Chapter 2: Nucleic Acids

Chapter Organization

1) Structure and Forces: Stabilizing DNA & RNA
2) Biochemistry/Molecular Biology of Nucleic Acids
3) Manipulation – Molecular Bio Techniques

Learning Objectives

- *Analyze the structures of nucleic acids at the chemical level.*
- *Illustrate when and how nucleic acids function in replication of DNA, transcription of DNA into RNA, regulation of transcription, and translation of RNA into proteins.*
- *Describe how alterations to nucleic acids in the cell can facilitate biochemical studies.*

DNA & RNA – Chemical Structure and Nature

More than just ATGC (plus some RNA stuff)…

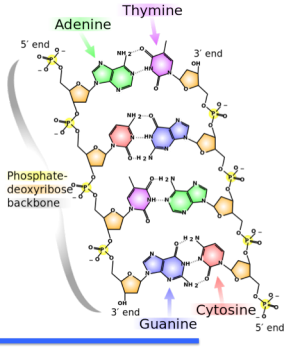
Points to consider:

- What is the chemical nature of the bases?
- What are the forces that create and stabilize the shapes of DNA and RNA

DNA & RNA Basics

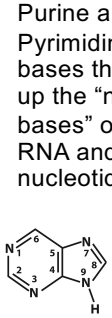
DNA – Right handed helix made of opposite oriented strands of nucleotides

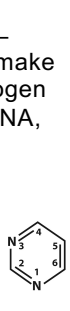
RNA – single strands of nucleotides, often forming helix within the same strand and bound to proteins (diverse set of molecules)

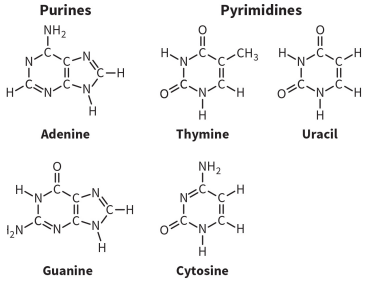


Bases

Purine and Pyrimidine – bases that make up the "nitrogen bases" of DNA, RNA and nucleotides





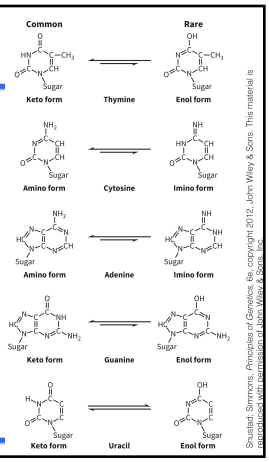


Bases

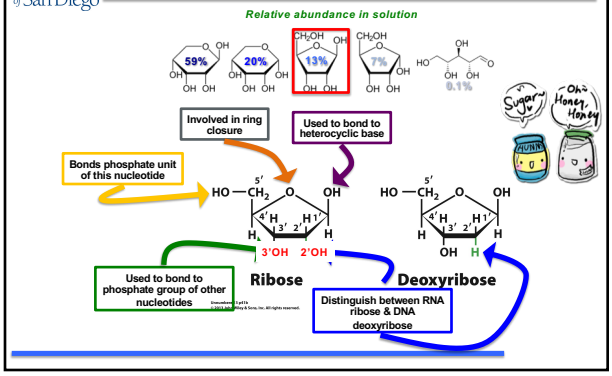
Purine and Pyrimidine – bases that make up the "nitrogen bases" of DNA, RNA and nucleotides

Exist as tautomers

- *Facilitated by resonance of the ring*
- *Enol/Keto changes possible IMF-bonding*
- *The R-OH act as catalytic centers that are not found in keto forms*



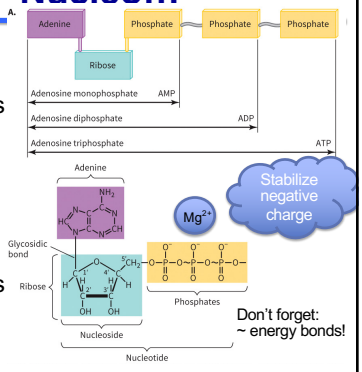
Sugar (ribose) backbone



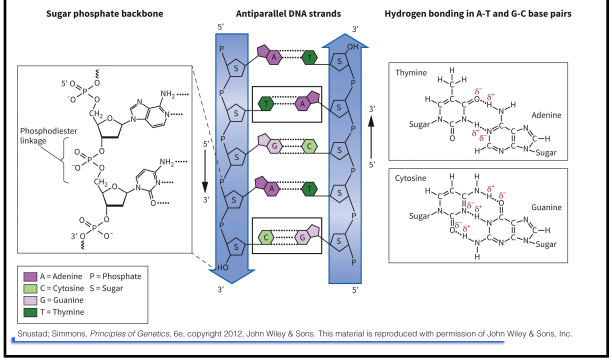
Nucleo…

Nucleotides are biomolecules that contain a nitrogenous base, a ribose sugar, and a phosphate group.

Nucleosides are biomolecules that contain a nitrogenous base and a ribose sugar.



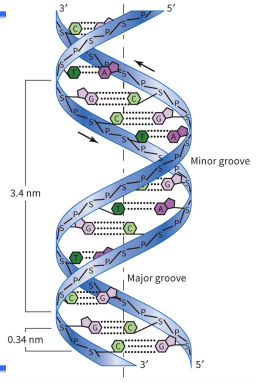
Polymeric Nucleotides



Basic DNA Structure

DNA Structure

- Antiparallel
- Right or left handed around common axis
- Sugar-phosphate backbone exposed to solvent
- Bases occupy hydrophobic core
- Minor and Major grooves (where protein aa side chains will bind

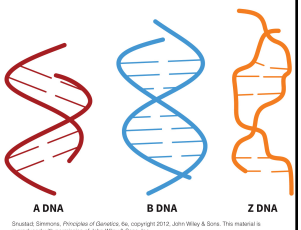


Three forms of DNA

B DNA – *First characterized and most "native" of commonly observed in vivo*

A DNA – *A dehydrated form found in ethanol ppt*

Z DNA – *Not common but may form when binding to protein – left handed and "twisted"*



Many Forces Create the Double Helix

Formation vs Stabilization – recognize the difference(s)

- H bonds between A-T and G-C do not "glue" the strands together. They form the specificity and impact the hybridization
- H bonds between bases are skewed/bent and less strong than linear H bonds
- No H bonds with water – entropy driving
- Helix forms because of hydrophobic effect caused by the mostly non-polar base stacking

P-Backbone: aligns/pulls to opposite sides – this is balanced by H bonding water and ions

Charge-charge repulsion (coulombic interactions) between phosphates in the backbone repels the two strands away from each other.

H bonding is required for specificity of base paring but contributes little to stability of DNA structure…

The alignment of base pairs on top of one another, termed base stacking, is a result of a combination of dipole-dipole, π–π stacking, and London-dispersion forces.

Base pairs exhibit the hydrophobic effect: more water molecules are "free" from having to cage hydrophobic parts of bases.

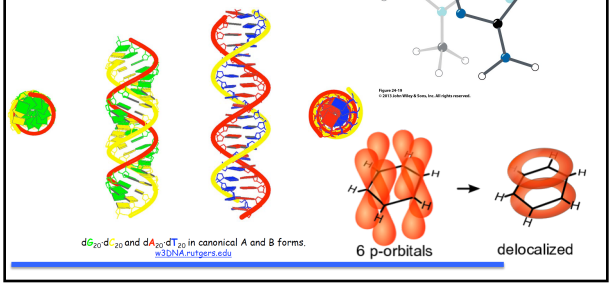
Phosphates and carbohydrate moieties interact with water through hydrogen bonding, ion-dipole, and dipole-dipole interactions.

Hydrogen bonding between base pairs and steric interactions of bases directs specificity of base pairing.

Stabilization of double helix

The purine and pyrimidine structure and the delocalized p orbitals stacked one over the other.

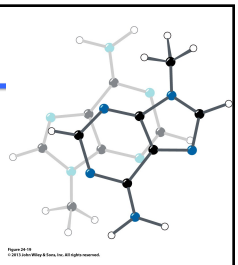
Overlap of stack 9-methyladenine

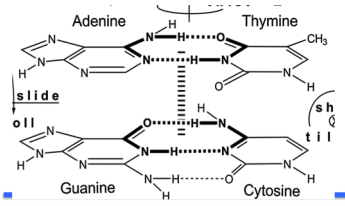


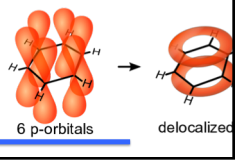
$d(G)_{20}\cdot d(C)_{20}$ and $d(A)_{20}\cdot d(T)_{20}$ in canonical A and B forms.
w3DNA.rutgers.edu

Stacking!!!!

Think of the purine and pyrimidine structure and the delocalized P orbitals stacked one over the other.



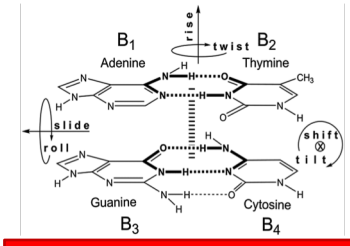




Stabilization of double helix

The stacking interactions of the atoms also form van der Waals interactions

The sequence influences the strength.



Base stacking is enthalpically favored, entropically opposed.

TABLE 24-2 Stacking Energies for the Ten Possible Dimers in B-DNA

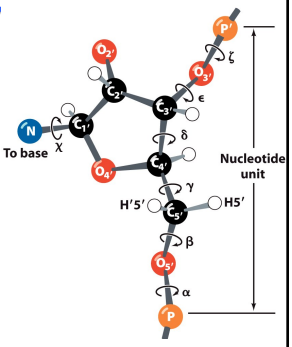
Stacked Dimer	Stacking Energy (kJ · mol⁻¹)
C · G G · C	−61.0
C · G A · T	−44.0
C · G T · A	−41.0
G · C C · G	−40.5
G · C G · C	−34.6
G · C A · T	−28.4
T · A A · T	−27.5
G · C T · A	−27.5
A · T A · T	−22.5
A · T T · A	−16.0

Source: Ornstein, R.L., Rein, R., Breen, D.L., and MacElroy, R.D., *Biopolymers* **17**, 2356 (1978).

Table 24-2

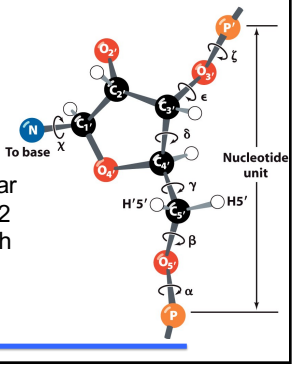
Flexibility & Torsional Angles

- DNA forms transition from one form to another based on conditions (water, ions or binding proteins)
- 7 torsional angles of rotational freedom
  - restricted due to internal constraints
- α and ζ are restricted due to interactions between ribose and P groups.
- Two conformations χ syn and anti conformation
  - Most are in anti except Z DNA or Hoogsteen base pairing …

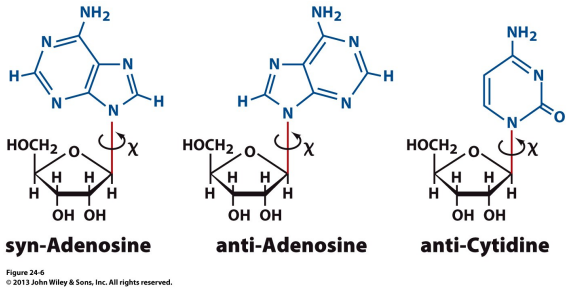


DNA has limited flexibility

- Rotation within the ribose is limited to accommodate phosphate repulsion and substituents of ribose atoms
- The ring can not remain planar due to crowding and either C2 or C3 will be out of plane (with C5)



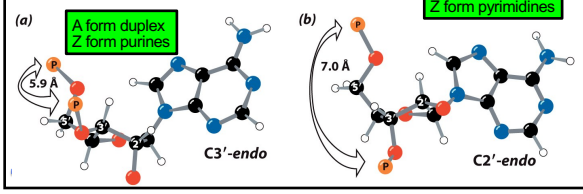
Sterically Allowed Base Orientations Anti or Syn



C2' or C3' Endopucker

Rotation within the ribose is limited to accommodate phosphate repulsion and substituents of ribose atoms

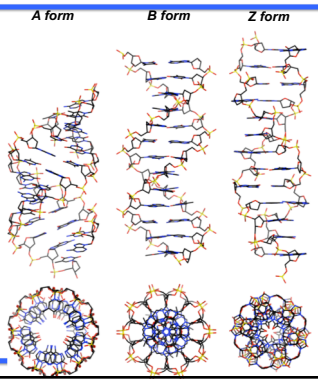
- Endopucker – the shape of the 2nd or 3rd carbon of Ribose – One will be slightly above plain with C5 carbon of ribose
- Results in a large shift in distance between phosphate groups – impacts the type of DNA forms



DNA Forms

A Form DNA-

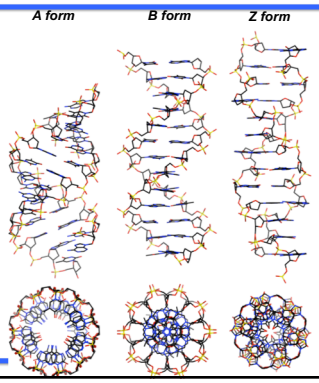
- Right handed, minor form, dehydrated
- Found in RNA-DNA interaction sites
- Wider and flatter helix than B DNA
- Major Groove narrow and deep, Minor Groove wide and shallow
- Bases are tilted 20o to axis
- Anti for both purines and pyrimadines
- C3' endopucker
- This is the double helical form of double stranded RNA



DNA Forms

B Form DNA-

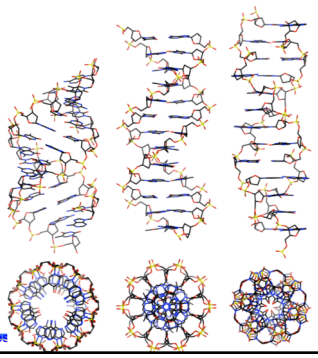
- Right handed, major form, hydrated
- Major Groove wide and deep, Minor Groove narrow and deep
- Bases are perpendicular to axis
- Anti for both purine and pyrimidine
- 11.6 bp/turn
- C2' endopucker



DNA Forms

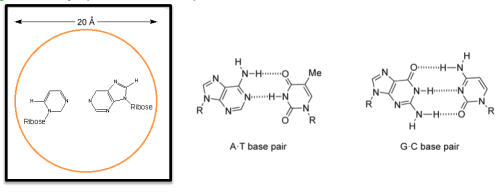
Z Form DNA-

- Left handed, minor form, dehydrated
- Evidence for Z DNA is binding proteins
- Found near start of genes – may alter protein binding to proteins for transcription
- Found with high alternating purine-pyrimidine sequences (dGC), during negative supercoiling and high salt or select cation binding (reducing electrostatic repulsion from phos groups.
- Not a stable helix
- Flat or missing major groove, narrow and deep minor groove
- Bases are nearly flat
- Alt anti – syn
- ~~Endopucker C2' Pyrimidines C3' Purines~~



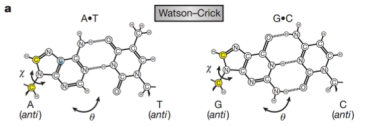
Why only AT and GC?

Base pairing constrained by duplex diameter & H-bond pattern

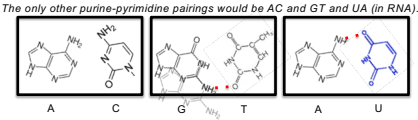


Space between backbone only permits a purine-pyrimidine pairing

Why only AT and GC?

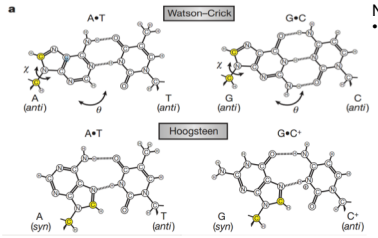


The only other purine-pyrimidine pairings would be AC and GT and UA (in RNA).



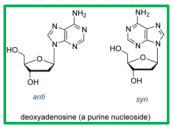
Pairings are mismatches because the pattern of hydrogen donors and acceptors do not correspond

Hoogsteen Base Pairing



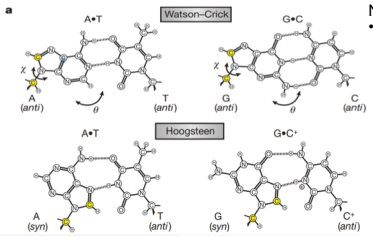
Non-WC are possible
- purine rotation around the glycosidic bond (χ) and base-flipping (θ), affecting simultaneously C8 and C1′ (yellow)

Conformation of base



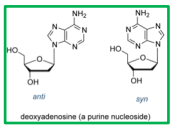
Normal: Watson-Crick pairing
H-bonds are 180°, short (strong)
Alternate H bonds: Hoogsteen pairing
Requires *syn* conformation
H-bonds are not 180°, GC reduced from 3 to 2 H bonds

Hoogsteen Base Pairing



Non-WC are possible
- purine rotation around the glycosidic bond (χ) and base-flipping (θ), affecting simultaneously C8 and C1′ (yellow)

Conformation of base



Hoogsteen BP is an alternative BP sometimes seen in transition between DNA forms with distorted backbones

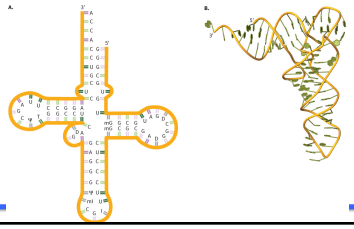
Normal: Watson-Crick pairing
H-bonds are 180°, short (strong)
Alternate H bonds: Hoogsteen pairing
Requires *syn* conformation
H-bonds are not 180°, GC reduced from 3 to 2 H bond

RNA: A Complex Molecule

Generally single stranded but can be double or triple stranded

Contains same weak forces used to stabilize DNA

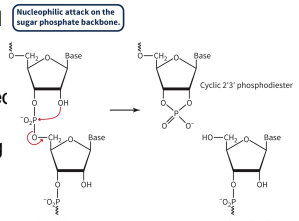
Provides information for protein synthesis

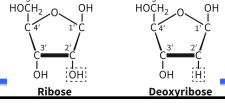


Complex RNA Structures

The extra OH (2' of the ribose) makes RNA less stable

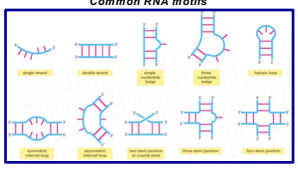
- In alkaline conditions, 2'OH can attack the P bond.
- ALSO many RNases (ribonucleases) are secreted in large amounts – attack requires the 2'OH – making the degradation specific to RNA not DNA



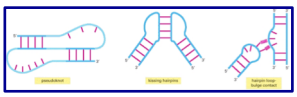


Nucleic acids – Other base paired structures

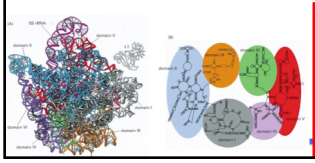
Common RNA motifs



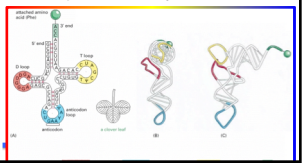
Tertiary & Quaternary structures of RNA



Ribosome: 23S subunit



Example of RNA Tertiary structure: tRNA

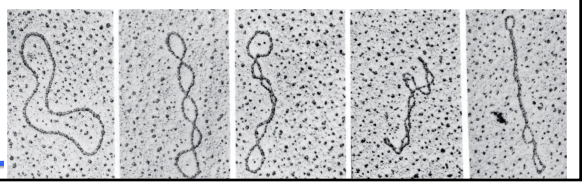


DNA can be Supercoiled

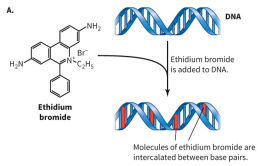
Consider a closed, circular double stranded DNA.

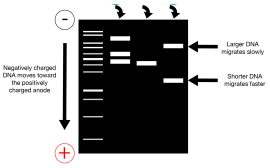
-twisting one strand (over or under) will cause kinks leading to a "supercoil"

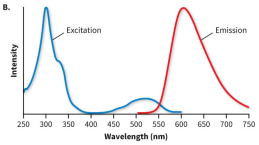
- The number of coils cannot be altered without first cleaving at least one or both of the strands



Monitoring DNA and RNA





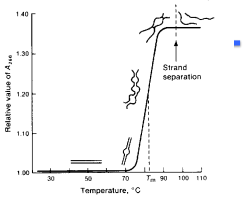


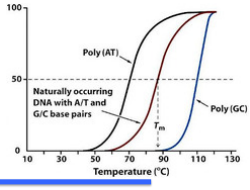
Shape of DNA can influence Electrophoretic Migration

- Large vs small
- Supercoiled vs relaxed vs open
- Protein bound to DNA/RNA

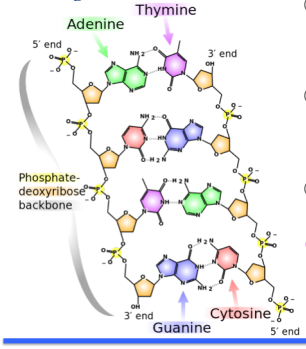
Thermal denaturation = melting

1) Denaturation occurs through disruption of hydrophobic stacking of bases and hydrogen bonding between bases
2) Melting temperature (T_m) defined as the temperature at which half of the DNA strands are in the random coil or single-stranded (ssDNA) state.
3) T_m depends on the length and sequence.
   a) ↑GC content = higher T_m
   b) Mismatches = ↓T_m
5) Hybridization = process of establishing a non-covalent, sequence-specific interaction between two or more complementary strands of nucleic acids into a single complex
6) Annealing = pairing of complementary sequences via hydrogen bonding of bases – usual refers to short sequences, ie primers, used in *in vitro* rxns.





Nucleic acids – implications of base pairing



① The regular structure and data redundancy provided by the DNA double helix makes DNA well suited to the storage of genetic information.

② The base-pairing between DNA and incoming nucleotides provides the mechanism through which DNA polymerase replicates DNA.

③ The base-pairing of RNA to DNA allows RNA polymerase to synthesize mRNA.

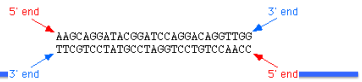
④ DNA-binding proteins can recognize specific base pairing patterns that identify particular regulatory regions of genes.

Nucleic acids – DNA sequence conventions

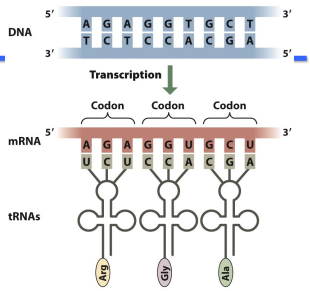
1) In a DNA duplex, strands are anti-parallel.
2) Individual strands are synthesized from the 5' OH of the ribose to the free 3'OH at the terminus of the strand. Thus, DNA sequences have directionality going from 5' to 3'.
3) The anti-parallel strands are said to be complimentary due to the hydrogen bonding occurring between heterocyclic bases across strands of duplex.
4) DNA sequences are written 5' to 3'
5) When a complimentary strand is given for the 5' to 3' strand, this is the reverse complement as it is the 3' strand sequence rewritten from 5' to 3'

If sequence 5'-TGC-3', then what is the complement?
a) 3'-TGC-5'
b) 3'-ACG-5'
c) 3'-GCA-5'
d) 3'-CGT-5'

If sequence 5'-ATCCG-3', then what is the complement?
a) 5'-GCCAT-3'
b) 5'-GCCTA-3'
c) 5'-TAGGC-3'
d) 5'-CGGAT-3'



Replication, Transcription and Translation



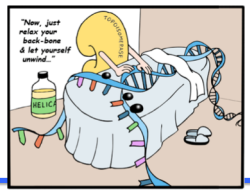
DNA Replication

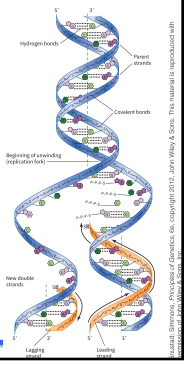
When/Why is DNA replicated?

How does DNA duplex separate?

How is synthesis of new DNA initiated?

The strands are running opposite directions. How does this affect synthesis of the two strands?

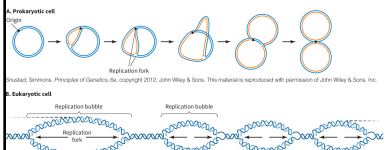


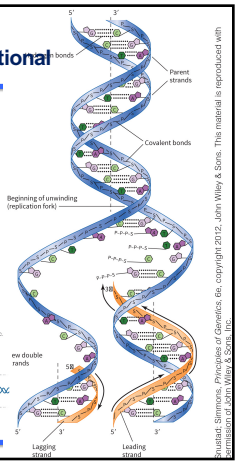


DNA Replication Is Bidirectional

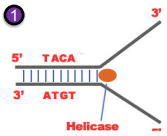
A replication bubble forms where two strands of DNA unwind

Two replication forks are formed
- Leading strand
- Lagging strand

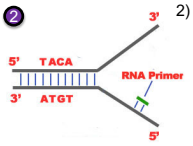




DNA replication – Steps in the process

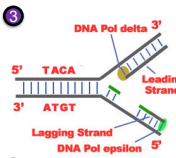


1) Breaking open the DNA duplex
   a) A-T rich regions
   b) **DNA gyrase** (enzyme) nicks DNA strand
   c) **Helicase** (enzyme) splits two strands
   d) Each strand serves as a template (semi-conservative replication)
   e) Initiation point = "origin of replication"
   f) structure formed where DNA opens = "Replication Fork"

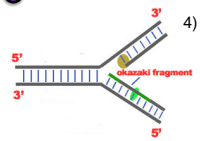


2) Priming the system
   a) **RNA Primase** (enzyme) binds in the initiation point of the 3'-5' parent chain.
   b) **RNA Primase** can attract RNA nucleotides which bind to the DNA nucleotides of the 3'-5' strand due to the hydrogen bonds between the bases.
   c) RNA nucleotides are the primers (starters) for the binding of DNA nucleotides.

DNA replication – Steps in the process

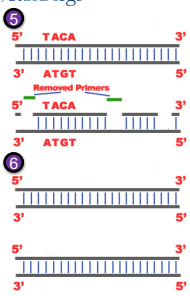


3) Elongation Process
   a) DNA polymerase III (DNA Pol III – enzyme) synthesizes complementary strand
   b) 5' - 3' template
      i. Leading stranding
      ii. Continuous addition of nucleotides
   c) 3' – 5' template
      i. Lagging strand
      ii. Additional RNA primers required
      iii. Replicated segments = Ozaki fragments



4) Sewing together the pieces
   a) DNA Pol I examines Ozaki fragments and removes RNA primers
   b) DNA Pol III fills gaps with complementary nte.
   c) DNA ligase (enzyme) connects fragments by completing phosphodiester linkages.

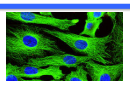
DNA replication – Steps in the process



5) Termination
   a) DNA polymerase III reaches end of strand
   b) Lagging strand cannot replicate to end of strand

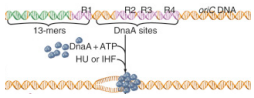
6) Repair
   a) Nucleases excise any errors in replicating DNA (mismatches)
   b) DNA Pol III fills gaps with complementary nte.

Regulation of DNA Replication



Bacteria/Prokaryotes

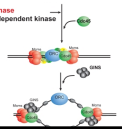
Signaled via energy levels –DnaA
- DnaA binds and is activated by high ATP then acts at the replication fork and unwinds DNA
- Followed by other binding proteins



Eukaryotes

Coordinated by cell cycle and cyclin dependent kinases (CDKs)
- Yeast have specific origin sites
- Mammalian cells are non-specific.

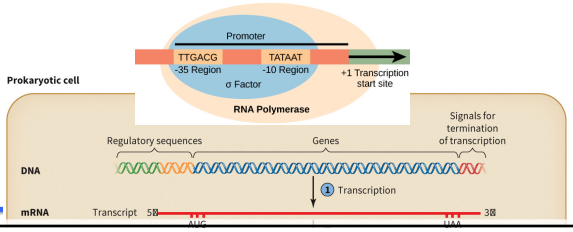
Several kinases initiate many replication sites – Replication Licensing: happens prior to replication.



Transcription: DNA-> RNA

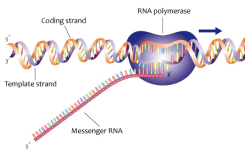
Prokaryotes - Genes organized into groups or operons. Proteins with a common purpose. i.e. metabolizing glucose

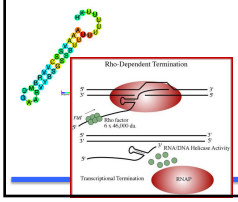
- Gene codes for all proteins without breaks and translated into one big mRNA
- Initiated at the promoter by sigma factors that recruits RNA polymerase



Elongation & Termination

Once RNA polymerase is initiated, elongation proceeds unwinding the DNA template and adding to the 3' end of the growing mRNA strand. Nucleosides provide the new bases and energy. Error is less problematic as the proteins coded will have a shorter half life than DNA and are not used to make new, mutated proteins.





Two mechanisms of termination

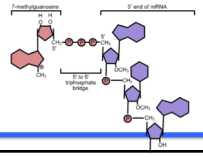
- Rho-independent/intrinsic termination: a looping of the DNA (palindrome) which interferes, stalls and falls the RNA Pol to fall off.
- Rho-dependent / ρ factor: A second protein binds to the RNA Pol which then binds to a series of G bp and stalls the Pol

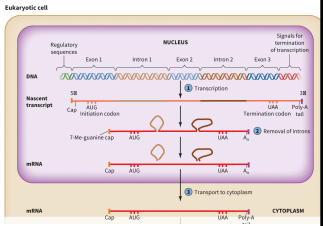
Transcription: DNA-> RNA

Eukaryotes – One Gene one mRNA (sort of… splice variant).

- Genes (DNA) are broken into coding (exon) and non-coding, intervening bp (intron).
- Nacent (new) mRNA is spliced and modified.

- 5' mRNA is capped with a methylated guanosine (5' cap)
- Unusual 5' to 5' triphosphate bond
- Capping blocks exonuclease activity – stabilizes mRNA lifetime



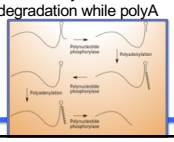


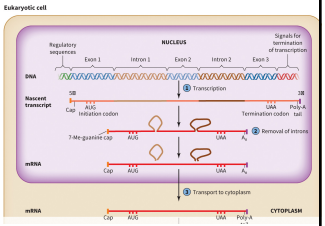
Transcription: DNA-> RNA

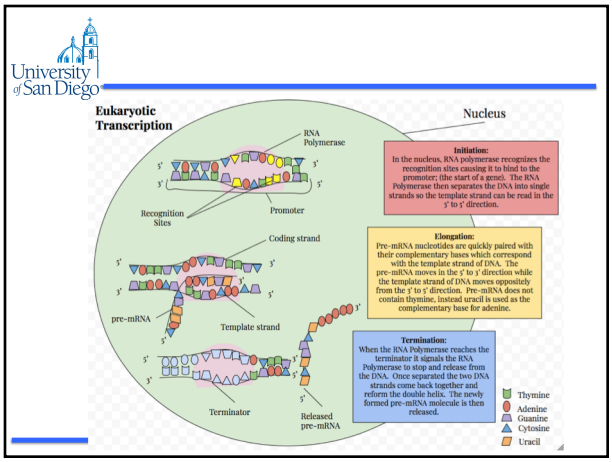
Eukaryotes – One Gene one mRNA (sort of… splice variant).

- Genes (DNA) are broken into coding (exon) and non-coding, intervening bp (intron).
- Nacent (new) mRNA is spliced and modified.

- 3' mRNA polyadenylated
- Last few 3'bp are removed and replaced by A_n bp sometimes at different sites
- Is used for nuclear export and translation – slowly removed but blocks degradation while polyA tailed

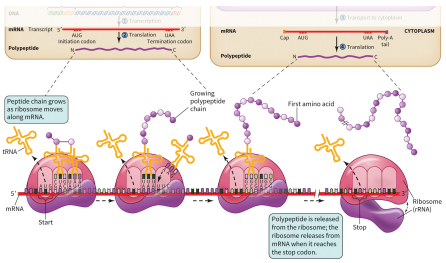




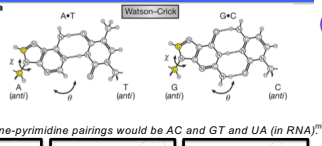


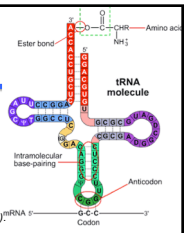
Translation

Synthesis of protein using mRNA, tRNA and rRNA

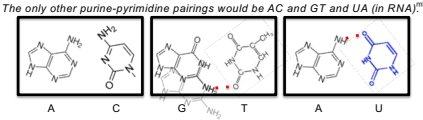


Base pairing & codons





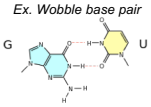
The only other purine-pyrimidine pairings would be AC and GT and UA (in RNA)



Pairings are mismatches because the pattern of hydrogen donors and acceptors do not correspond

Wobble hypothesis: *Crick postulated that the 5' base on the anticodon, which binds to the 3' base on the mRNA, was not as spatially confined as the other two bases, and could, thus, have non-standard base pairing*

Ex. Wobble base pair



A wobble base pair is a pairing between two nucleotides in RNA molecules. The thermodynamic stability of a wobble base pair is comparable to that of a Watson-Crick base pair. Wobble base pairs are fundamental in RNA secondary structure and translation.

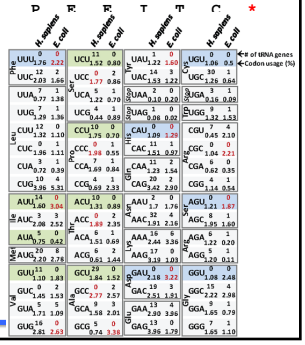
Start codon – sequence – Stop codon

ATG-AGT-GGG-CAT-AAA-CGT- XXX$_N$ -CCC-GAA-GAA-AUU-ACC-UGU-**TAA**

M S G H K R X P E E I T C *

Second Letter

1st letter	U	C	A	G	3rd letter
U	UUU Phe; UUC Phe; UUA Leu; UUG Leu	UCU Ser; UCC Ser; UCA Ser; UCG Ser	UAU Tyr; UAC Tyr; UAA Stop; UAG Stop	UGU Cys; UGC Cys; UGA Stop; UGG Trp	U C A G
C	CUU Leu; CUC Leu; CUA Leu; CUG Leu	CCU Pro; CCC Pro; CCA Pro; CCG Pro	CAU His; CAC His; CAA Gln; CAG Gln	CGU Arg; CGC Arg; CGA Arg; CGG Arg	U C A G
A	AUU Ile; AUC Ile; AUA Ile; AUG Met	ACU Thr; ACC Thr; ACA Thr; ACG Thr	AAU Asn; AAC Asn; AAA Lys; AAG Lys	AGU Ser; AGC Ser; AGA Arg; AGG Arg	U C A G
G	GUU Val; GUC Val; GUA Val; GUG Val	GCU Ala; GCC Ala; GCA Ala; GCG Ala	GAU Asp; GAC Asp; GAA Glu; GAG Glu	GGU Gly; GGC Gly; GGA Gly; GGG Gly	U C A G



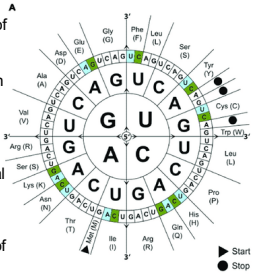
Codon Bias / Rare Codons

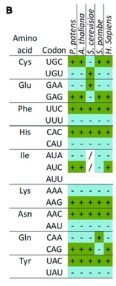
Codon bias is the frequency of occurrence of coding DNA/RNA.

tRNA populations for each wobble codon are not equal

Very important when expressing a non-bacterial gene in E. coli

Can lead to low expression or truncation of protein.





Retroviruses and Dogma