Building blocks of life MSE 493

Prof. Tiffany Abitbol

2024

Last week

- If you missed it – it's on Moodle, ask me, a friend, a TA…

- ELMs are:

Living cells in a hospitable environment that sustains life and provides a responsive function related to life

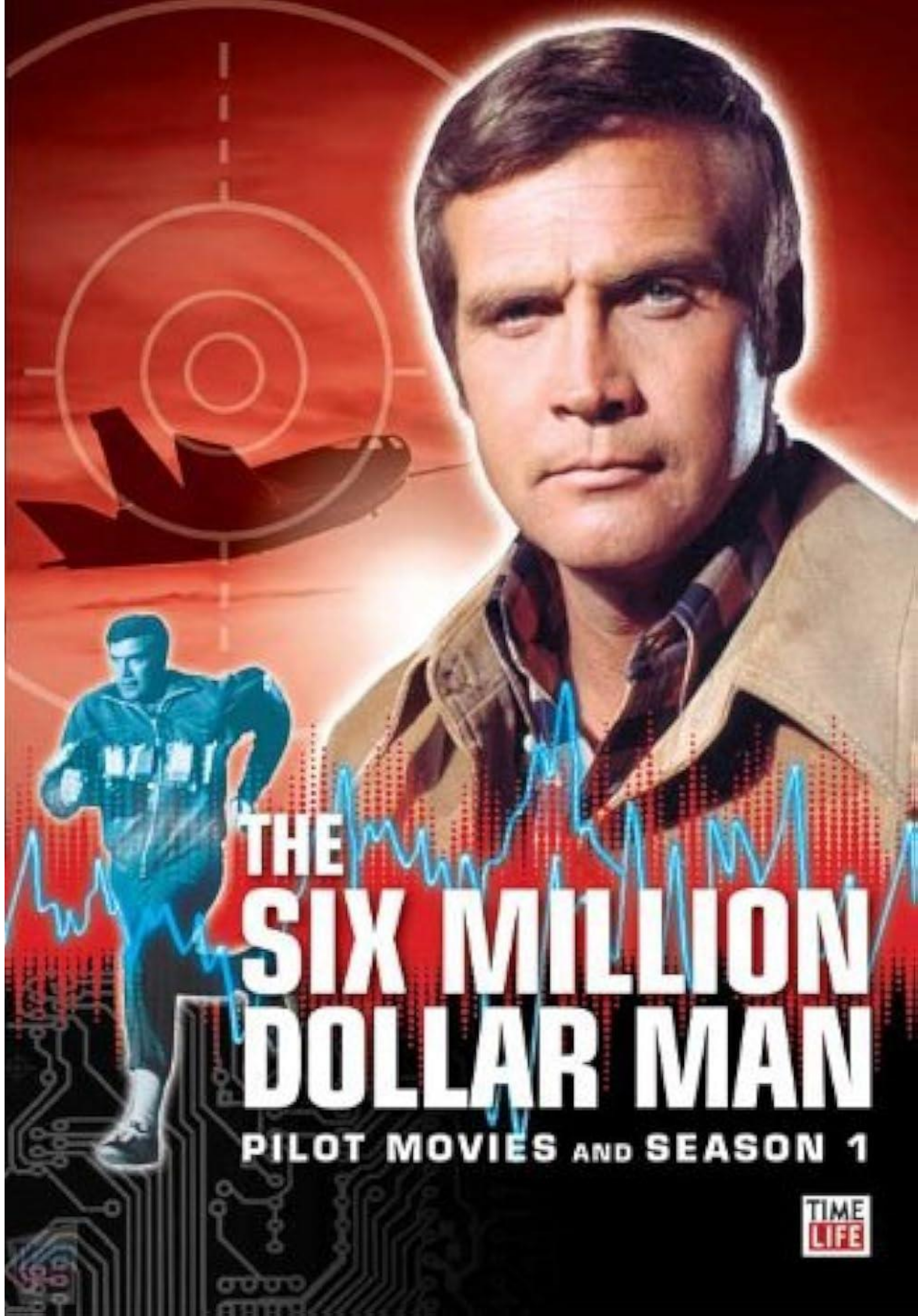
OR

A cell that produces a useful material

OR

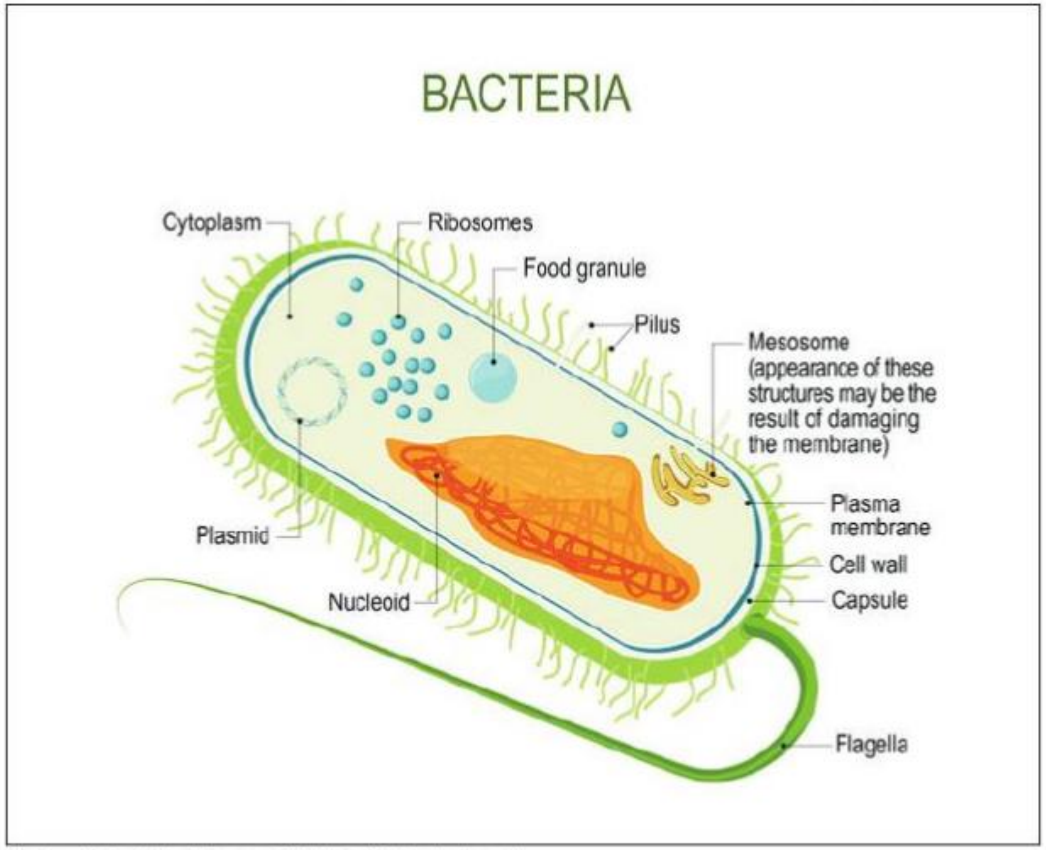
Fundamentally, any material with a biology part and a materials part

- Overview of grading scheme and course calendar
- Upcoming event: Lab activity on Sept 30, with reading homework/quiz uploaded on Sept 23, deadline Sept 30 at noon
- To biology or not? We decided to biology a bit…



Cells

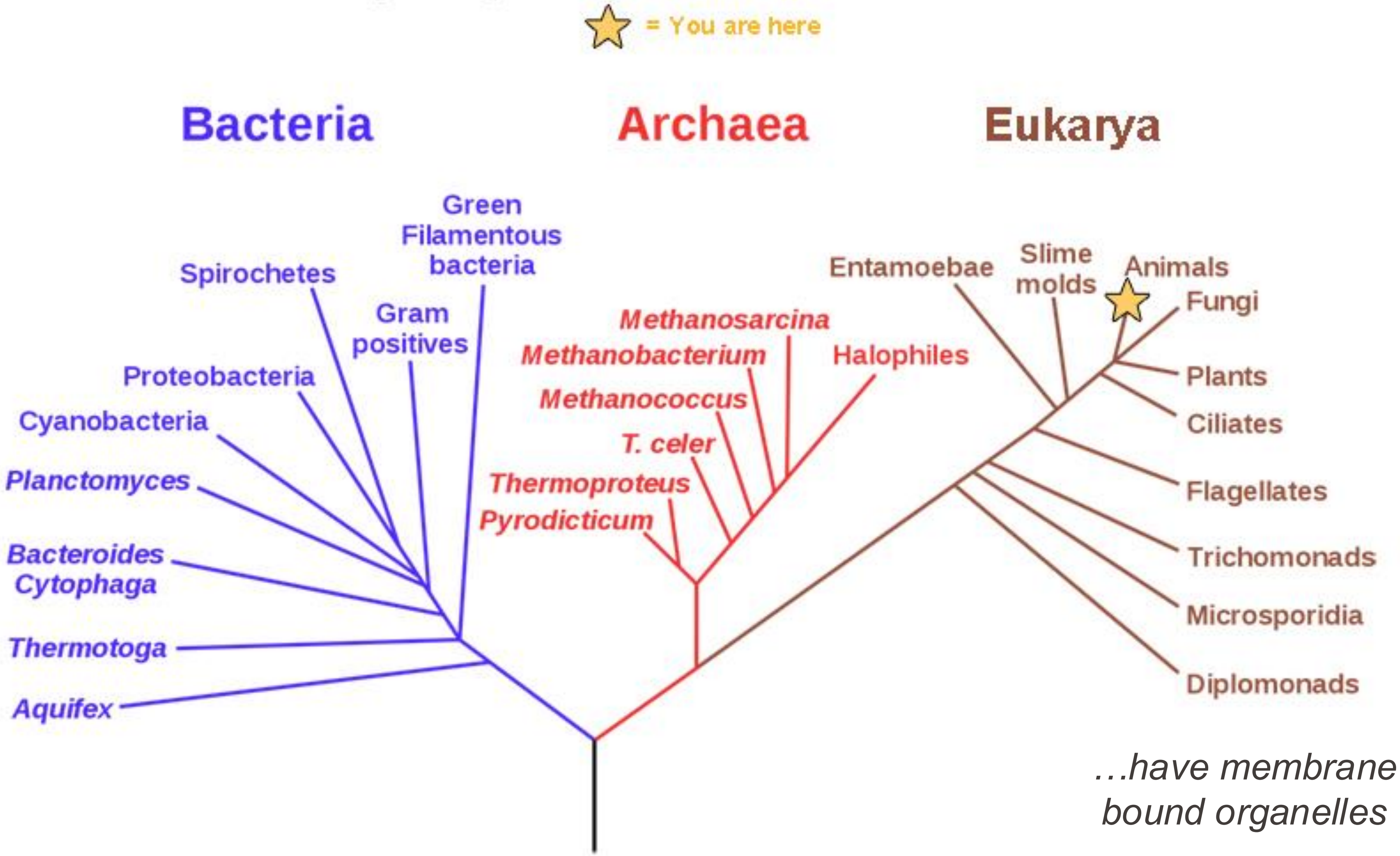
- Basic unit of organization in biological systems
- Many different types, but some common features
- Cell membrane
- Sometimes a cell wall (bacteria, yeasts, plant cells)
- Contains genetic material that provides instructions for function



Structure of a bacterial cell. Image via iStock/ttsz. [Used under license.]

https://doi.org/10.1017/CBO9780511800986.004

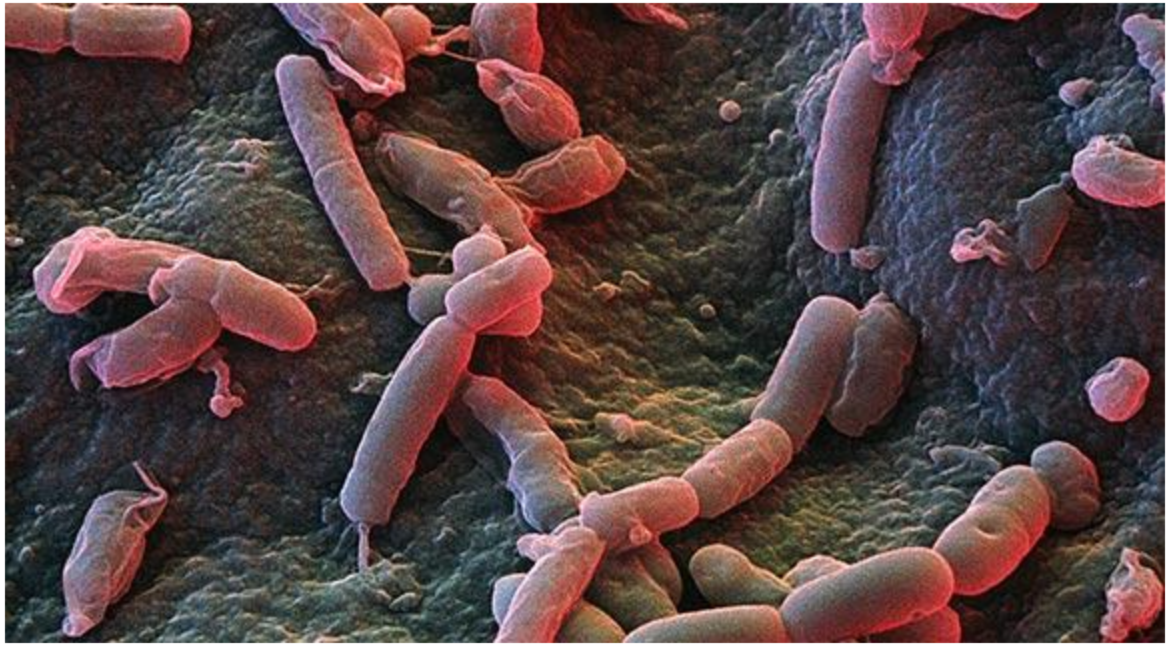
Phylogenetic Tree of Life



…have membrane bound organelles

[Three domains]()

Bacteria and Archaea



Prokaryotes:
- No nucleus, no organelles, single cell
- Genetic information (DNA) is not compartmentalized
- Bacterial cell 1-10 µm

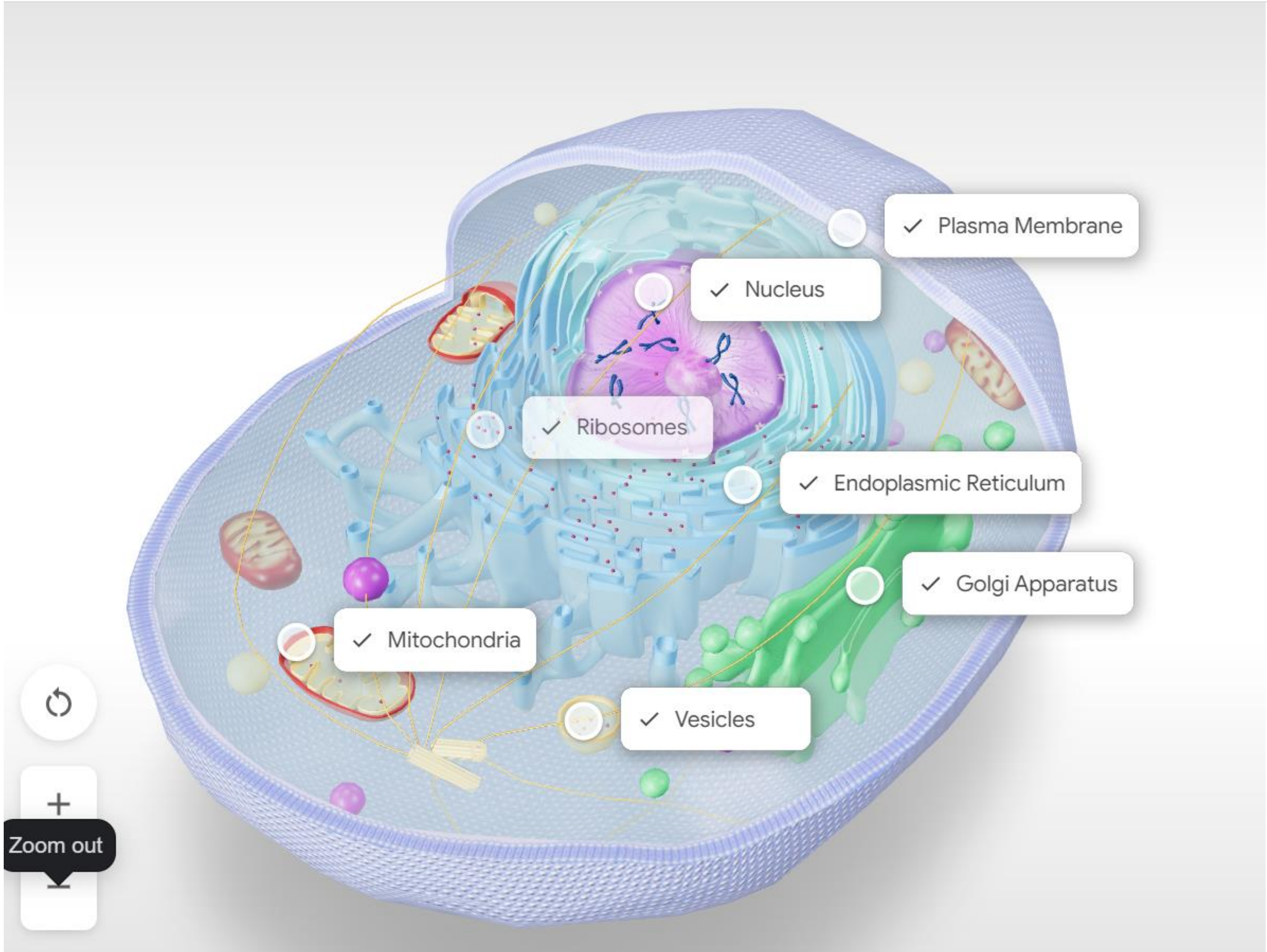
Three domains 2

Eukarya

Eukaryotes:
- Nucleus, organelles, multicellular, more complex
- DNA located in membrane-bound nucleus
- Plant/animal cells 10-100 µm

Eukaryote

From Visible Body



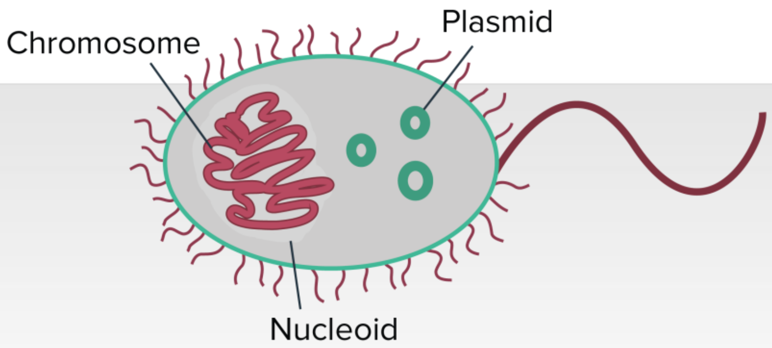
Notable mentions:

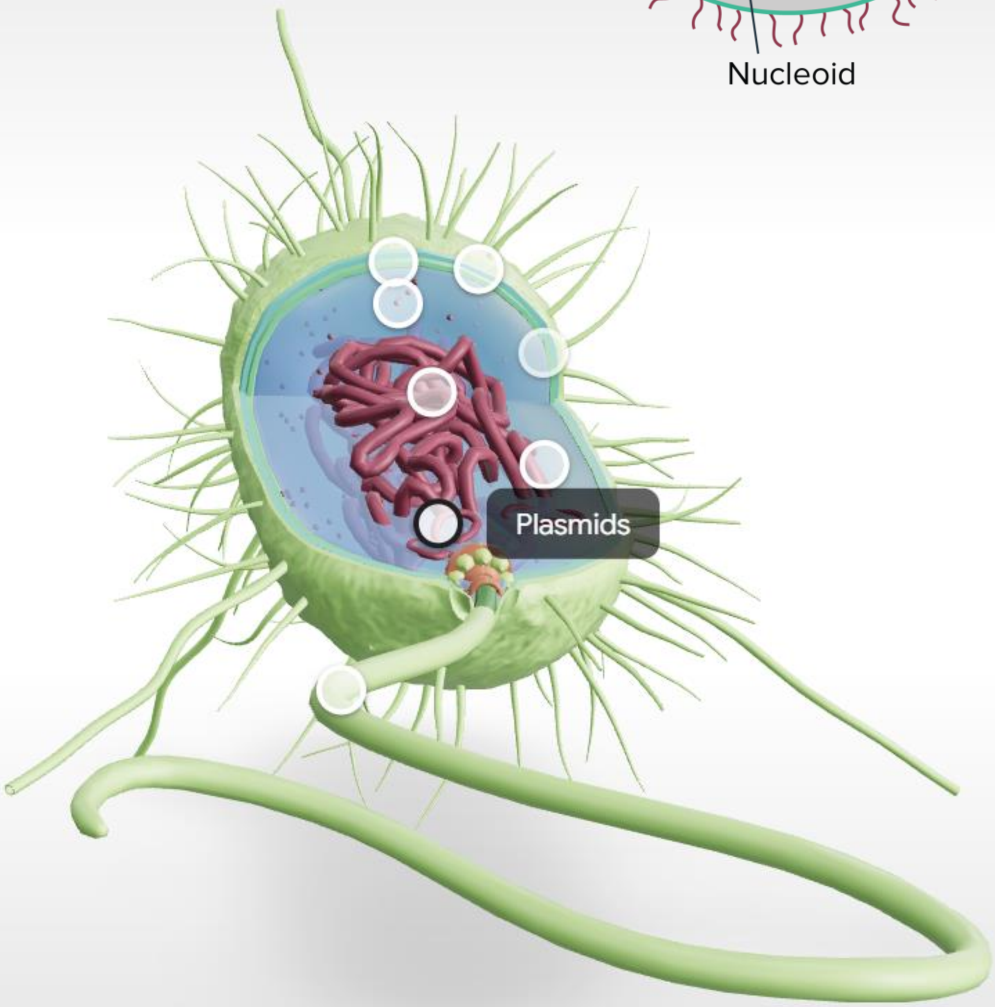
- Nucleus (chromosomes, DNA)
- Ribosomes (protein "factory")
- Mitochondria ("powerhouse" of cell) – cellular respiration where glucose and oxygen are converted to ATP – the "energy currency" of the cell

Prokaryote

From Visible Body

Read   Play





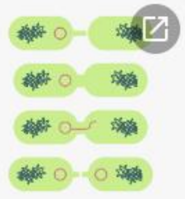
Fimbria

Flagellum

Nucleoid

Plasma Membrane

Plasmids



In addition to the DNA located in the nucleoid, bacterial cells have small, circular DNA molecules called plasmids.

Function: Plasmids carry nonessential genes that help bacteria survive in extraneous conditions.

Found in: Bacterial cells

Source Visible Body · Natallia Yatskova

Plasmids

Biological molecules

Banana
Fruit ⋮

Nutrition facts | Recipes | Types

Bananas
Sources include: USDA

Amount Per 100 grams			
Calories 89			
			% Daily Value*
Total Fat 0.3 g			0%
Saturated fat 0.1 g			0%
Cholesterol 0 mg			0%
Sodium 1 mg			0%
Potassium 358 mg			10%
Total Carbohydrate 23 g			7%
Dietary fiber 2.6 g			10%
Sugar 12 g			
Protein 1.1 g			2%
Vitamin C	14%	Calcium	0%
Iron	1%	Vitamin D	0%
Vitamin B6	20%	Cobalamin	0%
Magnesium	6%		

*Per cent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

Macronutrients to Micronutrients

- Lipids – fatty acids
- Carbohydrates – glucose
- Protein – amino acids

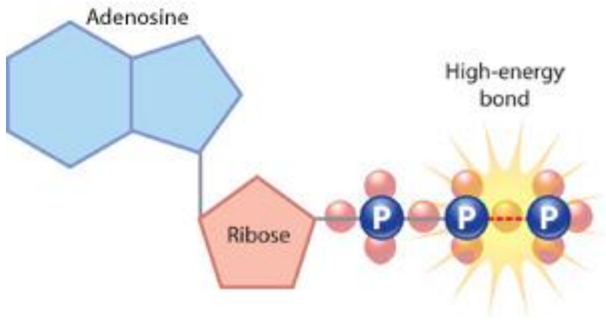
↓ *digestion*

Different cells metabolize the products of digestion to gain energy (ATP)

Krebs cycle

EPFL

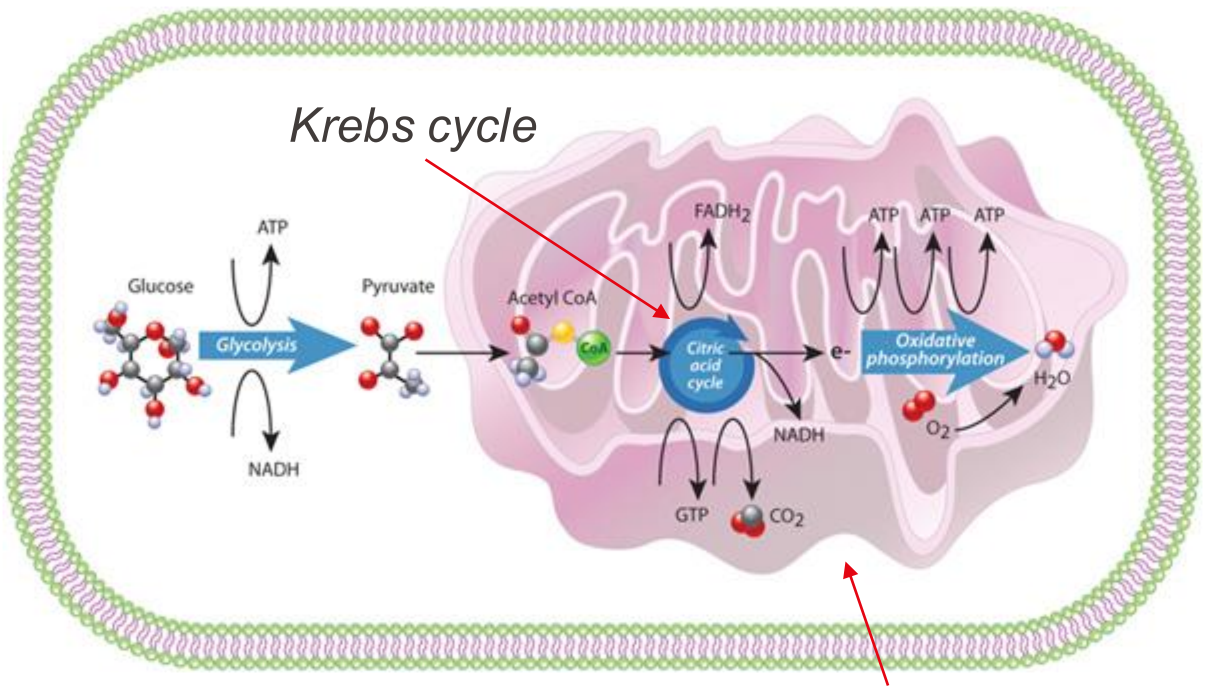
Glucose metabolism in eukaryotes



ATP: Adenosine 5'-triphosphate

- Cells transform the energy in the chemical bonds of food molecules to more readily usable forms – energy-rich carrier molecules
- ATP is the most abundant energy carrier molecule in cells

Glucose metabolism in a eukaryotic cell



Mitochondria

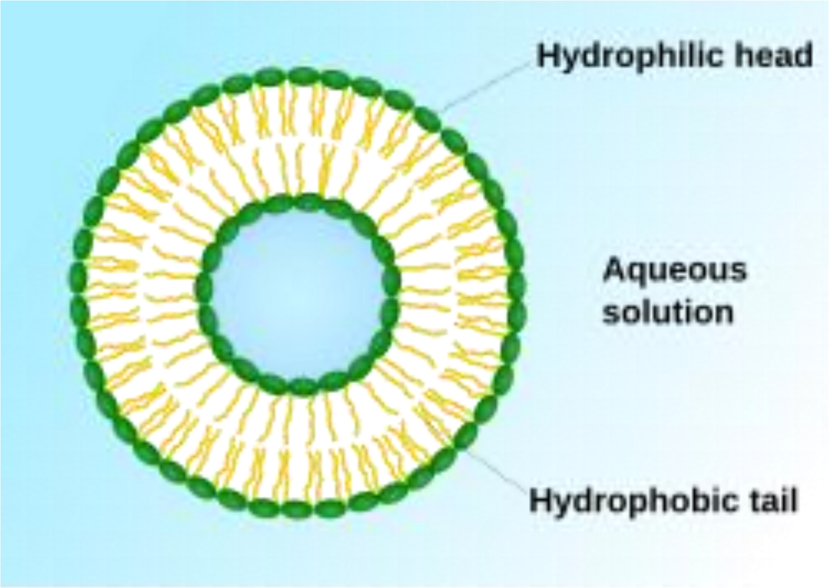
- *Cell's powerhouse*

1. Glycolysis to turn sugars into pyruvate
2. Krebs cycle
3. Oxidative phosphorylation

1 glucose molecule = 36-38 ATP molecules

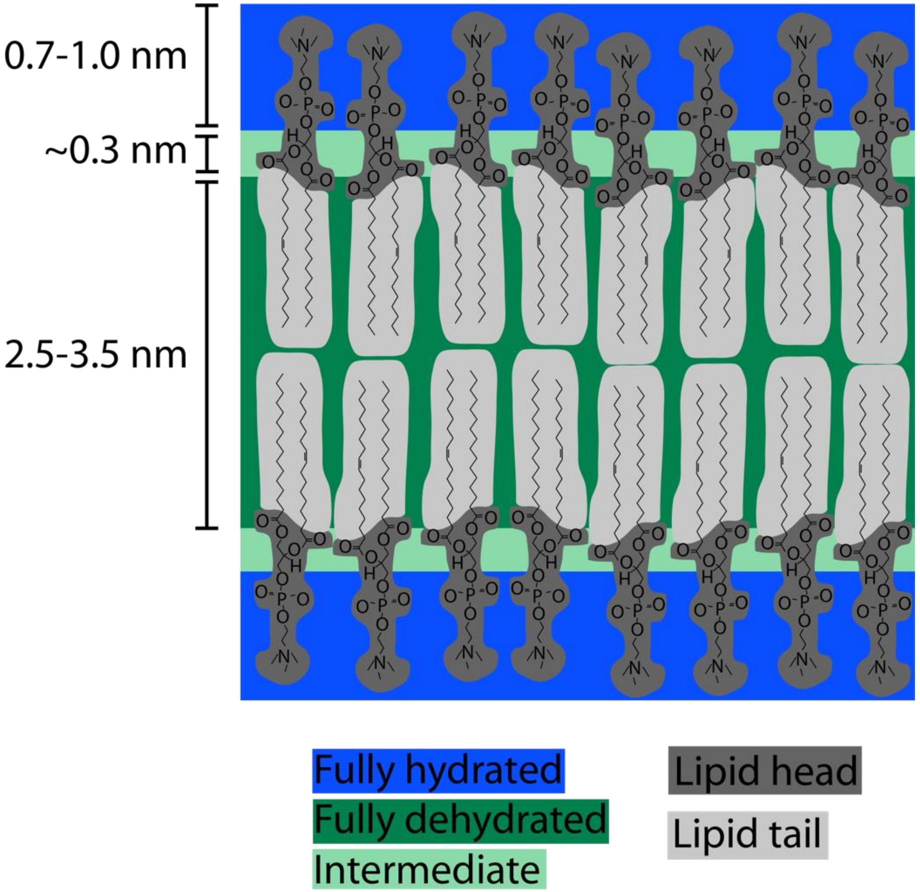
Cell energy

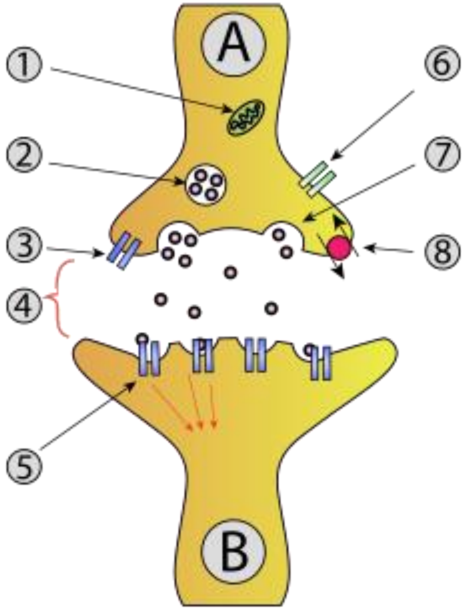
Lipids



Vesicle

Lipid bilayer

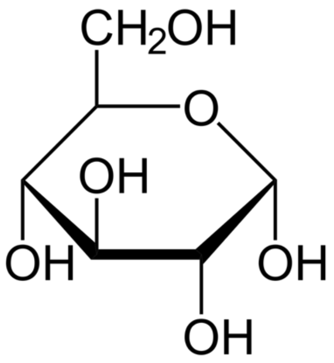




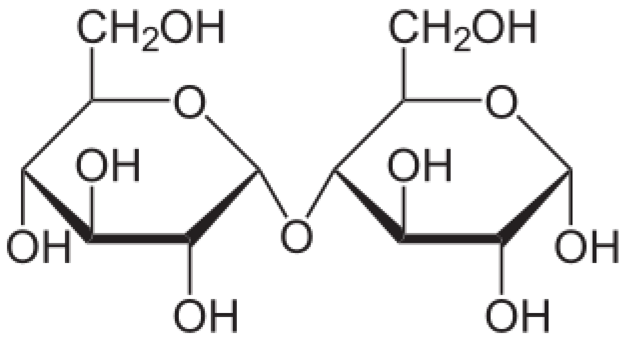
Synaptic vesicle

Carbohydrates

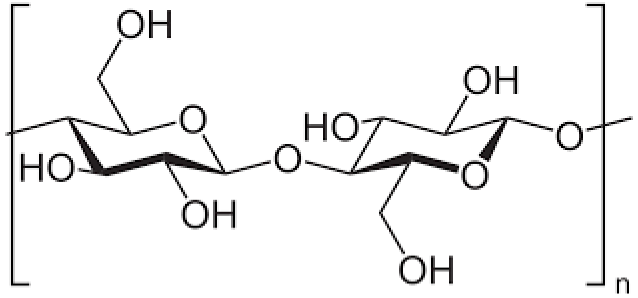
- $(CH_2O)_n$
- C:H:O = 1:2:1
- Subtypes: Monosaccharide, disaccharide, polysaccharide
- Link via dehydration



glucose



maltose



cellulose

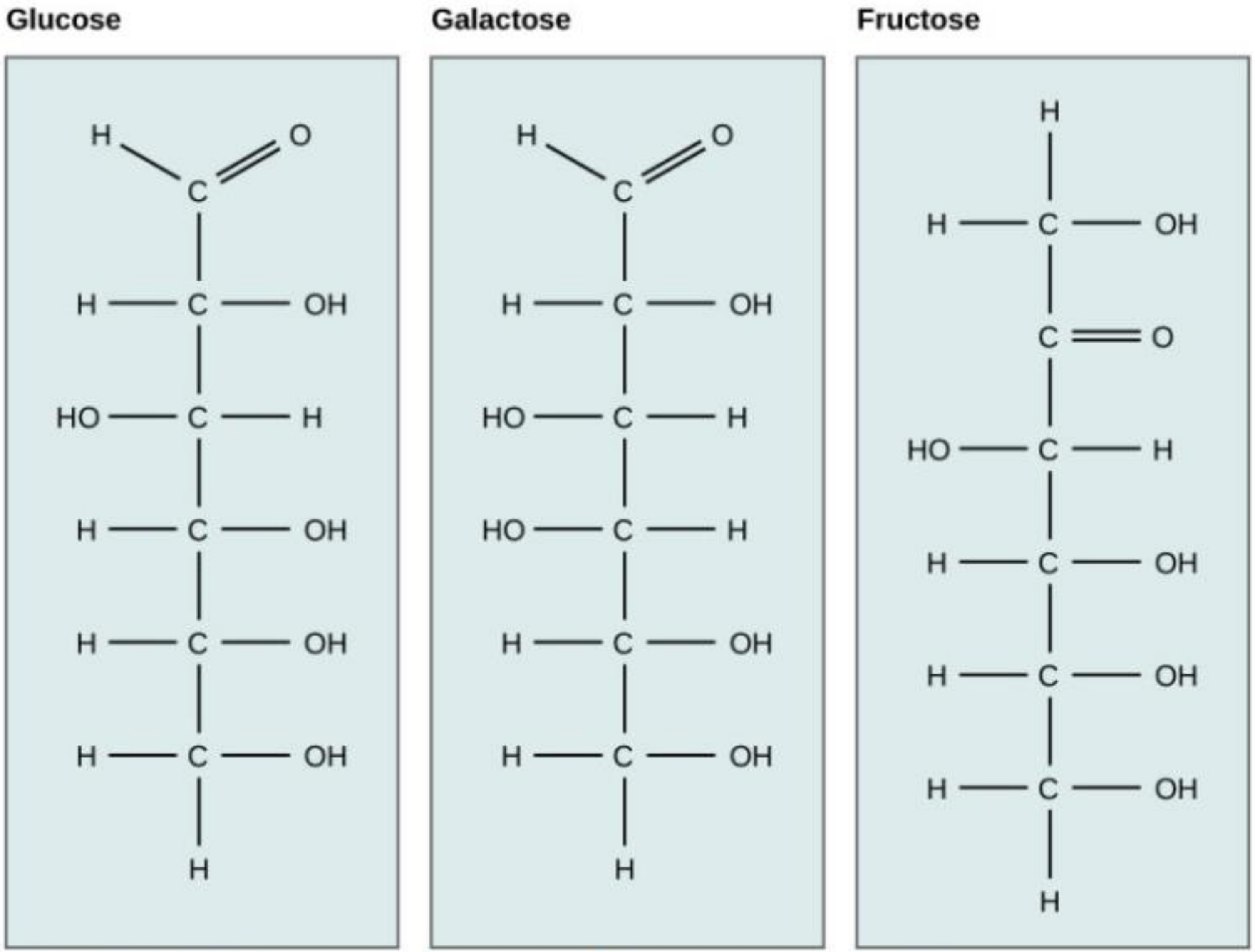
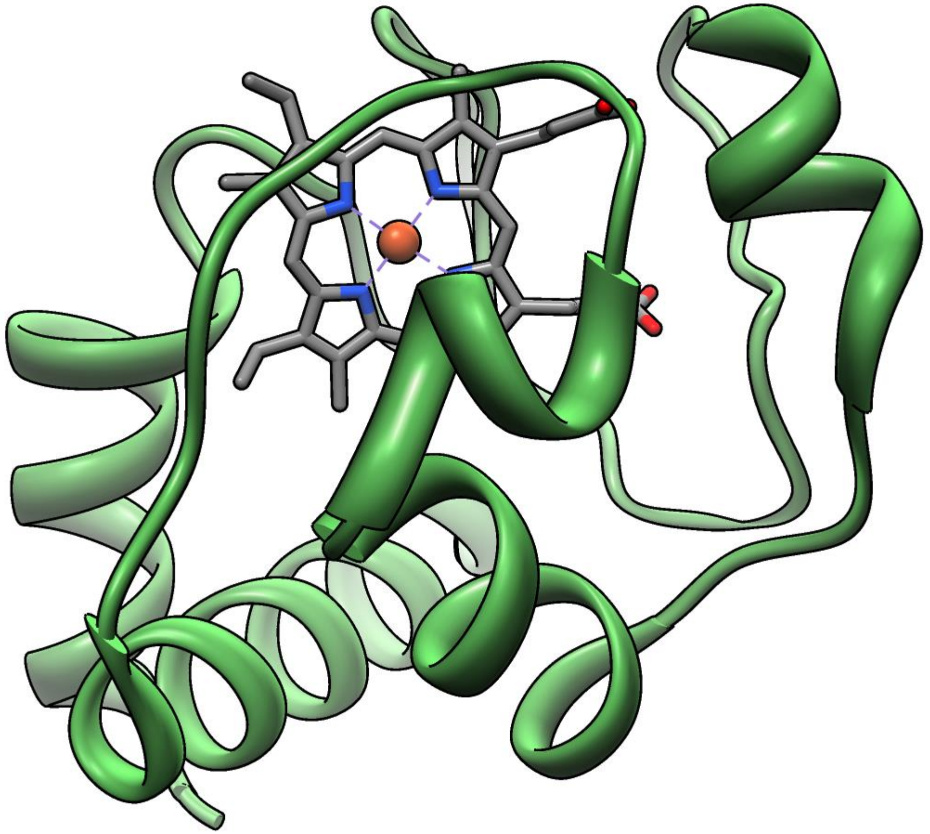


Figure 2.14 Glucose, galactose, and fructose are isomeric monosaccharides, meaning that they have the same chemical formula but slightly different structures.

[Biological molecules]

Proteins

- Expression of genetic information is achieved via proteins (especially enzymes)
- Living systems contain 1000s of different proteins, each with a unique function
- Transport, storage, membranes, toxins, enzymes, hormones
- Linear polymer of amino acids (AA's)
- Different molecular weights
- Different shapes, e.g., globular or fibrillar
- Shape is critical to function (determined by sequence and number of AA's)

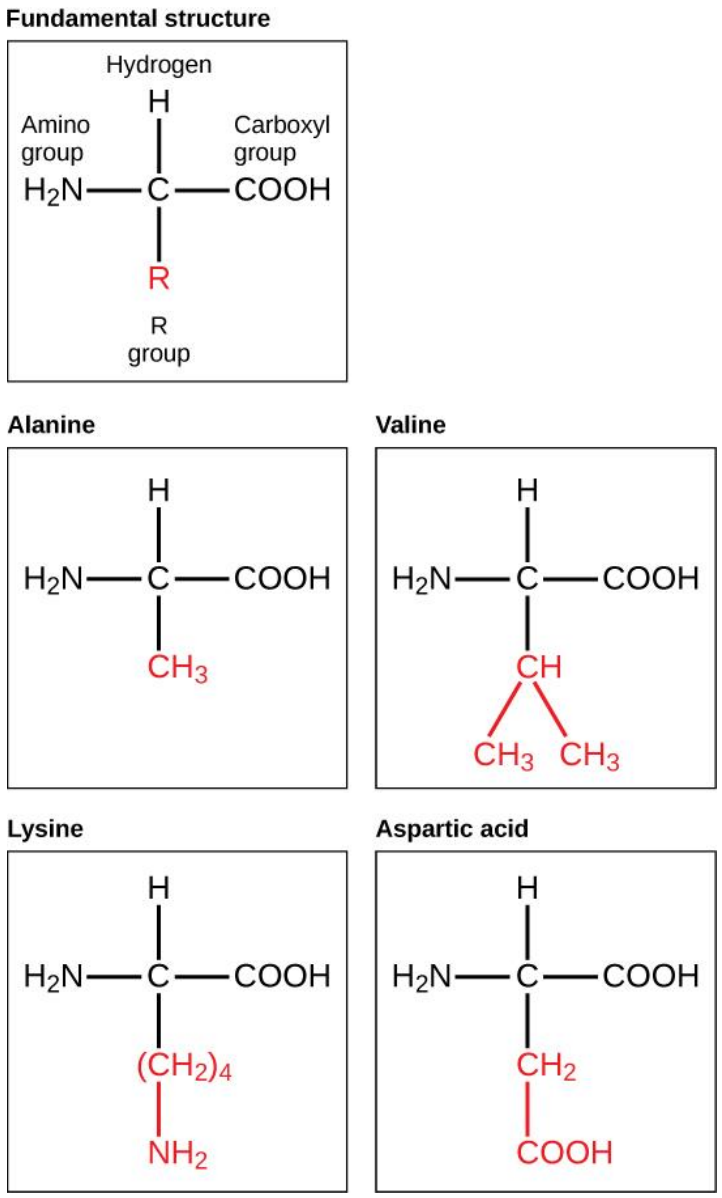


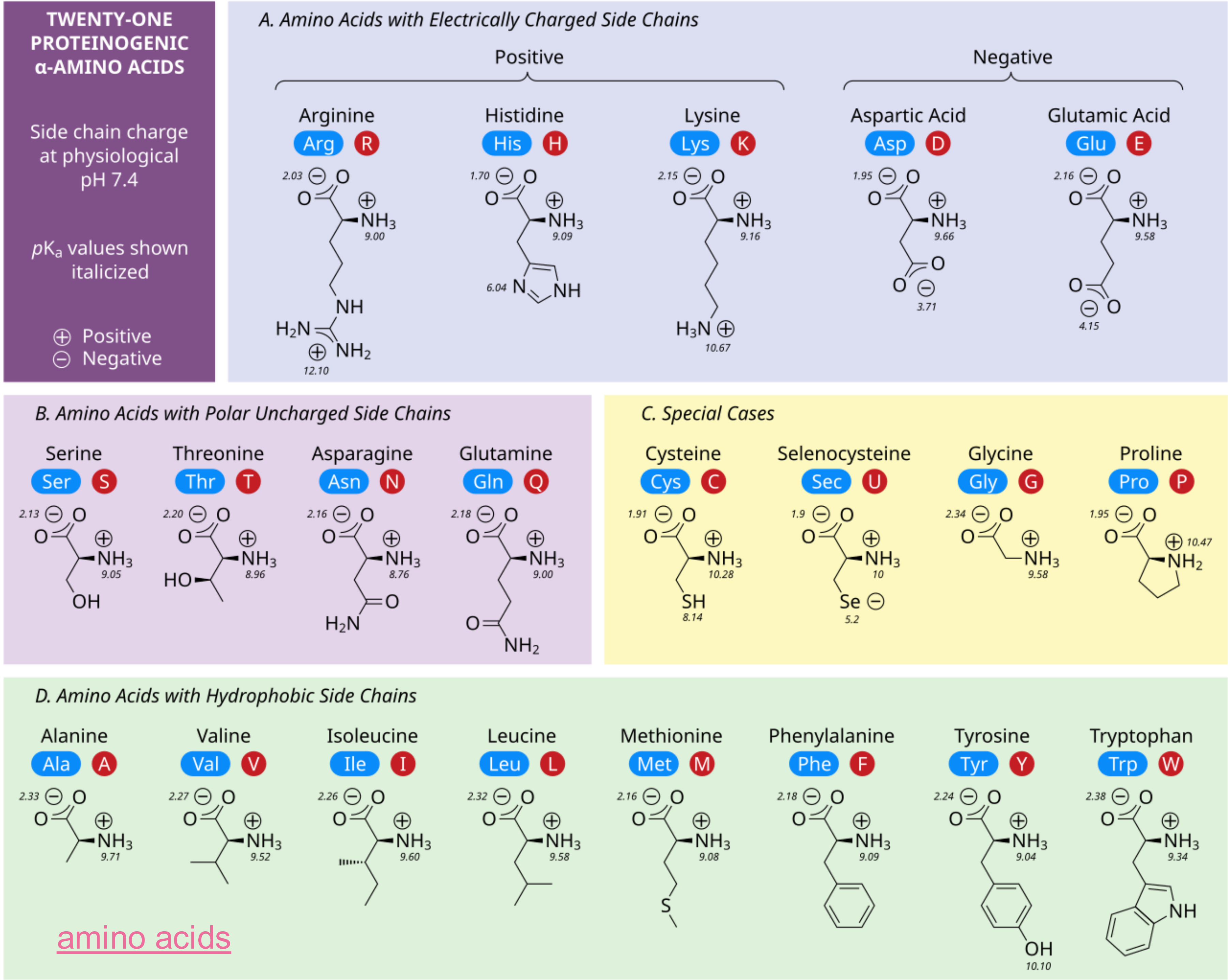
Cytochrome C

Biological molecules

Amino acids

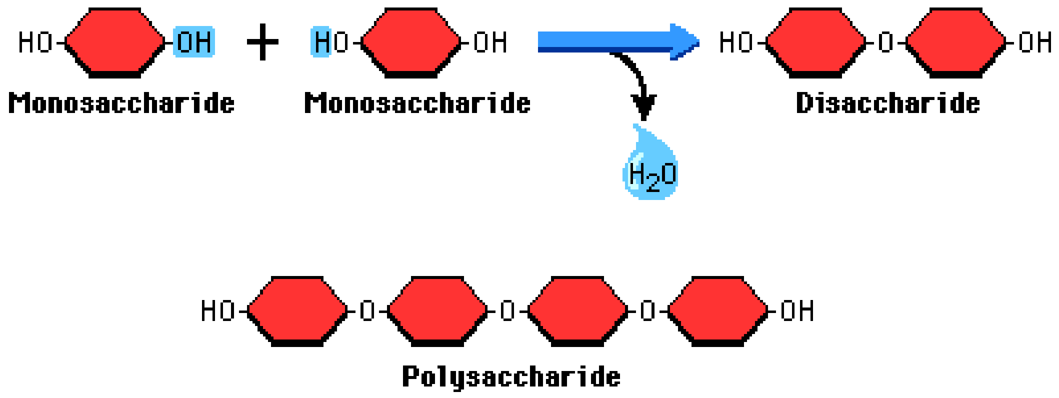
- 20 main AA's
- Same fundamental structure
- Different R-groups: acidic, basic, polar, non-polar
- AA's connect through peptide bonds via dehydration reactions
- A polypeptide is an AA polymer
- A protein is also an AA polymer, with a distinct shape and function
- Proteins can consist of several polypeptides

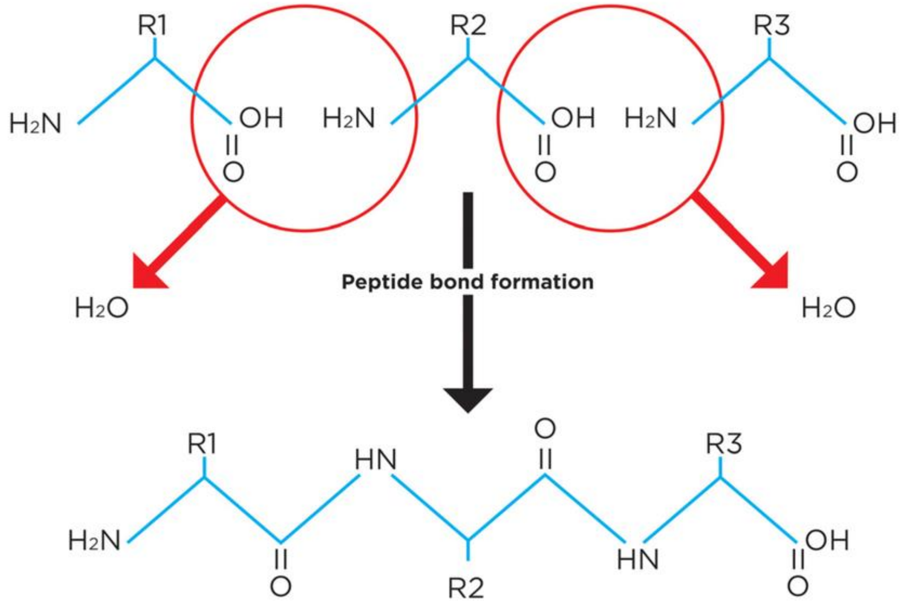



TWENTY-ONE PROTEINOGENIC α-AMINO ACIDS
Side chain charge at physiological pH 7.4
pKa values shown italicized
⊕ Positive
⊖ Negative
A. Amino Acids with Electrically Charged Side Chains
Positive
Negative
Arginine Arg R
Histidine His H
Lysine Lys K
Aspartic Acid Asp D
Glutamic Acid Glu E
B. Amino Acids with Polar Uncharged Side Chains
Serine Ser S
Threonine Thr T
Asparagine Asn N
Glutamine Gln Q
C. Special Cases
Cysteine Cys C
Selenocysteine Sec U
Glycine Gly G
Proline Pro P
D. Amino Acids with Hydrophobic Side Chains
Alanine Ala A
Valine Val V
Isoleucine Ile I
Leucine Leu L
Methionine Met M
Phenylalanine Phe F
Tyrosine Tyr Y
Tryptophan Trp W

amino acids

Dehydration reactions





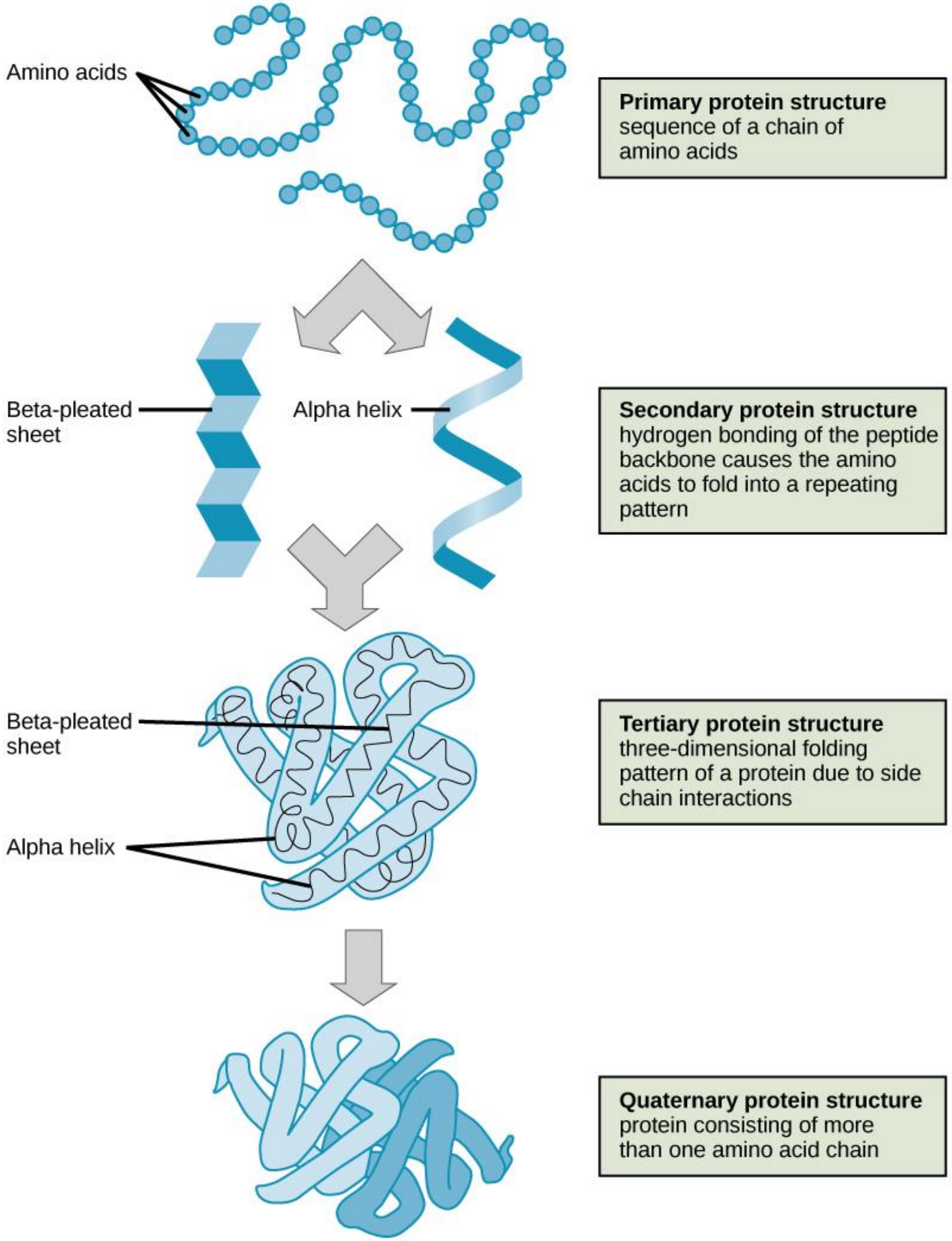
Glycosidic bonds

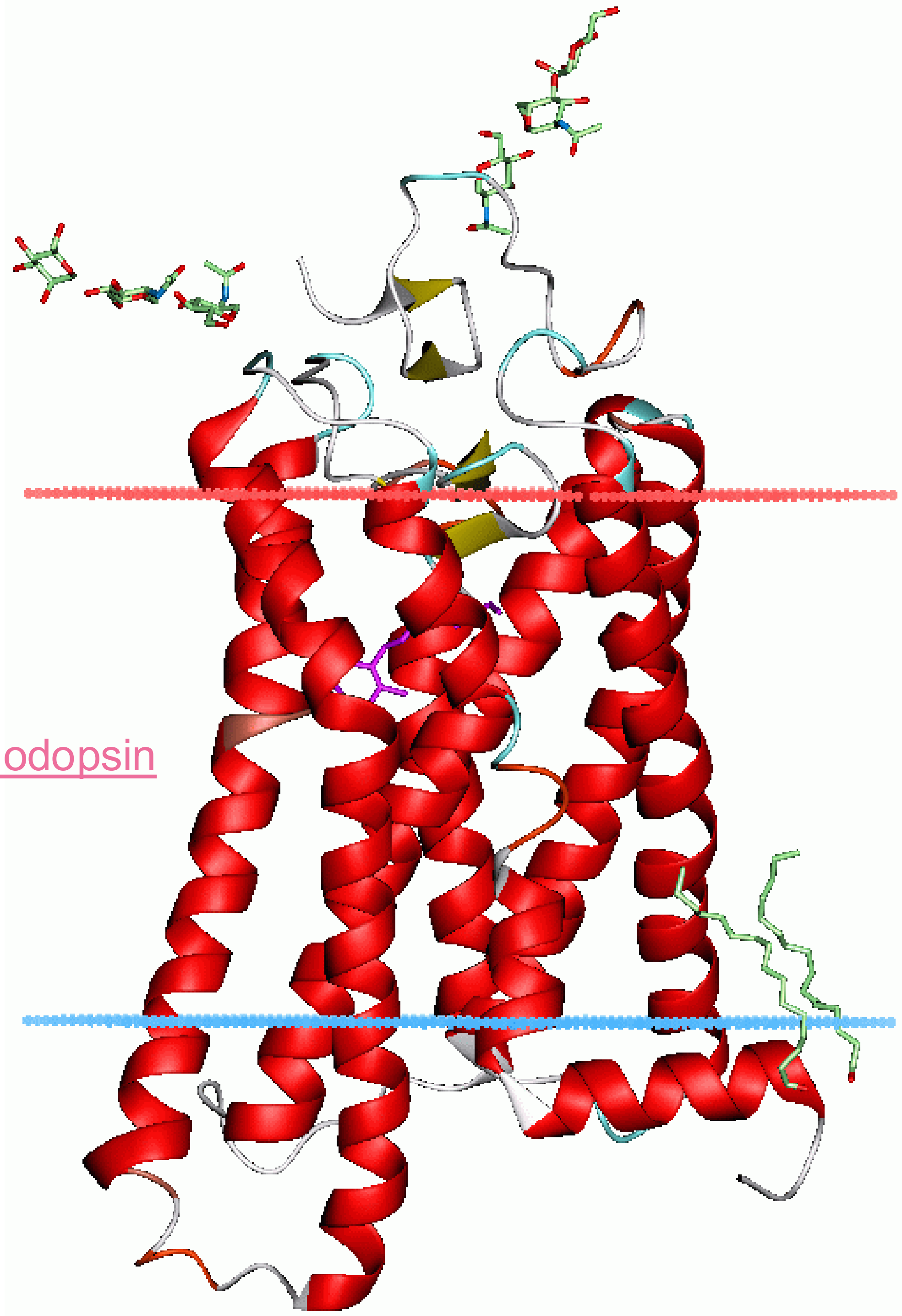
Peptide bonds

Protein structure

Biological molecules

- Primary – sequence and number of amino acids, determined by gene that encodes the protein
- Secondary – folding patterns from interactions between non-R groups, held together by H-bonds, e.g., α-helix and β-sheet
- Tertiary – 3D structure, mainly from interactions between R groups
- Quaternary – proteins formed from several polypeptides (subunits), functional form of many proteins

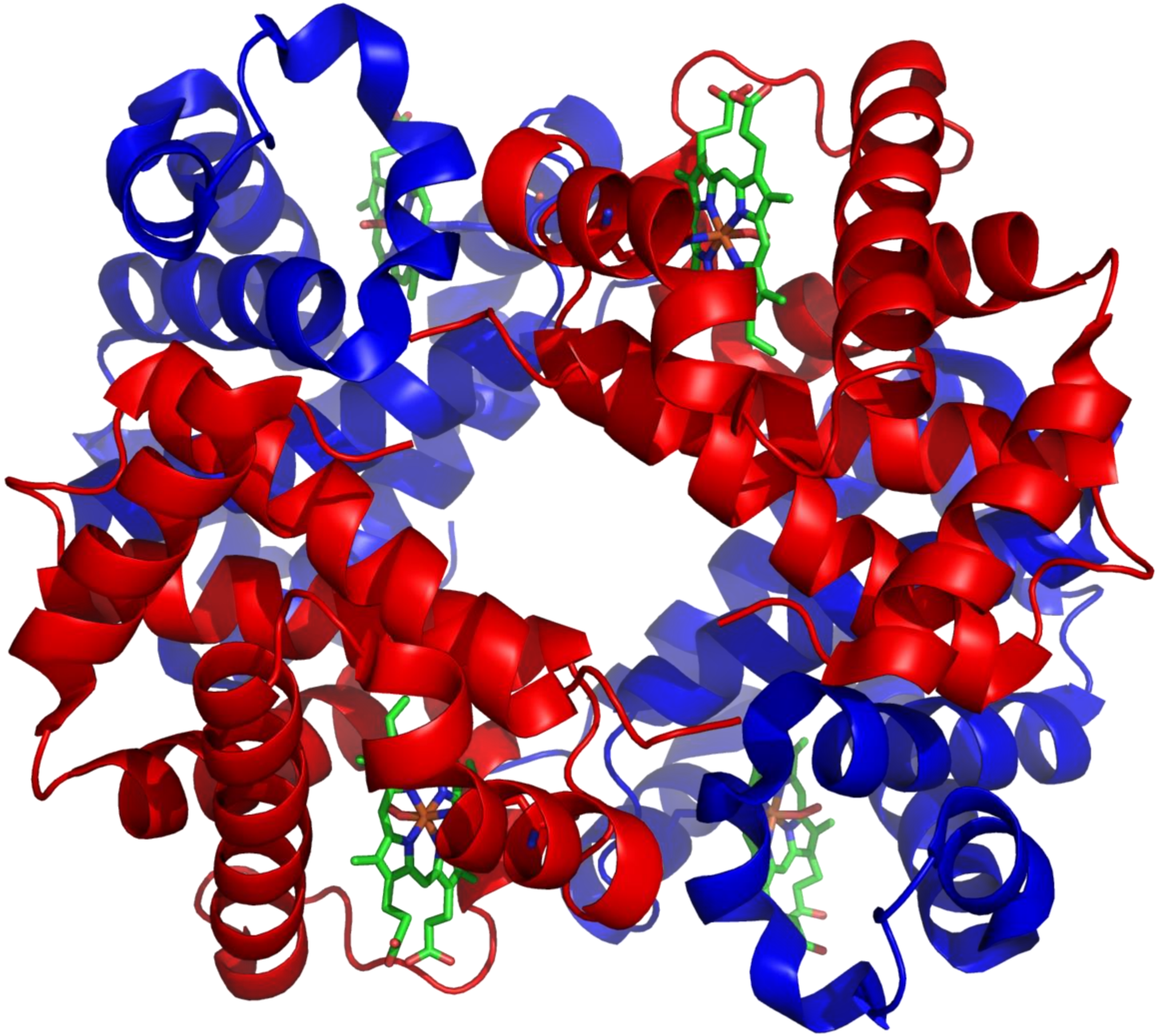




bovine rhodopsin

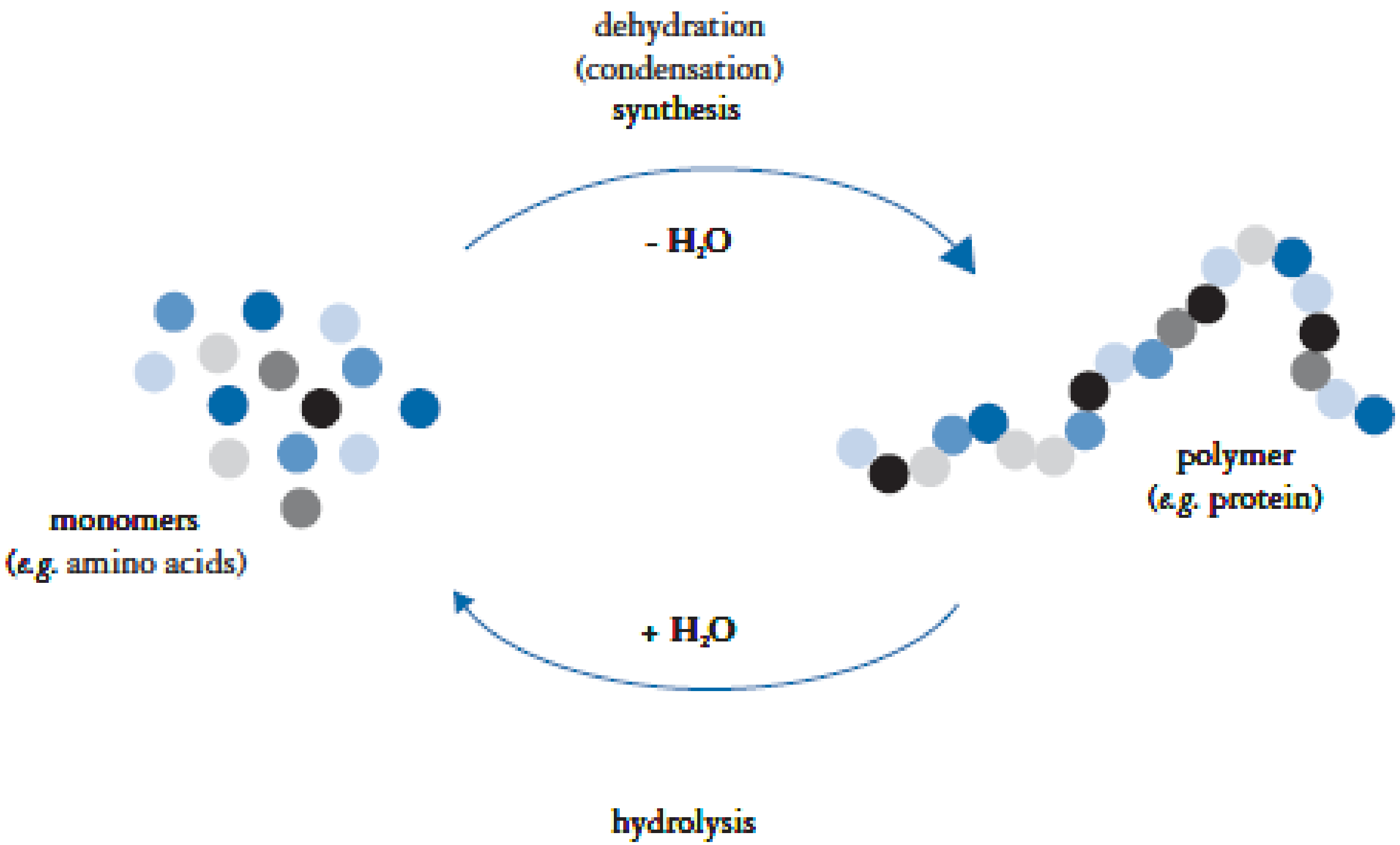
beta meander

Hemoglobin



- 4 subunits (quaternary structure)
- Oxygen transport in RBCs
- Contains 4 iron-containing heme groups = can bind up to 4 oxygen molecules
- Tense state: lower affinity to oxygen, promotes *oxygen release* in tissues where oxygen is low
- Relaxed state: upon binding 1 oxygen molecule, easier to bind other 3 (*cooperative binding*)

Monomer-polymer cycle…how are proteins made?

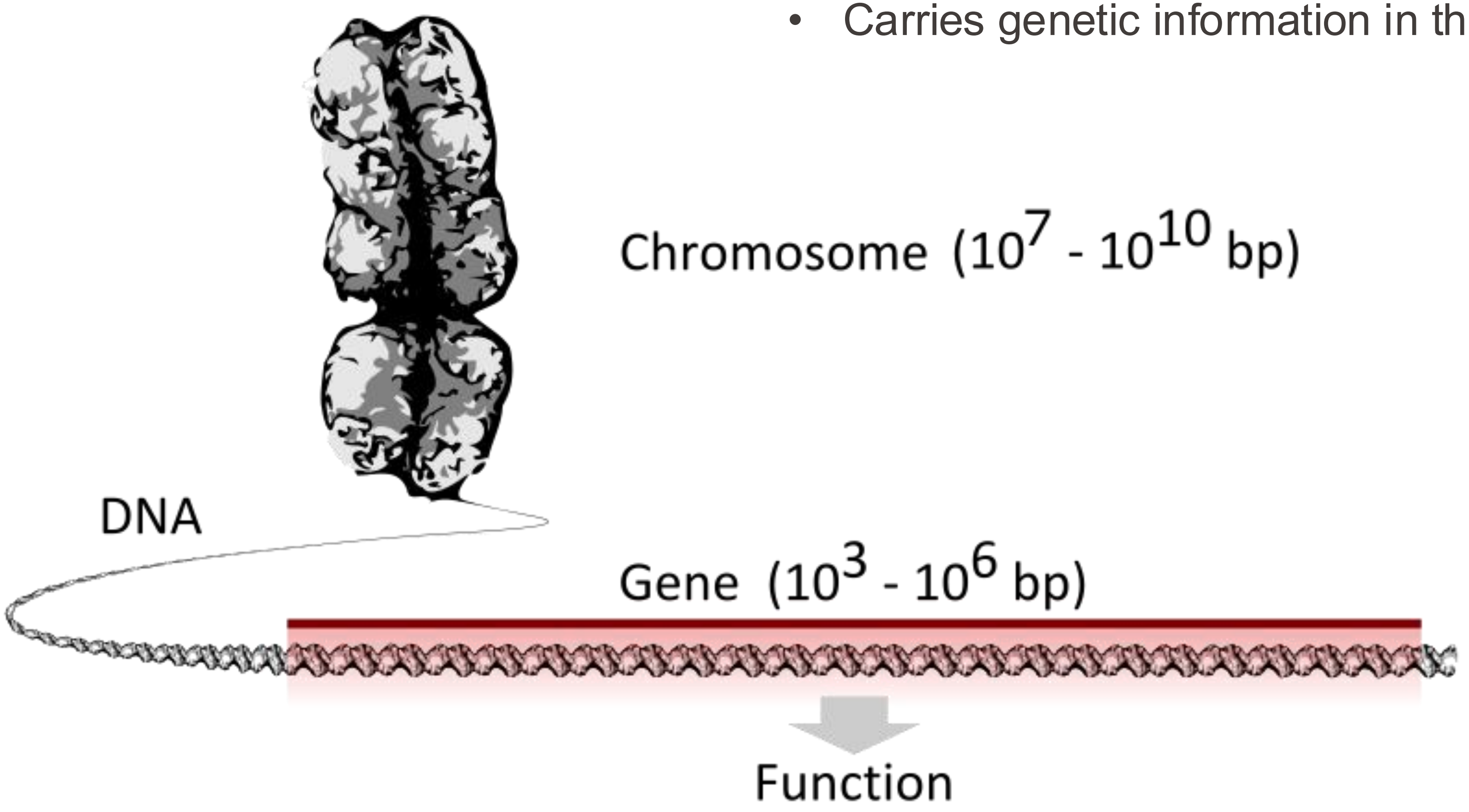


- Protein is a condensation heteropolymer (dehydration)
- Looks simple but protein synthesis requires many components whose functions are coordinated during the complex process of ***translation***

https://doi.org/10.1017/CBO9780511800986.004

So, how are proteins made?

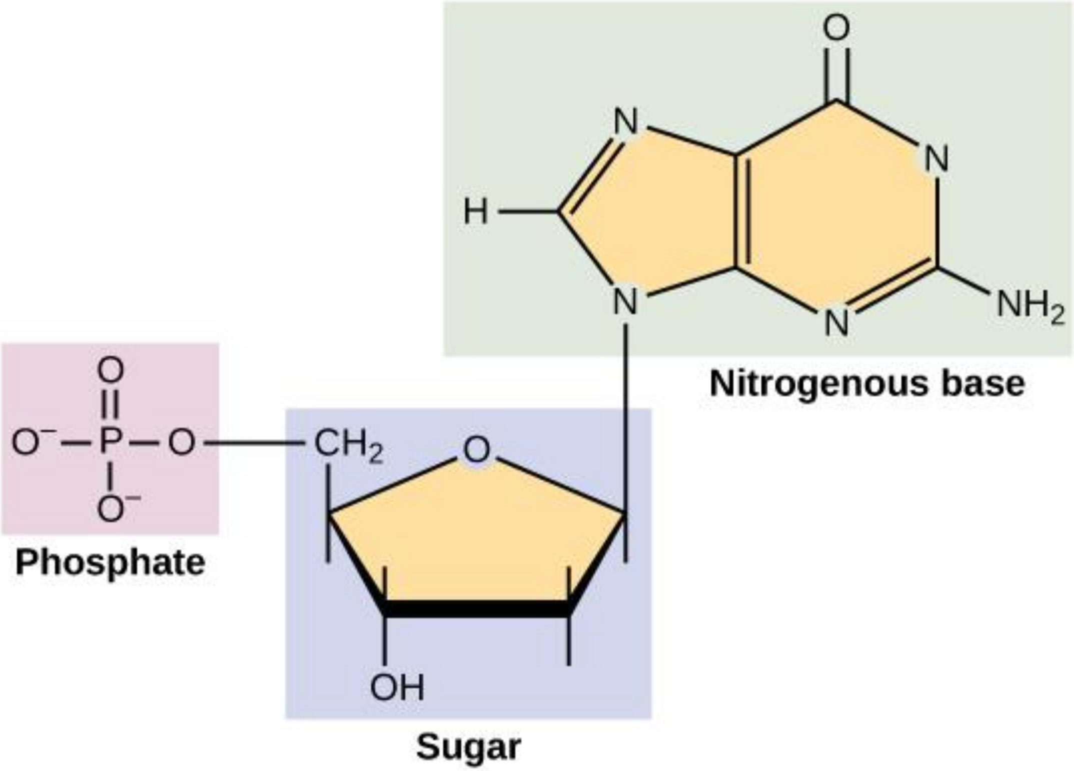
- Thread-like structure
- Made of DNA and proteins
- Carries genetic information in the form of genes

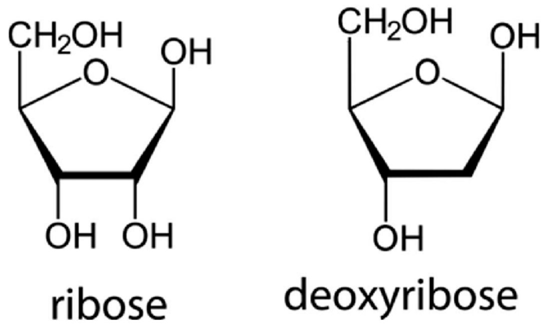


Chromosome

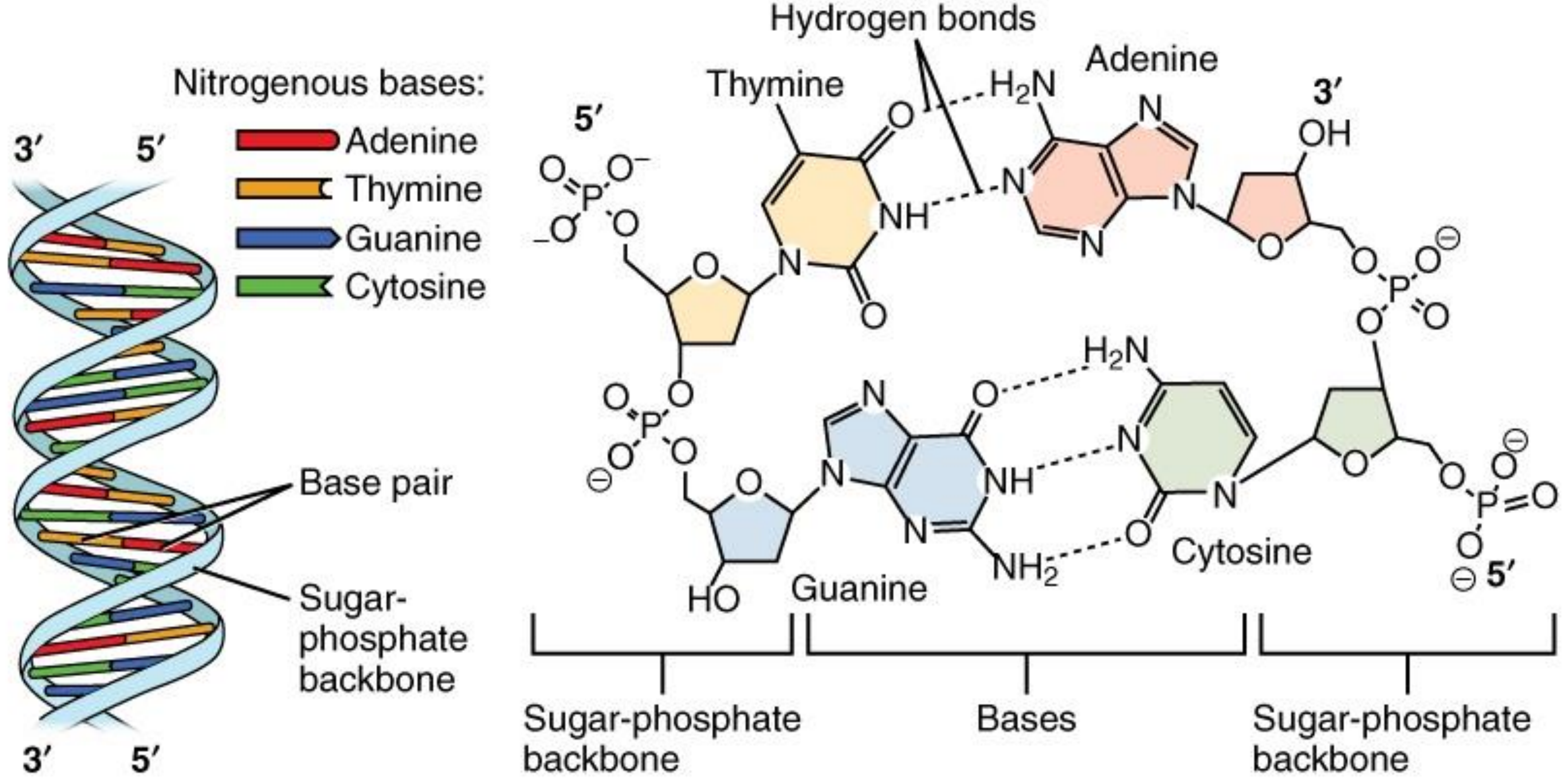
Nucleic acids (DNA/RNA)

- Carry genetic blueprint of life
- Two main types: DNA (deoxyribonucleic acid) and RNA (ribonucleic acid)
- DNA = genetic material of all living organisms
- RNA = mostly involved in protein synthesis
- **Nucleotides** are made of 3 components: nitrogenous base, pentose sugar, and a phosphate group
- Nitrogenous bases (**A**, **G**, **C**, **T**) have coding function in DNA



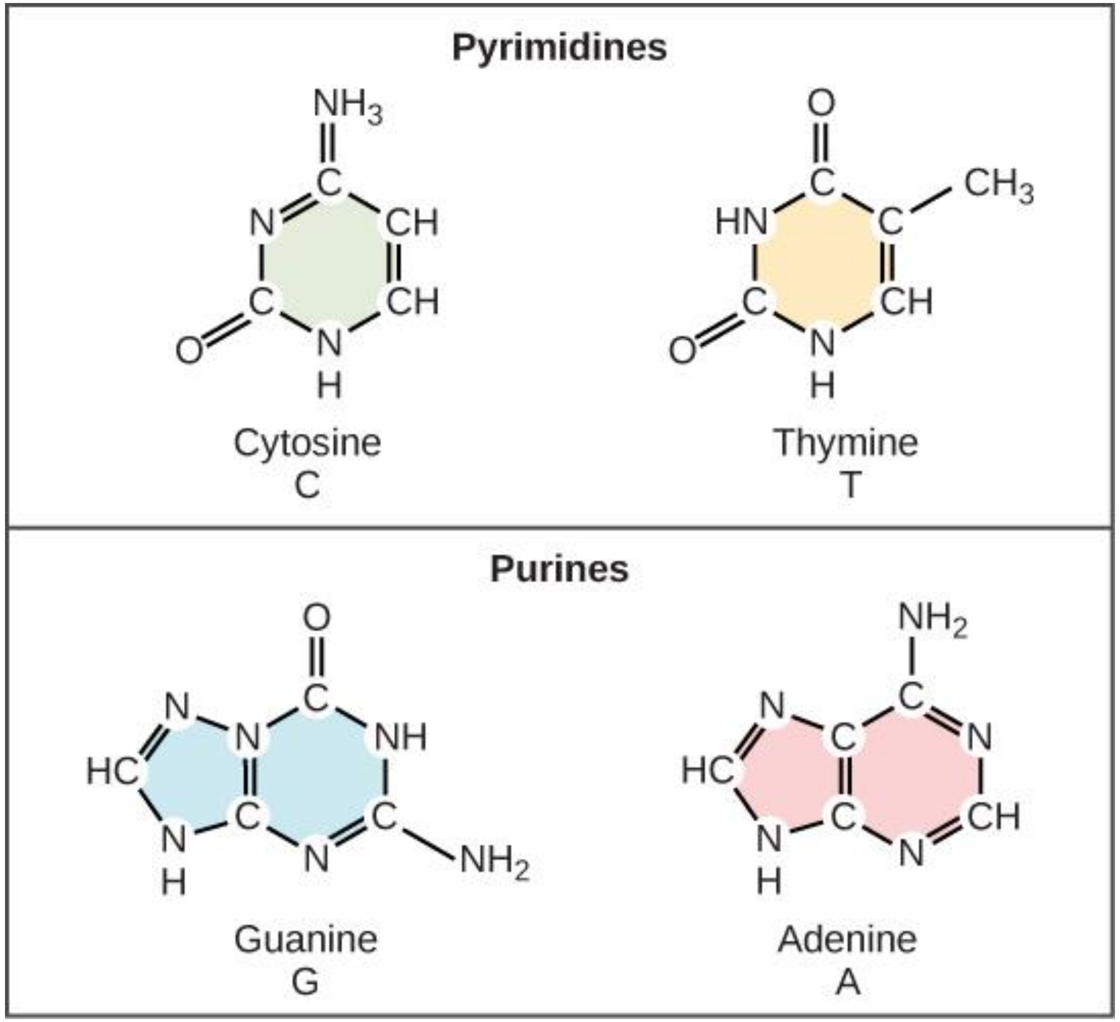


DNA

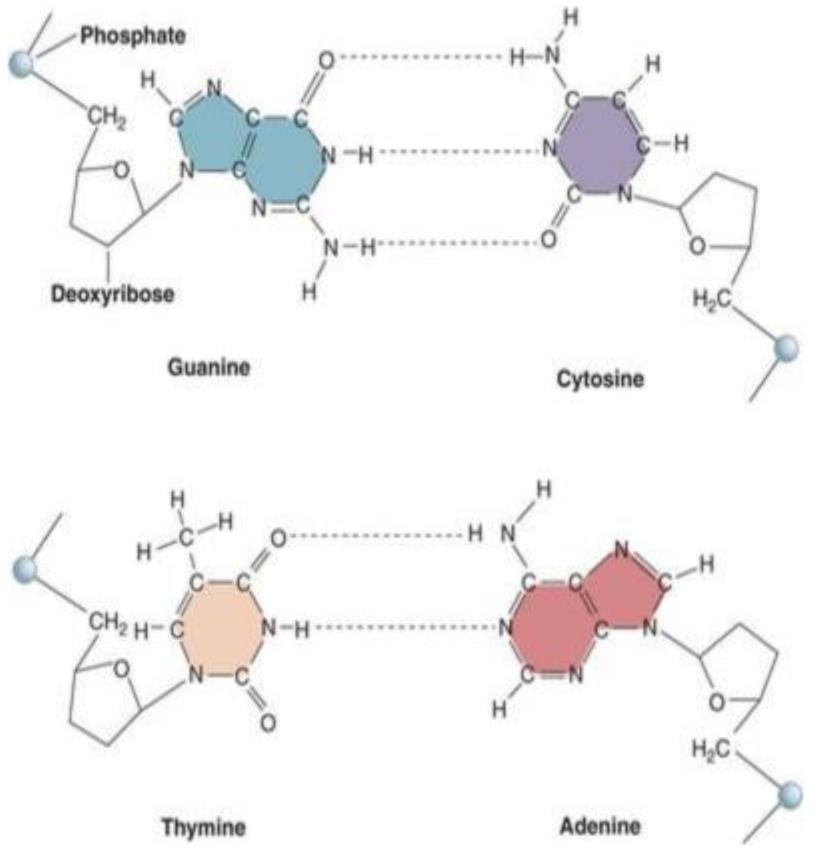


- Nucleotides join via 5'→3' phosphodiester linkage that confers directionality
- 5' end has a free phosphate, 3' end has a free hydroxyl

All genetic information – 4 letters only

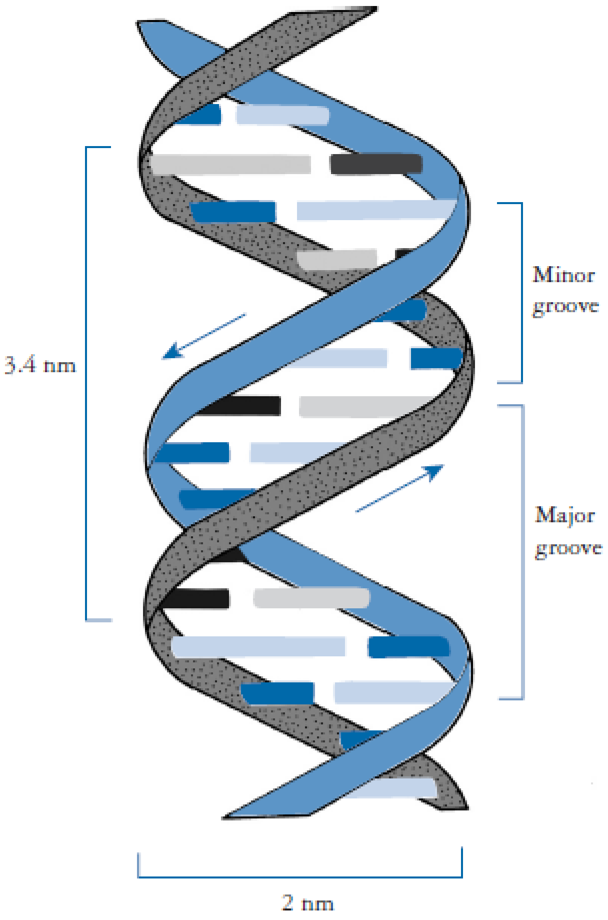


In RNA: Uracil (U) instead of thymine (T)



DNA bases

DNA

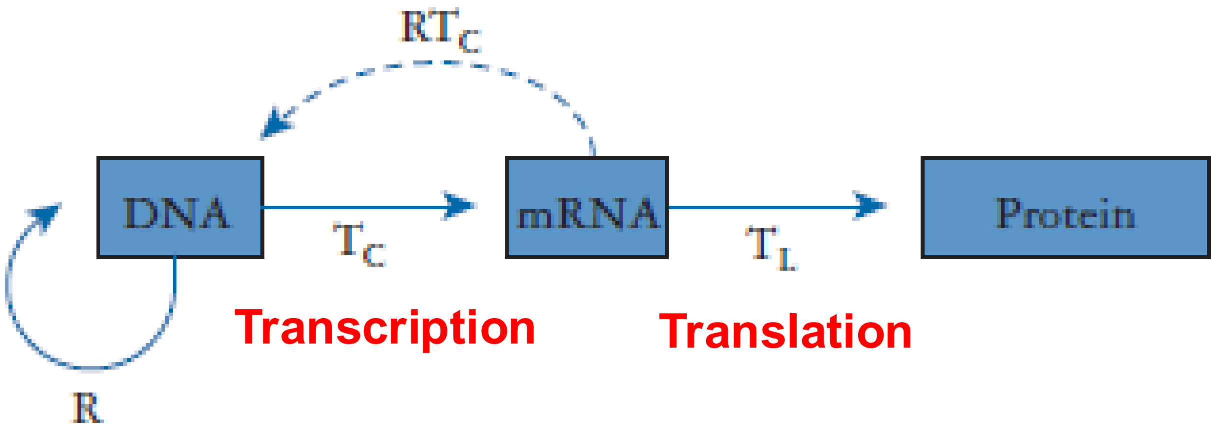


- Double-stranded DNA - primary genetic material of most organisms
- Must be stable (>100 years)
- Must be able to replicate during growth and development
- Limited potential for alteration of genetic materials (mutation) to enable evolutionary pressures to exert their effects

STABLE, REPLICABLE, MUTABLE

https://doi.org/10.1017/CBO9780511800986.004

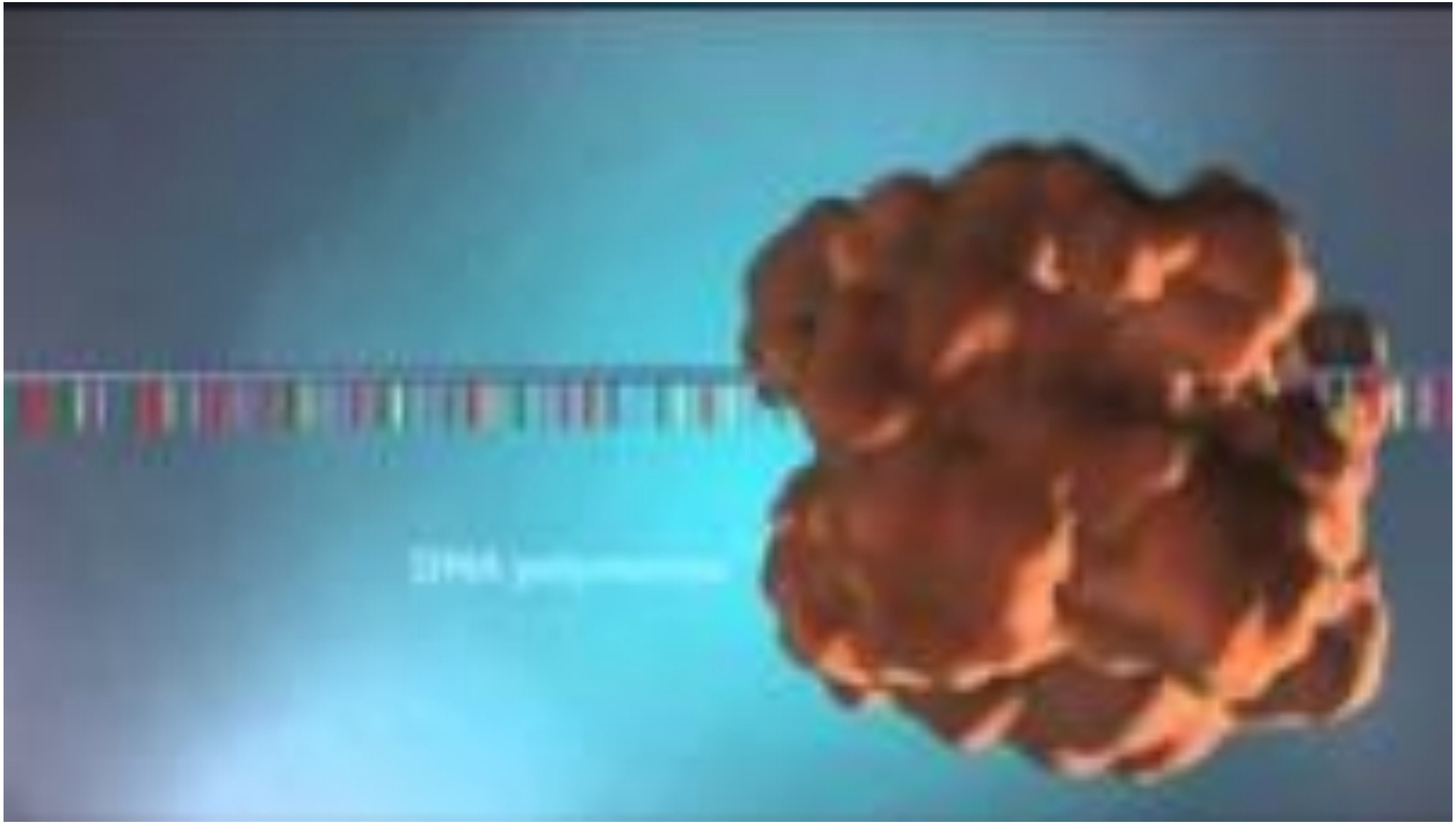
The central dogma



- Unidirectional transfer of information
- DNA's A,G, T, C codes are transferred to RNA's A,G, U, C codes
- Encoding proteins is done in **groups of 3** or ***codons***

Fig. 2.2 The Central Dogma states that information flow is unidirectional, from DNA to mRNA to protein. The processes of transcription (T_C), translation (T_L), and DNA replication (R) obey this rule. An exception is found in retroviruses (RNA viruses), which have an RNA genome and carry out a process known as reverse transcription (RT_C) to produce a DNA copy of the genome following infection of the host cell.

https://doi.org/10.1017/CBO9780511800986.004

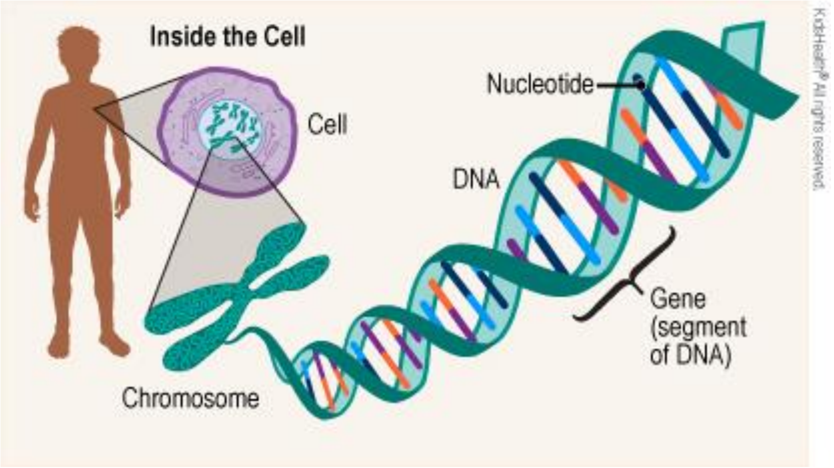


DNA Replication

[From DNA to protein]

Genes

- Basic units of genetic information
- **Stretches of DNA** that contain the instructions for building biological molecules, like protein
- Determine specific traits by encoding proteins, which carry out various functions in the body
- 'Gene' as a term usually represents the genetic information ***transcribed* into *one* molecule of RNA**, which is in turn ***translated* into *one* protein**



- Prokaryotes – in cytoplasm
- Eukaryotes – in nucleus

https://doi.org/10.1017/CBO9780511800986.004

The genetic code

- 20 AA's but only 4 letters (A,G, C, T)
- Triplet combinations of 4 letters: $4^3 = 64$ possibilities
- Codons are sets of 3 nucleotides (trinucleotides) that specify a given amino acid
- More codons than needed – 3 are 'STOP' codons, some amino acids are specified by more than 1 codon (redundancy in code)

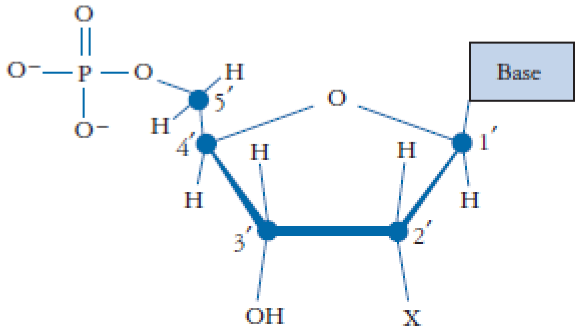
https://doi.org/10.1017/CBO9780511800986.004

The genetic code

Table 2.1. The genetic code

First base (5′ end)	Second base U	C	A	G	Third base (3′ end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

https://doi.org/10.1017/CBO9780511800986.004



- Read 5’ → 3’
- RNA uses uracil (U) to replace thymine (T)

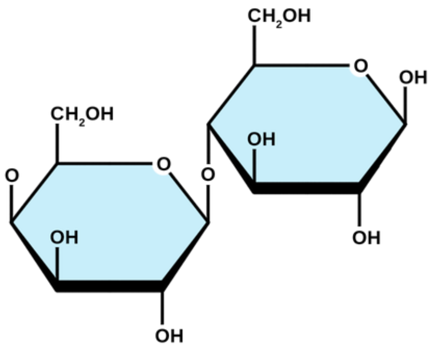
- Genes are made of sequences of trinucleotides that encode for a specific AA acid sequence, e.g. a protein

A protein made up of 300 amino acids needs how many nucleotides at a minimum?

A. 4
B. 900
C. 600
D. 64

Provide the codons for the 'STOP' command

Gene structure in prokaryotes (*lac* operon example)



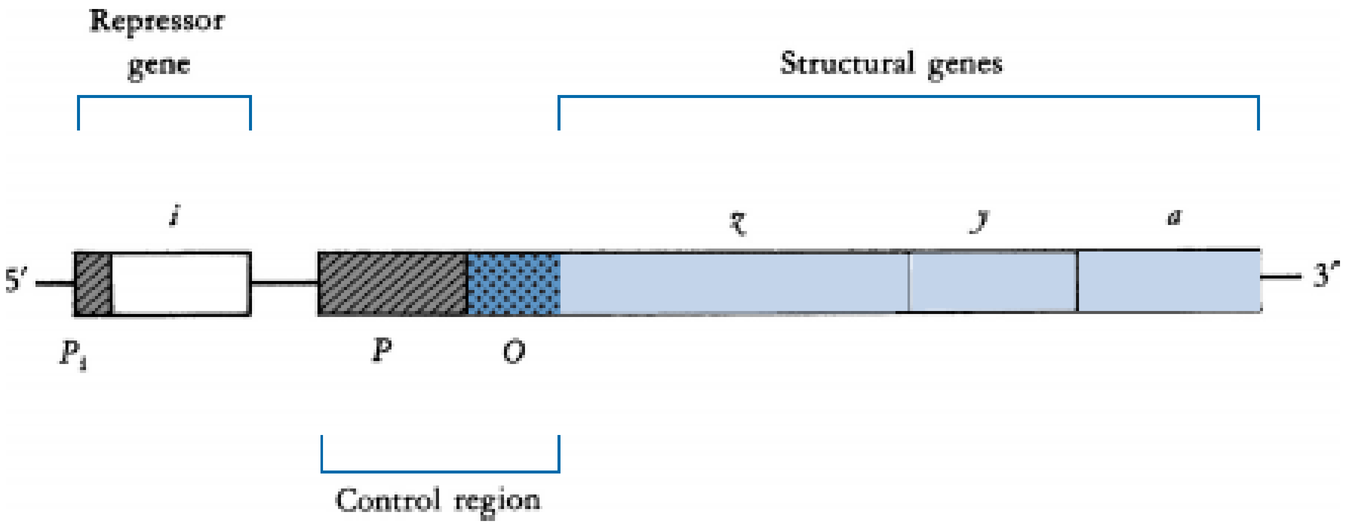


Fig. 2.7 The *lac* operon. The structural genes *lacZ*, *lacY*, and *lacA* (noted as *z*, *y*, and *a*) encode ß-galactosidase, galactoside permease, and a transacetylase, respectively. The cluster is controlled by a promoter (*P*) and an operator region (*O*). The operator is the binding site for the repressor protein, encoded by the *lacI* gene (*i*). The repressor gene lies outside the operon itself and is controlled by its own promoter, P_i.

- Operon is a cluster of genes usually found in prokaryotes, can code for more than 1 protein
- *Lac* operon - codes the enzymes for **lactose catabolism**
- 3 structural genes that code for protein that digests lactose
- Gene circuit that only lets cell use lactose in the absence of glucose
- Repressor gene codes for LacI repressor protein (off switch)

https://doi.org/10.1017/CBO9780511800986.004

EPFL

Why are we learning lac operon in ELM's?

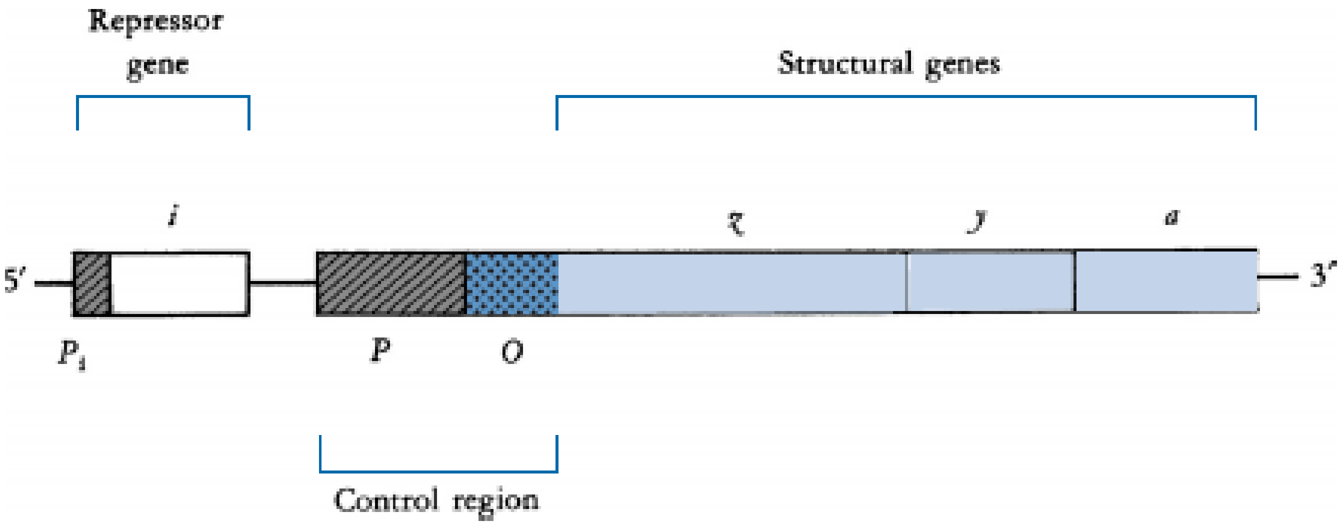


Fig. 2.7 The *lac* operon. The structural genes *lacZ*, *lacY*, and *lacA* (noted as *z*, *y*, and *a*) encode ß-galactosidase, galactoside permease, and a transacetylase, respectively. The cluster is controlled by a promoter (*P*) and an operator region (*O*). The operator is the binding site for the repressor protein, encoded by the *lacI* gene (*i*). The repressor gene lies outside the operon itself and is controlled by its own promoter, P_i.

https://doi.org/10.1017/CBO9780511800986.004

- ON/OFF Switch for genes (and their response)
- LacI is a repressor protein that keeps the switch off
- IPTG is a chemical inducer
- To turn the switch ON: IPTG binds to LacI, leading to its release
- Now transcription and translation can happen – stays ON
- Remove IPTG, LacI binds again, off-switch (LacI is not metabolized)
- You can engineer the structural gene part to turn on/off the expression of different proteins!

Recombinant DNA technology

General idea:

- Used to alter the genetic make up of an organism
- Combines DNA from different sources to create sequences that do not naturally occur
- Can be used to get a host to produce a non-native trait or protein

Steps:

(1) Identify target gene

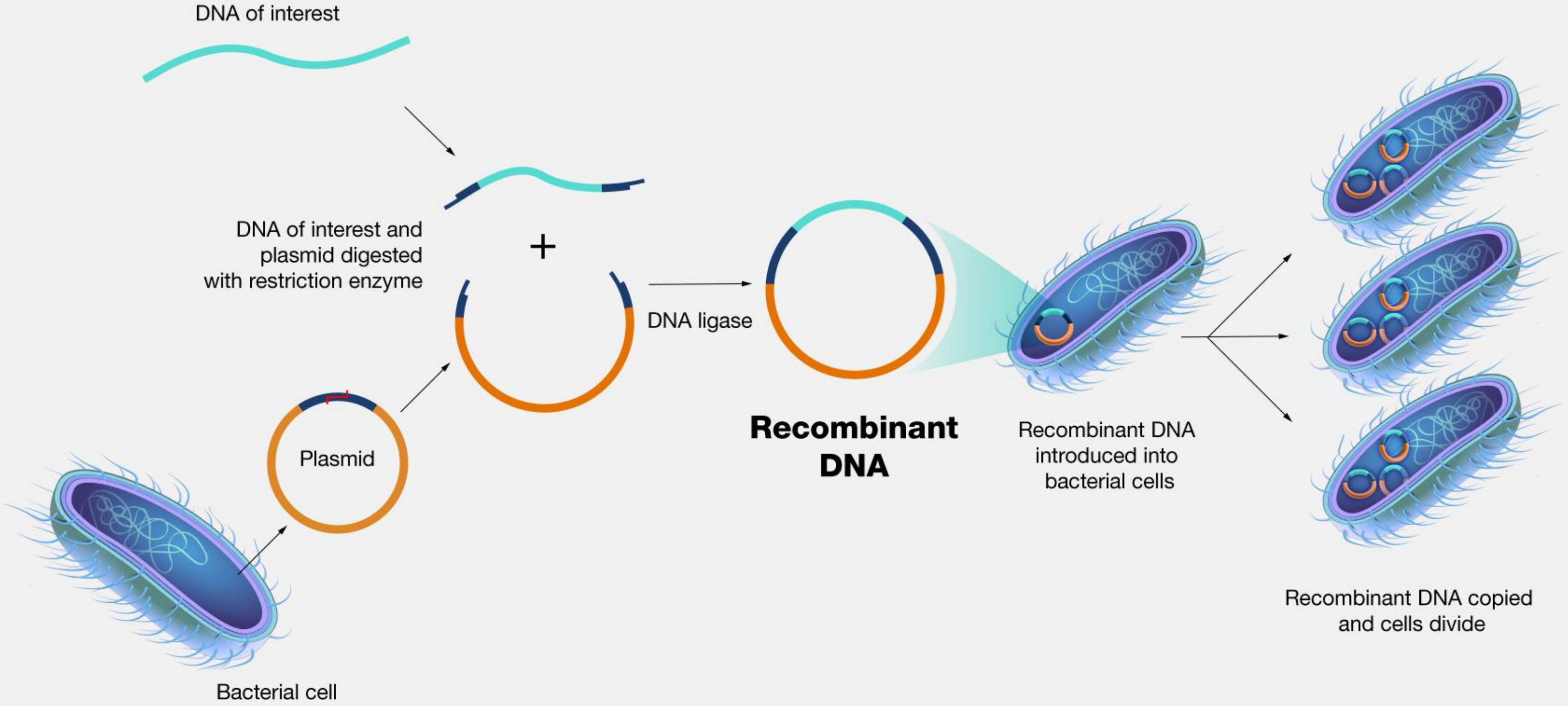
(2) Cut DNA with restriction enzymes

(3) Insert target gene into vector (carrier molecule, usually a plasmid), join with DNA ligase; now you have recombinant DNA

(4) Introduce rDNA into host organism

(5) Host organism can express new gene to produce, e.g., a protein of interest

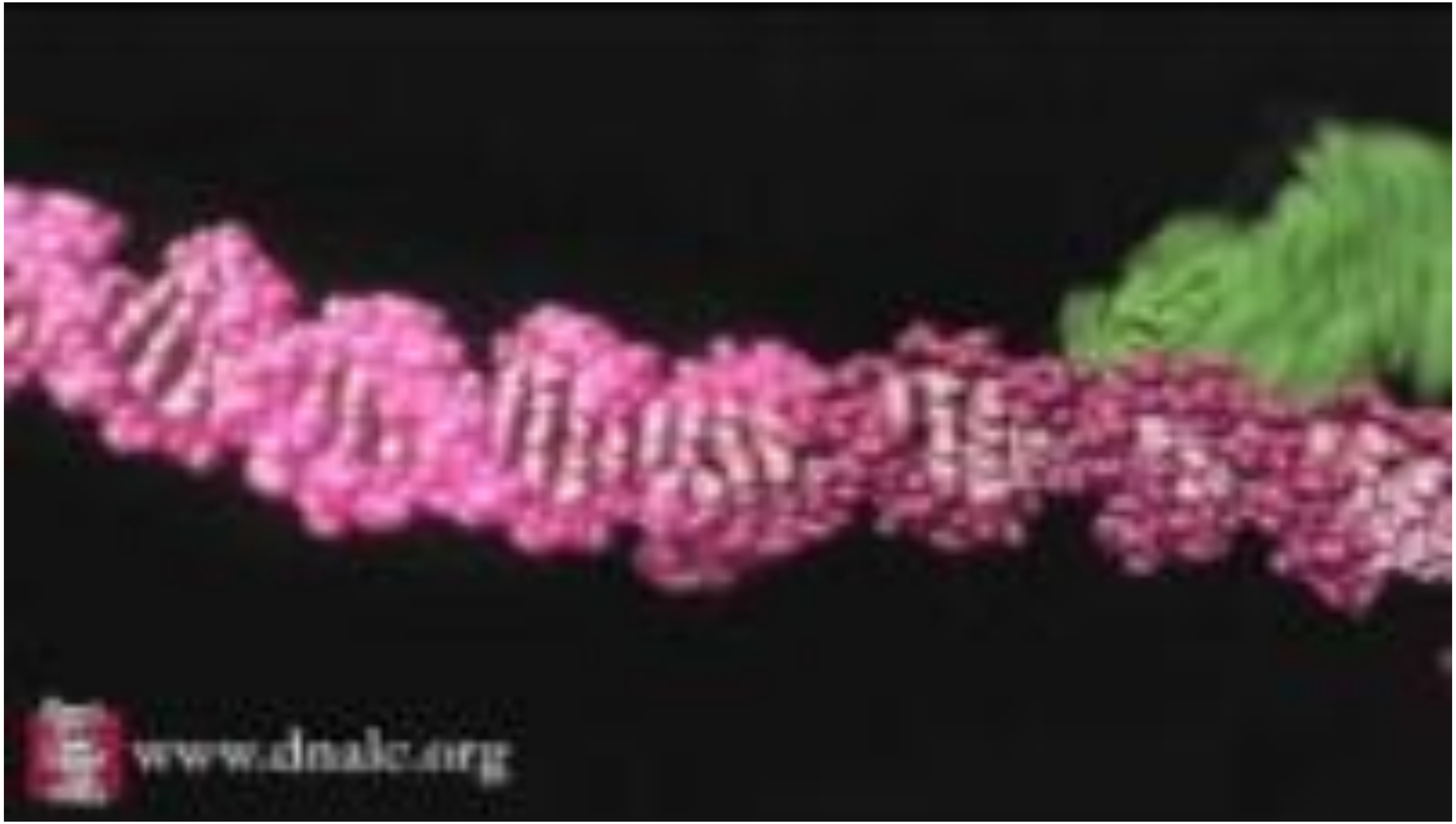
Recombinant DNA technology



[Recombinant DNA technology]

Tools for gene editing - Recombinant DNA Technology

Recombination



What is CRISPR?

- CRISPR – “Clustered Regularly Interspersed Short Palindromic Repeats”
- DNA sequences in bacteria that are from past viral infection
- CRISPR DNA is coupled to a molecular scissor (Cas9)
- CRISPR-Cas9 defends bacteria against subsequent viral infection by recognizing the viral sequence, cutting the viral DNA, and deactivating the virus
- Part of the adaptive immune system of bacteria

https://www.nobelprize.org/uploads/2020/10/popular-chemistryprize2020.pdf

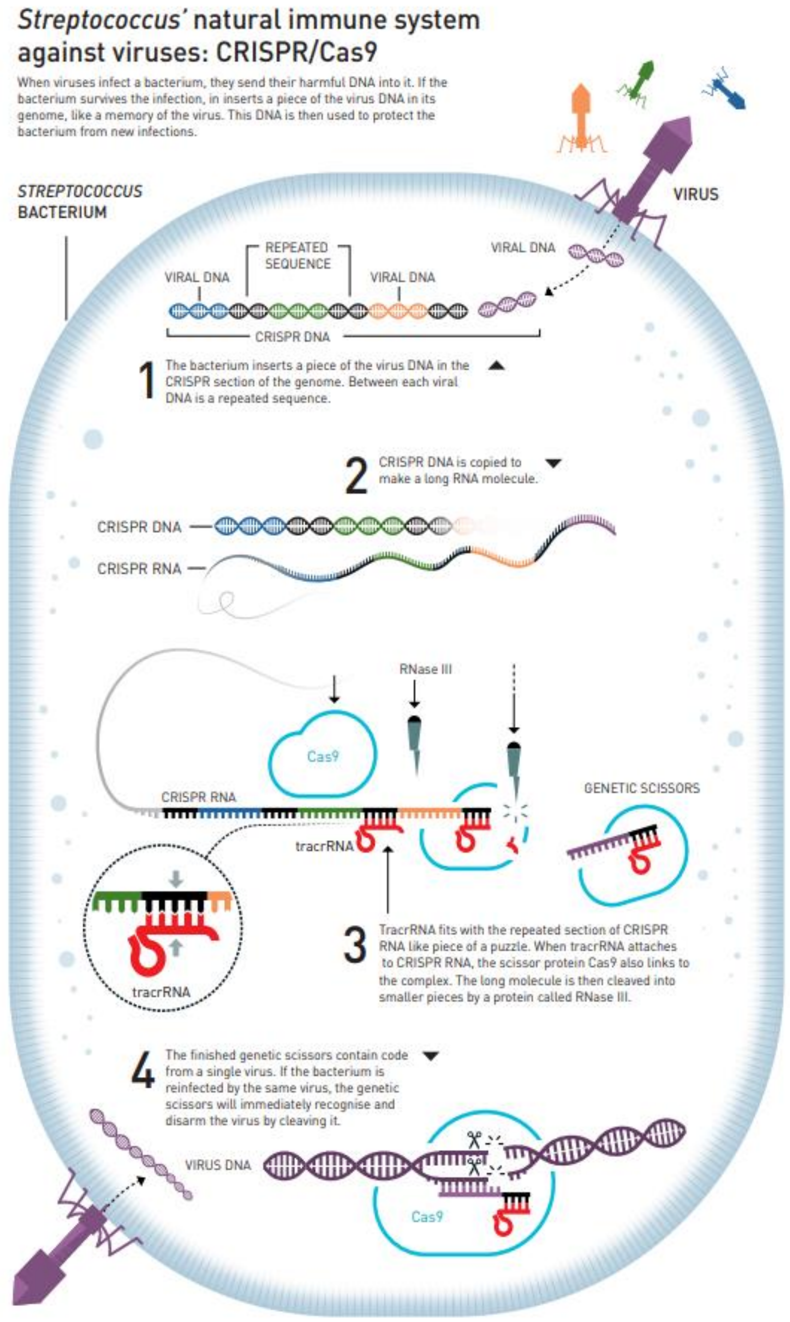


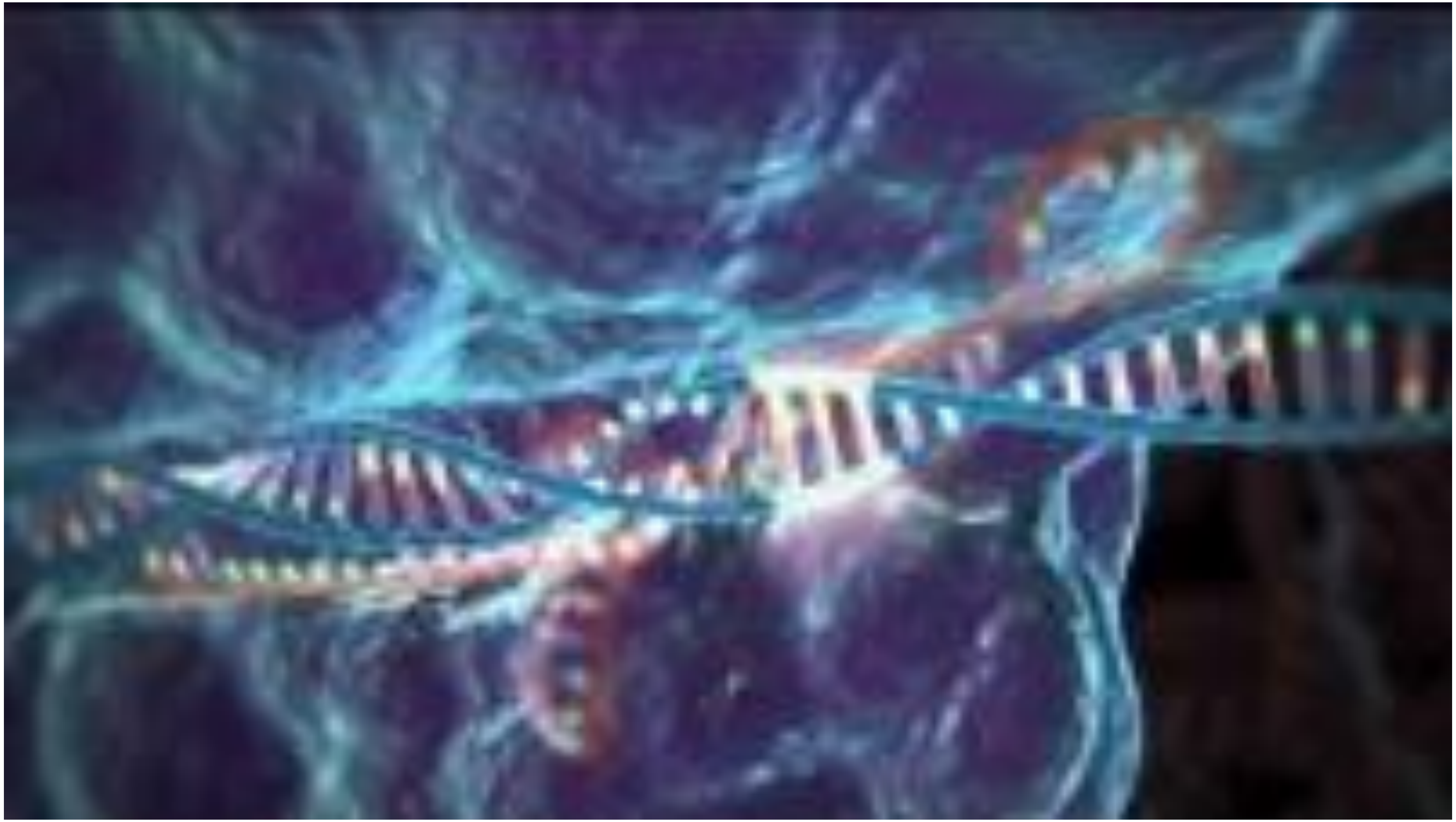
Figure 2

2020 Nobel Prize: CRISPR-Cas9, Emmanuelle Charpentier Jennifer A. Doudna

CRISPR-Cas9 Story

Tools for gene editing - CRISPR-CAS9

https://www.youtube.com/watch?v=2pp17E4E-O8



Take aways

- Biological ingredients
- Some glimmer of tools of genetic engineering
- Can create designer microbes that elicit a specific response, on demand
- This is the worst it will get in terms of the biology that we will study