Individual variation and drug interaction

56

OVERVIEW

This chapter addresses sources of variation between individuals (interindividual variation) in their responses to drugs. Genetic variation in pharmacokinetic processes and pharmacodynamic response has been discussed in Chapter 11. In this chapter, we mention briefly some other important factors responsible for pharmacological variation, including age, pregnancy and disease, and describe in more detail the mechanisms underlying drug interaction (i.e. modfication of the action of one drug by another).

INTRODUCTION

Therapeutics would be a great deal easier if responses to the same dose of drug were always the same. In reality, inter- and even intraindividual variation is often substantial. Physicians need to be aware of the sources of such variation to prescribe drugs safely and effectively. Variation can be caused by different concentrations at sites of drug action or by different responses to the same drug concentration. The first kind is called pharmacokinetic variation and can occur because of differences in absorption, distribution, metabolism or excretion (Chs 8 and 9). The second kind is called pharmacodynamic variation.

Variation is usually quantitative in the sense that the drug produces a larger or smaller effect, or acts for a longer or shorter time, while still exerting qualitatively the same effect. In other cases, the action is qualitatively different. These are known as 'idiosyncratic' reactions (the *Oxford English Dictionary* defines idiosyncrasy as 'the physical constitution peculiar to an individual or class') and are often caused by genetic or immunological differences between individuals.

Individual variation

- Variability is a serious problem; if not taken into account, it can result in:
  - lack of efficacy
  - unexpected side effects.
- Types of variability may be classified as:
  - pharmacokinetic
  - pharmacodynamic
  - idiosyncratic.
- The main causes of variability are:
  - age
  - genetic factors
  - immunological factors (Ch. 57)
  - pathological states (e.g. kidney or liver disease)
  - drug interactions.

FACTORS RESPONSIBLE FOR QUANTITATIVE INDIVIDUAL VARIATION

ETHNICITY

Ethnic means 'pertaining to race', and many anthropologists are sceptical as to the value of this concept (see, for example, Cooper et al., 2003). Citizens of several modern societies are asked to define their race or ethnicity from a list of options (e.g. 'white', 'black', 'mixed', 'Chinese', 'Asian' or 'other' were the options provided by the UK Office of National Statistics for the 2001 National Census). Members of such self-defined groups share some characteristics on the basis of common genetic and cultural heritage, but there is obviously also enormous diversity within each group.

Despite the crudeness of such categorisation, it can give some pointers to drug responsiveness (Wood, 2001). One example is the evidence discussed in Chapter 22 that African-Americans with heart failure gain a mortality benefit from treatment with a combination of hydralazine plus a nitrate, whereas white Americans may not.

Some adverse effects may also be predicted on the basis of race; for example, many Chinese subjects differ from Europeans in the way that they metabolise ethanol, producing a higher plasma concentration of acetaldehyde, which can cause flushing and palpitations (Chs 48 and 57). Chinese subjects are considerably more sensitive to the cardiovascular effects of **propranolol** (Ch. 14) than white Europeans, whereas Afro-Caribbean individuals are less sensitive. Despite their increased sensitivity to β-adrenoceptor antagonists, Chinese subjects metabolise propranolol faster than white people, implying that the difference relates to pharmacodynamic differences in sensitivity at or beyond the β-adrenoceptors.

Overall effectiveness of **gefitinib** (Ch. 55) in treating patients with advanced lung tumours has been disappointing, but in about 10% of patients lung tumours shrink rapidly in response to this drug. Japanese patients are three times as likely as whites to respond in this way. The underlying difference is that patients who respond well have specific mutations in the receptor for epidermal growth factor (see Wadman, 2005). It is probable that many such ethnic differences are genetic in origin, but environmental factors, for example relating to distinctive dietary habits, may also contribute. It is important not to abandon the much more sophisticated search for ways to individualise medicine on the basis of pharmacogenomics (Ch. 11) just because the much simpler and cheaper process of asking patients to define their ethnic group has had some success: this should rather act as a spur. If such a crude and imperfect approach has had some success, think how much better we ought to be able to do with genomic testing!

AGE

The main reason that age affects drug action is that drug elimination is less efficient in newborn babies and in old people, so that drugs commonly produce greater and more prolonged effects at the extremes of life. Other age-related factors, such as variations in pharmacodynamic sensitivity, are also important with some drugs. Physiological factors (e.g. altered cardiovascular reflexes) and pathological factors (e.g. hypothermia), which are common in elderly people, also influence drug effects. Body composition changes with age, fat contributing a greater proportion to body mass in the elderly, with consequent changes in distribution volume of drugs. Elderly people consume more drugs than do younger adults, so the potential for drug interactions (see below) is also increased. For fuller accounts of drug therapy in paediatrics and in the elderly, see, respectively, Fox & Balis (Ch. 23) and Abernethie (Ch. 24) in Atkinson et al., 2006.

EFFECT OF AGE ON RENAL EXCRETION OF DRUGS

Glomerular filtration rate (GFR) in the newborn, normalised to body surface area, is only about 20% of the adult value, and tubular function is also less. Accordingly, plasma elimination half-lives of renally eliminated drugs are longer in neonates than in adults (Table 56.1). In babies born at term, renal function increases to values similar to those in young adults in less than a week, and continues to increase to a maximum of approximately twice the adult value at 6 months of age. Improvement in renal function occurs more slowly in premature infants. Renal immaturity in premature infants can have a substantial effect on drug elimination. For example, in premature newborn babies, the antibiotic **gentamicin** has a plasma half-life of ≥ 18 h, compared with 1–4 h for adults and approximately 10 h for babies born at term. It is therefore necessary to reduce and/or space out doses to avoid toxicity in premature babies.

Glomerular filtration rate declines slowly from about 20 years of age, falling by about 25% at 50 years and by 50% at 75 years. Figure 56.1 shows that the renal clearance of **digoxin** in young and old subjects is closely correlated with creatinine clearance, a measure of GFR. Consequently, chronic administration over the years of the same daily dose of digoxin to an individual as he or she ages leads to a progressive increase in plasma concentration, and this is a common cause of glycoside toxicity in elderly people (see Ch. 21).

▼ The age-related decline in GFR is not reflected by an increase in plasma creatinine concentration, as distinct from creatinine clearance. Plasma creatinine typically remains within the normal adult range in elderly persons despite substantially diminished GFR. This is because creatinine synthesis is reduced in elderly persons because of their reduced muscle mass. Consequently, a 'normal' plasma creatinine in an elderly person does not indicate that they have a normal GFR. Failure to recognise this and reduce the dose of drugs that are eliminated by renal excretion can lead to drug toxicity.

EFFECT OF AGE ON DRUG METABOLISM

Several important enzymes, including hepatic microsomal oxidase, glucuronyltransferase, acetyltransferase and plasma esterases, have low activity in neonates, especially if premature. These enzymes take 8 weeks or longer to reach the adult level of activity. The relative lack of conjugating activity in the newborn can have serious consequences, as in *kernicterus* caused by drug displacement of bilirubin from its binding sites on albumin (see below) and in the 'grey baby' syndrome caused by the antibiotic **chloramphenicol** (see Ch. 50). This sometimes fatal condition, at first thought to be a specific biochemical sensitivity to the drug in young babies, actually results simply from accumulation of very high tissue concentrations of chloramphenicol because of slow hepatic conjugation. Chloramphenicol is no more toxic to babies than to adults provided the dose is reduced to make allowance for this. Slow conjugation is also one reason why **morphine** (which

Table 56.1 Effect of age on plasma elimination half-lives of various drugs

Drug	Mean or range of half-life (h)		
	Term neonate[a]	Adult	Elderly person
Drugs that are mainly excreted unchanged in the urine			
Gentamicin	10	2	4
Lithium	120	24	48
Digoxin	200	40	80
Drugs that are mainly metabolised			
Diazepam	25–100	15–25	50–150
Phenytoin	10–30	10–30	10–30
Sulfamethoxypyridazine	140	60	100

[a] Even greater differences from mean adult values occur in premature babies.
(Data from Reidenberg 1971 Renal function and drug action. Saunders, Philadelphia; and Dollery 1991 Therapeutic drugs. Churchill Livingstone, Edinburgh.)

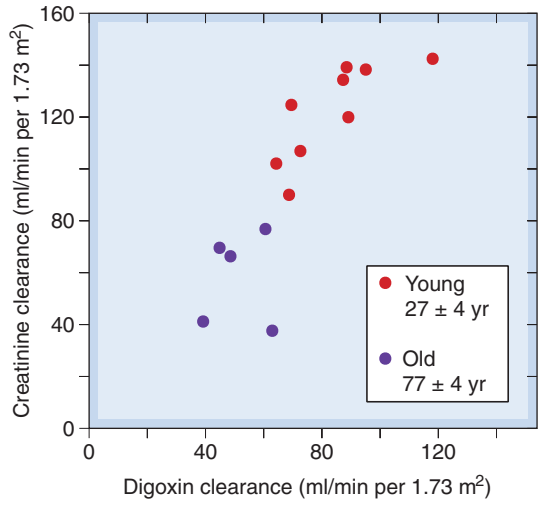

Fig. 56.1 Relationship between renal function (measured as creatinine clearance) and digoxin clearance in young and old subjects. (From Ewy G A et al. 1969 Circulation 34: 452.)

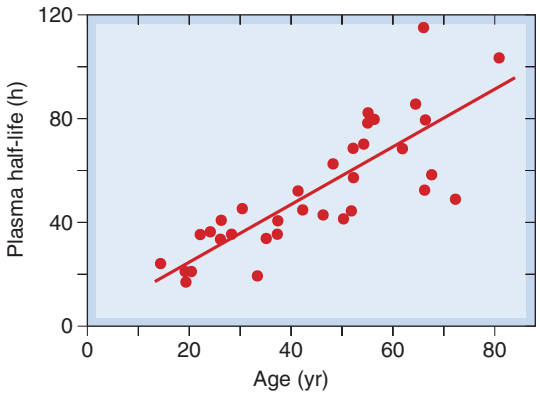

Fig. 56.2 Increasing plasma half-life for diazepam with age in 33 normal subjects. Note the increased variability as well as increased half-life with ageing. (From Klotz U et al. 1975 J Clin Invest 55: 347.)

is excreted mainly as the glucuronide, see Ch. 41) is not used as an analgesic in labour, because drug transferred via the placenta has a long half-life in the newborn baby and can cause prolonged respiratory depression.

The activity of hepatic microsomal enzymes declines slowly (and very variably) with age, and the distribution volume of lipid-soluble drugs increases, because the proportion of the body that is fat increases with advancing age. The increasing half-life of the anxiolytic drug **diazepam** with advancing age (Fig. 56.2) is one consequence of this. Some other benzodiazepines and their active metabolites show even greater age-related increases in half-life. Because half-life determines the time course of drug accumulation during repeated dosing (Ch. 10), insidious effects, developing over days or weeks, can occur in elderly people and may be misattributed to age-related memory impairment rather than to drug accumulation. The effect of age is less marked for many other drugs, but even though the mean half-life may not change much, there is often a striking increase in the variability of half-life between individuals with increasing age. This is important, because a population of old people will contain some individuals with grossly reduced rates of drug metabolism, whereas such extremes do not occur so commonly in young adult populations. Drug regulatory authorities therefore usually require studies in elderly patients as part of drug evaluation.

AGE-RELATED VARIATION IN SENSITIVITY TO DRUGS

The same plasma concentration of a drug can cause different effects in young and old subjects. Benzodiazepines (Ch. 43) exemplify this, producing more confusion and less sedation in elderly than in young subjects; similarly, hypotensive drugs (Ch. 22) cause postural hypotension more commonly in elderly than in younger adult patients.

PREGNANCY

Pregnancy causes physiological changes that influence drug disposition (Ch. 8) in mother and fetus. Maternal plasma albumin concentration is reduced, influencing drug protein binding. Cardiac output is increased, leading to increased renal blood flow and GFR, and increased renal elimination of drugs. Lipophilic molecules rapidly traverse the placental barrier, whereas transfer of hydrophobic drugs is slow, limiting fetal drug exposure following a single maternal dose. The placental barrier excludes some drugs (e.g. low-molecular-weight heparins; Ch. 24) so effectively that they can be administered chronically to the mother without causing effects in the fetus. However, drugs that are transferred to the fetus are eliminated more slowly than from the mother. The activity of most drug-metabolising enzymes in fetal liver is much less than in the adult. Furthermore, the fetal kidney is not an efficient route of elimination because excreted drug enters the amniotic fluid, which is swallowed by the fetus. For a fuller account, see Striker & Frederiksen (Ch. 22) in Atkinson et al., 2006.

DISEASE

Therapeutic drugs are prescribed to patients, so effects of disease on drug response are very important in clinical pharmacology. Detailed consideration is beyond the scope of this book, and interested readers should refer to a clinical text such as the chapters on renal and hepatic disease in Atkinson et al., 2006. Disease can cause pharmacokinetic or pharmacodynamic variation. Common disorders such as impaired renal or hepatic function predispose to toxicity by causing unexpectedly intense or prolonged drug effects as a result of increased drug concentration following a standard dose. Drug absorption is slowed in conditions causing gastric stasis (e.g. *migraine*, *diabetic neuropathy*) and may be incomplete in patients with malabsorption owing to ileal or pancreatic disease or to oedema of the ileal mucosa caused by heart failure or nephrotic syndrome. *Nephrotic syndrome* (characterised by heavy proteinuria, oedema and a reduced concentration of albumin in plasma) alters drug absorption because of oedema of intestinal mucosa; alters drug disposition through changes in binding to plasma albumin; and causes insensitivity to diuretics such as **furosemide** that act on ion transport mechanisms on the lumenal surface of tubular epithelium (Ch. 28), through binding to albumin in tubular fluid. *Hypothyroidism* is associated with increased sensitivity to several widely used drugs (e.g. **pethidine**), for reasons that are poorly understood. *Hypothermia* (to which elderly persons, in particular, are predisposed) markedly reduces the clearance of many drugs.

Other disorders affecting receptors and signal transduction mechanisms (see Ch. 3), although uncommon, illustrate mechanisms that may prove to be of more general applicability. Examples include:

- diseases that influence receptors:
  - *myasthenia gravis*, an autoimmune disease characterised by antibodies to nicotinic acetylcholine receptors (Ch. 13) and increased sensitivity to neuromuscular blocking agents (e.g. **vecuronium**) and other drugs that may influence neuromuscular transmission (e.g. *aminoglycoside antibiotics*, Ch. 50)
  - *X-linked nephrogenic diabetes insipidus*, characterised by abnormal antidiuretic hormone (ADH, vasopressin) receptors (Ch. 28) and insensitivity to ADH
  - *familial hypercholesterolaemia*, an inherited disease of low-density-lipoprotein receptors (Ch. 23); the (very rare) homozygous form is relatively resistant to treatment with statins (which work mainly by increasing expression of these receptors), whereas

Genetic factors (see Ch. 11)

- Genetic variation is an important source of pharmacokinetic variability.
- There are several clear examples where genetic variation influences drug response, including:
  - fast/slow acetylators (**hydralazine**, **procainamide**, **isoniazid**)
  - plasma cholinesterase variants (**suxamethonium**)
  - hydroxylase polymorphism (**debrisoquine**).
- In future, profiling an individual's DNA (e.g. for combinations of single nucleotide polymorphisms) could provide a way to anticipate drug responsiveness.

Variation due to disease

Pharmacokinetic alterations in:

- Absorption:
  - gastric stasis (e.g. migraine)
  - malabsorption (e.g. steatorrhoea from pancreatic insufficiency)
  - oedema of ileal mucosa (e.g. heart failure, nephrotic syndrome).
- Distribution:
  - altered plasma protein binding (e.g. of **phenytoin** in chronic renal failure)
  - impaired blood–brain barrier (e.g. to **penicillin** in meningitis).
- Metabolism:
  - chronic liver disease
  - hypothermia.
- Excretion:
  - acute and/or chronic renal failure.

Pharmacodynamic alterations in:

- Receptors (e.g. myasthenia gravis, familial hypercholesterolaemia).
- Signal transduction (e.g. pseudohypoparathyroidism, familial precocious puberty).
- Unknown mechanisms (e.g. increased sensitivity to **pethidine** in hypothyroidism).

the much commoner heterozygous form responds well to statins.

- diseases that influence signal transduction mechanisms:
  - *pseudohypoparathyroidism*, which stems from impaired coupling of receptors with adenylyl cyclase
  - *familial precocious puberty* and *hyperthyroidism* caused by functioning thyroid adenomas, which are each caused by mutations in G-protein-coupled receptors that result in the receptors remaining 'turned on' even in the absence of the hormones that are their natural agonists.

Idiosyncratic reactions

- Harmful, sometimes fatal, reactions that occur in a small minority of individuals.
- Reactions may occur with low doses.
- Genetic factors may be responsible (e.g. **primaquine** sensitivity, malignant hyperthermia), although often the cause is poorly understood (e.g. bone marrow depression with **chloramphenicol**).
- Immunological factors are also important (see Ch. 57).

IDIOSYNCRATIC REACTIONS

An idiosyncratic reaction is a qualitatively abnormal, and usually harmful, drug effect that occurs in a small proportion of individuals. For example, **chloramphenicol** causes aplastic anaemia in approximately 1 in 50 000 patients (Ch. 50). In many cases, genetic anomalies are responsible. Glucose 6-phosphate dehydrogenase (G6PD) deficiency and the hepatic porphyrias are well-understood examples of this (Ch. 11). Malignant hyperthermia is a metabolic reaction to drugs including **suxamethonium** and various *inhalational anaesthetics* and *antipsychotic drugs*. Susceptibility to these drugs in affected individuals is caused by an inherited abnormality in the Ca^{2+} release channel known as the *ryanodine receptor* located in the sarcoplasmic reticulum of striated muscle (Ch. 4).

Immunological mechanisms underlie many idiosyncratic reactions. Propensity to these is genetically determined (Ch. 11).They are considered further in Chapter 57.

DRUG INTERACTIONS

Many patients, especially elderly ones, are treated continuously with one or more drugs for chronic diseases such as hypertension, heart failure, osteoarthritis and so on. Acute events (e.g. infections, myocardial infarction) are treated with additional drugs. The potential for drug interactions is therefore substantial, and drug interactions account for 5–20% of adverse drug reaction. These may be serious (approximately 30% of fatal adverse drug reactions are estimated to be the consequence of drug interaction) and may be misattributed to the natural history of disease (e.g. rejection of a transplanted kidney may be attributed to this when it was actually caused by loss of effectiveness of immunosuppressant medication as a result of drug interaction; see below). Drugs can also interact with chemical entities in other dietary constituents (e.g. grapefruit juice, which downregulates expression of CYP3A4 in the gut) and herbal remedies (such as St John's wort; Ch. 46). The administration of one chemical entity (A) can alter the action of another (B) by one of two general mechanisms:[1]

[1]A third category of pharmaceutical interactions should be mentioned, in which drugs interact in vitro so that one or both are inactivated. No pharmacological principles are involved, just chemistry. An example is the formation of a complex between **thiopental** and **suxamethonium**, which must not be mixed in the same syringe. **Heparin** is highly charged and interacts in this way with many basic drugs; it is sometimes used to keep intravenous lines or cannulae open and can inactivate basic drugs if they are injected without first clearing the line with saline.

1. Modifying the pharmacological effect of B without altering its concentration in the tissue fluid (pharmacodynamic interaction).
2. Altering the concentration of B at its site of action (pharmacokinetic interaction).

For such interactions to be important clinically, it is necessary that the therapeutic range of drug B is narrow (i.e. that a small reduction in effect will lead to loss of efficacy and/or a small increase in effect will lead to toxicity). For pharmacokinetic interactions to be clinically important, it is also necessary that the concentration–response curve of drug B is steep (so that a small change in plasma concentration leads to a substantial change in effect). For many drugs, these conditions are not met: even quite large changes in plasma concentrations of relatively non-toxic drugs such as **penicillin** are unlikely to give rise to clinical problems, because there is usually a comfortable safety margin between plasma concentrations produced by usual doses and those resulting either in toxicity or in loss of efficacy. Several drugs do have steep concentration–response relationships and a narrow therapeutic margin and, for these, drug interactions can cause major problems, for example with *antithrombotic*, *antidysrhythmic*, *antiviral* and *antiepileptic* drugs; **lithium**; and several *antineoplastic* and *immunosuppressant* drugs.

PHARMACODYNAMIC INTERACTION

Pharmacodynamic interaction can occur in many different ways (including those discussed under *Drug antagonism* in Ch. 2). There are many mechanisms, and some examples of practical importance are probably more useful than attempts at classification:

- β-Adrenoceptor antagonists diminish the effectiveness of β-adrenoceptor agonists such as **salbutamol** (Ch. 14).
- Many diuretics lower plasma K^+ concentration (see Ch. 28), and thereby predispose to **digoxin** toxicity and to toxicity with *type III antidysrhythmic drugs* (Ch. 21).
- **Sildenafil** inhibits the isoform of phosphodiesterase (type V) that inactivates cGMP (Chs 20 and 34); consequently, it potentiates organic nitrates, which activate guanylyl cyclase, and can cause severe hypotension in patients taking these drugs.
- *Monoamine oxidase inhibitors* increase the amount of noradrenaline stored in noradrenergic nerve terminals and interact dangerously with drugs, such as **ephedrine** or **tyramine**, that release stored noradrenaline. This can also occur with tyramine-rich foods—particularly fermented cheeses such as Camembert (see Ch. 46).
- **Warfarin** competes with vitamin K, preventing hepatic synthesis of various coagulation factors (see Ch. 24). If vitamin K production in the intestine is inhibited (e.g. by antibiotics), the anticoagulant action of warfarin is increased.
- The risk of bleeding, especially from the stomach, caused by warfarin is increased by drugs that cause bleeding by different mechanisms (e.g. **aspirin**, which inhibits platelet thromboxane A_2 biosynthesis and which can damage the stomach; Ch. 26).
- *Sulfonamides* prevent the synthesis of folic acid by bacteria and other microorganisms; **trimethoprim** inhibits its reduction to tetrahydrofolate. Given together, the drugs have a synergistic action of value in treating *Pneumocystis* infection (Ch. 53).
- *Non-steroidal anti-inflammatory drugs* (NSAIDs; Ch. 26), such as **ibuprofen** or **indometacin**, inhibit biosynthesis of prostaglandins, including renal vasodilator/natriuretic prostaglandins (prostaglandin E_2, prostaglandin I_2). If administered to patients receiving treatment for hypertension, they increase the blood pressure. If given to patients being treated with diuretics for chronic heart failure, they cause salt and water retention and hence cardiac decompensation.[2]
- Histamine H_1 receptor antagonists, such as **promethazine**, commonly cause drowsiness as an unwanted effect. This is more troublesome if such drugs are taken with alcohol, leading to accidents at work or on the road.

PHARMACOKINETIC INTERACTION

All the four major processes that determine pharmacokinetics—absorption, distribution, metabolism and excretion—can be affected by drugs. Some of the more important mechanisms are given here, with examples.

ABSORPTION

Gastrointestinal absorption is slowed by drugs that inhibit gastric emptying, such as **atropine** or *opiates*, or accelerated by drugs that hasten gastric emptying (e.g. **metoclopramide**; see Ch. 29). Alternatively, drug A may interact with drug B in the gut in such a way as to inhibit absorption of B (cf. pharmaceutical interactions; see footnote 1). For example, Ca^{2+} or Fe^{2+} each form insoluble complexes with **tetracycline** that retard its absorption; **colestyramine**, a bile acid-binding resin, binds several drugs (e.g. **warfarin**, **digoxin**), preventing their absorption if administered simultaneously. Another example is the addition of adrenaline (epinephrine) to local anaesthetic injections; the resulting vasoconstriction slows the absorption of the anaesthetic, thus prolonging its local effect (Ch. 42).

DRUG DISTRIBUTION

One drug may alter the distribution of another, by competing for a common binding site on plasma albumen or tissue protein, but such interactions are seldom clinically important unless accompanied by a separate effect on drug elimination (see below). Displacement of a drug from binding sites in plasma or tissues transiently increases the concentration of free (unbound) drug, but this is followed by increased elimination, so a new steady state results in which total drug concentration in plasma is reduced but the free drug concentration is similar to that before introduction of the second 'displacing' drug. Consequences of potential clinical importance are as follow:

- Toxicity from the transient increase in concentration of free drug before the new steady state is reached.
- If dose is being adjusted according to measurements of total plasma concentration, it must be appreciated that the target therapeutic concentration range will be altered by co-administration of a displacing drug.

[2]The interaction with diuretics may involve a pharmacokinetic interaction in addition to the pharmacodynamic effect described here, because NSAIDs compete with weak acids, including diuretics, for renal tubular secretion; see below.

- When the displacing drug additionally reduces elimination of the first, so that the free concentration is increased not only acutely but also chronically at the new steady state, severe toxicity may ensue.

Although many drugs have appreciable affinity for plasma albumin (Ch. 8), and therefore might potentially be expected to interact in these ways, there are rather few instances of clinically important interactions of this type. Protein-bound drugs that are given in large enough dosage to act as displacing agents include various *sulfonamides* and **chloral hydrate**; trichloracetic acid, a metabolite of chloral hydrate, binds very strongly to plasma albumin. Displacement of bilirubin from albumin by such drugs in jaundiced premature neonates can have clinically disastrous consequences: bilirubin metabolism is undeveloped in the premature liver, and unbound bilirubin can cross the immature blood–brain barrier and cause kernicterus (staining of the basal ganglia by bilirubin). This causes a distressing and permanent disturbance of movement known as choreoathetosis, characterised by involuntary writhing and twisting movements in the child.

Phenytoin dose is adjusted according to measurement of its concentration in plasma, and such measurements do not routinely distinguish bound from free phenytoin (that is, they reflect the total concentration of drug). Introduction of a displacing drug in an epileptic patient whose condition is stabilised on phenytoin (Ch. 44) reduces the total plasma phenytoin concentration owing to increased elimination of free drug, but there is no loss of efficacy because the concentration of unbound (active) phenytoin at the new steady state is unaltered. If it is not appreciated that the therapeutic range of plasma concentrations has been reduced in this way, an increased dose may be prescribed, resulting in toxicity.

There are several instances where drugs that alter protein binding additionally reduce elimination of the displaced drug, causing clinically important interactions. **Phenylbutazone** displaces **warfarin** from binding sites on albumin, and more importantly selectively inhibits metabolism of the pharmacologically active (*S*) isomer (see below), prolonging prothrombin time and resulting in increased bleeding (Ch. 24). *Salicylates* displace **methotrexate** from binding sites on albumin and reduce its secretion into the nephron by competition with the organic anion transporter (OAT; Ch. 9). **Quinidine** and several other antidysrhythmic drugs including **verapamil** and **amiodarone** (Ch. 21) displace **digoxin** from tissue-binding sites while simultaneously reducing its renal excretion; they consequently can cause severe dysrhythmias through digoxin toxicity.

DRUG METABOLISM

Drugs can either induce (Table 56.2) or inhibit (Table 56.3) drug-metabolising enzymes.

Enzyme induction

Enzyme induction (e.g. by anticonvulsants, ethanol or **rifampicin**; see Ch. 9) is an important cause of drug interaction. The slow onset of induction and slow recovery after withdrawal of the inducing agent together with the potential for selective induction of one or more CYP isoenzymes contributes to the insidious nature of the clinical problems that induction presents. Adverse clinical outcomes from such interactions are very diverse including graft rejection as a result of loss of effectiveness of immunosuppressive treatment, seizures due to loss of anticonvulsant effectiveness, unwanted pregnancy and thrombosis (from loss of effectiveness of warfarin) or bleeding (from failure to recognise the need to reduce warfarin dose when induction wanes). Over 200 drugs cause enzyme induction and thereby decrease the pharmacological activity of a range of other drugs. Some examples are given in Table 56.2. Because the inducing agent is often itself a substrate for the induced enzymes, the process can result in slowly developing tolerance. This pharmacokinetic kind of tolerance is generally less marked than pharmacodynamic tolerance, for example to opioids (Ch. 41), but it is clinically important when starting treatment with **carbamazepine** (Ch. 44). This is initiated at a low dose to avoid toxicity (because liver enzymes are not induced initially) and gradually increased over a period of a few weeks, during which it induces its own metabolism.

Figure 56.3 shows how the antibiotic **rifampicin**, given for 3 days, reduces the effectiveness of **warfarin** as an anticoagulant. Conversely, enzyme induction can increase toxicity of a second drug if the toxic effects are mediated via an active metabolite. **Paracetamol** (**acetaminophen**) toxicity is a case in point (see Fig. 57.1): this is caused by its CYP metabolite *N*-acetyl-*p*-benzoquinone imine. Consequently, the risk of serious hepatic injury following paracetamol

Table 56.2 Examples of drugs that induce drug-metabolising enzymes

Drugs inducing enzyme action	Drugs with metabolism affected
Phenobarbital	Warfarin
Rifampicin	Oral contraceptives
Griseofulvin	Corticosteroids
Phenytoin	Ciclosporin
Ethanol	Drugs listed in left-hand column will also be affected
Carbamazepine	

Table 56.3 Examples of drugs that inhibit drug-metabolising enzymes

Drugs inhibiting enzyme action	Drugs with metabolism affected
Allopurinol	Mercaptopurine, azathioprine
Chloramphenicol	Phenytoin
Cimetidine	Amiodarone, phenytoin, pethidine
Ciprofloxacin	Theophylline
Corticosteroids	Tricyclic antidepressants, cyclophosphamide
Disulfiram	Warfarin
Erythromycin	Ciclosporin, theophylline
Monoamine oxidase inhibitors	Pethidine
Ritonavir	Saquinavir

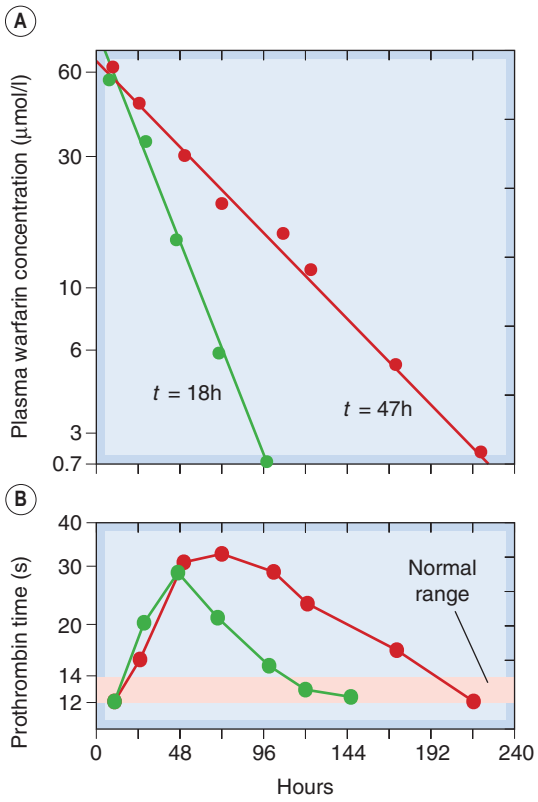

Fig. 56.3 Effect of rifampicin on the metabolism and anticoagulant action of warfarin. [**A**] Plasma concentration of warfarin (log scale) as a function of time following a single oral dose of 5 μmol/kg body weight. After the subject was given rifampicin (600 mg daily for a few days), the plasma half-life of warfarin decreased from 47 h (red curve) to 18 h (green curve). [**B**] The effect of a single dose of warfarin on prothrombin time under normal conditions (red curve) and after rifampicin administration (green curve). (Redrawn from O'Reilly 1974 Ann Intern Med 81: 337.)

overdose is increased in patients in whom CYP has been induced, for example by chronic alcohol consumption. Variability in rates of drug metabolism between individuals results partly from varying exposure to environmental chemicals, some of which are powerful enzyme inducers.

Enzyme induction is exploited therapeutically by administering **phenobarbital** to premature babies to induce glucuronyltransferase, thereby increasing bilirubin conjugation and reducing the risk of kernicterus (see above).

Enzyme inhibition

Enzyme inhibition, particularly of CYP enzymes, slows the metabolism and hence increases the action of other drugs inactivated by the enzyme. Such effects can be clinically important and are major considerations in the treatment of patients with HIV infection with triple and quadruple therapy, because several protease inhibitors are potent CYP inhibitors (Ch. 51). Other examples of drugs that are enzyme inhibitors are shown in Table 56.3. To make life even more difficult, several inhibitors of drug metabolism influence the metabolism of different stereoisomers selectively. Examples of drugs that inhibit the metabolism of the active (*S*) and less active (*R*) isomers of warfarin in this way are shown in Table 56.4.

Table 56.4 Stereoselective and non-stereoselective inhibition of warfarin metabolism

Inhibition of metabolism	Drug(s)
Stereoselective for (*S*) isomer	Phenylbutazone Metronidazole Sulfinpyrazone Trimethoprim–sulfamethoxazole Disulfiram
Stereoselective for (*R*) isomer	Cimetidine[a] Omeprazole[a]
Non-stereoselective effect on both isomers	Amiodarone

[a] Minor effect only on prothrombin time.
From Hirsh 1991 N Engl J Med 324: 1865–1875.

The therapeutic effects of some drugs are a direct consequence of enzyme inhibition (e.g. the xanthine oxidase inhibitor **allopurinol**, used to prevent gout; Ch. 26). Xanthine oxidase metabolises several cytotoxic and immunosuppressant drugs, including **mercaptopurine** (the active metabolite of **azathioprine**), the action of which is thus potentiated and prolonged by allopurinol. **Disulfiram**, an inhibitor of aldehyde dehydrogenase used to produce an aversive reaction to ethanol (see Ch. 48), also inhibits metabolism of other drugs, including **warfarin**, which it potentiates. **Metronidazole**, an antimicrobial used to treat anaerobic bacterial infections and several protozoal diseases (Chs 50 and 53), also inhibits this enzyme, and patients prescribed it are advised to avoid alcohol for this reason.

In other instances, inhibition of drug metabolism is less expected because enzyme inhibition is not the main mechanism of action of the offending agents. Thus, *glucocorticosteroids* and **cimetidine** enhance the actions of a range of drugs including some antidepressant and cytotoxic drugs.

When a drug works through an active metabolite, inhibition of its metabolism can result in *loss* of activity. An example of topical concern is an interaction between proton pump inhibitors (such as **omeprazole**, Ch. 29) and the antiplatelet drug **clopidogrel** (Ch. 24). These have been widely co-prescribed (because clopidogrel is often used with other antithrombotic drugs so there is a high risk of bleeding from the stomach—omeprazole reduces this). Clopidogrel works through an active metabolite formed by CYP2C19 (people who carry a genetic variant where CYP2C19 is less active have an increased risk of thrombosis during treatment with clopidogrel). Omeprazole inhibits CYP2C19, and also reduces the antiplatelet action of clopidogrel. It is not yet clear how clinically important this may be, but the Food and Drug Administration has warned against concomitant use of these drugs (http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm190848.htm). It is anticipated that other enzyme inhibitors will have a similar deleterious effect.

As with induction, interactions caused by enzyme inhibition are hard to anticipate from first principles. If in doubt about the possibility of an interaction, it is best to look it up (e.g. in the *British National Formulary*, which has

Table 56.5 Examples of drugs that inhibit renal tubular secretion

Drug(s) causing inhibition	Drug(s) affected
Probenecid, Sulfinpyrazone, Phenylbutazone, Sulfonamides, Aspirin, Thiazide diuretics, Indometacin	Penicillin, Azidothymidine, Indometacin
Verapamil, Amiodarone, Quinidine	Digoxin
Indometacin	Furosemide (frusemide)
Aspirin, Non-steroidal anti-inflammatory drugs	Methotrexate

an invaluable appendix on drug interactions indicating which are of known clinical importance).

Haemodynamic effects

Variations in hepatic blood flow influence the rate of inactivation of drugs that are subject to extensive presystemic hepatic metabolism (e.g. **lidocaine**, **propranolol**). A reduced cardiac output reduces hepatic blood flow, so drugs that reduce cardiac output (e.g. propranolol) reduce the rate of metabolism of lidocaine by this mechanism. Extraction of lidocaine by liver approaches 100% and measurement of lidocaine clearance has been used to estimate hepatic blood flow in the same way as clearance of *p*-aminohippuric acid (PAH) has been used to estimate renal blood flow (Ch. 9).

DRUG EXCRETION

The main mechanisms by which one drug can affect the rate of renal excretion of another are by:

Drug interactions

- These are many and varied: if in doubt, look it up.
- Interactions may be pharmacodynamic or pharmacokinetic.
- Pharmacodynamic interactions are often predictable from the actions of the interacting drugs.
- Pharmacokinetic interactions can involve effects on:
  - absorption
  - distribution (e.g. competition for protein binding)
  - hepatic metabolism (induction or inhibition)
  - renal excretion.

- altering protein binding, and hence filtration
- inhibiting tubular secretion
- altering urine flow and/or urine pH.

Inhibition of tubular secretion

Probenecid (Ch. 28) was developed to inhibit **penicillin** secretion and thus prolong its action. It also inhibits the excretion of other drugs, including **zidovudine** (see Ch. 51). Other drugs have an incidental probenecid-like effect and can enhance the actions of substances that rely on tubular secretion for their elimination. Table 56.5 gives some examples. Because diuretics act from within the tubular lumen, drugs that inhibit their secretion into the tubular fluid, such as NSAIDs, reduce their effect.

Alteration of urine flow and pH

Diuretics tend to increase the urinary excretion of other drugs and their metabolites, but this is seldom immediately clinically important. Conversely, loop and thiazide diuretics indirectly increase the proximal tubular reabsorption of **lithium** (which is handled in a similar way as Na^+), and this can cause lithium toxicity in patients treated with lithium carbonate for mood disorders (Ch. 46). The effect of urinary pH on the excretion of weak acids and bases is put to use in the treatment of poisoning with *salicylate* (see Ch. 8), but is not a cause of accidental interactions.

REFERENCES AND FURTHER READING

Further reading

Barry, M., Mulcahy, F., Merry, C. et al., 1999. Pharmacokinetics and potential interactions amongst antiretroviral agents used to treat patients with HIV infection. Clin. Pharmacokinet. 36, 289–304. (*Multidrug combinations have transformed the outlook for patients with HIV infection; drug interactions are one of the main problems associated with these*)

Carmichael, D.J.S., 2005. Handling of drugs in kidney disease. In: Davison, A.M., Cameron, J.S., Grunfeld, J.P., et al. (Eds.), Oxford textbook of clinical nephrology, third ed. Oxford University Press, Oxford, pp. 2599–2618. (*Principles and practice of dose adjustment in patients with renal failure*)

Cooper, R.S., Kaufman, J.S., Ward, R., 2003. Race and genomics. N. Engl. J. Med. 348, 1166–1170. (*Scholarly and appropriately sceptical analysis*)

Fugh-Berman, A., Ernst, E., 2001. Herb–drug interactions: review and assessment of report reliability. Br. J. Clin. Pharmacol. 52, 587–595. (*Warfarin the most common drug, St John's wort the most common herb; more data needed! See also Fugh-Berman, A., 2000. Lancet 355, 134–138*)

Hanratty, C.G., McGlinchey, P., Johnston, G.D., Passmore, A.P., 2000. Differential pharmacokinetics of digoxin in elderly patients. Drugs Aging 17, 353–362. (*Reviews pharmacokinetics of digoxin in relation to age, concomitant disease and interacting drugs*)

Lin, J.H., Liu, A.Y.H., 2001. Interindividual variability in inhibition and induction of cytochrome P450 enzymes. Annu. Rev. Pharmacol. Toxicol. 41, 535–567. (*Examines sources of interindividual variability in inhibition and induction of P450 enzymes*)

Rowland, M., Tozer, T.N., 2010. Clinical pharmacokinetics and pharmacodynamics. Concepts and applications. Wolters Kluwer/ Lippincott, Williams & Wilkins, Baltimore. (*Online simulations by H Derendorf and G Hochhaus*)

Sproule, B.A., Hardy, B.G., Shulman, K.I., 2000. Differential pharmacokinetics in elderly patients. Drugs Aging 16, 165–177. (*Reviews age-related changes in pharmacodynamics as well as pharmacokinetics and drug interactions, all of which are clinically important*)

Westphal, J.F., 2000. Macrolide-induced clinically relevant drug interactions with cytochrome P450A (CYP) 3A4: an update focused on clarithromycin, azithromycin and dirithromycin. Br. J. Clin. Pharmacol. 50, 285–295. (*Review: theophylline, ciclosporine, warfarin, involvement of P-glycoprotein as well as metabolism*)

Xie, H.-G., Kim, R.B., Wood, A.J.J., Stein, C.M., 2001. Molecular basis of ethnic differences in drug disposition and response. Annu. Rev. Pharmacol. Toxicol. 41, 815–850. (*Recent developments in understanding genetic variations that may underlie ethnic differences in drug-metabolising enzymes, transporters, receptors and second messenger systems*)

Zevin, S., Benowitz, N.L., 1999. Drug interactions with tobacco smoking—an update. Clin. Pharmacokinet. 36, 425–438. (*Polycyclic aromatic hydrocarbons in tobacco smoke induce various P450 enzymes. 'Cigarette smoking should be specifically studied in clinical trials of new drugs'*)

References

Atkinson, A.J. Jr, Abernethie, D.R., Daniels, C.E., et al., 2006. Principles of clinical pharmacology, second ed. Academic Press, San Diego. (*Includes detailed accounts of clinical aspects including effects of renal and liver disease on pharmacokinetics, of effects of age and on drug therapy in pregnant and nursing women*)

Wadman, M., 2005. Drug targeting: is race enough? Nature 435, 1008–1009. (*No!*)

Wood, A.J.J., 2001. Racial differences in response to drugs—pointers to genetic differences. N. Engl. J. Med. 344, 1393–1396.

57 Harmful effects of drugs

OVERVIEW

This chapter addresses *harmful* effects of drugs, both in the context of therapeutic use—so-called *adverse drug reactions*, and of deliberate *overdose*. The classification of adverse drug reactions is considered, followed by aspects of drug toxicity: toxicity testing in drug development, mechanisms of toxin-induced cell damage, mutagenesis and carcinogenicity, teratogenesis and allergic reactions.

INTRODUCTION

Paracelsus, a 16th-century alchemist, is credited with the aphorism that all drugs are poisons: '… the dosage makes it either a poison or a remedy'. Today, toxic effects of drugs remain clinically important in the context of deliberate overdose (self-poisoning accounts for approximately 10% of the workload of emergency medicine departments in the UK; by contrast, homicidal poisoning, while obviously important, is extremely uncommon). Some susceptible individuals may experience dose-related toxicity even during therapeutic dosing; some of this susceptibility is genetically determined, and genomic testing as a means of avoiding such harms is beginning to make its way into the clinic (Ch. 11).

Rigorous toxicity testing in animals (see below), including tests for carcinogenicity, teratogenicity and organ-specific toxicities, is carried out on potential new drugs during development (see Ch. 60), and in many cases leads to abandonment of the compound before it is tested in humans. Such animal toxicity studies form part of the package of information routinely submitted to drug regulatory agencies when seeking approval to market a new drug. Such studies do sometimes usefully focus attention on a particular organ, the function of which can be monitored prospectively during human studies. Nevertheless, harmful effects are often encountered during therapeutic use, often the result of misprescribing, but also due to the emergence of toxic effects not detected in animals. These harms are usually referred to as 'adverse drug reactions' (ADRs) and are of great concern to drug regulatory authorities, which are charged with establishing the safety as well as the efficacy of drugs. Unpredictable events are of particular concern. Some ADRs are a consequence of the main pharmacological effect of the drug but some (e.g. immunological reactions), are not. *Safety* (as distinct from *toxicity*) of new drugs can only be established during drug development and therapeutic use in humans (Walker, 2004).

Clinically important ADRs are common, costly and avoidable (see Pirmohamed et al., 2004).[1] Any organ can be the principal target, and several systems can be involved simultaneously. The time course helps to recognise a clinical event as an ADR. Several patterns are recognised. The symptoms sometimes closely shadow drug administration and discontinuation, but in other cases adverse effects only occur during prolonged use (*osteoporosis* during continued high-dose glucocorticoid therapy [Ch. 32], or *tardive dyskinesia* during continuous use of antipsychotic drugs [Ch. 45], for example). Some adverse effects occur on ending treatment, either within a few days (e.g. tachycardia on abrupt discontinuation of β-adrenoceptor blockade) or after a delay, first appearing months or years after treatment is discontinued, as in the case of some second malignancies following successful chemotherapy. Consequently, anticipating, avoiding, recognising and responding to adverse drug reactions are among the most challenging and important parts of clinical practice.

CLASSIFICATION OF ADVERSE DRUG REACTIONS

Harmful effects of drugs are either related or unrelated to the principal pharmacological action of the drug. Aronson & Ferner (2003) have suggested that ADRs be described according to the **do**se, **t**ime course and **s**usceptibility (DoTS).

ADVERSE EFFECTS RELATED TO THE MAIN PHARMACOLOGICAL ACTION OF THE DRUG

Many adverse effects related to the main pharmacological action of the drug are predictable, at least if this action is well understood. They are sometimes referred to as type A ('augmented') adverse reactions (Rawlins & Thomson, 1985) and are related to dose and susceptibility. Many such reactions have been described in previous chapters. For example, postural hypotension occurs with α_1-adrenoceptor antagonists, bleeding with anticoagulants, sedation with anxiolytics and so on. In many instances, this type of unwanted effect is reversible, and the problem can often be dealt with by reducing the dose. Such effects are sometimes serious (e.g. intracerebral bleeding caused by anticoagulants, hypoglycaemic coma from insulin), and occasionally they are not easily reversible, for example drug dependence produced by opioid analgesics (see Ch. 48).

Some adverse effects related to the main action of a drug result in discrete events rather than graded symptoms, and can be difficult to detect. For example, drugs that block cyclo-oxygenase (COX)-2 (including 'coxibs', for example **rofecoxib**, **celecoxib**, **valdecoxib**, as well as some conventional non-steroidal anti-inflammatory drugs, NSAIDs) increase the risk of myocardial infarction in a dose-dependent manner (Ch. 26). This potential was apparent from the pharmacology of these drugs, in particular their ability to inhibit prostacyclin biosynthesis as well as to

[1] 6.5% of hospital admissions were due to ADRs at a projected annual cost of £466 million in the UK. Antiplatelet drugs, diuretics, non-steroidal anti-inflammatory drugs and anticoagulants between them accounted for 50% of the ADRs. Most events were avoidable and 2.3% of the patients died.

increase arterial blood pressure, and early studies gave a hint of such problems. The effect was difficult to prove because of the high background incidence of coronary thrombosis, and it was only when placebo-controlled trials were performed for another indication (in the hope that COX-2 inhibitors could prevent bowel cancer) that this effect was confirmed unequivocally.

ADVERSE EFFECTS UNRELATED TO THE MAIN PHARMACOLOGICAL ACTION OF THE DRUG

Adverse effects unrelated to the main pharmacological effect may be predictable when a drug is taken in excessive dose, for example **paracetamol** hepatotoxicity (see below) or **aspirin**-induced tinnitus; or when susceptibility is increased, for example during pregnancy or by a predisposing disorder such as glucose 6-phosphate dehydrogenase deficiency or a mutation in the mitochondrial DNA that predisposes to aminoglycoside ototoxicity (Ch. 11).

Unpredictable idiosyncratic reactions are often initiated by a chemically reactive metabolite rather than the parent drug. Examples of such ADRs, which are often immunological in nature, include drug-induced hepatic or renal necrosis, bone marrow suppression, carcinogenesis and disordered fetal development. Uncommon but severe unpredictable adverse effects that have been mentioned in earlier chapters include aplastic anaemia from **chloramphenicol** and anaphylaxis in response to **penicillin**. These idiosyncratic reactions are termed type B ('bizarre') in the Rawlins & Thomson (1985) classification. They are usually severe—otherwise they would go unrecognised—and their existence is important in establishing the safety of medicines.

▼ If the incidence of an adverse reaction is 1 in 6000 patients exposed, approximately 18 000 patients would have to be exposed to the drug for three events to occur, and approximately double that number for three events to be detected and their possible relationship to the drug recognised and reported, even if there were no background incidence of the event in question. Consequently, such reactions cannot be excluded by preapproval clinical trials (which might typically expose only a few thousand individuals to the drug), and the association may come to light only after years of use, so there is a need for continued monitoring by regulatory authorities after drugs have been licensed and marketed. An example is the association between pulmonary hypertension and valvular heart disease with **fenfluramine**, an appetite suppressant that had been used for several years, and with **dexfenfluramine**, its pharmacologically active isomer. Such experiences call for a balanced approach to prescribing new drugs if there are adequate existing alternatives.[2] This conflicts with the culture of drug marketing, especially when this involves advertising the product direct to the consumer.

DRUG TOXICITY

TOXICITY TESTING

Toxicity testing in animals is carried out on new drugs to identify potential hazards before administering them to humans. It involves the use of a wide range of tests in different species, with long-term administration of the drug, regular monitoring for physiological or biochemical abnormalities, and a detailed postmortem examination at the end of the trial to detect any gross or histological abnormalities. Recently, use of non-mammalian species, notably the transparent zebra fish, has shown promise as an intermediate stage between toxicity studies on cells and tissues in vitro and mammalian toxicity testing (see Parng, 2005, for a review). Toxicity testing is performed with doses well above the expected therapeutic range, and establishes which tissues or organs are likely 'targets' of toxic effects of the drug. Recovery studies are performed to assess whether toxic effects are reversible, and particular attention is paid to irreversible changes such as carcinogenesis or neurodegeneration. The basic premise is that toxic effects caused by a drug are similar in humans and other animals. This is inherently reasonable in view of the similarities between higher organisms at the cellular and molecular levels. There are, nevertheless, wide interspecies variations, especially in metabolising enzymes; consequently, a toxic metabolite formed in one species may not be formed in another, and so toxicity testing in animals is not always a reliable guide. **Pronethalol**, the first β-adrenoceptor antagonist synthesised (by James Black) at ICI, was not developed because it caused carcinogenicity in mice; it subsequently emerged that carcinogenicity occurred only in the ICI strain—but by then other β-blockers were already in development.

Toxic effects can range from negligible to so severe as to preclude further development of the compound. Intermediate levels of toxicity are more acceptable in drugs intended for severe illnesses (e.g. AIDS or cancers), and decisions on whether or not to continue development are often difficult. If development does proceed, safety monitoring can be concentrated on the system 'flagged'

Types of drug toxicity

- Toxic effects of drugs can be:
  - related to the principal pharmacological action (e.g. bleeding with anticoagulants)
  - unrelated to the principal pharmacological action (e.g. liver damage with **paracetamol**).
- Some adverse reactions that occur with ordinary therapeutic dosage are unpredictable, serious and uncommon (e.g. agranulocytosis with **carbimazole**). Such idiosyncratic reactions are almost inevitably detected only after widespread use of a new drug.
- Adverse effects unrelated to the main action of a drug are often caused by reactive metabolites and/or immunological reactions.

[2]Hesitation in prescribing a newly licensed drug may delay recognition of an ADR without reducing the total number of patients harmed, so a 'cautious physician' is one who prefers to let others take the risk. Grant of a product licence to a company to market a new drug does not require evidence of superiority over existing treatments, so from the patient's perspective a physician who is neither at the forefront of fashion nor the last to adopt a genuine advance may be the best bet in an uncertain world.

[3]The value of toxicity testing is illustrated by experience with **triparanol**, a cholesterol-lowering drug marketed in the USA in 1959. Three years later, a team from the Food and Drug Administration, acting on a tip-off, paid the manufacturer a surprise visit that revealed falsification of toxicology data demonstrating cataracts in rats and dogs. The drug was withdrawn, but some patients who had been taking it for a year or more also developed cataracts. Regulatory authorities now require that toxicity testing is performed under a tightly defined code of practice (Good Laboratory Practice), which incorporates many safeguards to minimise the risk of error or fraud.

as a potential target of toxicity by the animal studies.[3] *Safety* of a drug (as distinct from toxicity) can be established only during use in humans.

GENERAL MECHANISMS OF TOXIN-INDUCED CELL DAMAGE AND CELL DEATH

Toxic concentrations of drugs or drug metabolites can cause necrosis; however, programmed cell death (apoptosis; see Ch. 5) is increasingly recognised to be of paramount importance, especially in chronic toxicity (see, for example, Pirmohamed, 2003).

Chemically reactive drug metabolites can form covalent bonds with target molecules as well as damage tissue by non-covalent mechanisms. The liver is of great importance in drug metabolism (Ch. 9), and hepatocytes are exposed to high concentrations of nascent metabolites. Drugs and their polar metabolites are concentrated in renal tubular fluid as water is reabsorbed, so renal tubules are exposed to higher concentrations than are other tissues. Several hepatotoxic drugs (e.g. **paracetamol**) are also nephrotoxic. Consequently, hepatic or renal damage are common reasons for abandoning development of drugs during toxicity testing.

General mechanisms of cell damage and cell death

- Drug-induced cell damage/death is usually caused by reactive metabolites of the drug, involving non-covalent and/or covalent interactions with target molecules. Cell death is often 'self-inflicted', via triggering apoptosis.
- Non-covalent interactions include:
  - lipid peroxidation via a chain reaction
  - generation of cytotoxic reactive oxygen species
  - depletion of reduced glutathione
  - modification of sulfhydryl groups on key enzymes (e.g. Ca^{2+}-ATPase) and structural proteins.
- Covalent interactions, for example adduct formation between a metabolite of paracetamol (NAPBQI: *N*-acetyl-*p*-benzoquinone imine) and cellular macromolecules (Fig. 57.1). Covalent binding to protein can produce an immunogen; binding to DNA can cause carcinogenesis and teratogenesis.

NON-COVALENT INTERACTIONS

▼ Reactive metabolites of drugs are implicated in several potentially cytotoxic, non-covalent processes, including:

- lipid peroxidation
- generation of toxic reactive oxygen species
- depletion of reduced glutathione (GSH)
- modification of sulfhydryl groups.

Lipid peroxidation

▼ Peroxidation of unsaturated lipids can be initiated either by reactive metabolites or by reactive oxygen species (see below). Lipid peroxyradicals ($ROO^{\bullet}$) can produce lipid hydroperoxides (ROOH), which produce further lipid peroxyradicals. This chain reaction—a peroxidative cascade—may eventually affect much of the membrane lipid. Defence mechanisms, for example GSH peroxidase and vitamin E, protect against this. Cell damage results from alteration of membrane permeability or from reactions of the products of lipid peroxidation with proteins.

Reactive oxygen species

▼ Reduction of molecular oxygen to superoxide anion ($O_2^{-\bullet}$) may be followed by enzymic conversion to hydrogen peroxide (H_2O_2), hydroperoxy ($HOO^{\bullet}$) and hydroxyl ($OH^{\bullet}$) radicals or singlet oxygen. These reactive oxygen species are cytotoxic, both directly and through lipid peroxidation (see above), and are important in excitotoxicity and neurodegeneration (Ch. 39, Fig. 39.1).

Depletion of glutathione

▼ The GSH redox cycle protects cells from oxidative stress. GSH can be depleted by accumulation of normal oxidative products of cell metabolism, or by the action of toxic chemicals. GSH is normally maintained in a redox couple with its disulfide, GSSG. Oxidising species convert GSH to GSSG, GSH being regenerated by NADPH-dependent GSSG reductase. When cellular GSH falls to about 20–30% of normal, cellular defence against toxic compounds is impaired and cell death can result.

Modification of sulfhydryl groups

▼ Modification of sulfhydryl groups can be produced either by oxidising species that alter sulfhydryl groups reversibly or by covalent interaction. Free sulfhydryl groups have a critical role in the catalytic activity of many enzymes. Important targets for sulfhydryl modification by reactive metabolites include the cytoskeletal protein actin, GSH reductase (see above) and Ca^{2+}-transporting ATPases in the plasma membrane and endoplasmic reticulum. These maintain cytoplasmic Ca^{2+} concentration at approximately 0.1 µmol/l in the face of an extracellular Ca^{2+} concentration of more than 1 mmol/l. A sustained rise in cell Ca^{2+} occurs with inactivation of these enzymes (or with increased membrane permeability; see above), and this compromises cell viability. Lethal processes leading to cell death after acute Ca^{2+} overload include activation of degradative enzymes (neutral proteases, phospholipases, endonucleases) and protein kinases, mitochondrial damage and cytoskeletal alterations (e.g. modification of association between actin and actin-binding proteins).

COVALENT INTERACTIONS

Targets for covalent interactions include DNA, proteins/peptides, lipids and carbohydrates. Covalent bonding to DNA is a basic mechanism of mutagenic chemicals; this is dealt with below. Several non-mutagenic chemicals also form covalent bonds with macromolecules, but the relationship between this and cell damage is incompletely understood. For example, the cholinesterase inhibitor paraoxon (the active metabolite of the insecticide parathion) binds acetylcholinesterase at the neuromuscular junction (Ch. 13) and causes necrosis of skeletal muscle. One toxin from an exceptionally poisonous toadstool, *Amanita phalloides*, binds actin, and another binds RNA polymerase, interfering with actin depolymerisation and protein synthesis, respectively.

HEPATOTOXICITY

Many therapeutic drugs cause liver damage, manifested clinically as hepatitis or (in less severe cases) only as laboratory abnormalities (e.g. increased activity of plasma aspartate transaminase, an enzyme released from damaged liver cells). **Paracetamol**, **iproniazid** and **halothane** cause hepatotoxicity by the mechanisms of cell damage outlined above. Genetic differences in drug metabolism (see Ch. 11) have been implicated in some instances (e.g. **isoniazid**, **phenytoin**). Mild drug-induced abnormalities of liver function are not uncommon, but the mechanism of liver injury is often uncertain (e.g. *statins*; Ch. 23). It is not always necessary to discontinue a drug when such mild laboratory abnormalities occur, but the occurrence of cirrhosis as a

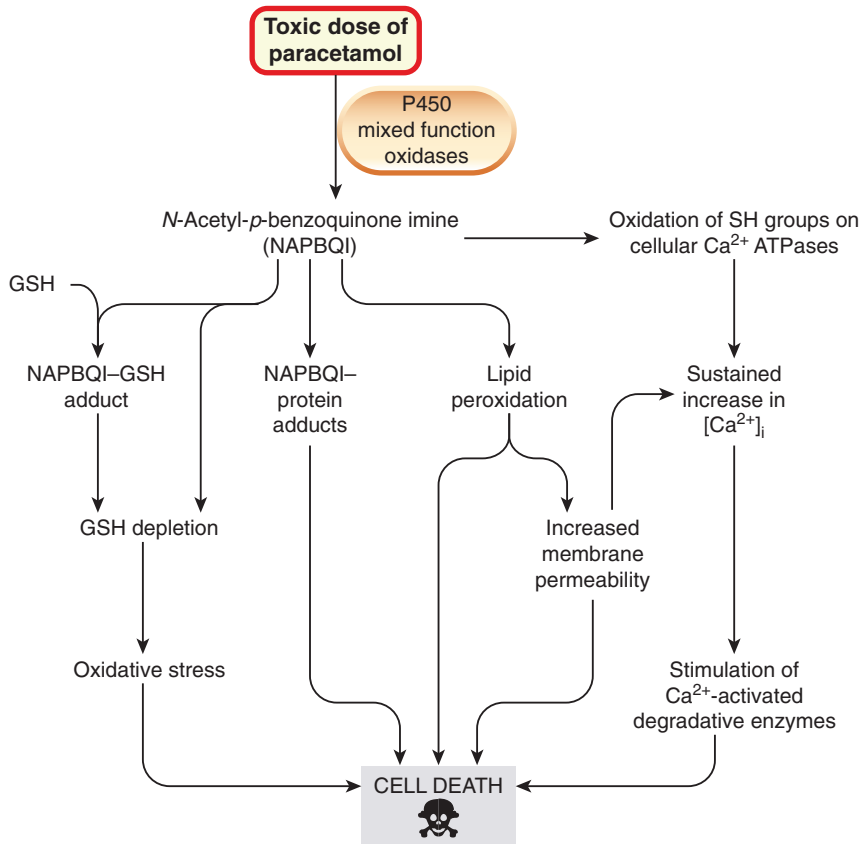

Fig. 57.1 Potential mechanisms of liver cell death resulting from the metabolism of paracetamol to *N*-acetyl-*p*-benzoquinone imine (NAPBQI). GSH, glutathione. (Based on data from Boobis A R et al. 1989 Trends Pharmacol Sci 10: 275–280 and Nelson S D, Pearson P G 1990 Annu Rev Pharmacol Toxicol 30: 169.)

result of long-term low-dose **methotrexate** treatment for arthritis or psoriasis (a chronic scaling skin disease of unknown cause that is usually mild, if tiresome, but can rarely be very severe[4]) argues for caution. Hepatotoxicity of a different kind, namely reversible obstructive jaundice, occurs with **chlorpromazine** (Ch. 45) and androgens (Ch. 34).

Hepatotoxicity caused by **paracetamol** overdose remains a common cause of death following self-poisoning. An outline is given in Chapter 26. Because the body's handling of this drug exemplifies many of the general mechanisms of cell damage outlined above, the story is taken up again here. With toxic doses of paracetamol, the enzymes catalysing the normal conjugation reactions are saturated, and mixed-function oxidases convert the drug to the reactive metabolite *N*-acetyl-*p*-benzoquinone imine (NAPBQI). As explained in Chapters 9 and 56, paracetamol toxicity is increased in patients in whom P450 enzymes have been induced, for instance by chronic excessive consumption of alcohol. NAPBQI initiates several of the covalent and non-covalent interactions described above and illustrated in Figure 57.1. Oxidative stress from GSH depletion is important in leading to cell death. Regeneration of GSH from GSSG depends on the availability of cysteine, the intracellular availability of which can be limiting. *Acetylcysteine* or *methionine* can substitute for cysteine, increasing GSH availability and reducing mortality in patients with paracetamol poisoning.

Liver damage can also be produced by immunological mechanisms (see below), which have been particularly implicated in **halothane** hepatitis (see Ch. 40).

Hepatotoxicity

- Hepatocytes are exposed to reactive metabolites of drugs as these are formed by P450 enzymes.
- Liver damage is produced by several mechanisms of cell injury; **paracetamol** exemplifies many of these (see Fig. 57.1).
- Some drugs (e.g. **chlorpromazine**) can cause reversible cholestatic jaundice.
- Immunological mechanisms are sometimes implicated (e.g. **halothane**).

NEPHROTOXICITY

Drug-induced nephrotoxicity is a common clinical problem: NSAIDs (Table 57.1) and angiotensin-converting enzyme (ACE) inhibitors are among the commoner precipitants of

[4]Aficionados of Dennis Potter will recall the protagonist in the television drama *The Singing Detective*; Potter was himself afflicted by the most severe form of the disease.

Table 57.1 Adverse effects of non-steroidal anti-inflammatory drugs on the kidney

Cause	Adverse effects
Principal pharmacological action (i.e. inhibition of prostaglandin biosynthesis)	Acute ischaemic renal failure Sodium retention (leading to or exacerbating hypertension and/or heart failure) Water retention Hyporeninaemic hypoaldosteronism (leading to hyperkalaemia)
Unrelated to principal pharmacological action (allergic-type interstitial nephritis)	Renal failure Proteinuria
Unknown whether or not related to principal pharmacological action (analgesic nephropathy)	Papillary necrosis Chronic renal failure

Adapted from Murray & Brater 1993.

acute renal failure. This is usually caused by the principal pharmacological actions of these drugs, which, although well tolerated in healthy people, cause renal failure in patients with diseases that jeopardise glomerular filtration. In patients with heart or liver disease, glomerular filtration rate (GFR) depends critically on vasodilator prostaglandin biosynthesis. This is inhibited by NSAIDs (Ch. 26), and hence these drugs reduce renal perfusion in such patients. Similarly, in patients with bilateral renal artery stenosis (i.e. narrowings of the renal arteries, most often caused by fibromuscular tissue in young women or by atheromatous disease in older people), GFR depends on angiotensin II-mediated efferent arteriolar vasoconstriction (which is inhibited by ACE inhibitors; Ch. 22); acute renal impairment occurs on starting treatment with an ACE inhibitor and is reversible if the drug is discontinued promptly. Additionally, NSAIDs indirectly depress renin and aldosterone secretion by inhibiting renal prostaglandin I_2 biosynthesis, and ACE inhibitors depress angiotensin II-stimulated aldosterone secretion, leading to low renin/low aldosterone states ('hyporeninaemic hypoaldosteronism') that are particularly notable in diabetic patients. Reduced aldosterone can cause hyperkalaemia, especially if GFR is also reduced.

In addition to effects related to their main pharmacological action, NSAIDs can also cause interstitial nephritis through an immunological mechanism. This presents several months to 1 year after starting treatment as acute renal failure, often accompanied by eosinophil leukocytes in the urine and proteinuria, or as nephrotic syndrome (heavy proteinuria, hypoalbuminuria and oedema). **Fenoprofen** is particularly liable to cause this type of renal damage, possibly because its metabolites bind irreversibly to albumin. Penicillins (Ch. 50), especially **meticillin**, also cause interstitial nephritis.

Analgesic nephropathy is a third kind of renal damage in which NSAIDs are implicated (see also Ch. 26). This consists of renal papillary necrosis[5] and chronic interstitial nephritis. The clinical course is typically insidious but leads ultimately to end-stage chronic renal failure. It is associated with prolonged and massive overuse of analgesics. **Phenacetin** has been incriminated, but other NSAIDs have not been exonerated. The role of **caffeine** (often included with analgesics and NSAIDs in combined preparations for migraine) is uncertain but could be important. It is possible that such analgesic-associated nephropathy is causally related to inhibition of renal prostaglandin synthesis, but its pathogenesis is not understood.

Captopril, in higher doses than are currently recommended, can cause heavy proteinuria (Ch. 22). This is the result of glomerular injury, which is also caused by some other drugs that, like captopril, contain a sulfhydryl group (e.g. **penicillamine**; Ch. 26). It is therefore believed that it is this chemical feature rather than ACE inhibition per se that is responsible for this adverse effect.

Ciclosporin, used to prevent transplant rejection (Ch. 26), causes renal damage via renal vasoconstriction, which reduces GFR and causes hypertension.

Nephrotoxicity

- Renal tubular cells are exposed to high concentrations of drugs and metabolites as urine is concentrated.
- Renal damage can cause papillary and/or tubular necrosis.
- Inhibition of prostaglandin synthesis by non-steroidal anti-inflammatory drugs causes vasoconstriction and lowers glomerular filtration rate.

MUTAGENESIS AND CARCINOGENICITY

Chemical agents cause mutation by covalent modification of DNA. Certain kinds of mutation result in carcinogenesis, because the affected DNA sequence codes for a protein that regulates cell growth. It usually requires more than one mutation in a cell to initiate the changes that result in malignancy, mutations in proto-oncogenes (which regulate cell growth) and tumour suppressor genes (which code for products that inhibit the transcription of oncogenes) being particularly implicated (see Chs 5 and 55).

[5]The renal papilla is the part of the kidney exposed to the highest concentration of solutes, including drug metabolites; it also has a lower effective blood flow than other parts as a result of counter-current exchange in the vasa recta.

Mutagenesis and carcinogenicity

- Mutagenesis involves modification of DNA.
- Mutation of proto-oncogenes or tumour suppressor genes leads to carcinogenesis. More than one mutation is usually required.
- Drugs are relatively uncommon (but not unimportant) causes of birth defects and cancers.

BIOCHEMICAL MECHANISMS OF MUTAGENESIS

▼ Most chemical carcinogens act by modifying bases in DNA, particularly guanine, the O6 and N7 positions of which readily combine covalently with reactive metabolites of chemical carcinogens. Substitution at the O6 position is the more likely to produce a permanent mutagenic effect, because N7 substitutions are usually quickly repaired.

The accessibility of bases in DNA to chemical attack is greatest when DNA is in the process of replication (i.e. during cell division). The likelihood of genetic damage by many mutagens is therefore related to the frequency of cell division. The developing fetus is particularly susceptible, and mutagens are also potentially teratogenic (see below). This is also important in relation to mutagenesis of germ cells, particularly in girls, because in humans the production of primary oocytes occurs by a rapid succession of mitotic divisions very early in embryogenesis. Each primary oocyte then undergoes only two further divisions much later in life, at the time of ovulation. It is consequently during early pregnancy that germ cells of the developing female embryo are most likely to undergo mutagenesis, the mutations being transmitted to progeny conceived many years after exposure to the mutagen. In the male, germ cell divisions occur throughout life, and sensitivity of germ cells to mutagens is continuously present.

CARCINOGENESIS

Alteration of DNA is the first step in the complex, multistage process of carcinogenesis (see Ch. 5). Carcinogens are chemical substances that cause cancer, and can interact directly with DNA (genotoxic carcinogens) or act at a later stage to increase the likelihood that mutation will result in a tumour (epigenetic carcinogens; Fig. 57.2).

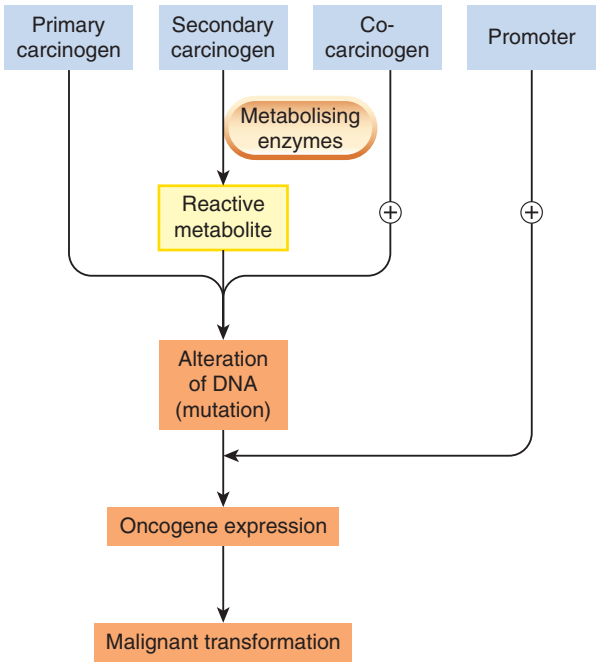

Fig. 57.2 Sequence of events in mutagenesis and carcinogenesis.

MEASUREMENT OF MUTAGENICITY AND CARCINOGENICITY

Much effort has gone into developing assays to detect mutagenicity and carcinogenicity. In vitro tests for *mutagenicity* are used for screening large numbers of compounds but as predictors of carcinogenicity can give false positive or false negative results. Whole-animal tests for carcinogenicity tests are expensive and time-consuming but are usually required by regulatory authorities before a new drug is licensed for use in humans. The main limitation of this kind of study is that there are important species differences, mainly to do with the metabolism of the foreign compound and the formation of reactive products.

The most widely used in vitro tests are variations on the *Ames test* for mutagenicity, which measures the rate of back-mutation (i.e. reversion from mutant to wild-type form) in *Salmonella typhimurium*.

▼ The wild-type strain can grow in a medium containing no added amino acids, because it can synthesise all the amino acids it needs from simple carbon and nitrogen sources. The test makes use of the fact that a mutant form of the organism cannot make histidine in this way and therefore grows only on a medium containing this amino acid. The test involves growing the mutant form on a medium containing a small amount of histidine, the drug to be tested being added to the culture. After several divisions, the histidine becomes depleted, and the only cells that continue dividing are those that have back-mutated to the wild type. A count of colonies following subculture on plates deficient in histidine gives a measure of the mutation rate.

Primary carcinogens cause mutation by a direct action on bacterial DNA, but most carcinogens have to be converted to an active metabolite (see above). Therefore it is necessary to include, in the culture, enzymes that catalyse the necessary conversion. An extract of liver from a rat treated with **phenobarbital** to induce liver enzymes is usually employed. There are many variations based on the same principle.

Other short-term in vitro tests for genotoxic chemicals include measurements of mutagenesis in mouse lymphoma cells, and assays for chromosome aberrations and sister chromatid exchanges in Chinese hamster ovary cells. However, all the in vitro tests give some false positive and some false negative results.

In vivo tests for carcinogenicity entail detection of tumours in groups of test animals. Carcinogenicity tests are inevitably slow, because there is usually a latency of months or years before tumours develop. Furthermore, tumours can develop spontaneously in control animals, and the results often provide only equivocal evidence of carcinogenicity of the test drug, making it difficult for industry and regulatory authorities to decide on further development and possible licensing of a product. None of the tests so far described can reliably detect epigenetic carcinogens. To do this, it is necessary to measure the effect of the test substance on tumour production with a threshold dose of a genotoxic agent. Such tests are being evaluated.

Few therapeutic drugs are known to increase the risk of cancer, the most important groups being drugs that act on DNA, i.e. cytotoxic and immunosuppressant drugs (Chs 55 and 26, respectively), and sex hormones (e.g. *oestrogens*, Ch. 34). **Pyrimethamine** (Ch. 53) is mutagenic in high concentrations, and carcinogenicity testing in strain A mice (but not other strains or species) was positive for a three-fold increase in lung tumours. **Methoxsalen** (a psoralen used

Carcinogens

- Carcinogens can be:
  - genotoxic, i.e. causing mutations directly (primary carcinogens) or after conversion to reactive metabolites (secondary carcinogens)
  - epigenetic, i.e. increasing the possibility that a mutagen will cause cancer, although not themselves mutagenic.
- New drugs are tested for mutagenicity and carcinogenicity.
- The main test for mutagenicity measures back-mutation, in histidine-free medium, of a mutant *Salmonella typhimurium* (which, unlike the wild-type, cannot grow without histidine) in the presence of:
  - the chemical to be tested
  - a liver microsomal enzyme preparation for generating reactive metabolites.
- Colony growth indicates that mutagenesis has occurred. The test is rapid and inexpensive, but some false positives and false negatives occur.
- Carcinogenicity testing:
  - involves chronic dosing of groups of animals
  - is expensive and time-consuming
  - does not readily detect epigenetic carcinogens.

together with ultraviolet light, PUVA, in specialist skin disease centres for treatment of psoriasis) is both mutagenic and carcinogenic in animal models and may increase the incidence of skin cancer in humans.

TERATOGENESIS AND DRUG-INDUCED FETAL DAMAGE

Teratogenesis signifies the production of gross structural malformations during fetal development, in distinction from other kinds of drug-induced fetal damage such as growth retardation, dysplasia (e.g. iodide-associated goitre) or the asymmetrical limb reduction resulting from vasoconstriction caused by **cocaine** (see Ch. 48) in an otherwise normally developing limb. Examples of drugs that affect fetal development adversely are given in Table 57.2.

It has been known that external agents can affect fetal development since the 1920s, when it was discovered that X irradiation during pregnancy causes fetal malformation. The importance of rubella infection was recognised two decades later, but it was not until 1960 that drugs were implicated as causative agents in teratogenesis: the shocking experience with **thalidomide** led to a widespread reappraisal of many other drugs in clinical use, and to the setting up of drug regulatory bodies in many countries. Most birth defects (about 70%) occur with no recognisable causative factor. Drug or chemical exposure during

Table 57.2 Some drugs reported to have adverse effects on human fetal development

Agent	Effect(s)	Teratogenicity[a]	See Chapter
Thalidomide	Phocomelia, heart defects, gut atresia, etc.	K	This chapter
Penicillamine	Loose skin etc.	K	26
Warfarin	Saddle nose; retarded growth; defects of limbs, eyes, central nervous system	K	24
Corticosteroids	Cleft palate and congenital cataract—rare	–	32
Androgens	Masculinisation in female	–	34
Oestrogens	Testicular atrophy in male	–	34
Stilbestrol	Vaginal adenosis in female fetus, also vaginal or cervical cancer	20+ years later	34
Phenytoin	Cleft lip/palate, microcephaly, mental retardation	K	44
Valproate	Neural tube defects (e.g. spina bifida)	K	44
Carbamazepine	Retardation of fetal head growth	S	44
Cytotoxic drugs (especially folate antagonists)	Hydrocephalus, cleft palate, neural tube defects, etc.	K	55
Aminoglycosides	Deafness	–	50
Tetracycline	Staining of bones and teeth, thin tooth enamel, impaired bone growth	S	50
Ethanol	Fetal alcohol syndrome	K	48
Retinoids	Hydrocephalus etc.	K	56
Angiotensin-converting enzyme inhibitors	Oligohydramnios, renal failure	K	22

[a]K, known teratogen (in experimental animals and/or humans); S, suspected teratogen (in experimental animals and/or humans).
Adapted from Juchau 1989 Annu Rev Pharmacol Toxicol 29: 165.

Table 57.3 The nature of drug effects on fetal development

Stage	Gestation period in humans	Main cellular process(es)	Affected by
Blastocyst formation	0–16 days	Cell division	Cytotoxic drugs, ?alcohol
Organogenesis	17–60 days approximately	Division Migration Differentiation Death	Teratogens Teratogens Teratogens Teratogens
Histogenesis and functional maturation	60 days to term	As above	Miscellaneous drugs (e.g. alcohol, nicotine, antithyroid drugs, steroids)

pregnancy is estimated to account for only approximately 1% of all fetal malformations. Fetal malformations are common, so the absolute numbers of children affected are substantial.

MECHANISM OF TERATOGENESIS

The timing of the teratogenic insult in relation to fetal development is critical in determining the type and extent of damage. Mammalian fetal development passes through three phases (Table 57.3):

1. blastocyst formation
2. organogenesis
3. histogenesis and maturation of function.

Cell division is the main process occurring during blastocyst formation. During this phase, drugs can kill the embryo by inhibiting cell division, but provided the embryo survives, its subsequent development does not generally seem to be compromised. Ethanol is an exception, affecting development at this very early stage (Ch. 48).

Drugs can cause gross malformations if administered during organogenesis (days 17–60 in humans). The structural organisation of the embryo occurs in a well-defined sequence: eye and brain, skeleton and limbs, heart and major vessels, palate, genitourinary system. The type of malformation produced thus depends on the time of exposure to the teratogen.

The cellular mechanisms by which teratogenic substances produce their effects are not at all well understood. There is a considerable overlap between mutagenicity and teratogenicity. In one large survey, among 78 compounds, 34 were both teratogenic and mutagenic, 19 were negative in both tests and 25 (among them thalidomide) were positive in one but not the other. Damage to DNA is important but, as with carcinogenesis, is not the only factor. The control of morphogenesis is poorly understood; vitamin A derivatives (retinoids) are involved and are potent teratogens (see below). Known teratogens also include several drugs (e.g. **methotrexate** and **phenytoin**) that do not react directly with DNA but which inhibit its synthesis by their effects on folate metabolism (see Ch. 25). Administration of **folate** during pregnancy reduces the frequency of both spontaneous and drug-induced malformations, especially neural tube defects.

The fetus depends on an adequate supply of nutrients during the final stage of histogenesis and functional maturation, and development is regulated by a variety of hormones. Gross structural malformations do not arise from exposure to mutagens at this stage, but drugs that interfere with the supply of nutrients or with the hormonal milieu may have deleterious effects on growth and development. Exposure of a female fetus to androgens at this stage can cause masculinisation. **Stilbestrol** was commonly given to pregnant women with a history of recurrent miscarriage during the 1950s (for unsound reasons) and causes dysplasia of the vagina of the infant and an increased incidence of carcinoma of the vagina in the teens and twenties. Angiotensin II plays an important part in the later stages of fetal development and in renal function in the fetus, and ACE inhibitors and angiotensin receptor antagonists ('sartans') cause oligohydramnios and renal failure if administered during later stages of pregnancy and fetal malformations if given earlier.

TESTING FOR TERATOGENICITY

The thalidomide disaster dramatically brought home the need for routine teratogenicity studies on new therapeutic drugs. Assessment of teratogenicity in humans is a particularly difficult problem for various reasons. One is that the 'spontaneous' malformation rate is high (3–10% depending on the definition of a significant malformation) and highly variable between different regions, age groups and social classes. Large-scale studies are required, which take many years and much money to perform, and they usually give suggestive, rather than conclusive, results.

▼ Studies using embryonic stem cells in assessing developmental toxicity are showing some promise (see Bremer & Hartung, 2004, for a review from a regulatory perspective). In vitro methods, based on the culture of cells, organs or whole embryos, have, however, not so far been developed to a level where they satisfactorily predict teratogenesis in vivo, and most regulatory authorities require teratogenicity testing in a rodent plus in one non-rodent species (e.g. rabbit). The visceral yolk sac and development of the chorioallantoic placenta of the rabbit resemble those of humans more so than do those of rodents, in some respects (Foote & Carney, 2000). Pregnant females are dosed at various levels during the critical period of organogenesis, and the fetuses are examined for structural abnormalities. However, poor cross-species correlation means that tests of this kind are not reliably predictive in humans, and it is usually recommended that new drugs are not used in pregnancy unless it is essential.

SOME DEFINITE AND PROBABLE HUMAN TERATOGENS

Although many drugs have been found to be teratogenic in varying degrees in experimental animals, relatively few are known to be teratogenic in humans (see Table 57.2). Some of the more important are discussed below.

Thalidomide

Thalidomide is virtually unique in producing, at therapeutic dosage, virtually 100% malformed infants when taken in the first 3–6 weeks of gestation. It was introduced in 1957 as a hypnotic and sedative with the special feature that it was much less hazardous in overdosage than barbiturates, and it was even recommended specifically for use in pregnancy (with the advertising slogan 'the safe hypnotic'). It had been subjected to toxicity testing only in mice, which are resistant to thalidomide teratogenicity (probably because mouse embryonic cells have higher glutathione levels than humans; Knobloch et al., 2008). Thalidomide was marketed energetically and successfully, and the first suspicion of its teratogenicity arose early in 1961 with reports of a sudden increase in the incidence of phocomelia. This abnormality ('seal limbs') consists of an absence of development of the long bones of the arms and legs, and had hitherto been virtually unknown. At this time, approximately 1 000 000 tablets were being sold daily in West Germany. Reports of phocomelia came simultaneously from Hamburg and Sydney, and the connection with thalidomide was made.[6] The drug was withdrawn late in 1961, by which time an estimated 10 000 malformed babies had been born (Fig. 57.3 illustrates the use of data linkage in detecting delayed ADRs). Despite intensive study, its mechanism remains poorly understood, although epidemiological investigation showed very clearly the correlation between the time of exposure and the type of malfunction produced (Table 57.4).

Cytotoxic drugs

Many alkylating agents (e.g. **chlorambucil** and **cyclophosphamide**) and antimetabolites (e.g. **azathioprine** and **mercaptopurine**) cause malformations when used in early pregnancy but more often lead to abortion (see Ch. 55). Folate antagonists (e.g. **methotrexate**) produce a much higher incidence of major malformations, evident in both live-born and stillborn fetuses.

Retinoids

Etretinate, a retinoid (i.e. vitamin A derivative) with marked effects on epidermal differentiation, is a known teratogen and causes a high proportion of serious abnormalities (notably skeletal deformities) in exposed fetuses. Dermatologists use retinoids to treat skin diseases including several, such as acne and psoriasis, that are common in young women. Etretinate accumulates in subcutaneous fat and is eliminated extremely slowly, detectable amounts persisting for many months after chronic dosing is discontinued. Because of this, women should avoid pregnancy for at least 2 years after treatment. **Acitretin** is an active metabolite of etretinate. It is equally teratogenic, but tissue accumulation is less pronounced and elimination may be more rapid.

[6]A severe peripheral neuropathy, leading to irreversible paralysis and sensory loss, was reported within a year of the drug's introduction and subsequently confirmed in many reports. The drug company responsible was less than punctilious in acting on these reports (see Sjöström & Nilsson, 1972), which were soon eclipsed by the discovery of teratogenic effects, but the neurotoxic effect was severe enough in its own right to have necessitated withdrawal of the drug from general use. Today, use of thalidomide has had a resurgence related to several highly specialised applications. It is prescribed by specialists (in dermatology, oncology and in HIV infection, among others) under tightly controlled and restricted conditions.

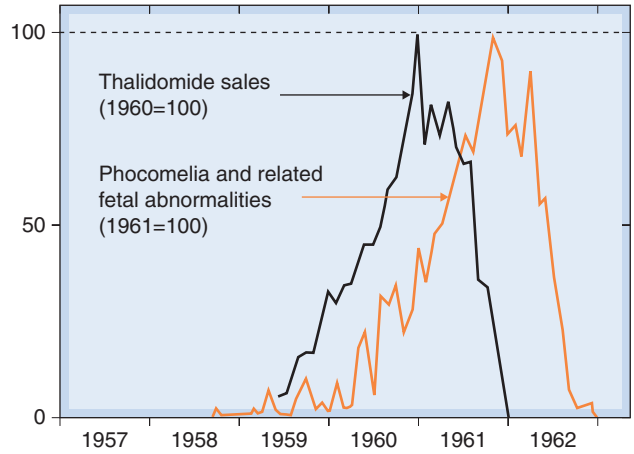

Fig. 57.3 Incidence of major fetal abnormalities in Western Europe following the introduction and withdrawal of thalidomide linked to sales data for thalidomide.

Table 57.4 Thalidomide teratogenesis

Day of gestation	Type of deformity
21–22	Malformation of ears Cranial nerve defects
24–27	Phocomelia of arms
28–29	Phocomelia of arms and legs
30–36	Malformation of hands Anorectal stenosis

Heavy metals

Lead, *cadmium* and *mercury* all cause fetal malformation in humans. The main evidence comes from Minamata disease, named after the locality in Japan where an epidemic occurred when the local population ate fish contaminated with methylmercury that had been used as an agricultural fungicide. This impaired brain development in exposed fetuses, resulting in cerebral palsy and mental retardation, often with microcephaly. Mercury, like other heavy metals, inactivates many enzymes by forming covalent bonds with sulfhydryl and other groups, and this is believed to be responsible for these developmental abnormalities.

Antiepileptic drugs

Congenital malformations are increased two- to three-fold in babies of epileptic mothers. Interestingly, all existing antiepileptic drugs have been implicated, including **phenytoin** (particularly cleft lip/palate), **valproate** (neural tube defects) and **carbamazepine** (spina bifida and hypospadias, a malformation of the male urethra), as well as newer agents including **lamotrigine** (Ch. 44).

Warfarin

Administration of **warfarin** (Ch. 24) in the first trimester is associated with nasal hypoplasia and various central nervous system abnormalities, affecting roughly 25% of exposed babies. In the last trimester, it must not be used because of the risk of intracranial haemorrhage in the baby during delivery.

Teratogenesis and drug-induced fetal damage

- Teratogenesis means production of gross structural malformations of the fetus (e.g. the absence of limbs after thalidomide). Less comprehensive damage can be produced by several drugs (see Table 57.2). Less than 1% of congenital fetal defects are attributed to drugs given to the mother.
- Gross malformations are produced only if teratogens act during organogenesis. This occurs during the first 3 months of pregnancy but after blastocyst formation. Drug-induced fetal damage is rare during blastocyst formation (exception: fetal alcohol syndrome) and after the first 3 months (exception: angiotensin-converting enzyme inhibitors and sartans).
- The mechanisms of action of teratogens are not clearly understood, although DNA damage is a factor.
- New drugs are usually tested in pregnant females of at least one rodent and one non-rodent (e.g. rabbit) species.

ASSESSMENT OF GENOTOXIC POTENTIAL

Registration of pharmaceuticals requires a comprehensive assessment of their genotoxic potential. Because no single test is adequate, the usual approach recommended by the International Conference on Harmonisation (ESRA Rapporteur 1997 4: 5–7) is to carry out a battery of in vitro and in vivo tests for genotoxicity. The following battery is often used:

- a test for gene mutation in bacteria
- an in vitro test with cytogenetic evaluation of chromosomal damage
- an in vivo test for chromosomal damage using rodent haemopoietic cells
- reproductive toxicity testing
- carcinogenicity testing.

ALLERGIC REACTIONS TO DRUGS

Allergic reactions of various kinds are a common form of adverse response to drugs. Most drugs, being low-molecular-weight substances, are not immunogenic in themselves. A drug or its metabolites can, however, act as a *hapten* by interacting with protein to form a stable conjugate that is immunogenic (Ch. 6). The immunological basis of some allergic drug reactions has been well worked out, but often it is inferred from the clinical characteristics of the reaction, and direct evidence of an immunological mechanism is lacking. Suggestive features are as follow:

- The time course differs from the main action of the drug; it is either delayed in onset or occurs only with repeated exposure to the drug.
- Allergy may result from doses that are too small to elicit pharmacodynamic effects.
- The reaction conforms to one of the clinical syndromes associated with allergy—types I, II, III and IV of the Gell and Coombs classification (see below and Ch. 6)—and is unrelated to the main action of the drug.

The overall incidence of allergic drug reactions is variously reported as being between 2% and 25%. Most are minor skin eruptions. Serious reactions (e.g. anaphylaxis, haemolysis and bone marrow depression) are rare. Penicillins, which are the commonest cause of drug-induced anaphylaxis, produce this response in an estimated 1 in 50 000 patients exposed. Rashes can be severe, and fatalities occur with Stevens–Johnson syndrome (provoked, for example, by sulfonamides) and toxic epidermal necrolysis (TEN, which can be caused by **allopurinol**). The association between **cabamazepine**-induced TEN and the gene for a particular HLA allele *HLAB*1502* in people of Asian ancestry is mentioned in Chapter 11. Susceptibility to severe rashes in response to **abacavir** is closely linked to the human leukocyte antigen (HLA) variant *HLAB*5701* and this forms the basis of a clinically useful genomic test (Ch. 11).

IMMUNOLOGICAL MECHANISMS

The formation of an immunogenic conjugate between a small molecule and an endogenous protein requires covalent bonding. In most cases, reactive metabolites, rather than the drug itself, are responsible. Such reactive metabolites can be produced during drug oxidation or by photoactivation in the skin. They may also be produced by the action of toxic oxygen metabolites generated by activated leukocytes. Rarely (e.g. in drug-induced lupus erythematosus), the reactive moiety interacts to form an immunogen with nuclear components (DNA, histone) rather than proteins (see below). Conjugation with a macromolecule is usually essential, although penicillin is an exception because it can form sufficiently large polymers in solution to elicit an anaphylactic reaction in a sensitised individual even without conjugation to protein, although penicillin–protein conjugates can also act as the immunogen.

CLINICAL TYPES OF ALLERGIC RESPONSE TO DRUGS

In the Gell and Coombs classification of hypersensitivity reactions (Ch. 6), types I, II and III are antibody-mediated reactions and type IV is cell mediated. Unwanted reactions to drugs involve both antibody- and cell-mediated reactions. The more important clinical manifestations of hypersensitivity include anaphylactic shock, haematological reactions, allergic liver damage and other hypersensitivity reactions.

ANAPHYLACTIC SHOCK

Anaphylactic shock—see also Chapter 27—is a type I hypersensitivity response. It is a sudden and life-threatening reaction that results from the release of histamine, leukotrienes and other mediators. The main features include urticarial rash, swelling of soft tissues, bronchoconstriction and hypotension.

Penicillins account for about 75% of anaphylactic deaths, reflecting the frequency with which they are used in clinical practice. Other drugs that can cause anaphylaxis include various enzymes, for example **asparaginase** (Ch. 55); therapeutic monoclonal antibodies (Ch. 59), hormones, for example **corticotropin** (adrenocorticotrophic hormone; Ch. 32); **heparin** (Ch. 24); dextrans; radiological contrast agents; vaccines; and other serological products. Anaphylaxis with local anaesthetics (Ch. 42), the antiseptic chlorhexidine and

with many other drugs (sometimes as a consequence of contaminants such as latex used to seal reusable vials or of excipients and colouring agents rather than the drug itself) can occur. Treatment of anaphylaxis is given in Chapter 27.

It is sometimes feasible to carry out a skin test for the presence of anaphylactic hypersensitivity, which involves injecting a minute dose intradermally. A patient who reports that she or he is allergic to a drug such as penicillin may actually be allergic to fungal contaminants in early preparations rather than to penicillin itself. The use of penicilloylpolylysine as a skin test reagent for penicillin allergy is an improvement over the use of penicillin itself, because it bypasses the need for conjugation of the test substance, thereby reducing the likelihood of a false negative. Other specialised tests are available to detect the presence of specific immunoglobulin E in the plasma, or to measure histamine release from the patient's basophils, but these are not used routinely.

Other drug-induced type I hypersensitivity reactions include bronchospasm (Ch. 27) and urticaria.

HAEMATOLOGICAL REACTIONS

Drug-induced haematological reactions can be produced by type II, III or IV hypersensitivity. Type II reactions can affect any or all of the formed elements of the blood, which may be destroyed by effects either on the circulating blood cells themselves or on their progenitors in the bone marrow. They involve antibody binding to a drug–macromolecule complex on the cell surface membrane. The antigen–antibody reaction activates complement, leading to lysis, or provokes attack by killer lymphocytes or phagocytic leukocytes (Ch. 6). *Haemolytic anaemia* has been most commonly reported with sulfonamides and related drugs (Ch. 50) and with an antihypertensive drug, **methyldopa** (Ch. 14), which is still widely used to treat hypertension during pregnancy. With methyldopa, significant haemolysis occurs in less than 1% of patients, but the appearance of antibodies directed against the surface of red cells is detectable in 15% by the Coombs test. The antibodies are directed against Rh antigens, but it is not known how methyldopa produces this effect.

Drug-induced *agranulocytosis* (complete absence of circulating neutrophils) is usually delayed 2–12 weeks after beginning drug treatment but may then be sudden in onset. It often presents with mouth ulcers, a severe sore throat or other infection. Serum from the patient lyses leukocytes from other individuals, and circulating antileukocyte antibodies can usually be detected immunologically. Drugs associated with agranulocytosis include NSAIDs, especially **phenylbutazone** (Ch. 26), **carbimazole** (Ch. 33) and **clozapine** (Ch. 45) (increased genetic susceptibility associated with *HLA-DQB1*0201* is mentioned in Ch. 11) and **sulfonamides** and related drugs (e.g. *thiazides* and *sulfonylureas*). Agranulocytosis is rare but life-threatening. Recovery when the offending drug is stopped is often slow or absent. Antibody-mediated leukocyte destruction must be distinguished from the direct effect of cytotoxic drugs (see Ch. 55), which cause granulocytopenia that is rapid in onset, predictably related to dose and reversible.

Thrombocytopenia (reduction in platelet numbers) can be caused by type II reactions to **quinine** (Ch. 53), **heparin** (Ch. 24) and thiazide diuretics (Ch. 28).

Some drugs (notably **chloramphenicol**) can suppress all three haemopoietic cell lineages, giving rise to *aplastic anaemia* (anaemia with associated agranulocytosis and thrombocytopenia).

The distinction between type III and type IV hypersensitivity reactions in the causation of haematological reactions is not clear-cut, and either or both mechanisms can be involved.

ALLERGIC LIVER DAMAGE

Most drug-induced liver damage results from the direct toxic effects of drugs or their metabolites, as described above. However, hypersensitivity reactions are sometimes involved, a particular example being **halothane**-induced hepatic necrosis (see Ch. 40). *Trifluoracetylchloride*, a reactive metabolite of halothane, couples to a macromolecule to form an immunogen. Most patients with halothane-induced liver damage have antibodies that react with halothane–carrier conjugates. Halothane–protein antigens can be expressed on the surface of hepatocytes. Destruction of the cells occurs by type II hypersensitivity reactions involving killer T cells, and type III reactions can also contribute.

OTHER HYPERSENSITIVITY REACTIONS

The clinical manifestations of type IV hypersensitivity reactions are diverse, ranging from minor skin rashes to generalised autoimmune disease. Fever may accompany these reactions. Rashes can be antibody mediated but are usually cell mediated. They range from mild eruptions to fatal exfoliation. Stevens–Johnson syndrome is a very severe generalised rash that extends into the alimentary tract and carries an appreciable mortality. In some cases, the lesions are photosensitive, probably because ultraviolet light converts the drug to reactive products.

▼ Some drugs (notably **hydralazine** and **procainamide**) can produce an autoimmune syndrome resembling systemic lupus

Allergic reactions to drugs

- Drugs or their reactive metabolites can bind covalently to proteins to form immunogens. Penicillin (which can also form immunogenic polymers) is an important example.
- Drug-induced allergic (hypersensitivity) reactions may be antibody mediated (types I, II, III) or cell mediated (type IV). Important clinical manifestations include the following:
  - anaphylactic shock (type I): many drugs can cause this, and most deaths are caused by **penicillin**
  - haematological reactions (type II, III or IV): including haemolytic anaemia (e.g. **methyldopa**), agranulocytosis (e.g. **carbimazole**), thrombocytopenia (e.g. **quinine**) and aplastic anaemia (e.g. **chloramphenicol**)
  - hepatitis (types II, III): for example, **halothane**, **phenytoin**
  - rashes (type I, IV): are usually mild but can be life-threatening (e.g. Stevens–Johnson syndrome)
  - drug-induced systemic lupus erythematosus (mainly type II): antibodies to nuclear material are formed (e.g. **hydralazine**).

erythematosus. This is a multisystem disorder in which there is immunological damage to many organs and tissues (including joints, skin, lung, central nervous system and kidney) caused particularly, but not exclusively, by type III hypersensitivity reactions. The prodigious array of antibodies directed against 'self' components has been termed an 'autoimmune thunderstorm'. The antibodies react with determinants shared by many molecules, for example the phosphodiester backbone of DNA, RNA and phospholipids. In drug-induced systemic lupus erythematosus, the immunogen may result from the reactive drug moiety interacting with nuclear material, and joint and pulmonary damage is common. The condition usually resolves when treatment with the offending drug is stopped.

REFERENCES AND FURTHER READING

Adverse drug reactions

Aronson, J.K., Ferner, R.E., 2003. Joining the DoTS: a new approach to classifying adverse drug reactions. Br. Med. J. 327, 1222–1225. (*Description of ADRs in terms of dose, time course and susceptibility*)

Pirmohamed, M., James, S., Meakin, S. et al., 2004. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. Br. Med. J. 329, 15–19. (*There were 1225 admissions related to an adverse drug reaction. The median bed stay was 8 days, accounting for 4% of the hospital bed capacity. The projected annual cost is £466 million. Most reactions were avoidable. Drugs most commonly implicated were aspirin and other NSAIDs, diuretics, warfarin; the most common reaction was gastrointestinal bleeding*)

Rawlins, M.D., Thomson, J.W., 1985. Mechanisms of adverse drug reactions. In: Davies, D.M. (Ed.), Textbook of adverse drug reactions, third ed. Oxford University Press, Oxford, pp. 12–38. (*Type A/type B classification*)

Drug toxicity: general and mechanistic aspects

Bhogal, N., Grindon, C., Combes, R., Balls, M., 2005. Toxicity testing: creating a revolution based on new technologies. Trends Biotechnol. 23, 299–307. (*Reviews current and likely future value of new technologies in relation to toxicological evaluation*)

Bremer, S., Hartung, T., 2004. The use of embryonic stem cells for regulatory developmental toxicity testing in vitro—the current status of test development. Curr. Pharm. Des. 10, 2733–2747. (*Summarises requirements for an in vitro embryotoxicity test needed for regulatory toxicity testing*)

Parng, C., 2005. In vivo zebrafish assays for toxicity testing. Curr. Opin. Drug Discov. Devel. 8, 100–106. (*Effective in vivo toxicity screening early in development can reduce the number of compounds that progress to laborious and costly late-stage animal testing. The transparent zebra fish provides accessibility to internal organs, tissues and even cells, and has emerged as a model organism for toxicity testing. Straightforward in vivo zebra fish assays can serve as an intermediate step between cell-based and mammalian testing, thus streamlining the drug development timeline*)

Pirmohamed, M., 2003. Drug-induced apoptosis: clinical significance. Drug Metab. Rev. 35, 24–24; 48 (Suppl. 1).

Pirmohamed, M., 2004. Role of the immune system in idiosyncratic drug reactions. Drug Metab. Rev. 36, 29–29; 58 (Suppl. 1).

Timbrell, J.A., 2009. Principles of biochemical toxicity. Informa Healthcare, New York.

Uetrecht, J., 2003. Screening for the potential of a drug candidate to cause idiosyncratic drug reactions. Drug Discov. Today 8, 832–837. (*Highlights current mechanistic hypotheses of idiosyncratic drug reactions and discusses future directions in the development of better predictive tests*)

Walker, D.K., 2004. The use of pharmacokinetic and pharmacodynamic data in the assessment of drug safety in early drug development. Br. J. Clin. Pharm. 58, 601–608. (*Pharmacokinetic profile is a factor in assessing safety during early drug development, especially in relation to safety parameters such as QT interval prolongation, where free plasma concentrations are predictive; procedures are available that allow this on the microdose scale—potential limitations are discussed*)

Drug toxicity: carcinogenesis, teratogenesis

Briggs, G.G., Freeman, R.K., Yaffe, S.J., 2008. Drugs in pregnancy and lactation, eighth ed. Lippincott, Williams & Wilkins, Philadelphia. (*Invaluable reference guide to fetal and neonatal risk for clinicians caring for pregnant women*)

Collins, M.D., Mayo, G.E., 1999. Teratology of retinoids. Annu. Rev. Pharmacol. Toxicol. 39, 399–430. (*Overviews principles of teratology as they apply to the retinoids, describes signal transduction of retinoids and toxikinetics*)

Foote, R.H., Carney, E.W., 2000. The rabbit as a model for reproductive and developmental toxicity studies. Reprod. Toxicol. 14, 477–493. (*Discusses the use of the rabbit in developmental toxicity and teratology studies*)

Knobloch, J., Reimann, K., Klotz, L.-O., Rüther, U., 2008. Thalidomide resistance is based on the capacity of the glutathione-dependent antioxidant defense. Mol. Pharm. 5, 1138–1144.

Sjöström, H., Nilsson, R., 1972. Thalidomide and the power of the drug companies. Penguin Books, London.

Drug toxicity: organ involvement

Murray, M.C., Brater, D.C., 1993. Renal toxicity of the nonsteroidal anti-inflammatory drugs. Annu. Rev. Pharmacol. Toxicol. 33, 435–465.

Park, B.K., Kitteringham, N.R., Maggs, J.L., et al., 2005. The role of metabolic activation in drug-induced hepatotoxicity. Annu. Rev. Pharmacol. Toxicol. 45, 177–202. (*Reviews evidence for reactive metabolite formation from hepatotoxic drugs such as paracetamol, tamoxifen, diclofenac and troglitazone, and the current hypotheses of how this leads to liver injury*)

Ritter, J.M., Harding, I., Warren, J.B., 2009. Precaution, cyclooxygenase inhibition, and cardiovascular risk. Trends Pharmacol. Sci. 30, 503–514.

Svensson, C.K., Cowen, E.W., Gaspari, A.A., 2001. Cutaneous drug reactions. Pharmacol. Rev. 53, 357–380. (*Covers epidemiology, clinical morphology and mechanisms. Assesses current knowledge of four types of cutaneous drug reaction: immediate-type immune mediated, delayed-type immune mediated, photosensitivity and autoimmune. Also reviews the role of viral infection as predisposing factor*)

Valentin, J.-P., 2010. Reducing QT liability and proarrhythmic risk in drug discovery and development. Br. J. Pharmacol. 159, 5–11. (*See also accompanying articles in this themed section on QT safety*)

58 Lifestyle drugs and drugs in sport

OVERVIEW

The term *lifestyle drugs* refers to an eclectic group of drugs that are used for non-medical purposes. It includes drugs of abuse, drugs used to enhance athletic or other performance, as well as those taken for cosmetic purposes or for purely social reasons. Alcohol, nicotine and various abused drugs are covered in Chapter 48. Many lifestyle drugs are also used as conventional therapeutics and the pharmacology is dealt with elsewhere in this book. In this chapter, we present an overall summary of the classes of drugs that are used for non-medical purposes, and discuss some of the social and medicolegal problems associated with their growing use. Drugs, officially prohibited, that are used to enhance sporting performance represent a special category of lifestyle drugs. Once again, a wide range of agents are used for this purpose, and their pharmacological properties are described in other chapters. Here we discuss specific issues relating to their use in competitive sports.

WHAT IS A LIFESTYLE DRUG?

Lifestyle drugs is a term of fairly recent origin and not precisely defined. The most commonly accepted definition refers to a drug or medicine[1] that is used to satisfy an aspiration or a non-health-related goal. Examples include the use of the antihypertensive **minoxidil** for treating baldness or **sildenafil** for erectile difficulties in the absence of underlying disease. Oral contraceptives, which clearly lie in the domain of mainstream medicine, should also be considered lifestyle drugs. The term is sometimes also used to describe medicines that are used to treat 'lifestyle illnesses', that is to say diseases that arise through 'lifestyle choices' such as smoking, alcoholism or overeating, and there are many other shades of meaning as well. Also included in the lifestyle category are food supplements and other related preparations that are taken by the general public from choice because of some claimed benefit, even when there is no good evidence that they are effective.

CLASSIFICATION OF LIFESTYLE DRUGS

To classify all the different drugs or medicines that might fall into the lifestyle category, and provide a standard universally acceptable definition, is therefore difficult, and cuts across the pharmacological classification used throughout this book. The classification scheme summarised in Table 58.1 is based largely on the work of Gilbert et al. (2000) and Young (2003). This scheme embraces drugs that have been used for lifestyle choices based on historical precedent, such as oral contraceptives, as well as agents used to manage potentially debilitating lifestyle illnesses such as addiction to smoking (e.g. **bupropion**). It also includes drugs such as **caffeine** and **alcohol** that are consumed on a mass scale around the world, and drugs of abuse such as **cocaine** as well as nutritional supplements. Particularly controversial are drugs aimed at improving intellectual performance, such as **modafinil** and **methylphenidate** (Ch. 47), that are gaining favour as a route to academic success.[2]

Drugs can, over time, switch from 'lifestyle' to 'mainstream' use. For example, **atropine** (Ch.13) was first used as a beauty aid based on its ability to dilate the pupil. Cocaine was first described as a lifestyle drug in use by the South American Indians. Early explorers commented that it 'satisfies the hungry, gives new strength to the weary and exhausted and makes the unhappy forget their sorrows'. Subsequently assimilated into European medicine as a local anaesthetic (Ch. 42), it is now largely returned to lifestyle drug status and, regrettably, is the basis of an illegal multimillion dollar international drugs industry. **Cannabis** is another good example of a drug that has been considered (in the West at least) as a purely recreational drug but which is now (as **tetrahydrocannabinol**) in trial for various clinical uses (see Chs 18 and 48).

Many widely used lifestyle 'drugs' or 'sports supplements' consist of natural products (e.g. ginkgo extracts, melatonin, St John's wort, cinchona extracts), whose manufacture and sale has not generally been controlled by regulatory bodies.[3] Their composition is therefore highly variable, and their efficacy and safety generally untested. Many contain active substances that, like synthetic drugs, can produce adverse as well as beneficial effects.

DRUGS IN SPORT

The use of drugs to enhance sporting performance[4] is evidently widespread although officially prohibited. The World Anti-Doping Agency (http://www.wada-ama.org), which was established partly in response to some high-profile doping cases and drug-induced deaths among

[1]We use the terms 'drug' and 'medicine' interchangeably in this chapter for discussion and classification purposes.

[2]Drugs intended to give a competitive advantage in sport are, of course, considered unfair, banned and very actively policed. Will there come a time when taking drugs to improve examination performance will become illegal, with similar surveillance methods and sanctions? See Bostrom & Sandberg (2009) for a discussion of this ethical minefield.

[3]In truth, it would be more accurate to refer to *lifestyle uses* for drugs and medicines rather than categorise these agents separately. Things are changing. In the UK, the Medicines and Healthcare Products Regulatory Agency now has a Herbal Medicines Advisory Committee.

[4]Some of these drugs are used 'legitimately' by the military to improve battlefield effectiveness.

Table 58.1 Lifestyle drugs and medicines, excluding drugs in sport

Category	Example(s)	Primary clinical use	'Lifestyle' use	See Chapter
Medicines approved for specific indications but that can also be used for other 'lifestyle' purposes	Sildenafil	Erectile dysfunction	Erectile enhancement	34
	Oral contraceptives	Preventing conception	Preventing conception	34
	Orlistat	Obesity	Weight loss	31
	Sibutramine	Anorectic agent	Weight loss	31
	Bupropion	Managing nicotine addiction	Managing nicotine addiction	43
	Methadone	Managing opiate addiction	Managing opiate addiction	41
Medicines approved for specific indications that can also be used to satisfy 'lifestyle choices' or to treat 'lifestyle diseases'	Minoxidil	Hypertension	Regrowth of hair	22
	Methylphenidate	Attention-deficit hyperactivity disorder (ADHD)	Improving academic performance	47
	Modafinil	Treatment of ADHD	Cognitive enhancement	47
	Opiates	Analgesia	'Recreational' usage	41
Drugs that have slight or no current clinical use but which fall into the lifestyle category	Alcohol	None as such	Widespread component of drinks	48
	Botulinum toxin	Relief of muscle spasm	Cosmetic alteration	13
	Caffeine	Migraine treatment	Widespread component of drinks	47
	Cannabis	Managing chronic pain, nausea and possibly muscle spasm	'Recreational' usage	18, 48
Drugs (generally illegal) that have no clinical utility but which are used to satisfy lifestyle requirements	Methylenedioxymethamphetamine (MDMA, 'ecstasy')	None as such	'Recreational' usage	47
	Tobacco (nicotine)	Patches for tobacco addiction	'Recreational' usage	48
	Cocaine (some formulations)	Local anaesthesia (largely obsolete)	'Recreational' usage	42
Natural products, largely unregulated with claimed (often anecdotal and unsubstantiated) effects but which cater to lifestyle needs or desires	Fish oils	Slight Nutritional supplement	Widespread, for many conditions	—
	Ascorbic acid	Slight Nutritional supplement	Widespread, for many conditions	—
	Melatonin	None	Widespread, for many conditions	—
	Numerous herbal preparations	None	Widespread, for many conditions	—

(After Gilbert et al., 2000 and Young, 2003).

athletes, publishes an annually updated list of prohibited substances that may not be used by sportsmen or sportswomen either in or out of competition. Drug testing is based mainly on analysis of blood or urine samples according to strictly defined protocols. The chemical analyses, which rely mainly on gas chromatography/mass spectrometry or immunoassay techniques, must be carried out by approved laboratories.

Table 58.2 summarises the main classes of drugs that are prohibited for use in sports. Athletes are easily persuaded of the potential of a wide variety of drugs to increase their chances of winning, but it should be emphasised that in very few cases have controlled trials shown that the drugs actually improve sporting performance among trained athletes, and indeed many such trials have proved negative. However, marginal improvements in performance (often 1% or less), which are difficult to measure experimentally, may make the difference between winning and losing, and the competitive instincts of athletes and their trainers generally carry more weight than scientific evidence.

A brief account of some of the more important drugs in common use follows. For a broader and more complete coverage, see British Medical Association (2002) and Mottram (2005).

ANABOLIC STEROIDS

Anabolic steroids (Ch. 34) include a large group of compounds with testosterone-like effects, including about 50 named compounds on the prohibited list. New chemical derivatives ('designer steroids') such as **tetrahydrogestrinone** (THG) are regularly developed and offered illicitly to

Table 58.2 Drugs used in sport

Drug class	Example(s)	Effects	Detection	Notes
Anabolic agents	Androgenic steroids (testosterone, nandrolone and many others; Ch. 34)	Increased muscle development Increased aggression and competitiveness Serious long-term side effects (see text)	Urine or blood samples	Many are endogenous hormones, so results significantly above normal range are required Human chorionic gonadotrophin is sometimes used to increase androgen secretion
	Clenbuterol (Ch. 14)	Combined anabolic and agonist action on β_2 adrenoceptors may increase muscle strength		
Hormones and related substances	Erythropoietin (Ch. 25)	Increased erythrocyte formation and hence increased oxygen transport Increased blood viscosity causes hypertension and risk of strokes and coronary attacks Used mainly for endurance sports[a]	Plasma half-life is short, so detection is difficult	Use of other plasma markers indicating erythropoietin administration may be possible
	Human growth hormone (Ch. 32)	Increased lean body mass and reduced fat May accelerate recovery from tissue injury Causes cardiac hypertrophy, acromegaly, liver damage and increased cancer risk	Blood testing	Distinguishing endogenous (highly variable) from exogenous human growth hormone is difficult
	Insulin (Ch. 30)	Sometimes used (with glucose so as to avoid hypoglycaemia) to promote glucose uptake and energy production in muscle Probably ineffective in improving performance	Plasma samples	—
β_2-Adrenoceptor agonists	Salbutamol and others (Ch. 14)	Used by runners, cyclists, swimmers, etc. to increase oxygen uptake (by bronchodilatation) and cardiac function Controlled studies show no improvement in performance	Urine samples	—
β-Adrenoceptor antagonists	Propranolol, etc. (Ch. 14)	Used to reduce tremor and anxiety in 'precision' sports (e.g. shooting, gymnastics, diving)	Urine samples	Not banned in most sports where they actually impair performance
'Stimulants'	Ephedrine and derivatives Amphetamines, cocaine caffeine (Ch. 47)	Many trials show slight increase in muscle strength and performance in non-endurance events (sprint, swimming, field events, etc.)	Urine samples	The most widely used group, along with anabolic steroids
Diuretics	Thiazides, furosemide (Ch. 28)	Used mainly to achieve rapid weight loss before weighing in Also to mask presence of other agents in urine by dilution	Urine samples	—
Narcotic analgesics	Codeine, morphine, etc. (Ch. 41)	Used to mask injury-associated pain	Urine samples	—

[a] 'Blood doping' (removal of 1–2 l of blood in advance, followed by retransfusion immediately before competition) has a similar effect and is even more difficult to detect.

Lifestyle drugs

- Comprise a group of drugs and medicines taken mainly for non-medical reasons. Should more accurately be called 'lifestyle uses'.
- Include prescription drugs such as **sildenafil** and **methylphenidate**, substances such as **alcohol** and **caffeine**, drugs of abuse and various nutritional preparations.
- Are linked to the concepts of 'self-diagnosis' and 'non-disease'.
- Are a growing sector of the pharmaceutical market.
- Are often brought to the consumer's attention through the Internet or direct marketing of drugs.

athletes, which represents a continuing problem to the authorities charged with detecting and identifying them. A further problem is that some of these drugs are endogenous compounds or their metabolites, making it difficult to prove that the substance had been administered illegally. Isotope ratio techniques, based on the fact that endogenous and exogenous steroids have slightly different ^{12}C:^{13}C ratios, may enable the two to be distinguished analytically.

Anabolic steroids produce long-term effects and are normally used throughout training, rather than during competition, so out-of-competition testing is necessary.

Although anabolic steroids, when given in combination with training and high protein intake, undoubtedly increase muscle mass and body weight, there is little evidence that they increase muscle strength over and above the effect of training, or that they improve sporting performance. On the other hand, they have serious long-term effects, including male infertility, female masculinisation, liver and kidney tumours, hypertension and increased cardiovascular risk, and in adolescents premature skeletal maturation causing irreversible cessation of growth. Anabolic steroids produce a feeling of physical well-being, increased competitiveness and aggressiveness, sometimes progressing to actual psychosis. Depression is common when the drugs are stopped, sometimes leading to long-term psychiatric problems.

Clenbuterol, a β-adrenoceptor antagonist (see Ch. 14), has recently come into use by athletes. Through an unknown mechanism of action, it produces anabolic effects similar to those of androgenic steroids, with apparently fewer adverse effects. It can be detected in urine and is banned for use in sport.

HUMAN GROWTH HORMONE

The use of **human growth hormone** (hGH; see Ch. 32) by athletes followed the availability of the recombinant form of hGH, used to treat endocrine disorders. It is given by injection and its effects appear to be similar to those of anabolic steroids. hGH is also reported to produce a similar feeling of well-being, although without the accompanying aggression and changes in sexual development and behaviour. It increases lean body mass and reduces body fat, but its effects on muscle strength and athletic performance are unclear. It is claimed to increase the rate of recovery from tissue injury, allowing more intensive training routines to be followed.

The main adverse effect of hGH is the development of acromegaly, causing overgrowth of the jaw and thickening of the fingers (Ch. 32), but it may also lead to cardiac hypertrophy and cardiomyopathy, and possibly also an increased cancer risk.

Detection of hGH administration is difficult because physiological secretion is pulsatile, so normal plasma concentrations vary widely. The plasma half-life is short (20–30 min), and only trace amounts are excreted in urine. However, secreted hGH consists of three isoforms varying in molecular weight, whereas recombinant hGH contains only one, so measuring the relative amounts of the isoforms can be used to detect the exogenous material.

Growth hormone acts partly by releasing insulin-like growth factor from the liver, and this hormone itself is coming into use by athletes.

Another hormone, erythropoietin, which increases erythrocyte production (see Ch. 25) is given by injection for days or weeks to increase the erythrocyte count and hence the O_2-carrying capacity of blood. The development of recombinant erythropoietin has made it widely available, and detection of its use is difficult. It carries a risk of hypertension, neurologic disease and thrombosis.

STIMULANT DRUGS

The main drugs of this type used by athletes and officially prohibited are **ephedrine** and **methylephedrine**; various amphetamines and similar drugs, such as **fenfluramine** and **methylphenidate**;[5] cocaine; and a variety of other CNS stimulants such as **nikethamide**, **amiphenazole** (no longer used clinically) and **strychnine** (see Ch. 47). Caffeine is also used: some commercially available 'energy drinks' contain taurine as well as caffeine. However, taurine is an agonist at glycine and extrasynaptic $GABA_A$ receptors (see Ch. 38). Its effects on the brain are therefore likely to be inhibitory rather than stimulatory. In this regard, taurine may be responsible for the post-energy-drink low that is experienced once the stimulatory effect of caffeine has worn off.

In contrast to steroids, some trials have shown these drugs to improve performance in events such as sprinting and weightlifting, and under experimental conditions they increase muscle strength and reduce muscle fatigue significantly. The psychological effect of stimulants is probably more relevant than their physiological effects. Surprisingly, caffeine appears to be more consistently effective in improving muscle performance than other more powerful stimulants.

Several deaths have occurred among athletes taking amphetamines and ephedrine-like drugs in endurance events. The main causes are coronary insufficiency, associated with hypertension; hyperthermia, associated with cutaneous vasoconstriction; and dehydration.

CONCLUSION

The recent lifestyle drugs debate is one aspect of the broader long-standing question of what actually constitutes 'disease' and how far medical science should go in attempting to alleviate human distress and dysfunction in the absence of pathology, or to satisfy the needs and aspirations of otherwise healthy individuals. Discussion of these issues is beyond the scope of this book but can

[5]Also used to improve academic performance!

Drugs in sport

- Many drugs of different types are commonly used by sportsmen and sportswomen with the aim of improving performance in competition.
- The main types used are:
  - *anabolic agents*, mainly androgenic steroids and **clenbuterol**
  - *hormones*, particularly erythropoietin and human growth hormone
  - *stimulants*, mainly **amphetamine** and **ephedrine** derivatives and **caffeine**
  - *β-adrenoceptor antagonists*, to reduce anxiety and tremor in 'accuracy' sports.
- The use of drugs in sport is officially prohibited—in most cases, in or out of competition.
- Detection depends mainly on analysis of the drug or its metabolites in urine or blood samples. Detection of abuse is difficult in the case of endogenous hormones such as erythropoietin, growth hormone and testosterone.
- Controlled trials have mostly shown that drugs produce little improvement in sporting performance. Anabolic agents increase body weight and muscle volume without clearly increasing strength. The effect of stimulants is often psychological rather than physiological.

be found in articles cited at the end of this chapter (see Smith, 2002).

There are several reasons why the lifestyle drug phenomenon—no matter how we choose to define it—is of increasing concern. The increasing availability drugs from Internet 'e-pharmacies', coupled with direct advertising by the pharmaceutical industry to the public that occurs in some countries, will ensure that demand is kept buoyant, and the pharmaceutical sector will undoubtedly develop more lifestyle agents. The lobbying power of patients advocating particular drugs, regardless of the potential costs or proven utility, is causing major problems for drug regulators and those who set healthcare priorities for state-funded systems of social medicine. The use of drugs that improve short-term memory to treat patients with dementia (Ch. 39) is generally seen as desirable (even though current drugs are only marginally effective). Extending the use of existing and future drugs to give healthy children and students a competitive advantage in tests is much more controversial. Further off is the prospect of drugs that retard senescence and prolong life—another social and ethical minefield in an overpopulated world.

From a pharmacological perspective, it is fair to say that the use of drugs to enhance sporting performance carries many risks and is of very doubtful efficacy. Its growing prevalence reflects many of the same pressures as those driving the introduction of lifestyle drugs, namely the desire to improve on human attributes that are not impaired by disease, coupled with disregard for scientific evidence relating to efficacy and risk.

REFERENCES AND FURTHER READING

Lifestyle drugs and medicines, and general reading

Atkinson, T., 2002. Lifestyle drug market booming. Nat. Med. 8, 909. (*Interesting comments on the financial value of this sector*)

Bostrom, N., Sandberg, A., 2009. Cognitive enhancement: methods, ethics, regulatory challenges. Sci. Eng. Ethics 15, 343–349. (*Interesting discussion of a complex problem which will soon have to be faced*)

Flower, R.J., 2004. Lifestyle drugs: pharmacology and the social agenda. Trends Pharmacol. Sci. 25, 182–185. (*Accessible review that enlarges on some of the issues raised in this chapter*)

Gilbert, D., Walley, T., New, B., 2000. Lifestyle medicines. BMJ 321, 1341–1344. (*Short but focused review dealing mainly with the clinical implications of the 'lifestyle medicine' phenomenon*)

Lexchin, J., 2001. Lifestyle drugs: issues for debate. CMAJ 164, 1449–1451. (*Excellent review highlighting many important issues*)

Smith, R., 2002. In search of 'non-disease'. BMJ 324, 883–885. (*A fascinating and amusing piece that seeks to define 'health' and 'disease'. It's more tricky than you think! Forms a fascinating background to the lifestyle drugs debate*)

Walley, T., 2002. Lifestyle medicines and the elderly. Drugs Aging 19, 163–168. (*Excellent review of the whole area and its relevance to the treatment of the elderly*)

Young, S.N., 2003. Lifestyle drugs, mood, behaviour and cognition. J. Psychiatry Neurosci. 28, 87–89.

Drugs in sport

Avois, L., Robinson, N., Saudan, C., et al., 2006. Central nervous system stimulants and sport practice. Br. J. Sports Med. 40 (Suppl. 1), 16–20. (*Deals mainly with the illegal use of stimulants such as ephedrine, amfetamine and cocaine in sport. Particularly highlights the dangers of misuse*)

British Medical Association, 2002. Drugs in sport: the pressure to perform. BMJ Publications, London. (*Useful coverage of the whole topic*)

Catlin, D.H., Fitch, K.D., Ljungqvist, A., 2008. Medicine and science in the fight against doping in sport. J. Intern. Med. 264, 99–114. (*Very interesting review of the whole area and the establishment of antidoping agencies*)

Mottram, D.R. (Ed.), 2005. Drugs in sport, fourth ed. Routledge, London. (*Comprehensive description of pharmacological and regulatory aspects of drugs in sport, with balanced discussion of evidence relating to efficacy and risk*)

Spedding, M., Spedding, C., 2008. Drugs in sport: a scientist–athlete's perspective: from ambition to neurochemistry. Br. J. Pharmacol. 154, 496–501. (*Very accessible review written by two brothers, one of whom is an Olympic athlete and the other a pharmacologist. Unique insights. Highly recommended*)

Biopharmaceuticals and gene therapy

59

OVERVIEW

In this chapter, we review the impact of two therapeutic concepts based on our growing understanding and skill in manipulating genes. *Biopharmaceuticals* is an umbrella term applied to the use of nucleic acids or 'engineered' proteins and antibodies in medicine, while *gene therapy* refers specifically to attempts to use those nucleic acids to reprogram cells to prevent, alleviate or cure disease. Of the two, the former has already proved itself in the clinic, whereas the latter has not yet led to licensed products,[1] although there are many ongoing trials, and some convincing successes. It is clear that once the last remaining technical hurdles have been surmounted, it will hold great promise. In addition to introducing the central concepts in this chapter, we consider the considerable problems associated with developing these therapies, discuss safety issues and review the progress made to date.

INTRODUCTION

The 'molecular biology revolution', which had its roots in the discovery of the structure of DNA in the 1950s, and the advances in cell biology that followed in its train, has enabled us to manipulate the genetic material from cells in ways that are useful in practical therapeutics. The seductive notion that a gene of interest can be expressed in vitro to generate useful proteins that could not be prepared synthetically or, more daringly, that a gene could be directly introduced in vivo and persuaded to synthesise some crucial cellular component, has driven this field at breakneck speed.

Biopharmaceuticals (considered for the purposes of this chapter to comprise genetically engineered proteins and monoclonal antibodies) are already a well-recognised part of therapy, and we have already encountered them elsewhere in this book (see, for example, the anti-tumour necrosis factor [TNF] antibodies in Ch. 26). We still face many problems, not the least of which is the cost of manufacture, but the technology is established and maturing fast. Reviewing the area in 2004, Walsh noted that some 140 biopharmaceuticals had been licensed around the world by the previous year, and that 250 million patients were receiving these products at a cost of some US$30 billion.[2]

While the same basic concepts and technologies underpin both these approaches, gene therapy is the more considerable challenge. However, the idea commands such appeal that vast resources (both public and private) have been committed to its development. There are several reasons why it is so attractive. First, the approach offers the potential for radical cure of single-gene diseases such as *cystic fibrosis* and the *haemoglobinopathies*, which are collectively responsible for much misery throughout the world. Second, many other more common conditions, including malignant, neurodegenerative and infectious diseases, have a large genetic component. Conventional treatment of such disorders is, as readers of this book will have appreciated by now, woefully inadequate, so the promise of a completely new approach has enormous attraction. Finally, an ability to control gene expression could even revolutionise the management of diseases in which there is no genetic component at all.

The gurus are emphatic that 'the conceptual part of the gene therapy revolution has indeed occurred …'—so where are the therapies? The devil, of course, is in the detail: in this case, the details of:

- *pharmacokinetics*: delivery of the gene to appropriate target cells (especially in the CNS)
- *pharmacodynamics*: the controlled expression of the gene in question
- *safety*
- *clinical efficacy* and *long-term practicability*.

But perhaps the most fundamental hurdle is the delivery problem; here, modern virology has helped with techniques borrowed from viruses that can be used to introduce functional nucleic acids into mammalian cells. The principle is so simple that any broadsheet reader can apprehend it, and the potential rewards (humanitarian, scientific and commercial) so great that it has led inevitably to great expectations and, perhaps equally inevitably, to frustration at the lack of practical progress.

There is a broad consensus that the *Weismann barrier*[3] should not be breached and so gene therapy trials have focused on somatic cells. A moratorium has been agreed on therapies intended to alter the DNA of germ cells, which could influence future generations.

BIOPHARMACEUTICALS

We consider first the use of proteins as therapeutic agents. Of course, this in itself is not a novel idea; insulin, extracted from animal pancreas tissue (Ch. 30), and human growth hormone, extracted from human cadaver pituitary glands (Ch. 32), were among the first therapeutic proteins to be used, and for many years provided the only option for treating hormone deficiency disorders. However, there were problems. First, there were difficulties in extraction

[1]At least not in Western countries. A gene therapy product for treating cancer was licensed in China in 2003.

[2]Biopharmaceuticals currently comprise about 25% of new drugs approved.

[3]Named after August Weismann (1834–1914) who formulated the concept that inheritance utilises only germ, and not somatic, cells.

and disappointingly low yields. Second, in the case of insulin, administration of animal hormones to humans could evoke an immune response. Third, there was always a danger of the transmission of infectious agents across species, or between people. This was highlighted in the 1970s, when cases of *Creutzfeldt–Jakob disease* (see Ch. 39) occurred in patients treated with human growth hormone obtained from cadavers. This serious problem was later traced to contamination of the donor pituitary glands with infectious prions (Ch. 39). The advent of 'genetic engineering' techniques offered a new way to deal with these perennial problems.

Biopharmaceuticals and gene therapy: definition and potential uses

- *Biopharmaceuticals* include proteins, antibodies (and oligonucleotides) used as drugs:
  - *first-generation* biopharmaceuticals are mainly copies of endogenous proteins or antibodies, produced by recombinant DNA technology
  - *second-generation* biopharmaceuticals have been 'engineered' to improve the performance of the protein or antibody.
- Applications:
  - therapeutic monoclonal antibodies
  - recombinant hormones.
- *Gene therapy* is the genetic modification of cells to prevent, alleviate or cure disease.
- Potential applications:
  - radical cure of monogenic diseases (e.g. cystic fibrosis, haemoglobinopathies)
  - amelioration of diseases with or without a genetic component, including many malignant, neurodegenerative and infectious diseases.

PROTEINS AND POLYPEPTIDES

The biopharmaceuticals in use today are generally classified as 'first-' or 'second-' generation agents. *First-generation* biopharmaceuticals are usually straightforward copies of human hormones or other proteins prepared by *transfecting* the human gene into a suitable *expression system* (a cell line that produces the protein in good yield), harvesting and purifying the recombinant protein produced and using this as the drug. The first agent to be produced in this way was human recombinant insulin in 1982.

Second-generation biopharmaceuticals are those that have been engineered; that is to say, either the gene has been deliberately altered prior to transfection such that the structure of the expressed protein is changed, or some alteration is made to the purified end product. The reasons for making these changes are generally to improve some aspect of the protein's activity profile. Human recombinant insulins designed to act faster or last longer were among the first in this class to be marketed; Table 59.1 contains other examples. *Third-generation agents* would be those in which proteins are designed from scratch to do a particular biological function. This technology is still some way off.

PROBLEMS IN MANUFACTURE

There are several problems associated with the manufacture of any type of recombinant protein, and one of the most pressing is the choice of expression system. Many recombinant proteins are expressed in bacterial systems (*Escherichia coli*, for example), which are useful because cultures grow quickly and they are generally easy to manipulate. Disadvantages include the fact that they may contain bacterial endotoxins, which must be scrupulously removed before administration to patients, and that bacterial cells do not accomplish the same type of *post-translational processing* (e.g. glycosylation) as mammalian cells. This could pose problems if the protein's action is crucially dependent on this modification. To circumvent these problems, mammalian (e.g. Chinese hamster ovary, CHO) cells are also used as expression systems, although here the problem is often one of yield. Such cells require more careful culture, grow more slowly and produce less

Table 59.1 Some examples of 'second-generation' biopharmaceuticals

Type of change	Protein	Indication	Reason for change
Altered amino acid sequence	Insulin	Diabetes	Faster-acting hormone
	Tissue plasminogen activator analogues	Thrombolysis	Longer circulating half-life
	Interferon analogue	Antiviral	Superior antiviral action
	Factor VIII analogue	Haemophilia	Smaller molecule, better activity
	Diphtheria toxin–interleukin-2 fusion protein	T-cell lymphoma	Targets toxin to appropriate cells
	Tumour necrosis factor receptor–human immunoglobulin G F_c fusion protein	Rheumatoid disease	Prolongs half-life
Altered carbohydrate residues	Glucocerebrosidase enzyme	Gaucher's disease	Promotes phagocyte uptake
	Erythropoietin analogue	Anaemia	Prolongs half-life
Covalent attachment to polyethylene glycol	Interferon	Hepatitis C	Prolongs half-life
	Human growth hormone	Acromegaly	Prolongs half-life

Modified from Walsh, 2004.

product, all of which contribute to the cost of the final medicine.

There are, however, a number of emergent technologies that could revolutionise the production process. The use of plants to produce recombinant proteins has attracted considerable interest (see Daniell et al., 2001, and Fischer et al., 2004). Several species have shown promise, including the tobacco plant. Human genes of interest can readily be transfected into the plant by using tobacco mosaic virus as a vector; the crop grows rapidly (yields a high *biomass*) and offers a number of other advantages. But attention has also focused on edible plants such as lettuce and bananas. The advantage here is that some orally active proteins, such as vaccines, expressed in the plant could be consumed directly without the need for prior purification. Several such proteins have already been produced in plants, and some are in clinical trial.

Another technology that could dramatically increase the yield of human recombinant proteins is the use of transgenic cattle. A dairy cow can produce some 10 000 litres of milk per year, and recombinant proteins introduced into the genome, and under the control of promoters that regulate production of other milk proteins, can generate yields as high as 1 g/l (see Brink et al., 2000).

ENGINEERED PROTEINS

There are several ways in which proteins can be altered prior to expression. Alteration of the nucleotide sequence of the gene coding for the protein in question can be used to change single amino acids or, indeed, whole regions of the polypeptide chain. There are good reasons why it is an advantage to 'engineer' proteins prior to expression:

1. Modification of pharmacokinetic properties.
2. Generation of novel *fusion* or other proteins.
3. Reducing immunogenicity, e.g. by *humanising*.

It is frequently advantageous to modify the pharmacokinetic properties of recombinant proteins. Changes in the structure of human insulin, for example, provided diabetics with a form of the hormone that did not self-associate during storage and was thus faster acting and easier to manage. The half-life of proteins in the blood can often be extended by *PEGylation* (see Ch. 10), the addition of polyethylene glycol to the molecule. This *post-translational engineering* approach has been applied to some human hormones, such as recombinant growth hormone, interferons and others. Prolonging half-life is not merely a convenience to patients; it also reduces the overall cost of the treatment, and economic factors are important in the adoption of this type of therapy.

Fusion proteins comprise two or more proteins engineered to be expressed as one single polypeptide chain, sometimes joined by a short linker. An example is **etanercept**, an anti-inflammatory drug used in the treatment of rheumatoid arthritis and other conditions (see Ch. 26). This consists of the ligand-binding domain taken from the tumour necrosis factor receptor, joined to the F_c domain of a human immunoglobulin G antibody. The latter moiety increases its persistence in the blood. Reduction of immunogenicity through bioengineering is discussed below.

MONOCLONAL ANTIBODIES

Although antibodies are used to confer *passive immunity*, there are a number of disadvantages inherent in their production and use that limit their utility. Conventionally, antisera are produced from the blood of immunised humans (e.g. to collect antitetanus serum) or from animals immunised with the antigen in question (e.g. with inactivated bacterial toxins). These are used to prepare antiserum containing high levels of specific antibodies, which can then be used clinically to neutralise pathogens or other dangerous substances in the blood of the patient.

Such preparations contain *polyclonal antibodies*—that is, a mixture of antibodies from all the plasma cell clones that reacted to that particular antigen. The actual composition and efficacy of these varies over time, and obviously there is a limit to how much plasma can be collected on any one occasion. The Nobel Prize-winning discovery by Milstein and Köhler, in 1975, of a method of producing from immunised mice an immortalised *hybridoma*, a fusion of one particular lymphocytic clone with an immortalised tumour cell, provided us for the first time with a method of producing *monoclonal antibodies*, comprising a single species of defined antibody at high abundance in vitro. Because these hybridomas were immortal, the cell line could be retained indefinitely and expanded to any density while preserving the integrity of its product.

Monoclonal antibodies can be classified into first- or second-generation reagents along similar lines to other proteins discussed above. First-generation monoclonals were essentially murine monoclonals (or fragments thereof), but these suffered from several drawbacks. As mouse proteins, they provoked an immune response in 50–75% of all recipients. Other limiting factors were the short half-life in the circulation and the inability of the mouse antibodies to activate human complement.

Most of these problems have been surmounted by using either *chimeric* or *humanised* monoclonals. The two terms refer to the degree to which the monoclonals have been engineered. Figure 59.1 shows how this is done; the antibody molecule consists of a *constant* domain (F_c) and the *antibody-binding* domain (Fab), with *hypervariable* regions that recognise and bind to the antigen in question. The genes for chimeric monoclonals are engineered to contain the cDNA of the murine Fab domain coupled with the human F_c domain sequences. This greatly (around fivefold) extends the plasma half-life and improves the functionality of the antibody in human medicine. A further development (and now the preferred approach) is to replace the entire F_c and Fab region with the human equivalent with the exception of the hypervariable regions, giving a molecule that, while essentially human in nature, contains the murine antibody-binding sites. The anticancer monoclonal **herceptin** (**trastuzumab**; see Ch. 55) is an example of such a therapeutically useful antibody, and some others are given in Table 59.2.

SAFETY ISSUES

We are, by now, accustomed to the concept of using proteins therapeutically, and many of the risks associated with (for example) anti-TNF therapy are well understood (see Ch. 26). For the most part, therapeutic proteins do not cause the range of toxic effects encountered with small molecules discussed in Chapter 57, but there are still very real dangers.

In 2006, for example, a UK clinical trial of a new monoclonal antibody (TGN 1412) designed to activate T cells (see Ch. 6) and thus treat B-cell lymphocytic leukaemia went

Table 59.2 Some examples of 'second-generation' therapeutic monoclonal antibodies

Antibody	Type	Target	Use	See Chapter
Infliximab	Chimeric Mab	Tumour necrosis factor	Crohn's disease, rheumatoid disease	26
Adalimumab	Humanised Mab	Tumour necrosis factor	Rheumatoid disease	26
Etanercept	Fusion protein	Tumour necrosis factor	Rheumatoid disease	26
Trastuzumab	Humanised Mab	HER2 epidermal growth factor receptor	Breast cancer	55
Palivizumab	Humanised Mab	Respiratory syncytial virus	Respiratory infections in young children	—
Omalizumab	Humanised Mab	Immunoglobulin E	Immunoglobulin E-mediated asthma	27
Abatacept	Fusion protein	B7 epitope on antigen presenting cells	Rheumatoid disease	26

Mab, monoclonal antibody. Therapeutic monoclonal antibody names all end in '-mab', prefixed by an indication of their species nature: -umab (human), -omab (mouse), -ximab (chimera), -zumab (humanised).
Source: Walsh, 2004 and British National Formulary.

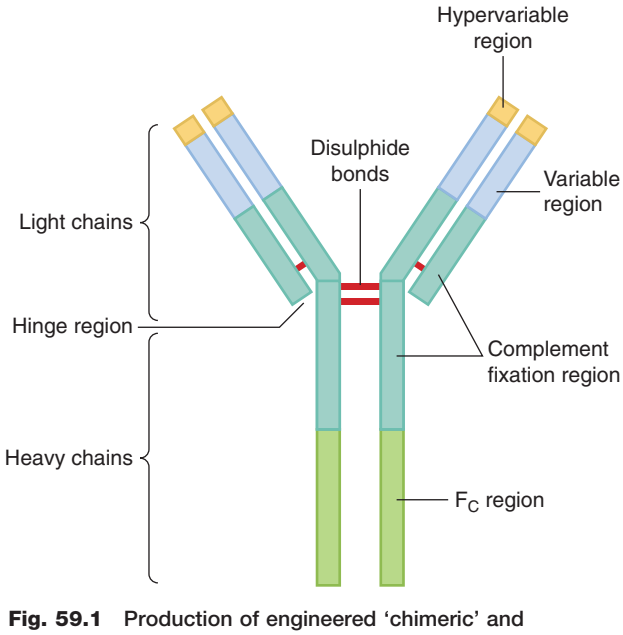

Fig. 59.1 Production of engineered 'chimeric' and 'humanised' monoclonal antibodies. The Y-shaped antibody molecule consists of two main domains: the F_c (constant) domain and the Fab (antibody-binding) domain. At the tip of the Fab regions (on the arms of the 'Y') are the hypervariable regions that actually bind the antigen. Chimeric antibodies are produced by replacing the murine F_c region with its human equivalent by altering and splicing the gene. For humanised antibodies, only the murine hypervariable regions are retained, the remainder of the molecule being human in origin. (After Walsh, 2004.)

badly wrong. All 6 subjects became severely ill following a 'cytokine storm' and suffered lasting damage. The incident provoked wide media publicity[4] and, while the subsequent investigation blamed an 'unpredictable' biological reaction, it caused many to think hard about how such trials should be conducted in the future (see Muller & Brennan, 2009). Highly specific reagents such as monoclonals pose a particular problem as they may not cross-react with the corresponding antigens of other species, thus evading detection in the usual preclinical animal safety screens.

[4]One tabloid headline read 'We saw human guinea pigs explode' (quoted by Stobbart et al., 2007).

GENE THERAPY

GENE DELIVERY

The transfer of recombinant nucleic acid into target cells—a special instance of the 'drug distribution' problem—is critical to the success of gene therapy. Nucleic acid must pass from the extracellular space across the plasma and nuclear membranes, and it must then be incorporated into the chromosomes. Because DNA is negatively charged and single genes have molecular weights around 10^4 times greater than conventional drugs, the problem is of a different order from the equivalent stage of routine drug development.

There are several important considerations in choosing a delivery system; these include:

- the *capacity* of the system (e.g. how much DNA it can carry)
- the *transfection efficiency* (its ability to enter and become utilised by cells)
- the *lifetime* of the transfected material (determined by the lifetime of the targeted cells)
- the *safety issue*, especially important in the case of viral delivery systems.

Various approaches have been developed (see Table 59.3) in an attempt to produce the optimal system.

There are two main strategies for delivering genes into patients: the in vivo and ex vivo approach. Using the *in vivo strategy*, the vector containing the therapeutic gene is injected into the patient, either intravenously (in which case some form of organ or tissue targeting is required) or directly into the target tissue (e.g. a malignant tumour). The *ex vivo strategy* is to remove cells from the patient (e.g. stem cells from bone marrow or circulating blood, or myoblasts from a biopsy of striated muscle), treat them with the vector and inject the genetically altered cells back into the patient.

Table 59.3 Characteristics of some delivery systems for gene therapy

Vector	Advantages	Disadvantages
Liposomes	Virus-free, cheap to produce	Low efficiency, sometimes cytotoxic
DNA cassettes	Virus-free	Low efficiency, expression temporary
Herpes simplex virus type I	Highly infective, persistent expression	No integration with host DNA, cytotoxic, difficult to handle
Adenovirus	Highly infective in epithelia	Immunogenic and transient, requires readministration
Adeno-associated virus	Stable	Low capacity
Retrovirus	Efficient, permanent	Low capacity, unstable, must integrate into host DNA, requires dividing cells

After Wolf & Jenkins, 2002.

An ideal vector should be *safe*, highly *efficient* (i.e. insert the therapeutic gene into a high proportion of target cells) and *selective* in that it would lead to expression of the therapeutic protein in the target cells but *not* to the expression of viral proteins. Provided that the cell into which it is inserted is itself long-lived, the vector should ideally cause persistent expression, avoiding the need for repeated treatment. The latter consideration can be a problem in some tissues. In the autosomal recessive disorder *cystic fibrosis*, for example, the airway epithelium malfunctions because it lacks a membrane Cl^- transporter known as the *cystic fibrosis transport regulator* (CFTR). Epithelial cells in the airways are continuously dying off and being replaced, so even if the CFTR gene were stably transfected into the epithelium, there would still be a periodic need for further treatment unless the gene could be inserted into the progenitor (stem) cells. Similar problems are anticipated in other cells that turn over continuously, such as gastrointestinal epithelium and skin.

VIRAL VECTORS

Many contemporary gene delivery strategies aim to capitalise on the capacity of viruses to subvert the transcriptional machinery of the cells they invade and their ability (in some cases) to fuse with the host genome. While producing a tantalising glimpse of the possible, there remain substantial practical problems with this approach, partly because as viruses have evolved the means to invade human cells, so humans have evolved immune responses and other protective mechanisms to thwart them. Although irritating in some respects, this is not all bad news from the point of view of safety.

As many of these viruses are pathogenic, they are usually modified such that they are 'replication defective' to avoid toxicity.

Retroviruses

If introduced into stem cells, retroviral vectors have the attraction that their effects are persistent because they are incorporated into, and replicate with, host DNA, and so the 'therapeutic' gene is passed down to each daughter cell during division. Against this, the retroviral integrase randomly inserts the construct into chromosomes, so it may cause damage (see below). Also, since many retroviruses show little specificity, they could infect germ or non-target cells and produce undesired effects if administered in vivo. For this reason, retroviruses have been used mainly for ex vivo gene therapy. The life cycle of naturally occurring retroviruses may be exploited to create useful vectors for gene therapy (see Fig. 59.2).

Many viruses are equipped to infect specific cell types, though not necessarily the target cell of interest. It is possible to alter the retroviral envelope to alter specificity, such that the vector could be administered systemically but would target only the desired cell population. An example of this approach with a *lentivirus* (a type of retrovirus) is the substitution of the envelope protein of a non-pathogenic vector (e.g. mouse leukaemia virus) with the envelope protein of human vesicular stomatitis virus, in order specifically to target human epithelial cells.

Most retrovirus vectors are unable to penetrate the nuclear envelope, and because the nuclear membrane dissolves during cell division, they only infect dividing cells and not non-dividing cells such as adult neurons.

Adenovirus

Adenovirus vectors are popular because of the high transgene expression that can be achieved. They transfer genes to the nucleus of the host cell, but (unlike retroviruses) these are not inserted into the host genome and so do not produce effects that outlast the lifetime of the transfected cell. This property also obviates the risk of disturbing the function of other cellular genes and the theoretical risks of carcinogenicity and germ cell transfection, although at the cost of producing only a temporary effect. Because of these favourable properties, adenovirus vectors have been used for in vivo gene therapy. The vectors are genetically modified by making deletions in the viral genome, rendering it unable to replicate or cause widespread infection in the host while at the same time creating space in the viral genome for the therapeutic transgene to be inserted.

One of the first adenoviral vectors to be used lacked part of a growth-controlling region called E_1. This defective virus was grown in a cell line that substitutes for the missing E_1 function. Recombinant virus was produced by infecting target cells with a *plasmid* containing the cloned DNA of therapeutic interest plus an expression cassette and portions of adenoviral DNA. Recombination between this and the 'backbone' of the E_1-deficient adenoviral genome resulted in a virus encoding the desired transgene. This approach led to seemingly spectacular results, demonstrating gene transfer to cell lines and animal models of disease, but it has been disappointing (e.g. in cystic fibrosis) in humans. The main problem is that low doses (administered by aerosol to patients with this disease) produce only a very low-efficiency transfer, whereas higher doses cause inflammation, a host immune response and short-lived gene expression. Furthermore, treatment cannot be repeated because of neutralising antibodies. This has led to recent attempts to manipulate adenoviral vectors to

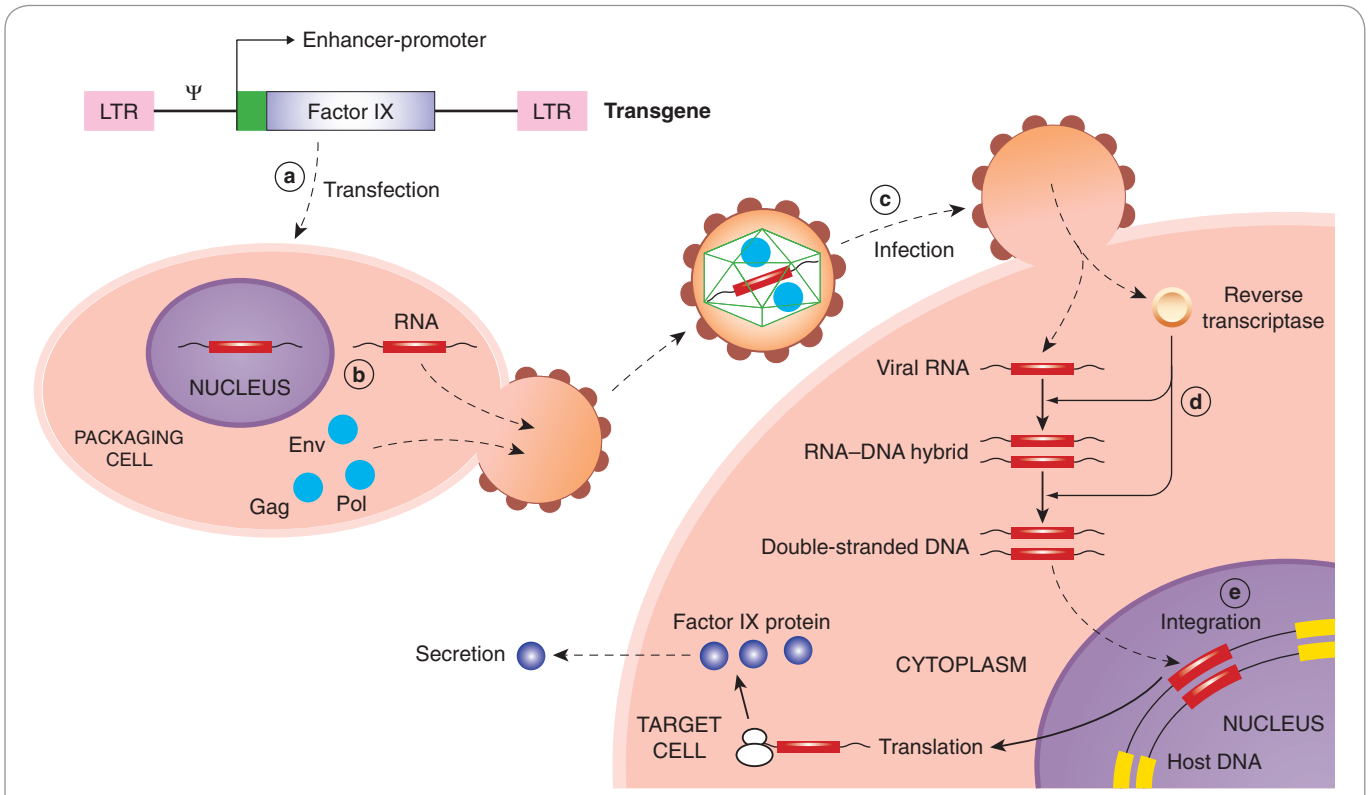

Fig. 59.2 Strategy for making retroviral vectors. The transgene (the example shows the gene for factor IX) in a vector backbone is introduced (a) into a packaging cell, where it is integrated into a chromosome in the nucleus, and (b) transcribed to make vector mRNA, which is packaged into the retroviral vector and shed from the packaging cell. It then infects the target cell (c). Virally encoded reverse transcriptase (d) converts vector RNA into an RNA–DNA hybrid, and then into double-stranded DNA, which is integrated (e) into the genome of the target cell. It can then be transcribed and translated to make factor IX protein. (Redrawn from Verma I M, Somia N 1997 Nature 389: 239–242.)

mutate or remove the genes that are most strongly immunogenic.

Other viral vectors

Other potential viral vectors under investigation include *adeno-associated virus*, *herpes virus* and disabled versions of *human immunodeficiency virus* (HIV). Adeno-associated virus associates with host DNA but is not activated unless the cell is infected with an adenovirus. It is less immunogenic than other vectors but is hard to mass produce and cannot be used to carry large transgenes. Herpes virus does not associate with host DNA but is very long lived in nervous tissue and could have a specific application in treating neurological disease. HIV, unlike most other retroviruses (see above), can infect non-dividing cells such as neurons. It is possible to remove the genes from HIV that control replication and substitute other genes. Alternatively, it may prove possible to transfer to other non-pathogenic retroviruses those genes that permit HIV to penetrate the nuclear envelope.

NON-VIRAL VECTORS

Liposomes

Non-viral vectors include a variant of liposomes (Ch. 8). Plasmids (diameter up to approximately 2 μm) are too big to package in regular liposomes (diameter 0.025–0.1 μm), but larger particles can be made from positively charged lipids ('lipoplexes'), which interact with both negatively charged cell membranes and DNA, improving delivery into the cell nucleus and incorporation into the host chromosome. Such particles have been used to deliver the genes for HLA-B7, interleukin-2 and CFTR. They are much less efficient than viruses, and attempts are currently under way to improve this by incorporating various viral signal proteins (membrane fusion proteins, for example) in their outer coat. Direct injection of these complexes into solid tumours (e.g. melanoma, breast, kidney and colon cancers) can, however, achieve high local concentrations within the tumour.

Microspheres

Biodegradable microspheres made from polyanhydride co-polymers of fumaric and sebacic acids (see Ch. 8) can be loaded with plasmid DNA. A plasmid with bacterial β-galactosidase activity formulated in this way and given by mouth to rats has resulted in systemic absorption and expression of the bacterial enzyme in the rat liver, raising the possibility of oral gene therapy.

Plasmid DNA

Surprisingly, plasmid DNA itself ('naked DNA') enters the nucleus of some cells and is expressed, albeit much less

efficiently than when it is packaged in a vector. Such DNA carries no risk of viral replication and is not usually immunogenic (although autoantibodies to DNA do occur in systemic lupus erythematosus), but it cannot be targeted to a cell of interest. There is considerable interest in the possibility of using naked DNA for vaccines, because even very small amounts of foreign protein can stimulate an immune response. Such a vaccine for influenza is in clinical development, and more ambitious long-term targets include malaria, tuberculosis, *Chlamydia*, *Helicobacter* and hepatitis.

CONTROLLING GENE EXPRESSION

To realise the full potential of gene therapy, it is not enough to transfer the gene selectively to the desired target cells and maintain acceptable expression of its product—difficult though these goals are. It is also essential that the activity of the gene is controlled. Historically, it was the realisation of the magnitude of this task that diverted attention from the haemoglobinopathies (which were the first projected targets of gene therapy). Correction of these disorders demands an appropriate balance of normal α- and β-globin chain synthesis to be effective, and for this, and many other potential applications, precisely controlled gene expression will be essential.

It has not yet proved possible to control transgenes in human recipients, but there are techniques that may eventually enable us to achieve this goal. One hinges on the use of an *inducible expression system*. This is a fairly standard technique whereby the inserted gene also includes a doxycycline-inducible promoter such that expression of the gene can be switched on or off by treatment with, or withdrawal of, doxycycline.

The control of transfected genes is important in gene targeting as well. By splicing the gene of interest with a tissue-specific promoter, it should be possible to restrict expression of the gene to the target tissue. Such an approach has been used in the design of gene therapy constructs for use in ovarian cancer, the cells of which express several proteins at high abundance, including the proteinase inhibitor SLP1. In combination with the SLP1 promoter, plasmids carrying various genes were successfully and selectively expressed in ovarian cancer cell lines (Wolf & Jenkins, 2002).

SAFETY ISSUES

Gene therapy raises a number of specific concerns that generally relate to the use of viral vectors. These are usually selected because they are non-pathogenic, or modified to render them innocuous, but there is a concern that such agents might still acquire virulence during use. Retroviruses, which insert randomly into host DNA, could damage the genome and interfere with the protective mechanisms that normally regulate the cell cycle (see Ch. 5), and if they happen to disrupt essential cellular functions, this could increase the risk of malignancy. This risk is more than a theoretical possibility; several children treated for *severe combined immunodeficiency* (SCID; see below) with a retrovirus vector developed a leukaemia-like illness (Woods et al., 2006). The retroviral vector was shown to have inserted itself into a gene called *LMO-2*. Mutations of *LMO-2* are associated with childhood cancers.

Gene delivery and expression

- Gene delivery is one of the main hurdles to practical gene therapy.
- Recombinant genes are transferred using a vector, often a suitably modified virus.
- There are two main strategies for delivering genes into patients:
  - in vivo injection of the vector directly into the patient (e.g. into a malignant tumour)
  - ex vivo treatment of cells from the patient (e.g. stem cells from marrow or circulating blood), which are then returned to the patient.
- An ideal vector would be safe, efficient, selective and produce long-lasting expression of the therapeutic gene.
- Viral vectors include retroviruses, adenoviruses, adeno-associated virus, herpesvirus and disabled human immunodeficiency virus (HIV):
  - *retroviruses* infect many different types of dividing cells and become incorporated randomly into host DNA
  - *adenoviruses* are genetically modified to prevent replication and accommodate the therapeutic transgene. They transfer genes to the nucleus but not to the genome of the host cell. Problems include a strong host immune response, inflammation and short-lived expression. Treatment cannot be repeated because of neutralising antibodies
  - *adeno-associated* virus associates with host DNA and is non-immunogenic but is hard to mass-produce and has a small capacity
  - herpesvirus does not associate with host DNA but persists in nervous tissue and may be useful in treating neurological disease
  - disabled versions of HIV differ from most other retroviruses in that they infect non-dividing cells, including neurons.
- Non-viral vectors include:
  - a variant of liposomes, made using positively charged lipids and called 'lipoplexes'
  - biodegradable microspheres, which may offer orally active gene therapy
  - plasmid DNA ('naked DNA'), which can be used as a vaccine.
- A *tetracycline-inducible expression system* or similar technique can control the activity of the therapeutic gene.

Another problem is that immunogenic viral proteins may be expressed that elicit an inflammatory response, and this could be harmful in some situations (e.g. in the airways of patients with cystic fibrosis). Initial clinical experience was reassuring, but the tragic death of Jesse Gelsinger, an 18-year-old volunteer in a gene therapy trial for the non-fatal disease *ornithine decarboxylase deficiency* (which can be controlled by diet and drugs), led to the appreciation that safety concerns related to immune-mediated responses to vectors are very real (see Marshall, 1999).

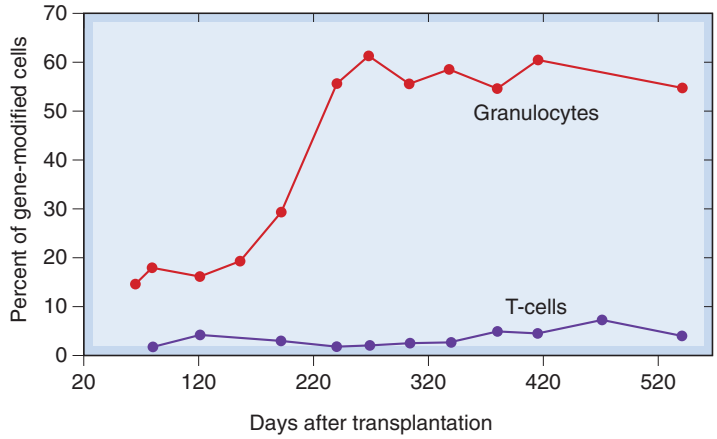

Fig. 59.3 Correcting an inherited defect using gene therapy. In this clinical trial, two patients with X-linked chronic granulomatous disease were transfused with GMCSF-treated peripheral blood cells that had been genetically modified with a retroviral vector bearing the intact gp91^{phox} gene ('in vitro protocol'—see text). The graph shows that the number of gene-modified peripheral blood leukocytes remained high for well over a year and this was accompanied by good levels of superoxide production in these cells—a clinical 'cure'. (Data redrawn from Ott et al., 2006.)

Safety

- There are those safety concerns that are *specific* to any particular therapy (e.g. polycythaemia from overexpression of erythropoietin) and additional *general* concerns relating, for example, to the nature of vectors.
- Viral vectors:
  - might acquire virulence during use
  - contain viral proteins, which may be immunogenic
  - can elicit an inflammatory response
  - could damage the host genome and interfere with the cell cycle, provoking malignancy.
- The limited clinical experience to date has not so far provided evidence of insurmountable problems.

But despite safety concerns, there have been some encouraging successes. We will finish by glancing at some of the areas where gene therapy has already proved its worth—as well as several areas of particular promise for the future.

THERAPEUTIC APPLICATIONS

SINGLE-GENE DEFECTS

Single-gene (*monogenic*) disorders were the obvious starting point for gene therapy trials. The haemoglobinopathies were the first projected targets, but early attempts (in the 1980s) were put 'on hold' because of the problem, mentioned above, posed by the need to control precisely the expression of the genes encoding the different polypeptide chains of the haemoglobin molecule. Patients with thalassaemia (the commonest monogenic disease) exhibit enormous phenotypic diversity and hence variable clinical symptoms because, even in monogenic disorders, other genes as well as environmental factors are also important.

Attention then shifted to a rare genetic disorder called *adenine deaminase deficiency*, which results in SCID. This led to the first therapeutic gene transfer protocol to be approved by the US National Institutes of Health, and subsequently a French team has treated 11 children with another form of SCID. The results provided the first proof that gene therapy can cure a life-threatening disease but also, less happily, evidence that retroviral vectors can cause malignancy.

Another early target was cystic fibrosis. Progress here has been slow: Atkinson (2008) has reviewed this area and explains the many problems associated with this approach.

More recently, however, there have been several successes. For example, X-linked *chronic granulomatous disease* (see Ch. 17) has been successfully treated using a retroviral technique to deliver a functional version of the mutated NADPH oxidase protein (Ott et al., 2006 and Fig. 59.3) and a form of inherited blindness, *Leber's congenital amaurosis*, associated with a mutation in a gene that produces retinal pigment, has been rectified using an adeno-associated virus vector bearing a cDNA coding for the intact gene (Maguire et al., 2008).

GENE THERAPY FOR CANCER

Many current clinical gene therapy trials relate to its use in cancer. The first gene transfer experiment to be approved by the National Institutes of Health was a non-therapeutic protocol in the late 1980s designed to introduce a marker gene (conferring resistance to an analogue of neomycin) into a class of lymphocytes that infiltrate various tumours. Gene transfer was performed ex vivo and the cells reinjected into the patient in order to track their subsequent redistribution. This strategy was useful in tracking other cells and hence identifying the cause of relapse following bone marrow transplantation for various leukaemias. Several therapeutic approaches are under investigation. Promising approaches include:

- restoring 'protective' proteins such as the tumour suppressor gene (see Ch. 5)
- inactivating oncogene expression (e.g. by using a retroviral vector bearing an antisense transcript RNA to the *k-ras* oncogene; see below)
- delivering a gene to malignant cells that renders them sensitive to drugs (e.g. thymidylate kinase, which activates **ganciclovir**)—the so-called 'suicide gene' approach
- delivery of proteins to healthy host cells in order to protect them (e.g. addition of the multidrug resistance channel to bone marrow cells ex vivo, thereby rendering them resistant to drugs used in chemotherapy)

Gene therapy for cancer

- Promising approaches include:
  - restoring protective proteins such as p53
  - inactivating oncogenes
  - delivering a gene to malignant cells that renders them sensitive to drugs
  - delivering a gene to healthy host cells to protect them from chemotherapy
  - tagging cancer cells with genes that make them immunogenic.

- tagging cancer cells with genes expressing proteins that render malignant cells more visible to the immune system (e.g. for antigens such as HLA-B7 or cytokines such as granulocyte macrophage colony-stimulating factor and interleukin-2).

Ovarian cancer is considered to be a good target for gene therapy because the vector can be directly introduced into the peritoneal cavity, where it is retained in a 'closed' environment. Several clinical trials are in progress or have been completed (see Wolf & Jenkins, 2002) with a variety of genes including p53 and the multidrug resistance gene, and utilising retroviral, adenoviral and liposome vectors. For a recent review of gene therapy in breast cancer, see Takahashi et al. (2006).

GENE THERAPY AND INFECTIOUS DISEASE

In addition to DNA vaccines mentioned above, there is considerable interest in the potential of gene therapy for HIV infection. Some 10% of all clinical gene therapy research is focused on this area and, by rendering stem cells (which differentiate into immune cells) resistant to HIV before they mature, aims to prevent HIV replication as well as its spread to uninfected cells. Various strategies are under investigation, including the use of genes that code for variants of HIV-directed proteins that serve as blocking agents (so-called 'dominant-negative' mutations, e.g. *rev*, which began clinical testing in 1995), RNA decoys and soluble forms of CD4 (the cellular receptor used by HIV to enter lymphocytes; Ch. 51) that will bind, and it is hoped inactivate, HIV extracellularly.

GENE THERAPY AND CARDIOVASCULAR DISEASE

Vascular gene transfer is attractive not least because cardiologists and vascular surgeons routinely perform invasive studies that offer the opportunity to administer gene therapy vectors ex vivo (e.g. to a blood vessel that has been removed to use as an autograft) or locally in vivo (e.g. by injection through a catheter directly into a diseased coronary or femoral artery). Vascular gene transfer offers potential new treatments for several cardiovascular diseases (see Ylä-Herttuala & Martin, 2000). The nature of many vascular disorders, such as restenosis following angioplasty (stretching up a narrowed artery using a balloon that can be inflated via a catheter), is such that transient gene expression might be all that is needed therapeutically. Extension of vein graft patency by gene therapy approaches has been reviewed by Chandiwal & Balasubramanian (2005). There is no shortage of attractive candidates for therapeutic overexpression in blood vessels, including nitric oxide synthase, prostacyclin synthase, thymidylate kinase, homeobox proteins and many others. Some of these have been studied in animal models of restenosis, finding that overexpression of vascular endothelial growth factor and fibroblast growth factor increases blood flow and collateral vessel growth in ischaemic leg muscle and myocardium. This is a promising area; for further details of angiogenic gene therapy, see Hammond & McKirnan (2001) and of peripheral vascular disease Ghosh et al. (2008).

Many trials are ongoing and these can be viewed online at *Gene Therapy Review* (http://www.genetherapyreview.com) and other sites (see Further Reading). Other uses of gene therapy include conditions as diverse as *uterine leiomyoma* (fibroids; Al-Hendy & Salama, 2006) and *periodontal disease* (Karthikeyan & Pradeep, 2006).

OTHER GENE-BASED APPROACHES

So far, we have largely been considering the addition of entire genes, but there are other, related nucleic acid-based therapeutic strategies. One such attempt is to correct a gene that has been adversely altered by mutation. This has the enormous theoretical advantage that the corrected gene would remain under physiological control, avoiding many of the problems discussed above. This approach is in its infancy and is beyond the scope of this book.

Other therapeutic approaches that are, in effect, gene therapies are conventionally excluded from this category. These include organ transplantation to correct a gene deficiency (e.g. liver transplantation to correct low-density-lipoprotein receptor deficiency in homozygous familial hypercholesterolaemia; Ch. 23).

Another approach is the use of *antisense oligonucleotides.* These are short (15–25mer) oligonucleotides that are complementary to part of a gene or gene product that it is desired to inhibit. These snippets of genetic material can be designed to influence the expression of a gene either by forming a triplex (three-stranded helix) with a regulatory component of chromosomal DNA, or by complexing a region of mRNA. Oligonucleotides can cross plasma and nuclear membranes by endocytosis as well as by direct diffusion, despite their molecular size and charge. However, there are abundant enzymes that cleave foreign DNA in plasma and in cell cytoplasm, so *methylphosphorate* analogues have been synthesised in which a methyl group substitutes for an oxygen atom in the nucleotide backbone. Another approach is the use of *phosphothiorate* analogues in which a negatively charged sulfur atom substitutes for oxygen (so-called 'S oligomers'). This increases water solubility as well as conferring resistance to enzymic degradation. The oligomer needs to be at least 15 bases long to confer specificity and tight binding.

Following parenteral administration, such oligomers distribute widely (although not to the central nervous system) and work in part by interfering with the transcription of mRNA and in part by stimulating its breakdown by ribonuclease H, which cleaves the bound mRNA. This approach is being used in clinical studies in patients with viral disease (including HIV infection) and malignancy (including the use of *Bcl-2* antisense therapy administered

Other gene-based approaches

- Correction of a mutated gene. This is in its infancy.
- *Antisense oligonucleotides* are short (15–25) oligonucleotides that are complementary to part of the target gene and influence expression by forming a triplex (three-stranded helix) with a regulatory component of chromosomal DNA or by complexing a region of mRNA. siRNA, which acts by a different mechanism, can be used in the same way.
- Oligonucleotides can cross plasma and nuclear membranes but there are abundant enzymes that cleave foreign DNA, so water-soluble methylphosphorate or phosphothiorate analogues, which are resistant to enzymic degradation, are used. This approach is being used in clinical trials in HIV infection and malignancy.

subcutaneously in patients with non-Hodgkin's lymphoma). A related approach (see Castanatto & Rossi, 2009), which provides more efficient gene silencing than antisense oligonucleotides, is the use of *short interfering RNA* (siRNA),[5] whereby short lengths of double-stranded RNA recruit an enzyme complex, known as *RISC*, which selectively degrades the corresponding mRNA produced by the cell, thereby blocking expression. Clinical trials of siRNA therapeutics are in progress.

[5]Discovered when it was found by plant scientists, to their surprise, that introducing RNA that encoded the colour-producing enzyme in petunias made the flowers less colourful, not more so. Subsequently siRNA has emerged as an important physiological mechanism for controlling gene expression, leading to the 2006 Nobel Prize award to Mello and Fire.

REFERENCES AND FURTHER READING

General reviews on biopharmaceuticals, gene therapy and utilities

Scientific American published an issue devoted to gene therapy in June 1997, which is an excellent introduction, including articles by T Friedmann (on 'overcoming the obstacles to gene therapy'), P L Felgner (on non-viral strategies for gene therapy), R M Blaese (on gene therapy for cancer) and D Y Ho and R M Sapolsky (on gene therapy for the nervous system).

Brink, M.F., Bishop, M.D., Pieper, F.R., 2000. Developing efficient strategies for the generation of transgenic cattle which produce biopharmaceuticals in milk. Theriogenology 53, 139–148. (*A bit specialised, as it focuses mainly on the husbandry of transgenic cattle, but interesting nonetheless*)

Castanatto, D., Rossi, J.J., 2009. The promises and pitfalls of RNA-interference-based therapeutics. Nature 457, 426–433. (*Useful review of the mechanism, current status and potential applications of RNAi as a means of controlling gene expression*)

Daniell, H., Streatfield, S.J., Wycoff, K., 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci. 6, 219–226. (*Interesting paper with some good examples*)

Fischer, N., Leger, O., 2007. Bispecific antibodies: molecules that enable novel therapeutic strategies. Pathobiology 74, 3–14. (*Interesting new twist to the therapeutic monoclonal story*)

Fischer, R., Stoger, E., Schillberg, S., et al., 2004. Plant-based production of biopharmaceuticals. Curr. Opin. Plant Biol. 7, 152–158. (*Interesting general review on the use of plants for the production of biopharmaceuticals*)

Guttmacher, A.E., Collins, F.S., 2002. Genomic medicine: a primer. N. Engl. J. Med. 347, 1512–1520. (*First in a series on genomic medicine*)

Verma, I.M., Somia, N., 1997. Gene therapy—promises, problems and prospects. Nature 389, 239–242. (*The authors, from the Salk Institute, describe the principle of getting corrective genetic material into cells to alleviate disease, the practical obstacles to this and the hopes that better delivery systems will overcome them*)

Walsh, G., 2004. Second-generation biopharmaceuticals. Eur. J. Pharm. Biopharm. 58, 185–196. (*Excellent overview of therapeutic proteins and antibodies; some good tables and figures*)

Weatherall, D.J., 2000. Single gene disorders or complex traits: lessons from the thalassaemias and other monogenic diseases. BMJ 321, 1117–1120. (*Argues that relating genotype to phenotype is the challenge for genetic medicine over the next century*)

Problems

Check, E., 2002. A tragic setback. Nature 420, 116–118. (*News feature describing efforts to explain the mechanism underlying a leukaemia-like illness in a child previously cured of SCID by gene therapy*)

Marshall, E., 1999. Gene therapy death prompts review of adenovirus vector. Science 286, 2244–2245. (*Deals with the tragic 'Gelsinger affair'*)

Muller, P.Y., Brennan, F.R., 2009. Safety assessment and dose selection for first-in-human clinical trials with immunomodulatory monoclonal antibodies. Clin. Pharmacol. Ther. 85, 247–258. (*A sober and, at times, very technical assessment of the safety precedures required for the 'first-in-man' testing of therapeutic monoclonals. Written in the wake of the TGN 1412 affair*)

Stobbart, L., Murtagh, M.J., Rapley, T., et al., 2007. 'We saw human guinea pigs explode'. BMJ 334, 566–567. (*Analysis of the press coverage of the above clinical trial*)

Woods, N.B., Bottero, V., Schmidt, M., von Kalle, C., Verma, I.M., 2006. Gene therapy: therapeutic gene causing lymphoma. Nature 440, 1123.

Therapeutic uses

Al-Hendy, A., Salama, S., 2006. Gene therapy and uterine leiomyoma: a review. Hum. Reprod. Update 12, 385–400.

Athanasopoulos, T., Fabb, S., Dickson, G., 2000. Gene therapy vectors based on adeno-associated virus: characteristics and applications to acquired and inherited diseases (review). Int. J. Mol. Med. 6, 363–375. (*Good review*)

Atkinson, T.J., 2008. Cystic fibrosis, vector-mediated gene therapy, and relevance of toll-like receptors: a review of problems, progress, and possibilities. Curr. Gene Ther. 8, 201–207.

Bauerschmitz, G.J., Barker, S.D., Hemminki, A., 2002. Adenoviral gene therapy for cancer: from vectors to targeted and replication competent agents (review). Int. J. Oncol. 21, 1161–1174. (*Superb review: very comprehensive*)

Chandiwal, A., Balasubramanian, V., Baldwin, Z.K., Conte, M.S., Schwartz, L.B., 2005. Gene therapy for the extension of vein graft patency: a review. Vasc. Endovascular Surg. 39, 1–14.

Ghosh, R., Walsh, S.R., Tang, T.Y., Noorani, A., Hayes, P.D., 2008. Gene therapy as a novel therapeutic option in the treatment of peripheral vascular disease: systematic review and meta-analysis. Int. J. Clin. Pract. 62, 1383–1390.

Hammond, H.K., McKirnan, M.D., 2001. Angiogenic gene therapy for heart disease: a review of animal studies and clinical trials. Cardiovasc. Res. 49, 561–567. (*Comprehensive review spanning animal and human trials of gene therapy for myocardial ischaemia*)

Karthikeyan, B.V., Pradeep, A.R., 2006. Gene therapy in periodontics: a review and future implications. J. Contemp. Dent. Pract. 7, 83–91.

Li, F., Hayes, J.K., Wong, K.C., 2000. Gene therapy: a novel method for the treatment of myocardial ischemia and reperfusion injury—mini-review. Acta Anaesthesiol. Sin. 38, 207–215. (*The title is self-explanatory*)

Maguire, A.M., Simonelli, F., Pierce, E.A., et al., 2008. Safety and efficacy of gene transfer for Leber's congenital amaurosis. N. Engl. J. Med. 358, 2240–2248 (*Clinical trial of gene therapy to correct a cause of congenital blindness*)

Nathwani, A.C., Davidoff, A.M., Linch, D.C., 2005. A review of gene therapy for haematological disorders. Br. J. Haematol. 128, 3–17. (*The title is self-explanantory; easy to read and comprehensive in scope*)

Ott, M.G., Schmidt, M., Schwarzwaelder, K., et al., 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat. Med. 12,

401–409. (*Clinical trial of gene therapy to correct hereditary neutrophil dysfunction*)

Roth, J.A., Grammer, S.F., 2004. Gene replacement therapy for non-small cell lung cancer: a review. Hematol. Oncol. Clin. North Am. 18, 215–229. (*Useful and readable paper on 'replacement therapy' in cancer therapy*)

Takahashi, S., Ito, Y., Hatake, K., Sugimoto, Y., 2006. Gene therapy for breast cancer. Review of clinical gene therapy trials for breast cancer and MDR1 gene therapy trial in Cancer Institute Hospital. Breast Cancer 13, 8–15.

Wolf, J.K., Jenkins, A.D., 2002. Gene therapy for ovarian cancer (review). Int. J. Oncol. 21, 461–468. (*Excellent review and broad introduction to gene therapy in general*)

Ylä-Herttuala, S., Martin, J.F., 2000. Cardiovascular gene therapy. Lancet 355, 213–222. (*Reviews rationale, vectors, delivery, therapeutic targets, human trials, ethics and future directions*)

Useful Web resources

http://www.genetherapynet.com (*Gene Therapy Net – a fantastic resource for both patients and professionals. It is a veritable clearing house for information and up-to-date news on all aspects of gene therapy. It even advertises for volunteers and has a 'jobs' section, in case you are tempted! Has links to other related sites*)

60 Drug discovery and development

OVERVIEW

With the development of the pharmaceutical industry towards the end of the 19th century, drug discovery became a highly focused and managed process. Discovering new drugs moved from the domain of inventive doctors to that of scientists hired for the purpose. Today, the bulk of modern therapeutics, and of modern pharmacology, is based on drugs that came from the laboratories of pharmaceutical companies, without which neither the practice of therapeutics nor the science of pharmacology would be more than a pale fragment of what they have become.

In this chapter, we describe in outline the main stages of the process, namely (i) the discovery phase, i.e. the identification of a new chemical entity as a potential therapeutic agent; and (ii) the development phase, during which the compound is tested for safety and efficacy in one or more clinical indications, and suitable formulations and dosage forms devised. The aim is to achieve registration by one or more regulatory authorities, to allow the drug to be marketed legally as a medicine for human use.

Our account is necessarily brief and superficial, and more detail can be found elsewhere (Rang, 2006).

THE STAGES OF A PROJECT

Figure 60.1 shows in an idealised way the stages of a 'typical' project, aimed at producing a marketable drug that meets a particular medical need (e.g. to retard the progression of Parkinson's disease or cardiac failure, or to prevent migraine attacks).

Broadly, the process can be divided into three main components:

1. **Drug discovery**, during which candidate molecules are chosen on the basis of their pharmacological properties.
2. **Preclinical development**, during which a wide range of non-human studies (e.g. toxicity testing, pharmacokinetic analysis and formulation) are performed.
3. **Clinical development**, during which the selected compound is tested for efficacy, side effects and potential dangers in volunteers and patients.

These phases do not necessarily follow in strict succession as indicated in Figure 60.1, but generally overlap.

THE DRUG DISCOVERY PHASE

Given the task of planning a project to discover a new drug to treat—say, Parkinson's disease—where does one start? Assuming that we are looking for a novel drug rather than developing a slightly improved 'me too' version of a drug already in use,[1] we first need to choose a new molecular target.

TARGET SELECTION

As discussed in Chapter 2, drug targets are, with few exceptions, functional proteins (e.g. receptors, enzymes, transport proteins). Although, in the past, drug discovery programmes were often based—successfully—on measuring a complex response in vivo, such as prevention of experimentally induced seizures, lowering of blood sugar or suppression of an inflammatory response, without the need for prior identification of a drug target, nowadays it is rare to start without a defined protein target, so the first step is target identification. This most often comes from biological intelligence. It was known, for example, that inhibiting angiotensin-converting enzyme lowers blood pressure by suppressing angiotensin formation, so it made sense to look for antagonists of the vascular angiotensin II receptor—hence the successful 'sartan' series of antihypertensive drugs (Ch. 22). Similarly, the knowledge that breast cancer is often oestrogen sensitive led to the development of aromatase inhibitors such as **anastrazole**, which prevents oestrogen synthesis. Current therapeutic drugs address about 120 distinct targets (see Hopkins & Groom, 2002; Rang, 2006), but there are still many proteins that are thought to play a role in disease for which we still have no cognate drug, and many of these represent potential starting points for drug discovery. Estimates range from a few hundred to several thousand potential drug targets that remain to be exploited therapeutically (see Betz, 2005). Selecting *valid* and '*druggable*' targets from this plethora is a major challenge.

Conventional biological wisdom, drawing on a rich fund of knowledge of disease mechanisms and chemical signalling pathways, remains the basis on which novel targets are most often chosen. However, genomics is playing an increasing role by revealing new proteins involved in chemical signalling and new genes involved in disease. Space precludes discussion here of this burgeoning area; interested readers are referred to more detailed accounts (Lindsay, 2003; Kramer & Cohen, 2004; Betz, 2005; Rang, 2006).

Overall, it is evident that in the foreseeable future there is ample biological scope in terms of novel drug targets for therapeutic innovation. The limiting factor is not the biology and pharmacology, but other factors, such as the emergence of unexpected adverse effects during clinical

[1]Many commercially successful drugs have in the past emerged from exactly such 'me too' projects, examples being the dozen or so β-adrenoceptor-blocking drugs developed in the wake of propranolol, or the plethora of 'triptans' that followed the introduction of sumatriptan to treat migraine. Quite small improvements (e.g. in pharmacokinetics or side effects), coupled with aggressive marketing, have often proved enough, but the barriers to registration are getting higher, so the emphasis has shifted towards developing innovative (first in class) drugs aimed at novel molecular targets.

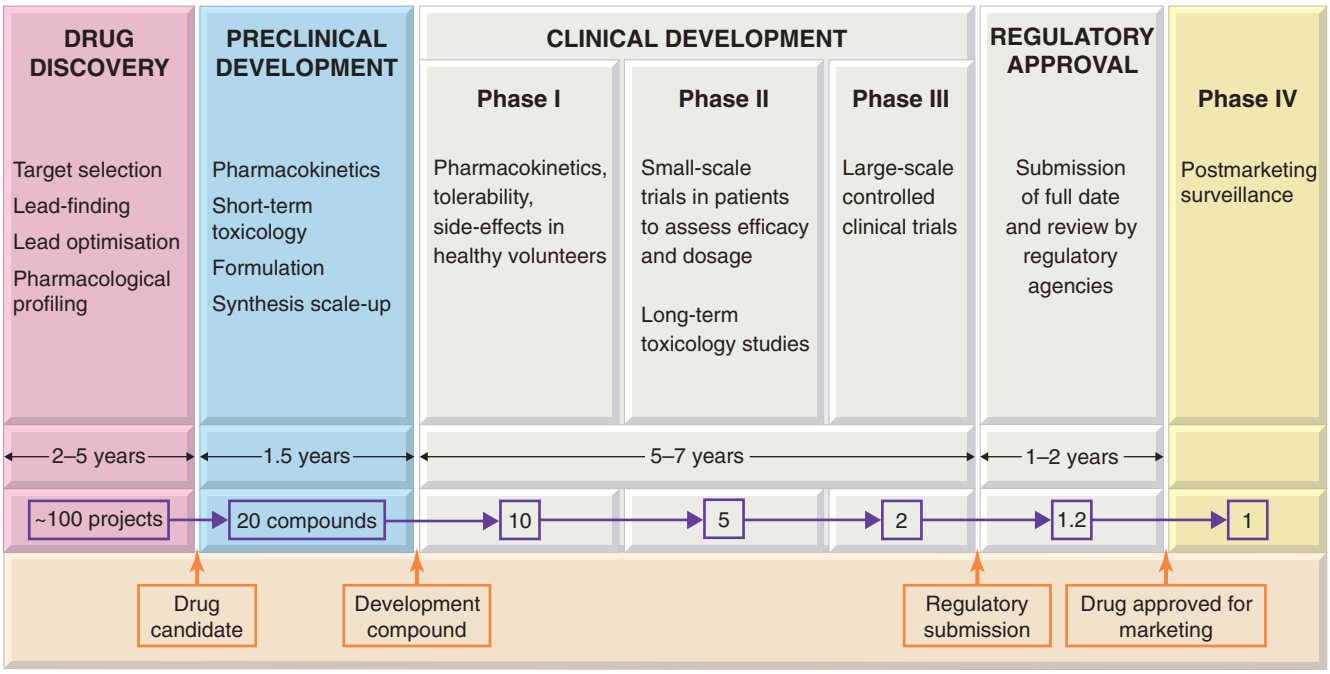

Fig. 60.1 The stages of development of a 'typical' new drug, i.e. a synthetic compound being developed for systemic use. Only the main activities undertaken at each stage are shown, and the details vary greatly according to the kind of drug being developed.

testing, and the cost and complexity of drug discovery and development in relation to healthcare economics and increasing regulatory hurdles.

LEAD FINDING

When the biochemical target has been decided and the feasibility of the project has been assessed, the next step is to find *lead compounds*. The usual approach involves cloning of the target protein—normally the human form, because the sequence variation among species is often associated with pharmacological differences, and it is essential to optimise for activity in humans. An assay system must then be developed, allowing the functional activity of the target protein to be measured. This could be a cell-free enzyme assay, a membrane-based binding assay or a cellular response assay. It must be engineered to run automatically, if possible with an optical read-out (e.g. fluorescence or optical absorbance), and in a miniaturised multiwell plate format for reasons of speed and economy. Robotically controlled assay facilities capable of testing tens of thousands of compounds per day in several parallel assays are now commonplace in the pharmaceutical industry, and have become the standard starting point for most drug discovery projects. For details on high-throughput screening, see Sundberg (2000) and Hüser (2006).

To keep such hungry monsters running requires very large compound libraries. Large companies will typically maintain a growing collection of a million or more synthetic compounds, which will be routinely screened whenever a new assay is set up. Whereas, in the past, compounds were generally synthesised and purified one by one, often taking a week or more for each, the present tendency is to use combinatorial chemistry, which allows families of several hundreds or thousands of related compounds to be made simultaneously. By coupling such high-speed chemistry to high-throughput assay systems, the time taken over the initial lead-finding stage of projects has been reduced to a few months in most cases, having previously often taken several years. Despite the apparent mindlessness of the high-throughput random screening approach, it is often successful in identifying lead compounds that have the appropriate pharmacological activity and are amenable to further chemical modification. Building and maintaining huge compound libraries is, however, a costly business, and it has to be realised that even the largest practicable compound collection represents only a minute fraction of the number of 'drug-like' molecules that exists in theory—estimated at about 10^{60}.

One problem with random screening is that many of the 'hits' detected in the initial screen turn out to be molecules that have features undesirable in a drug, such as too high a molecular weight, excessive polarity and possession of groups known to be associated with toxicity. Computational 'prescreening' of compound libraries is often used to eliminate such compounds.

The hits identified from the primary screen are used as the basis for preparing sets of homologues by combinatorial chemistry so as to establish the critical structural features necessary for binding selectively to the target. Several such iterative cycles of synthesis and screening are usually needed to identify one or more lead compounds for the next stage.

Natural products as lead compounds

Historically, natural products, derived mainly from fungal and plant sources, have proved to be a fruitful source of new therapeutic agents, particularly in the field of anti-infective, anticancer and immunosuppressant drugs.

Familiar examples include **penicillin**, **streptomycin** and many other antibiotics; vinca alkaloids; **paclitaxel**; **ciclosporin**; and **sirolimus** (**rapamycin**). These substances presumably serve a specific protective function, having evolved so as to recognise with great precision vulnerable target molecules in an organism's enemies or competitors. The surface of this resource has barely been scratched, and many companies are actively engaged in generating and testing natural product libraries for lead-finding purposes. Fungi and other microorganisms are particularly suitable for this, because they are ubiquitous, highly diverse, and easy to collect and grow in the laboratory. Compounds obtained from plants, animals or marine organisms are much more troublesome to produce commercially. The main disadvantage of natural products as lead compounds is that they are often complex molecules that are difficult to synthesise or modify by conventional synthetic chemistry, so that lead optimisation may be difficult and commercial production very expensive.

LEAD OPTIMISATION

Lead compounds found by random screening are the basis for the next stage, lead optimisation, where the aim (usually) is to increase the potency of the compound on its target and to optimise it with respect to other properties, such as selectivity and metabolic stability. In this phase, the tests applied include a broader range of assays on different test systems, including studies to measure the activity and time course of the compounds in vivo (where possible in animal models mimicking aspects of the clinical condition; see Ch. 7), and checking for unwanted effects in animals, evidence of genotoxicity and usually for oral absorption. The objective of the lead optimisation phase is to identify one or more *drug candidates* suitable for further development.

As shown in Figure 60.1, only about one project in five succeeds in generating a drug candidate, and it can take up to 5 years. The most common problem is that lead optimisation proves to be impossible; despite much ingenious and back-breaking chemistry, the lead compounds, like antisocial teenagers, refuse to give up their bad habits. In other cases, the compounds, although they produce the desired effects on the target molecule and have no other obvious defects, fail to produce the expected effects in animal models of the disease, implying that the target is probably not a good one. The virtuous minority proceed to the next phase, preclinical development.

PRECLINICAL DEVELOPMENT

The aim of preclinical development is to satisfy all the requirements that have to be met before a new compound is deemed ready to be tested for the first time in humans. The work falls into four main categories:

1. Pharmacological testing to check that the drug does not produce any obviously hazardous acute effects, such as bronchoconstriction, cardiac dysrhythmias, blood pressure changes and ataxia. This is termed *safety pharmacology*.
2. Preliminary toxicological testing to eliminate genotoxicity and to determine the maximum non-toxic dose of the drug (usually when given daily for 28 days, and tested in two species). As well as being checked regularly for weight loss and other gross changes, the animals so treated are examined minutely postmortem at the end of the experiment to look for histological and biochemical evidence of tissue damage.
3. Pharmacokinetic testing, including studies on the absorption, metabolism, distribution and elimination (ADME studies) in laboratory animals.
4. Chemical and pharmaceutical development to assess the feasibility of large-scale synthesis and purification, to assess the stability of the compound under various conditions and to develop a formulation suitable for clinical studies.

Much of the work of preclinical development, especially that relating to safety issues, is done under a formal operating code, known as *Good Laboratory Practice* (GLP), which covers such aspects as record-keeping procedures, data analysis, instrument calibration and staff training. The aim of GLP is to eliminate human error as far as possible, and to ensure the reliability of the data submitted to the regulatory authority, and laboratories are regularly monitored for compliance to GLP standards. The strict discipline involved in working to this code is generally ill-suited to the creative research needed in the earlier stages of drug discovery, so GLP standards are not usually adopted until projects get beyond the discovery phase.

Roughly half the compounds identified as drug candidates fail during the preclinical development phase; for the rest, a detailed dossier is prepared for submission to the regulatory authority such as the European Medicines Evaluation Agency or the US Food and Drugs Administration, whose permission is required to proceed with studies in humans. This is not lightly given, and the regulatory authority may refuse permission or require further work to be done before giving approval.

Non-clinical development work continues throughout the clinical trials period, when much more data, particularly in relation to long-term toxicity in animals, has to be generated. If a drug is intended for long-term use in the clinic, the toxicology studies may have to be extended for up to 2 years, and may include time-consuming studies for possible effects on fertility and fetal development. Failure of a compound at this stage is very costly, and considerable efforts are made to eliminate potentially toxic compounds much earlier in the drug discovery process by the use of in vitro, or even in silico, methods.

CLINICAL DEVELOPMENT

Clinical development proceeds through four distinct phases (see Friedman et al., 1996, for details):

- *Phase I studies* are performed on a small group (normally 20–80) of normal healthy volunteers, and their aim is to check for signs of any potentially *dangerous effects*, for example on cardiovascular, respiratory, hepatic or renal function; *tolerability* (does the drug produce any unpleasant symptoms, for example headache, nausea, drowsiness?); and *pharmacokinetic properties* (is the drug well absorbed? What is the time course of the plasma concentration? Is there evidence of cumulation or non-linear kinetics?). Phase I studies may also test for pharmacodynamic effects in volunteers (e.g. does a novel analgesic compound block experimentally induced pain in humans? How does the effect vary with dose?).

- *Phase II studies* are performed on groups of patients (normally 100–300) and are designed to test for efficacy in the clinical situation, and if this is confirmed, to establish the dose to be used in the definitive phase III study. Often, such studies will cover several distinct clinical disorders (e.g. depression, anxiety states and phobias) to identify the possible therapeutic indications for the new compound and the dose required. When new drug targets are being studied, it is not until these phase II trials are completed that the team finds out whether or not its initial hypothesis was correct, and lack of the expected efficacy is a common reason for failure.
- *Phase III studies* are the definitive double-blind, randomised trials, commonly performed as multicentre trials on thousands of patients, aimed at comparing the new drug with commonly used alternatives. These are extremely costly, difficult to organise and often take years to complete, particularly if the treatment is designed to retard the progression of a chronic disease. It is not uncommon for a drug that seemed highly effective in the limited patient groups tested in phase II to look much less impressive under the more rigorous conditions of phase III trials.

  ▼ The conduct of trials has to comply with an elaborate code known as Good Clinical Practice, covering every detail of the patient group, data collection methods, recording of information, statistical analysis and documentation.[2]

  Increasingly, phase III trials are being required to include a *pharmacoeconomic analysis* (see Ch. 1), such that not only clinical but also economic benefits of the new treatment are assessed.

  At the end of phase III, the drug will be submitted to the relevant regulatory authority for licensing. The dossier required for this is a massive and detailed compilation of preclinical and clinical data. Evaluation by the regulatory authority normally takes a year or more, and further delays often arise when aspects of the submission have to be clarified or more data are required. Eventually, about two-thirds of submissions gain marketing approval. Overall, only 11.5% of compounds entering Phase I are eventually approved (see Munos, 2009). Increasing this proportion by better compound selection at the laboratory stage is one of the main challenges for the pharmaceutical industry.
- *Phase IV studies* comprise the obligatory postmarketing surveillance designed to detect any rare or long-term adverse effects resulting from the use of the drug in a clinical setting in many thousands of patients. Such events may necessitate limiting the use of the drug to particular patient groups, or even withdrawal of the drug.[3]

BIOPHARMACEUTICALS

'Biopharmaceuticals', i.e. therapeutic agents produced by biotechnology rather than conventional synthetic chemistry, are discussed in Chapter 59. Such therapeutic agents comprise an increasing proportion—currently about 30%—of new products registered each year. The principles underlying the development and testing of biopharmaceuticals are basically the same as for synthetic drugs. In practice, biopharmaceuticals generally run into fewer toxicological problems than synthetic drugs,[4] but more problems relating to production, quality control and drug delivery. Walsh (2003) covers this specialised field in more detail.

COMMERCIAL ASPECTS

Figure 60.1 shows the approximate time taken for such a project and the attrition rate (at each stage and overall) based on recent data from several large pharmaceutical companies. The key messages are (i) that it is a high-risk business, with only about one drug discovery project in 50 reaching its goal of putting a new drug on the market, (ii) that it takes a long time—about 12 years on average, and (iii) that it costs a lot of money to develop one drug (currently a mind-boggling $3.9 billion in 2008, see Munos, 2009).[5] For any one project, the costs escalate rapidly as development proceeds, phase III trials and long-term toxicology studies being particularly expensive. The time factor is crucial, because the new drug has to be patented, usually at the end of the discovery phase, and the period of exclusivity (20 years in most countries) during which the company is free from competition in the market starts on that date. After 20 years, the patent expires, and other companies, which have not supported the development costs, are free to make and sell the drug much more cheaply, so the revenues for the original company decrease rapidly thereafter. Many profitable drugs will come to the end of their patents between 2010 and 2015, adding to the industry's problems. Reducing the development time after patenting is a major concern for all companies, but so far it has remained stubbornly fixed at around 10 years, partly because the regulatory authorities are demanding more clinical data before they will grant a licence. In practice, only about one drug in three that goes on the market brings in enough revenue to cover its development costs. Success for the company relies on this one drug generating enough profit to pay for the rest.[6]

FUTURE PROSPECTS

Since about 1990, the drug discovery process has been in the throes of a substantial methodological revolution, following the rapid ascendancy of molecular biology, genomics and informatics, amid high expectations that this would bring remarkable dividends in terms of speed, cost and success rate. High-throughput screening has undoubtedly emerged as a powerful lead-finding technology, but overall the benefits are not yet clear: costs have risen steadily, the

[2]Similar highly detailed codes must be followed in laboratory tests to determine safety (Good Laboratory Practice; see text) and drug manufacture (Good Manufacturing Practice).

[3]Recent high-profile cases include the withdrawal of **rofecoxib** (a cyclo-oxygenase-2 inhibitor; see Ch. 26) when it was found to increase the frequency of heart attacks, and of **cerivastatin** (Ch. 23), a cholesterol-lowering drug found to cause severe muscle damage in a few patients.

[4]The serious toxicity caused to human volunteers in the 2006 Phase I trials of the monoclonal antibody TGN 1412 (see Ch. 59) showed that this could not be relied on, and has led to substantial tightening of standards (and slowdown of the development of biopharmaceuticals).

[5]These cost estimates have been strongly challenged by commentators (see Angell, 2004) who argue that the pharmaceutical companies overestimate their costs several-fold in order to justify high drug prices.

[6]Actually, companies spend about twice as much on marketing and administration as on research and development.

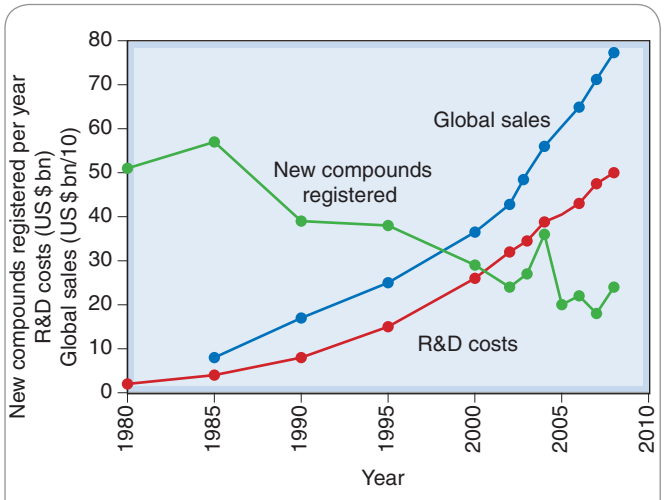

Fig. 60.2 Research and development (R&D) spend, sales and new drug registrations, 1980–2010. Registrations refer to new chemical entities (including biopharmaceuticals, excluding new formulations and combinations of existing registered compounds). (Data from various sources, including the Centre for Medicines Research, Pharmaceutical Research and Manufacturers Association of America.)

success rate has not improved (Fig. 60.2) and development times have not decreased.

Figure 60.2 illustrates the steady decline in the number of new drugs launched in the major markets worldwide, despite escalating costs and improved technology. There has been much speculation as to the causes, the optimistic view (see below) being that fewer but better drugs are being introduced, and that the recent technological jump has yet to make its impact.

If the new drugs that are being developed improve the quality of medical care, there is room for optimism. In recent ('prerevolutionary') years, synthetic drugs aimed at new targets (e.g. selective serotonin reuptake inhibitors, statins and the kinase inhibitor **imatinib**) have made major contributions to patient care. Even if the new technologies do not improve productivity, we can reasonably expect that their ability to make new targets available to the drug discovery machine will have a real effect on patient care.

Trends to watch include the growing armoury of biopharmaceuticals, particularly monoclonal antibodies such as **trastuzumab** (an oestrogen receptor antibody used to treat breast cancer) and **infliximab** (a tumour necrosis factor antibody used to treat inflammatory disorders; see Ch. 26); these are successful recent examples, and more are in the pipeline. Another likely change will be the use of genotyping to 'individualise' drug treatments, so as to reduce the likelihood of administering drugs to 'non-responders' (see Ch. 11, which summarises the current status of 'personalised medicine'). The implications for drug discovery will be profound, for the resulting therapeutic compartmentation of the patient population will mean that markets will decrease, bringing to an end the reliance on the 'blockbusters' referred to earlier. At the same time, clinical trials will become more complex (and expensive), as different genotypic groups will have to be included in the trial design. The hope is that therapeutic efficacy will be improved, not that it will be a route to developing drugs more cheaply and quickly. However, there is general agreement that the current modus operandi is commercially unsustainable (see Munos, 2009). Costs and regulatory requirements are continuing to rise, and the anticipated use of genomics to define subgroups of patients likely to respond to particular therapeutic agents (see Ch. 11) will mean fragmentation of the market, as we move away from the 'one-drug-suits-all' approach that has encouraged companies to focus their efforts on blockbuster drugs. More niche products targeted at smaller patient groups will be needed, though each costs as much to develop as a blockbuster and carries a similar risk of failure.

A FINAL WORD

The pharmaceutical industry in recent years has attracted much adverse publicity, some of it well deserved, concerning drug pricing and profits, non-disclosure of adverse clinical trials data, reluctance to address major global heath problems such as tuberculosis and malaria, aggressive marketing practices and much else (see Angell, 2004). It needs to be remembered though that, despite its faults, the industry has been responsible for most of the therapeutic advances of the past half-century, without which medical care would effectively have stood still.

REFERENCES AND FURTHER READING

Angell, M., 2004. The truth about the drug companies. Random House, New York. (*A powerful broadside directed against the commercial practices of pharmaceutical companies*).

Betz, U.A.K., 2005. How many genomics targets can a portfolio afford? Drug Discov. Today 10, 1057–1063. (*Interesting analysis – despite its odd title – of approaches to target identification in drug discovery programmes*)

Drews, J., 1998. In quest of tomorrow's medicines. Springer, New York. (*Thoughtful and non-technical account of the history, principles and future directions of drug discovery*)

Evans, W.E., Relling, M.V., 2004. Moving towards individualised medicine with pharmacogenomics. Nature 429, 464–468. (*Good review article discussing the likely influence of pharmacogenomics on therapeutics*)

Friedman, L.M., Furberg, C.D., DeMets, D.L., 1996. Fundamentals of clinical trials, third ed. Mosby, St Louis. (*Standard textbook*)

Hopkins, A.L., Groom, C.R., 2002. The druggable genome. Nat. Rev. Drug Discov. 1, 727–730. (*Interesting analysis of the potential number of drug targets represented in the human genome*)

Hüser, J. (Ed.), 2006. High throughput screening in drug discovery. Vol. 35 of Methods and principles in drug discovery. Wiley-VCH, Weinheim. (*Comprehensive textbook covering all aspects of this technology*)

Kramer, R., Cohen, D., 2004. Functional genomics to new drug targets. Nat. Rev. Drug Discov. 3, 965–972. (*Describes the various approaches for finding new drug targets, starting from genomic data*)

Lindsay, M.A., 2003. Target discovery. Nat. Rev. Drug Discov. 2, 831–836. (*Well-balanced discussion of the application of genomics approaches to discovering new drug targets; more realistic in its stance than many*)

Munos, B., 2009. Lessons from 60 years of pharmaceutical innovation. Nat. Rev. Drug Discov. 8, 959–968. (*Informative summary of the current status of the drug discovery industry, making clear that the modus operandi that has been successful in the past is no longer sustainable*)

Rang, H.P. (Ed.), 2006. Drug discovery and development. Elsevier, Amsterdam. (*Short textbook describing the principles and practice of drug discovery and development at the beginning of the 21st century*)

Sundberg, S.A., 2000. High-throughput and ultra-high-throughput screening: solution and cell-based approaches. Curr. Opin. Biotechnol. 11, 47–53.

Walsh, G., 2003. Biopharmaceuticals, second ed. Wiley, Chichester. (*Comprehensive textbook covering all aspects of discovery, development and applications of biopharmaceuticals*)