BIOLOGY 101

A Timeline of Molecular Biology

TIMELINE OF MOLECULAR BIOLOGY

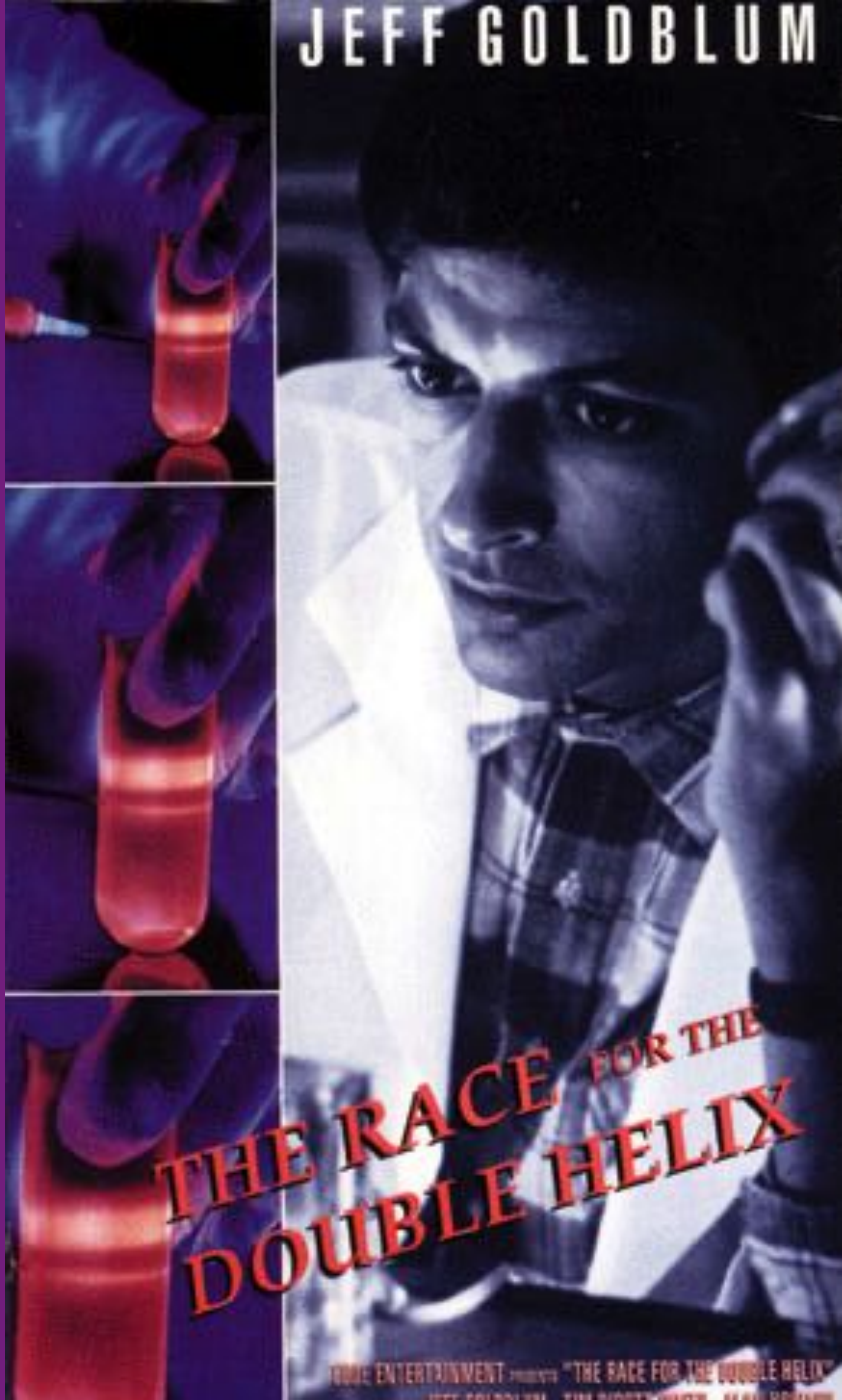
- 1665
  - Robert Hooke discovered organisms are made up of cells
- 1859
  - Charles Darwin published the "On the Origin of Species"
- 1865
  - Gregor Mendel investigated "traits" passed from parents to prodigy and coined the terms dominant and recessive traits

TIMELINE OF MOLECULAR BIOLOGY

- 1900
    - Chemical structures of all 20 amino acids had been identified
- 1902
    - Emil Hermann Fischer wins Nobel prize: showed amino acids are linked and form proteins
- 1941
    - George Beadle and Edward Tatum identify that genes make proteins

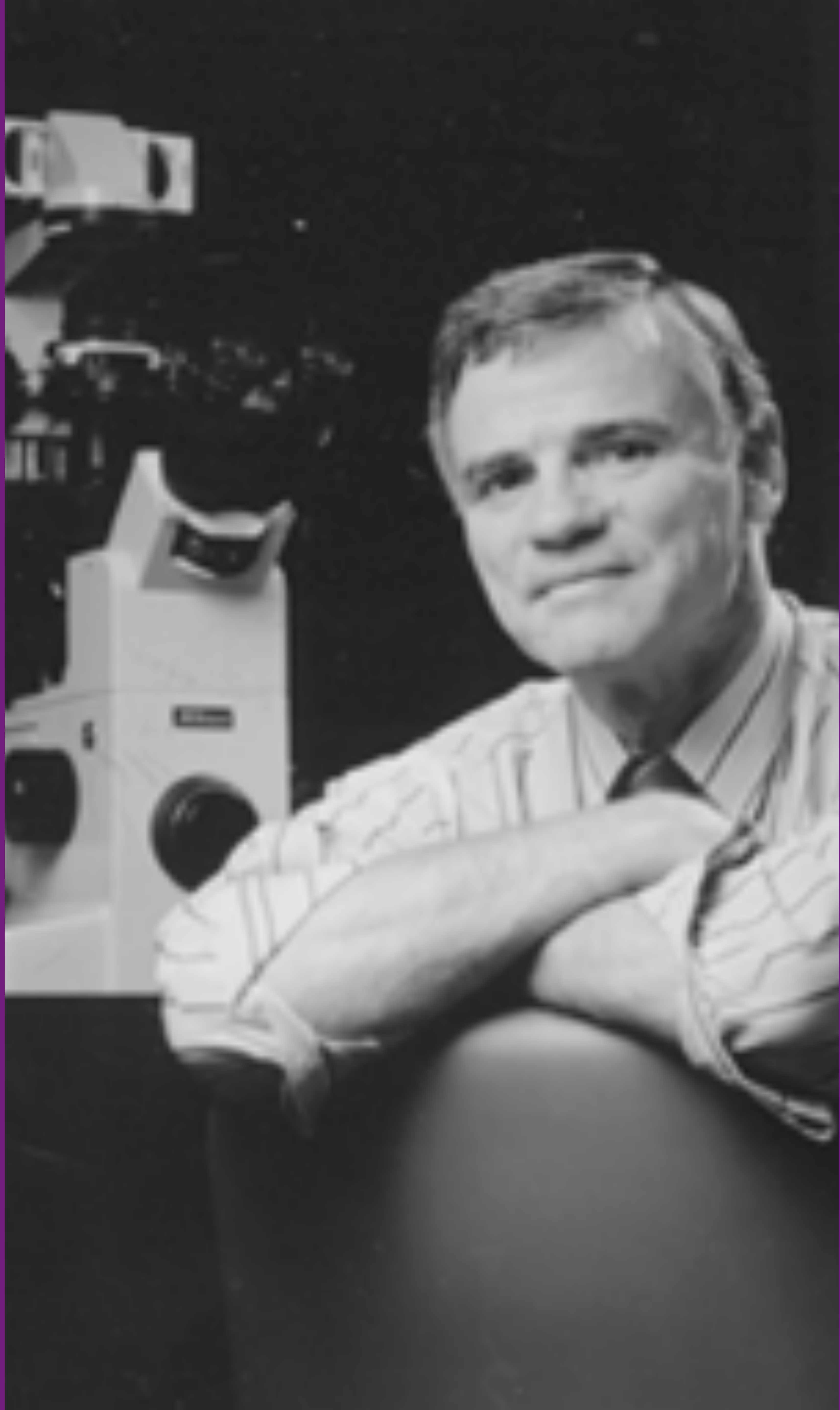
TIMELINE OF MOLECULAR BIOLOGY

- 1952-1953
  - Rosalind Franklin produce X-ray diffraction of DNA
  - James D. Watson and Francis H. C. Crick deduced the double helical structure of DNA (Race for the Double Helix movie)
- 1956
  - George Emil Palade showed the site of enzymes manufacturing in the cytoplasm is made on RNA organelles called ribosomes
- 1977
  - Phillip Sharp and Richard Roberts demonstrated that pre-mRNA is processed by the excision of introns and exons are spliced together



TIMELINE OF MOLECULAR BIOLOGY

- 1977
  - Phillip Sharp and Richard Roberts demonstrated that pre-mRNA is processed by the excision of introns and exons are spliced together
- 1986
  - Leroy Hood developed automated sequencing mechanism
- 1986
  - Human Genome Initiative announced



TIMELINE OF MOLECULAR BIOLOGY

- 1990
    - The 15 year Human Genome project is launched
- 1995
    - John Craig Venter: First bacterial genomes sequenced
    - Automated fluorescent sequencing instruments and robotic operations

TIMELINE OF MOLECULAR BIOLOGY

- 1996
  - First eukaryotic genome-yeast-sequenced
- 1997
  - E. Coli sequenced
- 1999
  - First human chromosome (22) sequenced

Nature. 1999 Dec 2;402(6761):489-95.

The DNA sequence of human chromosome 22.

Dunham I[1], Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JE, Bruskiewich R, Beare DM
Smink LJ, Ainscough R, Almeida JP, Babbage A, Bagguley C, Bailey J, Barlow K, Bates K
Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J, Burrill WD, O'Brien KP, et al.

Author information

Erratum in
Nature 2000 Apr 20;404(6780):904.

Abstract
Knowledge of the complete genomic DNA sequence of an organism allows a syster
to defining its genetic components. The genomic sequence provides access to the
structures of all genes, including those without known function, their control elemen
inference, the proteins they encode, as well as all other biologically important seque
Furthermore, the sequence is a rich and permanent source of information for the de
biological studies of the organism and for the study of evolution through cross-speci
comparison. The power of this approach has been amply demonstrated by the dete
the sequences of a number of microbial and model organisms. The next step is to o
complete sequence of the entire human genome. Here we report the sequence of tl
part of human chromosome 22. The sequence obtained consists of 12 contiguous s
spanning 33.4 megabases, contains at least 545 genes and 134 pseudogenes, and
first view of the complex chromosomal landscapes that will be found in the rest of th

Comment in
Do we need a huge new centre to annotate the human genome? [Nature. 2000]
Tiny chromosome is rich in genes and medical promise. [Nature. 1999]
'Finishing' success marks major genome sequencing milestone...as researchers po
data. [Nature. 1999]
The book of genes. [Nature. 1999]

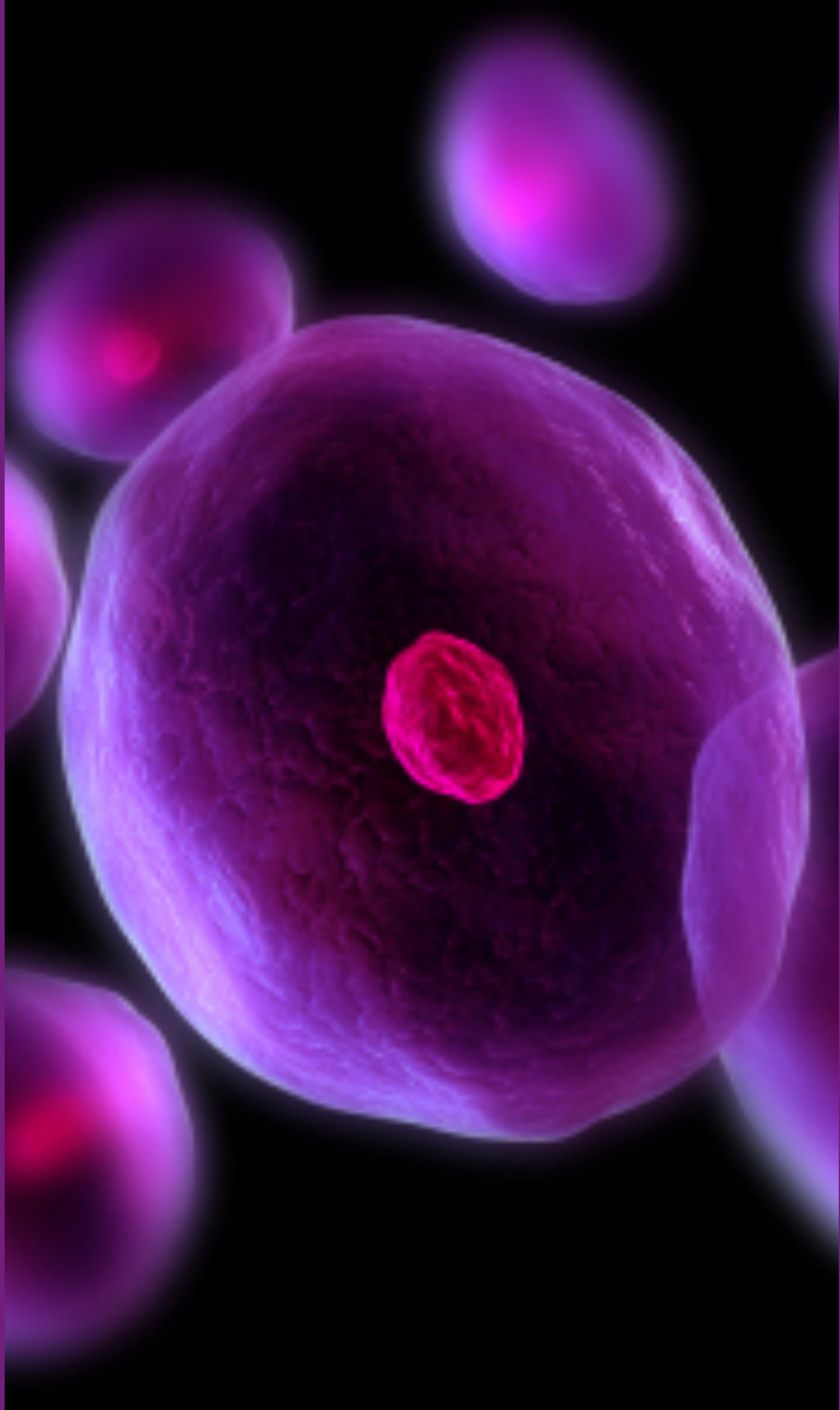
TIMELINE OF MOLECULAR BIOLOGY

- 2000
  - Complete sequence of the Drosophila melanogaster genome
- 2001
  - International Human Genome Sequencing publishes first draft of the sequence of the human genome
- 2003
  - Human Genome Project "completed"

Cells

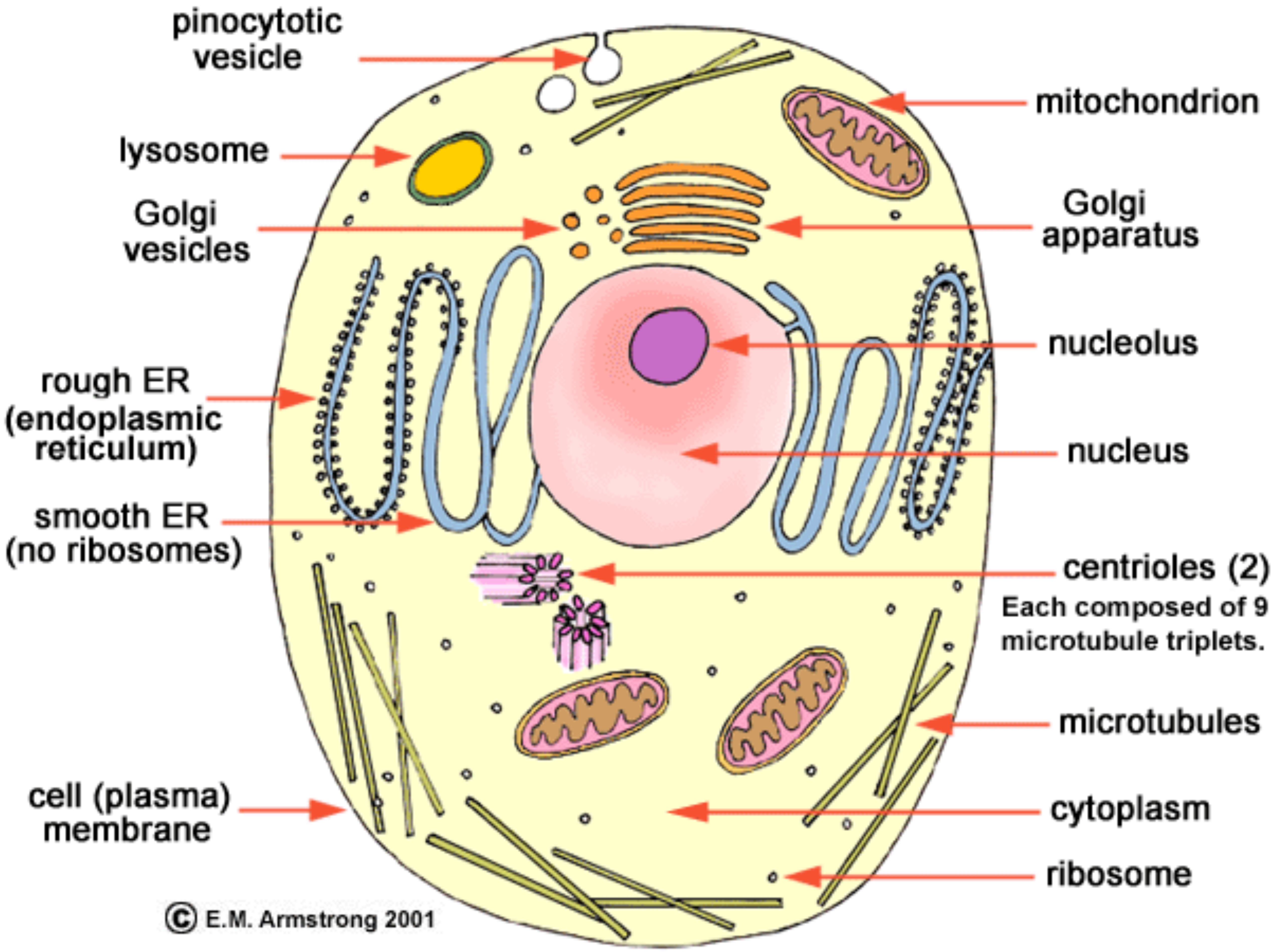
CELLS

- Cell Theory
  - All living organisms are composed of one or more cells
  - The cell is the most basic unit of life
  - All cells arise from pre-existing, living cells
  - Cells contain heredity information that is passed during cell division

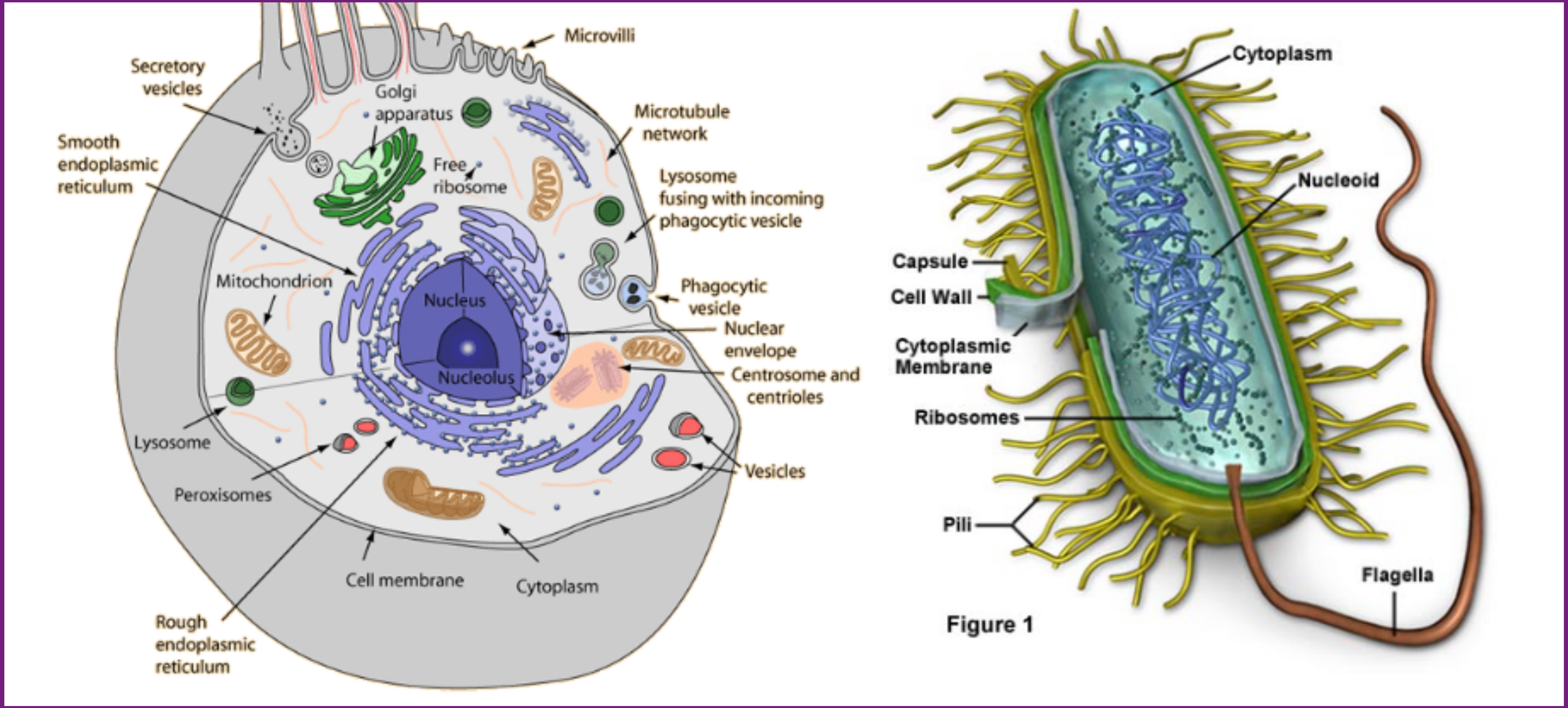


CELLS

- Organisms can be of single cells or multiple cells
  - Most living organisms are single cells (e.g. E. coli, yeast)
  - Multicellular organisms have trillions of cells
- Cells have many parts
  - Organelles - specialized structures that perform certain tasks within the cell



CELLS

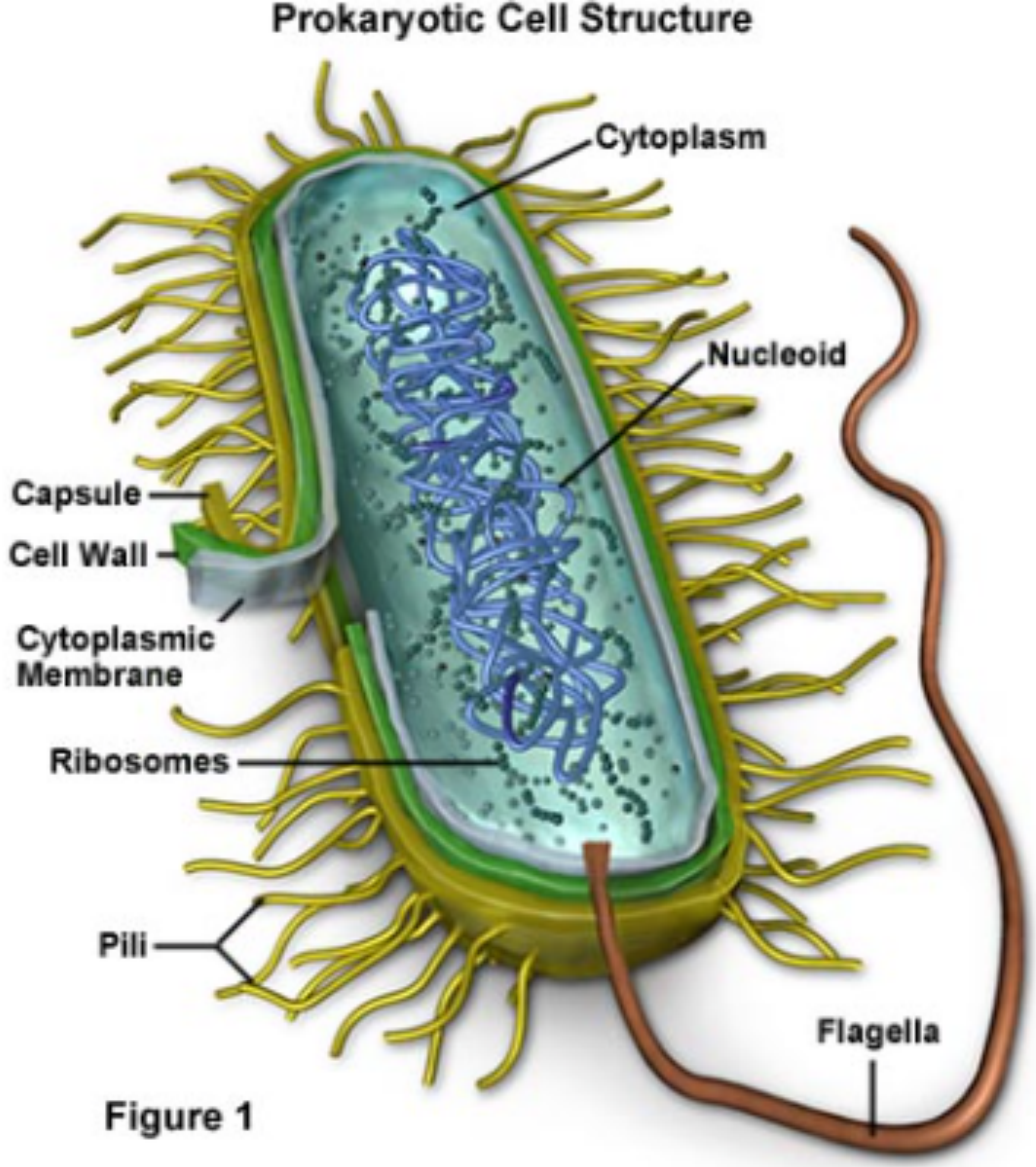


- Organisms classified by type of cells: prokaryotic, eukaryotic

CELLS

PROKARYOTES

- Lack a cell nucleus
- Lack membrane bound organelles
- DNA in a single loop
- Mostly single-cellular (some multi-cellular)
- Two groups
  - Bacteria
  - Archaea (extremophiles)

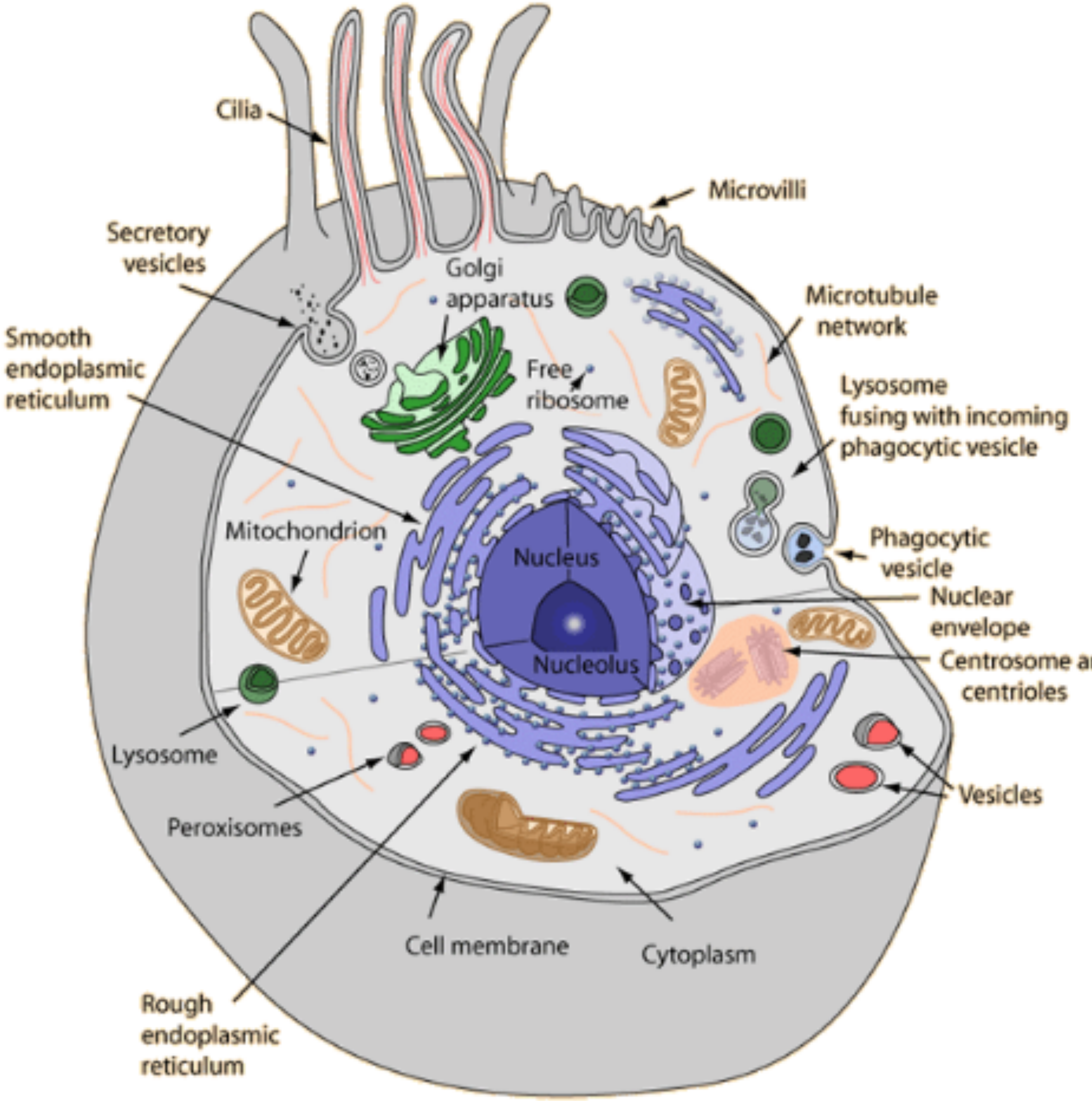


Prokaryotic cell structure

CELLS

EUKARYOTES

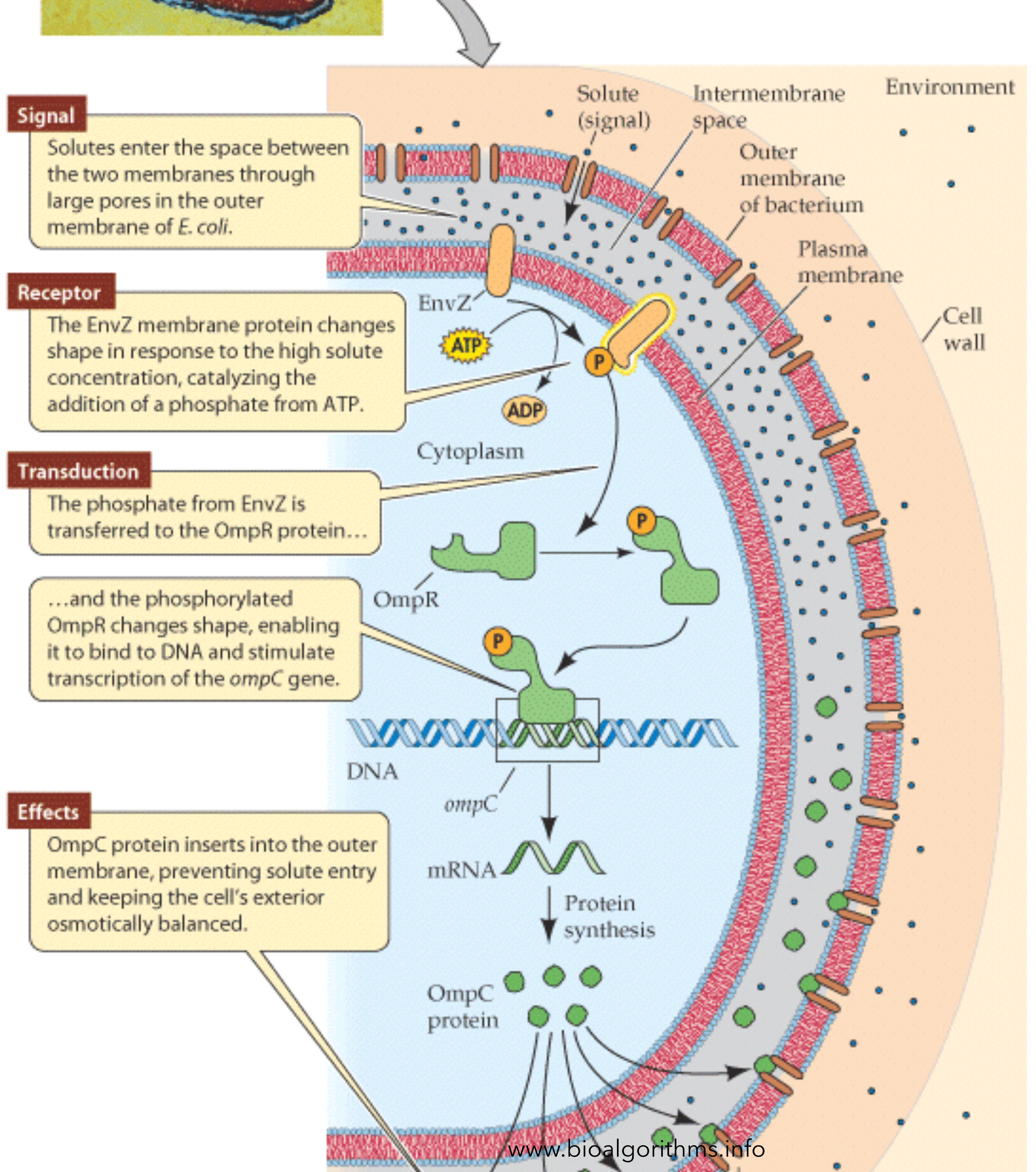
- Cells posses membrane-bound organelles
- Genetic material contained in nucleus
- DNA organized into chromosomes
- Can be single or multi-cellular
- Examples:
  - Animals, plants, fungi



http://hyperphysics.phy-astr.gsu.edu/hbase/biology/imgbio/cellhlabel.gif

CELLS

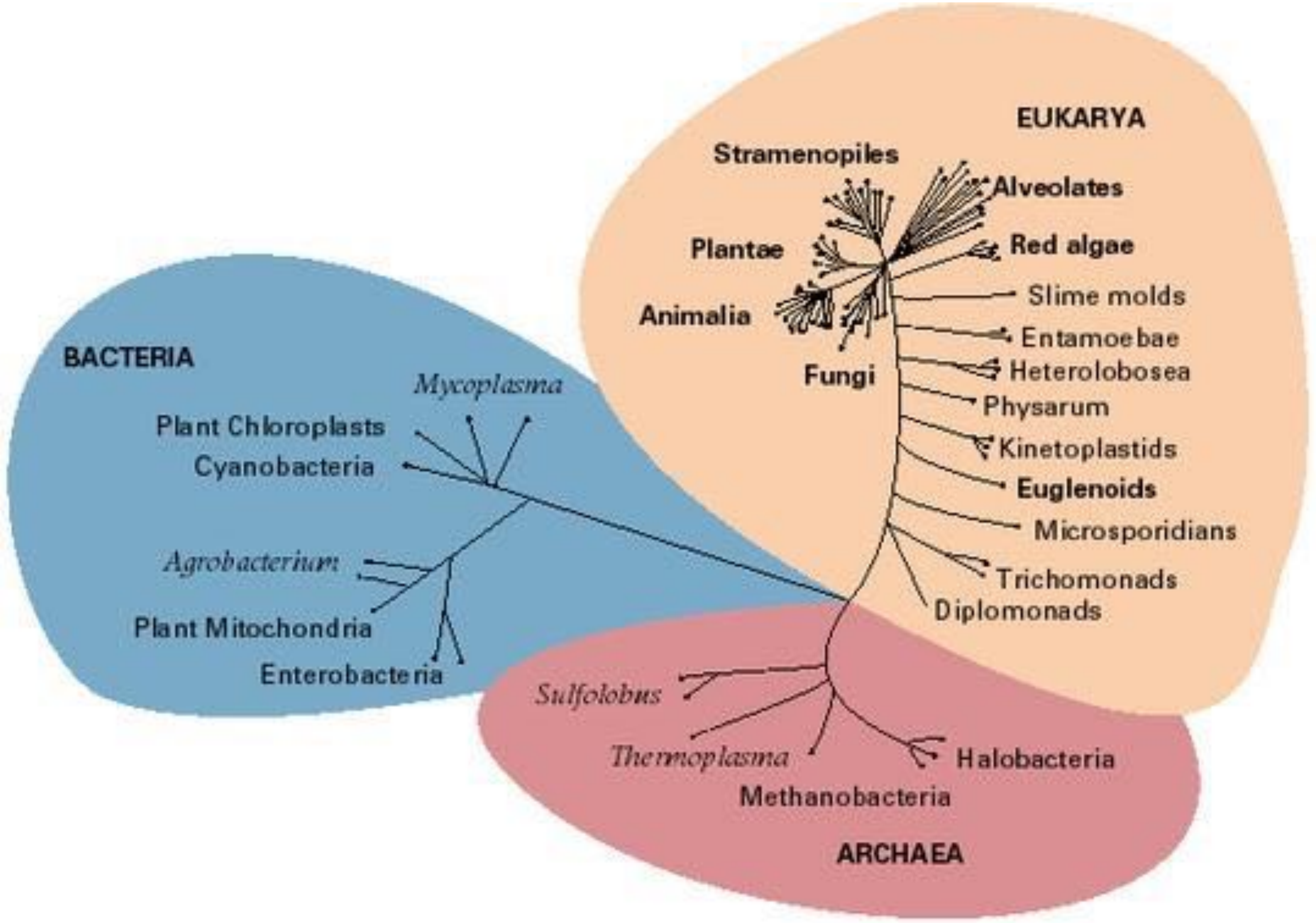
- Cells make decision through complex networks of chemical reactions, called pathways
    - Synthesize new materials
    - Break other materials down for spare parts
    - Signal to eat or die



CELLS

EVOLVING MODEL AS WE LEARN MORE (NEW ORGANISMS, SEQUENCES)

- According to the most recent evidence, there are three main branches to the tree of life



CELLS

- Global oceanic sampling expedition

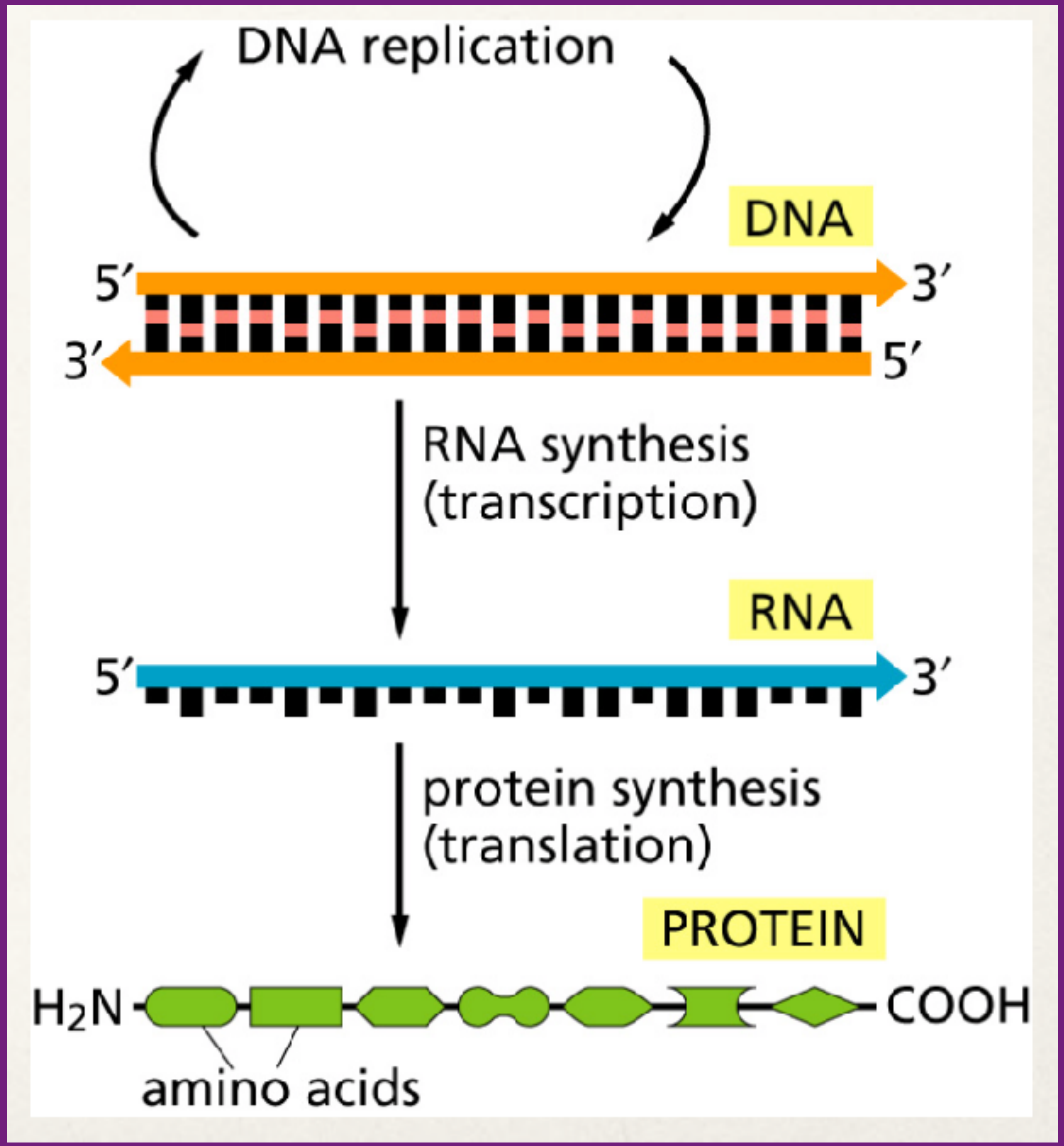
Central Dogma of Molecular Biology

CENTRAL DOGMA OF MOLECULAR BIOLOGY

DNA CAN REPLICATE

INFORMATION IN DNA IS PASSED TO RNA

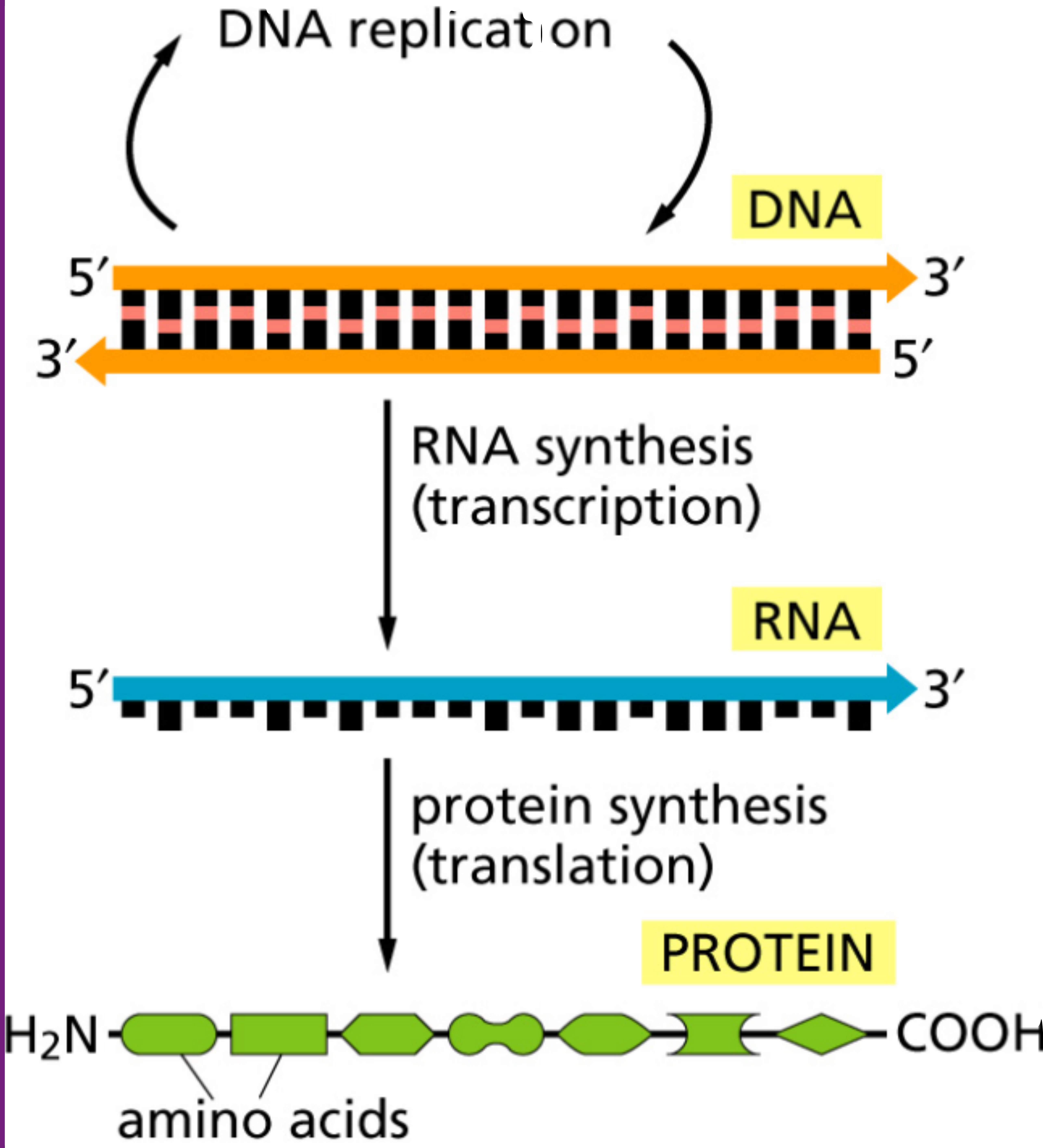
INFORMATION IN RNA IS PASSED TO PROTEINS



- Describes the flow of information in cells

CENTRAL DOGMA OF MOLECULAR BIOLOGY

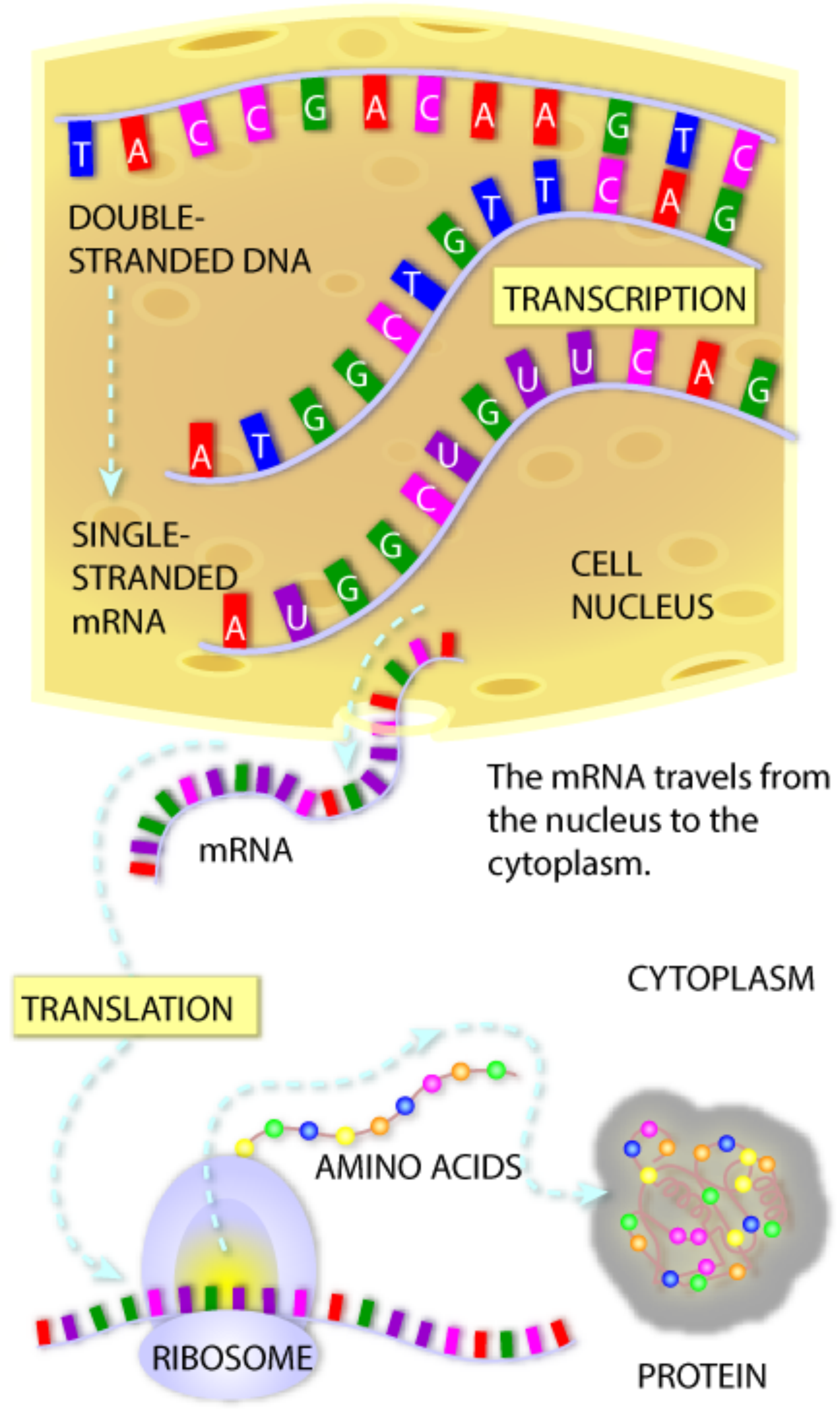
- All forms of life follow this scheme
  - Slight variations between organisms
- Genomic DNA encodes all the molecules necessary for life of an organism



CENTRAL DOGMA OF MOLECULAR BIOLOGY

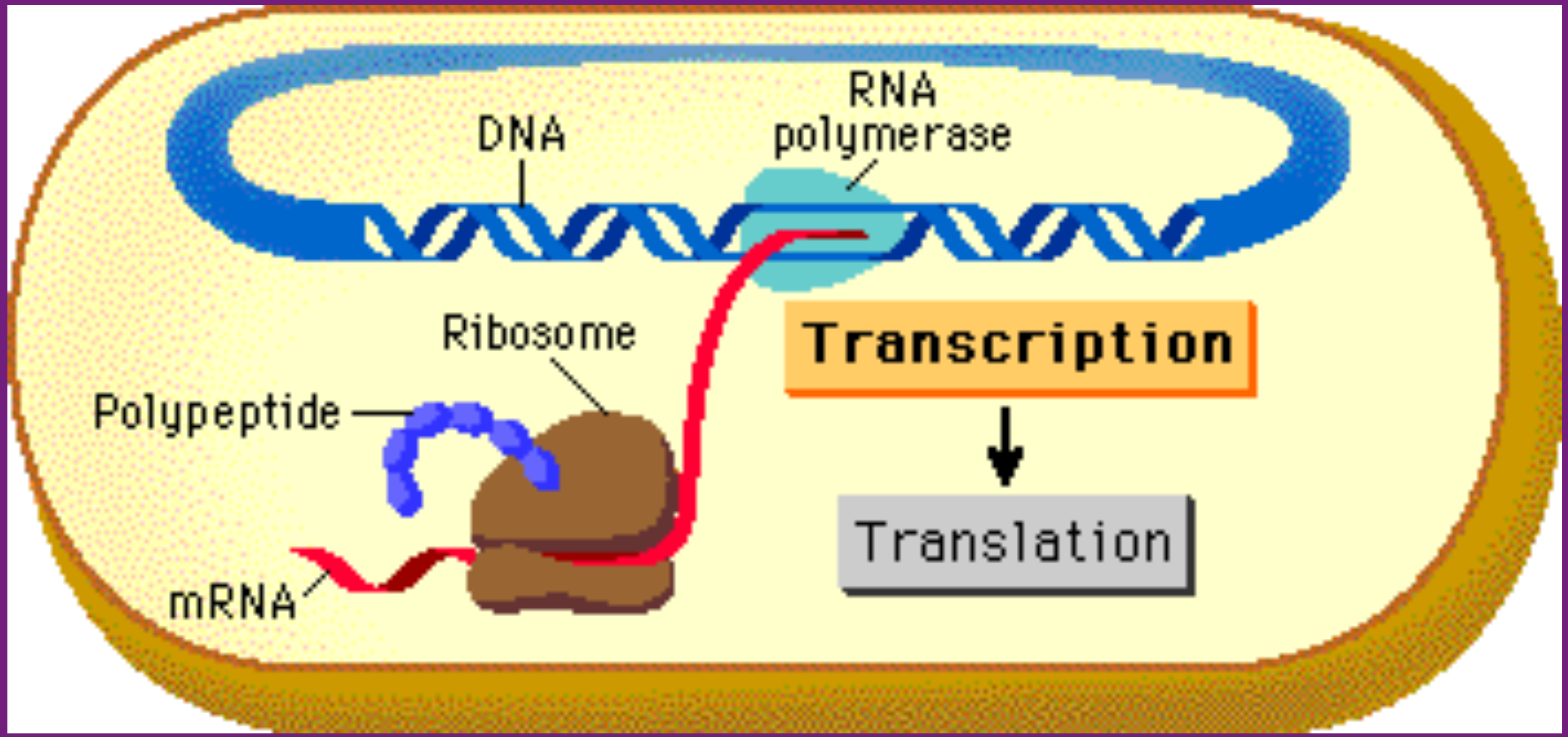
CENTRAL DOGMA OF MOLECULAR BIOLOGY

- Transcription
  - Synthesis of an RNA copy of a segment of DNA
  - RNA is synthesized by the enzyme RNA polymerase
- Translation
  - Ribosome reads the mRNA sequence
  - Translates to amino acid sequence of the protein



http://learn.genetics.utah.edu/content/molecules/transcribe/

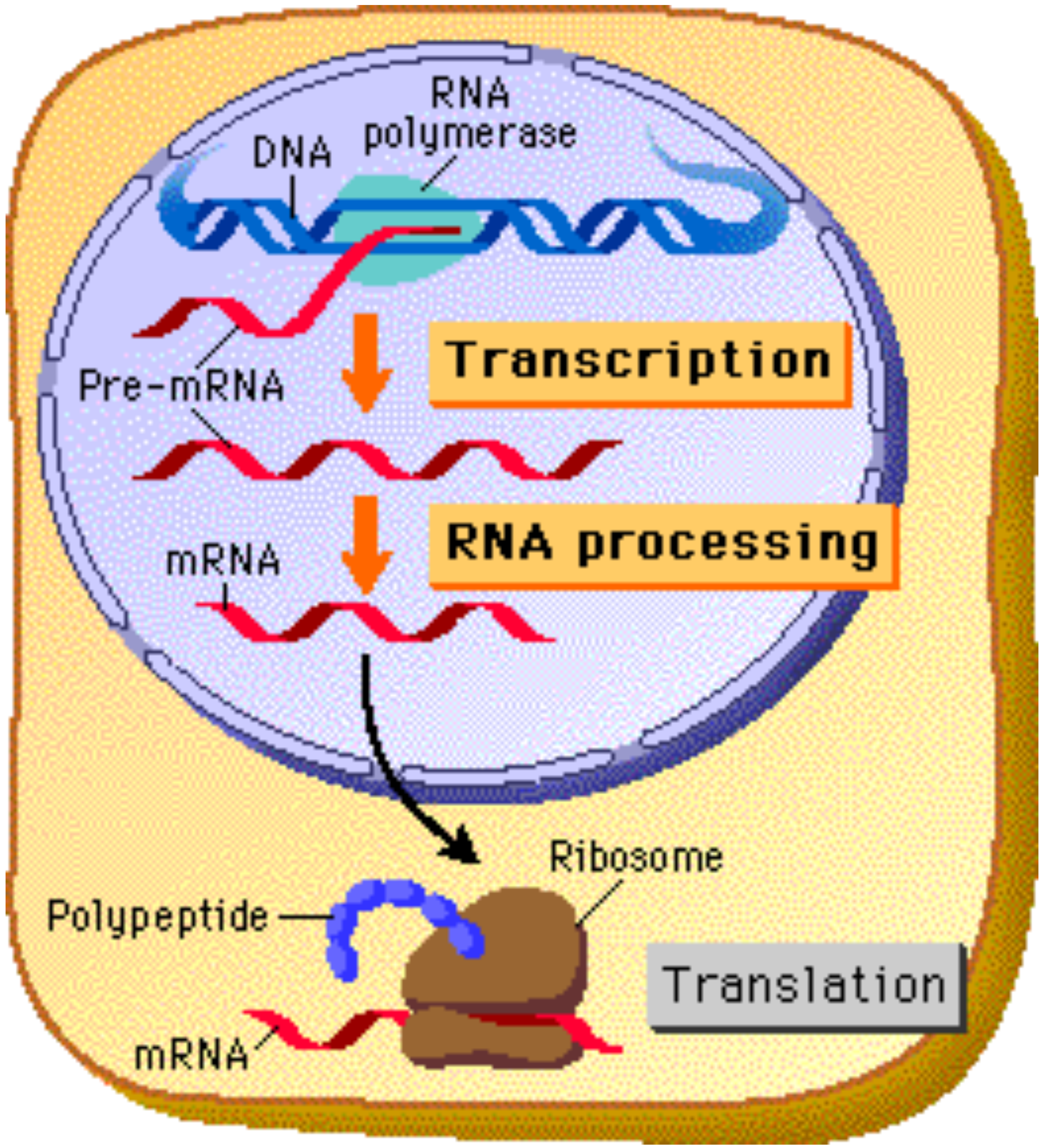
CENTRAL DOGMA OF MOLECULAR BIOLOGY



- Prokaryotic cell
  - Transcripts are immediately translated (no nucleus to cross)

CENTRAL DOGMA OF MOLECULAR BIOLOGY

- Eukaryotic cell
  - Transcription occurs in nucleus
  - Pre-mRNA produced
  - RNA processing
  - Mature mRNA exits
  - Translated in the cytoplasm



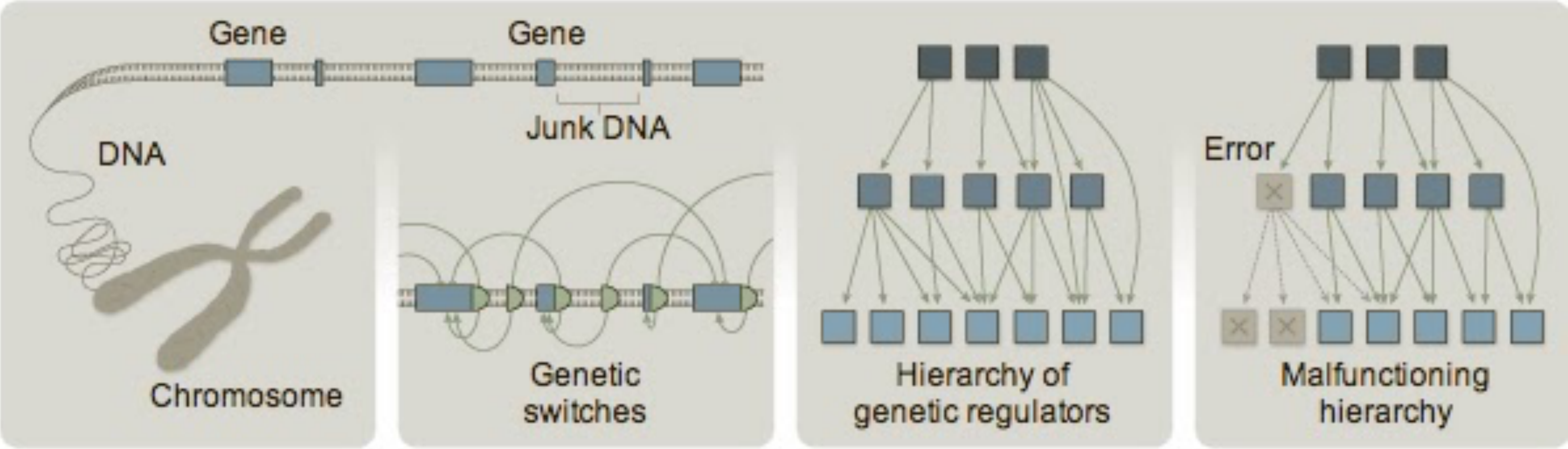
CENTRAL DOGMA OF MOLECULAR BIOLOGY

- Not all genetic information in DNA encodes proteins
- Non-coding "junk" DNA
  - 98% is non-coding in human genome
  - Play other roles
    - Regulation
    - Transcription factor site
    - Operators, promoters
    - Undiscovered functionality

CENTRAL DOGMA OF MOLECULAR BIOLOGY

Rethinking 'Junk' DNA

A large group of scientists has found that so-called junk DNA, which makes up most of the human genome, does much more than previously thought. Related Article »

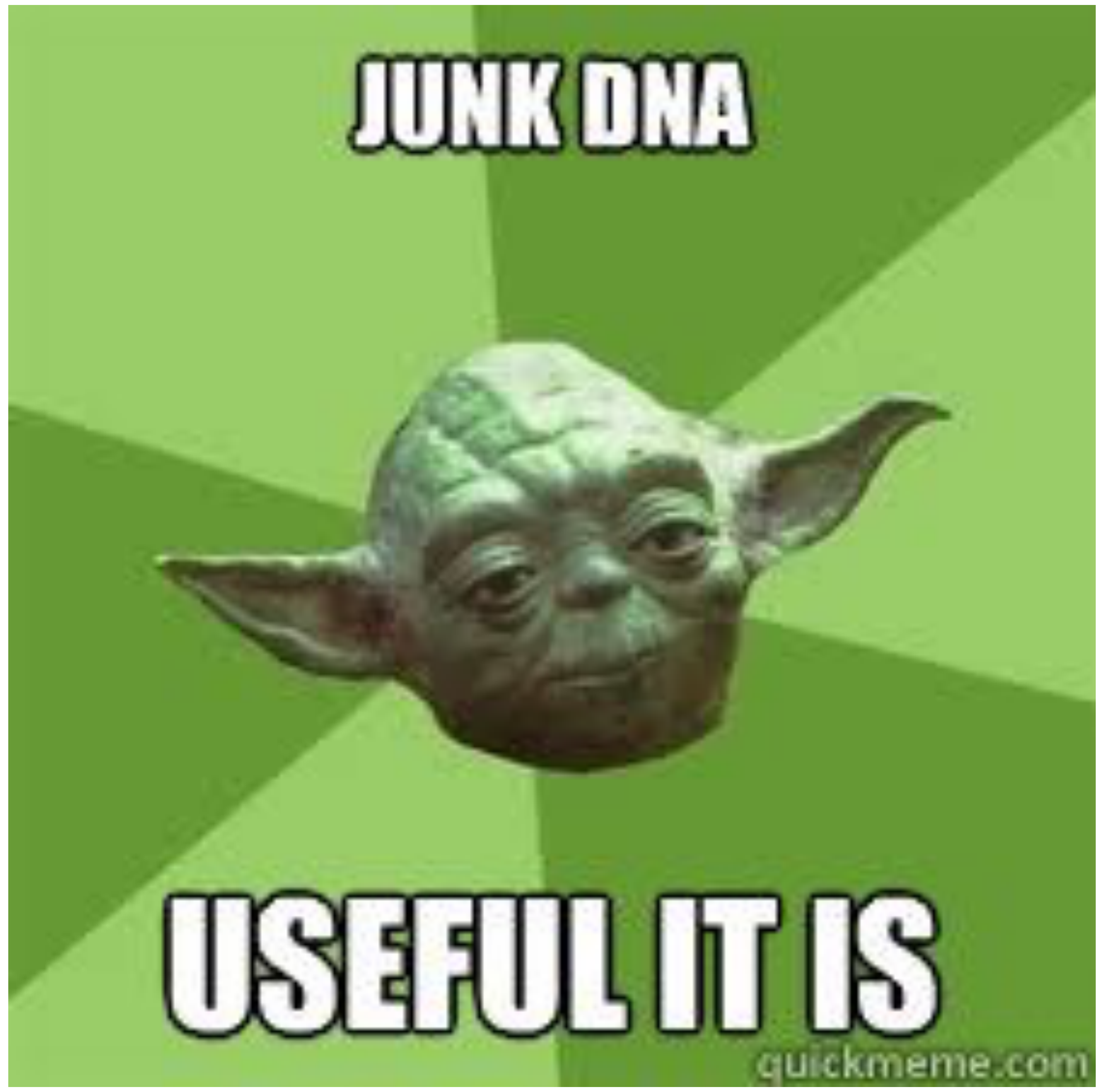


GENES
Each human cell contains about 10 feet of DNA, coiled into a dense tangle. But only a very small percentage of DNA encodes genes, which control inherited traits like eye color, blood type and so on.

JUNK DNA
Stretches of DNA around and between genes seemed to do nothing, and were called junk DNA. But now researchers think that the junk DNA contains a large number of tiny genetic switches, controlling how genes function within the cell.

REGULATION
The many genetic regulators seem to be arranged in a complex and redundant hierarchy. Scientists are only beginning to map and understand this network, which regulates how cells, organs and tissues behave.

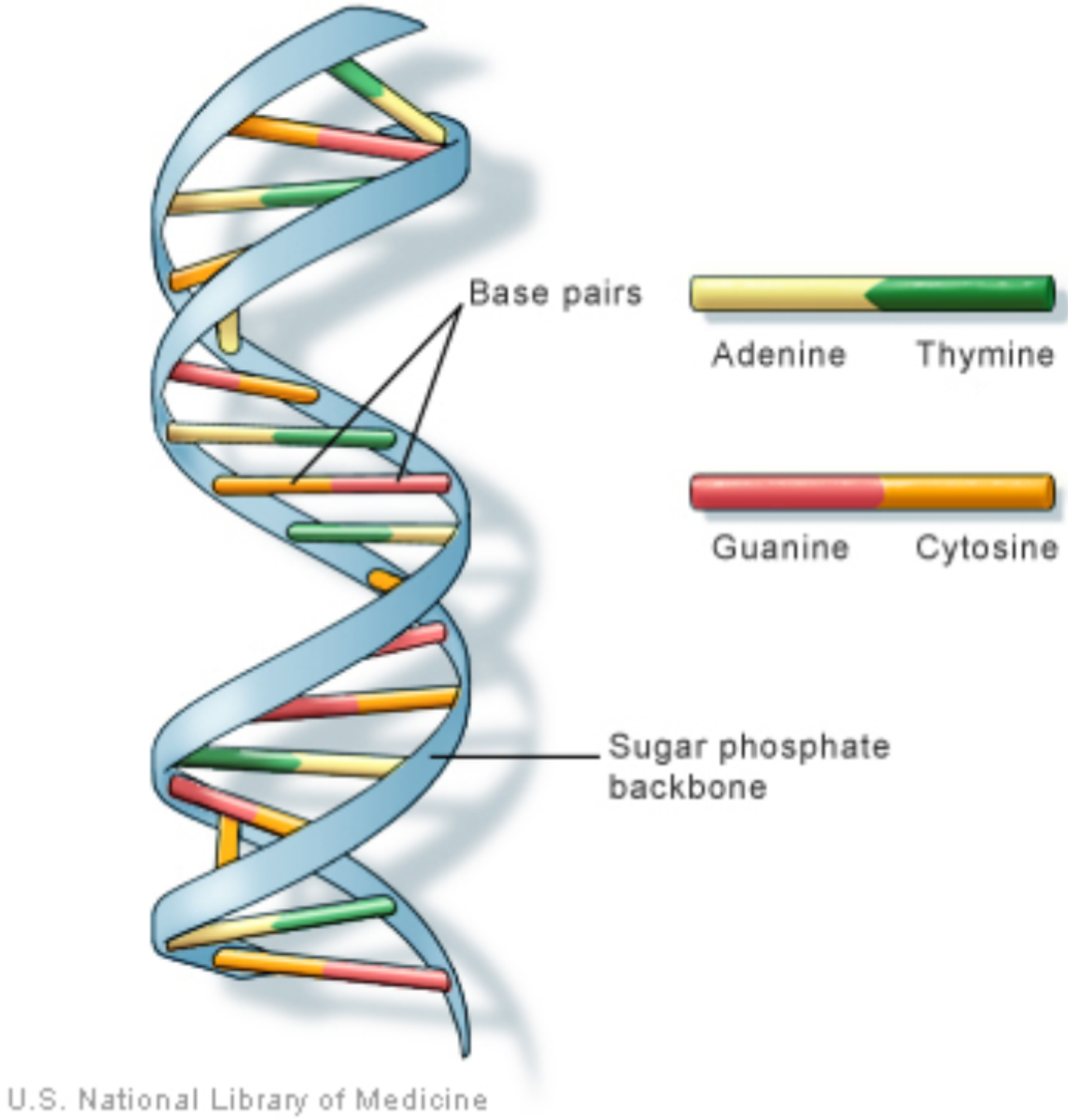
DISEASE
Errors or mutations in genetic switches can disrupt the network and lead to a range of diseases. The new findings will spur further research and may lead to new drugs and treatments.



DNA

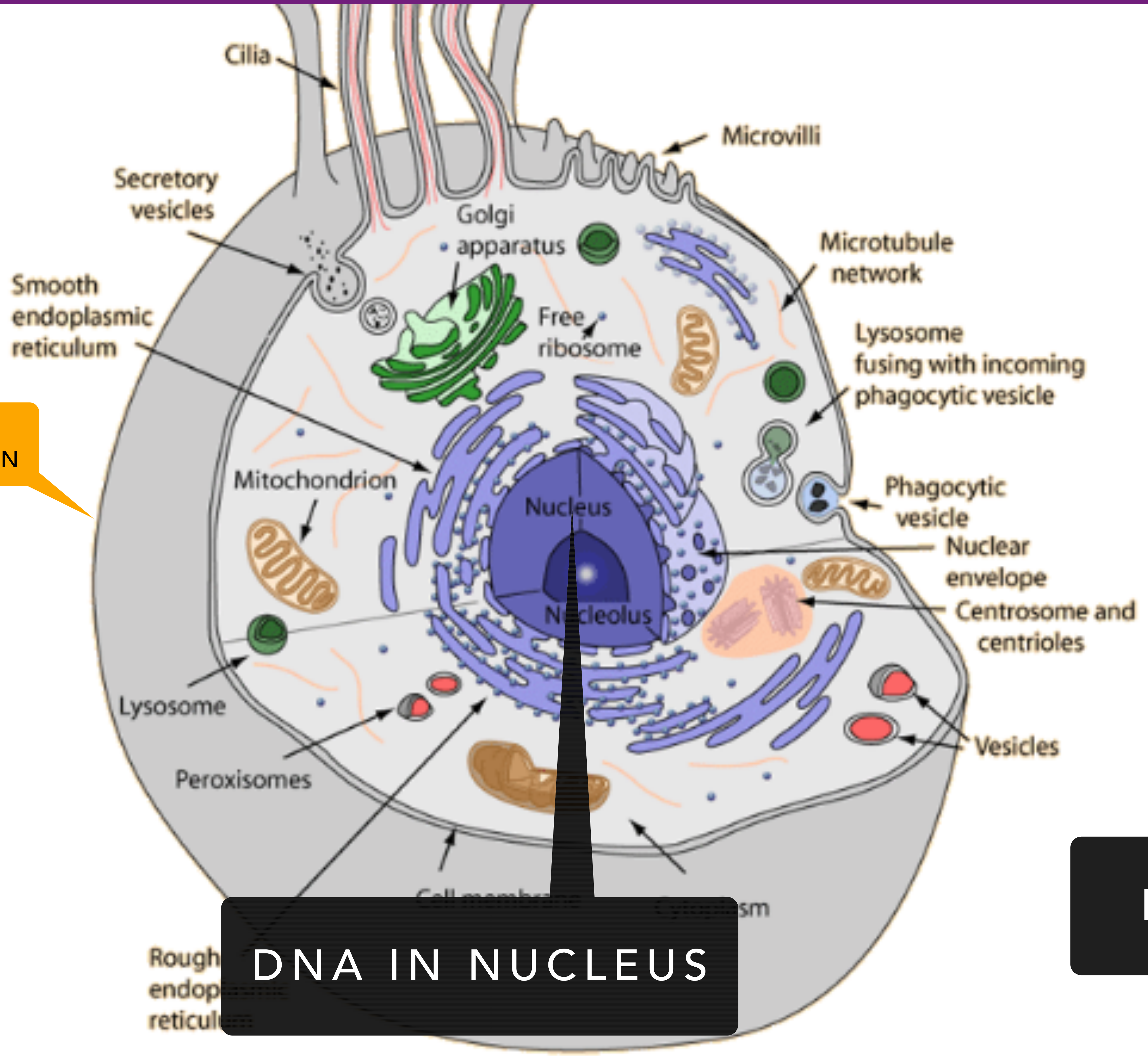
DNA

- DNA (**d**eoxyribo**n**ucleic **a**cid)
  - Hereditary material in organisms
- Nearly every cell in a person's body has the same DNA



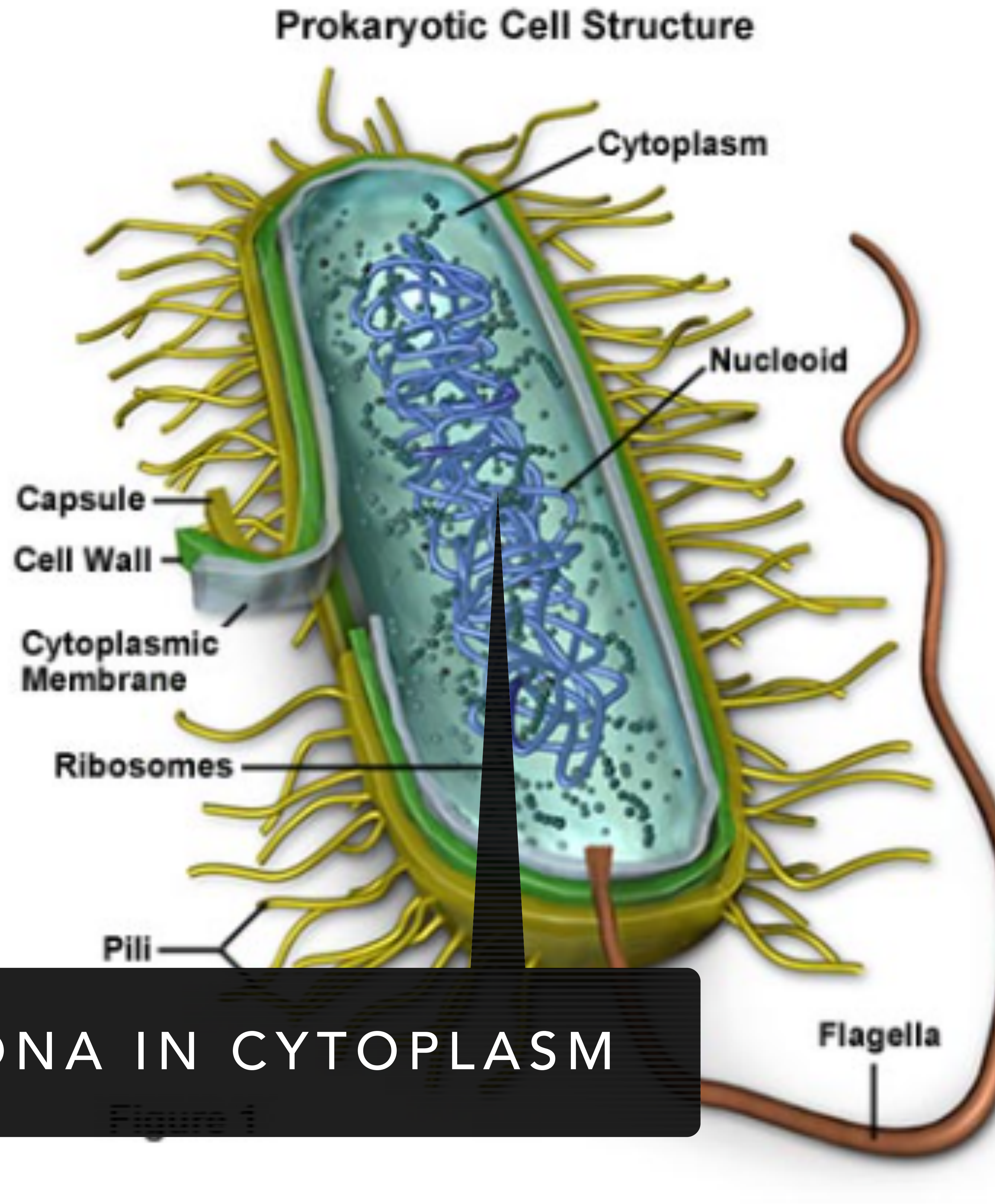
DNA

DNA IN MITOCHONDRION



DNA IN NUCLEUS

Animal cell structure

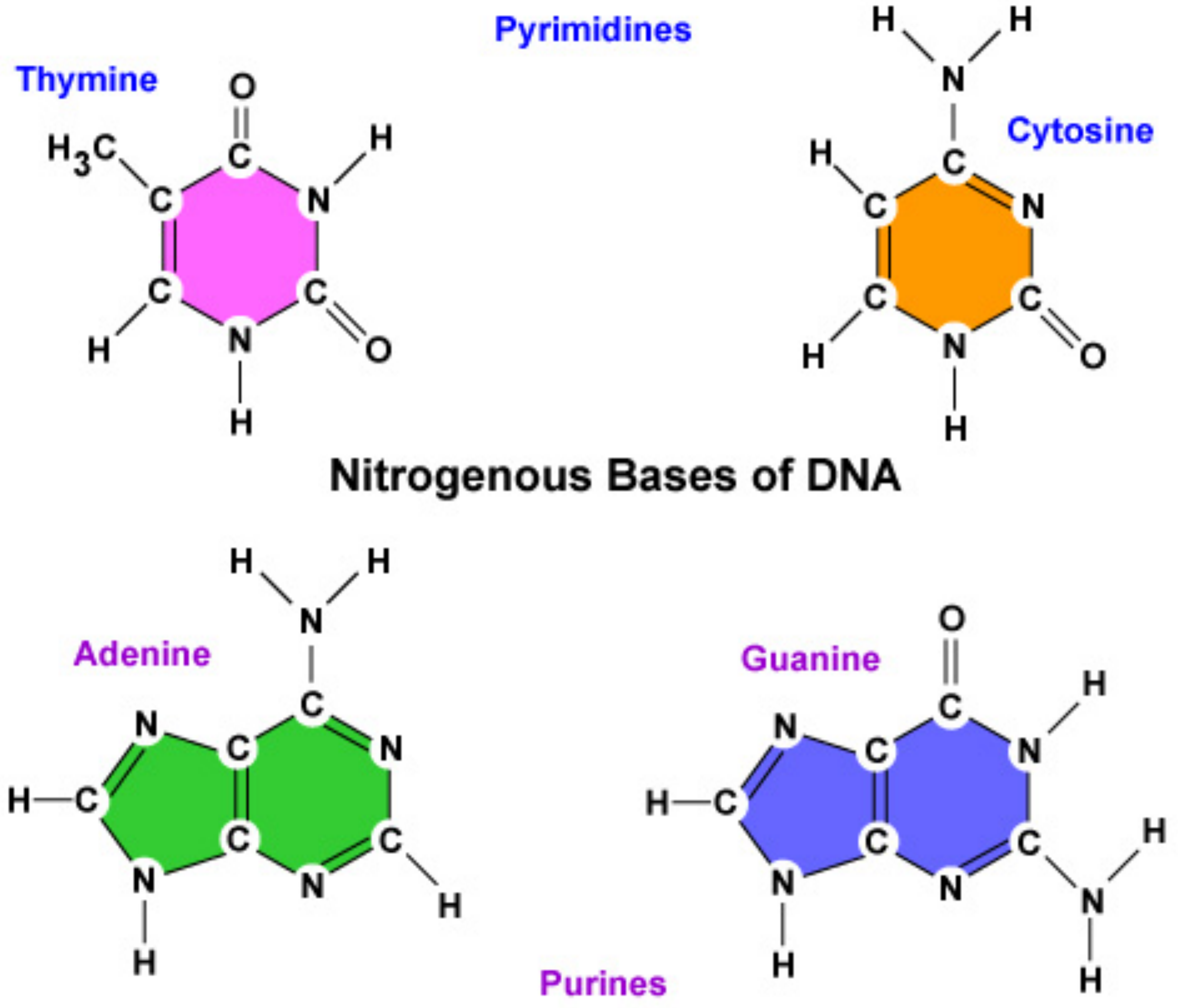


DNA IN CYTOPLASM

Prokaryotic cell structure

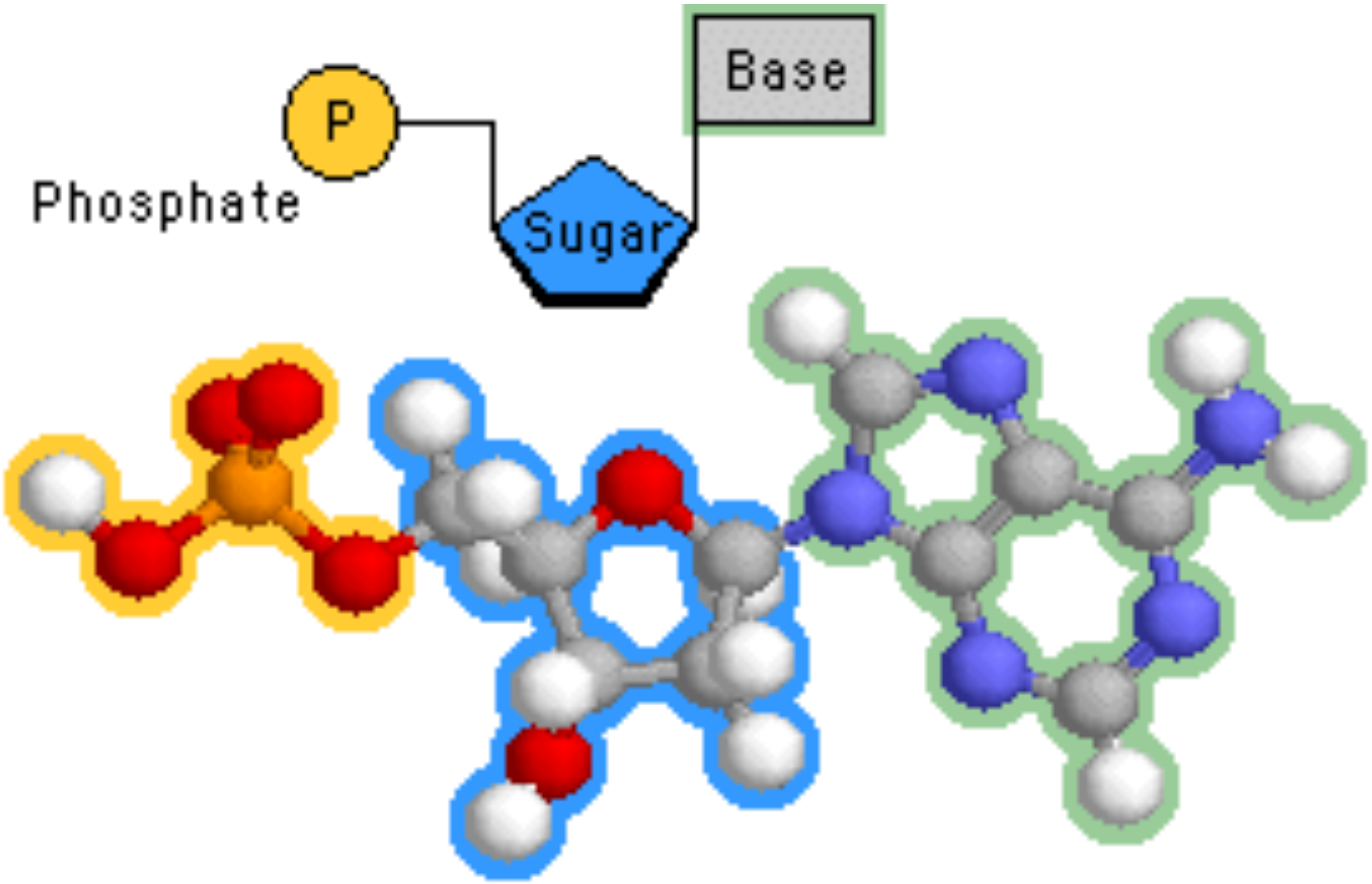
DNA

- The information in DNA is stored as a code made up of four chemical bases
  - adenine (A)
  - guanine (G)
  - cytosine (C)
  - thymine (T)

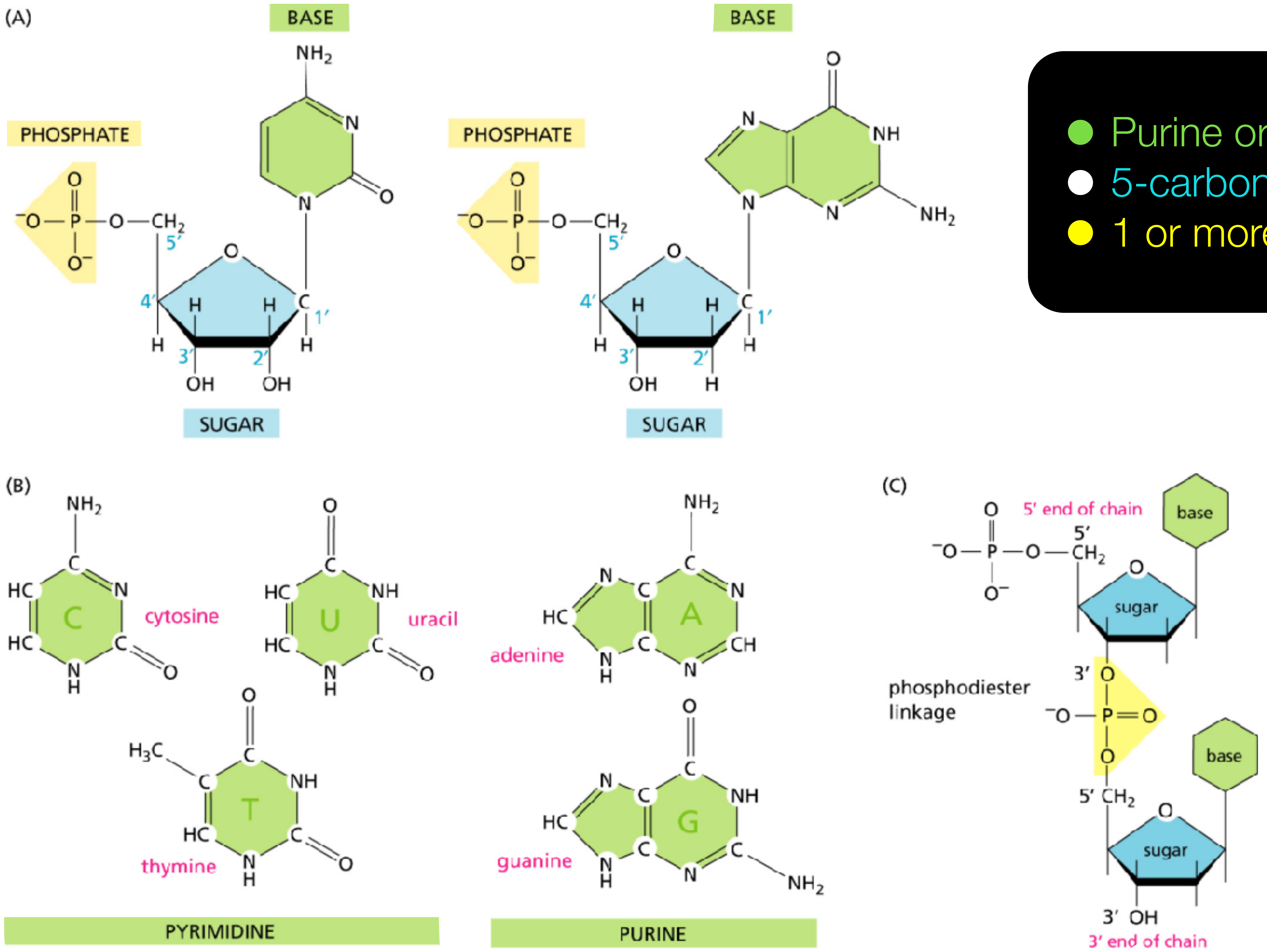


DNA

- Each nucleotide has three components
  - 1 or more phosphate groups
  - 5-carbon, or pentose, sugar
  - Nucleotide



DNA

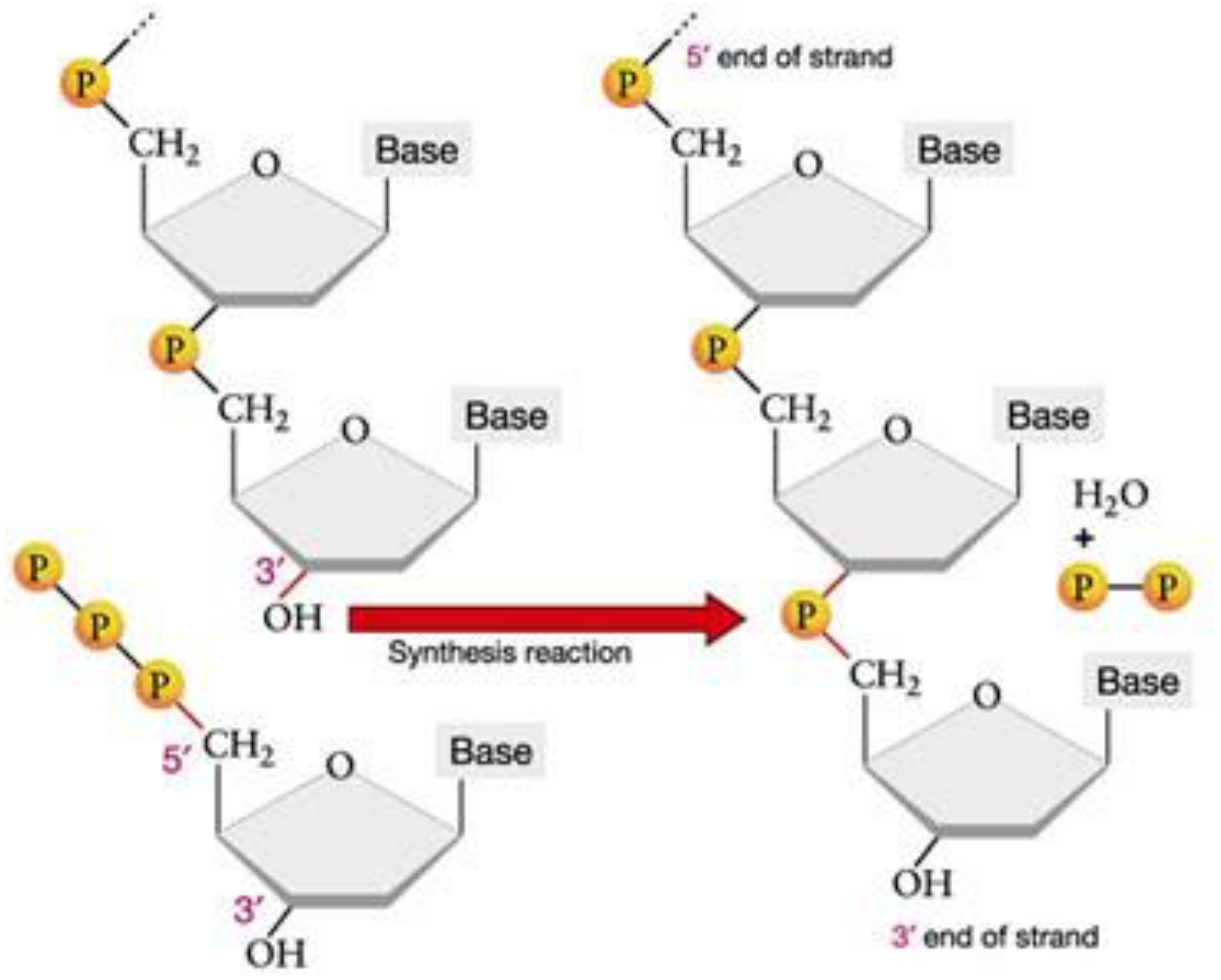


- Purine or pyrimidine bases
- 5-carbon (pentose) sugar
- 1 or more phosphate groups

POLYMERIZATION INTO CHAIN

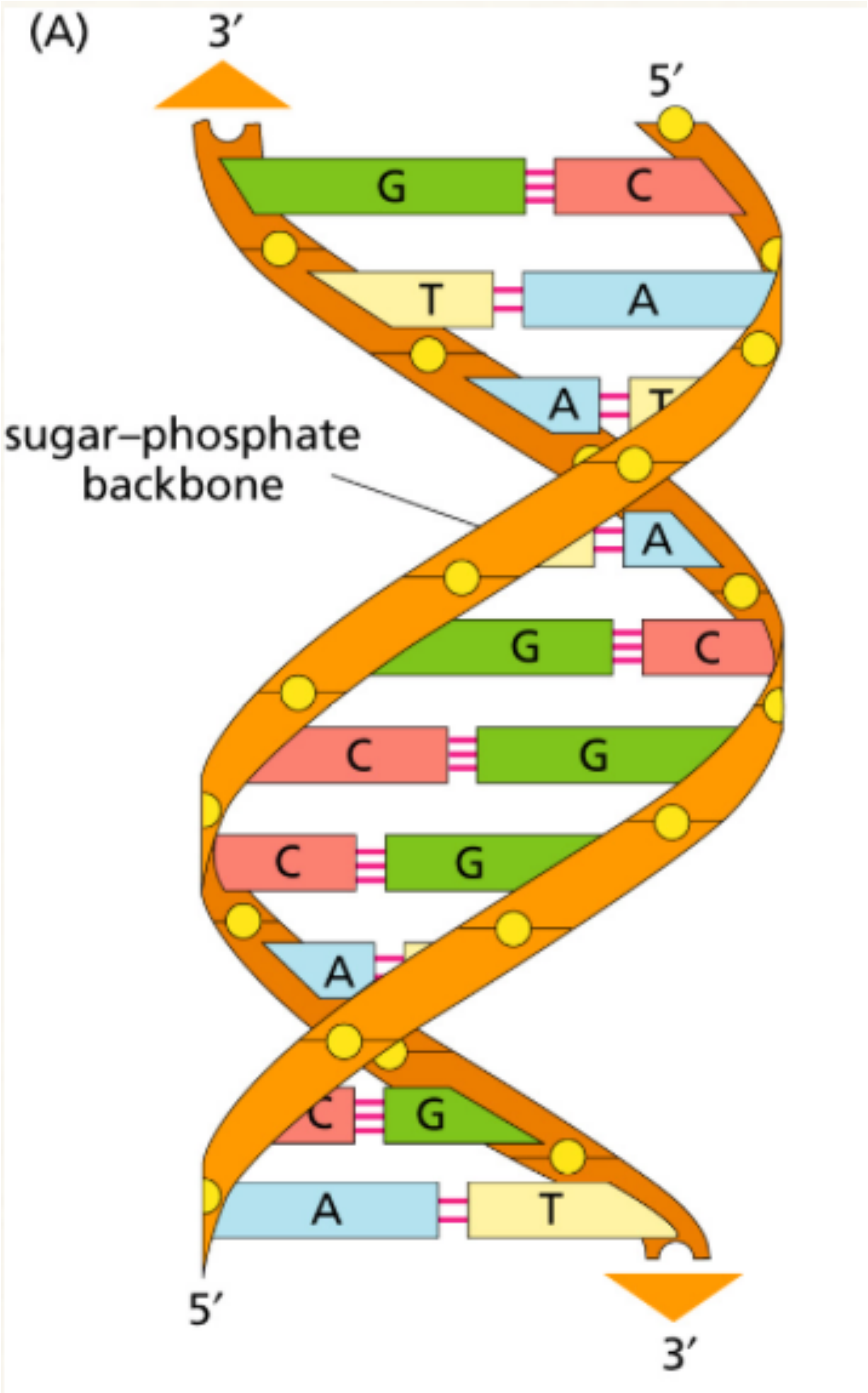
DNA

- Nucleic acids are built by polymerizing nucleotides
  - Millions of bases
  - Sequence encodes information



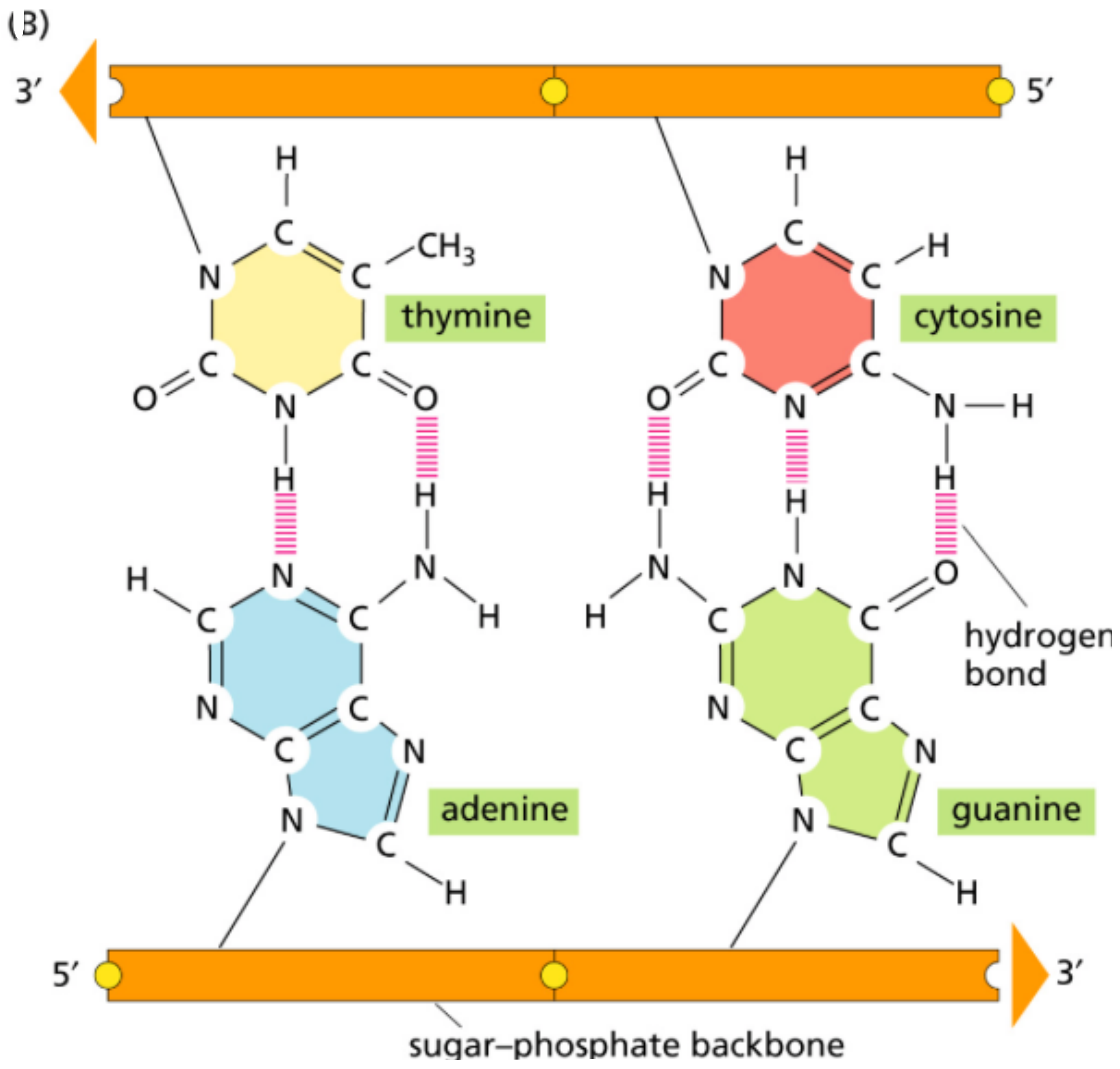
DNA

- DNA has a double helix structure
  - The backbone consists of alternating deoxyribose and phosphate groups
  - Each strand has a base sequence that is complementary to its partner strand



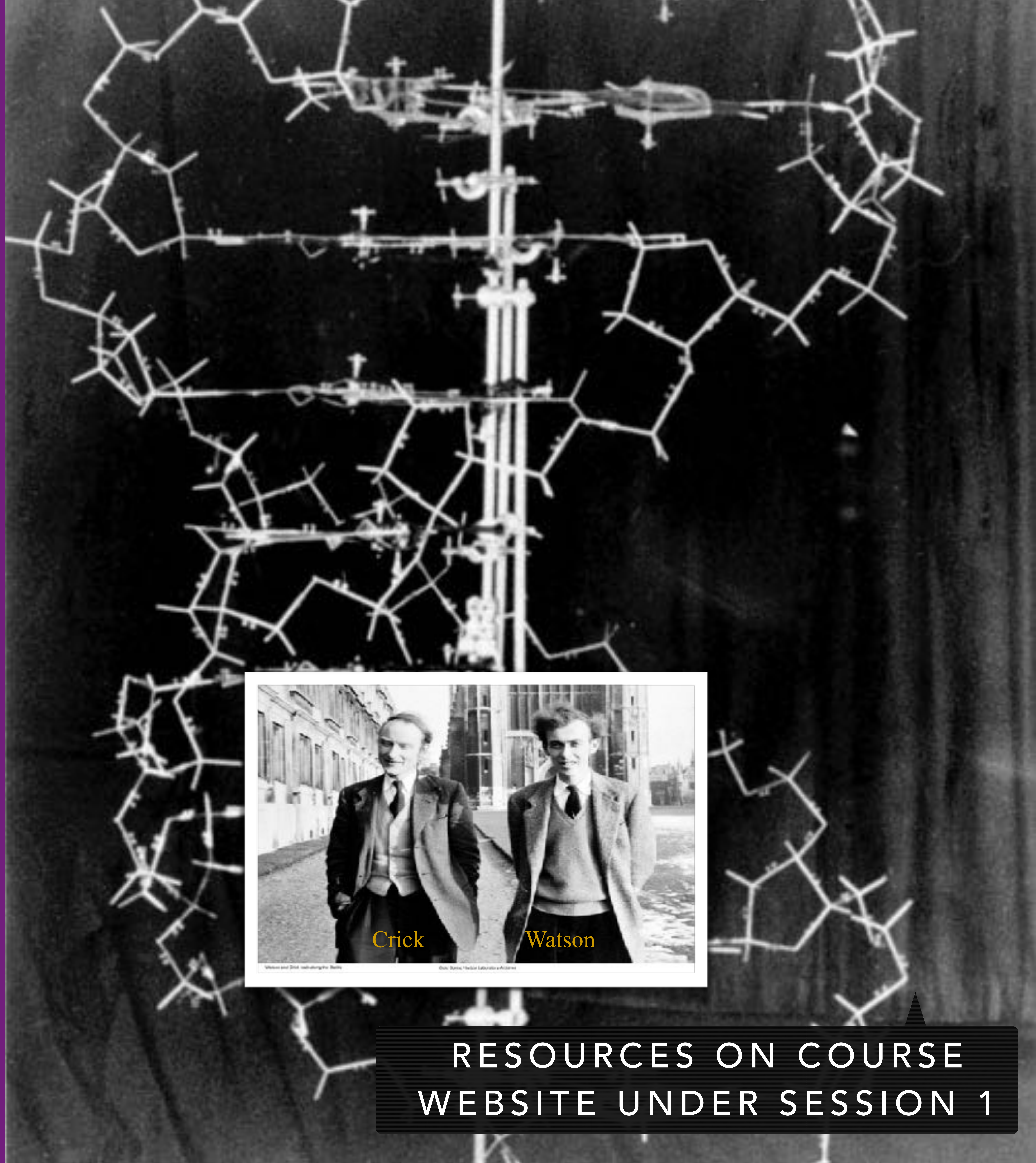
DNA

- Watson-Crick base-pairing
  - A will only base-pair with T
  - C will only base-pair with G
- Base-pairs
  - A and T contain two H-bonds
  - G and C contain three H-bonds (more stable that AT)



DISCOVERY OF DNA

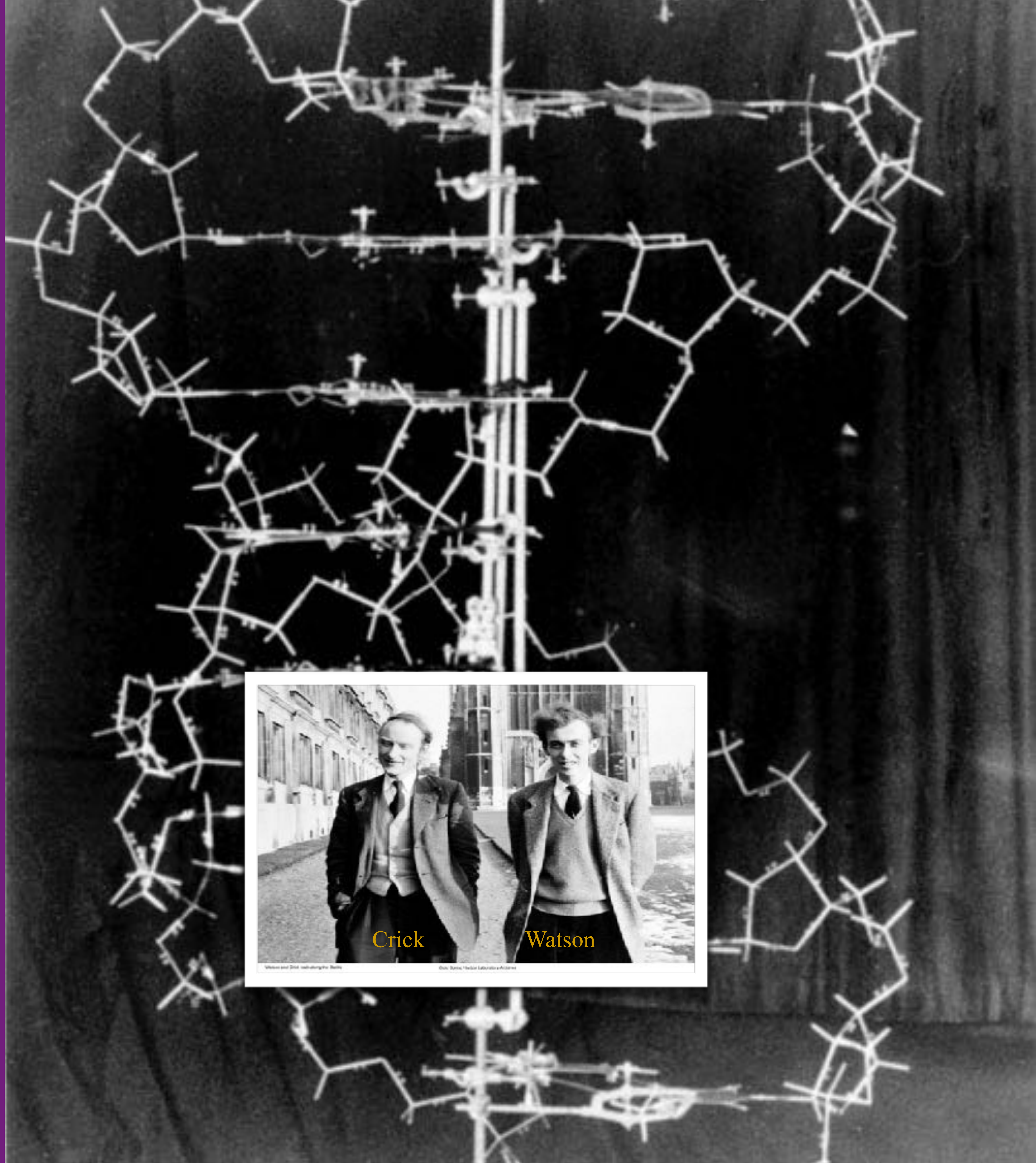
- DNA Sequences
  - Chargaff and Vischer, 1949
    - DNA consisting of A, T, G, C
  - Chargaff Rule (#A≈#T and #G≈#C)
    - A "strange but possibly meaningless" phenomenon.



RESOURCES ON COURSE WEBSITE UNDER SESSION 1

DISCOVERY OF DNA

- DNA Structure
  - Watson and Crick, Nature, April 25, 1953
  - Rich, 1973l Structural biologist at MIT
    - DNA’s structure in atomic resolution



DNA

- The players
  - Watson: a zoologist
  - Crick: a physicist “In 1947 Crick knew no biology and practically no organic chemistry or crystallography”
  - Rosalind Franklin - Xray crystallographer
  - Watkins - Showed an X-ray to Watson
- The Discovery
  - Applying Chagraff’s rules and the X-ray image from Rosalind Franklin, they constructed a “tinkertoy” model showing the double helix
- 1962 Nobel Prize for Watson, Crick, Wilkins
  - Rosalind Franklin died in 1958

Watson & Crick with DNA model

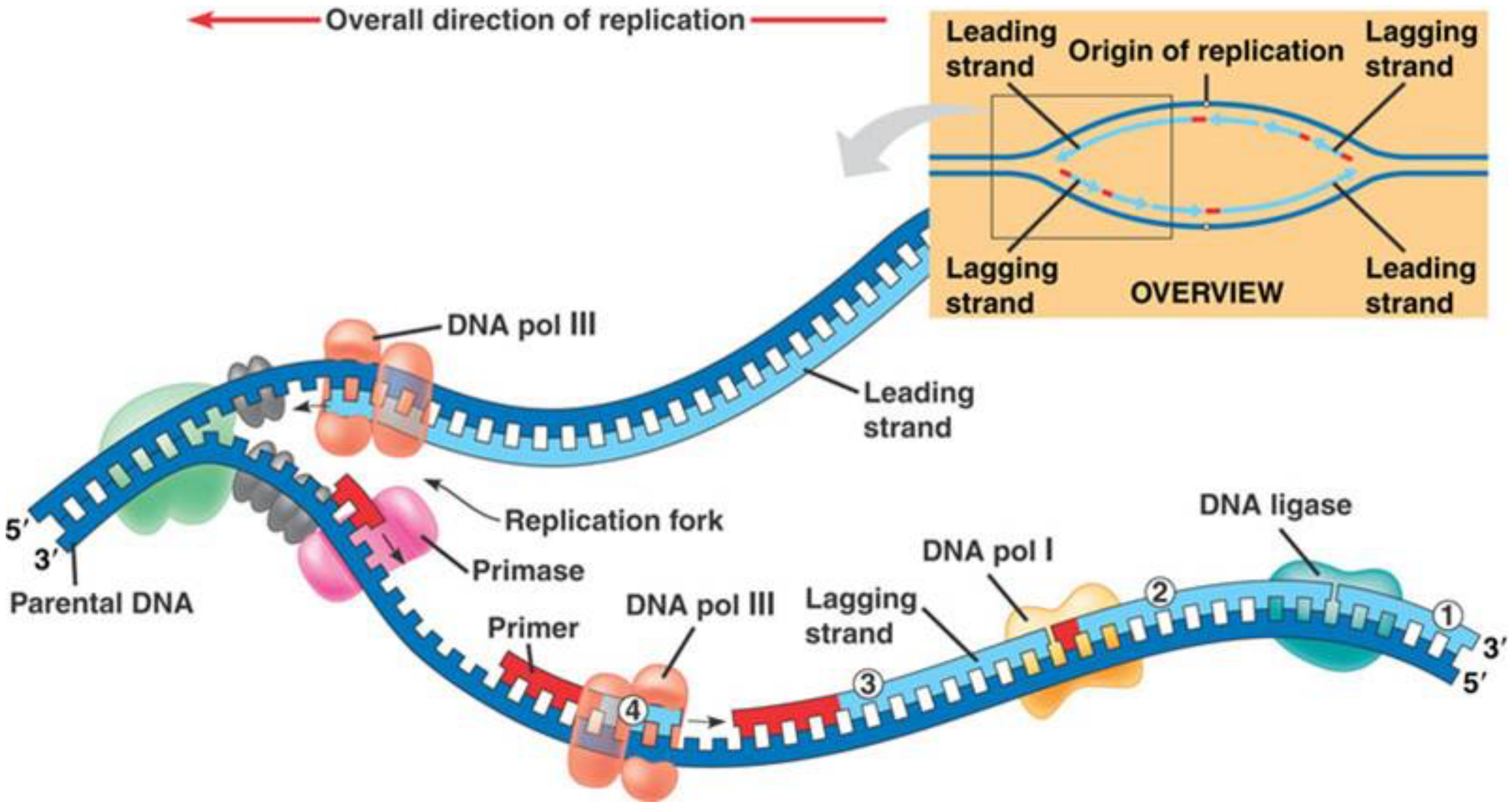
Rosalind Franklin with X-ray image of DNA

DNA

- Life Story: The Race for the Double Helix

DNA

- DNA Replication
  - Process of producing two identical replicas from one original DNA molecule

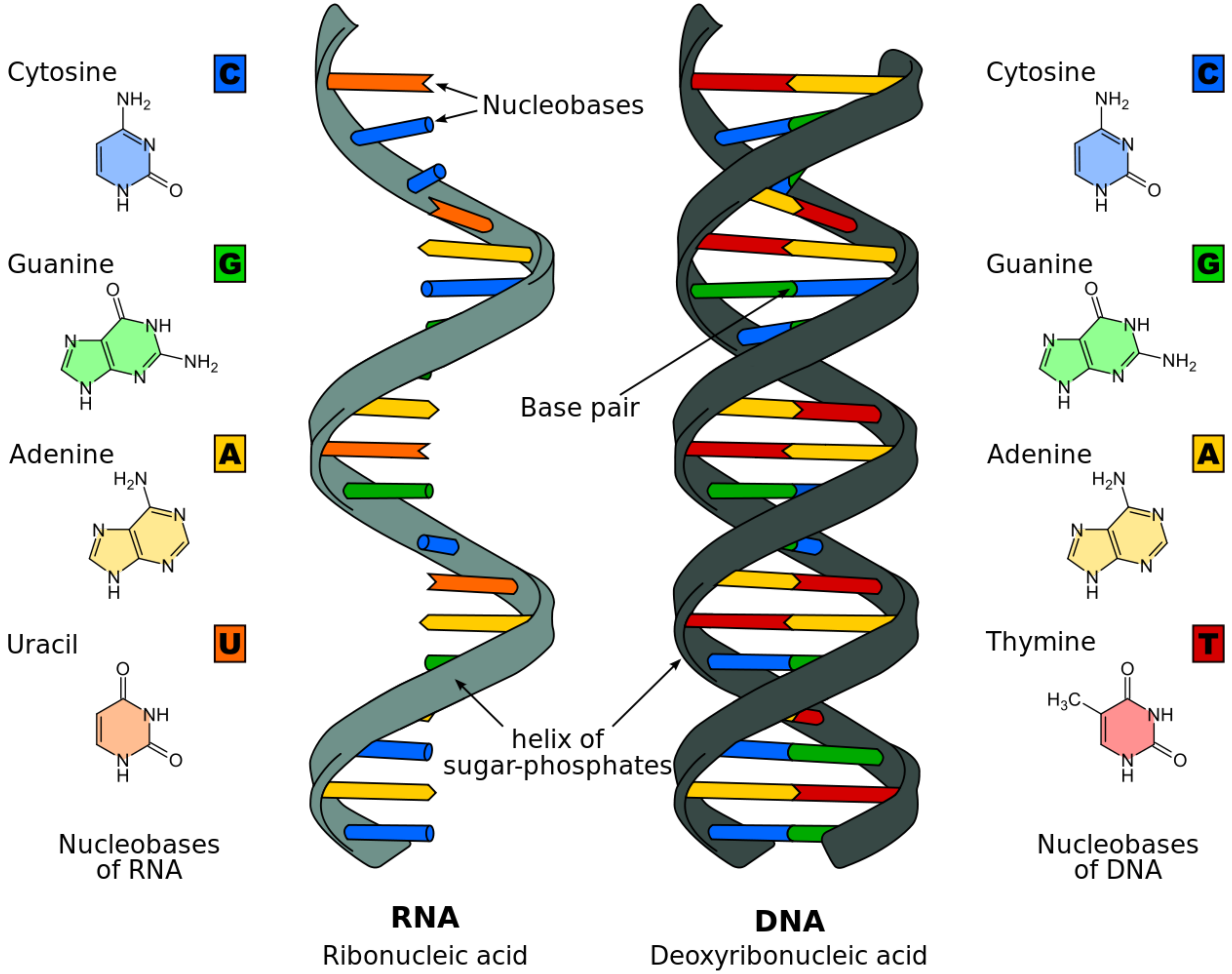


RESOURCES ON COURSE WEBSITE UNDER SESSION 1

RNA

RNA

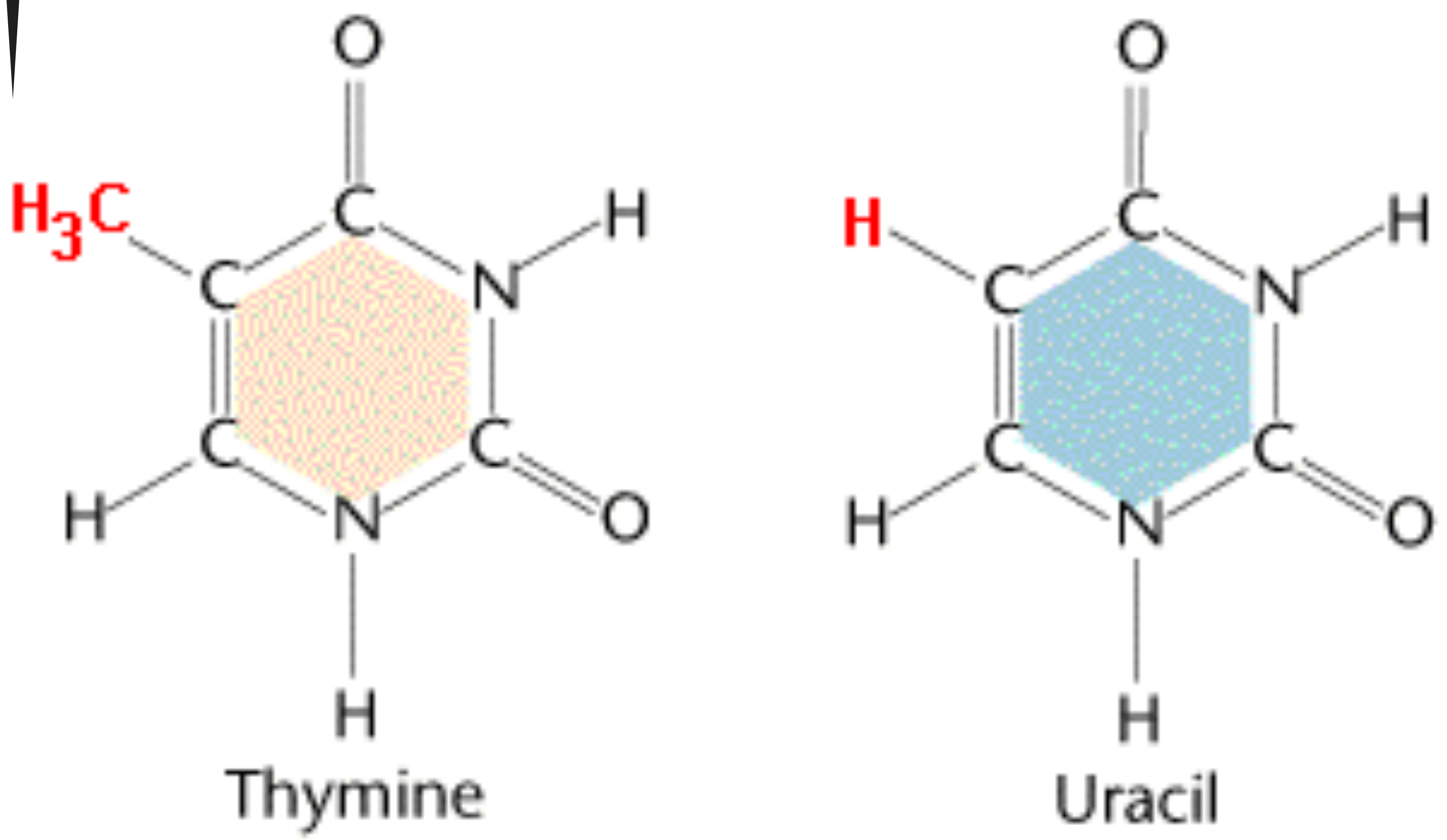
- Ribonucleic acid
- RNA is chemically similar to DNA chemically
  - Different base
  - Different sugar



RNA

- T(hyamine) is replaced by U(racil)
- Methylation of DNA
  - More structural stability
  - More efficient for replication

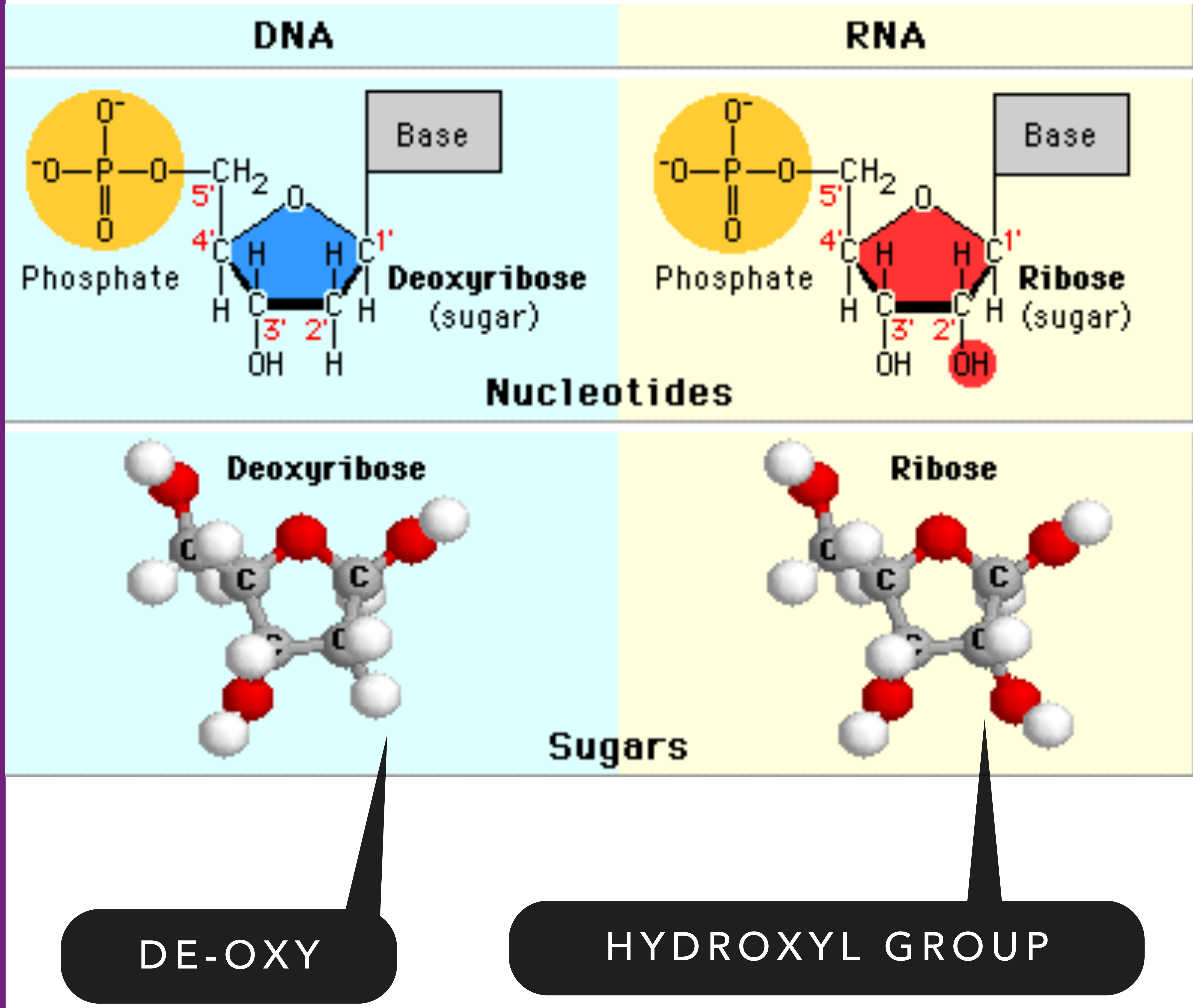
METHYL GROUP



(Klug & Cummings 1997)

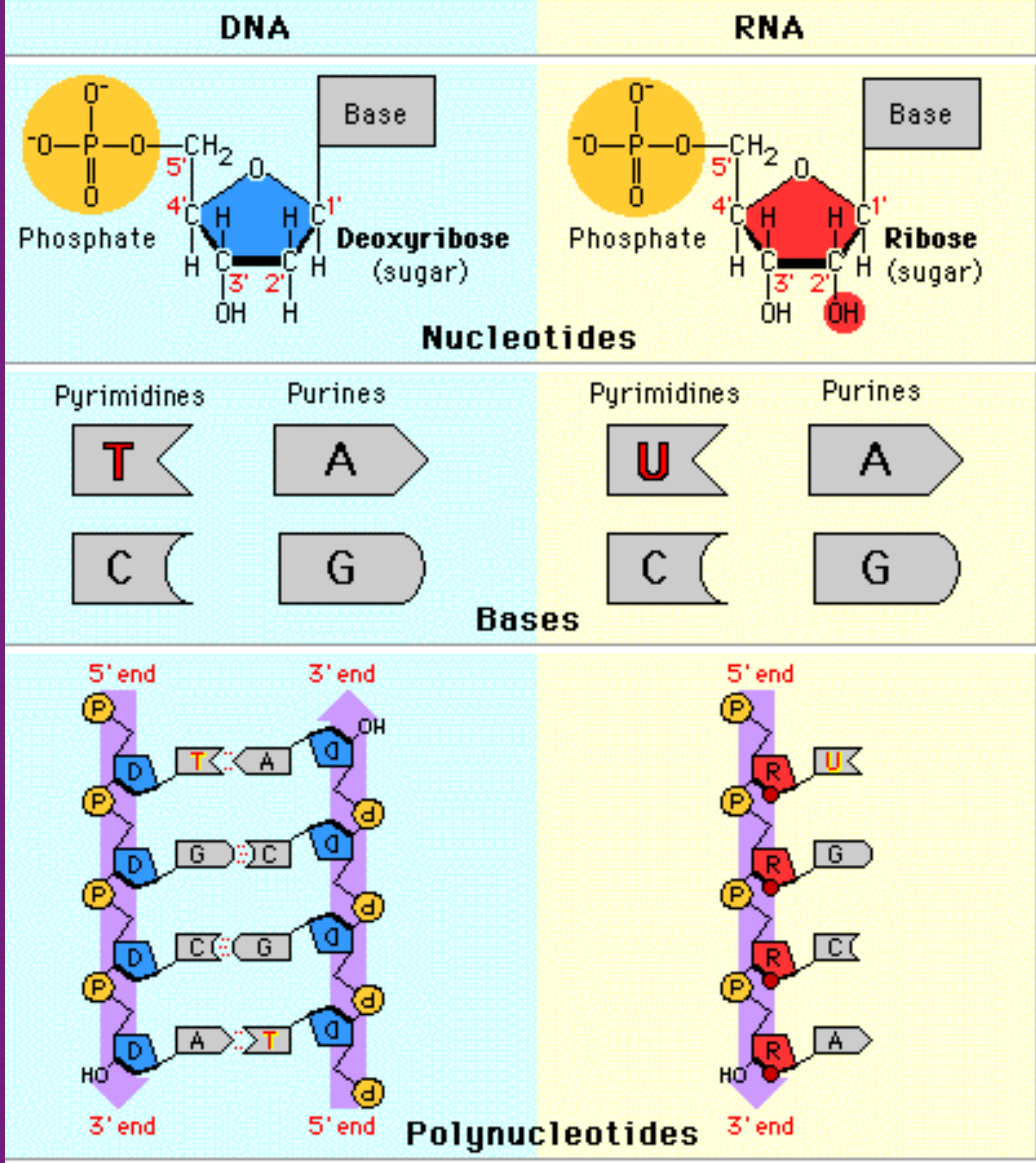
RNA

- RNA sugar is ribose
- DNA sugar is Deoxyribose (minus one oxygen atom)
    - Important for the enzymes that recognize DNA and RNA inside organisms



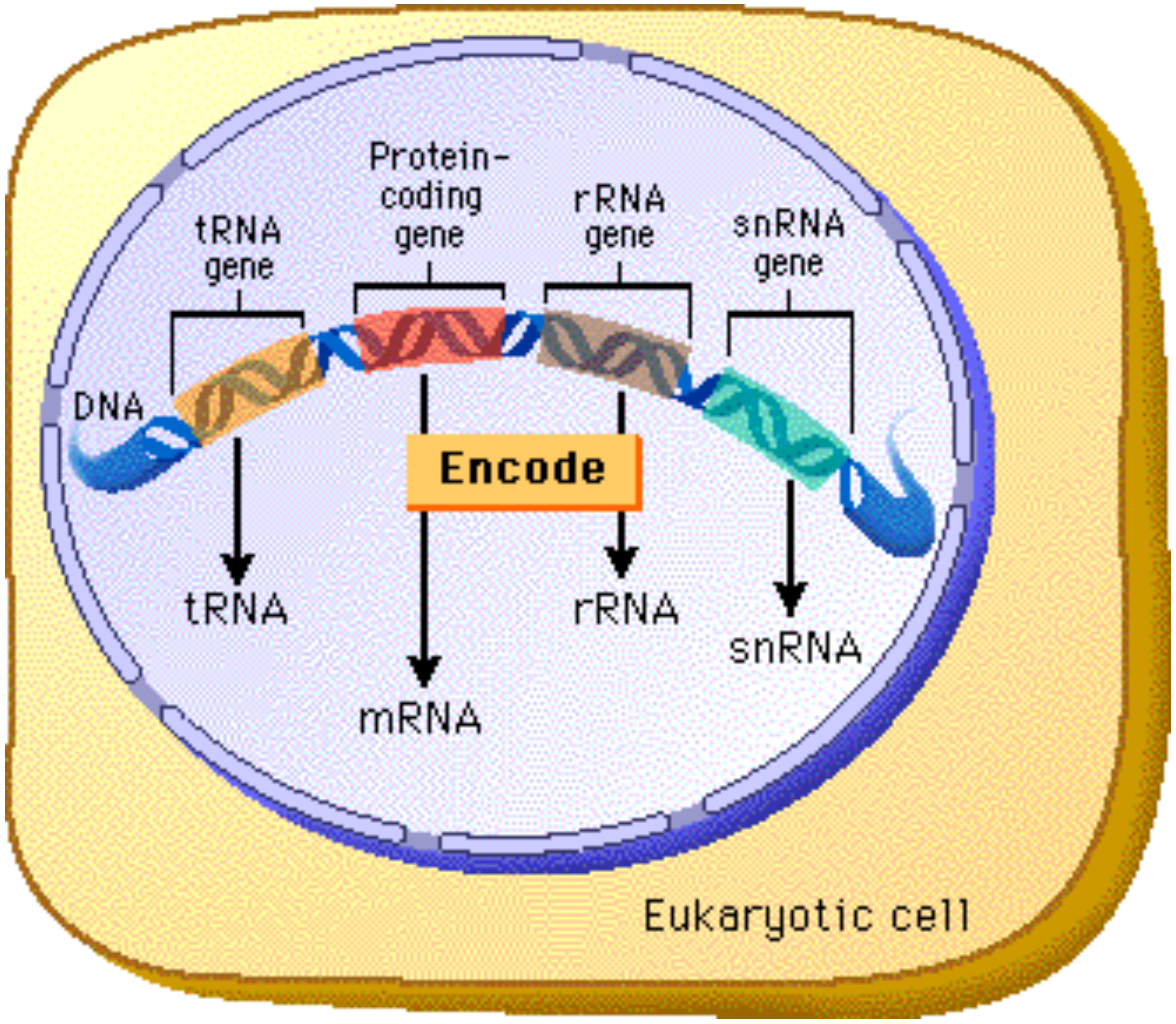
RNA

- Properties of RNA
  - More structurally flexible than DNA
  - Less stable than DNA
  - More reactive (OH group)
  - Almost all RNA is single stranded



RNA

- 4 types of RNA, each encoded by a different gene
  - **mRNA** - Messenger RNA: Encodes amino acid sequence of a polypeptide
  - **tRNA** - Transfer RNA: Brings amino acids to ribosomes during translation
  - **rRNA** - Ribosomal RNA: With ribosomal proteins, makes up the ribosomes
  - **snRNA** - Small nuclear RNA; combine with proteins for post-transcriptional processing (in eukaryotes only)

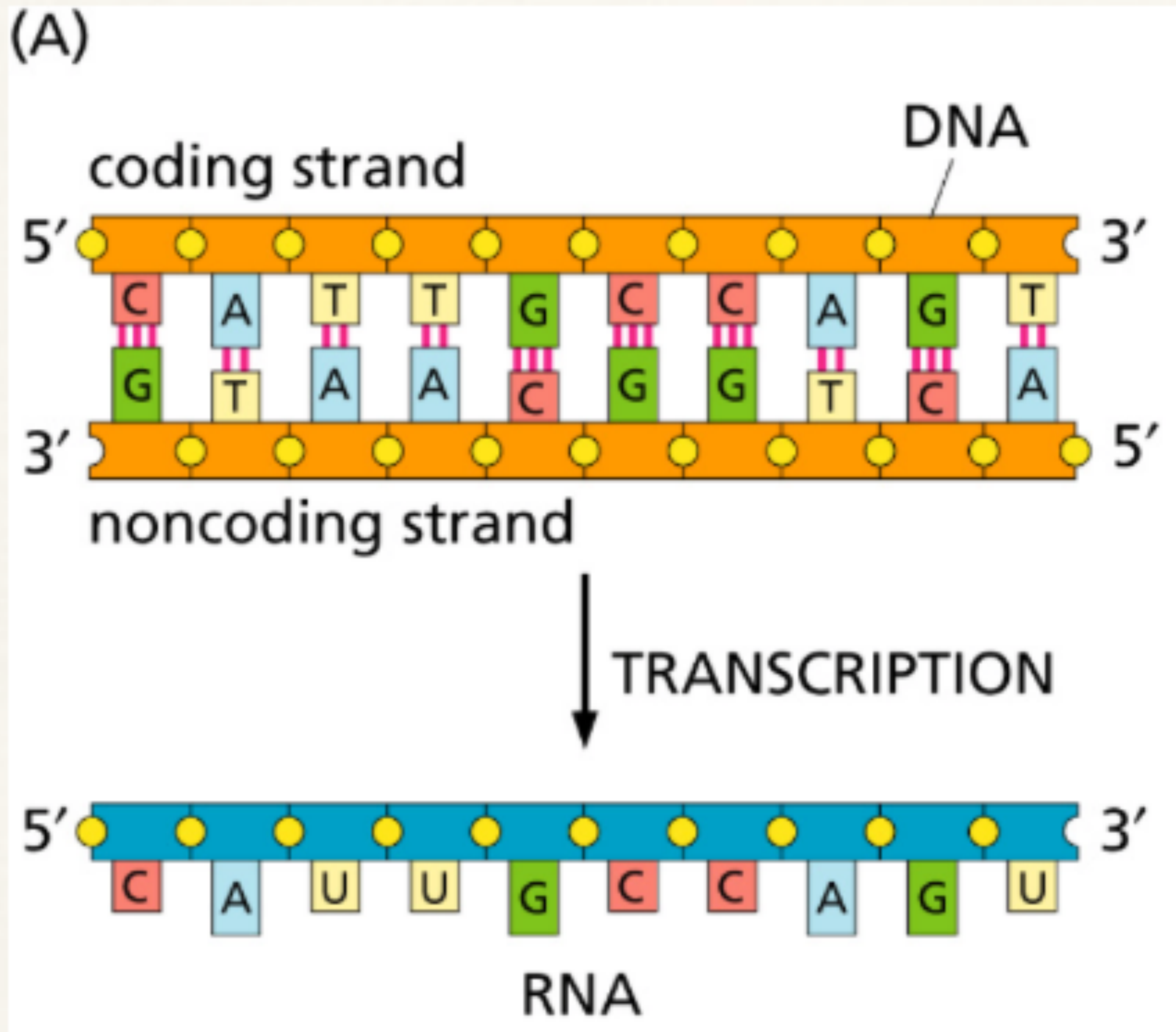


Transcription

TRANSCRIPTION

- The synthesis of an RNA copy of a segment of DNA
  - When a gene is being transcribed into RNA, the gene is said to be expressed

GENE IS TRANSCRIBED FROM DNA

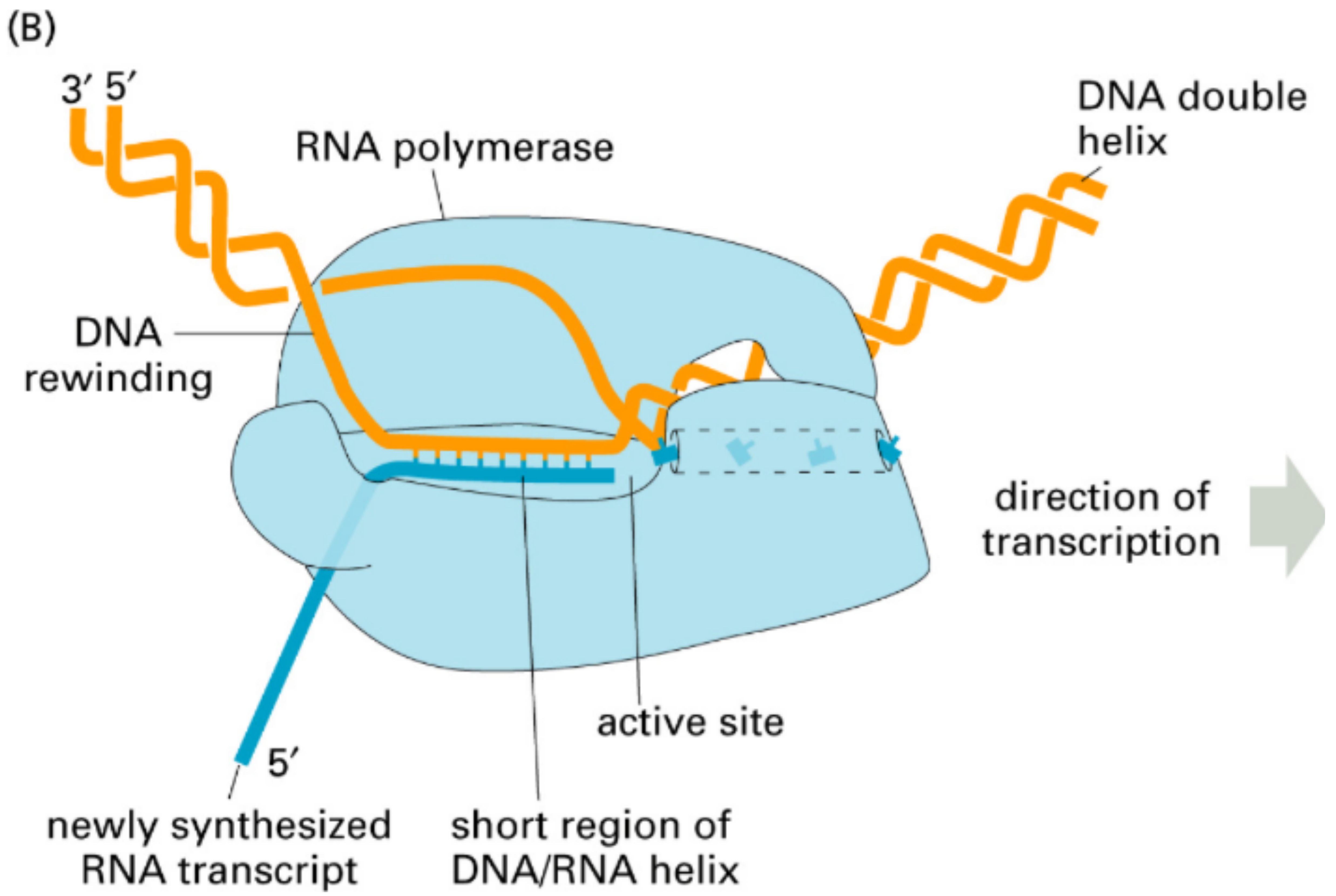


TRANSCRIPTION

- Basic: https://www.youtube.com/watch?v=5MfSYnItYvg
- Advanced: https://www.youtube.com/watch?v=SMtWvDbfHLo

TRANSCRIPTION

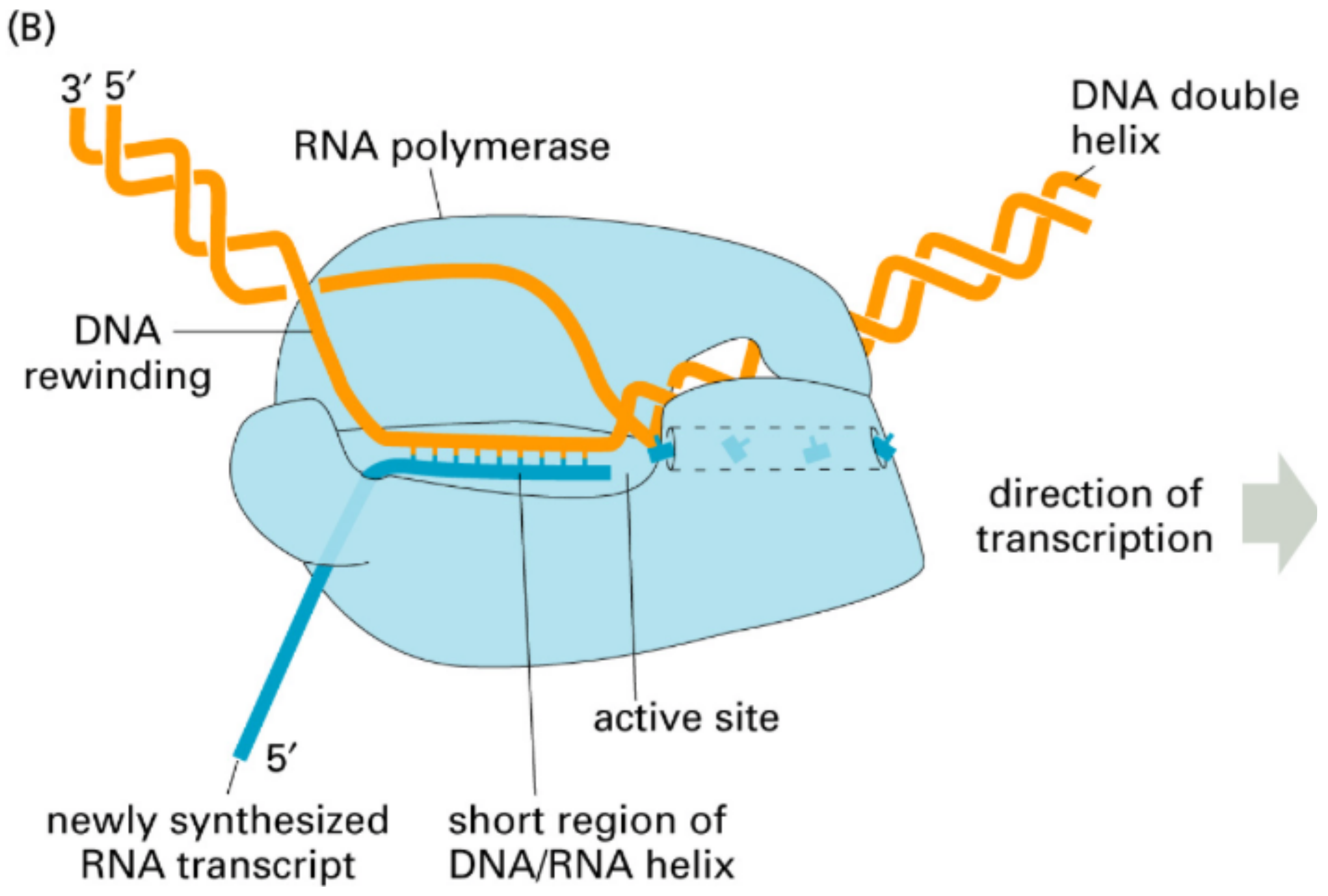
- RNA synthesis
  - Separation of the DNA strands
  - Use one DNA strand as a template
  - Synthesis of RNA in the 5' to 3' direction by **RNA polymerase**



http://www.cs.rice.edu/~nakhleh/COMP572/Slides/MolCellBiolReview.pdf

TRANSCRIPTION

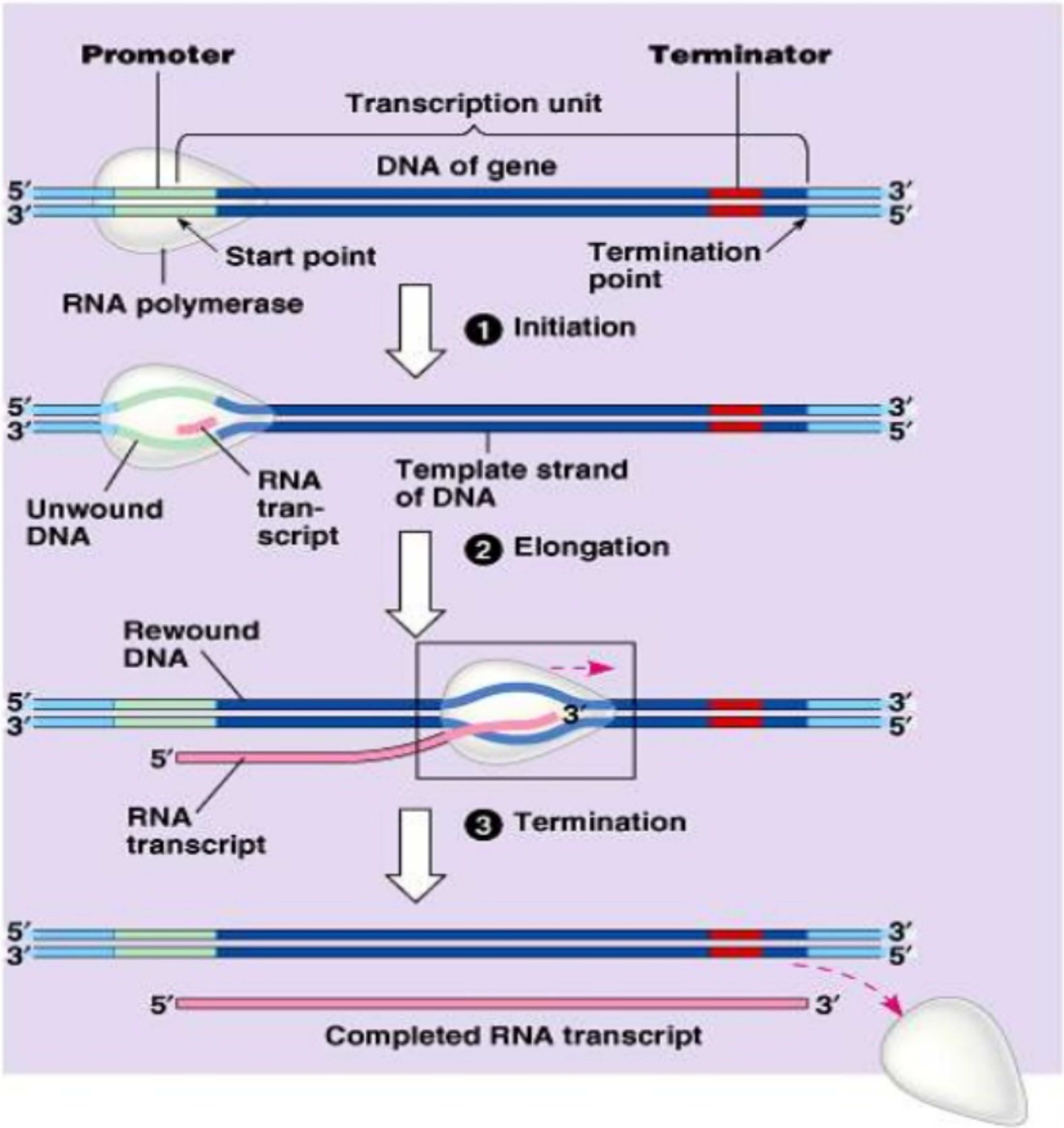
- RNA Polymerase
  - Multi-subunit enzyme
  - Reads the DNA and recruits the correct building blocks of RNA to string them together based on the DNA code



http://www.cs.rice.edu/~nakhleh/COMP572/Slides/MolCellBiolReview.pdf

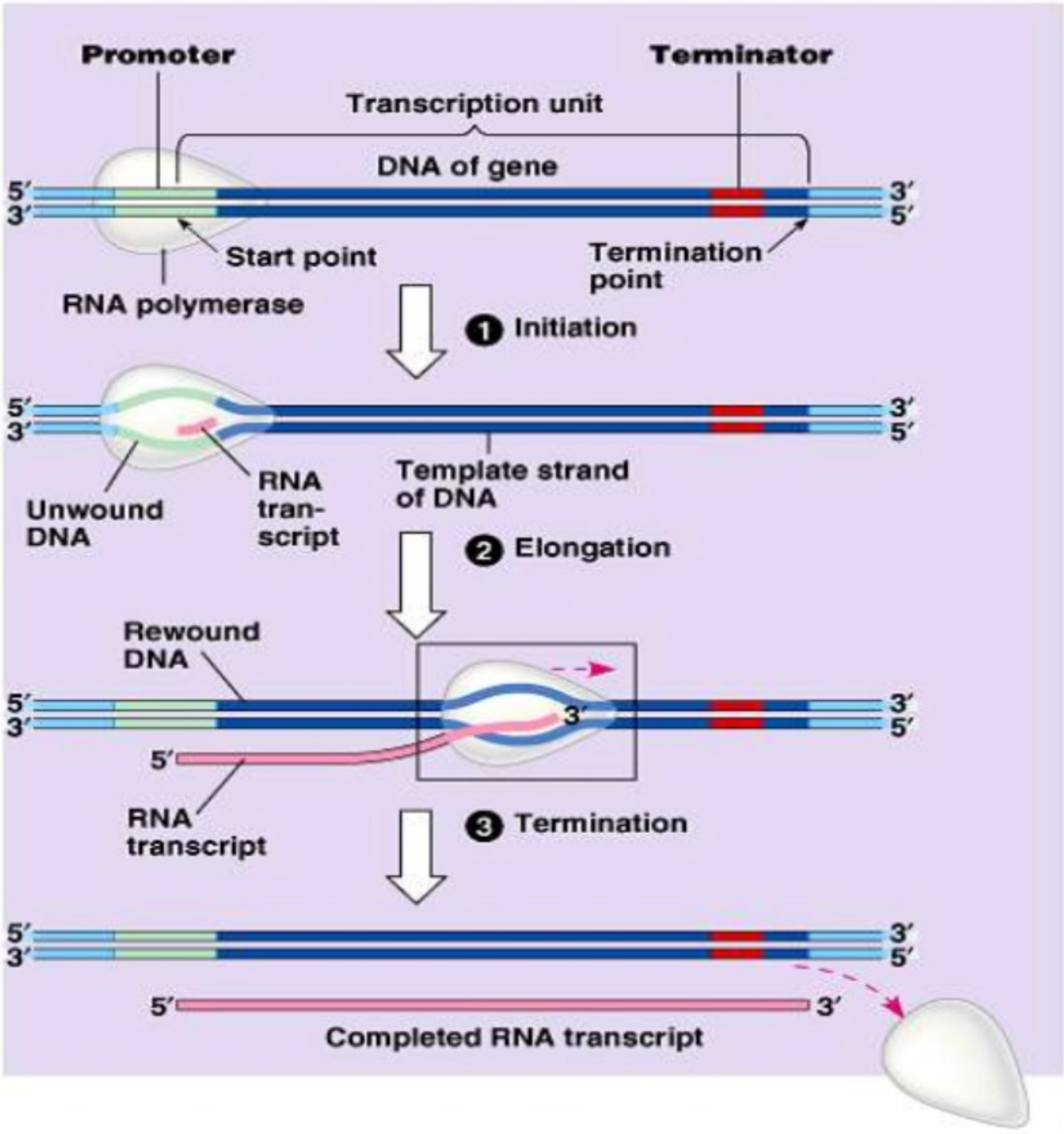
TRANSCRIPTION

- Initiation Steps
  - Transcription factors
    - Proteins bind together on the promoter region to form a transcription initiation complex
  - Promoter
    - Special sequence of nucleotides indicating the starting point for RNA synthesis
  - RNA polymerase
    - Binds the transcription initiation complex

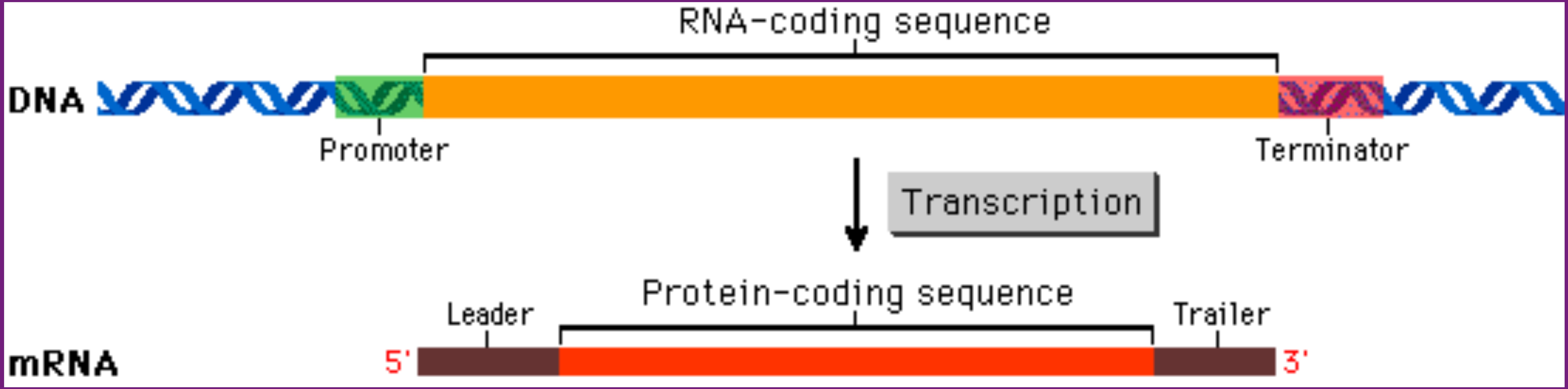


TRANSCRIPTION

- Elongation
  - RNA polymerase starts moving along the template strand from the 5′ to 3′ adding free RNA to make mRNA
- Termination
  - RNA polymerase stops at a termination codon
  - Terminator
    - Signal in DNA that halts transcription

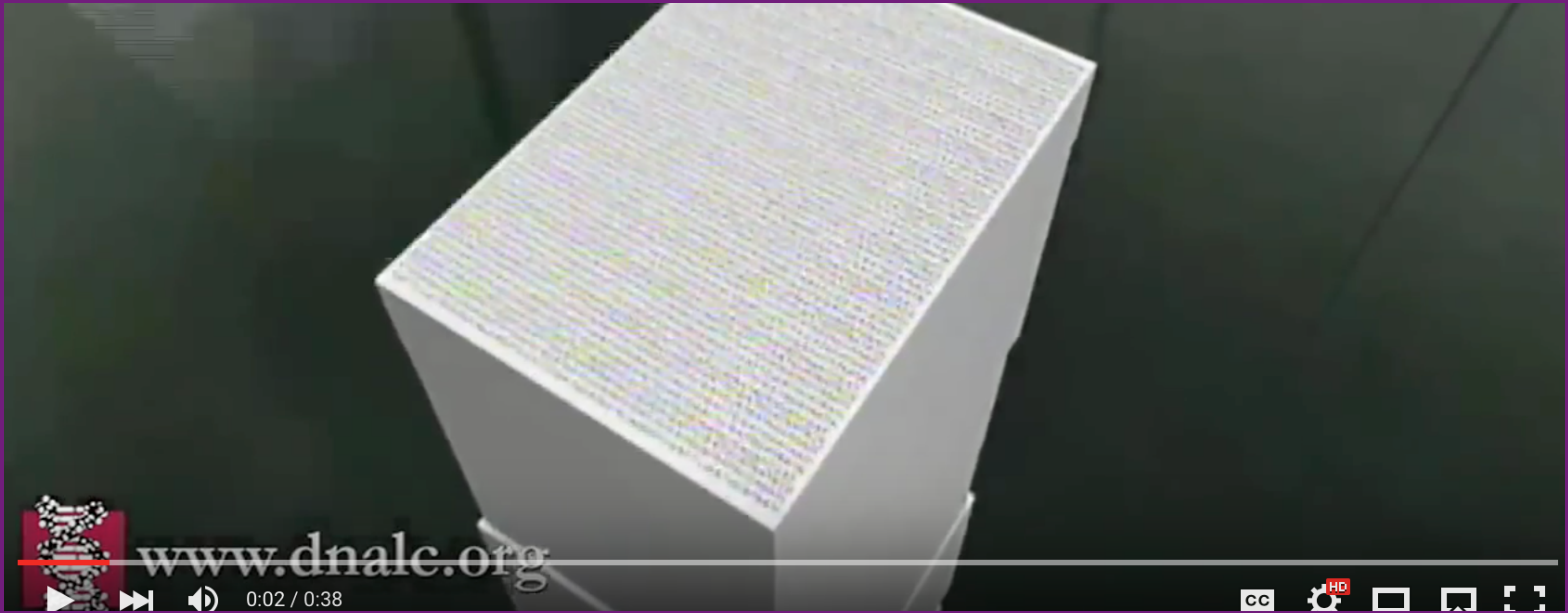


TRANSCRIPTION



- mRNA in prokaryotes
    - Colinear with the translated mRNA
    - The transcript of the gene is the molecule that is translated directly to polypeptide

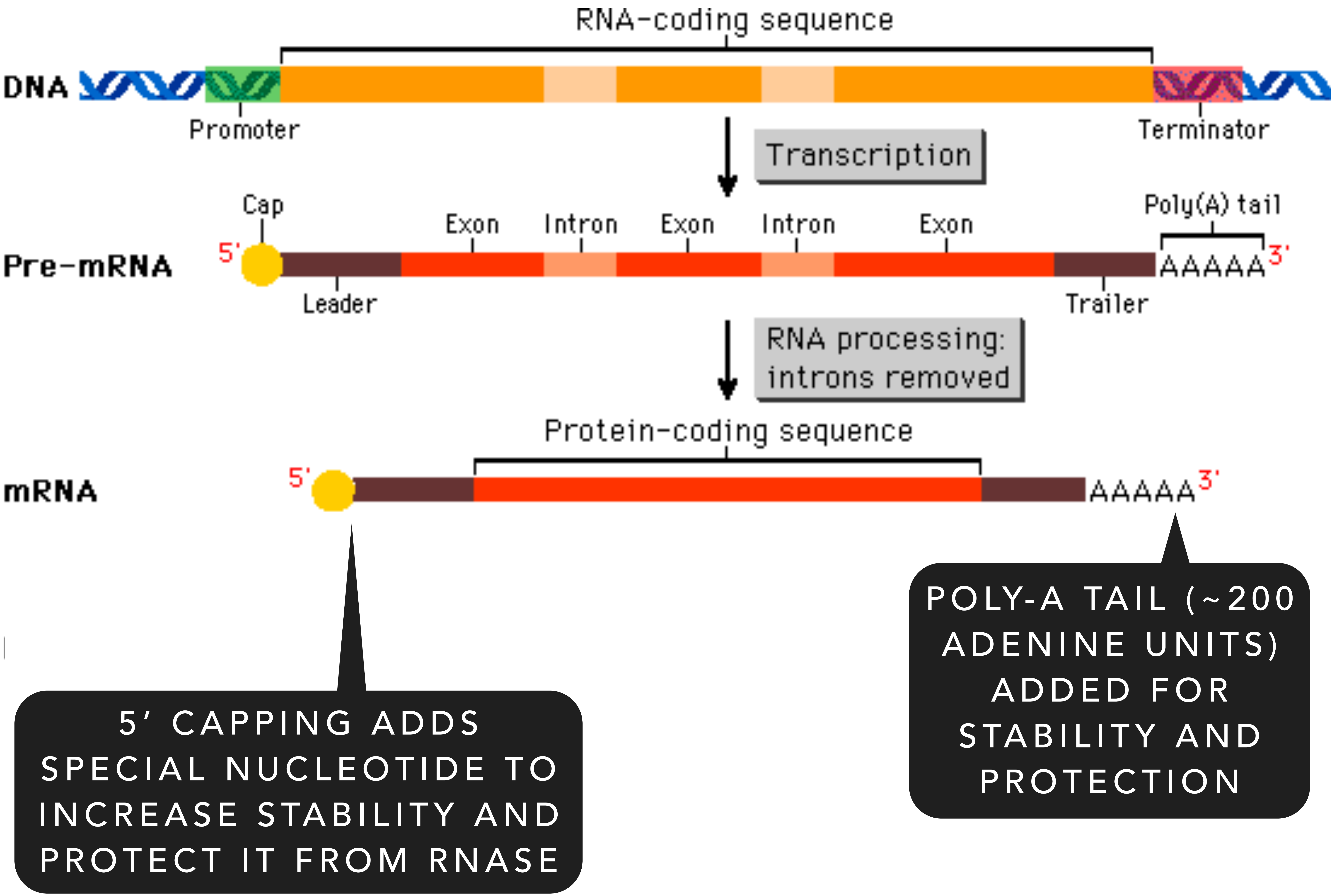
TRANSCRIPTION



- https://www.youtube.com/watch?v=hV6NSHjTR1s

TRANSCRIPTION

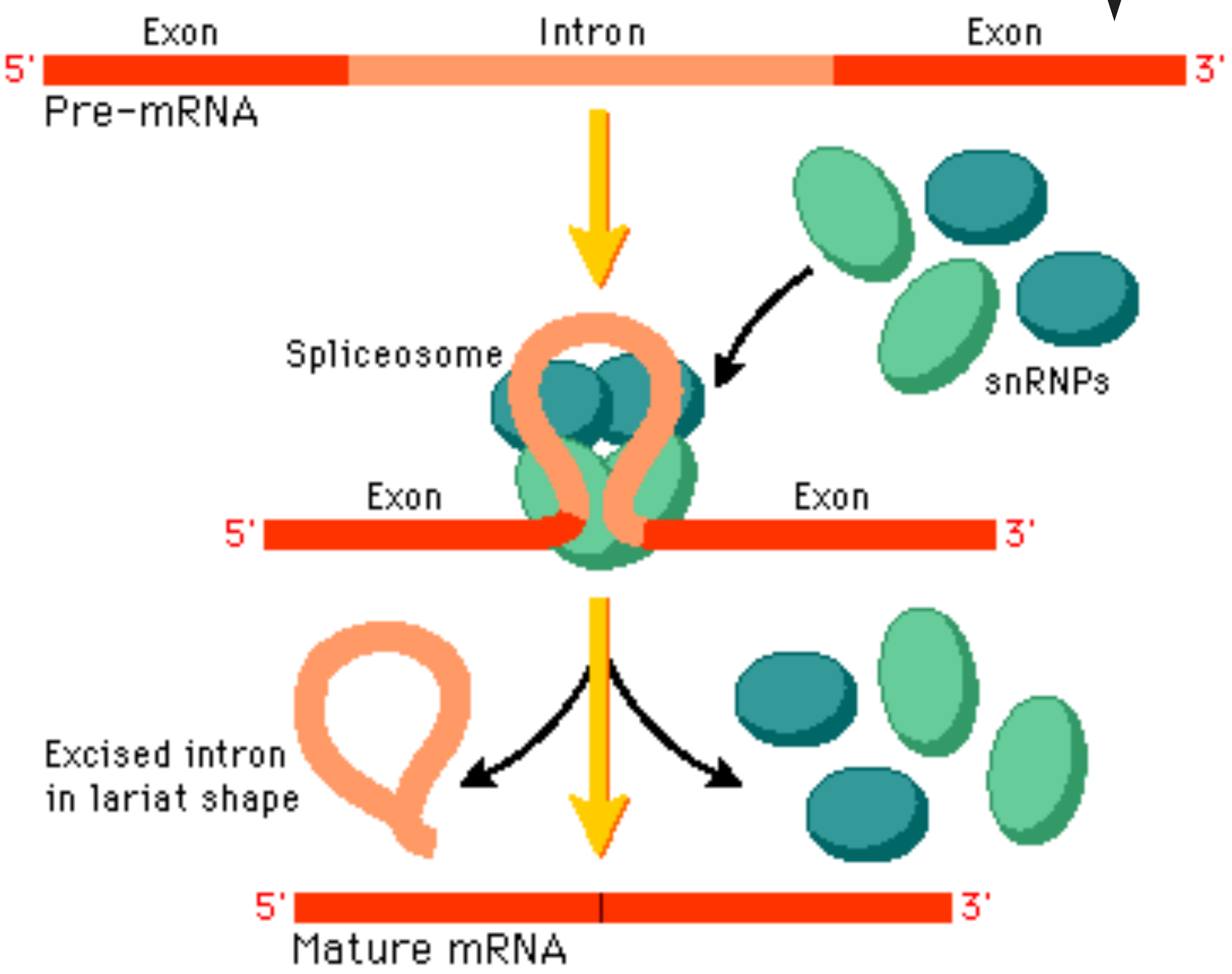
- mRNA in eukaryotes
  - Post-transcriptional modification
  - Spliceosome processes gene to remove extra sequences (introns)
  - Exits nucleus to cytoplasm



TRANSCRIPTION

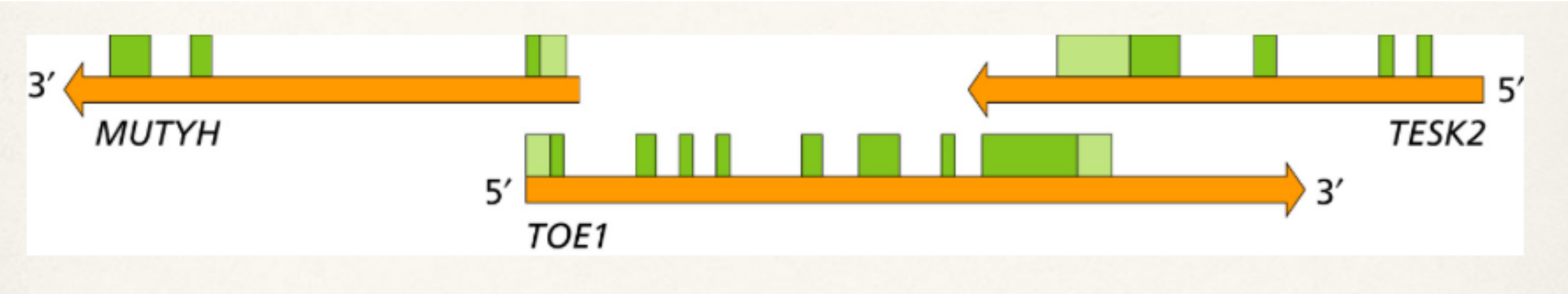
SNRNP - SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES; COMPLEXES OF SNRNAS AND PROTEINS; FORM THE SPLICEOSOME

- Pre-mRNA Processing (splicing)
  - The intron loops out as snRNPs bind to form the spliceosome
  - Intron is excised
  - Exons are spliced together
  - Resulting mature mRNA exits the nucleus



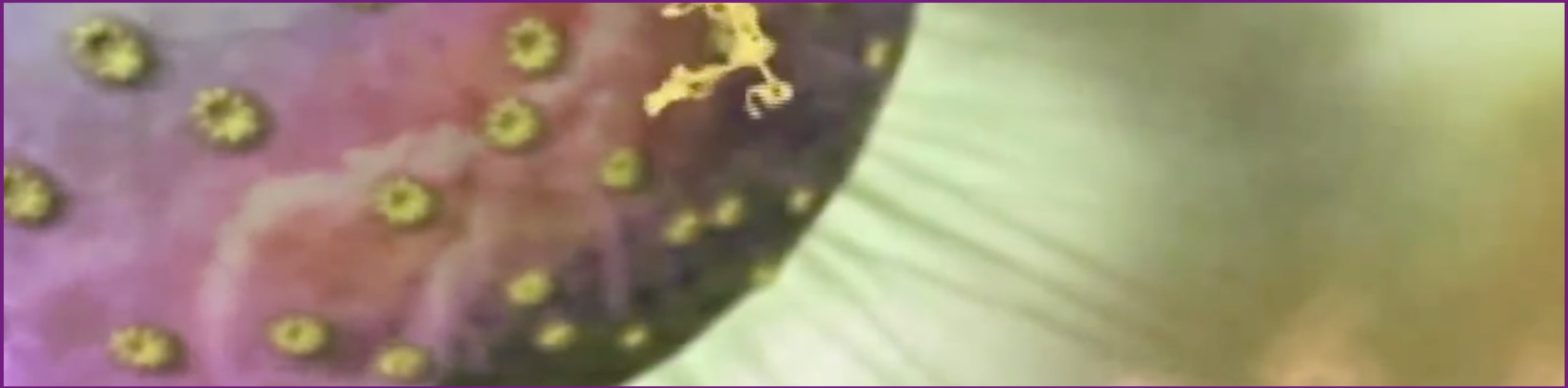
TRANSCRIPTION

- Only one segment of DNA is transcribed for any given gene
  - Genes can overlap so that one (or both) strands encode different parts of proteins
  - Efficient for small genomes
- At any given time a cell only expresses a fraction of the genes in its genome
  - Regulated gene expression



Translation

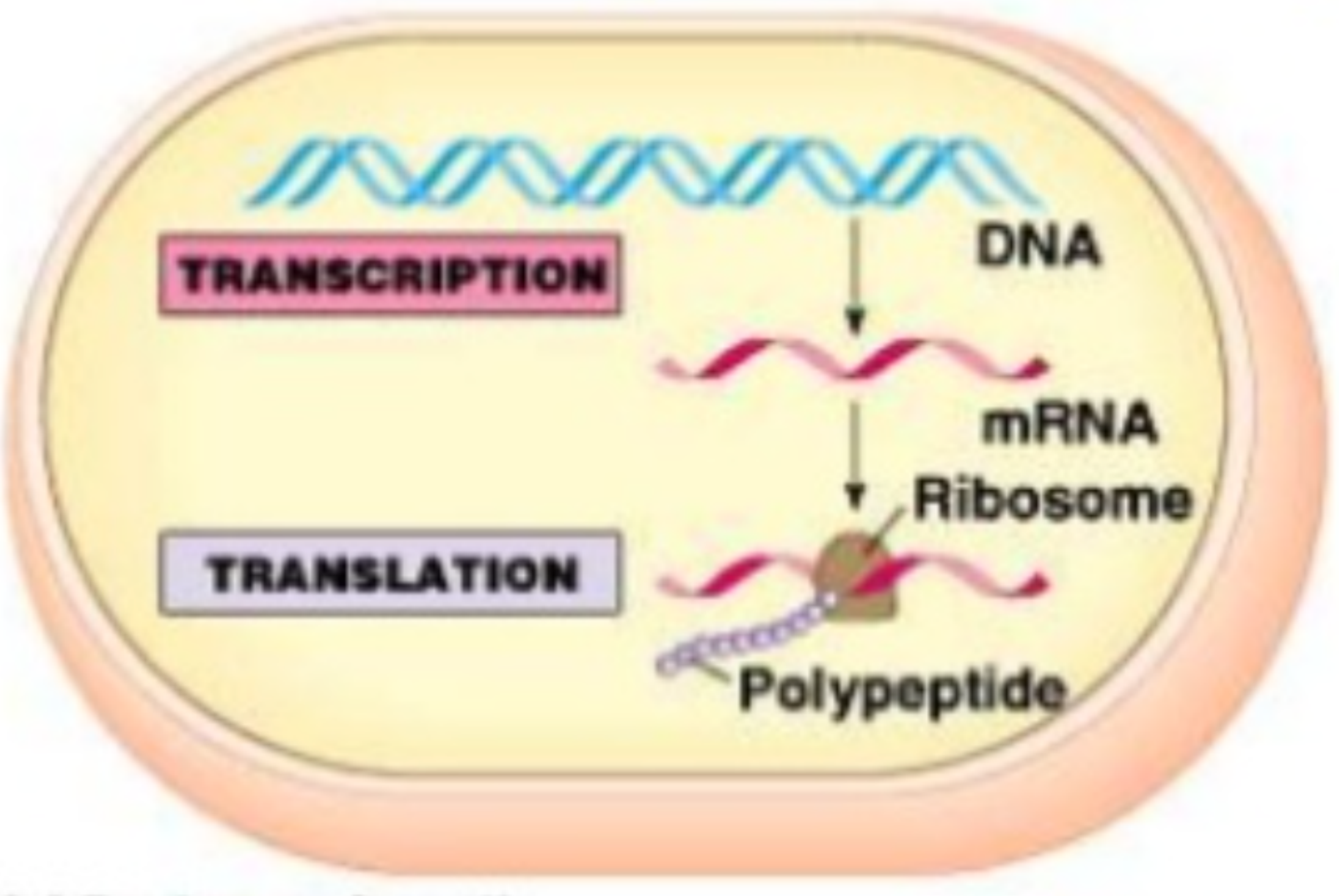
TRANSLATION



- https://www.youtube.com/watch?v=8dsTvBaUMvw Basic
- https://www.youtube.com/watch?v=TfYf_rPWUdY Advanced

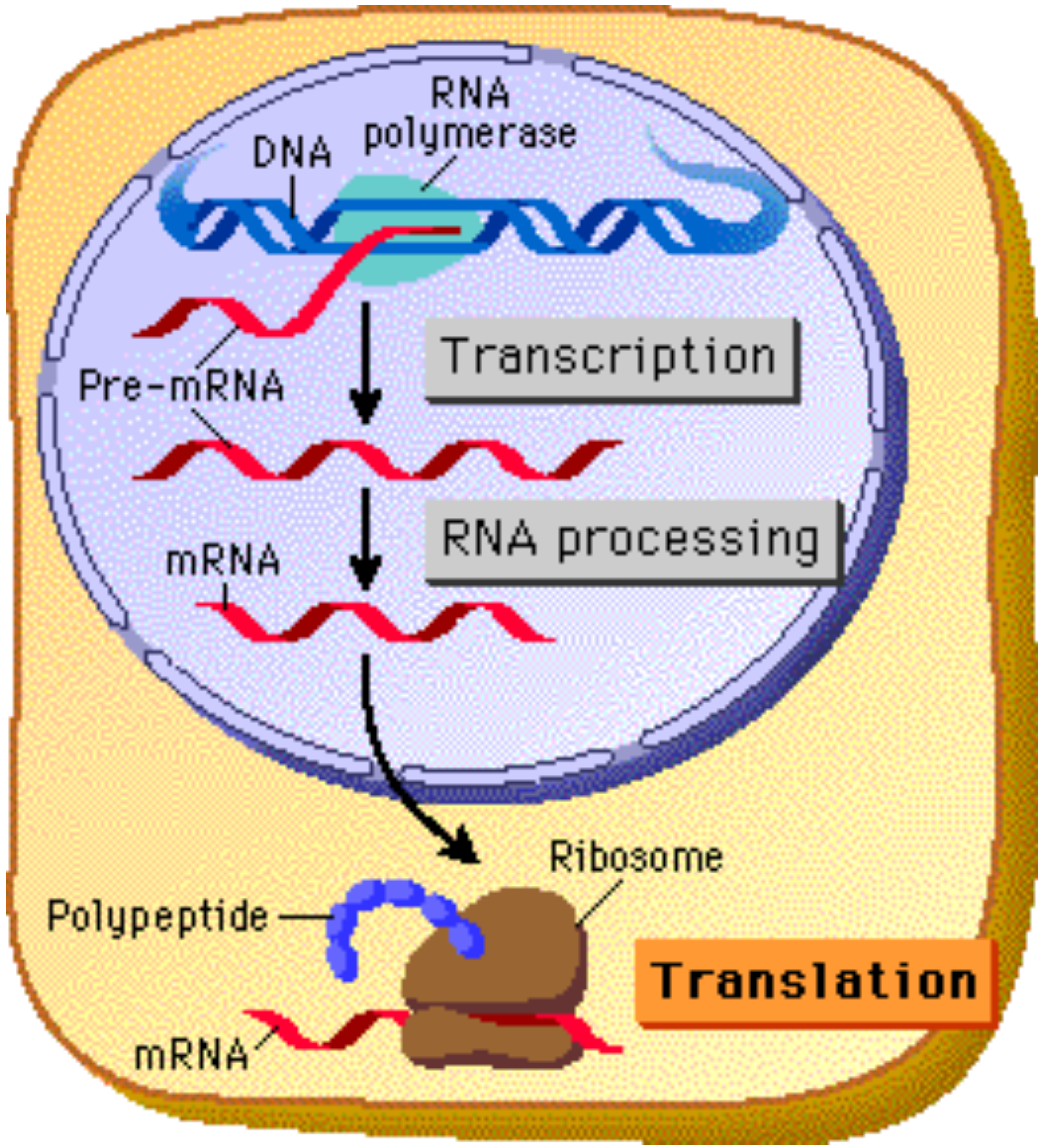
TRANSLATION

- Translation in prokaryotic cell
  - mRNA is translated by ribosomes to produce polypeptide chain

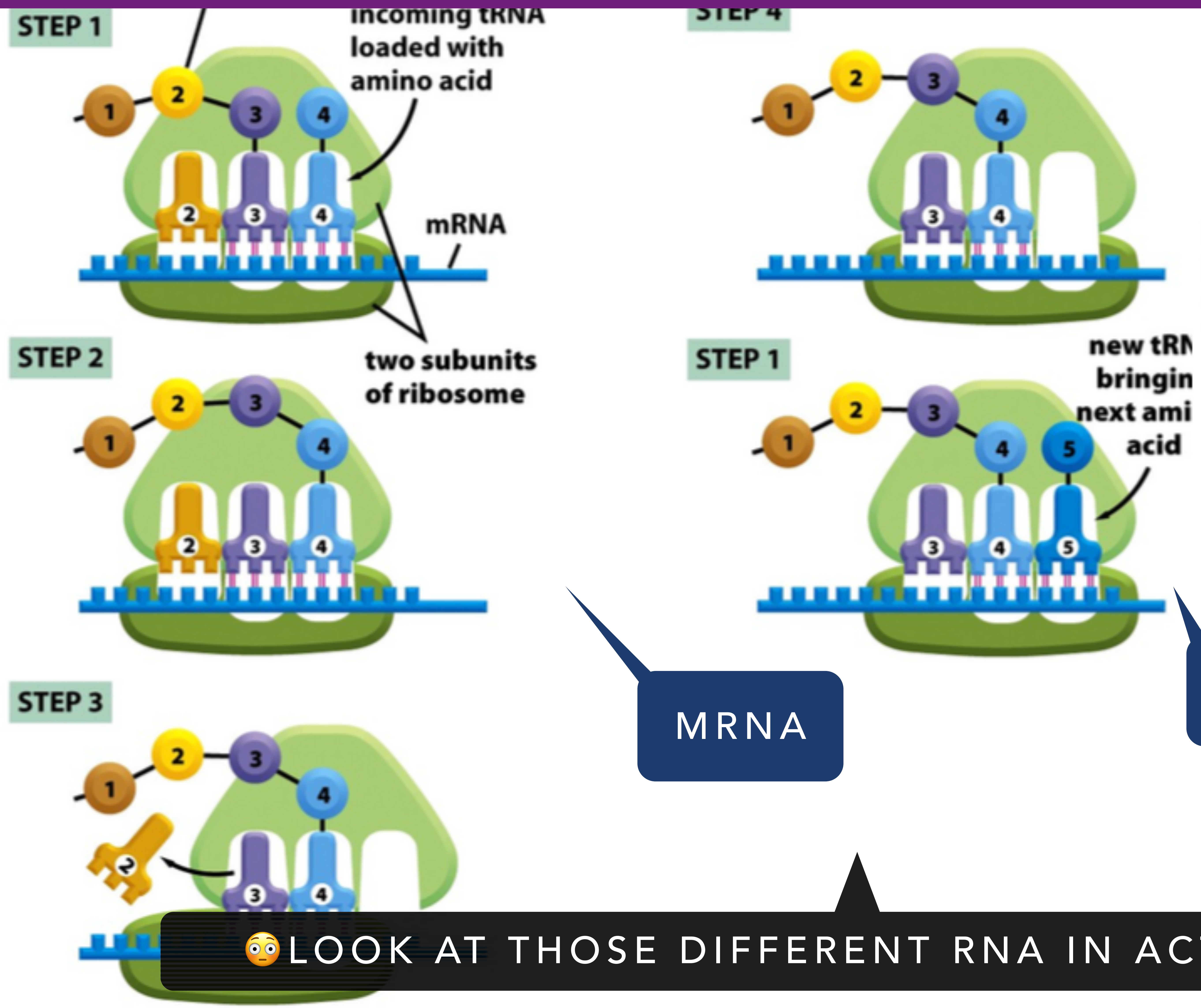


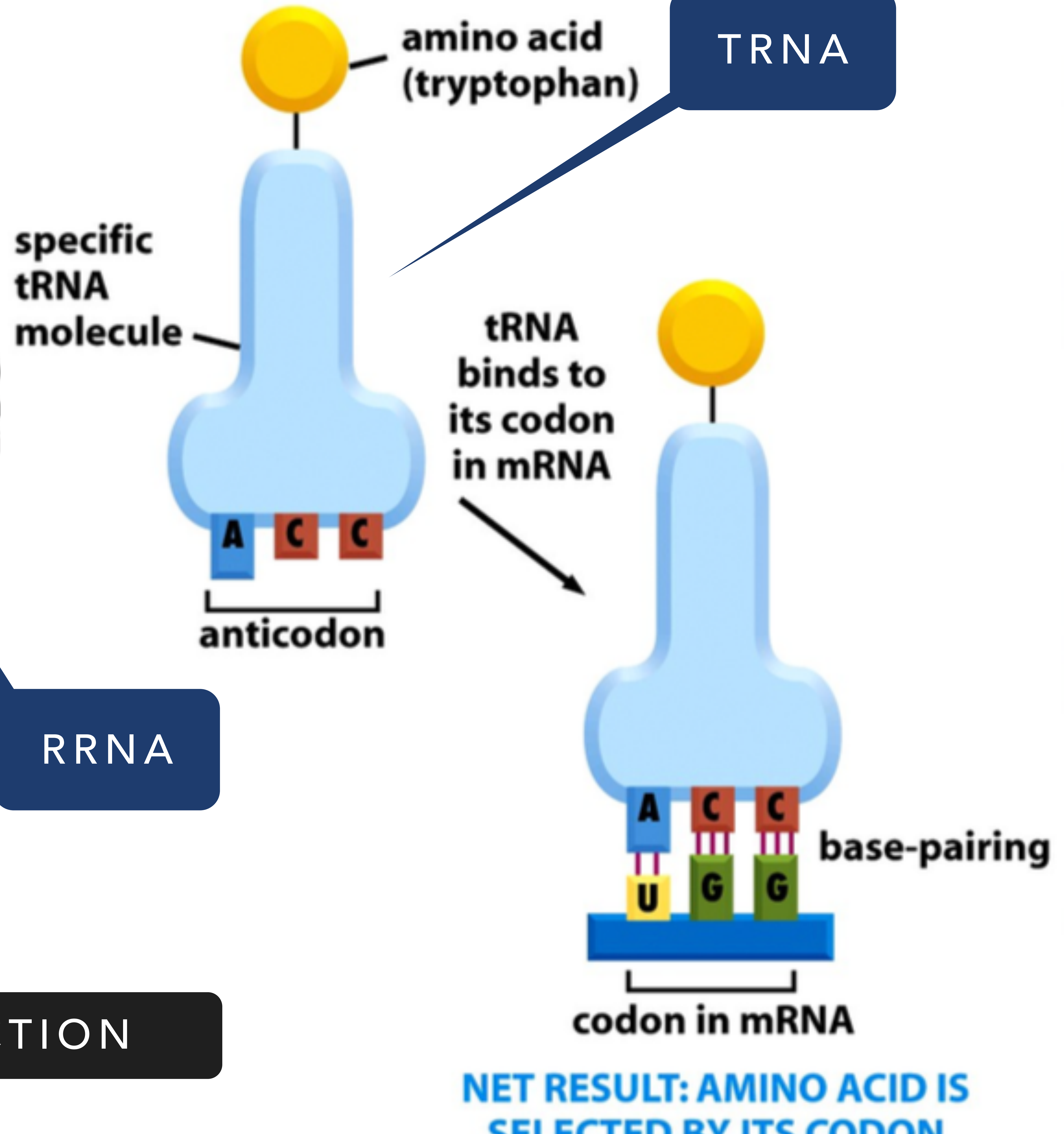
TRANSLATION

- Translation in a Eukaryotic Cell
  - A protein-coding gene is transcribed into a pre-mRNA
  - Pre-mRNA is processed into a mature mRNA
  - mRNA exits the nucleus
  - mRNA is translated on ribosomes to produce the polypeptide chain



TRANSLATION





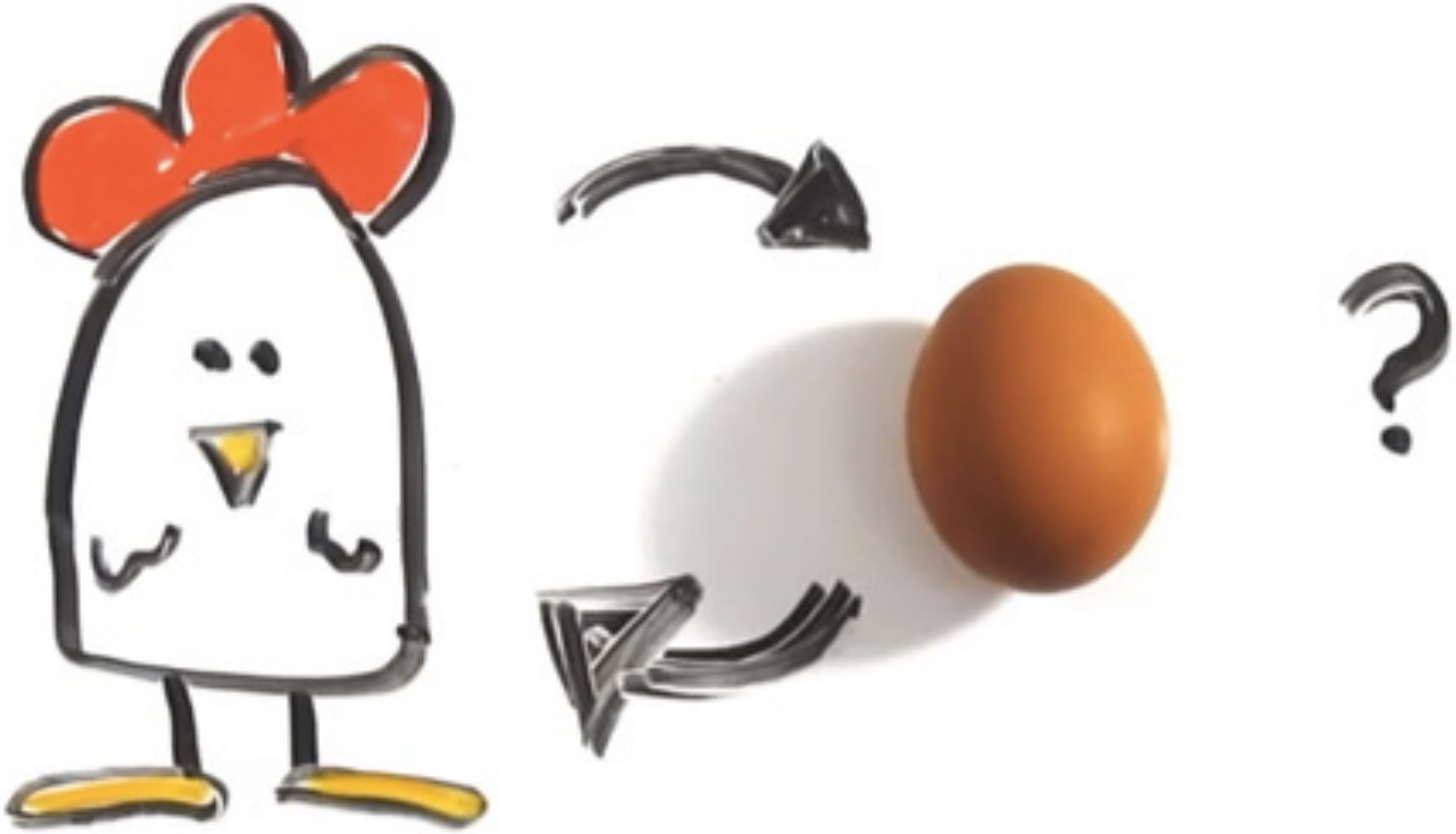
TRNA

MRNA

RRNA

😳LOOK AT THOSE DIFFERENT RNA IN ACTION

TRANSLATION



TRANSLATION

- If ribosomes are needed to make proteins but they are also made of proteins, which came first?

From left, Venkatraman Ramakrishnan of the MRC Laboratory of Molecular Biology in Cambridge, England; Thomas A. Steitz of Yale University; and Ada E. Yonath of the Weizmann Institute of Science in Rehovot, Israel, will share the 2009 Nobel Prize in Chemistry.

THE ANSWER, DR. BERG SAID, IS THAT THE ACTIVE CORE OF THE RIBOSOME IS MADE OF RNA. THE PROTEIN SEEMS TO HAVE BEEN ADDED LATER, WHICH MEANS THE RIBOSOME IS "AN RNA-BASED MACHINE THAT EVOLVED THE ABILITY TO MAKE PROTEINS."

TRANSLATION

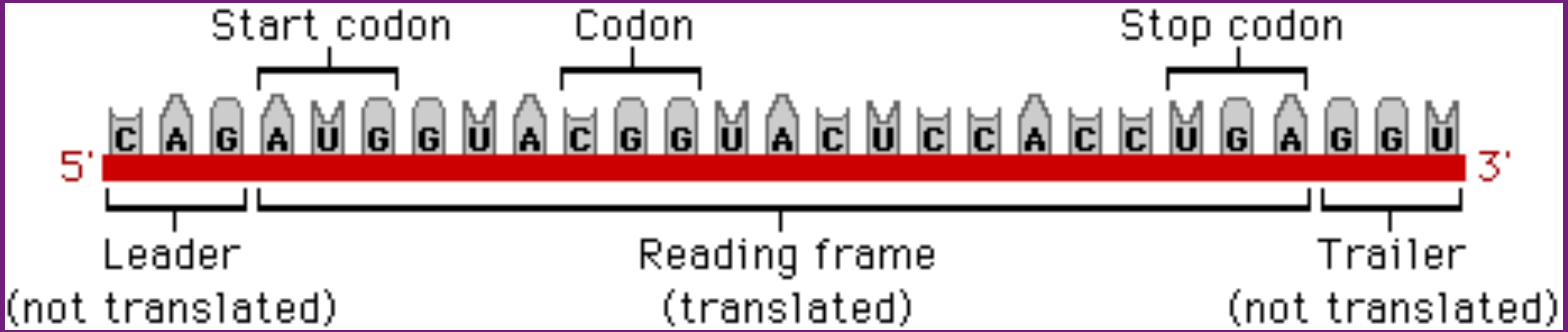
- mRNA is translated into a protein according the genetic code
- Codon
  - 3 nucleotides read at a time
  - Start codon (AUG)
  - Stop codon (UAA,UAG,UGA) tell the ribosome the protein is complete

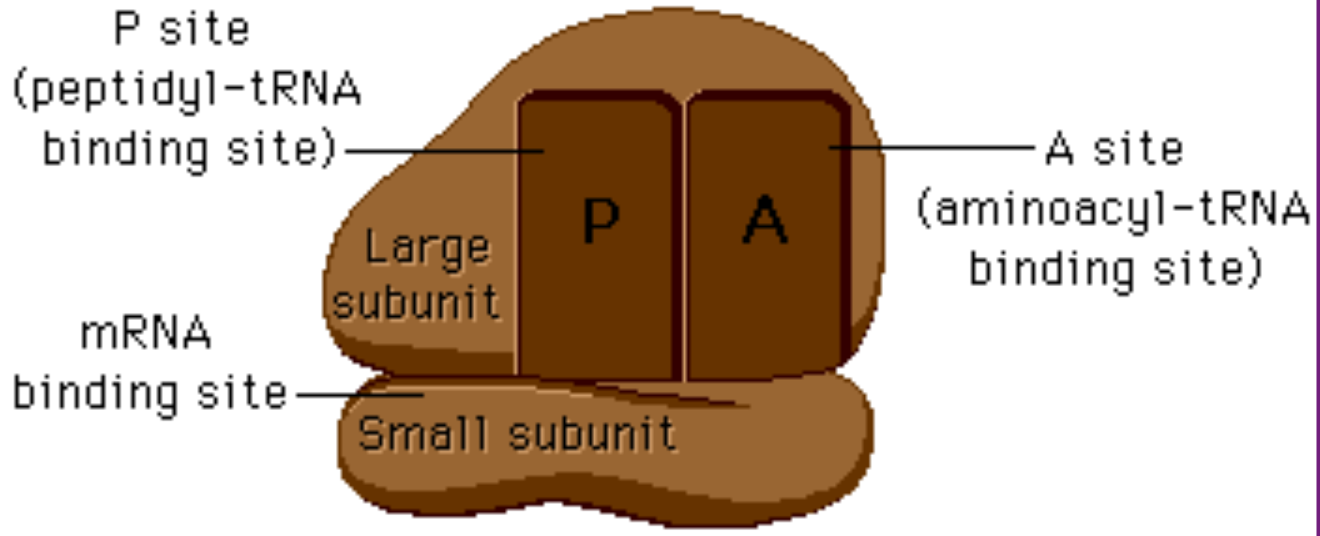
Second letter

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

STANDARD GENETIC CODE

TRANSLATION

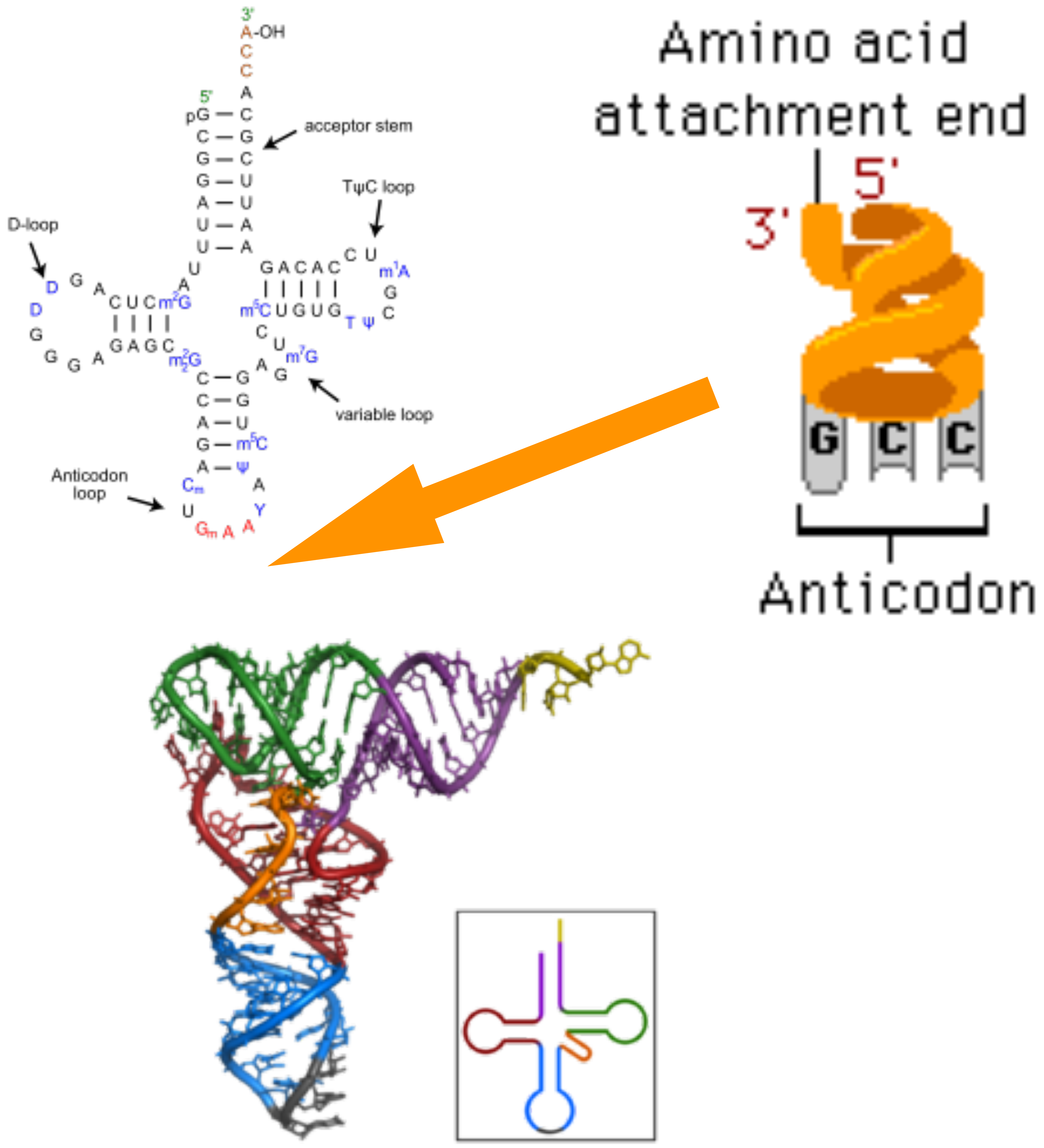




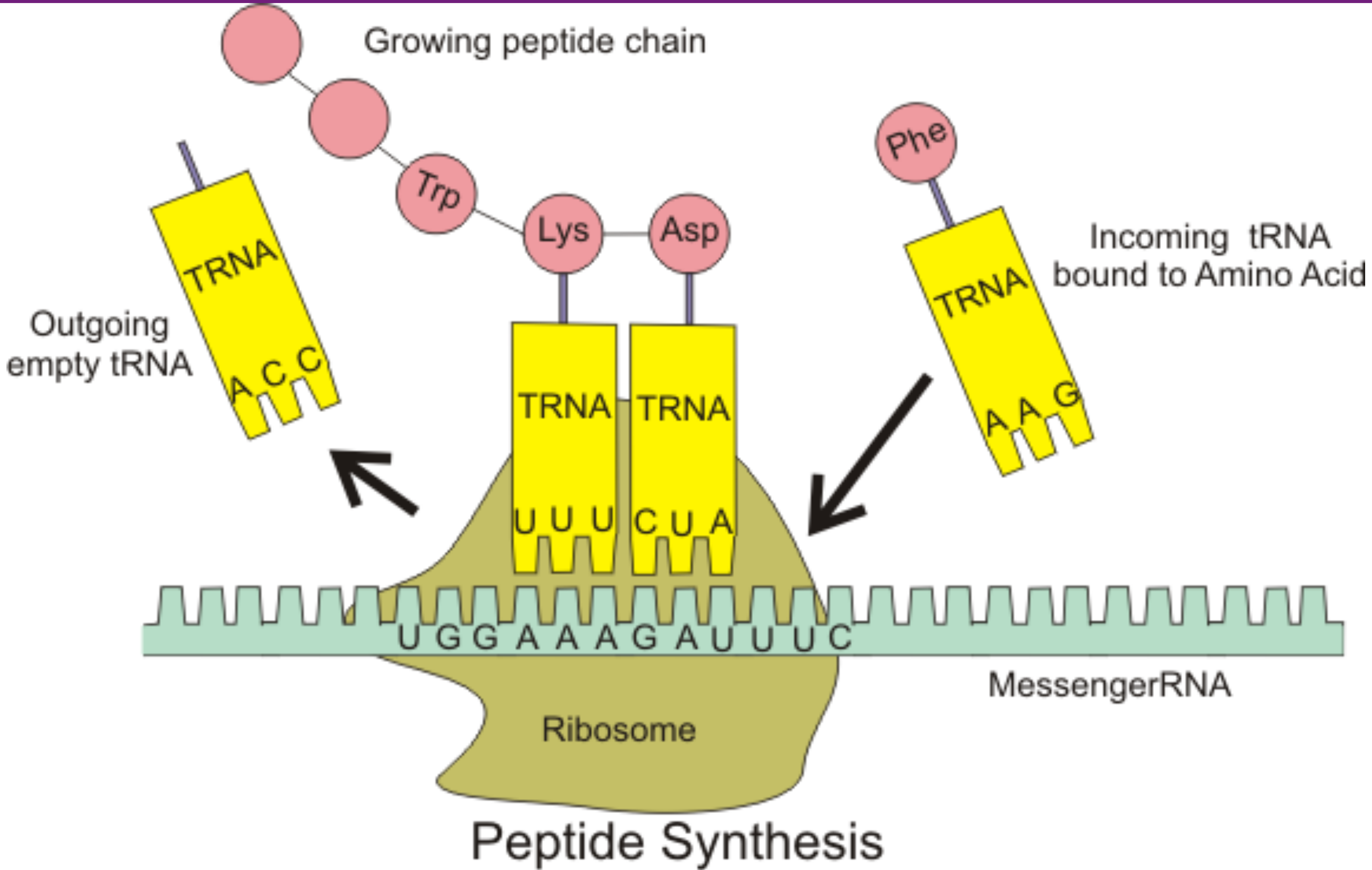
- mRNA passes through the ribosome
- Codons are recognized by tRNAs carrying the specified amino acids
- Each ribosomal subunit consists of rRNA (ribosomal RNA, encoded by rRNA genes) and ribosomal proteins

TRANSLATION

- Transfer RNA
  - Encoded by tRNA genes
  - All tRNA molecules are similar in size and shape (cloverleaf)
  - All tRNAs have CCA at the 3' end to which the amino acid attaches
  - Anticodon “reads" the matching codon on the mRNA

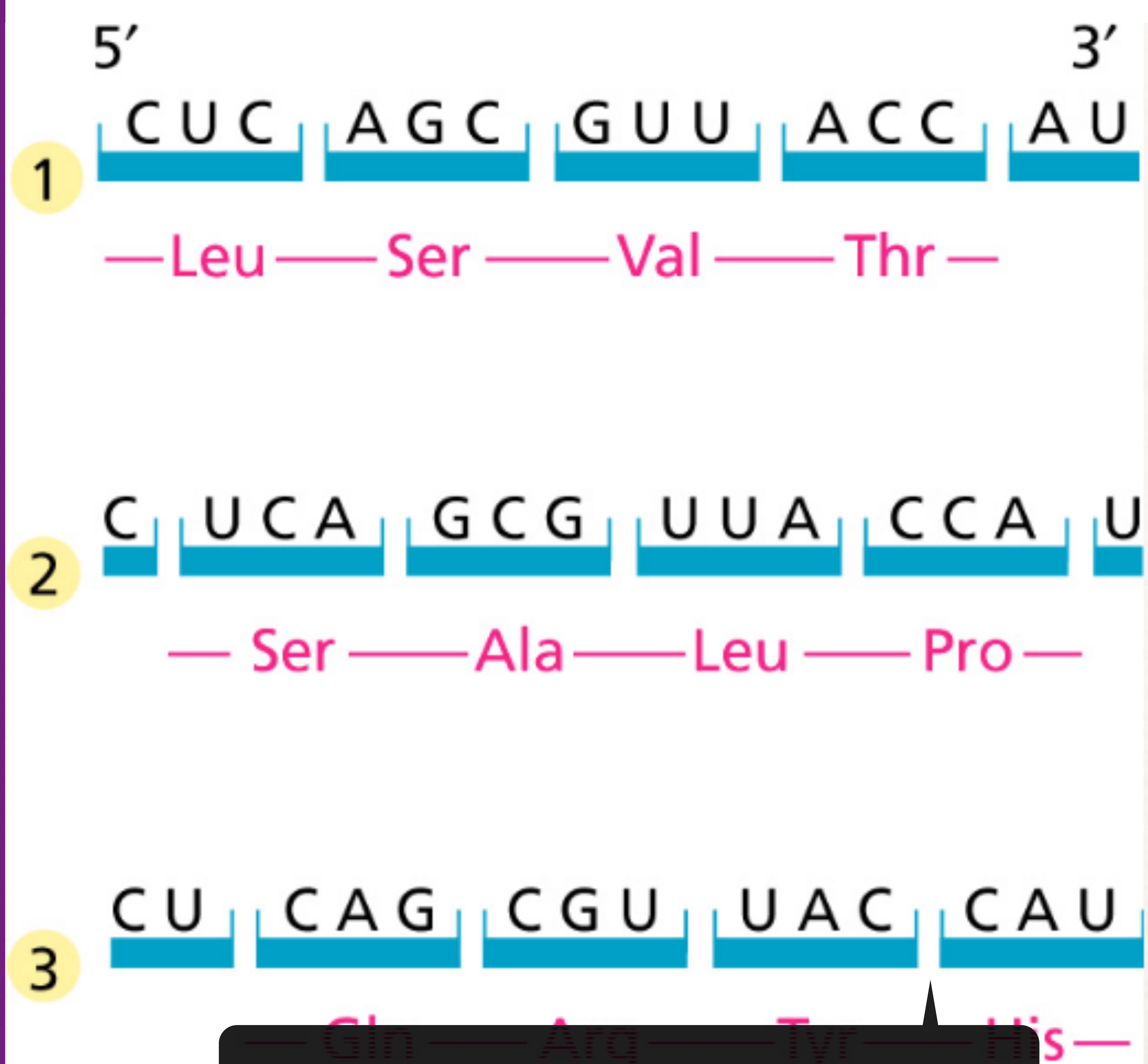


TRANSLATION



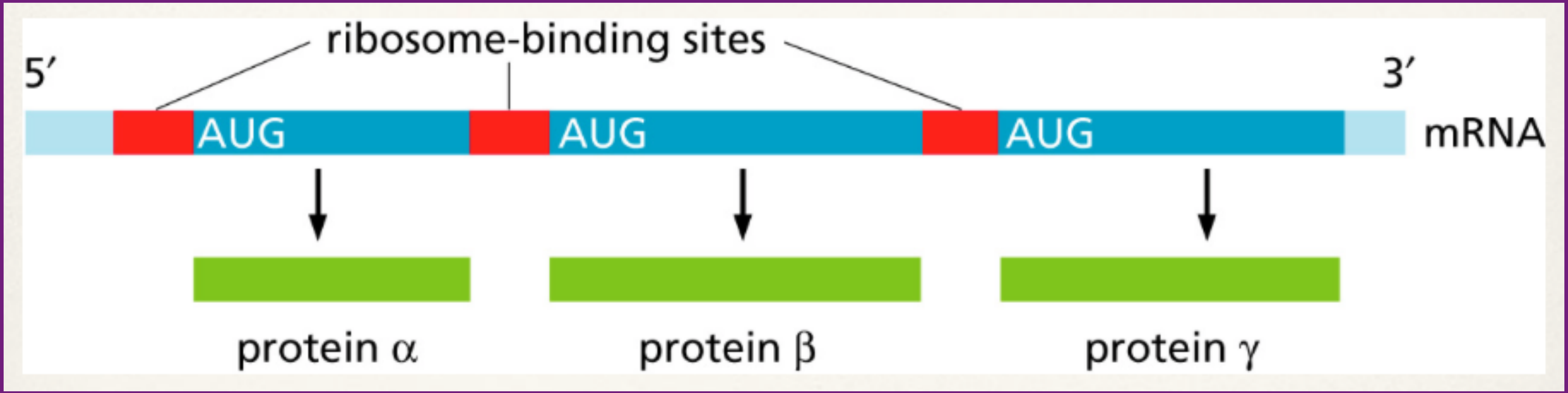
TRANSLATION

- Reading frames
  - Translation occurs in nonoverlapping sets of three bases
  - There are thus three possible ways to translate any nucleotide sequence
  - These three reading frames give three different protein sequences
- Detailed control signals ensure that only the appropriate reading frame is translated into protein



3 DIFFERENT READING FRAMES

TRANSLATION



- Functionally related protein-coding sequences are often clustered together into **operons**
- Each operon is transcribed as a single mRNA transcript; proteins are separately translated from this one long molecule
  - A single **operator** activate the simultaneous expression of all genes in the operon
- Allows efficient, coordinated protein synthesis

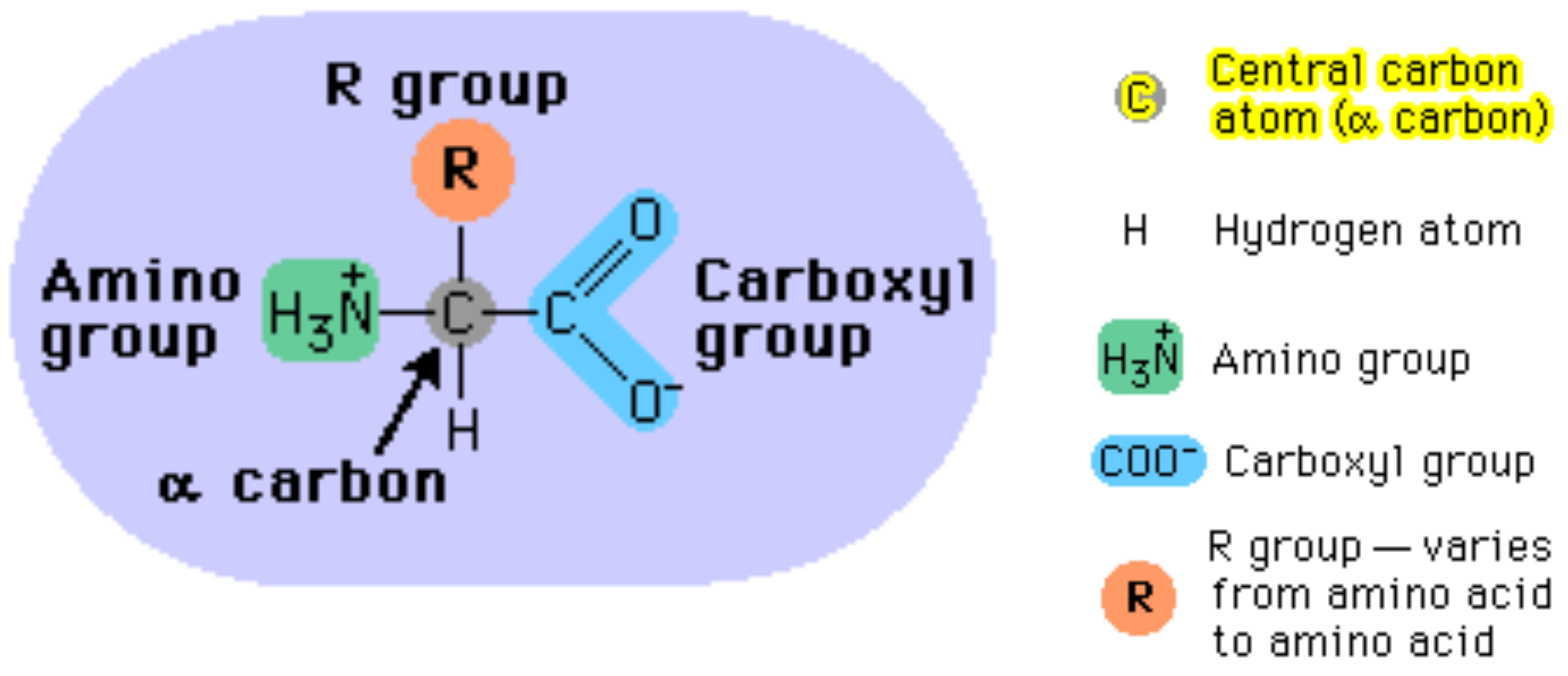
Proteins

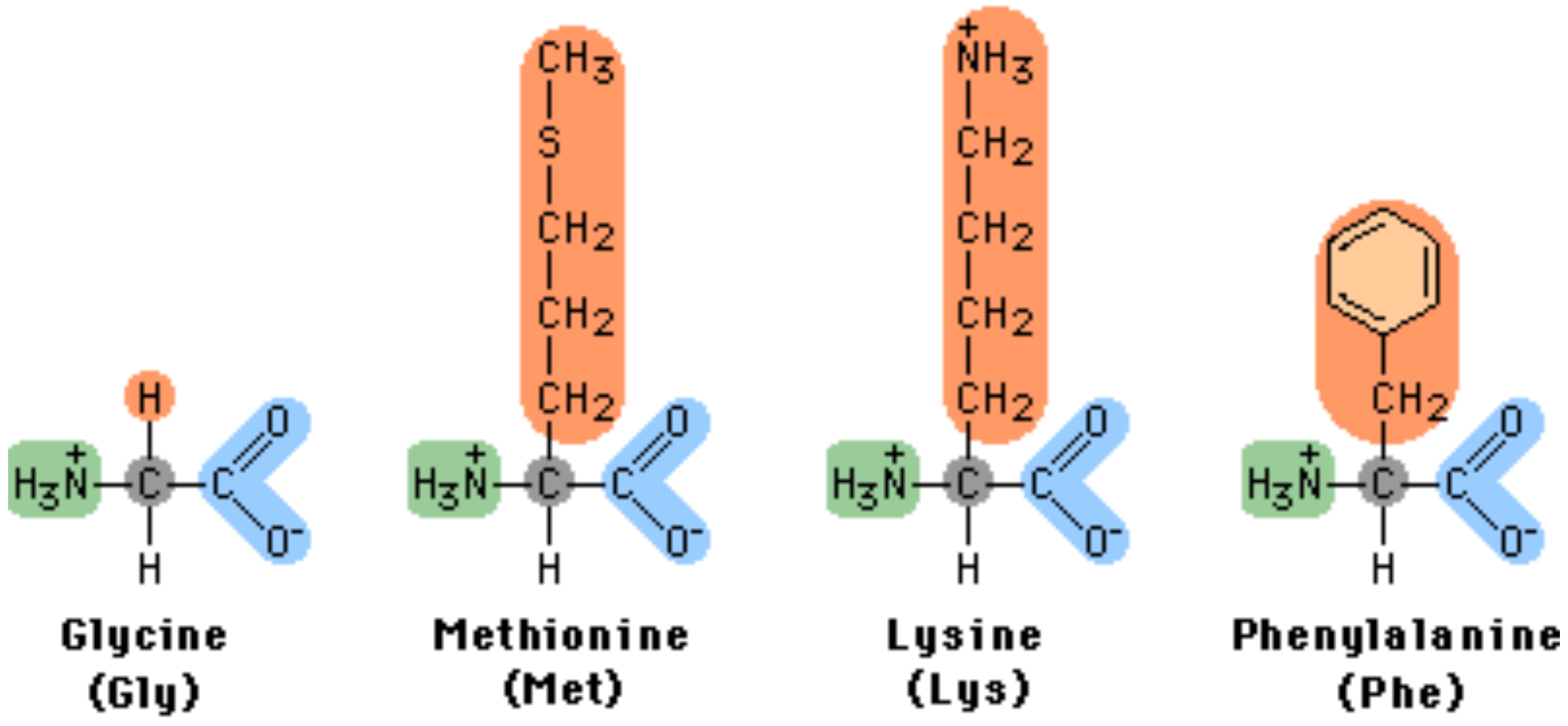
PROTEINS

- Proteins are large, complex molecules made of amino acid residues
- Workhorse of cells
- Required for the structure, function, and regulation of the body's tissues and organs

PROTEINS

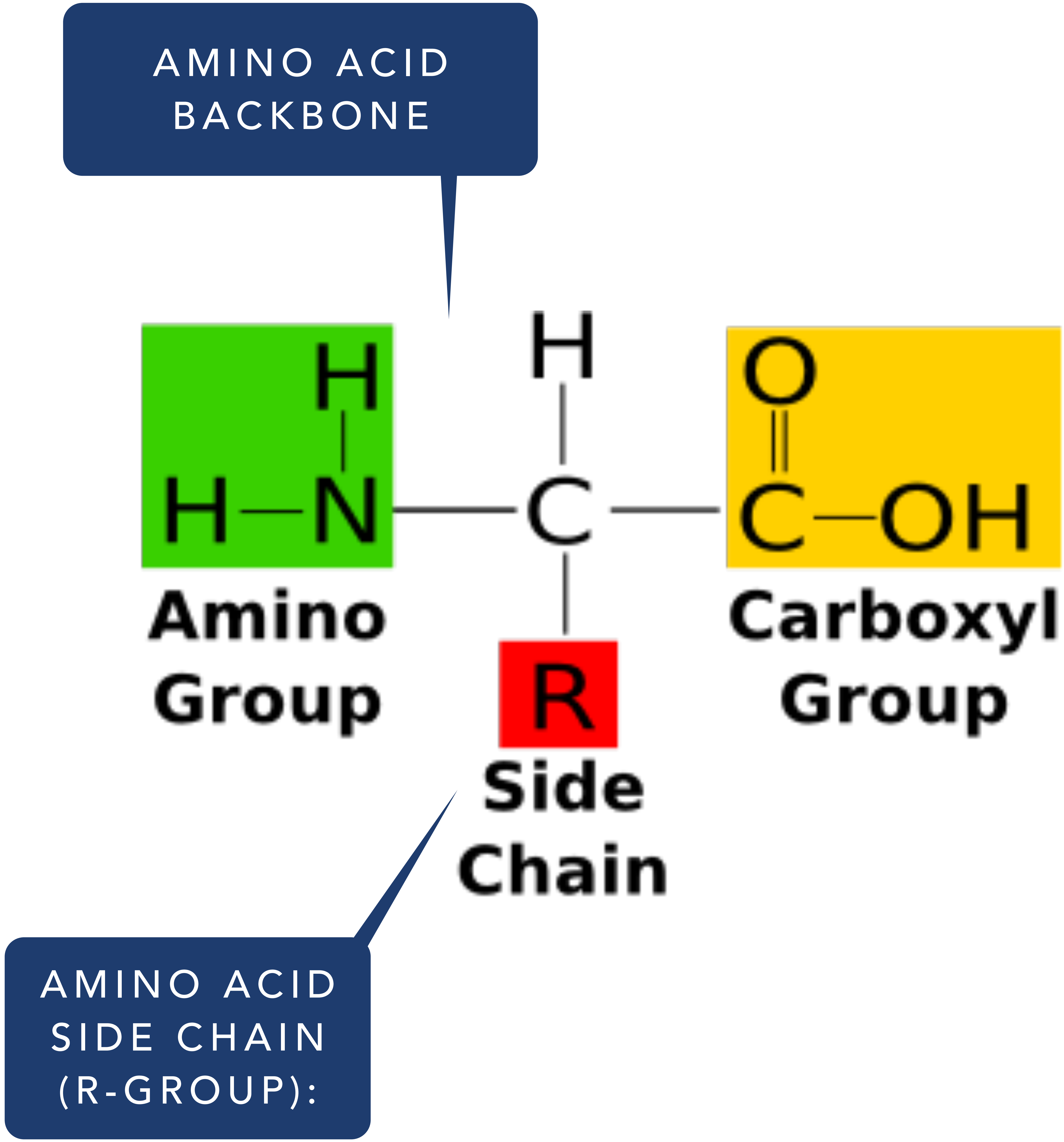
- Proteins structure
  - Building blocks are amino acids
  - There are 20 standard amino acids
    - Differ only by R group





PROTEINS

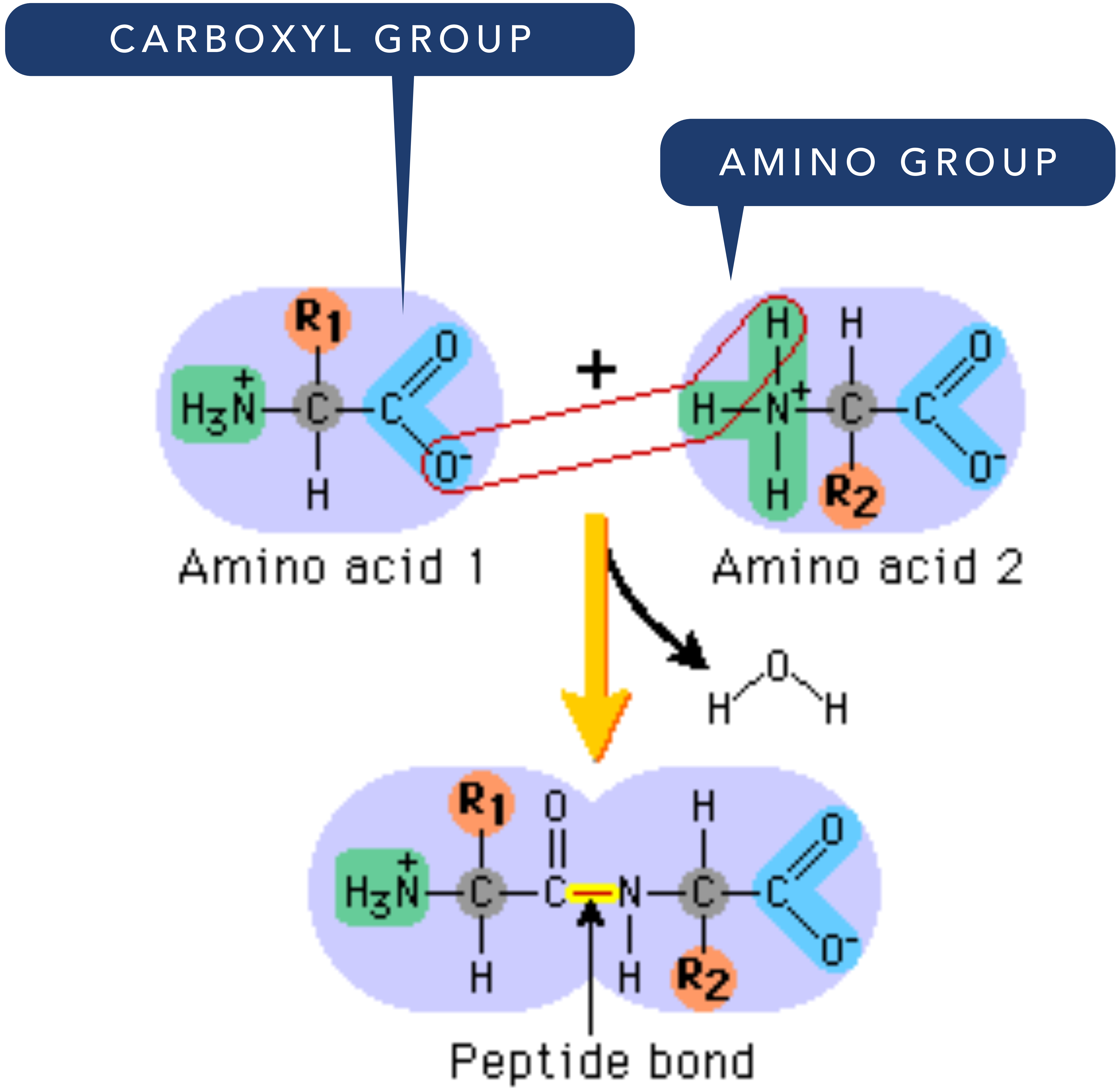
- Amino acid side chains
- Differ by their side chains (R-groups)
- Determine chemical properties



PROTEINS

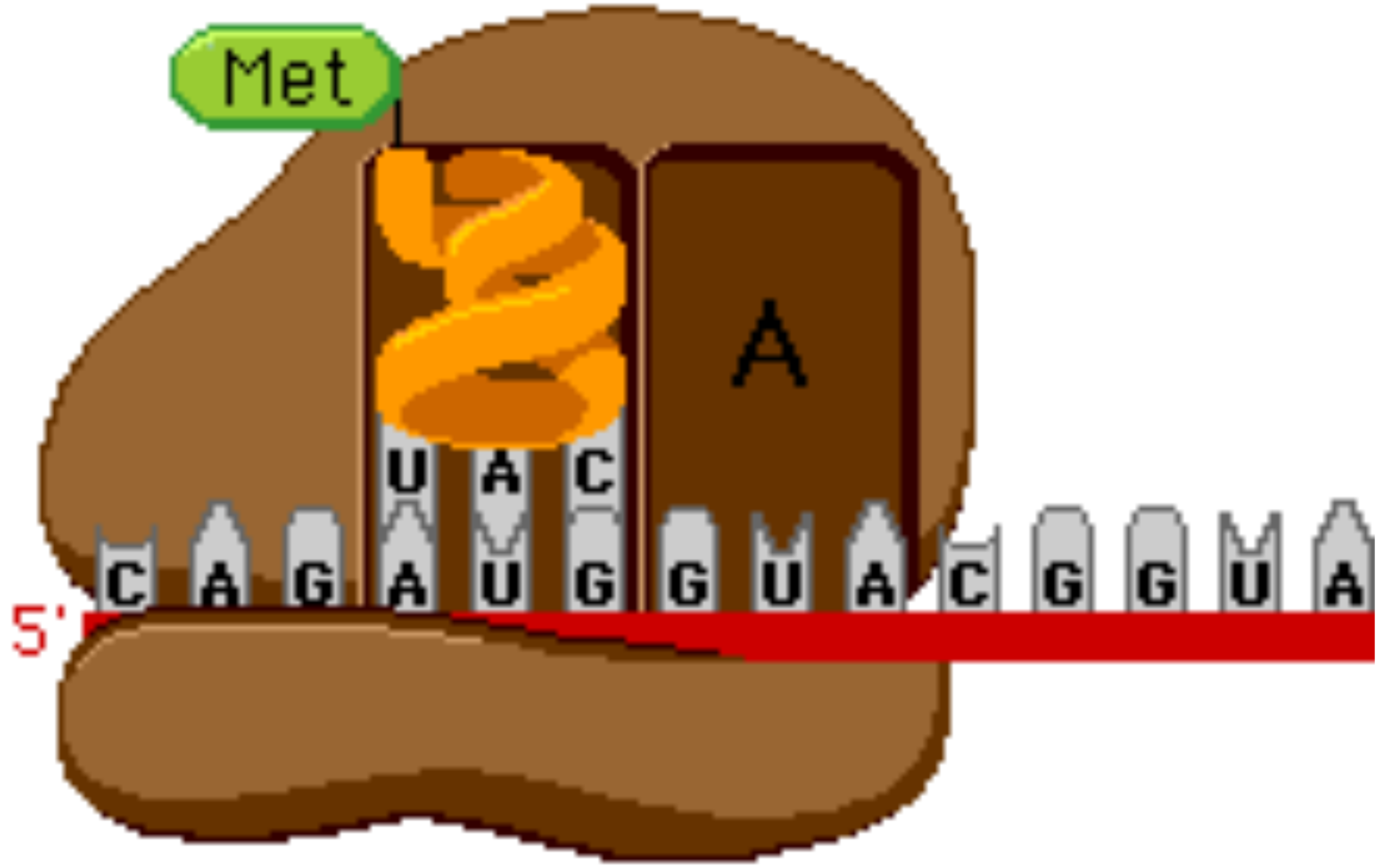
PROTEINS

- The Peptide Bond
  - Joins amino acids in proteins
  - Forms between the carboxyl group of one amino acid and the amino group of the adjacent amino acid



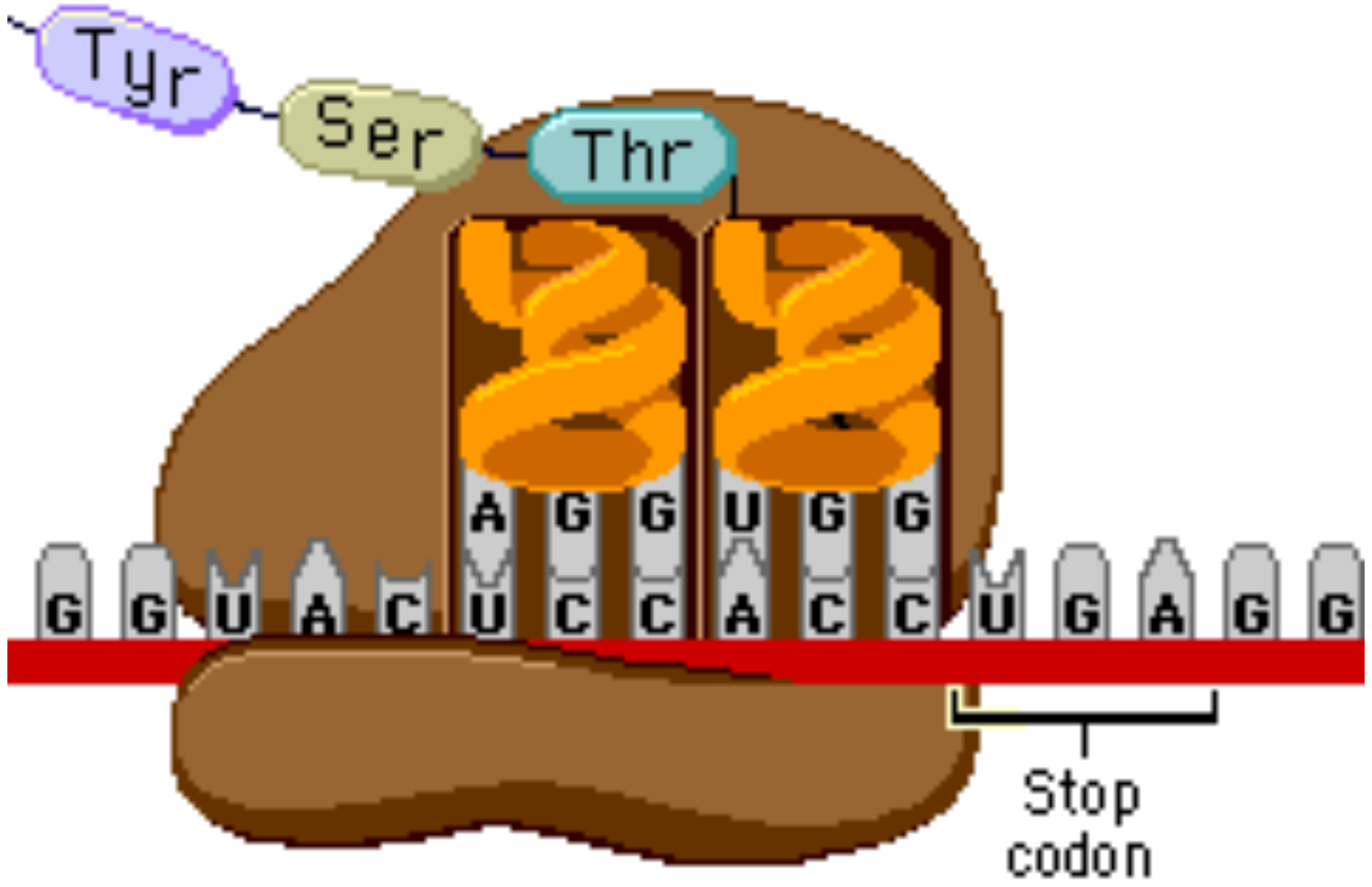
PROTEINS

- Elongation of the peptide chain
  - tRNA anti-codon recognized the complementary mRNA codon
  - Amino acids are joined by peptide bond



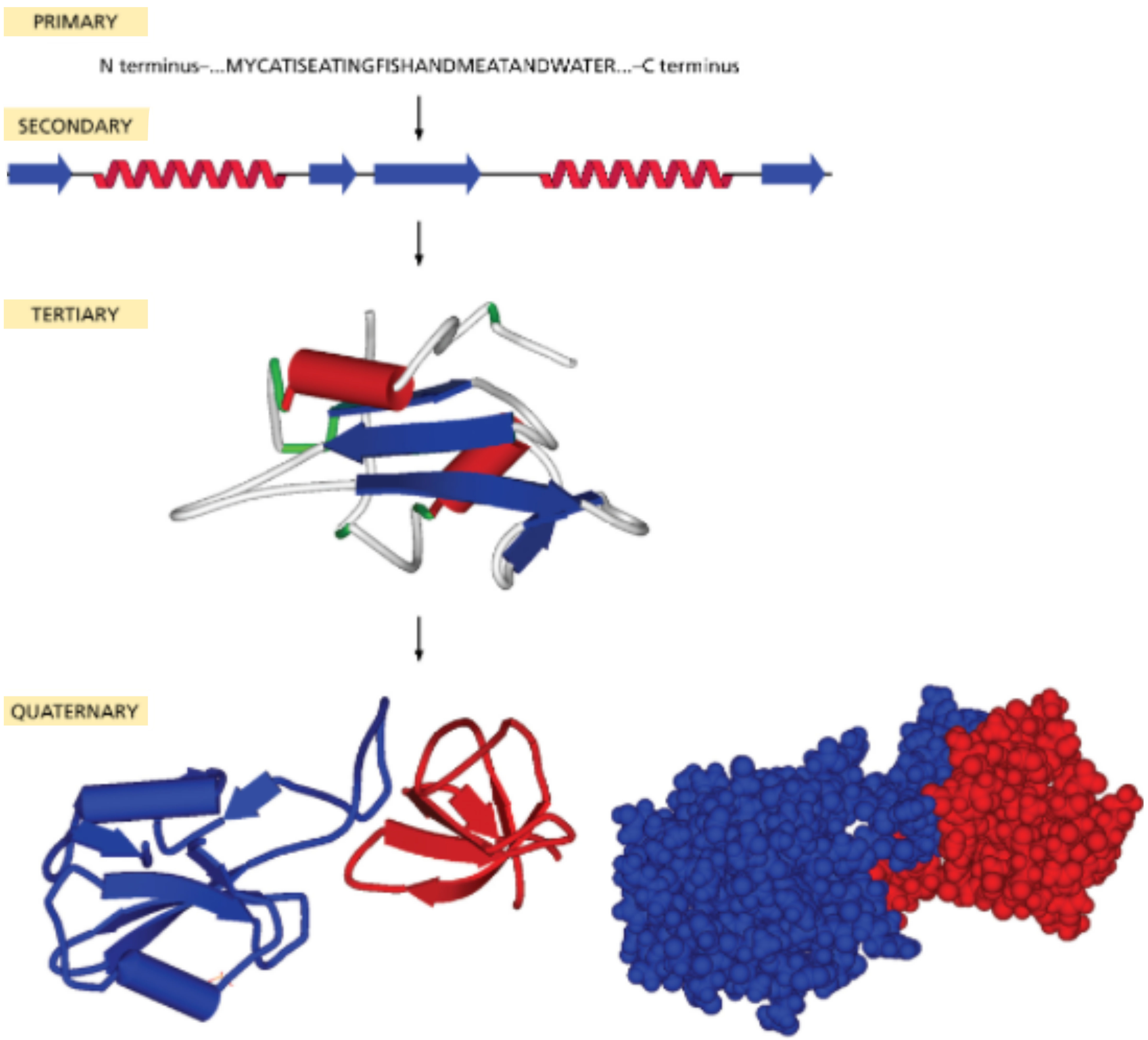
PROTEINS

- Termination of the peptide chain
  - Release factor - protein that recognizes the stop code in an mRNA sequence



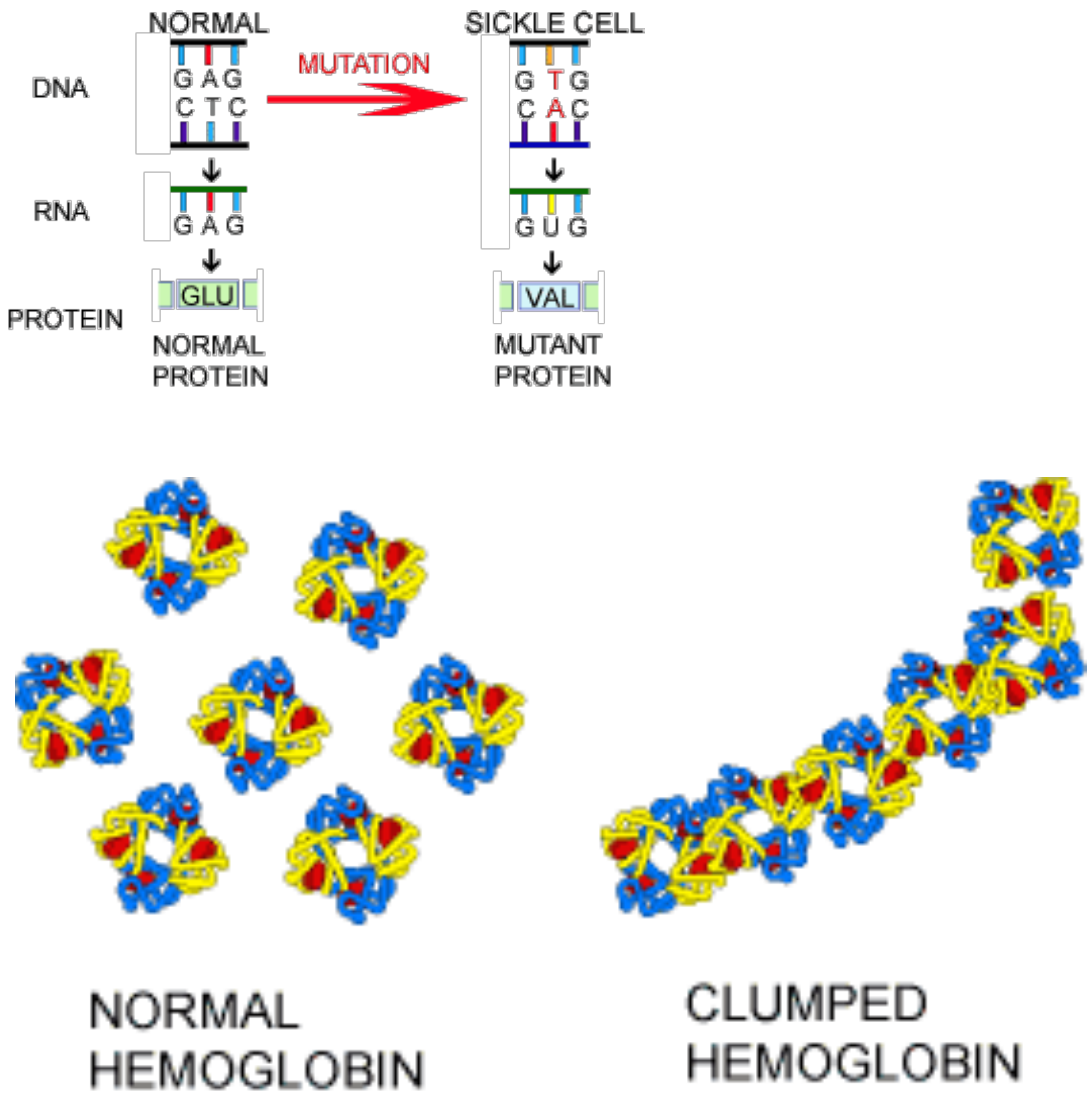
PROTEINS

- Primary sequence
- Secondary structure
- Tertiary structure
  - 3 dimensional
  - X-ray crystallography
- Quarternary
  - Assembled molecular units



PROTEINS

- Sickle cell anemia
  - Mutation in hemoglobin
    - Carries oxygen in red blood cells
  - Deprived of oxygen cells become sickle-shaped
  - Carrier experiences pain and fatigue
  - Carrier are resistant to malaria
    - Parasites are killed in blood cells



Mutations in DNA

MUTATIONS

- Change of nucleotide or amino acid at a given position
- Driven by evolution or environmental factors

TRANSCRIPTIONS & TRANSLATION

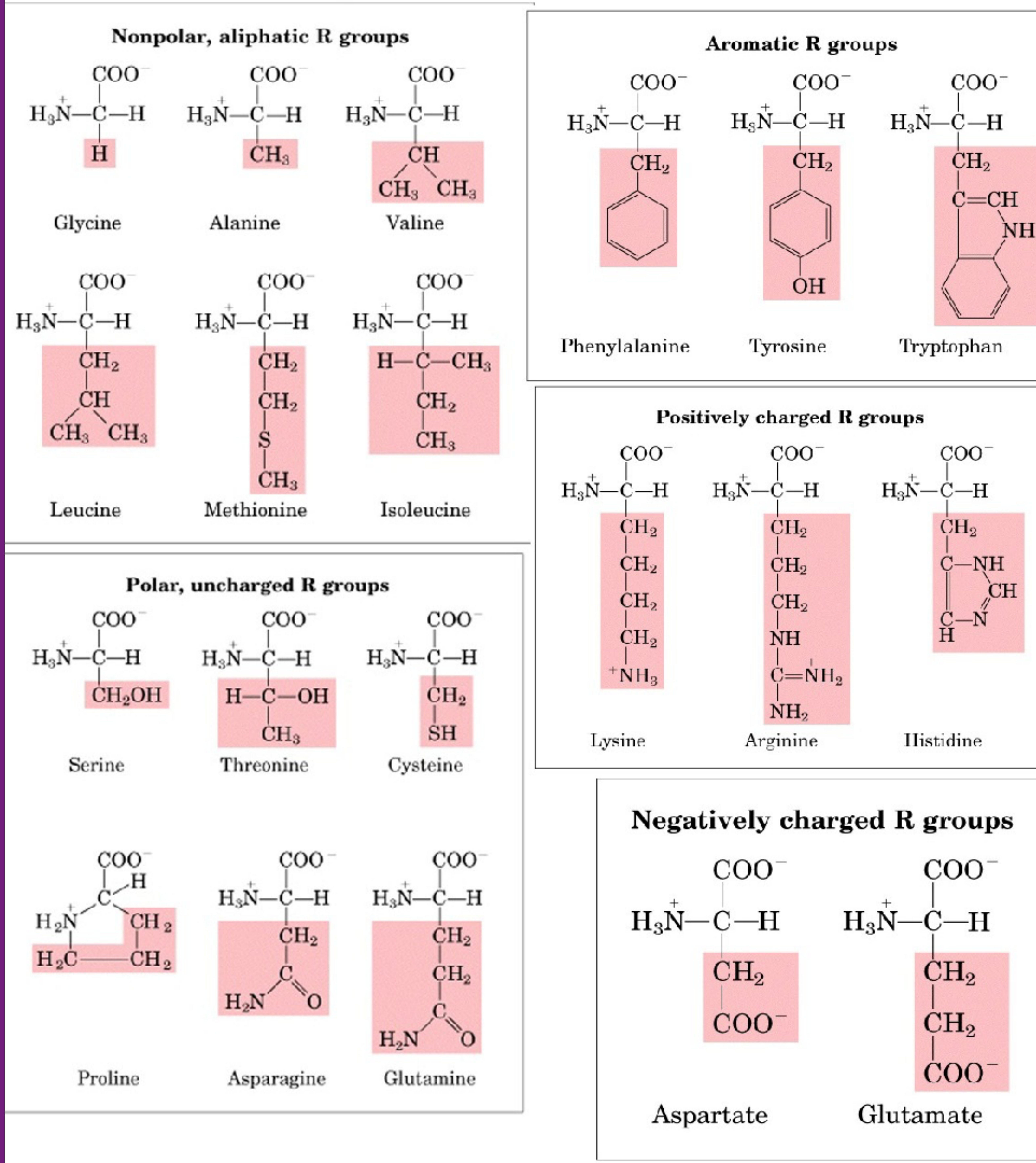
- Mutations in DNA
  - Substitution
    - Non-synonymous mutation - amino acid sequence is changed (“sense”)
    - Synonymous mutation - do not change an amino acid sequence (“silent mutation”)
    - Nonsense mutation - base changes to stop codon, premature termination of translation
  - Insertion
  - Deletion
  - Frameshift

Second Position			
U	C	A	G
UUU F UUC F UUA L UUG L	UCU S UCC S UCA S UCG S	UAU Y UAC Y UAA stop UAG stop	UGU C UGC C UGA stop UGG W
CUU L CUC L CUA L CUG L	CCU P CCC P CCA P CCG P	CAU H CAC H CAA Q CAG Q	CGU R CGC R CGA R CGG R
AUU I AUC I AUA I AUG M	ACU T ACC T ACA T ACG T	AAU N AAC N AAA K AAG K	AGU S AGC S AGA R AGG R
GUU V GUC V GUA V GUG V	GCU A GCC A GCA A GCG A	GAU D GAC D GAA E GAG E	GGU G GGC G GGA G GGG G

http://www.nature.com/scitable/topicpage/genetic-mutation-441

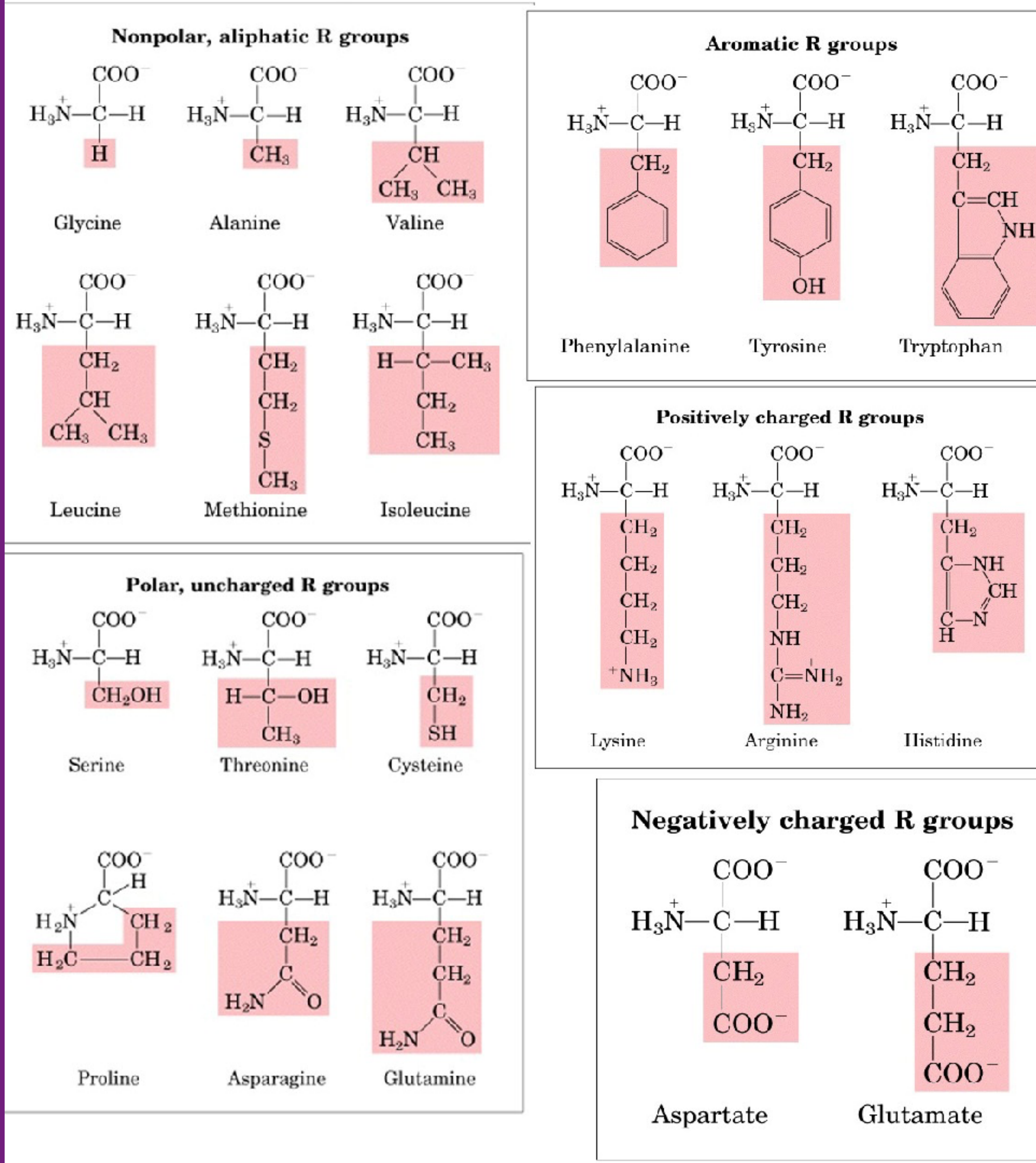
MUTATIONS

- Amino acids differ by their side chains (R-groups)
- Grouped by properties
  - Hydrophobic or non polar
    - which are amino acids with side chains that repel water
  - Hydrophilic or polar
    - which are amino acids with side chains that are attracted to water
  - Acidic or negatively-charged
    - Amino acids that contain carboxyl groups (-COO-) as side chains
  - Basic or positively-charged side chains
    - Amino acids that contain amine groups (-NH3+) as side chains



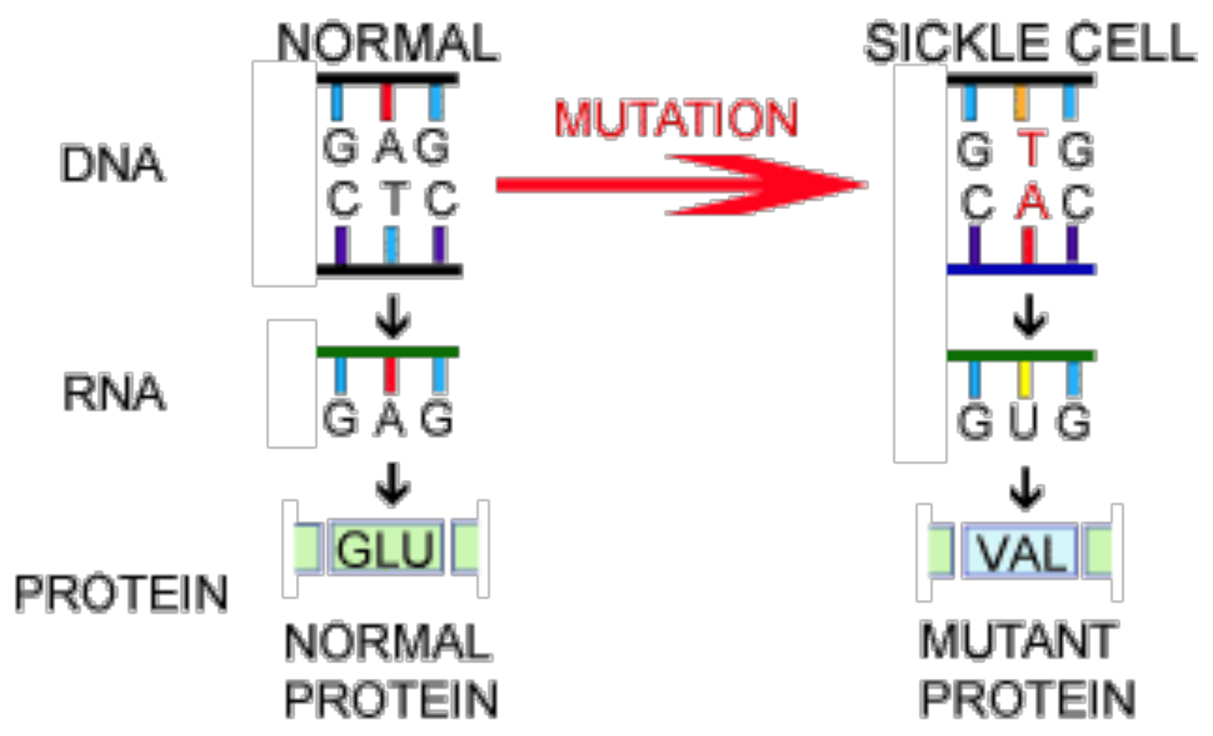
MUTATIONS

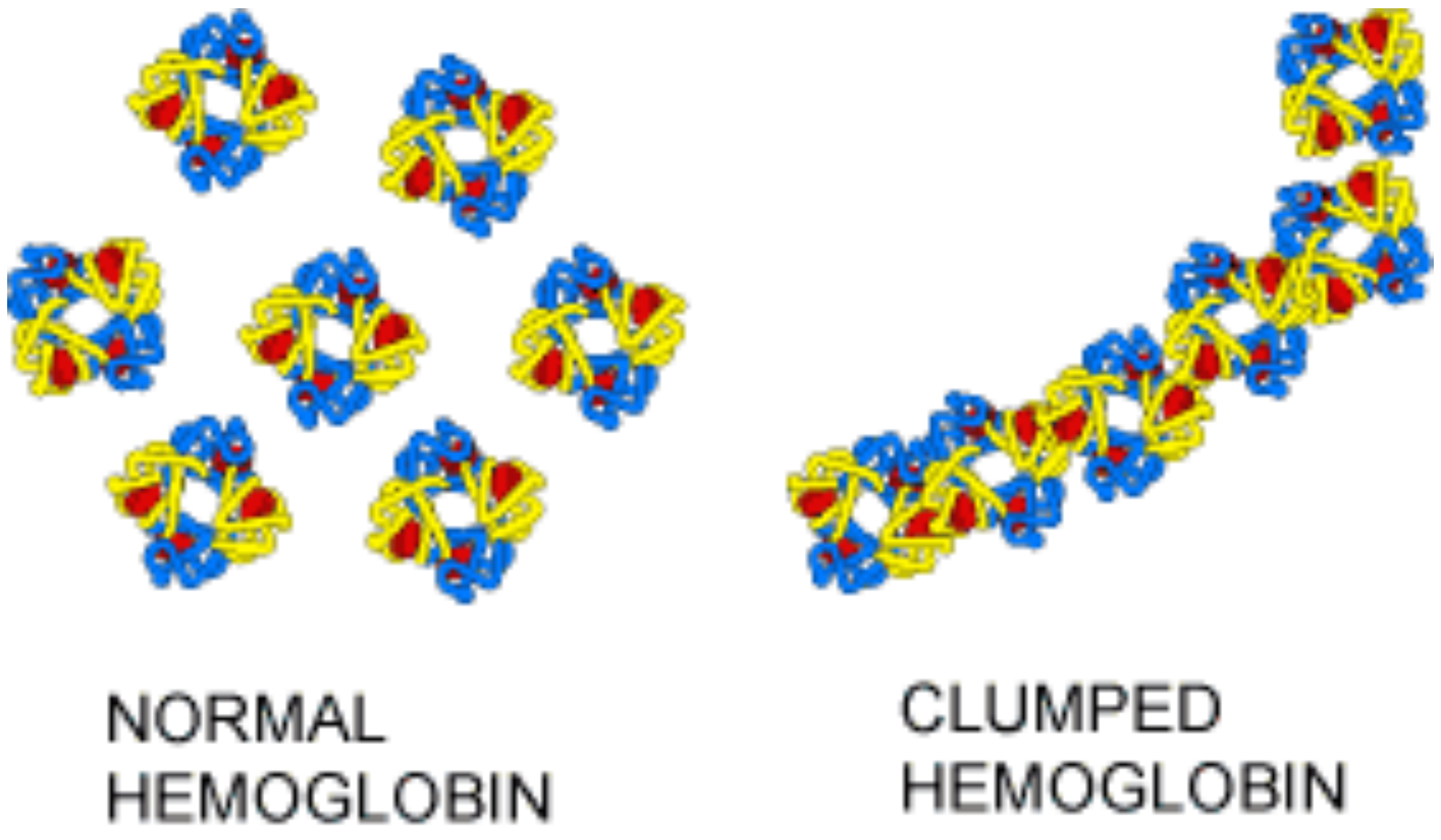
- Mutations in protein sequence
  - Non-conserved
    - results in different side chain
  - Conserved
    - results in same side chain



TRANSCRIPTIONS & TRANSLATION

- Sickle cell anemia
  - Mutation in hemoglobin
    - Carries oxygen in red blood cells
  - Deprived of oxygen cells become sickle-shaped
  - Carrier experiences pain and fatigue
  - Carrier are resistant to malaria
    - Parasites are killed in blood cells





BIOINFORMATICS
(FOR COMPUTER SCIENTISTS)

MPCS56420
AUTUMN 2020
SESSION 1A

THE UNIVERSITY OF CHICAGO