and B(a)P (80 mg kg^{-1}) (Umemura *et al.*, 2006).

Studies attempting to verify the hypothesis that the c-Ha-*ras* transgene in the *rasH2* mouse is a target for mutational activation have yielded some positive results. When genotoxic carcinogen-induced tumours from this line were investigated for mutations at codon 12 and codon 61, the results with some compounds supported the hypothesis that the human c-Ha-*ras* transgene was a mutational target, but it appeared that other carcinogens were acting by a different mechanism, possibly involving cooperation between the overexpressed transgene and endogenous murine oncogenes and tumour suppressor genes.

The *rasH2* mouse line is very susceptible to genotoxic carcinogens but this line actually has reduced susceptibility to some compounds (Mitsumori, 2003). In 6-month studies, *rasH2* mice have reduced sensitivity to phenolphthalein compared with the results in B6C3F1 (2-year assay) and $p53^{+/-}$ (6-month

assay) models. Furthermore, ethinylestradiol promotes ENU-induced endometrial carcinogenesis in wild-type mice but actually inhibits it in *rasH2* mice. These findings, which have not been fully explained, raise doubts concerning the usefulness of the *rasH2* line for testing novel compounds and should be borne in mind when interpreting the results obtained. However, on the other hand, it should be noted that none of the conventional *in vivo* bioassays gives 100% predicted results either.

Overall, despite these limitations, this model seems to identify known and predicted human carcinogens as well as conventional assays but with fewer false positives, and it is now specifically mentioned as an acceptable test in international guidelines. The recommended protocol involves studies of 6 months' duration with groups of 20–25 mice and the preferred positive control is MNU (75 mg kg^{-1}), which induces forestomach papillomas within 8 weeks and thymic neoplasms within 12 weeks leading to death by week 20 (Morton *et al.*, 2002). Indeed, it may become possible to reduce the study duration even further by taking advantage of toxicogenomic biomarkers, since common and significant gene expression changes due to 2-AAF and DEN have been identified in *rasH2* mice as early as 7 days after the start of daily dosing and are sustained for 13 weeks (Park *et al.*, 2011).

9.5.2 Tg.AC

The Tg.AC mouse line[10] contains ~40 copies and one tandem repeat of a v-Ha-*ras* oncogene carrying mutations at codon 12 and codon 61 on chromosome 11 of a mouse of FVB genetic background. The expression of the transgene is regulated by the ζ-globin promoter, meaning that it is normally not expressed in adult tissues except for the bone marrow. The skin of this mouse model behaves as if it has already been treated with a genotoxic initiating agent in that the papillomas develop rapidly in response to stimuli which induce cell proliferation (e.g. at sites of wounding/abrasion or in response to tumour promoters such as TPA).

For assay purposes, the positive control promoter used is TPA (three dermal treatments weekly) which leads to papilloma development within 12–14 weeks. This model is limited in that it was designed for dermal application of the test compound (although positive results have been reported following oral dosing) and it has been criticised on the grounds that the mechanism of papilloma induction is poorly understood.

Another issue which arose during the early stages of evaluation of the Tg.AC line is its genetic instability, which gives rise to a non-responsive phenotype (i.e. no tumours in response to TPA). The transgenic locus in this line has a high background mutation rate (1.1% in females and 3.3% in males) and the number of copies of the transgene increased from 4 to 40 over the first few years of breeding. However, these issues have now, to some extent, been resolved and rigorous QC by the supplier ensures that populations of mice supplied for carcinogenicity testing are now >95% TPA responders.[11]

[10] http://www.taconic.com/wmspage.cfm?parm1=905

[11] For further information, see http://www.taconic.com/wmspage.cfm?parm1=308..

9.5.3 p53 models

Loss of tumour suppressor genes is a key event in human tumour development, and great effort has therefore been devoted to the attempt to detect tumour suppressor genes in animal tumours. Tumour suppressor genes are subject to various types of genetic alteration, including point mutations, loss of heterozygosity and chromosome aberrations/rearrangements during tumour development. Depending on their nature, these changes can be detected by PCR-based methods, loss of heterozygosity, karyotyping or fluorescent *in situ* hybridisation. Most of this work has involved the p53 gene, mutation of which is a very common genetic event in human tumours.

Despite all this effort, very few p53 mutations have been detected in liver and lung tumours from wild-type rats and mice treated with various carcinogens, the only convincingly positive result being in a study of liver tumour induction by 3′-methyl-4-dimethylaminoazobenzene in F344 rats. Despite assiduous searching, very little evidence has been obtained to indicate that p53 plays a role in carcinogenesis in the mouse, although in one SENCAR mouse study, p53 mutations were detected in 1/8 skin papillomas and 6/12 carcinomas, suggesting that p53 may be involved in progression in this model. However, while subsequent studies confirmed that murine skin tumours induced by UV light or PAHs do often contain p53 mutations (~40% incidence), such mutations are still very rarely detected in murine liver and lung tumours.

Although the data concerning activation of the endogenous murine p53 gene remains inconclusive, the p53 gene has proved useful in the context of creating a transgenic mouse model for carcinogenicity testing. A p53 null rat model has also recently become available (Tong *et al*., 2010).

The $p53^{+/-}$ mouse model Mice in which the p53 gene has been inactivated by homologous recombination and transgenic mice with germline mutations in p53 develop normally but are susceptible to lymphomas and sarcomas.

- Homozygotes (−/−) die within 10 months of birth whereas heterozygotes (+/−) can live for at least 18 months although they develop tumours well before this time.
- p53 +/− mice are highly susceptible to DEN-induced haemangiosarcoma development, but the −/− mice die of other causes before the carcinogen takes effect.

The $p53^{+/-}$ (B6.129-Trp53tm1Brd) mouse line (Donehower *et al*., 1992) carries a targeted inactivating deletion of intron 4 and exon 5 of the p53 gene and is used in the hemizygous state so that it contains only one functional copy of p53. When bred onto a C57BL/6 background it has a low spontaneous tumour incidence for the first 6–9 months of life although in old age about half the animals develop spontaneous tumours (lymphomas, osteosarcomas and haemangiosarcomas). The spontaneous tumours observed around 7 months of age include malignant lymphomas (1.7% in males, 2.9% in females) and subcutaneous sarcomas

(0.4%/1.5%) (Floyd *et al*., 2002). Interestingly, the incidence of subcutaneous sarcomas increases markedly in studies in which mice are identified by means of electronic transponders, suggesting that physical factors can influence carcinogenesis in this line.

The low incidence of spontaneous tumours in $p53^{+/-}$ mice between 7 weeks and 9 months of age allows treatment-induced and spontaneous tumours to be distinguished, at least in short-term bioassays. In contrast, completely p53 deficient ($p53^{-/-}$) mice are less useful because they develop tumours very rapidly. It is assumed that the mechanism of tumour induction in hemizygous mice is loss of the remaining functional copy of p53. However, this may not be the case for all carcinogens: there is some evidence that accelerated tumourigenesis in $p53^{+/-}$ mice may be a function of gene dosage and haploinsufficiency meaning that a second loss of heterozygosity event is not required.

The utility of this line is thought to be primarily in detecting genotoxic carcinogens; early examples tested were benzene, cyclophosphamide and *p*-cresidine (which has been used as a positive control for the $p53^{+/-}$ bioassay). The model has limited utility for the detection of carcinogens which act via epigenetic mechanisms, at least within a 6-month assay protocol.

In order for a short-term bioassay to be valid, it is essential that an appropriate control which is easy to administer and consistently yields positive results is available. The positive control used in the $p53^{+/-}$ mouse assay, NMU (75 mg kg^{-1}), is administered as a single intraperitoneal dose at about 6–8 weeks of age and the mice are then observed for 6 months (Morton *et al*., 2002). This treatment induces thymic lymphoma with a high frequency (80–90%). The advantages of MNU as a positive control are that it only involves administration of a single intraperitoneal dose (hence reducing the exposure of animal staff to the carcinogen as well as simplifying the assay), it consistently yields a specific treatment-related tumour type and the high incidence of tumours observed means that a small group size (e.g. 10/group/sex) can be used.

Unlike the *rasH2* model, the $p53^{+/-}$ model may have value in the detection of endocrine disrupting chemicals. An initiation-promotion protocol has been developed for this purpose using initiation with ENU (120 mg kg^{-1} i.p.) followed by promotion with the putative endocrine disruptor for 26 weeks and evaluation of the incidence of uterine tumours (Mitsumori, 2002).

Humanisation of p53 The $p53^{+/-}$ mouse model has proved itself to be of value in short-term carcinogenicity testing. However, the mechanistic role of p53 in murine carcinogenesis is problematic because of the difference in incidence of p53 mutations between human and murine tumours. This leads to problems in interpreting the results obtained in either wild-type or $p53^{+/-}$ mice. Furthermore, sequence differences between the human and murine p53 genes make it difficult to interpret any mutations which are observed in mouse tumours.

In order to address this problem, a mouse line humanised for the DNA-binding domain of p53 has been generated (Zielinski *et al*., 2002). In this humanised model (formal name $Trp53^{tm/Holl}$; referred to as the *Hupki* mouse), exons 4–9 of the endogenous murine p53 gene were replaced with the corresponding human

sequences. The substituted segment encodes the polyproline and DNA-binding domains of normal human p53 (with the key amino acid resides Ser46, Pro72, Thr76 and Thr86). The chimeric p53 gene remains under normal transcriptional control via the surrounding mouse regulatory sequences. The line was originally generated on a C57BL/6 background and subsequently backcrossed to C3H yielding a line which would be predicted to be susceptible to hepatocarcinogenesis. The mice develop normally, exhibit no apparent physiological defects, are fertile and show no increase in the incidence of spontaneous lymphoma, sarcoma or other neoplasms compared with wild type mice (Besaratinia and Pfeifer, 2010). The incidence of liver tumours in C3H Hupki mice 35 weeks after treatment with DEN (10 mg kg^{-1} i.p at 2 weeks of age) was comparable to that in wild-type mice (Jaworski *et al*., 2005). The resulting tumours had a 40–50% incidence of H-*ras* mutations but no p53 mutations, consistent with expectations for a murine liver tumour-susceptible model.

The Hupki model has utility *in vitro* as well as *in vivo*; for example, an *in vitro* p53 mutation assay using Hupki embryonic fibroblasts has been reported (vom Brocke *et al*., 2009). These are reported to be metabolically competent (Besaratinia and Pfeifer, 2010), although the range of enzymes they have been shown to express is quite limited (Cyp1b1, Cyp1a2, microsomal epoxide hydrolase and NADPH:quinone oxidoreductase).

A number of questions still remain regarding the Hupki mouse: it is unclear whether the human p53 construct can perform optimally (e.g. its ability to participate normally in DNA repair and recombination) in the transgenic mouse environment, and the chimeric p53 protein may undergo a different pattern of post-translational modifications (e.g. phosphorylation) compared with the normal human or mouse proteins. Furthermore, appropriate DNA methylation has not been demonstrated in the transgene; this may be important because methyl CpG-containing codons may be preferential mutation targets in human p53.

9.5.4 $XPC^{-/-}$, $XPA^{-/-}$ and $XPA^{-/-}/p53^{+/-}$ null mouse models

The $XPC^{-/-}$, $XPA^{-/-}$ and $XPA^{-/-}/p53^{+/-}$ null strains, which have defects in DNA repair homologous to those observed in human xeroderma pigmentosum patients, are maintained on a C57BL/6 background. The $XPA^{-/-}$ models carry a defect which affects both global genomic repair and transcription-coupled repair whereas the defect in the $XPC^{-/-}$ model is specific to global genomic repair. These models respond in a satisfactory manner to human carcinogens and genotoxic rodent/human carcinogens but they are less fully developed than the p53 and H-*ras*-based models described previously and are not yet used routinely in regulatory testing.

9.5.5 Comparative evaluation of models

During the late 1990s, the Alternatives to Carcinogenicity Testing Committee of the International Life Sciences Institute (ILSI) initiated a large-scale multinational collaborative programme to evaluate short-term *in vivo* carcinogenicity

bioassays which use transgenic (and other novel) models (MacDonald *et al*., 2004; Alden, Smith, and Morton, 2002; Cohen, 2001; Cohen, Robinson, and MacDonald, 2001). The models tested were as follows:

- *rasH2* transgenic mouse model (Central Institute for Experimental Animals, Japan);
- Tg.AC transgenic mouse model (Taconic Farms);
- $p53^{+/-}$ null mouse model (Taconic Farms);
- $XPA^{-/-}$ and $XPA^{-/-}/p53^{+/-}$ null mouse models (RIVM);
- neonatal mouse assay (CD-1 or B6C3F1 mice);
- SHE assay.

The protocol used involved groups of 15 mice/sex/group dosed daily (7 times weekly, oral (plus dermal for Tg.AC)) for 26 weeks. Three dose levels per compound plus vehicle were administered. Subsequent evaluation of the results obtained indicated that this group size was, in some cases, too small to provide sufficient statistical power and the recommended group size for assays using transgenic models is now 20–25.

The ILSI collaboration evaluated 21 chemicals comprising a range of genotoxic, non-genotoxic, carcinogenic and non-carcinogenic chemicals. The criteria for compound selection were that a broad range of mechanisms should be represented, each compound should have a comprehensive existing database, and that the compounds should be non-proprietary and readily available. Several compounds which cause tumours in rodents but are not believed to be human carcinogens were included.

The key results are summarised in Table 9.3. These results indicated that the $p53^{+/-}$, *rasH2*, $XPA^{-/-}$ and $XPA^{-/-}/p53^{+/-}$ deliver fewer positive results than do conventional assays, but this may be a function of the shorter duration of the assays using transgenic lines (6 vs 24 months).

A comparison of these studies and other reported studies using the *rasH2*, Tg.Ac and $p53^{+/-}$ models with the results of conventional 2-year bioassays (Pritchard *et al*., 2003) indicated that individual transgenic models made the correct prediction for 74–81% of the 99 chemicals considered, whereas the corresponding value for the 2-year bioassay was 69%. However, the transgenic assays did miss some known carcinogens whereas the 2-year bioassay did not miss any. A combined approach using the 2-year rat bioassay together with data from transgenic models yielded ~85% correct determinations and no false negatives.

9.5.6 Regulatory status

The genetically modified mouse models most commonly used for risk assessment today are the $p53^{+/-}$, *rasH2* and Tg.Ac models. The views of the FDA, the European Committee for Proprietary Medical Products (CPMP) Safety Working Party and the Japanese Ministry of Health, Labour and Welfare (MHLW) on these models, as summarised by MacDonald et al. (2004) are as follows.

- **The FDA position** is that the $p53^{+/-}$ assay is considered appropriate for clearly/equivocally genotoxic drug candidates while the *rasH2* model may be

used for either genotoxic or non-genotoxic agents. It notes that the Tg.AC assay is generally used for testing dermally applied products, although it acknowledges that some drugs intended for systemic administration have been tested in this system using the dermal route. In line with this FDA view, the $p53^{+/-}$ model is now the most commonly used short-term *in vivo* bioassay using a transgenic model in FDA submissions.

- **The CPMP Safety Working Party** expressed the view that the $p53^{+/-}$ and *rasH2* models are acceptable for regulatory use but are not definitive. In contrast, because the Tg.AC model reacts inconsistently and incompletely to some known human carcinogens, its utility for regulatory purposes is limited to the testing of dermally applied products.
- **The MHLW** opined that the $p53^{+/-}$ model is the most responsive to complete carcinogens and the *rasH2* model responds well to both genotoxic and nongenotoxic carcinogens but the Tg.AC model is of restricted use due to its limited phenotype.

The key areas of agreement between all three agencies were that the *rasH2* model is suitable for the short-term *in vivo* testing of both genotoxic and non-genotoxic compounds and the $p53^{+/-}$ model is useful for testing clearly and

Table 9.3 Comparison of transgenic models for the detection of carcinogens

Class of chemical	Compounds tested	Results
Non-genotoxic Non-carcinogens	Ampicillin; D-mannitol; sulfisoxazole	All negative in all transgenic models
Genotoxic carcinogens	Cyclophosphamide	Positive in $p53^{+/-}$; equivocal in *rasH2* and Tg.AC
	Melphalan	Positive in $p53^{+/-}$and Tg.AC; equivocal in *rasH2*.
	Phenacetin	Positive in *rasH2* only
Immunosuppressants hormonal carcinogens	Cyclosporin A	Positive in $p53^{+/-}$; equivocal in *rasH2* and Tg.AC.
	Diethylstilbestrol	Positive in all except oral Tg.AC model
	Estradiol	Variable results
Nongenotoxic rodent carcinogens	PB; methapyrilene; reserpine; dieldrin; haloperidol; chlorpromazine; chloroform; metaproterenol; sulfamethoxazole.	Negative in all assays except equivocal result with chloroform in $p53^{+/-}$.
	Clofibrate; DEHP; WY-14643	Highly variable; all gave positive/equivocal results in at least one model and all models gave positive/equivocal results with at least one peroxisome proliferator

Data from MacDonald *et al.* (2004).

equivocally genotoxic compounds (although the CPMP Safety Working Party does not consider that its use needs to be restricted to this group). The FDA and Committee for Proprietary Medical Products Safety Working Party considered the Tg.AC model suitable for testing dermally applied products, but the MHLW still had doubts concerning the stability of the phenotype in this line.

The ILSI programme also attempted to evaluate the $XPA^{-/-}$ and $XPA^{-/-}/p53^{+/-}$ models, but the agreed view of the regulatory bodies represented at the workshop was that these models still required further development before they could be used in a regulatory setting.

The conclusions drawn by MacDonald *et al.* (2004) are as follows:

- The alternative assays currently under consideration have value in carcinogen identification.
- These assays can serve as an alternative to the standard mouse 2-year bioassay in a testing program.
- The testing paradigm is, in general, accepted for the testing of pharmaceuticals in the three major regulatory arenas (United States, Europe, and Japan).
- The results of these assays should not be considered on their own, but rather, integrated with other available data and considered as part of a weight of evidence approach for risk assessment purposes.

9.5.7 Limitations of the assays

A key limitation of any transgenic approach is that it only addresses one or a limited number of pathways, whereas carcinogenesis is a complex, multifactorial process. In the case of the models described previously, single genes with incompletely understood relevance to human carcinogenesis have been modified. In some cases the results obtained have not been consistent with the predicted role of the transgene in tumour development, and this has led to criticism of the assays. Furthermore, the insertion of a human gene into the mouse genome may interfere with the activities of endogenous pathways in ways which cannot be predicted, leading to further uncertainty regarding the validity of the results obtained.

When evaluating the results obtained using transgenic mouse models it is essential to take into account the genetic background of the strains used, since differences between inbred strains have a profound effect on susceptibility to carcinogenesis in particular tissues and crossing to different genetic background may have a marked effect on tumour incidences. The $p53^{+/-}$ and *rasH2* lines are maintained on a C57BL/6 background, but they may be interbred with other lines for use in carcinogenicity testing. The *rasH2* model, for example, is actually used on a CB6F1 background while the $p53^{+/-}$ transgene is available on C57BL/6, SJL, Balb/c, CBA and FVB backgrounds.

In addition, the method by which the line is generated may have a profound effect on its responsiveness to carcinogens. The responsiveness of the $p53^{+/-}$ line described here is well defined, but several other $p53^{+/-}$ lines generated by different methods have been generated and these have differing susceptibilities to carcinogens depending upon the method used for transgenesis and the genetic background of the mice used.

All the models described in this section were created by random insertion of the transgene. The problems experienced during validation of the Tg.AC line illustrate the disadvantages of this approach, and now that targeted integration strategies are available these are preferable in order to ensure generation of a stable line in which the transgene is present at an appropriate, defined location. However transgenic lines are generated, it is essential that they are characterised fully in terms of site of integration, number of copies and identification of any endogenous genes disrupted by insertion of the transgene.

The problems experienced with the Tg.AC model also illustrate the importance of rigorous QC when breeding and using transgenic mouse lines. The consensus view is that the *rasH2* and $p53^{+/-}$ models were adequately characterised by their developers and are being appropriately controlled and monitored by their suppliers, while recent action taken by the supplier of the Tg.AC line has resolved the problems experienced with this model.

While the transgenic lines described here are able to detect the majority of carcinogens, there are some exceptions. Both the *rasH2* and $p53^{+/-}$ models are, for some reason, resistant to the induction of malignant lymphomas by phenolphthalein. In the case of the $p53^{+/-}$ model this is independent of strain background, since the phenomenon is observed on both CBA and CIEA backgrounds.

9.5.8 Evaluation

While the limitations outlined previously must always be borne in mind, initial concerns that transgenic strains with defects in oncogenes and tumour suppressor genes would be excessively susceptible to tumour development and oversensitive to chemical carcinogenesis have not been borne out in practice. In fact, the transgenic lines are less prone to false positives than are conventional strains, at least over a 6-month assay period. These assays represent an appropriate pragmatic approach to reducing the cost, labour intensity and timescale of regulatory carcinogenesis testing and are well suited to the hazard identification stage of risk assessment, especially in a pharmaceutical setting.

In general, the transgenic models appear to have greater specificity with regard to human carcinogenicity than the conventional 2-year bioassay, but this is not really surprising given that the 2-year bioassay was designed to maximise sensitivity even at the cost of reduced specificity.

The $p53^{+/-}$ model is well suited for detecting genotoxic carcinogens while the *rasH2* strain detects the majority of both genotoxic and non-genotoxic carcinogens. The Tg.AC strain is of more limited usefulness because it can only detect nongenotoxic carcinogens (and tumour promoters) administered via the skin; indeed, it has now been more-or-less abandoned because of difficulties in interpretation and its sensitivity to effects due to irritation.

From the regulatory perspective, short-term *in vivo* bioassays such as the $p53^{+/-}$ assay may be used as an adjunct to conventional bioassays, and regulatory guidelines now allow for the use of a short-term alternative bioassay in mice to substitute for a second long-term bioassay (Storer *et al*., 2003). The *rasH2* model is

currently favoured for general screening and testing of non-genotoxic compounds, while the $p53^{+/-}$ model tends to be used to assess genotoxicity (Long *et al.*, 2010). Based upon the validation studies carried out to date, the combination of a 2-year rat bioassay and a 6-month transgenic study is expected to result in fewer false positives and no increase in false negatives; however, there is a requirement that the use of an alternative assay will provide additional information (e.g. on genotoxic mechanisms) that would not be generated by a second long-term bioassay. The key advantage of these assays is shorter assay duration, but they still do not provide information about modes of action or relevance to humans. Furthermore, concern has been expressed regarding the fact that candidate drugs which are positive in clastogenicity assays are often negative in the P53+/− assay (Storer *et al.*, 2010), and also on the basis that a reduction in the number of lifetime bioassays conducted may lead to non-genotoxic carcinogens escaping detection (Hernandez *et al.*, 2009). The industry perspective, however, is a pragmatic one: the use of transgenic models will help to reduce the high costs of drug development, and on that basis it is likely to continue. The higher up-front cost of purchasing the transgenic mice is outweighed by savings in manpower, reductions in the number of animals required, smaller amount of test item required and lower resource costs (formulation, diet and facility overheads). Indeed, it has been predicted that, in future, a single carcinogenicity study in a short-term transgenic model could be deemed acceptable for regulatory assessment of compounds which have not raised concerns in an integrated genotoxicity test battery and chronic rat toxicity studies (Storer *et al.*, 2010).

9.6 Conclusions

While the lifetime bioassay remains the gold standard in identifying potential carcinogens, transgenic models have made a huge contribution in recent years, particularly in shortening the time taken to establish the carcinogenic liabilities of novel compounds. If used judiciously, keeping in mind the fact that the deletion of a key gene or introduction of an exogenous one may perturb homeostasis in unpredictable ways, these models are valuable tools with significant potential to contribute to the process of carcinogenic risk assessment.

In summary, carcinogens may act via a multitude of mechanisms which include direct and indirect genotoxicity as well as a variety of non-genotoxic mechanisms. When thinking about this issue it is important to bear in mind that carcinogenesis always involves both genotoxic and non-genotoxic processes: all tumours arise via a combination of mutational and epigenetic mechanisms, and there must be some DNA damage, but this may not be caused directly by the carcinogen. Conversely, DNA damage alone is not sufficient for carcinogenesis; common epigenetic processes which occur in human tumours include global hypomethylation early in the progression phase and regional hypermethylation of normally unmethylated CpG islands, both of which can result in potentially oncogenic changes in gene expression.

Self-assessment questions

- How would you evaluate the risk associated with exposure to a chemical which is non-genotoxic but causes a doubling of liver weight in rodents during the first 2 weeks of dosing?
- How can transgenic mouse models contribute to the identification of mutagenic site of contact carcinogens?
- Will short-term assays in transgenic mice ever replace the NTP 2-year bioassay?

References

Alden, C., Smith, P. and Morton, D. (2002) Application of genetically altered models as replacement for the lifetime mouse bioassay in pharmaceutical development. *Toxicologic Pathology*, **30**, 135–138.

Besaratinia, A. and Pfeifer, G.P. (2010) Applications of the human p53 knock-in (Hupki) mouse model for human carcinogen testing. *FASEB Journal*, **24**, 2612–2619.

Boobis, A.R., Cohen, S.M., Doerrer, N.G. *et al.* (2009) A data-based assessment of alternative strategies for identification of potential human cancer hazards. *Toxicologic Pathology*, **37**, 714–732.

Boverhof, D.R., Chamberlain, M.P., Elcombe, C.R. *et al.* (2011) Transgenic animal models in toxicology: historical perspectives and future outlook. *Toxicological Sciences*, **121**, 207–233.

Buchmann, A., Karcier, Z., Schmid, B. *et al.* (2008) Differential selection for B-*raf* and Ha-*ras* mutated liver tumors in mice with high and low susceptibility to hepatocarcinogenesis. *Mutation Research*, **638**, 66–74.

Cheung, C., Akiyama, T.E., Ward, J.M. *et al.* (2004) Diminished hepatocellular proliferation in mice humanized for the nuclear receptor peroxisome proliferator-activated receptor alpha. *Cancer Research*, **64**, 3849–3854.

Cohen, S.M. (2001) Alternative models for carcinogenicity testing: weight of evidence evaluations across models. *Toxicologic Pathology*, **29** (Suppl), 183–190.

Cohen, S.M. and Arnold, L.L. (2011) Chemical carcinogenesis. *Toxicological Sciences*, **120** (Suppl 1), S76–S92.

Cohen, S.M., Robinson, D. and MacDonald, J. (2001) Alternative models for carcinogenicity testing. *Toxicological Sciences*, **64**, 14–19.

Corvi, R., Aardema, M.J., Gribaldo, L. *et al.* (2012) ECVAM prevalidation study on *in vitro* cell transformation assays: general outline and conclusions of the study. *Mutation Research*, **744**, 12–19.

Creton, S., Aardema, M.J., Carmichael, P.L. *et al.* (2012) Cell transformation assays for prediction of carcinogenic potential: state of the science and future research needs. *Mutagenesis*, **27**, 93–101.

Donehower, L.A., Harvey, M., Slagle, B.L. *et al.* (1992) Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature*, **356**, 215–221.

Floyd, E., Mann, P., Long, G. and Ochoa, R. (2002) The Trp53 hemizygous mouse in pharmaceutical development: points to consider for pathologists. *Toxicologic Pathology*, **30**, 147–156.

Hernandez, L.G., Van Steeg, H., Luijten, M. and Van Benthem, J. (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutation Research*, **682**, 94–109.

Hoenerhoff, M.J., Pandiri, A.R., Lahousse, S.A. *et al.* (2011) Global gene profiling of spontaneous hepatocellular carcinoma in B6C3F1 mice: similarities in the molecular landscape with human liver cancer. *Toxicologic Pathology*, **39**, 678–699.

Inselman, A.L., Hansen, D.K., Lee, H.Y. *et al.* (2011) Assessment of research models for testing gene-environment interactions. *European Journal of Pharmacology*, **668** (Suppl 1), S108–S116.

Jacobson-Kram, D., Sistare, F.D. and Jacobs, A.C. (2004) Use of transgenic mice in carcinogenicity hazard assessment. *Toxicologic Pathology*, **32** (Suppl 1), 49–52.

Jaworski, M., Hailfinger, S., Buchmann, A. *et al.* (2005) Human p53 knock-in (hupki) mice do not differ in liver tumor response from their counterparts with murine p53. *Carcinogenesis*, **26**, 1829–1834.

Kenzelmann Broz, D. and Attardi, L.D. (2010) *In vivo* analysis of p53 tumor suppressor function using genetically engineered mouse models. *Carcinogenesis*, **31**, 1311–1318.

King-Herbert, A. and Thayer, K. (2006) NTP workshop: animal models for the NTP rodent cancer bioassay: stocks and strains--should we switch? *Toxicologic Pathology*, **34**, 802–805.

King-Herbert, A.P., Sills, R.C. and Bucher, J.R. (2010) Commentary: update on animal models for NTP studies. *Toxicologic Pathology*, **38**, 180–181.

Krieken, J. H. J. M., Jung, A., Kirchner, T., Carneiro, F., *et al.* 2008. *KRAS* mutation testing for predicting response to anti-EGFR therapy for colorectal carcinoma: proposal for an European quality assurance program. *Virchows Archiv*, **453**, 417–431.

Lemieux, C.L., Douglas, G.R., Gingerich, J. *et al.* (2011) Simultaneous measurement of benzo[a]pyrene-induced Pig-a and lacZ mutations, micronuclei and DNA adducts in Muta™Mouse. *Environmental and Molecular Mutagenesis*, **52**, 756–765.

Long, G.G., Morton, D., Peters, T. *et al.* (2010) Alternative mouse models for carcinogenicity assessment: industry use and issues with pathology interpretation. *Toxicologic Pathology*, **38**, 43–50.

MacDonald, J., French, J.E., Gerson, R.J. *et al.* (2004) The utility of genetically modified mouse assays for identifying human carcinogens: a basic understanding and path forward. The Alternatives to Carcinogenicity Testing Committee ILSI HESI. *Toxicological Sciences*, **77**, 188–194.

Maire, M.A., Pant, K., Phrakonkham, P. *et al.* (2012a) Recommended protocol for the Syrian hamster embryo (SHE) cell transformation assay. *Mutation Research*, **744**, 76–81.

Maire, M.A., Pant, K., Poth, A. *et al.* (2012b) Prevalidation study of the Syrian hamster embryo (SHE) cell transformation assay at pH 7.0 for assessment of carcinogenic potential of chemicals. *Mutation Research*, **744**, 64–75.

Mitsumori, K. (2002) Evaluation on carcinogenicity of chemicals using transgenic mice. *Toxicology*, 181–182, 241–244.

Mitsumori, K. (2003) Possible mechanism on enhanced carcinogenesis of genotoxic carcinogens and unsolved mechanisms on lesser carcinogenic susceptibility to some carcinogens in *rasH2* mice. *Journal of Toxicological Sciences*, **28**, 371–383.

Morton, D., Alden, C.L., Roth, A.J. and Usui, T. (2002) The Tg *rasH2* mouse in cancer hazard identification. *Toxicologic Pathology*, **30**, 139–146.

Nakamura, T., Ito, Y., Yanagiba, Y. *et al.* (2009) Microgram-order ammonium perfluorooctanoate may activate mouse peroxisome proliferator-activated receptor alpha, but not human PPARalpha. *Toxicology*, **265**, 27–33.

Pant, K., Bruce, S.W., Sly, J.E. *et al.* (2012) Prevalidation study of the Syrian hamster embryo (SHE) cell transformation assay at pH 6.7 for assessment of carcinogenic potential of chemicals. *Mutation Research*, **744**, 54–63.

Park, H.J., Oh, J.H., Park, S.M. *et al.* (2011) Identification of biomarkers of chemically induced hepatocarcinogenesis in *rasH2* mice by toxicogenomic analysis. *Archives of Toxicology*, **85**, 1627–1640.

Parsons, B.L., Beland, F.A., Von Tungeln, L.S. *et al.* (2005) Levels of 4-aminobiphenyl-induced somatic H-*ras* mutation in mouse liver DNA correlate with potential for liver tumor development. *Molecular Carcinogenesis*, **42**, 193–201.

Pitot, H.C. (2007) Adventures in hepatocarcinogenesis. *Annual Review of Pathology*, **2**, 1–29.

Prior, I.A., Lewis, P.D. and Mattos, C. (2012) A comprehensive survey of *Ras* mutations in cancer. *Cancer Research*, **72**, 2457–2467.

Pritchard, J.B., French, J.E., Davis, B.J. and Haseman, J.K. (2003) The role of transgenic mouse models in carcinogen identification. *Environmental Health Perspectives*, **111**, 444–454.

Sasaki, K., Bohnenberger, S., Hayashi, K. *et al.* (2012) Recommended protocol for the BALB/c 3T3 cell transformation assay. *Mutation Research*, **744**, 30–35.

Schechtman, L.M. (2012) Rodent cell transformation assays-a brief historical perspective. *Mutation Research*, **744**, 3–7.

Stanley, L.A. (1995) Molecular aspects of chemical carcinogenesis: the roles of oncogenes and tumour suppressor genes. *Toxicology*, **96**, 173–194.

Stanley, L. A., Devereux, T. R., Foley, J., Lord, P. G., *et al.* 1992. Proto-oncogene activation in liver tumors of hepatocarcinogenesis-resistant strains of mice. *Carcinogenesis*, **13**, 2427–33.

Storer, R.D., French, J.E., Donehower, L.A. *et al.* (2003) Transgenic tumor models for carcinogen identification: the heterozygous Trp53-deficient and *RasH2* mouse lines. *Mutation Research*, **540**, 165–176.

Storer, R.D., Sistare, F.D., Reddy, M.V. and Degeorge, J.J. (2010) An industry perspective on the utility of short-term carcinogenicity testing in transgenic mice in pharmaceutical development. *Toxicologic Pathology*, **38**, 51–61.

Tanaka, N., Bohnenberger, S., Kunkelmann, T. *et al.* (2012) Prevalidation study of the BALB/c 3T3 cell transformation assay for assessment of carcinogenic potential of chemicals. *Mutation Research*, **744**, 20–29.

Taneja, P., Zhu, S., Maglic, D. *et al.* (2011) Transgenic and knockout mice models to reveal the functions of tumor suppressor genes. *Clinical Medicine Insights: Oncology*, **5**, 235–257.

Tong, C., Li, P., Wu, N.L. *et al.* (2010) Production of p53 gene knockout rats by homologous recombination in embryonic stem cells. *Nature*, **467**, 211–213.

Umemura, T., Kodama, Y., Hioki, K. *et al.* (2002) The mouse *rasH2*/BHT model as an *in vivo* rapid assay for lung carcinogens. *Japanese Journal of Cancer Research*, **93**, 861–866.

Umemura, T., Kodama, Y., Nishikawa, A. *et al.* (2006) Nine-week detection of six genotoxic lung carcinogens using the *rasH2*/BHT mouse model. *Cancer Letters*, **231**, 314–318.

Vanparys, P., Corvi, R., Aardema, M. *et al.* (2011) ECVAM prevalidation of three cell transformation assays. *ALTEX*, **28**, 56–59.

Vasseur, P. and Lasne, C. (2012) OECD Detailed Review Paper (DRP) Number 31 on "Cell Transformation Assays for Detection of Chemical Carcinogens": main results and conclusions. *Mutation Research*, **744**, 8–11.

Vom Brocke, J., Krais, A., Whibley, C. *et al*. (2009) The carcinogenic air pollutant 3-nitrobenzanthrone induces GC to TA transversion mutations in human p53 sequences. *Mutagenesis*, **24**, 17–23.

Wahli, W. and Michalik, L. (2012) PPARs at the crossroads of lipid signaling and inflammation. *Trends in Endocrinology and Metabolism*, **23**, 351–363.

Yang, Q., Nagano, T., Shah, Y. *et al*. (2008) The PPAR alpha-humanized mouse: a model to investigate species differences in liver toxicity mediated by PPAR alpha. *Toxicological Sciences*, **101**, 132–139.

Young, R.R., Rogers, B.J., Provost, G.S. *et al*. (1995) Interlaboratory comparison: liver spontaneous mutant frequency from lambda/*lacI* transgenic mice (Big Blue®) (II). *Mutation Research*, **327**, 67–73.

Zielinski, B., Liu, Z., Hollstein, M. *et al*. (2002) Mouse models for generating P53 gene mutation spectra. *Toxicology Letters*, **134**, 31–37.

10
Emerging Techniques

10.1 What's next?

As we move forward into the twenty-first century, the discipline of toxicology is undergoing many changes driven by a desire (and legislative mandate) to reduce the use of mammalian *in vivo* models, the need for biomarkers which will allow more reliable extrapolation between animals and humans and the increasing availability of sophisticated approaches to improve our understanding of mechanisms of toxicity. In this chapter, we will look at some recent developments in each of these areas, including: alternative non-mammalian models with reduced welfare issues and the scope for HTS; novel circulating biomarkers which can be used in minimally invasive studies allowing real-time monitoring of adverse effects and direct comparison of effects in humans and animals; and the application of systems biology approaches which permit the detailed analysis of complex processes such as systemic toxicity.

10.2 Novel model organisms

Mammalian (usually rodent) models have traditionally been used in toxicology and testing because they are considered to be biologically similar to humans and are therefore assumed to respond to toxic insults in a manner which replicates effects in humans. However, it will be clear from the preceding discussions in this book that this is not necessarily the case. Furthermore, animal testing is now discouraged for welfare reasons and from a practical perspective mammalian organisms do not lend themselves to HTS (not to put too fine a point on it, they are too big to be kept in the large numbers required).

At present, the most commonly used lower organisms used in toxicity testing are the nematode *Caenorhabditis elegans*, the fruit fly *Drosphila melanogaster* and the zebrafish, *Dario rerio*. The first two of these are limited in that they are

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology

evolutionarily very distant from humans, although they do exhibit many of the pharmacological and physiological properties required for toxicity testing. Their main advantages in relation to mammalian models are their ready availability, low cost and amenability to HTS methods.[1] However, it is the zebrafish which is becoming the organism of choice for many types of toxicity testing because it represents a convenient halfway house between simple invertebrate organisms such as *C. elegans* and *D. melanogaster* and complex, ethically compromised mammalian models.

10.2.1 The zebrafish

The zebrafish, *Danio rerio* (Figure 10.1), is a tropical freshwater fish belonging to the minnow family (Cyprinidae) of order Cypriniformes. It is a popular aquarium fish, frequently sold under the trade name zebra danio, was introduced as a genetic model organism in the late 1960s and is now an important vertebrate model organism, particularly in ecotoxicology, as well as in other areas of genetic and developmental biology and toxicology.

Example: Use of zebrafish to study ecotoxicological interactions between endocrine disrupting chemicals and nanomaterials

One of the obvious applications of an aquatic organism such as the zebrafish is in ecotoxicology. Among the hottest topics in ecotoxicology at the moment are the potential environmental effects of endocrine disrupting chemicals and nanomaterials. A series of studies addressing the potential interactions between these two categories of materials is currently under way using the C_{60} compound fullerene as a representative nanomaterial and 17α-ethinylestradiol as a representative endocrine disrupting chemical. The ability of the insoluble nanomaterial

Figure 10.1 **Zebrafish** (source: Dr Theodore Henry, Centre for Marine Biodiversity and Biotechnology, Heriott-Watt University. With permission of Dr Theodore Henry)

[1] The advantages and limitations of *C. elegans* and *D. melanogaster* are itemised in Van Vliet (2011).

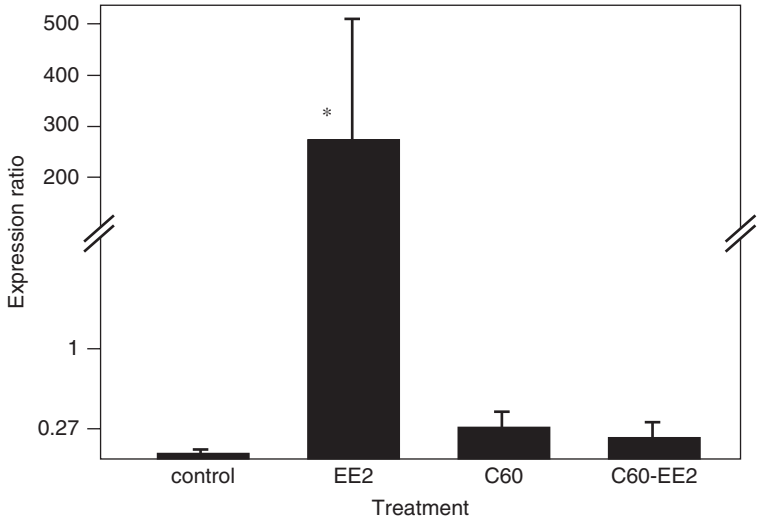


Figure 10.2 No bioavailability of 17α-ethinylestradiol when associated with nC60 aggregates during dietary exposure in adult male zebrafish (*Danio rerio*). Induction of vitellogenin genes (vtg1A/B) in zebrafish fed control brine shrimp (BS) or BS that had accumulated nC60, nC60 + 17α-ethinylestradiol (EE2), or EE2 only. Target gene expression was normalized with internal control gene (b actin). All data are expressed as mean ± SEM ($n = 3$) and asterisk indicates significant difference in vtg1A/B expression from all other treatment (one-way ANOVA, followed by Tukey's post hoc test at $p < 0.05$) (source: Park *et al.* (2010); figure 3. Reproduced with permission from Elsevier)

C_{60} to form suspended colloidal aggregates in aqueous environments had led to concerns that this material might have toxic effects in aquatic organisms, although after careful study it was shown that C_{60} has little or no intrinsic toxicity in fish (Henry *et al.*, 2007). Another suggestion was that C_{60} aggregates might exacerbate the adverse effects of endocrine disrupting chemicals such as 17α-ethinylestradiol by acting as a carrier, thus increasing the effective exposure of aquatic organisms. This has been addressed in studies in which zebrafish were exposed to either C_{60} alone, 17α-ethinylestradiol alone or C_{60} plus 17α-ethinylestradiol, using up-regulation of the gene encoding the yolk lipoprotein vitellogenin to assess the bioavailablity of 17α-ethinylestradiol (Henry *et al.*, 2009). When delivered via the diet (by the fish consuming brine shrimp which had accumulated it), 17α-ethinylestradiol caused significant induction of vitellogenin gene expression (representing bioavailability of 17α-ethinylestradiol); however, when it was combined with C_{60} in the shrimps' feed, no induction was observed; indeed, the level of expression in fish exposed to 17α-ethinylestradiol plus C_{60} was similar to that in controls and in fish treated with C_{60} alone (Figure 10.2) (Park *et al.*, 2010). It seemed that the C_{60} aggregates were resistant to the digestive processes of the zebrafish, since they remained intact during passage through the zebrafish digestive tract, and they appeared to be unable to cross the mucosal epithelium. Furthermore, it appeared likely that the 17α-ethinylestradiol was adsorbed onto the C_{60} aggregates, protecting it from being absorbed. This led to

the conclusion that concern about the potential association of toxic substances with C_{60} aggregates may be misplaced and that, in fact, C_{60} might sequester such substances, thus reducing their bioavailability. This result was not unique to the slightly unusual dietary exposure route used; in another study, very similar results were observed when the zebrafish were exposed to C_{60} and 17α-ethinylestradiol via the water they swam in (Park *et al.*, 2011). It seems, therefore, that C_{60} aggregates may be able to behave like activated carbon, removing compounds such as 17α-ethinylestradiol from water by sorption and potentially exerting a protective effect against their adverse effects.

The zebrafish has a number of attributes which enable it to bridge the gap between simple *in vitro* assays and experiments in whole animal models, causing it to become increasingly popular as a model for medium to HTS toxicity studies (Lieschke and Currie, 2007; McCollum *et al.*, 2011; Sipes, Padilla, and Knudsen, 2011).

- Zebrafish are small (an adult is about 3 cm long, while an embryo at the stage used for toxicity testing is 2–3 mm in length), cheap to keep and develop rapidly outside the mother's body. A zebrafish embryo can survive and grow for 4–5 days in a single well of a 96 or even 384-well multiwall plate and a single female zebrafish can generate 10 000 offspring per annum, meaning that a ready supply of test organisms can be produced even in a relatively small laboratory (Peterson and MacRae, 2012).
- Zebrafish have a genome and body plan very similar to those of other vertebrates and are anatomically and histologically surprisingly similar to their mammalian counterparts (Menke *et al.*, 2011). Their tissues are largely homologous to their mammalian equivalents at the anatomical, physiological and molecular levels. Furthermore, because of its small size, the complete histology of a zebrafish embryo can be reviewed using a minimal number of microscope slides although the down-side of this is that a zebrafish embryo provides insufficient material for many conventional biochemical assays.
- The zebrafish genome has been fully sequenced. It comprises 25 chromosomes (making it about half the size of the human genome) and, despite the fact that the zebrafish and human lines diverged 450 million years ago, contains more than 1300 genes having ≥86% homology with corresponding human genes.
- The fact that during early development the zebrafish embryo is transparent means that this model lends itself very well to microscopic analysis and is particularly well suited to time-lapse fluorescent imaging of transgenic reporter lines. The small size and large number of individual organisms available, coupled with the suitability of this model for fluorescent microscopy, means that zebrafish assays lend themselves very well to use in HTS or even ultra-HTS mode.

These characteristics enable the zebrafish to act as a complement to existing *in vitro* technologies and established preclinical models that lends itself to scaling to HTS mode; to put it another way the zebrafish model represents the sweet spot between HTS and full representation of vertebrate development (Peterson and MacRae, 2012). The use of zebrafish assays is closely aligned with drives

to introduce more HTS tests into toxicology since it is amenable to automation without the loss of detailed morphological information (Sipes, Padilla, and Knudsen, 2011).

Zebrafish are generally used at an early stage of development, conventionally between 72 and 120 h post-fertilisation (though they may be considered to be embryonic up to 30 days post-fertilisation, when they become juveniles).[2] Zebrafish embryos from fertilisation to about 120 h post-fertilisation are not considered to be animals for welfare purposes in most countries because at this stage of development (defined by EC Directive 2010/68/EU as being the period before they are able to feed independently) they are assumed to experience no/little pain, suffering, distress or lasting harm as a result of experimental manipulations (Strahle *et al*., 2012). However, it is quite difficult to define the end of the embryonic period unequivocally in fish, and this has led to ambiguity in some jurisdictions: for example, in Germany the embryonic period is considered by some authorities to last to 120 h post-fertilisation while in others embryos after 48 h post-fertilisation are considered to be animals requiring protection. Indeed, depending on whether one defines independent feeding as the time the yolk sac is no longer visible, the time of maturation of the digestive organs, the time at which free active swimming starts or the time the fish are actually able to incorporate food, the cut-off point can vary by several days (from about 96 to 144 h post-fertilisation), and this varies further with the temperature at which the fish are kept. Most people, however, take it that free feeding starts at about 120 h post-fertilisation, and that prior to that time point use of the zebrafish falls into the Replacement and Refinement categories of the 3Rs. The advantage of this for toxicity testing purposes is that it creates an alternative model which resembles an adult animal in terms of complexity and organisation; in other words an alternative whole organism model. Nevertheless, although this view is strictly correct and complies with the letter of the law, zebrafish are free-living sentient organisms and it should be recognised that use of such a test system could be considered contrary to the spirit of the 3Rs.

With regard to developmental toxicity, zebrafish are becoming increasingly popular for embryological screening techniques, for which they are well suited because they can survive major malformations and loss of organ function well into the developmental period as a consequence of their ability to survive on the contents of the yolk sac and diffusion of oxygen from the surrounding medium. The availability of thousands of developmental mutants has also helped to establish the zebrafish as a mainstream model in developmental biology. However, there is still a need for more standardisation in terms of husbandry (nutrition, lighting, microflora), scoring time with respect to developmental stage, strains used and endpoints measured (de Esch *et al*., 2012; Peterson and MacRae, 2012).

One test which is well established and becoming increasingly popular is the zebrafish embryotoxicity test (Piersma, 2011), which uses embryos up to 72 h after fertilisation, during which time they develop into completely hatched, swimming larvae. This is a very rapid developmental process compared with that undergone by mammalian embryos, and it allows the study of vertebrate development *in situ* throughout embryogenesis, although the developmental process is so fast that it

[2] For general background on the zebrafish and its use in all aspects of life sciences, see http://zfin.org/.

has been described as 'like screening on a runaway train'. This means that great care has to be taken to dose during the correct developmental window, which can be very short, and to take into account the fact that individual embryos develop at different rates. It is, however, facilitated by the fact that the eggs are transparent so that the embryos are easy to visualise, and it is legally considered to be a non-animal test method for the reasons outlined previously.

Zebrafish embryology methods are now covered by OECD and EPA guidelines, and various non-regulatory methods are also available (McCollum *et al*., 2011). The endpoints covered include hatching, body size, curvature, defects in swim bladder inflation and yolk sac oedema. In addition, a standardised scoring system, the General Morphology Score, has been developed at RIVM to facilitate the evaluation of outcomes in the zebrafish assay (Piersma, 2011). One note of caution, however, is that the chorion which surrounds the developing embryo is impermeable to some classes of chemicals, so penetration should be evaluated as part of the processing of results. Small molecules such as ethanol can diffuse through the chorion, though slowly; however, in order to permit the entry of larger molecules it is necessary either to disrupt the membrane (e.g. by pronase digestion) or bypass it by injecting compounds directly into the peri-vitteline space (Ali *et al*., 2011). Disruption of the chorion is ill-advised since it can affect both the integrity and behaviour of the embryo, but injection is less difficult than it sounds if the embryos are grown in multiwall plates, and a robotic microinjection method for dosing zebrafish embryos is available.

Zebrafish are becoming increasingly popular for developmental neurotoxicity testing (de Esch *et al*., 2012), being used to measure endpoints which include altered CNS development, visual impairment, olfactory toxicity and ototoxicity. Neural cells first become evident in zebrafish embryos at the end of gastrulation, although specification of neural fate occurs earlier than this, and interestingly cyp26 (a member of a CYP gene family which is inducible by retinoic acid in mammals (Ray *et al*., 1997)) is a neuronal marker in zebrafish. Furthermore, one advantage the zebrafish has over purely *in vitro*, cell-based methods is that it engages in behaviours which can be used to evaluate neurotoxicological effects. The endpoints assayed can include seizures, involuntary movements, learning/memory and various behavioural readouts.

The zebrafish lends itself particularly well to the use of fluorescent readouts because, in addition to being transparent, it is easy to engineer with fluorescent reporters using Tol2 transposon technology (Suster *et al*., 2009a; Suster, Sumiyama, and Kawakami, 2009b). This method allows large BAC constructs to be introduced intact into the zebrafish genome at specific integration sites which can be mapped accurately, leading to robust expression of the transgene and avoiding the problems associated with concatameric integration. Fragmentation methods, including LoxP-mediated deletion, can be used to dissect the transgenes and identify *cis*-regulatory elements, and the fact that zebrafish embryos develop outside the dam's body means that the consequences of any alteration can be observed *in vivo* throughout development. The method does have a few limitations, however: the rate of germ line transmission in zebrafish tends to be low, and zebrafish transgenes are susceptible to silencing and other alterations in expression as a result of perturbations in the genomic landscape (although this is,

to some extent, alleviated by the use of transposon-based methods) (Chatterjee and Lufkin, 2011).

Transgenic zebrafish lines labelled for specific tissues, cellular subtypes, developmental stages and processes such as apoptosis are readily available (McCollum *et al*., 2011), and the wide variety of different coloured labels they carry allows multiple readouts to be measured. Several GFP-based reporter models of relevance to human toxicology, including an Hsp70:eGFP line which responds to heavy metals and a CYP1A1:GFP line which responds to Ah-receptor ligands, have also been reported (Yang *et al*., 2009). In addition, it is easy to knock down the expression of specific genes in zebrafish using antisense RNA and they are also amenable to reverse genetics approaches such as TILLING and zinc finger nuclease mediated gene disruption (Doyon *et al*., 2008). More than 5000 lines of transgenic reporter zebrafish are now available (Strahle *et al*., 2012), but their usefulness in toxicology remains to be demonstrated.

Zebrafish are now commonly used in areas of drug safety assessment (McGrath and Li, 2008). Further development of such assays depends on the identification of zebrafish phenotypes which are direct parallels for the equivalent processes in mammals in terms of, for example, cardiac physiology and drug responses (although it has been stated that at 48 h post-fertilisation, zebrafish drug responses are more like those of humans than those of existing rodent preclinical models) (Peterson and Macrae, 2012). The range of applications currently under investigation includes hepatotoxicity (Hill *et al*., 2012), cardiotoxicity, nephrotoxicity, GI toxicity, myotoxicity and neurotoxicity.

It is, however, necessary to remember some caveats before becoming too enthusiastic about replacing all existing tests with zebrafish tests:

- **Anatomy and histology:** Despite the general similarity between zebrafish organs and their mammalian counterparts, there are some important differences (Menke *et al*., 2011). The zebrafish intestine is essentially a single tube which starts off with a wide lumen and gradually gets narrower, folding twice within the abdominal cavity. It is not possible to identify a small or large intestine or a stomach within this structure, although there is some evidence of functional differentiation along the length of the tube. Furthermore, the zebrafish does not have a discrete pancreas and its liver, though functionally similar to that of a mammal, does not have the lobular and trabecular arrangement of hepatocytes which is typical of mammalian liver.
- **Energy metabolism:** Metabolic processes, including energy homeostasis, are generally conserved across evolution but zebrafish use lipids as a primary energy source rather than carbohydrates, so they tend to be hyperlipidaemic and hypercholesterolaemic in comparison with mammals (McCollum *et al*., 2011). Their PPARs (α, β and γ) are, however, highly homologous with those of mammals.
- **ADME:** The processes involved in ADME are much less fully characterised in zebrafish than in mammalian systems, partly because they are difficult to study in such a small organism. The fact that compound uptake from aqueous medium is strongly pH dependent should also be borne in mind when considering the zebrafish for use in ADME testing. It is also important to recognise that exposure via the medium is effectively dermal in zebrafish embryos since they

do not have functioning gills or take in medium by mouth; the main alternative route of exposure is by partitioning into the yolk.

  - Absorption and biodistribution processes in zebrafish are, as yet, poorly understood. However, there is evidence that zebrafish do have a BBB which is functionally homologous to the human one and expresses the tight junction protein Claudin 5 as well as the zebrafish homologue of MDR1, which has been identified by mining of the zebrafish genome and localised to the BBB by dual immunofluorescence (staining for Mdr1 and endothelial cell markers) (Umans and Taylor, 2012). Further work is still needed to determine when and how the zebrafish BBB develops and characterise the permeability characteristics of the mature and immature BBB. Furthermore, there are unanswered questions about the relationship between angiogenesis and barriergenesis and the signals which regulate these processes. These may be amenable to investigation using the many fluorescent reporter lines which are now available for various cell types and signalling pathways, taking advantage of the transparency of the embryo and the availability of time lapse fluorescence microscopy techniques.
  - Zebrafish do have a full complement of Phase I and Phase II xenobiotic metabolising enzymes, but these are, as yet, incompletely characterised and expression has mostly been characterised only at the mRNA level, with limited functional characterisation. CYP expression is first detected in fertilised eggs, suggesting that the mRNA detected is predominantly of maternal origin at this stage of development, although Cyp1 expression is detectable in embryos at 8 h post-fertilisation. Further characterisation of xenobiotic metabolism in zebrafish is still necessary.

- **DNA repair:** The zebrafish genome contains orthologues of genes involved in all the DNA repair processes found in higher eukaryotes, although the actual functions of the corresponding proteins have not yet been characterised (Pei and Strauss, 2013). Subject to proper characterisation of these functions, the zebrafish may be of value as a model in which to study DNA repair processes at time points which are inaccessible to study in mammalian embryos.
- **Chronic toxicity:** Long-term effects such as carcinogenicity have not yet been characterised in the zebrafish and in these cases the zebrafish has few advantages over mammalian systems because speed, its main asset, is not relevant over the long timescales involved. However, the amenability of the zebrafish to *in vivo* fluorescent methods could be of great benefit in the molecular analysis of pathways leading to tumour development.

10.2.2 Evaluation

Overall, it is clear (forgive the pun!) that the zebrafish has much to offer the toxicologist, both in developmental and other settings. As well as its scientific suitability for many applications, the implementation of zebrafish assays could lead to significant savings in terms of cost, time and animal usage. To achieve this objective, though, it is important to identify the specific areas where the use of zebrafish can improve on existing models in terms of sensitivity, specificity, cost and efficiency. In particular, the questions ‘How similar is zebrafish development

to that of humans and rodents?' and 'How predictive are zebrafish findings of mammalian developmental toxicity?' still await a definitive answer. Although the processes involved in cleavage and blastula/gastrula formation are highly conserved in evolution, implying that they are essential and any alterations are likely to be lethal, few studies detailing early changes in embryonic phenotypes following chemical treatment have been conducted to date. Most developmental processes are quite well understood, but there is a lack of well-established test methods to relate these to toxic effects and more evaluation is needed before the zebrafish is recognised as a valid alternative model for the prediction of human developmental toxicity (McCollum *et al.*, 2011).

10.3 Less invasive methods

Where the use of animals cannot be avoided, one way to reduce the numbers used and the suffering of those which are used is to implement minimally invasive methods which allow repeated sampling from the same animal (Reduction) and, ideally, are sensitive to effects at lower doses (Refinement). One approach to this is to take advantage of circulating and/or excreted biomarkers, some of which have been in routine use in the clinic for many years.

10.3.1 Use of biomarkers

Biomarkers have been defined as measurable internal indicators of molecular and/or cellular alterations that may appear in an organism after or during exposure to a toxicant or possible disease (Roux *et al.*, 2011). The ideal biomarker is sensitive, specific, easily measurable, accurate, reproducible, cost-effective and easy to interpret. It should readily discriminate between healthy and disease states and should be detectable early in the disease process, or even before disease arises (in the case of biomarkers of exposure and susceptibility). Additional desirable characteristics include low cost, rapidity of onset and quantification, close correlation with the severity of injury and target organ specificity.

When necrosis occurs in response to tissue damage, intracellular enzymes are released into the circulation. Their detection can provide information about the location and extent of the damage and organ-specific enzymes, isozymes and enzyme patterns can be used in non/minimally invasive methods to identify the organs involved in toxic damage. Thus, measurement of the concentrations of different marker enzymes in plasma allows a diagnosis to be made of the involvement of a specific organ. In humans, this provides the possibility of rapid diagnosis, selection of an appropriate treatment and monitoring the effects of therapy, while in experimental studies it allows toxicity to be monitored in real time *in vivo*.

10.3.2 Liver

No biomarker taken alone is a definitive indicator of liver damage, so the evaluation of liver damage and disease usually involves quantitating a panel of biomarkers (Amacher, 2002). In the clinical context this is referred to as a

liver function test; the classic combination of biomarkers used includes aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP) and total plasma bilirubin. It should be noted that among these biomarkers the three enzymes reflect cellular injury but only bilirubin is a marker for actual loss of hepatic functions. Some liver function test panels also include γGT. Serum albumin and globulins can also be used as markers for liver damage; the latter are observed in chronic liver disease as a consequence of inflammatory reactions in the reticuloendothelial system.

- **AST** is a mitochondrial enzyme found in heart, liver, skeletal muscle and kidney. It catalyses the reaction:

  α-ketoglutaric acid + aspartic acid → oxaloacetic acid + glutamic acid

  AST can be used as a marker for either hepatocellular necrosis or myocardial infarction; it is usually included in liver function test panels, but is really redundant since its release from the liver usually parallels that of the other biomarkers.
- **Alanine aminotransferase** is more specific to the liver than AST. It is a cytosolic enzyme which catalyses the reaction:

  α-ketoglutaric acid + alanine → pyruvic acid + glutamic acid

  ALT is a marker for hepatocellular necrosis and has been found to be a useful preliminary indicator of potential post-marketing hepatotoxicity liabilities during drug development (Moylan *et al.*, 2012), although the results obtained should be treated with caution because of the lack of a standardised method of detection and the fact that the definition of normal depends on the population being studied. Causes of ALT release include hepatotoxins, infections, diabetes mellitus, liver tumours and drug-induced liver damage (corticosteroids, oestrogens, androgens, chloramphenicol, erythromycin and salicylates); an elevation to more than three times the upper limit of normal, as defined in the population being studied, is considered to be markedly abnormal.
- **Alkaline phosphatase** is present in many tissues. In the liver it is located on the cell membranes of the bile canaliculi. It is not tissue-specific but in the presence of liver damage (as indicated, for example, by a concomitant increase in ALT) it is a marker for damage to the bile ducts because accumulation of bile acids has a detergent-like effect on lipid membranes leading to cholestatic changes which are accompanied by a marked increase in serum alkaline phosphatase. Elevated serum levels of alkaline phosphatase can indicate disease of the biliary system (e.g. intrahepatic cholestasis, biliary obstruction or primary biliary cirrhosis).
- **γ-Glutamyl transpeptidase** is a sensitive indicator of liver damage but non-specific as to cause. It is an inducible enzyme so can be used as a marker for microsomal enzyme induction as well as liver damage. The presence of γGT-positive foci in the liver of rodents is used as a method for early detection of hepatocarcinogenesis.

Circulating levels of liver biomarkers vary markedly between individuals, so it is important to define an appropriate threshold to use in the interpretation of liver function tests. In the clinical setting, levels of these enzymes are defined as abnormal when they are raised to more than twice the upper limit of the normal reference range (ULN). In toxicology, a useful rule-of-thumb is that mild hepatotoxicity is often associated with a threefold elevation in circulating liver enzyme levels compared with control values while an elevation greater than about 10-fold indicates severe hepatotoxicity.

Abnormalities in liver function tests may indicate the presence of liver disease, and may also be predictive of risk of HCC (Wen *et al.*, 2012). The pattern of changes observed may help to determine whether a particular drug is implicated, although a diagnosis cannot be undertaken on the basis of these tests alone. Furthermore, elevations in circulating liver enzymes are common with certain drugs, but this is not always associated with significant hepatic injury. For example, serum transaminase levels rise in up to 40% of patients taking isoniazid but this gives rise to overt hepatotoxic symptoms in fewer than 5%. When potentially hepatotoxic drugs are prescribed, it is important to measure circulating liver enzymes prior to the initiation of treatment in order to provide a baseline with which subsequent measurements can be compared. Continued monitoring during the first few weeks of treatment with drugs which are known to cause hepatic toxicity can provide an early warning of an impending severe reaction.

Drug-induced liver injury Hepatotoxicity is a common cause of failure during drug development. Drugs whose development or marketing has had to be discontinued due to hepatotoxicity include troglitazone, temafloxacin, ibufenac, bromfenac, nomifensin, benoxaprofen, tienilic acid and perhexilin. The use of other drugs (e.g. zileuton, trovafloxacin, tolcapone and felbamate) has had to be severely restricted. However, it is difficult to predict which cases of DILI will become severe in comparison with those in which only transient ALT/AST elevations will occur and then resolve spontaneously. Furthermore, ALT and AST do not distinguish between different aetiologies of DILI, and ALT can also be elevated in response to non-toxic chemicals such as dexamethasone.

Drug-induced liver injury can be caused by intrinsic hepatotoxins or idiosyncratic drug reactions. In the United States, 10% of all liver toxicity is believed to be drug induced.[3] The symptoms of DILI include icterus (yellowing of the skin due to high levels of bilirubin in the blood), pruritis (itching), fatigue, weight loss, right upper abdominal discomfort, dark urine and pale stools. However, these are observed both in drug-induced hepatotoxicity and unrelated liver disease, making the differential diagnosis of DILI very problematic; indeed, in practice DILI is often diagnosed by the elimination of all other possibilities. Current FDA guidance regarding the evaluation of DILI during the pre-marketing phase of drug development[4] recommends the following criteria:

- An ALT level more than 3× ULN is an initial warning sign.

[3] For the latest information, see the DILI Network website, https://dilin.dcri.duke.edu/
[4] http://www.fda.gov/downloads/Drugs/ … /Guidances/UCM174090.pdf

- An ALT level more than 5× ULN indicates serious injury.
- An ALT level more than 8× ULN indicates the need to discontinue treatment.

These cut-off points represent a convenient rule-of-thumb for the preliminary diagnosis of DILI, but there is still an urgent need for new biomarkers that are specific for serious acute hepatocellular injury and that truly predict clinical outcomes. As things stand at the moment, elevations in serum enzyme levels do not, on their own, indicate a prognosis: an elevated ALT reading is an indicator of what has happened, not what will happen. The best current biomarker of actual liver injury is the combination of ALT and total bilirubin, which provides a measure of both cellular damage and loss of function. The best way to view the FDA cut-off points is as a measure of the urgency, in an individual case, of further action: repeating the tests, confirming the findings, establishing the direction/rapidity of change, initiating studies to establish the cause of the injury and preserving serum samples for retrospective testing (Senior, 2012).

Predictable hepatotoxic reactions Intrinsic hepatotoxicity is predictable in the sense that it is dose dependent and explicable. It can have either a direct or indirect mechanism, either of which can lead to cytotoxic or cholestatic liver injury.

Example: Acetaminophen

The most common cause of intrinsic hepatotoxicity is acetaminophen. It is often assumed that acetaminophen fatalities are all suicides, but in fact not all acetaminophen fatalities are due to suicidal intent – in fact less than 50% of cases are attempted suicides. A very clear-cut example of inadvertent fatal hepatotoxicity due to acetaminophen is the tragic death of Danielle Welsh.[5] Danielle was born on 5 June 1989. She had an undefined condition which gave rise to spondyloepiphyseal dysplasia, short stature, mild but longstanding learning difficulties, problems with hearing and chronic pain particularly in the limbs and joints. In June 2008 she weighed 35 kg. Danielle became unwell on the 15 June and her parents took her to the Southern General Hospital, Glasgow. Intravenous acetaminophen (paracetamol) was prescribed on the basis that Danielle was chronologically an adult. However, the dose prescribed (1000 mg four times daily i.v.) was way in excess of the appropriate dosage for someone of Danielle's weight and no-one had appreciated that the intravenous dosage had different parameters from the oral dosage: she should have received no more than 525 mg per dosage. Danielle was seen by 11 nurses and 12 different doctors and received 20 doses of acetaminophen from doctors at all levels, but not one of them noted the overdose. As a consequence, Danielle suffered a fatal cardiac arrest 9 days after admission.

[5] See the report of the Fatal Accident Inquiry into her death: http://www.scotland-judiciary.org.uk/10/715/Fatal-Accident-Inquiry-into-the-death-of-Danielle-Welsh

Example: Panadiplon

Panadiplon is a non-benzodiazepine anxiolytic which was originally developed for the treatment of generalised anxiety disorder and panic disorder. No evidence of toxicity emerged during preclinical development, but serum transaminase elevation was observed during Phase I clinical trials. The rabbit was identified as the best animal model for this toxic effect. In the rabbit, hepatotoxicity was characterised by hepatic triglyceride deposition and centrilobular necrosis. Studies using cultured hepatocytes indicated inhibition of mitochondrial fatty acid β-oxidation and transport systems (as measured by rhodamine 123 transport). These effects were also observed in human, but not rat, hepatocytes and were associated with metabolism to form cyclopropane carboxylic acid. Inhibition of mitochondrial β-oxidation was not, in itself, sufficient to induce toxicity: an additional insult such as hypoxia was also required. On the basis of these data, development of this drug was discontinued.

Example: Diclofenac

Diclofenac is a widely used non-steroidal anti-inflammatory drug which is safe when used within the therapeutic dose range. However, overdosing and long-term treatment can result in severe hepatotoxic effects. Mild liver dysfunction is observed in 15% of patients and, in a minority of these, can become severe. In 4% of patients ALT/AST levels reach 3× ULN and in 1% levels 8× ULN are reached. In very rare cases, severe hepatocellular necrosis and inflammation associated with haematological toxicity are observed but it is unclear whether this is due to idiosyncratic toxicity, direct cytotoxicity or immune-mediated hypersensitivity.

Diclofenac is metabolised to 4′-hydroxy and 5-hydroxy metabolites by CYP2C9 and CYP3A4, respectively. It also undergoes glucuronidation. In an effort to characterise the species-specific hepatic metabolism of diclofenac, precision-cut liver slices from rats, monkeys and humans were exposed to [^{14}C]diclofenac for 2, 24, 48 or 72 h. The results indicated that diclofenac was metabolised to a greater extent in rat liver slices than in slices from human or monkey liver, and rat liver slices were correspondingly more susceptible to protein binding and the diclofenac cytotoxicity. In precision-cut human liver slices the major product was the 4′-hydroxy metabolite. The extent of specific protein binding was lower in human liver slices and less cytotoxicity was observed in the human liver slices. Thus experiments using an *in vitro* liver slice model demonstrated concentration- and time-dependent increases in specific protein binding, which correlated with the extent of diclofenac metabolism, and altered liver slice function. This suggests that the hepatotoxic effects of diclofenac are explicable in terms of its tissue- and species-specific metabolism leading to the formation of reactive intermediates which induce oxidative stress and cellular toxicity.

Idiosyncratic toxicity The liver is the major target organ for idiosyncratic drug toxicity, whereby individuals exhibit an unpredictable, severe and often delayed

reaction to a drug.[6] The rare incidence of idiosyncratic effects (1 in 10 000 to 1 in 100 000) demonstrates the multifactorial nature of hepatotoxicity and variation of individual response. Idiosyncratic reactions occur at a rate of between 1 : 1000 and 1 : 100 000, usually with a latency of 5–90 days, and lead to liver transplantation or death in 75% of cases (Leitner, Graninger, and Thalhammer, 2010).

Although idiosyncratic hepatotoxicity is a rare event on a population basis, it is currently the main reason for FDA-mandated warnings, restrictions of use and withdrawal of drugs from the market. It is observed in less than 5% of exposed subjects and is usually unrelated to the drug's pharmacological effects. It is unpredictable, exhibits no clear dose-dependency, is often associated with polypharmacy and may be either metabolic or immunological in aetiology. Additional factors (e.g. metabolic polymorphisms) may influence risk, as can the presence of specific histocompatibility antigens. This is consistent with the observation that many idiosyncratic hepatotoxic reactions are immunologically mediated hypersensitivity reactions which have a host-dependent component.

There are some warning signs that a drug candidate could induce idiosyncratic toxicity. These include the following:

- Structural alerts
- GSH depletion and the formation of glutathione-drug/metabolite conjugates
- Covalent binding
- Bioaccumulation in the liver
- DDIs
- Induction of toxicity genes
- Ames/micronucleus assay positive
- Liver injury in the rat and/or higher species without a safety margin at steady-state concentrations of drug in the liver
- CYP induction and liver hypertrophy without a safety margin

Idiosyncratic hepatotoxicity is, by definition, rare and difficult to predict from animal studies. It tends to exhibit erratic temporal and dose–response relationships and can have a long latency. Early onset tends to suggest a direct effect of the drug or a metabolite, whereas long latency may indicate an indirect mechanism. In man, unpredictable severe idiosyncratic reactions may occur against a background of mild, asymptomatic liver injury.

The most common causes of idiosyncratic DILI are antibacterial drugs (particularly tuberculostatic agents) and non-steroidal anti-inflammatory drugs. However, medications with an idiosyncratic potential frequently cause only mild and harmless increases in transaminases; only about 15% of these extend beyond 3× the ULN.

Example: Trovafloxacin

The fluoroquinolone antibiotic trovafloxacin is a broad-spectrum antibacterial agent which is mainly excreted via the faeces in rats, dogs and humans. It only undergoes minimal CYP-mediated metabolism but is subject to glucuronidation,

[6] Reviewed by Peters (2005).

N-acetylation and metabolism via other minor conjugative pathways. Its major pathway of excretion, as suggested by the route of elimination, is biliary excretion of the glucuronide.

In preclinical testing, trovafloxacin only caused minimal to mild fatty change in male rat and dog liver, and this was reversible even at 10× the highest human dose. In human clinical trials liver enzyme release or cholestasis was observed in a few patients, but only one case of hepatic necrosis occurred. However, following the marketing of trovafloxacin (as Trovan) in the United States in 1998, the FDA began receiving reports of patients who experienced serious liver reactions in association with use of the product. Some of these patients developed serious liver injury leading to the need for a liver transplant and/or death.[7]

Follow-up studies indicated that, in mice, non-hepatotoxic inflammatory stress induced by TNFα synergises with trovafloxacin to induce liver injury. Trovafloxacin prolonged the plasma peak of TNFα due to reduced clearance, which was independent of kidney dysfunction, consistent with a generic mechanism of idiosyncratic toxicity in which drugs can synergise with a subclinical inflammatory stress, precipitating an idiosyncratic response. Trovafloxacin can synergise either with a non-hepatotoxic dose of LPS or with directly injected TNFα to precipitate TNFα-mediated hepatic injury. When male C57BL/6J mice were dosed with trovafloxacin (150 $mg kg^{-1}$ p.o.) with or without TNFα (50 $\mu g\, kg^{-1}$ i.p.), neither substance induced ALT release when dosed on its own but the combined dose led to marked elevations in ALT within 3 h of dosing (Shaw *et al.*, 2009). This was associated with midzonal hepatocellular necrosis and apoptosis, and *in vitro* studies in Hepa1c1c7 cells suggested that trovafloxacin sensitises hepatocytes to TNFα-induced apoptosis by activating intracellular apoptotic signals such as caspase activation. Trovafloxacin pretreatment exacerbated TNF-induced increases in plasma cytokine concentrations and hepatic neutrophil accumulation, but toxicity was prevented by the TNF-blocking agent Etanercept (which is used to treat autoimmune disease). On the other hand, dual injections of TNFα (to mimic the delayed clearance in response to trovafloxacin) prolonged the TNFα peak and mimicked TNFα/trovafloxacin toxicity.

Example: Troglitazone

The PPARγ agonist troglitazone, which was used to control Type II diabetes, was withdrawn in 2000 because it caused rare life-threatening idiosyncratic hepatotoxicity in humans. This toxicity was thought to have arisen as a consequence of the production of reactive quinone, quinone epoxide, quinone methide and/or reactive isocyanate metabolites. The major form of troglitazone detected in humans is the sulfate conjugate, which can reach 7–10 times the concentration of the parent compound. Other metabolites are found at lower concentrations, the quinone metabolite being found at a concentration similar to that of the parent compound while the glucuronide is found at very low concentrations. Troglitazone induces its own CYP-mediated metabolism catalysed by CYP3A4.

[7] http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/DrugSafetyInformationforHeathcareProfessionals/PublicHealthAdvisories/UCM053104

The results of preclinical studies were as follows.

- In monkeys, troglitazone was cholestatic and caused bile duct hyperplasia at doses close to those experienced by humans. It was not possible to define a NOAEL for this effect.
- In dogs, elevated ALT levels, plasma volume expansion and cardiomegaly were observed.
- Increased fatty marrow was observed in all species.

In clinical trials, 1.3% of trial subjects developed ALT levels more than 3x ULN and in 1% the level observed was greater than 8× ULN. Even after approval, several cases of severe hepatotoxicity occurred and a number of patients died as a result of acute liver failure. Liver toxicity was associated with elevated serum enzymes, hyperbilirubinaemia and clinical signs of hepatocellular necrosis/cholestasis. These phenomena appeared at inconsistent times after the initiation of therapy with no consistent dose–response and had no obvious mechanism. Troglitazone hepatotoxicity was particularly problematic because its onset was too rapid even for regular monitoring to help in early diagnosis or identifying individuals at risk. Following notification from the United States and Japan of 135 cases of serious liver toxicity and six deaths, GlaxoWellcome voluntarily withdrew troglitazone with effect from 1 December 1997.

Studies using human hepatocytes demonstrated that metabolic activation by CYPs was central to the hepatotoxicity of troglitazone, and mechanistic studies indicated that high concentrations of troglitazone have the capacity to induce mitochondrial stress in cultured cells. Results obtained using $Sod2^{+/-}$ mice (which are unable to regenerate NADPH from $NADH/NADP^+$, thus exacerbating mitochondrial oxidative stress) suggested that therapeutic doses of troglitazone cause subclinical stress to liver mitochondria and superimpose oxidant stress, potentiating mitochondrial damage and inducing delayed hepatic necrosis in mice with genetically compromised mitochondrial function (Jaeschke, 2007; Ong, Latchoumycandane, and Boelsterli, 2007). Inherited or acquired mitochondrial abnormalities may exacerbate susceptibility to troglitazone-induced liver injury, and by extension an underlying genetic defect may predispose some individuals to DILI because liver and mitochondrial targeting drugs may superimpose a second hit in a pre-existing condition within the same subcellular compartment (Ong, Latchoumycandane, and Boelsterli, 2007). Thus, underlying genetic or acquired mitochondrial defects may represent a risk factor for the adverse effects of mitochondrial-targeted drugs; further use of animal models with endogenous antioxidant deficiencies may uncover potential liabilities at therapeutic doses (Jaeschke, 2007), but cannot provide a complete explanation of the idiosyncratic toxicity observed in humans. A defect in oxidant defences may contribute to susceptibility but a single genetic defect is insufficient, on its own, to precipitate actual liver failure.

Example: Flutamide

Another drug whose idiosyncratic toxicity seems to be associated with mitochondrial dysfunction is flutamide, a non-steroidal antiandrogen which is used in the treatment of prostate cancer. Flutamide is associated with

idiosyncratic toxicity involving cholestasis or overt hepatocellular injury which sometimes develops to fulminant hepatitis, and mitochondrial dysfunction seems to be the most likely mechanism.

When wild-type and $Sod2^{+/-}$ mice were treated with flutamide at 30 or 100 $mgkg^{-1}$/day for 28 days (sufficient to induce clinically silent oxidative stress in wild-type animals), 5/10 $Sod2^{+/-}$ mice exhibited small foci of hepatocellular necrosis and an increased number of apoptotic hepatocytes (Kashimshetty *et al.*, 2009). This was associated with GSH depletion, protein carbonyl modification and serum lactate levels consistent with oxidative stress and mitochondrial dysfunction. In contrast, no hepatic changes were observed at 50 $mgkg^{-1}$/day either in wild-type or $Sod2^{+/-}$ mice with the flutamide analogue bicalutamide, for which there has only been one report of DILI. These observations suggest that flutamide hepatotoxicity is only triggered when the antioxidant defence system of the liver has been compromised. The transcriptional changes which occurred in response to flutamide in wild-type and $Sod2^{+/-}$ mice were similar, suggesting that the initial oxidative challenge was similar in the two strains but the outcome (oxidative damage to mitochondrial DNA, membrane lipids and proteins and exacerbation of mitochondrial oxidative stress) is determined by the status of antioxidant defences. The source of flutamide-associated oxidative stress may be increased generation of superoxide, redox cycling and/or production of reactive metabolites by CYP1A2 (possibly induced via AhR activation), CYP3A4 and CYP2C19.

The fact that several examples of idiosyncratic toxicity have been shown to involve mitochondrial dysfunction suggests that this could be a generic mechanism. However, it should by no means be assumed that it is universal; it is more likely that initial observations, plus the availability of a suitable genetically modified mouse model, has led investigators to focus on this particular mechanism. Metabolic idiosyncrasies and immune-mediated reactions can also lead to idiosyncratic toxicity, and no doubt other mechanisms are yet to emerge.

10.3.3 Kidney

Considerable effort has been devoted to the identification of clinical biomarkers for the diagnosis and management of acute kidney injury (Molitoris *et al.*, 2008). There is an acute need for biomarkers to allow more reliable diagnosis of drug-induced acute kidney injury, easier monitoring of therapeutic strategies and differential diagnosis of specific forms of injury (Ozer *et al.*, 2010).

Conventional biomarkers of renal damage Conventional biomarkers of renal toxicity include creatinine and blood urea nitrogen (BUN) as well as biomarkers of functional deficit and proximal tubule dysfunction (e.g. albumin) and biomarkers lost from dead or injured cells (e.g. GSTα)

- **Creatinine** is commonly used as a surveillance marker for acute kidney injury, but it correlates poorly with histological injury in rats and humans and many factors that are independent of changes in renal function (including age,

race, muscle mass, metabolic state, dehydration and muscle injury) can affect plasma creatinine. Creatinine values may be compromised by loss of muscle in critically ill patients and are affected by systemic production and tubular secretion, meaning that it is not specific to tubular injury.
Basal serum creatinine correlates with glomerular filtration but it becomes more and more dependent upon tubular secretion as glomerular filtration rate decreases. In addition, some drugs and organic compounds can block creatinine excretion. Furthermore, the onset of changes in creatinine may be delayed even after significant injury. Even large changes in glomerular filtration rate may be associated with only small changes in serum creatinine during the first 24–48 h after injury, meaning that critical therapeutic windows may be missed. Conversely, when renal function improves, serum creatinine underestimates glomerular filtration rate until a new steady state is reached and can be retarded further by enhancing hydration status. Finally, there is considerable interindividual variation in the relationship between serum creatinine and basal glomerular filtration rate, the size of the renal reserve and rates of creatinine synthesis. In addition, basal levels of creatinine in rodents are close to the limit of detection, meaning that a large increase is required before the change becomes detectable.

- **Blood urea nitrogen** is affected by many factors. Urea is freely filtered by the glomerulus, but urea is reabsorbed to differing extents by other parts of the nephron and changes in BUN can be observed due to volume depletion in the absence of tubular injury. Increases in BUN can occur in response to increased production of urea, for example, during protein supplementation, catabolic states or when blood is present in the GI tract.

The standard criteria for acute kidney injury are based on serum creatinine and urine output, and injury is staged according to the so-called RIFLE criteria (risk, injury, failure, loss, end-stage). However, the routine assessment of kidney function based on serum creatinine and BUN associated with a fall in urine output is now considered outmoded (Basnakian, 2008). In particular, serum creatinine is a poor biomarker of acute kidney injury because the patients are not in steady state and increases in serum creatinine lag behind changes in renal function.

Better biomarkers of kidney injury are required in both translational and clinical settings. Sensitive, specific and quantitative biomarkers are required in order to facilitate early diagnosis because the early stages of kidney injury lack specific signs and symptoms and the available biochemical tests tend to be non-specific and insensitive. Translational biomarkers (i.e. markers which are detectable in both humans and experimental animals) should be able to predict toxicity sensitively both in preclinical and clinical situations in order to guide drug development.

Second-generation biomarkers The ideal biomarker for drug-induced kidney injury would:

- identify kidney injury early, well before the renal reserve is dissipated and levels of serum creatinine increase;
- reflect the degree of toxicity and be able to characterise dose dependencies;

- display similar reliability across multiple species including humans;
- localise the site of kidney injury;
- track the progression of injury and recovery;
- have well-characterised (and minimal) limitations;
- be accessible in readily available body fluids or tissues.

Second-generation biomarkers have the potential to transform the way we detect and quantify nephrotoxicity, although it must be noted that a biomarker for one kind of kidney injury may not be informative regarding other types of injury (e.g. early injury vs delayed repair; inflammation vs proximal tubular damage).[8] Among the most promising of the new generation of biomarkers are actively secreted proteins which are induced or repressed as a result of injury. They include the following:

- **KIM-1:** an extracellular protein which is normally anchored in the membranes of proximal convoluted tubule cells but can be cleaved by metalloprotease and excreted into urine. This biomarker responds both sensitively and dynamically to proximal convoluted tubule injury.
- **Clusterin:** a glycoprotein secreted by a variety of cell types, notably the tubular cells of the kidney.
- **Trefoil factor 3:** a small secreted mucin with hormonal activity which promotes the survival and differentiation of epithelial cells and shows reduced urinary excretion in response to acute kidney injury.
- **Cystatin c:** This non-glycosylated low-molecular weight (13 kDa) protein is a serum marker of renal function. This housekeeping protein is produced by all nucleated cells and undergoes free and direct glomerular filtration. It is a good marker for glomerular filtration rate because it is not affected to the same extent as creatinine by age, sex, muscle mass, dehydration state and circadian rhythm. Its filtration mechanism is simple and it exhibits no tubular secretion or extrarenal clearance. The FDA has approved an assay for cystatin c.

A comparative study on urinary biomarkers in renal tubular injury induced by cisplatin and a glomerular injury model induced by puromycin aminoglucoside has analysed the accuracy and sensitivity of different markers (Tonomura *et al.*, 2010). Histopathological analysis was used to confirm typical tubular injury at the corticomedullary junction in response to cisplatin and glomerular injury in response to puromycin aminoglucoside. The results were as follows.

- KIM-1 exhibited perfect accuracy and sensitivity and also exhibited the earliest and largest time-dependent increase in urine. KIM-1 protein production is associated with dedifferentiation of the renal epithelium and is an early response to injury and the onset of repair processes. It is a good marker for histological damage to the proximal convoluted tubule.
- β2-microglobulin and cystatin c are functional biomarkers whose low molecular weight means that they are normally almost completely reabsorbed. Thus, elevated levels of these proteins in the urine suggest that reabsorption in the convoluted tubules has been compromised. This tends to be a delayed response compared with KIM-1.

[8] For further discussion of potential second-generation biomarkers of kidney injury, see Bonventre *et al.* (2010).

- GSTM also gave perfect accuracy and sensitivity with a time-dependent increase in urinary levels, but the magnitudes of the changes observed were smaller than for KIM-1, while GSTα exhibited the characteristics of a specific marker for proximal tubular damage.
- LDH is expressed all the way along the nephron and can detect a broad range of nephrotoxicity along its whole length.
- Albumin can be increased via two mechanisms, loss of the charge/size restriction of the glomerular capillary wall and impaired reabsorption by the proximal convoluted tubule due to dysfunction or toxic injury.

In this study, KIM-1 was found to be a reliable and sensitive biomarker for proximal convoluted tubule damage and LDH proved to be useful for broad range screening of damage along the nephron. Although often used as markers for glomerular toxicity, β2-microglobulin and cystatin c were found to be poor markers for glomerular damage.

In 2008, the EMA indicated that the urinary biomarkers KIM-1, albumin, total protein, β2-microglobulin, clusterin, trefoil factor 3 and cystatin c were acceptable markers of nephrotoxicity in the context of non-clinical drug development.[9] They can be used to detect acute drug-induced nephrotoxicity, either tubular or glomerular with associated tubular involvement, and are considered complementary to BUN and creatinine. It was, however, noted that while these markers may merit further exploration as clinical biomarkers for acute drug-induced kidney injury, at present their use for monitoring nephrotoxicity in the clinical setting cannot be recommended.

Additional biomarkers have also been suggested. One example is L-type FABP, which is localised to proximal convoluted tubule cells in the human kidney but has been shown to be released to the urinary space during ischaemia-reperfusion injury in kidney transplantation. A humanised FABP mouse model has been used to evaluate the potential of FABP as a biomarker (Negishi *et al*., 2009) of renal injury. Mice were treated with cisplatin (5, 10 or 20 $mgkg^{-1}$ i.p.) or ischaemia-reperfusion. In the cisplatin model, BUN started to increase 48–72 h after dosing with cisplatin, whereas urinary FABP increased within 2 h and could discriminate between doses at all subsequent time points. In the ischaemia-reperfusion model, BUN did not increase until 24 h but urinary FABP started to increase after 5 min and peaked at 3 h. The dynamic range for FABP was sufficient to discriminate between different periods of ischaemia and, in both test systems, FABP gave a good correlation with histological alterations and changes in glomerular filtration rate.

One aspect of renal biomarker characterisation which has only recently been considered is reversibility: do biomarker levels decline as the injury is repaired? This is important because acute necrosis of moderate numbers of proximal convoluted tubule cells is reversible and regeneration of the proximal convoluted tubule layer restores the integrity of the tubule. Thus, limited tubular injury is a reversible form of kidney injury associated with a limited inflammatory response. Comparison of a broad panel of urinary biomarkers has identified those which allow renal

[9] http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004205.pdf

injury to be monitored both at the onset of toxicity and when it begins to reverse after withdrawal of the renal toxin (Ozer *et al.*, 2010). GSTα was shown to be an excellent early biomarker for epithelial necrosis, whereas KIM-1 and clusterin levels persisted during regeneration, reflecting repair processes. In the same study, serum cystatin c shows improved sensitivity and specificity compared with BUN and creatinine, allowing drug-induced injury in various compartments of the kidney to be measured. In addition, the fact that cystatin c is measured in serum rather than urine means that it is convenient for preclinical studies (although conversely this may limit its usefulness in a clinical setting).

The protein netrin-1 has been proposed as a universal biomarker of renal injury and recovery (Reeves, Kwon, and Ramesh, 2008). Netrins were initially identified as neuronal guidance cues. They are laminin-like molecules with a distinctive domain organisation which belong to the laminin-related family of axon-guidance molecules. Netrins are also involved in the development of the mammary gland, lung, pancreas and blood vessels. It is usually expressed in peritubular capillaries and is highly induced in post-ischaemic kidney tubular epithelial cells. Netrin-1 is highly induced in tubular epithelial cells as early as 3 h after the onset of renal ischaemia-reperfusion injury and peaks at 24 h; it is also excreted in the urine of patients with acute renal failure (Reeves, Kwon, and Ramesh, 2008). In mice, netrin-1 was also released in response to cisplatin (20 $mgkg^{-1}$ i.p.; 30-fold after 6 h), LPS (5 $mgkg^{-1}$ i.p.; 160-fold after 6 h) and folic acid (250 $mgkg^{-1}$ i.p.; 40-fold after 3 h).

Urinary netrin-1 levels peaked and fell rapidly with all treatments except for folic acid, which induced sustained release. The mechanism responsible for netrin-1 release appears to be post-transcriptional, involving changes in translation and/or release of presynthesised protein. The fact that urinary levels of netrin-1 return to normal during reperfusion indicates that it can be used as a prognostic marker for renal recovery in clinical settings such as delayed graft function following kidney transplantation, hypotension in dialysis patients and following cardiovascular surgery and nephrotoxicant or sepsis induced renal injury (Basnakian, 2008).

10.3.4 Circulating mRNA biomarkers

It has recently been shown that RNA molecules are released into biofluids in various forms. This can occur via various mechanisms including passive leakage of cellular mRNA following necrosis, during which RNA molecules enter the circulation either in native form or in association with cellular debris, or encapsulated in membrane vesicles which are actively released from cells. These vesicles include the following.

- Exosomes (40–100 nm diameter) are formed by inward budding of endosomal vesicles which are subsequently packaged into larger multivesicular bodies. These then release their contents into the extracellular environment by exocytosis.
- Shedding vesicles (<200 nm) are released from live cells by direct budding from the plasma membrane.
- Apoptotic blebs (100–1000 nm) bud directly from the plasma membrane during apoptosis.

After release from the cell, these vesicles circulate in the extracellular space. Most of them are broken down within minutes because they have inside–out orientation, leading them to be targeted for degradation because of the presence of phosphatidylserine on the outside of the boundary membrane, but a small proportion moves into the circulation and appears in the biofluids.

Release of RNA does occur under homeostatic conditions, possibly playing a role in cell–cell communication, and circulating RNA is found in healthy individuals, but the pattern and amount changes during toxicity. Circulating RNA has been shown to be a useful biomarker for a number of clinical endpoints, including as a predictor of mortality in acute trauma patients and as a diagnostic marker in pre-eclampsia, complications of diabetes and cancer.

The potential of extracellular RNAs as biomarkers of liver injury has been investigated in rats using two hepatotoxins (galactosamine and acetaminophen) and a non-liver toxin (bupivacaine, a skeletal muscle toxin) (Wetmore *et al.*, 2010). Liver-specific mRNA-containing materials in plasma were characterised by quantitative RT-PCR, sucrose density gradient centrifugation and electron microscopy. Serum enzyme levels, histopathology and transcriptional profiling were used to evaluate effects in the liver itself. The mRNAs encoding albumin, fibrinogen, haptoglobulin and β-polypeptide were assayed and found to be released in response to galactosamine and acetaminophen, but not bupivacaine. The changes in circulating mRNA levels were more sensitive than either liver enzyme release tests or histopathological analysis. They were mostly found in the exosomal fraction although at high doses they were associated with cell debris. Analysis of a wider range of extracellular RNAs by whole genome microarray analysis indicated treatment-specific profiles of RNA release from the liver.

The presence of detectable circulating mRNA levels in control animals indicated that mRNA release is a normal physiological process, but increased levels were seen in treated animals prior to the onset of overt toxicity. This was actually associated with decreased transcription in the liver, indicating that the increased plasma levels were due to increased release rather than up-regulation of expression. These results suggest that circulating mRNA has the potential to be a valuable biomarker of liver damage, but the interpretation of the data is complicated by the mixture of passive and active mechanisms of release and the potential for the same mRNA to be released from different tissues at different rates, so further work is required before their routine use can begin.

In a follow-up study of human DILI, 11 patients suffering from DILI were tested 72 h after reaching peak ALT/AST values and RNA release was compared with that in 10 controls. Albumin, fibrinogen and haptoglobulin mRNAs were elevated although the results obtained were more variable than with ALT, suggesting that further development of RNA markers will be required before they can be used clinically or in mechanistic studies. However, it does seem that RNA markers may ultimately be more sensitive and specific than enzyme markers and they have a number of potential strengths as biomarkers.

- Convenience – they are found in the circulation so do not require a tissue biopsy.
- Sensitivity – they can be detected by quantitative RT-PCR.

- Organ specificity – tissue-specific RNA markers can be identified and assayed specifically.
- Treatment specificity – markers of the effects of individual drugs can be identified.

In the long term, RNA biomarkers may provide a virtual biopsy in that tissue-specific microvesicles can be isolated and subjected to RNA characterisation, for example, by microarray analysis.

10.3.5 Evaluation

The traditional approach to biomarkers has been to attempt to identify individual proteins which are found circulating in the blood or excreted via the urine. Various groups are now extending this concept by looking at other types of molecule (e.g. circulating mRNA) or by incorporating biomarkers into a systems biology approach using, for example, MS-based metabolomic analysis of urine (Kumar *et al*., 2012) and, while it is considered unlikely that the use of biomarkers will ever completely replace traditional histopathological evaluation in the identification and characterisation of adverse responses (Boekelheide and Schuppe-Koistinen, 2012), the non-invasive nature and quantifiability of biomarkers from body fluids mean that they will have an increasing role to play in the identification of adverse effects in the laboratory and in translation to the clinic.

10.4 The systems biology approach

Systems biology is concerned with the identification of molecular networks as key components of cellular (dys)functioning in metabolism, signalling and genetic networks. It addresses entire biological systems and uses a wide range of approaches, integrating biological knowledge, clinical information, mathematical modelling, computer simulations, imaging and 'omics technologies. Thus it is complementary to the conventional reductionist approach of breaking the system down into its component parts.

This concept is by no means new: over 70 years ago Beadle and Tatum (1941) stated that '....an organism consists essentially of an integrated system of chemical reactions controlled in some manner by genes' and predicted that the levels at which such a system might be controlled would range 'from simple one-to-one relations to relations of great complexity'. Indeed, systems biology has been practised in various forms since the 1940s, although the discipline, as it is now recognised, began in earnest at the beginning of the twenty-first century. The factors triggering this expansion have included increases in computing power, the introduction of genome-wide measurement methods for important biological molecules and the publication of the entire sequence of the human genome.

Systems biology is a multidisciplinary approach which takes advantage of developments in fields as diverse as molecular analysis, imaging technology, mathematics, physics, engineering and mathematical modelling and involves the

acquisition and integration of multiple types of data, including but not limited to 'omics profiles.

The fundamental premise of systems biology is that a complete understanding of a molecular network requires not just knowledge of its parts but also an understanding of their dynamic interactions. By adopting a variety of perspectives, systems biology can generate a multiplex picture of a complete biological system. The results thus obtained are subjected to detailed statistical analysis and arranged into networks which are used to generate new hypotheses which can then be tested experimentally.

Systems biology may adopt either a bottom-up or a top-down approach to the analysis of data (Cooke, Savage, and Wild, 2009).

- The *bottom-up* approach is consistent with the conventional scientific method in that it begins with analysis and description of the biochemical processes involved in the process being studied. The resulting model is usually expressed as a series of differential equations describing, for example, a regulatory pathway. Thus, the 'bottom-up' approach involves combining the results of experimental research with mathematical modelling.
- The *top-down*, or data-driven, approach involves collecting global information (e.g. data about all genes, all proteins or all metabolic interactions) and assimilating them into a comprehensive database. The process being studied (e.g. a regulatory network) is then reverse-engineered from these HTS data. In this top-down approach, a model of cellular behaviour is reverse-engineered based upon inference, allowing the generation of models based purely upon the information available in the data, without the bias of prior knowledge or preconceived ideas about the network structure.

The vast amount of data generated in systems biology programmes and their scope for generating new biological insights creates a dual need, for open access data-handling and visualisation software (e.g. Cytoscape[10] (Smoot *et al.*, 2011); Figure 10.3) and for investigators with in-depth biological knowledge that allows them to define the most appropriate questions and hypotheses.[11] Furthermore, there is a critical need for international data standards to allow results to be shared and integrated between investigators and in the public arena. Systems biology strives to understand the composition, structural organisation and dynamic behaviour of cells under different conditions, addressing complex networks of physical and functional interactions that go beyond the traditional emphasis on linear pathways. It focuses on the so-called 'emergent properties' of a particular system (Short, 2009). These properties are difficult, if not impossible, to predict using the traditional approach of studying a few molecules at a time.

Example: Cancer

Systems biology has huge potential in the area of cancer research because it is uniquely placed to enhance our understanding of the many complex ways in which

[10] http://www.cytoscape.org/

[11] For a detailed review of approaches to the visualisation of complex data, see Afshari, Hamadeh, and Bushel (2011).

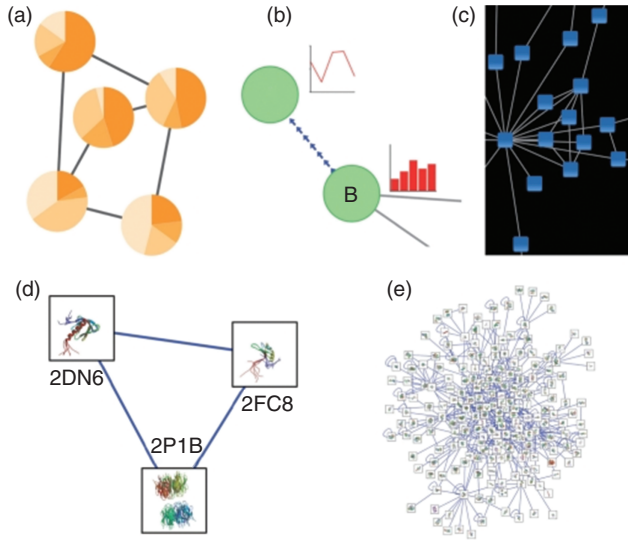


Figure 10.3 Rich network visualizations enabled by the new Cytoscape features. Simple networks are shown with custom node images based on (a) pie chart displays or (b) line plots and bar charts generated using Google's Chart API. (c) Nodes have a transparent custom graphic to give the appearance of shading. (d and e) Protein–protein interaction networks in which each node contains a 3D image of the protein structure of the protein represented by the node. (source: Smoot *et al.* (2011); figure 1. Reproduced by permission of Oxford University Press)

things can go wrong in the regulation of cellular homeostasis and cell number. Indeed, cancer can be considered to be a systems biology disease.

The problem has been neatly summarised as follows: 'In order to understand how a cancer cell is functionally different from a normal cell it is necessary to assess the complex network of pathways involving gene regulation, signalling and cell metabolism and the alterations in its dynamics caused by the several different types of mutation leading to malignancy' (Laubenbacher *et al.*, 2009). The key is to consider the overall network and changes in its dynamics rather than focusing on individual pathways.

Taking a systems biology view of the hallmarks of cancer, the key properties of cancer are the following:

- Independence from external growth signals.
- Insensitivity to anti-growth signals and evasion of apoptosis.
- Limitless replicative potential.
- Sustained angiogenesis and metastasis.

The most productive way to think about this is as a process of dysregulation of a multi-pathway network which can make connections within and beyond the tumour cell, including to other cells in the local environment and the components of the extracellular matrix (Kreeger and Lauffenburger, 2010).

It will still be a long time before systems biology fulfils the prediction of Hanahan and Weinberg in the year 2000 that:

> Two decades from now, having fully charted the wiring diagrams of every cellular signalling pathway, it will be possible to lay out the complete "integrated circuit of the cell" .... We will then be able to apply the tools of mathematical modelling to explain how specific genetic lesions serve to reprogram this integrated circuit in each of the constituent cell type so as to manifest cancer.

Indeed, in email correspondence during the preparation of this book, Prof Weinberg said:

> As an aside, this statement of ours seemed to be a reasonable estimate of when this task will be accomplished, but now, 13 years later and 7 years before the 2020 deadline, I am skeptical that we will have wiring diagrams of most of the critical regulatory circuits in the cell. I suspect that Doug Hanahan would share my skepticism.[12] So, if you wish to quote us, you may or may not wish to position this as an overly optimistic assessment of the rate of future progress!

However, considerable progress has been made and Laubenbacher *et al.* (2009), at least, are optimistic that this is within our grasp.

10.4.1 Systems biology in toxicology

Systems biology approaches have a number of potential applications in toxicology (Marchan, Bolt, and Hengstler, 2012).[13] They can help researchers to integrate experimental measurements of genes, proteins, metabolites and drug compounds under consistent biological conditions. The construction of molecular connectivity maps makes it possible to integrate experimental measurements of genes, proteins, metabolites and drug compounds under consistent biological conditions, allowing researchers to compare the molecular, therapeutic and toxicological profiles of many drug candidates and potential targets simultaneously (Li, Zhu, and Chen, 2009). This can open up new possibilities in terms of toxicogenomics and drug discovery, including the following:

- Helping to define modes and mechanisms of action.
- Providing a rational basis for species-species and *in vitro–in vivo* extrapolation.
- Defining the mechanistic basis for differences in susceptibility to toxic insults.

Example: The Liver Toxicity Biomarker Study

This collaboration between the National Center for Toxicological Research and BG Medicine Inc., which was supported by 10 pharma and technology companies, focused on the biochemical effects of drugs. This US programme, which also involved a number of pharmaceutical companies, addressed compounds which did not exhibit adverse effects in preclinical studies but turned out to be hepatotoxic in the clinic. Its aim was to establish whether predictable and idiosyncratic hepatotoxicity is part of a continuum or qualitatively different.

[12] He emailed to say that he does!

[13] This has been the subject of special issues of *Archives of Toxicology* (introduced by Marchan, Bolt, and Hengstler (2012) and *Mutation Research* (introduced by Van Ravenzwaay *et al.* (2012)).

In order to address this question, a comprehensive molecular comparison of the events produced *in vivo* by pairs of compounds was undertaken. The compound pairs chosen were:

- Similar in structure and mechanism of action.
- Associated with few or no signs of toxicity in preclinical studies.
- Markedly different in terms of hepatotoxic potential.

In each pair, one compound was identified as the Clean Compound and the other as the Toxic Compound. In each pair, the compounds had similar chemical structures and identical primary target activities but the Toxic Compound was a withdrawn drug or a marketed drug which carries a warning label for hepatotoxicity, whereas the clean compound has no known hepatotoxicity. The ultimate aim of the project was to compare five such pairs by means of a comprehensive molecular systems analysis of liver, plasma and urine samples from a standard 28-day, three-dose Sprague-Dawley rat toxicity study with an additional sacrifice on day 4 and urine collections on days 0–4 and day 27.

Phase I, a single pairwise comparison, was published in 2009 (McBurney *et al*., 2009). This addressed the compounds entacapone and tolcapone, which are used for adjunct treatment of Parkinson's disease. In this pair, tolcapone is the Toxic Compound while entacapone is the Clean Compound. Conventional indicators of toxicity are all negative for tolcapone, but it caused four cases of acute hepatotoxicity among 60 000 patients. This compound was initially withdrawn then reissued with a warning label.

In the Liver Toxicity Biomarker Study, the effects of tolcapone and entacapone were analysed using eleven analytical approaches based upon clinical chemistry, transcriptional profiling, proteomics, lipid analysis and various LC–MS techniques (Figures 10.4 and 10.5). The authors state that 'Molecular analysis of the rat liver and plasma samples combined with statistical analysis of the resulting datasets yielded marker analytes, illustrating the value of the broad spectrum, molecular systems analysis approach to studying pharmacological or toxicological effects' (McBurney *et al*., 2009).

In 2007 the US National Research Council published a report entitled 'Toxicity testing in the 21st Century: A Vision and a Strategy'[14], which encouraged the discipline to focus on *in vitro* techniques and to adopt an approach based upon the concept of toxicity pathways. This represents a significant challenge since such pathways have not yet been defined!

One of the problems with this approach, which focuses on the identification of markers and endpoints at earlier stages than traditional outcomes, is that it can be very difficult to distinguish between a true adverse effect and a reversible adaptive response. An expert workshop held in May 2011 attempted to come up with clear definitions of these terms and agreed on the following, at least as working definitions (Keller *et al*., 2012):

- **Adverse effect:** A change in morphology, physiology, growth, development, reproduction or life span of a cell or organism, system or (sub)population that

[14] Available via http://www.nap.edu/catalog.php?record_id=11970

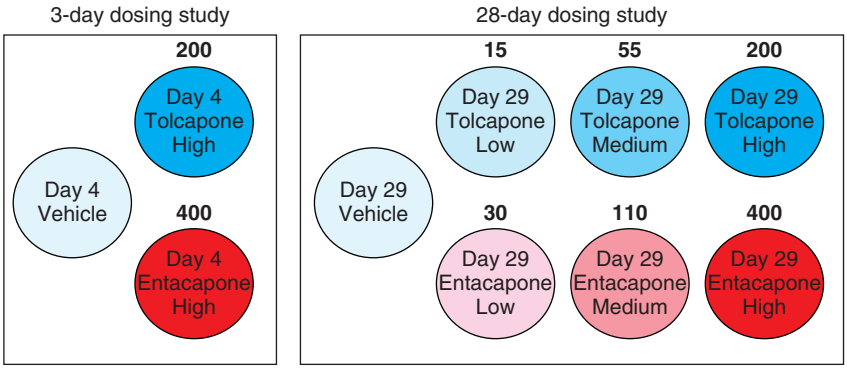


Figure 10.4 **A schematic of the experimental design of the 3-day dosing and 28-day dosing parts of the Liver Toxicity Biomarker Study.** The coloured circles represent the day of sacrifice samples that were generated for gene expression analysis, proteomic analyses, and metabolomic analyses. The $mgkg^{-1}$/day dosing for each drug for the study cohort represented by each coloured circle is shown in bold above the coloured circle. (source: McBurney *et al.* (2009); figure 1. Copyright © 2009 by Society of Toxicologic Pathology. Reprinted by permission of SAGE Publications)

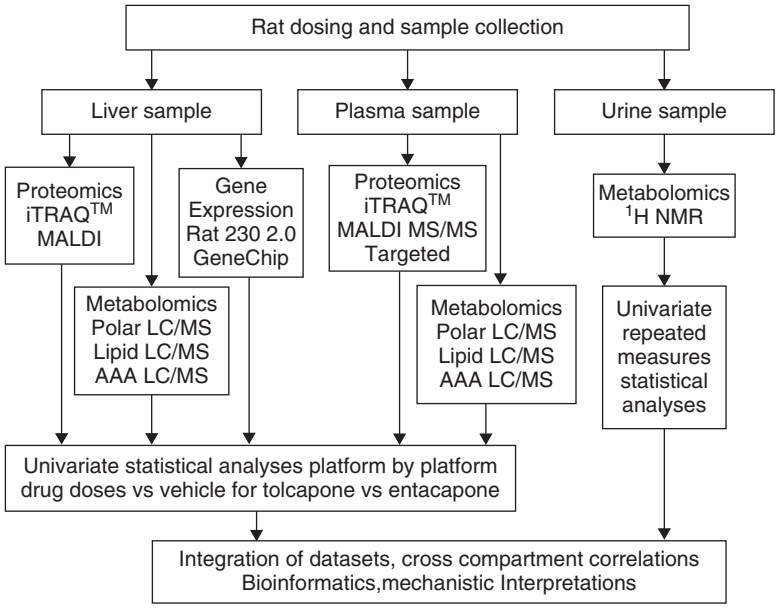


Figure 10.5 **Workflow of the Liver Toxicity Biomarker Study.** A schematic that illustrates the overall workflow of the LTBS beginning with the rat dosing and sample collection stage and ending with the integration of datasets and mechanistic interpretations of the results of the statistical analysis. The methods employed in each of the bio-analytical platforms applied to liver, plasma, or urine samples are described in (McBurney *et al.*, 2009) (source: McBurney *et al.* (2009); figure. Copyright © 2009 by Society of Toxicologic Pathology. Reprinted by permission of SAGE Publications)

results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences.

- **Adaptive response:** In the context of toxicology, the process whereby a cell or organism responds to a xenobiotic so that the cell or organism will survive in the new environment that contains the xenobiotic without impairment of function.

Adaptive responses are often early homeostatic adjustments to particular exposure such as a change in metabolism or gene expression; they may be sustained, an example being liver enlargement as a consequence of CYP induction and proliferation of the endoplasmic reticulum, but are ultimately reversible if the causative stimulus is removed. They can, however, still have toxic consequences, an example being thyroid carcinogenesis resulting from metabolic changes leading to altered regulation of thyroid hormone synthesis.

By considering two examples, dimethylarsinic acid and acetaminophen, the workshop participants identified the importance of defining toxicity pathways and proposed the term 'Relevant Pathway of Toxicological Concern'. They identified these as discrete biological mechanisms associated with toxicity, expressing the hope that there will only be relatively small number of them and that they will describe the majority of toxic responses. They noted, however, that the Relevant Pathway of Toxicological Concern concept is really only an indicator of hazard, and the workshop participants further proposed the concept of the Relevant Response for Regulation, a prescribed effect on which regulatory action can be based. The Relevant Response for Regulation is based upon the evaluation of a Relevant Pathway of Toxicological Concern in the context of exposure, dose and reversibility. The conclusion of the workshop was that 'understanding the spectrum of adaptation and adversity as it applies to risk assessment will ultimately inform regulatory decisions' and recommended a systematic effort to define and characterise Relevant Pathways of Toxicological Concern, focussing on the existence of dose transitions (i.e. thresholds).

The elucidation of toxicity pathways will inevitably involve the application of a systems biology approach in order to identify mechanisms of action and develop a quantitative framework for the interpretation of *in vitro* toxicity data. Furthermore, the application of systems biology should facilitate extrapolation from *in vitro* studies to human risk assessment (Edwards and Preston, 2008). The way in which systems biology data can be incorporated into the conventional risk assessment parallelogram in order to assist in this process is illustrated in Figure 10.6. The necessary steps in this procedure are the following:

- Screen representative compounds *in vitro* and suitable animal models *in vivo*.
- Undertake global biology measurements on target cells and tissues.
- Identify key events, for example, oxidative stress, growth factor signalling and cytotoxicity associated with cell proliferation.
- Create a key event network in which each mode of action becomes a subset of connected nodes associated with an adverse event at the molecular level.
- Use the key event network to make predictions which can be tested by direct measurement.

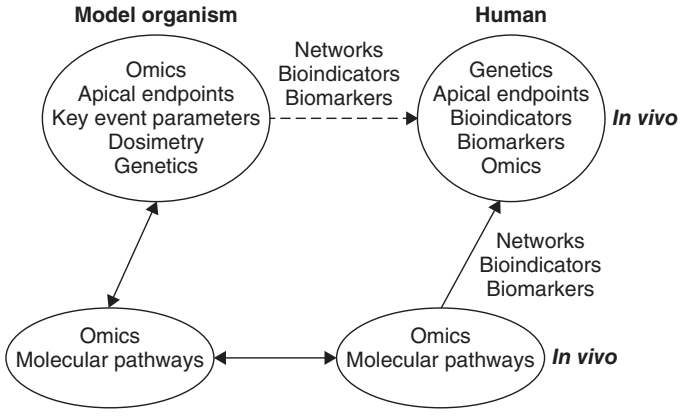


Figure 10.6 Schematic illustration of how systems approaches fit into the traditional parallelogram for risk assessment. Experimental approaches/measurements within circles represent data collected from that approach. Networks, bioindicators, and biomarkers represent outputs from model organisms and *in vitro* approaches which can be used for human risk assessment. The goal is to reduce the need for *in vivo* animal data as the knowledgebase containing network information and *in vitro* assays providing critical key event parameter values are developed (source: Edwards and Preston (2008); figure 1)

Once the key modes of action have been defined, the next step is to develop dose–response models using animal and, where available, human data. In addition, the information derived should be used to develop improved cell-based assays which addresses the key events in the *in vivo* mode of action and to identify biomarkers which can be used to evaluate risks to humans.

The application of systems biology will enhance the risk assessment process in a number of ways:

- Systems approaches contribute to the definition of modes of action by providing global biology measurements which can be correlated with disease parameters and known key events.
- The use of mathematical modelling approaches will help to prevent bias in data interpretation due to the application of preconceived ideas.
- A systems biology approach will allow the assessment of susceptibility to be explicitly linked to genetic polymorphisms and mechanisms of toxicity.

Overall, systems biology represents a new way of tackling problems in toxicology and has enormous potential in this context. This complex approach to the identification of markers and mechanisms is clearly expensive and technically challenging, but the resulting datasets are extremely information rich and the approach offers a feasible way of addressing fundamental questions in target organ toxicity. However, it does require a new funding mechanism because it is far too expensive and resource intensive to be undertaken by a single organisation. We are still a long way from having all the information needed in order to put the National Research Council recommendations into practice and work to implement them is still very much in progress (Stephens *et al.*, 2012).[15]

[15] See, for example http://humantoxicologyproject.org/about-2/; Stephens *et al.* (2012).

10.5 Collaborative programmes

The cost and complexity of all kinds of systems biology analysis make it impossible for individual companies and academic groups to generate the amount of reference and validation data required to engender confidence in the use of these methods to predict the toxicity and ADME properties of drugs and xenobiotics. The need to integrate systems biology results with conventional measures of toxicity and pathological processes makes this task yet more onerous. Furthermore, as the amount of primary target organ toxicity data and mechanistic information increases it is becoming more and more important to organise the available information and make it available for use by stakeholders. The cost and amount of work involved, together with the issues of confidentiality associated with the release of proprietary information, mean that these initiatives often take the form of collaborative programmes between government, industry and academia. Probably the most impressive examples are the InnoMed PredTox programme and its successor, the EU-sponsored Innovative Medicines Initiative (IMI) programme by the EU) and the Tox21 Program sponsored by the US Government.

10.5.1 Europe

The InnoMed PredTox programme, funded as part of the EU Framework Six programme and coordinated by the European Federation of Pharmaceutical Industries and Associations, involved 14 pharma companies, three academic institutions and two technology providers (Suter *et al*., 2011). The aim of InnoMed PredTox was to facilitate more informed decision making earlier in preclinical safety evaluation by combining 'omics results with conventional toxicology methods. Its secondary objective was the education of a population of scientists giving them the appropriate skills to take advantage of the systems biology approach. The 3-year project ran from October 2005 to January 2009 and focussed on three common findings in preclinical development: renal toxicity, bile duct necrosis and hepatocellular hypertrophy. As part of the project, the partners collected data on 14 failed drug candidates together with two reference compounds, troglitazone and gentamycin. Each was tested in a 2-week systemic study in Wistar rats at three doses and analysed using transcriptional profiling (using Affymetrix RAE 230-2.0 arrays), proteomics (2D-fluorescence difference gel electrophoresis of urine and SELDI), metabolomics (^{1}H–NMR and LC-MS/MS), histopathology and clinical chemistry (Figure 10.7). The compounds chosen had originally been developed for their potential to treat significant diseases such as diabetes and CNS disorders but then had to be excluded for therapeutic use due to serious adverse effects in liver and kidney tissues. The results indicated that the best combination of methods was transcriptional profiling combined with histopathology, supported by additional evidence from 2D-gel electrophoresis and LC-MS/MS as appropriate. These studies helped to provide further support for second-generation biomarkers such as clusterin and KIM-1, identifying them as sensitive, early non-invasive markers of renal injury in the rat and thus facilitating the process of extrapolation from experimental toxicity to real-life effects in the clinic (Hoffmann *et al*., 2010).

The results of the InnoMed PredTox programme were described as encouraging by senior project members who indicated that, while traditional methods such as

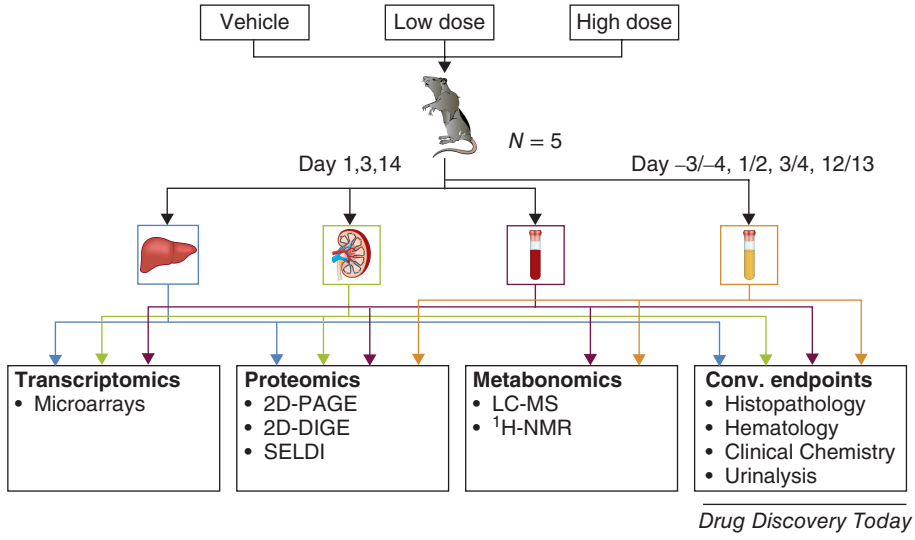


Figure 10.7 Experimental design overview of the InnoMed PredTox project. Sixteen compounds with known hepatotoxicity and/or nephrotoxicity were used to treat groups of male Wistar rats ($n = 5$) at three dose levels (vehicle alone, low dose and high dose) and with target tissue (liver and kidney) and biological fluid samples (blood and urine) taken at several evaluation time points (24 h, 3 days and 14 days). Investigations performed on these materials included the use of classical assessment approaches (histopathology, clinical chemistry) and three omics approaches, namely transcriptomics (Affymetrix GeneChips), proteomics (2D-DIGE, 2D-PAGE, and SELDI) and metabolomics (LC/MS and NMR) (source: Gallagher *et al.* (2009); figure 2. Reproduced with permission of Elsevier)

histopathology are unlikely to be replaced in the near future, molecular profiling methods are likely to move to the centre stage within the next few years. The reason for this 'lies in the benefits of understanding the mechanistic foundations of adverse effects and the high level of measurement specificity'.[16]

The InnoMed PredTox programme served as a pilot project for a much larger initiative funded under the EU Framework Seven Programme, IMI. At the time of writing, IMI was still in progress and represented Europe's largest public–private partnership aiming to improve the drug development process by supporting a more efficient discovery and development of better and safer medicines for patients. Indeed, it claimed to be the world's largest public-private partnership in health research and development.

The overall goal of the IMI programme is to build a more collaborative environment for pharmaceutical R&D by creating unique, large-scale networks of innovation in pharmaceutical research. To that end, it supported a wide range of projects, three of which were directly relevant to toxicology:

[16] http://www.genedata.com/news-events/press-releases/

- **eTOX:** Integrating bioinformatics and chemoinformatics approaches for the development of expert systems allowing the *in silico* prediction of toxicities (Briggs *et al.*, 2012).
- **MARCAR:** Biomarkers and molecular tumour classification for non-genotoxic carcinogenesis. One example of this effort is the MARCAR consortium project,[17] whose aim is to identify transcriptional and epigenetic biomarkers of non-genotoxic carcinogenesis and human disease. The hope of this programme is that pharmacoepigenetics will allow the identification of epigenetic biomarkers for mechanistic studies and patient stratification. Initial studies have focussed on exemplification of the approach using an experimental models of non-genotoxic carcinogenesis (PB hepatocarcinogenesis in mouse liver) to identify novel early mechanism-based biomarkers (Lempiainen *et al.*, 2010).
- **MIP-DILI:** Mechanism-based integrated systems for the prediction of DILI.

The scope of the IMI extends well beyond toxicology; the projects in progress as of September 2012 addressed many aspects of human disease including infectious diseases, diabetes, cancer, neurodegenerative disorders and mental illness/impairment. The programme is at the halfway point at the time of writing, having been established in 2007 with a budget running to the end of 2017.

10.5.2 USA

The US Government's response to the National Research Council report was to establish the Tox21 Program,[18] aimed at moving toxicity testing from traditional methods which depend upon animal models treated at the MTD with extrapolation to predicted human health outcomes at lower doses to higher throughput *in vitro* methods which can be used to prioritise compounds for further study, identify mechanisms of action and ultimately develop predictive methods. The Memorandum of Understanding for Tox21 was signed in 2008 between the NIH Chemical Genomics Center (providing *in vitro* screening and informatics), the NTP (experimental toxicology) and the EPA (computational toxicology). The library of chemicals for testing was provided by the NIH Chemical Genomics Center and gave access to more than 400 000 chemicals from the NIH Molecular Libraries Small Molecule Repository as well as ~1400 compounds each from the NTP and EPA collections and 2816 drugs in clinical use. At the time of writing, Tox21 is well under way, with intense effort going into the determination of a range of toxicity endpoints in cell-based HTS assays using robotic systems (Shukla *et al.*, 2010). The future aims of Tox21 include the development of new quantitative HTS platforms which take more sources of variability into account, the identification and prioritisation of key cellular targets, establishment of a link between *in vitro* findings and human pathology and the creation of publicly accessible relational databases and computational tools to allow interested parties to interrogate the vast amount of data collected.

[17] www.imimarcar.eu
[18] http://www.ncats.nih.gov/research/reengineering/tox21/tox21.html

In order to support this effort, the EPA has made a significant commitment to the concept of computational toxicology, including the creation of the Aggregated Computational Toxicology Resource (ACToR), which comprises core ACToR (chemical identifiers, structures, hazard, exposure and use), ToxRefDB (the Toxicity Reference Database, a compilation of detailed *in vivo* toxicity data), ExpoCastDB (detailed human exposure data) and ToxCastDB (data from high-throughput screening programs). ToxRefDB[19] was set up (Martin *et al*., 2009) with the aim of developing novel approaches to the prediction of chemical toxicity in order to facilitate rapid screening of thousands of chemicals, improve mechanistic understanding and reduce animal use. It captures animal testing results generated over the last 30 years, storing study designs, dosing information and treatment-related effects using a standardised vocabulary. Such a controlled vocabulary is necessary because of the long time period the database covers.

Together with its partner database, ToxCastDB, which focuses on HTS data (Judson *et al*., 2010), the ToxRefDB provides an anchoring point for toxicity data from the EPA ToxCast research programme, facilitates the prioritisation of chemicals for future studies and provides a model for other efforts to capture quantitative, tabular toxicology data from legacy and new studies in a format suitable for cross-chemical computational toxicology analysis. Data from the ToxCast program has been used to predict steady-state blood concentrations which are likely to cause toxicity and compare them with estimated human exposure, thus starting to move beyond pure hazard identification and starting to get a handle on actual risk (Wetmore *et al*., 2012).

As of May 2012, the ToxRefDB database contained information on 474 chemicals, mostly data-rich active ingredients of pesticides. For this class of compounds, it was found that non-neoplastic toxicities dominated the determination of systemic NOAELs and LOAELs. More than 90% of the toxic chemicals caused toxicities in the liver, kidney, thyroid, lung, testis or spleen. It was not possible to classify compounds according to very specific pathological processes, but the use of broader terms (e.g. encompassing liver adenoma, combined adenoma/carcinoma and carcinoma, possibly extending even further to include all proliferative lesions of the liver) was effective. The ToxRefDB is now being extended to include data from other chemical classes and study types, including multigenerational reproductive toxicity and prenatal developmental toxicity studies. The strength of this approach is the broad spectrum of biology addressed, but future success in predicting target organ toxicity will depend on numerous factors including target, species and dose–response.

10.5.3 Evaluation

It has been stated that 'It is open to question as to whether the new approaches of systems biology are the start of a paradigm shift that will eventually spread to all other fields of biology as well, or whether they will stay within a subfield' (Cedersund and Roll, 2009). However, it is unarguable that systems biology approaches are now established alternatives to conventional hypothesis-driven

[19] http://www.epa.gov/ncct/toxrefdb/

experimentation. The rapid development of new HTS technologies and year-on-year increases in computer power will undoubtedly continue to fuel their evolution.

An important issue which needs to be tackled in the near future is how to store and manage the vast amounts of data which are emerging from toxicology studies using 'omics and systems biology approaches. Not only is it essential to establish databases in which relevant data may easily be identified and downloaded, but the pragmatic question of the actual form in which data are stored must always be kept in mind.[20]

10.6 Final word

Actually, this is not the moment for a final word on molecular and cellular toxicology. As I hope I have illustrated in this book, the discipline is moving so fast that by the time it is published more new developments will have occurred and a new understanding of toxicology will have been achieved. I hope, however, that the book will still be useful as an introduction, providing a place marker as we move into an new era of testing in which alternative methods begin to mitigate the welfare implications of our work and provide answers to key regulatory questions while sophisticated molecular methods continue to enhance our understanding of complex toxicological processes.

In other words, that's (not) all folks!

Self-assessment questions

- Under what circumstances would you use the zebrafish in preference to conventional mammalian models?
- To what extent are novel biomarkers enhancing our understanding of hepatic and renal toxicity?
- What are the relative merits and demerits of the systems biology approach as compared with the conventional scientific method?
- Will systems biology fundamentally change the discipline of toxicology?

References

Afshari, C.A., Hamadeh, H.K. and Bushel, P.R. (2011) The evolution of bioinformatics in toxicology: advancing toxicogenomics. *Toxicological Sciences*, **120** (Suppl 1), S225–S237.

Ali, S., Champagne, D.L., Spaink, H.P. and Richardson, M.K. (2011) Zebrafish embryos and larvae: a new generation of disease models and drug screens. *Birth Defects Research. Part C, Embryo Today*, **93**, 115–133.

Amacher, D.E. (2002) A toxicologist's guide to biomarkers of hepatic response. *Human and Experimental Toxicology*, **21**, 253–262.

[20] For a current example of a similar issue, see http://www.bbc.co.uk/news/technology-13367398

Basnakian, A.G. (2008) Netrin-1: a potential universal biomarker for acute kidney injury. *American Journal of Physiology. Renal Physiology*, **294**, F729–F730.

Beadle, G.W. and Tatum, E.L. (1941) Genetic control of biochemical reactions in *Neurospora*. *Proceedings of the National Academy of Sciences of the United States of America*, **27**, 499–506.

Boekelheide, K. and Schuppe-Koistinen, I. (2012) SOT/EUROTOX debate: biomarkers from blood and urine will replace traditional histopathological evaluation to determine adverse responses. *Toxicological Sciences*, **129**, 249–255.

Bonventre, J.V., Vaidya, V.S., Schmouder, R. *et al*. (2010) Next-generation biomarkers for detecting kidney toxicity. *Nature Biotechnology*, **28**, 436–440.

Briggs, K., Cases, M., Heard, D.J. *et al*. (2012) Inroads to predict *in vivo* toxicology-an introduction to the eTOX project. *International Journal of Molecular Sciences*, **13**, 3820–3846.

Cedersund, G. and Roll, J. (2009) Systems biology: model based evaluation and comparison of potential explanations for given biological data. *FEBS Journal*, **276**, 903–922.

Chatterjee, S. and Lufkin, T. (2011) Fishing for function: zebrafish BAC transgenics for functional genomics. *Molecular BioSystems*, **7**, 2345–2351.

Cooke, E.J., Savage, R.S. and Wild, D.L. (2009) Computational approaches to the integration of gene expression, ChIP-chip and sequence data in the inference of gene regulatory networks. *Seminars in Cell and Developmental Biology*, **20**, 863–868.

De Esch, C., Slieker, R., Wolterbeek, A. *et al*. (2012) Zebrafish as potential model for developmental neurotoxicity testing: a mini review. *Neurotoxicology and Teratology*, **34**, 545–553.

Doyon, Y., McCammon, J.M., Miller, J.C. *et al*. (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nature Biotechnology*, **26**, 702–708.

Edwards, S.W. and Preston, R.J. (2008) Systems biology and mode of action based risk assessment. *Toxicological Sciences*, **106**, 312–318.

Gallagher, W. M., Tweats, D. and Koenig, J. (2009). Omic profiling for drug safety assessment: current trends and public-private partnerships. *Drug Discov Today*, **14**, 337–42.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**, 57–70.

Henry, T.B., McPherson, J.T., Rogers, E.D. *et al*. (2009) Changes in the relative expression pattern of multiple vitellogenin genes in adult male and larval zebrafish exposed to exogenous estrogens. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **154**, 119–126.

Henry, T.B., Menn, F.M., Fleming, J.T. *et al*. (2007) Attributing effects of aqueous C60 nano-aggregates to tetrahydrofuran decomposition products in larval zebrafish by assessment of gene expression. *Environmental Health Perspectives*, **115**, 1059–1065.

Hill, A., Mesens, N., Steemans, M. *et al*. (2012) Comparisons between *in vitro* whole cell imaging and *in vivo* zebrafish-based approaches for identifying potential human hepatotoxicants earlier in pharmaceutical development. *Drug Metabolism Reviews*, **44**, 127–140.

Hoffmann, D., Adler, M., Vaidya, V.S. *et al*. (2010) Performance of novel kidney biomarkers in preclinical toxicity studies. *Toxicological Sciences*, **116**, 8–22.

Jaeschke, H. (2007) Troglitazone hepatotoxicity: are we getting closer to understanding idiosyncratic liver injury? *Toxicological Sciences*, **97**, 1–3.

Judson, R.S., Houck, K.A., Kavlock, R.J. *et al*. (2010) *In vitro* screening of environmental chemicals for targeted testing prioritization: the ToxCast project. *Environmental Health Perspectives*, **118**, 485–492.

Kashimshetty, R., Desai, V.G., Kale, V.M. *et al.* (2009) Underlying mitochondrial dysfunction triggers flutamide-induced oxidative liver injury in a mouse model of idiosyncratic drug toxicity. *Toxicology and Applied Pharmacology*, **238**, 150–159.

Keller, D.A., Juberg, D.R., Catlin, N. *et al.* (2012) Identification and characterization of adverse effects in 21st century toxicology. *Toxicological Sciences*, **126**, 291–297.

Kreeger, P.K. and Lauffenburger, D.A. (2010) Cancer systems biology: a network modeling perspective. *Carcinogenesis*, **31**, 2–8.

Kumar, B.S., Chung, B.C., Kwon, O.S. and Jung, B.H. (2012) Discovery of common urinary biomarkers for hepatotoxicity induced by carbon tetrachloride, acetaminophen and methotrexate by mass spectrometry-based metabolomics. *Journal of Applied Toxicology*, **32**, 505–520.

Laubenbacher, R., Hower, V., Jarrah, A. *et al.* (2009) A systems biology view of cancer. *Biochimica et Biophysica Acta*, **1796**, 129–139.

Leitner, J.M., Graninger, W. and Thalhammer, F. (2010) Hepatotoxicity of antibacterials: Pathomechanisms and clinical. *Infection*, **38**, 3–11.

Lempiainen, H., Muller, A., Brasa, S. *et al.* (2010) Phenobarbital mediates an epigenetic switch at the Constitutive Androstane Receptor (CAR) target gene Cyp2b10 in the liver of B6C3F1 Mice. *PLoS One*, **6**, e18216.

Li, J., Zhu, X. and Chen, J.Y. (2009) Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts. *PLoS Computational Biology*, **5**, e1000450.

Lieschke, G.J. and Currie, P.D. (2007) Animal models of human disease: zebrafish swim into view. *Nature Reviews. Genetics*, **8**, 353–367.

Marchan, R., Bolt, H.M. and Hengstler, J.G. (2012) Systems biology meets toxicology. *Archives of Toxicology*, **86**, 1157–1158.

Martin, M.T., Judson, R.S., Reif, D.M. *et al.* (2009) Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. *Environmental Health Perspectives*, **117**, 392–399.

McBurney, R.N., Hines, W.M., Von Tungeln, L.S. *et al.* (2009) The liver toxicity biomarker study: phase I design and preliminary results. *Toxicologic Pathology*, **37**, 52–64.

McCollum, C.W., Ducharme, N.A., Bondesson, M. and Gustafsson, J.A. (2011) Developmental toxicity screening in zebrafish. *Birth Defects Research. Part C, Embryo Today*, **93**, 67–114.

McGrath, P. and Li, C.Q. (2008) Zebrafish: a predictive model for assessing drug-induced toxicity. *Drug Discovery Today*, **13**, 394–401.

Menke, A.L., Spitsbergen, J.M., Wolterbeek, A.P. and Woutersen, R.A. (2011) Normal anatomy and histology of the adult zebrafish. *Toxicologic Pathology*, **39**, 759–775.

Molitoris, B.A., Melnikov, V.Y., Okusa, M.D. and Himmelfarb, J. (2008) Technology insight: biomarker development in acute kidney injury–what can we anticipate? *Nature Clinical Practice. Nephrology*, **4**, 154–165.

Moylan, C.A., Suzuki, A., Papay, J.I. *et al.* (2012) A pre-marketing ALT signal predicts post-marketing liver safety. *Regulatory Toxicology and Pharmacology*, **63**, 433–439.

Negishi, K., Noiri, E., Doi, K. *et al.* (2009) Monitoring of urinary L-type fatty acid-binding protein predicts histological severity of acute kidney injury. *American Journal of Pathology*, **174**, 1154–1159.

Ong, M.M., Latchoumycandane, C. and Boelsterli, U.A. (2007) Troglitazone-induced hepatic necrosis in an animal model of silent genetic mitochondrial abnormalities. *Toxicological Sciences*, **97**, 205–213.

Ozer, J.S., Dieterle, F., Troth, S. *et al.* (2010) A panel of urinary biomarkers to monitor reversibility of renal injury and a serum marker with improved potential to assess renal function. *Nature Biotechnology*, **28**, 486–494.

Park, J.W., Henry, T.B., Ard, S. *et al.* (2011) The association between nC60 and 17alpha-ethinylestradiol (EE2) decreases EE2 bioavailability in zebrafish and alters nanoaggregate characteristics. *Nanotoxicology*, **5**, 406–416.

Park, J.W., Henry, T.B., Menn, F.M. *et al.* (2010) No bioavailability of 17alpha-ethinylestradiol when associated with nC60 aggregates during dietary exposure in adult male zebrafish (Danio rerio). *Chemosphere*, **81**, 1227–1232.

Pei, D.S. and Strauss, P.R. (2013) Zebrafish as a model system to study DNA damage and repair. *Mutation Research*, 743–744, 151–159.

Peters, T.S. (2005) Do preclinical testing strategies help predict human hepatotoxic potentials? *Toxicologic Pathology*, **33**, 146–154.

Peterson, R.T. and MacRae, C.A. (2012) Systematic approaches to toxicology in the zebrafish. *Annual Review of Pharmacology and Toxicology*, **52**, 433–453.

Piersma, A. 2011. Development and implementation of alternative methods in reproductive toxicology. RIVM 340700005/2011. http://www.rivm.nl/bibliotheek/rapporten/340700005.pdf

Ray, W.J., Bain, G., Yao, M. and Gottlieb, D.I. (1997) CYP26, a novel mammalian cytochrome P450, is induced by retinoic acid and defines a new family. *Journal of Biological Chemistry*, **272**, 18702–18708.

Reeves, W.B., Kwon, O. and Ramesh, G. (2008) Netrin-1 and kidney injury. II. Netrin-1 is an early biomarker of acute kidney injury. *American Journal of Physiology. Renal Physiology*, **294**, F731–F738.

Roux, A., Lison, D., Junot, C. and Heilier, J.F. (2011) Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: a review. *Clinical Biochemistry*, **44**, 119–135.

Senior, J.R. (2012) Alanine aminotransferase: a clinical and regulatory tool for detecting liver injury-past, present, and future. *Clinical Pharmacology and Therapeutics*, **92**, 332–339.

Shaw, P.J., Beggs, K.M., Sparkenbaugh, E.M. *et al.* (2009) Trovafloxacin enhances TNF-induced inflammatory stress and cell death signaling and reduces TNF clearance in a murine model of idiosyncratic hepatotoxicity. *Toxicological Sciences*, **111**, 288–301.

Short, B. (2009) Cell biologists expand their networks. *Journal of Cell Biology*, **186**, 305–311.

Shukla, S.J., Huang, R., Austin, C.P. and Xia, M. (2010) The future of toxicity testing: a focus on *in vitro* methods using a quantitative high-throughput screening platform. *Drug Discovery Today*, **15**, 997–1007.

Sipes, N.S., Padilla, S. and Knudsen, T.B. (2011) Zebrafish: as an integrative model for twenty-first century toxicity testing. *Birth Defects Research. Part C, Embryo Today*, **93**, 256–267.

Smoot, M.E., Ono, K., Ruscheinski, J. *et al.* (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, **27**, 431–432.

Stephens, M.L., Barrow, C., Andersen, M.E. *et al.* (2012) Accelerating the development of 21st-century toxicology: outcome of a Human Toxicology Project Consortium workshop. *Toxicological Sciences*, **125**, 327–334.

Strahle, U., Scholz, S., Geisler, R. *et al.* (2012) Zebrafish embryos as an alternative to animal experiments–a commentary on the definition of the onset of protected life stages in animal welfare regulations. *Reproductive Toxicology*, **33**, 128–132.

Suster, M.L., Kikuta, H., Urasaki, A. *et al.* (2009a) Transgenesis in zebrafish with the *tol2* transposon system. *Methods in Molecular Biology*, **561**, 41–63.

Suster, M.L., Sumiyama, K. and Kawakami, K. (2009b) Transposon-mediated BAC transgenesis in zebrafish and mice. *BMC Genomics*, **10**, 477.

Suter, L., Schroeder, S., Meyer, K. *et al.* (2011) EU Framework 6 project: predictive toxicology (PredTox)--overview and outcome. *Toxicology and Applied Pharmacology*, **252**, 73–84.

Tonomura, Y., Tsuchiya, N., Torii, M. and Uehara, T. (2010) Evaluation of the usefulness of urinary biomarkers for nephrotoxicity in rats. *Toxicology*, **273**, 53–59.

Umans, R.A. and Taylor, M.R. (2012) Zebrafish as a model to study drug transporters at the blood-brain barrier. *Clinical Pharmacology and Therapeutics*, **92**, 567–570.

Van Ravenzwaay, B., Brunborg, G., Kleinjans, J. *et al.* (2012) Use of 'omics to elucidate mechanism of action and integration of 'omics in a systems biology concept. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, **746**, 95–96.

Van Vliet, E. (2011) Current standing and future prospects for the technologies proposed to transform toxicity testing in the 21st century. *ALTEX*, **28**, 17–44.

Wen, C.P., Lin, J., Yang, Y.C. *et al.* (2012) Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. *Journal of the National Cancer Institute*, **104**, 1599–1611.

Wetmore, B.A., Brees, D.J., Singh, R. *et al.* (2010) Quantitative analyses and transcriptomic profiling of circulating messenger RNAs as biomarkers of rat liver injury. *Hepatology*, **51**, 2127–2139.

Wetmore, B.A., Wambaugh, J.F., Ferguson, S.S. *et al.* (2012) Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicological Sciences*, **125**, 157–174.

Yang, L., Ho, N.Y., Alshut, R. *et al.* (2009) Zebrafish embryos as models for embryotoxic and teratological effects of chemicals. *Reproductive Toxicology*, **28**, 245–253.

Index

Molecular and Cellular Toxicology: An Introduction, First Edition. Lesley A. Stanley.

Companion Website: www.wiley.com/go/stanley/moleculercellulartoxicology