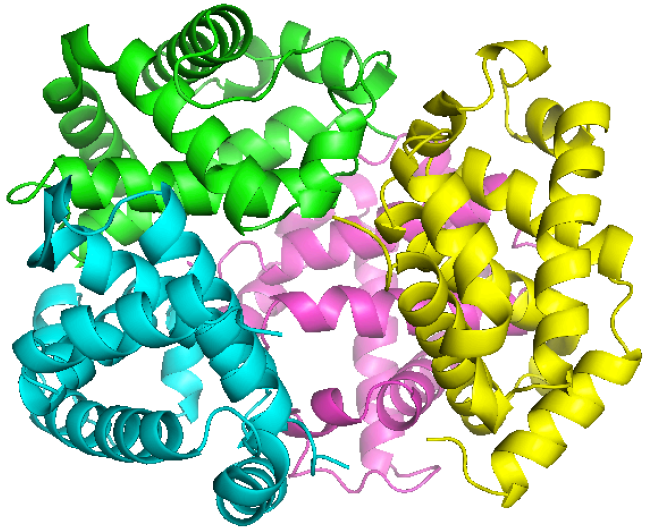
An Introduction to the Chemistry of Biomolecules & Mechanisms in Biological Chemistry



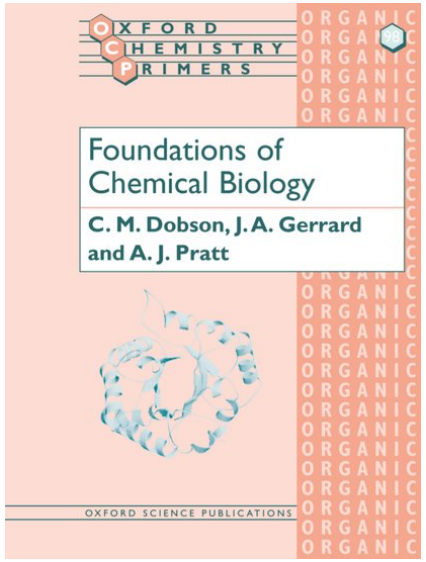
Course aims:

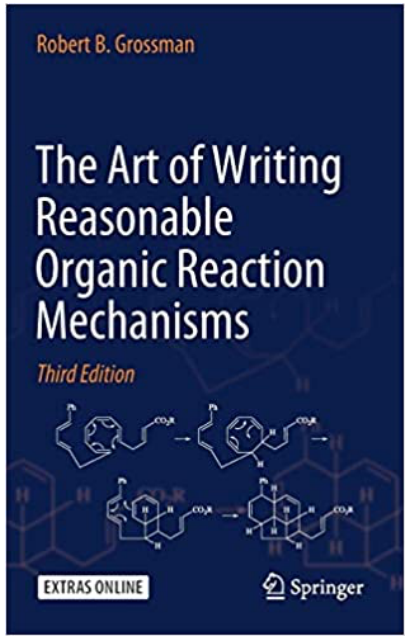
- To be familiar with the structures of important classes of biomolecules (carbohydrates, nucleic acids, proteins, lipids).
- To understand the importance of supramolecular interactions in biological chemistry (Van der Waals, hydrogen bonding, hydrophobic).
- To be able to draw reasonable mechanisms for chemical reactions in nature.
- To understand how enzymes catalyse chemical reactions.
- To be familiar with synthetic approaches to proteins and understand the biosynthesis of proteins.
- To understand the basics of drug action *via* case studies.

Recommended books:

Foundations of Chemical Biology by Dobson, Gerrard and Pratt, published by Oxford University Press (primer) ISBN: 9780199248995

The Art of Writing Reasonable Organic Reaction Mechanisms by Grossman, published by Springer ISBN: 9783030287320





Reaction Mechanisms – How to Use Curly Arrows

Mechanisms in biological chemistry

- Life runs on chemistry.
- There is only one chemistry. The chemical principles that frame your understanding of reactions carried out in the chemistry lab apply to reactions in nature. **There is no magic in biological chemistry!**
- All the mechanisms that you have studied previously will help you draw mechanisms for biological reactions.

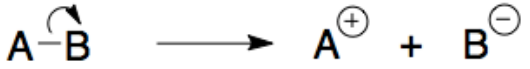
Mechanisms in biological chemistry

- Grossman, Professor of chemistry at the University of Kentucky, defines the structures we draw to describe elements and molecules as the vocabulary of chemistry and reaction mechanisms as the stories we tell with this vocabulary.
- Mechanisms are stories which describe the process of a chemical reaction.
- Reaction mechanisms give step-by-step pictures of how reactants are converted into products.
- A reaction mechanism shows the position of attack by a reagent on a reactant, together with an indication of which bonds break and which bonds form.

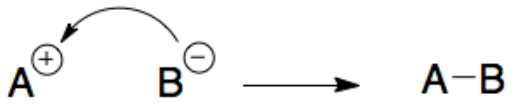
Mechanisms in biological chemistry

The conversion of reactants into products involves the breaking and making of bonds (the pairs of electrons in bonds move to different atoms). Chemists show this using **curly arrows**.

Heterolysis (Heterolytic cleavage)



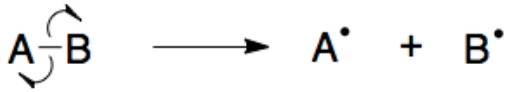
Heterolytic bond formation



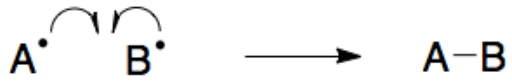
Mechanisms in biological chemistry

Organic molecules also undergo radical reactions (a radical is an atom or molecule containing a single electron). Single headed curly arrows show the movement of a single electron.

Homolysis (Homolytic cleavage)

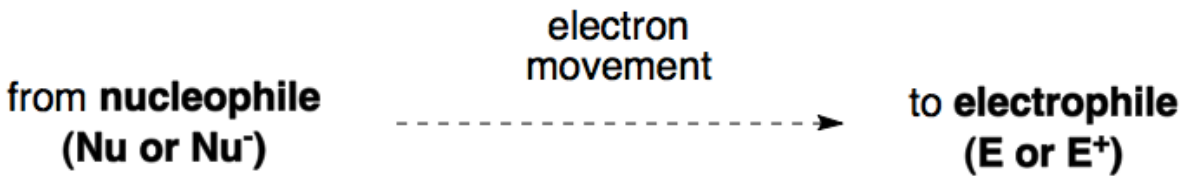


Homolytic bond formation



Mechanisms in biological chemistry

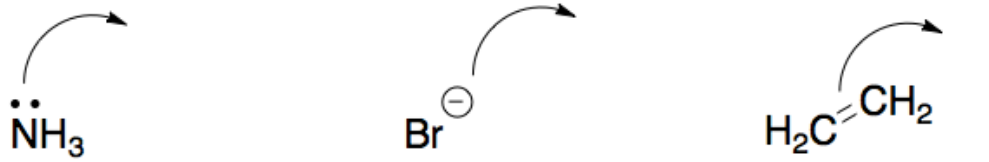
A bond forms when electrons move from a nucleophile to an electrophile.



- The **nucleophile** donates electrons
- The **electrophile** accepts electrons

Mechanisms in biological chemistry

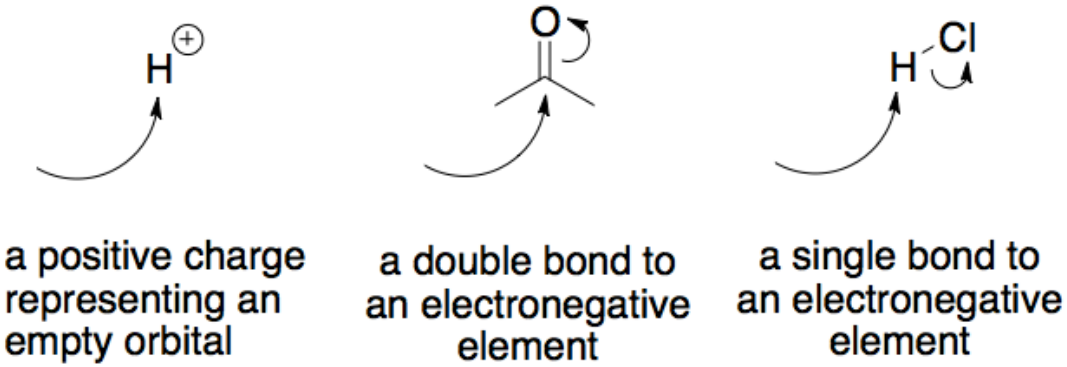
Nucleophiles donate electrons from **available high energy orbitals**. You will meet the following nucleophiles:



a lone pair a negative charge a double bond

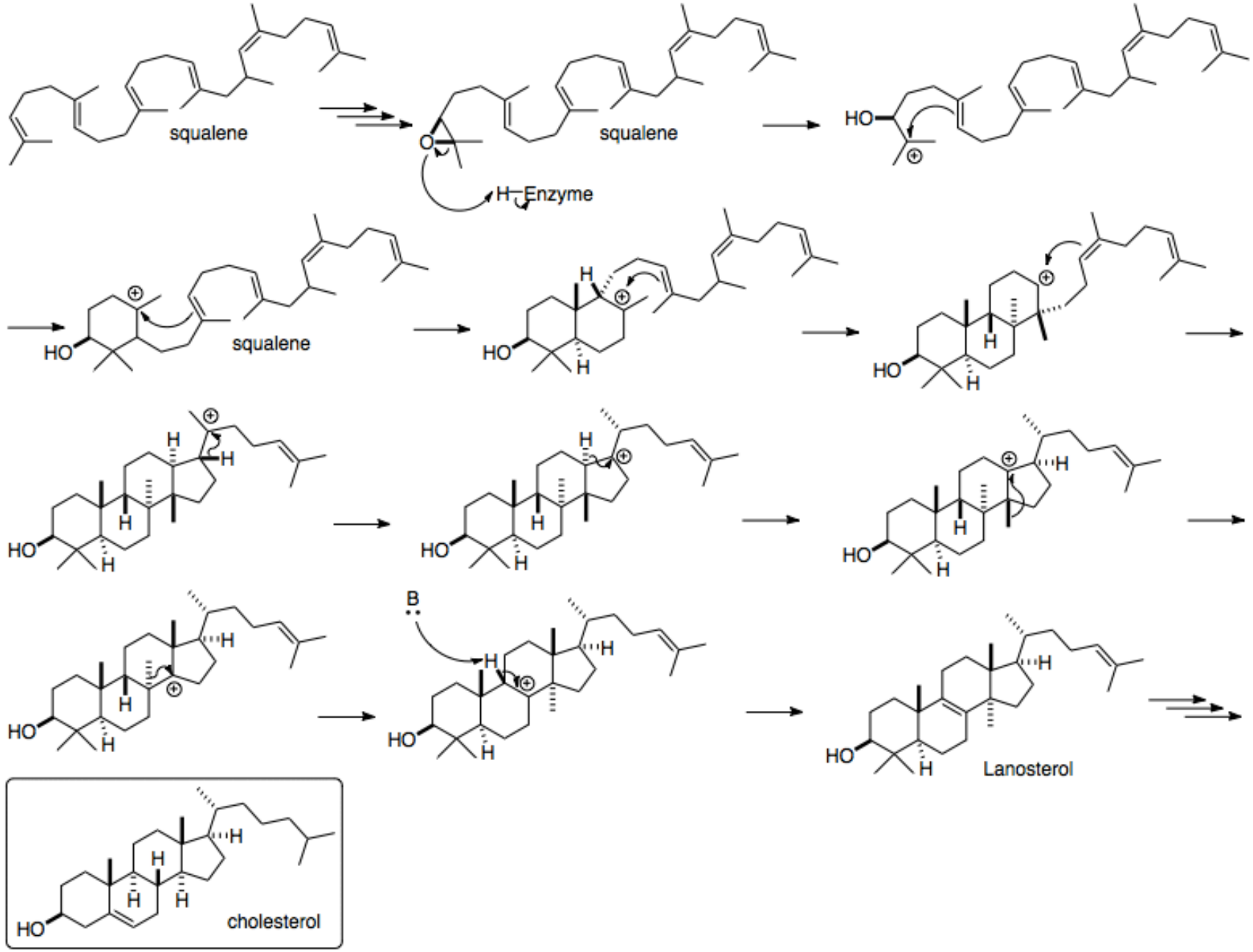
Mechanisms in biological chemistry

Electrophiles accept electrons into **empty low-energy orbitals**. You will meet the following electrophiles:



Mechanisms in biological chemistry

- Cholesterol has many important biological functions. It is the precursor in the biosynthesis of most other steroids including the sex hormones.
- The mechanism for the biosynthesis of cholesterol tells the story of how squalene is converted to cholesterol.
- The mechanism shows pi bonds acting as nucleophiles, carbocations acting as electrophiles, bases removing acidic protons etc. **There is no magic in biological chemistry!**


squalene
squalene
H—Enzyme
HO
squalene
HO
HO
HO
HO
HO
HO
B
HO
HO
Lanosterol
HO
cholesterol

Mechanisms in biological chemistry

- Chemical reactions are often classified by mechanism and you should review the mechanisms covered in L1.
- Commonly encountered reaction mechanisms in biological chemistry include:

1. Nucleophilic addition to the carbonyl group.
2. Nucleophilic substitution at the carbonyl group.
3. Nucleophilic substitution at the carbonyl group with loss of the carbonyl oxygen (e.g. acetal formation, imine formation).
4. Formation and reaction of enols (e.g. aldol reactions, Claisen ester condensations).

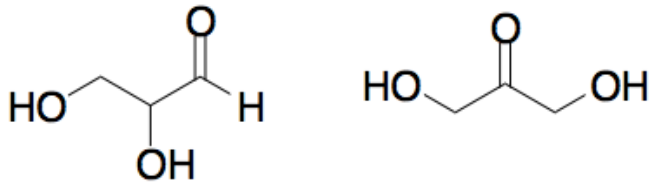
Carbohydrates

Carbohydrates

- Carbohydrates (sugars) are found in the food we eat and they provide our bodies with energy to carry out their biochemical processes.
- Carbohydrates also have a structural role for living organisms, for example cellulose in the cell walls of plants.
- Nucleotides, the subunits of DNA, are derivatives of carbohydrates.

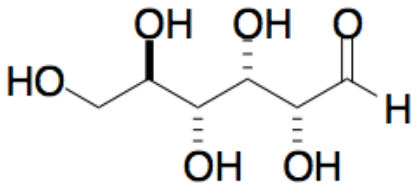
Monosaccharides

- Simple sugars are called **monosaccharides**.
- A monosaccharide is an aldehyde or ketone containing at least two hydroxyl groups.
- Aldehydic monosaccharides are classified as **aldoses** and those with a ketone function are called **ketoses**.
- The two simplest monosaccharides are 2,3-dihydroxypropanal (glyceraldehyde) and 1,3-dihydroxyacetone.

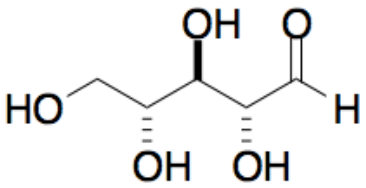


Monosaccharides

On the basis of chain length, we call monosaccharides **trioses** (three carbons), **tetroses** (four carbons), **pentoses** (five carbons), **hexoses** (six carbons), etc.



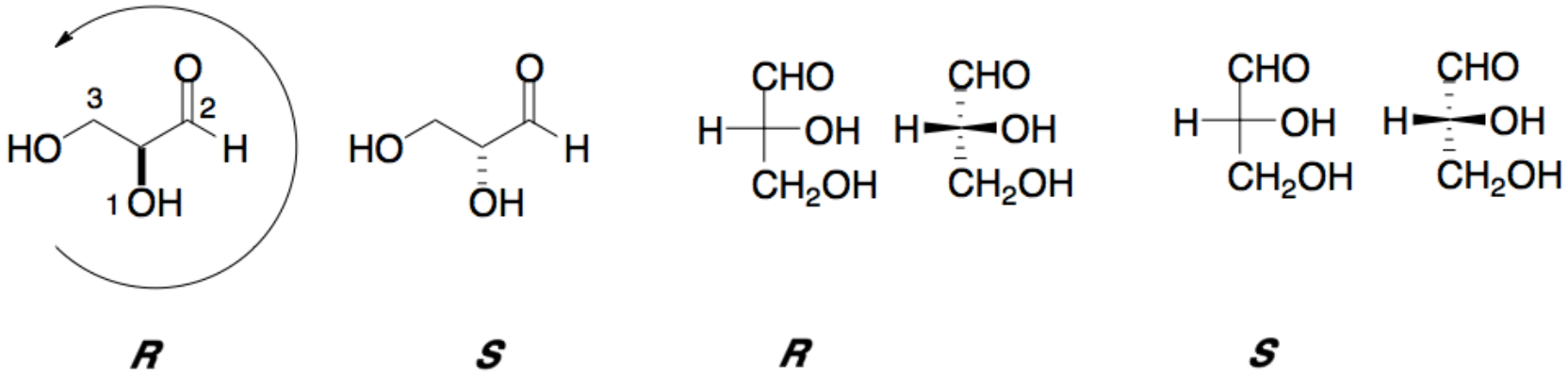
D-Glucose is a hexose



D-Ribose is a pentose

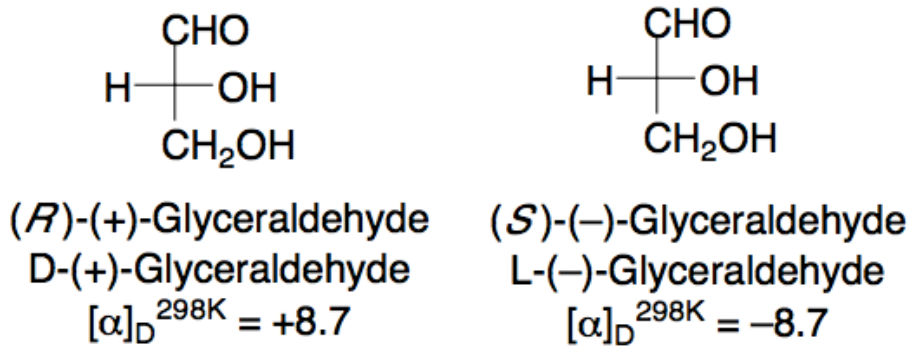
Stereochemistry - *R* & *S* nomenclature

- The simplest chiral monosaccharide is glyceraldehyde, with one stereogenic centre.
- Fischer projections are commonly used to indicate the configuration at each stereogenic center of a monosaccharide.



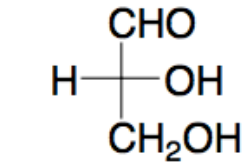
Stereochemistry - D & L nomenclature

- The D and L nomenclature relates all monosaccharides to glyceraldehyde.
- Instead of *R* and *S*, it uses the prefixes D for the (+) enantiomer of glyceraldehyde and L for the (-) enantiomer.

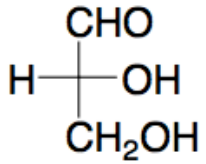


Stereochemistry - D & L nomenclature

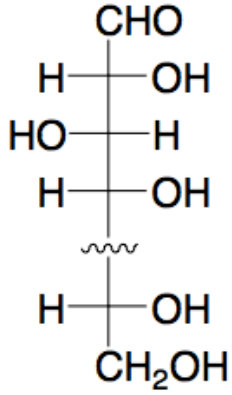
- Monosaccharides whose highest-numbered stereocenter (the stereocentre furthest from the aldehyde or ketone group) have the same absolute configuration as that of D-(+)-glyceraldehyde are labelled D and those with the opposite configuration at that stereocentre are labelled L.
- Why use the D, L nomenclature? Almost all naturally occurring sugars have the D-configuration.



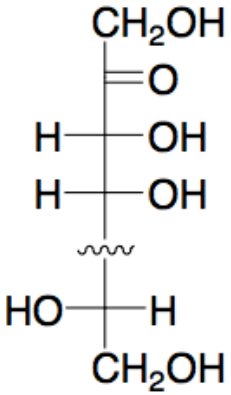
D-Glyceraldehyde



L-Glyceraldehyde



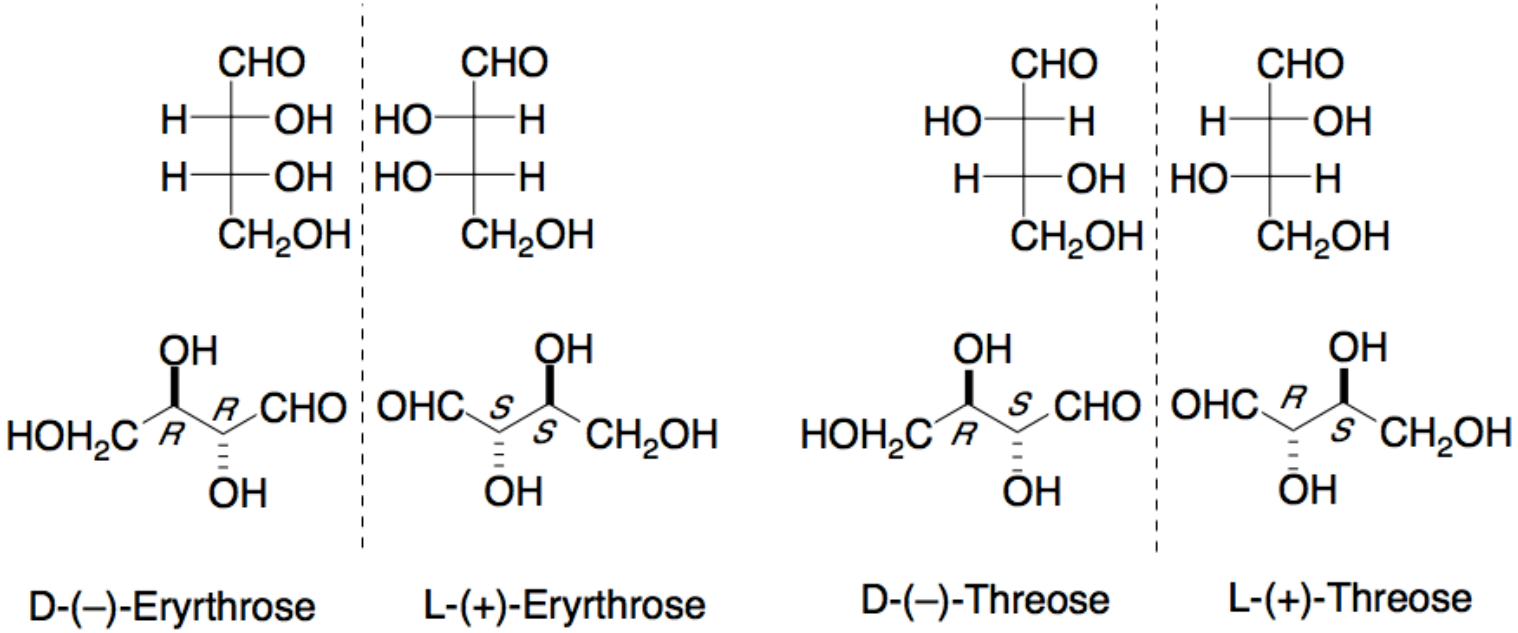
D-Aldose



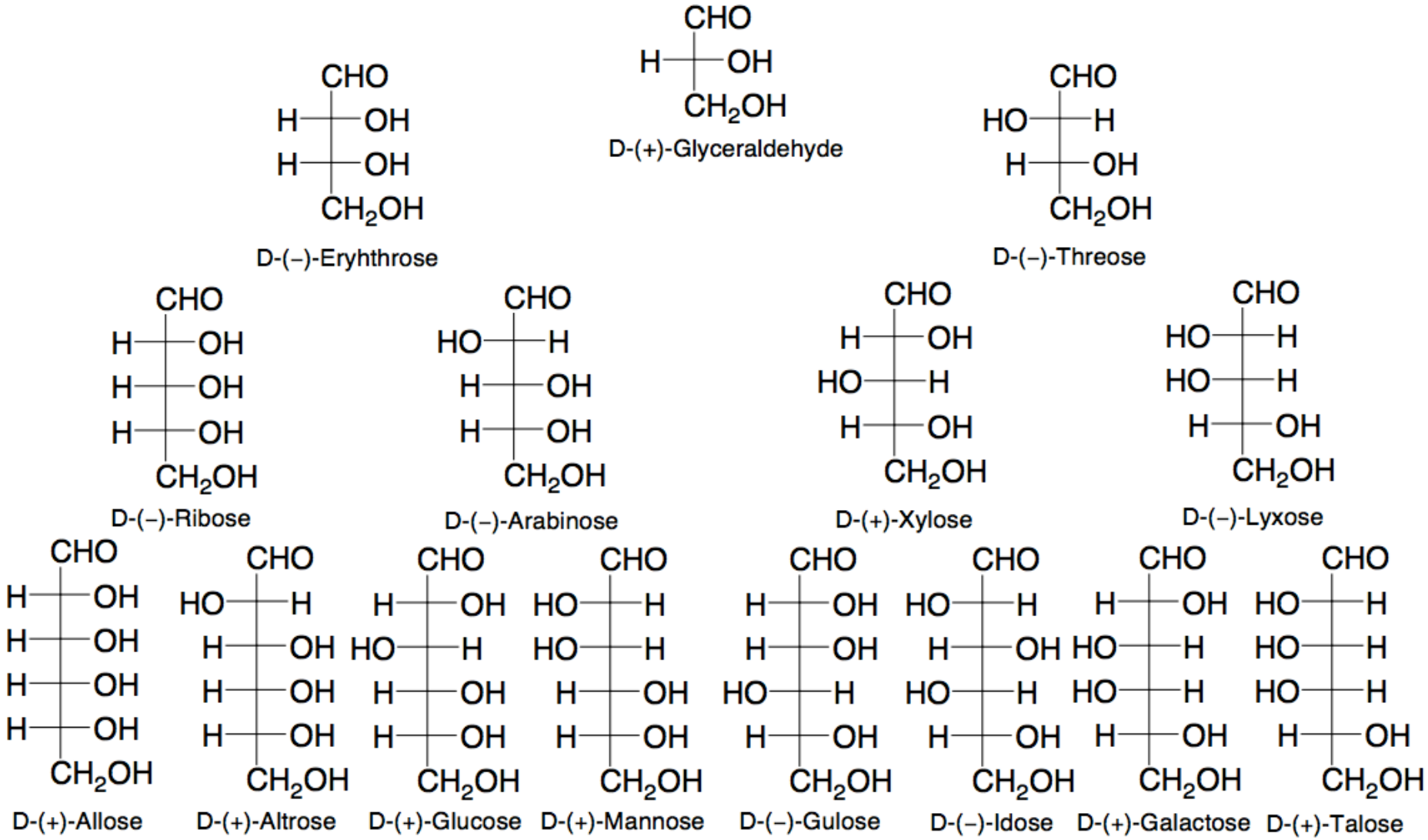
L-Ketose

Stereochemistry - Diastereomers

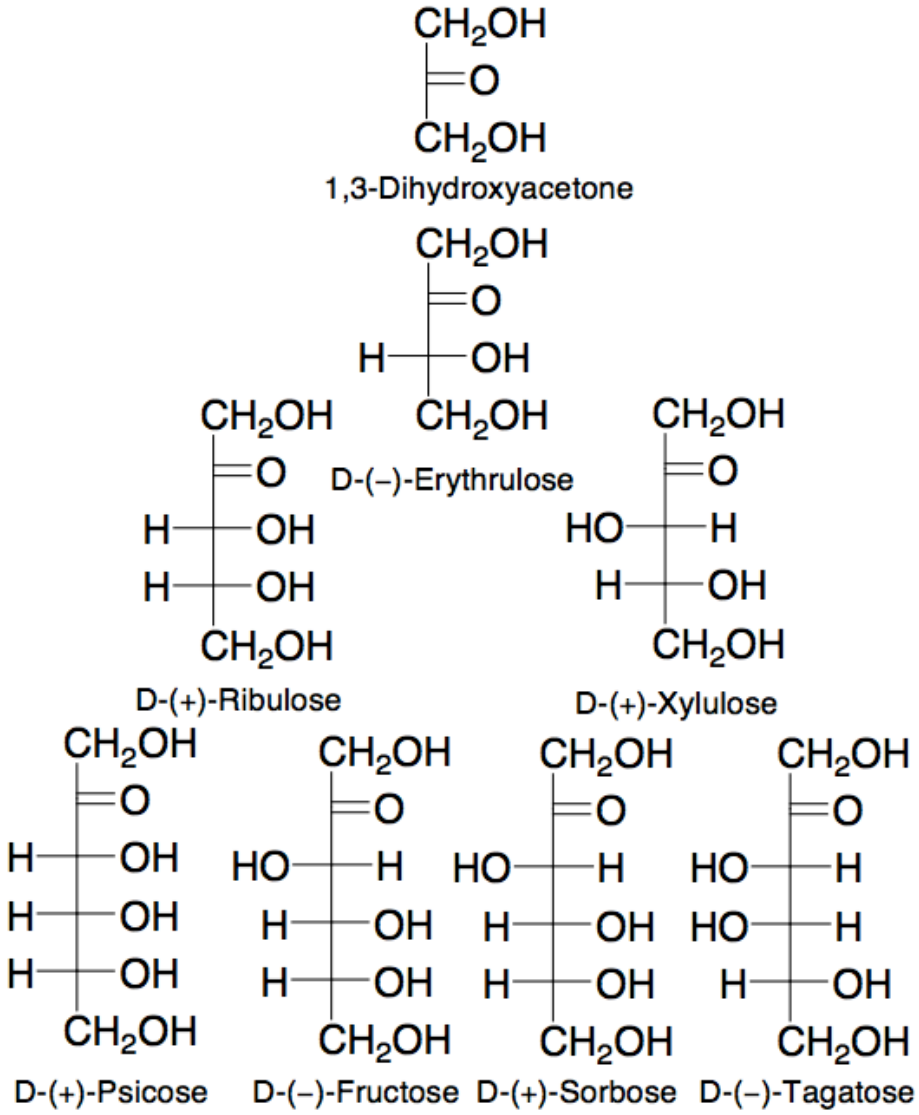
- As the number of stereocentres increase, the number of stereoisomers increase.
- The aldotetrose 2,3,4-trihydroxybutanal exists as four stereoisomers, two diastereomers, each as a pair of enantiomers.



Stereochemistry – D-Aldoses (up to the aldohexoses)

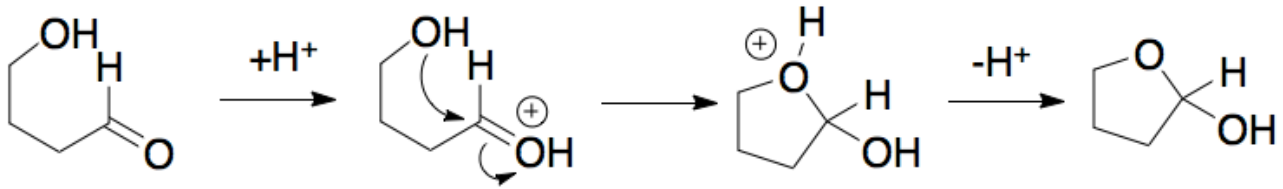


Stereochemistry – D-Ketoses (up to the ketohexoses)



Sugars form intramolecular hemiacetals

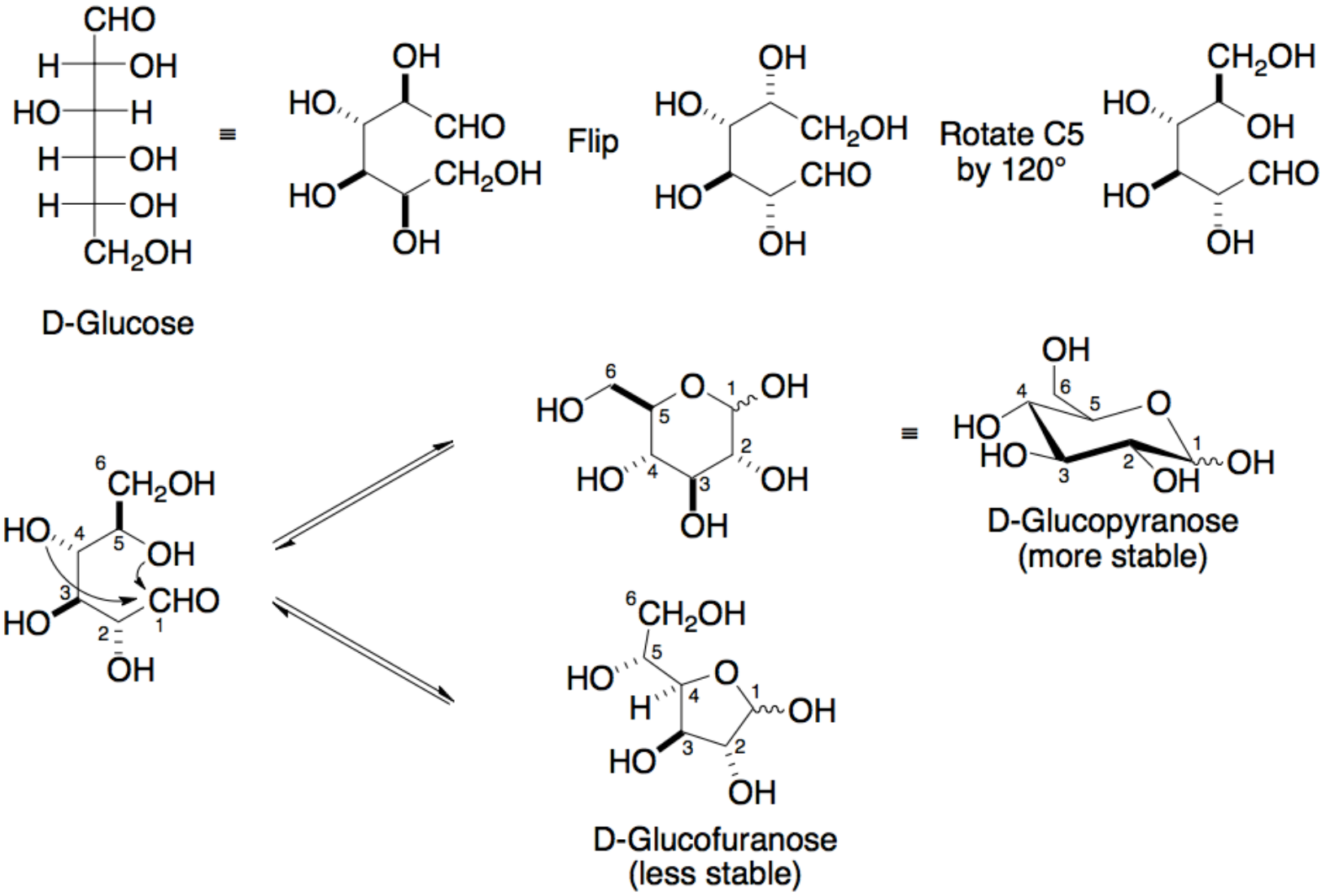
Monosaccharides are hydroxycarbonyl compounds and are capable of intramolecular hemiacetal formation.



Sugars form intramolecular hemiacetals

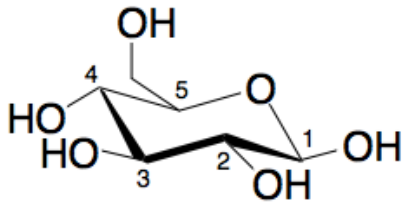
- The hexoses and pentoses exist as an equilibrium mixture with the cyclic hemiacetal in which the hemiacetal predominates.
- Three- and four-membered rings have high ring strain. Five- and six-membered rings are stable with six membered rings usually the most stable.
- A six-membered ring monosaccharide is called a **pyranose**.
- A five-membered ring monosaccharide is called a **furanose**.

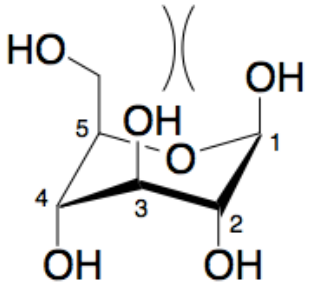
D-Glucose – Hemiacetal formation



The chair conformation – 1,3 diaxial interactions

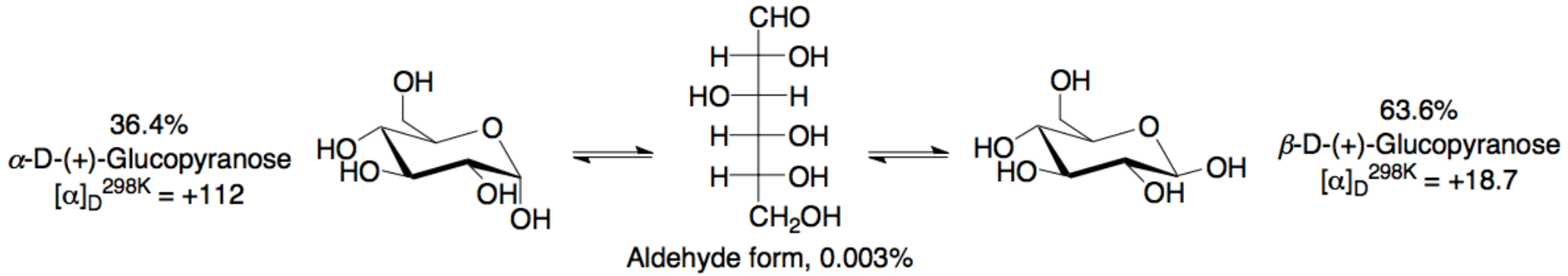
Most aldohexoses adopt the chair conformation with the bulky hydroxymethyl group at C5 in the equatorial position, thus minimising unfavourable 1,3 diaxial interactions.





The anomeric carbon

- Upon cyclisation, the carbonyl carbon turns into a new stereocentre.
- Hemiacetal formation leads to the formation of two diastereomers differing in the configuration of the hemiacetal group.
- For a D sugar, if the configuration is *S*, the diastereomer is α and if it's *R* the diastereomer is β.
- This type of diastereomer is unique to sugars and they have been given their own name, **anomers**. The new stereocentre is called the **anomeric centre**.



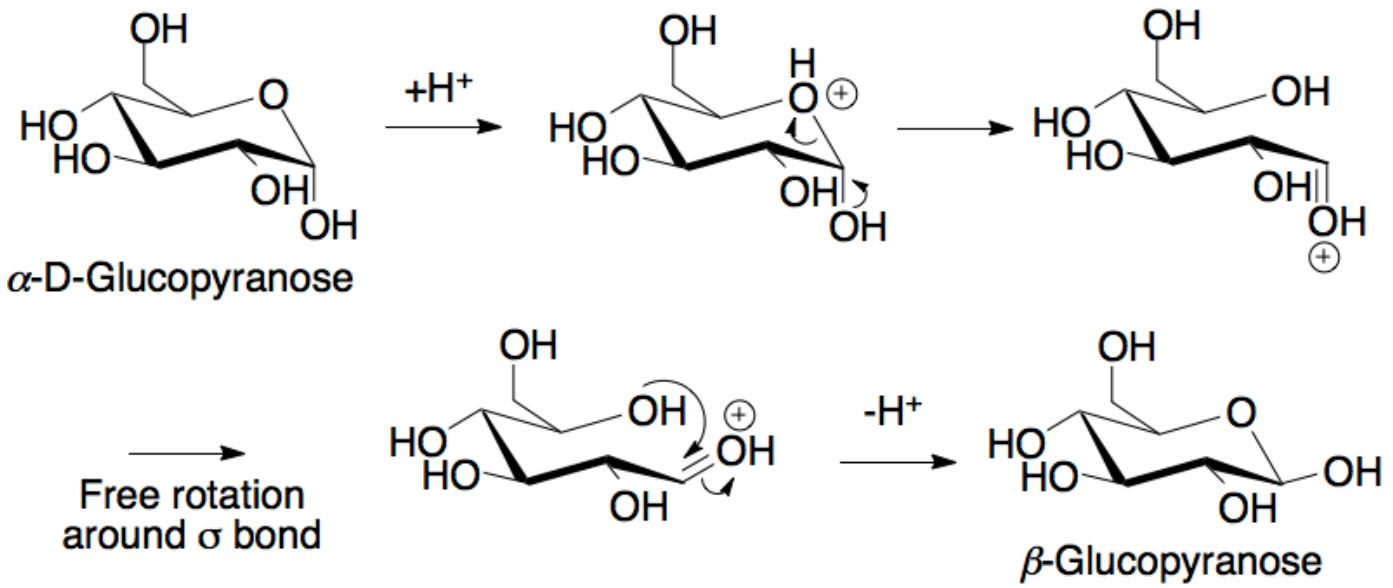
Mutarotation

- When pure α-D-(+)-glucopyranose is dissolved in water, its optical rotation, $[\alpha]_D^{298\ K}$ is +112. With time this decreases until it reaches a constant +52.7.

- **Mutarotation** is the change in optical rotation observed when pure α or β anomers equilibrate.

- Interconversion of α and β anomers is a general property of sugars capable of existing as cyclic hemiacetals.

Mutarotation - Mechanism

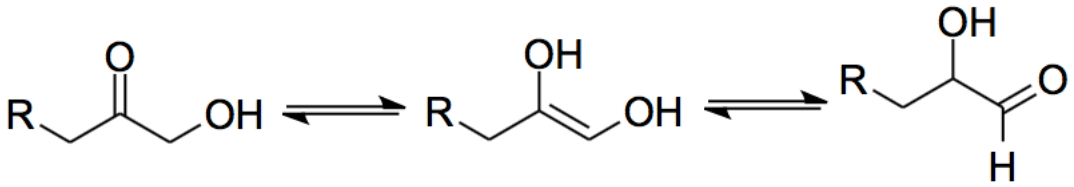


Reducing Sugars

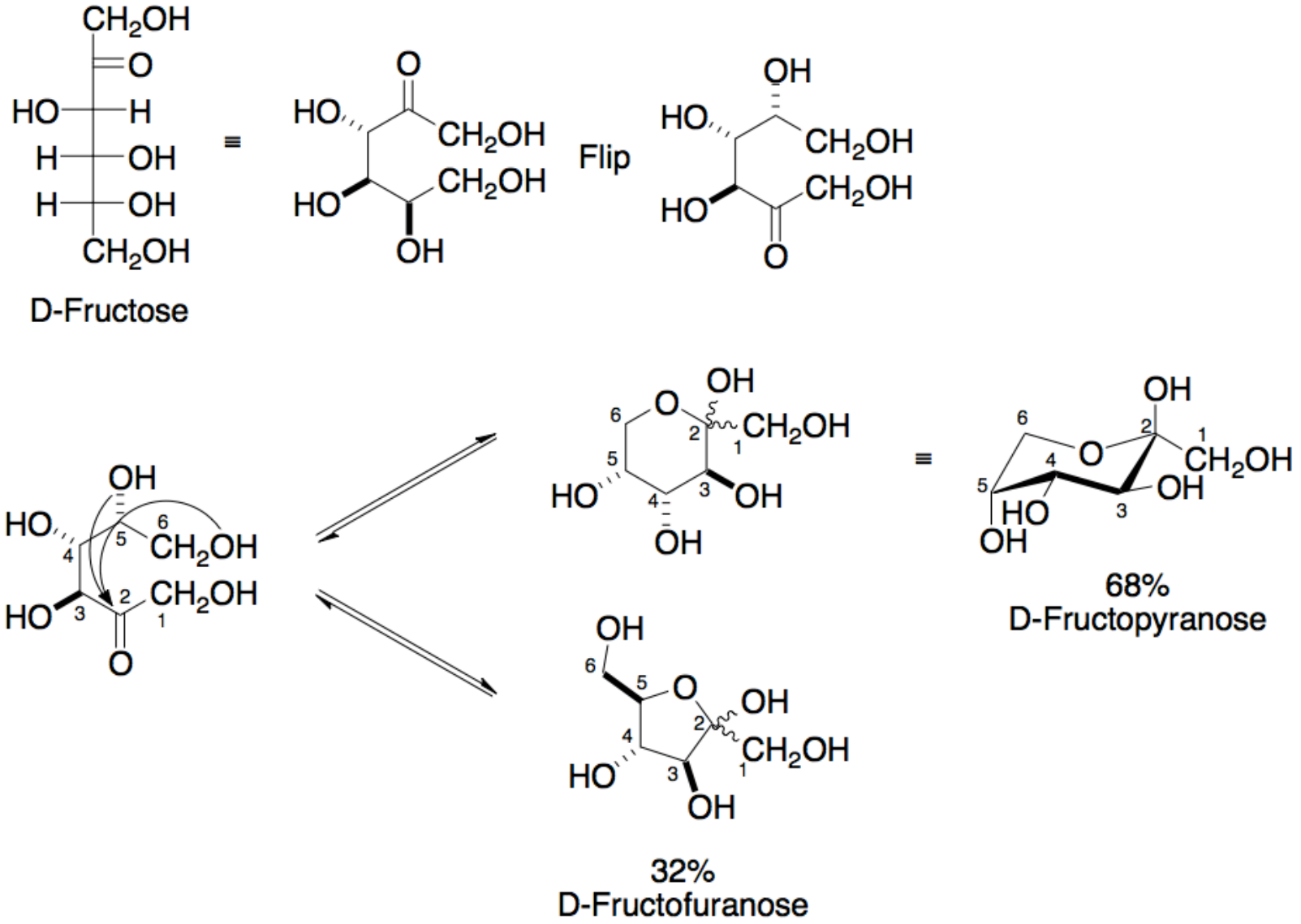
- We've seen that glucopyranose is in equilibrium with a ring opened form containing an aldehyde.
- Aldehydes can be oxidised to carboxylic acids (e.g with chromic acid).
- A visual test for the oxidation of aldehydes uses Fehling's solution where an aldehyde **reduces** the blue Cu(II) solution to red Cu(I) as Cu_2O.
- We saw that the concentration of glucose in the open-chain aldehyde form at any given time is small but is sufficiently long lived to react with the Fehling's reagent.
- We have a visual test for aldose monosaccharides.

Reducing Sugars

- Ketones are not oxidised by Fehling's reagent but ketoses are also reducing sugars and Fehling's solution can be used as a test for ketoses.
- Alpha-hydroxy ketones are in equilibrium with aldehydes via tautomerism and so ketose monosaccharides are also reducing sugars.

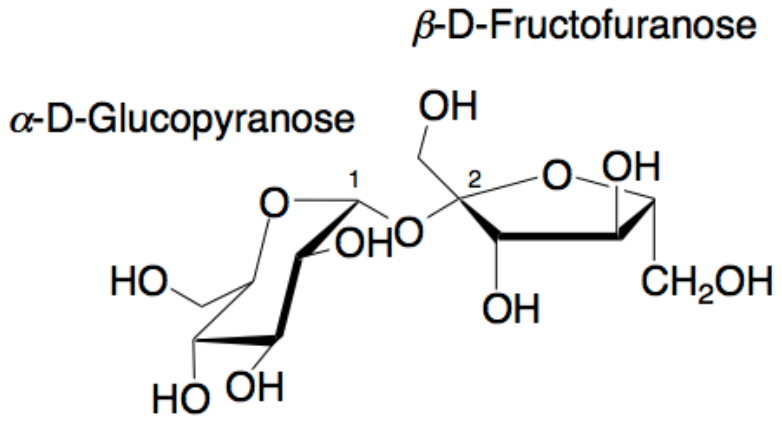


D-Fructose – Hemiacetal formation



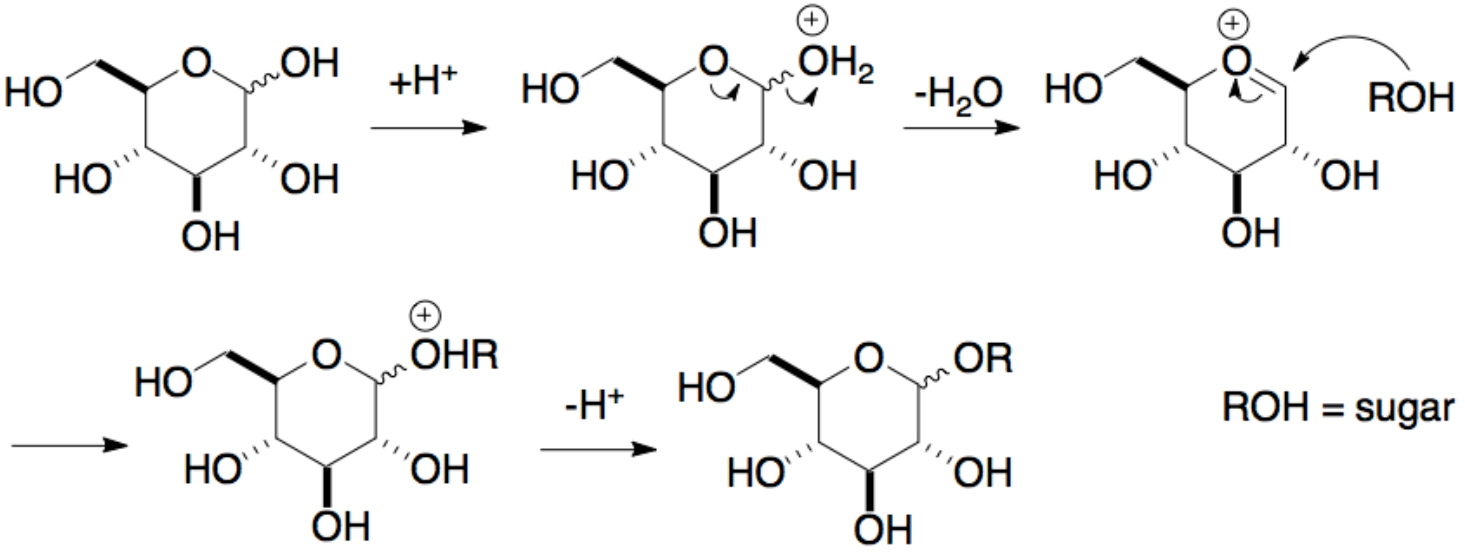
Complex sugars in nature - Sucrose

- A substantial fraction of natural sugars exist as dimers, trimers, oligomers (2-10 sugar units) and polymers.
- Table sugar, sucrose, is a dimer composed of two units, glucose and fructose.
- Sucrose contains an acetal linkage (glycosidic bond) between the anomeric carbons of the constituent sugars. Sucrose isn't a reducing sugar.



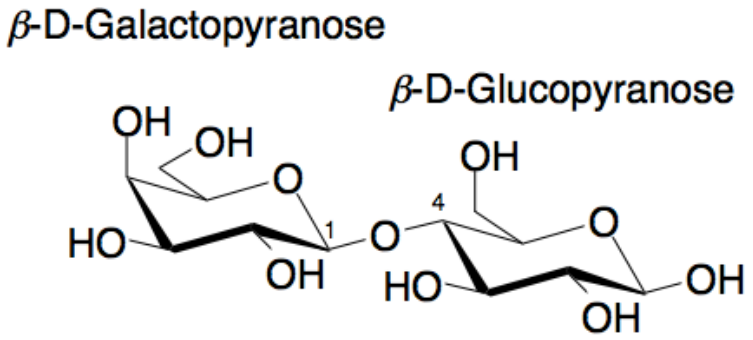
1α, 2β glycosidic bond

Complex sugars in nature – Acetal formation



Complex sugars in nature - Lactose

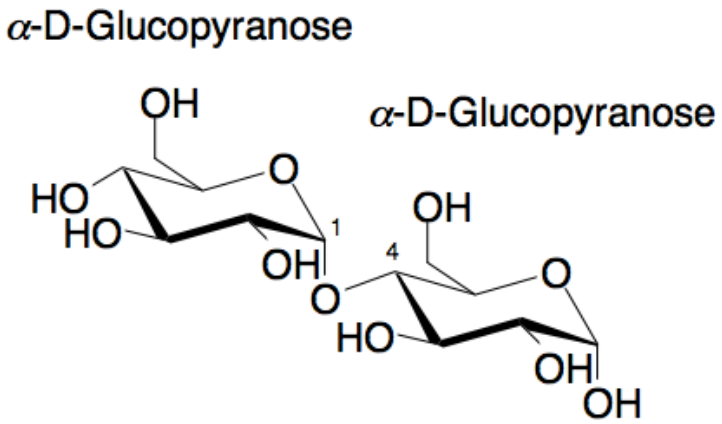
- The most abundant natural disaccharide after sucrose is **lactose** (milk sugar).
- Lactose contains an unprotected hemiacetal unit (capabale of opening and closing). As such it is a reducing sugar and undergoes mutarotation.
- The enzyme lactase, catalyses hydrolysis of the glycosidic bond yielding galactose and glucose. Some people are lactose intolerant because they do not produce enough lactase to break down lactose.



1β, 4 glycosidic bond

Complex sugars in nature - Maltose

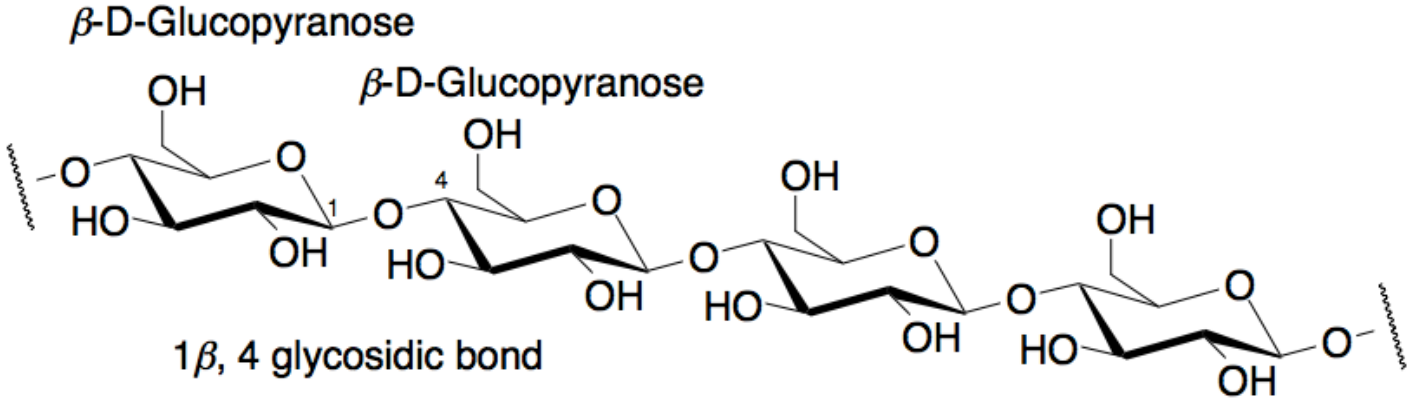
- Maltose (malt sugar) is a dimer of glucose obtained by enzymatic degradation of starch by amylase.
- The hemiacetal oxygen of one glucose is bound to C4 of another glucose.
- Maltose has a hemiacetal unit and as such is a reducing sugar and undergoes mutarotation.
- The hydrolysis of maltose to two molecules of glucose is catalysed by aqueous acid or maltase.



1α, 4 glycosidic bond

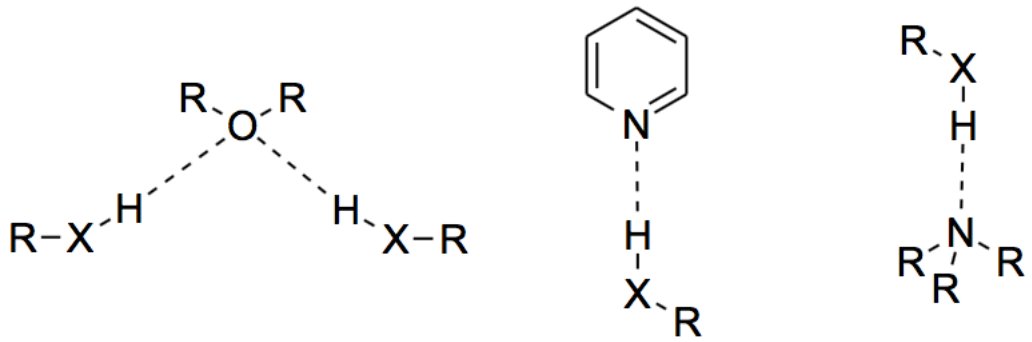
Complex sugars in nature - Cellulose

- Cellulose is a poly-β-glucopyranoside linked at C4, containing about 3000 monomeric units.
- Individual chains of cellulose tend to align and connect *via* multiple hydrogen bonds.
- The hydrogen bonds are responsible for the rigid structure of cellulose. It serves a structural role in cell walls in trees and plants.



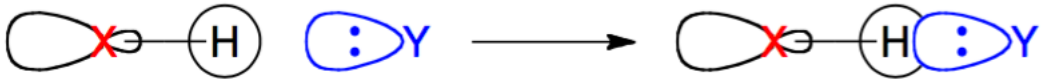
Hydrogen bonding

- Hydrogen bonding is a non-covalent interaction. It is very important in biological chemistry.
- Hydrogen bonding takes place between an electron rich heteroatom (***hydrogen bond acceptor***) and an electron deficient hydrogen atom (***hydrogen bond donor***).
- The hydrogen atom is usually covalently linked to an electronegative atom, such as O or N giving the hydrogen a slight positive charge.
- There is a significant electrostatic component to hydrogen bonding.



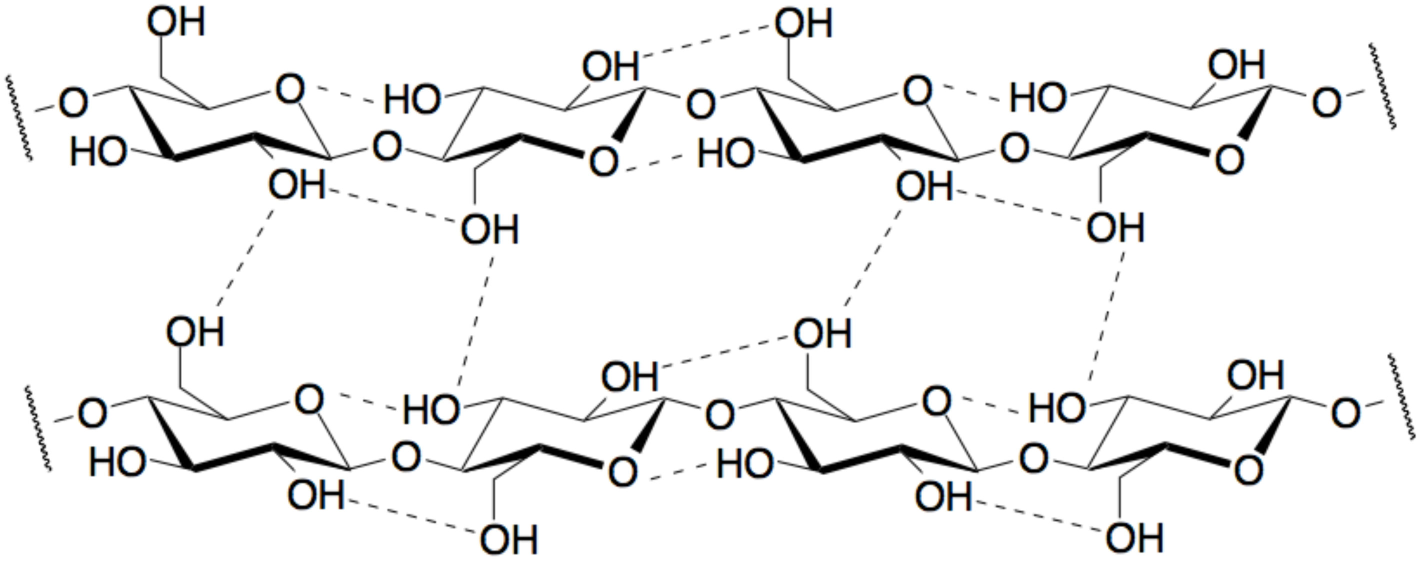
Hydrogen bonding

- Orbital interactions are an important component of hydrogen bonds.
- Hydrogen bonds can be viewed as having a σ-bonding component.
- Hydrogen bonds are directional and there is an optimum orientation for hydrogen bonding, where the X-H bond points directly to the lone pair on the heteroatom (180° bond angle).



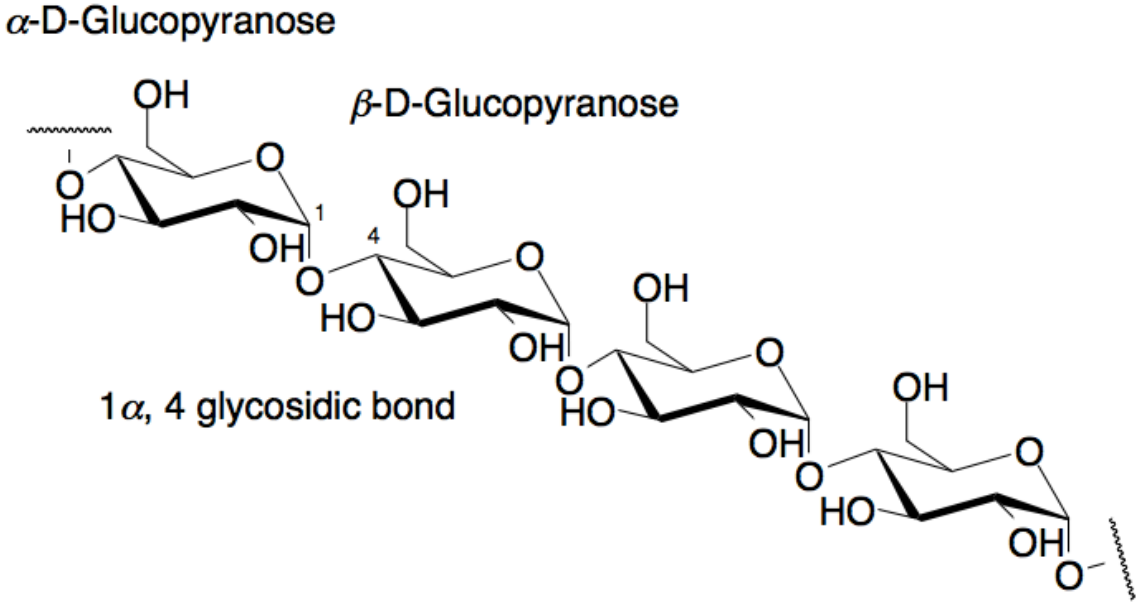
- Hydrogen bonds are typically 1.5-2.2 Å compared to 1.0-1.5 Å for covalent bonds.

Complex sugars in nature – Cellulose – Hydrogen bonding



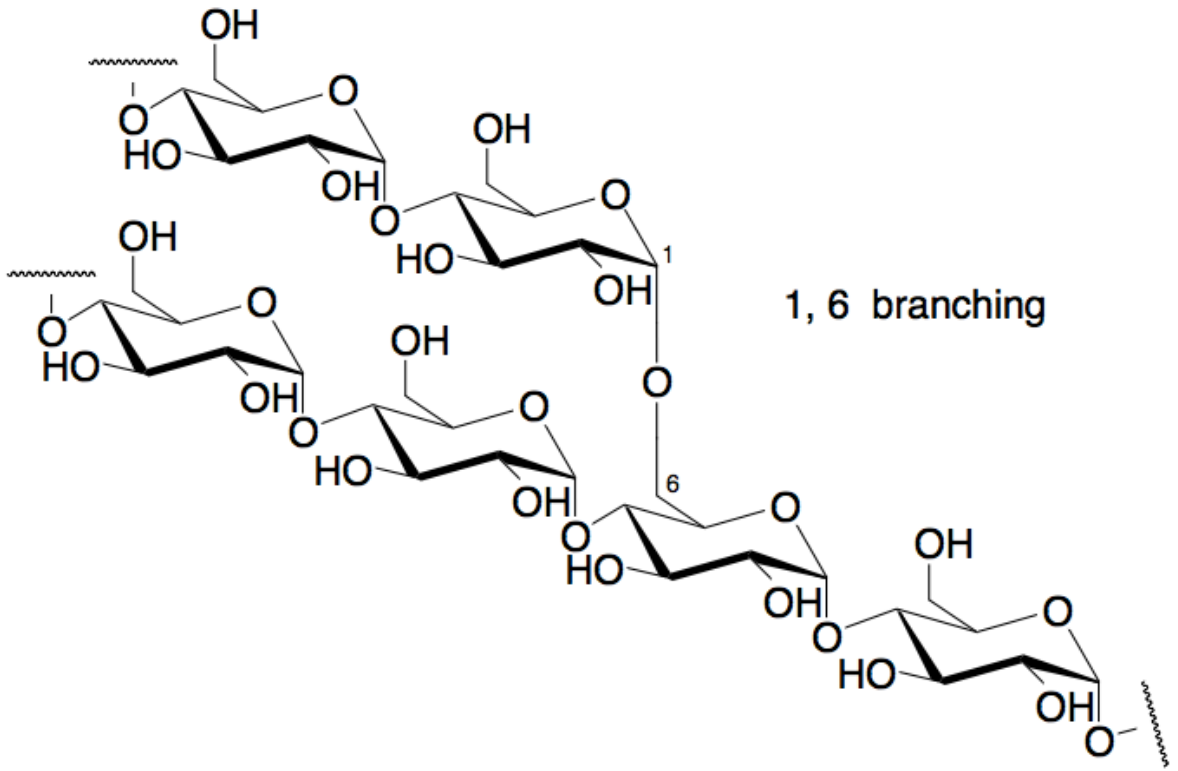
Complex sugars in nature – Starch - Amylose

- Starch is a poly-α-glucopyranoside made up of amylose (20%) and amylopectin (80%).
- It serves as a food reserve in plants. Major sources of starch are corn, potatoes, wheat and rice.
- Amylose contains a few hundred glucose units per molecule (Molecular weight = 150 000 – 600 000 g mol^{-1})



Complex sugars in nature – Starch - Amylopectin

- Amylopectin is branched at C6 approximately once every 20 – 25 glucose units.
- Molecular weight (g mol^{-1}) is variable but can be in the millions.

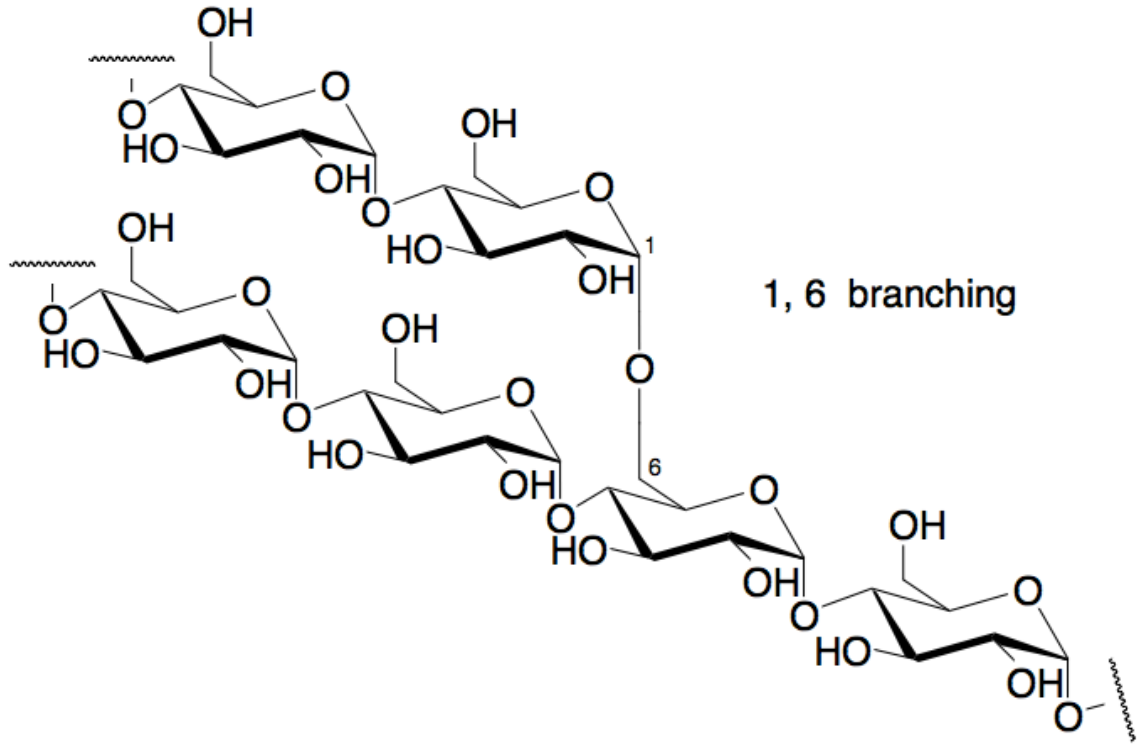


Complex sugars in nature – Starch

- The hydrolysis of starch is catalysed by highly selective glycosidase enzymes.
- Humans do not possess enzymes which catalyse the break down of cellulose which is why we can eat potatoes but not grass!
- Interestingly, cows cannot synthesise the enzymes necessary to digest cellulose either. However, their stomachs contain microorganisms that produce enzymes capable of catalysing the hydrolysis of cellulose. As such, cows can digest grass!

Complex sugars in nature - Glycogen

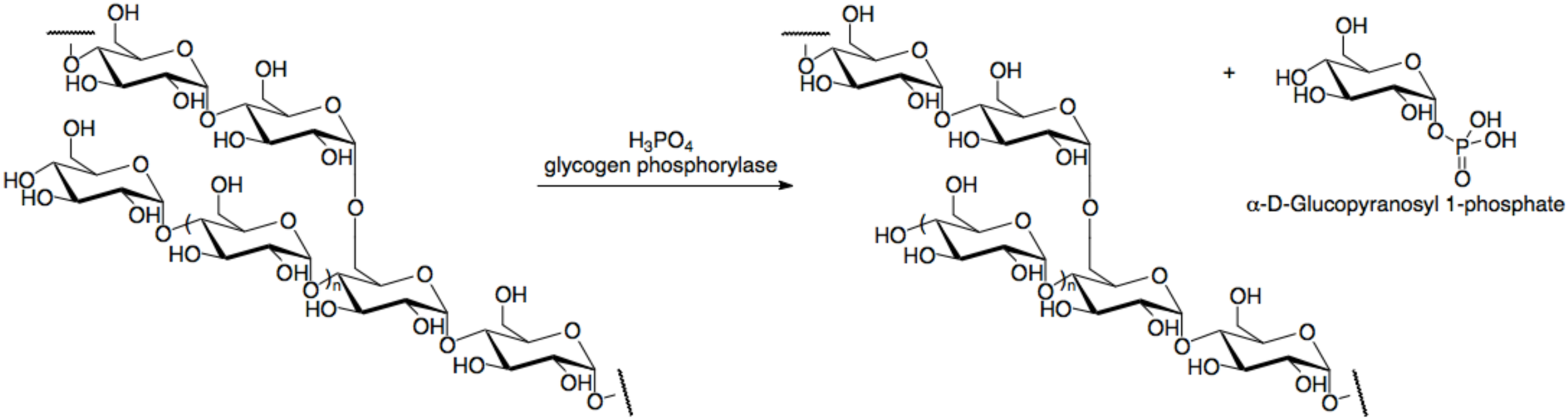
- Glycogen is a polysaccharide similar in structure to amylopectin.
- It has greater branching than amylopectin (approximately one branch every ten glucose units) and it is much bigger (molecular weight can reach 100 million g mol^{-1})



Complex sugars in nature - Glycogen

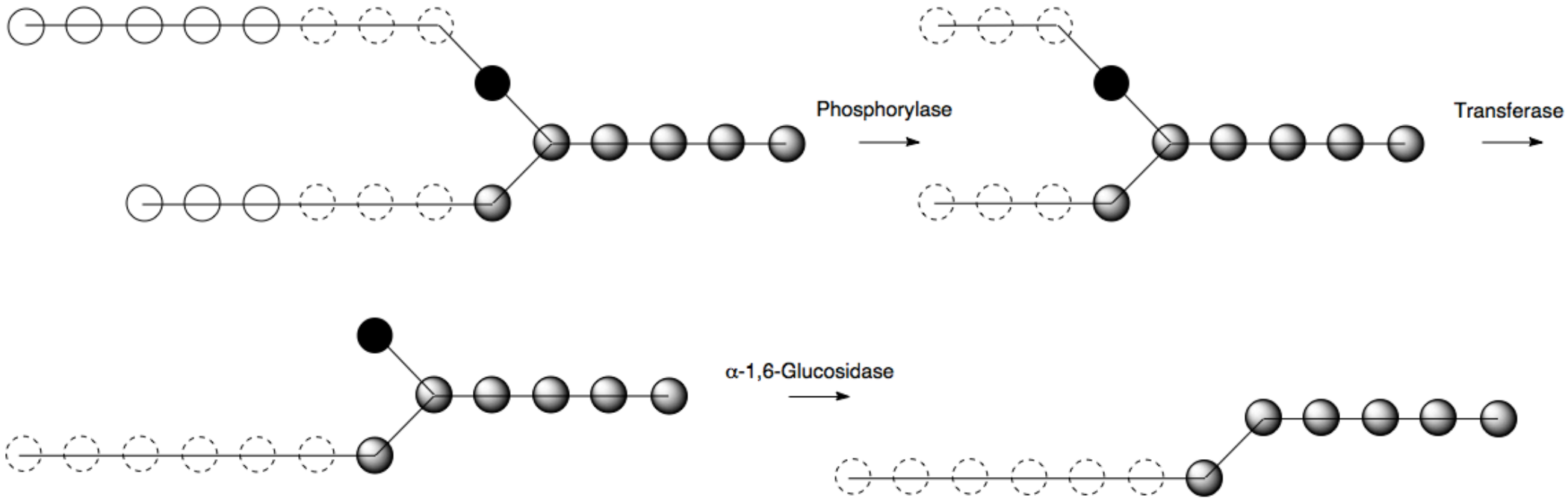
- Glycogen is one of the major energy storage polysaccharides in animals.
- It is stored in the liver and skeletal muscles and provides an immediate source of glucose for use during physical activity.
- It is broken down to α-D-glucopyranosyl 1-phosphate enzymatically by glycogen phosphorylase.

Complex sugars in nature - Glycogen



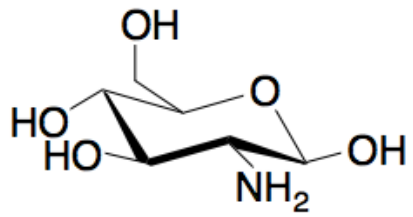
Complex sugars in nature - Glycogen

- Glycogen phosphorylase cannot break the α-1,6-glycosidic bonds.
- Two other enzymes, transferase and α-1,6-glucosidase, catalyse the degradation of glycogen at the branching points.



Amino sugars

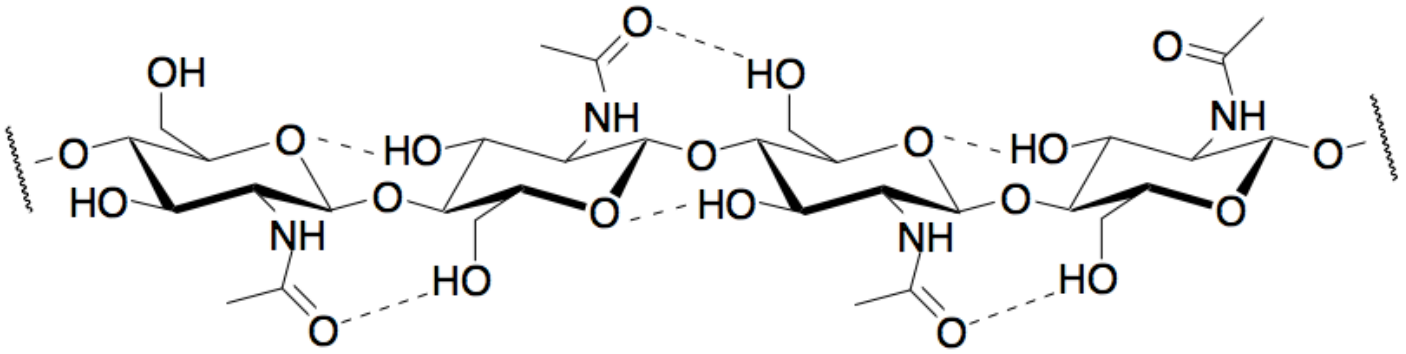
- Amino sugars are carbohydrate derivatives into which nitrogen is incorporated.
- An example is β-D-glucosamine, which is biosynthesized from D-glucose.



β-D-Glucosamine

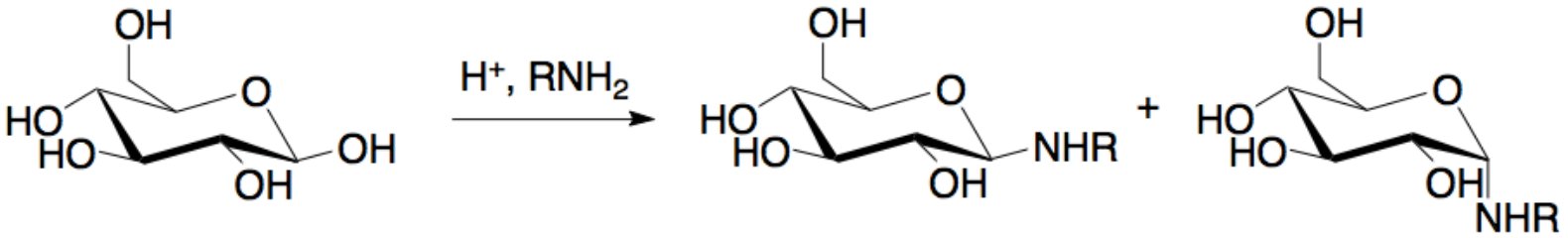
Amino sugars - Chitin

- The hard outer skeletons of insects and crustaceans contain chitin, a polymer similar to cellulose but made up of repeating acetyl glucosamine monomers instead of glucose.
- Like cellulose, there is significant intramolecular hydrogen bonding and intermolecular hydrogen bonding between neighboring strands, rendering the polymer extremely strong.



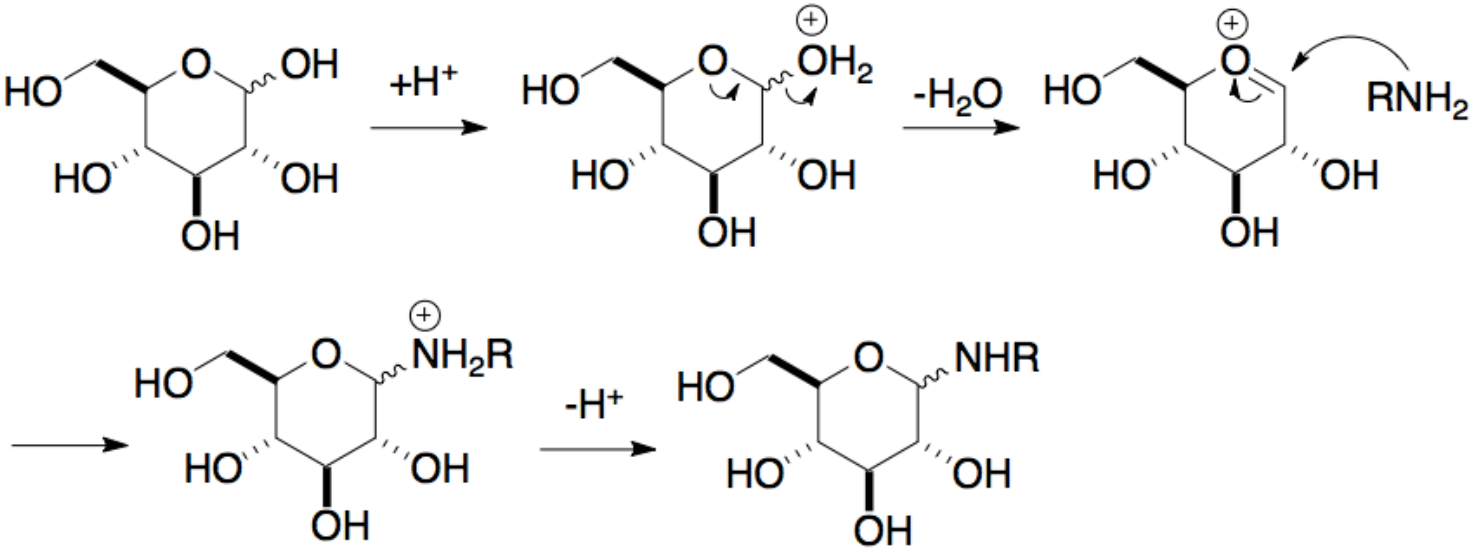
N-Glycosides

- If we treat a monosaccharide with an amine in the presence of an acid catalyst, the corresponding *N*-glycoside is formed.



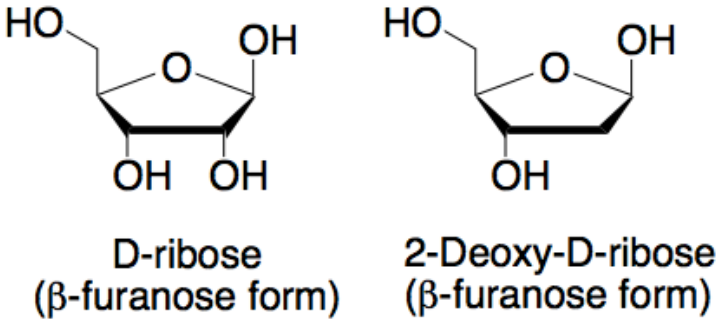
β-D-Glucopyranose A β-*N*-glycoside An α-*N*-glycoside

N-Glycosides



N-Glycosides - Nucleosides

Two carbohydrates, D-ribose and 2-deoxy-D-ribose, form very important *N*-glycosides.

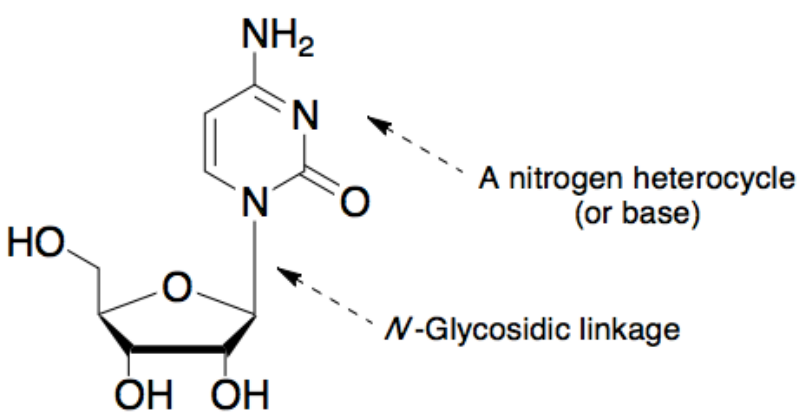


D-ribose (β-furanose form)

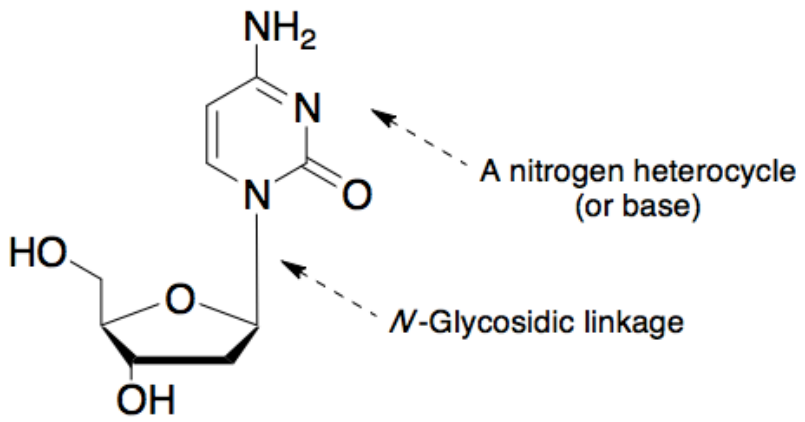
2-Deoxy-D-ribose (β-furanose form)

N-Glycosides - Nucleosides

- D-ribose and 2-deoxy-D-ribose serve as the building blocks for RNA and DNA.
- They are coupled with nitrogen heterocycles (called bases) to yield special *N*-glycosides, called **nucleosides**.
- The β anomers are formed exclusively.



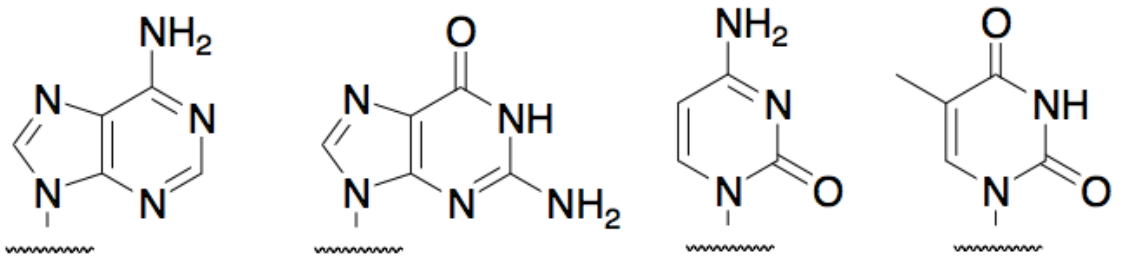
A ribonucleoside



A deoxyribonucleoside

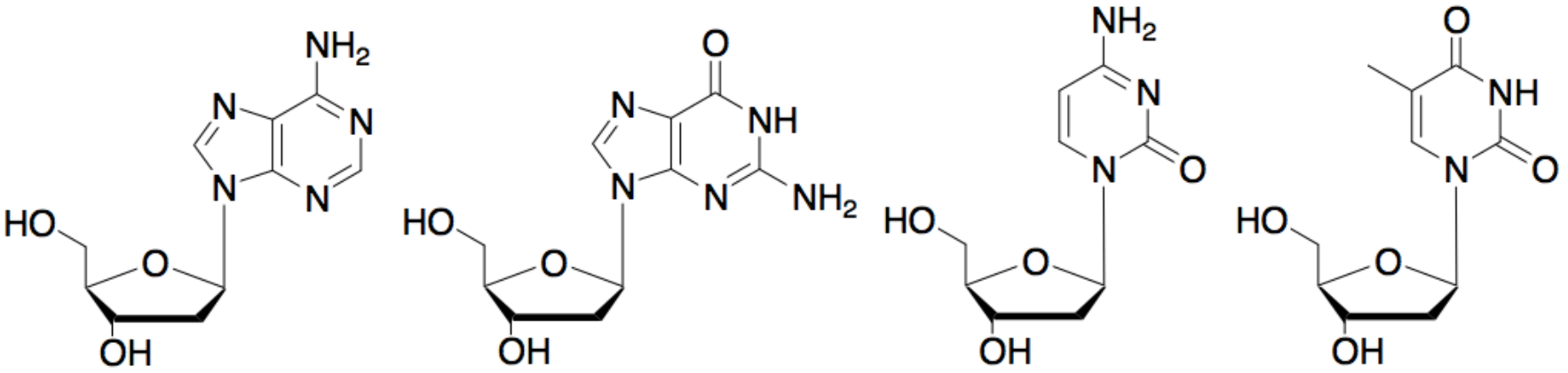
Heterocyclic amines in DNA

Four heterocyclic amines are found as bases in DNA:



Deoxyribonucleosides in DNA

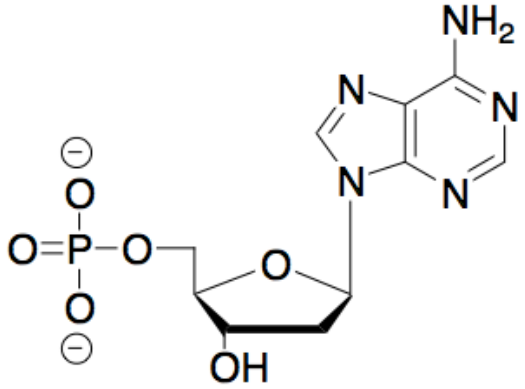
These four bases can couple to 2-deoxyribose to give four deoxyribonuleosides.



Deoxy-adenosine Deoxy-guanosine Deoxy-cytidine Deoxy-thymidine

Deoxyribonucleotides in DNA

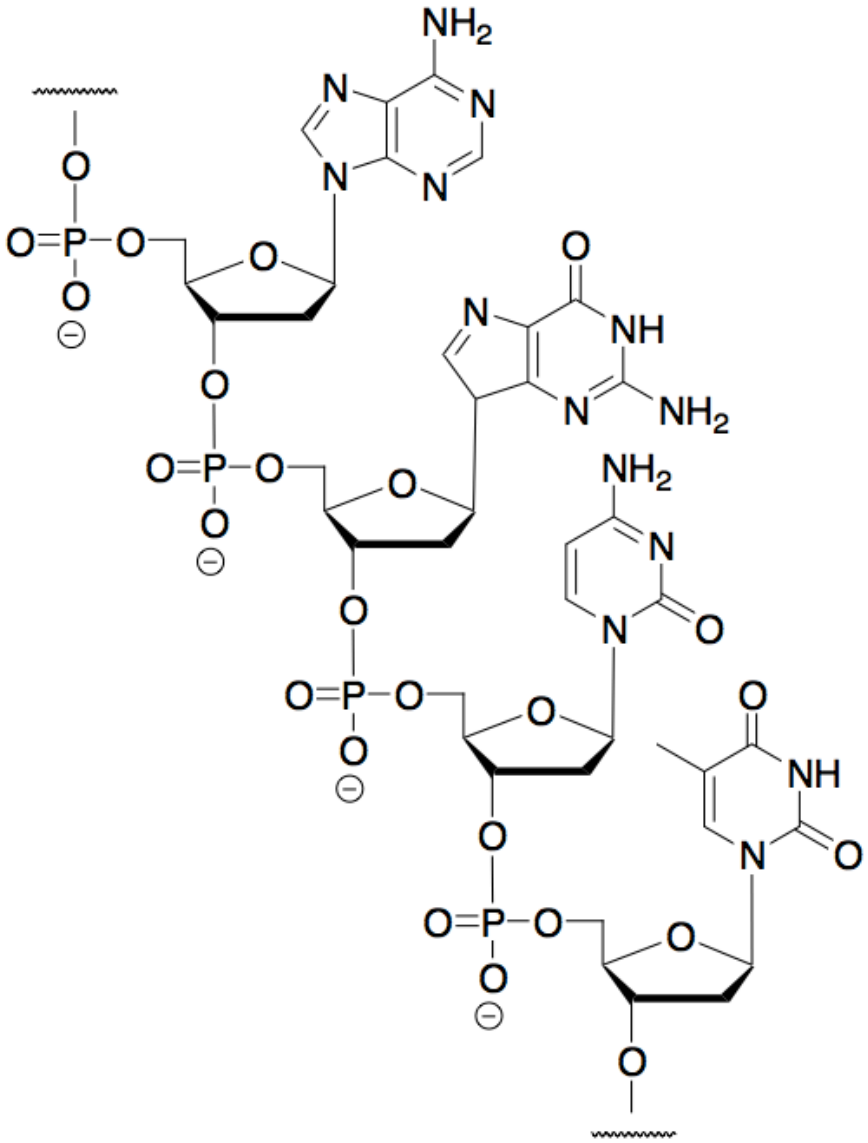
- A nucleoside can be coupled to a phosphate group to give a nucleotide.
- A deoxyribonucleotide is composed of three parts: deoxyribose, a nitrogen-containing base and a phosphate group



A deoxyribonucleotide
Deoxy-adenosine monophosphate

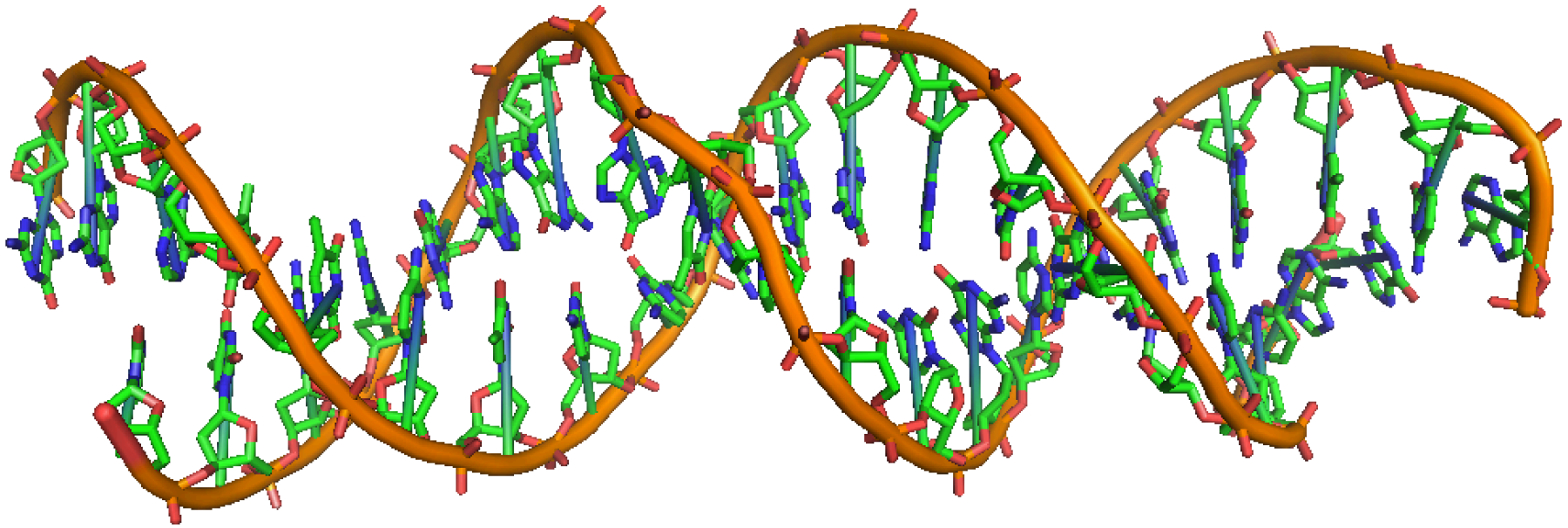
Polynucleotides in DNA

In DNA, nucleotides link together in a polymer (a polynucleotide). Each sugar is connected to two phosphate groups at C3' and C5'.



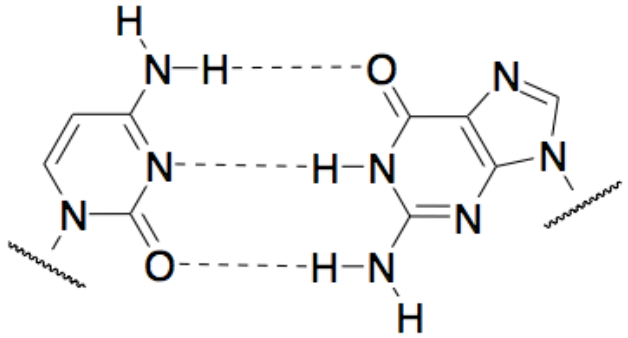
Double helix in DNA

In 1953, Crick and Watson proposed that DNA is a double helix composed of two polynucleotides with complementary base sequences.

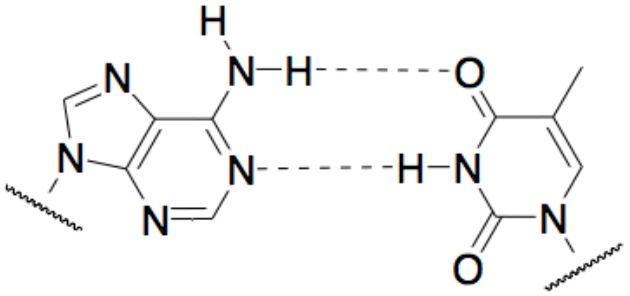


Hydrogen bonding between complementary base pairs in DNA

- Each base pairs specifically with another (C-G) and (A-T).
- There is hydrogen bonding between an amine hydrogen and a carbonyl oxygen and between an amide hydrogen and an imine nitrogen.
- In each case a purine is paired with a pyrimidine as only the combination of smaller pyrimidine and larger purine bridges the gap between the polynucleotides in the DNA helix.



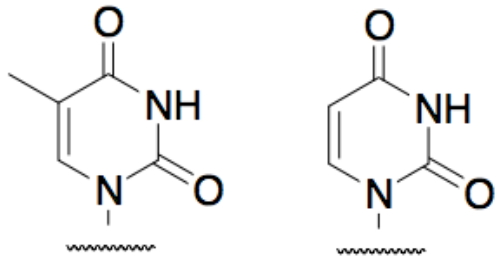
Cytosine Guanine



Adenine Thymine

Ribonucleosides in RNA

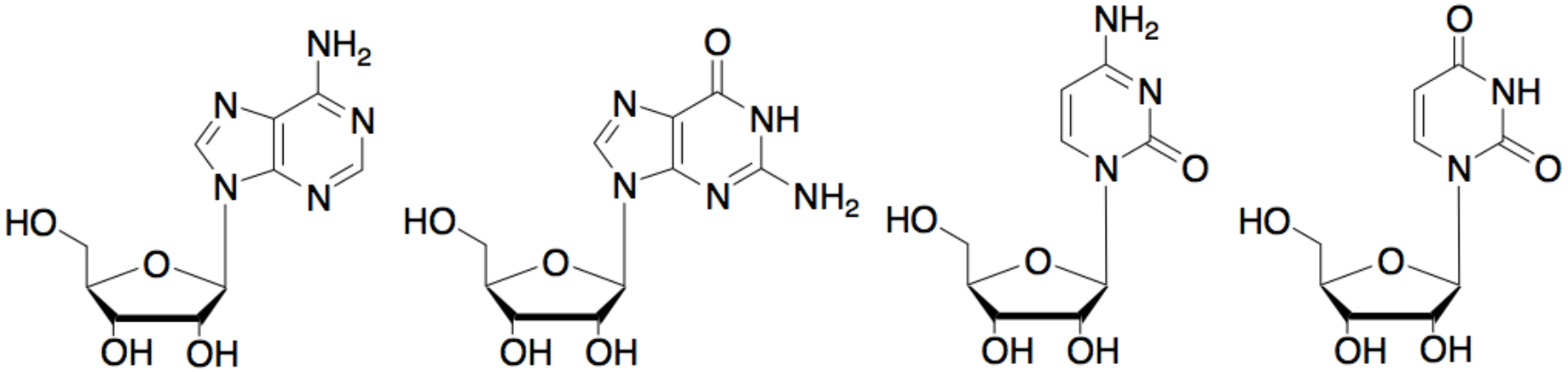
- DNA encodes our genetic information and serves as a template for RNA.
- A strand of RNA is similar in structure to a single strand of DNA but there are two important differences: The sugar in RNA is D-ribose and RNA contains the base uracil instead of thymine.



Thymine Uracil

Ribonucleosides in RNA

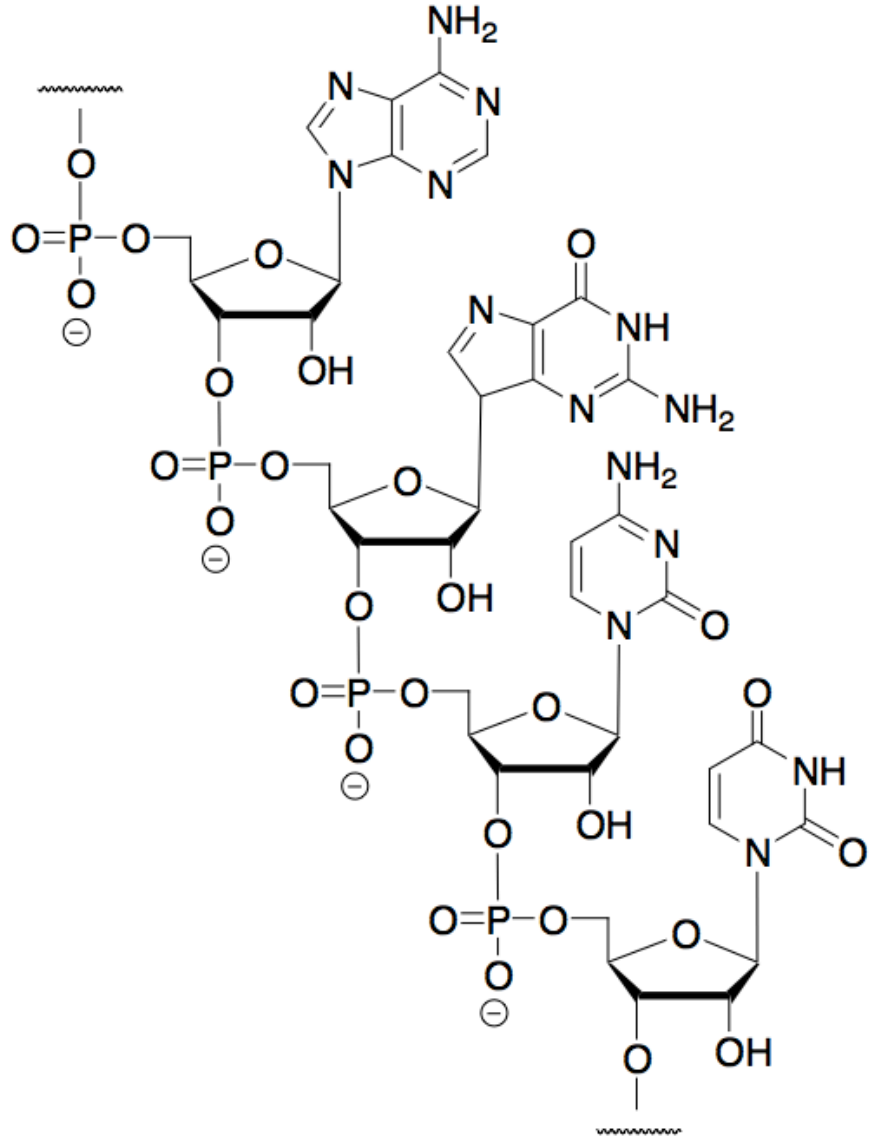
The four bases of RNA give rise to the following four RNA nucleosides:



Adenosine Guanosine Cytidine Uridine

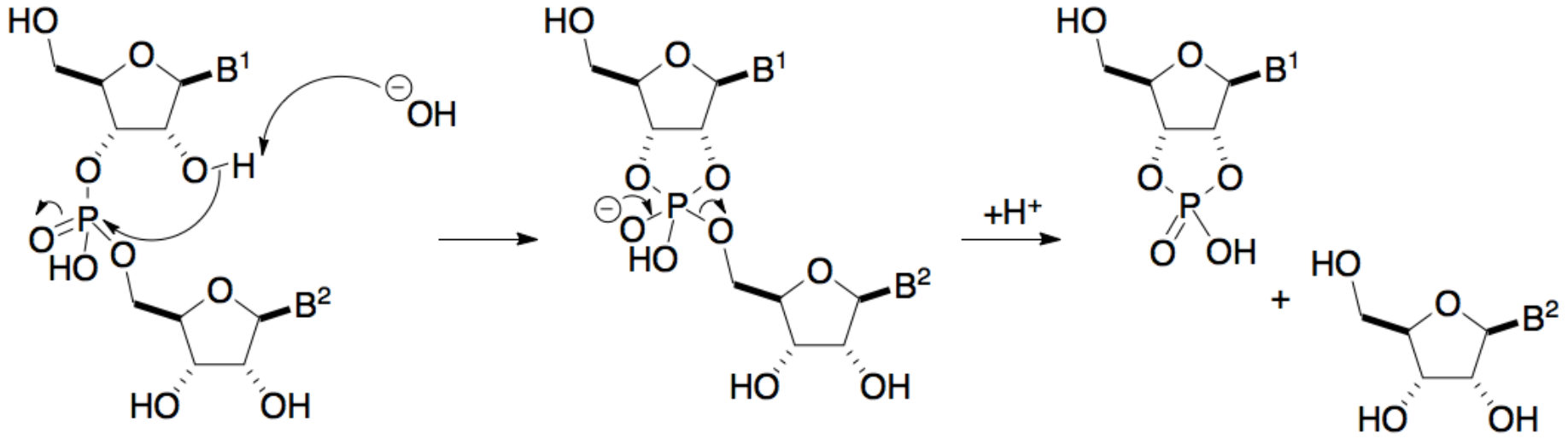
Polynucleotides in RNA

- Like in DNA, RNA nucleotides link together in a polymer (a polynucleotide). Each sugar is connected to two phosphate groups at C3’ and C5’.
- RNA directs the assembly of proteins and enzymes used to catalyse biochemical reactions.



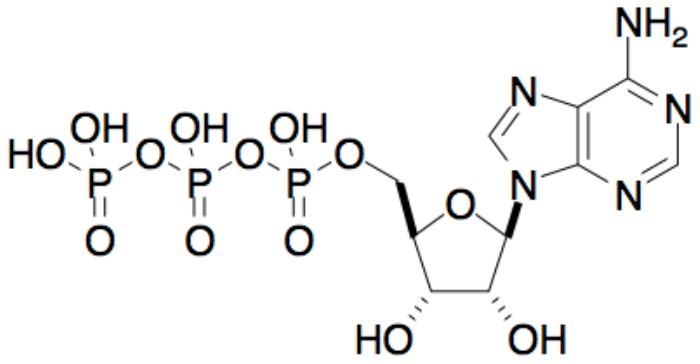
DNA is more stable than RNA

- DNA is more stable than RNA because its sugars lack the 2′ hydroxyl group.
- In RNA the 2′- and 3′-OH groups are on the same side of the ring and so alkaline hydrolysis is exceptionally rapid by intramolecular nucleophilic catalysis.



ATP

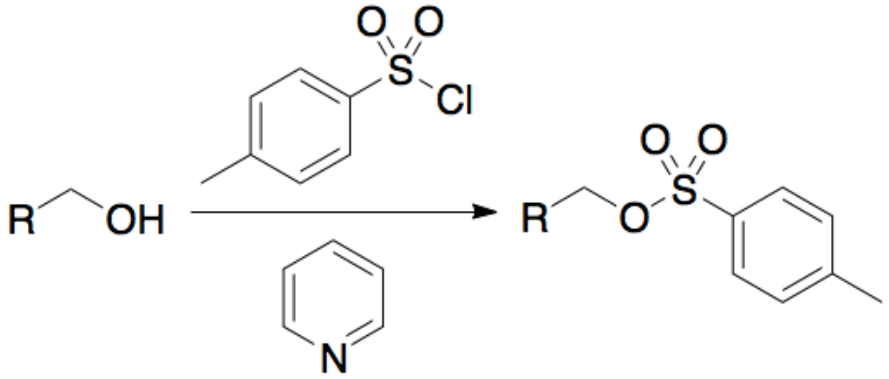
- The nucleotide, adenosine triphosphate (ATP) is one of the most important molecules in nature.
- It is highly reactive as phosphates are stable anions and thus good leaving groups.
- ATP can be attacked by soft nucleophiles at C5’ on the sugar and by hard nucleophiles at a phosphate group.

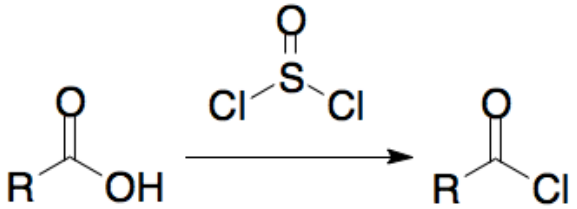


Adenosine triphosphate (ATP)

ATP

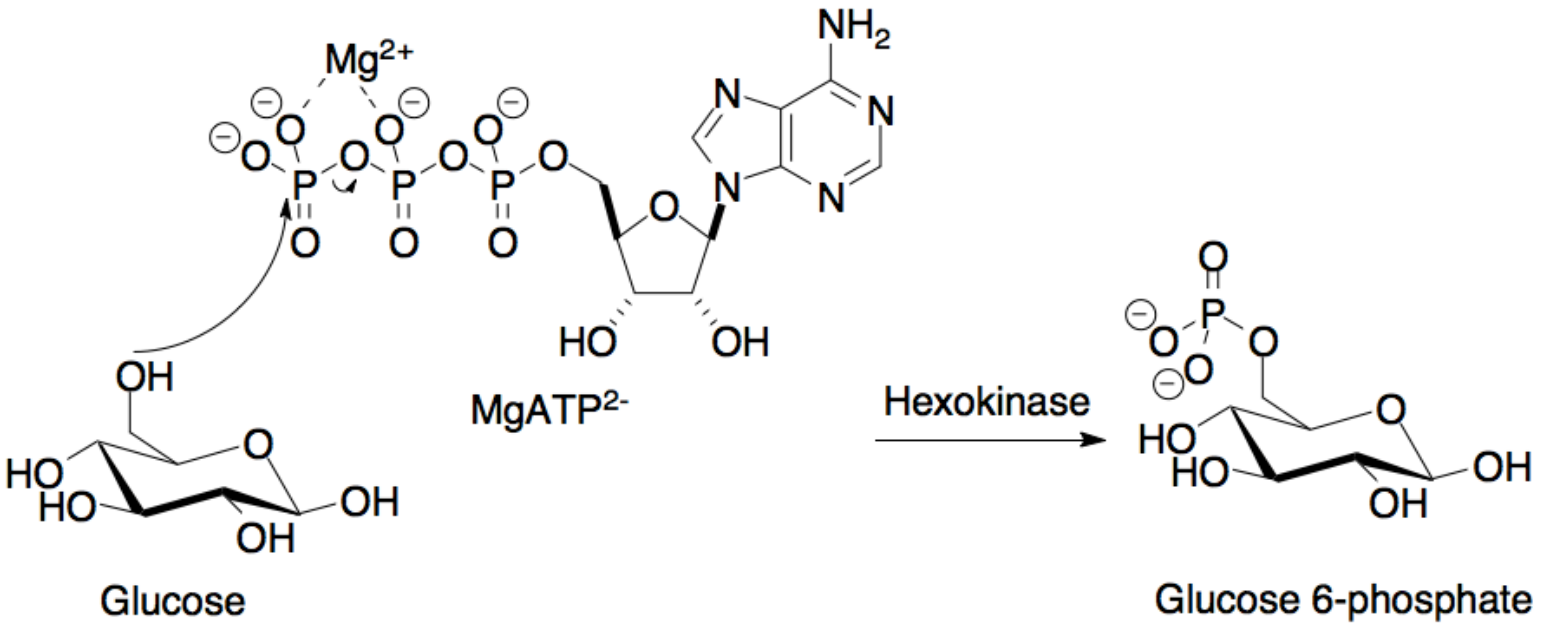
ATP is used in nature in a similar way to how chemists use tosyl chloride to make alcohols more reactive or use thionyl chloride to make carboxylic acids more reactive.





ATP

- Glucose is phosphorylated at the hydroxyl group on C6 by ATP to give glucose 6-phosphate.
- This is the first reaction in the glycolytic pathway.



ATP

- The standard Gibbs energy for the hydrolysis of ATP^{4-} to ADP^{3-} at pH 7 is −30.5 kJ mol^{-1}.

$$\mathbf{ATP^{4-} + H_2O \rightarrow ADP^{3-} + H^+ + HPO_4^{2-}}\ \Delta_r G^{\circ\prime} = -30.5\ \text{kJ mol}^{-1}$$

- The standard Gibbs energy for the phosphorylation of glucose to glucose 6-phosphate at pH 7 is +13.8 kJ mol^{-1}.

$$\mathbf{Glucose + H^+ + HPO_4^{2-} \rightarrow Glucose\ 6\text{-}phosphate + H_2O}\ \Delta_r G^{\circ\prime} = +13.8\ \text{kJ mol}^{-1}$$

- Coupling the endergonic process to the exergonic hydrolysis of ATP^{4-} makes the reaction take place spontaneously, $\Delta_r G^{\circ\prime} = -17.2$ kJ mol^{-1}.

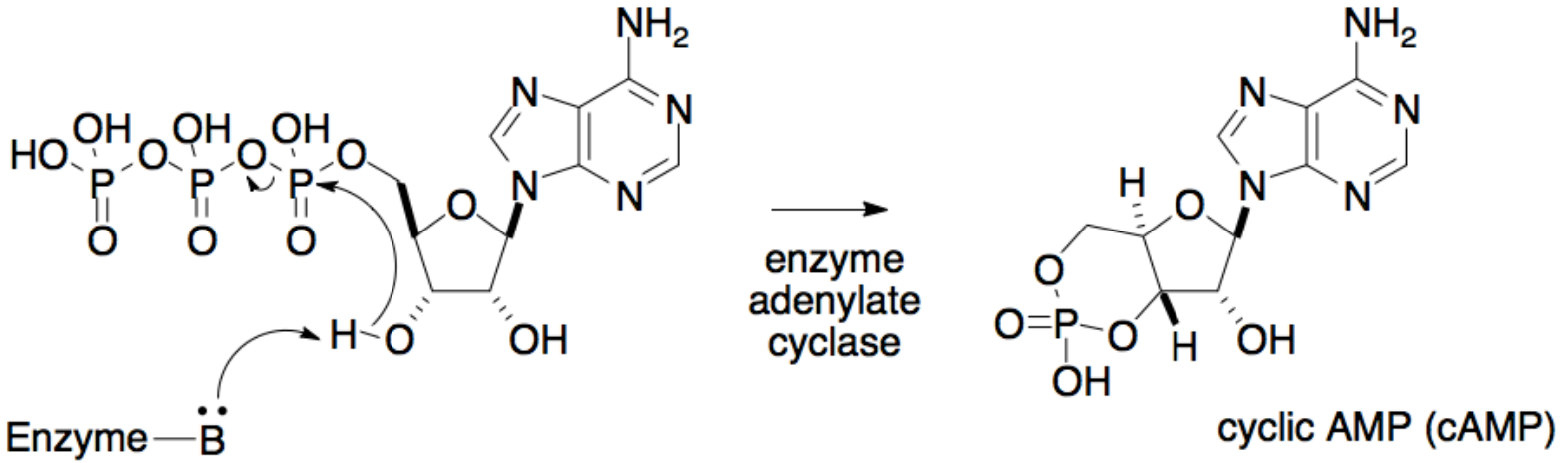
ATP

The Gibbs energy change for the hydrolysis of ATP^{4-} to ADP^{3-} is negative because:

- Electrostatic repulsion between the negative charges raises the energy of ATP^{4-}.
- ADP^{3-} and $H_3PO_4^{2-}$ have a lower energy than ATP because they possess more resonance structures than ATP.
- On hydrolysis, there is a release of steric crowding among oxygen atoms on adjacent phosphate groups in ATP.

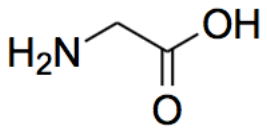
Cyclic AMP

- Cyclic AMP (cAMP) is an important biological messenger which helps control blood clotting and acid secretion in the stomach.
- cAMP is formed enzymatically from ATP.

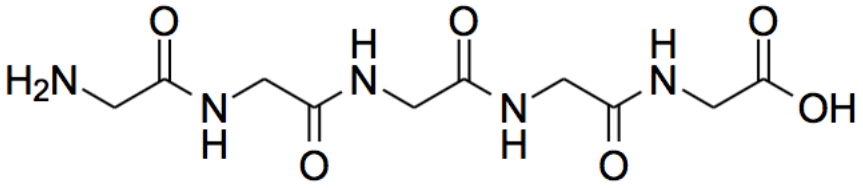


Amino acids

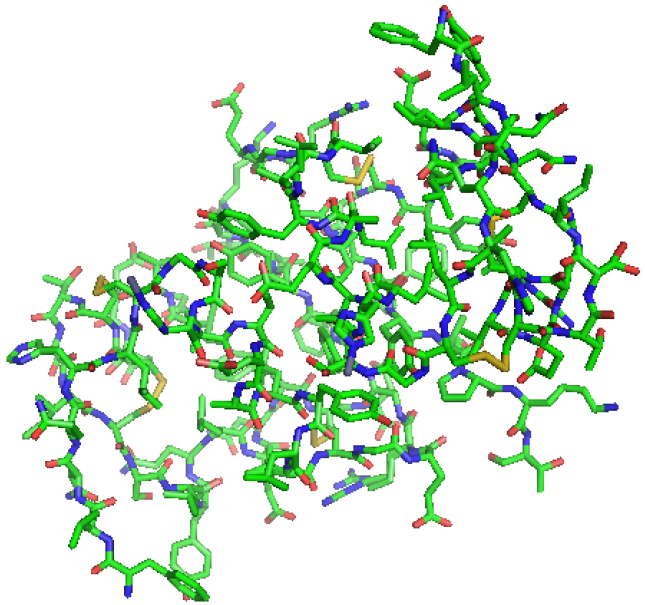
Amino acids, peptides and proteins

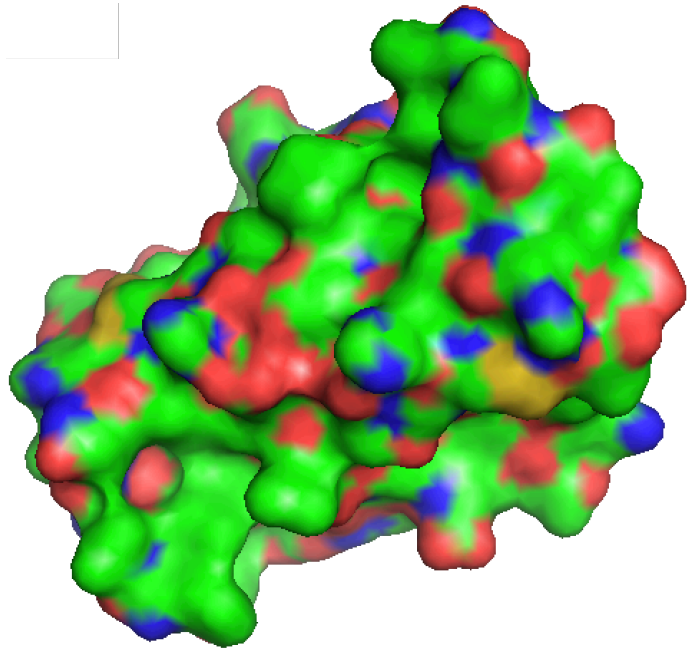


amino acid



peptide

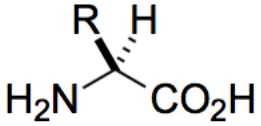




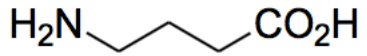
Protein (insulin)

Amino acids, peptides and proteins

- Proteins make up 15% of the mass of cells and serve an enormous variety of biological functions.
- The basic building blocks of proteins are amino acids.
- Proteins are linear polymers comprised of α-amino acids linked by amide (peptide) bonds.
- Amino acids are carboxylic acids that bear an amine group.
- An amino acid can have any number of carbon atoms separating the two functional groups but the most common are the α-amino acids.
- There are hundreds of amino acids found in nature but most proteins consist mainly of only 20.

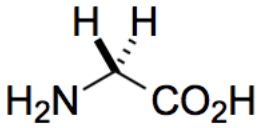


α-amino acid



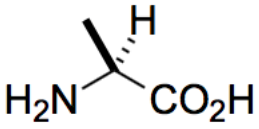
γ-aminobutyric acid, GABA

Amino acids - Nomenclature

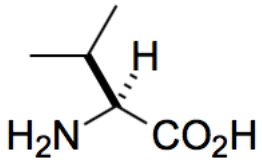


Glycine
Gly
G

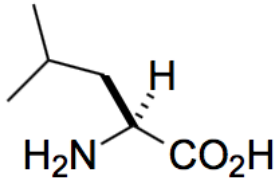
Alkyl



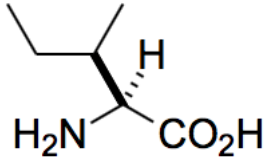
Alanine
Ala
A



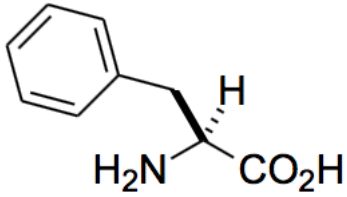
Valine
Val
V



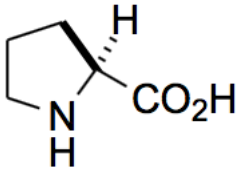
Leucine
Leu
L



Isoleucine
Ile
I



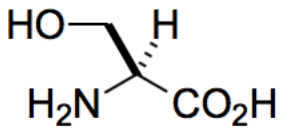
Phenylalanine
Phe
F



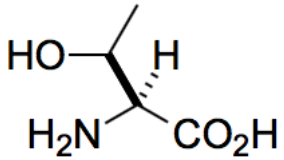
Proline
Pro
P

Amino acids - Nomenclature

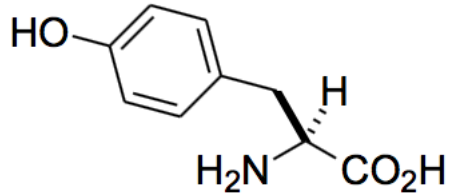
Hydroxy containing



Serine
Ser
S

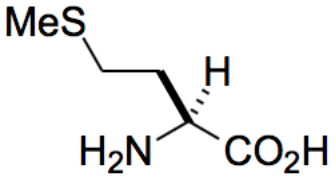


Threonine
Thr
T

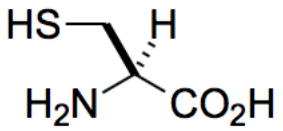


Tyrosine
Tyr
Y

Mercapto or sulfide containing



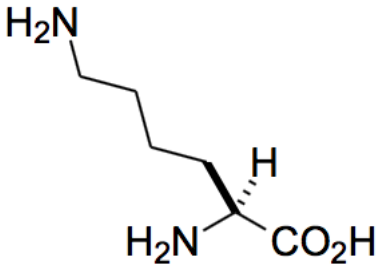
Methionine
Met
M



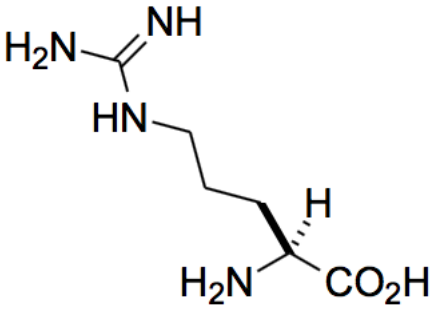
Cysteine
Cys
C

Amino acids - Nomenclature

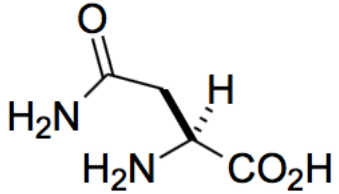
Amino containing



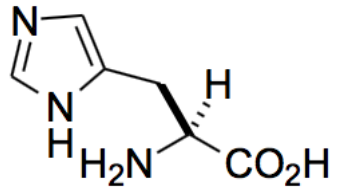
Lysine
Lys
K



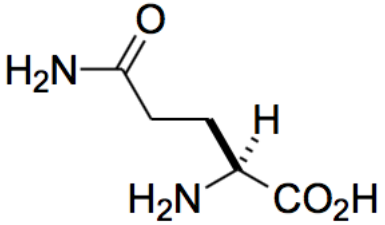
Arginine
Arg
R



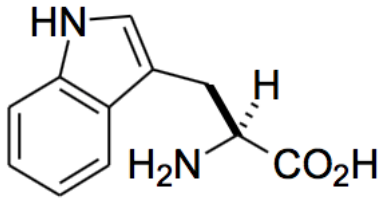
Asparagine
Asn
N



Histidine
His
H



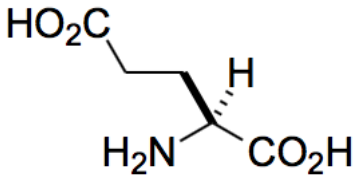
Glutamine
Gln
Q



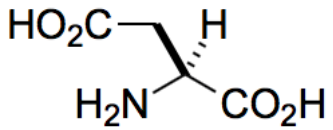
Tryptophan
Trp
W

Amino acids - Nomenclature

Carboxy containing



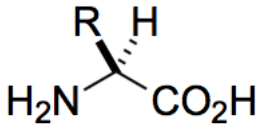
Glutamic acid
Glu
E

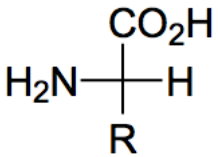


Aspartic acid
Asp
D

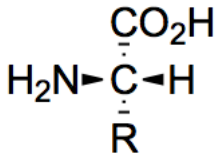
Amino acids - Stereochemistry

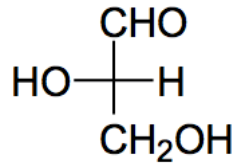
- Amino acids can be drawn as hashed wedge line structures or as Fischer diagrams.
- With the exception of glycine, the α carbon is a stereocentre.
- Like for sugars, the D and L nomenclature is still in use and natural α-amino acids usually have the L configuration.
- The absolute configuration depends on the side chain but natural α-amino acids usually have the (*S*) configuration.



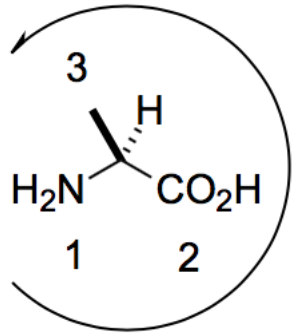


L



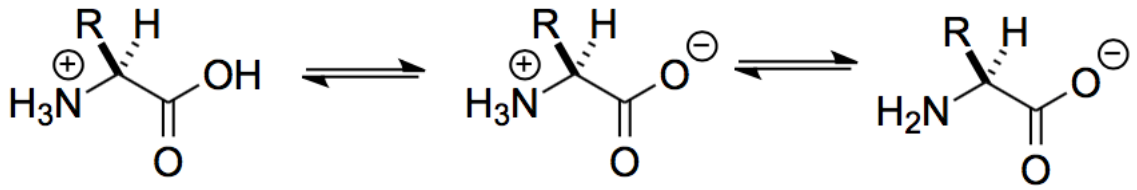


L-Glyceraldehyde



(*S*)

Amino acids are acidic and basic: Zwitterions



An ammonium ion predominates at pH < 1

A neutral zwitterion predominates at pH ≈ 6

A carboxylate ion predominates at pH > 13

Increasing pH

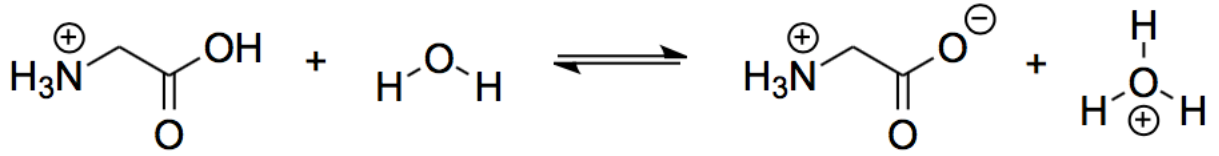
pK_a of an ammonium ion $\approx$ 10–11

pK_a of a carboxylate ion $\approx$ 2–5

Amino acids are acidic and basic: Isoelectric point

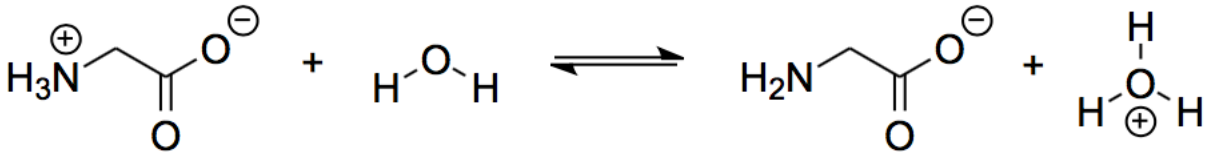
Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Glycine	2.3	9.6	-	6.0

The pK_a of –COOH refers to the following equilibrium:



$$K_1 = \frac{[H_3\overset{+}{N}CH_2COO^-][H^+]}{[H_3\overset{+}{N}CH_2COOH]} = 10^{-2.3}$$

The pK_a of –$^+NH_3$ refers to the following equilibrium:



$$K_2 = \frac{[H_2NCH_2COO^-][H^+]}{[H_3\overset{+}{N}CH_2COO^-]} = 10^{-9.6}$$

Amino acids are acidic and basic: Isoelectric point

Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Glycine	2.3	9.6	-	6.0

The pH at which a molecule carries no net charge is called the isoelectric point.

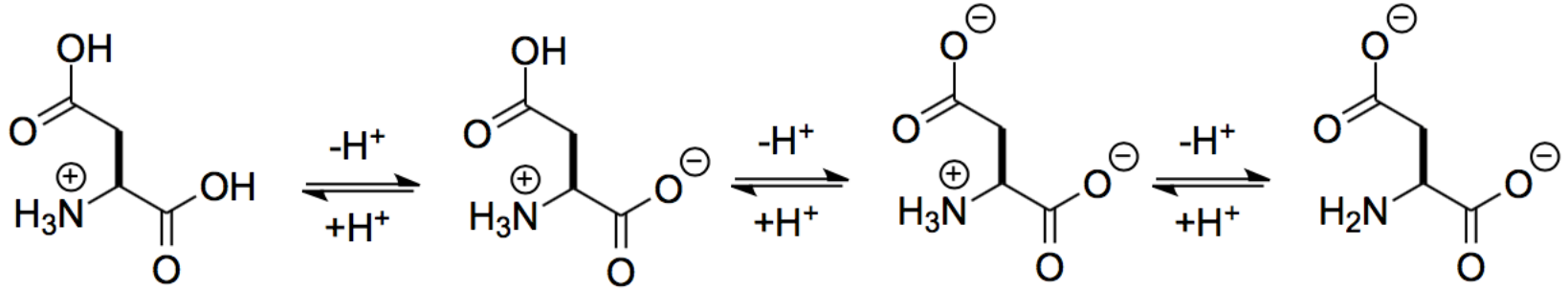
The isoelectric point for glycine is:

$$pI = \frac{pK_a\ {-COOH} + pK_a\ \text{of}\ {-^+NH_3}}{2} = 6.0$$

Amino acids are acidic and basic: Isoelectric point

Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Aspartic acid	1.9	9.6	3.7	2.8

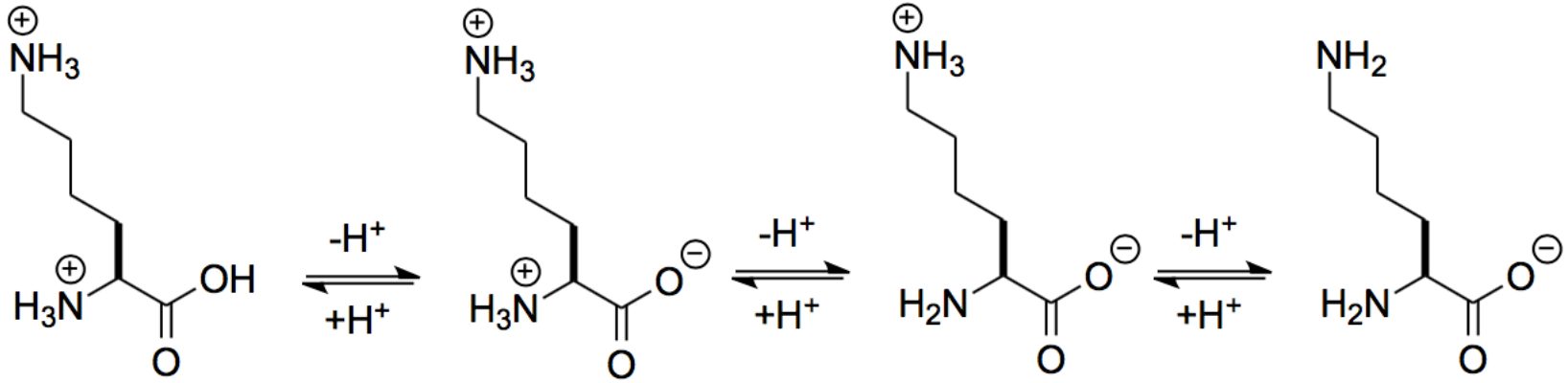
- When the side chain of the amino acid bears an additional acidic or basic group, this has an effect on the isoelectric point.
- When the side chain is acidic, the p*I* is the average of its two lowest pK_a values.



Amino acids are acidic and basic: Isoelectric point

Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Lysine	2.2	9.0	10.5	9.7

When the side chain is basic, the isoelectric point is the average of its two highest pK_a values.



Amino acids are acidic and basic: Isoelectric point

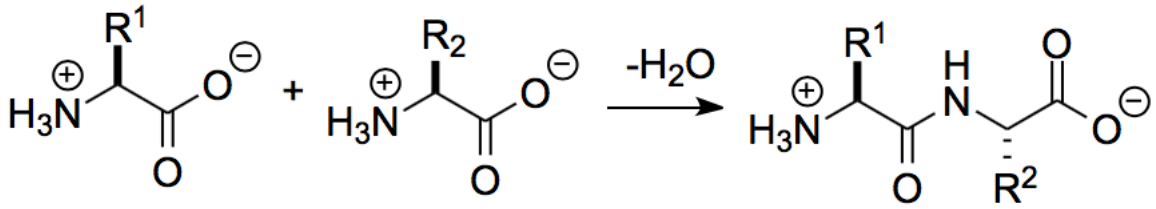
Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Glycine	2.3	9.6	-	6.0
Alanine	2.3	9.7	-	6.0
Valine	2.3	9.6	-	6.0
Leucine	2.4	9.6	-	6.0
Isoleucine	2.4	9.6	-	6.0
Phenylalanine	1.8	9.1	-	5.5
Proline	2.0	10.6	-	6.3
Serine	2.2	9.2	-	5.7
Threonine	2.1	9.1	-	5.6
Tyrosine	2.2	9.1	10.1	5.7

Amino acids are acidic and basic: Isoelectric point

Amino acid	pK_a of -COOH	pK_a of -$^+NH_3$	pK_a of acidic function in R	Isoelectric point, p*I*
Asparagine	2.0	8.8	-	5.4
Glutamine	2.2	9.1	-	5.7
Lysine	2.2	9.0	10.5	9.7
Arginine	2.2	9.0	12.5	10.8
Tryptophan	2.8	9.4	-	5.9
Histidine	1.8	9.2	6.1	7.6
Cysteine	2.0	10.3	8.2	5.1
Methionine	2.3	9.2	-	5.7
Aspartic acid	1.9	9.6	3.7	2.8
Glutamic acid	2.2	9.7	4.3	3.2

Peptides

- Amino acids can be linked together to form peptides via an amide bond (peptide bond).
- This is formally a dehydration reaction (water is lost).

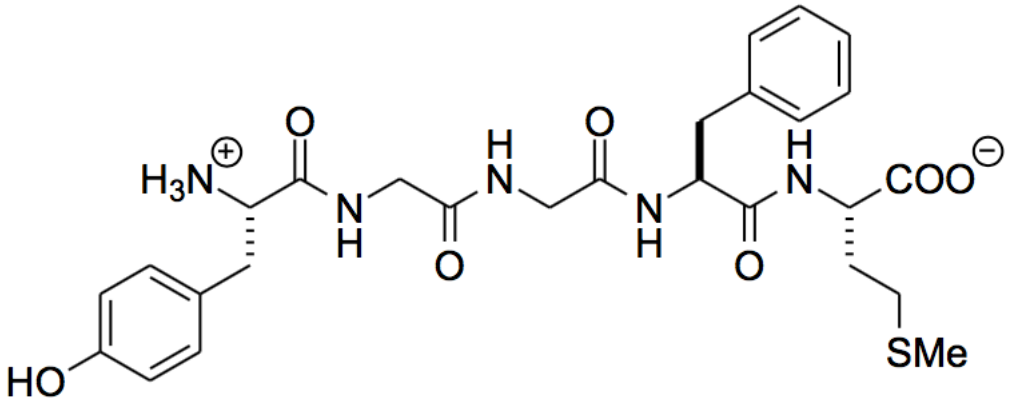


- Long polypeptide chains (>50 amino acids) are called proteins.

Proteins – Primary structure

- The primary structure of proteins is the order in which amino acids are linked together through amide bonds.

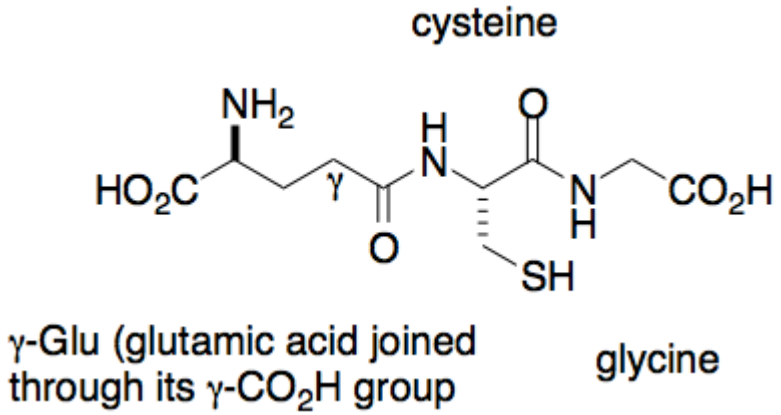
- The primary structure of Met-enkephalin (one of the body's painkillers):



Tyr-Gly-Gly-Phe-Met
YGGFM

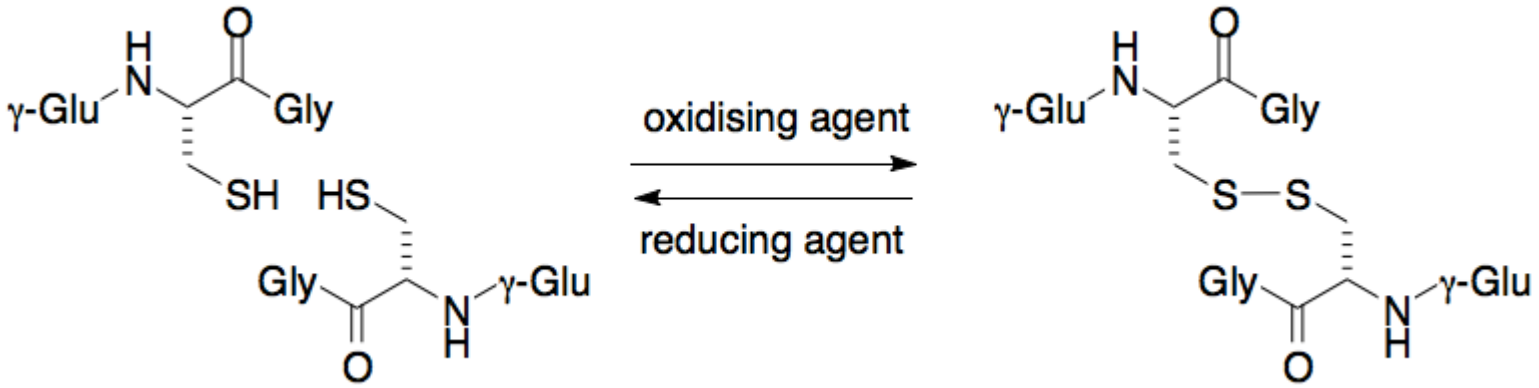
Proteins – Primary structure – Glutathione

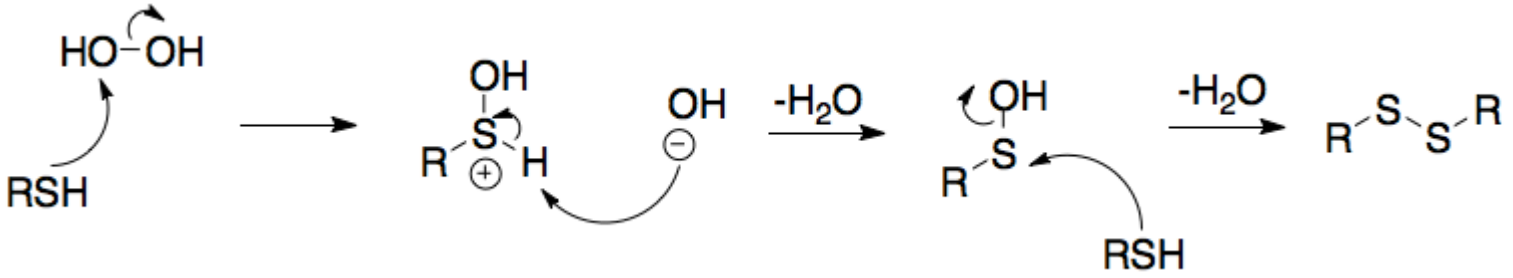
- Glutathione is an important tripeptide, found in animals and plants.
- It is an atypical tripeptide in that glutamic acid bonds to cysteine through the γ-CO_2H group instead of the α-CO_2H group.



Proteins – Glutathione

Glutathione is an antioxidant, it removes dangerous oxidising agents by giving itself up to be oxidised to a disulfide.

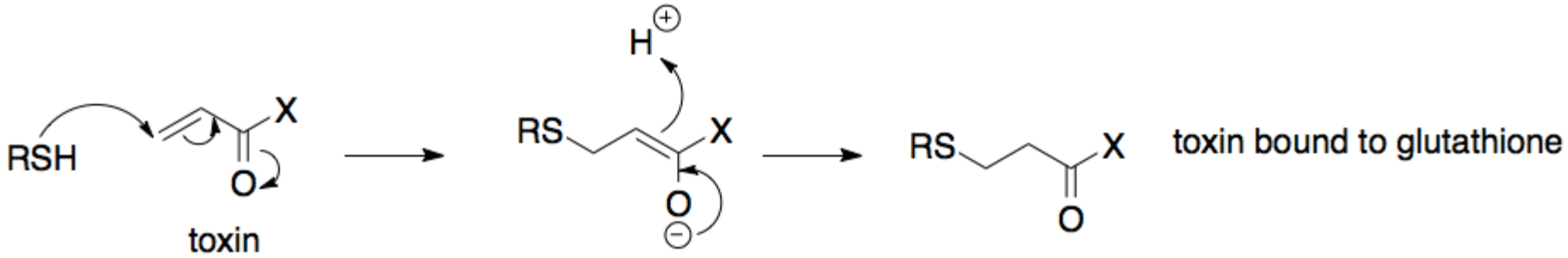




Proteins – Glutathione

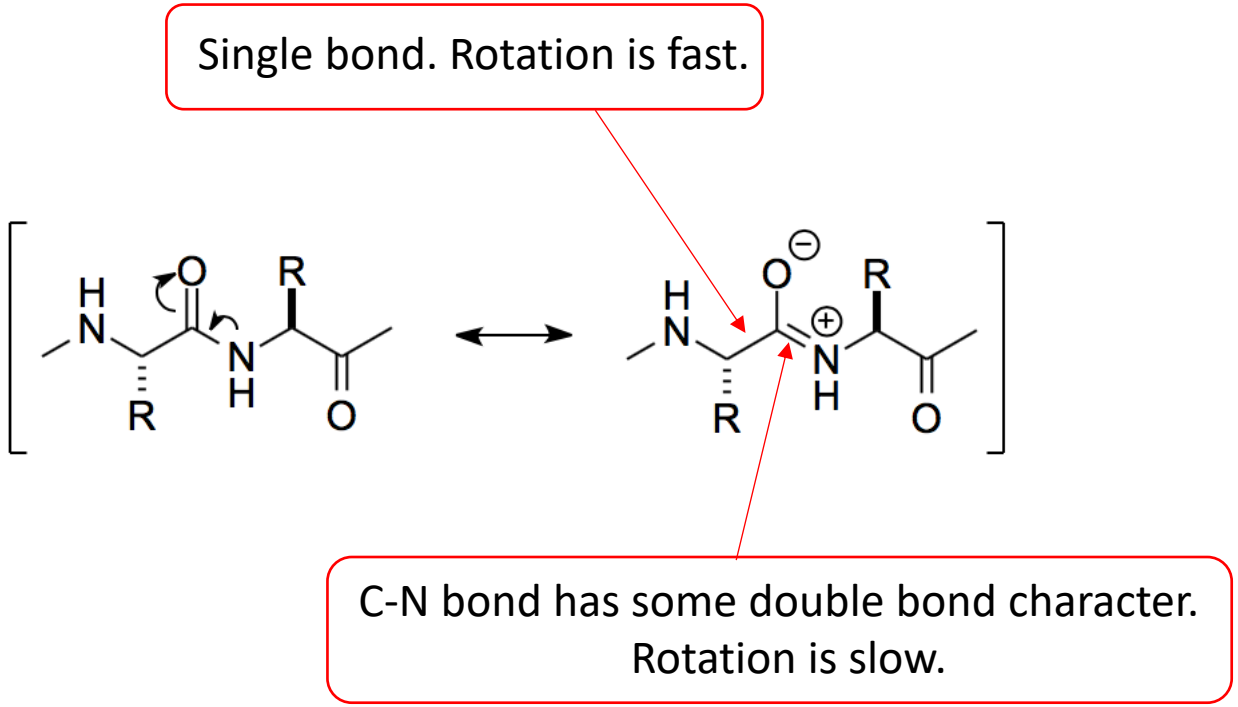
- Glutathione also detoxifies Michael acceptors which are carcinogenic because they react with the enzyme DNA polymerase, an important enzyme needed for DNA synthesis and repair.

- Glutathione reacts with Michael acceptors *via* nucleophilic addition and once bound to glutathione they are harmless and can be excreted.



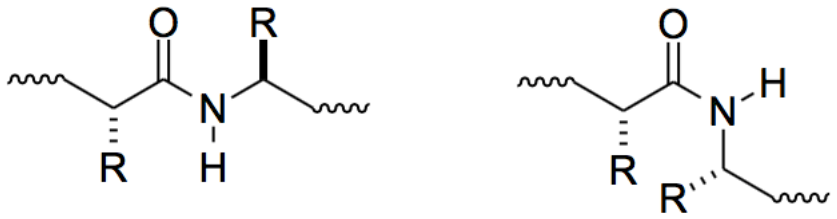
Proteins – Primary structure – The planar peptide bond

The amide bond is planar due to conjugation and this gives the polypetide structure a certain rigidity.



Proteins – Primary structure – The planar peptide bond

- The partial double bond character of the amide bond means there are two possible conformations that can be adopted.

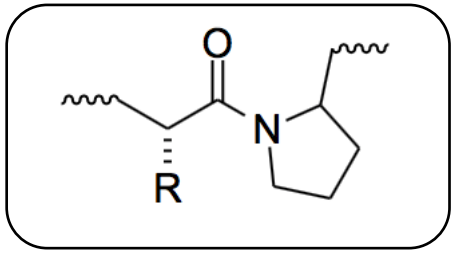


trans *cis*

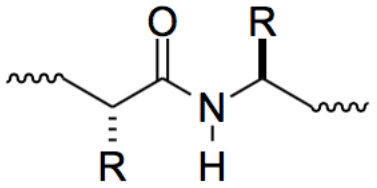
- The *trans* confirmation is preferred in most cases as it avoids an unfavourable steric clash between C-atoms in the protein chain and between R groups.

Proteins – Primary structure – The planar peptide bond

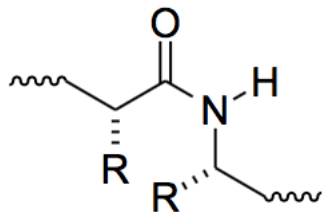
- In the peptide bond to the nitrogen of proline, the *trans* conformation, has a greater steric clash than other *trans* peptides.



trans proline peptide



trans



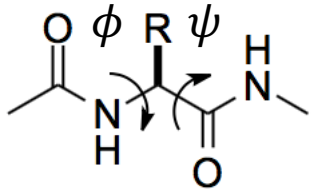
cis

- The steric clash present in the *trans* conformation is similar to that in a *cis* peptide bond and so the *cis* peptide conformation of proline is not disfavoured.

Three Dimensional Protein Structure

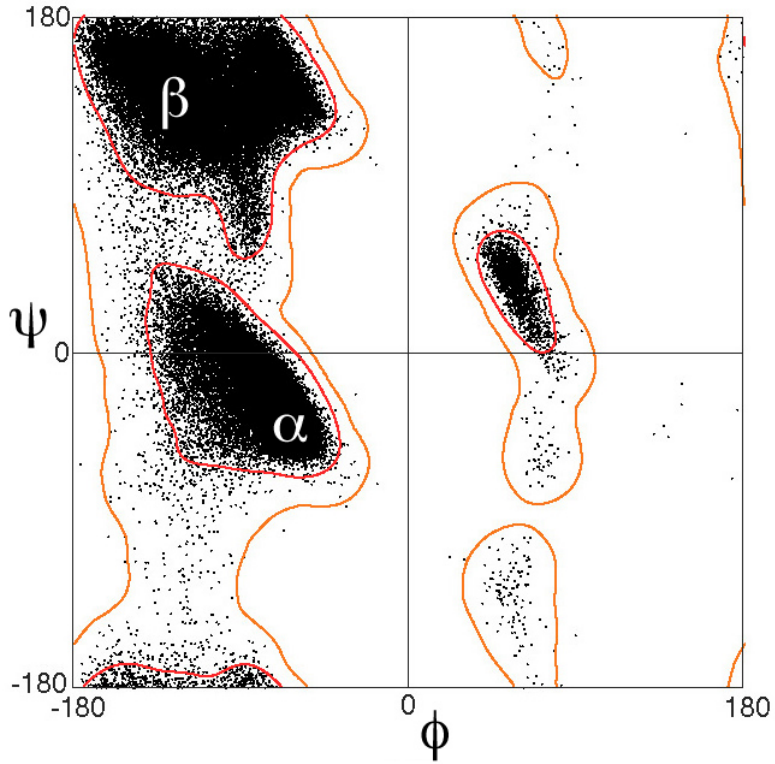
Proteins – Secondary structure – Phi and Psi angles

- The amide bond is planar but there is free rotation around the N-C_α and C-C_α bonds.
- The N-C_α dihedral angle is labelled ϕ (Phi).
- The C-C_α dihedral angle is labelled ψ (Psi).



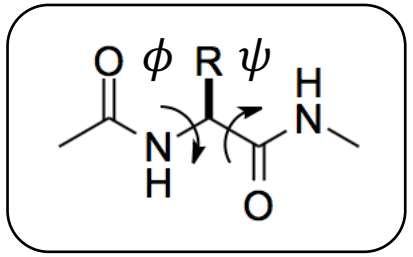
Proteins – Secondary structure – Ramachandran plots

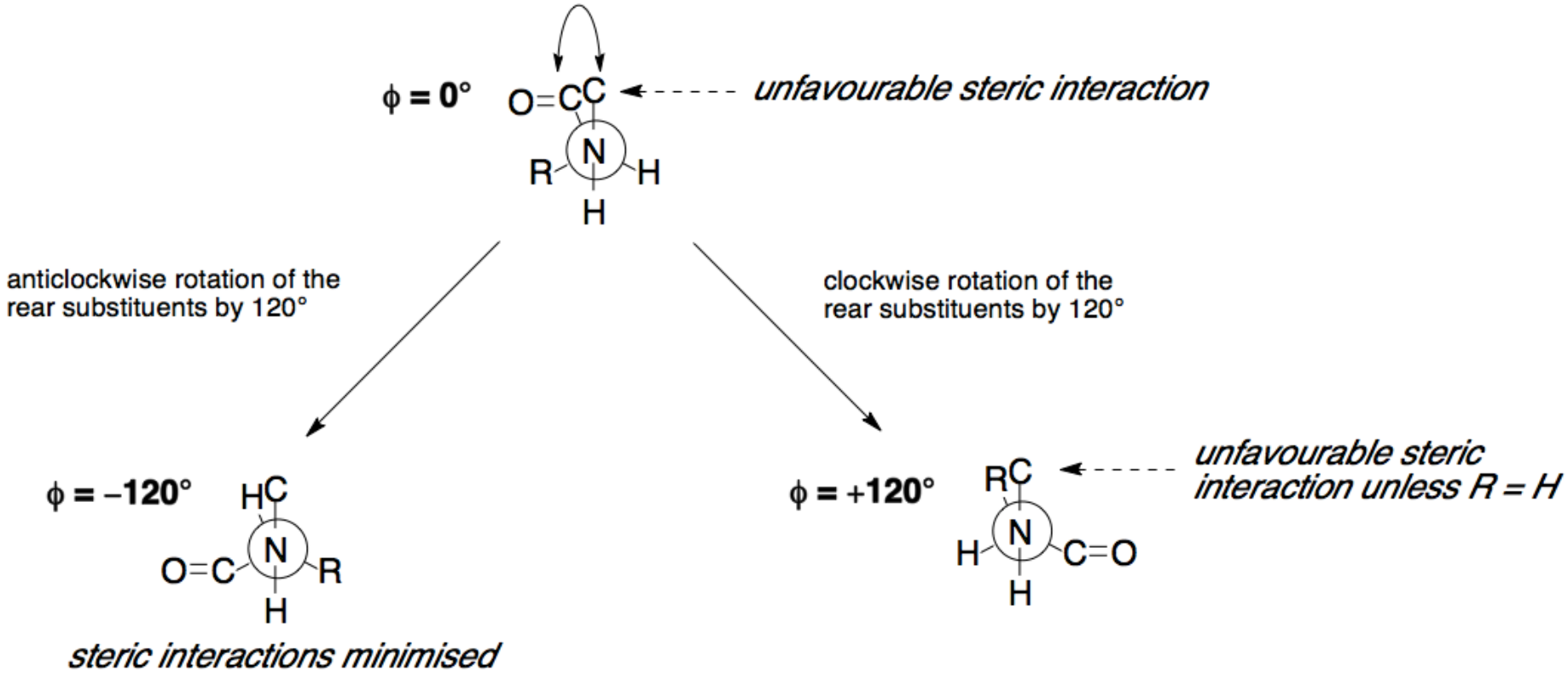
- G.N. Ramachandran, realised that not all dihedral angles are favoured due to steric clashes.
- Using molecular models, Ramachandran tested all possible angles and found that there are only three regions that contain favoured values of ϕ and ψ angles for normal natural proteins.
- To form the two most common secondary structures, alpha helices and beta sheets, we need several amino acid residues in series within one given region of this Ramachandran plot.



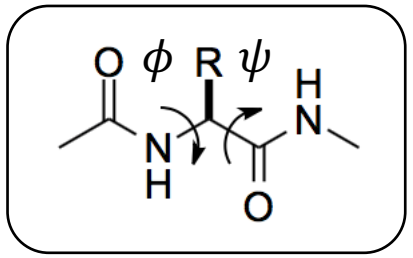
Ramachandran plot general, https://commons.wikimedia.org/wiki/File:Ramachandran_plot_general_100K.jpg#/media/File:Ramachandran_plot_general_100K.jpg, Licence: Creative commons

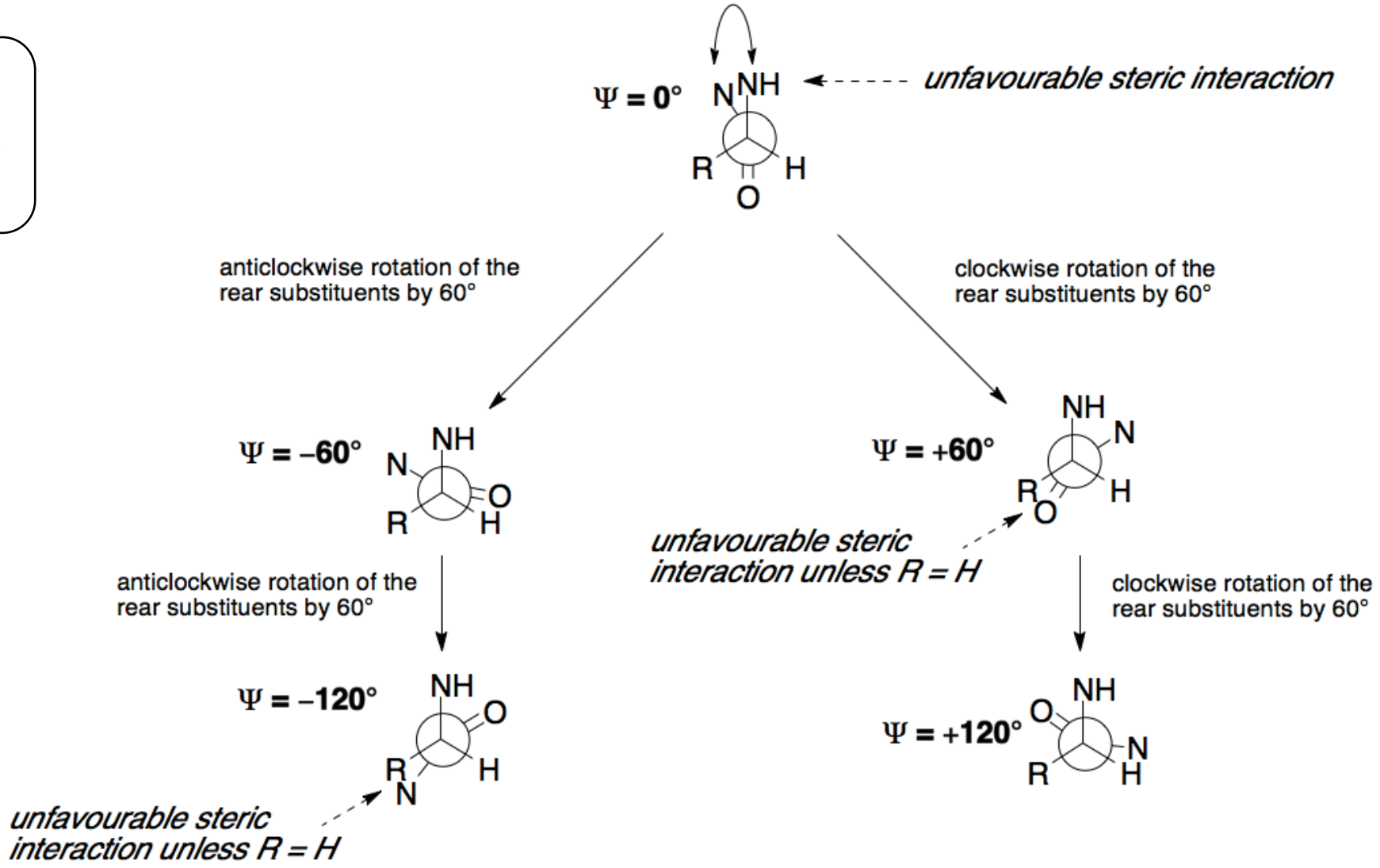
In order to minimise unfavourable steric interactions, L-amino acids tend to adopt conformations where ϕ (Phi) ≈ −60° to −180°.



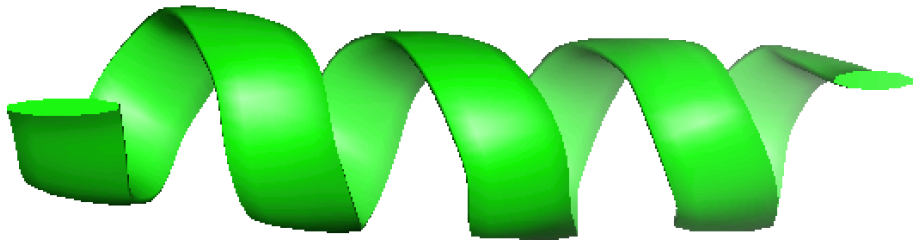


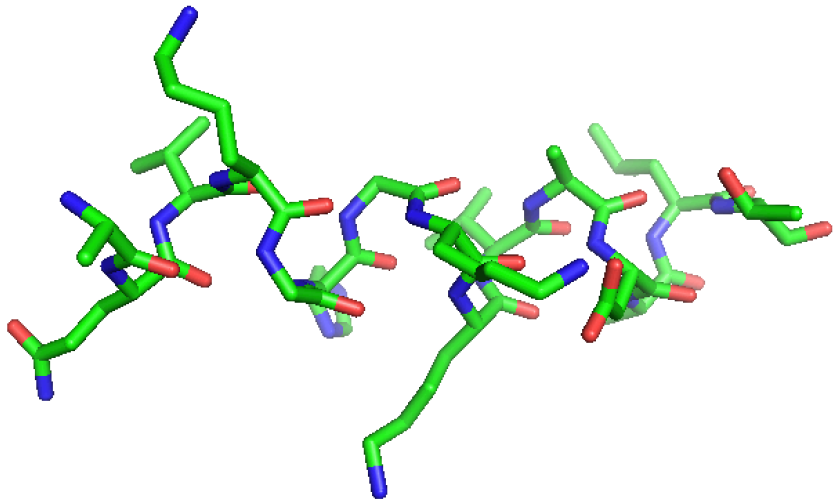
In order to minimise unfavourable steric interactions, L-amino acids tend to adopt conformations where ψ (Psi) ≈ −60° or +120° to +180°.



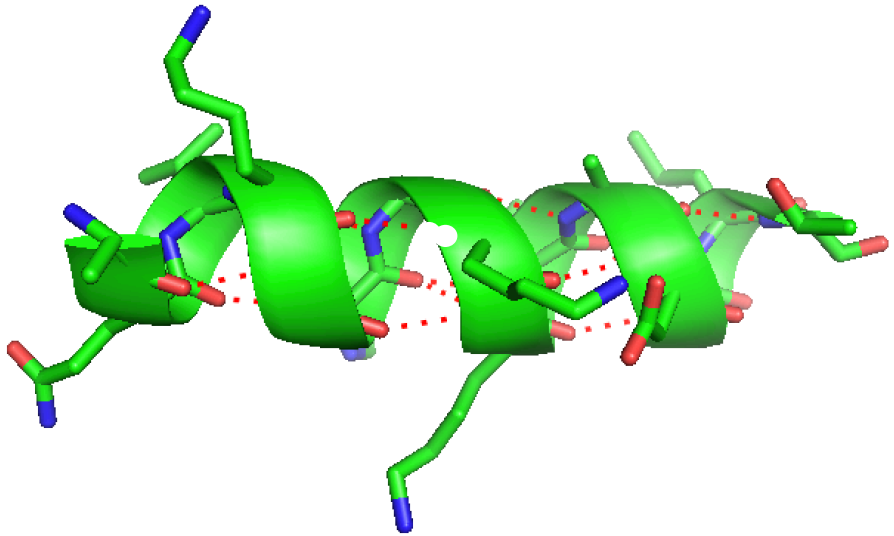


Proteins – Secondary structure – Alpha helix





- The alpha helix is a spiral held in place by intramolecular hydrogen bonding.
- Alpha helices in most proteins are *right-handed* because of *amino acid chirality* (L-amino acids).
- An alpha helix has 3.6 residues (amino acids) per turn and a pitch of 5.4 Å.

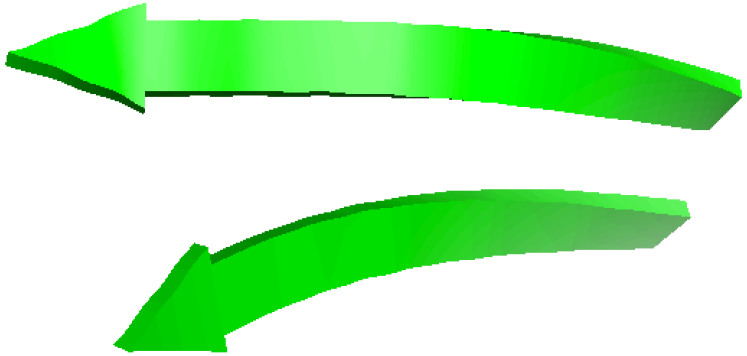


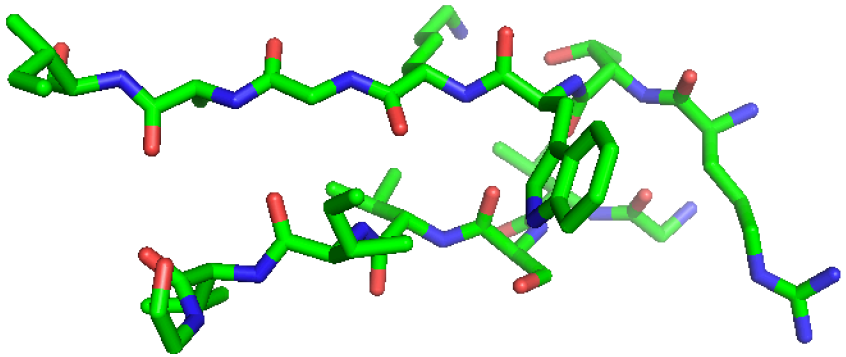
- Each amide oxygen hydrogen bonds to an amide hydrogen, four residues ahead.
- The R groups point away from the helix.

Proteins – Secondary structure – Alpha helix

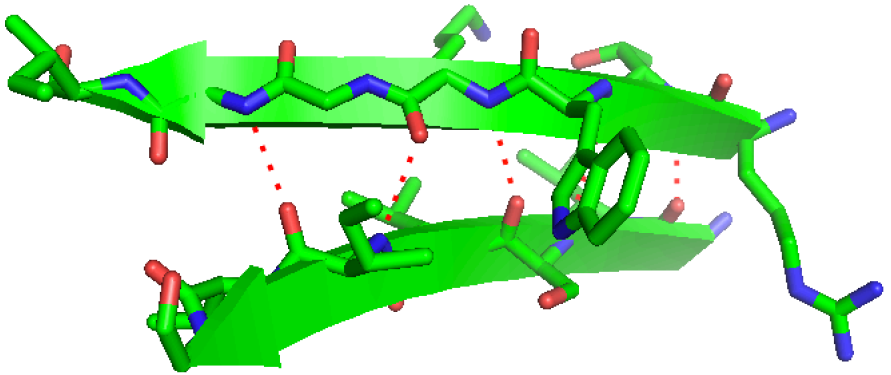
- Interactions between side chains can stabilise or destabilise the alpha-helix.
- Adjacent negatively charged residues destabilise so a long block of Glu residues at pH 7 will repel each other and prevent formation of an alpha-helix.
- Adjacent positively charged residues destabilise. Adjacent Lys and/or Arg residues at pH 7.0 will repel each other and prevent formation of an alpha-helix.
- If the bulky residues Asn, Ser, Thr, and Cys residues are close together in the chain they can destabilise an alpha-helix for steric reasons.
- Proline cannot rotate about the N-C_α bond and so it introduces a destabilsing kink in the structure. In addition, it cannot participate in hydrogen bonding as it does not have a hydrogen bonded to nitrogen.

Proteins – Secondary structure – Beta sheet

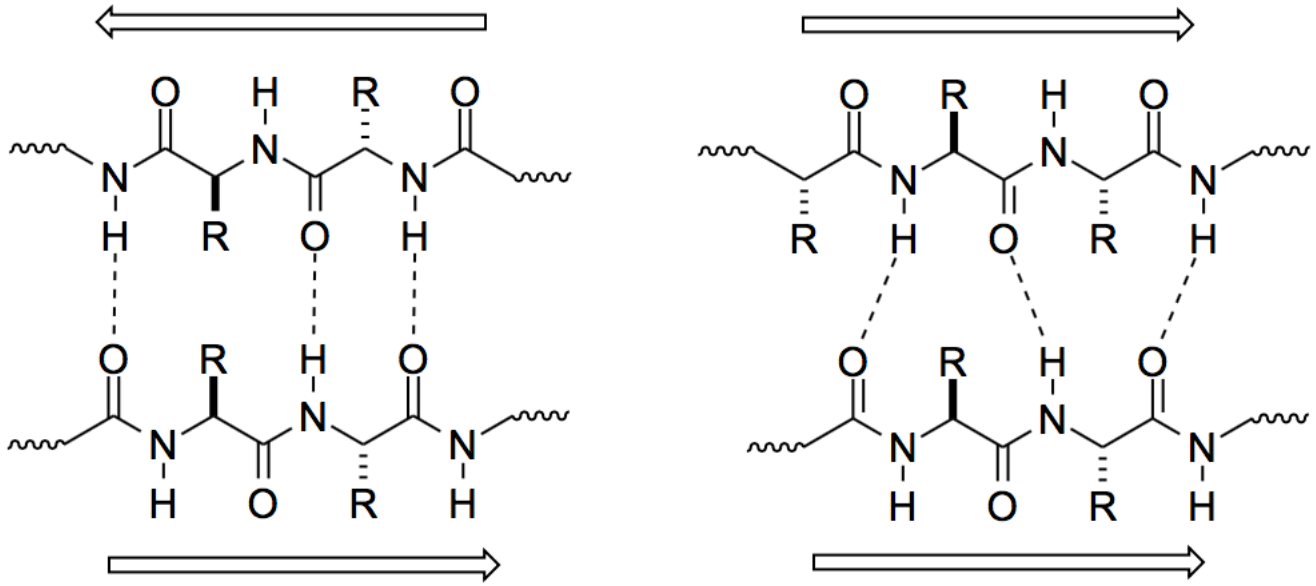




- Beta sheets are held in place by hydrogen bonding as two chains line up with the amide oxygen atoms opposite the amide hydrogen atoms.
- The side chains are situated at right angles to the sheets.



Proteins – Secondary structure – Beta sheet

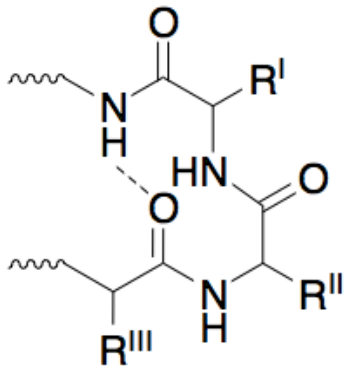


Hydrogen bonding in antiparallel and parallel β-sheets (the arrows are pointing to the *C*-terminal end of the chain).

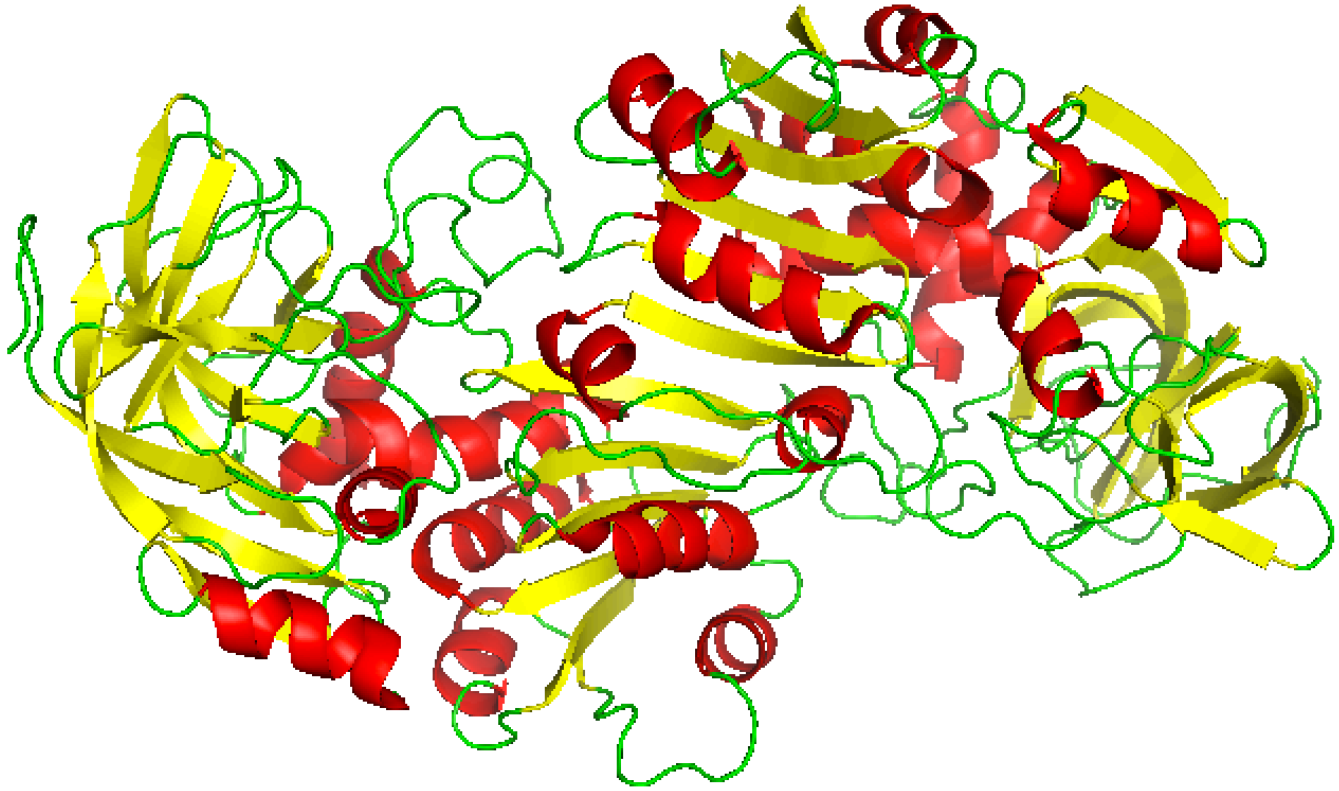
- Beta sheets can form within a single chain if it loops back on itself.
- The chains can run in the same direction (parallel) or in opposite directions (antiparallel).

Proteins – Secondary structure – Beta turn

- A β-turn allows the polypeptide chain to turn abruptly and go in the opposite direction.
- This allows the protein to adopt a more globular compact shape.
- The turn is stabilised by a hydrogen bonding interaction between the hydrogen of the first peptide bond and the carbonyl oxygen of the third peptide bond.
- Residue 2 of a β-turn is often proline as the *cis* conformation accommodates the turn.



Proteins – Secondary structure



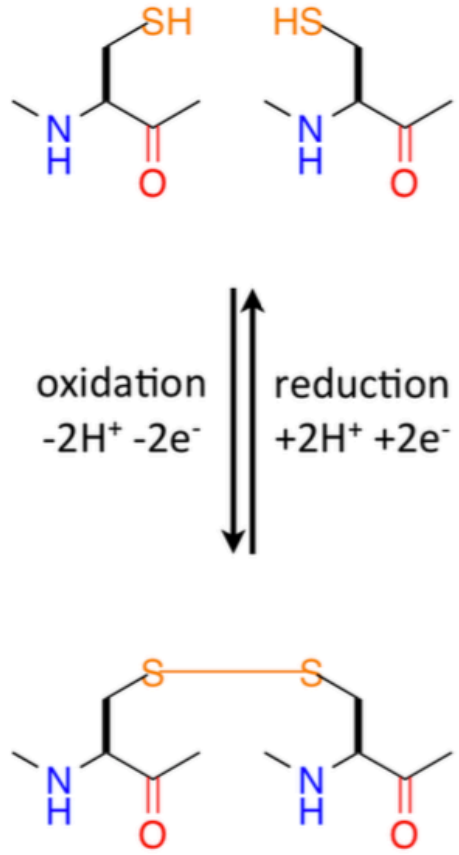
Horse liver alcohol dehydrogenase
(alpha helices in red and beta sheets in yellow)

Proteins – Tertiary structure

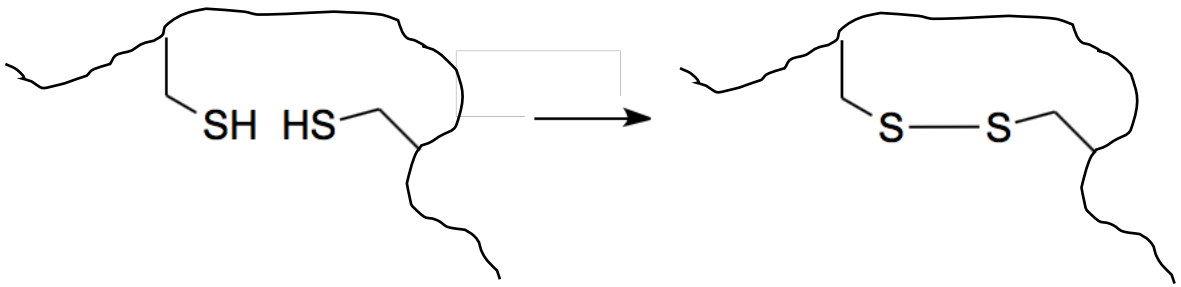
- The secondary structure refers to regions of ordered structure in the protein.
- The tertiary structure refers to the overall shape of the protein.
- It results from interactions between side chains.
- Interactions include covalent disulfide bonds, ionic interactions, hydrogen bonds and Van der Waals interactions.

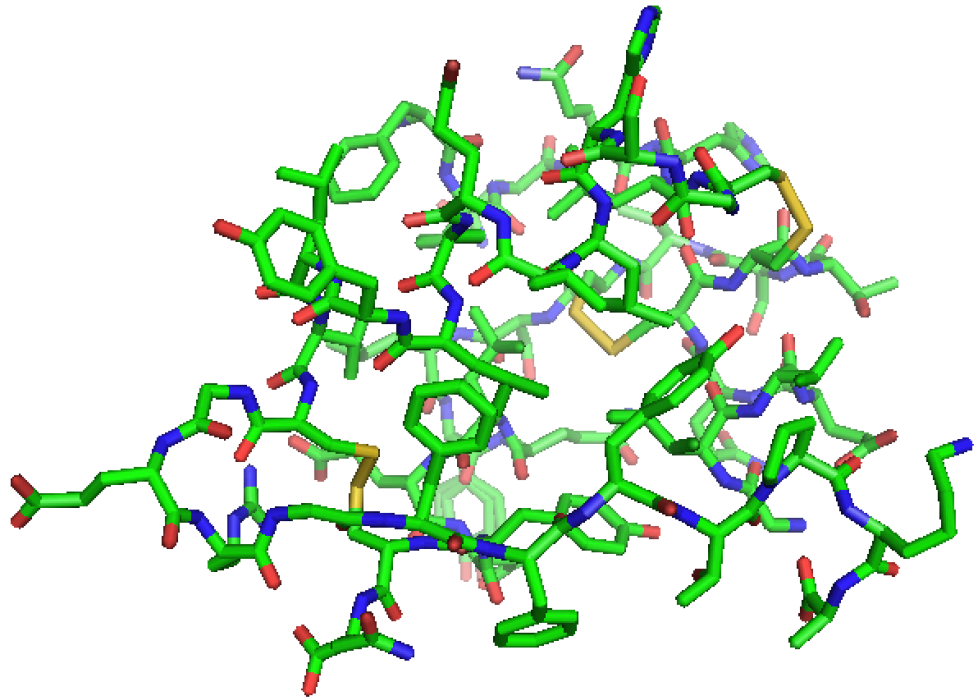
Proteins – Tertiary structure – Covalent disulfide bonds

- Thiols do **not** form strong hydrogen bonds as sulfur has the same electronegativity as carbon (the C-S bond is not polar).
- The thiol is the most readily oxidised of all the side chains.
- When two thiols are oxidised, a disulfide bond results.
- In strongly reducing conditions, disulfide bonds can be broken.



Proteins – Tertiary structure – Covalent disulfide bonds

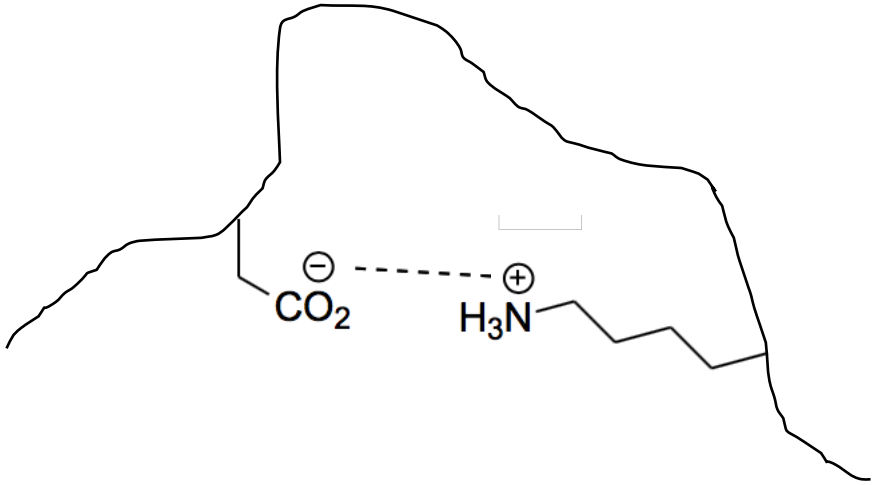




- The side chain of cysteine contains a thiol group.
- When two cysteine residues are close together, a covalent disulfide bond, known as a cystine, can be formed by oxidation.

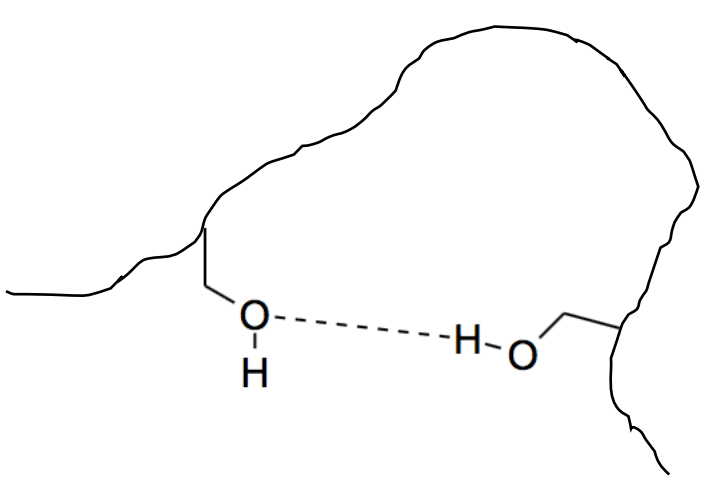
- The side chain of human insulin is made up of two polypeptide chains linked covalently by three disulfide bonds (shown in yellow).

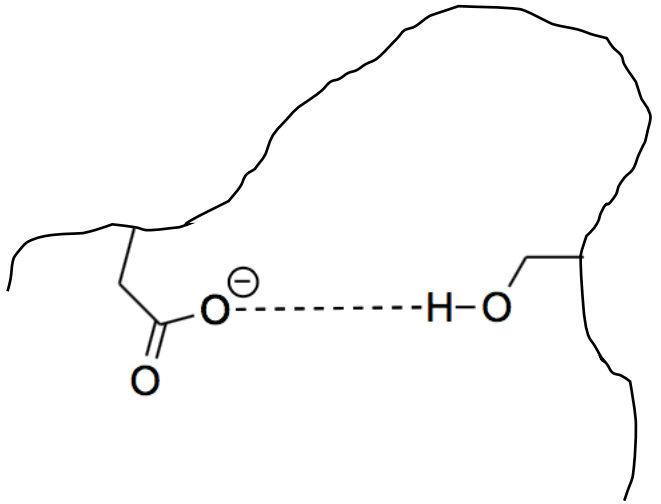
Proteins – Tertiary structure – Ionic interactions



- An ionic bond can be formed between the carboxylate anion of an acidic amino acid residue (aspartic acid, glutamic acid) and the ammonium ion of a basic residue (lysine, arginine or histidine).
- Unlike hydrogen bonding interactions, ionic interactions are non-directional (purely electrostatic interactions).

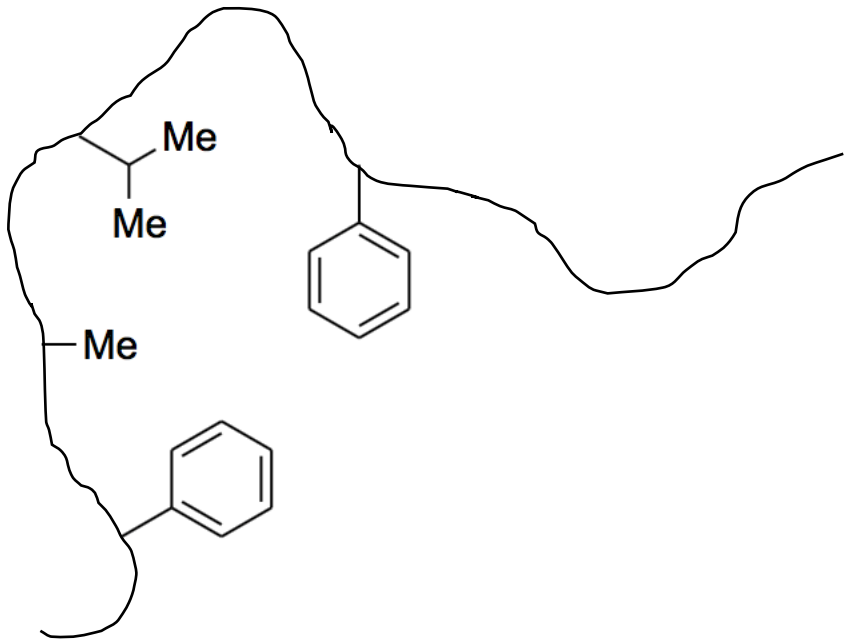
Proteins – Tertiary structure – Hydrogen bonding





Hydrogen bonds can form between a large number of amino acid side chains, such as serine, threonine, aspartic acid, glutamic acid, glutamine, lysine, arginine, histidine, tryptophan, tyrosine, and asparagine.

Proteins – Tertiary structure – Van der Waals interactions



- Van der waals interactions are weaker than hydrogen bonds.
- They take place between hydrophobic groups on the protein (e.g the side chains of alanine, valine, leucine, isoleucine, phenylalanine, and proline).
- Some of the polar side chains also have hydrophobic character (methionine, tryptophan, threonine, and tyrosine).

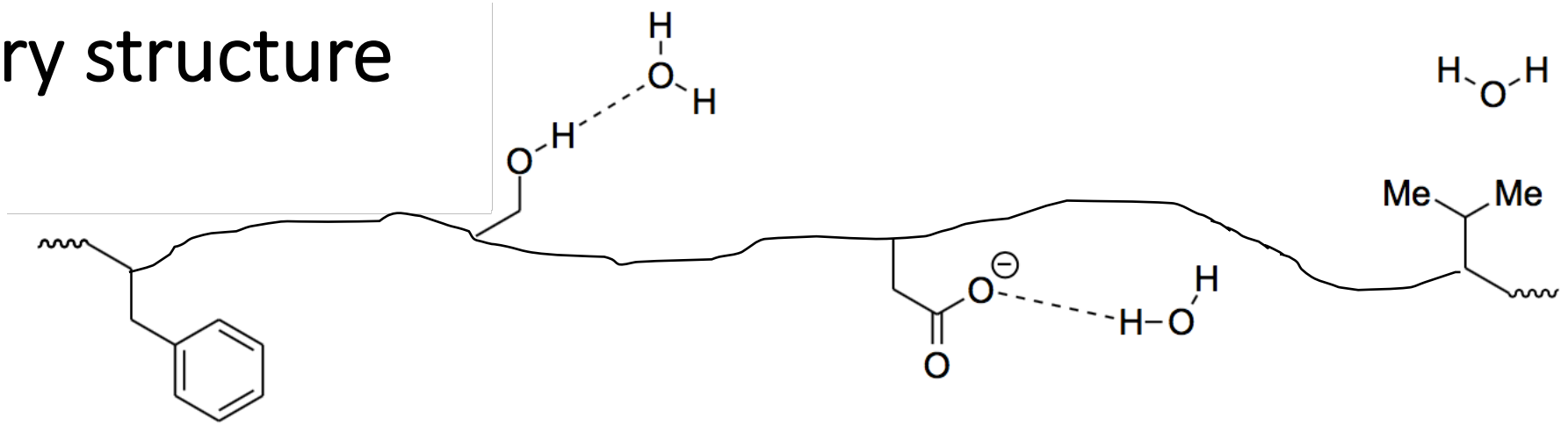
Proteins – Tertiary structure – Van der Waals interactions

- Van der Waals interactions arise as a result of **London forces** and the **hydrophobic effect**.
- **London forces** refer to the weak interaction between temporary dipoles. The electronic distribution in neutral, non-polar regions is never totally symmetrical, leading to temporary dipoles. These dipoles induce dipoles in neighbouring regions, leading to a weak attractive interaction.
- **The hydrophobic effect** refers to the fact that water cannot solvate non-polar regions. Protein folding minimises contact between non-polar regions and the aqueous environment.

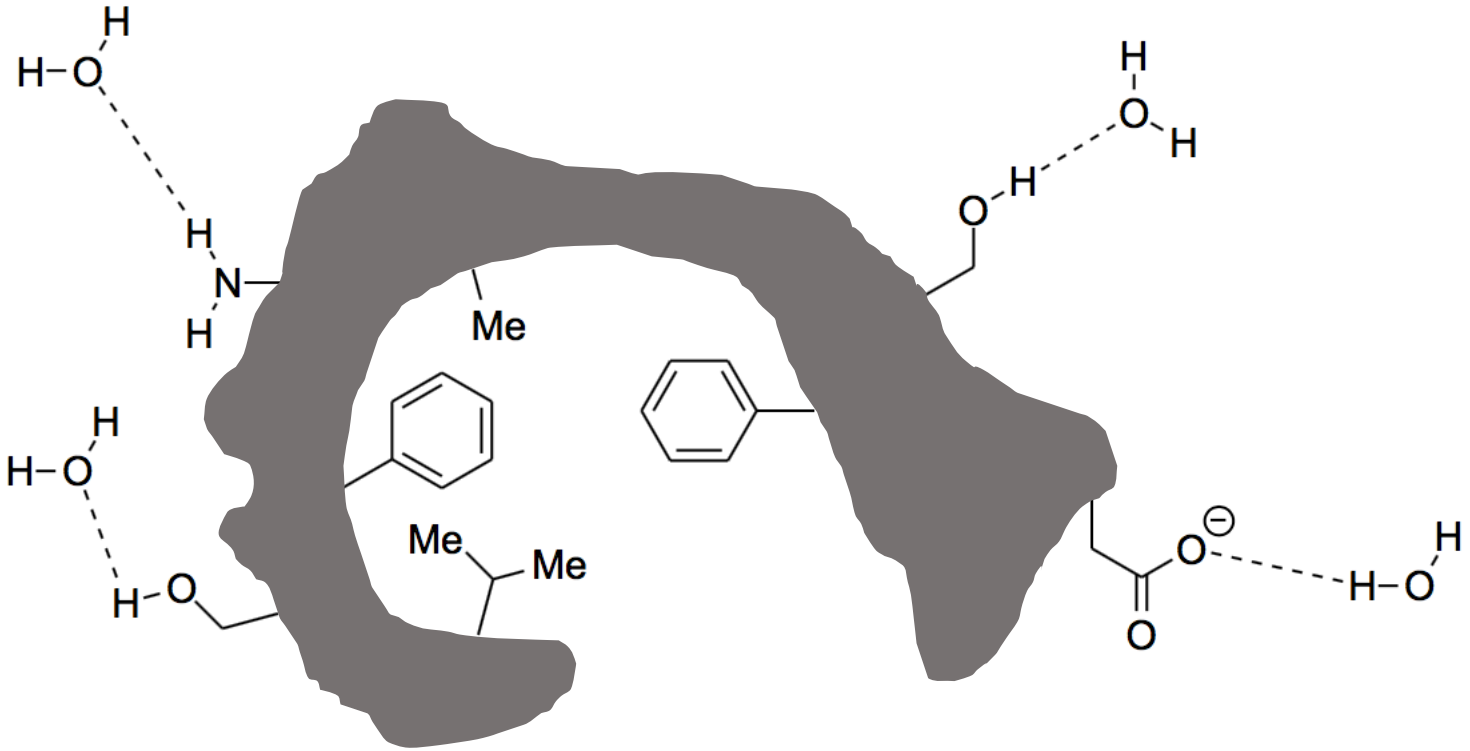
Proteins – Tertiary structure

- The relative strength of interactions between R groups is: disulfide bonds (240 – 260 kJ mol^{-1}) > ionic interactions (20 – 40 kJ mol^{-1}) > hydrogen bonds (15 – 30 kJ mol^{-1}) > Van der Waals interactions (2 – 4 kJ mol^{-1})
- However, the most important interactions in determining tertiary structure are Van der Waals interactions and hydrogen bonds.
- This is a statistical phenomena. The only amino acid which can form a disulfide bond is cysteine whereas many amino acids can interact via hydrogen bonding or Van der Waals interactions.

Proteins – Tertiary structure



Protein folding



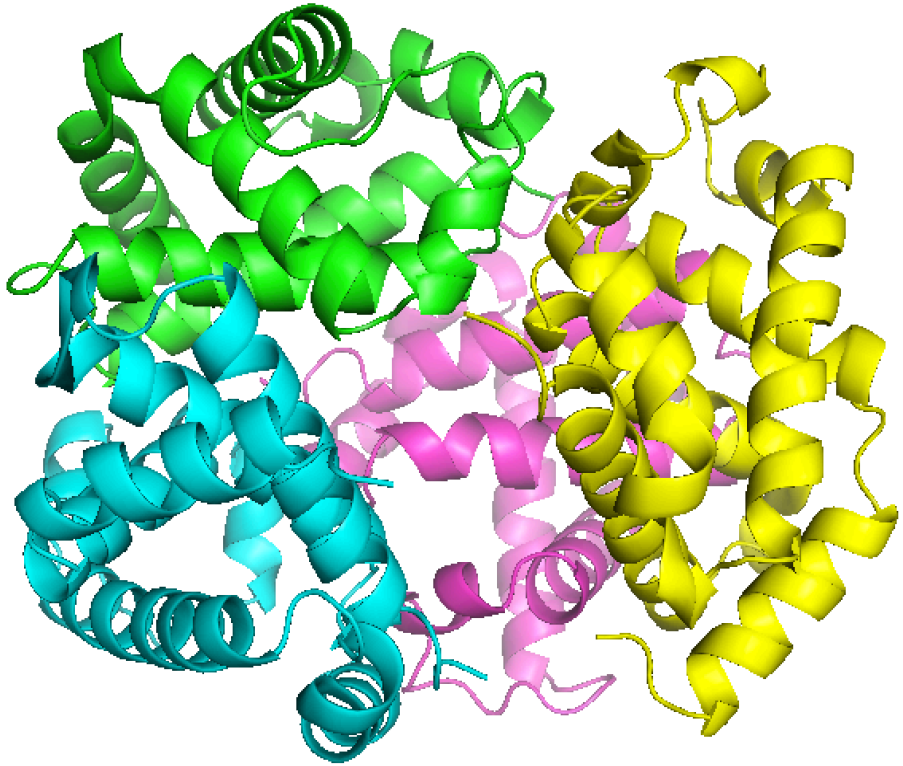
- Proteins do not exist in a vacuum, they are surrounded by water.
- The tertiary structure ensures most hydrophilic groups are on the surface and hydrophobic groups are in the centre.

Proteins – Tertiary structure

- The tertiary structure explains why enzymes (nature's catalysts) are able to catalyse reactions that would otherwise be impossible in an aqueous environment.
- Enzymes contain a hollow on the surface, the active site, which tends to be hydrophobic and provides a non-aqueous environment for the reaction to take place.

Proteins – Quartenary structure

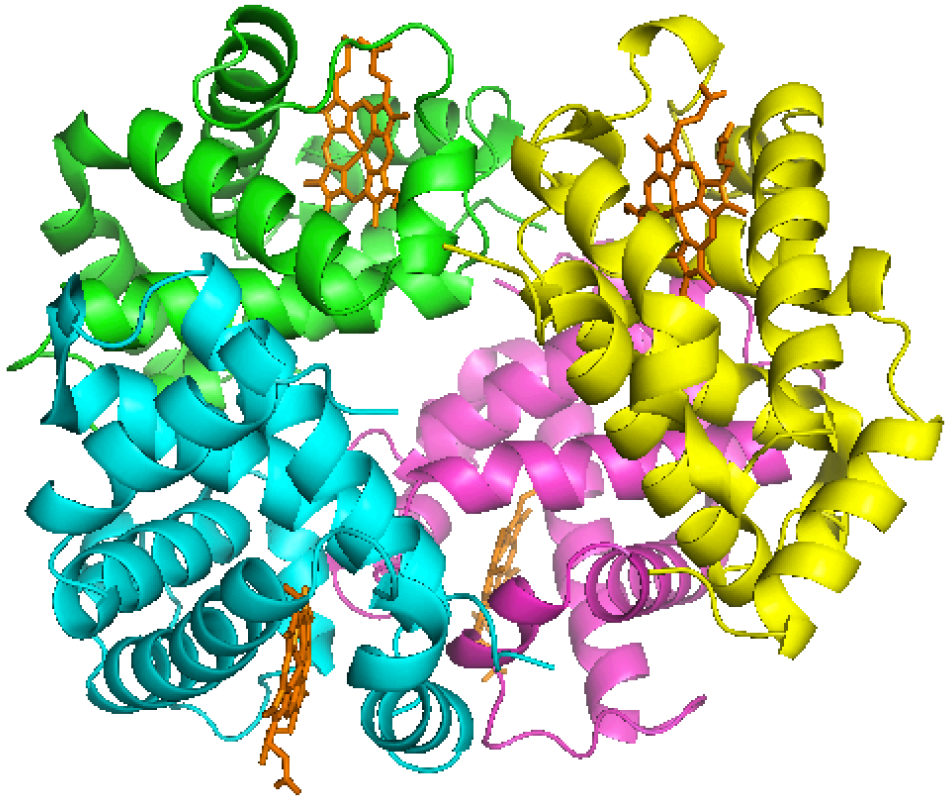
- Some molecules, such as haemoglobin adopt a **quaternary structure** in which two or more polypeptide chains combine to form a larger assembly.
- The quaternary structure refers to the way in which the chains associate with each other.
- Ionic bonding has an important role to play but Van der Waals interactions between exposed hydrophobic groups also facilitate the coming together of different chains.



Haemoglobin is made up of four protein units

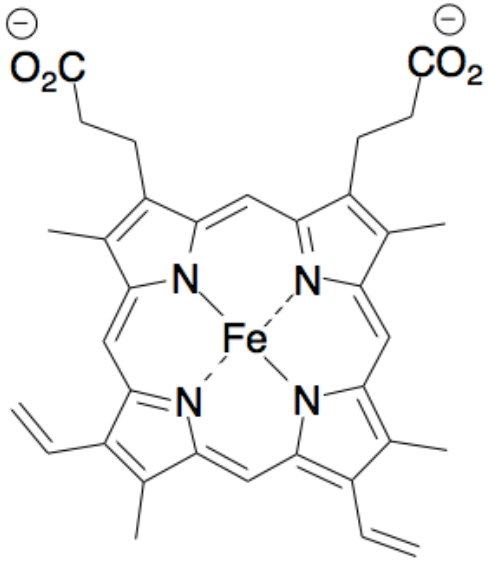
Haemoglobin – A transport protein

- **Transport proteins** are used to transport molecules or ions from one place to another.
- Haemoglobin transports molecular oxygen from the lungs to all the tissues of the body.
- Each protein unit of haemoglobin is bound to a non protein unit (prosthetic group).
- The prosthetic group in haemoglobin is haeme, shown in orange.

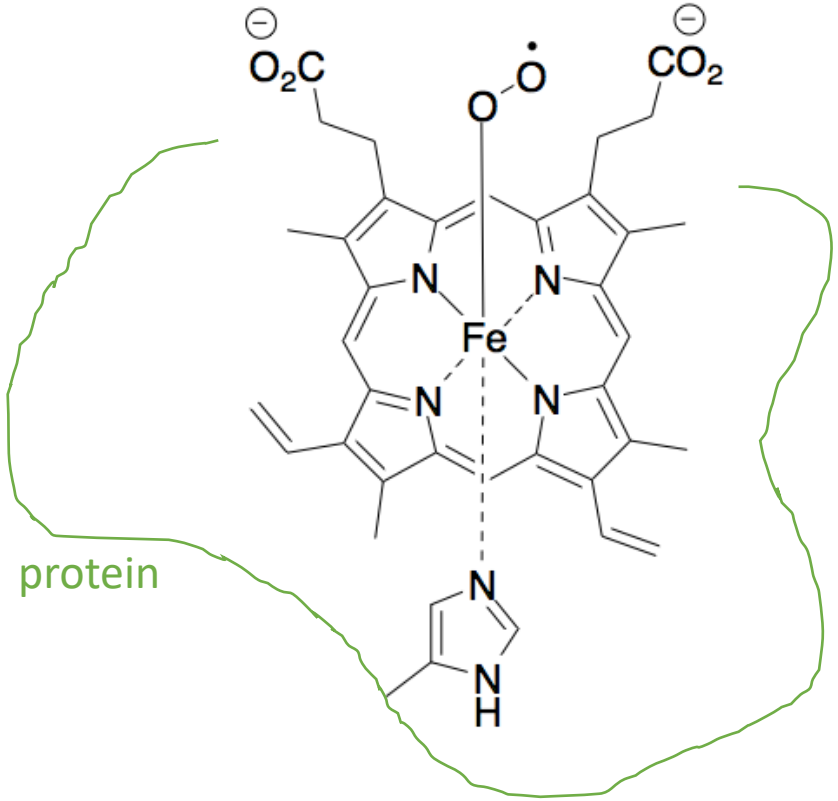


Haemoglobin – A transport protein

- Haeme is a porphyrin. The hydrophilic carboxylate groups interact with the surrounding environment.
- The majority of the haeme is embedded in a hydrophobic cleft in the protein (made up of amino acids such as leucine and valine).



Haem

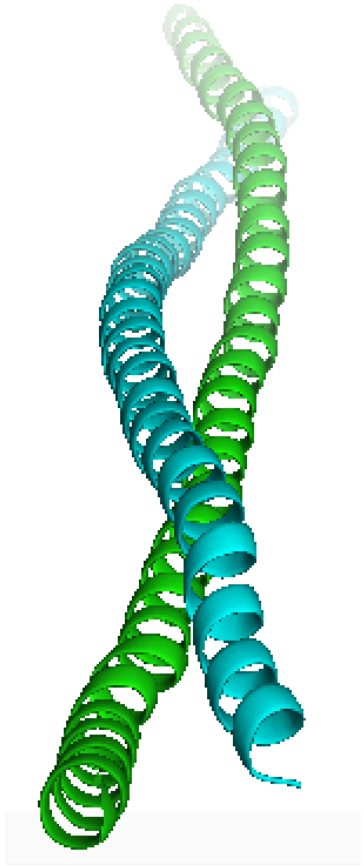


Oxyheamoglobin

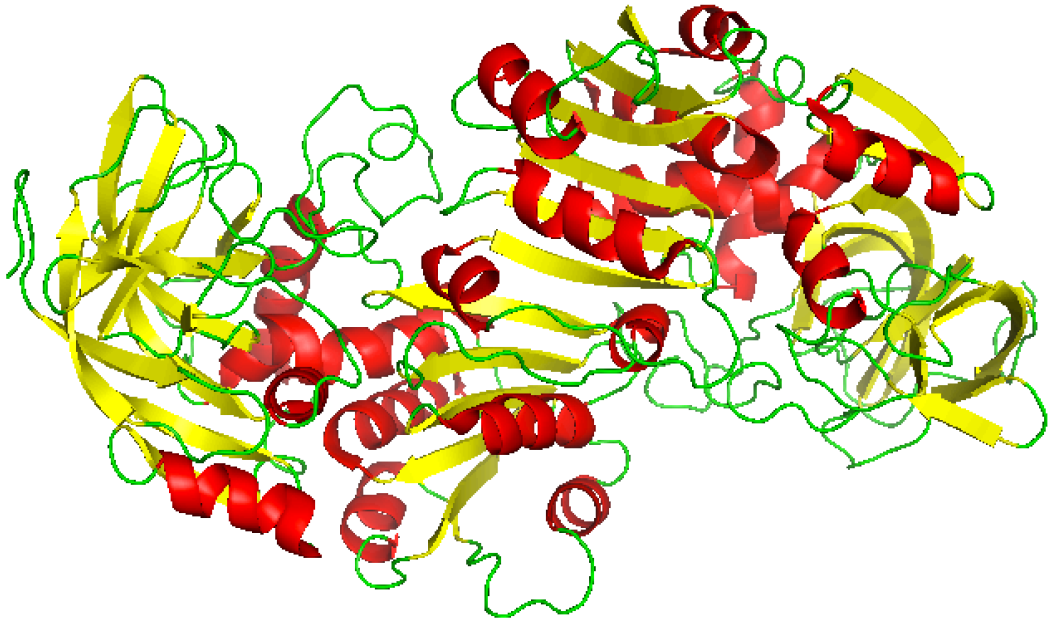
Fibrous and Globular Proteins

- When considering a proteins overall shape (tertiary and quaternary structure) proteins are often classified as either **fibrous** or **globular**.
- **Fibrous proteins** consist of polypeptide chains arranged in long strands or sheets.
- **Globular proteins**, consist of polypeptide chains folded into a spherical or globular shape.

Fibrous and Globular Proteins



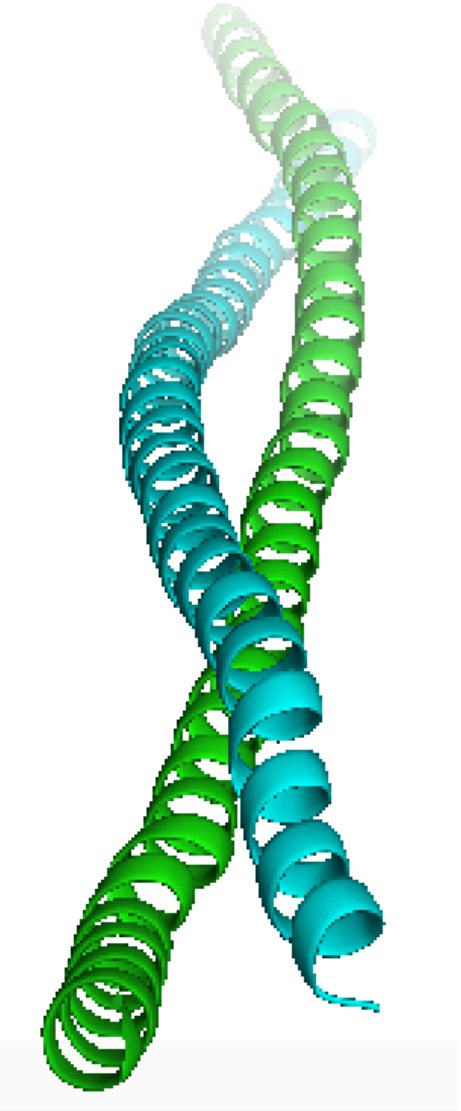
Fibrous proteins often consist of mostly one type of secondary structure and their tertiary structure is often relatively simple



Globular proteins often contain several types of secondary structure.

Fibrous proteins – α-Keratin

- α-Keratin is mechanically durable (strong)
- It is a relatively unreactive protein that is found in all higher vertebrates.
- It is the main component of hair, wool, nails, claws, horns and the outer layer of skin.
- α-Keratin has a rope-like structure (a coiled coil). Two polypeptide strands are wrapped about each other to form a supertwisted coiled coil. Much like in rope, the supertwisting strengthens the overall structure.
- The intertwining of the two polypeptides is another example of quaternary structure.

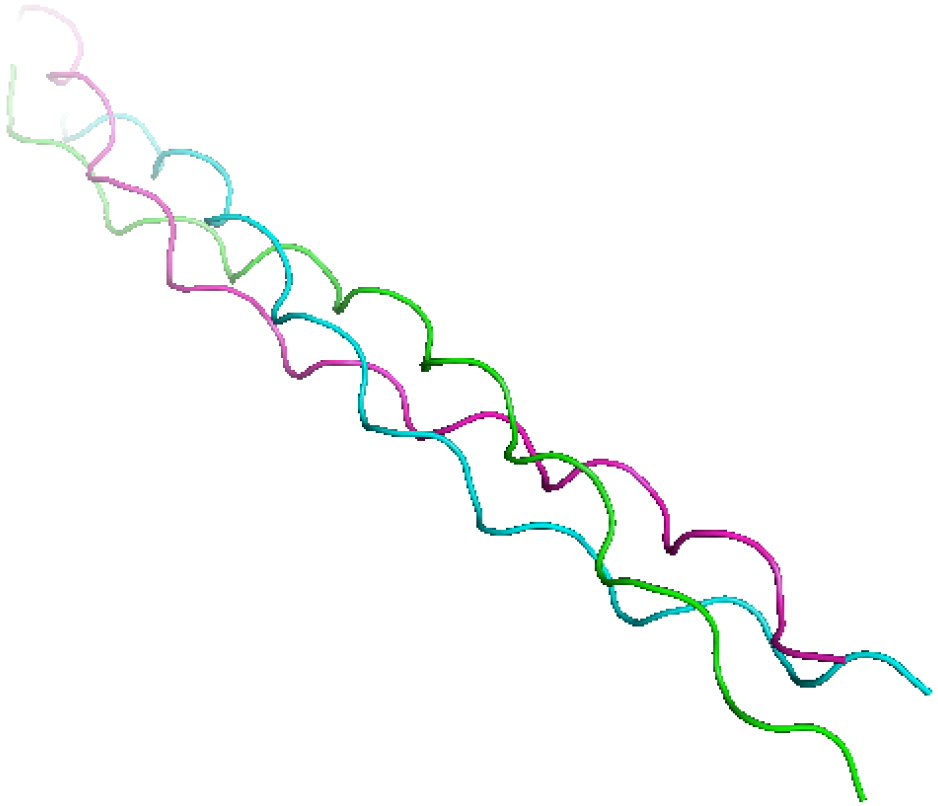


Fibrous proteins – α-Keratin

- α-keratin is rich in the hydrophobic residues Ala, Val, Leu, Ile, Met, and Phe and the surfaces where the two helices are in close proximity are held together by Van der Waals interactions.
- The quaternary structure of α-keratin is also stabilised by disulfide bonds which cross-link adjacent polypeptide chains.
- The hardest, strongest α-keratins, (e.g. rhinoceros horn) contain up to 18% cysteine residues.

Fibrous proteins – Collagen

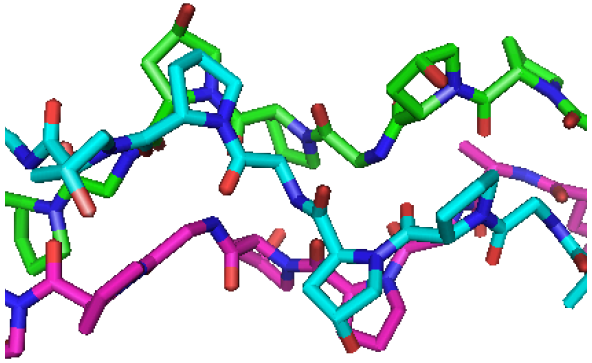
- Collagen is also mechanically durable (strong).
- Its strength is put to structural use in bone, teeth, cartilage and tendon.
- A single collagen molecule consists of three polypeptide chains arranged as a triple helix.



Fibrous proteins – Collagen

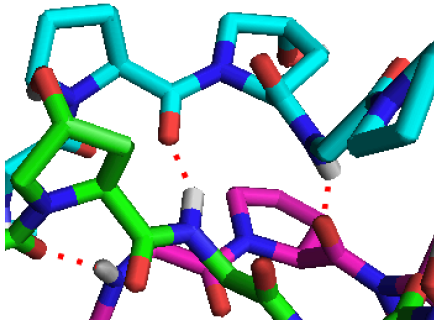
- Approximately a third of collagen's amino acid residues are glycine and again about a third are either proline or hydroxyprolyl residues.
- Synthesis of hydroxyprolyl residues requires vitamin C and without it collagen cannot form. This is why vitamin C deficiency causes scurvy.

Fibrous proteins – Collagen



- The amino acid sequence of a collagen polypeptide typically consists of repeating triplets of sequence Gly-X-Y where X is often proline and Y is often hydroxyprolyl.
- The proline residues prevent the polypeptide from forming an alpha-helix.
- The glycine residues are the only residues small enough to fit in the centre of the triple helix (each polypeptide adopts a helical conformation with three residues per turn).

Fibrous proteins – Collagen

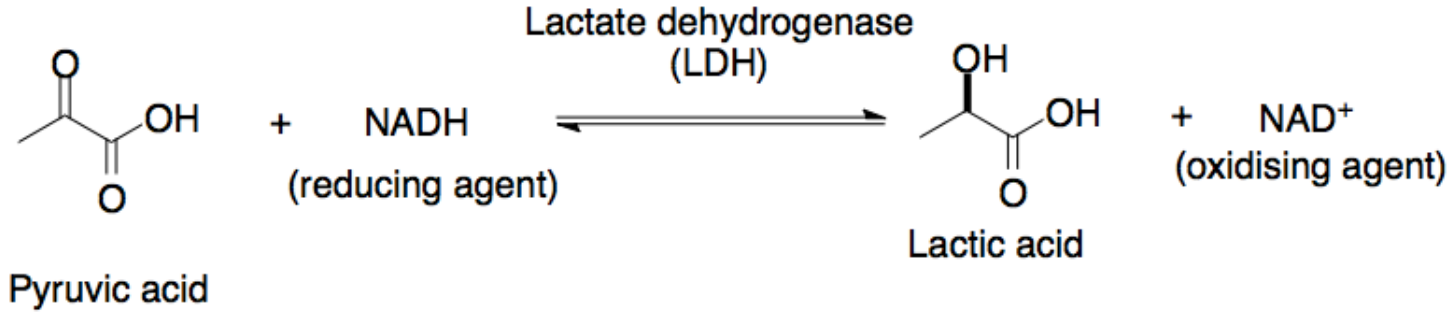


- The three polypeptide chains are oriented such that the N-H of each glycine makes a strong hydrogen bond with the carbonyl oxygen of a proline residue on one of the other polypeptide chains.
- The proline and hydroxyprolyl residues provide rigidity to the entire assembly.

Enzymes

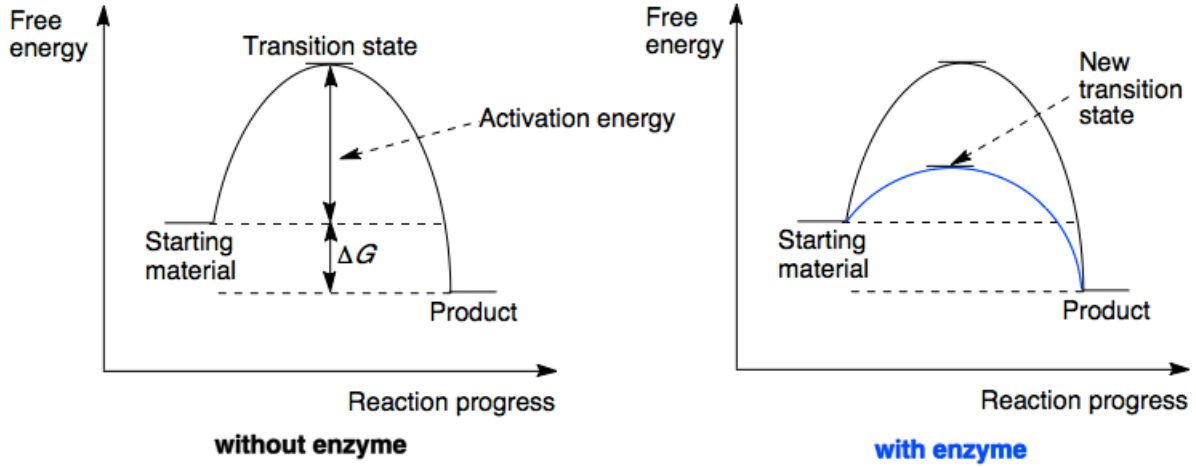
Enzymes

- Enzymes are proteins which act as nature's catalysts.
- Catalysts increase the rate of a reaction without being used up in the reaction.
- Catalysts increase the rate of reaction in both directions. They speed up the approach to equilibrium.
- The reduction of pyruvic acid to lactic acid, which takes place when muscles are over-worked, is catalysed by the enzyme lactate dehydrogenase.



Enzymes

- Catalysts increase the rate of a reaction by reducing the activation energy.
- An enzyme lowers the activation energy by stabilising (lowering the energy) of the transition state.
- ΔG is unaffected by a change in the energy of the transition state and so the equilibrium position (ratio of substrate to product) is unaffected.



Enzymes

- Substrates bind to, and react at, a specific site on the enzyme called the active site.
- The active site is found on or near the surface of the enzyme so that it is accessible for the substrate.
- It is often a groove or a hollow on the surface of the enzyme.
- It is often more hydrophobic in nature than the surface of the enzyme, allowing reactions to occur that would otherwise be difficult or impossible in an aqueous environment.

Enzymes

The amino acids in the active site allow for catalysis either by:

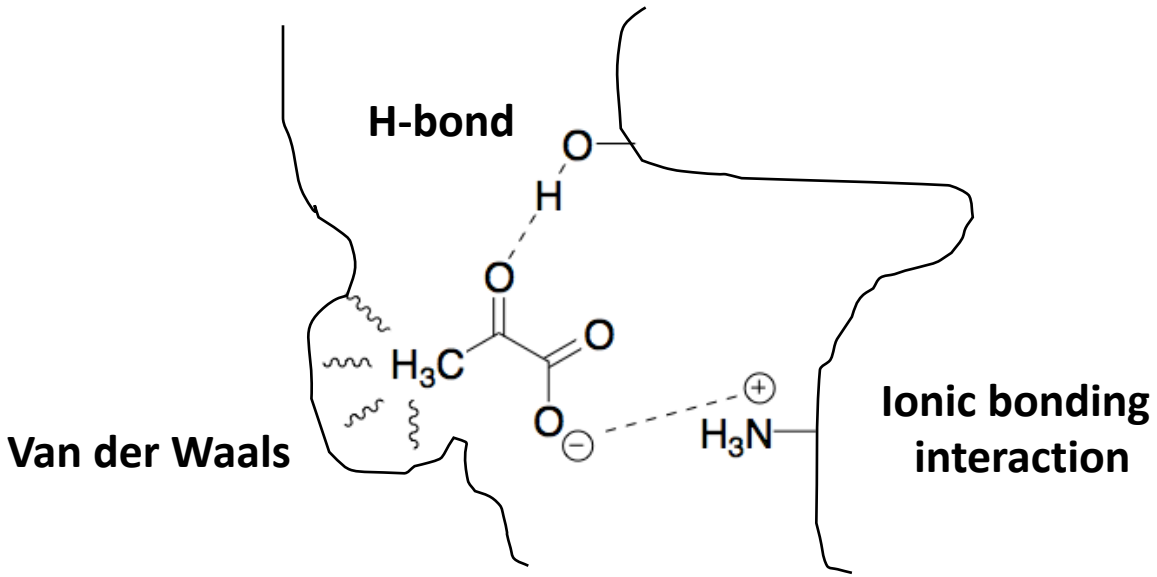
1. binding the substrate or a cofactor (additional non-protein substances) to the active site.
2. taking part in the mechanism of the reaction.

Enzymes – 1

- Intermolecular forces (ionic interactions, hydrogen bonds, dipole–dipole and ion–dipole interactions, London forces and hydrophobic interactions) bind the substrate to the enzyme in the active site.
- Ionic bonding interactions are often particularly important in substrate binding.

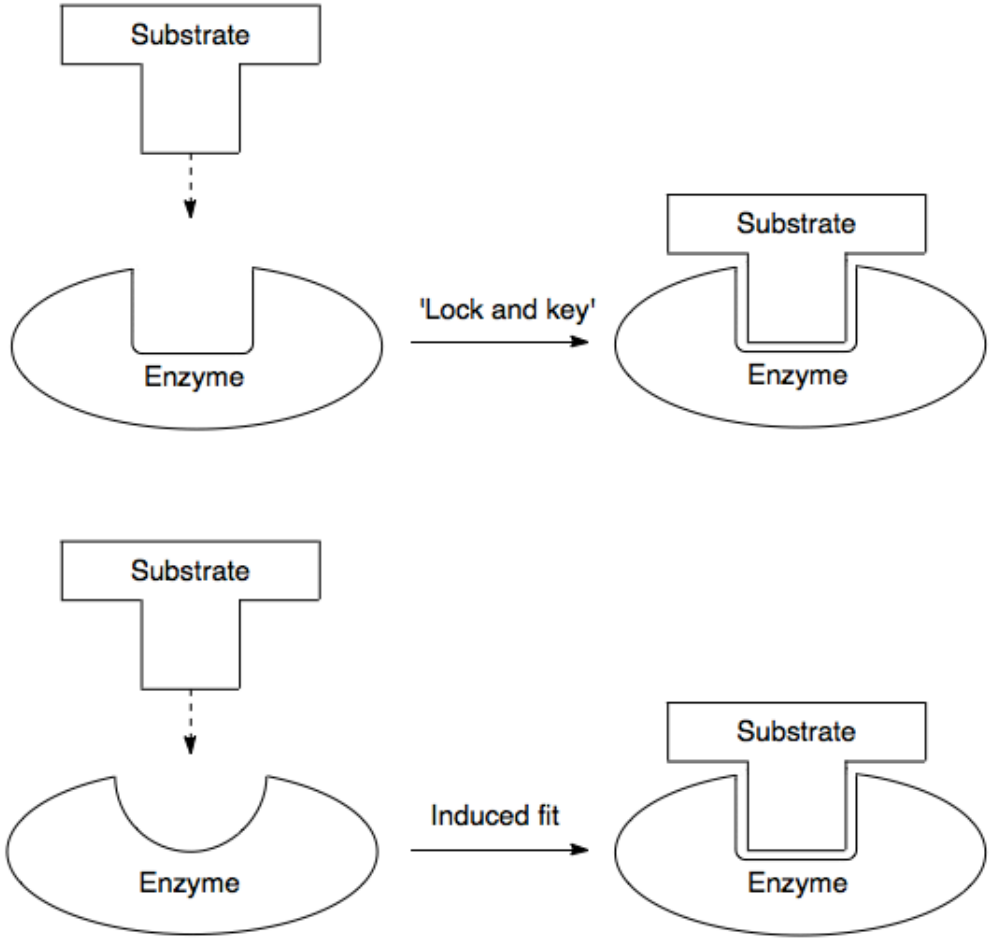
Enzymes – 1

- We can often determine the most likely binding interactions between a substrate and the active site from the structure of the substrate.
- Given the structure of pyruvic acid, it is reasonable to postulate that it binds to lactate dehydrogenase *via* Van der Waals, hydrogen bonding and ionic bonding interactions.



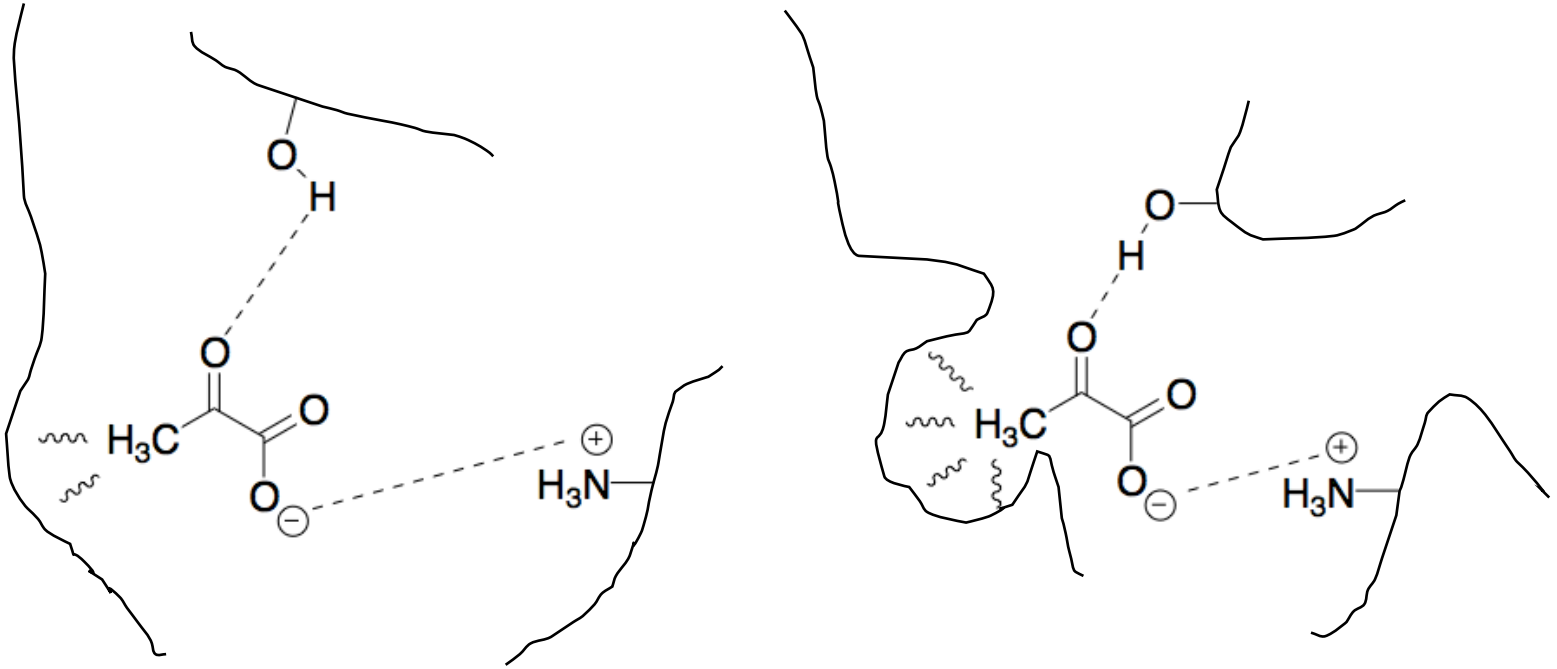
Enzymes – 1

- It was once thought that a substrate fitted the active site of an enzyme in much the same way a key fits a lock (lock and key hypothesis).
- It is now thought an enzyme changes shape slightly to accommodate the substrate (theory of induced fit). This explains how some enzymes can catalyse a reaction on a range of substrates.



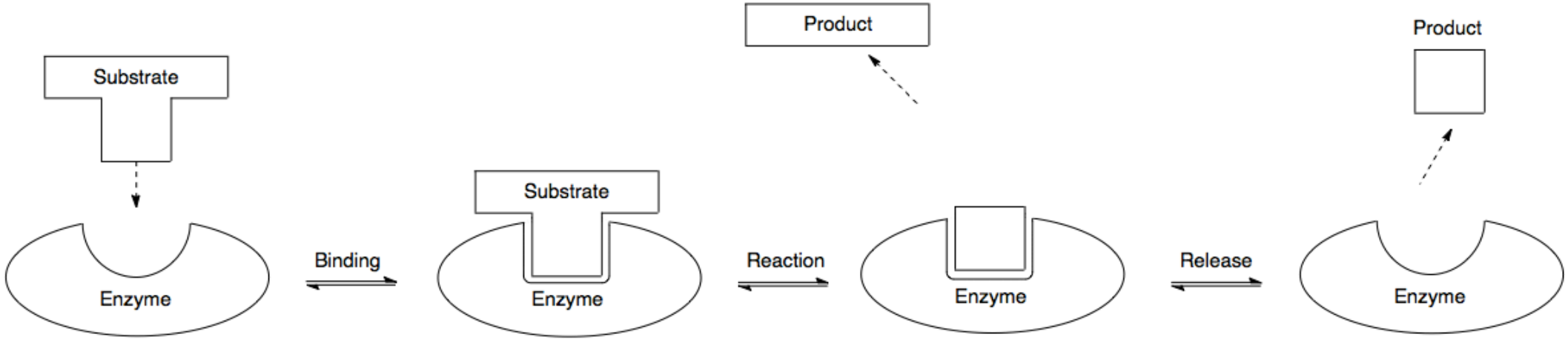
Enzymes – 1

- In the induced fit model, the enzyme changes shape in order to maximise the strength of the binding interactions.
- Each substrate induces the active site into a shape that is ideal for the structure of the substrate.



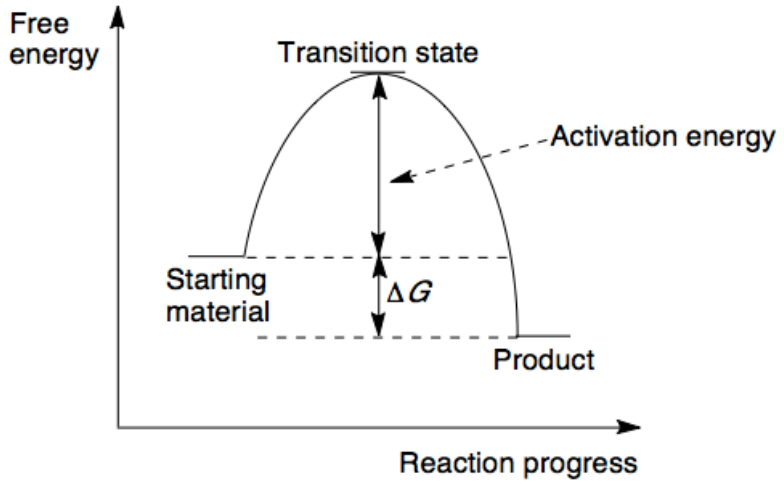
Enzymes – 1

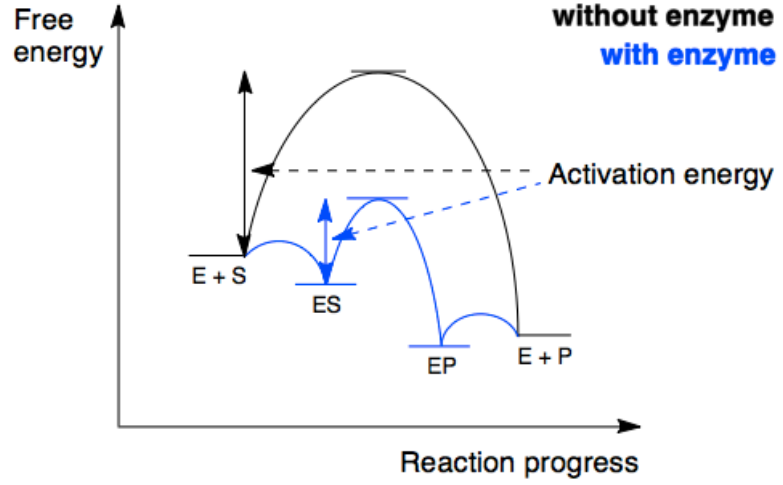
- The range of substrates that can bind to the active site of a particular enzyme is still dependent on the substrates having the appropriate size and binding groups in appropriate positions.
- The substrate can also adapt to fit the active site by changing its conformation. Bonds in the substrate may also be weakened.



Enzymes – 1

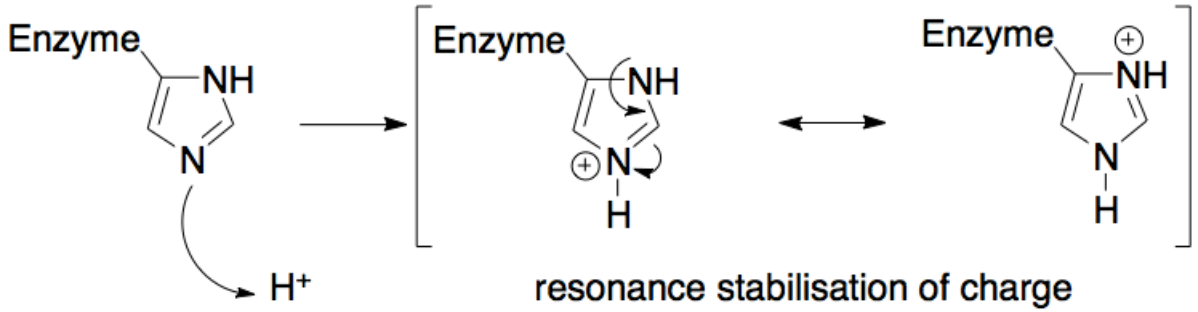
- The binding interaction with the product must not be so strong that the product does not leave the active site of the enzyme after the reaction, preventing the enzyme from catalysing another reaction.
- The enzyme also binds to the transition state for the reaction. The binding interactions with the transition state must be stronger than with the substrate so that the energy of the transition state is lowered more compared to the substrate thus lowering the activation energy.





Enzymes – 2

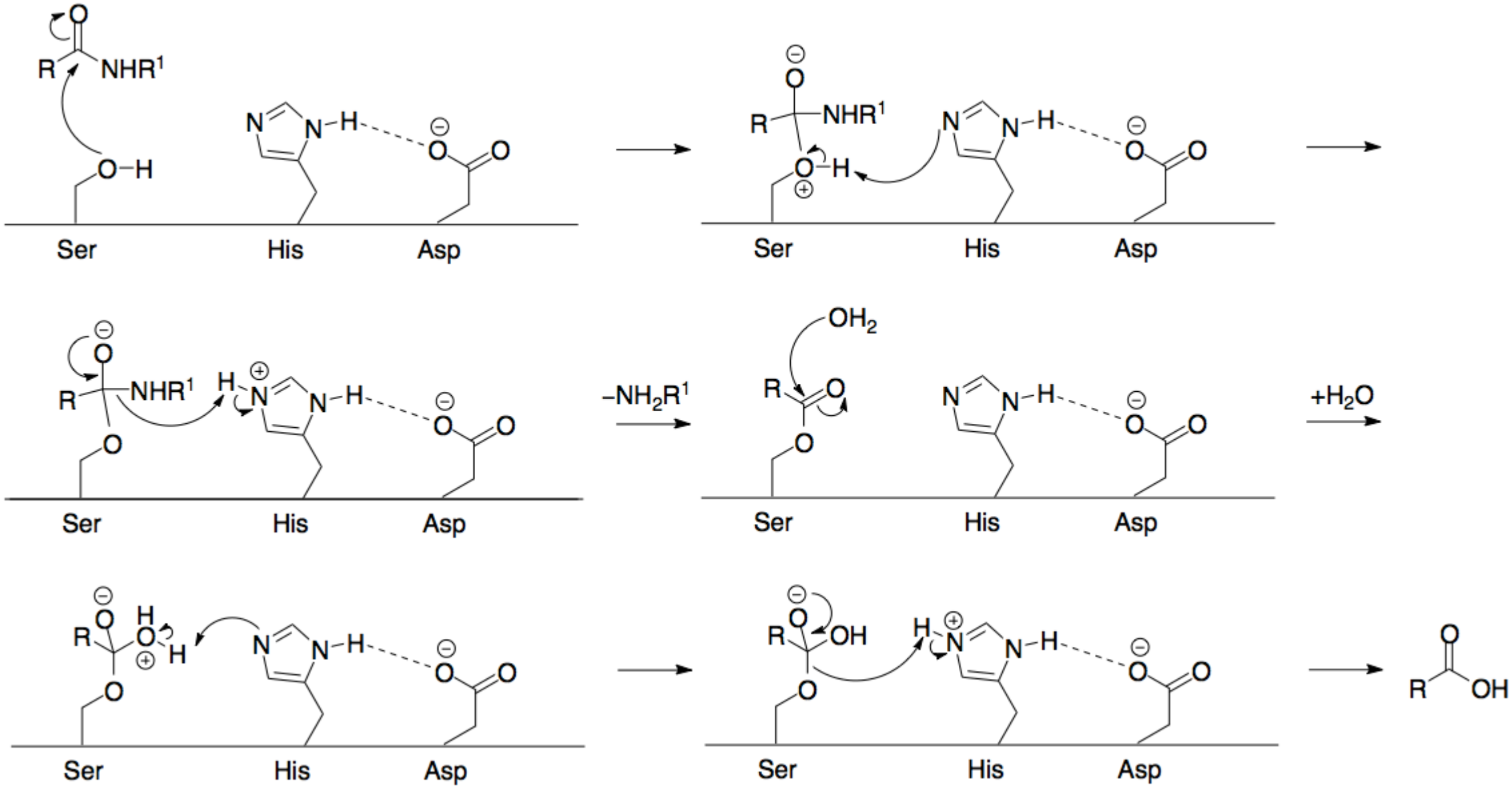
- The amino acids in the active site may take part in the mechanism of the reaction.
- Acid/base catalysis is often provided by the amino acid histidine.
- The imidazole ring of histidine is a weak base and so it exists in equilibrium with its protonated form. It can accept and donate protons during a reaction mechanism.



- Other amino acid residues such as tyrosine, glutamic acid and aspartic acid can also take part in acid/base catalysis.

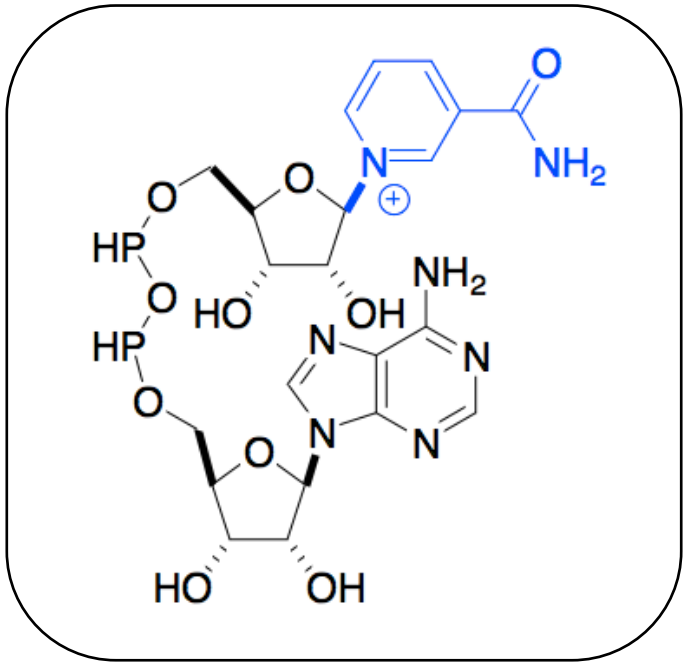
Enzymes – 2

- Serine (-OH) and cysteine (-SH) have nucleophilic residues which can take part in a reaction mechanism.
- Chymotrypsin catalyses the hydrolysis of peptide bonds.
- The mechanism involves a catalytic triad of amino acids (serine, histidine, and aspartic acid).


Ser
His
Asp
$-NH_2R^1$
$+H_2O$

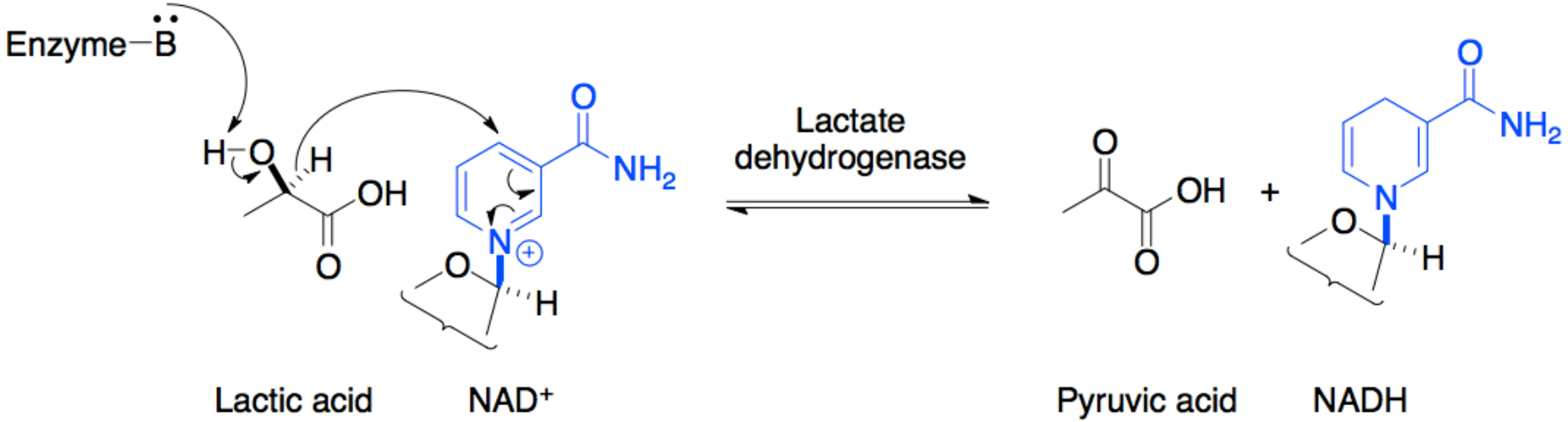
Enzymes - Cofactors

- Cofactors are either metal ions (e.g. zinc) or small organic molecules called coenzymes (e.g. NAD^+, pyridoxal phosphate)
- Most coenzymes are bound to enzymes *via* non-covalent bonding interactions. Some are bound covalently and are called prosthetic groups.
- Cofactors act as helper molecules/ions assisting the enzyme in the catalysis of biochemical reactions.



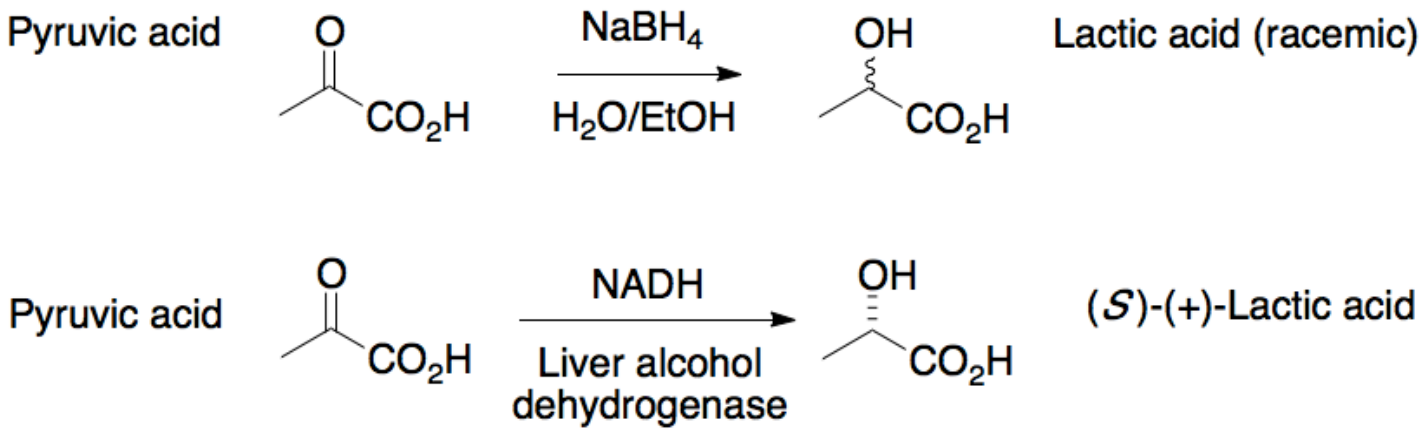
Enzymes – NAD^+/NADH

- Lactate dehydrogenase needs the coenzyme nicotinamide adenine dinucleotide (NAD^+) to catalyse the oxidation of lactic acid to pyruvic acid.
- NAD^+ is bound to the active site along with lactic acid, and the mechanism shows how it acts as the oxidizing agent.



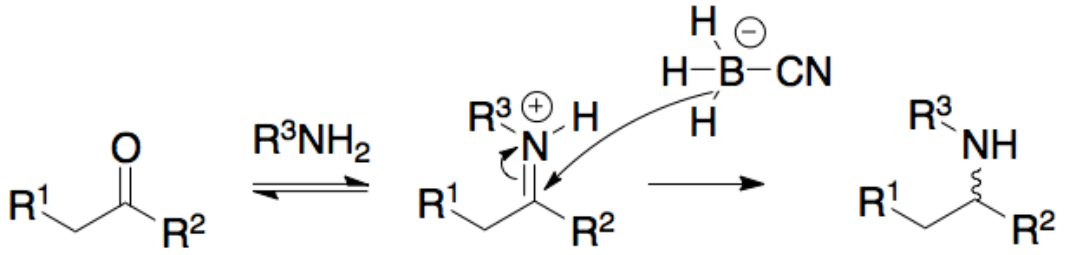
Enzymes – NAD^+/NADH

- The reduction of NAD^+ is reversible, and NADH is itself a reducing agent.
- NADH is nature's $NABH_4$ although the product of an enzymatic reaction is optically pure. The two faces of pyruvic acid's carbonyl group are enantiotopic. The chiral and enantiomerically pure enzyme controls the addition so that it occurs from one face only.

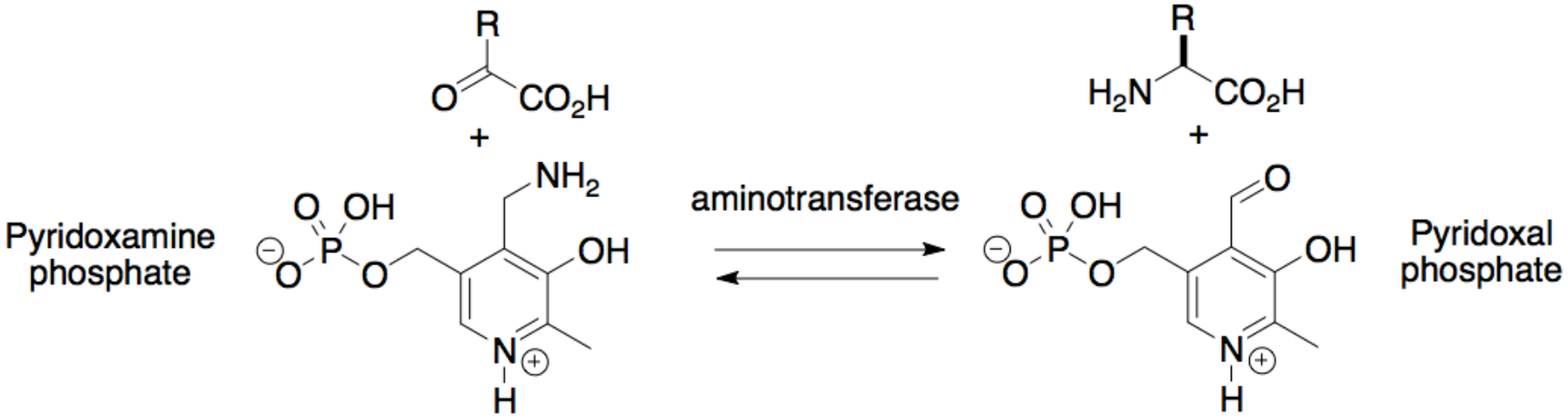


Enzymes – Pyridoxal phosphate/Pyridoxamine phosphate

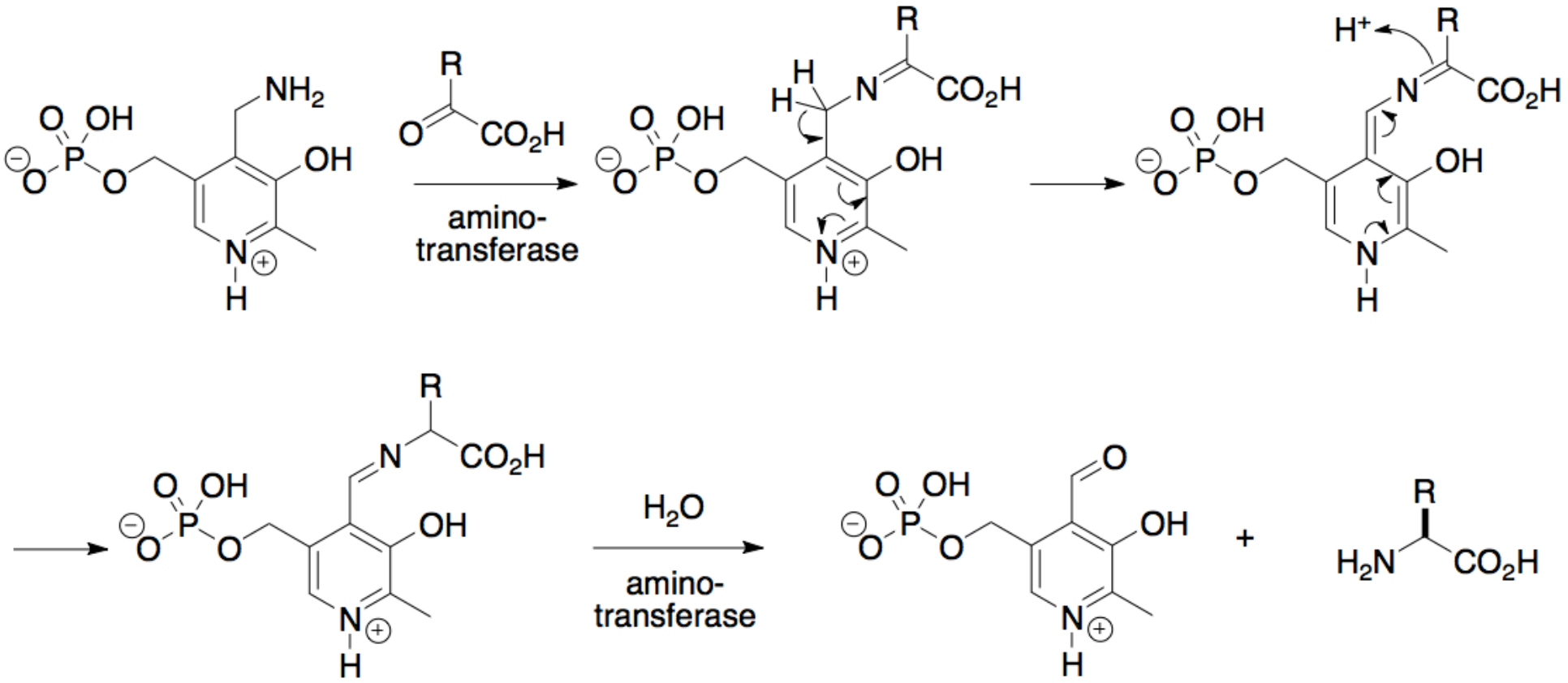
- One of the best methods available to synthetic organic chemists for making amines is reductive amination.



- Nature carries out this reaction enantioselectively, using an aminotransferase enzyme and the pyridoxamine phosphate cofactor.

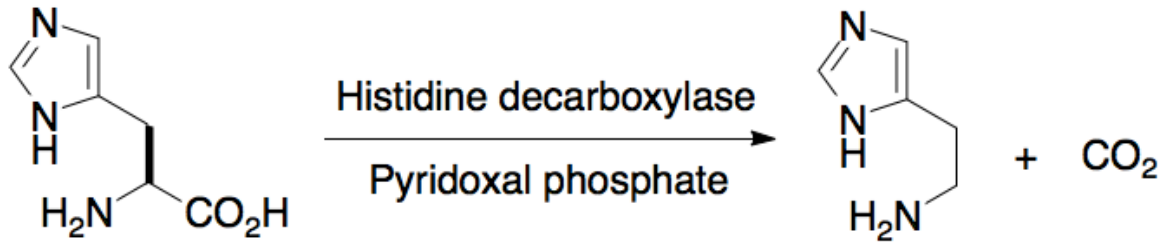


Enzymes – Pyridoxal phosphate/Pyridoxamine phosphate

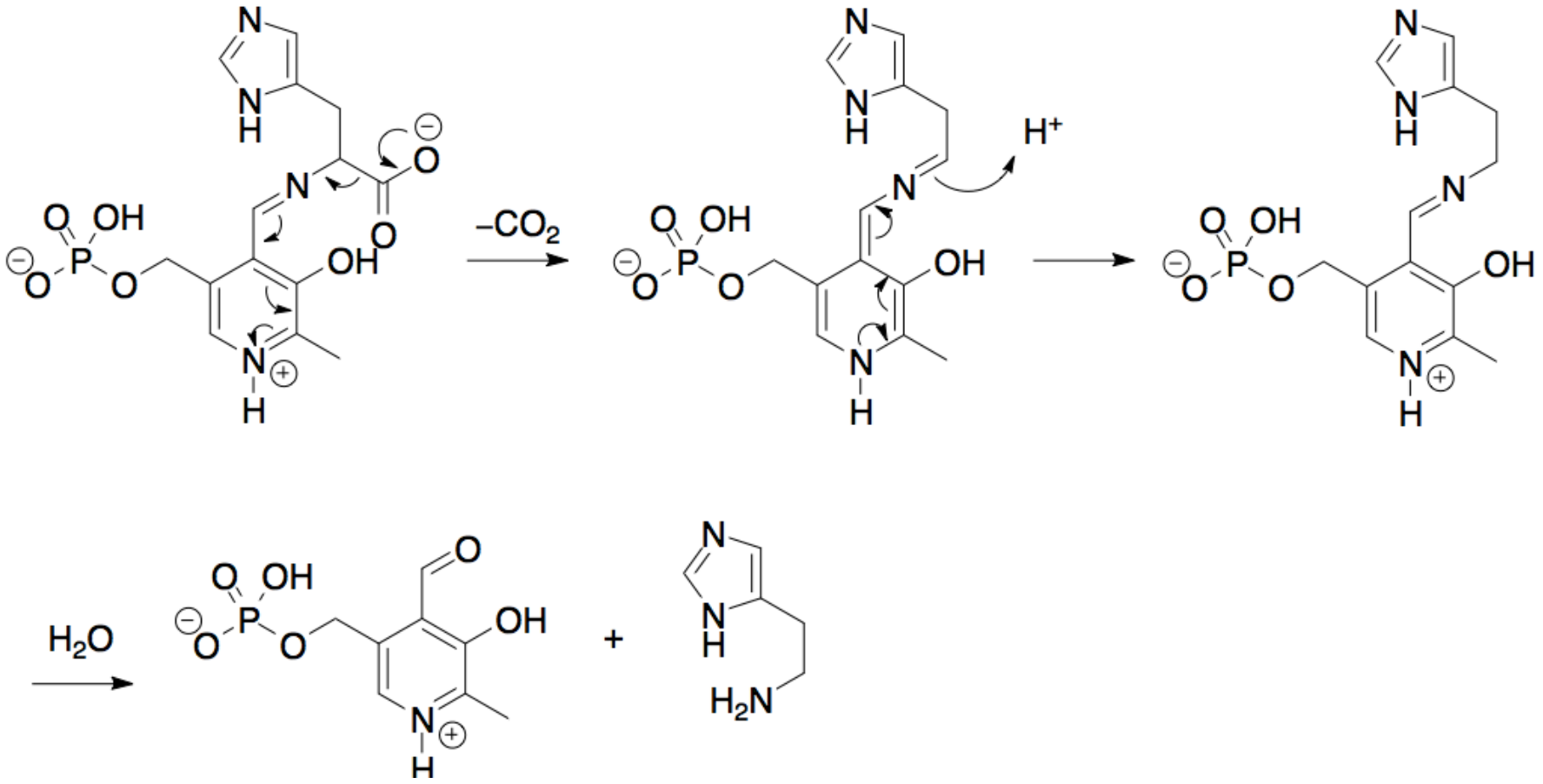


Enzymes – Pyridoxal phosphate/Pyridoxamine phosphate

- Decarboxylase enzymes use the coenzyme pyridoxal phosphate to catalyse decarboxylation.
- Histamine (responsible for promoting acid secretion in the stomach and involved in the inflammatory response) is produced in the body by decarboxylation of histidine.

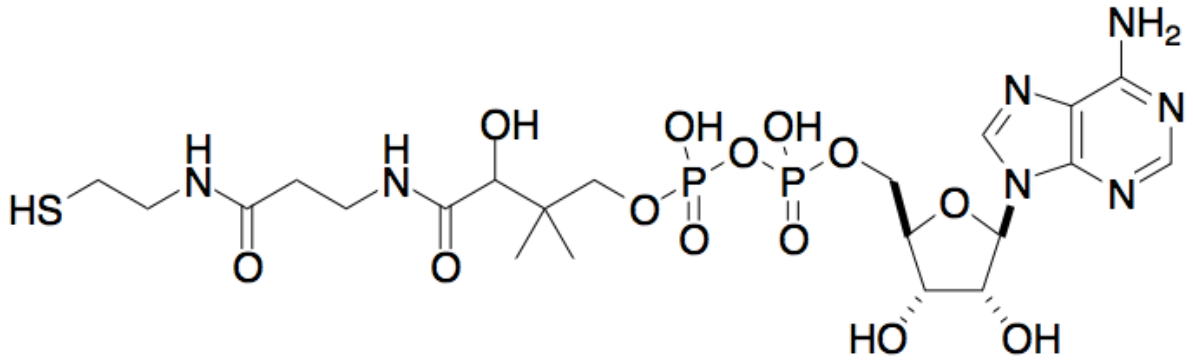


Enzymes – Pyridoxal phosphate/Pyridoxamine phosphate



Enzymes – Coenzyme A

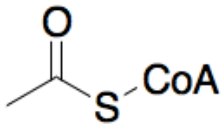
- Coenzyme A is an adenine nucleotide, linked by a 5’-pyrophosphate to pantothenic acid and then to an amino thiol.
- Most of the molecule is there to allow interaction with the various enzymes that catalyse the reactions of coenzyme A. Coenzyme A can be abbreviated to CoASH, where the thiol, SH is the important functional group.



Coenzyme A

Enzymes – Coenzyme A

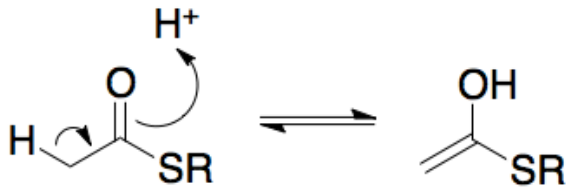
- Acetyl CoA is the acetyl thioester of CoASH.



CoA represents the rest of the coenzyme A molecule

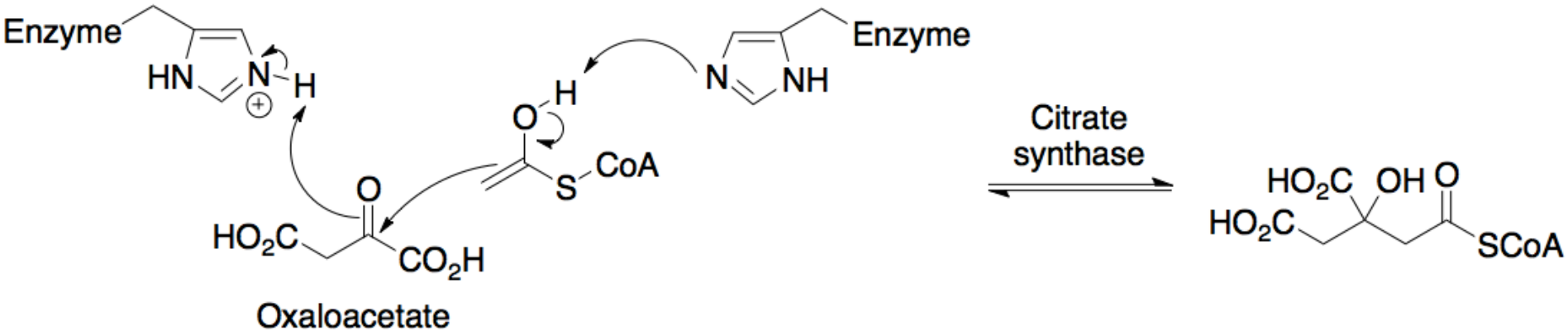
Acetyl CoA

- Thioesters are more electrophilic than ordinary esters because they are less conjugated. Because the thioester is more electrophilic, enolisation is easier.



Enzymes – Coenzyme A

- The key step in the citric acid cycle is the synthesis of citric acid from oxaloacetate and the enol of acetyl CoA.
- The enol of acetyl CoA attacks the reactive ketone of oxaloacetate to give citryl CoA. Two histidine residues provide acid/base catalysis.
- Hydrolysis of citryl CoA yields citric acid.

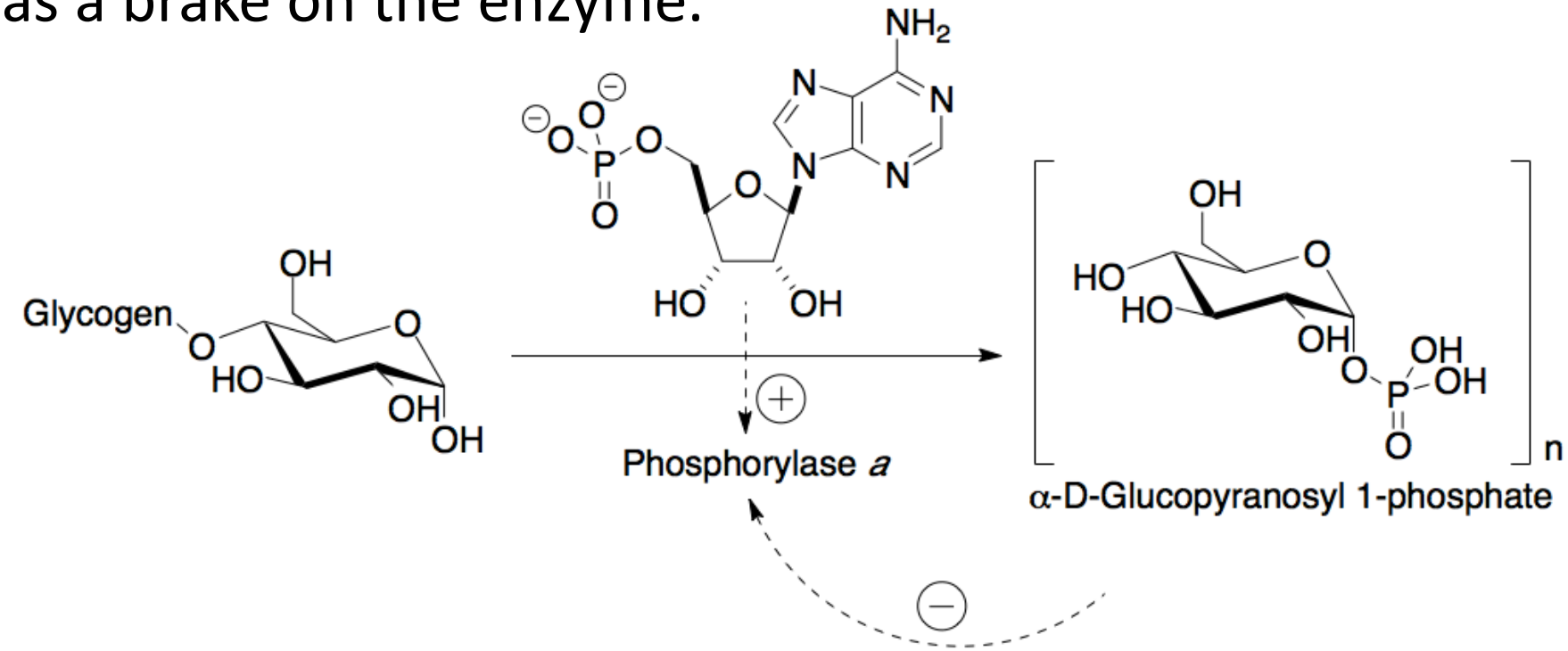


Enzymes are stereospecific

- Enzymes are highly specific in binding chiral substrates.
- Enzymes are **stereospecific** because enzymes consist of only L-amino acids and so have asymmetric active sites.

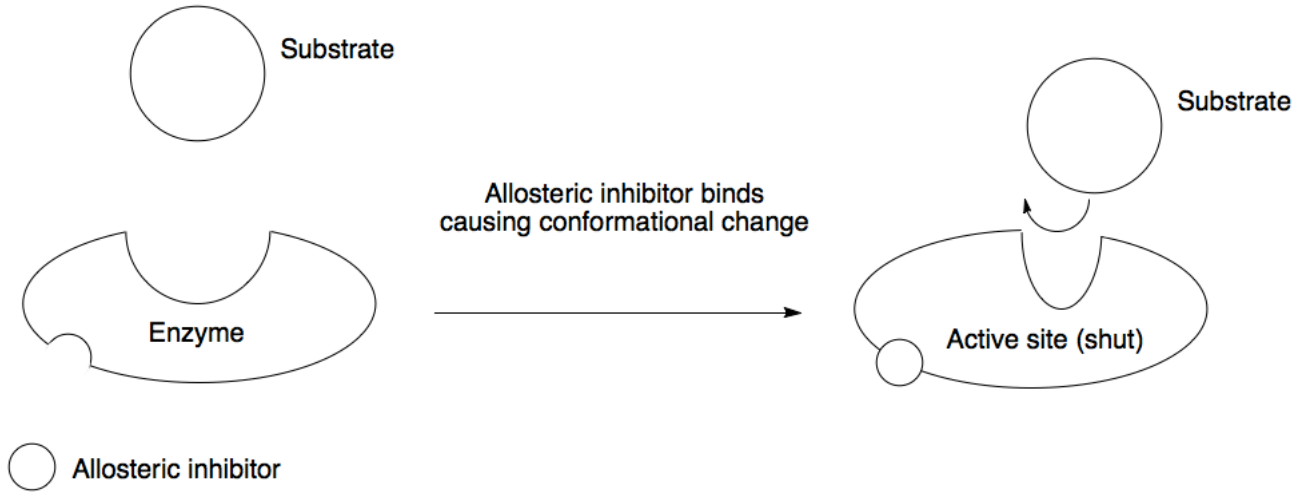
Enzymes – Regulation

- The enzyme phosphorylase *a* catalyses the breakdown of glycogen to glucose 1-phosphate.
- It is stimulated by AMP and inhibited by glucose-1-phosphate.
- This is an example of self regulation as increasing levels of product act as a brake on the enzyme.



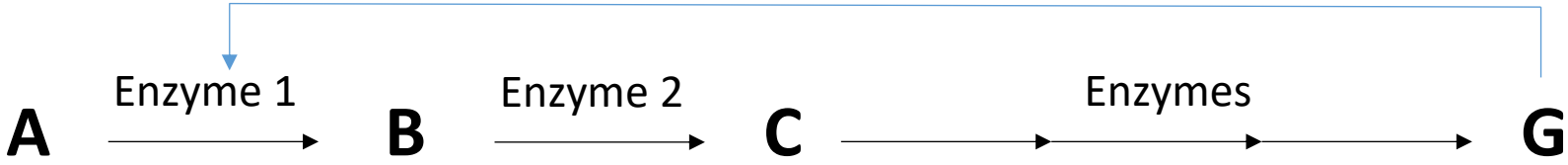
Enzymes – Regulation – Allosteric binding site

- Agents which control the activity of an enzyme bind at the allosteric binding site and are called allosteric inhibitors.
- Allosteric inhibitors produce an induced fit which alters the shape of the active site and renders it unrecognizable to the substrate.



Enzymes – Regulation

- The most common control mechanism is feedback control, where the final product in a biosynthetic pathway controls its own synthesis by inhibiting the first enzyme in the biosynthetic pathway by binding at an allosteric binding site.



- A separate allosteric binding site is an efficient feedback mechanism as binding of the final product to the active site would require the product to compete with the enzyme's substrate. If the concentration of the substrate increased, the inhibitor would be displaced and feedback control would fail.

Enzymes – Regulation

Competitive & non-competitive inhibition

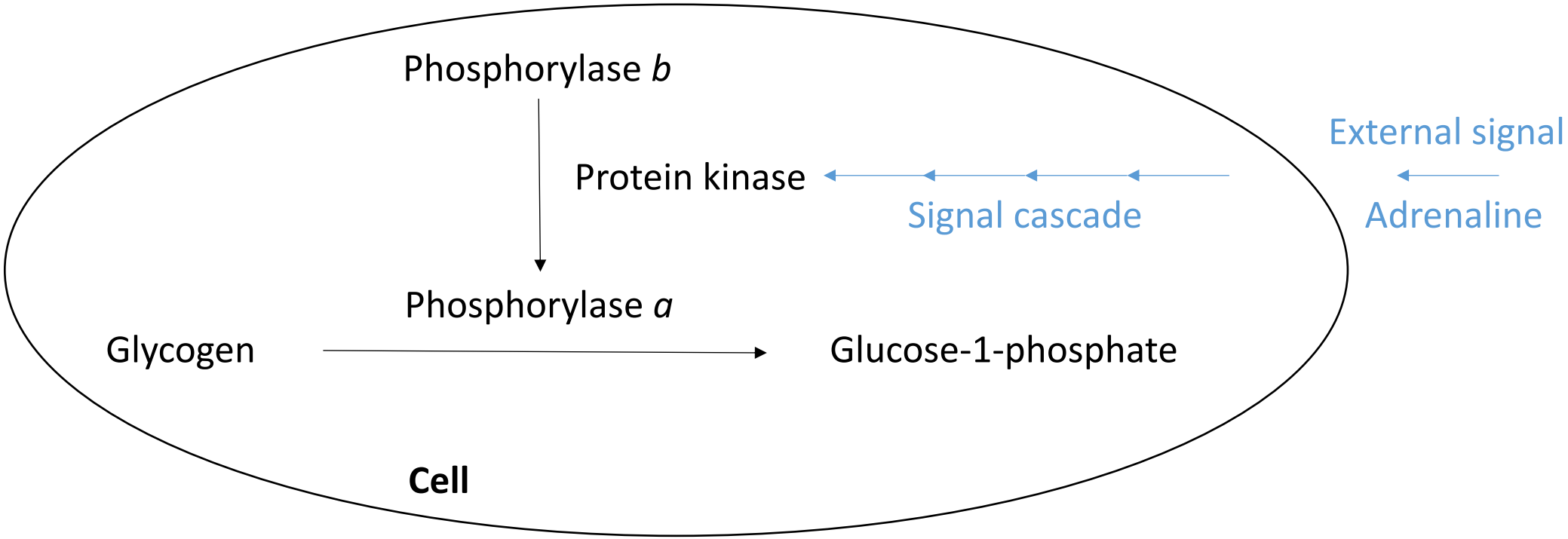
- **Inhibitors** are substances that slow down the rate of an enzyme catalysed reaction.
- Allosteric inhibitors do not compete with the substrate at the active site and are often referred to as **non-competitive inhibitors**.
- **Competitive inhibitors** have a similar shape to the substrate and so can bind at the active site, slowing down the rate of the reaction by occupying the active sites of the enzyme molecules making them less available for occupation by the substrate.

Enzymes – Regulation

- Many enzymes are regulated externally. A cell receives chemical messages from their environment setting off a cascade of signals within the cell which activate a special set of enzymes known as protein kinases.
- Kinases phosphorylate amino acids on the enzyme such as serine, threonine or tyrosine.
- The degradation of glycogen to glucose-1-phosphate is regulated externally by the hormone adrenaline.

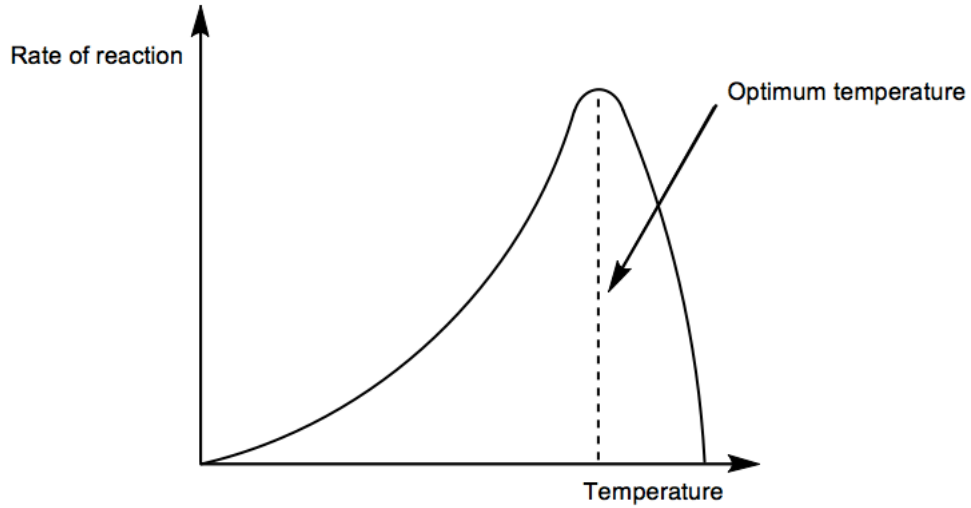
Enzymes – Regulation

Adrenaline triggers a signalling sequence which activates a protein kinase enzyme which phosphorylates phosphorylase *b* (inactive) to phosphorylase *a* (active).



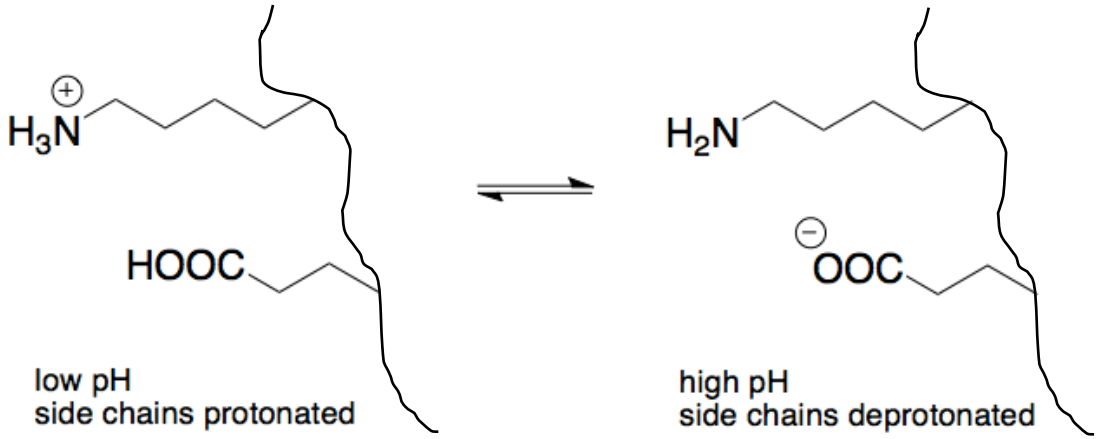
Effect of temperature on enzyme activity

- Initially increasing the temperature will increase the rate of an enzyme-catalysed reaction as more of the substrate molecules will posses the necessary acitivation energy.
- Above a certain temperature, the shape of the protein is changed as there is enough energy to overcome the intermolecular forces responsible for the tertiary structure. As such, the enzyme is no longer able to bind the substrate at the active site.



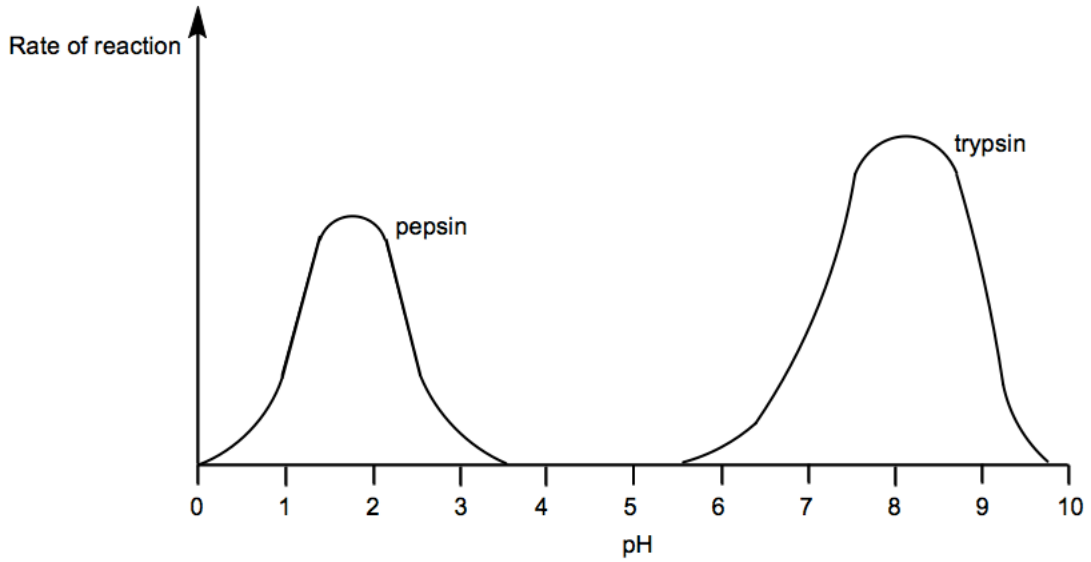
Effect of pH on enzyme activity

- Enzymes typically only work within a narrow pH range.
- The pH of a solution directly influences the state of ionisation of acidic or basic side chains in the protein affecting the intermolecular forces responsible for the enzyme's tertiary structure.
- Changes in the ionisation of side chains at the active site affect the enzyme's ability to form a complex with the substrate.



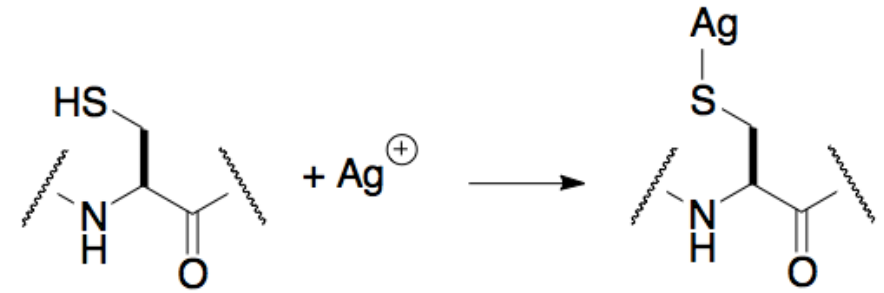
Effect of pH on enzyme activity

- Changes in the ionisation of side chains at the active site affect the enzyme's ability to form a complex with the substrate.
- Unlike for temperature, different enzymes in the same organism can have a wide variation in the optimum values of their pH.



Effect of heavy metal ions on enzyme activity

- Heavy metals such as lead, copper, mercury, and silver are poisonous
- These metal ions react with the thiol group, –SH, in the side chain of cysteine residues in the protein, forming a covalent bond with the sulfur atom and displacing a hydrogen ion.
- The folding of the protein is disrupted and the shape of the active site can change affecting the enzyme's ability to bind the substrate.



Classification of enzymes

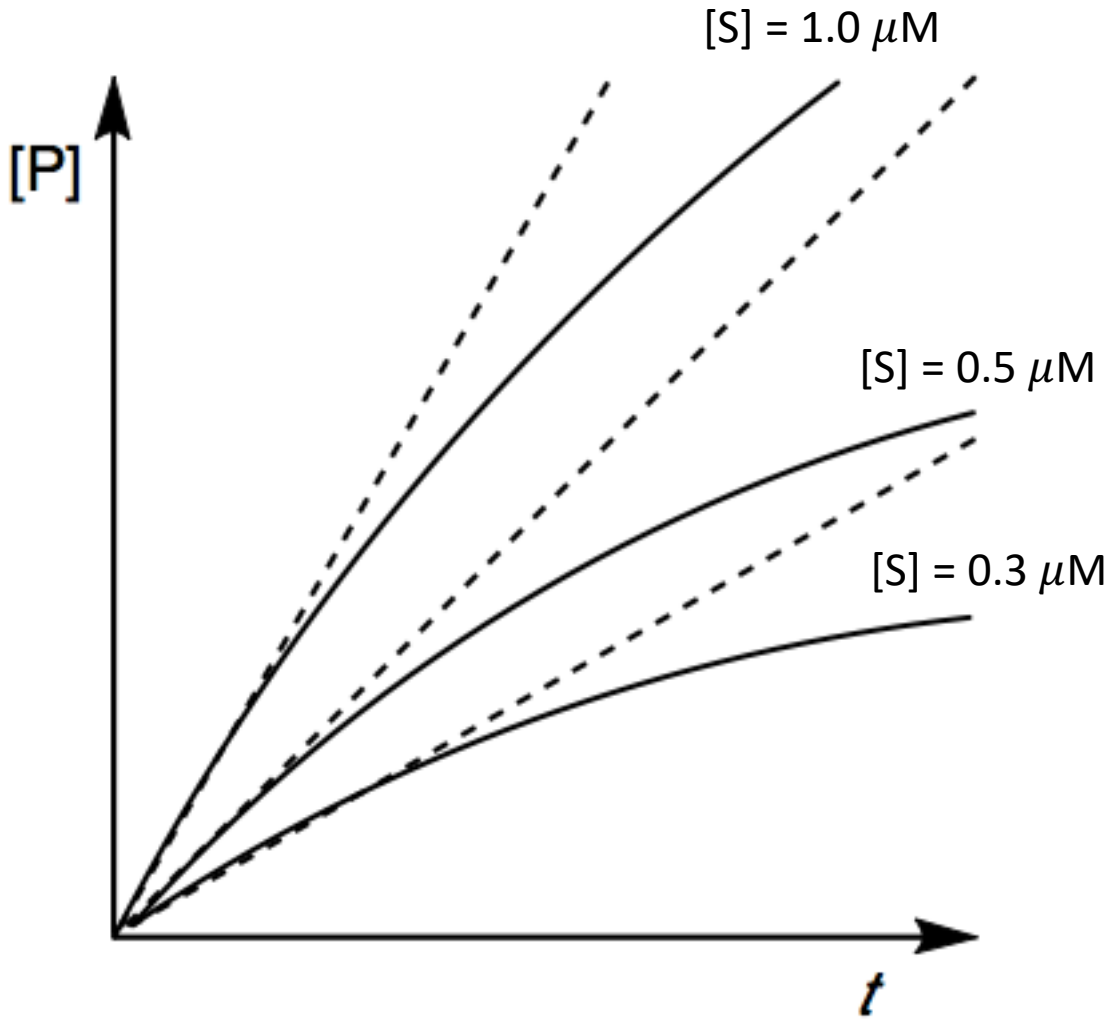
- Enzymes have the suffix ‘ase’.
- An oxidase enzyme catalyses oxidation. Enzymes are catalysts and so catalyse both the forward and reverse reactions. An oxidase enzyme can thus also catalyse reduction reactions.

Enzyme class	Type of reaction
Oxidoreductases	Oxidations and reductions
Transferases	Group transfer reactions
Hydrolases	Hydrolysis reactions
Lyases	Addition or removal of groups to form double bonds
Isomerases	Isomerisations and intramolecular group transfers
Ligases	Joining two substrates at the expense of ATP hydrolysis

Enzyme Kinetics

- A key factor affecting the rate of an enzyme catalysed reaction is the concentration of substrate, [S]
- In an in vitro reaction, [S] changes during the reaction as S → P
- A simplifying approach in kinetics experiments is to measure the initial rate, V_0
- V_0 can be investigated as a function of [S]

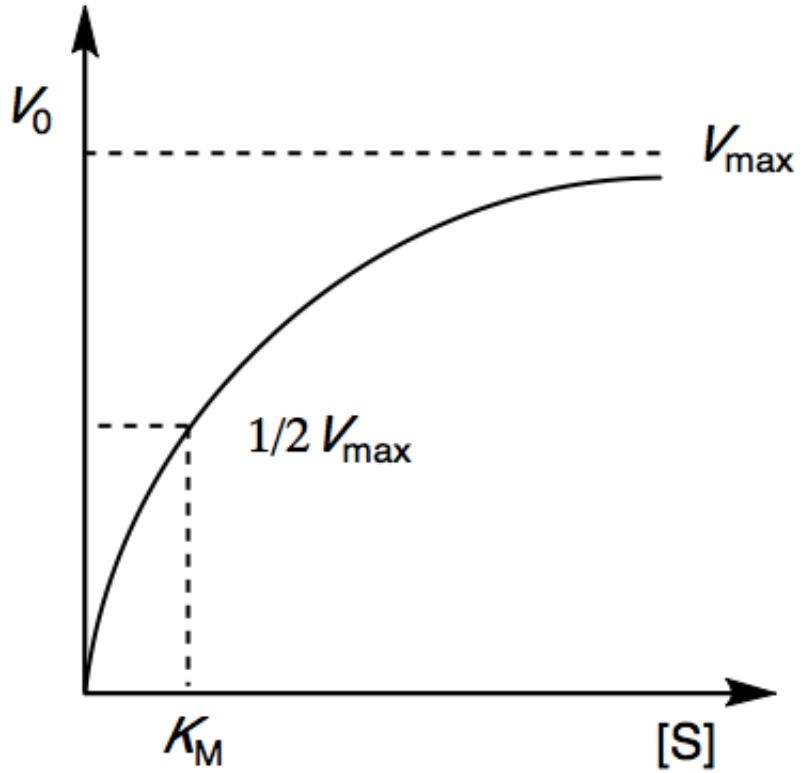
Enzyme Kinetics



Enzyme Kinetics

- At relatively low [S], V_0 increases almost linearly

- At higher [S], V_0 increases less and less in response to increases in [S] until at a given [S], the increases in V_0 are very small. This plateau like region is close to the maximum rate, V_{max}

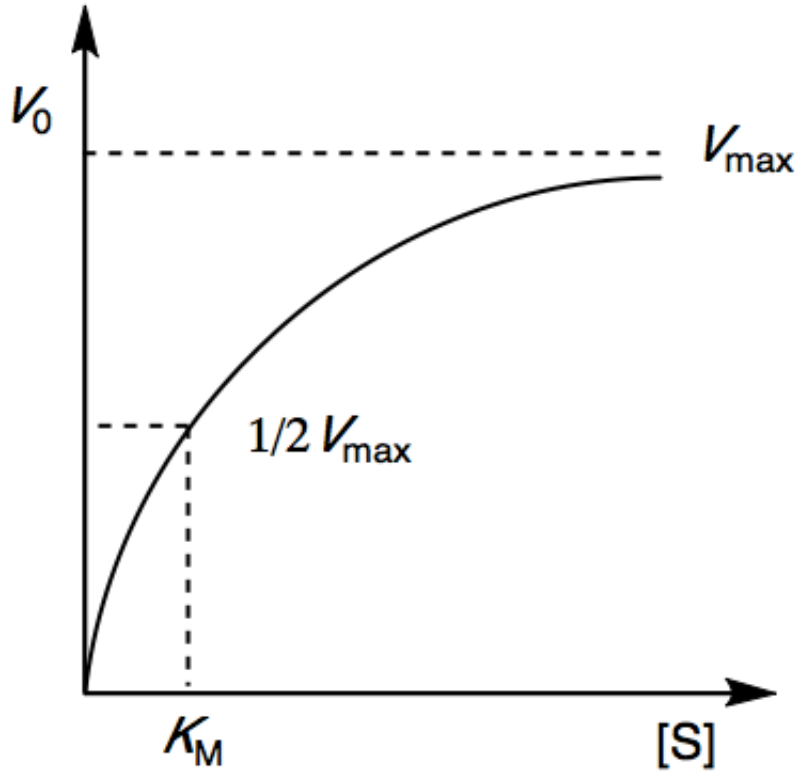


Enzyme Kinetics

- The free enzyme, E, combines with the substrate to form an enzyme-substrate complex, ES.
- The ES complex then breaks down to yield E and the reaction product, P.

$$E + S \underset{\text{Relatively fast}}{\rightleftharpoons} ES \underset{\text{Slower}}{\rightarrow} P + E$$

The maximum initial rate, V_{max} is observed when [S] is high enough that basically all the enzyme exists in the ES form. This saturation effect is characteristic of enzyme catalysis.



At low [S], most of the enzyme is in the E form.
Rate is proportional to [S] as the following equilibrium, E + S ⇌ ES, is pushed towards ES.

Enzyme Kinetics

- On mixing the enzyme with a large excess of substrate, there is an initial period when [ES] builds up, the **pre–steady state**. It is usually too short to be observed, lasting only microseconds.

- The reaction is then in the **steady state** in which [ES] remains approximately constant over time.

- The measured V_0 reflects the steady state, even though V_0 is limited to the early part of the reaction.

Enzyme Kinetics

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

$$\boldsymbol{V_0 = k_2[ES]}$$

[ES] is not easily measured experimentally but if we introduce the term, $[E_t]$, the total enzyme concentration can be expressed as: $[E_t] = [ES] + [E]$

Rate of ES formation $= k_1([E_t] - [ES])[S]$

$= k_1[E_t][S] - k_1[ES][S]$

Rate of ES breakdown $= k_{-1}[ES] + k_2[ES]$

$= (k_{-1} + k_2)[ES]$

Enzyme Kinetics

The steady-state assumption:

Rate of ES formation = Rate of ES breakdown

$k_1[E_t][S] - k_1[ES][S] = (k_{-1} + k_2)[ES]$

Enzyme Kinetics

$k_1[E_t][S] - k_1[ES][S] = (k_{-1} + k_2)[ES]$ // $+ k_1[ES][S]$

$k_1[E_t][S] = (k_1[S] + k_{-1} + k_2)[ES]$ // $\div (k_1[S] + k_{-1} + k_2)$

$$\frac{k_1[Et][S]}{k_1[S] + k_{-1} + k_2} = [ES]$$

$$\frac{[Et][S]}{[S] + (k_{-1} + k_2)/k_1} = [ES]$$

$(k_{-1} + k_2)/k_1$ is defined as the **Michaelis contant, K_m**

$$\frac{[Et][S]}{[S] + K_m} = [ES]$$

Enzyme Kinetics

We can now express V_0 in terms of [ES]

$$\frac{[Et][S]}{[S] + K_m} = [ES] \qquad V_0 = k_2[ES]$$

$$V_0 = \frac{k_2[Et][S]}{[S] + K_m}$$

The maximum rate, V_{max} occurs when the enzyme is saturated (when [ES] = [E_t]) and is equal to k_2[Et]

$$V_0 = \frac{V_{max}[S]}{[S] + K_m}$$

This is the **Michaelis-Menten equation**, the rate equation for a one substrate enzyme catalysed reaction

Enzyme Kinetics

The Michaelis constant is equal to the substrate concentration at which the reaction rate is half of its maximum value.

$$V_0 = \frac{V_{max}}{2} = \frac{V_{max}[S]}{[S] + K_m} \qquad // \div V_{max}$$

$$\frac{1}{2} = \frac{[S]}{[S] + K_m} \qquad // \text{ take reciprocals then } \times [S]$$

$$2[S] = [S] + K_m \qquad // - [S]$$

$$[S] = K_m$$

Enzyme Kinetics

- The K_m of an enzyme is significant as it gives the substrate concentration at which half the active sites in the enzyme are occupied.
- Importantly, it gives an indication of the substrate concentration required for significant catalysis to occur.

Enzyme Kinetics

- It should be noted that many enzymes follow Michaelis-Menten kinetics but the Michaelis-Menten equation is not dependent on the simple two-step reaction mechanism proposed by Michaelis and Menten:

$$E + S \overset{\text{Relatively fast}}{\rightleftharpoons} ES \overset{\text{Slower}}{\rightarrow} P + E$$

- We have seen that there are often many steps in the mechanism of an enzyme catalysed reaction and so the real meaning of V_{max} and K_m can differ from one enzyme to the next.
- However, Michaelis-Menten kinetics is the standard used by biochemists to compare and characterise the catalytic efficiencies of enzymes.

Enzyme Kinetics – Double reciprocal (Lineweaver-Burk) plot

- A problem with Michaelis-Menten kinetics is that if there are not sufficient data points, it can be hard to determine whether the curve of the Michaelis-Menten plot has reached a maximum value meaning that values for V_{max} and K_m are likely to be inaccurate.
- Plotting the reciprocals of the rate and the substrate concentration gives a Lineweaver-Burk plot.

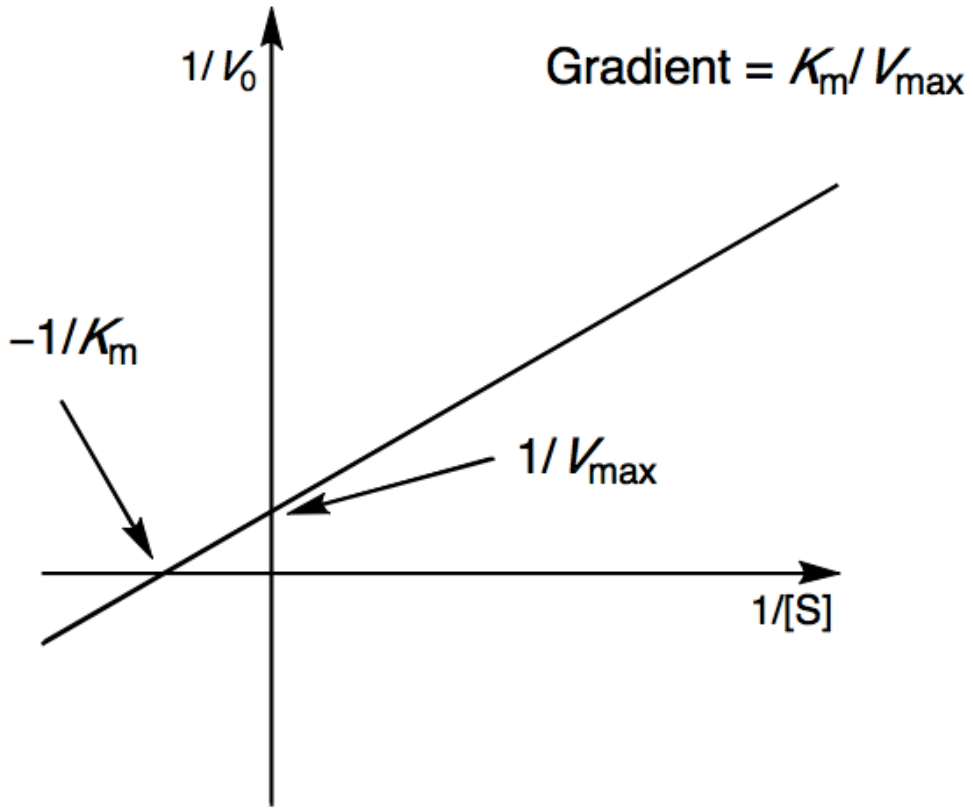
$$V_0 = \frac{V_{max}[S]}{[S] + K_m}$$

$$\frac{1}{V_0} = \frac{K_m}{V_{max}[S]} + \frac{[S]}{V_{max}[S]} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}}$$

Enzyme Kinetics – Double reciprocal (Lineweaver-Burk) plot

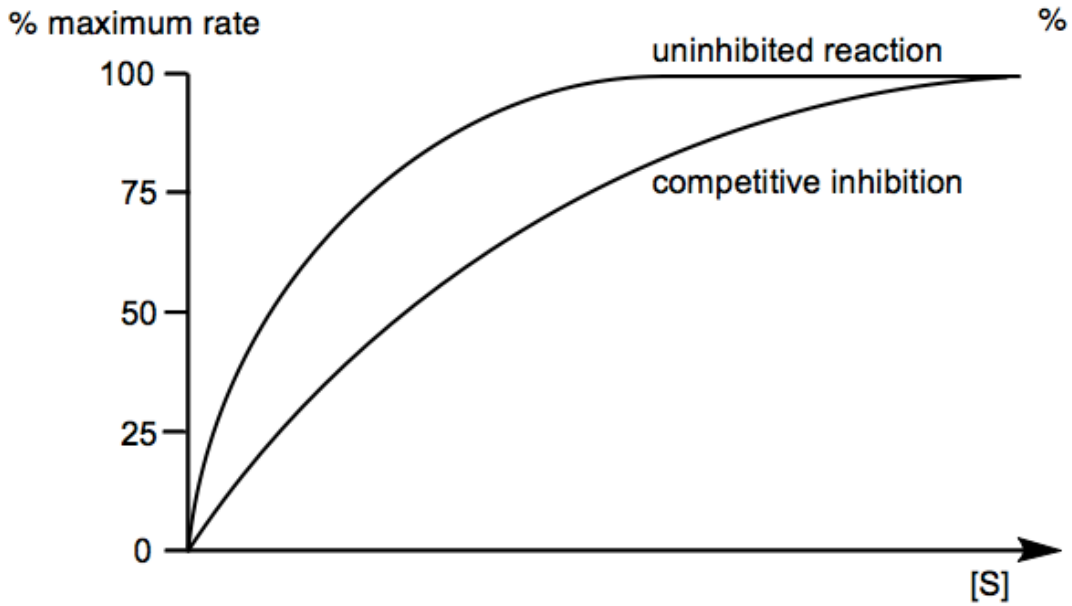
$$\frac{1}{V_0} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}}$$

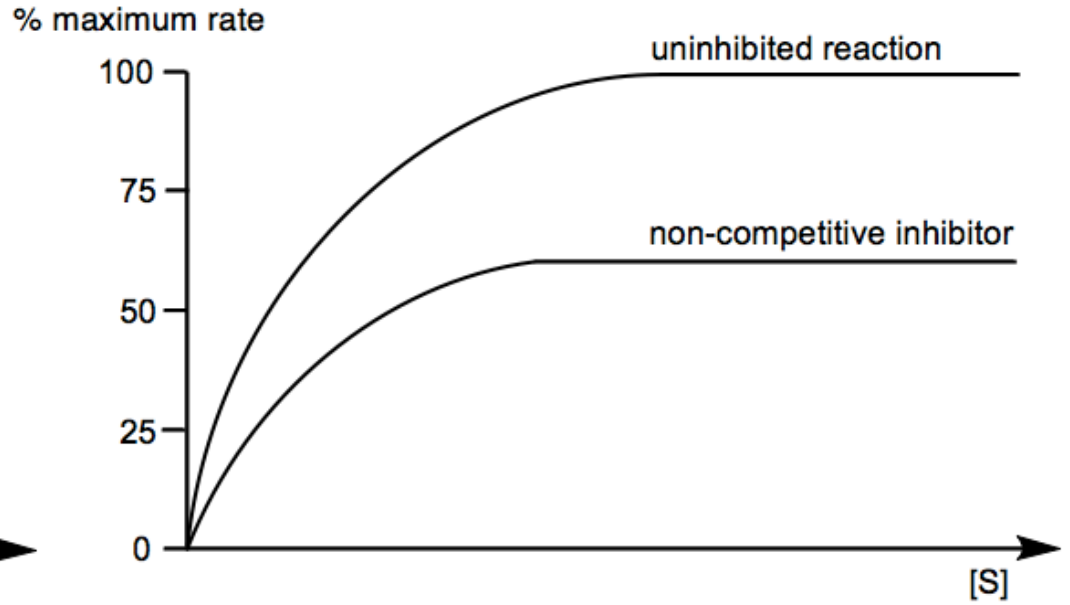
V_{max} can be obtained from the y-intercept and K_m can be obtained from the gradient or the x-intercept.



Enzyme Kinetics – Inhibitors

The following graphs shows the effects of different inhibitors on an enzyme catalysed reaction.





Enzyme Kinetics – Inhibitors

Competitive inhibitors:

- V_{max} is unchanged as there is still a substrate concentration where full enzyme activity can be achieved.
- However, it takes a higher substrate concentration to reach V_{max}, and so K_m is increased.

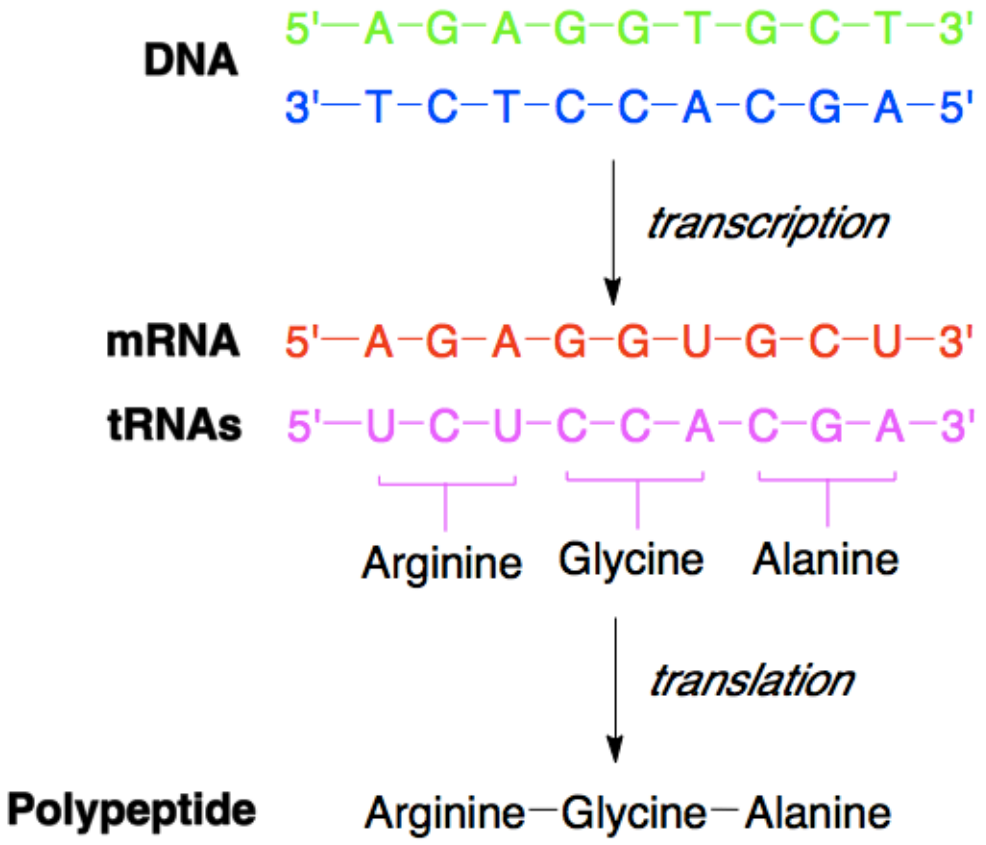
Non-competitive inhibitors:

- V_{max} is decreased because no matter how high the substrate concentration, a certain proportion of the enzymes have been inhibited by allosteric binding.
- A non-competitive inhibitor essentially lowers the concentration of the enzyme available for catalysis. V_{max} will decrease but the ability of the remaining enzymes to bind substrate is unaffected and so K_m is the same.

Protein Synthesis

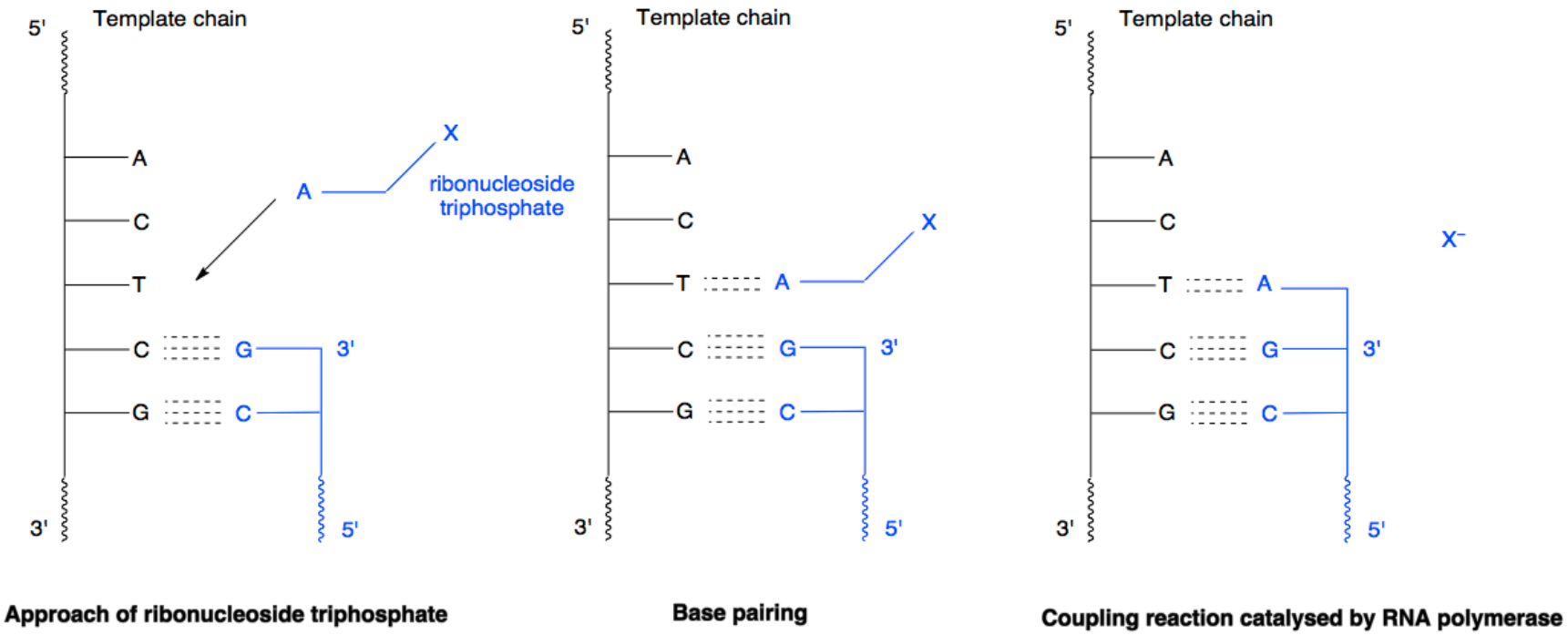
Peptide synthesis – Biosynthesis

- DNA directs transcription to yield RNA which is translated to form proteins.
- **Transcription** indicates that the "language" of the information is the same.
- **Translation** indicates that the "language" of the information changes from base sequence to amino acid sequence.

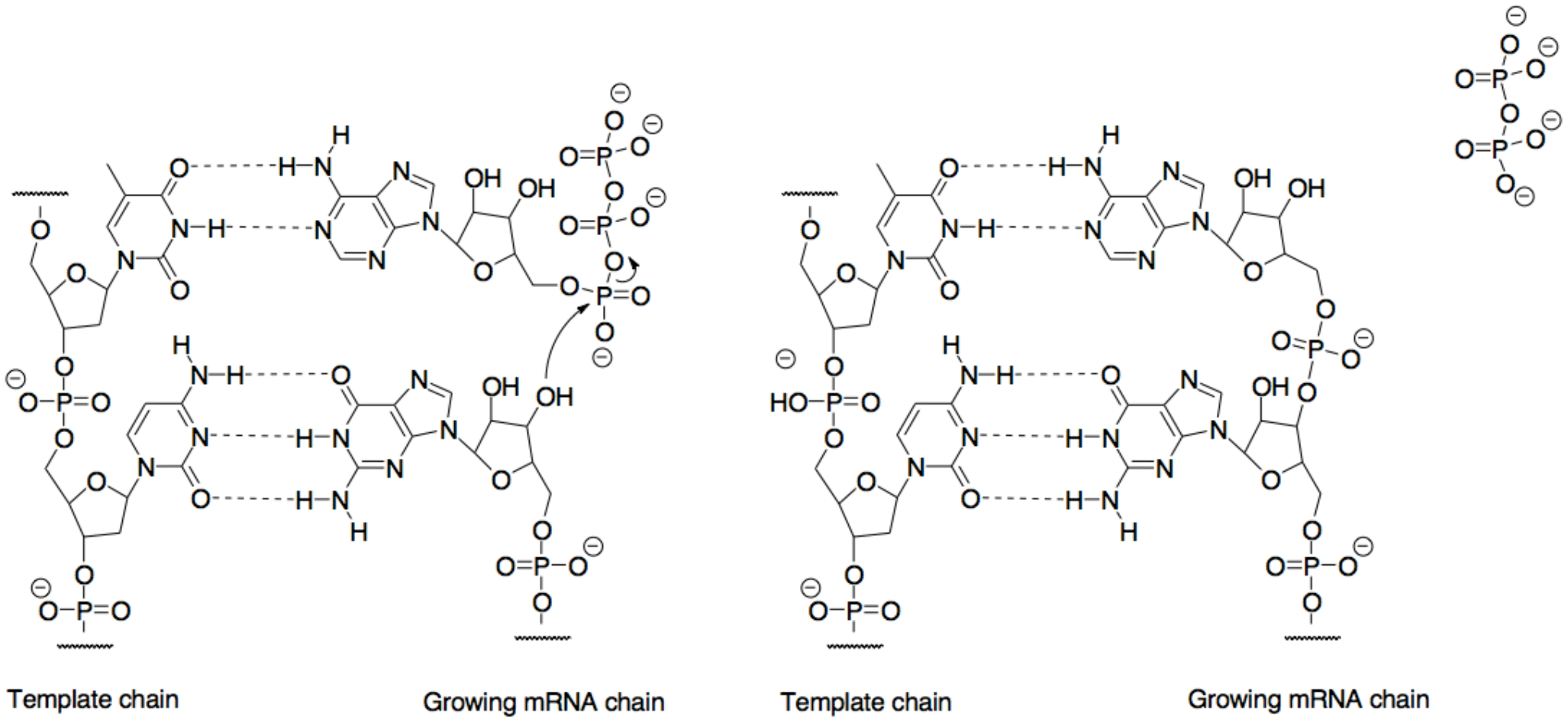


Peptide synthesis – Biosynthesis – Transcription

- mRNA synthesis proceeds in a stepwise manner in the 5’ → 3’ direction.
- The enzyme that synthesises RNA is called RNA polymerase.



Peptide synthesis – Biosynthesis – Transcription



Peptide synthesis – Biosynthesis – Translation

CODES WITHOUT COMMAS

BY F. H. C. CRICK, J. S. GRIFFITH, AND L. E. ORGEL

MEDICAL RESEARCH COUNCIL UNIT, CAVENDISH LABORATORY, AND DEPARTMENT OF THEORETICAL CHEMISTRY, CAMBRIDGE, ENGLAND

"The problem of how a sequence of four things can determine a sequence of twenty things is known as the coding problem."

F. H. C. Crick, J. S. Griffith, L. E. Orgel, *Proc. Natl. Acad. Sci. U.S.A.,* **1957**, 43 (5), 416-21

Peptide synthesis – Biosynthesis – Translation

- 4 bases in DNA code for 20 amino acids and so a group of several bases (a **codon**) is needed to code for a single amino acid.
- A triplet code (3 bases per codon) is needed since a doublet code would only allow $4^2 = 16$ codons whereas there are $4^3 = 64$ different triplets of four bases.

Peptide synthesis – Biosynthesis – Translation

Second position

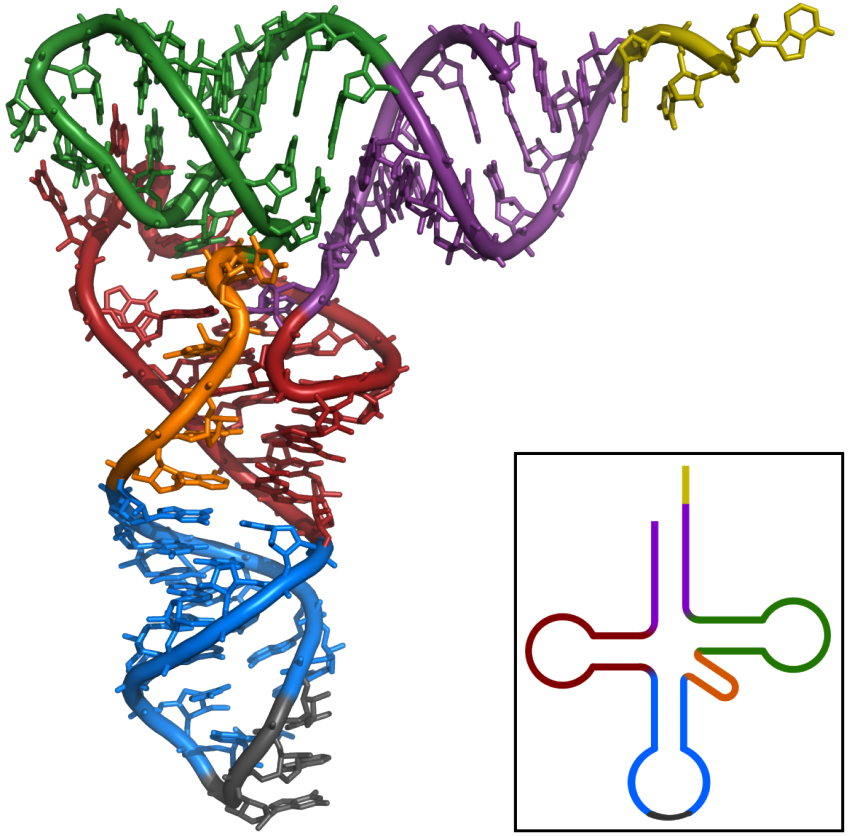
First position (5')	U	C	A	G	Third position (3')
U	UUU UUC Phe UUA UUG Leu	UCU UCC UCA UCG Ser	UAU UAC Tyr UAA UAG Stop	UGU UGC Cys UGA Stop UGG Trp	U C A G
C	CUU CUC CUA CUG Leu	CCU CCC CCA CCG Pro	CAU CAC His CAA CAG Gln	CGU CGC CGA CGG Arg	U C A G
A	AUU AUC AUA Ile AUG Met	ACU ACC ACA ACG Thr	AAU AAC Asn AAA AAG Lys	AGU AGC Ser AGA AGG Arg	U C A G
G	GUU GUC GUA GUG Val	GCU GCC GCA GCG Ala	GAU GAC Asp GAA GAG Glu	GGU GGC GGA GGG Gly	U C A G

UAG, UAA and UGA are "**stop**" codons, which act as a signal to the ribosome to terminate protein synthesis.

They are the only codons that do not specify amino acids.

Peptide synthesis – Biosynthesis – Translation

- Transfer RNA is the unit which links the triplet code on mRNA to a specific amino acid.
- Each tRNA molecule has two different binding regions at opposite ends.
- One binding region is for the amino acid, where a specific amino acid is covalently linked to a terminal adenosyl residue (shown in yellow).
- One binding region is a set of three nucleic acid bases, **anticodon** (shown in grey) which base-pair with the complementary triplet on the mRNA molecule.

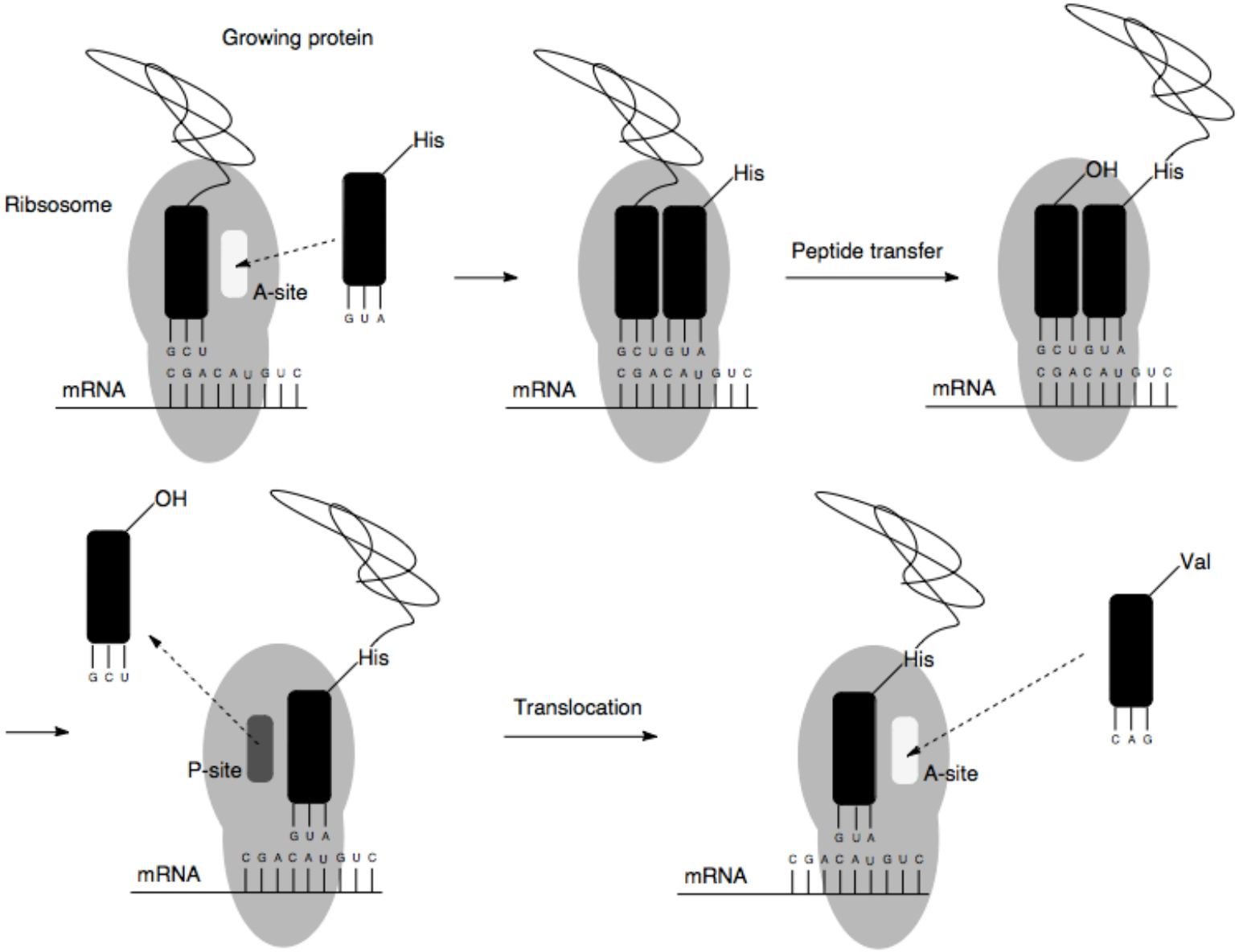


TRNA-Phe yeast, https://commons.wikimedia.org/wiki/File:TRNA-Phe_yeast_1ehz.png, Licence: Creative commons

Peptide synthesis – Biosynthesis – Translation

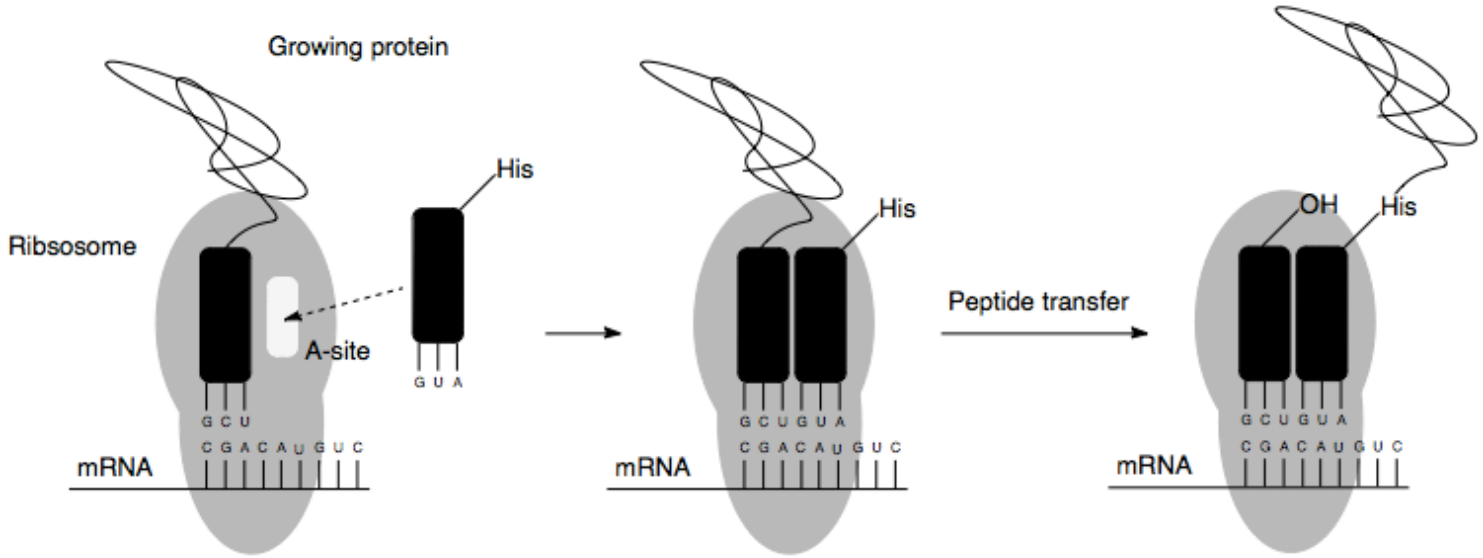
- The ribosome binds to one end of the mRNA molecule.
- The ribosome moves along the mRNA, allowing the triplet code to be read and catalysing the construction of the protein molecule one amino acid at a time.
- The genetic code is highly degenerate: Three amino acids (Leu, Arg, Ser) are each specified by six codons.
- Only Met and Trp, two of the least common amino acids in proteins, are specified by a single codon.

Peptide synthesis – Biosynthesis – Translation



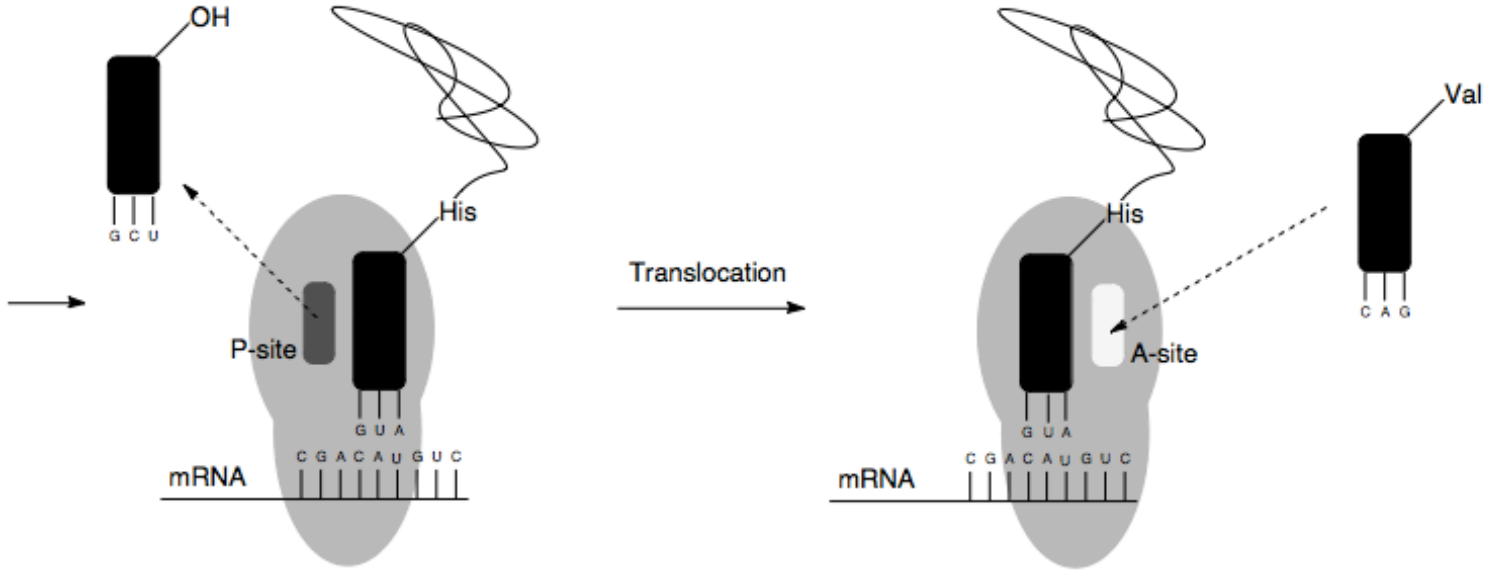
Peptide synthesis – Biosynthesis – Translation

- As the ribosome moves along the mRNA, it reveals the triplet codes on mRNA one by one along with an associated binding site called the A site.
- The tRNA molecule which base pairs with the exposed triplet on the mRNA enters the site.
- The peptide chain that has been synthesised so far is attached to a tRNA molecule bound to the P binding site.

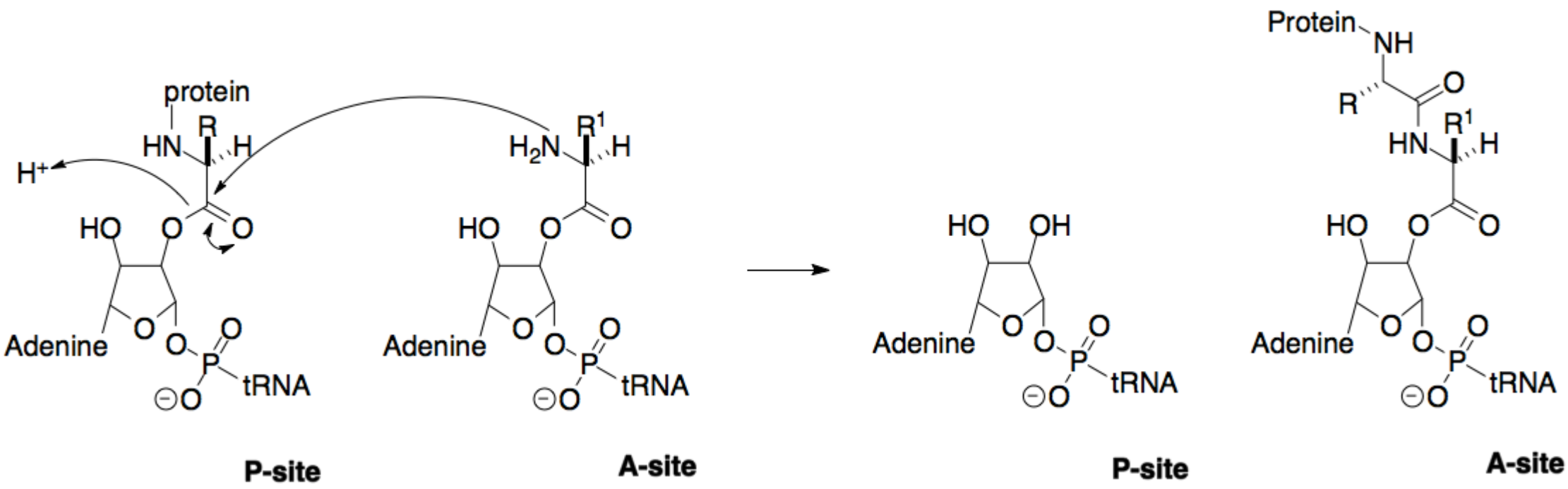


Peptide synthesis – Biosynthesis – Translation

- A peptide bond is formed, catalysed by the ribosome. The peptide chain is transferred to the tRNA occupying the A site.
- The ribosome enhances the rate of peptide bond formation by properly positioning and orienting its substrates.
- The tRNA occupying the P binding site departs and the ribosome shifts along the mRNA to reveal the next triplet (**translocation**)

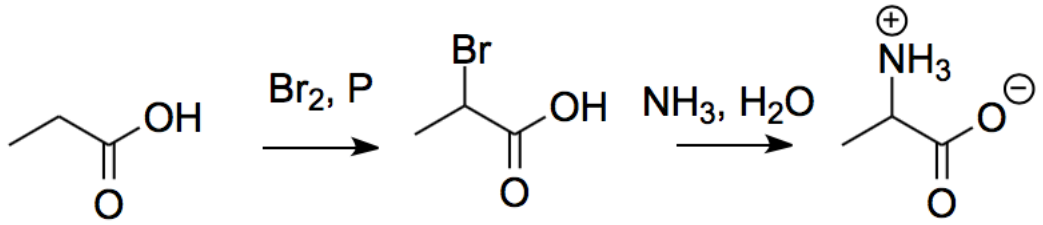


Peptide synthesis – Biosynthesis – Translation



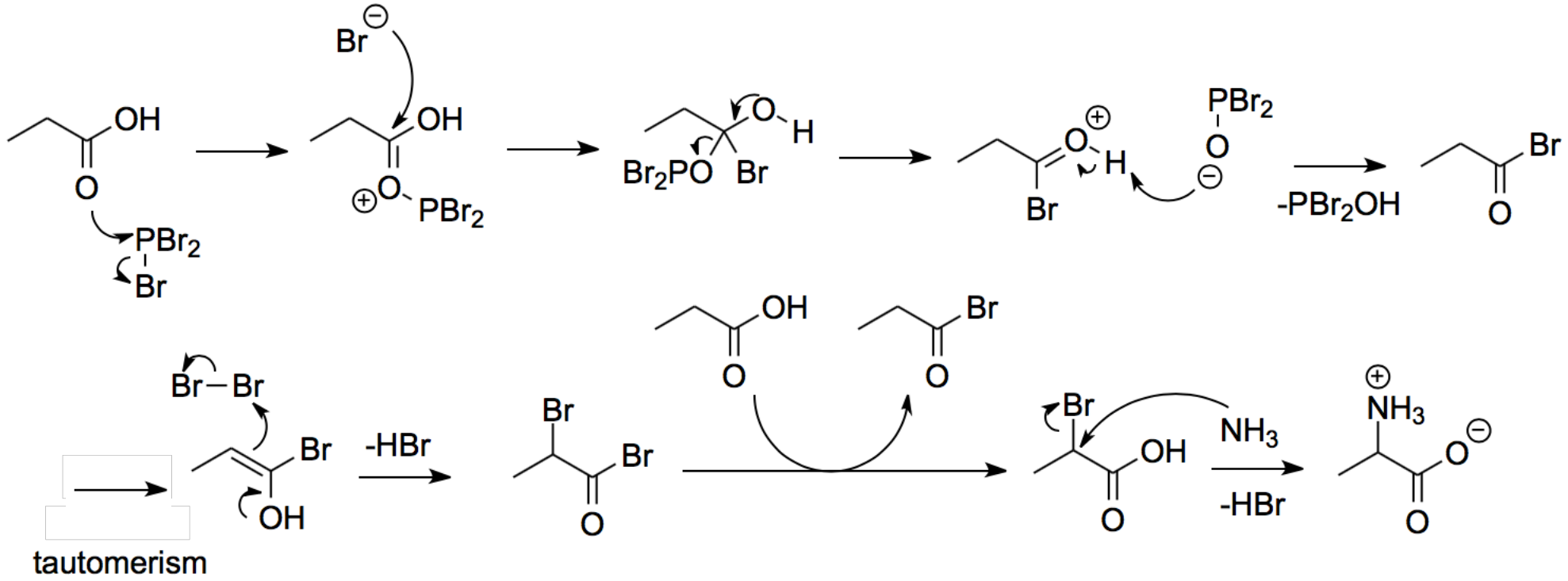
Amino acid synthesis

Hell-Volhard-Zelinsky bromination followed by amination



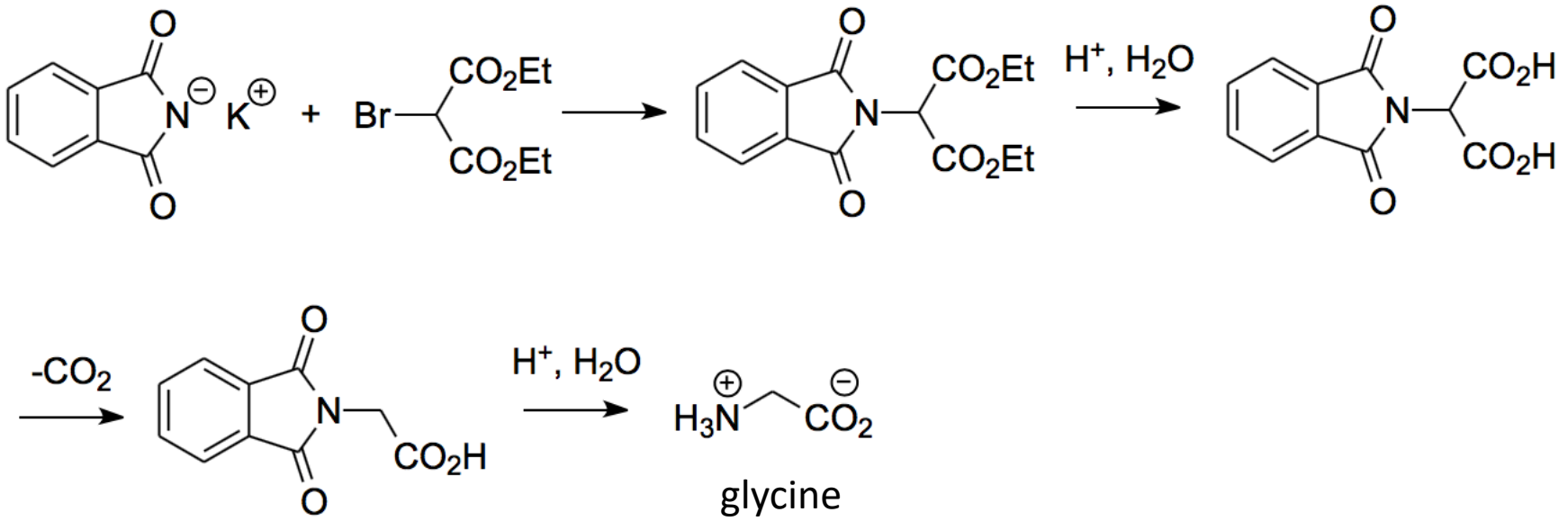
alanine

Amino acid synthesis

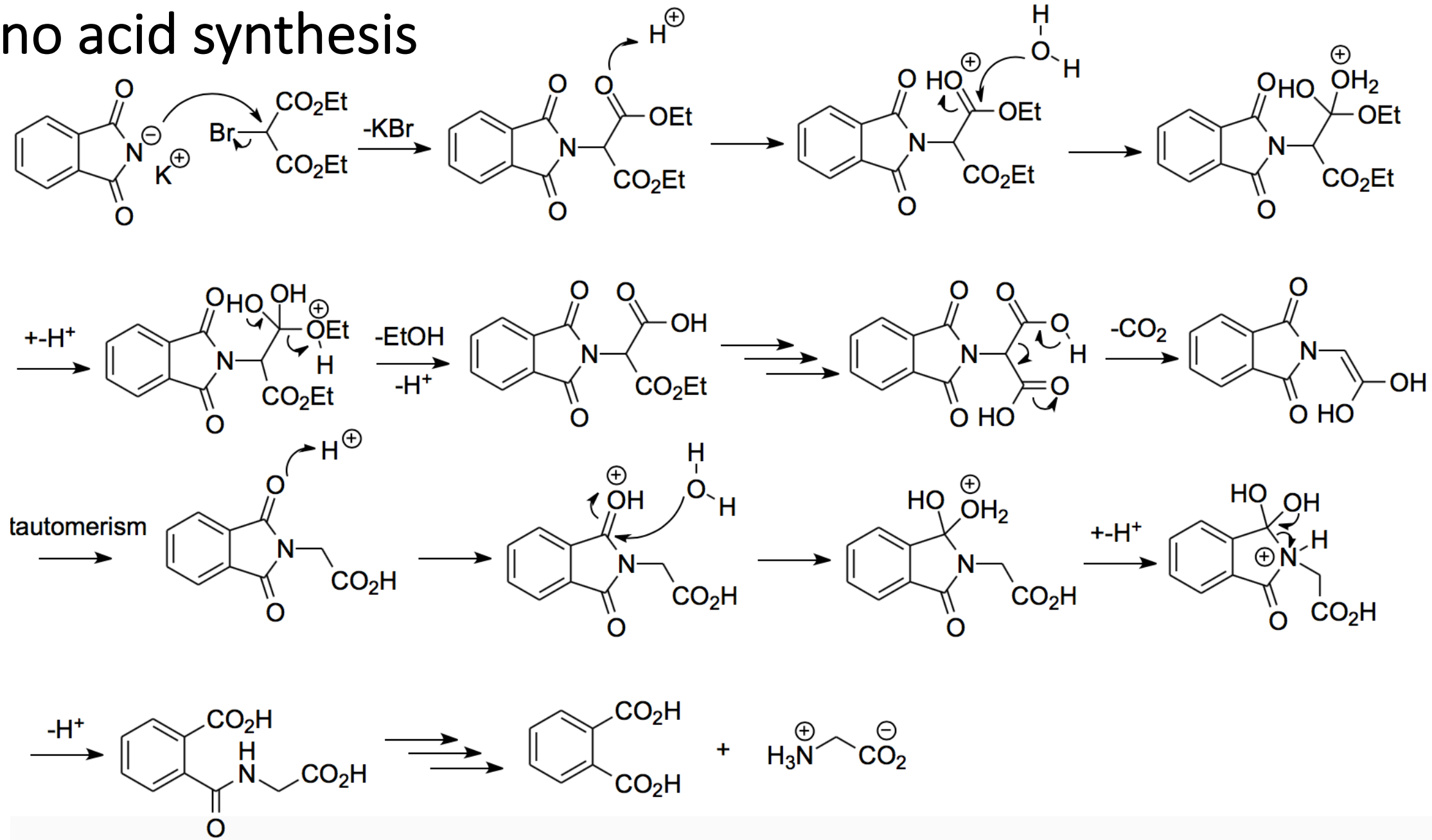


Amino acid synthesis

The Gabriel synthesis approach

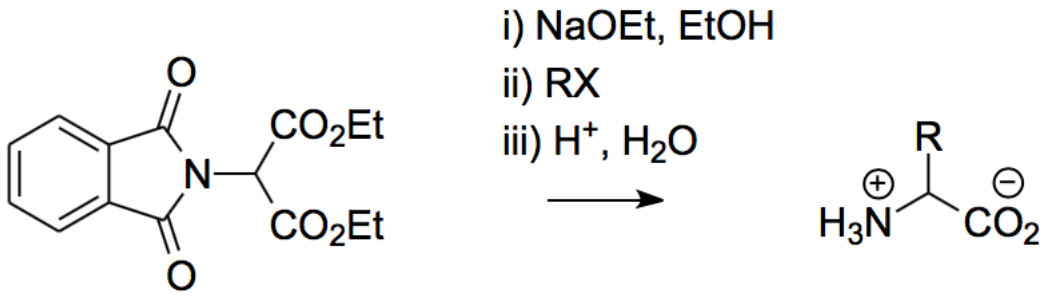


Amino acid synthesis

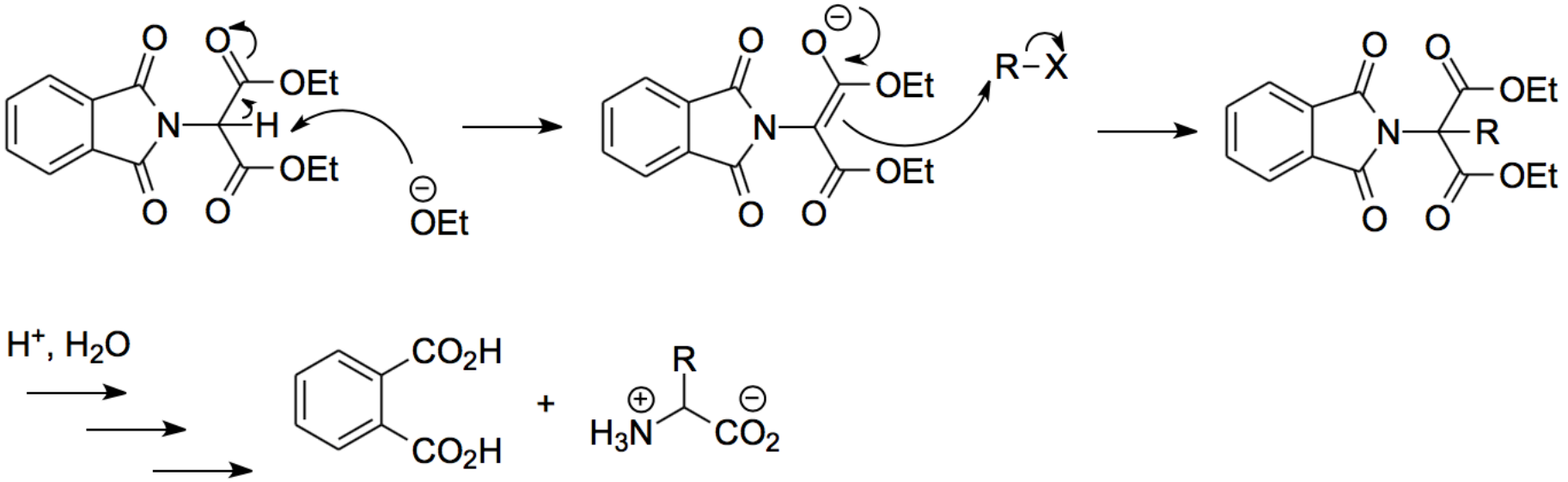


Amino acid synthesis

The Gabriel synthesis approach is versatile as we can alkylate the α-substituted malonate and so a variety of substituted amino acids can be prepared.

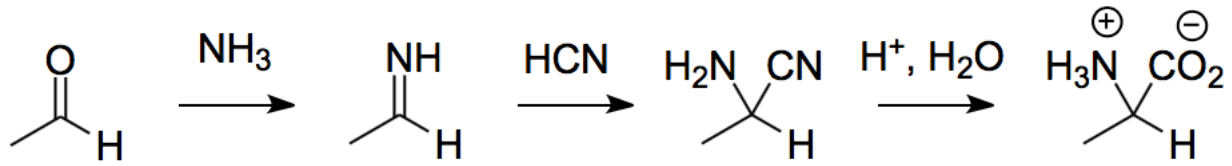


Amino acid synthesis

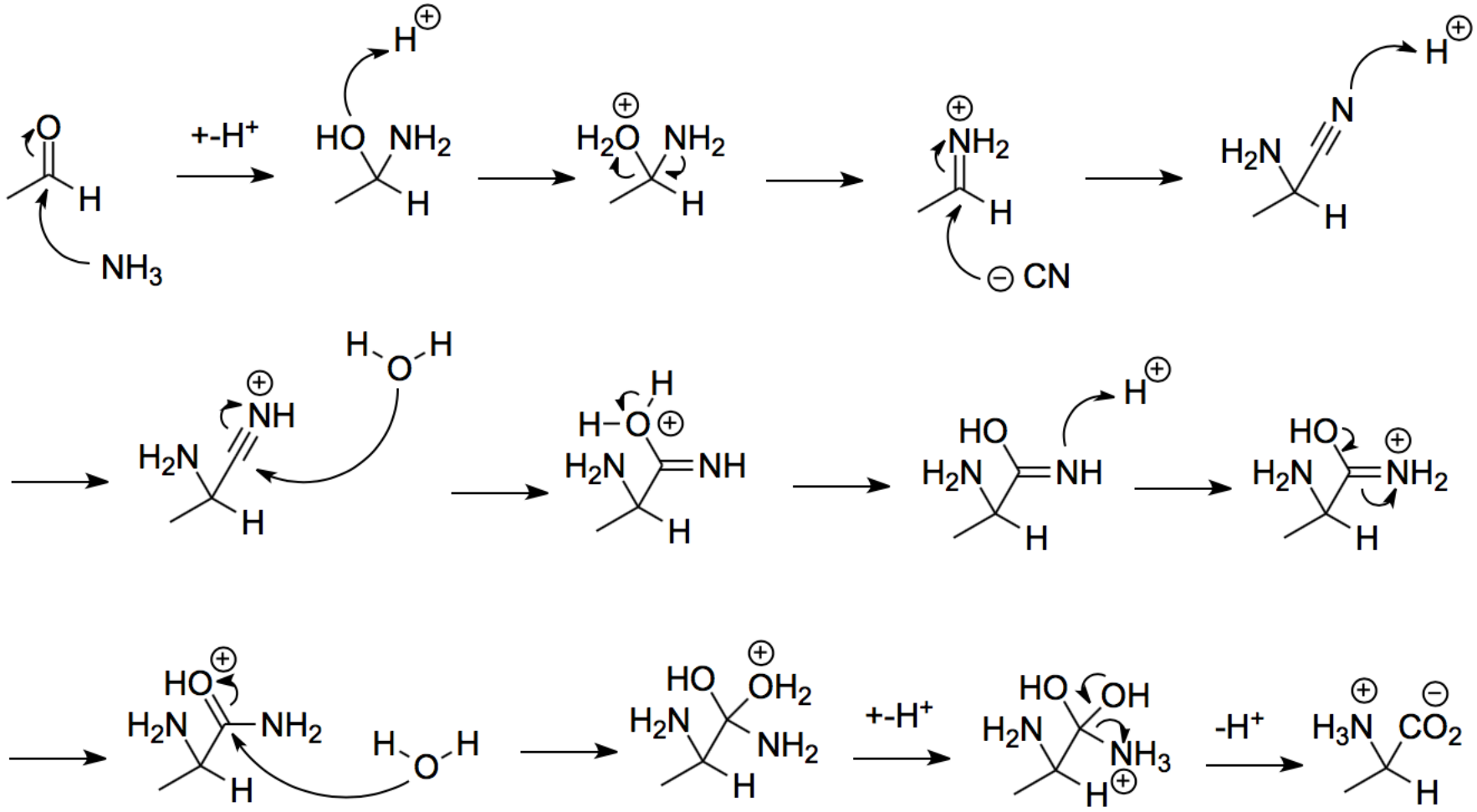


Amino acid synthesis

The Strecker synthesis



Amino acid synthesis

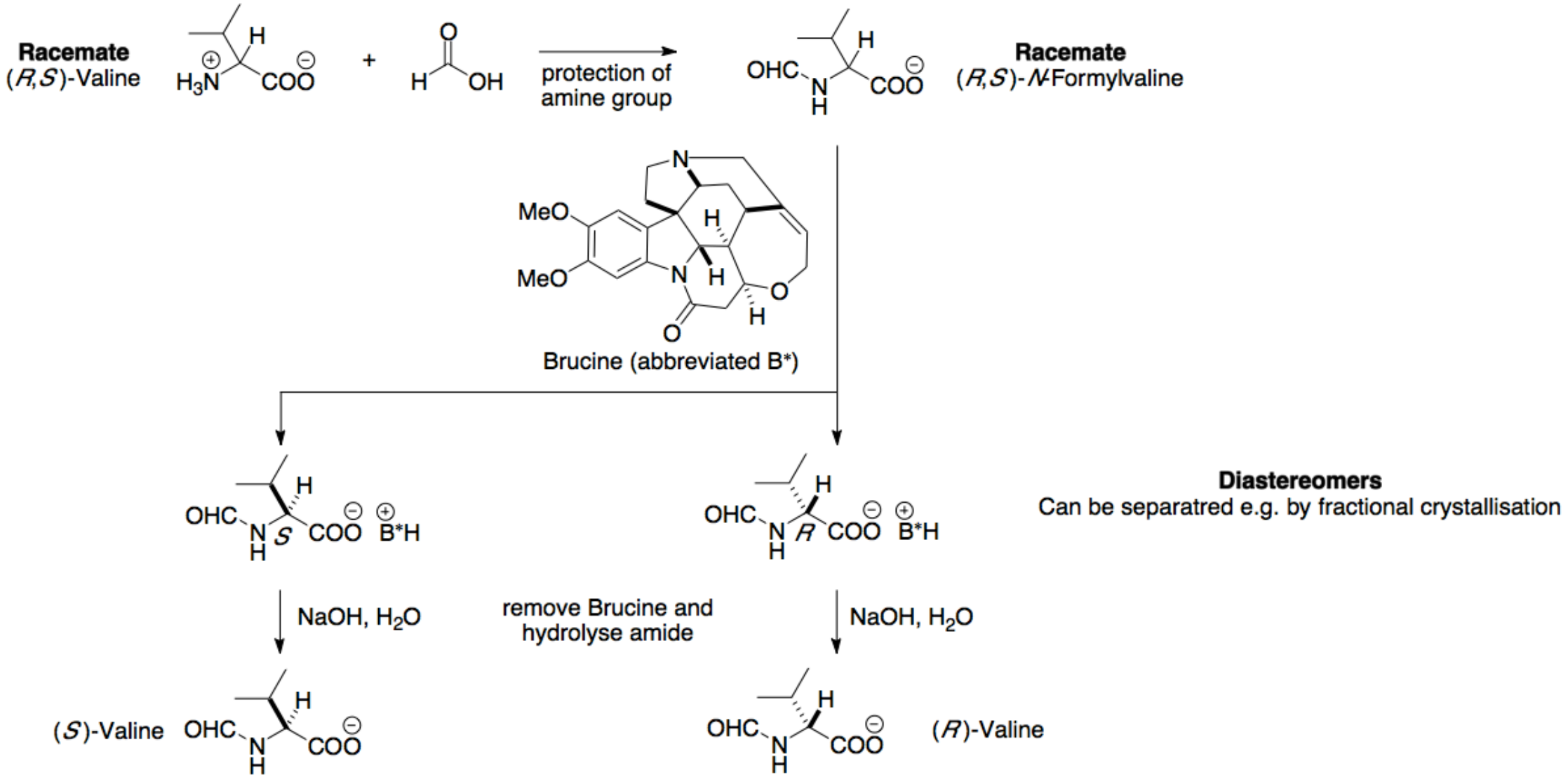


Enantiopure amino acid synthesis

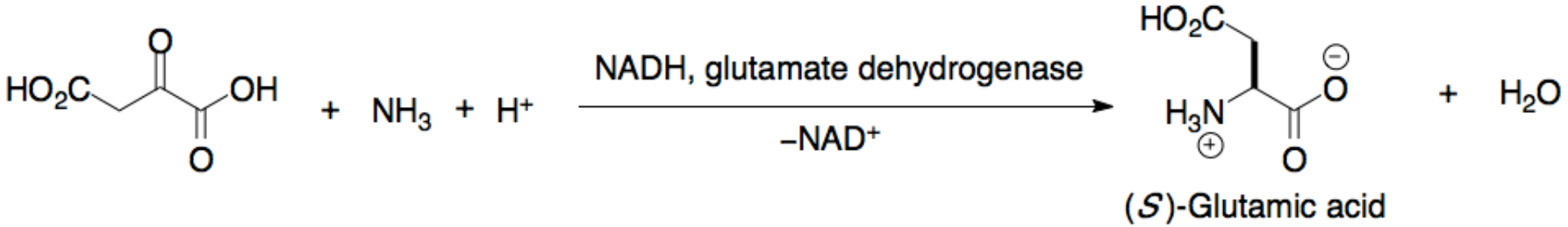
- '*L'univers est dissymétrique*'. Louis Pasteur, Comptes Rendu
- Chemists make use of nature's enantiomerically pure compounds (the chiral pool) to prepare enantiopure products.
- Strategies used to prepare enantiopure amino acids include:

1. **Resolution of diastereomeric salts** involves bringing two stereogenic centres together *via* an ionic interaction. Separable diastereomers are created from inseparable enantiomers.
2. **Assymetric biocatalysis**, where an achiral substrate is converted into a chiral product *via* enzymatic catalysis. The chirality is introduced into the product by influence of the enzyme, which is chiral.
3. **Alkylation of derivatives of glycine** in the presence of chiral ammonium salts.

Enantiopure amino acid synthesis – Resolution

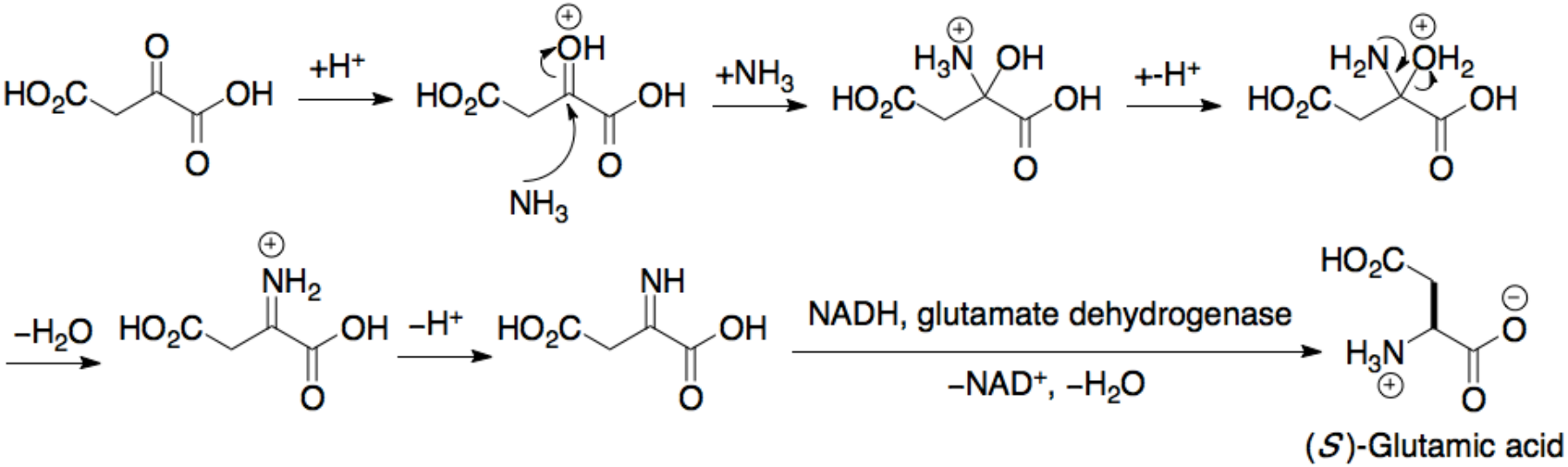


Enantiopure amino acid synthesis – Assymetric biocatalysis

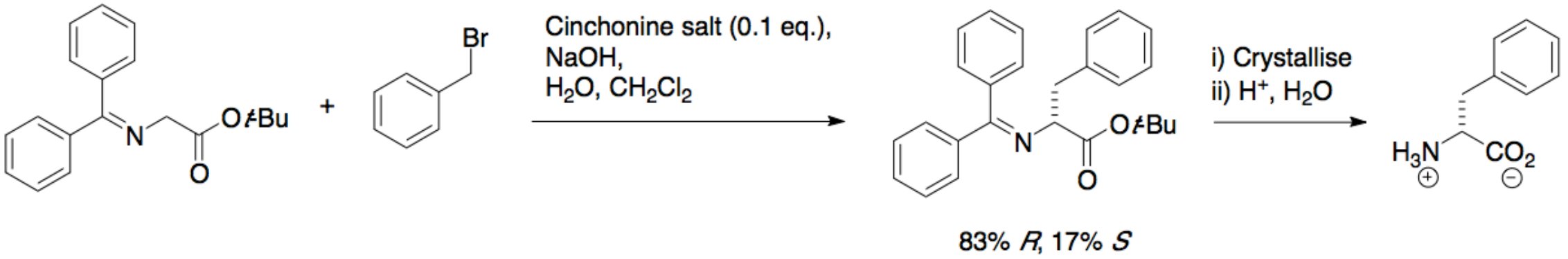


(*S*)-Glutamic acid

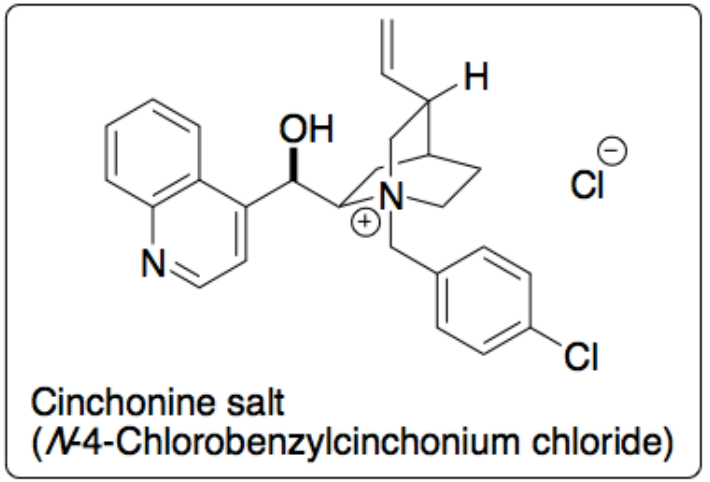
Enantiopure amino acid synthesis – Assymetric biocatalysis



Enantiopure amino acid synthesis – Enantioslective alkylation

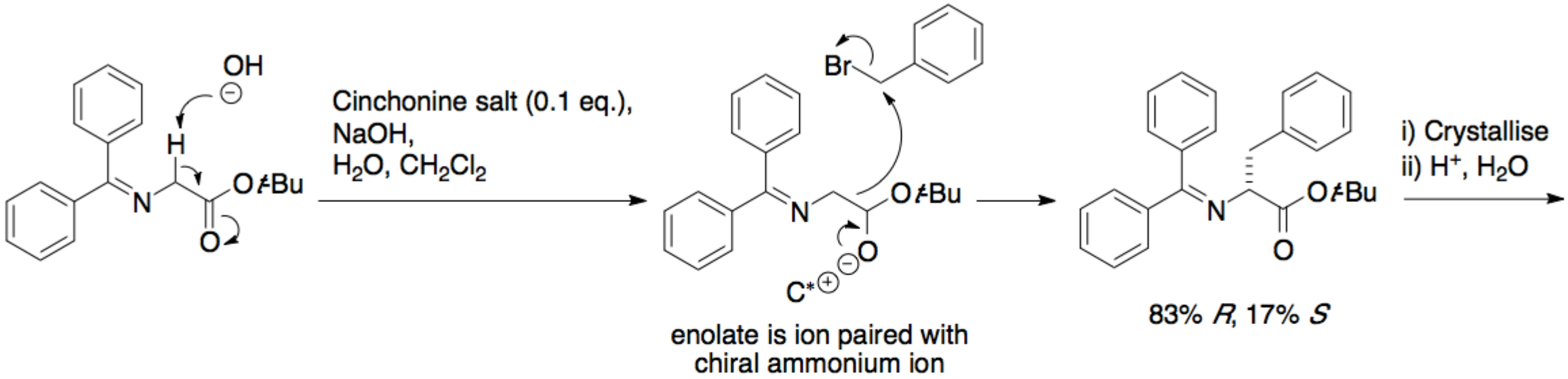


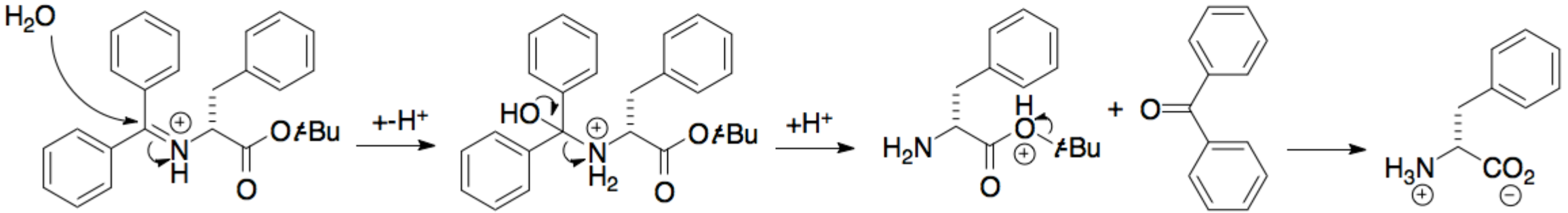
83% *R*, 17% *S*



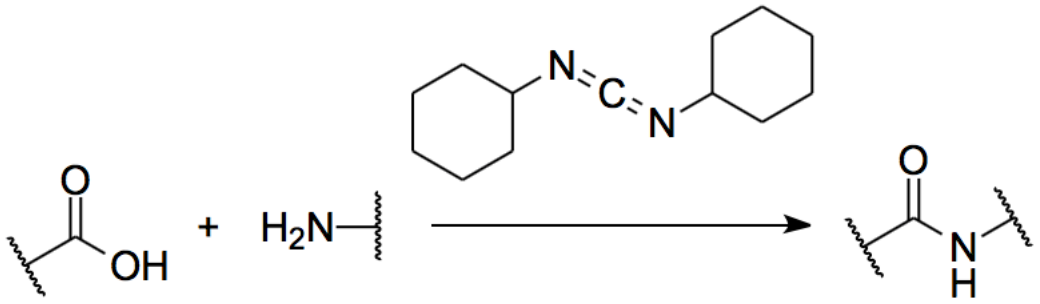
Cinchonine salt
(*N*-4-Chlorobenzylcinchonium chloride)

Enantiopure amino acid synthesis – Enantioslective alkylation



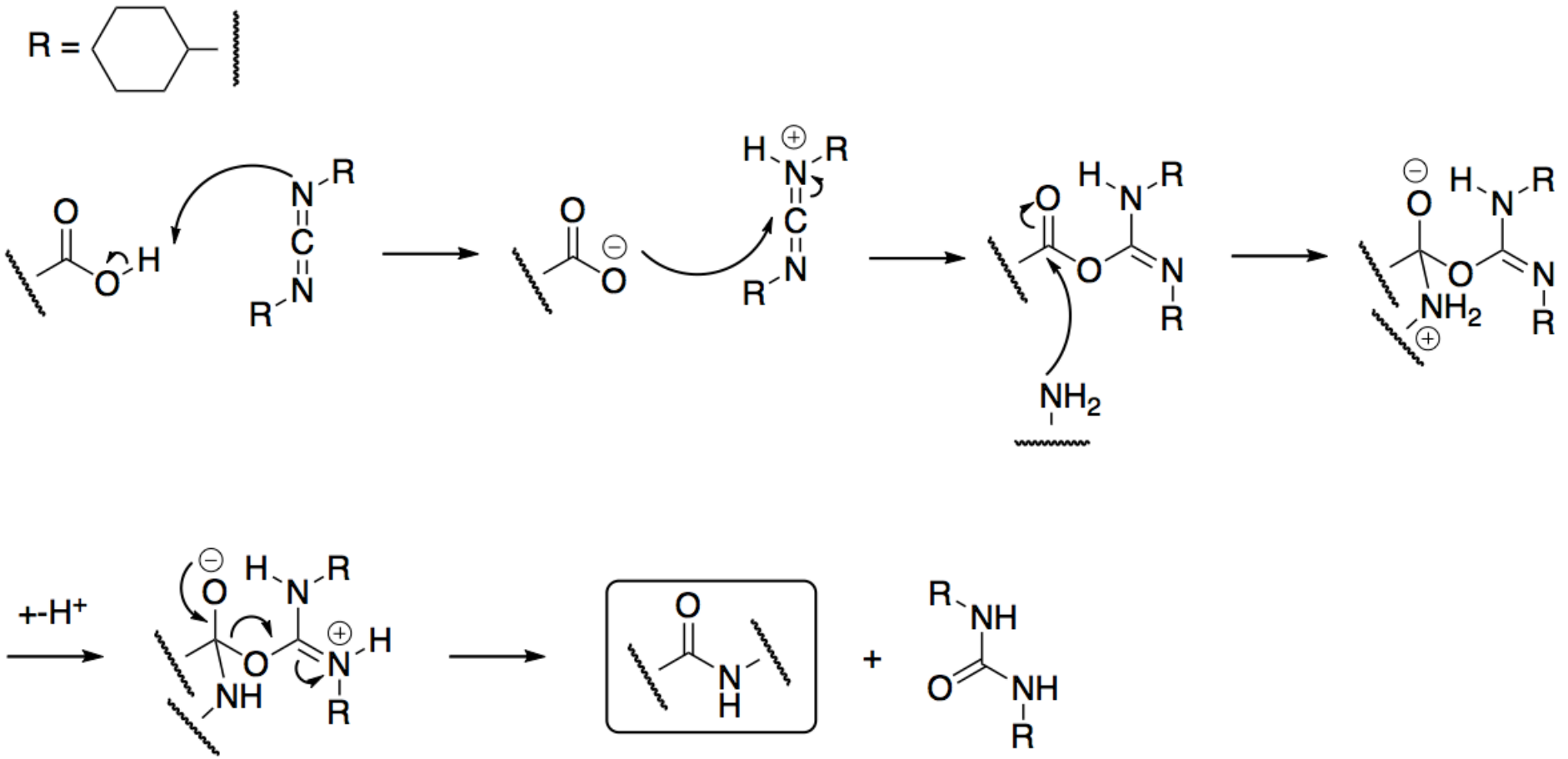


Peptide synthesis – Carboxyl group activation – DCC



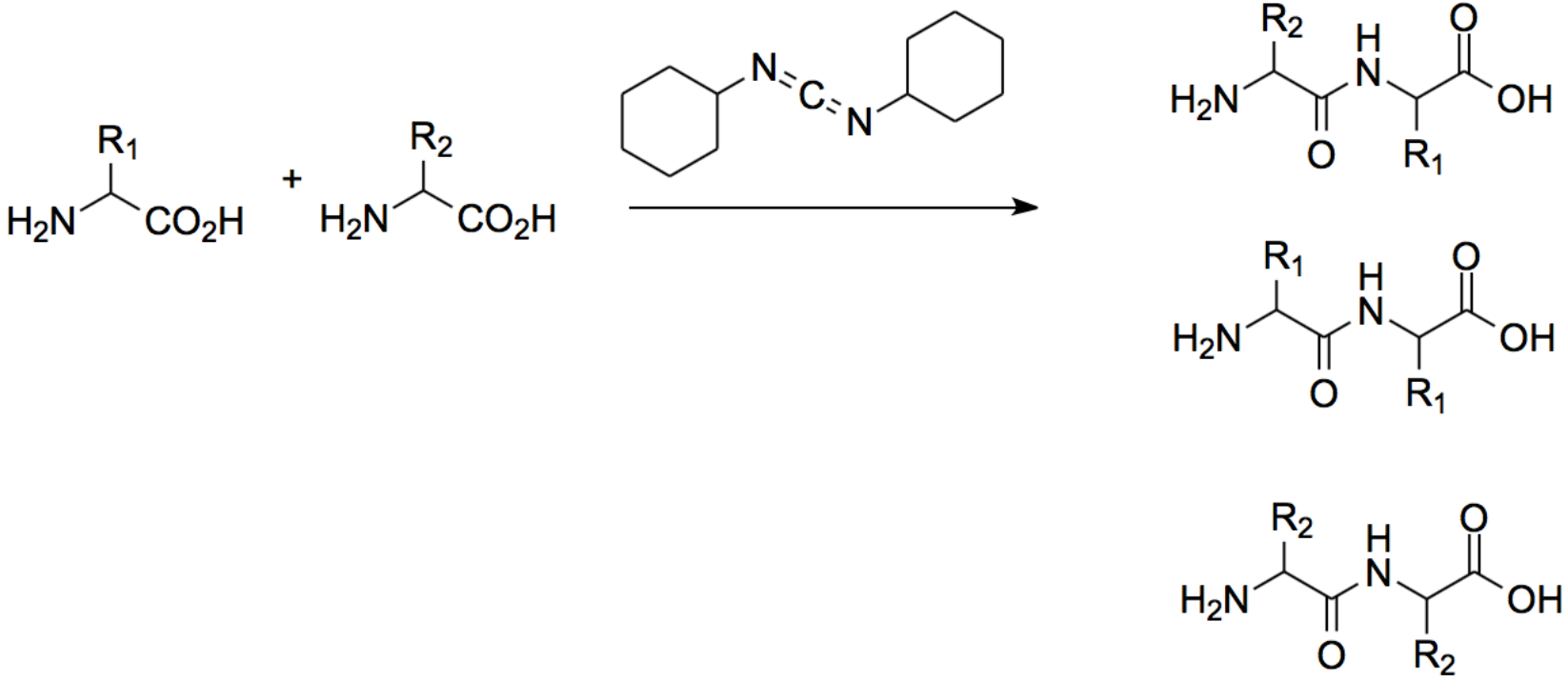
Dicyclohexylcarbodiimide (DCC) activates the carbonyl group of the acid to nucleophilic attack *via* formation of an *O*-acycl isourea.

Peptide synthesis – Carboxyl group activation – DCC



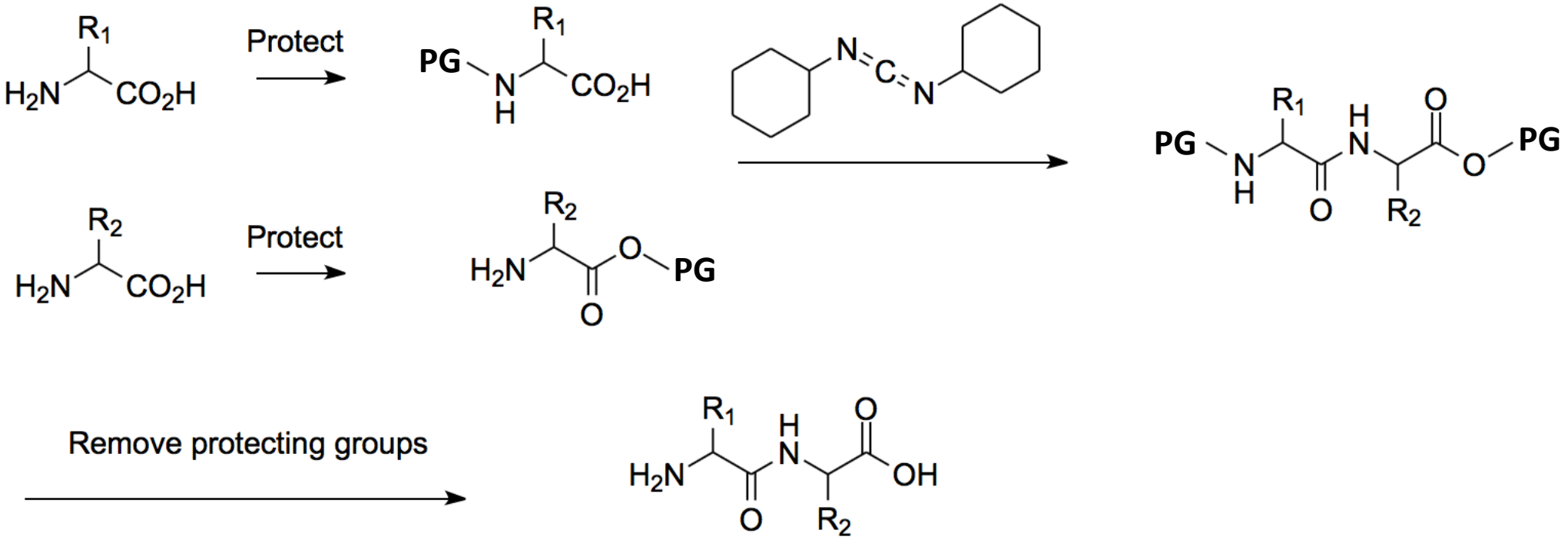
Peptide synthesis- Regioselectivity

If we try and couple two amino acids using DCC, we observe a lack of regioselectivity.



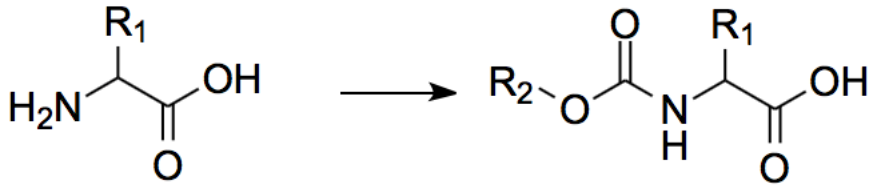
Peptide synthesis – Protecting group chemistry

The regioselectivity problem can be solved by protecting the amino group of one amino acid and the acid group of the other amino acid.

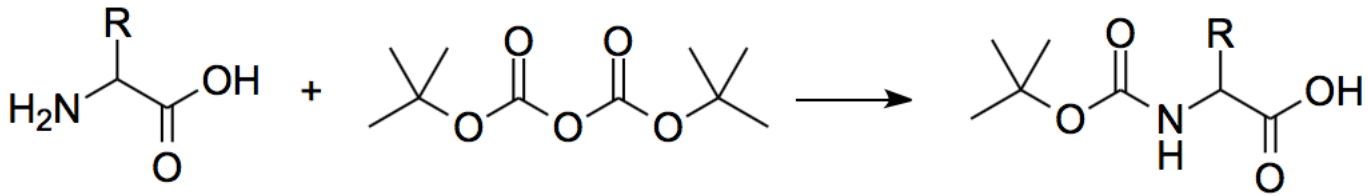


Peptide synthesis – Protecting groups – Boc protection

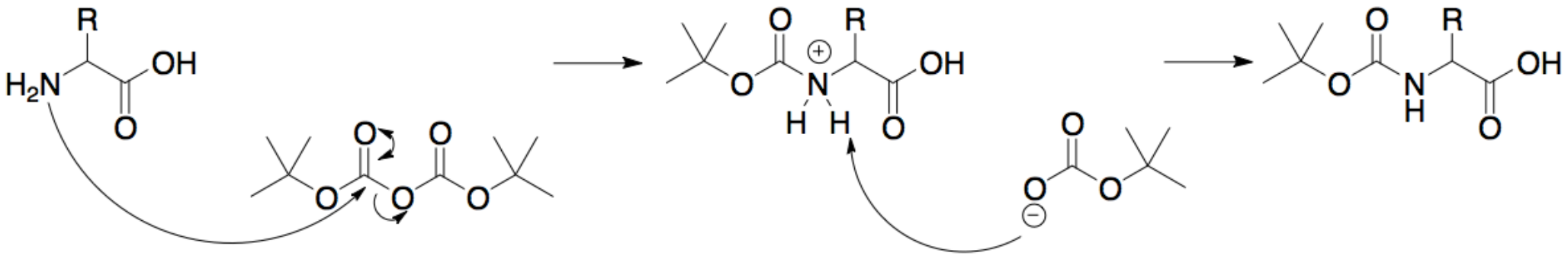
The amino group can be protected as its carbamate. The nitrogen is less nucleophilic as the lone pair is delocalised into the neighbouring carbonyl.



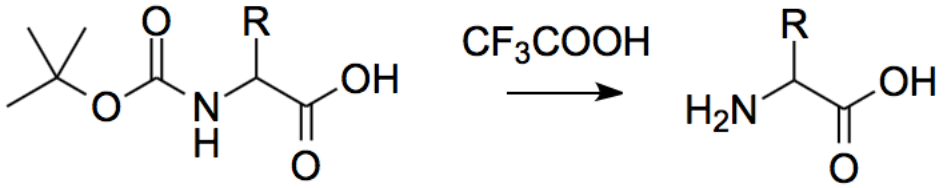
A common carbamate protecting group is the *tert*-butoxycarbonyl group (Boc).



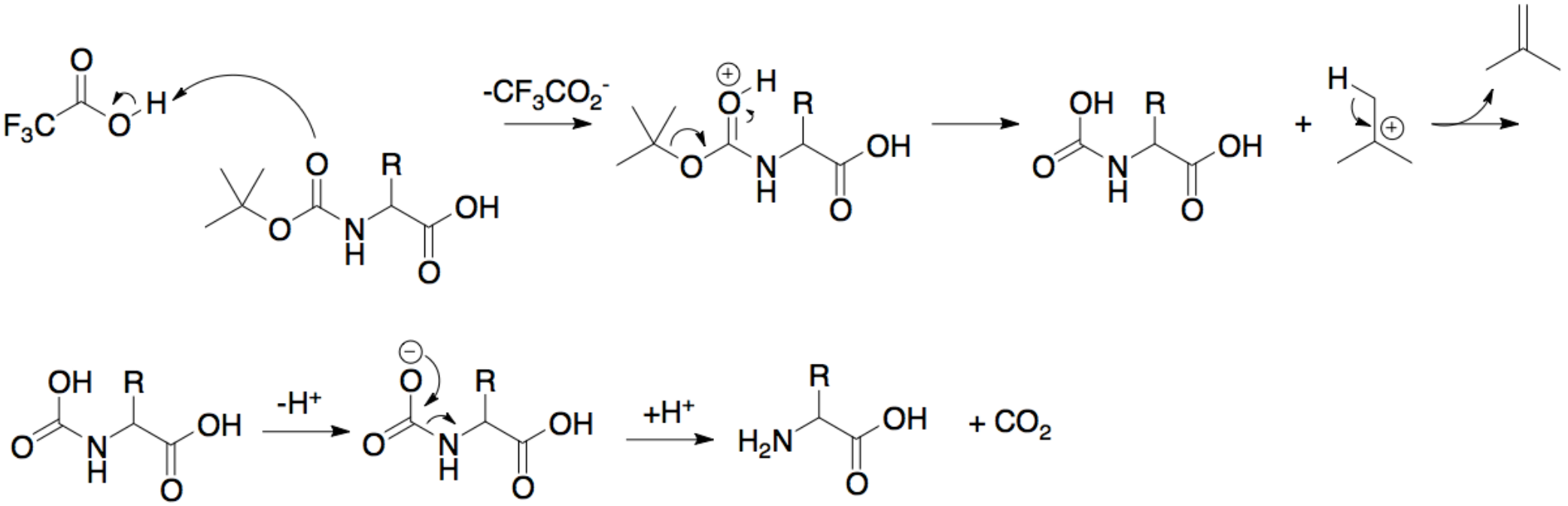
Peptide synthesis – Protecting groups – Boc protection



Peptide synthesis – Protecting groups – Boc deprotection

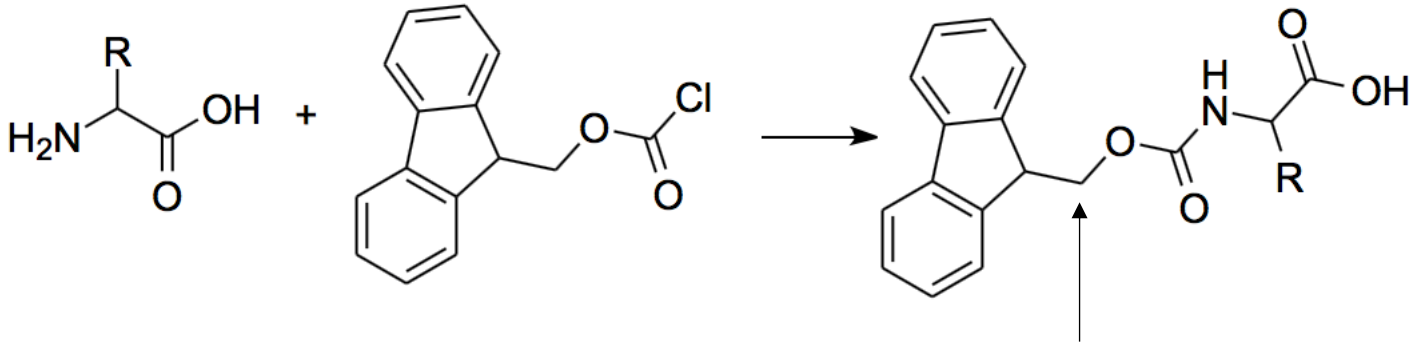


The Boc group is commonly used because it can be removed easily.



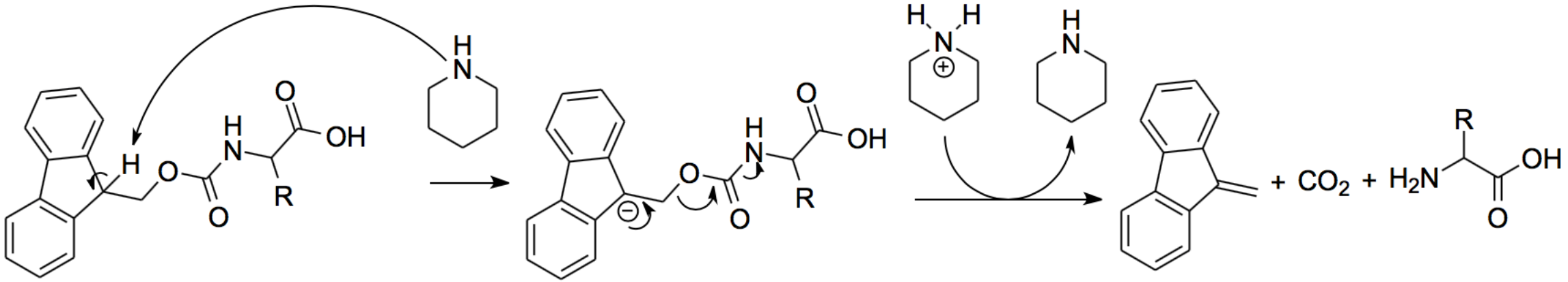
Peptide synthesis – Protecting groups – Fmoc protection

An alternative protecting group for the amino group is the fluorenylmethyloxycarbonyl (Fmoc).



The Fmoc group cannot be removed via an S_N1 or S_N2 mechanism as the carbon atom indicated with an arrow is both primary and hindered.

Peptide synthesis – Protecting groups – Fmoc deprotection

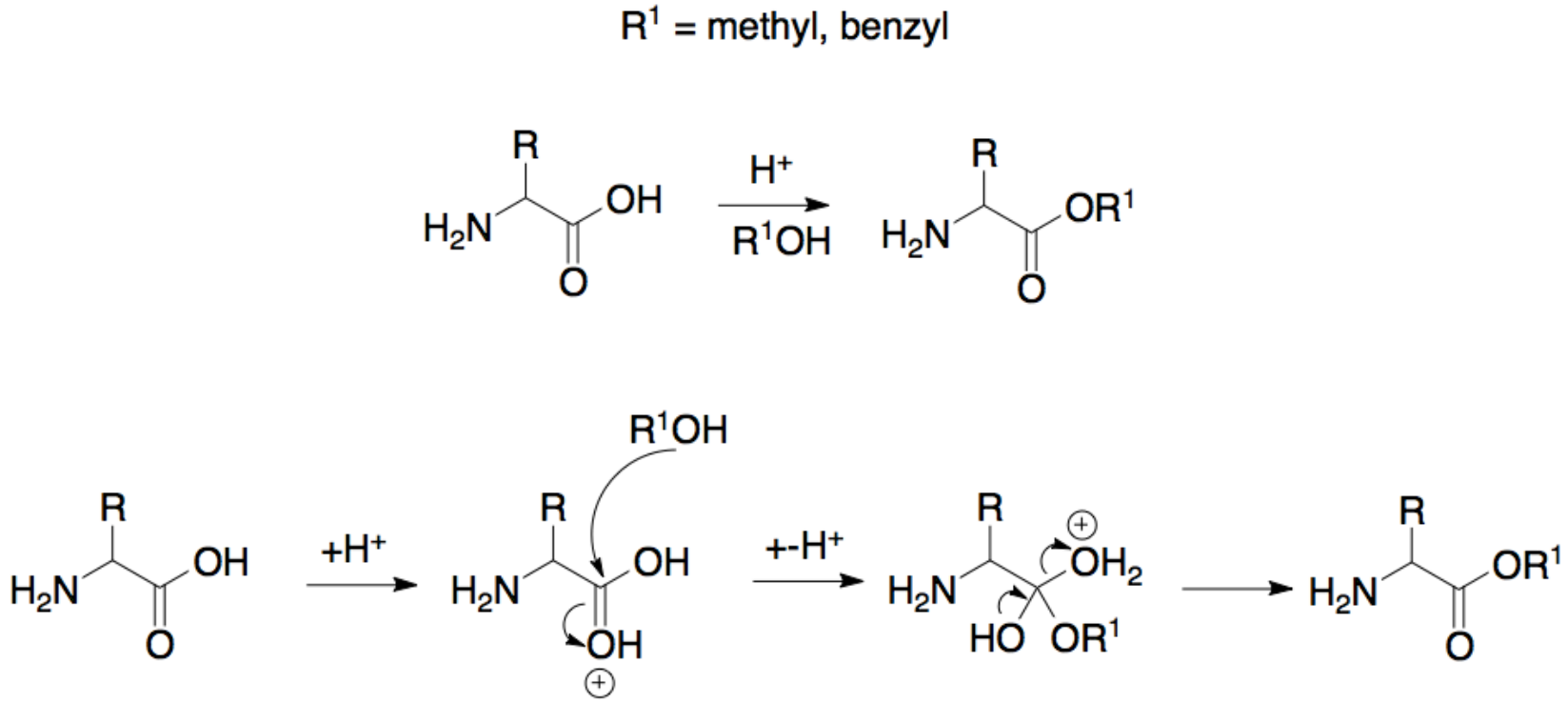


The Fmoc group can be removed as it has a fairly acidic proton, p$K_a \approx 25$

This proton is acidic because the conjugate base is aromatic.

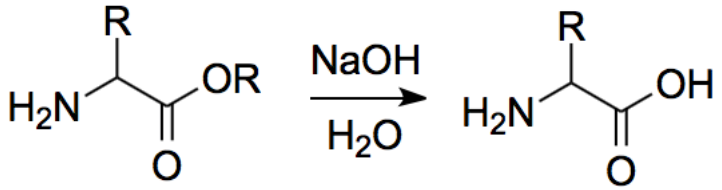
Peptide synthesis – Protecting groups – Esters

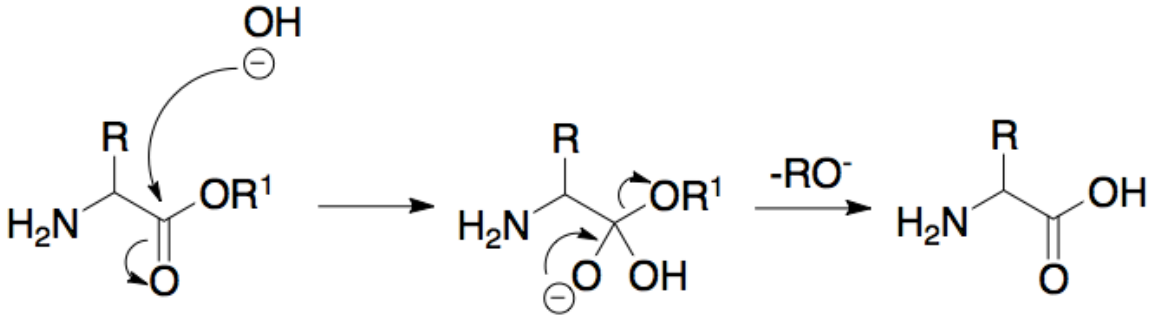
Protecting the carboxylic acid as an ester is the simplest way to prevent the CO_2H groups acting as an acid or a nucleophile.



Peptide synthesis – Protecting groups – Esters

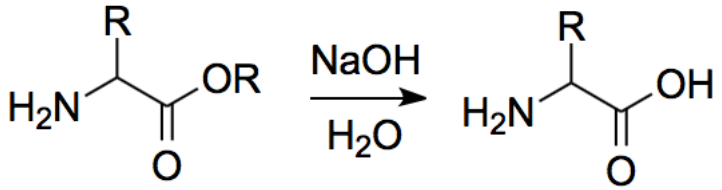
An ester group is easily deprotected by hydrolysis.

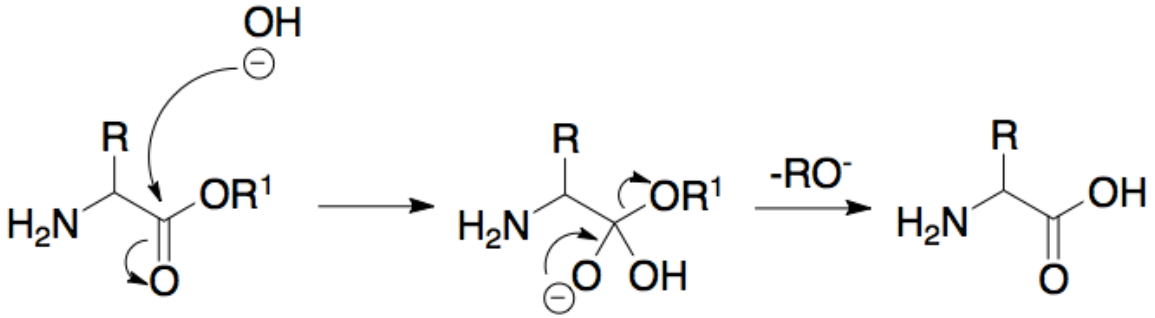




Peptide synthesis – Protecting groups – Esters

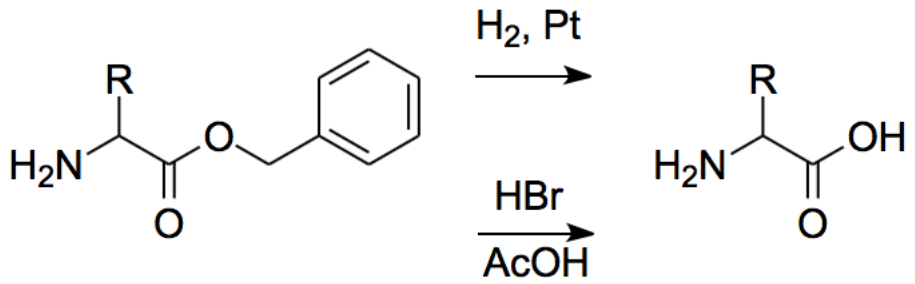
An ester group is easily deprotected by hydrolysis.

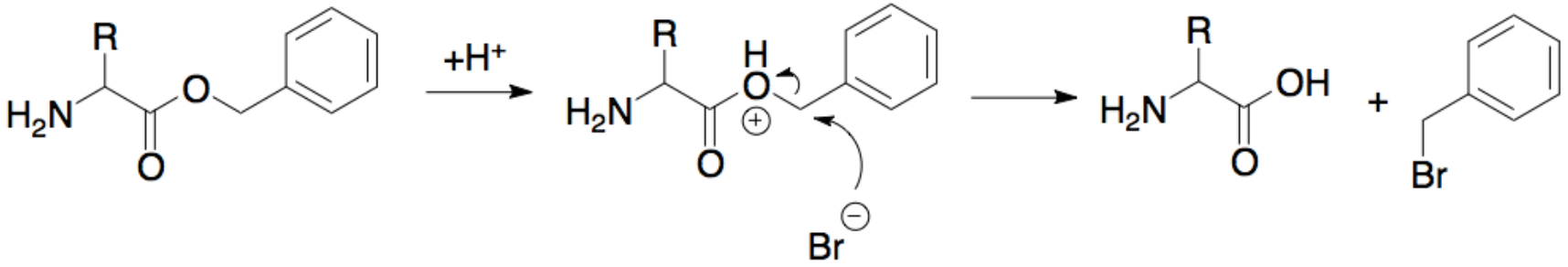




Peptide synthesis – Protecting groups – Esters

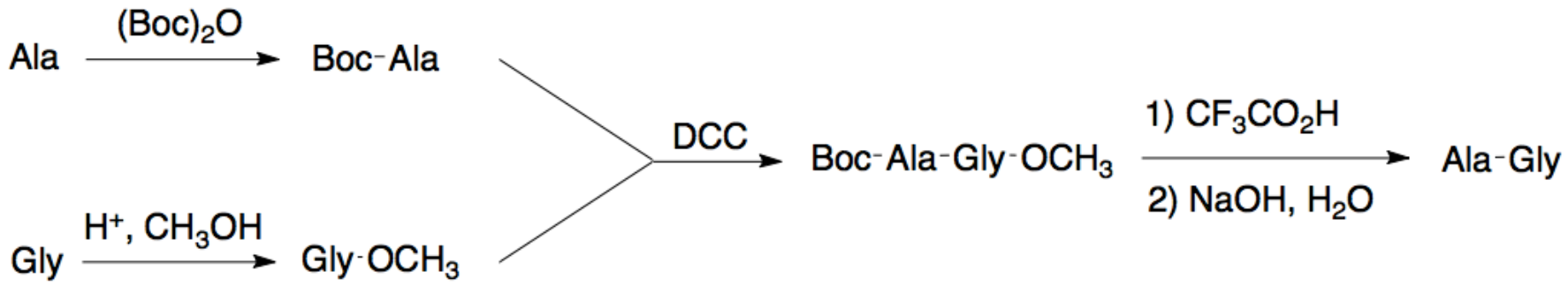
A benzyl ester group can be deprotected by hydrogenolysis or treatment with HBr/AcOH.





Peptide synthesis – Synthesis of short peptide chains

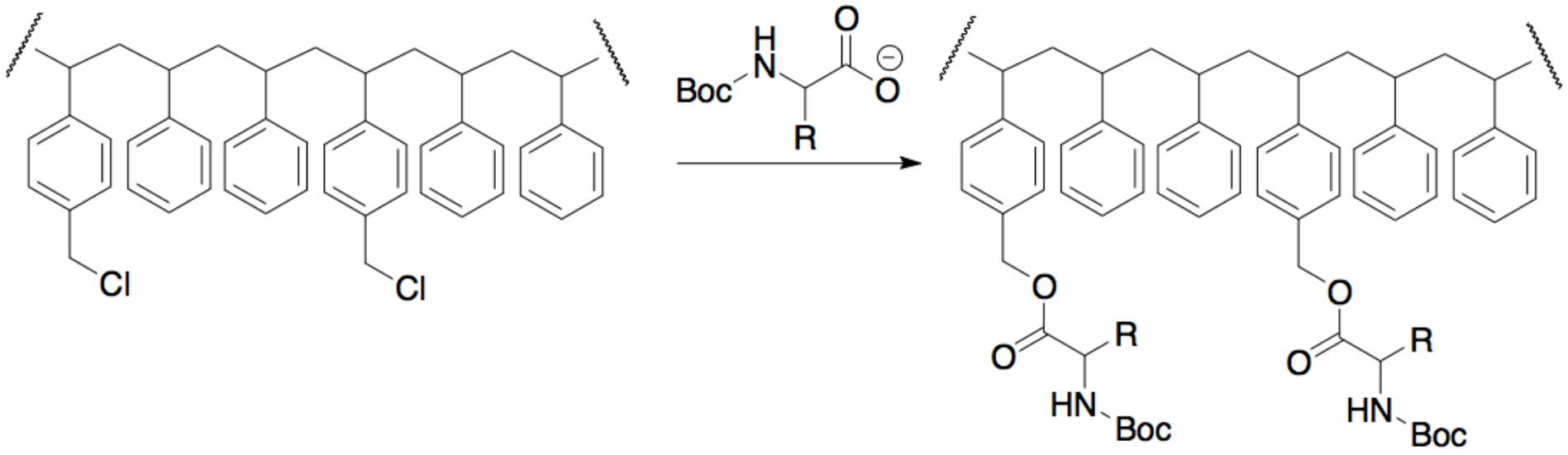
Protecting group chemistry allows for the preparation of short peptide chains such as di-, tri- and tetra-peptides.



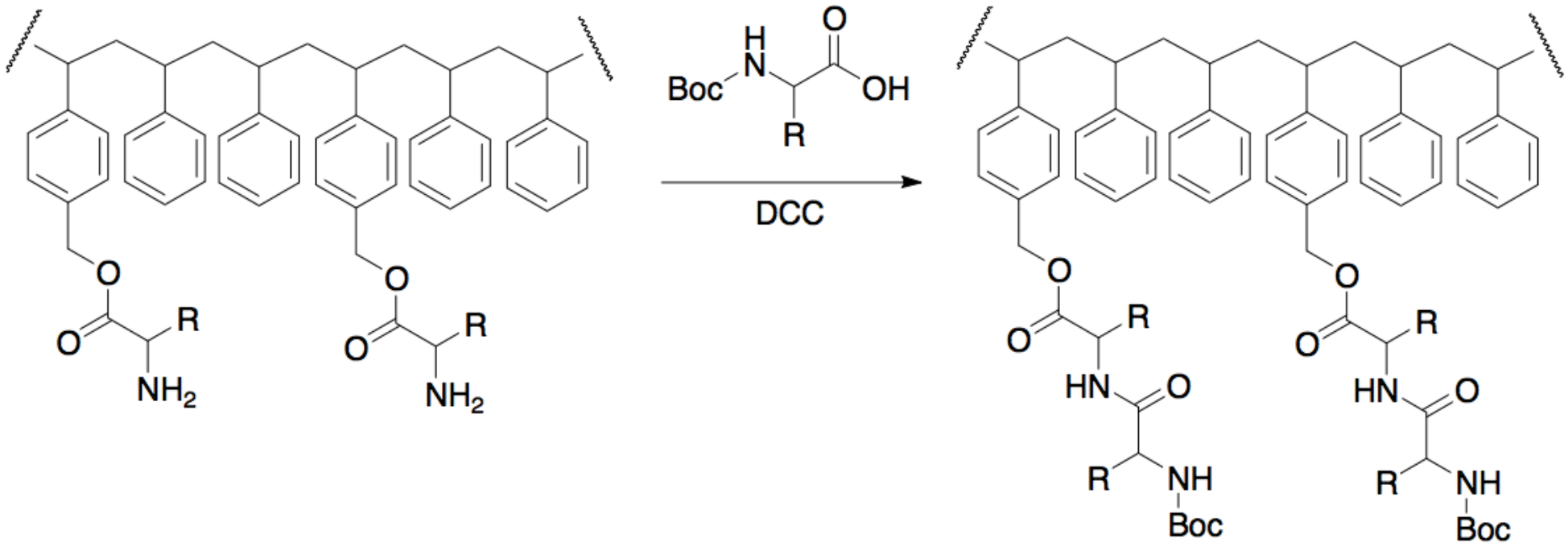
Each step requires isolation and purification and as such this approach is inappropriate for the synthesis of larger peptides.

Peptide synthesis – The Merrifield Synthesis

- The Merrifield synthesis allows for the synthesis of large peptides.
- A protected amino acid is tethered to beads of an insoluble polymer via an S_N2 reaction.

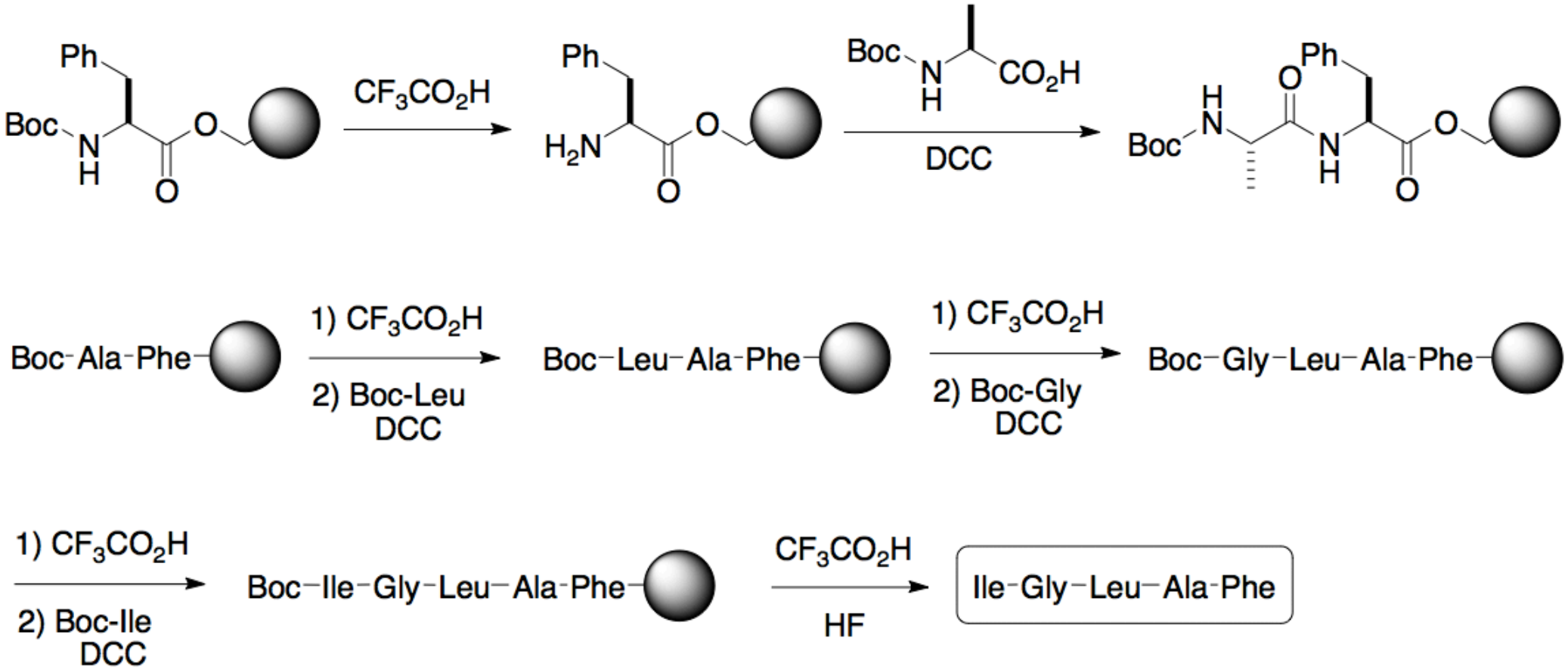


Peptide synthesis – The Merrifield Synthesis



- The Boc protecting group is removed, and DCC is used to couple the next Boc-protected amino acid.
- Following the coupling, impurities and by products are simply washed away as the peptide chain is tethered to the polymer.
- The process can be repeated until the desired polypetide has been made.

Peptide synthesis – The Merrifield Synthesis



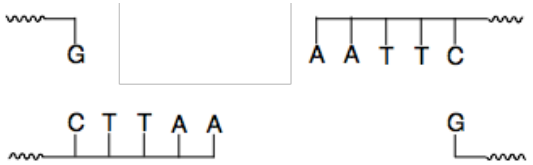
Peptide synthesis – Genetic engineering

- It is possible to incorporate specific genes into the DNA of fast-growing cells so that the proteins encoded by them are synthesised in the modified cell.
- Because the cells are fast-growing a significant quantity of the desired protein is synthesised allowing for its isolation and purification.

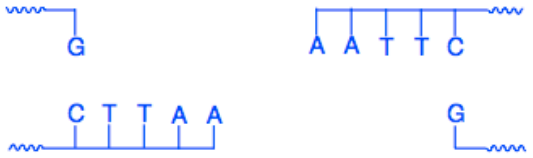
Peptide synthesis – Recombinant DNA technology

- **Recombinant DNA technology** allows scientists to produce modified DNA. Natural enzymes called **restriction enzymes** and **ligases** are used.
- Restriction enzymes recognise particular sequences of bases in DNA and split specific sugar phosphate bonds. Restriction enzymes are used to split the DNA of the fast growing cells (vector DNA) and to prepare the DNA to be incorporated.

Vector DNA

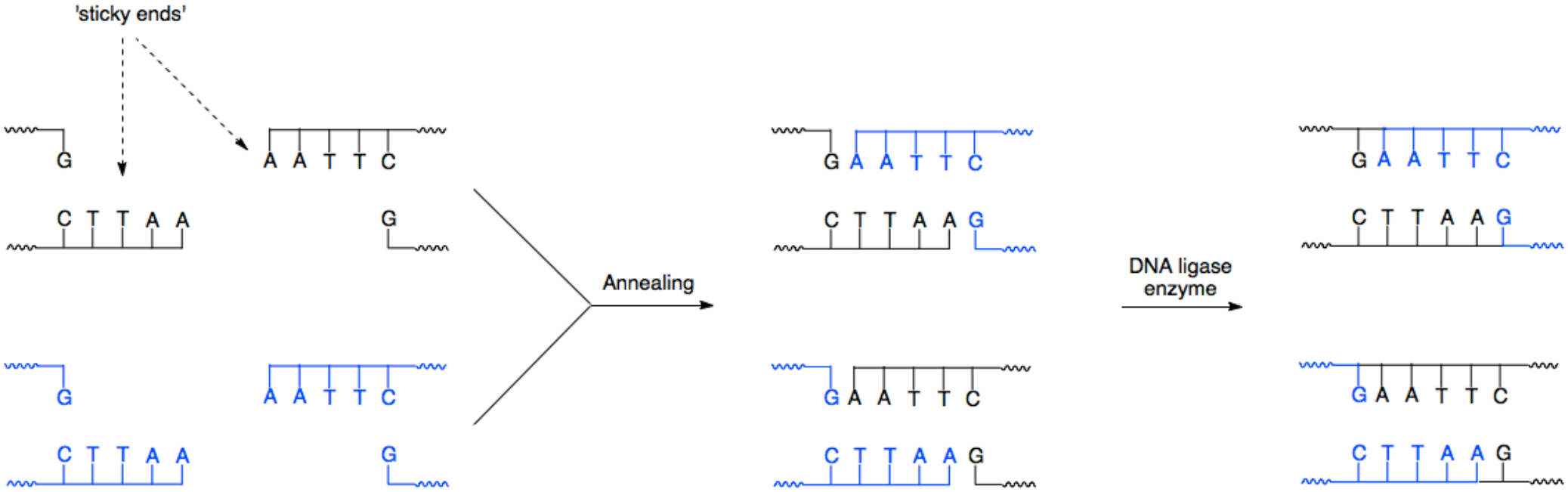


DNA to be incorporated

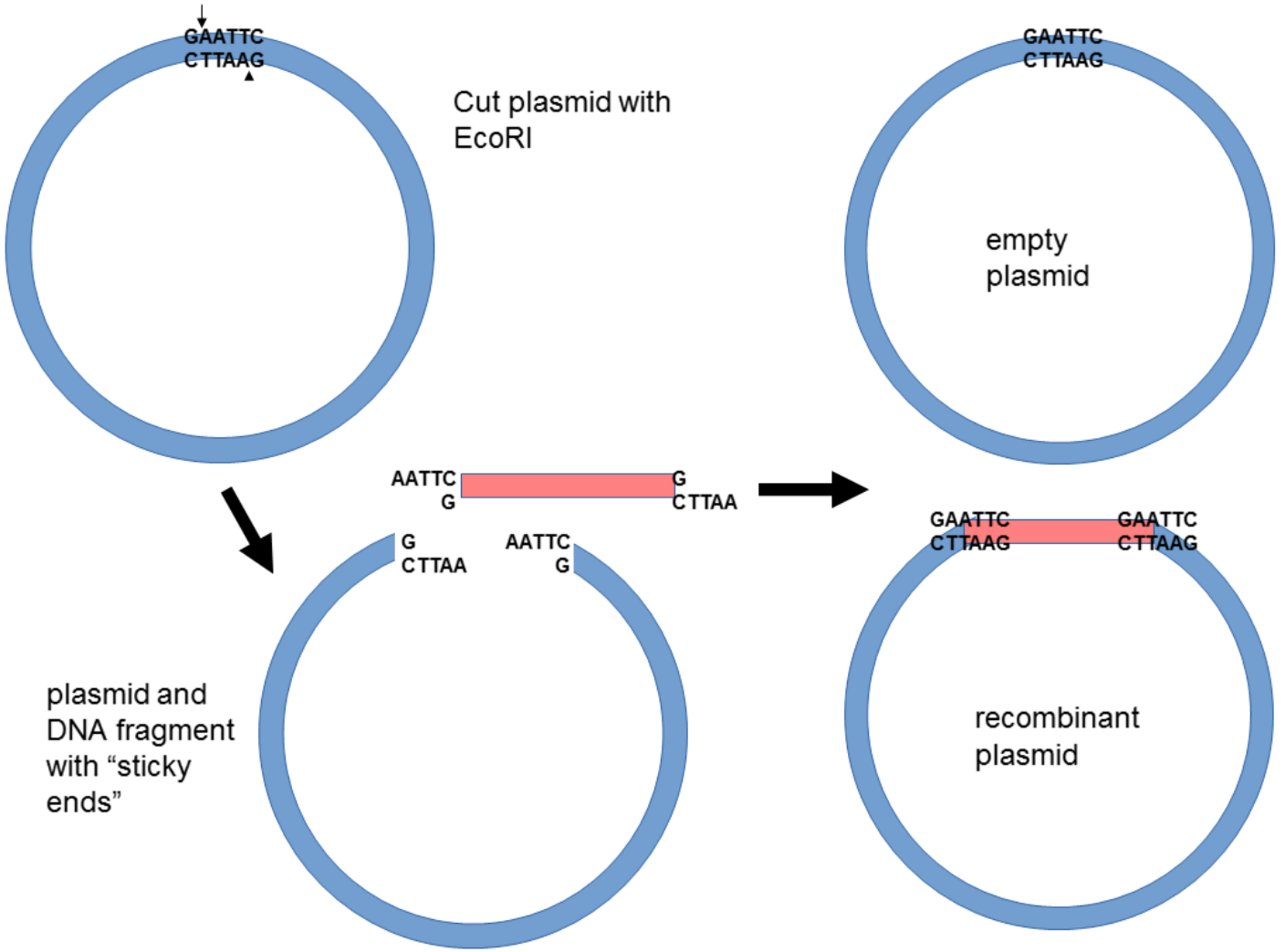


Peptide synthesis – Recombinant DNA technology

- The DNA molecules are mixed together. Because they have the same sticky ends, base pairing takes place in a process called **annealing**.
- Treatment with the ligase enzyme then repairs the sugar phosphate backbone and a new DNA molecule is formed.



Peptide synthesis – Recombinant DNA technology



Basic recombinant DNA techniques, http://bio1510.biology.gatech.edu/wp-content/uploads/2012/11/RecombinantPlasmid-1.png, Licence: Creative commons

Peptide synthesis – Uses of genetic engineering

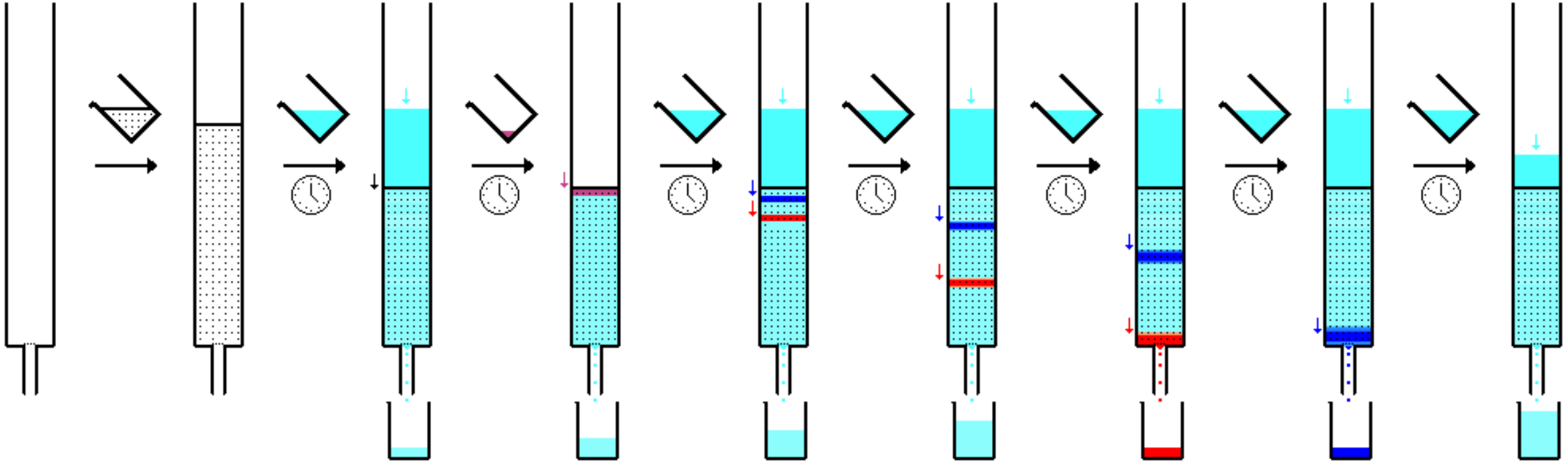
- Genetic engineering has been used to harvest important proteins such as insulin and human growth factor.
- The **Human Genome Project** (the mapping of human DNA which was completed in 2000) has led to the discovery of previously unknown proteins which can be synthesised in large quantities for further study.
- Many drug targets are proteins. Genetic engineering allows the structure of particular drug targets and its active site to be studied.
- Changing the DNA that encodes a given protein can shed light on the role of particular amino acids in enzyme catalysis or in binding interactions between a drug and the target protein.

Protein purification & analysis

Protein purification – Chromatography

- In most chromatographic processes there is a liquid phase (the mobile phase) and a solid phase (stationary phase).
- The solute mixture to be purified is dissolved in the liquid phase and this is passed through a column containing a porous solid phase.
- For chromatographic purification to be successful, the components of the mixture must move through the column at different speeds (have different affinities for the stationary phase).

Protein purification – Chromatography



Column chromatography sequence, https://commons.wikimedia.org/wiki/File:Column_chromatography_sequence.png, Licence: Creative commons

Protein purification – Ion exchange chromatography

- In ion exchange chromatography, the stationary phase is made up of charged molecules bound to a matrix (e.g. cellulose).
- Anions bind to cationic groups on an anion exchanger.
- Cations bind to anionic groups on a cation exchanger.
- Diethylaminoethyl, DEAE, is the most commonly used anion exchanger and CM, carboxymethyl is the most commonly used cation exchanger.

DEAE: matrix-$CH_2CH_2NH(CH_2CH_3)_2^+$

CM: matrix-$CH_2CO_2^-$

Protein purification – Hydrophobic interaction chromatography

- Hydrophobic interaction chromatography is used to purify non-polar molecules.
- The stationary phase is made up of octyl or phenyl groups attached to a matrix.
- Proteins with more hydrophobic residues move through the column more slowly.

Protein purification – Gel filtration chromatography

- In gel filtration chromatography, the stationary phase is made up of gel beads containing pores that span a relatively narrow size range.
- The molecules in the mixture are separated according to their size.
- Molecules which are too large to pass through the pores cannot pass through the beads and so traverse the column more quickly as they pass in between the beads.

Protein purification – Affinity chromatography

- In affinity chromatography, the stationary phase is made up of molecules which bind specifically to the protein of interest attached to an inert matrix.
- The molecule could be an unreactive analogue of an enzyme substrate.

Protein purification – Electrophoresis

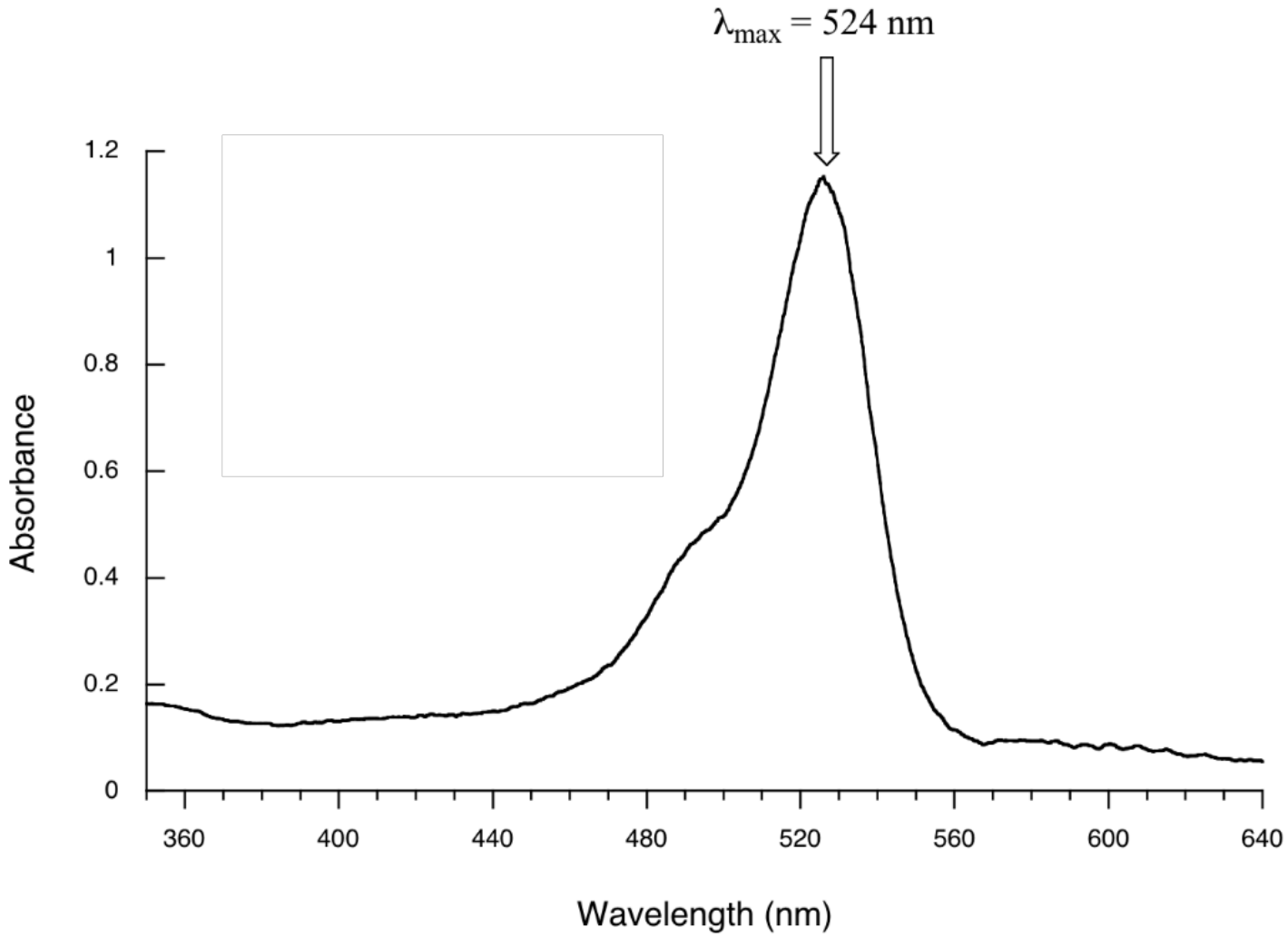
- Electrophoresis separates molecules according to size and charge.
- In gel electrophoresis the medium is a gel (typically polyacrylamide).
- The mixture is placed in wells in the centre of the gel and a voltage is applied.
- Cations move to the cathode and anions move to the anode. It should be noted that the pH of the gel is usually about 9 and so nearly all proteins have a net negative charge.

Protein assays – UV-vis absorbance spectroscopy

- Protein concentrations can be determined by **absorbance spectroscopy**.
- The technique can be divided into two separate parts:

1. A UV-vis spectrum is run to measure the absorbance of the protein or of the protein after a dye has been attached to make it a coloured compound. From the resulting spectrum, the wavelength of maximum absorbance, λ_{max} can be determined.

Protein assays – Maximum absorbance, λ_{max}



UV-vis spectrum, https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Map%3A_Organic_Chemistry_(McMurry)/14%3A_Conjugated_Compounds_and_Ultraviolet_Spectroscopy/14.09%3A_Interpreting_Ultraviolet_Spectra-_The_Effect_of_Conjugation,

Protein assays – Beer-Lambert Law

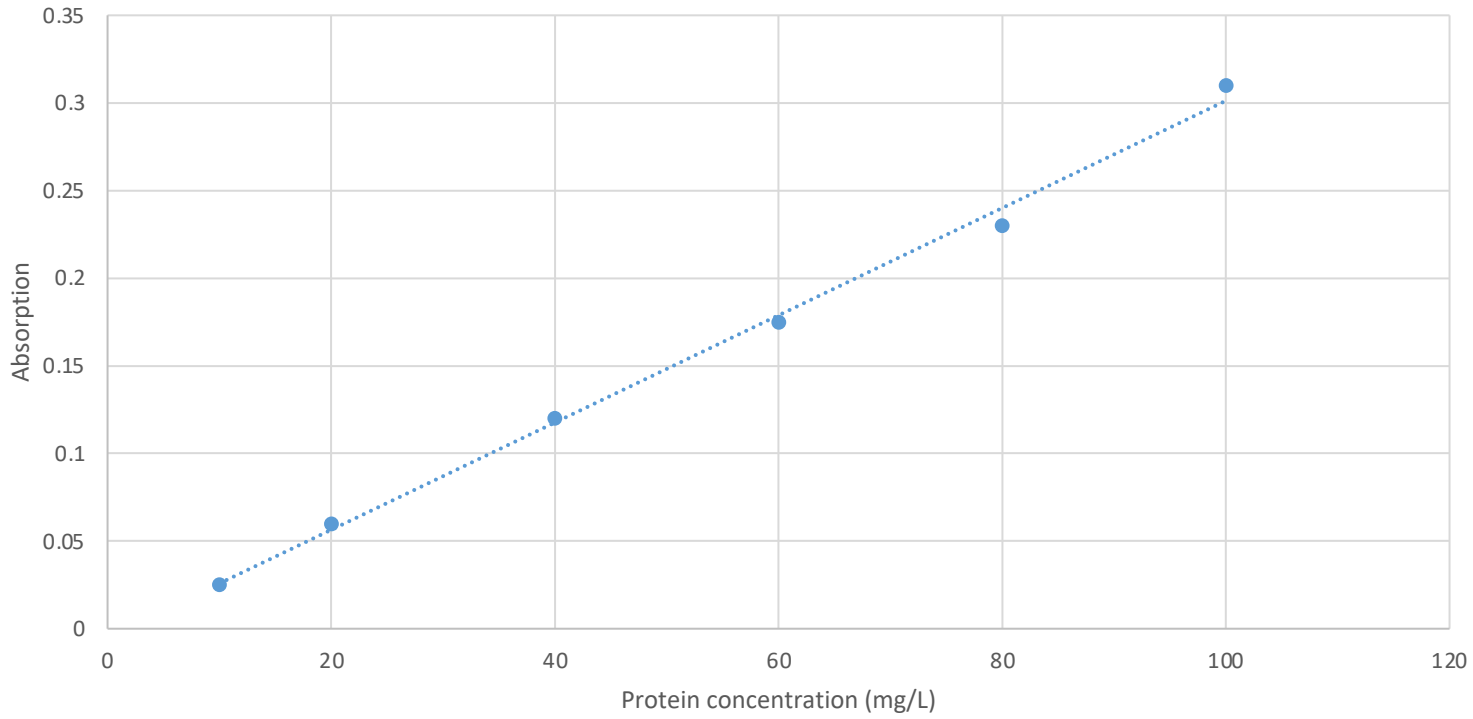
2. Solutions of known different concentrations of the protein are prepared and their absorbance measured at the wavelength of λ_{max}. Dilute solutions absorb light according to the Beer-Lambert Law where I_0 is the intensity of the incident radiation, I is the intensity of the transmitted radiation, ε is the molar absorption coefficient (a constant for the absorbing substance), c is the concentration and l is the path length:

$$A = \log\left(\frac{I_0}{I}\right) = \varepsilon c l$$

Because the path length is kept constant and ε is a constant for the particular protein, the absorbance is directly proportional to the concentration and a callibration curve can be constructed.

Protein assays – Calibration curve

From the calibration curve, the concentration of the protein in a solution of unknown concentration can be determined by interpolation.



Protein sequencing

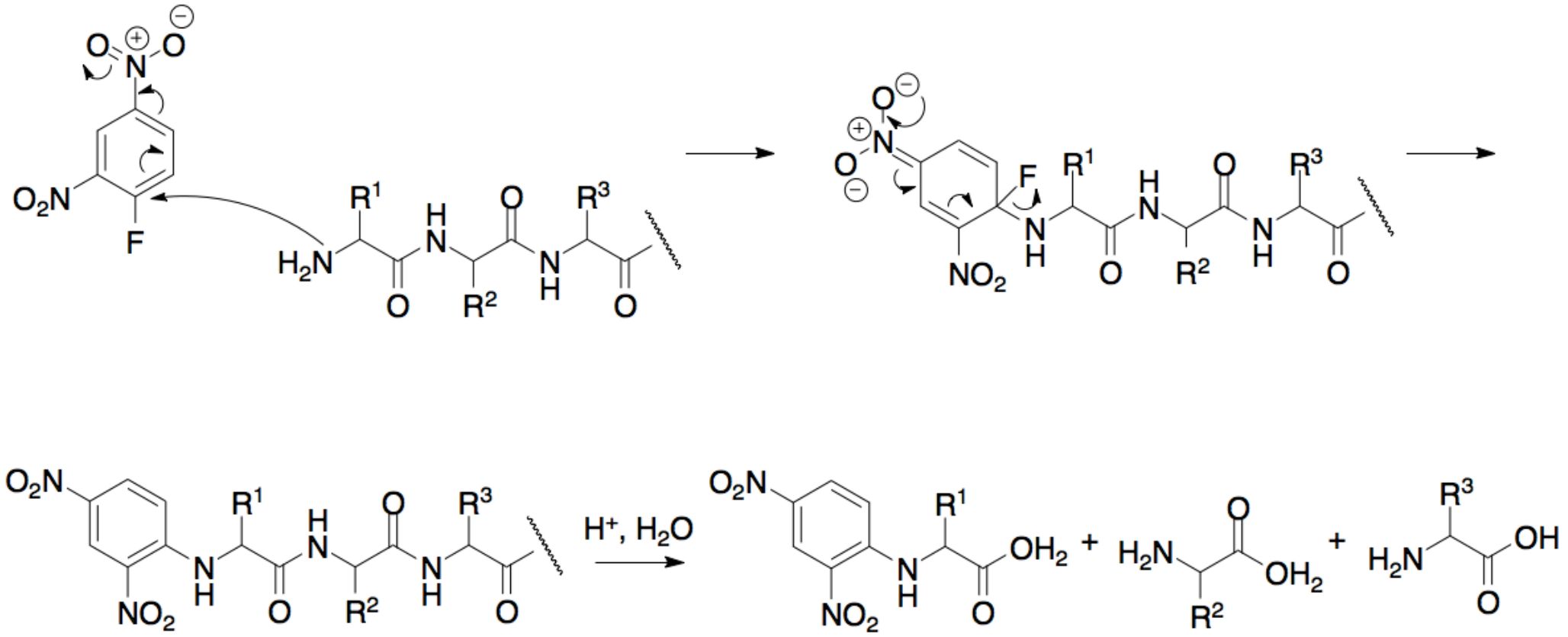
Structure determination – The primary sequence

Structure determination of proteins involves two steps:

1. The protein is first purified. Methods include chromatography or electrophoresis.
2. The polypeptide can then be sequenced using a variety of methods. The Sanger method and Edman degradation are widely used.

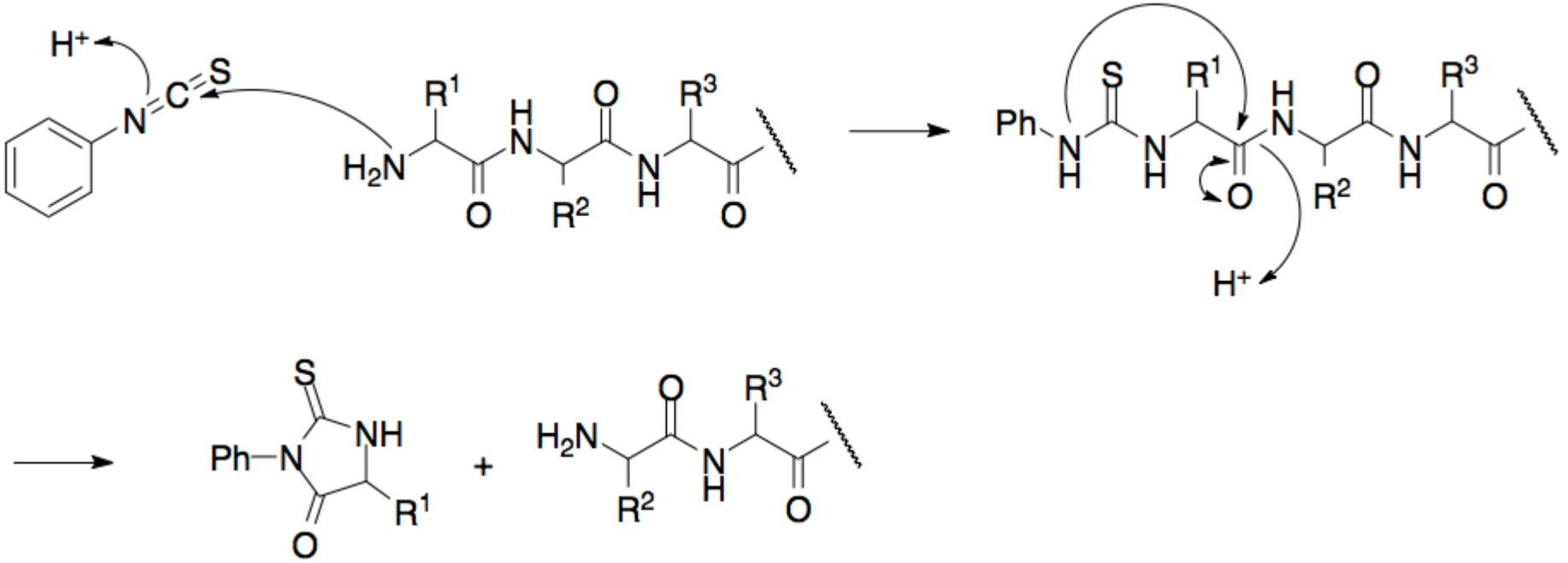
Determining the primary sequence – Sanger method

The Sanger method can be used to determine the amino acid at the *N*-terminus of the polypetide chain.



Determining the primary sequence – Edman degradation

- The Edman degradation can be used to determine the sequence of a polypeptide chain.
- The process has been fully automated, and peptide sequencers can sequence peptide chains with up to 50 amino acid residues.

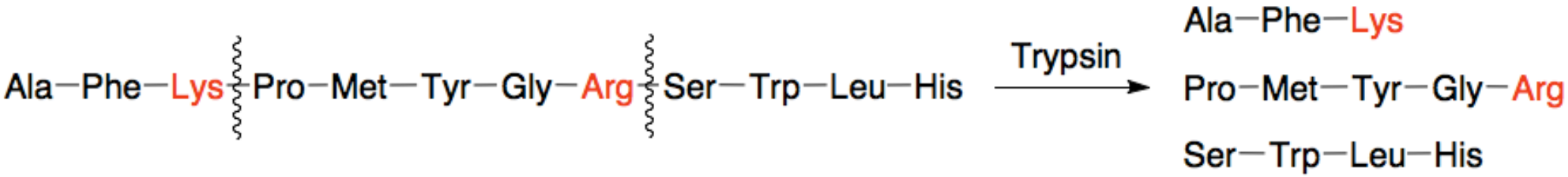


Determining the primary sequence – Enzymatic cleavage

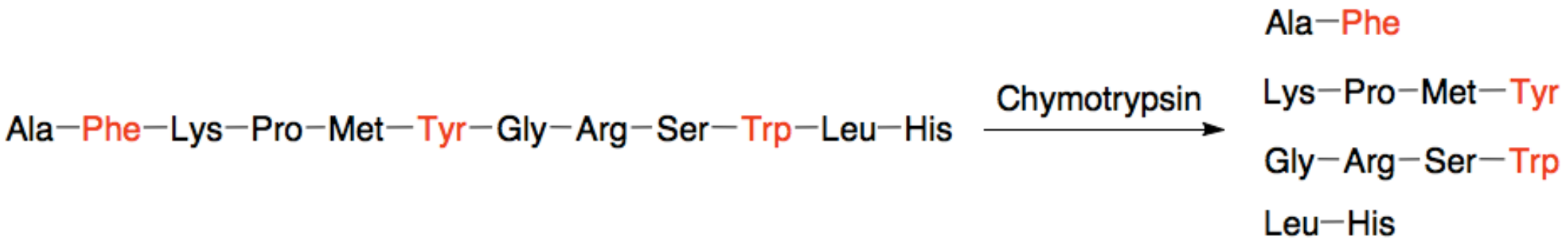
- For peptides with more than 50 residues, the removal of each residue one by one is impractical.
- Larger peptides are first cleaved into smaller fragments before being sequenced.
- Enzymes called peptidases, selectively hydrolyse specific peptide bonds.

Determining the primary sequence – Enzymatic cleavage

- Trypsin catalyses peptide hydrolysis at the carboxyl side of arginine and lysine amino acids.



- Chymotrypsin catalyses hydrolysis at the carboxyl end of amino acids containing aromatic side chains (phenylalanine, tyrosine, and tryptophan).

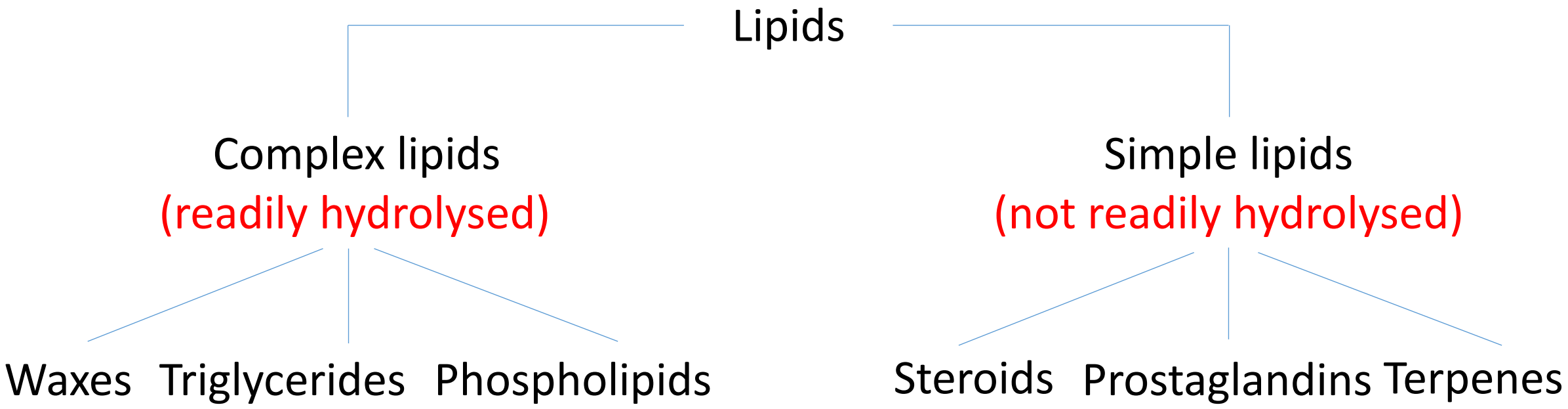


Lipids

Lipids

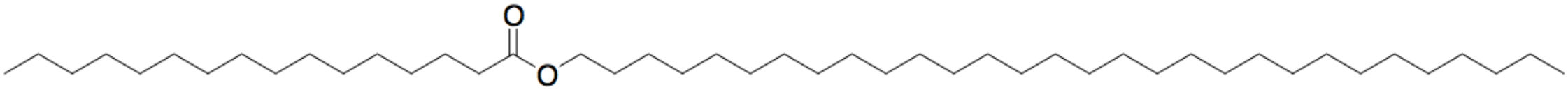
- Lipids are defined by their solubility. They are naturally occurring compounds that can be extracted from cells using non polar organic solvents.
- Many biological compounds are classified as lipids and so it is useful to categorise them.
- **Complex lipids** readily undergo hydrolysis in aqueous acid or base
- **Simple lipids** do not readily undergo hydrolysis.

Lipids - Categorisation



Lipids – Waxes

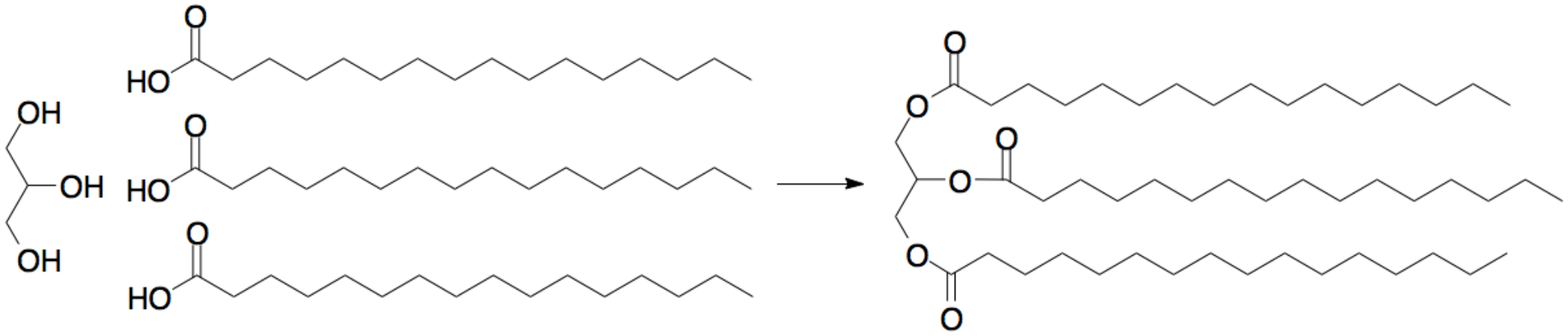
- Waxes are high molecular weight esters (produced from carboxylic acids and alcohols).
- Waxes have high melting points due to intermolecular London forces between hydrocarbon tails.
- Waxes are present on the surface of leaves reducing water loss by evaporation.
- Birds have waxes on their feathers for waterproofing as does the fur of some mammals (e.g. sheep).



Triacontyl hexadecanoate
(a major component of beeswax)

Lipids – Triglycerides

- Triglycerides are triesters produced from glycerol and three long-chain carboxylic acids (fatty acids).
- Used by plants and mammals for long term energy storage. They are less oxidised than sugars and so have greater potential energy (≈38 kJ/g as opposed to ≈17 kJ/g for carbohydrates).

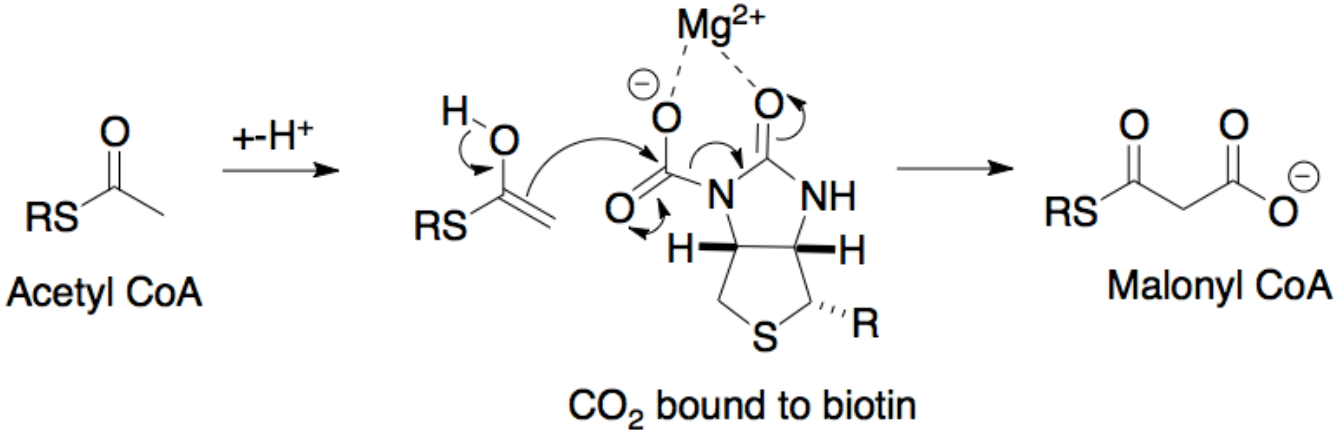


Lipids – Fatty acids

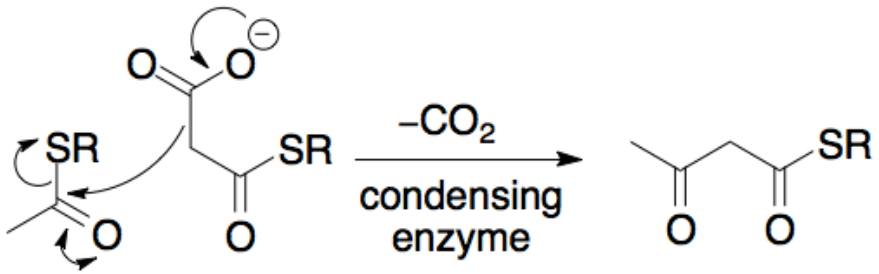
- The fatty acid residues of naturally occurring triglycerides are unbranched carboxylic acids typically containing 12 – 20 carbon atoms.
- They have even numbers of carbon atoms because they are biosynthesized from building blocks containing two carbon atoms.
- Some fatty acids are saturated and some are unsaturated, usually *cis.*

Lipids – Fatty acid biosynthesis

- Malonyl CoA is is biosynthesized by acylation of acetyl CoA with carbon dioxide (CO_2 is carried around the body as a covalent compound with the coenzyme biotin).

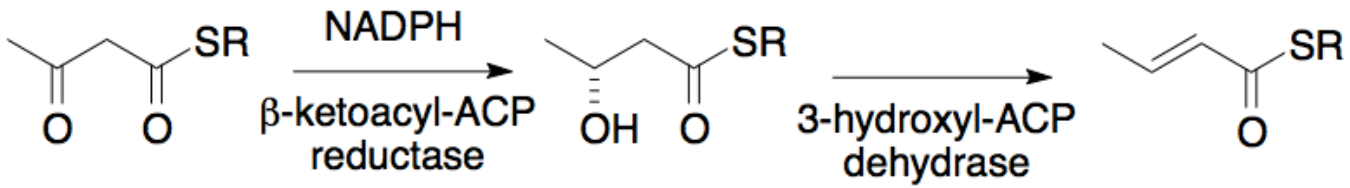


- The first stage in fatty acid biosynthesis is a condensation between acetyl CoA and malonyl CoA.

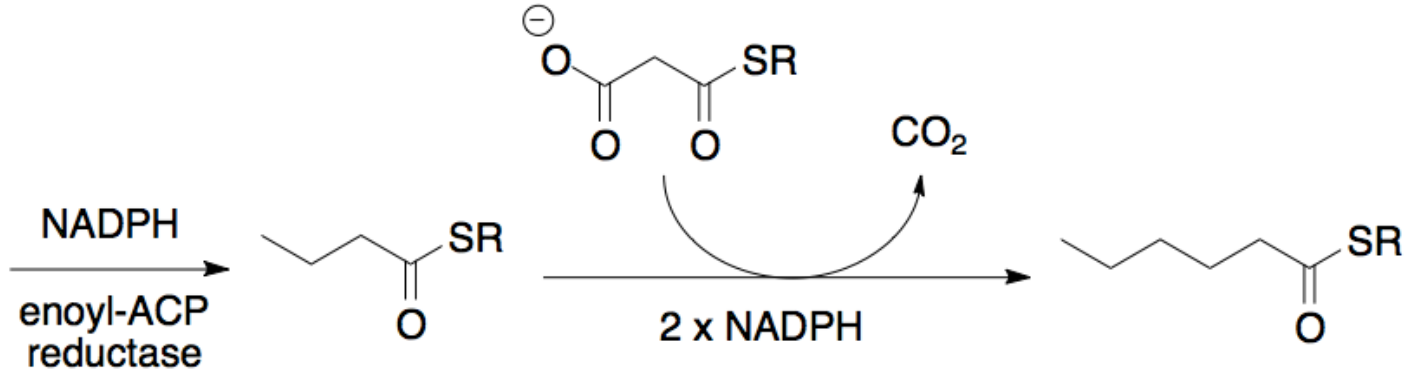


Lipids – Fatty acid biosynthesis

- Reduction of the ketone group by NADPH and dehydration *via* an enzyme catalysed E1cB reaction follows.

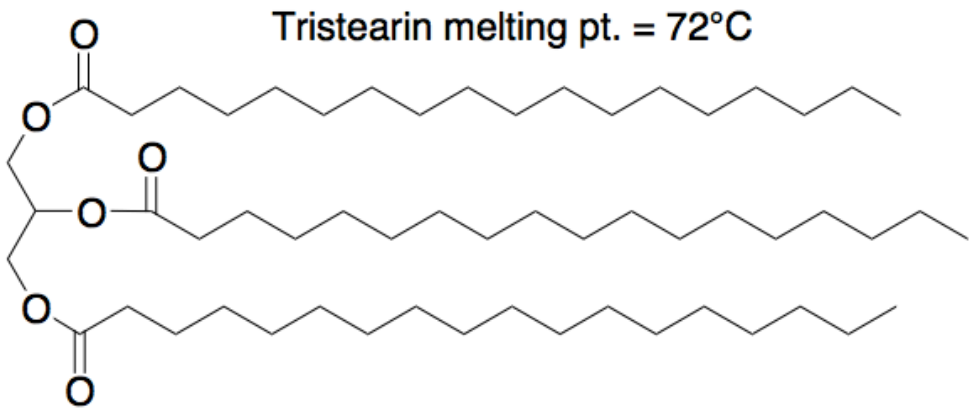


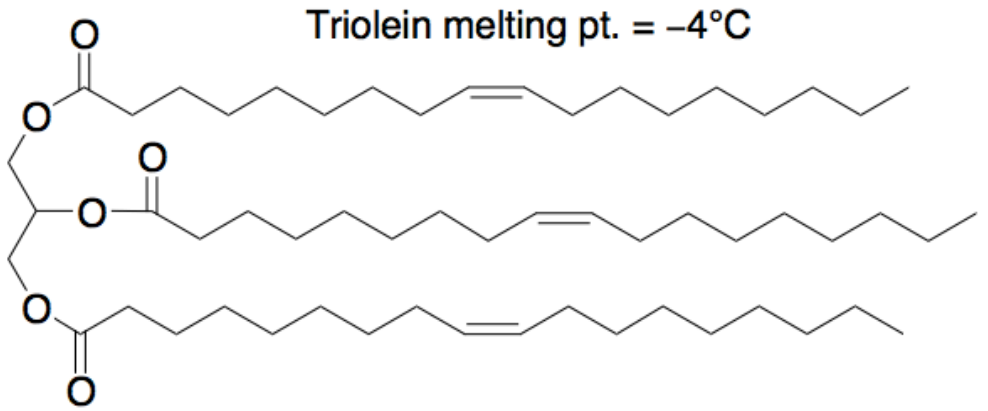
- The resulting double bond is then reduced by NADPH leaving a C4 unit. The cycle can be repeated to make a C6 unit etc. 2 carbon atoms are added to the acyl end each time.



Lipids – The melting points of fatty acids and triglycerides

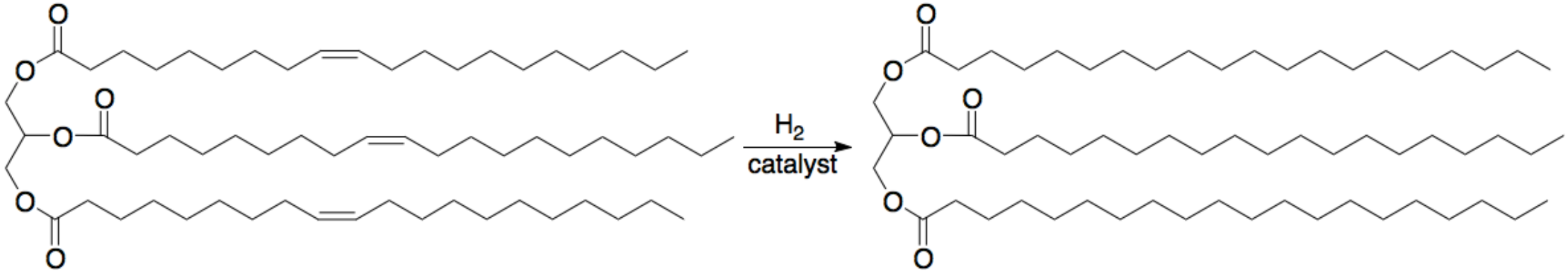
- The melting point of fatty acids increases with increasing chain length as the strength of intermolecular London forces increases with increasing molecular weight.
- The presence of a *cis* double bond in the chain introduces a kink in the chain reducing the ability of the fatty acids to pack tightly. This decreases the strength of the intermolecular London forces and lowers the melting point.
- These two trends are also observed for triglycerides.



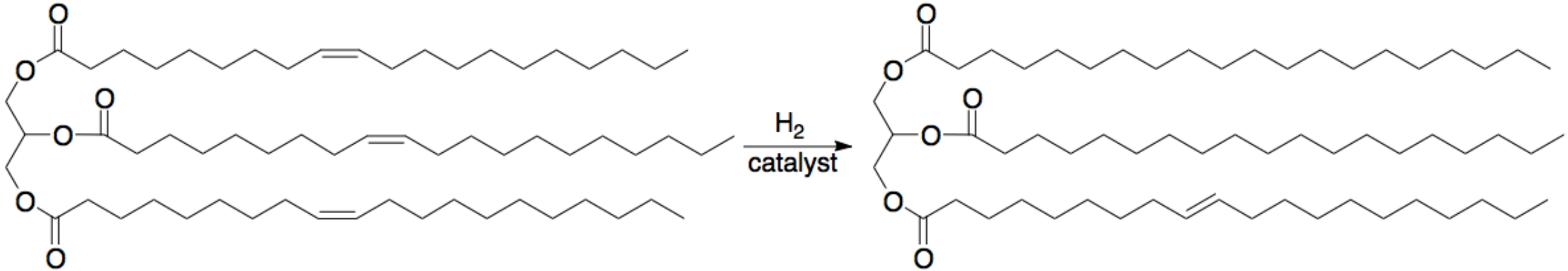


Lipids – Hydrogenation of triglycerides

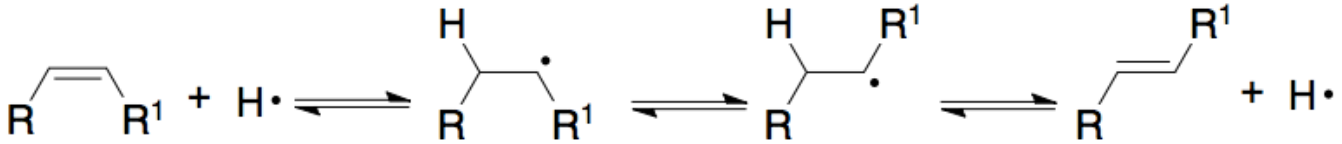
- Triglycerides containing unsaturated fatty acid residues undergo hydrogenation in the presence of a catalyst such as nickel.



- Instead of being hydrogenated, some of the double bonds can isomerise (*via* a half hydrogenated intermediate) to give *trans* π bonds.

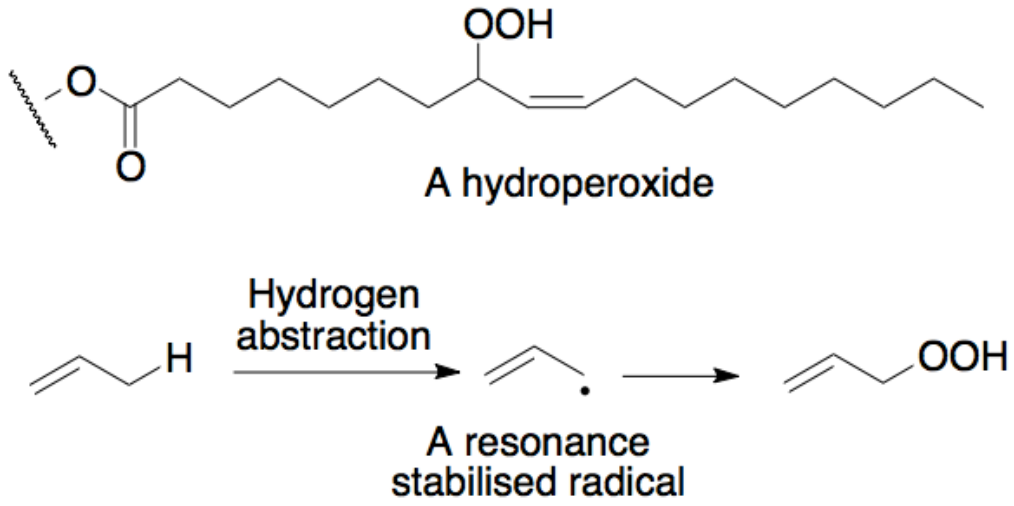


Lipids – *trans* fats *via* a half hydrogenated intermediate

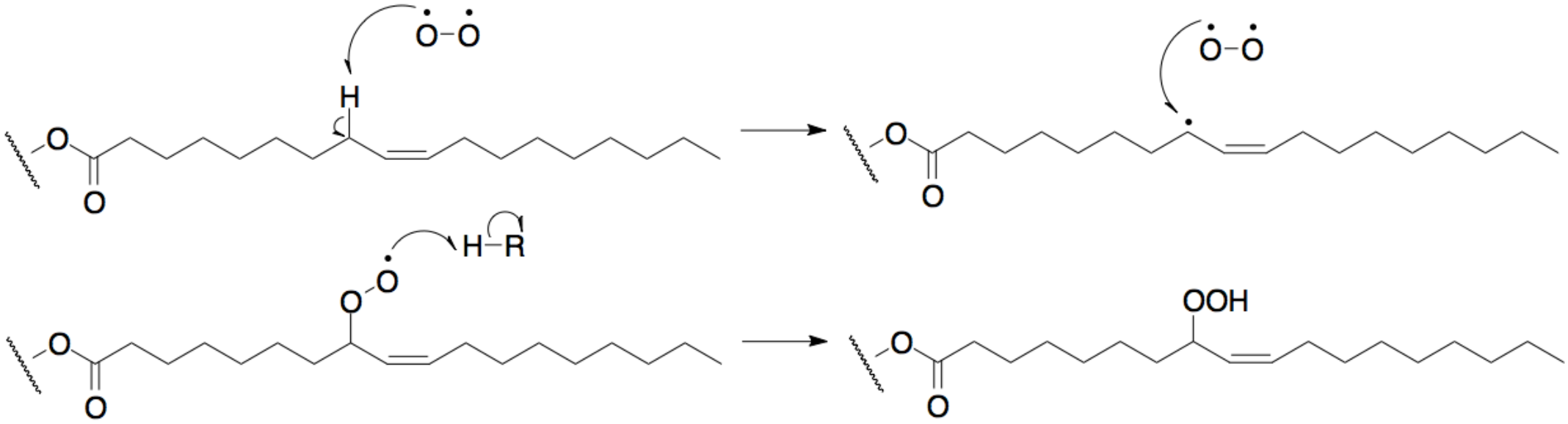


Lipids – Auto-oxidation of triglycerides

- In the presence of molecular oxygen, O_2, triglycerides with unsaturated fatty acid residues are susceptible to auto-oxidation.
- Auto-oxidation takes place at the allylic position as the allyl radical is resonance stabilised.

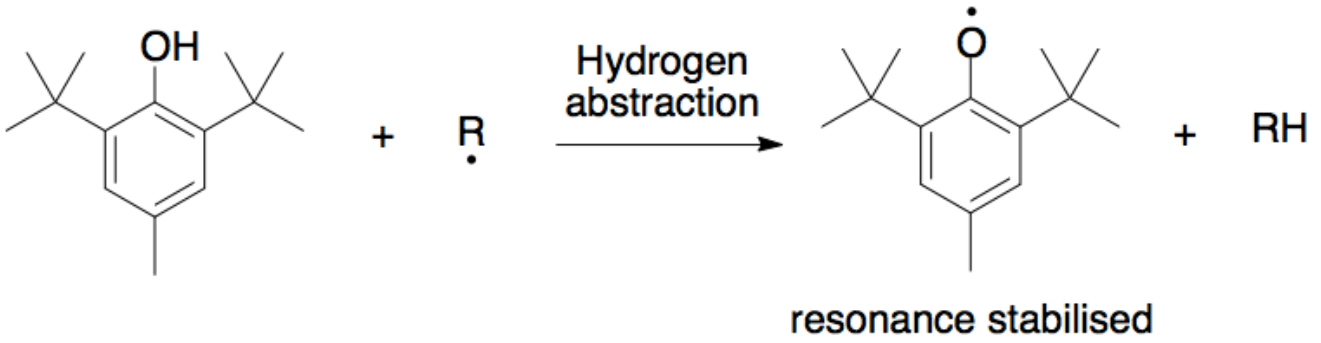


Lipids – Auto-oxidation of triglycerides



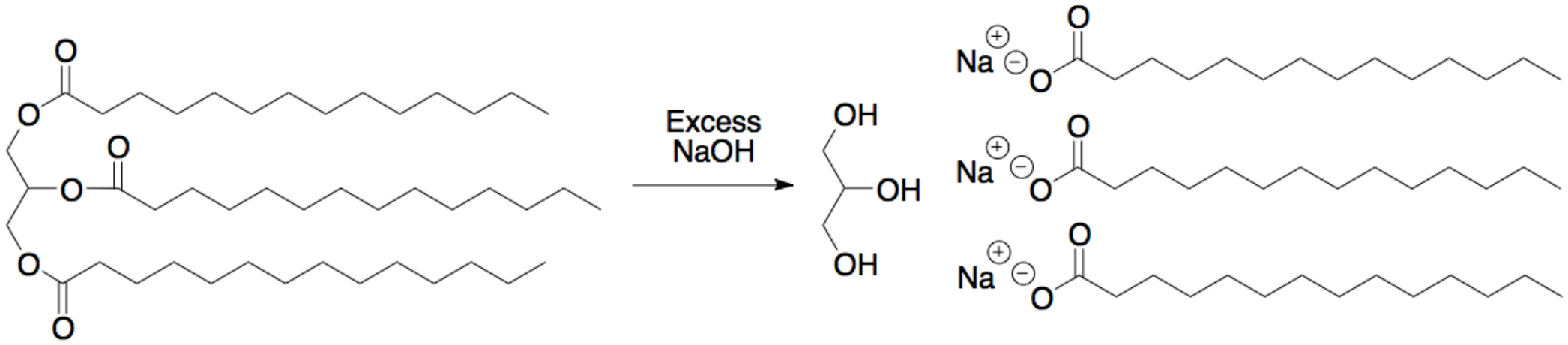
Lipids – Auto-oxidation of triglycerides

- The hydroperoxides contribute to the rancid smell that develops in foods containing unsaturated oils.
- These hydroperoxides are also toxic.
- Radical inhibitors such as BHT are added as food preservatives to slow the auto-oxidation process.
- These radical inhibitors react with radicals to generate persistent radicals which are less reactive due to resonance stabilisation and the steric hindrance provided by the *tert*-butyl groups.

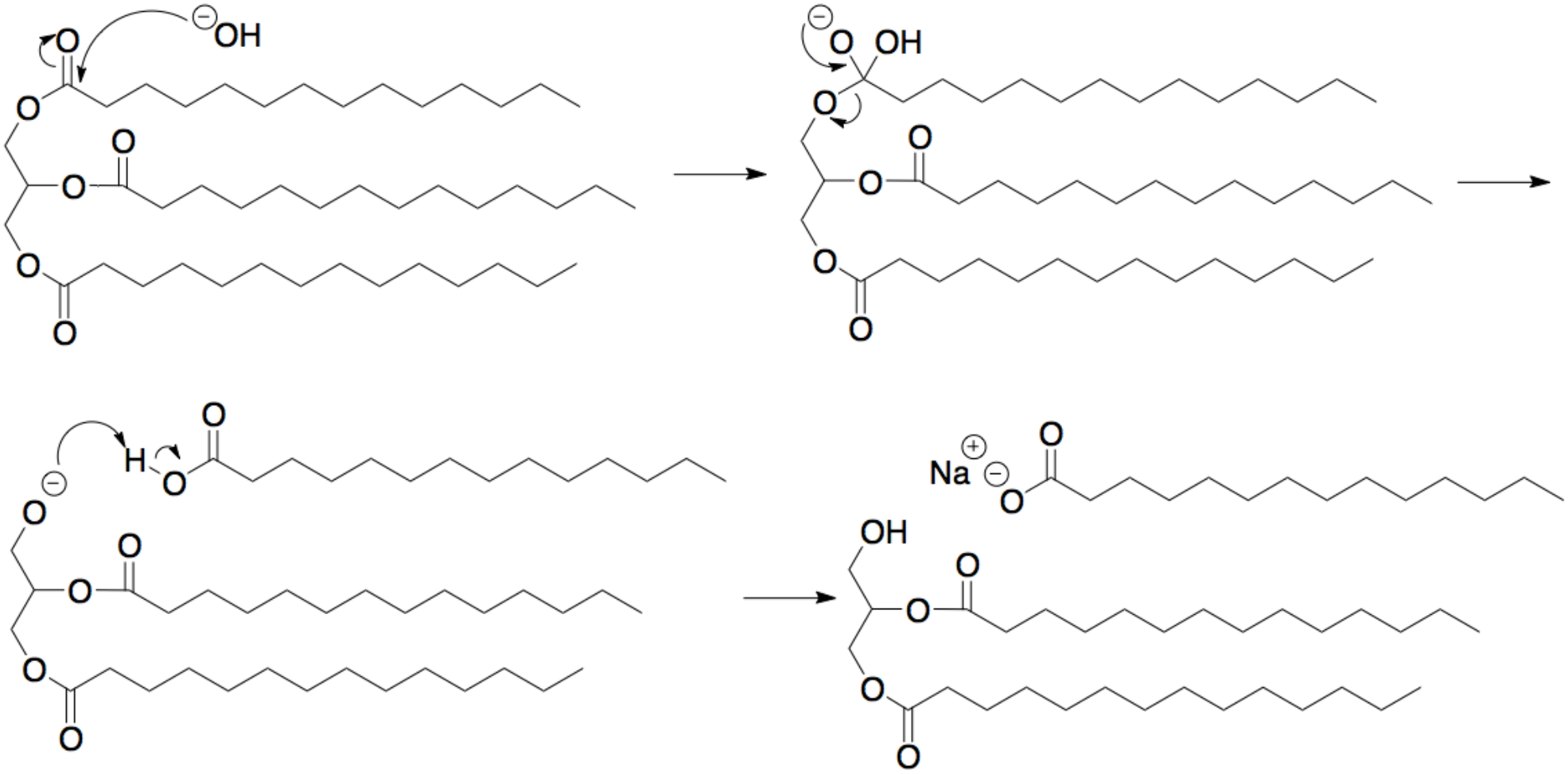


Lipids – Hydrolysis of triglycerides – Soap

- When triglycerides are boiled under alkaline conditions, hydrolysis of the esters takes place resulting in a mixture of carboxylate salts and glycerol.
- Soap is made by boiling either animal fat or vegetable oil together with a strong alkaline solution, such as aqueous sodium hydroxide.

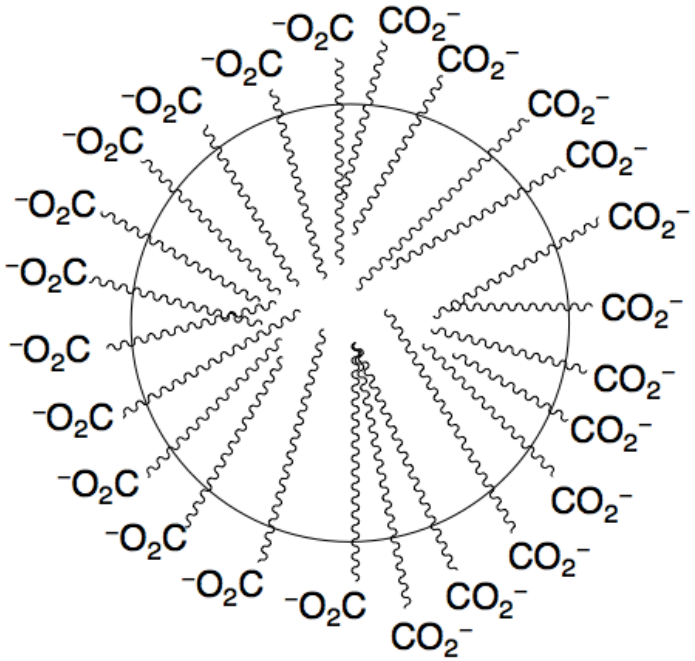


Lipids – Hydrolysis of triglycerides – Soap



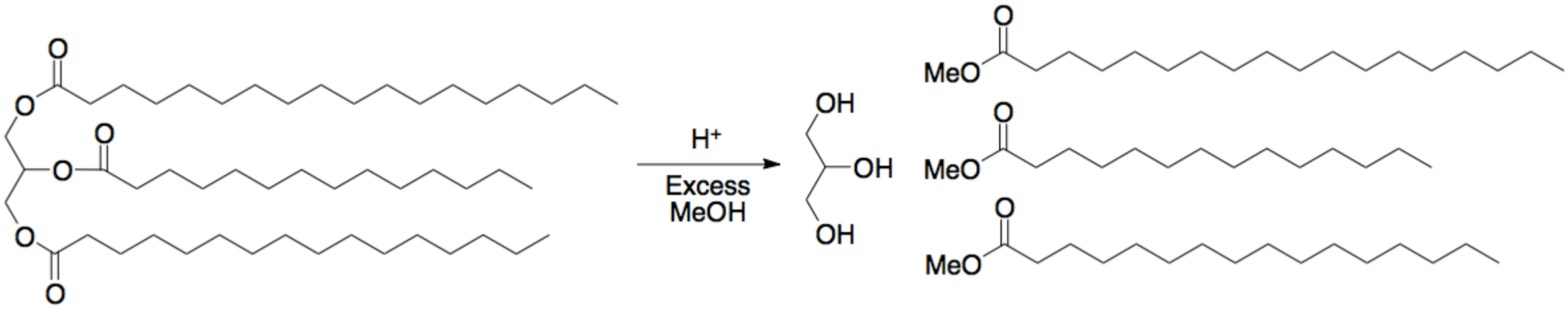
Lipids – Hydrolysis of triglycerides – Soap

- When suspended in water, soaps assemble around nonpolar substances such as oily globules to form spheres called *micelles.*
- The nonpolar substance interacts with the nonpolar ends of the soap molecules *via* intermolecular London forces.
- The surface of the micelle is comprised of the polar carboxylate groups, which interact with the water.
- The micelle is thus solvated by the water and so soap molecules can solvate nonpolar substances, such as grease, in water.

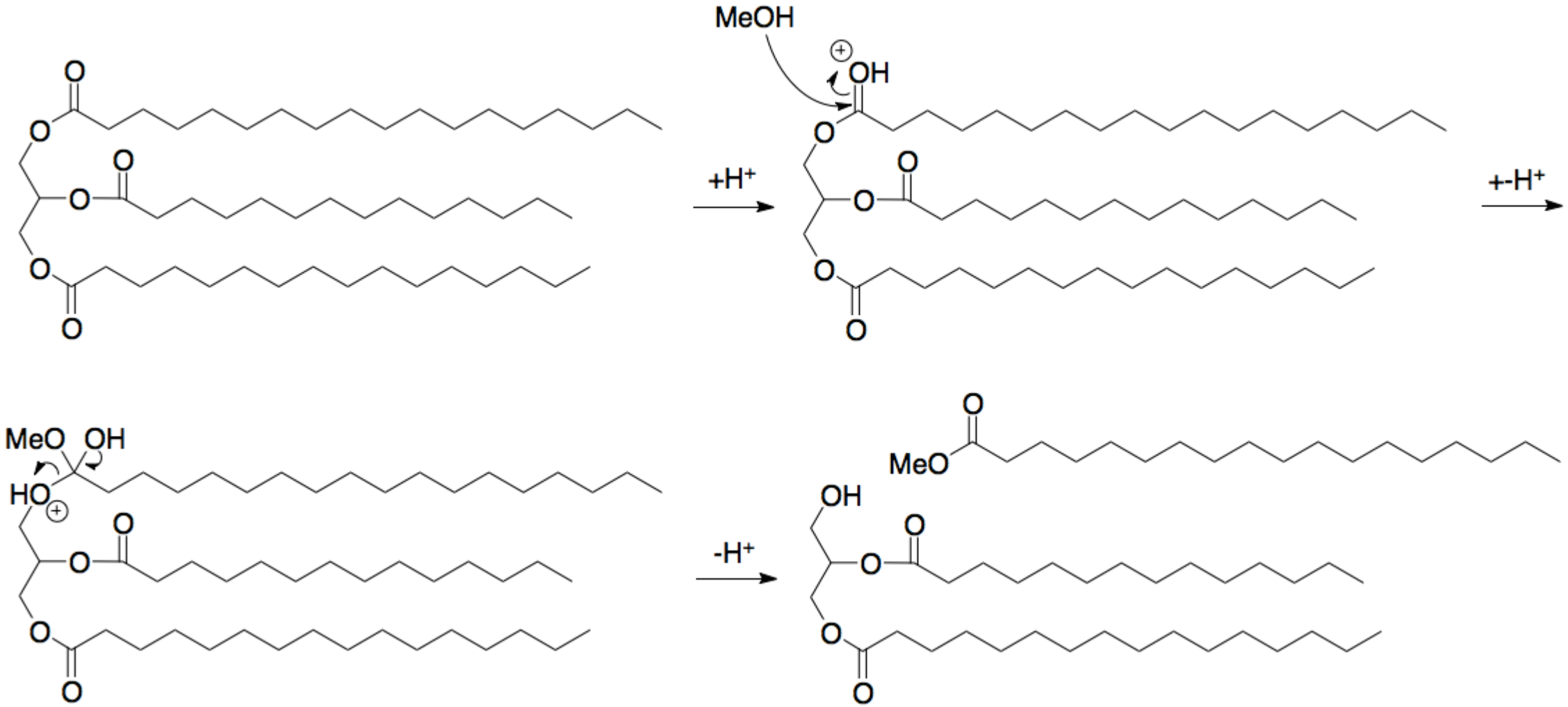


Lipids – Transesterification of triglycerides – Biodiesel

- Diesel engines can be made to run on biodiesel.
- Biodiesel is the result of transesterification of triglycerides.
- Transesterification can be achieved with either acid or base catalysis

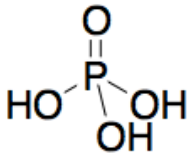


Lipids – Transesterification of triglycerides – Biodiesel

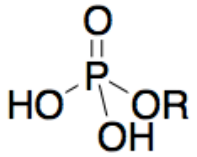


Phospholipids

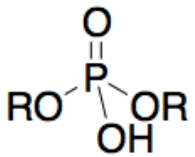
- Phospholipids are ester like derivatives of phosphoric acid.



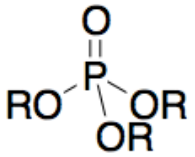
Phosphoric acid



A phosphoric acid monoester

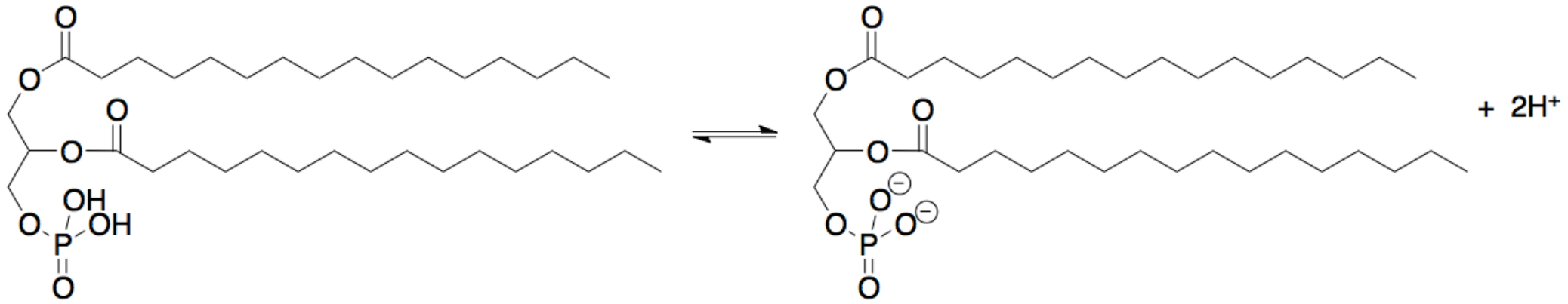


A phosphoric acid diester



A phosphoric acid triester

- Phosphoglycerides are structurally similar to triglycerides. One fatty acid residue on the triglyceride is replaced by a phosphoester group in a phosphoglyceride.

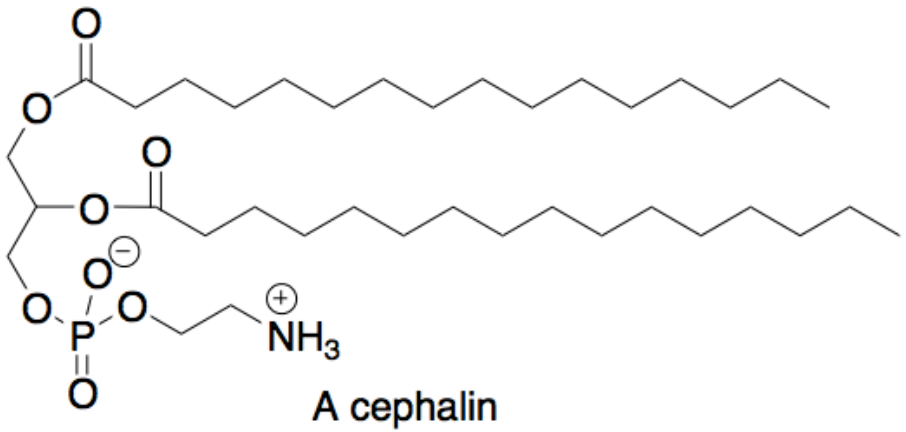


A phosphitidic acid

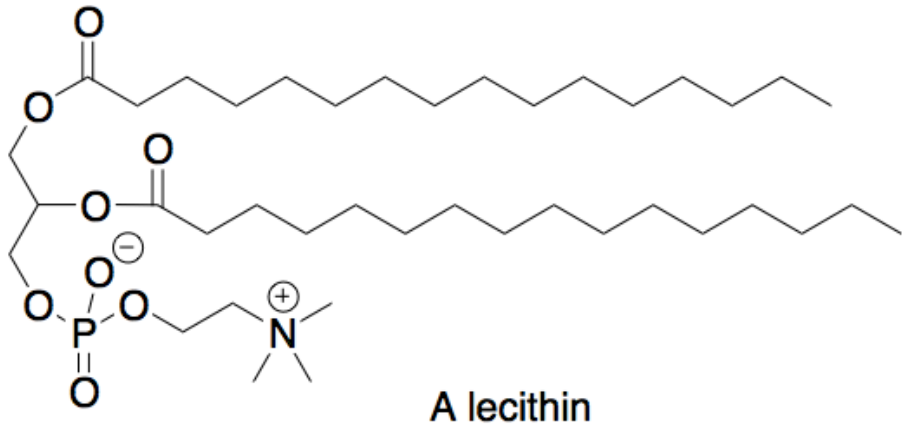
Ioniside form

Phospholipids – cepahlins and lecithins

- Phosphoglycerides containing ethanolamine are called cephalins, and those containing choline are called lecithins.



A cephalin

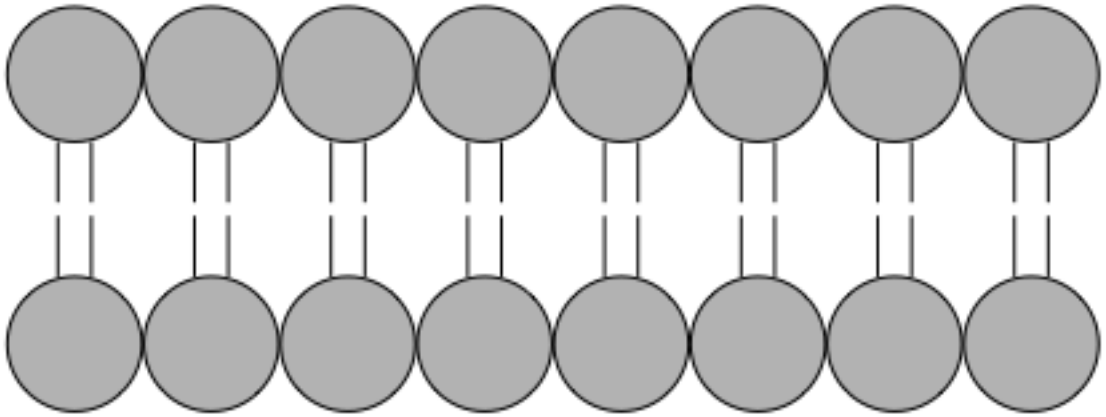


A lecithin

- The C2 of the glycerol unit is chiral and generally has the *R* configuration.

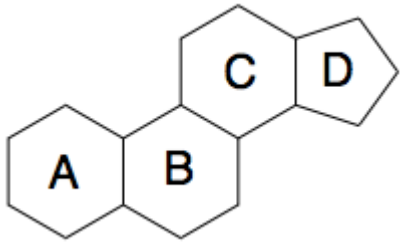
Lipids – Lipid bilayers

- In water, phosphoglycerides self-assemble into a lipid bilayer.
- This arrangement allows the hydrophobic tails to avoid contact with water and interact with each other *via* London forces.
- The phospholipid bilayer is the main constituent of the cell membrane where it functions as a barrier that restricts the flow of water and ions.



Lipids – Steroids

- Most steroids act as chemical messengers, or hormones.
- They are secreted by endocrine glands and move through the bloodstream to their target organs.
- Steroids have a tetracyclic ring system composed of three six-membered rings and one five-membered ring.

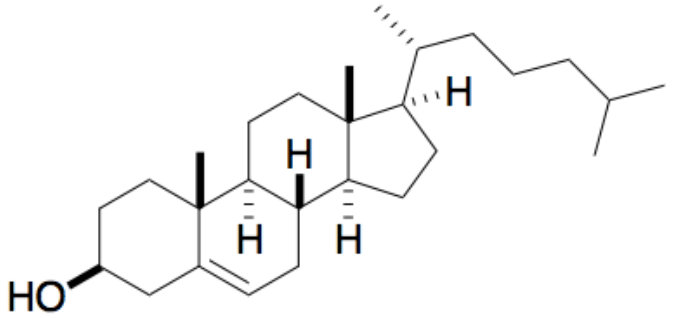


Lipids – Steroids

- In order to understand the configuration and conformation of the tetracyclic ring structure of steroids, we can learn from the decalins.
- When six-membered rings are fused with a *cis* configuration the two rings can both undergo ring flipping.
- *trans*-Decalin does not exhibit ring flipping and as such is a more rigid structure.

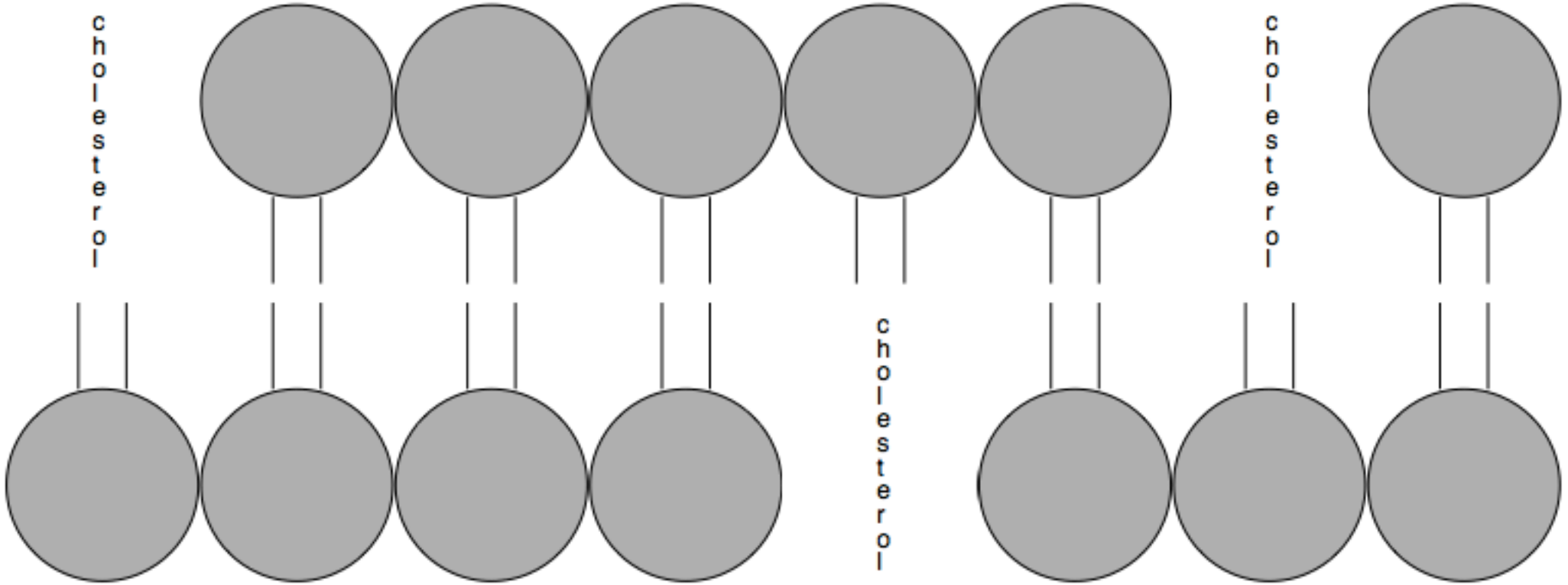
Lipids – Steroids

The ring fusions are all *trans* in the vast majority of steroids leading to a fairly rigid and planar structure.

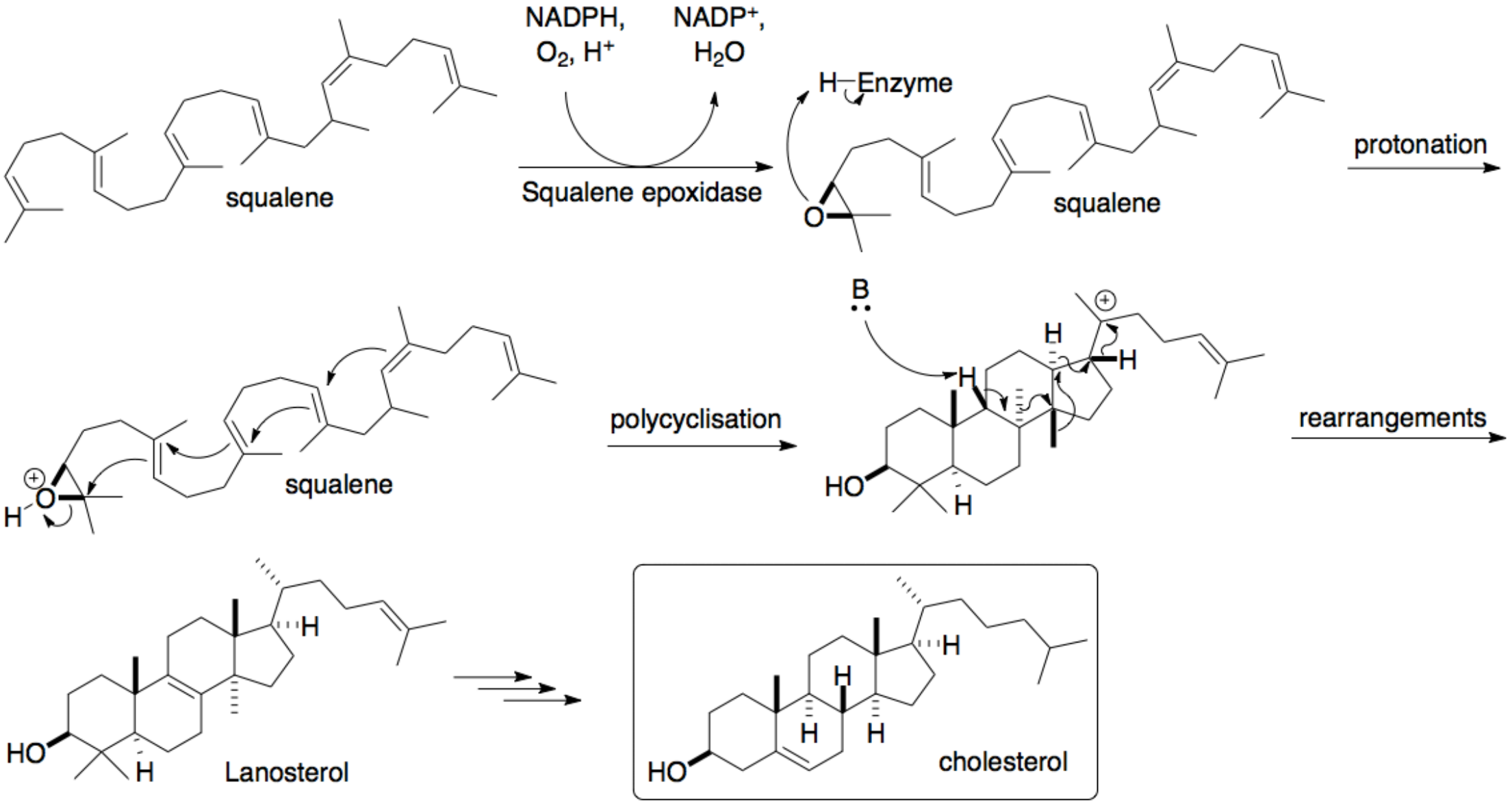


Lipids – Cholesterol as a lipid bilayer stiffening agent

The rigid structure of cholesterol allows it to function as a stiffening agent in lipid bilayers.



Lipids – Biosynthesis of cholesterol

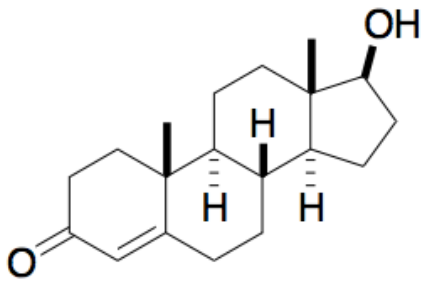


Lipids – Cholesterol and heart disease

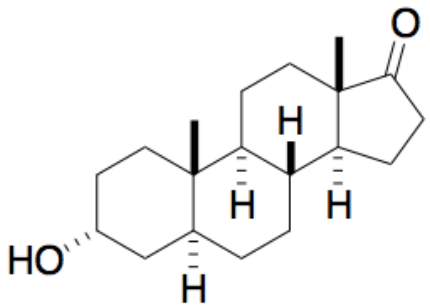
- Cholesterol is a lipid and so by definition is not water soluble.
- The biosynthesis of cholesterol takes place in the liver and it is transported in the bloodstream by large lipoproteins (proteins and lipids bound together by intermolecular interactions).
- LDL (low density lipoproteins) have a low proportion of proteins and transport cholesterol from the liver to the cells. HDL (high density lipoproteins) transport excess cholesterol from the cells to the liver to serve as a precursor in the biosynthesis of other steroids.
- If the LDL levels in the blood are too high, cholesterol will not only be unloaded where needed and deposits of cholesterol accumulate in the artery walls.
- The ratio of LDL to HDL can be used to estimate the risk of stroke or heart attack.

Lipids – Sex hormones

- Human sex hormones are steroids. They regulate tissue growth and reproductive processes.
- The male sex hormones are called **androgens**.
- There are two types of female sex hormones called **estrogens** and **progestins**.

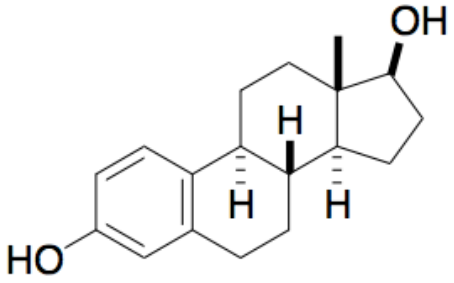


Testosterone

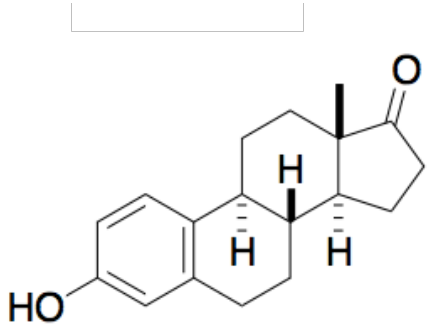


Androsterone

Androgens

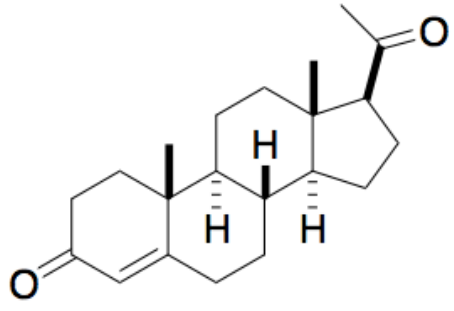


Estradiol



Estrone

Estrogens



Progesterone

A progestin

Lipids – Sex hormones – Androgens

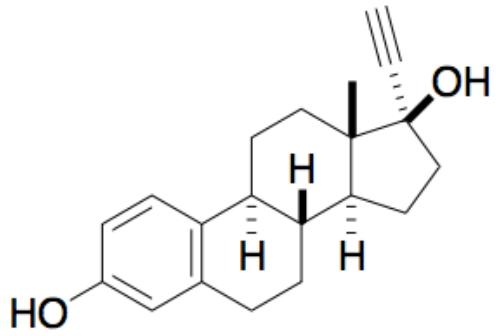
- Testosterone, androsterone, estradiol, estrone and progesterone are all found in both males and females. However, estrogens and progestins are produced in greater concentrations in women, while androgens are produced in greater concentrations in men.
- The androgens control the development of secondary sex characteristics in males.

Lipids – Sex hormones – Estrogens and Progestins

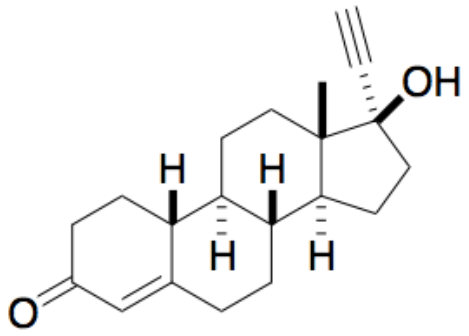
- Estradiol and estrone (estrogens) are synthesised in the ovaries from testosterone.
- They are involved in regulating the menstrual cycle and control the development of secondary sex characteristics in females.
- Progesterone is a progestin that prepares the uterus for nurturing a fertilized egg during pregnancy.

Lipids – Sex hormones – Birth control

- During pregnancy, the body inhibits ovulation by releasing estrogens and progestins. This is mimicked by birth control formulations containing a mixture of a synthetic estrogen and progestin.



Ethynyl estradiol



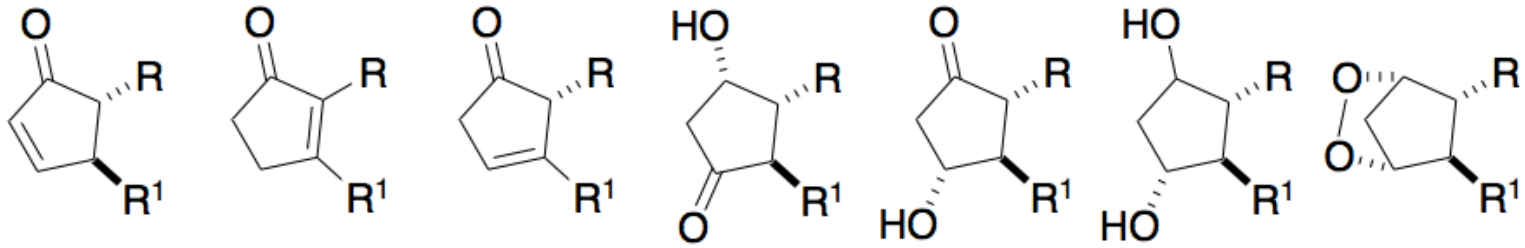
Norethindrone

Lipids – Prostaglandins

- **Prostaglandins** are ‘local hormones’ or **autacoids**.
- They are biochemical regulators and control a number of processes, such as blood pressure, blood clotting, gastric secretions, inflammation, kidney function, and reproductive systems.

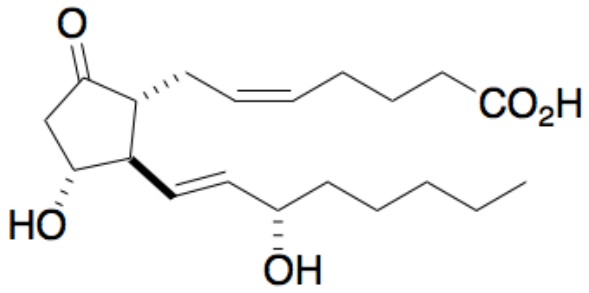
Lipids – Prostaglandins

- Prostaglandins contain 20 carbon atoms and are made up of a five-membered ring with two side chains.
- Prostaglandins with different substitution patterns are found in nature:



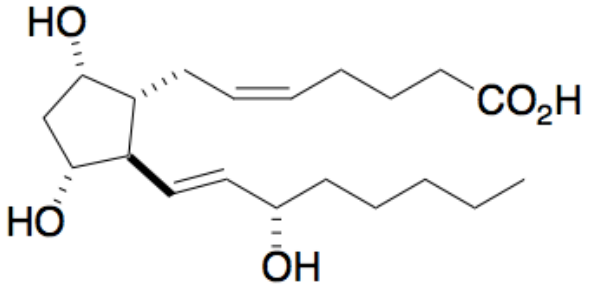
Lipids – Prostaglandins

- The number of carbon-carbon π bonds in the side chains is indicated with a subscript.

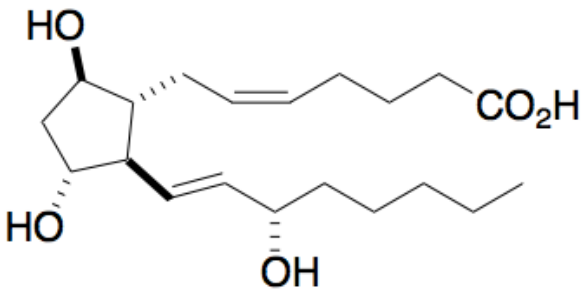


PGE_2

- For the PGF substitution pattern, the configuration of the OH groups is indicated. A *cis* diol is designated α and a *trans* diol is designated β.



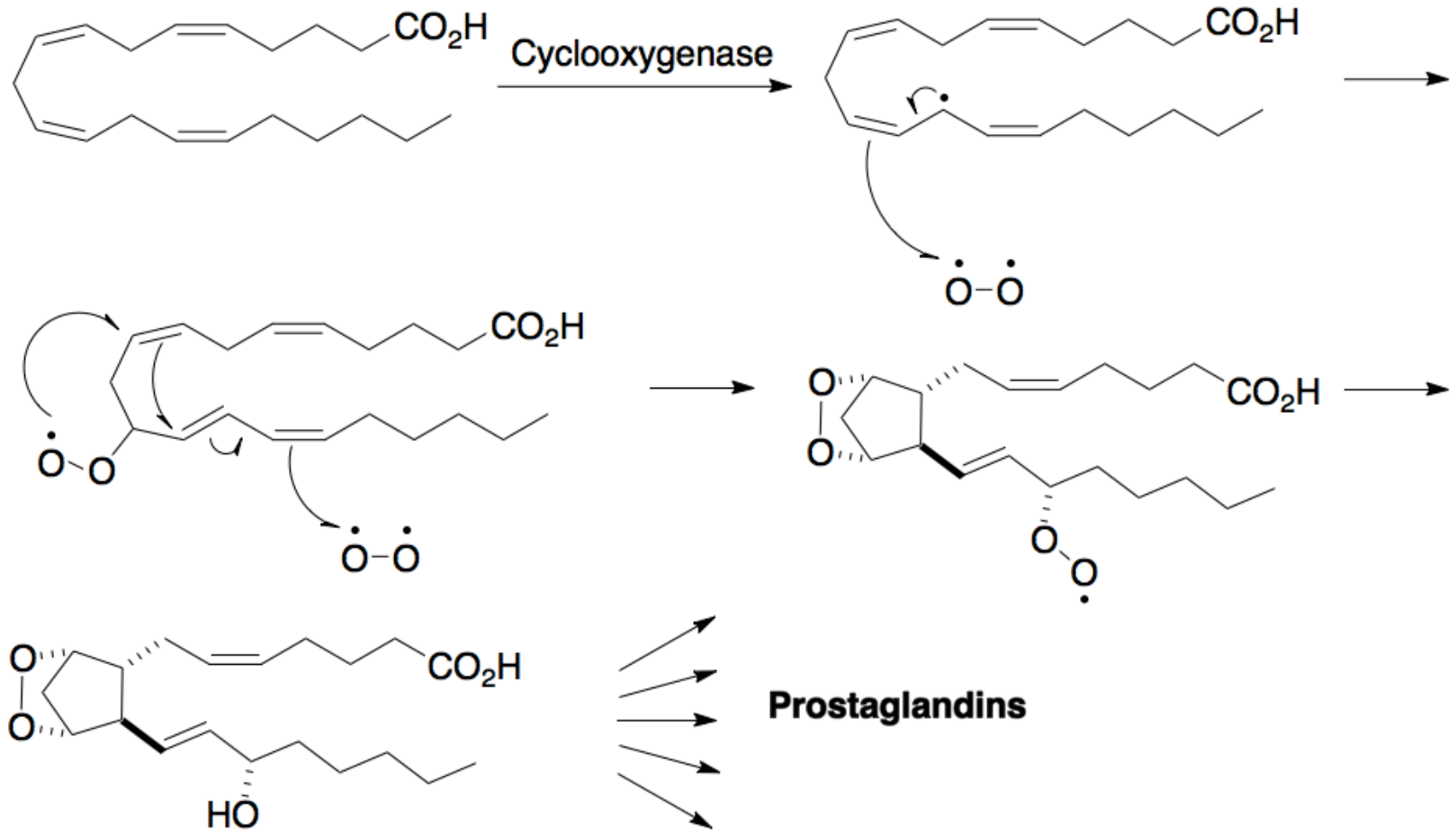
$PGF_{2\alpha}$



$PGF_{2\beta}$

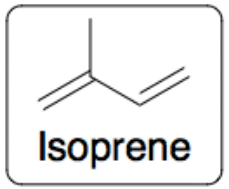
Lipids – Prostaglandins

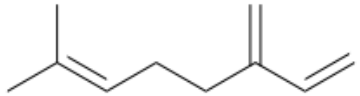
Prostaglandins are biosynthesized from arachidonic acid catalysed by enzymes called cyclooxygenases.



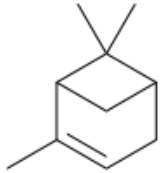
Lipids – Terpenes

- Terpenes were originally named after turpentine.
- Terpenes can be thought of as being made up of five carbon isoprene units.

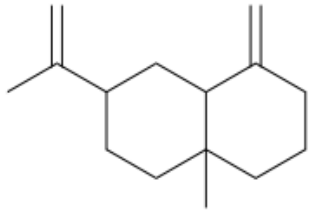




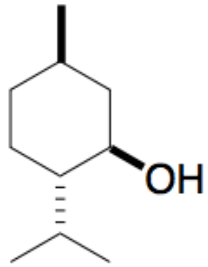
Myrcene
(isolated from bay and myrcia plants)



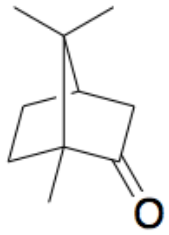
α-Pinene
(isolated from pine trees)



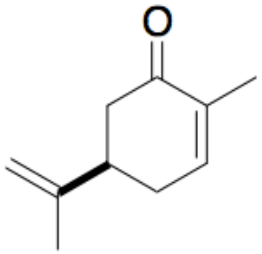
β-Selenine
(isolated from celery)



Menthol
(isolated from peppermint oil)



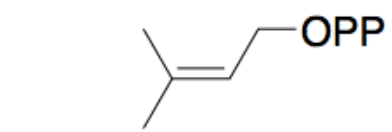
Camphor
(isolated from evergreen trees)



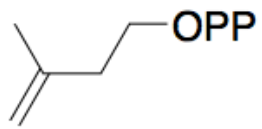
R-Carvone
(isolated from spearmint)

Lipids – Terpenes

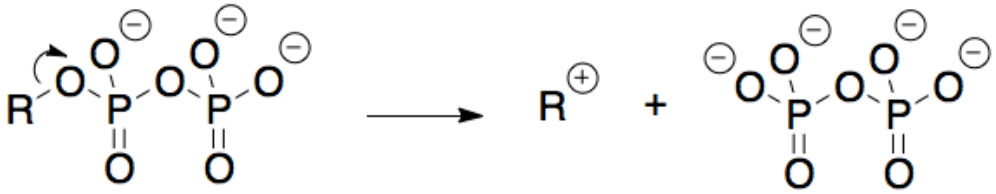
- The isoprene rule considers terpenes as being made up isoprene units. However, the actual building blocks are dimethylally pyrophosphate and isopentenyl pyrophosphate.
- OPP, pyrophosphate, is a good leaving group (weak base).



Dimethylally pyrophosphate



Isopentyl pyrophosphate

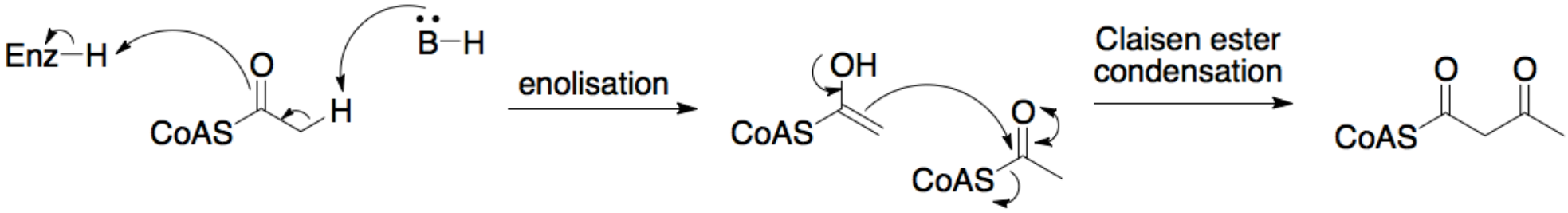


R-OPP

Weak base
(good leaving group)

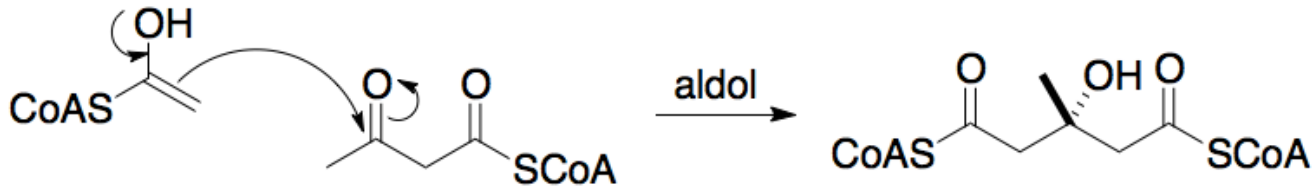
Lipids – Terpenes

- The building blocks for terpenes, dimethylally pyrophosphate and isopentenyl pyrophosphate are both made from the same key intermediate, mevalonic acid which is prepared from 3 molecules of acteyl CoA.
- The first step in the biosynthesis of mevalonic acid is a Claisen ester condensation between two molecules of acteyl CoA.

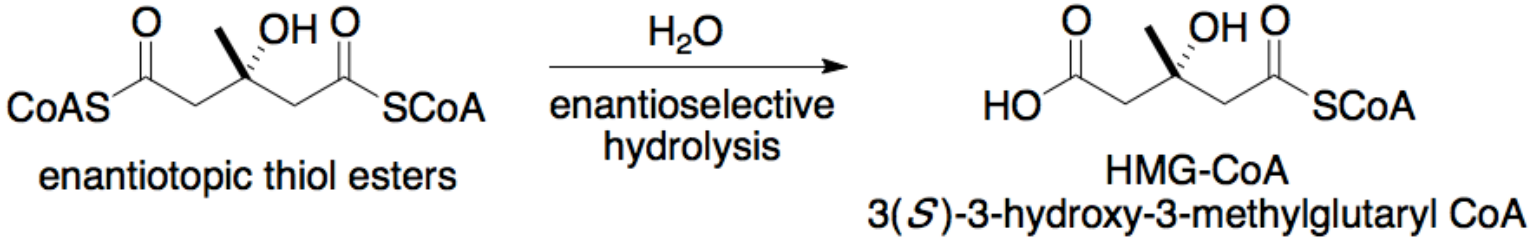


Lipids – Terpenes

- An aldol reaction between the acetoacetyl CoA (formed in the previous step) and the third molecule of acetyl CoA yields an achiral dithiol ester.

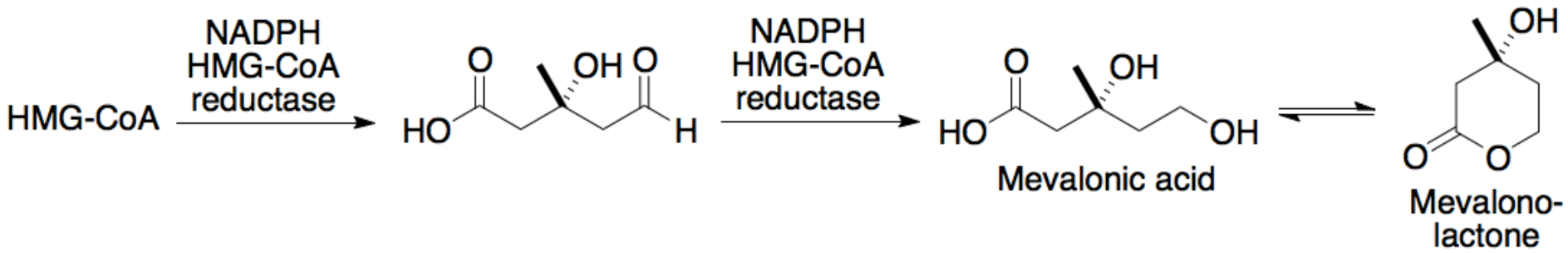


- The dithiol ester has two enantiotopic thiol esters. Because it is still bound to the enzyme, hydrolysis of one of these thiol esters yields a single enantiomer of the half-acid/half-thiol ester, HMG-CoA.



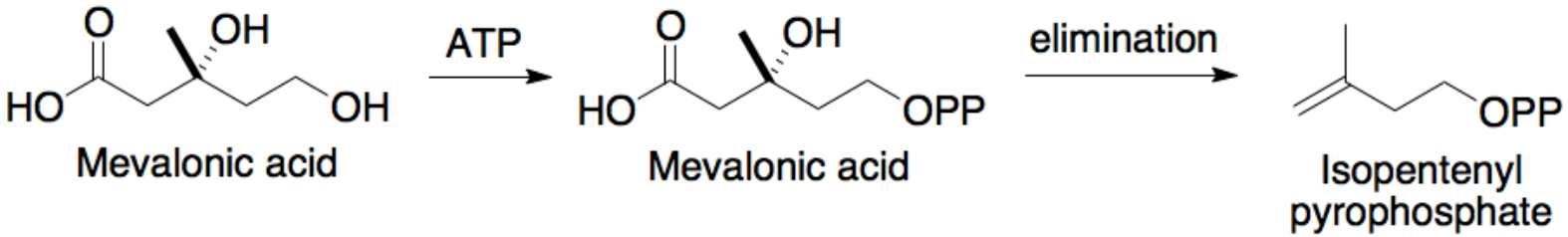
Lipids – Terpenes

Reduction of of the thiol ester in HMG CoA to a hydroxyl group by NADPH yields mevalonic acid.

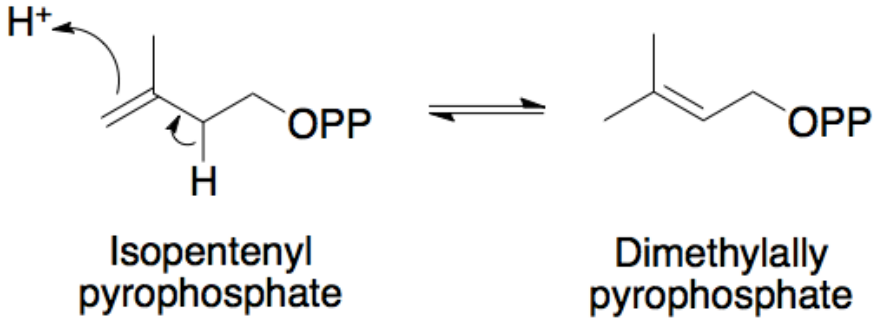


Lipids – Terpenes

- Mevalonic acid is a C6 compound and so must lose a carbon atom. The primary alcohol is pyrophosphorylated with ATP and then CO_2 and water are lost in a concerted elimination yielding the isopentyl pyrophosphate building block.

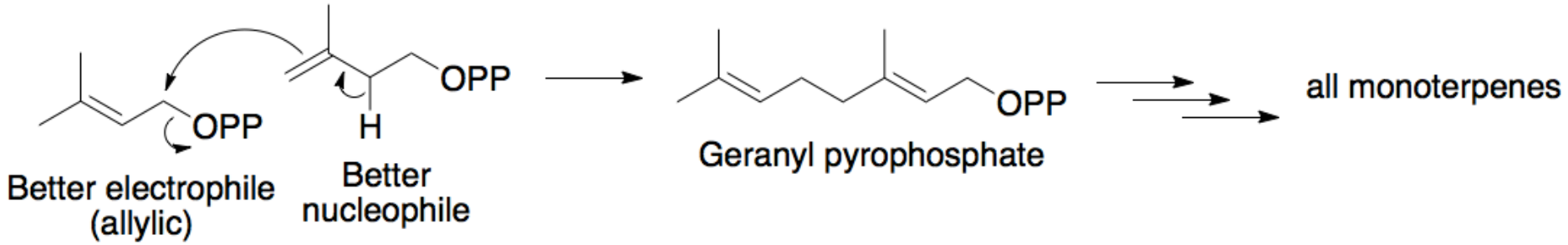


- Isopentenyl pyrophosphate is in equilibrium with dimethylallyl pyrophosphate by a simple allylic proton transfer.



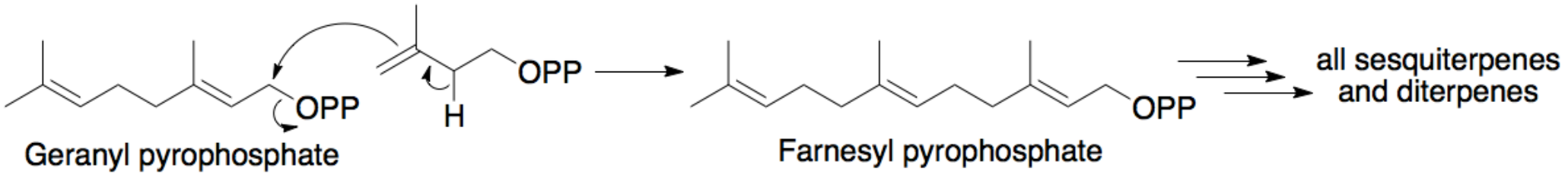
Lipids – Terpenes

The reaction between dimethylallyl pyrophosphate and isopentenyl pyrophosphate gives a monoterpene called geranyl pyrophosphate, the starting material for all other monoterpenes.



Lipids – Terpenes

The same mechanistic steps are repeated (addition of another isoprene unit to geranyl pyrophosphate) in the biosynthesis of the sesquiterpene farnesyl pyrophosphate.



Lipids – Terpenes

Squalene is biosynthesised from the coupling of two molecules of farnesyl pyrophosphate. We have already seen that squalene is the biological precursor for all steroids.