Chapter 14 Lecture Notes: Nucleic Acids

Educational Goals

1. Know the *three* chemical components of a **nucleotide**: a *monosaccharide residue* (either ribose or deoxyribose), at least one *phosphate group*, and an "*organic base*."
2. Identify **phosphoester** bonding patterns and ***N*-glycosidic bonds** within *nucleotides*.
3. Compare and contrast **ribonucleotides** and **deoxyribonucleotides**.
4. Understand the bonding patterns within **nucleotide diphosphates** and **nucleotide triphosphates**.
5. Predict the products for the hydrolysis of **ATP** and **ADP**.
6. Describe the bonding patterns within **cyclic nucleotides**.
7. Describe the chemical structure of **polynucleotides** (DNA and RNA) and identify the **phosphodiester bonding patterns** within *polynucleotides*.
8. Draw the structural formula of a **dinucleotide** that is formed by combining two specified nucleotides.
9. Compare and contrast the monosaccharide residues present in DNA vs. RNA.
10. Know the names of the **organic bases** that are present in DNA and RNA.
11. Given the structure of a DNA or RNA strand, identify the **sugar-phosphate backbone**, the **3' terminus**, and the **5' terminus**.
12. Describe the DNA **double helix structure**, and understand and define the term "**complementary base pairing**."
13. Understand how DNA folds back on itself and wraps around *histones* to form **chromatin**. Define, compare, and contrast **chromatin**, **genes**, and **chromosomes**.
14. Understand and explain how **DNA replication** takes place and how **DNA polymerase** is involved in the replication process.
15. Describe the **polymerase chain reaction**.
16. Know how **DNA fingerprinting** works and how it is used in forensic science.
17. Understand and explain **transcription** and how **RNA polymerase** is involved in the transcription process.
18. Understand and explain **translation** and define the term "**codon**."
19. Given the primary structure of DNA or mRNA, use the **genetic code** table to predict the sequence of amino acids in the polypeptide that would be produced in translation.
20. Describe the ***three types of RNA*** and understand the role of each in translation.
21. Define the term "**gene expression**."
22. Define the term "**operon**." Describe how the ***lac*** **operon** is *regulated* in order to control *gene expression* in *E. coli*.
23. Describe the structure of **viruses** and understand the way in which they can cause infections/diseases. Understand how the **reverse transcriptase** enzyme is used by HIV viruses.
24. Define and understand the terms "**mutation**," "**genome**," and "**gene therapy**." Explain how mutations can lead to **genetic diseases**. Give an example of a **monogenic disease**.
25. Define the terms: **recombinant DNA**, **genetically modified organism (GMO)**, **transfection**, and **bioethics**.

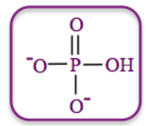
____________________ ____________ is a term used for the class of biological polymers consisting of **d**eoxy**r**ibonucleic **a**cid (DNA) ***and*** **r**ibo**n**ucleic **a**cid (RNA).

- Nucleic acids are ***polymers*** that consist of _________________________residues.

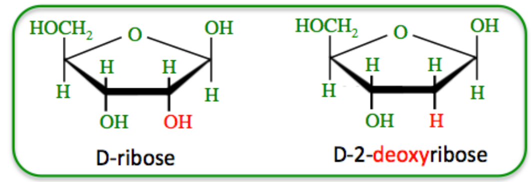
Nucleotides

A ***nucleotide*** is formed from *three* chemical components:

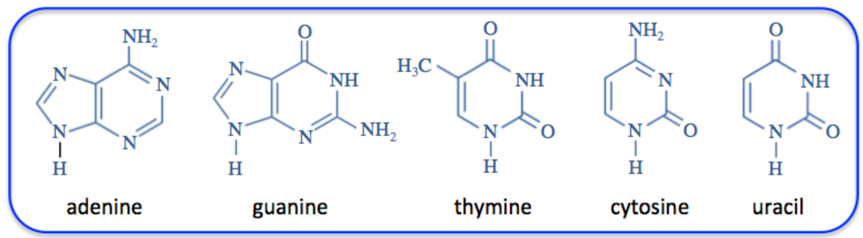
1) at least one ____________________ (hydrogen phosphate ion)
2) a ___________________________ (either ribose *or* deoxyribose)
3) an ______________________ base



a phosphate
(hydrogen phosphate ion)

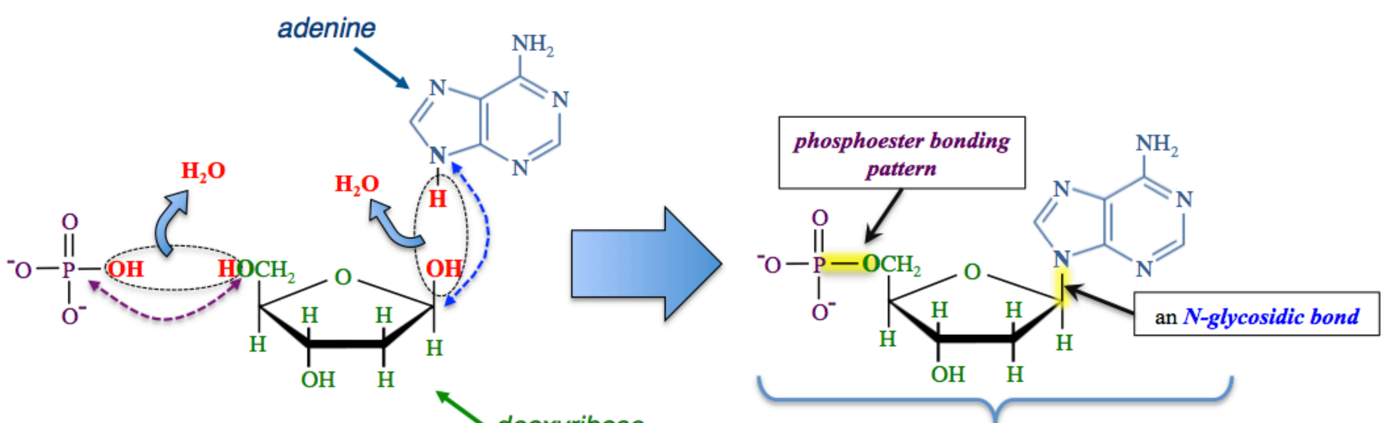


nucleotides contain one of these monosaccharides



nucleotides contain one organic base
(the five most common organic bases are shown)

Example: a deoxyribonucleotide



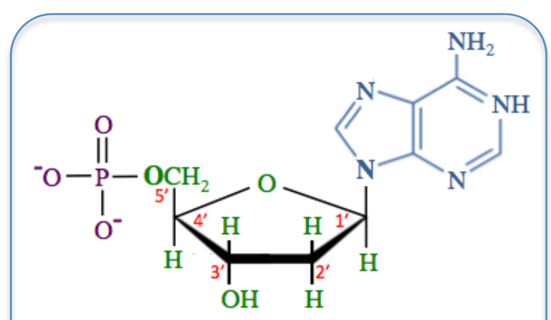
a nucleotide

I used *adenine* as the *organic base* and *deoxyribose* as the *monosaccharide.*

The ___________________ bonding pattern occurs when a monosaccharide bonds with the phosphate group.

An ____________________is made when the monosaccharide bonds to a nitrogen in the *organic base*.

In this example, the *phosphate residue* is bonded to the carbon in position number **5'** of the monosaccharide ring.



***deoxyadenosine* 5'-monophosphate**

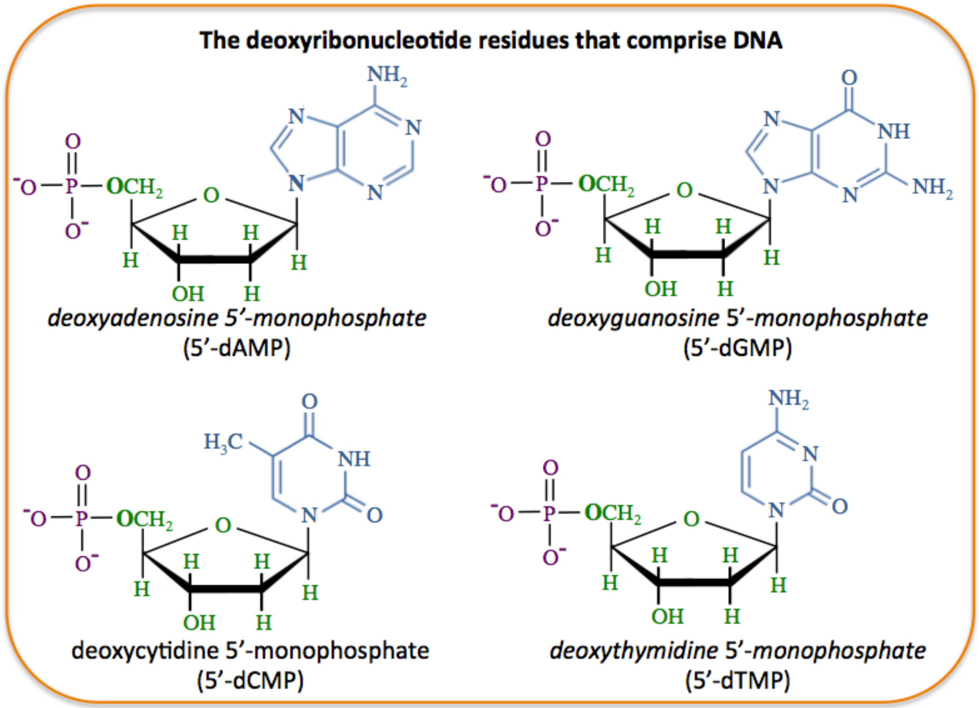
Organisms synthesize DNA from **5'**-nucleotides.

Because this nucleotide is formed from one *phosphate* residue, a *deoxyribose* residue, and an *adenine* organic base, we write its name as "*deoxyadenosine* **5'**-monophosphate."

It is common to use abbreviations when naming nucleotides. In this case, we would abbreviate the name as "**5'-d**AMP," the lowercase "**d**" indicates "*deoxy*."

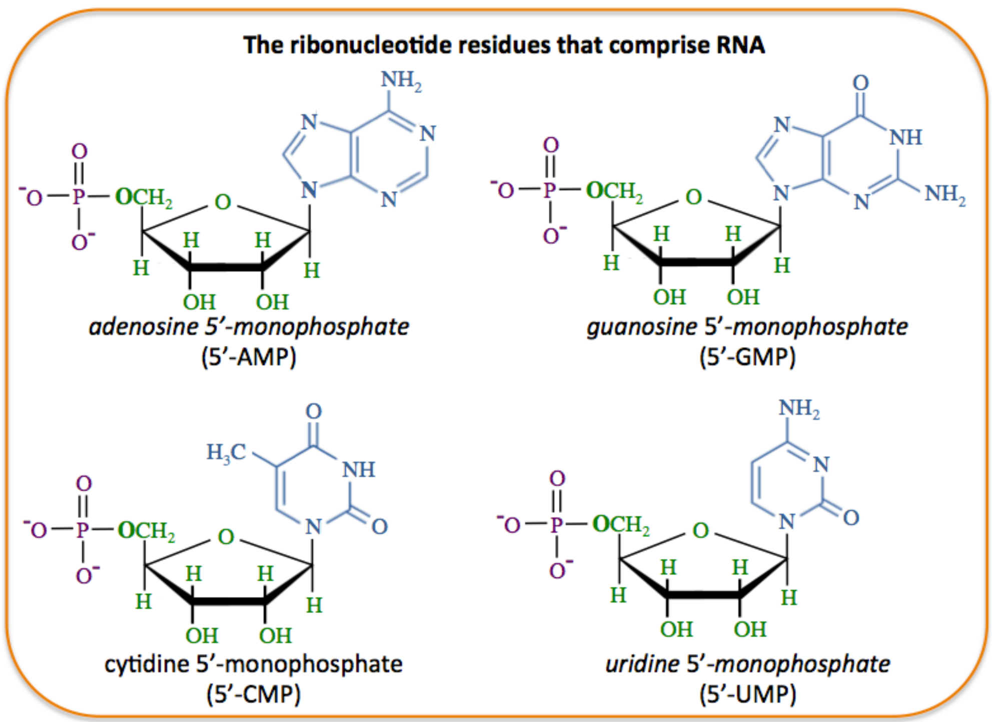
Deoxyribonucleotides

The nucleotides that make up **DNA** contain a _______________ residue, a *phosphate* residue, and either *adenine*, *guanine*, *cytosine*, or *thymine* organic base residues.



Ribonucleotides

The nucleotides that make up **RNA** contain a ____________ residue, a *phosphate* residue, and either *adenine*, *guanine*, *cytosine*, or *uracil* organic base residues.



Both DNA and RNA have *adenine, guanine*, and *cytosine* organic bases, **however** *thymine* is present only in DNA and *uracil* is present only in RNA.

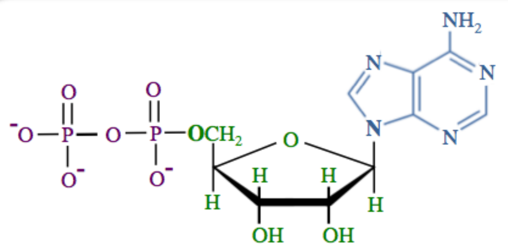
These **5'**-nucleotides are so frequently encountered in biology that the "**5**-" is usually omitted in the names and abbreviations. **3'**-nucleotides do exist, but they are rare.

Nucleotide Diphosphates and Triphosphates

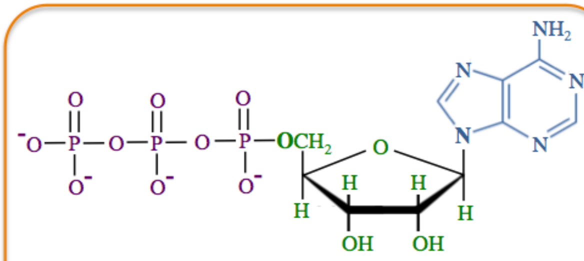
Nucleotides have biological roles other than forming DNA and RNA.

One of these roles involves ______________ ________________.

The two most important energy-transfer nucleotides are *adenosine triphosphate* (ATP) and *adenosine diphosphate* (ADP).



adenosine diphosphate (**ADP**)



adenosine triphosphate (**ATP**)

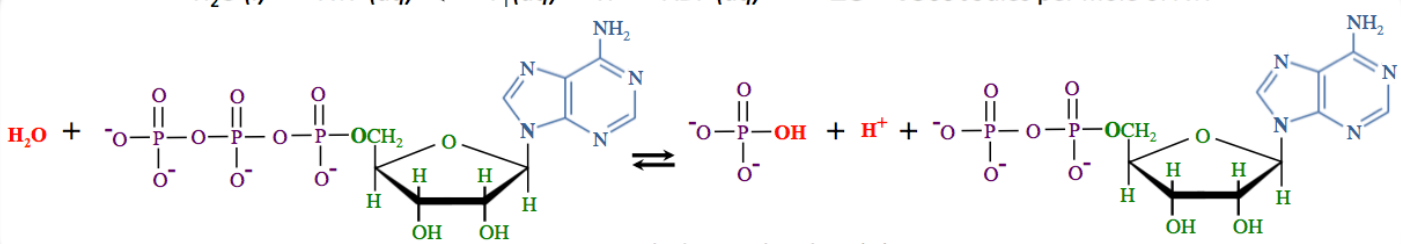
ATP contains ***more*** *chemical potential energy* than does ADP.

Organisms ***obtain energy*** from their environment (sunlight or chemical potential energy in food) when they use it to convert ADP to ATP.

Energy is ***released*** from ATP when it is converted to back to ADP.

- Organisms can do mechanical ***work,*** or drive chemical reactions that require energy by converting ATP to ADP. This is analogous to burning fuel in order to move a car, or discharging a battery while powering an electrical device. *One* way that energy can be released from ATP is by reacting it with H_2O to form ADP, hydrogen phosphate (abbreviated as P_i), and an H^+ ion. The chemical equation for this reaction is shown below.

$H_2O\ (l)$ + ATP (aq) $\rightleftharpoons$ $P_i\ (aq)$ + H^+ + ADP (aq) **ΔG = -7300** Joules per mole of ATP



adenosine triphosphate (**ATP**) *hydrogen phosphate* (**P_i**) *adenosine diphosphate* (**ADP**)

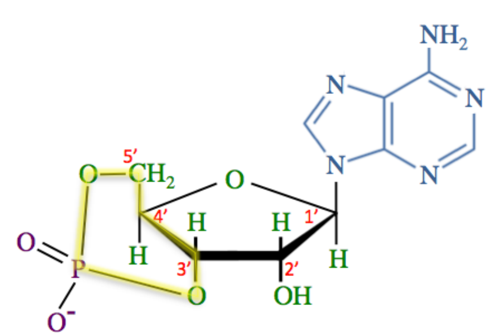
Cyclic Nucleotides

Another biological role of nucleotides is to act as _______________ ______________compounds.

A type of nucleotide, called a ________________ **nucleotide**, is often used for this purpose.

Cyclic nucleotides are nucleotides that contain a phosphate group *that bonds to the monosaccharide reside at two locations*, thereby forming a *phosphorus-containing* ring.

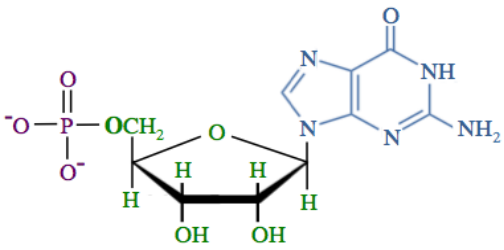
The lower-case "c" in the abbreviation (cAMP) indicates that it is a cyclic nucleotide.



cyclic adenosine monophosphate (cAMP)

Understanding Check: Draw the structural formula of ***cyclic*** guanosine monophosphate (cGMP) by making the phosphoester bonds to the **3'** carbon and the **5'** carbon of the ribose monosaccharide residue. Also, label the N-glycosidic bond.

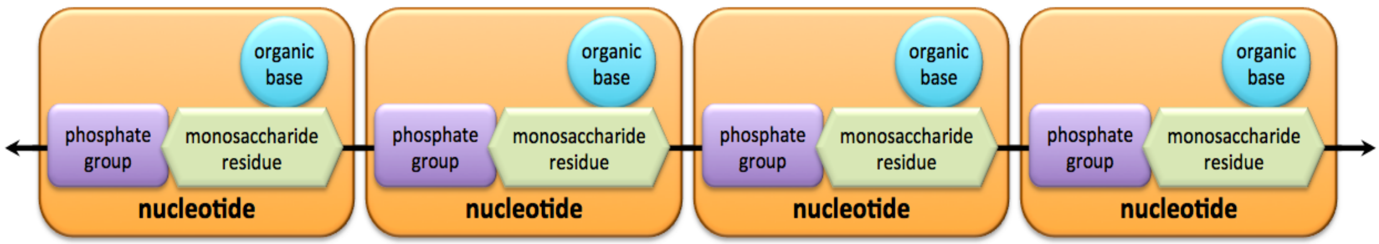
HINT: Look at the image of guanosine monophosphate (below) to see how *guanine* is connected to *ribose*.



guanosine 5'-monophosphate
(5'-GMP)

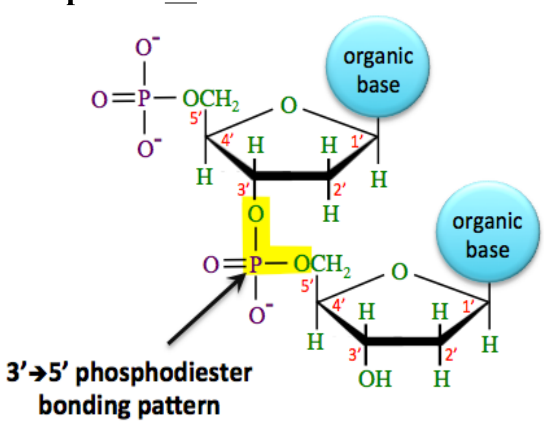
Polynucleotides: DNA and RNA

A ________________________ is a polymer made from ***nucleotide*** residues.



A chain of covalently bonded nucleotides is referred to as a "_______________."

Example: A <u>Di</u>nucleotide



Nucleotides are covalently bonded to each other by a _______________________ bonding pattern between a **3'** carbon on one monosaccharide and **5'** carbon on the other monosaccharide.

- We call this bonding pattern a 3'→ 5' *phosphodiester* because the phosphate group is between the 3' and 5' carbons.

Understanding Check: Starting with the dinucleotide, draw a *trinucleotide* by adding another nucleotide to the 3' carbon of the *bottom nucleotide residue*. Highlight the *phosphodiester* bonding patterns. You may use a circle to represent the organic base structures, as I did in the dinucleotide.

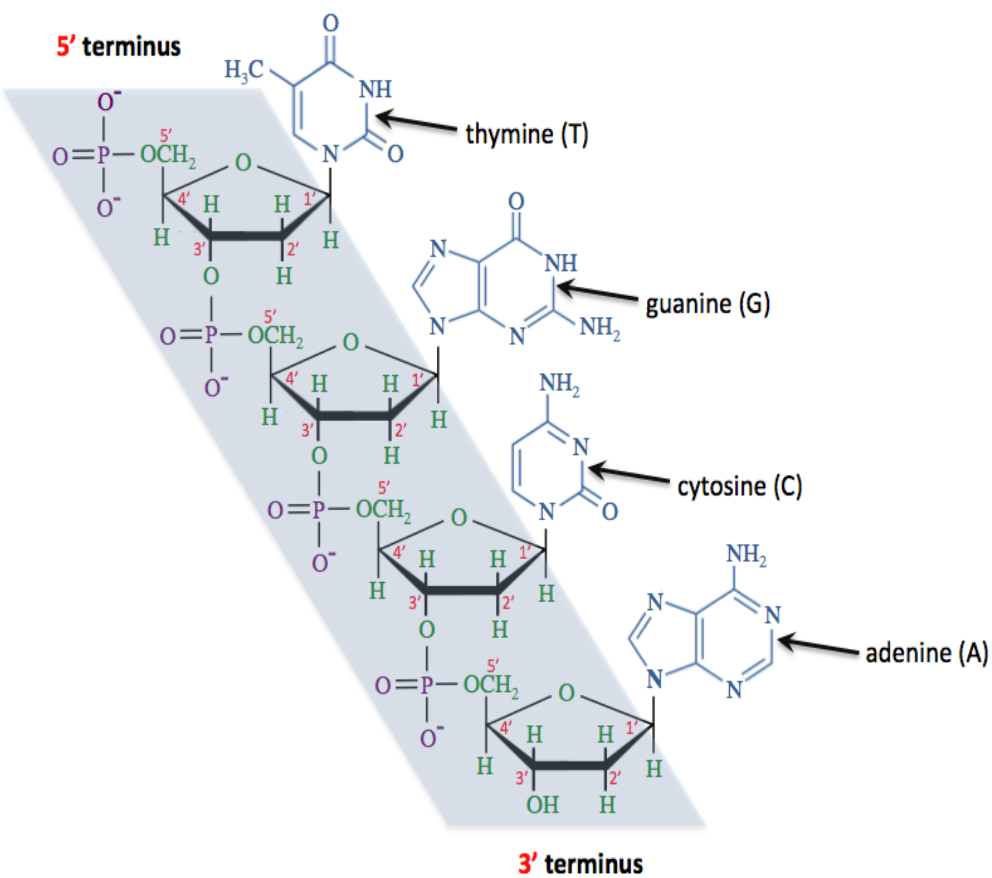
DNA Structure

DNA contains the "information" needed for life. This information enables cells to grow and divide. It is responsible for your physical characteristics, such as your height, skin tone, and eye color. Human DNA contains about three million deoxyribonucleotide residues. Eukaryotic cells, which are cells having membrane-bound nuclei, have most of their DNA in the nucleus, and small amounts in other organelles such as chloroplasts and mitochondria. The DNA of prokaryotic cells is contained in the cytoplasm.

Genetic information, the information used to make the various proteins and thereby enabling life, is contained in the ___________________ ***of nucleotides*** in DNA.

The sequence of nucleotides in DNA is referred to as **DNA's** ______________ ________________.

DNA is composed of a combination of ***deoxy****ribonucleotides* that contain either adenine, guanine, thymine, or cytosine organic bases.



For example:

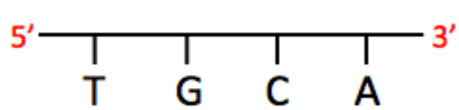
Nucleotides are given a one-letter abbreviation based on the **first letter in the name of their *organic base***, as shown in structure on the left.

In order to describe the primary structure of a DNA strand, it is customary to list the nucleotides' one-letter abbreviations in the order that the nucleotide residues appear in the strand.

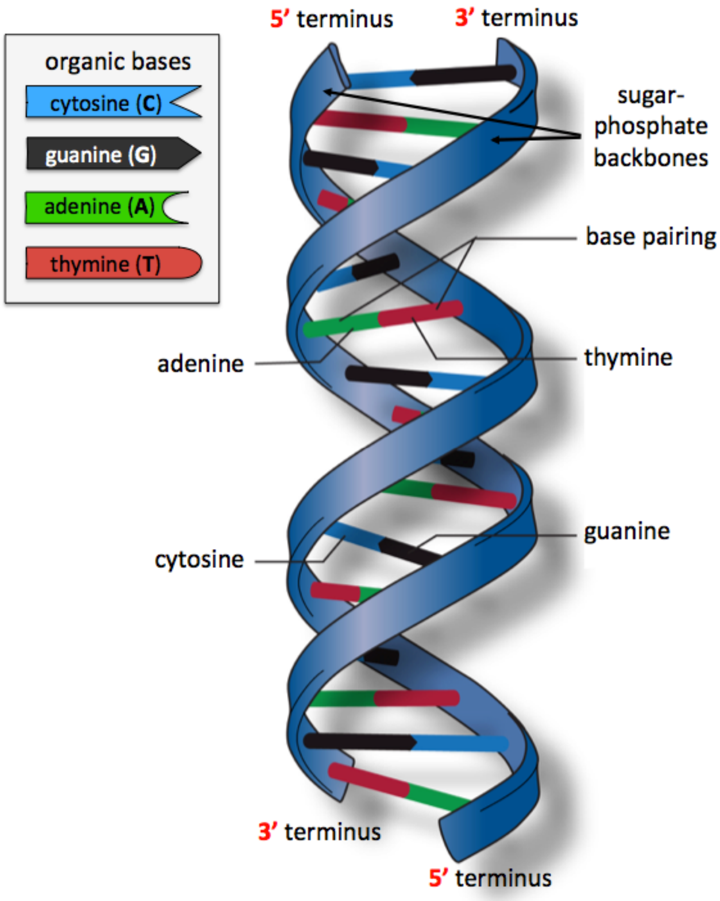
For example, the primary structure of the DNA strand drawn on the left is written as:

5'-TGCA-3'

Another common method used to represent the primary structure of DNA and RNA is to draw the one-letter abbreviations of the organic bases branching from a line that represents the phosphate-sugar backbone. For example, the DNA strand shown above is represented as:



DNA exists as a _____________ ____________, which is made from ***two** polynucleotide strands*.



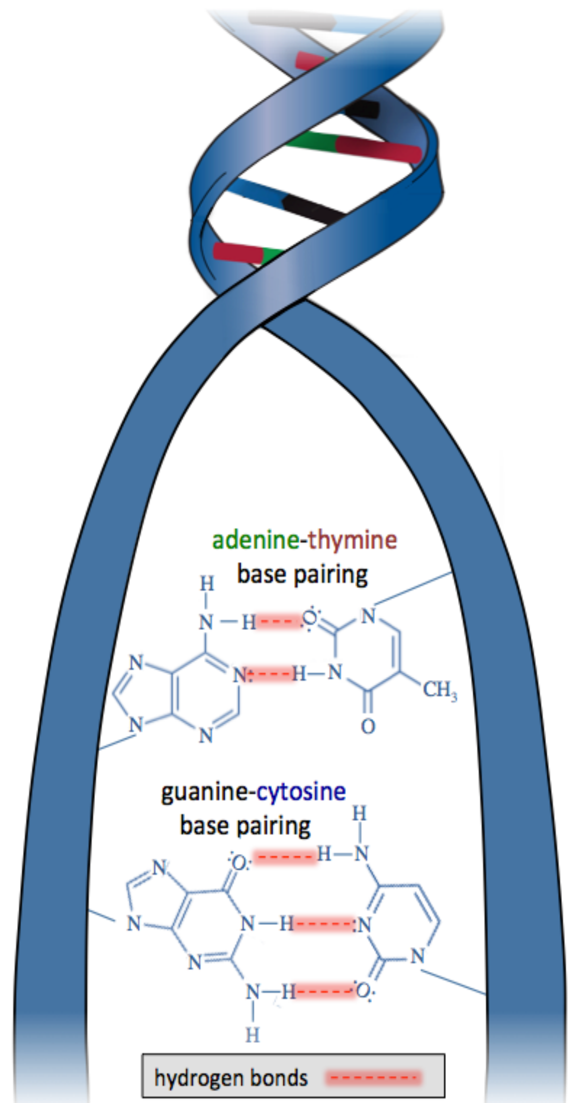
The sugar-phosphate backbones for each of the two DNA strands are illustrated as blue ribbons.

The two DNA strands of a DNA double helix have their **5'** terminuses and **3'** terminuses with opposite orientations. We call this opposing orientations of the DNA strands an "**antiparallel**" arrangement.

The two DNA strands are held together by especially strong ________________ ____________ between ***specific*** pairs of organic bases.

Adenine (**A**) bases hydrogen bond with *thymine* (**T**) bases.

Guanine (**G**) bases hydrogen bond with *cytosine* (**C**) bases.

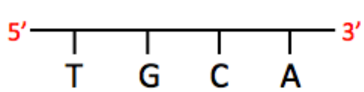


The reason for the strong hydrogen bonding between these organic base pairs is that they have _________________ *shapes*.

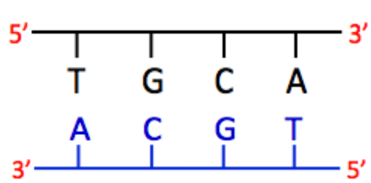
The hydrogen bonding between these pairs of organic bases is referred to as _________ _____________ *or* **complementary base pairing**.

- We say that the two polynucleotide strands in a double-stranded DNA particle are "***complementary***."

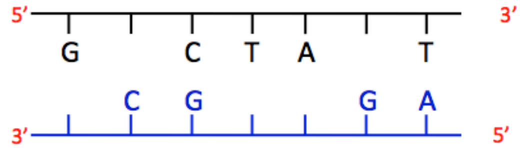
Example: Find the complement to this short DNA strand.



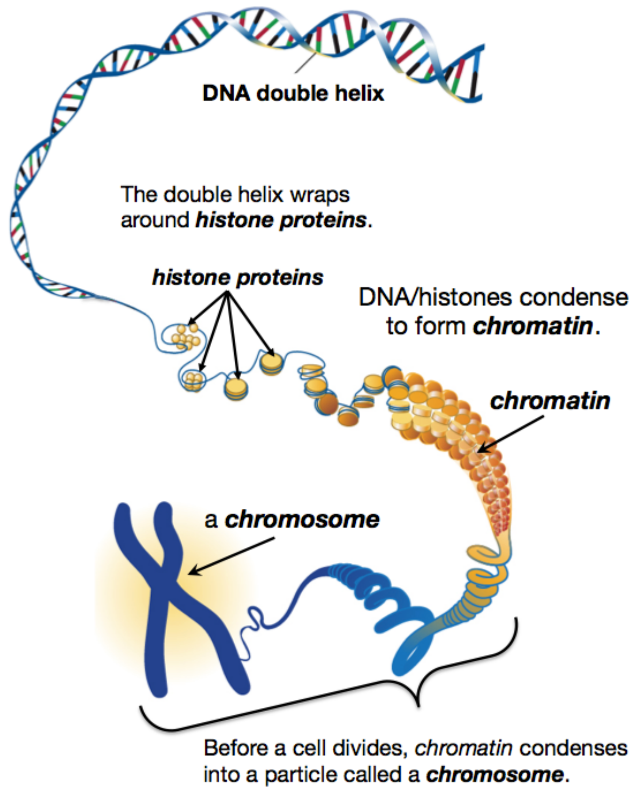
Solution:



Understanding Check: Fill in the missing one-letter abbreviations for the organic bases in the image below to show the correct complementary base pairing.



DNA Structure: Chromatin and Chromosomes



In eukaryotic organisms (plants, animals, and fungi), the DNA double helix coils into a more compact structure, as illustrated in the figure on the right.

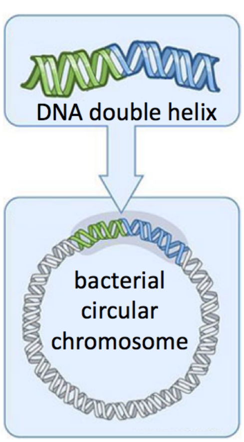
In the coiling process, the double helix wraps around ____________ proteins, as shown in figure on the right.

The DNA/histone further condenses into a fiber that is called __________________.

Depending on the cell life cycle, *chromatin* can undergo further coiling. For example, before a cell divides chromatin is coiled into its tightest, *most compact shape*, which is called a
______________________.

The amount of *chromosomes* that a cell contains depends on the organism. Humans have a total of 46 chromosomes per cell, 23 from each parent. Dogs have a total of 78 chromosomes, 39 from each parent. Fruit flies have a total of 8 chromosomes. The smallest human chromosome contains about 50 million base pairs, and the largest one contains about 250 million base pairs.

Bacterial DNA



In *bacteria*, which do not have nuclei, chromosomes exist as circular units, called __________________ ***chromosomes***, (illustrated on the left).

- Circular chromosomes can twist/fold into more compact shapes.

In addition to *circular chromosomes*, bacteria contain smaller, *yet still circular*, double-stranded DNA units called ________________.

Denaturation of the DNA Double Helix

Disruption of hydrogen bonding between complementary base pairs is called _____ ______________.

- If enough disruption occurs, the DNA strands of a double helix will completely ______________ from each other.
- DNA can be denatured by the same agents that are responsible for the denaturation of proteins.

Unlike *protein* denaturation, when *DNA* is denatured by heat and then allowed to cool to room temperature, the DNA strands will _______________ into the original double helix shape.

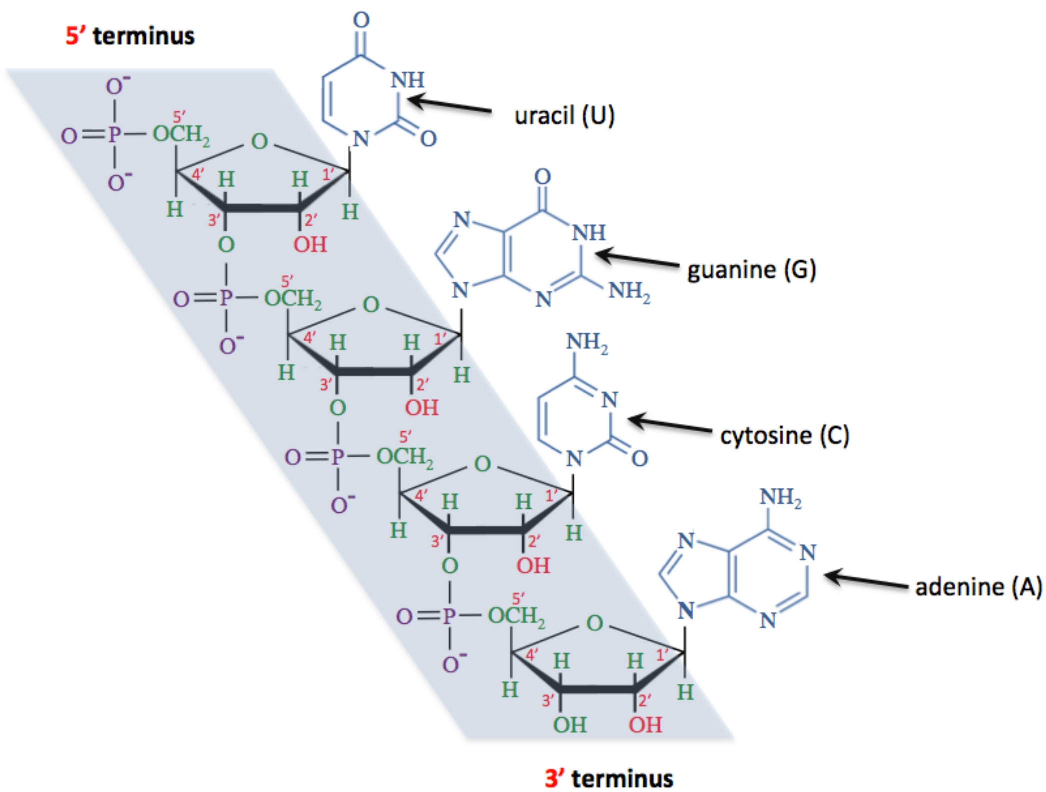
- When complementary DNA strands re-form their double helix structure, it is referred to as _________.
- Annealing of DNA strands is possible because of complementary base pairing.

RNA Primary Structure

RNA is a polynucleotide that contains ***ribose*** residues, phosphate groups, and adenine, guanine, cytosine and uracil organic bases.

The sequence of nucleotides in RNA is referred to as **RNA's primary structure**.

The primary structure of a small section of RNA is shown below.

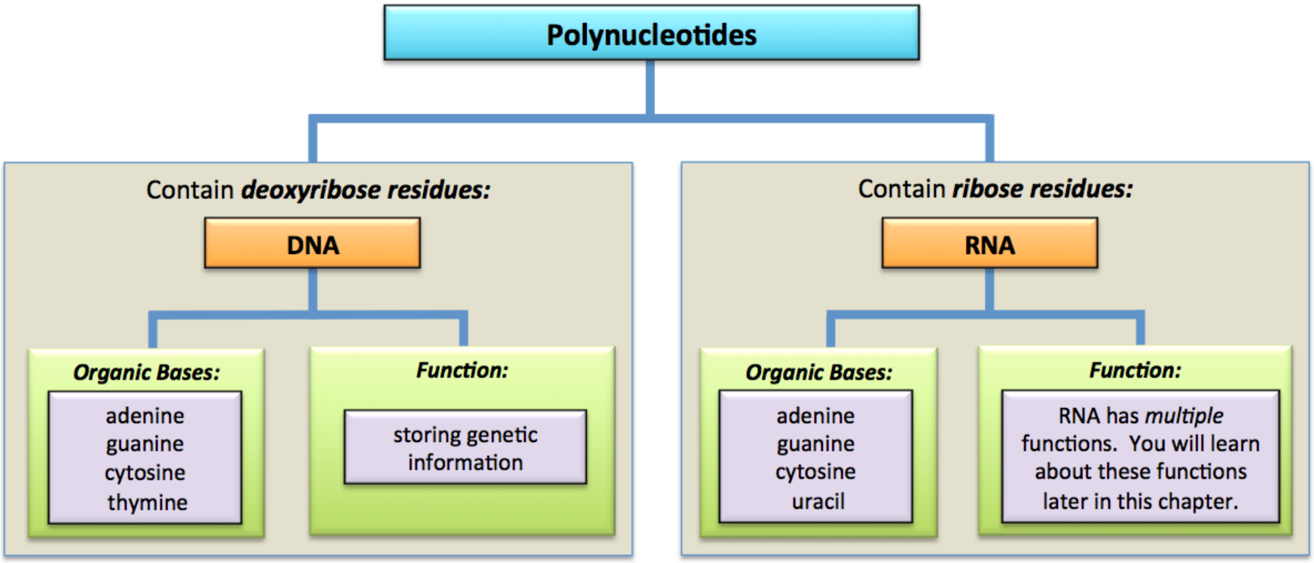


DNA contains ***thymine*** organic bases, however **RNA** does not; it contains _______ organic bases instead.

DNA has just one function - storing genetic information; **RNA** has __________ functions.

- The overall shape of an RNA particle depends on its ______________.
- You will learn about the three types of RNA that are involved in protein synthesis in a later section of this chapter.

Differences Between the Two Classes of Polynucleotides: RNA vs. DNA



DNA Replication

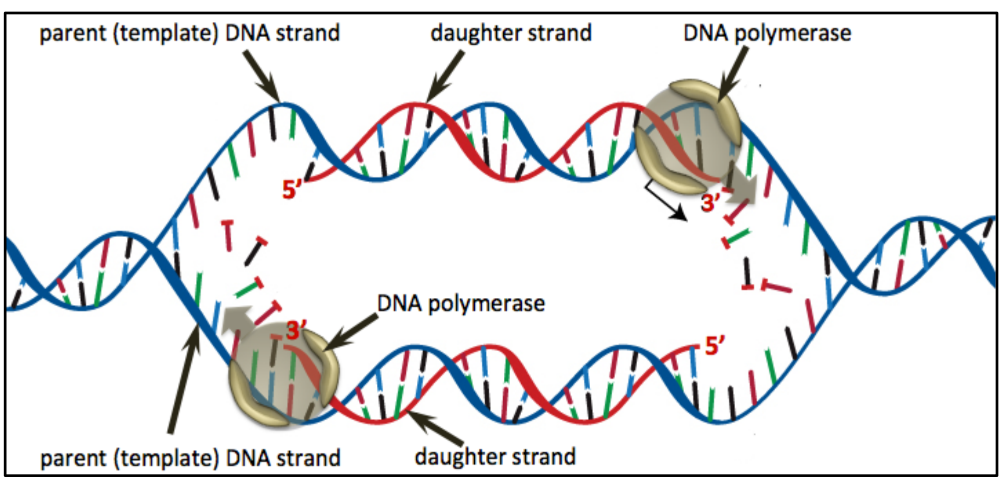
When a cell divides, the original cell, called the *parent cell*, forms two new cells that are called the *daughter cells*.

Before a somatic cell (non germ/sex cell) physically splits into two daughter cells, an exact ___________ of all of the parent cell's DNA must be made so that a copy can be placed in each daughter cell.

The process in which a duplicate copy of a DNA double helix is made is called _____ _______________.

In DNA replication, each of the two original DNA strands (called the ***parent strands***) is used as a "template" for the formation of a new "___________ ______________."

- In eukaryotic cells, proteins called "helicases" unwind and then open sections of the double helices by disrupting the hydrogen bonds between base pairs.



Once a double helix has opened, DNA polymerase enzymes catalyze the formation of complementary daughter strands *on* _________*of the parent strands*.

The daughter strands are formed by adding _________ deoxyribonucleotides, one-by-one to the growing daughter strands.

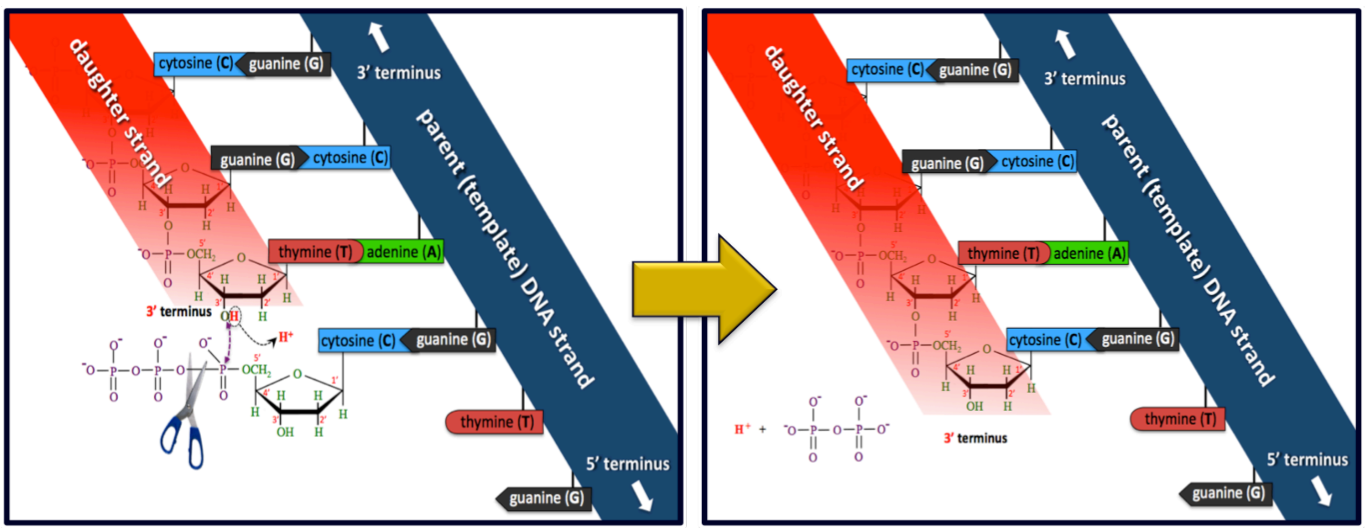
Because each daughter strand is assembled, nucleotide by nucleotide, with organic bases that are *complementary* to one of the parent strands, then both of the double-stranded DNA particles that result from the replication process are ***identical to the original DNA particle***.

DNA polymerase enzymes catalyze the addition of complimentary nucleotides to the _____ terminus of the growing ***daughter strand***.

In order to do this, DNA polymerase moves along the ***parent (template) strand*** in the direction from **3'** terminus to **5'** terminus of the ***parent*** strand.

- In plants and animals, this process occurs simultaneously *at multiple places along each parent strand*. Although ***not shown*** *in the image on the previous page*, these multiple sections of a daughter strand are linked/bonded together with the help of *DNA ligase enzymes*.

Chemical Bonding in DNA Replication



DNA polymerase enzymes catalyze the formation of ____________________ ***bonds*** between free nucleotide *tri*phosphates and the **3'** terminus of a growing daughter strand.

The daughter strands that are produced are ____________________ to their template parent strands because ***DNA polymerase directs the addition of nucleotide triphosphates that are complementary to the adjacent parent strand bases***.

- In the image shown above, a triphosphate with the *cytosine* (C) base was added because of the adjacent complementary *guanine* (G) base in the template parent strand.

Proofreading and Repair

When adding new deoxyribonucleotides to a growing DNA strand, DNA polymerase inserts the wrong residue slightly less than once every 10,000 times.

The enzyme proofreads its work to see if the correct deoxyribonucleotide residue has been added; if a mistake has been made, DNA polymerase clips the residue and tries again.

If proofreading does not catch the error, then other DNA repair enzymes are likely to catch the problem. After proofreading and repair, the error rate during replication falls to *less than one in 1 billion bases*.

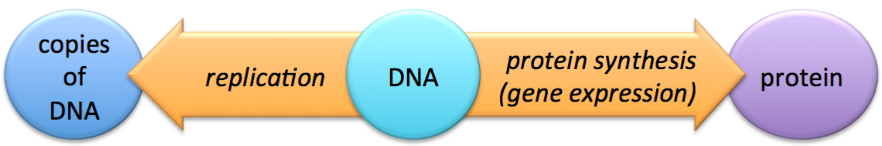
Understanding Check: Using the illustration on shown in the viewgraph, name the next *nucleotide triphosphate* that would be added to the growing daughter strand.

Understanding Check: Label each of the statements below as either **TRUE** or **FALSE**.

a. A double helix can be converted to two identical double helices.
b. If the sequence within a parent template strand is **5'**-CGCGTTA-**3'**, then the sequence of its daughter strand would be **5'**-GCGCAAT-**3'**.
c. If the sequence within a parent template strand is **5'**-CGCGTTA-**3'**, then the sequence of its daughter strand would be **3'**-GCGCAAT-**5'**.
d. One of the double helices produced in DNA replication contains the two parent strands, and the other double helix contains the two daughter fragments.
e. The two parent strands are complementary to each other and therefore, the two daughter strands are complementary to each other.
f. DNA polymerase moves in the direction from **5'** terminus of the parent (template) DNA strands toward the **3'** terminus of the parent strands.
g. DNA polymerase catalyzes the addition of nucleotides to the **3'** terminus of the growing daughter strand.

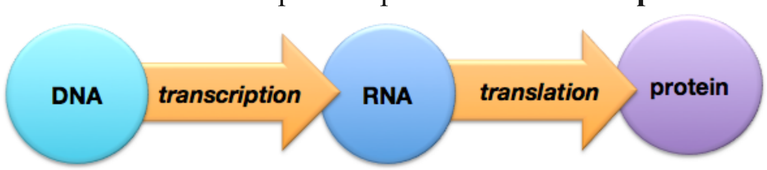
Protein Synthesis

The information contained in the sequence of DNA nucleotides is used to generate proteins. This process is central to the existence of all known life forms because proteins are critical in directing and controlling cell growth and function, and in regulating an organism's metabolism.



The process of making proteins from the information in DNA is called ____________ _____________ or ____________ _________________.

Protein synthesis can be divided into two sequential processes: **transcription** and **translation**.



Transcription: DNA to RNA

A ___________ is generally defined as *a region of DNA that carries the information needed to produce a protein*.

- Human DNA contains about ***25,000*** genes.

The first step in *protein synthesis* is called _____________________.

In the transcription process, the information (sequence of nucleotides) in a gene is used to create a specific sequence of *ribo*nucleotides in a single-stranded _______________ _______ (**mRNA**) particle.

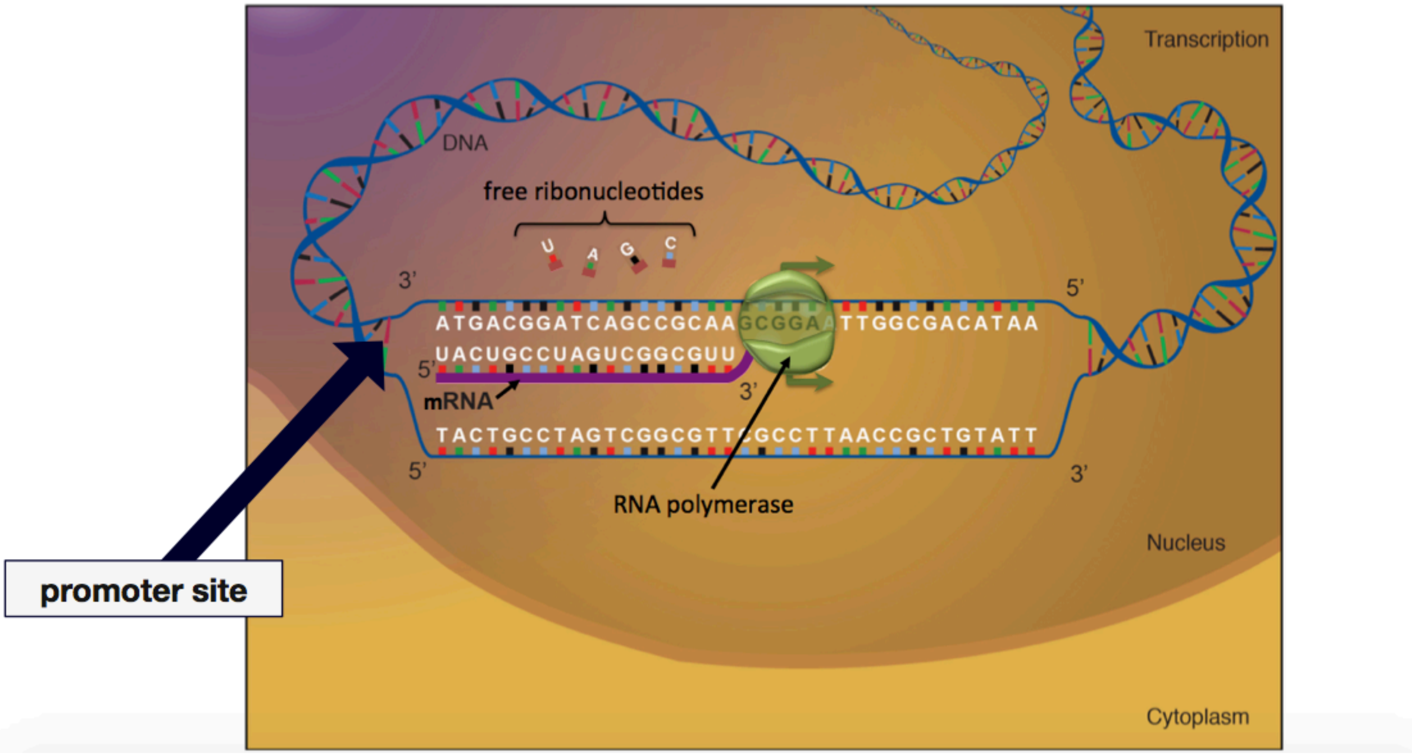
The *transcription process* is similar to DNA replication, with three main differences:

1) RNA is produced in transcription, whereas DNA is produced in DNA replication.
2) In transcription, only ***one*** of the DNA double helix strands is used as a template to produce ***one mRNA strand***, whereas in DNA replication, ***each*** of the two DNA strands in a double helix produce a daughter strand.
3) Different enzymes are used.

Transcription begins when an ______ ________________ **enzyme** binds to a "**promoter site**" on the gene.

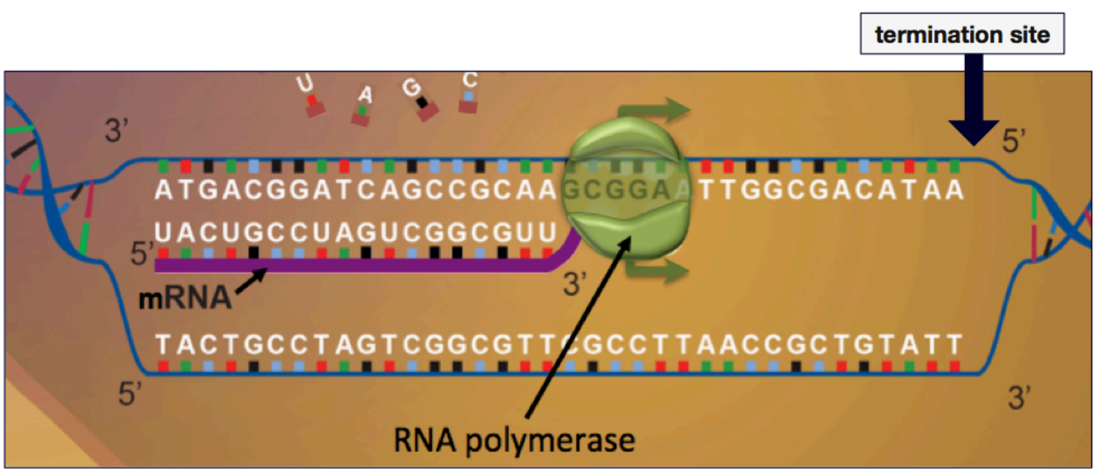
- ***Promoter sites*** are specific sequences of nucleotides at the beginning of a gene. When RNA polymerase binds to the promoter site, the hydrogen bonding between base pairs within the DNA double helix is disrupted. This unwinds and opens a section of the gene's double helix.

Next, *RNA polymerase* moves along one strand of the gene while catalyzing the addition of *free ribonucleotides,* that are complementary to the DNA template, to the ___ end of a growing **mRNA** strand.



The DNA-to-RNA base pairing is a bit different from the DNA-to-DNA base pairing because RNA contains *uracil* (U) and does not contain *thymine* (T).

- *Adenine* (A) DNA organic bases base pair with *uracil* (U) RNA bases.
- DNA organic bases C,G, and T, base pair with RNA bases G, C, and A, respectively.



When the RNA polymerase reaches a nucleotide sequence in the gene, called the **termination site**, the mRNA strand is ________________.

Transcription occurs in the ________ of eukaryotic cells and in the _______________ of prokaryotic cells.

Although not shown in the illustration on the previous page, in eukaryotic cells, there are regions of RNA that undergo additions and deletions of nucleotides in a process called **co-transcriptional modifications**. A discussion of the details of co-transcriptional modifications is beyond the scope of this course.

Messenger RNA contains the ______________ (message) from a gene that is required to produce a protein.

In eukaryotic cells, mRNA exits the nucleus through pores in the nuclear membrane and enters the cytoplasm where the *protein synthesis* process will be completed.

About 1.5% of human DNA is in the form of a gene; the remaining 98.5% is called **noncoding DNA**. Until recently, the noncoding DNA was called "junk DNA." It was thought that all noncoding DNA is extraneous DNA that has been passed down, through evolution, from ancestral species. More recently, research has shown that not all noncoding DNA is "junk." For example, noncoding DNA has been found to play a role in the regulation of transcription.

Translation: mRNA to Protein

The next step in *protein synthesis* is called __________________.

- In this process, the information that was in a gene and is now contained in mRNA is used to construct the polypeptide chains that make up proteins.

In *translation*, the sequence of nucleotides in mRNA, is converted (translated) to a sequence of

__________ ____________residues in a *polypeptide*.

This is done using the ____________ ____________.

The *genetic code* is based on three-nucleotide sequences, called _____________.

A *codon* directs the addition of a ____________ amino acid residue to a polypeptide that is being formed.

The *genetic code* was completely known by the late 1950's and is shown in the table below.

- Example: the mRNA three-nucleotide sequence (codon) **GAU** codes for the addition of an *aspartic acid* (Asp) amino acid residue to a growing polypeptide chain.

Genetic Code Table: The mRNA Codons and the Amino Acids for which they Code

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

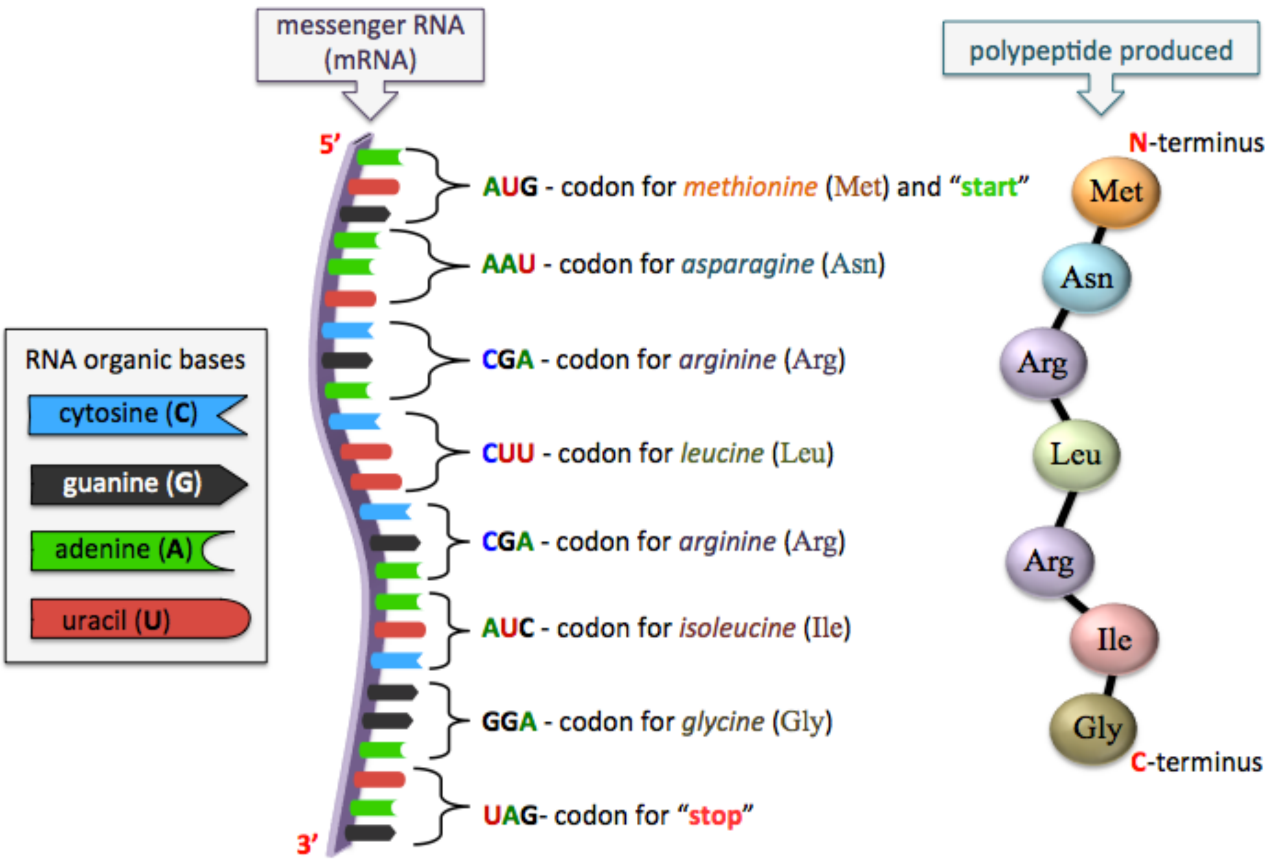
The mRNA nucleotide sequences in the codons are listed in the direction of the mRNA **5'** terminus toward the **3'** terminus.

Since there are ***twenty-one*** common amino acids, and only ***four*** types of RNA nucleotides (**C**, **G**, **A**, **U**), it was necessary for nature to use ***three nucleotides per codon***. ***If*** there were only *one* nucleotide per codon, then only *four* amino acids could have been specified. ***If*** there were *only two nucleotides per codon*, then only *sixteen amino acids* could have been specified.

In nature, mRNA sequences contain hundreds to many thousands of nucleotide residues, however, for simplicity, I will use a very short mRNA sequence in the following example.

The "start" codon (**AUG**) determines the first amino acid residue to be used in the polypeptide.

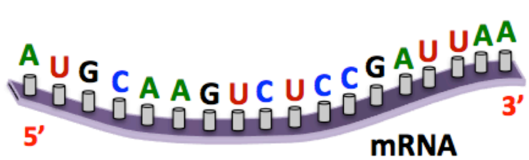
Peptides are synthesized in the direction from their **N**-terminus to their **C**-terminus.



Nucleotides are added to the growing peptide chain until the stop codon is reached.

Understanding Check: Write the three-letter abbreviations for the amino acid residues, in order from **N**-terminus to **C**-terminus, of the polypeptide that would be produced in the translation of the mRNA shown below.

- Recall that in translation, polypeptides are formed in the direction from their **N**-terminus to the **C**-terminus.

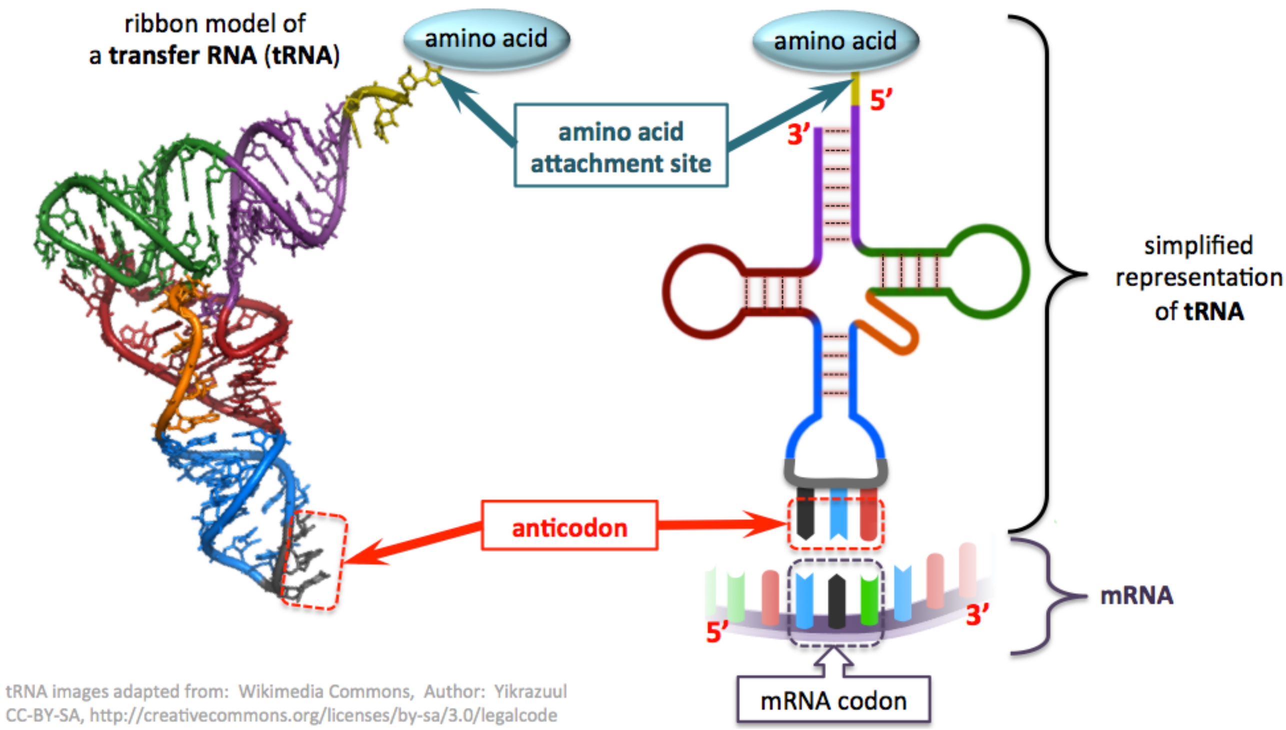


The Role of Transfer RNA (tRNA) in Translation

In order for you to understand how the codon information is used to direct the addition of specific amino acids to a growing peptide chain, we will consider another type of RNA, called **transfer RNA (tRNA)**.

The function of a tRNA particle is to _____________ and then ______________ the amino acid that is specified by an mRNA codon to a growing peptide chain.

A ribbon model and a simplified representation of a tRNA particle are shown below.



tRNA particles are single-stranded polynucleotides that contains 73 to 93 ribonucleotides. Base pairing causes the single tRNA strand to *fold back on itself*. Hydrogen bonding between tRNA base pairs is represented by the dotted lines that are highlighted in the simplified tRNA representation.

One region of tRNA has a site that attaches to an *amino acid residue,* and another region contains a three-nucleotide sequence called an ________________.

The particular amino acid that a tRNA carries depends on the three-nucleotide sequence in its *anticodon*.

Anticodons are ________________ to mRNA codons (as illustrated in the bottom-right of the image above).

In the illustration above, the mRNA codon (in the **5'** to **3'** direction) is **CGA**, therefore the tRNA anticodon (in the **3'** to **5'** direction) is **GCU**. Using the mRNA codon in this example (**CGA**), and the *genetic code table*, we find that this tRNA would carry an *arginine* (Arg) amino acid residue.

The Translation Mechanism

Translation occurs in ***three*** steps:

1) ____________________ 2) ____________________ 3) ____________________

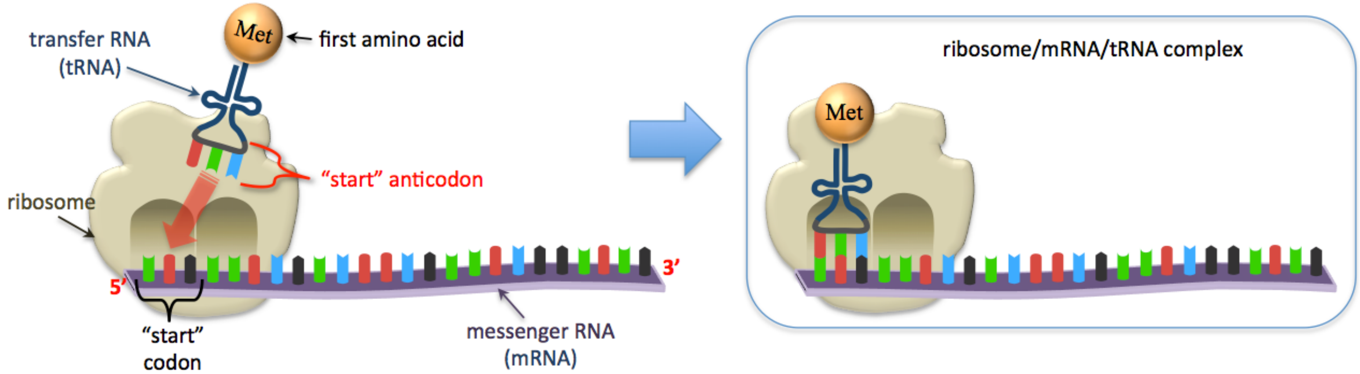
1) Initiation

Once the *messenger RNA* leaves the nucleus and enters the cytoplasm, a ____________ attaches to its "start" codon (**AUG**).

Ribosomes are relatively large particles that contain protein *and* very long RNA strands called **ribosomal RNA** (________).

- The function of *ribosomes* is to provide a structure upon which polypeptides can be produced.

Next, a *transfer* RNA with a "_________" *anticodon* (**UAC**) binds to the mRNA start codon to form a ribosome/mRNA/tRNA complex. The formation of this complex, as illustrated below, is called the **initiation step**.

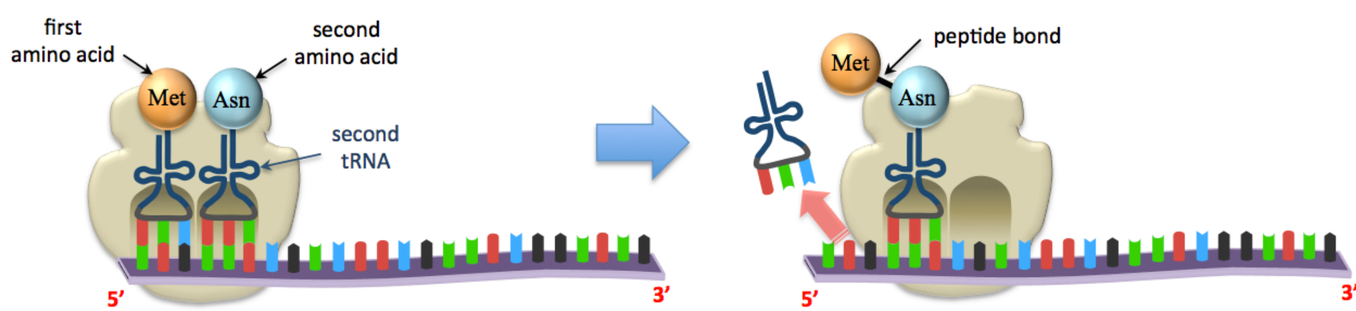


2) Elongation

Elongation begins when a second tRNA binds to the ribosome/mRNA/tRNA complex (as illustrated below).

- The second tRNA delivers the amino acid that corresponds to the codon that follows the start codon. In this example, the second tRNA delivers an asparagine (Asn) amino acid.

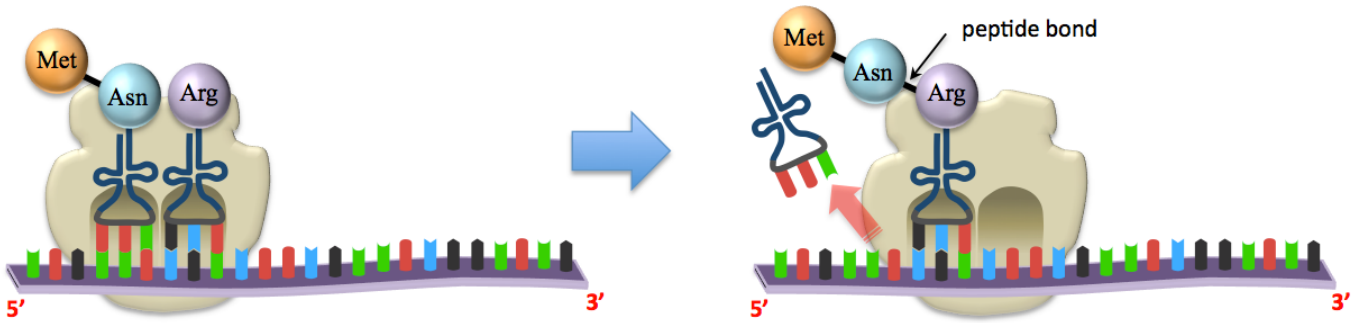
In elongation, a polypeptide chain is made by the formation of __________ *bonds* between the amino acid residues that are specified by the genetic code in mRNA.



A part of the ribosome, called the *transferase center*, catalyzes the formation of *peptide bonds* between the amino acid residues.

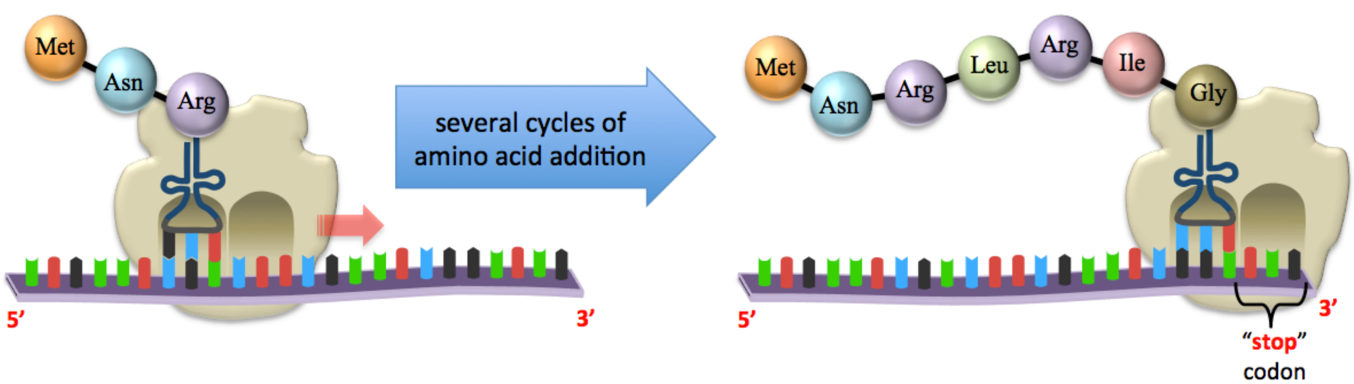
Next, the ribosome moves a distance of three bases (a codon) along the mRNA strand in the **5'** to **3'** direction. The tRNA that was attached to the first amino acid detaches from the mRNA and diffuses throughout the cytoplasm where it will encounter an enzyme that catalyzes its reattachment *to an amino acid that matches its anticodon*.

This entire process is repeated when a third amino acid binds to the next mRNA codon, as illustrated below.



Peptides are synthesized in the direction from the **N**-terminus to the **C**-terminus because the new amino acid residues form peptide bonds with the **C**-terminus of a growing peptide chain.

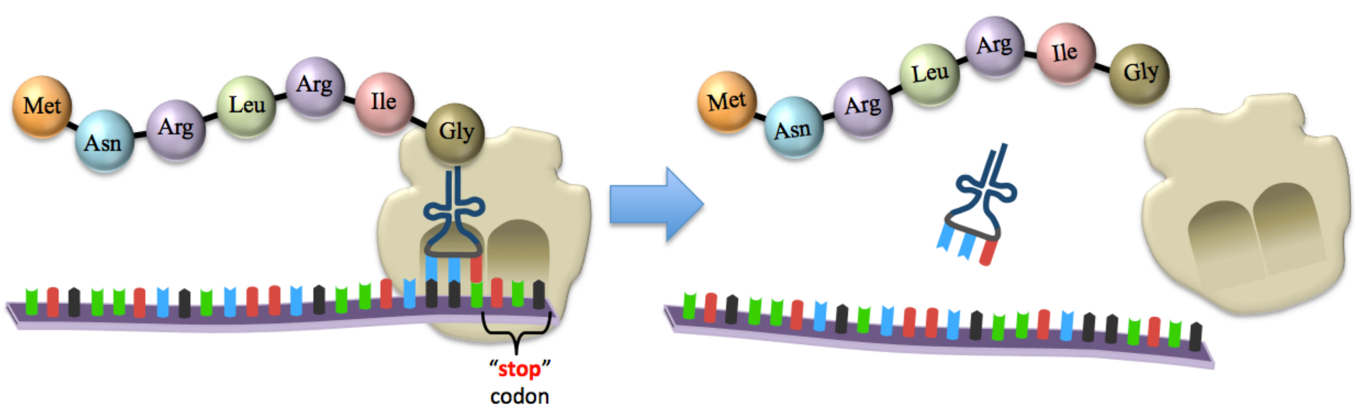
After many cycles of amino acid addition, the ribosome will reach a "***stop***" codon as illustrated below.



Termination

The ***termination*** step occurs when the ribosome reaches a __________ codon (**UAA**, **UGA**, or **UAG**).

When this happens, the polypeptide, mRNA, tRNA, and ribosome become separated, as illustrated below.



In some cases, the polypeptide that is produced in translation is a fully-functional protein. In other cases, the polypeptides undergo further folding to acquire the secondary and tertiary structures that they need in order to function. In many cases, multiple polypeptide subunits must assemble into a quaternary structure in order to form a fully-functional protein.

Some proteins require covalent bonding modifications, called **post-translational modifications**, after they are translated. An example of a *post-translational modification* is the addition of *heme groups* to heme-containing proteins, such as hemoglobin and myoglobin.

Although not discussed in this section, there are some differences in how prokaryotic and eukaryotic cells synthesize proteins. Many antibiotics take advantage of these differences in order to inhibit protein synthesis in prokaryotic pathogens while not harming the human (or animal) hosts. The table below lists some of these antibiotics and their effects on prokaryotic protein synthesis.

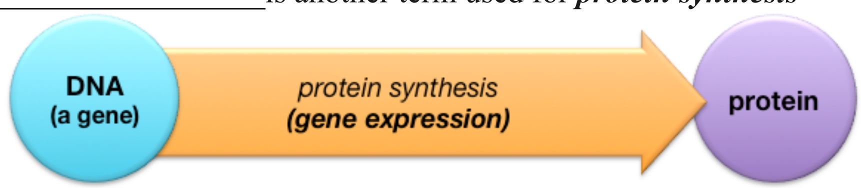
Antibiotic	Mode of Action
chloroamphenicol	inhibits peptide bond formation
erythromycin	prevents ribosome from moving along mRNA
streptomycin	inhibits initiation and cause misreading of mRNA
tetracycline	inhibits the binding of tRNA to mRNA/ribosome

Understanding Check: Write the three-letter abbreviations for the amino acid residues, in order from **N-**terminus to **C-**terminus, of the polypeptide that would be produced by the **transcription** *and* **translation** of **DNA** with a sequence of **3'**-TACGGGGTACACACT-**5'**.

- **CAUTION:** In this question, you were given the sequence of DNA, *not* mRNA.

Control of Gene Expression

_________ ______________________is another term used for ***protein synthesis***



The DNA of each living thing contains *thousands* of ***genes***.

These genes are not continually ________________________ (*read to make proteins*), because the production of unneeded proteins would be an inefficient use of resources.

In chapter 13, you learned how organisms often use several reactions in series, called **metabolic pathways**, in order to carry out the chemical changes they require to meet their physiological needs. You also learned that many of the reactions in metabolic pathways require enzymes and that organisms can regulate (slow down or speed up) a metabolic process, according to their needs, by *inhibiting* or *activating* one (or more) of the enzymes involved in the metabolic pathway.

Another way to control metabolic pathways, or other conditions that involve proteins, is to **regulate the** ______ ____ ________________ **of a protein/enzyme**.

This is called ______________________ ***gene expression***.

Increasing the rate of gene expression is called ____________________of the gene; decreasing the rate of gene expression is called ______________________of the gene.

Example: Control of the *lac* operon.

Recall that *transcription*, the creation of mRNA from the information in a DNA, involves the *RNA polymerase enzyme*. In order for transcription to begin, *RNA polymerase* must bind to a "**promoter site**" on the DNA to be transcribed.

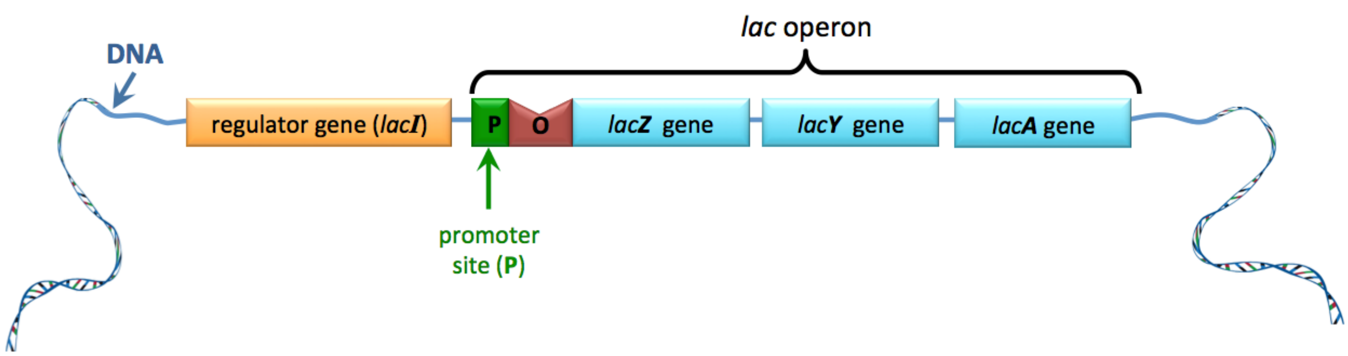
Sometimes, ***several genes*** are transcribed from the **same** *promoter site*.

A section of DNA made up from genes that are *transcribed from the same promoter site* is called an _________________.

Escherichia coli (*E. coli*) contain an operon comprised of three genes that code for proteins involved in the metabolism or transport of lactose.

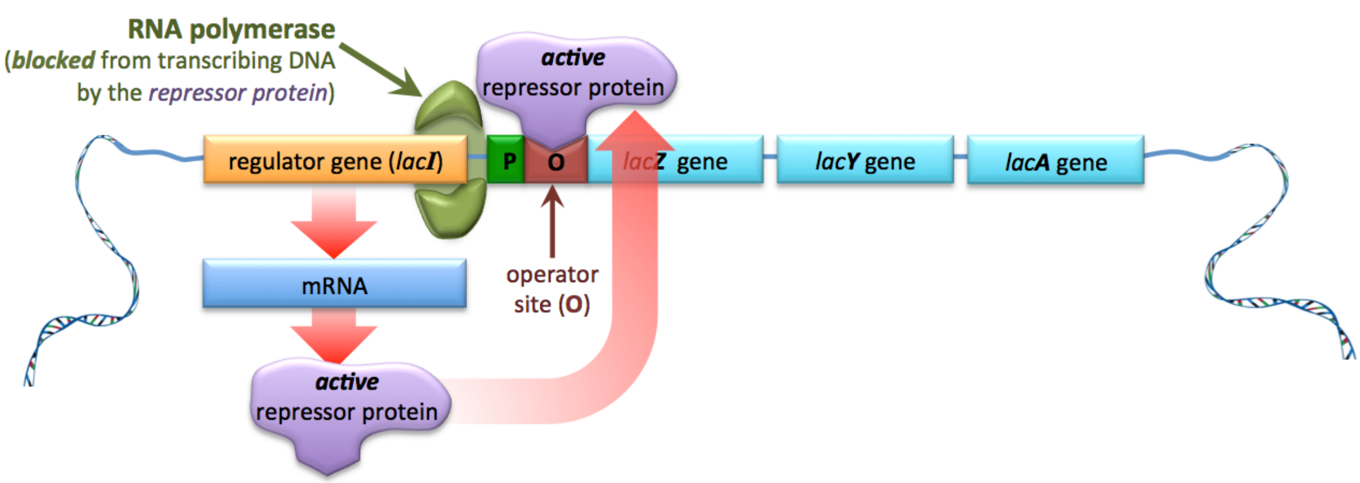
This *operon* is called the ***lac* operon**.

The three genes of the *lac* operon are called *lac*Z, *lac*Y, and *lac*A. The *lac* operon is preceded by a *lac regulator* (*lacI*) *gene*, as illustrated below.



In the presence of glucose, lactose is not needed as an energy source and almost zero lactose is allowed into the bacterial cell, therefore it is not necessary for the genes of the *lac* operon to be expressed. It is for this reason that nature has provided a regulator gene (*lacI* gene).

The *lacI* gene produces an _______________ ***repressor protein***, which binds to a segment of DNA called an **operator site**, as illustrated below.



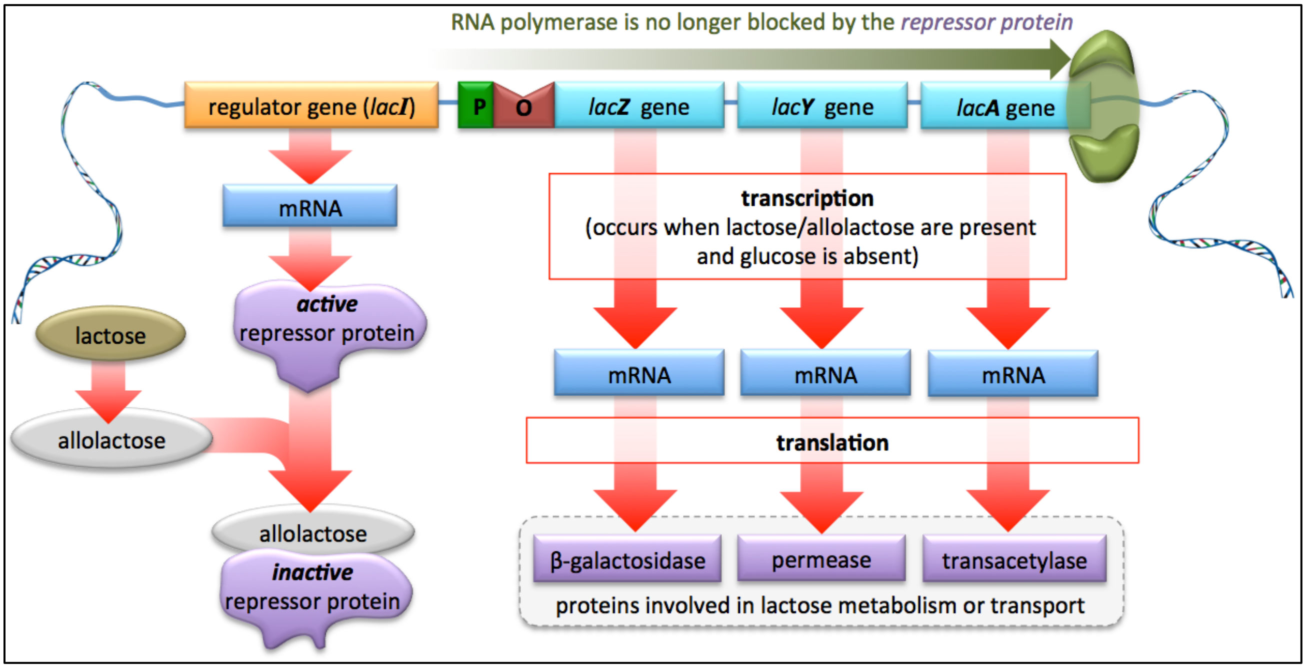
When the active repressor protein is bound to the operator site, ***RNA polymerase is*** ___________ ***from moving along the lac operon* to transcribe the *lacZ*, *lacY*, and *lacA* genes**.

- Under normal conditions (glucose present), this is ideal for the *E. coli* because there is plenty of glucose present, and lactose is *not* being taken in, therefore the bacteria do not need the lactose metabolism or transport proteins that the genes of the *lac* operon express.

Let's consider the scenario in which *lactose* is present ***in the absence of glucose***. In this case, lactose is allowed to enter the *E. coli*. Whenever lactose is present in the bacteria, some of it is converted to a compound called *allolactose*.

Allolactose will bind to, and thereby ______________ *the repressor protein*.

The *inactive* repressor protein cannot ___________ to the operator site, therefore RNA polymerase is no longer blocked and will move along the *lac* operon to transcribe the *lac* operon genes, as illustrated below.



The mRNAs from the *lac* operon genes are *translated* to the three proteins: *β-galactosidase*, *permease*, and *transacetylase.*

- *β-galactosidase* is an enzyme involved in lactose metabolism.
- *Permease* is a transmembrane protein that transports lactose into cells.
- *Transacetylase* is an enzyme thought to be involved in the breakdown of some non lactose species that are transported into cells by *permease*.

The presence of lactose within *E. coli*, and the subsequent presence of allolactose, **upregulate** the genes of the *lac* operon. Allolactose and lactose are continuously being used up by the bacteria. Whenever lactose is no longer obtained from the surroundings, the concentration of allolactose within the cells will decrease, and then *active regulator proteins* become available again to **downregulate** the *lac* operon.

Understanding Check: Determine whether each of the following conditions would result in the genes of the *lac* operon being ***upregulated*** or ***downregulated***.

a. RNA polymerase is blocked from moving along the *lac* operon
b. Allolactose binds to the repressor protein
c. Lactose concentration within a cell is decreased
d. Allolactose concentration within a cell is increased
e. Active repressor proteins are present in the absence of allolactose
f. Glucose is available.

Viruses

Viruses are small particles that are not able to ________________ on their own.

- A typical virus is about 1/100 the size of a bacteria.

In order for a virus to reproduce, it must invade a cell of *another* organism.

A cell that is invaded by a virus is called a ____________ **cell**.

Viruses have much less-complicated structures and contents than do cells; they do not have the components, such as nucleotides and certain enzymes, that are needed in order to self-replicate. Viruses come in a variety of shapes.

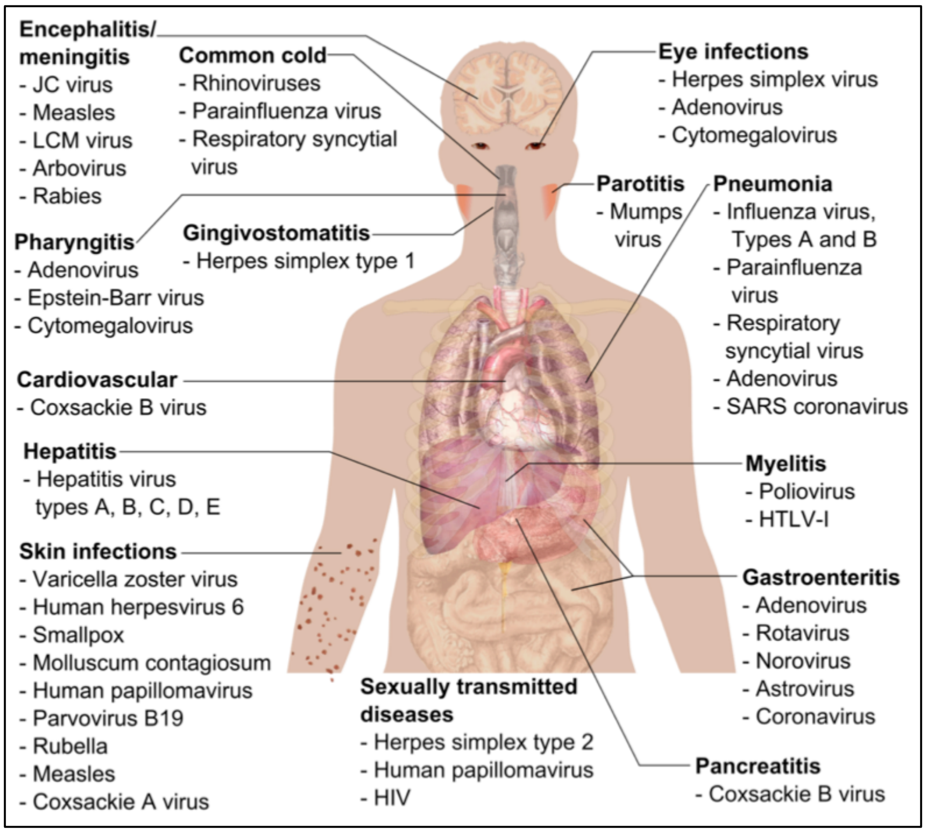
All viruses have a protein shell that encapsulates *either* _________ *or* __________, and some viruses contain a small set of enzymes.

- A virus that contains DNA is called a **DNA virus**. A virus that contains RNA is called an **RNA virus**. Some viruses have a lipid coating that surrounds their protein shell.

All forms of life can have their cells invaded by viruses.

When viruses take over plant or animal cells, we call this a __________ **infection**.

- Several human *infectious diseases*, some of which are *chronic diseases*, are caused by viral infection. Some of these diseases, the location of the infections, and the most common viral species involved are shown on the right.

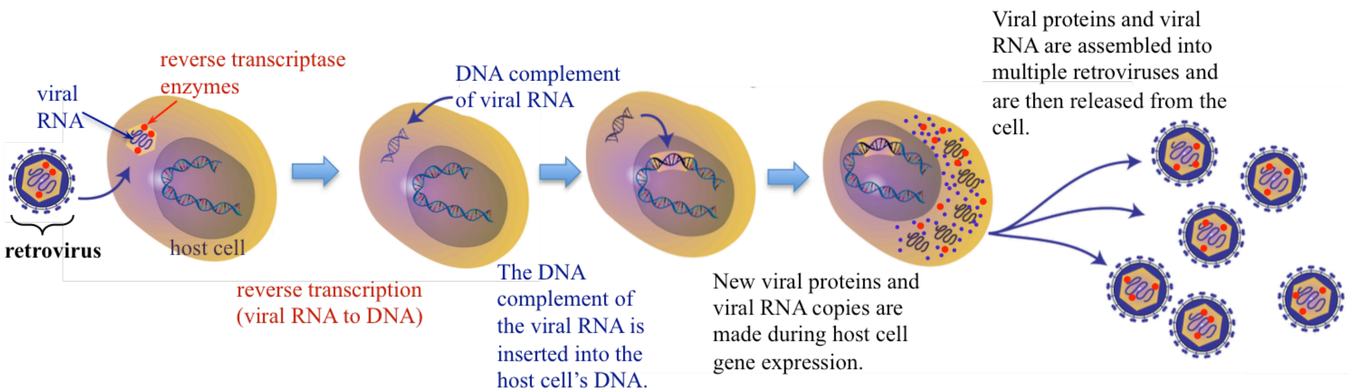


Viruses infect cells by introducing their DNA or RNA and, in some cases, a few types of enzymes, into the host cell. This is done in various ways depending on the particular virus. Once the viral DNA or RNA is introduced to the host cell, enzymes, nucleic acids, and amino acids ***from the host cell*** are employed to make more viral DNA or viral RNA, and viral proteins. The viral DNA or RNA and viral proteins are ***reassembled into multiple, new virus particles*** within the host cell. These new viral particles cause the host cell to burst or release the viral particles using the exocytosis process. The infection spreads as the new viruses infect other cells.

Viral DNA or RNA self-replicates and expresses viral proteins in several different ways, depending of the type of virus. A detailed description of each of these processes is beyond the scope of this course. Instead, I will give one example by describing how the *human immunodeficiency virus* (HIV) infects cells.

Viral infection example: the *human immunodeficiency virus* (HIV)

HIV is in a category of viruses called **retroviruses**. When taken into the host cell, the information in the *retroviral* RNA is converted to complementary DNA. This is called **reverse transcription** and is catalyzed by a viral enzyme called ***reverse transcriptase***. The name "*reverse transcriptase*" is applied because the chemical process is the *reverse* of transcription. The newly formed DNA is inserted into the host cell's DNA, and then *transcribed* into new *retroviral* RNA and viral proteins during the host cell's normal gene expression. The *retroviral* RNA and viral proteins are ***assembled into multiple retroviruses*** and then released from the cell. This process is illustrated below.



Animals are capable of producing immune responses that eliminate most viral infections. Vaccinations are effective as anti-viral agents because they allow vaccinated individuals to produce artificially acquired immune responses to many viral infections. Some viruses, such as HIV, herpes simplex virus (HSV), and hepatitis C virus (HCV), are capable of evading immune responses and therefore result in chronic (persistent or long-lasting) disease.

Understanding Check: Enzymes Involved in DNA and RNA Formation

Which enzyme (DNA polymerase, RNA polymerase, or reverse transcriptase) is involved in each of the processes below:

a. Transcription
b. Replication
c. Using a DNA template to make RNA
d. Using an RNA template to make DNA

Genetics

_______________ is the study of genes, variation in genes, and heredity.

The information contained in the DNA of an organism is called its ______________.

Individuals, other than identical twins, have different ____________ _________ because the DNA of each person is not exactly the same as that of another person.

- Even identical twins have about *one hundred* incidents of differences in nucleotide sequences in their genome, although that is not enough to make differences in their appearances. Small differences in the appearances of identical twins are attributed to environmental, not genetic, differences.
- Examples of inherited traits in humans include eye, skin, and hair color.
- Sometimes, _____________ result from an abnormal DNA sequence in a parent being passed to the next generation.

Mutations

Any permanent change in the primary structure of (sequence of nucleotide residues in) DNA is called a ______________________________.

- Mutations might involve the switching of one base pair for another or the addition or deletion of base pairs.
- Changes in the nucleic acid sequence of a gene (a mutation) can result in changes in the __________ ___________ sequence of the protein that is expressed, and can thereby give rise to partially or completely *nonfunctional proteins*.
- A chemical or physical agent that induces mutations is called a _____________.
- Errors in replication and exposure to mutagens (including x-rays, UV radiation, nuclear radiation, and chemicals) are the common causes of mutations.
- If the mutated gene results in a partially or completely nonfunctional protein that is important for an individual to function, the health of the individual may be diminished. *Some mutations result in premature death.*

Types of Mutations

Mutations that occur in multicellular organisms can be categorized as either ______________ **mutations** or _________ _______________ **mutations**.

Somatic mutations, also called *acquired mutations*, are mutations that occur within cell types that *are not involved in reproduction* (**somatic cells**).

- When a mutation arises in a *somatic cell*, then the mutation is limited to that cell or any of its cell-division descendants ***within the organism***. Cancer is a disease that results from *somatic mutations* in genes that are responsible for cell growth or cell differentiation. Such mutations cause the formation of a mass of mutated cells (cancer cells) called a **tumor**. For example, a skin cell can acquire a mutation, very often from exposure to UV radiation, which causes it to begin to rapidly divide/replicate and form a tumor.

Germ line mutations are mutations that occur within germ line cells.

- Germ line cells are the cells that are involved in reproduction (e.g. sperm and ovarian cells).
- If a mutated germ line cell is passed to an offspring, then it results in a _______________________ **mutation**, which is a mutation occurring in the nuclei of ***every cell of the offspring***.
- When a *constitutional mutation* causes a negative health condition, it is referred to as a _____________ _________________.
    - Note that mutations in germ line cells can be transmitted to offspring, whereas somatic cell mutations cannot.

Small-Scale Mutations vs. **Large-Scale Mutations**

Mutations can be classified as **small-scale** or **large-scale**, depending on the ____________ *of nucleotides involved in the mutation*.

- *Small-scale mutations* involve a change in a small number of nucleotides (usually one to three nucleotides) within a single gene.
- *Large-scale mutations*, sometimes called ***chromosomal mutations***, involve changes in large sections of chromosomes.

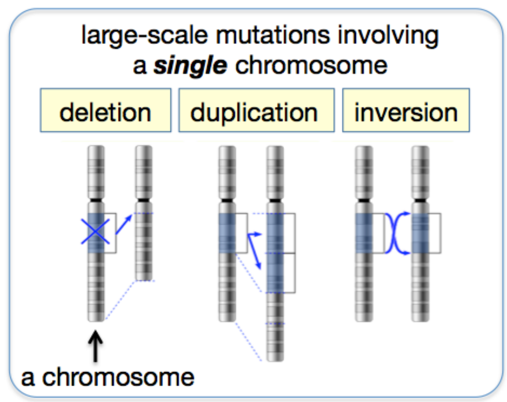
Small-Scale Mutations:

1) **Point:** A single, incorrect nucleotide takes the place of the original nucleotide. This often happens during DNA replication.
2) **Insertions**: One or more extra nucleotides are inserted between two of the original nucleotides in a gene.
3) **Deletions**: A single nucleotide or a short sequence (usually 2 or 3 nucleotides) are removed from a gene.

Large-Scale Mutations:

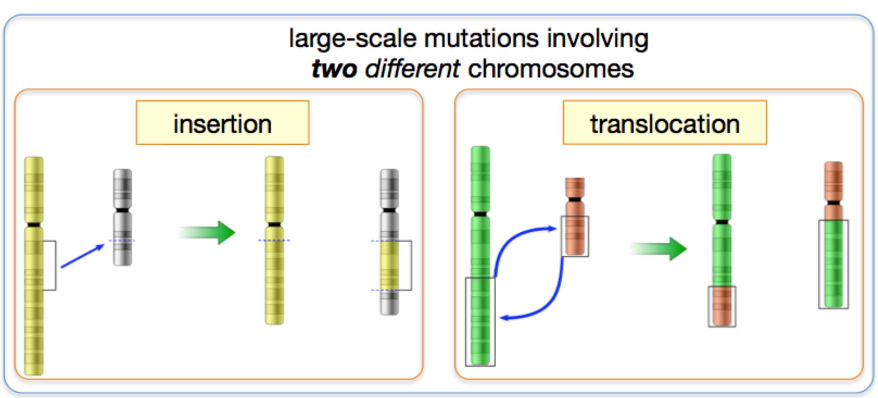
Large-scale mutations involving a **single chromosome** (illustrated on the right):

1) **Deletions:** A large section of DNA within a single chromosome is removed/deleted.
2) **Duplication:** A large section of DNA within a single chromosome is duplicated and reinserted.
3) **Inversion:** A large section of DNA within a single chromosome is inverted (3' and 5' directions reversed).



Large-scale mutations involving **two different chromosomes**:

1) **Insertion:** A large section of DNA within one chromosome is inserted into another chromosome (illustrated below, left).
2) **Translocation:** Large sections of DNA in two different chromosomes are exchanged (illustrated below, right).



Human Genetic Diseases

The inheritance of one or more mutated genes can result in a **genetic disease**.

- Some genetic diseases involve a particular gene, others involve multiple genes.

Normally, individuals have **two copies** of each gene in the nucleus of every somatic cell in their bodies. One copy of the gene was inherited from mother, and one copy from father.

A ___________________ **disease** occurs when *one or both* copies ***of a particular gene*** contains a harmful mutation.

- "**Dominant diseases**" are monogenic diseases that occur when _________ ***copies*** of a gene are mutated.
- "**Recessive diseases**" are monogenic diseases that occur when ________ copy of the gene *is not* mutated *and* __________ copy *is* mutated.
  - The exceptions to these dominant/recessive disease definitions are *some genes* on the human X chromosome.
  - Example of a monogenic recessive disease: cystic fibrosis

Approximate Prevalence of Some Monogenic Diseases

monogenic disease	prevalence
familial hypercholesterolemia	1 in 500
polycystic kidney disease	1 in 1,250
Huntington's disease	1 in 15,000
sickle cell disease	1 in 625
cystic fibrosis	1 in 2,000
Tay-Sachs disease	1 in 3,000
Phenylketonuria (PKU)	1 in 12,000
glycogen storage diseases	1 in 50,000
Duchenne muscular dystrophy	1 in 7,000
hemophilia	1 in 10,000

The prevalences shown here are approximate values and can vary significantly between populations with different ancestral ethnicities as well as gender.

More than 7,000 different *monogenic diseases* have been identified. Despite this large number of diseases, monogenic diseases are quite rare. The table on the right lists some monogenic diseases and their prevalences.

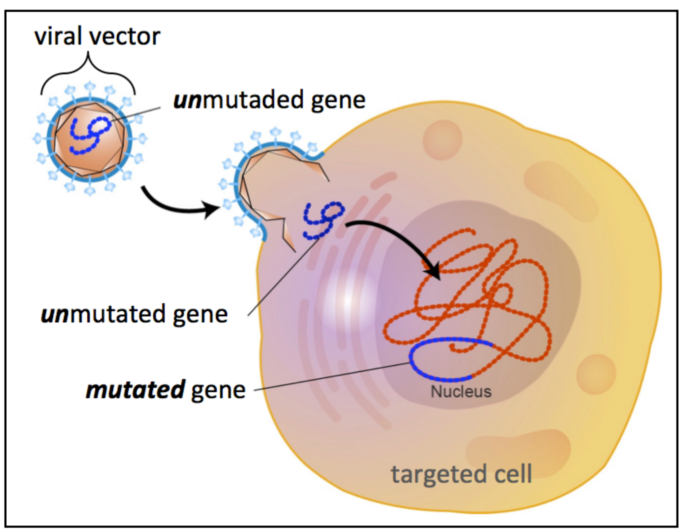
A ______________ **disease** occurs when *one or both* copies ***of multiple, different genes*** contain a harmful mutation.

- Examples: hypertension, coronary heart disease, and diabetes.

Treating Genetic Diseases: Gene Therapy

Some genetic diseases can be treated by ___________ _______________, which involves the delivery of functional (***un***mutated) genes to the cells of an individual who has a monogenic disease.

In some cases, the genes are introduced to the body by "infecting" the patient with inert DNA viruses that have had their genomes replaced with the gene that is to be delivered. A biological agent (a virus in this case) that is used to artificially deliver DNA to a cell is called a "**vector.**" An overview of this process is illustrated on the right.



Recombinant DNA: Genetically Modified Organisms

It is possible to extract copies of a gene from one organism and then insert that gene into another organism's genome.

- This process forms "_________________ **DNA**," which is ***DNA made from two or more sources***.
- An organism that contains recombinant DNA is called a **genetically modified organism (_____)** or a _____________.

GMO Examples:

1) Glyphosate Resistant Crops
2) Human Insulin from Bacteria
3) Measuring Gene Expression using Green Fluorescent Protein Mutants

Bioethics: Using What We Know

___________ is a term used to describe *the study of ethical issues that arise from biological technologies*.

Some of the bioethical concerns that involve human genetics include the following questions:

- What individuals or corporate entities can have access to personal genetic information?
- Is genetic information something that can be patented and owned?
- Who should be genetically tested and at what age should genetic testing be performed?
- Should we re-engineer the genes we pass on to our children?

These are complicated questions. The best solutions for complicated questions are usually not found by simple yes/no, or "one-size-fits-all" approaches. Furthermore, new research and emerging technologies make it necessary to frequently reconsider prior policies. Complicated questions often require complicated solutions. In order to reasonably address these bioethical questions, it is necessary to have an understanding of the fundaments of genetics and nucleic acid chemistry.   It was my intention to provide you with such an understanding in this chapter.