18 Errors of translation

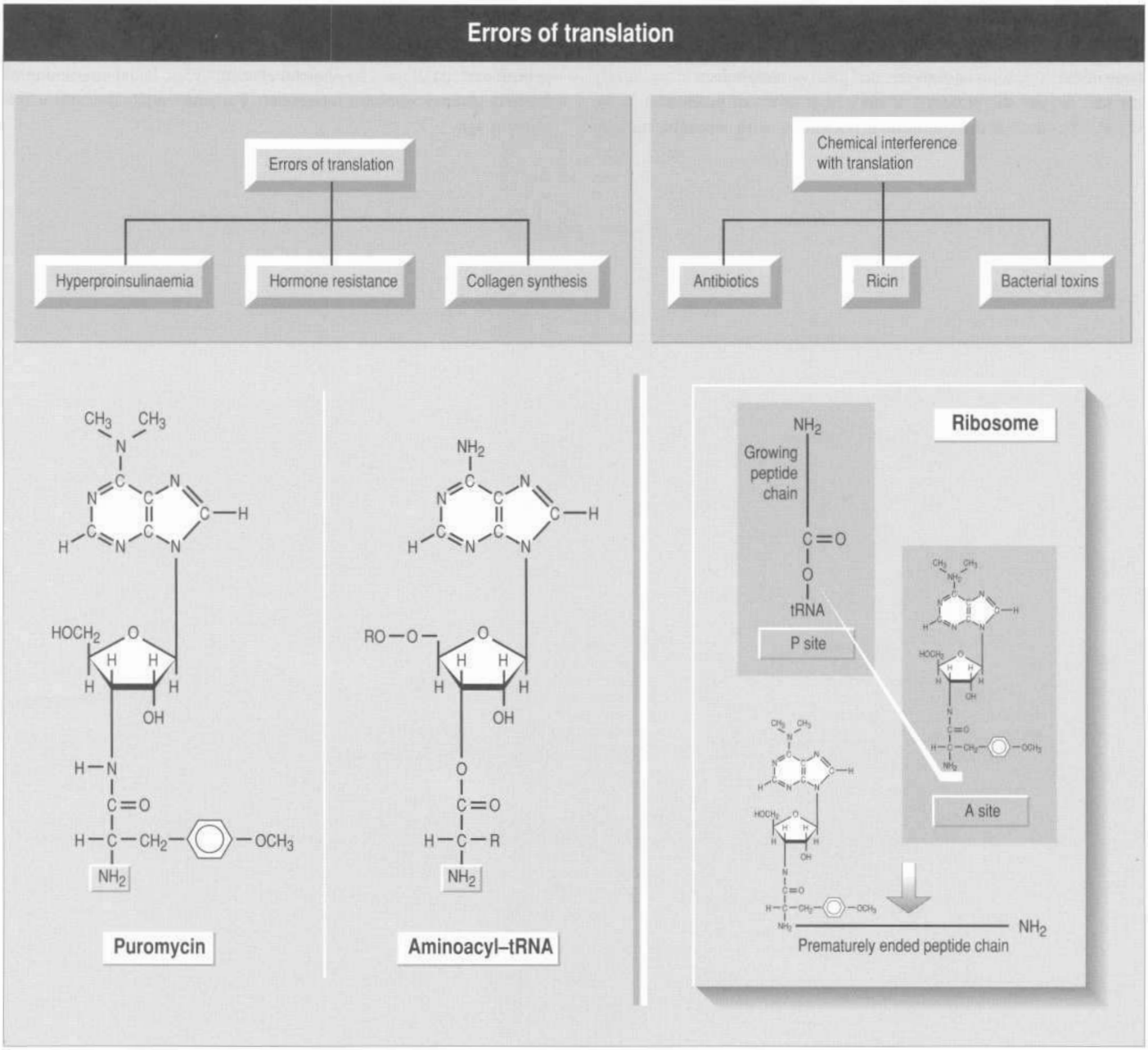

Fig. 18.1

Protein synthesis can be blocked by drugs, termed **antibiotics**. This term encompasses agents that have actions other than translation. They may interfere with cell wall synthesis or with intermediary metabolism. They are useful tools for elucidating the mechanisms of translation, and some are extremely useful therapeutically because they target the prokaryote translation machinery specifically. Defects in the translation process may also be responsible for several diseases in humans.

INHIBITORS OF PROTEIN SYNTHESIS

Antibiotics

Streptomycin and **neomycin** are aminoglycoside antibiotics extensively used therapeutically. Streptomycin inhibits initiation by binding to the 30S ribosomal subunit and interfering with the functioning of a single ribosomal protein called S12, which mediates the binding of

Inhibitor	Site of inhibition	Process inhibited
Chloramphenicol	Prokaryote 50S subunit	Peptidyl transferase
Cycloheximide	Eukaryote 80S ribosome	Elongation
Erythromycin	Prokaryote 50S subunit	Translocation
Fusidic acid	Prokaryote EF-G action	Translocation
Neomycins	Prokaryotes many sites	Translation
Puromycin*	Ribosome	Peptide transfer
Ricin	Eukaryote	Many processes
Streptomycin	Prokaryote 30S subunit	Initiation Elongation
Tetracyclines	Prokaryote 30S subunit	Aminoacyl–tRNA Binding

* Eukaryotes and prokaryotes.

fMet–aminoacyl–tRNA and mRNA to the ribosome. It also causes misreading of the mRNA codons. Neomycin does not appear to act through S12. **Tetracyclines** also bind to the ribosome and block aminoacyl–tRNA binding. They were heavily prescribed in children until it was discovered that they stained their permanent teeth yellow. **Chloramphenicol** is toxic to a wide variety of microorganisms, but is also seriously toxic to bone marrow, and is reserved for dangerous diseases such as typhoid fever, or for localized (topical) application in, for example eye drops.

Antibiotics, such as **cycloheximide** and **puromycin**, which attack both prokaryotic and eukaryotic translation, are virtually useless therapeutically, but are very powerful experimental tools. Puromycin resembles the terminal aminoacyl–adenosine part of the aminoacyl–tRNA and therefore competes with it for the binding site on the A site of the ribosome. Puromycin becomes incorporated into the growing peptide chain at the carboxyl end because it has an α-amino group recognized by peptidyl transferase. **Ricin** is a plant toxin from castor beans, that has *N*-glycosidase activity and attacks the ribosome directly, cleaving an adenine base from the large subunit. Recently, in Spain, many hundreds of people died after using cooking oil contaminated with ricin.

Translation can also be inhibited by **bacterial toxins**. For example, a protein toxin produced by the diphtheria-producing bacterium *Corynebacterium diphtheria*, enzymatically inactivates EF2 by converting it to ADP–ribosyl EF2.

TRANSLATION AND DISEASE

Several, usually familial, diseases are caused by defects in the translation process. These may involve alterations to the **cleavage** of newly synthesized peptide chains, or defects in the **targeting** of proteins or from **mutations** in the genes, and thus in the mRNA, resulting in the incorporation of the wrong amino acid into the protein.

Cleavage

Defects of cleavage are responsible for **familial hyperproinsulinaemia**, in which affected individuals have abnormally high circulating levels of the insulin precursor proinsulin. Normally, the insulin gene codes for a large precursor, preproinsulin, which is released into the lumen of the ER. The signal sequence is cleaved by a signal peptidase to yield a smaller protein, proinsulin. This is transported to the Golgi apparatus, where it is packaged into vesicles for export. Inside the vesicle, the molecule is cleaved further to remove the so-called C peptide, an internal peptide chain, and bioactive insulin is eventually released by exocytosis.

Receptor defects and hormone resistance

Errors of translation can result in **hormone resistance**. This is a failure of the target cell to respond to certain hormones, for example androgens, glucocorticoids, thyroid hormones and vitamin D, which all act through intracellular receptors

In patients with **androgen** resistance, these receptors may be: (i) absent altogether; (ii) grossly altered in structure due to deletion of segments of the carboxyterminal end of the polypeptide chain; or (iii) altered in just one amino acid (a **point mutation**, which results from a mutation of a single base pair in the DNA). All these are due to mutations of the androgen receptor gene, and the condition is therefore familial. As a result, the patient does not respond to his own androgen, and is infertile.

Mutations of different regions of the **low-density lipoprotein (LDL) receptor gene** can cause a disruption of cholesterol metabolism, and lead to hypercholesterolaemia and premature vascular disease. The mutation may result in: (i) a reduced number of receptors or none at all; (ii) a reduced rate of transport of the newly synthesized receptor from the ER to the Golgi apparatus; (iii) the failure of the LDL receptor to bind LDL; or (iv) failure of recycling of LDL receptors. These diseases are usually familial.

There are defects of translation of the proteins of **collagen**, resulting in structural weaknesses of collagen, an important supporting tissue in the body.

19 Collagen

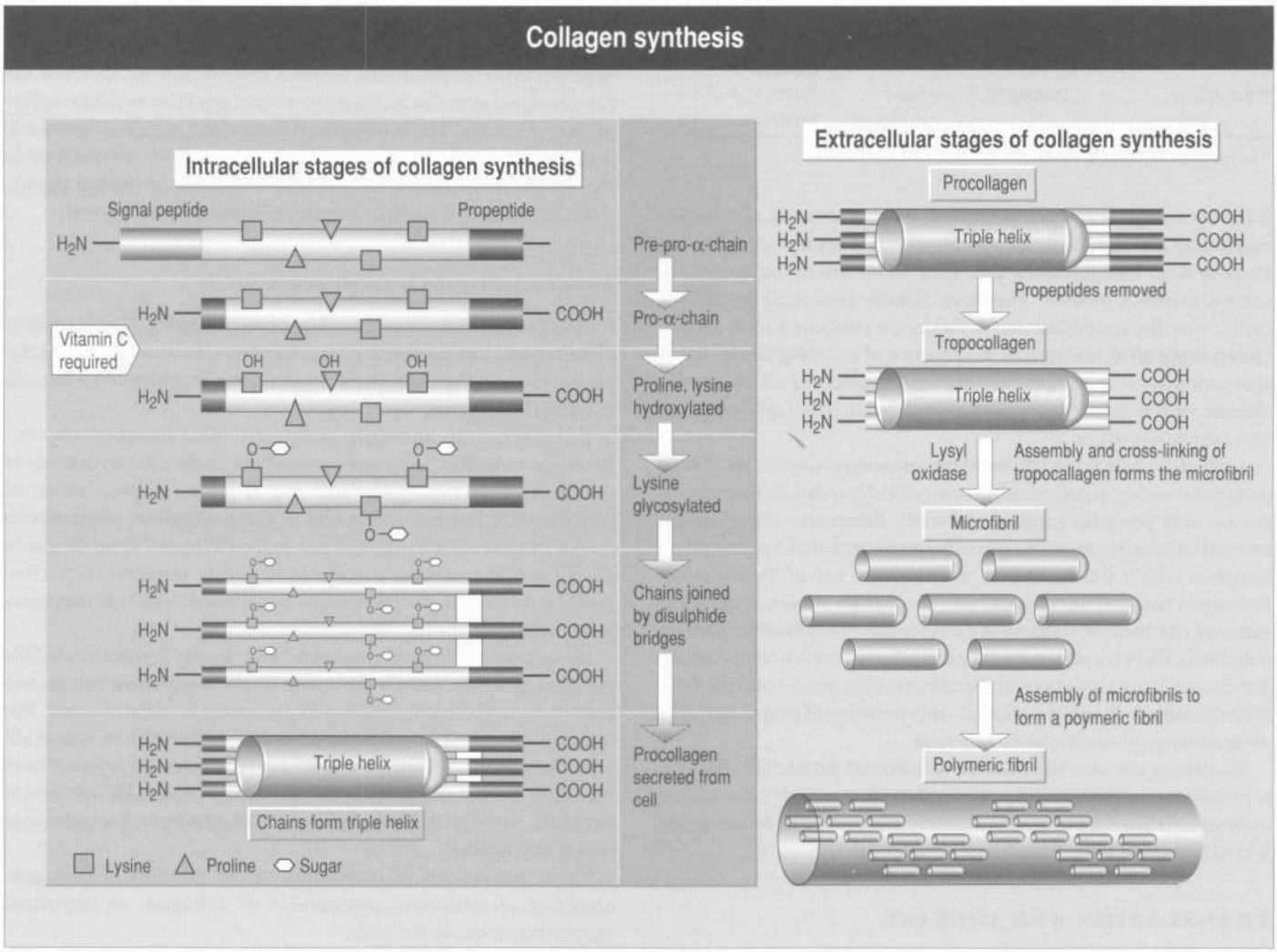

Fig. 19.1

Collagen is used in this chapter as an example of the events in protein synthesis described in previous chapters.

NATURE OF COLLAGEN

Occurrence and structure

Collagen is a fibrous, secretory protein, the most abundant protein in the human body, and occurs in all tissues that demand a framework or support in order to give them structural strength and retention of their characteristic shape or form. Other examples of fibrous proteins are tropomyosin and α-keratin. Structurally, fibrous proteins possess a high proportion of regular secondary structure and a rod-like cylindrical shape, and are relatively insoluble in H_2O.

Each collagen peptide, referred to as an α-chain, occurs as a left-handed helical polypeptide, in which every third residue is a glycine, and is about 1000 residues long. Three α-chains intertwine to form a right-handed triple helix, and the glycine residues are at the centre of the helix. This helical structure is termed **tropocollagen**, and it is the fundamental building block or repeating unit of collagen.

Tropocollagen

Tropocollagen can be expressed in terms of an approximate structural formula: $(Gly–A–B)_{333}$, where proline occupies approximately one-third of the A positions, and about one-third of the B positions are occupied by hydroxyproline. Tropocollagen molecules spontaneously combine through a crosslinking between lysine and hydroxylysine residues, to form the so-called polymeric fibrils, which can be seen under the electron microscope, and these aggregate to form the light microscope-visible polymeric collagen fibres.

The rigidity of collagen is due to the presence of proline residues. The strength of collagen is provided by the fact that tropocollagen molecules are adjacent to each other along approximately 75% of the length of each molecule. Because of the way they are attached, they also resist stretching, or tensing, and will rupture if subjected to excessive tensile forces.

Types of collagen

Different tissues possess different types of collagen, of which there are at least 13, called type I, II, etc., up to XIII. Collagen types differ in terms of the α-chains, which are termed α1 and α2. Type I, the first to be characterized, is present in the largest abundance in the human body.

SYNTHESIS OF COLLAGEN

Collagen is synthesized by several different specialized cell types, for example by the osteoblasts in bone, fibroblasts in the tendons and by chondroblasts in cartilage. Synthesis may be thought of in terms of intracellular and extracellular events.

Intracellular events

In the cell, the first step is the synthesis of an α-chain by the ER-attached ribosomes, which translate the α-chain sequence from the mRNA as well as the propeptide sequences, and the N-terminal signal peptide, which directs the newly formed protein into the lumen of the ER. The propeptide sequences ensure that the precursors remain soluble, as they do not form a helix. In the lumen, the enzyme signal peptidase removes the signal sequence, and the pro-α-chain moves along the smooth ER and the Golgi apparatus towards the plasma membrane.

During this voyage, –OH groups are added to proline and lysyl residues which use **vitamin C** as a cofactor. The sugar residues glucose and galactose are added to hydroxylysine; the extent of glycosylation determining the thickness of the resultant collagen fibrils. The sugar residues reduce the degree of packing of the tropocollagen microfibrils into polymeric fibrils. The extent of the glycosylation depends on the tissue in which collagen is formed.

During the late stages of glycosylation, three pro-α-chains are formed into a unit by thiol (SH) bonds between the propeptides at the C terminal, and the linked chains coil into the characteristic triple helix of tropocollagen. The Golgi apparatus packages the soluble precursor (procollagen) into vesicles, which are secreted from the cell by pinocytosis.

Extracellular events

Once outside the cell, the fibre-forming collagens (types I, II and III) lose the propeptides by enzymatic cleavage to yield tropocollagen. (Type IV tropocollagen does not lose the propeptides.) These molecules spontaneously aggregate to form fibrils, and as the polypeptide chains build up a network of crosslinkages, so the strength of the collagen fibres increases.

PATHOPHYSIOLOGY

Scurvy

Scurvy is the result of a deficiency of vitamin C (ascorbic acid), whose lack in the diet causes a decrease in the synthesis of hydroxyproline. (Note, unlike many other animals, humans lack the enzymes required to synthesize ascorbic acid from glucose.) Hydroxyproline contributes additional hydrogen bonding for stabilization of collagen helices. As a result, collagen loses its stability at body temperature, and structures cannot adhere to connective tissues. The consequences include suppression of growth in children, capillary fragility and delayed and deficient wound healing. Teeth become dislodged from the gums, and sudden death can result in extreme cases if the patient changes posture.

Other disorders

There are, in addition, a number of genetic disorders that result in:
1 failure of collagen fibrils to crosslink, for example, **osteogenesis imperfecta**, where point mutations in the α-chain prevent helix formation, and enzymatic destruction of the tropocollagen molecules occurs; and
2 deficiency of certain enzymes, such as lysylhydroxylase, which is called **Ehlers–Danlos syndrome type VI**. Here, synthesis of hydroxylysine is suppressed.

These diseases are characterized by poor wound healing, multiple fractures and hyperextensible skin and joints.

20 Control of gene expression in prokaryotes

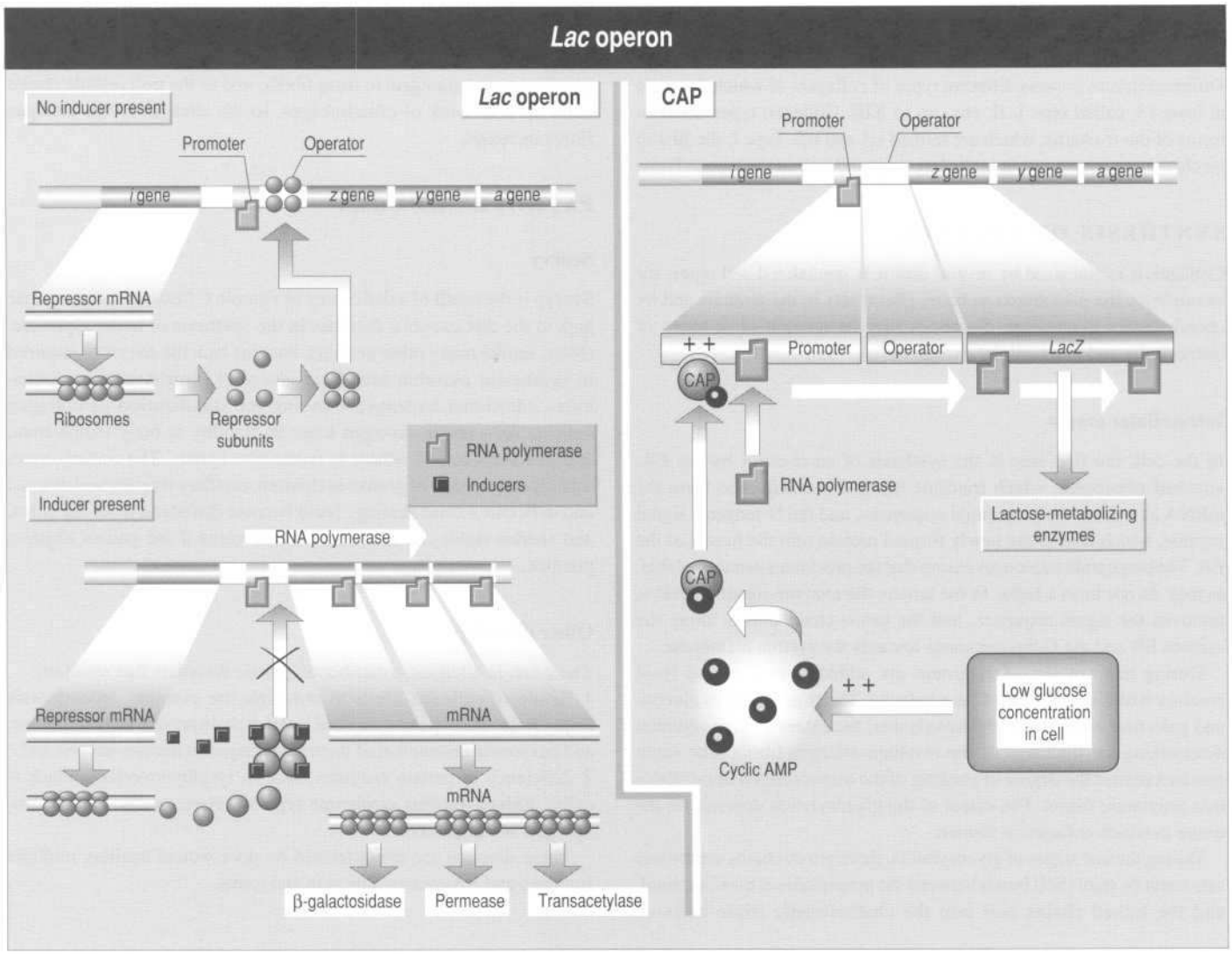

Fig. 20.1

PROKARYOTES

E. coli is a bacterium in which gene expression has been extensively studied and characterized. *E. coli* has a single, circular chromosome, composed of a double-stranded DNA molecule of about 4×10^6 base pairs. There are about 3000 genes, which are clustered according to function. For example, the genes coding for enzymes of a particular metabolic pathway are clustered, as are those coding for structural proteins. These clustered genes are usually under co-ordinate control, and are transcribed together to form a single strand of mRNA that codes for several different proteins. Such an mRNA strand is termed **polycistronic mRNA**, and a complete set of functionally clustered genes, together with their operator and regulatory genes, is termed an **operon**.

Regulatory genes code for proteins that in turn control expression of the genes by binding to **control elements** at sites on the DNA near to the structural gene. Regulatory proteins control the degree of access that the enzyme **RNA polymerase** has to its binding site on the gene. Two types of regulatory protein have been found: (i) negatively acting, which repress the operon by binding to the operator; and (ii) positively acting, which enhance the affinity of RNA polymerase for its binding sites on the gene. A good example of an operon is the lactose operon of *E. coli*.

THE LACTOSE OPERON

Expression of the operon is regulated by an **inducer** (lactose), and by a **repressor protein**, expressed by the ***i* gene**. The *i* gene (also called *lacI*), is situated just before the controlling elements for the cluster of

genes coding for three enzymes, at least two of which are important in the splitting of the disaccharide lactose into galactose and glucose. These three genes, called the *lacZYA* cluster, code for: (i) **β-galactosidase**, which acts on lactose; (ii) **β-galactoside permease**, a membrane-bound protein which forms part of the transport system for taking lactose into the cell; and (iii) **β-galactoside transacetylase**, whose precise function is unknown, but whose expression is essential for the metabolism of lactose. The mRNA transcribed by the *lac* operon is extremely unstable, having a half-life of approximately 3 min, which means that expression of the operon can change rapidly. As soon as inducer concentrations fall, expression of the gene ceases.

The *lac* repressor

The *i* gene codes for a repressor regulatory protein, called the ***lac* repressor**, and the *i* gene itself is not regulated but continues to produce the repressor at a low level, independent of other cellular events. The *lac* repressor is expressed as a monomer of 360 amino acids, which are associated to form a tetramer, and there are usually about 10 tetramers in the cell at any one time. The tetramer binds with high affinity to a specific DNA sequence situated between the promoter called ***lacP***, and the operator, called ***lacO***, for the *z* gene. This binding reaction blocks the binding of RNA polymerase to the promoter. The operon is said to be **repressed**.

Derepression of the operon

Lactose induces or derepresses the operon by binding to specific high-affinity sites on the tetramer subunits. The binding reaction causes an allosteric change in the tetramer, which drastically lowers its affinity for the DNA sequence to which it usually binds.

Lactose is not the only inducer. A number of so-called gratuitous inducers have been found, including isopropylthiogalactose, which bind to the tetramer but are not themselves metabolized by β-galactosidase, and are therefore useful in the study of the *lac* operon. The mechanism described above is an example of **negative control**. But, lactose metabolism can also be under **positive control**.

POSITIVE CONTROL

E. coli prefers glucose to lactose as an energy substrate; if there is plenty of glucose in the cell, the *lac* operon is repressed, even if there is plenty of lactose present. This is known as **catabolite repression**, since it happens only when glucose is being metabolized.

The cell will turn to lactose as a substrate only when glucose concentrations fall. When concentrations of glucose are high, those of the second messenger cAMP are low. When glucose concentrations fall, concentrations of cAMP in the cell rise, and cAMP binds to a protein called **catabolite activator protein** (CAP), which is an allosteric protein. CAP undergoes a conformational change as a result of the binding reaction. This enables CAP to bind to the promoter just before the RNA polymerase binding site, and this in turn facilitates the binding of RNA polymerase to the promoter. CAP synthesis is regulated by a gene that is not a component of the *lac* operon.

CAP is an activator of a number of other genes, including the galactose and arabinose operons, and is probably a co-ordinator for the general control of enzyme synthesis whose activity is unwanted when high concentrations of the preferred energy substrate (glucose) are high.

21 Control of gene expression in eukaryotes

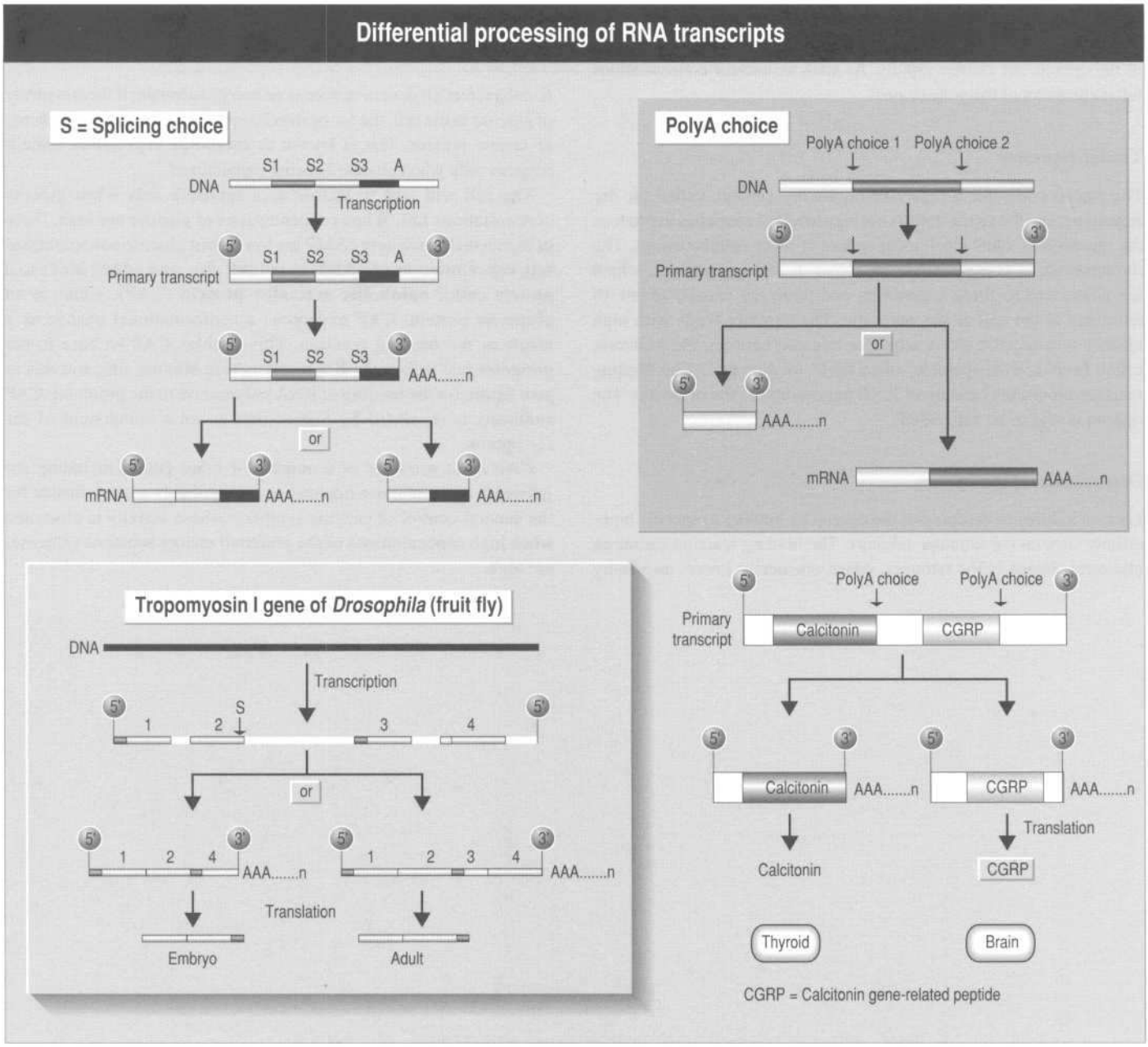

Fig. 21.1

COMPONENTS OF GENE CONTROL

Gene control has three main components: (i) signals; (ii) levels; and (iii) mechanisms. Signals include hormones, protein factors and environmental conditions such as heat shock.

Levels of regulation

Control is effected at three levels: (i) nuclear RNA synthesis; (ii) differential processing of primary transcripts; and (iii) altering mRNA stability in the cytoplasm. Control of nuclear RNA synthesis is effected mainly at the **initiation** stage. Initiation is activated by **transcription factors (activators)**, which may interact with genomic **promoters**, in order to guide RNA polymerase II to the correct site for expression of an mRNA species.

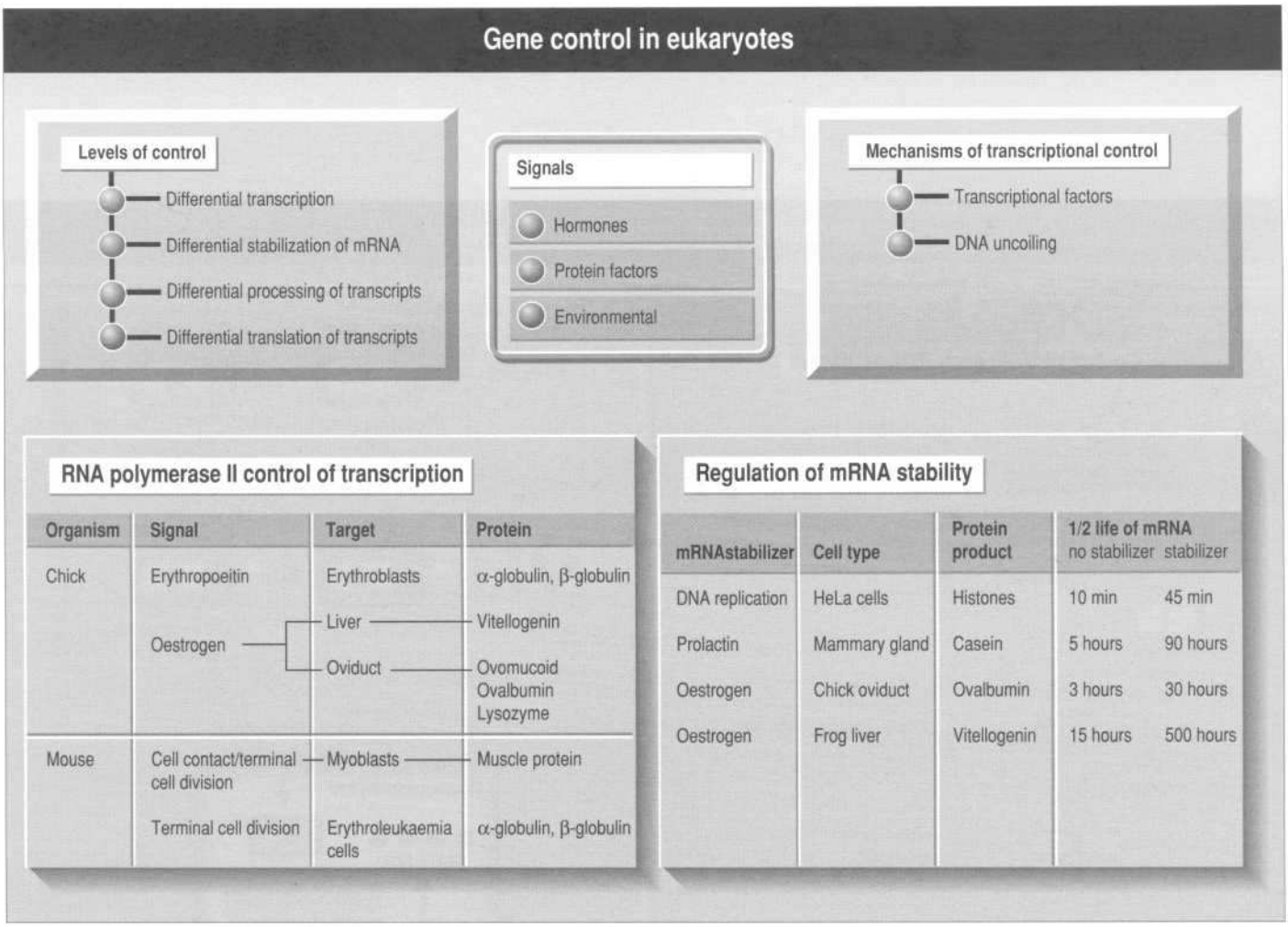

Fig. 21.2

Nuclear RNA synthesis

Cell or organism	Transcription factor	Genomic target	Cell or organism	Transcription factor	Genomic target
Mammalian cells	Sp1	Promoters containing GC boxes, e.g. dihydrofolate reductase promoter	Glucocorticoid-responsive cells	Glucocorticoid receptor	5′-GGTACA*nnn*TGTTCT-3′ consensus sequence (*n*, any nucleotide); genes coding for, e.g. chicken lysozome
	CTF	CAAT box			
Drosophila	B protein	TATA box	Yeast	gal4 protein	Specific upstream sequences in promoters of genes expressing galactose-metabolizing enzymes
	HSTF	Promoter of heat shock genes			
B lymphocytes	Immunoglobulin enhancer	Enhancer sequence for immunoglobulin gene expression			

CTF, CCAAT-binding transcription factor; HSTF, heat shock transcription factor.

Differential nuclear processing of transcripts. Differential choice of **polyA sites** on the primary transcripts determines tissue specificity of gene expression. For example, in rats, the transcript encoding the hormone **calcitonin** also codes for a brain peptide, calcitonin gene-related peptide (**CGRP**), and in the thyroid gland the cells involved produce calcitonin, while in the brain CGRP is produced.

Cytoplasmic gene control. The rate of protein synthesis may be affected by: (i) the rate of transport of mRNA into the cytosplasm; (ii) half-life of mRNA; (iii) frequency of mRNA translation; and (iv) post-translational control. Control of **half-life of mRNA** may be enhanced by hormone, the occurrence of DNA replication, tissue regeneration (liver) and by certain viral proteins.

22 Mechanisms of transcriptional control

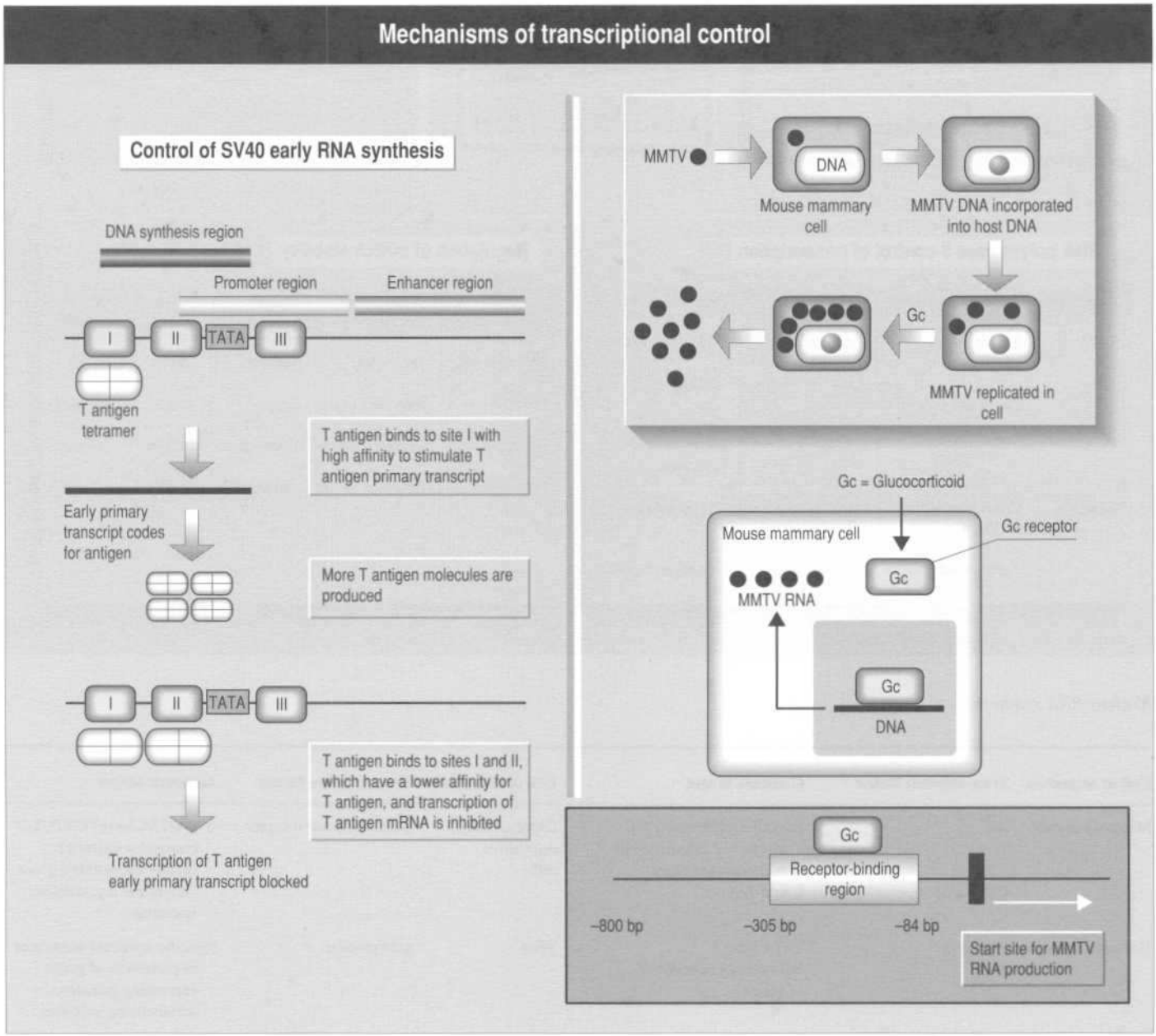

Fig. 22.1

INTRODUCTION

This is a brief, but more detailed look at some of the mechanisms underlying the processes whereby transcription is controlled. Examples used are the following.

1 The role of the T antigen in the control of transcription in the SV40 virus, which is a papovavirus. These are a family of small, non-enveloped double-stranded DNA viruses, including the papilloma viruses which produce the common wart, some of which can be oncogenic.

2 The role of glucocorticoids in the control of transcription of the MMTV.

SV40 VIRUS

When SV40 virus infects cells, its DNA is transcribed by host RNA polymerase II. SV40 DNA contains two transcription units: one termed 'early', and the other termed 'late'. This is because the early unit is preferentially transcribed soon after infection, and the late unit is

preferentially transcribed later during infection. The shift from early to late stages is mediated by a protein called T antigen, which is produced during early transcription. T antigen, which functions as a tetramer, has three binding sites on SV40 DNA: I, II and III. Binding to I may increase the affinity of T antigen for sites II and III, thus inhibiting further transcription of the early primary transcript. T antigen may therefore be an autoregulatory protein, which inhibits the synthesis of its own mRNA. T antigen was the first of the proteins to be discovered in eukaryotes that binds to specific sites on the DNA.

MMTV

Steroids are lipophilic substances which pass easily through the plasma membrane and combine with specific receptor proteins. These complexes bind to specific sites on the DNA to alter transcription. For example, glucocorticoids are substances that bind to intracellular receptors. Glucocorticoids are known to stimulate the production of increased numbers of MMTV molecules in cells infected with the retrovirus.

Note: retroviruses are enveloped, single-stranded RNA viruses, including human immunodeficiency virus (HIV), which infect cells, are converted into DNA by the enzyme reverse transcriptase. The DNA is incorporated into a host chromosome and may preferentially be transcribed back into the virus or remain dormant for several cell generations.

23 Growth

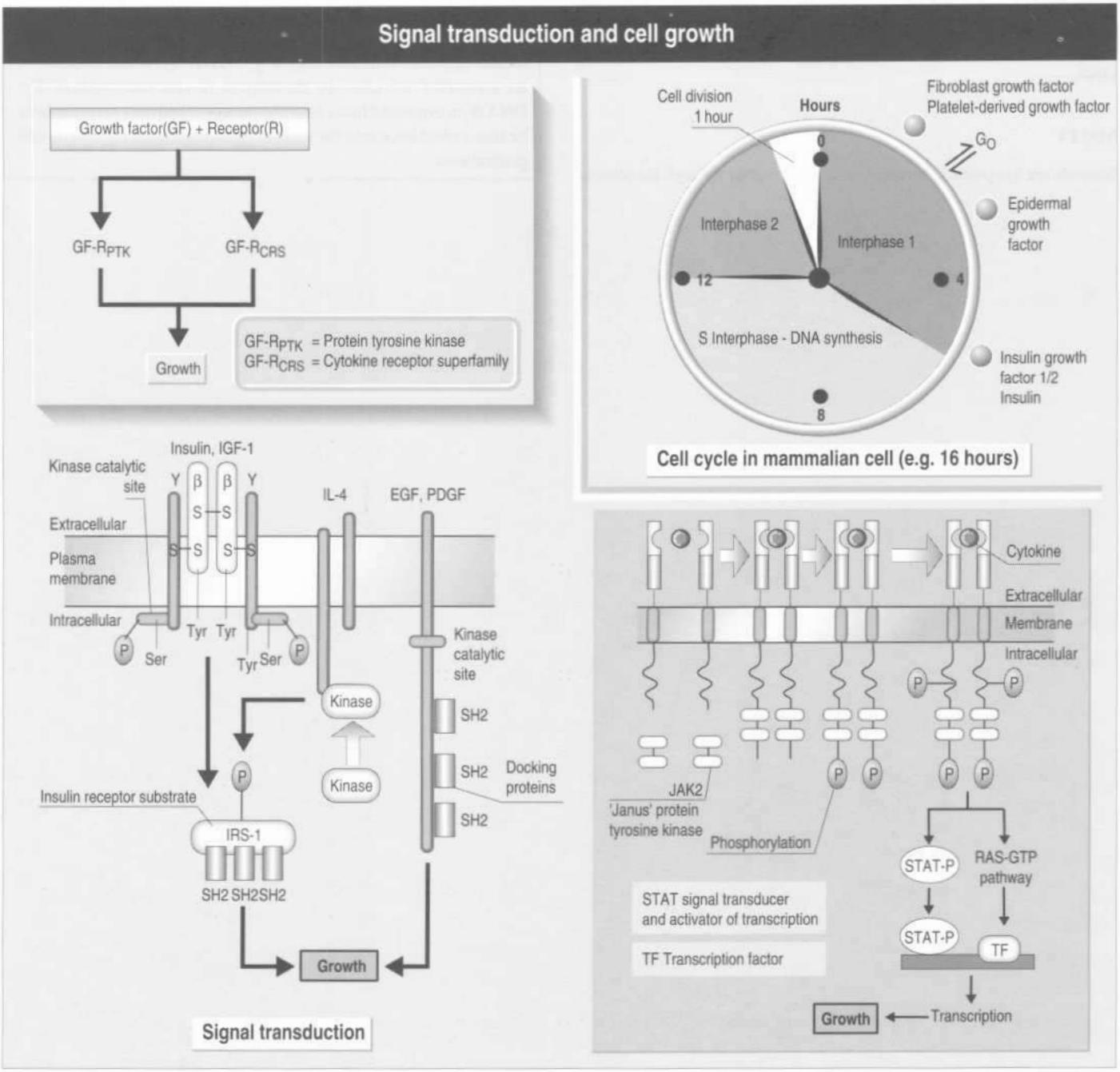

Fig. 23.1

THE CELL CYCLE

The eukaryotic cell has a life cycle characterized by four distinct successive phases, called G_1, S, G_2 and M. G stands for gap, S for synthesis and M for mitosis. DNA is synthesized during the S phase, while RNA and protein are synthesized during G_1, S and G_2. The S, G_2 and the M phases in different growing cells remain relatively constant in duration (7, 3 and 1 h, respectively). During mitosis, RNA synthesis ceases and protein synthesis is greatly reduced.

Mitosis is a typical process of nuclear division, when two daughter nuclei are formed, each having the same chromosomal complement as the parent nucleus. **Cell** division occurs after a full cycle has been

Growth factor or hormone*	Main sources	Target tissue or cell
EGF	Mouse submaxillary gland	Epidermal cells, fibroblasts
Transforming growth factor α (TGF-α)	Embryonic and cancerous cells	e.g. fibroblasts
TGF–β	Most cells, cancer cells	e.g. fibroblasts
Erythropoeitin	Kidney	Red blood cell precursors (erythroblasts)
Insulin	Pancreatic islet	Liver, muscle
Nerve growth factor (NGF)	Many cells	Sympathetic nerves
PDGF	Platelets	Arterial smooth muscle cells (repair function)
IGF–II; somatomedin A	Liver	Mediates GH action on growing bone; mitogenic in some cells
GH	Anterior pituitary	Liver–stimulates somatomedin production

*This list is not comprehensive.

completed. Some cells, such as muscle and nerve, may never divide after formation, while others, such as liver, skin or gut, will be active in cell division. When the cell enters a phase called G_0, it is dormant until a trigger (usually a growth factor) switches it into active cell division. Certain growth factors, notably platelet–derived growth factor (PDGF), fibroblast growth factor (FGF), epidermal growth factor (EGF), and insulin-like growth factor (IGF–1) and IGF–2 (see below), stimulate the cell cycle at various points during G_1.

GROWTH FACTORS

Growth factors are peptides that stimulate cellular proliferation. They regulate normal growth and development. They may also cause what is known as phenotypic transformation of cells, which may result in cancer. They produce these effects through an interaction with a specific receptor on the cell surface. They fall into several arbitrarily classified groups.

1 Hormones, such as **insulin,** secreted by the B cells of the pancreas, **growth hormone** (GH) and **prolactin**, both of which are secreted by the anterior pituitary gland.

2 Cytokines, which are not necessarily hormones, and which are made by many different cells. They affect growth and division of cells.

3 Lymphokines, which are polypeptides released by activated macrophages and T cells of the immune system. Some influence white cell migration, while others also act as growth factors. Interleukin 1 (IL-1), which is released by macrophages, stimulates proliferation and differentiation of the B lymphocytes.

Mechanicsm of action of growth factors

Growth factors exert their effects on cells through **membrane receptors**. When the factor binds to the receptor, the signal is transferred to the cell in the form of a cascade of several reactions, especially phosphorylation reactions. In certain cases, the intracellular domain of the receptor contains a protein kinase that enables the receptor to autophosphorylate itself (see below). The final result is initiation of transcription, protein synthesis and growth.

The growth factor receptors can be classified in terms of the intracellular phosphorylation cascade.

Insulin receptor substrate 1 (IRS-1) signalling system. Some hormones and growth factors, for example insulin and IGF-I, bind to receptors that contain intrinsic tyrosine kinase activity and autophosphorylate themselves. This results in the tyrosine phosphorylation of a protein called IRS-1. This enables a group of so-called SH2 proteins (proteins with *src* homology) to 'dock' at IRS-1, resulting in consequent intracellular events culminating in the mitogenic response. Some growth factors, e.g. **IL-4**, bind to a receptor that does not contain intrinsic kinase activity, but which recruits a cytoplasmic tyrosine kinase that phosphorylates IRS-1 on tyrosine to create the SH2 binding sites.

> *Note*: the insulin receptor consists of two α-chains, each of which binds a molecule of insulin, and two β-chains., which span the membrane and contain the catalytic sites. The chains are held together by disulphide bonds.

Receptor kinase–SH2 system. Certain growth factors, for example EGF and PDGF, bind to receptors that possess intrinsic kinase activity, which allows them to become autophosphorylated, and these phosphorylated sites become docking sites for SH2 proteins.

Cytokine receptor superfamily (CRS). This is a group of hormone and growth factor receptors that do not themselves have any kinase activity, but which activate a group of cytoplasmic protein kinase kinases, called the **Janus kinase (JAK)** family. At least four members of the JAK family have been identified: JAK1, JAK2, JAK3 and TYK2, and there are probably more. The JAK proteins bind to the membrane proximal area of the activated receptor, become tyrosine phosphorylated and in turn they phosphorylate the receptor causing the activation of the following.

1 A group of cytoplasmic proteins that form part of a family of transcription activators, called **signal transducers and transcription activators (STATS).**

2 The phosphorylated STATS activate transcription at the level of the DNA — the **Ras** pathway. The phosphorylated receptor provides binding sites for a group of SH2 docking proteins, which activates the Ras–GTP system, resulting in activation of a group of mitogen-activated protein kinase (MAPK), and a transcription activator TF.

Termination of action

Growth factor action will be terminated by the dissociation of the factor from its binding site on the receptor. It may also be terminated by the inactivation of the intracellular MAPK. There is evidence that the MAPK may be down-regulated by dephosphorylating enzyme phosphatases. Furthermore, it seems that when MAPK are activated this triggers the expression of the genes encoding the phosphatases, which would provide a tightly and elegantly controlled mechanism for the regulation of growth factor activity. Any disturbance of this control system might conceivably contribute to the factors that cause cancer through unrestrained growth factor action.

24 Cancer

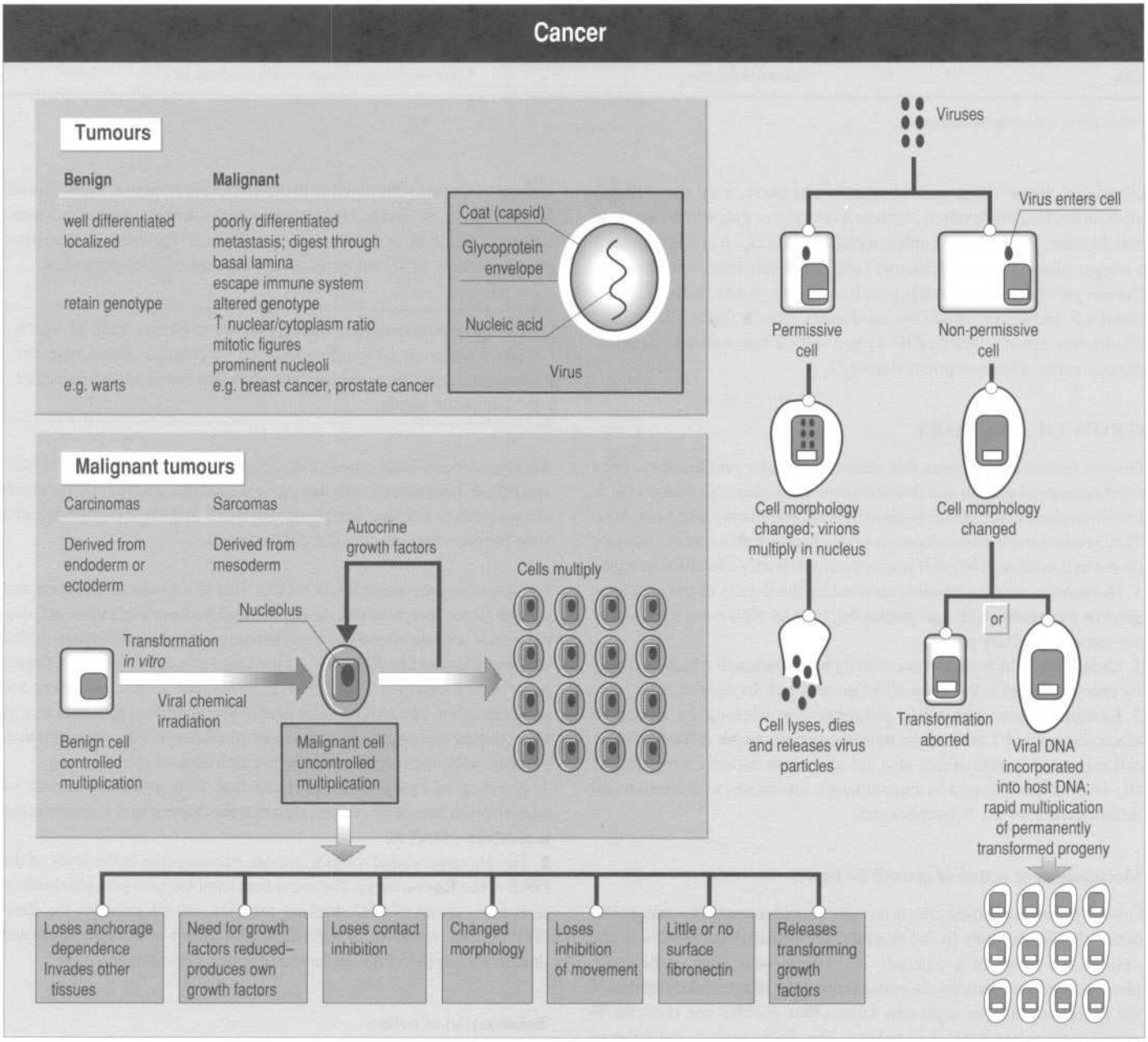

Fig. 24.1

Cancer is the uncontrolled, apparently autonomous growth of cells, and their invasion of the rest of the body. Such cells are termed **malignant**, and their invasion is termed **metastasis**. Normal cells become transformed through **mutagenic** agents.

MALIGNANCY

Knowledge of the molecular basis of cancer stems largely from the study of **viral** activation of **oncogenes**.

Cancer-producing viruses

Certain viruses transform healthy cells to malignancy. Viruses are

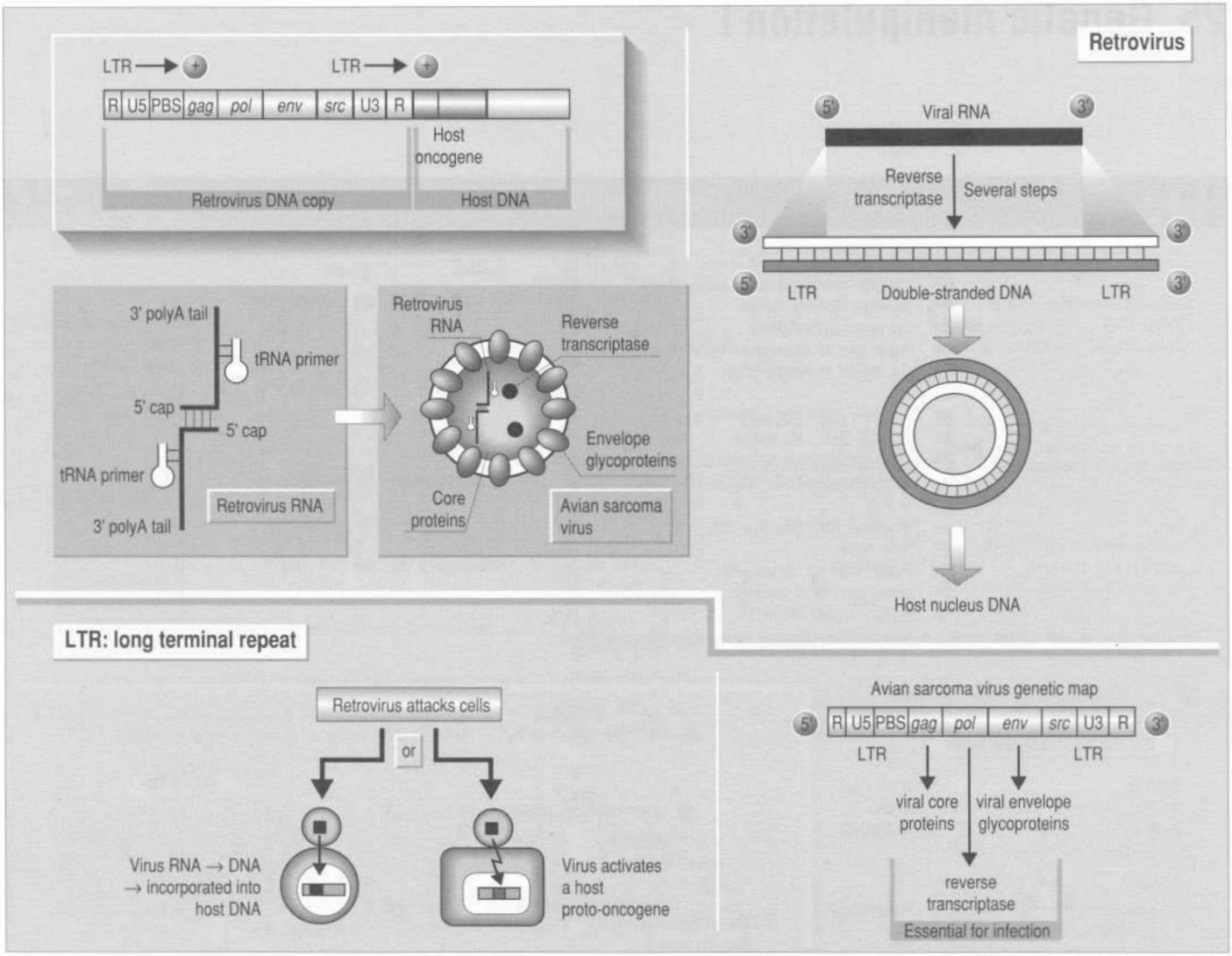

Fig. 24.2

infectious DNA or RNA, which may be single or double stranded, surrounded by a protective coat consisting of a large number of repeating protein subunits.

Viruses must invade other cells to multiply, and their own nucleic acid instructs the host (permissive) cell to synthesize more viral nucleic acid protein. This invasion usually kills the cell, and the complete **virion**, or virus particle, is released. Occasionally, the cell is non-permissive, and may destroy the invading virus, or incorporate the viral genetic information into its own genome.

Some viruses that contain RNA are termed **retroviruses**. They have a polymerase enzyme termed **reverse transcriptase**, which, when they have entered a cell, produces a double-stranded DNA copy which is incorporated into host DNA. The cell is transformed either because the viral DNA copy contains the oncogene, or because the viral genes activate a host oncogene.

The avian sarcoma retrovirus binds to specific cell membrane receptors, and inserts its contents into the cell. Reverse transcriptase produces a DNA copy which is longer than the viral RNA template because it contains **long terminal repeats** (**LTR**). The LTR contains the enhancers, polyadenylation sites and promoters. The viral DNA circularizes, enters the host nucleus, and is spliced into the host DNA at TCAG sites.

The 10 kbp of avian sarcoma virus contains four genes, three of which are necessary for infection. These are *gag*, *pol* and *env*. The other, *src*, is necessary for transformation to occur. Three viral mRNA species are expressed: one unspliced mRNA coding for both *gag* and *pol*; and two unspliced mRNAs encoding *env* and *src*. The *src* product is a tyrosine kinase. The viral proteins and the unspliced primary RNA transcript migrate to the cell membrane and are incorporated into it. Part of the altered membrane buds off a new viral particle. Retroviruses do not usually kill the host. The HIV is a retrovirus.

Oncogenes

Oncogenes are cancer-producing genes. They may be present on the viral genome, imported into the host genome or already present as a cellular gene which is activated through viral infection. Cellular oncogenes (also called **proto-oncogenes**) are usually silent, or expressed normally under cellular control. But, if a viral DNA copy is spliced into the cellular genome adjacent to a proto-oncogene, the viral LTR may stimulate its expression, thereby transforming the cell.

Oncogenes are named after the species and disease produced by the infective virus.

25 Genetic manipulation I

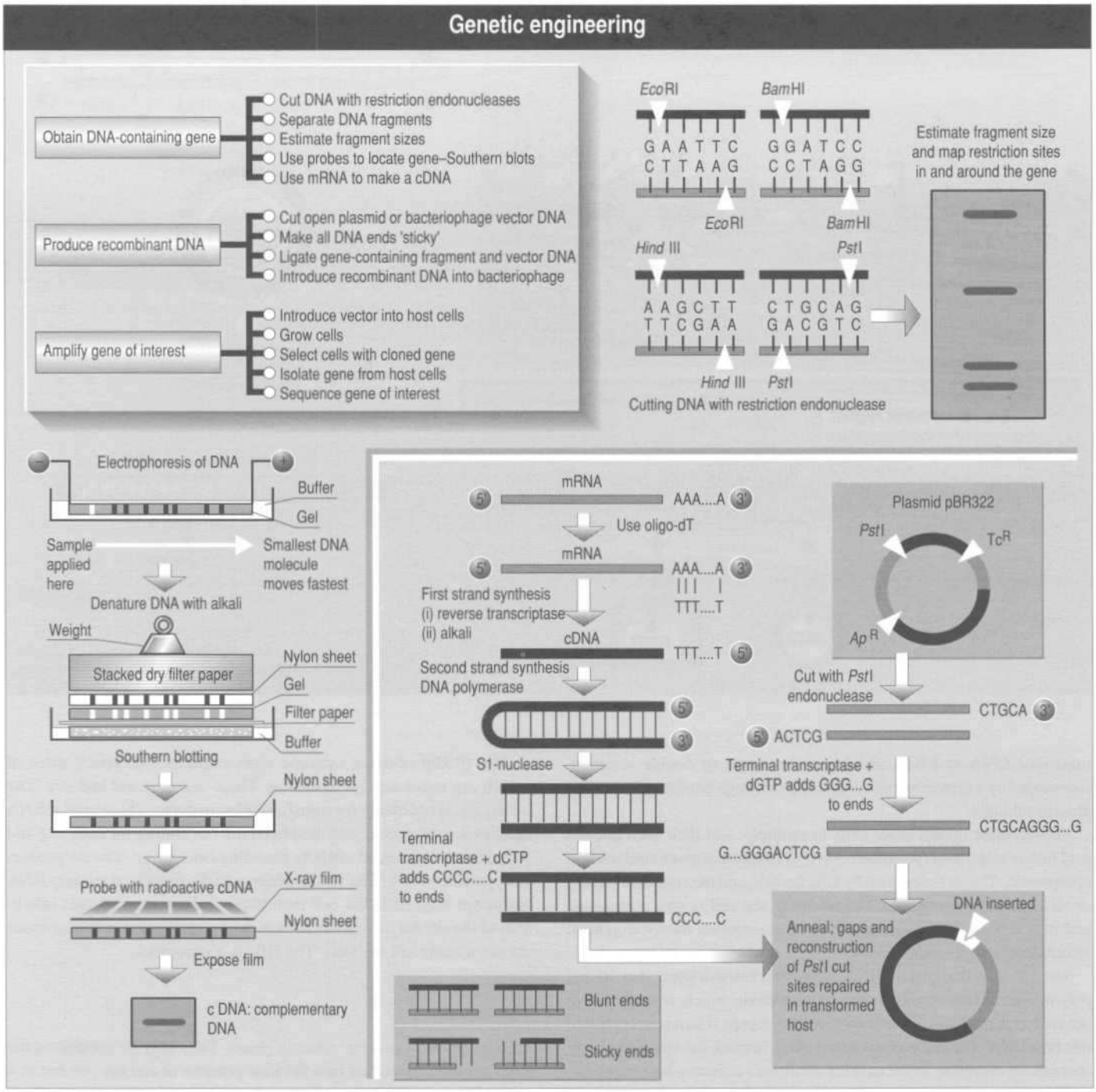

Fig. 25.1

INTRODUCTION

Genetic manipulation is the alteration of DNA, which is transmitted in a vector system into host organisms, in which the manipulated gene(s) are amplified and recovered for further use. The genes implanted into the host may cause it to express the gene products directed by the implanted gene. The vector may be a bacterial plasmid, a bacteriophage, or the DNA may be introduced directly into mammalian embryonic

tissues using surgical techniques.

Techniques involve the introduction of the chosen DNA nucleotide sequence into a vector, the introduction of the vector into a host in which the sequence will be amplified, the cloning of the altered host and the recovery of the amplified sequence.

CREATING THE DNA SEQUENCE

DNA from mRNA

The production of a mixture of **cDNA** from small quantities of mRNA is made possible using the enzyme **reverse transcriptase**, obtained from retroviruses. Double-stranded DNA for insertion into a vector can be prepared from the mRNA expressed by the gene. The **procedure** involves the following.

1 The mRNA is isolated and given an oligo (d)T primer, after which it is incubated with reverse transcriptase and the four deoxynucleotides to make the first cDNA strand.

2 The first strand has at its 3′ end a hairpin bend which is used, together with a primer called the **Klenow fragment** and DNA polymerase to make a second DNA strand complementary to the first, after which the bend is removed with an enzyme, S1 nuclease.

3 The 3′-terminal ends of the DNA are given homopolymeric tracts of CCC..., etc., which combine non-covalently with tracts of GGG..., etc., added to ends of a cut plasmid. After the plasmid with its new addition of the sequence of interest is annealed, it is introduced into the host.

It is possible to amplify the gene starting with minutely small quantities of mRNA. A disadvantage is that many thousands of different cDNA molecules will be cloned.

Restriction endonucleases

Cutting DNA at selected sites is possible through the use of restriction endonucleases. The physiological role of these enzymes is to destroy unwanted cellular DNA. Once DNA has been extracted, it can be cut with enzymes that recognize certain sequences, and which cut the tracts in predictable places. The enzymes share in common the ability to read palindromic sequences. (A pure **palindrome**, e.g. AATTAA, is identical when read from left to right or from right to left.) Some enzymes (e.g. AluI) cut the DNA leaving 'blunt' ends. This is not ideal, since a meeting between ends to be joined depends purely on chance. Some enzymes (e.g. *Eco*RI) make 'sticky' ends, with sequences extending from the chain, allowing complementary base pairing with another chain.

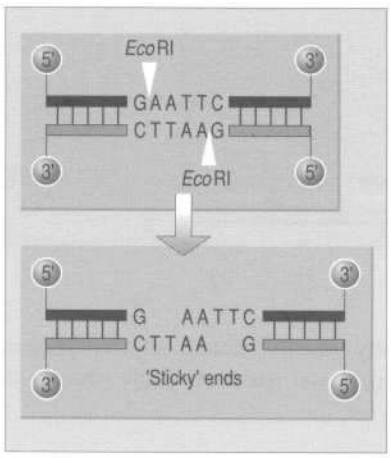

Fig. 25.2

SEPARATION OF DNA FRAGMENTS

Restriction endonucleases will cut the DNA into several fragments of differing sizes. These are separated by gel electrophoresis. The negatively charged DNA fragments migrate to the positive electrodes, with the smallest moving fastest. In order to estimate the sizes, a separate lane is run containing standard DNA fragments of known size. After electrophoresis, the bands of separated DNA are visualized by staining the gel with ethidium bromide, which intercalates in the DNA and fluoresces under UV light. The gel can be photographed. This enables the estimation of the number of fragments cut. As little as 25 ng of DNA can be seen this way. If the starting DNA material is radioactively labelled with phosphorus-32 (^{32}P), 1–2 ng of DNA can be visualized using X-ray film. The size of the band is obtained by comparison with the migration rate of the standards.

Once the size and number of fragments are known, and the sites where they have been cut, it is possible to start constructing a map of the DNA molecule — so-called gene maps.

SOUTHERN BLOTTING AND GENE PROBING

The technique of blotting involves transferring the separated DNA bands from the gel to a nitrocellulose or nylon sheet, when the DNA bands are transferred by capillary action. Once on the sheet, the DNA is fixed to it by heating or chemical reaction, and the bands can be probed using a radioactively labelled gene probe, whose DNA sequence is complementary to that of the gene to be probed for. The probe hybridizes to the gene, the unbound probe is washed away and the hybridized probe visualized. The experiment confirms that a gene has indeed been isolated. The DNA of interest can be recovered and inserted into a vector.

RNA can be electrophoresed instead of DNA, and transferred to a nitrocellulose or nylon sheet (**Northern blotting**); the separated bands of RNA are probed with labelled cDNA complementary to the mRNA expressed. This confirms that a particular gene has been expressed, and that its mRNA is among the species of mRNA electrophoresed. Similarly, proteins can be transferred and probed using immunocytochemical techniques, when the process is termed **Western blotting**.

POLYMERASE CHAIN REACTION (PCR)

PCR enables us to amplify very rapidly a DNA sequence in a small biological sample. This allows the identification of potential disease-producing genes in a prenatal sample, the forensic identification of DNA for legal purposes and the amplification of sequences for insertion into vectors or for sequencing purposes.

Principle

The method was made possible by the discovery of organisms living near boiling geysers. Their enzymes, including DNA polymerase, operate at high temperatures. PCR depends on the use of this DNA polymerase, a supply of the four (d) nucleotides and primers which flank the DNA to be amplified, after it has been denatured to give single-stranded DNA: (i) the DNA is denatured at 90°C; (ii) the primers anneal to the sequence to be amplified at 50°C; (iii) the primer sequences are extended at 70°C; (iv) the cycle is repeated several times; and (v) the amplified DNA is recovered from the reaction mix. Theoretically, it is possible to create over 250 million copies after 30 cycles.

26 Genetic manipulation II

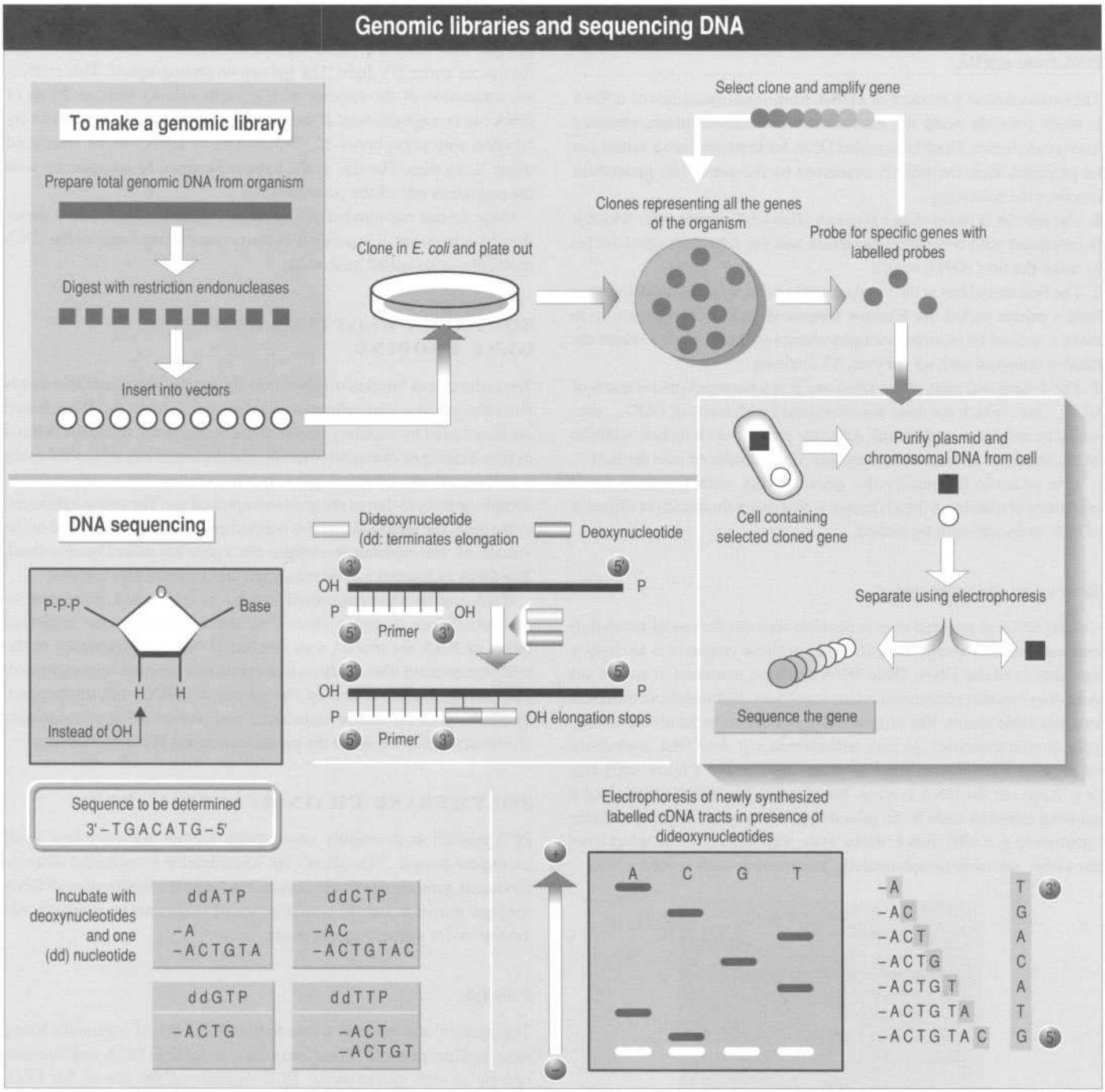

Fig. 26.1

VECTORS

A vector must be able to enter a cell, be stable and able to replicate within the cell. The ideal vector has a low molecular weight, enters host cells easily, readily confers selected phenotype characteristics on host cells and possesses single sites for many restriction endonucleases.

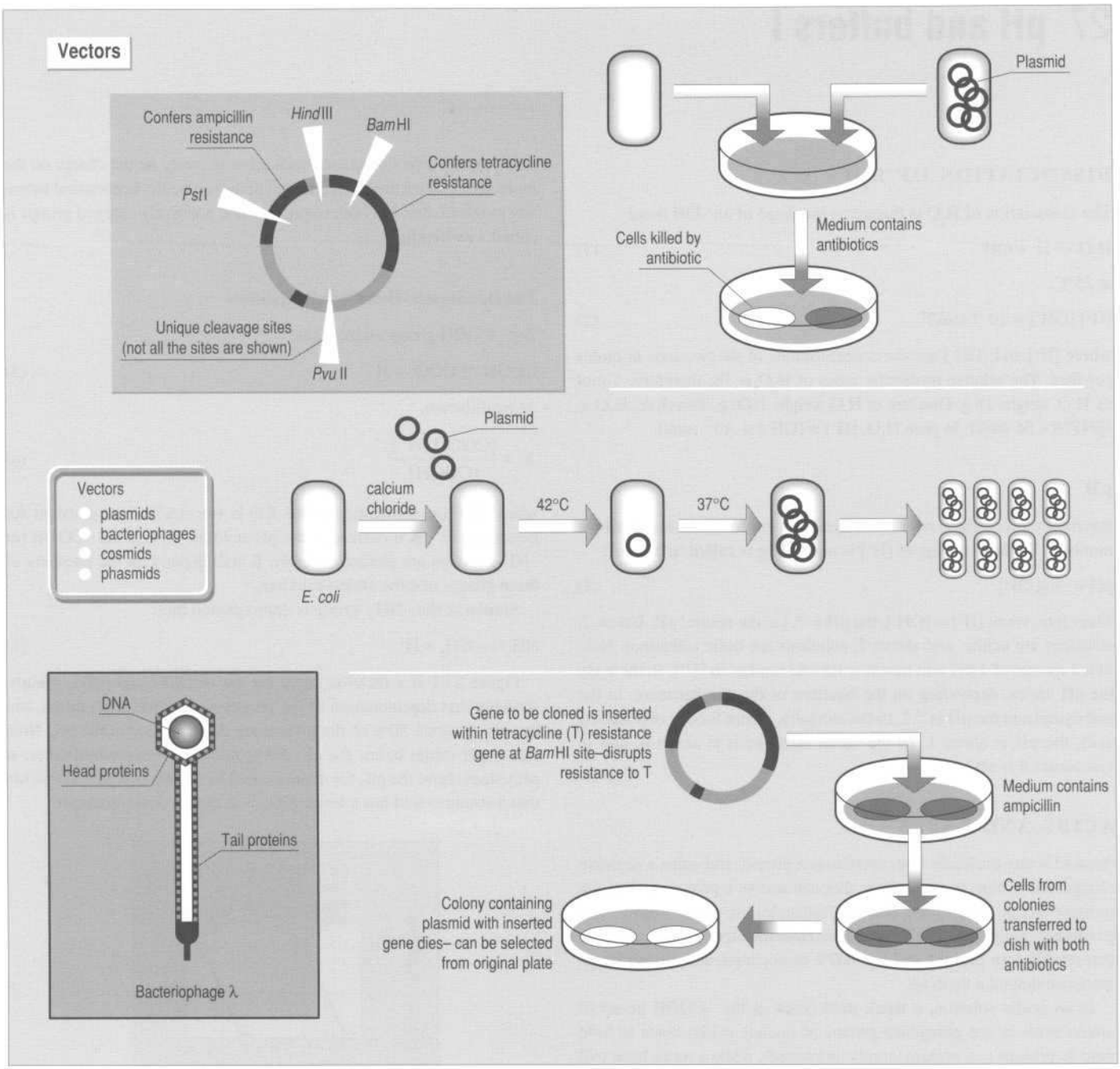

Fig. 26.2

Plasmids

Plasmids are small circular DNA, occurring naturally in bacteria. Inside bacteria, plasmids replicate independently of the chromosomal DNA. Plasmids enter bacterial cells more readily if the cells are first treated with $CaCl_2$ at low temperatures, or exposed to an electric field ('electroporation'). $CaCl_2$ causes the plasmid to bind to the cell membrane, and if the cell is heated briefly to 40°C, the plasmid rapidly enters the cell. Transformed cells, i.e. those containing foreign DNA, will replicate in culture to form a **clone** of daughter cells. Not all the bacteria in a culture will take up the plasmid that contains the gene of interest; these have to be distinguished and separated from those that have taken up the gene. Plasmids, such as pBR322, have been designed to achieve this separation.

pBR322. Plasmid pBR322 is an artificial circular DNA molecule containing 4363 base pairs, and contains sequences which are the genes which confer resistance to ampicillin and tetracycline (genes Ap^R and Tc^R, respectively). The plasmid also has several sites for cleavage by restriction endonucleases. When the plasmid is cleaved, the size of every fragment produced can be calculated.

Other vectors

Bacteriophages are viruses that infect bacteria. The phage binds to the bacterial membrane and injects its nucleic acid into the cell, where it is replicated. The phage can be 'primed' with the recombinant DNA one wants to infect the cell with.

27 pH and buffers I

DISSOCIATION OF H_2O

The dissociation of H_2O is through a breakage of an –OH bond:

$$H_2O \rightleftharpoons H^+ + OH^- \quad (1)$$

at 25°C:

$$[H^+]\,[OH^-] = 10^{-14}\ \text{mol/l}^2 \quad (2)$$

where $[H^+]$ and $[OH^-]$ are the concentrations of the two ions in moles per litre. The relative molecular mass of H_2O is 18; therefore, 1 mol of H_2O weighs 18 g. One litre of H_2O weighs 1000 g. Therefore, H_2O is $1000/18 \approx 56$ mol/l. In pure H_2O, $[H^+] = [OH^-] = 10^{-7}$ mol/l.

pH

It is more convenient to refer to the hydrogen ion concentration in whole numbers. To do this, –log of $[H^+]$ is used (–log is called 'p'). Thus:

$$pH = -\log\,[H^+] \quad (3)$$

Therefore, when $[H^+] = [OH^-]$, the pH = 7, i.e. the neutral pH. Below 7, solutions are acidic, and above 7, solutions are basic (alkaline). Note that a change of 1 pH unit means a 10-fold change in $[H^+]$. In the body the pH varies, depending on the function of the compartment. In the cell cytoplasm the pH is 7.2. In the stomach, where food is digested by acid, the pH is about 1. In the small intestine it is about 8, and in lysosomes it is about 5.

ACIDS AND BASES

An acid is any molecule that can release a proton, and gains a negative charge, and a base is a substance that can accept a proton, and gain a positive charge. Negatively charged molecules are called **anions**, and positively charged molecules are called **cations**. **Strong acids** are acids that readily lose protons and are 100% dissociated, and **strong bases** are those that take them up.

In an acidic solution, a **weak acid** (such as the –COOH group of amino acids or the phosphate groups of nucleic acids) tends to hold onto its protons (i.e. remain largely unionized), while a **weak base** will take up protons (i.e. ionize). In a basic solution, weak acids will ionize, while weak bases are only partially ionized.

Nucleic acids ionize by releasing protons, and become negatively charged. Amino acids contain both basic ($-NH_2$) and acidic (–COOH) groups, and can become positively or negatively charged, or both.

Ionization of amino acids

Amino acids have –COOH groups which can release protons, and they have $-NH_2$ groups which can accept protons. Therefore, depending on the pH of the solution, amino acids can exist as weak acids or bases. Since proteins may have unequally balanced numbers of –COOH and $-NH_2$ groups, changes in pH of the solution in which they are dissolved will cause changes in the ratio of charged acidic and basic groups. For different amino acids, the pH at which the –COOH and $-NH_2$ groups exactly balance each other to create no net charge on the molecule is called the isoelectric pH of the molecule. In chemical terms, any molecule that has both negatively and positively charged groups is called a **zwitterion**.

The Henderson–Hasselbalch equation

The –COOH group ionizes thus:

$$COOH \rightleftharpoons COO^- + H^+ \quad (5)$$

At equilibrium:

$$K = \frac{[COO^-]\,[H^+]}{[COOH]} \quad (6)$$

where **K** (also sometimes called Ka) is the equilibrium constant for Equation (5). **pK** is defined as the pH at which 50% of the –COOH (or $-NH_2$) groups are ionized. Clearly, K will depend on the numbers of these groups that the amino acid has.

Similarly, the $-NH_3^+$ group is deprotonated thus:

$$NH_3^+ \rightleftharpoons NH_2 + H^+ \quad (7)$$

Figure 27.1 is a titration curve for the –COOH and $-NH_2$ groups, showing that deprotonation of the groups occurs over a pH range, and the pH at which 50% of the groups are deprotonated is the pK. Note that at pH values below the pK, the protonated form predominates; at pH values above the pK, the deprotonated forms predominate. It follows that a stronger acid has a lower pK, i.e. it readily loses protons.

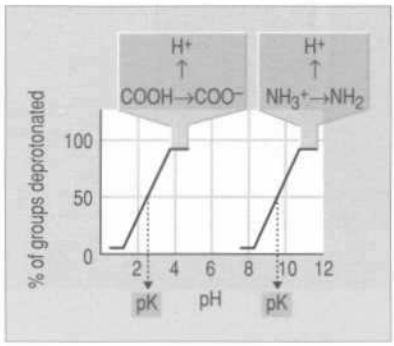

Fig. 27.1

Using Equation (6), we can derive one which enables us to predict the state of ionization of a given amino acid if we know K and the pH of the solution.

1 Rearrange and take the log of both sides

$$\log K = \log\,[H^+] + \log \frac{[COO^-]}{[COOH]} \quad (8)$$

2 Convert to –log and rearrange

$$-\log\,[H^+] = -\log\,[K] + \log \frac{[COO^-]}{[COOH]} \quad (9)$$

3 Express in terms of p(−log): the **Henderson–Hasselbalch equation**

$$pH = pK + \log \frac{[COO^-]}{[COOH]} \qquad (10)$$

Equation (10) allows us to predict, for example, the degree of ionization of the –COOH group of drugs for absorption through biological membranes, which are lipophilic and allow only the unionized form of the drug to pass through easily. For example, aspirin is a weak acid, with a pK of 3.5; the reader is invited to use Equation (10) to calculate the degree of ionization, i.e. the ratio of ionized to unionized groups in the stomach (pH 1.5), and in the small intestine (pH 8). From this result, the theoretical site of greater absorption may be predicted.

BUFFERS

Buffer solutions are those that resist a change in pH even when H^+ ions are added to, or removed from the solution. Thus, they protect the solutes within the buffer from sharp changes in pH that could, for example, inhibit a chemical reaction. In the absence of a buffering mechanism, the pH of a solution will change much more when acids or alkalis are added to the solution.

Mechanism of buffer action

The weak acid, acetic acid (CH_3COOH) (found in vinegar and bad wine), and its salt sodium acetate (CH_3COONa) provide an example of a buffering system. The acid has a pK of 4.75. A change of 2 pH units in the solution from 5.75 to 3.75 causes a change from about 10% CH_3COOH in the unionized form to about 90% unionized CH_3COOH. The ability of a weak acid and its salt to buffer a solution is greatest over the pH range pK −1 to pK +1. When CH_3COOH and CH_3COONa are present together in solution, they ionize as follows:

$$CH_3COOH \rightleftharpoons CH_3COO^- + H^+ \qquad (11)$$

$$CH_3COONa \rightleftharpoons CH_3COO^- + Na^+ \qquad (12)$$

Although the acid ionizes only partially, salts ionize virtually completely. Therefore, there will be a large concentration of CH_3COO^- and Na^+ ions in solution. The increased concentration of CH_3COO^- ions from the salt suppresses even further the ionization of CH_3COOH. If more H^+ ions are added to the solution, they will combine with CH_3COO^- ions to form even more of the largely undissociated CH_3COOH. A new equilibrium is established, and the resulting liberation of H^+ ions is relatively slight.

If OH^- ions are added, they will combine with H^+ ions to form neutral H_2O. Thus, at pH values close to its pK, a weak acid is a useful buffering agent when mixed with its salt.

The system will lose its buffering capacity sharply at pH values more than 1 pH value away from the pK. In a strongly basic solution the weak acid itself ionizes virtually completely, so cannot exist in the unionized form, and in strongly acidic solutions it cannot exist in the ionized form.

In cells, buffering of physiological fluids is achieved largely through the ionization of phosphoric acid (H_3PO_4), to form phosphates. H_3PO_4 can exist in three forms, depending on the pH of the solution:

$$pK \approx 2.1 \qquad H_3PO_4 \rightleftharpoons H_2PO_4^- + H^+ \qquad (13)$$

$$pK \approx 7.2 \qquad H_2PO_4^- \rightleftharpoons HPO_4^{2-} + H^+ \qquad (14)$$

$$pK \approx 12.7 \qquad HPO_4^{2-} \rightleftharpoons PO_4^{3-} + H^+ \qquad (15)$$

At cytoplasmic pH, phosphates can act as buffering systems.

28 pH and buffers II

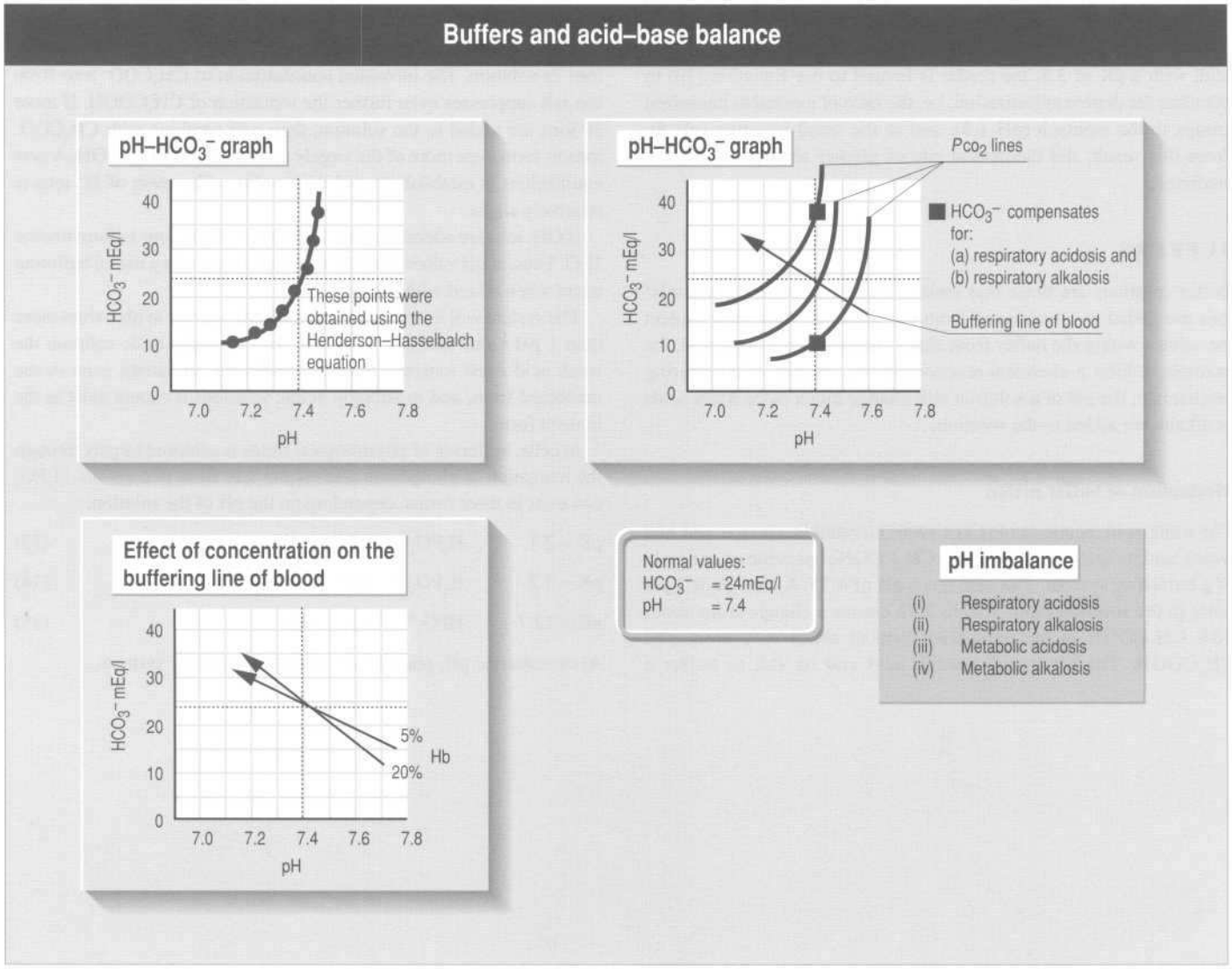

Fig. 28.1

PHYSIOLOGICAL BUFFER SYSTEMS

The major fluid compartments in the body are the intracellular fluid (ICF), and the extracellular fluid (ECF), which consists of the plasma and the interstitial fluid. All are bounded by semi-permeable membranes, whose properties depend on their function. All need buffering systems, which depend on the major ions of the compartment.

Plasma buffering system

In plasma and interstitial fluid, the CO_2–bicarbonate (HCO_3^-) system is very important. It prevents the development of dangerous acid or base imbalance, and works as follows:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \quad (1)$$

Carbonic acid (H_2CO_3) is a weak acid, and ionizes:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \quad (2)$$

H_2CO_3 ionizes so rapidly that for our purposes the reaction of importance can be considered to be:

$$CO_2 + H_2O \rightleftharpoons H^+ + HCO_3^- \quad (3)$$

Recalling the Henderson–Hasselbalch equation:

$$pH = pKa + \log \frac{[HCO_3^-]}{[CO_2]\,[H_2O]} \quad (4)$$

The pH in blood is 7.4 (actually ranges from 7.35 to 7.45); the pKa for

HCO_3^- is 6.1. Therefore, from the previous spread, we know that this is a buffering system that operates over the pH range of about 5–7. [H_2O] is taken as unity. The concentrations of gases in fluids such as plasma are expressed as partial pressures (e.g. the P_{CO_2} in plasma ranges from 4.5 to 6.1 kPa). To convert partial pressures into concentration terms, it is necessary to use a conversion factor. For CO_2, the correction factor, at 37°C, is 0.23 mEq/l per kPa. For our purposes, let P_{CO_2} = 5.0 kPa.

From Equation (4):

$$0.4 = 6.1 + \log \frac{[HCO_3^-]}{[0.23 \times 5.0]} \quad (5)$$

Therefore, under these conditions, $[HCO_3^-]$ = 22.94 mEq/l.

The reader may have noticed that HCO_3^- is not, theoretically, a good buffering system above pH 7.1. Yet, it is efficient at buffering plasma at pH values as high as 7.9. This is because the body eliminates CO_2 through respiration. In other words, undissociated weak acid is being eliminated, and so Equation (3) is driven to the left. This enhances the buffering capacity of the system.

If CO_2 built up in the plasma and other tissues, the body would suffer an acidosis, but CO_2 is lost through the lungs. The pH–HCO_3^- graph shows how HCO_3^- ion changes with pH at a given P_{CO_2}. When P_{CO_2} changes, the graph shifts in a line called the **blood buffering line**. The buffering action of HCO_3^- is supplemented by proteins and phosphates, which assist to 'mop up' H^+ ions as they are formed. The steeper the slope of the buffering line, the better the buffering action. The slope reflects the concentration of Hb in blood.

Hb is an important buffering agent in blood through its transporting of CO_2, and its ionization. Hb removes about 60% of H^+ ions produced through normal CO_2 transport. Therefore, in disease states involving a depletion of Hb, the buffering capacity of the blood will be reduced.

ACID–BASE BALANCE

The body generates acids through metabolism and respiration. The major respiratory acid is CO_2, and important metabolic acids are lactic acid, and the ketoacids β-hydroxybutyric acid and acetoacetic acid. In addition, acids may be taken in the form of drugs such as aspirin (acetylsalicylic acid). In the disease state diabetes mellitus (see below) excess ketoacids are produced.

Acid–base imbalance

There are four main types of imbalance.

1 Respiratory acidosis. Here, CO_2 is retained, either because of hypoventilation or intrinsic lung disease interfering with gas exchange. For example, hypoventilation may result from depression of the respiratory centre by drugs, while intrinsic lung disease could include conditions such as chronic bronchitis. In the latter condition, mucosal thickening and airway plugging with mucus may lead to poor alveolar ventilation and CO_2 retention with low arterial P_{O_2} values.

2 Respiratory alkalosis. This results from hyperventilation. CO_2 is blown off through the lungs too rapidly, and blood pH rises. Hyperventilation may be caused by poisoning with acids such as aspirin, by fever or by anxiety.

3 Metabolic acidosis. This may arise from ingestion of acids or substances metabolized to acids (e.g. methanol intoxication, when methanol is oxidized to formic acid); from overproduction of endogenous acids such as ketoacids in diabetes mellitus; from failure to excrete non-volatile acids in certain types of renal disease, including acute and chronic renal failure; from loss of base (HCO_3^-) reserve, for example in severe diarrhoea; or from loss of alkaline upper gastrointestinal contents after surgery (e.g. fistula formation).

4 Metabolic alkalosis. This can occur through ingestion of bases such as sodium bicarbonate. It can also occur if certain diuretics are taken. Diuretics are drugs that promote the flow of urine.

The body compensates for metabolic acidosis by hyperventilation to blow off CO_2, and by increased renal (kidney) excretion of H^+ and HCO_3^- regeneration. The body compensates for metabolic alkalosis by hypoventilation and increased excretion of bicarbonate through the kidneys, although hypoventilation is limited by the fall in arterial P_{O_2} which would otherwise occur.

29 Chemical reactions I

CHEMICAL EQUILIBRIUM

A reversible chemical reaction can be characterized thus:

$$A + B \rightleftharpoons C + D \quad (1)$$

where A and B are reactants, and C and D are products. Initially, the concentrations of C and D are low, but as their concentrations build up, so the reaction slows down, until there is no further net change in the concentrations of any of the chemicals. The reaction has reached a **chemical equilibrium** for the conditions, i.e. temperature and pressure, under which the reaction occurred, and the ratio of reactants and products is constant, defined by the **equilibrium constant** (K_{eq}):

$$K_{eq} = \left(\frac{[C]\,[D]}{[A]\,[B]}\right)_{eq} \quad (2)$$

where [A] and [B] are the molar concentrations of reactants, and [C] and [D] are the molar concentrations of the products.

K_{eq} is useful, since it defines the reaction system under given conditions, but it does not tell whether the reaction is going to occur or not. For that, we need to know the **energy level** at the start of the reaction, and the energy level of the system at the end of the reaction.

FREE ENERGY

Free energy is the energy available for a reaction, and the relative free energies in a system at the start and end of a reaction will determine whether a reaction will take place or not. In other words, it is the **change in free energy** (ΔG) that is important to measure. ΔG is defined by:

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

where Δ means the change, G is free energy, H is heat energy (also called enthalpy) of the system, T is the absolute temperature and S is the entropy of the system (see below). A chemical reaction can occur only if ΔG is negative.

Equation (3) was derived from the laws of thermodynamics, which were formulated to predict the directions of chemical reactions. There are several laws, but only some need concern us here.

Laws of thermodynamics

The **first law of thermodynamics** states that energy is conserved in a chemical system, i.e. the **total energy** of a system and its surroundings is a constant. The first law is in effect the conservation of energy. The energy may be converted from one form to another, for example from chemical bond energies to heat, and vice versa, but the total energy within the system is conserved. There is an equation derived from the first law:

$$\Delta E = E_2 - E_1 \quad (4)$$

where ΔE is the energy change, E_2 is the energy of the system at the end of the reaction and E_1 is the energy at the start.

ΔE and ΔG are related by the equation:

$$\Delta G \approx \Delta E - T\Delta S \quad (5)$$

The first law cannot predict if a reaction will occur spontaneously. For this, another law is summoned, namely the **second law of thermodynamics**, which states that a reaction can occur spontaneously only if there is a net increase in the sum of the **entropies** of the system and its surroundings.

Entropy is a term that means, quite simply, the degree of **disorder** or randomness of a system. An example of an increase in entropy is the diffusion of a solute such as a lump of sugar in a cup of tea. A negative entropy condition would be required to sustain the sugar in the hot tea as a lump. Similarly, negative entropy is a measure of the 'holding together' of biomolecules in their characteristic shape. Notice that ΔG gives no information about the **rate** of a reaction, only if it can occur spontaneously.

Types of reaction

In an **exothermic** reaction, heat is given off. Therefore, ΔH is negative. In an **endothermic** reaction, ΔH is positive since heat is absorbed by the system from its surroundings. According to the first law, no energy can be lost from the system during the chemical reaction, and in an **exergonic** reaction, the energy lost during the reaction is conserved as heat.

Standard free energy changes

Changes in free energy during a reaction are influenced by the pressure, temperature and the initial concentrations of the reactants and products. Biological reactions are also influenced by pH. In order to standardize free energy changes, standard conditions have been adopted: the temperature is taken to be 25°C (298 K); the pressure is 1 atm (1.1325×10^5 Pa); the initial concentrations of reactants and products is 1 mol/l; the pH is taken as 7; the molar concentration of H_2O is taken as unity (1 mol/l). ΔG under these conditions is expressed as $\Delta G^{0\prime}$.

There is a formula for $\Delta G^{0\prime}$:

$$\Delta G = \Delta G^{0\prime} + R\log_e \frac{[C]\,[D]}{[A]\,[B]} \quad (6)$$

where ΔG is the free energy change for a reaction, R is the gas constant, T is the absolute temperature and $\log_e$ is the natural logarithm.

From Equations (2) and (6), it can be worked out that:

$$\Delta G^{0\prime} = 2.303\ RT\log_{10} K_{eq} \quad (7)$$

Note: the **gas constant (R)** is also called the universal molar gas constant, and has a value of 8.134 J/mol per K. It means that all gases have the same kinetic energy for a fixed number of gas molecules and at a given temperature. The **absolute temperature** is measured in Kelvins (K). Theoretically, the lowest possible temperature is –273 K.

If K_{eq} is less than 1, $\Delta G^{0\prime}$ is positive. This means the reaction will not occur unless energy is applied. The reaction is said to be **endergonic**. If K_{eq} is greater than 1, the reaction will occur spontaneously because $\Delta G^{0\prime}$ is negative. The reaction is said to be **exergonic** (see above).

30 Chemical reactions II

OXIDATION–REDUCTION REACTIONS

Oxidation–reduction **(redox)** reactions are important in biochemistry and the mitochondrial electron transport chain, for example is better understood if the basic principles of redox reactions are known beforehand. Any substance that donates an electron to another in a chemical reaction is called a **reductant**. Any substance that accepts an electron is called an **oxidant**. Consider the equation:

$$2A^- + 2H^+ \rightarrow 2A + H_2 \quad (1)$$

which can also be written as two half-equations:

$$2A^- \rightarrow 2A + 2e^- \quad (2)$$

$$2H^+ + 2e^- \rightarrow H_2 \quad (3)$$

Compound A has been oxidized, and H^+ ions have been reduced to gaseous H_2.

Electron flow is part of electrochemical change, and is associated with an electromotive force (emf) that drives the reaction. This force can be measured, and is expressed as the **redox potential** (E), and is a measure of the likelihood that a redox reaction will occur. E can be measured for reactions under standard conditions (1 atm of pressure; 298 K (25°C); pH 7) when it is termed E'_0 and expressed in volts.

Measurement of E'_0

E'_0 is measured as an electron flow between two half-cells, one of which is the sample cell, containing a solution of 1 mol/l concentration of the oxidant to be tested, linked by a bridge to a standard reference half-cell, which contains a 1 mol/l solution of H^+ ions in equilibrium with H_2 gas at 1 atm. The voltage of the reference solution is taken as 0 V under these conditions. Electrons will flow from reductant to oxidant through electrodes immersed in the solution, and the emf is measured in a voltmeter.

Note: although the potential of a standard H_2 electrode is set at 0 V, this is taken for pH = 0. When pH = 7.0, the E'_0 is –0.42 V.

Substance A in Equation (1) has a lower affinity for electrons than does H_2, and the E'_0 will be negative. A positive E'_0 means that the substance has a higher affinity for electrons than does H_2. A strong oxidizing agent, such as O_2, has a positive redox potential, while a strong reducing agent, such as nicotinamide adenine dinucleotide (NADH), has a negative redox potential.

The larger the positive E'_0, the stronger the oxidant, i.e. the higher is its affinity for electrons; in other words, it will tend to be reduced. The more negative the E'_0 is, the more easily will the reductant release electrons.

The tendency of the overall redox reaction to happen can be calculated simply by subtracting the redox potentials of the two half-reactions. For example, consider the exercise below.

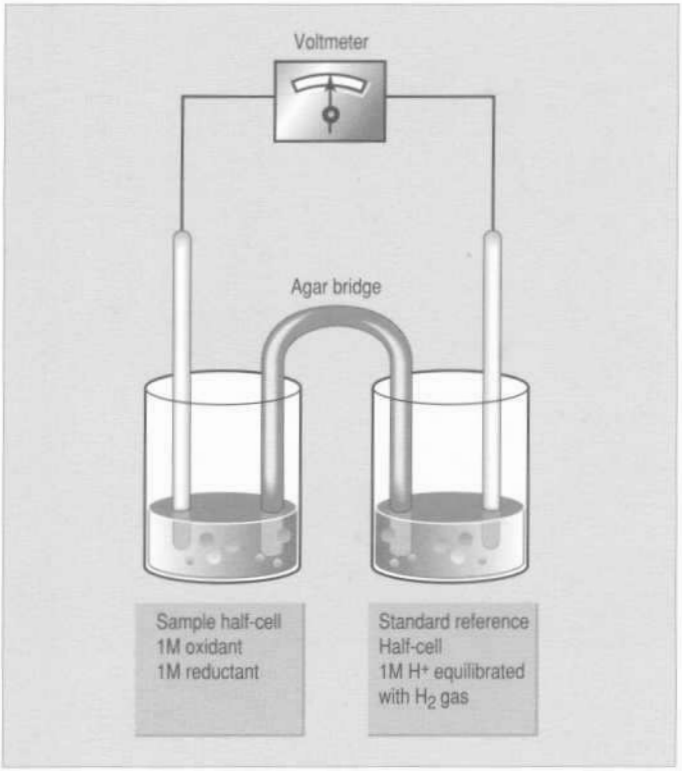

Fig. 30.1

Assume that we dissolve in 1 l of distilled H_2O, at 25°C, 1-g molecule (1 mol) of each of oxaloacetate, acetate, malate and acetaldehyde. Predict which substance will be oxidized, and which will be reduced. Which will be the oxidant and which the reductant?

oxaloacetate + $2H^+ + 2e^-$ → malate — where $E'_0 = -0.10$ V (i)
acetate + $2H^+ + 2e^-$ → acetaldehyde — where $E'_0 = -0.60$ V (ii)

Equation (i) has the more positive redox potential. Therefore, it will proceed more readily as a reduction, i.e. as it is written. Therefore, Equation (ii) will proceed in reverse as an oxidation.

Solving:

oxaloacetate + $2H^+ + 2e^-$ → malate
acetaldehyde → acetate + $2H^+ + 2e^-$
oxaloacetate + acetaldehyde → malate + acetate

Oxaloacetate gains electrons and is reduced to malate, and acetaldehyde loses electrons and is oxidized to acetate. Therefore, acetaldehyde is the reducing agent, and oxaloacetate is the oxidizing agent.

Also:

$$\Delta E'_0 = (-0.10) - (-0.60) = -0.10 + 0.60$$
$$= +0.5 \text{ V}$$

FREE ENERGY OF OXIDATION

The free energy of oxidation, $\Delta G^{0\prime}$, is given by:

$$\Delta G^{0\prime} = -n\mathrm{F}\Delta E'_0 \qquad (4)$$

where n is the number of electrons transferred per mole and F is the Faraday constant (96 500 J/V equiv.) Note that $\Delta E'_0$ must be positive in order for $\Delta G^{0\prime}$ to be negative.

THE NERNST EQUATION

The Nernst equation can be used to calculate the redox potential under non-standard conditions, in which the concentrations of the reactants are present at non-standard concentrations:

$$E = E_0 + \frac{\mathrm{RT}}{n\mathrm{F}} \log_e \frac{[\text{oxidized form}]}{[\text{reduced form}]} \qquad (5)$$

Expressed in terms of $\log_{10}$:

$$E = E_0 + \frac{2.303\mathrm{RT}}{n\mathrm{F}} \log_{10} \frac{[\text{oxidized form}]}{[\text{reduced form}]} \qquad (6)$$

Therefore, the Nernst equation can be expressed as:

$$E = E_0 + \frac{0.059}{n} \log_{10} \frac{[\text{oxidized form}]}{[\text{reduced form}]} \qquad (7)$$

COUPLED REACTIONS

Biochemical reactions very often require energy; they may not 'go' by themselves, for example the formation of macromolecules such as nucleic acids and proteins from their respective nucleotide and amino acid subunits. This sort of elaboration requires work, which in turn requires energy. The energy is supplied by **exergonic** reactions. Exergonic reactions do the work that drive the **endergonic** reactions.

The cell has enzymes that catalyse exergonic reactions; some of the energy produced by the reactions is trapped by coupling the reaction to an endergonic reaction, which generates so-called **energy-rich compounds** such as ATP.

31 Enzymes I

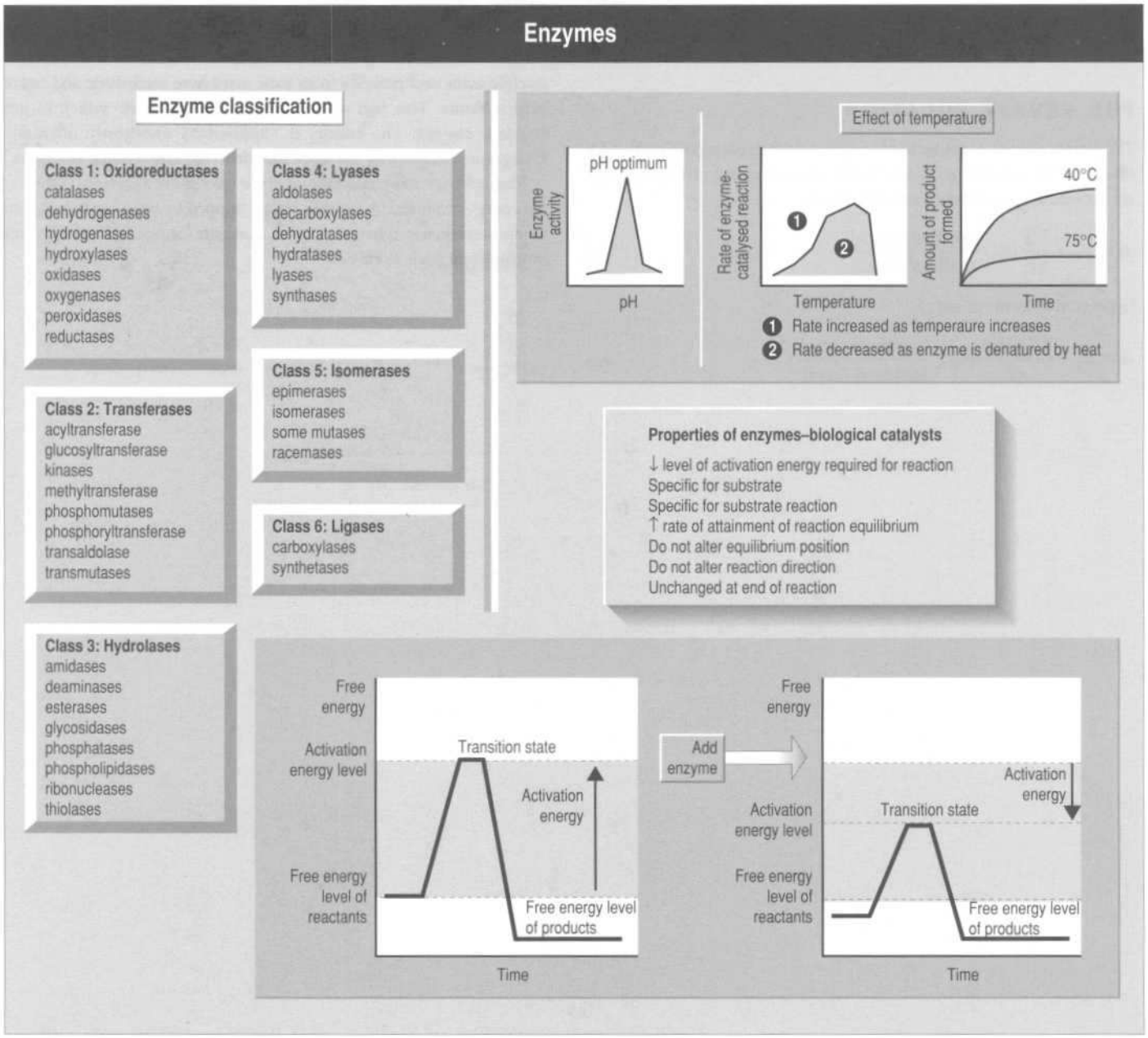

Fig. 31.1

Enzymes are proteins which catalyse chemical reactions. They can operate in the living cell, or in the test tube under the correct conditions, and are even put into washing powders for digesting food stains in clothing. Enzymes increase the rate at which a reaction reaches equilibrium, although they do not alter the thermodynamic properties of the reaction, for example the equilibruim constant of the reaction.

ACTIVATION STATES

Enzymes, like other catalysts, are not changed themselves after the reaction, although they may be temporarily altered in structure while the reaction is proceeding. Enzymes make reactions possible in the body, at rates which allow the cells to live. Normally, at 37°C and at a pH of 7, most of the reactions necessary for life would go too slowly.

During a chemical reaction, the reacting chemicals pass through an energy state higher than that of either the reactants or of the products. The reactants need to attain a certain **activation energy** in order to reach a so-called **transition state,** and the rate of the reaction will depend on the rate at which individual reacting molecules build up activation energy to the transition state.

Enzymes, like other catalysts, actually reduce the amount of activation energy required. They achieve this by providing alternative pathways for the reaction; these pathways require less activation energy than would be needed in the absence of the enzyme.

PROPERTIES OF ENZYMES

Substrate and reaction specificity

Unlike inorganic catalysts, such as platinum, which can catalyse a whole host of reactions in the test tube, the biological catalysts, the enzymes, are highly **specific**. They will recognize one or a few chemically closely related substrates. Substrates may be defined as substances on which enzymes act in biochemical reactions. Similarly, enzymes will catalyse specific **types** of reactions.

Chemical composition

Enzymes are macromolecules. They are almost all proteins (although it has been shown that RNA can catalyse certain reactions). They are composed of amino acid chains, whose specific sequences determine the folding, shape and function of the enzyme.

FACTORS AFFECTING ENZYME ACTIVITY

pH

The pH can affect biochemical reactions in a variety of ways:
1 extremes of pH may radically alter enzyme structure, and thus denature the enzyme;
2 the pH may affect the degree to which the substrate is ionized, and thus affect the rate of the reaction;
3 the pH can affect the binding of the enzyme and substrate;
4 the pH can alter the reactivity of the enzyme during the catalytic process.
Only (3) and (4) will be considered here.

Most enzymes are active only within a fairly limited pH range, and they have an optimum pH at which their activity is greatest. The pH optimum will depend on where in the organism the enzyme is physiologically active. For example, pepsin, a digestive enzyme, operates in the stomach in the presence of hydrochloric acid, and its pH optimum is around 2. Lysosomal enzymes have a pH optima of around 5, the pH in the lysosome.

Temperature

All chemical reactions are increased as the temprerature is increased, including enzyme-catalysed reactions. But, when enzymes are heated above 40°C many of them begin to become denatured, which results in a fall in activity and a fall in the rate of the reaction. Once the enzyme is denatured, comparatively little more of the product will be formed, no matter how long the reaction is left to proceed.

CLASSIFICATION OF ENZYMES

The Commission on Enzyme Nomenclature of the International Union of Biochemistry established a classification of enzymes into six classes.
1 Oxidoreductases: these catalyse oxidation–reduction reactions. An example is alcohol dehydrogenase, which oxidizes alcohol to acetaldehyde.
2 Transferases: these catalyse group transfer reactions. The groups transferred include methyl, ketone, nitrogenous or phosphorus groups. Examples of a transferase are hexokinase and methyl-transferase.
3 Hydrolases: these catalyse hydrolytic reactions. They cleave O–P, C–N and C–O bonds. Examples are the peptide hydrolases, which cleave peptide bonds.
4 Lyases: these catalyse the reversible addition of groups to double bonds, or formation of double bonds by removal of groups. For example, they may remove ammonia, CO_2 or H_2O during the reaction. An example is pyruvate decarboxylase, which decarboxylates a keto-acid to yield an aldehyde with release of CO_2.
5 Isomerases: these enzymes catalyse different kinds of isomerization, which involves the rearrangement of a molecule to yield one with different physical and/or chemical properties. There are different types of isomerase. The **epimerases** and **racemases** catalyse inversion at asymmetrical carbon atoms. For example, lactate racemase catalyses the conversion of L-lactate to D-lactate. The **mutases** catalyse the transfer of groups within a molecule. For example, phosphoglycerate mutase produces 3-phosphoglycerate from 2-phosphoglycerate.
6 Ligases: these are also termed synthetases. They catalyse the condensation of two molecules, and the reaction is coupled to the cleavage of a high-energy phosphate bond, such as is found in ATP. An example is pyruvate carboxylase, which condenses pyruvate and bicarbonate to yield oxaloacetate, which is coupled to the conversion of ATP to ADP.

32 Enzymes II

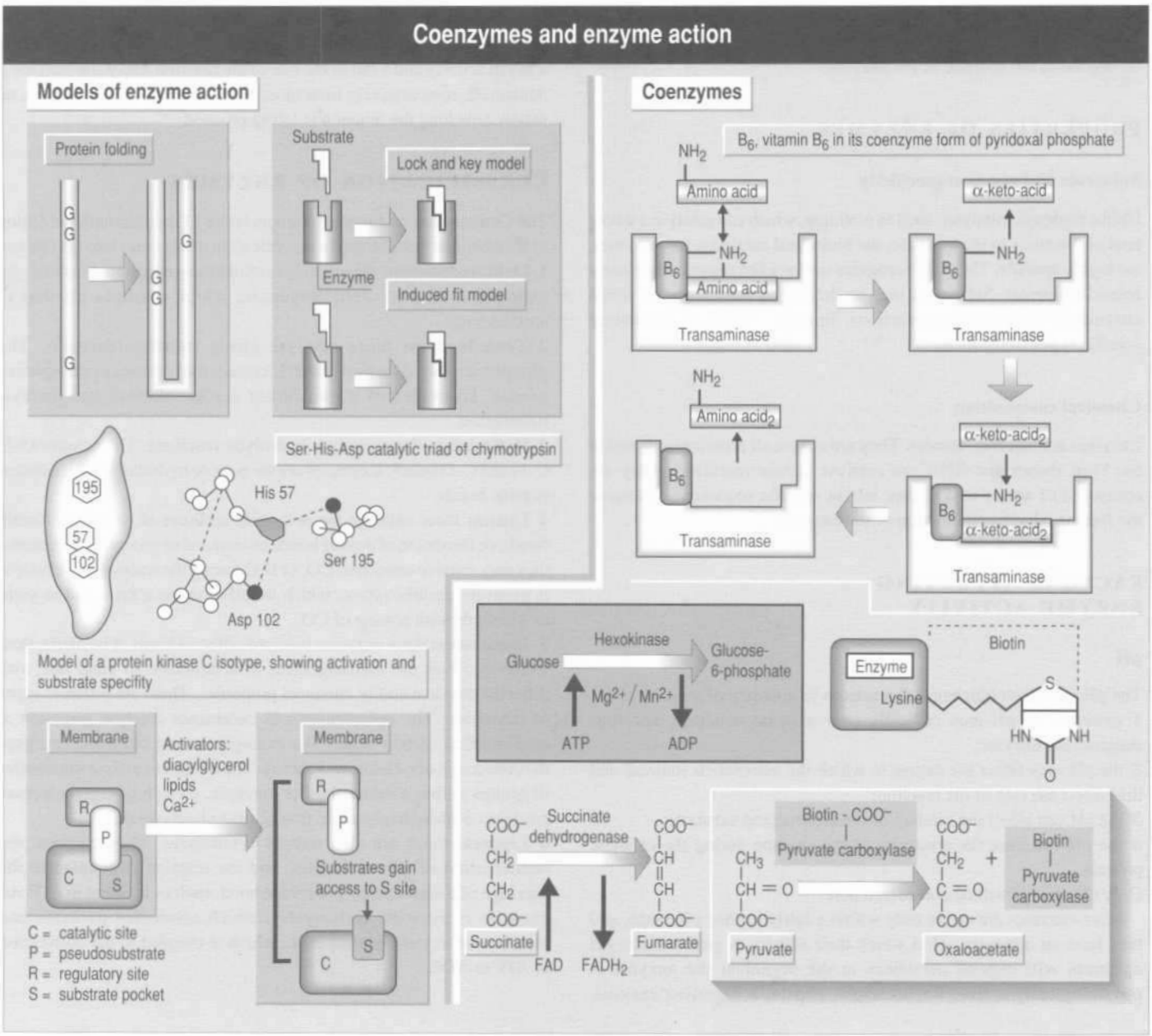

Fig. 32.1

MECHANISM OF ACTION

Protein folding

Enzymes are globular proteins consisting of one or more polypeptide chains which fold into a three-dimensional structure, depending on the amino acid sequence. The amino acids forming an active site may be far apart, but are brought into proximity through chain folding. Active sites on the enzyme surface include the substrate-binding, regulatory and catalytic sites. The nature and arrangement of the amino acids at a site determine, largely, the functional specificity of the active site.

Substrate–enzyme complex formation

The substrate is held at its binding site by non-covalent forces, including hydrogen and hydrophobic bonds, and electrostatic and van der Waals forces. The binding site is so specific for a substrate that often it will bind only one isomer of a diastereomeric pair. Some enzymes are not

as specific. Glucokinase will bind only glucose, while hexokinase will bind and catalyse the phosphorylation of 2-deoxyglucose, fructose, glucosamine, glucose and mannose, although not all at the same rate.

It is believed that when the substrate makes contact with the active site, the site changes its conformation to accommodate the substrate. This is the **induced fit** model. Possibly, the induced fit mechanism places strains on the substrate, thereby lowering the activation energy required for the reaction to occur.

Mechanism of catalysis

The substrate-binding site is usually where some or all of the catalytic reaction occurs. **Chymotrypsin** is a digestive enzyme catalysing the hydrolysis of dietary proteins in the small intestine, cleaving peptide bonds at the carboxyl side of aromatic side chains of phenylalanine, tryptophan and tyrosine. There are two main steps in the reaction.

1 The substrate binds to the enzyme at a chemically active site dominated by a triad of three amino acids. These are Asp 102, His 57 and Ser 195, which form a **catalytic triad.** The substrate and the triad become bound by hydrogen bonding.

2 The susceptible peptide bond of the substrate is cleaved through the action of His 57 and Ser 195, and the peptide substrate is hydrolysed to yield an acid and an amine.

Regulatory sites on enzymes

Enzyme activity may be regulated by sites on the enzyme itself. Examples are enzymes that phosphorylate proteins, the **protein kinases.** Protein kinase C (PKC), a serine/threonine kinase, is activated by DAG, phospholipids and calcium ions, and is believed to mediate cellular events following activation of cells by hormones and second messengers. PKC appears to be a member of a family of PKC isotypes.

The enzyme has functionally distinct regions: a substrate pocket; an amino-terminal regulatory site; a pseudosubstrate site within the regulatory domain; and a carboxy-terminal catalytic site. The pseudosubstrate site consists of a sequence of amino acids resembling the substrate, but without a serine/threonine residue which can be phosphorylated. Instead, it contains an alanine residue. In the inactive state, the pseudosubstrate site occupies the substrate pocket within the catalytic domain and inactivates it. Activation of the enzyme causes a conformational change and dissociation of the pseudosubstrate from the substrate pocket. But, the pseudosubstrate still interacts with the substrate pocket, allowing only certain substrates to gain access to the pocket, thus playing a part in determining substrate specificity.

COFACTORS

Cofactors are chemicals that assist or are necessary for enzyme action. Cofactors become attached to the enzyme, usually at the catalytic site, and may enable the binding of the substrate and/or the catalytic process. There are two main groups: (i) **coenzymes;** and (ii) **prosthetic groups.**

Coenzymes

Coenzymes may be metals, e.g. cobalt, copper, iron, Mg^{2+}, manganese (Mn^{2+}) or Zn^{2+}. For example, Mg^{2+} or Mn^{2+} are needed for the reduction of the high negative charges of ATP during the kinase-catalysed phosphorylation reaction. The phosphorylation of glucose to glucose-6-phosphate provides an example of this type of reaction. Coenzymes may be organic molecules which are derived from vitamins.

Mechanism of coenzyme action. Coenzymes, such as flavine adenine dinucleotide (FAD) may be required as proton and electron acceptors. FAD is reduced to $FADH_2$ during dehydrogenation reactions, and in a separate reaction is oxidized back to FAD, which is ready to act as coenzyme again. Therefore, FAD is not only a coenzyme but a substrate too. Since the coenzyme is altered during the reaction, it is sometimes referred to as a **second substrate** for the enzyme.

Vitamin B_6 (pyridoxine) is an α-amino group acceptor which can transfer the δ-amino group to an α-keto-acid, thus forming a new amino acid.

Prosthetic groups

Prosthetic groups are non-proteins, which bind covalently to the enzyme at its active site. These include metal ions and organic molecules such as biotin. Biotin is vitamin B_4, and is required for the incorporation of CO_2 into organic compounds. It does this by acting as a carrier for CO_2.

Biotin binds covalently to a lysine residue at the catalytic site of the enzyme, and accepts $-COO^-$, usually from HCO_3^-. The $-COO^-$ is then passed rapidly to the substrate to form a carboxylated compound. Biotin needs ATP to bind CO_2.

Metals such as Zn^{2+} may function as cofactors by binding the substrate, and/or by promoting catalysis. Proteins that contain covalently bound metal ions are termed **metalloproteins.**

MULTISUBSTRATE REACTIONS

Enzyme reactions in which one enzyme may be able to bind more than one substrate may be **sequential** or what is sometimes called **'ping-pong'**. In a sequential reaction, one substrate may need to be bound before another, for example alcohol dehydrogenase first binds ethyl alcohol prior to binding the cofactor NAD^+. The reaction is said to be **ordered,** whereas one in which two substrates can bind in any order is said to be **random.** In a **ping-pong** reaction, an enzyme binds a substrate, converts it to a product which is released and then binds another substrate. An example is the transaminase reaction, where a keto-acid is released as product and another keto-acid bound as a substrate.

33 Enzymes III

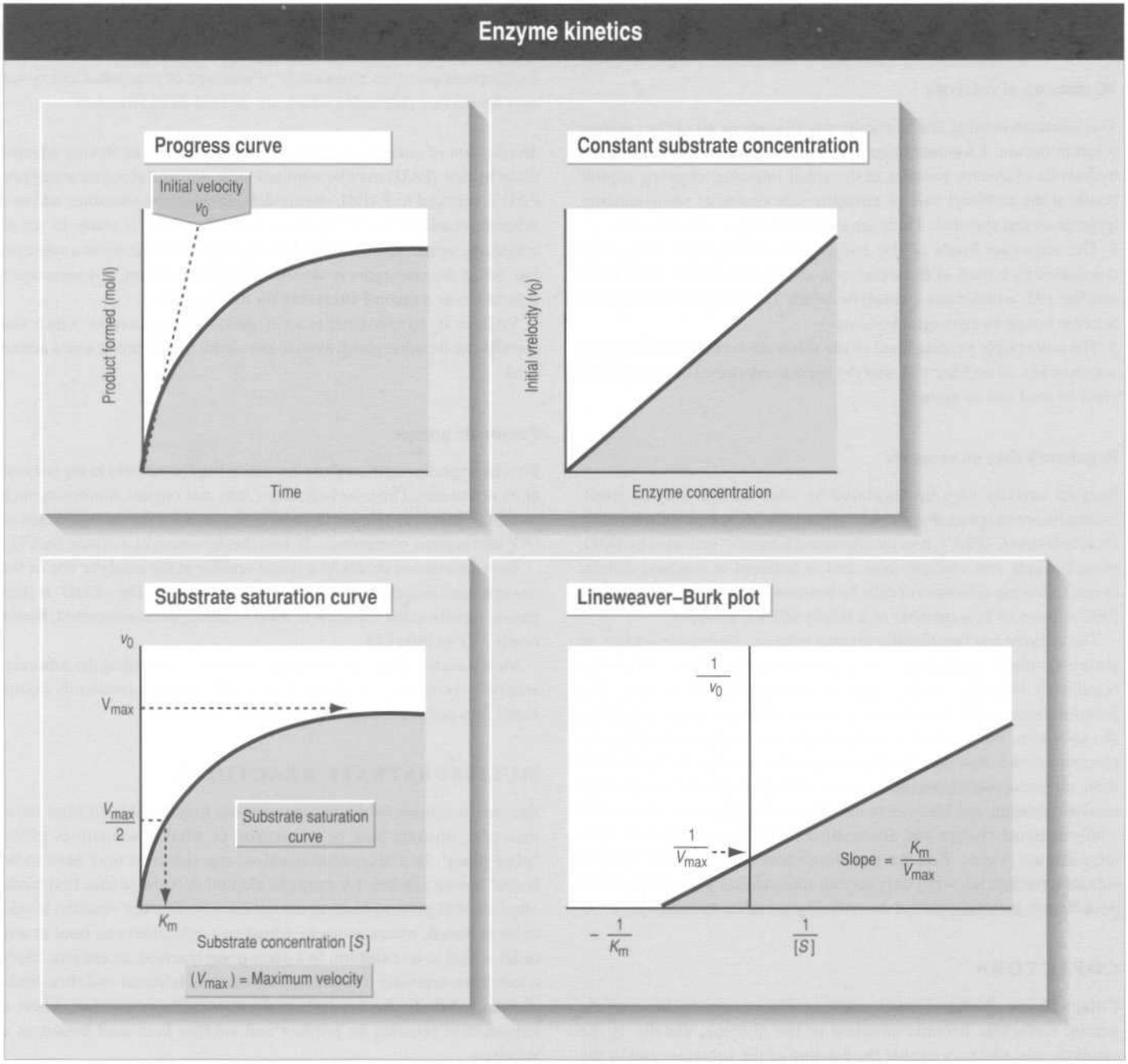

Fig. 33.1

ENZYME KINETICS

Enzyme–substrate interaction

When enzyme and substrate react together, the rate of the reaction will increase at first then slow down with time as the enzyme becomes degraded, or as equilibrium is approached. From the graph obtained, called a **progress curve**, we can derive the initial velocity, v_0. There is a direct relationship between v_0 and the enzyme concentration in the reaction mixture, in that the rate of the reaction will double when the enzyme concentration is doubled, for a given concentration of substrate.

Substrate saturation curve

If v_0 is plotted against substrate concentration at a given concentration of enzyme, then a curve is obtained whose shape we call a rectangular hyperbole. Initially, the rate of the reaction v_0 is directly proportional to the substrate concentration $[S]$. (Mathematicians call this 'first-order' kinetics.) But, as $[S]$ increases, the rate eventually reaches a limiting maximum, V_{max}. The maximum reflects the fact that all the binding sites on the enzyme are taken up or saturated, and the rate could be increased only by adding more enzyme. This curve is called the **substrate saturation curve**. (Mathematicians call the 'bend' in the curve 'mixed' or first- and zero-order kinetics.) The flat part of the curve reflects the fact that there is no further increase in rate no matter how much substrate is added. (This is called 'zero-order' kinetics.)

Enzyme units

Enzyme activity is measured in units established by The Commission on Enzyme Nomenclature of the International Union of Biochemistry (SI). The unit is the **katal** (kat). One katal is the quantity of enzyme in the presence of which 1 mol of substrate is converted per second. It is more common, however, to express enzyme activity in millikatals or microkatals, since 1 kat is equivalent to the activity of about 1 kg of the pure enzyme. In many textbooks, enzyme activity is expressed as micromoles of substrate that are converted per minute to product under given assay conditions. The **standard unit** of enzyme activity, U, is the amount of enzyme catalysing the formation of 1 μmol of substrate per minute. The **specific activity** of an enzyme preparation is the number of enzyme units per milligram of protein. This gives an indication of the purity of the preparation.

The formula of the substrate saturation curve

The substrate saturation curve allows us to derive equilibrium constants that characterize the enzyme–substrate reaction for given conditions of pH and temperature. We can arbitrarily define a constant, the K_m, or **Michaelis constant**, which is the substrate concentration when the rate v_0 is half the maximum rate (see below). The equation that describes the curve was derived by Michaelis and Menten, and the equation they derived is called the **Michaelis–Menten equation**:

$$V_0 = \frac{V_{max}[S]}{K_m + [S]} \tag{1}$$

If v_0 is plotted against $[S]$, a rectangular hyperbole is obtained. It is easy to calculate that when $v_0 = V_{max}/2$, then $K_m = [S]$.

In order to obtain values of K_m and V_{max}, the Michaelis–Menten equation can be linearized.

1 Invert Equation (1):

$$\frac{1}{v_0} = \frac{K_m + [S]}{V_{max}[S]} \tag{2}$$

2 Separate and simplify to give the Lineweaver–Burk plot:

$$\frac{1}{v_0} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}} \tag{3}$$

The Lineweaver–Burk plot is the equation of a straight line ($y = mx + c$), when $1/v_0$ (y) is plotted against $1/[S]$ (x), with gradient (m) $= K_m/V_{max}$, and the intercept (c) on the ordinate is $1/V_{max}$. $1/K_m$ can be read directly from the graph where the line crosses the abscissa.

Although not presented here, there are other methods of linearizing the Michaelis–Menten equation, which are generally favoured by biochemists.

34 Enzymes IV

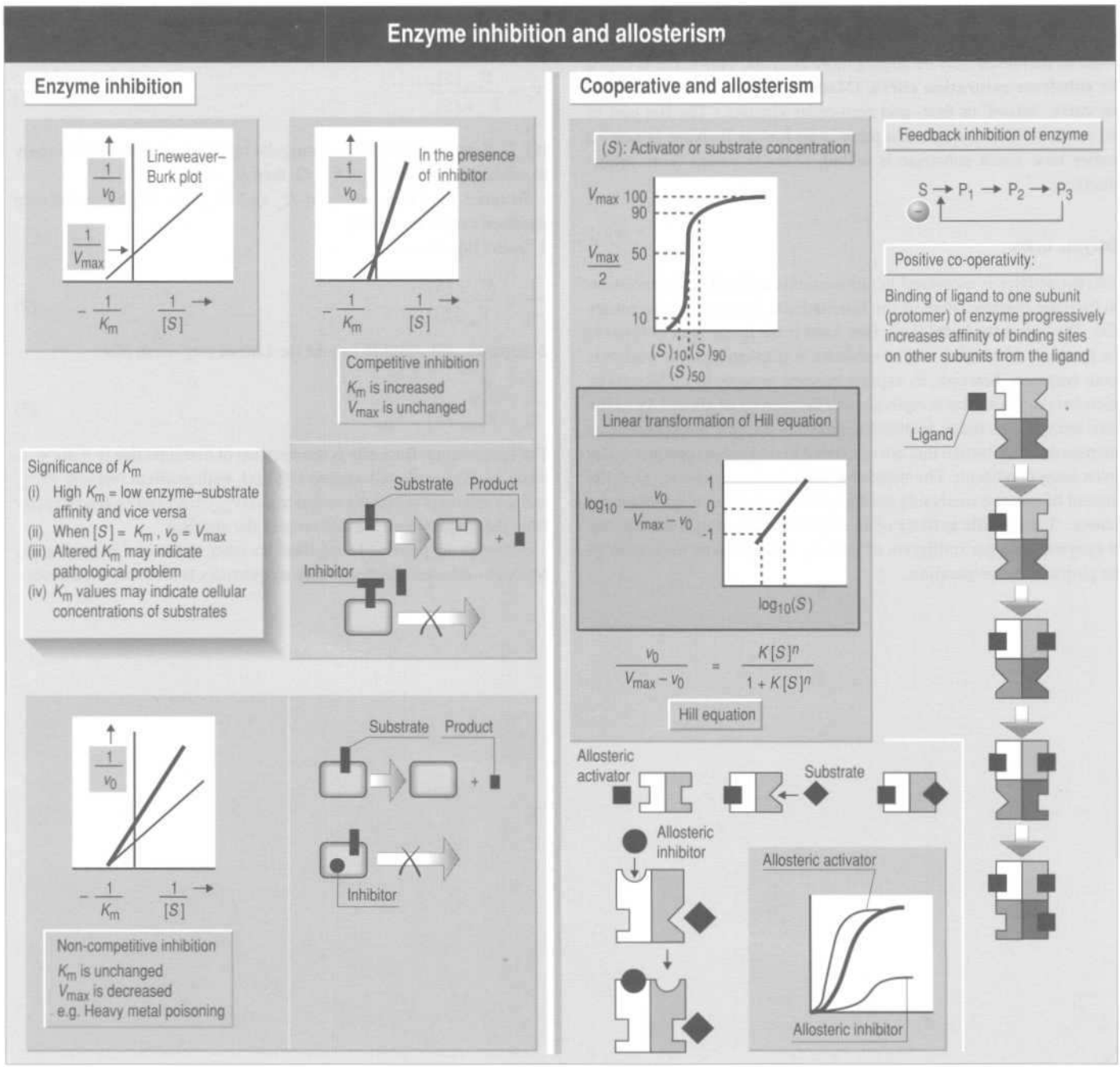

Fig. 34.1

KINETICS OF ENZYME INHIBITION

Enzymes may be inhibited from acting by small ions or molecules which form part of regulatory control systems or by drugs. Inhibition of enzyme action may be **irreversible** or **reversible**. Reversible inhibition may be **competitive** or **non-competitive**; the nature of reversible inhibition can be determined using enzyme kinetics and the Lineweaver–Burk plot. Some small molecules may control enzyme activity by binding to **allosteric sites**, which inhibit or activate the enzyme. Many enzymes are **oligomers**, being formed of identical subunits or **protomers**, each with a substrate-binding site. Activation of an allosteric site on one protomer may increase substrate–enzyme affinity on other protomers, a process called **co-operativity**.

Irreversible inhibition

Irreversible inhibition is permanent enzyme inhibition, usually due to the covalent attachment of a chemical to the enzyme at one or more of its active sites, or at another site which alters the conformational shape of the enzyme. The organism has to produce more enzyme to replace it. Examples of irreversible inhibitors are heavy metal ions such as mercury ions. Nerve gases such as the organophosphorus compound di-isopropylphosphofluoridate (DFP) bind irreversibly to the specific serine in the active centre of a large group of hydrolases, for example acetylcholinesterase.

Reversible inhibitors

Reversible inhibitors bind to enzymes non-covalently and are able to dissociate, thus leaving them free to catalyse substrates. Reversible inhibitors can be removed by dialysis.

Competitive inhibition

A competitive inhibitor competes with the substrate for its binding site on the enzyme, and is often structurally similar to the substrate. Once bound, the inhibitor may itself be converted to a product, or occupy the site until it dissociates from it. The action of the inhibitor can be overcome by increasing the concentration of the normal substrate at the binding site. The Lineweaver–Burk plot in the presence of a fixed concentration of the inhibitor reflects competitive inhibition through an apparent increase in the K_m, and an unaltered V_{max}.

An example of a competitive inhibitor is malonate, which competes with succinate at its binding site on the enzyme succinate dehydrogenase, where succinate is normally converted to fumarate. If the concentration of succinate is increased, it will displace malonate from the enzyme. Examples of metabolized inhibitors are the antibiotic sulphonamides, such as **sulphanilamide**, which bind to dihydropterate synthetase, a bacterial enzyme that synthesizes folic acid from *p*-aminobenzoate, which is necessary for bacterial growth. The enzyme converts the sulphonamide to a compound that cannot be metabolized to folic acid, and the bacteria die.

Non-competitive inhibition

Non-competitive inhibitors usually bind to the enzyme at sites other than that which binds the substrate, and the substrate does not compete with the inhibitor. Therefore, although the substrate may still be able to occupy its own binding site on the enzyme, it is not converted to product. The inhibitor–enzyme–substrate complex is sometimes called a 'dead-end' complex, since it is catalytically inactive. Non-competitive inhibitors in effect remove enzyme from the available pool, and the Lineweaver–Burk plot reflects this in that the K_m is unchanged but the V_{max} is reduced.

ALLOSTERISM

Allosterism is a word used to describe enzymes that have binding sites for molecules, usually of low molecular weight, other than those where substrates bind and are converted to product. These small molecules are termed **ligands**. A ligand is a molecule that binds to a binding site on a macromolecule, such as an enzyme or a receptor. These allosteric binding sites are very often where enzyme activity is controlled. Ligands that bind to allosteric binding sites may be **allosteric activators** or **allosteric inhibitors**. The product of an enzyme may itself be an allosteric inhibitor, binding to and inhibiting an enzyme further back in the chain of metabolic pathways that produced it. This is an example of a physiological control mechanism whereby metabolic processes are regulated.

Classes of allosteric enzymes

Allosteric enzymes have been classified as **K class** and **V class**, depending on how their allosteric ligands affect the kinetic constants K_m and V_{max}. In the presence of their allosteric ligands, K-class enzymes yield plots showing an altered K_m and unchanged V_{max} (as is seen with competitive inhibitors of enzyme action). In other words, an allosteric inhibitor binds to the allosteric site, and the enzyme reacts by losing affinity for the substrate.

In the presence of their allosteric inhibitors, V-class enzymes yield plots showing an unchanged K_m and lower V_{max} (as is seen with non-competitive enzyme inhibitors).

Mechanism of allosterism

The mechanism of allosterism is not well understood, but it is believed that when a ligand binds to an allosteric site it alters the conformation of the enzyme so that the affinity of the substrate (or some other ligand) for its binding site on the enzyme is either increased or decreased, depending on whether the ligand is an allosteric activator or inhibitor, respectively.

If a ligand increases the affinity of another protomer for the same ligand, this is termed a **homotrophic** interaction. If the ligand increases the affinity of another protomer for a different ligand, this is termed a **heterotrophic** interaction.

Co-operativity is the term used to describe the effect that one protomer can exert on another, as happens when a ligand binds. Co-operativity reveals itself in the substrate saturation curve, which is sigmoidal, and the formula of the line is given by the **Hill equation**, which can be linearized. Allosteric activators will shift the curve to the left, while inhibitors will shift the curve to the right. The binding of O_2 to Hb provides another example of positive co-operativity and a sigmoidal curve. Note, however, that sigmoidal curves will always be obtained with multisubunit enzymes, even in the absence of co-operativity.

Allosteric effects can occur on one protomer **without** a co-operative effect on other protomers. For example, alcohol dehydrogenase is a zinc-containing metalloprotein which reduces acetaldehyde to alcohol, and consists of protomers which are independently activated allosterically.

35 Digestion: basic principles and cell types

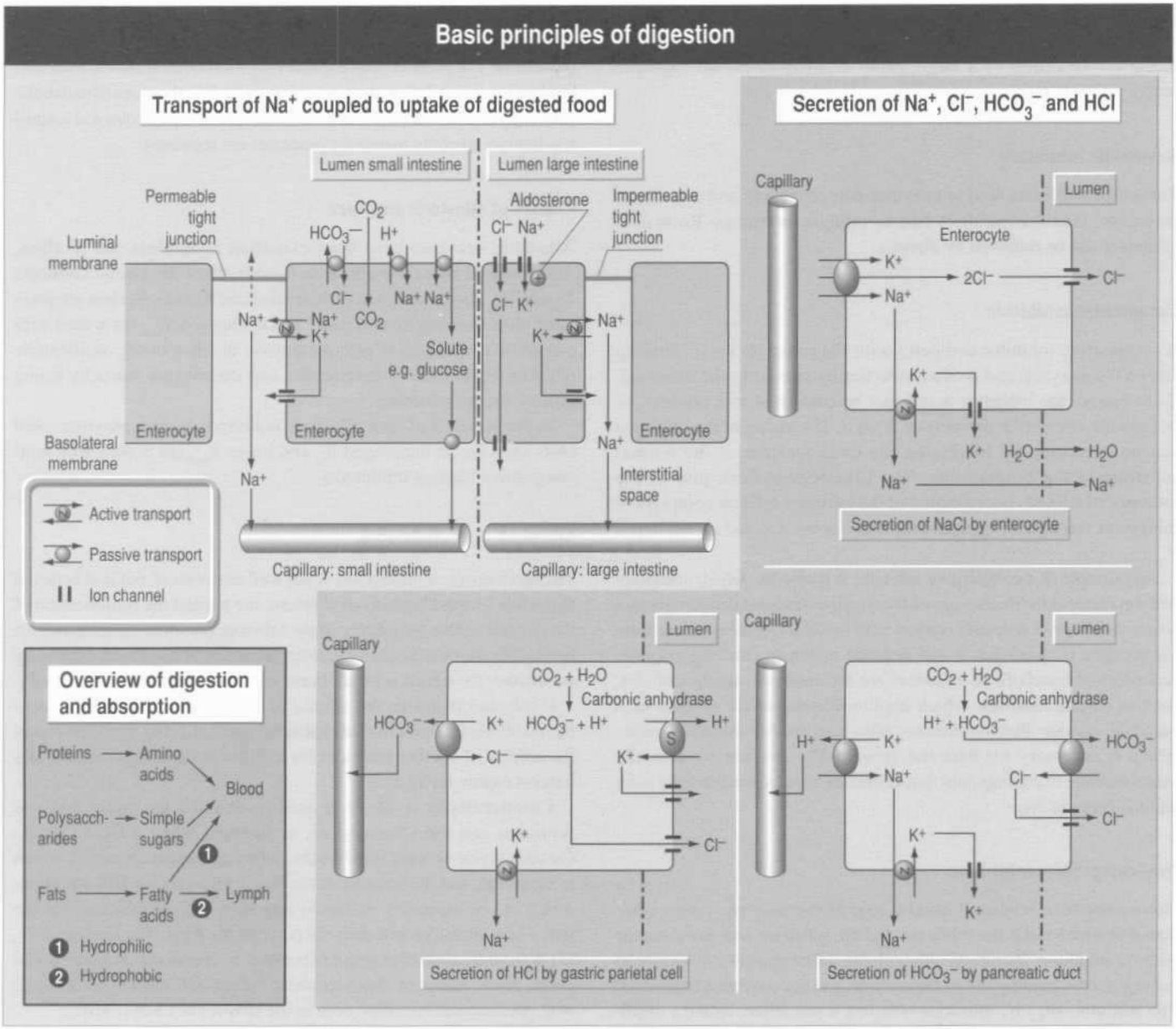

Fig. 35.1

INTRODUCTION

When food is ingested, it is not in a form that can be readily absorbed by the cells of the alimentary tract. For efficient absorption to occur, the polymeric foodstuffs are hydrolysed into their constitutive residues, which can be taken up more easily into the epithelial cells.

Digestive enzymes are released in controlled secretions from specialized organs and cells of the alimentary tract. These secretions also contain compositions of solutes to maximize the activity of the enzymes. Daily, 1–2 l of saliva are secreted into the mouth, 2.5 l of gastric juice are secreted into the stomach, 1–3 l are secreted into the small intestine by the epithelial cells of the intestinal wall and 1 l of bile and 1 l of pancreatic juice drain directly into the duodenum.

Due to the differences in the composition of the secretions, the pH of the tract varies widely. In the stomach, the pH is usually 4.0 due to the secretion of hydrochloric acid, but in the duodenum the pH rises to 6.5–6.8 with the addition of bicarbonate ions in the pancreatic juice.

The epithelial cells of the intestinal wall, therefore, are not only adapted to the absorption of digested foods but also to the maintenance

of a luminal environment in which digestion and absorption can occur at maximal rates.

Solute transport

There are two distinct pathways for solute transport: (i) the solutes may pass through the tight junctions which connect the epithelial cells of the intestinal wall to each other (**paracellular route**); or (ii) they may pass directly through the epithelial cells themselves, so passing through the luminal membrane and the basolateral membrane (**transcellular route**).

The luminal membrane contains many protein transport molecules that are specialized for uptake or secretion of solutes and of digestive enzymes. The basolateral membrane is more typical of the plasma membrane of most cells, while also showing transport systems for the exit of nutrients absorbed into the cell from the lumen.

The dominant driving force for almost all of the transcellular solute movement against an electrochemical gradient is the Na^+/K^+–ATPase in the basolateral membrane. The actions of this energy-consuming transport protein are responsible for low Na^+ and high K^+ concentrations in the cytosol and an electrical potential of the cytoplasm of –60 mV compared with the extracellular fluid. The potential difference is caused by the asymmetrical transport of the ions by the protein (three Na^+ ions are transported out, while two K^+ ions are transported in).

MECHANISMS FOR THE ABSORBTION OF SOLUTES

The Na^+ in the lumen of the small and large intestines orginates from both the dietary intake and the secretions of the exocrine glands which drain into the intestinal tract. The uptake of Na^+ into the enterocytes along with Cl^- is therefore crucial to the maintenance of overall electrolyte balance in the body. The transport of Na^+ is also intimately involved in the uptake of digested food molecules such as glucose and amino acids (see Fig. 35.1).

The tight junctions connecting the epithelial cells in the large intestine are much less permeable to Na^+ and H_2O than are those in the small intestine. This correlates with the Na^+-scavenging function of the large intestine and the ability of the small intestine to secrete Na^+.

Uptake of solutes

The **luminal transporter** involves the uptake of solutes such as glucose or amino acids: Na^+ flows into the cell down the electrochemical gradient, established by the Na^+/K^+–ATPase, through a cotransport protein which carries the solute across the luminal membrane against its concentration gradient. This cotransporter can transport both ways, but is influenced in one direction by the negative potential in the cell.

In the small intestine, six specific carrier-mediated cotransporter systems for L-amino acids have been identified: (i) acidic amino acids, e.g. aspartate; (ii) basic amino acids, e.g. arginine; (iii) uncharged amino acids with polar or short side chains, e.g. threonine; (iv) uncharged amino acids with hydrophobic or aromatic side chains, e.g. methionine; (v) imino acids, e.g. proline; and (vi) β-amino acids, e.g. taurine.

D-fructose, D-glucose and D-galactose are the main monosaccharides resulting from digestion of carbohydrates. The latter two are absorbed by a monosaccharide cotransporter on the luminal membrane as described above, while D-fructose is absorbed through the luminal membrane by a facilitated diffusion mechanism which is Na^+ independent.

The **basolateral transporter** is a different type of transport system which facilitates the passage of the absorbed solute from the epithelial cell into the bloodstream. The basolateral transporter is not a cotransporter.

MECHANISMS FOR THE SECRETION OF SOLUTES

The secretion of solutes other than the digestive enzymes themselves is important for a number of reasons.

1 The changes in osmotic pressure that result from the movement of the solutes from the epithelial cell layer cause the movement of H_2O into the lumen of the tract. The H_2O is necessary to provide a more effective medium for digestion.

2 Enzymes in different areas of the digestive tract require differing conditions for their maximal efficiency, and the controlled secretion of acidic and basic ions can provide the enzymes with the environment that they need.

Na^+ and Cl^- can be secreted by most of the epithelial cells of the intestinal tract, but the pH of lumen is controlled by specialized cell types that secrete bicarbonate ions into the pancreatic juice and hydrochloric acid into the stomach.

PATHOPHYSIOLOGY

Vibrio cholerae is an infection of the gastrointestinal tract. Toxins that are produced by the bacteria cause excessive secretion of electrolytes by the stimulation of a cAMP regulatory pathway. The resulting diarrhoea can be life-threatening due to loss of essential electrolytes and H_2O. Treatment is based on the fact that the Na^+–glucose cotransporter is not regulated by a cAMP pathway and so administration of oral glucose will cause absorption of Na^+.

36 Digestion of protein and carbohydrates

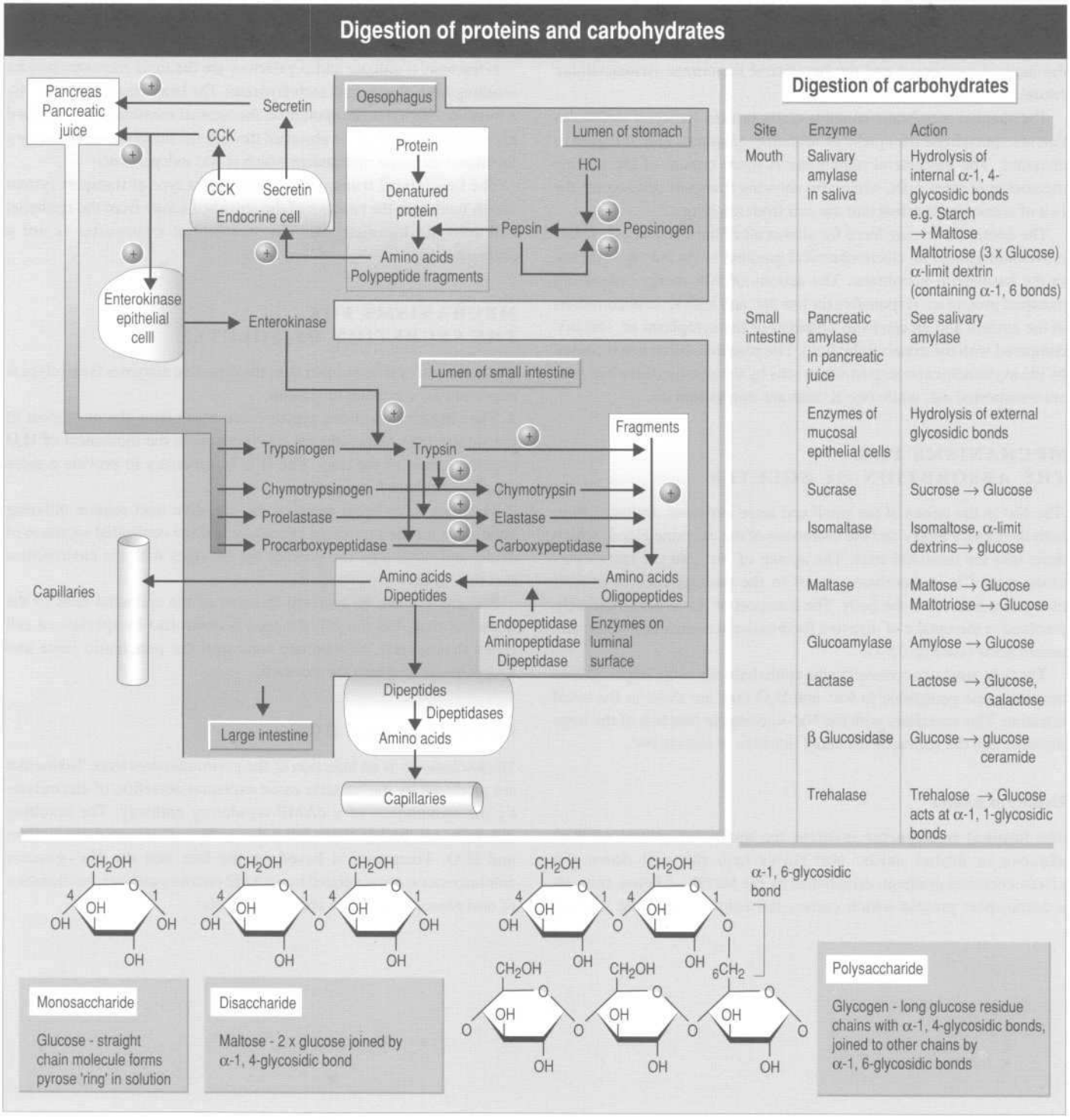

Fig. 36.1

DIGESTION OF PROTEINS

On average, a 70-kg man eats 80–100 g of protein daily. In addition to this, 10–20 g of protein are secreted as enzymes and about 20 g of protein as mucosal cells that are sloughed off from the intestinal surface into the lumen of the intestine. Virtually all of this protein is digested and absorbed.

The first stage of protein digestion occurs in the mouth with the mechanical breakdown of the food by the teeth. This provides a larger surface area for the later stages.

Protein digestion in the stomach

In the stomach, hydrochloric acid secreted by the parietal cells kills invading bacteria and causes unravelling of the protein chains or **denaturation**, providing an increased surface area for digestion. The first stage of enzymatic digestion also occurs in the stomach. When digestive enzymes are released into the lumen, they are in an inactive form (**zymogens**) so that they will not damage the mucosal surfaces of the epithelial cells that line the intestine. **Pepsinogen** is released by zymogen-releasing cells in the stomach, and is the precursor for the digestive enzyme **pepsin**. Cleavage of the peptide bond between residues 44 and 45 of pepsinogen to release pepsin can occur spontaneously at a pH more acidic than 5 (as provided by the hydrochloric acid (**auto-activation**)), or by active cleavage of the peptide bond by pepsin itself (**autocatalysis**). Pepsin is stable only in the acidic environment of the stomach and cleaves peptide bonds on the $-NH_2$ side of aromatic amino acids, e.g. tyrosine (Tyr), phenylalanine (Phe). The large peptide fragments and amino acids which result stimulate secretion of digestive enzymes into the small intestine.

Digestion of proteins in the small intestine

Digestion of proteins in the small intestine is triggered by the controlled release of **enterokinase** from duodenal epithelial cells, and is dependent on the release of bicarbonate ions, which neutralize the acid from the stomach. Enterokinase cleaves a hexapeptide from **trypsinogen**, one of the zymogens released from the pancreas, to form **trypsin**. Trypsin, in addition to its own autocatalytic powers, cleaves peptide fragments off other pancreatic zymogens to activate them. The activated pancreatic enzymes hydrolyse peptide bonds at different sites along the polypeptide chains. There are two carboxypeptidases that release amino acids from the carboxyl terminal of the protein. Trypsin, **chymotrypsin** and **elastase** are endopeptidases and will digest the protein from within the chain.

Oligopeptides resulting from the action of the pancreatic enzymes are further digested by endopeptidases, aminopeptidases and dipeptidases present on the luminal surface of epithelial cells of the intestine. The final products of the luminal digestion are amino acids, dipeptides and tripeptides, which can be taken up into the epithelial cells. Further hydrolysis of peptide bonds occurs in the epithelial cells before the final transport of amino acids into the portal blood. In general, all dipeptides and tripeptides are broken down into their constituent amino acids in the epithelial cell. Exceptions to this rule include peptides that contain proline, hydroxyproline or unusual amino acids.

DIGESTION OF CARBOHYDRATES

Nature of carbohydrates

The three most common carbohydrate products of digestion, **glucose**, **galactose** and **fructose** are all examples of hexose monosaccharides. Monosaccharides are the simplest form of carbohydrates as they are comprised of one sugar unit, which is the basis for the structure of more complex carbohydrates. Disaccharides contain two sugar units, oligosaccharides contain a few units and polysaccharides contain many units. The general formula for monosaccharides is $C_nH_{2n}O_n$. The simplest monosaccharides contain three carbon atoms. Those containing four, five, six or seven carbons are called tetroses, pentoses, hexoses or heptoses, respectively. The carbon atoms in the monosaccharide are numbered so that an aldehyde or keto group is at the low number end and a primary alcohol is at the high number end. Almost all the carbohydrates in the body exist in the R–configuration. This refers to the configuration of atoms around C_{n-1}, which determines whether a R-isomer or a S-isomer is formed. Two further isomers are formed when a monosaccharide forms a ring structure in solution: α- and β-isomers. α and β refer to the arrangement of the –H and –OH groups around C_1.

The bond that connects sugar units to each other is called a glycosidic bond, and when formed between two monosaccharides involves the loss of H_2O, i.e. it is a condensation reaction.

Carbohydrates are stored in the form of long-chain polysaccharides whose units are joined by glycosidic bonds. The existence of long chains of units avoids the generation of osmotic pressures in a cell, which would occur with the equivalent number of units stored separately or in smaller chains. Glycogen is the major storage form of carbohydrates in animals. In plants, the equivalent storage molecule is **starch**. Starch is made up of **amylopectin**, which has the same structure as glycogen except that the branch points are less common, and amylose which is made up of glucose residues joined by α-1,4-glycosidic links and has no branch points.

The dietary carbohydrates are made up mainly of starch, sucrose and lactose. In the mouth, the efficiency of the salivary amylase is limited by the extent that the food has been chewed. Acid in the stomach inactivates salivary amylase, but may hydrolyse some of the glycosidic bonds.

PATHOPHYSIOLOGY

Lactase deficiency or milk intolerance is the most frequent of the disaccharidase deficiency syndromes. The inability to digest lactose in the upper small intestine results in bacterial fermentation of the sugar in the lower small intestine. The products of the fermentation include gas, which causes abdominal distension and flatulence, and osmotically active solutes which cause diarrhoea. Treatment includes prior digestion of the milk by commercially purified lactase.

37 Digestion and absorption of lipids

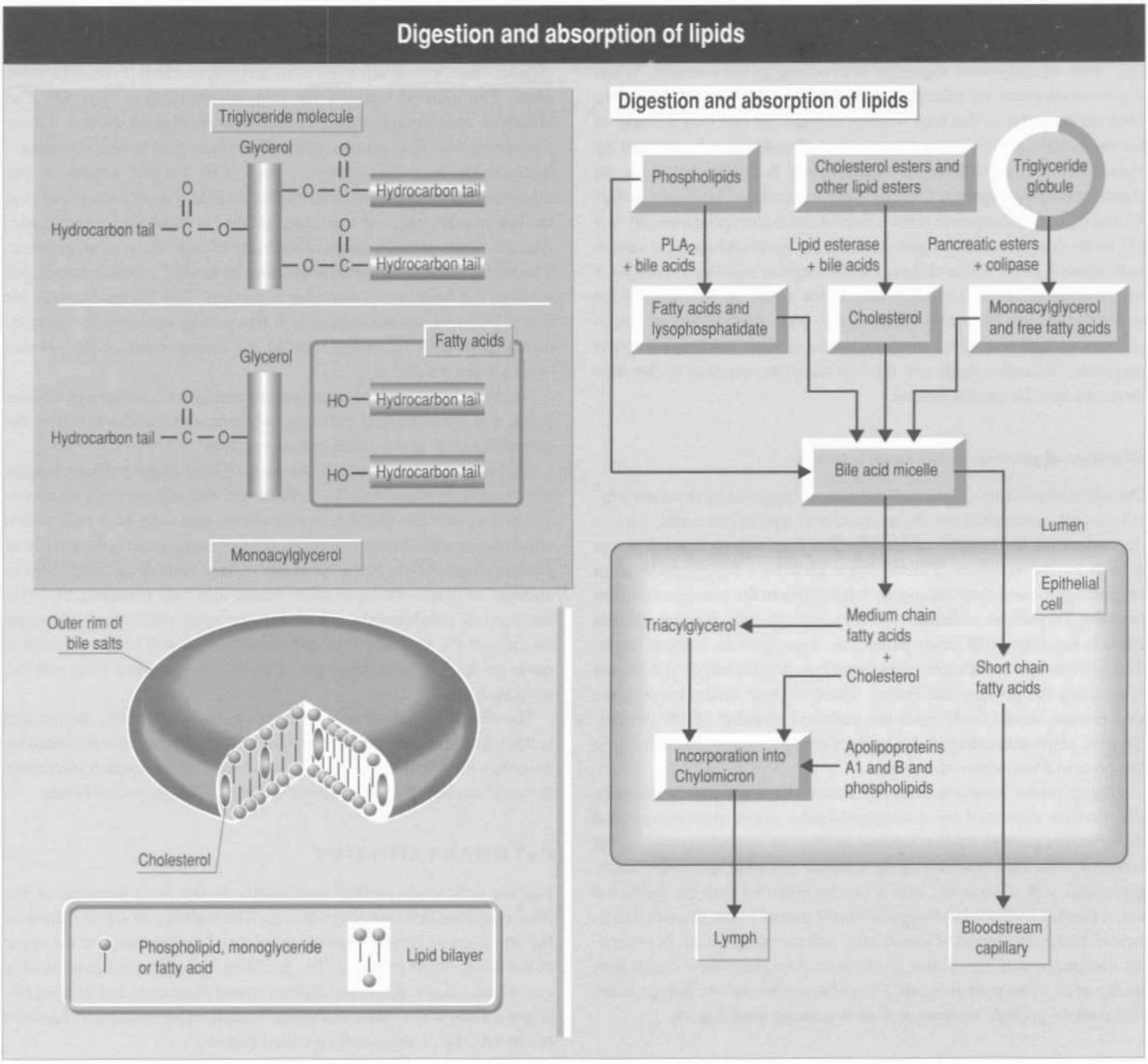

Fig. 37.1

INTRODUCTION

Lipids (or fats) are found in nuts, cheese, milk and oils, and their daily intake in a healthy normal adult is between 60 and 150 g. Most of this is in the form of **triacylglycerols (triglycerides)**. Phospholipids, cholesterol, cholesterol esters and free fatty acids are also present.

Triglycerides and fatty acids

Triglycerides are the most common form of lipid and they are composed of three fatty acids joined by an ester linkage to a molecule of glycerol. Fatty acids are characterized by a long hydrocarbon tail with a carboxyl tail. The most common length for these is between 16 and 18 carbon molecules, although lengths may range from four up to

24 molecules. Each fatty acid has a carboxyl and a methyl terminal. The carbon atom that comprises the carboxyl terminal is numbered C_1. The adjacent carbon is numbered C_2, but is also known as the α-carbon. The carbon that comprises the methyl terminal is known as the ω-carbon.

DIGESTION OF LIPIDS

Lipids are hydrophobic in nature and in an aqueous solution are insoluble and aggregate together. This applies to both the digestive substrates and also to most of the products. The result is that a comparatively large volume of lipids has a small surface area. The digestive process acts to: (i) increase the surface area of lipids; and (ii) 'solubilize' the digestive products so that they can be absorbed.

Gastric digestion

Digestion of lipids begins in the stomach with the action of **lingual lipase**. This is released by glands at the back of the tongue and is active at the acidic pH of the stomach. The action of the lipase is to hydrolyse triglyceride molecules to form free fatty acids, monoacylglycerols, diacylglycerols and glycerol. The action of gastric lipase is aided by churning movements of the stomach, which act to disperse the lipids into a fine emulsion increasing the surface area for digestion.

Intestinal digestion

The main part of the digestion of the fats occurs in the small intestine and is co-ordinated by components of both the pancreatic juice (bicarbonate ions and enzymes) and bile (bile acids). Bicarbonate ions from the pancreatic juice neutralize the acid from the stomach and provide optimal pH conditions for the digestive enzymes. These include pancreatic lipase, lipid esterase and phospholipase A_2, all of which are secreted from the pancreas. **Pancreatic lipase** has the same action as the gastric lipase, but preferentially acts on lipids whose fatty acids contain over 10 atoms of carbon. The enzyme shows inhibition by bile acids but this is overcome by **colipase**, which binds to and activates lipase while holding it to the H_2O–lipid interface. Colipase is secreted by the pancreas as procolipase and is activated by hydrolysis at its amino terminal by trypsin. **Lipid esterase** hydrolyses most lipid esters, e.g. cholesterol esters and monoglycerides, and **phospholipase A_2** hydrolyses phospholipids. Both the lipid esterase and the phospholipase A_2 require bile acids for efficient function.

Bile salts. Bile salts are synthesized in the liver from cholesterol and are released into the duodenum where they aid the digestion and absorption of fats. The bile salts are amphipathic molecules and in solution aggregate to form a **micelle** with the hydrophilic parts of the salts facing the solution, while the hydrophobic parts of the salts face inward. Free fatty acids and monoacylglycerols are slightly more H_2O-soluble than are lipids, and as they equilibrate with the aqueous surroundings they will be incorporated into the bile salt micelle. Cholesterol and phospholipids are also incorporated into the micelle.

ABSORPTION OF FAT

The bile salt micelles transport their contents to the wall of the small intestine and so concentrate the digested fats next to the cells into which they are to be absorbed. The contents of the luminal area adjacent to these cells are very poorly mixed and without the micelles the digested fats would not be able to achieve the steep concentration gradients that are necessary for efficient absorption into the epithelial cells. The absorption of fats occurs mostly in the jejunum, whereas the absorption of bile acids occurs via a Na^+–bile salt cotransport system in the ileum. The salts then pass to the liver where they are recycled.

Inside the cells, the fate of fatty acids depends on the length of their hydrocarbon tail. Fatty acids whose tails are less than 12 carbon atoms long, pass directly into the bloodstream. Fatty acids whose tails are over 12 carbon atoms are conjugated to a binding protein and transported to the smooth ER. Here, they are incorporated into triglycerides which form lipid globules. Modification in the ER and Golgi apparatus results in the formation of specific lipoproteins called chylomicrons.

Transport of absorbed lipids: the exogenous pathway

Chylomicrons in the epithelial cells pass into the intestinal lymph vessels. These lead ultimately to the thoracic duct which empties into the left subclavian vein. This blood passes through the lungs before reaching the capillaries of the peripheral tissues.

Most of the chylomicron constituents are taken up by peripheral tissues, especially adipose and muscle tissues. At the surface of these tissues the triglycerides are hydrolysed by lipoprotein lipase; the fatty acids are then taken up and reassembled into triglycerides inside the cell. The lipoprotein that remains after depletion of most of the chylomicron is a chylomicron remnant and is taken up in the liver by receptor-mediated endocytosis.

PATHOPHYSIOLOGY

Cholesterol and phospholipids are secreted in equal proportions by the liver into the bile. It is important that there are equal proportions of cholesterol and phospholipids because cholesterol is insoluble in aqueous solution and the two are only solubilized in the bile by their incorporation into a bile acid/phospholipid micelle. When too much cholesterol is secreted into the bile, it precipitates out of solution to form crystals. **Gallstones** can eventually form from the crystals and these may become lodged in the cystic duct, causing abdominal pain, vomiting, steatorrhoea and jaundice.