

Topic 7

Nucleotides and Nucleic Acids

Nelson & Cox (2016)

Chapter 8: p281-297

Chapter 24: p979-1003

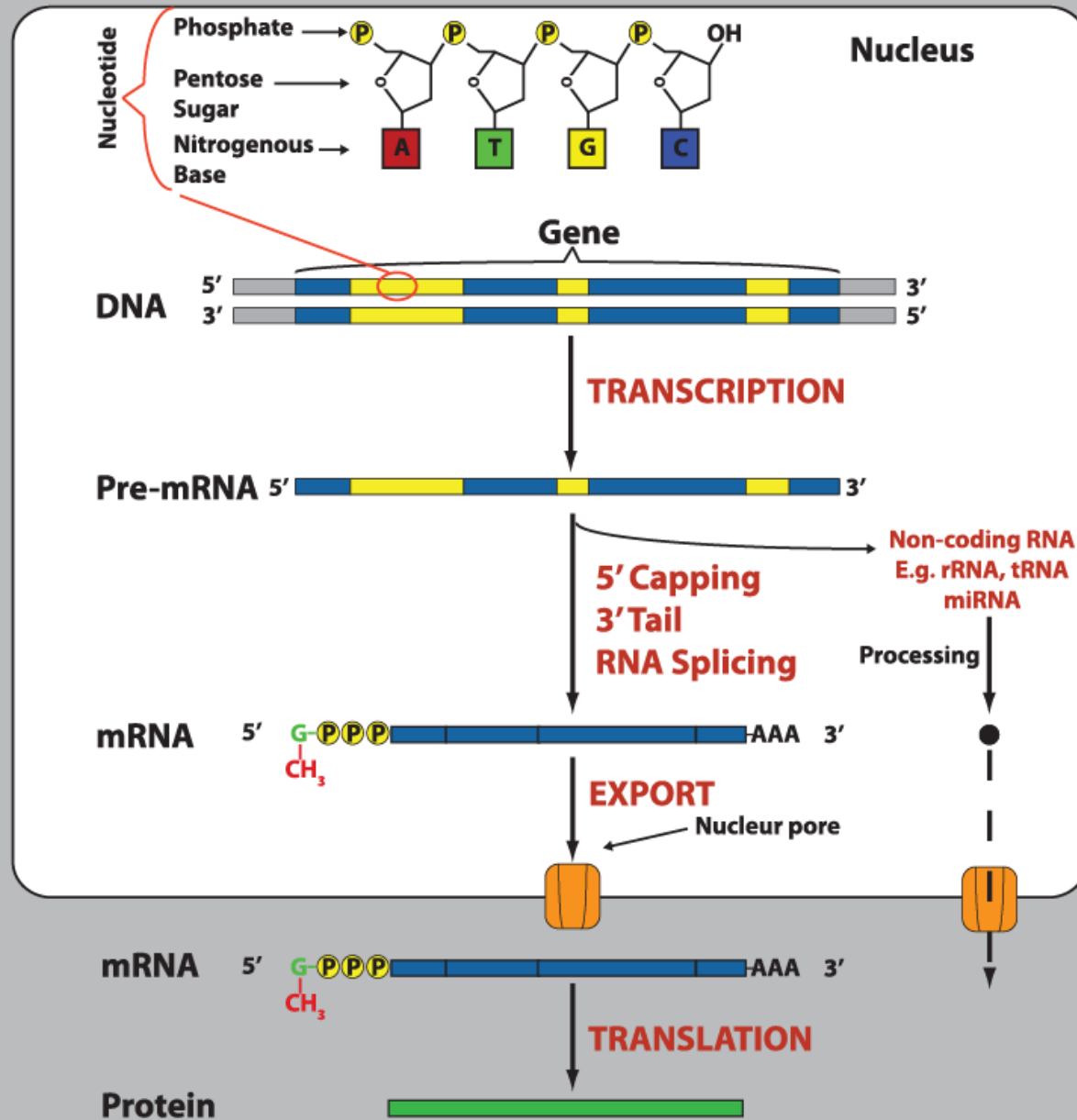
Nelson & Cox (2017)

Chapter 8: p279-295

Chapter 24: p957-982

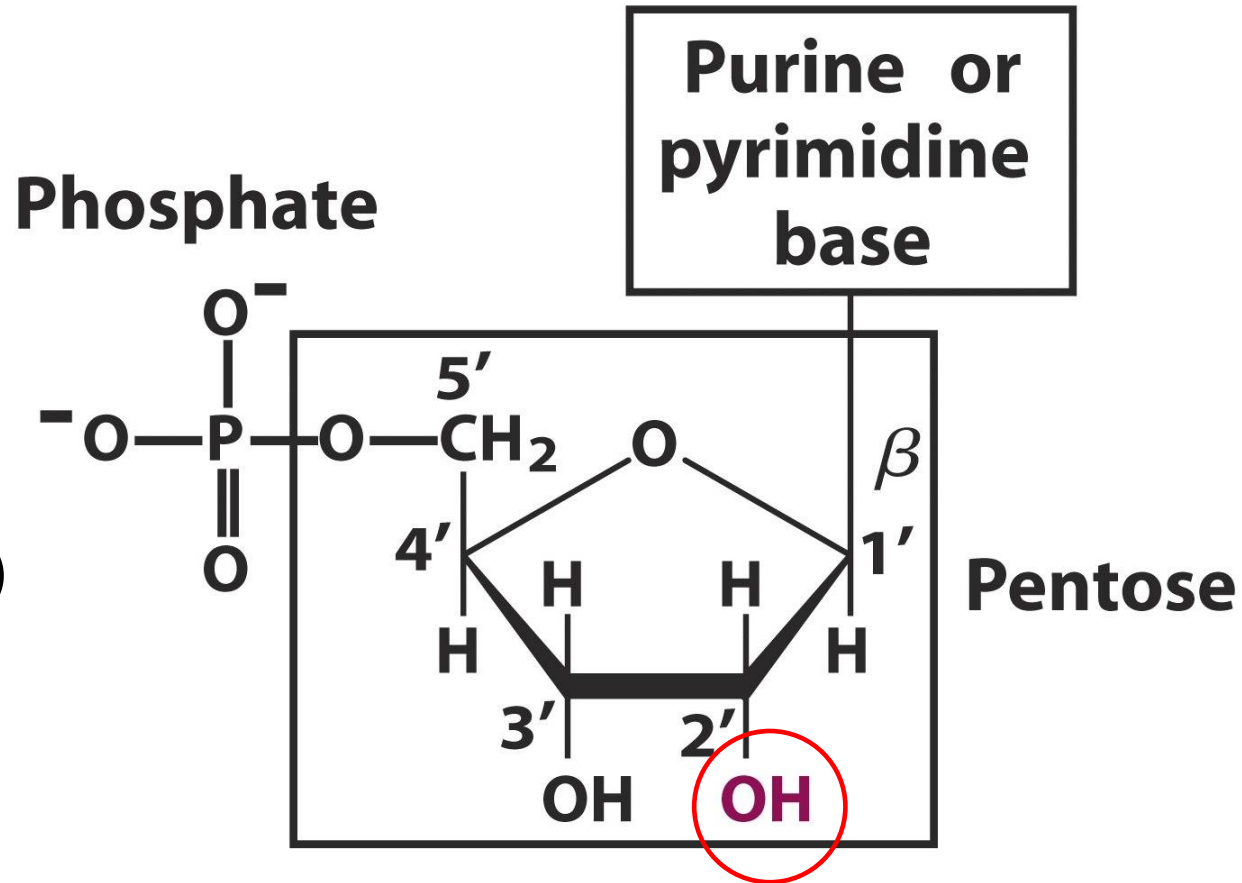
- Exons: expressed sequences - coding sequence
- Introns: intervening sequences - non-coding sequence

Cytoplasm

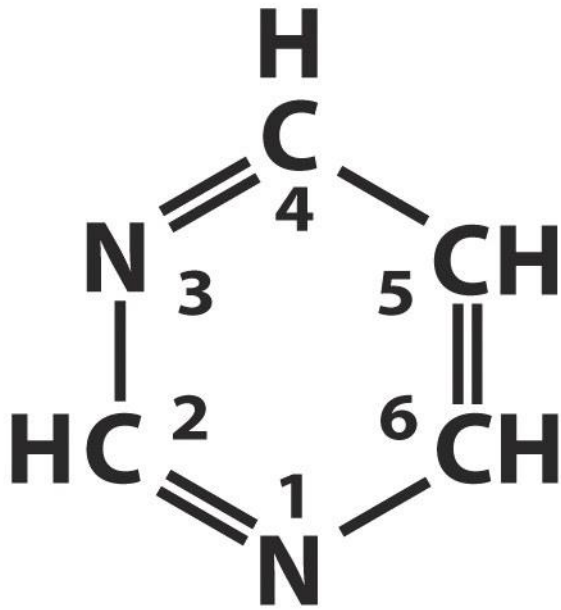


Nucleotides – Monomeric Units

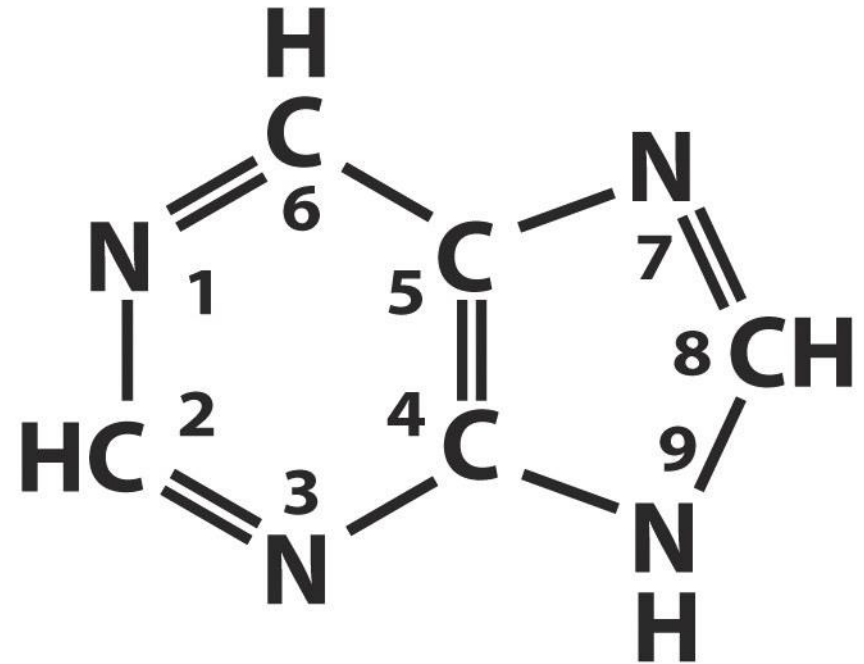
1. Phosphate Group
2. Pentose Sugar
 - Ribose: RNA
 - Deoxyribose: DNA
3. Nitrogenous Base
 - Pyrimidines (C, T & U)
 - Purines (G & A)



Nitrogenous Base

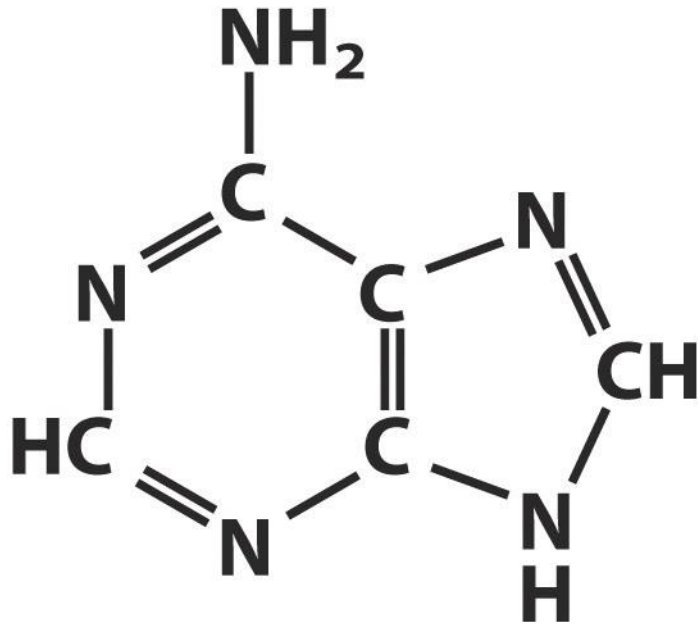


Pyrimidine



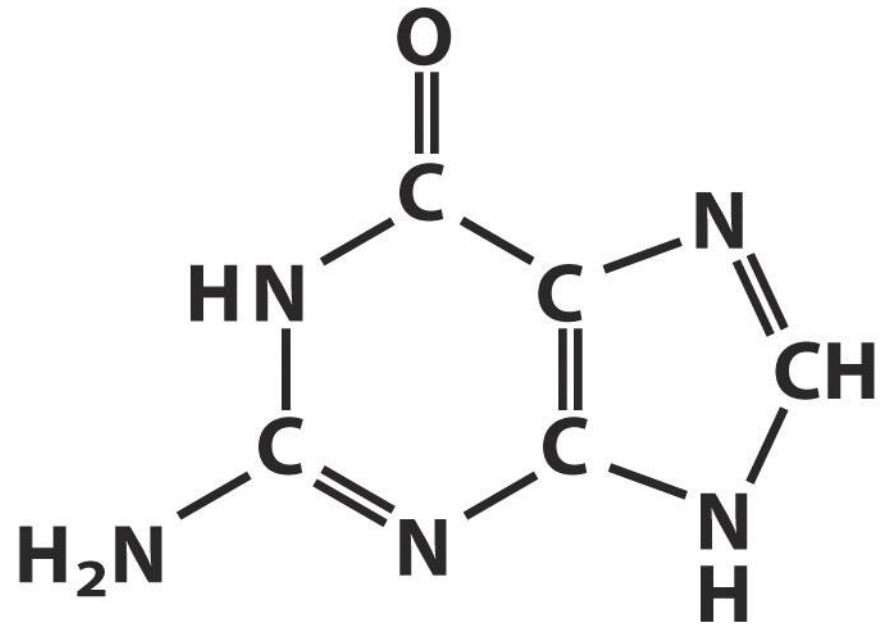
Purine

Nitrogenous Base - Purines



Adenine

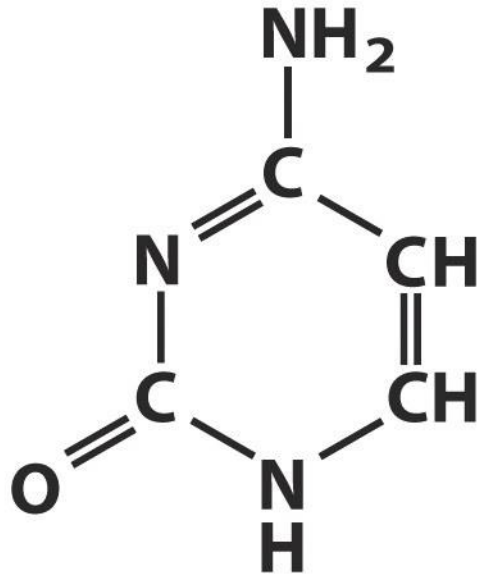
(A)



Guanine

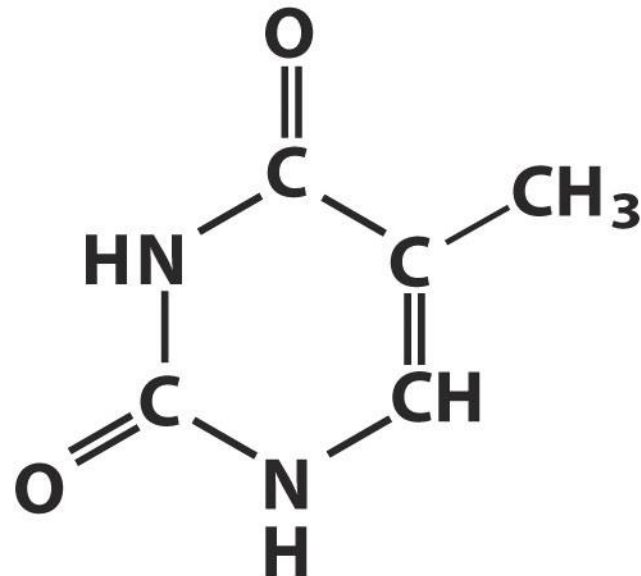
(G)

Nitrogenous Base - Pyrimidines



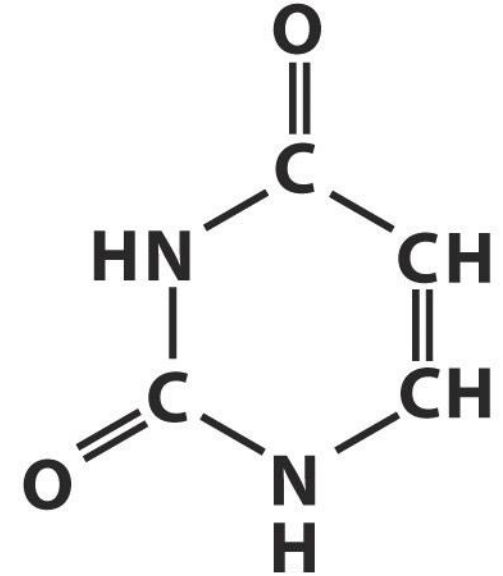
Cytosine

(C)



**Thymine
(DNA)**

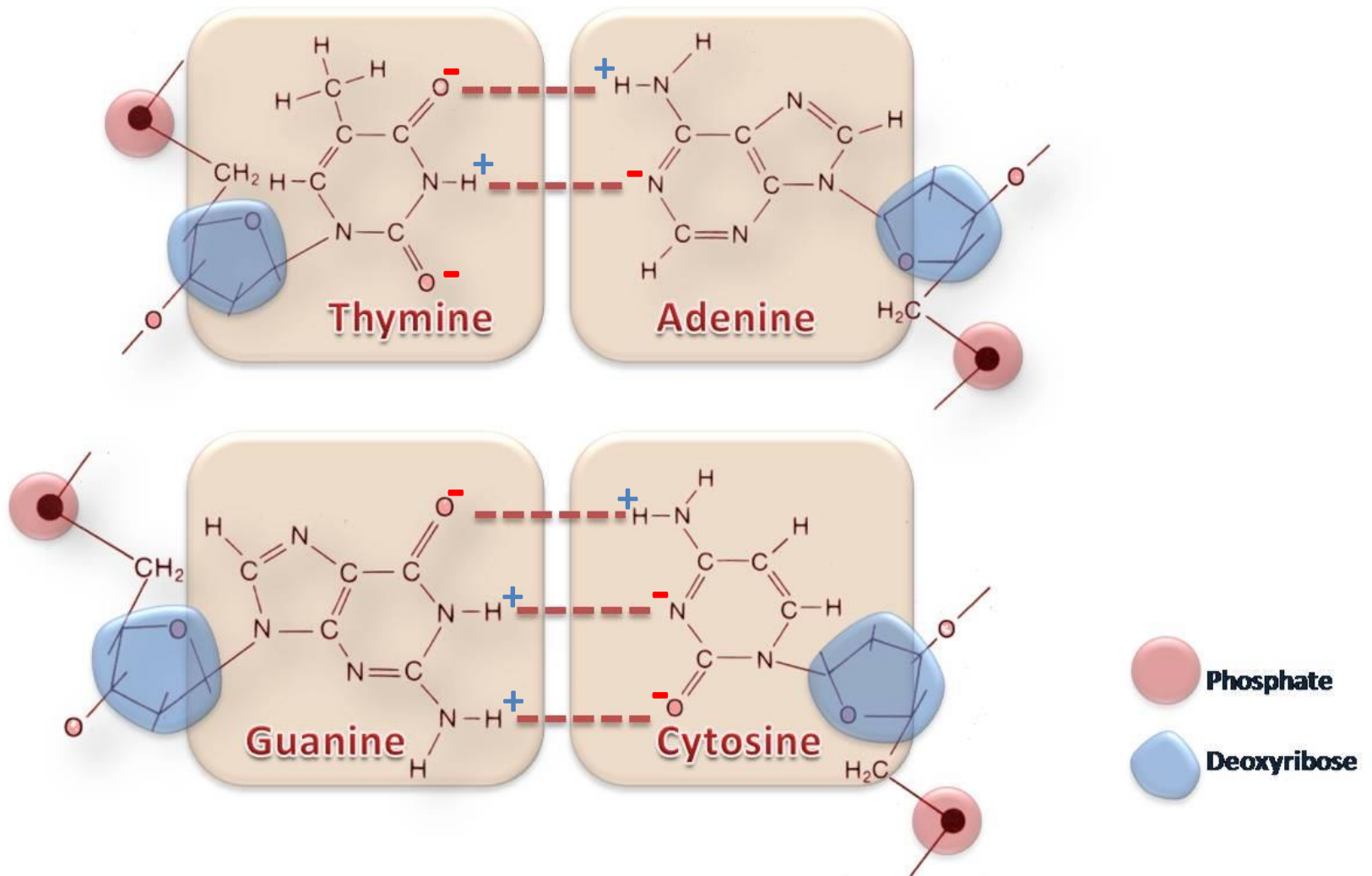
(T)



**Uracil
(RNA)**

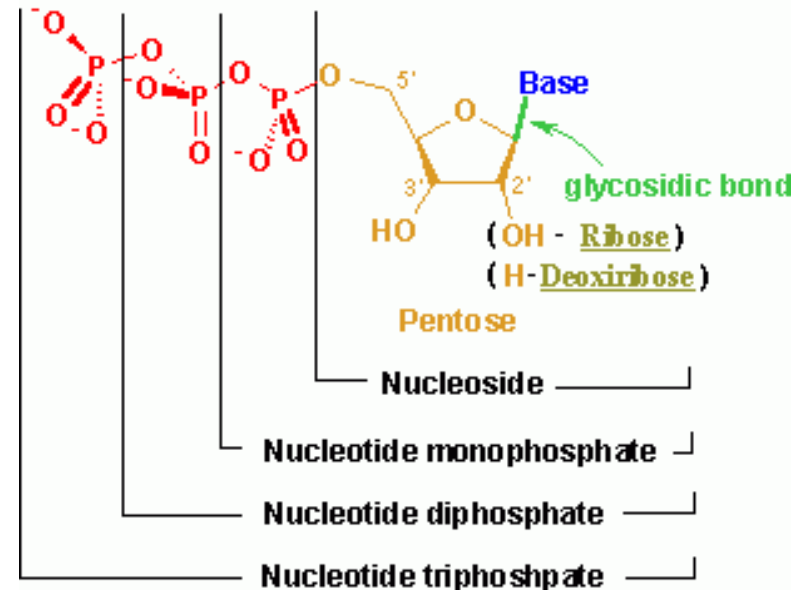
(U)

Nitrogenous Base-Pairing



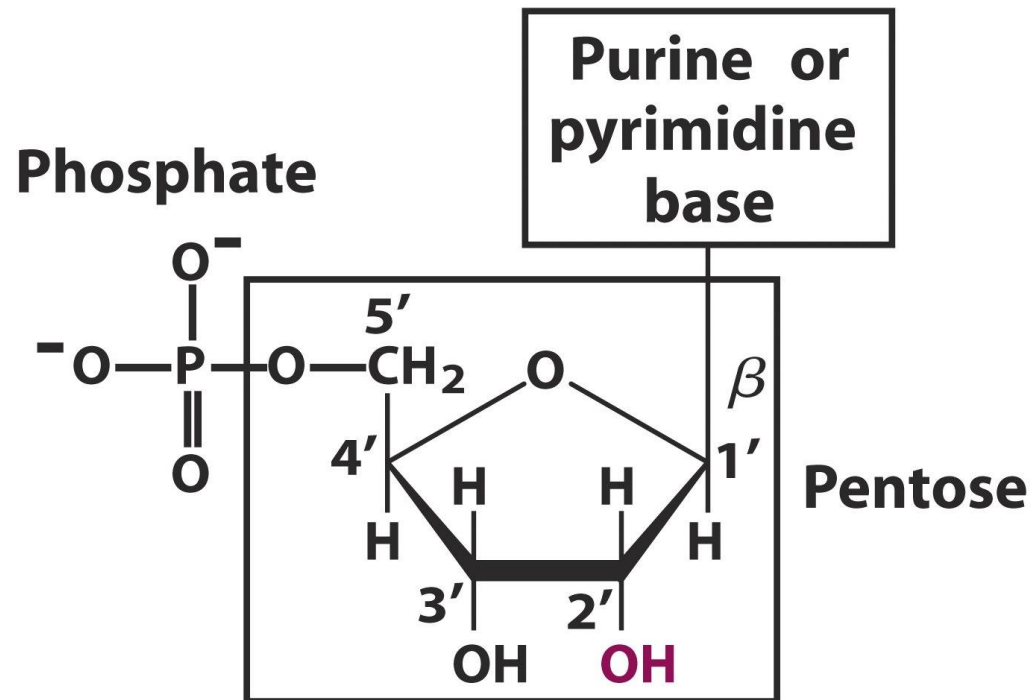
Nucleotides

- Nucleotide
 - Monomeric unit
 - Contains: phosphate, pentose sugar (ribose/deoxyribose) & nitrogenous base (A, T, G, C or U)
- Nucleoside
 - Monomeric unit
 - Contains: as above, minus the phosphate group
- Nucleic acid
 - Polymeric unit
 - DNA: deoxyribonucleic acid
 - No hydroxyl group on C2' of the ribose sugar
 - Thymine as 2nd pyrimidine
 - RNA: ribonucleic acid
 - Hydroxyl group on C2' of the ribose sugar
 - Uracil as 2nd pyrimidine
 - Oligonucleotide: 2-50 nucleotides long
 - Polynucleotide: >50 nucleotides long

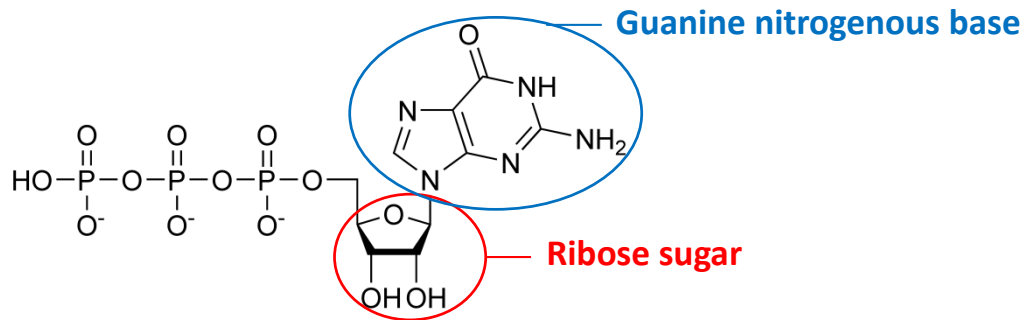


Nucleotides

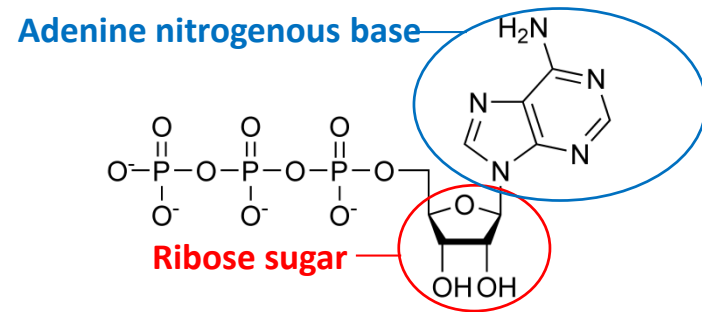
- Phosphate group attached by phosphoester bond to C5' of the ribose ring.
- Nitrogenous base attached by an *N*- β -glycosidic bond to C1' of the ribose ring to N-1 (pyrimidine), N-9 (purines).
 - *N*: indicates nitrogen has replaced oxygen in this bond
 - β : above the plane of the ring



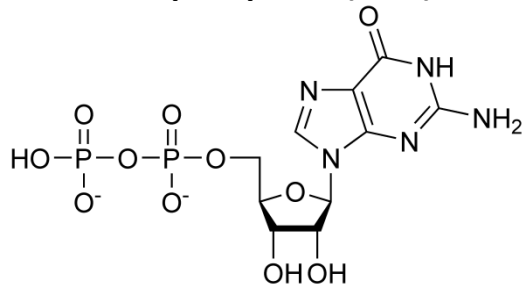
Guanosine 5'-triphosphate (GTP)



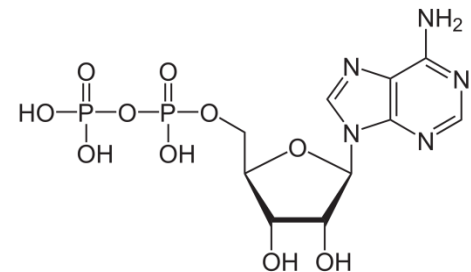
Adenosine 5'-triphosphate (ATP)



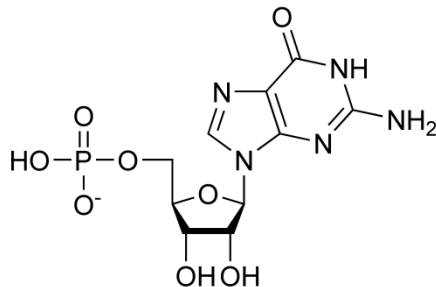
Guanosine 5'-diphosphate (GDP)



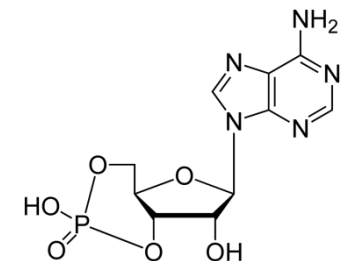
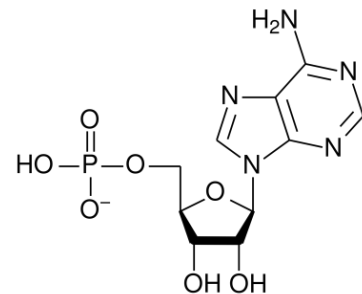
Adenosine 5'-diphosphate (ADP)



Guanosine 5'-monophosphate (GMP)



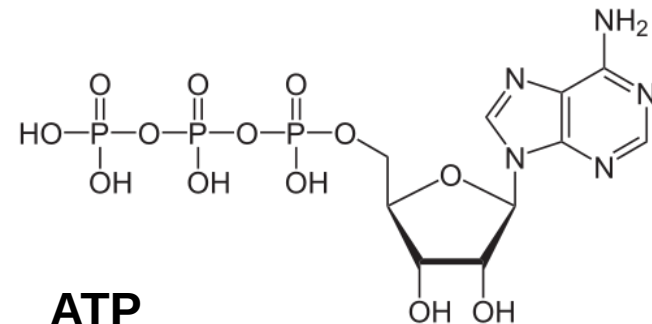
Adenosine 5'-monophosphate (AMP)



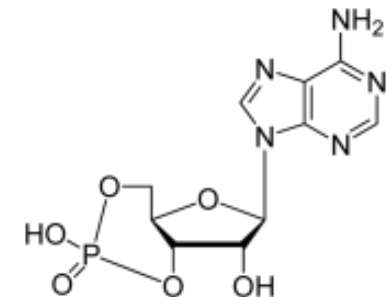
Cyclic adenosine 5'-monophosphate (cAMP)

Nucleotides – Other Roles

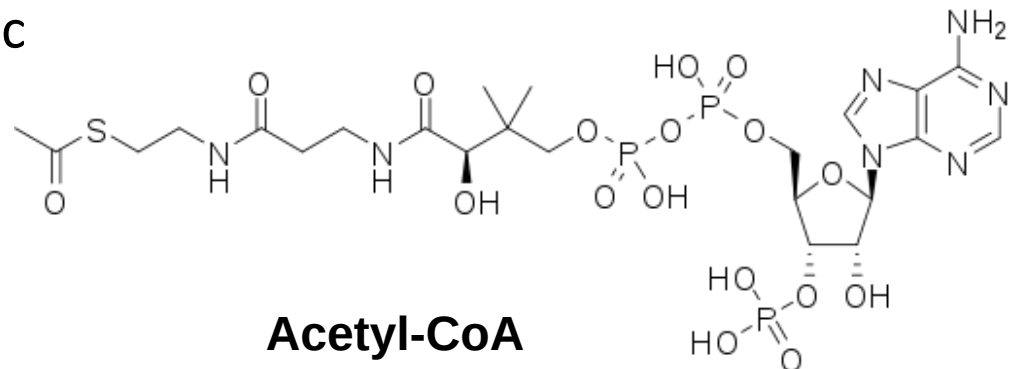
- **As a source of energy**
 - Adenosine triphosphate (ATP), ADP, AMP, cAMP
 - Guanosine triphosphate (GTP), GDP, GMP, cGMP
- **Chemical signal** (signalling cascade)
 - E.g. cAMP/cGMP or ADP
- **Structural component** of coenzymes and metabolic intermediates



ATP



cAMP



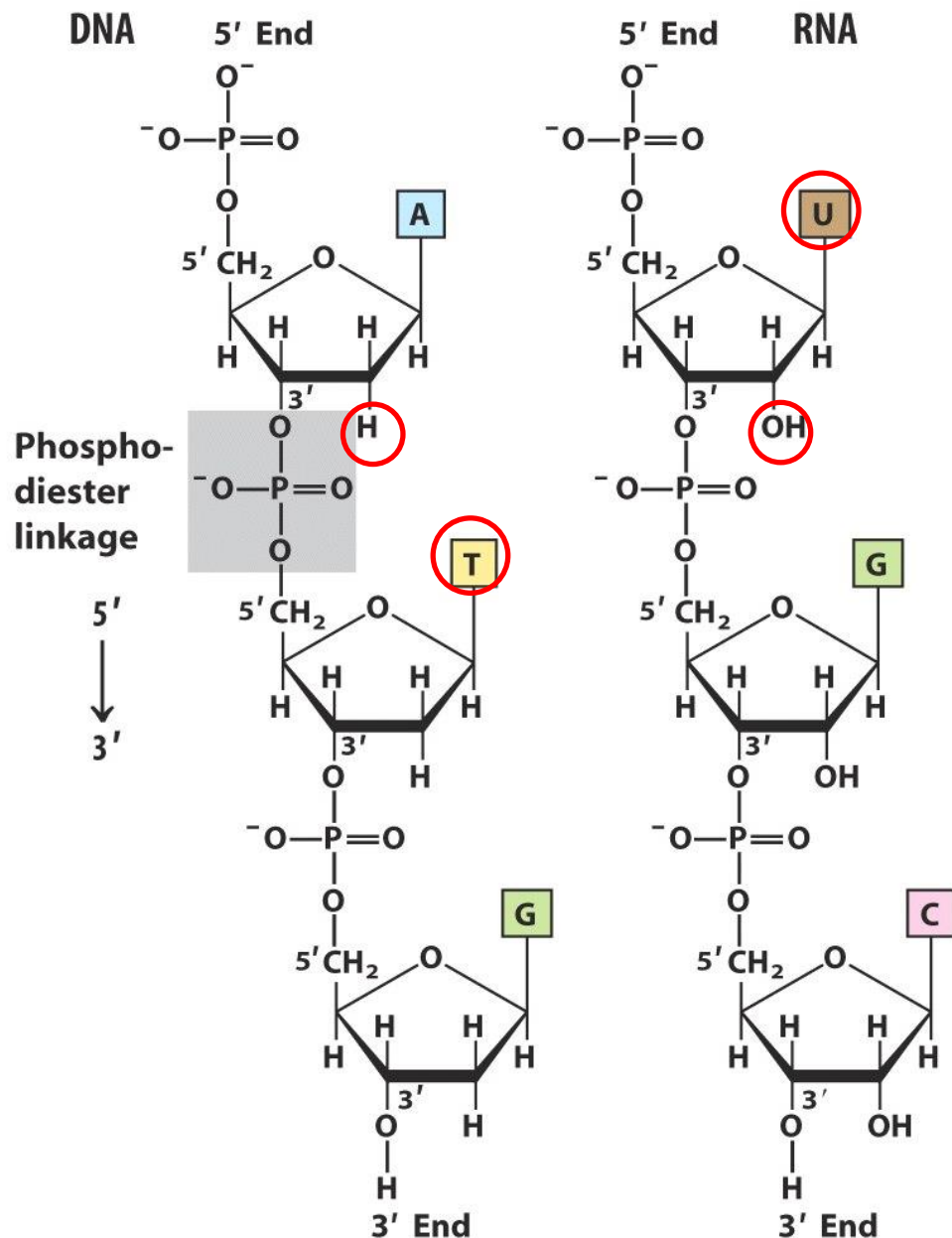
Acetyl-CoA

Formation of Polymers – Nucleic Acids

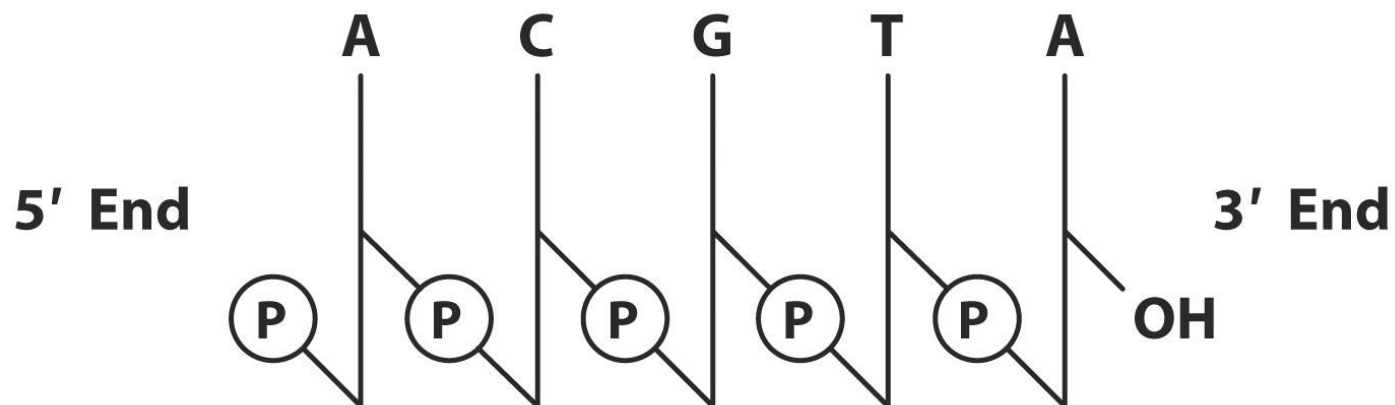
- Polymers have three layers of structure:
 - Primary: polymer of nucleotides
 - Secondary: double helix
 - Tertiary: chromosomes
- How does this compare to the primary, secondary and tertiary structure of proteins?

Primary Structure of DNA and RNA

- Polymer
- Comprised of:
 - Deoxyribonucleotides
 - Ribonucleotides
- Joined by phosphodiester bonds: (5') C3' → C5' (3')
- Has directionality
 - 5' → 3'



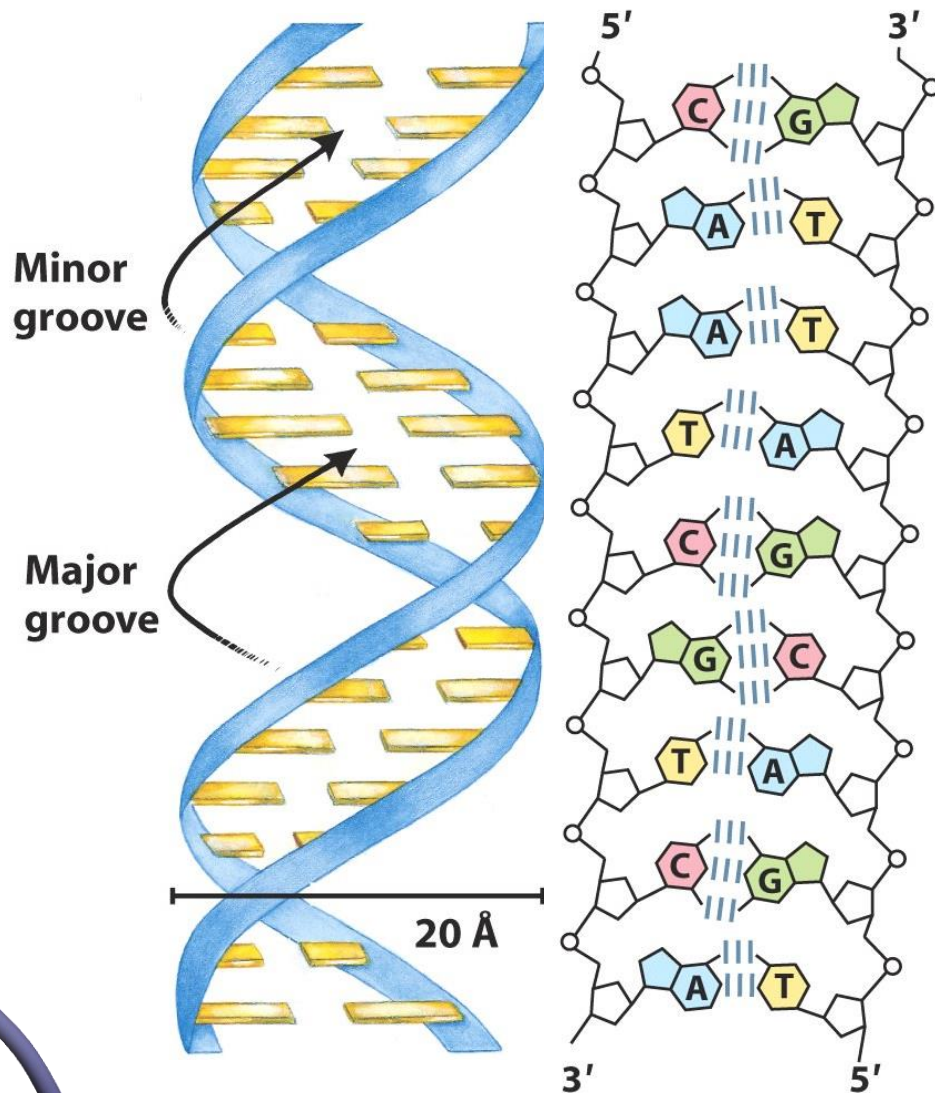
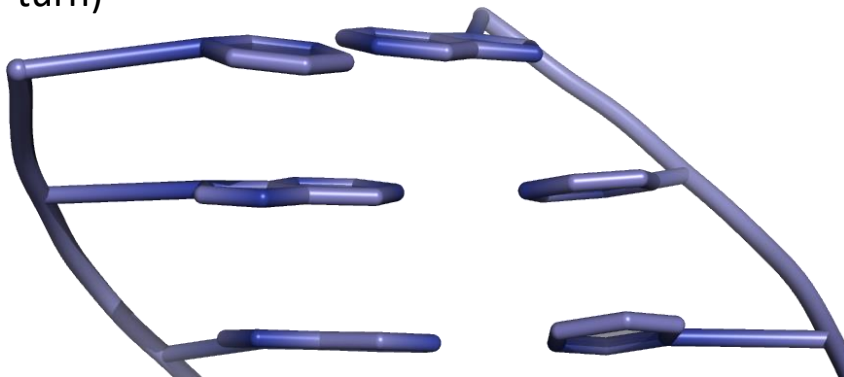
Nucleic Acids have Distinct Polarity & Directionality



- 5' end represents a free phosphate (PO_4) group
- 3' end represents a free hydroxyl (OH) group
- Sugar/phosphate backbone: hydrophilic, ionic (neg. charged phosphate groups)
- DNA/RNA has an overall negative charge
- Nitrogenous bases (R-groups): upper/lower surface hydrophobic, while edges are hydrophilic

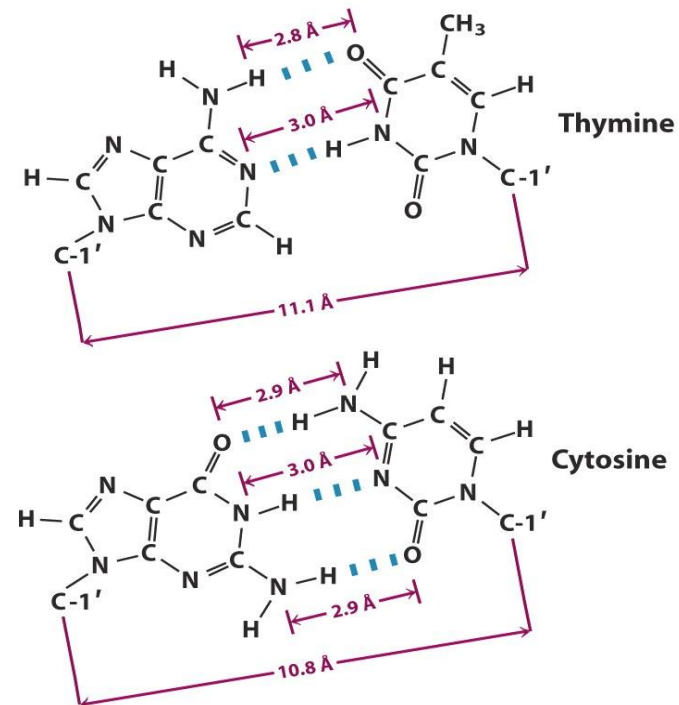
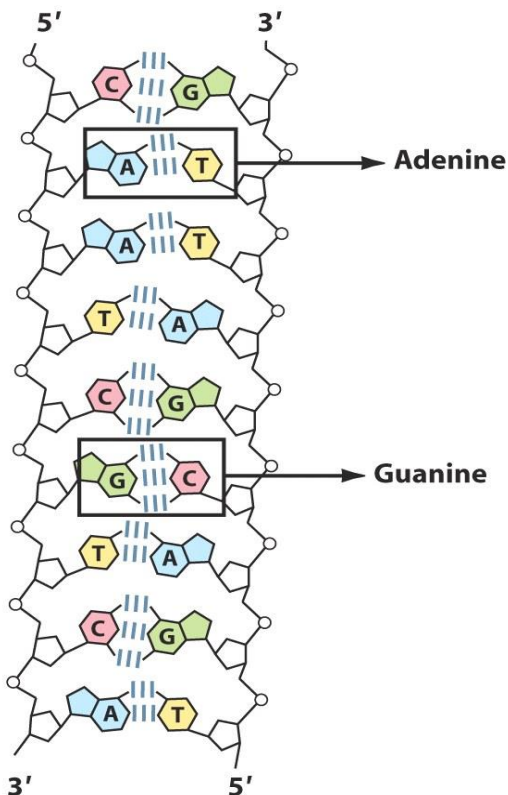
Secondary Structure – DNA

- Anti-parallel right-handed double helix of two polynucleotide strands
- Held together by hydrogen bonds between base-pairs that sit in the same plane, as well as base-stacking interactions
- Each strand is complementary to the other, rather than being identical.
- Helical shape creates a major & minor groove
- Vertically stacked bases are 3.4 Å (0.34nm) apart, with there being 10.5 bases per turn (36 Å per turn)



Secondary Structure – DNA

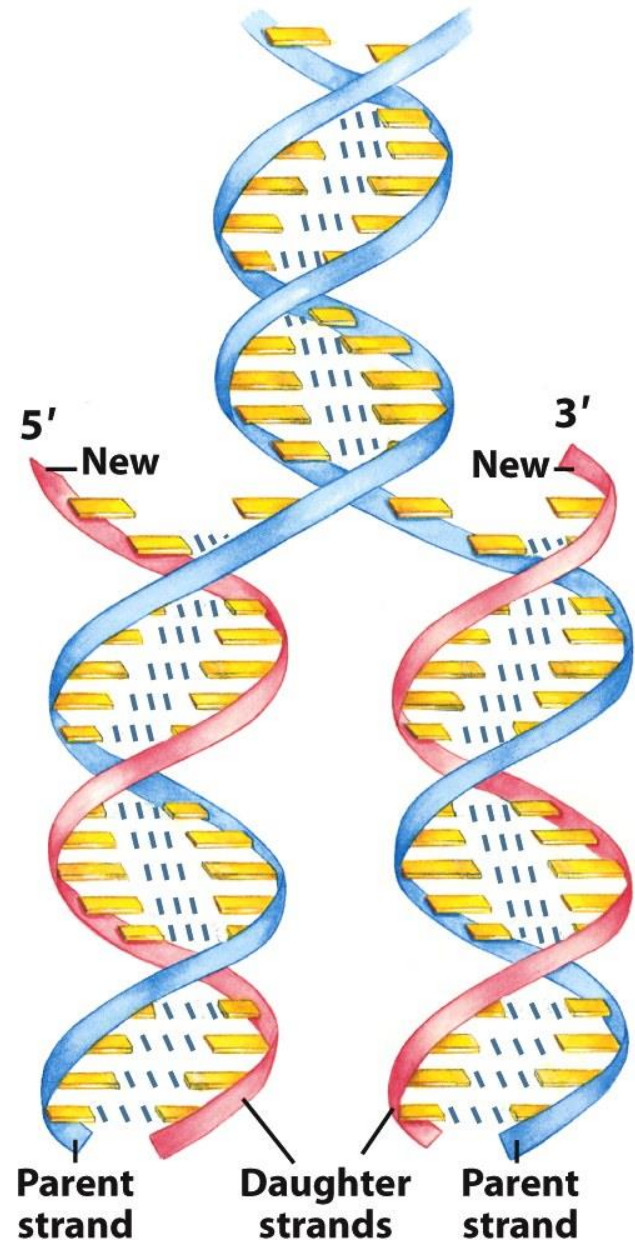
- In 1869, Friedrich Miescher isolated "nuclein" (DNA with associated proteins). He was the first to identify DNA as a distinct molecule.
- However, until the mid 20th century, scientists believed proteins carried our hereditary information
- Erwin Chargaff (1950) noted that the number of C = G, and the number of A = T in any given sequence.
- Watson and Cricks model of DNA showed:
 - C≡G (3 H-bonds)
 - T=A (2 H-bonds)
- 2 purines, or 2 pyrimidines will not base-pair due to size
- Note: C≡G is a stronger bond, therefore, higher GC content makes it harder to disrupt the DNA secondary structure



QUESTION: If cytosine (C) was to make up 28% of your genome, what percentage of your genome is composed of A, T & G?

DNA Replication

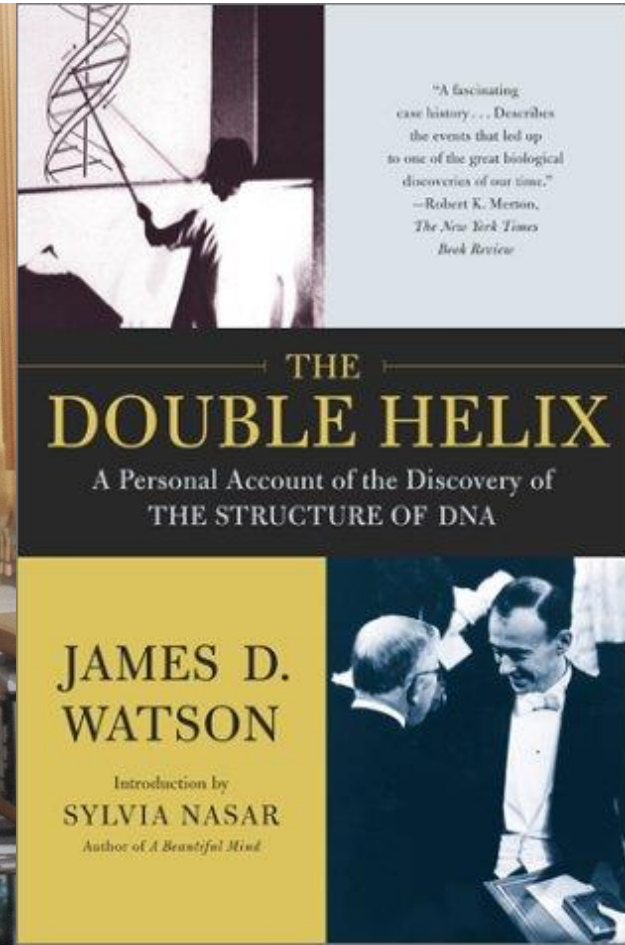
- The anti-parallel nature of DNA, with complementary base pairing suggested a logical mechanism for the transmission of information
- This diagram shows a model for DNA replication





Cold Spring Harbor Laboratory

18

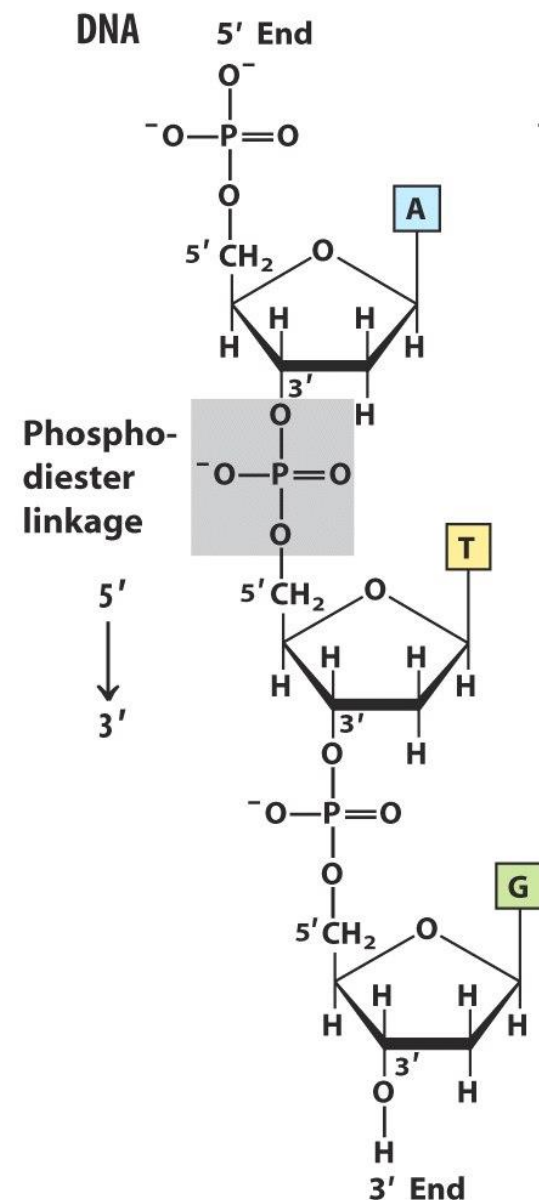


Advanced Techniques in Molecular Neuroscience 2011

Week 9: Nucleotides and Nucleic Acids

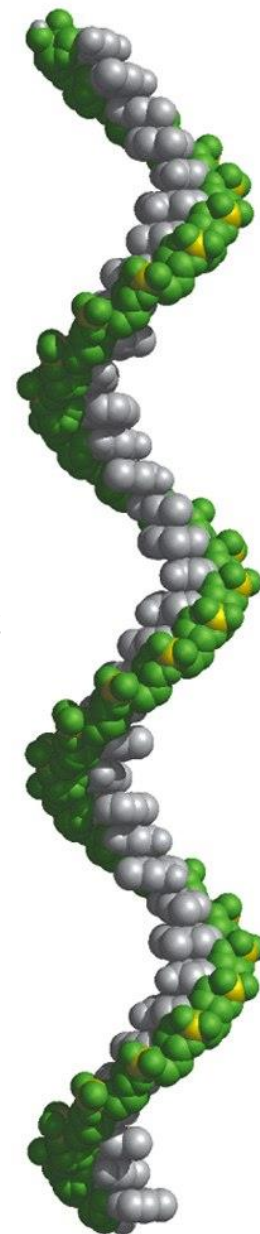
Summary of DNA

- In Eukaryotes: found in the cell nucleus and mitochondria
- Sugar backbone: deoxyribose, no OH group on C2' of the sugar ring
- Nitrogenous bases: A, G, C and T(U) ($C \equiv G$, $A = T$)
- Sugar/phosphate backbone: hydrophilic, ionic (neg. charged phosphate groups)
- DNA/RNA has an overall negative charge
- Nitrogenous bases (R-groups): generally hydrophobic
- Secondary structure: anti-parallel double stranded right-handed α -helix
- Tertiary structure: chromosomes

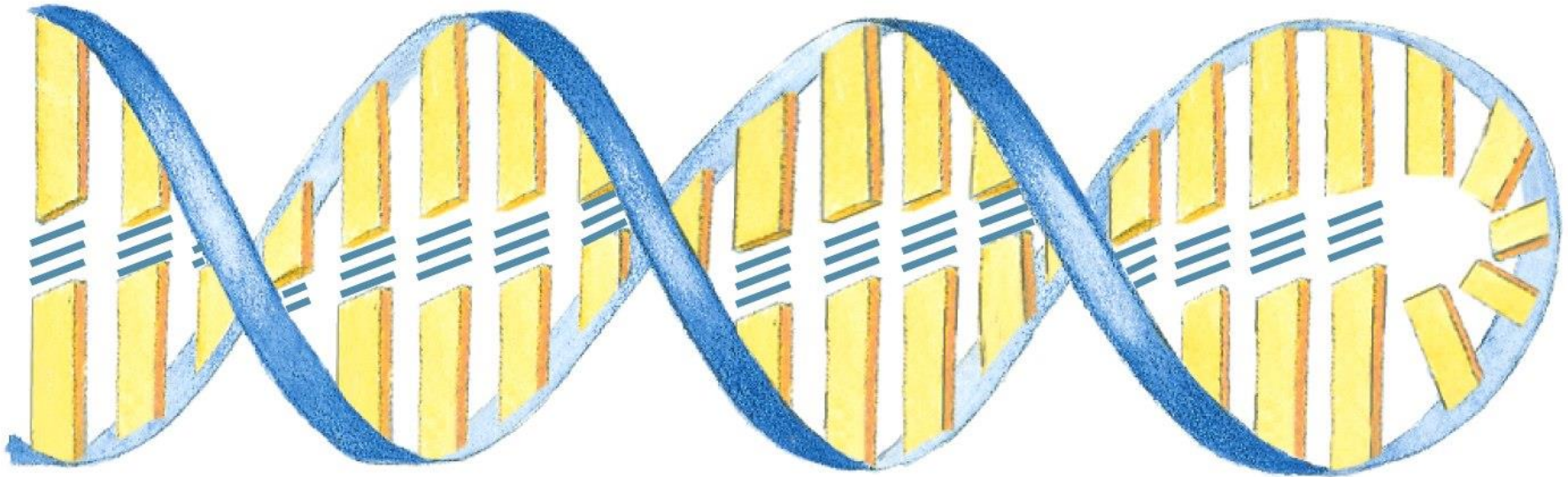


How is RNA Different?

- Is a nucleic acid, like DNA
- Normally, single stranded (exceptions: certain viruses and some regulatory units)
- Uracil (U) instead of Thymine (T)
- Ribose instead of deoxyribose
- 4 main types:
 - mRNA (messenger RNA): 'transcribed' from DNA, carries genetic information from DNA to ribosomes where they are 'translated' into proteins
 - tRNA (transfer RNA): role in translation to protein
 - rRNA (ribosomal RNA): form ribosomes (enzymatic units for the synthesis of proteins)
 - Non-coding RNA: regulatory abilities (e.g. microRNA, siRNA and piRNA etc.)



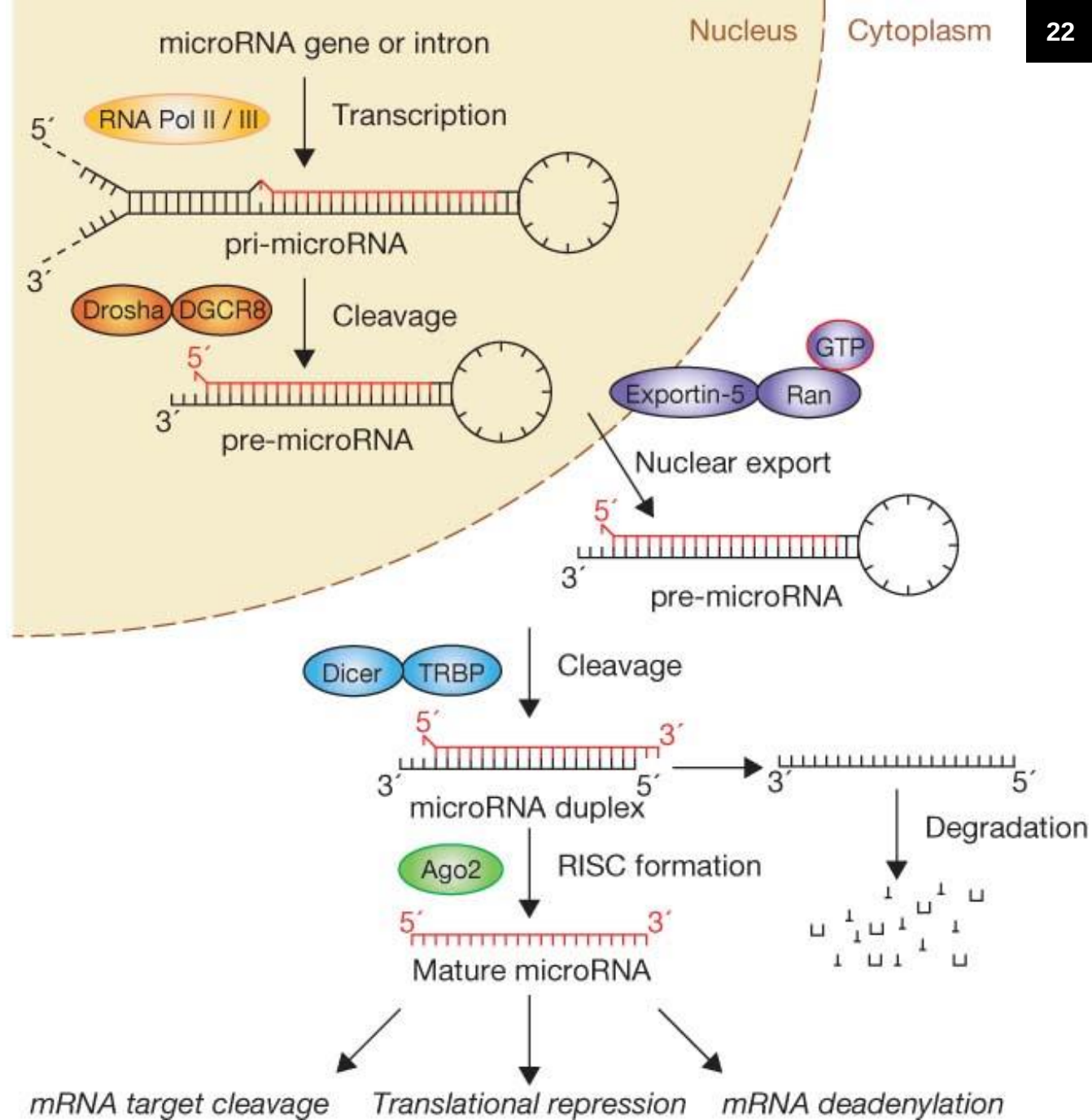
RNA can form Secondary Structures



Hairpin double helix

RNA Secondary Structure:

MicroRNAs



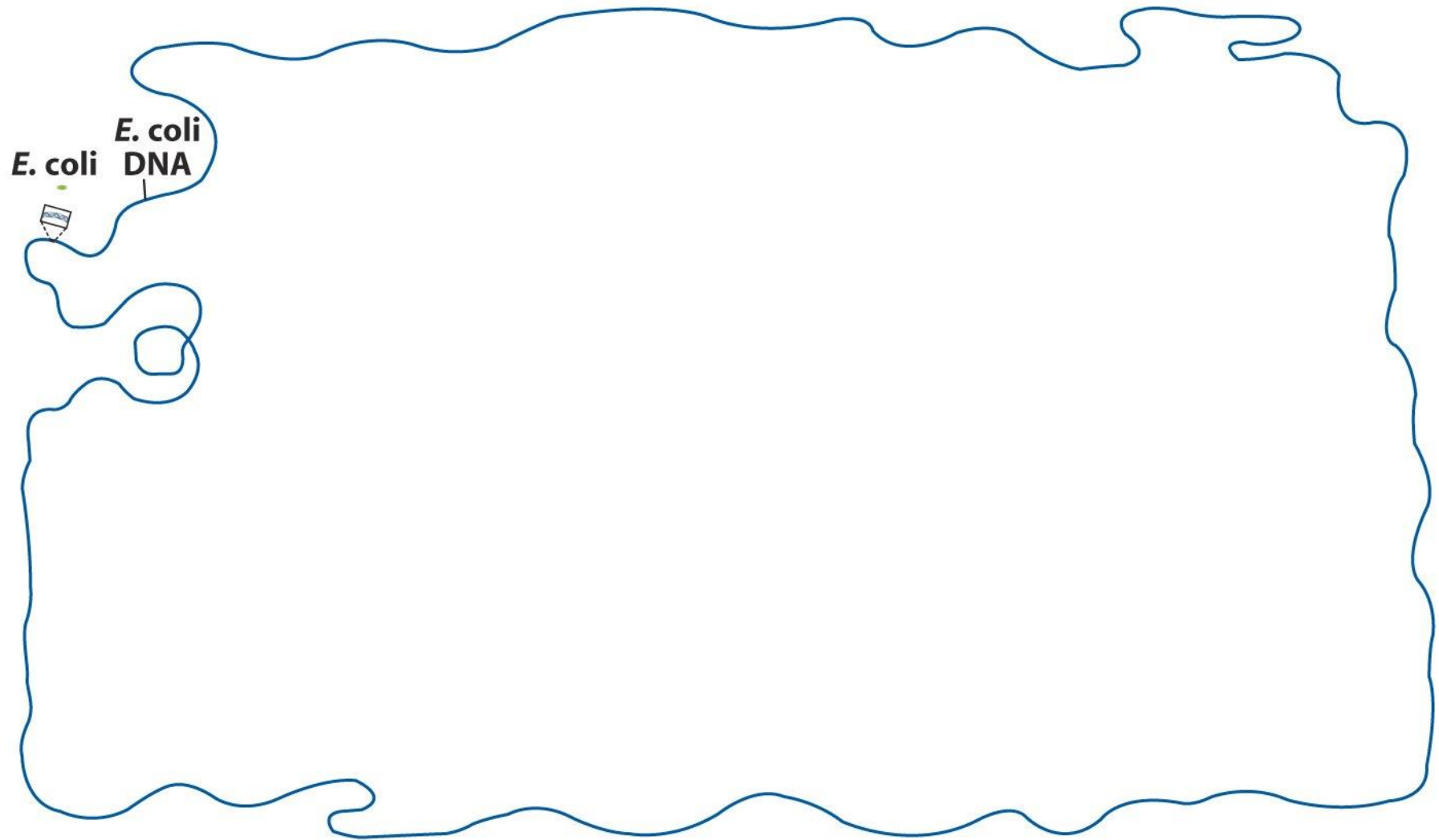
DNA has Tertiary Structure

Unwound DNA is greater in length than the cells that contain it!

TABLE 24-1 The Sizes of DNA and Viral Particles for Some Bacterial Viruses (Bacteriophages)

<i>Virus</i>	<i>Size of viral DNA (bp)</i>	<i>Length of viral DNA (nm)</i>	<i>Long dimension of viral particle (nm)</i>
ϕ X174	5,386	1,939	25
T7	39,936	14,377	78
λ (lambda)	48,502	17,460	190
T4	168,889	60,800	210

Note: Data on size of DNA are for the replicative form (double-stranded). The contour length is calculated assuming that each base pair occupies a length of 3.4 Å (see Fig. 8-15).



Human Genome

- **Genome:** entirety of the hereditary genetic information of a cell
- **Chromosome:** in eukaryotes, the genome is divided into sets of linear segments of DNA referred to as chromosomes (46 for humans, 23 pairs (somatic cells))
- **Chromatin:** the complex of DNA and proteins found together
- **Nucleosome:** the most basic unit of DNA packing involving DNA wound around histone protein complexes
- **Centromere:** the point that links sister chromatids and allows separation into daughter cells
- **Telomeres:** specialised DNA sequences found at the end of the chromosome to allow complete replication and to protect against degradation of the chromosomal ends. Consists of thousands of tandem 5'-TTAGGG-3' repeats (up to 2500 repeats).

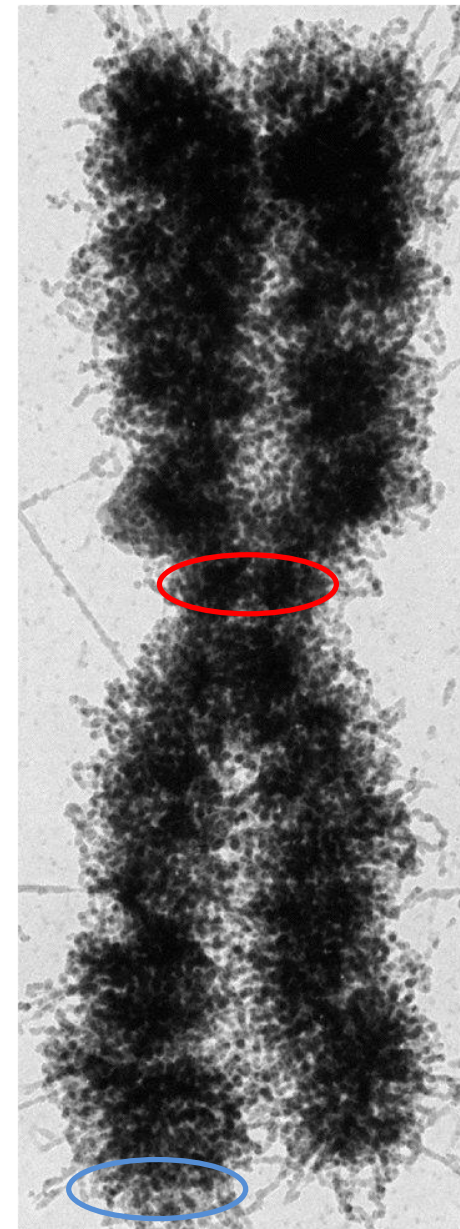


Image taken during mitosis

TABLE 24-2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes ^a	Approximate number of genes
<i>Escherichia coli</i> K12 (bacterium)	4,641,652	1	4,494 ^b
<i>Saccharomyces cerevisiae</i> (yeast)	12,157,105	16 ^c	6,340 ^b
<i>Caenorhabditis elegans</i> (nematode)	90,269,800	12 ^d	23,000
<i>Arabidopsis thaliana</i> (plant)	119,186,200	10	33,000
<i>Drosophila melanogaster</i> (fruit fly)	120,367,260	18	20,000
<i>Oryza sativa</i> (rice)	480,000,000	24	57,000
<i>Mus musculus</i> (mouse)	2,634,266,500	40	27,000
<i>Homo sapiens</i> (human)	3,070,128,600	46	20,000
Note: This information is constantly being refined. For the most current information, consult the websites for the individual genome projects. ^aThe diploid chromosomes number is given for all eukaryotes except yeast. ^bIncludes known RNA-coding genes. ^cHaploid chromosomes number. Wild yeast strains generally have eight (octoploid) or more sets of these chromosomes. ^dNumber for females, with two X chromosomes. Males have an X but no Y, thus 11 chromosomes in all.			

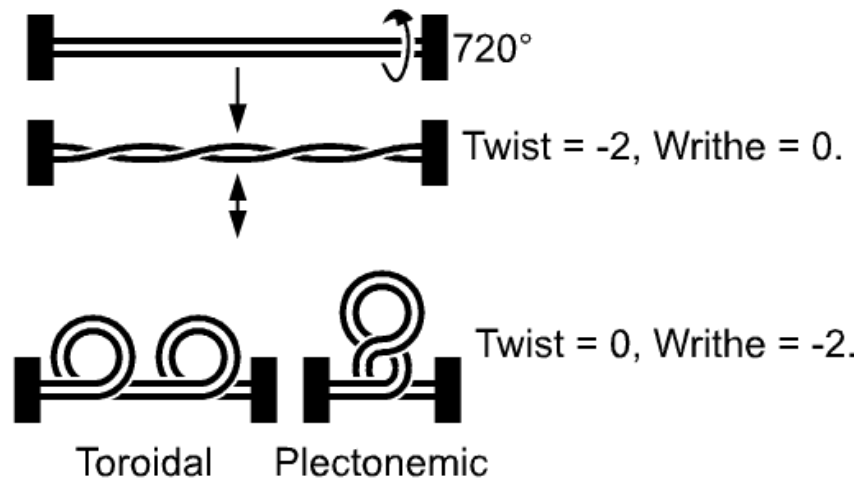
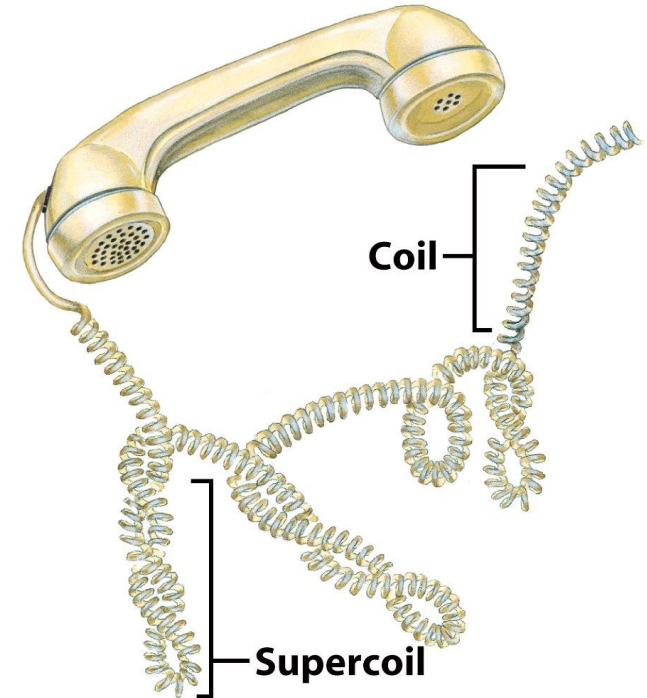
Human Genome

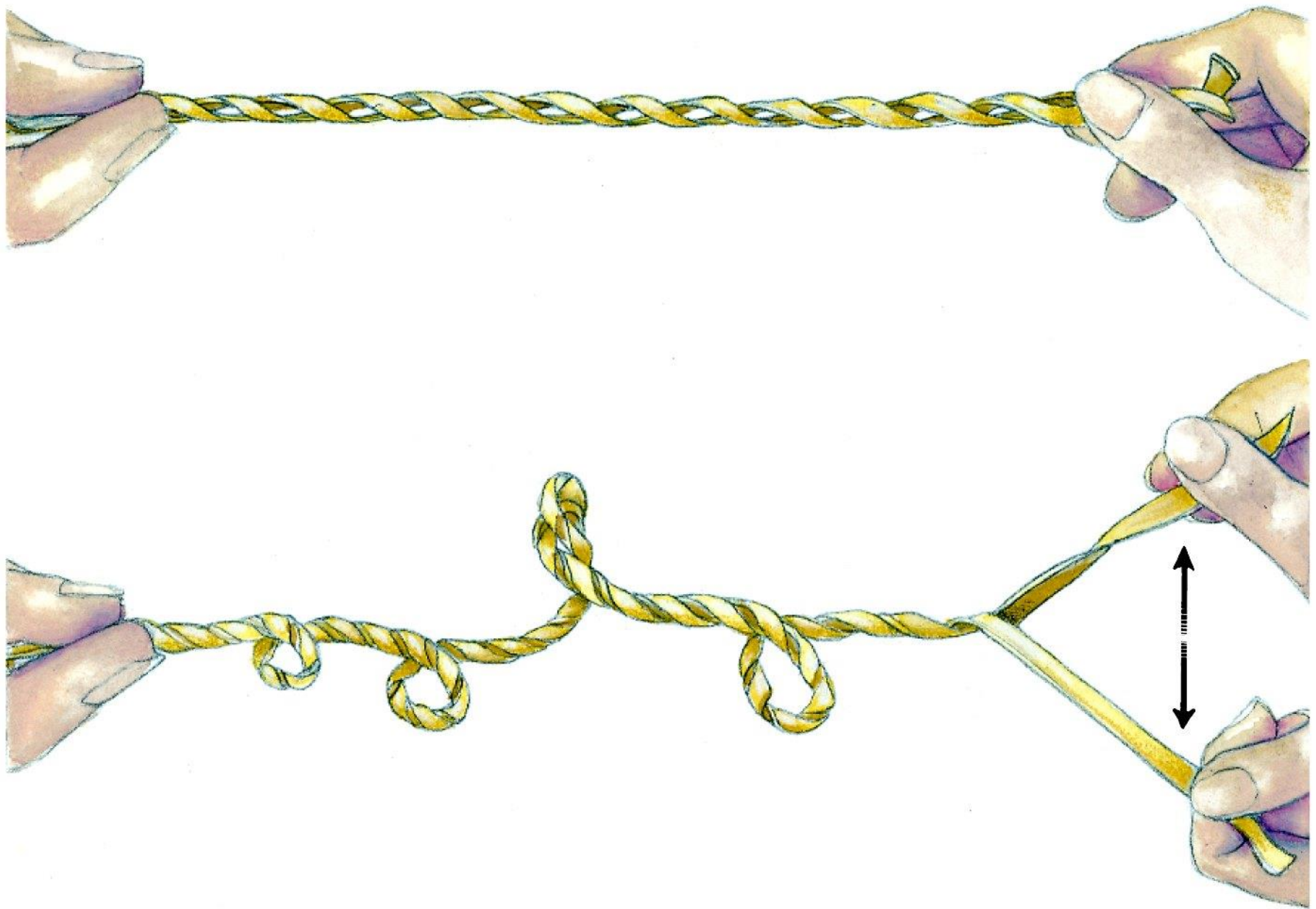
- Every somatic cell has
 - 46 chromosomes
 - From Mum: 22 non-sex chromosomes (autosomes) + X (allosome)
 - From Dad: 22 non-sex chromosomes (autosomes) + X or Y (allosome)
- Somatic cells are diploid ($2n$, 2×23)
- DNA of one human genome (23 chromosomes) would stretch for about a meter, so diploid cells have 2m of DNA



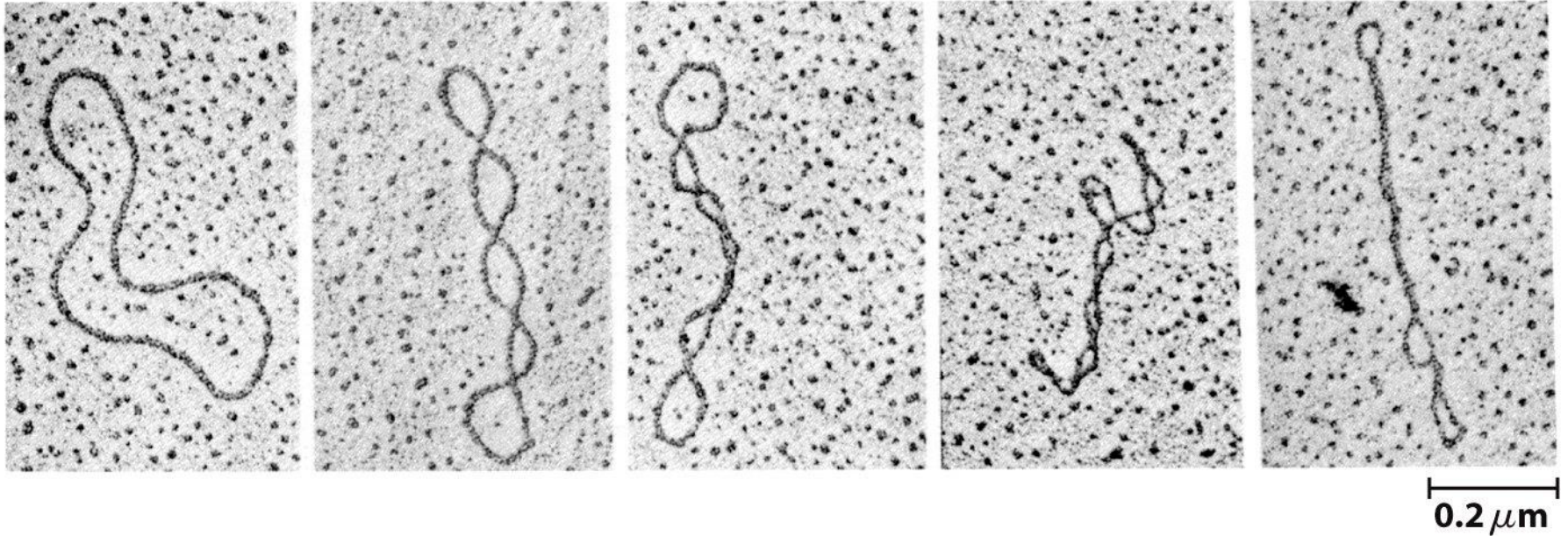
DNA Packaging

- Cellular DNA is extremely compact
- Supercoiled
 - Positive supercoil
 - Extra helical twist
 - Negative supercoil
 - Subtractive twisting
- In nature, negatively supercoiled DNA is more common
- In eukaryotes, supercoils exist predominantly as plectonemic and solenoidal (nucleosomes)





Relaxed vs Supercoiled Plasmid DNA



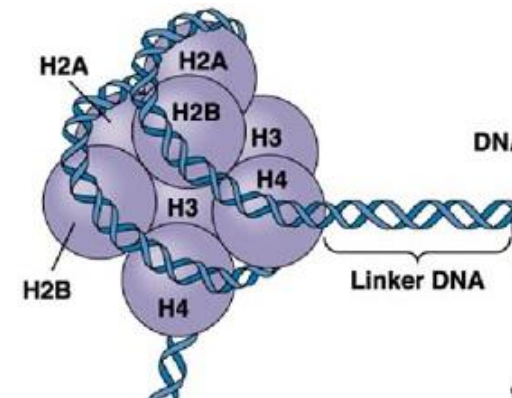
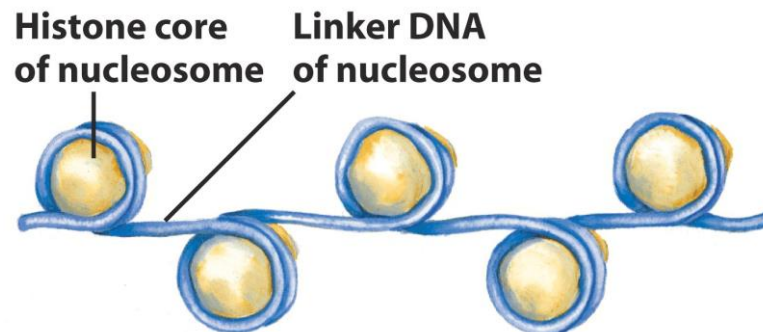
- DNA can be unwound in the cell using enzymes (topoisomerases)
- DNA needs to be unwound in order for replication & transcription to take place

DNA Packaging in Eukaryotes

- In the human nucleus, genomic DNA must be tightly packed (~3.3 billion bp's, or 2m of DNA must fit into a cell nucleus 6 μ m in diameter)
- This is achieved through
 - General supercoiling, discussed before
 - Solenoidal supercoiling, which involves histone proteins

Histones and Nucleosomes

- Five major classes:
 - H1, H2A, H2B, H3, H4
 - H3 & H4: Highly conserved (e.g. similar DNA/protein sequence between species)
 - H1, H2A & H2B: Less conserved
 - High Lys and Arg content
 - Cationic
 - Interacts with anionic DNA backbone
- The basic level of DNA packing (nucleosomes - beads on a string)
 - Composed of: a complex of 8 histones (2 X each of histones H2A, H2B, H3 & H4), and a region of double stranded DNA that is 146bp long wound around it
 - Each nucleosome is separated by a region of linker DNA (3-80bp long)
 - H1: linker histone to allow nucleosomes and DNA between the nucleosomes to form higher order chromatin structures.

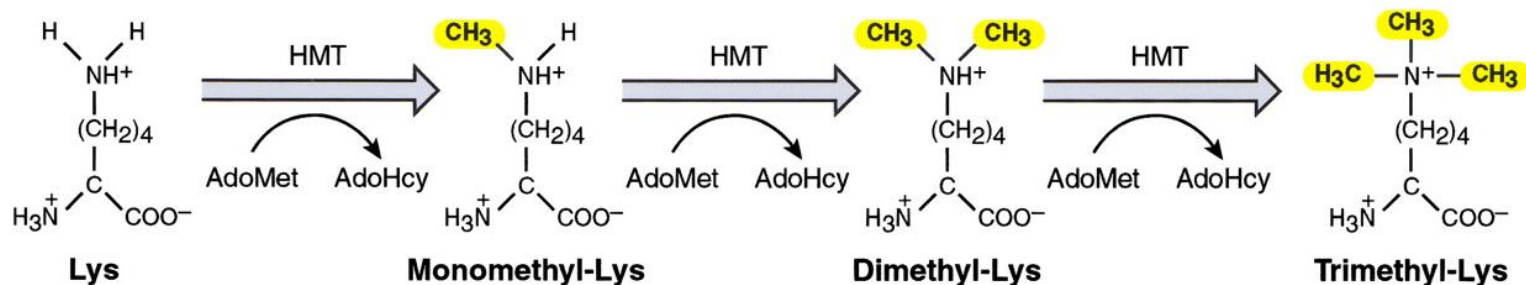
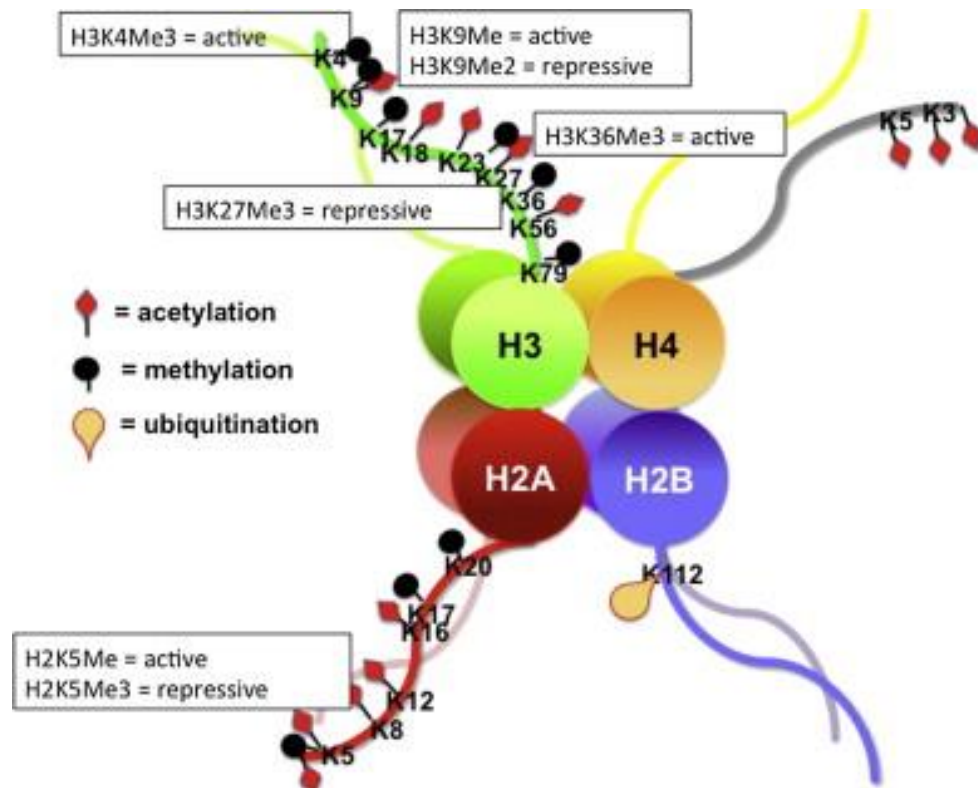


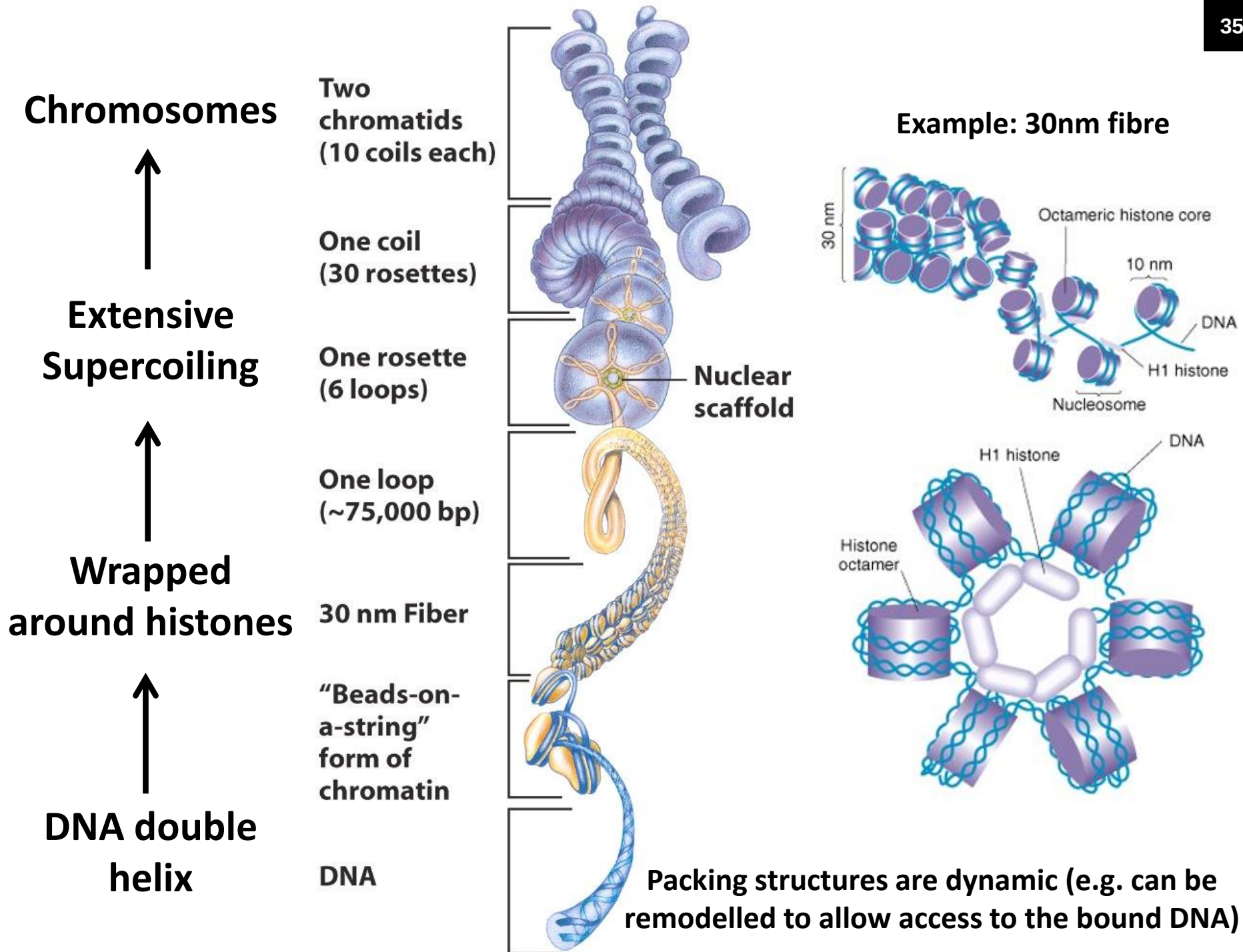
Nucleosomes

- The interface between the DNA and histones in a nucleosome are extensive:
 - ~142 hydrogen bonds (~50% between the amino acid backbone of the histones & the phosphodiester backbone of the DNA)
 - Hydrophobic interactions
 - As discussed, ionic interactions between the –ve charged DNA backbone and +ve charged histones
- As might be expected from their fundamental role in DNA packing, a number of histones are highly conserved, i.e.
 - Only 2 of the 102 aa's differ between peas and cows for H4.
 - Only 8 of the 102 aa's differ between humans and yeast for H4.

Histone Modifications

- Chromatin is not an inert structure, rather dynamic changes in the chromatin state, in response external cues, alters the packaging of DNA.
 - This regulates all aspects of genome function by altering the accessibility of DNA to copy, repair, express and splice the genome.
- A principle component of this dynamic regulation is a diverse number of modifications that can be made to histones.
- Modifications to the amino acid tail of each histone or replacement with a different form (i.e. H2AZ for H2A) regulates chromatin tertiary and quaternary structure but also recruits remodelling enzymes.
- For example, tri-methylation of the lysine residue at position 27 in the protein histone H3 (H3K27me3), is associated with gene repression.

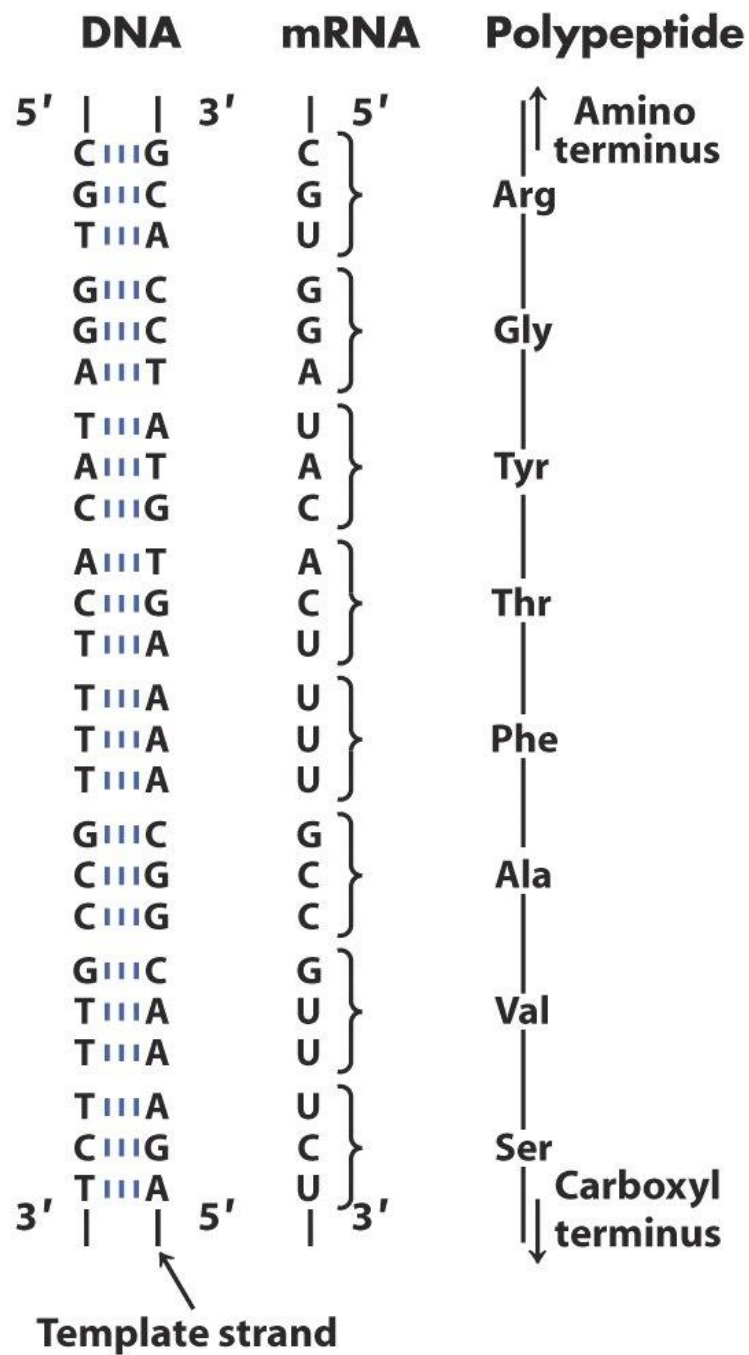




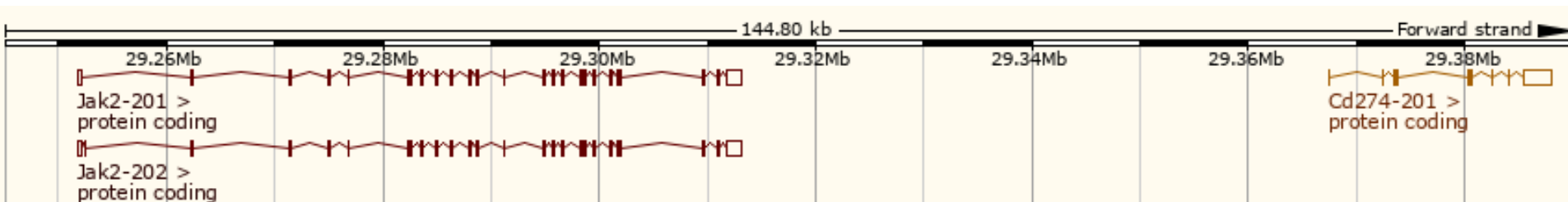
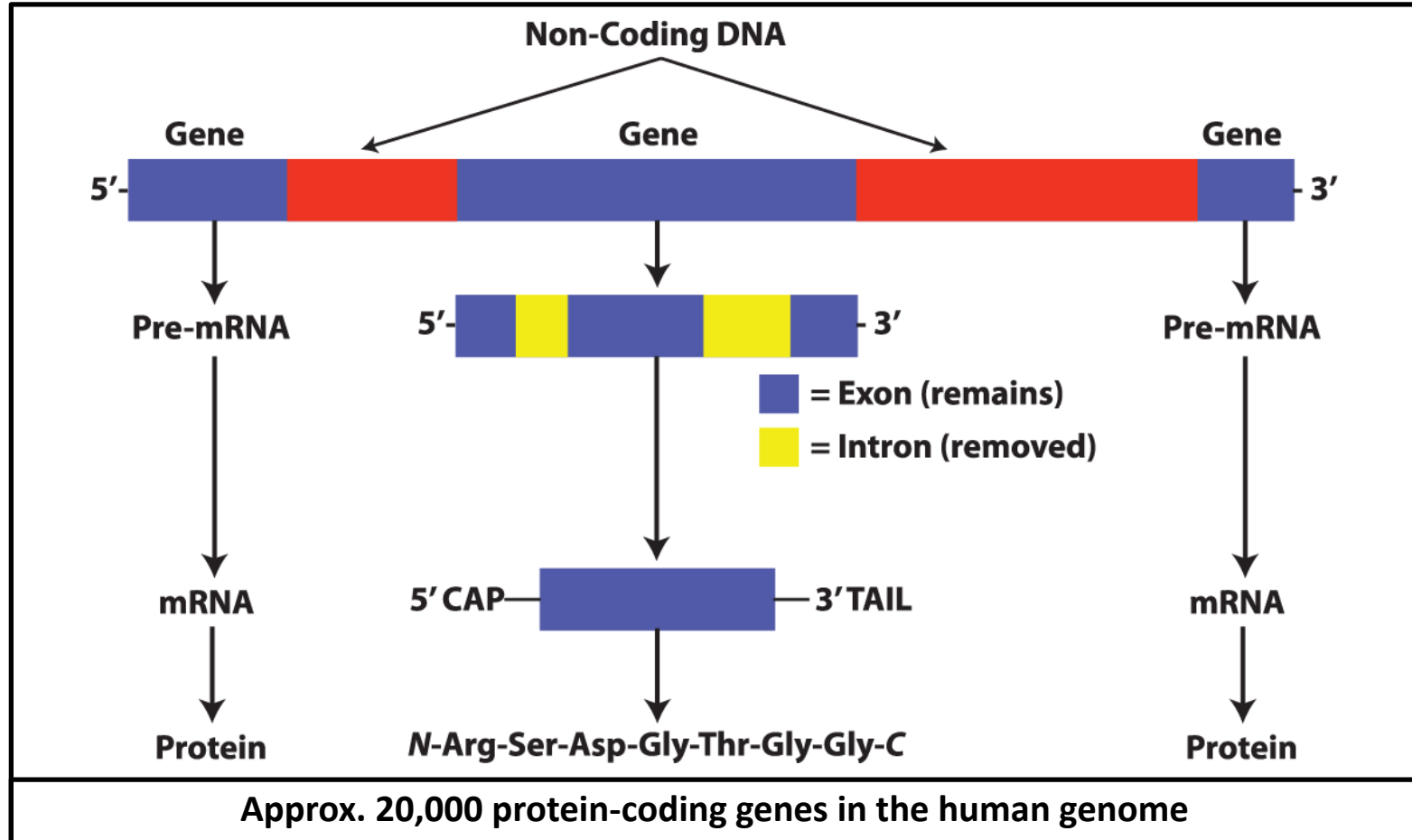
Genes and Chromosomes

- The entire DNA (or genetic) content in an organism is called the 'genome'
- The genome encodes the information necessary to exist
 - The 'gene' is the fundamental unit of information

- Definition of a gene:
 - A chromosomal segment of DNA that **CAN** encode for a specific biologically functional polypeptide or RNA molecule, or set of closely related isoforms
- Old theories:
 - One gene - one single character or phenotype
 - One gene - one polypeptide or protein



DNA is Composed of Multiple Genes



Mouse Chromosome 19 – 144,800 base-pair region of roughly 61 million base-pairs

Non-Coding DNA

- Once called 'junk DNA' (~98% of your genome)
- Does not code for a final polypeptide product (this includes introns (~25% of the genome))
- Contains numerous repeat sequences from transposable elements (regions of DNA that can move from one location in the genome to another (~45% of the genome))
- Codes for functional non-coding RNA molecules (e.g. tRNA, rRNA, miRNA, siRNA, piRNA, snoRNA, long ncRNA, snRNA etc.)
- Contains many regulatory sequences involved in the regulation of transcription
- **IMPORTANTLY: IT IS AN EVER-CHANGING AREA OF RESEARCH!**

You should be able to...

- Draw and describe the primary and secondary chemical structure of DNA and RNA
- Explain what is meant by the complementarity and polarity of nucleic acids
- List and explain differences between RNA and DNA
- Describe the different types of RNA
- Describe the 3D structure of chromosomes
- Describe the different stages of condensation of DNA into chromosomes
- Explain what is meant by the term “gene”

What's Next?

- ✓ • Topic 1 – Foundations in Biochemistry
- ✓ • Topic 2 – Amino Acids
- ✓ • Topic 3 – Proteins
- ✓ • Topic 4 – Lipids and Membranes

- ✓ • Topic 5 – Signal Transduction
- ✓ • Topic 6 – Enzymes
- ✓ • Topic 7 – Nucleotides and Nucleic Acids
 - Topic 7 – DNA Replication
 - Topic 8 – Transcription
 - Topic 9 – Translation

