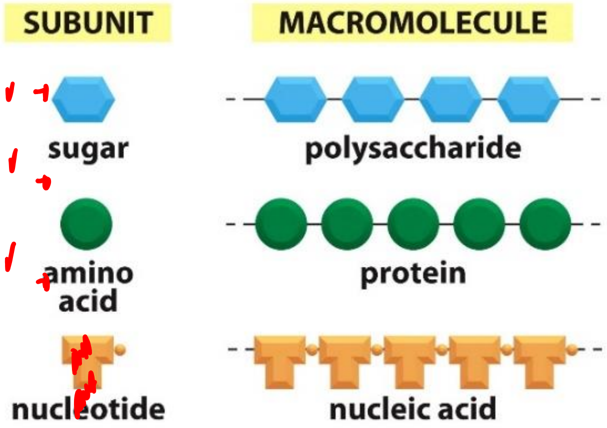
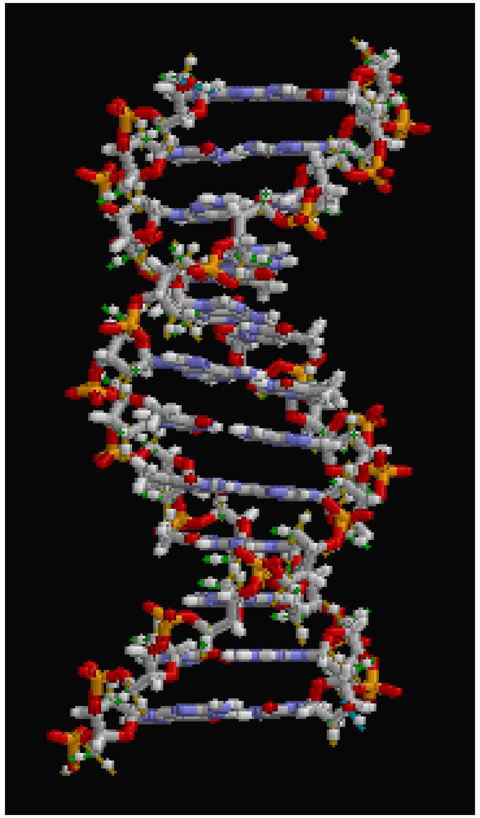
Lecture 3 – Pre-lecture Material
Nucleic acid Structure and DNA Polymerases

Nucleic Acids & Central Dogma



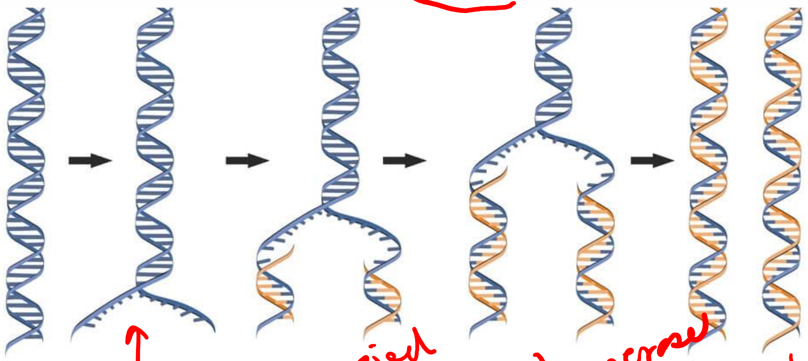


Double stranded DNA

Introduction to Central Dogma (information flow)

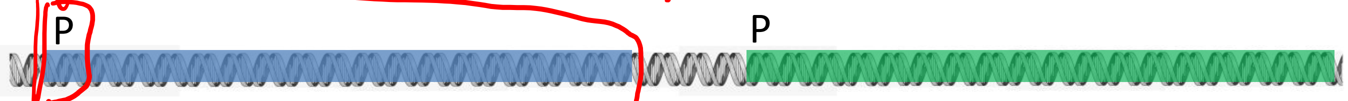
Genome: Entire DNA content of an organism, contains all of the instructions for life. Single circular molecule in Proks, multiple linear molecules (chromosomes) in Euks. The genome is *replicated* when cells divide.

promoter. dsDNA



-AGCT-

Gene – a segment of DNA that is converted (*transcribed*) to RNA. A *promoter* (P) sequence on the DNA is the minimal requirement for the production of RNA.



RNA: non-coding

Transcription

mRNA

Possible functions

miRNA (euk)

Splicing (euk)

Telomerase (euk)

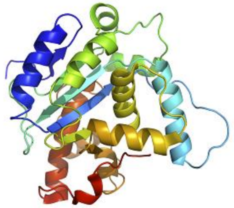
tRNA (p syn)

rRNA (p syn)

mRNA Processing (Euks)

G AAAAA

Translation



protein.

RNA molecules are often processed in **Eukaryotic cells** before they are functional

Many RNAs are functional on their own

mRNA are *translated* to a protein.

The Genetic Code

codon table.

First base	Second base: U	Second base: C	Second base: A	Second base: G	Third base
U	UUU, UUC – Phenyl-alanine	UCU, UCC, UCA, UCG – Serine	UAU, UAC – Tyrosine	UGU, UGC – Cysteine	U, C
	UUA, UUG – Leucine		UAA – Stop codon; UAG – Stop codon	UGA – Stop codon; UGG – Tryptophan	A, G
C	CUU, CUC, CUA, CUG – Leucine	CCU, CCC, CCA, CCG – Proline	CAU, CAC – Histidine	CGU, CGC, CGA, CGG – Arginine	U, C
			CAA, CAG – Glutamine		A, G
A	AUU, AUC, AUA – Isoleucine	ACU, ACC, ACA, ACG – Threonine	AAU, AAC – Asparagine	AGU, AGC – Serine	U, C
	AUG – Methionine (start codon)		AAA, AAG – Lysine	AGA, AGG – Arginine	A, G
G	GUU, GUC, GUA, GUG – Valine	GCU, GCC, GCA, GCG – Alanine	GAU, GAC – Aspartic acid	GGU, GGC, GGA, GGG – Glycine	U, C
			GAA, GAG – Glutamic acid		A, G

. . . ATATGCCCATGTGGTAA . .

(DNA Sequence)

codons.

. . . AUAUGCCCAUGUGGUAA . .

(mRNA Sequence)

. . U–**AUG**–CCC–AUG–UGG–**UAA**

(Punctuated RNA sequence – how the ribosome interprets the sequence)

Met - Pro - Met - Trp

(Protein Sequence)

Note:

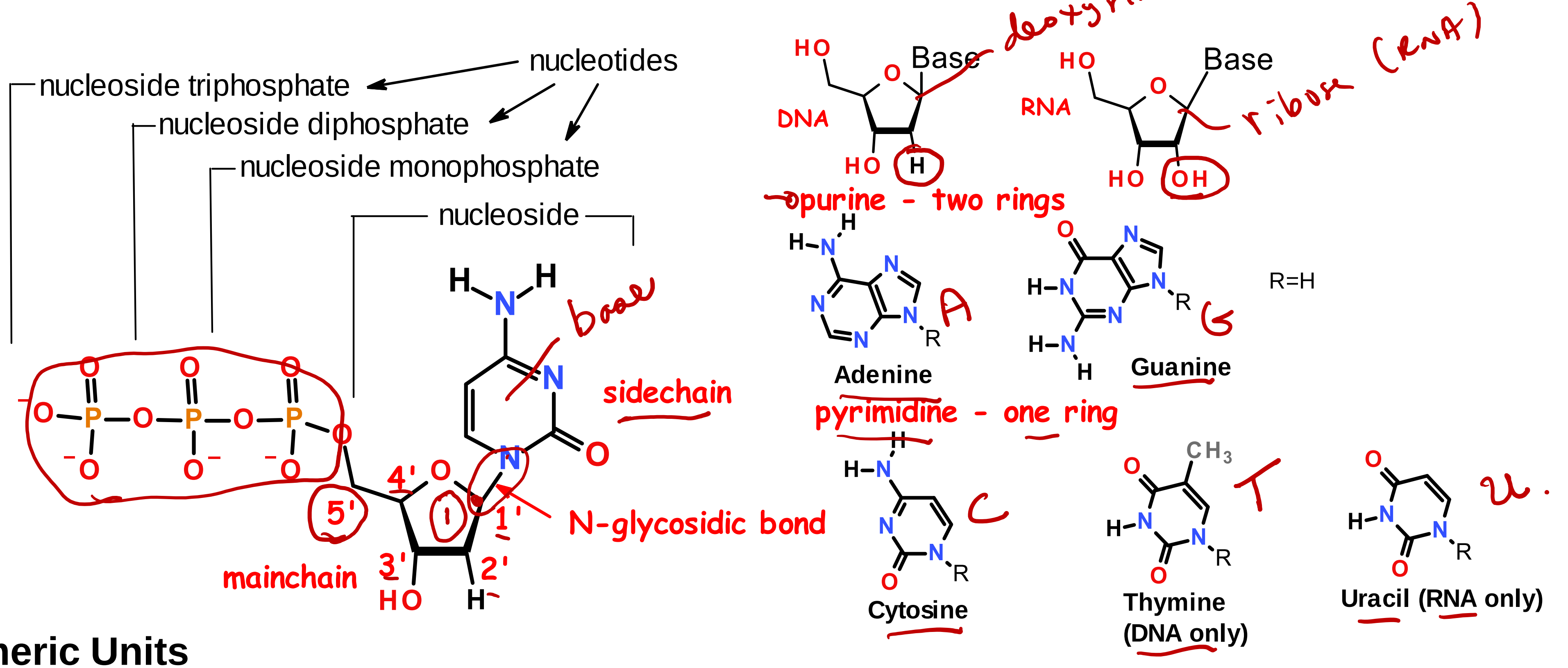
- Each codon codes for one amino acid.
- Many amino acids are coded by more than one codon.
- Most organisms use the same codon table – some codons have different meanings in some organisms.

Special Codons:

AUG = Is used to begin almost all proteins that are synthesized on the ribosome, it also codes for methionine when found internally.

UAA, UAG, UGA = stop codons (do not code for any amino acids), terminate synthesis

Nucleic Acid Structure



1. Monomeric Units

a) Nucleoside triphosphates are the building blocks of nucleic acids (**dNTP = dATP, dGTP, dCTP, dTTP**)

b) The base ("sidechain") is connected to the C1' of the sugar ("mainchain") by an **N-linked glycosidic** bond.

Base + sugar = **nucleoside**.

Base + sugar + n-phosphates = **nucleotide**

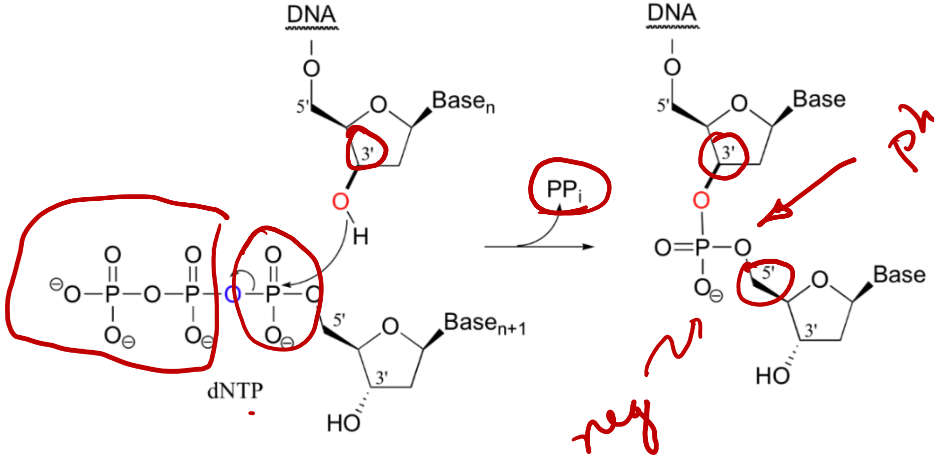
c) The carbon atoms on the sugar are numbered 1' to 5'. The primes distinguish the atoms on the sugar from those on the base.

d) DNA differs from RNA in the sugar (deoxyribose versus ribose) and one base.

e) Four different monomers, A, G, C, T in DNA. U replaces T in RNA.

DNA (& RNA) are directional

How Triphosphates are added to the polymer.

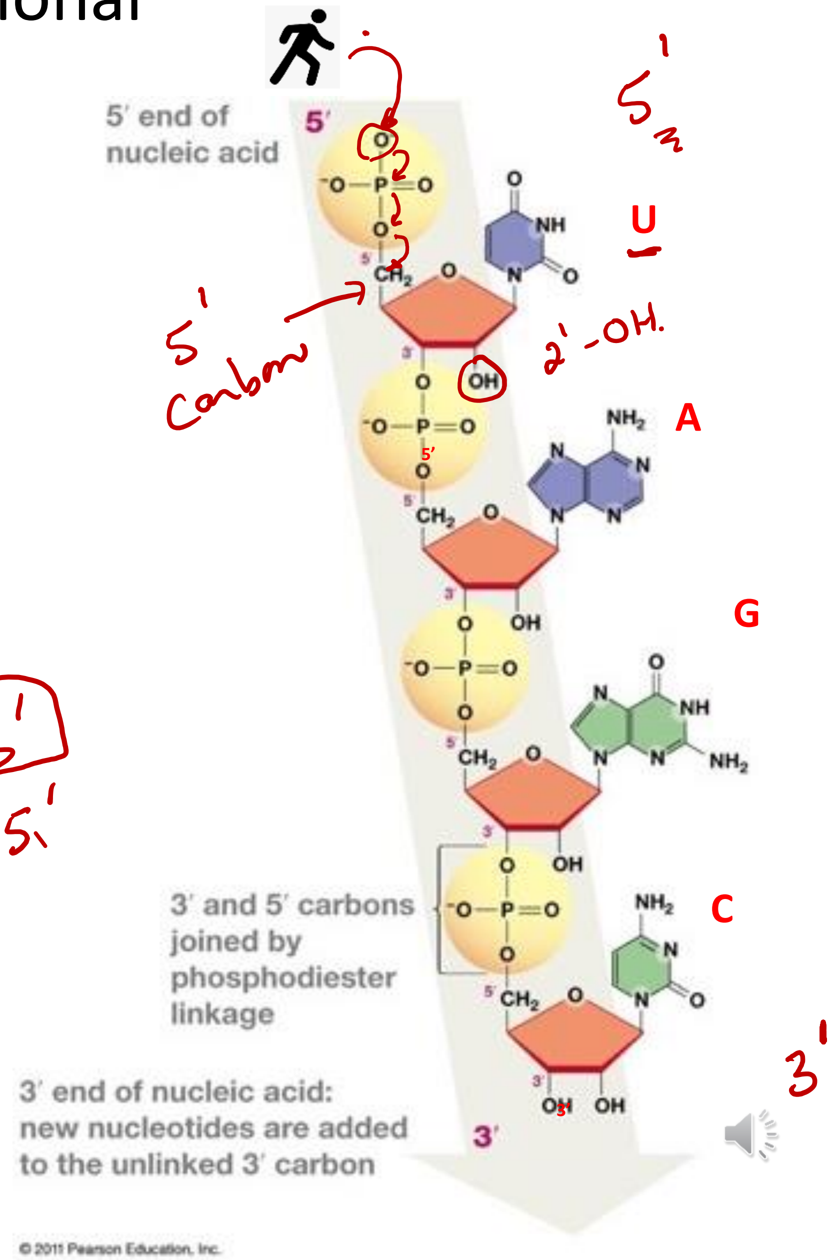


What are the two different ways we could write the sequence of this nucleic acid?

$\boxed{5' \rightarrow 3'}$

$3' \rightarrow 5'$

5' U A G C 3'

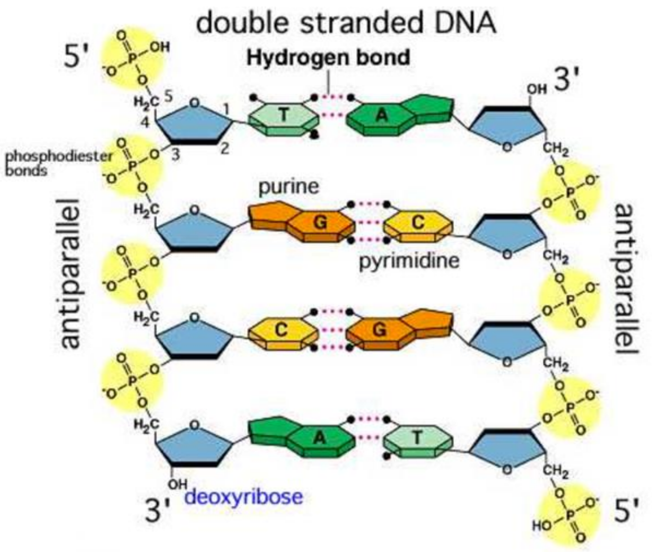


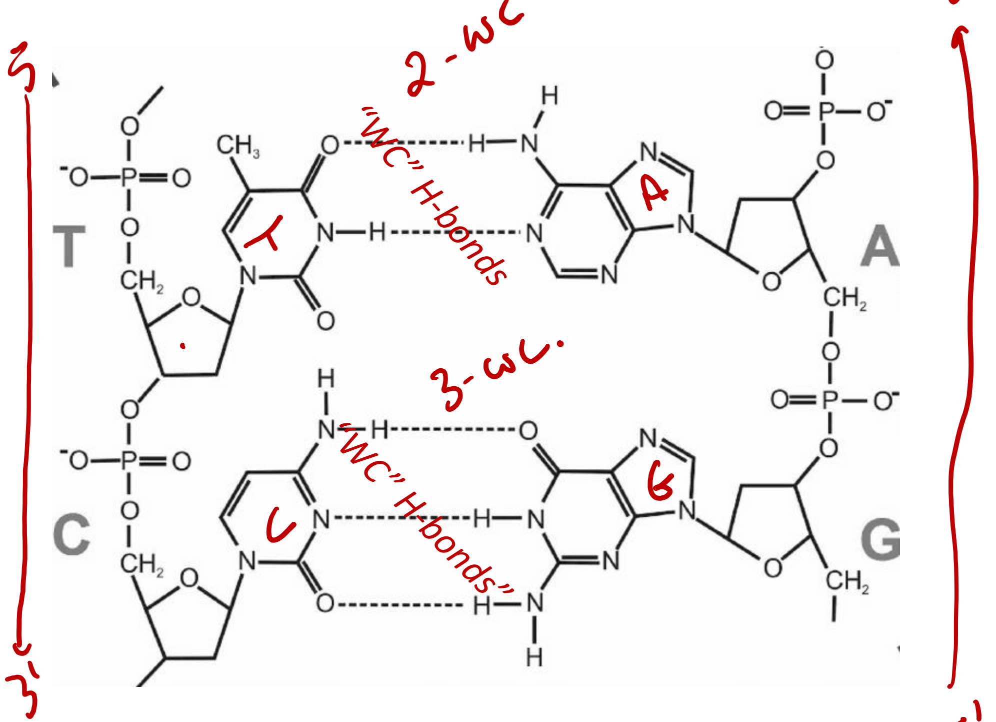
Double Stranded DNA structure

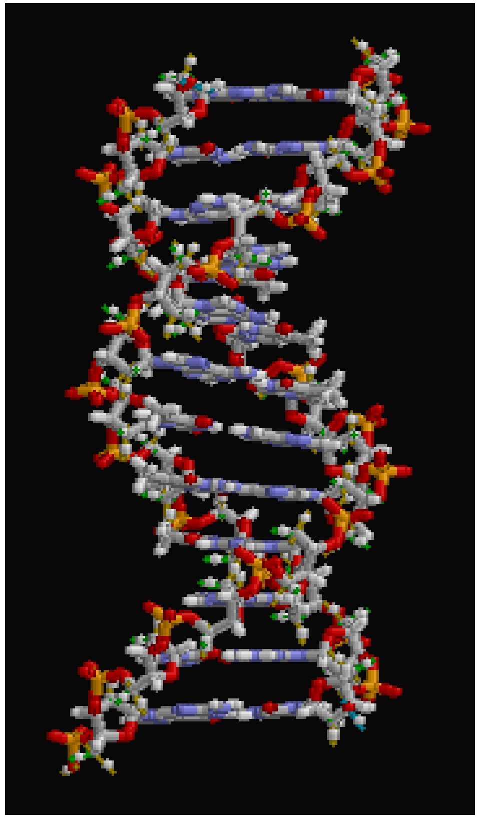
Complementary base pairing: Hydrogen bonds form between bases, thus linking the 2 stands with weak non-covalent interactions.

DNA twisted into double helix

- Strands anti-parallel
- Sugar-phosphate backbone outside
- Nucleotide bases project inward.
- Basepairs are stacked on each other.
- Uniform width
- H-bonds between bases:
    - A=T (two h-bonds)
    - G Ξ C (three h-bonds)







How to indicate the sequence of ssDNA & dsDNA?

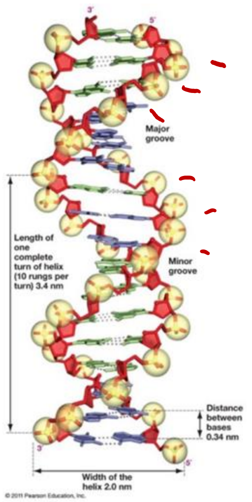
5' → 3'

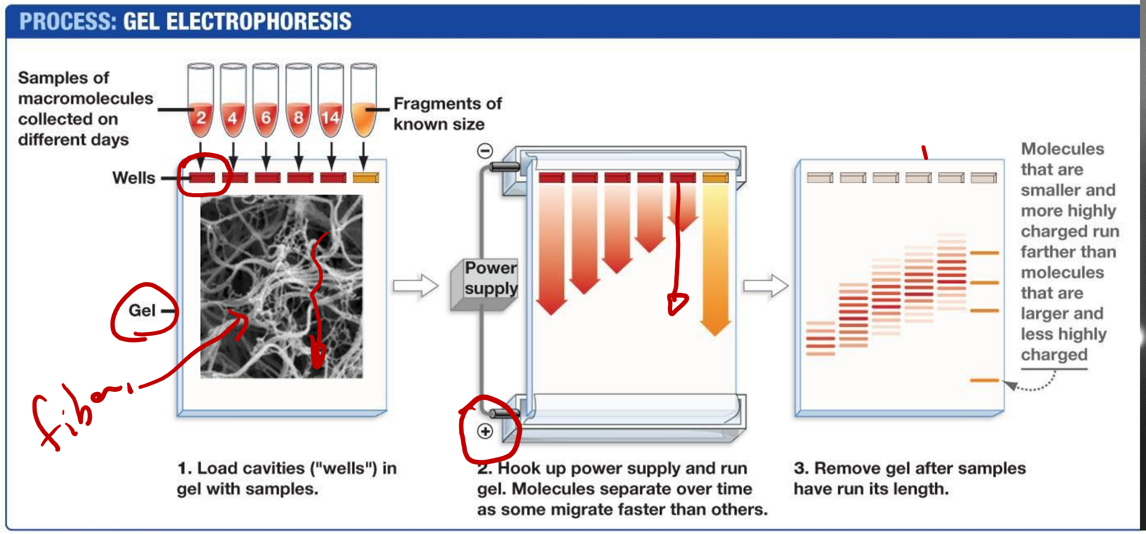
5' T C 3'
3' A G 5'

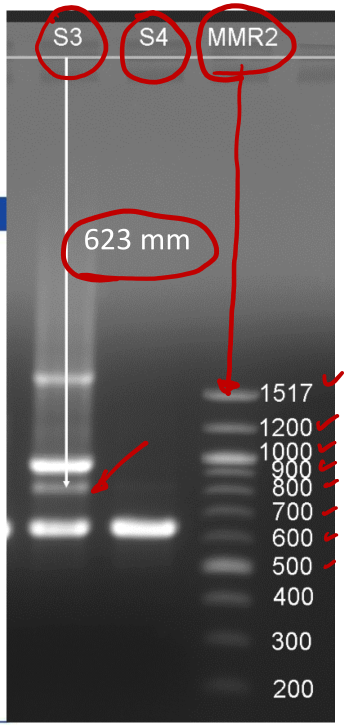
5 → 3
3' ← 5'

Size Determination of DNA - Agarose Gel Electrophoresis.

DNA has a neg charge on each phosphate
Separation is by size as the DNA strands are forced through the gel.



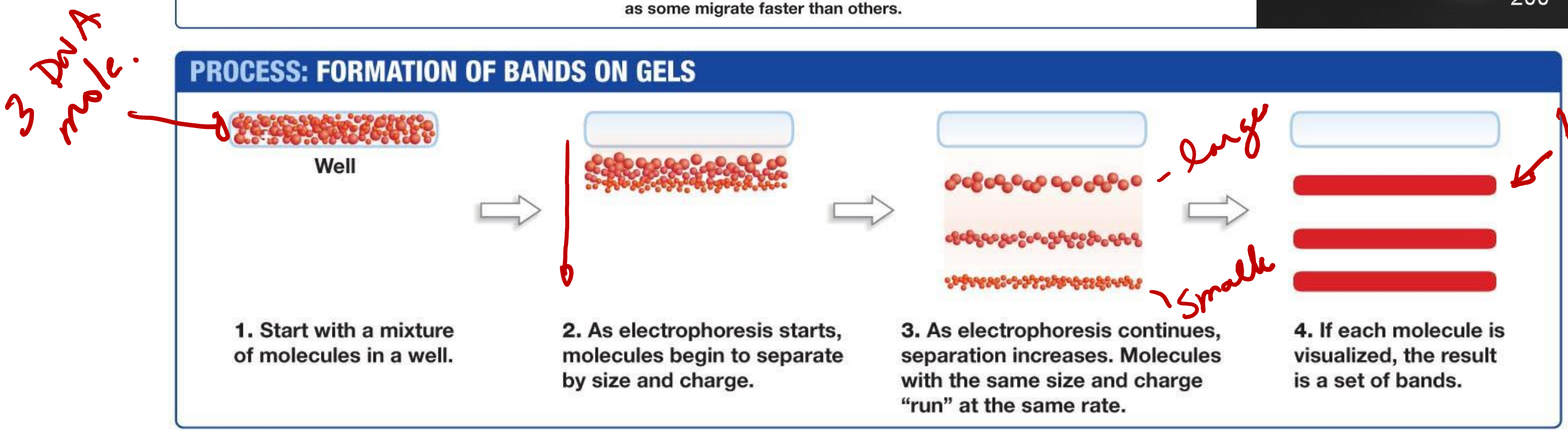


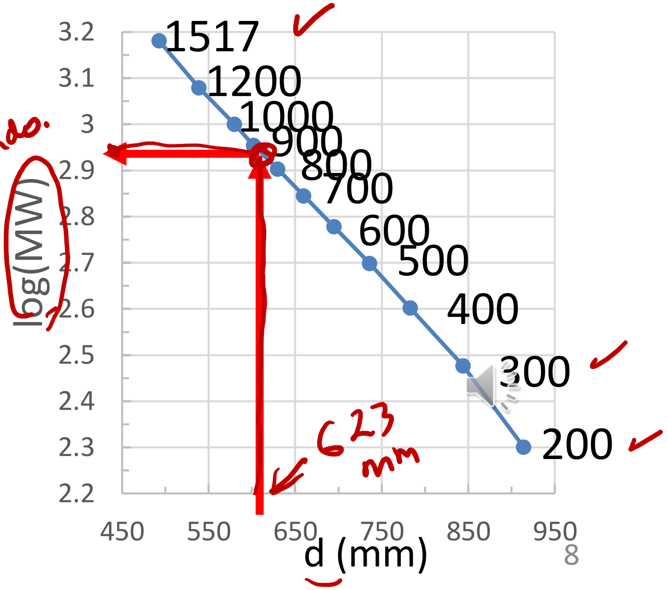


S3, S4 = unknowns.
MMR2 = standards, DNA fragments of known length.

Obtaining MW (length) of DNA fragments.

1. Plot log(basepair) versus distance for standards.
2. Obtain equation of curve.
3. Use distance for unknown to find its log(MW) (red arrows)
4. S3 3rd fragment migrated 623 mm, Log(MW)~2.92, MW ~820 bp

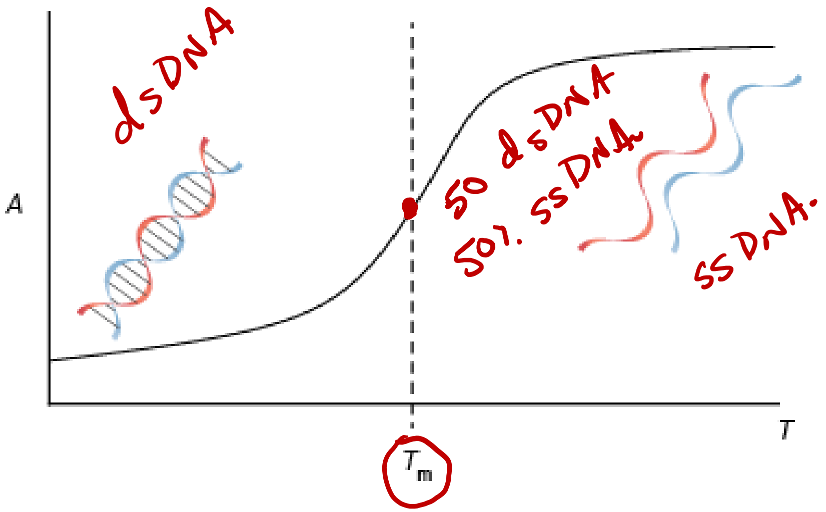




https://dnalc.cshl.edu/resources/animations/gelelectrophoresis.html

Thermal Stability of Double Stranded DNA (dsDNA)

5’ A-T-C-A
3’A-T-G-C-C-G-T-A-G-T-C-G-T-A-C-A-G-T-A-C-G-T-G-C-A



dsDNA

5’ A-T-C-A
3’A-T-G-C-C-G-T-A-G-T-C-G-T-A-C-A-G-T-A-C-G-T-G-C-A

1. How does the DNA sequence affect T_m? Explain the difference in Tm for the following two DNAs:

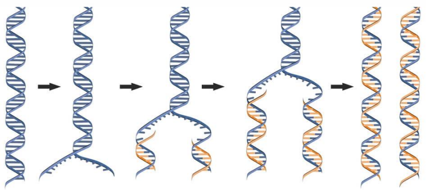
Duplex A: Tm=16 C

AAAATTTT
TTTTAAAA

Duplex B: Tm=32 C

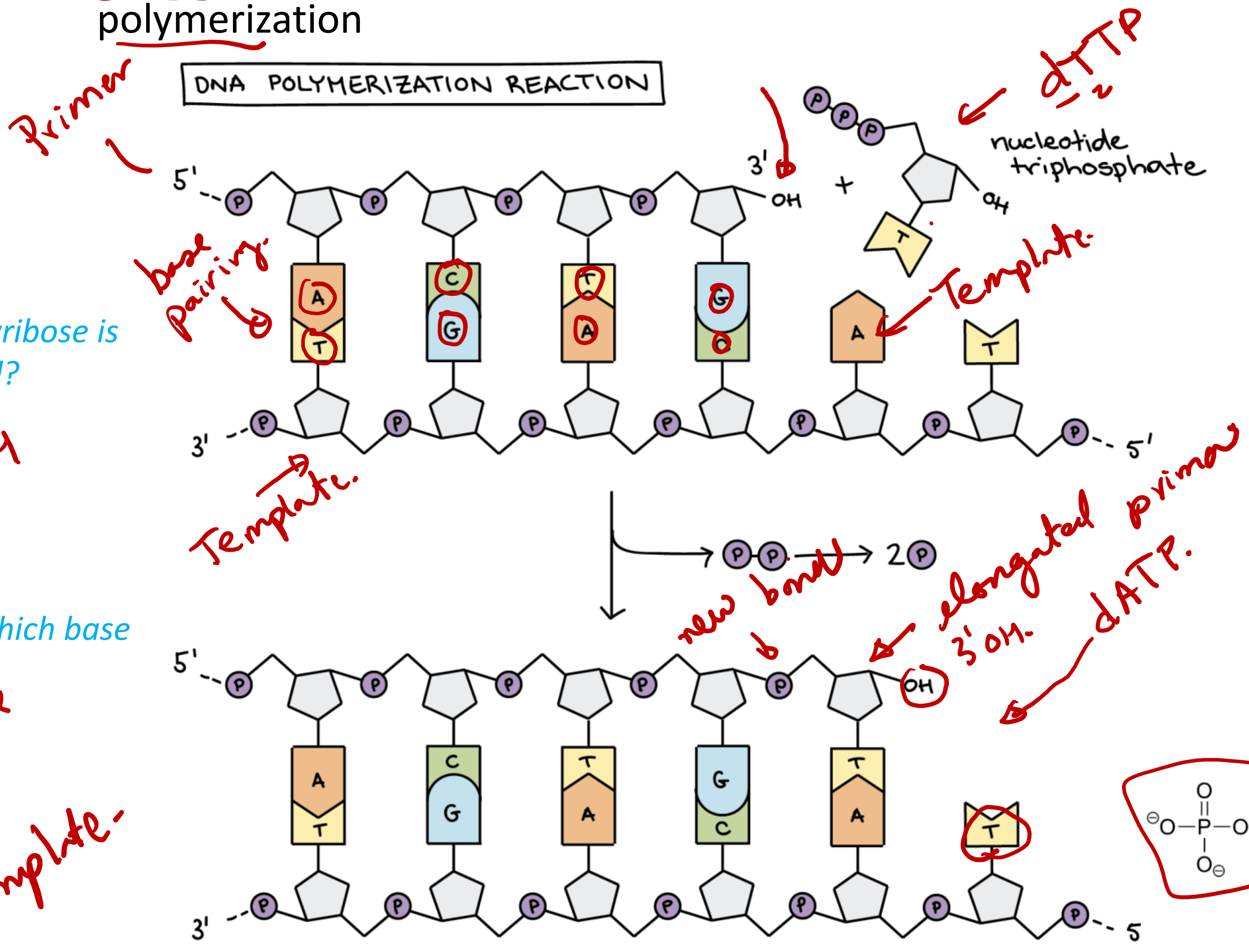
GCGCGCGC
CGCGCGCG

more H-bonds ∴ more heat to convert dsDNA to ssDNA.



DNA Polymerase – Fundamental Activity.

5' to 3' polymerization

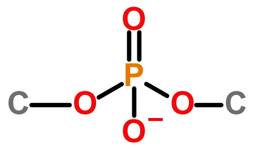


1. *Where on the deoxyribose is the new base added?*

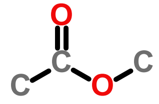
-3' -OH

2. *What determines which base is to be added?*

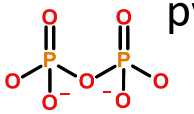
Comp base in the template.



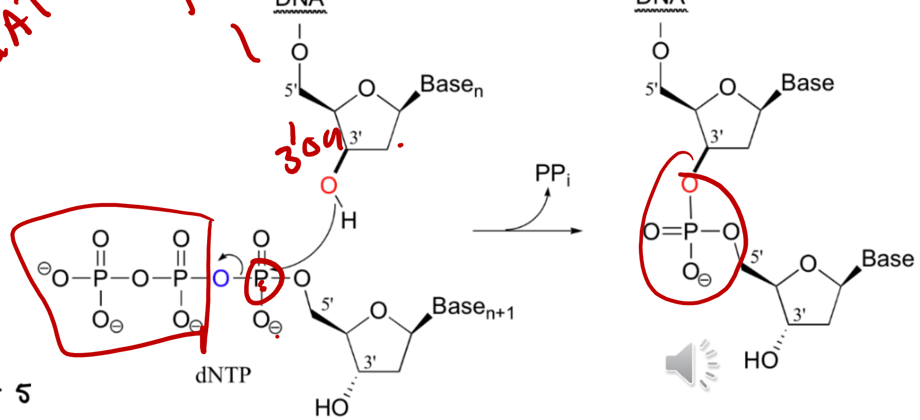
Phosphodiester linkage



ester linkage

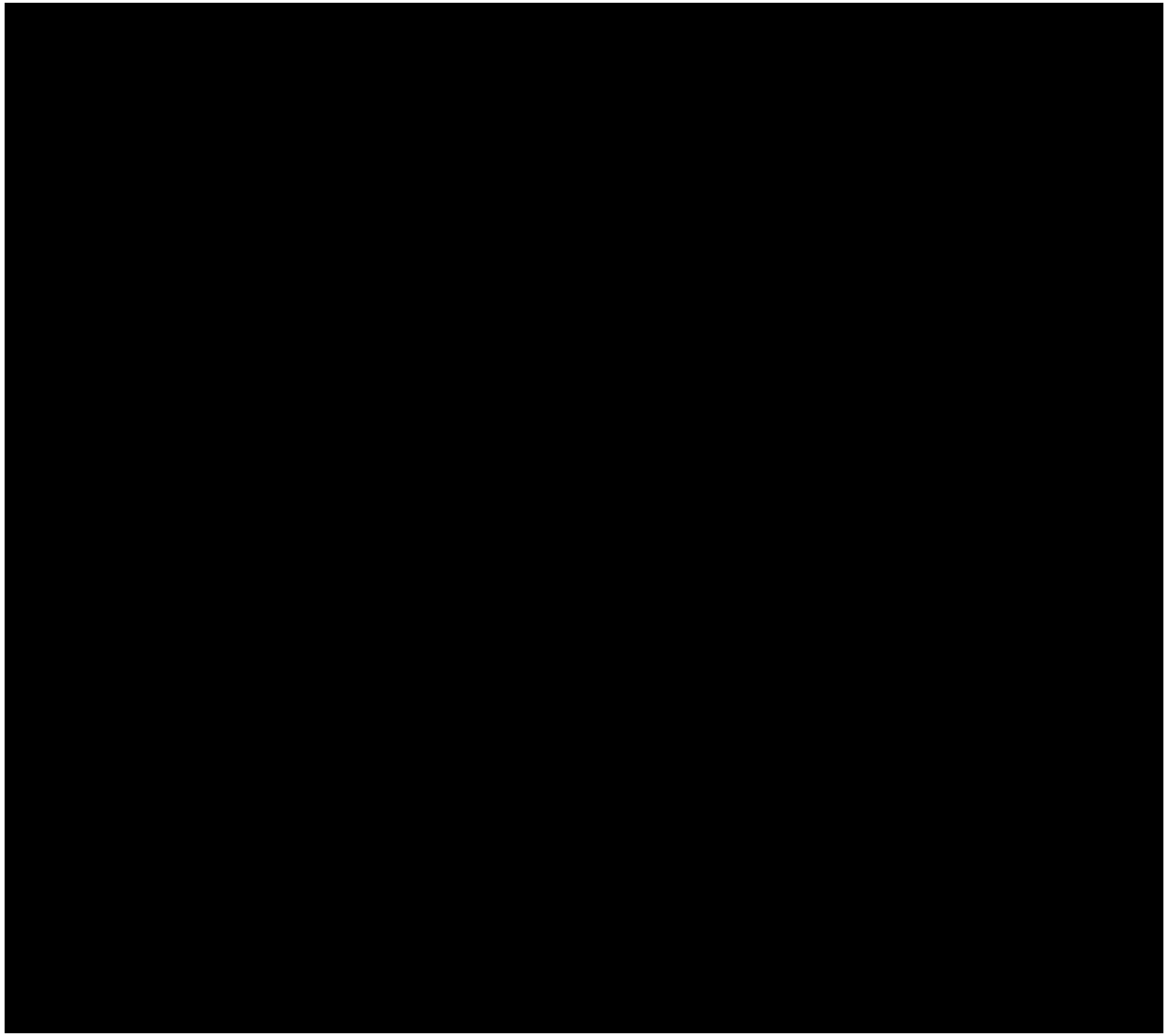


pyrophosphate



DNA Polymerase – Fundamental Activity.

- Synthesize new polymers of DNA.
- Require a short region of double stranded DNA to start synthesis – primer-template junction.
  - Primer can be a short DNA or RNA oligonucleotide (oligo) that is complementary to the DNA template.
  - RNA primers are used in DNA replication in the cell
  - DNA primers are used in other biotechnology applications (PCR, DNA Sequencing)
- Require single stranded template to provide information on which base to add.
- Add new dNTPs to 3’-OH of the primer, elongating in the **5’ to 3’** direction.
- Elongation will go to the end of the template.



(1:48 syn starts)

DNA Polymerase – Fundamental Activity.

Primer

5′ A-T-C-A

dGTP

Template

5′ A-T-C-A-G 3′ ... T

3′A-T-G-C-C-G-T-A-G-T-C-G-T-A-C-A-G-T-A-C-G-T-G-C-A

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5

1 (under position 10), 2 (under position 20)

1. *Where (what position) will this primer (ATCA) anneal?*
2. *What base will be added first?* ✓
3. *What is the last base added?*

DNA Polymerase – Error Correction – 3’ Exonuclease

- Incorrectly incorporated bases are removed by a 3’ exonuclease activity.
- Most DNA polymerases have this activity.
- The polymerase used by the HIV virus has no proofreading activity
- The polymerase used by Covid-19 has limited proofreading activity.

Reflection: What are the consequences of poor error correction in HIV and Covid viruses?

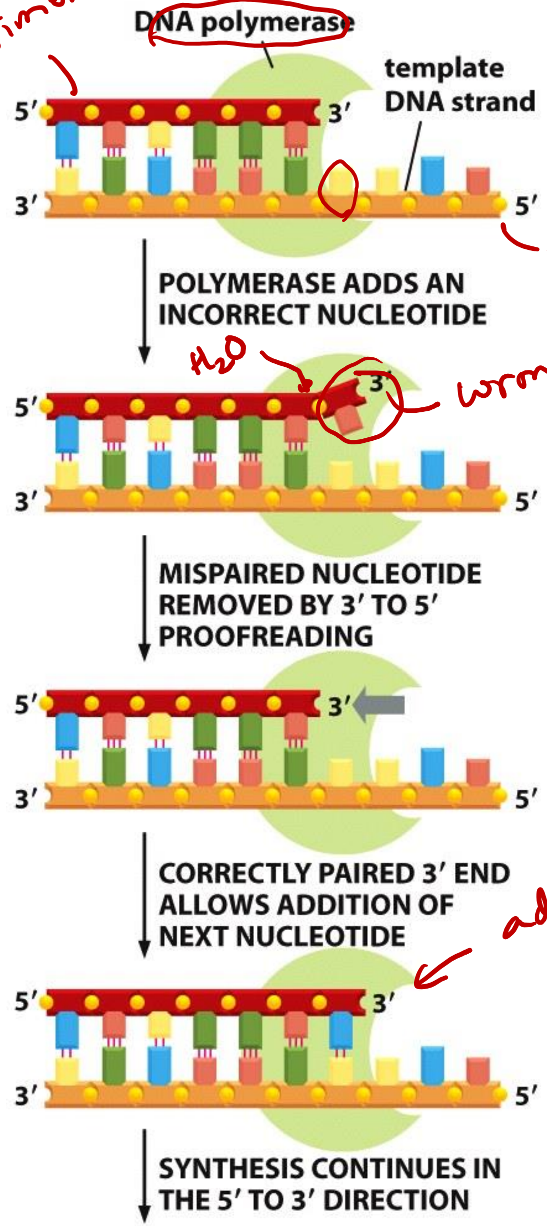
→ changes in genetic seq

→ mutation in virus

Polymerase Expectations:

1. Identify where primer anneals to the template.
2. Predict order of base addition.
3. Explain the mechanism of dNTP addition by polymerases (addition of dNTP to 3’OH, release of P-P)
4. Explain how polymerases correct errors (3’ -> 5’ exonuclease)

Primer.



Template

H_2O

wrong base

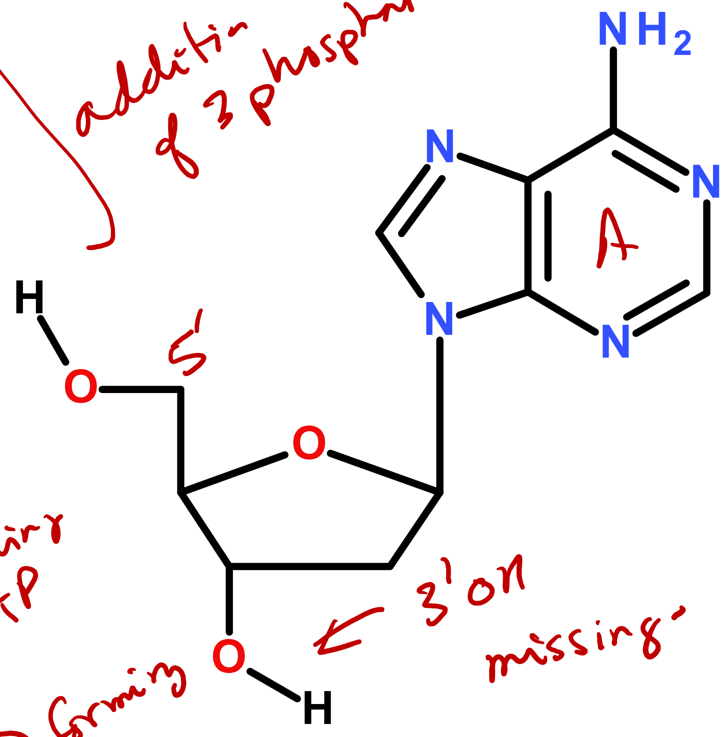
add correct base.

Exercise

A cell acquires deoxyadenosine from the environment.

A. Indicate the (three) steps that have to occur before this base can get incorporated into DNA

addition of 3 phosphates

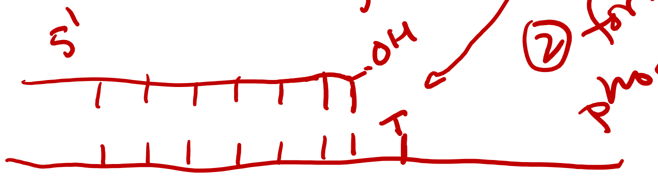


3' OH missing.

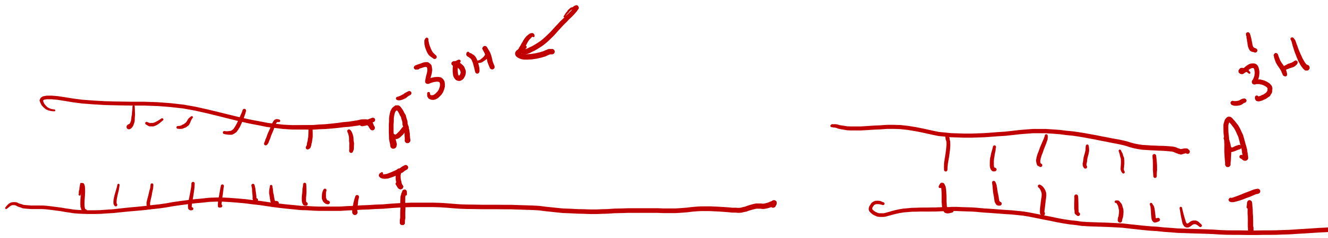
B. Indicate the ***two*** steps that will result in this base becoming part of a DNA strand, after the events in part A.

① binding dATP

② forming phosphodiester bond

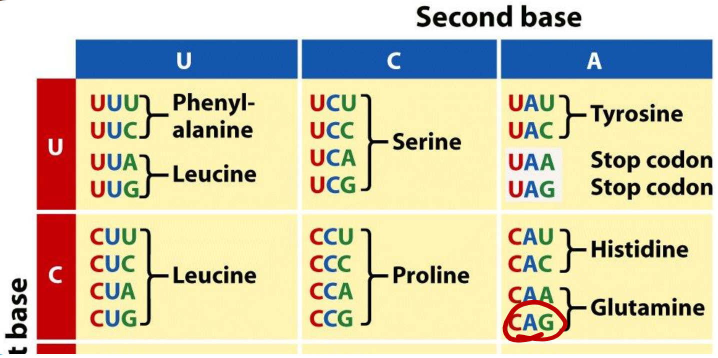


C. What will happen to DNA synthesis if the base is missing the 3'-OH?



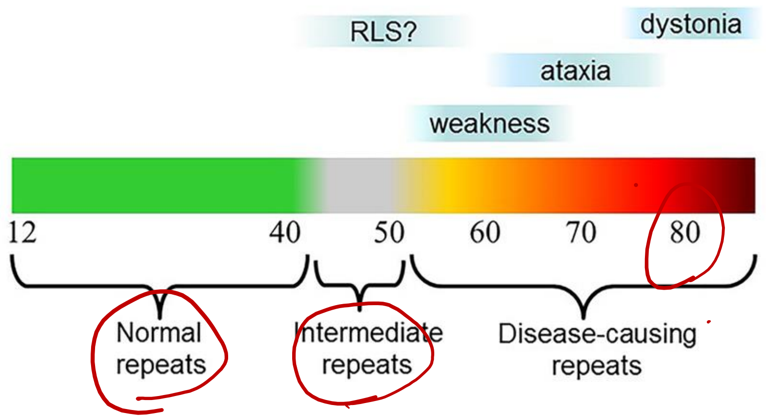
No addition because a 3' OH is required

Repeat Expansions Related to Diseases



- CAG – at least 10 diseases (Huntington disease, spinal and bulbar muscular atrophy, dentatorubral-pallidoluysian atrophy and seven SCAs)
- CGG – fragile X, fragile X tremor ataxia syndrome, other fragile sites (GCC, CCG)
- CTG – myotonic dystrophy type 1, Huntington disease-like 2, spinocerebellar ataxia type 8, Fuchs corneal dystrophy
- GAA – Friedreich ataxia
- GCC – FRAXE mental retardation
- GCG – oculopharyngeal muscular dystrophy
- CCTG – myotonic dystrophy type 1

CAGCAGCAG...
↓
Glu



- A small number of repeats is "normal"
- If the number of repeats increases the individual may show disease symptoms.
- These repeats can grow during replication in the cell due to "primer slippage".
- Longer repeats can be inherited from the parent
- Longer repeats can also occur within cells of the individual

Why do they cause disease?

- Additional amino acids if in protein coding region.
- Affect binding of DNA regulatory proteins if outside the coding region.

Treatment:

None yet, except genetic counseling.