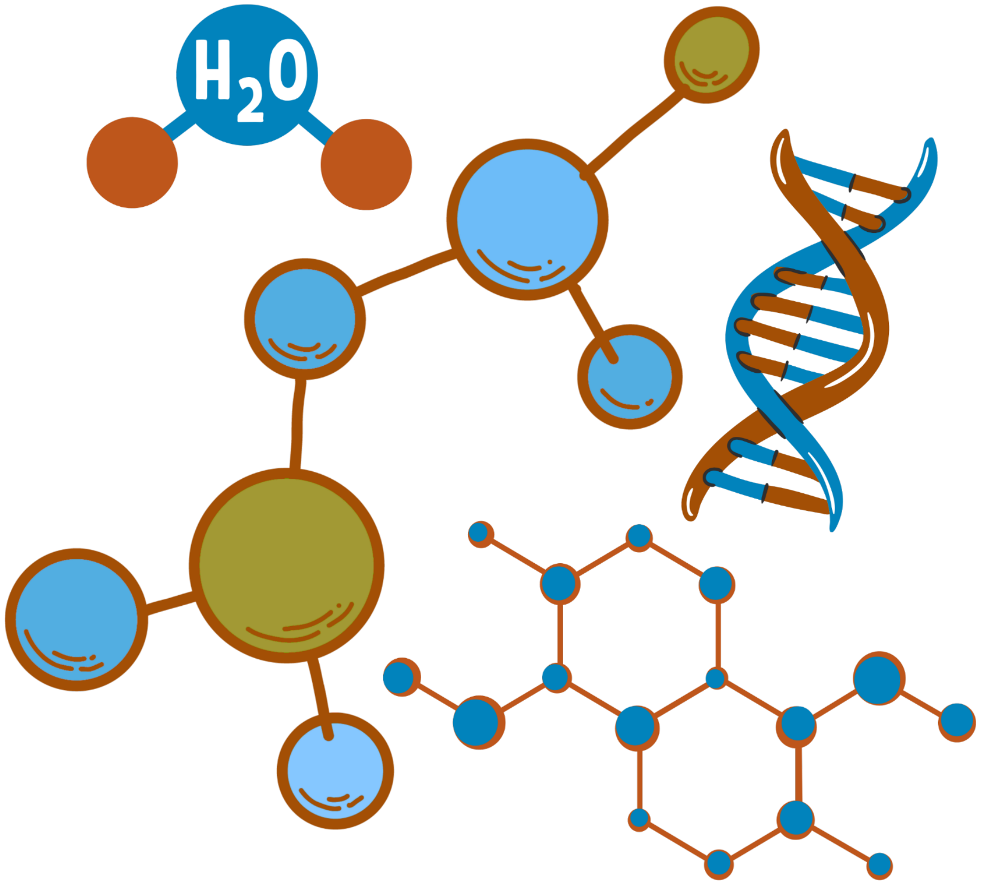
جامعة الملك سعود
King Saud University

COLLEGE OF SCIENCES
CHEMISTRY DEPARTMENT

Introduction to Organic Chemistry
CHEM 109



Chapter 6: Biomolecules and Their Biological Relevance

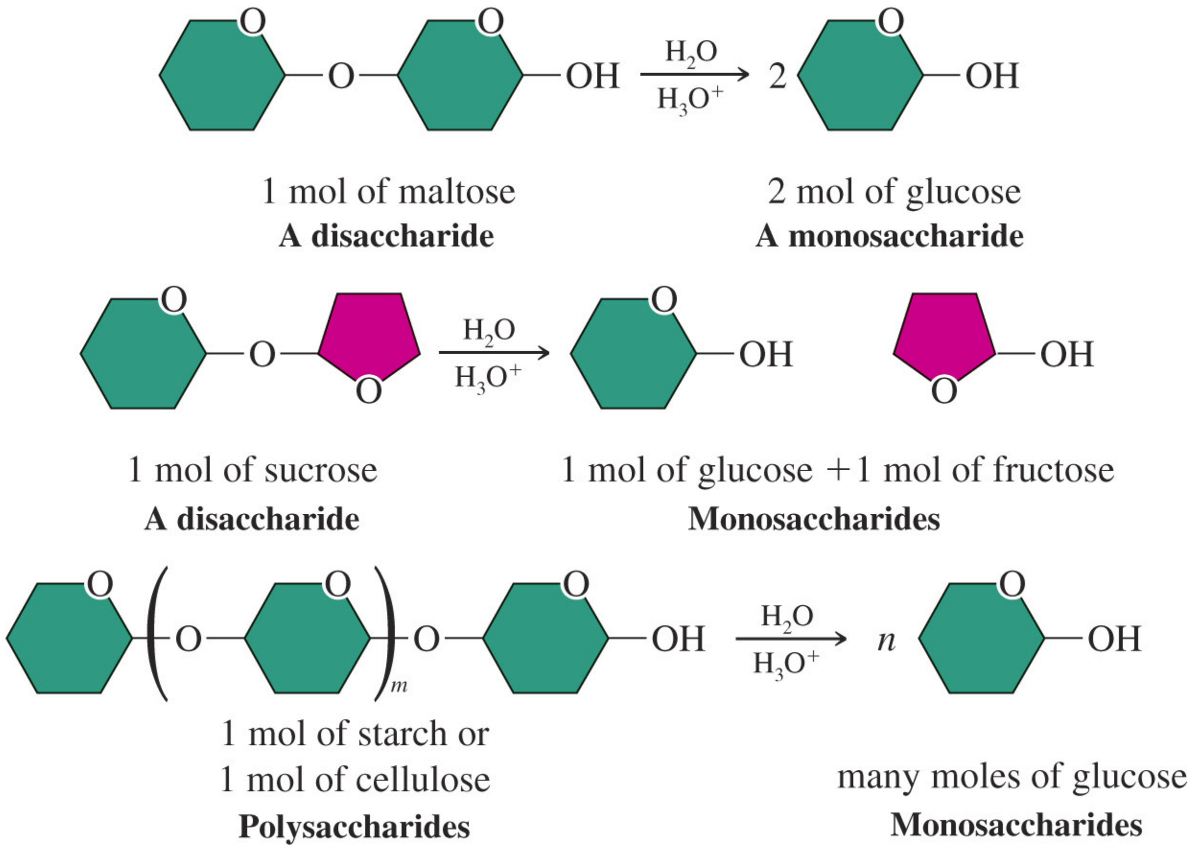
Carbohydrates

6.1. Introduction

- **Carbohydrates** are polyhydroxy aldehydes or ketones, or compounds that can be hydrolyzed to produce such molecules.
- Carbohydrates typically follow the general chemical formula $Cx(H_2O)y$.

6.1.1. Classification of Carbohydrates

- **Monosaccharides** are the simplest carbohydrates and cannot be further hydrolyzed into smaller carbohydrate units.
- **Disaccharides** can be broken down by hydrolysis into two monosaccharide units.
- **Oligosaccharides** produce between two and ten monosaccharides upon hydrolysis.
- **Polysaccharides** yield more than ten monosaccharides when hydrolyzed.



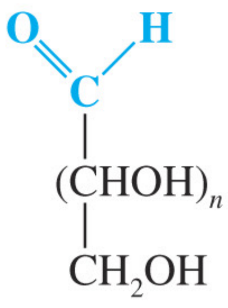
6.2. Monosaccharides

6.2.1. Classification of Monosaccharides

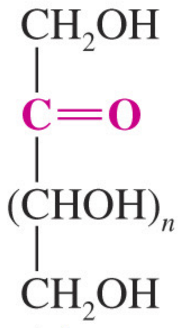
- Monosaccharides are classified based on two main features:
  1. The number of carbon atoms they contain.
  2. The type of carbonyl group present — either an aldehyde or a ketone group.

6.2.2. D and L Designations of Monosaccharides

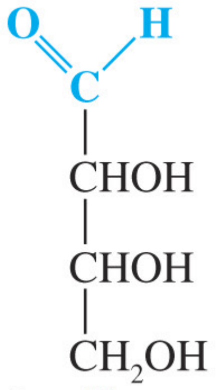
- **Chirality** is the property of a molecule that makes it non-superimposable on its mirror image, usually due to the presence of a chiral (asymmetric) carbon atom bonded to four different groups.
- The simplest carbohydrates are glyceraldehyde, which is chiral and assigned as L or D-configuration, and dihydroxyacetone, which is achiral.



An aldose

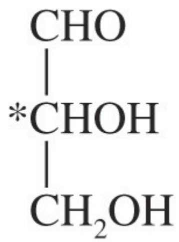


A ketose

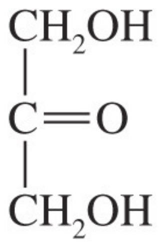


An aldotetrose
C_4

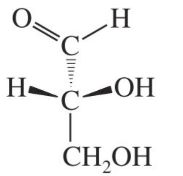
A ketopentose
C_5



Glyceraldehyde
(an aldotriose)

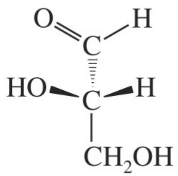


Dihydroxyacetone
(a ketotriose)



(+)-Glyceraldehyde

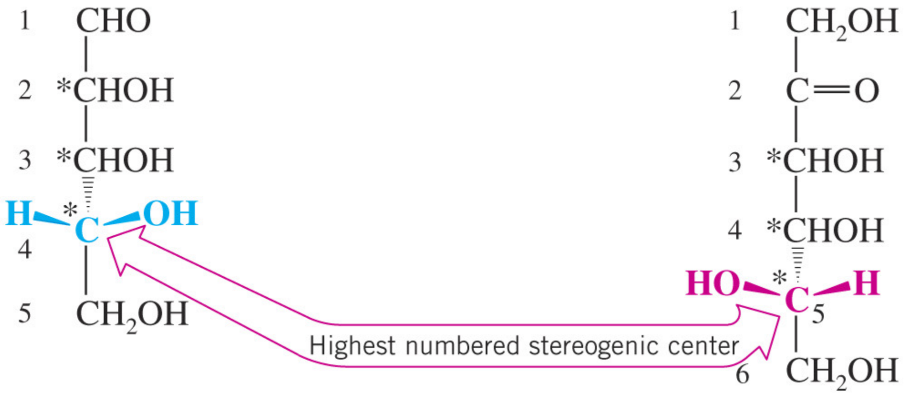
and



(−)-Glyceraldehyde

6.2. Monosaccharides

- A monosaccharide is classified as a D-sugar if its highest-numbered chiral center has the same configuration as D-(+)-glyceraldehyde.
- A monosaccharide is identified as an L-sugar if its highest-numbered chiral center shares the same configuration as L-(–)-glyceraldehyde.

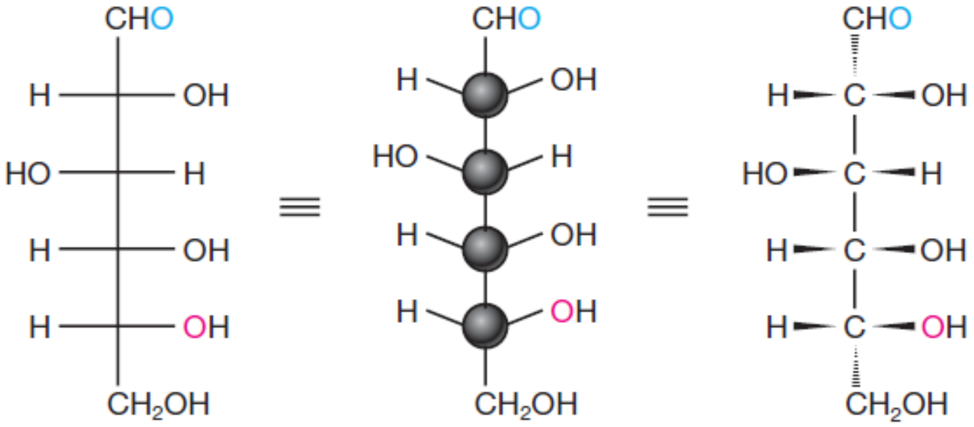


A D-aldopentose **An L-ketohexose**

6.2.3. Structural Formulas for Monosaccharides

Fischer projection

- Represents sugar stereochemistry in the open-chain form.
- Horizontal lines → bonds projecting out of the plane (toward the viewer).
- Vertical lines → bonds behind the plane (away from the viewer).
- The –CHO group is placed at the top, and the –CH$_2$OH group at the bottom.

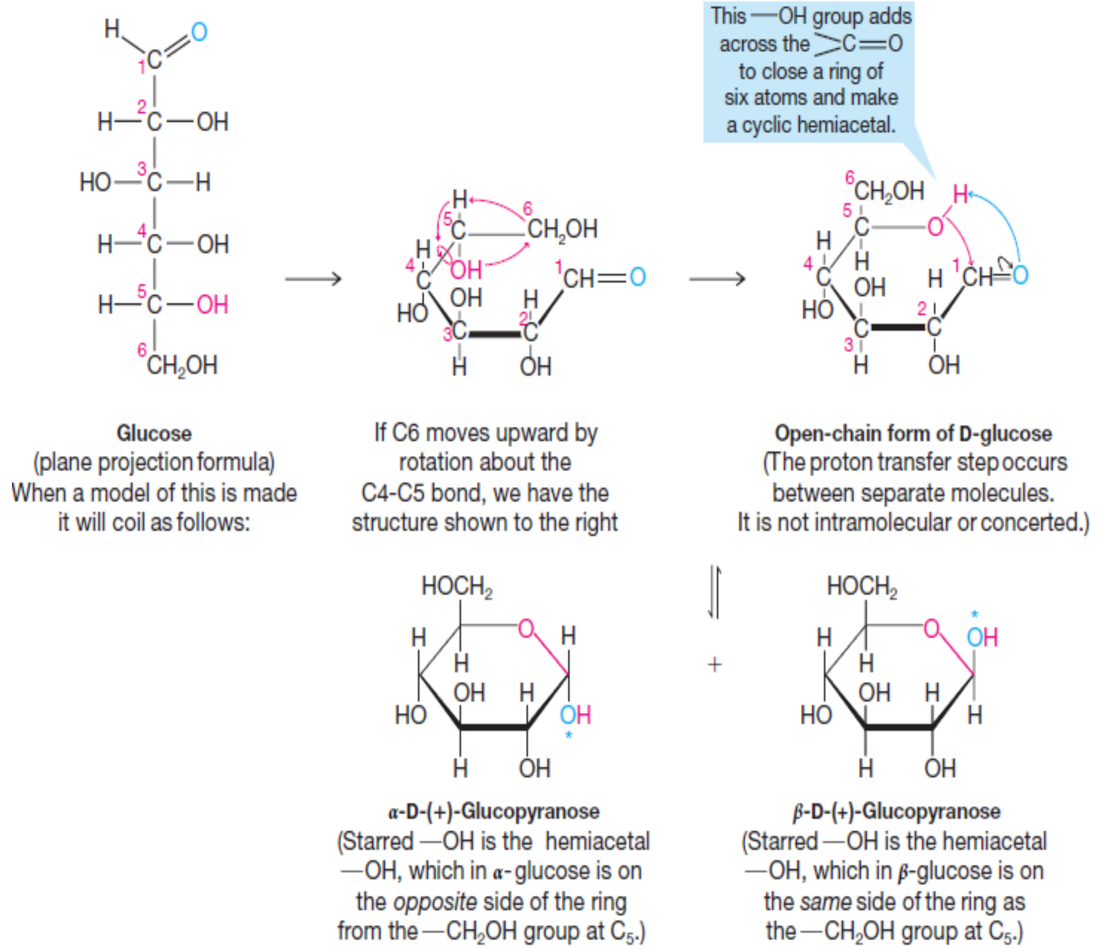


Fischer projection formula **Circle-and-line formula** **Wedge–line–dashed wedge formula**

6.2.3. Structural Formulas for Monosaccharides

Haworth formulas

- Represent sugar stereochemistry in the cyclic form, which exists in equilibrium with the open-chain form.
- The cyclic (hemiacetal) forms are produced by an intramolecular reaction between the –OH group at C5 and the aldehyde group at C1.
- Cyclization creates a new chiral center at C1, explaining the existence of two cyclic forms.
- These two forms are diastereomers, differing only in the configuration at C1 (the anomeric carbon).
    - α-anomer → the hemiacetal –OH is *trans* to the –CH_2OH group.
    - β-anomer → the hemiacetal –OH is *cis* to the –CH_2OH group.



This —OH group adds across the >C=O to close a ring of six atoms and make a cyclic hemiacetal.

Glucose
(plane projection formula)
When a model of this is made it will coil as follows:

If C6 moves upward by rotation about the C4-C5 bond, we have the structure shown to the right

Open-chain form of D-glucose
(The proton transfer step occurs between separate molecules. It is not intramolecular or concerted.)

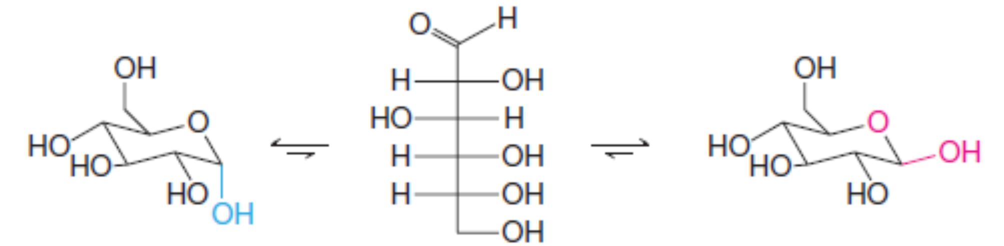
α-D-(+)-Glucopyranose
(Starred —OH is the hemiacetal —OH, which in α-glucose is on the *opposite* side of the ring from the —CH_2OH group at C_5.)

β-D-(+)-Glucopyranose
(Starred —OH is the hemiacetal —OH, which in β-glucose is on the *same* side of the ring as the —CH_2OH group at C_5.)

6.2.4. Mutarotation

Mutarotation: a change in specific rotation toward an equilibrium value.

- It results from the equilibrium established between the open-chain form of D-(+)-glucose and its cyclic α- and β-hemiacetal forms.
- In solutions: The concentration of open-chain is negligible, α anomer 36%, and β anomer 64%, reflecting a greater stability for β-d-(+)-glucopyranose due to the equatorial OH group on C1.



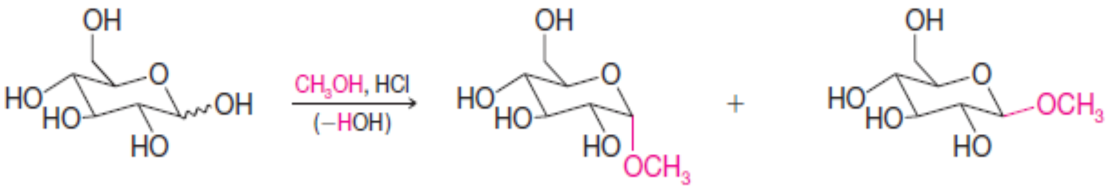
α-D-(+)-Glucopyranose (mp 146 °C; $[\alpha]_D^{25} = +112$)

Open-chain form of D-(+)-glucose

β-D-(+)-Glucopyranose (mp 150 °C; $[\alpha]_D^{25} = +18.7$)

6.2.4.1. Glycoside Formation

- Carbohydrate acetals are generally called glycosides.
- Reaction of D-(+)-glucose with methanol in presence of anhydrous (gaseous) hydrogen chloride yields anomeric methyl acetal.
- An acetal of glucose is called a glucoside, acetals of mannose are mannosides, acetals of fructose are fructosides, and so on.)



D-(+)-Glucose

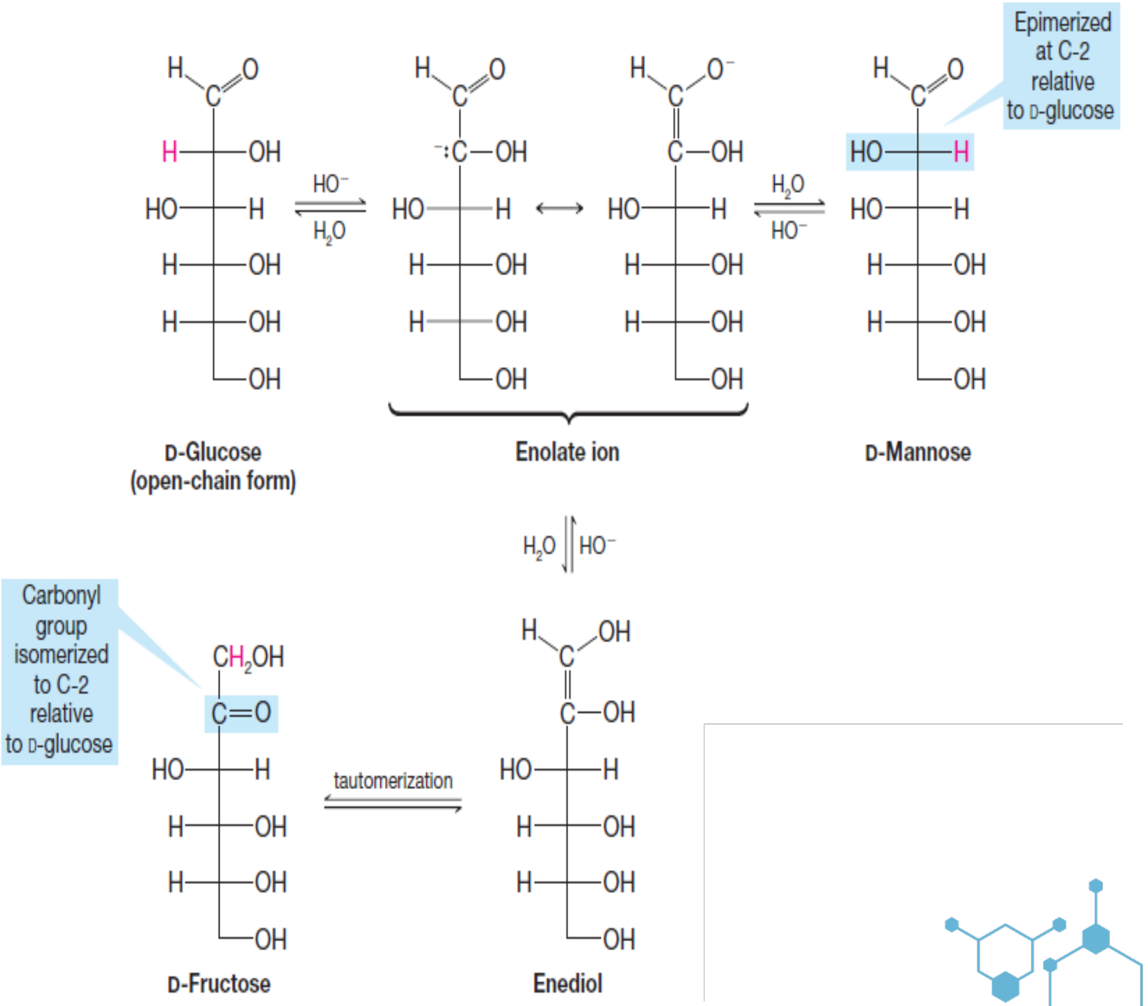
Methyl α-D-glucopyranoside (mp 165 °C; $[\alpha]_D^{25} = +158$)

Methyl β-D-glucopyranoside (mp 107 °C; $[\alpha]_D^{25} = -33$)

6.2.4.2. Other Reactions of Monosaccharides

a. Enolization, Tautomerization, and Isomerization

- Dissolving monosaccharides in aqueous base causes them to undergo a series of enolizations and keto–enol tautomerizations that lead to isomerizations.
- For example, if a solution of D-glucose containing calcium hydroxide is allowed to stand for several days, a number of products can be isolated, including D-fructose and D-mannose.
- Converting a monosaccharide to a methyl glycoside protects the anomeric carbon by preventing isomerization and preserving stereochemistry under basic conditions.



6.2.4.2. Other Reactions of Monosaccharides

b. Use of Protecting Groups in Carbohydrate Synthesis

- Protecting groups temporarily block reactive sites in a molecule to allow selective reactions elsewhere.
- After desired transformations, they are removed.
- Carbohydrates have multiple reactive sites, so protecting groups are often necessary during reactions.
- Forming a glycoside (acetal) protects the anomeric carbon from unwanted reactions.
- Typical protecting groups for alcohols in carbohydrates include ethers, esters, and acetals.

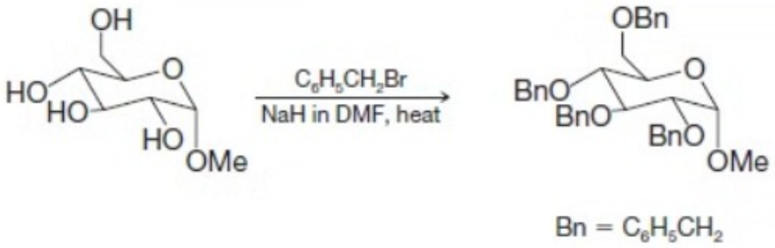
6.2.4.2. Other Reactions of Monosaccharides

C. Formation of Ether

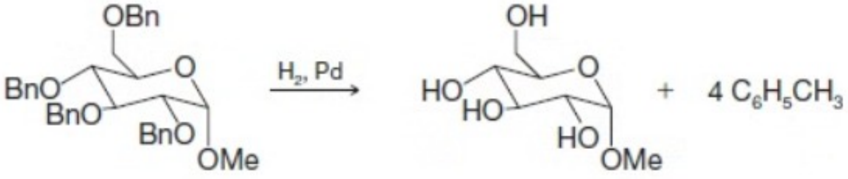
- Hydroxyl groups of sugars can be converted to ethers using a base and an alkyl halide (Williamson ether synthesis).
- **Benzyl ethers** are commonly used to protect hydroxyl groups in sugars in presence of sodium hydride as a base and DMF or DMSO as a solvent.
- The benzyl groups can later be easily removed by hydrogenolysis using a palladium catalyst.

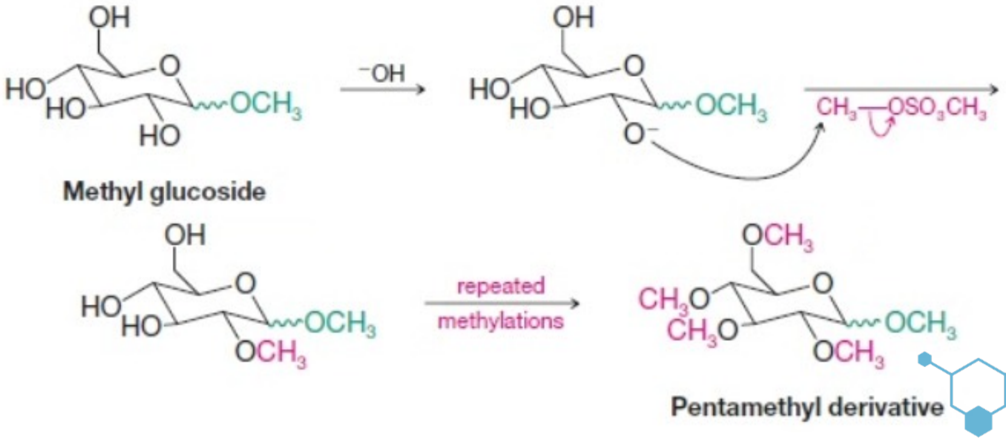
- **The pentamethyl derivative** is synthesized by treating methyl glucoside with excess dimethyl sulfate in aqueous sodium hydroxide.
- In aqueous NaOH the hydroxyl groups are all eventually converted to alkoxide ions, and each of these, in turn, reacts with dimethyl sulfate to yield a methyl ether.

Benzyl Ether Formation



Benzyl Ether Cleavage

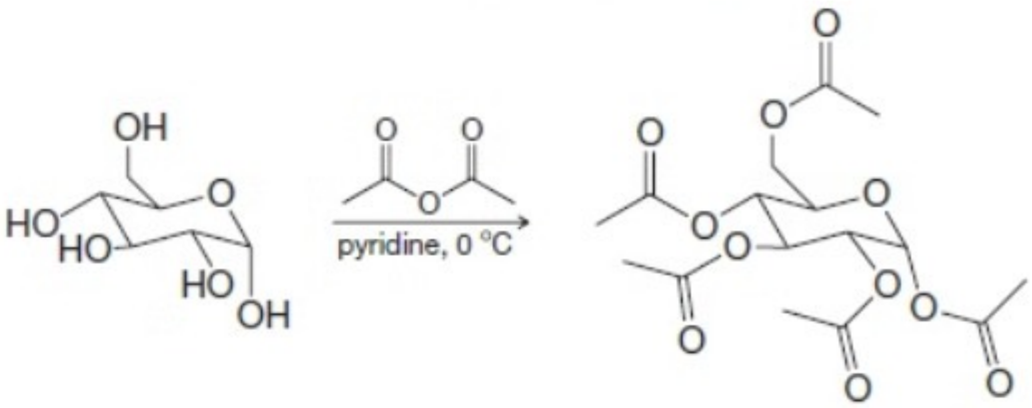




6.2.4.2. Other Reactions of Monosaccharides

d. Conversion to Esters

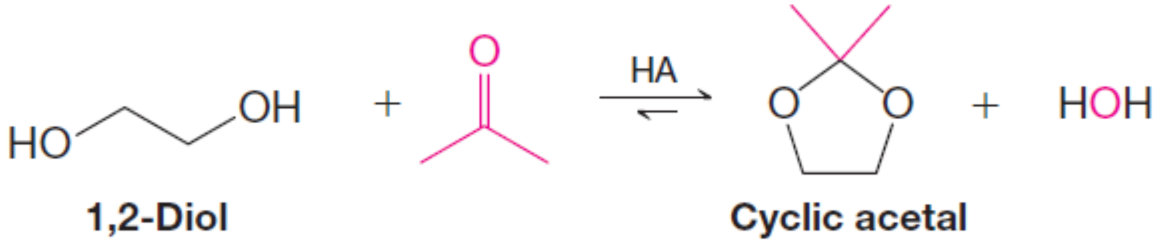
- Treating a monosaccharide with excess acetic anhydride and a weak base (such as pyridine or sodium acetate) converts all of the hydroxyl groups, including the anomeric hydroxyl, to ester groups.
- At low temperature (≈0 °C), the reaction is stereospecific: the α-anomer forms the α-acetate, and the β-anomer forms the β-acetate.
- Acetate esters serve as common protecting groups for carbohydrate hydroxyls.

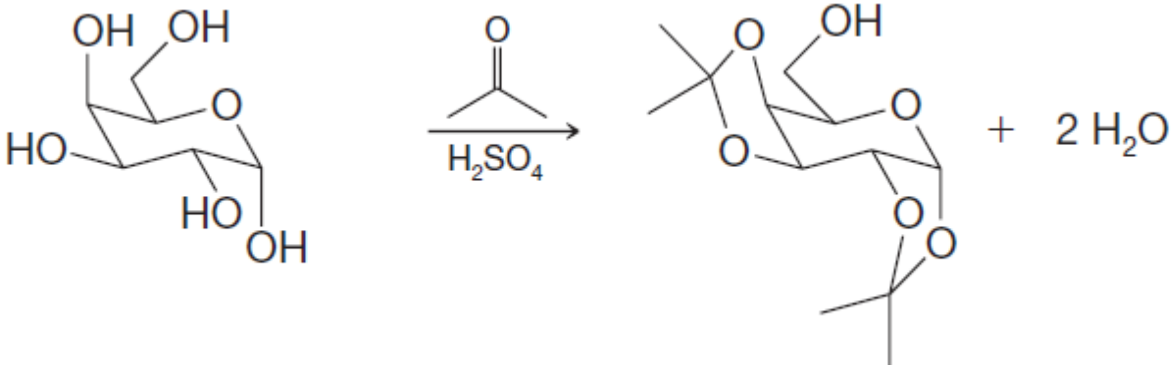


6.2.4.2. Other Reactions of Monosaccharides

e. Conversion to Cyclic Acetals

- As was mentioned before alcohol react with aldehyde or ketone to form acetal.
- Formation of the cyclic acetals in a monosaccharide occurs only when the vicinal hydroxyl groups (1,2-diol) on a ring are cis to each other.
- Cyclic acetals are commonly used to protect vicinal cis hydroxyl groups of a sugar while reactions are carried out on other parts of the molecule.
- When acetals such as these are formed from acetone, they are called acetonides.

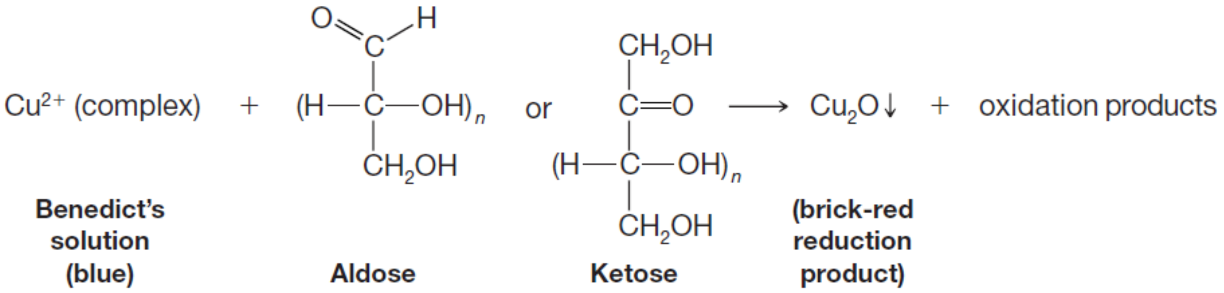


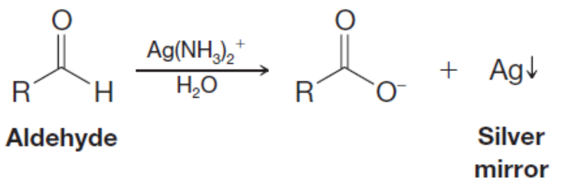


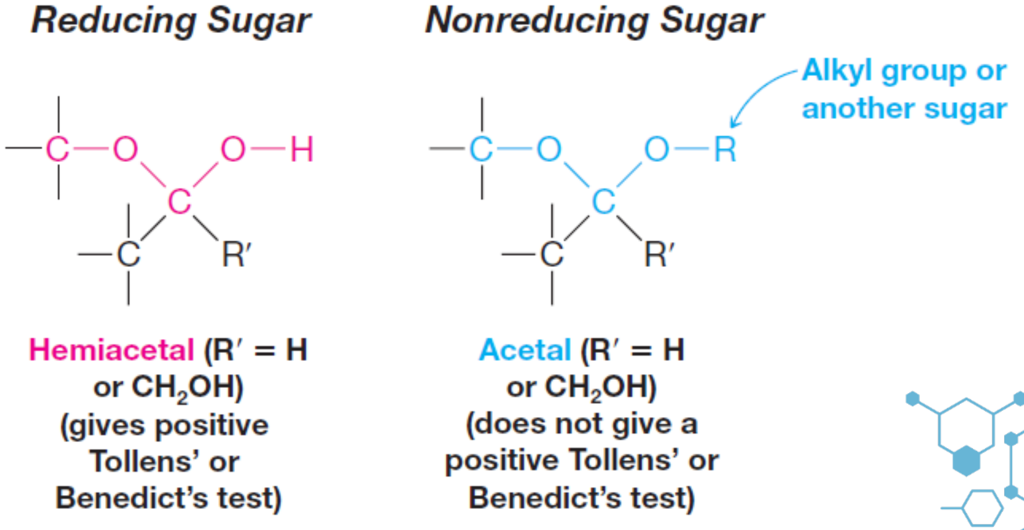
6.2.5. Oxidation Reactions of Monosaccharides

a. Benedict's or Tollens' Reagents: Reducing Sugars

- **Benedict's reagent:** an alkaline solution containing a cupric citrate complex ion.
- **Tollens' solution:** $[\overset{+}{Ag}(NH_3)_2\overset{-}{O}H]$
- They both oxidize and give positive tests with aldoses and ketoses even those that exist primarily as cyclic hemiacetals.
- In aqueous solution, the hemiacetal form of sugars is in equilibrium with small amounts of open-chain aldehyde or *α*-hydroxy ketone forms, which undergo oxidation until the reactants are consumed.
- Carbohydrates that contain only acetal groups do not give positive tests with Benedict's or Tollens' solutions, and they are called *nonreducing sugars*.
- Acetals do not exist in equilibrium with aldehydes or *α*-hydroxy ketones in the basic aqueous media of the test reagents.



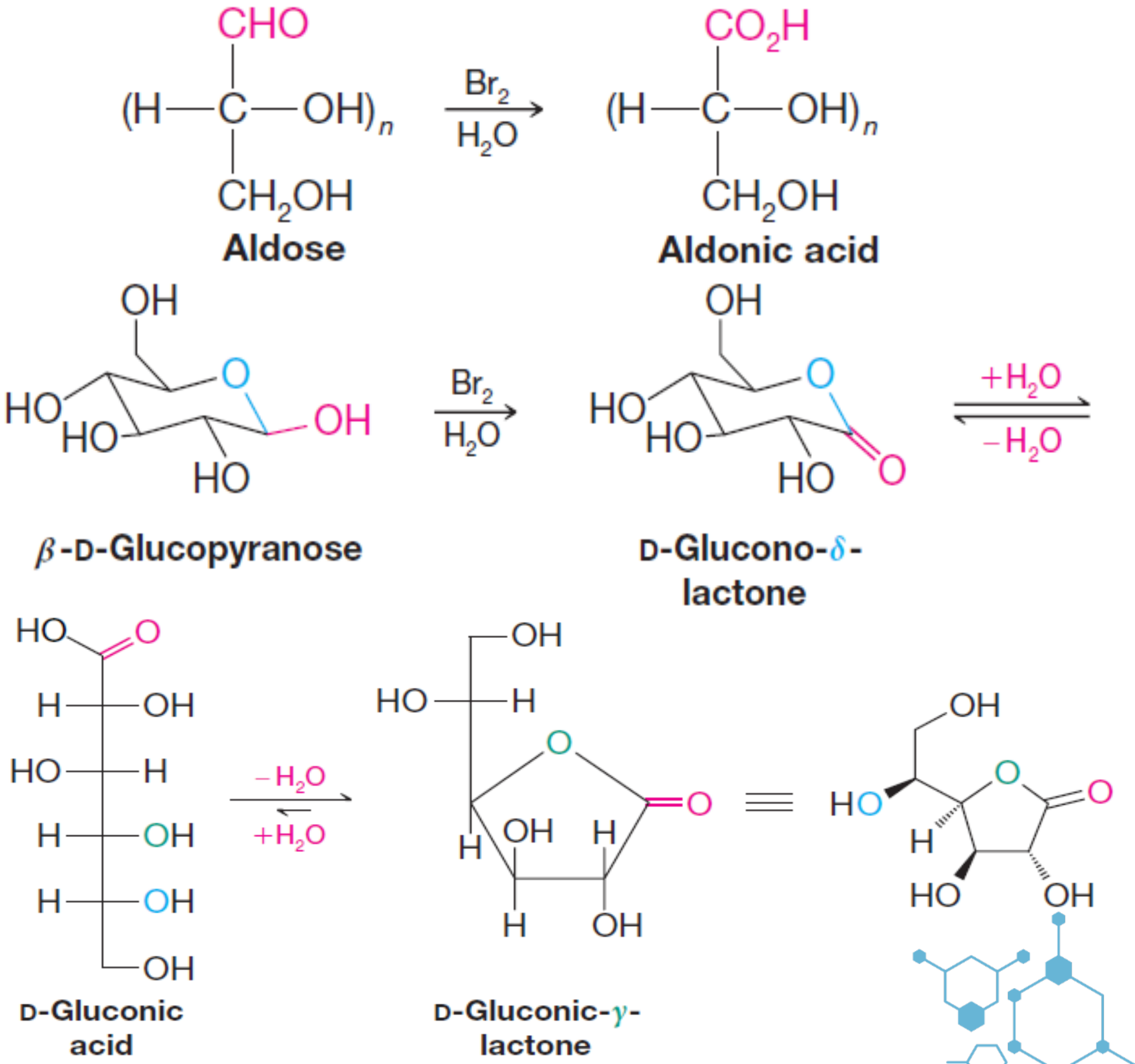




6.2.5. Oxidation Reactions of Monosaccharides

b. Bromine Water: The Synthesis of Aldonic Acids

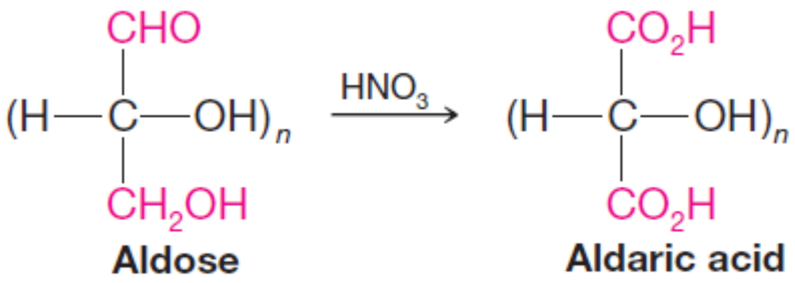
- Bromine water is a general reagent that selectively oxidizes the −CHO group to a $-CO_2H$ group, thus converting an aldose to an **aldonic acid**.
- Experiments show bromine water selectively oxidizes the β-anomer of aldopyranoses to form δ-aldonolactone, which hydrolyzes to aldonic acid and may cyclize again to yield γ-aldonolactone.

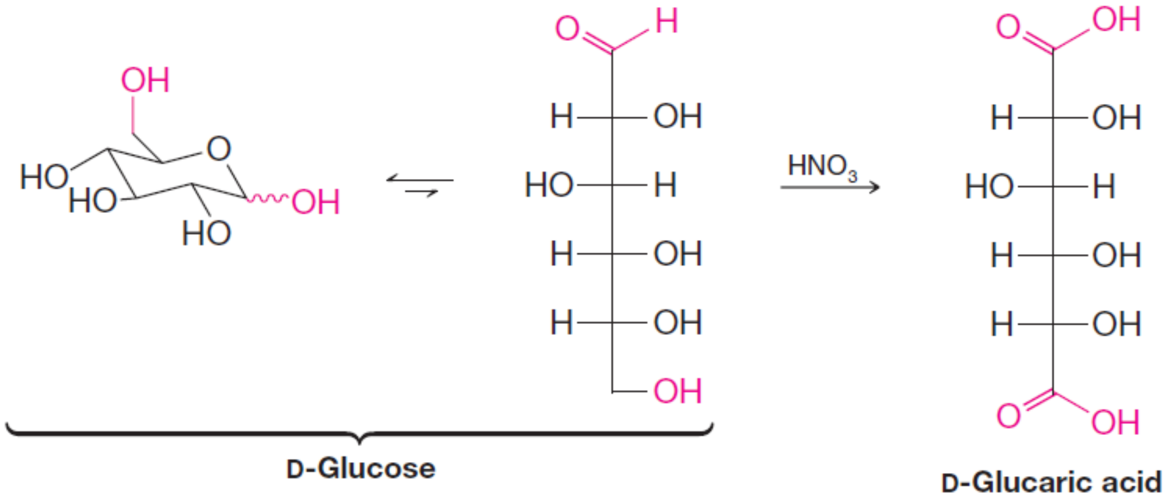


6.2.5. Oxidation Reactions of Monosaccharides

c. Nitric Acid Oxidation: Aldaric Acids

- Dilute nitric acid is a stronger oxidizing agent than bromine water , it oxidizes both the −CHO group and the terminal $-CH_2OH$ group of an aldose to $-CO_2H$ groups, forming dicarboxylic acids are known as **aldaric acids**.

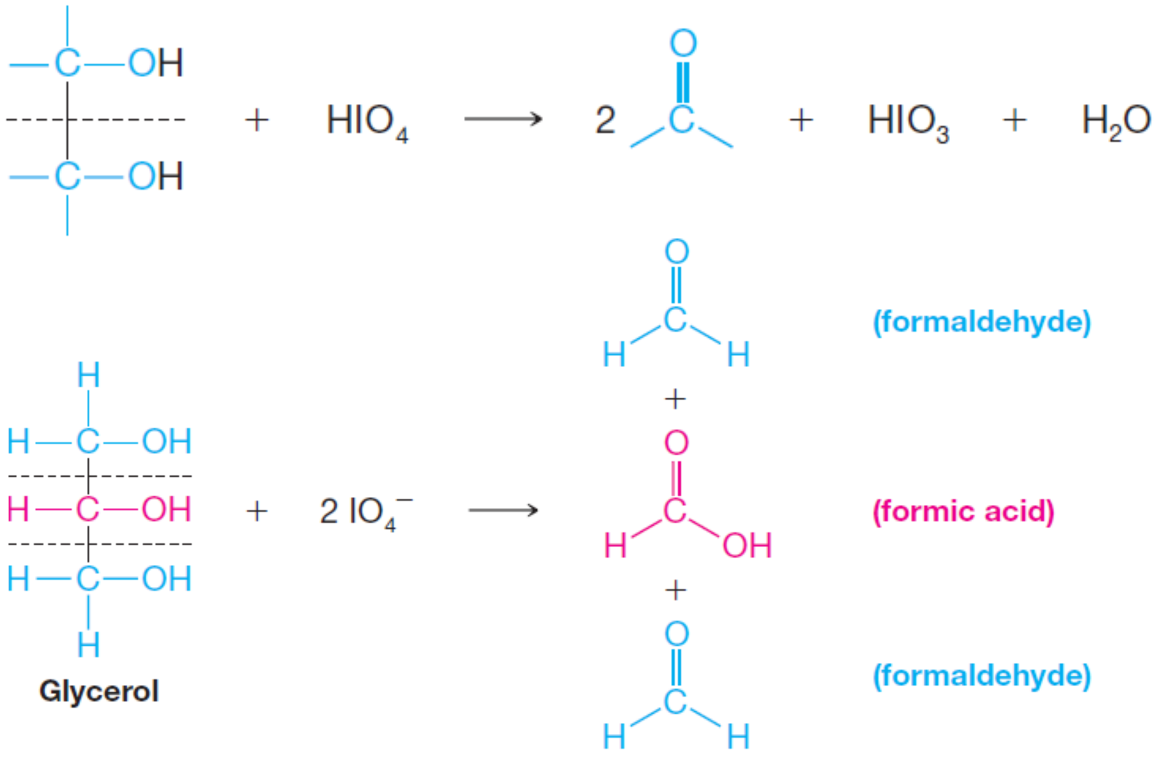




6.2.5. Oxidation Reactions of Monosaccharides

d. Periodate Oxidations: Oxidative Cleavage of Polyhydroxy Compounds

- Compounds with adjacent hydroxyl groups undergo oxidative cleavage with aqueous periodic acid (HIO_4), breaking C–C bonds to form carbonyl compounds (aldehydes, ketones, or acids).
- When three or more −CHOH groups are contiguous, the internal ones are obtained as formic acid.
- For example: Periodate oxidation of glycerol gives two molar of formaldehyde and one molar of formic acid.
- Oxidative cleavage also takes place when an −OH group is adjacent to the carbonyl group of an aldehyde or ketone (but not that of an acid or an ester).



6.2.5. Oxidation Reactions of Monosaccharides

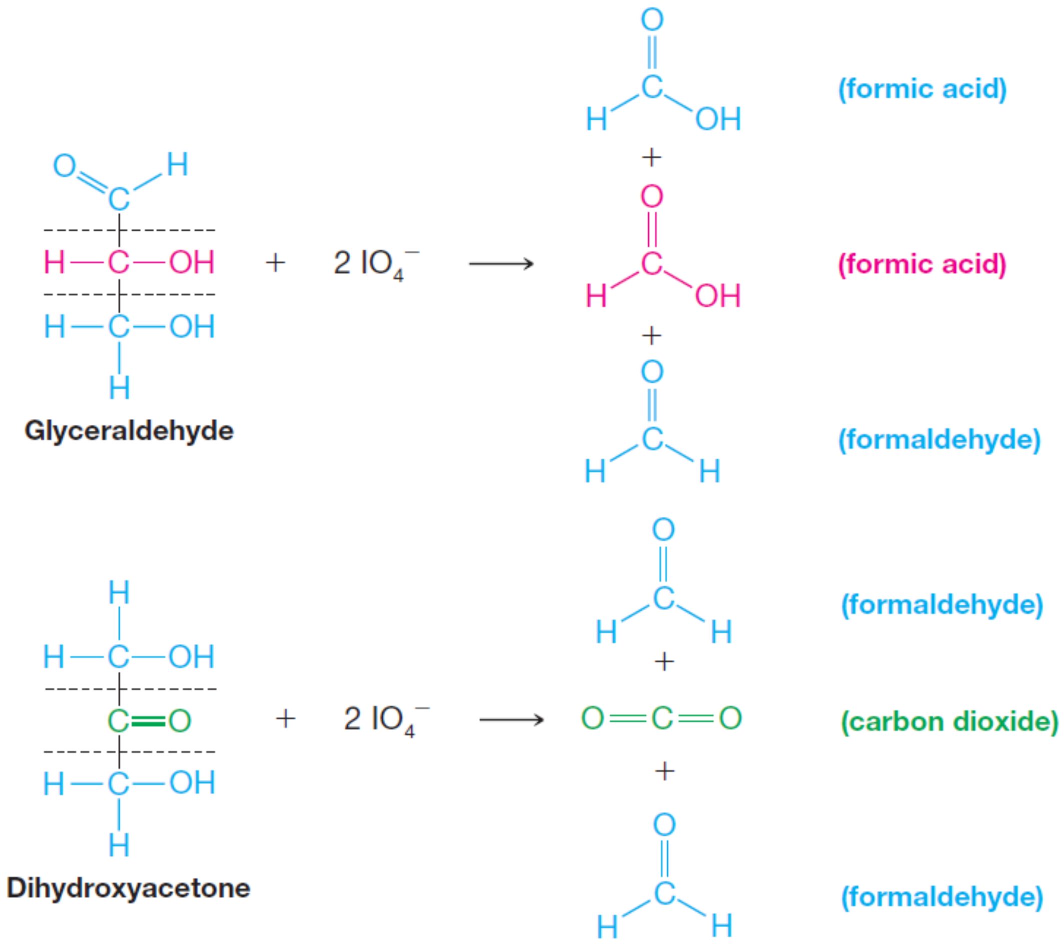
d. Periodate Oxidations: Oxidative Cleavage of Polyhydroxy Compounds

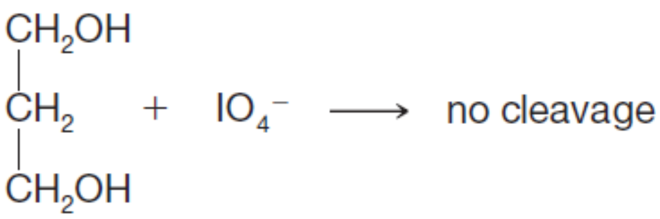
Oxidative cleavage also takes place when an −OH group is adjacent to the carbonyl group of an aldehyde or ketone (but not that of an acid or an ester).

Glyceraldehyde yields two molar of formic acid and one molar of formaldehyde.

Dihydroxyacetone gives two molar of formaldehyde and one molar of carbon dioxide.

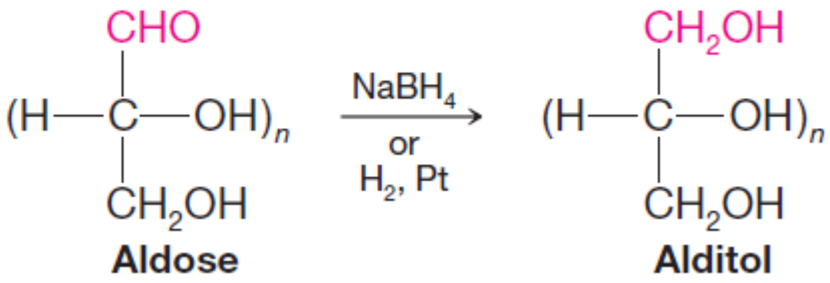
Periodic acid does not cleave compounds in which the hydroxyl groups are separated by an intervening −CH₂−



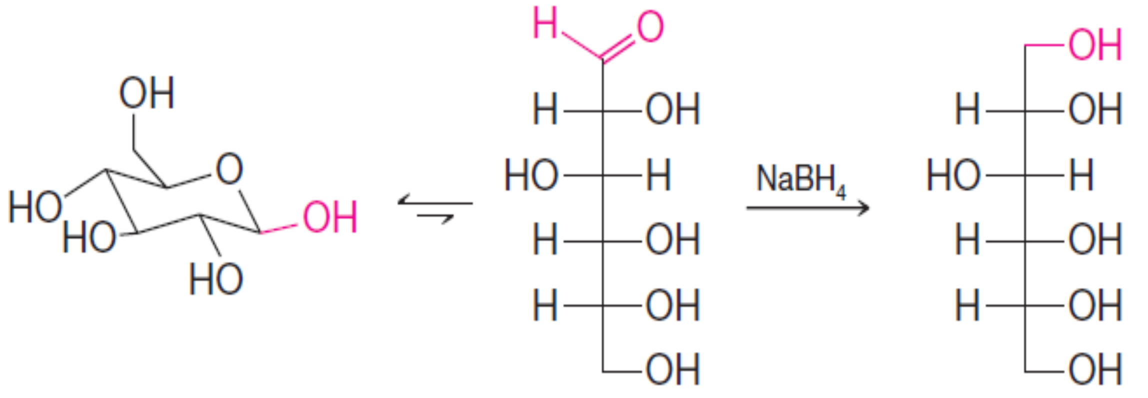


6.2.5. Reduction of Monosaccharides: Alditols

- Aldoses (and ketoses) can be reduced with sodium borohydride to give compounds called **alditols**.



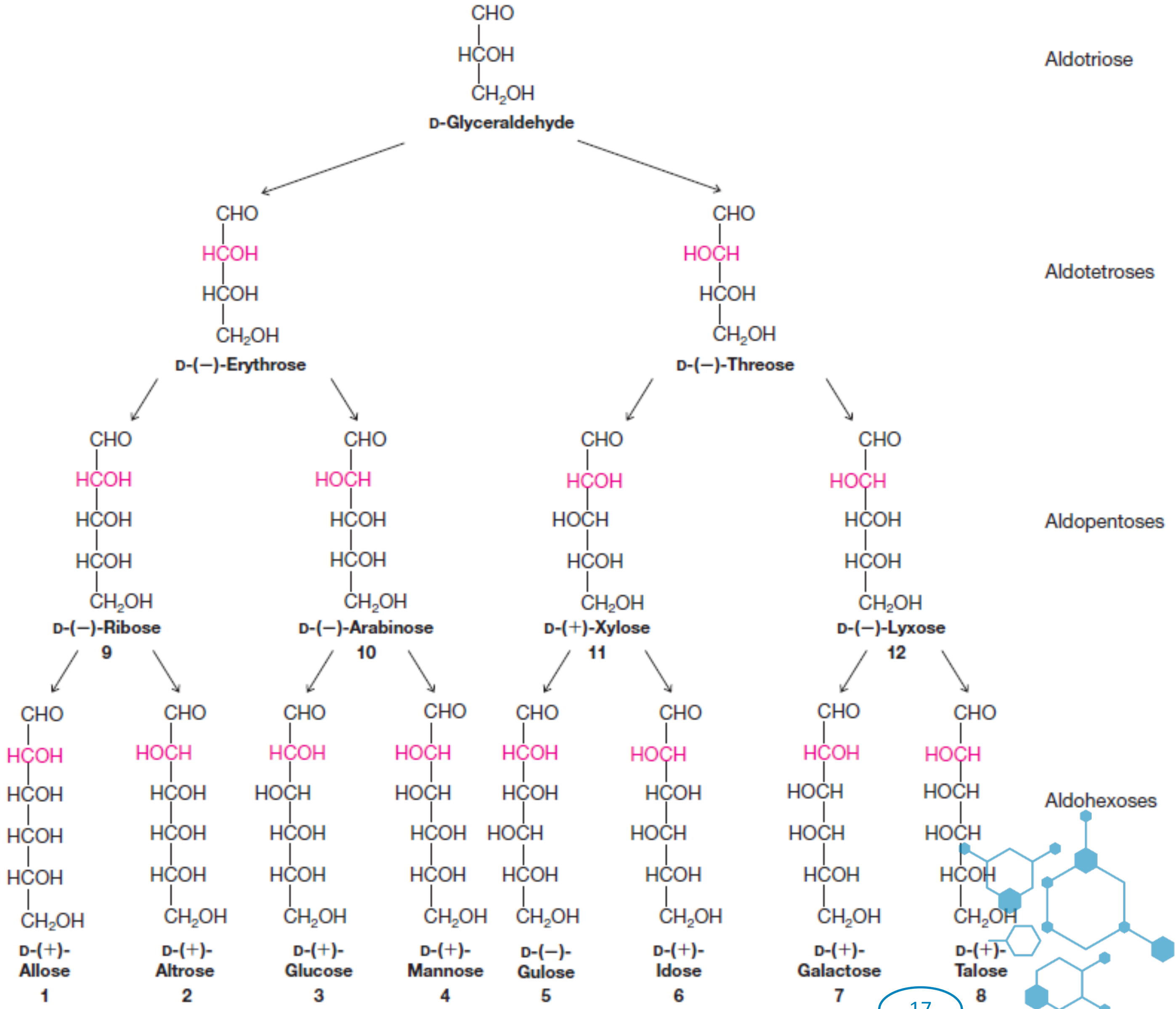
- Reduction of D-glucose, for example, yields D-glucitol



D-Glucitol
(or D-sorbitol)

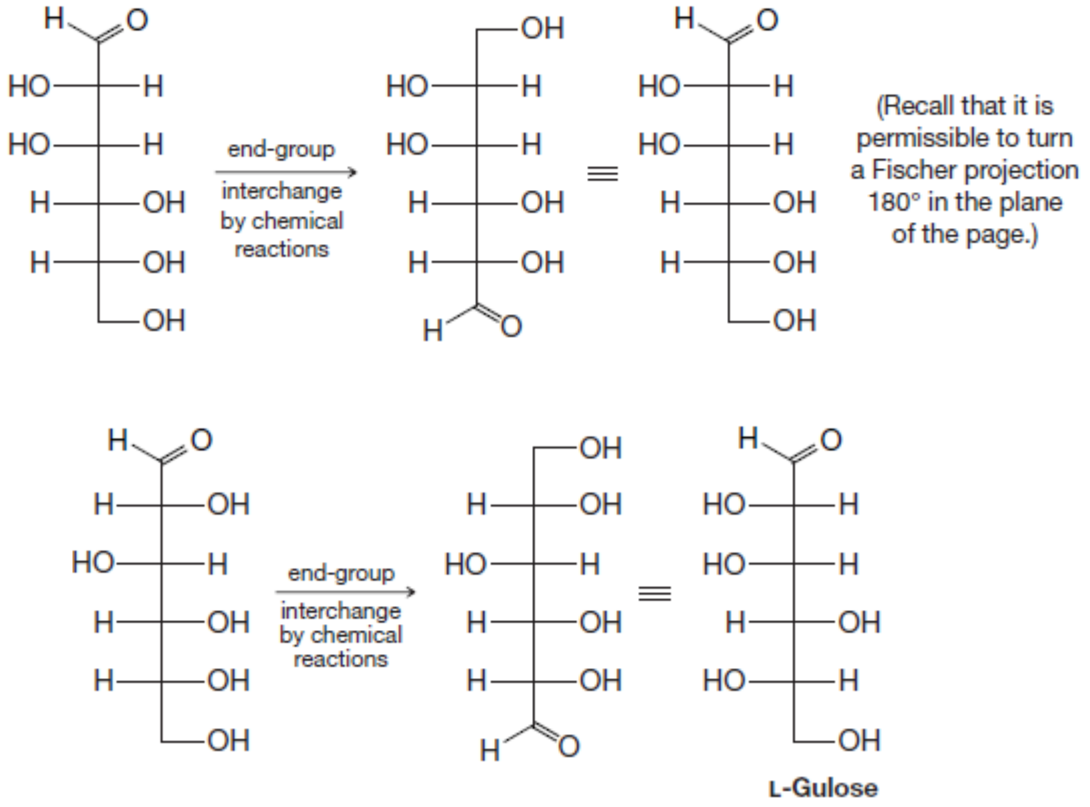
6.2.6. The D Family of Aldoses

- The Ruff degradation and the Kiliani–Fischer synthesis allow us to place all of the aldoses into families or “family trees” based on their relation to D- or L-glyceraldehyde.
- Most, but not all, of the naturally occurring aldoses belong to the d family, with D-(+)-glucose being by far the most common.



6.2.7. Fischer’s Proof of the Configuration of D-(+)-Glucose

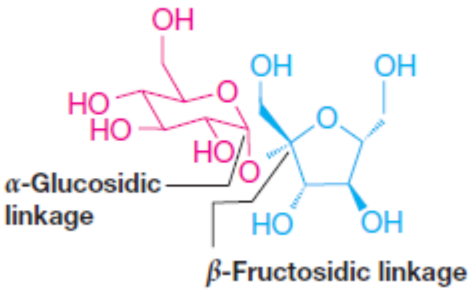
- Emil Fischer began his work on the stereochemistry of (+)-glucose in 1888.
- Glucose is an aldohexose with 4 chiral centers, giving many possible stereoisomers.
- Fischer focused on 8 D-configured structures for simplicity.
- Assignment of (+)-glucose structure was based on:
- Nitric acid oxidation and degradation to arabinose to eliminate unlikely structures.
- Kiliani–Fischer synthesis to confirm relationships between glucose, mannose, and arabinose.
- Epimer analysis and end-group interchange to assign the correct structure of D-(+)-glucose.



6.2.8. Disaccharides

a. Sucrose

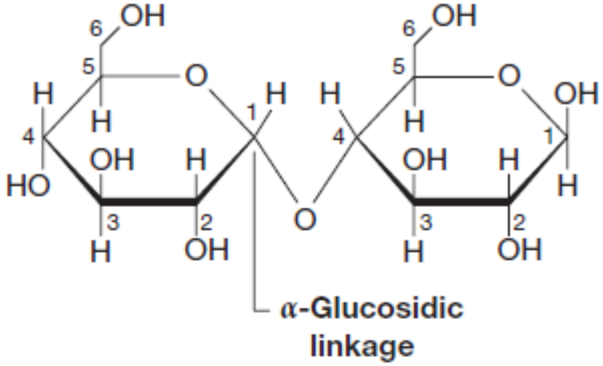
- Sucrose is the most widely occurring disaccharide, obtained commercially from sugarcane or sugar beets.
- It is a nonreducing sugar and has the molecular formula $C_{12}H_{22}O_{11}$.
- Acid-catalyzed hydrolysis of 1 mole of sucrose yields 1 mole of D-glucose and 1 mole of D-fructose



(+)-Sucrose
2-O-(α-D-glucopyranosyl)-β-D-fructofuranoside

b. Maltose

- Partial hydrolysis of starch gives Maltose.
- When 1 mol of maltose is subjected to acid-catalyzed hydrolysis, it yields 2 mol of D-(+)-glucose.
- Maltose is a reducing sugar; it gives positive tests with Benedict's, and Tollens' solutions.



(+)-maltose,
4-O-(α-D-glucopyranosyl)-β-D-glucopyranose.

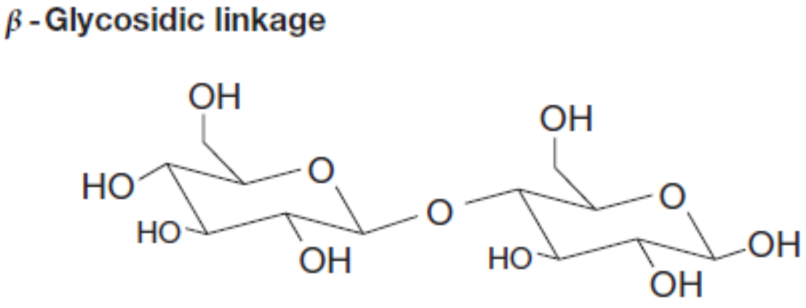
6.2.8. Disaccharides

c. Cellobiose

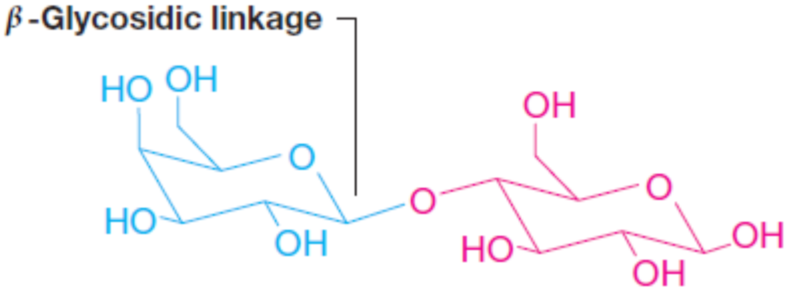
- Partial hydrolysis of cellulose gives cellobiose.
- Cellobiose resembles maltose in every respect except one: the configuration of its glycosidic linkage.
- Cellobiose is a reducing sugar that, on acid-catalyzed hydrolysis, yields two molar equivalents of D-glucose

d. Lactose

- Lactose is a disaccharide present in the milk of humans, cows, and almost
- all other mammals.
- Lactose is a reducing sugar that hydrolyzes to yield d-glucose and
- D-galactose; the glycosidic linkage is β.



Cellobiose,
4-O-(β-D-glucopyranosyl)-β-D-glucopyranose.



Lactose,
4-O-(β-D-galactopyranosyl)-β-D-glucopyranose.

6.2.9. Polysaccharides

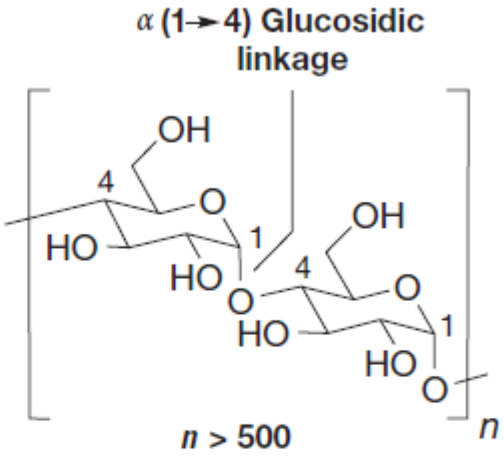
- Polysaccharides are polymers consisting of monosaccharides joined together by glycosidic linkages.
- Homopolysaccharides are polymers made of a single type of monosaccharide.
- Heteropolysaccharides are polymers made of more than one type of monosaccharide.

a. Starch

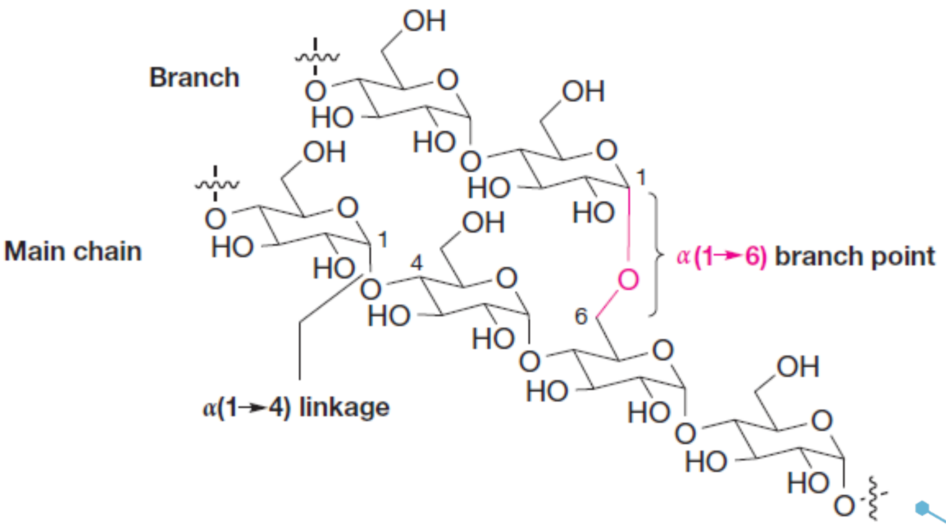
- Starch occurs as microscopic granules in roots, tubers, and seeds (e.g., corn, potatoes, wheat, rice).
- Heating starch with water swells the granules, forming a colloidal suspension.
- Starch is a nonreducing polysaccharide composed of amylose (10–20%) and amylopectin (80–90%).

6.2.9. Polysaccharides

- **Amylose** typically consists of more than 1000 D-glucopyranoside units connected in α linkages between C1 of one unit and C4 of the next.
- **Amylopectin** has a structure similar to amylose, with $\alpha(1\rightarrow4)$ glycosidic links, but its chains are branched.
- The branching occurs between C6 of one glucose unit and C1 of another.
- Branch points appear approximately every 20–25 glucose units.
- Amylopectin has a molecular weight of 1–6 million, consisting of hundreds of interconnected chains, each containing about 20–25 glucose units.



Partial structure of Amylose



Partial structure of Amylopectin

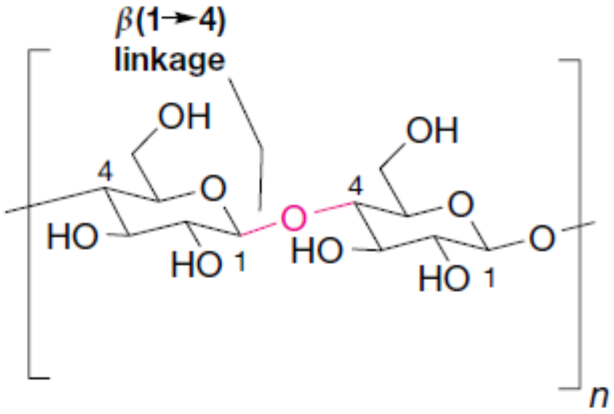
6.2.9. Polysaccharides

b. Glycogen

- Glycogen has a structure very much like that of amylopectin; however, in glycogen the chains are much more highly branched.
- Glycogen has a very high molecular weight, and samples isolated under conditions minimizing hydrolysis show molecular weights as high as 100 million.

c. Cellulose

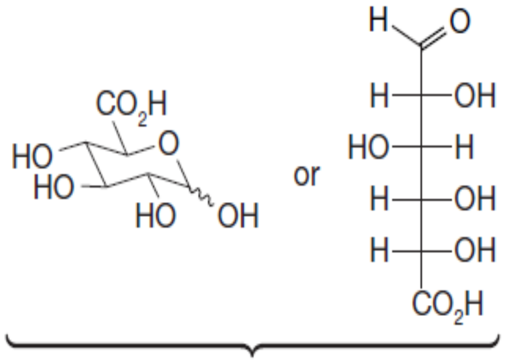
- Cellulose consists of D-glucopyranose units linked by β(1→4) glycosidic bonds forming very long, unbranched chains.
- The β-linkages make cellulose chains linear, unlike the helical α(1→4) chains in starch and glycogen.
- Hydroxyl (–OH) groups on adjacent chains form extensive hydrogen bonds, producing a rigid, insoluble, fibrous structure.
- The precise stereochemistry of D-glucose allows the formation of strong linear fibers ideal for plant cell walls.



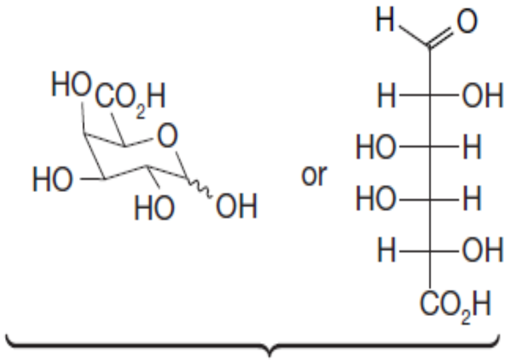
Partial structure of Cellulose

6.2.10. Other Biologically Important Sugars

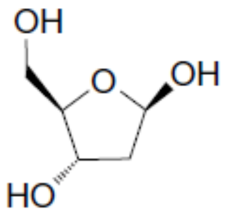
- Monosaccharide derivatives in which the –CH_2OH group at C6 is oxidized to a carboxyl group are called uronic acids.
- Their names are derived from the parent monosaccharide; for example, oxidation of C6 in glucose forms glucuronic acid, while oxidation of C6 in galactose yields galacturonic acid.
- Monosaccharides in which an **–OH** group is replaced by **–H** are called **deoxy sugars**; the most important is deoxyribose (found in DNA), while L-rhamnose and L-fucose are common in polysaccharides.



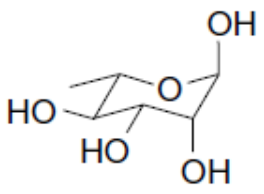
D-Glucuronic acid



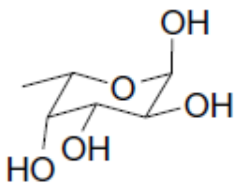
D-Galacturonic acid



β-2-Deoxy-D-ribose



α-L-Rhamnose
(6-deoxy-L-mannose)

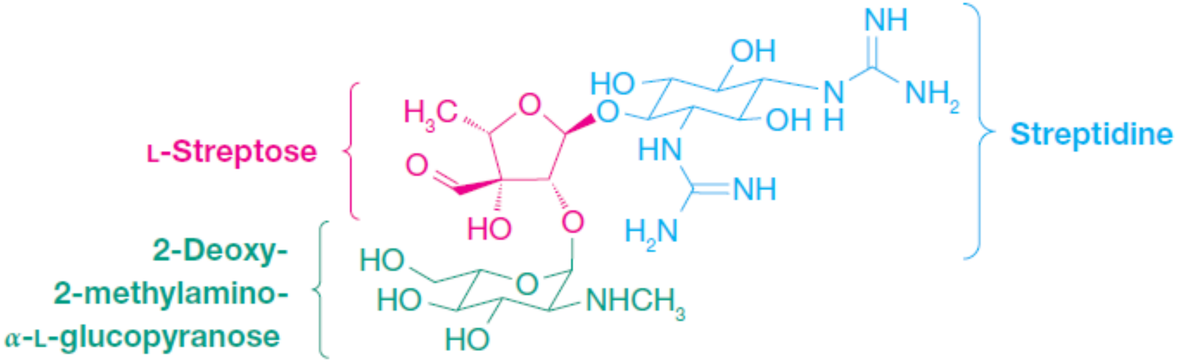


α-L-Fucose
(6-deoxy-L-galactose)

6.2.10. Carbohydrate Antibiotics

- One of the important discoveries in carbohydrate chemistry was the isolation (in 1944) of the carbohydrate antibiotic streptomycin.
- Streptomycin disrupts bacterial protein synthesis.
- Composed of three subunits:

❑ An amino sugar based on L-glucose.
❑ A branched-chain monosaccharide (streptose).
❑ An amino cyclitol (streptidine), a cyclohexane derivative.



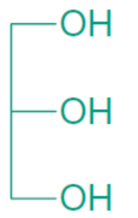
Lipids

6.2. Lipids

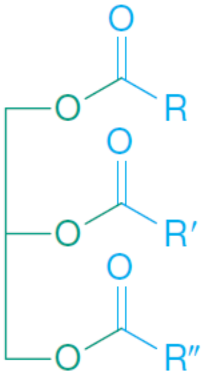
- **Lipids** are compounds of biological origin that dissolve in nonpolar solvents, such as chloroform and diethyl ether.
- The name lipid comes from the Greek word *lipos*, for fat.

6.2.1. Fatty Acids and Triacylglycerols

- **Triacylglycerols** (**Triglycerides**) are the oils of plants and the fats of animal origin.
- **Triacylglycerols** are carboxylic acids of biological origin that are found as esters of glycerol.
  - The groups R, R′, and R″ are usually long chain alkyl groups.
  - R, R′, and R″ may also contain one or more carbon–carbon double bonds.
  - R, R′, and R″ may all be different.
- **Triacylglycerols** include such common substances as peanut oil, soybean oil, corn oil, sunflower oil, butter, lard, and tallow.



Glycerol

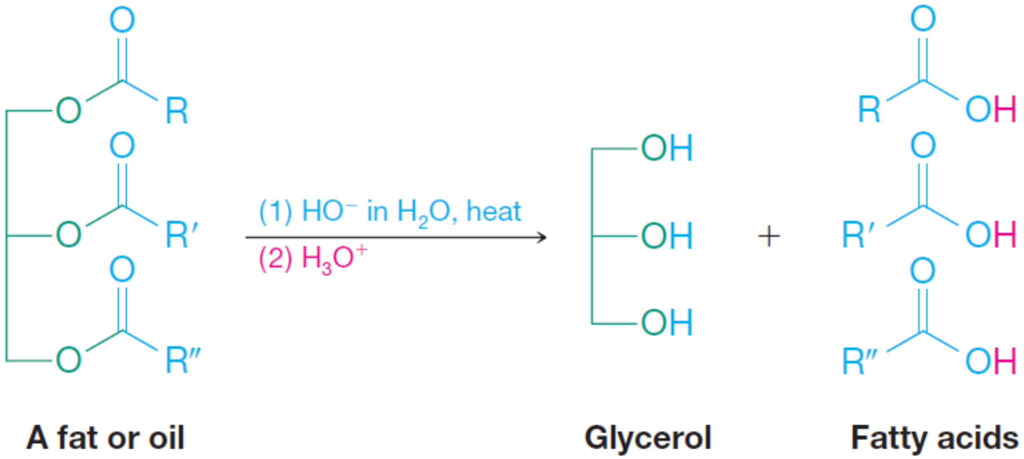


A triacylglycerol

6.2. Lipids

6.2.1. Fatty Acids and Triacylglycerols

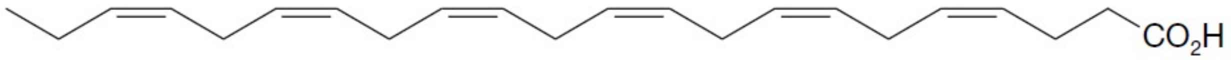
- **Triacylglycerols are called**
  - **Oils;** liquids at room temperature.
  - **Fats;** solids at room temperature.
- **Fats;** Triacylglycerols made up of largely saturated fatty acids have high melting points and are solids at room temperature.
- **Oils;** Triacylglycerols with a high proportion of unsaturated and polyunsaturated fatty acids have lower melting points.
- Triacylglycerols can be
  - **Simple triacylglycerols**; all three acyl groups are the same.
  - **Mixed triacylglycerols**; the acyl groups are different.
- Hydrolysis of a fat or oil produces a mixture of fatty acids
  - Most natural fatty acids have unbranched chains
  - Fatty acids have an even number of carbon atoms..



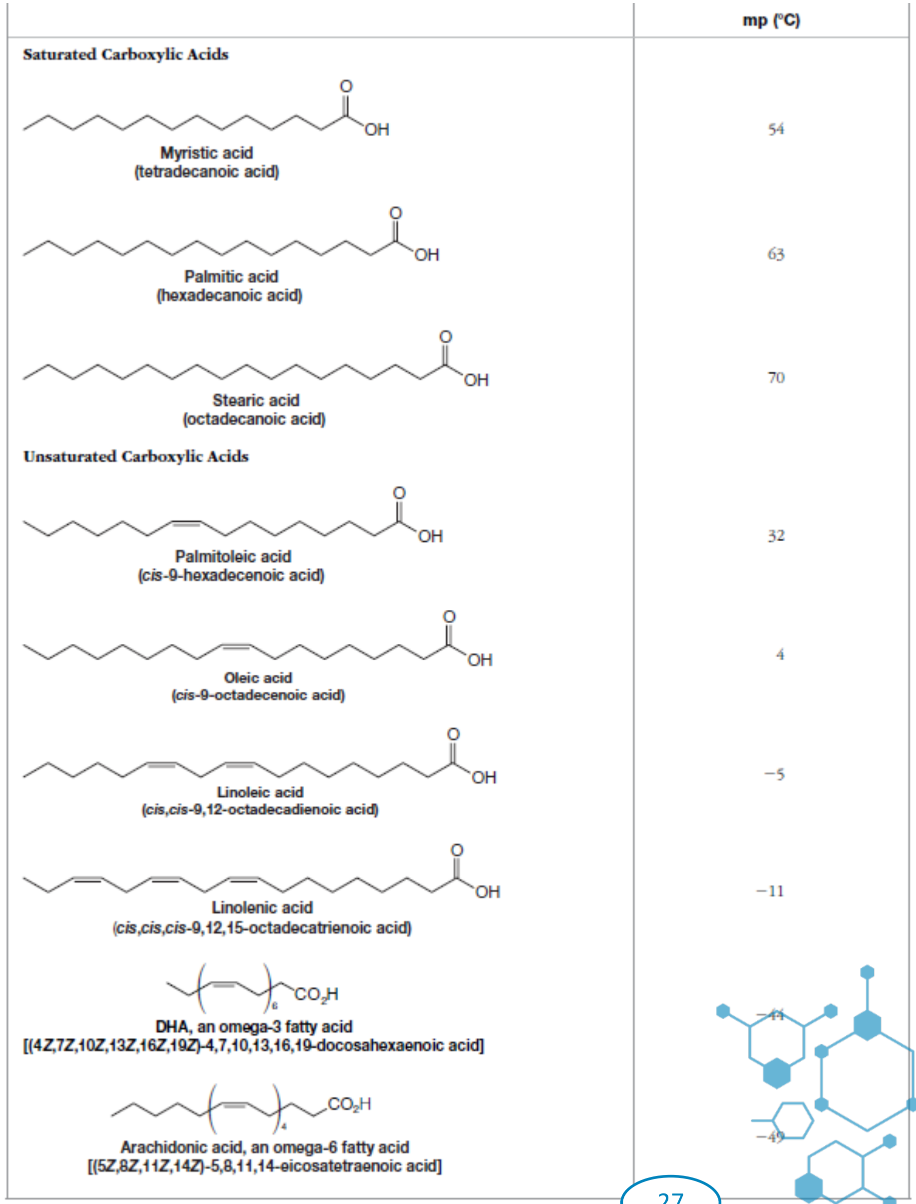
6.2. Lipids

6.2.1. Fatty Acids and Triacylglycerols

- **Unsaturated fatty acids,**
  - The double bonds are all *cis*.
  - Many naturally occurring fatty acids contain two or three double bonds.
  - The fats or oils that these come from are called polyunsaturated fats or oils.
  - The first double bond of an unsaturated fatty acid commonly occurs between C9 and C10
- **Saturated fatty acids** have relatively high melting points
- **Omega-3 fatty acids** are those where the third to last carbon in the chain is part of a carbon–carbon double bond



(4Z,7Z,10Z,13Z,16Z,19Z)-4,7,10,13,16,19-Docosahexaenoic acid
(DHA, an omega-3 fatty acid)

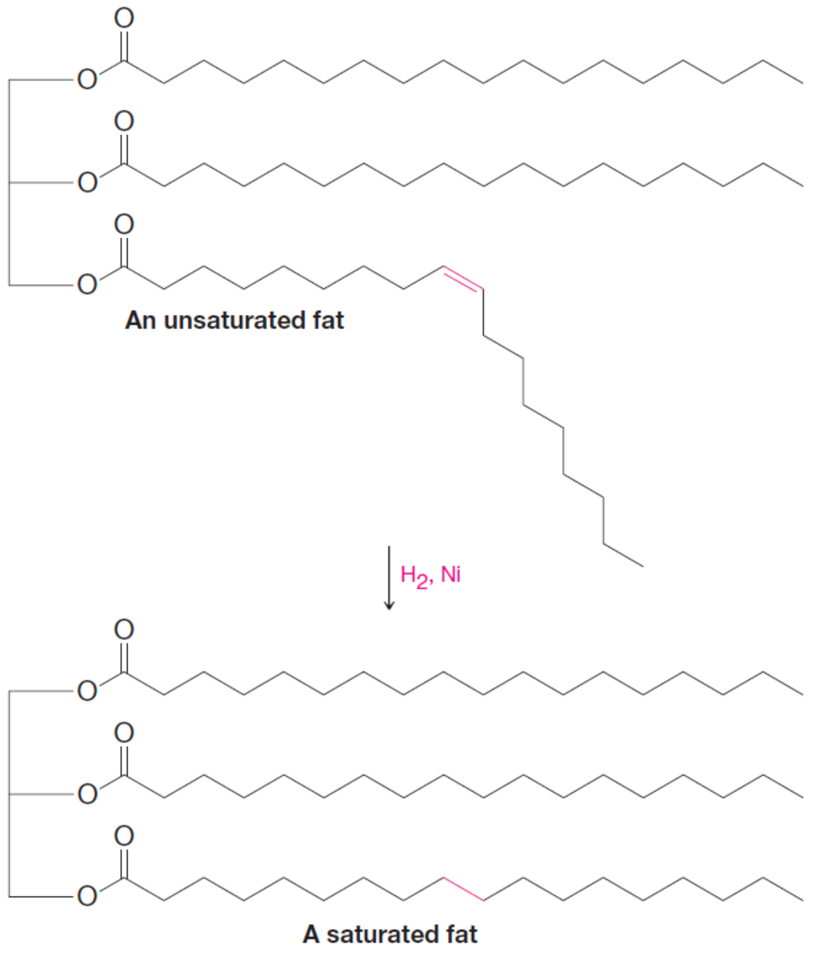


	mp (°C)
Saturated Carboxylic Acids	
Myristic acid (tetradecanoic acid)	54
Palmitic acid (hexadecanoic acid)	63
Stearic acid (octadecanoic acid)	70
Unsaturated Carboxylic Acids	
Palmitoleic acid (*cis*-9-hexadecenoic acid)	32
Oleic acid (*cis*-9-octadecenoic acid)	4
Linoleic acid (*cis*,*cis*-9,12-octadecadienoic acid)	−5
Linolenic acid (*cis*,*cis*,*cis*-9,12,15-octadecatrienoic acid)	−11
DHA, an omega-3 fatty acid [(4Z,7Z,10Z,13Z,16Z,19Z)-4,7,10,13,16,19-docosahexaenoic acid]	−44
Arachidonic acid, an omega-6 fatty acid [(5Z,8Z,11Z,14Z)-5,8,11,14-eicosatetraenoic acid]	−49

6.2. Lipids

6.2.1.1. Hydrogenation of Triacylglycerols

- **Fats;** Triacylglycerols made up of largely saturated fatty acids have high melting points and are solids at room temperature.
- **Oils;** Triacylglycerols with a high proportion of unsaturated and polyunsaturated fatty acids have lower melting points.
- Catalytic hydrogenation can be used to convert an unsaturated triacylglycerol into a saturated one.
- **Partially hydrogenated fat;** Solid commercial cooking fats are manufactured by partial hydrogenation of vegetable oils.
- **Complete hydrogenation** of the oil is avoided because a completely saturated triacylglycerol is very hard and brittle.



6.2. Lipids

6.2.1.2. Biological Functions of Triacylglycerols

- The primary function of triacylglycerols in animals is as an energy reserve

When triacylglycerols are metabolized, they yield more than twice as many kilocalories per gram as do carbohydrates or proteins.

This is largely because of the high proportion of carbon–hydrogen bonds per molecule and they are converted to carbon dioxide and water by biochemical reactions.

Some of the unsaturated ones, can be synthesized from carbohydrates and proteins.

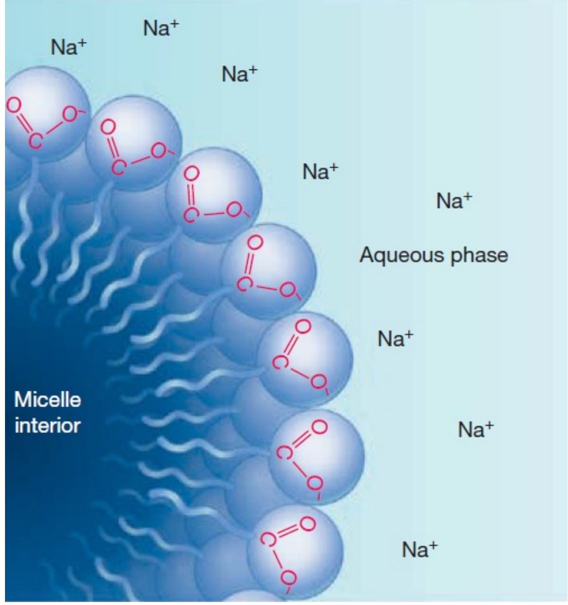
Certain polyunsaturated fatty acids are essential in the diets of higher animals.

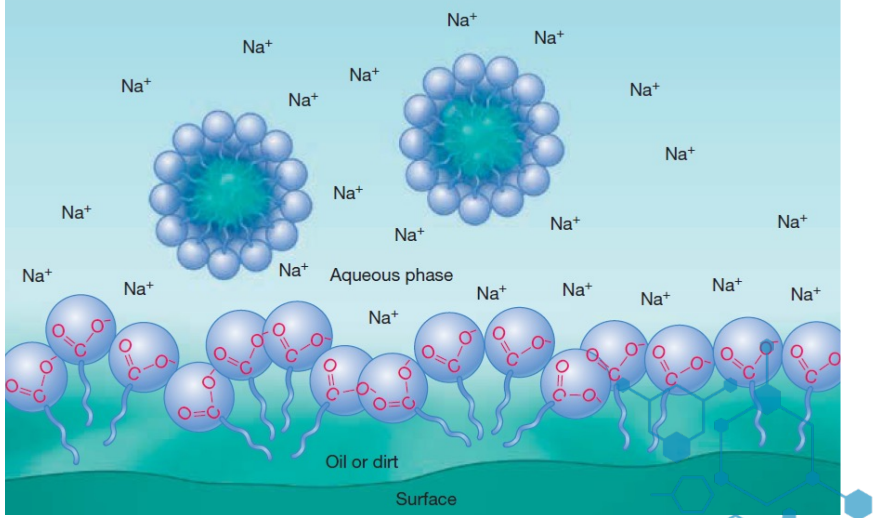
The amount of fat in the diet, especially the proportion of saturated fat, has been a health concern for many years.

6.2. Lipids

6.2.1.3. Saponification of Triacylglycerols

- **Saponification** is the alkaline hydrolysis of triacylglycerols, leading to glycerol and a mixture of salts of long-chain carboxylic acids.
- **Soap;**
  - **Soap** are salts of long-chain carboxylic acids.
  - The sodium salts of long-chain carboxylic acids (soaps) are almost completely miscible with water.
  - They do not dissolve as we might expect, that is, as individual ions.
  - Soaps exists as ***micelles*** in very dilute solutions
  - Micelle formation accounts for the fact that soaps dissolve in water.
  - The nonpolar (and thus hydrophobic) alkyl chains of the soap remain in a nonpolar environment—in the interior of the micelle.
  - The polar (and therefore hydrophilic) carboxylate groups are exposed to a polar environment—that of the aqueous phase.

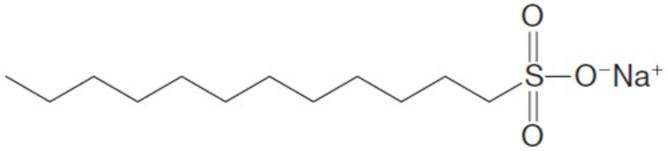




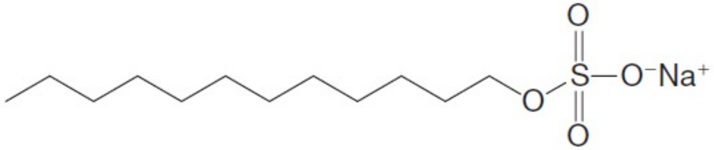
6.2. Lipids

6.2.1.3. Saponification of Triacylglycerols

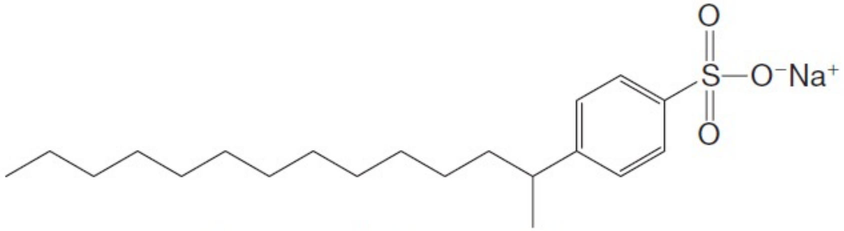
- **Synthetic detergents ;**
  - **Synthetic detergents** function in the same way as soaps; they have long nonpolar alkane chains with polar groups at the end..
  - The polar groups of most synthetic detergents are sodium sulfonates or sodium sulfates
  - These detergents proved to be nonbiodegradable.
  - Synthetic detergents offer an advantage over soaps; they function well in "hard" water, that is, water containing Ca^{2+}, Fe^{2+}, Fe^{3+}, and Mg^{2+} ions.



A sodium alkanesulfonate



A sodium alkyl sulfate

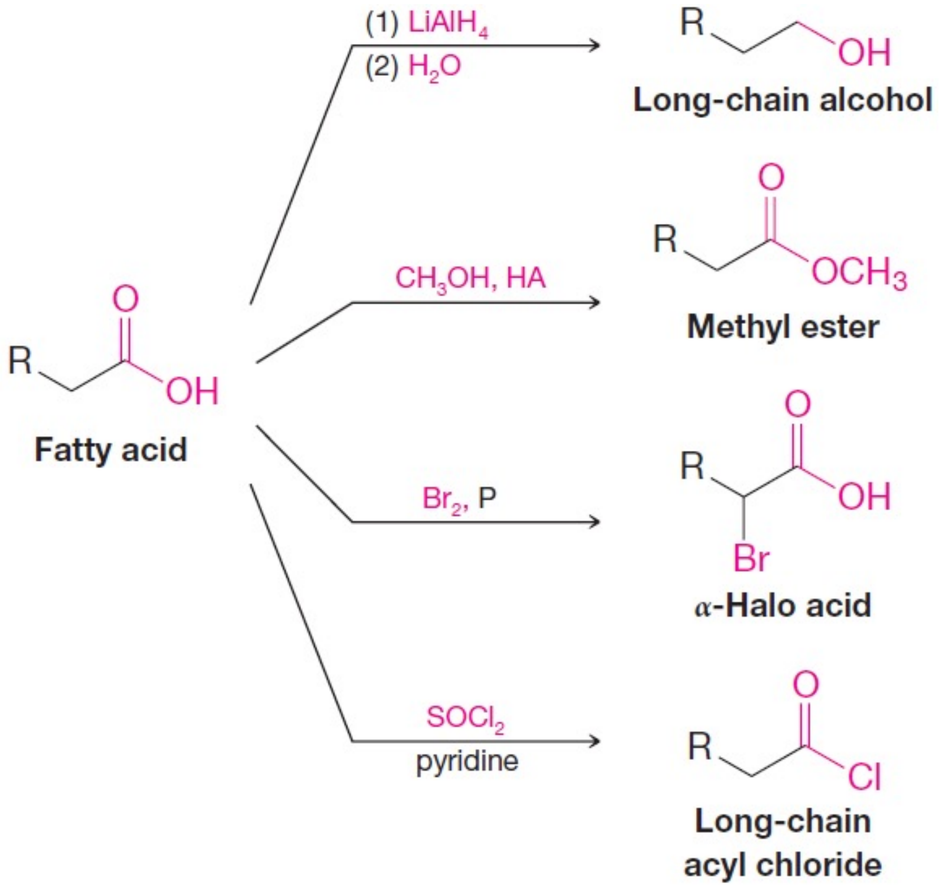


A sodium alkylbenzenesulfonate

6.2. Lipids

6.2.1.4. Reactions of the Carboxyl Group of Fatty Acids

- Fatty acids undergo reactions typical of carboxylic acids.
  - They react with LiAlH4 to form alcohols,
  - They react with alcohols and mineral acid to form esters,
  - They react with bromine and phosphorus to form α-halo acids,
  - They react with thionyl chloride to form acyl chlorides

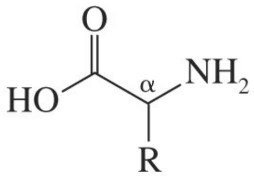


Amino acids and Proteins

6.3. Amino acids and Proteins

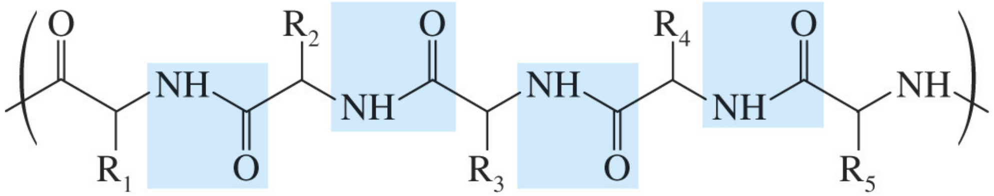
6.3.1. Introduction:

- The three major groups of biological polymers are polysaccharides (carbohydrates), proteins and nucleic acids.
- Proteins have many diverse functions; they are major components of the following biomolecules.
- Enzymes and hormones which catalyze and regulate biological reactions.
- Muscles and tendons which provide the body with means for movement.
- Hemoglobin which carries oxygen to all parts of the body.
- Antibodies they are integral parts of the immune system.
- All proteins are polyamides: Their monomeric units are one of about 20 α-amino acids.



An α-amino acid

R is a side chain at the α carbon that determines the identity of the amino acid (Table 24.1).



A portion of a protein molecule

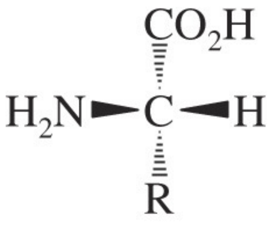
Amide (peptide) linkages are shaded.

R_1–R_5 may be any of the possible side chains.

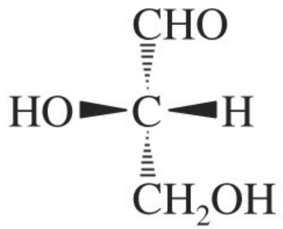
6.3. Amino acids and Proteins

Structure and Names

- **All natural occurring amino acids** are chiral (except glycine which is achiral) and have the L-configuration (as related to glyceraldehyde) at the *α*-carbon.
- **Nomenclature and Classification of amino acids:**
- The 20 amino acids are known by their common names (IUPAC nomenclature could be applied) and abbreviated by **three letter** based on its common name, which is used when writing the formulas of peptides, and a **one-letter** abbreviation used to describe the amino acid sequence in a protein (see the following Table). **For example; Glycine = Gly or (G); Alanine = Ala or (A); Valine = Val or (V), etc..**



An L-*α*-amino acid [usually an (*S*)-*α*-amino acid]



L-Glyceraldehyde [(*S*)-glyceraldehyde]

Fischer projections for an L-*α*-amino acid and L-glyceraldehyde

6.3. Amino acids and Proteins

Classification of amino acids

- Amino acids are classified according to the structure of R group:

1. The acidity of R group: neutral or acidic or basic as shown in the Table.

- Neutral R Group: Are those which containing one amino group and one carboxylic group. The majority of amino acids contain neutral R group.
- Acidic R group: Are those amino acids that contain more than one carboxylic group such as

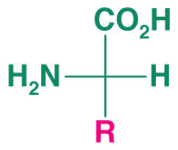
2. Polarity of the R group: Amino acid could be classified into polar (hydrophilic) or non-polar (hydrophobic).

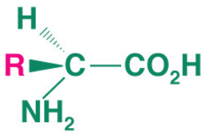
- *Essential and non-essential amino acid:*

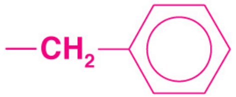
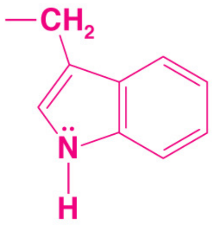
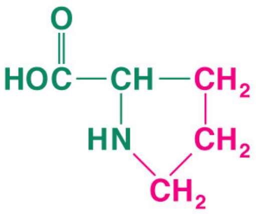
1. Essential amino acid are those amino acids which cannot be synthesized by adult humans and therefore must be included in the diet in the form of proteins. There are eight amino acids Valine. Leucine, Isoleucine, Threonine, Methionine, Phenylalanine, Tryptophan, and Lysine.

2. Non-essential amino acids: These can be synthesized in the body from other foods and include Glycine, Alanine, Serine, Cysteine, Proline, Tyrosine, Aspartic acid, Glutamic acid, Asparagine, Glutamine, Arginine, and Histidine.

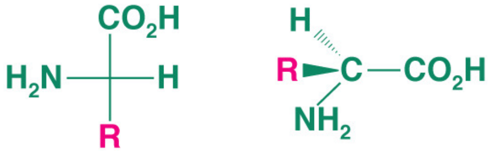
Chapter 6: Biomolecules and their Biological Relevance

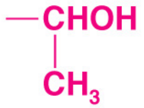
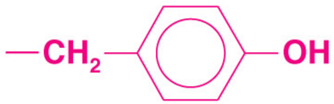
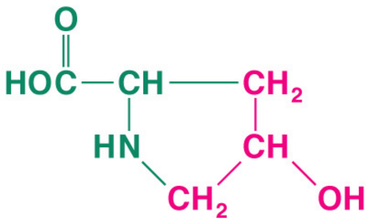
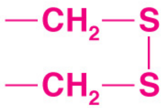


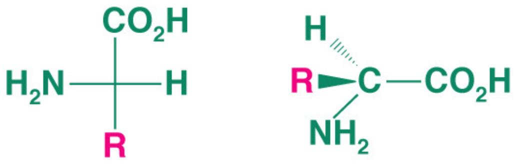


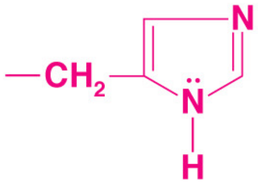
Structure of R	Name	Abbreviations[a]	pK_{a_1} α-CO_2H	pK_{a_2} α-NH_3^+	pK_{a_3} R group	pI
Neutral Amino Acids						
$-H$	Glycine	G or Gly	2.3	9.6		6.0
$-CH_3$	Alanine	A or Ala	2.3	9.7		6.0
$-CH(CH_3)_2$	Valine[b]	V or Val	2.3	9.6		6.0
$-CH_2CH(CH_3)_2$	Leucine[b]	L or Leu	2.4	9.6		6.0
 $-\underset{\displaystyle CH_3}{\underset{\vert}{C}H}CH_2CH_3$	Isoleucine[b]	I or Ile	2.4	9.7		6.1
	Phenylalanine[b]	F or Phe	1.8	9.1		5.5
$-CH_2CONH_2$	Asparagine	N or Asn	2.0	8.8		5.4
$-CH_2CH_2CONH_2$	Glutamine	Q or Gln	2.2	9.1		5.7
	Tryptophan[b]	W or Trp	2.4	9.4		5.9
 **(complete structure)**	Proline	P or Pro	2.0	10.6		6.3

Chapter 1: Introduction to Organic Chemistry



Structure of R	Name	Abbreviations[a]	pK_{a_1} α-CO_2H	pK_{a_2} α-NH_3^+	pK_{a_3} R group	p*I*
Neutral Amino Acids						
$—CH_2OH$	Serine	S or Ser	2.2	9.2		5.7
 $—\underset{\displaystyle CH_3}{CHOH}$	Threonine[b]	T or Thr	2.6	10.4		6.5
	Tyrosine	Y or Tyr	2.2	9.1	10.1	5.7
 (complete structure)	Hydroxyproline	Hyp	1.9	9.7		6.3
$—CH_2SH$	Cysteine	C or Cys	1.7	10.8	8.3	5.0
	Cystine	Cys-Cys	1.6 2.3	7.9 9.9		5.1
$—CH_2CH_2SCH_3$	Methionine[b]	M or Met	2.3	9.2		5.8



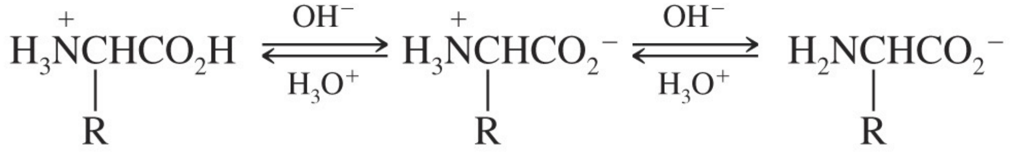
Structure of R	Name	Abbreviations[a]	pK_{a_1} α-CO_2H	pK_{a_2} α-NH_3^+	pK_{a_3} R group	p*I*
R Contains an Acidic (Carboxyl) Group						
$-CH_2CO_2H$	Aspartic acid	D or Asp	2.1	9.8	3.9	3.0
$-CH_2CH_2CO_2H$	Glutamic acid	E or Glu	2.2	9.7	4.3	3.2
R Contains a Basic Group						
$-CH_2CH_2CH_2CH_2NH_2$	Lysine[b]	K or Lys	2.2	9.0	10.5[c]	9.8
 $-CH_2CH_2CH_2NH-\overset{\overset{NH}{\parallel}}{C}-NH_2$	Arginine	R or Arg	2.2	9.0	12.5[c]	10.8
	Histidine	H or His	1.8	9.2	6.0[c]	7.6

[a] Single-letter abbreviations are now the most commonly used form in current biochemical literature.
[b] An essential amino acid.
[c] pK_a is of protonated amine of R group.

Amino acids

Amino acids as dipolar ions

- In the dry solid state amino acids exist as dipolar ions (zwittrions or inner salts).
- In aqueous solutions an equilibrium exists between the dipolar ion, the cationic and the anionic forms of the amino acid.
- The predominant form depends on the pH of the solution.
- At low pH, the amino acid exists primarily in the cationic form.
- At high pH, the amino acid exists primarily in the anionic form.
- At some intermediate pH, called the P*I* (*isoelectric point*), the concentration of the dipolar ion is at maximum and the concentrations of anionic and cationic forms are equal
- Each individual amino acid has a characteristic P*I*.



$$\mathrm{H_3\overset{+}{N}\underset{\substack{|\\ R}}{C}HCO_2H} \underset{\mathrm{H_3O^+}}{\overset{\mathrm{OH^-}}{\rightleftharpoons}} \mathrm{H_3\overset{+}{N}\underset{\substack{|\\ R}}{C}HCO_2^-} \underset{\mathrm{H_3O^+}}{\overset{\mathrm{OH^-}}{\rightleftharpoons}} \mathrm{H_2N\underset{\substack{|\\ R}}{C}HCO_2^-}$$

Cationic form (predominant in strongly acidic solutions, e.g., at pH 0)

Dipolar ion

Anionic form (predominant in strongly basic solutions, e.g., at pH 14)

Amino acids

Synthesis of Amino acids

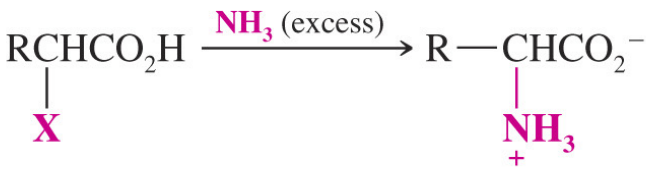
- There are many methods to prepare amino acids, two methods will be mentioned here.

1. Direct Ammonolysis of *α*-halo-carboxylic acid

   With this method, yields tend to be poor.

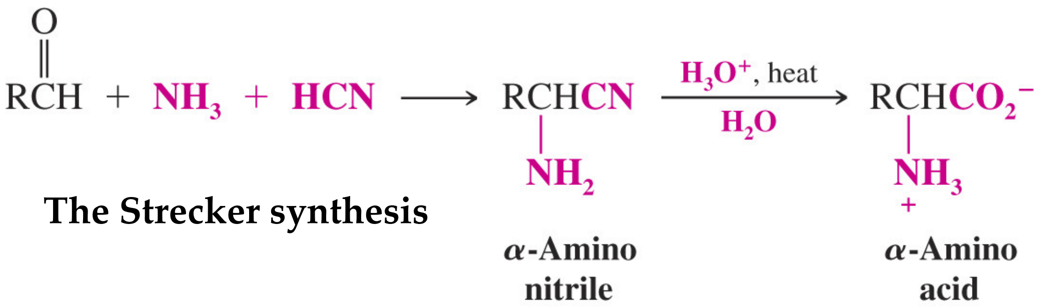
2. The Strecker Synthesis (Reductive amination)

   Treatment of an aldehyde with ammonia and hydrogen cyanide yields an *α*-aminonitrile which is hydrolyzed to the *α*-amino acid. The mechanism proceeds via an intermediate imine



$$\underset{X}{\underset{|}{RCHCO_2H}} \xrightarrow{NH_3\text{ (excess)}} R{-}\underset{\underset{+}{NH_3}}{\underset{|}{CHCO_2^-}}$$

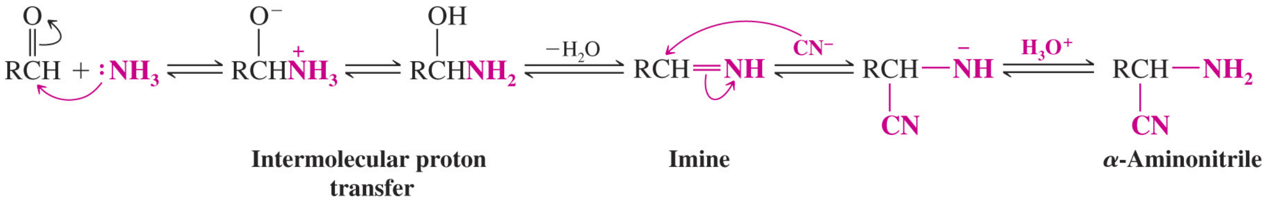
Direct ammonolysis of *α*-halocarboxylic acids



$$R\overset{O}{\overset{\|}{C}}H + NH_3 + HCN \longrightarrow \underset{NH_2}{\underset{|}{RCHCN}} \xrightarrow[H_2O]{H_3O^+\text{, heat}} \underset{\underset{+}{NH_3}}{\underset{|}{RCHCO_2^-}}$$

The Strecker synthesis

***α*-Amino nitrile** **α-Amino acid**

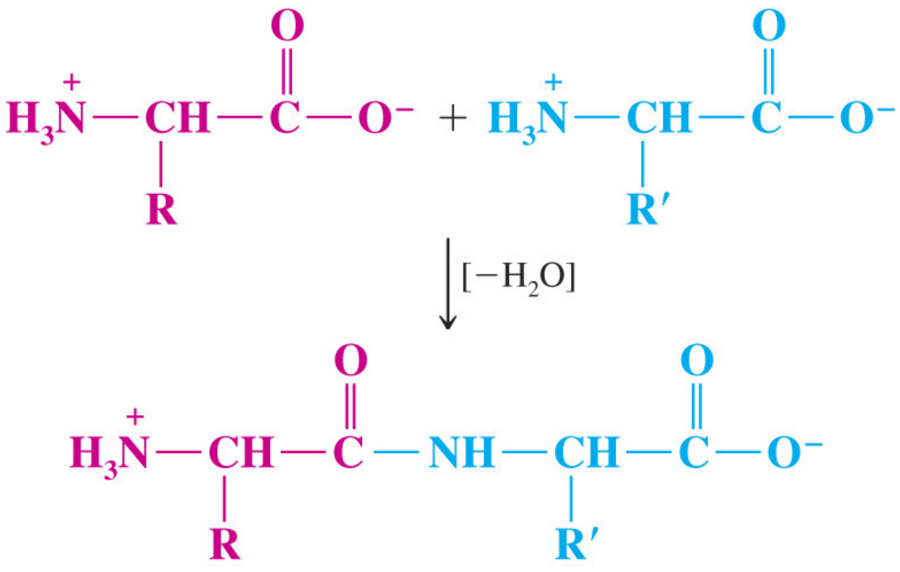


$$R\overset{O}{\overset{\|}{C}}H + :NH_3 \rightleftharpoons R\overset{O^-}{\overset{|}{C}}H\overset{+}{N}H_3 \rightleftharpoons R\overset{OH}{\overset{|}{C}}HNH_2 \xrightleftharpoons{-H_2O} RCH{=}NH \xrightleftharpoons{CN^-} \underset{CN}{\underset{|}{RCH}}{-}\overset{-}{N}H \xrightleftharpoons{H_3O^+} \underset{CN}{\underset{|}{RCH}}{-}NH_2$$

Intermolecular proton transfer **Imine** ***α*-Aminonitrile**

Polypeptides and Proteins

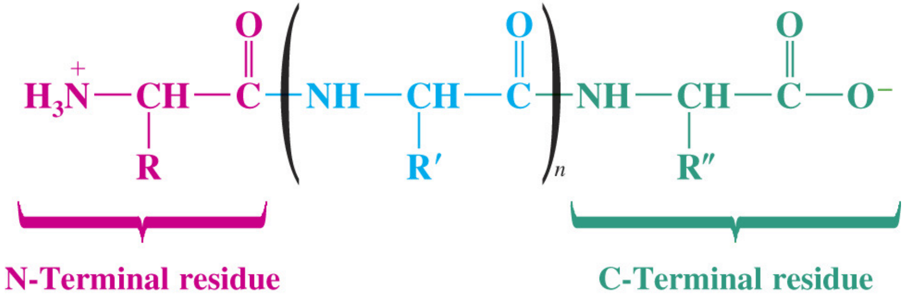
- Enzymes polymerize amino acids by forming amide linkages.
- The polymer is called a peptide and the amide linkages are called peptide bonds or peptide linkages.
- Each amino acid in the peptide is called an amino acid residue.
- Proteins can contain one or more peptide chains and others associated molecules or metal ions.

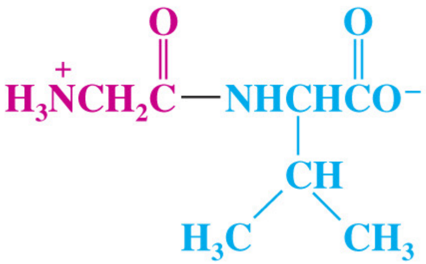


A dipeptide

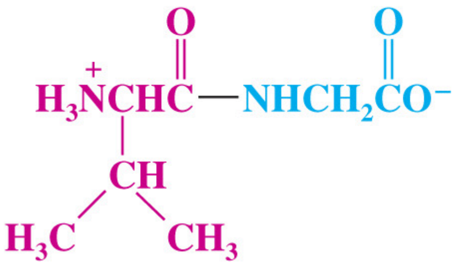
Polypeptides and Proteins

- Polypeptides are linear polymers. One end of a polypeptide chain in an amino acid residue has a free amino group $-\overset{+}{N}H_3$, the other terminate in amino acid residue with free carboxylic group $–CO_2^-$. These two groups are called the N-terminal and the C-terminal residues, respectively.
- Polypeptides are usually written with the N-terminal residue to the left and the C-terminal residue to the right.





Glycylvaline
(Gly · Val)

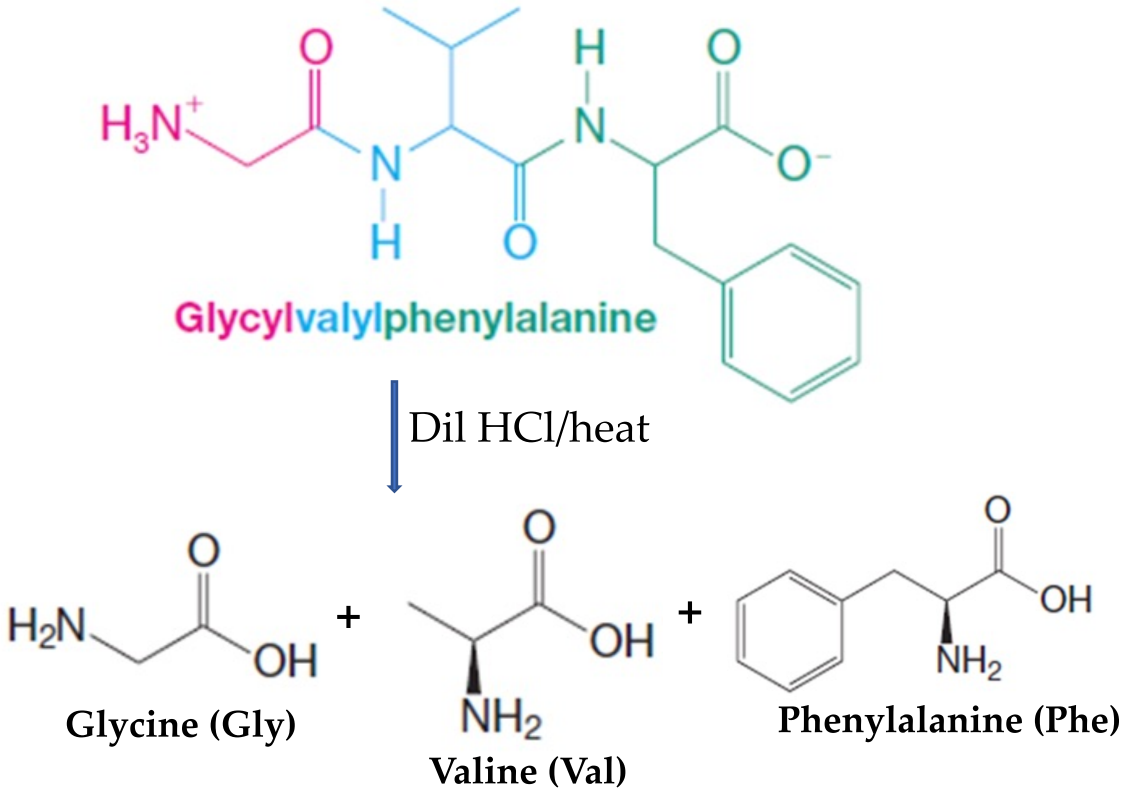


Valylglycine
(Val · Gly)

Polypeptides and Proteins

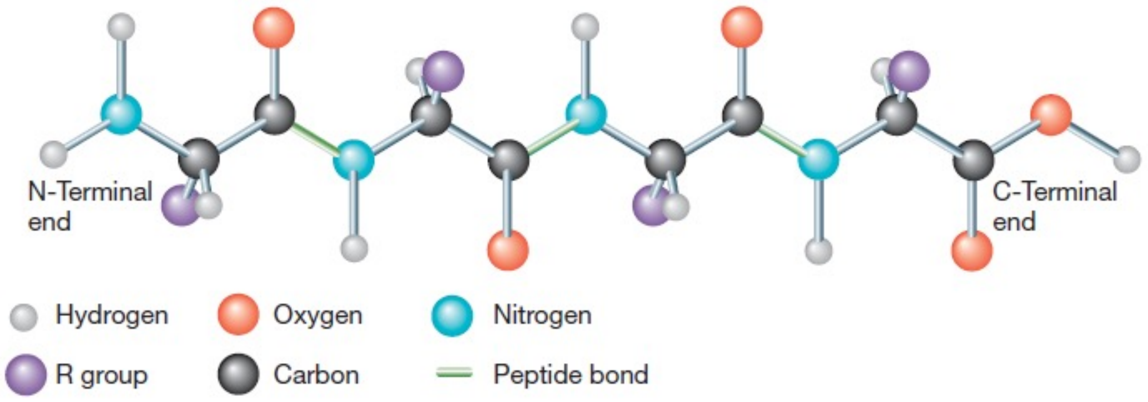
Hydrolysis of Polypeptides

- Hydrolysis of all amide linkages of a protein or polypeptide usually takes place by refluxing with 6M HCl for 24 hours. This hydrolysis results in the liberation of amino acids. The individual amino acids can be separated from each other using chromatographic methods.
- For example, the complete hydrolysis of the tripeptide Glycylvalylpheneylalanine (Gly-Val-Phe) will give the three amino acids glycine, valine, and phenylalanine.



Primary structure of polypeptides and proteins

- The sequence of amino acid residues in a polypeptide or protein is called its primary structure (the covalent structure of a polypeptide or protein). A simple peptide composed of three amino acids can have 6 different sequences. The maximum possible number of arrangements of a peptide consisting of "n amino acids" equals n! (the factorial of the number of amino acids n). For example a tetrapeptide can have as many as 24 different sequences (4! = 4x3x2x1 = 24).
- It is very important to determine the correct sequence of amino acids in the polypeptide or the protein.
- The first step in determining the sequence is to determine the terminal residues (N-terminal and C-terminal residues). For example a representation of the primary structure of a tetrapeptide is:



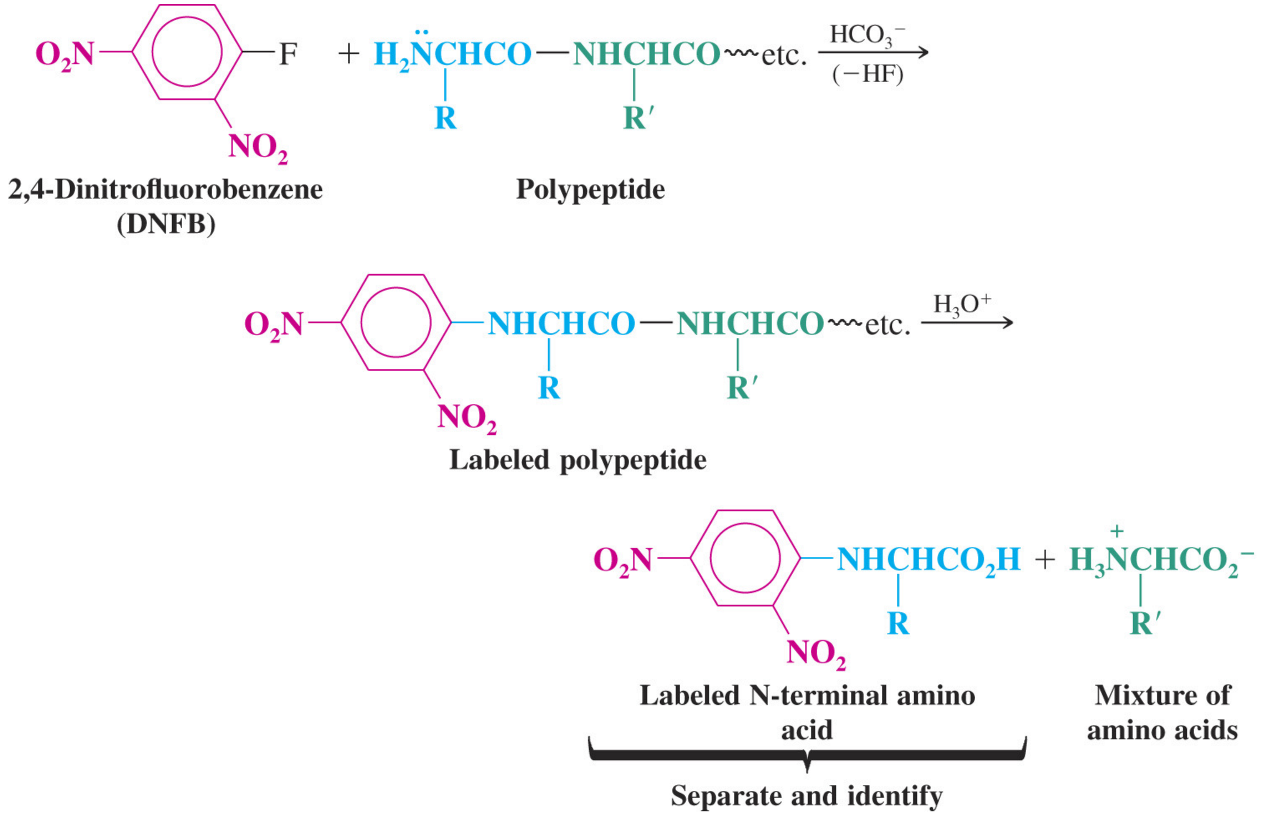
Primary structure of polypeptides and proteins

Sanger N-Terminal Analysis

- The *N*-terminal end of the polypeptide is labeled with 2,4-dinitrofluorobenzene and the polypeptide is hydrolyzed. Then, the labeled N-terminal amino acid is separated from the mixture and identified.

- The C-terminal end

Enzymes called carboxypeptidase hydrolyze C-terminal amino acids selectively. The enzyme continues to release each newly exposed C-terminal amino acid as the peptide is hydrolyzed.



Complete Sequence Analysis

- The complete sequence of polypeptides could be done by partial hydrolysis. In this approach, the sample is subjected to partial hydrolysis by dilute acid to give a random assortment of shorter polypeptides which are then analyzed.
- The smaller polypeptides are sequenced, and regions of overlap among them allow the entire polypeptide to be sequenced.
- For example: A pentapeptide is known to contain the following amino acids $\boxed{\text{Val}_2\text{, Leu, His, Phe}}$

To determine the correct sequence for this pentapeptide, we should assign the N-terminal residue and the C-terminal residue by Sanger and carboxypeptidase methods. This results in $\boxed{\text{Val (Val, His, Phe) Leu}}$

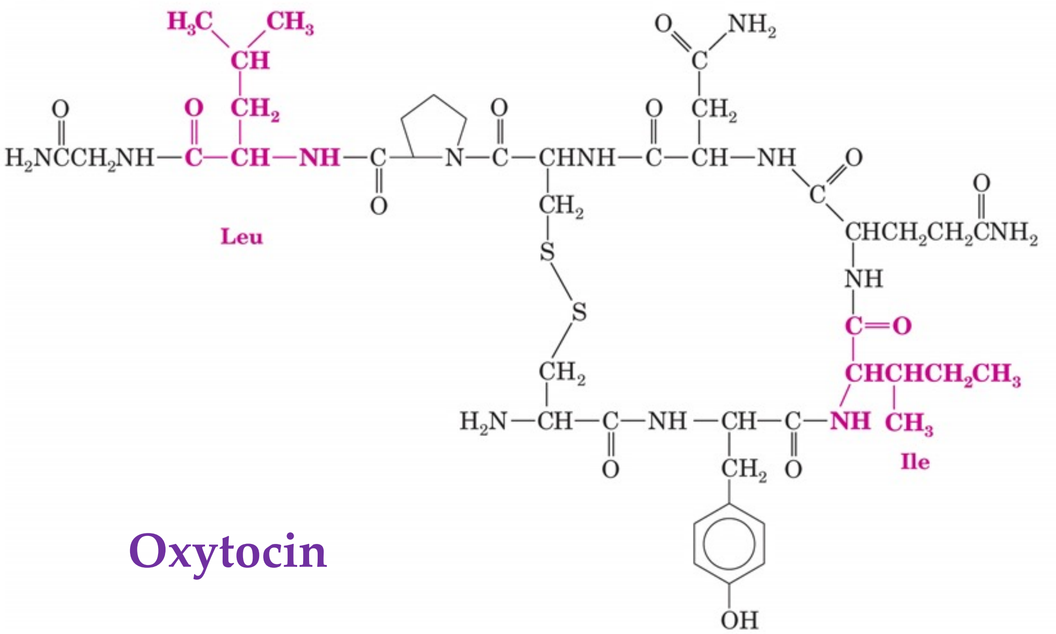
Then the pentapeptide is subjected to partial and the following dipeptides are obtained

$$\boxed{\text{Val}\cdot\text{His} + \text{His}\cdot\text{Val} + \text{Val}\cdot\text{Phe} + \text{Phe}\cdot\text{Leu}}$$

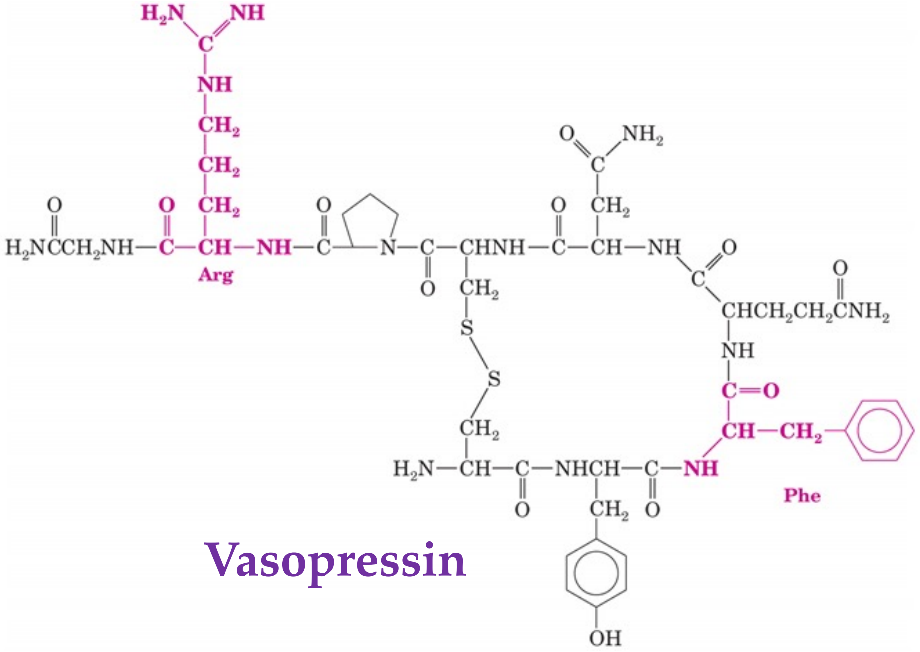
From this partial hydrolysis, the correct sequence the pentapeptide is $\boxed{\text{Val}\cdot\text{His}\cdot\text{Val}\cdot\text{Phe}\cdot\text{Leu}}$

Examples of polypeptide and protein primary structure

- Polypeptides and proteins have important biological and physiological functions in living organisms. Here is three important examples: Oxytocin, Vasopressin, and insulin.
- Oxytocin and Vasopressin are two small polypeptides strikingly similar structures.
- The two polypeptides are nonapeptides and differ in only two amino acid residues as shown with red colors.
- Oxytocin stimulates uterine contractions during childbirth.
- Vasopressin causes contraction of peripheral blood vessels and result ant increase in blood pressure.



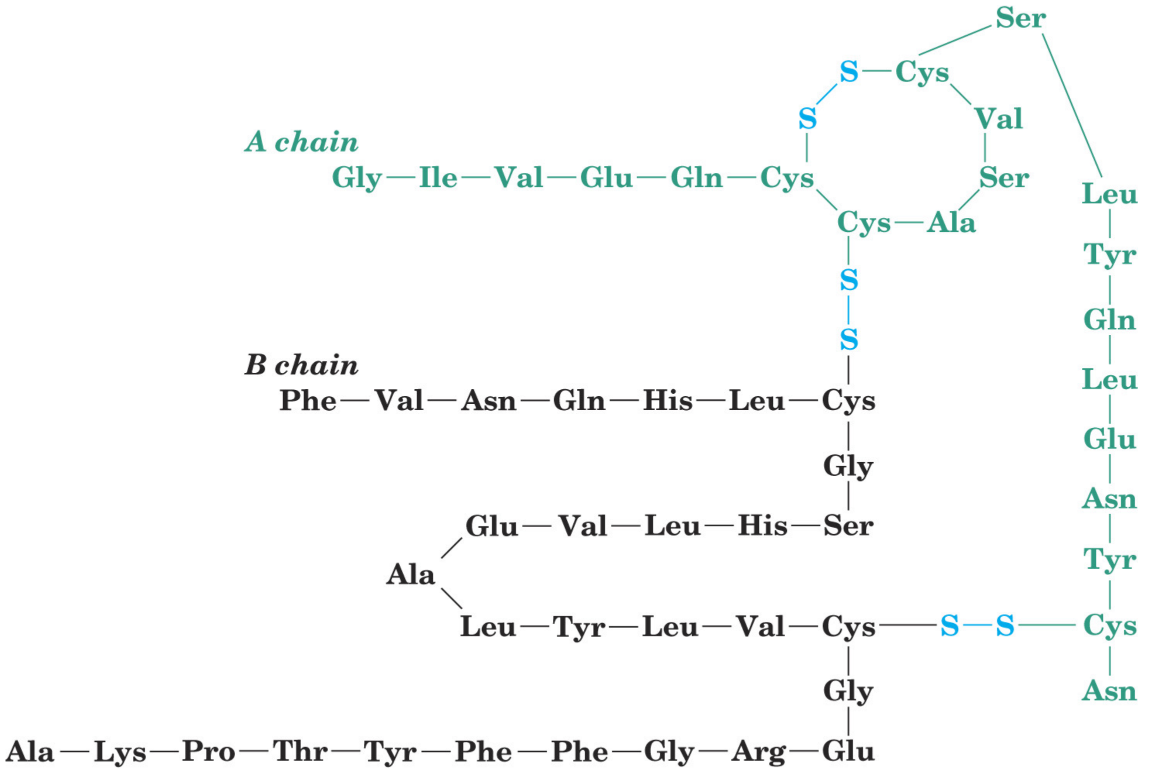
Oxytocin



Vasopressin

Examples of polypeptide and protein primary structure

- Insulin is a hormone which regulates glucose metabolism.
- The structure of bovine insulin was determined in 1953 by Sanger.
- Bovine insulin contains 51 amino acids and consists of two chains called A-chain (21amino acids) and B-chain (30 amino acids).
- The two chains are connected by disulfide linkage.
- Insulin deficiency in humans is the major cause of diabetes mellitus.
- Bovine insulin differs from human insulin at only three amino acids in its sequence.



Insulin