Chromosome organization and molecular structure

Presented by: Dr.Asma

INTRODUCTION

- Chromosomes are the structures that contain the genetic material
  - They are complexes of DNA and proteins
- The genome comprises all the genetic material that an organism possesses
  - In bacteria, it is typically a single circular chromosome
  - In eukaryotes, it refers to one complete set of *nuclear* chromosomes
  - Note:
    - Eukaryotes possess a mitochondrial genome
    - Plants also have a chloroplast genome

INTRODUCTION

- The main function of the genetic material is to store information required to produce an organism
  - The DNA molecule does that through its base sequence

- DNA sequences are necessary for
  1. Synthesis of RNA and cellular proteins
  2. Proper segregation of chromosomes
  3. Replication of chromosomes
  4. Compaction of chromosomes
     - So they can fit within living cells

10.1 Viral Genomes

- Viruses are small infectious particles containing nucleic acid surrounded by a capsid of proteins
  - Refer to Figure 10.1
- For replication, viruses rely on their host cells
  - ie., the cells they infect
- Most viruses exhibit a limited host range
  - They typically infect only specific types of cells of one host species

- Bacteriophages may also contain a sheath, base plate and tail fibers
  - Refer to Figure 9.4

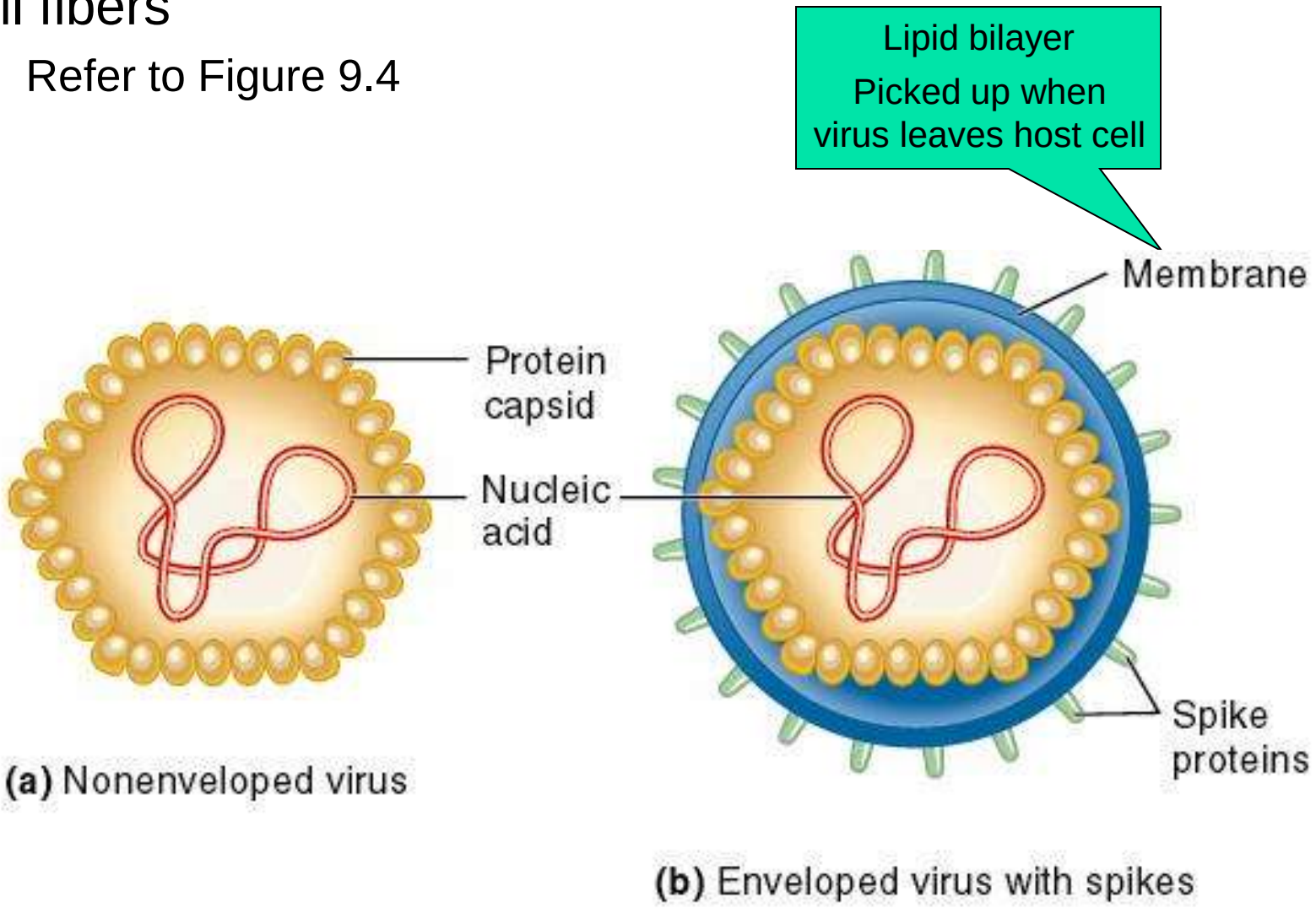


Figure 10.1 **General structure of viruses**

Viral Genomes

- A viral genome is the genetic material of the virus
  - Also termed the viral chromosome
- The genome can be
  - DNA or RNA
  - Single-stranded or double-stranded
  - Circular or linear
- Viral genomes vary in size from a few thousand to more than a hundred thousand nucleotides

TABLE 10.1

Characteristics of Selected Viral Genomes

Virus	Host	Type of Nucleic Acid*	Size**	Number of Genes
Parvovirus	Mammals	ssDNA	5.0	5
Phage fd	*E. coli*	ssDNA	6.4	10
Lambda	*E. coli*	dsDNA	48.5	36
T4	*E. coli*	dsDNA	169	>190
Qβ	*E. coli*	ssRNA	4.2	4
TMV	Many plants	ssRNA	6.4	6
Influenza virus	Mammals	ssRNA	13.5	12

*ss refers to single stranded and ds refers to double stranded.

**Number of thousands of nucleotides or nucleotide base pairs.

10.2 Bacterial Chromosomes

- The bacterial chromosome is found in a region called the nucleoid
  - Refer to Figure 10.3
- The nucleoid is not membrane-bounded
  - So the DNA is in direct contact with the cytoplasm
- Bacteria may have one to four identical copies of the same chromosome
  - The number depends on the species and growth conditions

- Bacterial chromosomal DNA is usually a circular molecule that is a few million nucleotides in length
  - *Escherichia coli* → ~ 4.6 million base pairs
  - *Haemophilus influenzae* → ~ 1.8 million base pairs

- A typical bacterial chromosome contains a few thousand different genes
  - Structural gene sequences (encoding proteins) account for the majority of bacterial DNA
  - The nontranscribed DNA between adjacent genes are termed intergenic regions

- Figure 10.4 summarizes the key features of bacterial chromosomes

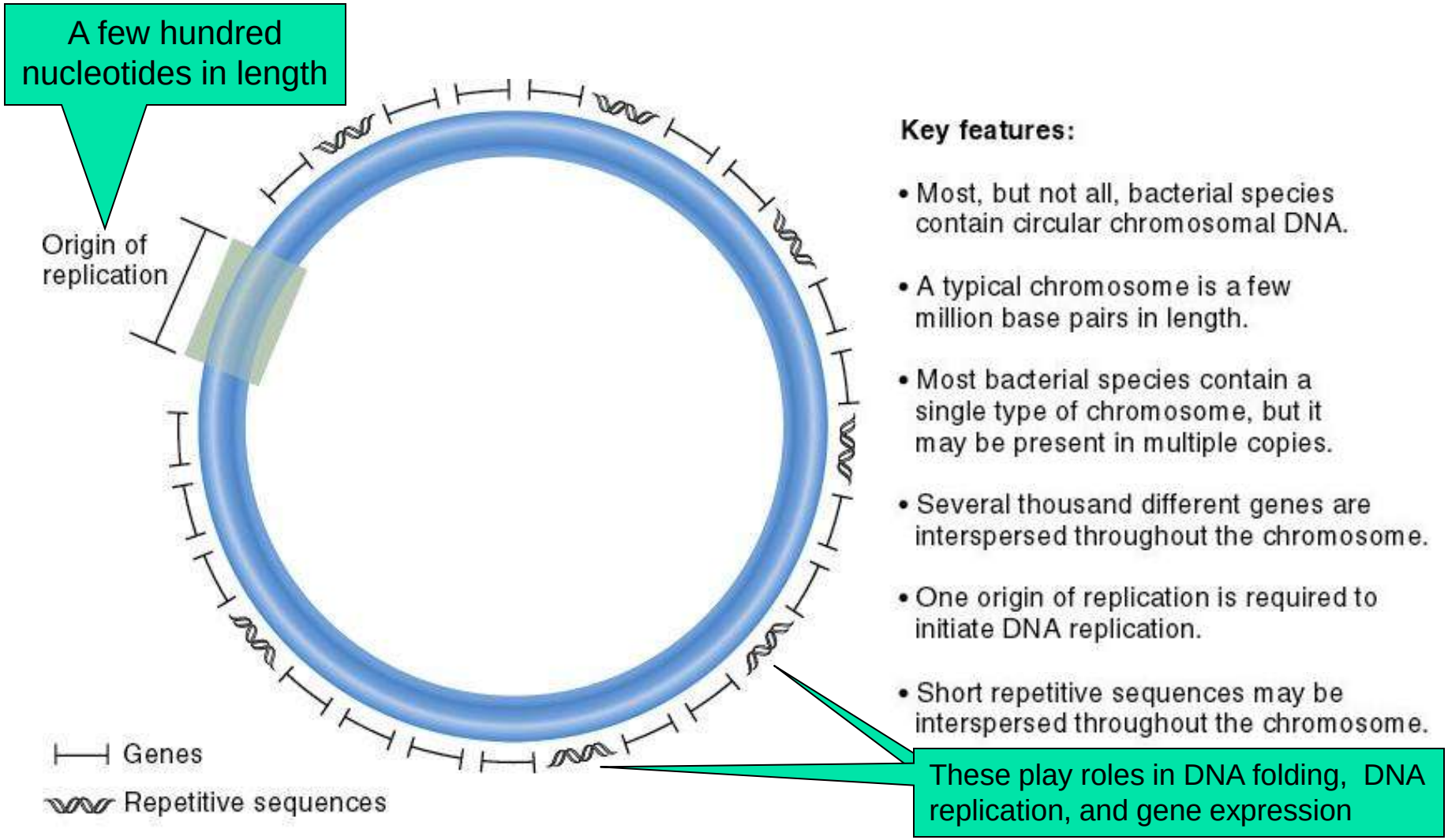

A few hundred nucleotides in length
Origin of replication
Genes
Repetitive sequences
Key features:
• Most, but not all, bacterial species contain circular chromosomal DNA.
• A typical chromosome is a few million base pairs in length.
• Most bacterial species contain a single type of chromosome, but it may be present in multiple copies.
• Several thousand different genes are interspersed throughout the chromosome.
• One origin of replication is required to initiate DNA replication.
• Short repetitive sequences may be interspersed throughout the chromosome.
These play roles in DNA folding, DNA replication, and gene expression

Figure 10.4

- To fit within the bacterial cell, the chromosomal DNA must be compacted about a 1000-fold
  - This involves the formation of loop domains

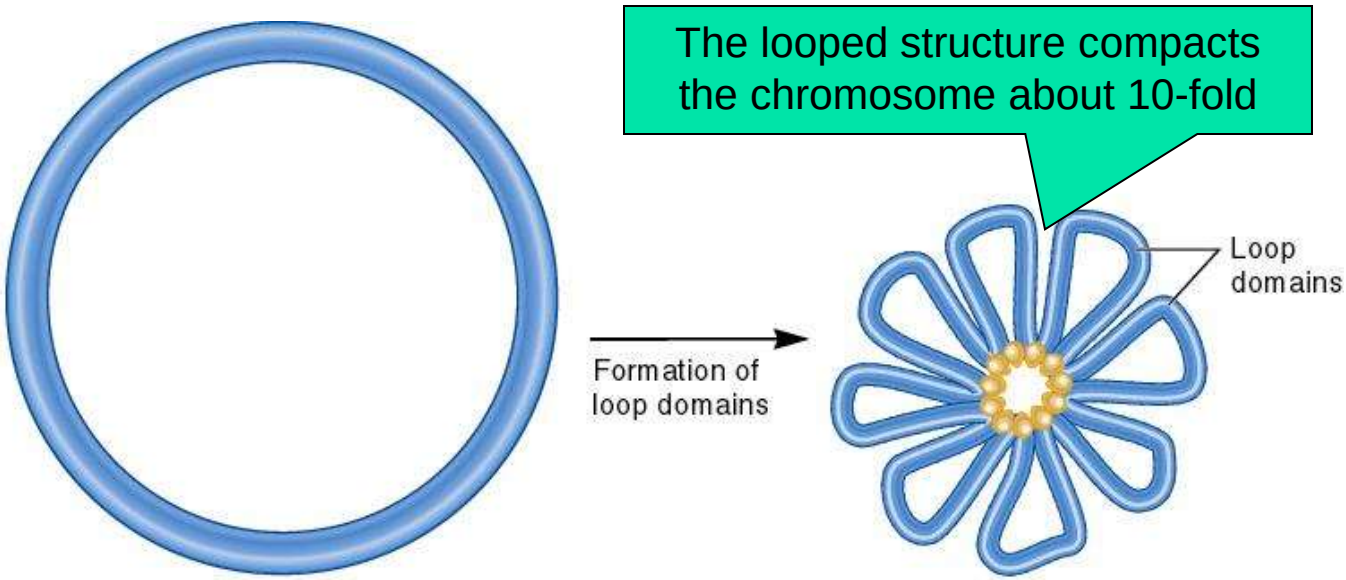


Figure 10.5

- The number of loops varies according to the size of the bacterial chromosome and the species
  - *E. coli* has 50-100 with 40,000 to 80,000 bp of DNA in each

- DNA supercoiling is a second important way to compact the bacterial chromosome

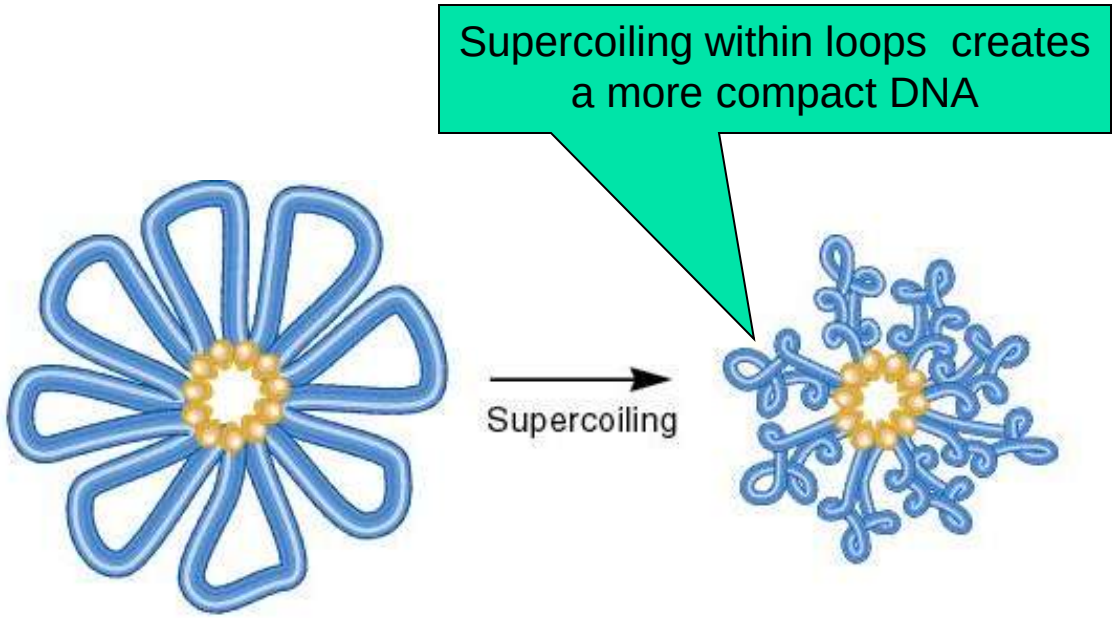


Figure 10.6

- Figure 10.7 provides a schematic illustration of DNA supercoiling

- The control of supercoiling in bacteria is accomplished by two main enzymes
  - 1. DNA gyrase (also termed DNA topoisomerase II)
    - Introduces negative supercoils using energy from ATP
      - Refer to Figure 10.9
    - It can also relax positive supercoils when they occur
  - 2. DNA topoisomerase I
    - Relaxes negative supercoils
- The competing action of these two enzymes governs the overall supercoiling of bacterial DNA

10.3 Eukaryotic Chromosomes

- Eukaryotic species contain one or more sets of chromosomes
  - Each set is composed of several different linear chromosomes
- The total amount of DNA in eukaryotic species is typically greater than that in bacterial cells
- Chromosomes in eukaryotes are located in the nucleus
  - To fit in there, they must be highly compacted
    - This is accomplished by the binding of many proteins
    - The DNA-protein complex is termed chromatin

- Eukaryotic genomes vary substantially in size
  - Refer to Figure 10.10*a*

- In many cases, this variation is not related to complexity of the species
  - For example, there is a two fold difference in the size of the genome in two closely related salamander species
  - Refer to Figure 10.10*b*

  - The difference in the size of the genome is not because of extra genes
    - Rather, the accumulation of repetitive DNA sequences
      - These do not encode proteins

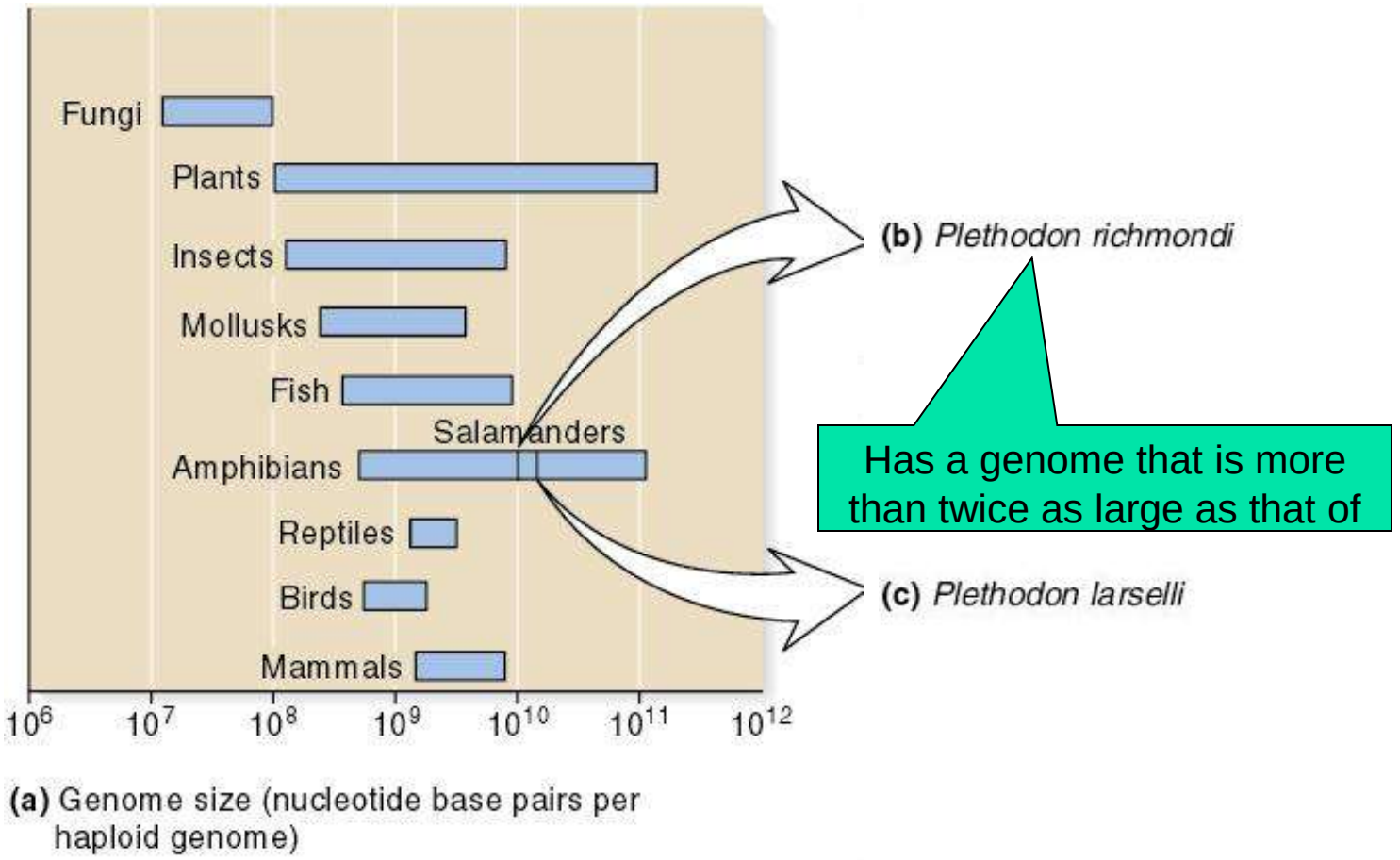
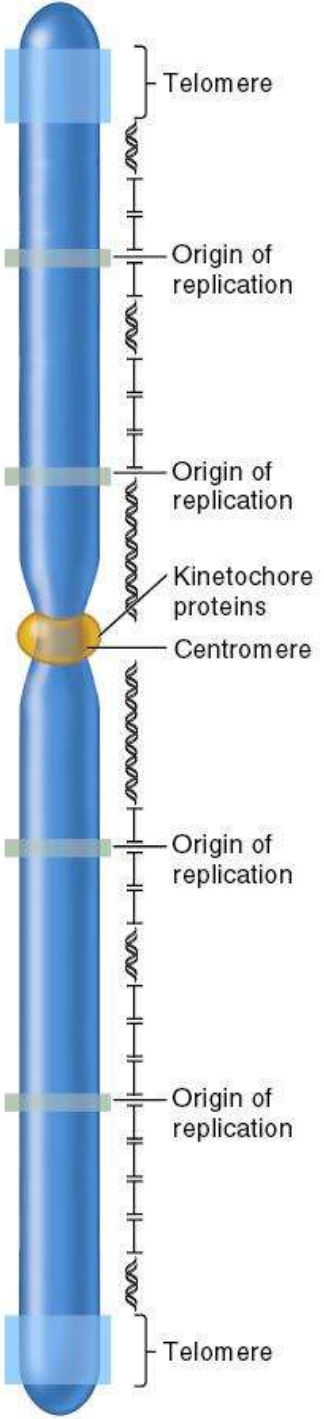

Fungi
Plants
Insects
Mollusks
Fish
Salamanders
Amphibians
Reptiles
Birds
Mammals
10^6
10^7
10^8
10^9
10^{10}
10^{11}
10^{12}
(a) Genome size (nucleotide base pairs per haploid genome)
(b) *Plethodon richmondi*
Has a genome that is more than twice as large as that of
(c) *Plethodon larselli*

Figure 10.10

Organization of Eukaryotic Chromosomes

- A eukaryotic chromosome contains a long, linear DNA molecule
  - Refer to Figure 10.11

- Three types of DNA sequences are required for chromosomal replication and segregation
  - Origins of replication
  - Centromeres
  - Telomeres



Key features:

- Eukaryotic chromosomes are usually linear.
- A typical chromosome is tens of millions to hundreds of millions of base pairs in length.
- Eukaryotic chromosomes occur in sets. Many species are diploid, which means that somatic cells contain 2 sets of chromosomes.
- Genes are interspersed throughout the chromosome. A typical chromosome contains between a few hundred and several thousand different genes.
- Each chromosome contains many origins of replication that are interspersed about every 100,000 base pairs.
- Each chromosome contains a centromere that forms a recognition site for the kinetochore proteins.
- Telomeres contain specialized sequences located at both ends of the linear chromosome.
- Repetitive sequences are commonly found near centromeric and telomeric regions, but they may also be interspersed throughout the chromosome.

Genes

Repetitive sequences

Figure 10.11

- Genes are located between the centromeric and telomeric regions along the entire chromosome
  - A single chromosome usually has a few hundred to several thousand genes

- In lower eukaryotes (such as yeast)
  - Genes are relatively small
    - They contain primarily the sequences encoding the polypeptides
    - ie: Very few introns are present

- In higher eukaryotes (such as mammals)
  - Genes are long
    - They tend to have many introns

Repetitive Sequences

- Sequence complexity refers to the number of times a particular base sequence appears in the genome

- There are three main types of repetitive sequences
  - Unique or non-repetitive
  - Moderately repetitive
  - Highly repetitive

Repetitive Sequences

- Unique or non-repetitive sequences
  - Found once or a few times in the genome
  - Includes structural genes as well as intergenic areas
- Moderately repetitive
  - Found a few hundred to a few thousand times
  - Includes
    - Genes for rRNA and histones
    - Origins of replication
    - Transposable elements

Repetitive Sequences

- Highly repetitive
  - Found tens of thousands to millions of times
  - Each copy is relatively short (a few nucleotides to several hundred in length)
  - Some sequences are interspersed throughout the genome
    - Example: *Alu* family in humans
      - Discussed in detail in Chapter 17
  - Other sequences are clustered together in tandem arrays
    - Example: AATAT and AATATAT sequences in *Drosophila*
    - These are commonly found in the centromeric regions

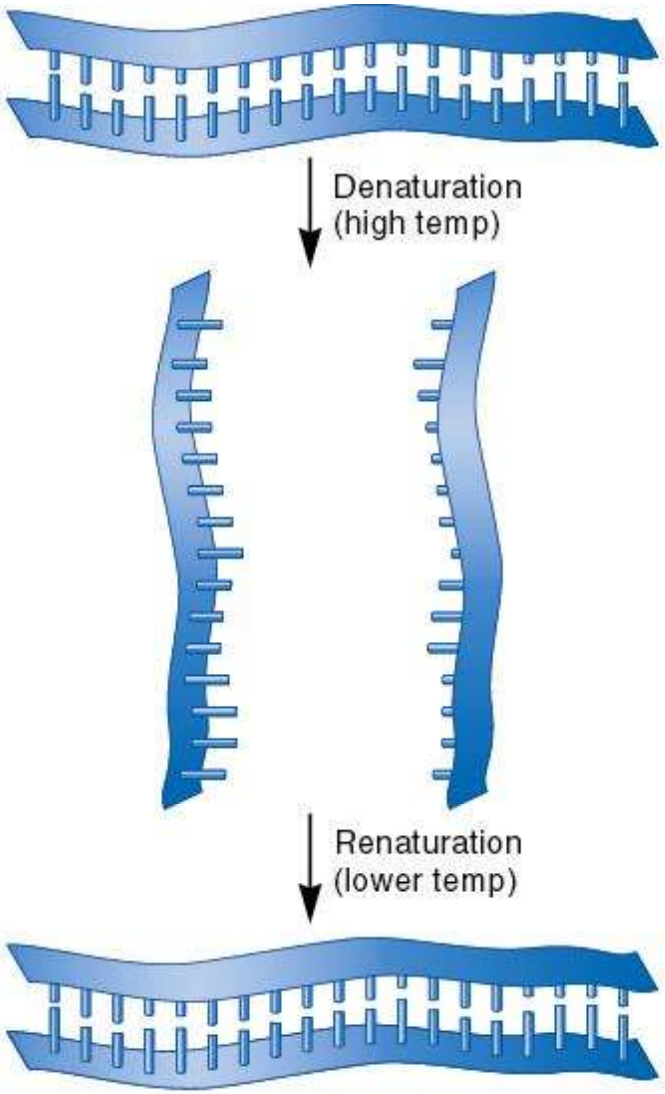


Figure 10.13 **(a)** Renaturation of DNA strands

Renaturation Experiments

- The rate of renaturation of complementary DNA strands provides a way to distinguish the three different types of repetitive sequences

- The renaturation rate of a particular DNA sequence depends on the concentration of its complementary partner
  - Highly repetitive DNA will be the fastest to renature
    - Because there are many copies of complementary sequences
  - Unique sequences will be the slowest to renature
    - It takes added time for these sequences to find each other

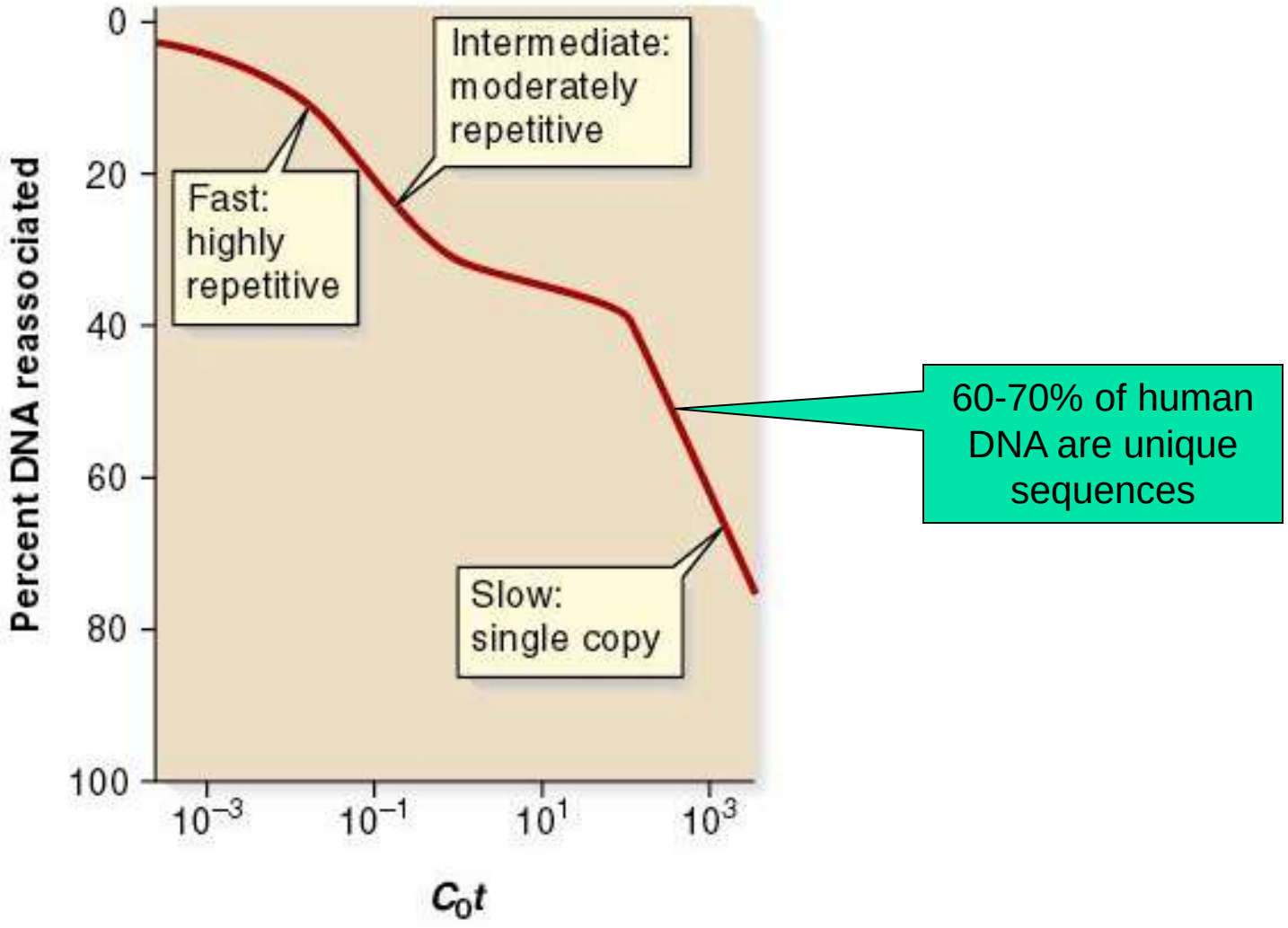


Figure 10.13 **(b)** Human chromosomal DNA C_0t curve

Eukaryotic Chromatin Compaction

- If stretched end to end, a single set of human chromosomes will be over 1 meter long!
  - Yet the cell's nucleus is only 2 to 4 μm in diameter
    - Therefore, the DNA must be tightly compacted to fit
- The compaction of linear DNA in eukaryotic chromosomes involves interactions between DNA and various proteins
  - Proteins bound to DNA are subject to change during the life of the cell
    - These changes affect the degree of chromatin compaction

NUCLEOSOMES

- The repeating structural unit within eukaryotic chromatin is the nucleosome
- It is composed of double-stranded DNA wrapped around an octamer of histone proteins
  - An octamer is composed two copies each of four different histones
  - 146 bp of DNA make 1.65 negative superhelical turns around the octamer
- Refer to Figure 10.14*a*

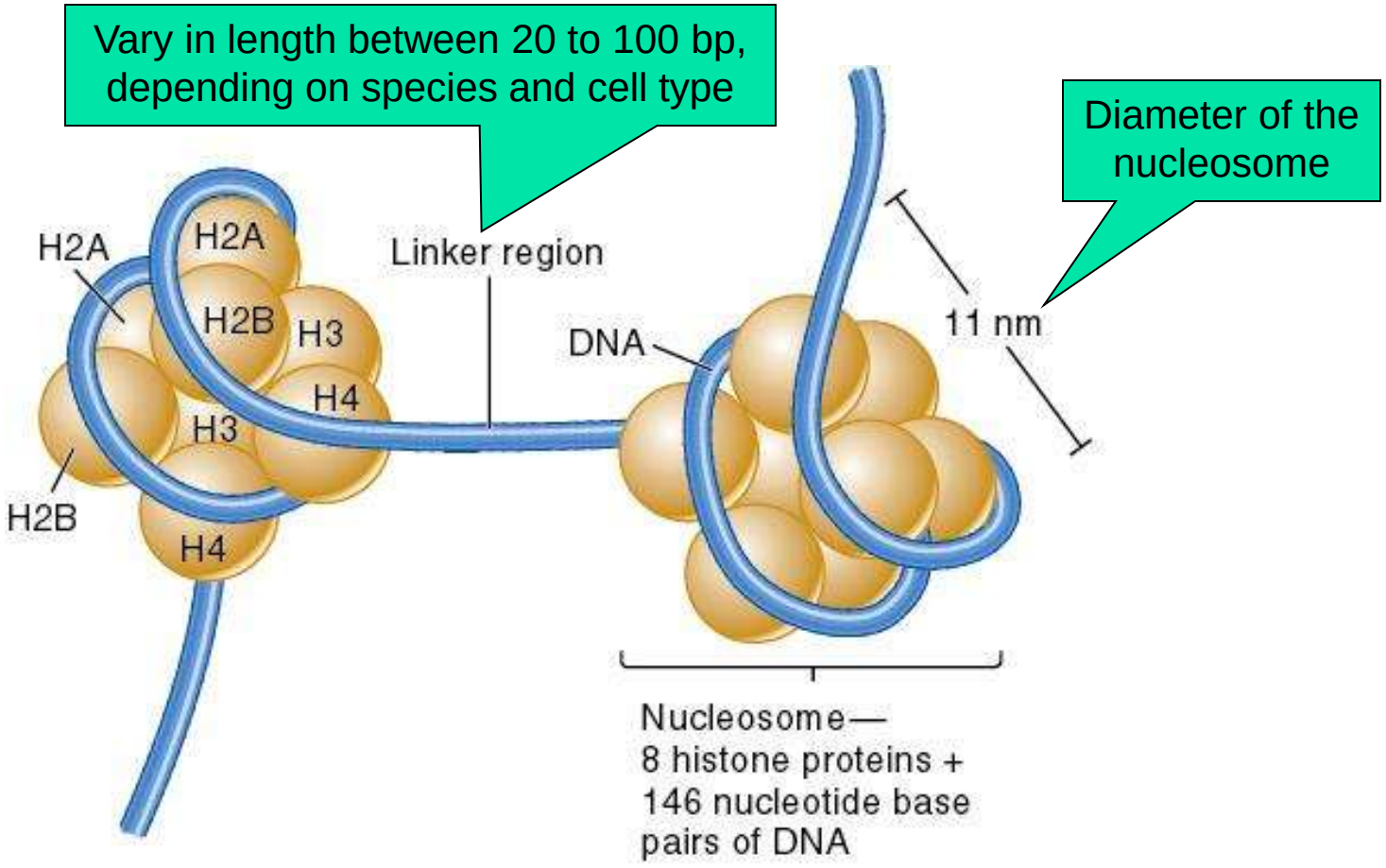


Figure 10.14 **(a)** Nucleosomes showing core histones

- Overall structure of connected nucleosomes resembles “beads on a string”
  - This structure shortens the DNA length about seven-fold

- Histone proteins are basic
  - They contain many positively-charged amino acids
    - Lysine and arginine
  - These bind with the phosphates along the DNA backbone

- There are five types of histones
  - H2A, H2B, H3 and H4 are the core histones
    - Two of each make up the octamer
  - H1 is the linker histone
    - Binds to linker DNA
    - Also binds to nucleosomes
      - But not as tightly as are the core histones
  - Refer to Figure 10.14*b*

Play a role in the organization and compaction of the chromosome

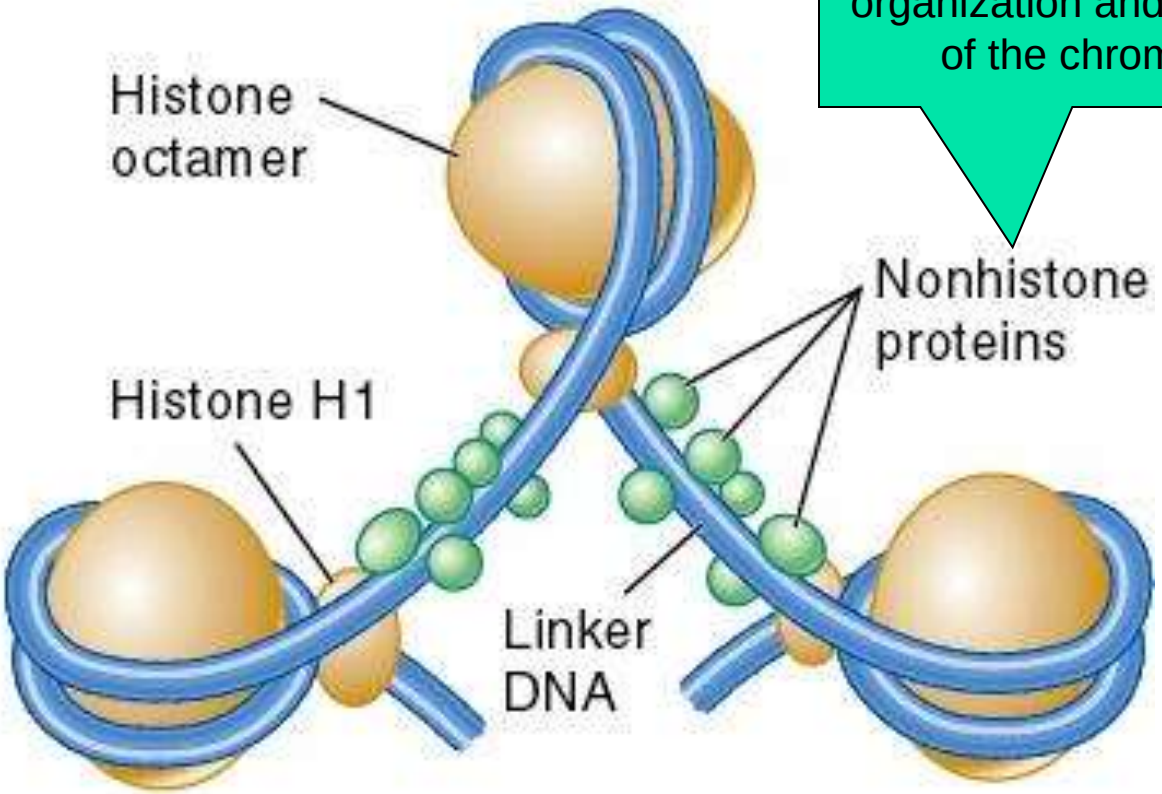


Figure 10.14 **(b)** Nucleosomes showing linker histones and nonhistone proteins

Nucleosomes join to form a 30 nm fiber

- Nucleosomes associate with each other to form a more compact structure termed the 30 nm fiber
- Histone H1 plays a role in this compaction
  - At moderate salt concentrations, H1 is removed
    - The result is the classic beads-on-a-string morphology
  - At low salt concentrations, H1 remains bound
    - Beads associate together into a more compact morphology
  - Refer to Figure 10.16

- Fig. 10.17*a* shows a micrograph of the 30 nm fiber
- The 30 nm fiber shortens the total length of DNA another seven-fold
- Its structure has proven difficult to determine
  - The DNA conformation may be substantially altered when extracted from living cells
  - Two models have been proposed
    - Solenoid model
    - Three-dimensional zigzag model
  - Refer to Figure 10.17*b* and *c*

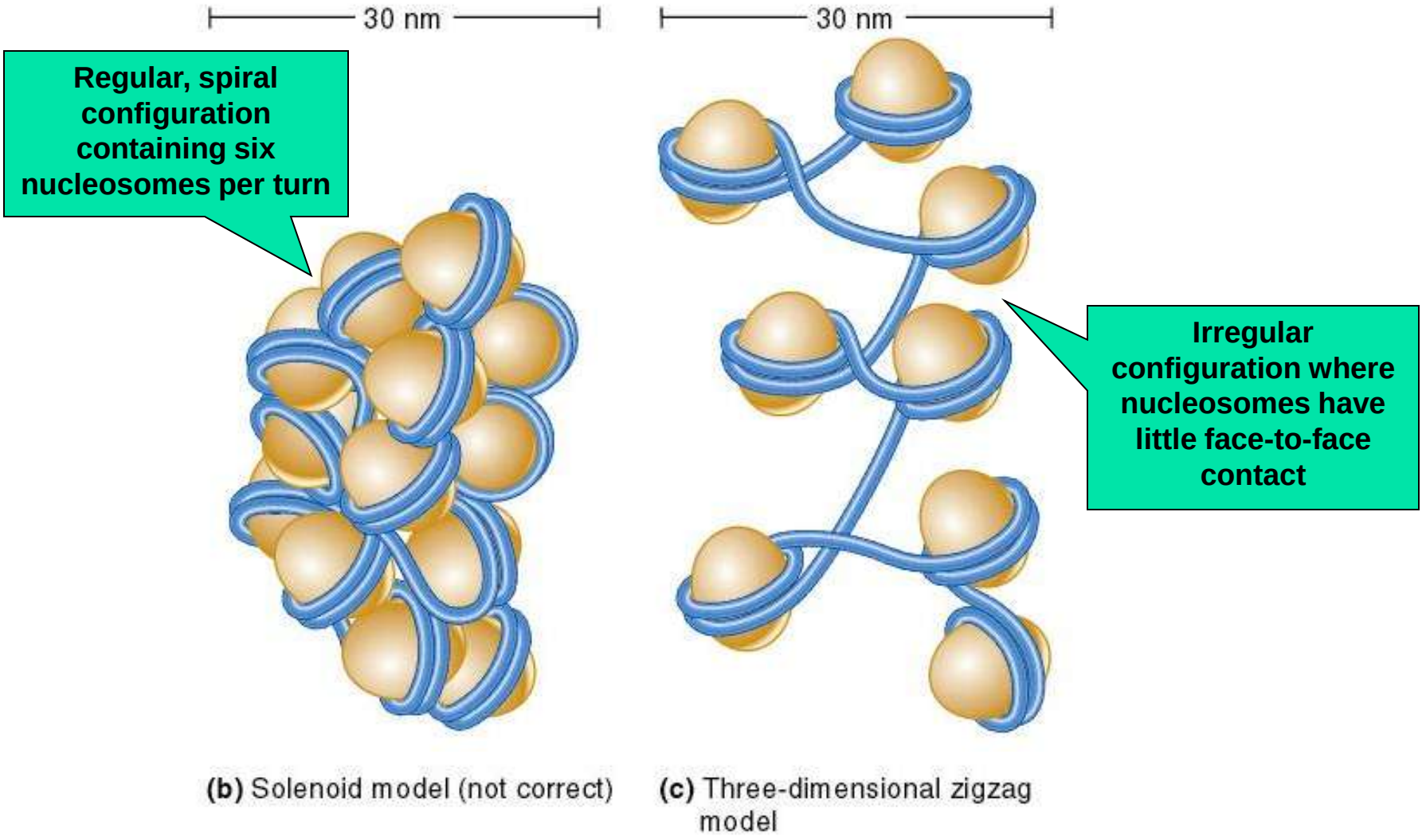


Figure 10.17

Further Compaction of the Chromosome

- The two events we have discussed so far have shortened the DNA about 50-fold
- A third level of compaction involves interaction between the 30 nm fiber and the nuclear matrix
- The nuclear matrix is composed of two parts
  - Nuclear lamina
  - Internal matrix proteins
    - 10 nm fiber and associated proteins

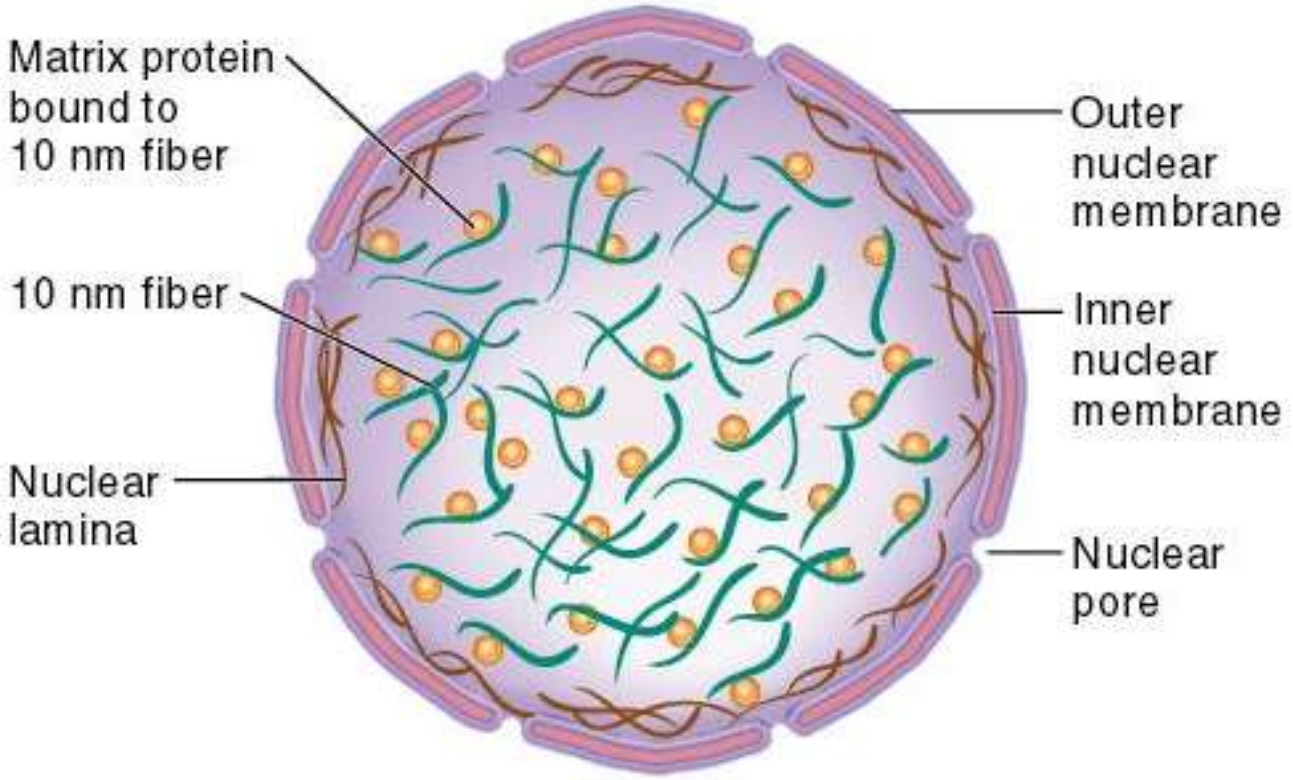
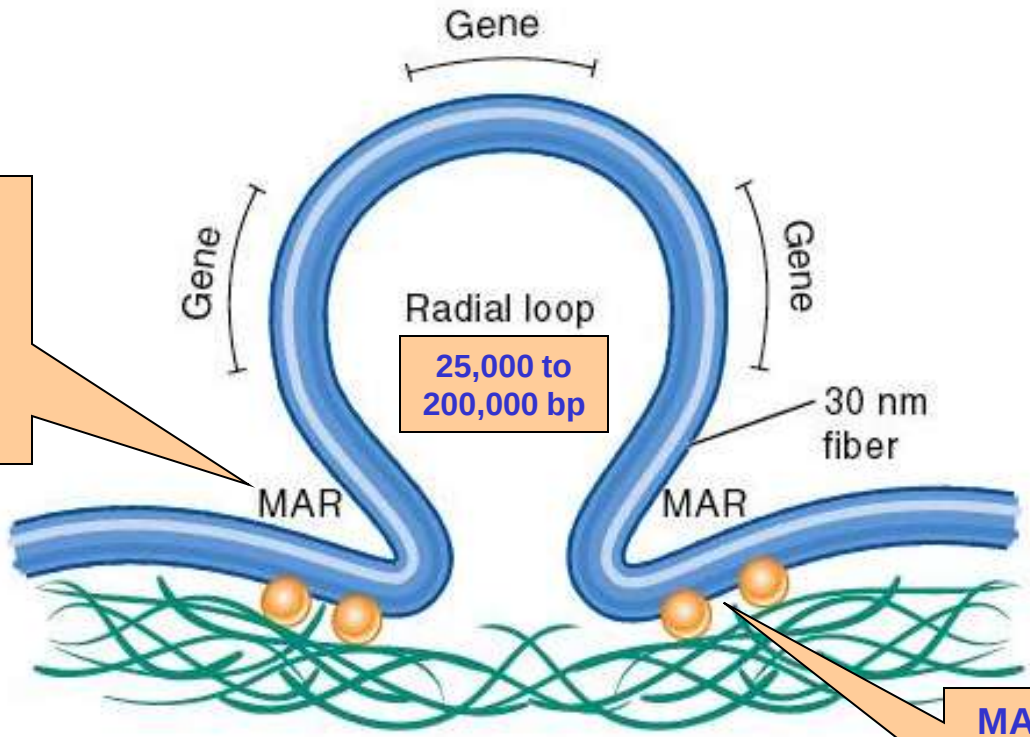


Figure 10.18 (a)

- The third mechanism of DNA compaction involves the formation of radial loop domains



Matrix-attachment regions

or

Scaffold-attachment regions (SARs)

MARs are anchored to the nuclear matrix, thus creating radial loops

Figure 10.18 (d)

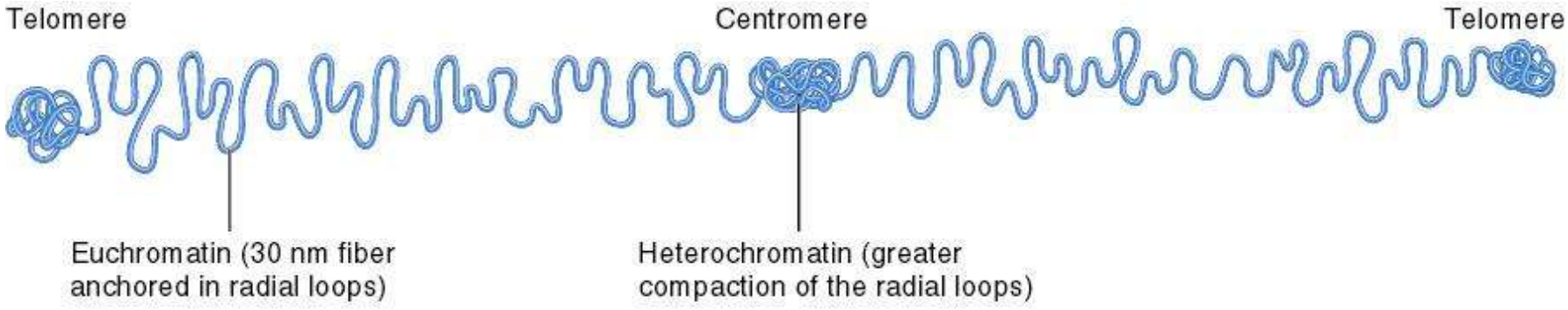
Further Compaction of the Chromosome

- The attachment of radial loops to the nuclear matrix is important in two ways
  - 1. It plays a role in gene regulation
    - Discussed in Chapter 15
  - 2. It serves to organize the chromosomes within the nucleus
    - Each chromosome in the nucleus is located in a discrete and nonoverlapping chromosome territory
    - Refer to Figure 10.19

Heterochromatin vs Euchromatin

- The compaction level of interphase chromosomes is not completely uniform
  - Euchromatin
    - Less condensed regions of chromosomes
    - Transcriptionally active
    - Regions where 30 nm fiber forms radial loop domains
  - Heterochromatin
    - Tightly compacted regions of chromosomes
    - Transcriptionally inactive (in general)
    - Radial loop domains compacted even further

Figure 10.20



- There are two types of heterochromatin
  - Constitutive heterochromatin
    - Regions that are always heterochromatic
    - Permanently inactive with regard to transcription
  - Facultative heterochromatin
    - Regions that can interconvert between euchromatin and heterochromatin
    - Example: Barr body

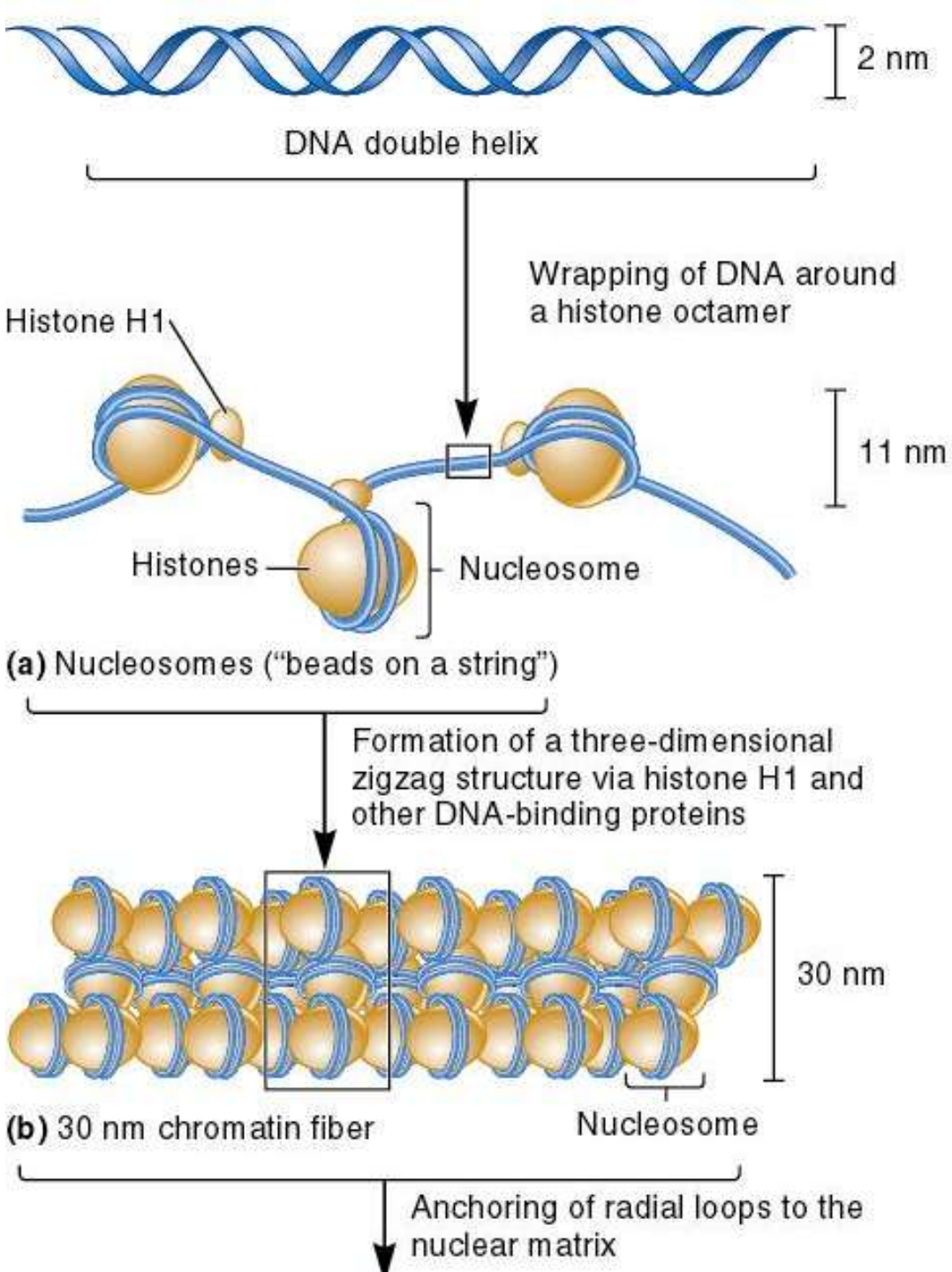

2 nm
DNA double helix
Wrapping of DNA around a histone octamer
Histone H1
11 nm
Histones
Nucleosome
(a) Nucleosomes ("beads on a string")
Formation of a three-dimensional zigzag structure via histone H1 and other DNA-binding proteins
30 nm
(b) 30 nm chromatin fiber
Nucleosome
Anchoring of radial loops to the nuclear matrix

Figure 10.21

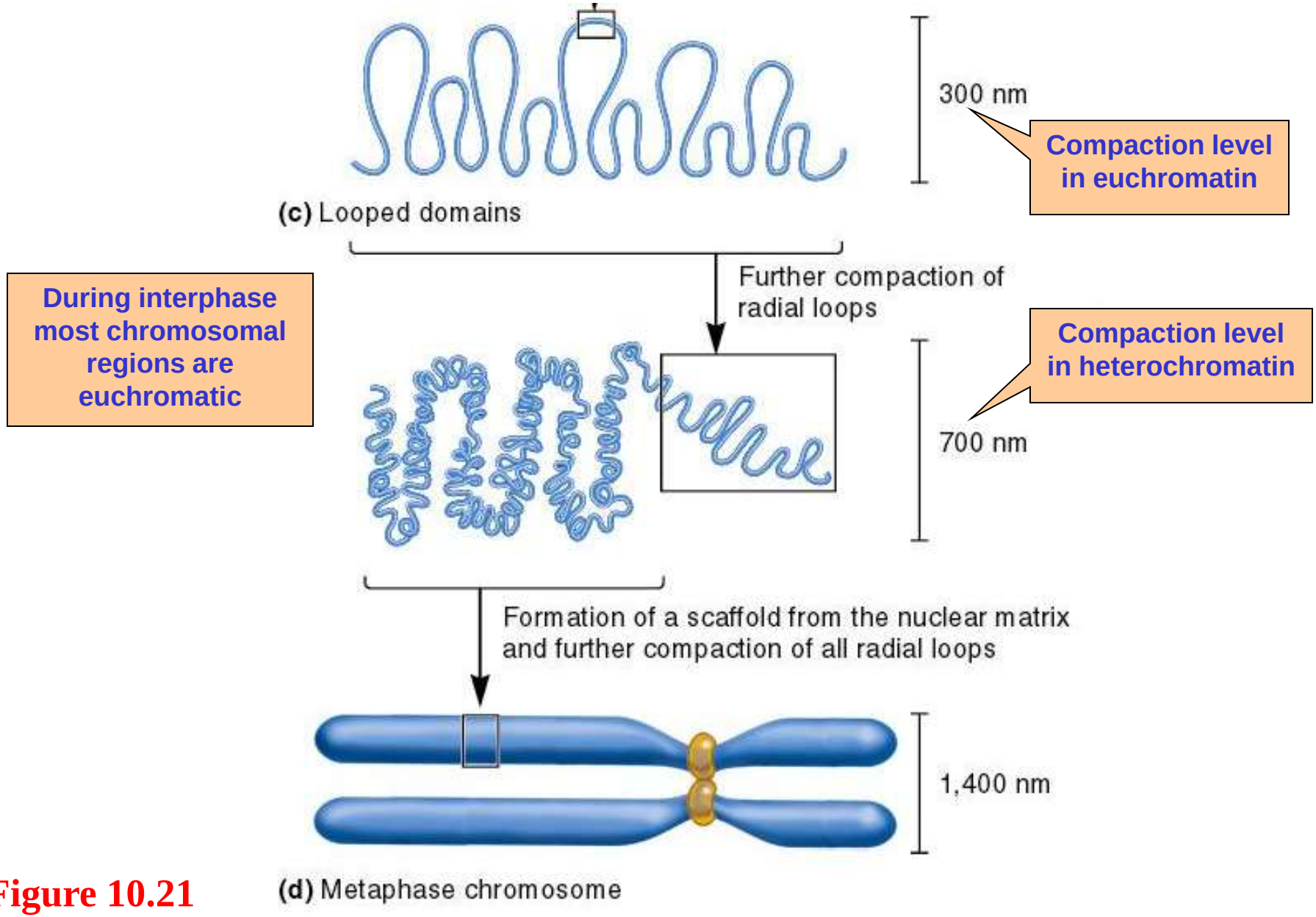

300 nm
Compaction level in euchromatin
(c) Looped domains
Further compaction of radial loops
During interphase most chromosomal regions are euchromatic
Compaction level in heterochromatin
700 nm
Formation of a scaffold from the nuclear matrix and further compaction of all radial loops
1,400 nm
(d) Metaphase chromosome

Figure 10.21

METAPHASE CHROMOSOMES

- As cells enter M phase, the level of compaction changes dramatically
  - By the end of prophase, sister chromatids are entirely heterochromatic
  - Two parallel chromatids have an overall diameter of 1,400 nm

- These highly condensed metaphase chromosomes undergo little gene transcription

Metaphase Chromosomes

- In metaphase chromosomes the radial loops are highly compacted and stay anchored to a scaffold
  - The scaffold is formed from the nuclear matrix
- Histones are needed for the compaction of radial loops
- Refer to Figure 10.22

- Two multiprotein complexes help to form and organize metaphase chromosomes
  - Condensin
    - Plays a critical role in chromosome condensation
  - Cohesin
    - Plays a critical role in sister chromatid alignment
- Both contain a category of proteins called SMC proteins
  - Acronym = Structural maintenance of chromosomes
  - SMC proteins use energy from ATP and catalyze changes in chromosome structure

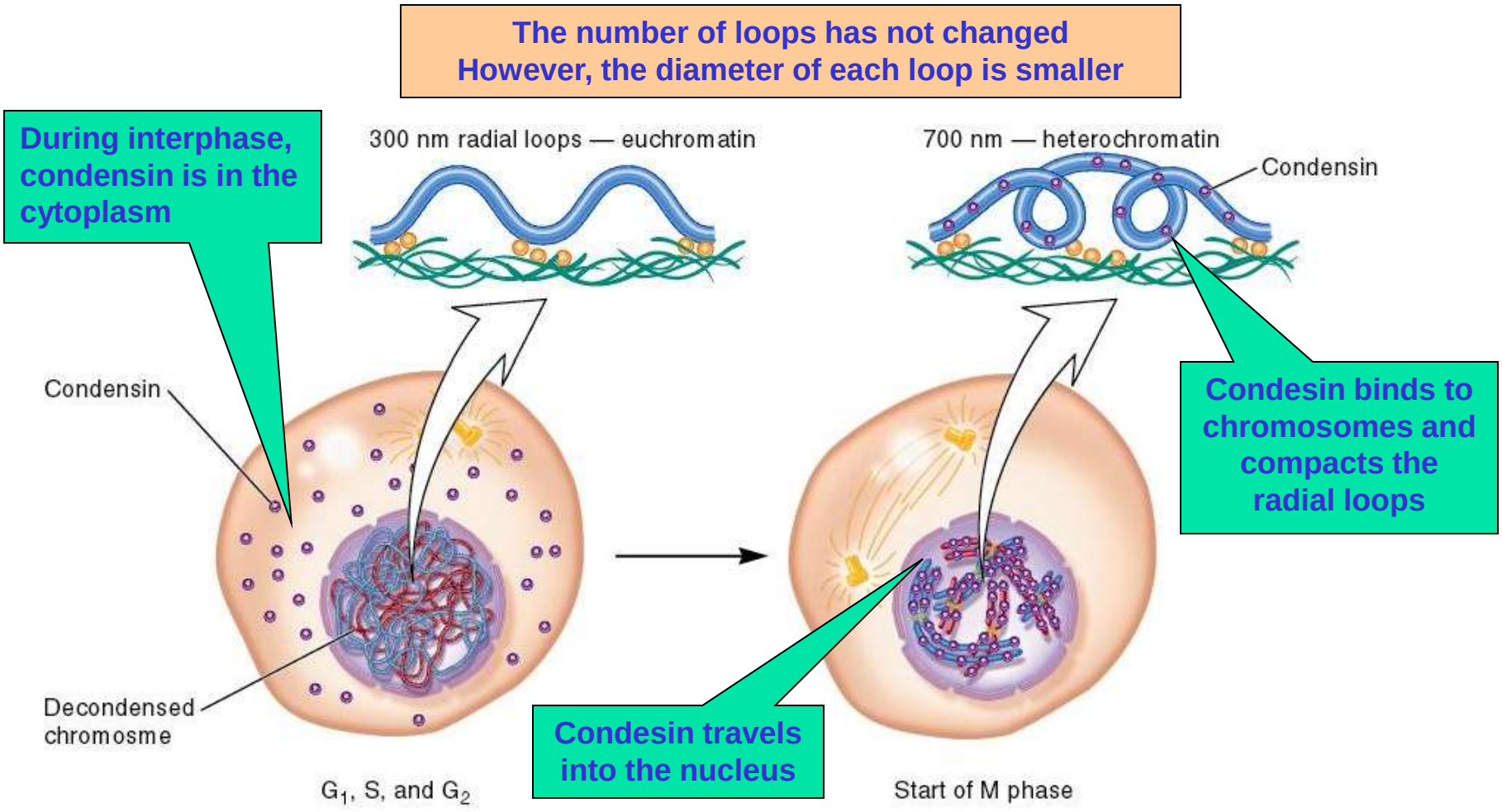


Figure 10.23 The condensation of a metaphase chromosome by condensin

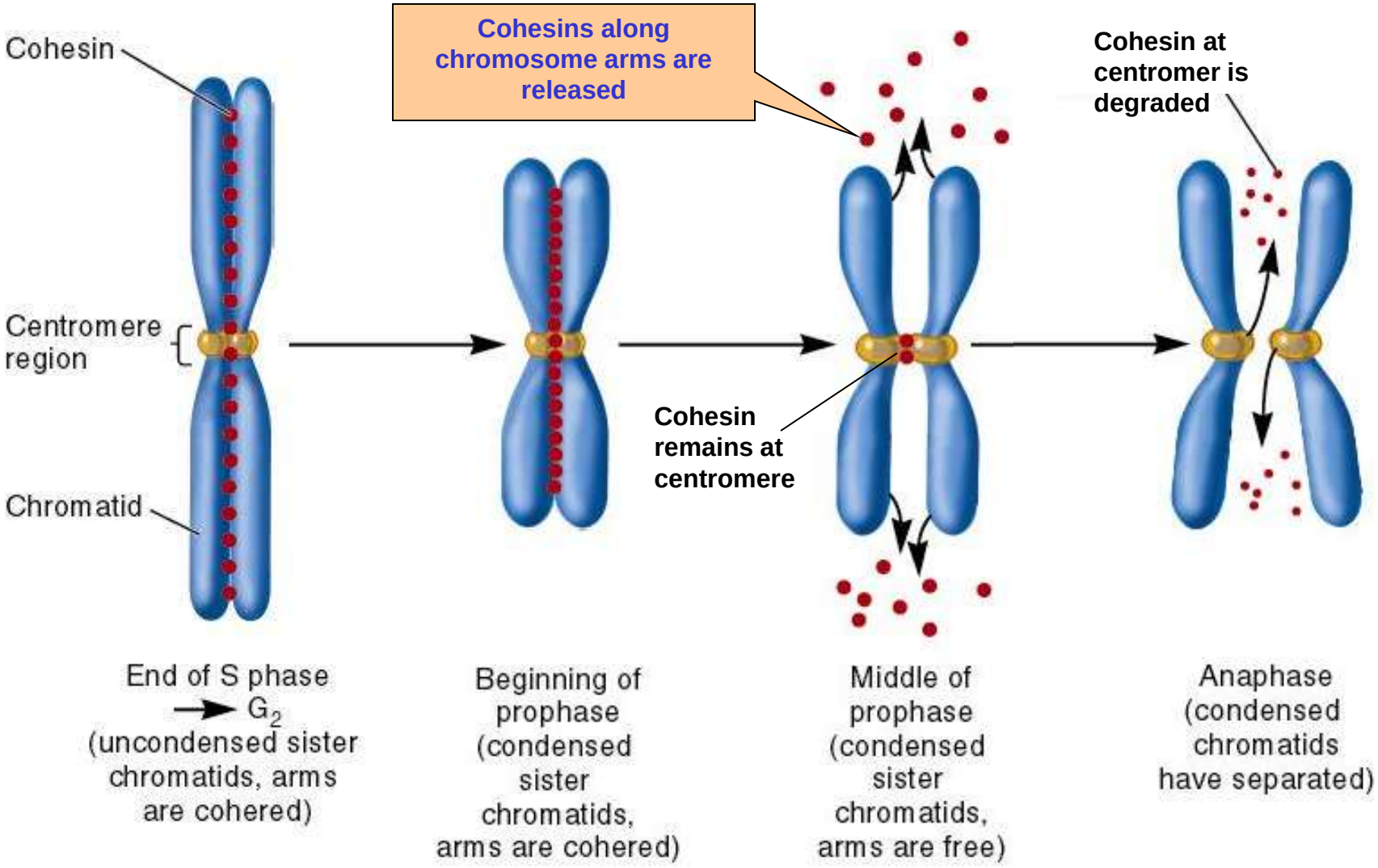


Figure 10.24 The alignment of sister chromatids via cohesin