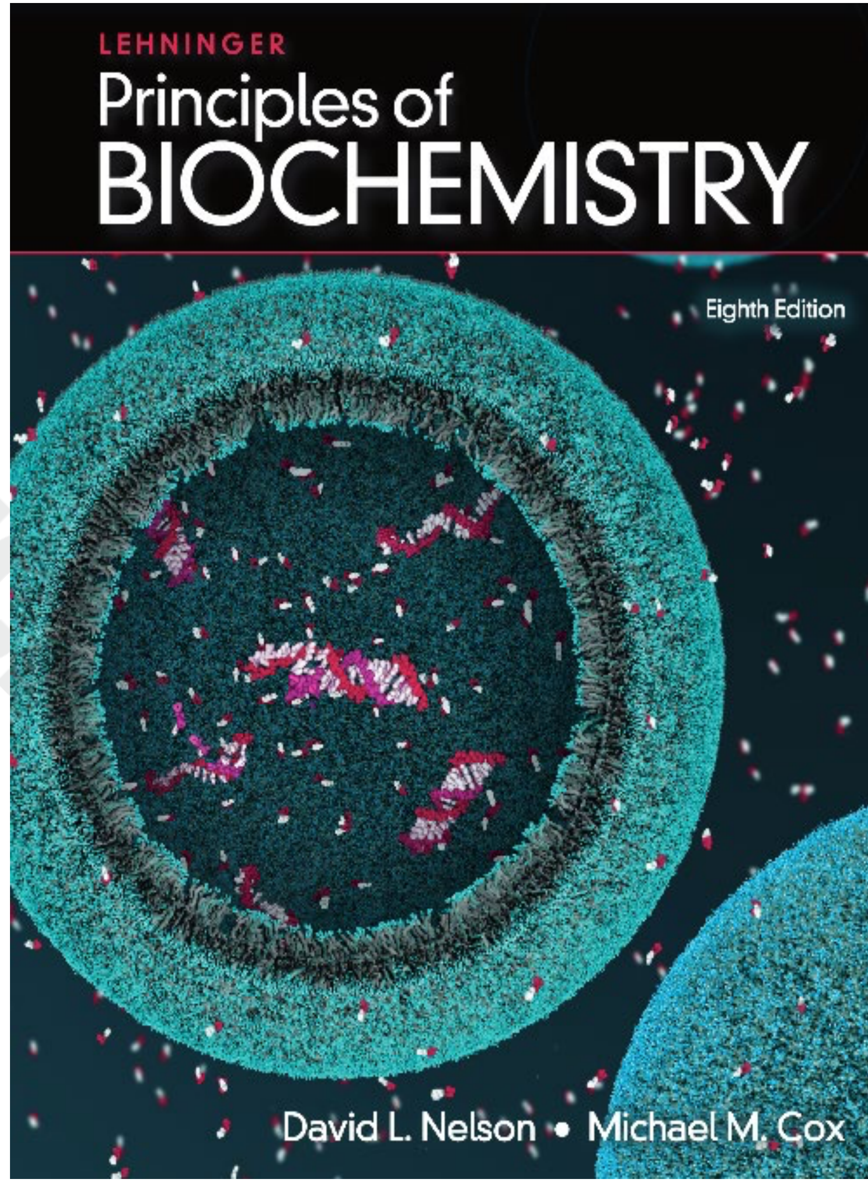
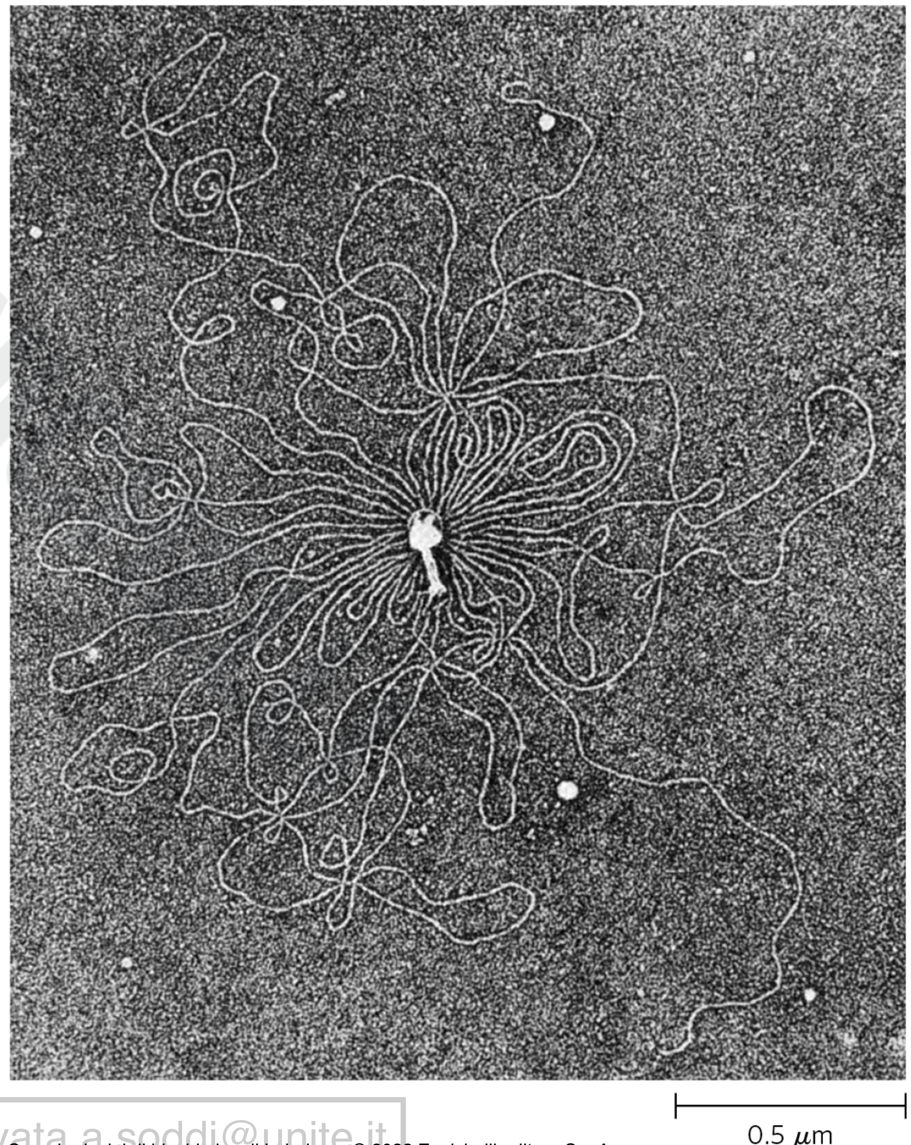
24 Genes and Chromosomes



© 2021 Macmillan Learning

Chromosomes

- **chromosomes** = tertiary packaging of DNA
  - serve as repositories of genetic information



P1 Principle 1 (1 of 4)

Chromosomes include dedicated sequences that ensure their replication, transcription, packaging, and transmission from one generation to the next. They are more than a long stretch of protein-encoding genes.

P2 Principle 2 (1 of 6)

Chromosomes are large. To constrain them in a small space can require multiple layers and multiple modes of tertiary structure.

P3 Principle 3 (1 of 7)

Chromosomes in all cells are maintained in a state of torsional stress. DNA is underwound relative to the stable B-form structure, facilitating both the packaging of DNA and access to the genetic information contained within it.

P4 Principle 4

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

24.1 Chromosomal Elements

Genes Are Segments of DNA That Code for Polypeptide Chains and RNAs

- **phenotype** = visible property
- classically, a gene was defined as a portion of a chromosome that determines or affects a single phenotype

Molecular Definitions of a Gene

- **mutations =** alterations in DNA sequence
- experiments by Beadle and Tatum found that mutant fungal strains lacked one or another specific enzyme
- the **one gene-one enzyme** hypothesis = a gene is a segment of genetic material that determines, or codes for, one enzyme
  - later broadened to **one gene-one polypeptide** hypothesis

The Modern Biochemical Definition of a Gene

- **gene** = all the DNA that encodes the primary sequence of some final gene product (polypeptide or RNA with a structural or catalytic function)

Regulatory Sequences

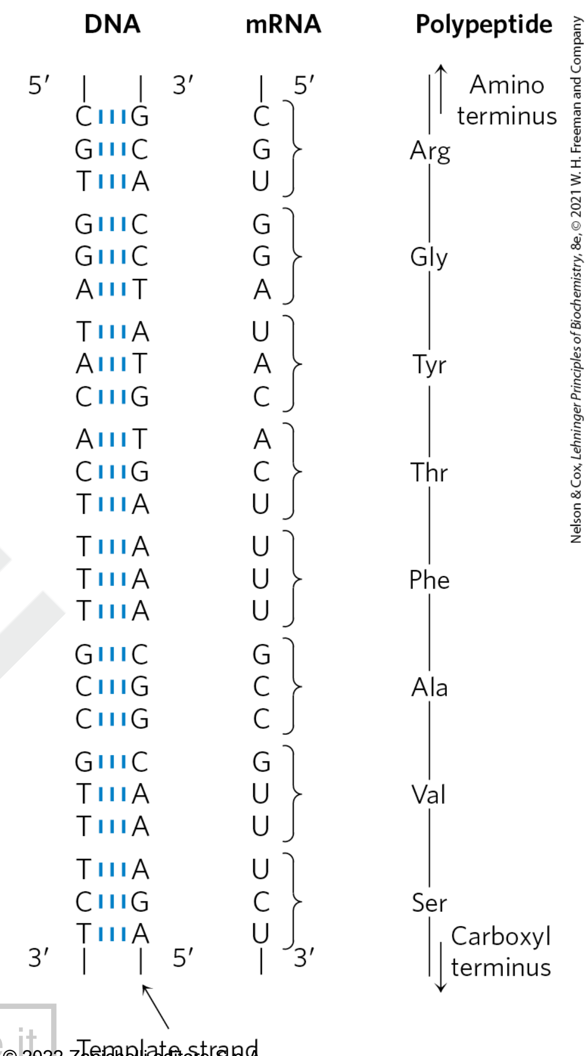
- **regulatory sequences** = segments of sequences of DNA that have a purely regulatory function

- functions of regulatory sequences:
  - denote the beginning or the end of genes
  - influence the transcription of genes
  - initiation points for replication or recombination

Colinearity of DNA, mRNA, and Protein

- each amino acid of a polypeptide chain is coded for by three consecutive nucleotides in a single strand of DNA (“codon”)

- a polypeptide chain of 350 amino acid residues (an average-size chain) corresponds to 1,050 base pairs (bp) of coding DNA



Number of Genes in a Single Chromosome

- the *Escherichia coli* chromosome is a circular DNA molecule with 4,641,652 bp
  - encodes ~4,300 genes for proteins and >200 genes for structural or catalytic RNAs

- the human genome is ~3.1 billion bp
  - encodes approximately 20,000 genes across 24 different chromosomes

DNA Molecules Are Much Longer than the Cellular or Viral Packages That Contain Them

- chromosomal DNAs are many orders of magnitude longer than the cells or viruses in which they are located

Table 24-1 The Sizes of DNA and Viral Particles for Some Bacteria Viruses (Bacteriophages)

Virus	Size of viral DNA (bp)	Length of viral DNA (nm)	Long dimension of viral particles (nm)
ϕX174	5,386	1,939	25
T7	39,936	14,377	78
λ (lambda)	48,502	17,460	190
T4	168,889	60,800	210

Viruses

- viruses = infectious parasites that use resources of a host cell to carry out processes required to propagate
  - many consist of just a single RNA or DNA molecule and a protein coat

P2 Principle 2 (2 of 6)

Chromosomes are large. To constrain them in a small space can require multiple layers and multiple modes of tertiary structure.

RNA and DNA Viruses

- almost all plant viruses and some bacterial and animal viruses have RNA genomes
  - tend to be particularly small

- DNA viral genomes vary greatly in size and tend to be circular for at least part of their life

- the contour lengths of their DNA are typically hundreds of times longer than the long dimensions of the viral particles that contain them

Replicative Forms Appear During Viral Replication

- **replicative forms** = specific types of viral DNA that appear during viral replication within a host cell

- examples of replicative forms:
  - many linear DNAs become circular
  - all single-stranded DNAs become double-stranded

Bacteria

- an *E. coli* cell contains ~100x as much DNA as a bacteriophage λ particle

- the *E. coli* chromosome is a single, double-stranded circular DNA molecule of 4,641,652 bp
  - the genome is ~850 times the length of the cell

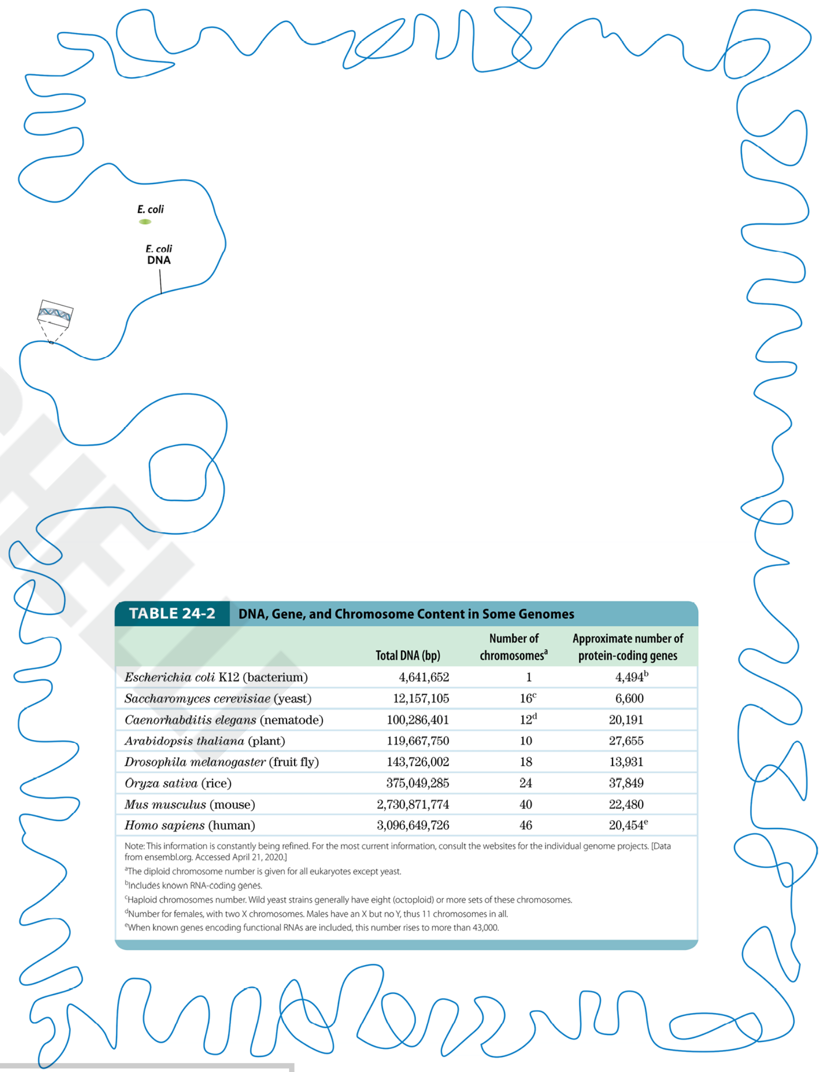


TABLE 24-2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes[a]	Approximate number of protein-coding genes
Escherichia coli K12 (bacterium)	4,641,652	1	4,494[b]
Saccharomyces cerevisiae (yeast)	12,157,105	16[c]	6,600
Caenorhabditis elegans (nematode)	100,286,401	12[d]	20,191
Arabidopsis thaliana (plant)	119,667,750	10	27,655
Drosophila melanogaster (fruit fly)	143,726,002	18	13,931
Oryza sativa (rice)	375,049,285	24	37,849
Mus musculus (mouse)	2,730,871,774	40	22,480
Homo sapiens (human)	3,096,649,726	46	20,454[e]

Note: This information is constantly being refined. For the most current information, consult the websites for the individual genome projects. [Data from ensembl.org. Accessed April 21, 2020.]

[a]The diploid chromosome number is given for all eukaryotes except yeast.

[b]Includes known RNA-coding genes.

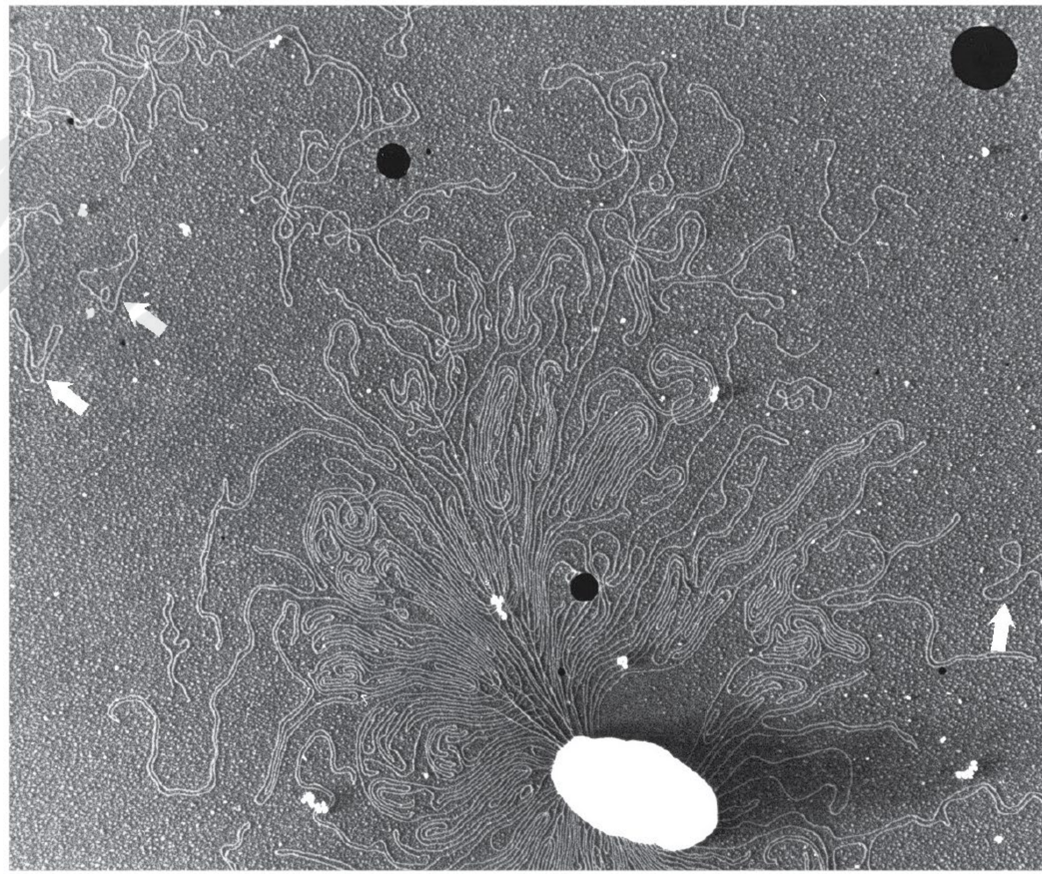
[c]Haploid chromosomes number. Wild yeast strains generally have eight (octoploid) or more sets of these chromosomes.

[d]Number for females, with two X chromosomes. Males have an X but no Y, thus 11 chromosomes in all.

[e]When known genes encoding functional RNAs are included, this number rises to more than 43,000.

Bacterial Plasmids

- **plasmids** = small, circular DNA molecules that are in the cytosol of many bacteria
  - undergo replication
  - some carry useful genes, such as antibiotic resistance genes



Huntington Potter, University of Colorado School of Medicine, and David Dressler, Balliol College, Oxford University.

Eukaryotes

- a yeast cell contains 2.6x as much DNA as an *E. coli* cell.

- a *Drosophila* (fruit fly) cell contains >35x as much DNA as an *E. coli* cell

- a human cell contains ~700x as much DNA as an *E. coli* cell
  - cells of many plants and amphibians contain even more

DNA, Gene, and Chromosome Content in Some Genomes

Table 24-2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes	Approximate number of protein-coding genes
Escherichia coli K12 (bacterium)	4,641,652	1	4,494[b]
Saccharomyces cerevisiae (yeast)	12,157,105	16[c]	6,600
Caenorhabditis elegans (nematode)	100,286,401	12[d]	20,191
Arabidopsis thaliana (plant)	119,667,750	10	27,655
Drosophila melanogaster (fruit fly)	143,726,002	18	13,931
Oryza sativa (rice)	375,049,285	24	37,849
Mus musculus (mouse)	2,730,871,774	40	22,480
Homo sapiens (human)	3,096,649,726	46	20,454[e]

Eukaryotic Chromosomes

- genetic material of eukaryotic cells is distributed into chromosomes
  - the diploid (2*n*) number depends on the species
  - each chromosome contains a single, very large, duplex DNA molecule that carries a characteristic set of genes

- human somatic cells have 46 chromosomes

(a)

Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

(b)

Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

P2 Principle 2 (3 of 6)

Chromosomes are large. To constrain them in a small space can require multiple layers and multiple modes of tertiary structure.

The Length of Human DNA Molecules

- placed end to end, DNA molecules of one human genome would extend for about a meter
  - human cells are diploid, so each cell has 2 m of DNA

- the adult human body contains ~10^{14} cells and a total DNA length of 2×10^{11} km
  - greater than the the circumference of the earth (4×10^4 km)
  - greater than the distance between the earth and the sun (1.5×10^8 km)

Mitochondria Contain DNA

- mitochondrial DNA (mtDNA) molecules are much smaller than nuclear chromosomes
  - codes for mitochondrial tRNAs and rRNAs and for a few mitochondrial proteins
- animal mtDNA contains <20,000 bp in a circular duplex
  - each mitochondrion has 2-10 copies
  - number of copies rises in some cells in an embryo during cell differentiation
- plant mtDNA ranges from 200,000 to 2,500,000 bp

Chloroplasts Contain DNA

- chloroplast DNA (cpDNA) exists as circular duplexes from 120,000 to 160,000 bp

P1 Principle 1 (2 of 4)

Chromosomes include dedicated sequences that ensure their replication, transcription, packaging, and transmission from one generation to the next. They are more than a long stretch of protein-encoding genes.

Eukaryotic Genes and Chromosomes Are Very Complex

- **introns** = nontranslated, intervening DNA segments that do not code for the amino acid sequence of the polypeptide product

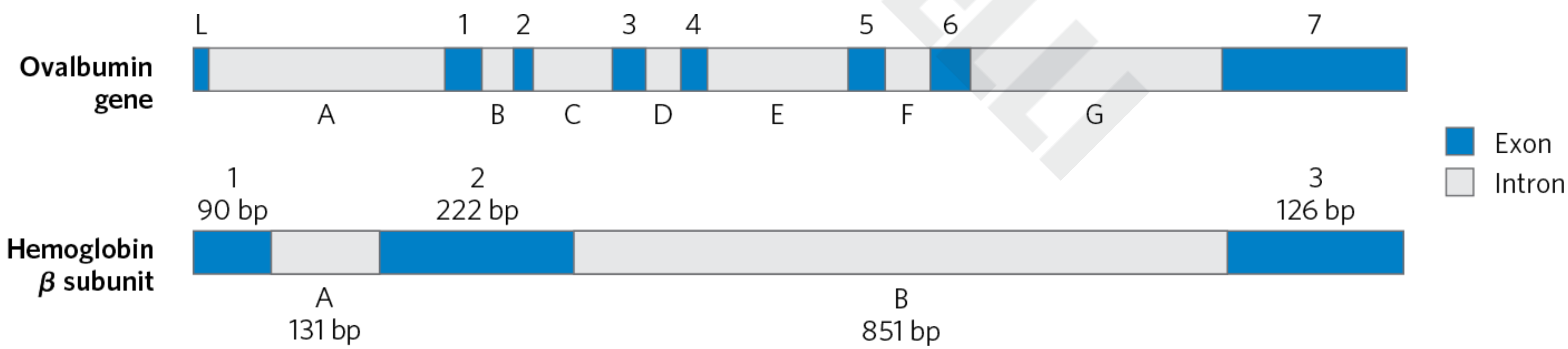
- **exons** = coding DNA segments
  - makes up only ~1.5% of human DNA

- the colinearity of the eukaryotic DNA and amino acid sequence is broken by introns
  - few bacterial genes contain introns

Introns in Eukaryotic Genes

- in higher eukaryotes, the typical gene has more intron sequences than exon sequences

- together, introns and exons make up as much as 30% of the human genome



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Highly Repetitive Sequences

- **highly repetitive sequences** (**simple-sequence DNA** or **simple sequence repeats (SSR**)) = short sequences, generally less than 10 bp long, that are sometimes repeated millions of times per cell
  - make up ~3% of the human genome
  - do not encode proteins or RNAs
  - associated with centromeres and telomeres

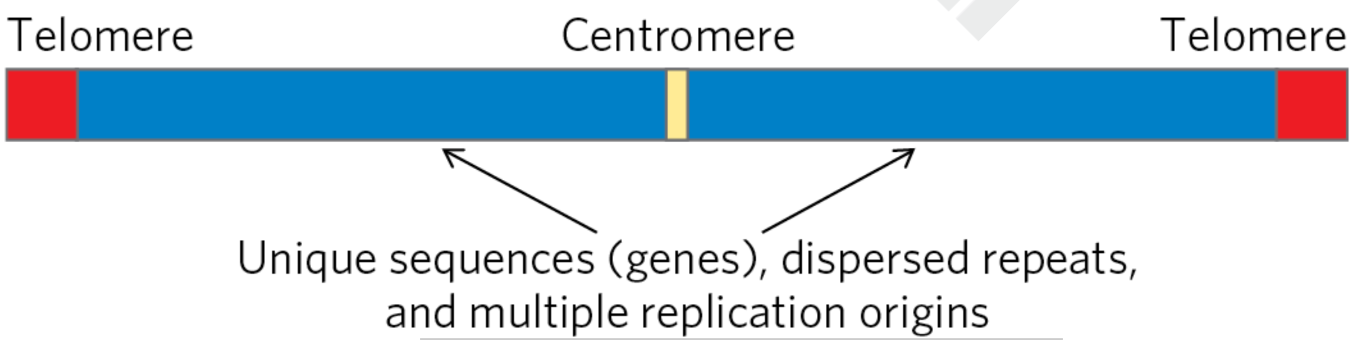
- **satellite DNA** = another name for simple-sequence DNA
  - named as such because it migrates as “satellite” bands in a cesium chloride density gradient

P1 Principle 1 (3 of 4)

Chromosomes include dedicated sequences that ensure their replication, transcription, packaging, and transmission from one generation to the next. They are more than a long stretch of protein-encoding genes.

Centromeres

- **centromere** = a DNA sequence that functions during cell division as an attachment point for proteins that link the chromosome to the mitotic spindle
  - essential sequences in yeast are ~130 bp long and are very rich in A=T pairs
  - sequences of higher eukaryotes are longer and generally consist of thousands of tandem copies of one or several sequences of 5 to 10 bp



P1 Principle 1 (4 of 4)

Chromosomes include dedicated sequences that ensure their replication, transcription, packaging, and transmission from one generation to the next. They are more than a long stretch of protein-encoding genes.

Telomeres

- **telomeres** = sequences at the ends of eukaryotic chromosomes that help stabilize the chromosome
  - shortened after each round of replication
  - end with multiple repeated sequences of the form

$$(5')(T_xG_y)_n$$
$$(3')(A_xC_y)_n$$

where *x* and *y* are generally between 1 and 4 and *n* is in the range of 20 to 100 for single-celled eukaryotes and >1,500 in mammals

Telomere Sequences

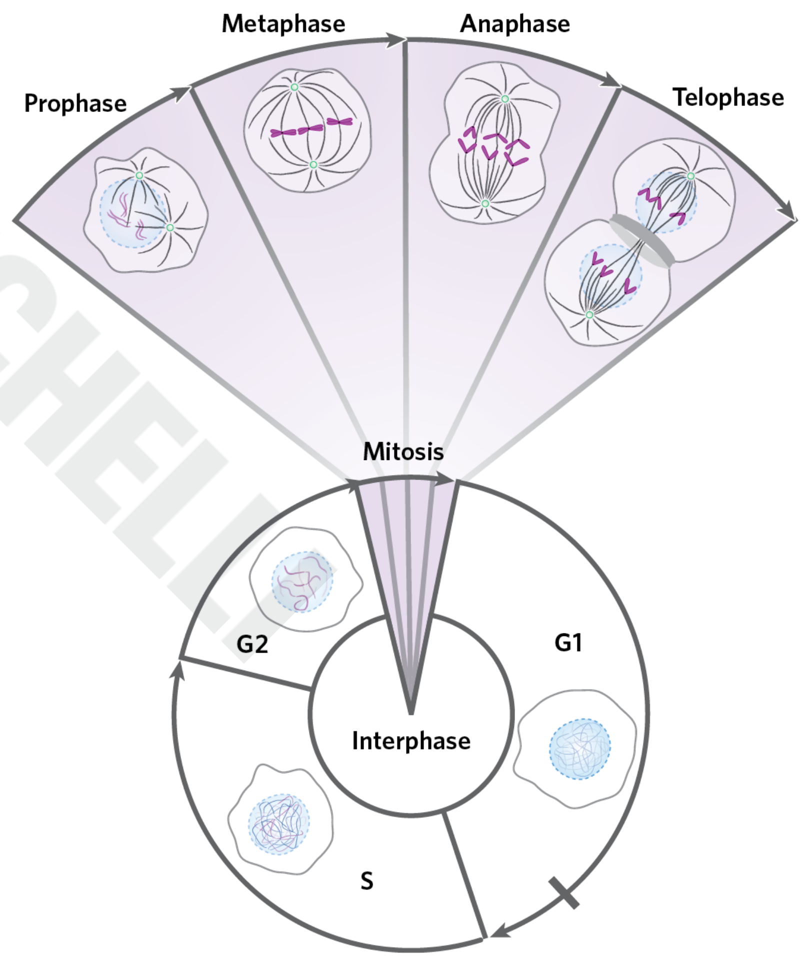
Table 24-3 Telomere Sequences	
Organism	**Telomere repeat sequence**
Homo sapiens (human)	$(TTAGGG)_n$
Tetrahymena thermophile (ciliated protozoan)	$(TTGGGG)_n$
Saccharomyces cerevisiae (yeast)	$(T(G)_{1–3}(TG)_{2–3})_n$
Arabidopsis thaliana (plant)	$(TTTAGGG)_n$

24.3 The Structure of Chromosomes

Chromatin Consists of DNA, Proteins, and RNA

- **chromatin** = eukaryotic chromosomal material composed of DNA, RNA, and proteins
  - amorphous in G0 and interphase (G1, S, and G2) phases

- in the S phase of interphase, the amorphous DNA replicates to produce sister chromatids



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

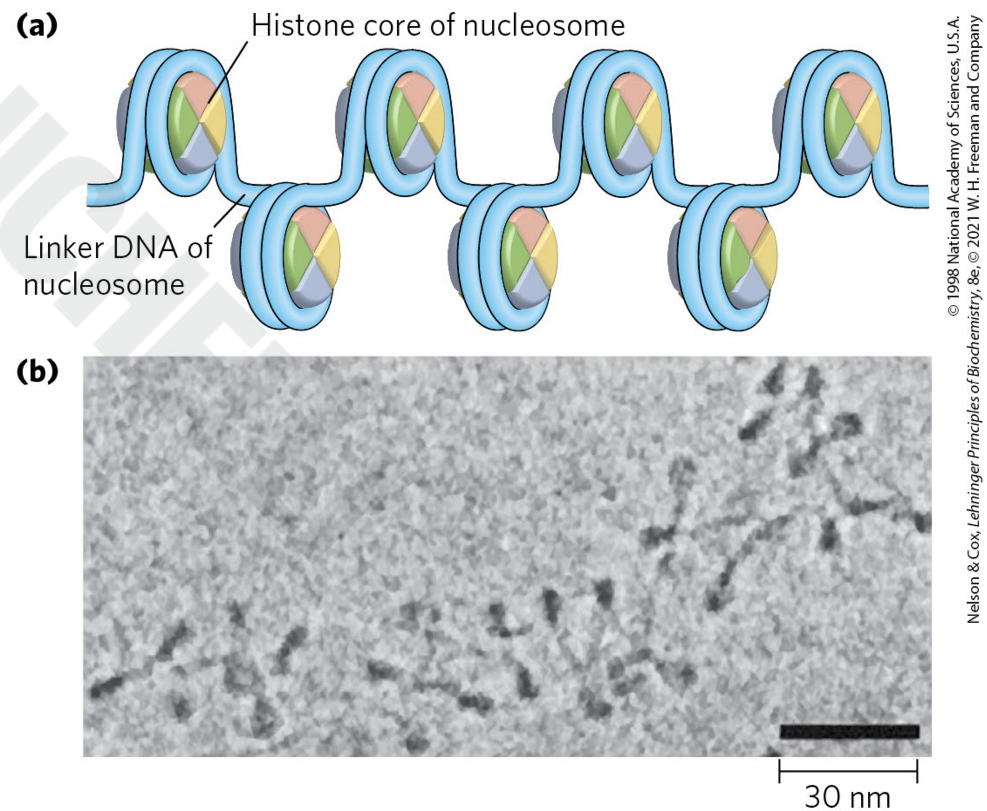
P4 Principle 4 (2 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

Histones Package and Order DNA into Nucleosomes

- **histones** = proteins that are tightly associated with chromatin and function to package and order the DNA

- **nucleosomes** = the fundamental structural unit of chromatin
  - composed of core histone proteins bound to DNA



Histones Are Small, Basic Proteins

- all eukaryotic cells have five major classes of histones
  - H3 and H4 are nearly identical in all eukaryotes
  - H1, H2A, and H2B show less sequence similarity across eukaryotic species

Table 24-5 Types and Properties of the Common Histones

Histones	Molecular weight	Number of amino acid residues	Content of basic amino acids (% of total): Lys	Content of basic amino acids (% of total): Arg
H1[a]	21,130	223	29.5	11.3
H2A[a]	13,960	129	10.9	19.3
H2B[a]	13,774	125	16.0	16.4
H3	15,273	135	19.6	13.3
H4	11,236	102	10.8	13.7

P4 Principle 4 (3 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

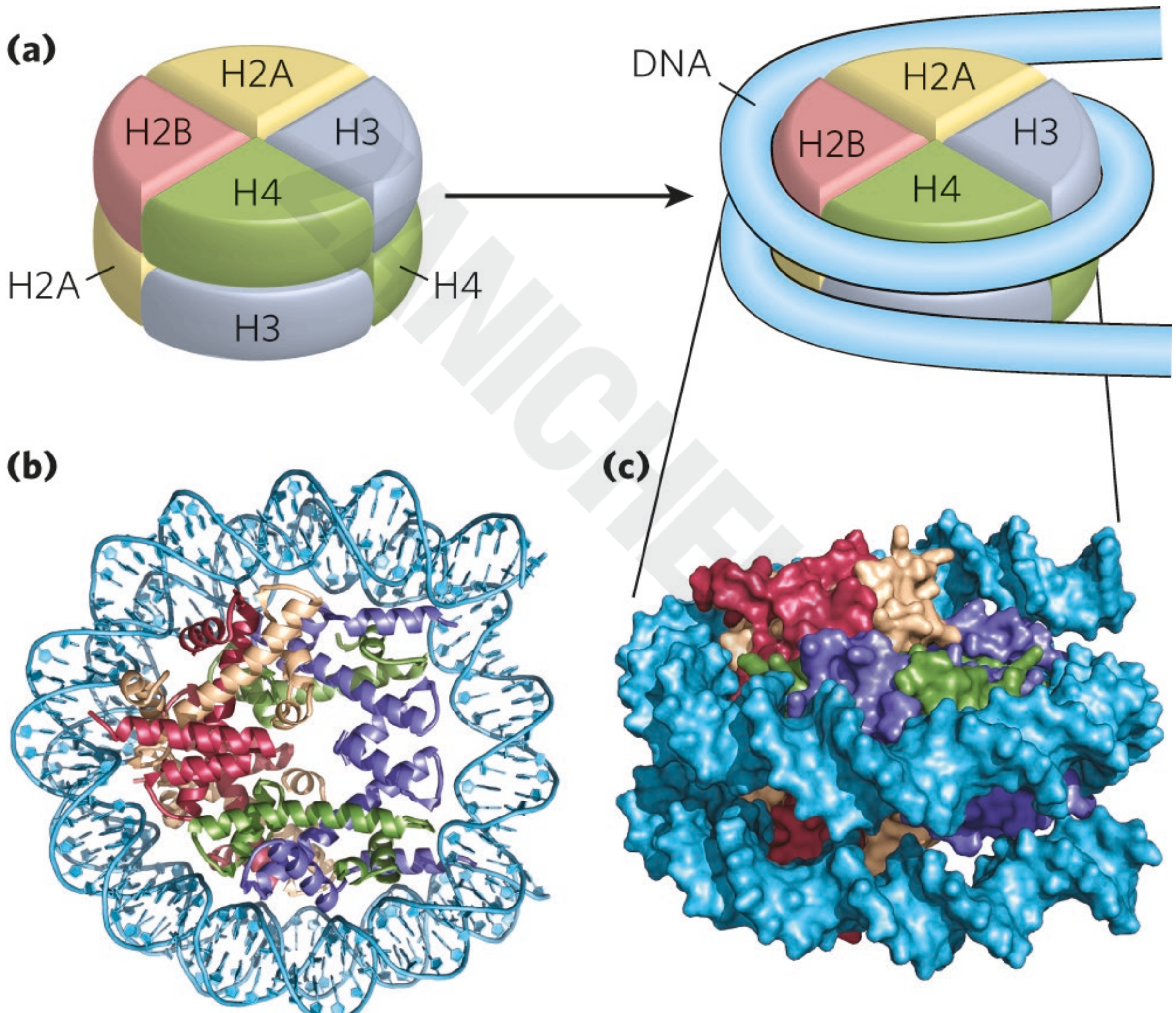
Nucleosomes Are the Fundamental Organizational Units of Chromatin

- chromosomes are arranged as “beads-on-a-string”

- the beads are complexes of histones and DNA
  - nucleosome contains eight histone molecules: two copies each of H2A, H2B, H3, and H4
  - contains ~200 bp of DNA, of which 146 bp are bound tightly around the histone core in a left-handed solenoidal supercoil

- the nucleosome is the bead plus the connecting DNA that leads to the next bead

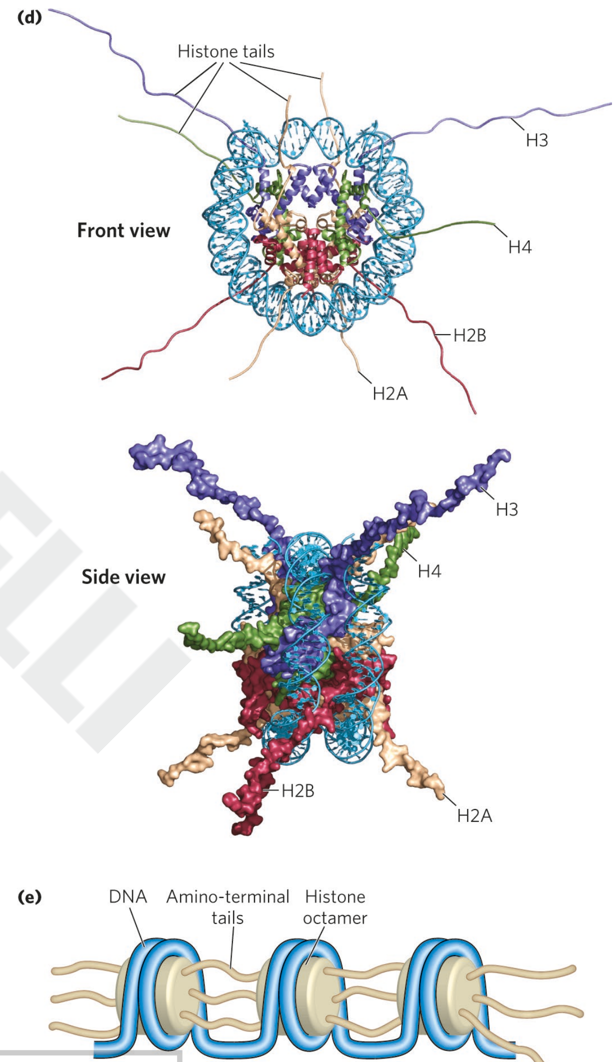
DNA Wrapped Around a Histone Core



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Front and Side Views of Histone Amino-Terminal Tails

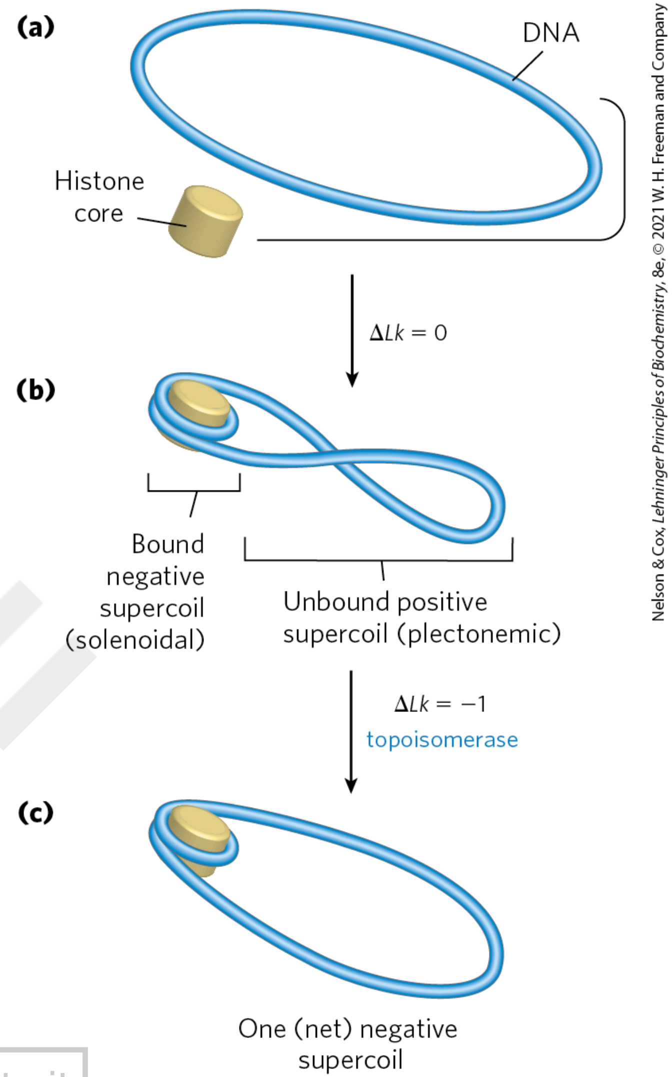
- the amino-terminal tails of the histones:
  - are intrinsically disordered
  - are where most of the histone modifications occur
  - play a key role in forming contacts between nucleosomes in the chromatin



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

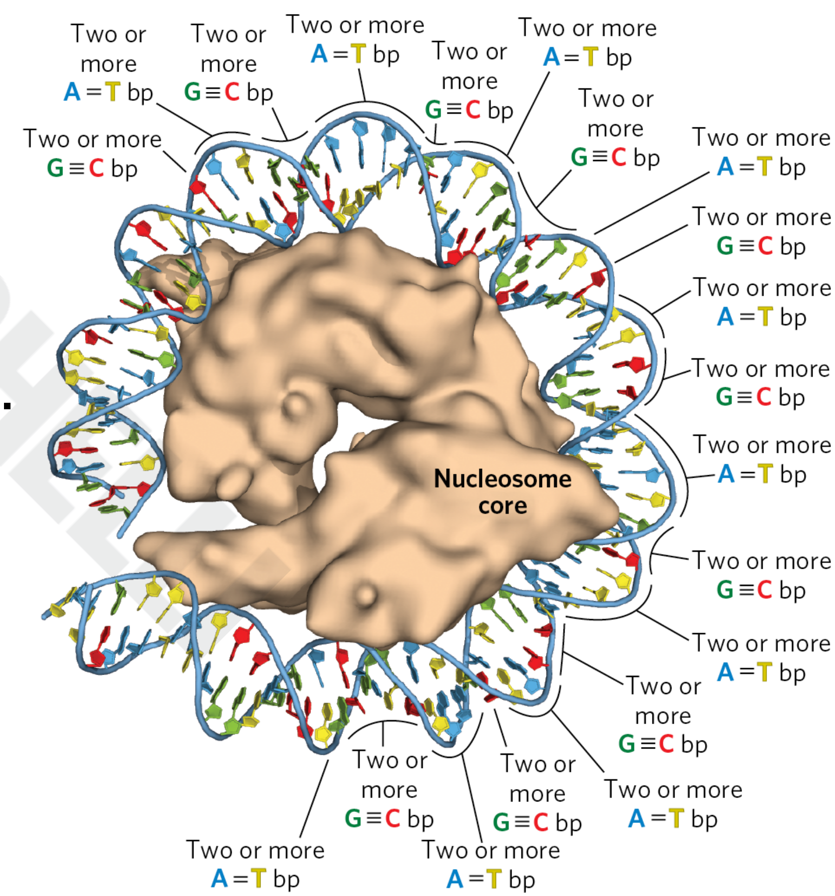
Chromatin Assembly

- DNA wrapping around the histone core requires removal of ~1 helical turn in the DNA
- binding of the histone core to a relaxed closed-circular DNA introduces a negative supercoil
  - accompanied by a positive supercoil in the unbound region
- when eukaryotic topoisomerases relax positive supercoils, *Lk* decreases



The Effect of DNA Sequence on Nucleosome Binding

- histone binding is not random
- binding of the histone core seems to depend on a local abundance of A=T base pairs.
- staggering AA, AT, or TT dinucleotides at 10 bp intervals facilitates DNA binding around the histone core



Depositing Nucleosome Cores on DNA

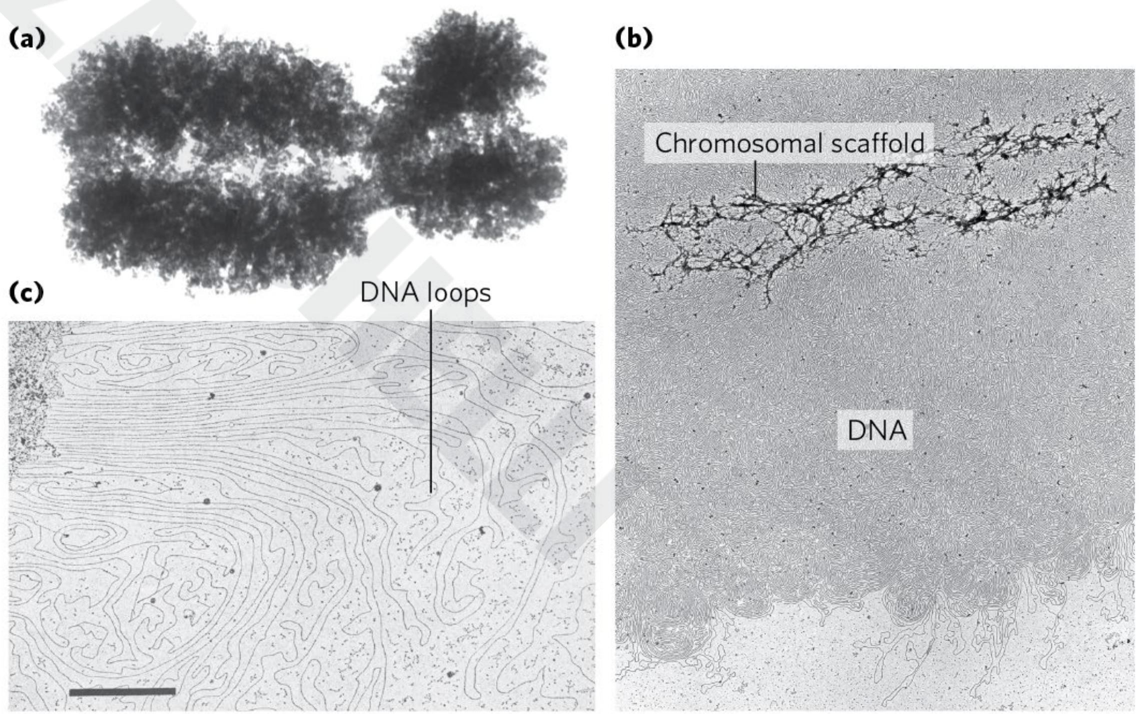
- deposited stepwise: a tetramer of two H3 and two H4 histones binds first, followed by two H2A–H2B dimers
- **histone chaperones** = mediate the incorporation of nucleosomes into chromosomes after chromosomal replication
  - conserved in all eukaryotes
- **histone exchange factors =** permit the substitution of histone variants for core histones in contexts other than postreplication

P4 Principle 4 (4 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

Nucleosomes Are Packed into Highly Condensed Chromosome Structures

- certain regions of DNA seem to associate with a chromosomal scaffold

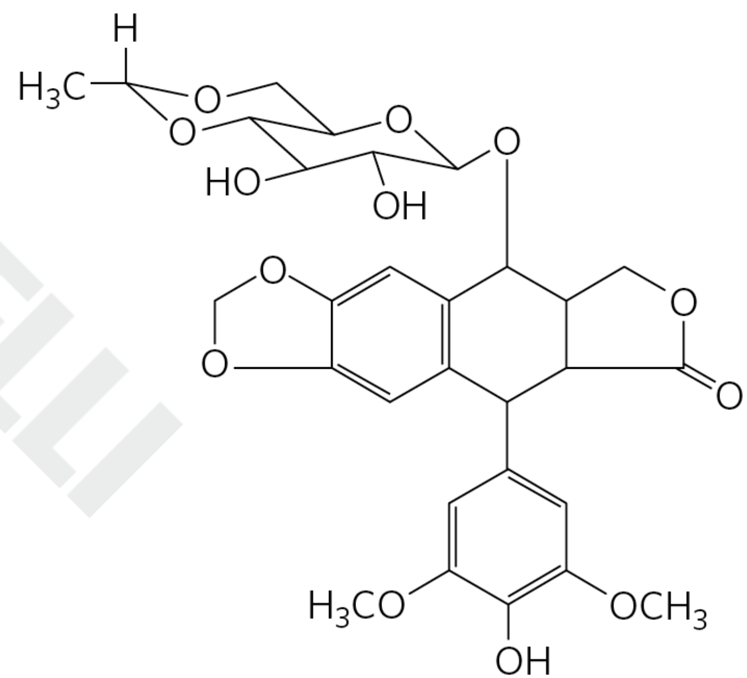


P3 Principle 3 (6 of 7)

Chromosomes in all cells are maintained in a state of torsional stress. DNA is underwound relative to the stable B-form structure, facilitating both the packaging of DNA and access to the genetic information contained within it.

Topoisomerase II Is One of the Most Abundant Proteins in the Chromosome

- topoisomerase II is important to the maintenance of chromatin structure
- inhibitors of topoisomerase II can kill rapidly dividing cells
  - several chemotherapy drugs are topoisomerase II inhibitors

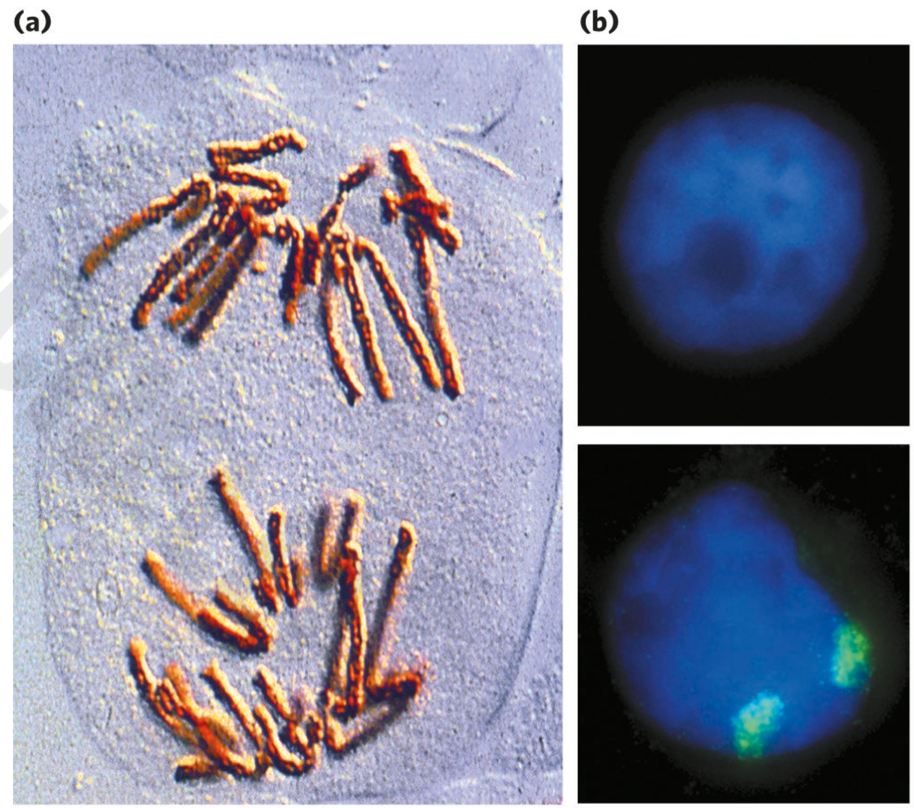


Etoposide

Chromosomal Organization in Eukaryotic Nucleus

- chromosomes are highly condensed and organized just before cell division

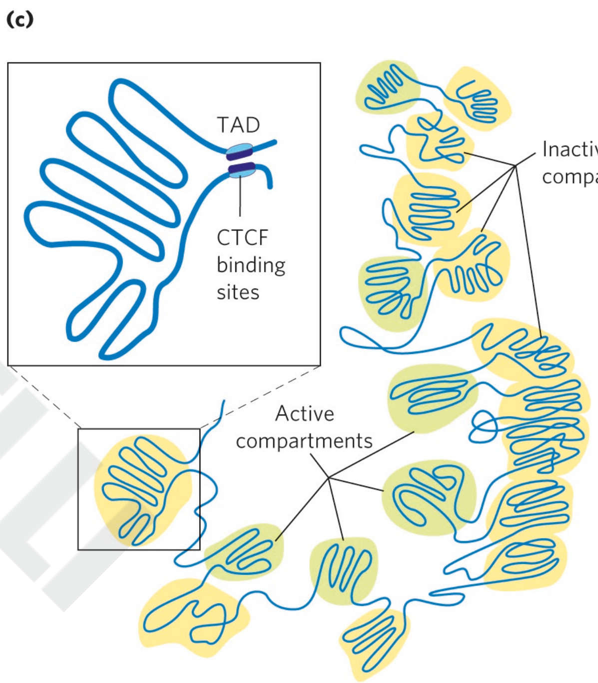
- chromosomes appear dispersed during interphase
  - each chromosome is still organized, with two sets of compartments



(a) Pr. G. Giménez-Martín/Science Source. (b) Karen Meaburn and Tom Misteli/National Cancer Institute.
Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Active and Inactive Compartments

- active compartments have reduced chromatin condensation
- inactive compartments **(heterochromatin)** are highly condensed
- **topologically associating domains (TADs)** = large segments of DNA are organized in loops
  - binding of CTCF to bordering sites brings DNA together



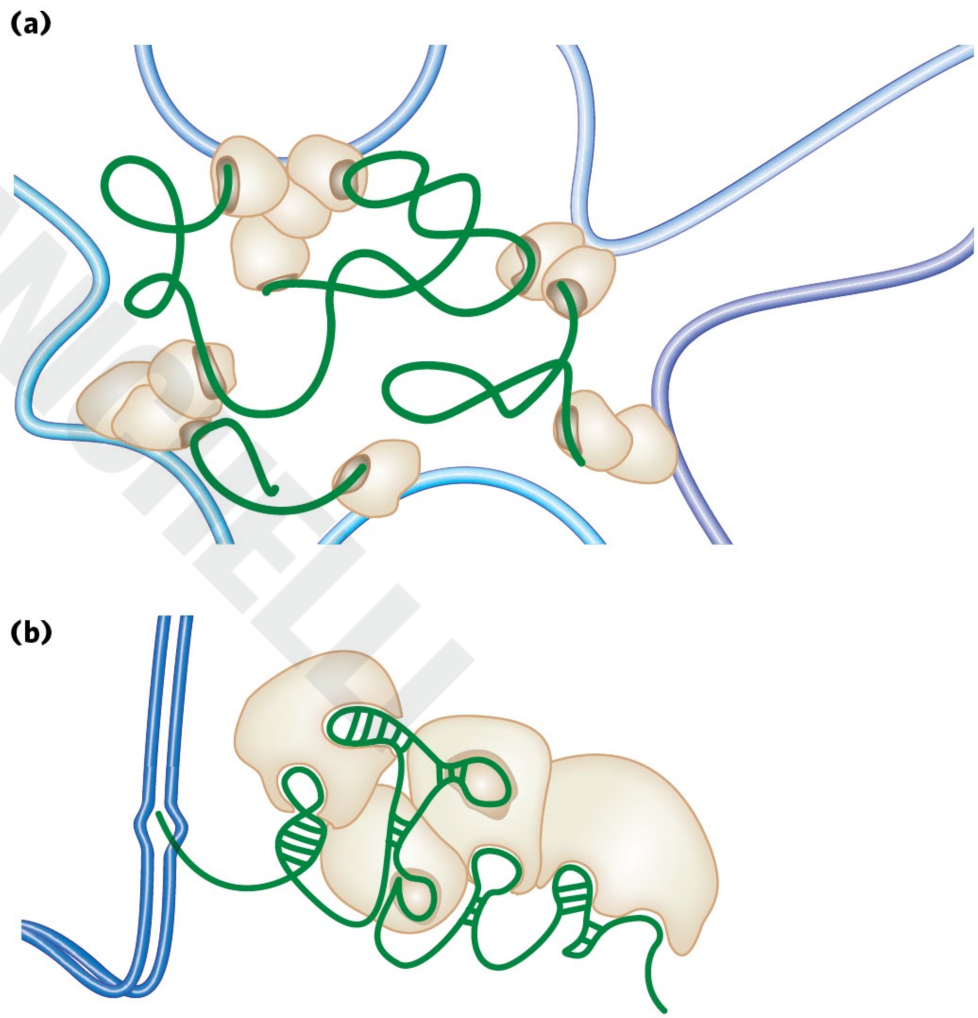
(a) Pr. G. Giménez-Martín/Science Source. (b) Karen Meaburn and Tom Misteli/National Cancer Institute.
Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

P4 Principle 4 (5 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

Long Noncoding RNAs (lncRNAs)

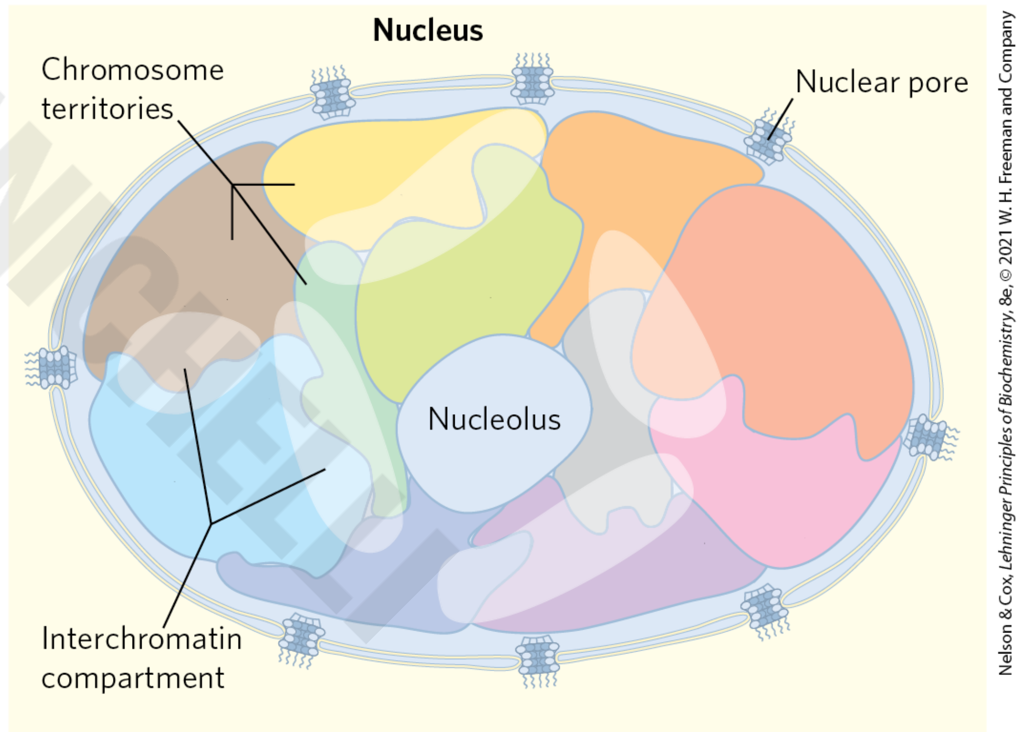
- **long noncoding RNAs (lncRNAs)** = play a functional role in defining the chromosome structure
  - many provide a scaffold for proteins



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Chromosome Territory

- **chromosome territory** = subnuclear domain that constrains the entire structure of each chromosome
  - little or no intermingling of DNA in different territories

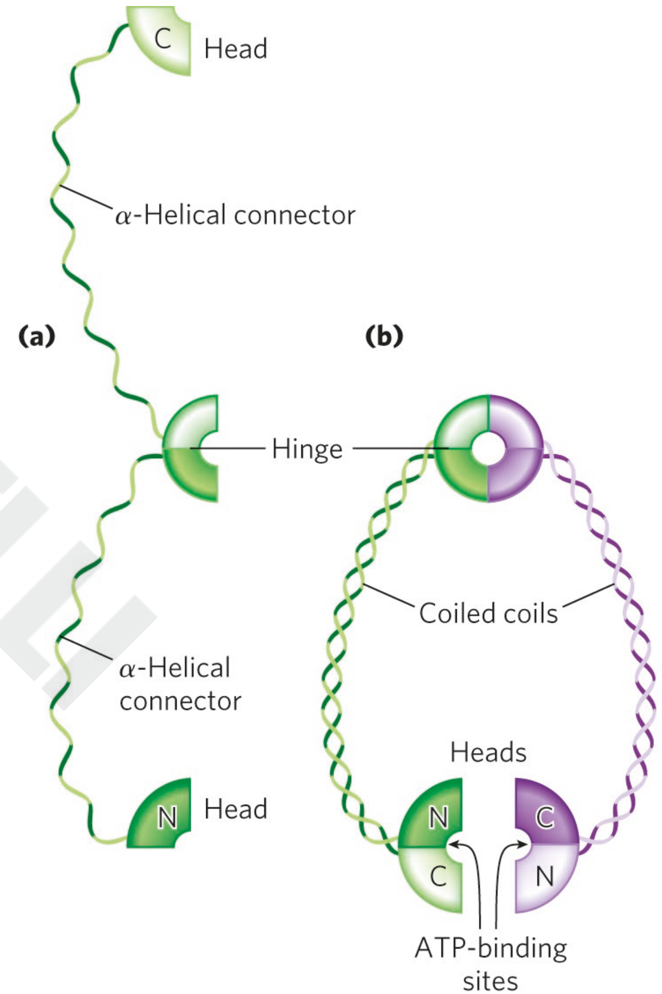


P4 Principle 4 (6 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

Condensed Chromosome Structures Are Maintained by SMC Proteins

- **SMC proteins** (*s*tructural *m*aintenance of *c*hromosomes) = responsible for maintaining the structure and integrity of chromosomes following replication
  - consist of 5 distinct domains

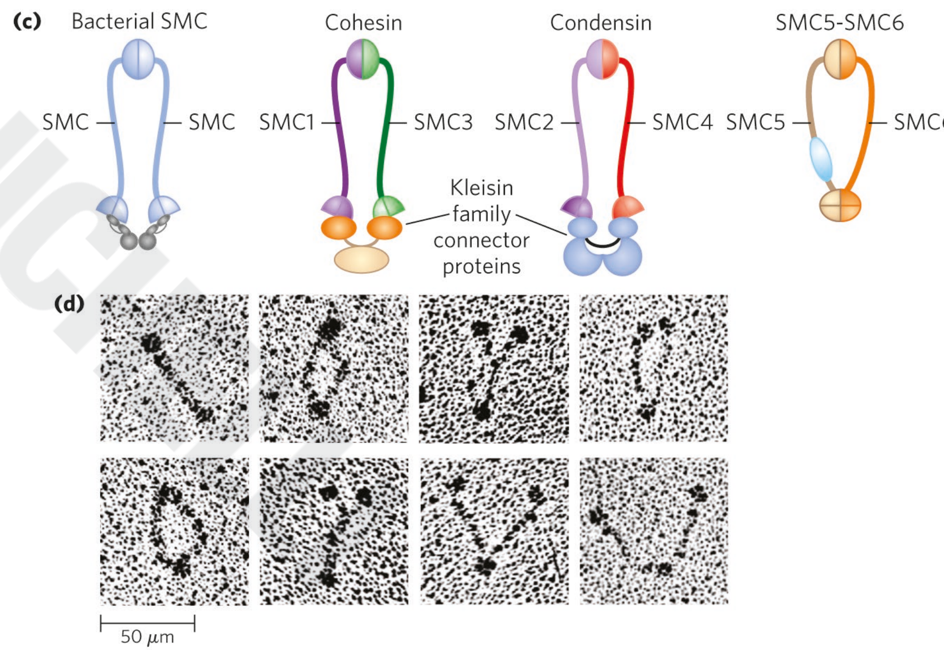


Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Eukaryotes Have Two Major Types of SMC Proteins

- **cohesins** = link sister chromatids together after replication and keep them together as the chromosomes condense to metaphase

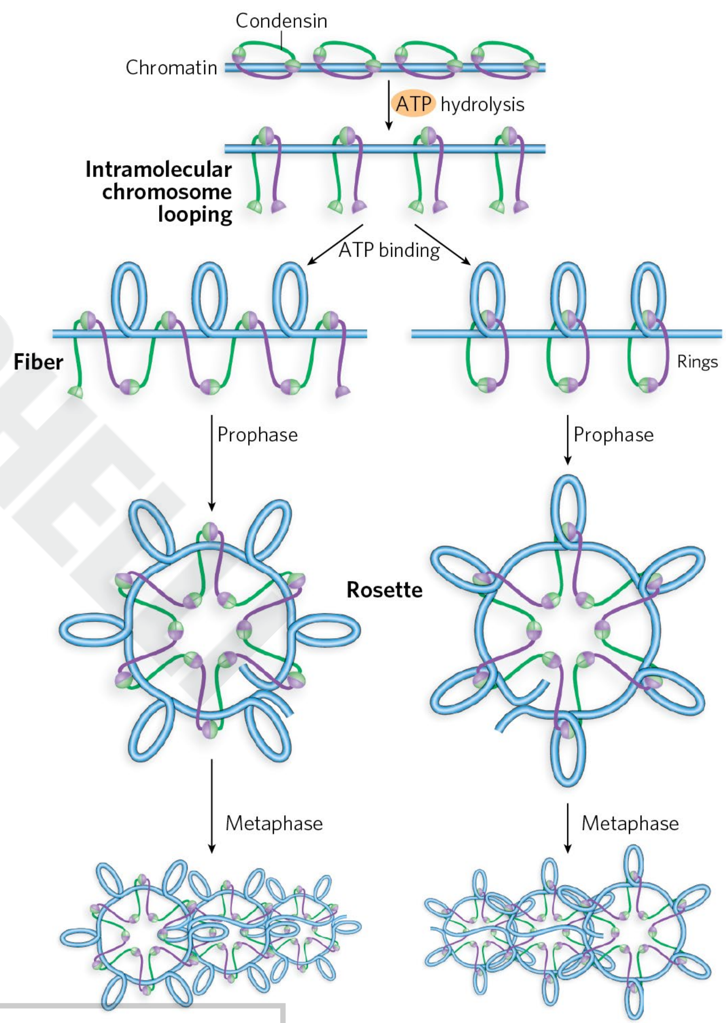
- **condensins** = essential to chromosomal condensation as cells enter mitosis
  - create positive supercoils



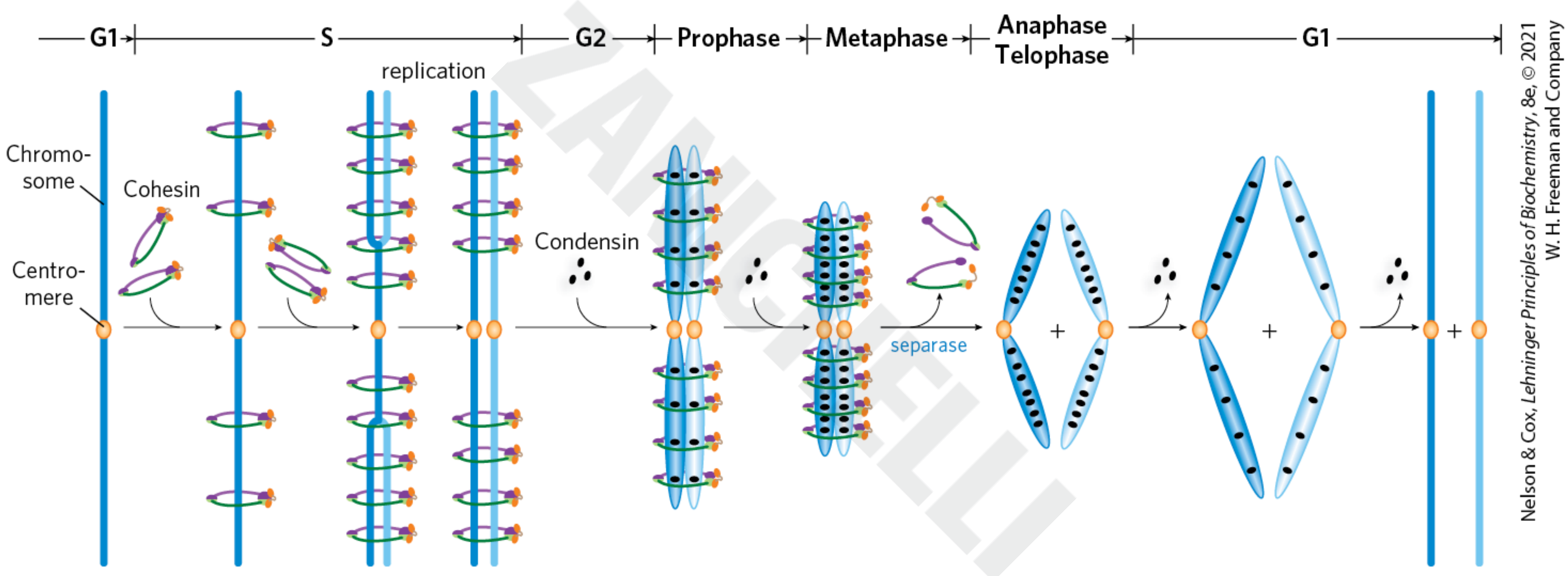
Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

Models of the Role of Condensins in Chromatin Condensation

- as DNA is compacted to form tighter and tighter loops, condensins stabilize the loops by binding at the base of each one



The Role of Cohesins and Condensins in the Cell Cycle



P2 Principle 2 (6 of 6)

Chromosomes are large. To constrain them in a small space can require multiple layers and multiple modes of tertiary structure.

Bacterial DNA Is Also Highly Organized

- bacterial DNA is compacted in a structure called the **nucleoid**
  - can occupy a significant fraction of the cell volume
  - appears attached at 1+ points in the inner surface of the plasma membrane

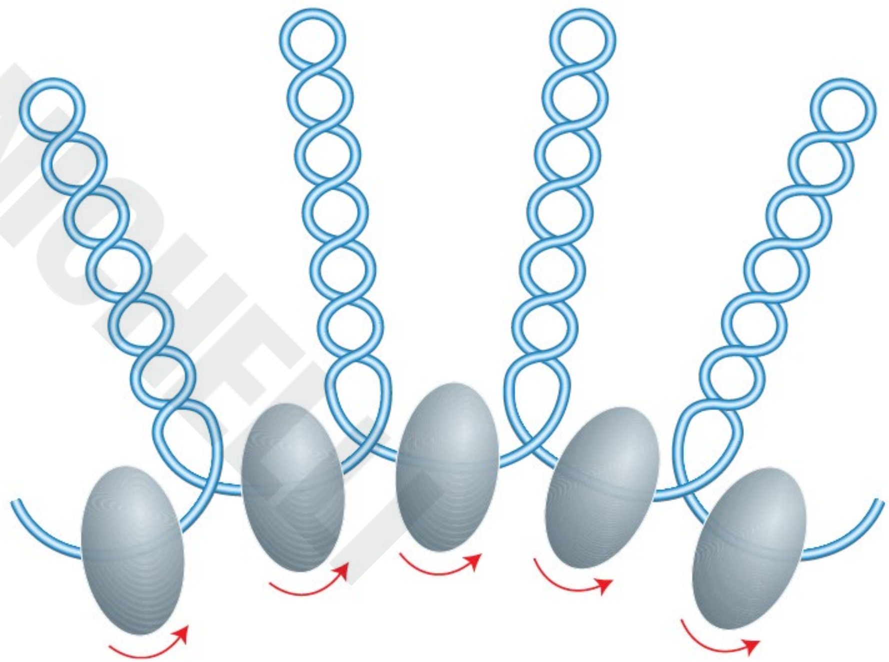
Lars Renner.

P3 Principle 3 (7 of 7)

Chromosomes in all cells are maintained in a state of torsional stress. DNA is underwound relative to the stable B-form structure, facilitating both the packaging of DNA and access to the genetic information contained within it.

Looped Domains of the *E. coli* Chromosome

- a scaffoldlike structure seems to organize the circular chromosome into a series of ~500 looped domains
  - domains are topologically constrained
  - domains do not have fixed end points



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

P4 Principle 4 (7 of 7)

Specialized proteins and RNAs maintain chromosome structure. Topoisomerases control DNA underwinding. Histones, condensins, cohesins, and other DNA-binding proteins provide scaffolds to organize chromosome structure. Certain long, noncoding RNAs also play important roles in chromosome structure and function.

Histonelike Proteins Are Abundant in *E. coli*

- bacterial DNA does not seem to have comparable structure to eukaryotic nucleosomes
- histonelike proteins = proteins that bind and dissociate within minutes
  - no regular, stable DNA-histone structure has been found
- the dynamic and irregular structure of the bacterial chromosome reflects:
  - the shorter cell cycle
  - very active metabolism