Lecture 3
Protein Structure and Function, Carbohydrates, Nucleic Acids

- Protein Structure and Stability
- Ligand Binding
- Proteins as enzymes (PKU disease)
- Carbohydrates
- Nucleic Acid Technologies

Summary - Interactions that Stabilize Folded Proteins.

- **Hydrogen bonds** form between hydrogen atoms (NH) and the carbonyl group in the peptide backbone (mainchain), and between and donors and acceptors on sidechains. *Mainchain-mainchain H-bonds are responsible for secondary structures.*
- **Hydrophobic interactions** within a protein increase stability of the folded state by *increasing entropy due to the release of water that was ordered by the exposed non-polar groups in the unfolded protein.*
- **van der Waals interactions** are *optimized in the well packed core of the protein*.

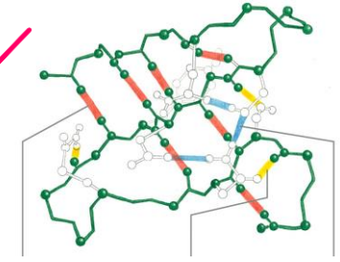
- **Covalent disulfide bonds** form between sulfur-containing cysteine residues stabilizing them (usually only exported, secreted proteins).

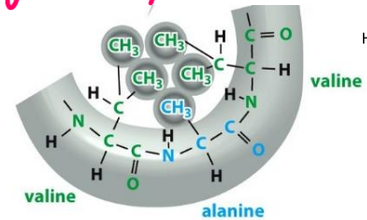
H-bonds
van der Waals
Hydrophobic effect

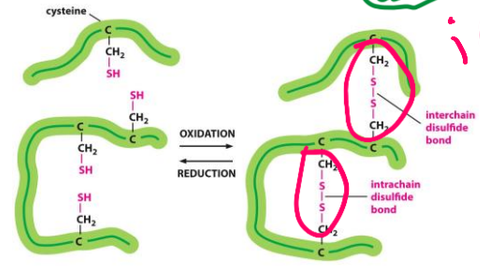
Chain disorder

Native

Unfolded







A single change in the amino acid sequence can change the function of a protein, and often affecting how it folds – Producing Inactive Proteins.

frontiers in Molecular Biosciences

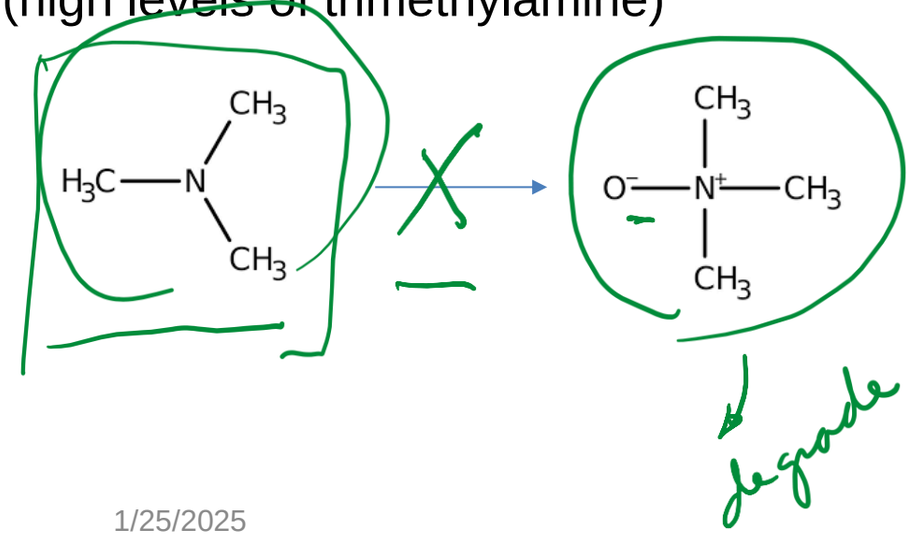
ORIGINAL RESEARCH
published: 07 January 2021
doi: 10.3389/fmolb.2020.626363

Solvent Accessibility of Residues Undergoing Pathogenic Variations in Humans: From Protein Structures to Protein Sequences

Castrense Savojardo [1], Matteo Manfredi [1], Pier Luigi Martelli [1] and Rita Casadio [1,2]*

[1] Biocomputing Group, Department of Pharmacy and Biotechnologies, University of Bologna, Bologna, Italy, [2] Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies of the National Research Council, Bari, Italy

Mutations in dimethylaniline monooxygenase 3 cause trimethylaminuria (high levels of trimethylamine)



degrade

dimethylaniline monooxygenase 3

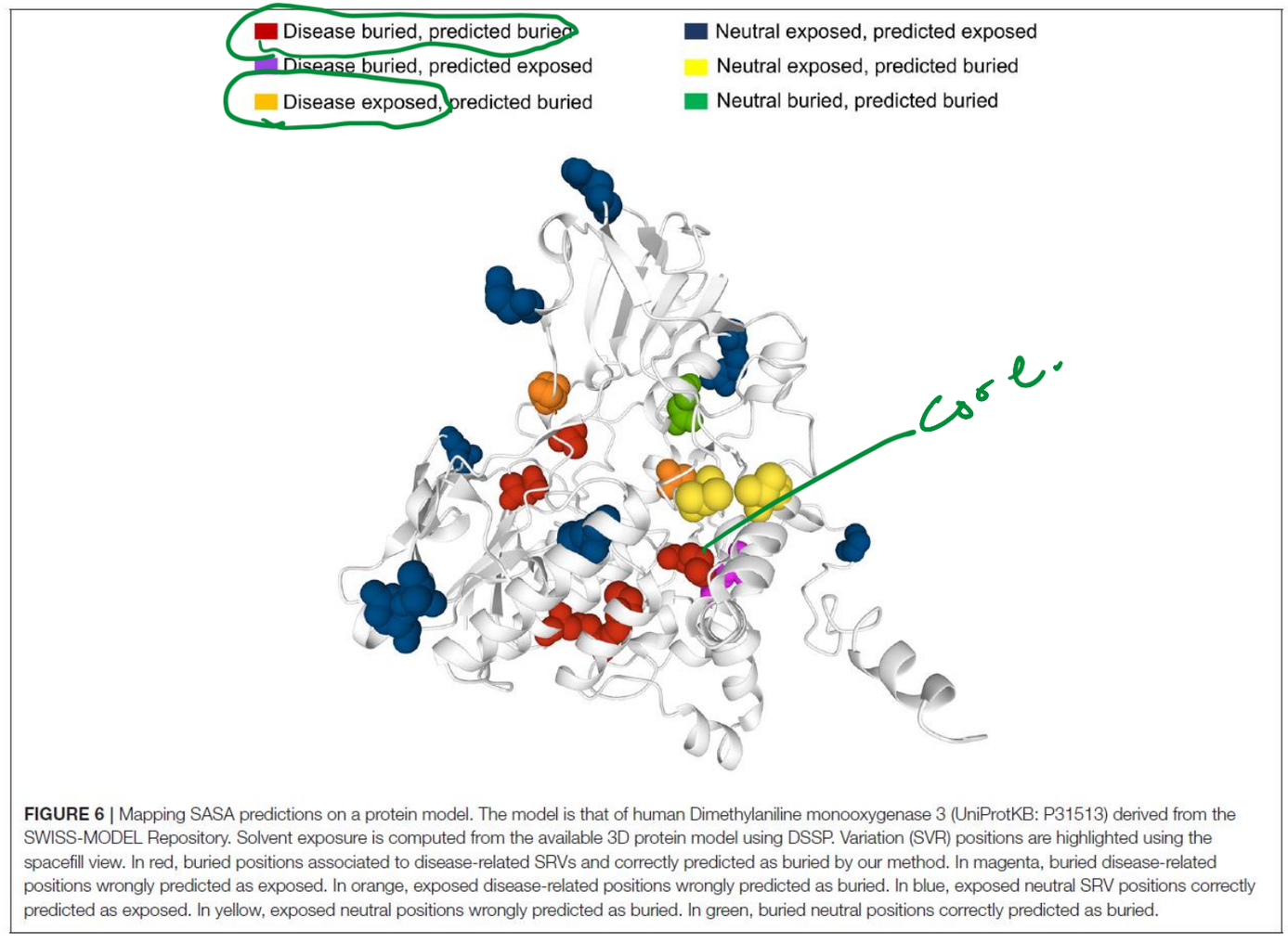
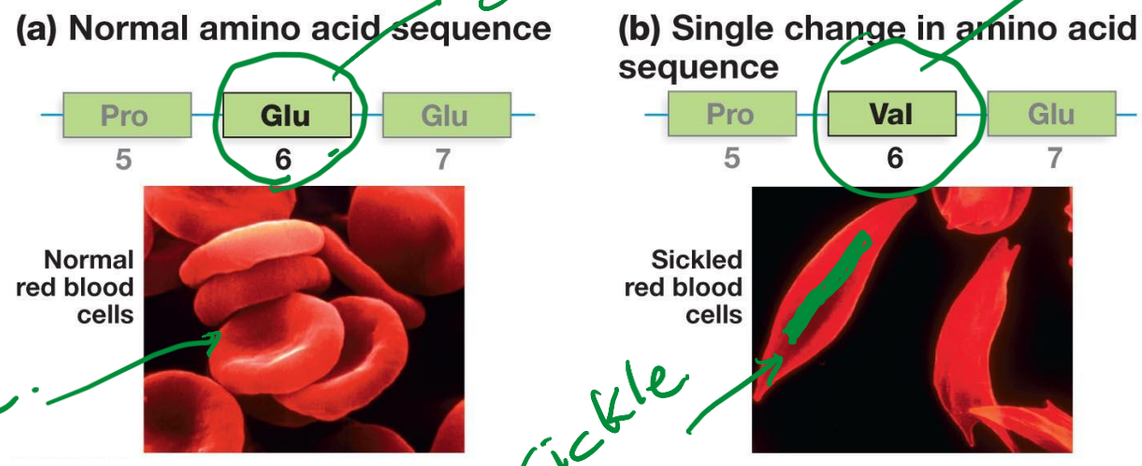


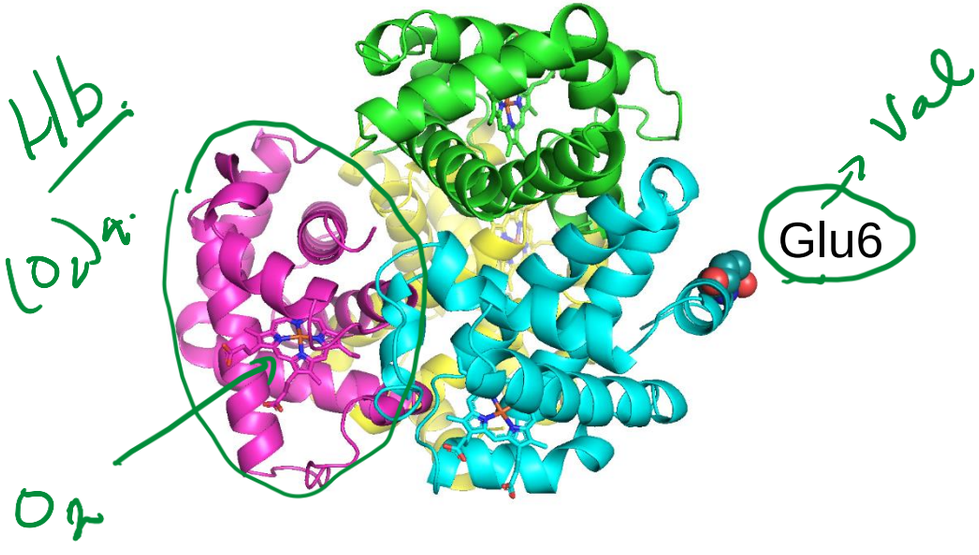
FIGURE 6 | Mapping SASA predictions on a protein model. The model is that of human Dimethylaniline monooxygenase 3 (UniProtKB: P31513) derived from the SWISS-MODEL Repository. Solvent exposure is computed from the available 3D protein model using DSSP. Variation (SVR) positions are highlighted using the spacefill view. In red, buried positions associated to disease-related SRVs and correctly predicted as buried by our method. In magenta, buried disease-related positions wrongly predicted as exposed. In orange, exposed disease-related positions wrongly predicted as buried. In blue, exposed neutral SRV positions correctly predicted as exposed. In yellow, exposed neutral positions wrongly predicted as buried. In green, buried neutral positions correctly predicted as buried.

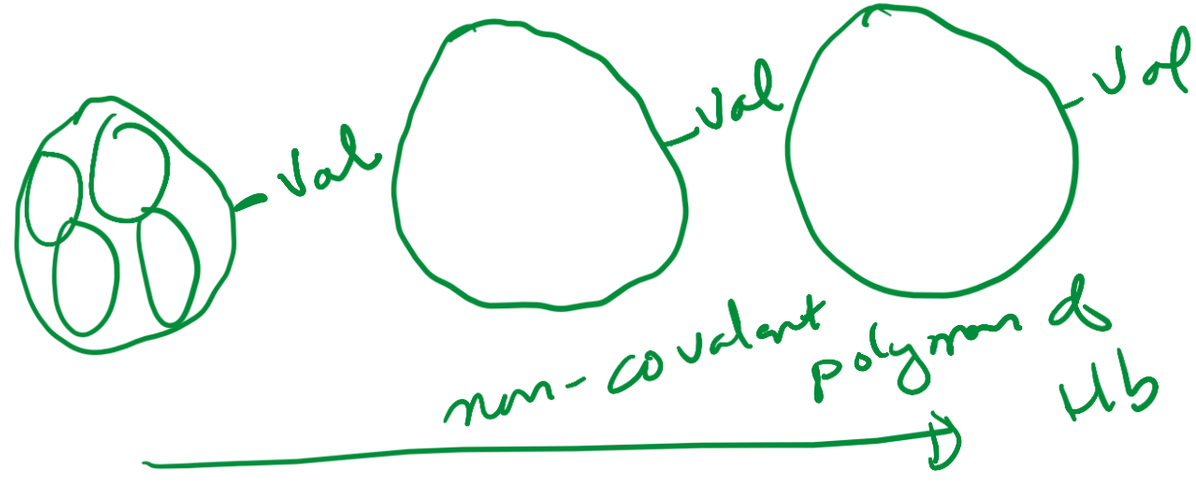
Surface Mutations May Also Lead to Disease

Effect of mutations on protein folding – sickle cell anemia



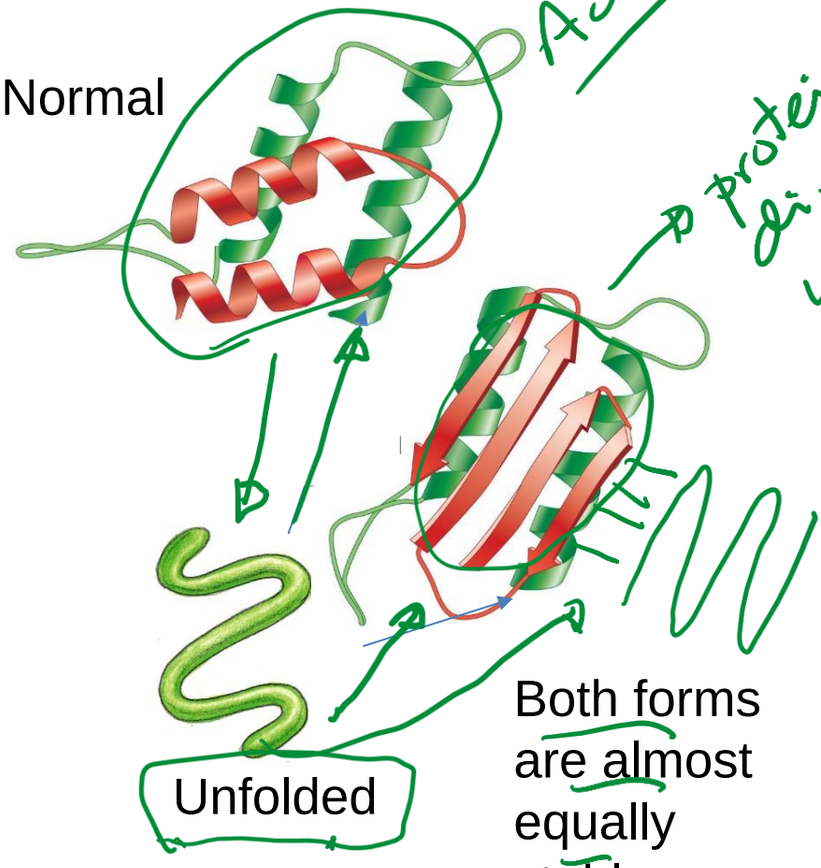
A single change in the amino acid sequence can change the function of a protein



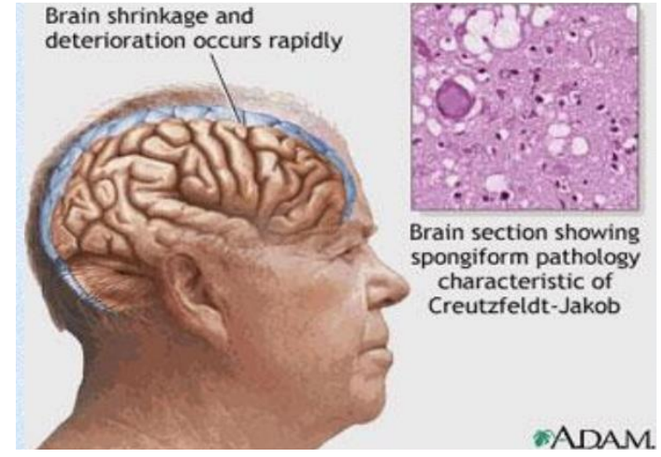


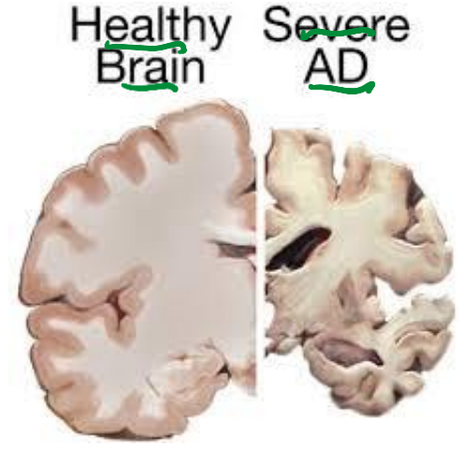
What Happens When Proteins Fold Into Different Structures?

Prions are improperly folded proteins that cause neurodegenerative diseases

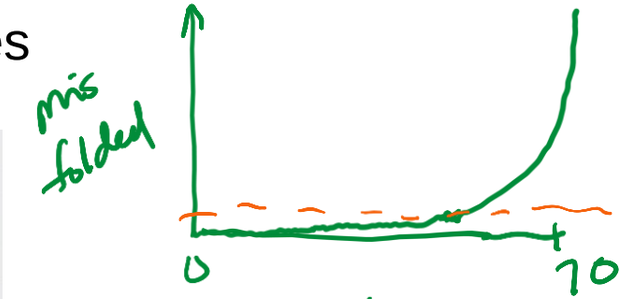


Both forms are almost equally stable.





Misfolded prion protein



Unfolded protein response (UPR): The presence of unfolded proteins can trigger the UPR, which can turn off protein synthesis in the cell, leading to cell death.

What is the effect on the brain?

Why do the brain cells die?

Protein Structure - Summary and Expectations

Primary Structure:

- Can you describe the mechanism of peptide bond formation
- Can you draw structure of peptides.
- Can you identify amino terminus and give the sequence of amino acids, N -> C

Secondary structure:

- Identify helical and sheet secondary structures,
- know that they are stabilized by **mainchain** hydrogen bonds between N-H and O=C.
- Location of H-bonds and sidechains

Tertiary Structure:

- Can you describe and identify role of the following in stabilizing the folded state.
  - H-bonds,
  - van der Waals,
  - hydrophobic effect
- Can you predict, based on sidechain, which amino acids are found in the core of the protein and which are found on the surface.

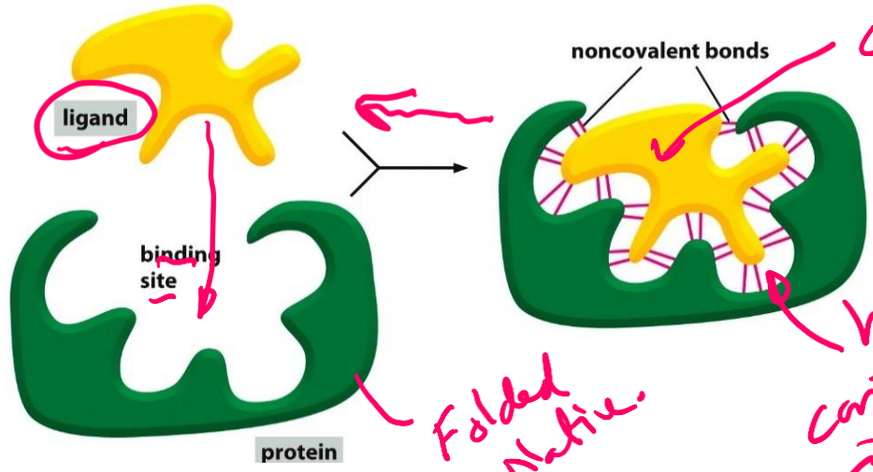
Quaternary Structure:

- Multiple chains, stabilized by non-covalent and covalent (disulfide bonds) interactions.

Diseases related to protein structure:

- Core mutations - affect folding
- Surface mutations - affect protein-protein (and Protein-DNA) interactions
- Stable isoforms – toxic to the cell

Ligand Binding: Most Proteins Bind to Other Molecules in Biological Interactions:



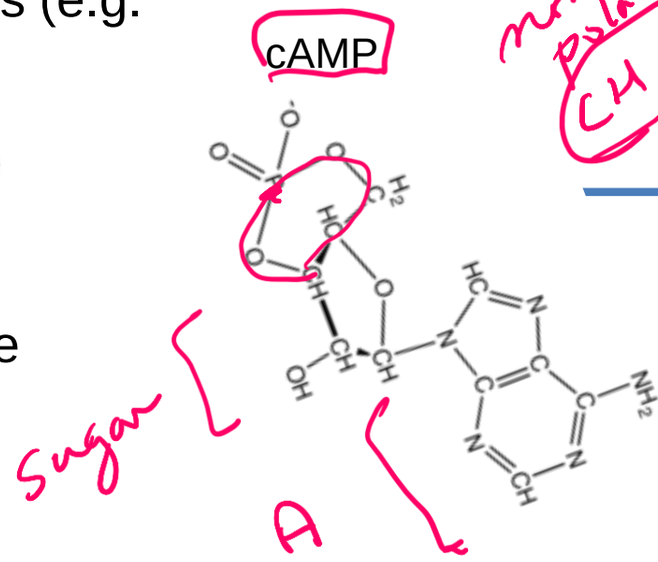
Ligand: Something that binds to a protein, usually small molecules (e.g. cyclicAMP, cAMP).

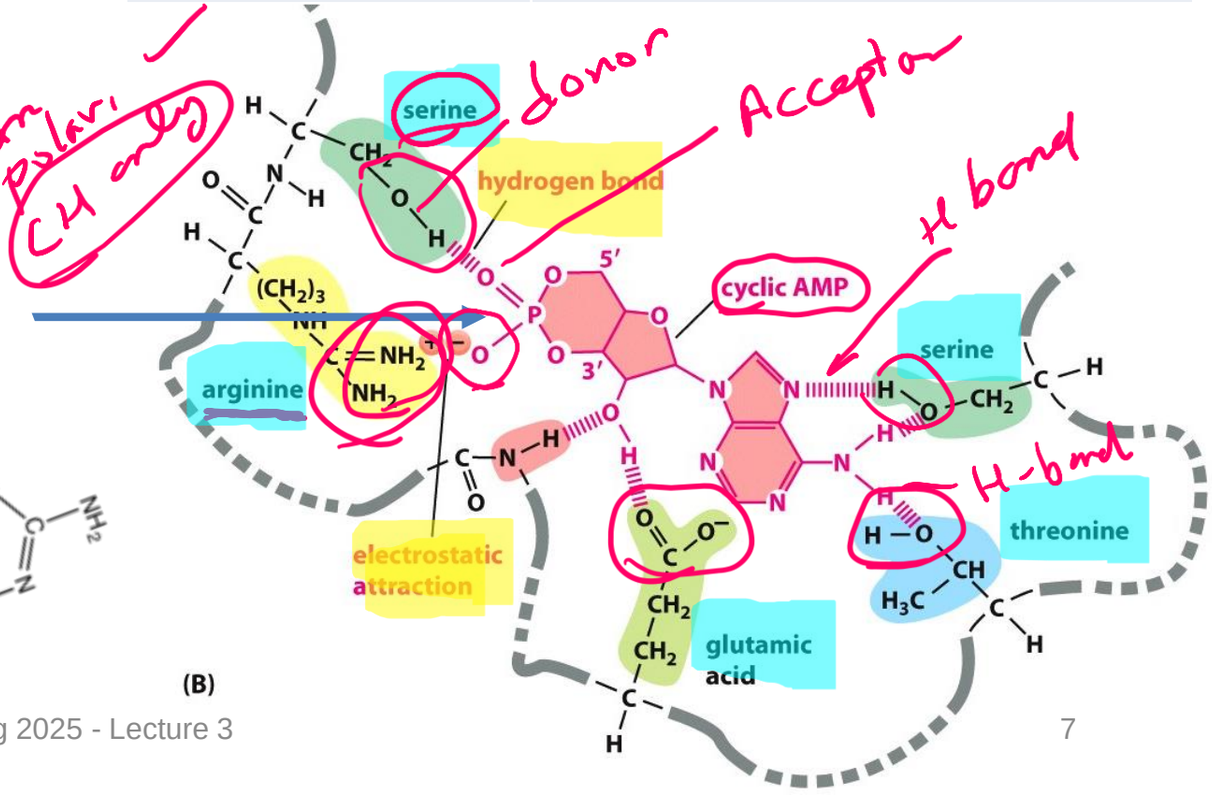
Binding site allow a protein to interact with specific **ligands**

Binding site is generated by the **folded** form of the protein.

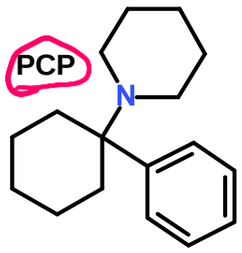
The bound ligand can be stabilized by any and all of the following interactions:

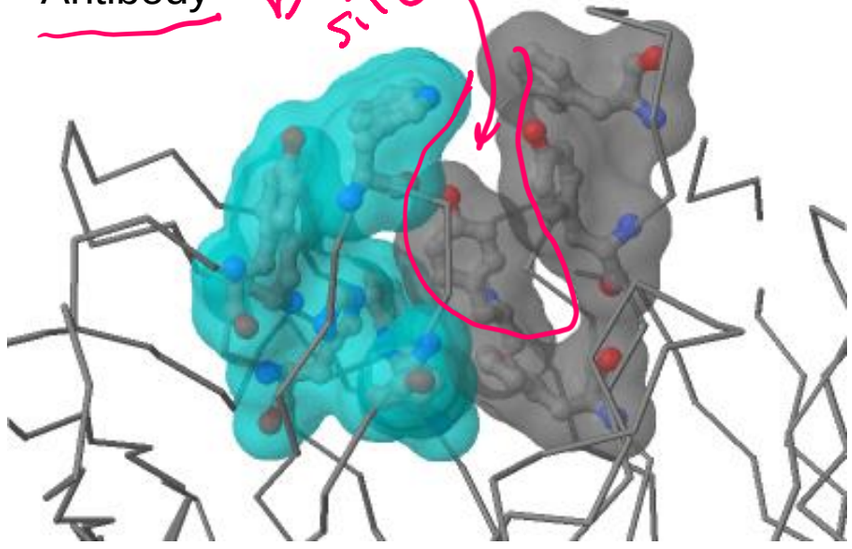
Interaction	Which stabilize cAMP Binding?
Electrostatic	✓
van der Waals	✓
H-Bonding	✓
Hydrophobic effect	✗



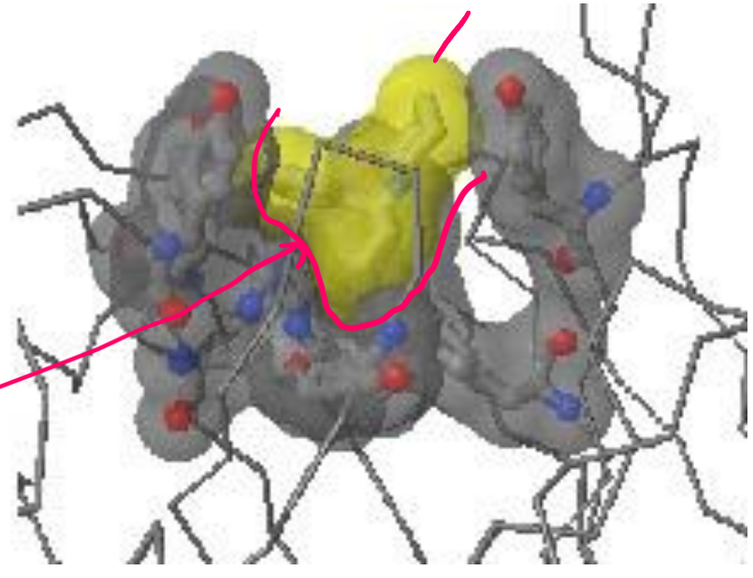


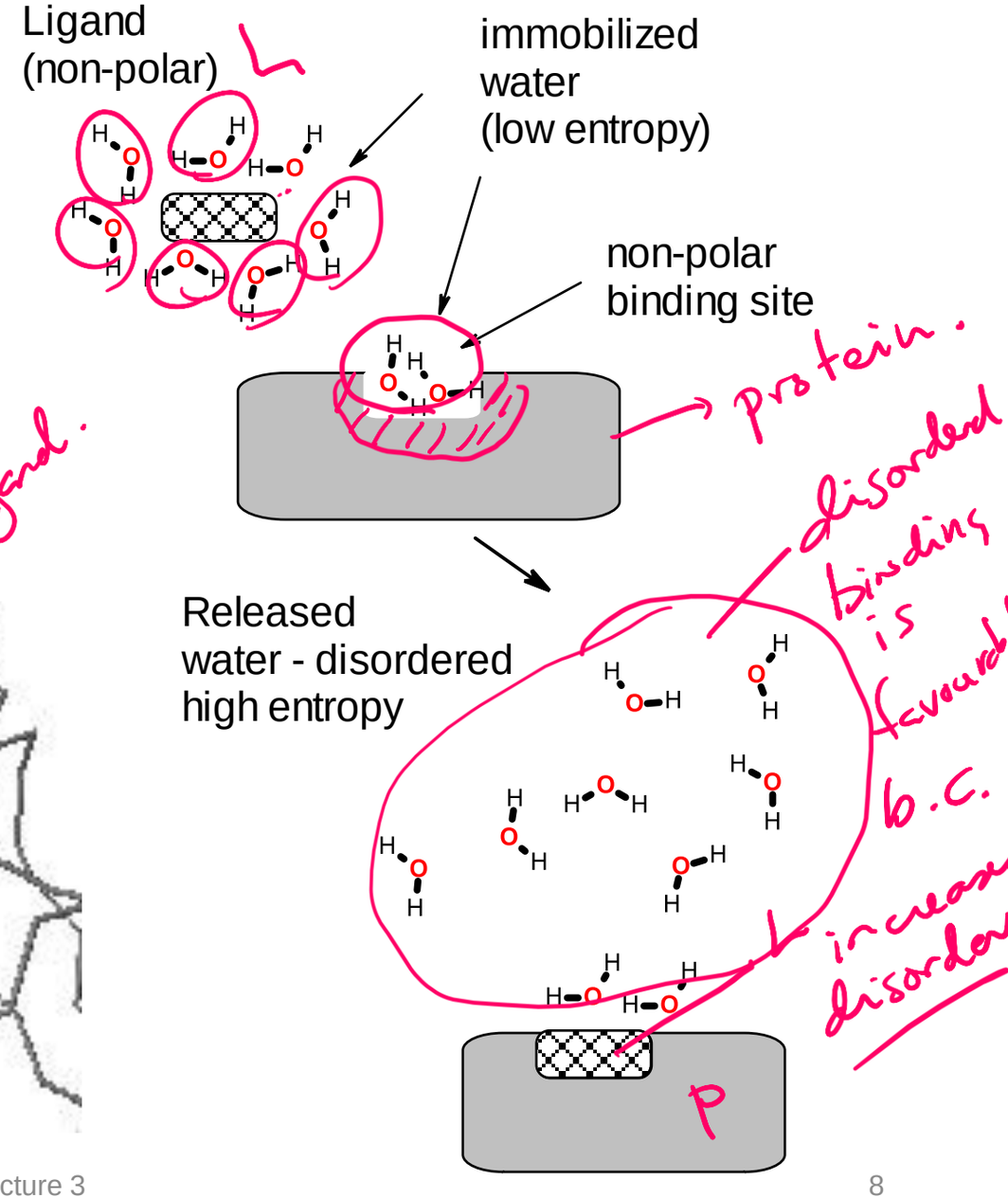
Hydrophobic Effect Drives Binding of Non-polar Ligands





Antibody-PCP complex





Ligand Binding & Saturation:

Define fraction saturated: $Y = \frac{[ML]}{[M]+[ML]}$

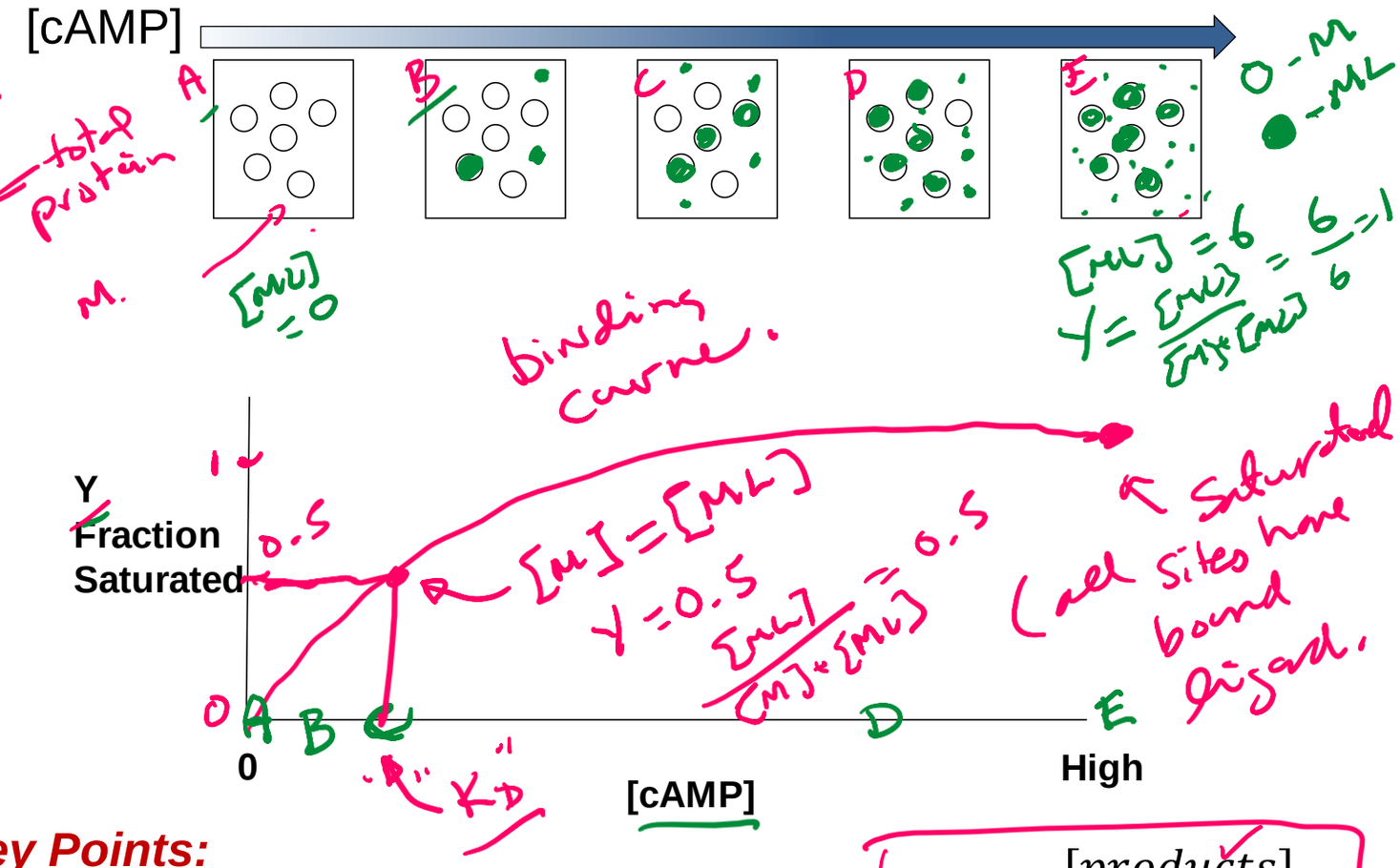
[M] = free macromolecule (e.g. antibody with no antigen).

[ML] = macromolecule with ligand bound (e.g. antibody with antigen bound).

The boxes with circles represent proteins with no cAMP bound, each box (left to right) is at a higher [cAMP]. Filled circles indicate bound ligand.

1. How will the number of filled circles depend on the cAMP concentration?

2. Plot the fraction saturated data point for each box.



Key Points:

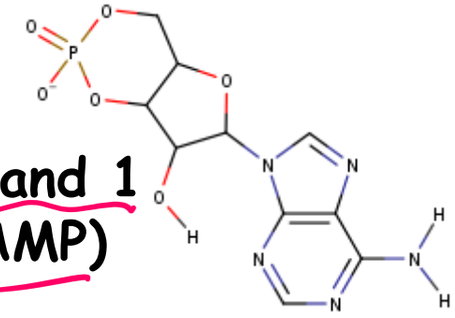
1. The binding sites saturate, when all are full no more ligand can bind.
2. There is a ligand concentration, [L], where ½ the sites are full. This [L] is K_D
3. K_D is the equilibrium constant for ligand dissociation:

$$(ML) \rightleftharpoons (M) + (L)$$

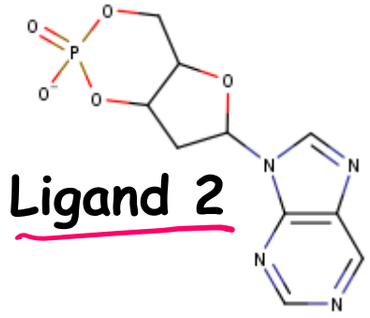
$$K_{Eq} = \frac{[products]}{[reactants]}$$

$$K_D = \frac{[M][L]}{[ML]}$$

Using K_D to Compare Ligand Binding



Ligand 1 (cAMP)



Ligand 2

The binding of two different molecules to the same protein was measured and the data is shown on the right. L1 is cAMP, L2 is similar to cAMP

Which ligand has a K_D of 1? L1 or L2?

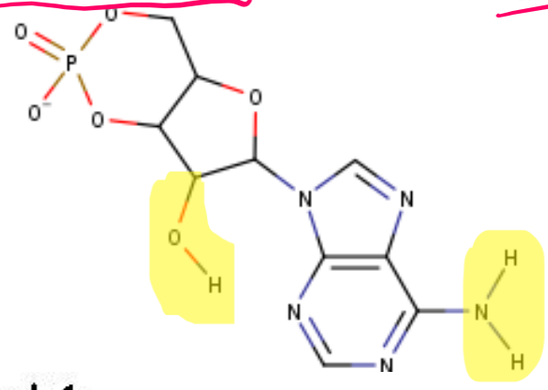
Which ligand has a K_D of 10? L1 or L2?

Which ligand binds more tightly to the protein (higher affinity)? L1 or L2?

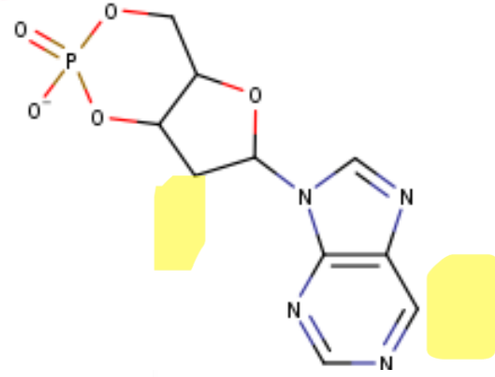
***Why** does L1 bind more tightly (higher affinity)?*

1. What are the chemical differences between L1 and L2 (Upper diagram)

Ligand 1 (cAMP)



Ligand 2

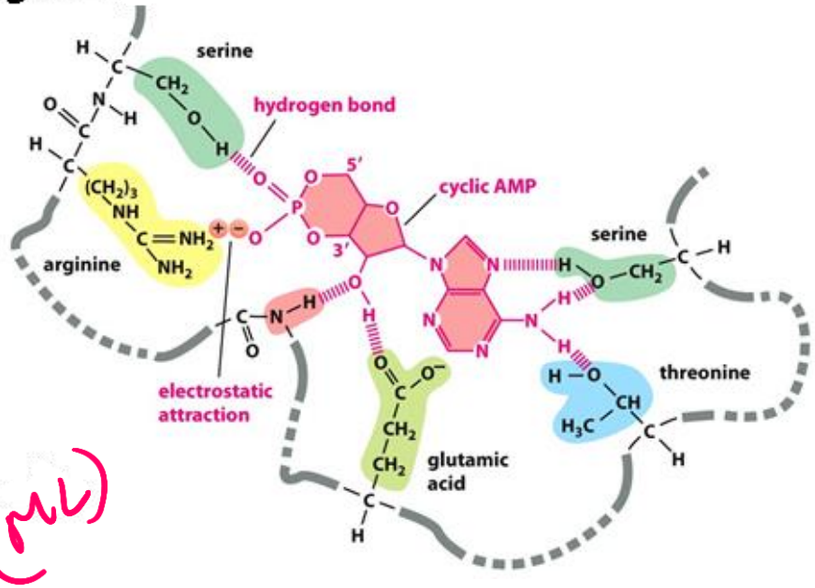


2. How do these differences affect the interactions with the protein (lower diagram)?

What interactions are lost.

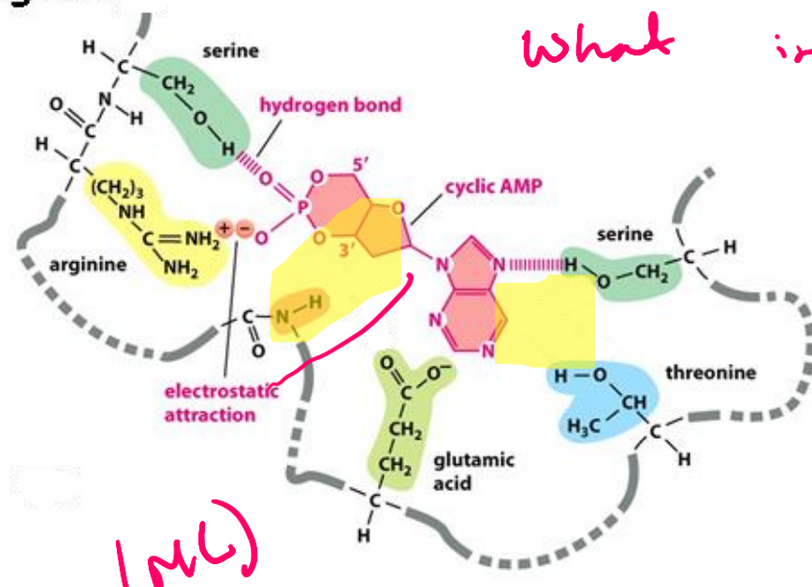
H-bonds.

Ligand 1



(ML)

Ligand 2



(ML)

3. How do the differences affect K_D?

increased the K_D.

→ more interactions = higher affinity = lower K_D

Key Points:

Binding:

Folded proteins have ***binding sites*** that recognize other molecules (***ligands***) using ***any and all*** of the following:

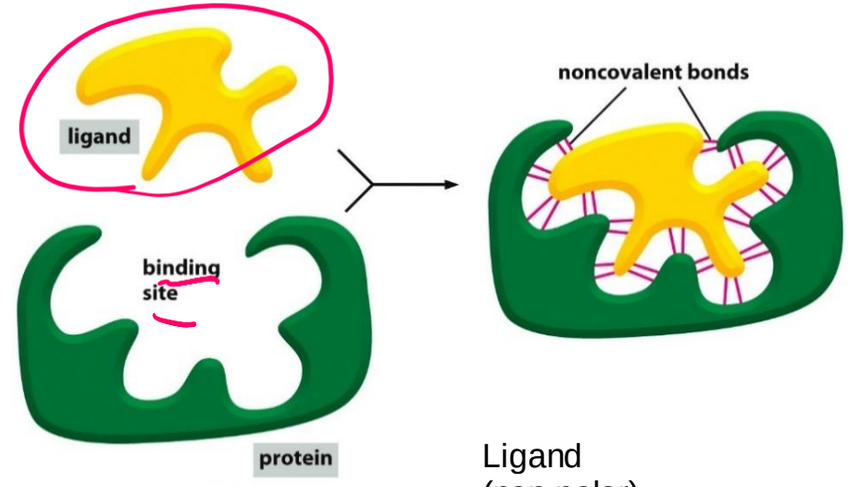
- H-bonds,
- van der Waals,
- Electrostatic,
- Non-polar interactions (hydrophobic)

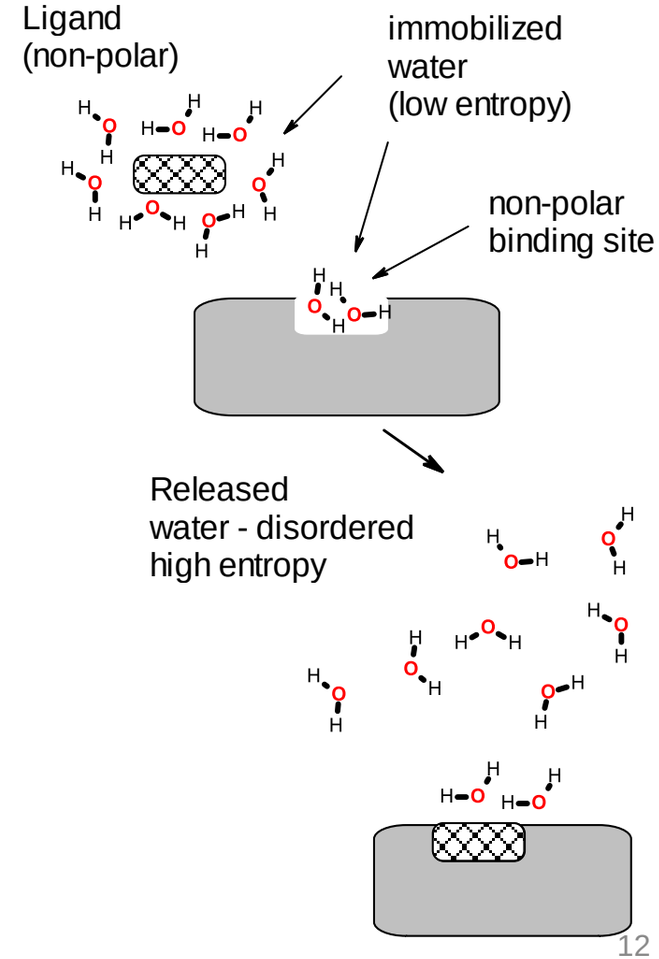
Binding is **reversible**

Binding is **saturable**

Binding ½ point (Y=0.5) occurs at K_D

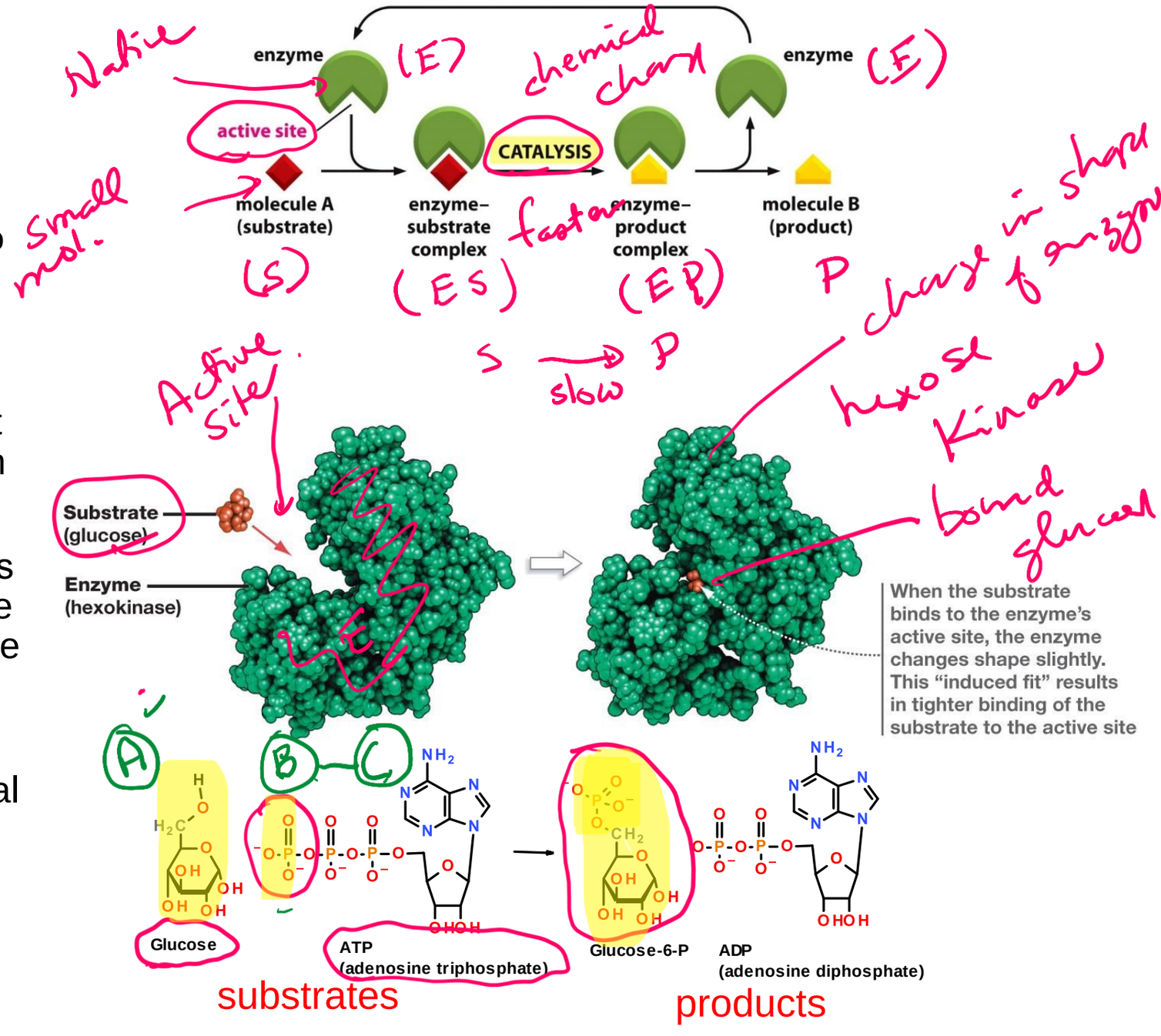
The higher the affinity (strength of interaction), the lower the K_D





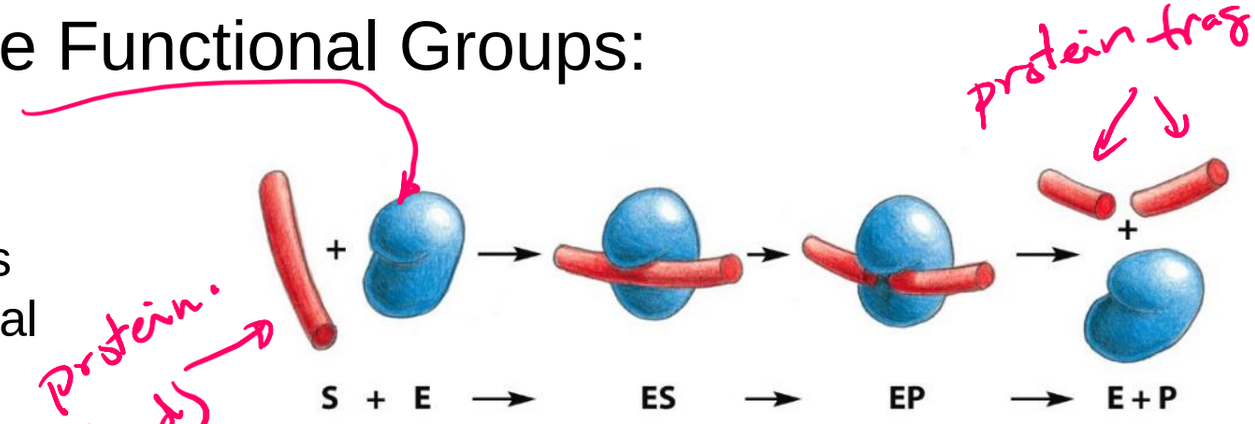
Enzymes

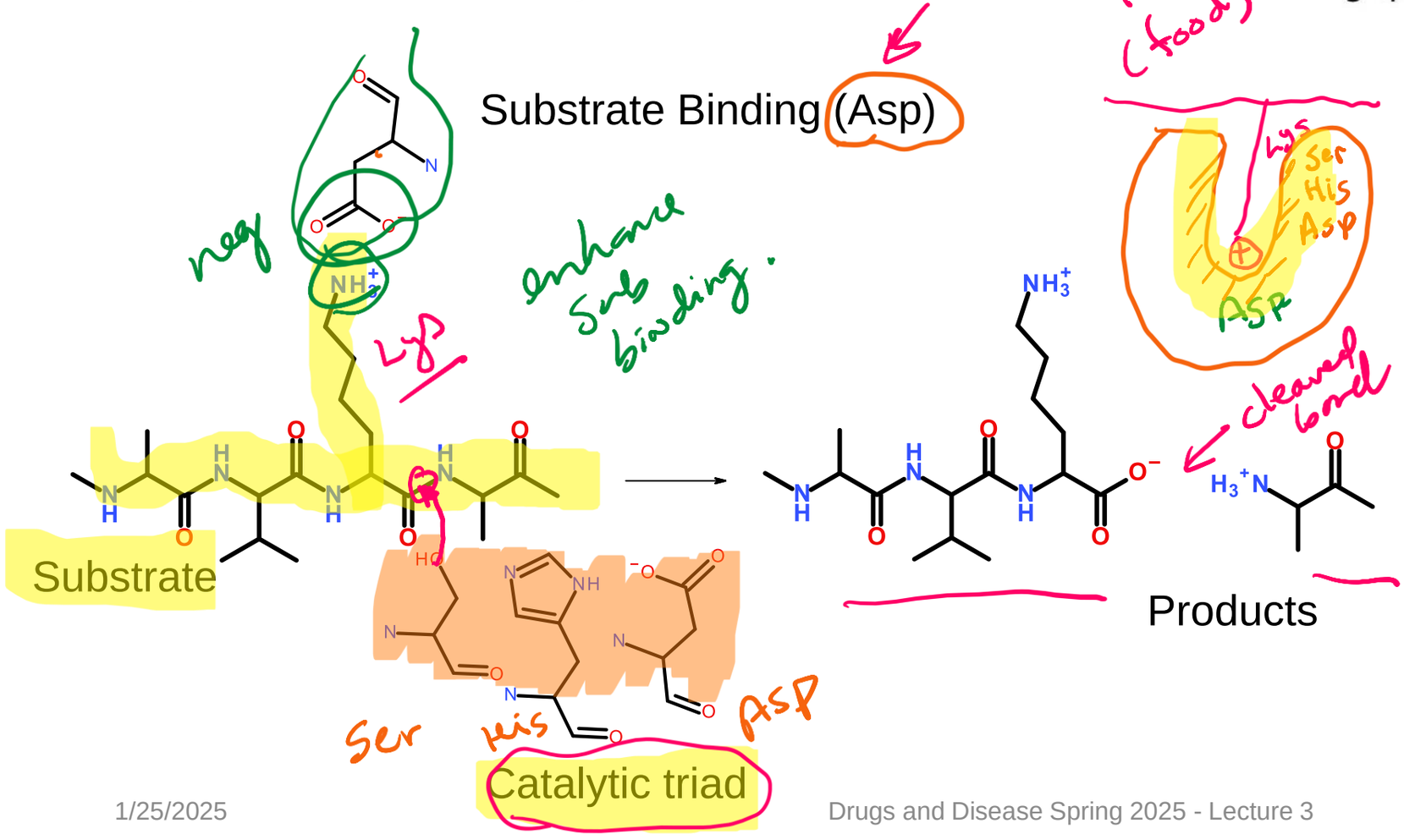
- **Enzymes** are protein or RNA catalysts. They increase the rate of the reaction.
- They bind “substrates” and convert them to “products”. Usually, the substrate undergoes a chemical reaction and is changed in its structure.
- Most biological chemical reactions occur at meaningful rates only in the presence of an enzyme.
- Substrates bind specifically to the enzyme’s **active site,** interacting with amino acid side chains (or RNA bases). Usually, an enzyme binds one substrate.
- The chemical change caused by the enzyme is catalyzed by additional functional groups in the active site.
- Many enzymes undergo a conformational change when the substrates are bound to the active site; this change is called an **induced fit**.

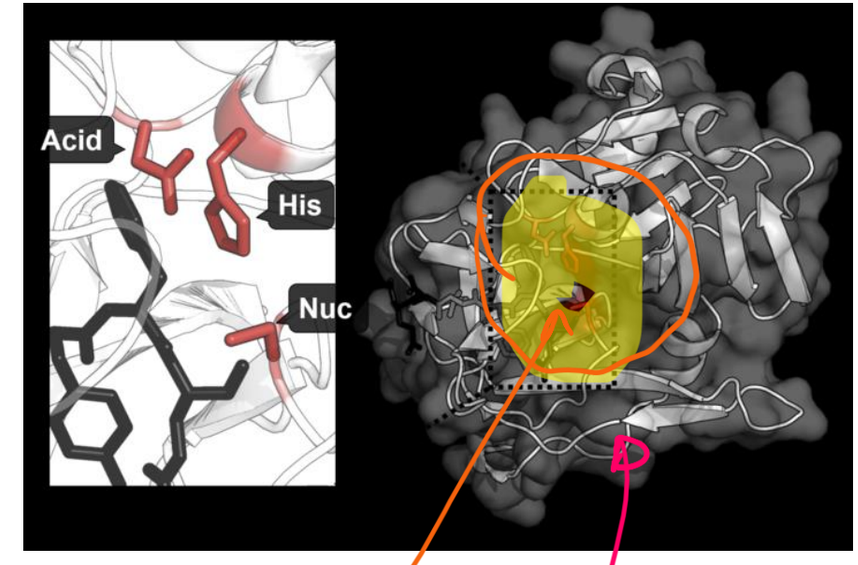


Example of Active Site Functional Groups:

- Catalytic triad (Asp, His, Ser) in Protease Trypsin cleaves the peptide bond.
- More active with Lys and Arg containing substrates because of a favorable interaction with an additional Asp residues in the enzyme.







https://shirleychemproject.weebly.com/

Active site.

trypsin

Enzyme – Chemical Diversity

TABLE 4–1 SOME COMMON FUNCTIONAL CLASSES OF ENZYMES

ENZYME CLASS	BIOCHEMICAL FUNCTION
Hydrolase	General term for enzymes that catalyze a hydrolytic cleavage reaction.
Nuclease	Breaks down nucleic acids by hydrolyzing bonds between nucleotides.
Protease	Breaks down proteins by hydrolyzing peptide bonds between amino acids.
Synthase	General name used for enzymes that synthesize molecules in anabolic reactions by condensing two molecules together.
Isomerase	Catalyzes the rearrangement of bonds within a single molecule.
Polymerase	Catalyzes polymerization reactions such as the synthesis of DNA and RNA.
Kinase	Catalyzes the addition of phosphate groups to molecules. Protein kinases are an important group of kinases that attach phosphate groups to proteins.
Phosphatase	Catalyzes the hydrolytic removal of a phosphate group from a molecule.
Oxido-reductase	General name for enzymes that catalyze reactions in which one molecule is oxidized while the other is reduced. Enzymes of this type are often called oxidases, reductases, or dehydrogenases.
ATPase	Hydrolyzes ATP. Many proteins with a wide range of roles have an energy-harnessing ATPase activity as part of their function, including motor proteins such as myosin and membrane transport proteins such as the sodium–potassium pump.

- Most enzyme names end in "-ase"
- Usually named by their substrates and the reactions they catalyse, i.e. glucose kinase

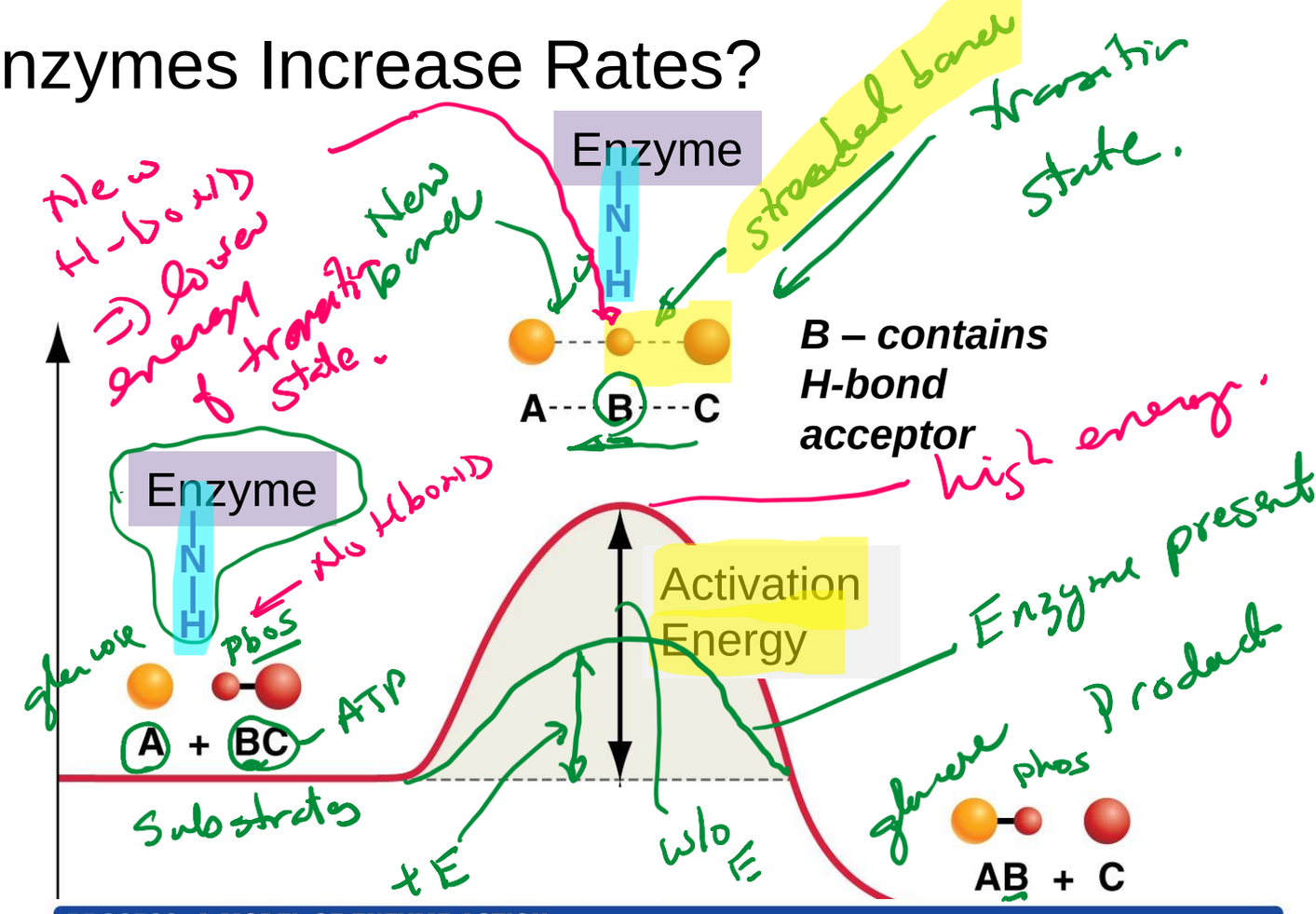
How Do Enzymes Increase Rates?

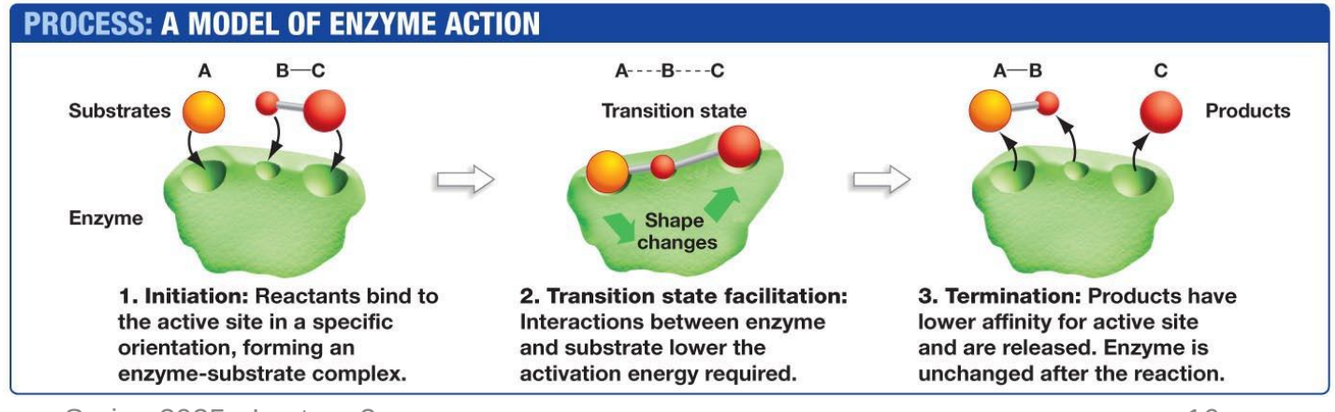
- **Transition state** = high energy intermediate that occurs during the reaction.
- Energy barrier is called the activation energy.
- Rate of product formation depends on the concentration of the transition state.

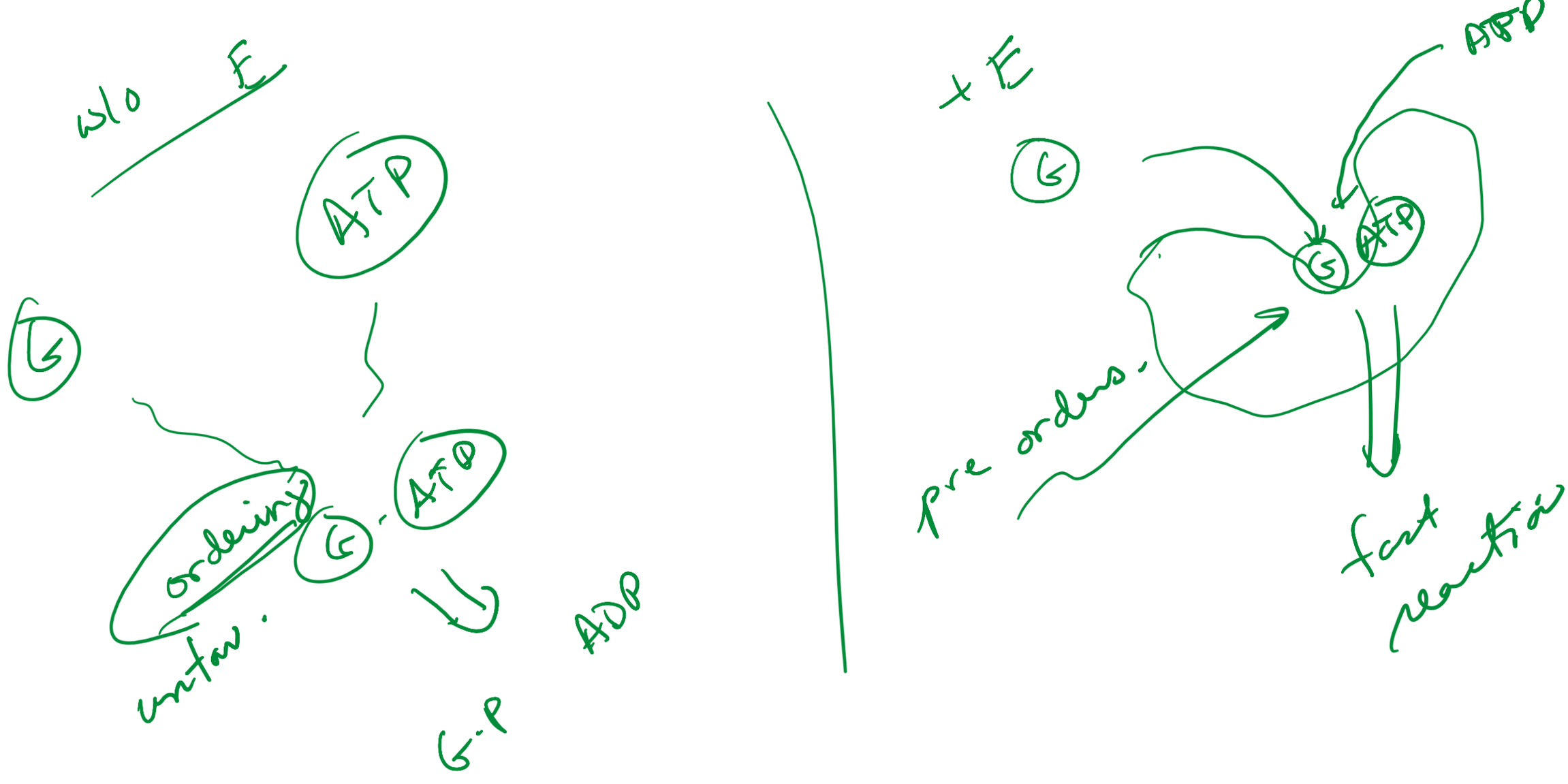
Low [X] = Slow reaction

Higher [EX] = Faster reaction

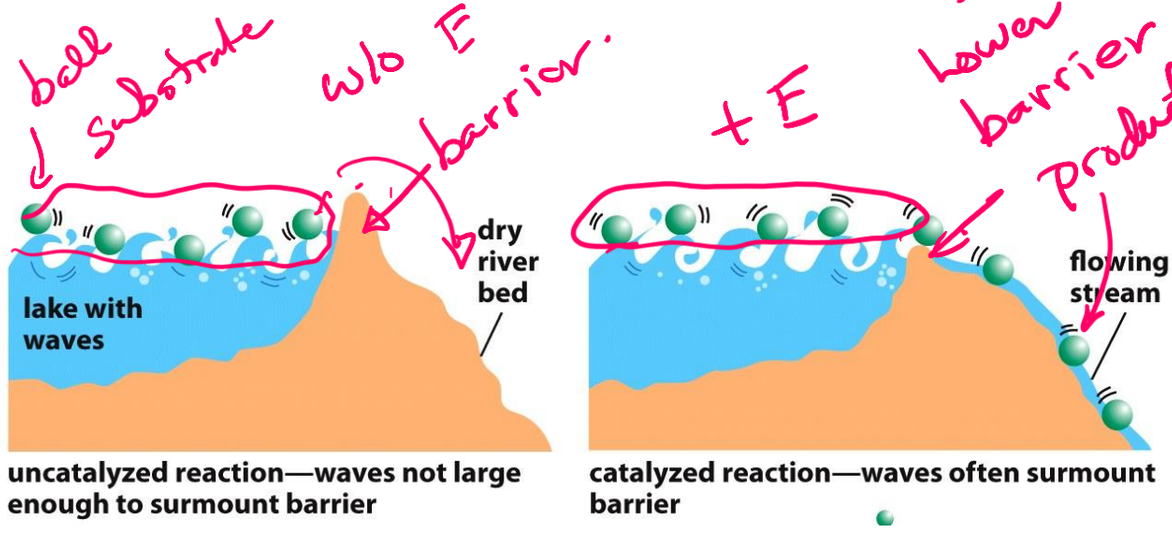
- Interactions between the enzyme and the substrate stabilize the **transition state** (X) and lower the activation energy required for the reaction to proceed.
- Stabilization can include:
  - Pre- alignment of key groups in the active site, reducing entropy cost of organizing groups.
  - Direct interactions with just the transition state, e.g. formation of new H-bonds.

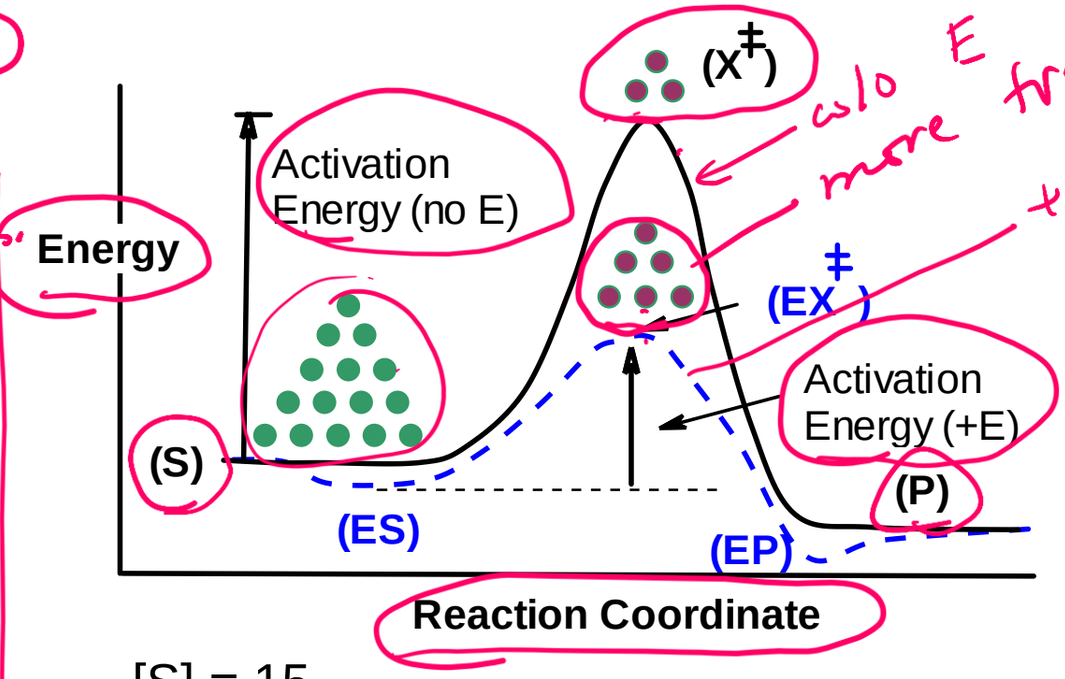





w/o F
ATP
G
ordering
G-ATP
G-P
ADP
+F
G
ATP
G
ATP
pre orders.
fast reaction

A model of transition state stabilization.





- Lower energy of transition state allows more substrates to reach transition state due to their thermal energy.

[S] = 15

[X] = 3

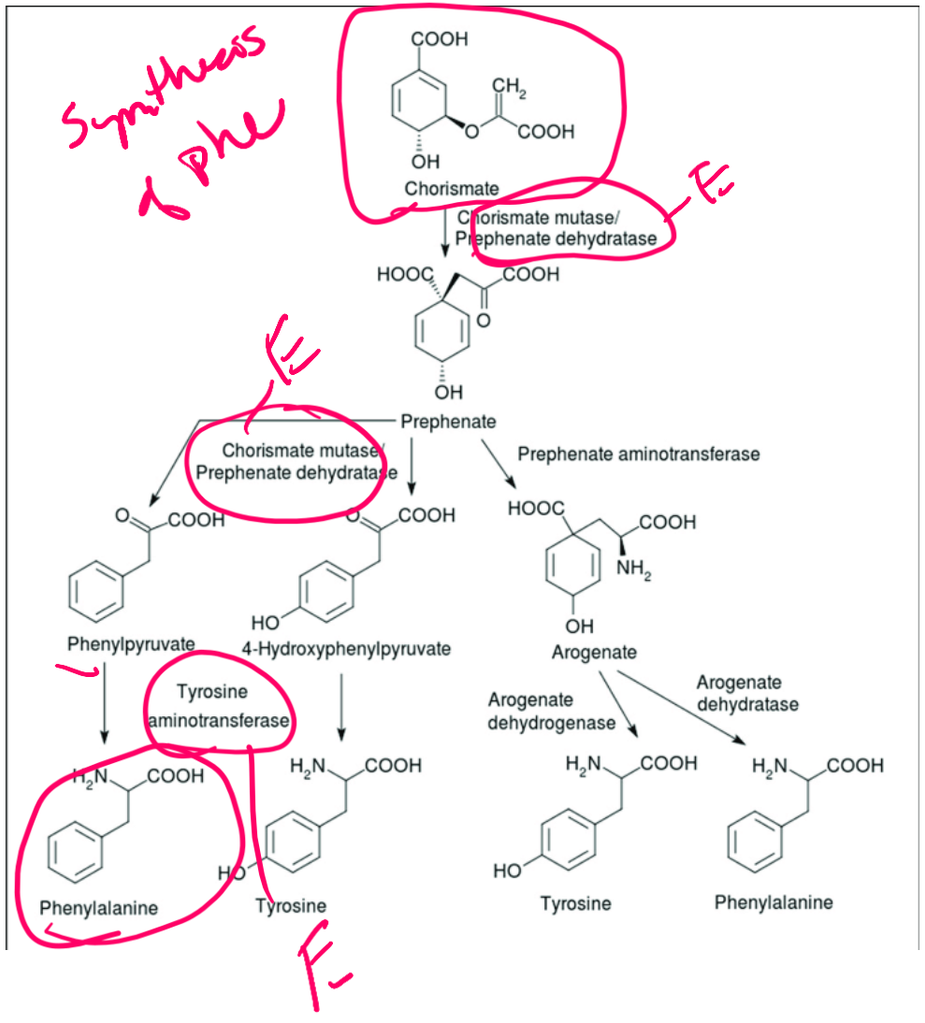
[EX] = 6

How much faster will the rate be when the enzyme is present?

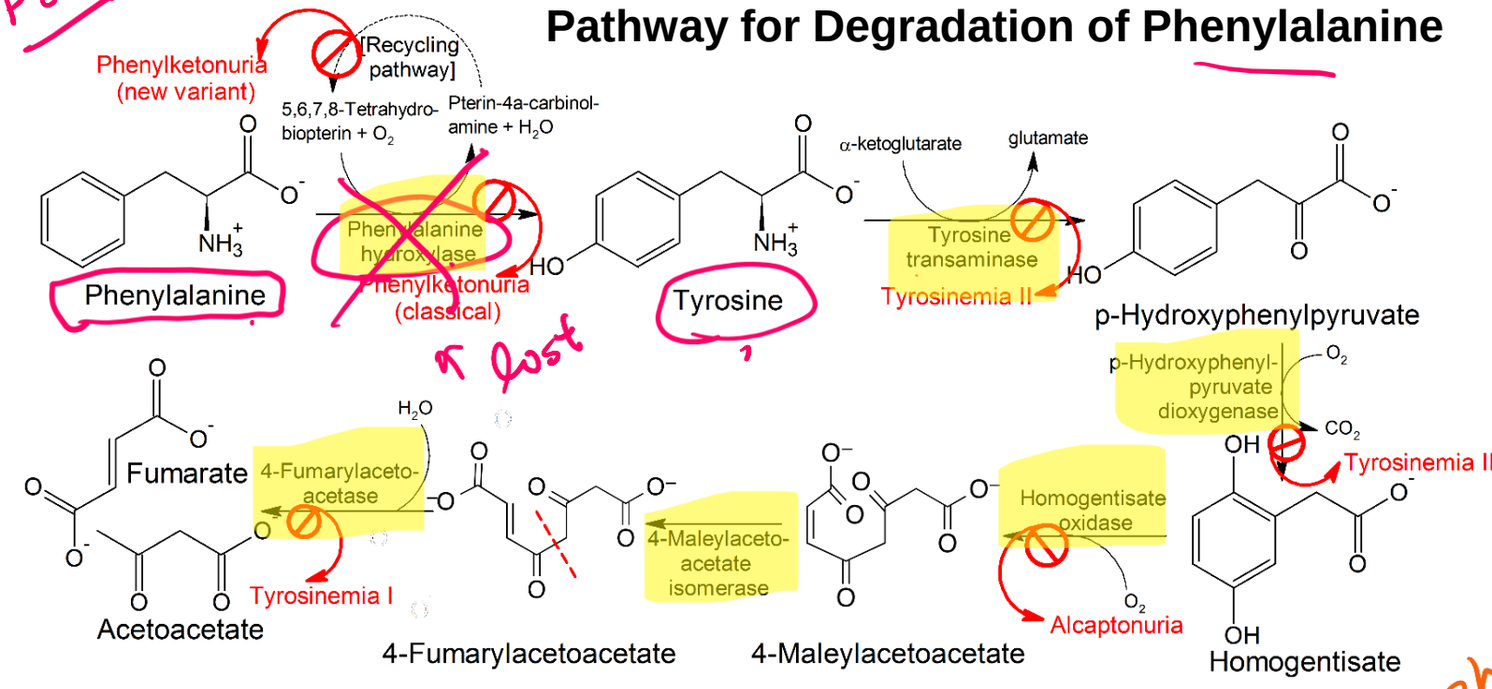
Enzymes, Metabolic Pathways, and Diseases

Synthetic Pathway for Phe, Tyr
(beginning with chorismite)

- Each step catalyzed by an enzyme



Pathway for Degradation of Phenylalanine

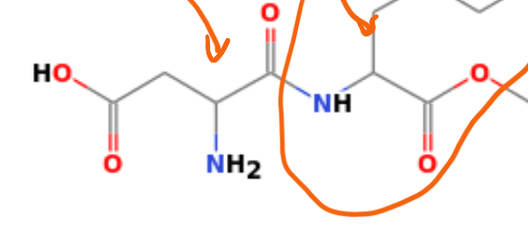


PKU Disease:

- Inactive phenylalanine hydroxylase
- Phe levels become toxic:
- Neurological problems
  - Intellectual disability
  - Developmental delays
  - Mental health disorders.

Aspartame (artificial sweetener)

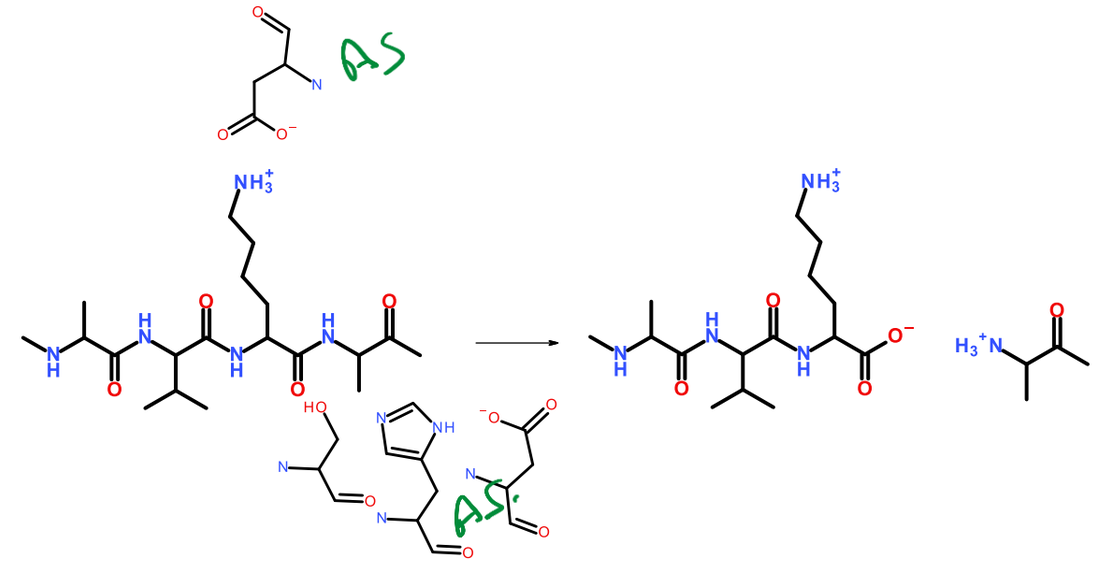
Asp-Phe-CH_3

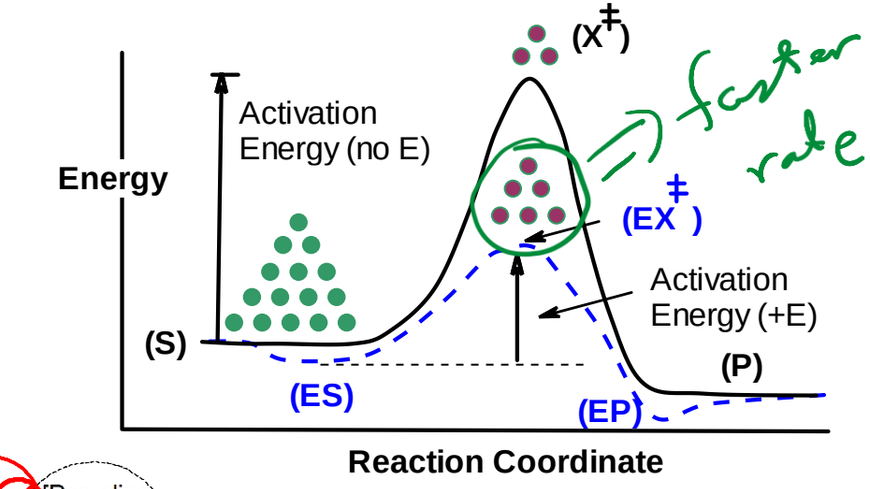


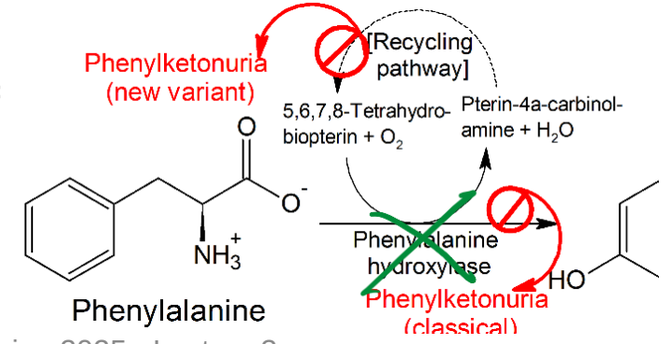
Key Points:

Enzymes:

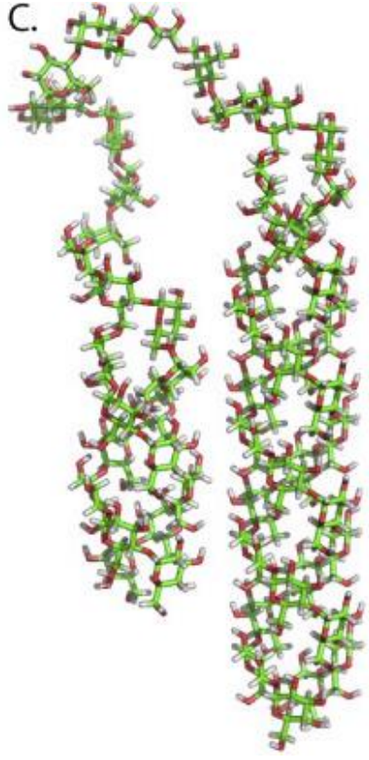
- Enzymes bind substrates (S), forming (ES) complex in active site, converting to P, releasing P.
- Rate enhancement since the transition state complex (EX) forms more readily with enzymes due to:
  - Bringing substrates and functional groups on the enzyme together by binding (less entropy change)
  - Directly lowering energy of transition state (X) through favorable interactions that are unique to the transition state, such as forming unique hydrogen bonds.
- Genetic diseases that lead to inactive metabolic enzymes can cause disease due to the build-up of toxic intermediates.



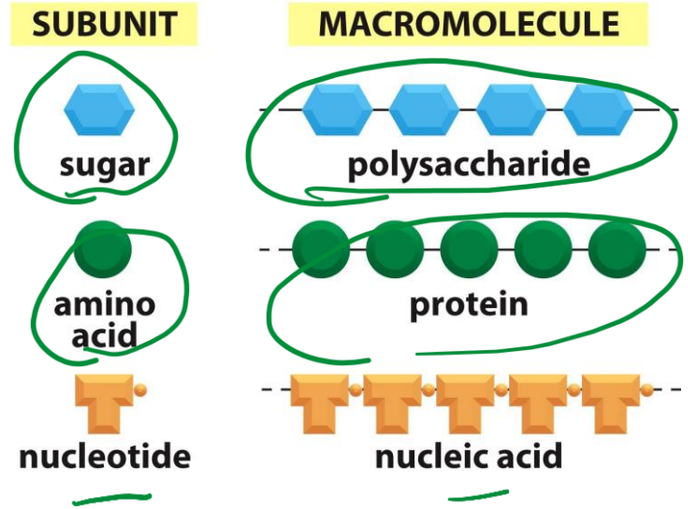




Carbohydrates



Polysaccharide



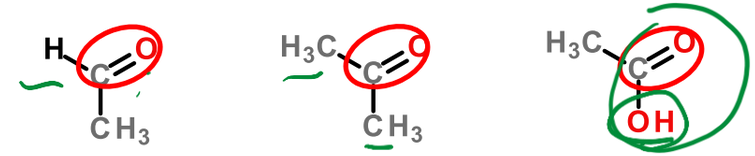
- General structure of monomers
- Disaccharides (e.g. lactose)
- Glycogen (glucose storage)
- Bacterial cell wall structure (antibiotic target)

- Lactose intolerance
- Glycogen storage disease

Carbohydrates

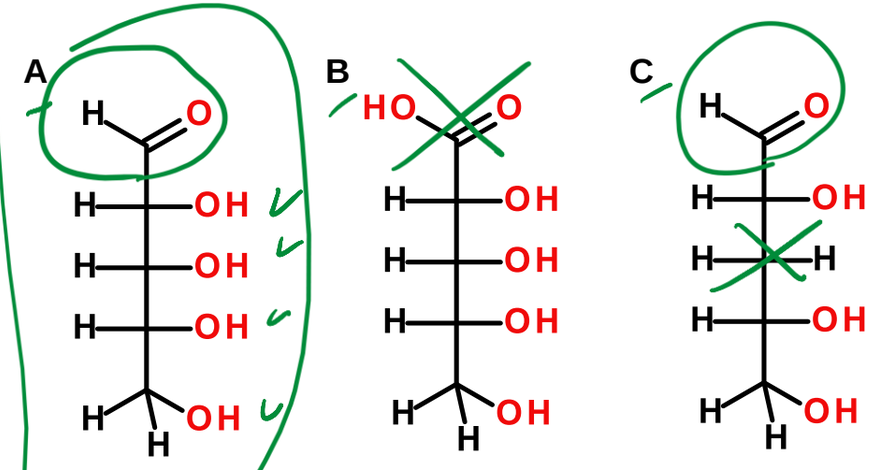
- Monosaccharides (one sugar),
- oligosaccharides (few sugars)
- polysaccharides (many sugars)
- Chemical formula is $(CH_2O)_n$ (e.g. hydrated carbon)
- They are molecules with:
  - one aldehyde or ketone group, on 1st or 2nd carbon
  - -OH group on *all* other carbons, leading to a chiral carbon for most carbons.

Functional groups:



aldehyde ketone carboxylic acid

Carbonyl group –> C=O



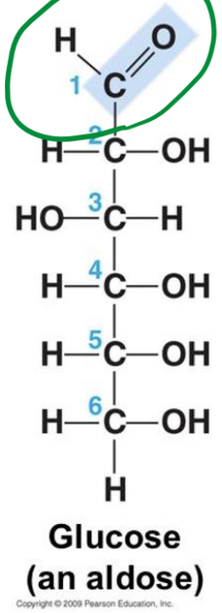
Only one of these is a carbohydrate, which one?

A B C

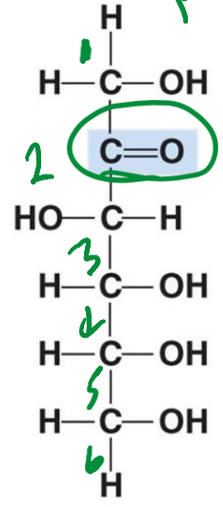
3 ways simple sugars (monosaccharides) differ from each other

1. **Location of the carbonyl group**
2. Number of carbons
3. Spatial arrangement of atoms (the position of the OH groups)

Aldose: Carbonyl group is located on C_1



Glucose (an aldose)

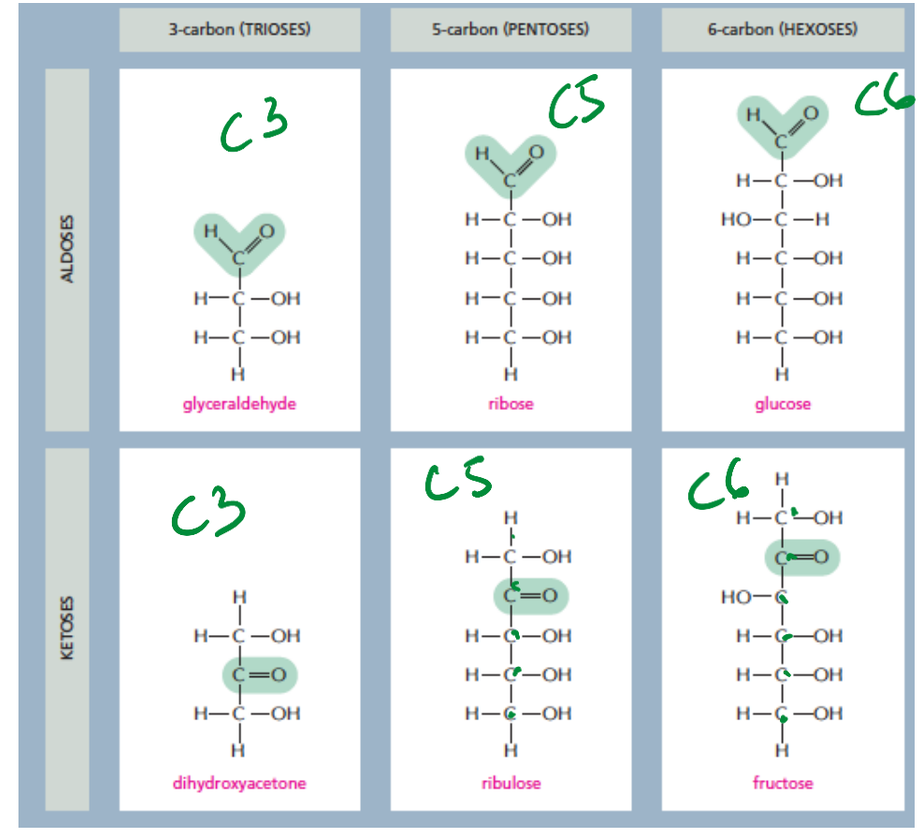


Fructose (a ketose)

Numbering carbons: Carbon 1 is at the end closest to the C=O group.

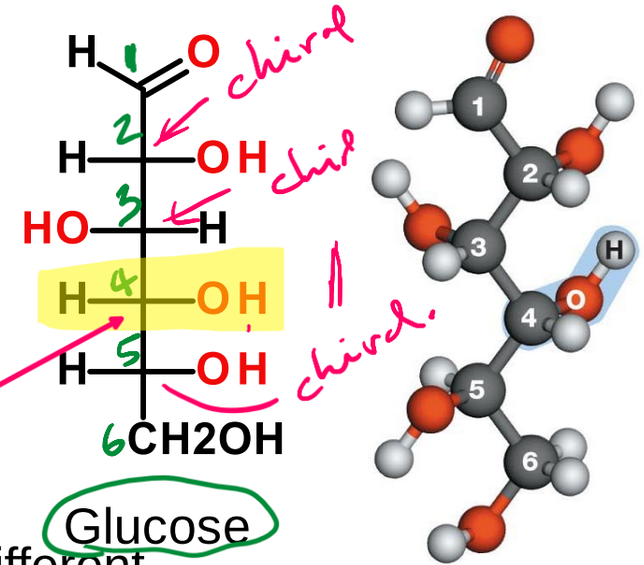
What carbon is the carbonyl?

1. Location of the carbonyl group
2. **Number of carbons**
3. Spatial arrangement of atoms (the position of the OH groups)

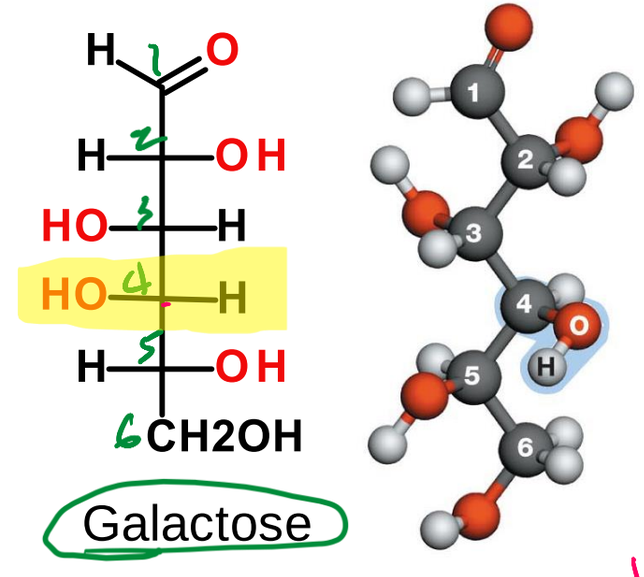


3 ways simple sugars (monosaccharides) differ from each other

1. Location of the carbonyl group
2. Number of carbons
3. **Spatial arrangement of atoms (the position of the OH groups)**



Glucose



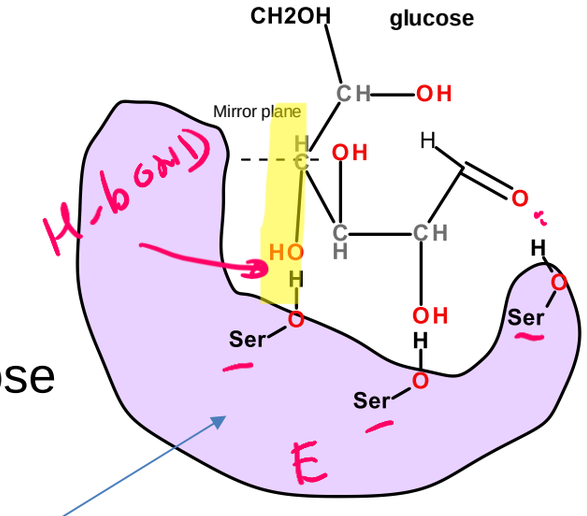
Galactose

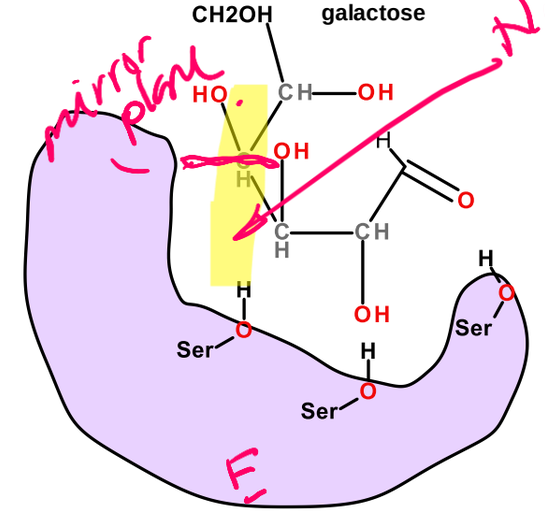
Both have the same chemical formula $C_6H_{12}O_6$. Both are aldose sugars with 6 carbons.
Yet their functions are different.

- Glucose can be used for energy immediately.
- Galactose has to be converted to glucose before it can be used for energy.

They have different interactions with enzymes due to the different chirality at carbon 4.

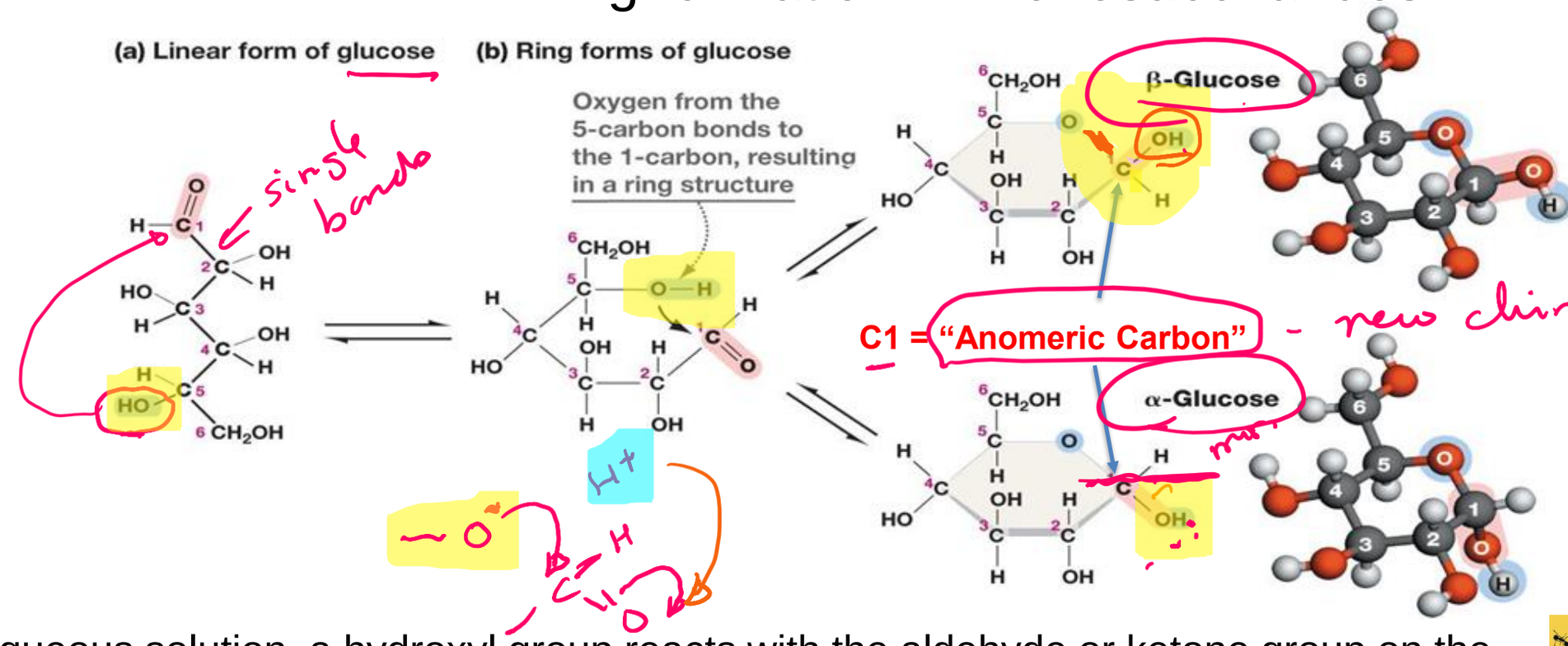
- OH is down in glucose
- OH is up in galactose





Enzyme specific for a-glucose

Ring formation in Monosaccharides:



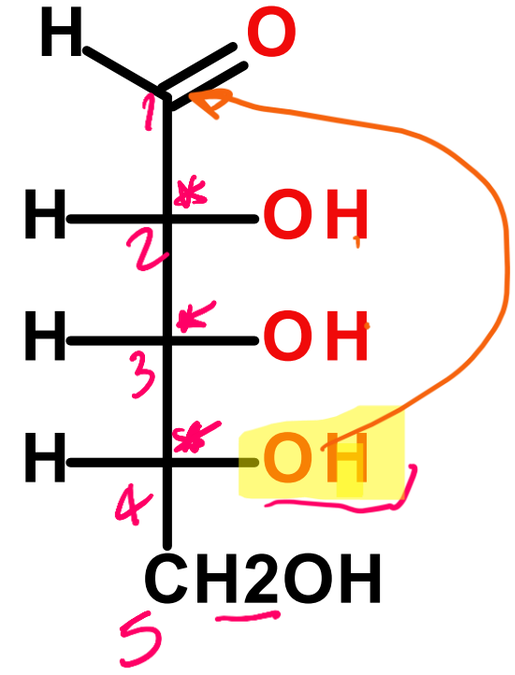
- In aqueous solution, a hydroxyl group reacts with the aldehyde or ketone group on the same molecule, closing the molecule into a ring, with a bridging oxygen
- It is usually the 2nd to last -OH group, i.e. C5 in glucose, C4 in ribose.
- Stable ring sizes are 5 atoms or 6 atoms
- No atoms are lost or gained in this reaction.
- The carbonyl carbon becomes chiral and is called the ***anomeric carbon*****.**
- The rings with different chirality at C1 are different:

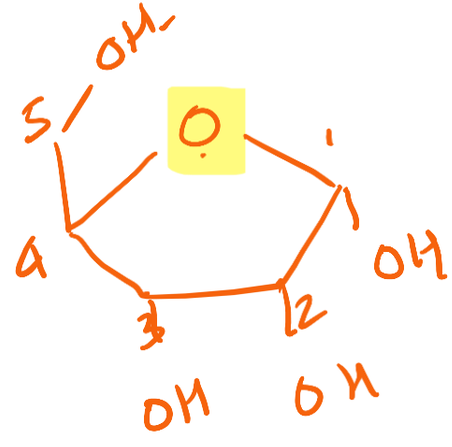
  α (new OH is down), β (new OH is up) *"(ants are down, birds are up)"*

Example Problem:

The linear form of ribose, a 5 carbon aldose is shown on the right. This sugar is found in RNA (ribonucleic acid).

1. Number the carbons.
2. Which carbons are chiral? Mark them with a *.
3. Draw the cyclic form of α-ribose





α-ribose.

Disaccharides

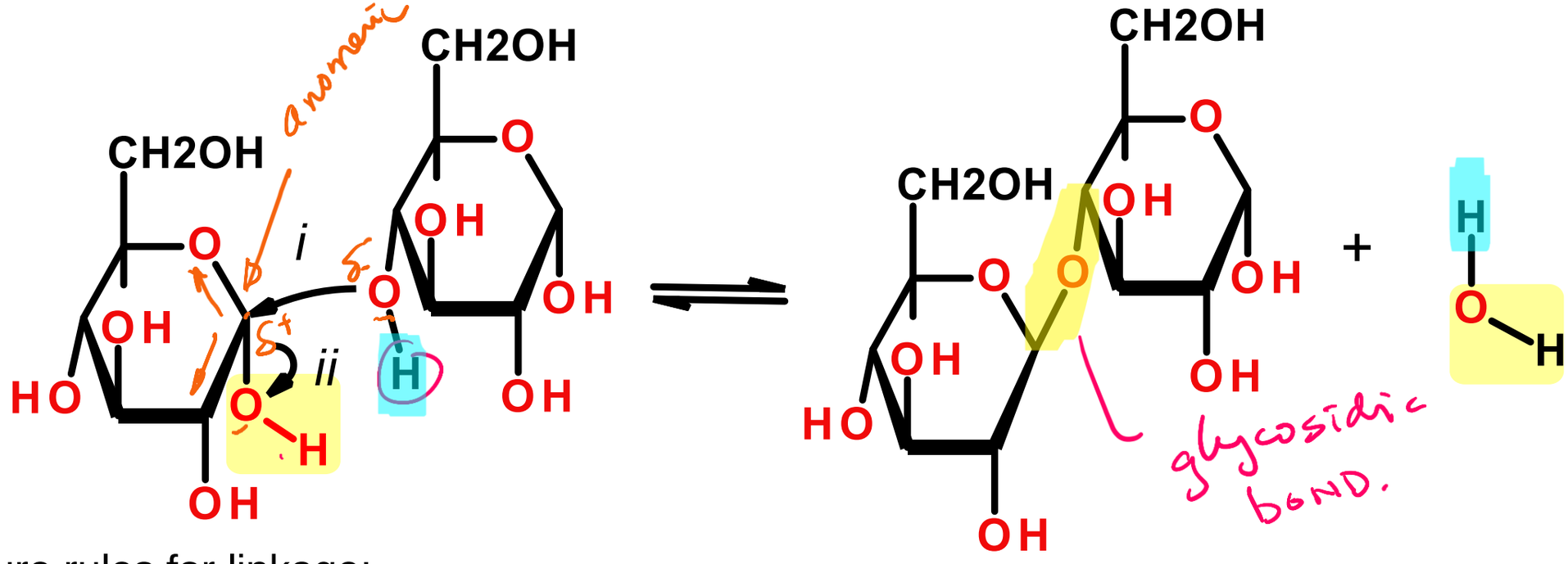
Linkage of the anomeric carbon of one monosaccharide to the OH of another monosaccharide via a *condensation* reaction.

The bond is termed a ***glycosidic bond:***

i) The anomeric carbon is the site of attack by another -OH group.

ii) A water is released

Why is the anomeric carbon the preferred site for nucleophilic attack?



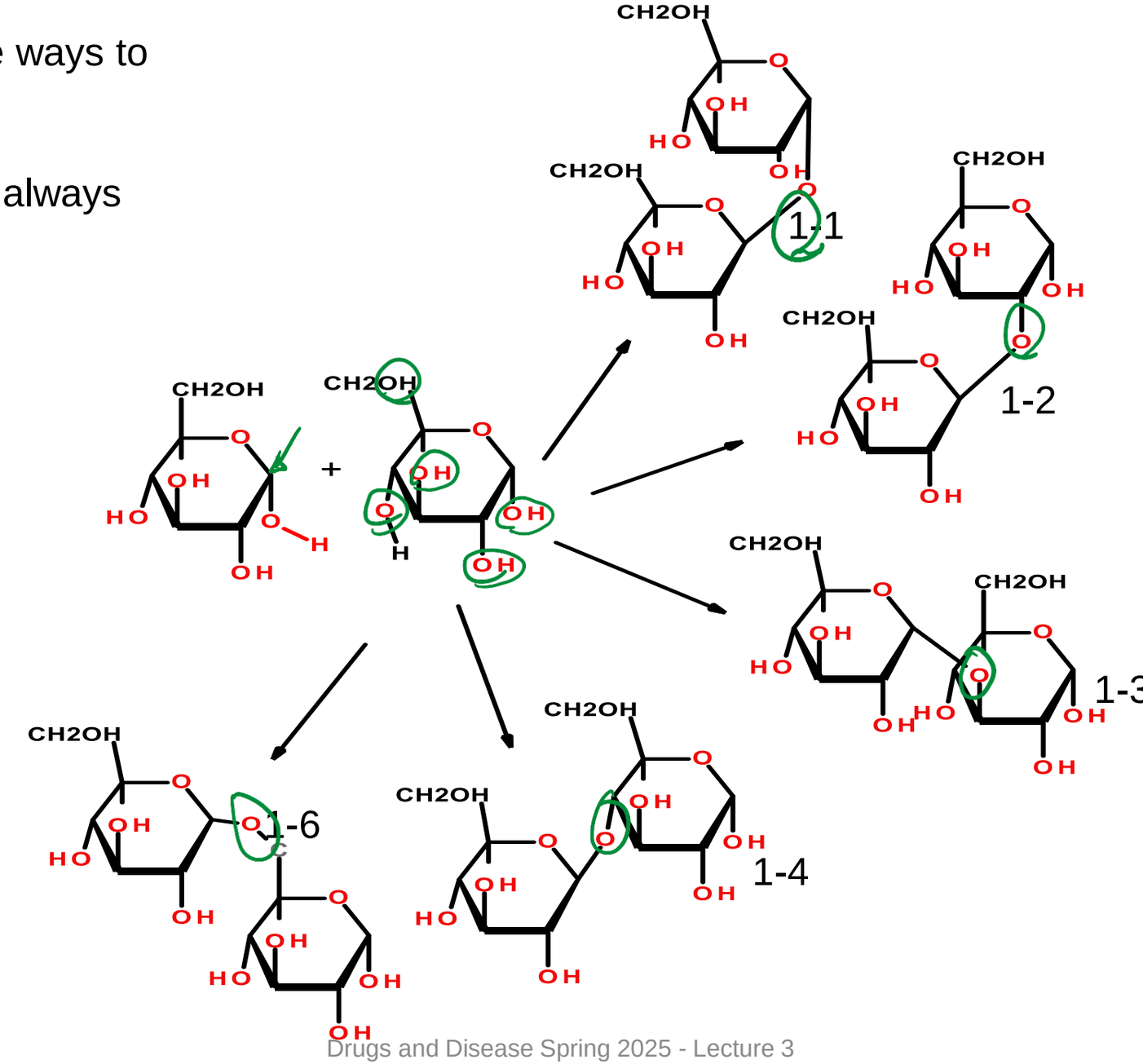
Nomenclature rules for linkage:

- Orientation of the **anomeric** involved in the linkage (α oxygen is down, β oxygen is up)
- Carbons involved in the linkage (e.g. 1-4)

Disaccharides

There are many possible ways to connect two sugars.

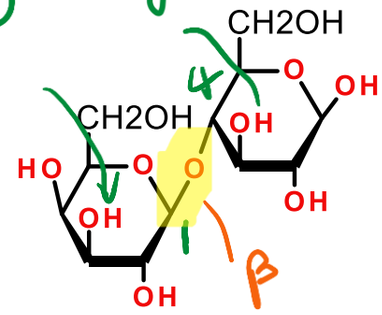
At least one anomeric is always involved.

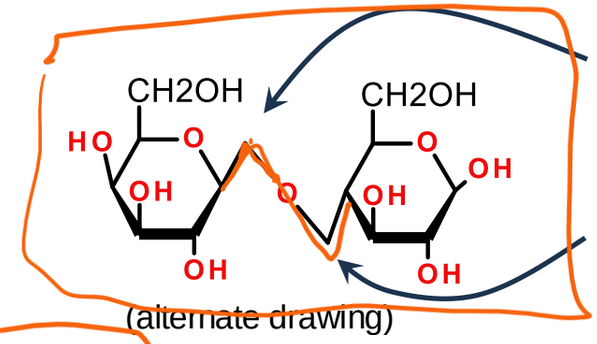


many ways to link two sugars together.

Disaccharides

Lactose (milk sugar)



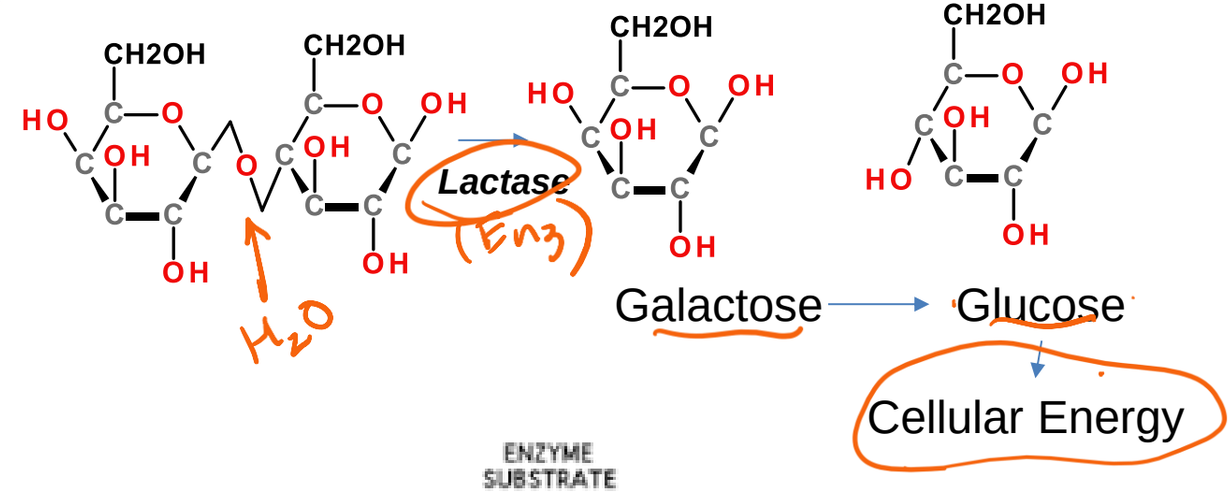


(alternate drawing)

These kinks are not carbons but are drawn in this way to indicate that the chirality of the anomeric is beta (pointing up). The kinks allow the line to reach the downward pointing –OH on C4 in glucose.

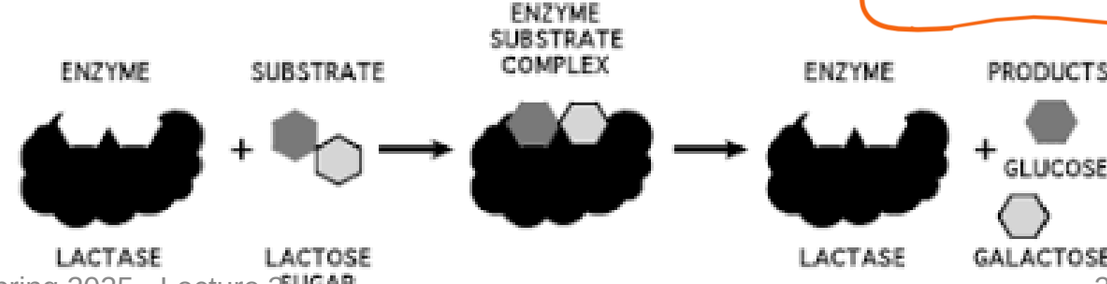
β-galactopyranosyl-(1→4)-β-glucopyranose

Metabolism of Lactose

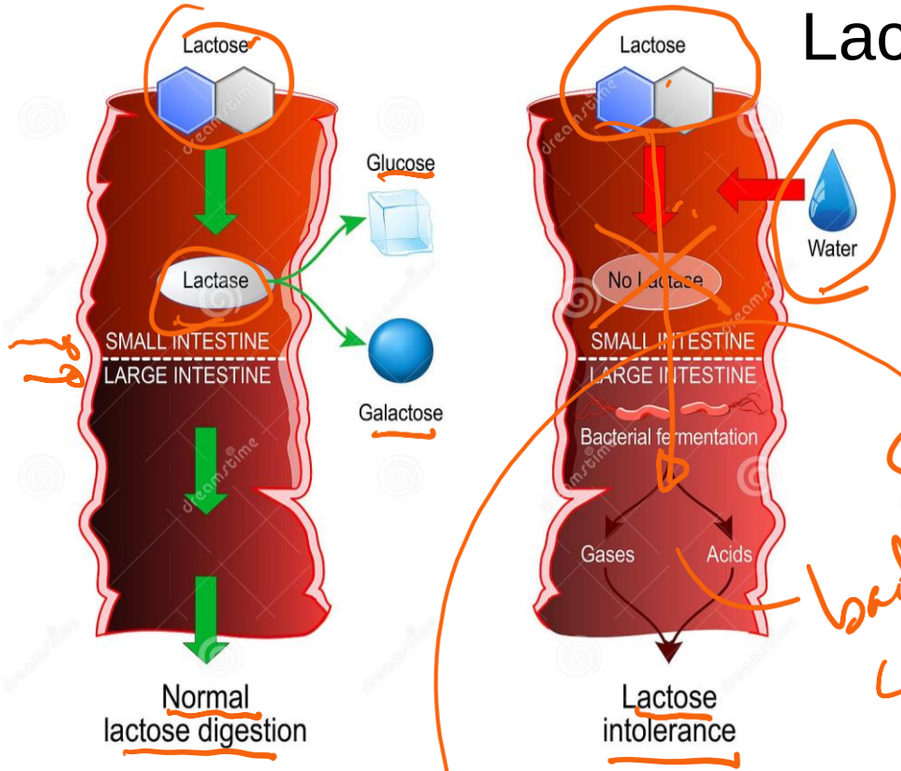


Lactose is the major sugar in mammalian milk.

- Infants produce the enzyme ***lactase*** to hydrolyze the disaccharide to monosaccharides.
- Lactase expression is turned off in some adults, depending on their genetic background.



Lactose Intolerance

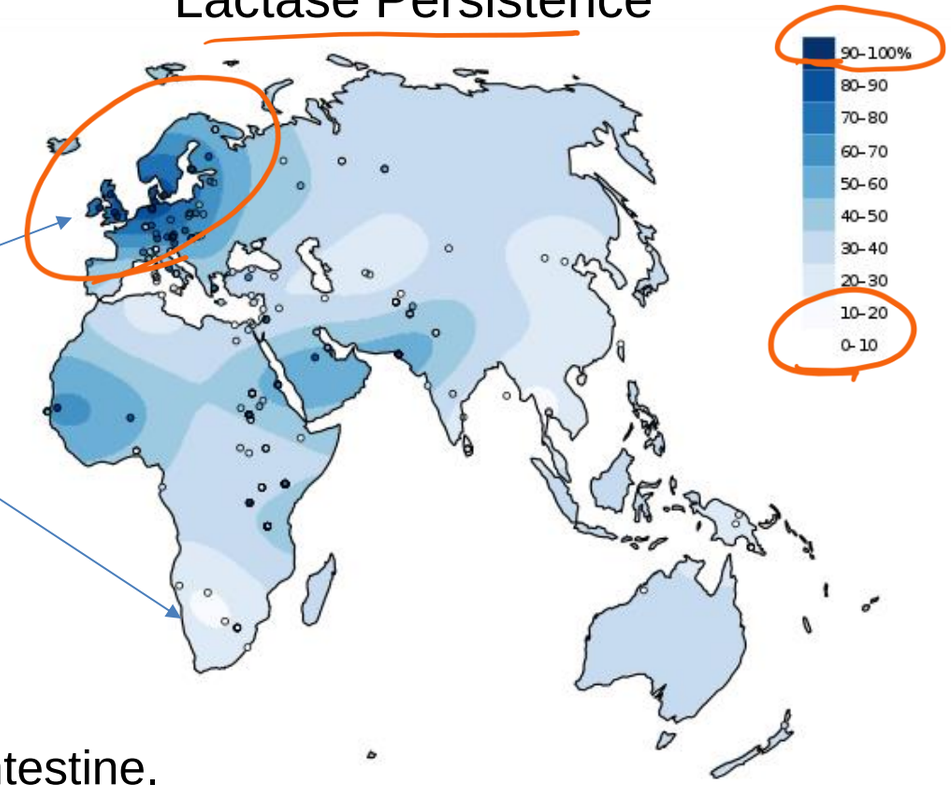


In which region do most of the people still produce lactase as adults?

A) UK

B) South Africa

Lactase Persistence



In an infant (lactase +):

- lactose is broken down to glucose and galactose in the small intestine.
- The two sugars are absorbed and used for energy

In a lactose intolerant individual (lactase -)

- The lactose is not absorbed in the small intestine, but instead draws water into the intestine due to osmosis – leading to bloating and diarrhea.
- Lactose enters the large intestine where gut bacteria use it as a carbon source, generating gas.

Lactose Intolerance

What to do if you are lactose intolerant:

A. Consume less lactose

Most individuals with lactose maldigestion can tolerate up to 12g of lactose as a single dose with no, or minor, symptoms
The European Food Safety Authority (EFSA)

Whole/low-fat Milk
200ml
9-10g lactose

Cheddar Cheese
25g
0.03g lactose

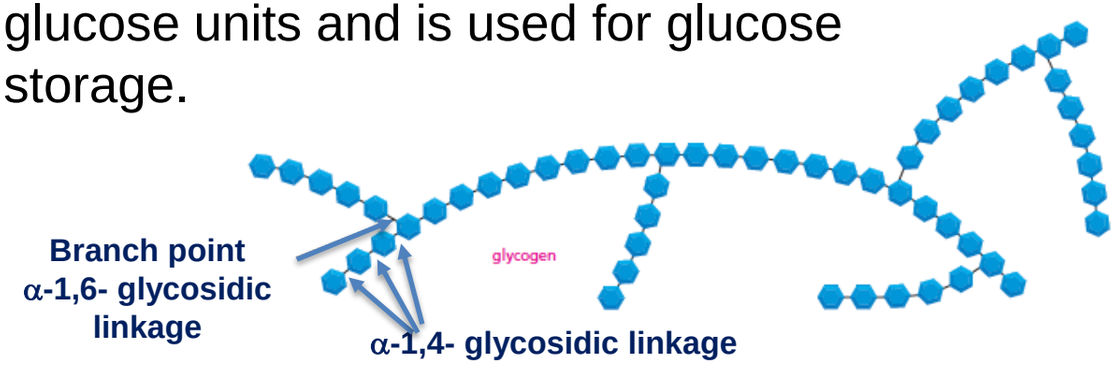
Whole Milk Yogurt
125g
6g lactose

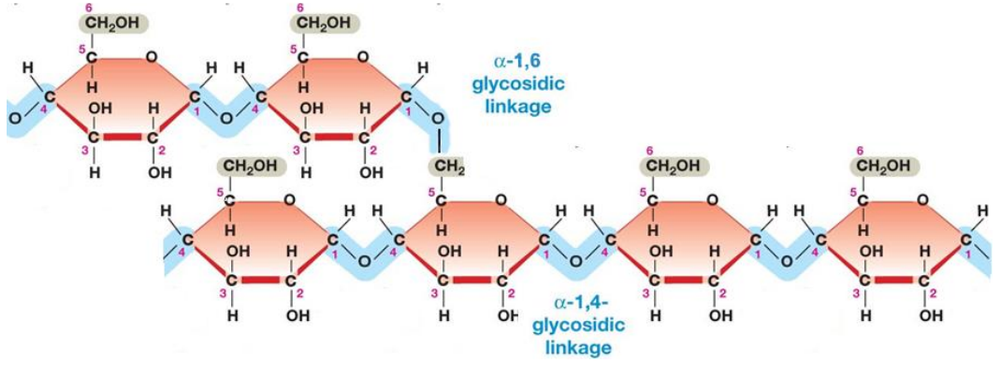
Lactose-free Milk
200ml
0g lactose

B. Hydrolyze the lactose to glucose and galactose before consumption.

Polysaccharides as Energy Storage – Glycogen Storage Disease

Glycogen and is made entirely of glucose units and is used for glucose storage.



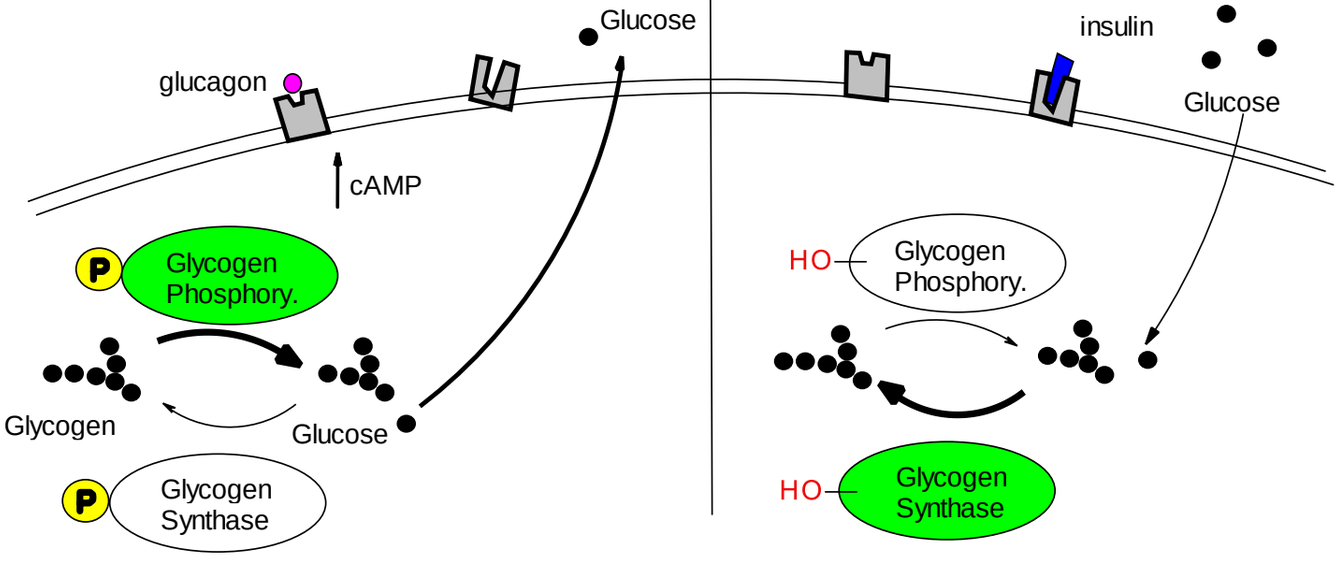


Glycogen Levels are regulated by hormones secreted due to blood glucose levels.

- Glucagon – low blood sugar
- Insulin – high blood sugar

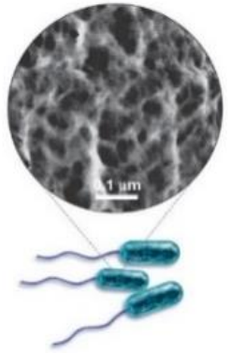
Two enzymes degrade or synthesize glycogen

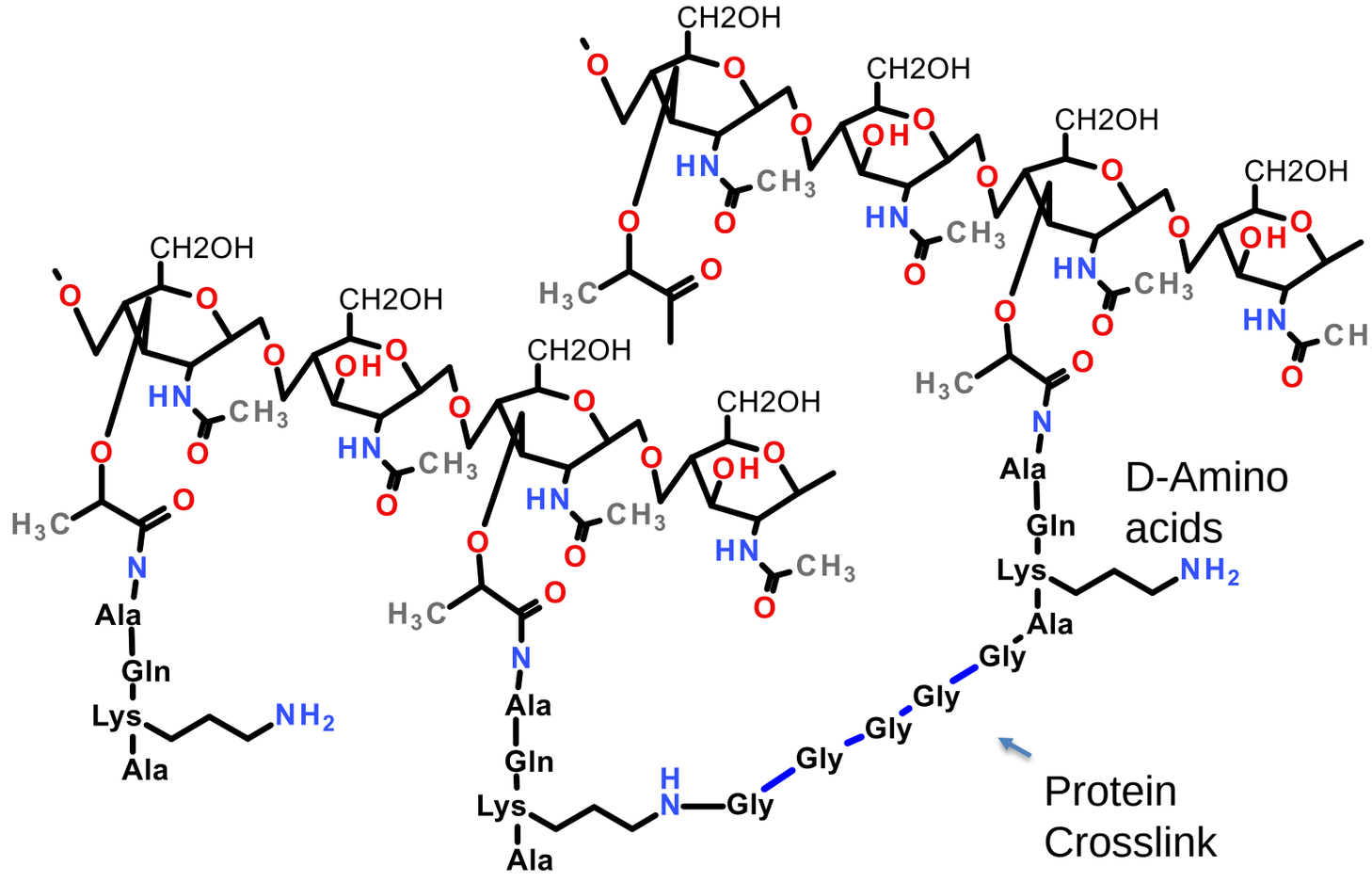
- Glycogen phosphorylase – releases glucose from glycogen
- Glycogen synthase – stores glucose in glycogen



Polysaccharides as Structural Molecules

Peptidoglycan (protein + sugar) in bacterial cell wall





Peptidoglycan (Bacterial Cell Wall)

Many antibiotics interfere with cell wall synthesis (e.g. penicillin)

Summary and Expectations for Carbohydrates

Key Points:

- General structure of monosaccharides - be able to distinguish between aldose and ketose (and identify compounds that are not sugars).
- Know how to number carbons on aldoses and ketoses
- Be able to describe the linkage between two monosaccharides (configuration at the anomeric carbon, atoms linked)
- Treatments for lactose intolerance
- Be able to describe the linkage between glucose molecules in:
  - Glycogen (glucose storage)
- Be able to describe the overall structure of the peptidoglycan in bacterial cell walls.