精要速览系列

Instant Notes

先·锋·版

P.C. Turner, A.G. McLennan, A.D. Bates & M.R.H. White

Molecular Biology

分子生物学

（第二版）

影印本

科学出版社
www.sciencep.com

精要速览系列——先锋版

Instant Notes in

Molecular Biology

(Second Edition)

分子生物学

（第二版）

（影印版）

P. C. Turner, A. G. McLennan,
A. D. Bates & M. R. H. White

科学出版社
北京

图字:01-2002-5258

内容简介

“精要速览系列(*Instant Notes Series*)”是国外教材“Best Seller”榜的上榜教材。该系列结构新颖,视角独特;重点明确,脉络分明;图表简明清晰;英文自然易懂,被国内多所重点院校选用作为双语教材。先锋版是继“现代生物学精要速览”之后推出的跨学科的升级版本。

本书是该系列中的《分子生物学(第二版)》分册,全书共19章。新版在原版的基础上内容进行了全面调整和扩充,增加了若干新专题,如中心法则、细胞周期和细胞凋亡等,书后并附有练习题。

本书是指导大学生快速掌握分子生物学基础知识的优秀教材,也是辅助教师授课的极佳教学参考书,同时可供生命科学相关专业的研究生参考。

P. C. Turner, A. G. McLennan, A. D. Bates & M. R. H. White
Instant Notes in Molecular Biology (Second Edition)
Original edition published in the United Kingdom under the title of Instant Notes in Molecular Biology (Second Edition)

图书在版编目(CIP)数据

分子生物学:第2版/(英)特纳(Turner, P. C.)等著.—影印本.—北京:科学出版社,2003.4
(精要速览系列——先锋版)

ISBN 7-03-010959-7

Ⅰ.分… Ⅱ.特… Ⅲ.分子生物学-英文 Ⅳ.Q7

中国版本图书馆CIP数据核字(2002)第093752号

责任编辑:周 辉 黄 方
责任印制:刘士平/封面设计:王 浩

科学出版社 出版
北京东黄城根北街16号
邮政编码:100717
http://www.sciencep.com
新蕾印刷厂 印刷
科学出版社发行 各地新华书店经销

*

1999年3月第 一 版 开本:B5(720×1000)
2003年4月第 二 版 印张:22 1/2
2003年4月第七次印刷 字数:435 960
印数:24 001—28 000

定价:40.00元

(如有印装质量问题,我社负责调换〈环伟〉)

ABBREVIATIONS

ADP	adenosine 5´-diphosphate
AIDS	acquired immune deficiency syndrome
AMP	adenosine 5´-monophosphate
ARS	autonomously replicating sequence
ATP	adenosine 5′-triphosphate
BAC	bacterial artificial chromosome
BER	base excision repair
bp	base pairs
BRF	TFIIB-related factor
BUdR	bromodeoxyuridine
bZIP	basic leucine zipper
CDK	cyclin-dependent kinase
cDNA	complementary DNA
CHEF	contour clamped homogeneous electric field
CJD	Creutzfeld–Jakob disease
CRP	cAMP receptor protein
CSF-1	colony-stimulating factor-1
CTD	carboxyl-terminal domain
Da	Dalton
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxynucleoside triphosphate
DMS	dimethyl sulfate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DOP-PCR	degenerate oligonucleotide primer PCR
dsDNA	double-stranded DNA
EDTA	ethylenediamine tetraacetic acid
EF	elongation factor
ENU	ethylnitrosourea
ER	endoplasmic reticulum
ESI	electrospray ionization
ETS	external transcribed spacer
FADH	reduced flavin adenine dinucleotide
FIGE	field inversion gel electrophoresis
β-gal	β-galactosidase
GMO	genetically modified organism
GTP	guanosine 5′-triphosphate
HIV	human immunodeficiency virus
HLH	helix–loop–helix
hnRNA	heterogeneous nuclear RNA
hnRNP	heterogeneous nuclear ribonucleoprotein
HSP	heat-shock protein
HSV-1	herpes simplex virus-1
ICE	interleukin-1 β converting enzyme
IF	initiation factor
IHF	integration host factor
IPTG	isopropyl-β-D-thiogalactopyranoside
IS	insertion sequence
ITS	internal transcribed spacer
JAK	Janus activated kinase
kb	kilobase pairs in duplex nucleic acid, kilobases in single-stranded nucleic acid
kDa	kiloDalton
LAT	latency-associated transcript
LINES	long interspersed elements
LTR	long terminal repeat
MALDI	matrix-assisted laser desorption/ionization
MCS	multiple cloning site
MMS	methylmethane sulfonate
MMTV	mouse mammary tumor virus
mRNA	messenger RNA
NAD^+	nicotinamide adenine dinucleotide
NER	nucleotide excision repair
NLS	nuclear localization signal
NMN	nicotinamide mononucleotide
NMR	nuclear magnetic resonance
nt	nucleotide
NTP	nucleoside triphosphate
ORC	origin recognition complex
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PAP	poly(A) polymerase
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PFGE	pulsed field gel electrophoresis
PTH	phenylthiohydantoin
RACE	rapid amplification of cDNA ends
RBS	ribosome-binding site
RER	rough endoplasmic reticulum
RF	replicative form

CONTENTS

---

[a] Contributed by Dr M. Bennett, Department of Veterinary Pathology, University of Liverpool, Leahurst, Neston, South Wirral L64 7TE, UK.

[b] Contributed by Dr C. Green, School of Biological Sciences, University of Liverpool, PO Box 147, Liverpool L69 7ZB, UK.

A1 CELLULAR CLASSIFICATION

Key Notes

Eubacteria	Structurally defined as prokaryotes, these cells have a plasma membrane, usually enclosed in a rigid cell wall, but no intracellular compartments. They have a single, major circular chromosome. They may be unicellular or multicellular. *Escherichia coli* is the best studied eubacterium.
Archaea	The Archaea are structurally defined as prokaryotes but probably branched off from the eukaryotes after their common ancestor diverged from the eubacteria. They tend to inhabit extreme environments. They are biochemically closer to eubacteria in some ways but to eukaryotes in others. They also have some biochemical peculiarities.
Eukaryotes	Cells of plants, animals, fungi and protists possess well-defined subcellular compartments bounded by lipid membranes (e.g. nuclei, mitochondria, endoplasmic reticulum). These organelles are the sites of distinct biochemical processes and define the eukaryotes.
Differentiation	In most multicellular eukaryotes, groups of cells differentiate during development of the organism to provide specialized functions (e.g. as in liver, brain, kidney). In most cases, they contain the same DNA but transcribe different genes. Like all other cellular processes, differentiation is controlled by genes. Co-ordination of the activities of different cell types requires communication between them.

Related topics

Subcellular organelles (A2)
Prokaryotic and eukaryotic chromosome structure (Section D)
Bacteriophages and eukaryotic viruses (Section R)

Eubacteria

The **Eubacteria** are one of two subdivisions of the **prokaryotes**. Prokaryotes are the simplest living cells, typically 1–10 μm in diameter, and are found in all environmental niches from the guts of animals to acidic hot springs. Classically, they are defined by their structural organization (*Fig. 1*). They are bounded by a **cell (plasma) membrane** comprising a lipid bilayer in which are embedded proteins that allow the exit and entry of small molecules. Most prokaryotes also have a rigid cell wall outside the plasma membrane which prevents the cell from swelling or shrinking in environments where the osmolarity differs significantly from that inside the cell. The cell interior (**cytoplasm** or **cytosol**) usually contains a single, circular chromosome compacted into a **nucleoid** and attached to the membrane (see Topic D1), and often plasmids [small deoxyribonucleic acid (DNA) molecules with limited genetic information, see Topic G2], ribonucleic acid (RNA), ribosomes (the sites of protein synthesis, see Section Q) and most of the proteins which perform the metabolic reactions of the cell. Some of these proteins are attached to the plasma membrane, but there are no distinct **subcel-**

lular organelles as in **eukaryotes** to compartmentalize different parts of the metabolism. The surface of a prokaryote may carry **pili**, which allow it to attach to other cells and surfaces, and **flagella**, whose rotating motion allows the cell to swim. Most prokaryotes are unicellular; some, however, have multicellular forms in which certain cells carry out specialized functions. The **Eubacteria** differ from the **Archaea** mainly in their biochemistry. The eubacterium *Escherichia coli* has a **genome size** (DNA content) of 4600 kilobase pairs (kb) which is sufficient genetic information for about 3000 proteins. Its molecular biology has been studied extensively. The genome of the simplest bacterium, *Mycoplasma genitalium*, has only 580 kb of DNA and encodes just 470 proteins. It has a very limited metabolic capacity.

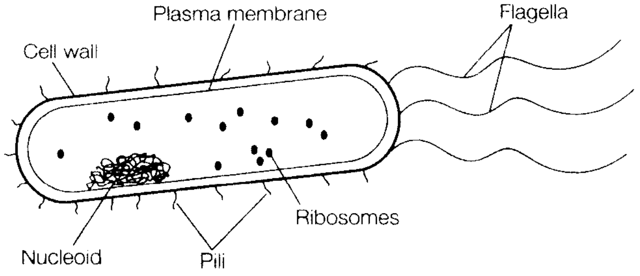

Fig. 1. Schematic diagram of a typical prokaryotic cell.

Archaea

The **Archaea**, or **archaebacteria**, form the second subdivision of the prokaryotes and tend to inhabit extreme environments. Structurally, they are similar to eubacteria. However, on the basis of the evolution of their ribosomal RNA (rRNA) molecules (see Topic O1), they appear as different from the eubacteria as both groups of prokaryotes are from the eukaryotes and display some unusual biochemical features, for example ether in place of ester linkages in membrane lipids (see Topic A3). The 1740 kb genome of the archaeon *Methanococcus jannaschii* encodes a maximum of 1738 proteins. Comparisons reveal that those involved in energy production and metabolism are most like those of eubacteria while those involved in replication, transcription and translation are more similar to those of eukaryotes. It appears that the Archaea and the eukaryotes share a common evolutionary ancestor which diverged from the ancestor of the Eubacteria.

Eukaryotes

Eukaryotes are classified taxonomically into four kingdoms comprising **animals**, **plants**, **fungi** and **protists** (algae and protozoa). Structurally, eukaryotes are defined by their possession of membrane-enclosed organelles (*Fig.* 2) with specialized metabolic functions (see Topic A2). Eukaryotic cells tend to be larger than prokaryotes: 10–100 μm in diameter. They are surrounded by a plasma membrane, which can have a highly convoluted shape to increase its surface area. Plants and many fungi and protists also have a rigid cell wall. The cytoplasm is a highly organized gel that contains, in addition to the organelles and ribosomes, an array of protein fibers called the **cytoskeleton** which controls the shape and movement of the cell and which organizes many of its metabolic functions. These fibers include **microtubules**, made of **tubulin**, and **microfilaments**, made of **actin** (see Topic A4). Many eukaryotes are multicellular, with groups of cells undergoing **differentiation** during development to form the specialized tissues of the whole organism.

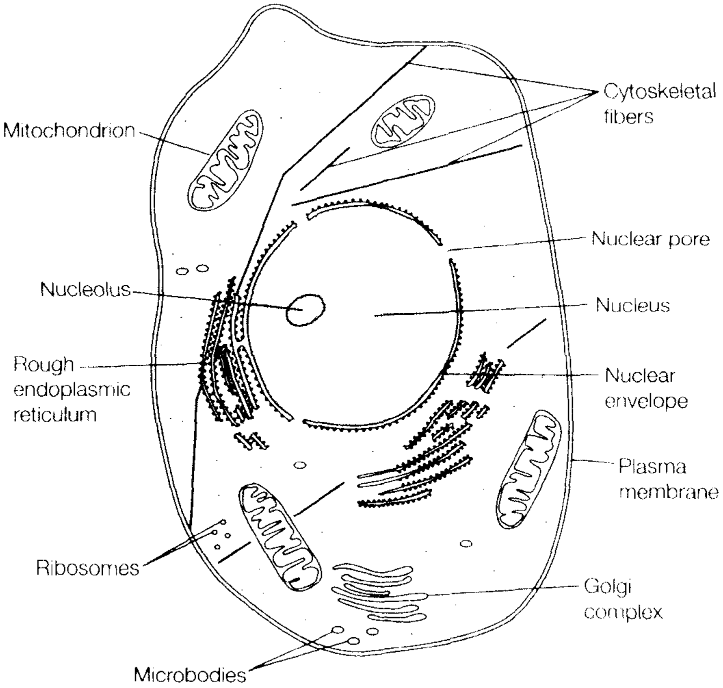

Fig. 2. Schematic diagram of a typical eukaryotic cell.

Differentiation

When a cell divides, the daughter cells may be identical in every way, or they may change their patterns of gene expression to become functionally different from the parent cell.

Among prokaryotes and lower eukaryotes, the formation of spores is an example of such **cellular differentiation** (see Topic L3). Among complex multicellular eukaryotes, embryonic cells differentiate into highly specialized cells, for example muscle, nerve, liver and kidney. In all but a few exceptional cases, the DNA content remains the same, but the genes which are transcribed have changed. Differentiation is regulated by developmental control genes (see Topic N2). Mutations in these genes result in abnormal body plans, such as legs in the place of antennae in the fruit fly *Drosophila*. Studying such gene mutations allows the process of embryonic development to be understood. In multicellular organisms, co-ordination of the activities of the various tissues and organs is controlled by communication between them. This involves signaling molecules such as neurotransmitters, hormones and growth factors which are secreted by one tissue and act upon another through specific cell-surface receptors.

A2 SUBCELLULAR ORGANELLES

Key Notes

Nuclei	The membrane-bound nucleus contains the bulk of the cellular DNA in multiple chromosomes. Transcription of this DNA and processing of the RNA occurs here. Nucleoli are contained within the nucleus.
Mitochondria and chloroplasts	Mitochondria are the site of cellular respiration where nutrients are oxidized to CO_2 and water, and adenosine 5′-triphosphate (ATP) is generated. They are derived from prokaryotic symbionts and retain some DNA, RNA and protein synthetic machinery, though most of their proteins are encoded in the nucleus. Photosynthesis takes place in the chloroplasts of plants and eukaryotic algae. Chloroplasts have a basically similar structure to mitochondria but with a thylakoid membrane system containing the light-harvesting pigment chlorophyll.
Endoplasmic reticulum	The smooth endoplasmic reticulum is a cytoplasmic membrane system where many of the reactions of lipid biosynthesis and xenobiotic metabolism are carried out. The rough endoplasmic reticulum has attached ribosomes engaged in the synthesis of membrane-targeted and secreted proteins. These proteins are carried in vesicles to the Golgi complex for further processing and sorting.
Microbodies	The lysosomes contain degradative, hydrolytic enzymes; the peroxisomes contain enzymes which destroy certain potentially dangerous free radicals and hydrogen peroxide; the glyoxysomes of plants carry out the reactions of the glyoxylate cycle.
Organelle isolation	After disruption of the plasma membrane, the subcellular organelles can be separated from each other and purified by a combination of differential centrifugation and density gradient centrifugation (both rate zonal and isopycnic). Purity can be assayed by measuring organelle-specific enzymes.

Related topics

Cellular classification (A1)
rRNA processing and ribosomes (O1)
Translational control and post-translational events (Q4)

Nuclei

The eukaryotic nucleus carries the genetic information of the cell in multiple chromosomes, each containing a single DNA molecule (see Topics D2 and D3). The nucleus is bounded by a lipid double membrane, the nuclear envelope, containing pores which allow passage of moderately large molecules (see Topic A1, *Fig. 2*). Transcription of RNA takes place in the nucleus (see Section M) and the processed RNA molecules (see Section O) pass into the cytoplasm where translation takes place (see Section Q). **Nucleoli** are bodies within the nucleus where rRNA is synthesized and ribosomes are partially assembled (see Topics M2 and O1).

Mitochondria and chloroplasts

Cellular respiration, that is the oxidation of nutrients to generate energy in the form of adenosine 5′-triphosphate (ATP), takes place in the **mitochondria**. These organelles are roughly 1–2 μm in diameter and there may be 1000–2000 per cell. They have a smooth outer membrane and a convoluted inner membrane that forms protrusions called **cristae** (see Topic A1, *Fig. 2*). They contain a small circular DNA molecule, mitochondrial-specific RNA and ribosomes on which some mitochondrial proteins are synthesized. However, the majority of mitochondrial (and chloroplast) proteins are encoded by nuclear DNA and synthesized in the cytoplasm. These latter proteins have specific **signal sequences** that target them to the mitochondria (see Topic Q4). The **chloroplasts** of plants are the site of photosynthesis, the light-dependent assimilation of CO_2 and water to form carbohydrates and oxygen. Though larger than mitochondria, they have a similar structure except that, in place of cristae, they have a third membrane system (the **thylakoids**) in the inner membrane space. These contain chlorophyll, which traps the light energy for photosynthesis. Chloroplasts are also partly genetically independent of the nucleus. Both mitochondria and chloroplasts are believed to have evolved from prokaryotes which had formed a symbiotic relationship with a primitive nucleated eukaryote.

Endoplasmic reticulum

The **endoplasmic reticulum** is an extensive membrane system within the cytoplasm and is continuous with the nuclear envelope (see Topic A1, *Fig. 2*). Two forms are visible in most cells. The **smooth** endoplasmic reticulum carries many membrane-bound enzymes, including those involved in the biosynthesis of certain lipids and the oxidation and detoxification of foreign compounds (**xenobiotics**) such as drugs. The **rough** endoplasmic reticulum (**RER**) is so-called because of the presence of many ribosomes. These ribosomes specifically synthesize proteins intended for secretion by the cell, such as plasma or milk proteins, or those destined for the plasma membrane or certain organelles. Apart from the plasma membrane proteins, which are initially incorporated into the RER membrane, these proteins are translocated into the interior space (**lumen**) of the RER where they are modified, often by **glycosylation** (see Topic Q4). The lipids and proteins synthesized on the RER are transported in specialized **transport vesicles** to the **Golgi complex**, a stack of flattened membrane vesicles which further modifies, sorts and directs them to their final destinations (see Topic A1, *Fig. 2*).

Microbodies

Lysosomes are small membrane-bound organelles which bud off from the Golgi complex and which contain a variety of digestive enzymes capable of degrading proteins, nucleic acids, lipids and carbohydrates. They act as recycling centers for macromolecules brought in from outside the cell or from damaged organelles. Some metabolic reactions which generate highly reactive free radicals and hydrogen peroxide are confined within organelles called **peroxisomes** to prevent these species from damaging cellular components. Peroxisomes contain the enzyme catalase, which destroys hydrogen peroxide:

$$2H_2O_2 \rightarrow 2H_2O + O_2$$

Glyoxysomes are specialized plant peroxisomes which carry out the reactions of the glyoxylate cycle. Lysosomes, peroxisomes and glyoxysomes are collectively known as **microbodies**.

Organelle isolation

The plasma membrane of eukaryotes can be disrupted by various means including osmotic shock, controlled mechanical shear or by certain nonionic detergents. Organelles displaying large size and density differences, for example nuclei and mitochondria, can be separated from each other and from other organelles by **differential centrifugation** according to the value of their **sedimentation coefficients** (see Topic A4). The **cell lysate** is centrifuged at a speed which is high enough to sediment only the heaviest organelles, usually the nuclei. The supernatant containing all the other organelles is removed then centrifuged at a higher speed to sediment the mitochondria, and so on (*Fig. 1a*). This technique is also used to fractionate suspensions containing cell types of different sizes, for example red cells, white cells and platelets in blood. These crude preparations of cells, nuclei and mitochondria usually require further purification by **density gradient centrifugation**. This is also used to separate organelles of similar densities. In **rate zonal centrifugation**, the mixture is layered on top of a pre-formed concentration (and, therefore, density) gradient of a suitable medium in a centrifuge tube. Upon centrifugation, bands or zones of the different components sediment at different rates depending on their sedimentation coefficients, and separate (*Fig. 1b*). The purpose of the density gradient of the supporting medium is to prevent convective mixing of the components after separation (i.e. to provide stability) and to ensure linear sedimentation rates of the components (it compensates for the acceleration of the components as they move further down the tube). In **equilibrium (isopycnic) centrifugation**, the density gradient extends to a density higher than that of one or more components of the mixture so that these components come to equilibrium at a point equal to their own density and stop moving. In this case, the density gradient can either be pre-formed, and the sample layered on top, or self-forming, in which case the sample may be mixed with the gradient material (*Fig. 1c*). Density gradients are made from substances such as sucrose, Ficoll (a synthetic polysaccharide), metrizamide (a synthetic iodinated heavy compound) or cesium chloride (CsCl), for separation of nucleic acids (see Topics C2 and G2). Purity of the subcellular fraction can be determined using an electron microscope or by assaying enzyme activities known to be associated specifically with particular organelles, for example succinate dehydrogenase in mitochondria.

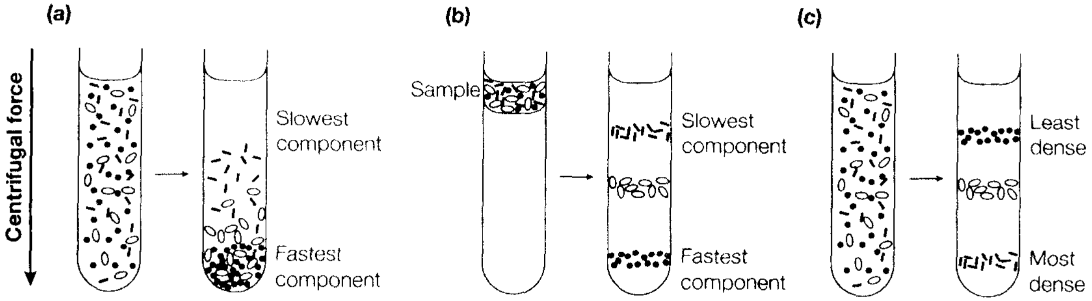

Fig. 1. Centrifugation techniques. (a) Differential, (b) rate zonal and (c) isopycnic (equilibrium).

A3 MACROMOLECULES

Key Notes

Proteins and nucleic acids

Proteins are polymers of amino acids, and the nucleic acids DNA and RNA are polymers of nucleotides. They are both essential components of the machinery which stores and expresses genetic information. Proteins have many additional structural and functional roles.

Polysaccharides

α-Amylose and cellulose are polymers of glucose linked $\alpha(1\rightarrow4)$ and $\beta(1\rightarrow4)$ respectively. Starch, a storage form of glucose found in plants, contains α-amylose together with the $\alpha(1\rightarrow6)$ branched polymer amylopectin. Cellulose forms strong structural fibers in plants. Glucose is stored as glycogen in animals. Chitin, a polymer of *N*-acetylglucosamine, is found in fungal cell walls and arthropod exoskeletons. Mucopolysaccharides are important components of connective tissue.

Lipids

Triglycerides containing saturated and unsaturated fatty acids are the major storage lipids of animals and plants respectively. Structural differences between the two types of fatty acid result in animal triglycerides being solid and plant triglycerides (oils) liquid. Phospholipids and sphingolipids have polar groups in addition to the fatty acid components, and are important constituents of all cell membranes.

Complex macromolecules

Nucleoproteins contain both nucleic acid and protein, as in the enzymes telomerase and ribonuclease P. Glycoproteins and proteoglycans (mucoproteins) are proteins with covalently attached carbohydrate and are generally found on extracellular surfaces and in extracellular spaces. Lipid-linked proteins and lipoproteins have lipid and protein components attached covalently or noncovalently respectively. Glycolipids have both lipid and carbohydrate parts. Mixed macromolecular complexes such as these provide a wider range of functions than the component parts.

Related topics

Large macromolecular assemblies (A4)
Protein structure (Section B)
Properties of nucleic acids (Section C)

Proteins and nucleic acids

Proteins are polymers of amino acids linked together by peptide bonds. The structures of the amino acids and of proteins are dealt with in detail in Section B. Proteins have both structural and functional roles. The nucleic acids DNA and RNA are polymers of nucleotides, which themselves consist of a nitrogenous base, a pentose sugar and phosphoric acid. Their structures are detailed in Section C. There are three main types of cellular RNA: messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA). Nucleic acids are involved in the storage and processing of genetic information (see Topic D5), but the expression of this information requires proteins.

Polysaccharides

Polysaccharides are polymers of simple sugars covalently linked by glycosidic bonds. They function mainly as nutritional sugar stores and as structural materials. **Cellulose** and **starch** are abundant components of plants. Both are glucose polymers, but differ in the way the glucose monomers are linked. Cellulose is a linear polymer with β(1→4) linkages (*Fig. 1a*) and is a major structural component of the plant cell wall. About 40 parallel chains form horizontal sheets which stack vertically above one another. The chains and sheets are held together by hydrogen bonds (see Topic A4) to produce tough, insoluble fibers. Starch is a sugar store and is found in large intracellular granules which can be hydrolyzed quickly to release glucose for metabolism. It contains two components: α-amylose, a linear polymer with α(1→4) linkages (*Fig. 1b*), and amylopectin, a branched polymer with additional α(1→6) linkages. With up to 10^6 glucose residues, amylopectins are among the largest molecules known. The different linkages in starch produce a coiled conformation that cannot pack tightly, hence starch is water-soluble. Fungi and some animal tissues (e.g. liver and muscle) store glucose as **glycogen**, a branched polymer like amylopectin. **Chitin** is found in fungal cell walls and in the exoskeleton of insects and crustacea. It is similar to cellulose, but the monomer unit is *N*-acetylglucosamine. **Mucopolysaccharides** (glycosaminoglycans) form the gel-like solutions in which the fibrous proteins of connective tissue are embedded. Determination of the structures of large polysaccharides is complicated because they are heterogeneous in size and composition and because they cannot be studied genetically like nucleic acids and proteins.

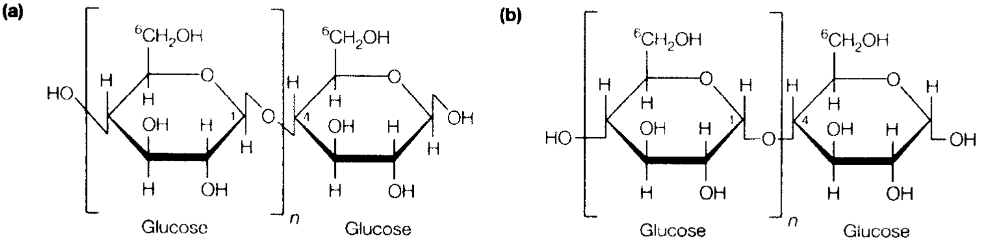

Fig. 1. The structures of (a) cellulose, with β(1→4) linkages and (b) starch α-amylose, with α(1→4) linkages. Carbon atoms 1, 4 and 6 are labeled. Additional α(1→6) linkages produce branches in amylopectin and glycogen.

Lipids

While individual lipids are not strictly macromolecules, many are built up from smaller monomeric units and they are involved in many macromolecular assemblies (see Topic A4). Large lipid molecules are predominantly hydrocarbon in nature and are poorly soluble in water. Some are involved in the storage and transport of energy while others are key components of membranes, protective coats and other cell structures. **Glycerides** have one, two or three long-chain fatty acids esterified to a molecule of glycerol. In animal triglycerides, the fatty acids have no double bonds (saturated) so the chains are linear, the molecules can pack tightly and the resulting fats are solid. Plant oils contain **unsaturated** fatty acids with one or more double bonds. The angled structures of these chains prevent close packing so they tend to be liquids at room temperature. Membranes contain **phospholipids**, which consist of glycerol esterified to two fatty acids and phosphoric acid. The phosphate is also usually esterified to a small molecule such as serine, ethanolamine, inositol or choline (*Fig. 2*). Membranes also contain **sphingolipids** such as **ceramide**, in which the long-chain amino alcohol sphingosine has a fatty acid linked by an amide bond. Attachment of phosphocholine to a ceramide produces **sphingomyelin**.

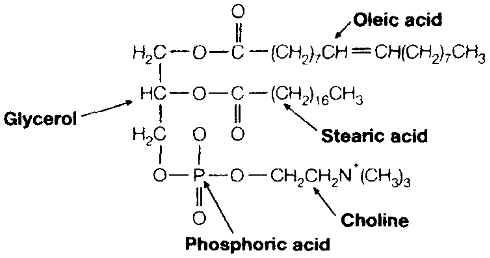

Fig. 2. A typical phospholipid: phosphatidylcholine containing esterified stearic and oleic acids.

Complex macromolecules

Many macromolecules contain covalent or noncovalent associations of more than one of the major classes of large biomolecules. This can greatly increase the functionality or structural capabilities of the resulting complex. For example, nearly all enzymes are proteins, but some have a noncovalently attached RNA component which is essential for catalytic activity. Associations of nucleic acid and protein are known as **nucleoproteins**. Examples are **telomerase**, which is responsible for replicating the ends of eukaryotic chromosomes (See Topics D3 and E3) and **ribonuclease P**, an enzyme which matures transfer RNA (tRNA). In telomerase, the RNA acts as a template for telomere DNA synthesis, while in ribonuclease P, the RNA contains the catalytic site of the enzyme. Ribonuclease P is an example of a **ribozyme** (see Topic O2).

Glycoproteins contain both protein and carbohydrate (between $<1\%$ and $>90\%$ of the weight) components; glycosylation is the commonest form of post-translational modification of proteins (see Topic Q4). The carbohydrate is always covalently attached to the surface of the protein, never the interior, and is often variable in composition, causing microheterogeneity (*Fig. 3*). This has made glycoproteins difficult to study. Glycoproteins have functions that span the entire range of protein activities, and are usually found extracellularly. They are important components of cell membranes and mediate cell–cell recognition.

Proteoglycans (mucoproteins) are large complexes ($>10^7$ Da) of protein and mucopolysaccharide found in bacterial cell walls and in the extracellular space in connective tissue. Their sugar units often have sulfate groups, which makes

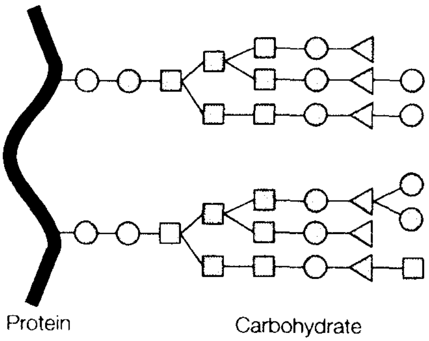

Fig. 3. Glycoprotein structure. The different symbols represent different monosaccharide units (e.g. galactose, N-acetylglucosamine).

them highly hydrated. This, coupled with their lengths (> 1000 units), produces solutions of high viscosity. Proteoglycans act as lubricants and shock absorbers in extracellular spaces.

Lipid-linked proteins have a covalently attached lipid component. This is usually a fatty acyl (e.g. myristoyl or palmitoyl) or isoprenoid (e.g. farnesyl or geranylgeranyl) group. These groups serve to anchor the proteins in membranes through hydrophobic interactions with the membrane lipids and also promote protein–protein associations (see Topic A4).

In **lipoproteins**, the lipids and proteins are linked noncovalently. Because lipids are poorly soluble in water, they are transported in the blood as lipoproteins. These are basically particles of triglycerides and cholesterol esters coated with a layer of phospholipids, cholesterol and protein (the **apolipoproteins**). The structures of the apolipoproteins are such that their hydrophobic amino acids face towards the lipid interior of the particles while the charged and polar amino acids (see Topic B1) face outwards into the aqueous environment. This renders the particles soluble.

Glycolipids, which include cerebrosides and gangliosides, have covalently linked lipid and carbohydrate components, and are especially abundant in the membranes of brain and nerve cells.