WILEY-VCH

Edited by Michael Wink

An Introduction to Molecular Biotechnology

Fundamentals, Methods and Applications

Third, Completely Revised Edition

An Introduction to Molecular Biotechnology

An Introduction to Molecular Biotechnology

Fundamentals, Methods and Applications

Edited by Michael Wink

Third, Completely Revised Edition

WILEY-VCH

Editor
Michael Wink
Universität Heidelberg
Institut für Pharmazie und Molekulare
Biotechnologie (IPMB)
Im Neuenheimer Feld 329
69120 Heidelberg
Germany

Cover credits
Shutterstock #287352494 / paulista

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Print ISBN: 978-3-527-34414-7
ePDF ISBN: 978-3-527-81286-8
ePub ISBN: 978-3-527-81288-2

Cover Design Formgeber, Mannheim, Germany
Typesetting SPi Global, Chennai, India
Printing and Binding

Printed on acid-free paper

10 9 8 7 6 5 4 3 2 1

Contents

Abbreviations

1 Å	=0.1 nm
aa-tRNA	aminoacyl-tRNA
AAV	adeno-associated virus
ABC	ATP-binding cassette
Acetyl CoA	acetyl coenzyme A
AcNPV	*Autographa californica* nuclear polyhedrosis virus
ACRS	amplification-created restriction sites
ACTH	adrenocorticotropic hormone
ADA	adenosine deaminase
ADEPT	antibody-directed enzyme prodrug therapy
ADME-T	absorption, distribution, metabolism, excretion, and toxicity
ADP	adenosine diphosphate
ADRs	adverse drug reactions
AEC	aminoethylcysteine
AFLP	amplified fragment length polymorphism
AFM	atomic force microscope
AIDS	acquired immunodeficiency syndrome
ALS	amyotrophic lateral sclerosis
AMP	adenosine monophosphate
AMPA	α-amino-3-hydroxyl-5-methyl-4-isoxazol-propionate
Amp^r	ampicillin resistance gene
AMV	avian myeloblastosis virus
ANN	artificial neural network
AO	acridine orange
AOX1	alcohol oxidase 1
APC	anaphase-promoting complex
ApoB100	apolipoprotein B100
ApoE	apolipoprotein E
APP	amyloid precursor protein
ARMS	amplification refractory mutation system
ARS	autonomously replicating sequence
ATP	adenosine triphosphate
att	attachment site
BAC	bacterial artificial chromosome
bcl2	B-cell leukemia lymphoma 2 (protein protecting against apoptosis)
BfArM	German Bundesinstitut für Arzneimittel und Medizinprodukte
β-Gal	β-galactosidase
BHK-21	baby hamster kidney cells
BLA	biologics licence application
BLAST	Basic Local Alignment Search Tool
BMP	bone morphogenetic proteins
bp	base pairs
BrdU	bromodeoxyuridine
CA	correspondence analysis
CAD	coronary artery disease
CaM-Kinase	Ca^{2+}/calmodulin-dependent protein kinase
cAMP	cyclic AMP
cap	AAV gene mediating encapsulation
CARS	coherent anti-Raman scattering
CAT	Committee for Advanced Therapies
CBER	Center for Biologics Evaluation and Research
CC	chromatin remodeling complex
CCD	charge-coupled device
CDER	Center for Drug Evaluation and Research
CDK	cyclin-dependent kinase
cDNA	copy DNA
CDR	complementarity-determining region
CDRH	Center for Devices and Radiological Health
CEO	chief executive officer
CFP	cyan fluorescent protein

CFTR	cystic fibrosis transmembrane regulator
CGAP	Cancer Genome Anatomy Project
CGH	comparative genome hybridization
CHMP	Committee for Medicinal Products for Human Use
CHO	Chinese hamster ovary
CIP	calf intestinal phosphatase
CML	chronic myeloid leukemia
CMN	*Corynebacterium–Mycobacterium–Nocardia* group
CaMV	cauliflower mosaic virus
CMV	cytomegalovirus
CNS	central nervous system
COMP	Committee for Orphan Medicinal Products
COS-1	simian cell line, CV-1, transformed by origin-defective mutant of SV40
cpDNA	chloroplast DNA
CPMV	cowpea mosaic virus
cPPT-sequence	central polypurine tract – regulatory element in lentiviral vectors that facilitates double strand synthesis and the nuclear import of the pre-integration complex
CSF	colony-stimulating factor
CSO	contract service organization
CTAB	cetyltrimethylammonium bromide
CVM	Center for Veterinary Medicine
CVMP	Committee for Medicinal Products for Veterinary Use
2D	two-dimensional
Da	Dalton
DAG	diacylglycerol
DAPI	4,6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
DBD	DNA-binding domain
DAC	divide-and-conquer strategy
DD	differential display
DDBJ	DNA Data Bank of Japan
ddNTP	dideoxynucleotide triphosphate
DEAE	diethylaminoethyl
dHPLC	denaturing HPLC
DIC	differential interference contrast
DIP	Database of Interacting Proteins
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
Dox	doxycycline
ds diabodies	disulfide-stabilized diabodies
dsDNA	double-stranded DNA
dsFv-fragment	disulfide-stabilized Fv fragment
dsRNA	double-stranded RNA
DtxR	diphtheria toxin repressor
Ebola-Z	envelope protein of the Ebola-Zaire virus, which has a high affinity to lung epithelial cells
EC_{50}	effective concentration, the dose or concentration that produces a 50% effect in the test population within a specified time
ECD	electron capture dissociation
EDTA	ethylenediaminetetraacetic acid
ee	enantiomeric excess
EF2	elongation factor 2
EF-Tu	elongation factor Tu
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EGTA	ethylene glycol bis(2-aminoethyl)tetraacetic acid
EIAV	equine infectious anemia virus
ELISA	enzyme-linked immunosorbent assay
EM	electron microscope
EMA	European Medicines Agency
EMBL	European Molecular Biology Laboratory
EMCV	encephalomyocarditis virus
EMSA	electrophoretic mobility shift assay
EMEA	European Agency for the Evaluation of Medicinal Products
ENU	*N*-ethyl-*N*-nitrosourea
env	retroviral gene coding for viral envelope proteins
EPO	European Patent Office
EPR effect	enhanced permeability and retention effect
EPC	European Patent Convention
ER	endoplasmic reticulum
ESI	electrospray ionization
EST	expressed sequence tags
ES cells	embryonic stem cells
EtBr	ethidium bromide
Fab-fragment	antigen-binding fragment
FACS	fluorescence-activated cell sorter

FAD	flavin adenine dinucleotide
FBA	flux balance analysis
FCS	fluorescence correlation spectroscopy
FDA	Food and Drug Administration
FFL	feed-forward loop
FGF	fibroblast growth factor
FISH	fluorescence *in situ* hybridization
FIV	feline immunodeficiency virus
FKBP	FK506-binding protein
FLIM	fluorescence lifetime imaging microscopy
FLIPR	fluorescent imaging plate reader
FMN	flavin mononucleotide
FPLC	fast performance liquid chromatography
FRAP	fluorescence recovery after photobleaching
FRET	fluorescence resonance energy transfer
FT-ICR	Fourier transformation cyclotron resonance, method in mass spectroscopy
FtsZ	prokaryotic cell division protein
Fur	ferric uptake regulator
Fv-fragment	variable fragment
FWHM	full width at half maximum
GABA	gamma-aminobutyric acid
Gag	retroviral gene coding for structural proteins
Gal	galactose
GAP	GTPase-activating protein
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
Gb	gigabases
GCC	German cDNA consortium
GCG	Genetics Computer Group
GCP	good clinical practice
ΔG_d	free enthalpy
GDH	glutamate dehydrogenase
GDP	guanosine diphosphate
GEF	guanine exchange factor
GEO	gene expression omnibus
GFP	green fluorescent protein
GM-CSF	granulocyte/macrophage colony-stimulating factor
GO	Gene Ontology
GOI	gene of interest
GPCR	G-protein-coupled receptor
GPI anchor	glycosylphosphatidylinositol anchor
GRAS	generally regarded as safe
GST	glutathione-S-transferase
GTC	guanidinium isothiocyanate
GTP	guanosine triphosphate
GUS	glucuronidase
GMO	genetically modified organism
HA	hemagglutinin
HCM	hypertrophic cardiomyopathy
HCV	hepatitis C virus
HEK	human embryonic kidney
HeLa cells	human cancer cell line (isolated from donor Helene Larsen)
HER 2	human epidermal growth factor 2
HGH	human growth hormone
HIC	hydrophobic interaction chromatography
His_6	hexahistidine tag
HIV	human immunodeficiency virus, a retrovirus
HIV 1	human immunodeficiency virus 1
HLA	human leukocyte antigen
hnRNA	heterogeneous nuclear RNA
HPLC	high-performance liquid chromatography
HPT	hygromycin phosphotransferase
HPV	human papillomavirus
HSP	high-scoring segment pairs
HSP	heat shock protein
HSV-1	herpes simplex virus
HTS	high-throughput analysis
HUGO	Human Genome Organization
HV	herpesvirus
IAS	international accounting standard
ICDH	isocitric dehydrogenase
ICH	International Council for Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICL	isocitric lyase
ICP-MS	inductively coupled plasma mass spectrometry
ICR-MS	ion cyclotron resonance mass spectrometer
IDA	iminodiacetic acid
IEF	isoelectric focusing
Ig	immunoglobulin
IHF	integration host factor
IMAC	immobilized metal affinity chromatography
IND-Status	investigational new drug status
IP_3	inositol-1,4,5-triphosphate
IPO	initial public offering
IPTG	isopropyl-β-D-thiogalactoside

IRs	inverted repeats
IR	investor relations
IRES	internal ribosome entry site
ISAAA	International Service for the Acquisition of Agri-biotech Applications
ISH	*in situ* hybridization
ISSR	inter-simple sequence repeats
ITC	isothermal titration calorimetry
ITR	inverse terminal repeats – regulatory elements in adenoviruses and AAV
i.v.	intravenous
k_a	second-order velocity constant in bimolecular association
Kan^r	kanamycin resistance gene
K_{av}	specific distribution coefficient
Kb	kilobases
k_d	first-order velocity constant in unimolecular dissociation
$K_d = k_d/k_a$	velocity constant in dissociation/K_a in association
KDa	kilodalton
KDEL	amino acid sequence for proteins remaining in the ER
KDR receptor	kinase insert domain-containing receptor
KEGG	Kyoto Encyclopedia of Genes and Genomes
Lac	lactose
LASER	light amplification by stimulated emission of radiation
LB	left border
LB	Luria-Bertani medium
LCR	ligation chain reaction
LDL	low-density lipoprotein
LIMS	laboratory information management systems
LINE	long interspersed elements
LSC	laser scanning cytometer
LTQ	linear trap quadrupole
LTQ-FT-ICR	linear trap quadrupole–Fourier transformation ion cyclotron resonance
LTR	long terminal repeats; regulatory elements in retroviruses
LUMIER	LUMInescence-based Mammalian intERactome
MAC	mammalian artificial chromosome
mAChR	muscarinic acetylcholine receptor
MAGE-ML	microarray gene expression markup language
MALDI	matrix-assisted laser desorption/ionization
6-MAM	6-monoacetylmorphine
MAP	microtubule-associated protein
MAP	mitosis-activating protein
Mb	megabases
MBP	maltose-binding protein
MCS	multiple cloning site
M-CSF	macrophage colony-stimulating factor
MDR	multidrug resistance protein
MDS	multidimensional scaling
MGC	Mammalian Gene Collection
MHC	major histocompatibility complex
MIAME	minimum information about a microarray experiment
miRNA	microRNA
MIT	Massachusetts Institute of Technology
MoMLV	Moloney murine leukemia virus
Mowse	molecular weight search
MPF	M-phase promotion factor
MPSS	Massively Parallel Signature Sequencing
Mreb/Mbl	proteins of prokaryotic cytoskeleton
mRNA	messenger RNA
MRSA	methicillin-resistant *Staphylococcus aureus*
MS	mass spectrometry
MSG	monosodium glutamate
MS-PCR	mutationally separated PCR
MTA	material transfer agreement
mtDNA	mitochondrial DNA
MULVR	Moloney murine leukemia virus
MW	molecular weight
µF	µFarad
nAChR	nicotinic acetylcholine receptor
NAD	nicotinamide adenine dinucleotide
NAPPA	nucleic acid programmable protein array
NCBI	National Center for Biotechnology Information
NDA	new drug application
NDP	nucleoside diphosphate
NDPK	nucleoside diphosphates kinase
NFjB	nuclear factor jB
NIH	National Institutes of Health
NK cell	natural killer cell
NMDA receptor	*N*-methyl-D-aspartate receptor

NMR	nuclear magnetic resonance
NPTII	neomycin phosphotransferase II
NSAID	nonsteroidal anti-inflammatory drug
NTA	nitrilotriacetic acid
NTP	nucleoside triphosphate
OD	optical density
ODE	ordinary differential equation
ODHC	2-oxoglutarate dehydrogenase
OMIM	Online Mendelian Inheritance in Man
ORF	open reading frame
ori	origin of replication
OXA complex	membrane translocator in mitochondria
PAC	P1-derived artificial chromosome
PAGE	polyacrylamide gel electrophoresis
PAZ domain	*PIWI/Argonaute/Zwille domain*
PCA	principal component analysis
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PFAM	protein families database of alignments and HMMs
PFG	pulsed-field gel electrophoresis
PI	propidium iodide
PIR	protein information resource
piRNA	piwi-interacting RNA
PKA	protein kinase A
PKC	protein kinase C
PK data	pharmacokinetic data
PLoS	Public Library of Science
PMSF	phenylmethylsulfonyl fluoride
PNA	peptide nucleic acid
PNGaseF	peptide *N*-glycosidase F
PNK	T4 polynucleotide kinase
pol	retroviral gene coding for reverse transcriptase and integrase
P_{PH}	polyhedrin promoter
PR	public relations
psi	retroviral packaging signal
PTGS	posttranscriptional gene silencing
PTI	pancreatic trypsin inhibitor
Q-FT-ICR	q-Fourier transform ion cyclotron resonance
Q-TOF	quadrupole time-of-flight
RACE	rapid amplification of cDNA ends
Ran	protein involved in nuclear import
RAPD	random amplification of polymorphic DNA
RAP-PCR	RNA arbitrarily primed PCR
RB	right border
RBD	RNA-binding domain
Rb gene	retinoblastoma gene
RBS	ribosome-binding site
RDA	representative difference analysis
RdRp	RNA-dependent RNA polymerase
rep	*AAV* gene mediating replication
RES	reticuloendothelial system
RFLP	restriction fragment length polymorphism
R_f-value	retention factor
RGS	regulator of G-protein signaling
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
rpm	revolutions per minute
RRE	regulatory element in a lentiviral vector, enhancing the nuclear export of viral RNA
rRNA	ribosomal RNA
RSV	respiratory syncytial virus
RSV	promoter of the Rous sarcoma virus
RT	reverse transcriptase
rtTA	tetracycline-sensitive regulatory unit
SAGE	serial analysis of gene expression
SALM	spectrally assigned localization microscopy
SAM	*S*-adenosylmethionine
sc diabodies	single-chain diabodies
scFab	single-chain Fab fragment
scFv/sFv fragment	single-chain Fv fragment
SCID	severe combined immunodeficiency
SCOP	structural classification of proteins
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELEX	systematic evolution of ligands by exponential enrichment
SEM	scanning electron microscope
Sf cells	*Spodoptera frugiperda* cells
SFM	scanning force microscope
SFV	Semliki Forest virus

SH1	Src homology domain 1 = kinase domain
SH2	Src homology domain 2
SH3	Src homology domain 3
SHG	second harmonic generation
SIM	single input
SIN	self-inactivating lentiviral vectors, due to a 3′ LTR mutation
SINE	scattered or short interspersed elements
siRNA	small interfering RNA
SIV	simian immunodeficiency virus
SNARE proteins	SNAP receptor proteins
SNP	single nucleotide polymorphism
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SOP	stock option program
SP function	sum-of-pairs function
SPA	scintillation proximity assay
SPDM	spectral precision distance microscopy
SPF	S-phase promotion factor
SRP	signal recognition particle
SSB	single-strand binding proteins
SSCP	single-strand conformation polymorphism
ssDNA	single-stranded DNA
SSH	suppression subtractive hybridization
SssI methylase	methylase from *Spiroplasma*
ssRNA	single-stranded RNA
STED	stimulated emission depletion
STEM	scanning transmission electron microscope
stRNA	small temporal RNA
STS	sequence-tagged site
SV40	Simian virus type 40
TBP	TATA-binding protein
T_c	cytotoxic T cells
Tc	tetracycline
T-DNA	transfer DNA
TEM	transmission electron microscope
TEV	tobacco etch virus
T_H	T helper cell
THG	third harmonic generation
TIGR	The Institute for Genome Research
TIM	translocase of inner membrane
T_m	melting temperature of dsDNA
TNF	tumor necrosis factor
TOF	time of flight
TOM	translocase of outer membrane
t-PA	tissue plasminogen activator
TRE	tetracycline-responsive element
TRIPs	Trade-Related Aspects of Intellectual Property Rights
tRNA	transfer RNA
Trp	tryptophan
t-SNARE	protein in target membrane to which v-SNARE binds
TSS	transformation and storage solution
tTA	tetracycline-controlled transactivator
TY	transposon from yeast
UPOV	Union for the Protection of New Varieties of Plants
US-GAAP	US generally accepted accounting principle
UV	ultraviolet
V_0	empty volume
VC	venture capital
V_e	elution volume
VEGF	vascular endothelial growth factor
VIP	vasoactive peptide
VNTR	variable number tandem repeats
v-SNARE	protein in vesicular membrane, binding to t-SNARE
VSV-G	envelope protein of vesicular stomatitis virus, great affinity to a wide range of cells
V_t	total volume
wNAPPA	modified nucleic acid programmable protein array
WPRE	woodchuck hepatitis virus posttranscriptional regulatory element
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YAC	yeast artificial chromosome
YEp	yeast episomal plasmid
YFP	yellow fluorescence protein
YIp	yeast-integrating plasmid
YRp	yeast-replicating plasmid
Yth	yeast two-hybrid

Part I

Fundamentals of Cellular and Molecular Biology

1

The Cell as the Basic Unit of Life

Michael Wink

Heidelberg University, Institute of Pharmacy and Molecular Biotechnology (IPMB), Im Neuenheimer Feld 329, 69120 Heidelberg, Germany

The base unit of life is the **cell**. Cells constitute the base element of all **prokaryotic cells** (cells without a cell nucleus, e.g. **Bacteria** and **Archaea**) and **eukaryotic cells** (or **Eukarya**) (cells possessing a nucleus, e.g. protozoa, fungi, plants, and animals). Cells are small, membrane-bound units with a diameter of 1–20 µm and are filled with concentrated aqueous solutions. Cells are not created *de novo*, but possess the ability to copy themselves, meaning that they emerge from the division of a previous cell. This means that all cells, since the beginning of life (around 4 billion years ago), are connected with each other in a continuous lineage. In 1885, the famous cell biologist Rudolf Virchow conceived the law of *omnis cellula e cellula* (all cells arise from cells), which is still valid today.

The structure and composition of all cells are very similar due to their shared evolution and phylogeny (Figure 1.1). We see an astonishing constancy in fundamental structures and mechanisms. Owing to this, it is possible to limit the discussion of the general characteristics of a cell to a few basic types (Figure 1.2):

- Bacterial cells
- Plant cells
- Animal cells

Nucleotide sequences from 16S rRNA, amino acid sequences of cytoskeleton proteins, and characteristics of the cell structure were used to reconstruct this phylogenetic tree. Prokaryotes are divided into **Bacteria** and **Archaea**. Archaea form a sister group with eukaryotes; they share important characteristics (Tables 1.1 and 1.2). Many monophyletic groups can be recognized within the eukaryotes (diplomonads/trichomonads, Euglenozoa, Alveolata, Stramenopilata [heterokonts], red algae and green algae/plants, fungi and animals; see Tables 6.3–6.5 for details).

A highly resolved tree of life is based on completely sequenced genomes (Ciccarelli 2006). The image was generated using Interactive Tree Of Life (iTOL) (Letunic 2007), an online phylogenetic tree viewer and Tree of Life resource. Eukaryotes are colored red, archaea green, and bacteria blue.

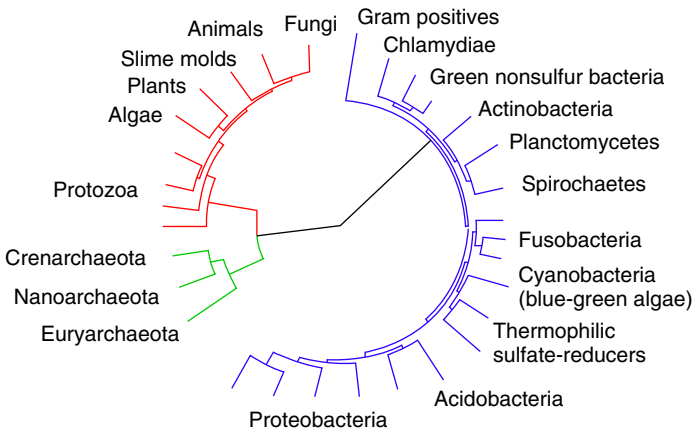

Figure 1.1 Tree of life – phylogeny of life domains.

An Introduction to Molecular Biotechnology: Fundamentals, Methods and Applications, Third Edition. Edited by Michael Wink.

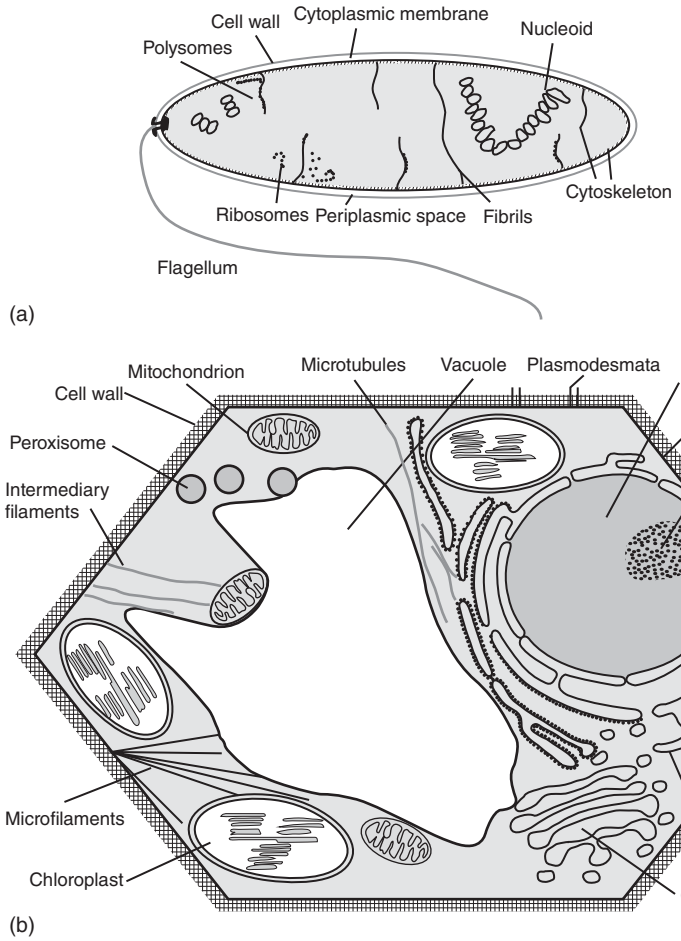

Figure 1.2 Schematic structure of prokaryotic and eukaryotic cells. (a) Bacterial cell, (b) plant mesophyll cell, and (c) animal cell.

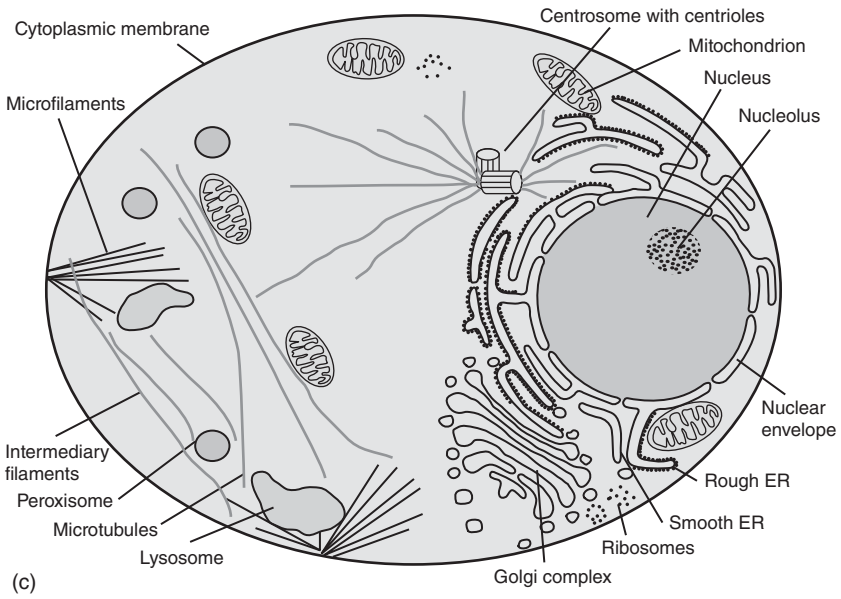

Figure 1.2 (*Continued*)