UNIT 2
Biology For Engineers

DNA Synthesis Proceeds in a 5' to 3' Direction and Is Semidiscontinuous

A new strand of DNA is always synthesized in the 5' to 3' direction, with the free 3' OH as the point at which the DNA is elongated

Because the two DNA strands are antiparallel, the strand serving as the template is read from its 3' end toward its 5'end

Okazaki found that one of the new DNA strands is synthesized in short pieces, called **Okazaki fragments.**

This work ultimately led to the conclusion that one strand is synthesized continuously and the other discontinuously

The continuous strand, or **leading strand,** is the one in which 5' to 3 synthesis proceeds in the *same* direction as replication fork movement.

The discontinuous strand, or **lagging strand,** is the one in which 5n3 synthesis proceeds in the direction *opposite* to the direction of fork movement.

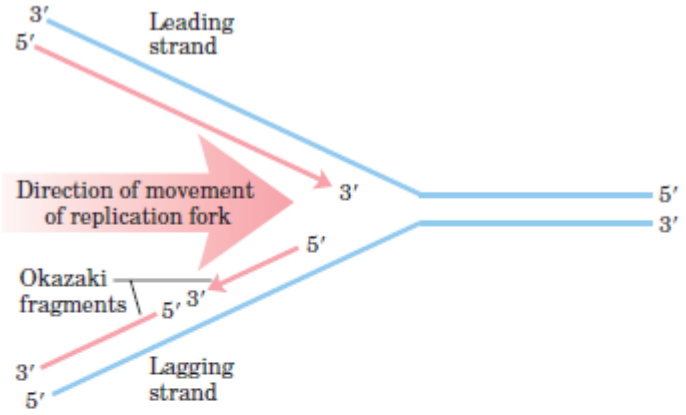


FIGURE 25–4 Defining DNA strands at the replication fork. A new DNA strand (red) is always synthesized in the 5′→3′ direction. The template is read in the opposite direction, 3′→5′. The leading strand is continuously synthesized in the direction taken by the replication fork. The other strand, the lagging strand, is synthesized discontinuously in short pieces (Okazaki fragments) in a direction opposite to that in which the replication fork moves. The Okazaki fragments are spliced together by DNA ligase. In bacteria, Okazaki fragments are ~1,000 to 2,000 nucleotides long. In eukaryotic cells, they are 150 to 200 nucleotides long.

DNA Is Degraded by Nucleases

Enzymes that degrade DNA are known as **nucleases,** or **DNases** if they are specific for DNA rather than RNA.

Every cell contains several different nucleases, belonging to two broad classes: **exonucleases and endonucleases.**

Exonucleases degrade nucleic acids from one end of the molecule. Many operate in only the 5’ to 3 or the 3n5 direction, removing nucleotides only from the 5 or the 3 end, respectively, of one strand of a doublestranded nucleic acid or of a single-stranded DNA.

Endonucleases can begin to degrade at specific internal sites in a nucleic acid strand or molecule, reducing it to smaller and smaller fragments.

A few exonucleases and endonucleases degrade only single-stranded DNA.

There are a few important classes of endonucleases that cleave only at specific nucleotide sequences (such as the restriction endonucleases that are so important in biotechnology).

DNA Polymerase

The fundamental reaction is a phosphoryl group transfer. The nucleophile is the 3'-hydroxyl group of the nucleotide at the 3' end of the growing strand.

Nucleophilic attack occurs at the phosphorus of the incoming deoxynucleoside 5'triphosphate. Inorganic pyrophosphate is released in the reaction. The general reaction is

$$\underset{\text{DNA}}{(\text{dNMP})_n} + \text{dNTP} \longrightarrow \underset{\substack{\text{Lengthened}\\ \text{DNA}}}{(\text{dNMP})_{n+1}} + \text{PP}_\text{i} \qquad (25\text{–}1)$$

where dNMP and dNTP are deoxynucleoside 5-monophosphate and 5-triphosphate

DNA polymerases synthesize only from 5' to 3

The two central requirements for DNA polymerization are as follows:

1. First, all DNA polymerases require a **template.** The polymerization reaction is guided by a template DNA strand according to the base-pairing rules predicted by Watson and Crick: where a guanine is present in the template, a cytosine deoxynucleotide is added to the new strand, and so on.
2. Second, the polymerases require a **primer.** A primer is a strand segment (complementary to the template) with a free 3-hydroxyl group to which a nucleotide can be added; the free 3 end of the primer is called the **primer terminus.**

In other words, part of the new strand must already be in place: all DNA polymerases can only add nucleotides to a preexisting strand. Most primers are oligonucleotides of RNA rather than DNA, and specialized enzymes synthesize primers when and where they are required.

Requirements for DNA polymerase activity

Template	[Basis for heredity]
dNTPs (not ATP, not NDPs, not NMPs)	[Building blocks]
Mg^{2+}	[Promotes reaction]
Primer - (complementary bases at 3' end, removed by fractionation and added back)	[DNA pol can't start!]

DNA Replication Requires Many Enzymes and Protein Factors

Replication in *E. coli* requires not just a single DNA polymerase but 20 or more different enzymes and proteins, each performing a specific task. The entire complex has been termed the **DNA replicase system** or **replisome.**

The Enzymes involved in this system are as follows:

Access to the DNA strands that are to act as templates requires separation of the two parent strands. This is generally accomplished by **helicases,** enzymes that move along the DNA and separate the strands, using chemical energy from ATP.

Strand separation creates topological stress in the helical DNA structure (see Fig. 24–12), which is relieved by the action of **topoisomerases.**

The separated strands are stabilized by **DNA-binding proteins**

As noted earlier, before DNA polymerases can begin synthesizing DNA, primers must be present on the template—generally short segments of RNA synthesized by enzymes known as **primases.**

Ultimately, the RNA primers are removed and replaced by DNA; in *E. coli,* this is one of the many functions of DNA polymerase I. After an RNA primer is removed and the gap is filled in with DNA, a nick remains in the DNA backbone in the form of a broken phosphodiester bond.

These nicks are sealed by **DNA ligases.**

Replication of the *E. coli* Chromosome Proceeds in Stages

Initiation The *E. coli* replication origin, *oriC,* consists of 245 bp; it bears DNA sequence elements that are highly conserved among bacterial replication origins.

The general arrangement of the conserved sequences is illustrated in Figure 25–11.

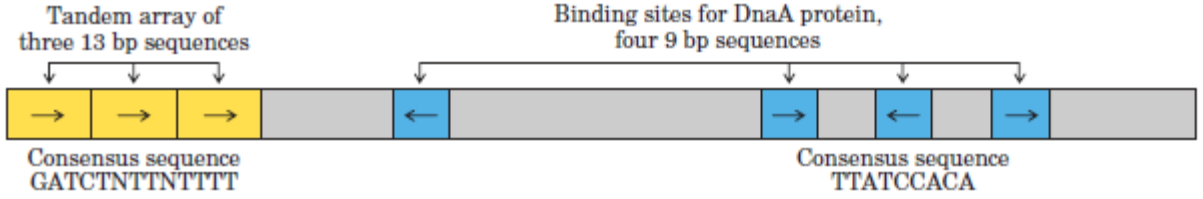


FIGURE 25–11 Arrangement of sequences in the *E. coli* replication origin, *oriC*. Although the repeated sequences (shaded in color) are not identical, certain nucleotides are particularly common in each position, forming a consensus sequence. In positions where there is no consensus, N represents any of the four nucleotides. The arrows indicate the orientations of the nucleotide sequences.

The key sequences of interest here are two series of short repeats: **three repeats of a 13 bp sequence and four repeats of a 9 bp sequence.**

At least nine different enzymes or proteins (summarized in Table 25–3) participate in the initiation phase of replication.

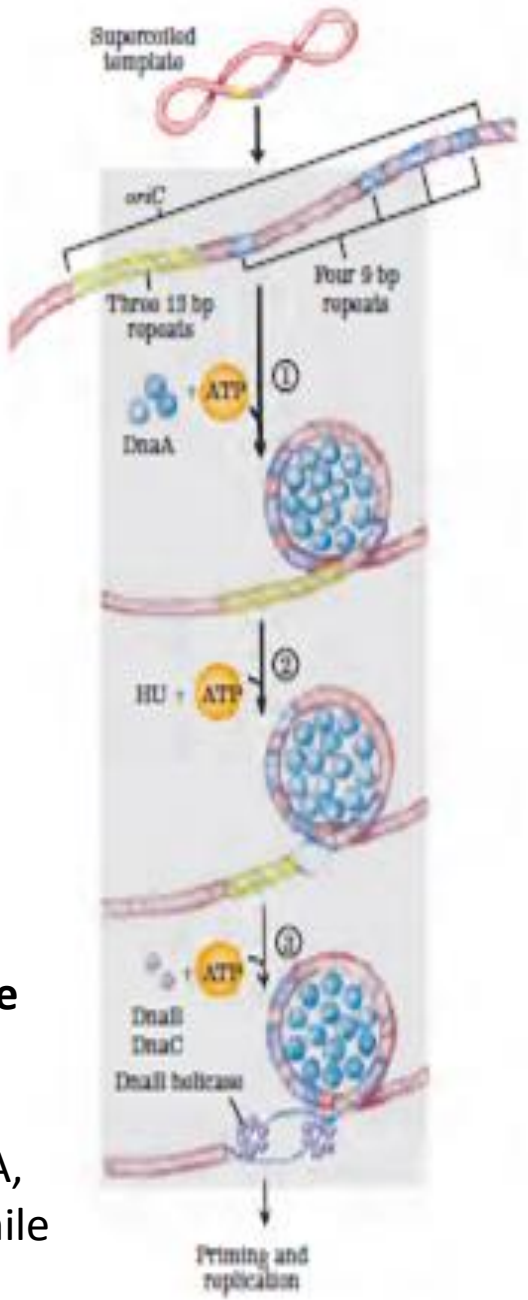
They open the DNA helix at the origin and establish a prepriming complex for subsequent reactions.

The crucial component in the initiation process is the DnaA protein.

Step 1: A single complex of four to five **DnaA protein** molecules binds to the four 9 bp repeats in the origin

Step 2:then recognizes and successively denatures the DNA in the region of the three 13 bp repeats, which are rich in AT pairs. This process requires ATP and the bacterial histone like **protein HU.**

Step **3:The DnaC** protein then loads the **DnaB** protein onto the unwound region. Two ringshaped hexamers of DnaB, one loaded onto each DNA strand, act as helicases, unwinding the DNA bidirectionally and creating two potential replication



Strand seperation is followed by binding of E. coli single-stranded DNA–binding **protein (SSB) and DNA gyrase (DNA topoisomerase II).** This enables unwinding of thousands of base pairs by the DnaB helicase, proceeding out from the origin. Many molecules of SSB bind cooperatively to singlestranded DNA, stabilizing the separated strands and preventing renaturation while gyrase relieves the topological stress produced by the DnaB helicase.

Elongation

Leading strand and lagging strand synthesis.

Several enzymes at the replication fork are important to the synthesis of both strands.

Parent DNA is first unwound by DNA helicases, and the resulting topological stress is relieved by topoisomerases.

Each separated strand is then stabilized by SSB. From this point, synthesis of leading and lagging strands is sharply different.

Leading strand synthesis, begins with the synthesis by **primase** (DnaB + DnaG protein) of a short (10 to 60 nucleotide) RNA primer at the replication origin. Deoxyribonucleotides are added to this primer by **DNA polymerase III.**

Leading strand synthesis then proceeds continuously, keeping pace with the unwinding of DNA at the replication fork

Elongation

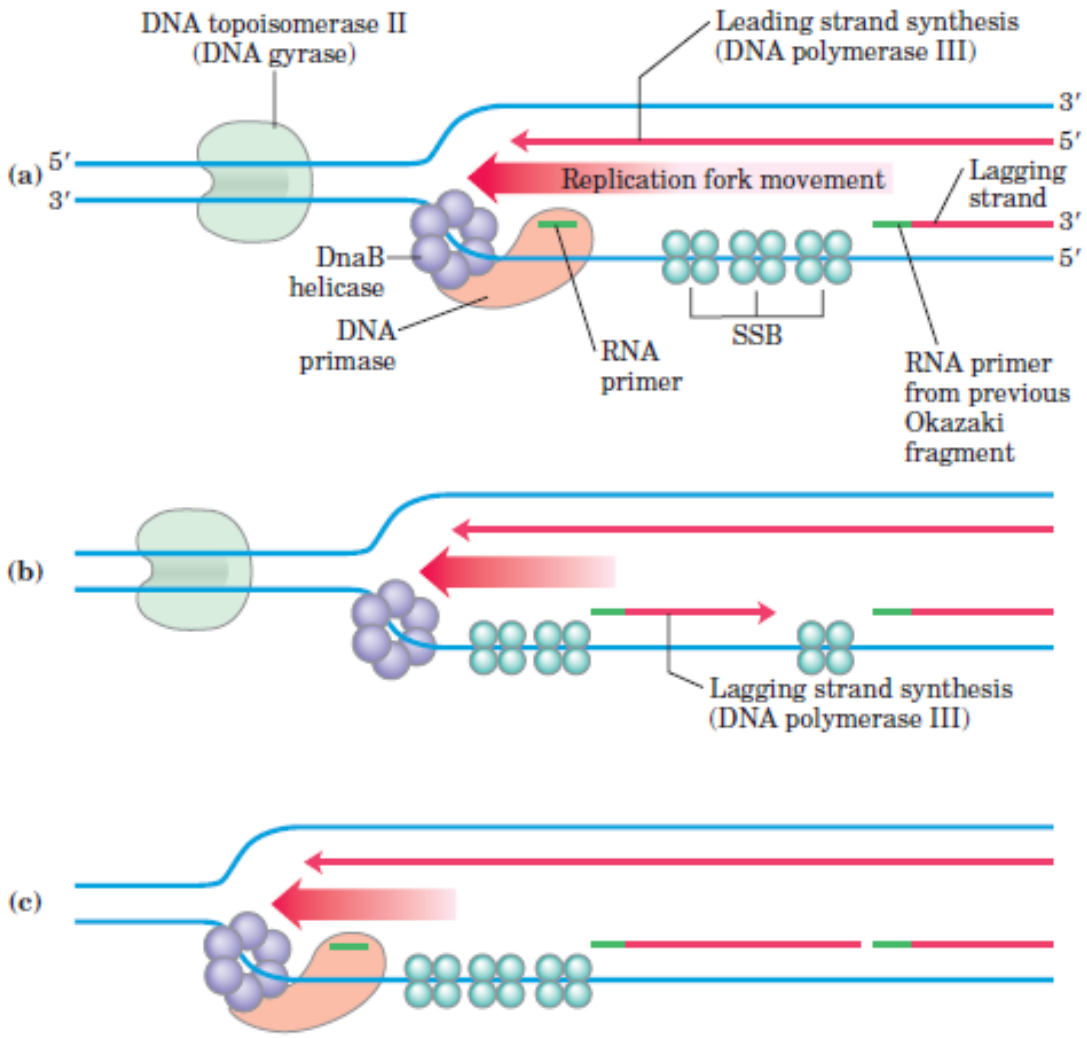


FIGURE 25–13 Synthesis of Okazaki fragments. **(a)** At intervals, primase synthesizes an RNA primer for a new Okazaki fragment. Note that if we consider the two template strands as lying side by side, lagging strand synthesis formally proceeds in the opposite direction from fork movement. **(b)** Each primer is extended by DNA polymerase III. **(c)** DNA synthesis continues until the fragment extends as far as the primer of the previously added Okazaki fragment. A new primer is synthesized near the replication fork to begin the process again.

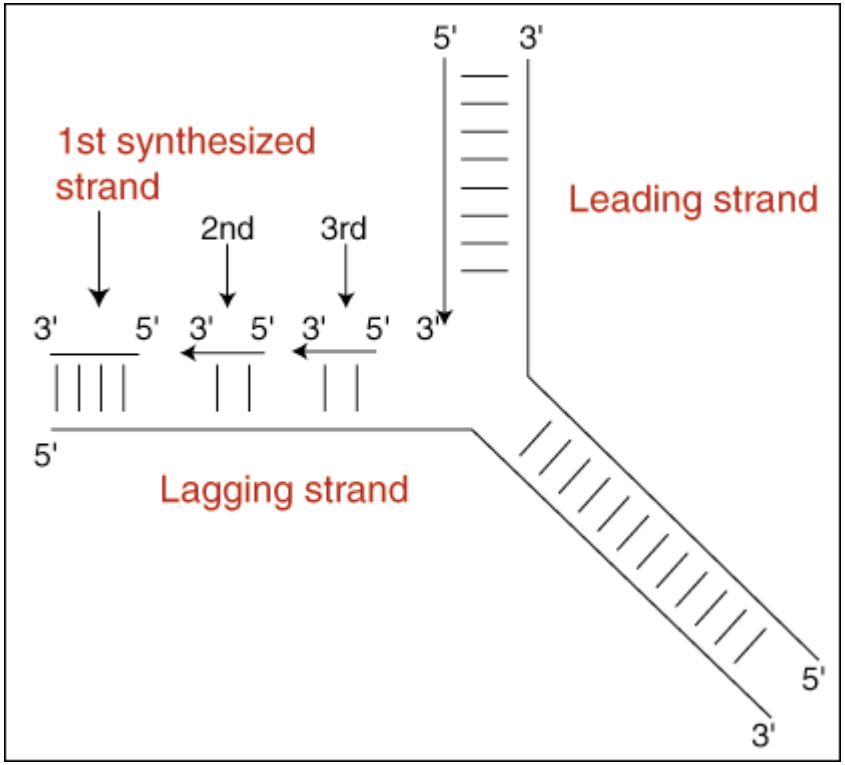
The lagging strand replicates in small segments, called Okazaki fragments.

These fragments are synthesized in the 5' to 3' direction away from the replication fork.

Yet while each individual segment is replicated away from the replication fork, each subsequent Okazaki fragment is replicated more closely to the receding replication fork than the fragment before.

These fragments are then stitched together by DNA ligase, creating a continuous strand. This type of replication is called *discontinuous*

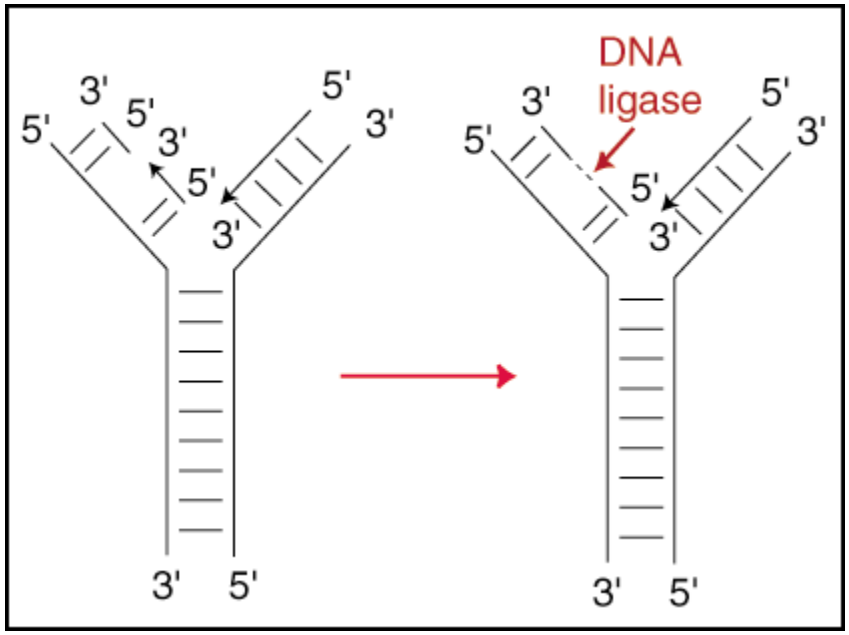
http://www.sparknotes.com/biology/molecular/dnareplicationandrepair/section1.rhtml



As you can see in the figure above, the first synthesized Okazaki fragment on the lagging strand is the furthest away from the replication fork, which is itself receding to the right.

Each subsequent Okazaki fragment starts at the replication fork and continues until it meets the previous fragment. The two fragments are then stitched together by DNA ligase

In figure , we can also see how replication on the lagging strand remains slightly behind that on the leading strand.



Because synthesis on the lagging strand takes place in a "backstitching" mechanism, its replication is slightly delayed in relation to synthesis on the leading strand.

The lagging strand must wait for a patch of the parent helix to open up a short distance in front of the newly synthesized strand before it can begin its synthesis back to the end of the daughter strand.

This "lag" time does not occur in the leading strand because it synthesizes the new strand by following right behind as the helix unwinds at the replication fork.

http://www.sparknotes.com/biology/molecular/dnareplicationandrepair/section1.rhtml

Another complication to replication on the lagging strand is the initiation of replication.

Whereas the **RNA primer on** the leading strand only has to trigger the initiation of the strand once, on the **lagging strand each individual Okazaki fragment must be triggered**.

On the lagging strand, then, an enzyme called primase that moves with the replication fork synthesizes numerous RNA primers, each of which triggers the growth of an Okazaki fragment.

The RNA primers are eventually removed leaving gaps that are filled by the replication machinery.

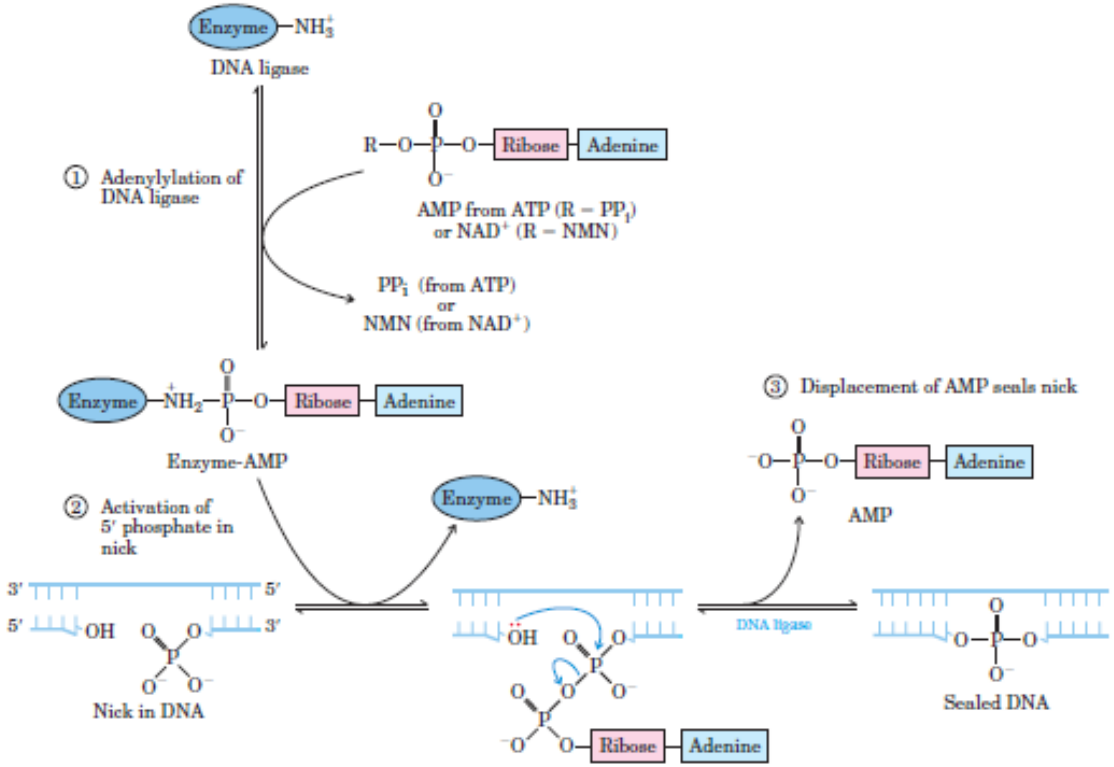
The entire complex responsible for coordinated DNA synthesis at a replication fork is a **replisome.**

DNA Ligases

DNA ligase catalyzes the formation of a phosphodiester bond between a 3 hydroxyl at the end of one DNA strand and a 5 phosphate at the end of another strand.

The phosphate must be activated by adenylylation. DNA ligases isolated from viruses and eukaryotes use ATP for this purpose. DNA ligases from bacteria generally use NAD—a cofactor
as the source of the AMP activating group .

DNA ligase is another enzyme of DNA metabolism that has become an important reagent in recombinant DI



Terminatio

Eventually, the two replication forks of the circular *E. coli* chromosome meet at a terminus region **containing multiple copies of a 20 bp sequence called Ter (for *ter*minus**)

The Ter sequences are arranged on the chromosome to create a sort of trap that a replication fork can enter but cannot leave.

The Ter sequences function as binding sites for a protein called Tus (*t*erminus *u*tilization *s*ubstance). The **Tus-Ter** complex can arrest a replication fork from only one direction.

Only one Tus-Ter complex functions per replication cycle

So, when either replication fork encounters a functional Tus-Ter complex, it halts; the other fork halts when it meets the first (arrested) fork.

The final few hundred base pairs of DNA between these large protein complexes are then replicated (by an as yet unknown mechanism), completing two topologically interlinked (catenated) circular chromosomes.

DNA circles linked in this way are known as **catenanes.** Separation of the catenated circles in *E. coli* requires topoisomerase IV (a type II topoisomerase). The separated chromosomes then segregate into daughter cells at cell division.

Terminatio n

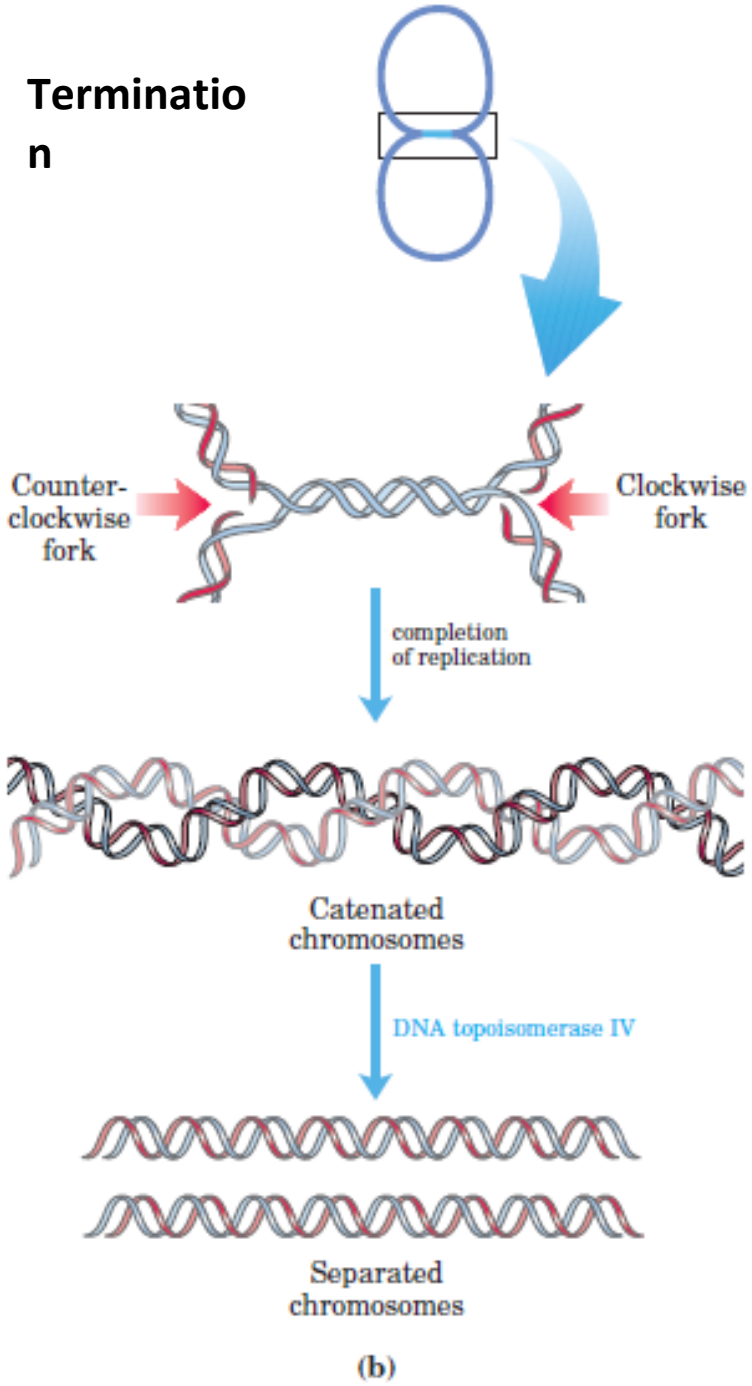


FIGURE 25–17 Termination of chromosome replication in *E. coli*. (a) The Ter sequences are positioned on the chromosome in two clusters with opposite orientations. (b) Replication of the DNA separating the opposing replication forks leaves the completed chromosomes joined as catenanes, or topologically interlinked circles. The circles are not covalently linked, but because they are interwound and each is covalently closed, they cannot be separated—except by the action of topoisomerases. In *E. coli*, a type II topoisomerase known as DNA topoisomerase IV plays the primary role in the separation of catenated chromosomes, transiently breaking both DNA strands of one chromosome and allowing the other chromosome to pass through the break.

Animation for replication

http://sites.fas.harvard.edu/~biotext/animations/replication1.swf

Transcription and Translation

Transcription

- Transcription resembles replication in its fundamental chemical mechanism, its polarity (direction of synthesis), and its use of a template.

- And like replication, transcription has initiation, elongation, and termination phases—though in the literature on transcription, initiation is further divided into discrete phases of DNA binding and initiation of RNA synthesis.

- Transcription differs from replication in that
  - it does not require a primer
  - generally, involves only limited segments of a DNA molecule.
  - Additionally, within transcribed segments only one DNA strand serves as a template

RNA Is Synthesized by RNA Polymerases

The discovery of DNA polymerase and its dependence on a DNA template spurred a search for an enzyme that synthesizes RNA complementary to a DNA strand. By 1960, four research groups had independently detected an enzyme in cellular extracts that could form an RNA polymer from ribonucleoside 5'-triphosphates. Subsequent work on the purified *Escherichia coli* RNA polymerase helped to define the fundamental properties of transcription (Fig. 26–1).

DNA-dependent RNA polymerase requires, in addition to a DNA template, all four ribonucleoside 5'-triphosphates (ATP, GTP, UTP, and CTP) as precursors of the nucleotide units of RNA, as well as Mg2+ . The protein also binds one Zn2+.

The chemistry and mechanism of RNA synthesis closely resemble those used by DNA polymerases.

RNA polymerase elongates an RNA strand by adding ribonucleotide units to the 3'-hydroxyl end, building RNA in the 5' to 3' direction. The 3'-hydroxyl group acts as a nucleophile, attacking the phosphate of the incoming ribonucleoside triphosphate (Fig. 26–1b) and releasing pyrophosphate. The overall reaction is

$$\underset{\text{RNA}}{(\text{NMP})_n} + \text{NTP} \longrightarrow \underset{\text{Lengthened RNA}}{(\text{NMP})_{n+1}} + \text{PP}_\text{i}$$

- **RNA polymerase requires DNA for activity** and is most active when bound to a double-stranded DNA.

- Only one of the two DNA strands serves as a template. **The template DNA strand is copied in the 3' to5' direction** (antiparallel to the new RNA strand), just as in DNA replication.

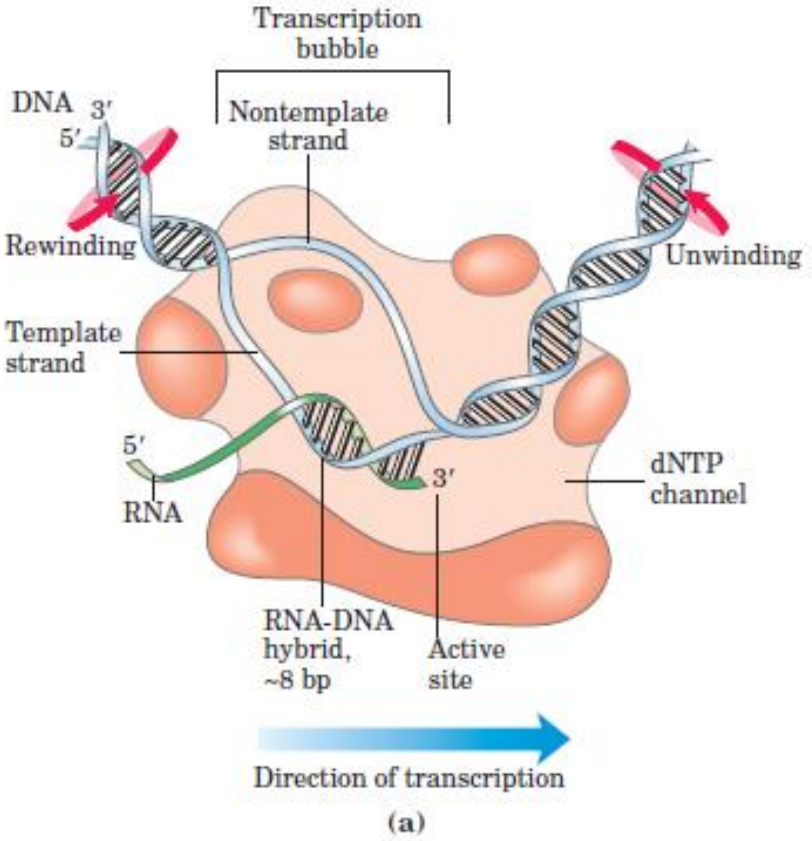
- Each nucleotide in the newly formed RNA is selected by Watson-Crick base-pairing interactions; U residues are inserted in the RNA to pair with A residues in the DNA template, G residues are inserted to pair with C residues, and so on. Base-pair geometry may also play a role in base selection.

- **Unlike DNA polymerase, RNA polymerase does not require a primer to initiate synthesis.** Initiation occurs when RNA polymerase binds at specific DNA sequences called **promoters.**

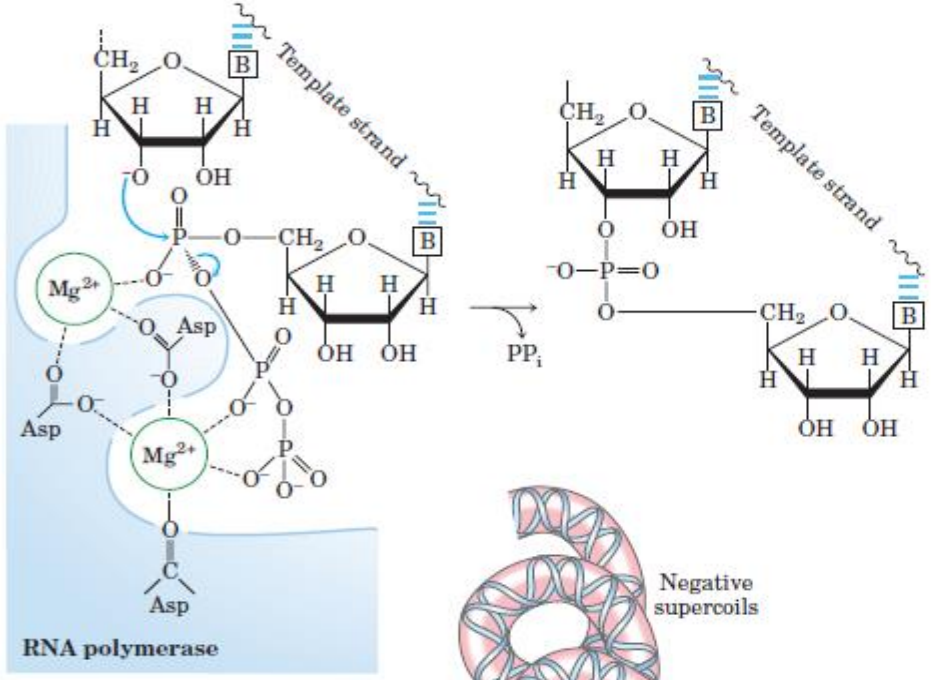
- **The 5'-triphosphate group of the first residue in a nascent (newly formed) RNA molecule is not cleaved to release PPi, but instead remains intact throughout the transcription process.**

- During the elongation phase of transcription, the growing end of the new RNA strand base-pairs temporarily with the **DNA template to form a short hybrid-RNA-DNA double helix, estimated to be 8 bp long** (Fig.26–1a). The RNA in this hybrid duplex "peels off" shortly after its formation, and the DNA duplex re-forms.

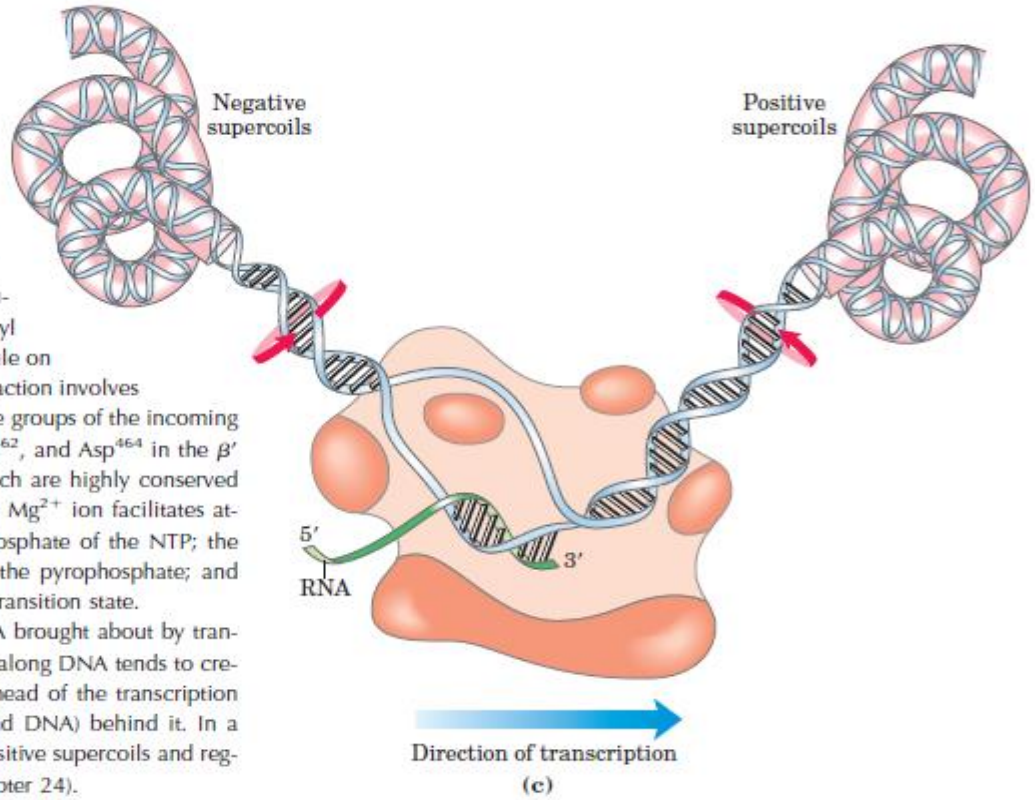
Transcription by RNA Polymerase in E coli



MECHANISM FIGURE 26–1 **Transcription by RNA polymerase in *E. coli*.** For synthesis of an RNA strand complementary to one of two DNA strands in a double helix, the DNA is transiently unwound. **(a)** About 17 bp are unwound at any given time. RNA polymerase and the bound transcription bubble move from left to right along the DNA as shown; facilitating RNA synthesis. The DNA is unwound ahead and rewound behind as RNA is transcribed. Red arrows show the direction in which the DNA must rotate to permit this process. As the DNA is rewound, the RNA-DNA hybrid is displaced and the RNA strand extruded. The RNA polymerase is in close contact with the DNA ahead of the transcription bubble, as well as with the separated DNA strands and the RNA within and immediately behind the bubble. A channel in the protein funnels new nucleoside triphosphates (NTPs) to the polymerase active site. The polymerase footprint encompasses about 35 bp of DNA during elongation.



(b) Catalytic mechanism of RNA synthesis by RNA polymerase. Note that this is essentially the same mechanism used by DNA polymerases (see Fig. 25–5b). The addition of nucleotides involves an attack by the 3′-hydroxyl group at the end of the growing RNA molecule on the α phosphate of the incoming NTP. The reaction involves two Mg^{2+} ions, coordinated to the phosphate groups of the incoming NTP and to three Asp residues (Asp^{460}, Asp^{462}, and Asp^{464} in the β' subunit of the *E. coli* RNA polymerase), which are highly conserved in the RNA polymerases of all species. One Mg^{2+} ion facilitates attack by the 3′-hydroxyl group on the α phosphate of the NTP; the other Mg^{2+} ion facilitates displacement of the pyrophosphate; and both metal ions stabilize the pentacovalent transition state.



(c) Changes in the supercoiling of DNA brought about by transcription. Movement of an RNA polymerase along DNA tends to create positive supercoils (overwound DNA) ahead of the transcription bubble and negative supercoils (underwound DNA) behind it. In a cell, topoisomerases rapidly eliminate the positive supercoils and regulate the level of negative supercoiling (Chapter 24).

- To enable RNA polymerase to synthesize an RNA strand complementary to one of the DNA strands, the DNA duplex must unwind over a short distance, forming a transcription “bubble.”

- During transcription, the *E. coli* RNA polymerase generally keeps about 17 bp unwound.

- The 8 bp RNA-DNA hybrid occurs in this unwound region. Elongation of a transcript by *E. coli* RNA polymerase proceeds at a rate **of 50 to 90 nucleotides/s.**

- Because DNA is a helix, movement of a transcription bubble requires considerable strand rotation of the nucleic acid molecules.

- DNA strand rotation is restricted in most DNAs by DNA-binding proteins and other structural barriers. As a result, a moving RNA polymerase generates waves of positive supercoils ahead of the transcription bubble and negative supercoils behind (Fig.26–1c).

The two complementary DNA strands have different roles in transcription. The strand that serves as template for RNA synthesis is called the **template strand.**

The DNA strand complementary to the template, the **nontemplate strand,** or **coding strand,** is identical in base sequence to the RNA transcribed from the gene, with U in the RNA in place of T in the DNA (Fig. 26–2).

The coding strand for a particular gene may be located in either strand of a given chromosome (as shown in Fig. 26–3 for a virus). The regulatory sequences that control transcription are by convention designated by the sequences in the coding strand.

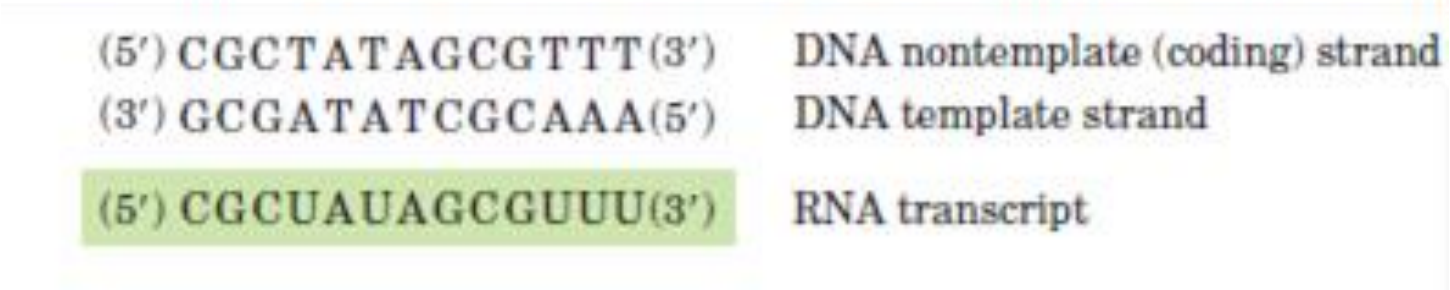


FIGURE 26–2 Template and nontemplate (coding) DNA strands. The two complementary strands of DNA are defined by their function in transcription. The RNA transcript is synthesized on the template strand and is identical in sequence (with U in place of T) to the nontemplate strand, or coding strand.

DNA dependent RNA Polymerases

The DNA-dependent RNA polymerase of *E. coli* is a large, complex enzyme with **five core subunits** ($\alpha_2\beta\beta'\omega$; *M*r 390,000) **and a sixth subunit**, one of a group designated σ , with variants designated by size (molecular weight).

The σ **subunit binds transiently to the core and directs the enzyme to specific binding sites on the DNA.** These six subunits constitute the RNA polymerase holoenzyme .

The RNA polymerase holoenzyme of *E. coli* thus exists in several forms, depending on the type of σ subunit. The most common subunit is σ^{70} (*M*r 70,000).

RNA polymerases lack a separate proofreading 3' to 5' exonuclease active site (such as that of many DNA polymerases), and the error rate for transcription is higher than that for chromosomal DNA replication— approximately one error for every 104 to 105 ribonucleotides incorporated into RNA.

- Because many copies of an RNA are generally produced from a single gene and all RNAs are eventually degraded and replaced, a mistake in an RNA molecule is of less consequence to the cell than a mistake in the permanent information stored in DNA.

- Many RNA polymerases, including bacterial RNA polymerase and the eukaryotic RNA polymerase II (discussed below), do pause when a mispaired base is added during transcription, and they can remove mismatched nucleotides from the 3’ end of a transcript by direct reversal of the polymerase reaction.

- But it is not yet know whether this activity is a true proofreading function and to what extent it may contribute to the fidelity of transcription

RNA Synthesis Begins at Promoters

Initiation of RNA synthesis at random points in a DNA molecule would be an extraordinarily wasteful process.

Instead, an RNA polymerase binds to specific sequences in the DNA called **promoters, which direct the transcription of adjacent segments of DNA (genes**).

The sequences where RNA polymerases bind can be quite variable, and much research has focused on identifying the particular sequences that are critical to promoter function.

In *E. coli,* RNA polymerase binding occurs within a region stretching from about 70 bp before the transcription start site to about 30 bp beyond it.

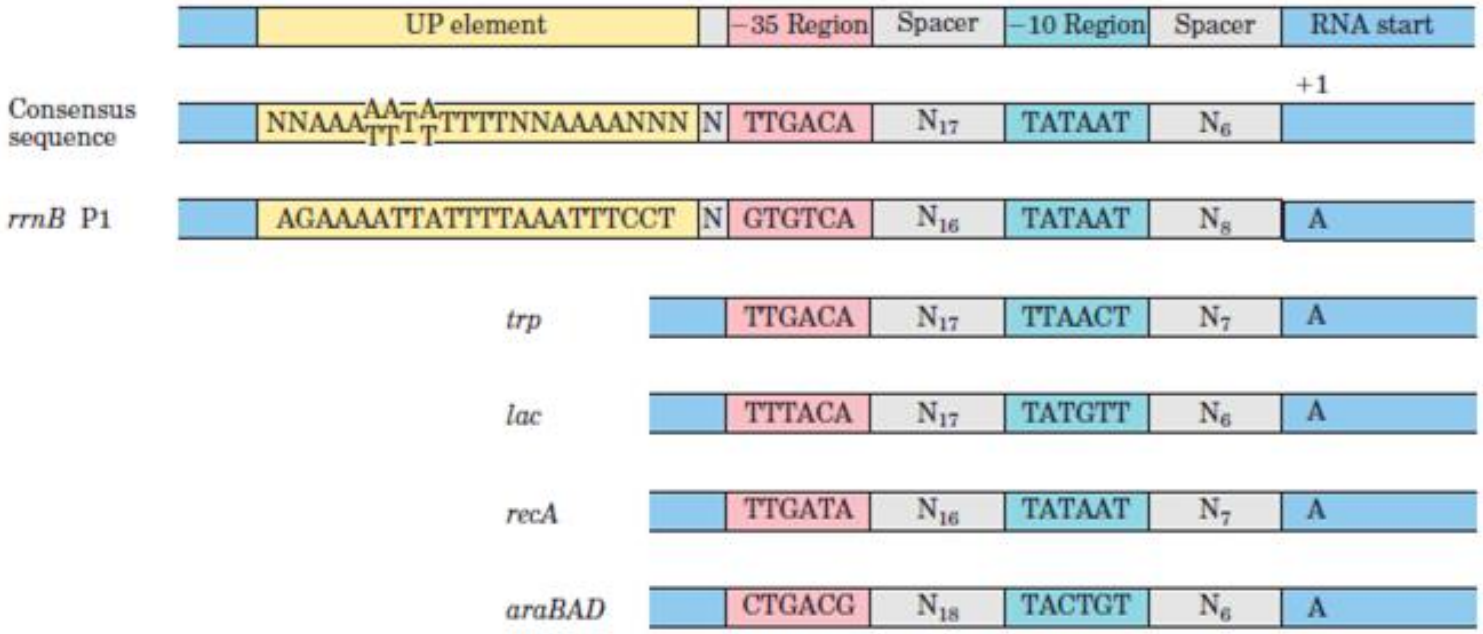
By convention, the DNA base pairs that correspond to the beginning of an RNA molecule are given positive numbers, and those preceding the RNA start site are given negative numbers.

The promoter region thus extends between positions -70 and +30

E Coli Promoters

Analyses and comparisons of the most common class of bacterial promoters (those recognized by an RNA polymerase holoenzyme containing σ^{70}) **have revealed similarities in two short sequences centered about positions -10 and -35** (Fig. 26–5).

These sequences are important interaction sites for the σ^{70} subunit. Although the sequences are not identical for all bacterial promoters in this class, certain nucleotides that are particularly common at each position form a **consensus sequence**



- The consensus sequence at **the -10 region is (5')TATAAT(3'); the consensus sequence at the -35 region is (5')TTGACA(3'**).

- A **third AT-rich recognition element, called the UP (upstream promoter**) element, occurs between positions -**40 and -60 in the promoters** of certain highly expressed genes.

- The UP element is bound by **the α subunit of RNA polymerase**. The efficiency with which an RNA polymerase binds to a promoter and initiates transcription is determined **in large measure by these sequences, the spacing between them, and their distance from the transcription start site.**

- Many independent lines of evidence attest to the functional importance of the sequences in the -35 and -10 regions. Mutations that affect the function of a given promoter often involve a base pair in these regions.

Variations in the consensus sequence also affect the efficiency of RNA polymerase binding and transcription initiation. A change in only one base pair can decrease the rate of binding by several orders of magnitude.

The promoter sequence thus establishes a basal level of expression that can vary greatly from one *E. coli* gene to the next

- The pathway of transcription initiation consists of two major parts, binding and initiation, each with multiple steps.

- **First, the polymerase binds to the promoter, forming, in succession, a closed complex (in which the bound DNA is intact) and an open complex** (in which the bound DNA is intact and partially unwound near the 10 sequence).

- **Second, transcription is initiated within the complex, leading to a conformational change that converts the complex to the elongation form, followed by movement of the transcription complex away from the promoter** (promoter clearance).

- Any of these steps can be affected by the specific makeup of the promoter sequences. The subunit sigma dissociates as the polymerase enters the elongation phase of transcription (Fig. 26–6a).

- *E. coli* has other classes of promoters, bound by RNA polymerase holoenzymes with different σ subunits. An example is the promoters of the heat-shock genes. The products of this set of genes are made at higher levels when the cell has received an insult, such as a sudden increase in temperature. RNA polymerase binds to the promoters of these genes only when σ^{70} is replaced with the σ^{32} (*M*r 32,000) subunit, which is specific for the heat-shock promoters (see Fig. 28–3). By using different subunits the cell can coordinate the expression of sets of genes, permitting major changes in cell physiology.

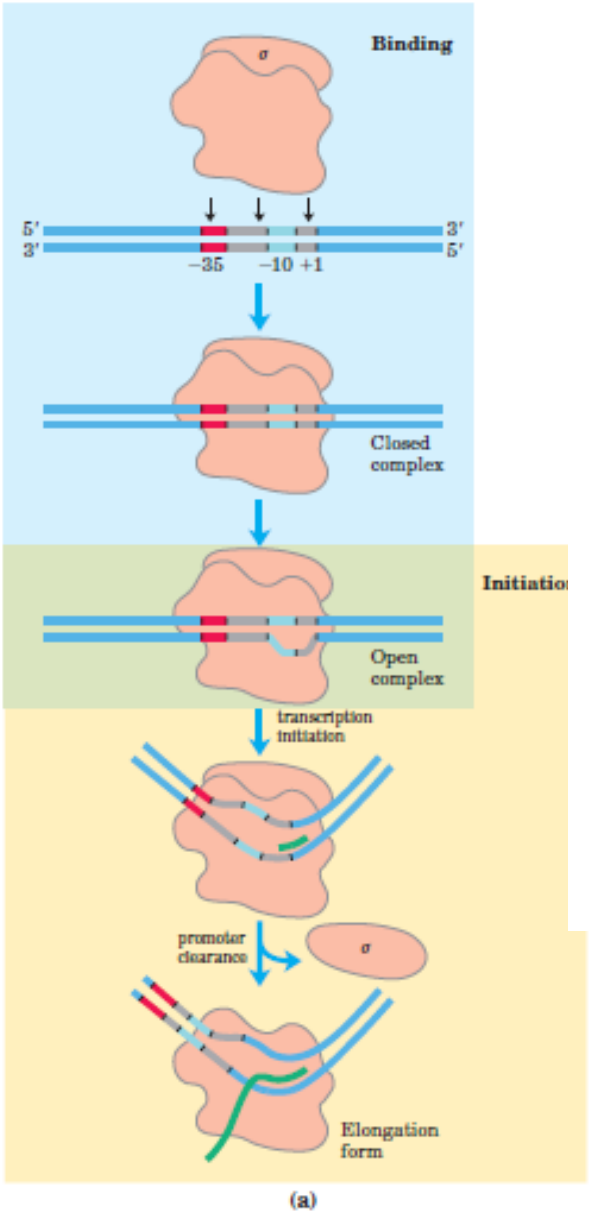


FIGURE 26–6 Transcription initiation and elongation by *E. coli* RNA polymerase. (a) Initiation of transcription requires several steps generally divided into two phases, binding and initiation. In the binding phase, the initial interaction of the RNA polymerase with the promoter leads to formation of a closed complex, in which the promoter DNA is stably bound but not unwound. A 12 to 15 bp region of DNA—from within the −10 region to position +2 or +3—is then unwound to form an open complex. Additional intermediates (not shown) have been detected in the pathways leading to the closed and open complexes, along with several changes in protein conformation. The initiation phase encompasses transcription initiation and promoter clearance. Once the first 8 or 9 nucleotides of a new RNA are synthesized, the σ subunit is released and the polymerase leaves the promoter and becomes committed to elongation of the RNA.

Transcription Is Regulated at Several Levels

- Requirements for any gene product vary with cellular conditions or developmental stage, and transcription of each gene is carefully regulated to form gene products only in the proportions needed.
- Regulation can occur at any step in transcription, including elongation and termination.
- However, much of the regulation is directed at the polymerase binding and transcription initiation steps outlined in Figure 26–6.
- Differences in promoter sequences are just one of several levels of control.
- The binding of proteins to sequences both near to and distant from the promoter can also affect levels of gene expression.

.

- Protein binding **can *activate* transcription by facilitating either RNA polymerase binding or steps further along in the initiation process, or it can *repress* transcription by blocking the activity of the polymerase.**

- In *E. coli,* one protein that activates transcription is the **cAMP receptor protein (CRP), which increases the transcription of genes coding for enzymes that metabolize sugars other than glucose when cells are grown in the absence of glucose.**

- **Repressors** are proteins that block the synthesis of RNA at specific genes. In the case of the Lac repressor , transcription of the genes for the enzymes of lactose metabolism is blocked when lactose is unavailable

Specific Sequences Signal Termination of RNA Synthesis

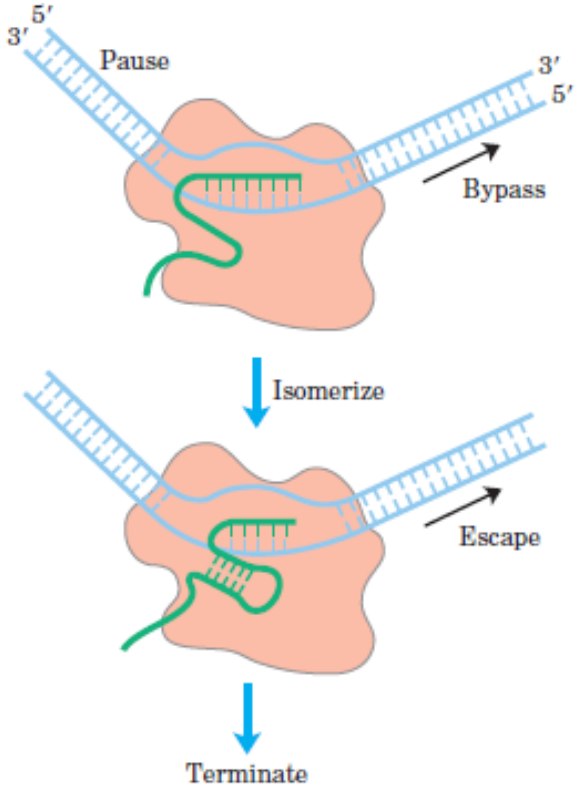


FIGURE 26–7 Model for ρ-independent termination of transcription in *E. coli*. RNA polymerase pauses at a variety of DNA sequences, some of which are terminators. One of two outcomes is then possible: the polymerase bypasses the site and continues on its way, or the complex undergoes a conformational change (isomerization). In the latter case, intramolecular pairing of complementary sequences in the newly formed RNA transcript may form a hairpin that disrupts the RNA-DNA hybrid and/or the interactions between the RNA and the polymerase, resulting in isomerization. An A═U hybrid region at the 3′ end of the new transcript is relatively unstable, and the RNA dissociates completely, leading to termination and dissociation of the RNA molecule. This is the usual outcome at terminators. At other pause sites, the complex may escape after the isomerization step to continue RNA synthesis.

RNA synthesis is processive (that is, the RNA polymerase has high processivity)—necessarily so, because if an RNA polymerase released an RNA transcript prematurely, it could not resume synthesis of the same RNA but instead would have to start over.

However, an encounter with certain DNA sequences results in a pause in RNA synthesis, and at some of these sequences transcription is terminated.

E. coli has at least two classes of termination signals: one class relies on a protein factor called ρ (rho) and the other is rho independent.

- Most **rho -independent** terminators have **two d**istinguishing features.

✓ The first is **a region that produces an RNA transcript with self-complementary sequences**, permitting the formation of a **hairpin structure** centered 15 to 20 nucleotides before the projected end of the RNA strand.

✓ The second feature is a highly conserved string of **three A residues** in the **template strand that are transcribed into U residues near the 3' end of the hairpin.**

When a polymerase arrives at a termination site with this structure, it pauses (Fig. 26–7).

Formation of the hairpin structure in the RNA disrupts several AUU base pairs in the RNA-DNA hybrid segment and may disrupt important interactions between RNA and the RNA polymerase, facilitating dissociation of the transcript.

- The rho-dependent terminators lack the sequence of repeated A residues in the template strand but usually include a **CA-rich** sequence called **a *rut* (*r*ho *ut*ilization)** element.

- The **ρ protein associates** with **the RNA** at specific binding sites and migrates in the **5’ to 3’ direction until it reaches the transcription complex that is paused at a termination site**. Here it contributes to release of the RNA transcript.

- The ρ **protein has an ATP-dependent RNA-DNA helicase activity that promotes translocation of the protein along the RNA, and ATP is hydrolyzed by protein during the termination process.**

- The detailed mechanism by which the protein promotes the release of the RNA transcript is not known

Animation of prokaryotic transcription

http://highered.mheducation.com/sites/dl/free/0072835125/126997/animation21.html

http://www.dnalc.org/resources/3d/16-translation-advanced.html

RNA Processing

- Many of the RNA molecules in bacteria and virtually all RNA molecules in eukaryotes are processed to some degree after synthesis.

- Several of the enzymes that catalyze these reactions consist of RNA (ribozymes) rather than protein.

- A newly synthesized RNA molecule is called a **primary transcript.**

- Perhaps the most extensive processing of primary transcripts occurs in **eukaryotic mRNAs** and in **tRNAs of both bacteria and eukaryotes**.

- Noncoding tracts that break up the coding region of the transcript are called introns, and the coding segments are called exons.

- In a process called **splicing,** the introns are removed from the primary transcript and the exons are joined to form a continuous sequence that specifies a functional polypeptide.

- Eukaryotic mRNAs are also modified at each end.
- A modified residue called a **5' cap is added at the 5' end.**
- The **3' end is cleaved, and 80 to 250** A residues are added to create a poly(A) "tail."
- The sometimes **elaborate protein complexes that carry out each of these three mRNA-processing reactions do not operate independently**.
- **They appear to be organized in association with each other and with the phosphorylated CTD of Pol II;** each complex affects the function of the others.
- Other proteins involved in mRNA transport to the **cytoplasm are also associated with the mRNA in the nucleus, and the processing of the transcript is coupled to its transport**.
- The composition of the complex changes as the primary transcript is processed, transported to the cytoplasm, and delivered to the ribosome for translation.

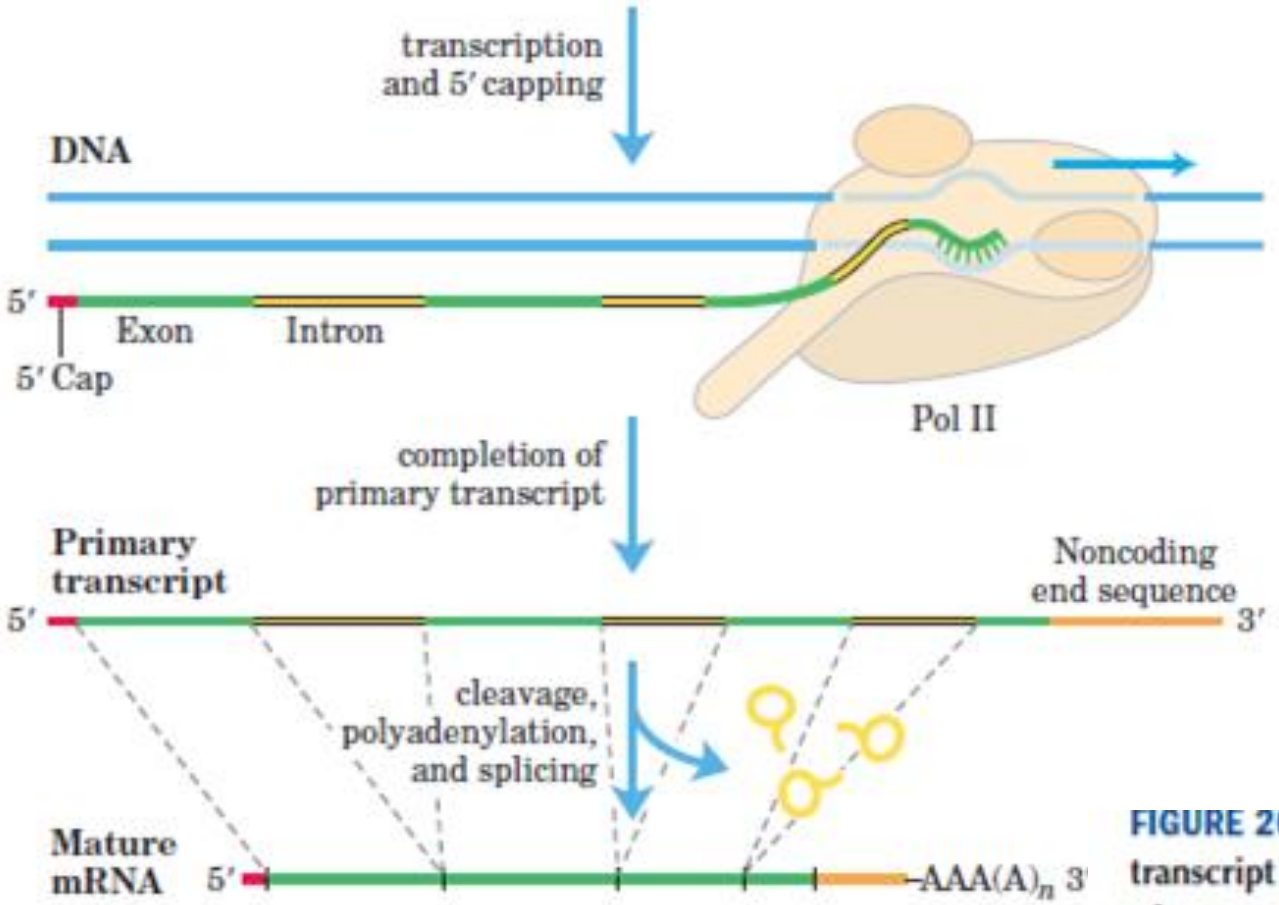


FIGURE 26–11 Formation of the primary transcript and its processing during maturation of mRNA in a eukaryotic cell. The 5′ cap (red) is added before synthesis of the primary transcript is complete. A noncoding sequence following the last exon is shown in orange. Splicing can occur either before or after the cleavage and polyadenylation steps. All the processes shown here take place within the nucleus.

- Proteins are the end products of most information pathways.

- A typical cell requires thousands of different proteins at any given moment. These must be **synthesized in response to the cell's current needs, transported (targeted) to their appropriate cellular locations, and degraded when no longer needed**.

- An understanding of protein synthesis, the most complex biosynthetic process, has been one of the greatest challenges in biochemistry.

- Eukaryotic protein synthesis involves more than **70 different ribosomal proteins**; **20 or more enzymes to activate the amino acid precursors; a dozen or more auxiliary enzymes and other protein factors for the initiation, elongation, and termination of polypeptides; perhaps 100 additional enzymes for the final processing of different proteins; and 40 or more kinds of transfer and ribosomal RNAs.**

- Overall, **almost 300 different macromolecules cooperate to synthesize polypeptides.**

- **Many of these macromolecules are organized into the complex three-dimensional structure of the ribosome.**

- Despite the great complexity of protein synthesis, proteins are made at exceedingly high rates.

- A **polypeptide of 100 residues is synthesized in an *Escherichia coli* cell (at 37 C) in about 5 seconds.**

- Synthesis of the thousands of different **proteins in a cell is tightly regulated, so that just enough copies are made to match the current metabolic circumstances**.

- To maintain the **appropriate mix and concentration of proteins, the targeting and degradative processes must keep pace with synthesis**.

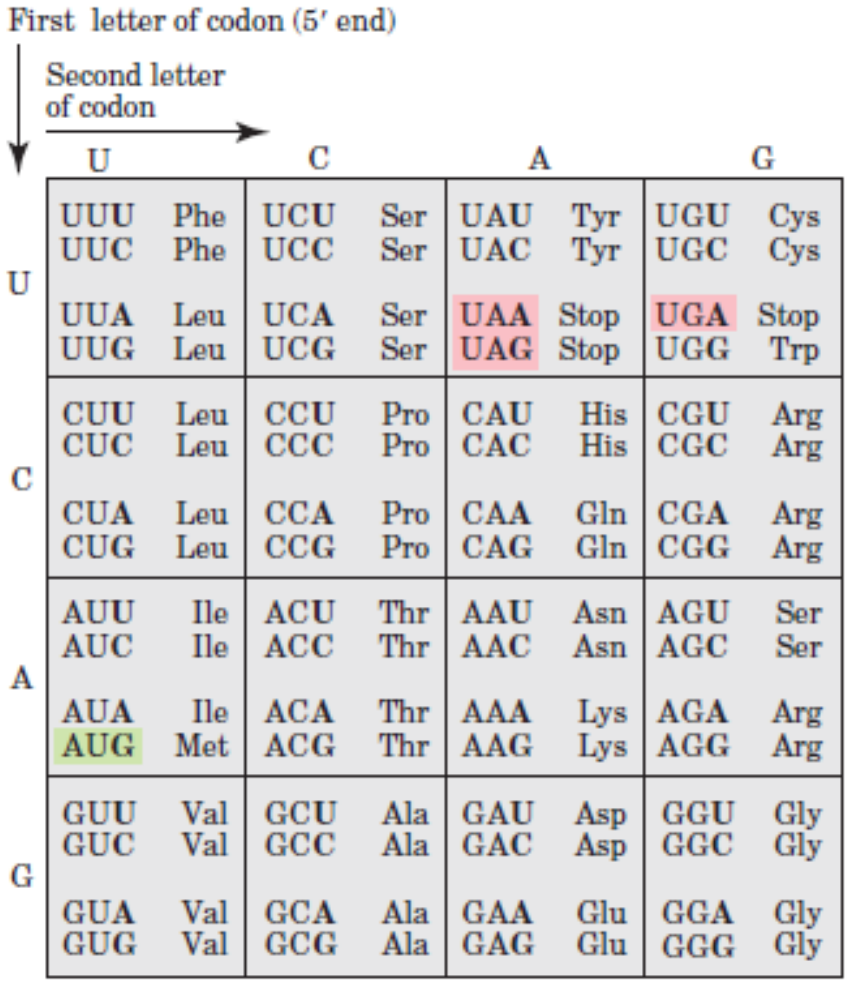
- Research is gradually uncovering the finely coordinated cellular choreography that guides each protein to its proper cellular location and selectively degrades it when it is no longer required.

- Researchers have elucidated the structure of bacterial ribosomes, revealing the workings of cellular protein synthesis in beautiful molecular detail. And what did they find? **Proteins are synthesized by a gigantic RNA enzyme!**

By the 1960s it had long been apparent that at least three nucleotide residues of DNA are necessary to encode each amino acid.

The four code letters of DNA (A, T, G, and C) in groups of two can yield only $4^2 = 16$ different combinations, insufficient to encode 20 aminoacids.

Groups of three, however, yield $4^3 = 64$ different combinations



First letter of codon (5′ end)	Second letter of codon: U		Second letter of codon: C		Second letter of codon: A		Second letter of codon: G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

FIGURE 27–7 **"Dictionary" of amino acid code words in mRNAs.** The codons are written in the 5′⟶3′ direction. The third base of each codon (in bold type) plays a lesser role in specifying an amino acid than the first two. The three termination codons are shaded in pink, the initiation codon AUG in green. All the amino acids except methionine and tryptophan have more than one codon. In most cases, codons that specify the same amino acid differ only at the third base.

Several key properties of the genetic code were established in early genetic studies .

A **codon** is a triplet of nucleotides that codes for a specific amino acid. Translation occurs in such a way that these nucleotide triplets are read in a successive, nonoverlapping fashion.

A specific first codon in the sequence establishes the **reading frame,** in which a new codon begins every three nucleotide residues.

There is no punctuation between codons for successive amino acid residues.

The amino acid sequence of a protein is defined by a linear sequence of contiguous triplets.

In principle, any given single-stranded DNA or mRNA sequence has three possible reading frames.

Each reading frame gives a different sequence of codons (Fig. 27–5), but only one is likely to encode a given protein.

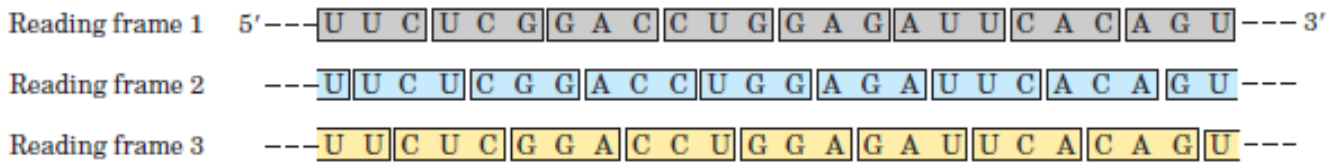


FIGURE 27–5 Reading frames in the genetic code. In a triplet, nonoverlapping code, all mRNAs have three potential reading frames, shaded here in different colors. The triplets, and hence the amino acids specified, are different in each reading frame.

- Consolidation of the results from many experiments permitted the **assignment of 61 of the 64 possible codons.**

- The **other three were identified as termination codons,** in part because they disrupted amino acid coding patterns when they occurred in a synthetic RNA polymer

- Meanings for all the triplet codons (tabulated in Fig. 27–7) were established by 1966 and have been verified in many different ways.

- The cracking of the genetic code is regarded as one of the most important scientific discoveries of the twentieth century

- **Codons are the key to the translation of genetic information, directing the synthesis of specific proteins.**

- **The reading frame is set when translation of an mRNA molecule begins, and it is maintained as the synthetic machinery reads sequentially from one triplet to the next.**

- **If the initial reading frame is off by one or two bases, or if translation somehow skips a nucleotide in the mRNA, all the subsequent codons will be out of register; the result is usually a "missense" protein with a garbled amino acid sequence.**

- Several codons serve special functions (Fig. 27–7).

- The **initiation codon AUG** is the most common signal for the beginning of a polypeptide in all cells , **in addition to coding for Met residues in internal positions of polypeptides.**

- The **termination codons** (**UAA, UAG, and UGA**), also called stop codons or nonsense codons, normally signal the end of polypeptide synthesis and do not code for any known amino acids.

- In a random sequence of nucleotides, 1 in every 20 codons in each reading frame is, on average, a termination codon.

- In general, a reading frame without a termination codon among 50 or more codons is referred to as an **open reading frame (ORF).**

- A striking feature of the genetic code is that an amino acid may be specified by more than one codon, so the code is described as **degenerate.**

- This does *not* suggest that the code is flawed: although an amino acid may have two or more codons, each codon specifies only one amino acid.

- The degeneracy of the code is not uniform.

- Whereas methionine and tryptophan have single codons, for example, three amino acids (Leu, Ser, Arg) have six codons, five amino acids have four, isoleucine has three, and nine amino acids have two (Table 27–3).

TABLE 27–3 Degeneracy of the Genetic Code

Amino acid	*Number of codons*	*Amino acid*	*Number of codons*
Met	1	Tyr	2
Trp	1	Ile	3
Asn	2	Ala	4
Asp	2	Gly	4
Cys	2	Pro	4
Gln	2	Thr	4
Glu	2	Val	4
His	2	Arg	6
Lys	2	Leu	6
Phe	2	Ser	6

- The genetic code is nearly universal.

- With the intriguing exception of a few minor variations in mitochondria, some bacteria, and some single-celled eukaryotes, amino acid codons are identical in all species examined so far.

- Human beings, *E. coli,* tobacco plants, amphibians, and viruses share the same genetic code.

- Thus it would appear that all life forms have a common evolutionary ancestor, whose genetic code has been preserved throughout biological evolution.

TABLE 27–5 Components Required for the Five Major Stages of Protein Synthesis in *E. coli*

Stage	*Essential components*
1. Activation of amino acids	20 amino acids 20 aminoacyl-tRNA synthetases 32 or more tRNAs ATP Mg^{2+}
2. Initiation	mRNA *N*-Formylmethionyl-$tRNA^{fmet}$ Initiation codon in mRNA (AUG) 30S ribosomal subunit 50S ribosomal subunit Initiation factors (IF-1, IF-2, IF-3) GTP Mg^{2+}
3. Elongation	Functional 70S ribosome (initiation complex) Aminoacyl-tRNAs specified by codons Elongation factors (EF-Tu, EF-Ts, EF-G) GTP Mg^{2+}
4. Termination and release	Termination codon in mRNA Release factors (RF-1, RF-2, RF-3)
5. Folding and posttranslational processing	Specific enzymes, cofactors, and other components for removal of initiating residues and signal sequences, additional proteolytic processing, modification of terminal residues, and attachment of phosphate, methyl, carboxyl, carbohydrate, or prosthetic groups

Protein Biosynthesis Takes Place in Five Stages

Stage 1: Activation of Amino Acids

For the synthesis of a polypeptide with a defined sequence, two fundamental chemical requirements must be met:

(1) the carboxyl group of each amino acid must be activated to facilitate formation of a peptide bond, and

(1) a link must be established between each new amino acid and the information in the mRNA that encodes it.

Both these requirements are met by attaching the amino acid to a tRNA in the first stage of protein synthesis.

Attaching the right amino acid to the right tRNA is critical. This reaction takes place in the **cytoso**l, not on the ribosome.

Each of the 20 amino acids is covalently attached to a specific tRNA at the expense of ATP energy, using Mg2- dependent activating enzymes known as **aminoacyltRNA synthetases.**

When attached to their amino acid (aminoacylated) the tRNAs are said to be “charged.

Stage 2: Initiation The mRNA bearing the code for the polypeptide to be made binds to the **smaller of two ribosomal subunits and to the initiating aminoacyl-tRNA**.

The **large ribosomal subunit then binds to form an initiation complex.**

The initiating **aminoacyl-tRNA basepairs with the mRNA codon AUG that signals the beginning of the polypeptide.**

This process, which **requires GTP, is promoted by cytosolic proteins called initiation factors.**

Stage 3: Elongation

The nascent polypeptide is lengthened by covalent attachment of successive amino acid units, each carried to the ribosome and correctly positioned by its tRNA, which base-pairs to its corresponding codon in the mRNA.

Elongation requires cytosolic proteins known as elongation factors. The binding of each incoming aminoacyl-tRNA and the movement of the ribosome along the mRNA are facilitated by the hydrolysis of GTP as each residue is added to the growing polypeptide

Stage 4: Termination and Release

Completion of the polypeptide chain **is signaled by a termination codon in the mRNA.**

The new polypeptide is released from the ribosome, aided by proteins called **release factors.**

Stage 5: Folding and Posttranslational Processing

In order to achieve its biologically active form, the new polypeptide must fold into its proper three-dimensional conformation.

Before or after folding, the new **polypeptide may undergo enzymatic processing**, including removal of one or more amino acids (usually from the amino terminus); addition of acetyl, phosphoryl, methyl, carboxyl, or other groups to certain amino acid residues; proteolytic cleavage; and/or attachment of oligosaccharides or prosthetic groups.

Basics of Genetics, Concepts and terms

- What is genetics?

– “Genetics is the study of **heredity**, the process in which a parent passes certain **genes** onto their children.”

- What does that mean?

– Children **inherit** their biological parents’ genes that express specific **traits**, such as some physical characteristics, natural talents, and genetic disorders.

Cell

- When a cell divides, one of its main jobs is to make sure that each of the two new cells gets a full, perfect copy of genetic material. Mistakes during copying, or unequal division of the genetic material between cells, can lead to cells that are unhealthy or dysfunctional (and may lead to diseases such as cancer).

But what exactly is this genetic material, and how does it behave over the course of a cell division?

- **DNA and genomes**
- **DNA** (**deoxyribonucleic acid**) is the genetic material of living organisms. In humans, DNA is found in almost all the cells of the body and provides the instructions they need to grow, function, and respond to their environment.

- When a cell in the body divides, it will pass on a copy of its DNA to each of its daughter cells. DNA is also passed on at the level of organisms, with the DNA in sperm and egg cells combining to form a new organism that has genetic material from both its parents.

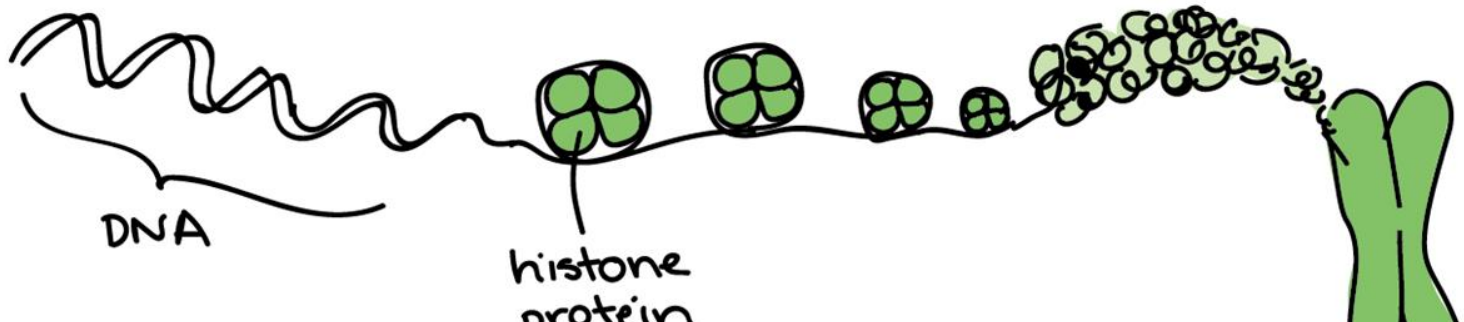
- Genes are small sections of the long chain of DNA.

- Genes typically provide instructions for making proteins, which give cells and organisms their functional characteristics.

- A cell's set of DNA is called its **genome**.

Chromatin

In a cell, DNA does not usually exist by itself, but instead associates with specialized proteins that organize it and give it structure. In eukaryotes, these proteins include the **histones**, a group of basic (positively charged) proteins that form “bobbins” around which negatively charged DNA can wrap. In addition to organizing DNA and making it more compact, histones play an important role in determining which genes are active. **The complex of DNA plus histones and other structural proteins is called chromatin.**



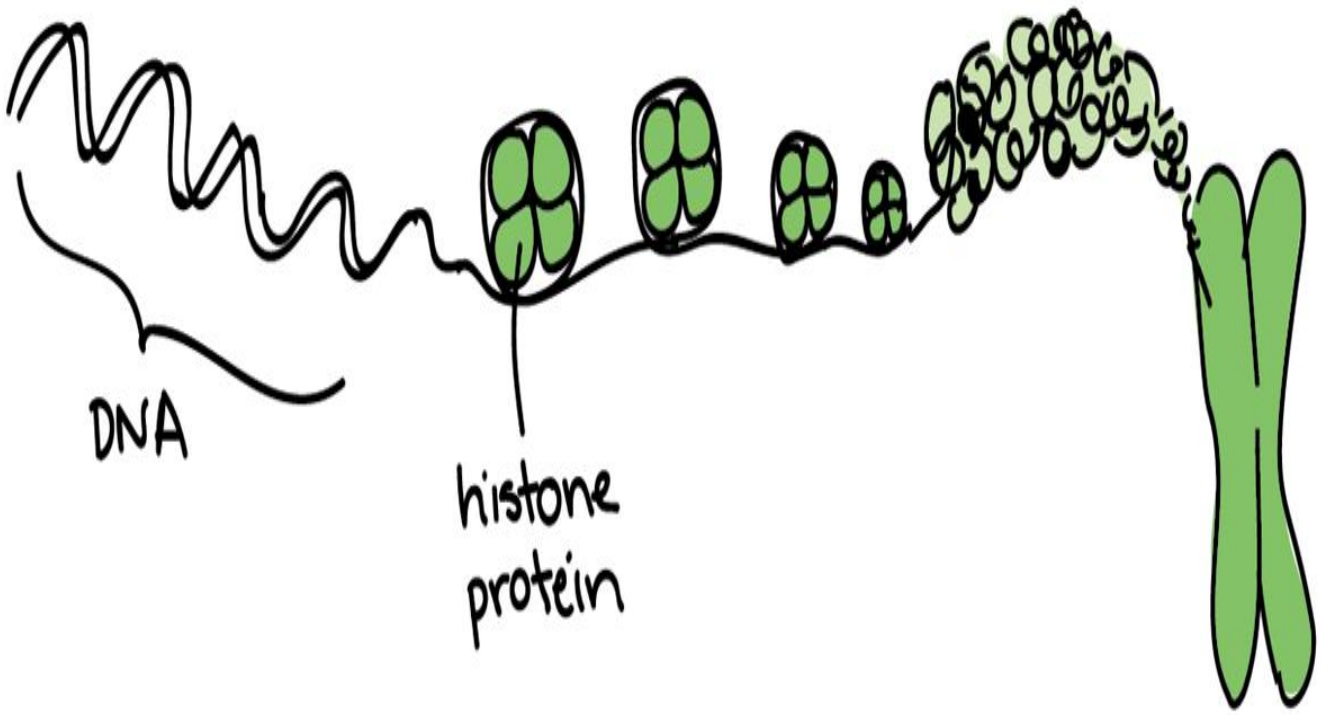
Chromosomes

- Each species has its own characteristic number of chromosomes. Humans, for instance, have 46 chromosomes in a typical body cell (somatic cell), while dogs have 78.

- Like many species of animals and plants, humans are **diploid** (**2*n***), meaning that most of their chromosomes come in matched sets known as **homologous pairs**.
- The 46 chromosomes of a human cell are organized into 23 pairs, and the two members of each pair are said to be **homologues** of one another.

- Human sperm and eggs, which have only one homologous chromosome from each pair, are said to be **haploid** (**1*n***). When a sperm and egg fuse, their genetic material combines to form one complete, diploid set of chromosomes. So, for each homologous pair of chromosomes in your genome, one of the homologues comes from your mom and the other from your dad.

Chromosomes



Chromosomes

- The two chromosomes in a homologous pair are very similar to one another and have the same size and shape. Most importantly, they carry the same type of genetic information: that is, they have the same genes in the same locations. However, they don't necessarily have the same versions of genes. That's because you may have inherited two different gene versions from your mom and your dad.

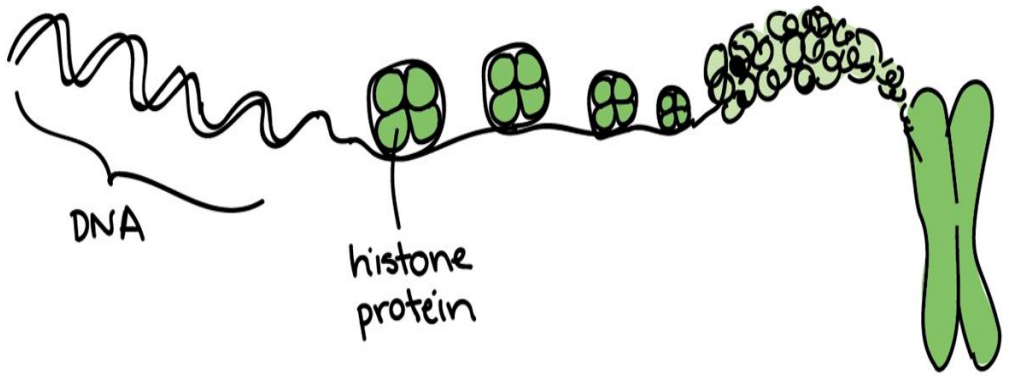
- As a real example, let's consider a gene on chromosome 9 that determines blood type (A, B, AB, or O). It's possible for a person to have two identical copies of this gene, one on each homologous chromosome—for example, you may have a double dose of the gene version for type A. On the other hand, you may have two different gene versions on your two homologous chromosomes, such as one for type A and one for type B (giving AB blood).

sex chromosomes

- The **sex chromosomes**, X and Y, determine a person's biological sex: XX specifies female and XY specifies male. These chromosomes are not true homologues and are an exception to the rule of the same genes in the same places. Aside from small regions of similarity needed during meiosis, or sex cell production, the X and Y chromosomes are different and carry different genes. The 44 non-sex chromosomes in humans are called **autosomes**.

Chromosomes and cell division

- As a cell prepares to divide, it must make a copy of each of its chromosomes. The two copies of a chromosome are called **sister chromatids**. The sister chromatids are identical to one another and are attached to each other by proteins called **cohesins**. The attachment between sister chromatids is tightest at the **centromere**, a region of DNA that is important for their separation during later stages of cell division.
- As long as the sister chromatids are connected at the centromere, they are still considered to be one chromosome. However, as soon as they are pulled apart during cell division, each is considered a separate chromosome.



Cell Cycle

- The **cell cycle** can be thought of as the life cycle of a cell. In other words, it is the series of growth and development steps a cell undergoes between its “birth”—formation by the division of a mother cell—and reproduction—division to make two new daughter cells.

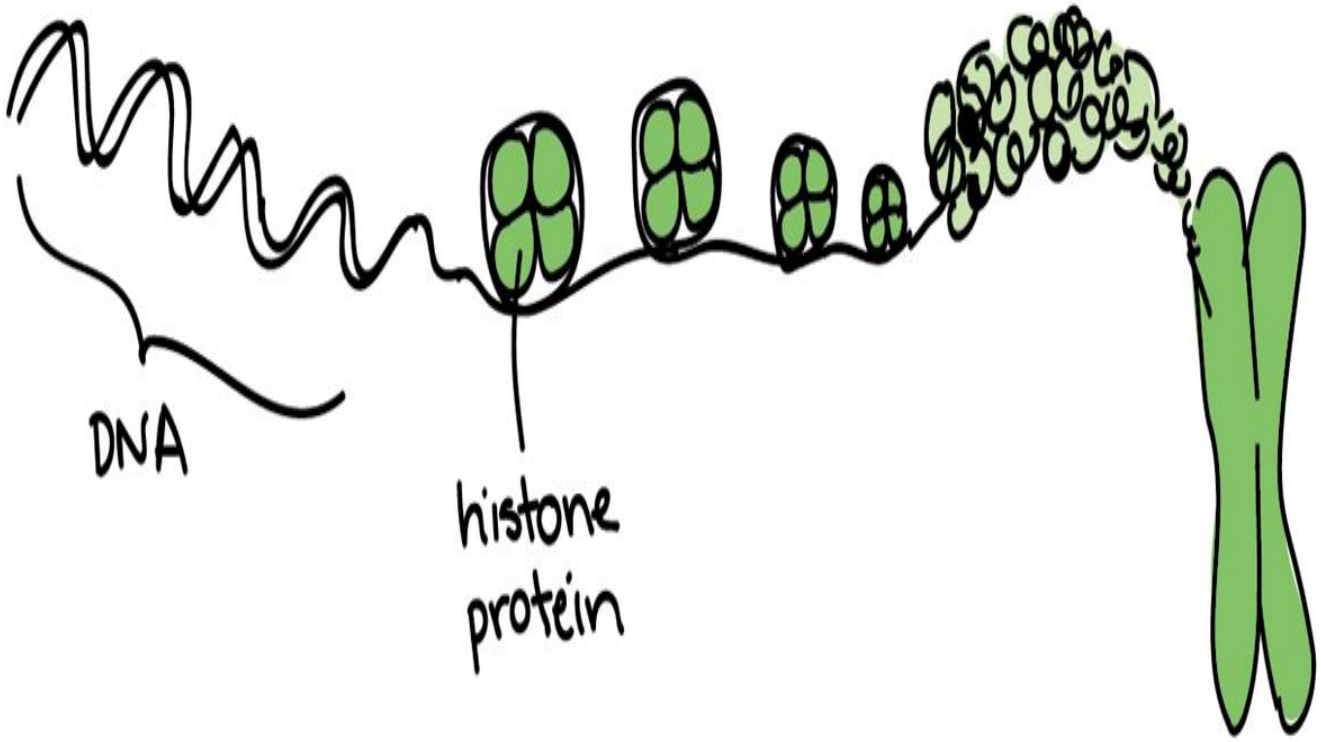
Stages of the cell cycle

- To divide, a cell must complete several important tasks: it must grow, copy its genetic material (DNA), and physically split into two daughter cells. Cells perform these tasks in an organized, predictable series of steps that make up the cell cycle. The cell cycle is a cycle, rather than a linear pathway, because at the end of each go-round, the two daughter cells can start the exact same process over again from the beginning.

In eukaryotic cells, or cells with a nucleus, the stages of the cell cycle are divided into two major phases: **interphase** and the **mitotic (M) phase**.

- During *interphase*, the cell grows and makes a copy of its DNA.
- During the *mitotic (M) phase*, the cell separates its DNA into two sets and divides its cytoplasm, forming two new cells

Cell Cycle



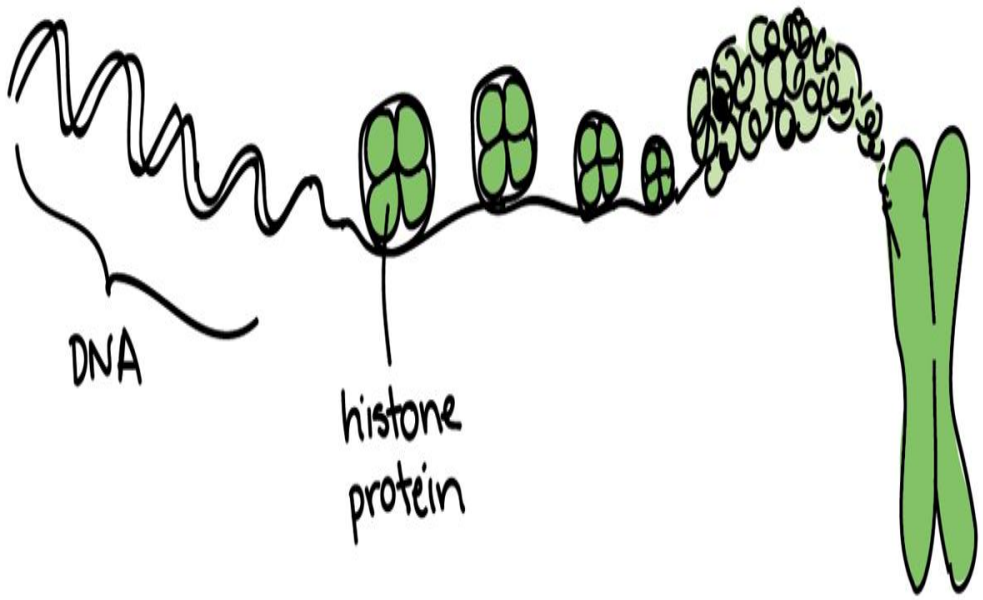
Interphase

- Let's enter the cell cycle just as a cell forms, by division of its mother cell. What must this newborn cell do next if it wants to go on and divide itself? Preparation for division happens in three steps:
- **G1 phase. During G1 phase, also called the first gap phase, the cell grows physically larger, copies organelles, and makes the molecular building blocks it will need in later steps.**
- **S phase. In S phase, the cell synthesizes a complete copy of the DNA in its nucleus. It also duplicates a microtubule-organizing structure called the centrosome. The centrosomes help separate DNA during M phase.**
- **G2 phase. During the second gap phase, the cell grows more, makes proteins and organelles, and begins to reorganize its contents in preparation for mitosis. G2 phase ends when mitosis begins.**
- **The G1, S and G2 phases together is called interphase. The prefix *inter*- means between, reflecting that interphase takes place between one mitotic (M) phase and the next.**

Mitosis

- **Mitosis** is a type of cell division in which one cell (the **mother**) divides to produce two new cells (the **daughters**) that are genetically identical to itself. In the context of the cell cycle, mitosis is the part of the division process in which the DNA of the cell's nucleus is split into two equal sets of chromosomes.
- During the mitotic (M) phase, the cell divides its copied DNA and cytoplasm to make two new cells. M phase involves two distinct division-related processes: mitosis and cytokinesis.
- In **mitosis**, the nuclear DNA of the cell condenses into visible chromosomes and is pulled apart by the mitotic spindle, a specialized structure made out of microtubules. Mitosis takes place in four stages: prophase (sometimes divided into early prophase and prometaphase), metaphase, anaphase, and telophase. You can learn more about these stages in the video on mitosis.
- In **cytokinesis**, the cytoplasm of the cell is split in two, making two new cells. Cytokinesis usually begins just as mitosis is ending, with a little overlap. Importantly, cytokinesis takes place differently in animal and plant cells.

- This cell is in interphase (late G2 phase) and has already copied its DNA, so the chromosomes in the nucleus each consist of two connected copies, called **sister chromatids**. You can't see the chromosomes very clearly at this point, because they are still in their long, stringy, decondensed form.
- This animal cell has also made a copy of its **centrosome**, an organelle that will play a key role in orchestrating mitosis, so there are tw[illegible]

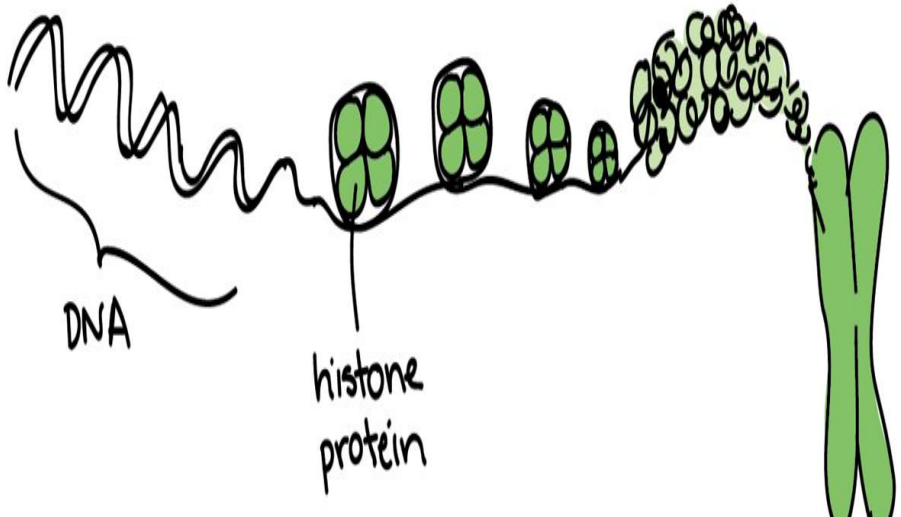


Phases of Mitotic phase

- Mitosis consists of four basic phases: prophase, metaphase, anaphase, and telophase.
- Some textbooks list five, breaking prophase into an early phase (called prophase) and a late phase (called prometaphase).
- These phases occur in strict sequential order, and cytokinesis - the process of dividing the cell contents to make two new cells - starts in anaphase or telophase.

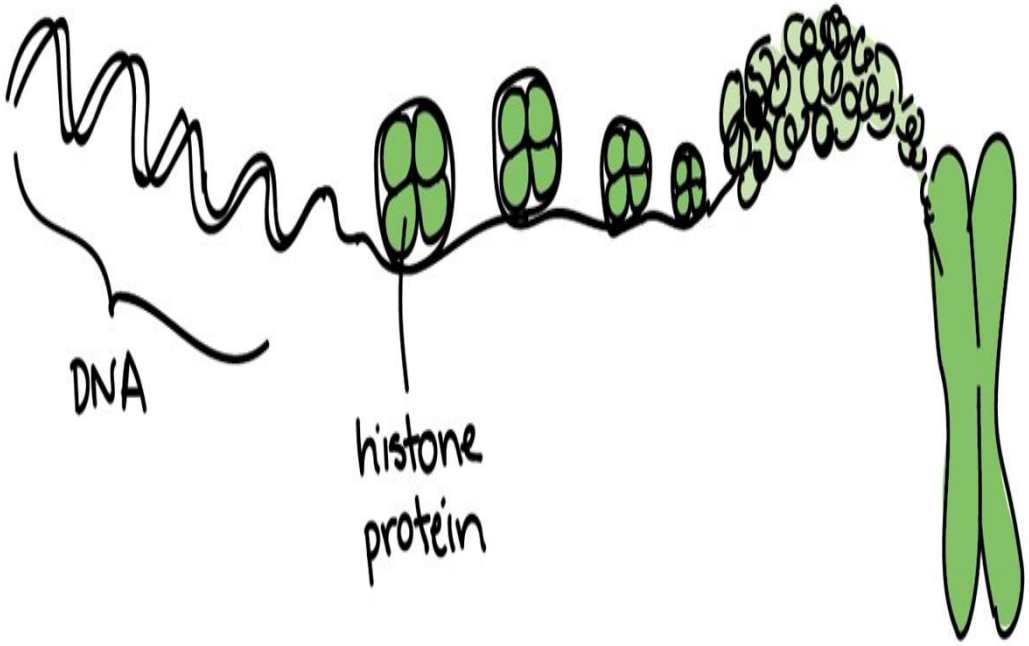
Early Prophase

- In early **prophase**, the cell starts to break down some structures and build others up, setting the stage for division of the chromosomes.
- The chromosomes start to condense (making them easier to pull apart later on).
- The **mitotic spindle** begins to form. The spindle is a structure made of microtubules, strong fibers that are part of the cell's "skeleton." Its job is to organize the chromosomes and move them around during mitosis. The spindle grows between the centrosomes as they move apart.
- The **nucleolus** (or nucleoli, plural), a part of the nucleus where ribosomes are made, disappears. This is a sign that the nucleus is getting ready to bre
-



Prometaphase

- In late prophase (sometimes also called **prometaphase**), the mitotic spindle begins to capture and organize the chromosomes.
- The chromosomes finish condensing, so they are very compact.
- The nuclear envelope breaks down, releasing the chromosomes.
- The mitotic spindle grows more, and some of the microtubules start t

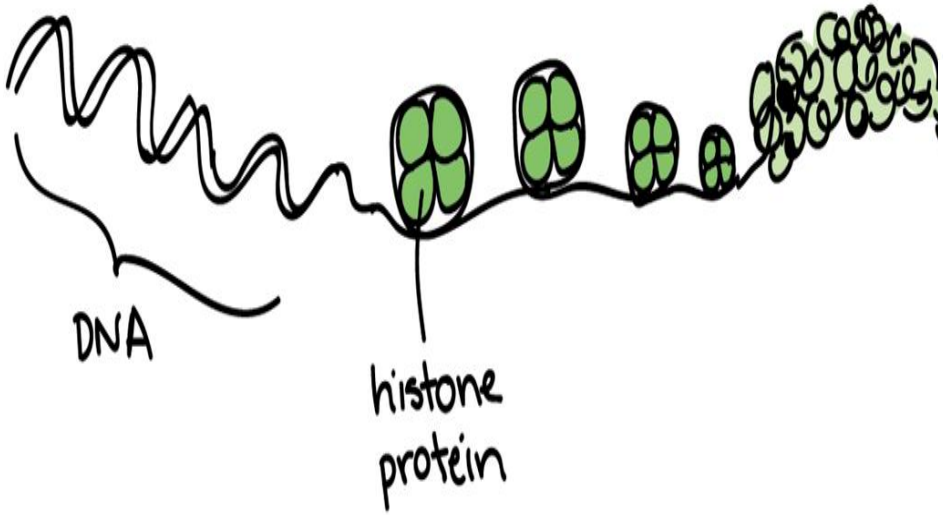


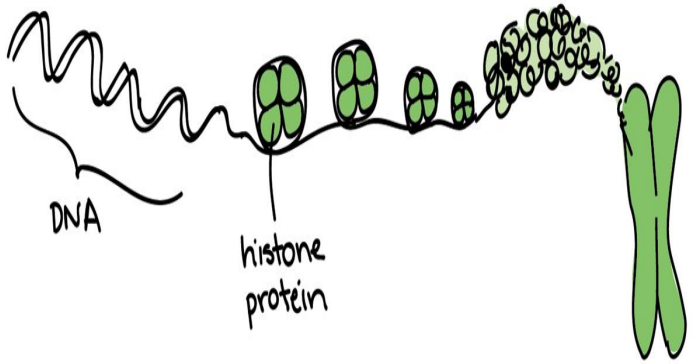
Metaphase

- Microtubules can bind to chromosomes at the **kinetochore**, a patch of protein found on the centromere of each sister chromatid. (**Centromeres** are the regions of DNA where the sister chromatids are most tightly connected.)
- Microtubules that bind a chromosome are called **kinetochore microtubules**. Microtubules that don't bind to kinetochores can grab on to microtubules from the opposite pole, stabilizing

each centrosome

ture called





Metaphase and checkpoint

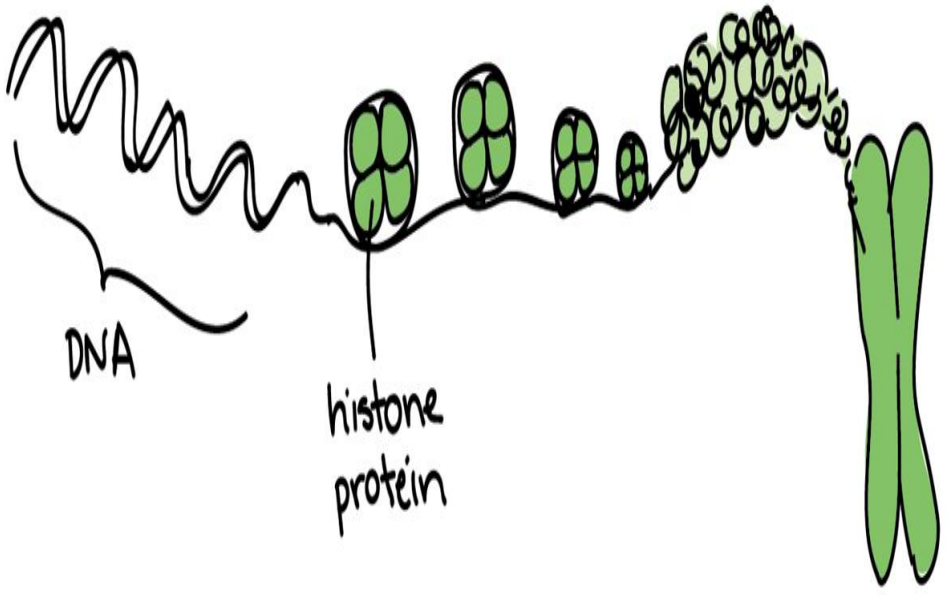
In **metaphase**, the spindle has captured all the chromosomes and lined them up at the middle of the cell, ready to divide.

- All the chromosomes align at the **metaphase plate** (not a physical structure, just a term for the plane where the chromosomes line up).
- At this stage, the two kinetochores of each chromosome should be attached to microtubules from opposite spindle poles.
- Before proceeding to anaphase, the cell will check to make sure that all the chromosomes are at the metaphase plate with their kinetochores correctly attached to microtubules. This is called the **spindle checkpoint** and helps ensure that the sister chromatids will split evenly between the two daughter cells when they separate in the next step. If a chromosome is not properly aligned or attached, the cell will halt division until the problem is fixed.

Anaphase

In **anaphase**, the sister chromatids separate from each other and are pulled towards opposite ends of the cell.

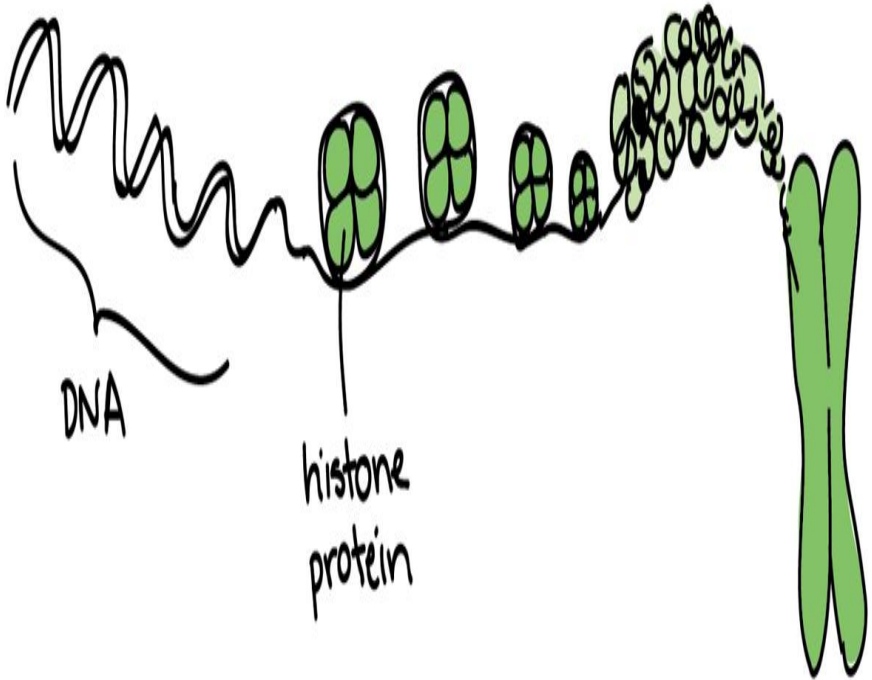
- The protein “glue” that holds the sister chromatids together is broken down, allowing them to separate. Each is now its own chromosome. The chromosomes of each pair are pulled towards opposite ends of the cell.
- Microtubules not attached to chromosomes elongate and push apart, separating the poles and making the cell longer.
- All of these processes are driven by **motor proteins**, molecular machines that can “walk” along microtu
  proteins carry chromosomes



Telophase

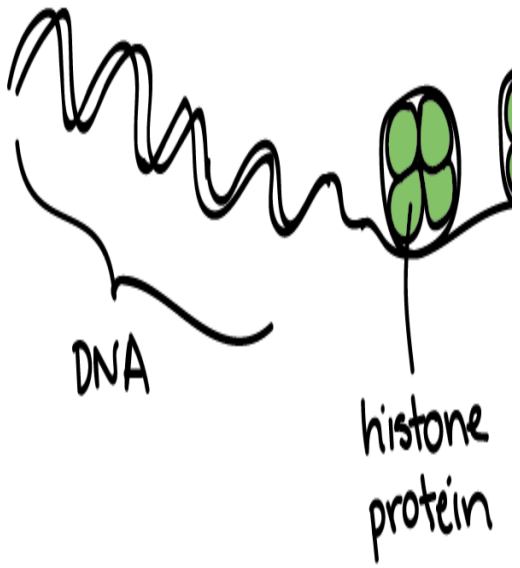
In **telophase**, the cell is nearly done dividing, and it starts to re-establish its normal structures as cytokinesis (division of the cell contents) takes place.

- The mitotic spindle is broken down into its building blocks.
- Two new nuclei form, one for each set of chromosomes. Nuclear membranes and nucleoli reappear.
- The chr to their "stringy"



Cytokinesis

- **Cytokinesis**, the division of the cytoplasm to form two new cells, overlaps with the final stages of mitosis. It may start in either anaphase or telophase, depending on the cell, and finishes shortly after telophase.
- In animal cells, cytokinesis is contractile, pinching the cell in two like a coin purse with a drawstring. The “drawstring” is a band of f protein called actin, and the pi as the **cleavage furrow**.



- When cytokinesis finishes, we end up with two new cells, each with a complete set of chromosomes identical to those of the mother cell. The daughter cells can now begin their own cellular “lives,” and – depending on what they decide to be when they grow up – may undergo mitosis themselves, repeating the cycle

Meiosis

- Mitosis is used for almost all of your body's cell division needs. It adds new cells during development and replaces old and worn-out cells throughout your life. The goal of mitosis is to produce daughter cells that are genetically identical to their mothers, with not a single chromosome more or less
- Meiosis, on the other hand, is used for just one purpose in the human body: the production of **gametes**—sex cells, or sperm and eggs. Its goal is to make daughter cells with exactly half as many chromosomes as the starting cell.
- To put that another way, **meiosis** in humans is a division process that takes us from a diploid cell—one with two sets of chromosomes—to haploid cells—ones with a single set of chromosomes. In humans, the haploid cells made in meiosis are sperm and eggs. When a sperm and an egg join in fertilization, the two haploid sets of chromosomes form a complete

Phases of meiosis

- In many ways, meiosis is a lot like mitosis. The cell goes through similar stages and uses similar strategies to organize and separate chromosomes. In meiosis, however, the cell has a more complex task. **It still needs to separate sister chromatids (the two halves of a duplicated chromosome), as in mitosis. But it must also separate homologous chromosomes, the similar but nonidentical chromosome pairs an organism receives from its two parents.**

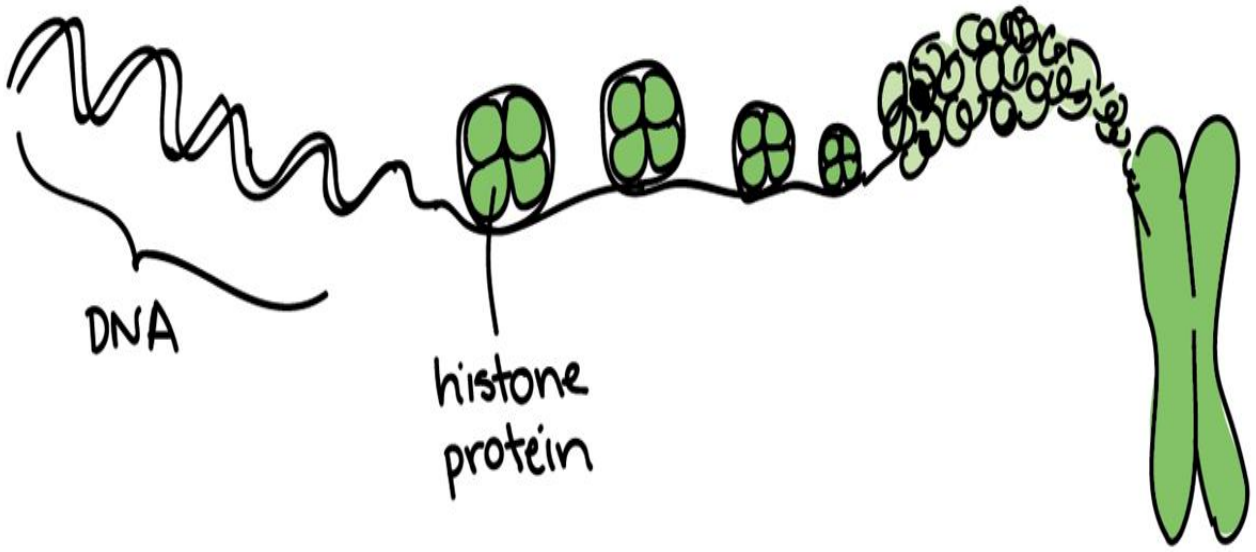
- These goals are accomplished in meiosis using a two-step division process. Homologue pairs separate during a first round of cell division, called **meiosis I**. Sister chromatids separate during a second round, called **meiosis II**.

- Since cell division occurs twice during meiosis, one starting cell can produce four gametes (eggs or sperm). In each round of division, cells go through four stages: prophase, metaphase, anaphase, and telophase.

Meiosis I

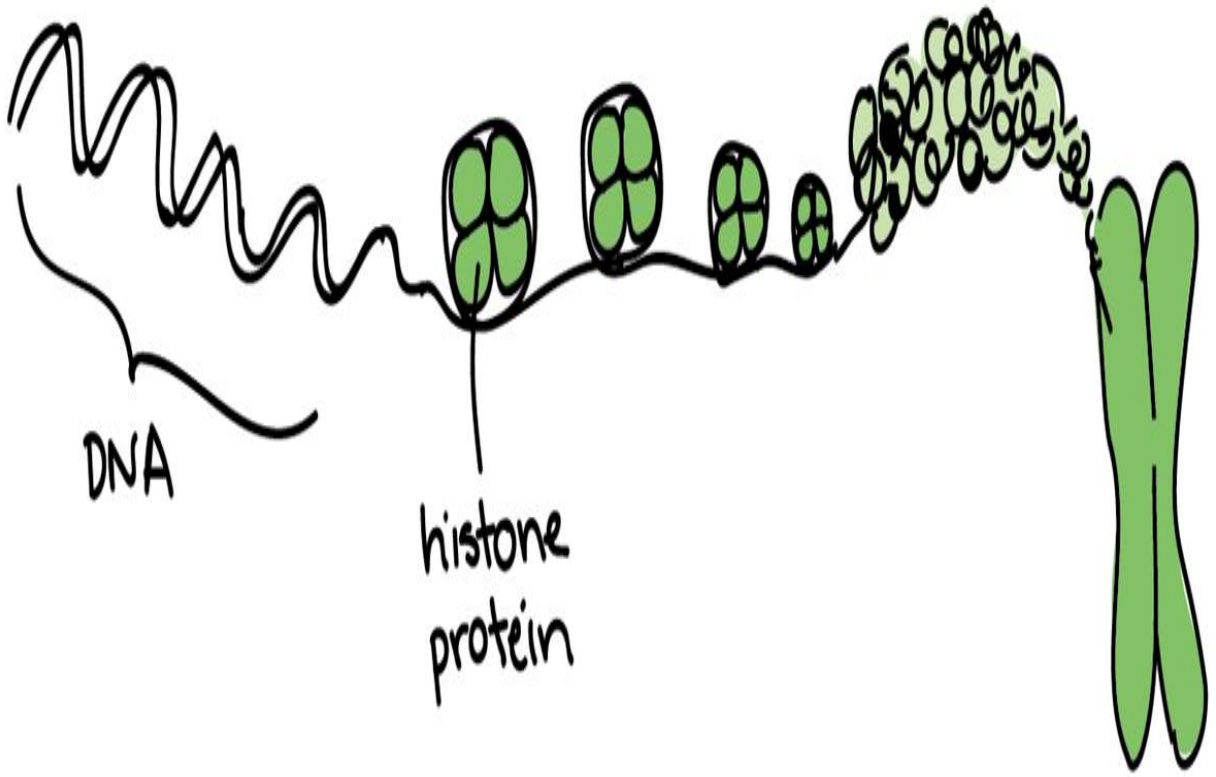
- Before entering meiosis I, a cell must first go through interphase. As in mitosis, the cell grows during G1phase, copies all of its chromosomes during S phase, and prepares for division during G2 phase.
- During **prophase I**, differences from mitosis begin to appear. As in mitosis, the chromosomes begin to condense, but in meiosis I, they also pair up. Each chromosome carefully aligns with its homologue partner so that the two match up at corresponding positions along their full length.

- For instance, in the image below, the letters A, B, and C represent genes found at particular spots on the chromosome, with capital and lowercase letters for different forms, or alleles, of each gene. The DNA is broken at the same spot on each homologue—here, between genes B and C—and reconnected in a criss-cross pattern so that the homologues exchange part of their DNA.
- This process, in which homologous chromosomes trade parts, is called **crossing over**. It's helped along by a protein structure called the **synaptonemal complex** that holds the homologues together. The chromosomes would actually be positioned one on top of the other—throughout crossing over; they're only shown side-by-side in the image below so that it's easier to see the exchange of genetic m
-



- After crossing over, the spindle begins to capture chromosomes and move them towards the center of the cell (metaphase plate). This may seem familiar from mitosis, but there is a twist. During **metaphase I**, homologue pairs—not individual chromosomes—line up at the metaphase plate for separation.

- In **anaphase I**, the homologues are pulled apart and move apart to opposite ends of the cell. The sister chromatids of each chromosome, however, remain attached to one another and don't come apart.
- Finally, in **telophase I**, the chromosomes arrive at opposite poles of the cell. In some organisms, the nuclear membrane re-forms and the chromosomes decondense, although in others, this step is skipped—since cells will soon go through another round of division, meiosis II.
- Cytokinesis usually occurs at the same time as telophase I, forming two haploid daughter cells.

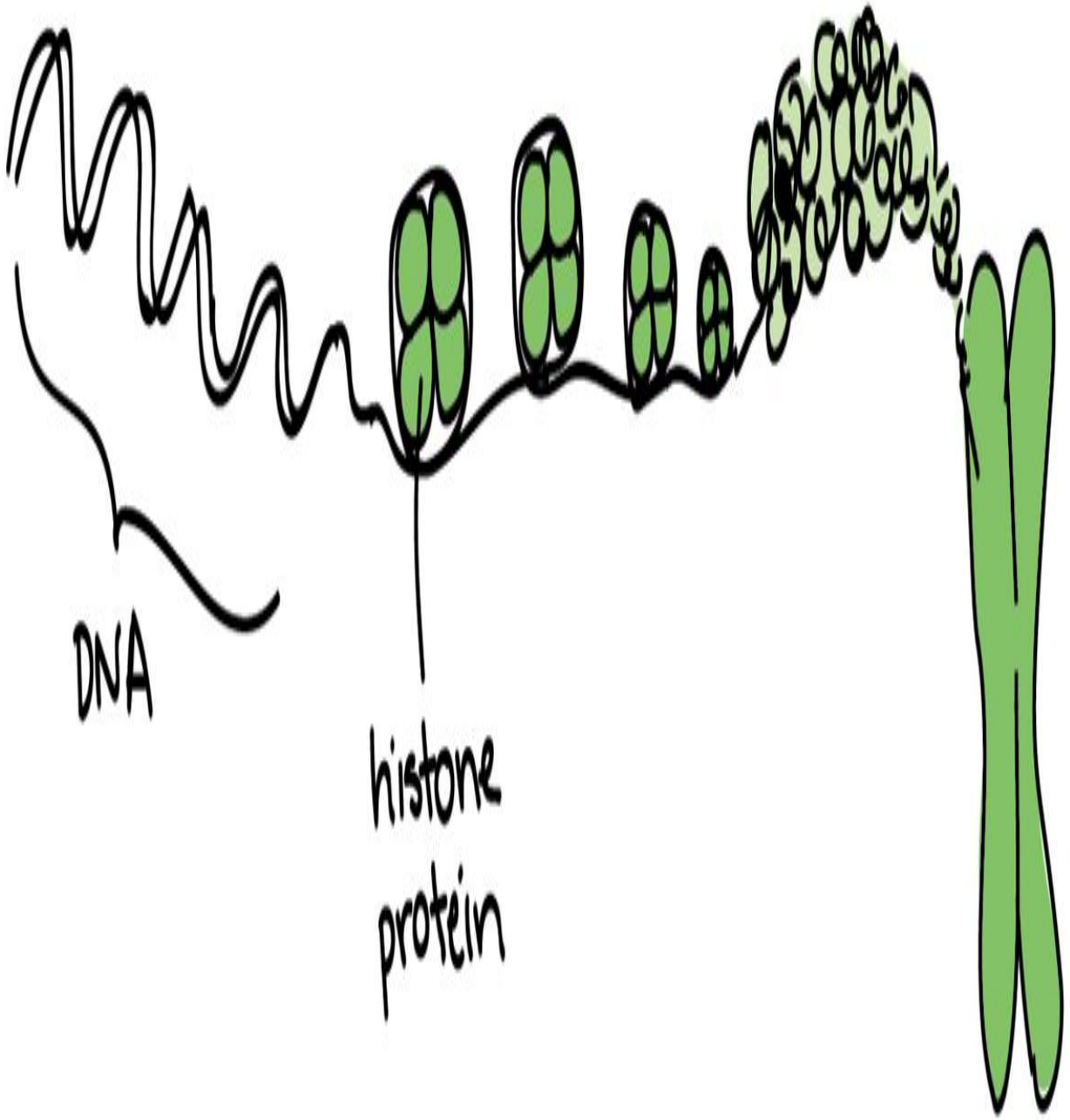

DNA
histone protein

Meiosis II

- Cells move from meiosis I to meiosis II without copying their DNA. Meiosis II is a shorter and simpler process than meiosis I, and you may find it helpful to think of meiosis II as “mitosis for haploid cells."
- The cells that enter meiosis II are the ones made in meiosis I. These cells are haploid—have just one chromosome from each homologue pair—but their chromosomes still consist of two sister chromatids. In meiosis II, the sister chromatids separate, making haploid cells with non-duplicated chromosomes.

Phases of Meiosis II

- During **prophase II**, chromosomes condense and the nuclear envelope breaks down, if needed. The centrosomes move apart, the spindle forms between them, and the spindle microtubules begin to capture chromosomes. *[When did the centrosomes duplicate?]*
- 3cubed5start superscript, 5, end superscriptThe two sister chromatids of each chromosome are captured by microtubules from opposite spindle poles. In **metaphase II**, the chromosomes line up individually along the metaphase plate. In **anaphase II**, the sister chromatids separate and are pulled towards opposite poles of the cell.
- In **telophase II**, nuclear membranes form around each set of chromosomes, and the chromosomes decondense. Cytokinesis splits the chromosome sets into new cells, forming the final products of meiosis: four haploid cells in which each chromosome has just one chromatid. In humans, the products of meiosis are sperm or egg cells


DNA
histone
protein

Inheritance

Today, we know that many of people's characteristics, from hair color to height to risk of diabetes, are influenced by genes. We also know that genes are the way parents pass characteristics on to their children (including things like dimples, or—in the case of me and my father—a terrible singing voice). In the last hundred years, we've come to understand that genes are actually pieces of DNA that are found on chromosomes and specify proteins.

But did we always know those things? Not by a long shot! About 150 years ago, a monk named Gregor Mendel published a paper that first proposed the existence of genes and presented a model for how they were inherited. Mendel's work was the first step on a long road, involving many hard-working scientists, that's led to our present understanding of genes and what they do.

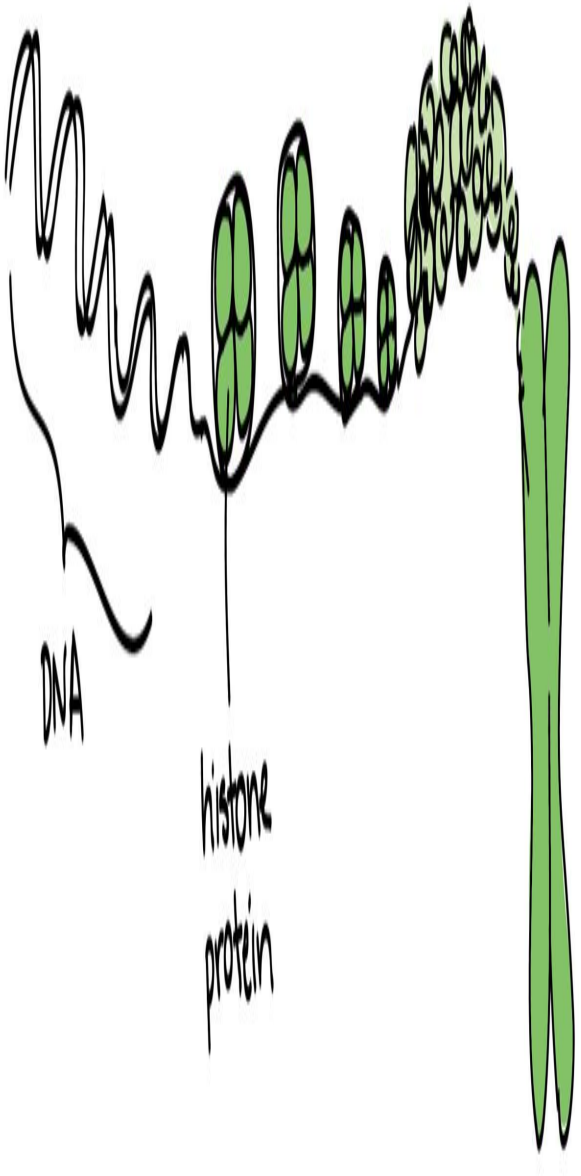
Mendel's Model

- Mendel studied the genetics of pea plants, and he traced the inheritance of a variety of characteristics, including flower color, flower position, seed color, and seed shape. To do so, he started by crossing **pure-breeding** parent plants with different forms of a characteristic, such as violet and white flowers. Pure-breeding just means that the plant will always make more offspring like itself, when self-fertilized over many generations.
- In self-fertilization, sperm and eggs from the same pea plant combine inside a closed flower, producing seeds with a single plant as both mother and father.

What results did Mendel find in his crosses for flower color?

- In the parental, or P generation, Mendel **crossed a pure-breeding violet-flowered plant to a pure-breeding white-flowered plant.** When he gathered and planted the seeds produced in this cross, Mendel found that **100 percent of the plants in the next generation, or F1 generation, had violet flowers.**
- Conventional wisdom at that time would have predicted that the hybrid flowers should be pale violet—that is, that the parents' traits should blend in the offspring. Instead, Mendel's results showed that the white flower trait had completely disappeared. He called the trait that was visible in the F1 generation (violet flowers) the **dominant trait**, and the trait that was hidden or lost (white flowers) the **recessive trait**.

- Importantly, Mendel did not stop his experimentation there. Instead, he let the F1 plants self-fertilize. Among their offspring, called the F2 generation, he found that 705 plants had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers to one white flower, or approximately 3:1
- This 3:1 ratio was no fluke. For the other six characteristics that Mendel examined, both the F1 and F2 generations behaved in the same way they did for flower color.
- One of the two traits would disappear completely from the F1, only to reappear in the F2 generation in a ratio of roughly 3:1


DNA
histone
protein

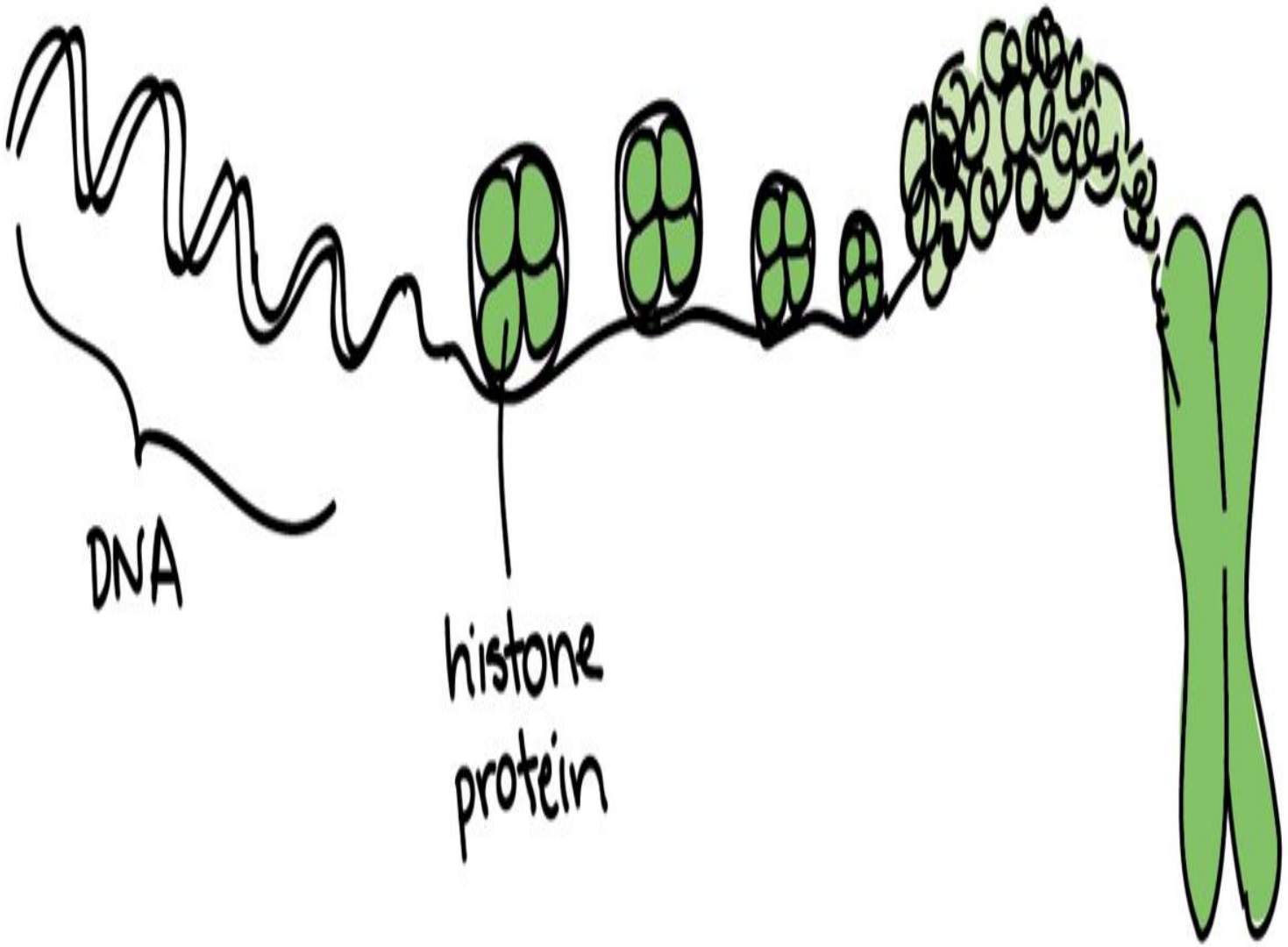
Mendel's model of inheritance

- Based on his results, Mendel came up with a model for the inheritance of individual characteristics, such as flower color.

- In Mendel's model, parents pass along “heritable factors," which we now call **genes**, that determine the traits of the offspring. Each individual has two copies of a given gene, such as the gene for seed color (*Y* gene) shown in figure in next slide. If these copies represent different versions, or **alleles**, of the gene, one allele—the **dominant** one—may hide the other allele—the **recessive** one. For seed color, the dominant yellow allele *Y* hides the recessive green allele *y*.

Genotype and Phenotype

- The set of alleles carried by an organism is known as its **genotype**.
- Genotype determines **phenotype**, an organism's observable features. When an organism has two copies of the same allele (say, *YY* or *yy*), it is said to be **homozygous** for that gene. If, instead, it has two different copies (like *Yy*), we can say it is **heterozygous**.
- Phenotype can also be affected by the environment in many real-life cases, though this did not have an impact on Mendel's work.

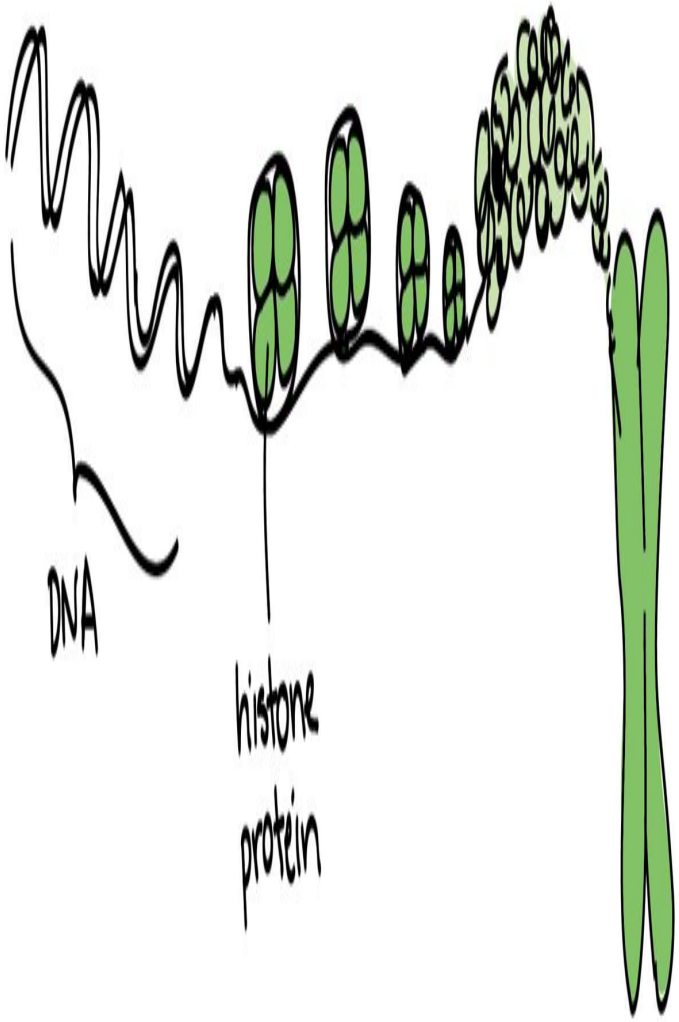

DNA
histone protein

Mendel's model: The law of segregation

- So far, so good. But this model alone doesn't explain why Mendel saw the exact patterns of inheritance he did. In particular, it doesn't account for the 3:1ratio. For that, we need Mendel's law of segregation.

- According to the **law of segregation**, **only one of the two gene copies present in an organism is distributed to each gamete (egg or sperm cell) that it makes, and the allocation of the gene copies is random**. When an egg and a sperm join in fertilization, they form a new organism, whose genotype consists of the alleles contained in the gametes. The diagram in the next slide illustrates this idea:

The law of segregation



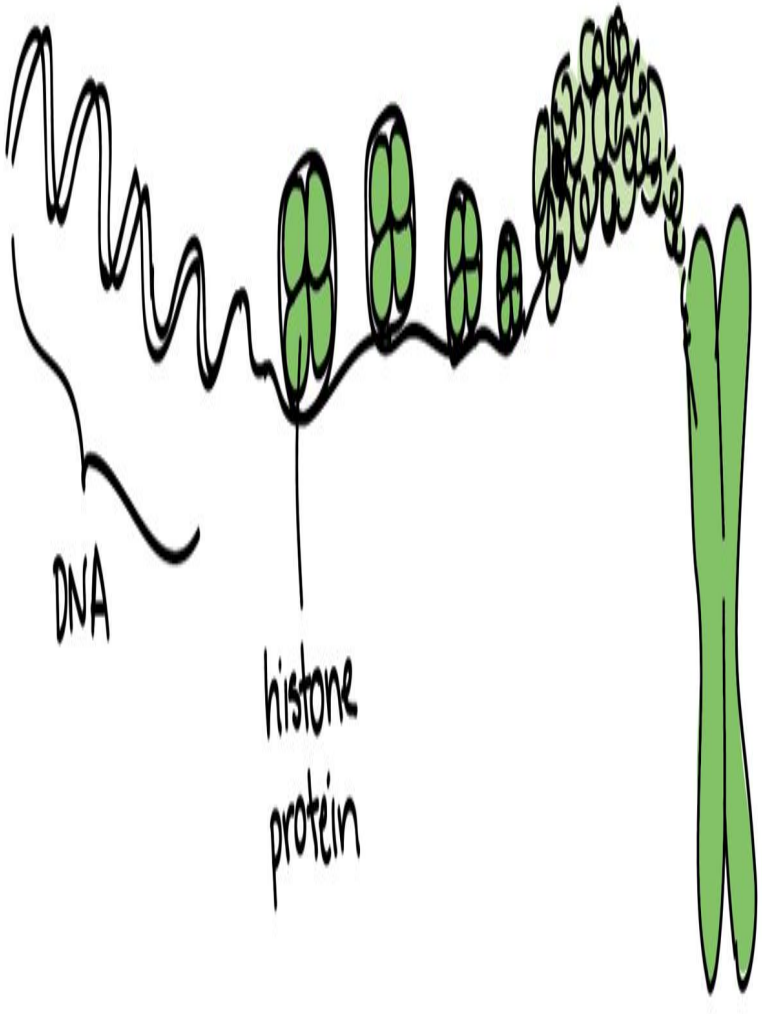
- The four-squared box shown for the F2 generation is known as a **Punnett square**.
- To prepare a Punnett square, all possible gametes made by the parents are written along the top (for the father) and side (for the mother) of a grid. Here, since it is self-fertilization, the same plant is both mother and father.
- The combinations of egg and sperm are then made in the boxes in the table, representing fertilization to make new individuals. Because each square represents an equally likely event, we can determine genotype and phenotype ratios by counting the squares.

- A key point of the law of segregation is that a parent's two gene copies are randomly distributed to its gametes. Thus, for a *Yy* heterozygote, *Y* and *y*gametes are equally likely to be made: 50% of the sperm and eggs will have a *Y* allele, 50% will have a *y* allele, and the same will be true for eggs.
- Since each type of gamete is equally common, each fertilization event (meeting of gametes, corresponding to a square of the table) also has an equal chance of happening. Thus, the boxes of the table represent four equal-probability events.
- Since the table contains 1 box with a *YY* genotype, 2 boxes with a *Yy* genotype, and 1 box with a *yy* genotype, we'd expect to see *YY*, *Yy*, and *yy*plants in a ratio of 1:2:1 in the F2 generation. Since both *YY* and *Yy* plants are yellow, this genotype ratio translates into a phenotype ratio of 3:1 yellow-seeded to green-seeded plants, almost exactly what Mendel observed.

The test cross

- Mendel also came up with a way to figure out whether an organism with a dominant phenotype (such as a yellow-seeded pea plant) was a heterozygote (*Yy*) or a homozygote (*YY*). This technique is called a **test cross** and is still used by plant and animal breeders today.
- In a test cross, the organism with the dominant phenotype is crossed with an organism that is homozygous recessive (e.g.,

Test cross



- If the organism with the dominant phenotype is homozygous, then all of the F1 offspring will get a dominant allele from that parent, be heterozygous, and show the dominant phenotype.
- If the organism with the dominant phenotype organism is instead a heterozygote, the F1 offspring will be half heterozygotes (dominant phenotype) and half recessive homozygotes (recessive phenotype).
- The fact that we get a 1:1 ratio in this second case is another confirmation of Mendel's law of segregation.

The law of independent assortment

- The law of segregation lets us predict how a single feature associated with a single gene is inherited. In some cases, though, we might want to predict the inheritance of two characteristics associated with two different genes. How can we do this?
- To make an accurate prediction, we need to know whether the two genes are inherited independently or not. That is, we need to know whether they "ignore" one another when they're sorted into gametes, or whether they "stick together" and get inherited as a unit.
- When Gregor Mendel asked this question, he found that different genes were inherited independently of one another, following what's called the **law of independent assortment**.

What is the law of independent assortment?

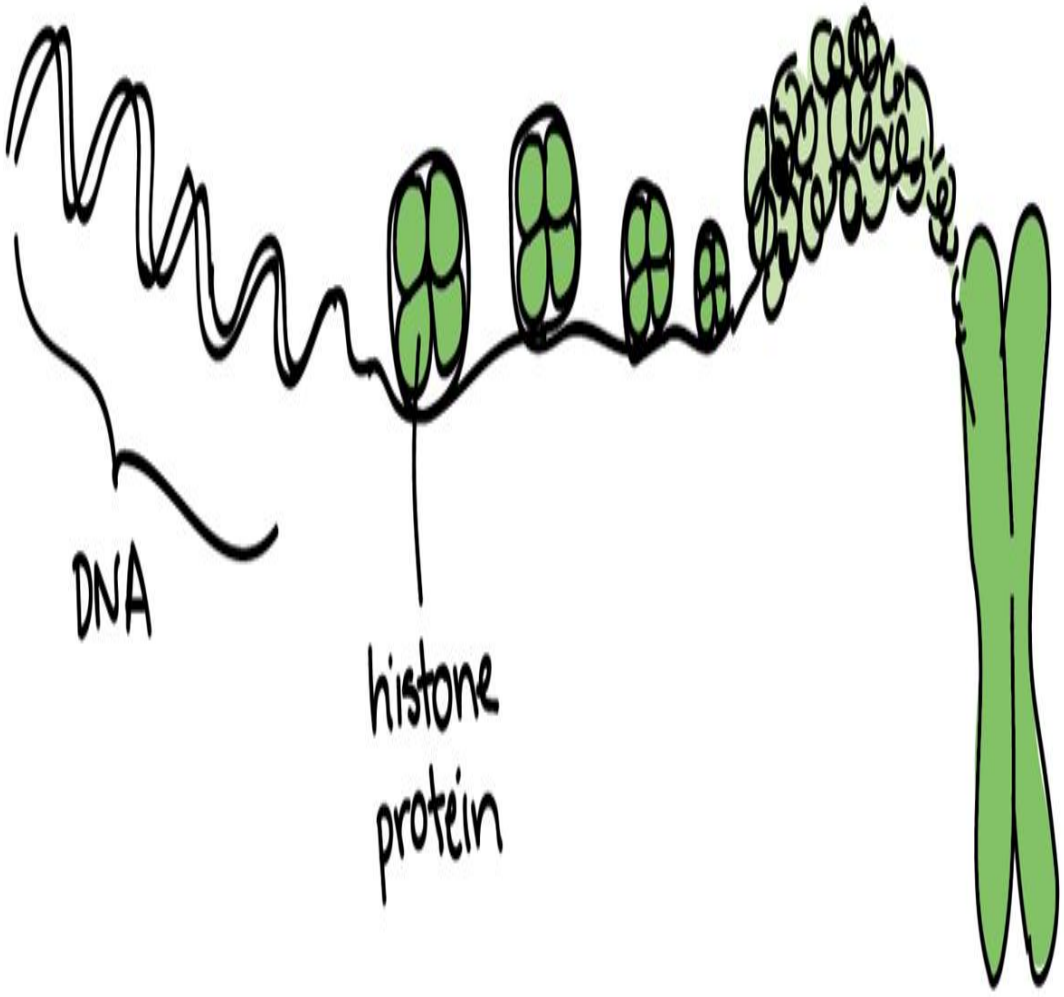
- Mendel's **law of independent assortment** states that the alleles of two (or more) different genes get sorted into gametes independently of one another. In other words, the allele a gamete receives for one gene does not influence the allele received for another gene.

Example: Pea color and pea shape genes

- Let's look at a concrete example of the law of independent assortment. Imagine that we cross two pure-breeding pea plants: one with yellow, round seeds (*YYRR*) and one with green, wrinkled seeds (*yyrr*).
- Because each parent is homozygous, the law of segregation tells us that the gametes made by the wrinkled, green plant all are *ry*, and the gametes made by the round, yellow plant are all *RY*. That gives us F1 offspring that are all *RrYy*.

Dihybrid cross

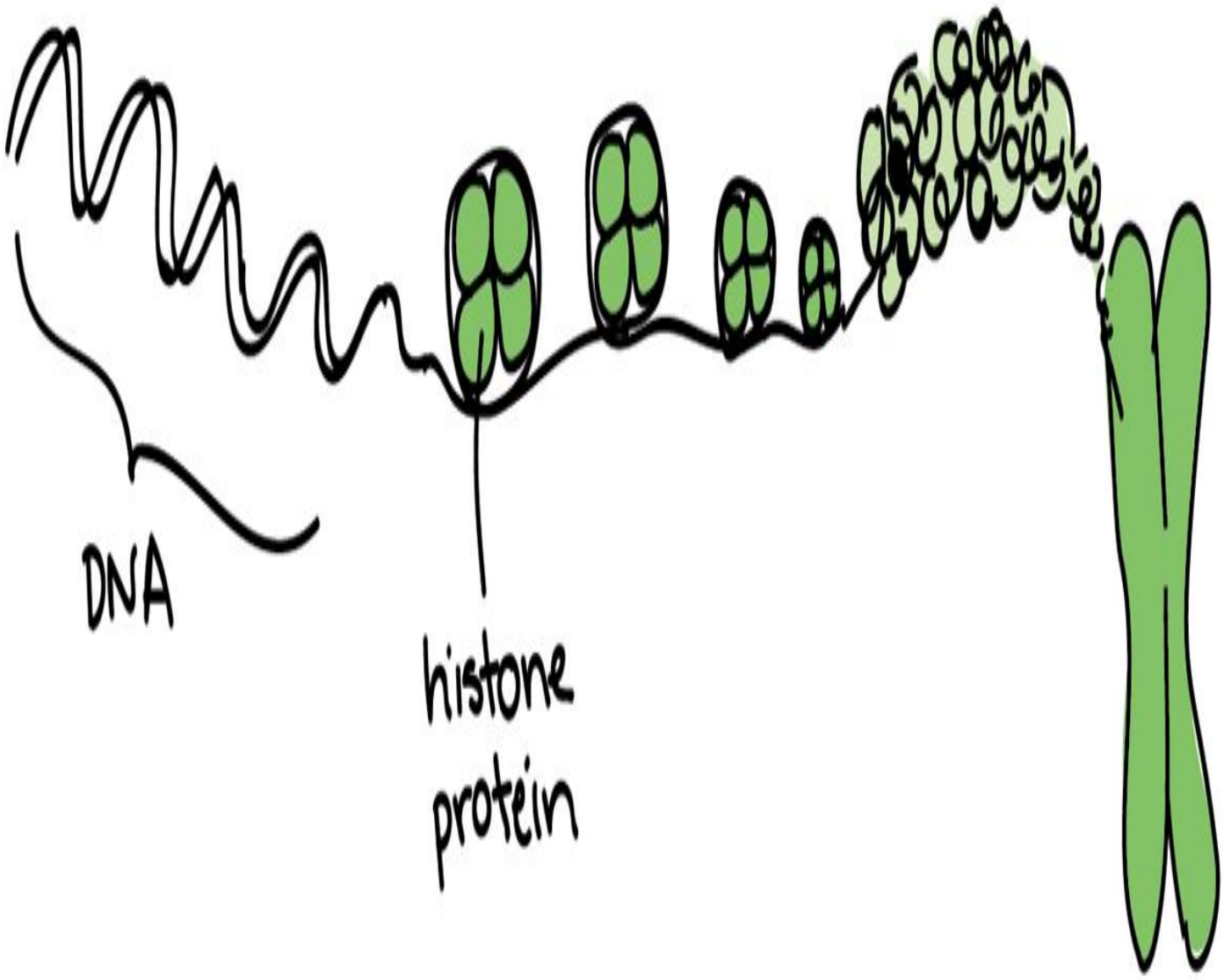
- The allele specifying yellow seed color is dominant to the allele specifying green seed color, and the allele specifying round shape is dominant to the allele specifying wrinkled shape, as shown by the capital and lower-case letters. This means that the F1 plants are all yellow and round. Because they are heterozygous for two genes, the F1 plants are called **dihybrids** (*di-* = two, *-hybrid* = heterozygous).
- A cross between two dihybrids (or, equivalently, self-fertilization of a dihybrid) is known as a **dihybrid cross**. When Mendel did this cross and looked at the offspring, he found that there were four different categories of pea seeds: yellow and round, yellow and wrinkled, green and round, and green and wrinkled. These **phenotypic** categories (categories defined by observable traits) appeared in a ratio of approximately 9:3:3:1


DNA
histone protein

- This ratio was the key clue that led Mendel to the law of independent assortment. That's because a 9:3:3:1 ratio is exactly what we'd expect to see if the F1 plant made four types of gametes (sperm and eggs) with equal frequency: *YR*, *Yr*, *yR*, and *yr*.
- In other words, this is the result we'd predict if each gamete randomly got a *Y* or *y* allele, and, in a separate process, also randomly got an *R* or *r* allele (making four equally probable combinations).
- We can confirm the link between the four types of gametes and the 9:3:3:1 ratio using the Punnett square. To make the square, we first put the four equally probable gamete types along each axis. Then, we join gametes on the axes in the boxes of the chart, representing fertilization events. The 16 equal-probability fertilization events that can occur among the gametes are shown in the 16 boxes. The offspring genotypes in the boxes correspond to a 9:3:3:1, ratio of phenotypes, just as Mendel observed.

Independent assortment vs. linkage

- The section above gives us Mendel's law of independent assortment in a nutshell, and lets us see how the law of independent assortment leads to a 9:3:3:1 ratio. But what was the alternative possibility? That is, what would happen if two genes *didn't* follow independent assortment? In the extreme case, the genes for seed color and seed shape might have always been inherited as a pair. That is, the yellow and round alleles might always have stayed together, and so might the green and wrinkled alleles.
- To see how this could work, imagine that the color and shape genes are physically stuck together and cannot be separated, as represented by the boxes around the alleles in the diagram below. For instance, this could happen if the two genes were located very, very close together on a chromosome


DNA
histone protein

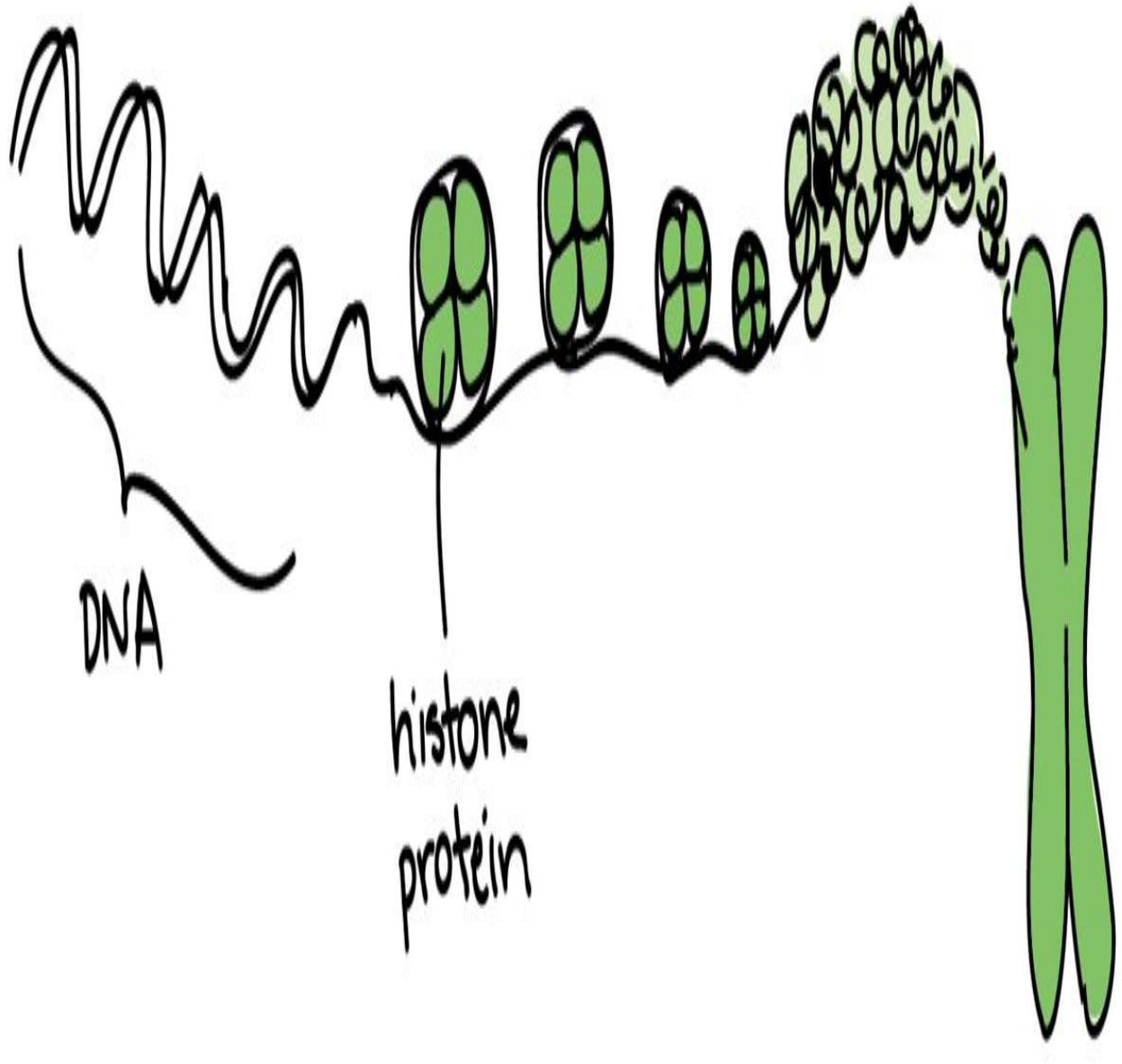
- Rather than giving a color allele and, separately, giving a shape allele to each gamete, the F1 F1start text, F, end text, start subscript, 1, end subscript dihybrid plant would simply give one “combo unit” to each gamete: a *YR* allele pair or a *yr* allele pair.
- We can use a Punnett square to predict the results of self-fertilization in this case as shown in previous slide. If the seed color and seed shape genes were in fact always inherited as a unit, or **completely linked**, a dihybrid cross should produce just two types of offspring, yellow/round and green/wrinkled, in a 3:1 ratio.
- Mendel's actual results were quite different from this telling him that the genes assorted independently.

The reason for independent assortment

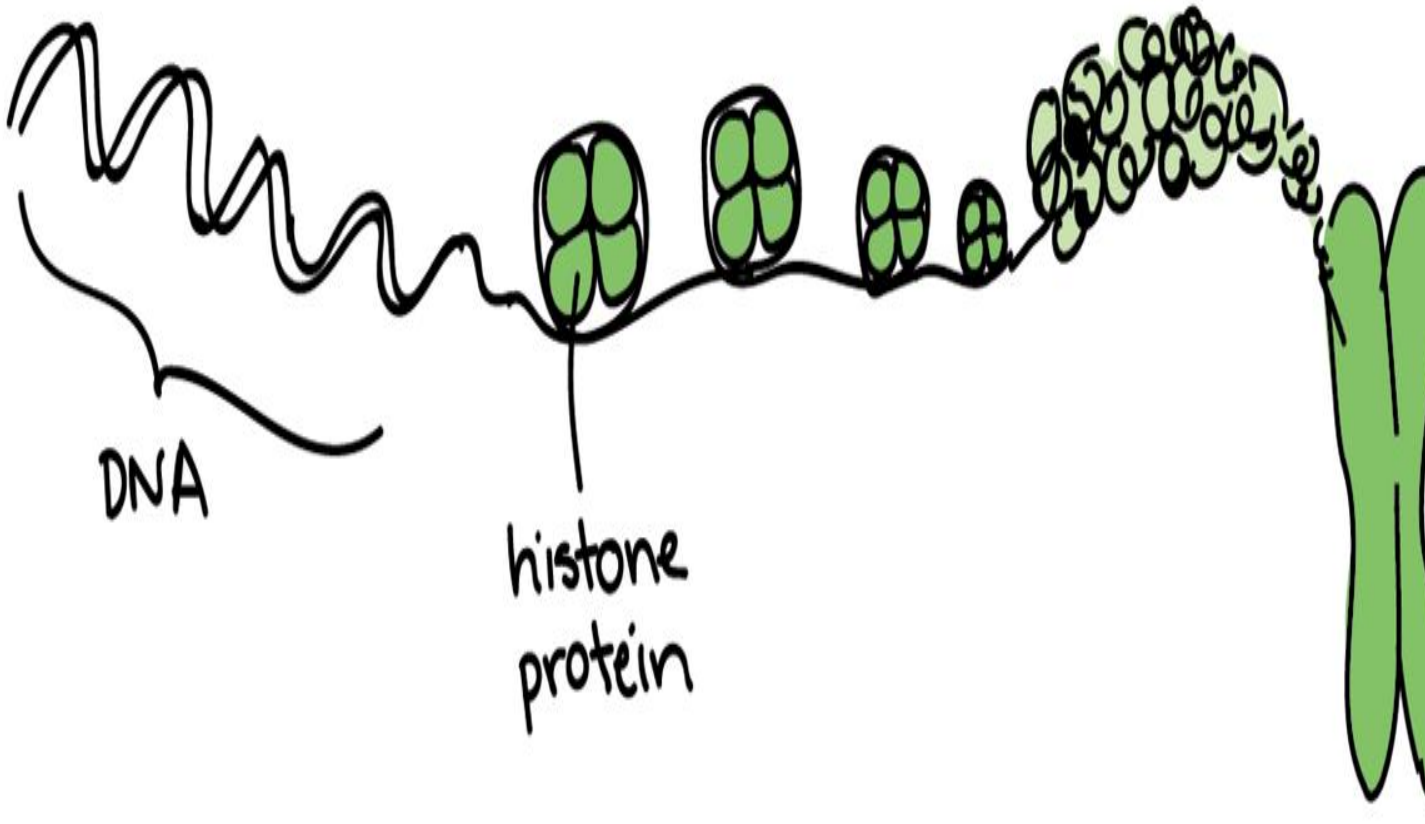
- To see why independent assortment happens, we need to fast-forward half a century and discover that genes are physically located on chromosomes. To be exact, the two copies of a gene carried by an organism (such as a *Y* and a *y*allele) are located at the same spot on the two chromosomes of a **homologous pair**. Homologous chromosomes are similar but non-identical, and an organism gets one member of the pair from each of its two parents.

- The physical basis for the law of independent assortment lies in meiosis I of gamete formation, when homologous pairs line up in random orientations at the middle of the cell as they prepare to separate. We can get gametes with different combos of "mom" and "dad" homologues (and thus, the alleles on those homologues) because the orientation of each pair is random.

- To see what this means, compare chromosome arrangement 1 (top) and chromosome arrangement 2 (bottom) at the stage of metaphase I in the diagram below. In one case, the red "mom" chromosomes go together, while in the other, they split up and mix with the blue "dad" chromosomes. If meiosis happens many times, as it does in a pea plant, we will get both arrangements—and thus *RY*, *Ry*, *rY*, and *ry* classes of gametes—with equal frequency.
- Genes that are on different chromosomes (like the *Y* and *R* genes) assort independently. The seed color and seed shape genes are on chromosomes 1 and 7 of the pea genome, respectively, in real life11start superscript, 1, end superscript. Genes that are far apart on the same chromosome also assort independently thanks to the crossing over, or exchange of homologous chromosome bits, that occurs early in meiosis I.
- There are, however, gene pairs that do not assort independently. When genes are close together on a chromosome, the alleles on the same chromosome tend to be inherited as a unit more frequently than not. Such genes do not display independent assortment and are said

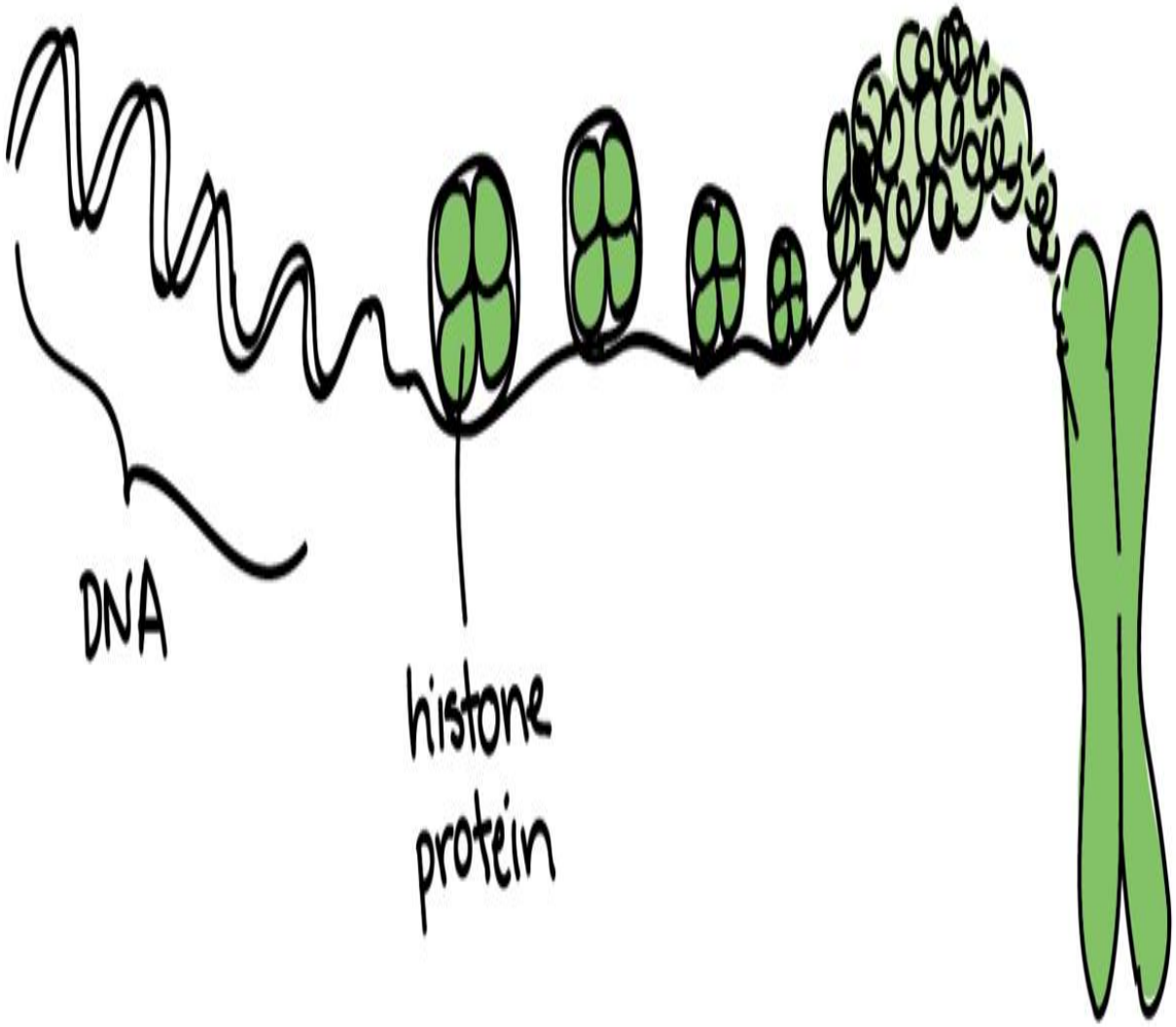

DNA
histone
protein

- Suppose you cross a pure-breeding, black-coated dog with curly fur to a pure-breeding, yellow-coated dog with straight fur. In the F1 generation, all the puppies have straight, black coats. Next, you interbreed the F1 dogs with one another to get an F2 generation. If coat color and coat texture are controlled by two genes that assort independently, what fraction of the F2 puppies are expecte



- Since all of the F1 dogs are black and straight-furred, we know that black fur color and straight fur texture are dominant over yellow fur color and curly fur texture. If we call the color gene *B/b* and the texture gene *C/c*, and use capital letters for the dominant form of each gene and lowercase letters for the recessive form, we can assign the two parental dogs genotypes of *BBcc* (black and curly) and *bbCC* (yellow and straight-furred). When the parental dogs are crossed, they produce black, straight-furred F1 dogs that are dihybrids: *BbCc*.

- A cross between two F1 dihybrid dogs results in the Pur



- The F1 dogs can make four different types of gametes, which are represented along the two axes of the Punnett square. The squares of the table represent fertilization events in which the gametes on the axes combine. Since all of the gamete types are equally likely to be produced (because the genes assort independently, i.e., do not influence each other's inheritance), all the squares in the table represent equal-probability events, ones that occur 1/16 of the time.

- Now, we need to find the squares that correspond to the outcome we are interested in: a puppy with yellow, straight fur. To have yellow, straight fur, the puppy must get two recessive alleles for fur color (*bb*genotype) and at least one dominant allele for fur texture (*Cc* or *CC*genotype).
- If we go through the table and circle the genotypes that match these requirements, we'll find that 3 out of the 16 boxes correspond to yellow, straight-furred puppies. Thus, we would expect 3/16 F2 puppies to have yellow, straight fur.

- Imagine that you are a rabbit breeder with two purebred rabbits, a male with black fur and a female with tan fur. When you cross your rabbits, all of the F1F1start text, F, end text, start subscript, 1, end subscript kits (baby rabbits) have tan fur. **Which trait is dominant, and which is recessive?**

- When two pure-breeding organisms with different forms of a characteristic are crossed, the offspring of the cross may show just one form of the characteristic, while the other form may be hidden.
- The trait that is visible in the offspring of the cross is called the dominant trait, while the trait that is hidden is called the recessive trait.
- In this example, tan fur is the dominant trait, because it is visible in the offspring of the cross. Black fur is the recessive trait, because it is hidden in the offspring of the cross.
- **Tan fur is dominant, while black fur is recessive.**

- **Which statement best describes the relationship between genotype and phenotype?**
- Genotype always completely determines phenotype
- Phenotype always completely determines genotype
- Genotype determines phenotype, but with influence from the environment
- Phenotype determines genotype, but with influence from the environment