

Contents

Introduction: Cells as macromolecular assemblies 1

1: Proteins	3
Macromolecules are assembled by polymerizing small molecules	6
Proteins consist of chains of amino acids	8
Protein conformation depends on the aqueous environment	13
Protein structures are extremely versatile	17
How do proteins fold into the correct conformation?	19
2: Compartments	27
Cellular compartments are bounded by membranes	29
The cytoplasm contains networks of membranes	33
Cell shape is determined by the cytoskeleton	36
Some organelles are surrounded by an envelope	39
The environment of the nucleus and its reorganization	41
The role of chromosomes in heredity	43

Part 1: DNA as information 49

3: Genes are mutable units	51
Discovery of the gene	54
Genes lie in a linear array on chromosomes	56
One gene—one protein	61
The cistron	63
Mapping mutations at the molecular level	65
The nature of multiple alleles	67
4: DNA is the genetic material	71
The discovery of DNA	74
DNA is the (almost) universal genetic material	76
The components of DNA	79
DNA is a double helix	82

DNA replication is semiconservative	82
The genetic code is read in triplets	86
Mutations change the sequence of DNA	89
Mutations are concentrated at hotspots	91
The rate of mutation	94

5: Nucleic acid structure 97

DNA can be denatured and renatured	98
Nucleic acids hybridize by base pairing	99
Single-stranded nucleic acids may have secondary structure	101
Inverted repeats and secondary structure	105
Duplex DNA has alternative double-helical structures	107
Closed DNA can be supercoiled	109
Supercoiling influences the structure of the double helix	111

6: Isolating the gene 115

A restriction map is constructed by cleaving DNA into specific fragments	117
Restriction sites can be used as genetic markers	122
Obtaining the sequence of DNA	129
Prokaryotic genes and proteins are colinear	134
<i>cis</i> -acting sites and <i>trans</i> -acting molecules	136
Eukaryotic genes are often interrupted	139
Some DNA sequences code for more than one protein	141
Genetic information can be provided by DNA or RNA	143
The scope of the paradigm	146

Part 2: From gene to protein

151

7: Messenger RNA 153

Transfer RNA is the adaptor	155
Messenger RNA is translated by ribosomes	159
The life cycle of messenger RNA	162
Most bacterial genes are expressed via polycistronic messengers	166
Translation of eukaryotic mRNA	168
Eukaryotic mRNAs are polyadenylated at the 3' end	170
Eukaryotic mRNAs have a methylated cap at the 5' end	171
Processing and stability of mRNA	173

8: Protein synthesis 179

Organization of the ribosome	181
The stages of protein synthesis	183
Initiation in bacteria needs 30S subunits and accessory factors	186
A special initiator tRNA starts the polypeptide chain	187
Initiation involves base pairing between mRNA and rRNA	191
Small subunits migrate to initiation sites on eukaryotic mRNA	192
Elongation factor T brings aminoacyl-tRNA into the A site	196

Translocation moves the ribosome	198
Three codons terminate protein synthesis	202
Ribosomes have several active centers	204
The role of ribosomal RNA in protein synthesis	207
9: Interpreting the genetic code	213
Codon-anticodon recognition involves wobbling	215
tRNA contains modified bases that influence its pairing properties	217
The genetic code is altered in mitochondria	221
tRNAs are charged with amino acids by individual synthetases	223
Accuracy depends on proofreading	228
Suppressor tRNAs have mutated anticodons that read new codons	231
The accuracy of translation	235
tRNA may influence the reading frame	237
10: Protein localization	244
Chaperones may be required for protein folding	246
Post-translational membrane insertion depends on leader sequences	251
A hierarchy of sequences determines location within organelles	254
Signal sequences initiate co-translational transfer through ER membranes	257
How do proteins enter and leave membranes?	258
The translocation apparatus interacts with signal and anchor sequences	263
Anchor signals are needed for membrane residence	266
Bacteria use both co-translational and post-translational translocation	269
Pores control nuclear ingress and egress	271
Protein degradation by proteasomes	278
Part 3: Prokaryotic gene expression	285
11: Transcription	287
Transcription is catalyzed by RNA polymerase	289
RNA polymerase consists of multiple subunits	294
Sigma factor controls binding to DNA	297
Promoter recognition depends on consensus sequences	302
RNA polymerase binds to one face of DNA	305
Substitution of sigma factors may control initiation	309
Sporulation utilizes a cascade of many sigma factors	312
Bacterial RNA polymerase has two modes of termination	317
How does rho factor work?	320
Antitermination depends on specific sites	323
More subunits for RNA polymerase	329
12: The operon	335
Structural gene clusters are coordinately controlled	338
Repressor is controlled by a small molecule inducer	340
Mutations identify the operator and the regulator gene	343

Repressor protein binds to the operator and is released by inducer	348
The specificity of protein-DNA interactions	354
Repression can occur at multiple loci	357
Distinguishing positive and negative control	359
Catabolite repression involves positive regulation at the promoter	361
Adverse growth conditions provoke the stringent response	365
Autogenous control may occur at translation	368
Alternative secondary structures control attenuation	374
Small RNA molecules can regulate translation	380
Regulation by cleavage of mRNA	385
Cleavages are needed to release prokaryotic and eukaryotic rRNAs	387

13: Phage strategies **395**

Lytic development is controlled by a cascade	397
Functional clustering in phages T7 and T4	400
The lambda lytic cascade relies on antitermination	403
Lysogeny is maintained by an autogenous circuit	407
The DNA-binding form of repressor is a dimer	410
Repressor binds cooperatively at each operator using a helix-turn-helix motif	412
How is repressor synthesis established?	417
A second repressor is needed for lytic infection	421
A delicate balance: lysogeny versus lysis	423

Part 4: Perpetuation of DNA

427

14: The replicon **429**

Origins can be mapped by autoradiography and electrophoresis	431
The bacterial genome is a single circular replicon	433
Each eukaryotic chromosome contains many replicons	436
Isolating the origins of yeast replicons	438
D loops may be maintained at mitochondrial origins	440
The problem of linear replicons	442
Rolling circles produce multimers of a replicon	445
Single-stranded genomes are generated for bacterial conjugation	449
Connecting bacterial replication to the cell cycle	453
Cell division and chromosome segregation	455
Multiple systems ensure plasmid survival in bacterial populations	460
Plasmid incompatibility is connected with copy number	463

15: DNA replication **471**

DNA polymerases: the enzymes that make DNA	472
DNA synthesis is semidiscontinuous and primed by RNA	477
The primosome initiates synthesis of Okazaki fragments	480
Coordinating synthesis of the lagging and leading strands	484
The replication apparatus of phage T4	491
Creating the replication forks at an origin	493

Common events in priming replication at the origin	496
Does methylation at the origin regulate initiation?	498
Licensing factor controls eukaryotic rereplication	500
16: Restriction and repair	505
The consequences of modification and restriction	506
Type II restriction enzymes are common	508
The alternative activities of type I enzymes	510
The dual activities of type III enzymes	513
Dealing with injuries in DNA	515
Excision repair systems in <i>E. coli</i>	518
Controlling the direction of mismatch repair	521
Retrieval systems in <i>E. coli</i>	523
RecA triggers the SOS system	525
Eukaryotic repair systems	527
17: Recombination	531
Breakage and reunion involves heteroduplex DNA	534
Double-strand breaks initiate recombination	537
Double-strand breaks may initiate synapsis	539
Bacterial recombination involves single-strand assimilation	542
Gene conversion accounts for interallelic recombination	548
Topological manipulation of DNA	550
Gyrase introduces negative supercoils in DNA	553
Specialized recombination involves breakage and reunion at specific sites	555
18: Transposons	563
Insertion sequences are simple transposition modules	565
Composite transposons have IS modules	567
Transposition occurs by both replicative and nonreplicative mechanisms	569
Common intermediates for transposition	572
Replicative transposition proceeds through a cointegrate	574
Nonreplicative transposition proceeds by breakage and reunion	576
TnA transposition requires transposase and resolvase	578
Transposition of Tn10 has multiple controls	580
Controlling elements in maize cause breakage and rearrangements	583
Controlling elements in maize form families of transposons	586
<i>Spm</i> elements influence gene expression	588
The role of transposable elements in hybrid dysgenesis	589
19: Retroviruses and retroposons	597
The retrovirus life cycle involves transposition-like events	598
Retroviruses may transduce cellular sequences	607
Yeast Ty elements resemble retroviruses	609
Many transposable elements reside in <i>D. melanogaster</i>	611
Retroposons fall into two classes	613

Part 5: The eukaryotic genome

621

20: DNA biotechnology	623
Any DNA sequence can be cloned in bacteria or yeast	624
Constructing the chimeric DNA	626
Copying mRNA into cDNA	629
Isolating individual genes from the genome	631
Walking along the chromosome	636
Eukaryotic genes can be expressed in prokaryotic systems	640
21: Genomes	645
The C-value paradox describes variations in genome size	646
Reassociation kinetics depend on sequence complexity	648
Eukaryotic genomes contain several sequence components	650
Nonrepetitive DNA complexity can estimate genome size	651
Eukaryotic genomes contain repetitive sequences	652
Most structural genes lie in nonrepetitive DNA	654
How many nonrepetitive genes are expressed?	657
Genes are expressed at widely varying levels	659
22: Exons and introns	663
Organization of interrupted genes may be conserved	665
Genes show a wide distribution of sizes	668
One DNA sequence may code for multiple proteins	672
Exon sequences are conserved but introns vary	674
Genes can be isolated by the conservation of exons	676
How did interrupted genes evolve?	679
23: Gene numbers	687
Essential genes and total gene number	689
Globin genes are organized in two clusters	692
Unequal crossing-over rearranges gene clusters	694
Gene clusters suffer continual reorganization	698
Sequence divergence is the basis for the evolutionary clock	699
Pseudogenes are dead ends of evolution	703
Genes for rRNA comprise a repeated tandem unit	704
An evolutionary dilemma: how are multiple active copies maintained?	708
24: Organelle genomes	713
Organelle genomes are circular DNAs that code for organelle proteins	715
The chloroplast genome codes for ~100 proteins and RNAs	719
The mitochondrial genome is large in yeast but small in mammals	720
Recombination and rearrangement of organelle DNA	723
25: Simple sequence DNA	727
Satellite DNAs often lie in heterochromatin	729
Arthropod satellites have very short identical repeats	730

Mammalian satellites consist of hierarchical repeats	730
Evolution of hierarchical variations in the satellite	734
The consequences of unequal crossing-over	736
Crossover fixation could maintain identical repeats	738
Minisatellites are useful for genetic mapping	739
26: Chromosomes	743
Condensing viral genomes into their coats	744
The bacterial genome is a nucleoid with many supercoiled loops	747
Loops, domains, and scaffolds in eukaryotic DNA	750
The contrast between interphase chromatin and mitotic chromosomes	753
The extended state of lampbrush chromosomes	756
Transcription disrupts the structure of polytene chromosomes	757
The eukaryotic chromosome as a segregation device	760
Telomeres seal the ends of chromosomes	763
27: Nucleosomes	769
The nucleosome is the subunit of all chromatin	770
DNA is coiled in arrays of nucleosomes	773
DNA structure varies on the nucleosomal surface	777
Supercoiling and the periodicity of DNA	780
The path of nucleosomes in the chromatin fiber	782
Organization of the histone octamer	784
Reproduction of chromatin requires assembly of nucleosomes	787
Do nucleosomes lie at specific positions?	791
Are transcribed genes organized in nucleosomes?	794
DNAase hypersensitive sites change chromatin structure	798
Domains define regions that contain active genes	801
Heterochromatin is created by interactions with histones	803

Part 6: Eukaryotic gene expression

809

28: Initiation of transcription	811
Eukaryotic RNA polymerases consist of many subunits	814
Promoter elements are defined by mutations and footprinting	815
RNA polymerase I has a bipartite promoter	817
RNA polymerase III uses both downstream and upstream promoters	819
The basal apparatus consists of RNA polymerase II and general factors	822
A connection between transcription and repair	829
Promoters for RNA polymerase II have short sequence elements	831
Enhancers contain bidirectional elements that assist initiation	835
Independent domains bind DNA and activate transcription	839
Interaction of upstream factors with the basal apparatus	842

29: Regulation of transcription	847
Response elements identify genes under common regulation	848
There are many types of DNA-binding domains	850
A zinc finger motif is a DNA-binding domain	852
Steroid receptors have several independent domains	855
Homeodomains bind related targets in DNA	859
Helix-loop-helix proteins interact by combinatorial association	862
Leucine zippers are involved in dimer formation	864
Dynamic versus pre-emptive models for gene activation	866
Long range regulation and insulation of domains	871
Gene expression is associated with demethylation	875
Methylation is responsible for imprinting	878
30: Nuclear splicing	885
Nuclear splice junctions are interchangeable but are read in pairs	887
Nuclear splicing proceeds through a lariat	891
SnRNAs are required for splicing and form a spliceosome	893
Group II introns autosplice via lariat formation	901
Alternative splicing involves differential use of splice junctions	904
<i>cis</i> -splicing and <i>trans</i> -splicing reactions	907
Yeast tRNA splicing involves cutting and rejoining	911
3' ends are generated by termination and by cleavage reactions	913
31: Catalytic RNA	921
Group I introns undertake self-splicing by transesterification	922
Group I introns form a characteristic secondary structure	926
Ribozymes have various catalytic activities	928
Some introns code for proteins that sponsor mobility	931
RNA can have ribonuclease activities	935
RNA editing utilizes information from several sources	937
32: Rearrangement of DNA	947
The mating pathway is triggered by signal transduction	948
Yeast can switch silent and active loci for mating type	952
Silent cassettes at <i>HML</i> and <i>HMR</i> are repressed	956
Unidirectional transposition is initiated by the recipient <i>MAT</i> locus	958
Regulation of <i>HO</i> expression	960
Trypanosomes rearrange DNA to express new surface antigens	962
Interaction of Ti plasmid DNA with the plant genome	967
Selection of amplified genomic sequences	975
Exogenous sequences can be introduced into cells and animals by transfection	979
33: Immune diversity	989
Clonal selection amplifies lymphocytes that respond to individual antigens	992
Immunoglobulin genes are assembled from their parts in lymphocytes	994
The diversity of germline information	1000
Recombination between V and C genes generates deletions and rearrangements	1002
Allelic exclusion is triggered by productive rearrangement	1007
DNA recombination causes class switching	1009
Early heavy chain expression can be changed by RNA processing	1011

Somatic mutation generates additional diversity	1012
T-cell receptors are related to immunoglobulins	1015
The major histocompatibility locus codes for many genes of the immune system	1019

Part 7: Cell growth, cancer, and development 1025

34: Protein trafficking	1027
Oligosaccharides are added to proteins in the ER and Golgi	1030
Coated vesicles transport both exported and imported proteins	1033
Protein localization depends on further signals	1042
Receptors recycle via endocytosis	1044
35: Signal transduction	1053
Carriers and channels form water-soluble paths through the membrane	1056
G proteins may activate or inhibit target proteins	1061
Protein tyrosine kinases induce phosphorylation cascades	1064
The Ras pathway	1070
Activating MAP kinase pathways	1076
Cyclic AMP and activation of CREB	1081
The JAK-STAT pathway	1082
36: Cell cycle and growth regulation	1089
Cycle progression depends on discrete control points	1090
M phase kinase is a dimer that regulates entry into mitosis	1095
Protein phosphorylation and dephosphorylation control the cell cycle	1098
p34 (cdc2 or CDC28) is the key regulator in yeasts	1100
<i>CDC28</i> acts at both START and mitosis in <i>S. cerevisiae</i>	1108
Many cdk-cyclin complexes are found in animal cells	1111
Functions of cdc2-cyclin and cdk-cyclin dimers	1113
G0/G1 and G1/S transitions involve cdk inhibitors	1116
Reorganization of the cell at mitosis	1119
Apoptosis	1122
37: Oncogenes and cancer	1131
Transforming viruses carry oncogenes	1135
Retroviral oncogenes have cellular counterparts	1139
Ras proto-oncogenes can be activated by mutation	1141
Insertion, translocation, or amplification may activate proto-oncogenes	1144
Oncogenes code for components of signal transduction cascades	1149
Growth factor receptor kinases and cytoplasmic tyrosine kinases	1151
Oncoproteins may regulate gene expression	1156
RB is a tumor suppressor that controls the cell cycle	1160
The tumor suppressor p53 suppresses growth or triggers apoptosis	1162
Immortalization and transformation	1167

38: Gradients and cascades 1173

A gradient must be converted into discrete compartments 1175

Maternal gene products establish gradients in early embryogenesis 1177

Anterior-posterior development uses localized gene regulators 1180

Dorsal-ventral development uses localized receptor-ligand interactions 1184

Cell fate is determined by compartments that form by the blastoderm stage 1191

Complex loci are extremely large and involved in regulation 1198

The homeobox is a common coding motif in homeotic genes 1205

Epilogue: Landmark shifts in perspectives 1213

Glossary 1217

Index 1241

35: Signal transduction

Carriers and channels form water-soluble paths through the membrane 1250

G proteins may activate or inhibit target proteins 1250

Protein tyrosine kinases induce phosphorylation cascades 1250

The Ras pathway 1250

Activating MAP kinase pathways 1250

Cyclic AMP and activation of CREB 1250

The JAK-STAT pathway 1250

36: Cell cycle and growth regulation

Cell cycle progression depends on discrete control points 1250

M phase kinase is a dimer that regulates entry into mitosis 1250

Protein phosphorylation and dephosphorylation control the cell cycle 1250

p34 (cdc2 or CDC28) is the key regulator in yeasts 1250

CDC28 acts at both START and mitosis in *S. cerevisiae* 1250

Many cdk-cyclin complexes are found in animal cells 1250

Functions of cdc2-cyclin and cdk-cyclin dimers 1250

G1/S and G2/M transitions involve cdk inhibitors 1250

Reorganization of the cell at mitosis 1250

37: Oncogenes and cancer

Transforming viruses carry oncogenes 1250

Retroviral oncogenes have cellular counterparts 1250

Proto-oncogenes can be activated by mutation 1250

Insertion, translocation, or amplification may activate proto-oncogenes 1250

Oncogenes code for components of signal transduction cascades 1250

Growth factor receptor kinases and cytoplasmic tyrosine kinases 1250

Oncoproteins may regulate gene expression 1250

Rb is a tumor suppressor that controls the cell cycle 1250

The tumor suppressor p53 suppresses growth or triggers apoptosis 1250

Immunization and transfection 1250

38: Immune diversity

Gene rearrangement amplifies lymphocyte antigen receptors 1250

The immunoglobulin genes are assembled with many building blocks 1250

The immunoglobulin genes are assembled with many building blocks 1250

Recombination between V and C genes generates antibody diversity 1250

Allelic exclusion is triggered by productive rearrangement 1250

DNA recombination causes class switching 1250

Early heavy chain expression can be changed by RNA processing 1250