Genes and Chromosomes

Tong-Jin Zhao

School of Life Sciences, Xiamen University

Outlines

Part I. Chromosomal Elements

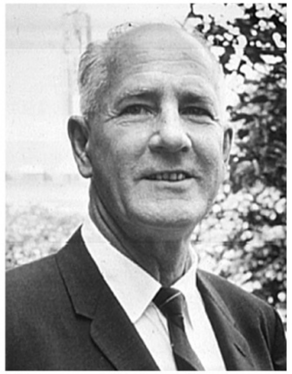
Part II. DNA Supercoiling

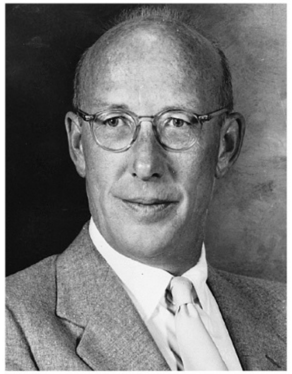
Part III. The Structures of Chromosomes

Part I . Genes and Chromosomes

Part I

Genes


George W. Beadle,
1903–1989


Edward L. Tatum,
1909–1975

one gene--one enzyme

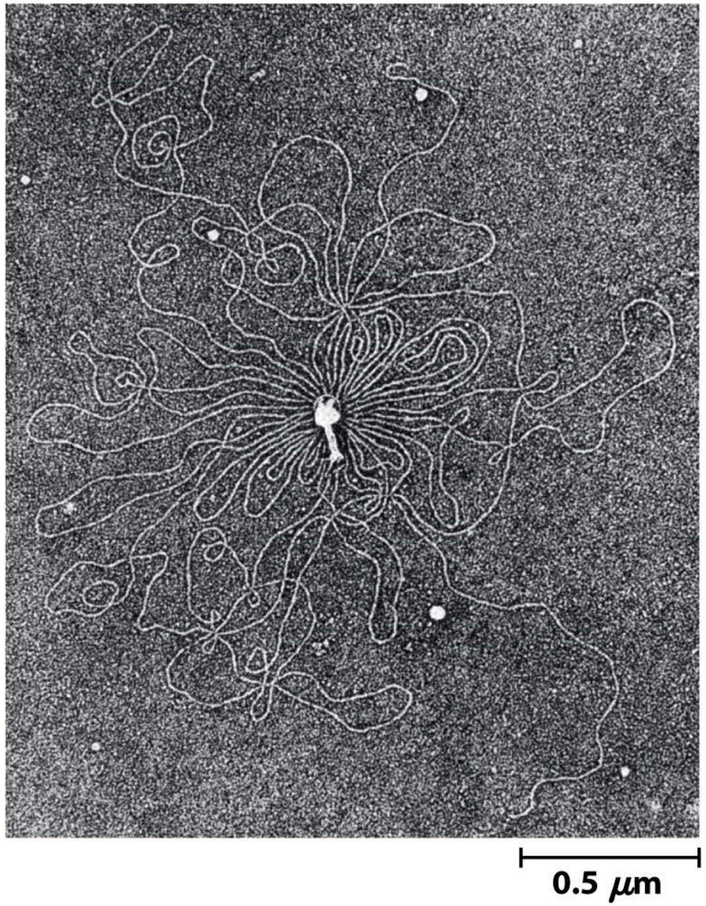
⇩

one gene--one polypeptide

⇩

A gene is all the DNA that encodes the primary sequence of some final product, which can be either a polypeptide or an RNA with a structural or catalytic function.

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



Bacteriophage T2 protein coat surrounded by its single, linear molecule of DNA.

Viruses

RNA viruses: Almost all plant viruses, and some bacterial and animal viruses.
These genomes tend to be particularly small.
HIV, ~9000 nt.

DNA Viruses: Vary greatly in size.
Replicative forms (linear to circular; single to double stranded)

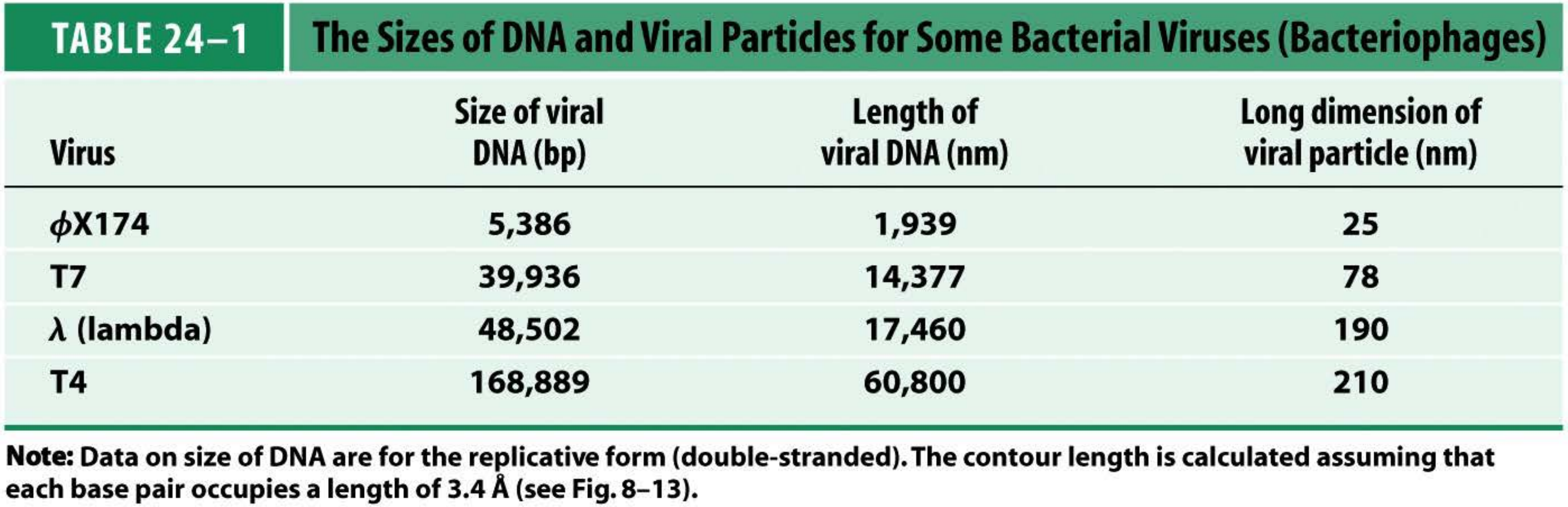
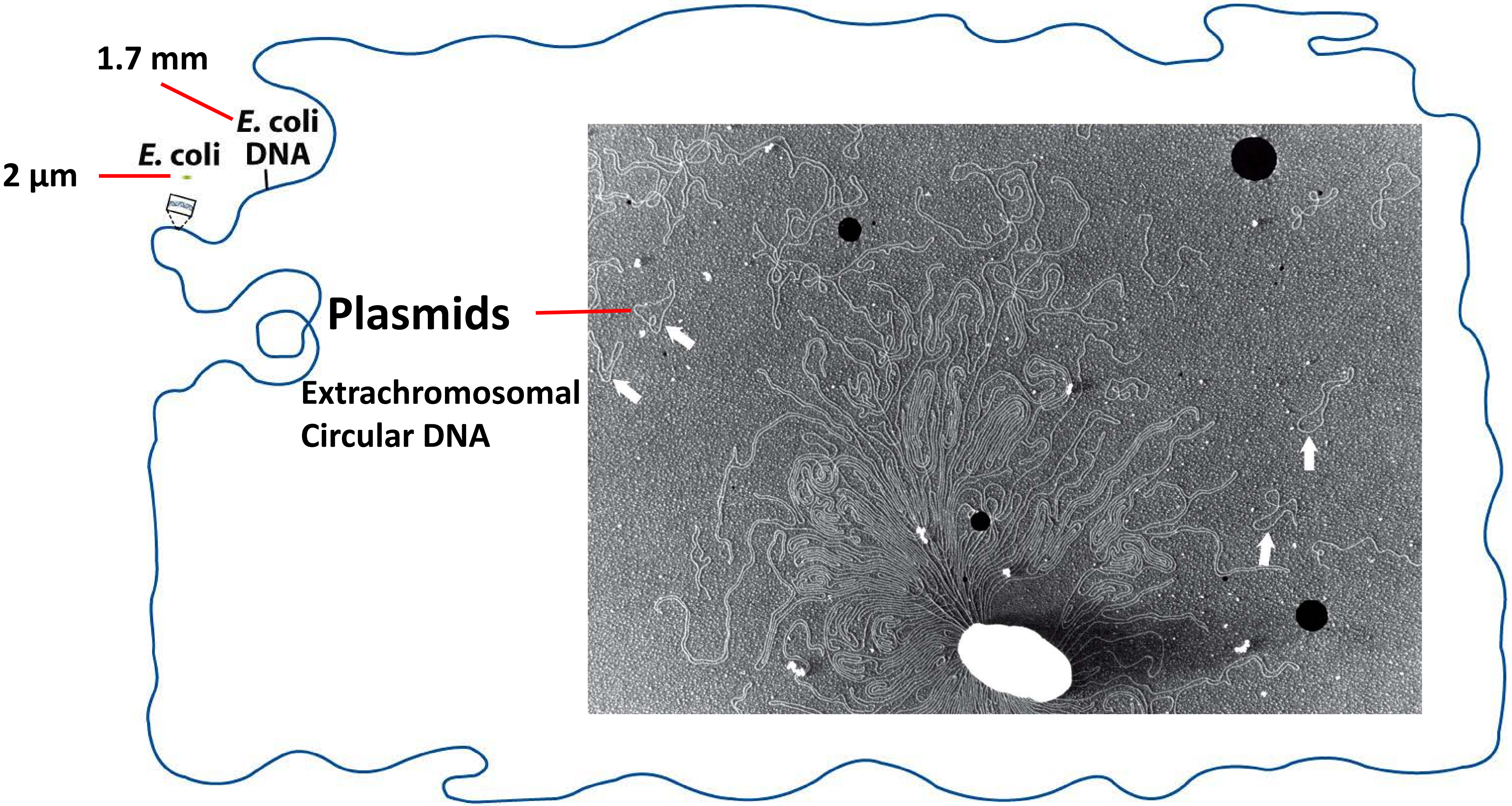


TABLE 24–1 The Sizes of DNA and Viral Particles for Some Bacterial Viruses (Bacteriophages)

Virus	Size of viral DNA (bp)	Length of viral DNA (nm)	Long dimension of viral particle (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

Note: Data on size of DNA are for the replicative form (double-stranded). The contour length is calculated assuming that each base pair occupies a length of 3.4 Å (see Fig. 8–13).

Bacteria



1.7 mm

E. coli DNA

2 µm

E. coli

Plasmids

Extrachromosomal Circular DNA

Part I

Eukaryotes

Much larger genome sizes

Other than chromosomal DNA, mitochondrial and/or chloroplast DNA

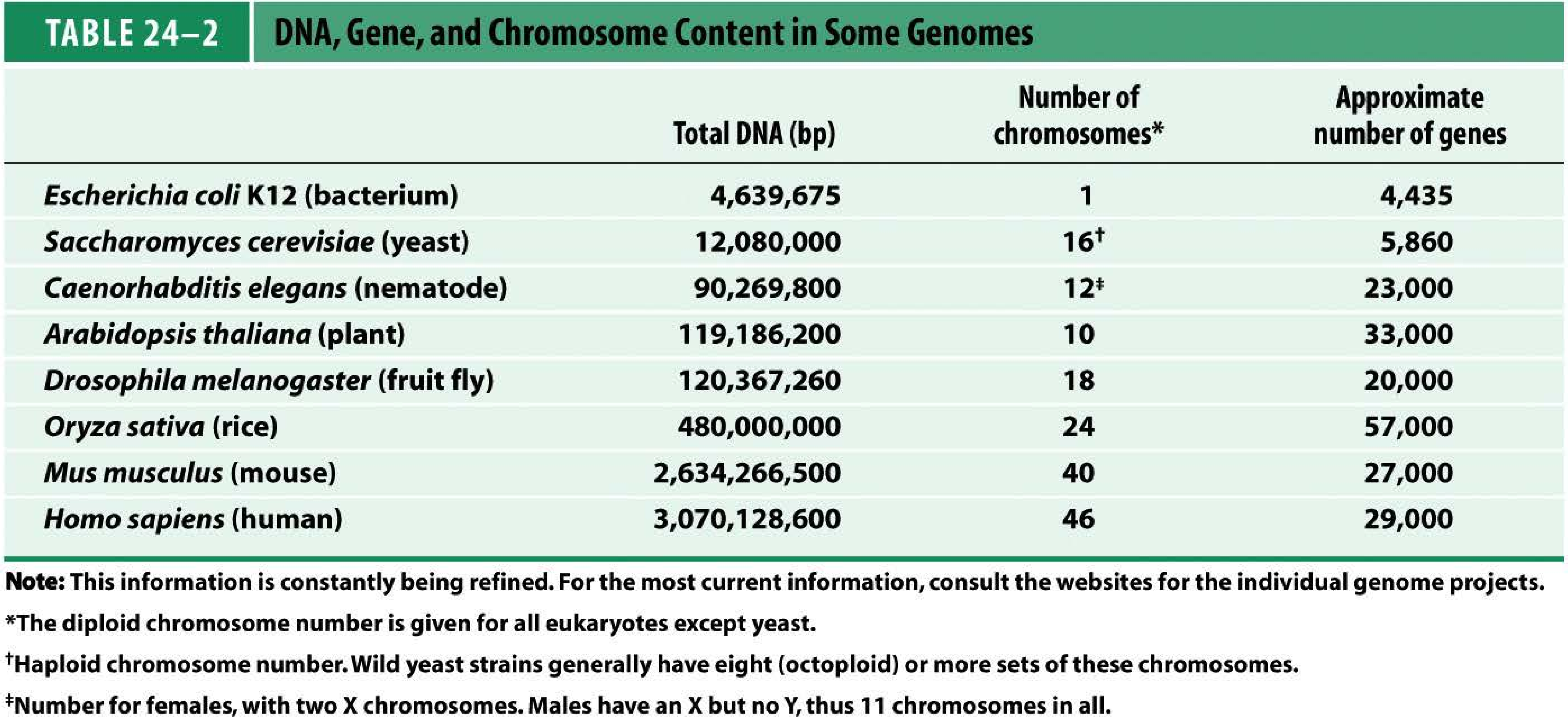


TABLE 24–2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes*	Approximate number of genes
Escherichia coli K12 (bacterium)	4,639,675	1	4,435
Saccharomyces cerevisiae (yeast)	12,080,000	16†	5,860
Caenorhabditis elegans (nematode)	90,269,800	12‡	23,000
Arabidopsis thaliana (plant)	119,186,200	10	33,000
Drosophila melanogaster (fruit fly)	120,367,260	18	20,000
Oryza sativa (rice)	480,000,000	24	57,000
Mus musculus (mouse)	2,634,266,500	40	27,000
Homo sapiens (human)	3,070,128,600	46	29,000

Note: This information is constantly being refined. For the most current information, consult the websites for the individual genome projects.

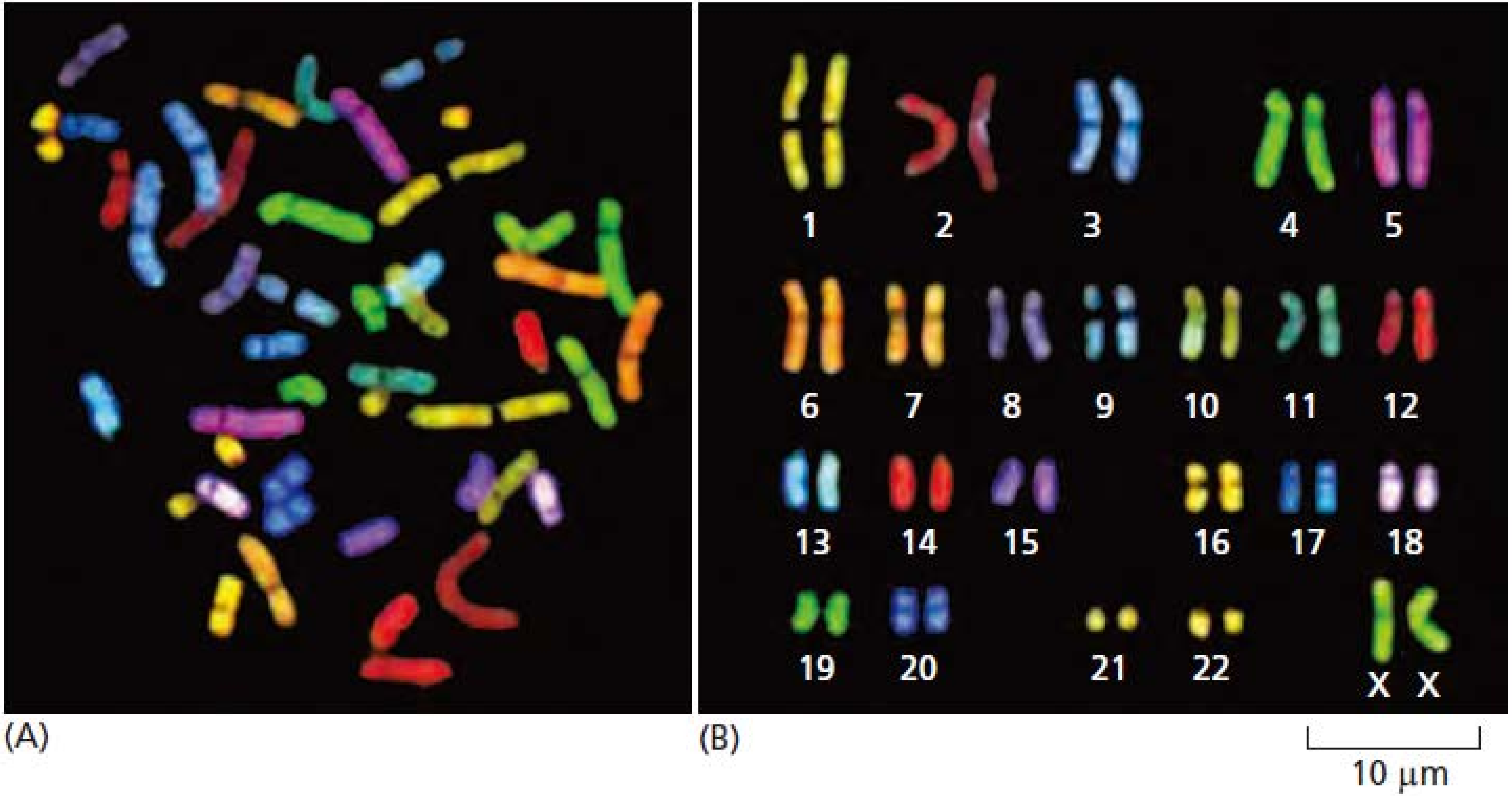
*The diploid chromosome number is given for all eukaryotes except yeast.

†Haploid chromosome number. Wild yeast strains generally have eight (octoploid) or more sets of these chromosomes.

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

Part I

Eukaryotic DNA Is Packaged into a Set of Chromosomes

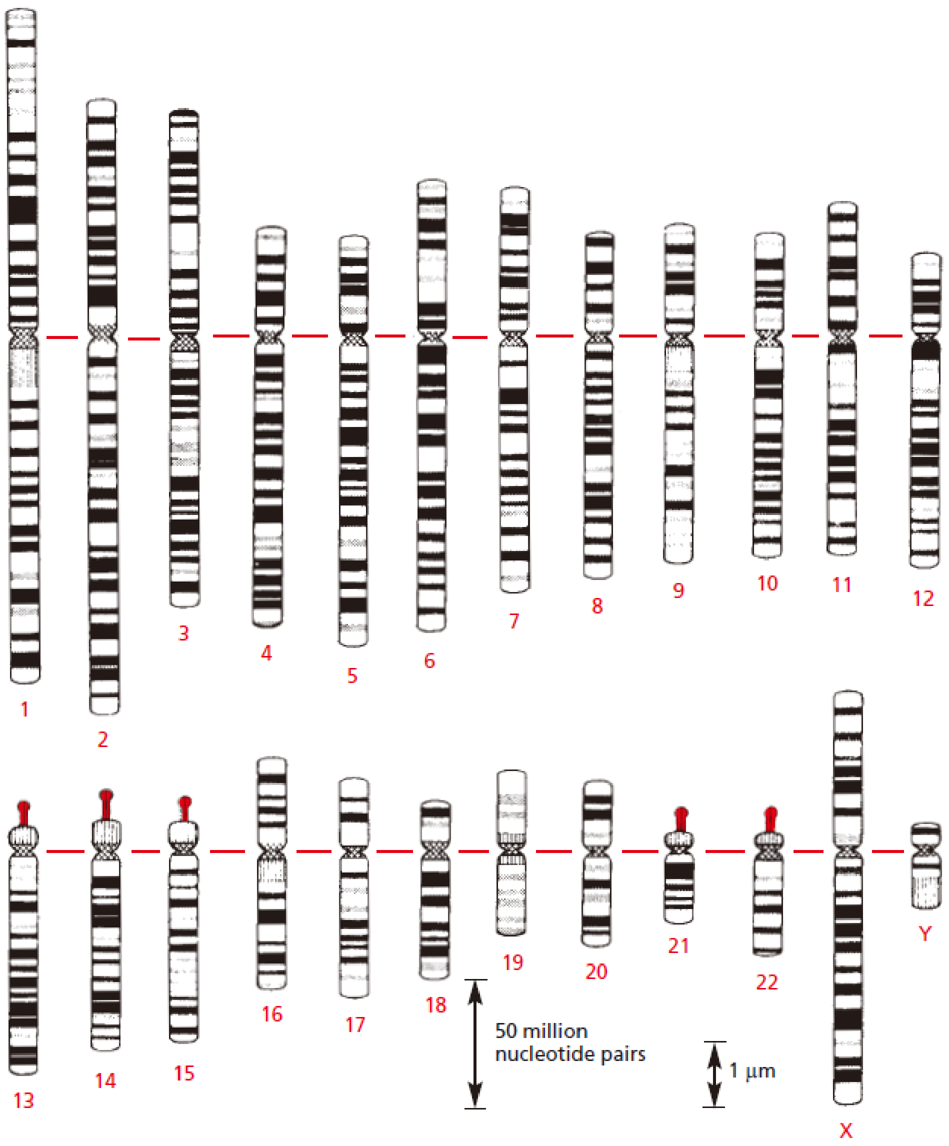


The human **Karyotype**.

DNA hybridization.

Part I

The Banding Patterns of Human Chromosomes



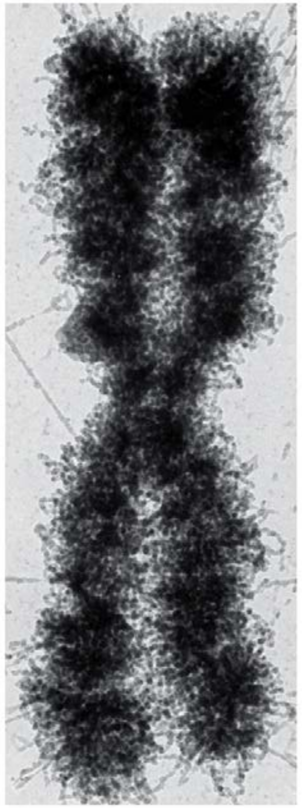
Giemsa staining

Chromosomes

Each chromosome in a eukaryotic cells consists of a single, enormously long linear DNA molecule along with proteins that fold and pack the fine DNA thread into a more compact structure.

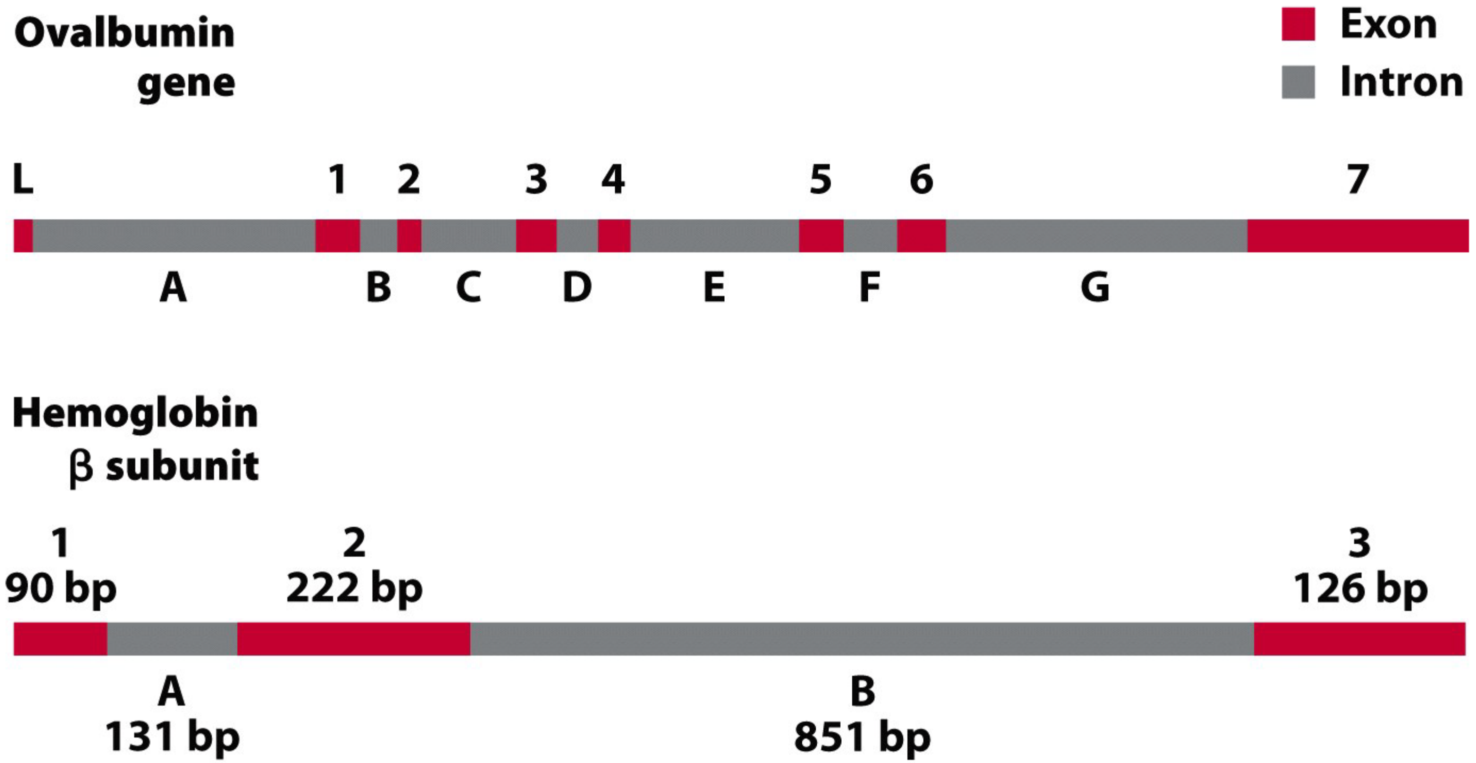
The complex of DNA and tightly bound protein is called chromatin.

Each cell, except gametes and a few highly specialized cells, has two copies of each chromosome, called homologous chromosomes (homologs).



Part I

Eukaryotic Genes and Chromosomes Are Very Complex

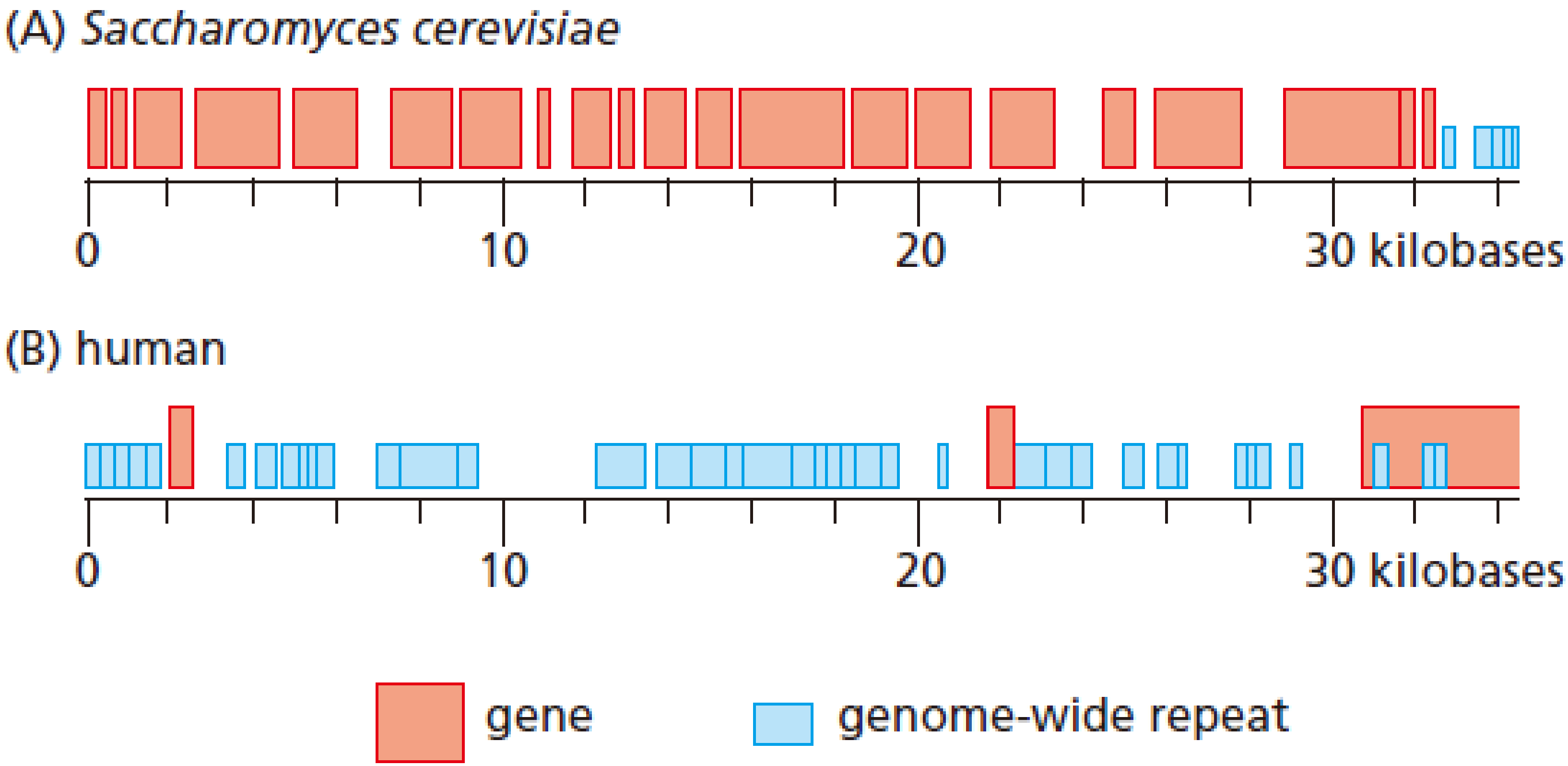


In higher eukaryotes, the typical gene has much more intron sequences than sequences devoted to exons.

Titin: 178 introns; Histones: no introns.

Part I

Chromosomes Contain Long String of Genes

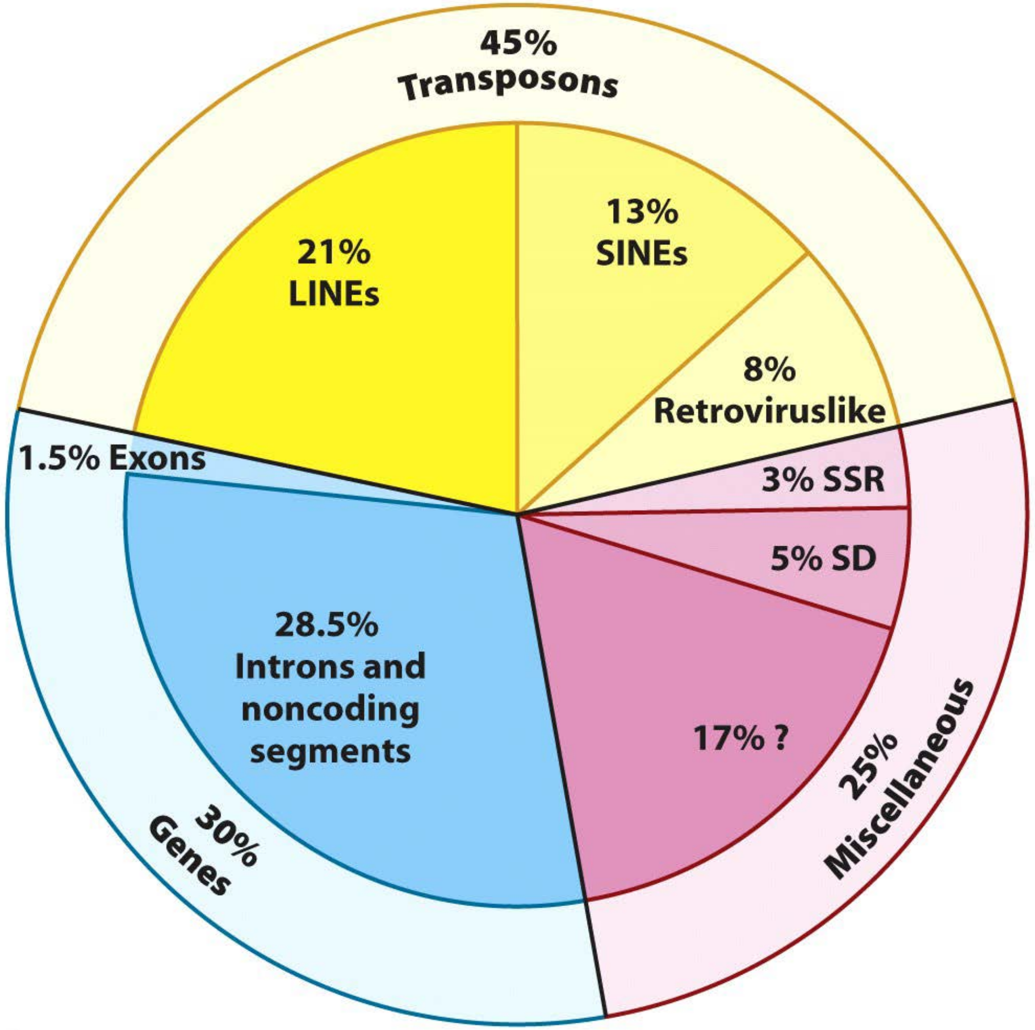


Yeast: High density of genes

Human: Genes are less densely packed and the amount of interspersed DNA sequence is far greater

Part I

Types of Sequences in Human Chromosomes



Introns

Exons

LINES: Long Interspersed Elements

SINES: Short Interspersed Elements

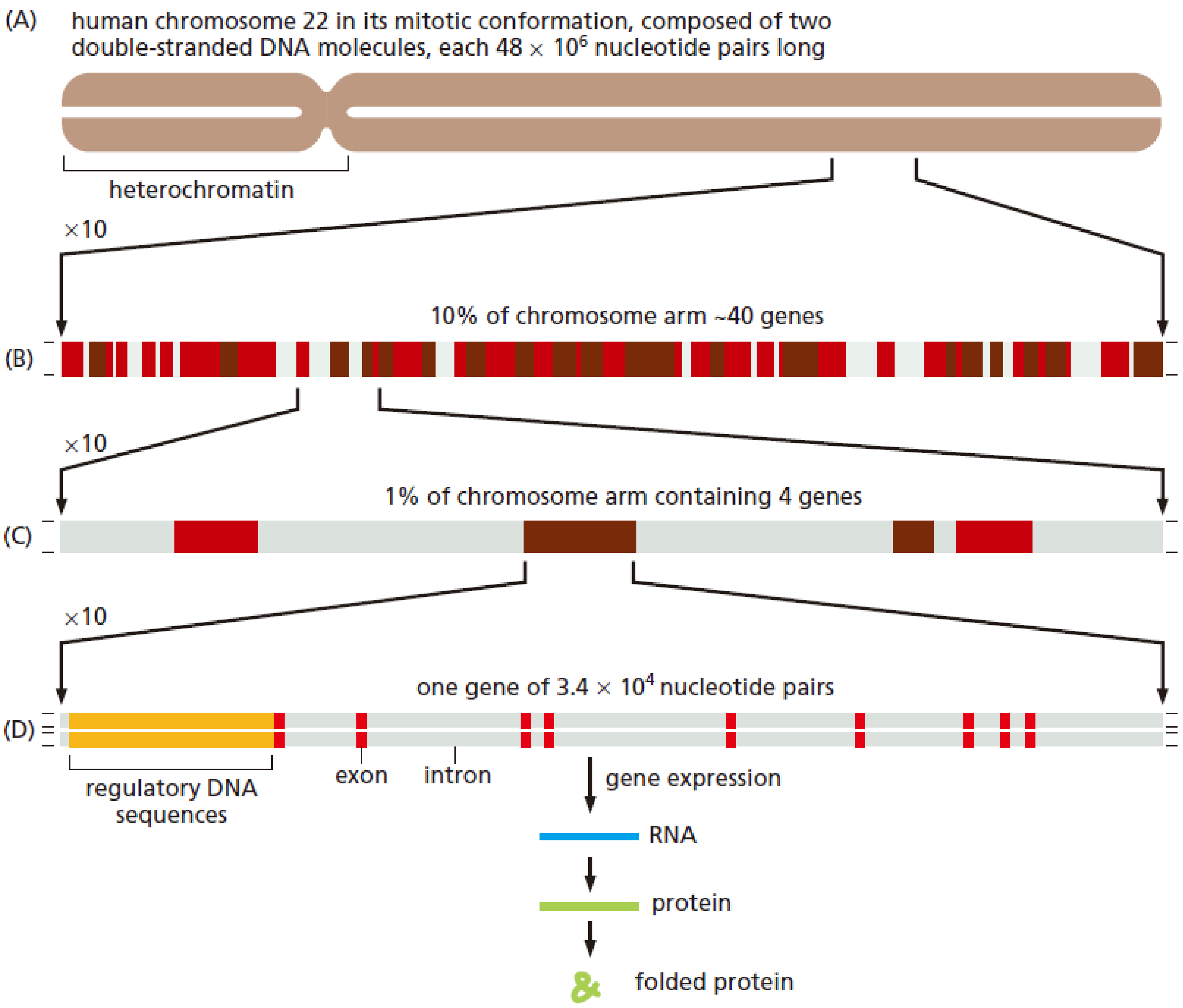
SSR: Simple Sequence Repeat

SD: Large Segmental Repeat

Completion of human genome sequence in 2004.

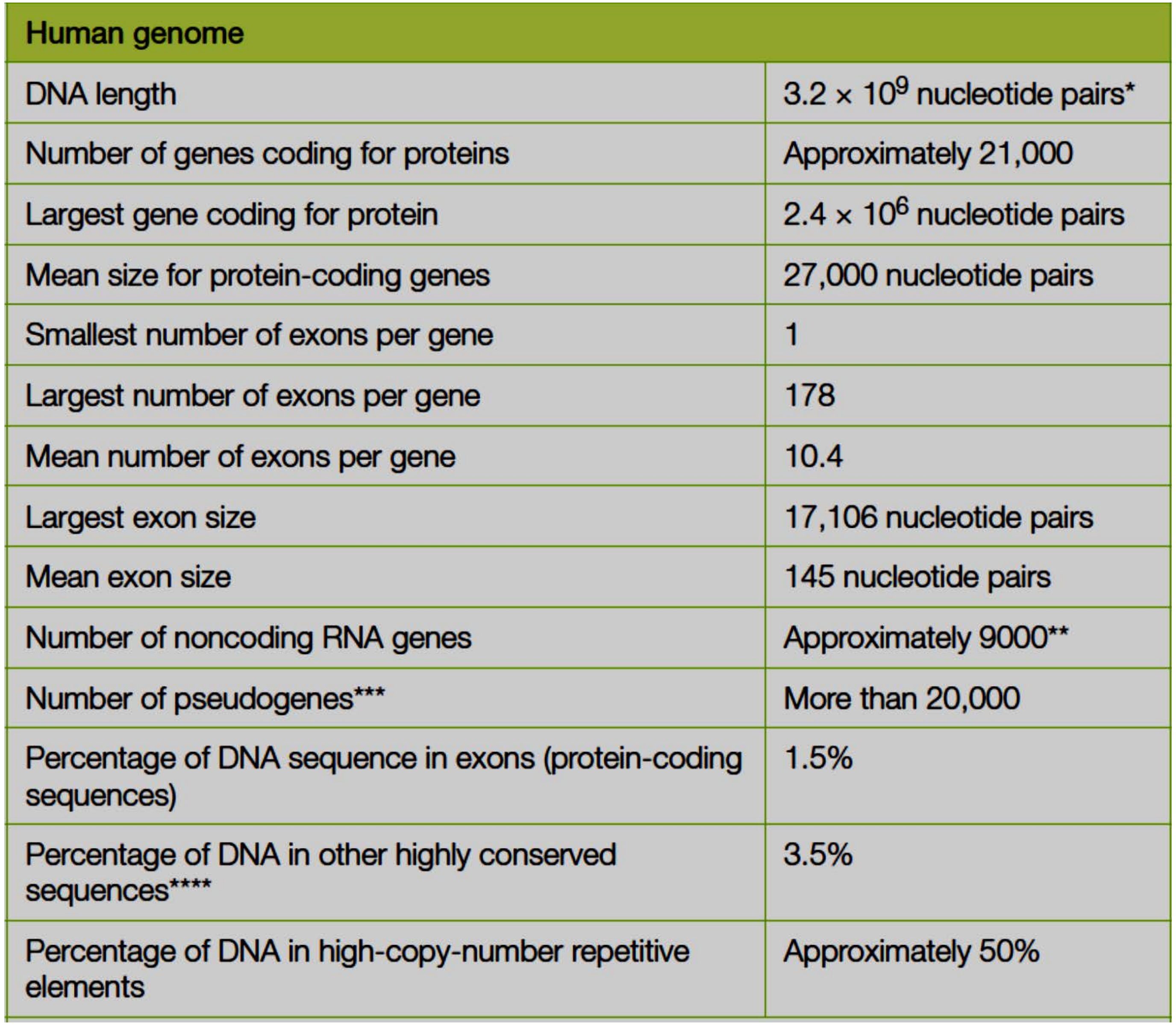
Part I

The Organization of Genes on a Human Chromosome



Part I

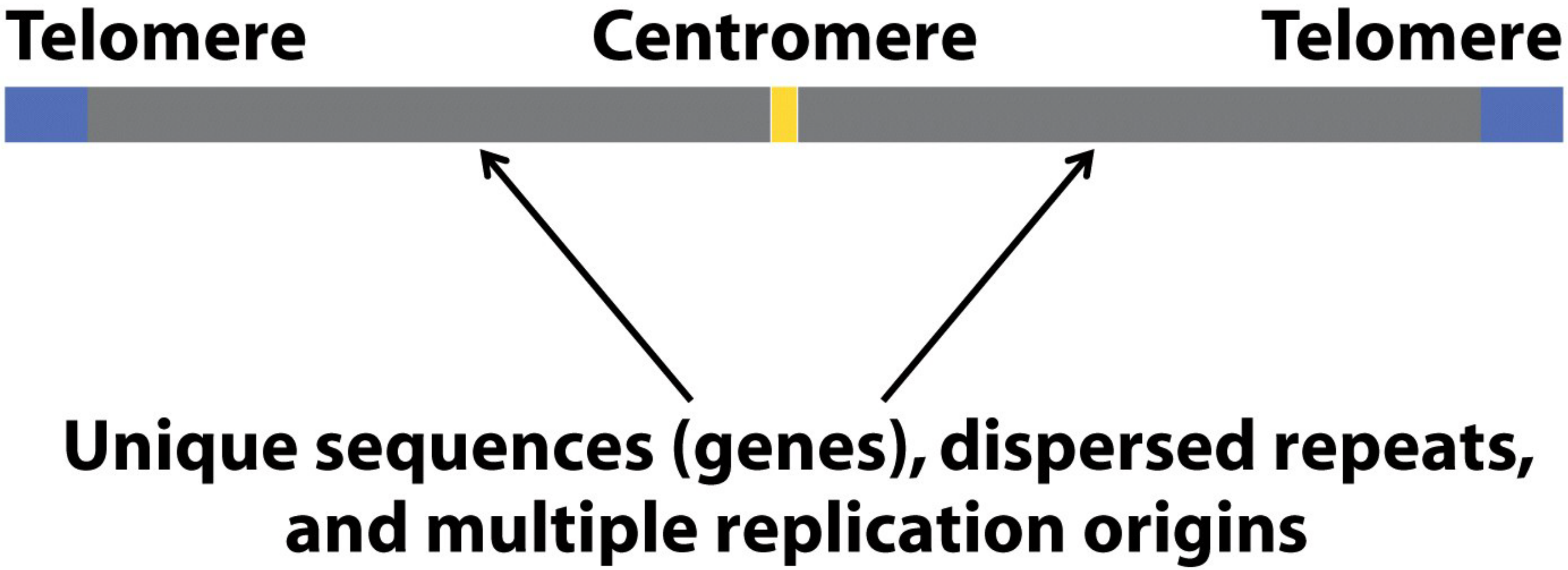
Some Vital Statistics for the Human Genome



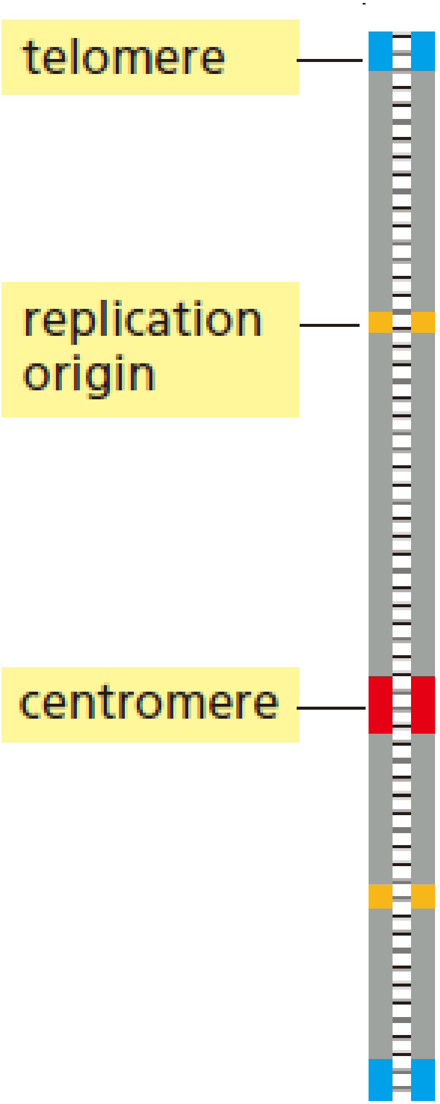
Human genome	
DNA length	3.2×10^9 nucleotide pairs*
Number of genes coding for proteins	Approximately 21,000
Largest gene coding for protein	2.4×10^6 nucleotide pairs
Mean size for protein-coding genes	27,000 nucleotide pairs
Smallest number of exons per gene	1
Largest number of exons per gene	178
Mean number of exons per gene	10.4
Largest exon size	17,106 nucleotide pairs
Mean exon size	145 nucleotide pairs
Number of noncoding RNA genes	Approximately 9000**
Number of pseudogenes***	More than 20,000
Percentage of DNA sequence in exons (protein-coding sequences)	1.5%
Percentage of DNA in other highly conserved sequences****	3.5%
Percentage of DNA in high-copy-number repetitive elements	Approximately 50%

Structural Elements to Form a Functional Chromosome

—— A Centromere, Two Telomeres and Replication Origins



Part I

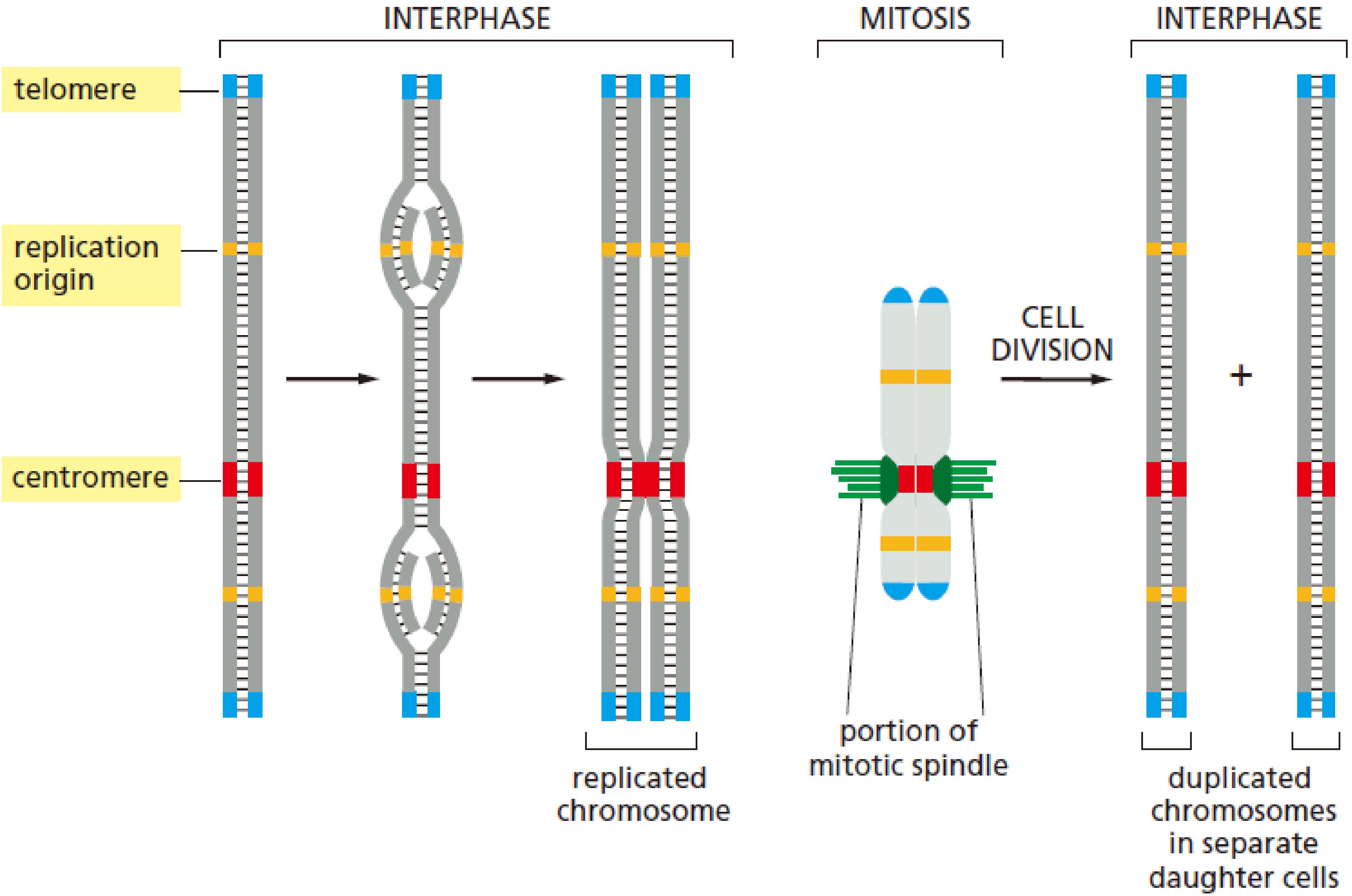


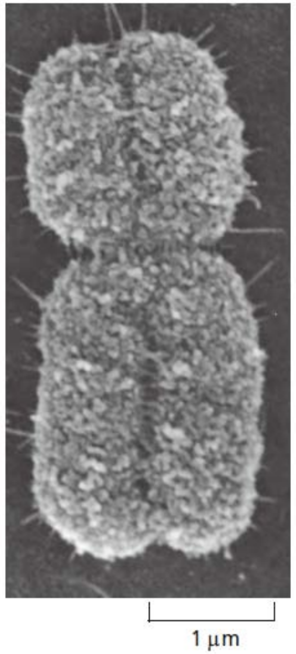
Telomeres: Sequences at the ends of eukaryotic chromosomes that help stabilize the chromosome.

Centromere: A sequence of DNA that functions during cell division as an attachment point of proteins that links chromosomes to mitotic spindle.

Part I

Three DNA Sequences Required to Produce a Eukaryotic Chromosome





Part I

Genes and Chromosomes

What are genes and chromosomes?

What do you know about introns and exons?

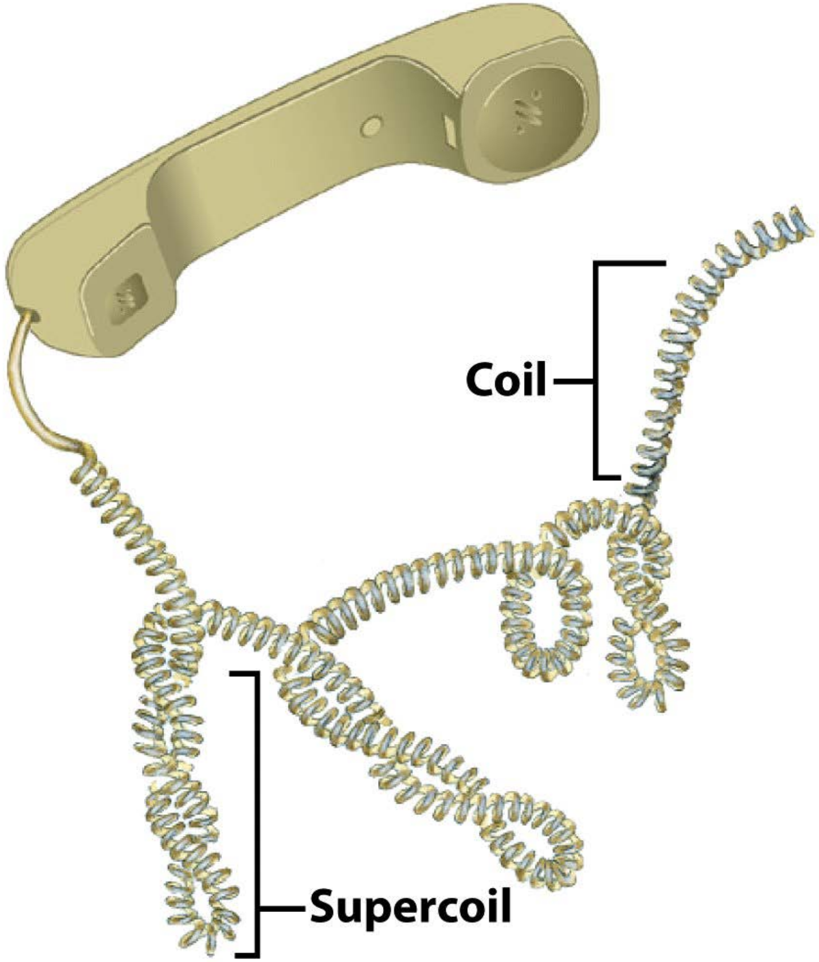
The human genome, surprising compositions.

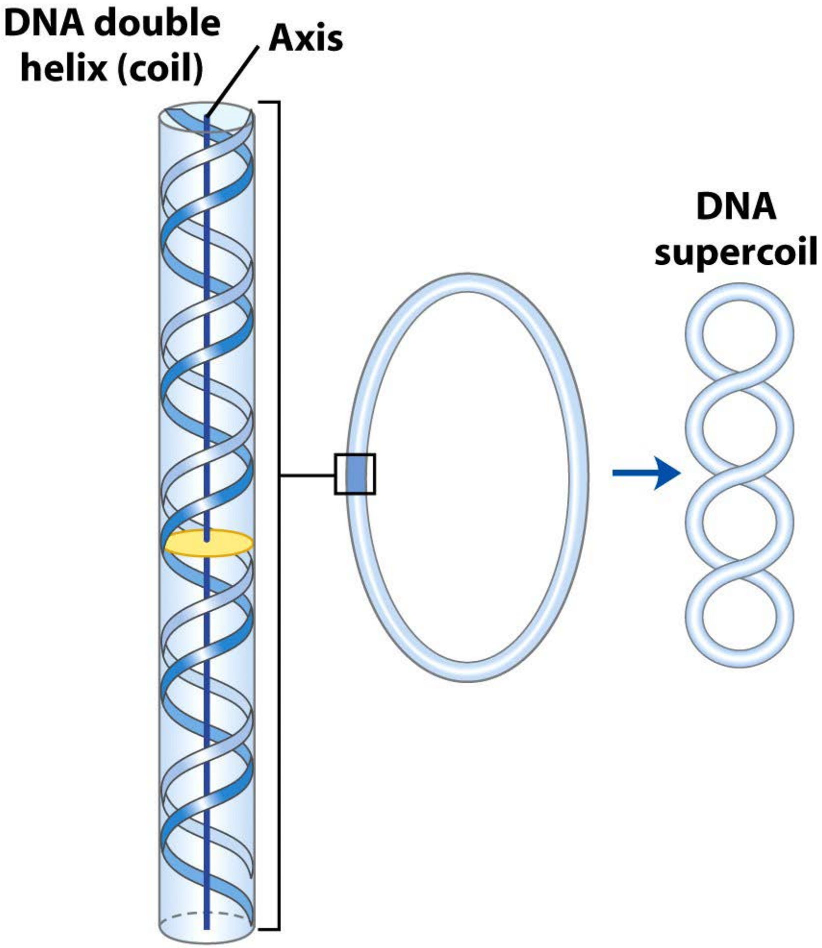
Features of a chromosome.

How do the Long DNA Molecules Fit into their cells?

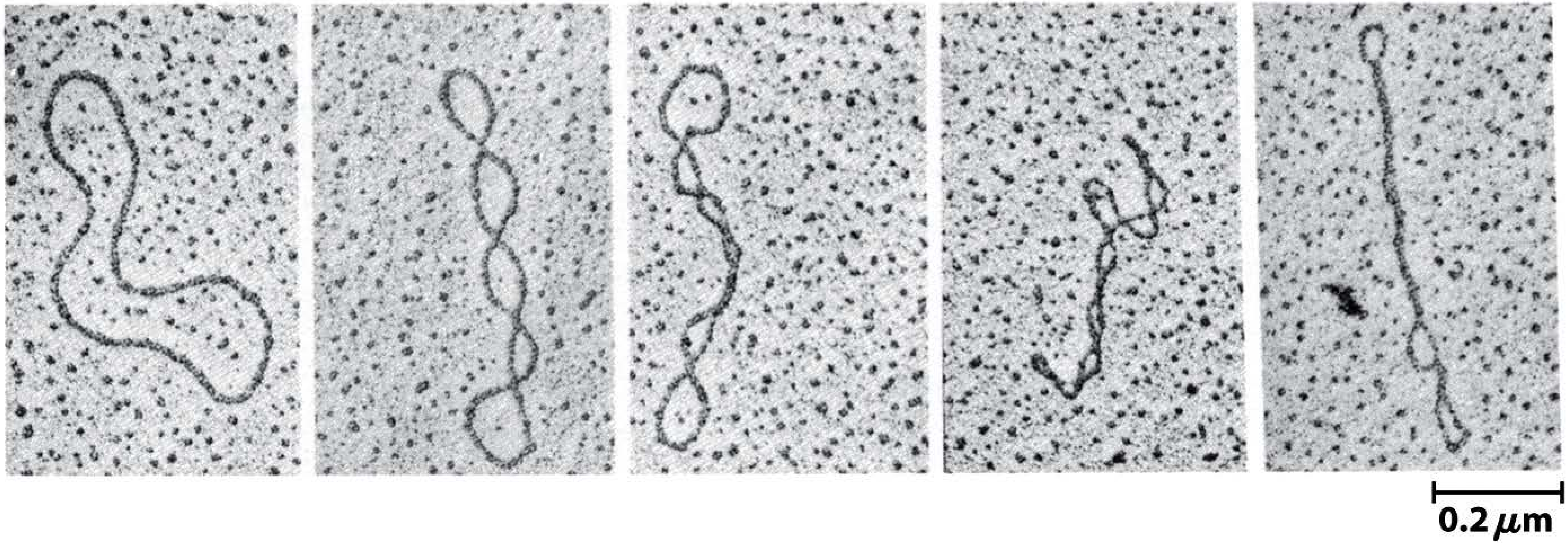
Part II. DNA supercoiling

Supercoiling of DNA





Relaxed and Supercoiled Plasmid DNAs



Relaxed state: no bending of the DNA axis upon itself

Topology

Underwinding of DNA

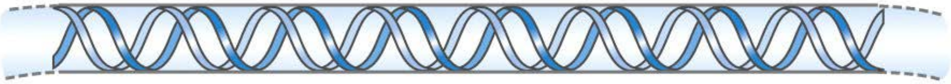
Supercoiling as a result of some form of structural strain, often times underwinding.

Underwinding: DNA has fewer helical turns than B-form structure.

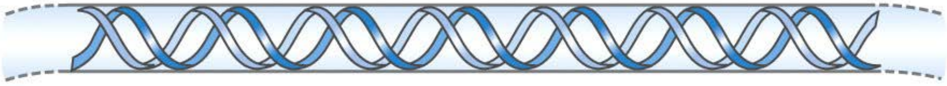
Underwinding of DNA makes separation of DNA strand easier.

Cells maintain DNA in an underwound state to facilitate coiling.

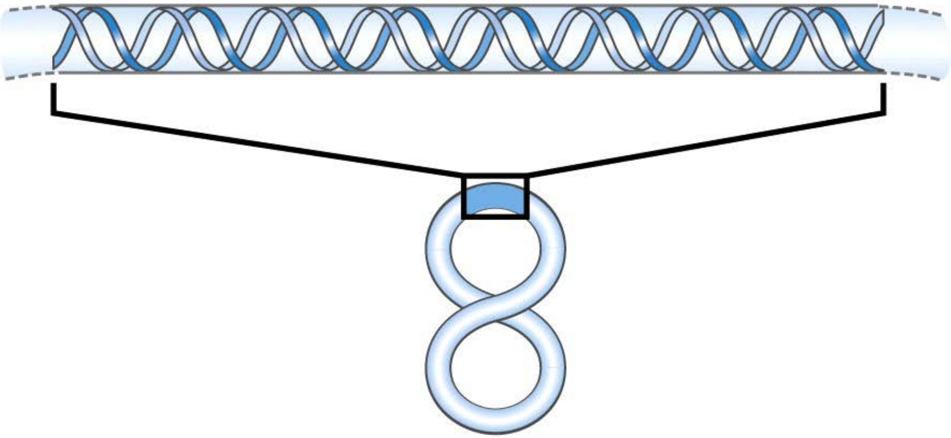
Most Cellular DNA Is Underwound



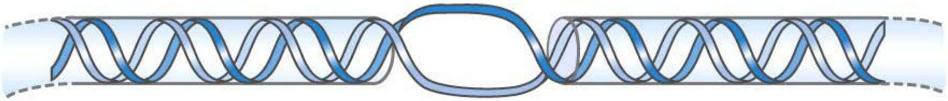
(a) Relaxed (8 turns)



(b) Strained (7 turns)

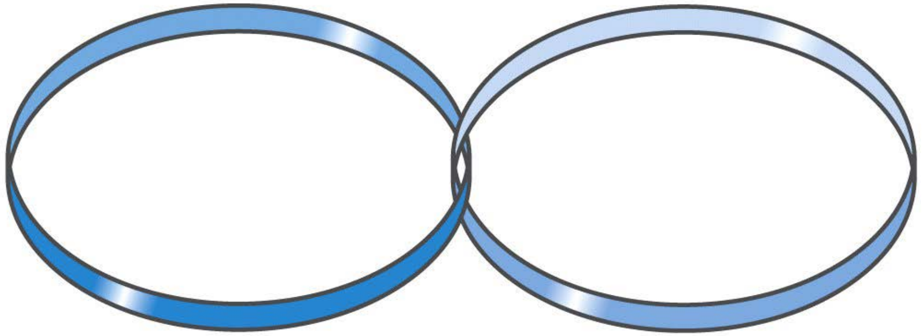


(c) Supercoil

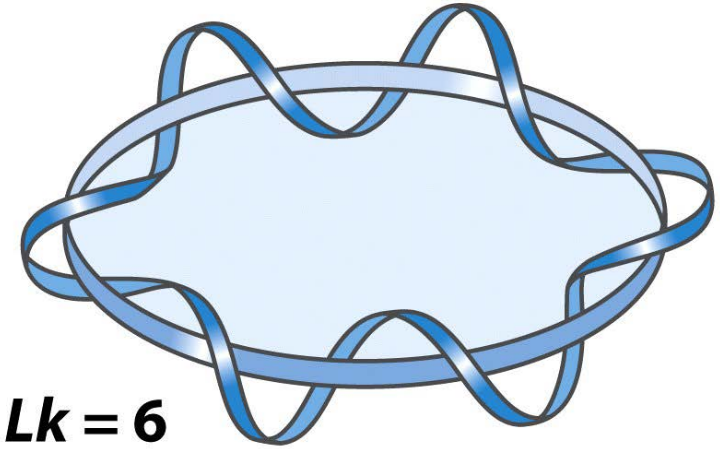


(d) Strand separation

DNA Underwinding Is Defined by Topological Linking Number



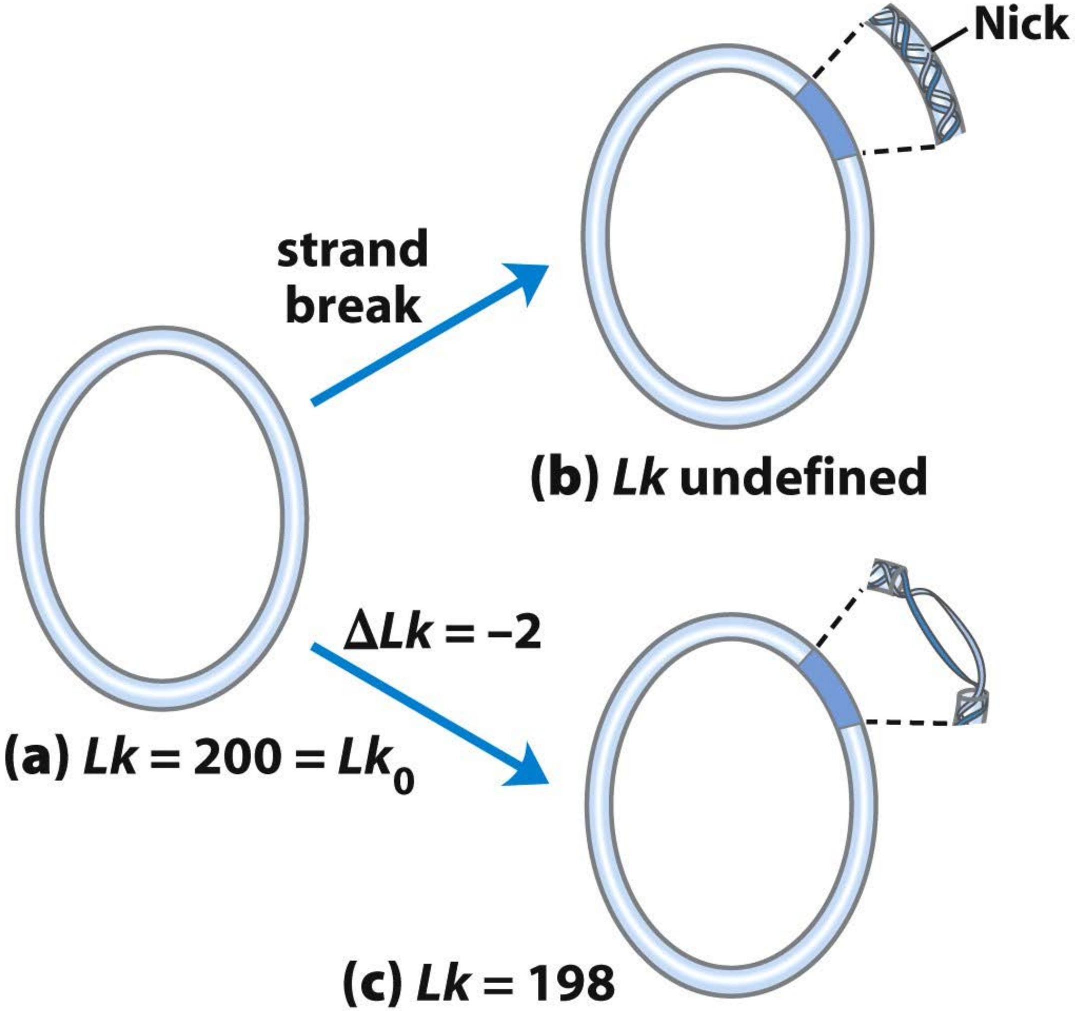
(a) *Lk* = 1



(b) *Lk* = 6

Linking number does not vary as long as both strands remain intact.

Linking Number Applied to Close-Circled Molecules



Specific Linking Difference (σ) or Superhelical Density

$$Lk_0 = \frac{\text{Number of Base Pairs}}{\text{10.5 (base pairs per turn)}}$$

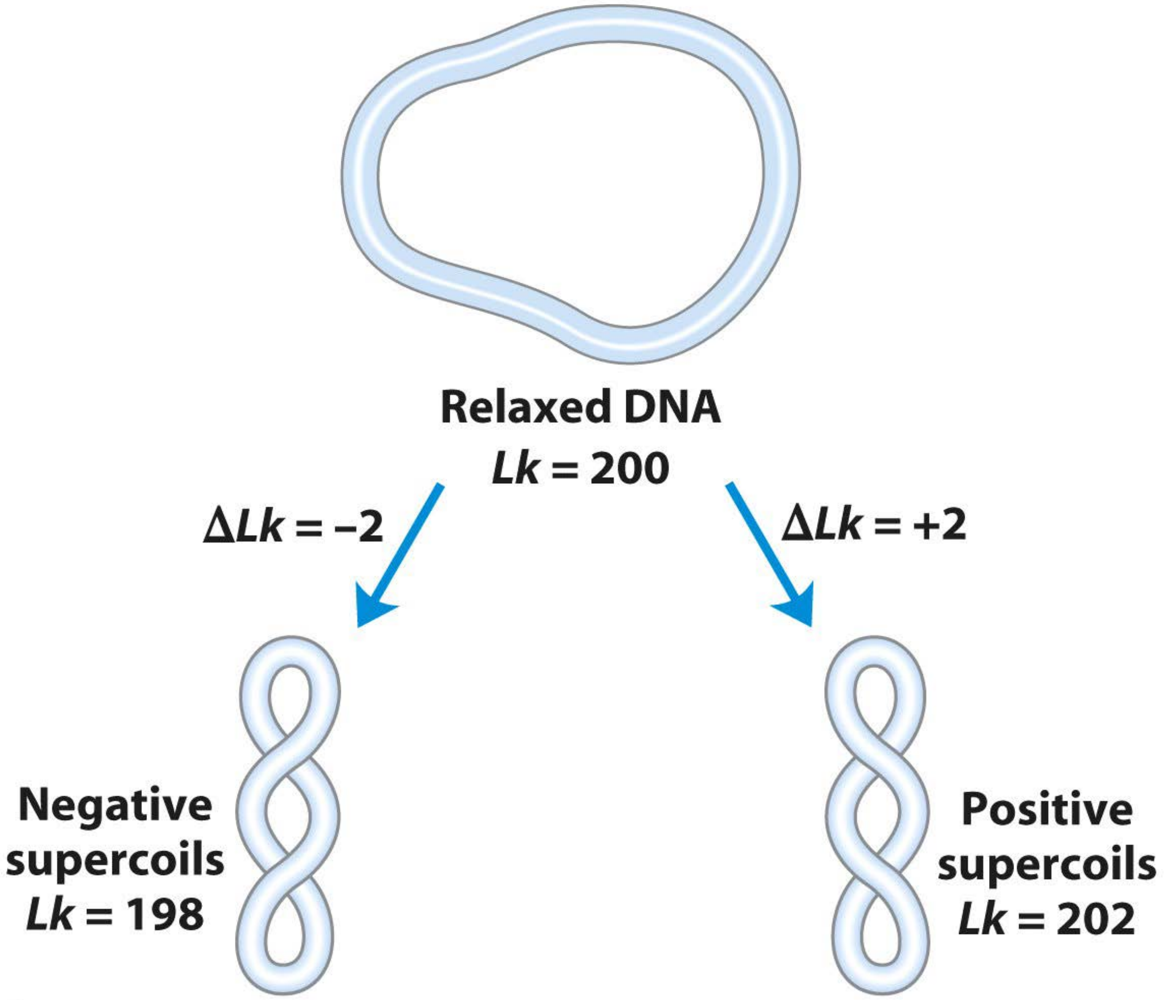
Lk : Actual linking number

$$\Delta Lk = Lk\text{-}Lk_0$$

$$\sigma = \frac{\Delta Lk}{Lk_0}$$

The degree of DNA underwinding in cellular DNA generally falls in the range of 5% to 7%.

Negative and Positive Supercoils



Right-handed

Left-handed

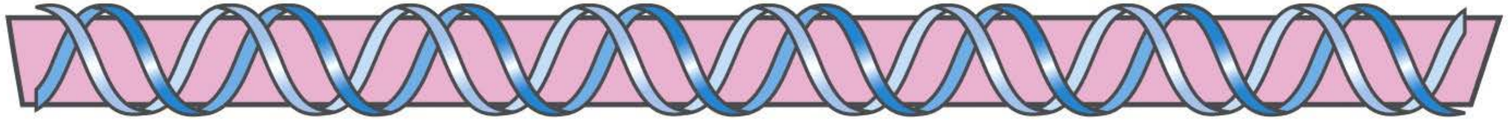
Topoisomers

Supercoil

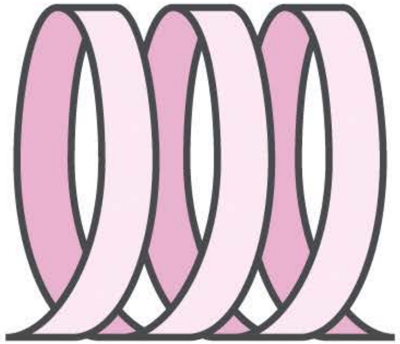
- Linking number can be broken down into two structural components, twist (Tw) and writhe (Wr).
- Writhe (Wr): it is a measure of the coiling of the axis of the double helix.
- In contrast, twist measures the helical winding of the DNA strands around each other.

$$Lk = Tw + Wr$$

Twist and Writhe



Straight ribbon (relaxed DNA)

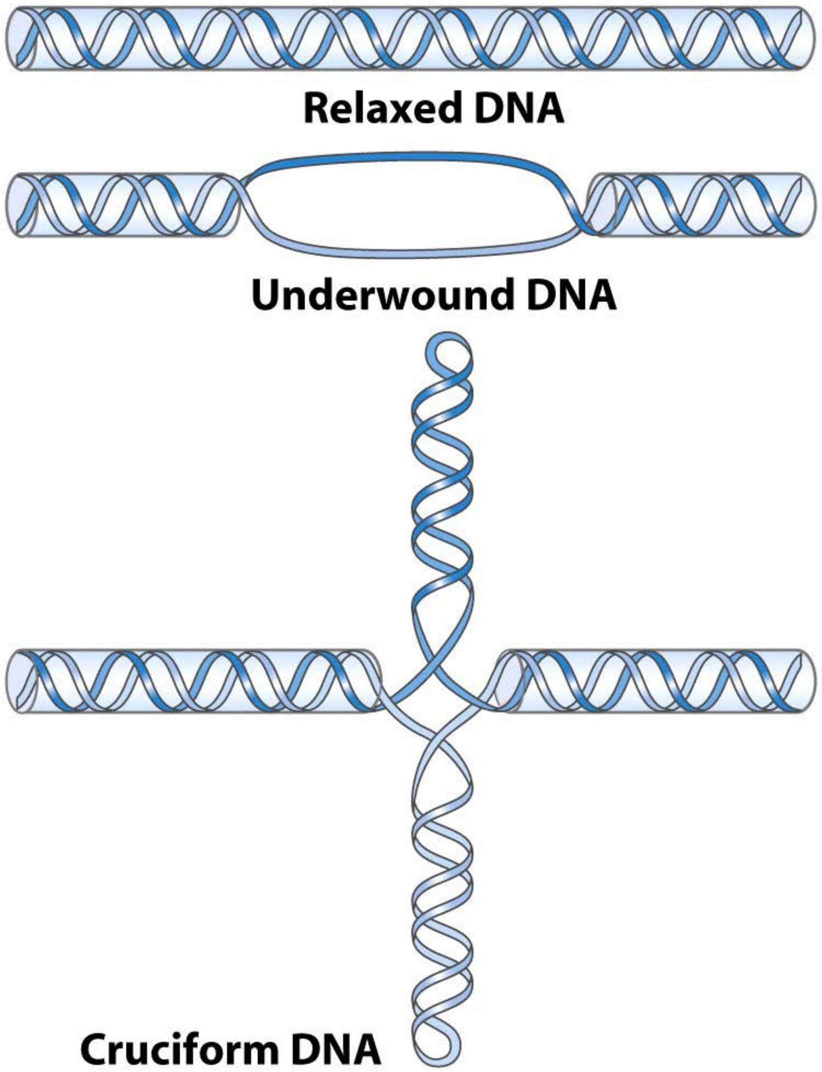


Large writhe, small change in twist

Zero writhe, large change in twist

Part II

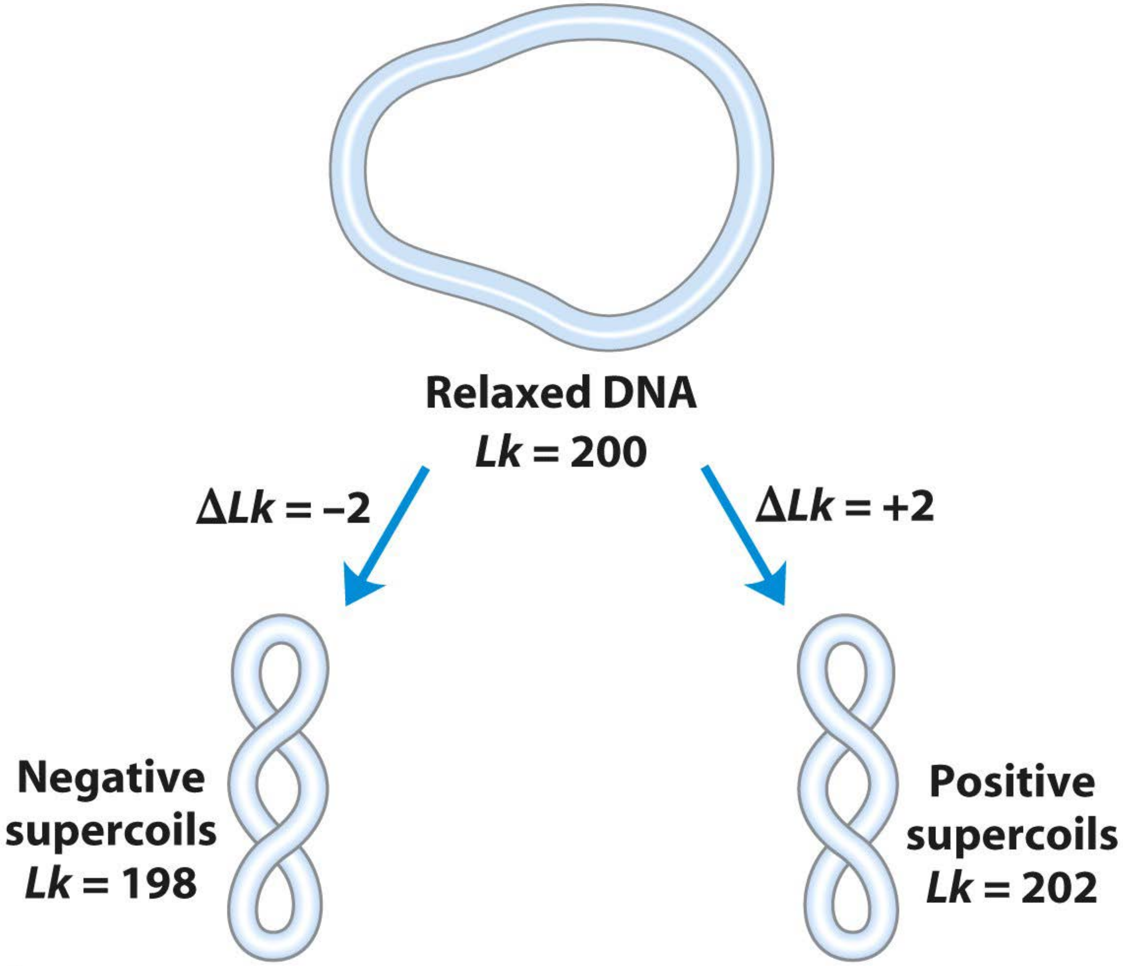
Promotion of Cruciform Structures by DNA Underwinding



Part II

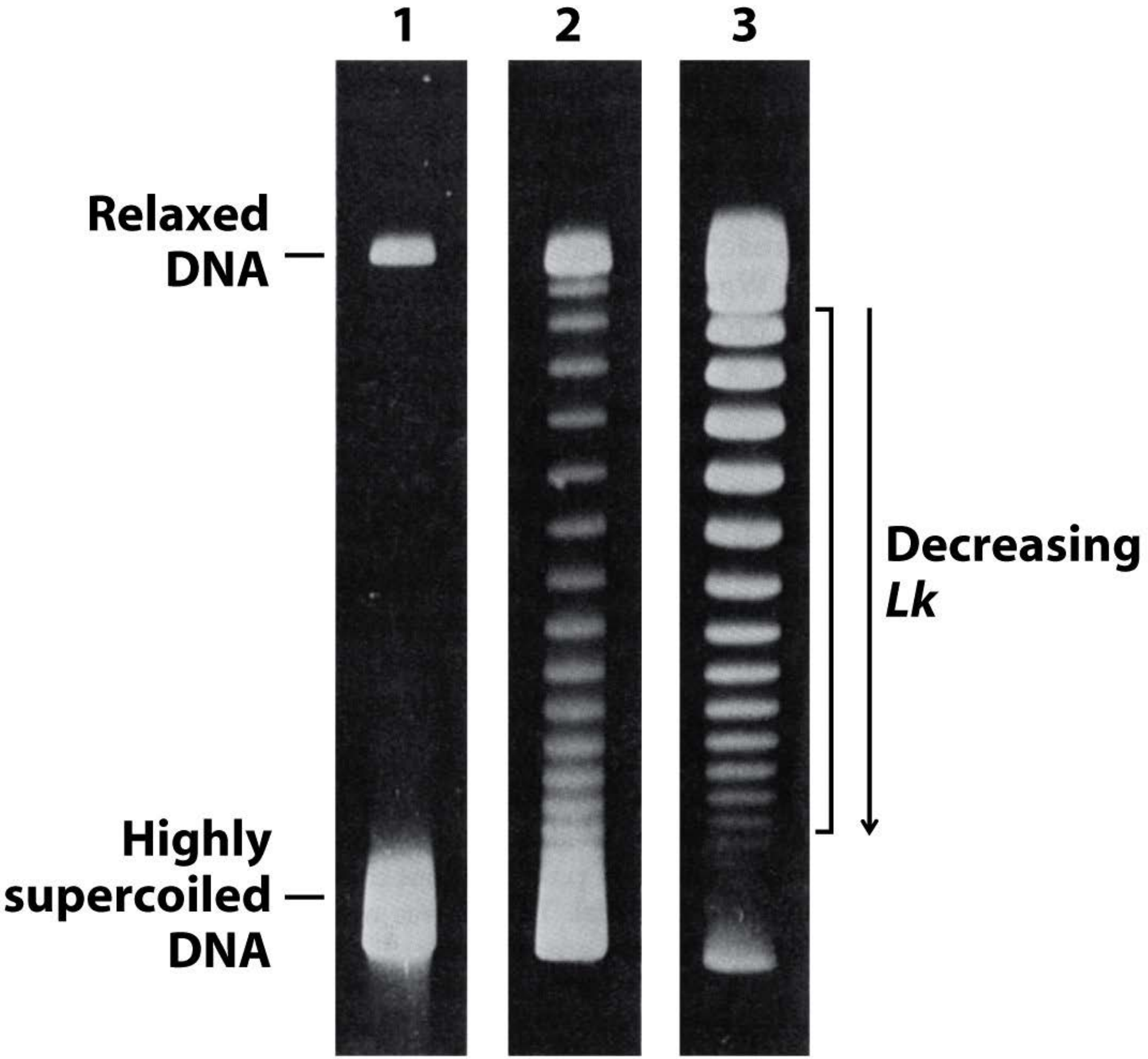
Topoisomerases Catalyze Changes in the Linking Number of DNA

- Enzymes that increase or decrease the extent of DNA underwinding
- E. Coli has at least 4 individual topoisomerases
- Type I and Type II topoisomerase



Part II

DNA Motility and Visualization of Topoisomers



Type I Topoisomerase

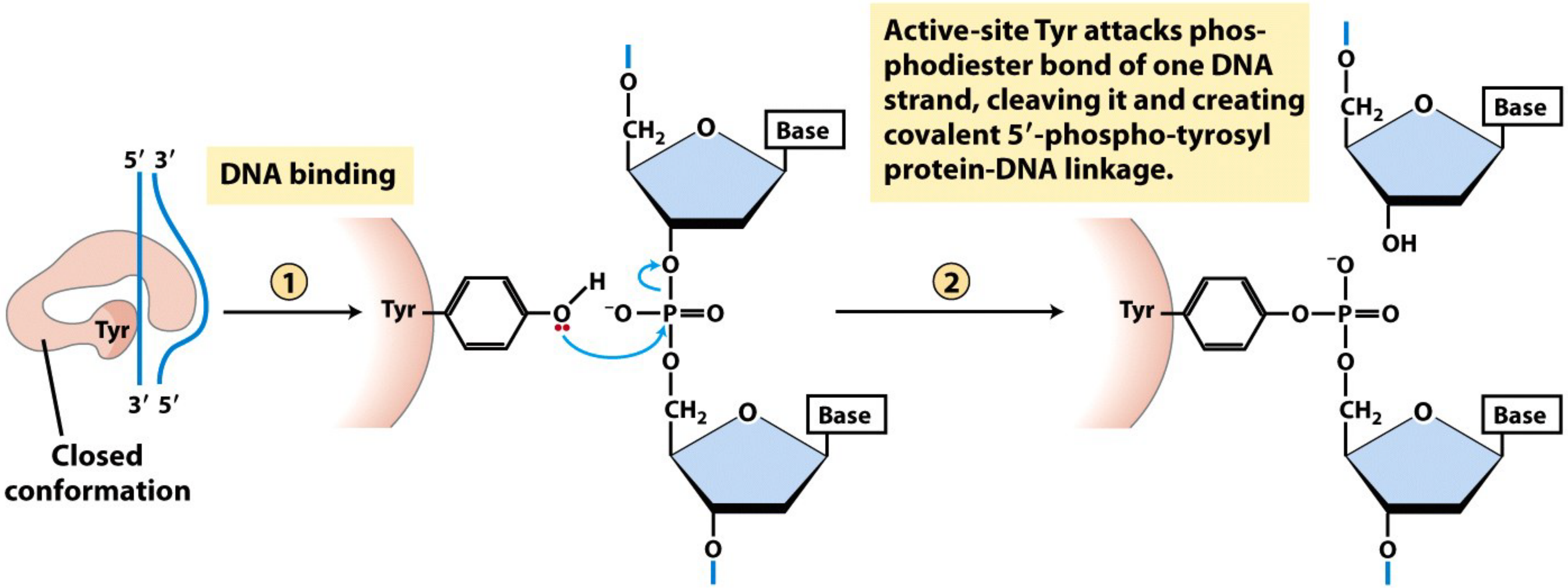
Transiently breaks one of the two DNA strands,
Passes unbroken strand through the break,
and rejoin the broken ends.

Change Lk in increments of 1.

Type I (topoisomerases I and III) generally relax DNA by removing negative supercoils (increasing Lk).

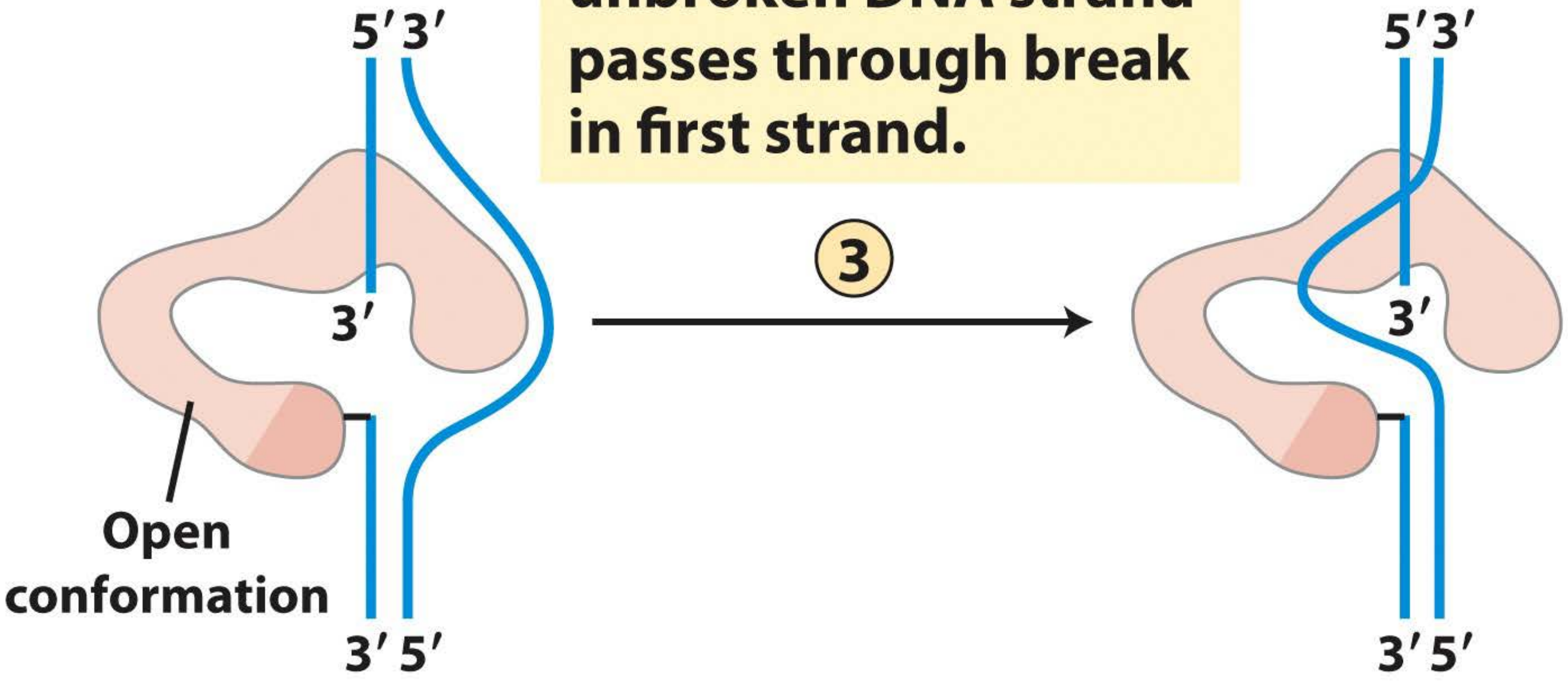
Part II

Action of Type I Topoisomerase



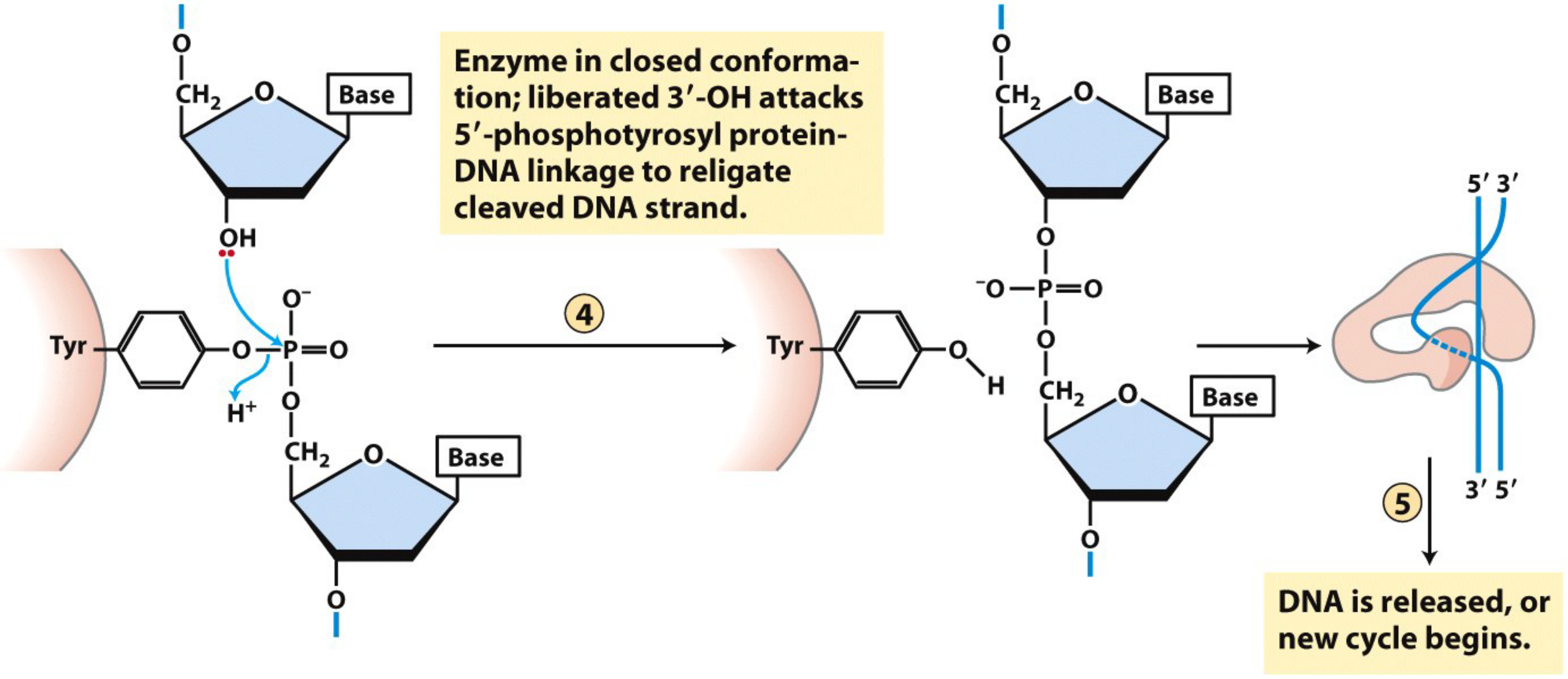
Active-site Tyr attacks phosphodiester bond of one DNA strand, cleaving it and creating covalent 5′-phospho-tyrosyl protein-DNA linkage.

Action of Type I Topoisomerase



Enzyme changes to open conformation; unbroken DNA strand passes through break in first strand.

Action of Type I Topoisomerase



Enzyme in closed conformation; liberated 3′-OH attacks 5′-phosphotyrosyl protein-DNA linkage to religate cleaved DNA strand.

DNA is released, or new cycle begins.

Part II

Type II Topoisomerase

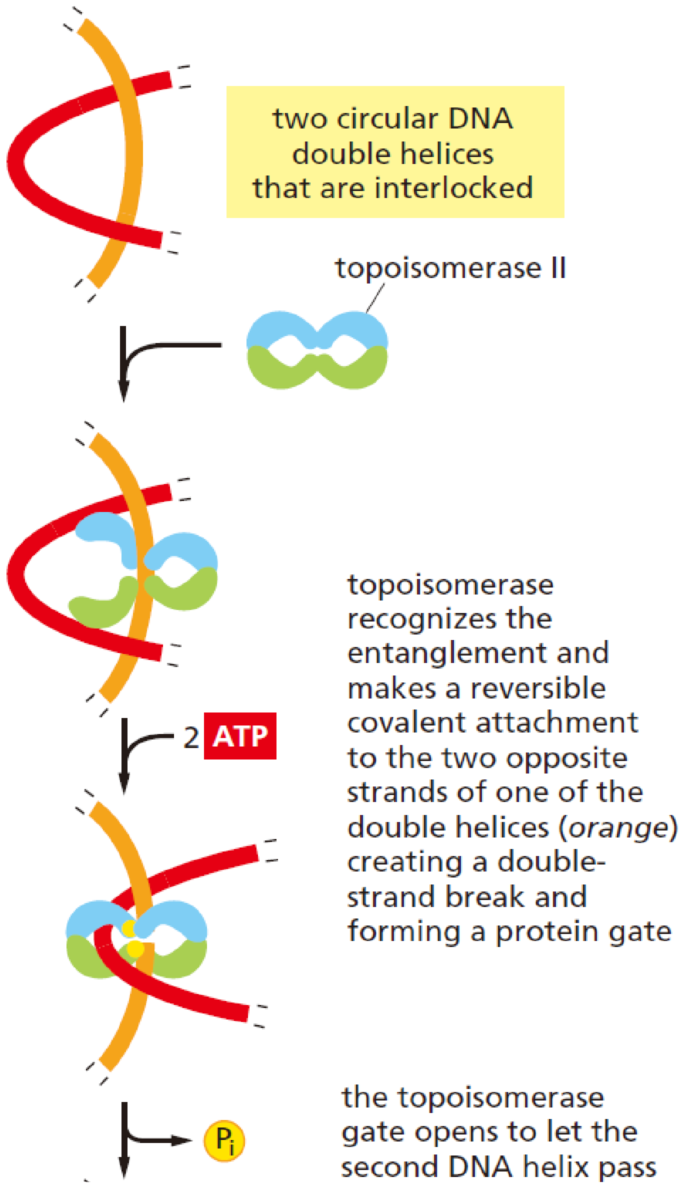
Transiently breaks both strands,

Change Lk in increments of 2.

A bacterial type II enzyme (topoisomerase II or DNA Gyrase) generally introduce negative supercoils (decrease Lk).

Part II

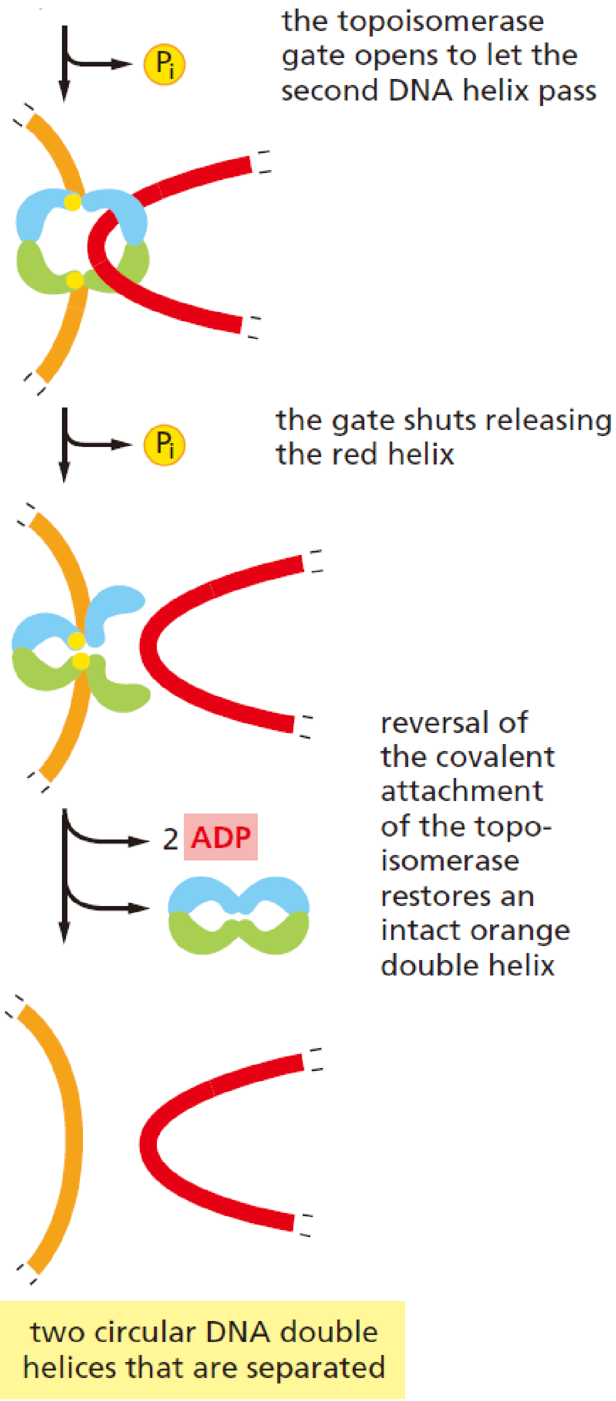
Action of Type II Topoisomerase



two circular DNA double helices that are interlocked

topoisomerase recognizes the entanglement and makes a reversible covalent attachment to the two opposite strands of one of the double helices (*orange*) creating a double-strand break and forming a protein gate

the topoisomerase gate opens to let the second DNA helix pass



the topoisomerase gate opens to let the second DNA helix pass

the gate shuts releasing the red helix

reversal of the covalent attachment of the topoisomerase restores an intact orange double helix

two circular DNA double helices that are separated

DNA Compaction Requires a Special Form of Supercoiling

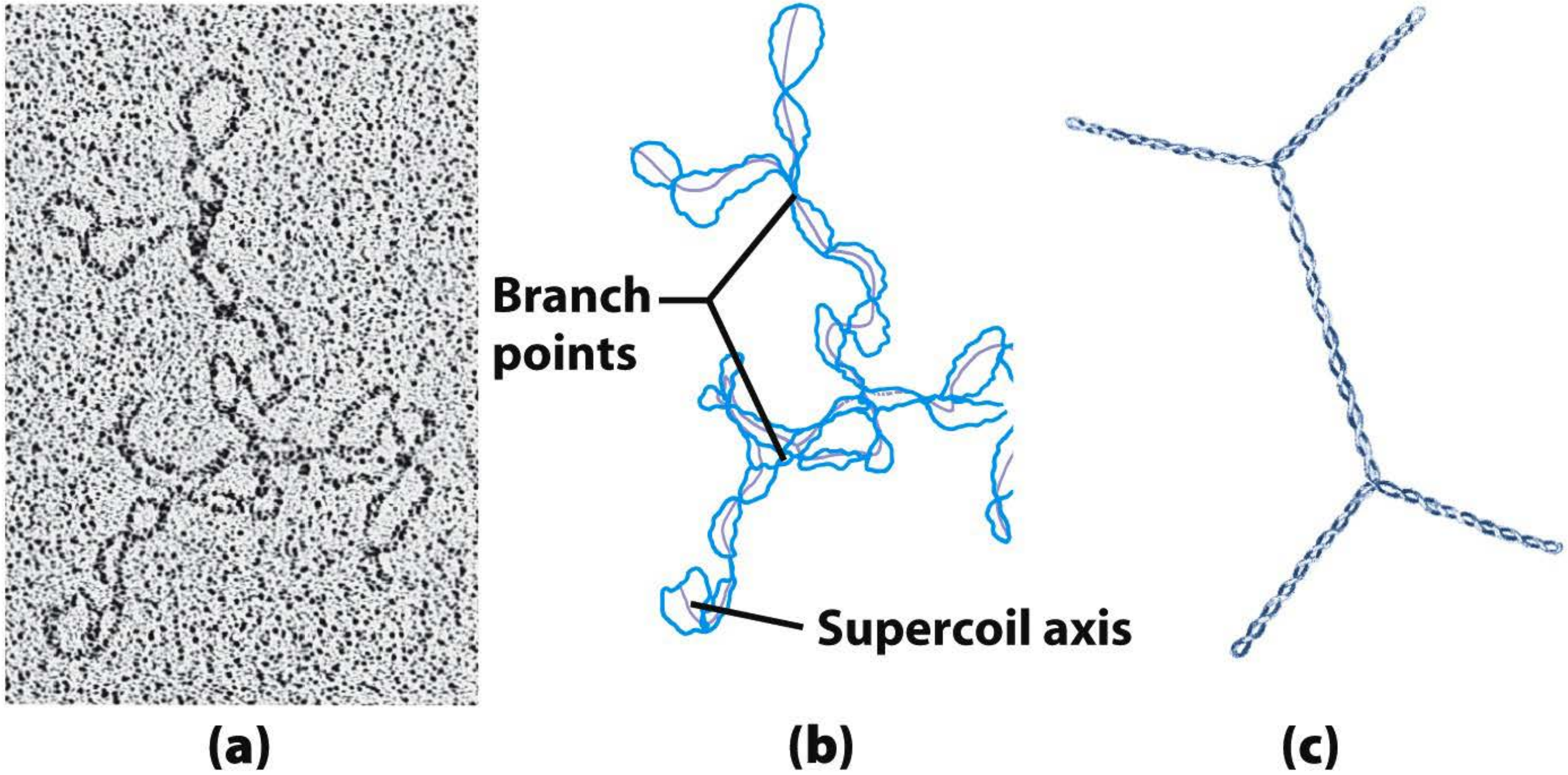
Supercoiled DNA are uniform in a number of respects:

The supercoils are right-handed in a negatively supercoiled DNA molecule.

They tend to be extended and narrow rather than compacted, often with multiple branches.

Part II

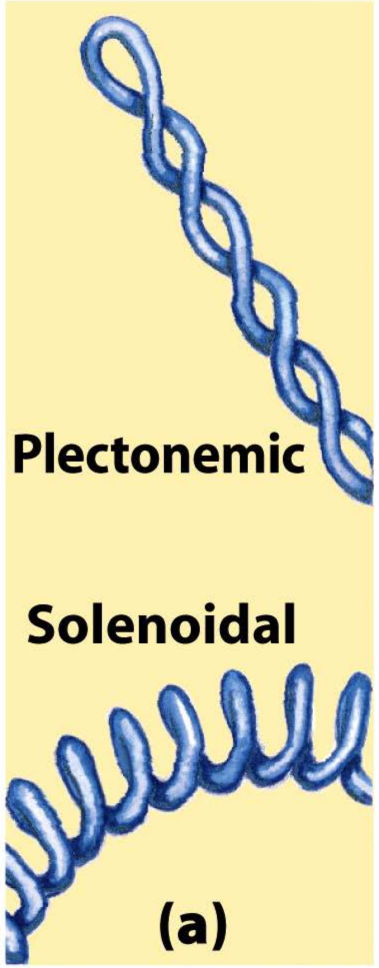
Plectonemic Supercoilng

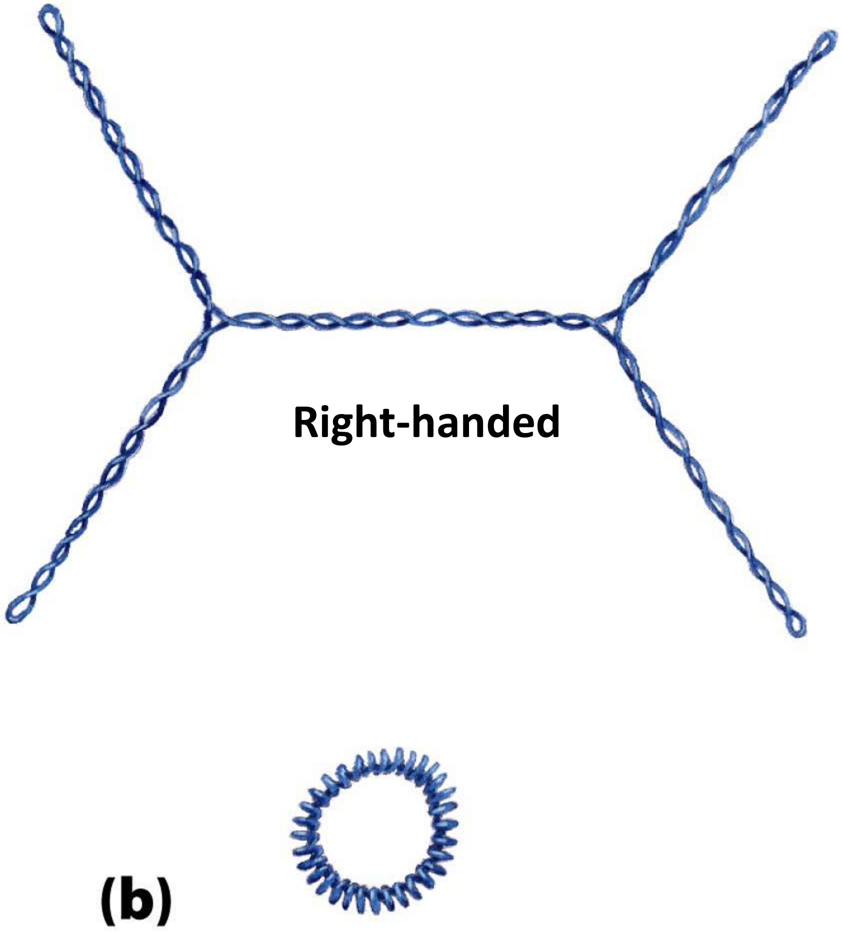


Plectonemic: from the Greek *plektos*, "Twisted" and *nema* "thread".

The length of the supercoiled axis, including branches, is about 40% of the length of the DNA.

Plectonemic and Solenoidal Supercoiling





Tight left-handed turns
Provides a greater degree of compaction

Both are negative supercoils.

Can be adopted by the same segment of underwound DNA.

Readily interconvertable.

Part II

DNA Supercoling

DNA underwinding

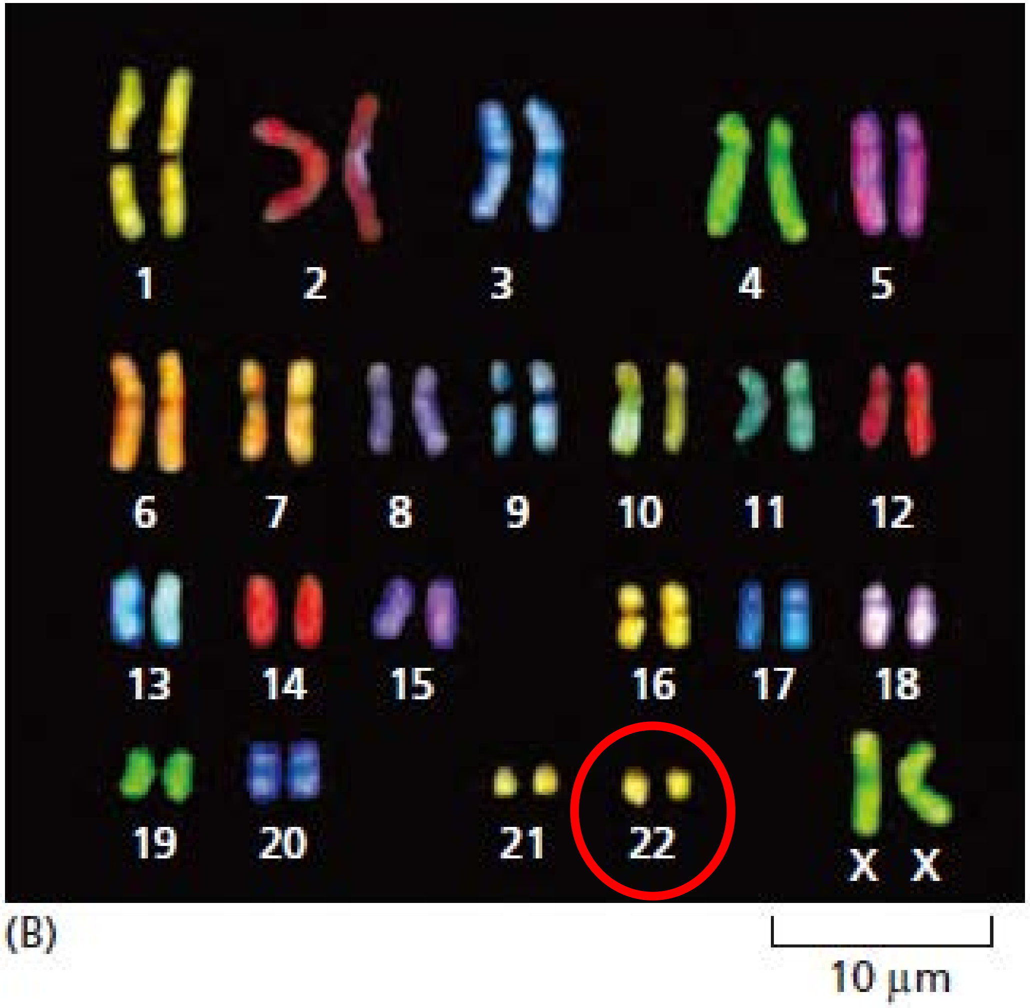
Linking number, Lk.

Specific linking difference, σ
(also called superhelical densitiy)

Type I and II Topoisomerases.

Part III. The Structures of Chromosomes

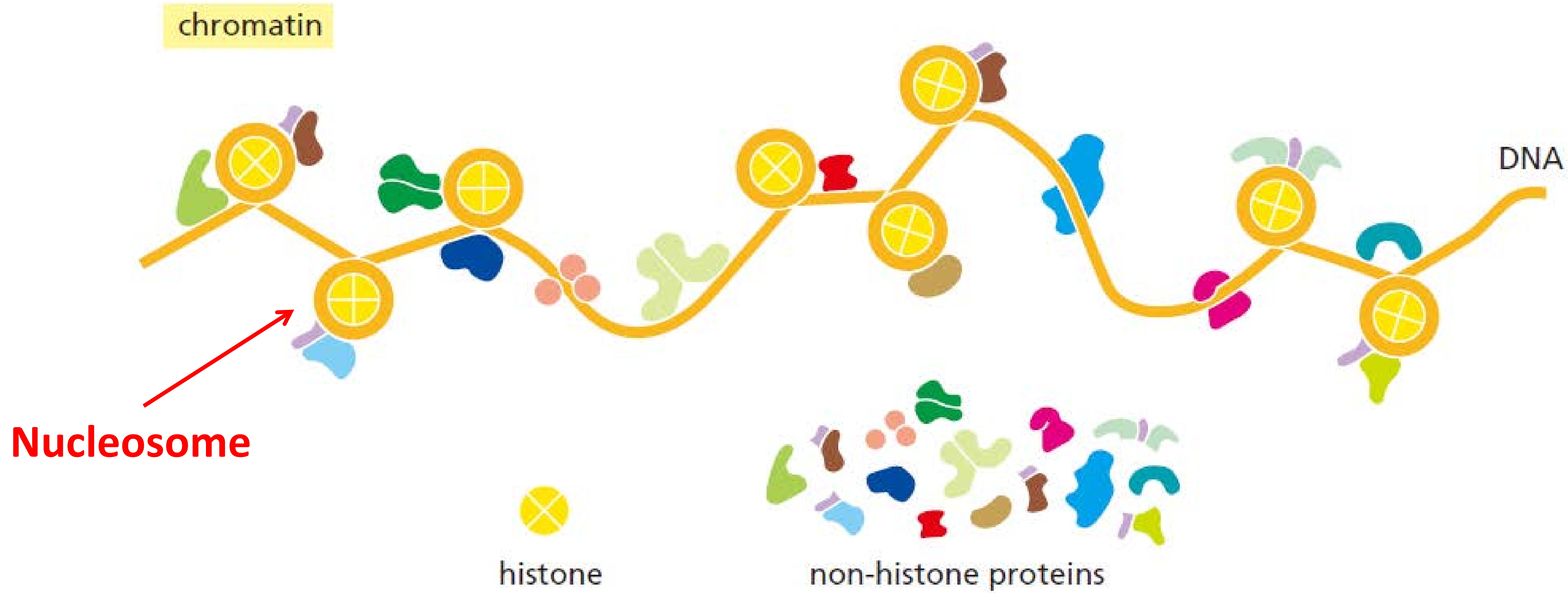
DNA Molecules Are Highly Condensed in Chromosomes



Chromosome 22

1.5 cm if stretched out
2 µm in mitosis

Chromatin Consists of DNA and Proteins



Nucleosome is a basic unit of chromosomal structure.

Histones Are Small, Basic Proteins

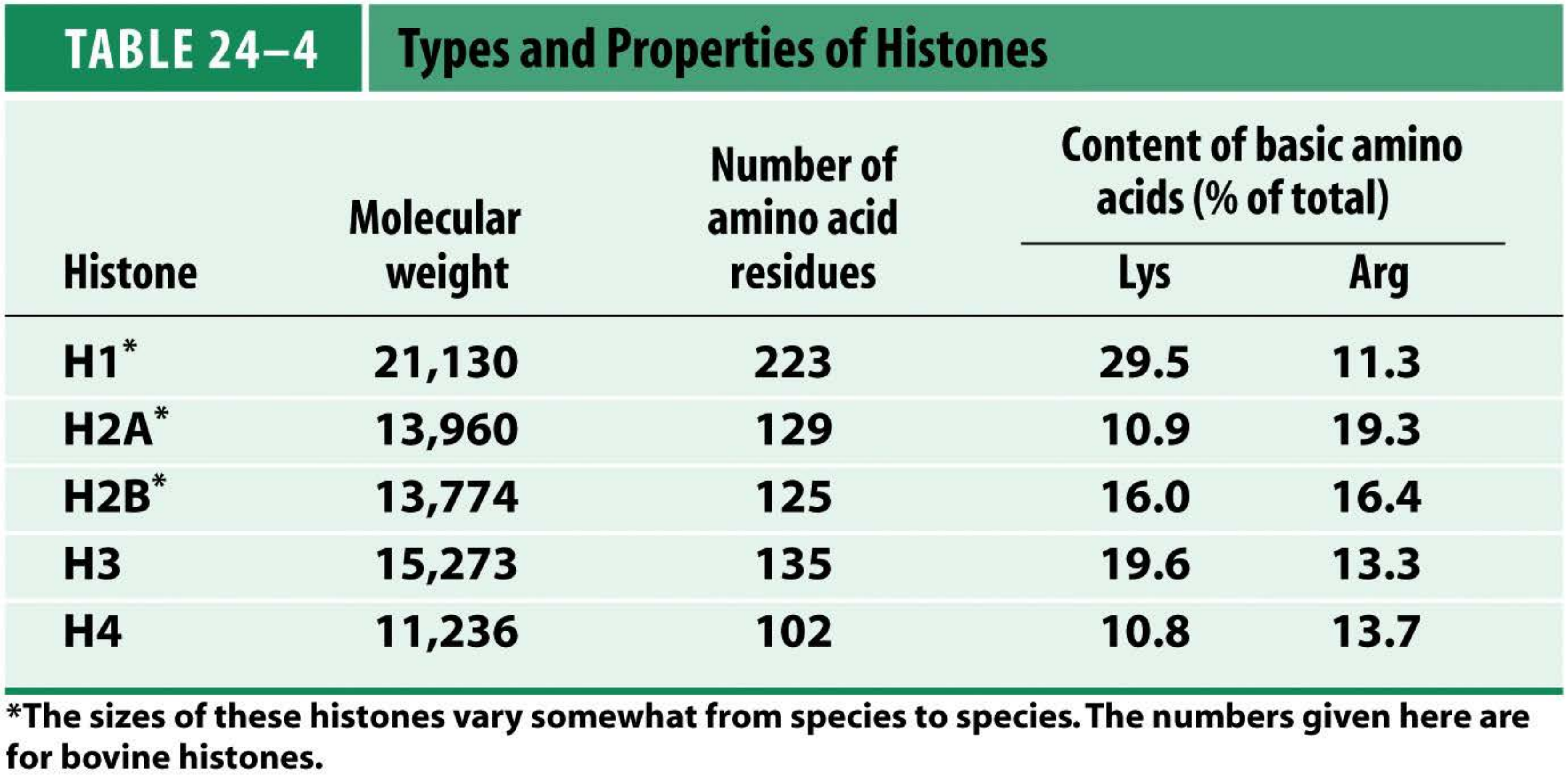


TABLE 24–4 Types and Properties of Histones

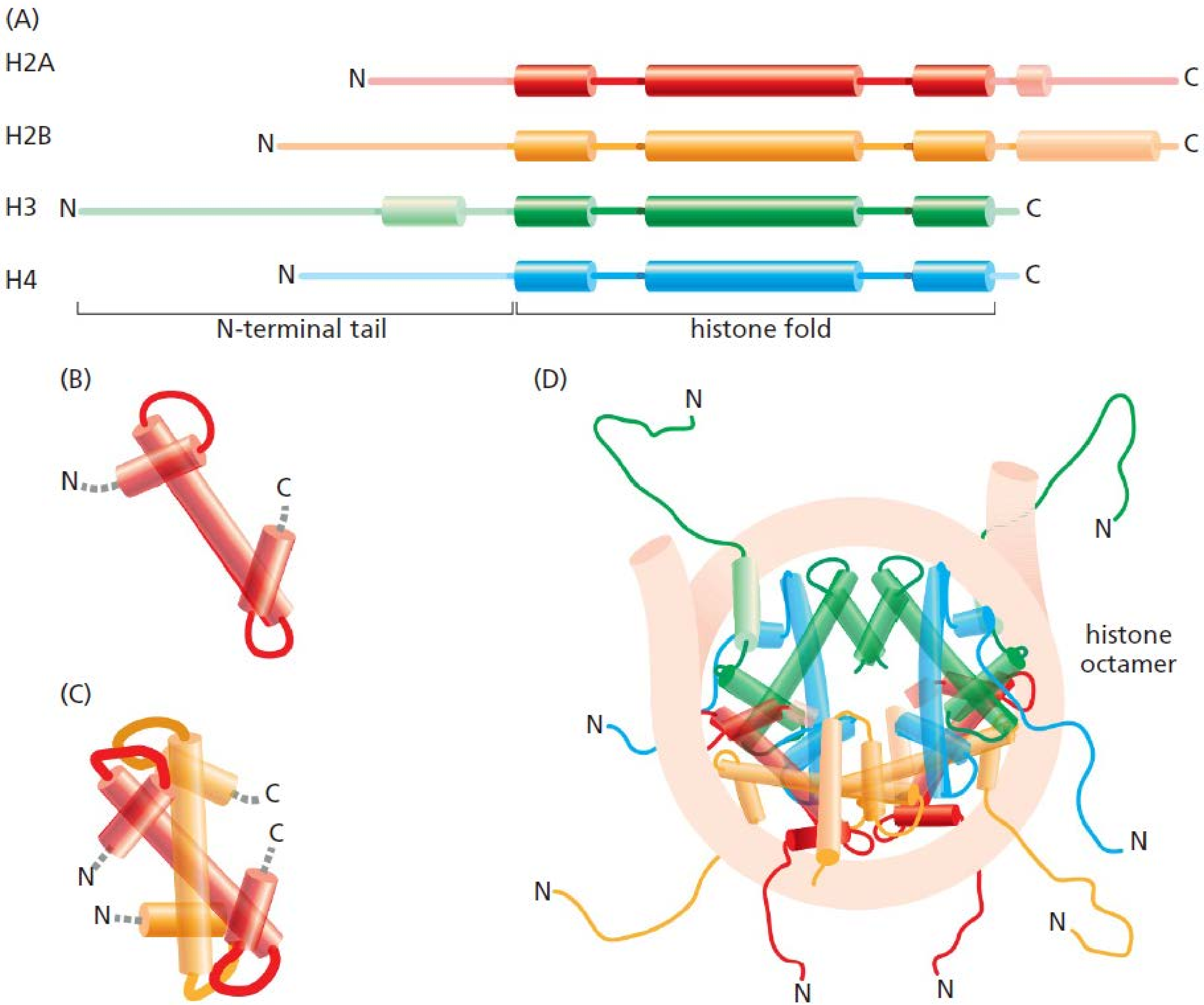
Histone	Molecular weight	Number of amino acid residues	Content of basic amino acids (% of total)	
			Lys	Arg
H1*	21,130	223	29.5	11.3
H2A*	13,960	129	10.9	19.3
H2B*	13,774	125	16.0	16.4
H3	15,273	135	19.6	13.3
H4	11,236	102	10.8	13.7

*The sizes of these histones vary somewhat from species to species. The numbers given here are for bovine histones.

Rich in lysines and arginines (1/4 of the amino acids).

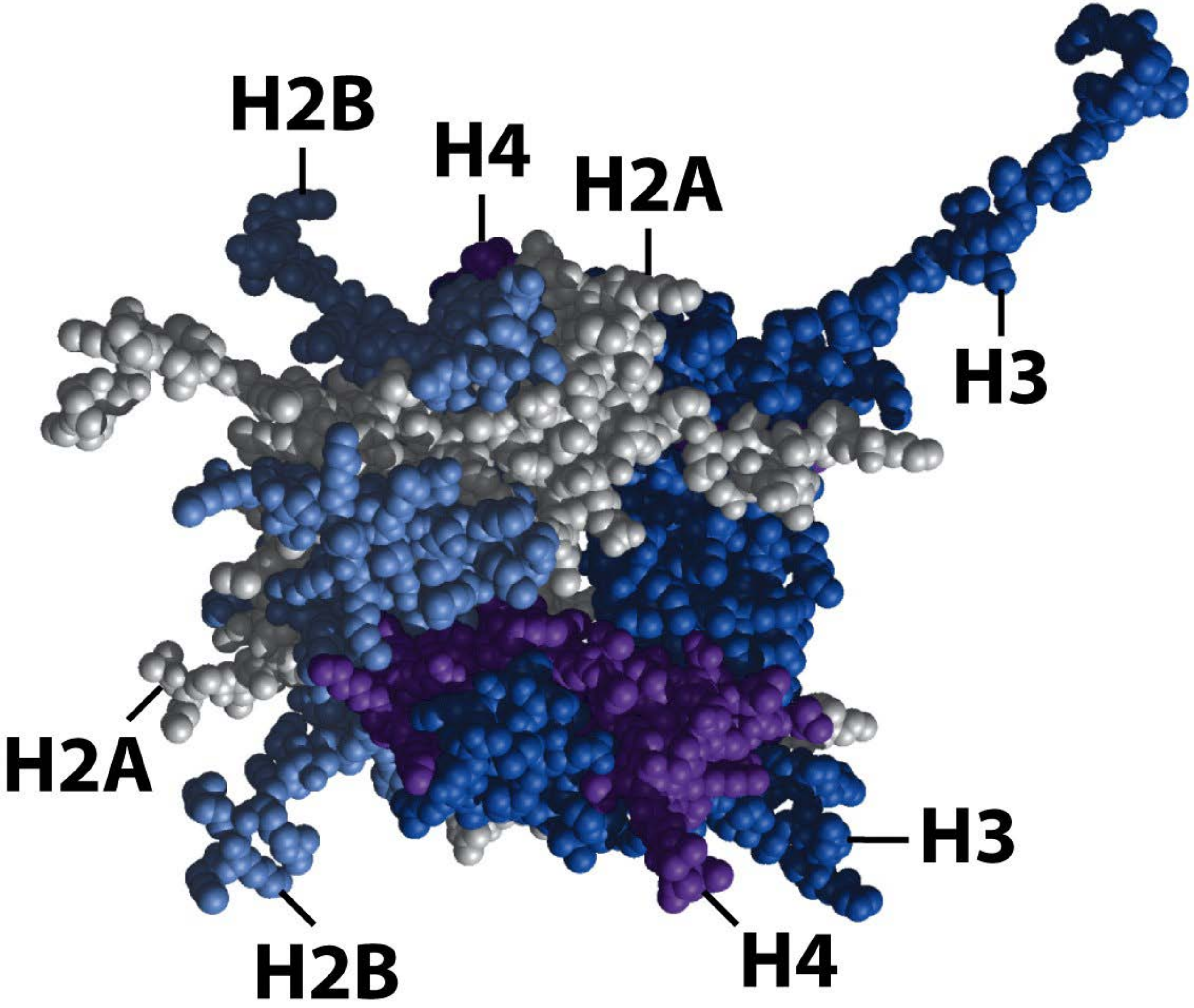
Part III

Overall Structural Organization of the Core Histones

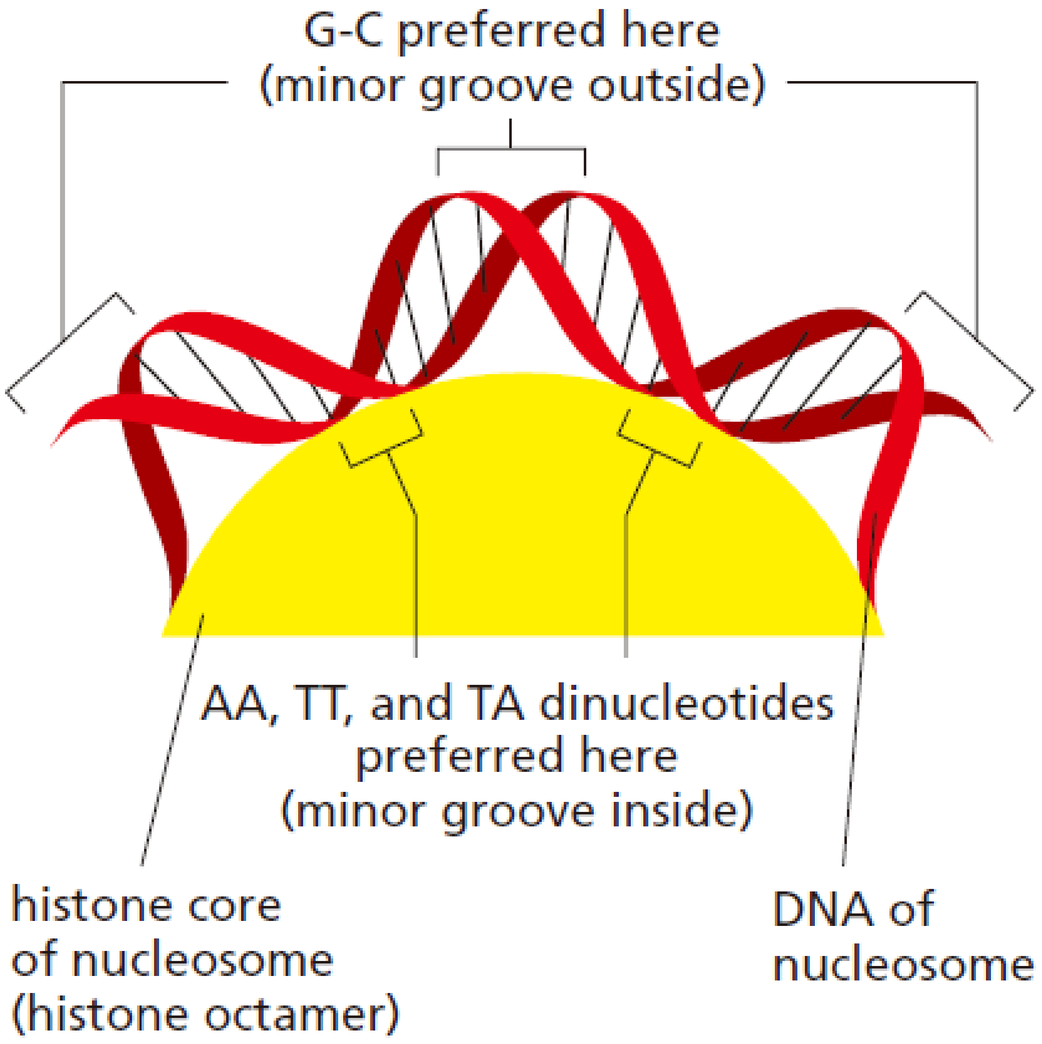


Core Histones of Nucleosome

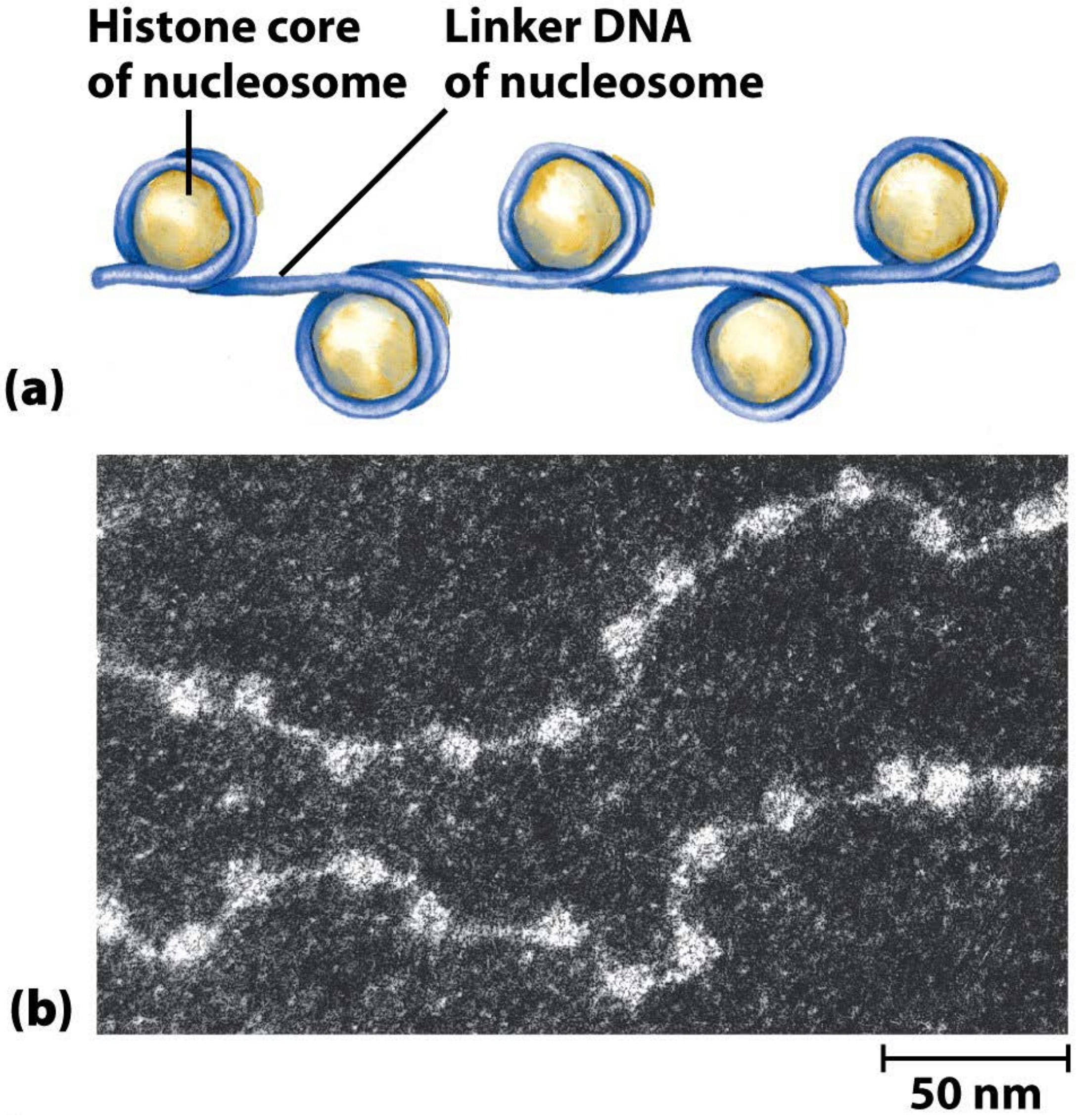
Octamer



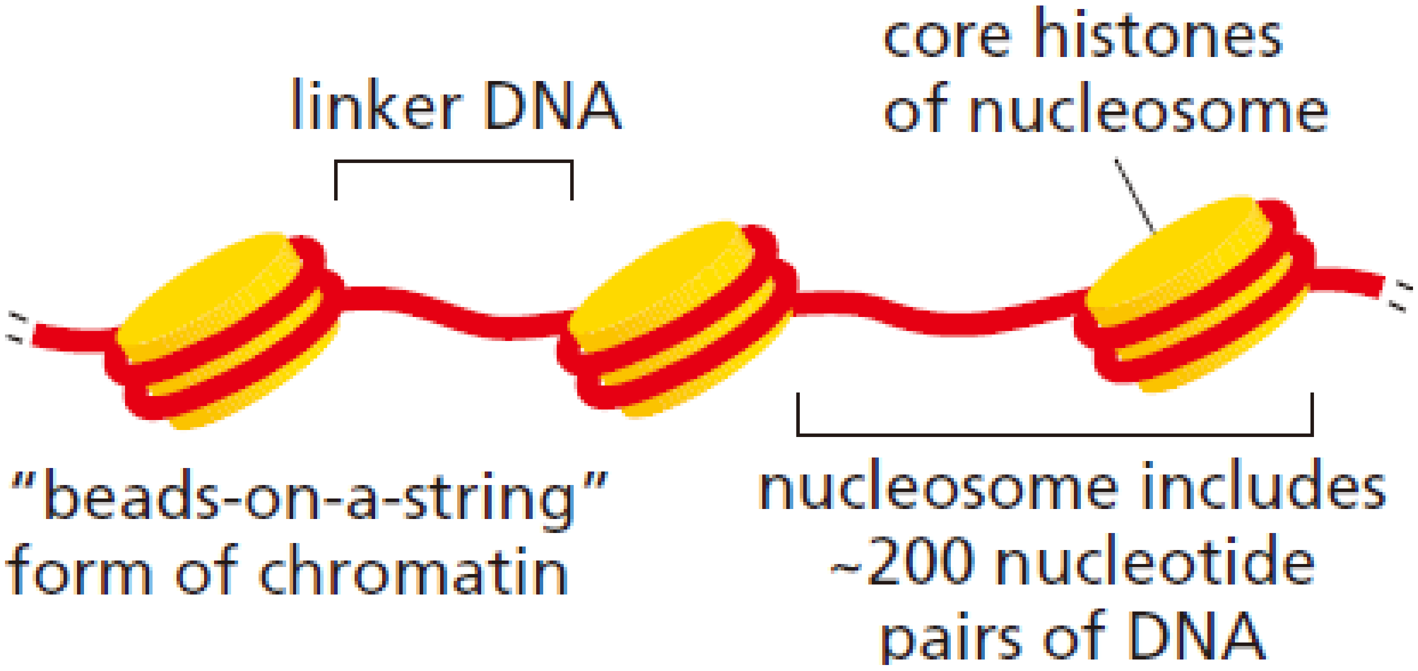
Bending of DNA in a Nucleosome



Nucleosome Is a Basic Unit of Chromosomes



A Close Look at a Nucleosome

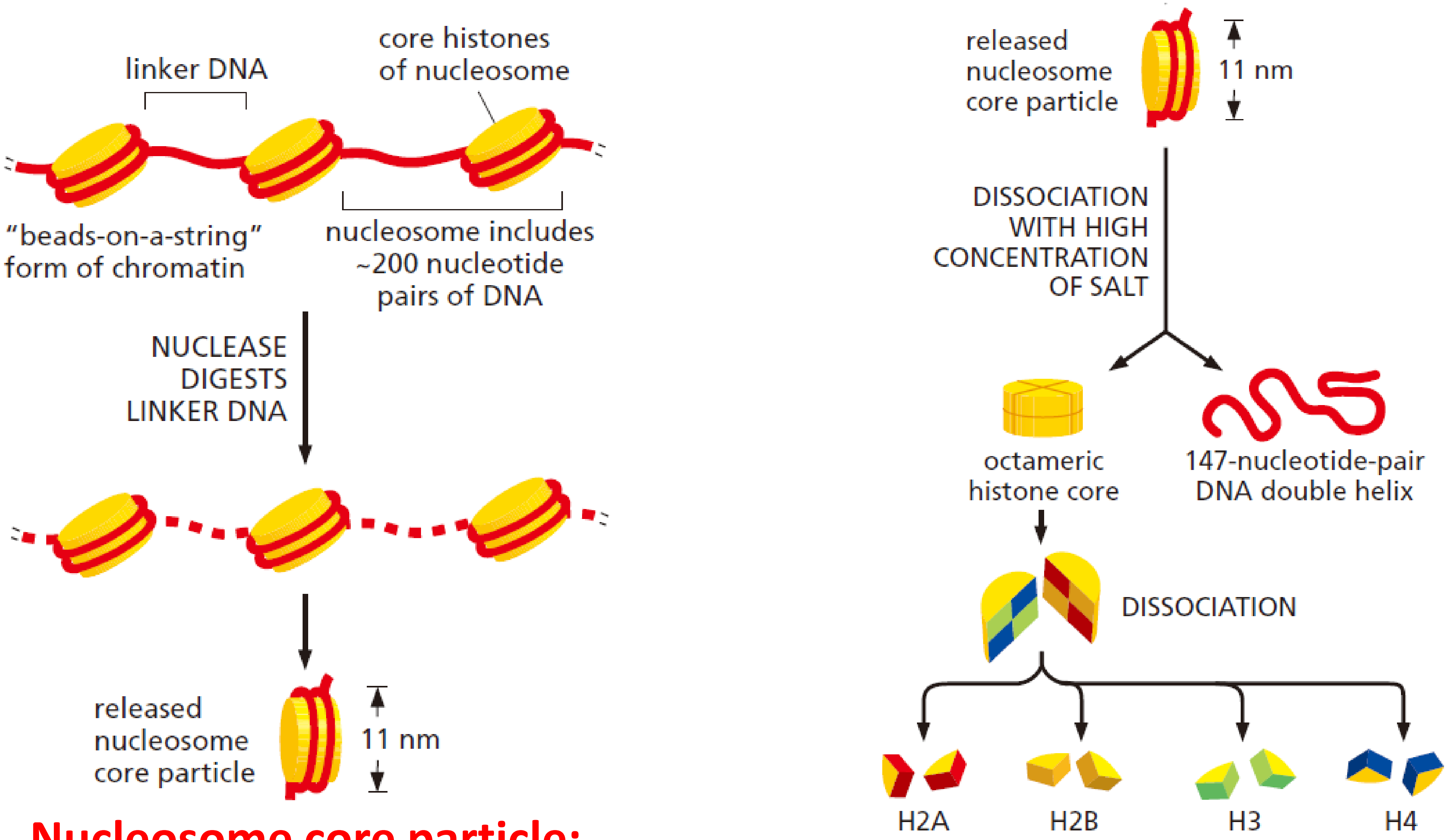


One nucleosome:

Histone Octamer

200 bp double-strand DNA

Nucleosome Core Particles



Nucleosome core particle:

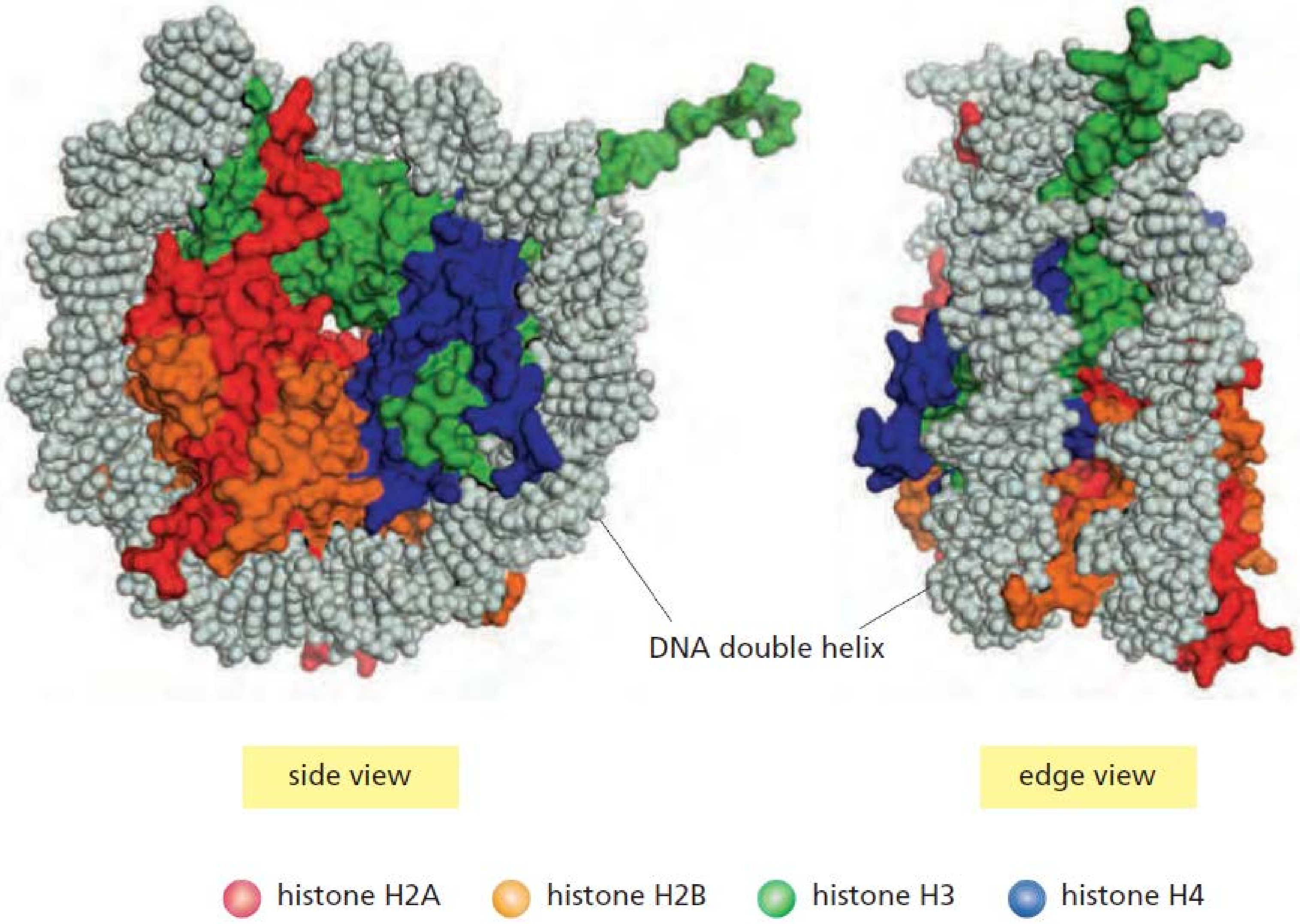
Histone Octamer (two molecules of H2A, H2B, H3 and H4)

147 bp double-strand DNA

DNA is tightly wrapped in a left-handed solenoidal supercoil of 1.7 turns

Part III

Structure of a Nucleosome Core Particle



Part III

Assembly of a Nucleosome

H3-H4 Dimer

H3-H4 Tetramer

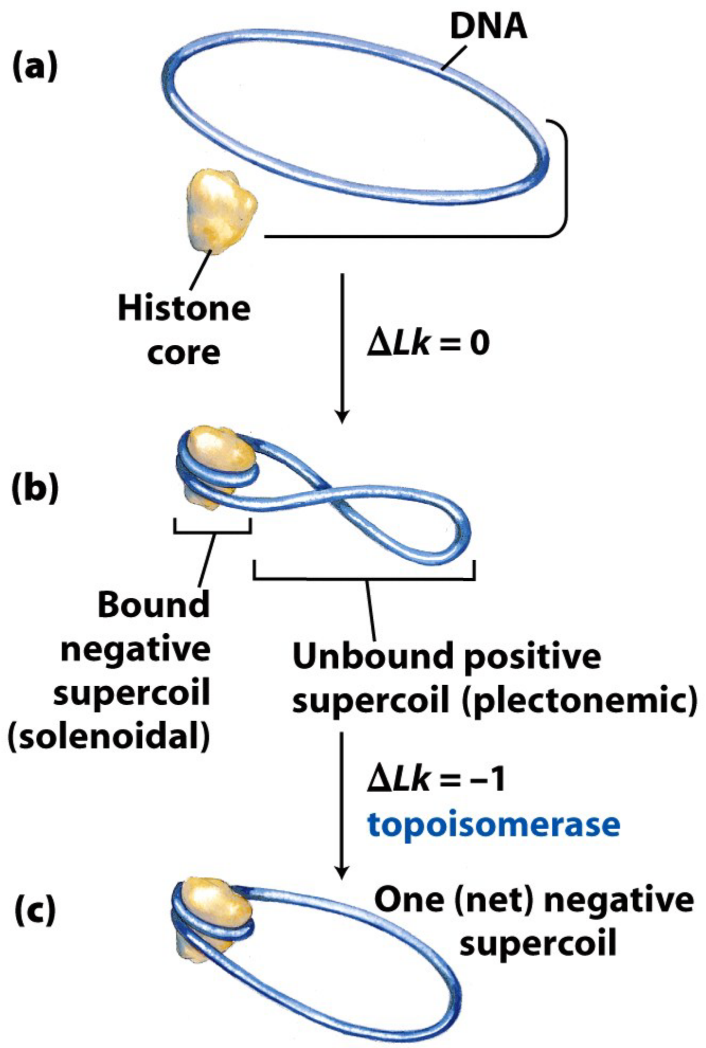
Two H2A-H2B Dimers

Histone Octamer Core

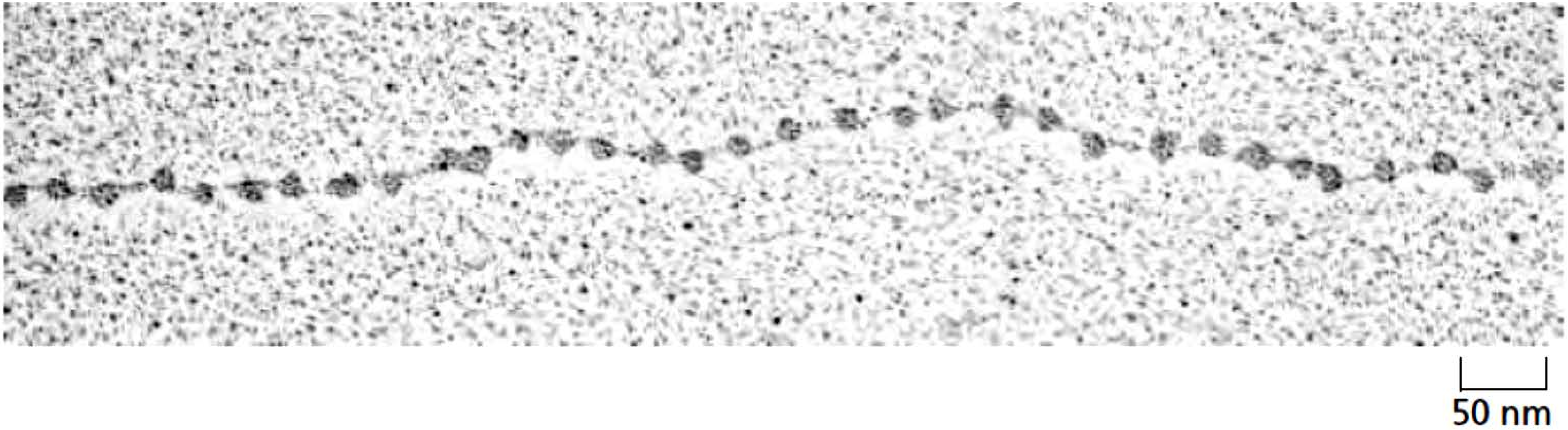
Wrapped with DNA

Nucleosome

Assembly of a Nucleosome

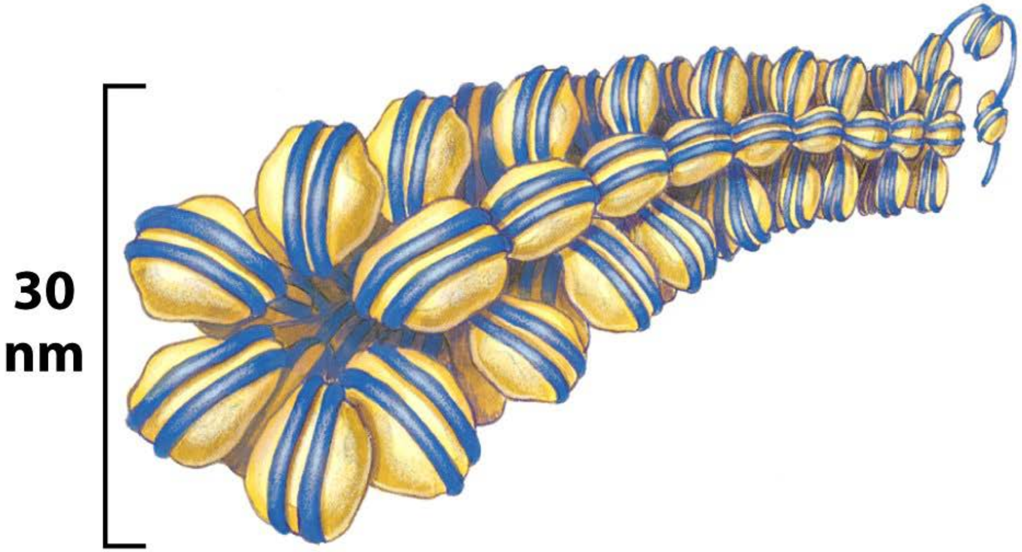


Beads on a String

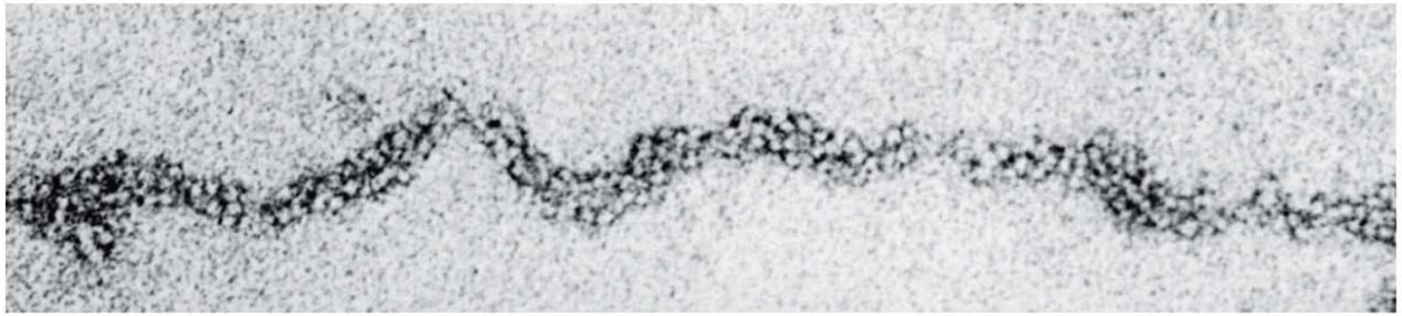


A living cell rarely adopts this form.

Nucleosomes Are Packed into Successively Higher-Order Structures

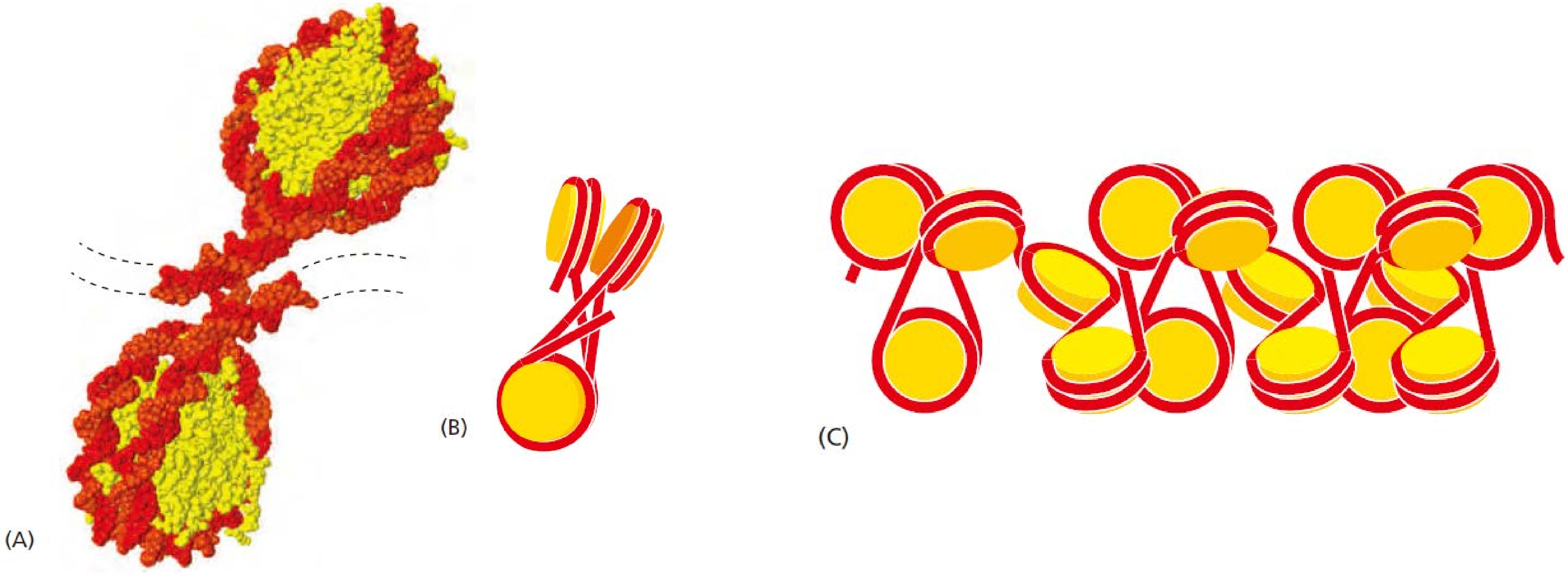


(a)



(b)

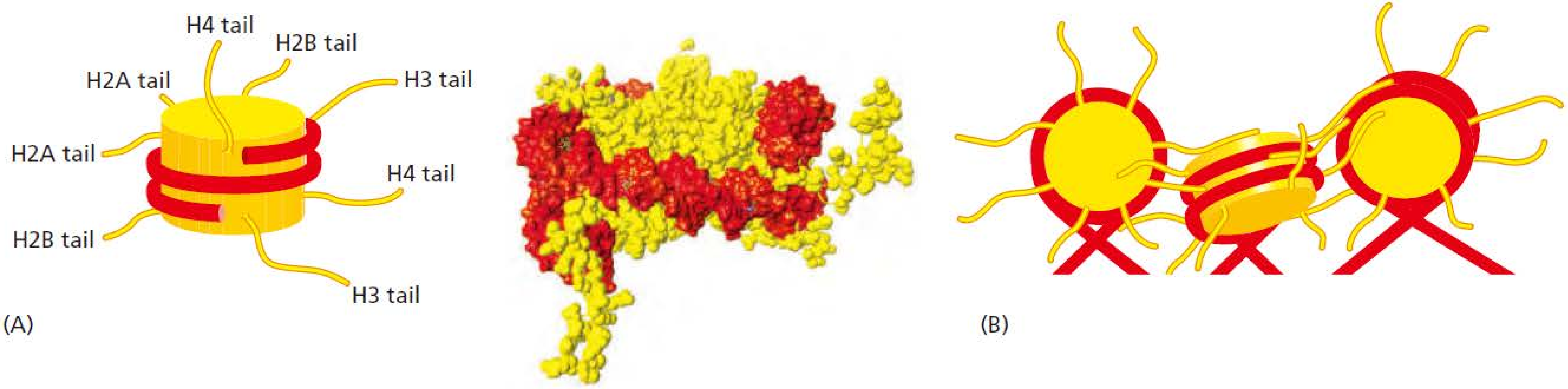
A Zigzag Model for the 30 nm Chromatin Fiber



What causes nucleosomes to stack so tightly on each other?

What Causes Nucleosomes to Stack So Tightly on Each Other?

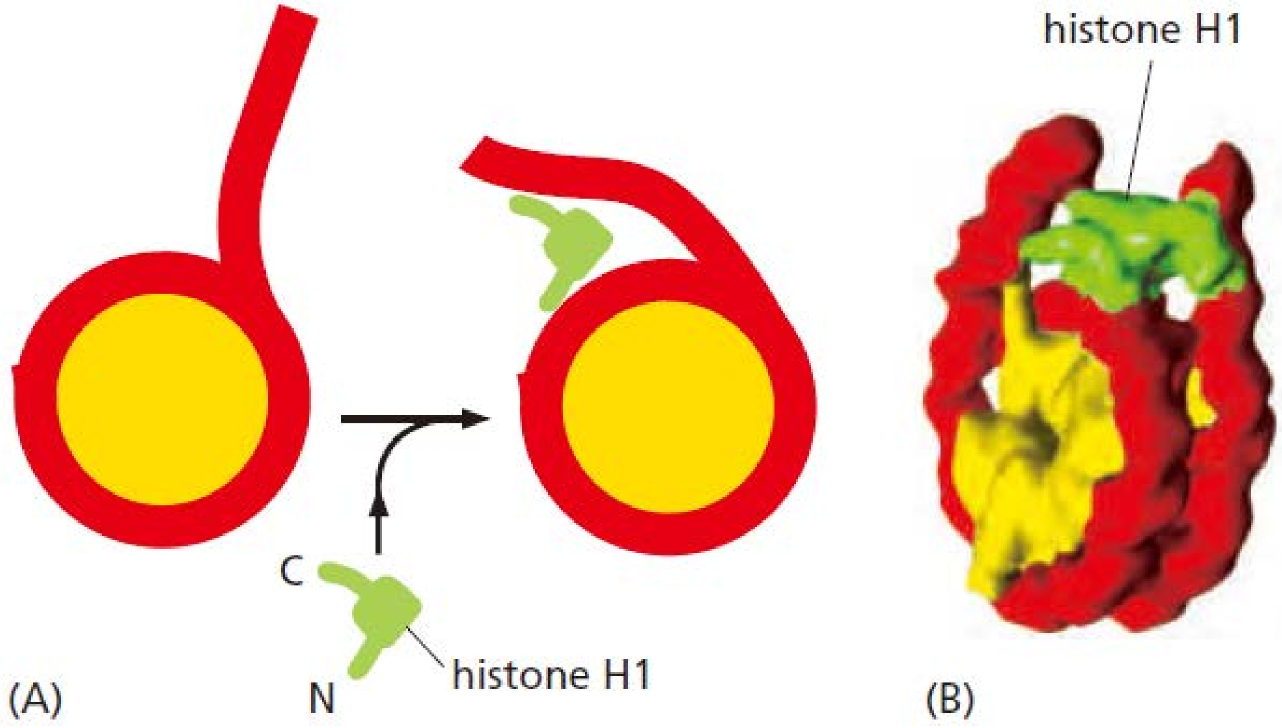
I: Nucleosome-to-nucleosome linkages that involved histone tails



Most notably, H4 tails might be more important.

What Causes Nucleosomes to Stack So Tightly on Each Other?

II. Linker Histone Binds to a Nucleosome

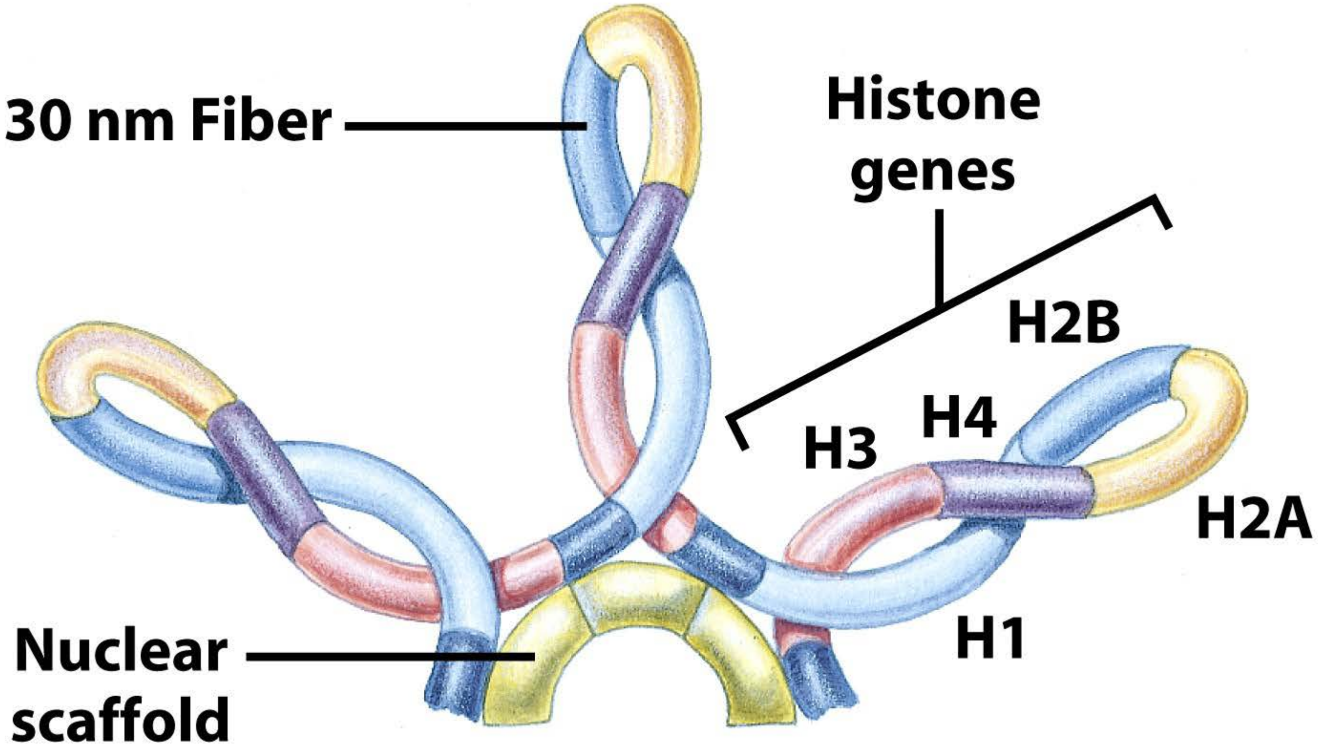


H1 Histone binds nucleosome in a 1:1 ratio.

Part III

Next Level of Organization

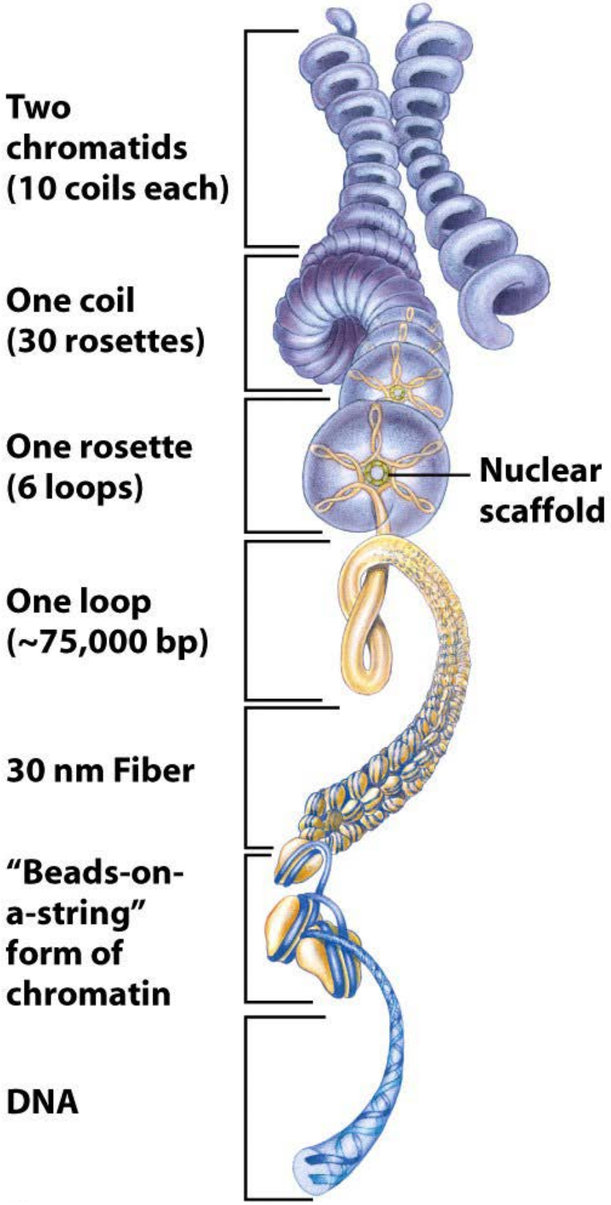
Loops of Chromosomal DNA Attached to a Nuclear Scaffold

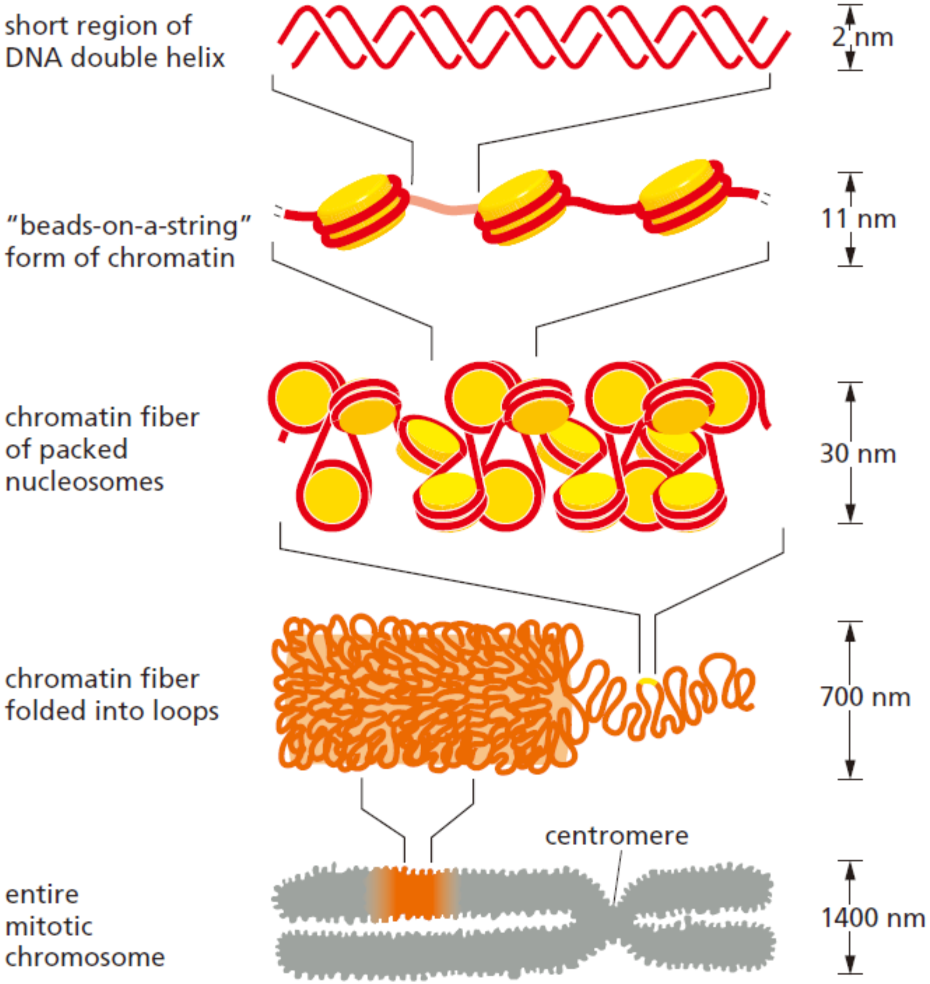


The scaffold may contain several proteins, notably large amounts of histone H1 and topoisomerase II.

Part III

Compaction of DNA in a Eukaryotic Chromosome





NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS FULLY EXTENDED LENGTH

Part III

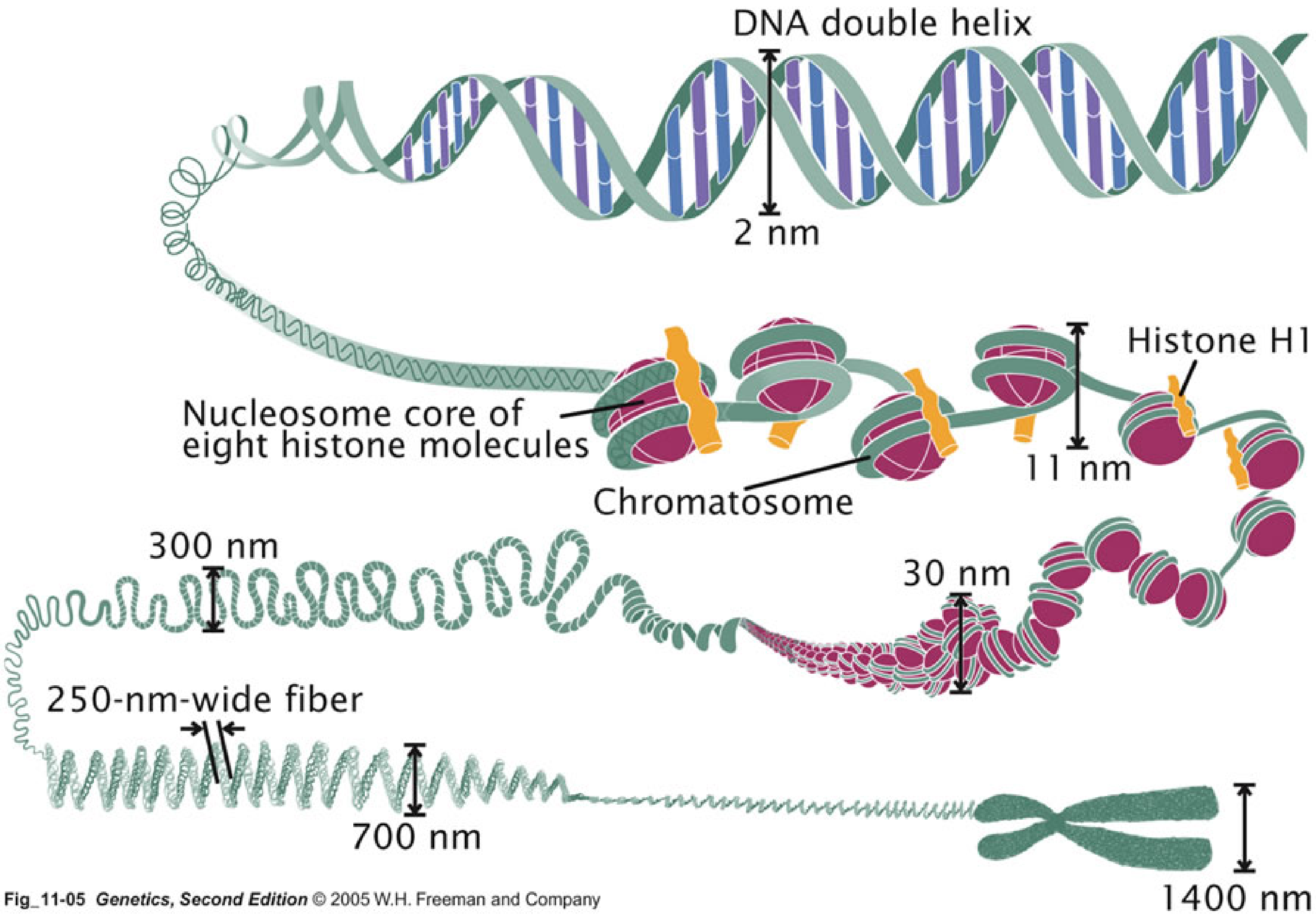


Fig_11-05 *Genetics, Second Edition* © 2005 W.H. Freeman and Company

One human genome: 1 m

Diploid: 2 m

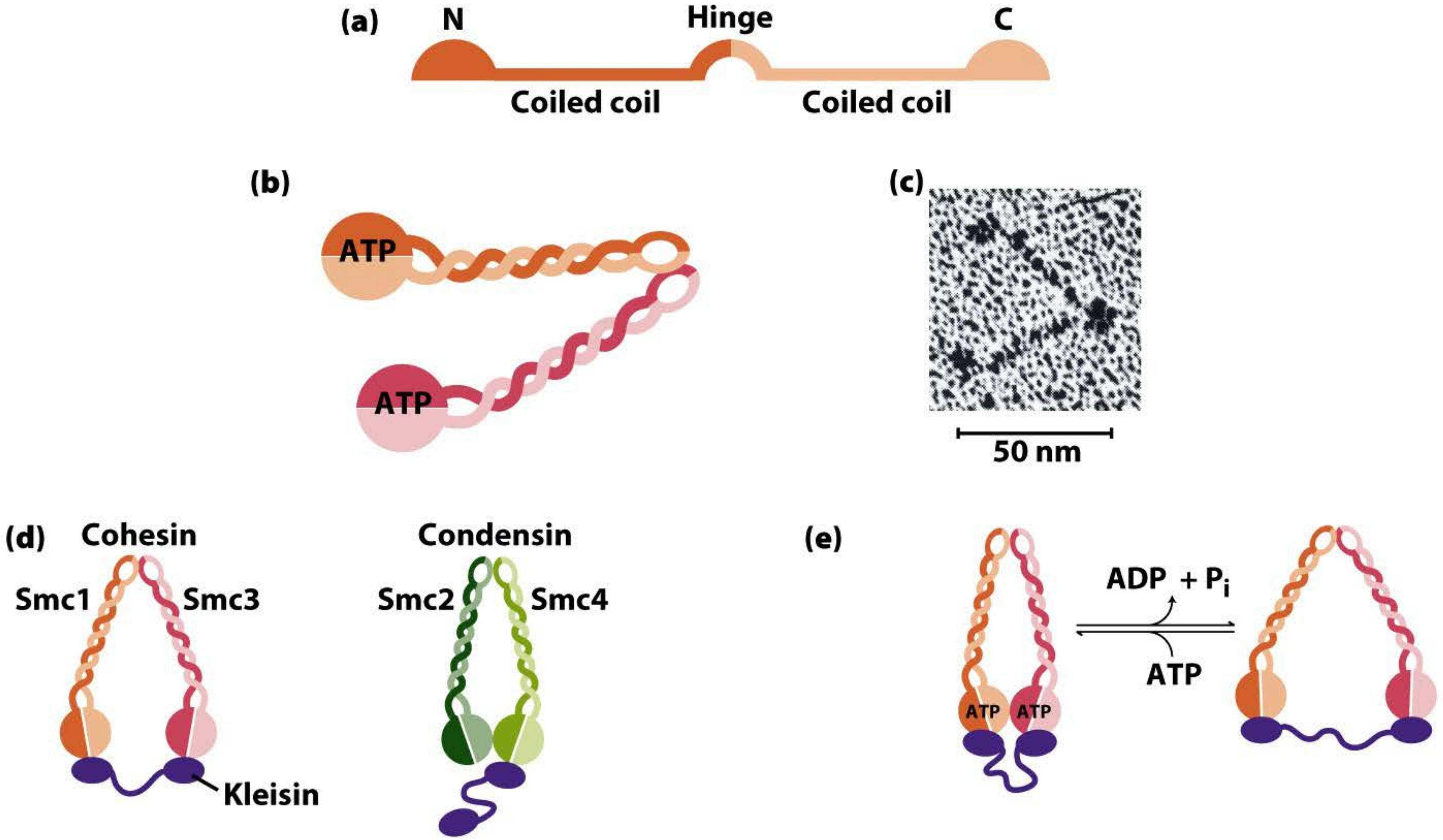
~10^{14} cells

Total DNA length: 2×10^{11} km

Circumference of the earth : 4×10^{4} km

Distance from earth to sun: 1.5×10^{8} km

Condensed Chromosome Structures Are Maintained by SMC Proteins



SMC: *S*tructural *M*aintenance of *C*hromosomes

Part III

SMC Proteins in Eukaryotes

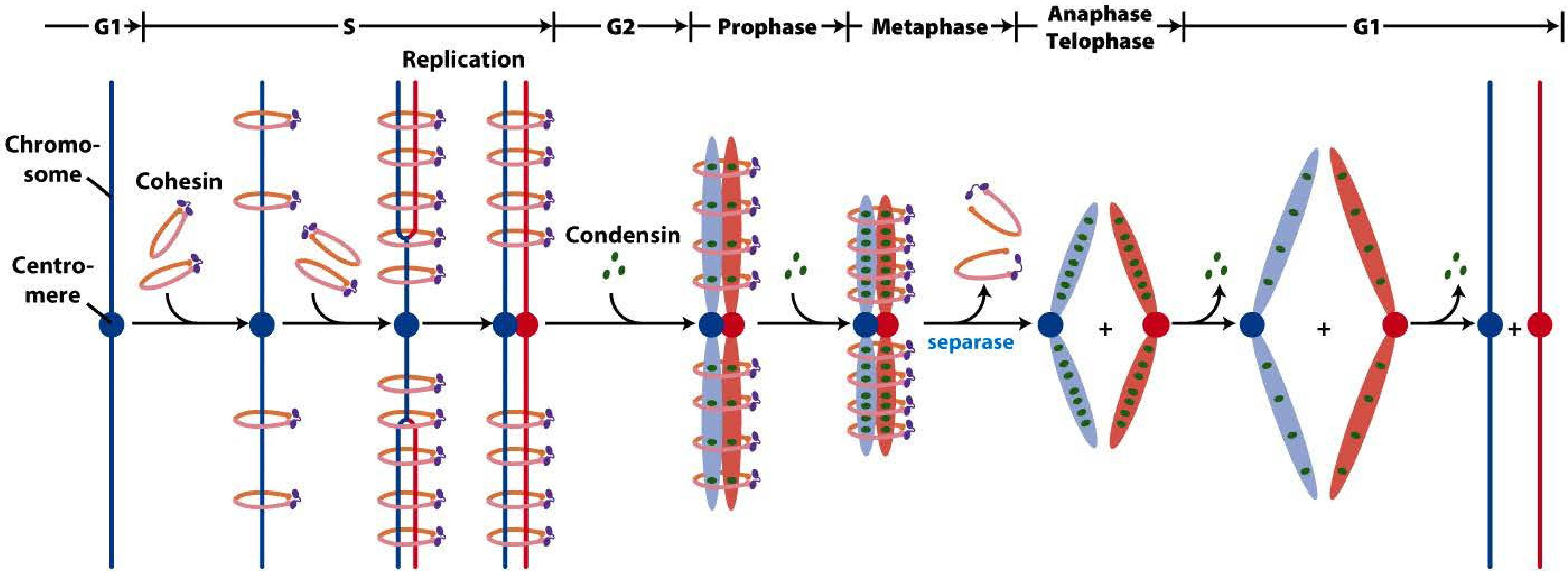
SMC: *S*tructural *M*aintenance of *C*hromosomes

Cohesins: Play a substantial role in linking together sister chromatids immediately after replication
Complex with Kleisin to form a ring around the replicated chromosome

Condensins: Essential to the condensation of chromosomes as cells enter mitosis
Create positive supercoils
Cause DNA to become overwound, in contrast to the underwinding induced by the binding of nucleosomes

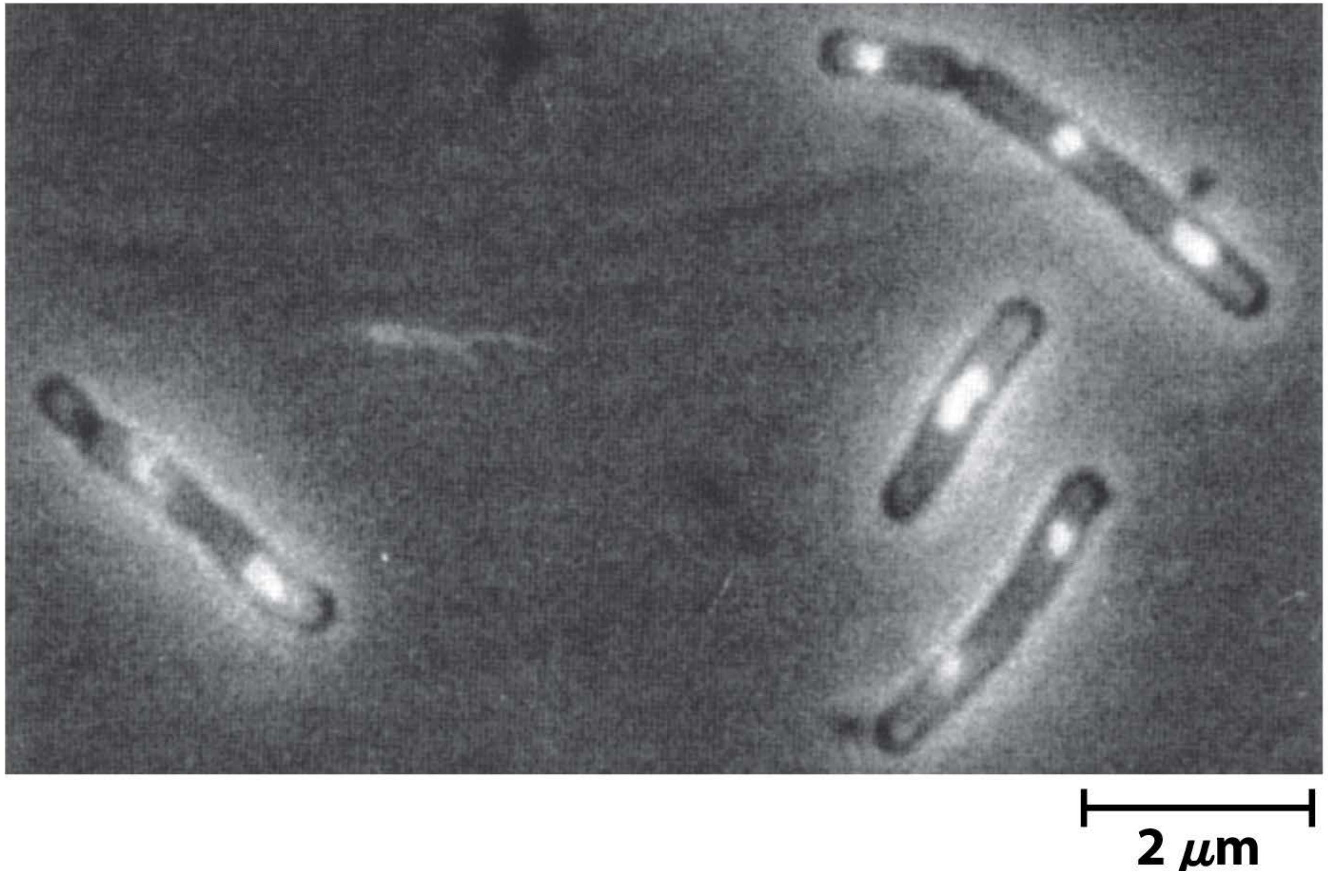
Part III

Model for the Roles of Cohesins and Codensins during Eukaryotic Cell Cycle

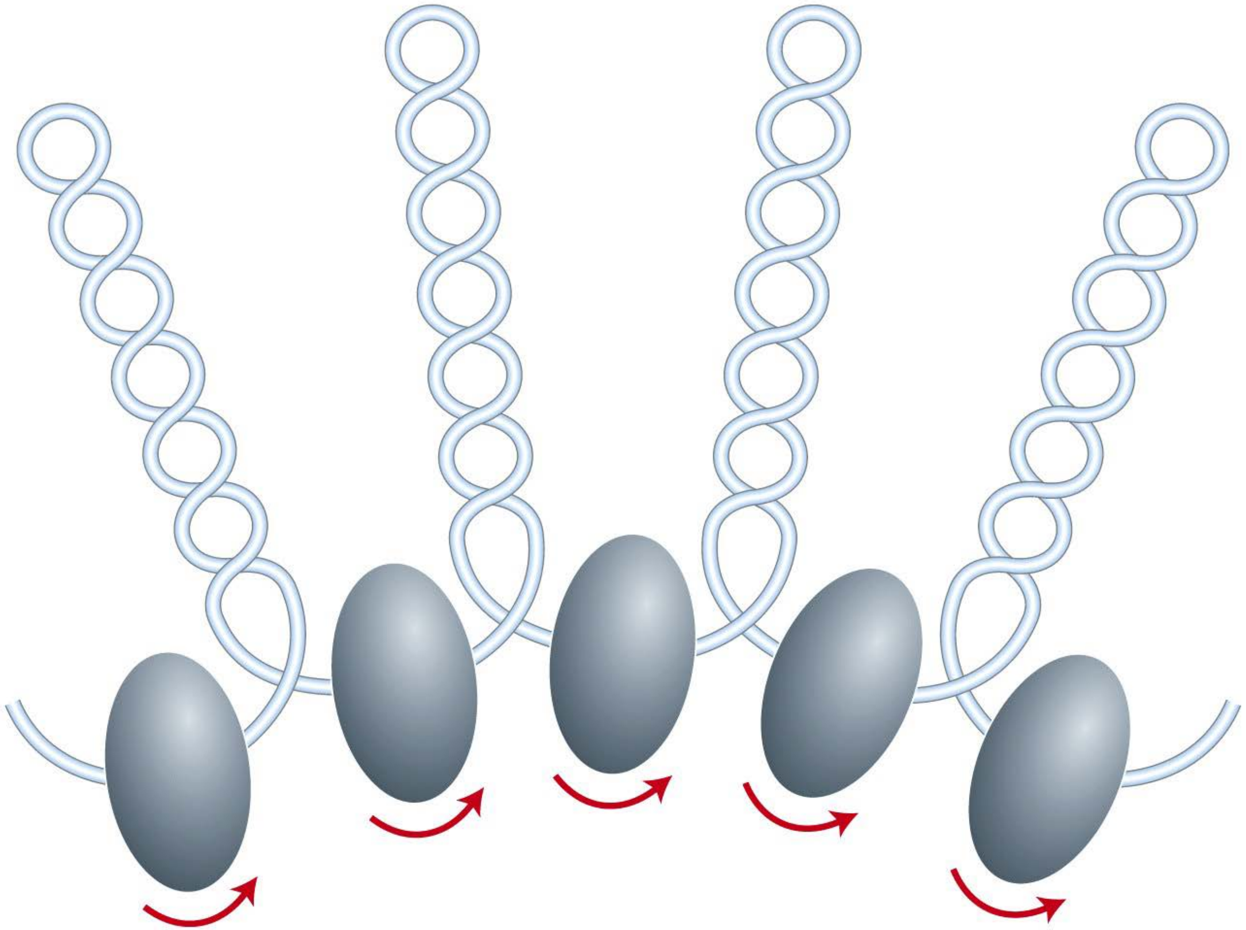


Bacterial DNA Is Also Highly Organized

Bacterial DNA is compacted into a structure called nucleoid.



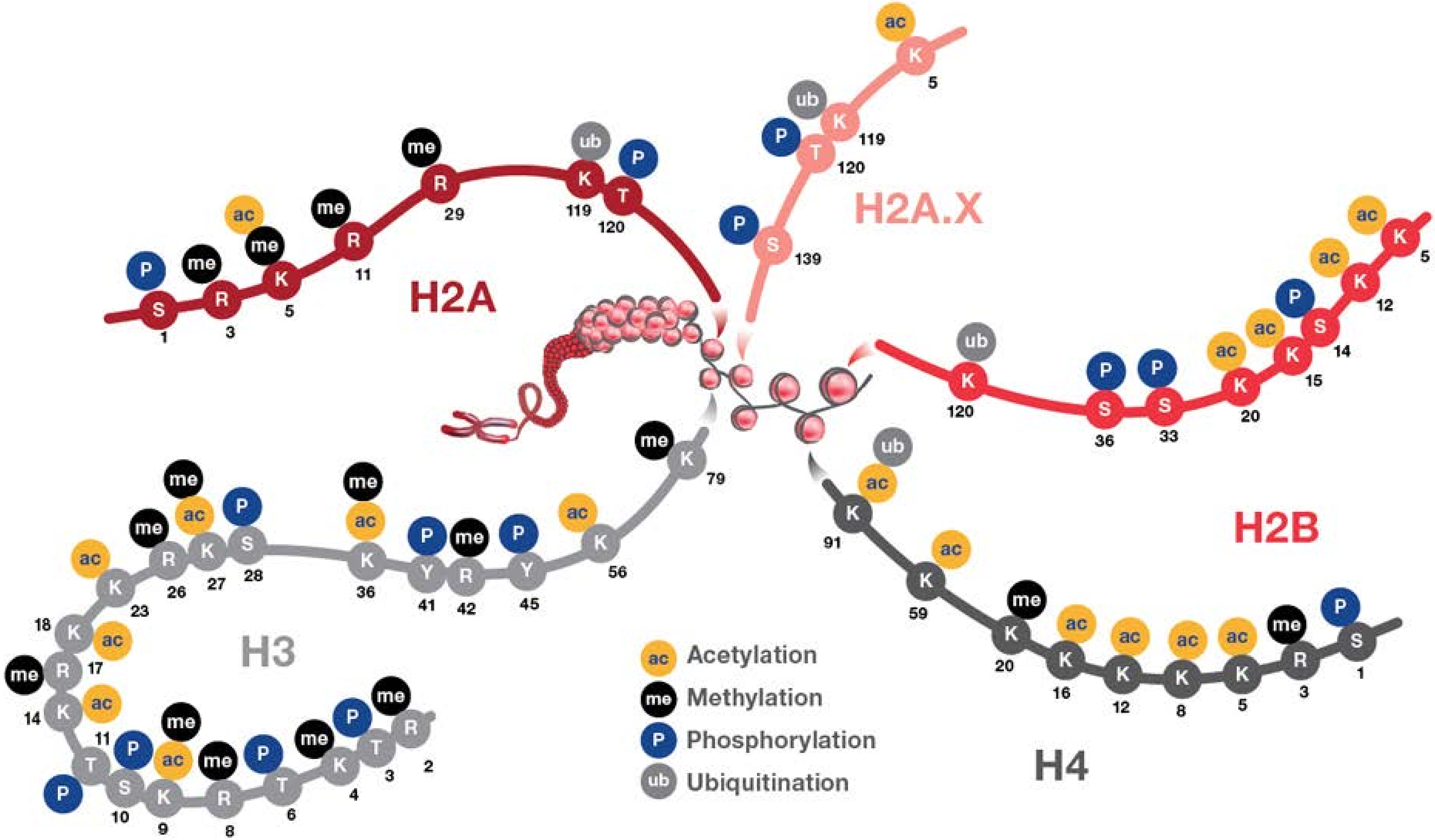
Looped Domain of the *E. Coli* Chromosome



About 500 looped domains, each encompassing 10,000 bp on average.

Part III

Epigenetics: Histone Modifications



Part III. The Structures of Chromosomes

The fundamental unit of organization in the chromatin of eukaryotic cells is the nucleosome.

Nucleosomes are organized into 30 nm fibers, which then fold extensively to provide 10,000 fold compaction required to fit a typical eukaryotic chromosome into a nucleus.

Nuclear scaffold proteins, Histone H1, topoisomerase and SMC proteins, are involved in higher-order folding events.

Bacterial chromosomes are extensively compacted into a nucleoid.

Folding of Double-Strand DNA into Chromosomes

Question Time

References:

Lehninger Principle of Biochemistry (fifth edition)
Molecular Biology of the Cell (sixth edition)

Find me @C328