

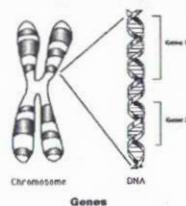
Protein Synthesis

What is it?

Making proteins using genetic code in DNA as instructions

In general:

- DNA codes for protein production
- Proteins are what do the work in a cell



What is a Gene?

Sequence of nucleotides on a chromosome that codes for proteins

The problem.....

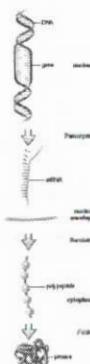
- Proteins must be synthesized in the cytoplasm by ribosomes
- The DNA molecule is too important to be allowed to exit the nucleus
 - It could be damaged or altered and this would render it useless (cell would die, organism may die)
- Often, proteins are needed in massive quantities in a cell → but there are only two copies of the DNA per cell

Cells have developed a way to avoid this problem...

The Central Dogma

DNA → mRNA → Protein

- A mechanism has evolved to pass the DNA's code into the cytoplasm
 - this occurs in a way such that the ribosome can read DNA in the cytoplasm without actually having direct access to DNA
- DNA is transcribed into an RNA message that can be passed to the cytoplasm to be used as a template by the ribosome (translation) to form proteins



RNA vs DNA

Table 1 Differences Between the Nucleic Acids

Deoxyribonucleic acid DNA	Ribonucleic acid RNA
• contains deoxyribose sugar	• contains ribose sugar
• double stranded	• single stranded
• adenine pairs with thymine guanine pairs with cytosine	• adenine pairs with uracil guanine pairs with cytosine
• resides in the nucleus	• resides both in the nucleus and in the cytoplasm

Types of RNA:

Table 2 Different Types of Ribonucleic Acid

Types of RNA	Characteristics and key functions
messenger RNA (mRNA)	<ul style="list-style-type: none"> varies in length, depending on the gene that has been copied acts as the intermediary between DNA and the ribosomes translated into protein by ribosomes RNA version of the gene encoded by DNA
transfer RNA (tRNA)	<ul style="list-style-type: none"> functions as the delivery system of amino acids to ribosomes as they synthesize proteins very short, only 70-90 base pairs long
ribosomal RNA (rRNA)	<ul style="list-style-type: none"> binds with proteins to form the ribosomes varies in length

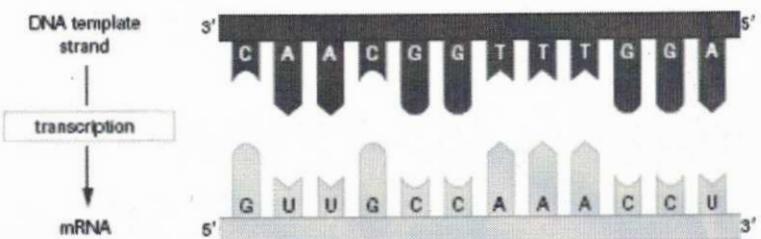
Transcription – an Overview

Goal: To make a mRNA copy from a DNA template (gene)

Divided into three main stages:

- Initiation: RNA polymerase binds to DNA at promoter (near beginning of gene)
- Elongation: RNA polymerase puts together complementary nucleotides
- Termination: recognizes stop signal at end of gene.
mRNA transcript is released & exits nucleus.

The mRNA Transcript that is formed is released and exits the nucleus into the cytoplasm

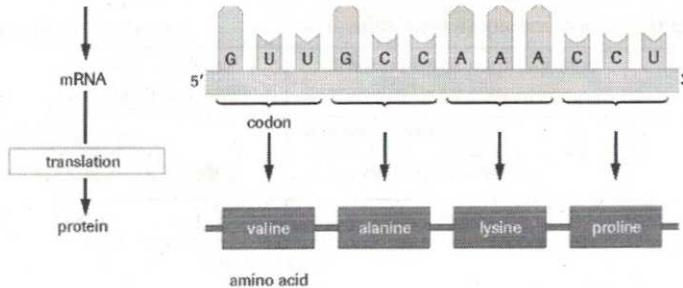


Translation – an Overview

Goal: to make polypeptide chain from mRNA made in transcription

Divided into three main stages:

1. Initiation: occurs when a ribosome recognizes a sequence of mRNA & binds to it
2. Elongation: ribosomes moves along three nucleotides at a time reading codons. Each codon codes for a particular amino acid that is delivered via tRNA
3. Termination: 3 nucleotide sequence is reached that does not code for an amino acid called a stop codon.



The Genetic Code:

- 20 different amino acids but only 4 bases
- Each triplet of nucleotides is called a codon
 - o Using three nucleotides results in 64 different possibilities.
 - o More than one codon can code for a single amino acid indicating a redundancy in the code.
- Redundancy minimizes errors
- Start codon → specific codon (AUG) that signals to the ribosome that the translation commences at that point
- Stop codon → specific codons that signal the end of translation to a ribosome

Second base of codon				
First base of codon				
U	C	A		
U	Phenylalanine Phe UUU UUC UUA UUG	UCU Serine UCC ser UCA Leucine UCC Ieu	UAU Tyrosine UAC Tyr UAA STOP codon UAG STOP codon UGC Cysteine UCC cysteine UGA STOP codon UGG Tyrosine UCC Tyr	U C A G
C	CYU Leucine CUU Ieu	CCC Proline CCA Pro CGC Glutamine CGA Gln CGG Gln	CAU Histidine CAC His CAA Glutamate CAG Glu	U C A G
A	AUU Isoleucine AUU Ile	ACU Threonine ACA Thr ACG Methionine Met (start)	AAU Asparagine AAC Asn AAA Lysine AAG Lys	U C A G
G	GUU Valine GUU Val	GCU Alanine GCC Ala GCA Glycine GCG Gly	GAU Aspartic acid GAC Asp GAA Glutamic acid GAG Glu	U C A G

© Clinical Tools, Inc.

Genetic Code

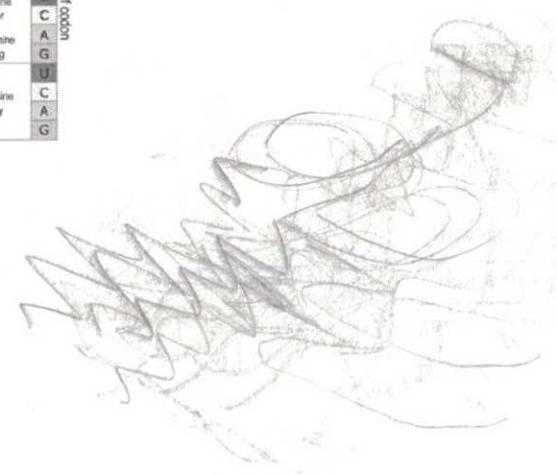
3' C G G, C A A T A 5'
codon

5' G C C G U U A U A 3'

↓
Protein amino acids

ala - val - ile

← use table/wheel



Protein Synthesis

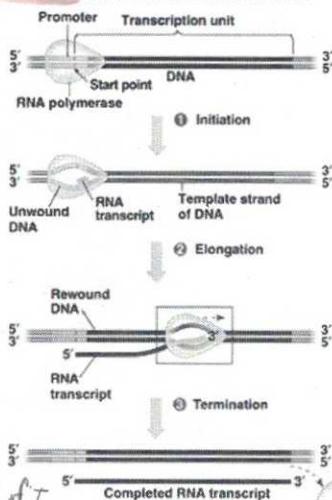
Transcription

Overview:

Where?	In the nucleus
Why?	To copy the information from DNA to mRNA To lead to the building of proteins (translation)
How?	Initiation, Elongation, Termination

Initiation: RNA Polymerase bonds to double helix at a promoter region at the 3' end of the strand to be copied (Template strand)

- RNA polymerase binds to DNA molecule on region 'upstream' of gene to be transcribed
- This region is called the Promoter region
 - RNA polymerase only recognizes the promoter region
 - Region high in Adenine & Thymine (called TATA box)
 - RNA polymerase uses less energy to open up the DNA molecule at this site because there are less hydrogen bonds between A and T (than C and G)
 - RNA polymerase unwinds the DNA as it moves along



Elongation:

- RNA polymerase moves along DNA, unwinding ~ 10 bases at a time
- RNA polymerase builds the new single-stranded mRNA in 5' to 3' direction
 - No primer required (as in DNA replication)
 - Elongation begins as soon as RNA polymerase binds to promoter
 - **Promoter region does not get transcribed**
- One strand from double-stranded DNA is used as template - called template strand
 - This would be the strand that runs 3'-5'
 - Other strand is called Coding strand
 - This would be the strand that runs 5'-3'
- RNA sequence is complementary to the template strand and identical to the coding strand **EXCEPT contains U instead of T**

Termination:

- A termination sequence is located at end of gene which RNA polymerase recognizes as 'stop' **NOT CODON** or stop is found on template strand
- Causes mRNA strand to disassociate from the DNA template strand

Post-Transcriptional Modifications

In eukaryotic cells, the mRNA produced from transcription needs to be 'modified' before it can leave the nucleus

- the mRNA produced in transcription is called the pre-RNA

Capping

- A 5' cap is added to the 5' end of the primary transcript
 - Protects the mRNA from digestion as it exits the nucleus and enters the cytoplasm
 - Will also play a role in translation
- ↳ Helps attach in translation

Tailing

- Approximately 200 adenine ribonucleotides are added to the 3' end
- it's called a poly-A tail
 - Added on by the enzyme poly-A polymerase
 - Protects the mRNA from digestion
 - Plays a role in translation termination

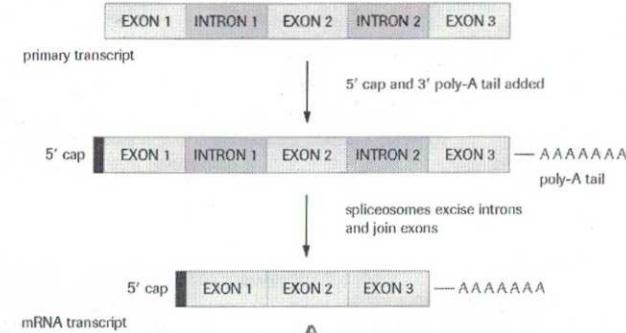
Eukaryotic genes are made up of coding and non-coding regions

- Exons are the coding regions
- Introns are the non-coding regions ↳ **Useless**
- The primary transcript contains both introns and exons
- Before mRNA leaves the nucleus, the Introns need to be removed
- If introns aren't removed, they will be translated and the protein will not fold properly (**USELESS!**)
- The introns are removed by spliceosomes
 - Made of RNA and proteins
 - Cut introns from primary transcript and joins together the remaining exons

Once the primary transcript has been capped and tailed and the introns removed, it is ready to be translated into a protein

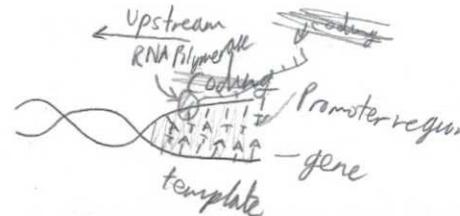
- This is now called the mRNA transcript

No proofreader needed following transcription (unlike DNA replication). Explain why this is the case (why is it not necessary):



Proteins can be replaced & code still exists in DNA

↑ can be switched around in different combinations (alternative splicing)



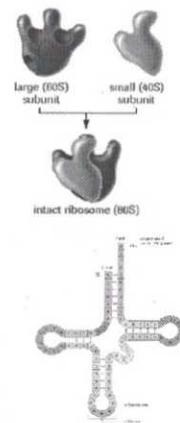
Translation

Overview:

Where?	In the cytoplasm
Why?	To synthesize proteins under the direction of mRNA
How?	Using mRNA, tRNA and ribosomes

Ribosomes:

- Two subunits: a large and small.
- These ribosomal subunits are a combination of protein and tRNA.
- The two subunits bind to mRNA so that the mRNA is clamped between them.
- Ribosome moves along the mRNA in the 5' to 3' direction.
- Ribosome must be positioned correctly to read mRNA correctly.

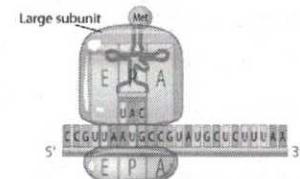


Transfer RNA

- Molecule that delivers amino acids.
- Small single stranded, resembles cloverleaf.
- Anticodon arm recognizes codon of mRNA.
 - Contains complementary RNA sequence
- Every tRNA carries a specific amino acid.

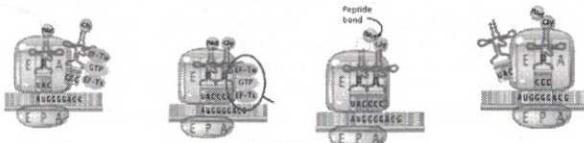
Initiation:

- Translation commences at the start codon (**AUG - methionine**) at the P site (peptide site).
- Appropriate tRNA delivers amino acid (methionine).



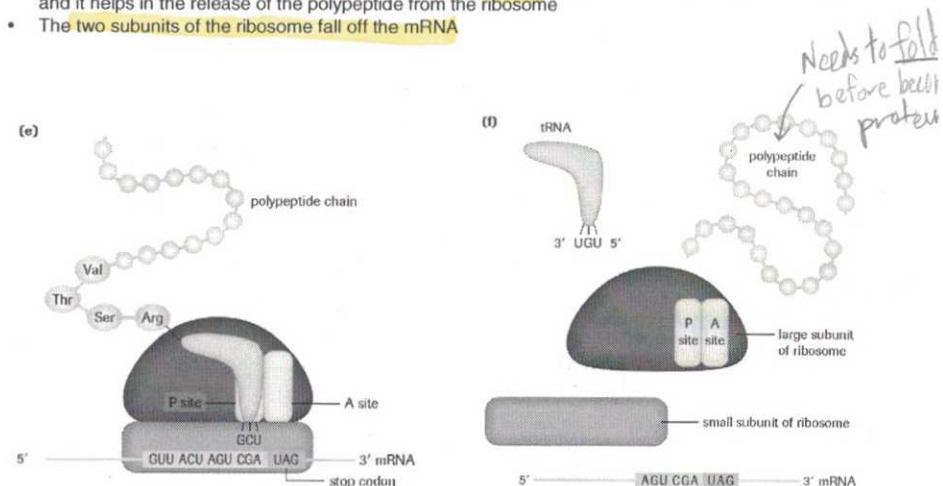
Elongation:

- Second tRNA enters at A site and brings in a second amino acid (Gly in this case)
- Peptide bonds are formed between amino acids attached to tRNAs in A and P site
- Ribosome shift over (5'-3')
 - tRNA that is now empty moves to 'E' site to be released back into the cytoplasm
 - aa that was in A site moves to P-site, leaving the A site now open to accept another aa....



Termination:

- Elongation continues until a stop codon is read in the A site. (UGA, UAG and UAA)
- Stop codons do not code for an aa, so there is no tRNA in the A site.
- Release factor (protein) recognizes the ribosome has stopped and it helps in the release of the polypeptide from the ribosome.
- The two subunits of the ribosome fall off the mRNA.

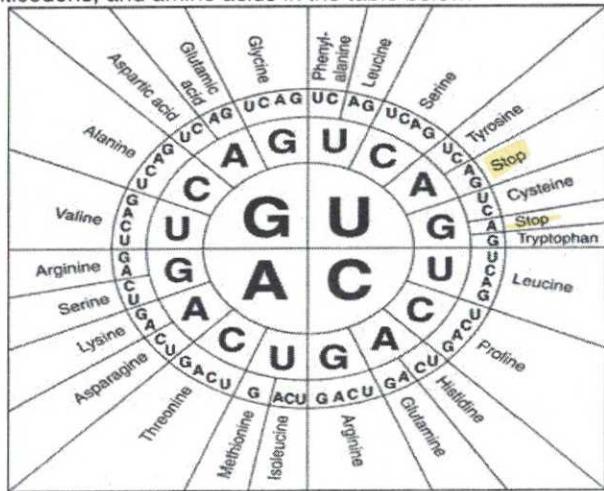


Extra Resources – Transcription and Translation

- Amoeba Sisters: <https://www.youtube.com/watch?v=oefAl2x2CQM>
- Crash Course: <https://www.youtube.com/watch?v=itsb2SqR-R0>

Codon Worksheet

Use the circular codon table to complete the DNA triplets, mRNA codons, tRNA anticodons, and amino acids in the table below.



DNA Triplet	mRNA Codon	tRNA Anticodon	Amino Acid
GGC	CCG	G ₁ G ₂ C	Proline
TTC	AAG	UUC	Lysine
GTC	CAG	GUC	Glutamine
TTA	AAU	UUA	Asparagine
AAA	UUU	AAA	Phenylalanine
GTA	CAU	GUA	Histidine
CTC	GAG	CUC	Glutamate Acid
TGT	ACA	UGU	Threonine
TAT	AUA	UAU	Isoleucine
TCG	AGC	UCG	Serine
ATT	UAA	AUU	Stop
CCA	GGU	CCA	Glycine
GGC	CCG	GGC	Proline

Questions:

1. What 3 codons act as termination signals?

UAA, UAG, UGA

2. What codon means start?

AUG

3. List ALL of the codons for leucine.

CUU CCA UUG
CCC CCG UUA

Name ALL the codons for these amino acids:

4. Phenylalanine –

UUU UUU

5. Serine –

AGC AGU

6. Isoleucine –

7. Valine –

9. Glycine –

10. Alanine –

11. If the DNA sequence is --- AAA TAT CCG TAG CAA ATG, write the mRNA sequence, tRNA anticodon sequence, and the amino acids for this.

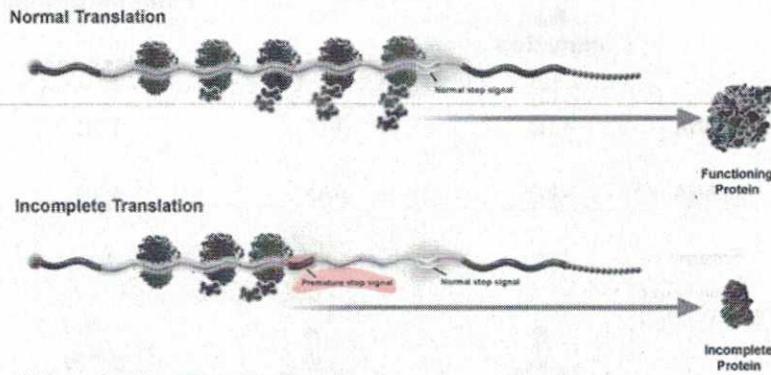
DNA: AAA TAT CCG TAG CAA ATG

mRNA: UUU AUA G₁ G₂ C AUC GUU UAC

Amino acids: phenyl, iso, gly, iso, val, tyro

Mutations

- Errors in the DNA sequence that are inherited.
- Mutations may have negative, positive, or no known side effects.
- Diploid organisms have two copies of each gene so error may be masked in the phenotype.



Point Mutations

- A mutation that occurs at a specific base pair in a genome
- Two general types:
 - Base substitution
 - Frameshift

Base Substitution Mutations

- The replacement of one base in a DNA sequence with another base
- Three types:
 - Silent
 - Missense
 - Nonsense

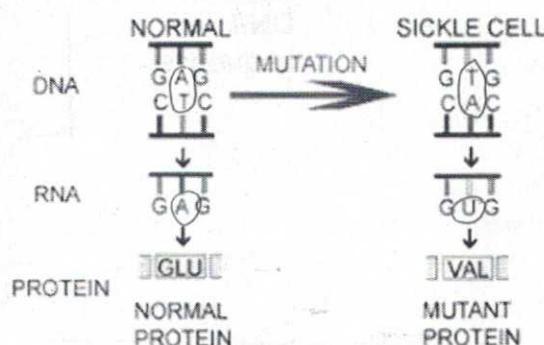
Next Page

Silent Mutation:

- Does not result in a change in the amino acid coded for.
- Does not result in a phenotypic change.
- Primarily occurs in introns.
- Could be due to redundancy of genetic code e.g. phenylalanine coded for by UUU and UUC... a change in the 3rd base does not change the amino acid.

Missense Mutation

- A single substitution of one base in a codon, resulting in a different amino acid.
- E.g., Sickle cell anemia.

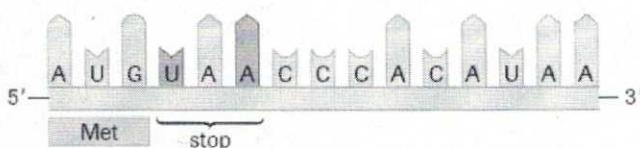


Nonsense Mutation

- A single substitution of one base in a codon that causes a stop codon to replace a codon specifying an amino acid
- Only part of protein produced during translation
- Often lethal to the cell.

Figure 1: Sickle Cell Anemia

nonsense mutation



No mutation		Point mutations		
		Silent	Missense	Nonsense
DNA	TTC	TTT	TCC	ATC
mRNA	AAG	AAA	AGG	UAG
Protein Amino Acid	Lys	Lys	Arg	STOP

Chemical structures of the four amino acids:

- Lysine: A large, positively charged amino acid with a long hydrocarbon side chain.
- Alanine: A small, neutral amino acid with a single methyl group.
- Arginine: A large, positively charged amino acid with a long hydrocarbon side chain and a guanidino group.
- Stop: A termination signal for protein synthesis.

Frameshift Mutation

- Causes changes in the reading frame.
- Caused by an insertion:
 - Addition of one or more base pairs in a DNA sequence.
- Caused by a deletion:
 - Removal of one or more base pairs in a DNA sequence.

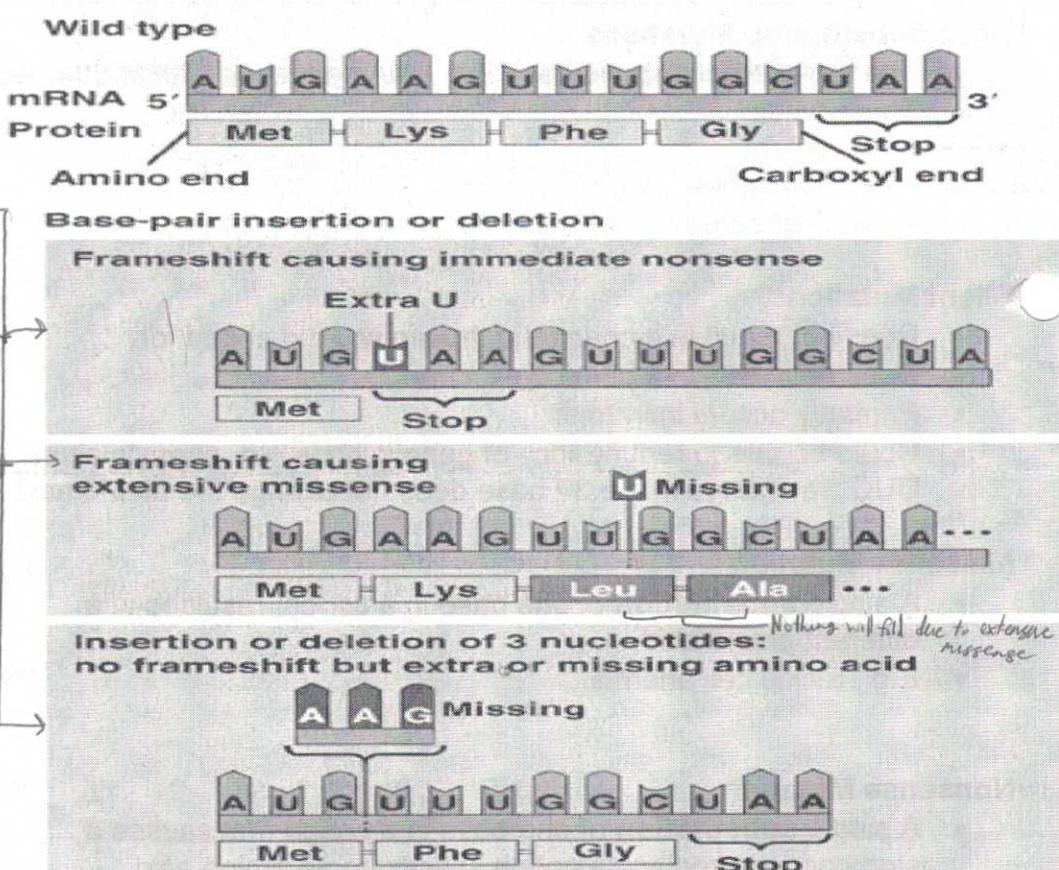
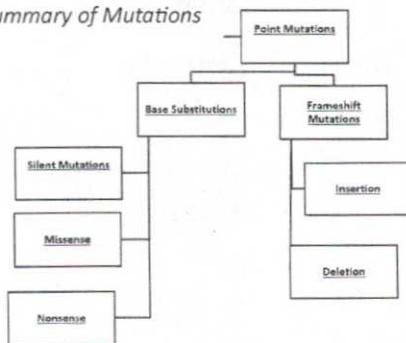


Figure 2: Summary of Mutations



Protein Synthesis and Codons Practice

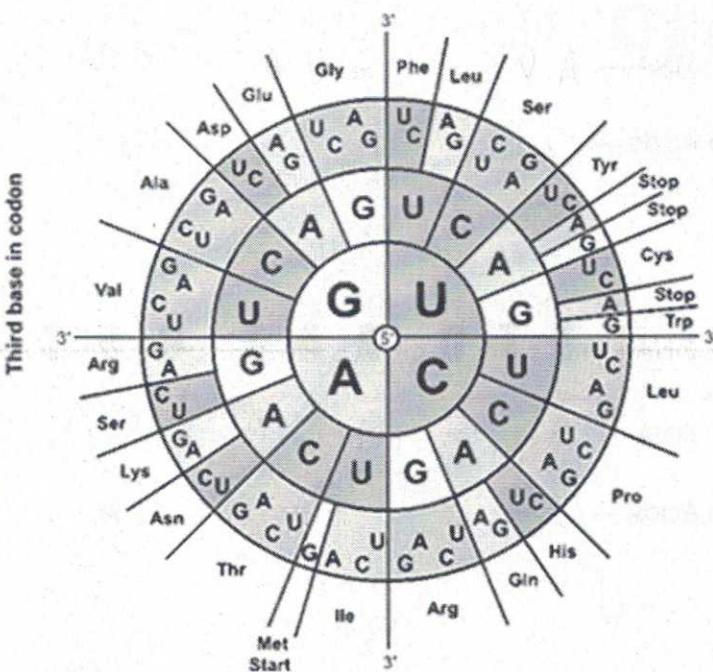
Protein synthesis is the process where a sequence of DNA is used to build a protein from individual amino acids. The first step in this process is called **TRANSCRIPTION**, where a coding region of DNA is converted to messenger RNA (mRNA). During transcription, mRNA is made from the DNA sequence following the base pair rule, except RNA does not contain the base Thymine, but instead has Uracil. The mRNA then leaves the nucleus and goes to a ribosome in the cell's cytoplasm. The ribosome reads the message three bases at a time, called a **CODON**. Each codon will specify a single amino acid. The amino acids are joined together and folded into a protein, a process called **TRANSLATION**.

Key Points

- DNA is used to make a copy of mRNA (transcription)
- mRNA leaves the nucleus and goes to ribosomes
- 3 bases = codon
- 1 codon = a single amino acid
- A chain of amino acids = a protein
- Protein synthesis is also called translation

Biologists use a codon chart or a codon wheel to determine the amino acids. Amino acids are usually abbreviated on these charts as three letter words, like Cys and Ser.

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



1. Use the codon chart to write the amino acid that corresponds to each codon found in mRNA:

CCC Pro
CAG Glu
GAA Glu
UUU Phe

AGU Ser
UAC Tyr
CGU Arg
CCA Pro

2. Write the CODON that corresponds with each amino acid. There may be more than one. The full names are written, but the codon chart only shows the first three letters.

proline CC (GACU)

glycine G G (UCAG)

valine C V (GACV)

phenylalanine UVU (UUC)

histidine CA (UC)

arginine VG (GUAC)

3. A single codon is used to signal the beginning of protein synthesis. It is commonly called the START CODON.

Locate the start codon on the chart. What are the three bases of this codon? A U G

4. There are three codons that signal the end of synthesis, these are called STOP codons.

What are the three stop codons? UAA, UGA, UAG

5. For each sequence of DNA is shown. Write the complementary RNA sequence underneath the letters, then use the codon chart to determine the amino acid sequence:

DNA → **T A C C A T G G A A T T A C T**

RNA → A U G G U A C C U U A A U G A

Amino Acids → Start Val Gly Stop Stop
 met

DNA → **T T C A A T G G T C T A G G G**

RNA → A A G U U A C C A G A U C C C

Amino Acids → Lys Leu Pro Asp Pro

DNA → **A C A T T T C A G A C C G T C**

RNA → U G U A A A G U C U G G C A G

Amino Acids → Cys Lys Val Try Glu

Mutations Worksheet

There are several types of mutation:

FRAMESHIFT – the reading ‘frame’ (i.e. codons) is changed, causing a change in the amino acid sequence

- **DELETION** (a base is lost)
- **INSERTION** (an extra base is inserted)

SUBSTITUTION - one base is substituted for another

- If a substitution **changes** the amino acid, it's called a **MISSENSE** mutation
- If a substitution **does not change** the amino acid, it's called a **SILENT** mutation
- If a substitution **changes the amino acid to a “stop,”** it's called a **NONSENSE** mutation

Complete the boxes below. Under ‘what type of mutation is this?’, classify each as either Frameshift or Substitution **AND** as either deletion, insertion, missense, silent or nonsense

Example #1

Original DNA Sequence: T A C A C C T T G G C G A C G A C T

mRNA Sequence: AUG UGG AAC CGC UG C U G A

Amino Acid Sequence: Start met Try Asp Arg Cys Stop

Mutated DNA Sequence #1: T A C A T C T T G G C G A C G A C T

What's the mRNA sequence? AVG UAG AAC CGC UGC UGA (Circle the change)

What will be the amino acid sequence? Met Stop Asp Arg Cys Stop

Will there likely be effects? Unfinished Protein What kind of mutation is this? Nonsense

Mutated DNA Sequence #2: T A C G A C C T T G G C G A C G A C T

What's the mRNA sequence? AVG CUG GAA CCC CUG CUG A (Circle the change)

What will be the amino acid sequence? Met Leu Glu Pro Leu Leu

Will there likely be effects? Yes What kind of mutation is this? Frame shift Insertion

Mutated DNA Sequence #3: T A C A C C T T A G C G A C G A C T

What's the mRNA sequence? AVG UGG AAU CGC UGC UGA (Circle the change)

What will be the amino acid sequence? Met Try Asp Arg Cys Stop

Will there likely be effects? No What kind of mutation is this? Silent

Mutated DNA Sequence #4: T A C A C C T T G G C G A C A A C T

What's the mRNA sequence? AUG UGG AAC CGC UGU UGA (Circle the change)

What will be the amino acid sequence? Met Try Asp Arg Cys Stop

Will there likely be effects? Yes What kind of mutation is this? Missense

Mutated DNA Sequence #5:	T A C A C C T T G G G A C T
What's the mRNA sequence?	<u>A U G U G G A A C C C U G A</u>
What will be the amino acid sequence?	<u>Met Trp Asn Pro</u>
Will there likely be effects?	<u>Yes</u>

What kind of mutation is this? Frameshift - Deletion

Example #2

Sickle Cell Anemia

Sickle cell anemia is the result of a type of mutation in the gene that codes for part of the hemoglobin molecule. Recall that hemoglobin carries oxygen in your red blood cells. The mutation causes the red blood cells to become stiff and sickle-shaped when they release their oxygen. The sickled cells tend to get stuck in blood vessels.

Analyze the DNA strands below to determine what amino acid is changed and what type of mutation occurred.

Normal hemoglobin DNA **C A C G T G G A C T G A G G A C T C C T C**

Normal hemoglobin mRNA G U G C A C C U G A C U C C U G A G G A G

Normal hemoglobin AA sequence

Sickle cell hemoglobin DNA **C A C G T G G A C T G A G G A C A C C T C**

Sickle cell hemoglobin mRNA

Sickle cell hemoglobin

What type of mutation is this? _____

DNA Mutation Simulation

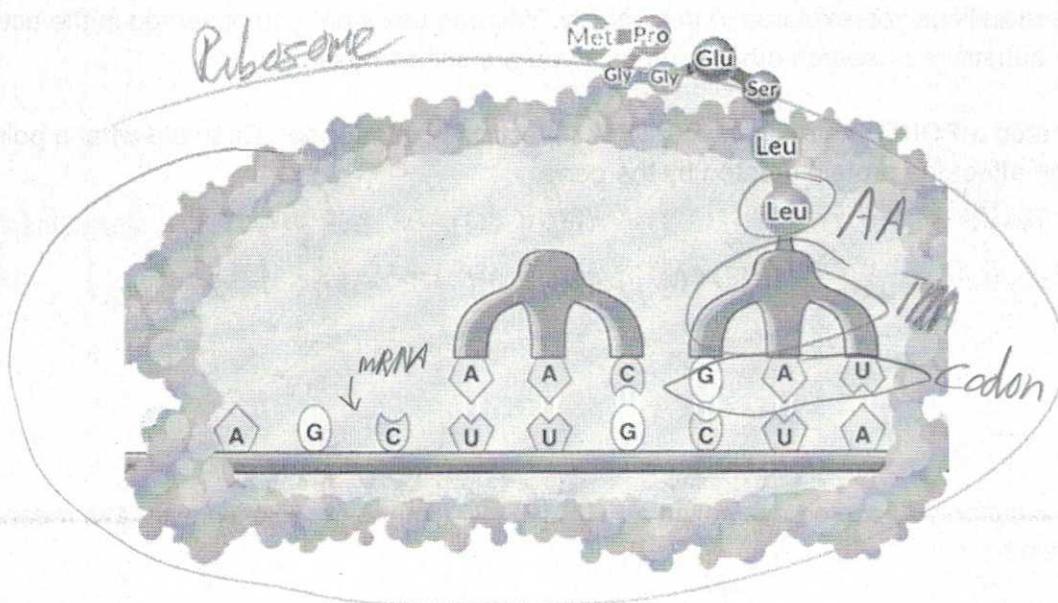
Access the simulation at: <https://www.biologycorner.com/worksheets/DNA-sim.html>

- 1) Transcribe and Translate your original DNA. Review those terms and write a short definition

Transcription: *Transcribe DNA into m-RNA*

Translation: *synthesise protein in the cytoplasm*

- 2) Identify the major players shown in the simulation: mRNA, Codon, Amino Acid, tRNA, anticodon, ribosome.
- Use the figure below to label these parts (if you are doing this digitally, double click the image to add labels in the Google Drawing).



3. When the protein is completed, write the sequence of amino acids shown, there are 11.

(Tip: click the "stop" button to make the model stop jiggling.)

4. Click on edit DNA, you will now see the original sequence used to make the protein.

ATGCCAGGCAGCGAGAGCTTGCTAATTGGCTTATAG

5. Edit the DNA by **changing the first triplet to AAA**

Check the new protein created by your new DNA. Describe how this changed the protein.

6. Return the triplet to its original state (ATG). Now place an additional A after the G, your strand will read ATGA. Check the new protein created by your new DNA. Describe how this changed the protein.

7. Return the triplet to its original state (ATG). Now change the second triplet from CCA to CCC. Check the new protein created by your new DNA. Describe how this changed the protein.

Final Analysis -

There are three mutations you explored in this activity. You can use what you observed in the activity to help you answer the questions or search other sources if you are still confused.

8. First, you created a **POINT** mutation in your DNA (specifically **missense**). Describe what a point mutation is and how this can affect the protein created by the gene.

A mutation in which nucleotide(s) are not properly transcribed & are replaced w/ different base pair - leads to different codons

9. The second mutation you explored is called a **FRAMESHIFT** mutation. Explain what this means and how it affects the protein.

nucleotides are added/removed, shifting the entire sequence.
Can lead to extensive missense or potentially nonsense

10. The third mutation you explored is a special kind of point mutation called a **SILENT** mutation. Explain what this means.

When a base pair is altered but still codes for same aa (redundancy)

Controlling Gene Expression in Prokaryotic Cells

Controlling Gene Expression:

What does it mean?

- Production of a functioning protein (DNA → mRNA → protein)
- Regulation → turn it on/off, speed up/slow down

Gene Expression → Turning DNA code into a functioning protein

Importance/Examples?

- Some genes/proteins are only expressed once during development (e.g. nose keeps growing)
- Some are needed during specific times (puberty)
- Some are produced by environmental stress (calluses)

How can it be regulated? (many ways)

- At the protein level (inhibition, allosteric activation/inhibition, etc.)
- At the DNA level (by regulating transcription/translation)

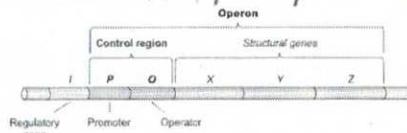
Operons:

Prokaryotic Cells use OPERONS to regulate gene expression

Operon → cluster of genes under control of a promoter & operator

Operator → Regulatory sequences of DNA to which a repressor protein binds

Only prokaryotes use operons to regulate genes



Example: The lac Operon:

Background on Lactose and Humans (not prokaryotic):

- In humans, lactase is an enzyme that breaks lactose into glucose and galactose
- Some humans cannot produce lactase after the age of five
 - Infants have high lactase levels so they can digest their mother's milk
 - Some time after weaning, lactase gene expression is turned off
 - This is lactose non-digestion, a trait that allows adults to continue to digest milk → results from a mutation in a region of the lactase gene

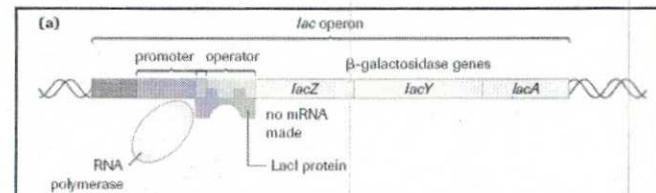
Bacteria and Lactose:

E. coli in the intestines of mammals use the energy supplied from the breakdown of lactose for growth

- Use the enzyme β -galactosidase

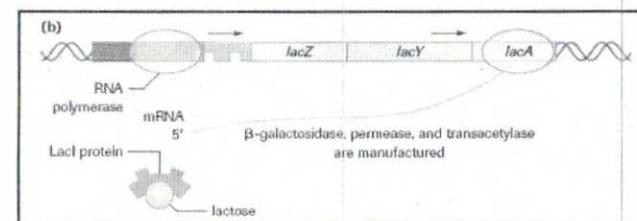
- It would not be economical for *E. coli* to produce the enzyme when lactose is not available (as mammal ages and milk intake decreases)
 - *E. coli* uses a negative regulation system to control transcription and translation of β -galactosidase.
- The genes for β -galactosidase are part of an Operon
- The *lac* operon consists of a cluster of three genes: *lacZ*, *lacY*, *lacA*

The lac Operon – the Details:



When lactose is absent, the *LacI* repressor protein blocks transcription of *lac* operon genes by binding to the *lac* operator and blocking RNA polymerase.

- This prevents RNA polymerase from reaching and transcribing the genes further down the strand



When lactose is present, the *LacI* repressor must be removed so that β -galactosidase can be made

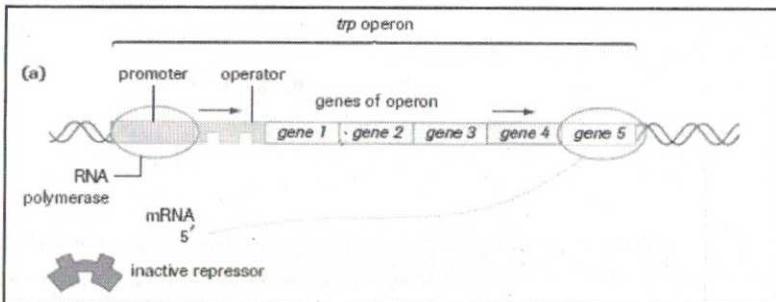
- Lactose acts as a signal molecule (inducer) by binding to *LacI* repressor protein.
- *LacI* changes its shape and can no longer bind to operator
- this allows transcription of the *lac* operon genes.

Example – The trp Operon:

Background info on Tryptophan:

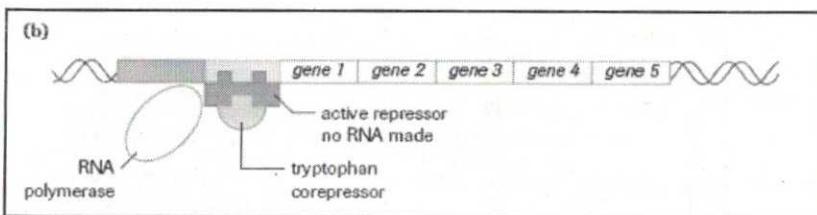
- *E. coli* uses the amino acid tryptophan to produce protein.
 - If *E. coli* can obtain tryptophan from the environment (mammal's diet) it does not need to make it
 - If it cannot, it must make its own
- The *trp* operon is used to regulate expression of tryptophan genes
- *trp* operon is active when high levels of tryptophan are present
- *trp* operon consists of 5 genes that code for 3 enzymes needed to synthesize tryptophan.

The trp Operon: The Details:



When tryptophan is absent, the shape of the *trp* repressor protein is such that it cannot bind to the *trp* operator.

- RNA polymerase transcribes *trp* operon genes and the enzymes that synthesize tryptophan are produced by *E. coli*



When tryptophan is present:

- Tryptophan binds to the *trp* repressor protein altering its shape.
- The *trp* repressor-tryptophan complex binds to the *trp* operator.
- RNA polymerase is unable to bind to promoter region and the genes are not expressed
- Since tryptophan is needed to inactivate the *trp* operon, it is called a corepressor.

Comparing the lac and trp Operons:

<i>lac</i> Operon	<i>trp</i> Operon
It regulates the production of <u>B-Galactose</u> (involved in lactose metabolism)	It regulates the production of the aa <u>tryptophan</u> (by regulating the production of enzymes involved in synthesizing trp)
It consists of a cluster of <u>3 genes</u> genes under the control of one promoter and one operator	It consists of a cluster of <u>5 genes</u> genes under the control of one promoter and one operator
The LacI repressor protein binds to the operator when lactose levels are <u>low</u>	The corepressor tryptophan binds to the <i>trp</i> repressor protein and the complex binds to the operator when tryptophan levels are <u>high</u>
High levels of lactose <u>activates</u> the operon	High levels of tryptophan <u>deactivates</u> the operon
- Therefore, lactose is an <u>inducer</u>	- Therefore, tryptophan is a <u>corepressor</u>

Extra Resource - Amoeba sisters - https://www.youtube.com/watch?v=h_1QLdtF8d0

Regulating Protein Synthesis - Eukaryotic Cells

Overview

- Eukaryotic cells, and their methods of protein synthesis, are more complex than prokaryotic cells.
 - (note that eukaryotic cells DO NOT have operons)
- In the Biochemistry unit, we discussed regulation at the protein level (with respect to enzymes)
- The Control Mechanisms for Eukaryotic cells fall into 4 main categories:
 - Transcriptional
 - Post-transcriptional
 - Translational
 - Post-translational

Transcriptional Regulation

- Regulation is at the level of transcription
- Most common
 - Ex 1: the gene must be activated (DNA released from histones) for transcription to occur
 - Uses activator molecules
 - if not activated, transcription does not occur
 - Ex 2: Methylation
 - A methyl group ($-CH_3$) is added to some bases in the promoter region of a gene
 - This is done by a group of enzymes called methylases
 - This prevents transcription (effectively puts it 'on hold' temporarily)
 - Acts as a 'roadblock' for RNA Polymerase
 - Called 'gene silencing'
 - Methylation has a lot of biotech applications as well

Activation of eukaryotic transcription

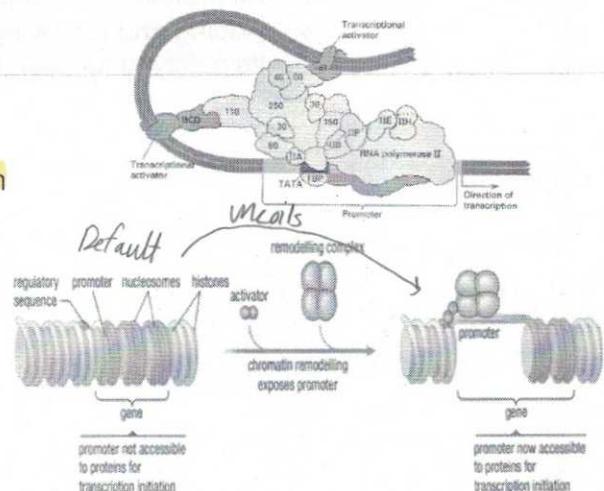
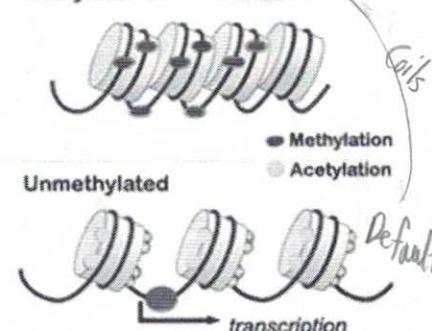


Figure 3 A chromatin remodelling complex exposes the promoter, and thus enables transcription.

Methylated DNA



Post-Transcriptional Regulation

- Regulation of the mRNA, and its ability to be translated
 - Ex 1 - Alternative Splicing
 - (Discussed in more detail during transcription)
 - Some primary transcripts can be used to make multiple polypeptides, and which introns are removed vs left in, and which order they are sequenced, determines which polypeptide will be translated
 - (Therefore, what is considered an 'intron' and an 'exon' will vary depending on the polypeptide)
 - Ex 2 - mRNA degradation
 - A regulatory molecule (eg a hormone) can impact the rate at which an mRNA is broken down
 - mRNA can be used repeatedly as the template for translation, so if the cell no longer wishes to make the protein, the regulatory molecule will start to have the mRNA degraded

Translational Regulation

- Regulation the rate of translation
 - Ex 1 - the availability of amino acids
 - The availability of an amino acid will impact the rate of translation

Post-Translational Regulation

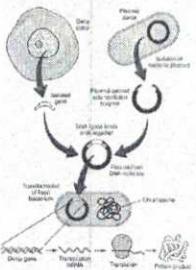
- Regulation the polypeptide
 - Ex 1 - **Processing**
 - Polypeptides must be processed after formation
 - Specific proteins, called **chaperones**, help it fold
 - Regulating these proteins can determine if the protein has the correct conformation
 - Ex 2 - **Chemical Modification**
 - Chemical groups, like methyl groups, are added to the polypeptide, which can impact its conformation, and therefore its ability to function
 - Ex 3 - **Degradation**
 - All proteins are subject to degradation in the cell
 - Adding and removing chemical tags that 'target' the protein for degradation can impact its ability to function in the cell

Genetic Engineering

Tools and Techniques for Recombinant DNA

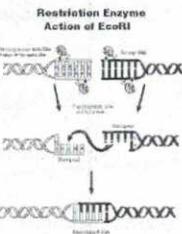
What is Recombinant DNA?

- Taking a piece of DNA, and combining it with another strand of DNA
- this creates sequences that would not otherwise be found in the genome



Tool - Restriction Enzymes (Endonucleases)

- Natural Role?
 - Found in bacterial DNA
 - Provides a crude immune system for bacteria against bacteriophage viruses
- What does it do?
 - Cuts double stranded DNA at a specific base pair sequence
 - Usually recognizes short, palindromic sequences
- Role in Genetic Engineering:
 - Cuts DNA in a predictable manner (uses a recognition site)
 - Can use DNA strands to form recombinant DNA



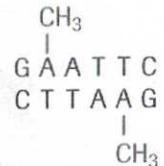
Sticky VS Blunt Ends:

- Restriction enzymes can cut DNA sequences leaving two different types of ends:
 - 'sticky' – ends have overhangs
 - 'blunt' – ends are flat/ no overhang
- In biotechnology, sticky ends are more useful
 - Ends can be more easily joined together because hydrogen bonds
 - Will form between base pairs in the overhang
 - Will not happen in blunt ends – rely solely on phosphodiester bonds to attach blunt ends

Microorganism of origin	Enzyme	Recognition site	After restriction enzyme digestion
<i>Escherichia coli</i>	EcoRI	5'-GAATTC-3' 3'-CTTAAAG-5'	5'-G AATTC -3' 3'-C TAA G-5'
<i>Serratia marcescens</i>	SmaI	5'-GGGCC-3' 3'-CCGGG-5'	5'-GGG CCC -3' 3'-CCG GG -5'
<i>Arthrobacter luteus</i>	AluA	5'-AGCT-3' 3'-TCGA-5'	5'-AG CT -3' 3'-TC GA -5'
<i>Streptomyces albus</i>	SafI	5'-GTGCAAC-3' 3'-CAGGTG-5'	5'-G TGCA C-3' 3'-C AGGT G-5'
<i>Haemophilus parainfluenzae</i>	HincII	5'-AAGCTT-3' 3'-TTGAA-5'	5'-A AAGCTT -3' 3'-T TGAA -5'

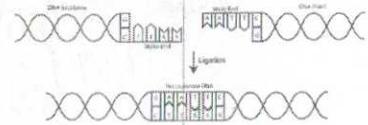
Tool - Methylases

- Natural Role?
 - Found in bacteria (also found in eukaryotes)
 - Recently restriction enzymes from clearing the bacterium's own DNA
 - Foreign DNA (viral) is not methylated, so the restriction enzymes can cleave their DNA
- What does it do?
 - Adds a methyl group (CH_3) to a nucleotide in the recognition site for a restriction enzyme
- Role in Genetic Engineering
 - Protect genes from being cut at an undesirable location



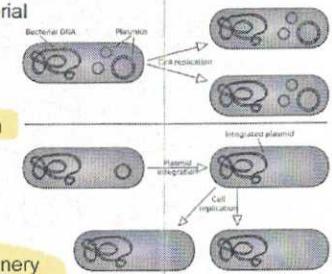
Tool - Ligase

- Natural Role?
 - Splice together sections of DNA by forming phosphodiester bonds
 - E.g. during DNA replication to join Okazaki fragments
 - Found in prokaryotic and eukaryotic cells
- Role in Genetic Engineering
 - DNA Ligase forms the phosphodiester bonds for the new sequences of DNA to be added into an existing genome



Plasmids

- Use _____ as the _____
- What is it?
 - Part of bacterial genetics
 - Found in bacterial cytoplasm (separate from bacterial DNA loop)
 - A small, circular piece of DNA that can exit and enter other bacterial cells
- Natural Role?
 - Carry genes for proteins that may be different than that of primary DNA
 - Plays a role in creating diversity (and antibiotic resistance, etc.) in bacteria
 - Bacteria normally reproduce asexually
- Role in Genetic Engineering
 - Bacterial cells are often used as the cellular machinery for expressing a genetically modified gene
 - Isolated genes are inserted into plasmids to form recombinant DNA
 - Plasmids can enter bacteria and express the new proteins
 - DNA strand is copied as the plasmid replicates



Gene Editing with CRISPR-Cas9

Why all the hype?

- It is faster, cheaper, and more accurate than previous gene editing technologies

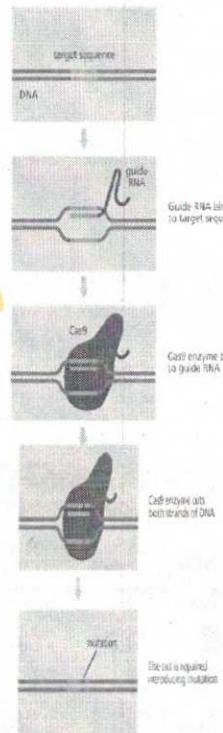
Natural Role?

- Used by bacteria to splice invading viral DNA

How does it work?

2 Components:

- Enzyme (Cas9) - 'molecular scissors' that cut the double stranded DNA at a very specific location so that bits of DNA can be added or removed. Cas9 can open the double helix as well as cut.
- Guide RNA (gRNA)- small piece of pre-designed RNA sequence located within a longer DNA scaffold. The scaffold binds to the DNA and the pre-designed sequence guides Cas9 to the right part of the genome to be cut. The gRNA is complementary to those in the targeted DNA sequence
- Once it is cut, the cell recognizes that the DNA is damaged and tries to repair it
- Scientists then use the DNA repair machinery to introduce changes to one or more genes in the cell of interest



Research and list two ways CRISPR has already been used:

DNA FINGERPRINTING

DNA Fingerprinting using RFLPs

Fingerprint □ DNA sequences are unique to individuals, just like 'fingerprints' are

RFLPs - Restriction Fragment Length Polymorphisms

- A population is polymorphic for the number of nucleotides between cuts from restriction enzymes
 - The differences between individual's DNA, when using the same restriction enzyme, are called RFLPs
- The fragments can be separated by **gel electrophoresis**
 - This provides a pattern of bands that is unique for the DNA being analyzed
 - This is the basis of **DNA Fingerprinting**
- Can also be used to identify groups of people at risk for genetic disorders

DNA Fingerprinting using STRs

Short tandem repeats - short sequence (2-10 bp) of nucleotides repeated many times

- Number of repeats in an STR varies from person to person
- The more STRs in common between two samples, the more likely they are a definitive match

agarose

- Placed into agarose
- Apply current
- DNA moves - to +
- Longer moves slow, shorter moves fast
- Nylon press on agarose & DNA dries.
- Add complementary radioactive DNA
- Photo plate on nylon
- Radioactivity affect plate rendering DNA & bands visible

Restriction Enzymes: How is DNA Manipulated?

Using restriction enzymes, foreign genes can be added to an existing organism (or an embryo). This organism has been genetically modified.

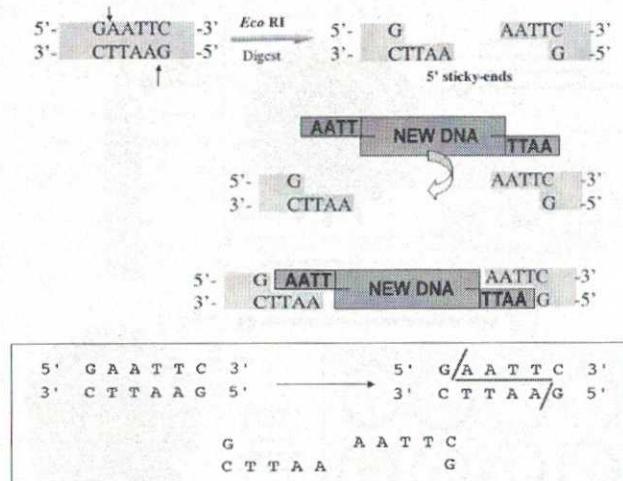
Adding new genes can create plants that are more resistant to pests or be more tolerant to weather patterns, such as drought. This technology can also be used to mass produce chemicals, such as human growth hormone, by inserting that gene into bacteria.

In order to combine the DNA, a chemical called a **restriction enzyme** is used to cut the DNA into fragments, exposing the gene of interest. On either side of the gene is an area of DNA called the "sticky end." The bases of the sticky end are ready to be paired to the new DNA following the base-pair rule. That gene is then spliced to the other organism's DNA using another enzyme called **ligase**, the joining procedure is called ligation.

Sometimes scientists want a gene removed without the sticky ends, so that it will not bind to other parts of DNA. These fragments are said to have "**blunt ends**".

*Recombinant DNA is often abbreviated as **rDNA** to denote that it has foreign genes (DNA) inserted into its genome.

This image shows a restriction enzyme called EcoRI being used to cleave a section of DNA. Different restriction enzyme will cleave DNA at different sequence points, called **recognition sites**, so scientists must isolate the right enzyme for the job – one that cuts around the desired gene.



You will often see restriction enzyme sites written like this: **G / A A T T C**.

How is DNA Manipulated?

1. Explain the following terms and their role in recombinant DNA technology

- a) Restriction enzyme
b) Recognition site

*cleaves dna at specific points
site that restriction enzymes
can cut*

- c) Sticky end
d) Ligation
e) rDNA

*has overhangs
base bonding nucleotides
recombinant dna*

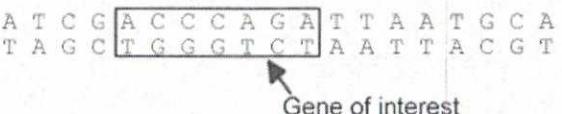
Here are some restriction enzymes and their sites. (Also on your ppt)

BAMHI	G / G A T T C	PstI	C T G C A / G
HindIII	A / A G C T T	Hhal	G C G / C
EcoRI	G / A A T T C	HpaII	C / C G G
Sall	G / T C G A C		
		HindII	G T C / G A C C A G / C T G (blunt ends)

2. On each of the sequences below, determine which restriction enzyme could be used to splice the DNA and indicate where the cut will be made and the enzyme used.

A. 5' T T T G A A T T C A G A T 3'	Enzyme: _____
B. 3' A A A C T T A A G T C T A 5'	Enzyme: _____
5' G T G G G A T C C C T T A 3'	Enzyme: _____
3' C A C C C T A G G G A T 5'	Enzyme: _____
C. 5' A C G C C T C C G G A G A 3'	Enzyme: _____
3' T G C G G A G G C C T C T 5'	Enzyme: _____
D. 5' T T A A G C T T A A G A A G C T T 3'	Enzyme: _____
3' A A T T C G A A T T C T C G A A 5'	Enzyme: _____
E. 5' S A A G C G C G T C G A C T T A T A 3'	Enzyme: _____
3' T T C G C G C A G C T G A A T A T 5'	Enzyme: _____

3. Create a restriction enzyme that will remove the gene of interest. Give it a name too!

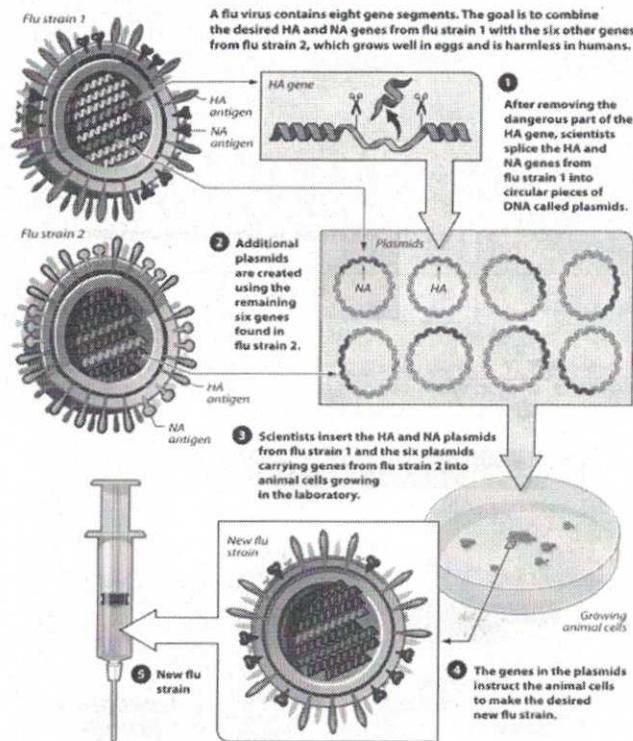


4. The following DNA sequence is from a virus that is dangerous, scientists want to use a restriction enzyme to cut the virus into bits. They do not need sticky ends because they do not plan to combine it with other DNA. Use HindII to show how this DNA would be cut. How many pieces would you have? _____

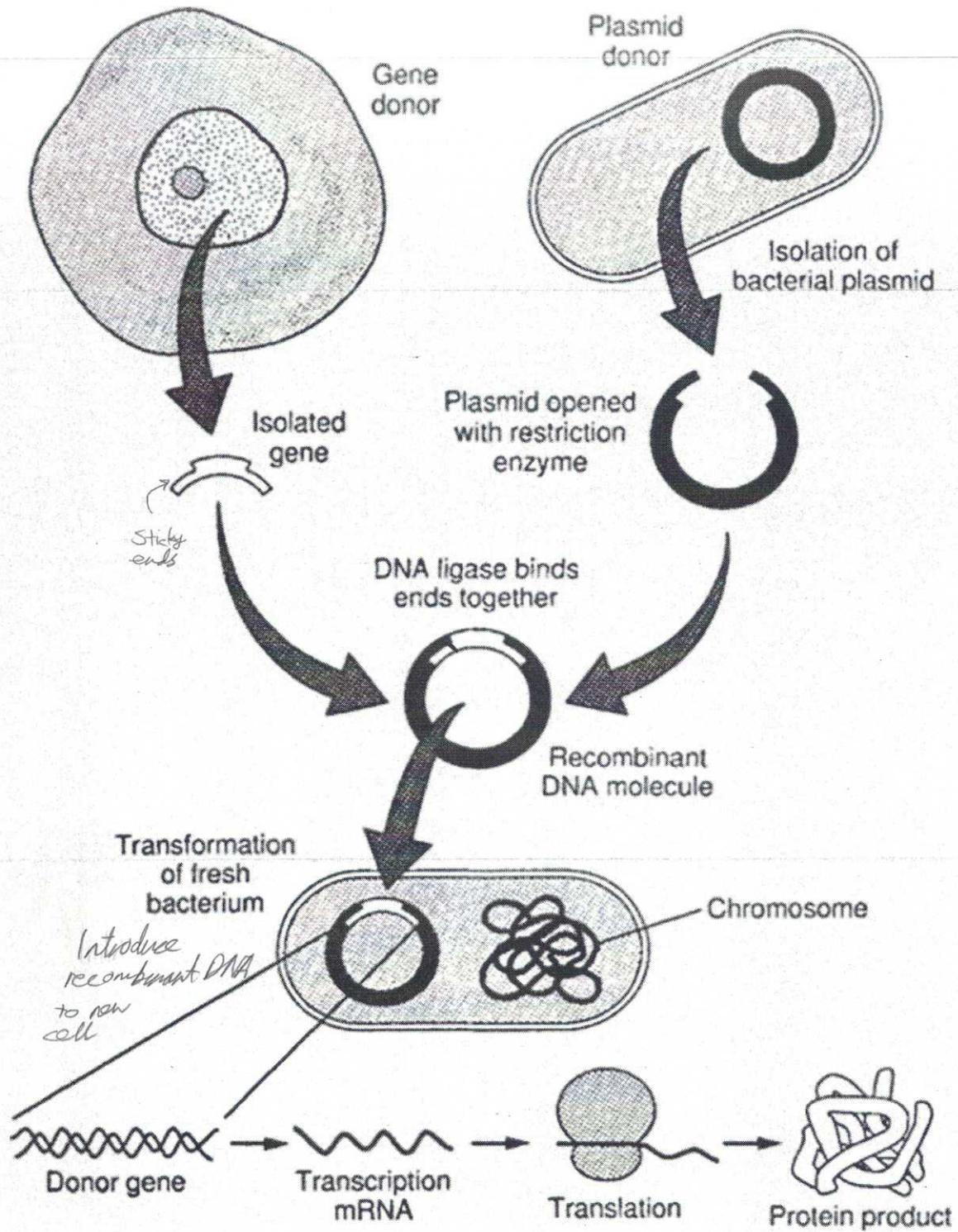
G A A A G T C G A C A A G G C A G T C G A C T T T T A A A A G T C G A C A T G C
C T T T C A G C T G T T C C G T C A G C T G A A A A T T T C A G C T G T A C G

How is the Recombinant DNA (rDNA) Used?

5. Check out this graphic showing how the AVIAN FLU vaccine was created:



- Which genes are being isolated from the flu virus?
- Where are the flu genes (HA and NA) placed?
- Where do plasmids come from and why are they important in this process?
- What is step 3 also known as (1 word)? Explain two different methods scientists use to accomplish step 3.
- What are the potential applications for making a new flu?



Recombinant Plasmid Activity Instructions

Instructions for making your own GMO:

1. Cut out the plasmid base sequence strips and tape them together in one long strip. The letters should all be in the same direction. Tape the two ends of the long strip together to form a circle – with the letters facing out. **THIS IS YOUR PLASMID DNA.**
2. Cut out the DNA base sequence strips and tape them together in numerical order. This is your **human DNA**, which contains the gene for insulin production. The gene area is shaded.
3. Cut out the restriction enzyme sequence cards. Each card shows a sequence where a particular restriction enzyme cuts DNA.
Eco RI HPA II Msp I
4. Compare the sequence of base pairs on an enzyme card with the sequences of the plasmid base pairs. If you find the same sequences of pairs in **both the enzyme card and the plasmid strip**, mark the location on the plasmid with a pencil and write the enzyme number in the marked area. Repeat this step for each enzyme card. Some enzyme sequences may not have a corresponding sequence on the plasmid, and that some enzymes sequences may have more than one corresponding sequence on the plasmid. In this step, you are simulating the process of choosing the correct restriction enzyme to recombine your DNA. With hundreds of restriction enzymes available, scientists must determine which one will work for the DNA they want to recombine.
Eco RI HPA II Msp I
5. Once you have identified all corresponding enzyme sequences on the plasmid, identify those enzymes which cut the plasmid once and only once. Discard any enzymes that cut the plasmid in the shaded plasmid replication sequence (you don't want to cut out this particular gene because it is necessary for the bacteria to replicate itself). Which enzymes fit these criteria?
6. Compare the enzymes you chose in step 5 against the human cell DNA strip. Find any enzymes that will make **two cuts in the DNA, one above the shaded insulin gene sequence and one below the shaded insulin gene sequence**. Mark the areas on the DNA strip that each enzyme will cut and make a note of which enzyme cuts in that spot.
7. Select one enzyme to use to make the cuts. The goal is to cut the DNA as closely as possible to the insulin gene sequence without cutting into the gene sequence. Make cuts on both the plasmid and the DNA strips. **Make the cuts in the staggered fashion indicated by the black line of the enzyme card (making sticky ends).**
8. Take the sticky ends of the plasmid to the sticky ends of the insulin gene to create their recombinant DNA. In the lab, DNA ligase is used to bind the strands together.

Congratulations!

You have successfully created a bacterial cell that contains the human insulin gene. This bacterium will reproduce and create more bacteria with the gene. Bacteria grown in cultures can now mass produce insulin for diabetes.

Discussion Questions:

1. Why was it important to find an enzyme that would cut the plasmid at only one site? What could happen if the plasmid were cut at more than one site?
Either the plasmid loses a chunk of DNA or only that gene is kept
2. Which restriction enzyme did you use? Ask other groups what they used and compare the final transgenic plasmids. Why might there be some different lengths?
Eco RI (should have been HIn dIII)
3. Why was it important to discard any enzymes that cut the plasmid at the replication site?
So the base pairs are complementary & properly bond
4. Why is it important to cut the plasmid and the human DNA with the same restriction enzyme?
Yes, serve as a crude immune system for plasmids
5. Do restriction enzymes exist naturally in organisms? If so, what is their purpose?
Blunt → Only phosphodiesters
6. Why would restriction enzymes that create 'blunt' ends not be as useful in recombination as those that create sticky ends?
Sticky → Hydrogen bonds + phosphodiester
7. In this activity, you simulated creating a recombinant bacteria organism. State what the following materials represented:
 - a. Scissors
 - b. Tape