Fundamentals of bioorganic chemistry

БиблиографияFundamentals of bioorganic chemistry = Основы биоорганической химии [Электронный ресурс] : учебник / Zurabyan S.E. - M. : ГЭОТАР-Медиа, 2015. - http://www.studentlibrary.ru/book/ISBN9785970434437.html
АвторыZurabyan S.E.
ИздательствоГЭОТАР-Медиа
Год издания2015
ПрототипЭлектронное издание на основе: Fundamentals of bioorganic chemistry = Основы биоорганической химии: учебник. Zurabyan S.E. 2015. - 304 с.: ил. - ISBN 978-5-9704-3443-7.

PREFACE

Оглавление

Fundamentals of bioorganic chemistry 1
PREFACE 2
PART 1. GENERAL ASPECTS OF CHEMICAL STRUCTURE AND REACTIVITY OF ORGANIC COMPOUNDS. 3
Chapter 1. CHEMICAL STRUCTURE AND BONDING 3
Chapter 2. CLASSIFICATION AND NOMENCLATURE OF ORGANIC COMPOUNDS 19
Chapter 3. ELECTRONIC STRUCTURE OF ORGANIC MOLECULES 32
Chapter 4. A BRIEF SURVEY OF ORGANIC REACTIONS 48
Chapter 5. ACIDITY AND BASICITY OF ORGANIC COMPOUNDS 59
Chapter 6. HYDROCARBONS 75
PART 2. MONOFUNCTIONAL ORGANIC COMPOUNDS OF BIOLOGICAL INTERESTS. 103
Chapter 7. ORGANIC HALIDES, ALCOHOLS, PHENOLS, ETHERS, AMINES, AND ORGANOSULFUR COMPOUNDS 103
Chapter 8. CARBONYL COMPOUNDS 124
Chapter 9. CARBOXYLIC ACIDS AND THEIR DERIVATIVES 139
PART 3. POLY- AND HETEROFUNCTIONAL COMPOUNDS IN LIVING SYSTEMS. Chapter 10. STEREOISOMERISM 153
Chapter 11. POLY- AND HETEROFUNCTIONAL COMPOUNDS 169
Chapter 12. LIPIDS 180
Chapter 13. TERPENOIDS AND STEROIDS 192
PART 4. BIOPOLYMERS AND THEIR STRUCTURAL CONSTITUENTS. 203

Chapter 14. CARBOHYDRATES 203
Chapter 15. α-AMINO ACIDS, PEPTIDES AND PROTEINS 229
Chapter 16. BIOLOGICALLY ACTIVE HETEROCYCLIC COMPOUNDS 245
Appendix 2. Glossary 285
Appendix 3. Answers to problems 297
References 329

References

Аннотация

The textbook is based on modern organic chemistry and considers the structure and chemical transformations of organic compounds, especially those that have biological importance. Special attention is given to the chemical reactions that have analogies in living systems. The book contains about 250 problems on all topics and solutions for them.

This book conforms to the Federal educational program on Bioorganic Chemistry for medical schools and universities. It is meant for students who study Bioorganic Chemistry during one term. The book may also be useful for teachers and students of specialized sec ondary schools with instruction conducted in English and colleges, whose main interest is medicine, pharmacy, biology and agriculture.
ГрифМинистерство образования и науки РФ

Рекомендовано ГОУ ВПО "Первый Московский государственный медицинский университет имени И.М. Сеченова" в качестве учебника для студентов учреждений высшего профессионального образования, обучающихся по специальностям 060101.65 "Лечебное дело", 060103.65 "Педиатрия", 060105.65 "Медико- профилактическое дело", 060201.65 "Стоматология" по дисциплине "Химия общая и биоорганическая"

PREFACE

Bioorganic chemistry is a relatively new science that appeared at the junction of organic chemistry and biology in the latter half of the 20th century. Bioorganic chemistry studies biologically important natural and synthetic compounds, mainly biopolymers (peptides and proteins, nucleic acids, polysaccharides) and low-molecular bio-regulators such as vitamins, hormones, alkaloids, and other biologically active compounds including drugs.

Bioorganic chemistry deals with the following problems:

- development of isolation and purification methods for naturally occurring compounds;
- determination of the exact structure of a compound to be investigated;
- chemical synthesis of natural biologically active compounds and their analogues;
- study of interrelations between the chemical structure and biological effect of a compound;
- study of chemical interaction between a biologically active compound and components of a living cell at the molecular level.

The two latter items attracted special attention during the last decades. This may be briefly described as follows:

Bioorganic chemistry concerns essentially with the application of the tools of organic chemistry to interpretation of biological processes.

Bioorganic Chemistry is a part of chemical education of modern physicians. Biological functions of organic compounds are traditionally problems of Biochemistry and Physiology in higher medical schools. Special attention in the course of Bioorganic Chemistry is given to the study of the structure and chemical transformations of biologically important compounds, especially those that help understanding biological phenomena. Wherever possible the author tried to draw a parallel with the reactions that occur in living systems and organic reactions carried out in the laboratory. Undoubtedly, biological chemistry and classical organic chemistry conform to the same rules.

PART 1. GENERAL ASPECTS OF CHEMICAL STRUCTURE AND REACTIVITY OF ORGANIC COMPOUNDS.

Chapter 1. CHEMICAL STRUCTURE AND BONDING

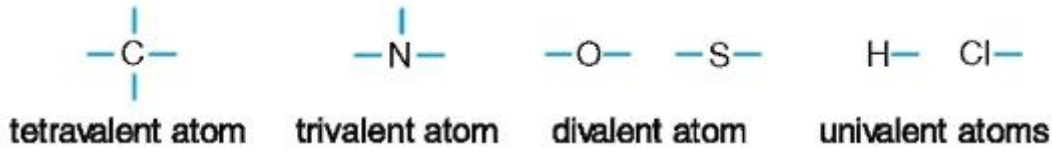
Robably many students have already acquired some knowledge in a beginner's chemistry course. Nevertheless, it would be expedient to recall in this opening chapter some information on general chemistry and the structural theory of organic chemistry that will be helpful in studying the entire course.

1.1. THE STRUCTURAL THEORY OF ORGANIC COMPOUNDS

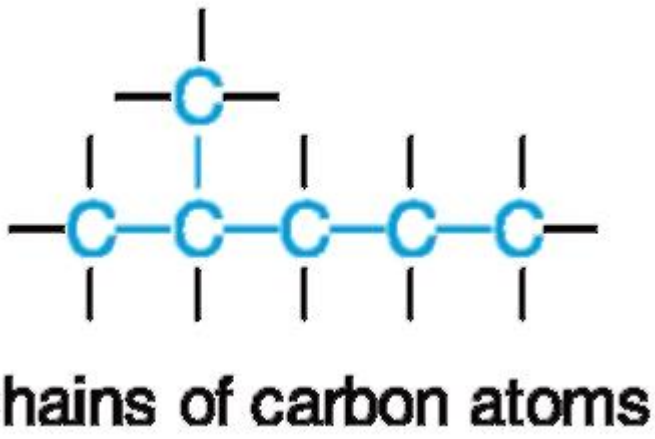
In the mid-nineteenth century (1858-1860), F.A. Kekule (in Germany) and A.S. Couper (in Scotland) independently formulated the basis of one of the most fundamental theories in modern organic chemistry, the structural theory. It comprises two main ideas.

- Atoms in organic compounds form a definite number of bonds. A measure of bonding ability of atoms is known today as *valence.* Thus, a carbon atom forms four bonds when it is linked to other atoms and is tetravalent. Nitrogen forms

usually three bonds and is trivalent, oxygen forms two bonds and is divalent, hydrogen and halogens are univalent.



- Carbon atoms can bond to each other to form extended chains of atoms linked



together.

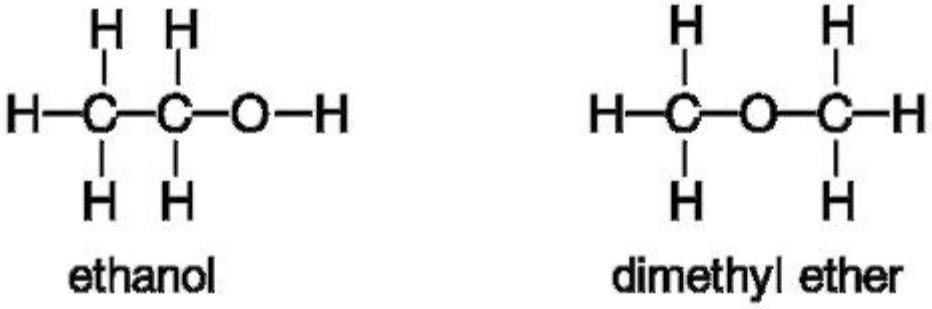
At that time, in 1861, the Russian chemist A. Butlerov made an outstanding contribution to organic chemistry when he developed new ideas in this field. The main of them stated:

The chemical nature of a complex particle is determined by the nature of its elementary constituent parts, their number and chemical structure.

Butlerov stated that all the atoms in a molecule interact with one another, and a mutual influence is the strongest between directly linked atoms. It was Butlerov who introduced the term *chemical structure* into chemistry.

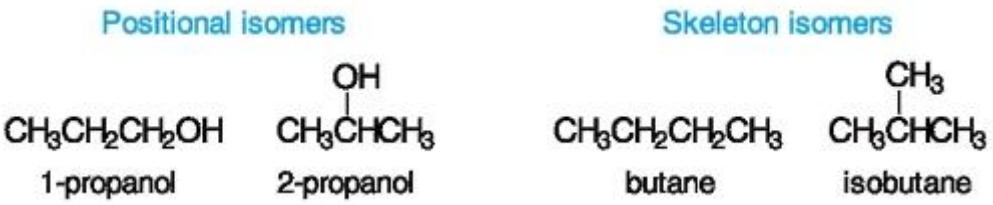
Butlerov's theory of chemical structure gave a rational explanation of isomerism.

There are many compounds that have the same molecular formula, but they differ in physical and chemical properties due to their different molecular structure. Such compounds are called *isomers* (from the Greek *isos* - equal, and *meros* part). One of the simplest examples of isomeric compounds is a pair of*ethanol* and *dimethyl ether* with the molecular formula C_2H_6O. The former is a liquid, the latter is a gas at room temperature (their boiling points are 78 and -25 °C, respectively). Ethanol reacts with sodium, dimethyl ether does not. It is quite enough to look at the structural formulas for these compounds to reveal their difference.



This is an example of constitutional isomerism when isomers differ from each other in their functional groups. Isomers can also differ in the position of the same functional group at the same carbon skeleton (positional isomers are, for example, isomeric alcohols 1-propanol and 2-propanol shown below). But the main reason for constitutional isomerism consists in the great variety of carbon skeletons. Thus, hydrocarbons C_4H_{10} represent the simplest example of skeleton isomerism.

The term *structural isomerism* is often used but it has a broader sense than constitutional isomerism. Really, structural isomers differ in *all* details of their structure including spatial arrangement of atoms (Chapter 10).



The structural theory originated by Kekule, Couper, and Butlerov was extended into three dimensions by the Dutch physico-chemist J.H. van't Hoff[1] and J.A. Le Bel (from France). In 1874, they independently proposed that four atoms to which carbon is bonded do not lie in one plane, but sit at the corners of a tetrahedron, with carbon in the centre (Fig. 1.1).

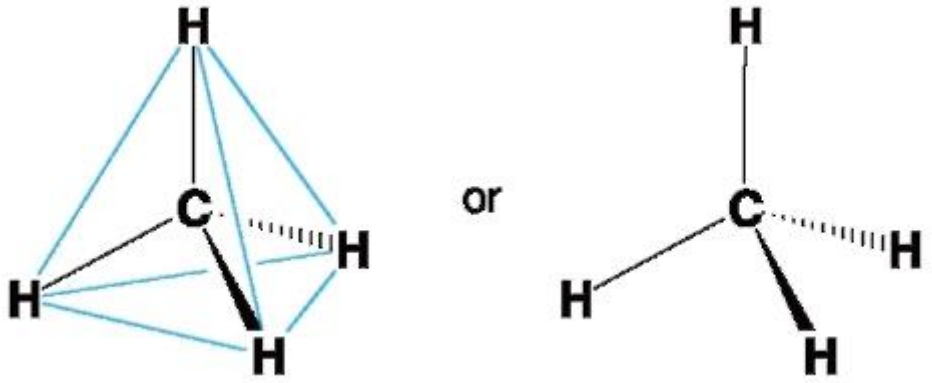


Figure 1.1. The tetrahedral structure of methane.

To show the three-dimensional arrangement of groups around the central carbon the following conventions are generally used: normal lines represent bonds in the plane of the paper, heavy wedged line represents a bond coming towards the viewer, and a dashed line represents a bond moving away from the viewer.

After the discovery of the electron (in 1897), the atomic theory penetrated into organic chemistry very intensively. Developments of quantum mechanics were also applied to organic chemistry, thus forming by the 1930's the basis of the modern understanding of the subject. These concepts are considered in greater detail in the following sections.

[1] Jacobus Hendricus van't Hoff (1852-1911), the first Nobel Prize winner in Chemistry (1901). His famous paper on stereochemistry, *The Position of Atoms in Space*, was published when he was 22 years old.

1.2. THE STRUCTURE OF ATOMS

An atom consists of a positively charged nucleus surrounded by negatively charged electrons. The nucleus, in its turn, consists of protons, carriers of the total positive charge, and neutral particles called neutrons (except the light isotope of hydrogen). Both particles are equal in mass, whereas the mass of the electron is about 1/1840 of the mass of the proton.

The number of protons determines the charge of the nucleus and identifies the atom as belonging to an individual element. In a neutral atom the positive charge of the nucleus is balanced by the negative charge of the electrons. The atomic number of an element is the number of protons (or electrons), and the mass number is the total of protons and neutrons.

Electrons are of the most interest to chemistry because their number and arrangement accounts for constitution and chemical properties of molecules. Above all, we will consider electron arrangement in carbon and hydrogen and then elements frequent in organic molecules, such as oxygen, nitrogen, sulfur, and halogens.

1.2.1. Atomic Orbitals

Electrons are located in certain regions of space called orbitals.

The orbital is a region of space where the probability of finding an electron

is large, about 95% of its time.

In other words, this is the region within which the electron density is the greatest. According to the *Pauli exclusion principle*, each orbital can contain a maximum of two electrons but they must be of opposite spin.

An electron spins about its axis and spin orientation is usually shown by arrows, either ↑ or ↓. Thus, two spin-paired electrons would be designated ↑ ↓.

Electrons occupy different shells (energy levels) designated by the numbers 1, 2, 3, and so on. The total number of orbitals in a given shell is $n^2$where n is the shell's

number. The farther the shell is from the nucleus, the greater the energy of those electrons is (Table 1.1).

Table 1.1. Distribution of electrons in the first three shells

Shell number	Energy increasing	Orbitals' number of each type			Electron capacity of the filled shell
		s	*p*	*d*	
1	⇩	1	0	0	2
2		1	3	0	8
3		1	3	5	18

Orbitals also differ in their shapes. There are four types of orbitals, specified *s, p, d,* and f. Only s and *p*orbitals will be concerned since they are the most important in organic chemistry. Taking into account both the size and the shape of orbitals, we designate them 1s, 2s, 2p, 3s, and so on. The s orbitals are spherical in shape. The three *p* orbitals are dumb-bell-shaped and oriented in such a way as to point mutually at right angles (Cartesian coordinates with the axes *x, y,* and z). They are sometimes called the $2p_x$, $2p_y$, and $2p_z$orbitals to emphasize their directional character. The shapes of the *s* and *p* orbitals are shown in Fig. 1.2.

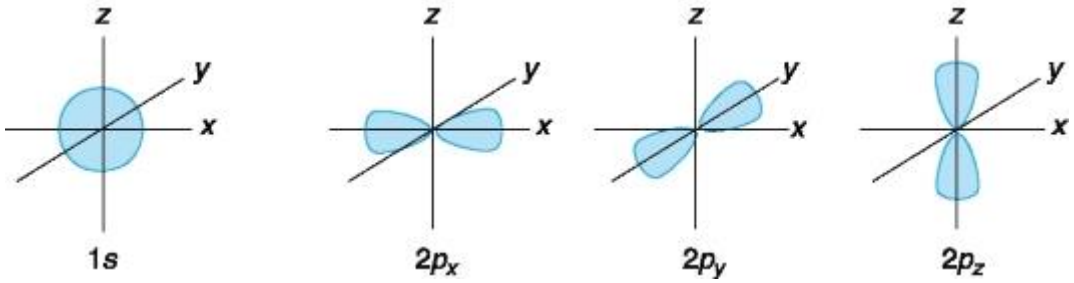


Figure 1.2. The shapes of the *s* and *p* atomic orbitals. The nucleus is at the origin of the three coordinate axes.

The 2*s* and 3*s* orbitals have the same shape as the 1 *s* orbital, but they are bigger and higher in energy. Similarly, the 3*p* orbital has the same shape as 2*p* one.

1.2.2. Electronic Configuration

A description of the orbitals that electrons occupy in an atom is called electronic configuration. The most stable state (so called ground-state) of an atom is that in which the electrons are at the lowest energy level, that is, in shells which are closest to the nucleus. To arrive at electronic configuration of any atom we should follow a few simple rules:

- Orbitals of the lowest energy are filled first. This is the *aufbau principle* (from the German *Aufbau* - building up);
- The Pauli exclusion principle mentioned above;
- The orbitals of a given shell are first filled with one electron each, and then with a second. This is known as the *Hund's rule.*

The application of these rules indicates a 1 *s* configuration of hydrogen. Electronic configuration of helium, the next element after hydrogen, will consequently be $1s^2$. Six electrons of a ground-state carbon atom are arranged to give a $1\,s^2 2s^2 2p_x 2p_y$orbital configuration (a simpler form is usually preferred, without sub-division of *p* orbitals, but it is less informative). These and other examples are shown in Table 1.2.

In representations of electronic configuration, a superscript is used to indicate the number of electrons at a particular orbital. Remember that numerals in front of the orbital symbols indicate the shell's number.

1.3. THE NATURE OF CHEMICAL BONDING

An understanding of the nature of a chemical bond is one of the major aims of chemistry. W. Kossel (a German physicist) and G.N. Lewis (an American physicochemist) developed the first, but still useful, theory of bonding in 1916. The point of the theory consists in high chemical stability of noble gases. This, in its turn, may be explained by very stable electronic configuration of the atoms that have a filled outermost shell (with two or eight electrons for helium and neon, respectively). According to this idea, other atoms tend to attain such configuration. This could be achieved either by complete transfer of electron (or electrons), called valence electrons, from one atom to another or by sharing electrons between atoms.

Table 1.2. Ground-state electronic configuration of some elements

Element	*Atomic number*	*Electronic configuration*	
		*Spin-orbital plotting**	*Written form*
Hydrogen	1	1*s* ↑	1*s*
Carbon	6	2*p* ↑ ↑ — 2*s* ↑↓ 1*s* ↑↓	$1s^2 2s^2 2p^2$
Nitrogen	7	2*p* ↑ ↑ ↑ 2*s* ↑↓ 1*s* ↑↓	$1s^2 2s^2 2p^3$
Oxygen	8	2*p* ↑↓ ↑ ↑ 2*s* ↑↓ 1*s* ↑↓	$1s^2 2s^2 2p^4$
Sulfur	16	3*p* ↑↓ ↑ ↑ 3*s* ↑↓ 2*p* ↑↓ ↑↓ ↑↓ 2*s* ↑↓ 1*s* ↑↓	$1s^2 2s^2 2p^6 3s^2 3p^4$

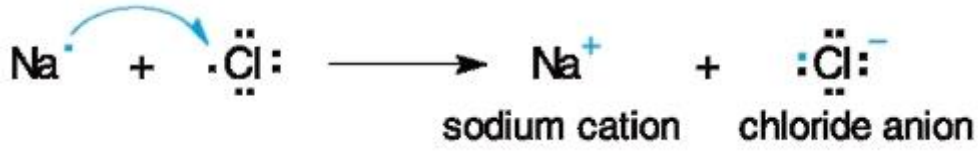
* Coloured lines represent orbitals. Empty 3d shell for the sulfur atom is not shown.

Thus, Kossel and Lewis suggested two main types of chemical bonds: the ionic and covalent bonds. Both will be reviewed.

1.3.1. Ionic Bonds

The transfer of one or more valence electrons from one atom to another results in ionic (or electrovalent) bond formation. This kind of bonding is not typical for organic compounds; nevertheless it will be briefly discussed.

The reaction between sodium and chlorine atoms to form sodium chloride is a simple example of the ionic bond formation:



The atom that gives up the electron(s) becomes positively charged, a cation. Such atoms are said to be electropositive. The atom that accepts the electron(s) is converted into an anion. Such atoms are said to be electronegative. In terms of the Lewis-Kossel theory both atoms acquire the electronic configuration of a noble gas (argon).
n general, an ionic bond is formed between two atoms that differ greatly in their electronegativity.

Electronegativity is the ability of an atom to attract valence electrons.

Electronegativity cannot be measured directly, but can be calculated from some atom's physical characteristics. Of different scales of electronegativity, the most widespread is the Pauling's[1] scale that is given in tabulated form (Table 1.3) where the principal elements of organic compounds (so-called*organogens)* are coloured.

Table 1.3. Electronegativities of some elements (after Pauling)

H 2.1							He
Li 1.0	Be 1.6	B 2.0	C 2.5*	N 3.0	O 3.5	F 4.0	Ne
Na 0.9	Mg 1.2	Al 1.5	Si 1.8	P 2.1	S 2.5	Cl 3.0	Ar
K 0.8	Ca 1.0					Br 2.8	Kr
						I 2.6	Xe

* This is the value for saturated carbon. Double-bonded and triple-bonded carbons have electronegativity values of 2.65 and 2.8, respectively.

Strictly speaking, the ionic bond may only conditionally be assigned to a bond. In a solid state, an ionic compound represents a crystal lattice with a definite arrangement of ions but it is impossible to say that any particular ion is connected to another particular ion. In solution, the ions exist separately being surrounded with molecules of a solvent.

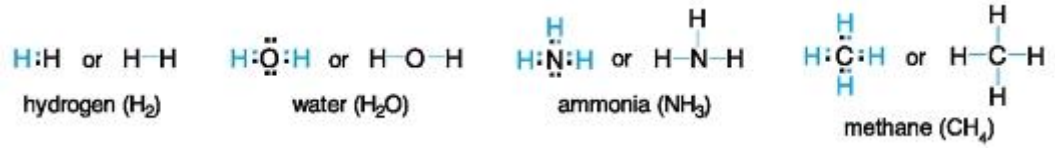
1.3.2. Covalent Bonds

A covalent bond is another kind of bonding when two atoms have the same electronegativity or they differ in electronegativity not too much. Such atoms are bound by *sharing* electron pairs instead of transferring electrons. A combination of atoms joined together by covalent bonds is called a *molecule.*

From this point of view, it is impossible to imagine a molecule of sodium chloride or other ionic compound. A crystal of NaCl represents a single giant «molecule».

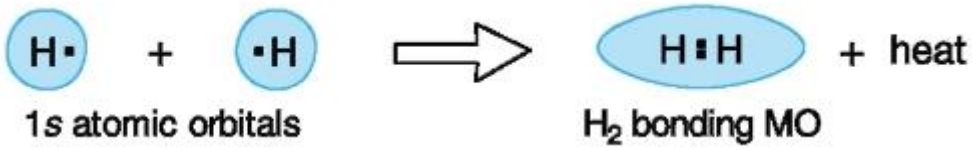
Molecular orbital theory. The nature of covalent bonds is the central concept in the study of organic chemistry. How are these bonds formed? The molecular orbital theory gives a satisfactory explanation for the question.

[1] Linus Pauling (1901-1994), an American physicist, chemist, and biochemist; Nobel Prize winner (1954 in chemistry, 1963 in World Peace).
Molecules may be represented by the electron-dot (or Lewis) structures where the valence electrons are shown as the dots. But more convenient are dash (or line-bond) structures in which a dash represents a pair of shared electrons (for more details see Section 1.5). Pairs of nonbonding valence electrons are often neglected, but it is necessary to keep them in mind. Examples are shown below:



In molecules, the atomic orbitals interact with each other by overlap to form a new set of orbitals, each of them is called a *molecular orbital* (MO). The total number of molecular orbitals always equals the number of atomic orbitals. Let us consider the covalent bond formation in the hydrogen molecule.

It is convenient to imagine that the MO is localized in the region of two nuclei. When filled with electrons, the MO represents a covalent bond and the shared electron pair is attracted to both nuclei rather than to a separate atom:



This situation is energetically favourable because this MO is lower in energy than two initial atomic orbitals. The low-energy orbital is called a *bonding* MO. The second, higher-energy MO is unfilled and no bond could result. This orbital is called an *antibonding* MO (Fig. 1.3).

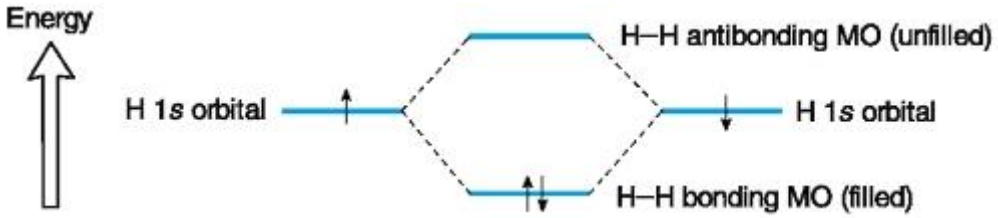


Figure 1.3. Energic diagram of the levels of the atomic and molecular orbitals for the H_2 molecule.

Problem 1.1. Draw a schematic representation of the formation of a bromine molecule from two bromine atoms.

The bond in the hydrogen molecule has maximum electron density along the line joining two nuclei (the internuclear axis) and is circularly symmetric with respect to it. Such a bond is referred to as a *sigma* (σ)*bond*. It may also be formed by similar head-on overlap of other atomic orbitals, for example, an *s* and a*p* or two *p* orbitals, as shown in Fig. 1.4.

These general considerations will now be applied to bonding in carbon compounds.

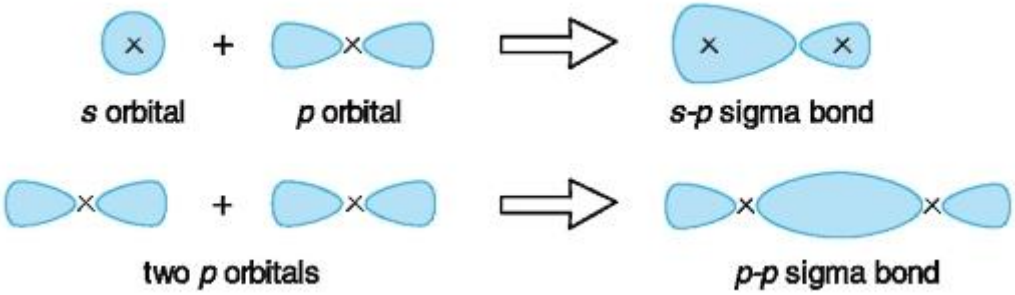


Figure 1.4. The formation of sigma bonds by head-on overlap of atomic *s* and *p* orbitals. The nuclei are shown with the sign x.

1.4. BONDING IN CARBON COMPOUNDS

The ground-state electronic configuration of the carbon atom is 1 $s^2 2s^2 2p_x 2p_y$, that means two of four electrons in its valence shell are paired (in the 2s orbital) and two are unpaired (in different 2p orbitals). We might expect the formation of only two covalent bonds between carbon and other atoms, for example hydrogens, giving rise to CH_2 molecule. Meanwhile, it is known for a long time that carbon is tetravalent and the simplest organic compound is CH_4 - methane.

1.4.1. sp^3 Hybridization

To form four bonds, carbon must adopt another electronic configuration in which all valence electrons would be unpaired. It can be achieved by the transfer of an electron from the 2*s* to the $2p_z$orbital to form an *excited-state* configuration, 1 $s^2 2s 2p_x 2p\ 2p_z$, as shown in Fig. 1.5. The transfer requires consumption of a certain amount of energy because of slightly higher energy of a *p* orbital than an sorbital. This energy loss is then compensated during formation of two additional σ bonds.

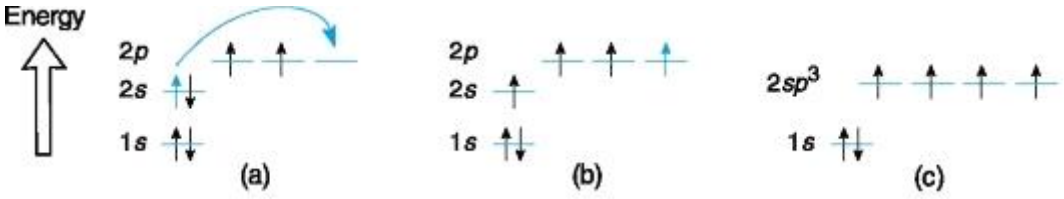


Figure 1.5. Electronic configuration of ground-state (a), excited-state (b), and sp^3-hybridized carbon (c).

Another problem requires an explanation: Why does an excited-state carbon form four *equivalent*bonds in spite of the fact that its orbitals with valence electrons (a 2*s* and three 2*p)* differ in shape and energy?

L. Pauling (in 1931) gave the answer by the concept of the *orbital hybridization.* It was suggested that an*s* orbital and three *p* orbitals are mathematically mixed, or hybridized, to form four equivalent atomic orbitals called *sp^3 hybrids:*

one 2*s* orbital + three 2*p* orbitals ⟶ four 2*sp*³ hybrid orbitals

New orbitals have the same energy: less than that of the 2*p* orbitals but greater than that of the 2*s*orbital (Fig. 1.5, *c).* The four hybrid orbitals are of the same shape resembling *p* orbitals, except that the dumb-bell is lop-sided and one of the lobes is much larger than the other (Fig. 1.6). This gives rise to better overlap with another orbital in bond formation and, as a result, sp^3 hybrid orbitals form stronger bonds than unhybridized s or *p* orbitals do.

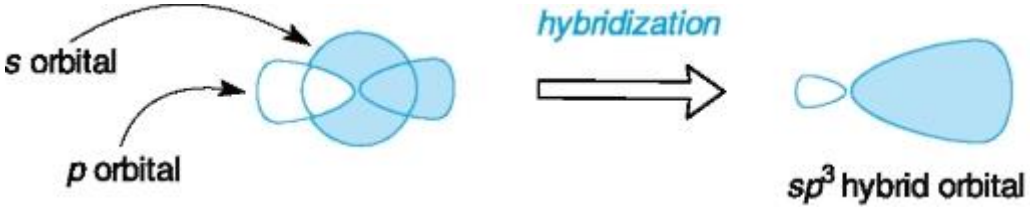


Figure 1.6. The formation of an sp^3 hybrid orbital by overlap of a *p* orbital with part of an *s* orbital.

Hybrid orbitals are located symmetrically in space at the vertices of a regular tetrahedron as shown in Fig. 1.7. All of the angles between any two bonds formed in this case are around 109.5°. This is the so-called *tetrahedral angle* that can slightly deviate (±2°) from the standard value when different atoms are attached to carbon.

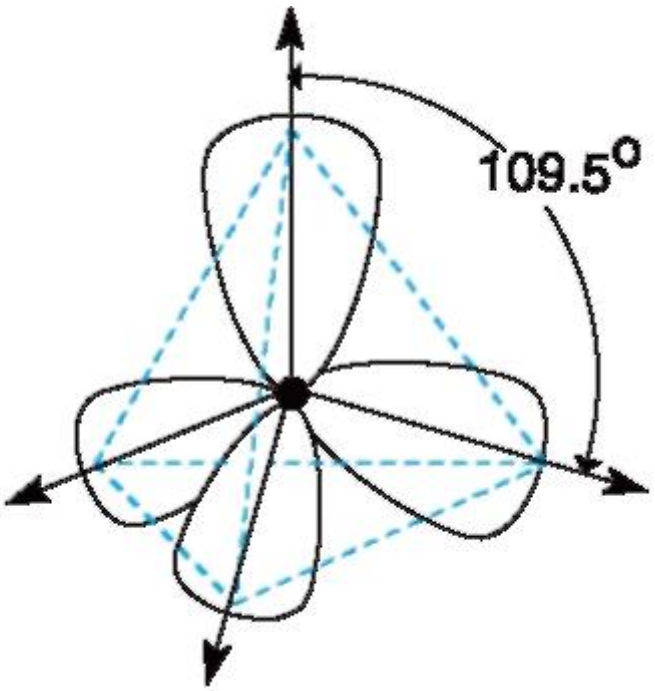


Figure 1.7. The tretrahedral configuration of an sp^3-hybridized carbon. A dot in the center represents a nucleus of the carbon atom. Small lobes of the orbitals are omitted for simplicity.

This type of hybridization is the most common electronic state of carbon, which forms σ bonds to other atoms: hydrogen, carbon, oxygen, nitrogen, and so on.

Problem 1.2. Draw a three-dimensional structure of chloroform, $CHCl_3$. What bond angles would you expect for all of the bonds in the molecule?

1.4.2. sp^2 Hybridization

There is another possibility to hybridize carbon atomic orbitals. For example, the 2s orbital and only two of the three 2p orbitals can be combined to produce three equivalent *sp^2 hybrid* orbitals:

one 2*s* orbital + two 2*p* orbitals ⟶ three 2*sp*^{2} hybrid orbitals

These hybrid orbitals lie in a plane at angles of 120° to each other. The remaining 2p orbital is perpendicular to the plane of the sp^2 orbitals, as shown in Fig. 1.8, a. The figure also shows how two such hybridized carbon atoms in ethylene molecule form not only a σ bond by sp^2-sp^2 head-on overlap, but also an additional bond by a *p-p* lateral, or sideways overlap. The latter produces a bonding orbital with maximum electron density outside the line joining the two carbon nuclei. This bond is called a *pi* (π)*bond.* A combination of the σ and π bonds results in the formation of a carboncarbon *double bond.*

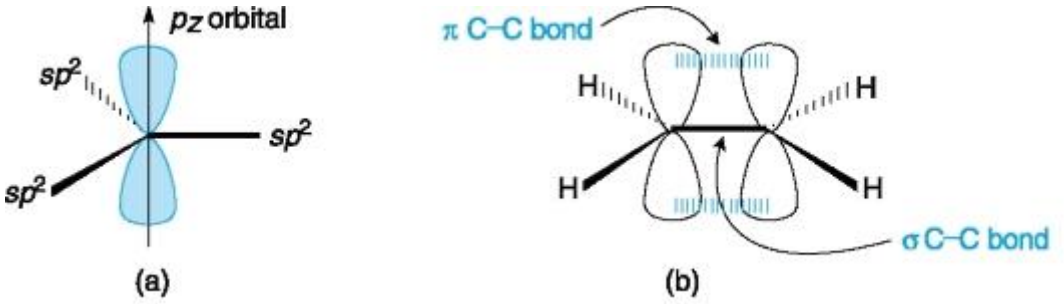


Figure 1.8. The sp^2-hybridized carbon (a) and the schematic orbital formula of ethylene (b). Hybridized orbitals are shown by their axes only.

Problem 1.3. Write a dash formula for propene, $CH_3CH=CH_2$. Indicate the type of hybridization for each carbon and predict each bond angle.

1.4.3. *sp* Hybridization

The third type of carbon hybridization consists in mixing the 2s orbital and only one 2p orbital to produce two equivalent *sp hybrid* orbitals:

one 2*s* orbital + one 2*p* orbital ⟶ two 2*sp* hybrid orbitals

The two *sp* hybrid orbitals are oriented at 180° to each other, while the remaining unhybridized two *p*orbitals are both mutually perpendicular and also perpendicular to the hybrid orbitals (Fig. 1.9, *a).*

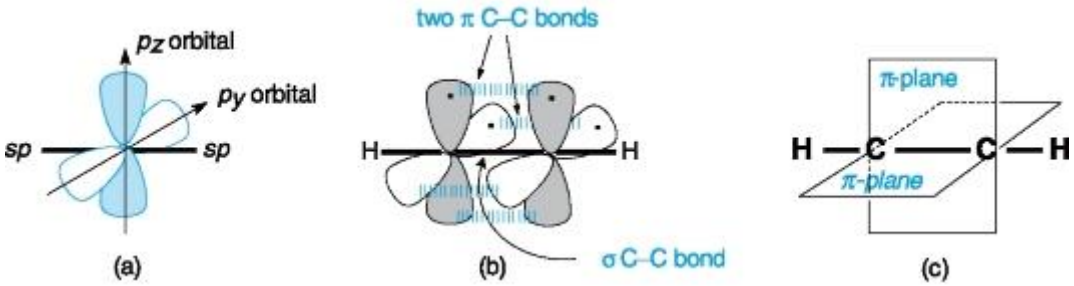


Figure 1.9. The sp-hybridized carbon (a), the schematic orbital formula of acetylene (b), and two mutually perpendicular π planes that formed two π bonds. The orbitals involved in the C-H and C-C single bonds are omitted for clarity.

In acetylene, C_2H_2, the *sp* orbitals of each carbon head-on overlap to form a σ bond. In addition, two pairs of the *p* orbitals overlap to form two π bonds, as shown in Fig. 1.9, *b* and c. Thus, the net effect is formation of a carbon-carbon *triple bond.*

1.4.4. Hybridization of Other Atoms

The description of covalent bonding can be applied to other elements, first of all nitrogen and oxygen being present very often in organic compounds. Their hybridization will be described by the simplest examples.

Nitrogen hybridization. Ground-state nitrogen atom has the electronic configuration $1s^2 2s^2 2p_x 2p_y 2p_z$. It might form three σ bonds with hydrogens, for example, using its half-filled 2p orbitals. In this case, ammonia would have all N-H

bonds at right angles which are typical for unhybridized *p* orbitals. In fact, all H-N-H bond angles in ammonia are 107° (Fig. 1.10, a), differing slightly from the tetrahedral value of 109.5°.

Such geometry of ammonia can be explained by sp^3 hybridization of nitrogen. The molecule is pyramidal, nearly tetrahedral, where *non-bonding* electrons (called also a *lone pair* of electrons) occupy one of the four sp^3 hybrid orbitals; the remaining three orbitals form three σ bonds to hydrogen atoms, as illustrated in Fig. 1.10. Another reason for hybridization lies in better overlap of a hybrid orbital compared to unhybridized one.

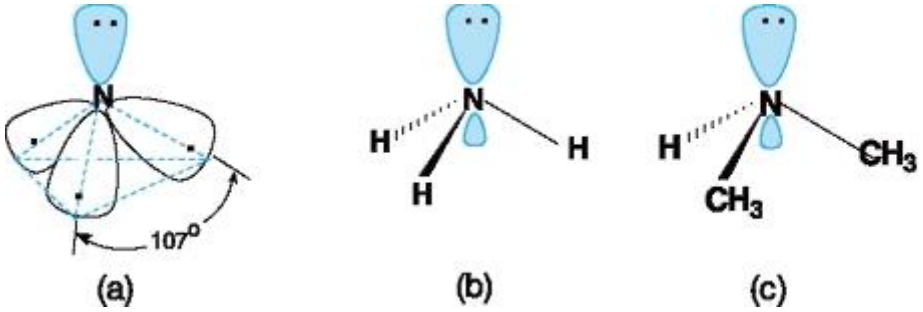


Figure 1.10. The sp^3-hybridized nitrogen (a), the pyramidal structures of ammonia (b) and seconddary amine (c). Hybrid orbitals filled with a lone pair of electrons are shown in colour.

Nitrogen hybridization and bonding in amines, from the primary, RNH_2, to the tertiary ones, R_3N, are quite similar to those of ammonia, except that one, two, or three sp^3-hybridized carbons of the R groups form σ bonds to nitrogen (Fig. 1.10, *c).*

Oxygen hybridization. An oxygen atom has the ground-state electronic configuration $1s^2 2s^2 2p_x^2 2p_y 2p_z$ and is therefore divalent. For the same reasons described for nitrogen in ammonia, oxygen in water and some organic compounds, such as alcohols, ROH, or ethers, ROR', is also sp^3-hybridized (Fig. 1.11).

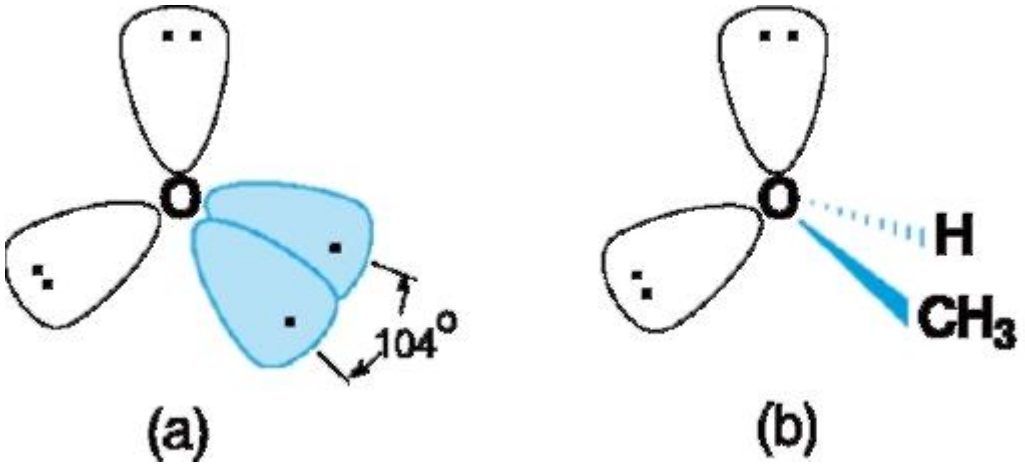


Figure 1.11. The sp^3-hybridized oxygen (a) and the structure of methanol (b). The bonding orbitals and σ bonds are shown in colour.

Important classes of organic compounds, such as aldehydes, ketones, and carboxylic acids, comprise ansp^2-hybridized double-bonded oxygen atom. It will be discussed in appropriate chapters.

Problem 1.4. What geometry would you predict for the compounds: (a) trimethylamine, $(CH_3)_3N$; (b) dimethyl ether, CH_3OCH_3? Draw three-dimensional structures of them.

1.5. THE REPRESENTATION OF STRUCTURAL FORMULAS

A molecular formula gives, with a few exceptions, no information on a structure of an organic compound. Even if a molecular formula corresponds to only one structure, organic chemists prefer structural or condensed structural formulas to show the functional group present. Thus, you never find the formula of methanol as CH_4O, but CH_3OH, as well as formic acid is usually written as HCOOH rather than CH_2O_2.

Structural formulas are, of course, the most informative ones. They show all atoms and their arrangement in a molecule. There are several types representing the structure. In the *dash formula*covalent bonds are depicted by dashes (Fig. 1.12, *a).* Two dots are sometimes used in the *dot formula*replacing the dash for representation of valence electrons. But for complicated structures, drawing each bond and atom is too time-consuming, therefore simplified formulas are often used.

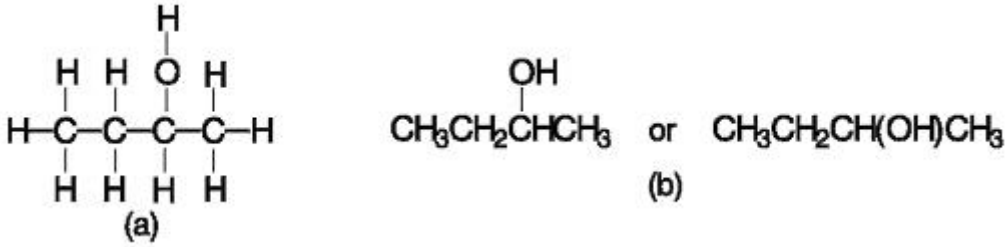


Figure 1.12. The dash formula (a) and the condensed formulas (b) for 2-butanol.

In the *condensed formula,* C-H and C-C bonds (but only single bonds) are not shown, and atoms attached to a particular carbon are written immediately after that carbon (Fig. 1.12, *b).* Such formulas are widely used in organic chemistry.

Problem 1.5. Write condensed formulas for all the structural isomers of $C_4H_{10}Br$.

The *skeletal formulas* are the easiest to write because they show only the molecular framework. In this case, carbon atoms and hydrogen atoms attached to them are omitted but other atoms (O, N, Cl, etc.) must be written. Each crossing and ending of the lines is understood to have a carbon. The number of hydrogens at each carbon can be easily calculated taking into account a valence of carbon that is always equal to four. This type of representation is most helpful in cyclic structures. Examples are given in Fig. 1.13. Of course, it is not forbidden to use a mixed type of representation as shown by the example of cumene in the figure.

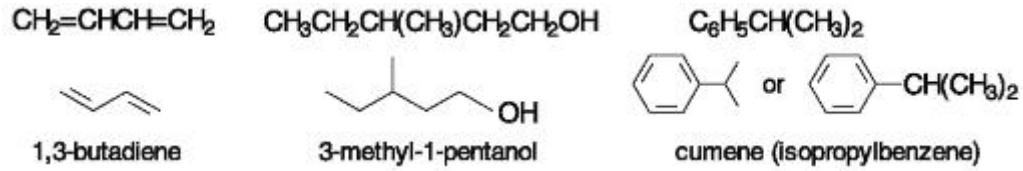


Figure 1.13. Different types of a structure representation: the condensed formulas (the upper row) and the skeletal formulas (the lower row).

Problem 1.6. Convert the following molecular formulas into any structural formula (more than one structure is possible for some compounds): (a) CH5N; (b) $C3H_6$; (c) C3H7Cl; (d) C4H10; (e) C3H8O.

1.6. SHAPE OF MOLECULES AND MOLECULAR MODELS

Organic molecules are generally not planar but three-dimensional objects. This results from tetrahedral configuration of saturated carbon atoms, that is, ones bonded to four other atoms (Fig. 1.7). Molecular shape often plays an important role in many chemical reactions; this role is crucial in biochemical processes.

Two parameters characterize the shape of a molecule, namely bond angles and bond lengths. The former have been discussed in Sec. 1.4. The bond length represents the distance between nuclei of atoms bonded. It depends on the atoms that form the bond and a type of bonding, but for the same type of bonding between the same atoms it remains nearly constant. For example, a C-H bond and a single C-C bond is 109 and 154 pm in length, respectively. Bond lengths for some covalent bonds are given in Appendix 1.

In the SI units, bond lengths are expressed in picometres (pm) or in nanometres (nm), i. e. 10^{-12} or 10^{-9}m, respectively. But most chemists (and textbooks) still use the angstrom unit (A), which is 10^{-10} m. The conversion gives 1 A = 100 pm = 0.1 nm.

Taking into account bond angles and bond lengths, it is possible to build threedimensional models of organic molecules, or *molecular models.* Several kinds of models are commercially available; three of them are essentially used (Fig. 1.14).

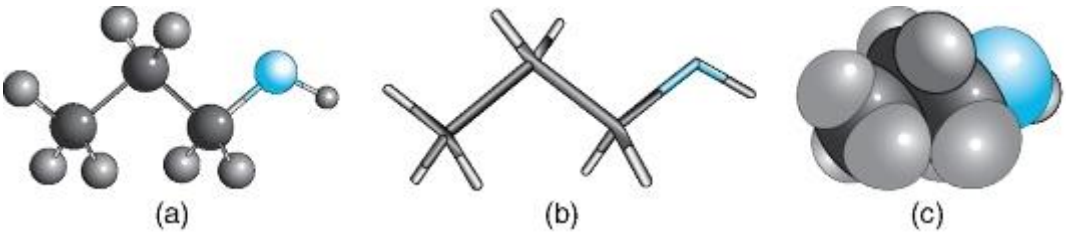


Figure 1.14. Three types of molecular models of 1-propanol, $CH_3CH_2CH_2OH$: ball-and-stick (a), stick (b), and space-filling (c). The oxygen atom is shown in colour.

Ball-and-stick models are visually effective for students' use. Stick, or skeletal models (called also the Dreiding models) show precise bond angles and lengths as well as interatomic distances in a molecule. Space-filling models, as it follows

from the name, demonstrate the space occupied by atoms and can thus show interaction and crowding of atoms in a molecule.

If you cannot find a set of models, you can make «ball-and-stick models» with your own hand from readily available materials. Use several matches for bonds and plasticine or any pliable material for atoms.

Additional Problems

1.7. Give the ground-state electronic configuration for the following: (a) Cl atom; (b) Na atom; (c) cation Na^+; (d) hydride ion, H^-.

1.8. Identity all the bonds in sodium ethoxide, CH_3CH_2ONa, that contains both ionic and covalent bonds.

1.9. Indicate a type of bonding (covalent or ionic) in the following compounds:

(a) C_2H_6;

(b) CH_3Cl;

(c) CH_3OH;

(d) $CaCl_2$;

(e) Br_2;

(f) HCl.

1.10. Indicate the type of hybridization for each carbon in the following compounds:

(a) CH2=CH2; (d) CH3CH2OH;

(b) (CH3)2CHCH(CH3)2; (e) CH2=CHCH=O;

(c) $CH_2=CHCH=CH_2$; (f) $CH_3CH_2C=N$.

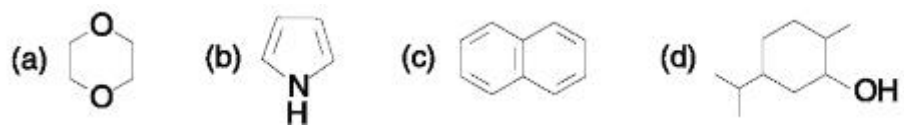
1.11. Select compound(s) in which *all* carbon atoms are sp^2-hybridized:

(a) $CH_3CH=CHCH=CH_2$; (d) $CH_2=CHC=CH$;

(b) $CH_2=CHCl$; (e) $CH_2=CHCH=O$;

(c) $CH_2=C(CH_3)-CH=CH_2$; (f) $CH_2=CCl-CH=CH_2$.

1.12. Convert the following skeletal formulas into molecular formulas:



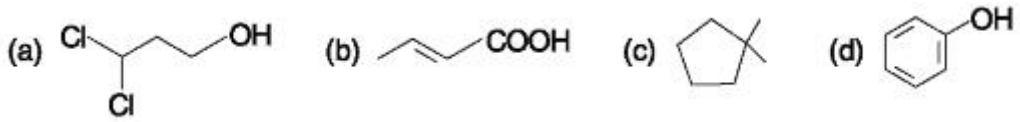
1.13. Rewrite each of the condensed formulas given below, as dash formulas and as skeletal formulas:

(a) CH3CH2CH(CH3)CH(CH3)2; (d) H2N(CH2)3NH2;

(b) $CH_3CHBrCHBrCH_3$; (e) $CH_2=C(CH_3)CH=CH_2$;

(c) $CH_3CH(OH)CH_2CH_3$; (f) $C_6H_5OCH_2CH_3$.

1.14. Write condensed formulas for each of the following skeletal formulas:



1.15. Convert the following molecular formulas into *cyclic* skeletal formulas (there is more than one possibility in each case): (a) C_4H_8; (b) C_3H_6O; (c) C_4H_7Cl.

Chapter 2. CLASSIFICATION AND NOMENCLATURE OF ORGANIC COMPOUNDS

Classification is important for all branches of science, especially for chemical sciences. Grouping of more than ten million known organic compounds into several dozen general families is exclusively helpful for studying organic chemistry. In addition, due to a big variety of organic compounds it is essential to be able to assign a definite name to each compound.

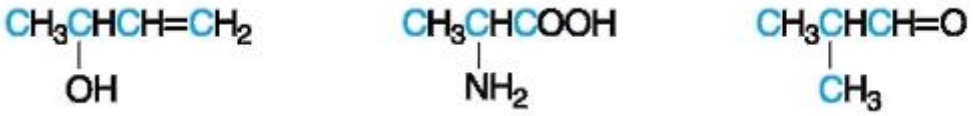
2.1. CLASSIFICATION

Organic compounds are classified according to the following features:

- a structure of molecular framework (sometimes called a molecular skeleton);
- the presence of functional groups in a molecule.

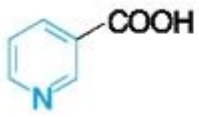
2.1.1. Classification According to the Molecular Framework

Organic compounds are subdivided into the following groups. Acyclic compounds. They have unbranched or branched carbon chain, but no rings. In the examples below, the first two represent compounds with unbranched carbon chain, whereas the third one is a compound with a branched chain:

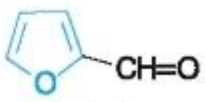


Carbocyclic compounds. They contain a ring (or rings) of carbon atoms only. The ring may contain multiple bonds and may have side carbon chains.

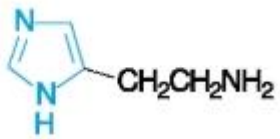
Heterocyclic compounds. They contain a cyclic skeleton having at least one heteroatom, an atom that is not carbon. The most common heteroatoms are nitrogen, oxygen, or sulfur. More than one heteroatom may be present and these atoms may be identical or different. The structures of some natural heterocyclic compounds are presented below:



nicotinic acid
(vitamin PP)



furfural
(responsible for the odour
of rye-bread)



histamine
(biologically important amine)

2.1.2. Classification According to Functional Groups

Hydrocarbons are parent compounds in organic chemistry, which, according to their name, consist of only carbon and hydrogen atoms. Most organic molecules involve *functional groups*, i. e. an atom or a group of atoms of non-hydrocarbon origin that determine chemical properties of a compound. Indeed, chemical changes occur in most reactions at the functional group whereas the molecular framework remains unchanged. Thus, the knowledge of properties of the functional groups will greatly help in the study of organic chemistry.
Organic compounds are divided into classes depending on the functional groups present. Some of the main functional groups and classes are listed in Table 2.1.

Table 2.1. Some of the functional groups and the corresponding classes of organic compounds

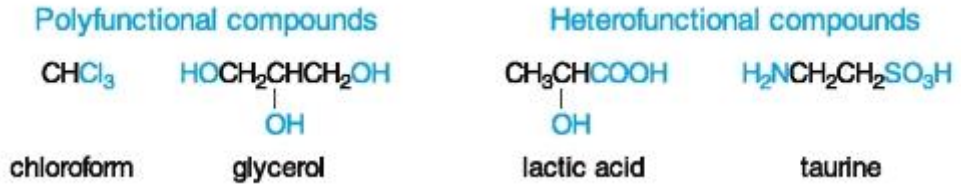
Functional group		*Name of the class*	*General formula of the class**
formula	*name*		
–	–	Hydrocarbons**	R–H
–F, –Cl, –Br, –I (–Hal)	Fluorine, chlorine, etc. (halogens)	Halogen compounds	R–Hal
–OH	Hydroxyl	Alcohols, phenols	R–OH Ar–OH
–O–	Oxy	Ethers	R–O–R′
–SH	Mercapto	Thiols	R–SH
$-NH_2$	Amino	Amines***	$R-NH_2$
$-N\begin{matrix}O\\O\end{matrix}$	Nitro	Nitro compounds	$R-NO_2$
$\begin{matrix}\\C=O\end{matrix}$	Carbonyl	Aldehydes, ketones	R–CH=O R–C(O)–R′
$-C\begin{matrix}O\\OH\end{matrix}$	Carboxyl	Carboxylic acids	R–COOH
$-S(=O)_2-OH$	Sulfo	Sulfonic acids	$R-SO_3H$

* The symbol R is usually used for any hydrocarbon radical, the symbol Ar - for an aromatic radical only.

** Multiple bonds in unsaturated compounds are sometimes related to the functional groups.

*** Only primary ones are shown.

Molecules with one functional group belong to *monofunctional* compounds. *Polyfunctional* compounds contain several identical functional groups, for example, chloroform and glycerol. Molecules with different functional groups are considered as *heterofunctional* compounds, they may be related to several classes. For example, lactic acid is both an alcohol and a carboxylic acid. Similarly, taurine belongs both to amines and sulfonic acids.



Classification characteristics form a foundation of the systematic chemical nomenclature of organic compounds.

2.2. NOMENCLATURE

At the earliest stage of organic chemistry, each new compound was usually named on the basis of its source (caffeine - from coffee-beans, urea - from urine) or its evident properties (glycerol and glucose - from the Greek *glykys*, sweet). Such names are known as trivial or common names. Trade names are widely used in pharmacy and medicine indicating some pharmaceutical effect (anesthesin, sarcolysin). Trivial and trade names are very convenient because of their brevity, but they give no information about the structure of a compound and cannot be systematized. Some trivial names went out of use with time; others have shown their viability and are used now in the systematic nomenclature.
Systematic nomenclature is an arrangement of terms that describes

complete structure of organic molecules.

The first systematic nomenclature appeared as far back as 1892 (Geneva Rules). It was then perfected by a commission of the International Union of Pure and Applied Chemistry (IUPAC) and is known now as the IUPAC rules or the IUPAC nomenclature.

2.2.1. General Principles of the IUPAC Nomenclature

To minimize confusion the following terms are used in the present rules.

Parent name: a part of the name used for the formation of a particular name according to the appointed rules. For example, the name *ethanol* is derived from ethane. The parent name may be both systematic(hexane from which *hexanal* is derived) and trivial (benzene and *nitrobenzene* from it).

Characteristic group: this term is practically equal to the term functional group, for example, the amino group -NH_2, the carbonyl group >C=O, the oxo group =O, the carboxyl group -COOH.

Principal (senior) group: the characteristic group chosen for expression as a suffix in a particular name. This group has no other advantages over remainder groups.

Substituent: any atom or group replacing hydrogen of a parent compound.

Radical: a part of a molecule that remains after removal of one or more hydrogen atoms from it. For example, the radicals, such as methyl, CH_3-, and methylene, -CH_2-, are derived from methane, CH_4.

Locant: a numeral or a letter showing a position of a substituent or a multiple bond in a parent structure.

Multiplying affix: syllables di-, tri-, tetra-, etc., which are used to indicate a set of identical substituents or multiple bonds.

Nomenclature Systems. There are eight basic nomenclature systems from which the most versatile and common therefore is the *substitutive* nomenclature. The next in prevalence is the *radicofunctional*nomenclature. These two nomenclatures, especially the former, will be considered in greater detail.
Substitutive nomenclature. The particular name of a polyfunctional compound represents a complex word that consists of a root (parent name), a suffix (principal group), and prefixes (other substituents). Fig. 2.1 demonstrates this approach.

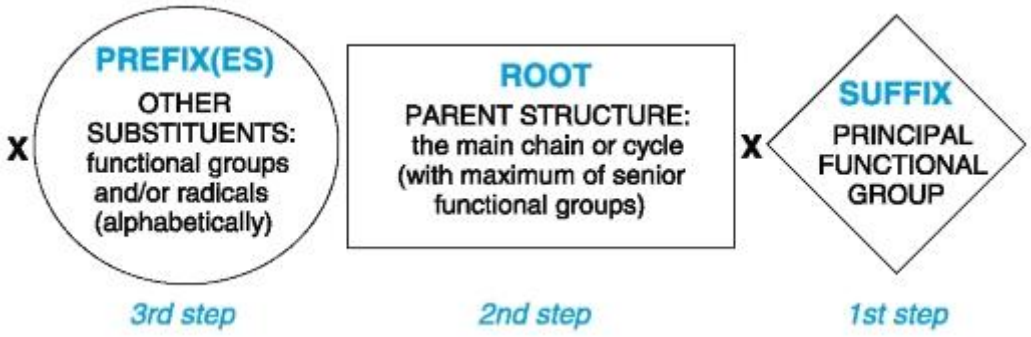


Figure 2.1. The scheme for constructing the IUPAC substitutive name. The symbol × represents multiplying affix(es).

There are two types of characteristic groups. One type is designated in a name only as prefixes. Nitro group, halogens, and some other groups belong to this type; they are listed in the lower part of Table 2.2.

Most of characteristic groups (the upper part of Table 2.2, beyond the coloured line) may be cited either as suffixes or as prefixes. But only one kind of group (principal group) is to be cited as a suffix. Within these groups, a conventional order of priority has been established (Table 2.2). It means the principal group is that which characterizes the class occurring as high as possible in Table 2.2. All other characteristic groups are then cited as prefixes. Multiplying affixes and locants are added as necessary.

Examples of using the substitutive nomenclature are given in Sec. 2.2.4.

Radicofunctional nomenclature. The principles of the radicofunctional nomenclature are identical with those of the substitutive nomenclature except that

suffixes are never used. Instead of the principal group being named as a suffix, the class name of a compound is expressed as one word and the remainder of the molecule as another.

Provided that the characteristic group is univalent (for example, an OH group of alcohols or a halogen atom of halogen derivatives) the remainder of the molecule attached to that group is expressed in its radical form as another word, which precedes the class name. When the class name refers to a characteristic group that is bivalent (for example, the fragment -Oof ethers), the two radicals attached to it are each named as separate words in alphabetic order.

Table 2.2. Suffixes and prefixes used for some important groups in the substitutive nomenclature IUPAC (in order of decreasing priority)

Class	*Characteristic group*		
	*formula**	*name*	
		prefix	*suffix*
Carboxylic acids	–COOH	—	-oic acid
	–COOH	carboxy-	-carboxylic acid
Sulfonic acids	$–SO_3H$	sulfo-	-sulfonic acid
Salts of carboxylic acids	–COOM	—	metal**…-oate
	–COOM	—	metal**...-carboxylate
Acid anhydrides	–C(O)–O–C(O)–	—	-oic anhydride
Esters	–COOR	—	R**...-oate
	–COOR	R-oxycarbonyl-	R**...-carboxylate
Acid halides (on example of chloride)	–C(O)Cl	—	-oyl chloride
	–C(O)Cl	chloroformyl-	-carbonyl chloride
Amides	$–C(O)NH_2$	—	-amide
	$–C(O)NH_2$	carbamoyl-	-carboxamide
Nitriles	–C≡N	—	-nitrile
	–C≡N	cyano-	-carbonitrile
Aldehydes	–CH=O	oxo-	-al
	–CH=O	formyl-	-carbaldehyde
Ketones		oxo-	-one
Alcohols	–OH	hydroxy-	-ol
Phenols	–OH	hydroxy-	-ol***
Thiols	–SH	mercapto-	-thiol
Amines	$–NH_2$	amino-	-amine
Imines	=NH	imino-	-imine
Ethers	–O–	…-oxy-****	—
Sulfides	–S–	…-thio-****	—
Halogen derivatives (on example of chloride)	–Cl	chloro-	—
Nitro compounds	$–NO_2$	nitro-	—

* Coloured carbon atoms are included in the name of parent structure and not in the suffix or prefix.

** Should be added in front of the name.

*** Phenols have usually common names.

**** Used only with the name of radical R, e. g. ROalkoxy- or RSalkylthio-.

This type of nomenclature is the most convenient one for such classes as ethers, sulfides, amines, and halogen compounds, especially for the compounds with simple radicals.

2.2.2. General Principles of Forming a Systematic Name

The formation of a name for a chemical compound usually involves the following steps in the order given below.

Step 1. From the nature of the compound determine the most pertinent type of nomenclature (substitutive, radicofunctional, or else).

Step 2. Determine the kind of characteristic group for use as the principal group, if any. It is this group that stipulates then the choice of a parent structure and its numbering.

Step 3[1]. Determine the parent structure (principal chain or parent ring system[2]). When in an acyclic compound there is a choice for principal chain, the following criteria are applied successively, in the order listed, until a decision is reached:

a) the maximum number of substituents of the highest priority from Table 2.2;

b) the maximum number of double and triple bonds considered together;

c) the maximum length of the chain;

d) the maximum number of substituents cited as prefixes.

Step 4. Name the parent structure and the principal group(s).

Step 5. Determine and name prefixes.

Step 6. Complete the numbering. The starting point and direction of numbering must be chosen so as to give the lowest number to the principal group. If this rule does not effect the choice, a principle of the*lowest locants* is used. It means that the locants of substituents must then be as low as possible.

Step 7. Assemble the partial names (according to steps 4 and 5) into a complete name, using the alphabetic, but not numerical, order. Multiplying affixes are not included in this order. Position of locants is not strictly determined by the IUPAC rules. In this respect, there are alternative versions in different languages and, what is more, there are some differences between American and British chemical English! But usually, locants are placed in front of prefixes and a suffix. Finally, locants are separated from each other by commas and separated from the letters by hyphens. All parts of a name are written without space, except for the word acid in the name of carboxylic acids.

[1] This step and the following one are related to the substitutive nomenclature.
[2] Seniority of ring systems is too much complicated and is not considered as well. It should be mentioned only that all heterocycles are senior to all carbocycles.

2.2.3. Names of Parent Structures

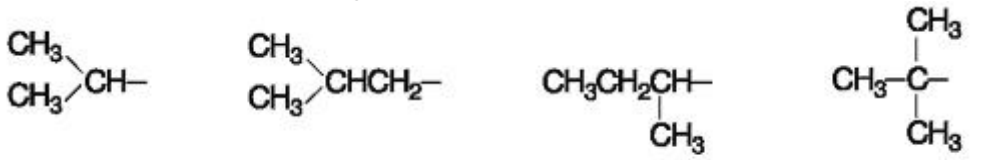
Parent structures are presented either by an open carbon chain, or carbocyclic framework, or heterocyclic one (Sec. 2.1).

Acyclic hydrocarbons. The generic name of saturated acyclic hydrocarbons (unbranched or branched) is*alkanes*. The first four saturated unbranched acyclic hydrocarbons are called methane, ethane, propane,and butane. Names of the higher members of this series consist of a numerical term (Greek mainly), followed by -ane with elision of the terminal -a from the numerical term. Examples are given in Table 2.3.

Table 2.3. Straight-chain saturated hydrocarbons and alkyl radicals

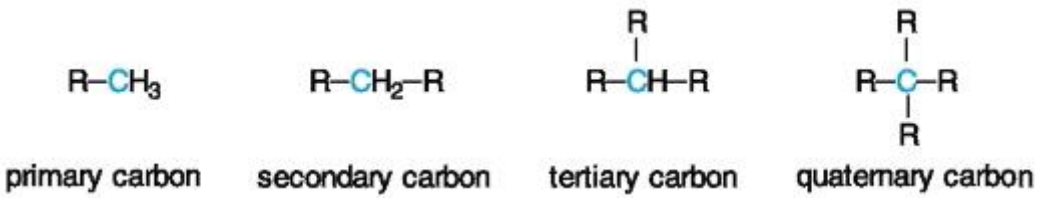
Alkanes		*Alkyl radicals*	
formula	*name*	*formula*	*name*
CH_4	Methane	CH_3–	Methyl
C_2H_6	Ethane	CH_3CH_2–	Ethyl
C_3H_8	Propane	$CH_3CH_2CH_2$–	Propyl
C_4H_{10}	Butane	$CH_3(CH_2)_3$–	Butyl
C_5H_{12}	Pentane	$CH_3(CH_2)_4$–	Pentyl
C_6H_{14}	Hexane	$CH_3(CH_2)_5$–	Hexyl
C_7H_{16}	Heptane	$CH_3(CH_2)_6$–	Heptyl
C_8H_{18}	Octane	$CH_3(CH_2)_7$–	Octyl
C_9H_{20}	Nonane	$CH_3(CH_2)_8$–	Nonyl
$C_{10}H_{22}$	Decane	$CH_3(CH_2)_9$–	Decyl
$C_{11}H_{24}$	Undecane	$CH_3(CH_2)_{10}$–	Undecyl
$C_{12}H_{26}$	Dodecane	$CH_3(CH_2)_{11}$–	Dodecyl
$C_{16}H_{34}$	Hexadecane	$CH_3(CH_2)_{15}$–	Hexadecyl
$C_{18}H_{38}$	Octadecane	$CH_3(CH_2)_{17}$–	Octadecyl
$C_{20}H_{42}$	Eicosane	$CH_3(CH_2)_{19}$–	Eicosyl

Univalent radicals derived from these hydrocarbons by removal of a hydrogen atom from a terminal carbon are named by replacing the suffix -ane in the name of the hydrocarbon by -yl (Table 2.3). The following names are used for the unsubstituted radicals only:

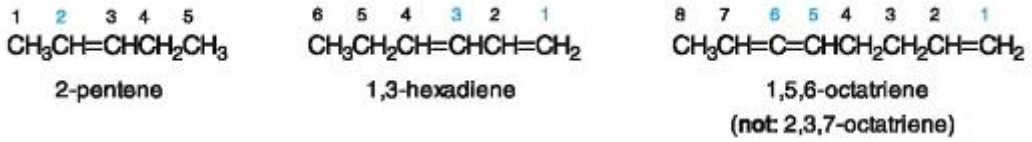


isopropyl (i-Pr) isobutyl (i-Bu) *sec*-butyl (*s*-Bu) *tert*-butyl (*t*-Bu)

The prefixes *sec* (secondary) and *tert* (tertiary) refer to the degree of alkyl substitution at a carbon atom. There are four types of carbons that differ in their alkyl environment. If the carbon atom is bonded to only one carbon, the former is referred to as a *primary* carbon. A *secondary* carbon has two other carbons bonded to it, and so on, as shown below:



Unsaturated unbranched hydrocarbons having one double bond are named by replacing the suffix -anein the name of the corresponding alkane by the suffix -ene. If there are two or more double bonds, the suffix will be -adiene, -atriene, and so on. The generic names of these hydrocarbons are *alkenes, alkadienes, alkatrienes,* etc. The chain is so numbered as to give the lowest possible numbers to the double bonds; only the lower locant is cited in the name.

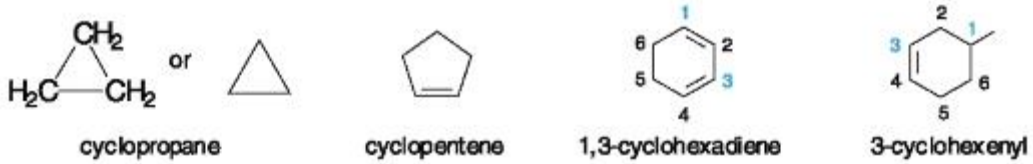


The generic name of unsaturated hydrocarbons with one triple bond is *alkynes.* Unbranched compounds of this series are named similarly to alkenes but using the suffix -yne. The position of the triple bond is indicated in the same way as for alkenes.

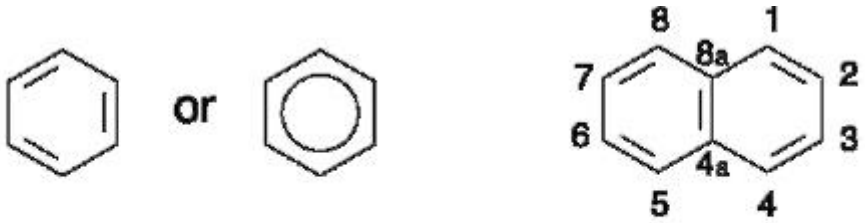
The following non-systematic names are retained: ethylene for $CH_2=CH_2$, acetylene for CH=CH, andisoprene for $CH_2=C(CH_3)-CH=CH_2$, as well as vinyl for the radical $CH_2=CH-$ and allyl for the radical $CH_2=CHCH_2-$.

It should be noted that the numbering system described above is applicable only to hydrocarbons (including a branched chain) and their derivatives with non-principal characteristic groups, such as halogens or a nitro group. The numbering of a chain may be changed in the presence of principal characteristic groups (see step 6 in Sec. 2.2.2 and the following section).

Cyclic hydrocarbons. The names of saturated monocyclic hydrocarbons (with no side chains) are formed by adding the prefix cyclo- to the name of unbranched alkane with the same number of carbon atoms. The generic name of these hydrocarbons (with or without side chains) is *cycloalkanes.*
Unsaturated monocyclic hydrocarbons are named and numbered similarly to acyclic analogues. This also concerns univalent radicals derived from cyclic hydrocarbons. The carbon with the free valence is numbered as 1 (but this locant may be omitted) regardless of a characteristic group present in the radical.



The generic name of monoand polycyclic aromatic hydrocarbons is *arenes.* The simplest representatives are called benzene and naphthalene.

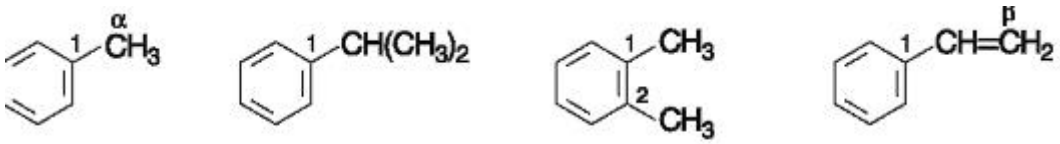


benzene
(**not** 1,3,5-cyclohexatriene)

naphthalene

The numbering in the benzene ring is applied only to derivatives having more than one substituent. The symbols *o- (ortho), m-* (meta), and *p- (para)* may be used in place of 1,2-, 1,3-, and 1,4-, respectively, when only two substituents are present.

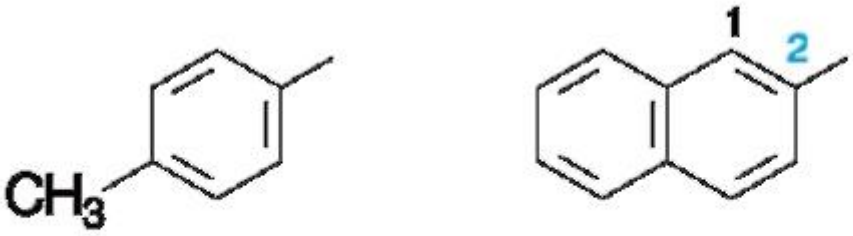
Several aromatic hydrocarbons with side chains may be used as parent structures for the names of compounds having non-principal characteristic groups. These are toluene, cumene, xylenes (three isomers), and styrene:



toluene cumene xylene (*ortho* isomer shown) styrene

α

The most widespread aromatic radicals are phenyl, C_6H_5-; benzyl, $C_6H_5CH_2$- ; tolyl *(ortho*, *meta,* and *para*isomers) derived from toluene; and naphthyl (1- and 2-isomers) - from naphthalene, for example:

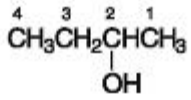


***p*-tolyl** **2-naphthyl**

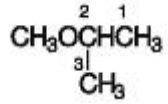
Heterocyclic compounds. Nomenclature of heterocycles will be discussed later (Chapter 16).

2.2.4. Examples of Constructing the Systematic Names

Naming of simple hydrocarbons and monofunctional compounds encounters usually few problems because the basic principles of nomenclature are logical and easy to understand. Difficulties arise with polyand heterofunctional compounds. Examples given below illustrate how the IUPAC rules (the substitutive names unless otherwise stated) are applied for a particular compound of different classes.

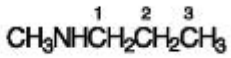
Example 1.  (1)

The name of the alcohol (1) is 2-butanol. The only carbon chain consists of four atoms (butane); the principal characteristic group (OH) is expressed by the suffix -ol. We number the chain from the right, starting closest to the OH group.
The radicofunctional name of the compound is *sec*-butyl alcohol, which is derived from the name of the radical *(sec*-butyl) and the name of a class. It should be noted that the names such as *sec*-butanol, as well as isopropanol and *tert*-butanol are incorrect because there are no hydrocarbons *sec*-butane, isopropane, and *tert*-butane to which the suffix -ol can be added.

Example 2.  (2)

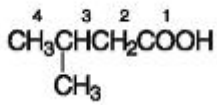
The main chain in the ether (2) is a three-carbon chain of propane. The group CH_3O- attached to C-2 represents a combined substituent methyl + oxy = methyloxy, or shortly methoxy that may be used only as a prefix. Together this gives the name 2-methoxypropane.

The radicofunctional name of the compound (2) is isopropyl methyl ether (two radicals are set in alphabetic order), which seems to be more convenient than the substitutive one.

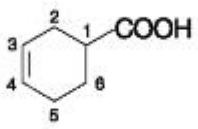
Example 3.  (3)

The name of the secondary amine (3) is N-methyl-1-propanamine. The locant *N* means that substitution with methyl group was performed at nitrogen atom but not at carbon (as is usually in the substitutive nomenclature).

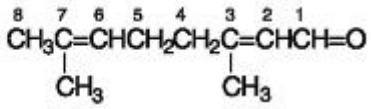
Using the radicofunctional nomenclature we get simpler name
- methylpropylamine.

Example 4.  (4)

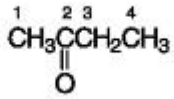
The systematic name of isovaleric acid, the compound (4), employed in pharmacy is 3-methylbutanoic acid. The suffix -oic acid is used because carbon of the carboxyl group is a member of the parent name (butane).

Example 5.  (5)

The name of the compound (5) is 3-cyclohexenecarboxylic acid *(not* 3-cyclohexenoic acid). Carbon of the carboxyl group cannot be a member of the cyclic parent structure (cyclohexene) in this example. The cycle is always numbered, starting from a carbon to which the principal characteristic group is attached.

Example 6.  (6)

The systematic name of citral, the compound (6), a component of lemon oil used in treatment of eye-diseases, is 3,7-dimethyl-2,6-octadienal. It should be mentioned here that the locant 1 is never used for suffixes -al and -oic acid.

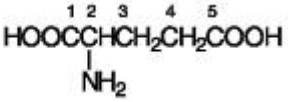
Example 7.  (7)

The compound (7) is named simply butanone. Pay attention to the absence of a locant for the oxo group. The atoms C-2 and C-3 are identical whereas the oxo group attached to C-1 gives rise not to a ketone, but an aldehyde.

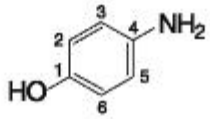
The radicofunctional name of the compound (7) is ethyl methyl ketone.

Example 8. $HOCH_2CH_2OH$ (8)

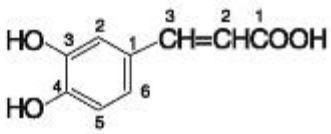
The simplest stable glycol is 1,2-ethanediol, the compound (8), known as an antifreeze component. The multiplying affix di- used in addition to the suffix -ol denotes that there are two hydroxyl groups. Locants for them must be used here to reject an isomeric structure of 1,1-diol $CH_3CH(OH)_2$. The latter does not exist in a free state but nomenclature rules ignore the problems of stability. This means that any structure written correctly, i. e. with appropriate valences of all atoms, can be named.

Example 9.  (9)

Glutamic acid, the compound (9), is one of natural (protein) amino acids. This trivial name is adopted by the IUPAC rules as the parent name. Systematically it may be named 2-aminopentanedioic acid. The multiplying affix di- is also used together with the suffix -oic acid as in Example 2.8.

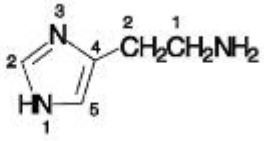
Example 10.  (10)

The heterofunctional compound (10) might be named as either a substituted phenol, C_6H_5OH, or a substituted aniline, $C_6H_5NH_2$. The former should be taken as the parent name due to priority of phenols (Table 2.2). Thus, the correct name is 4-aminophenol (or p-aminophenol).

Example 11.  (11)

The principal characteristic group in caffeic acid, the compound (11), is the carboxyl group, therefore the parent name (propene) is assigned to a chain of three carbons. A substituent at C-3 representsphenyl radical with two own substituents

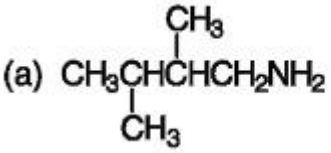
(hydroxyl groups). Together with locants, this leads to the full name: 3-(3,4-dihydroxyphenyl)propenoic acid.

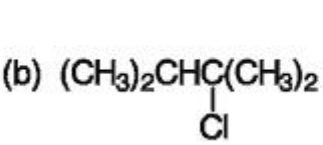
Example 12.  (12)

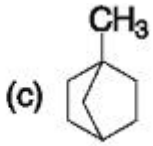
The principal characteristic group in biogenous amine histamine, the compound (12), is situated in a side chain. The heterocyclic substituent (imidazolyl radical) retains its numbering. Thus, the systematic name of histamine is 2-(4-imidazolyl)ethanamine.

Additional Problems

1.1. Identify the kind of each carbon (primary, secondary, and so on) in these molecules:







2.2. Write a structural formula for each of the following compounds:

(a) 2,2,3-trimethylbutane;

(b) 2-pentanol;

(c) 2-methyl-3-pentanamine;

(d) 3-methylbutanoic acid;

(e) 3-ethylbutanal;

(f) 4,5,5-trimethyl-2-pentene.

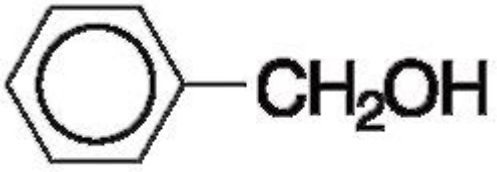
2.3. Write a structural formula for compounds from (a) to (d) and underline a parent structure in all the compounds:

(a) (CH3)2CHCH2NH2;

(b) $CH_3CH_2CH_2OCH{=}CH_2$;

(c) $CH_3CH_2CH(COOH)_2$;

(d) $CH_3CH(OH)CH(OH)CH_3$;

(e) 

.

Give the proper IUPAC name for all the compounds.

2.4. Write the condensed structural formula for each of the following compounds and name it by the IUPAC substitutive nomenclature:

(a) $(CH_3)_2CHC(CH_3)_3$;

(b) $CH_3CH(OH)CH_2CH_2CH_3$;

(c) $CH_3CH_2CH(NH_2)CH(CH_3)_2$;

(d) $(CH_3)_2CHCH_2COOH$;

(e) $CH_3CH_2CH(C_2H_5)CH{=}O$;

(f) $CH_3CH{=}CHCHClCHCl_2$.

2.5. The following names are *incorrect:*

(a) 1,1-dimethylpropane;

(b) fluorotribromomethane;

(c) 1-chloro-2-hydroxycyclohexane;

(d) 1-amino-4-carboxybenzene;

(e) 2-hydroxy-1-propanamine;

(f) *m*-aminotoluene.

Draw the structures they represent, explain why the names are incorrect, and give the correct substitutive IUPAC names.

2.6. Draw structural formulas for all the compounds containing the benzene ring and having molecular formula C_7H_8O. Give the substitutive IUPAC names for them.

Give the radicofunctional names where it is possible.

2.7. Propose suitable structures for the following:

(a) a five-carbon cycloalkane;

(b) a ketone with five carbons;

(c) a four-carbon carboxylic acid;

(d) an aromatic chlorocompound with seven carbons;

(e) an amino alcohol.

Write a structural formula and give the substitutive IUPAC name for each compound proposed.

2.8. Give an example of each of the following:

(a) an eight-carbon aromatic hydrocarbon;

(b) an aldehyde with four carbons;

(c) a three-carbon primary amine;

(d) a simplest diol;

(e) a hydroxy carboxylic acid.

Write structural formulas and substitutive IUPAC names for each example.

2.9. Draw structural formulas for all the compounds that fit these descriptions:

(a) hydrocarbons with formula C_4H_8;

(b) acyclic ethers with formula C_4H_8O;

(c) ketones with formula $C_5H_{10}O$;

(d) carboxylic acids with formula $C_4H_6O_2$.

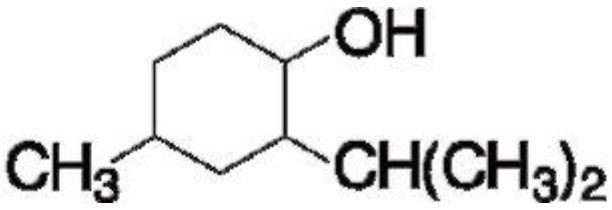
For one of the representatives from (a) to (d) give the proper IUPAC name.

2.10. Give the IUPAC substitutive names for the following compounds (the trivial or non-systematic names are given in parentheses):

(a) $CH_3CH=CHCH=CHCOOH$ (sorbic acid);

(b) $H_2N(CH_2)_4CH(NH_2)COOH$ (lysine);

(c) $HOCH_2CH(OH)CH=O$ (glyceraldehyde);

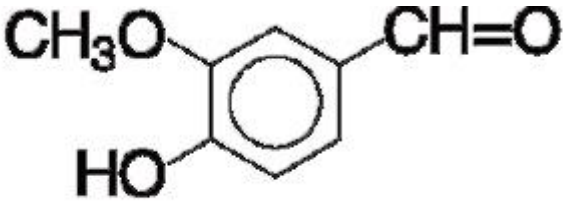
(d)  (menthol).

2.11. Name parent structures of the following compounds and name the compounds, using the IUPAC substitutive nomenclature (the trivial names are given in parentheses):

(a) $H_2NCH_2CH_2SO_3H$ (taurine);

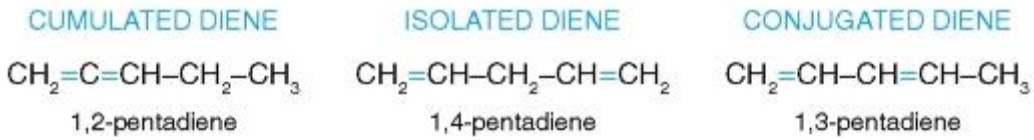
(b) $CH_3C(O)CH_2COOH$ (acetoacetic acid);

(c) $HSCH_2CH(NH_2)COOH$ (cysteine);

(d)  (vanillin).

Chapter 3. ELECTRONIC STRUCTURE OF ORGANIC MOLECULES

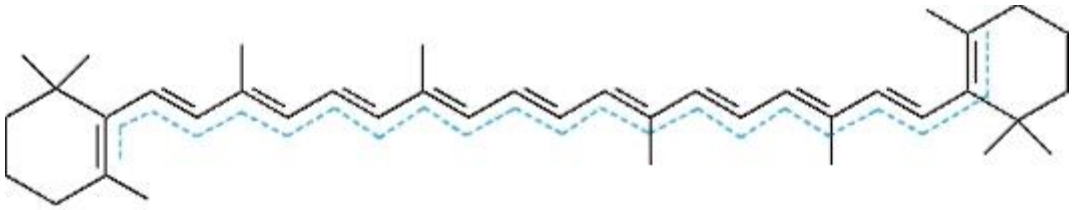
There are many organic compounds whose molecules contain more than one multiple (double or triple) bonds. According to mutual arrangement, the multiple bonds of polyunsaturated compounds are classified as being *cumulated*, *isolated*, and *conjugated*. These will be considered by the simplest examples, namely, alkadienes that are usually referred to simply as *dienes*.



The term *cumulated* means that one carbon participates in two double bonds; in other words, these bonds follow one after another. This type of bonding occurs very rarely among the natural products. If at least one saturated carbon atom intervenes between the double bonds of a diene, these bonds are regarded as isolated. Thus, 1,2-butadiene is an example of cumulated dienes, and 1,4-pentadiene represents isolated dienes.

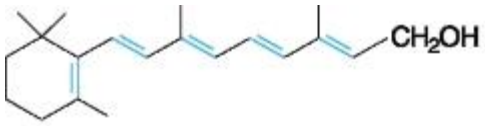
3.1. CONJUGATION AS STABILIZING FACTOR OF MOLECULES

The most interesting are dienes with conjugated double bonds, i. e. 1,3-dienes where the double and single bonds alternate in the chain. There are many polyunsaturated compounds with conjugated double bonds, which play an important role in nature and biology. For example, β-carotene is a yellow-orange pigment in carrots that involves eleven conjugated double bonds:

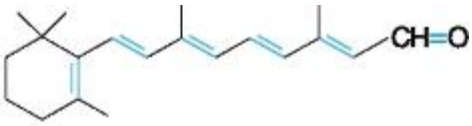


β-carotene
(conjugation chain is underlined in colour)

Other examples of highly conjugated systems among biologically active compounds are retinol (vitamin A) and retinal, the latter being a substance responsible for vision.



retinol (vitamin A)



retinal

3.1.1. π,π Conjugation

Conjugated dienes are similar to nonconjugated dienes (and alkenes) in many but not all of their chemical properties. The former are somewhat more stable than nonconjugated dienes. An explanation for the higher stability of conjugated dienes can be given in describing molecular orbitals of the conjugated systems by example of 1,3-butadiene, a simplest conjugated diene (Fig. 3.1, a).
All carbons in the molecule are sp^2-hybridized, which results in the *delocalization* of π electrons in the bonding molecular orbitals. According to this standpoint, *p* orbitals of the central carbon atoms are overlapped as well (Fig. 3.1, b). It should be stressed here that the interaction between unhybridized *p*orbitals of the atoms C-2 and C-3 can only occur when the molecule

is planar. This provides for parallel arrangement of the *p* orbitals and favours the effective orbital overlap. Delocalization of electron cloud leads to lower-energy orbitals and increased stability of the molecule.

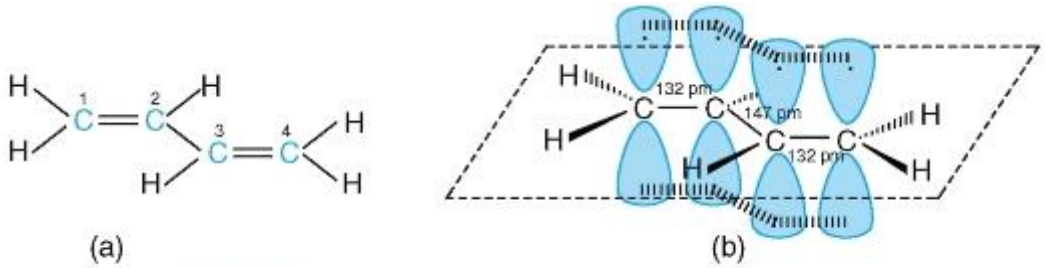


Figure 3.1. An orbital representation of the bonding in 1,3-butadiene.

Another observational proof for the peculiar nature of conjugated dienes arises from data on bond lengths (Fig. 3.1, *b).* Both double bonds of conjugated dienes are slightly longer than those of ethylene, whereas the central C-C bond of 1,3-butadiene is considerably shorter than the single bond of ethane. This signifies that the C-2-C-3 bond of 1,3-butadiene has an intermediate value (147 pm) between a pure single bond (154 pm) and a pure double bond (133 pm) and possesses, therefore, a partial double-bond character.

The type of orbital interaction when the *p* orbitals are delocalized over the

entire π system is called π,π conjugation.

Heteroatoms that form a multiple bond with carbon can be involved in π,π conjugation as well. In this case the C=O, C=N or C=N bonds are conjugated with the C=C double bond, as it is shown by examples of α,β-unsaturated carbonyl compound or nitriles, respectively:

$CH_2{=}CH{-}CH{=}O$ — propenal (acroleine)

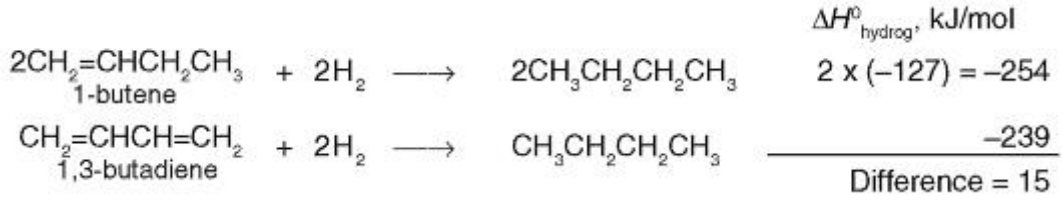
$CH_2{=}CH{-}C{\equiv}N$ — propenenitrile (acrylonitrile)

Problem 3.1. Which of the following molecules represents a conjugated system: (a) 2,4-hexadiene, (b) 3-pentenal, (c) 1,3-cyclohehadiene, (d) 1,4-cyclohehadiene? Draw structural formulas of the compounds and outline a conjugated portion of the molecule.

Energy of conjugation. The value of thermodynamic stability is expressed as the difference between the complete π-electron energy of the nonconjugated system (with localized double bonds) and the π-electron energy of the conjugated system: *E*conjug = *E* nonconjug. system - $E_{conjug.\ system.}$

Experimentally, conjugation energy can be calculated from measurements of heats of hydrogenation($\Delta H°_{hydrog}$). These values for monounsaturated alkenes are -126 ± 1 kJ/mol. We might expect that hydrogenation of dienes would liberate doubled amount of heat. Indeed, $\Delta H°_{hydrog}$ for isolated dienes (1,4-pentadiene, for

example) is approximately -254 kJ/mol. But 1,3-butadiene, a conjugated diene, liberates only 239 kJ/mol, which is markedly less than expected.



This difference represents conjugation energy *(E_{conjug})* for 1,3-butadiene. Highly conjugated compounds possess extra thermodynamic stability. It may be emphasized:

The longer the chain of conjugation, the greater conjugation energy at its formation.

3.1.2. *ρ,π* Conjugation

Another type of conjugation exists in compounds with a fragment >C=CH-X, where × is an atom possessing a lone pair of electrons, namely, oxygen of the hydroxyl and alkoxyl groups, nitrogen of the amino group, or halogens. In this case three orbitals are delocalized, two *p* orbitals of the double bond and one *p* orbital of the atom × (Fig. 3.2).

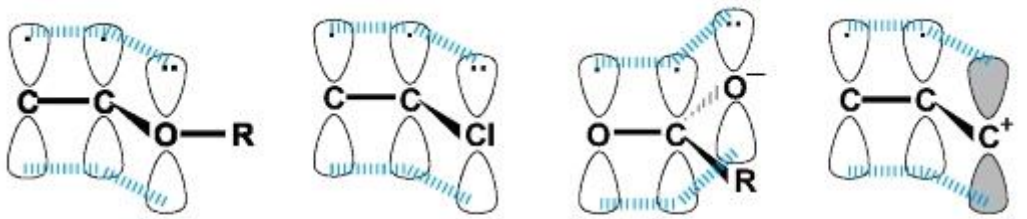


Figure 3.2. *p,π*-Conjugated systems.

The overlap of a *p* orbital on an atom adjacent to a double bond is called *p,π*conju

gation.

p,π-Conjugated systems can be either neutral compounds, or free radicals, or ions (both cations and anions). An allylic cation is exemplified in Fig. 3.2 showing conjugation of an empty *p* orbital of positively charged carbon with *p* orbitals of the double bond. We can now more accurately describe a conjugated system as follows:

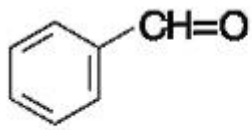
The conjugated system is one that consists of an extended series of overlapping *p* orbitals.

As we will see in further chapters, this phenomenon explains not only extra stability, but also specific chemical properties of conjugated molecules, ions, or radicals. The concept of conjugation is extremely useful in *understanding* chemical and biochemical processes.

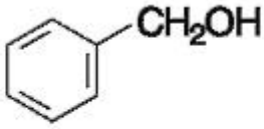
Problem 3.2. State a type of conjugation (the π,π or *p* ,π), *if any*, in the following molecules: (a) CH_3CH_2-CH=O; (b) CH_2=CH-O-CH_3; (c) CH_3CH=$CHCH_2CH$=O; (d) CH_2=CH-CH_2OH.

3.2. AROMATICITY

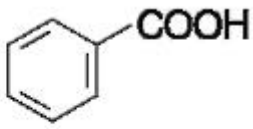
The term *aromatic* was originally applied to relatively simple compounds because of their specific odours. Most of the first aromatic compounds were isolated from balsams, resins, or essential oils, for example, benzaldehyde from bitter almonds, benzyl alcohol and benzoic acid from gum benzoin, and vanillin (see Problem 2.11 for the structure) from vanilla.



benzaldehyde



benzyl alcohol

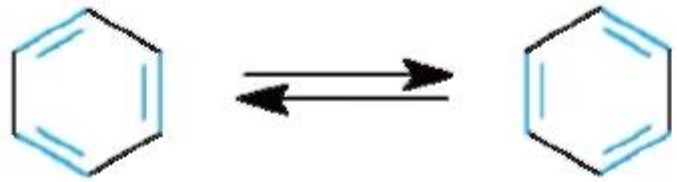


benzoic acid

All these compounds and many other flavoured substances are derivatives of the parent hydrocarbon*benzene,* C_6H_6. Nowadays, however, we use the term *aromatic compounds* not because of their odour, but because of their specific reactivity that can be explained by electronic structure of benzene.

3.2.1. Benzene

Benzene, according to the molecular formula C_6H_6, must be a highly unsaturated compound and we might expect a substantial similarity in the reactivity of benzene and alkenes. However, as we will see in Chapter 7, these hydrocarbons differ markedly in their reactivity. Nevertheless, F.A. Kekule (in 1865) proposed the first structure of benzene as a cyclic compound with three double bonds. To explain chemical diversity of benzene and unsaturated hydrocarbons, Kekule suggested equilibrium between two forms of benzene that results in *rapid* exchange of the positions of the double bonds:



the Kekulé structures for benzene

3.2.2. Modern Theories of the Structure of Benzene

Physical measurements show that benzene is a flat, symmetrical molecule with a shape of regular hexagon with all carbon-carbon bond lengths of 140 pm (compare this value with bond lengths of 147 and 134 pm for the C-C and C=C bonds in conjugated dienes, respectively). All carbons are sp^2-hybridized, as in ethylene. Two sp^2 orbitals form σ bonds with adjacent carbons and the third sp^2 orbital of each carbon forms the C-H bond (Fig. 3.3, a). In addition, each carbon has a *p* orbital containing the fourth valence electron. All six *p* orbitals are perpendicular to the plane of the six-membered carbon framework (Fig. 3.3, b).

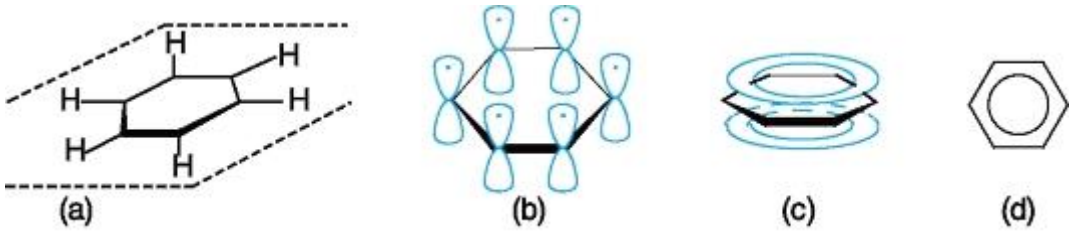
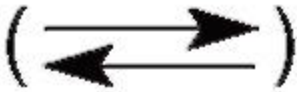


Figure 3.3. Benzene structure: a planar σ skeleton (a), *p* orbital view (b), cyclic conjugation (c), symbolic representation (d). The hydrogen atoms in (b)-(d) are omitted for clarity.

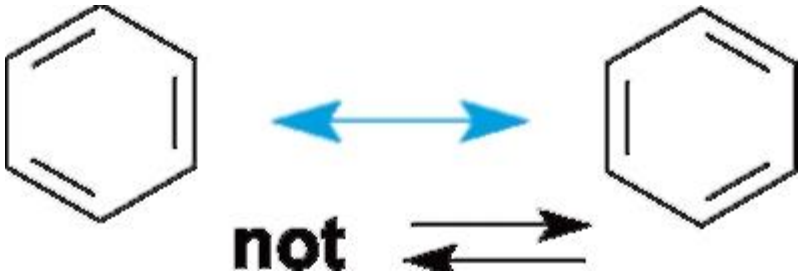
According to *molecular orbital theory* (Sec. 1.3.2), all the six unhybridized *p* atomic orbitals overlap so that it is impossible to define three localized π bonds of alkene type. Rather, each *p* orbital overlaps equally well with both neighboring orbitals to form a cloud of six π electrons completely delocalized around the ring. Thus, the benzene molecule represents a circular π,π-conjugated system with two doughnut-shaped clouds of electrons, one above and one below the ring (Fig. 3.3, *c)*. For this reason, a more satisfactory representation of the benzene molecule might be a hexagon with the inscribed circle (Fig. 3.3, d). However, the Kekule structures are still widely used, including this book, for depicting reaction mechanisms and examining electron distributions in benzene derivatives. Unfortunately, the structure with the inscribed circle does not indicate, first of all, the number of π electrons.

Electron delocalization results in enhanced stability of the benzene molecule. For example, conjugation energy for benzene is 151 kJ/mol; it means that «real» benzene is more stable than hypothetical 1,3,5-cyclohexatriene (the Kekule structure of benzene), which is shown by this value. Extra-stability of benzene may be confirmed by the fact that it remains unchanged on heating up to 700 °C (for comparison, decomposition of alkanes begins at the temperature about 400 °C).

The *resonance theory* developed by L. Pauling (in 1928) assumes *writing* (not the real existence!) two or more structures for a molecule with different arrangements of electrons but with the same arrangements of atoms. Resonance is completely different from isomerism that proposes different atoms arrangement. When resonance is possible, we can draw two or more contributing structures (or canonical forms) for a molecule. The real structure is said to be a *resonance hybrid*, or a blend of the canonical forms. A double-headed arrow  is used between contributing structures to distinguish resonance from an equilibrium, for which we use another sign  . It should be stressed once more:

 Resonance forms are imaginary, not real.

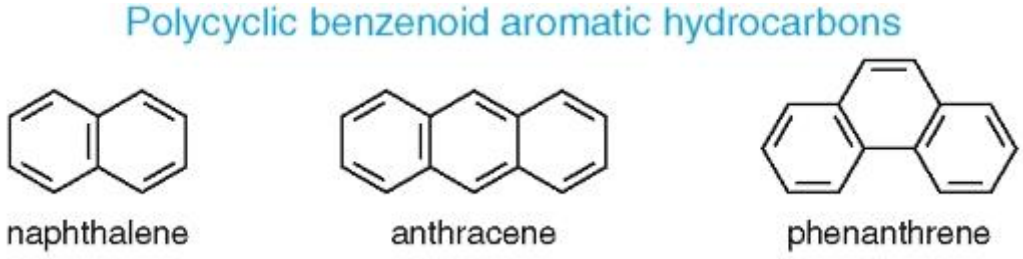
From this point of view, the benzene molecule can be presented as a resonance hybrid of two contributing Kekule structures.



Stabilities of other cyclic conjugated molecules can also be predicted. According to theoretical calculations, the German physicist E. Huckel formulated (in 1931) the following criterion of aromaticity:

A molecule can be aromatic only if it has a planar, cyclic system of  conjugation with a *p* orbital on each atom and only if the p-orbital system contains (4n + 2) π electrons, where *n* is an integer (0, 1, 2, 3, etc.).
It means that only molecules having 2, 6, 10, 14, ... π electrons can be aromatic. This is known today as the *Huckel's* (4n + 2) *rule.*

The concept of aromaticity can be extended beyond monocyclic compounds like benzene and its simple derivatives. In addition to those, there are other aromatic compounds called polycyclic *benzenoid*aromatic hydrocarbons.



These compounds consist of two or more fused benzene rings. The simplest example is *naphthalene,*but more complex compounds involving three fused benzene rings are also well known (see the structures above).

How is the aromaticity of naphthalene explained? Its molecule satisfies all the requirements of aromaticity. Naphthalene is a planar molecule in which all carbons are sp^2-hybridized. The orbital picture of naphthalene (Fig. 3.4) demonstrates a completely conjugated cyclic π electron system, where ten *p* orbitals overlap both around the periphery of the molecule and across the fusion part (the central carbon-carbon bond).

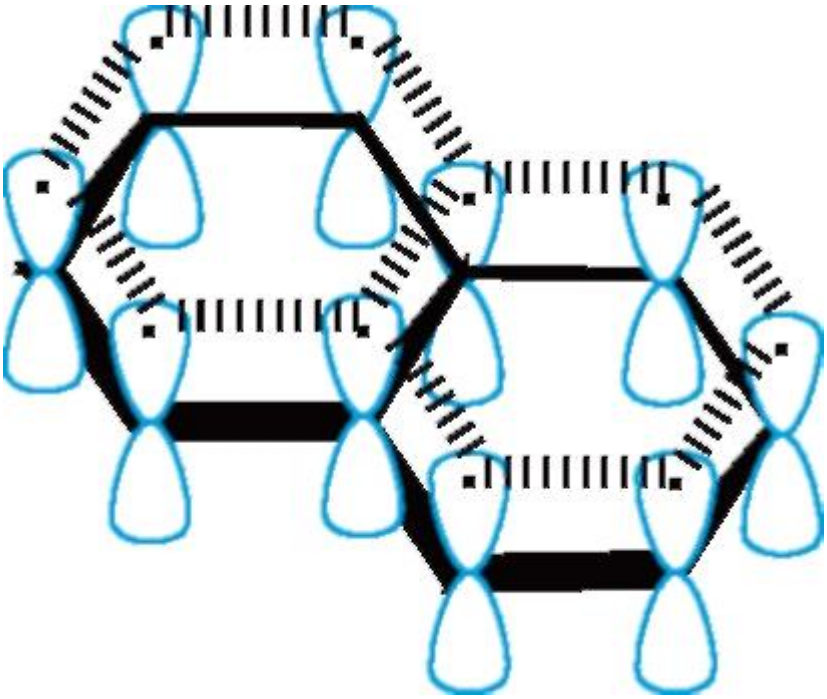


Figure 3.4. An orbital representation of the bonding in naphthalene. The orbital overlap below the carbon skeleton is omitted for clarity.

Ten π electrons is the Huckel's number (n = 2), therefore the naphthalene molecule reveals a high degree of π electrons delocalization and possesses high conjugation energy (250 kJ/mol). From the concept of resonance, three canonical structures can be proposed for naphthalene.

Problem 3.3. In addition to the resonance structure of naphthalene presented above, draw the two remaining resonance structures for the compound.

3.3. ELECTRONIC EFFECTS IN ORGANIC MOLECULES

As it has been shown earlier (Sec. 1.3), the main types of chemical bonds are either covalent or ionic. We generally meet with covalent bonds in organic compounds because their atoms do not differ greatly in electronegativity.

3.3.1. Polar and Nonpolar Covalent Bonds

When two atoms of the same electronegativity are combined together the bonding electrons are shared equally forming a *nonpolar* covalent bond. The simplest examples are molecules of hydrogen, oxygen, halogens. But, if two atoms of different electronegativity form a covalent bond, the bonding electrons are not shared equally. In this case electrons are attracted somewhat more strongly by one atom (which is more electronegative) than by the other. Such a bond, in which electron distribution is unsymmetrical, is called a *polar* covalent bond.

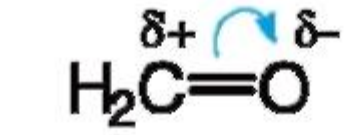
The hydrogen chloride molecule provides an example of a compound with the polar covalent bond. A chlorine atom is more electronegative than hydrogen atom (their electronegativities are 3.0 and 2.1, respectively; see Table 1.3). This means that the bonding electron pair is shifted to chlorine, which becomes partially negatively charged with respect to the hydrogen atom. A short *arrow,* instead of a dash, is used to indicate the direction of polarity. By convention, electrons move in the direction of the arrow. Partial charges designated as δ+ or δ- (should be read «delta plus» or «delta minus») may be shown by the examples of the hydrogen chloride and water molecules:

The C-H bond, which is so common in organic compounds, is relatively nonpolar because carbon and hydrogen have similar electronegativities. Another typical bond in organic chemistry, i. e. a carbon-carbon bond, is also nonpolar when two identical carbons form the bond. The carbons identity means the same type of hybridization or similar environment. For example, the carbon-carbon bond is nonpolar in ethane, CH_3-CH_3, in ethylene, CH_2=CH_2, or in ethylene glycol, $HOCH_2$-CH_2OH. There are, however, many instances of polar carbon-carbon bonds to be discussed later.

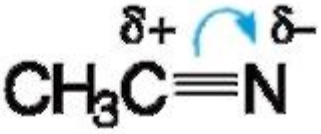
Other elements occurring in organic compounds are mostly situated in right side of the periodic table. Therefore, such elements as oxygen, nitrogen, and halogens are more electronegative than carbon (Table 1.3). Thus, the electrons in C-O, C-N, and C-Cl bonds are attracted stronger by a more electronegative atom, as it is shown in chloromethane and methanol molecules (the O-H bond in the latter is polarized too, of course).

There are few compounds having a bond between carbon and a less electronegative atom (except for hydrogen), e. g. metal. Inverse polarization of the carbon-metal bond is observed in so-called organometallic compounds. A well-known example is tetraethyllead, $(C_2H_5)_4Pb$, an antiknock additive in leaded petrol.

Localized π bonds of functional groups containing a double or triple bond, such as >C=O, >C=N, or -C≡N, are more readily polarized than the corresponding ordinary bonds (C-O or C-N). This can be explained by the fact that π electrons are less tightly held by the nuclei (recall the lateral overlap in the π bond formation). *Curved arrows* are used to show the shifting of the π electrons, as it is demonstrated in the following examples:



formaldehyde



acetonitrile

The curved arrow is always drawn alongside of the bond starting at its middle (the initial position of electrons).

3.3.2. Inductive Effect

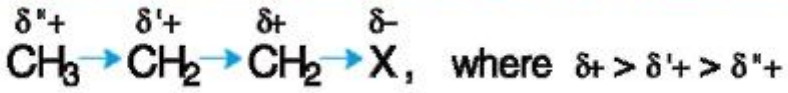
The presence of a polar σ or π bond in an organic molecule results in polarization of neighbouring sites.

The shifting of electrons in a bond in response to electronegativity of

nearby atoms is called the inductive effect.

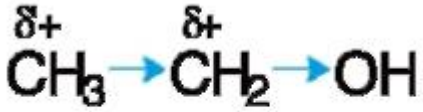
The inductive effect symbolized by the letter *I* can be *electron-withdrawing* (negative, designated *-I)* or*electron-donating* (positive, designated *+I).* In the first case electron density at the nearby site is decreased, in the second case it is increased. The inductive effect extends to include three (maximum four) bonds owing to low polarizability of C-C bonds. The effect of a substituent is the strongest on the neighbouring atom decreasing along the carbon chain in the following way:

The –*I* effect of an electronegative substituent X



The inductive effect is only a qualitative characteristic of a substituent since it is very difficult to take into consideration all particulars in electron density distribution. Nevertheless, functional groups with double-bonded oxygen reveal a

certain strong -*I* effect. Thus, a partial positive charge on the atom C-1 in acetaldehyde is greater than a charge on the atom C-1 in ethanol:



ethanol

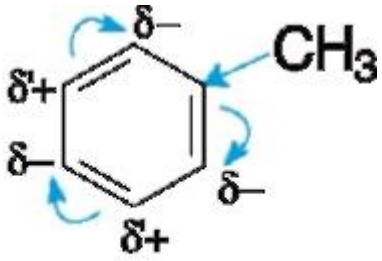
acetaldehyde

Other substituents that have strong -*I* effect are a sulfonyl group, $-SO_3H$, a cyano group, $-C\equiv N$, a nitro group, $-NO_2$, and a carboxyl group, -COOH.

Problem 3.4. Using arrow symbolism show the distribution of electron density in the following molecules: (a) ethylamine, (b) dimethyl ether, (c) propanoic acid, (d) acetone (propanone).

The positive inductive effect is rarely to be observed. It may be attributed to methyl (or other alkyl) group, if only this group is attached to sp^2-hybridized carbon. The polarity of the C-C bond in this case is accounted for the difference in electronegativity of the sp^3- and sp^2-hybridized carbons. So we can say that the methyl group in propene or toluene has the *+I* effect in regard to the double bond or the benzene ring, respectively:

propene



toluene

The *+I* effect of the methyl group results in appearance of partial charges on carbons of the double bond in propene and on carbons of the benzene ring. It should be noted that the electron-donating methyl group increases electron density on *all* carbon atoms but mainly on the atom C-1 in propene or the atoms C-2, C-4, and C-6 *(ortho* and *para* positions) of the benzene ring according to the arrow head, regardless of the signs δ+ used. This means the symbols δ+ and δ- are relative.

3.3.3. Mesomeric Effect

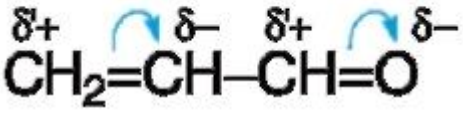
A more pronounced electronic effect is observed in molecules having conjugated fragments. The polar effect of a substituent extends in this case through the entire system of conjugation.

The shifting of electron density caused by a substituent in conjugated

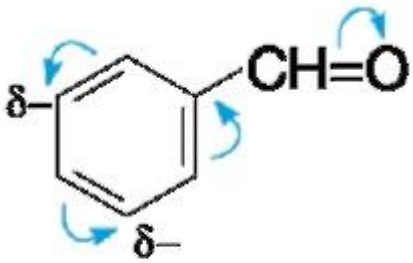
system through *p* orbital overlap is called the mesomeric (or resonance) effect.

The mesomeric effect symbolized by the letter *M* can be, like the inductive effect, electron-donating (designated *+M)* or electron-withdrawing (designated *-M)*. Let us consider the mesomeric effect in different π,π- and p,π-conjugated systems.

Acroleine (propenal) and bezaldehyde represent π,π-conjugated systems. The aldehyde group in these compounds withdraws the electron density from the C=C double bond or the benzene ring, respectively:



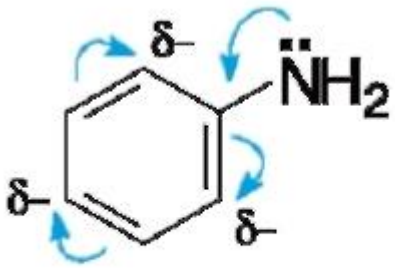
acroleine



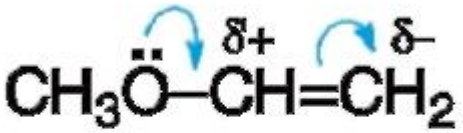
benzaldehyde

The -M effect of the functional group leads to decreased electron density on all the carbons of the remaining part of the molecule in comparison with unsubstituted compounds (ethylene and benzene). But less electron deficient are the atom C-2 in acroleine as well as the atoms C-3 and C-5 (both *meta*positions) in benzaldehyde as it follows from the heads of the arrows.

The positive mesomeric effect is observed in most p,π-conjugated systems, where a substituent with a lone pair of electrons donates electrons to the neighbouring benzene ring or a π bond. It is demonstrated in the following examples:



aniline

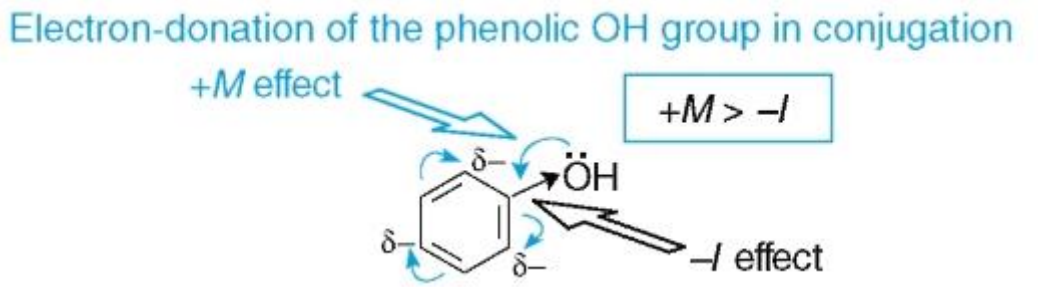


methyl vinyl ether

It should be mentioned once more that electron density is increased on all carbons of the benzene ring in aniline as well as on both double-bonded carbons of the unsaturated ether.

As is evident from the above, the mesomeric effect of a substituent can only be observed in the conjugated systems. The substituents with *+M* effect have an atom that possesses a lone pair of electrons and is *directly* attached to an aromatic ring or to a double bond. Examples of these substituents are hydroxyl, amino, alkoxyl (-OR) groups, substituted amino group (-NHR, -NR_2), and halogens. Conversely, the substituents with -M effect have the general structure -X=Y, where the atom Y is a more electronegative atom than the × one. Many important functional groups fit in this category, namely, a carbonyl group of aldehydes and ketones, a carboxyl group, a nitro group, a sulfonyl group, a cyano group (a single representative with a triple bond).

It may seem surprising that hydroxyl, amino and similar groups belong to electron-donating substituents. Indeed, both oxygen and nitrogen are highly electronegative and would be supposed to inductively withdraw electron density from the aromatic ring (or double bond). In reality, the electron-donating (+M) effect through *p,π* conjugation is stronger, than the electron-withdrawing *(-I)* effect through a σ bond.



Problem 3.5. Indicate the electronic effect of the methoxyl group in anisole, $C_6H_5OCH_3$. Show partial charges on all carbons of the benzene ring. Compare electron density on the cyclic carbons in anisole and benzene.

Thus, these substituents (with the exception of halogens) possess a net electrondonating property (Table 3.1).

Table 3.1. Electronic effects of substituents

Substituent		Electronic effects (I or/and M)		Electron donor (D) or withdrawer (W) (at sp^2-hybridized carbon)
Name	Formula			
Alkyl	CH_3, C_2H_5, etc.	+I	no M	D
Amino	NH_2, NHR, NR_2	–I	+M	D
Hydroxyl	OH	–I	+M	D
Alkoxyl	OR	–I	+M	D
Halogens	Cl, Br, I	–I	+M	W
Nitro	NO_2	–I	–M	W
Carboxyl	COOH	–I	–M	W
Carbonyl	>C=O	–I	–M	W
Cyano	–C≡N	–I	–M	W
Sulfo	SO_3H	–I	–M	W

Additional Problems

3.6. Which of the following statements is correct for the 1,3-butadiene molecule?

(a) it contains sp^2- and sp^3-hybridized carbons;

(b) all σ bonds lie in a plane;

(c) it is a p,π-conjugated system;

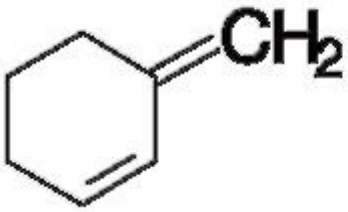
(d) conjugation results in the shortening of all carbon-carbon bonds;

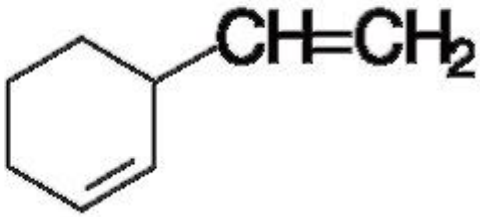
(e) it possesses enhanced thermodynamic stability.

3.7. Select the conjugated compounds from the following ones. Mark a conjugated portion of the molecule.

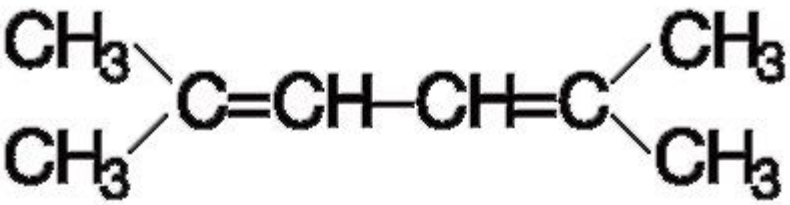
(a) CI-CH=CH-CH=CH_2

(b) CH_2=CH-COOH

(c) 

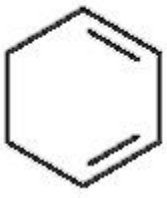
(d) 

3.8. Which of the following molecules is a conjugated one? Mark a conjugated portion of the molecule.

(a) 

(b) HO-CH_2-CH=O

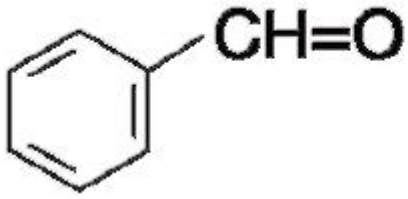
(c) CH_2=O=CH_2,

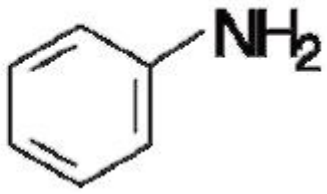
(d) 

3.9. What type of conjugation (π,π or p,π), *if any*, do you observe in the following molecules?

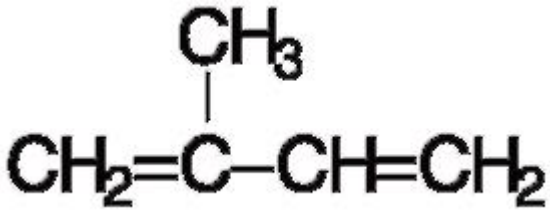
(a) CH_2=CH-CI

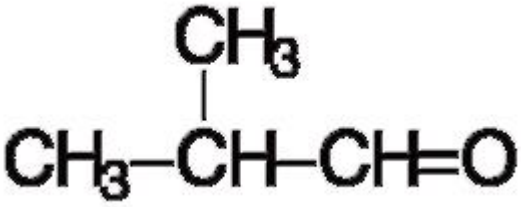
(b) CH3CH2-O-CH3

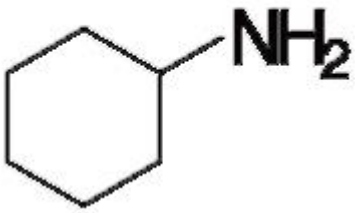
(c) 

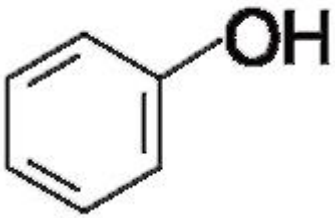
(d) 

3.10. Select the π,π- or p^-conjugated compounds from the following ones. Explain your choice.

(a) 

(b) 

(c) 

(d) 

3.11. Which of the following statements is correct for the phenol molecule?

(a) all carbons are sp^2-hybridized;

(b) all carbons lie in a plane;

(c) the hydroxyl group participates in conjugation;

(d) it is a p,π- and π,π-conjugated system;

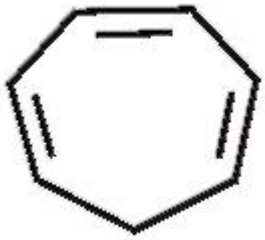
(e) it contains three C=C double bonds.

3.12. Which of the following compounds is defined as the conjugated system? Which of them is an aromatic compound?

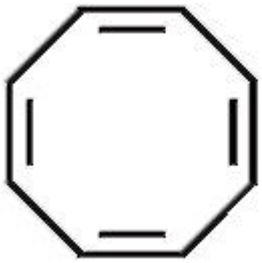
(a) CH_2=CH-CH=CH-CH=CH_2

(b)

(c)



(d)



3.13. Indicate the type of the electronic effects (as electron-donating or electronwithdrawing) of the following groups attached to an alkane part of the molecule:

(a) -COOH;

(b) -OH;

(c) -Cl;

(d) -NH_2;

(e) -CH=O; (f) -CH_3.

Explain your choice using arrow symbolism.

3.14. Indicate the type of the electronic effects (as electron-donating or electronwithdrawing) of the following groups attached to the benzene ring:

(a) -CH_3;

(b) -COOH;

(c) -OH;

(d) -NH_2;

(e) -Br;

(f) $-NH_3^+$.

Explain your choice using arrow symbolism.

3.15. Pick out molecule(s) in which the electron density on the ethylene fragment is higher than that on the ethylene molecule:

(a) $CH_3CH=CHCH_3$;

(b) $CH_3CH=CHCOOH$;

(c) $CH_2=CHOCH_3$;

(d) $CH_2=CHCl$;

(e) $CH_2=CHCH=O$;

(f) $CH_3CH=CH_2$.

Explain using arrow symbolism.

3.16. Pick out molecule(s) in which the electron density on the benzene ring is higher than that on the benzene molecule:

(a) benzoic acid;

(b) methoxybenzene;

(c) 3-chlorobenzaldehyde;

(d) 1,3-dinitrobenzene;

(e) 4-ethylphenol;

(f) 2-methylaniline.

Explain using arrow symbolism.

3.17. Show the order of increasing of the electron density on the aromatic ring for the following molecules:

(a) *p*-aminophenol;

(b) aniline;

(c) benzene;

(d) benzoic acid;

(e) chlorobenzene;

(f) toluene.

Use arrow symbolism in your answer for explanation.

Chapter 4. A BRIEF SURVEY OF ORGANIC REACTIONS

Millions of organic compounds undergo millions of chemical transformations and, at first sight, organic chemistry could be considered a mere

collection of infinite numbers of disconnected facts. There are only a few general concepts that can be applied to most chemical reactions. In fact modern organic chemistry is a very logical subject, therefore the best way to learn it consists in understanding concepts but not in memorization of numerous facts.

This chapter is concerned with fundamental types of organic reactions, classification of reactants, and general description of the reactions.

4.1. TYPES OF ORGANIC REACTIONS

While studying organic chemical reactions two aspects should be taken into account: what types of reactions exist and how a reaction proceeds. First it is more reasonable to look at the types of reactions. Organic reactions can be grouped into several types according to the overall result. They are:

- addition reactions;
- elimination reactions;
- substitution reactions;
- rearrangements (or isomerization reactions);
- oxidation and reduction reactions;
- acid-base interactions.

The latter reactions will be discussed in detail in Chapter 6. Other types of reactions listed above will be briefly characterized.

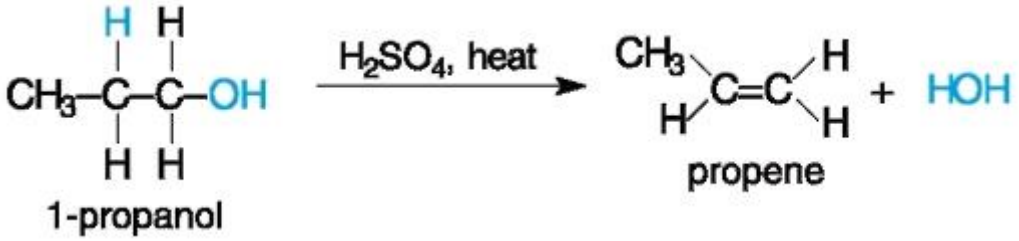
Addition reactions are those in which two substances react to form a single compound. In this case no atom is removed from both reactants. It can be generalized as follows:

An example of this reaction is the addition of a halogen to an alkene:

Elimination reactions are, to some extent, the reverse of addition reactions. The parts of a molecule are removed from adjacent atoms of the reactant to give two products:

One product is an unsaturated compound and the other is usually a small inorganic molecule such as hydrogen, water, ammonia, and so on. The reaction is

exemplified in dehydration (elimination of a water molecule) from an alcohol under appropriate conditions:

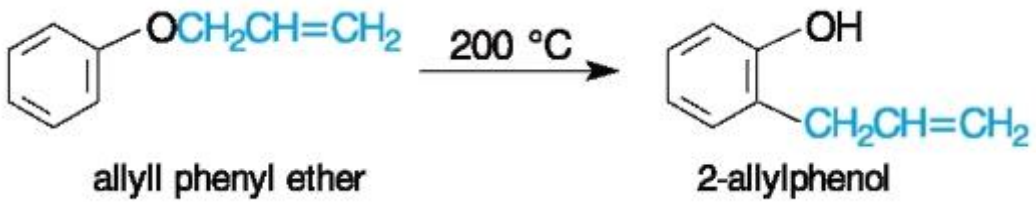


Substitution reactions occur when a group of atoms of one reactant replaces a group of atoms of another reactant. In other words, two reactants exchange their parts to give two new products:

These reactions are among the most common and most useful ones in organic chemistry. An example of such a reaction is demonstrated by the action of an aqueous alkali on a haloalkane to give an alcohol:

Rearrangements consist in reorganization of chemical bonds and atoms in a reacting molecule on retention of the molecular formula, i. e. in formation of an isomeric product.

This process often accompanies reactions of other types, but rearrangements in «pure form» are known as well, for example, on heating of aromatic allyl ethers:

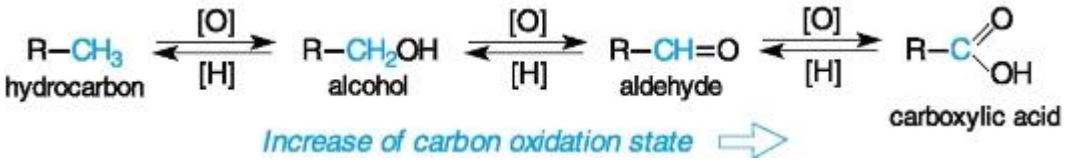


Oxidation and reduction reactions are defined in organic chemistry in a different manner as compared with definitions in inorganic chemistry.

Remember that in inorganic chemistry oxidation is defined as the loss of electrons, and reduction is defined as the gain of electrons. In organic reactions it is not easy to estimate whether electrons are removed from an atom or added to it. The notion of an oxidation number (or oxidation state) is necessary only for writing balanced equations.

Oxidation of an organic molecule usually corresponds to increasing its oxygen content or decreasing its hydrogen content. Conversely, reduction of the organic molecule is classified as decreasing the oxygen content of the molecule or increasing its hydrogen content. The symbols [O] and [H] are often used to

indicate oxidation and reduction processes, respectively. Oxidation-reduction reactions can be represented as



follows:

Problem 4.1. Classify the following reactions according to the types discussed before:

(a) $(CH_3)_3COH + HCl \longrightarrow (CH_3)_3CCl + H_2O$

(b) $CH_2=CH_2 + H_2 \longrightarrow CH_3CH_3$

(c)  $+ 3H_2$

(d) $C_6H_5CH=O + O_2 \longrightarrow C_6H_5COOH$

(e) $CH_3CH=O + H_2 \longrightarrow CH_3CH_2OH$

4.2. REACTION MECHANISMS

Chemical equations express the overall result of chemical transformation. They show the structures of reactants and products, but they tell us nothing how the reactants are turned to the products. Now we are interested in how the reaction takes place, i. e. its *reaction mechanism.*

A reaction mechanism is a step-by-step description of all changes in

reacting compounds that occur at the molecular level, as the reactants become products.

The mechanism describes which bonds are broken and which bonds are formed and in what order. It takes into account energy changes on the pathway from the reactants to the products including relative stability of intermediates formed in the reaction. The complete mechanism must also account for the relative rates of the steps and stereochemical result of the reaction.

4.2.1. Radical and Polar Processes

All reactions of organic compounds involve the breaking and making of covalent bonds. Generally speaking, there are two ways in which a covalent bond can break in a hypothetical molecule A:B. In the first one, the bond breaks in an electronically symmetrical way so that one electron is retained with each fragment:

Such bond splitting leads to the neutral species A' and B' that have an unpaired electron in one of its orbitals and are called *free radicals,* or often simply radicals. This type of bond breaking is known as*homolysis* (from the Greek *homo* - the same, and *lysis* - cleavage), or *homolytic* process.

It should be noted that we use a half-headed curved arrow (like a «fishhook») to show the movement of only one electron, whereas we use an ordinary curved arrow to show the movement of an electron pair. The tail of the arrow indicates the position of the electrons prior to the reaction, and the head of the arrow shows their location after the reaction.

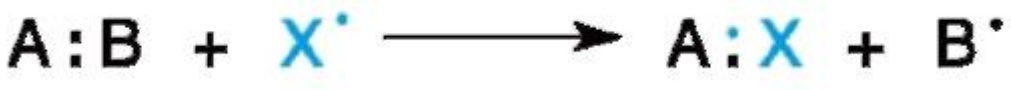
Remember, we have already used curved arrows to describe *p* electron shifting for representation of a mesomeric or inductive effect. To avoid confusion, the electron movement in the bond breaking and bond making are shown here and later in the text with *coloured* arrows. Black arrows will be used for representation of electronic effects.
The amount of energy required for the homolytic cleavage of a covalent bond is called the *bond dissociation energy,* or simply the bond energy. The same amount of energy is released when the radicals A' and B' combine to form the product A-B. Appendix 1 lists some data for bond dissociation energies.

Another type of bond breaking called *heterolysis* or heterolytic process results in formation of charged fragments or ions (for example, a cation A+ and an anion B^-) as it is shown below. This is a consequence of electronically unsymmetrical splitting in which both bonding electrons remain with one fragment, leaving the other fragment with an empty orbital:

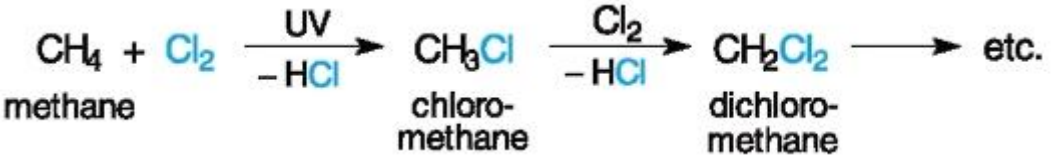
Similarly, there are two ways in which a covalent bond can form: a *homogenic* way when each reactant donates only one electron to the new bond, or a *heterogenic* way when both bonding electrons are donated to the new bond by one reactant.

With the foregoing considerations, the overwhelming majority of organic chemical transformations may be classified as *radical reactions* and *polar reactions.* The former involve symmetrical bond breaking and bond making with participation of a free radical reactant and intermediates. The latter involve unsymmetrical bond breaking and bond making with participation of species with an even number of valence electrons. Both types of the reactions will be exemplified here only schematically.

Radical reactions. A free radical X' can take away a part from a molecule A-B, for example, a species A' yielding a product A-X and leaving behind a new radical B'. The net result of the reaction is radical substitution:



These reactions are typical for nonpolar organic compounds, e. g. alkanes, subjected to reagents that bear free radicals under appropriate conditions (Sec. 4.2.2). Chlorination of methane demonstrates this:



Polar reactions. These reactions occur as a result of attractive forces between positive and negative charges (full or partial) on both reactants. Despite the fact that most organic substrates are electrically neutral, certain bonds within a molecule are polar. It will be recalled that bond polarity is a consequence of the difference in atom's electronegativity.

The term *substrate* originates from biochemical terminology and means a compound that is exposed to a*reagent* (the latter is an enzyme in biochemical reactions).

Polar reagents of different types can react with organic substrates. Examples of substitution reactions are represented below where both substrates A-B and polar reagents XY differ from each other in reactions (1) and (2):

Polar reactions are of very common occurrence in organic chemistry. No example will be given here because most organic reactions, especially those discussed in this book, belong to this type.

4.2.2. Types of Reagents

We saw in the preceding section that two types of reagents take part in a polar reaction, namely the electron-poor reagent X^+ in reaction (1) or the electron-rich reagent Y^- in reaction (2). The former has an electron-poor site and seeks electrons in the substrate. The latter has an electron-rich site and can form a bond by donating a pair of electrons to an electron-poor site in the substrate. Polar reagents are classified on this basis as either *electrophiles* or *nucleophiles* (literally, electron-loving or nucleus-loving, respectively).

An electrophile is an electron-poor reagent; a nucleophile, by contrast, is

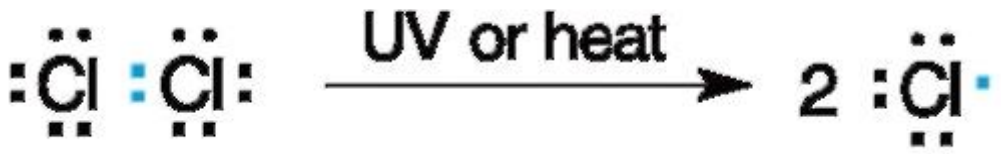
an electron-rich reagent.

These two types of reagents are illustrated in the addition reaction of ethylene and in the substitution reaction of ethyl bromide (detailed mechanisms of these reactions will be considered in further chapters).

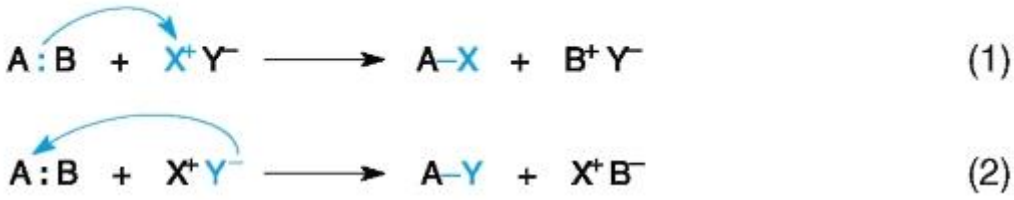
Electrophiles are often positively charged. Typical electrophiles are a proton, halonium ions (Cl^+ and Br^+), carbocations, or neutral molecules such as sulfur trioxide, SO_3, or compounds of the general formula R-X, where × is an electron-withdrawing group.

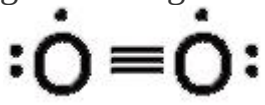
Nucleophiles are often, though not always, negatively charged. The most widely known nucleophiles are a hydroxide ion, alkoxide ions (RO^-), thiolate ions (RS^-), halide ions, a hydride ion (H^-), carbanions (particles with negatively charged carbon), and many neutral compounds such as water, alcohols, thiols, ammonia, and amines. Both electrophilic and nucleophilic reagents will be further considered in more details.

Free radicals. These neutral species contain an atom with an unpaired electron in its outer shell; they are usually highly reactive for this reason. Halogen radicals such as Cl' and Br' are of great importance in organic chemistry, particularly in some industrial processes. They can be produced from halogens on heating or by ultra-violet radiation. The process represents homolytic cleavage of relatively weak bond in halogen molecules:



POLAR SUBSTITUTION REACTIONS



A compound that possesses unpaired electrons in an outer shell is ordinary molecular oxygen. On the basis of paramagnetic properties and the interatomic distance (121 pm) which is much shorter than the length of a single bond O-O (148 pm), the O_2 molecule can be described as a biradical  in a ground state. As it is well known, oxygen is responsible for many oxidation reactions both *in vitro,*including industrial processes, and in living systems.

Thus, taking into consideration a type of reaction and the nature of reagents we can classify addition reactions as electrophilic, nucleophilic, and radical ones. Mechanisms of these reactions are designated by the symbols A_E, A_N, and A_R, respectively. Substitution reactions can be classified in a similar way as electrophilic, nucleophilic, and radical ones, using the symbols S_E, S_N, and S_R, respectively.

All these reactions are studied in this book. General principles will be demonstrated first with simple examples and then transferred to objects of biological importance.

Problem 4.2. Classify reactions (1) and (2) represented in Sec. 4.2.1 according to their mechanisms.

4.3. ENERGETICS OF CHEMICAL REACTIONS

Subject matters presented in this section are usually discussed in a course of General Chemistry. A cursory introduction to the question will be made here only for understanding a reaction mechanism.

4.3.1. Activation Energy and Reaction Energy Diagram

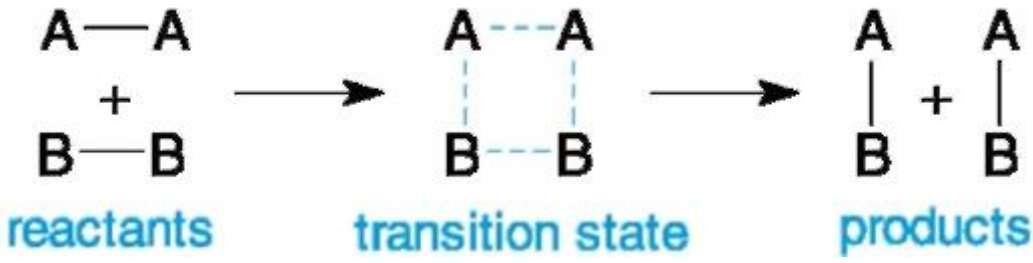
In order to react, reactant particles must collide with each other. Not all collisions between them are effective, however. A chemical reaction takes place when some conditions are provided for redistribution of the electron density of the colliding molecules. To form products, the molecules must have enough energy and the right orientation so that the breaking and making of the bonds can occur. The course of the reaction can be characterized by three consecutive states:

For a generalized one-step reaction between two gaseous compounds A_2 and B_2, the *transition state*corresponds to the formation of an activated complex A_2

•

•

• B_2. In this complex, the bonds A-A and B-B break simultaneously with the formation of new A-B bonds:



The activated complex exists for a very short time (less that 10^{-10} sec) and cannot be isolated, of course. Then it decomposes to form either a product A-B or the initial compounds A_2 and B_2. The energy required for the formation of the

activated complex is a barrier to reaction. It is called the*activation energy*, E_a, and it determines how rapidly the reaction occurs. The greater is the activation energy, the slower is the reaction, and *vice versa*[1].

[1] This resembles the situation of a traveller who climbs over a mountain pass. The traveller needs a lot of energy to overcome a high pass and will reach a top slowly. If the pass is low, the traveller needs less energy and will surmount the barrier quickly.

Most organic reactions have activation energies in the range of 170-650 kJ/mol. Reactions with activation energies less than 80 kJ/mol occur spontaneously at room temperature or below, whereas many reactions require heating to provide the energy necessary for the reactants to overcome the activation barrier. Some reactions, including highly exothermic ones, are either retarded or do not take place at all because of high-energy barriers. It is well known that wood, paper, petrol, and other inflammables can be oxidized and burn in air. But they do not ignite spontaneously under normal conditions because the reaction requires a considerable activation energy.

The energy changes occurring in the course of a reaction is usually depicted in the form of a *reaction energy diagram* (Fig. 4.1). The progress of the reaction from the beginning (left) to the end (right) is plotted against the X-axis that is called the reaction coordinate. The potential energy of the system is plotted against the ordinate.

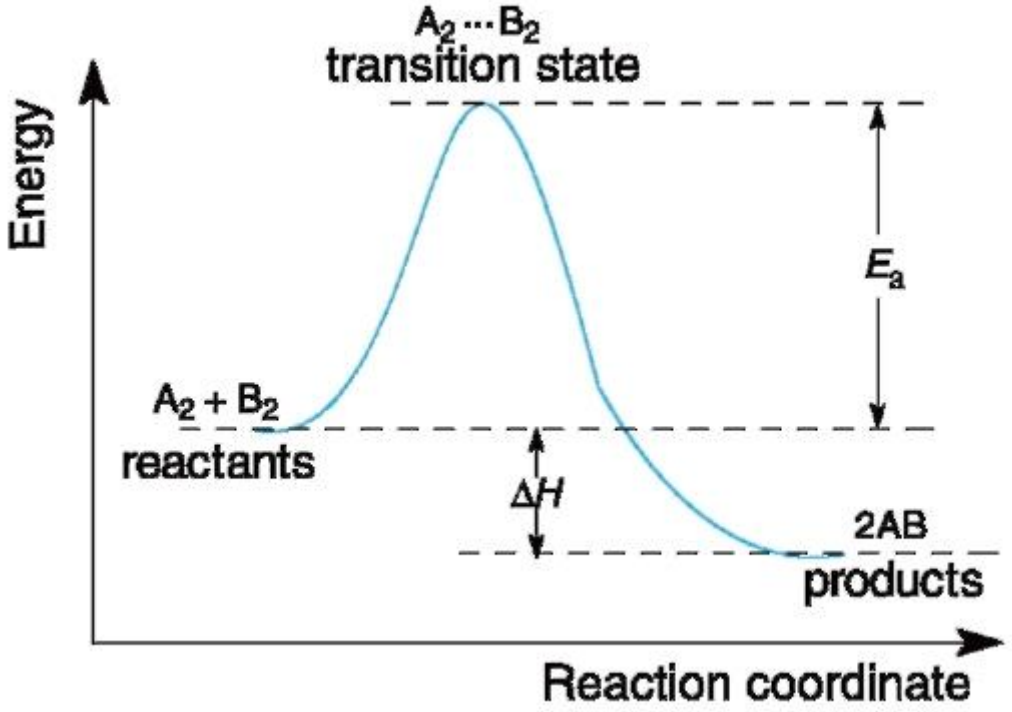


Figure 4.1. Reaction energy diagram for the one-step $A_2 + B_2 = 2AB$ reaction

The energy spent on activating the reactants is later, partly or wholly, liberated during the formation of the reaction products. If the amount of energy liberated during the decomposition of the activated complex is greater than the activation

energy, heat is given off in the reaction. Such a reaction is*exothermic,* as it is shown in Fig. 4.1. In other words, the products are lower in energy than the reactants, i. e. the difference in enthalpy between products and reactants is negative (AH < 0). Otherwise, when heat is absorbed (ΔH > 0), the reaction is *endothermic.*

4.3.2. Catalysis

One of the common methods of accelerating reactions is catalysis. A catalyst is a substance that increases the reaction rate by participating in an intermediate chemical interaction with reactants and is recovered unchanged after the reaction. The course of the overall reaction is altered in the presence of a catalyst and the rate of the reaction is altered as well. Rate increasing is associated with lower activation energy of a new pathway, which is usually multi-step in catalyzed reactions.

Let us consider a reaction between compounds A and B that forms, through an activated complex

• B, a product AB (the black curve in Figure 4.2). If its activation energy is high, the reaction if any proceeds at a very slow rate. Let us now assume that a catalyst C reacts readily with A to form an intermediate compound AC (the coloured curve in Figure 4.2):

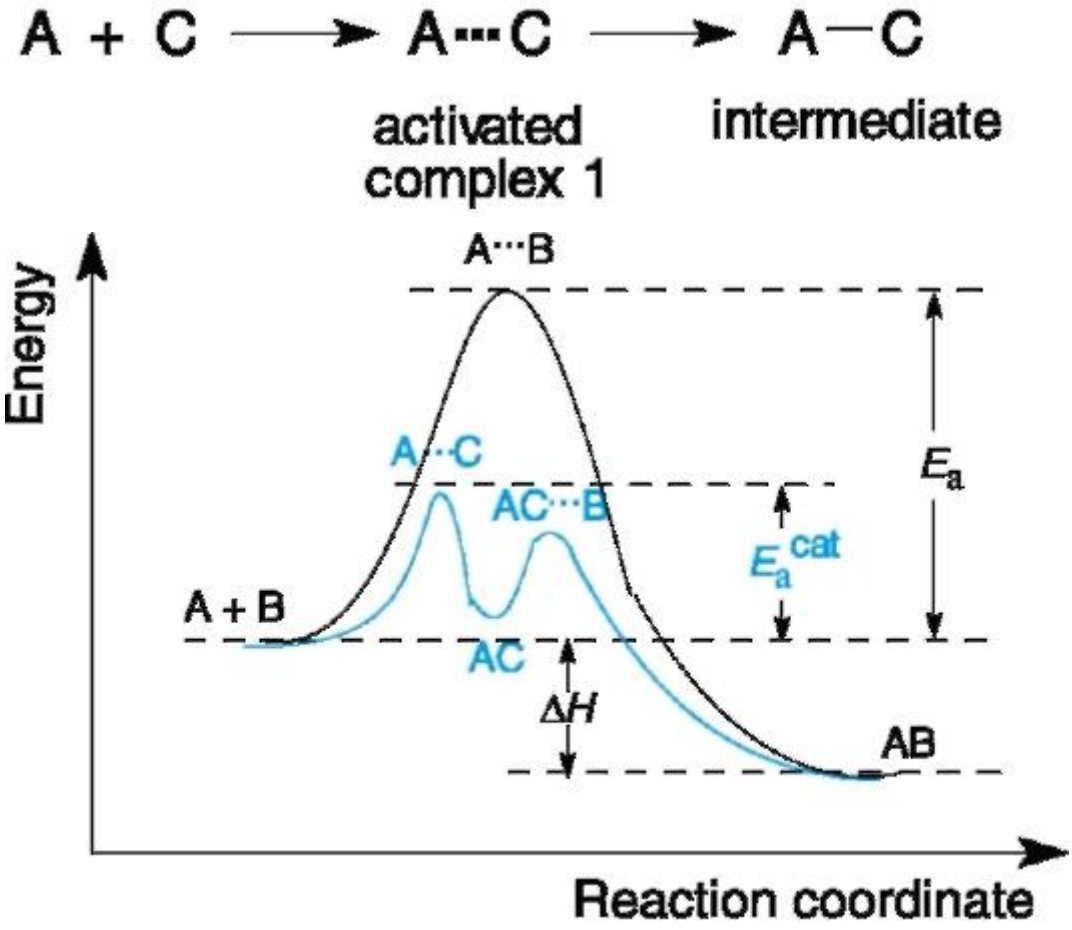
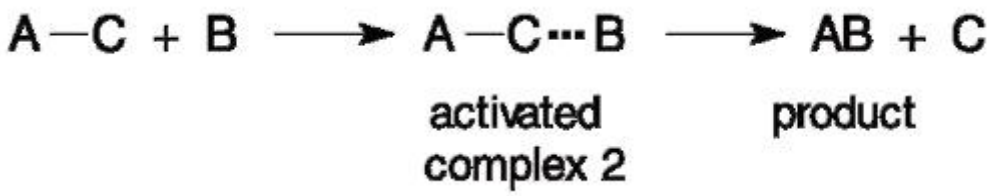


Figure 4.2. Reaction energy diagram for a noncatalyzed reaction (the black curve) and for a catalyzed reaction (the coloured curve)

The catalyzed route has the lower activation energy, E_a^{cat}, than the original, noncatalyzed path. The compound AC reacts then readily with the reactant B (for the same reasons) to form the product AB and a recovered catalyst C:



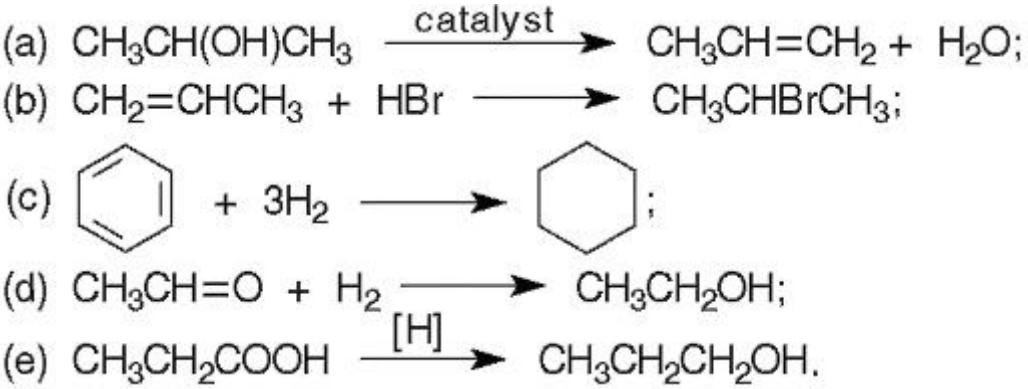
It should be noted that the activation energy for the reverse reaction is lowered exactly by the same amount as for the forward reaction. Thus the catalyst speeds up the *two* reactions and has no effect on chemical equilibrium. It only accelerates the attainment of equilibrium in the system. On the other hand, the value of ΔH does not change when the catalyst is used as it is shown in Figure 4.3. Only the reaction mechanism and the activation energy change.

Catalytic reactions are most varied in organic chemistry. As we will see later, a proton and a hydroxide ion have a great effect on the rate and direction of chemical reactions. The catalytic effect appears in many reactions in an indirect form. These are the reactions in solution where a solvent activates the reactants by means of polarization or ionization.

There is no doubt that the most important catalysts are *enzymes*, organic molecules responsible for catalyzing a broad spectrum of reactions in living systems. Biochemistry is a science which studies biological processes. The application of chemical principles and an understanding of molecular structures have been important in advances made in biochemistry.

Additional Problems

4.3. Classify the following reactions according to the types discussed in Section 4.1:



4.4. Which of the following species would you expect to belong to electrophiles and which to nucleophiles, *if any:*

(a) methane;

(b) carbon dioxide;

(c) water;

(d) a phenoxide ion, $C_6H_5O^-$;

(e) a hydronium ion, H_3O^+?

Chapter 5. ACIDITY AND BASICITY OF ORGANIC COMPOUNDS

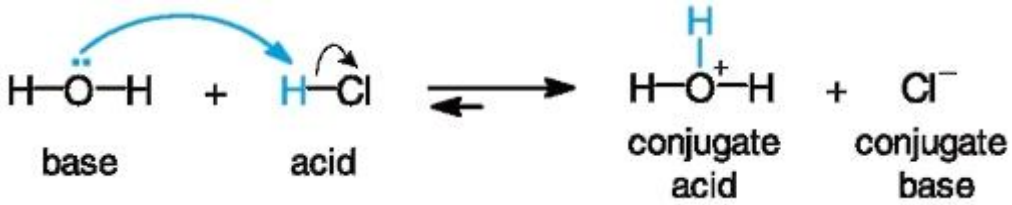
Acidity and basicity are the main notions determining many fundamental physico-chemical and biochemical properties of organic compounds. First of all, acid and basic catalyses are the most widespread enzymic reactions. Besides, the acid-base behaviour of organic compounds helps explain much of their chemistry.

5.1. GENERAL CONCEPTS OF ACIDS AND BASES

At present, there are two main concepts of acids and bases in organic chemistry. In the first one, independently proposed by the Danish physico-chemist J.N. Branstedt and the English chemist T.M. Lowry (in the 1920's), it is stated that:

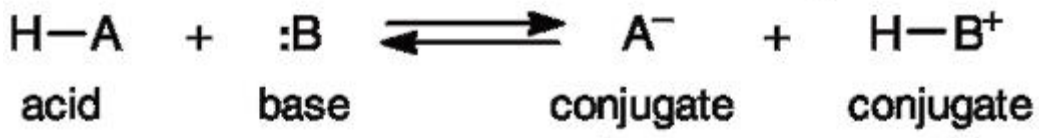
An acid is a neutral molecule or an ion that can donate a proton, and a base is a neutral molecule or an ion that can accept a proton.

For example, when gaseous hydrogen chloride dissolves in water, the latter accepts a proton from hydrogen chloride:

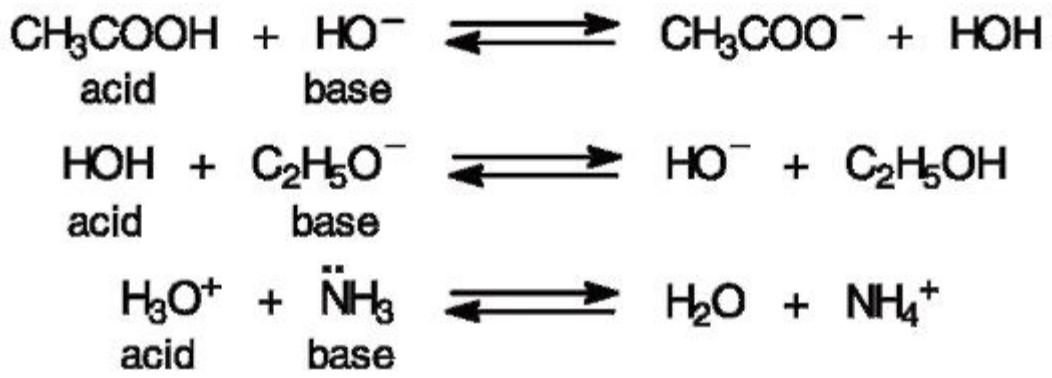


Here hydrogen chloride acts as a proton donor (an acid), and water acts as a proton acceptor (a base). The products of this reaction are the hydronium ion, H_3O^+ (the *conjugate acid* of water), and the chloride ion, Cl^- (the *conjugate base* of hydrogen chloride). Note that the term *conjugate* is used in another sense different from that in Chapter 3.

In a general sense, acid-base reaction can be expressed in the following way:



Other examples of acid-base interaction are presented below:



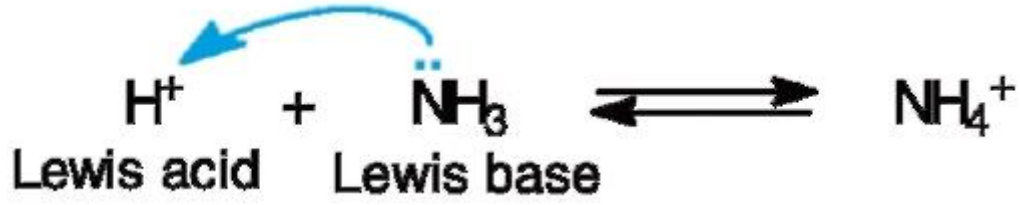
In the above examples water acts either as an acid or as a base, as well as it represents either a conjugate acid or conjugate base.

Another definition of acids and bases was proposed by G.N. Lewis (1926).

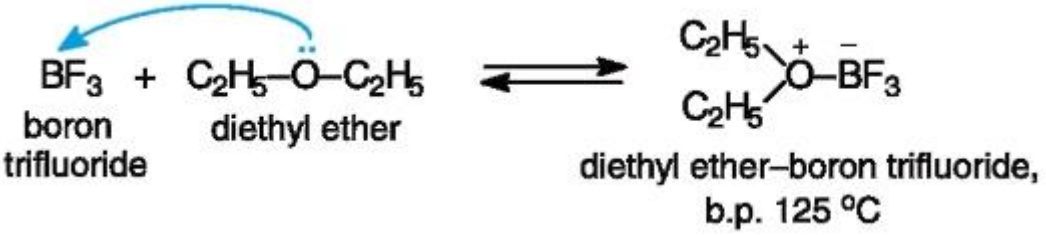
A Lewis acid is any substance that can accept an electron pair, and a Lewis

base is any substance that can donate an electron pair in forming a covalent bond.

According to this definition, a Lewis acid must have a vacant orbital for bonding. The simplest Lewis acid is a proton because it can accept an electron pair from a Lewis base to fill its 1 *s* shell, for example:



Not only proton donors belong to Lewis acids but also many other species with an atom whose valence shell is unfilled, such as various metal cations or compounds of Group 3A elements (BF_3, $AlCl_3$). Thus, boron trifluoride (a gaseous Lewis acid) reacts with diethyl ether (a liquid Lewis base, boiling point 36 °C) to form a stable addition product:



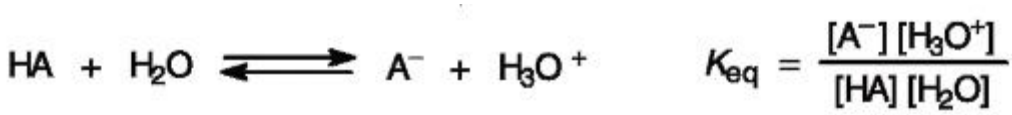
The Lewis definition of acids is much broader than the Bronstedt definition, whereas bases are defined very similar in both theories. The Lewis conception finds a use for a big variety of organic reactions. The only, but substantial, disadvantage of the Lewis theory consists in the absence of quantitative acidity and

basicity scales. The Bronstedt concept is therefore more widely used in organic chemistry.

5.2. ACIDS

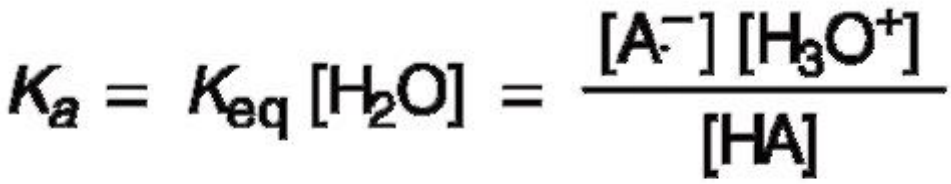
Acids differ greatly in their proton-donating properties. Stronger acids, such as sulfuric, nitric, hydrochloric and similar acids, react almost completely with water, whereas weaker acids, such as acetic acid (and most of organic acids), react only slightly. The exact strength of an acid can be measured quantitatively by its *acidity constant.*

Acid-base interaction in water solution can be expressed by means of the equilibrium constant for its dissociation *(K_{eq})*, where HA represents any acid[1]:



$$HA + H_2O \rightleftharpoons A^- + H_3O^+ \qquad K_{eq} = \frac{[A^-]\,[H_3O^+]}{[HA]\,[H_2O]}$$

Since the concentration of water, $[H_2O]$, is nearly constant (~55.5 M), it is combined with *K* to give the acidity constant, K_a:



$$K_a = K_{eq}\,[H_2O] = \frac{[A^-]\,[H_3O^+]}{[HA]}$$

Strong inorganic acids have their equilibrium shifted to the right, thus increasing the concentration of H_3O+ and the value of K_a. They have K_a's in the range from 10^2 to 10^{10} Organic acids belong to weaker acids and have their equilibrium shifted to the left that results in smaller acidity constants (less than $10^{-4}$for carboxylic acids).

[1] Remember that square brackets denote molar concentration (M) of the enclosed species.

In order to avoid using numbers with negative exponents, acid strengths are often expressed as pK, which is equal to the negative logarithm of the acidity constant:

$^{pK}a = {}^{-\log K}a$.

The relationship between $K_a$and pK_a values signifies that:

A stronger acid has a lower pK_a, and a weaker acid has a higher pK_a.

The Bronstedt definition of acidity is extremely useful in organic and bioorganic chemistry because almost all organic compounds contain hydrogen and are therefore potential acids. Usually, organic acids are classified into:

- OH acids (carboxylic acids, alcohols, and phenols);
- SH acids (thiols);
- NH acids (amines, amides, some heterocycles, and ammonium salts);
- CH acids (hydrocarbons and their derivatives).

A part of a molecule that involves hydrogen together with an atom attached to it is called an *acidic site.*

We can compare the strength of different acids if we know their p$K_a$values. These values were determined for most important organic compounds. Table 5.1 lists p$K_a$values for selected representatives of different organic classes; some inorganic compounds are given for reference.

Problem 5.1. The *K* of nitric acid is about 44. Find the pK's of acetic acid and phenol from Table 5.1 and rank all the three compounds on their acidity strength.

If pK_a values are not available, we nevertheless are able to compare the strength of Bronstedt acids in terms of stability of their conjugate bases (anions) according to the objective regularity:

The more stable is an anion, the stronger is an acid.

The following factors influence the stability of conjugate bases:

- electronegativity and polarizability of the atom in the acidic site;
- delocalization of a negative charge due to the effect of substituents in a molecule;
- solvation effects.

These factors will be discussed below.

5.2.1. Electronegativity and Polarizability of an Atom

The greater electronegativity of an atom in the acidic site, the more stable conjugate base is observed within one row of the Periodic Table of the elements. This is accounted for higher ability of oxygen (in comparison with nitrogen and carbon, elements of the second row) to hold a negative charge. For this reason, alcohols are stronger as acids than amines, and alkanes show extremely low acidity, as shown in Table 5.2.

Table 5.1. The values of $pK_a$for selected Bronstedt acids

Name	Formula	Conjugate base	pK_a*
Organic acids			
OH acids			
Ethanol	C_2H_5OH	$C_2H_5O^-$	16.0
Phenol	C_6H_5OH	$C_6H_5O^-$	10.0
Acetic acid	CH_3COOH	CH_3COO^-	4.8
Benzoic acid	C_6H_5COOH	$C_6H_5COO^-$	4.2
Lactic acid	$CH_3CH(OH)COOH$	$CH_3CH(OH)COO^-$	3.9
Citric acid	$HOOCCH_2\underset{\displaystyle COOH}{\overset{\displaystyle |}{C}}H(OH)CH_2COOH$	$HOOCCH_2\underset{\displaystyle COO^-}{\overset{\displaystyle |}{C}}H(OH)CH_2COOH$	3.1
SH acids			
Ethanethiol	C_2H_5SH	$C_2H_5S^-$	12
Thiophenol	C_6H_5SH	$C_6H_5S^-$	8
Thioacetic acid	$CH_3C(O)SH$	$CH_3C(O)S^-$	3.3
NH acids			
Acetamide	$CH_3C(O)NH_2$	$CH_3C(O)NH^-$	25
CH acids			
Methane	CH_4	CH_3^-	40
Acetylene	$CH{\equiv}CH$	$CH{\equiv}C^-$	25
Acetone	$CH_3C(O)CH_3$	$CH_3C(O)CH_2^-$	20
Chloroform	Cl_3CH	Cl_3C^-	15.7
Inorganic acids			
Strong acids			
Hydroiodic acid	HI	I^-	–11
Hydrobromic acid	HBr	Br^-	–9
Hydrochloric acid	HCl	Cl^-	–7
Sulfuric acid	H_2SO_4	HSO_4^-	–3
Nitric acid	HNO_3	NO_3^-	–1.6
Weak acids			
Phosphoric acid	H_3PO_4	$H_2PO_4^-$	2.1
Hydrofluoric acid	HF	F^-	3.4
Carbonic acid	H_2CO_3	HCO_3^-	6.4
Hydrogen sulfide	H_2S	HS^-	7.0
Ammonium ion	NH_4^+	NH_3	9.2
Water	H_2O	HO^-	15.7
Ammonia**	NH_3	NH_2^-	36

* Approximate values for very strong acids ($pK_a < -2$) and very weak acids ($pK_a > 16$).

** Do not be surprised to see ammonia in the list. The NH_3 molecule contains hydrogen atoms and is, therefore, a potential acid.

Table 5.2. Acid strength of compounds of different classes

Class of compounds	Formula	Element's electronegativity in the acidic site	pK_a
Alkanes	RCH_3	2.5	~44
Amines	RNH_2	3.0	~30
Alcohols	ROH	3.5	16–18
Thiols	RSH	2.6	11–12
Phenols	ArOH	3.5	~10
Carboxylic acids	RCOOH	3.5	~5

Another stabilizing factor is the *polarizability* (opposed to polarity) of an element in the acidic site. This term means the ability of the electrons to respond to a changing electric field, as a result of its interaction with solvent or with other polar reagents. Relative polarizability increases *within one group* of the Periodic Table from top to bottom because a larger atom holds electrons more loosely than a smaller atom with tightly held electrons. For example, iodine whose electrons are far from the nucleus is much more polarizable than fluorine whose electrons are close to the nucleus. Thus, polarizability of halogens increases in the following order: F < Cl < Br < I. Stability of the corresponding halide ions increases in the same order.

Problem 5.2. Fluorine is much more electronegative than iodine, yet hydroiodic acid is much stronger than hydrofluoric acid (by a factor of approximately 10^{14}). Explain this difference.

Similarly, the size of the sulfur atom is larger than that of oxygen. Therefore, the negative charge in a thiolate ion, RS^-, is delocalized more effectively in comparison with an alkoxide ion, RO^-. Indeed, thiols show evidently higher acidity than alcohols (Tables 6.1 and 6.2).

The difference in acidity of thiols and alcohols is displayed in the reactions with alkali. Thiols do react to give salts, whereas alcohols practically do not react:

$$\underset{\text{thiol}}{RSH} + NaOH \longrightarrow \underset{\text{sodium thiolate}}{RS^-Na^+} + H_2O$$

$$ROH + NaOH \longleftrightharpoons RO^-Na^+ + H_2O$$

Alkoxides (alcohol salts) can be obtained only in the reaction of alcohols with active metals or with extremely strong bases, such as sodium hydride, NaH, or sodium amide, $NaNH_2$. For example:

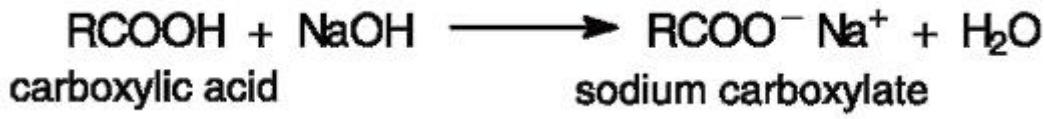
$$\underset{\text{alcohol}}{ROH} + NaH\ (\text{or } Na) \longrightarrow \underset{\text{sodium alkoxide}}{RO^-Na^+} + H_2\uparrow$$

Metal salts of alcohols are themselves strong bases used frequently as reagents in organic chemistry.

Problem 5.3. Show the order of acidity increase for the following compounds: (a) 1-butanol, (b) ethanethiol, (c) ethanol; (d) ethylamine. Explain the reasons for your choice. Write an equation for the salt formation for the most acidic compound.

5.2.2. Delocalization of a Charge in an Anion

As we have seen, there are three types of OH acids, namely carboxylic acids, alcohols, and phenols. As follows from their general name, carboxylic acids are acidic compounds. They therefore react with bases such as metal hydroxides to give metal carboxylate salts:



Unsubstituted carboxylic acids are much weaker than mineral acids; nevertheless they are much stronger than alcohols and phenols (compare the p$K_a$values in Tables 5.1 and 5.2). The question arises, why carboxylic acids are the most acidic organic compounds even though all the three types of the OH acids contain the same acidic site?

To answer this question let us consider, first of all, at the relative stability of an alkoxide ion, RO^-, a phenoxide ion, ArO^-, and a carboxylate ion, $RCOO^-$. The former is an oxygen anion in which the negative charge is localized on a single electronegative atom.

Phenols are stronger acids than alcohols by a factor of at least 10^5. The difference is to be accounted for by higher stability of the phenoxide ion that, in turn, is a result of a conjugation (Fig. 5.1, a). Stability of the phenoxide ion is also manifested by resonance theory. We can write several resonance structures for the anion, showing delocalization of the negative charge over the benzene ring (Fig. 5.1, b). At the same time, no analogous resonance ions are possible for an alkoxide ion.

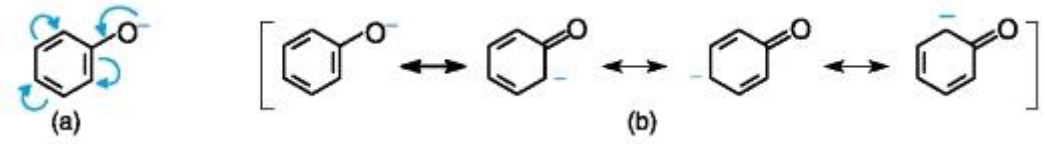


Figure 5.1. Representation of charge delocalization in the phenoxide ion: ρ,π conjugation (a), and resonance hybrid (b).

Unlike alcohols, phenols react with alkalis to give water-soluble salts, phenoxides:

The carboxylate ion is also an oxygen anion, but the negative charge is delocalized over both oxygen atoms through conjugation (Fig. 5.2, a). This results in stabilization of the anion. In resonance terms, the carboxylate ion is a stabilized resonance hybrid of two *equivalent* structures (Fig. 5.2, b) neither of which contains a localized charge (Fig. 5.2, *c).*

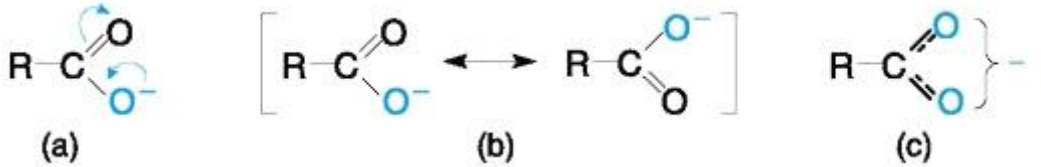
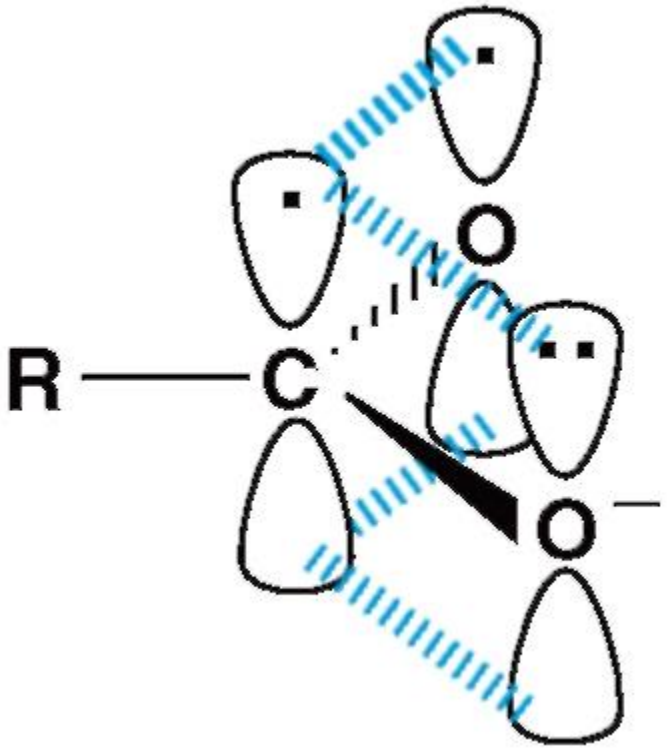


Figure 5.2. Representation of charge delocalization in the carboxylate ion: ρ,π conjugation (a), resonance hybrid (b), and equal charges on both oxygens (c).

Equivalence of both carbon-oxygen bonds becomes clear from an orbital picture of the carboxylate ion shown in Fig. 5.3. Four *p* electrons of a conjugated system are delocalized throughout three *p* orbitals. Consequently, the *p* orbital on the carboxylate carbon overlaps equally well with *p* orbitals of both oxygen atoms, thus making both carbon-oxygen bonds intermediate between single and double

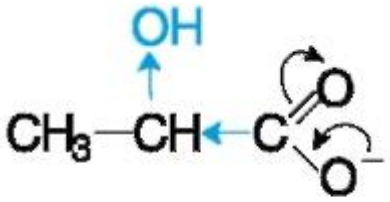


bonds (see also Fig. 5.2, c).

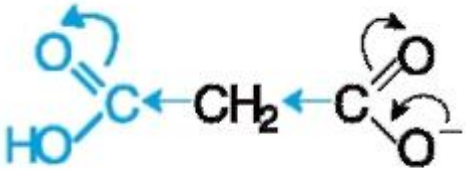
Figure 5.3. Orbital overlap in the carboxylate ion.

Effects of substituents on acidity. Other factors that stabilize a conjugate base (alkoxide, phenoxide, or carboxylate ion) result in increased acidity. It might be an

electron-withdrawing group disposed near an acidic site. Such groups in aliphatic series are, for example, halogens, the hydroxyl group, and an additional carboxyl group. Let us consider acidity of substituted carboxylic acids.
Electron-withdrawing substituents shift inductively electron density from the anionic site, delocalizing the negative charge on the carboxylate ion, stabilizing it, and increasing acidity. Lactic acid (pK_a 3.9) and malonic acid (pK_a 2.9) exemplify this consideration (compare these values with the value of 4.8 for unsubstituted propionic acid):



lactate ion

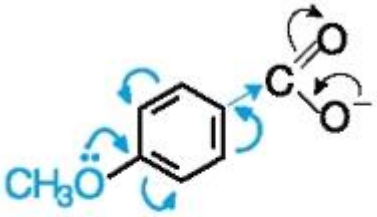


malonate ion

In contrast to the substituents mentioned just before, electron-donating groups should have the opposite effect by localizing the negative charge, destabilizing the carboxylate ion, and decreasing acidity. It is interesting that there are no carboxylic acids having electron-donating groups attached to an aliphatic chain.

Electron-donating substituents, however, can be present in aromatic carboxylic acids (recall positive inductive or mesomeric effect of alkyl, amino, hydroxyl, and alkoxyl groups attached to the benzene ring). But these substituents influence the negative charge of a carboxylate ion only inductively. Thus, a slight decrease of acidity is observed for 4-methylbenzoic and 4-methoxybenzoic acids (compare the values given below with the value of 4.2 for benzoic acid itself).

4-methylbenzoate ion,
pK_a 4.4 for conjugate acid

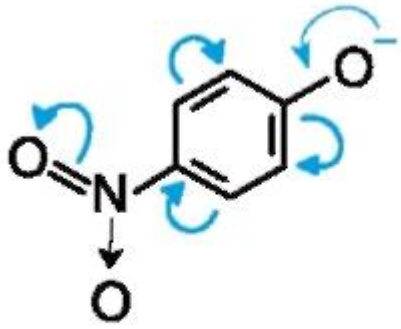


4-methoxybenzoate ion,
pK_a 4.5 for conjugate acid

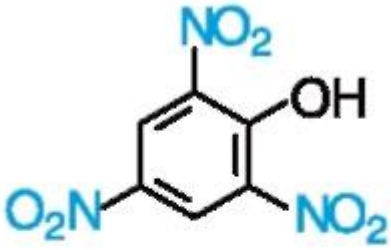
For an obvious reason, electron-withdrawing groups introduced in the benzene ring increase the acidity of substituted benzoic acids. Here we also notice a moderate acidity increase even though a substituent is a very strong electron-withdrawing nitro group. This is supported by the acidity data for 4-nitrobenzoic acid (pK_a3.4) and 4-chlorobenzoic acid (pK_a4.0).

Inductive and mesomeric effects of substituents are also important in determining acidity of substituted phenols. Very similar regularities are to be observed in this case. As a rule, phenols with an electron-withdrawing substituent are more acidic than phenol itself, and phenols with electron-donating substituents are less acidic

by the reasons that have just been discussed for substituted carboxylic acids. For example, 4-nitrophenol (pK_a 7.1) is a much stronger acid than phenol (pK_a 10.0). The nitro group acts in two ways: it stabilizes the anion not only through the inductive effect, but also through a long chain of conjugation as shown below. Additional electron-withdrawing substituents further increase acidity. Thus, 2,4,6-trinitrophenol, commonly called *picric acid*, is much stronger than most carboxylic acids.



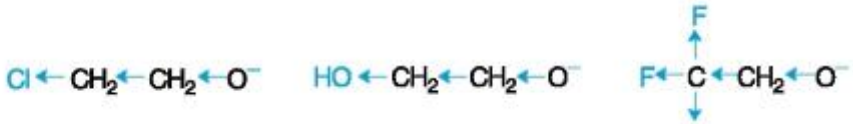
4-nitrophenoxide ion



2,4,6-trinitrophenol (picric acid), pK_a 0.8

Problem 5.4. Compare the acidic sites in the salicylic (2-hydroxybenzoic) acid molecule. Write an equation for its reaction with: (a) equimolar amount of sodium hydroxide; (b) excess sodium hydroxide.

In the alcohol series, electron-withdrawing substituents make a compound more acidic owing to the fact that the negative charge in a substituted alkoxide ion is spread out over a larger area. Usually, electron-withdrawing substituents increase acidity of alcohols slightly, by a factor within 1-2. Only several strongly electronegative fluorine atoms (as in a CF_3 group) increase acidity obviously. The p$K_a$values for 2-chloroethanol, ethylene glycol, and 2,2,2-trif luoroethanol (presented below) support this reasoning:



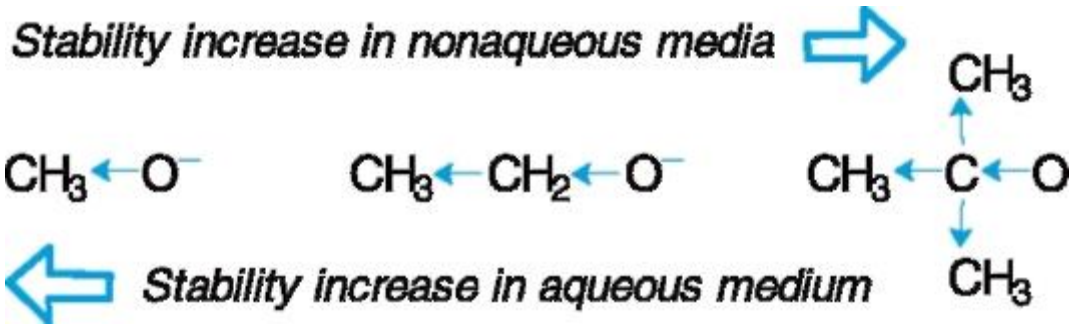
pK_a's of conjugate acids:	14.3	14.8	12.4

In general, alcohols are the weakest OH acids.

5.2.3. Solvation Effects

Solvation of an anion is a further factor that influences stability of the conjugate base. Water is a solvent in biological systems; therefore an effect of hydration (interaction of a substance with water) must be taken into consideration.

Lower alcohols, such as methanol and ethanol, are similar to water in acidity (Table 5.1), while tert-butyl alcohol, $(CH_3)_3COH$, is slightly less acidic *(pK_a*~18). It is surprising, because the ferf-butoxide ion is more stable than the ethoxide and methoxide ions, when we *only* compare the inductive effect of alkyl groups:



In fact, this is true for solutions in nonpolar solvents. In an aqueous solution, water surrounds the oxygen atom, thus helping to stabilize the anion. Small anions of unhindered lower alcohols (Fig. 5.4, a) are better hydrated than bulky alkoxide ions of tertiary alcohols (Fig. 5.4, b).

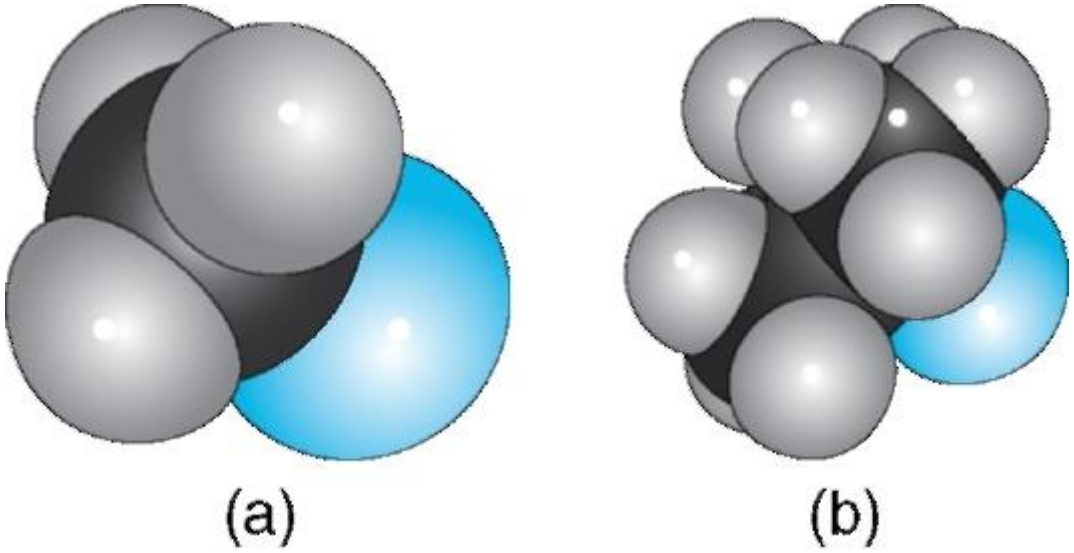


Figure 5.4. Space-filling models: a methoxide ion (a) and ferf-butoxide ion (b).

5.3. BASES

According to the Bronstedt-Lowry definition given earlier, any anion or neutral compound, containing a heteroatom with a lone-pair of electrons, can act as a base. This is the main group of bases called *n-bases* (having nonshared electrons). They are further classified into the following types, depending on the nature of heteroatom, which represents the *basic sife* (only neutral compounds are listed here and in the following Tables 5.3 and 5.4):

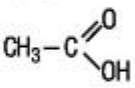
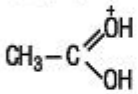
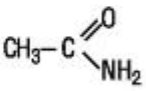
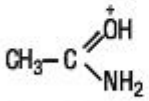
- N-bases (amines and many heterocycles);
- O-bases (alcohols, phenols, ethers, and compounds with the >C=O group);
- S-bases (thiols and sulfides).

A less significant group of bases constitutes *π-bases*, in which electrons of the localized π bond or π electrons of the conjugated system can accept a proton to form non-covalent complexes. Table 5.4 lists some representatives of n-bases and their strength, which will be discussed below.

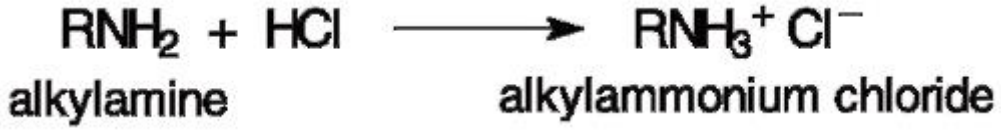
Table 5.3. Base strength of compounds of different classes

Class of compounds	Formula	Name of onium salts produced	pK_{BH^+}
Aliphatic amines	RNH_2	Ammonium	10–11
Aromatic amines	$ArNH_2$	Ammonium	4–5
Alcohols	ROH	Oxonium	from –2 to –5
Phenols	ArOH	Oxonium	–6
Ethers	ROR	Oxonium	from –3 to –6
Thiols	RSH	Sulfonium	–7

Table 5.4. The values of pK_{BH^+} for selected bases

Name	Formula	Conjugate acid	pK_{BH^+}
N-Bases			
Ammonia	NH_3	NH_4^+	9.2
Ethylamine	$C_2H_5NH_2$	$C_2H_5NH_3^+$	10.7
Diethylamine	$(C_2H_5)_2NH$	$(C_2H_5)_2NH_2^+$	10.9
Triethylamine	$(C_2H_5)_3N$	$(C_2H_5)_3NH^+$	10.9
Aniline	$C_6H_5NH_2$	$C_6H_5NH_3^+$	4.6
Diphenylamine	$(C_6H_5)_2NH$	$(C_6H_5)_2NH_2^+$	0.8
O-Bases			
Water	H_2O	H_3O^+	–1.7
Ethanol	C_2H_5OH	$C_2H_5OH_2^+$	–2.5
Diethyl ether	$(C_2H_5)_2O$	$(C_2H_5)_2OH^+$	–3.6
Phenol	C_6H_5OH	$C_6H_5OH_2^+$	–6.7
Acetic acid			–6
Acetamide			–0.5
Acetone	$(CH_3)_2C{=}O$	$(CH_3)_2C{=}OH^+$	–7
Urea	$(NH_2)_2C{=}O$	$(NH_2)_2C{=}OH^+$	0.1
S-Bases			
Methanethiol	CH_3SH	$CH_3SH_2^+$	–6.7
Dimethyl sulfide	$(CH_3)_2S$	$(CH_3)_2SH^+$	–5.3

Organic N-bases react with acids to yield stable ammonium salts, which, usually, are soluble in water. Many drugs are used as salts because of their better solubility.

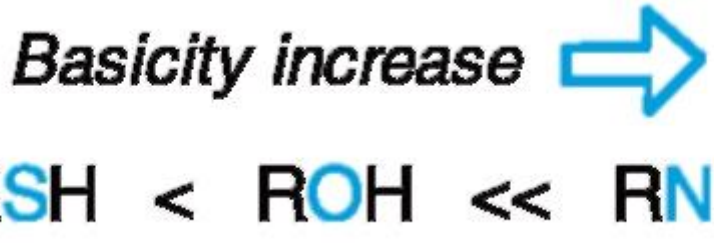


Bases differ too much in their proton-accepting ability. The strength of bases depends on electronegativity and polarizability of the atom, which represents the basic site. Within one row of the Periodic table, an atom with higher

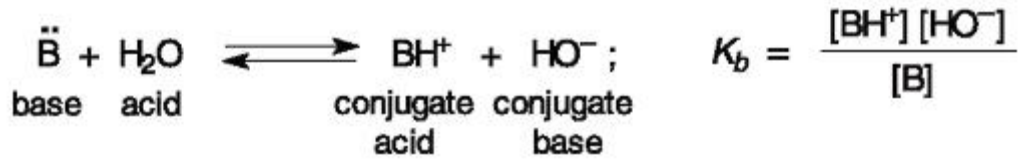
electronegativity is less capable of proton acceptance; therefore amines are more basic than alcohols (just as ammonia is a stronger base than water).

When we compare basicity of alcohols (as O-bases) and thiols (as S-bases), the difference in polarizability of oxygen and sulfur should be taken into consideration. The size of the sulfur atom is larger than that of the oxygen atom; therefore electron density is less on sulfur. For this reason thiols, as well as organic sulfides, are not able to form a strong bond with a proton, i.e. they are weaker bases than alcohols. Thus we can say, that the strength of bases with the same or similar R substituents increases in the following way[1]:

[1] This is in accordance with a mnemonics *SON* (sulfur, oxygen, nitrogen).



Acid-base interaction in aqueous medium can be used for quantitative determination of the base strength, where $K_b$is the basicity constant:



A more convenient way for basicity evaluation is to consider the acidity of the corresponding conjugate acid. For example, a primary amine RNH_2 and its ammonium ion RNH_3^+ are related to each other as a base and its conjugate acid, according to the following equilibrium:

The acidity constant of the conjugate acid designated in this case by $K_{BH}+$ is expressed as follows:

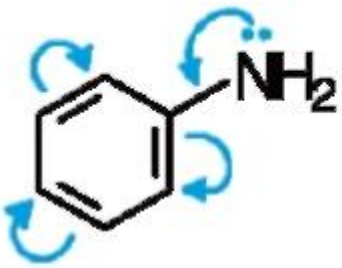
$$K_{BH^+} = \frac{[B]\,[H_3O^+]}{[BH^+]} \quad \text{and} \quad pK_{BH^+} = -\log K_{BH^+}$$

This constant is designated more often by pK_a. It should be differentiated whether it is really related to acidity of a compound or to acidity of its conjugate base.

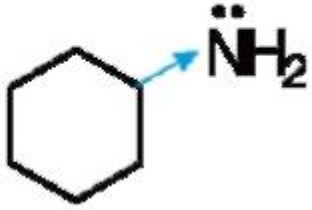
Table 5.3 and previous general consideration clearly show superior basicity of Nbases (with few exceptions) to O- and S-bases. Aliphatic amines are the strongest

of the N-bases, somewhat stronger than ammonia. They have K_{BH}+ values in the narrow range (10-11), regardless of their structure.

Aromatic amines, or arylamines, are less basic than aliphatic, because the nitrogen lone pair of electrons is delocalized by orbital overlap with the aromatic π-electron system through conjugation. They are, therefore, less available for proton acceptance. Cyclohexylamine, for example, is 10^6 times stronger as a base than



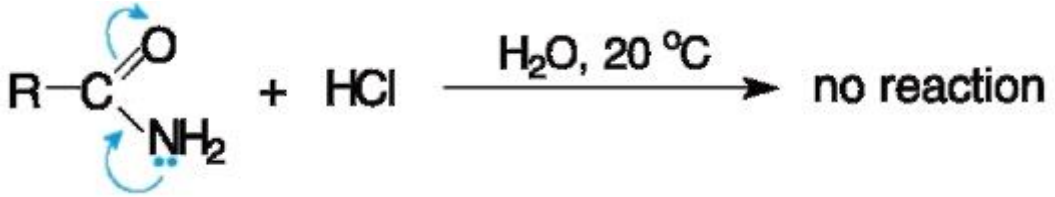
aniline,
pK_{BH}+ 4.6



cyclohexylamine,
pK_{BH}+ 10.6

aniline.

Amides, $RC(O)NH_2$, are much weaker bases on a similar reason. Here, the nitrogen lone pair of electrons is conjugated with the C=O double bond. Extremely low basicity of amides is confirmed by the fact that water insoluble amides do not dissolve in aqueous solutions of strong mineral acids.



Substituents in the alkyl or aryl part of an amine molecule affect basicity. Electronwithdrawing groups increase it; in other words, they decrease acidity of the conjugate acid. Conversely, electron-donating groups increase basicity.

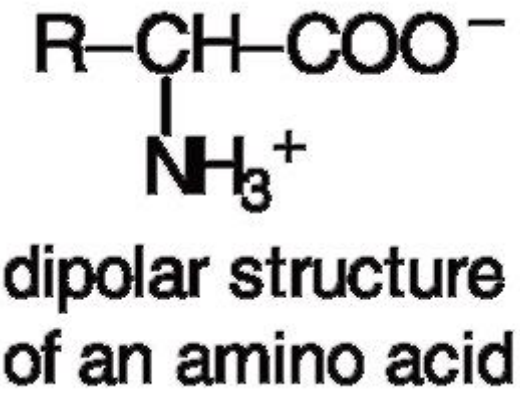
Problem 5.5. Show the order of basicity increase for the following compounds: (a) ammonia; (b) ethanethiol; (c) 1-propanol; (d) ethylmethylamine. Explain the reasons for your choice. Write an equation for the salt formation for the most basic compound.

Basic properties of heterocycles will be considered in Chapter 16.

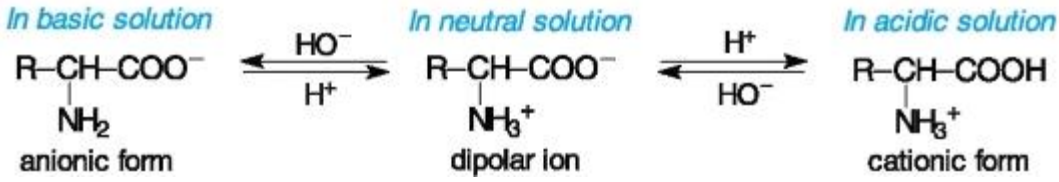
5.4. ACIDIC AND BASIC SITES IN A MOLECULE

Many organic compounds are *amphoteric.* That is, they are capable of functioning either as an acid or as a base, depending on the circumstances. Typical examples of these compounds are natural (or protein) amino acids. Their structure is usually written as $RCH(NH_2)COOH$ with two independent functional groups. The presence of an acidic (COOH) and a basic (NH_2) sites in the same molecule results in an acid-base interaction to produce a salt-like compound. Therefore, the real

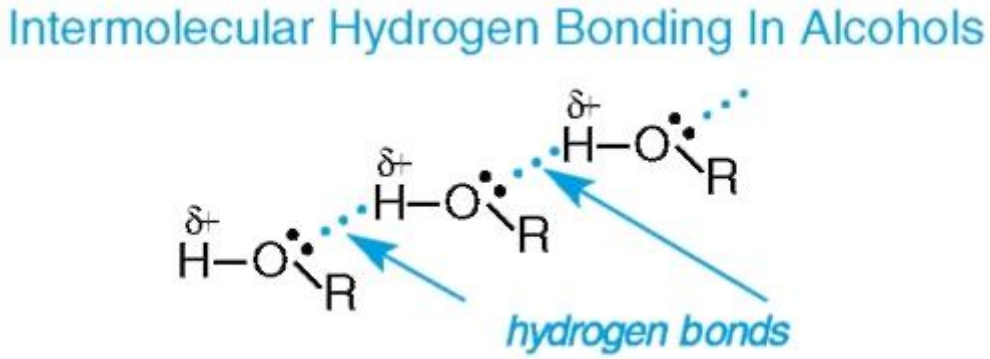
structure of amino acids in neutral solution and in crystalline state is a *dipolar ion* structure, sometimes called *zwitterion* (from the German *Zwitter*- hybrid).



The amino acids can be protonated in acidic solution, thus converting into a cationic (ammonium) form. In alkaline medium they exist as carboxylate ions with a free amino group:



Molecules containing a weak acidic and weak basic site in their structure can form an intramolecular or intermolecular *hydrogen bond.* Water is the simplest and most known amphoteric compound that forms intermolecular hydrogen bonds (shown usually by a dotted line).



Additional Problems

5.6. Show the order of acidity increase for the following compounds:

(a) 1,2-ethanediol;

(b) ethanol;

(c) 4-nitrophenol;

(d) phenol.

Explain the reasons for your choice. Write an equation for the salt formation using the most acidic compound.

5.7. Show the order of basicity increase for the following compounds:

(a) ammonia;

(b) aniline;

(c) cyclohexylamine;

(d) phenol.

Explain the reasons for your choice. Write an equation for the salt formation using the most basic compound.

5.8. Propose a chemical method for isolating phenol and aniline from their mixture in benzene solution.

5.9. Propose a chemical procedure for isolating 1-butanol and 1-butanethiol from the mixture of the two compounds. Notice that both compounds are liquids insoluble in water.

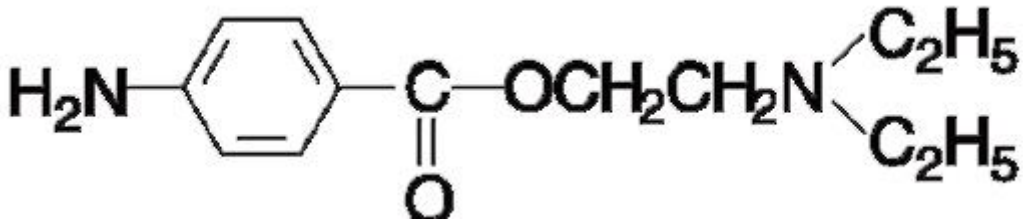
5.10. Explain the reason of the greater pK_a value for p-hydroxybenzoic acid (4.5) compared with that for benzoic acid (4.2).

Try to explain, why *o*-hydroxybenzoic (salicylic) acid is a stronger acid (pK_a 3.0) than benzoic acid. *(Hint:*Consider all factors that stabilize a conjugate base.)

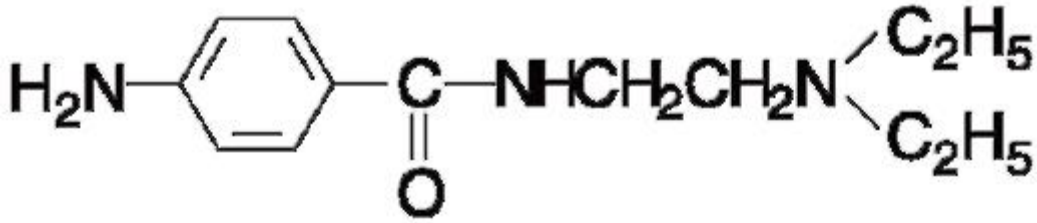
5.11. The pK_a values for phenol and 4-nitrophenol are 7.1 and 10.0, whereas the p$K_a$values for benzoic and 4-nitrobenzoic acids differ to a lesser extent (3.4 and 4.2).

Explain these differences, but first refer the p$K_a$values to each compound in the absence of reference data.

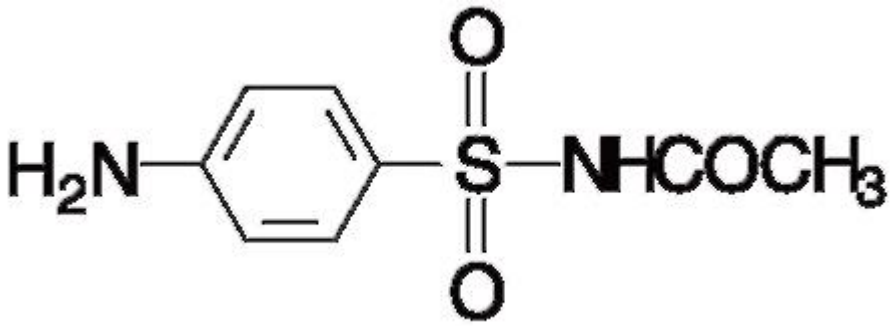
5.12. Discuss all the basicity sites in the Procaine molecule (see the structure). Draw the structure of Procaine hydrochloride (also called Novocain).



5.13. Procainamide, or Novocainamide (see the structure), is used in medicine in the form of hydrochloric acid salt. Discuss all the basicity sites in the Procainamide molecule and draw the structure of its hydrochloride salt.



5.14. The antibacterial drug (see the structure below) is used as sodium salt, Sulfacyl soluble. Discuss all the acidity sites in the molecule and draw the structure of its sodium salt.



Chapter 6. HYDROCARBONS

We proceed to a study of the chemistry of different classes of organic compounds beginning in this chapter from the simplest ones, hydrocarbons. As it follows from the name, hydrocarbons contain carbon and hydrogen only. They comprise a big series of natural and synthetic compounds. Many hydrocarbons are constituents of petroleum and natural gas. Methane (common name - *marsh gas)* is the end product of the anaerobic decomposition of organic matter.

6.1. CLASSIFICATION

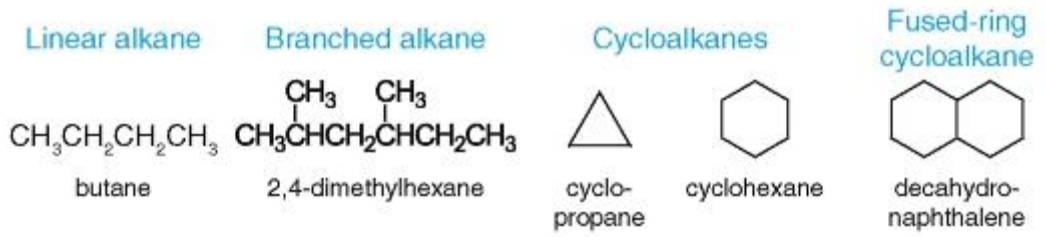
There are three main types of hydrocarbons, depending on the nature of carboncarbon bonds in a molecule:

- Saturated hydrocarbons that contain single C-C bonds only;
- Unsaturated hydrocarbons containing multiple carbon-carbon bonds: a double bond (or bonds), a triple bond (or bonds), or their combination. Single C-C bonds may be also present, of course;
- Aromatic hydrocarbons that are related in structure to benzene (Sec. 3.2). The first two types are sometimes referred to as *aliphatic* compounds (from the Greek *aleiphas* - fat), i. e. compounds of a nonaromatic series.

Hydrocarbons vary considerably in their chemical properties that depend on the nature of the carbon-carbon bonds. This is true also for compounds with various functional groups because hydrocarbons represent a framework of organic molecules.

6.2. SATURATED HYDROCARBONS

These hydrocarbons can be subdivided into two groups,
namely, *alkanes,* compounds with an open carbon chain (linear or branched), and *cycloalkanes,* compounds with a cyclic carbon skeleton. The rightmost structure in the examples below (decaline) represents a fused-ring system of two cyclohexane rings jointed together.



6.2.1. Conformational Isomerism

Conformations of a molecule of definite configuration are various spatial arrange

ments of its atoms that differ only after rotation about C-C single bonds.

This notion is important not only to relative small molecules, such as simple acyclic or cyclic compounds, but also to biologically significant polymers, such as proteins, nucleic acids, and polysaccharides. This chapter is concerned with only simple examples of conformational isomerism.

Conformations of Ethane. Ethane is the simplest molecule where free rotation about the C-C bond is possible at room temperature. There are two ways for representation of conformations. The first one uses perspective formulas resembling «saw-horse» (Fig. 6.1, a, c). The second one uses *Newman projections* (Fig. 6.1, b, *d),* which are easier to draw.

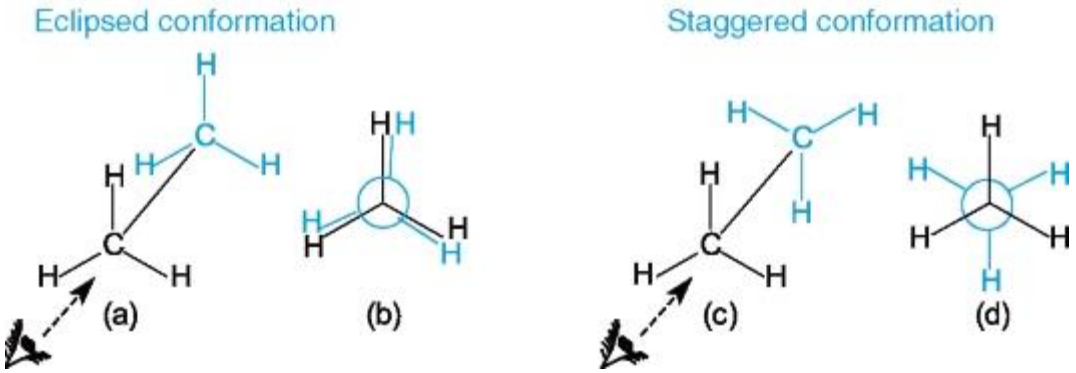


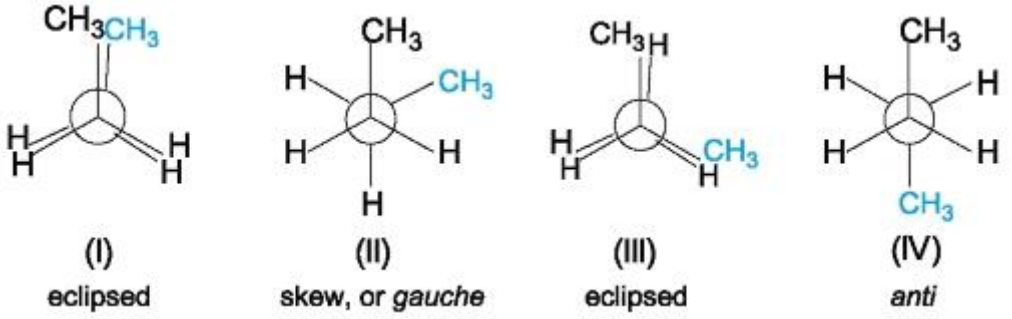
Figure 6.1. Perspective formulas (a) and Newman projections (b) of the eclipsed and staggered conformations of ethane.

To prepare a Newman projection a molecule is viewed along a bond between two carbons; a circle represents these carbons. The lines going to the centre of the circle represent bonds attached to the front carbon, and the lines going to the edge of the circle represent bonds on the rear carbon (coloured in Fig. 6.1). When two such bonds would be coincident in the projection, they are drawn at a small angle to each other.

An infinite number of conformations could result from rotation of one CH_3 group relative to another. These conformations slightly differ in stability that depends on mutual disposition of the CH_3 groups. The two extreme forms are called the *eclipsed* and *staggered* conformations.

In the eclipsed conformation the hydrogen atoms attached to each carbon are in direct opposition to each other when viewed along the C-C bond, whereas in the staggered conformation these hydrogens are perfectly staggered, i. e. each C-H bond on one carbon bisects the H-C-H angle on the other carbon.
The staggered conformation is the most stable of ethane conformations that is usually explained by repulsive interaction between electron clouds in the C-H bonds, so called *torsional strain.*

Conformations of Butane. The more complicated is the structure, the more complex is the conformational situation. For example, four extreme conformations are possible for the butane molecule, if we consider rotation about the C-2-C-3 bond (in addition to conformations which arise after rotation about the terminal C-1-C-2 and C-3-C-4 bonds). Here are two eclipsed conformations, (I) and (III), and two staggered conformations which are called *skew*, or *gauche* (II), and *anti* (IV) conformations.



The eclipsed conformations possess maximum potential energy. They have not only torsional strain but also van der Waals repulsions arising from the eclipsed pairs CH_3 and H in the conformation (III) or from two CH_3 groups in the conformation (I). This type of interaction when two groups are forced to be closer than their atomic radii allow is called *steric strain.* It is observed in the conformation (II) as well because the CH_3 groups are close enough to each other for repulsion. Obviously, the *anti* conformation is the most stable one because both torsional and steric strains are not revealed.

The same principles just considered for butane can be applied to all alkanes. The most favoured arrangement for any alkane is that, in which all C-C bonds occupy the opposite positions, i. e. are in the*anti* conformations. Thus long-chain molecules have a tendency to take a zigzag shape as it is shown on the example of butane (Fig. 6.2).

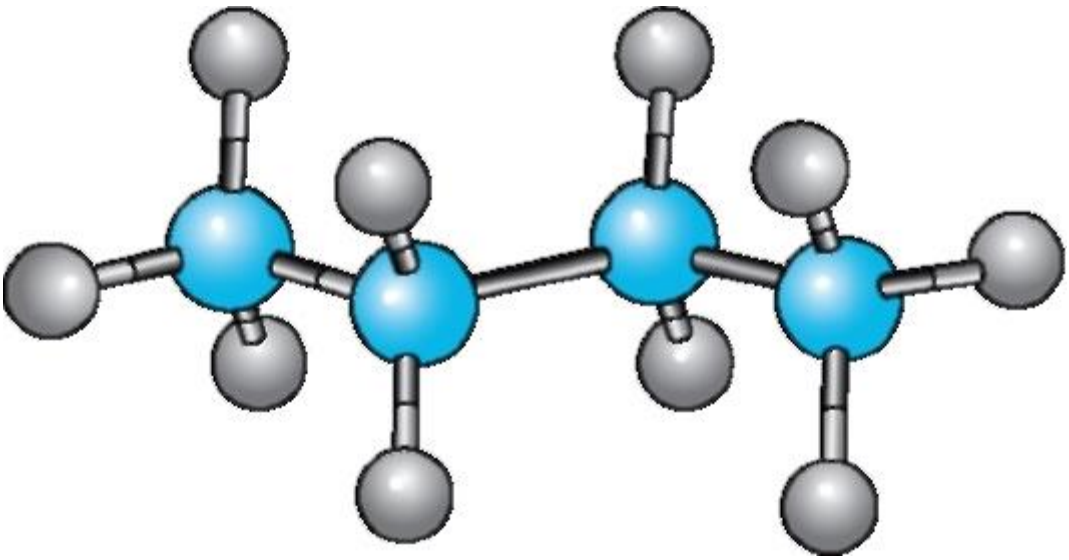


Figure 6.2. A ball-and-stick model of butane in a zigzag conformation.

In conclusion, it is important to stress that energy barriers caused by rotation are not large enough to prevent conformational equilibrium at room and much lower temperatures.

Conformational isomers cannot be isolated at normal temperatures.

Different energies of conformers mean only that at any instant, for example, about 70% of the butane molecules exist in the *anti* conformation and about 30% in the *gauche* conformation, as well as 99% of the ethane molecules possess the staggered conformation and 1% the eclipsed one.

Problem 6.1. Draw the Newman projections for the staggered and eclipsed conformations of the propane molecule.

Conformations of Cycloalkanes. There are some peculiarities in conformations of cyclic compounds in comparison with conformations of open-chain molecules. The first distinction is that free rotation about C-C bonds is restricted in cycloalkanes because of their relatively rigid structure. In fact, unlimited rotation is impossible without cleavage of the cycle. The second distinction consists in an additional factor that affects the shape and total strain energy of some rings. This is a concept of *angle strain*proposed by the German chemist A. Baeyer in 1885.

The Baeyer strain hypothesis. According to Baeyer's suggestion, each cycloalkane represents a planar regular polygon, a triangle for cyclopropane, a square for cyclobutane, a pentagon for cyclopentane, a hexagon for cyclohexane, and so on. Angle strain was defined as the resistance of a bond angle to deviation from the tetrahedral angle (near to value of 109°). From this point of view, cyclopropane, with a bond angle compression of 49° (109° - 60°) as it is shown in Fig. 6.3, should have maximum angle strain and must therefore be highly reactive. Cyclopentane was predicted to be the most stable of cycloalkanes because it is

nearly strain-free. A definite bond angle expansion (109° - 120° = -11°) could be observed in cyclohexane, thus making it less stable than cyclopentane. Experimental determination of strain energy showed later the highest stability of cyclohexane. A weak point of the Baeyer's hypothesis was an assumption of planar geometry of all rings. Actually, most cycloalkanes (except for cyclopropane) are not flat; they acquire puckered shapes (conformations) in which bond angles approximate to tetrahedral ones. Nevertheless, the concept of angle strain remains very useful for explaining reactivity of small rings, cyclopropane and cyclobutane.

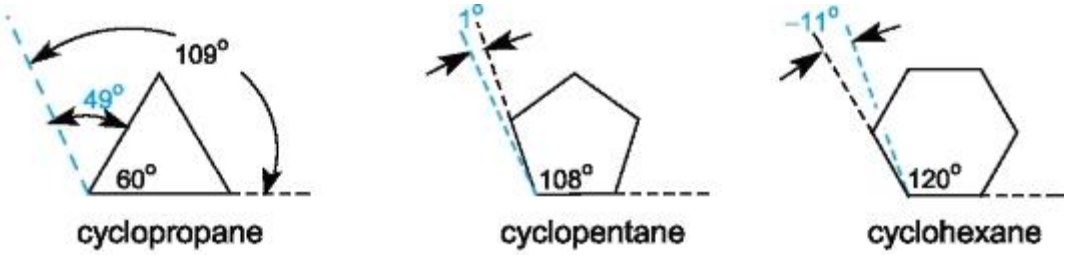


Figure 6.3. Bond angles and angle strains (shown in colour) in cycloalkanes.

Compounds containing six-membered rings are the most important of all cyclic saturated compounds due to their wide occurrence in nature. They have been studied in great detail, and main attention will be paid in this section to conformations of cyclohexane.

Cyclohexane. Cyclohexane ring is not flat as Baeyer supposed it to be. The most stable (and therefore favored) conformation of cyclohexane is the *chair* conformation[1] (Fig. 6.4) devoid of all strains. When viewed along the C-C bond on any side (for example, along the C-1-C-6 and C-3-C-4 bonds, as it is shown in Fig. 6.4, *b)* all hydrogens are seen to be staggered and the atoms C-2 and C-5 of the CH_2groups occupy *gauche* positions (Fig. 6.4, c). Thus the chair conformation is free of torsional and steric strains. Furthermore, in this nonplanar structure the C-C-C angles are 110.8° each, close to the ideal tetrahedral value, and are free of angle strain.

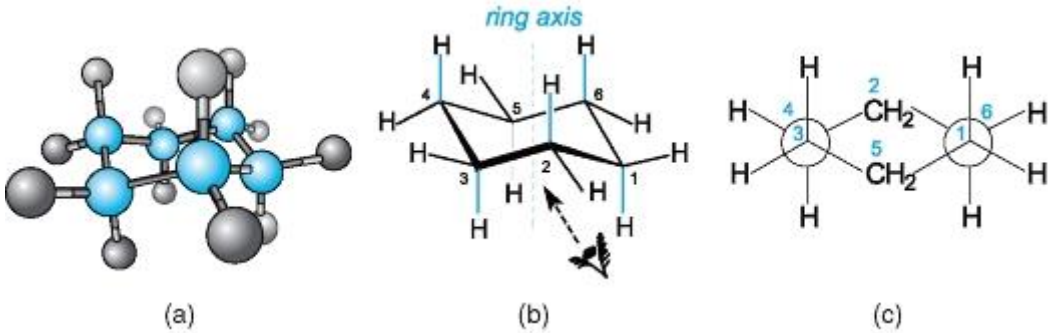


Figure 6.4. Representations of the chair conformation of cyclohexane: a ball-and-stick model (a); a line drawing (b); and a Newman projection (c).

[1] The term originates from similarity to a lounge chair, rather than an ordinary chair.

Look carefully at the arrangement of hydrogen atoms in the chair conformation of cyclohexane. Six C-H bonds, one at each carbon, are parallel to the vertical ring

axis (or perpendicular to the average plane of the ring). These bonds and hydrogens are called *axial* bonds and hydrogens. They are coloured in the structure (I) (Fig. 6.5). Note there are three axial hydrogens on each side of the ring and their arrangement (up or down) alternates from one carbon to the next. Six other bonds and hydrogens that lie more or less in the rough plane of the ring are called *equatorial*, by analogy with the equator of the Earth (see Fig. 6.5).

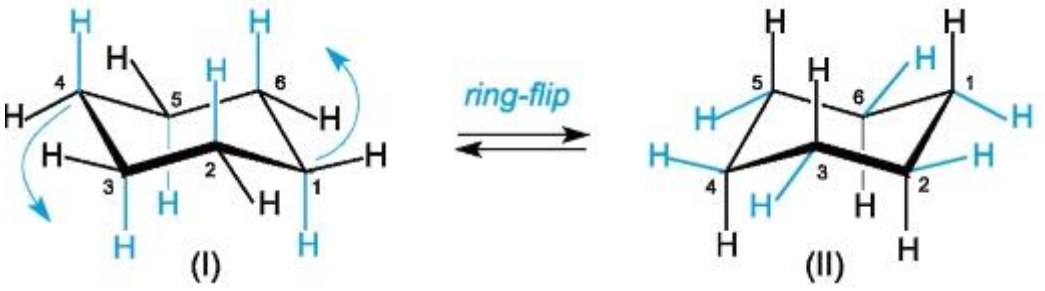


Figure 6.5. Interconversion of the chair conformations of cyclohexane. Axial hydrogens in (I) and equatorial hydrogens in (II) are shown in colour.

Cyclohexane ring is conformationally mobile owing to partial rotation about all C-C bonds. Two equivalent chair conformations can easily interconvert as it is shown in Fig. 6.5 by moving the atom C-1 up and the atom C-4 down in the conformation (I). This interconversion, usually called a *ring-flip* (or ring inversion), results in that all axial positions in one chair form become equatorial in the ring-flipped form, and *vice versa.*

How to draw a chair conformation? Here are some recommendations for proper drawing chair conformations of cyclohexane that involve the following steps:

1. Draw two parallel lines, slightly lifted to the right[1] and shifted from each other. The ends of the lines represent four carbons lying in a plane. This is a «chair-bottom».

2. Place the topmost carbon above and to the left of the plane. Draw two bonds to connect this carbon to the nearest carbons of the plane. So a «back» of the chair is added.

3. Place the bottommost carbon below and to the right of the plane and connect it to the remaining carbons of the plane. Thus construction of the chair is finished. Note that each pair of the opposite bonds has to be parallel.

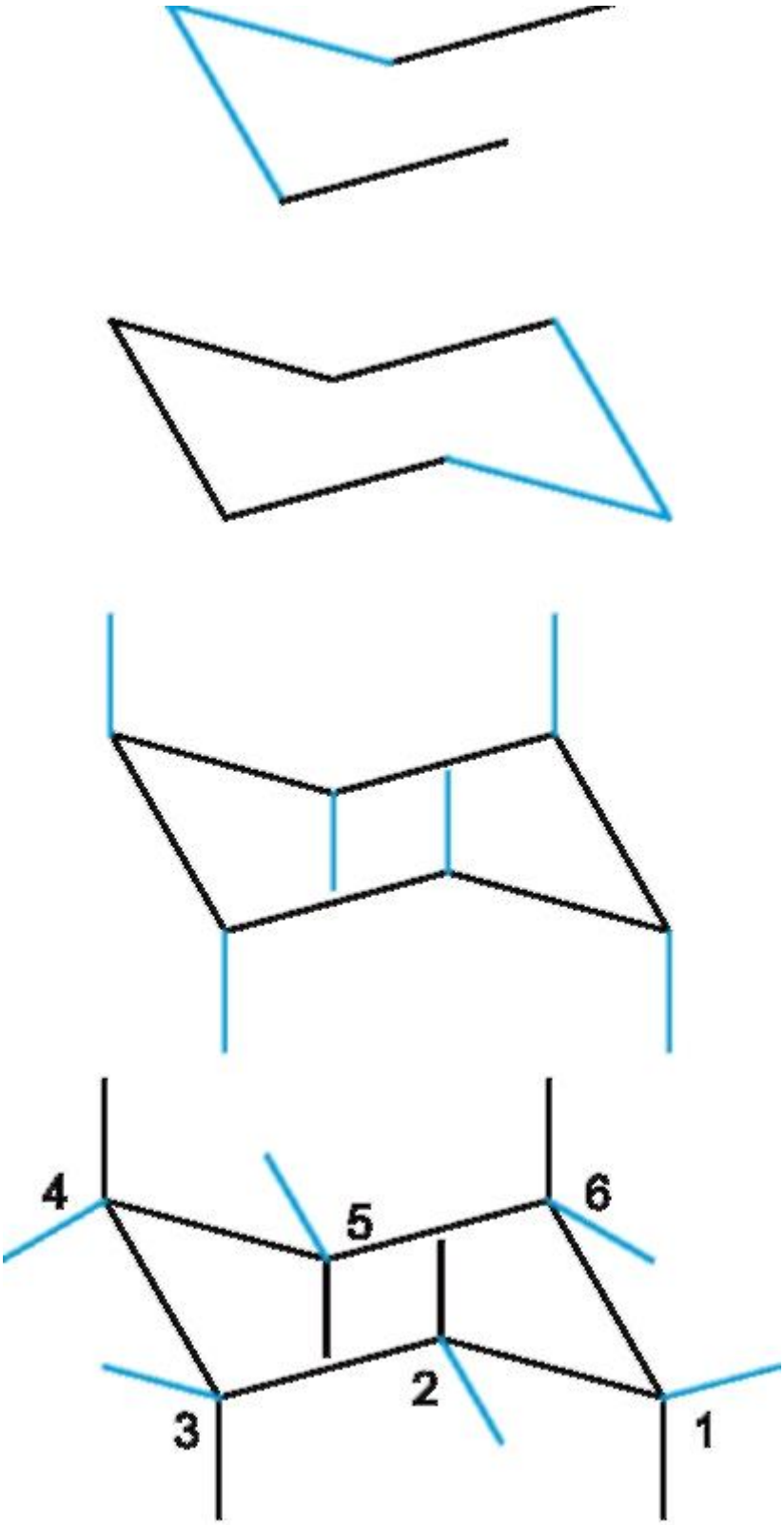
4. Draw six vertical axial bonds, one on each carbon, in the alternating up-down directions. Note that at the topmost carbons the axial bonds have to be directed «up» and at the bottommost carbons - «down».

5. Draw six equatorial bonds in such a way that each bond has to be parallel to the ring bond next to the nearest C-C bond and directed away. For example, the equatorial bond at the atom C-1 has to be parallel to the C-2-C-3 bond (as well as to the C-5-C-6 bond).

The second advice is to use a squared paper. A picture of the chair may be well inserted in a rectangle with size 3x6 squares as shown. Axial and equatorial bonds should be drawn in a usual manner.

Another possible conformation of cyclohexane is the *boat* conformation (Fig. 6.6, *a)*. Like chair cyclohexane it is also free of angle strain, but it is not free of torsional strain. When a molecule is viewed along the C-C bonds lying in a horizontal plane the four remaining C-C bonds and four C-H bonds are found to be eclipsed. Moreover, the inside hydrogens on the atoms C-1 and C-4 are close enough to each other to cause steric strain. As a result of these interactions, the boat conformation is less stable than the chair conformation. Both strains are partly reduced by twisting the boat to form a slightly more stable form called the *twist* (or skew-boat) conformation (Fig. 6.6, b).

[1] Alternatively, you may draw these lines slanted. But in this case continue drawing a «back» of the chair in step 2 from the right side of the plane. The resulting conformation will be a reflection (or a flipped form) of chair cyclohexane shown in the example.


4
5
6
3
2
1

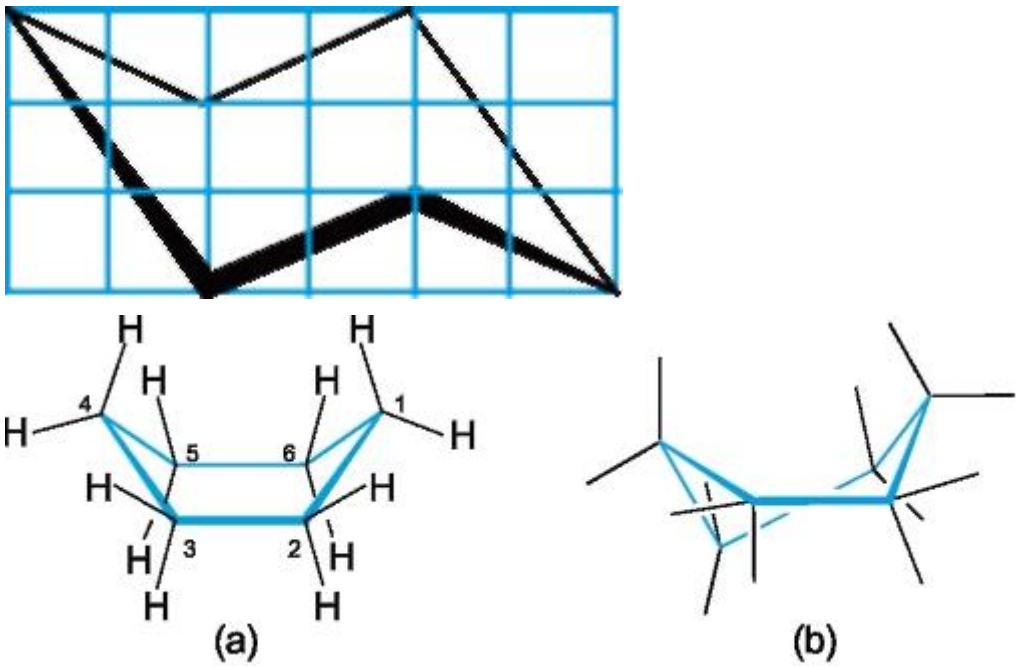
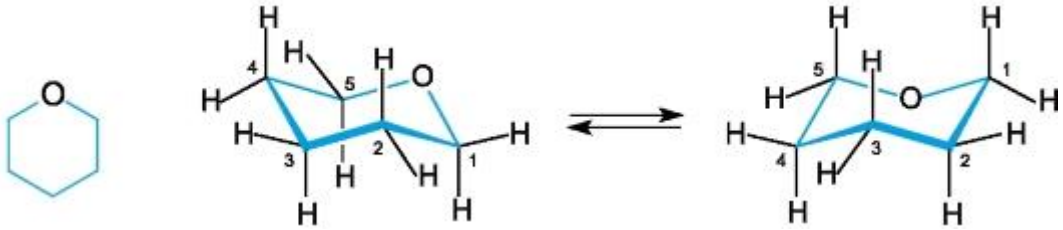


Figure 6.6. The boat (a) and the twist (b) conformations cyclohexane. Hydrogen atoms in (b) are omitted for clarity.

Because of great difference in energies of the chair, boat, and twist conformations of cyclohexane, over 99% of the molecules exist in the chair conformation at room temperature.

A similar consideration can be applied to tetrahydropyran, a saturated six-membered ring containing an oxygen atom that is a framework of many monosaccharides. The most stable conformations of tetrahydropyran are the chair conformations which are readily interconverted.



tetrahydropyran chair conformations of tetrahydropyran

Substituted Cyclohexanes. Cyclohexane and other six-membered rings (with heteroatoms such as oxygen or nitrogen in the ring, for example) are the most common rings found in naturally occurring organic compounds. What is a preferred conformation of a cyclohexane derivative carrying one or more substituents in the ring? Let us consider monosubstituted cyclohexane as the simplest example.

Methylcyclohexane can adopt two possible chair conformations interconvertible through partial rotation about the C-C bonds of the ring (Fig. 6.7). In the conformation (I) the substituent occupies an axial position (ax) whereas in the conformation (II) a position of the methyl group is equatorial (eg). The axial methyl group in the conformer (I) is too close to the axial hydrogens on the atoms

C-3 and C-5 so that the van der Waals forces between them are repulsive. This type of interference is called *1,3-diaxial interaction* (see Fig. 6.7) and is very similar to steric strain observed in conformational analysis of butane. Obviously, the conformer (II) is free of such repulsion. That means that most of methylcyclohexane molecules (about 95%) have the equatorial methyl group at any instant.

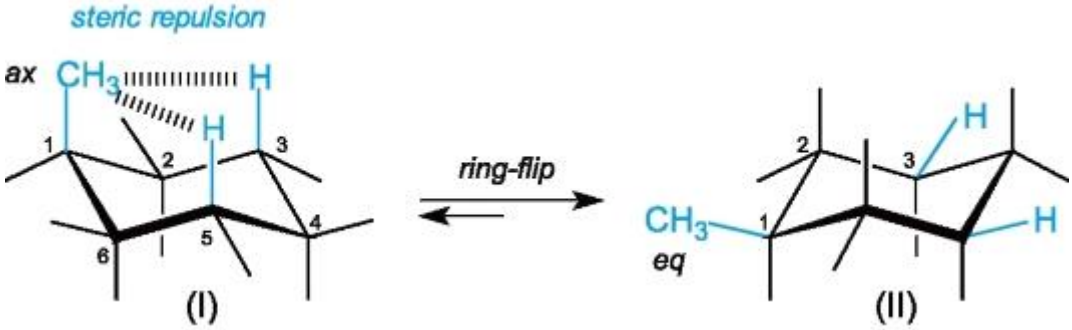


Figure 6.7. Conformations of methylcyclohexane with axial (I) and equatorial (II) methyl group and 1,3-diaxial interaction in (I). Most hydrogens are omitted for clarity.

In cyclohexane derivatives with larger substituents, 1,3-diaxial interactions are even more pronounced. It depends on the nature and size of the axial group. As it might be expected, the amount of steric strain in alkylcyclohexanes increases through the series $-CH_3 < -CH_2CH_3 < -CH(CH_3)_2 << -C(CH_3)_3$. Thus, virtually 100% of *tert*-butylcyclohexane molecules have the bulky *tert*-butyl group in the equatorial position.

In general, similar considerations can by applied to cyclohexanes with two and more substituents. All this leads to the following conclusion:

The preferred conformation of substituted cyclohexane is the chair conforma

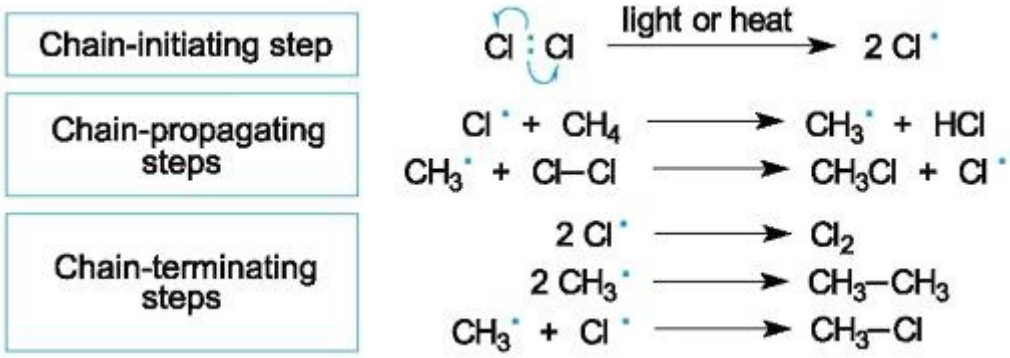
tion in which the most bulky groups are equatorial.

Problem 6.2. Draw the most stable conformation of cyclohexanol.

6.2.2. Chemical Properties

Alkanes and cycloalkanes (except for strained cycles of a small size, especially for cyclopropane which will be considered separately) possess very similar chemical characteristics. These compounds contain non-polar and strong C-H and C-C σ bonds only (Sec. 3.3.1 and Appendix 1) therefore they are relatively unreactive. They are unaffected by strong acids and alkalis as well as by strong oxidants in solution even on heating. On the basis of these properties, alkanes were originally called *paraffins* (from the Latin *parum affinis* - low affinity), and the Russian chemist of the 19th century M. Konovalov called them «chemical corpses».

Halogenation. Reactions of alkanes require either violent conditions (as in cracking process which is not considered in this book), or highly reactive species with an unpaired electron known as free radicals. One of not numerous, but industrially important reactions of alkanes is their halogenation, or introduction of a halogen atom in the alkane molecule. A principal course of the radical reaction of methane and chlorine was demonstrated in Sec. 4.2.1.
All steps in the mechanism of this typical radical substitution reaction *(S_R)* are presented below.

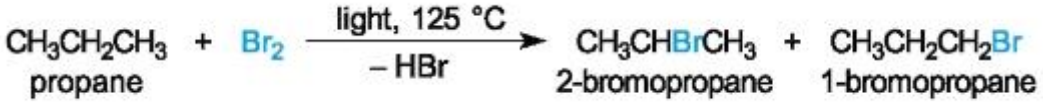


The first step is the *chain-initiating* step that lies in the breaking of the halogen molecule into two halogen atoms. The Cl-Cl bond is much weaker than the C-H bonds in methane (and C-C bonds in other alkanes), and is therefore the easiest bond to split, whereas the alkane remains unchanged.

In the second *chain-propagating* step, a very reactive chlorine atom can collide with a methane molecule, abstracting a hydrogen atom and producing a molecule of HCl and a highly reactive methyl radical, CH_3. This then reacts with a chlorine molecule to give the product chloromethane, CH_3Cl, and a new chlorine atom. The latter formed in this step can further react with methane to continue the chain. This type of a sequential mechanism is called a *chain reaction.*

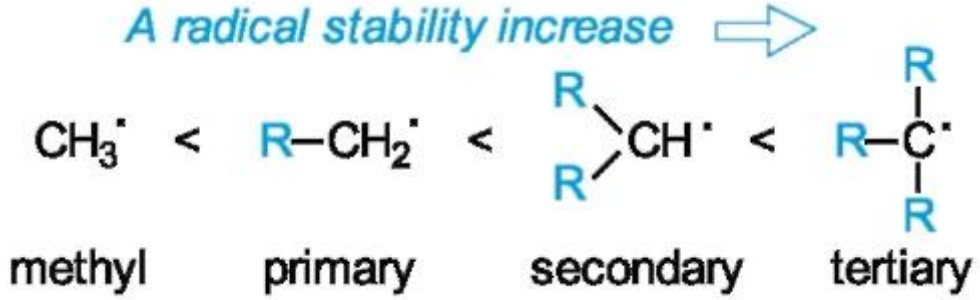
A chain reaction can be ceased if any two radicals combine. No new radicals are formed in these reactions as shown above in the *chain-terminating* steps.

Similar reactions take place when a halogen (chlorine or bromine) reacts with an alkane of a longer chain, though in these cases a mixture of mono-substituted products is formed. The less reactive bromine attacks predominantly a hydrogen atom attached to more substituted carbon, namely, a tertiary (if it present in a molecule), then a secondary, and finally a primary one. Bromination of propane illustrates this:

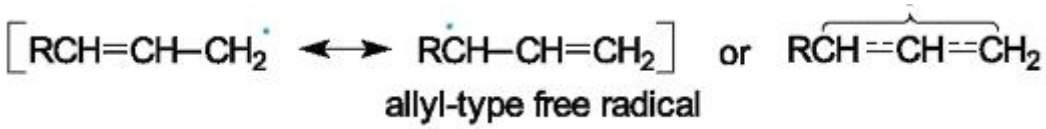


Organic chemists often do not write a balanced equation for some reactions, especially when several products are formed in the reaction. Moreover, non-principal inorganic products may be omitted from the equation.

Two bromides are produced in the reaction with great predominance (about 50:1) of a secondary bromide. Statistically, we would predict the contrary, a ratio 3:1 in favour of a primary bromide.
he prevalent formation of one of the possible products, which is
called *regioselectivity* of the reaction, can be explained by comparative stability of alkyl radicals as intermediates in radical reactions. Indeed, a tertiary radical is the best stabilized one among other alkyl radicals due to the assistance of three methyl groups (remember that carbon radicals are electron-deficient species). Thus the stability order of alkyl radicals is as follows:

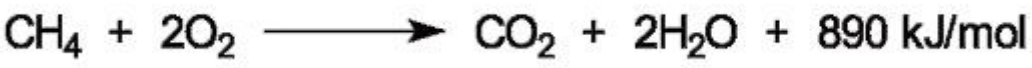


An increased stability is observed for the benzyl radical, $C_6H_5CH_2$', and the allyl-type radical, $RCH{=}CHCH_2$'. The unpaired electron in both radicals is effectively delocalized through resonance, for example:



Problem 6.3. Draw resonance stabilized structures for the benzyl radical.

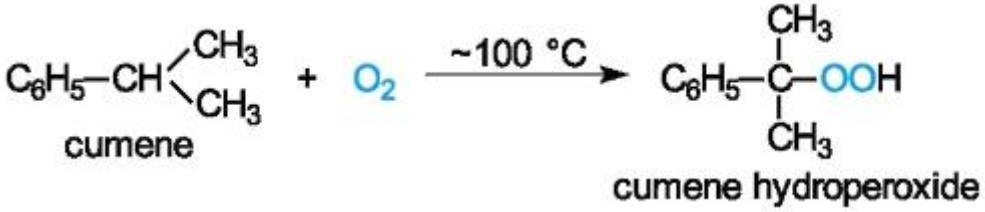
Oxidation. Vigorous oxidation of organic materials with oxygen is known as combustion, the oldest chemical process discovered by primitives. Similarly, combustion of alkanes (and other hydrocarbons) is a highly exothermic reaction used in power engineering. This oxidation, which results in breaking all C-H and C-C bonds, is of no interest to biological processes and organic chemistry including industrial processes.



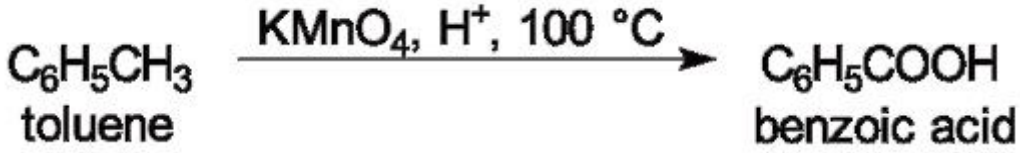
Nevertheless, molecular oxygen, being a radical reagent (see Section 4.2.2), can slowly oxidize an alkane moiety of some functional compounds. Two oxygen atoms intervene in this reaction in the C-H bond to form initially
a *hydroperoxide*, R-O-OH, and products of their further transformations. Thus, diethyl ether widely used in medicine is subjected to air oxidation with the

formation of an explosive hydroperoxide product:

A carbon atom in the allylic or benzylic position, i. e. nearest to the double bond or to the benzene ring, respectively, can also be oxidized with molecular oxygen under relatively mild conditions. The former relates rather to unsaturated compounds and will be discussed later. Benzylic oxidation is exemplified by the conversion of cumene to a hydroperoxide, which is a stage in the industrial production of phenol and acetone.

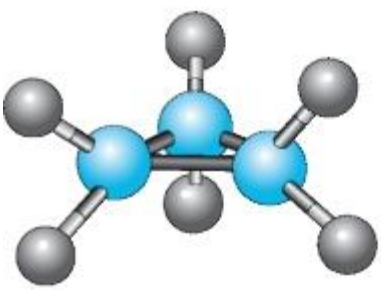


A further example of benzylic oxidation is oxidation of toluene with permanganate solution (remember that the benzene ring and alkanes taken separately are resistant to oxidation under these conditions).

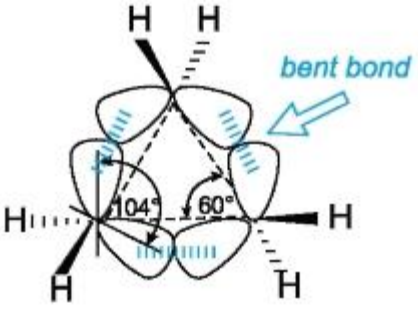


6.2.3. Cyclopropane

The chemical properties of cyclopropane differ greatly from those of alkanes and cycloalkanes of the «normal size», fiveand six-membered. This can be accounted for by the peculiarity of its structure. Cyclopropane as a flat molecule with the shape of regular triangle (Fig. 6.8, a) exhibits a great steric strain. To avoid this, the C-C bonds are formed not by a head-on overlap of the carbon orbitals, which is usual in alkanes, but by an overlap at a slight angle (Fig. 6.8, *b).* The bonds in cyclopropane are often called *bent bonds.* Such an overlap in cyclopropane is less effective than the ordinary overlap in alkanes and, as a result, the bent bonds are weaker than normal alkane bonds.



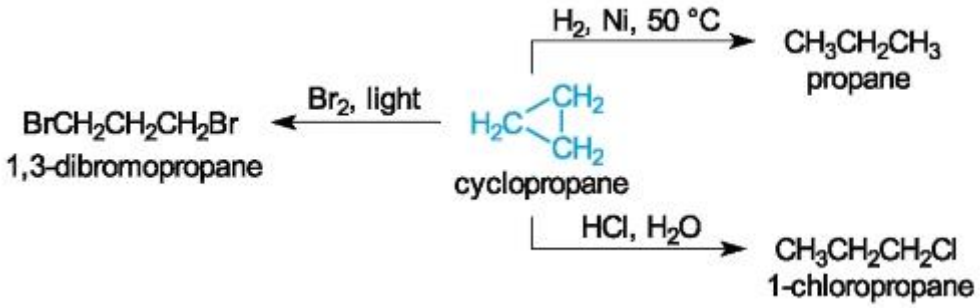
(a)



(b)

Figure 6.8. The cyclopropane molecule: a ball-and-stick model (a) and the C-C-C bond angles and orbital overlap (b).

The next examples show a bond splitting in cyclopropane in the reactions with hydrogen, hydrogen halides, and bromine. Such reactions are not typical of alkanes but rather of alkenes (see the next section).



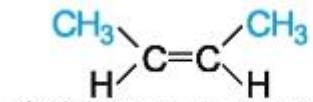
Cyclopropane in itself is a gas (boiling point -33 °C) applied in medicine as a general inhalation anesthetic.

6.3. UNSATURATED ALIPHATIC HYDROCARBONS

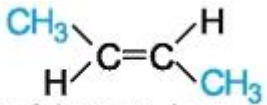
Two main types of unsaturated aliphatic hydrocarbons, alkenes and conjugated dienes, will be considered in this section.

6.3.1. Isomerism

The simplest isomeric alkenes are four-cabon compounds, 1-butene and 2-butene (as linear constitutional isomers). The structure of 2-butene demonstrates another type of isomerism,*stereoisomerism* (in more details see Chapter 10).



cis-2-butene, b. p. +1 °C



trans-2-butene, b. p. +4 °C

Stereoisomers are compounds that have the same order of atoms attachment, but differ only in the arrangement of their atoms or groups in space.

Stereoisomers are termed *cis-trans* isomers when they differ only in the positions of atoms or groups relative to a double bond (namely a π bond). The prefixes *cis* and *trans* have long been used for describing stereoisomers of this type. Atoms or groups are termed *cis* or *trans* to one another when they lie respectively on the same or on the opposite side of a reference plane (from the Latin *cis* - on this side, and *trans* - across). In the *cis* isomer both methyl groups are on the same side on the double bond (Fig. 6.9, a), in the *trans* isomer these groups occupy the opposite positions (Fig 6.9, *b).*

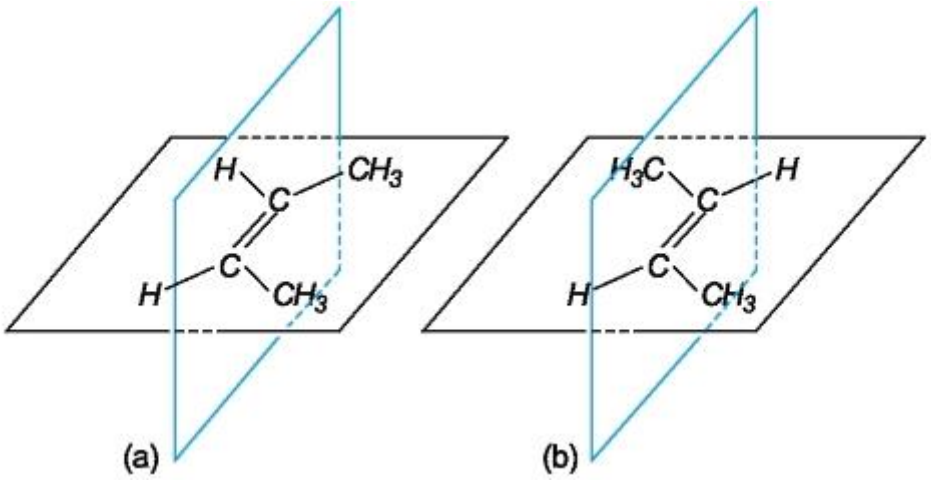


Figure 6.9. *Cis* (a) and *trans* (b) isomers of 2-butene.

Cis and *trans* isomers are distinct compounds, they have different physical properties (boiling points of isomeric 2-butenes see above) and can be separated from one another.

The main reason why *cis-trans* isomers exist is the absence of rotation about a C=C double bond. These isomers can be interconverted only by breaking the π bond, for example, at high temperatures.

Problem 6.4. Which of the following compounds can exist as a pair of *cis-trans* isomers? Draw the structural formulas for them.

(a) $CH_3CH=CH_2$

(b) $CH_3CH=CHCH_3$

(c) $CH_3CH=CHCl$

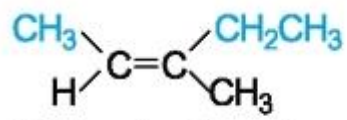
(d) $(CH_3)_2C=CHCOOH$

(e) $ClCH=CHCl$

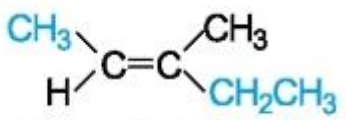
(f) $CH_2=CHCH=CHCH_3$

The E,Z designation. The *cis-trans* notation is successfully used for designating the stereochemistry of disubstituted unsaturated compounds. But this system is ambiguous or non-applicable at all for trisubstituted or tetrasubstituted double bonds. A general system of nomenclature for such stereoisomers, called the *E,Z* system, is based on the priorities of groups applied in the *R,S* convention (Sec. 10.1.5).

In the E,Z system, each of double-bonded carbons is considered *separately* to determine a priority order of the two groups attached to these carbons. If the higher-priority groups on each carbon are on the same side of the double bond, the isomer is designated *Z* (from the German *zusammen* - together). If these groups are on opposite sides of the double bond, the isomer is designated *E* (from the German*entgegen* - opposite). These assignments are illustrated by the following example where higher-priority groups are shown in colour (CH_3 > H and CH_3CH_2 > CH_3).



(*Z*)-3-methyl-2-pentene

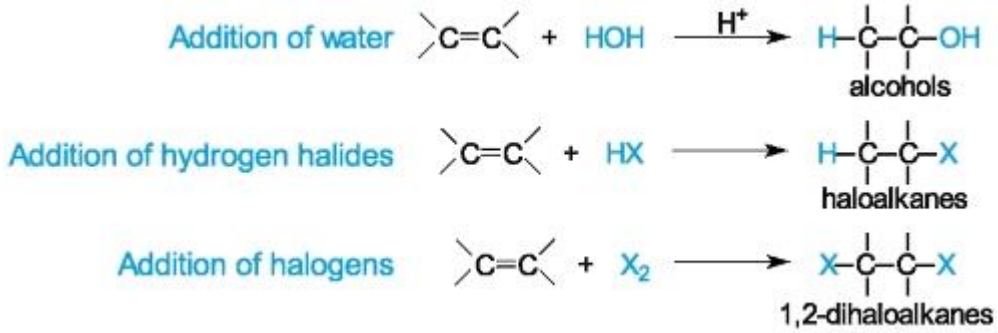


(*E*)-3-methyl-2-pentene

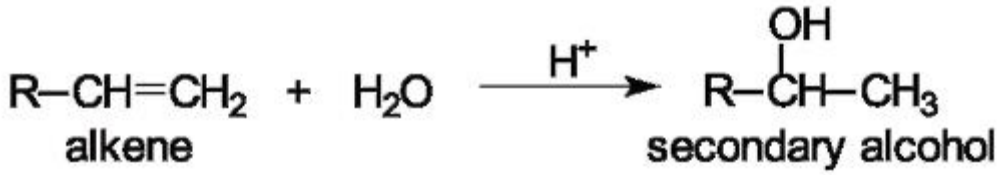
6.3.2. Addition Reactions of Alkenes

The electronic structure of a double bond has been discussed in the previous chapters (Sec. 1.4.2 and 3.1.1). The π bond in alkenes is weaker than the σ bond in alkanes because the electrons of the π bond are less tightly held. As a result unsaturated hydrocarbons are more reactive than alkanes and undergo addition reaction.

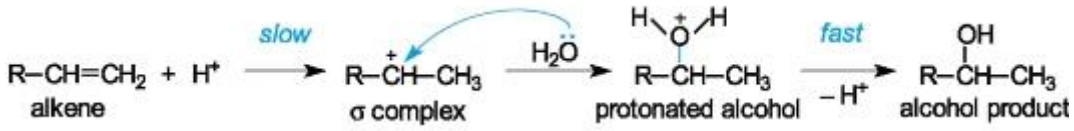
The electrons of the π bond of alkanes are susceptible to electron-poor reagents (electrophiles). Electrophiles include proton-containing compounds, e. g. strong acids, or neutral reagents, such as halogens, which can be polarized in the polar media. A reaction begins with an attack of an electrophilic molecule on the electrons of the π bond, and is therefore called an *electrophilic addition reaction* (A_E) The most important reactions of this type are shown below (where × is a halogen) and then will be considered in more detail.



Addition of water (hydration). This reaction is one of the most important reactions in biological transformation of unsaturated compounds. It takes place on treating an alkene with water in the presence of a strong acid catalyst.



In a first, slow step of the reaction a proton acts as an electrophile and forms a new C-H bond to one of the double-bonded carbon of a substrate. Because this bond uses both π electrons, the other carbon atom becomes positively charged. A carbocation intermediate thus formed is called a σ *complex.* It in itself is an extremely reactive electrophile that reacts rapidly with water as a nucleophile to yield a further intermediate, protonated alcohol. Loss of the proton from the latter in a final, fast step gives a neutral product and regenerates the catalyst.



Alkene hydration is used industrially, for example, in ethanol production.

In all reactions of the unsymmetrically substituted alkenes with polar reagents (like HOH or, in general, HX) a single addition product is formed rather than the mixture that might be expected. This means that the reaction proceeds *regioselectively.*

After studying a number of such reactions, the Russian chemist V. Markovnikov proposed (in 1869) what has become known as *the MarkovnikoVs rule:*

In the addition of HX to an unsymmetric alkene, the H attaches to the

carbon with fewer alkyl substituents (or with the greater number of H atoms), and the × attaches to the carbon with more alkyl substituents (or to the less hydrogenated carbon).

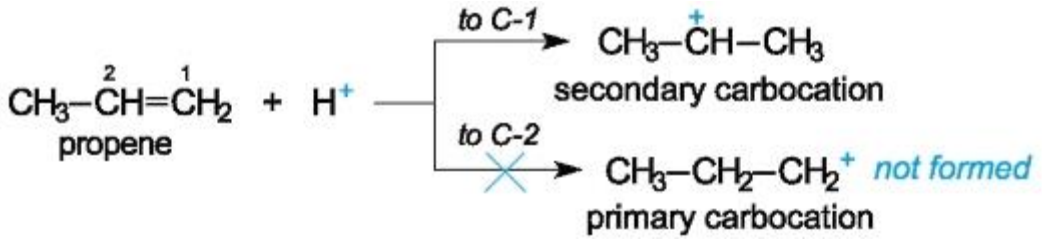
Problem 6.5. Use the Markovnikov's rule to predict the major product formed in the reaction of water with: (a) 2-methyl-2-butene; (b) 1-methylcyclohexene.

Much later, a rational explanation for the Markovnikov's rule has been given in terms of the modern chemical theory. The first approach consists in the distribution of partial charges in a non-reacting alkene molecule.

For example, the +/ effect of the methyl group in propene leads to the appearance of a partial negative charge on the C-1 atom. Just this carbon will be attacked first by an electrophile in the static state.

bonds polarization in propene

The second approach to explain the Markovnikov's rule involves carbocation stability. Propene might form either a carbocation with two alkyl substituents (a secondary ion), or a carbocation with one alkyl substituent (a primary ion):



Alkyl carbocations greatly differ in their stability. Carbocations have a planar configuration where a charged carbon is sp^2-hybridized (Figure 6.10). Alkyl substituents delocalize the positive charge due to their +/ effect. Thus, the more alkyl groups are attached to the cationic site, the more electron density shifts toward the charge and the more the cation is stabilized. The order of stability of carbocations is tertiary > > secondary > primary > methyl, i. e. $R_3C+ > R_2CH^+ > RCH_2+ > CH_3+$. Notice that this order is the same as that of free radicals (Sec. 5.1.1).

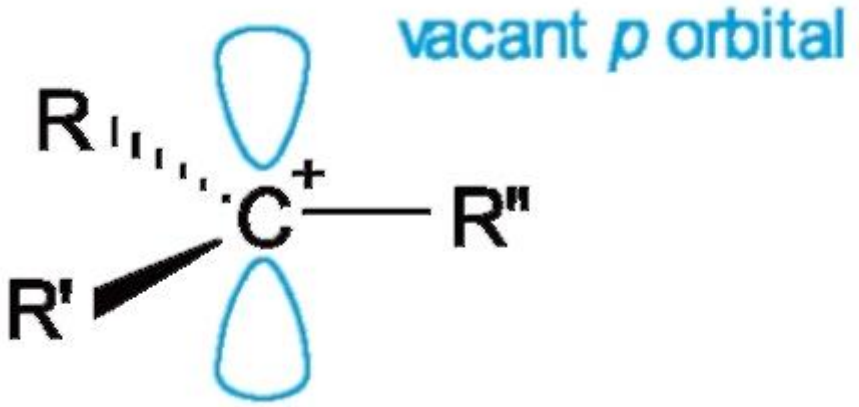
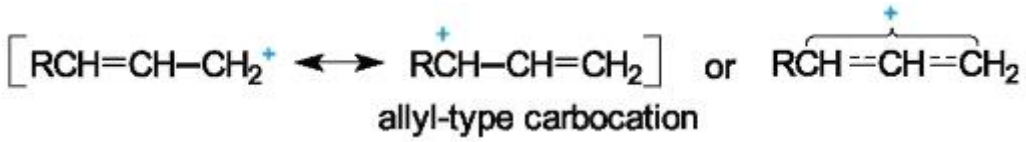
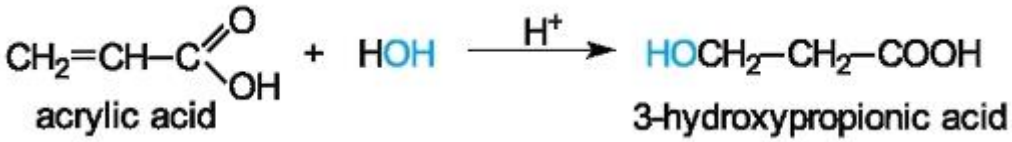


Figure 6.10. The electronic structure of a carbocation (sp^2-hybridized orbitals are shown by dashes).

It should be added that the increased stability is observed for two primary cations - the benzyl cation, $C_6H_5CH_2^+$, and an allyl-type cation, $RCH=CHCH_2^+$, where the positive charge is effectively delocalized through resonance, for example:



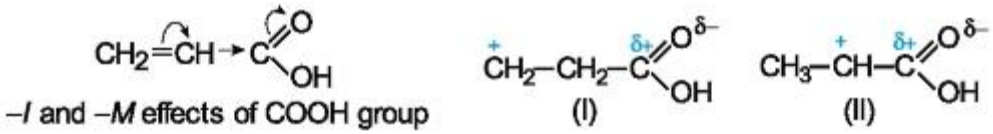
One cannot but say that some addition reactions do not follow Markovnikov's rule. This occurs when strong electron-withdrawing group, such as a carboxyl group or a carbonyl group, is attached to a double bond. For example:



After a rational interpretation, this reaction course is not considered as an exception, but as another way called *the anti-Markovnikov addition.*

Example 6.1. How to explain the anti-Markovnikov addition of water to acrylic acid?

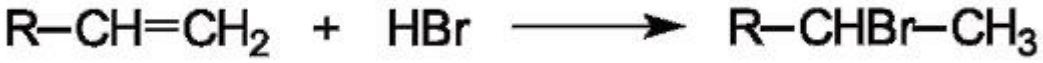
Solution. Polarization of the double bonds in a substrate results in the appearance of a partial negative charge on the α-carbon (nearest to the functional group) in α,β-unsaturated carboxylic acids due to -/ and -M effects of the substituent. This result can also be explained in terms of the stability of the σ complexes (I) and (II). The primary carbocation (I) is more stable of the two possible ones since it contains remote positive charges, whereas these charges are situated on the neighbouring carbons in the secondary cation (II).



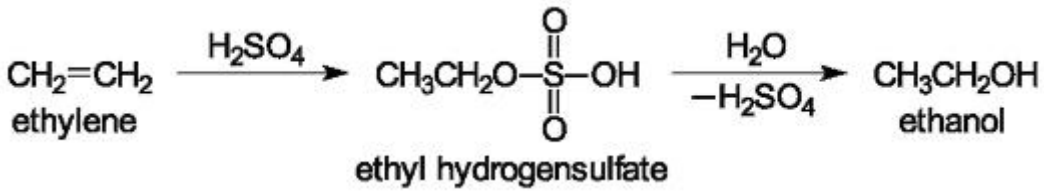
At present the Markovnikov's rule can be restated in a modern form as follows:

Electrophilic addition of unsymmetric reagent to an unsymmetric double bond proceeds through the formation of the most stable carbocation.

Addition of hydrogen halides and sulfuric acid. Alkenes add hydrogen halides to yield haloalkanes, or alkyl halides. Reactions of unsymmetric alkenes and cycloalkenes proceed regioselectively according to the Markovnikov's rule.



Reaction with concentrated sulphuric acid results in formation of hydrogensulfates, which can be readily converted to alcohols under hydrolysis.

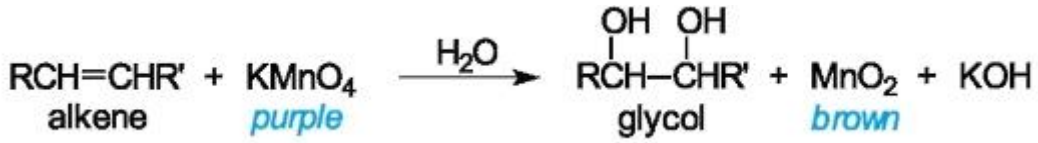


Addition of halogens (halogenation). Alkenes readily add chlorine and bromine, even at room temperature or below. The reaction is nearly instantaneous and is used therefore as a simple visual test for the presence of a double bond by decolouration of a brown-coloured bromine solution.

This reaction also belongs to electrophilic addition though bromine is a non-polar reagent. However, the bromine molecule easily becomes polar in a polar media, for example as Br^{5} +-Br^{5-} in water solution.

Oxidation and reduction. Alkenes are more easily oxidized than alkanes. The most employed chemical oxidizing agents are potassium permanganate, peroxides, and ozone.

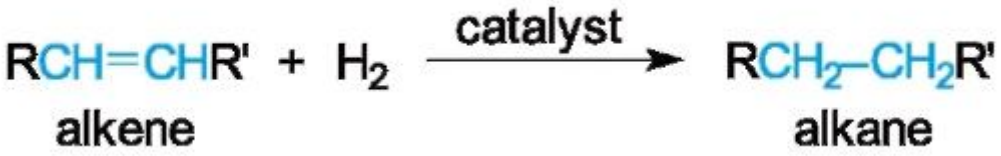
Potassium permanganate reacts with unsaturated compounds at room temperature in a neutral or alkaline solution to form *glycols,* or 1,2-diols (compounds with two adjacent hydroxyl groups). The distinctive purple color of the permanganate ion is lost during the reaction and the brown precipitate of manganese dioxide appears. The reaction is therefore widely used as a test to distinguish unsaturated compounds from saturated ones. It is often called the *Baeyer test* (in Russia it is known as the *Wagner test).*



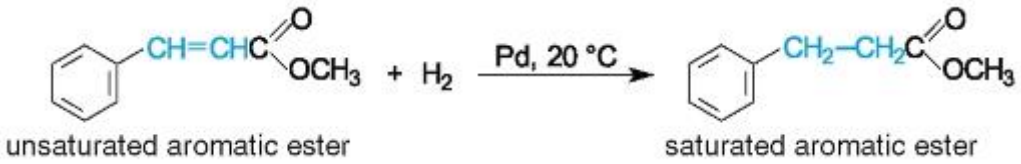
The permanganate oxidation of alkenes may also be performed in a slightly acidic medium at low temperatures with a formation of a colourless Mn^{2+} ion. The same oxidant cleaves the double bond in a strong acidic medium.

Since permanganate is able to oxidize other functional groups as well, this test is best used to complement other criteria of unsaturation such as bromine decolouration.

The most generally applicable procedure for the reduction of alkenes is *hydrogenation* (addition of hydrogen) in the presence of an appropriate catalyst, such as finely divided metals (usually platinum, palladium, or specially prepared nickel).



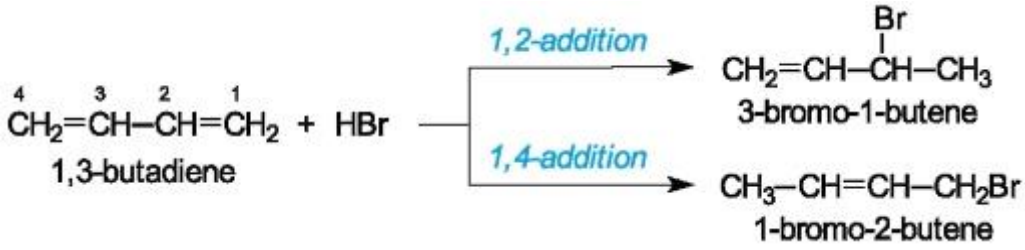
Many alkene derivatives that have potentially reducible functions (the benzene ring, the carbonyl, carboxyl, or ester group) can be hydrogenated selectively under mild conditions with retention of the mentioned groups. This is a result of higher reactivity of a double C=C bond. Note in the example below that the benzene ring is not affected even though it is a highly unsaturated compound.



Hydrogenation of the alkene double bond is of great importance in the laboratory and, especially, in industry, for example in the process of hardening liquid vegetable oils (Chapter 12).

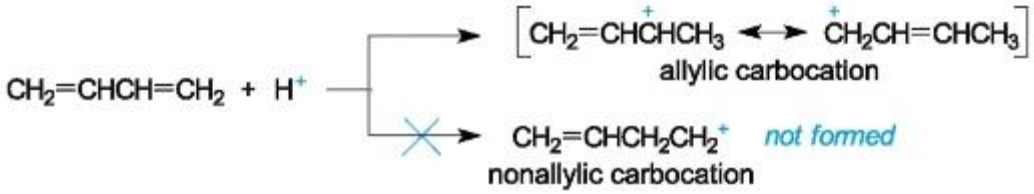
6.3.3. Addition Reactions to Dienes

Dienes undergo addition reactions in much the same manner as alkenes. One of the most substantial differences between conjugated dienes and alkenes or isolated dienes is in their electrophilic reactions. The former yields usually a mixture of addition products in the reaction with equimolar amount of electrophilic reagent, for example:



In one of the products (3-bromo-1-butene), hydrogen bromide has added to one of the double bonds to give a «normal» Markovnikov product, the so-called product of *1,2-addition.* But the second product, called the *1,4-addition* product, appears unusual. How can we account for its formation?

Recall the two-step mechanism $A_E$described in Sec. 6.3.2. In the first, slow step two intermediate carbocations are possible: an allylic cation, which is stabilized by resonance, and a primary nonallylic cation. The formation of a more stable intermediate requires less activation energy, thus the allylic cation is the only species in the reaction:



When in the second step the allylic cation reacts with the bromide ion, it can react either at the C-1 or C-3 cationic site to give a mixture of the 1,2- and 1,4-addition products. In general, 1,4-addition products are predominant when reaction is performed at higher temperatures.

6.4. AROMATIC HYDROCARBONS

Benzene and other aromatic hydrocarbons, including their homologues, are known generally as *arenes.*They are highly unsaturated compounds, but they do not behave like typical alkenes because of the peculiarities of their electronic structure discussed above (Sec. 3.2.2). The principal distinctions in the chemical behaviour between benzene and alkenes are:

• benzene reacts with halogens (bromine and chlorine) only in the presence of a catalyst by*substitution*, whereas alkenes readily undergo *addition* reaction with halogens without a catalyst;

• benzene is much more resistant to oxidation, whereas alkenes are easily oxidized.

6.4.1. Electrophilic Aromatic Substitution

The electron cloud of the benzene ring makes it susceptible to attack by electrophiles. The most important reactions of arenes are the *electrophilic aromatic substitution* reactions in which an electrophile substitutes for one of the hydrogen atoms (Fig. 6.11).

These reactions allow introducing a wide variety of groups into aromatic rings thus giving rise to a vast number of organic compounds of different classes. Using the proper conditions and reagents, it is possible to introduce an alkyl group R in an *alkylation* reaction, a halogen atom in a *halogenation*reaction, a nitro group -NO_2 in a *nitration* reaction, a sulfo group -SO_3H in a *sulfonation* reaction, or an acyl group -C(O)R in an *acylation* reaction. Reagents and reaction conditions will further be considered.

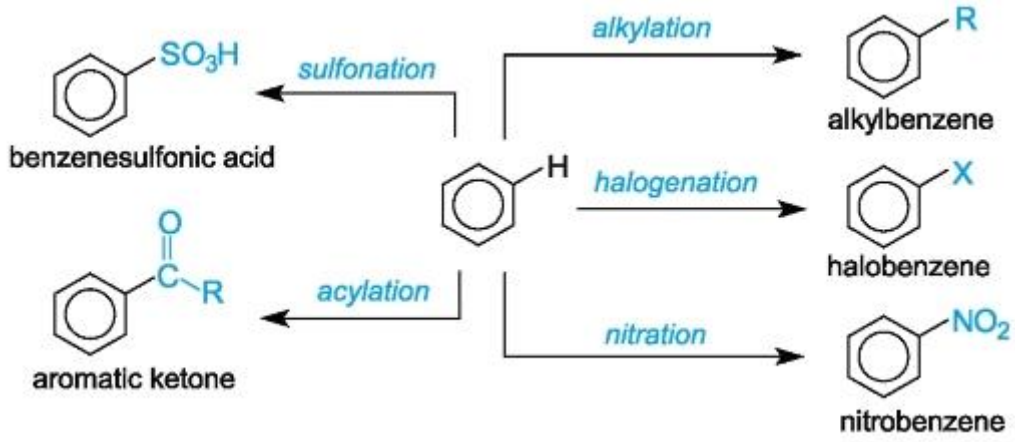
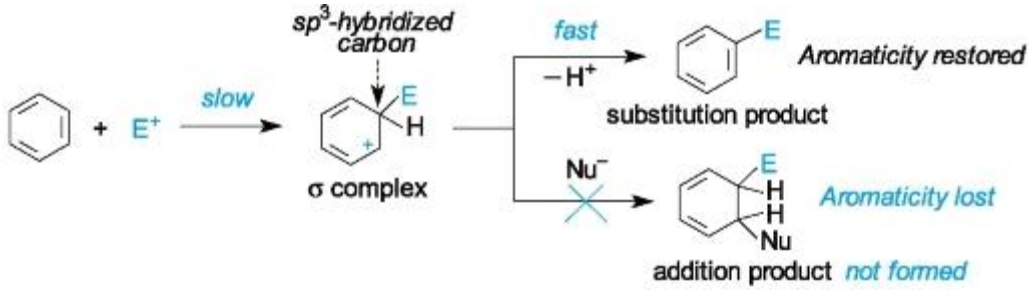


Figure 6.11. Electrophilic aromatic substitution in benzene.

Note that the term *halogenation* designates two different reactions, *addition* to alkenes and *substitution*in arenes.

To understand why electrophiles react by substitution instead of addition, let us consider a reaction mechanism.
An initial step in the reaction involves an attack on the π-electron system of the benzene ring by an electrophile (E+) just as in the first step of an electrophilic addition to alkenes. However, aromatic hydrocarbons are much less reactive than alkenes and a catalyst is always needed which makes the reagent molecule more electrophilic. The electrophile bonds to one carbon of the benzene ring, using two of the π electrons from the aromatic cloud to form a σ bond with this carbon that becomes sp^3-hybridized. The resulting nonaromatic carbocation (a σ complex) is stabilized by resonance, i. e. the positive charge is delocalized over five carbons. This is a slow, rate-limiting step having high activation energy because of disruption of the aromatic π system.

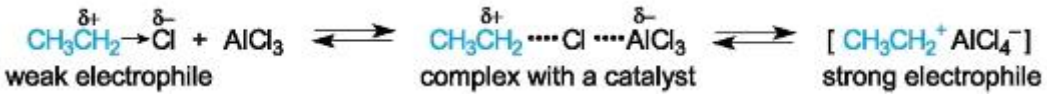


The reaction is completed by loss of a proton from the sp^3-hybridized carbon in the second, fast step in which the aromatic system is regenerated. Thus the net result of the reaction is substitution. Another possibility for stabilization of the σ complex by addition of a nucleophile is unfavourable because of loss of aromaticity.

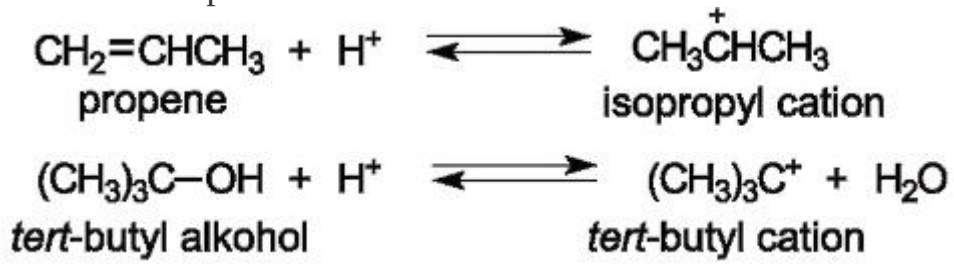
Several types of electrophilic aromatic substitutions will now be briefly considered.

Alkylation of benzene. This is one of the most useful of all aromatic substitution reactions, called after its discoverers the *Friedel-Crafts reaction* (1877). Alkyl halides serve as electrophilic reagents in a classic variant of the reaction.

Aluminium chloride, as a Lewis acid (Sec. 5.1), accepts an electron pair of the chlorine atom of alkyl chloride thus helping the latter to ionise.



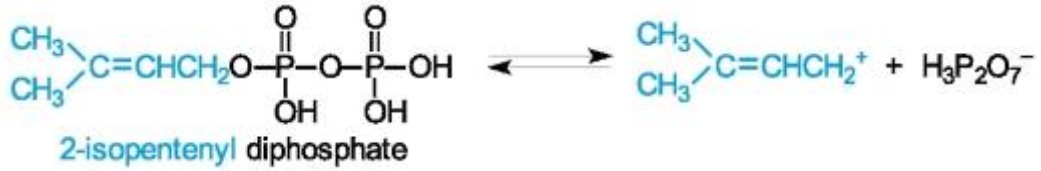
Alkenes and secondary or tertiary alcohols may be used as reagents for alkylation of the aromatic ring. In these cases electrophilic species (carbocations) are produced in the presence of an acid.



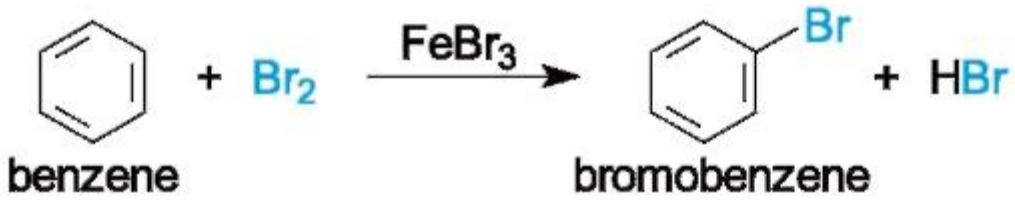
Problem 6.6. A first step in the industrial process of producing styrene, $C_6H_5CH=CH_2$, involves acid-catalyzed alkylation of benzene by an appropriate alkene. Suggest this alkene, write the equation for the reaction and the steps in the reaction mechanism.

In living systems, aromatic substrates can be alkylated by means of the amino acid methionine,$CH_3SCH_2CH_2CH(NH_2)COOH$ (for transfer of a methyl group), or with

an allylic type carbocation (for transfer of C_5 fragment), generated from 2-isopentenyl diphosphate, for example:



Halogenation of benzene. Chlorine or bromine reacts with benzene only in the presence of a Lewis acid. Note that only one atom of halogen is introduced in the benzene ring.



The catalyst functions in a similar way as in the Friedel-Crafts reaction, namely, it polarizes the bromine molecule to form the strong electrophilic complex Br_3Fe^{5-}

•

•

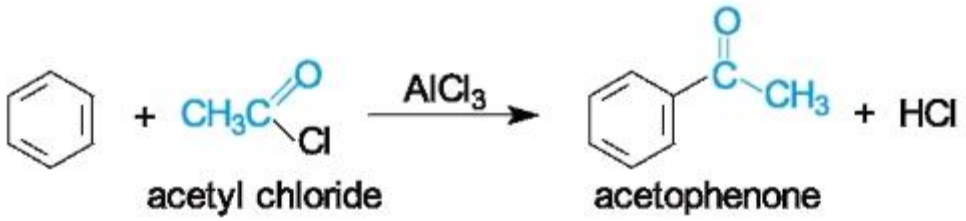
• Br

•

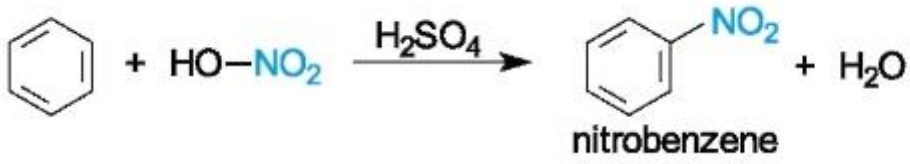
•

• Br^{5+}.

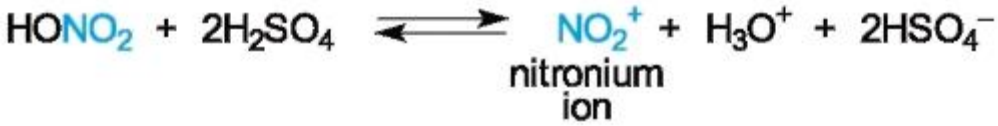
Acylation of benzene. Acylation is a reaction of introducing an acyl group, RC(O)-, into a substrate molecule. In the aromatic series this reaction is also called the Friedel-Crafts reaction and resembles the alkylation reaction in many respects. The electrophile is generated from a carboxylic acid derivative, usually an acyl chloride, RC(O)Cl, as in the following example:



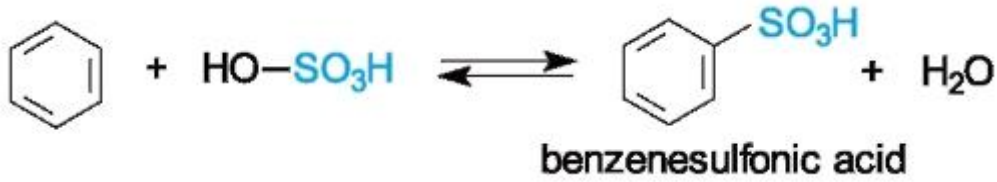
Nitration of benzene. A mixture of concentrated nitric and sulfuric acids is used for introducing a nitro group into the aromatic ring.



Sulfuric acid acts as a catalyst for producing the electrophilic *nitronium ion*, NO_2^+, according to the equation:



Sulfonation of Benzene. Incorporation of a sulfo group, $-SO_3H$, into the benzene ring occurs when an aromatic hydrocarbon is treated either with concentrated sulfuric acid or its mixture with sulfur trioxide, SO_3; for example:



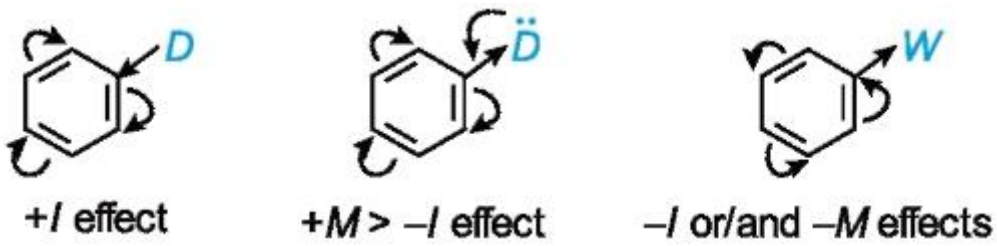
The sulfonation reaction is reversible; this means that the sulfo group can be removed, for example, on heating benzenesulfonic acid with steam at 120-150 °C.

Both sulfonation and nitration are important reactions because their products, nitro compounds and sulfonic acids, can be easily transformed into other classes of aromatic compounds such as phenols, amines, etc., which are valuable intermediates in the preparation of various biologically active compounds, including pharmaceuticals.

6.4.2. Substituent Effects in Electrophilic Aromatic Substitution

Substituents already present in the benzene ring have a marked influence on the electrophilic substitution. Firstly, they affect the *reactivity* of the ring and, secondly, they affect the *orientation* of substitution. Both the effects are controlled by the cooperation of two factors: inductive and mesomeric influences of the substituent.

Reactivity of aromatic rings. Substituents attached to the benzene ring can be divided into electron-donating and electron-withdrawing ones, as it has been shown above (Sec. 3.3). The electron-donating groups (designated *D* in the drawing below) increase electron density on the ring; the electron-withdrawing groups (W), on the contrary, decrease it.



Electron-donating groups facilitate an electrophilic attack on the benzene ring and are therefore referred to as *ring-activating* substituents. Electron-withdrawing groups are naturally *ring-deactivating*ones. It is clearly demonstrated by

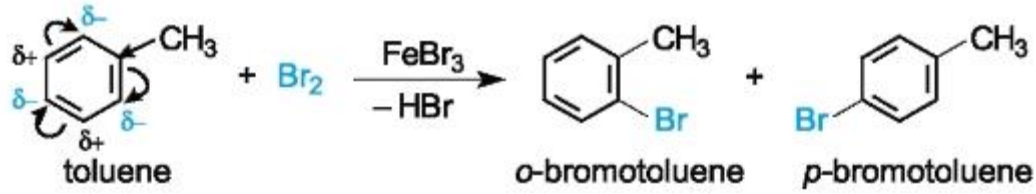
comparing the relative rates of nitration (under the same reaction conditions) of the following compounds:

C_6H_5OH	$C_6H_5CH_3$	C_6H_6	C_6H_5Cl	$C_6H_5NO_2$
phenol	toluene	benzene	chlorobenzene	nitrobenzene
103	2.5	1.0	0.03	<0.0001

Many aromatic compounds that contain strong deactivating groups are not reactive at all. For example, benzoic acid, benzaldehyde, and some others, do not undergo the Friedel-Crafts alkylation and acylation reactions.

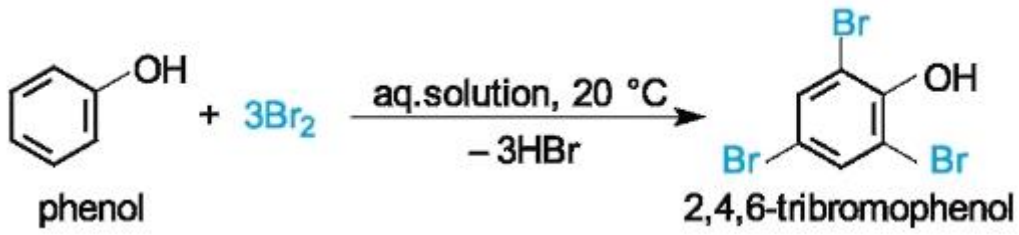
Problem 6.7. Arrange the following compounds according to the increase of their reactivity in electrophilic substitution reaction: (a) aniline, (b) benzene, (c) benzoic acid, (d) ethylbenzene. Write the equation for a bromination reaction of the most reactive compound.

Directing effects. Electron-donating groups increase electron density on all carbon atoms of the ring but especially on the atoms C-2, C-4, and C-6, i.e. on the *ortho* and *para* positions as it is indicated by the arrows head. Two products are formed in this case: *ortho* and (not «or»!) *para* disubstituted aromatic compounds:



A ratio of the substitution products depends on many factors and varies greatly. Substituents of this kind are called *orthoand para-directing.*

Phenol and aniline are the most reactive substrates for electrophilic substitution because of a strong mesomeric effect of the OH and NH_2 groups. Bromination of these compounds readily proceeds without a catalyst, giving rise to a water insoluble trisubstituted product, for example:



Bromination reaction is used in qualitative and quantitative analyses of phenol and aniline.

Electron-withdrawing groups attached to the benzene ring decrease electron density on all carbon atoms of the ring but to a less extent on the atoms C-3 and C-5. Such substituents are called *meta-directing.*

The main exception to this regularity is halogenobenzenes which undergo *ortho* and *para* substitution, although they are less reactive than benzene. Halogens are deactivating substituents because of their electron-withdrawing effect; but they are *ortho*- and para-directing due to their +M effect.

The results on ring reactivity and orientation in electrophilic substitution are summarized in Table 6.1.

Problem 6.8. Write equations for the following reactions, and name the products obtained:

(a) chlorobenzene + isopropyl chloride ($AlCl_3$ catalyst) -

(b) toluene + acetyl chloride ($AlCl_3$ catalyst) -

Table 6.1. Substituents effects in electrophilic aromatic substitution

*Substituent in the benzene ring**	*Electronic effect*	*Influence on reactivity***
***ortho*- and *para*-Directing**		
$-NH_2$, $-NHR$, $-NR_2$, $-OH$	$+M \gg -I$	+++
$-O^-$	$+M$, $+I$	+++
$-NHCOR$, $-OR$	$+M > -I$	++
$-CH_3$ and other alkyl groups	$+I$	+
$-Cl$, $-Br$, $-I$	$-I > +M$	–
***meta*-Directing**		
$-C\equiv N$, $-CHO$, $-COR$, $-COOH$, $-COOR$, $-SO_3H$	$-I$, $-M$	– –
$-NO_2$	$-I$, $-M$	– – –
$-NH_3^+$, $-NR_3^+$, $-CF_3$, $-CCl_3$	$-I$	– – –

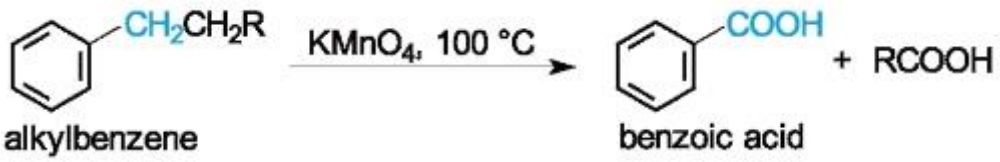
* R is an alkyl group.

** Activates strongly (+++), moderately (++), weakly (+); deactivates strongly (---), moderately (-), weakly (-).

6.4.3. Oxidation and Reduction of Arenes

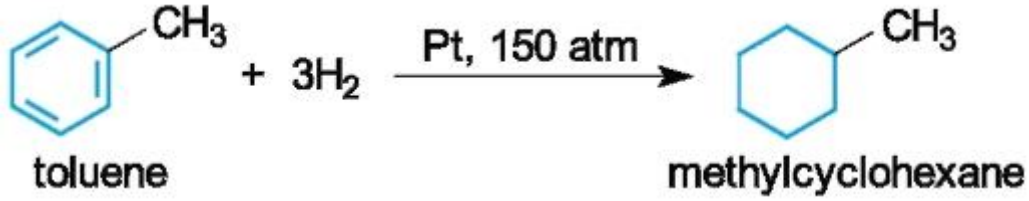
As it has been stated at the beginning of this section, benzene (and the benzene ring) is quite inert to strong oxidizing agents. Benzene homologues, on the contrary, are readily oxidized by aqueous potassium permanganate or chromic acid

on heating. The reaction involves attack of the oxidant on a C-H bond nearest to the benzene ring, thus converting an alkyl side-chain into a carboxyl group:



Similar oxidation is well known in biological systems where methyl groups are oxidized enzymically to COOH groups. This transformation explains a mechanism for the removal of toluene in the form of benzoic acid from a human body (for the equation see Sec. 6.2.2). For this reason toluene is less toxic than benzene which cannot be oxidized.

Aromatic rings are also inert to catalytic reduction under the conditions used for hydrogenation of alkene double bonds. Nevertheless, the benzene ring can be reduced into a cyclohexane ring using a more powerful catalyst (such as rhodium) or vigorous conditions (high temperature and pressure) in the presence of ordinary catalysts for hydrogenation, for example:



Additional Problems

6.9. Draw a Newman projection of the most stable conformation of 2,3-dimethylbutane, looking along the C-2-C-3 bond.

6.10. Write the equation for the monobromination reaction of cyclohexane. Name the product and state the mechanism of the reaction.

6.11. Arrange the following compounds according to the increase of their reactivity in electrophilic substitution reaction:

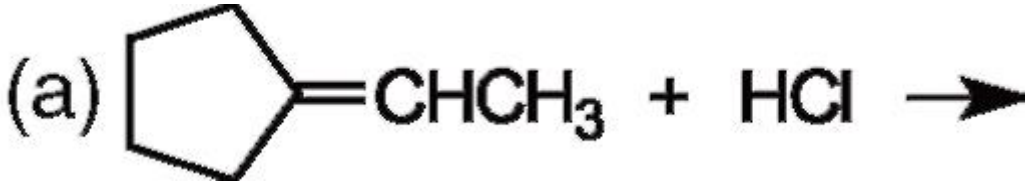
(a) acrylic (propenoic) acid;

(b) 2-butene;

(c) propene.

Write equations for hydrogen chloride addition to all three compounds.

6.12. Show the structures of carbocation intermediates you would expect to appear in the following reactions:



(b) 2-pentene + HBr → .

Name the products of the reactions.

6.13. Propose the structure of an alkene to yield 3-methyl-2-butanol by hydration. Show all steps in the reaction mechanism.

6.14. What *alkenes* would you start with to prepare 1-methylcyclohexanol in one step? Write equations for the reactions.

6.15. Give the structures of the likely products from the reaction of one equivalent of hydrogen chloride with 1,3-butadiene. Show both 1,2- and 1,4-addition products.

6.16. Write equations for the following reactions, and name the products obtained:

(a) toluene + Cl_2 (Fe catalyst) →

(b) toluene + Br_2 (UV light)→ .

6.17. Show how the following conversions may be carried out. More than one step may be required. Assume that isomers can be separated if necessary.

(a) toluene to p-bromobenzoic acid;

(b) benzene to m-chlorobenzoic acid.

6.18. An amino group is one of the strongest activating groups in electrophilic aromatic substitution (recall bromination of aniline). Aniline, however, undergoes sulfonation only at 180-190 °C to produce p-aminobenzenesulfonic acid (or sulfanilic acid, an intermediate in the preparation of sulfa drugs). Explain the experimental difficulty in the reaction.

PART 2. MONOFUNCTIONAL ORGANIC COMPOUNDS OF BIOLOGICAL INTERESTS.

Chapter 7. ORGANIC HALIDES, ALCOHOLS, PHENOLS, ETHERS, AMINES, AND ORGANOSULFUR COMPOUNDS

This chapter involves compounds of various classes that have the following functional groups: halogens, -OH, -O-, -SH, -S-, and -NH_2. General formulas of the classes are given below, where R is aliphatic or aromatic radical (except for alcohols and phenols):

HALIDES	ALCOHOLS	PHENOLS	ETHERS	AMINES	THIOLS	SULFIDES
R–Hal	R–OH	Ar–OH	R–O–R'	R–NH_2, R_2NH, R_3N	R–SH	R–S–R'

It is easy to see that all the functional groups are attached to a carbon framework by a single bond which is polar to a greater or a lesser extent. This provides some common chemical properties of the mentioned classes though they have substantial distinctions.

Compounds containing the amino and hydroxyl groups are of very common occurrence in nature. Organic halides, on the contrary, are rarely encountered in higher plants and animals. Most of them are creations of the laboratory chemists. They are employed as industrial solvents, refrigerants, pesticides in agriculture,

inhaled anesthetics in medicine and so on. Moreover, organic halides are versatile reagents in syntheses since they can be converted into many other classes of organic compounds, including the objects of this chapter. We begin the consideration with organic halogen compounds.

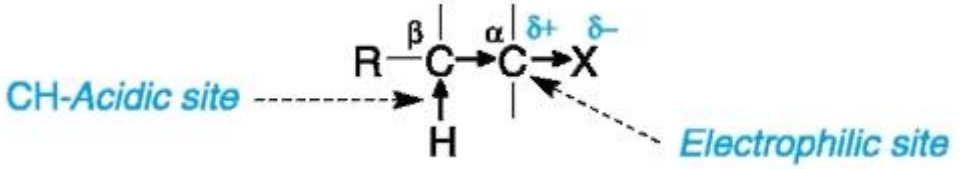
7.1. ORGANIC HALIDES

Organic halides are hydrocarbon derivatives in which one or more halogen at

oms replace hydrogen atom(s).

In this section the main attention will be paid to alkyl halides in which a halogen is attached to saturated,sp^3-hybridized carbon. Other types of organic halides, such as aryl halides and vinyl halides, RCH=CHHal, also exist. Halogen in these compounds is bonded to sp^2-hybridized carbon, therefore they differ greatly in their chemistry from alkyl halides.

Before discussing the chemistry of alkyl halides, let us consider reactive sites in a structure of the saturated compound of the general formula R-X, where × is an electron-withdrawing group. Polarization of a C-X bond results in the formation of an electrophilic site (the α carbon) that can be attacked by a nucleophilic reagent. The bonds C_a-C_β and C_β-H are also polarized due to the -/ effect of a substituent X. This results in the formation of CH-acidic site that can react with bases.

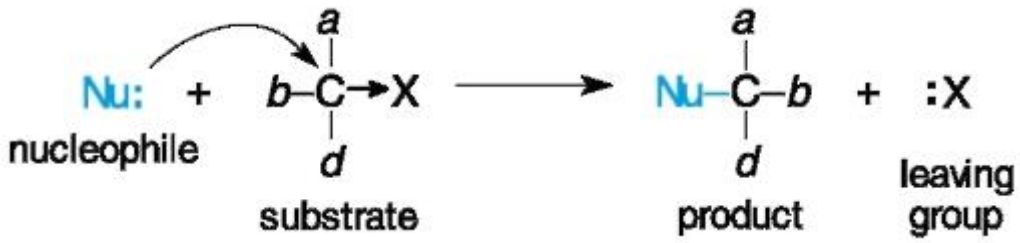


Both sites define the reactivity of the compound. They will be discussed below.

7.1.1. Nucleophilic Substitution Reactions

Nucleophilic substitution at saturated carbon is one of the simplest and, at the same time, the most important type of organic reactions. It may also be characterized as the *alkylation* reaction with reference to the nucleophile. Many reactions of biological importance represent nucleophilic substitution.

In this reaction, one covalent bond is broken, and a new bond is formed. The reaction can be expressed in the following general equation, where a, *b*, *d*, and × are atoms or groups attached to the electrophilic carbon:

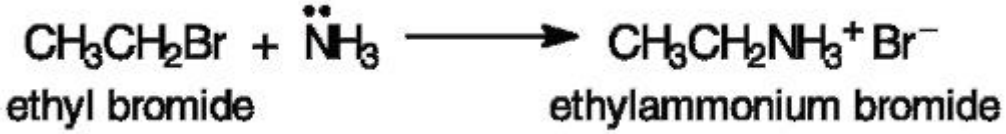


In the overall transformation, the C-X bond is ruptured in such a way that a pair of electrons that formed the bond becomes associated with a group × called the *leaving group.* The nucleophile possesses an unshared (non-bonding) pair of electrons and uses them to form a new bond to the carbon.

A typical nucleophilic substitution reaction is alkaline hydrolysis of halides, to give an alcohol and a halide ion, for example:

In this reaction a nucleophilic hydroxide ion attacks the substrate (methyl bromide) and expels a bromide ion as the leaving group.

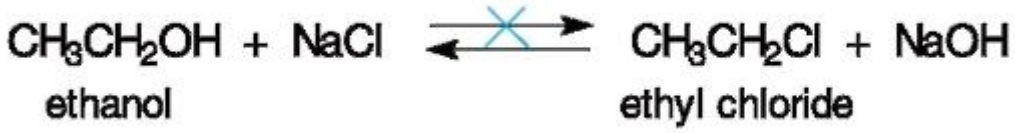
Nucleophiles and Substrates. Compounds of different classes may be considered as nucleophiles. Both neutral molecules and anions can serve as nucleophilic reagents. They are classified according to the kind of an atom that forms a new bond to the electrophilic carbon. The most common nucleophiles are oxygen, nitrogen, sulfur, halogen, and carbon nucleophiles. Table 7.1 lists some nucleophilic reagents and shows a vast variety of products that they form in the reaction with alkyl halides. It should be noted once more that aryl halides and vinyl halides undergo this type of nucleophilic substitution with difficulty, if at all.



Ethylamine as a free base can then be obtained from its salt by treatment with a stronger base, an alkali, for example:

Example 7.2. Is it possible to synthesize ethyl chloride from ethanol and sodium chloride as a nucleophile?

Solution.This reaction cannot be performed at any conditions since the hydroxide ion (the potential leaving group) is less stable than the nucleophile, the chloride ion. Rather a reversed reaction will proceed:

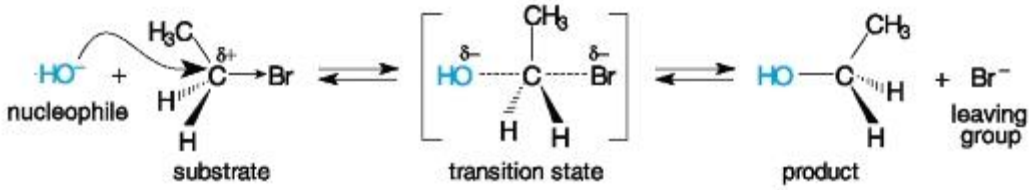


Nevertheless, it is possible to use sodium chloride in converting ethanol to ethyl chloride as we will see in Sec. 7.2.

Problem 7.1. Write an equation for the reaction between each of the following pairs of compounds (assume that the reactions are carried out in aqueous or aqueous-alcoholic solution): (a) benzyl chloride and sodium hydroxide; (b) ethyl iodide and potassium acetate; (c) 1-bromobutane and potassium hydrogensulfide; (d) methyl chloride and trimetylamine.

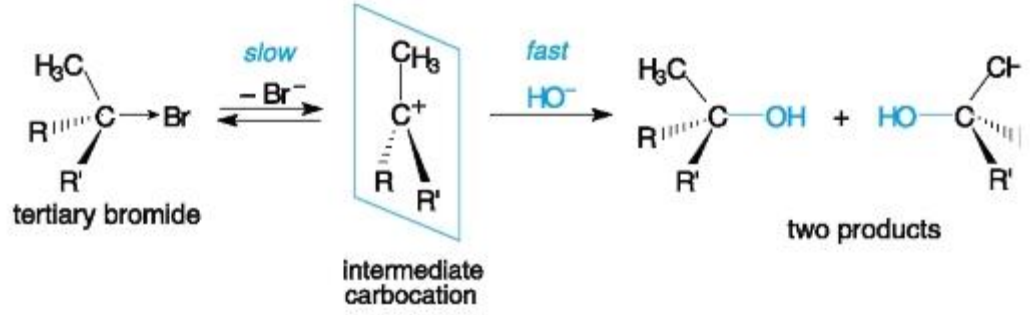
Nucleophilic substitution mechanisms. There are known two mechanisms of this reaction, *bimolecular*and *unimolecular* denoted by the symbols S_N2 and S_N1, respectively. The particular mechanism depends mainly on the structure of a substrate and the nature of a solvent used.

The S_N2 *mechanism* is a one-step process, exemplified below in the alkaline hydrolysis of ethyl bromide:



The hydroxide ion attacks the substrate from the backside of the C-Br bond. The reaction occurs through a transition state, at which the nucleophile and the leaving group are partly bound to the carbon atom. The bromide ion leaves the carbon and the nucleophilic oxygen forms a covalent bond to the same atom simultaneously.

In the S_N2 symbol the number 2 indicates that the reaction is bimolecular, i. e. two molecules are involved in the *only* step of the reaction mechanism. The rate of such reactions depends, as is known, on the concentration of both reactants and can be expressed as Rate = k [Substrate] [Nucleophile].
The S_N1 *mechanism* is a two-step process. It will be considered on alkaline hydrolysis of a tertiary bromide.



In the first, slow step the C-Br bond breaks heterolytically to form a carbocation. In the second, fast step the carbocation combines with the nucleophile. If the tertiary substrate has three different alkyl groups, two products are possible with opposite spatial orientation of the substituents (in more details see Sec. 10.3). This is a result of a planar configuration of sp^2-hybridized carbon in the cation and subsequent nucleophilic attack from either side of the plane.

The number 1 in the S_N1 symbol means that the rate-determining step in the mechanism is unimolecular. The reaction rate does not depend on the nucleophile concentration and is expressed as Rate = *k* [Substrate].

Comparison of the *S_N1* and mechanisms. Why is it necessary to know a substitution mechanism? Firstly, it is very important for substrates that can form configurationally different products (as in S_N1 reactions). Secondly, rate of a particular reaction depends on the mechanism too.

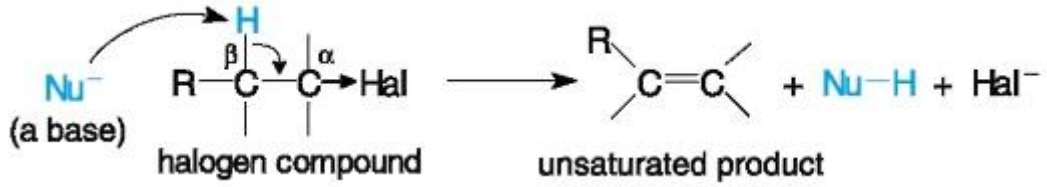
Primary alkyl halides usually react by the S_N2 mechanism. Reactions with tertiary halides proceed almost always by the S_N1 mechanism. This mechanism is also observed in substitution of allylic and benzylic substrates that form resonance stabilized carbocations $CH_2{=}CHCH_2^+$ and $C_6H_5CH_2^+$, respectively (Sec. 6.3.2). For this reason allyl and benzyl halides belong to the most reactive halogen derivatives.

Secondary halides may react by either mechanism depending on various factors. Secondary substrates that have substituents of a small volume react mainly by the S_N2 mechanism.

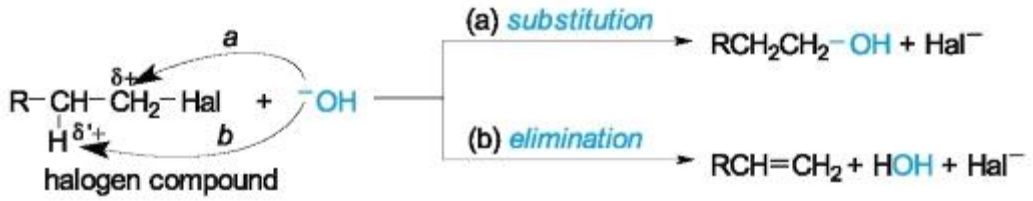
The solvent polarity strongly affects the reaction mechanism. Water, alcohols, and other polar solvents favour the S_N1 mechanism because solvation stabilizes the intermediate carbocation. On the other hand, these solvents are able to form hydrogen bonds to a nucleophile, decreasing its negative charge and thus stabilizing it.

7.1.2. Elimination Reactions

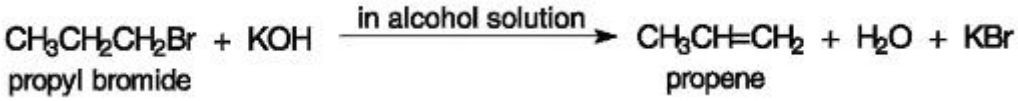
Now let us recall a CH-acidic site at the β-carbon of an alkyl halide. A nucleophile with strong basic properties is capable not only of expelling a halide ion from a fragment $-CH_2CHX-$, but also of splitting off a proton from the β-carbon. Elimination of HX, or *dehydrohalogenation,* is observed in this case with the formation of a double bond.



Often elimination and substitution reactions occur simultaneously.



Both paths of the competing reactions strongly depend on the structure of the substrate, the structure of the nucleophile, and the reaction conditions. Different products may be obtained even with the same pair of reactants. For instance, treatment of alkyl halides with aqueous alkali results mainly in alcohol formation (the nucleophilic substitution reaction, see above the path a) and only a negligible amount of an alkene is formed. When ethanol, a less polar solvent, is used for the same reactants, alkene is a predominant product (the elimination reaction, the path b):



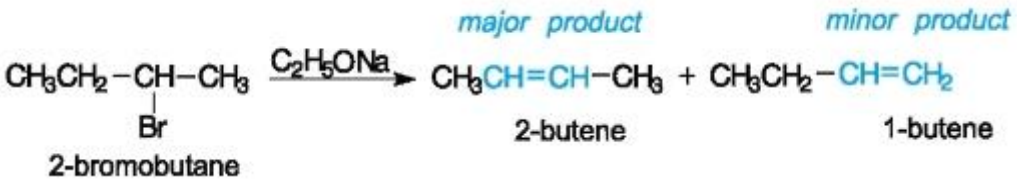
Unsymmetrical halides with non-terminal position of a halogen atom gave a mixture of elimination products. The Russian chemist A. Zaytsev formulated (in 1875) a rule known as *ZaytseVs rule:*

When alternative elimination products from alkyl halides (or alcohols) are possi

ble, the more highly substituted alkene predominates.

Example 7.3. Predict the products of the reaction of 2-bromobutane with sodium ethoxide in ethanol.

Solution. The ethoxide ion is a strong base, therefore it promotes the elimination reaction. According to Zaytsev's rule, 2-butene (the more substituted alkene) and 1-butene are formed with predominance of the former in the ratio of 4:1.



A trace product, not shown in the equation, is the ether $C_2H_5OCH(CH_3)CH_2CH_3$, a result of the competing substitution reaction.

7.2. ALCOHOLS, PHENOLS, AND THIOLS

Alcohols and phenols are compounds that have a hydroxyl group, -OH,

which is attached to a saturated carbon in alcohols or to an aromatic ring in phenols. They have the general formulas R-OH and Ar-OH, respectively.

Compounds of these classes occur widely in nature. A hydroxyl group is literally the most widespread functional group of naturally occurring compounds. The simplest alcohols have many industrial applications as solvents and raw materials. Ethanol is present in alcoholic drinks.

Thiols are sulfur analogues of alcohols of the general formula R-SH; their func

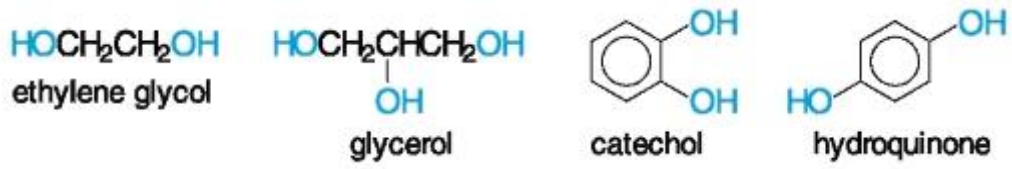
tional group is a mercapto group, -SH.

7.2.1. Classification and Nomenclature

Several classifications of alcohols exist. According to the first one, alcohols are classified as primary, secondary, or tertiary, depending on the number of organic groups bonded to the hydroxyl-bearing carbon atom. Another classification takes into account the nature of a hydrocarbon part of the alcohol molecule. According to this, alcohols fall into saturated, unsaturated, or aromatic ones (an aromatic ring in the latter is not directly bonded to the hydroxyl group). A peculiar type of unsaturated alcohols represents *enols*, unstable compounds which contain the hydroxyl group attached to an sp^2-hybridized carbon, i. e. a fragment =CH-OH. The examples below show various types of alcohols:

Primary	Secondary	Tertiary	Unsaturated	Aromatic	Enolic
$(CH_3)_2CHCH_2OH$	$(CH_3)_2CHOH$	$(CH_3)_3COH$	$CH_2{=}CHCH_2OH$	$C_6H_5CH_2OH$	$[CH_2{=}CHOH]$
2-methyl-1-propanol (isobutyl alcohol)	2-propanol (isopropyl alcohol)	2-methyl-2-propanol (*tert*-butyl alcohol)	2-propen-1-ol (allyl alcohol)	phenylmethanol (benzyl alcohol)	ethenol (vinyl alcohol)

There are also alcohols and phenols that have two or more hydroxyl groups. Alcohols containing two hydroxyl groups are classified as *glycols*. The generic systematic name *diol* is used for dihydroxyl alcohols. The most known examples and trivial names accepted in the IUPAC system are given below:

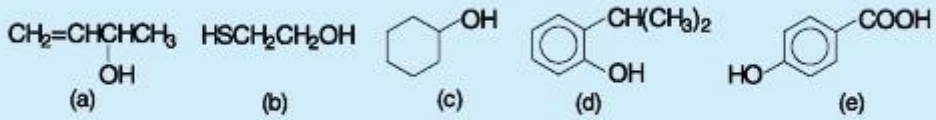


In the IUPAC substitutive nomenclature of alcohols, the hydroxyl group is indicated by the suffix -ol,with elision of terminal -e (if present) from the name of the parent compound. The radicofunctional nomenclature is also used to relatively simple alcohols as shown in the above examples (in parentheses). As regard to nomenclature of phenols, the word phenol is the name both of a class of

compounds and of a specific compound C_6H_5OH. The hydroxyl group is named as the prefix hydroxy-when it occurs in a heterofunctional molecule with substituents of higher priority (Table 2.2).

The same principles of classification and nomenclature are applied to thiols.

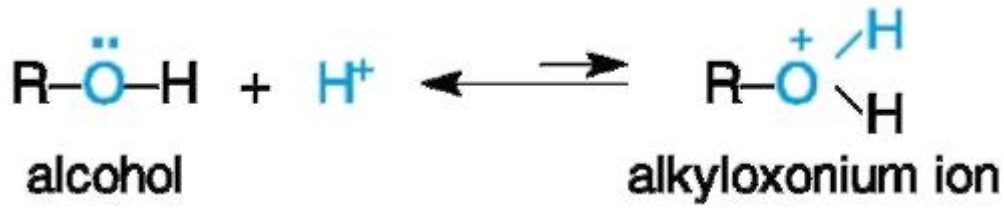
Problem 7.2. Give the substitutive IUPAC names for the following compounds:



Classify the compounds (a) – (c) as primary, secondary, or tertiary alcohol.

7.2.2. Acidic and Basic Properties

Acidic properties of alcohols, phenols and thiols were discussed quite enough in Chapter 5. In the same chapter these classes, especially thiols and phenols, were shown to be weak Bronstedt bases. Nevertheless, alcohols are reversibly protonated by strong acids to give a salt-like product, analogous to the hydronium ion, H_3O^+:



This interaction is the first step in two important reactions of alcohols, namely, nucleophilic substitution and elimination, which will be considered in the following sections.

Being weakly acidic and weakly basic, alcohols, phenols and thiols belong to *amphoteric* compounds which, like water, form hydrogen bonds. The most pronounced intermolecular hydrogen bonding is in alcohols (Sec. 5.4).

Hydrogen bonds account for exceptionally high boiling points of alcohols. Table 7.2 demonstrates boiling points of some lowest alcohols in comparison with those of alkanes, halogenoalkanes, and thiols. The molecules of a close size occupy one line in the table, yet alcohols have the boiling points by 20-60 °C higher than thiols, by 40-80 °C higher than chloroalkanes, and by 80-150 °C higher than alkanes.

Due to hydrogen bonding to water molecules, the lower alcohols (C_1-C_3 and *tert*-butyl alcohol) are completely miscible with it. Phenols show appreciable solubility in water. Solubility in water gradually decreases as the hydrocarbon portion of the molecule lengthens. Such compounds become less polar that results in better solubility in hydrocarbons and other nonpolar solvents[1]. Water solubility of organic substances is of interest to us since water is the medium of the cells of all living organisms.

Table 7.2. Boiling points of alkanes, chloroalkanes, alcohols, and thiols R-Z (°C)

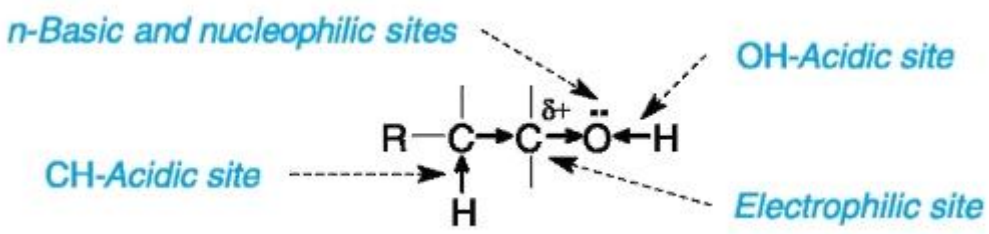
Alkyl group R	*Alkane* (Z = CH_3)	*Chloroalkane* (Z = Cl)	*Alcohol* (Z = OH)	*Thiol* (Z = SH)
CH_3–	–88.6	–23.8	64.5	6.0
CH_3CH_2–	–42.1	12.3	78.4	35.0
$CH_3CH_2CH_2$–	–0.5	46.6	97.4	67.6
$(CH_3)_2CH$–	–11.7	35.7	82.4	52.6
$CH_3CH_2CH_2CH_2$–	36.1	78.4	117.4	98.6

[1] It was well known to ancient Greeks that *similar subjects dissolve similar subjects.*

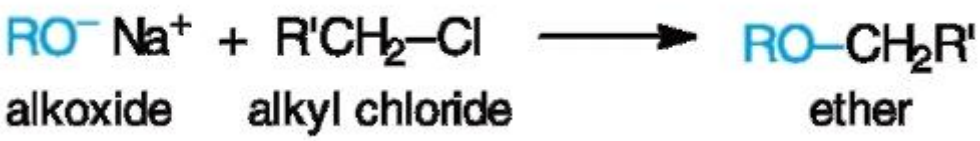
7.2.3. Electrophilic and Nucleophilic Properties

Though the functional groups of alcohols, phenols and thiols are similar in many respects, electrophilic and nucleophilic properties of these classes differ substantially and, therefore, are considered for each class separately.

Alcohols. In addition to the OH-acidic and basic sites discussed in the previous section, alcohols possess an electrophilic, nucleophilic, and CH-acidic sites shown below:

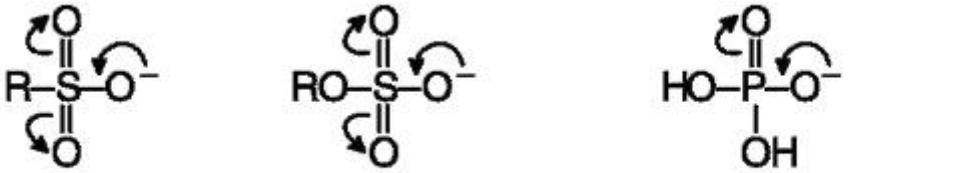


Alcohols are weak nucleophiles owing to high electronegativity of the oxygen atom. An alkoxide ion is a much stronger nucleophile than an alcohol; it reacts with electrophilic substrates such as alkyl halides to form substitution products called *ethers*. This is one of the oldest reactions in organic chemistry known as the *Williamson synthesis* (1851).



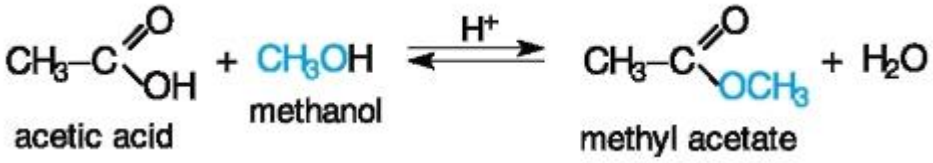
Primary alkyl halides work best, since competitive elimination of hydrogen halide can occur with secondary and tertiary halides (as shown in Example 7.3).

Besides halide ions there are other good leaving groups such as alkyl and aryl sulfonates RSO_3^-, alkyl sulphates $ROSO_3^-$, and dihydrogenphosphate $H_2PO_4^-$. The three anions are well stabilized due to conjugation:



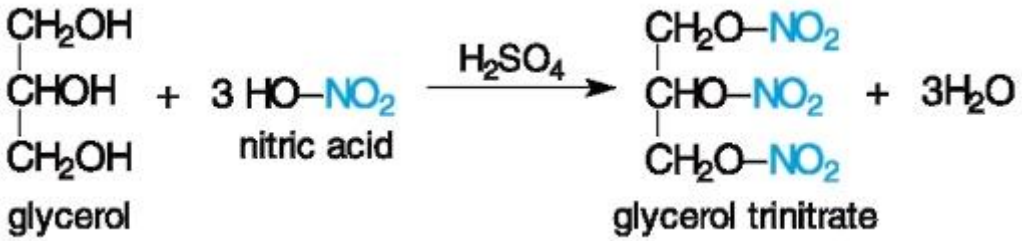
Sulfates and sulfonates are mostly used in laboratory practice; the phosphate ion is an excellent leaving group in biological substrates.

Alcohols as nucleophiles react with carboxylic acids and their derivatives by replacing the OH group of an acid to form *esters*. With respect to an alcohol this reaction is classified as *acylation*, or introduction of an *acyl group*, R-C(O)-. Only the principal equation is given here, the details of the reaction are dealt with in



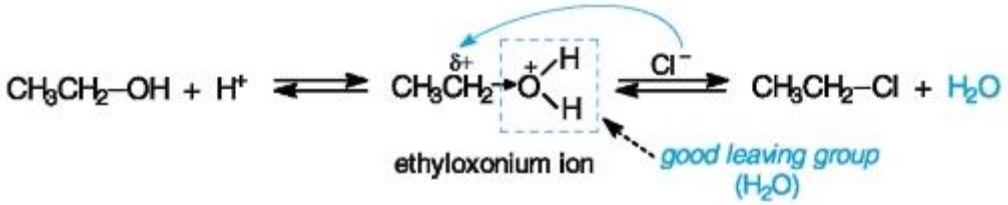
Chapter 9.

Alcohols also react with oxygen-containing inorganic acids (sulfuric, nitric, nitrous, boric), for example:



Glycerol trinitrate is, perhaps, the most known inorganic ester. It is a powerful explosive, the main brisant component of dynamite invented by Alfred Nobel (1865). Glycerol trinitrate is also used in medicine as a vasodilator.

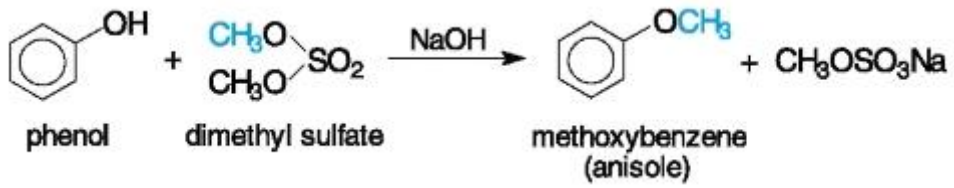
Like haloalkanes, alcohols are potential substrates in reactions with nucleophilic reagents. However, one must keep in mind that the hydroxyl group is a bad leaving group. Nevertheless, it is possible to substitute the OH group after its conversion into protonated (oxonium) form. Thus the protonated substrate is much more reactive than the neutral alcohol. A product is readily formed because of expelling a better leaving group, a water molecule.



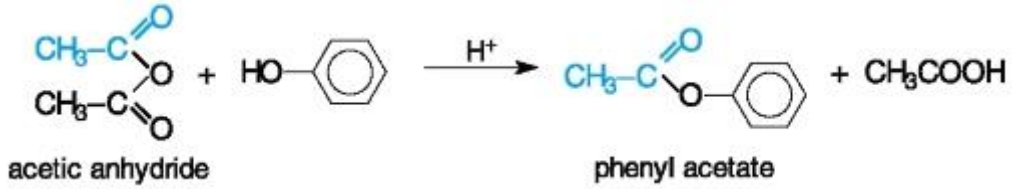
Primary alcohols undergo such substitution similarly to that shown in Sec. 7.1.1 for alkyl halides. The reaction proceeds slowly on heating an alcohol with

concentrated hydrochloric acid. Tertiary alcohols, on the contrary, easily react at room temperature to give the corresponding halide.

Phenols. Nucleophilic strength of phenols is even lower than that of alcohols in a neutral medium. But in alkaline media phenols are easily converted into phenoxide ions (Sec. 5.2.2) that possess an increased nucleophilic activity. Thus phenols can be alkylated with alkyl halides or other alkylating agents, for example:



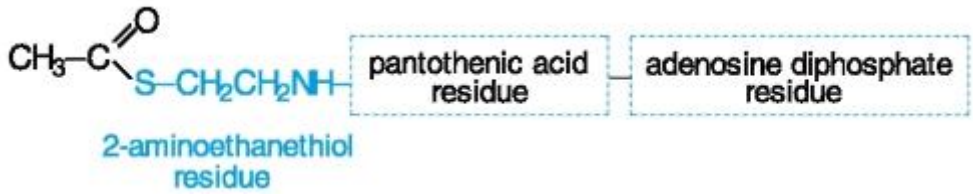
In contrast to alcohols, phenols do not form esters in the reaction with carboxylic acids. More powerful acylating agents, such as acid anhydrides, have been used for this purpose:



Again in contrast to alcohols, in phenols the C-O bond is difficult to break and replace the hydroxyl group, for instance, by a chlorine atom. The reason is that this bond is about 50 kJ/mol stronger than the Csp^3-O bond in alcohols (see Appendix 1). This can be explained by *p,π*conjugation of the oxygen lone pair of electrons with the benzene ring.

Thiols[1]. Compounds of this class are the most nucleophilic ones in the discussed group (alcohols, phenols, and thiols). This is the result of low electronegativity and high polarizability of the sulfur atom.

As strong nucleophiles, thiols readily react, especially in the presence of an alkali, with alkyl halides to form *sulfides,* RSR', earlier called thioethers. The reaction of thiols with acid anhydrides leads to*thioesters,* RC(O)SR'. One of the most important transfer reagents in biochemical processes is *acetyl coenzyme A,* a thioester of the following schematic structure (see the full structure in Sec. 17.4.2):

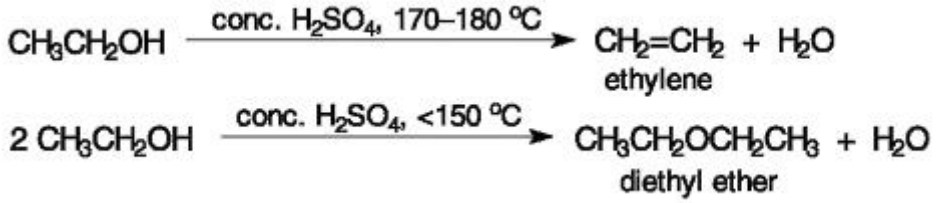


A distinctive organoleptic characteristic of thiols is their intense and extremely disagreeable odor. Negligible amounts of ethanethiol are added to odorless natural

gas for detection of a gas leak. A threshold feeling concentration for lower thiols is 2×10^{-6} mg/m^3 (if you know sharp odor of ammonia and unpleasant odor of hydrogen sulfide, the odor of rotten eggs, compare their values of 7 and 1.4 mg/m^3, respectively).

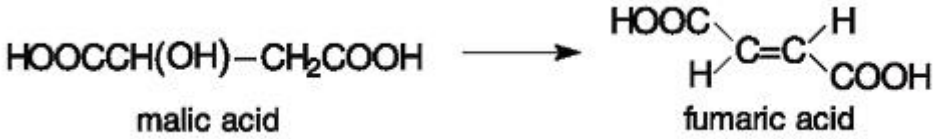
7.2.4. Elimination Reactions

Alcohols can eliminate a water molecule on heating with a strong acid to yield alkenes. The reaction is classified as *dehydration.* Ethers, as by-products of a competitive substitution reaction, are produced only in a small amount. Ethers are, however, the main products when the reaction is carried out at lower temperatures. For example:



Acid-catalyzed dehydration usually follows the Zaitsev's rule and yields predominantly the more substituted alkene. The ease of alcohol dehydration is the same as the order of carbocation stability, i. e. primary < secondary < tertiary.

Elimination is a well-known *in vivo* reaction. For example, enzymic dehydration of malic acid (a hydroxy acid) results in the formation of fumaric acid:

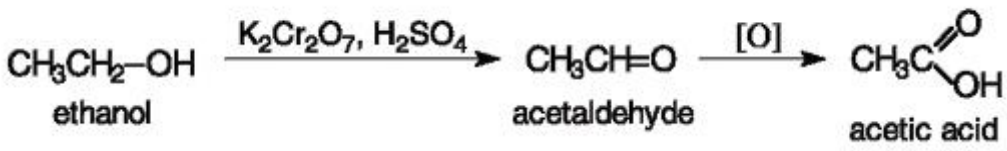


Problem 7.3. Indicate the major product of each of the following reactions:
(a) $CH_3CH_2CH_2CBr(CH_3)_2 + C_2H_5ONa \longrightarrow$
(b) $CH_3CH(OH)CH_2CH_3 + H_2SO_4 \longrightarrow$

[1] This term originates from the Greek *theion* - sulfur.

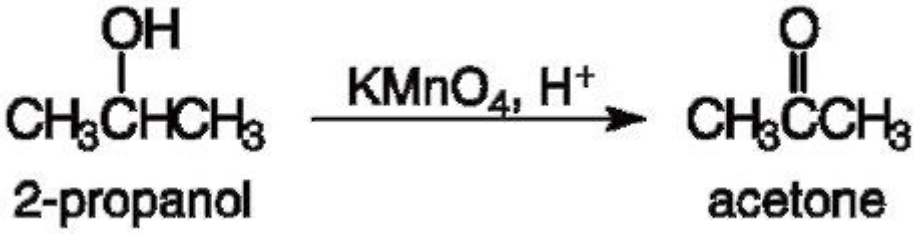
7.2.5. Oxidation Reactions

Oxidation of alcohols. Primary and secondary alcohols can be oxidized with potassium permanganate or potassium dichromate, $K_2Cr_2O_7$, in a dilute acid. Primary alcohols form aldehydes, but further oxidation to carboxylic acids usually occurs:

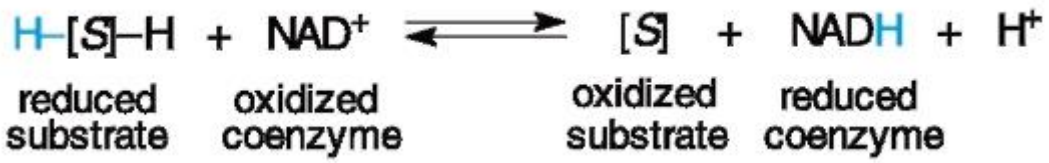


Secondary alcohols are oxidized to ketones. Further oxidation to carboxylic acid does not occur except under severe conditions. Tertiary alcohols are normally

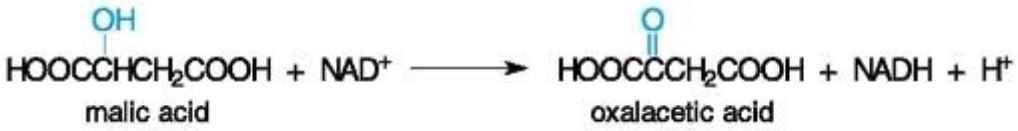
resistant to oxidation.



In plants and animals, similar oxidations are accomplished enzymically with a coenzyme *nicotinamide adenine dinucleotide,* abbreviated NAD+, and its phosphate, NADP+ (for their structures see Sec. 17.4.2). The role of NAD+ formally consists in abstraction of a hydride ion from a substrate schematically shown as H[S]H:

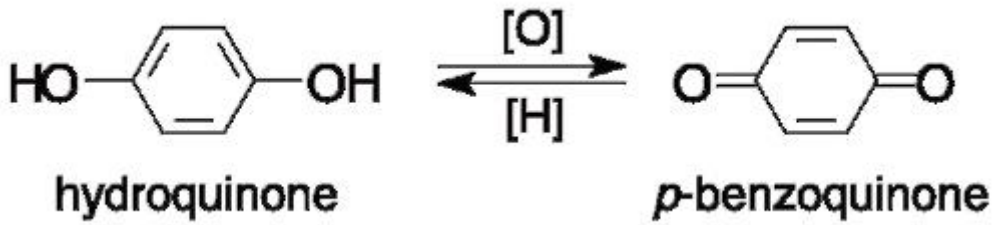


Oxidation of malic acid into oxalacetic acid (an oxo acid) with NAD^+ in the Krebs cycle is an example of numerous biochemical oxidation reactions:



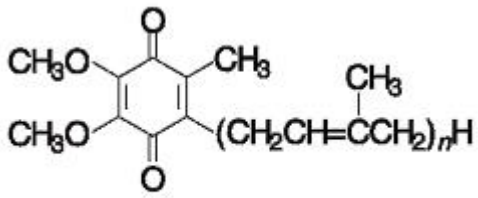
Problem 7.4. Which of the following alcohols will be oxidized with permanganate (without heating): (a) butyl alcohol; (b) isobutyl alcohol; (c) secbutyl alcohol; (d) tert-butyl alcohol? Illustrate with equations and name the products.

Oxidation of phenols. Phenolic compounds, especially those containing two or more hydroxyl groups, are susceptible to oxidation, even by atmospheric oxygen. Aromatic 1,2- and 1,4-dihydroxyl compounds are oxidized to cyclic unsaturated diketone known as *quinones.* The most significant property of quinones is their reversible reduction to the corresponding dihydroxybenzenes.

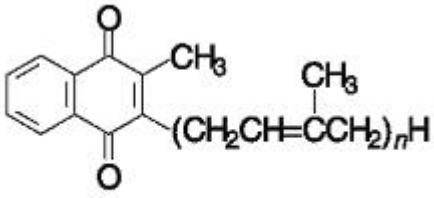


Quinones of many kinds are important compounds both because of their prevalence in nature as the products of plant and animal metabolism and because of their use in medicine. A group of quinones called *coenzymes Q* (or *ubiquinones* which are so named because of their ubiquitous, or widespread

occurrence in nature) serves as electron-carriers (oxidizing agents) to mediate the respiration process. Vitamins of group K, which are required for the normal clotting of blood, represent derivatives of 1,4-naphthoquinone.

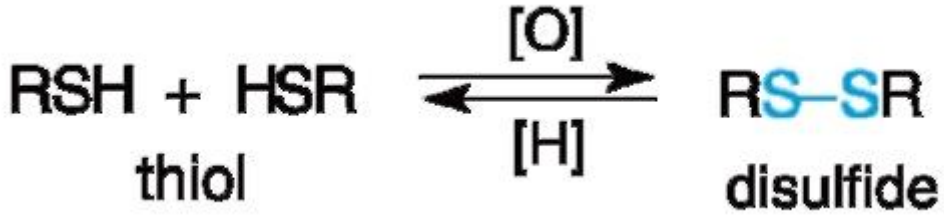


coenzymes Q (ubiquinones, $n = 6–10$)



vitamins K_2 ($n = 4–10$)

Oxidation of thiols. Mild oxidation of thiols (with hydrogen peroxide, bromine, iodine, or slowly with molecular oxygen) results in the formation of *disulfides*, RSSR. Note that the S-H bond is broken in this case, but not the C-H bond as it occurs in alcohols. The reaction is easily reversed, i. e. disulfides can be reduced to thiols.



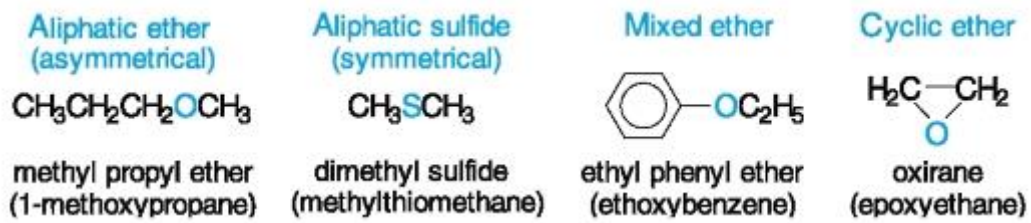
Easiness and reversibility of reactions in the system thiol-disulfide play a significant role in the formation of three-dimensional structure of proteins (Sec. 15.2.3).

Stronger oxidizing agents, such as nitric acid or permanganate, convert thiols to *sulfonic acids*, RSO_3H.

7.3. ETHERS AND SULFIDES

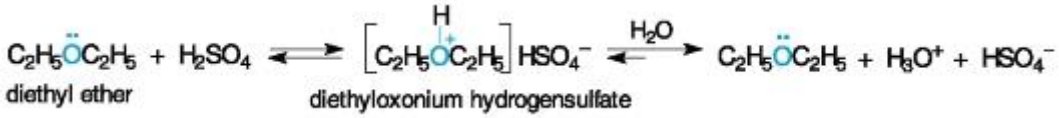
Ethers are compounds that have two organic groups connected to the same oxygen atom. Sulfides have two organic groups attached to a sulfur atom. The general formulas of ethers and sulfides are R-O-R' and R-S-R', respectively.

The radicals R and R' may be identical or different, aliphatic or aromatic. There are also cyclic representatives that contain a heteroatom as a part of the ring system. The most important of them are three-membered ethers, *oxiranes*, sometimes called *epoxides.*

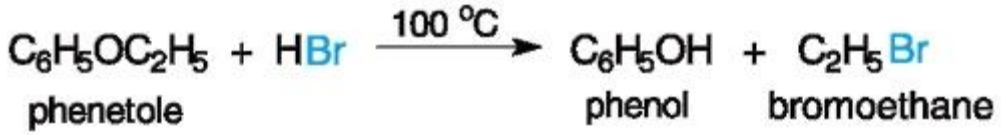


The radicofunctional IUPAC nomenclature is more convenient for relatively simple ethers and sulfides. In the substitutive nomenclature the prefix alkoxy- (shortening from *alkyloxy-)* shows the presence of the OR group at the parent structure of ethers; the prefix alkylthio- indicates the SR group in sulfides. Cyclic ethers and sulfides are usually named as heterocycles (Chapter 16). The examples of the name are given just above, the substitutive names in parentheses.

Ethers and sulfides are quite inert to most common reagents. They do not react with bases because they have very weak acidic sites. The oxygen atom of ethers makes them slightly basic. Ethers, being water-insoluble compounds, can be dissolved in concentrated mineral acids to form unstable*dialkyloxonium salts* that are easily hydrolyzed in the presence of water.



Ethers however react with concentrated hydroiodic or hydrobromic acid on heating to give cleavage products at one of the C-O bonds. If a mixed ether (as phenetole in the example below) is subjected to cleavage, one of the products is always phenol.



In this reaction aliphatic asymmetrical ethers ROR' give various alcohols and alkyl halides depending on the nature of radicals R and R'.

Example 7.4. Explain why dialkyl ethers are cleaved on heating with strong acids but are stable on treatment with weak acids or in neutral media.

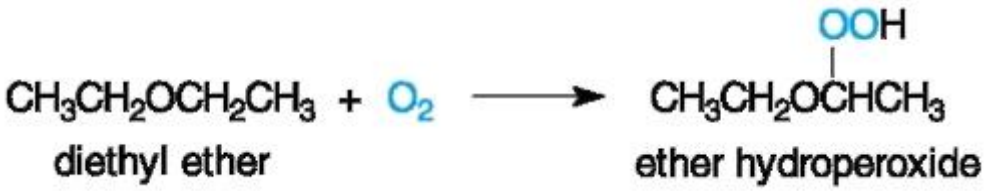
Solution. Ethers are protonated in strong acids as shown above. Thus, in a nucleophilic substitution reaction, a leaving group is an alcohol molecule.

In neutral media or in weak acids protonation does not occur, and a leaving group would have to be an alkoxide ion.

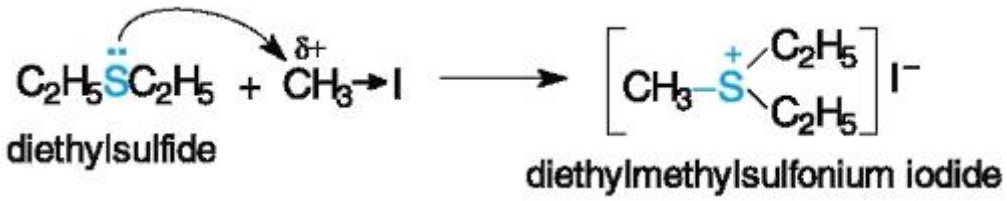
The ROH molecule is a good leaving group, whereas the RO^- ion is a very bad one.

Inertness and high solvation capacity of ethers make them excellent solvents for many organic reactions and for extracting organic compounds from natural sources (this does not relate to sulfides that, like thiols, have very unpleasant odor).

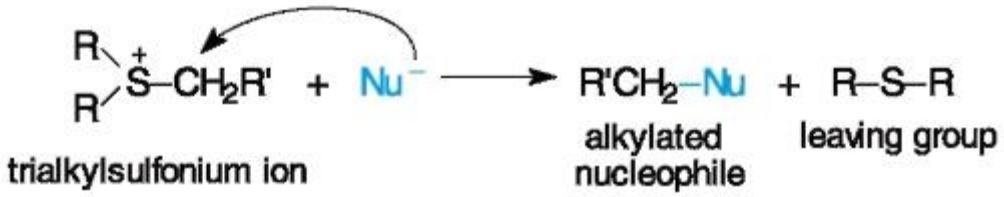
Two risks exist on handling ethers. Firstly, low-molecular weight ethers are highly flammable compounds. Care should be taken when diethyl ether (the most often applied in chemistry and medicine and called simply the ether) is used because open flames or even sparks from light switches can cause explosive combustion of its mixture with air. Secondly, most ethers that have been exposed to air for a long time are slowly oxidized to form organic peroxides, compounds with a -O-Obond; for example:



Organic peroxides are extremely explosive and, if they have been detected in ethers, must be removed before use. Treatment with Fe^{2+} or Sn^{2+} salts destroys the peroxides by reduction.
Sulfides have virtually no basic properties; they are at the same time much more nucleophilic than ethers. The reason is that sulfur is more polarizable than oxygen. Thus, sulfides are good nucleophiles that react with alkyl halides to yield stable crystalline *trialkylsulfonium salts*, R_3S+X^-:



Trialkylsulfonium salts can be subjected to nucleophilic attack at one of the alkyl groups attached to the positively charged sulfur. A nucleophile replaces a sulfide molecule as a leaving group. With respect to the nucleophile this is an alkylation reaction (examples of biochemical alkylation see in Sec. 7.5):



Problem 7.5. Give an explanation why HI and HBr are much more effective than HCl in ethers cleavage.

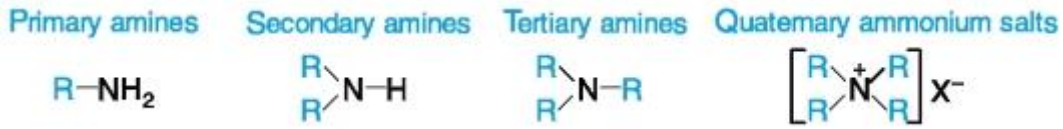
7.4. AMINES

Amines are organic derivatives of ammonia that contain one or more organic groups instead of hydrogen atom(s). The general formulas of amines are RNH_2, R_2NH, and R_3N.

Perhaps amines are of less occurrence in nature than oxygen-containing compounds, they are, however, of extreme importance in vital processes. A great number of biologically and medically important compounds belong to amines. Suffice it to recollect the term *vitamins* originated from the Latin *vita* (life) + amine[1].

7.4.1. Classification and Nomenclature

Amines are classified as primary, secondary, or tertiary depending on the number of organic groups replacing the hydrogen atoms of the ammonia molecule. Notice that the classification of amines differs from that previously used for alcohols and alkyl halides. There are also compounds with four substituents at a positively charged nitrogen. Such compounds mentioned in Sec. 7.1.1 are called*quaternary ammonium salts* (analogues of ammonium salts, $NH_4^+X^-$).



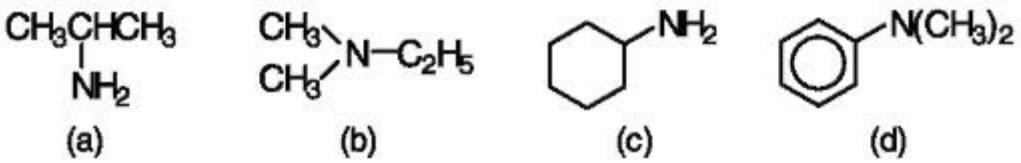
The R groups in these structures may be aliphatic or aromatic, identical or different (when two or more R groups are present).

[1] Though not all vitamins are amines, actually most of the first discovered vitamins were of this sort.

In the IUPAC nomenclature, the group $-NH_2$ is named by the suffix -amine, with elision of terminal -efrom the name of the parent compound. When a substituent of higher priority is also present, groups $-NH_2$, -NHR and $-NR_2$ are named by the prefixes amino-, alkylamino-, and dialkylamino-, respectively. Aromatic amines are named in the same manner as derivatives of the simplest representative $C_6H_5NH_2$ that have the parent name aniline.

Another principle, which is similar to the radicofunctional nomenclature, may be used to relatively simple amines. In this system, amines are named by adding the word -amine to the name of the radical(s).

Example 7.5. Give acceptable names for the following amines and classify them as primary, secondary or tertiary:



(a) (b) (c) (d)

Solution. The compound (a) is a secondary amine, the compounds (b) and (d) are tertiary amines, and the remaining (c) is a primary one.

The names are: (a) 2-propanamine; or isopropylamine; (b) ethyldimethylamine; (c) cyclohexanamine, or cyclohexylamine; (d) N,N-dimethylaniline. The symbol *N*- in the latter shows substitution at the nitrogen atom in aniline, but not at carbon as is common in substitutive nomenclature.

7.4.2. Chemical Properties

The most important chemical characteristics of amines are their basicity and nucleophilic properties. Both are a consequence of the presence of the nitrogen lone pair of electrons.

Basicity of organic compounds including amines was discussed in Sec. 5.3. Here it should be added that primary and secondary amines despite their very weak acidity can form a hydrogen bond of the type N-H

•

•

• N which is not so strong as O-H

•

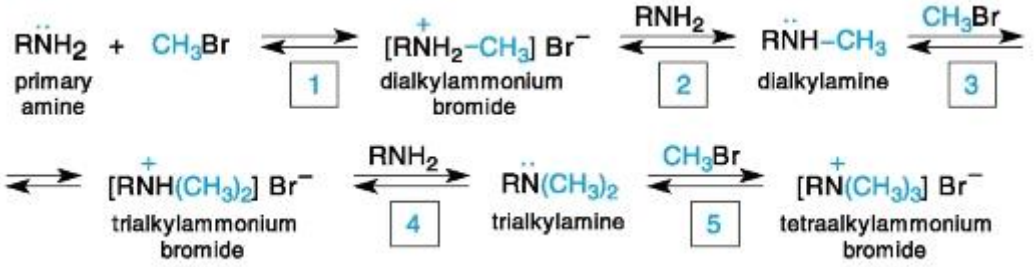
•

• O bonds, for example, in alcohols. Thus the boiling points of amines are much lower than those of comparable alcohols. Nevertheless, hydrogen bonds involving nitrogen are exclusively important in the formation of spatial structure of proteins and nucleic acids (Chapters 15 and 17).

Problem 7.6. Complete the following equation, and name the product:
$C_6H_5NH_2 + H_2SO_4$ (dilute) ⟶

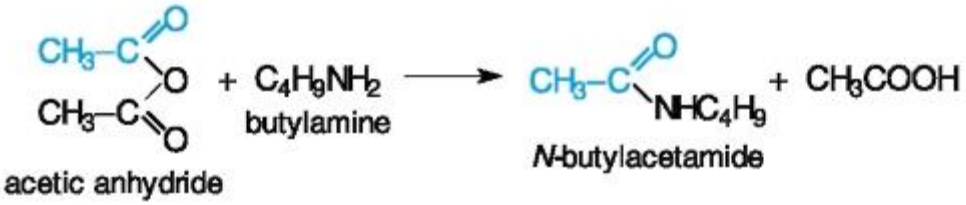
Lower amines dissolve in water to give alkaline solutions. Water acts as a protic acid and donates a proton to an amine in this reversible reaction:

Nucleophilic properties of amines appear in their reactions with electrophilic substrates such as alkyl halides. Treatment of a primary amine with an alkyl halide results in alkylation of the nitrogen to form a dialkylammonium salt (see below, reaction 1). The latter is partially deprotonated (reaction 2) by the starting amine (as a base). Because the resulting dialkylamine is also a nucleophile, it will react with methyl bromide (reaction 3), and so on (reactions *4* and 5). Thus, a mixture of products is formed, and the method has some limitations for the laboratory use.

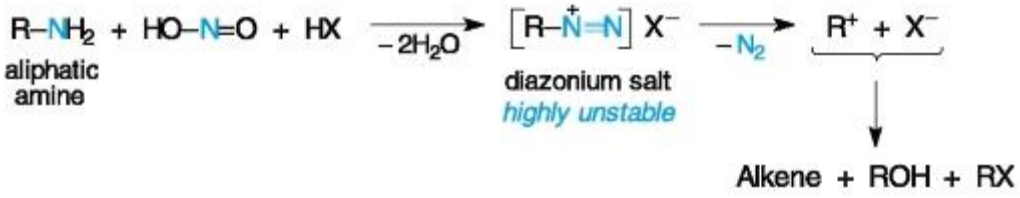


Other examples of amine alkylations were presented in Sec. 7.1.1, including Table 7.1. In a principally similar way proceeds biochemical alkylation of amines and other nucleophilic substrates that is considered in the following section.

Primary and secondary amines can also be acylated by reaction with active carboxylic acids derivatives. A product of the reaction is an N-substituted *amide* (for details see Chapter 9):



Important transformations undergo primary amines on treatment with nitrous acid, HNO_2. This acid is an unstable compound which is usually prepared *in situ* from sodium nitrite and a strong acid. Primary aliphatic amines react with nitrous acid to yield highly unstable *diazonium salts* that decompose spontaneously even at low temperature by removal of the nitrogen molecule. The resulting carbocation produces a mixture of alkenes, alcohols, and alkyl halides, when it eliminates a proton or reacts with water or with X^-. Such reactions are referred to as *deamination* reactions (the loss of an amino group).



Secondary amines react with nitrous acid to yield N-nitrosoamines, some of them are strong carcinogens.

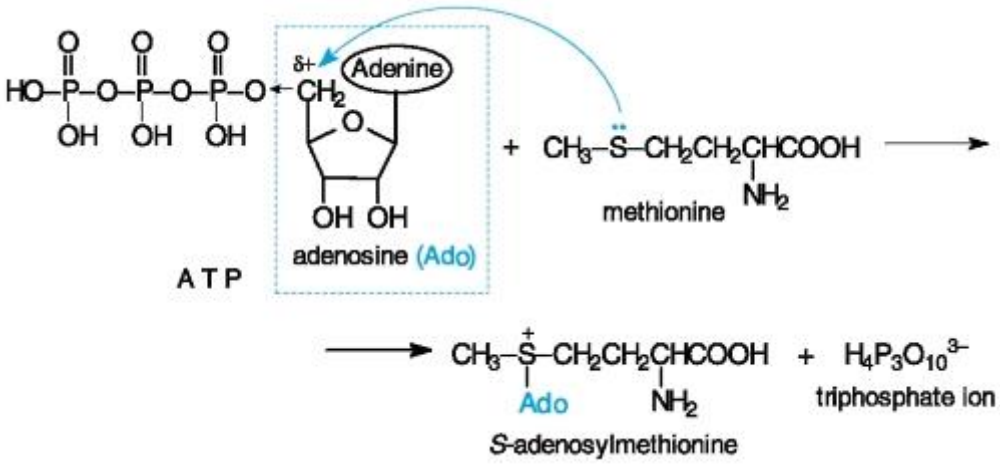
How may nitrites arise in the human organism? Nitrates, which are present to some degree in vegetables, can be reduced in the cells to nitrites. The latter can react with aminecontaining components of food as described above. Nitrous acid is also a dangerous reagent to primary amines present in animal cells. As we will see later (Chapter 17) nucleic acids contain aromatic amino groups, and their modification with nitrous acid can result in genetic distortions.

7.5. BIOCHEMICAL ALKYLATIONS

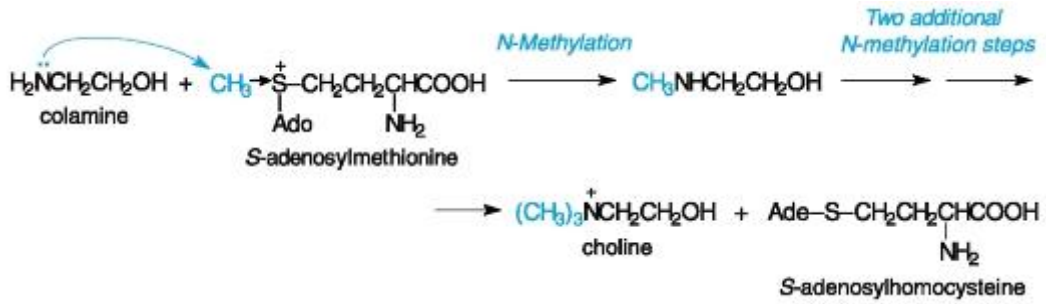
Many metabolic processes occur similarly to reactions carried out in the laboratory. Biochemical methylation is, perhaps, the most common of all substitution reactions in living matter.

In the cell, phosphoric acid esters participate as substrates in which good leaving groups, such as phosphate, diphosphate and triphosphate ions can easily be eliminated. Let us consider *in vivo* reaction with the assistance of adenosine triphosphate (ATP) and sulfur-containing amino acid *methionine.*

In the first step, methionine, as a sulfur nucleophile, attacks an electrophilic carbon of ATP with displacing a good leaving group, the triphosphate ion, by an S_N2 mechanism. This step represents a biological similarity of the reaction between a dialkyl sulfide R_2S and an alkyl halide RX to yield a trialkylsulfonium salt $R_3S^+X^-$. Methionine is S-alkylated in this step to form S-adenosylmethionine (SAM):



In the second step, the complex molecule SAM acts as a methyl-group donor in alkylation of N- or O-containing nucleophiles. For example, N-methylated product is formed in the reaction with a biogenous amine colamine (2-aminoethanol).



Thus, three similar alkylation steps with S-adenosylmethionine give *choline*, a participant of various metabolic processes. Choline is also the precursor of acetylcholine that plays an essential role in the transmission of nerve impulses.

At first sight, these reactions may seem complicated because of the complexity of the reagent structures. But these reactions well illustrate chemical principles discussed in this chapter and conceptually they are simple.

Additional Problems

7.7. How would you predict the result of treating methyl chloride with:

(a) diethylamine;

(b) tributylamine;

(c) diethyl sulfide.

7.8. Show, by giving equations, how the following conversions may be effected:

(a) ethyl bromide into ethanol;

(b) isopropyl bromide into propene;

(c) ethyl iodide into ethylmethylamine.

7.9. Why is benzyl chloride more reactive than hexyl chloride in nucleophilic substitution reactions? Exemplify by the hydrolysis reaction.

7.10. How might you effect the transformation of allyl chloride (3-chloropropene) to glycerol, using only inorganic reagents? More than one step is necessary.

7.11. Suggest a chemical test, using inorganic reactants, to distinguish between benzyl chloride and chlorobenzene.

7.12. Combine addition and substitution reactions to devise a two-stage synthesis of ethyl isopropyl sulfide, starting with propene.

7.13. Show how you might prepare tert-butyl ethyl ether from ethanol and tert-butyl alcohol, using only inorganic reagents? More than one step may be necessary.

7.14. Ethyleneglycol is a stronger acid than ethanol as it was shown in Sec. 5.1.2. Which is a stronger base, ethylenediamine, $H_2NCH_2CH_2NH_2$, or ethylamine? Explain your reasoning.

Chapter 8. CARBONYL COMPOUNDS

Aldehydes and ketones have the *carbonyl group* >C=O as a functional group and they are classified as carbonyl compounds. The carbonyl group is also a constituent of other functional groups generally expressed as -C(O)Z, where Z is a substituent that has an atom with the unshared electron pair, i. e. halogen or oxygen-, nitrogenor sulfur-containing groups. Such compounds are known rather as carboxylic acids and their derivatives, though they are sometimes referred to as carbonyl compounds too. We will discuss them later, and this chapter deals with aldehydes and ketones.

Aldehydes are compounds in which the carbonyl group is bonded to at least one hydrogen. Ketones contain the carbonyl group bonded to two hydrocarbon residues. The general formulas of aldehydes and ketones are RCH=O and RC(O)R' (or simply RCOR'), respectively.

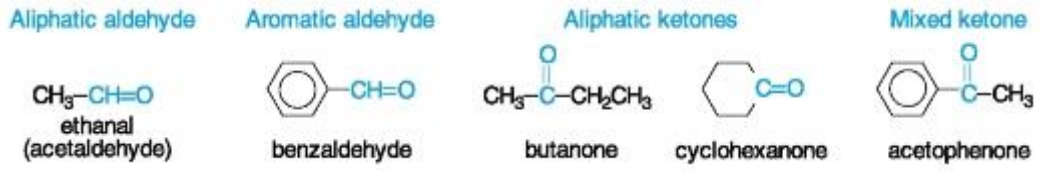
Carbonyl compounds are among the most important of all organic compounds, especially, in biological processes. Suffice it to say that carbohydrates contain a carbonyl group.

8.1. GENERAL CHARACTERISTICS OF ALDEHYDES AND KETONES

Chemical properties of aldehydes and ketones are very similar regardless of their specific structure. It is no wonder that both of the classes are considered together.

8.1.1. Classification and Nomenclature

Aldehydes may be classified as either aliphatic or aromatic depending on the R substituent. The only exception is the simplest representative - formaldehyde, HCH=O, where R is hydrogen in the general formula of aldehydes. Ketones are classified in a similar way, except for existence of cyclic ketones and a mixed type of the compounds.



In the IUPAC nomenclature system, the aldehyde group in aliphatic compounds is indicated by the suffix -al, with elision of terminal -e from the name of the parent hydrocarbon. Note that carbon of the aldehyde group is a part of the hydrocarbon structure. The parent name benzaldehyde is used for aromatic aldehydes. Many aldehydes have semi-trivial names derived from the corresponding carboxylic acids and accepted by the IUPAC nomenclature, for example *acetaldehyde* - from *acetic acid* (as shown above), *butyraldehyde*, $CH_3CH_2CH_2CH=O$ - from *butyric acid* (C_4 acid), and so on.

Ketones are named (see examples above) using the suffix -one which, actually, designates a group =O called the *oxo group*. If higher functional group is present in the same molecule, the prefix oxo- is used for the group =O. The simplest ketone $CH_3C(O)CH_3$ has the trivial name acetone.

8.1.2. Electronic Structure of the Carbonyl Group

Both carbon and oxygen atoms are sp^2-hybridized in the carbonyl group. Like a C=C double bond, a carbon-oxygen double bond represents a combination of a σ bond, formed by overlap of hybrid orbitals, and a π bond, formed by lateral overlap of unhybridized *p* orbitals (Fig. 8.1, a). Three σ bonds of the carbon atom, as well as two unshared electron pairs on the oxygen atom (shown in the figure by dots), lie in the same plane with bond angles of approximately 120°.

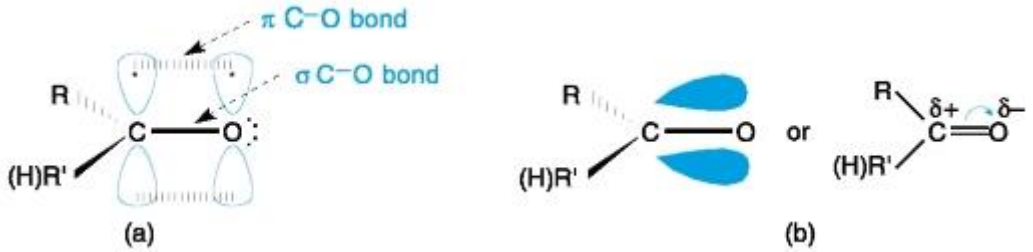
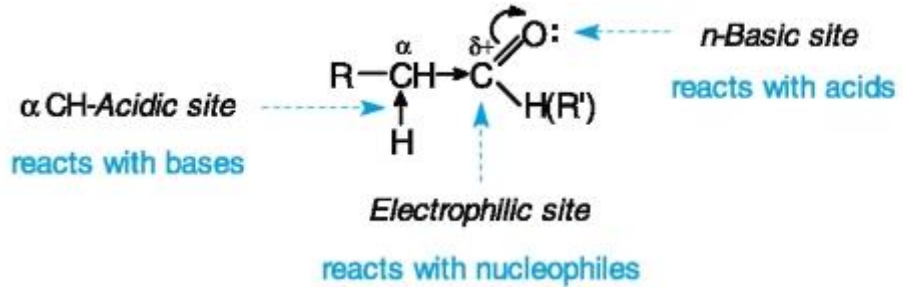


Figure 8.1. Bonding in the carbonyl group: the *p* orbital overlap (a) and polarization of the C=O bond (b).

Unlike alkenes, the C=O double bond, especially the π bond, is highly polarized because oxygen is much more electronegative than carbon. Thus, the carbonyl carbon carries a partial positive charge and, conversely, the oxygen atom carries a partial negative charge (Fig. 8.1, *b).*

Several reaction sites can be emphasized as a consequence of the carbonyl group polarization:

- the carbonyl carbon as an electrophilic site which can be attacked by nucleophiles;
- the oxygen atom as a weak *n*-basic site that can be protonated with strong acids;
- the α CH-acidic site is a weak acidic site that can be deprotonated with strong bases.



Additional reaction sites may arise when a double bond or an aromatic ring is present in the hydrocarbon portion of a carbonyl compound.

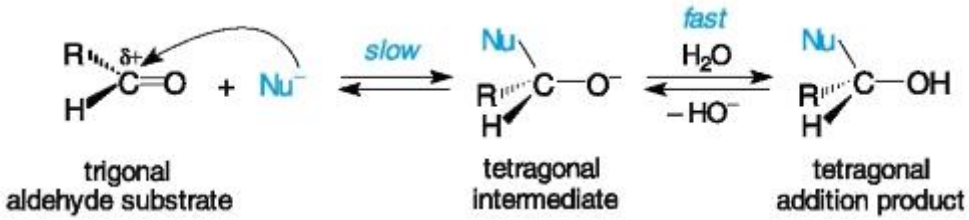
Aldehydes and ketones possess very weak acidity and basicity, therefore they cannot form intermolecular hydrogen bonds. Carbonyl compounds are more volatile than the corresponding alcohols. Compare, for example, boiling points of propanal (49 °C), acetone (56 °C), and 1-propanol (97 °C).

8.2. NUCLEOPHILIC ADDITION REACTIONS

Carbonyl compounds are susceptible to be attacked by nucleophiles. They undergo nucleophilic addition (A_N) reactions rather than substitution because they have very bad potential leaving group attached to the carbonyl carbon atom, i. e. anions H^- or R^-. The nucleophilic addition is the most important reaction of aldehydes and ketones.

A nucleophile that attacks the electrophilic carbon can be either negatively charged ($Nu:^-$) or neutral (Nu:). A hydroxide ion, alkoxide ions RO^-, and a hydride ion H^- are the examples of charged nucleophiles. To the neutral nucleophiles there belong water, alcohols, ammonia, and amines.

A nucleophile approaches the carbon atom from a direction approximately perpendicular to the plane of the carbonyl group. In the first, slow step of the reaction, the nucleophile uses its electron pair to form a new C-Nu bond. This step is accompanied by carbon rehybridization from $sp^2$to sp^3, resulting in a tetrahedral alkoxide ion intermediate. When the reaction is carried out in a protic solvent such as water or alcohols, the reaction is completed (in the second, fast step) by addition of a proton to the negatively charged oxygen.



In general, aldehydes are somewhat more reactive than ketones towards nucleophiles because organic groups R' in ketones are both larger and more electrondonating than the hydrogen atom in aldehydes. Thus, the carbonyl carbon is more hindered in ketones, and the partial positive charge on this atom is also reduced. For the same reasons aromatic aldehydes and ketones are less reactive than their aliphatic counterparts.

Problem 8.1. Explain, using arrow symbolism, higher reactivity of acetaldehyde in comparison with benzaldehyde.

Substituents with electron-donating effect (inductive or mesomeric) in the R portion of carbonyl compounds decrease reactivity in nucleophilic addition reactions. On the contrary, substituents with electron-withdrawing effect increase such reactivity. For example, trichloroacetaldehyde, $CCl_3CH=O$, known as chloral, is more reactive than acetaldehyde, whereas 4-methoxybenzaldehyde is less reactive than benzaldehyde.

Next, typical nucleophilic addition reactions will be considered according to the type of a new bond formed to the carbonyl carbon.

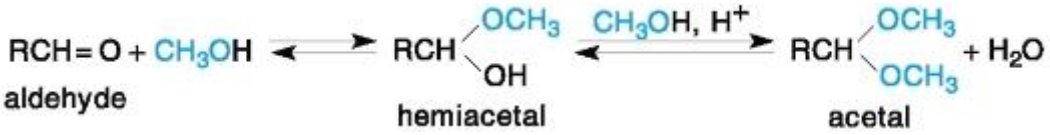
8.2.1. Addition of Alcohols: Hemiacetal and Acetal Formation

Aldehydes and some ketones react reversibly with alcohols to yield stepwise *hemiacetals* and *acetals.*

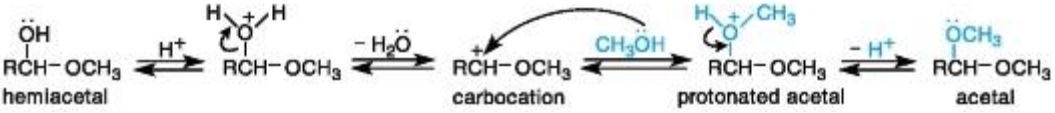
Hemiacetals are compounds in which a hydroxyl group and an alkoxyl

group are attached to the same carbon. Acetals have two alkoxyl groups at the same carbon.

The overall transformation of an aldehyde into an acetal involve two steps: a) nucleophilic addition of an alcohol and b) nucleophilic substitution of the OH group for an alkoxyl group of the second alcohol molecule. Most open-chain hemiacetals are not sufficiently stable to be isolated.

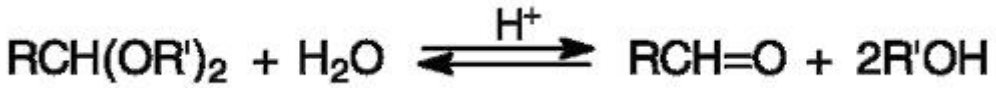


An acid catalyst is required in the second step of the reaction since alcohols are weak nucleophiles and the hydroxyl group is a poor leaving group.



Protonation converts the hemiacetal hydroxyl into a good leaving group, a water molecule. Dehydration yields an intermediate carbocation, that reacts with excess alcohol to give a protonated acetal. Loss of a proton results in a neutral acetal product.

Note that all the steps of acetal formation are reversible. To drive the process in the forward direction it is necessary to use a large excess of an alcohol or to remove water, a product of the forward reaction (remember one of the Le Chatelier's principles). On the other hand, the reverse reaction is favoured when acetal is treated with a large excess of water in the presence of an acid. This reaction, called acetal hydrolysis, results in the formation of the parent carbonyl compound and the alcohol component. Thus, acetals are inert to bases, but they are acid-sensitive and therefore can be hydrolyzed in acidic media.



Problem 8.2. How would you prepare 1,1-dimethoxypropane from the corresponding aldehyde? Write an equation for the reaction and show the steps in the reaction mechanism.

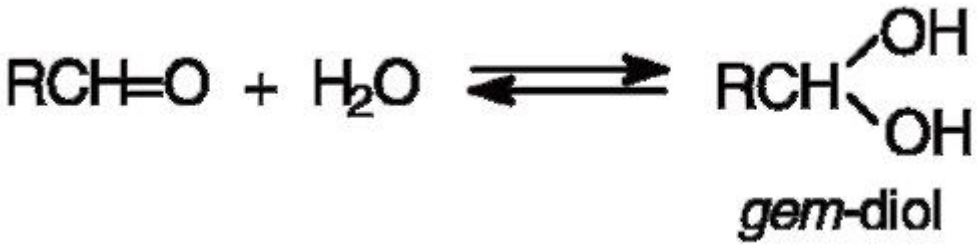
Aldehydes and some ketones similarly react with glycols to form cyclic acetals.

Example 8.1. Write the steps for the acid-catalyzed reaction of acetone with ethylene glycol (1,2-ethanediol).

Solution. In the first step an intermediate hemiacetal is formed in a usual way. The second step represents the *intramolecular* nucleophilic reaction, which is more probable than reaction between two molecules (of the hemiacetal and the second mole of the glycol). Thus, a product is a cyclic acetal:

8.2.2. Addition of Water: Hydration

Water, like alcohols, is an oxygen nucleophile and can react with some aldehydes and ketones by nucleophilic addition to yield so called hydrated forms of aldehydes, or gem-diols (1,1-diols). The reaction is reversible and hydrates of most aldehydes and ketones cannot be isolated from aqueous solution because they readily eliminate water to regenerate the carbonyl compound.

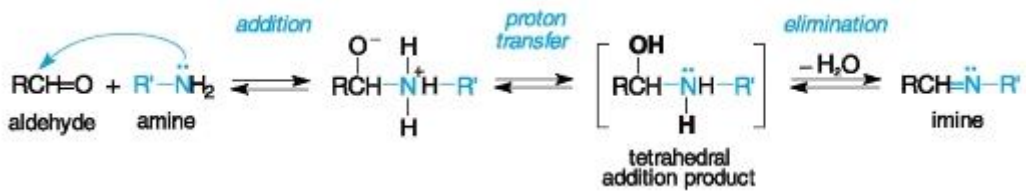


The equilibrium position depends on the structure of a compound. Formaldehyde, for example, exists in aqueous solution almost completely (over 99.9%) as the hydrate, $CH_2(OH)_2$. Acetaldehyde is approximately half hydrated in water, while acetone does not form a hydrate, though it is completely soluble in water due to hydrogen bonding.

Electron-withdrawing substituents in a hydrocarbon portion increase stability of the hydrated form of a carbonyl compound. Thus, trichloroacetaldehyde (chloral) forms a stable crystalline hydrate, $CCl_3CH(OH)_2$, which is used in medicine as a sedative and soporific.

8.2.3. Addition of Nitrogen Nucleophiles: Imines and Related Compounds

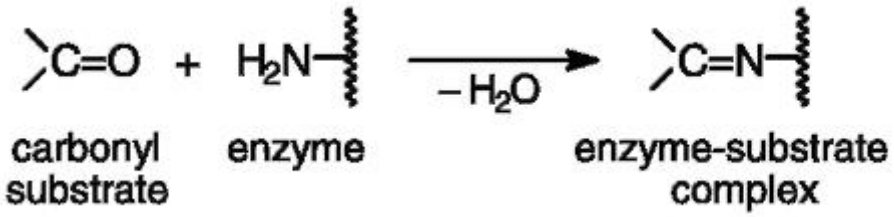
Amines and some related compounds (see below) act as nitrogen nucleophiles towards the carbonyl carbon. Primary amines, for example, add to aldehydes and ketones to yield N-substituted *imines.*The overall process involves two successive reactions, nucleophilic addition and elimination.



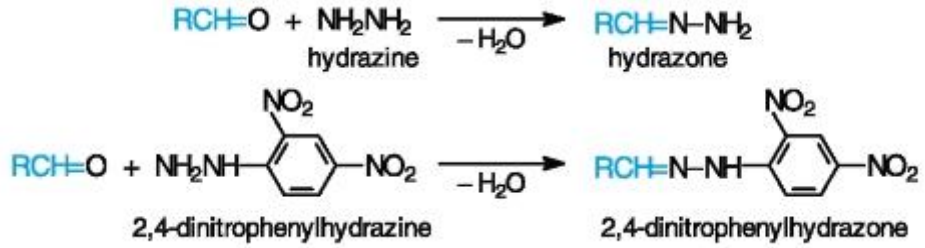
Nucleophilic attack on the carbonyl carbon by the lone-pair of electrons of an amine leads, after a proton transfer, to the tetrahedral addition product, a geminal amino alcohol that is similar to a hemiacetal. Such substances are normally unstable and cannot be isolated. They eliminate water to form a product with a carbonnitrogen double bond.

It should be noted that free imines, i. e. compounds with a fragment >C=NH, are usually unstable because of subsequent condensation reactions. On the contrary, if the radical R or R' in the above general formula is an aryl group, such imines called *Schiffs bases* are quite stable.

Imine compounds are important intermediates in a number of biochemical transformations and in biosynthesis of amino acids (Chapter 14). Imine formation represents sometimes an interaction mode in binding carbonyl compounds to the amino groups of proteins, including enzymes, as it is shown below in a generalized form:



In the laboratory, reactions of this type are useful in analytic and synthetic procedures. From other organic and inorganic compounds containing an amino group the most widely used are hydrazine and its derivatives. They react with aldehydes and ketones similarly to primary amines.



The products of these reactions are usually crystalline insoluble solids that have characteristic melting points. They are applied for isolation and identification of liquid aldehydes and ketones.

Imines and hydrazones can be subjected to acid-catalyzed or alkaline hydrolysis with formation of the initial compounds.

Problem 8.3. Write the structure of a product you would obtain in the reaction of p-methoxylbenzaldehyde with ethylamine. Which compound is more reactive in this reaction, benzaldehyde or p-methoxylbenzaldehyde?

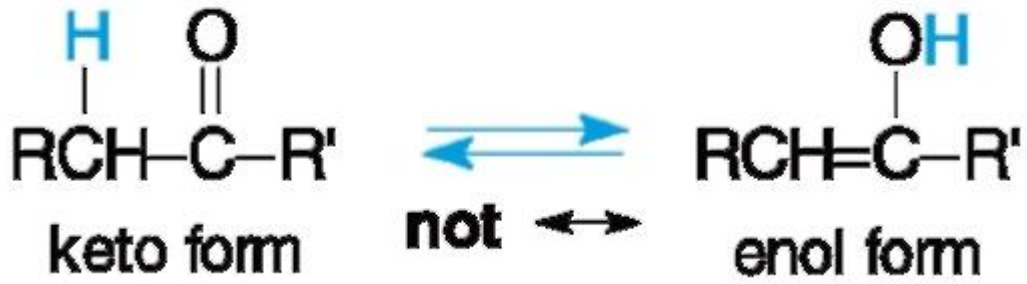
8.3. CH-ACIDIC PROPERTIES OF ALDEHYDES AND KETONES

The α-hydrogen in carbonyl compounds reveals increased acidity as compared to ordinary hydrogen atoms bonded to carbon. Really, the pK_a values for acetaldehyde and acetone are about 18 and 20 respectively, that are comparable to those of alcohols (Table 6.1). This is a result of the strong electron-withdrawing inductive effect of the carbonyl group. Another reason is a formation of the resonance-stabilized anion (as a conjugate base) that is produced when a carbonyl compound loses the α-proton on a base treatment.

This anion represents a resonance hybrid in which a negative charge is distributed between the α-carbon and the carbonyl oxygen. Because of its relation to enols, the resonance-stabilized ion is often called an *enolate ion.*

8.3.1. Keto-Enol Tautomerism

The resonance-stabilized ion can be reprotonated either on the carbon atom, giving back the carbonyl compound, or on the oxygen atom to give the isomeric enol. Both isomers are in equilibrium of two forms, called the *keto form* and the *enol form:*



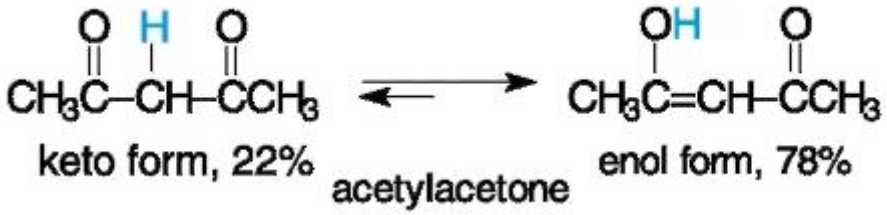
This special type of constitutional isomerism is described as *tautomerism,* and the isomers are said to be *tautomers* (form the Greek *tauto* - the same, and *meros* part). It should be stressed that tautomers are real isomers, not resonance contributing structures.

Tautomerism is a dynamic isomerism, which consists in migration of some groups (mostly a proton) within a molecule and is accompanied by redistribution of electron density.

Theoretically, any carbonyl compound having an α-hydrogen can exist as a tautomeric mixture. However, simple aldehydes and ketones exist mainly in the keto form. In acetone, for example, the enol form amounts to less than 0.001%. The main reason for the greater stability of the keto form is that energy of the C=O and C-H bonds present in the keto form is greater than that of C=C and O-H bonds of the enol form (control this, using Appendix 1).

Problem 8.4. Draw the structural formula for the enol form of the following compounds: (a) acetaldehyde; (b) acetone; (c) cyclohexanone.

There are carbonyl compounds whose molecules contain an additional electronwithdrawing group attached to the α-carbon (very often this is a second carbonyl group). The enol tautomer in such compounds is more abundant and may even predominate as, for example, in 2,4-pentanedione (a β-diketone commonly named acetylacetone).



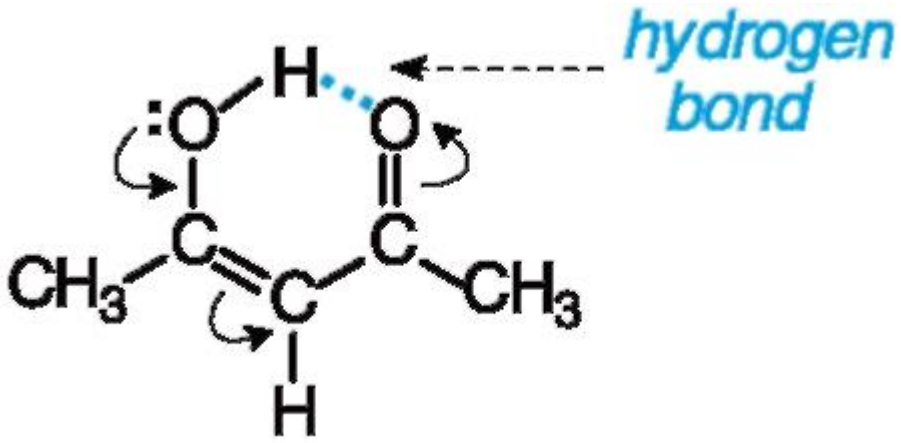
The data on keto-enol tautomerism are summarized in Table 8.1.

Table 8.1. Enol form contents (%) in some carbonyl compounds

Compound name	*Structure**	*Enol form*
Acetone	$CH_3–CO–CH_3$	0.0002
Malonic ester (diethylmalonate)	$C_2H_5OOC–CH_2–COOC_2H_5$	0.008
Butanone	$CH_3CH_2–CO–CH_3$	0.012
Cyclohexanone		0.02
Acetophenone	$C_6H_5–CO–CH_3$	0.035
Diacetyl (butanedione)	$CH_3–CO–CO–CH_3$	0.56
Acetoacetic ester (ethyl 3-oxobutyrate)	$CH_3–CO–CH_2COOC_2H_5$	7.5
Acetylacetone (2,4-pentanedione)	$CH_3–CO–CH_2–CO–CH_3$	78
Oxalacetic acid (oxobutanedioic acid)	$HOOC–CO–CH_2COOH$	80

* α-Hydrogen atoms are given in colour. When two or more equivalent α-hydrogens exist in a molecule, only one of them participates in tautomeric equilibrium.

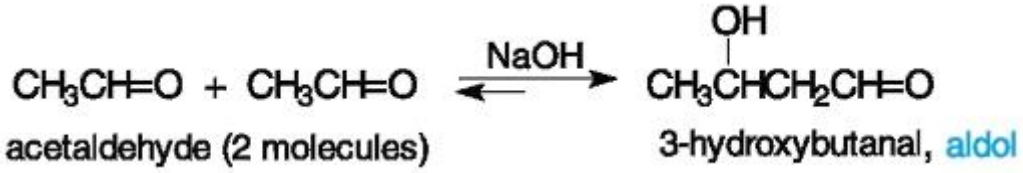
The greater stability of the enol form in some β-dicarbonyl compounds can be attributed to conjugation of the enolic OH group and to intramolecular hydrogen bonding.



Many reactions of carbonyl compounds, both in the laboratory and *in vivo,* involve enol forms.

8.3.2. The Aldol Condensation

When a carbonyl compound having an α-hydrogen is treated with an alkali in aqueous solution, a rapid dimerization takes place at room temperature (or below), for example:

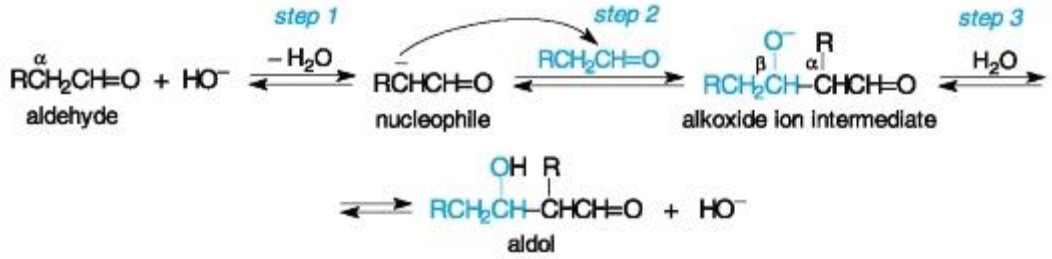


The product is commonly called an *aldol* (because it is both an *ald*ehyde and an alcoho/), and reactions of this type are known as *aldol condensations*[1]. These reactions are very useful in organic synthesis for the C-C bond formation and in living systems.

The mechanism of the aldol condensation shows two substantial characteristics of carbonyl compounds: the CH-acidity of the α position and the tendency of the carbonyl group to nucleophilic addition.

[1] The author believes this term being inaccurate because the reaction signifies not the condensation of an aldol but the condensation that results in the aldol formation.

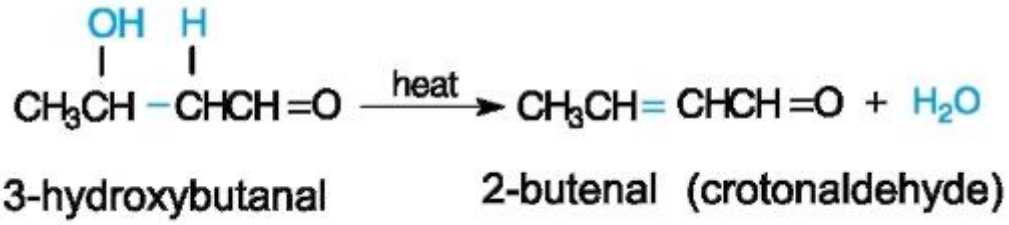
The reaction mechanism involves three steps. In the first one, a base removes the α-hydrogen to form an enolate ion (see above). In the second step, this anion as a nucleophile (actually as a carbanion) attacks another molecule of the neutral carbonyl compound in addition reaction to give an alkoxide ion intermediate. In the third step, the alkoxide ion abstracts a proton from the solvent (usually water), thus regenerating the catalyst (HO^-).



Note that α-carbon acts always as a nucleophile, therefore the aldol

product has the hydroxyl group always in the β position, regardless of the length of the carbon chain in the starting compound.

The aldol can be easily dehydrated on heating because it retains the acidic α-hydrogen. α,β-Unsaturated aldehydes and ketones are formed when the reaction is performed at elevated temperature. Crotonaldehyde is obtained from the simplest aldol in this way, therefore reactions of this type are called the *croton condensation.*

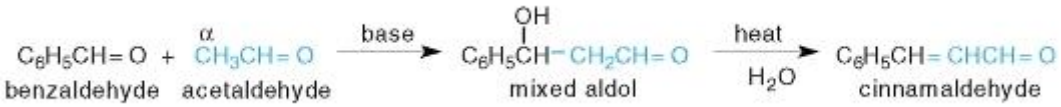


In some reactions dehydration proceeds so easily that it is impossible to isolate the product in the aldol form.

Problem 8.5. Give the structure of an aldol that may be obtained on treating butyraldehyde (butanal) with a base. Explain the role of a catalyst used.

The mixed aldol condensation. Two different carbonyl compounds can react in a mixed, or crossed manner to give four possible products. In general, base treatment of a mixture of carbonyl compounds A and B results in formation of a complex mixture of two «symmetrical» aldol products, A-A and B-B, along with two «mixed» aldol products, A-B and B-A. We omit the specific equation for this reaction because it is of little importance in the laboratory when both reactants, A and B, have α-hydrogens. A mixed aldol condensation is practical, provided that the reaction partners are carefully selected.
Example 8.2. Write the structure of the mixed aldol obtained from acetaldehyde and benzaldehyde. What is a dehydration product of this mixed aldol?

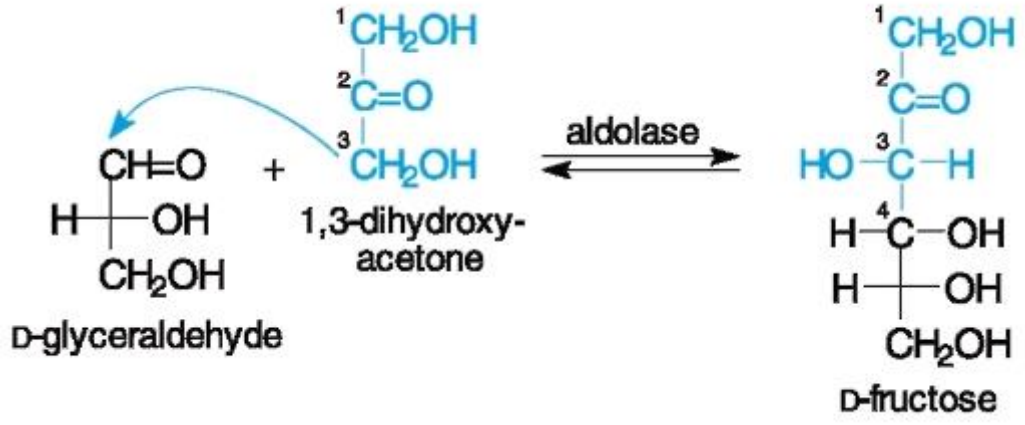
Solution. Of two aldehydes, only acetaldehyde has α-hydrogens and can form an enolate ion. The latter attacks the benzaldehyde carbonyl group in the crossreaction. The resulting mixed aldol eliminates water on heating to yield cinnamaldehyde, the flavour constituent of cinnamon.



To avoid self-condensation of acetaldehyde, as a side reaction, this aldehyde should be applied at very low concentration.

8.3.3. The Biochemical Aldol Condensation

A number of aldol-type reactions occur in the living cells of plants and animals. One *in vivo* reaction, which is a step in the biosynthesis of monosaccharides, involves a mixed aldol condensation of glyceraldehyde and 1,3-dihydroxyacetone to produce fructose (stereochemical aspects of the reaction are omitted here).

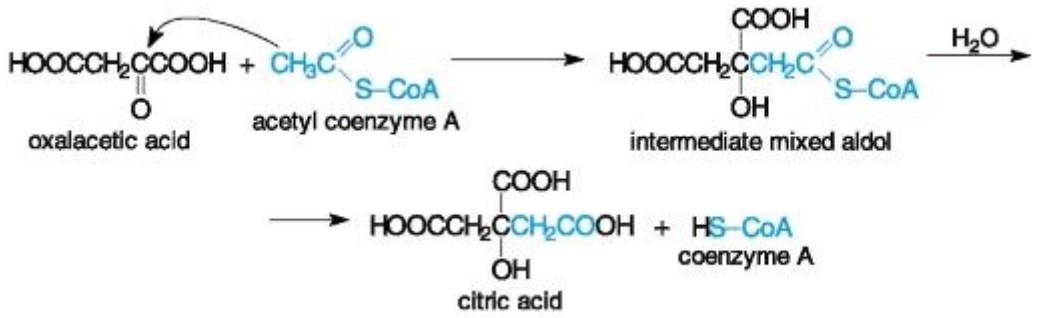


Actually, both reagents represent phosphoric acid esters, and the product is fructose 1,6-diphosphate. The phosphate groups are omitted in the equation for the sake of simplicity.

Note, that the reaction is regiospecific because both two carbonyl compounds possess α-hydrogens, and another mode of cross-reaction might be expected.

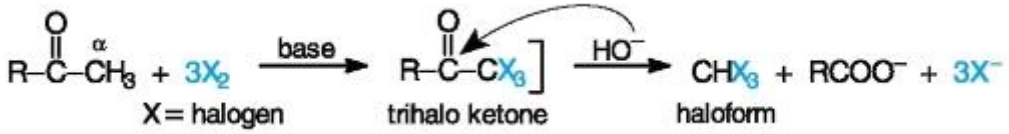
The reverse reaction, i. e. an aldol cleavage also called a *retro-aldol reaction*, is well known in living systems. The splitting of fructose 1,6-diphosphate into threecarbon units (as shown above) is the initial step in the glycolysis process.

Another example of the mixed aldol condensation is biosynthesis of citric acid from oxalacetic acid (an oxo dicarboxylic acid) in the Krebs cycle. Acetyl coenzyme A (Sec. 7.2.3 and 17.4.2), as a potential nucleophile, attacks the ketone carbonyl group of oxalacetic acid to give an intermediate aldol product. The subsequent hydrolysis of the thioester group completes the formation of citric acid.



8.3.4. The Haloform Reaction

In the presence of bases, aldehydes and ketones, that have α-hydrogen atoms, readily react with halogens to form α-halogeno products. If excess halogen and base are used in the reaction with methyl ketones, triple halogenations occur because the first substitution makes the remaining α-hydrogens more acidic due to the -/ effect of the halogen. The trihalo ketone thus formed is then cleaved under the alkaline conditions to a carboxylate ion and a *haloform* (chloroform, $CHCl_3$; bromoform, $CHBr_3$; or iodoform, CHI_3).



This reaction is known as the *haloform reaction.* When × is iodine, the characteristic yellow colour and odour of the crystalline iodoform product can be used as a diagnostic test for methyl ketones. Compounds which are easily converted under the used oxidizing conditions give a positive iodoform test as well. Alcohols that have the fragment $-CH(OH)CH_3$ belong to such compounds because they are first oxidized to methyl ketones $-C(O)CH_3$.

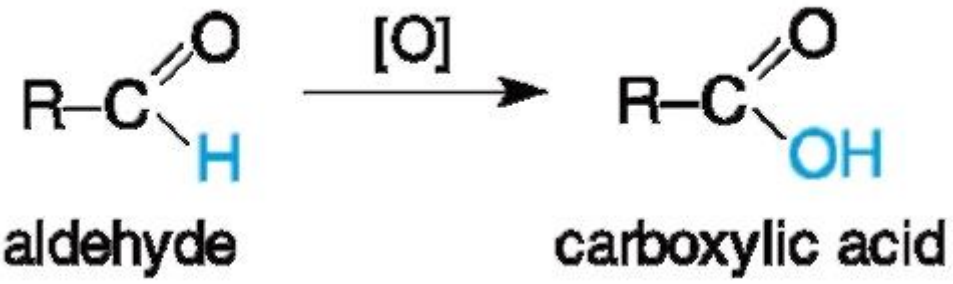
Problem 8.6. Which of the following compounds would give a positive iodoform test: (a) acetone; (b) butanal; (c) butanone; (d) 3-pentanone; (e) 2-propanol.

8.4. OXIDATION AND REDUCTION REACTIONS

The carbonyl carbon is regarded to be in an intermediate oxidation state; therefore a carbonyl group can be both reduced with increase in its hydrogen content or oxidized with increase in its oxygen content.

8.4.1. Oxidation of Aldehydes and Ketones

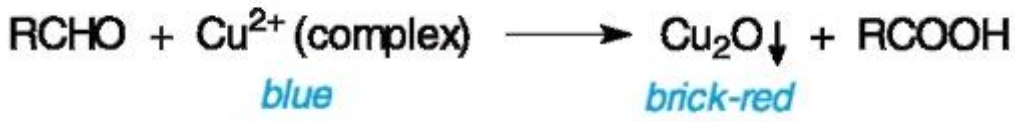
Aldehydes are very easily oxidized by most oxidizing agents, giving carboxylic acids. Ketones are more stable towards oxidation. Some aldehydes, especially aromatic ones, can be even oxidized on prolonged contact with air.



Several mild oxidizing agents provide simple visual tests that differentiate aldehydes from most ketones. One of them is *Tollens' reagent* (or Tollens' solution), a water-soluble complex silver hydroxide, $[Ag(NH_3)_2]OH$. Aldehydes reduce the silver diammin ion, $Ag(NH_3)_2^+$, to silver metal, which deposits on the glass surface as a mirror[1]. For this reason this test is called silver mirror test that can be expressed by the following equation:

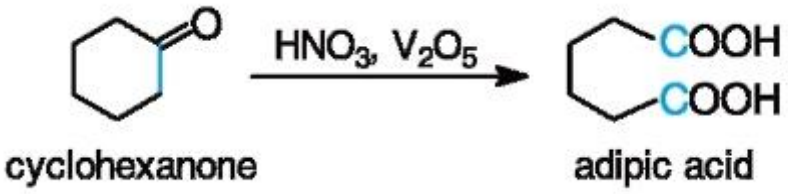
[1] If the walls of the vessel are not thoroughly cleaned, silver deposits as black precipitate.

Alternatively, *Fehling's reagent* (a blue solution of Cu^{2+} complexed with the tartrate ion) gives a brick-red precipitate of copper(I) oxide, Cu_2O, when it oxidizes aldehydes:



Both reagents are not specific for aldehydes. Other compounds that undergo easy oxidation, such as polyhydroxylic phenols, amino phenols, aromatic amines, and α-hydroxy ketones, also give positive Tollens' and Fehling's tests.

Ketones can be oxidized under specific, more vigorous conditions. Oxidation is accompanied by cleavage of the C-C bond next to the carbonyl group (coloured in the example below), both carbons become the COOH fragments. One of the simplest, but commercially important examples is oxidation of cyclohexanone to adipic acid (hexanedioic acid), which is used in manufacturing nylon.

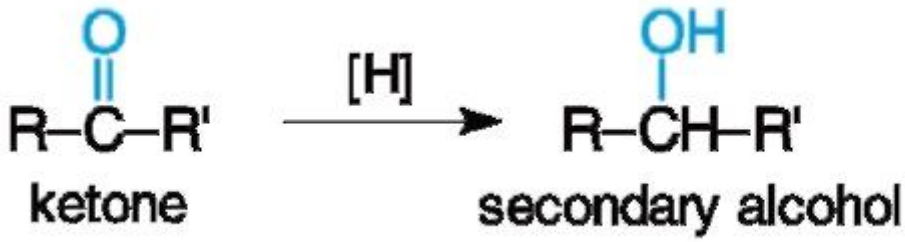


Problem 8.7. Explain, why cyclohexanone gives a single oxidation product, whereas butanone gives a mixture of carboxylic acids under similar conditions?

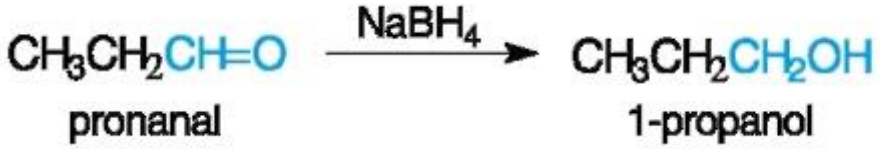
Thus, aldehydes and ketones can be easily distinguished by chemical tests, since aldehydes are oxidized readily to carboxylic acids under mild conditions, whereas ketones are relatively resistant to most oxidizing reagents.

8.4.2. Reduction of Aldehydes and Ketones

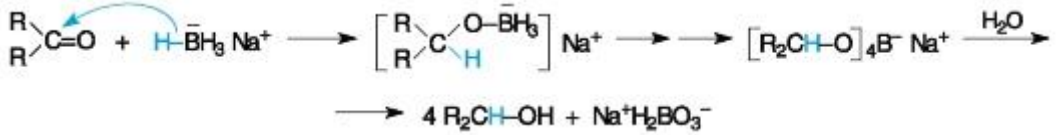
Aldehydes and ketones are readily reduced to primary and secondary alcohols, respectively.



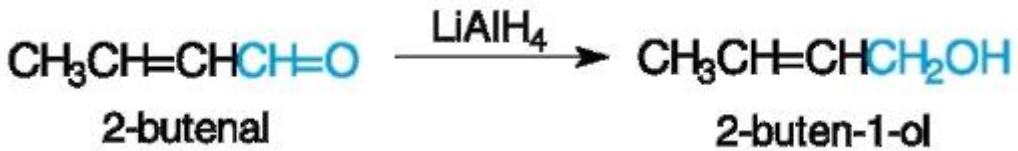
Various reducing agents may be applied for the reduction. These may be zinc in a dilute hydrochloric acid, sodium in ethanol, hydrogen over a nickel catalyst (catalytic hydrogenation). The most effective for the laboratory reductions are complex metal hydrides such as lithium aluminium hydride, $LiAlH_4$, or sodium borohydride, $NaBH_4$. The former reacts violently with water therefore thoroughly anhydrous conditions are required. Reductions with $NaBH_4$, by contrast, can be performed in aqueous or alcoholic solutions.



Reduction with complex metal hydrides represents irreversible nucleophilic addition of the hydride ion (H^- from $NaBH_4$) at the carbonyl carbon as shown below in a simplified form:



Since a carbon-carbon double bond is attacked by nucleophiles with difficulty, complex metal hydrides, especially more selective $LiAlH_4$, can reduce unsaturated aldehydes and ketones with retention of the C=C double bond, for example:

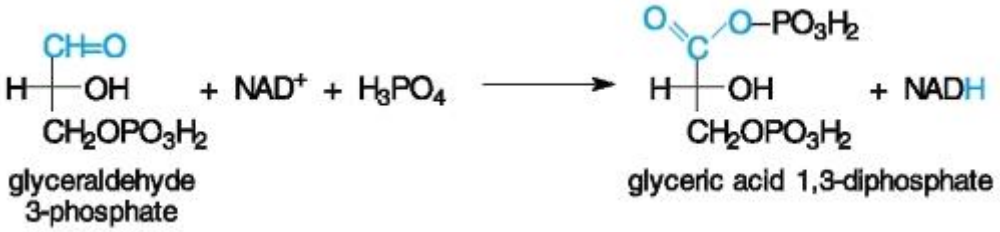


8.4.3. Biochemical Oxidation and Reduction of Carbonyl Compounds

Nicotinamide adenine dinucleotide (NAD+) is one of the most common oxidants of aldehydes and ketones in living systems, and the oxidant is converted to its

reduced form, NADH. The latter, in turn, serves as a reducing agent in enzymic reduction reactions, becoming oxidized to NAD^+. The two following examples demonstrate both types of the reactions.

In the first one, glyceraldehyde 3-phosphate is oxidized to a glyceric acid derivative in the presence of inorganic phosphate. This is one of the numerous steps in glycolysis.



glyceraldehyde 3-phosphate

glyceric acid 1,3-diphosphate

The second example shows enzymic reduction of an oxo acid (pyruvic acid) to a hydroxy acid (lactic acid):

pyruvic acid

lactic acid

Additional Problems

8.8. Which compound would you expect to be the more reactive and the less reactive towards nucleophilic reagents:

(a) benzaldehyde;

(b) p-methylbenzaldehyde;

(c) p-nitrobenzaldehyde?

Explain your reasoning.

8.9. Complete the following equations, and classify the products formed:

(a) $(CH_3)_2CHCH{=}O + 2C_3H_7OH \rightarrow$

(b) $C_6H_5CH{=}O + C_2H_5OH$ (1 mole) $\rightarrow$

(c) CH3C(O)CH3 + (CH3)2CHNH2 →

What catalyst, *ifany*, is used in these reactions?

8.10. Write an equation for the reaction of phenylacetaldehyde with these reagents:

(a) phenylhydrazine;

(b) excess methanol in the presence of an acid catalyst;

(c) Tollens' reagent.

8.11. Show all the steps in the reaction between propanal and 2,4-dinitrophenylhydrazine. Which compound acts as a nucleophile?

8.12. Predict a structure of the compound if it gives 3-methylbutanal and two moles of ethanol on hydrolysis. Write an equation for the reaction and show a catalyst.

8.13. Write an equation for the reaction of benzaldehyde with equimolar amount of propylene glycol (1,2-propanediol) in the presence of an acid catalyst.

8.14. Which of the following compounds exists as a tautomeric mixture with a marked amount of the enol form?

(a) $HOOC\text{-}CO\text{-}CH_2COOH$;

(b) $HOOCCH_2CH_2\text{-}CO\text{-}COOH$;

(c) $CH_3\text{-}CO\text{-}COOC_2H_5$;

(d) cyclopentanone.

Write the structures for enol forms of all compounds.

8.15. Give the structure of a mixed aldol obtained from acetone and formaldehyde.

8.16. Complete the equations for the following reactions. If no reaction occurs, say so.

(a) $CH_3CHO + NaBH_4\rightarrow$

(b) C6H5CHO + H2O2 →

(c) $CH_3C(O)CH_3$ + aqueous $KMnO_4\rightarrow$

8.17. Write equations for catalytic hydrogenation of the following compounds:

(a) cyclopentanone;

(b) 2-butenal;

(c) pyruvic (2-oxopropanoic) acid.

Name the products obtained.

8.18. What products would you obtain, *if any,* in a reaction with Fehling's reagent of the following compounds:

(a) cyclohexanol;

(b) butanal;

(c) propenal;

(d) glycerol?

Illustrate with equations and name the products.

Chapter 9. CARBOXYLIC ACIDS AND THEIR DERIVATIVES

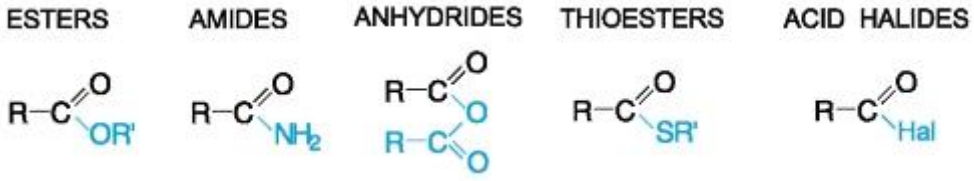
Carboxylic acids were among the earliest organic compounds studied by chemists. The Swedish pharmaceutical chemist C.W. Scheele was the first to discover many

carboxylic acids at the end of the 18th century, well before the theory of the chemical structure had been developed.

Not only carboxylic acids themselves but also their functional derivatives are of chemical and biochemical interest.

Carboxylic acids are compounds of the general formula RCOOH that have the  carboxyl group -COOH as a functional group. Acid derivatives have the general formula RC(O)Z, where Z is a substituent containing an unshared electron pair.

The most important acid derivatives are esters, amides, anhydrides, thioesters, and acid halides. Acid derivatives are called so because all of them can be derived from carboxylic acids and can be hydrolyzed to regenerate carboxylic acids.



9.1. GENERAL CHARACTERISTICS OF CARBOXYLIC ACIDS

Carboxylic acids and their numerous derivatives form a family of compounds whose chemistry is exclusively varied. It is no wonder because we observe various types of functions in their molecules.

9.1.1. Classification and Nomenclature

Carboxylic acids may be classified similarly to aldehydes (Sec. 8.1.1), i. e. as aliphatic or aromatic ones. The classification signs may be extended with unsaturated, heterocyclic, and dicarboxylic acids. These types are exemplified below and presented in Table 9.1.

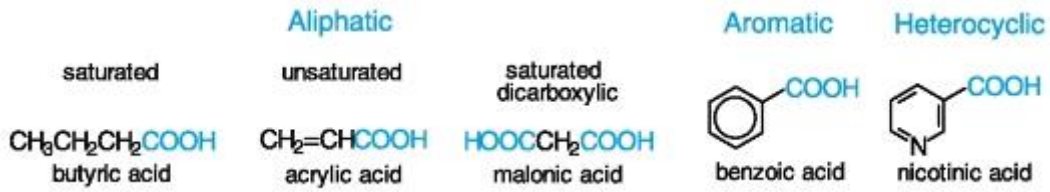
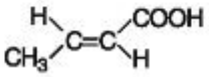
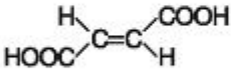
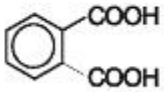
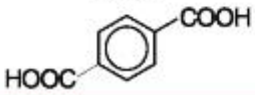


Table 9.1. Names and some physical constants of selected carboxylic acids

*Trivial name**	*Systematic name**	*Formula*	*Melting point, °C*	*Boiling point, °C*	*pK_a ***
Monocarboxylic					
Formic	Methanoic	$HCOOH$	8	101	3.8
Acetic	Ethanoic	CH_3COOH	17	118	4.8
Propionic	Propanoic	CH_2CH_3COOH	–21	141	4.9
Butyric	Butanoic	$CH_3(CH_2)_2COOH$	–5	164	4.8
Valeric	Pentanoic	$CH_3(CH_2)_3COOH$	–34	185	4.8
Isovaleric	3-Methylbutanoic	$(CH_3)_2CHCH_2COOH$	–29	177	4.8
Stearic	Octadecanoic	$CH_3(CH_2)_{16}COOH$	70	376	
Acrylic	Propenoic	$CH_2=CHCOOH$	12	141	4.2
Crotonic	*trans*-2-Butenoic		71	185	4.7
Benzoic	Benzenecarboxylic	C_6H_5COOH	122	249	4.2
Dicarboxylic					
Oxalic	Ethanedioic	$HOOC–COOH$	189		1.2; 4.2
Malonic	Propanedioic	$HOOCCH_2COOH$	136***		2.8; 5.7
Succinic	Butanedioic	$HOOCCH_2CH_2COOH$	186	300	4.2; 5.5
Glutaric	Pentanedioic	$HOOC(CH_2)_3COOH$	98	303***, 200 (20 Torr)	4.3; 5.4
Fumaric	*trans*-Butenedioic		296***	165 (1.7 Torr)	3.0; 4.4
Phthalic	1,2-Benzenedicarboxylic		200***		3.0; 5.4
Terephthalic	1,4-Benzenedicarboxylic		427***		3.5; 4,5

* The word *acid* is omitted in the table.

** For dicarboxylic acids the first number denotes pK_{a1}, the second one - pK_{a2}.

*** Melts or boils with decomposition.

For historical reasons mentioned above, many carboxylic acids received their trivial names long ago. Many of them (and all the names listed in Table 9.1) are allowed, and even recommended by the IUPAC rules, including the names of acid derivatives (Table 9.2). Trivial names are preferably used in the biochemical literature.

Table 9.2. Trivial names of carboxylc acid derivatives

Number of carbons in a chain	*Acyl group RC(O)–*	*Anion or a basis of the ester*	*Amide*
Monocarboxylic			
1	Formyl	Formate	Formamide
2	Acetyl	Acetate	Acetamide
3	Propionyl	Propionate	Propionamide
4	Butiryl	Butirate	Butyramide
5	Valeryl	Valerate	Valeramide
–	Benzoyl	Benzoate	Benzamide
Dicarboxylic*			
2	Oxalyl	Oxalate	Oxalamide
3	Malonyl	Malonate	Malonamide
4	Succinyl	Succinate	Succinamide
5	Glutaryl	Glutarate	Glutaramide

* For derivatives on both carboxyl groups.

In the IUPAC substitutive nomenclature, the carboxyl group in aliphatic representatives is indicated by the suffix -oic with the addition of the word acid. When the carbon atom of the COOH group is not a part of the parent structure, the ending -carboxylic acid is added (see Example 5 in Chapter 2). Acyl groups RC(O)- are named from the corresponding acids by changing the suffix -ic of the common or systematic name to the suffix -yl, for example, acetyl or ethanoyl for the group $CH_3C(O)$- (see Table 9.2).

Table 2.2 will also be useful in constructing the names of carboxylic acid deriva

tives.

Problem 9.1. Give the IUPAC names for the following compounds:

(a) $C_6H_5CH=CHCOOH$;

(b) CH_3COCl;

(c) $CH_2ClCOOH$;

(d) $HCOOCH_3$.

9.1.2. Electronic Structure of the Carboxyl Group

Although the carboxyl group appears to be a simple combination of the hydroxyl and carbonyl groups, the interaction between them generates some unique properties.

The carboxyl group represents a planar ρ,π-conjugated system, in which a lone pair ofelectrons of the hydroxyl oxygen overlaps with p or bitals of the C=O double bond (Fig. 9.1). This is confirmed by the data on carbon-oxygen bond lengths in related compounds. Thus, the C=O bond in carboxylic acids is longer than that in carbonyl compounds (124 *versus* 121 pm), whereas the C-OH bond in acids is shorter that that in alcohols (136 *versus* 143 pm).

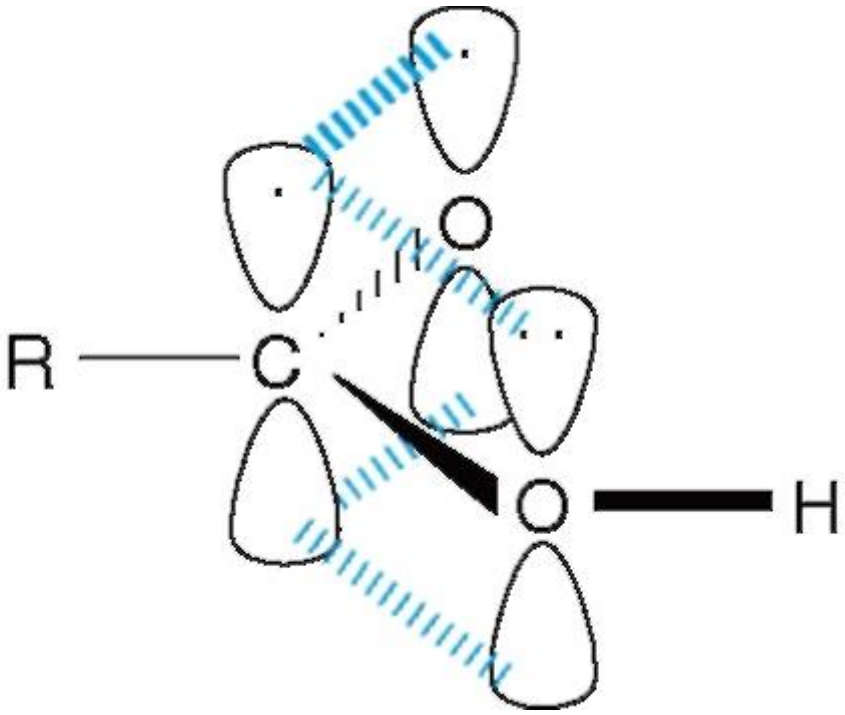
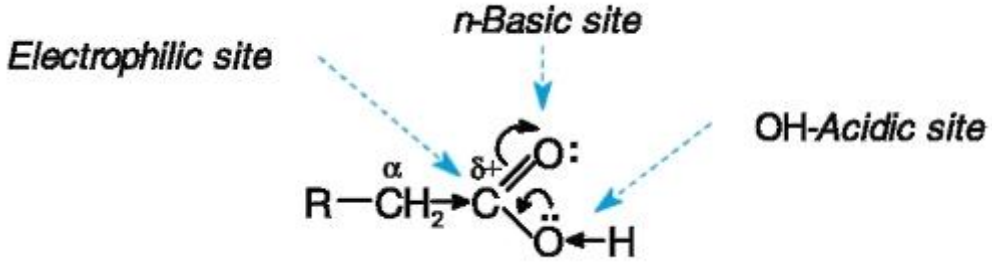


Figure 9.1. Orbital overlap in carboxylic acids.

The following reactive sites can be pointed out in the carboxylic acid molecule:

- the OH-acidic site that reacts with bases by deprotonation;
- the carbonyl carbon as an electrophilic site which can be attacked by nucleophiles by substitution;
- the oxygen atom as a weak *n*-basic site that can be protonated with strong acids.



Similarly to other classes of organic compounds, additional reaction sites may arise in the hydrocarbon portion of a molecule.

The α-hydrogen virtually is not acidic in carboxylic acids because a stronger acidic site is present. When the carboxyl group is ionized the α-hydrogen loses acidity completely. However α-CH acidity becomes appreciable in acid derivatives, for example, in esters (Sec. 9.4).

9.2. ACIDIC PROPERTIES

Before proceeding further, it would be helpful for you to review Sec. 5.1 in

which general principles of acidity were declared.

As their name suggests, carboxylic acids are *acidic*. Conjugation in the carboxyl group increases acidity of a carboxylic acid as compared with other OH-acids such as alcohols and phenols. Carboxylic acids are among the strongest acidic compounds in organic chemistry (only sulphonic acids, RSO_3H, exceed them in acidity). They partly dissociate in water, giving a carboxylate ion and a hydronium ion:

$$\underset{\text{carboxylic acid}}{RCOOH} + H_2O \rightleftharpoons \underset{\text{carboxylate ion}}{RCOO^-} + H_3O^+$$

Carboxylic acids react readily with metal hydroxides to form salts (Sec. 5.2.2). They also displace weaker acids from their salts to give metal carboxylate salts. The reaction serves as a simple visible test for carboxylic acids by a gas evolution. Carboxylate salts are converted into carboxylic acids on treatment with strong mineral acids.

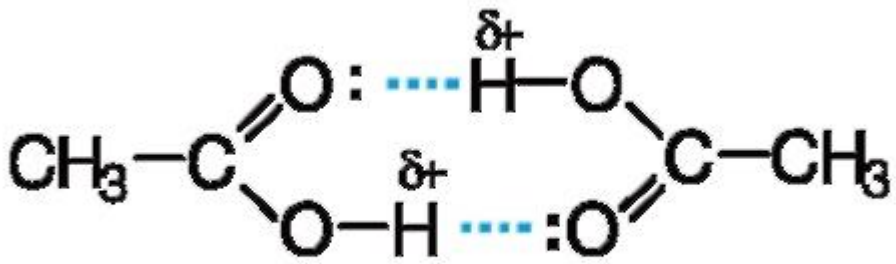
$$RCOOH + NaHCO_3 \longrightarrow \underset{\text{sodium carboxylate}}{RCOO^-Na^+} + CO_2\uparrow + H_2O$$

$$\underset{\text{sodium benzoate}}{C_6H_5COO^-Na^+} + HCl \longrightarrow \underset{\text{benzoic acid}}{C_6H_5COOH} + NaCl$$

Thus, carboxylic acids exist in acidic media only in an unionized form (RCOOH) whereas in alkaline media they are always ionized ($RCOO^-$).

Conjugation within the carbonyl group increases not only acidity of a compound but also basic properties of the double-bonded oxygen as compared with that of carbonyl compounds. This explains the fact that carboxylic acids exist normally in an associated form with strong intermolecular hydrogen bonds between the basic

and OH-acidic sites. For example, acetic acid exists as a dimer, even in a vapour



dimer of acetic acid

phase.

Hydrogen bonding with water also explains the complete water-solubility of the first four liquid monocarboxylic acids (including butyric acid) and the good solubility of the first four solid dicarboxylic acids (including glutaric acid).

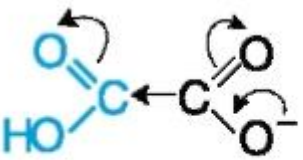
Acid strength. Most monocarboxylic acids of an aliphatic or aromatic series are acids of moderate strength with pK_a values in the range from 4 to 5 (Table 9.1). Substituent effects on acidity, including examples of carboxylic acids, have been discussed in Sec. 5.1.2. We turn to this question in the following example.

Example 9.1. How are the following facts explained?

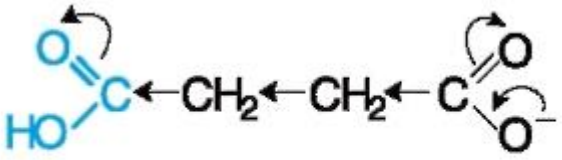
(I) The pK_{a1} for all dicarboxylic acids in Table 9.1 are higher than the pK_a for monocarboxylic acids with the same number of carbons.

(II) The acidity of the dicarboxylic acids decreases with the length of a carbon chain.

Solution. (I) The carboxyl group is an electron-withdrawing group. The inductive effect of such substituent tends to spread the negative charge in a carboxylate ion over more atoms and thus stabilizes an anion. Thus in dicarboxylic acids the second carboxyl group (unionized) increases the acidity of the other.



hydroxalate ion

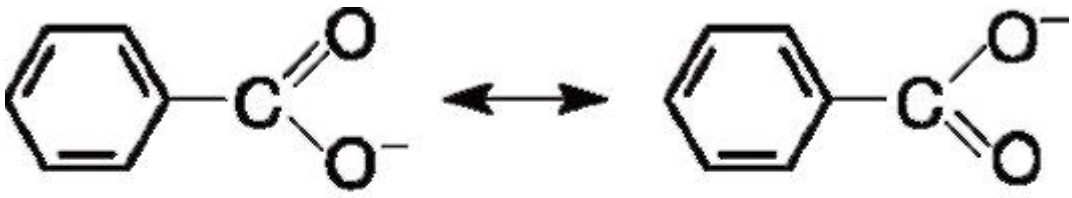


hydrogensuccinate ion

(II) The inductive effect falls off rapidly with distance. The carboxyl group exerts much smaller -/ effect in the hydrogensuccinate ion as compared with the hydroxalate ion. Succinic acid therefore only slightly exceeds in acidity monocarboxylic acids.

t might seem unexpected that a substituent affects stability of a carboxylate ion only inductively. (Recall delocalization of a negative charge in a phenoxide ion through ρ,π conjugation.) Conjugation of the oxygen lone pair of electrons is possible only with a C=O double bond regardless of other bonds present. Likewise,

it is impossible to draw a contributing resonance structure, using other but oxygen negatively charged atoms. Any attempt to do this with the benzoate ion will be unsuccessful.



resonance-stabilized benzoate ion

For this reason substitution in benzoic acid slightly affects acidity, depending on electron-donating or electron-withdrawing character of the substituent. For the same reason oxalic acid is the strongest unsubstituted carboxylic acid. The strongest of all carboxylic acids is trifluoroacetic acid, CF_3COOH, pK_a0.2.

Problem 9.2. Arrange the following compounds in the order of decreasing acidic strength:

(a) benzoic acid;

(b) phenol;

(c) p-nitrobenzoic acid;

(d) p-methylbenzoic acid.

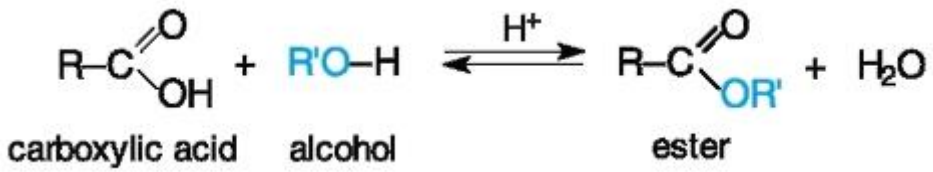
For the strongest acid write the equation for neutralization with ammonia and name the product.

9.3. NUCLEOPHILIC SUBSTITUTION AT ACYL CARBON

Nucleophilic substitution is, perhaps, the most common and important reaction of carboxylic acids and their derivatives. In carboxylic acids a partial positive charge (δ+) on the carboxyl carbon is decreased compared to that of aldehydes and ketones. This means, in general, that carboxylic acids are less reactive towards nucleophilic reagents than carbonyl compounds are. Moreover, a hydroxyl group belongs to poor leaving groups. However, it can be modified or transformed into other functions, which are good leaving groups. This approach is realized in the reaction of carboxylic acids with alcohols to form esters.

9.3.1. Esterification of Carboxylic Acids

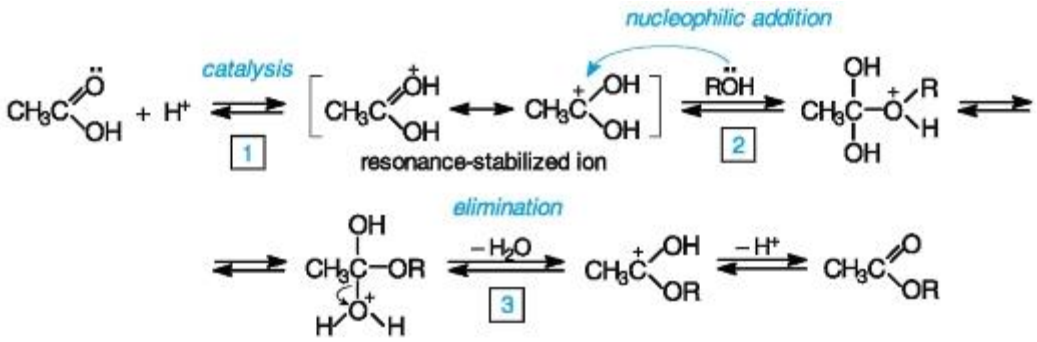
On heating a carboxylic acid with an alcohol in the presence of an acid catalyst (usually anhydrous H_2SO_4 or gaseous HCl), reversible ester formation occurs.



carboxylic acid alcohol ester

This reaction is called the *Fischer esterification.* The application of large excess of one of the reactants or removal of the ester or/and water can shift the equilibrium to the right.

The esterification mechanism. There are three main steps in a reaction mechanism. In the first step, protonation of a carboxylic acid increases the positive charge on the carboxyl carbon to give a resonance-stabilized cation. In the second, addition step an alcohol (as a nucleophile) attacks the carbocation with the formation of a new C-O bond. Then, after proton migration, water is eliminated. Finally, loss of a proton gives the ester product and regenerates the catalyst.



Thus, the overall reaction is outwardly similar to nucleophilic substitution that occurs by the S_N2 mechanism. But as it is seen from the reaction mechanism, esterification is not a direct substitution, rather it is addition-elimination. Since the net result of the reaction is substitution (OR for OH), the reaction is referred to as *nucleophilic acyl substitution.*

Acidic hydrolysis of esters is a reverse reaction to the ester formation. Esters can also be hydrolyzed with alkalis.

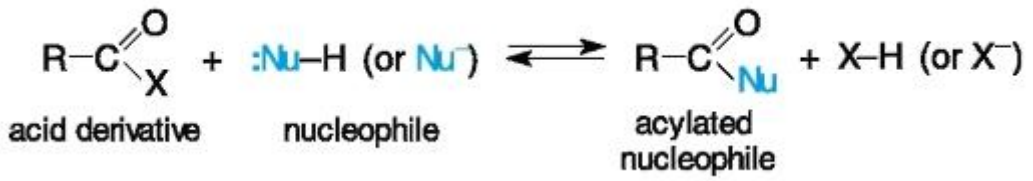
Alkaline hydrolysis is called *saponification* (from the Latin *sapo* - soap) because this type of reaction has been used and is used now to make soaps (alkali metal salts of long-chain acids) from fats (Sec. 12.2 and 12.3). Saponification is an irreversible reaction, and at least one equivalent of an alkali is required.

Problem 9.3. Write an equation for the reaction illustrating the mechanism for esterification of isopropyl alcohol with butyric acid. What would you suggest for shifting equilibrium to the direct reaction?

9.3.2. Acylation Reactions with Carboxylic Acid Derivatives

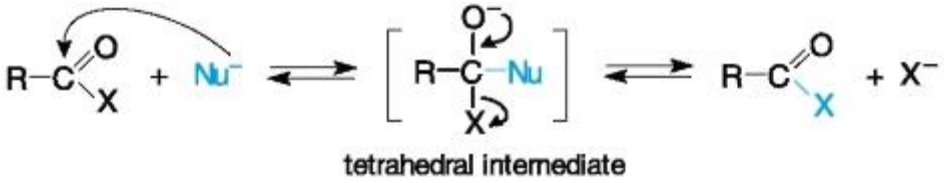
Numerous acid derivatives are known, but we will be concerned only with five of them: esters, amides, thioesters, anhydrides, and acid halides. Esters and amides occur widely in nature; anhydrides and, especially, acid halides are creatures of the laboratory chemists because of their high reactivity.

The general reaction can be presented as follows (where × is a nucleophilic group):



These reactions involving common nucleophiles, such as water, alcohols, ammonia, and amines, are usually designated hydrolysis, alcoholysis, ammonolysis, and aminolysis, respectively. With regard to nucleophiles, transformations of this kind are often referred to as acylation reactions, i. e. the acyl group is transferred from the group × in the acid derivative to nucleophile in the product. The term *acyl transfer* is used for such reactions in biochemistry.

Acylation reaction proceeds by a two-step addition-elimination pathway through a tetrahedral intermediate. Loss of a leaving group X^- regenerates the carbonyl group.



Relative reactivity of acid derivatives depends on stability of their leaving groups. The following anions are arranged in order of decreasing stability:

$Cl^- > RCOO^- > RS^- > HO^- > RO^- > NH_2^-$

For this reason, the reactivity order in acylation reactions for carboxylic acids and their derivatives is as follows:

$RCOCl > RCO\text{-}O\text{-}COR > RCOSR' > RCOOH \gg RCOOR' > RCONH_2 > RCOO^-$

This means that it is easy to transform a more reactive acid derivative into a less reactive one. Analysis of the reactivity order leads to the following conclusion shown in Fig. 9.2.

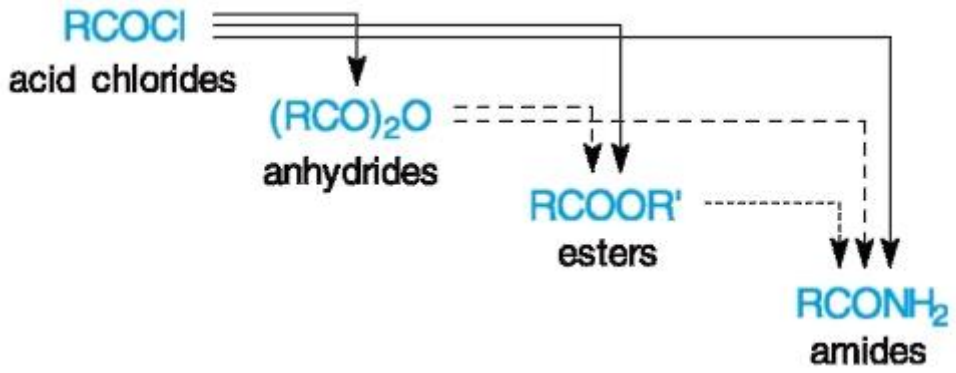
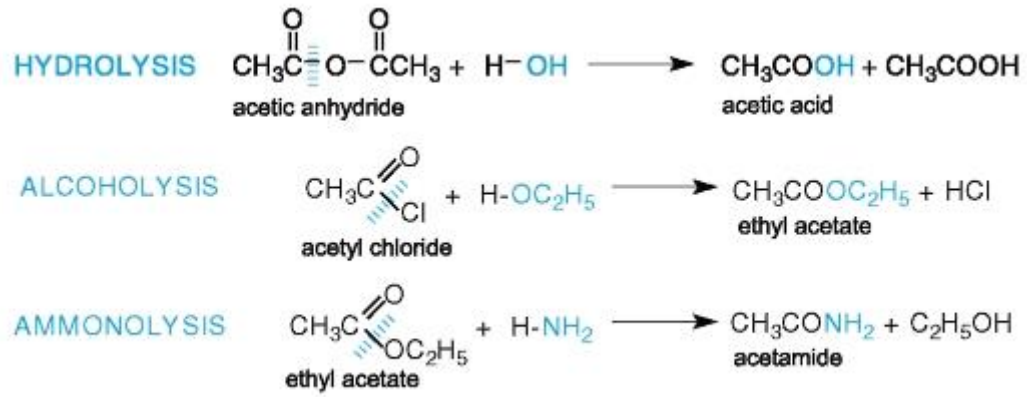
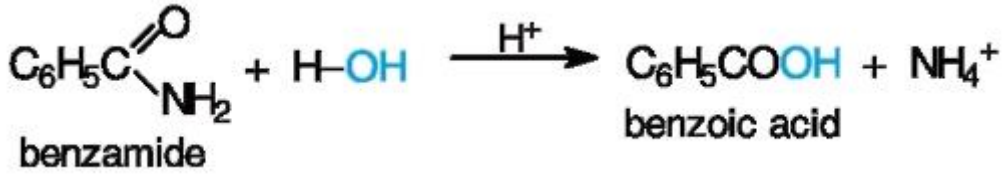


Figure 9.2. Interconversions of carboxylic acid derivatives.

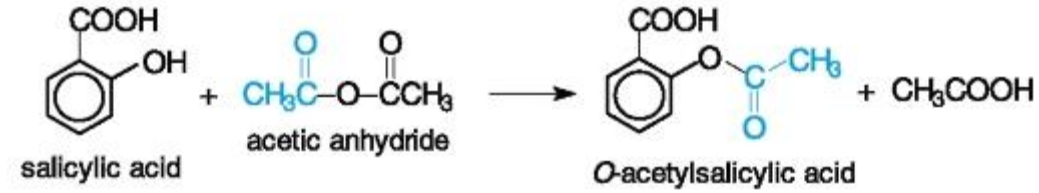
Some typical nucleophilic substitution reactions are illustrated below with various derivatives of acetic acid:



All acid derivatives can be hydrolyzed. Acid halides and anhydrides undergo hydrolysis most readily, whereas esters and amides are hydrolyzed only on heating in acidic or alkaline medium.

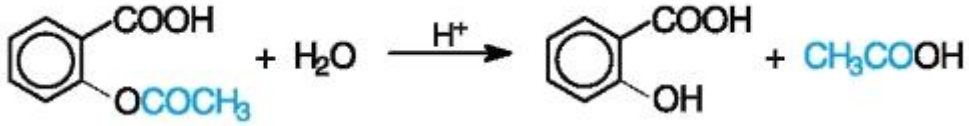


Amides are acid derivatives that resist to hydrolysis most. The reason is that the amino group is a very poor leaving group.
Many drugs are esters or amides from the chemical point of view. For example, aspirin (O-acetylsalicylic acid) is an ester manufactured from salicylic acid (2-hydroxybenzoic acid). The phenolic hydroxyl group undergoes acetylation in this reaction:

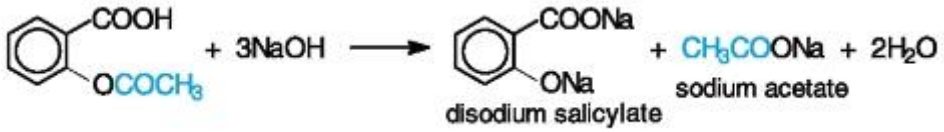


Aspirin can be hydrolyzed under acidic or alkaline conditions. That is why we should always bear in mind the possibility of ester hydrolysis in the acidic medium of the stomach or in the alkaline medium of the intestines.

Example 9.2. Write equations for the acidic and alkaline hydrolyses of aspirin. *Solution.* Acid-catalyzed hydrolysis yields, as usual, constituents of the ester, i. e. acetic acid plus salicylic acid as a phenolic component.



Alkaline hydrolysis results in the same principal products but in the ionized form (one should recall phenolic acidity too).

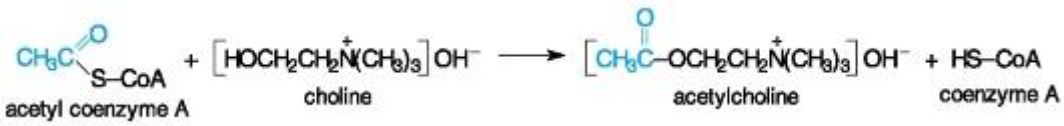


Two other esters of salicylic acid are known as remedies, they are the phenyl ester (the trade nameSalol) and the methyl ester.

Problem 9.4. Write the structural formulas for: (a) phenyl salicylate; (b) methyl salicylate. Show how methyl salicylate can be prepared from salicylic acid.

Esters are among the most widespread of all natural substances. Many simple esters are responsible for the pleasant odour of fruits and flowers. For example, pentyl acetate, $CH_3COO(CH_2)_4CH_3$, is a constituent of banana oil; octyl acetate, $CH_3COO(CH_2)_7CH_3$, has been isolated from orange oil; butyl butyrate, $CH_3CH_2CH_2COO(CH_2)_3CH_3$, has been found in pineapple oil; benzyl acetate, $CH_3COOCH_2C_6H_5$ smells of jasmine. As we will see further (Chapter 12), fats and oils from animal and vegetable sources are also esters.

Ester formation is an important type of reactions, which take place in living matter. Acyl halides and anhydrides are too reactive to be cell constituents because they are rapidly hydrolyzed by water. The best acylating agents *in vivo* are thioesters, acid derivatives of moderate reactivity. An example is *acetyl coenzyme A* (Sec. 8.2.3) usually written in a shortened form $CH_3C(O)S$-CoA. Acetyl coenzyme A serves as an acylating agent in the enzymic transformation of choline into acetylcholine by the following equation:



Acetycholine is then hydrolyzed in the cell. The direct and back reactions make up acetylcholine cycle that is the basis of the nervous conductivity.

Problem 9.5. Complete the following equations:

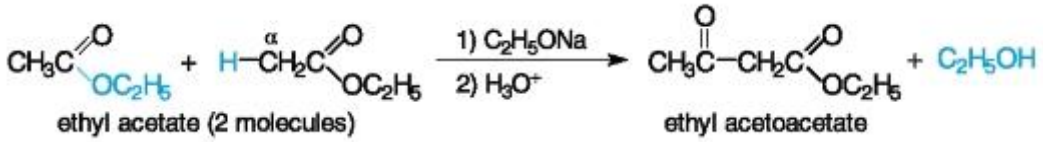
(a) $CH_3COOCH_3 + NaOH \rightarrow$

(b) $(CH_3CO)_2O + CH_3NH_2 \rightarrow$

(c) $CH_3CONHC_2H_5 + KOH \rightarrow$

9.4. ESTER CONDENSATION

Esters, like aldehydes and ketones, are weakly acidic when an α-hydrogen is present in a molecule. A reversible condensation reaction occurs on treatment of an ester with a strong base to give a β-keto ester, for example:

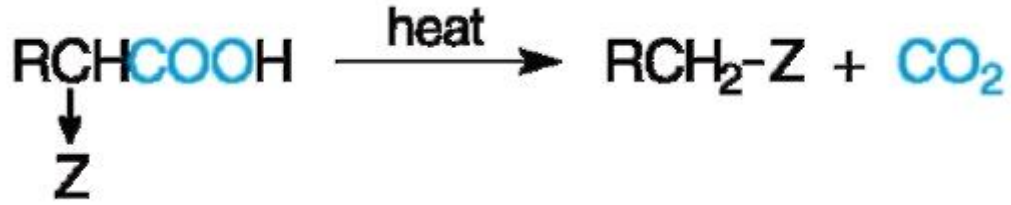


This reaction called the *Claisen condensation* is based on CH-acidic properties of the α-hydrogen thus resembling the aldol condensation in many respects (Sec. 8.3.2).

The Claisen condensation is very important both in nature and in the laboratory. Like the aldol condensation, it provides lengthening carbon chain in the biosynthesis of many natural products. For example, long-chain fatty acids, constituents of natural fats and oils, are produced biosynthetically from the simple two-carbon precursor, acetic acid, in the form of acetyl coenzyme A. It undergoes the Claisen condensation to form acetoacetyl CoA, a four-carbon unit that is similar to the product in the example above. Thus, in eight steps (which involve additional transformations) two acetic acid units combine into a butyric acid unit. Further repetition of the cycle yields a six-carbon unit, and so on.

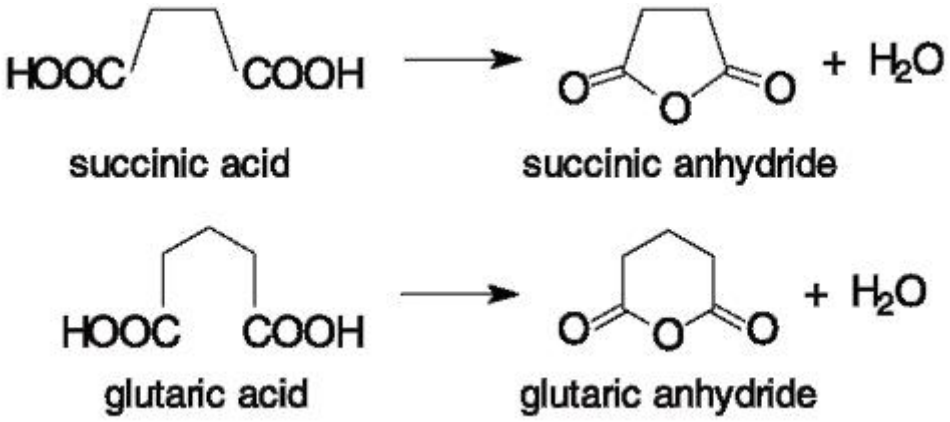
9.5. DECARBOXYLATION OF CARBOXYLIC ACIDS

Decarboxylation, or the loss of carbon dioxide, is one of important transformations of a carboxyl group. Monocarboxylic unsubstituted acids resist decarboxylation and can eliminate carbon dioxide only under severe conditions or by means of specific reagents. Electron-withdrawing substituents at α-carbon (Z is -COOH, -CO-R', -C≡N, $-NO_2$ in the equation below) facilitate decarboxylation.



For this reason, oxalic and malonic acids eliminate carbon dioxide on moderate heating (140-150 °C) thus converting to mono carboxylic acids having one less carbon atom:

Succinic and glutaric acids, dicarboxylic acids with a longer chain, are not decarboxylated on heating. Instead, they lose the water molecule on heating to give *cyclic anhydrides* with a stable fiveor six-membered ring, respectively; the same applies to phthalic acid.



Problem 9.6. Which isomer of benzenedicarboxylic acids can form a cyclic anhydride on heating? Write an equation for the reaction and name the product.

Decarboxylation of various heterofunctional carboxylic acids will be discussed in Chapters 11 and 15.

Additional Problems

9.7. Write structural formulas for each of the following compounds:

(a) 2-pentenoic acid;

(d) N-methylbutanamide;

(b) diethyl oxalate;

(e) benzoic anhydride;

(c) dimethylformamide;

(f) isobutyryl chloride.

9.8. Give the IUPAC names for these compounds:

(a) $C_6H_5CH_2CH=CHCOOH$;

(d) $CH_2=CHCONH_2$;

(b) C6H5CH2COOCH(CH3)2;

(e) (CH3CH2CO)2O;

(c) $CH_3OOCCH_2CH_2CH_2COOCH_3$;

(f) $BrCH_2C(O)Cl$.

9.9. Acetic acid boils at 118 °C, but its methyl ester boils at 57 °C. Why is the boiling point of the acid much higher, even though its molecule is smaller in size?

9.10. Write equations for the reactions of propionic acid with the following:

(a) isopropyl alcohol in acidic medium;

(b) aqueous sodium hydroxide;

(c) aqueous ammonia;

(d) ethanolic potassium hydroxide.

Name the products obtained.

9.11. Write equations for the reactions of methyl butyrate with the following:

(a) aqueous sodium hydrogencarbonate;

(b) dimethylamine;

(c) excess ethanol in the presence of an acid.

Name the products obtained.

9.12. Which of the following statements is correct for ethyl acetate?

(a) it contains sp^2- and sp^3-hybridized carbons;

(b) conjugation results in the shortening of all C-O bonds;

(c) it can be hydrolyzed *only* under acidic conditions;

(d) it reacts with methanol in the presence of an acid catalyst;

(e) in the reaction with methylamine it produces acetamide.

9.13. Write equations for the following reactions and name the products:

(a) neutralization of benzoic acid with aqueous ammonia;

(b) bromination of the same acid.

Compare acidity of benzoic acid and a product obtained in the reaction (b).

9.14. What would you predict as the product of the hydrogen bromide addition to acrylic (propenoic) acid? Write an equation for the reaction and explain the result.

PART 3. POLY- AND HETEROFUNCTIONAL COMPOUNDS IN LIVING SYSTEMS. Chapter 10. STEREOISOMERISM

As it has already been discussed in Chapter 1, constitutional (or structural) isomers are compounds that have the same molecular formula but different structural formula. There is another type of isomerism, namely stereoisomerism, which is important not only in organic chemistry but, especially, in bioorganic and biological chemistry. The branch of chemistry that concerns with the spatial aspects of molecules is called *stereochemistry*.

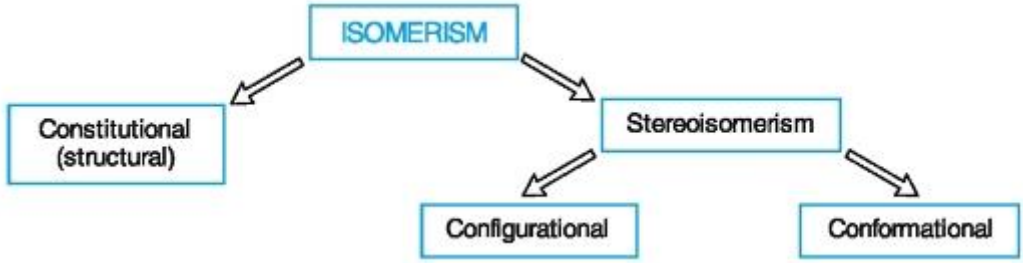
Stereoisomers are compounds that have the same order of atoms attachment,  but differ only in the arrangement of their atoms or groups in space.

In other words, stereoisomers are spatial isomers. They can be subdivided into two categories according to the principles by which they are interconverted:

• *configurational* isomers that can be interconverted only by breaking a covalent bond and creating a new covalent bond;

• *conformationa!* isomers that can be interconverted by rotation about single bonds.

The principal types of isomerism occurring in organic chemistry are summarized below:



Conformational isomerism has been discussed in Chapter 6; in this chapter configurational isomerism will be considered.

Molecules differing only in configuration are termed configurational isomers.

The configuration of a molecule of definite structure is the arrangement of its atoms or groups in space without regard to arrangements that differ only due to rotation about one or more single bonds.

10.1. CHIRAL AND ACHIRAL OBJECTS

A *chiral* object is one that exhibits the property of «handedness». The word *chiral* (pronounced «kai-ral») comes from the Greek word *cheir*, meaning hand. Many familiar objects are chiral, for example, hands and gloves, feet and shoes, helices (including the double helix of DNA), screws and other threaded objects. It is well known that the reflection of a left hand is not another left hand, but a right hand (Fig. 10.1, *a).* So both hands cannot be superposed (placed one upon another). Thus a left and right hand are said to be enantiomers. As we shall soon see, a similar relationship is correct with molecules.

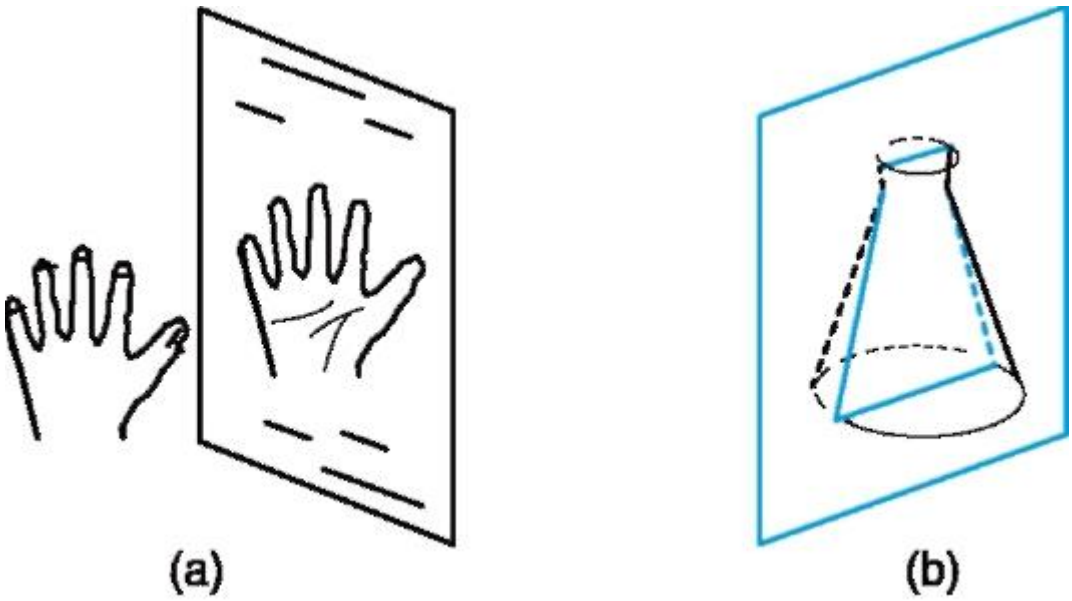


Figure 10.1. The mirror reflection of a chiral object (a) and a symmetry plane that cuts an achiral object (b).

An object and its nonsuperposable mirror image are called enantiomers.

Achira! objects do not have the property of handedness. Typical achiral objects are balls, spheres, cubes, etc. An achiral object can be recognized by the fact that its mirror image is superposable (or identical) with the object itself. It is evident without using a mirror that the mirror image of a laboratory flask is the same flask.

How is it possible to solve quickly whether an object (or molecule) is chiral or achiral? One way, of course, is to compare the object directly with its mirror image. The simpler way is to examine symmetry properties of an object. One of the features of achiral objects is the presence of a *plane of symmetry*[1].This is an imaginary plane that cuts an object in such a way that the two halves of the object are reflections of each other. For example, a spoon has a plane of symmetry and a flask has the infinite number of the planes (only one of them is shown in Fig. 10.1, b). Thus, any object with at least one plane of symmetry is achiral. On the contrary, chiral objects do not have a plane of symmetry.

Problem 10.1. Which of the following objects are chiral: (a) ear; (b) fork; (c) hammer; (d) snail; (e) coat; (f) screwdriver; (g) teacup.

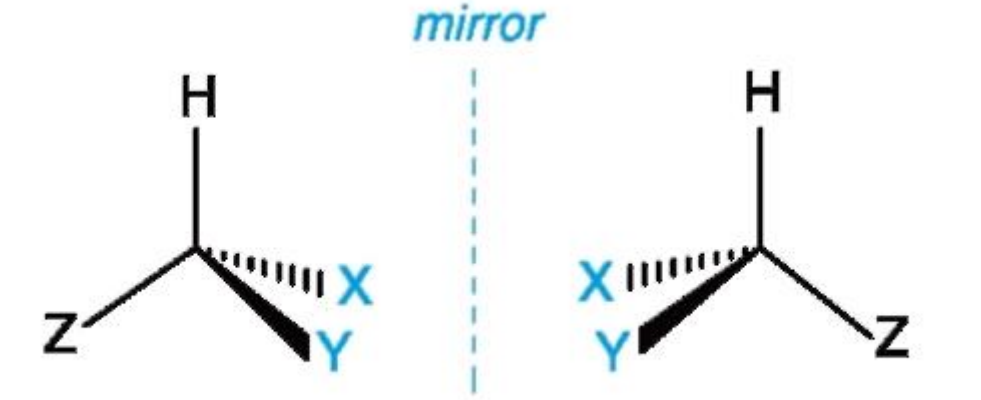
10.2. OPTICAL ACTIVITY

It was discovered at the early stage of organic chemistry (long before the development of the structural theory of organic chemistry) that some substances rotate the plane of polarized light. This property is known as *optical activity,* and the substances rotating plane-polarized light are called optically active.

[1] Other elements of symmetry are beyond the scope of this text.

This phenomenon is a subject matter of physics. Here, we state very briefly that planepolarized light represents electromagnetic waves vibrating in a definite plane, whereas an ordinary light beam consists of waves oscillating in random directions perpendicular to its path.

J. H. van't Hoff and J. A. Le Bel made a hypothesis about carbon (Sec. 1.1) that would explain the optical activity of some organic molecules. They noticed that when four *different* atoms or groups are attached to carbon having tetrahedral configuration, two arrangements are possible. These arrangements result in rightand left-handedness, i. e. in the objects which are related as nonsuperposable mirror images as it is presented below for the generalized molecule of the type CHXYZ (recall that lines of normal thickness show the bonds lying in the plane of the paper; the group Y lies in front of the paper and the group × lies behind it):



mirror images of the chiral molecule CHXYZ

The angle and direction of rotation can be measured with an instrument known as a polarimeter. Some optically active compounds rotate the polarized light clockwise; they are said to be dextrorotatory and are designated by the sign (+). Others, levorotatory compounds, rotate the light counterclockwise and are designated by the sign (-).

The observed rotation, α, depends, first of all, on the structure of a compound and also on the number of molecules in a sample tube, the wave length of the polarized light, the temperature, and a solvent (it is often used in polarimetric analysis). In order to standardize measurements, a value of the optical rotation is usually expressed as *specific rotation*, [α], which includes all the aforementioned factors. The specific rotation of a chiral compound is a physical characteristic of the substance like its melting point, boiling point, density, solubility, and others.

10.3. ENANTIOMERS

Now spatial relationships will be applied to molecules in more detail. Let us look at three-dimensional (3D) formulas of two alcohols, 2-propanol, $CH_3CH(OH)CH_3$, and 2-butanol, $CH_3CH(OH)CH_2CH_3$, as well as their mirror images. (It will be very helpful to use molecular models when studying this chapter. About self-made models see Sec. 1.6.)

The molecule of 2-propanol has a plane of symmetry, which is the plane of the paper in Fig. 10.2. The molecule is therefore achiral. The mirror image is superposable on the original molecule after 180-degree rotation about the C-H bond and both structures are identical.

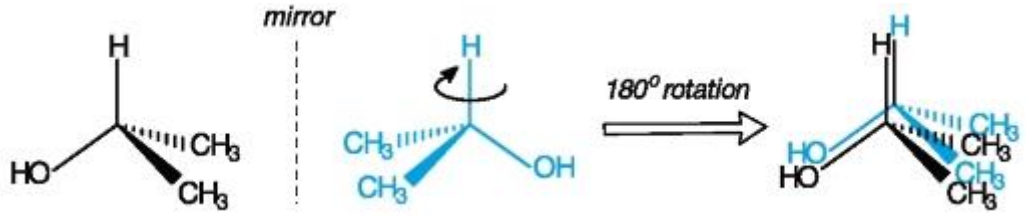


Figure 10.2. 3D structure of 2-propanol and its mirror image (in colour).

On the contrary, one of the possible structures of 2-butanol is not identical with its mirror image and we fail in our attempts to superpose them (Fig. 10.3). After partial superposition of the mirror image on the original molecule the positions of the methyl and ethyl groups are not superposed. Thus, the molecule of 2-butanol is chiral. There are two different structures of 2-butanol, which correspond to two *different compounds*, enantiomers. They are also called *optical isomers* because of their difference in relation to plane-polarized light. Being rightor left-handed, enantiomers rotate polarized light in opposite directions.

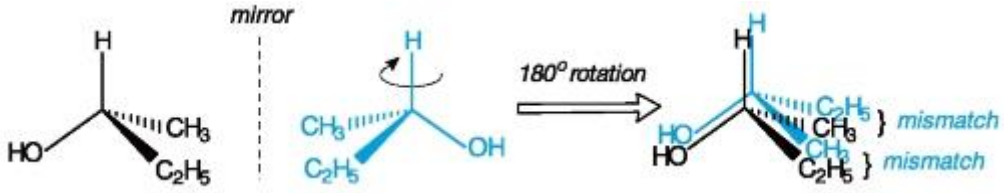
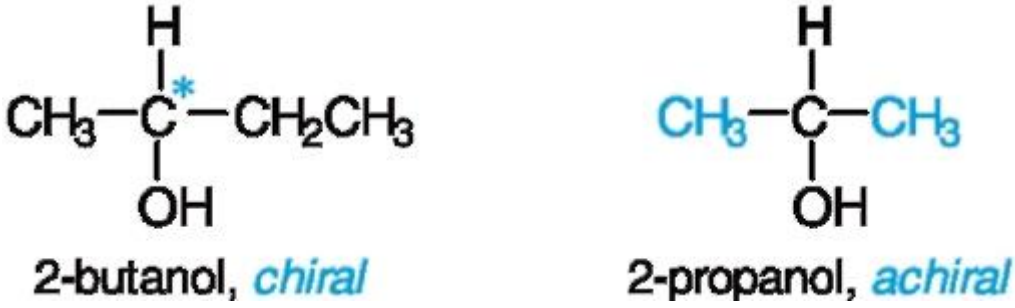


Figure 10.3. One of the 3D structures of 2-butanol (on the left) and its mirror image (in colour).

What is the difference in chirality between 2-propanol and 2-butanol? In 2-butanol, the C-2 atom has four different groups attached to it, namely H, OH, CH_3, and C_2H_5. By convention, such carbon is often marked by an asterisk. Long ago, Van't Hoff and Le Bel defined a carbon atom with four different groups attached to it as the asymmetric carbon. Nowadays chemists refer to such carbon as a *chiral centre*, or*stereogenic centre.* If at least two groups attached to a tetrahedral carbon are the same, a molecule has a plane of symmetry and is achiral. In the 2-propanol molecule such groups are two methyl groups.



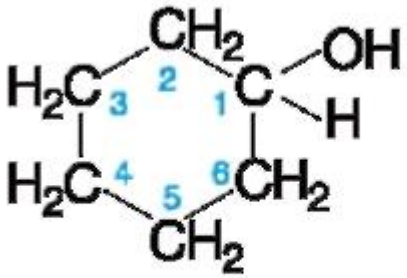
Detecting a chiral center in a complex molecule is the simplest way to solve whether the molecule is chiral or not. Evidently, carbon atoms of the CH_3 and CH_2 fragments cannot be chiral because they have at least two identical groups.

Problem 10.2. Pick out chiral centre(s), *if any*, in the following compounds: (a) 2,3-dimethylhexane; (b) 2,3-butanediol; (c) 2,3-dimethylbutane; (d) 1-bromo-1,2-dichlorobutane.

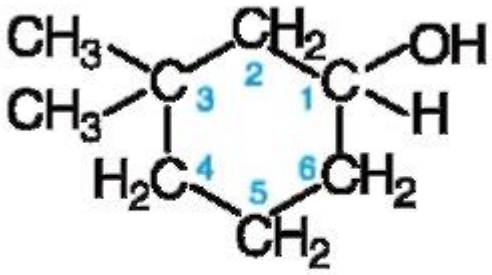
Some difficulties in determining a chiral centre in cyclic structures can arise. Let us consider an example.

Example 10.1. Are the molecules of cyclohexanol (I) and 3,3-dimethylcyclohexanol (II) chiral or achiral?

Solution. All carbons of CH_2 fragments in cyclohexanol are achiral, of course. As regards to the atom C-1, it has two identical substituents, namely, carbon chains from C-2 through C-6 and from C-6 through C-2. Thus, there are no chiral centers in the molecule and cyclohexanol



(I)



(II)

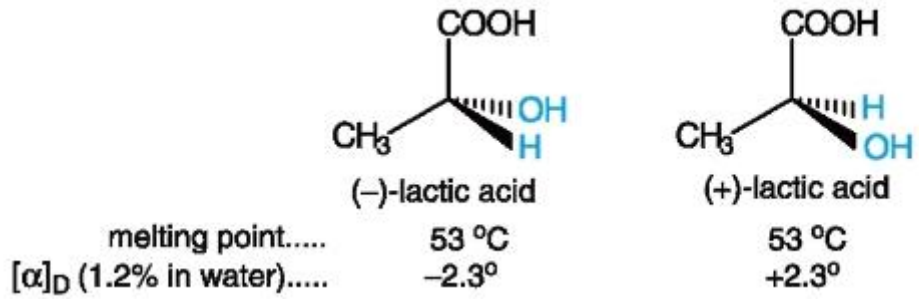
is therefore an achiral compound. We can arrive at the same conclusion after finding a plane of symmetry that passes through the OH group and the atoms C-1 and C-4.

In 3,3-dimethylcyclohexanol, the atom C-1 is bonded to four different groups: H, OH, and two *different*carbon chains in spite of the identity of the nearest CH_2 groups. The atom C-3 has two groups CH_3 and is therefore achiral. Nevertheless, the molecule is chiral as a whole.

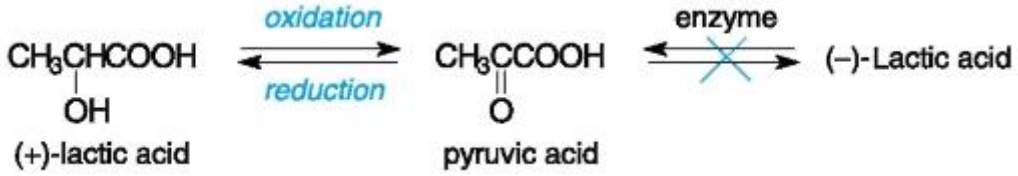
A pair of enantiomers has different properties only with respect to chirality. In all other respects their physical and chemical properties are identical. Thus, both enantiomeric compounds have identical melting and boiling points, solubilities in common solvents, densities, indexes of refraction, and various spectral characteristics. They react similarly and at the same rates with ordinary reagents. Enantiomers have, however, different chiral properties, for example the *direction* in which they rotate polarized light (clockwise or counterclockwise). Though enantiomers rotate a plane of polarization in opposite directions, they have identical specific rotation, except for sign. So polarimetric measurement is the easiest way to distinguish enantiomers from one another.

A mixture containing equal amounts of both enantiomers is called a *racemic mixture*. It is undoubtedly optically inactive and denoted by the symbol (±).

Similarities and differences in properties of enantiomers will be illustrated by lactic (2-hydroxypropanoic) acid. This acid is an optically active compound that takes part in several biochemical processes. It has one chiral centre, the atom C-2. A tetrahedral representation of lactic acid and some of its physical properties are shown below:

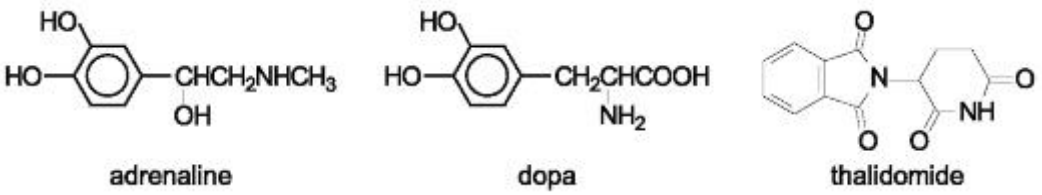


Enantiomers, as it has been mentioned, have similar chemical properties. But they have as a rule different biochemical properties. For example, the enzyme *!actic acid dehydrogenase* oxidizes (+)-lactic acid to pyruvic acid, but it does not oxidize (-)-lactic acid. Similarly, a reversed reaction, enzymic reduction of pyruvic acid, yields the only stereoisomer, i. e. (+)-lactic acid.



The enzyme-catalyzed reactions are said to be *stereose!ective* and *stereospecific* ones. In the stereoselective reaction, one of the possible stereoisomers is exclusively (or predominantly) produced regardless of stereochemistry of the reactant (as in the example above). In the stereospecific reaction, a particular stereoisomer gives a specific stereoisomeric product. The reason for high selectivity and specificity of reactions occurring in nature (so-called *in vivo* reactions) is that biochemical catalysts (enzymes) are chiral themselves.

Enantiomers differ in many types of biological activity. Only one enantiomer of natural amino acids, $RCH(NH_2)COOH$,participates in protein biosynthesis. Only one enantiomer serves usually as a drug, whereas its reflection may be ineffective. For example, only (-)-adrenaline is a cardiac stimulant. From two enantiomers of the amino acid dopa only one, namely levorotatory isomer, is used as an anti-Parkinsonian agent. It is worse when an enantiomer of a drug may be sometimes harmful as it happened with the sedative thalidomide.



Thus, chirality is a factor of great importance in the biological world.

Problem 10.3. Pick out the chiral centre in the molecules of adrenaline, dopa, and thalidomide.

10.4. CONFIGURATION AND THE D,L CONVENTION

Three-dimensional formulas are completely useful representation of chiral molecules. However, this formulas are inconvenient to use, especially, as we shall see later, for the molecules containing several chiral centres. E. Fischer[1] suggested a method of representing a tetrahedral model onto a plane. Two-dimensional formulas thus obtained are called The Fischer projection formulas, or simply the *Fischer projections.*

In the Fischer projection of a spatial formula, atoms or groups attached to a chiral centre are projected onto a plane of the paper in the following way. The atoms or groups appearing above and below the central atom lie *behind* the plane of the paper and those appearing to the left and to the right of the central atom lie *in front* of the plane of the paper. By convention, the carbon chain is disposed vertically with the lowest numbered carbon atom at the top. This approach is demonstrated in Fig. 10.4 for (+)-lactic acid.

[1] Emil Fischer (1852-1919), a great German chemist, the second Nobel Prize winner for chemistry (1902). He conducted monumental research in the chemistry of carbohydrates, amino acids, proteins, heterocycles, and in stereochemistry. When the author of this book, as a young scientist, was going to start his research in carbohydrate chemistry (in the 1960's), a prominent Russian chemist remarked: «What would you plan to study in this field? Fischer already investigated everything in carbohydrates long ago, you see.»

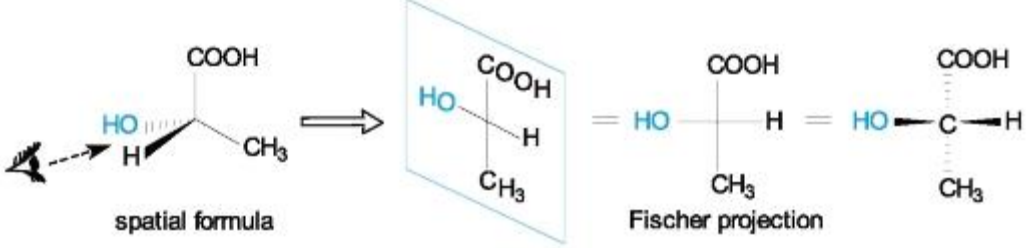


Figure 10.4. Projecting a 3D formula of (+)-lactic acid onto a plane to form the Fischer projection.

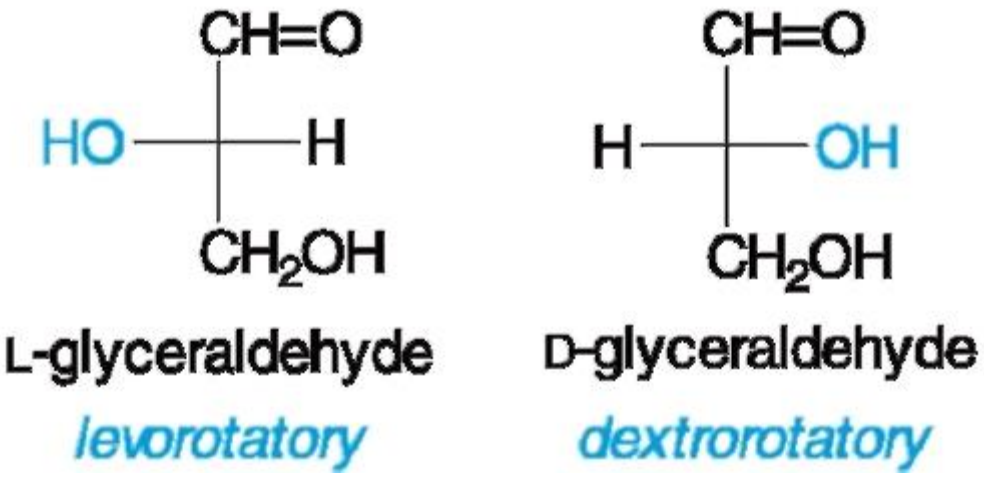
The chiral carbon is usually omitted in the Fischer projection and is represented by two crossed lines. All the bonds are shown by the lines of normal thickness but it should always be remembered that horizontal lines are directed towards the viewer and the vertical lines are directed away, as in the rightmost formula in Fig. 10.4.

A stereoisomer with a functional group at the chiral centre (as the OH group in lactic acid in Fig. 10.4) disposed to the left is called the L enantiomer. When a substituent is to the right an enantiomer belongs to the D series. Thus, the (+)-lactic acid shown in Fig. 10.4 is L-lactic acid. Note that it is *not* levorotatory acid, it is (+)- or dextrorotatory substance!

Relative and absolute configurations. At the beginning of the 20th century, glyceraldehyde was chosen as the configurational standard because it could be converted into other natural compounds. The (+) enantiomer was assigned arbitrary to the D series, and (-)-glyceraldehyde was assigned, of course, to the L series.

Other compounds with a structure similar to glyceraldehyde (lactic acid, natural amino acids, carbohydrates) were assigned to the D or L series after their chemical conversion into D- or L-glyceraldehyde, respectively. Such configurations were called *relative* ones.

Half a century later, it was established that



(+)-glyceraldehyde really had the D configuration. This means the hydroxyl group at atom C-2 is disposed to the right in a Fischer projection. Therefore, all configurations determined earlier as relative ones are, in fact, *absolute*configurations.

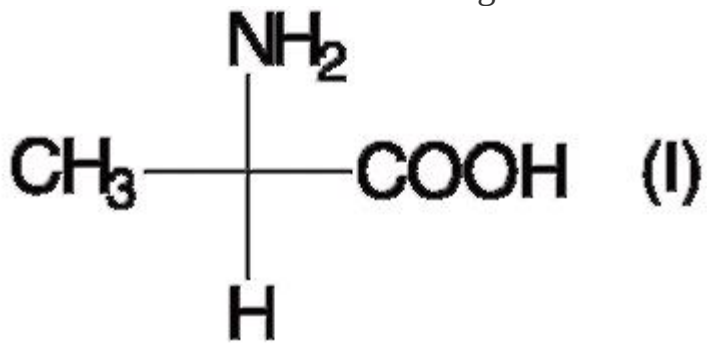
It should be noted, that if the Fischer projection is rotated through 180° in the plane of the paper, the upward and downward bonds still project behind the plane of the paper, and sideways bonds project in front of the plane. If, however, the Fischer projection is rotated through 90° in the plane of the paper, the upward and downward bonds now project in front of the plane of the paper, and the sideways bonds project behind that plane.
The following rules should be emphasized:

- the only interchange of any two substituents in the Fischer projection, like as in a space formula, leads to inversion of configuration;
- two interchanges do not alter the initial configuration.

This means that any Fischer projection can easily be converted into «correctly» oriented one to determine whether a compound belongs to the D or L series. This is illustrated by the following example.

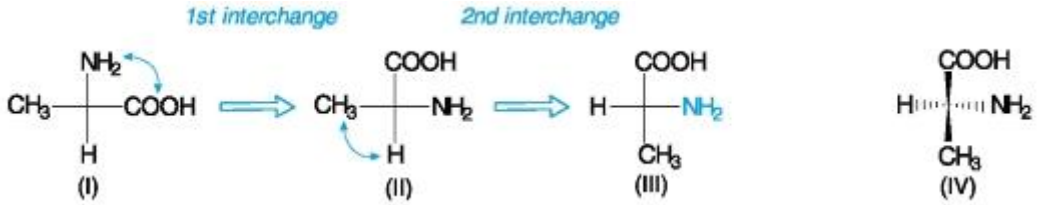
Example 5.2. What is the configuration of 2-aminopropionic acid (trivial name *alanine)* shown by the Fischer projection (I)?

Solution. After the first interchange in the structure (I) we



(I)

get the Fischer projection (II) in which a configuration is altered in comparison with (I), i. e. the projection (II) represents an enantiomer of the compound (I). The second interchange in (II) alters a configuration once more to give the Fischer projection (III) with the initial configuration of alanine. Looking at the projection (III) it is easy to say that it belongs to the Dseries (the amino group to the right). Thus, the Fischer projection (I) represents D-alanine.



Note, the rotation of the Fischer projection (I) through 90° (counterclockwise) is forbidden because the resulting projection (IV) will not be correctly oriented.

The D,L system is widely used for designating configuration of chiral molecules structurally related to glyceraldehyde, namely for such classes of natural organic compounds as amino acids, hydroxy acids, and, especially, carbohydrates (with some modifications). There are, however, lots of structures where it is impossible to use the projection principles just discussed. This relates to compounds structurally unlike the foregoing classes and to cyclic compounds containing chiral carbon in the cycle.

Problem 10.4. Draw formulas for stereoisomers of glyceric acid (2,3-dihydroxypropanoic acid) and malic acid (hydroxybutanedioic acid) by the Fischer projections and assign them to the D and L series.

10.5. CONFIGURATION AND THE R,S CONVENTION

A newer, general system for designating configuration of chiral molecules has been developed by R.S. Cahn, C.K. Ingold and V. Prelog in the 1960's. It is known as the *R,S* system. Here is how it works.

Let us consider again a 3D structure for the generalized molecule of the type C*abcd.* The four groups attached to a chiral centre are placed in a *priority* (or *preference)* order *a* > *b* > *c* > *d*, where the sign «>» denotes «is preferred to», by a system, which will be explained a little further. The molecule is then viewed from the side opposite the *lowest* priority group, *d*, like a driver looking at steering wheel (Fig. 10.5, a). If the remaining three groups (a, b, and c) are arranged clockwise, the configuration is symbolized by *R* (Fig. 10.5, b). If they form a counterclockwise array, the configuration is symbolized by *S*(Fig. 10.5, c). The symbols *R* and S came from the Latin words *rectus* and *sinister* meaning right (or for right hand) and left (or for left hand), respectively.

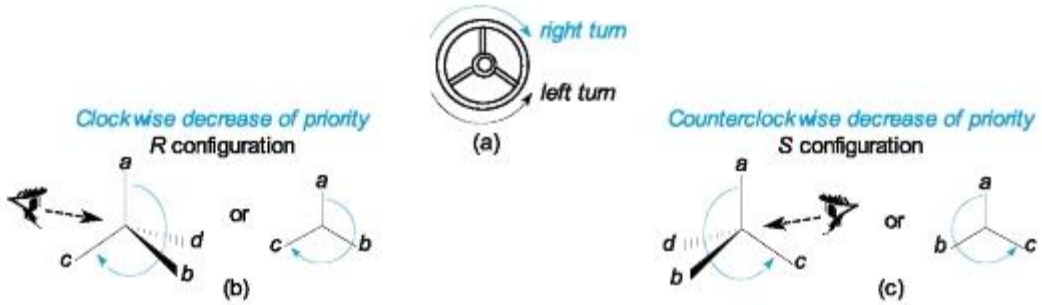


Figure 10.5. Assignment of configuration according to the *R,S* system (a) and (b), and turns of a steering wheel (c).

The priority order is set in the following way.

• Atoms attached *directly* to the chiral centre are first arranged according to decreased atomic number. The higher the atomic number, the higher the priority; for example:

$$_{35}Br > {}_{17}Cl > {}_{16}S > {}_{9}F > {}_{8}O > {}_{7}N > {}_{6}C > {}_{1}H$$

Lowering of priority ⇨

• If two or more of the atoms directly attached to the chiral centre are the same, then the priority order is determined by the next closest atom until a decision is reached. For example: $-CH_2OH > -CH_2NH_2 > -CH_2CH_3 > -CH_3$

• Double-bonded atoms are considered as two single-bonded ones, i. e. carbon of the aldehyde group, -CH=O, is conditionally bonded to «two» oxygens; in the same manner as carbon of the carboxyl group, -COOH, is bonded to «three» oxygens. This rule applies to triple-bonded atoms too. For this reason, the following order of decreased priority is seen for some groups: -COOH > -CHO > $-CH_2OH$ and $-C\equiv N > -CH_2NH_2$.

Problem 10.5. Assign a priority order to these sets of groups:

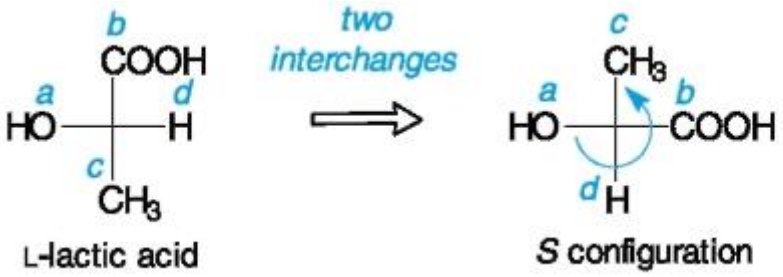
(a) -H, -Cl, $-CH_2OH$, $-CF_3$;

(b) -COOH, $-CH_2SH$, -OH, $-NH_2$;

(c) -CHO, $-OCH_3$, $-NHCOCH_3$, -SH.

Three-dimensional representation of a chiral compound was applied up till now for the *R,S* designation. The Fischer projections can also be useful to achieve this. First, assign a priority order to the four groups in the usual way. Next, by allowed interchanges (see the preceding section), place the lowest priority group either at the bottom or at the top of the projection. Finally, determine clockwise or counterclockwise decreasing of priority of the three remaining groups.

Example 10.3. Which is the configuration of L-lactic acid, *R* or *S? Solution.* Assign a priority order to all substituents at the chiral carbon of L-lactic acid (the Fischer projection to the left), which is HO > COOH > CH_3 > H.



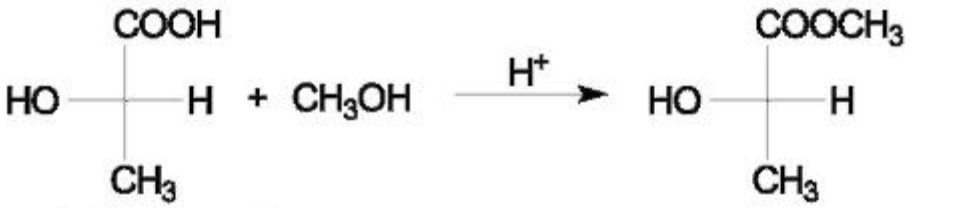
L-lactic acid S configuration

Next, bring the lowest priority group, hydrogen, to the bottom (interchange *d* and *c)* and make, to retain the initial configuration, the second interchange (for example, *b* and *c).* Finally, observe counterclockwise decreasing of priority for substituents *a, b,* and *c.* Thus L-lactic acid is *(S*)-lactic acid.

It is very important to stress once more:

There is no evident correlation between configuration (R or S, as well as D or L) and sign of rotation, (+) or (-).

As it was mentioned before, L-lactic acid (the *S* enantiomer) is a dextrorotatory compound. An obvious case is transformation of L-lactic acid into its levorotatory methyl ester. In this reaction the sign of rotation changes whereas the configuration of the chiral centre remains unchanged.



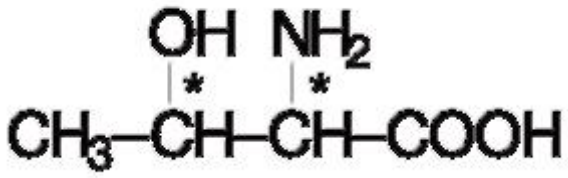
L-(+)-lactic acid (–)-methyl L-lactate

10.6. MOLECULES WITH MORE THAN ONE CHIRAL CENTRE

Many organic compounds, especially those important in biology, have more than one chiral centre. The first task is to determine how many stereoisomers exist in such cases and how the stereoisomers are related to one another. Each additional

chiral centre doubles the number of stereoisomers. Thus, if a molecule has *n* chiral centres, the total number of stereoisomers will not exceed 2^n or $2^n/2$ pairs of enantiomers.

Let us consider stereoisomers of 2-amino-3-hydroxybutyric acid,



Each of the two chiral centres marked with an asterisk could have the configuration *R* or S. Thus, four stereoisomers in all are possible: R,R; S,S; R,S; and S,R, where the first letter refers to the configuration at the atom C-2 and the second one at C-3. They can be drawn by the Fischer projections as shown in Fig. 10.6.

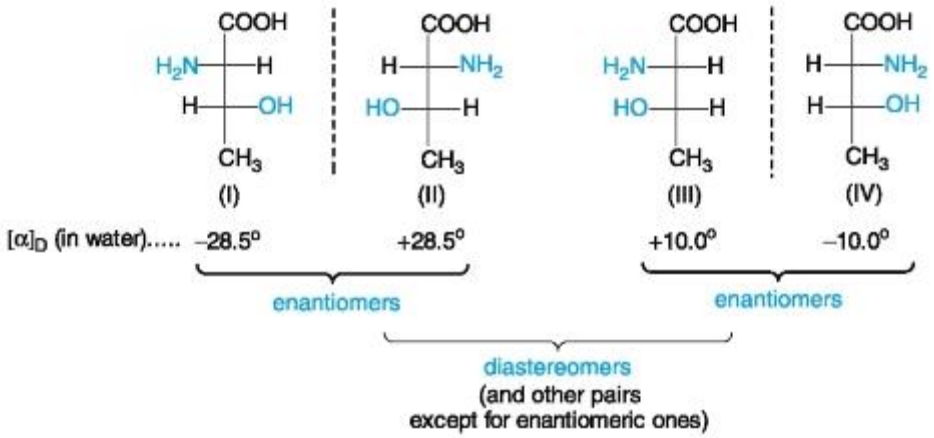


Figure 10.6. Four stereoisomers of 2-amino-3-hydroxybutyric acid.

The Fischer projection (I) represents the essential amino acid L-threonine, whereas the projection (II) is an enantiomer of L-threonine, i. e. D-threonine.

Amino acids, as well as hydroxy acids, containing two chiral centers are designated to the D or L series according to a configuration of the uppermost chiral carbon.

Besides, there are two additional stereoisomers shown by the projections (III) and (IV), which refer to each other as enantiomers too. They have own trivial names: L-allothreonine and D-allothreonine, respectively.
The pair (I) and (II), as well as the pair (III) and (IV), represent mirror images. But the pairs (I) and (III), (I) and (IV), (II) and (III), (II) and (IV) are *not* mirror images because in each pair the compounds have the same configuration at one carbon and opposite configuration at the other carbon. They are undoubtedly stereoisomers, but they are not enantiomers. Such pairs of stereoisomers are called *diastereomers.*

Diastereomers are stereoisomers that are not mirror images of one another.

Note the structural difference between enantiomers and diastereomers: the former have the opposite configurations at *all* chiral centers, whereas the latter have the opposite configurations only at*some* chiral centres, but the same configurations at the remaining centres.

Problem 10.6. Assign configurations at the both chiral centers in L-threonine according to the *R,S*system.

There is a very important difference in properties between enantiomeric pairs and diastereomeric pairs. It should be remembered that a pair of enantiomers has the same chemical and physical properties (except for sign of optical rotation) and they cannot be separated from one another by physical methods, such as crystallization, distillation or chromatography. In contrast to enantiomers, diastereomers differ more or less in all kinds of properties and behave as two different chemical substances in many respects. They can be separated from one another by ordinary means.

10.7. *meso* COMPOUNDS

Consider now stereoisomers of 2,3-dihydroxybutanedioic acid (common name *tartaric acid).* They can be written by means of the Fischer projections (I) - (IV) shown in Fig. 10.7.

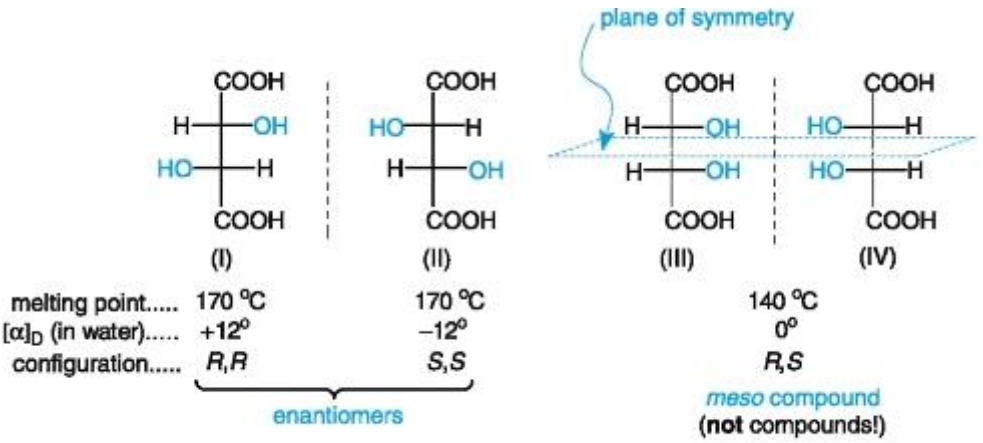


Figure 10.7. Stereoisomers of tartaric acid.

As in the previous example, the structures (I) and (II) represent a pair of nonsuperposable mirror images, i. e. enantiomers. The former is the *R,R* or D enantiomer, the latter is the *S,S* or L enantiomer.
There are some difficulties in assignment of stereoisomers of tartaric acid to the D or L series. Which chiral carbon should be chosen as a determinative one, the upper or the lower? If we use the «amino acid key» (as for threonine in this section), the Fischer projection (I) belongs to the D series (as in this text). But if we apply the «carbohydrate key» (Chapter 14), the projection (I) corresponds to

the L series. The use of the *R,S* convention removes uncertainty, nevertheless the D,L system is still applied.

Two other Fischer projections, (III) and (IV), represent in fact a single compound! It becomes clear after rotating one of them through 180° in the plane of the paper. Such a rotation interconverts these structures making them identical. Both structures have a symmetry plane that is perpendicular to the plane of the paper and bisects the central C-C bond. Each structure has therefore a mirror image superposable on itself and is not chiral even though it contains chiral carbons. Being achiral, this compound is optically inactive. Such structures are called meso compounds.

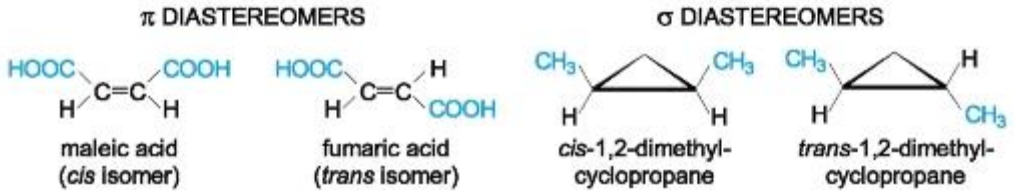
A *meso* compound is an optically inactive achiral stereoisomer containing

chiral centres.

The *meso* structure arises with compounds in which both chiral centres have the *same* four different groups around each centre. Thus, the third stereoisomer of tartaric acid is mesotartaric acid shown by the Fischer projection (III) or (IV) in Fig. 10.7.

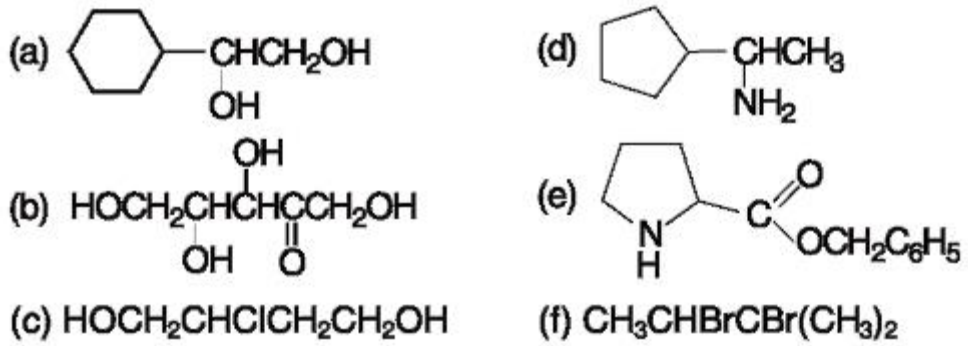
10.8. *Cis-trans* ISOMERISM

In addition to *cis-trans* isomerism of alkenes discussed in Sec. 6.3.1, there is a similar relationship for saturated cyclic structures. Both types of *cis-trans* isomers are related to diastereomers because they cannot be superposed. Stereoisomers around a double bond are classified as π *diastereomers* as well, whereas cyclic *cistrans* isomers called σ *diastereomers.*



Additional Problems

10.7. Mark with an asterisk a chiral centre (or centres) in the following structures:



10.8. Which of the following compounds is (are) chiral:

(a) 2-bromopropanoic acid;

(b) 3-hydroxypropanal;

(c) 1,3-butanediol. Draw formulas for stereoisomers by means of the Fischer projections.

10.9. Which of the following compounds is (are) chiral?

(a) glycerol (1,2,3-propanetriol);

(b) glyceraldehyde (2,3-dihydroxypropanal);

(c) 1,2-butanediol;

(d) 2-cyclohexenol;

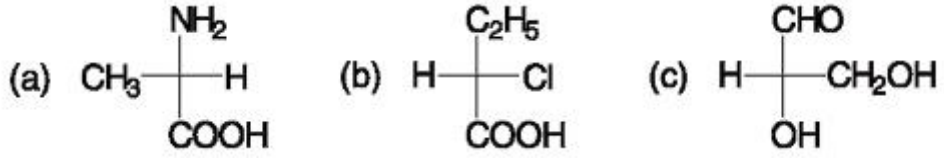
(e) alanine (2-aminopropanoic acid);

(f) β-alanine (3-aminopropanoic acid);

(g) *cis*-1,2-cyclopentanediol.

Pick out an asymmetrical centre (or centres) in chiral molecules.

10.10. Formulate the rules of transforming the Fischer projections. Name the presented compounds, using the d,l system:



10.11. Is it possible to assign configuration of the chiral centre according to the sign of optical rotation in the following compounds?

(a) (+)-alanine (2-aminopropanoic acid);

(b) (+)-glyceraldehyde (2,3-dihydroxypropanal);

(c) (±)-lactic acid;

(d) (-)-malic (hydroxybutanedioic) acid. Explain your reasoning.

10.12. Estimate whether the chiral centre in (-)-adrenalin (see the structure) has the *R* or *S*configuration? Is it possible to describe the configuration of the chiral centre in terms of the d,l system?

10.12. Is it possible to recognize, by polarimetric analysis, the stereoisomers in the following pairs?

(a) l-alanine (2-aminopropanoic acid) and β-alanine (3-aminopropanoic acid);

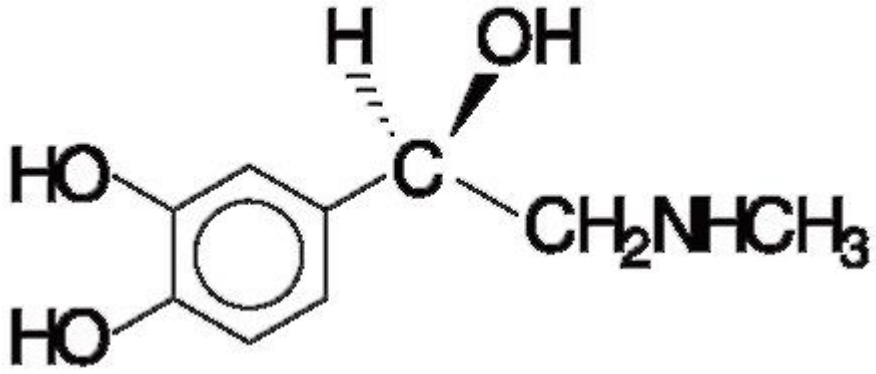
(b) *(R)*- and *(S)*-2-butanols;

(c) d- and l-glyceraldehydes (2,3-dihydroxypropanal);

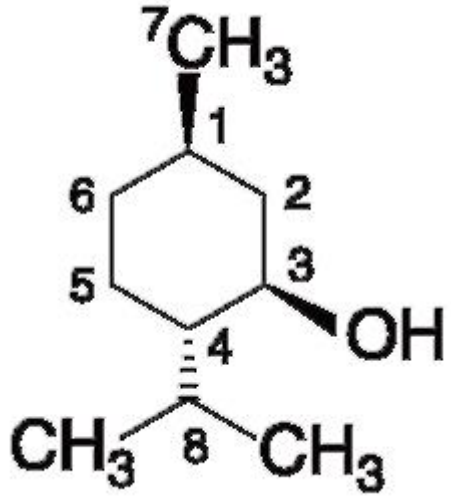
(d) d- and l-lactic acids. Explain your reasoning.

10.14. Draw a structural formula for optically active compound(s) with the molecular formula: (a) $C_4H_{10}O$ and (b) $C_5H_{11}Br$.
10.15. The systematic names of citric acid and isocitric acid are 2-hydroxy- and 1-hydroxy-1,2,3-propanetricarboxylic acid, respectively. Do stereoisomers of these acids exist? If so, mark chiral centres out with an asterisk.

10.16. Draw formulas for all stereoisomers of 2,3-pentanediol by the Fischer projections. Indicate the pairs of enantiomers and diastereomers.



10.17. Pick out the chiral centres in the structure of (-)-menthol presented here. How many stereoisomers of menthol are possible? Draw the structure of (+)-menthol. Suggest a method for distinguishing between (+)- and (-)-menthol.



10.18. Which of the following compounds can exist as *cis-trans* isomers and which as a pair of enantiomers? Draw structural formulas for stereoisomers. (a) 1-bromo-1-butene; (b) 2-bromo-1-butene; (c) 3-bromo-1-butene; (d) 4-bromo-1-butene.

10.19. Which is more stable, cisor trans-1,2-dibromocyclohexane? Comment on your answer by the chair conformations for each compound.

Chapter 11. POLY- AND HETEROFUNCTIONAL COMPOUNDS

Organic compounds of biological importance contain very often not one, but several functional groups. These groups can be identical or different. For example, ethylene glycol, $HOCH_2CH_2OH$, and glycerol, $HOCH_2CH(OH)CH_2OH$, contain two and three hydroxyl groups, respectively. Dicarboxylic acids, such as oxalic, malonic, succinic, have been discussed in the preceding chapter. These compounds are referred to as *polyfunctional* ones.

A significant importance in living systems belongs to *heterofunctional* compounds that involve different functional groups in the same molecule.

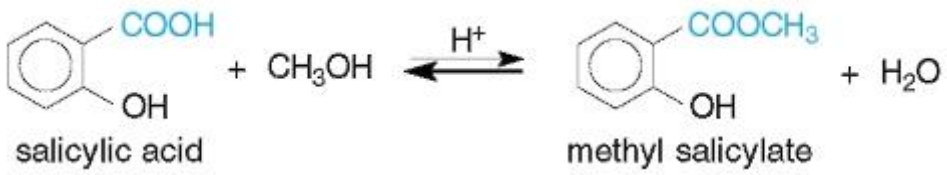
11.1. TYPES OF HETEROFUNCTIONAL COMPOUNDS

Hydroxyl, amino, oxo, and carboxyl groups are encountered most widely in heterofunctional compounds. A combination of different functional groups results in the formation of mixed classes of organic compounds, some of them are given in Table 11.1 (other combinations are possible, of course).

Table 11.1. Some types of combining functional groups in heterofunctional compounds

Heterofunctional classes	*Functional groups*		*Representatives*	
			formula	*trivial name*
Amino alcohols	NH_2	OH	$H_2NCH_2CH_2OH$	Colamine
Hydroxy carbonyl compounds	OH	$>C=O$	$HOCH_2CHCH{=}O$ (OH)	Glyceraldehyde
Hydroxy carboxylic acids	OH	COOH	$HOCH_2COOH$	Glycolic acid
Amino acids	NH_2	COOH	H_2NCH_2COOH	Glycine
Oxo acids	=O	COOH	$CH_3C(=O)COOH$	Pyruvic acid

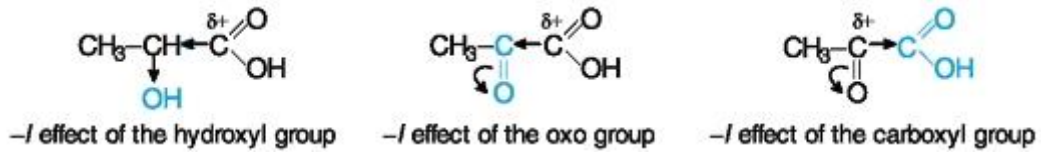
At the first approximation, the chemical behavior of heterofunctional compounds can be represented as a sum of properties of separate monofunctional classes. For instance, pyruvic acid (an oxo acid) can be esterified and transformed into derivatives on its carbonyl group. Salicylic and lactic acids (hydroxy acids) form esters in the reaction with alcohols, as well as their hydroxyl group can be acylated or alkylated. The reaction of salicylic acid with acetic anhydride is used to synthesize aspirin (Sec. 9.3.2). Esterification of the same acid with methanol results in the formation of methyl salicylate.



In consideration of various combinations of functional groups we will mainly attend to new properties arising in such combinations without resort to familiar reactions of individual functional groups.

When the functional groups are close to each other their interaction is more sharply pronounced. This may be illustrated by comparing acidic and electrophilic properties of some heterofunctional carboxylic acids.

In the aliphatic series, all groups listed in Table 11.1 are electron-withdrawing substituents, therefore one group has an influence on another. Thus, lactic and pyruvic acids are stronger (pK_a 3.9 and 2.4, respectively) than propionic acid (pK_a 4.9) for the reasons that were discussed earlier (Sec. 5.2.2). The hydroxyl group in lactic acid and the oxo group in pyruvic acid decrease an electron density on the carboxylic carbon (the leftmost and middle structures below).



On the other hand, the inductive effect of the carboxyl group results in a similar increase of δ+ on the atom C-2 as shown for pyruvic acid in the rightmost structure above. Both carbonyl carbons in pyruvic acid are stronger electrophilic sites as compared with monofunctional three-carbon analogues, i. e. acetone and propionic acid. Therefore pyruvic acid reacts with nucleophiles more readily by both nucleophilic addition and nucleophilic substitution reactions.

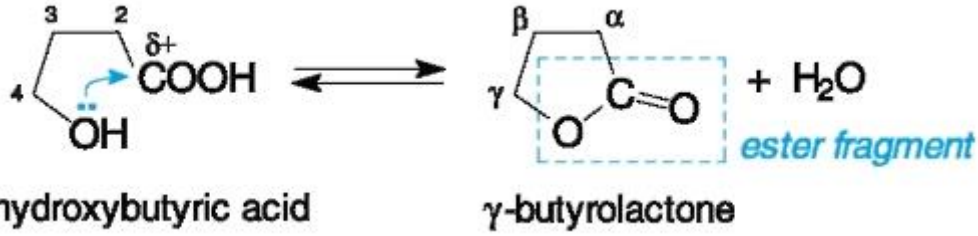
Problem 11.1. Classify, by writing the structural formulas, the following compounds as polyor heterofunctional ones, *if any:* (a) chloroform (trichloromethane); (b) glycolaldehyde (hydroxyethanal); (c) ethylene glycol (1,2-ethanediol); (d) phenetol (ethoxybenzene).

11.2. INTERACTION OF DIFFERENT GROUPS IN HETEROFUNCTIONAL COMPOUNDS

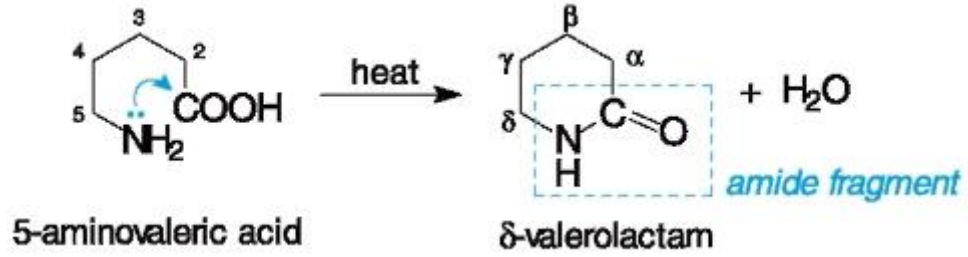
Many functional groups listed in Table 11.1 can affect each other, especially if one of these groups is a carboxyl or carbonyl group. Two types of interaction are possible: intermolecular or intramolecular; the latter occurs when two functional groups occupy favourable positions for such a reaction.

11.2.1. Intramolecular Reactions

Lactones and lactams. Let us begin consideration from a familiar reaction such as ester formation. Fundamentally, the nucleophilic OH group of hydroxy acids can react with its own carboxyl group. Cyclic esters having the generic name *lactones* would be expected in this case. Indeed, cyclization occurs if a thermodynamically stable fiveor six-membered cycle is formed. A molecule takes a claw-shaped conformation to make a better contact of both functional groups. This becomes possible for γ-hydroxy and δ-hydroxy carboxylic acids which undergo internal esterification (lactonization) in an acidic medium or on slight heating or even spontaneously, though slowly, in aqueous solution. For example:



Principally similar transformation proceeds with γ-amino and δ-amino acids. Cyclic amides called*lactams* are the cyclization products.

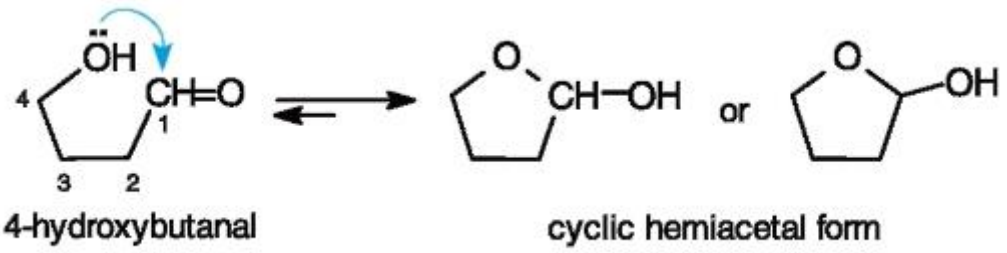


The main difference in these two cyclization reactions is that lactams are never obtained spontaneously or under acidic conditions because amino acids form ammonium salts in the presence of acids (Sec. 5.4). Lactam formation is a thermal reaction.

Problem 11.2. Write equations for the formation of: (a) δ-valerolactone; (b) γ-valerolactam from the corresponding hydroxy and amino acid.

Nomenclature of lactones and lactams. Some common names, derived from the trivial names of non-hydroxylated carboxylic acids, are permitted by the IUPAC rules. Such names were used in the text and in Problem 11.2. Systematically, lactones formed from aliphatic acids are named by adding the suffix -olide to the name of the parent hydrocarbon. A locant is added to define the position of ring closure. Thus, γ-butyrolactone should be named as «4-butanolide», and δ-valerolactone - as «5-pentanolide».
The same principles, but with the use of the suffix -lactam, are applied to nomenclature of lactams. For example, γ-valerolactam and δ-valerolactam have systematic names «4-pentanelactam» and «5-pentanelactam», respectively.

Cyclic hemiacetals. If an aldehyde or a ketone contains a hydroxyl group at an appropriate distance (at the C-4 or C-5 atom for aldehydes), the carbonyl and hydroxyl groups may react with each other. The molecule takes a conformation in which both functional groups are favourably located. The result of intramolecular nucleophilic addition is the formation of a *cyclic hemiacetal*. For example, 4-hydroxybutanal exists as an equilibrium mixture of two forms with predominance (about 94%) of the cyclic one.



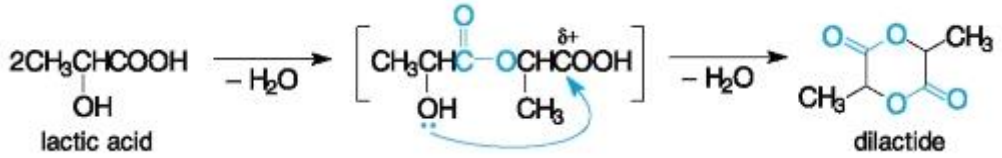
4-hydroxybutanal cyclic hemiacetal form

In general, fiveand six-membered cyclic hemiacetals are more stable than their open-chain counterparts. As we will see later (Chapter 14), a cyclic hemiacetal form is the element of a carbohydrate structure.

Problem 11.3. Write an equation for the formation of cyclic hemiacetal from 5-hydroxy-2-methylhexanal. Which form is predominant in the equilibrium mixture?

11.2.2. Intermolecular Reactions

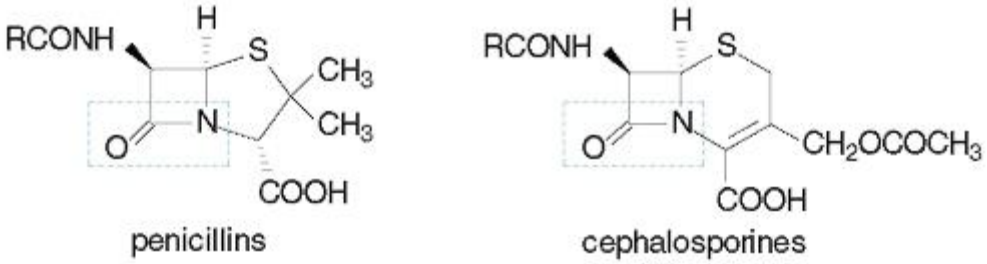
Cyclization reactions. The formation of a threeor four-membered cycle from the α- or β-substituted acid is unfavourable because of a great angle strain in small cycles. Nevertheless, α-hydroxy acids and α-amino acids can react intermolecularly. Lactic acid, for example, undergoes intermolecular esterification on heating with the formation of a dimeric product which, in its turn, converts into more stable six-membered cyclic diester (ester units are coloured in the equation below). The generic name of such cyclic diesters is *lactides.*



lactic acid dilactide

Problem 11.4. Explain why α-hydroxy and γ-hydroxy acids produce quite different products on heating. Illustrate this by the examples of 2-hydroxyand 4-hydroxypentanoic acids.

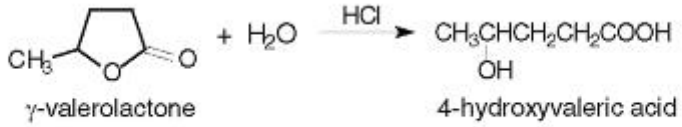
Four-membered cyclic esters and amides, β-lactones and β-lactams, cannot be prepared from the corresponding acids on heating for the reason of their low stability. Nevertheless, they are known and have been found in nature, for example, in antibiotics of *penicillin* and *cephalosporine* series. The general structures of the socalled β-lactam antibiotics are given below (β-lactam fragments are outlined):



penicillins cephalosporines

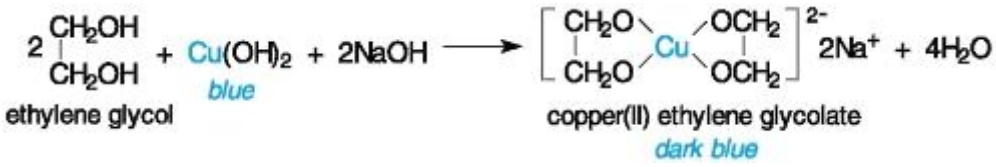
All the cyclic esters and amides mentioned above (i. e. lactones, lactides, and lactams) can be hydrolyzed just as «ordinary» carboxylic acid derivatives with the

ring opening and the formation of the starting heterofunctional compounds; for example:



β-Lactams and β-lactones are highly sensitive to hydrolysis.

Complexing properties. These are another characteristic of heterofunctional compounds. Their complexing (or chelating) ability is based on a tendency to form a stable fiveor six-membered cycle in the reaction with some metal ions (especially with Cu^{2+} and Ni^{2+}). For example, insoluble copper(II) hydroxide reacts with 1,2-diols with the formation of dark blue coloured solution.



Similar complex salts are produced when treating α-amino alcohols and α-amino acids with $Cu(OH)_2$(Sec. 15.1.2).

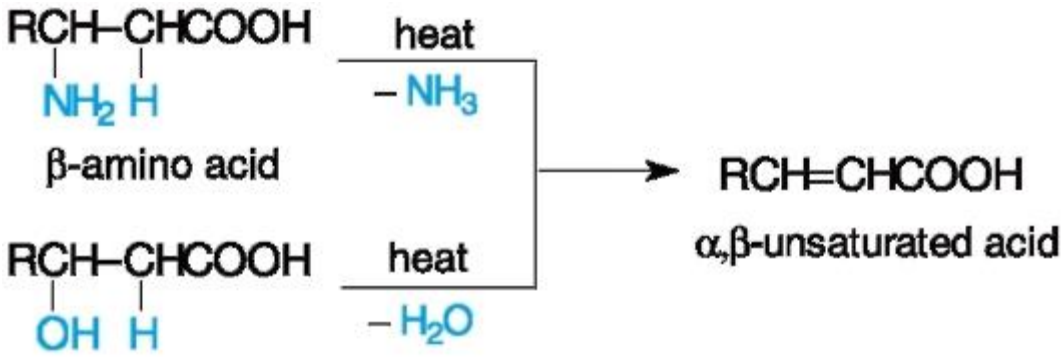
These reactions are used as a colour test for the detection of the vicinal diol fragment in the molecule.

11.3. CH-ACIDIC PROPERTIES OF HETEROFUNCTIONAL COMPOUNDS

Aliphatic compounds of the general formula X-CH_2-Y, in which the substituents × and Y represent electron-withdrawing groups, reveal the property of CH-acids. This is a result of polarization of the C-X and C-Y bonds with subsequent polarization of the C-H bond. Amino acids, hydroxy acids and oxo acids with the second functional group at the atom C-3 (or the β position) belong to compounds of this type.

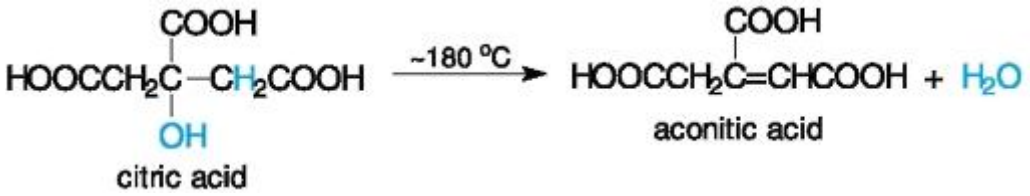
11.3.1. Elimination Reactions

Elimination reactions take place readily on heating β-hydroxy or β-amino acids. Both types of the acids form α,β-unsaturated acid, releasing water or ammonia, respectively:

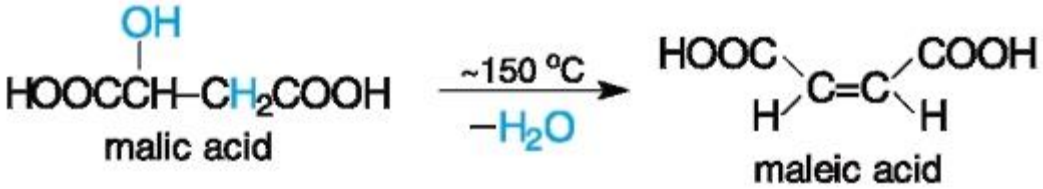


β-hydroxy acid

Citric acid which is simultaneously both the α and β hydroxy acid undergoes similar elimination. This reaction is a chemical analogue of one of the steps in the Krebs cycle.

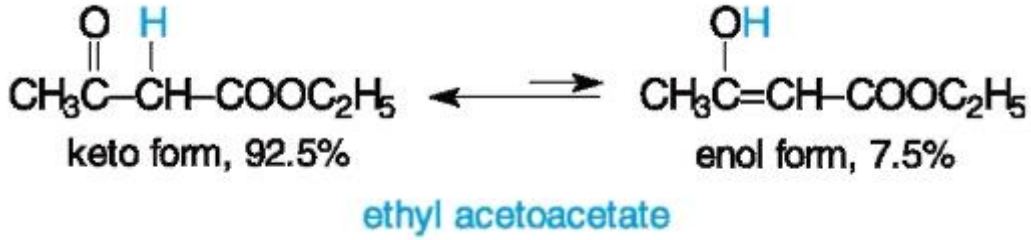


Thermal dehydration of malic acid analogously yields maleic acid (compare this result with enzymic dehydration of malic acid described in Sec. 7.2.4).



11.3.2. Keto-Enol Tautomerism

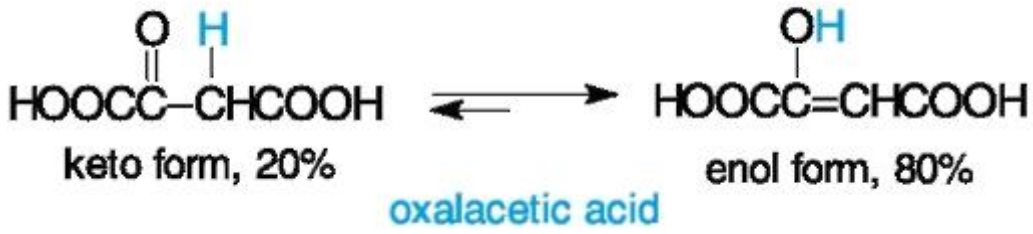
As it has already been shown (Sec. 8.3.1), keto-enol tautomerism is more pronounced for compounds that have a strong CH-acidic site. This is observed in β-oxo carboxylic acids and their derivatives. The most known representative in this series is ethyl acetoacetate, $CH_3C(O)CH_2COOC_2H_5$, commonly named *acetoacetic ester*. It has the pK_a of 10.7 (for a proton of the CH_2 group) which is comparable with phenolic acidity, and it is at least 10^6 times more acidic than alcohols.



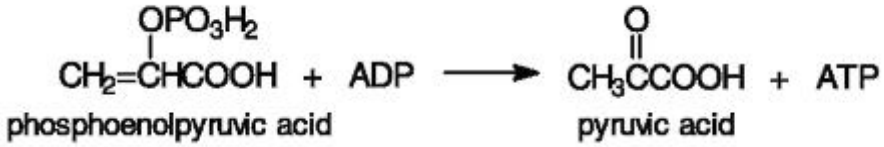
A structure of acetoacetic ester was an object of discussions of many years' standing. This compound first obtained in 1863 was assigned the structure of a hydroxy unsaturated ester (i. e. the enolic structure in modern terminology) because the product reacted with sodium like alcohols and with bromine like alkenes. Shortly after, it was shown that acetoacetic ester also reacted as a typical ketone. Both forms of acetoacetic ester were isolated under specific conditions some decades later. A long-standing controversy finished at the beginning of the 20th century when theconcept of tautomerism was adopted in organic chemistry.

It is interesting that A. Butlerov first predicted, on the basis of his structure theory, the possibility of a dual reactivity of some compounds and of tautomerism as a reversible isomerization. But this phenomenon was first demonstrated by A. Baeyer (in the 1880's) for another type of a reversible transformation, so-called lactim-lactam tautomerism that will be discussed in the next chapter.

Oxalacetic acid is one of the most important carboxylic acids in many biochemical processes. Despite its common name that corresponds to an oxo acid, this compound is a rather unsaturated acid because of the predominance of the enol form in the tautomeric equilibrium.



Phosphoenolpyruvic acid is an example of an enolic compound in living systems. It is produced in the glycolysis process and represents a phosphate of pyruvic acid in the enol form. Phosphoenolpyruvic acid is an energy-rich compound that eliminates energy on its transformation into pyruvic acid. The energy thus released is then accumulated in adenosine triphosphate (ATP) produced from adenosine diphosphate (ADP):



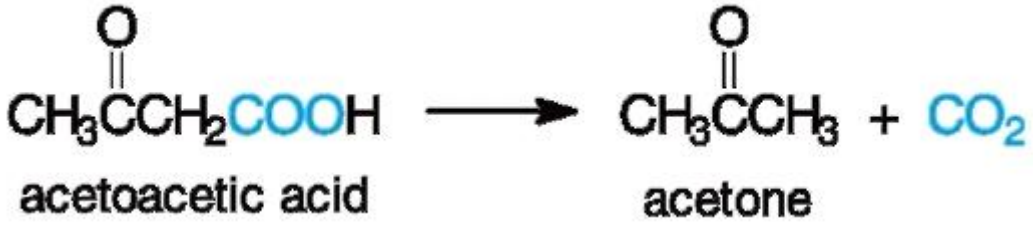
It should be mentioned that pyruvic acid itself contains only negligible quantities of the enol form.

Problem 11.5. Which of the following compounds can exist in the enol form: (a) 2-oxopentanoic acid; (b) 2-oxopentanedioc acid; (c) 3-oxopentanoic acid; (d) 3-oxopentanedioc acid; (e) 2,5-hexanedione? Write the enol structure(s).

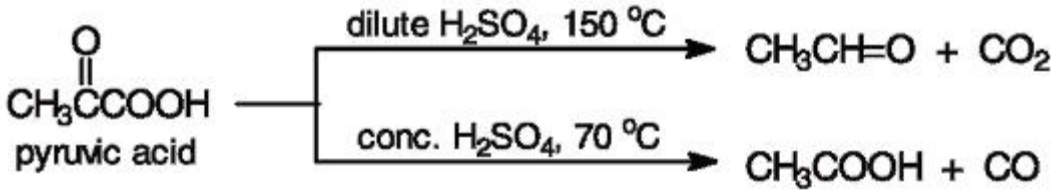
11.4. DECARBOXYLATION OF HETEROFUNCTIONAL CARBOXYLIC ACIDS

Carboxylic acids with strong electron-withdrawing group at the α or β position can be decarboxylated. Amino and hydroxy acids as well as dicarboxylic and tricarboxylic acids are subjected to this reaction.

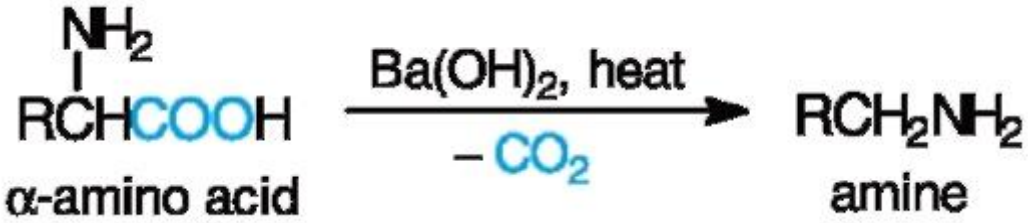
Acetoacetic acid is readily decarboxylated on slight heating to yield acetone.



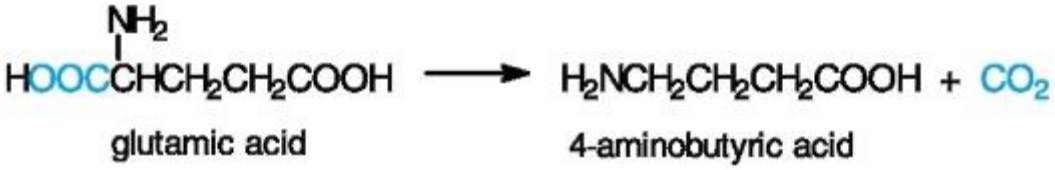
α-Oxo acids eliminate carbon dioxide on heating in the presence of a dilute sulfuric acid whereas a concentrated acid promotes the removal of carbon monoxide.



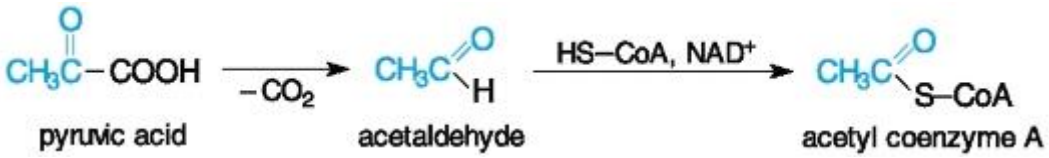
α-Amino acids can be hardly decarboxylated *in vitro* on heating in the presence of strong alkalis to give primary amines.



Decarboxylation plays an essential role in metabolic processes. So, natural α-amino acids are transformed in such reactions into *biogeneous amines.* For example, the dicarboxylic compound glutamic acid gives rise to 4-aminobutyric acid (or γ-aminobutyric acid, abbreviated GABA), which is a natural neuroregulator.



Enzymic decarboxylation of pyruvic acid results in the formation of so-called «active acetaldehyde». It is then oxidized and acylates coenzyme A (abbreviated HS-CoA) at sulfur atom to give acetyl coenzyme A. The latter is a versatile carrier of the acetyl group in biochemical reactions.



11.5. HETEROFUNCTIONAL COMPOUNDS IN CHEMOTHERAPY

Perhaps, the noblest aim of chemistry consists in creation of drugs. The German physician and scientist T. Paracelsus, the founder of medicinal chemistry in the 16th century, said that the real objective of chemistry was not the production of gold but the preparation of drugs. A new science - chemotherapy - that appeared in the 20th century is defined as the use of chemical substances for medical purposes.

The author is aware that the discussion of this problem is beyond the scope of this book. This is a subject matter of other sciences, first of all, pharmacology and pharmaceutical chemistry. Nevertheless, some relatively simple examples of heterofunctional compounds used in medicine are given here.

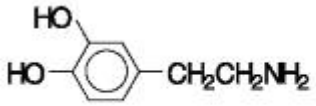
Only trivial or trade names are given in this section because of the complexity of most systematic names. Besides, as it is well known, many medicines have

numerous trade names. We give first a name applied in the Russian Pharmacopoeia and in Russian literature, then synonym(s) of the name in parentheses.

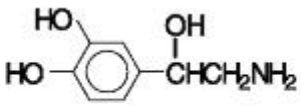
11.5.1. Derivatives of Amino Alcohols and Amino Phenols

A group of compounds called *catecholamines* (from the name *catechol* that means 1,2-dihydroxybenzene) represents biogeneous amines dopamine, noradrenaline (norepinephrine), and adrenaline (epinephrine). These substances are also known as remedies that stimulate adrenoreceptors.

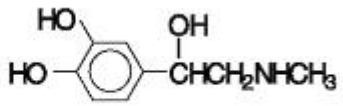
The prefix nor- designates in the IUPAC nomenclature the absence of the methyl group when a trivial name is employed.



dopamine



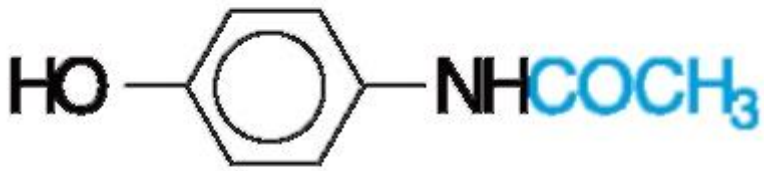
noradrenaline



adrenaline

These and subsequent compounds of basic nature, as well as compounds with acidic nature, are often used as salts (usually as hydrochlorides and sodium or other salts, respectively). These details are omitted mostly for simplicity, in structures given in this section.

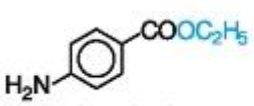
p-Aminophenol is a toxic matter but its N-aceylated derivative Paracetamol (Panadol, Tylenol, Efferalgan, and many other synonyms), is used as antipyretics and anesthetics.



Paracetamol

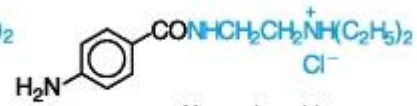
11.5.2. Derivatives of Amino Acids

Some esters of *p*-aminobenzoic acid are widely used as local anesthetics. The oldest of them is the ethyl ester, or Anesthesin (Benzocaine) which is known for more than a hundred years. More effective are an ester Procaine and its soluble salt Novocain. The related compound an amide Novocainamide also has an anesthetic effect but rather it is known as an antiarrhythmic drug.



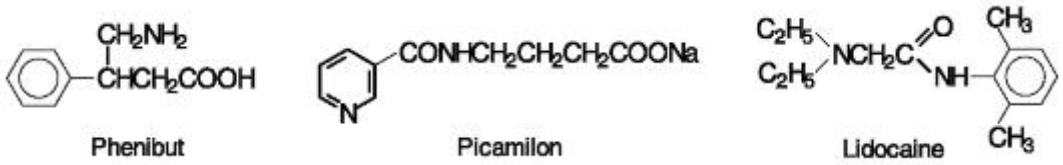
Anesthesin

Novocain



Novocainamide

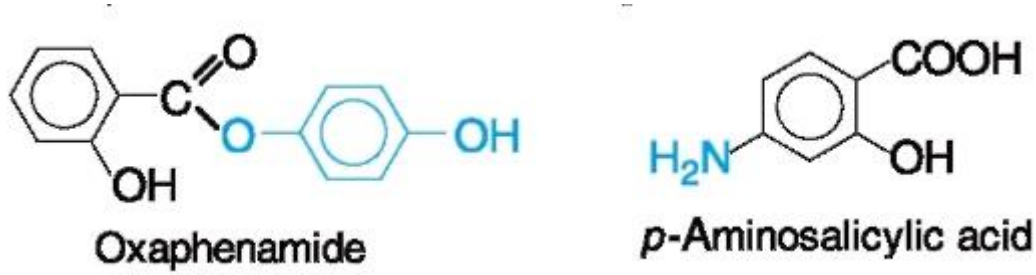
Among aliphatic amino acids, there should be mentioned Aminalon (Gammalon, GABA) and related compounds Phenibut and Picamilon that possess a nootropic activity. One of the most effective anesthetics Lidocaine (Xylocaine) is a derivative of glycine, the simplest amino acid.



11.5.3. Derivatives of Salicylic Acid

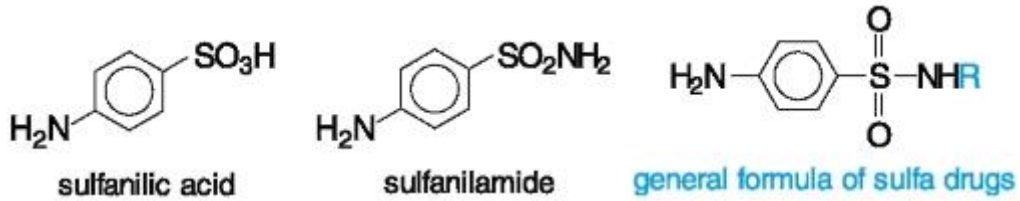
Three esters of salicylic acid were presented in Sec. 9.3.2. Acetylsalicylic acid (Aspirin and over sixty other synonyms) is the first synthetic antipyretic and anesthetic applied since the close of the 19th century. It is, undoubtedly, the most widely used derivative of salicylic acid, both by itself and in combination with other drugs.

Further derivatives are Oxaphenamide (Osalmid), a cholagogue, and p-Aminosalicylic acid (PAS-acid), a tuberculostatic drug.



11.5.4. Derivatives of Sulfanilic Acid (Sulfa Drugs)

Sulfa drugs were among the first synthetic antibacterial remedies. They are derivatives of sulfanilic acid, or *p*-aminobenzenesulfonic acid. Its amide, sulfanilamide, is the parent compound of all the sulfa drugs. In spite of the simple structure, it was found to be effective (by the trade name Streptocide white) against some bacterial infections.



Over 10,000 analogues of sulfanilamide were synthesized within a short time after its discovery. The best therapeutic results were obtained by variations in the amide portion of the molecule, especially by introducing various heterocycles as the R substituent in the general formula. Dozens of sulfa drugs are still used, though more effective remedies appeared.

Additional Problems

11.6. Write the structural formulas for the following compounds and classify them as polyor heterofunctional ones, *if any:*

(a) anisole (methoxybenzene);

(b) 2-chlorobutanoic acid;

(c) ethyl lactate;

(d) iodoform (triiodomethane);

(e) p-toluic acid (4-methylbenzoic acid);

(f) xylitol (1,2,3,4,5-pentanepentol).

11.7. Write two equations for C_{10}-ester formation for salicylic acid. Take into account the presence of different functional groups in the molecule. Name the products obtained.

11.8. Write equations for reactions of p-aminobenzoic acid with the following compounds: (a) sodium hydroxide; (b) hydrogen chloride; (c) acetic anhydride. What is the mechanism for the latter reaction?

11.9. Explain why β-amino and δ-amino acids produce quite different products on heating. Illustrate this by the examples of 3-amino- and 5-aminopentanoic acids.

11.10. Write equations for the formation of γ-butyrolactone and γ-valerolactam from the corresponding hydroxy or amino acid. How would you perform hydrolysis of these products?

11.11. What would you predict as the product obtained on heating ε-amino acids (by the example of 6-aminohexanoic acid)? Compare the result with that for γ-amino acid.

11.12. The enol-to-keto ratio strongly depends on the solvent used. So acetoacetic ester contains in aqueous solution only 0.4% of the enol form that is much lower than in a pure liquid. How can you account for such difference?

Chapter 12. LIPIDS

Lipids (from the Greek *lipos* - fat) are compounds of vegetable or animal origin that are characterized by their solubility properties[1]. They are practically insoluble in water but highly soluble in nonpolar organic solvents. Lipids can be extracted from cells and tissues by organic solvents, such as chloroform, ether, or hydrocarbons.

Lipids vary considerably in their chemical structure. In general they are considered as derivatives of long-chain carboxylic acids. The distinguishing feature of lipids lies in their *biphilic* properties, resulting from the presence in their molecules polar (hydrophilic) and non-polar (hydrophobic) regions. Thus, lipids have an affinity both to water and to non-aqueous phase.

12.1. CLASSIFICATION

Lipids are classified as either simple (two-component) or complex ones; the latter consist of three or more components (Table 12.1). Simple lipids represent esters in which carboxylic acids acylate trihydroxylic alcohol glycerol (as in fats and oils) or long-chain alcohols (as in waxes). Lipids of this group give only alcohols and carboxylic acids on hydrolysis.

Complex lipids may contain other components such as a substituted phosphate group (as in phosphatides) or carbohydrate units (as in glycolipids). Moreover, the amino alcohol sphingosine (instead of glycerol) is a constituent of the sphingolipid group.

Table 12.1. Classification of lipids

Simple		*Complex*		
waxes	*fats and oils**	*phosphatides*	*glycolipids*	*sphingolipids*
RCOOR'	$CH_2O–CO–R$ CHO–CO–R' $CH_2O–CO–R''$	$CH_2O–CO–R$ CHO–CO–R' $CH_2O–P(O)–OR''$ OH R'' = aminoalkyl	$CH_2O–CO–R$ CHO–CO–R' $CH_2O–X$ X = sugar residue	$CH=CH(CH_2)_{12}CH_3$ CHOH CHNH–CO–R $CH_2O–X$ X = phosphate or sugar residue

* Unstated R's are saturated or unsaturated long carbon chains.

[1] Note that this definition is based on a physical property (solubility) differing from that used for other classes. The latter were defined on the basis of their structures.

12.2. STRUCTURAL COMPONENTS OF LIPIDS

There are two obligatory components in all groups of lipids: long-chain carboxylic acids and alcohols.

12.2.1. Fatty Acids

Many of the most common carboxylic acids were first obtained from natural sources, particularly from fats and oils. Consequently, they are called *fatty acids.* Over 200 fatty acids are known today, that differ in the chain length, degree of unsaturation, degree of branching, and even in the presence of additional functional groups. Some of the commonly occurring fatty acids are listed in Table 12.2.

Table 12.2. Common fatty acids

Code symbol *	Structure	Common name	Melting point, °C
	Saturated		
$C_{4:0}$	$CH_3(CH_2)_2COOH$	Butyric	–5
$C_{12:0}$	$CH_3(CH_2)_{10}COOH$	Lauric	44
$C_{14:0}$	$CH_3(CH_2)_{12}COOH$	Myristic	54
$C_{16:0}$	$CH_3(CH_2)_{14}COOH$	Palmitic	63
$C_{18:0}$	$CH_3(CH_2)_{16}COOH$	Stearic	70
$C_{20:0}$	$CH_3(CH_2)_{18}COOH$	Arachidic	75
	Unsaturated**		
$C_{16:1}$	$CH_3(CH_2)_5CH=CH(CH_2)_7COOH$	Palmitoleic	32
$C_{18:1}$	$CH_3(CH_2)_7CH=CH(CH_2)_7COOH$	Oleic	13
$C_{18:2}$	$CH_3(CH_2)_4CH=CHCH_2CH=CH(CH_2)_7COOH$	Linoleic	–5
$C_{18:3}$	$CH_3CH_2CH=CHCH_2CH=CHCH_2CH=CH(CH_2)_7COOH$	Linolenic	–12
$C_{20:4}$	$CH_3(CH_2)_4(CH=CHCH_2)_4CH_2CH_2COOH$	Arachidonic	–50

* The first number represents total carbons, the second - the number of the C=C double bonds.

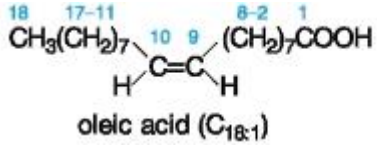
** All double bonds have the *cis* configuration.

In spite of a big variety of fatty acids, most of them are unbranched and contain an *even* number of carbon atoms. This is a consequence of their biosynthesis from the two-carbon fragment of acetyl coenzyme A (Sec. 9.4). The most widespread acids are those containing 16 and 18 carbon atoms. Unsaturated fatty acids usually have the *cis* (or *Z)* configuration and *are not* conjugated.

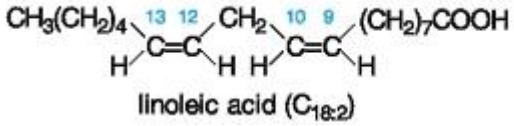
Example 12.1. Write the structures of oleic and linoleic acids, showing the geometry at each double bond.

Solution. All double bonds in both the acids have the *cis* configuration:

18 17-14 11 8-2 1

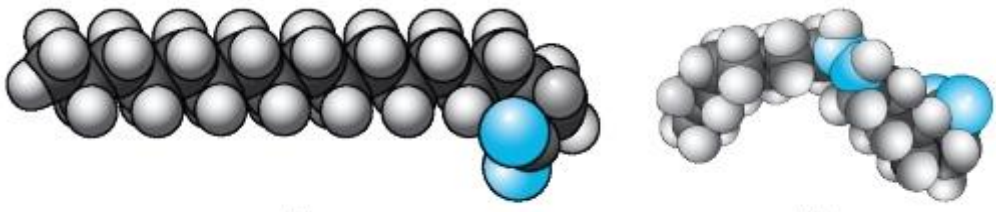


oleic acid ($C_{18:1}$)



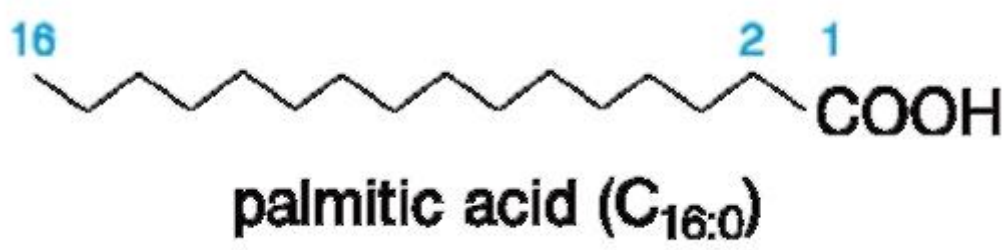
linoleic acid ($C_{18:2}$)

Saturated fatty acids and saturated portions of unsaturated acids exist in more stable staggered conformation because in this conformation the carbon atoms are as remote from one another as possible (Sec. 6.2.1). Thus, the preferred conformation of saturated acids is fully extended, whereas a carbon chain in unsaturated acids makes a bend at the position of the *cis* double bond (Fig. 12.1).



(a) (b)

Figure 12.1. Space-filling models of stearic acid (a) and oleic acid (b). The oxygen atoms are coloured, the carbons of the double bond are pale coloured in (b).
It is convenient to represent the carbon chain of fatty acids by the zigzag line in a skeletal formula that corresponds to a staggered conformation, for example:

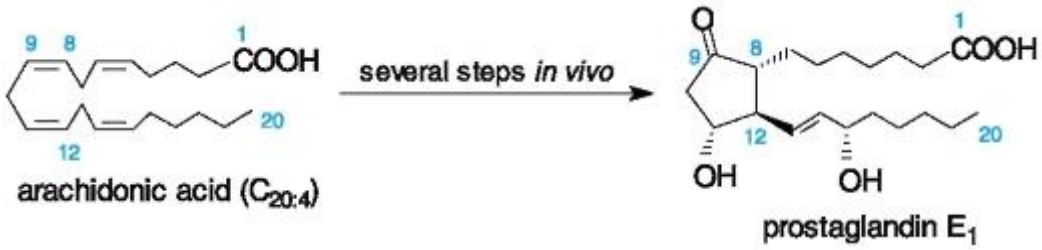


palmitic acid ($C_{16:0}$)

Certain polyunsaturated fatty acids are termed *essential* because their absence in the human diet leads to some dysfunction. These acids are also known to be precursors of *prostaglandins*, a family of physiologically potent lipid acids.

Prostaglandins. This name originates from the fact that prostaglandins were first thought to be produced by the prostate gland. It is known today that prostaglandins are widely distributed in many human tissues but in minute amounts. Most of them being in very small concentration possess a powerful physiological activity of a broad spectrum.

Prostaglandins are C_{20} carboxylic acids related to the unsaturated fatty acids. They contain a five-membered ring with two long side chains and differ from one another only in the number of hydroxyl and oxo groups present and in a degree of unsaturation. Prostaglandins are biosynthesized by oxidation and cyclization of the C_{204} fatty acid, arachidonic acid, for example:



arachidonic acid ($C_{20:4}$)

prostaglandin E_1

Problem 16.1. Write equations for reactions of oleic acid with: (a) aqueous sodium hydroxide; (b) bromine in CCl_4; (c) alkaline solution of $KMnO_4$. Name the products obtained.

12.2.2. Alcohols

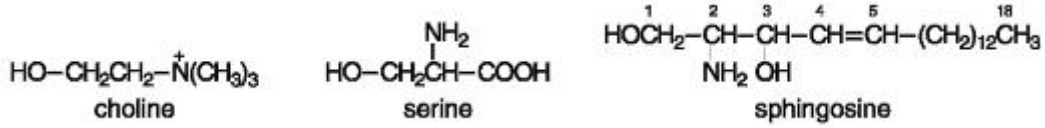
Several types of alcohols may serve as lipid components:

- monohydric long-chain alcohols;
- polyhydric alcohols;
- amino alcohols.

The first group is mainly presented in waxes that contain linear saturated alcohols with 16 and more carbons, such as cetyl alcohol (C_{16}) and melissyl alcohol (C_{30}).

The main representative of polyhydric alcohols is *glycerol,* though other polyols are sometimes encountered.

A group of amino alcohols includes *colamine* (2-aminoethanol), *choline*, an amino acid *serine,* and a long-chain unsaturated alcohol *sphingosine.*



12.3. SIMPLE LIPIDS

12.3.1. Waxes

Waxes are mixtures of esters formed by fatty acids and long-chain alcohols.

Both fatty acids and alcohols (usually monohydric alcohols) may each contain from 12 to 46 carbon atoms. Waxes are natural protective coatings on the skin, fur, or feathers of animals and on the leaves and fruits of plants. These solid substances are insoluble in water and in cold ethanol. *Beeswax* is largely myricyl palmitate, $C_{15}H_{31}COOC_{30}H_{61}$. *Spermaceti,* a flavoured substance from the head of the sperm whale, is mainly cetyl palmitate, $C_{15}H_{31}COOC_{16}H_{33}$.

It was much used in cosmetics until it was interdicted in the 1970's to save the whales population.

Waxes are generally harder and less greasy than fats. They find important applications in the production of cosmetics, ointments, and other pharmaceutical preparations.

12.3.2. Fats and Oils

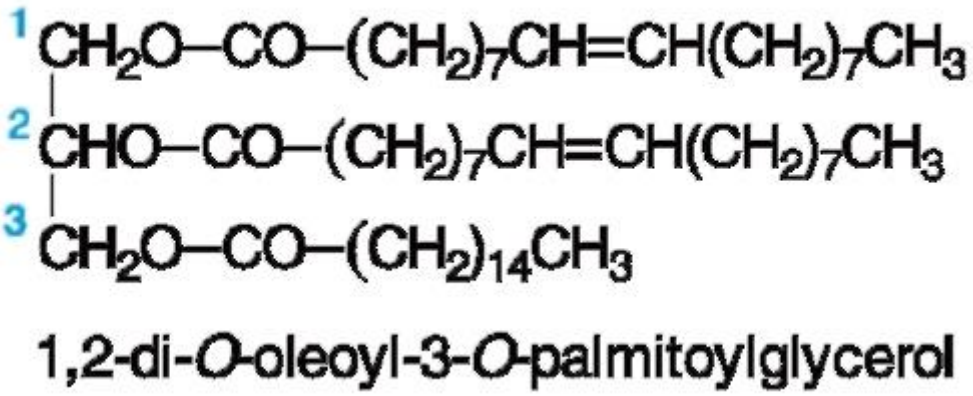
Fats and oils, along with carbohydrates and proteins, constitute a class of foodstuffs. We consume fats every day with milk, butter, animal fat, vegetable oil, and many other foods. Fats are the greatest and most concentrated storehouse of energy for animals. (Plants, on the contrary, usually store their energy in the form of starch.) Enough fat can be accumulated to last a person for months. Besides, fats and oils find application in various industries for manufacturing soap, paint, margarine, etc.

Fats and oils are mixtures of esters formed by fatty acids and glycerol.

These esters are known as *triacylglycerols;* the older term *triglycerides* is also used but it is not accepted by the IUPAC nomenclature.

Fats and oils have the same principal structure (Table 12.1) in which all three OH groups of glycerol are esterified with the same or different fatty acids. The esters that have two or three different fatty acid residues are referred to as *mixed* triacylglycerols.

Systematic names of triacylglycerols are formed from glycerol as its *O*-acyl derivatives. This is exemplified as follows:



The R radicals of fatty acids may be written in a shortened form such as $C_{15}H_{31}$ for palmitic acid, $C_{17}H_{35}$ - for stearic acid, $C_{17}H_{33}$ - for oleic acid, $C_{17}H_{31}$ - for linoleic acid, and $C_{17}H_{29}$ - for linolenic acid.

A particular fat or oil consists not of a single triacylglycerol, but of a multi-component mixture (up to 20) of glycerol esters. Therefore, the composition of a fat or oil is usually expressed as the percentage of various fatty acids obtained from it by hydrolysis. Fatty acid composition of some fats and oils is given in Table 12.3.

Table 12.3. Fatty acid composition of some fats and oils (% by mass)

Source	Saturated acids					Unsaturated acids		
	$\leq C_{10}$	C_{12}	C_{14}	C_{16}	C_{18}	$C_{18:1}$	$C_{18:2}$	$C_{18:3}$
Animal fats								
Butter	10–13	2–5	8–15	25–30	9–13	20–35	2–5	–
Lard	–	–	1–2	25–30	13–18	45–50	6–9	0–1
Beef tallow	–	–	2–5	24–34	15–25	35–45	2–5	0–1
Human fat	7	1	3	25	8	46	10	–
Vegetable oils								
Olive	–	–	1	5–10	2–4	70–84	7–17	–
Corn	–	–	1–2	7–10	3–4	30–45	41–50	–
Sunflower	–	–	1	6–9	2–5	24–40	45–70	1
Soybean	–	–	1–2	5–10	3–7	20–30	45–60	5–10
Peanut	–	–	–	6–11	25	40–60	20–35	–
Coconut	12–15	40–50	15–20	9–12	2–4	6–9	1	–

Now the question arises: Why are some triacylglycerols solids (fats) and the other are liquids (oils)? To answer this question, let us analyze the data in Tables 12.2 and 12.3. Unsaturated fatty acids have lower melting points than their saturated

counterparts do, because the double bond (or bonds) present in the unsaturated fatty acids prevents the molecules from packing together as tightly in the crystal lattice as the saturated acids do. As a result, the forces between the molecules of unsaturated fatty acids in a crystal are weaker than those between the molecules of saturated fatty acids, and this accounts for the lower melting points.

This is also true for triacylglycerols. Vegetable oils generally contain a high proportion of low-melting unsaturated acids in comparison to saturated ones than animal fats do (Table 12.3), thus oils are liquids at room temperature[1]. A high percentage of saturated fatty acids generally confers a semisolid or solid character of the fat.

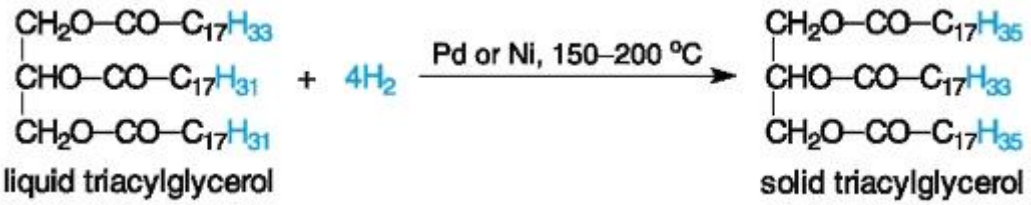
[1] Coconut oil and cocoa-bean oil, not listed in the table, are exceptions. Both of them contain a high proportion of high-melting saturated fatty acids and are therefore solids. This means that there is no strict definition for fats and oils. Fats are usually referred to as solid triacylglycerols, while oils are liquid ones. Sometimes they are classified according to their natural sources: fats are triacylglycerols of animal origin and oils are those of plants.

The degree of unsaturation in any fat or oil is conveniently expressed by *iodine value* (or iodine number). This is the mass of iodine in grams taken up by 100 g of a fat or oil. The method is based on electrophilic addition of iodine to the C=C double bond (Sec. 6.2.1).

The higher the iodine value, the greater the degree of unsaturation. Oils tend to have iodine numbers above 70, whereas those of fats are generally below 70. For example, butter has the iodine value of 25-30, olive oil - of 80-85, corn and sunflower oils - of 120-130, and linseed oil, one of the most unsaturated oils, - of 180-185.

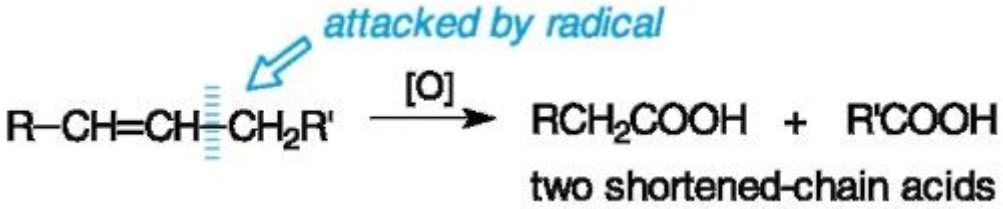
Problem 12.2. Triacylglycerols from olive oil and corn oil consist of approximately equal amounts of saturated (8-14%) and of unsaturated (86-92%) fatty acids, but their iodine values differ considerably. In the same order the freezing points of these oils decrease. Explain these facts.

Hydrogenation of oils. Because solid fats are often easier to use than liquid oils, the conversion of vegetable oils to solids by hydrogenation of some double bonds in the unsaturated fatty acid chains is done commercially on a large scale. *Margarine* and solid cooking fats are manufactured by hydrogenating vegetable oils until the proper consistency is obtained. A principal equation of the reaction may be expressed as follows:



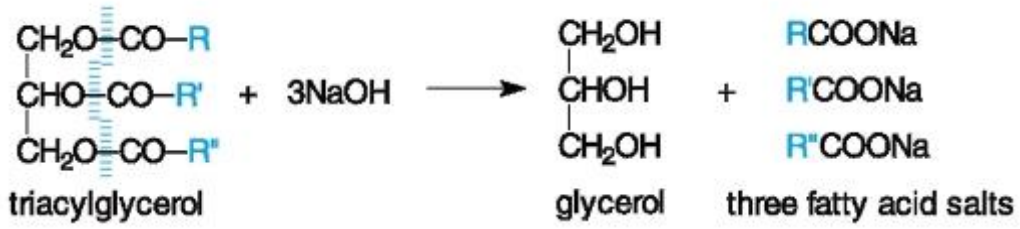
The process of hydrogenation of oils to fats is referred to as *hardening* of vegetable oils. Saturated fats have been implicated in arterial disease, and unsaturated fats have been thought to have a beneficial effect in preventing arterial deposits. The use of liquid oils in human diet has therefore increased.

Oxidation of triacylglycerols. This is one of the most important reactions of unsaturated triacylglycerols, including their transformation *in vivo*. Oxidation with molecular oxygen proceeds as a free radical process in which the radical HO' or HOO' attacks a CH_2 group neighboring to the C=C fragment. The process may be shown schematically as follows:



This reaction leads to rancidity of fats and oils because short-chain fatty acids have an unpleasant odor. In living cells the oxidation results in destroying cell membranes.

Hydrolysis of triacylglycerols. Either acid or base can catalyze hydrolysis of triacylglycerols in the same manner as with other esters (Sec. 10.3.1). When heated with aqueous alkali, fats and oils yield glycerol and fatty acids (as salts):



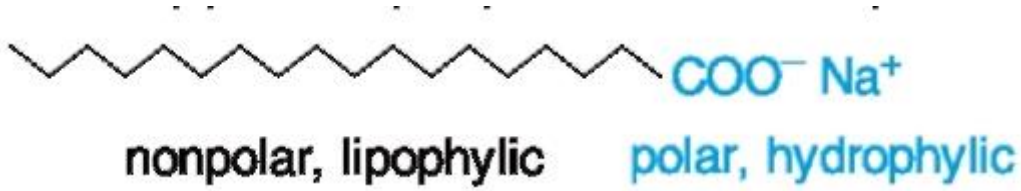
Such sodium (and potassium) salts are called *soaps*, from which the term *saponification* is derived. It is general practice to describe the alkaline hydrolysis of an ester as saponification even when the product is not soap.

12.3.3. Soaps and Detergents

One of the most ancient organic reactions is the boiling of lard with solution of caustic soda (NaOH) to produce soap. Modern commercial soaps contain perfumes and other additives to enhance their attractiveness.

Soap molecules have two dissimilar ends: a hydrocarbon chain that is non-polar and hydrophobic (repelled by water) but lipophilic (attracted to fats), and

carboxylate salt end, which has just the opposite properties and is polar and hydrophilic.



One of the undesirable properties of soaps is that they form insoluble salts with calcium, magnesium, or ferric ions present in hard water. For this reason one of the most useful innovations in cleansing has been the development of synthetic detergents. They have desirable properties of ordinary soaps but form water-soluble salts with the ions of hard water. Widely used detergents are alkylated benzenesulfonates, R-$C_6H_4SO_3Na$, and alkyl sulfates, RO-SO_3Na (R = C_{10}-C_{14} alkyl).

Problem 12.3. Complete equations for the following *possible* reactions:

(a) palmitic acid + NaCl →

(b) sodium stearate + HCl →

(c) linoleic acid + excess Br_2→

(d) potassium oleate + $MgSO_4$→

(e) stearic acid + H_2 (Pd catalyst) →

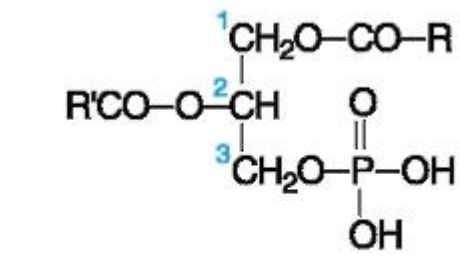
12.4. COMPLEX LIPIDS

From three main groups of complex lipids presented in Table 12.1 only two will be considered in separate sections since the third group - sphingolipids - may relate to either phospholipid or glycolipid group.

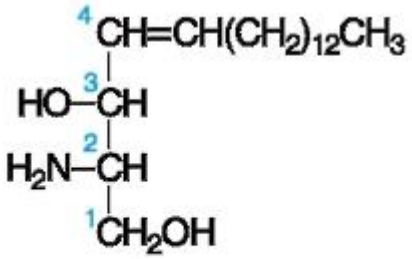
12.4.1. Phospholipids

Phospholipids form a class of great biochemical importance; they constitute about 40% of cell membranes (the remaining being proteins). There are two main types of phospholipids: *phosphatides* (or glycerophospholipids) and *phosphosphingolipids.*

Phosphatides are derivatives of *phosphatidic acid* which is related structurally to triacylglycerols, except that one of the three acyl groups is replaced by phosphoric acid residue. Phosphosphingplipids are phosphorylated and *N*-acylated derivatives of sphingosine. Note that the formulas below indicate the stereochemistry of the chiral centres, i. e. the *R* configuration at C-2 in phosphatidic acid and the 2*S*,3*R*configuration in sphingosine.

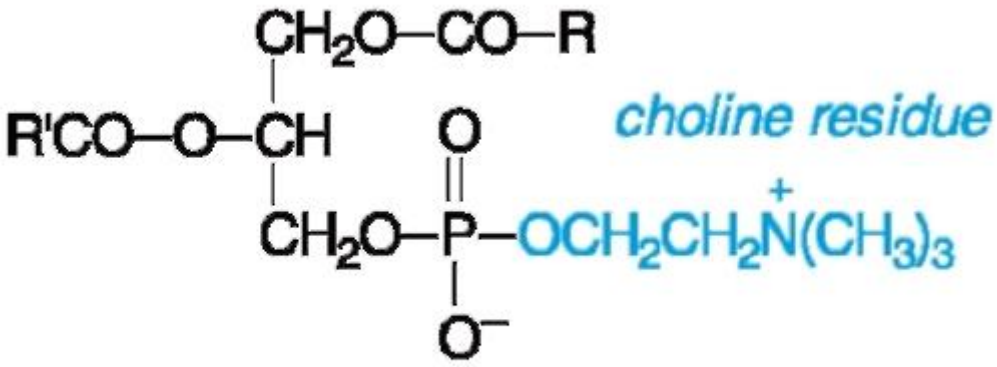


phosphatidic acid
(R and R' are fatty acid radicals)



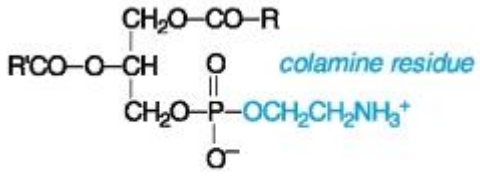
sphingosine

In phosphatides, the phosphate group is also linked by a separate ester bond to an amino alcohol.*Lecithins*, or phosphatidylcholines, are the most important representatives of phospholipids of higher animals and plants. They constitute up to 50% of total phospholipids.

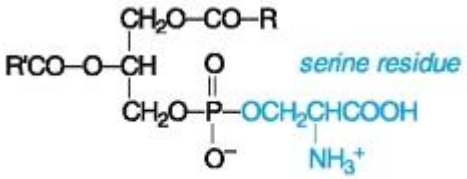


lecithin

Cephalins (from the Greek *kephalikos* - head) are isolated from the brain and spinal tissues. They form another important group of phospholipids in which nitrogenous component is colamine or serine.

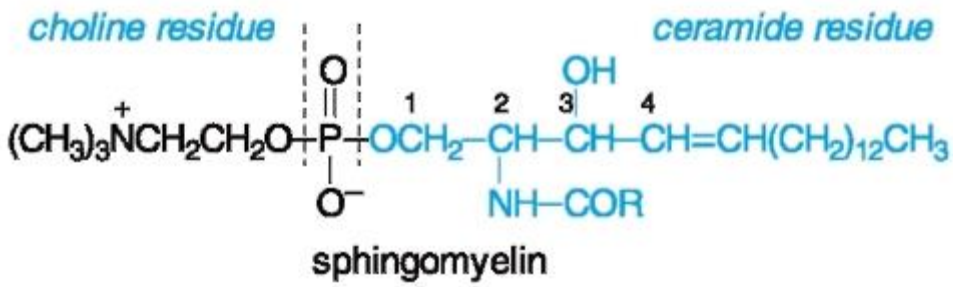


phosphatidylcolamine

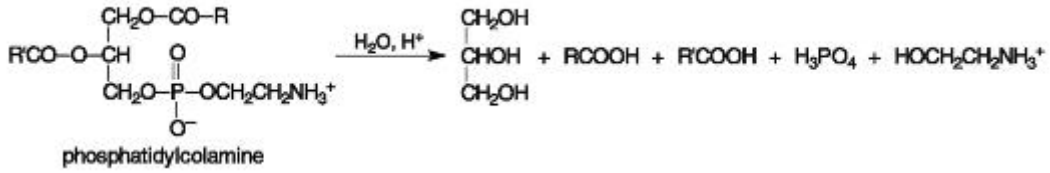


phosphatidylserine

Sphingomyelin, first found in the nervous tissues, is another example of cell membrane phospholipids. Its molecule consists of three parts: choline, phosphate, and *ceramide;* the latter is sphingosine acylated at the amino group with long-chain acids (mainly C240 and $C_{24:1}$).



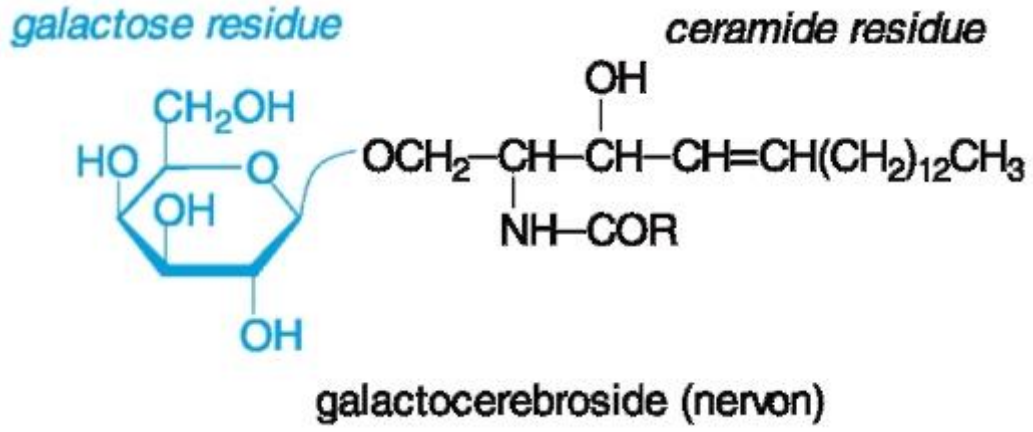
Like esters, phospholipids can be hydrolyzed under acidic or alkaline conditions. For example, complete hydrolysis of phosphatides releases fatty acids, glycerol, phosphoric acid, and choline, colamine, or serine, respectively.



Problem 16.4. Write the structures of the products that would be obtained by saponification of lecithin containing oleic and stearic acid residues.

12.4.2. Glycolipids

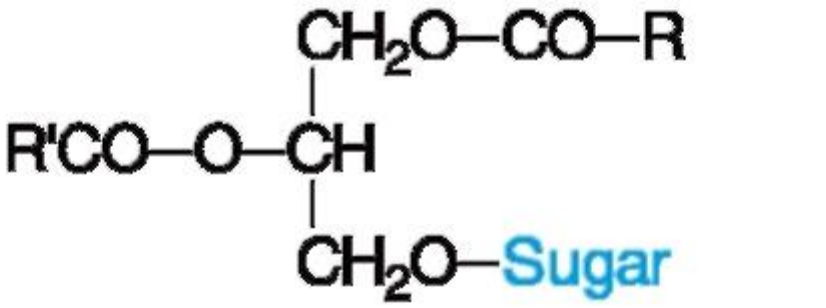
The most important glycolipids are *glycosphingolipids*, compounds in which monoor oligosaccharide is linked by a glycosidic bond to the terminal hydroxyl group of the ceramide molecule. Monoand oligosaccharide derivatives are called *cerebrosides* and *gangliosides*, respectively. The sugar in cerebrosides is mainly D-galactose in the β anomeric form as shown below. An oligosaccharide chain in gangliosides may contain 4 to 10 various monosaccharide units.



Both types of glycosphingolipids were first isolated from the brain and it is reflected in their names. Then they were found in other tissues.

Glycolipids of the glycerol type are also known but they are not widespread. They are present in microorganisms, higher plants, and nerve tissues of mammals. Structurally, they are glycosylated 1,2-di-*O*-acylglycerols, as shown in the general formula where a sugar unit may be D-glucose, D-galactose, an amino sugar, or a

disaccharide:



1,2-di-*O*-acyl-3-*O*-glycosylglycerol

A number of important functions have been found for complex lipids. Like soaps, phospholipids and glycolipids have a long, non-polar hydrocarbon tail bound to a polar head such as a phosphate group or sugar units. Cell membranes are composed mostly of phospholipids arranged in a *lipid bilayer* with polar heads exposed outside, as shown in Fig. 12.2. Proteins are incorporated in the bilayer.

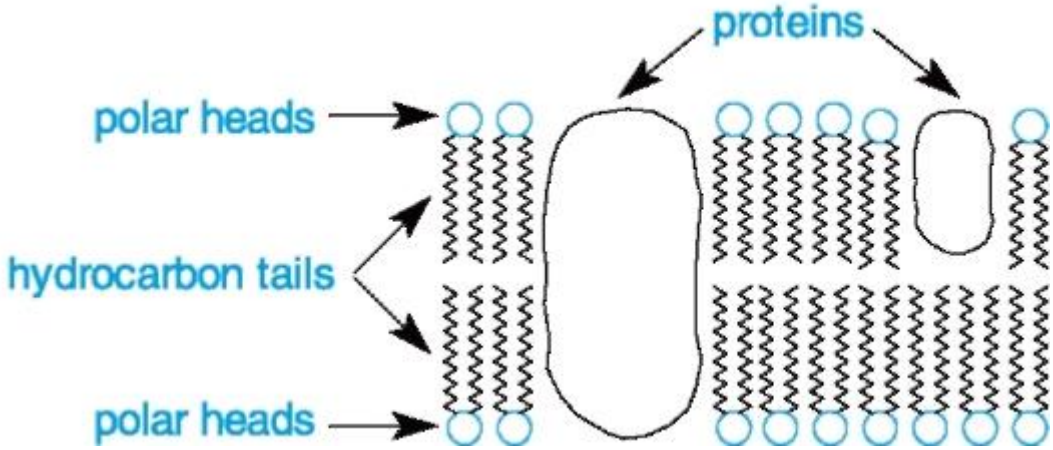


Figure 12.2. Schematic diagram of the lipid bilayer in the cell membrane.

Complex lipids play special roles in secretory processes, in ion transport, and in selective permeability of cell membranes.

Additional Problems

12.5. Classify each of the following compounds as a wax, or a fat, or an oil, *if any:*

(a) hexadecyl myristate (for myristic acid see Table 12.2);

(b) tri-*O*-palmitoylglycerol;

(c) 1,2-di-*O*-oleoyl-3-*O*-stearoylglycerol;

(d) $^{CH}3^{(CH}2^{)}16^{COOC}2^{H}5$;

(e) $^{C}25^{H}51^{COOC}30^{H}61$.

Draw the structural formulas for the compounds (a), (b), and (c).

12.6. Margarinic acid, $CH_3(CH_2)_{15}COOH$, wtas isolated from natural sources. Could you find some logical connection between this name and the name of butterlike product margarine?

Write an equation for principal preparation of margarine from a vegetable oil.

12.7. Punicic acid, a 9Z,11Z,13E-isomer of 9,11,13-hexadecatrienoic acid, is a constituent of pomegranate seeds. Draw its conformational formula. Indicate differences in the structures of punicic and polyunsaturated linoleic and linolenic acids.

12.8. How could you distinguish between stearic, oleic, and linoleic acids. *(Hint:* Not only qualitative analytical methods may be used.)

12.9. Which of the following compounds can be hydrolyzed in an alkaline medium and which of them forms glycerol on hydrolysis?

(a) $CH_2OC_{17}H_{35}$ | $CHOC_{17}H_{33}$ | $CH_2OC_{15}H_{31}$

(b) $CH_2OCOC_{15}H_{29}$ | $CHOCOC_{17}H_{33}$ | $CH_2OCOC_{17}H_{35}$

(c) $CH_2OCOC_{17}H_{35}$ | $CHOCOC_{15}H_{29}$ | $CH_2OC_{15}H_{31}$

(d) CH_2OH | $CHOH$ | $CH_2OPO_3H_2$

12.10. Show, by equations, how a wax might be converted into a synthetic detergent.

12.11. Design a chemical method to distinguish between beeswax and a solid triacylglycerol. (Hint: Use sodium hydroxide as a reagent.)

12.12. Using general formulas, show hydrophilic and hydrophobic portions, if any, in the following compounds:

(a) a liquid triacylglycerol;

(b) a phospholipid;

(c) a soap;

(d) a synthetic detergent.

12.13. Show hydrophobic and hydrophilic portions in the phosphatidylcolamine molecule. Write equations for its hydrolysis in:

(a) alkaline medium;

(b) acidic medium.

Chapter 13. TERPENOIDS AND STEROIDS

A large amount of naturally occurring substances belongs to *low-molecular bioregulators* that possess significant biological and pharmacological properties. These are alkaloids, antibiotics, flavonoids, steroids, terpenoids, vitamins and others. Some of them were or will be briefly mentioned in this book. This chapter deals with a group of low-molecular bioregulators called *isoprenoids.*

Isoprenoids are compounds whose framework consists of five-carbon isoprene  units.

A group of isoprenoids includes, first of all, terpenoids and steroids. The former are compounds of plant origin; the latter are mainly present in animals. Terpenoids and steroids differ considerably from one another in their carbon skeleton. Nevertheless, both the classes are biosynthesized from the same primary reactant acetyl coenzyme A.

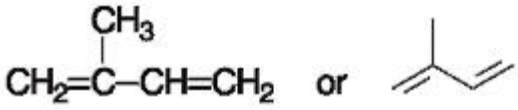
13.1. TERPENOIDS

Many terpenoids are compounds isolated from essential oils of plants by distilling the plant with water or by extraction. These oils are often responsible for a fragrant odour of the particular plant such as rose, geranium, peppermint, bay, citrus plants, conifers, and others. For centuries, the essential oils have been used as perfumes, spices, and medicines.

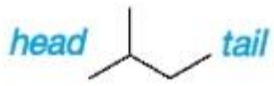
Chemically, terpenoids are relatively small molecules that have an enormous diversity of structure. A carbon skeleton of terpenes can be acyclic or, more often, cyclic. Some terpenoids are hydrocarbons (called simply - *terpenes*, but very often both terms are equivalent), but many contain oxygen, mostly as a constituent of a hydroxyl or carbonyl group.

The main structural peculiarity of terpenes is that they constitute of 10, 15, 20, 30, or 40 carbon atoms, i. e. a number divisible by five. Moreover, according to the *isoprene rule* proposed by the Swiss chemist L.S. Ruzicka (in the 1920's), terpene skeletons are built up from «head-to-tail» joining[1] of simple five-carbon blocks (isoprene units). The C-1 atom of isoprene is called the «head» of an isoprene unit, and C-4 is called the «tail».

[1] There are exceptions from the type of joining, especially for cyclic terpenes



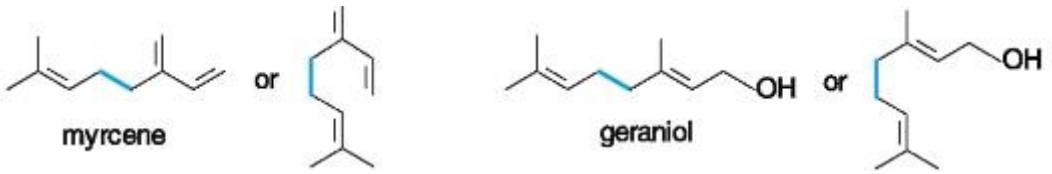
isoprene



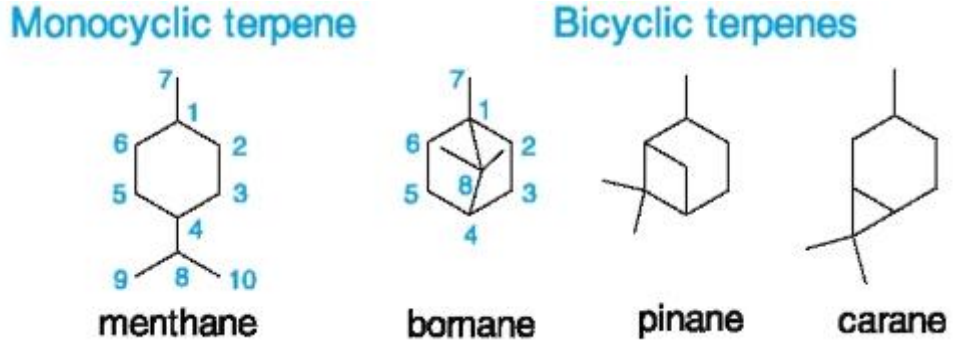
isoprene unit

Terpenes are classified according to the number of isoprene units present. Thus, monoterpenes are C_{10}compounds (two isoprene units); di-, tri-, and tetraterpenes contain 20, 30, and 40 carbons, respectively. We have already seen examples of diterpenes retinol and retinal, and the tetraterpene β-carotene (Sec. 3.1).

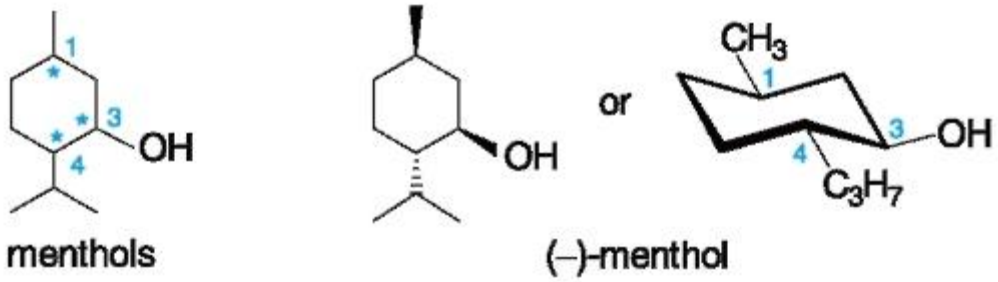
Acyclic monoterpenes *myrcene* (isolated from oil of bay leaves) and *geraniol* (present in roses and other flowers) exemplify the isoprene rule. Each molecule contains two isoprene units linked head-to-tail; the joint positions are given below in colour:



Most terpenes are derivatives of C_{10} monocyclic or bicyclic hydrocarbons whose names are adopted by the IUPAC nomenclature as parent names with a specific numbering. Menthane (more correctly p-menthane) relates to cyclohexane, whereas bicyclic hydrocarbons involve additional rings of a lesser size: cyclopentane ring - in bornane (the old name *camphane)*, cyclobutane ring - in pinane, and cyclopropane ring - in carane.

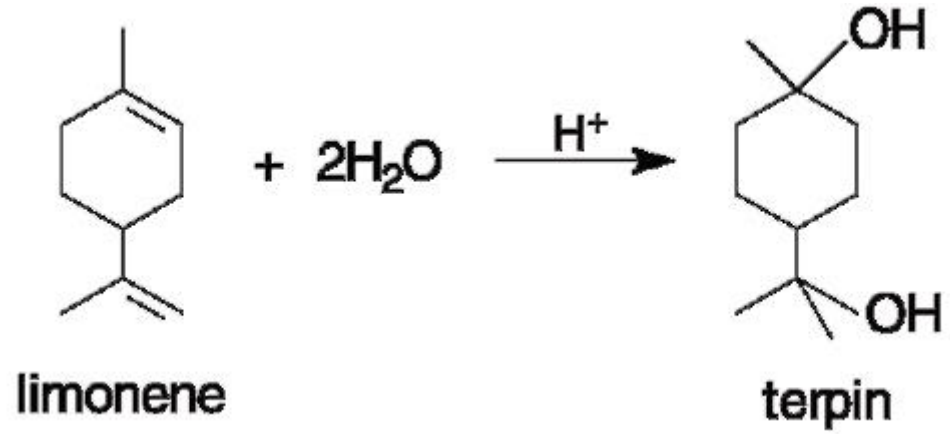


(-)-Menthol is the most known menthane derivative. It is the main terpene compound of peppermint oil and is used in medicine and especially as a flavouring and perfumery agent.

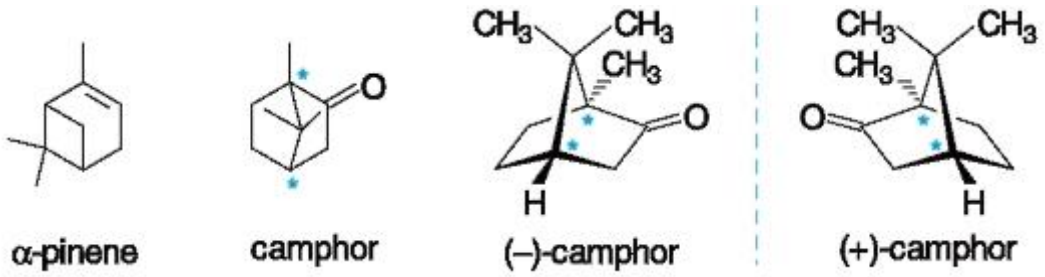


As a matter of fact, there are eight (2^3) stereoisomers of menthol since the molecule contains three chiral centres (C-1, C-3, and C-4). Only one pair of enantiomers is properly called *menthol* whose levorotatory isomer is shown below. Three other enantiomeric pairs have their own names - isomenthol, neomenthol, and neoisomenthol. Note that all substituents in the cyclohexane ring of the menthol molecule occupy the most preferable equatorial positions.

Limonene, a component of lemon and orange oils, is another representative of the monocyclic terpenes. Its hydration according to the Markovnikov's rule results in the formation of the diol *terpin*. In the form of a hydrate it is used in medicine as an expectorant.

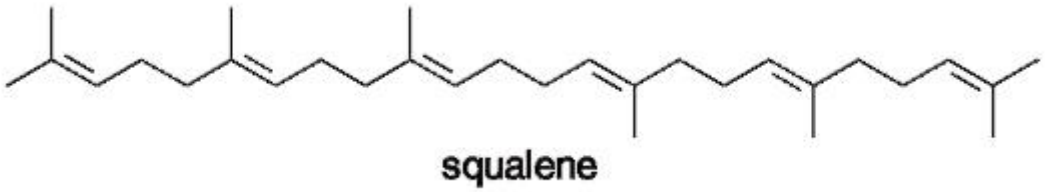


Bicyclic monoterpenes are exemplified by a hydrocarbon α-pinene, the main component of turpentine, and the ketone *camphor*, whose (+)-isomer was isolated from camphor-tree many years ago. The camphor molecule exists only as two enantiomeric forms (see below) though it has two chiral centres (C-1 and C-4). This is a consequence of rigidity of the bicyclic structure in which the carbon bridge $>C(CH_3)_2$ has always the *cis* arrangement. It should be stressed that the camphor molecule has a less stable *boat* conformation (an alternative *chair* conformation is impossible for the molecule).



Camphor is used for a long time in medicine as a cardiac stimulant. Racemic camphor is manufactured at present from available α-pinene.

Acyclic triterpene hydrocarbon *squalene* (from the Latin *squalus* - shark), first isolated from the shark liver, is the direct precursor of cholesterol and other steroids.



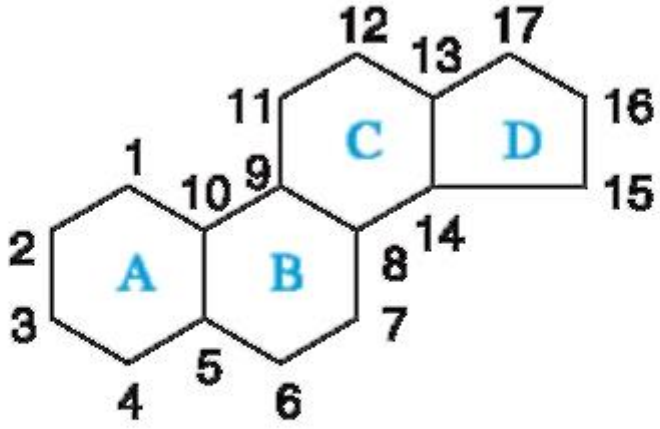
Problem 13.1. Show the isoprene units in the following terpenes: (a) geranial (3,7dimethyl-2,6-octadienal), a component of lemon oil; (b) pinane. Are stereoisomers of geranial possible? How many chiral centres could you find in the pinane molecule?

13.2. STEROIDS

Steroids are biologically active derivatives of a tetracyclic hydrocarbon that occur in all living cells. The lipid material from tissue that is not saponifiable by alkaline hydrolysis contains steroids. Thousands of steroids are known today, hundreds of them being made synthetically and are used in medical practice.

13.1.1. Structure of Steroids

Parent hydrocarbons. A principal hydrocarbon compound for all steroids is a fused tetracyclic system named in the IUPAC nomenclature gonane. The rings in its structure are marked for convenience by the letters A, B, C, and D.
The tetracyclic skeleton of steroid molecules usually has an additional side hydrocarbon chain (or chains) at the atoms C-10, C-13, and C-17, shown in Table 13.1.



gonane structure and numbering

Each parent hydrocarbon shown in the table represents a group of steroids possessing a definite physiological function. For example, estrane derivatives (estrogens) are female sex hormones,androstane derivatives (androgens) are male sex hormones, pregnane derivatives (corticoids) represent hormones of the adrenal cortex, cholane derivatives are the bile acids, and, finally, cholestanederivatives are the oldest steroids called sterols. All these groups will be considered in this chapter.

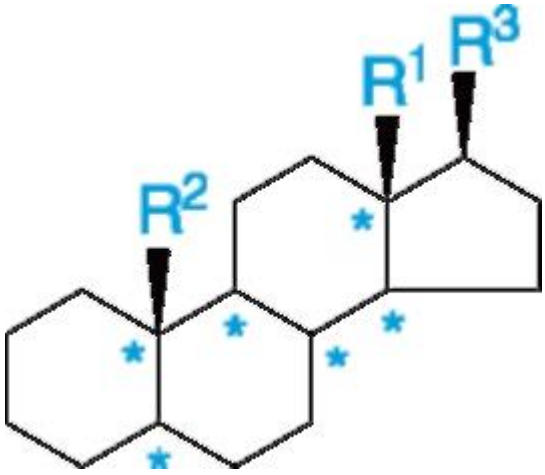
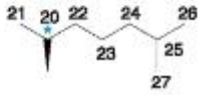
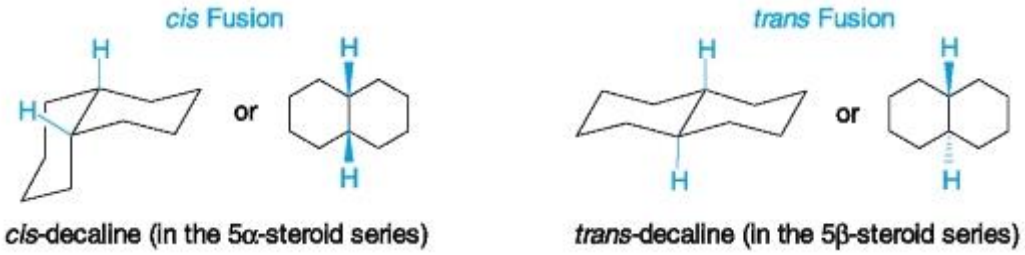


Table 13.1. Parent structures of steroids

R^1 (C-18)	R^2 (C-19)	R^3	Parent name	Biological function of derivatives
H	H	H	Gonane	No
CH_3	H	H	Estrane	Estrogens
CH_3	CH_3	H	Androstane	Androgens
CH_3	CH_3	$^{20}CH_2{}^{21}CH_3$	Pregnane	Corticoids, pregnancy hormones
CH_3	CH_3		Cholane	Bile acids
CH_3	CH_3		Cholestane	Sterols

Conformations and stereoisomerism. The number of stereoisomers in steroids is too great since six asymmetric carbon atoms are present in the gonane skeleton (they have been marked with the asterisks in the general formula).

Even two cyclohexane rings in chair conformation can be joined in either a *cis* or a *trans* manner. In c/s-decaline, both bonds to the hydrogen atoms at the ring junction positions are on the same side. In trans-decaline, on the contrary, both the C-H bonds lie on the opposite sides of the molecule.

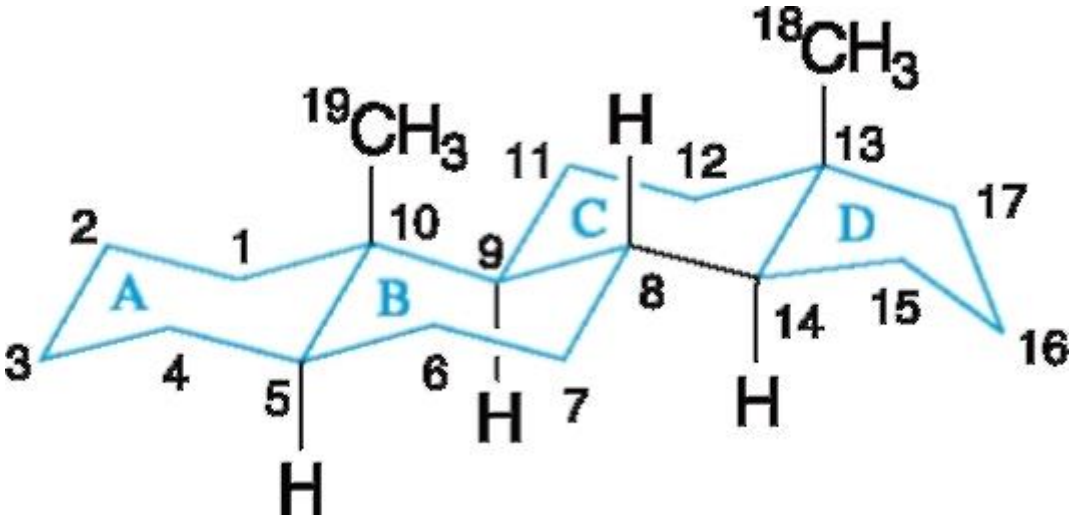


cis-decaline (in the 5α-steroid series) *trans*-decaline (in the 5β-steroid series)

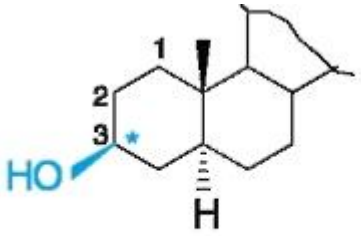
The stereochemistry in planar formulas is indicated by wedged lines (β-bonds, coming out of the plane of the paper) and by hashed wedges (α-bonds, going behind the plane of the paper), as shown above. Two types of the ring fusion increase the complexity of the stereochemistry.

Steroids can have either a *cis* or *trans* fusion of the A and B rings, but the other ring fusions (B-C and C-D) are usually *trans* ones unless otherwise stated. The A-B*trans* fused steroids have the C-19 angular methyl group «up» (denoted β), and the hydrogen atom at C-5 «down» (denoted α) giving the 5a-series.
Example 13.1. Draw the conformational formula for 5a-androstane.

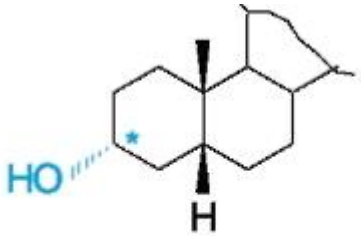
Solution. The rings A and B must be *trans* fused in the 5a-steroids, just as other rings.



Additional substituent at any carbon of the cyclic skeleton bears a pair of stereoisomers that relate to each other as diastereomers. They differ in configuration of a new chiral centre, which is the C-3 atom in the following example:



3β-alcohol of 5α-steroid



3α-alcohol of 5β-steroid

It should be mentioned here that the symbols of CH_3 groups are usually omitted in steroid structural formulas but the symbols of hydrogen are used to show a type of the ring fusion.

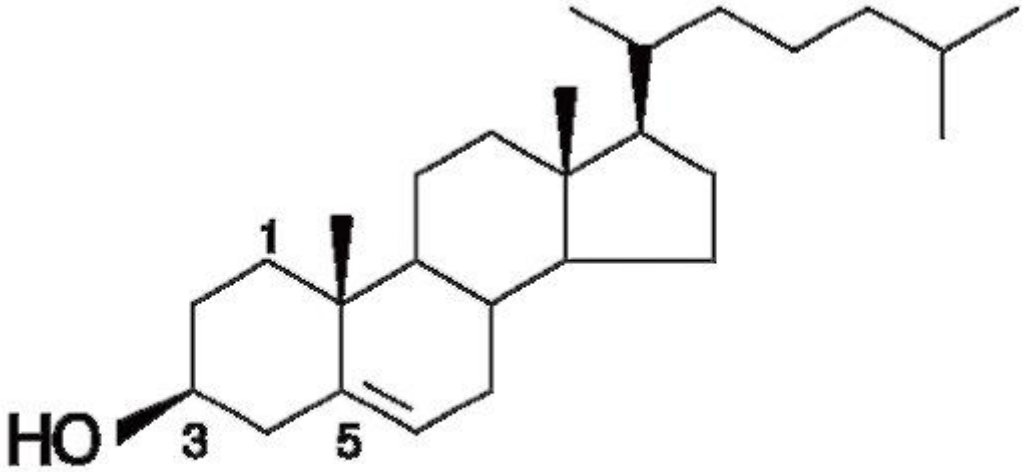
Finally, in any rigid cyclic system like gonane and its functional derivatives, stereochemical effects are marked, and often completely control the course of reaction, especially in biological processes.

Problem 13..2. Draw the conformational formulas for stereoisomers of estrane having: (a) all rings in*trans* fusion; (b) trans-fused B-C and C-D rings but *cis* fused A-B rings.

13.2.2. Steroid Groups

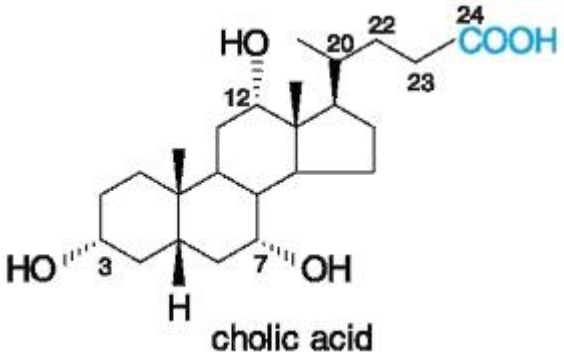
Several groups of steroids will now be considered.

Sterols. To this group belong steroid alcohols with the OH group at the position 3β. The parent hydrocarbon structure of sterols is cholestane. Sterol molecules often contain a C=C double bond. The most widespread sterol is cholesterol. Its systematic name is 5-cholesten^-ol. Here, general principles of the substitutive IUPAC nomenclature are used: the suffix -en(e) at C-5 indicates unsaturation in the parent structure, the OH group at C-3 is shown by the suffix -ol with the locant 3β, the letter β indicates the stereochemistry at C-3.



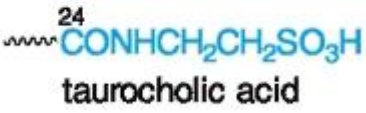
cholesterol

Bile acids. These are carboxylic derivatives of the hydrocarbon cholane. Salts of the bile acids are natural emulsifying agents found in the bile, a digestive fluid formed by the liver. Bile is stored in the gall bladder and released at intervals to assist in the digestion and adsorption of fats. The most important of bile acids is cholic acid. Note that cholic acid is a derivative of 5β-cholane (cis fusion of the A and B rings).



cholic acid

glycocholic acid

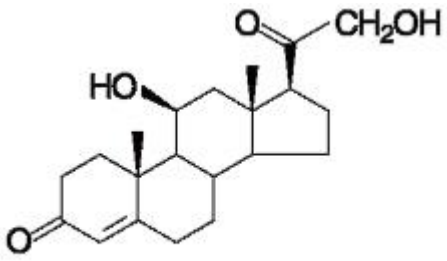


taurocholic acid

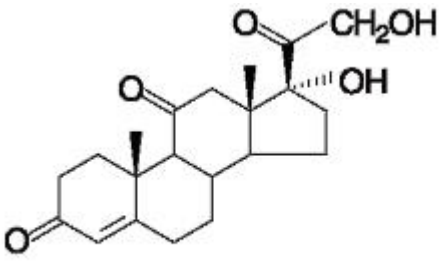
Cholic acid is usually combined with the amino acid glycine, H_2NCH_2COOH, or the amino sulfonic acid taurine, $H_2NCH_2CH_2SO_3H$, by an amide linkage to form glycocholic acid and taurocholic acid,respectively (see above). The sodium salts of the bile acids have an extensive hydrocarbon region (that is hydrophobic) and a highly polar region (hydrophilic) and function in the intestinal tract as emulsifying agents. Thus, they are biological «soaps» in a sense.

Steroid hormones. Hormones are compounds secreted by various endocrine glands in trace amounts. They function as chemical messengers and regulate a variety of physiological and metabolic activities in vertebrates. Many hormones have a steroid structure, including adrenal corticoids and sex hormones.

Hormones of the adrenal cortex. If the adrenal gland exhibits decreased function, carbohydrate and protein metabolism is unfavourably affected, electrolyte and water balance is abnormal, and a patient is more sensitive to cold and stress. Corticosterone and cortisone exemplify this group of hormones:

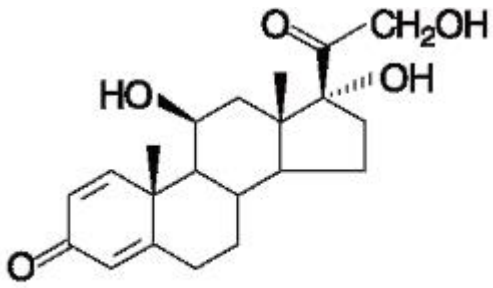


corticosterone

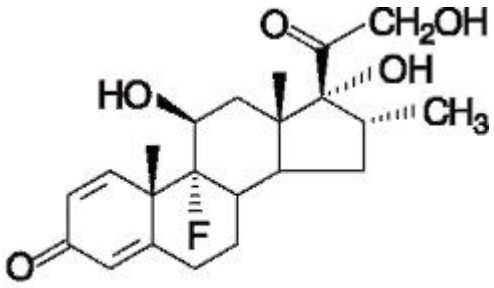


cortisone

Pharmaceutical industry prepared and tried many closely related compounds to increase the potency and decrease the side effects of some steroid drugs. Examples are Prednisolone and Dexamethasone:



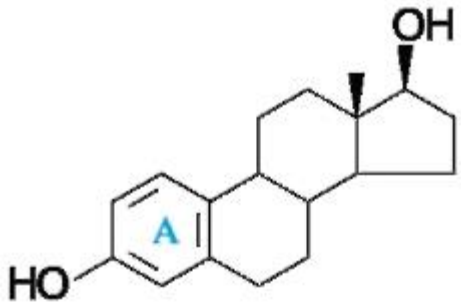
Prednisolone



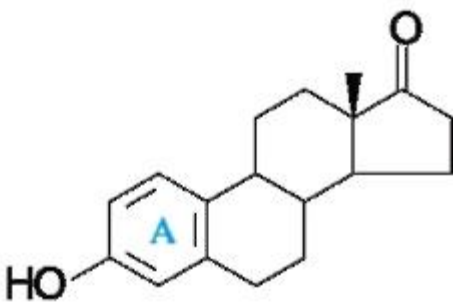
Dexamethasone

Problem 13.3. The synthetic drug Deltacortril shows pharmacological properties similar to that of prednisolone. Its systematic name is 11β,17,21trihydroxypregna-1,4-diene-3,20-dione. Are these compounds identical or not?

Sex hormones. The major estrogens, or female sex hormones,
are esradiol and estrone. They are derivatives of the hydrocarbon estrane.

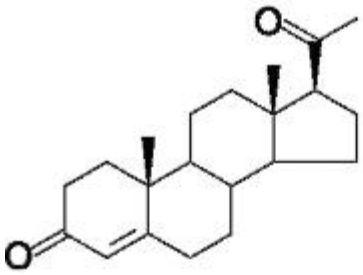


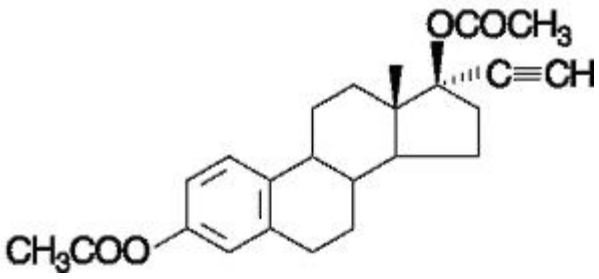
estradiol



estrone

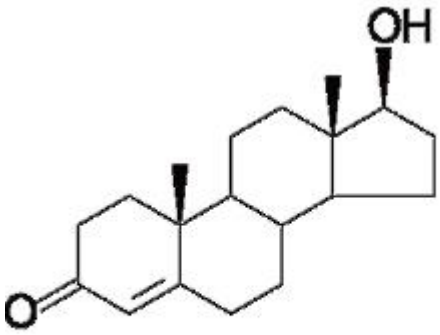
Note that the ring A in estrogen molecules is the aromatic one. This results in acidic character of the OH group, which is a phenolic hydroxyl.
The role of steroids in pregnancy had led to research into the application of these compounds as birth control agents. Initially, progesterone (the most important pregnancy hormone) itself was studied in this regard. It was found, however, that the dose preventing ovulation was much too large. One successful in this regard synthetic steroid is ethynyl estradiol diacetate.



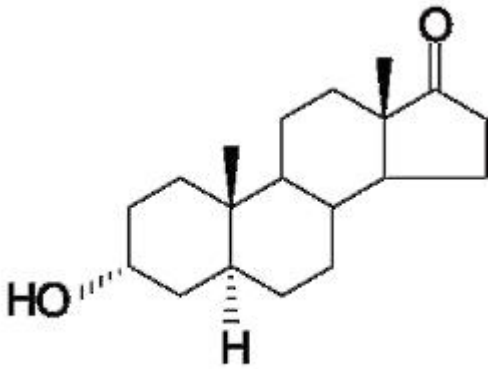


progesterone ethynyl estradiol diacetate

The major androgens, or male sex hormones, are testosterone and androsterone, the derivatives of androstane.



testosterone

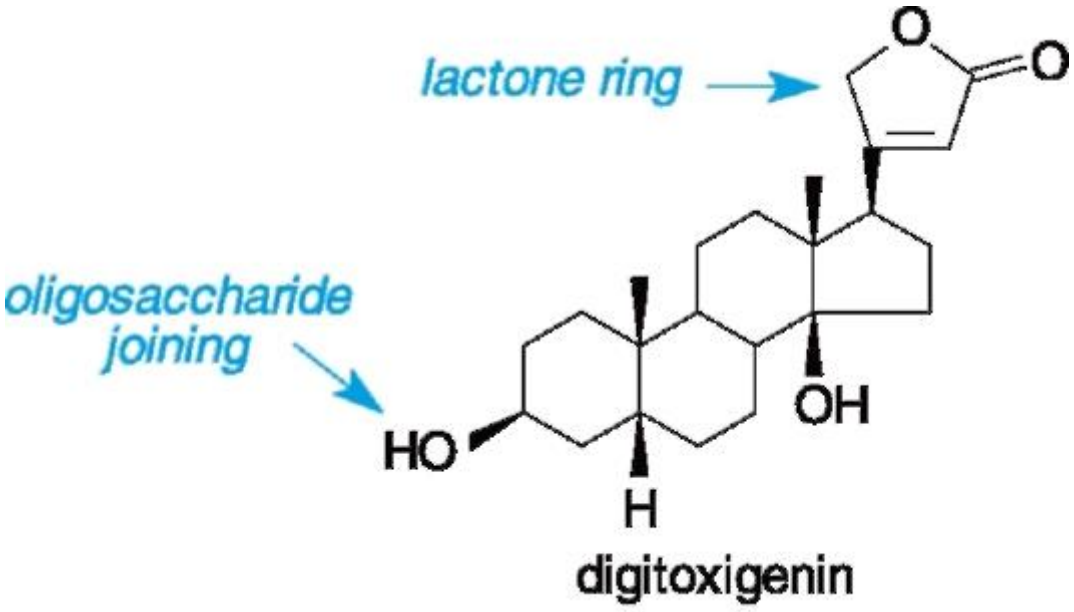


androsterone

The main function of the androgens in man is the development of masculine sexual characteristics, such as deepening of the voice, the growth of a beard, and others. They also control the function of the glands of reproduction.

Testosterone also belongs to a group of anabolic steroids, i. e muscle-building agents. Drugs of this type are sometimes administerted for maintaining body mass in recovering people. These drugs, however, are sometimes taken by healthy sportsmen (both males and females) to increase muscle mass. If used in high doses, anabolics have many side effects, including sexual disturbances.

Cardiac glycosides. Steroids also occur in the plant kingdom. One example is digitalis, a preparation made from the dried seeds and leaves of the purple foxglove. Historically, digitalis was used as a poison and as a medicine in heart therapy. The active agents of digitalis are cardiac glycosides, which are complex molecules built up from a steroid as an aglycone portion and several carbohydrates in the form of an oligosaccharide. Hydrolysis of digitoxin, one of cardiac glycosides, yields digitoxigenin.

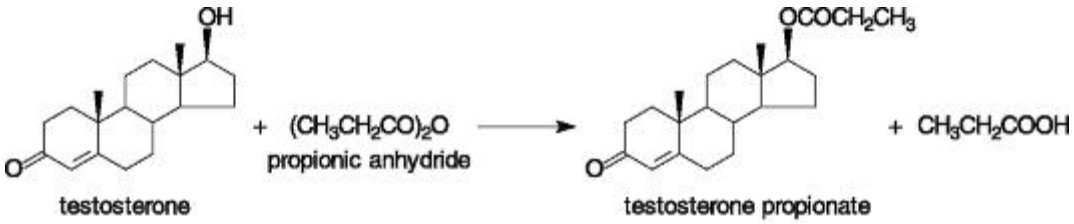


digitoxigenin

The main peculiarity of digitoxigenin and other cardiac glycosides is the presence of a lactone ring joined to the position C-17. It means that these substances are sensitive to both acids and alkalis, as it is typical for lactones. A trisaccharide unit presents carbohydrate portion of digitoxin. It has an influence on the rate and duration of cardiotonic action of the glycoside.

13.2.3. General Chemical Characteristic of Steroids

Chemical behaviour of steroids is conditioned by functional groups presented in a molecule. Steroid molecules often contain a hydroxyl group (or groups) therefore they can be oxidized or converted into esters or ethers (Sec. 7.2). Some drugs are employed as esters, for example, ethynyl estradiol diacetate, mentioned above, as well as testosterone propionate that can be prepared by a familiar acylation reaction:



testosterone testosterone propionate

Many steroids contain a carbonyl group, which can be reduced into a hydroxyl group or involved in the reaction with nucleophilic reagents such as alcohols, amines, etc. (Sec. 8.2 and 8.4). Unsaturated steroids can also participate in addition reaction such as hydrogenation, oxidation, and so on (Sec. 6.2.1). Reactivity of the phenolic hydroxyl in estrogens and of the lactone ring in cardiac glycosides has already been discussed in this chapter.

Problem 13.4. Show the products you would expect to obtain from reaction of estradiol with the following reagents: (a) excess acetyl chloride; (b) excess sodium hydroxide; (c) bromine water.

Additional Problems

13.5. Show the isoprene units in the following terpenes:

(a) borneol (3-bornanol), a component of various essential oils;

(b) retinol, or vitamin A (for the structure see Sec. 3.1).

How many chiral centres does borneol have? Are stereoisomers of retinol possible?

13.6. How could you distinguish by chemical methods between the compounds in the following pairs:

(a) retinol and retinal (for the structures see Sec. 3.1);

(b) pinane and α-pinene. Illustrate your answer by equations.

13.7. How could you distinguish between the following compounds:

(a) natural (-)-menthol;

(b) menthone, an oxidation product of (-)-menthol;

(c) menthane.

Hint: Draw the structure for the compound (b). To recognize menthane you do not need chemical methods.

13.8. The systematic name of lithocholic acid (found in human bile) is 3α-hydroxy5β-cholan-24-oic acid. Draw the structural formula for it. Indicate the types of ring fusion in the molecule. Is the hydroxyl group axial or equatorial?

13.9. The main metabolite of cholesterol is cholestanol named systematically 5acholestan-3β-ol. Draw its structural formula. How could you distinguish by chemical methods between cholesterol and cholestanol?

13.10. Draw the structures of products you would expect to obtain, *if any,* from reaction of estrone with the following reagents:

(a) dilute sodium hydroxide;

(b) dilute hydrochloric acid;

(c) excess methanol in the presence of an acid;

(d) phenylhydrazine.

13.11. Dehydrocholic acid (3,7,12-trioxo-5β-cholan-24-oic acid) can be obtained by oxidation of cholic acid. What product (or products) would you expect from reduction of the oxo groups in dehydrocholic acid? Illustrate your answer by equations.

PART 4. BIOPOLYMERS AND THEIR STRUCTURAL CONSTITUENTS.

Chapter 14. CARBOHYDRATES

Carbohydrates are among the most abundant substances on the Earth and perform many vital functions in both plants and animals. The most common of these are cellulose, starch, and various sugars. Cellulose is the main structural component of

plants, used to construct rigid cell walls, fibres, and woody tissue. Starch is the chief form for storing carbohydrates for later use as a food or energy source. In higher animals, the sugar glucose is an essential component of blood. Ribose and 2-deoxyribose are constitiuents of genetic material. Other carbohydrates are important components of coenzymes, antibiotics, cartilage, the shells of crustaceans, and bacteria cell walls. For the last decades a definite role played by carbohydrates in the processes of cell recognition and immunity was established.

Carbohydrates being the initial products of photosynthesis (from carbon dioxide and water) represent a certain «bridge» between inorganic and organic substances.

The term *carbohydrates,* which has been used since the mid-nineteen century, derives historically from the observation that many compounds in this class (for example ribose, $C_5H_{10}O_5$, glucose, $C_6H_{12}O_6$, and sucrose, $C_{12}H_{22}O_{11}$) have the empirical formula $C_x(H_2O)_y$ and are formally regarded as «hydrates of carbon». However many important compounds related to carbohydrates, such as deoxyribose, $C_5H_{10}O_4$, uronic acids, and amino sugars (see further in this chapter), are not covered by this definition. Nevertheless, this term is steadily employed along with the equivalent names *saccharides* or simply*sugars.*

It is difficult enough to give a clear definition of carbohydrates, because they include very different types of compounds, from small molecules with several carbon atoms to polymers whose molecular mass amounts to millions. Sometimes carbohydrates are defined as «polyhydroxy aldehydes, polyhydroxy ketones, or substances that give such compounds on hydrolysis». But in this book we have never used chemical properties in the definition of classes. It would be better therefore to divide carbohydrates into three big groups (or sub-classes), namely, monosaccharides, oligosaccharides, and polysaccharides, and to study them separately.

14.1. MONOSACCHARIDES

Monosaccharides are usually crystalline sweet substances. They are readily soluble in water and insoluble in ethanol, ether, and other solvents of low polarity. The most common definition of monosaccharides is the following:

Monosaccharides are polyhydroxy carbonyl compounds, namely,

polyhydroxy aldehydes or polyhydroxy ketones.

Although this definition focuses on the main functional groups of monosaccharides, it is not entirely satisfactory too. We will soon see that the carbonyl group is virtually absent in monosaccharides.

14.1.1. Classification, Stereoisomerism, and Nomenclature

Monosaccharides can be classified as *aldoses* or *ketoses* depending on whether they contain an aldehyde or ketone group. They can also be classified according to the number of carbon atoms they contain. A*pentose,* for example, contains five

carbons, a *hexose* is a six-carbon monosaccharide. These are the most widespread monosaccharides though trioses (C_3), tetroses (C_4), heptoses (C_7) etc. also exist. A combination of the two classifications gives rise, for example, to aldopentoses, ketohexoses, and so on (Fig. 14.1). Thus ribose (Fig. 14.1, a) is an aldopentose, i. e. it has five carbons and an aldehyde group, and fructose (Fig. 14.1, d) is a ketohexose (six carbons and a ketone group).

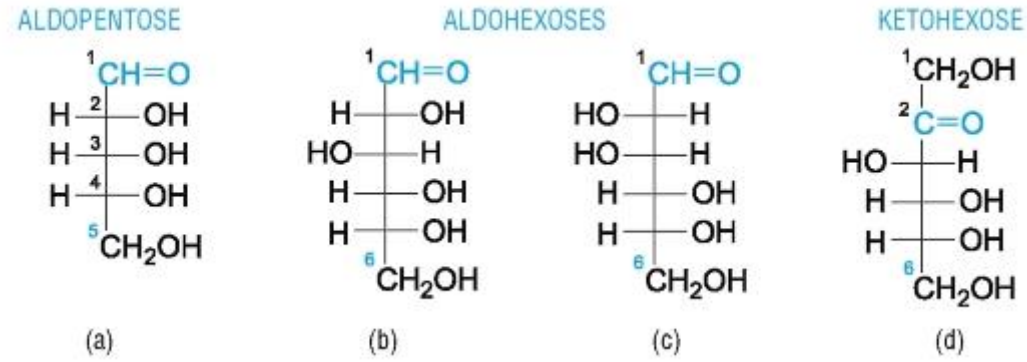


Figure 14.1. Classification of monosaccharides.

Monosaccharides have an abundance of chiral centres. Indeed, all carbons except for a carbonyl one and one of a terminal unit CH_2OH are chiral carbons. In other words, each carbon of the fragment >CHOH is asymmetric one. The Fischer projection formulas are very convenient for studying stereoisomerism of monosaccharides. Let us consider the monosaccharides presentrd in Fig. 14.1.
Example 14.1. How many stereoisomers are possible for each compound from (a) to (d) in Fig. 14.1?

Solution. It is easy to calculate the number of asymmetrical carbon atoms and the number of stereoisomers of each type of monosaccharides, which is 2^n.

The compounds (a) and (d) have three chiral centres each: C-2, C-3, and C-4 for (a), and C-3, C-4, and C-5 for (d); thus each compound can exist as eight (2^3) stereoisomers, i. e. four pairs of enantiomers.

Sixteen (2^4) stereoisomers, i. e. eight enantiomeric pairs, are possible for both (b) and (c) since each of them contains four chiral carbons, C-2, C-3, C-4, and C-5.

If the chiral carbon *farthest* from the carbonyl group has the D configuration (hydroxyl on the right), a compound is a D monosaccharide. If the remote carbon has the L configuration, the compound as a whole is an L monosaccharide. Notice that designations D and L refer *only* to the configuration of the highest numbered (or the lowest in the Fischer projection) chiral carbon. They are not used for other chiral centres.

Each monosaccharide gets a definite name according to the combination of chirality of *all* centres. The IUPAC nomenclature is based upon the trivial names with the suffix -ose. Figure 14.2 shows the Fischer projections for all aldoses through hexoses for the D series. Starting with D-glyceraldehyde, one chiral fragment CHOH (shown by the bold type) is inserted just after the aldehyde

carbon. In each step, the new chiral centre can have the OH group at the right or at the left in the Fischer projection.

Two stereoisomers that differ in configuration at only one chiral centre are called *epimers.* For example,D-glucose and D-mannose (Fig. 14.2) are epimers since they have the opposite configuration only at C-2. Note that D-glucose and D-mannose are diastereomers but not enantiomers. D-Fructose is neither diastereomer nor enantiomer but constitutional isomer of D-glucose.

Problem 14.1. Using Fig. 14.2, draw the remaining epimers of D-glucose in the Fischer projections.

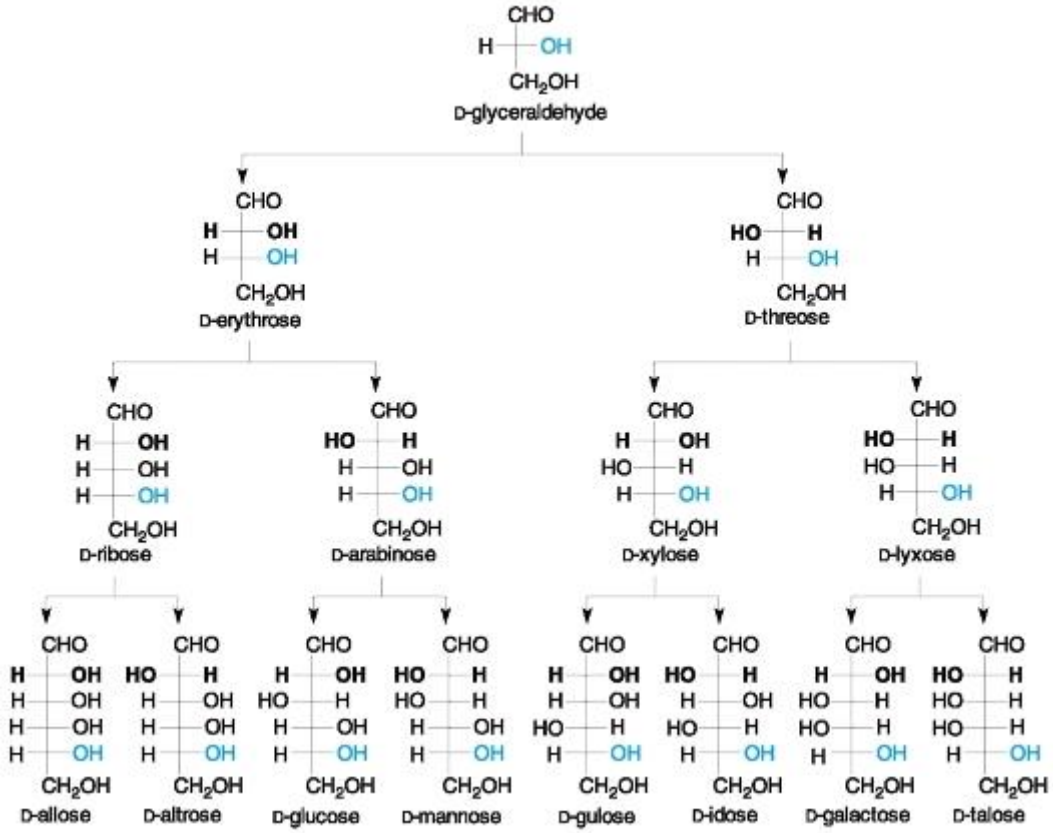


Figure 14.2. The family of the D-aldoses with up to six carbon atoms. The bold-faced H and OH show appearance of a new diastereomeric pair with additional carbon.

As you can see, all monosaccharides shown in Fig. 14.2 belong to the d series. In nature most monosaccharides, in fact, are d-sugars, but natural l-sugars are known as well. Here it should be stressed once more:

The D,L notation has no direct relationship to the sign of optical rotation of a given sugar.

Thus three of eight D-aldohexoses, three of four D-aldopentoses, and both D-aldotetroses are levorotatory substances. D-Fructose is also a levorotatory sugar, its

first name was *levulose.* Moreover, Kekule suggested for D-glucose the name *dextrose* because of its dextrorotation. The two terms remain in use in the German chemical literature up to the present time.

Only a few monosaccharides from Fig. 14.2 are virtually important, and their structures should be remembered. These are aldohexoses d-glucose, d-galactose, and d-mannose due to their wide occurrence in nature. The structures of the latter two sugars are easy to memorize once the structure of glucose is learned, since they are epimers of glucose. Mannose differs from glucose in configuration at C-2, galactose - at C-4.

The structure of ribose is very easy to memorize because all its OH groups are on the same side (for d-ribose on the right). Another aldopentose, xylose, is epimeric to ribose at C-3.

Finally, the most important ketose, d-fructose, has the same configuration at C-3, C-4, and C-5 as d-glucose.

Problem 14.2. What term would you use to describe the stereochemical relationship between: (a) D-glucose and D-mannose; (b) D-mannose and D-galactose; (c) D-arabinose and L-arabinose?

In closing the discussion of monosaccharide structures, some non-classical sugars of biological importance should be mentioned. They are not derivatives but rather analogues (or «relatives») of sugars. Their principal distinctions from classical monosaccharides are listed below:

Deoxy sugars - the absence of a hydroxyl group (or groups);

Amino sugars - replacement of an OH group (or groups) by the NH_2 and their N-acetylated derivatives group or -$NHCOCH_3$ group;

Uronic acids - the atom C-6 is oxidized to a carboxyl group;

Aldonic acids - the atom C-1 is oxidized to a carboxyl group;

Aldaric acids - both C-1 and C-6 are oxidized to carboxyl groups;

Alditols - a carbonyl group is reduced to an alcoholic group.

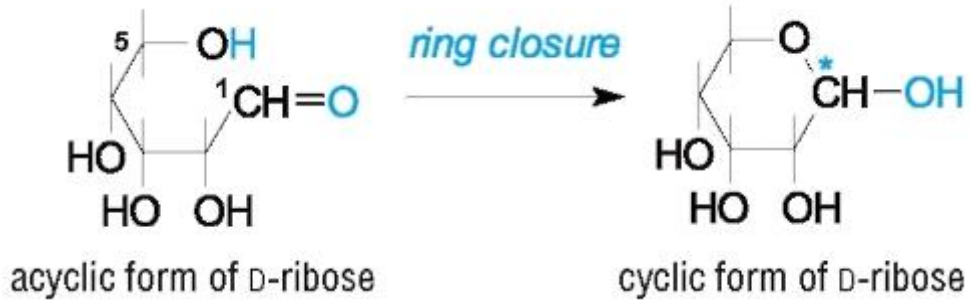
Specific examples of such unusual monosaccharides will be given hereafter.

14.1.2. The Cyclic Hemiacetal Structures

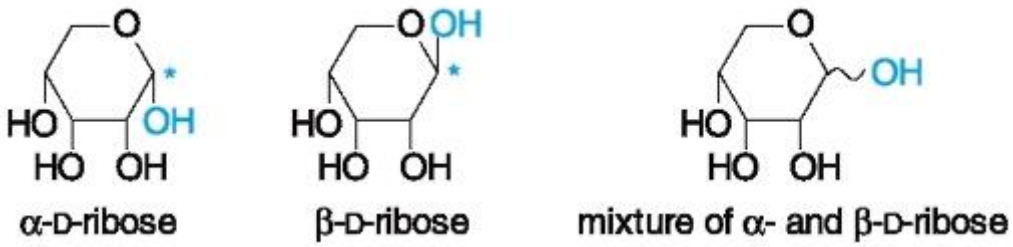
Although the structures of monosaccharides described so far are consistent with much of their chemistry, they do not respond to some physical and chemical properties of the compounds. It is time to examine the true monosaccharide structures.

Recall that γ- and δ-hydroxy aldehydes readily form cyclic hemiacetals (Sec. 11.2.1). The same is true of monosaccharides. The following scheme shows how the chain in D-ribose can be arranged so that the OH group on C-5 comes within reacting distance of the aldehyde carbon (C-1). Interaction of the two functional groups results in the cyclic, six-membered hemiacetal forms of the monosaccharide

(most H symbols are not shown for simplicity; both the hydrogens and bonds to them will be omitted hereafter).



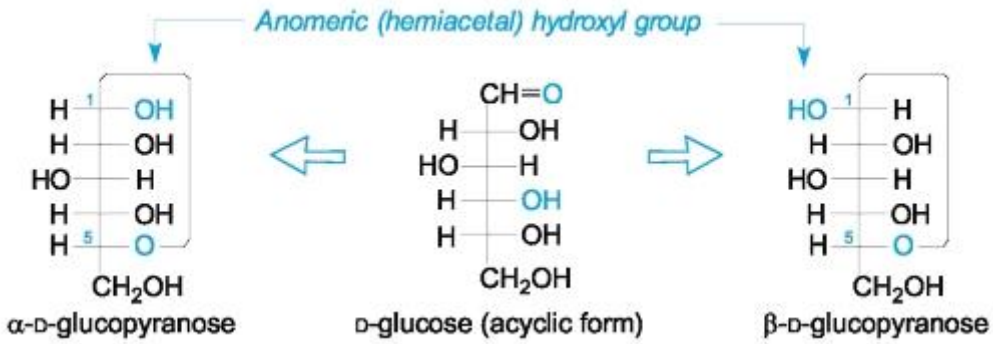
Anomeric centre. In the acyclic form of an aldose, the C-1 atom is achiral, but it becomes chiral in the cyclic structures. Thus, two hemiacetal structures are possible, differing in the configuration at the new chiral centre. This centre is called the anomeric centre, a new OH group at C-1 is also called anomeric (or hemiacetal) hydroxyl group. Two compounds that differ only in their configuration at the anomeric centre are said to be *anomers.* Anomers are attributed to the α or β type, depending on the position of the OH group (see below).



The anomeric mixture of monosaccharides can be depicted by a single structural formula with a wavy bond, like ~OH (as above).
Conventions for writing cyclic monosaccharide structures. The six-membered cyclic form of most monosaccharides is the preferred structure. These structures are called *pyranose* forms, after the six-membered oxygen heterocycle *pyran.* Thus, the full names of the anomeric forms of D-ribose in the example above are α-Dribopyranose and β-D-ribopyranose.

Three types of formulas commonly represent the cyclic structures of monosaccharides. *Conformational*formulas that will be discussed later give the closest representation of the true molecular geometry and shapes, but it is a little difficult to draw them. Two other types are the *Fischer projection formulas* and*Haworth formulas*[1].

The Fischer projection formulas demonstrate visually configurations at each asymmetric carbons. They can be adopted to show cyclic structures on the example of D-glucose:

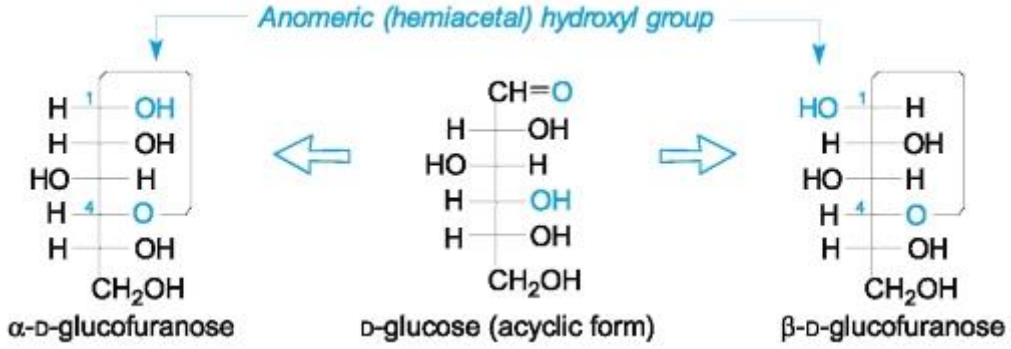


In the centre of the drawing we have the open-chain aldehyde form (sometimes designated *aldehydo*-or al-) of the monosaccharide. It is shown at the right and left that the OH group at C-5 is a part of hemiacetal structure. In the α-anomer the C-1 atom has the same configuration as configurational carbon (C-5 in hexoses and C-4 in pentoses), i. e. the anomeric OH group is at the right for D-sugars.

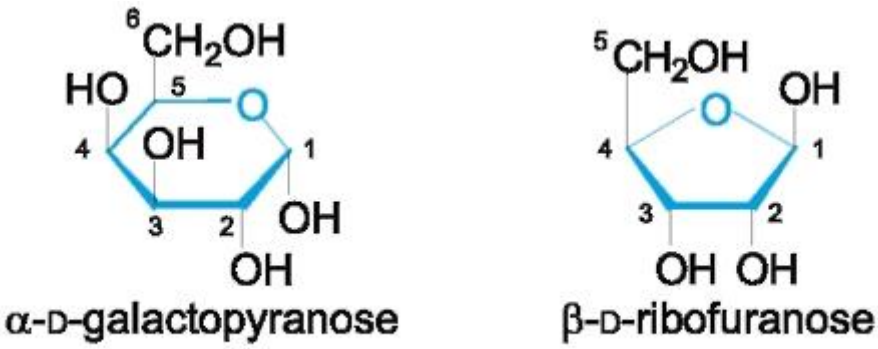
Although the Fischer conversion is fairly easy to use for cyclic structures, it contains long and distorted bonds that connect C-1 and C-5 through the ring oxygen.

There is another possibility for the cyclization. If the hydroxyl group at C-4 of an aldose participates in the cycle formation, the cyclic product represents a five-membered ring called a *furanose,* after the parent heterocycle *furan.*

[1] They are named after the British chemist W.N. Haworth, the Nobel Prize winner (1937).



The Haworth formulas that appeared in the 1920's use planar hexagon or pentagon to represent the cyclic structures, pyranose and furanose, respectively. The monosaccharide is depicted with the carbon chain horizontally, the anomeric C-1 atom being to the right. The cyclic oxygen is then depicted as being formed behind the plane of the paper. The ring is therefore located in a plane perpendicular to the plane of the paper and the groups attached to the carbons are above and below the ring. The groups, including the anomeric hydroxyl, which occur to the right in the Fischer projection then appear below the ring plane. The atom C-6 (usually CH_2OH group) will always be above the plane in the pyranose forms of D-sugars.



The Haworth perspective formulas clearly show the configuration at each chiral centre, but it should be kept in mind that they do not correspond to true geometry of the monosaccharide molecule either.

Mutarotation. Tautomerism of monosaccharides. What is the true structure of monosaccharides, open chain (aldehyde) or cyclic (hemiacetal)?

Let us consider the following experimental data. If D-glucose is crystallized from aqueous ethanol, the pure α form is obtained. On crystallization of D-glucose from acetic acid, another product was isolated, the β form. Both forms of D-glucose are *diastereomers* since they differ only in the configuration at C-1. Being diastereomers, the α and β forms have different physical properties such as melting point, specific rotation, solubility, etc.

The two forms of D-glucose interconvert in solution (Fig. 14.3). For example, if the pure α anomer is dissolved in water, the specific rotation decreases with time from an initial value of +112° to an equilibrium value of +53°. On the other hand, the same experiment with the pure β anomer results in a gradual increase of specific rotation from an initial +19° to the same equilibrium value of +53°. This phenomenon is known as *mutarotation* (mutation of rotation) and can be explained by the slow conversion of the α- and β-pyranose forms into the 36:64 equilibrium mixture (contribution of furanose forms may be ignored for D-glucose).

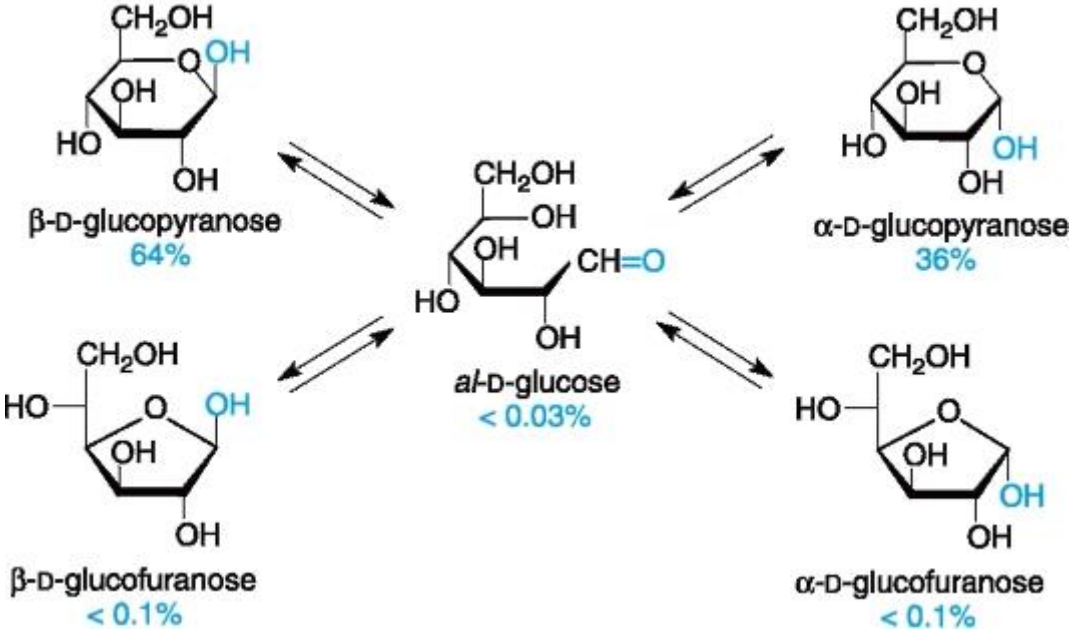


Figure 14.3. A tautomeric equilibrium of D-glucose.

Starting with either pure anomeric form, the ring opens to form an acyclic form, which then re-cyclizes to give both the α and β forms. Finally, an equilibrium mixture of five tautomers is obtained. At equilibrium, an aqueous solution of D-glucose contains approximately 64% of the β-pyranose form, 36% of the α-pyranose form, less than 0.1% of both furanose forms, and less than 0.03% of the open-chain form.

Interconversions of different forms of monosaccharides in a solution is called ring-chain tautomerism (or cyclo-oxo tautomerism).

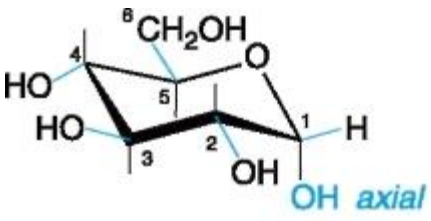
The ratio of tautomers in solution is specific for each monosaccharide. As we have seen, D-glucose in neutral solution at room temperature consists of over 99.9% of pyranose forms. D-Galactose and D-xylose have almost the same ratio of tautomers as D-glucose, whereas the α-pyranose form is predominant (68%) for D-mannose and the ratio between pyranose and furanose forms for D-ribose is about 3:1 in the same solution. Since any monosaccharide, especially in the crystalline form, is virtually a cyclic compound but not a carbonyl compound, a new definition for monosaccharide can be suggested:

Monosaccharides are cyclic hemiacetals of polyhydroxy carbonyl compounds.

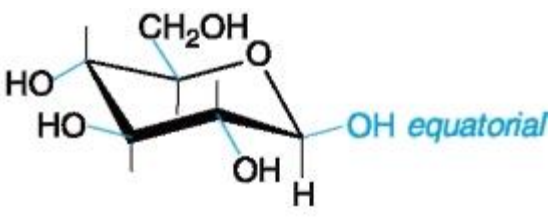
Problem 14.3. Draw the scheme of a tautomeric equilibrium for D-xylose in a solution.

Conformations. Finally, the most real structures for monosaccharides are conformational ones. It is known that a six-membered ring is nonplanar, but has the *chair* conformation, as being more stable (Section 6.2.1). The same is true of the pyranose ring.

Let us look at the conformations of α- and β-D-glucopyranose, in which the CH_2OH group and all OH groups (except for the hemiacetal OH in the α form) are in stable *equatorial* positions:



α-D-glucopyranose



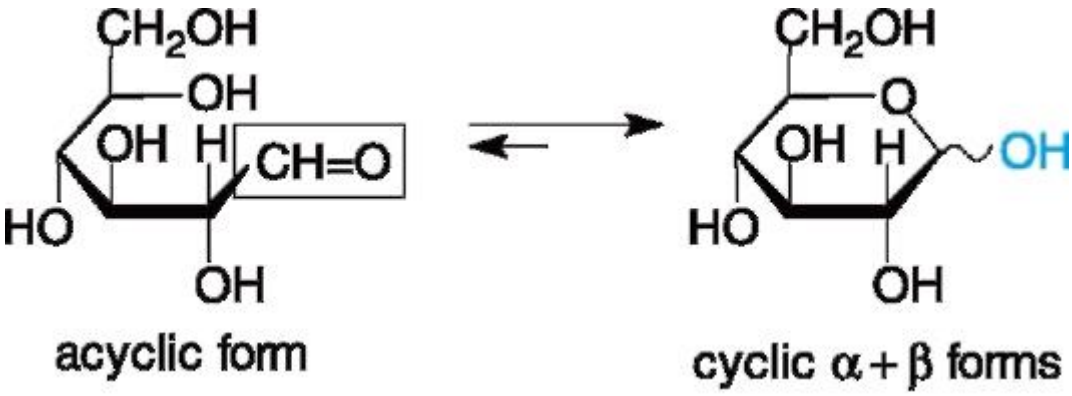
β-D-glucopyranose

Such arrangement of substituents in the ring accounts for the predominance of the β anomer over theα anomer in an equilibrium mixture of D-glucopyranose. Besides, this makes it clear why D-glucose is the most stable aldohexose. Indeed, any hexose, except for D-glucose, has at least one OH group oriented axially that diminishes its stability. Conformation of monosaccharides is very important for the space structure of polysaccharide chains.

14.1.3. Chemical Properties

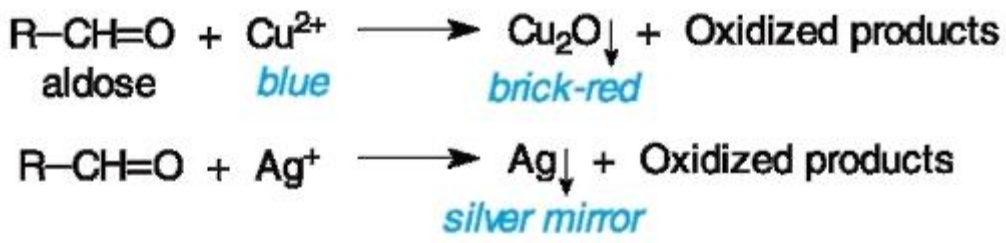
Monosaccharides are compounds of very high reactivity due to their polyand heterofunctionality. The following reactive sites can be noted in the monosaccharide molecule:

- the carbonyl group of an acyclic form shown in a rectangle;
- the hemiacetal hydroxyl group shown in colour;
- alcoholic hydroxyl groups (the remaining hydroxyls);
- a CH-acidic site (the atom C-2 in aldoses).



Let us consider all of them.

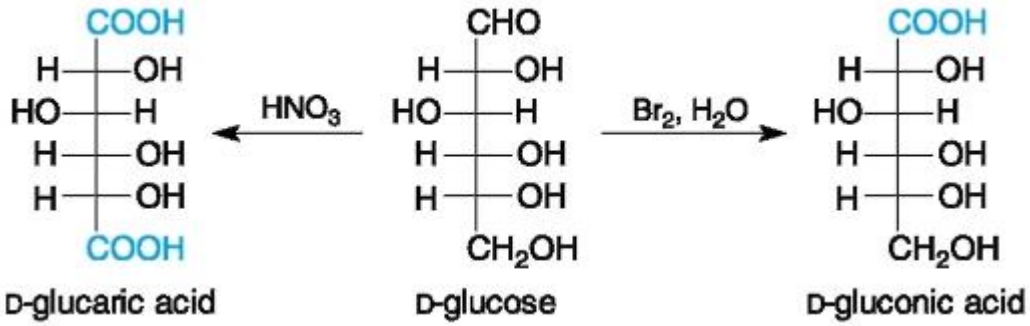
Reactions of the carbonyl group. In spite of a very low concentration of an open-chain form in a tautomeric mixture of monosaccharides, some reactions of monosaccharides as aldehydes can be carried out. For example, aldoses react with Tollens' reagent, Fehling's reagent (Section 8.4.1), or Benedict's reagent (Cu^{2+} complexed with citrate ion) to yield reducing metal species and an intricate mixture of oxidized products:



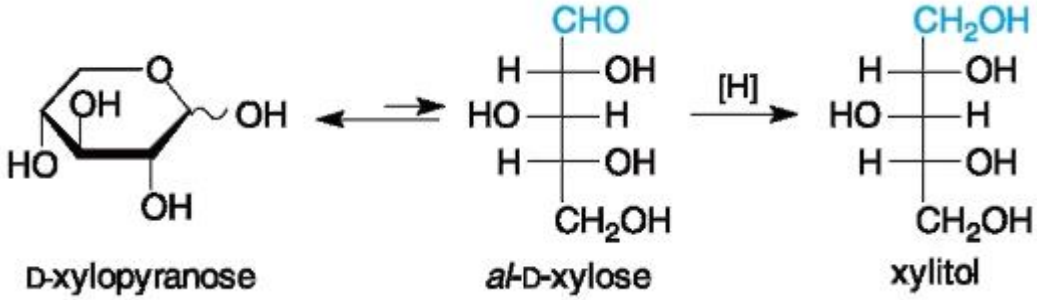
For this reason aldoses are classified as *reducing sugars.* It might be somewhat surprising that ketoses are also oxidized by the mentioned ions (recall that ketones do not normally reduce these reagents). Explanation is that ketoses readily convert into the isomeric aldoses in basic solution (for details see the end of this section).

The above reactions are used as diagnostic tools in quantitative determination of glucose in blood and urine.

Mild oxidation of aldoses with bromine water affects only the aldehyde group giving rise to *aldonic acids*(particular names are D-gluconic acid, D-galactonic acid, etc.). Stronger oxidants, such as dilute nitric acid, attack both the aldehyde group and the primary alcoholic group to form dicarboxylic acids known as *aldaric acids.*



Reduction of the carbonyl group into the CH_2OH fragment gives *alditols,* for example:



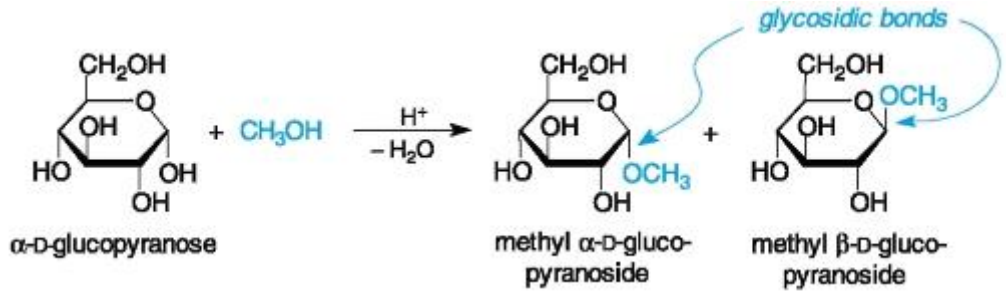
Some of them, such as xylitol and D-glucitol (the old name *sorbitol,* from D-glucose), are used as a sugar substitute for diabetics. Alditols are not involved in biochemical cycles because they do not belong to classical monosaccharides by their structure (they do not contain a carbonyl group).

Problem 14.4. Which of the following compounds belongs to reducing sugars: (a) D-glucitol; (b) D-gluconic acid; (c) 2-deoxy-D-glucose; (d) D-xylose. Explain the reason for your choice.

Problem 14.5. Propose a monosaccharide (or monosaccharides) that produces optically inactive mesotartaric acid after oxidation with nitric acid.

Reactions of the hemiacetal hydroxyl. This reactive site causes perhaps the most important chemical properties of monosaccharides. In plants and animals monosaccharides are rarely found in a free state, but mainly as acetals. (Here you should recollect the formation of an acetal in the reaction of an aldehyde with an alcohol through an intermediate hemiacetal; Sec. 8.2.1.)

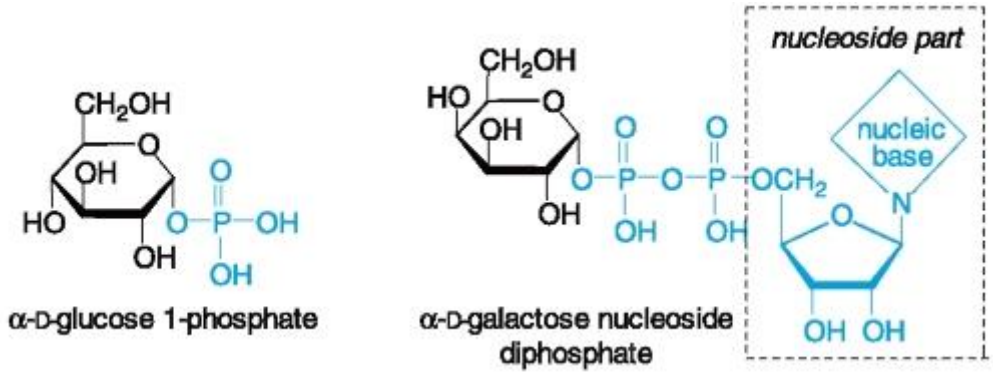
A similar reaction takes place between a monosaccharide and lower alcohols under anhydrous conditions in the presence of a proton as a catalyst, for example:



The method is known as the *Fischer glycosidation reaction* (1893). This is a familiar nucleophilic substitution reaction, in which a catalyst (H^+) converts the OH group at C-1 into a good leaving group (a water molecule).
It is interesting that the classic reaction of acetal formation was also devised by E. Fischer but several years later after his working out the glycosidation reaction.

The resulting sugar acetals are called *glycosides.* Particular glycosides are named from the respective monosaccharides, using the suffix -oside: glucosides from glucose, ribosides - from ribose, etc.; a name of the R group is placed in front the full name of a glycoside (see examples in the text). The bond from C-1 to the OR group of an alcohol is called the *glycosidic bond,* and the OR unit of the glycoside is called an *aglycone.*

Glycosides are also formed in living systems. Substrates in the reactions that occur in the organisms are sugar phosphates or more complex phosphates such as nucleoside diphosphates:

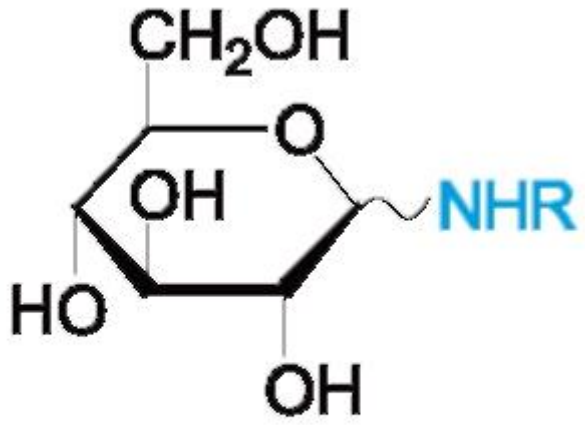


The main feature is that a phosphate and a nucleoside diphosphate (shown in colour) are excellent leaving groups.

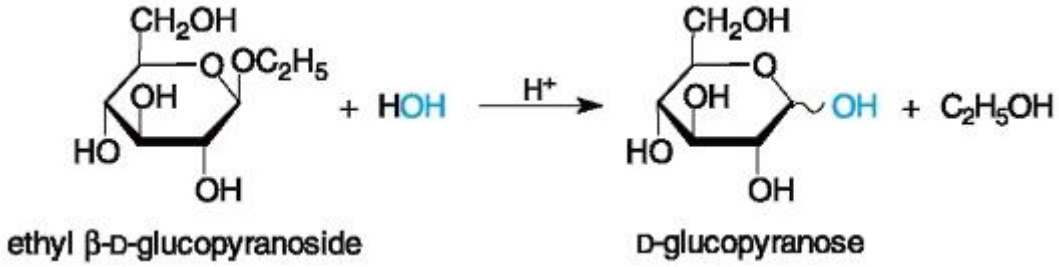
In biological systems monosaccharides form glycosidic bonds with an alcoholic OH group of another monosaccharide molecule thus giving a disaccharide, and so until a polysaccharide is formed.

The so-called *N-glycosides* relate to glycosides as their nitrogen analogues. *N*-Glycosides of D-ribose and 2-deoxy-D-ribose represent structural blocks of nucleic acids (RNA and DNA) and will be considered in Chapter 17.

The most important property of glycosides is their hydrolysis in acidic solution, whereas in dilute alkaline solution glycosides are quite stable. (Compare this with the behaviour of acetals under similar conditions.)



N-glycosides



ethyl β-D-glucopyranoside D-glucopyranose

This reaction is reversed to the formation of glycosides.
Problem 14.6. Draw an equation for the hydrolysis reaction for phenyl α-D-glucopyranoside. Under what conditions does this reaction proceed?

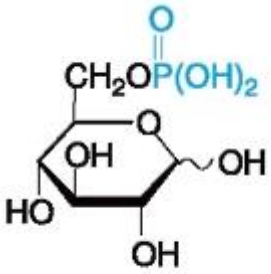
Reactions of alcoholic hydroxyls. Being polyhydroxylic compounds, monosaccharides can result in ester and ether formation. Sugar esters are formed in the reaction of monosaccharides with acylating agents such as acyl halides or acid anhydrides, for example:

α-D-xylopyranose tetra-*O*-acetyl-D-xylopyranose

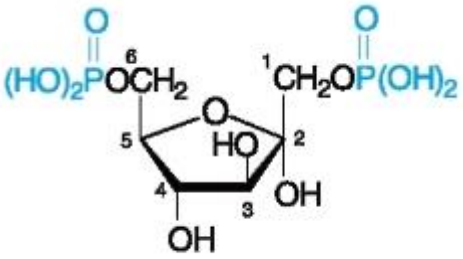
In biological transformations of carbohydrates, inorganic esters, namely, sugar phosphates, are very important.

Example 14.2. Draw the Haworth formulas for D-glucose 6-phosphate and α-Dfructose 1,6-diphosphate, both formed in the glycolysis process.

Solution. The structure of D-glucose 6-phosphate is very similar to that of 1-phosphate (see above). D-Fructose can exist in two cyclic forms, pyranose and furanose. 1,6-Diphosphate can be formed only from the furanose form that has the OH group at C-6 (to make sure of it draw the Haworth formula forD-fructopyranose).

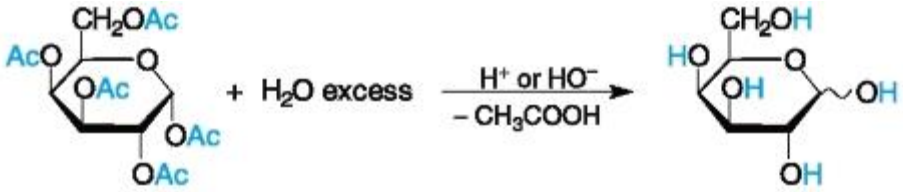


D-glucopyranose 6-phosphate



α-D-fructofuranose 1,6-diphosphate

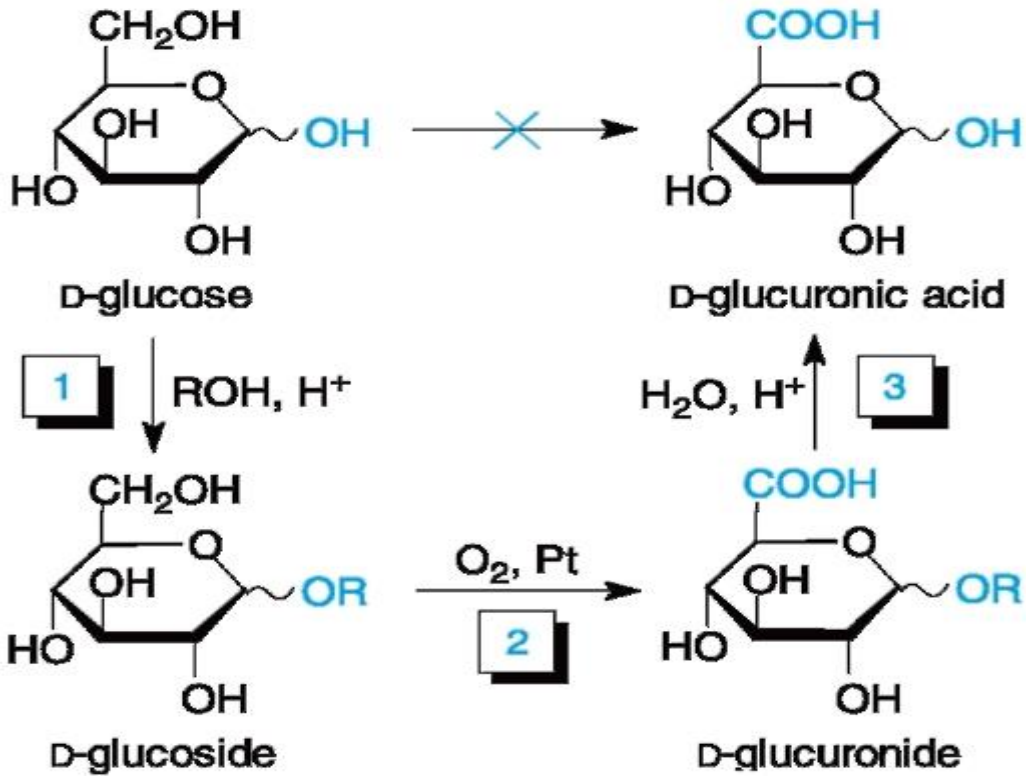
Monosaccharide esters can be hydrolyzed in an acidic or alkaline medium to the corresponding acid and alcohol, in our case it will be a monosaccharide:



penta-*O*-acetyl-α-D-galacto-pyranose

D-galactopyranose

As has already been shown, primary and secondary alcohols can be oxidized into carboxylic acids or ketones, respectively. Primary alcohols are oxidized easier than secondary ones. However, the atom C-1 (in the potential aldehyde group of an aldose) is the most sensitive site to oxidation. Therefore, if we want to oxidize a primary alcoholic group (the atom C-6 in hexoses), it is necessary to «protect» the hemiacetal group, for instance, by means of glycoside formation (step 1 in the scheme below).



D-glucose

D-glucuronic acid

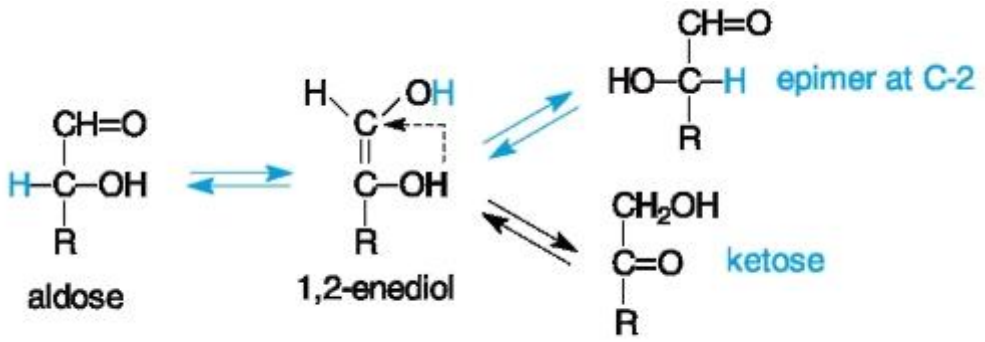
D-glucoside

D-glucuronide

Subsequent catalytic oxidation of the primary CH_2OH group in the glycoside gives a uronic acid glycoside (step 2), which then is hydrolyzed (step 3) to a *uronic acid.*

In living systems, a phosphate group at C-1 is used as a protection in the oxidation step. Uronic acids are components of connective tissue polysaccharides.
CH-Acidic site. This site is the C-2 atom of aldoses and the C-1 and C-3 atoms of ketoses, i. e. nearest to the carbonyl group.

Aldoses isomerize partly in basic solution at room temperature to give an equilibrium mixture of an epimeric at C-2 aldose and a ketose. Isomerization is thought to proceed through an intermediate 1,2-enediol.



D-Glucose thus yields a mixture with D-fructose (29%) and D-mannose (1%) on storage in calcium hydroxide solution. Such isomerization occurs also in an acidic medium, but slower, and can be catalyzed by enzymes in living systems.

14.2. OLIGOSACCHARIDES

Along with monosaccharides, other sugars that consist of several monosaccharide units are also present in the vegetable and animal sources. These are oligosaccharides, most frequently encountered are*disaccharides.*

Disaccharides are compounds that consist of two monosaccharide units

linked together by a glycosidic bond between the anomeric carbon of one unit and the hydroxyl oxygen of another monosaccharide.

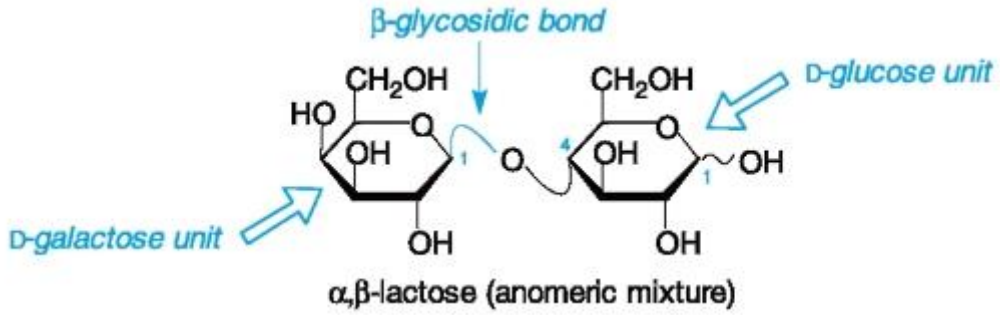
Two types of disaccharides (like all oligosaccharides) can be distinguished, namely reducing and nonreducing compounds. They differ in the mode of bonding monosaccharide units.

Most of the naturally occurring oligosaccharides have well established common names (for example, cellobiose, maltose, lactose, and sucrose) which were assigned before their complete structures were known.

14.2.1. Reducing Disaccharides

By definition, one hemiacetal hydroxyl group is preserved in the molecule of a reducing disaccharide. It becomes clear on consideration of the structure of *lactose.*

Lactose, or milk sugar, is a sugar found in human milk (6-8%) and, to a lesser degree, in cow's milk (about 4%). The monosaccharide composition of lactose is determined by the results of its acidic hydrolysis that gives equal amounts of D-glucose and D-galactose. Chemical and enzymic analyses showed that the anomeric carbon of the galactose unit has the β configuration and is linked to the OH group at C-4 of the glucose unit:



Note that the OH group at C-1 of the glucose unit in lactose is the hemiacetal one. Naturally, both cyclic hemiacetal forms exist in solution at equilibrium with an open-chain aldehyde form as shown in Fig. 14.4.

Lactose and other reducing disaccharides, like monosaccharides, give the positive Tollens' and Fehling's tests.

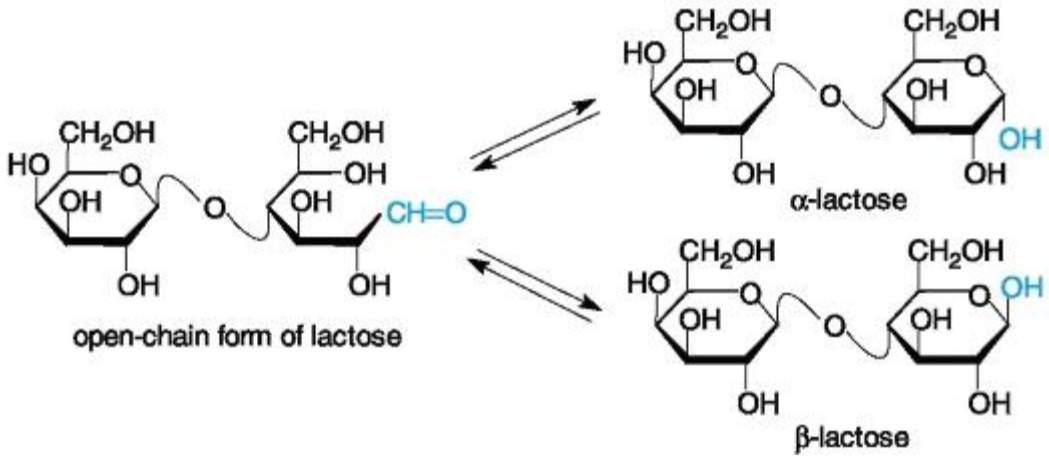


Figure 14.4. A tautomeric equilibrium of lactose.

Systematic nomenclature of disaccharides. Reducing disaccharides are in principle *O*-substituted derivatives of a monosaccharide to which another (nonoreducing) monosaccharide unit is bound by a glycosidic bond. This substituent is generally called glycosyl, in particular, glucosyl or to be more specificβ-D-glucopyranosyl depending on its complete stereochemistry. Thus the systematic name for α-lactose is 4-O-(β-D-galactopyranosyl)-α-Dglucopyranose.

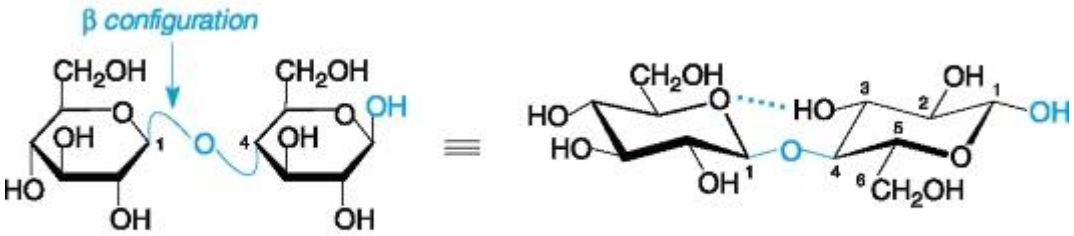
Alternatively, α-lactose may be named as O-β-D-galactopyranosyl-(1 →4)-α-D-glucopyranose, where locants in parentheses indicate the respective positions involved in the glycosidic bond. The locants are separated by an arrow pointing from the glycosyl carbon to the hydroxylic carbon involved. This method is especially useful for naming longer oligosaccharides. It may be simplified if a

three-letter abbreviation is applied for monosaccharide units, for example, β-D-Galp-(1→4)-α-D-Glcρ, where the italicized letterp designates the pyranose ring.

The most common monosaccharides are abbreviated as follows: Glc - for glucose, Gal - for galactose, Man - for mannose, Rib - for ribose, Fru - for fructose, etc.

Cellobiose is a disaccharide that gives on hydrolysis only D-glucose. Hence, cellobiose consists of two linked glucose units.
Example 14.3. Draw the structural formula for β-cellobiose, taking into account its systematic name O^-D-glucopyranosyl-(1 →4)-β-D-glucopyranose.

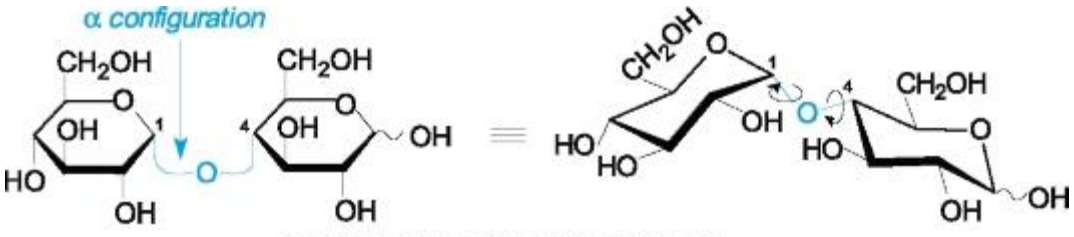
Solution. According to this name, the anomeric carbon of a nonreducing unit is linked to the hydroxyl group at C-4 of the other (reducing) unit and has the β configuration. Both units are pyranoses. The second symbol β in the name designates the configuration of the atom C-1 in the reducing unit.



β-cellobiose and conformational formula

In the conformational formula of cellobiose (see above), the oxygen atoms of each pyranose ring (but not glucose units as a whole) are reflection symmetric. This is a result of intramolecular hydrogen bonding between the OH group at C-3 and the cyclic oxygen of the left glucose unit.

Maltose, or malt sugar, is a disaccharide obtained by partial hydrolysis of starch or glycogen, and represents a stereoisomer of cellobiose. Maltose differs from cellobiose only in configuration of a glycosidic bond which is α, instead of β as in cellobiose.



maltose and conformational formula

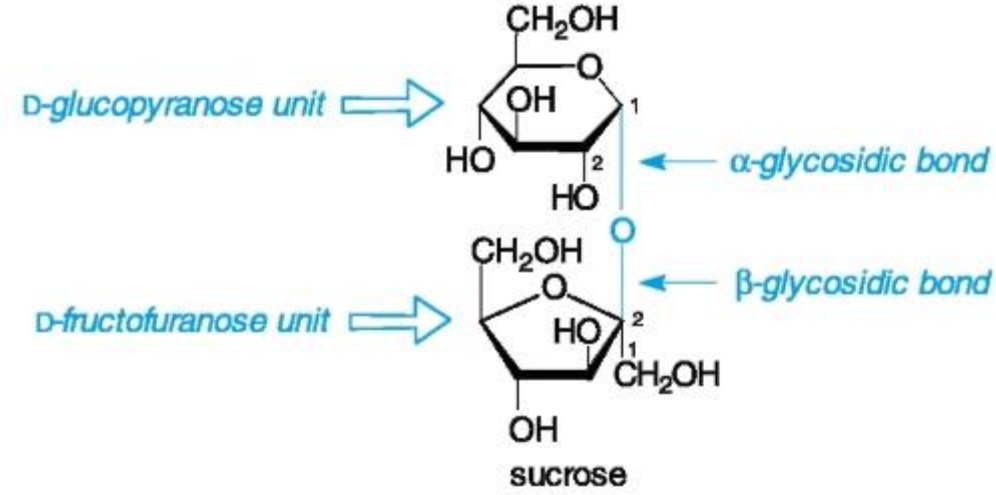
Spatial structure of maltose is more flexible as compared with that of cellobiose owing to free rotation around two C-O bonds. One of the possible conformations of maltose is shown above.

Problem 14.7. How many reducing disaccharides may be constructed from the D-glucopyranose units (tautomerism may be ignored)? Draw the structures for two examples except for those given in the text, using the Haworth formulas. Give systematic names for them.

14.2.2. Nonreducing Disaccharides

Only one representative of nonreducing disaccharides will be considered - *sucrose*, but it is perhaps the most abundant and important of all disaccharides. Sucrose is an ordinary table sugar which is produced commercially from sugar cane or sugar beets in the amount of over 100 million tons annually in the world. It occurs in all photosynthetic plants.

Hydrolysis of sucrose gives equal amounts of D-glucose and D-fructose. The principal difference of sucrose from the aforesaid disaccharides is that the anomeric carbons of both monosaccharide units are involved in the glycosidic bond. Namely, the atom C-1 of the glucose unit is linked to the atom C-2 of the fructose unit through an oxygen bridge. Note that the fructose unit is present in the furanose form. Thus, the structural formula for sucrose is:



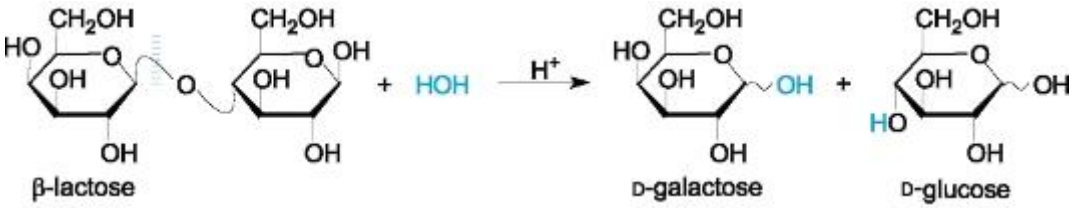
Both anomeric carbons in sucrose are linked; therefore no hemiacetal group remains in either monosaccharide unit. An acyclic form is impossible for sucrose, therefore this disaccharide cannot mutarotate. It therefore is referred to as a nonreducing disaccharide. In its chemical nature sucrose belongs rather to glycosides.

It should be kept in mind, however, that sucrose reduces the Tollens' and Fehling's reagents to a certain degree. This unexpected phenomenon is caused by partial hydrolysis of furanosides in an alkaline medium (not to mention an acidic medium). For this reason, sucrose is not a typical nonreducing disaccharide because it gives many other colour reactions characteristic of monosaccharides.

14.2.3. Chemical Properties

Many properties of oligosaccharides are very similar to those of monosaccharides. Reducing disaccharides can be oxidized and reduced. All disaccharides being polyhydroxyl compounds can produce corresponding esters.

Oligosaccharides differ from monosaccharides in the possibility to be hydrolyzed under acidic conditions resembling monosaccharide glycosides in this respect. This is illustrated by the reaction of hydrolysis of lactose:



Note that the products are not β anomeric monosaccharides but their tautomeric mixtures.

14.3. POLYSACCHARIDES

Polysaccharides constitute the dominant mass of organic matter in the Earth's biosphere. They have three important functions in the living organisms, such as an energy source, as structural components of cells and tissues, and as protective substances.

Polysaccharides are high-molecular carbohydrates built up of many monosaccharide units and therefore possess a high molecular weight. A principal structure of polysaccharides is similar to that of reducing oligosaccharides, i. e. one monosaccharide unit is bound with the next unit by means of glycosidic linkage. These units may form a linear chain or branched chains. Since polysaccharides have no free anomeric hydroxyls (except for only one at the end of the long chain), they are nonreducing compounds and do not mutarotate. Polysaccharides differ from simple sugars in some other characteristics. They do not have a sweet taste, and most of them are insoluble in water.
The monosaccharides occurring mostly in polysaccharide structures are D-glucose and, to a lesser extent, D-galactose, D-mannose, D-xylose, D-glucuronic acid, D-glucosamine, D-galactosamine and the N-acetates of these amino sugars. In spite of such a big variety of monosaccharides involved in polysaccharide structures, the latter are very regular in their constitution, in contrast to proteins, for instance.

There are two types of polysaccharides:

- *homopolysaccharides* - compounds built up from identical monosaccharide units;
- *heteropolysaccharides* - compounds consisting of different monosaccharide units.

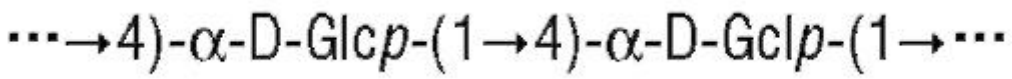
14.3.1. Homopolysaccharides

The most widespread homopolysaccharides are those composed of glucose units only and called*glucans.* The examples are starch, glycogen, and cellulose whose structures will be considered in this section.

Starch. This is the principal food reserve polysaccharide of plants, a major component of cereals, corn, and potatoes. In fact, starch consists of two fractions: *amylose* and *amylopectin.*

Amylose is the water-soluble linear polymer, composed of a chain of up to 3,000 glucose units joined by the α(1→4)-glycosidic bonds. It constitutes about 20-25%

of starch depending on its source. The abbreviated structure of amylose can be written as follows:



Amylose has helically ordered conformation (Fig. 14.5). Six glucose units constitute a coil of the helix. Amylose forms crystalline complexes (so-called inclusion compounds, or clathrates) with iodine and some polar compounds, in which «guest» molecules penetrate into a helix channel with axial disposition. Such a complex with iodine is deep-blue coloured, and its formation is used as the test for detection of both starch and iodine; this is the so-called *iodine-starch test.* Amylopectin represents the water-insoluble fraction of starch which is a branched polysaccharide with the α(1→4) bonds and α(1→6) bond in a branching point (Fig. 14.6).

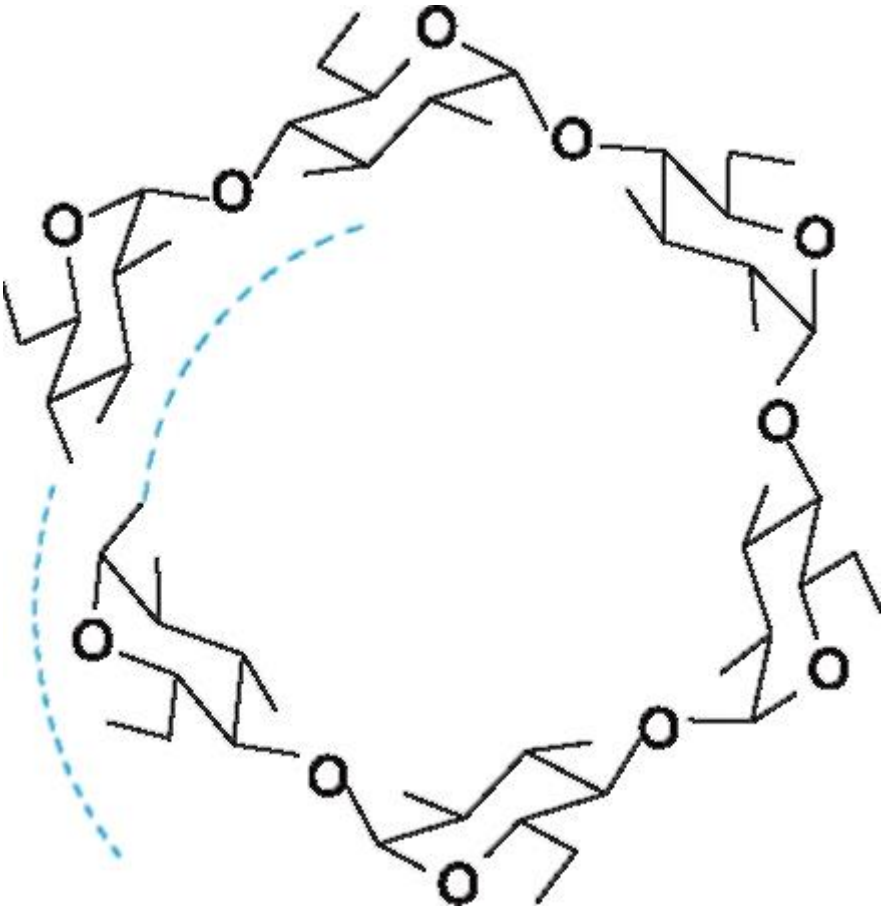


Figure 14.5. The helical structure of the amylose fragment.

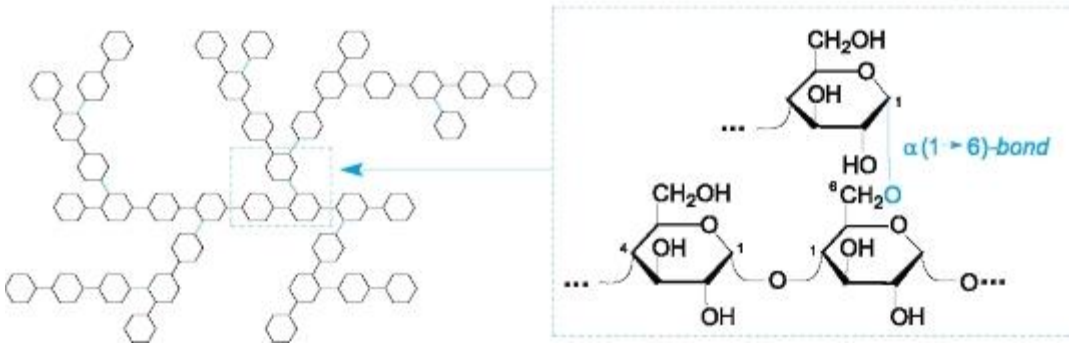


Figure 14.6. Schematic representation of the amylopectin fragment. The hexagons represent a-D-glucose units, the black dashes symlolize (1 →4) bonds, the coloured dashes - (1 →6) bonds.

Amylopectin has an average molecular mass about 10^6-10^7 (tens of thousands of units). The branching of the chain occurs about every 20 to 30 glucose units. Thus, amylopectin is organized in hundreds of relatively short chains and has a compact shape.

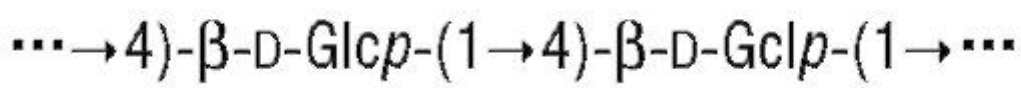
Animals have enzymes called *amylases* that cleave (hydrolyze) the α-glucosidic bonds of a starchy food such as bread, corn, or potatoes. The initial product of hydrolysis is the *disaccharide* maltose which is further hydrolyzed to glucose in the intestinal tract. In the laboratory, starch can be hydrolyzed either partially to shorter polysaccharides called *dextrins* or completely to yield glucose.

Problem 14.8. Write an equation for the complete hydrolysis of amylose, using the Haworth formula for a disaccharide fragment of the polysaccharide. Under which conditions this reaction proceeds?

Glycogen. This is a branched polymer of D-glucose with the molecular weight up to tens of millions. Its structure is very similar to that of amylopectin (Fig. 14.6), differing in a higher degree of branching. The branching in glycogen occurs about every 10-12 glucose units in the outer chains or even 3-4 units in the inner chains.

Glycogen is the storage form of glucose in animals. It is found in the liver and muscle tissues. When glycogen is hydrolyzed in the animal body, it forms glucose to help maintain the normal sugar content of the blood.
Cellulose. The most widespread organic substance on the Earth is cellulose. This polysaccharide occurs in various plants as a cell-wall material. Wood contains about 55% of cellulose, cotton and flax being almost pure (over 98%) cellulose.

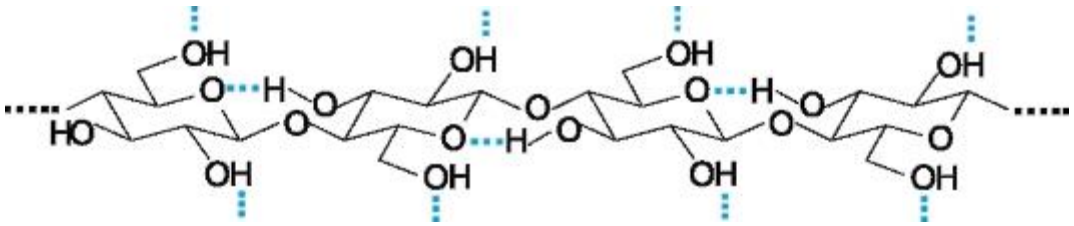
Like amylose, cellulose is composed of a straight chain of D-glucose units linked by the (1 ->4)-glycosidic bonds and numbers in 10,000 units. The main difference between the two glucans is in the configuration of the glycosidic bonds, in amylose it is α, while in cellulose it is β as shown in the abbreviated form:



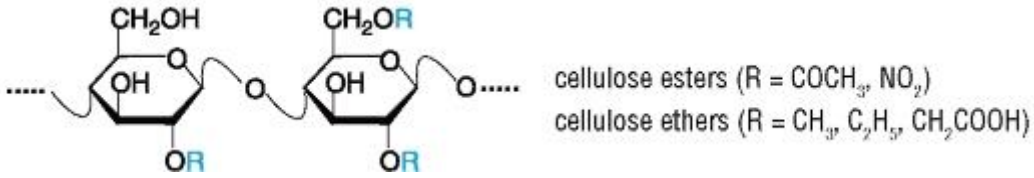
This stereochemical difference is, however, crucial for the biological fate of both the polysaccharides. Humans do not contain enzymes that catalyze the hydrolysis

of the β-glucosidic bond; such enzymes are present in many bacteria. Ruminants (cows, for example) can digest grass because they have the necessary microorganisms in their digestive system. We can eat bread and potatoes but not wood and grass.

In the linear conformation of the cellulose chain, the OH groups at C-3 are involved in intermolecular hydrogen bonding like in cellobiose (Section 14.2.1), and two remaining OH groups form intermolecular hydrogen bonds with hydroxyls of the adjacent chains (not shown in the drawing below). Such rigid structure of cellulose ensures its insolubility in water and considerable mechanical strength.

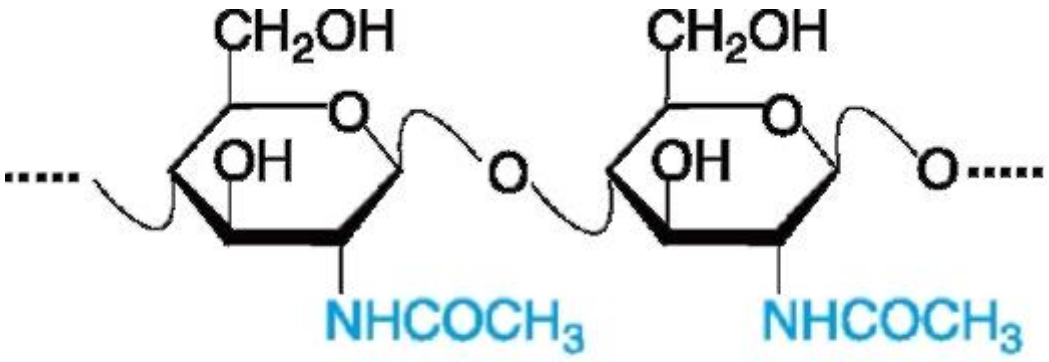


Cellulose is a readily available material for the preparation of commercially important derivatives. Hydroxyl groups of the polysaccharide are subjected to chemical modifications to give rise to cellulose esters (acetate and nitrates) and ethers (alkyl and carboxymethyl derivatives) with random disposition of the R substituents:



These derivatives are used in manufacturing textiles, fibers, coatings, plastics, and cosmetics.

Chitin. The repeating unit of this homopolysaccharide is N-acetyl-D-glucosamine (more precisely 2-acetamido-2-deoxy-D-glucose), the glucose analogue in which the OH group at C-2 is replaced by the acetamido group, $-NHCOCH_3$. In all other respects, the structure of chitin is similar to that of cellulose. Like cellulose in plants, chitin is a hard constructive polysaccharide; it forms the shells of crustaceans and insects.

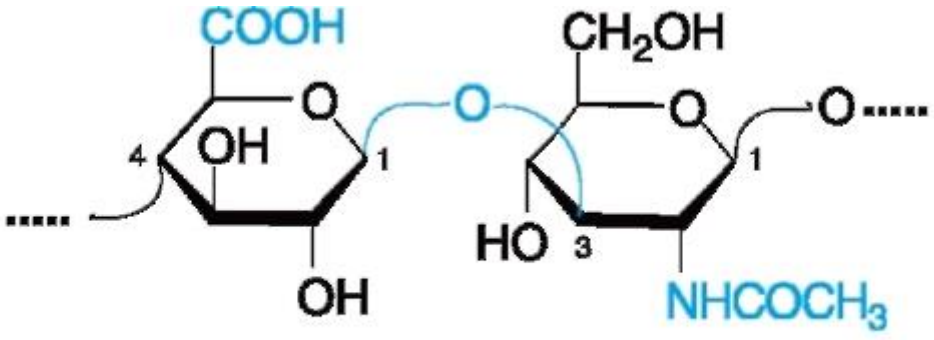


fragment of chitin molecule

14.3.2. Heteropolysaccharides

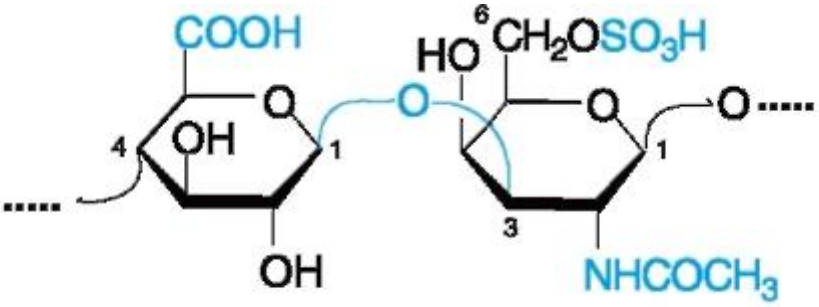
In living systems a significant role belongs to heteropolysaccharides that are composed mostly of repeating disaccharide blocks. Only two examples from numerous animal and plant heteropolysaccharides are presented in this section.

Hyaluronic acid is a structural polysaccharide found in higher animals. It is an essential component of the ground substances (or intercellular cement) of connective tissue. It has a high viscosity and a molecular weight in tens of millions. The molecule consists of repeating blocks of D-glucuronic acid and N-acetyl-D-glucosamine joined by the β(1→3) linkage. The disaccharide blocks, in turn, are joined by the β(1→4) linkage:



repeating disaccharide block of hyaluronic acid

Chondroitin sulfates are heteropolysaccharides important to the structure of the cartilage, skin, tendons, cornea, and heart valves. Their structures are similar to hyaluronic acid except that N-acetyl-D-galactosamine sulfated at the position O-6 (see below) or O-4 replaces N-acetyl-D-glucosamine:



repeating disaccharide block of chondroitin 6-sulfate

14.4. CARBOHYDRATES ON CELL SURFACES

Oligoand polysaccharide chains are often a component of complex biopolymers. When carbohydrate chains are linked covalently by O- or N-glycosidic bond with a protein molecule such biopolymers are called *glycoproteins.* Sites of binding are shown in Fig. 14.7.

Relatively short oligosaccharide chains of glycoproteins act as biochemical labels on cell surfaces. Immunoglobulins and so-called blood-group substances (antigens) belong to glycoproteins. In the schematical structure (Fig. 14.8), a great number of

carbohydrate chains (up to 80% by weight) are bound to the protein chain like bristles in a brush for cleaning bottles.

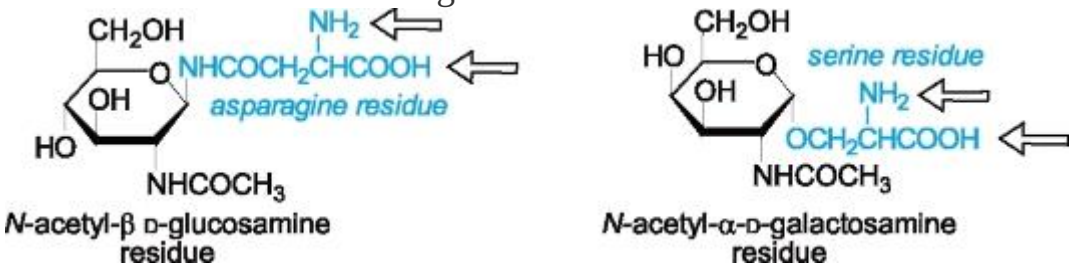


Figure 14.7. Sites of binding carbohydrate chains to protein: N-acetyl-D-glucosamine to asparagines (left) and N-acetyl-D-galactosamine to serine (right). Continuations of a protein chain are shown by arrows.

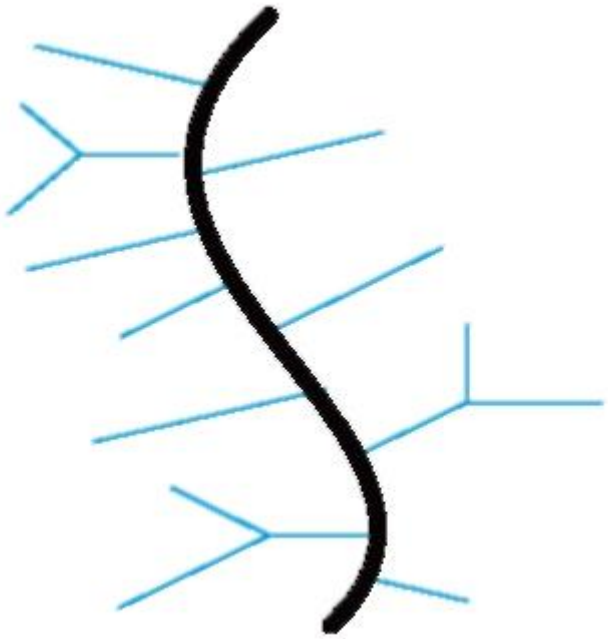
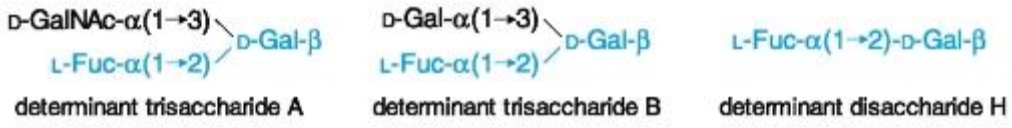


Figure 14.8. Schematic representation of glycoprotein molecule. Oligosaccharide chains are shown in colour.

The potentiality for cellular identification is exemplified by the structure proposed for oligosaccharide portions of glycoproteins responsible for the human blood-group antigens.

It is well known that human blood cannot be transfused successfully from one person to another unless it is of the proper type. This fact can be explained by the interaction of complementary proteins and glycoproteins on the cell surfaces. Erythrocytes thus carry surface-bound glycoproteins that identify them as belonging to the A, B, or H blood groups (the latter is sometimes called the «zero» group).

Each blood group in the ABH (or AB0) system is characterized by the definite oligosaccharide sequence at the nonreducing end of a carbohydrate chain called *antigenic determinants.* They contain usually no more than four monosaccharide units, for example:



determinant trisaccharide A | determinant trisaccharide B | determinant disaccharide H

Additional abbreviations of the monosaccharide units are also used: GalNAc - for *N*-acetyl-D-galactosamine and Fuc - for the deoxy sugar fucose that is 6-deoxygalactose. Note that fucose is presented as the L sugar.

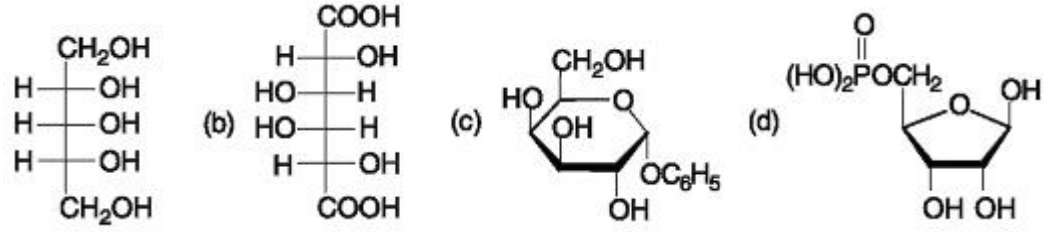
It is easy to see that all three blood-group determinants contain the disaccharide sequence shown above in colour. Interesting *in vitro* experiments were performed in the 1970's. The treatment of red blood cells of the B group with galactosidase, an enzyme that splits off the terminal D-galactose unit from the oligosaccharide chain, gave a new antigenic determinant of the H group. Thus, the B blood group was enzymically transformed into the H group.

This example demonstrates that the role of carbohydrates is not limited by the traditional function as energy sources and as structural materials. Elucidation of the role of carbohydrates in cell recognition is under active investigation.

Additional Problems

14.9. If you were observant you could see, the name *xylitol* was written without assignment to the D orL series, whereas we use it for D-glucitol (sorbitol). Draw the structures for both the compounds by the Fischer projection. Why is the symbol D or L omitted for xylitol?

14.10. Name the following compounds:



14.11. Draw the Fischer projections and Haworth formulas in a pyranose form for the following monosaccharides:

(a) *N*-acetyl-D-glucosamine (2-acetamido-2deoxy-D-glucose);

(b) D-fructose;

(c) D-fucose (6-deoxy-D-galactose);

(d) L-arabinose.

14.12. Draw the *chair* conformation for the most stable anomer of D-galactopyranose. Why do the anomeric forms of D-galactopyranose exist at equilibrium in a solution in nonequal amounts? Will the anomers have the same value of specific rotation, differing only in sign?

14.13. What reactions prove the presence in the structure of aldoses of:

(a) an aldehyde group;

(b) several hydroxyl groups?

Write equations, using D-glucose as an example.

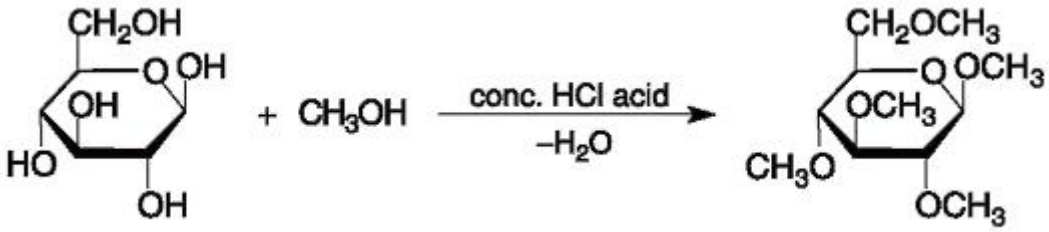
14.14. Which of the hydroxyl group of monosaccharide reacts with the excess of:

(a) methanol in acidic solution;

(b) acetic anhydride?

Illustrate your answer using D-mannose as an example and name the products obtained.

14.15. Show all mistakes in the following equation:



14.16. D-Gluconic acid forms cyclic compounds in a solution. Write equations for this transformation. Can that cyclization be referred to as tautomeric equilibrium?

14.17. Which of the following statements is (are) not correct for maltose:

(a) it consists of two D-glucopyranose units;

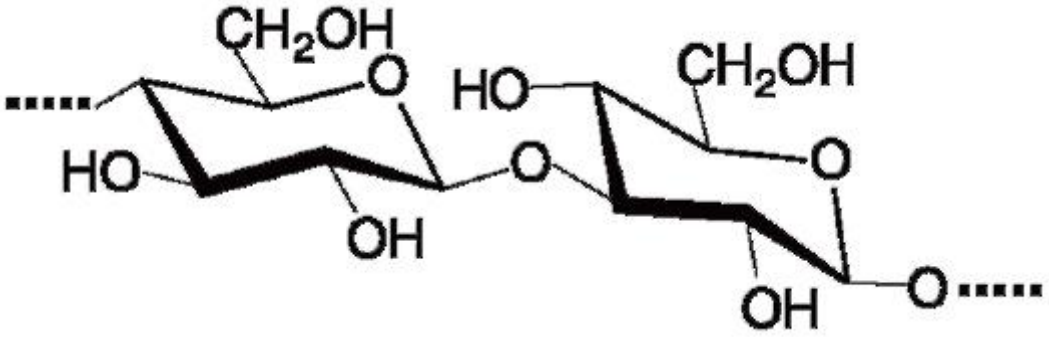
(b) it contains the β(1→4)-glycosidic bond;

(c) it is a nonreducing disaccharide;

(d) it can be hydrolyzed in alkaline solution;

(e) it can exist in ring-chain tautomeric forms.

14.18. Laminarin is a polysaccharide found in brown seaweeds of *Laminaria* species. The structure of its repeating disaccharide block is as follows:



Name:

(a) monosaccharide units of the structure;

(b) the type of glycosidic linkage between the monosaccharide units.

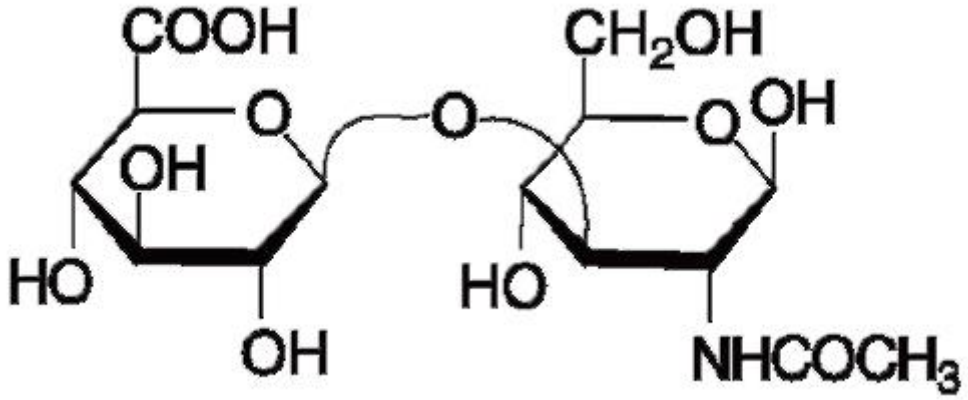
14.19. Is it possible to distinguish by polarimetry between substances in the following pairs (reference data are not available for you):

(a) α-lactose and sucrose;

(b) α-lactose and α-maltose.

14.20. Cellulose is resistant to dilute alkali, but its acetate is not. Explain this difference and write an equation for reaction, using α-cellulose acetate as an example.

14.21. Which of the following statement(s) corresponds to the disaccharide shown below:



(a) it contains the β(1→4)-glycosidic linkage;

(b) it consists of D-glucose and D-glucuronic acid units;

(c) it is a repeating block of hyaluronic acid;

(d) it contains an amide group;

(e) it contains an ester group;

(f) it can be completely hydrolyzed under acidic conditions.

Chapter 15. α-AMINO ACIDS, PEPTIDES AND PROTEINS

Peptides and proteins are polymeric compounds composed of amino acid units. It is difficult to overemphasize the role of proteins in living matter. They are the principal components of muscles, skin, hair, and blood. Proteins are constituents of antibodies and hemoglobin, having protective and transport functions. Many antibiotics and hormones are proteins, and all enzymes are proteins only. Therefore it is no wonder that the term *proteins* originates from the Greek *protos* the first.

In this chapter, we will mainly discuss the structure, chemical and some biochemical properties of amino acids. We will next consider briefly the structure and properties of peptides, and, finally, the main features of a protein structure.

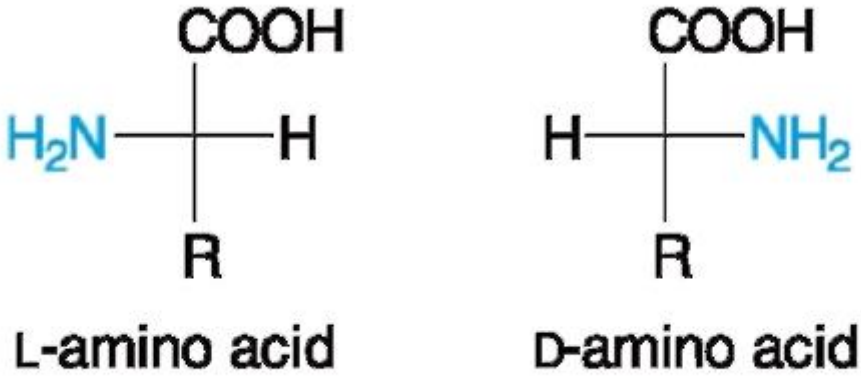
We studied amino acids to some extent in Chapter 11. From various types of these heterofunctional compounds, α-amino acids are the most important in biological

processes being the building blocks of proteins. Recall that α-amino acids are carboxylic acids with an amino group attached to the α-carbon atom; they may be represented by the general formula $RCH(NH_2)COOH$. Only α-amino acids will be the target of our consideration so the symbol α will further be omitted.

Hydrolysis of most animal proteins produces about twenty different amino acids listed in Table 15.1. With the exception of glycine (R = H in the general formula), all amino acids have four different groups attached to the α-carbon. This carbon is chiral, and two enantiomeric forms of each amino acid are therefore possible. Most natural amino acids belong to the same stereoisomeric family, namely, to the Lseries. D-Amino acids are also known in nature, but only as constituents of certain antibiotics and of proteins of bacterial cell walls. The generalized formulas for enantiomers are given in the Fischer projections.

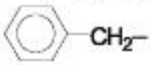
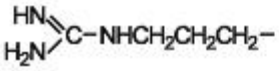
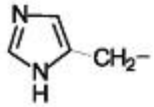
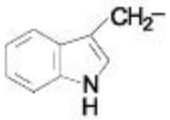
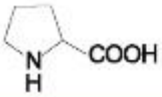
15.1. α-AMINO ACIDS

15.1.1. Structure and Classification



Amino acids are known by their trivial names, which are accepted by the IUPAC nomenclature (Table 15.1). They also have a three-letter abbreviation (mostly the first three letters are used), which are useful for writing the formulas of peptides and proteins. Systematic names of amino acids are, of course, possible but they are never used.

Table 15.1. Common amino acids $RCH(NH_2)COOH$ found in proteins

Trivial name	Abbreviation**	Structure of R	Isoelectric point
Glycine	Gly	H–	6.0
Alanine	Ala	CH_3–	6.0
Valine*	Val	$(CH_3)_2CH$–	6.0
Leucine*	Leu	$(CH_3)_2CHCH_2$–	6.0
Isoleucine*	Ile	$CH_3CH_2CH(CH_3)$–	6.0
Phenylalanine*	Phe		5.5
Serine	Ser	$HOCH_2$–	5.7
Threonine*	Thr	$CH_3CH(OH)$–	5.6
Tyrosine	Tyr		5.7
Aspartic acid	Asp	$HOOCCH_2$–	2.8
Glutamic acid	Glu	$HOOCCH_2CH_2$–	3.2
Asparagine	Asn	H_2NCOCH_2–	5.4
Glutamine	Gln	$H_2NCOCH_2CH_2$–	5.7
Lysine*	Lys	$H_2N(CH_2)_4$–	9.6
Arginine	Arg		10.8
Cysteine	Cys	$HSCH_2$–	5.0
Methionine*	Met	$CH_3SCH_2CH_2$–	5.7
Histidine	His		7.5
Tryptophan*	Trp		5.9
Proline***	Pro		6.3

* Essential amino acids.

** It may be applied only for shortened formulas of peptides (Sec. 15.2.1).

*** Full structure is presented.

Problem 15.1. Draw the Fischer projections for the following amino acids: (a)L-cysteine; (b) D-glutamine; (c) L-serine; (d) L-valine. Assign the configuration of the compounds (a) and (c), using the *R,S* system.

Amino acids can be classified as neutral, acidic, or basic, depending on the nature of their side chain, the R substituent. Most amino acids (fifteen of the twenty listed in Table 15.1) have the neutral R's. Two amino acids (aspartic and glutamic acids) have an extra carboxyl group and are acidic. Three compounds (lysine, arginine, and histidine) have an extra basic function in the side chains and are referred to as basic amino acids. In the largest group of neutral amino acids, there are compounds with polar groups in the R substituent, such as the OH groups of serine and threonine or even ionizable groups, such as a phenolic OH group of tyrosine and the SH group of cysteine.

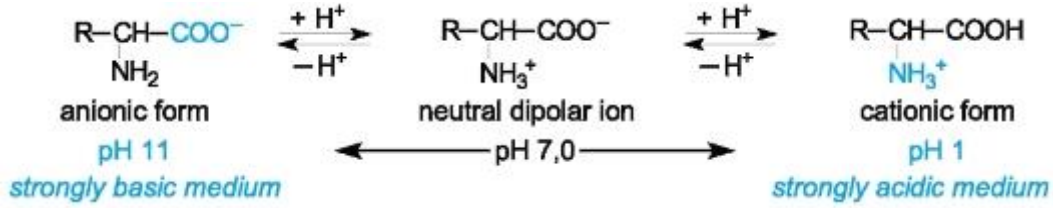
Alternatively, amino acids can be subdivided into several groups on the basis of the structural features of their side chains. These are aliphatic (the first five in the table), aromatic (phenylalanine and tyrosine), heterocyclic, hydroxyl-containing, and sulfur-containing amino acids (find representatives of the three last groups in the Table 15.1 on your own).

From a biological point of view, *essential amino acids* (Table 15.1) stand out because they, in contrast to other amino acids, cannot be synthesized in sufficient quantity by adult humans and therefore must be obtained from dietary sources. The lack of some essential amino acids in the diet can lead to severe deficiency diseases.

15.1.2. Chemical Properties

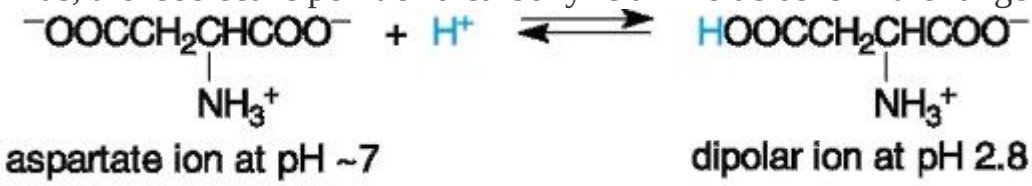
The acid-base properties. As has already been shown (Sec. 4.3), amino acids have a dipolar ion structure since they contain both an acidic group (COOH) and a basic group (NH_2) within the same molecule. The dipolar structure is in agreement with the salt-like properties of amino acids, which are crystalline compounds with high melting points (in the range of 220-340 °C) and are much better soluble in water than in organic solvents.

The predominant form of an amino acid in solution depends on the pH of the solution and on the nature of the amino acid (i. e. the R group in the general formula). In acidic solutions all amino acids exist mainly as cations; in basic solutions they are present as anions.



At some intermediate pH, the amino acid is present in an electrically neutral form. At this pH, called the*isoelectric point* (p*I)*, the amino acid exists almost exclusively in the dipolar form. The isoelectric point depends on the structure of an amino acid. Neutral amino acids have isoelectric points in the pH range of 5.0-6.3 (the pI values are listed in Table 15.1).

Aspartic and glutamic acids contain an extra carboxyl group and at neutral pH values they are mainly present in anionic form. To convert this anion into a neutral dipolar ion (in other words to reach the pl) some quantity of an acid must be added. Thus, the isoelectric point of dicarboxylic amino acids is in the range of 3.



For a similar reason, the isoelectric points of basic amino acids are in the basic region of pH.

Due to amphoteric nature of amino acids they are able to neutralize small quantities of acids or bases, thus maintaining a constant pH of the solution. Such compounds are termed *buffers* and are used in biochemical investigations.

Problem 15.2. Illustrate by equations the acid-base behaviour of the amino acids (a) - (c) in reactions with excess hydrochloric acid and of the amino acids (d) - (f) with excess sodium hydroxide:

(a) aspartic acid

(d) propline

(b) cysteine

(e) serine

(c) lysine

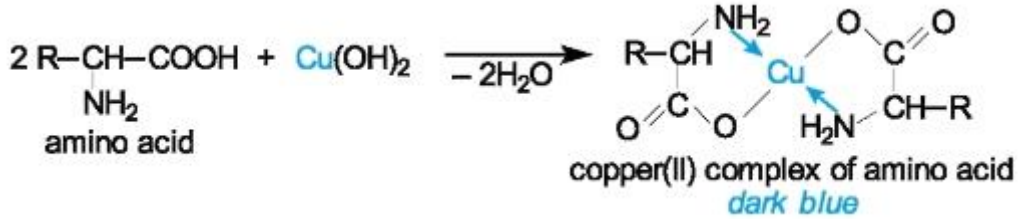
(f) tyrosine

General characteristics of functional group reactivity. Amino acids being heterofunctional compounds show the chemical behavior that would be expected due to the presence of the amino and carboxyl groups. They can be *N*-acylated or N-alkylated with participation of their amino group. They can also be esterified or transformed into amides and other carboxylic acid derivatives. Some reactions involve additional functional groups of amino acids present in the side chain. Generally, these reactions have been considered in the previous chapters.

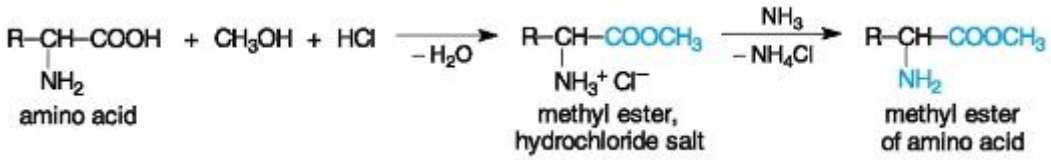
The reactions that can be useful in the amino acid analysis and identification are of interest in this section.

Complexing properties. Like 1,2-diols and α-amino alcohols (Sec. 11.2.2), α-amino acids form complex salts with some metal ions with the formation of dark blue coloured solution.

These are another characteristic of heterofunctional compounds. Their complexing (or chelating) ability is based on a tendency to form a stable fiveor six-membered cycle in the reaction with some metal ions (especially with Cu^{2+} and Ni^{2+}). For example, insoluble copper(II) hydroxide reacts with 1,2-diols with the formation of dark blue coloured solution.



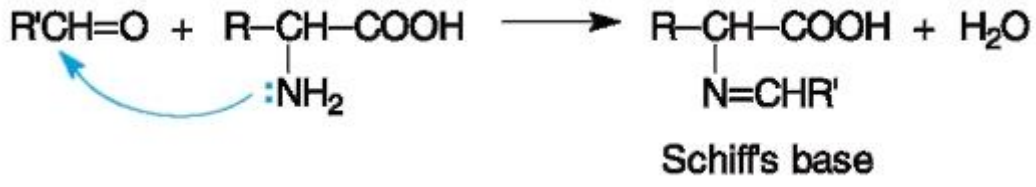
Esters formation. The familiar Fischer esterification of an amino acid yields an ester as a salt. A strong acid acts in this reaction not only as a catalyst but in the first place as an acid that converts an amino acid into protonated (cationic) form. An appropriate base (as ammonia in the example below) should be used in conversion of the product into a free base ester.



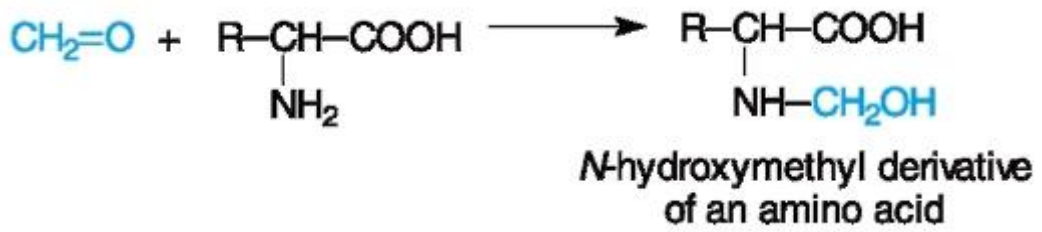
Amino acid esters, in contrast to amino acids themselves, are relatively volatile derivatives that can be distilled (sometimes in vacuum). This property is used in analysis of amino acid mixtures. Besides, amino acid esters are important intermediates in peptide synthesis.

Chemical synthesis of peptides is perhaps the most exciting field of the chemistry of amino acids over the whole 20th century; however, this question is beyond the scope of our course.

Reaction with carbonyl compounds. Most of carbonyl compounds react with the amino group of an amino acid giving Schiff's bases (Sec. 8.2.4). Such derivatives of some amino acids can be analyzed by spectral methods.



Formaldehyde reacts with amines and amino acids to form quite stable addition products (without elimination of water) in accordance with the equation:



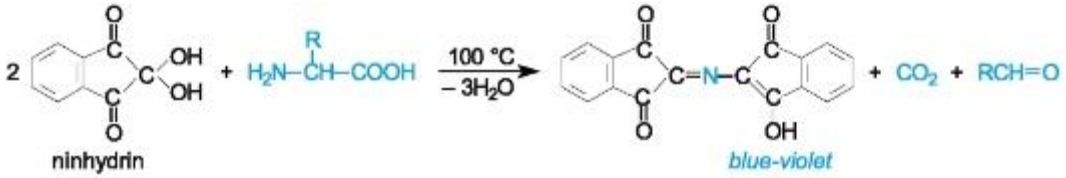
This reaction leads to the blocking of the amino group of an amino acid. In such a form an amino acid cannot exist as a dipolar ion because of low basicity of the nitrogen atom, and it is possible to determine its quantity by the titration method. This procedure is known as the *S0rensen method.*

Reaction with nitrous acid. Amino acids, like primary aliphatic amines, rapidly react with nitrous acid, HNO_2, producing alcohols and molecular nitrogen.

Measuring the volume of the nitrogen evolved it is possible to determine the quantity of amino groups in the tested sample of an amino acid. The reaction is a background of the *Van Slyke method.*

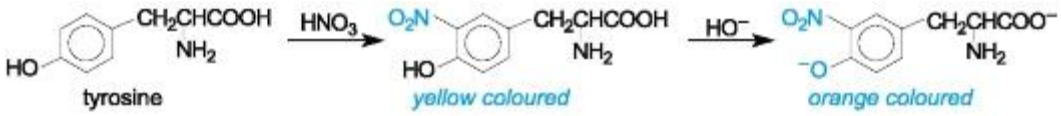
Problem 15.3. Write equations for reactions of alanine with: (a) methanolic hydrogen chloride; (b) formaldehyde; (c) nitrous acid. Name the products obtained.

Ninhydrin reaction. α-Amino acids (and primary amines) react with triketone ninhydrin to form blue-violet coloured product. Reaction is extremely sensitive and used in chromatographic analysis of amino acids.



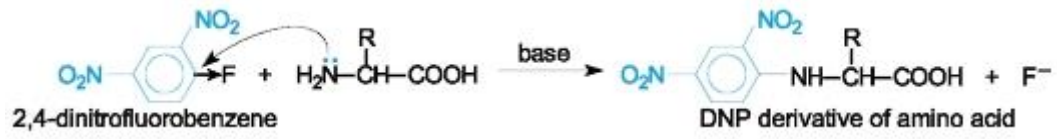
The reaction is also useful in criminalistics for visualization of fingerprints.

Xanthoproteic reaction. This reaction permits detection of aromatic amino acids: tyrosine, phenylalanine, histidine, and tryptophan. It is based on the nitration of the aromatic ring with the formation of a coloured nitro derivative, as shown by the example of tyrosine:



The colouration becomes deeper in the presence of an alkali, which ionizes the phenolic hydroxyl group.

Reaction with 2,4-dinitrofluorobenzene. The nucleophilic amino group of amino acids reacts with 2,4-dinitrofluorobenzene in mildly basic solution to give a yellow 2,4-dinitrophenyl (DNP) derivative:

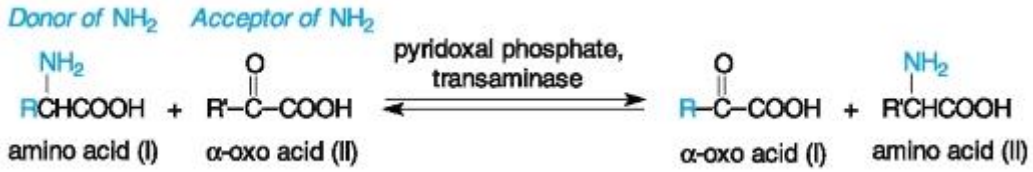


The reaction represents an example of nucleophilic aromatic substitution which is fairly rare in aromatic compounds. Two strong electron-withdrawing nitro groups facilitate substitution of the fluorine atom.

The method is mainly used in analysis of amino acids that compose a peptide molecule.

15.1.3. Biologically Important Reactions

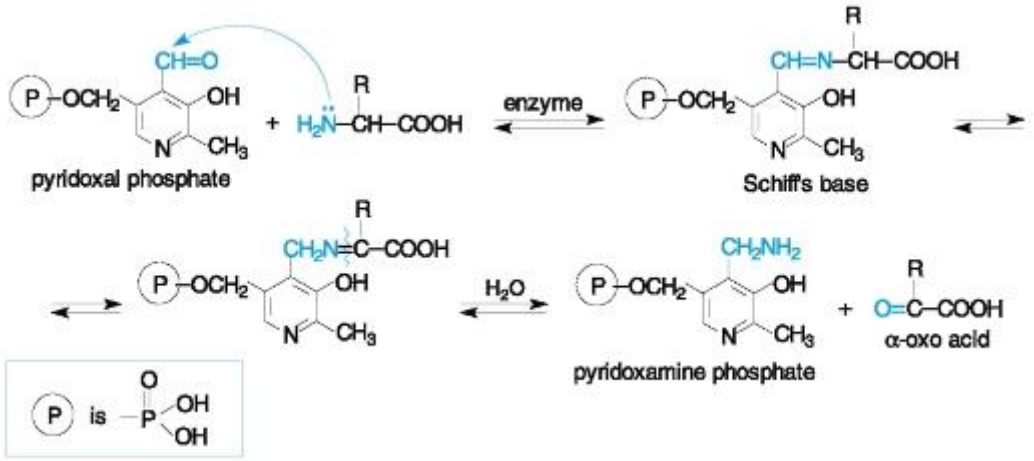
A number of reactions of amino acids occurring in biological systems proceed with the assistance of a coenzyme *pyridoxal phosphate.*
Transamination. This enzyme-catalyzed reaction is one of the most important metabolic transformations of amino acids. On the other hand, the transamination represents a pathway in which most of the naturally occurring α-amino acids are biosynthesised from α-oxo acids, according to the following general equation:



Example 15.1. Glutamic acid is produced in the organism along with oxalacetic acid. Which are their precursors in the transamination reaction?

Solution. Glutamic acid can be formed from the corresponding C_5 dicarboxylic oxo acid, i. e. 2-oxoglutaric acid. The precursor of oxalacetic acid is the corresponding C_4 dicarboxylic amino acid, i. e. aspartic acid.

In one of the initial steps of the reaction, pyridoxal phosphate (as an aldehyde) reacts with the NH_2group of an amino acid to form a Schiff's base. One of the possible transformations of the Schiff's base involves its hydrolysis that results in the formation of an α-oxo acid and pyridoxamine phosphate as shown below:

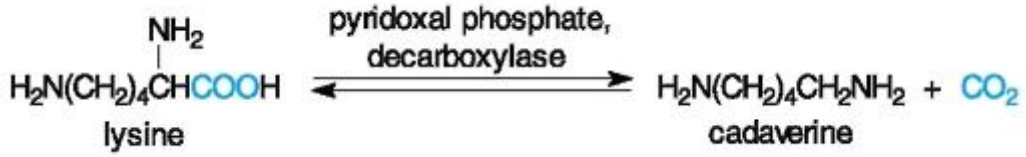


Pyridoxamine phosphate can react with another oxo acid in the reversed direction with the formation of another amino acid and regeneration of pyridoxal phosphate. A general equation shows that pyridoxamine phosphate functions as a reversible carrier of the amino group from an amino acid to an oxo acid.

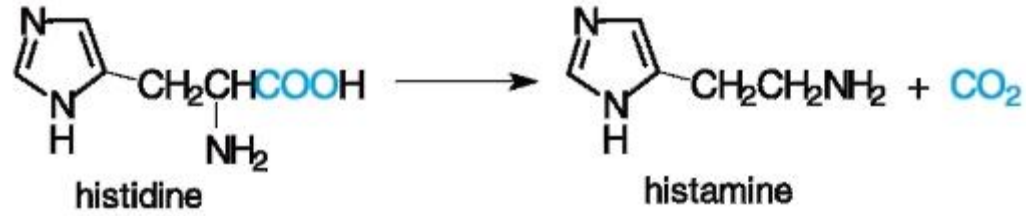
Problem 15.4. Write an equation for enzymic transamination between 2-oxopentanedioic acid and alanine. Name the products of the reaction.

Decarboxylation. This reaction also proceeds with the participation of pyridoxal phosphate and leads to the formation of naturally occurring amines. The simple diamines putrescine (1,4-butanediamine) andcadaverine (1,5-pentanediamine) occur (as their names suggest) in decomposing animal matter. Cadaverine is the

decarboxylation product of lysine:

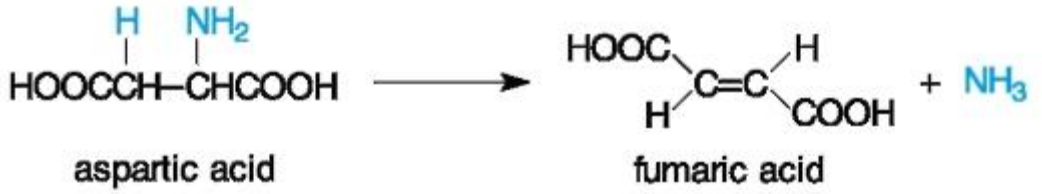


A similar reaction with histidine gives the biogenous amine, histamine.

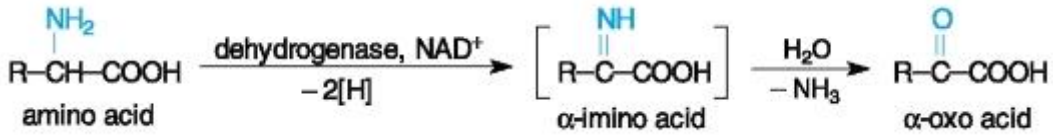


Decarboxylation of glutamic acid and tryptophan can be found in Sec. 11.4 and 16.3.

Deamination. Two types of the enzymic deamination (i. e. removal of an amino group) are known for amino acids. The first one is the *non-oxidative* deamination that takes place without the use of oxygen and leads to the formation of α,β-unsaturated carboxylic acids.



Another type of the reaction is the *oxidative* deamination which is a two-step process. The first step represents the enzymic oxidation of an amino acid into an intermediate α-imino acid in the presence of a coenzyme NAD+. Subsequent hydrolysis yields an α-oxo acid:



15.2. PEPTIDES AND PROTEINS

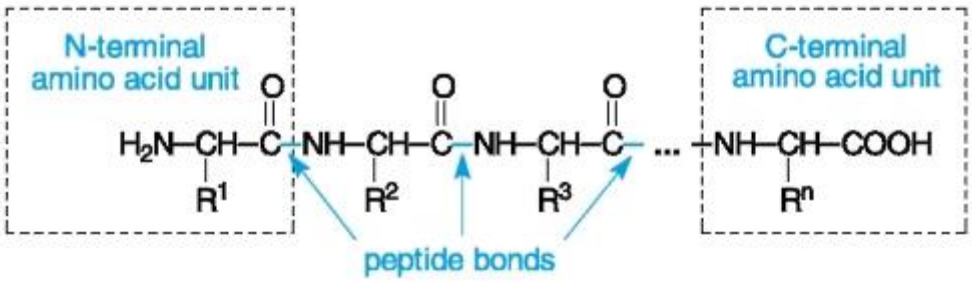
Peptides and proteins are compounds composed of many amino acids joined to one another through amide bonds called *peptide bonds.* Two terms are generally used for compounds made up of long sequences of amino acids: peptides and proteins.

The term *protein* is usually applied to naturally occurring polyamides that are derived from amino acids and have molecular mass greater than 10,000 (about 100 amino acid residues). The term *peptide* is used for natural or synthetic substances with a molecular mass less than 10,000. This limit is relative, of course. Peptides and proteins are very similar in their principal structure, but proteins, because of

their size, have more intricate, three-dimensional arrangement inherent in them than peptides do.

15.2.1. Primary Structure

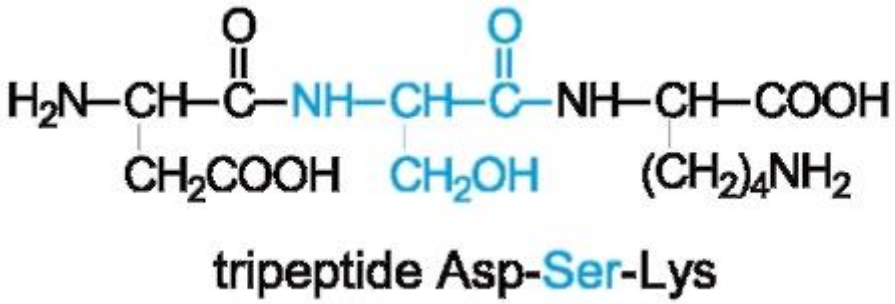
In the early 20th century, E. Fischer postulated that in peptides and proteins, peptide bonds are formed between the amino group of one amino acid and the carboxyl group of another as shown below:



If an amino acid contains two carboxyl groups or two amino groups, the α-amino and α-carboxyl groups are usually involved in the bonding.

By convention, the peptide chain is always written with the amino acid having a free NH_2 group at the left and the amino acid with a free COOH group at the right. These amino acids are called the *N-terminal* and *C-terminal* ones, respectively.

Formulas for peptides are often written in the abbreviated form for each amino acid listed in Table 15.1, starting with the *N*-terminus. Peptides are named as *N*-substituted C-terminal amino acids, using the suffix -yl for the amino acid substituents; for example: glycyl, aspartyl and glutamyl - for aspartic and glutamic acids, asparaginyl and glutaminyl - for amides. Thus, Asp-Ser-Lys is the shortened formula for the tripeptide aspartylseryllysine (which is written without hyphens):



A sequence of amino acids in the chain signifies the primary structure of pep

tides and proteins.

Problem 15.5. Draw the structural formulas for each of the following peptides: (a) glutaminylserine; (b) alanylphenylalanine; (c) Glu-His-Gly. Give the name of the latter peptide.

15.2.2. Secondary Structure

Proteins and long-chain peptides might be expected to have rather amorphous, or floppy structure. But many peptides and some proteins have been isolated in crystalline form that indicates well-defined shapes of their molecules. The shapes seem to be quite regular even in solution. To understand this, let us consider some structural features of peptide chains.

Examination of a peptide group geometry (Fig. 15.1, a) shows that the amide C-N bond (132 pm long) is much shorter than the usual C-N single bond (147 pm; compare with 127 pm for the C=N double bond), while the C=O bond is slightly longer (124 pm) than that of carbonyl compounds (121 pm). All bond angles around the nitrogen atom are practically 120°, which is typical of sp^2hybridization. All the data are the result of ρ,π conjugation or resonance in the amide group (Fig.15.1, *b* and c).

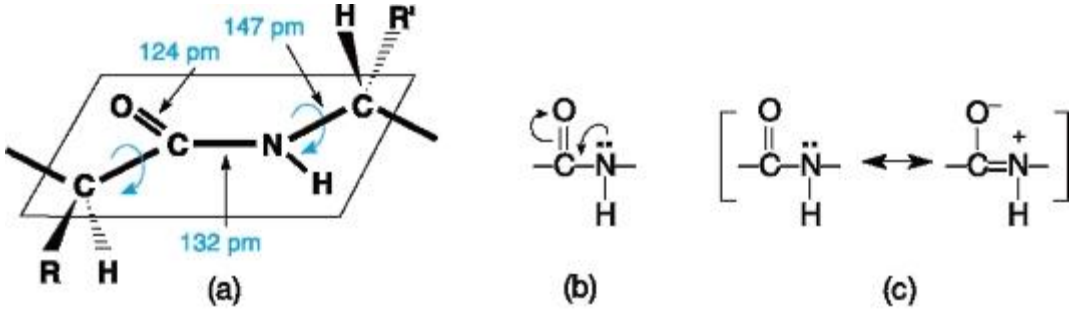


Figure 15.1. Peptide group: planar structure (a); *ρ,π* conjugation (b); resonance hybrid (c).

As a consequence, the C-N bond in the peptide group has a considerable double bond character and the peptide group is *planar*. But two adjacent peptide groups are not coplanar because of rotation about the other single bonds in the chain, i. e. the C-C and N-C bonds (Fig. 15.2). Thus, a peptide chain represents the series of planar portions divided by «joints», which are -CHRfragments.

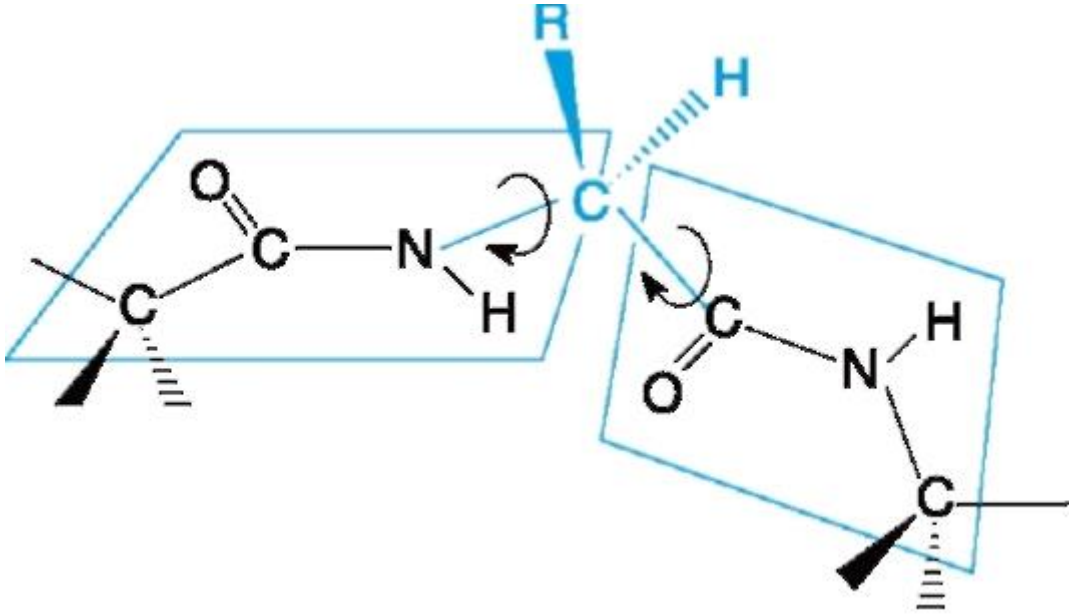
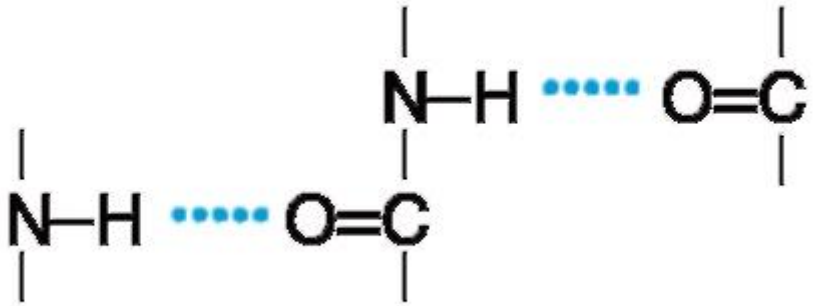


Figure 15.2. Planar portions and a -CHRjoint of the peptide chain.

Such geometry and restricted rotation of the peptide bond govern a definite shape of proteins. The actual proteins are more often coiled or folded to give helical or compact, often globular, molecules. The binding forces that give rise to the unique conformational characteristics of individual proteins are of several kinds.

Hydrogen bonding. The presence of the recurring -NHCOunit along the linear polypeptide chain provides for a structural organization common to the majority of proteins. The drawing represents parts of the same chain, or parts of separate chains. Note that the acidic site is the NH fragment and the basic site is the oxygen



of the carbonyl group.

The most distinctive structural consequence of hydrogen bonding is a coiling of the peptide chain about itself into a structure like a spiral staircase, which is called an *α-helix* (Fig. 15.3). The helix has a pitch of 0.54 nm, or about 3.6 amino acid units, and is maintained by hydrogen bonds that are approximately parallel with the axis of the helix. The bulky R groups of the amino acid residues are arranged outward, thus avoiding steric interaction between such R groups.

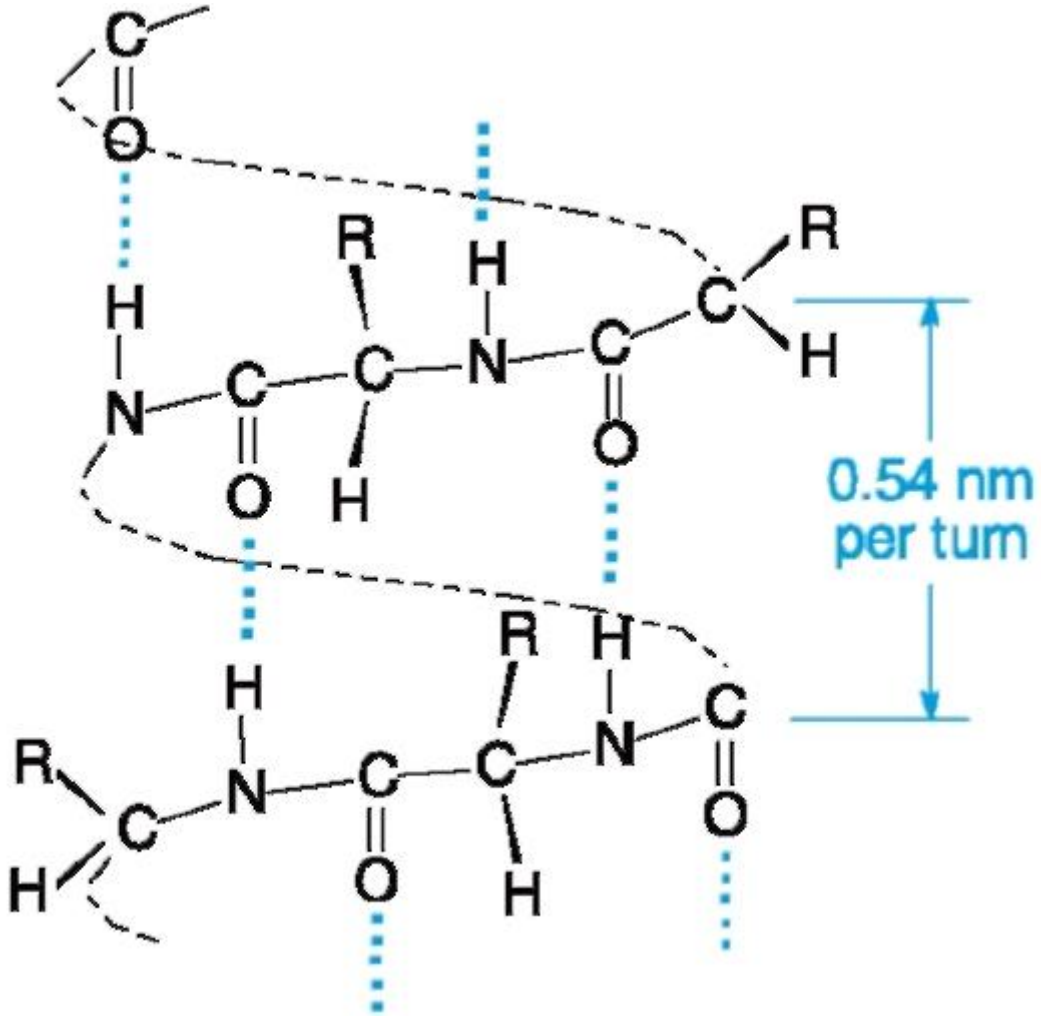


Figure 15.3. Hydrogen bonds (coloured dotted lines) in a segment of a polypeptide α-helix.

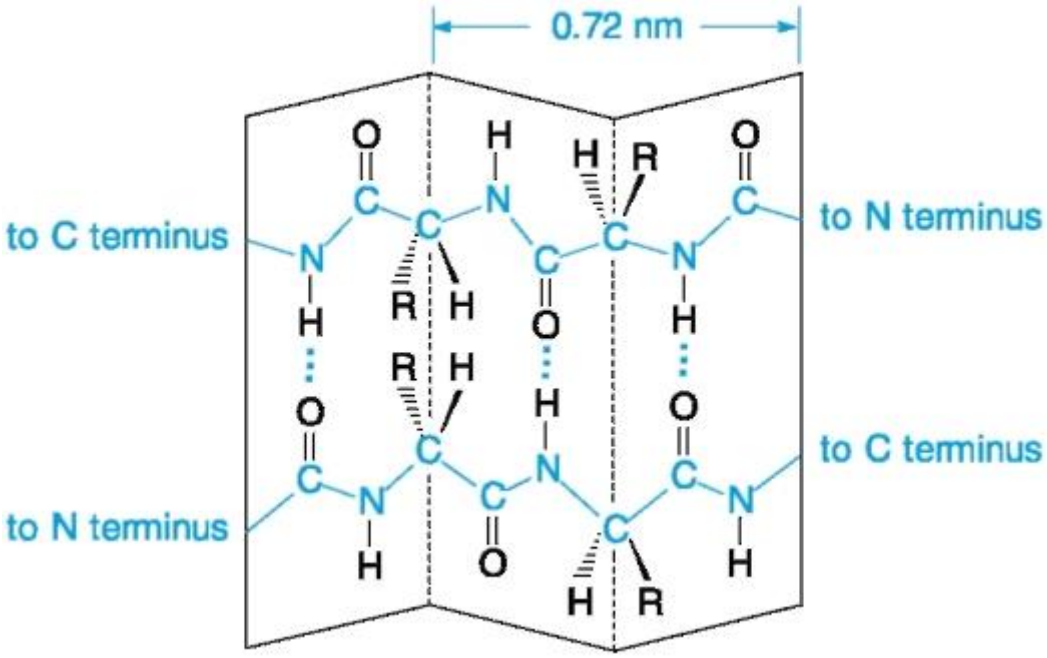


Figure 15.4. A segment of a pleated sheet structure (hydrogen bonds and peptide core are shown in colour).

Another type of hydrogen bonding is realized in a *pleated sheet* arrangement of the peptide chain (Fig. 15.4). In this case, some portions of the peptide chain lie side by side in opposite directions and form*interchain* hydrogen bonds. Such an arrangement is possible only in proteins having a high content of amino acid with small R groups. Otherwise, there will be appreciable steric repulsion between the R groups on adjacent chains. The pleated sheet structure is found in the structural protein β-keratin, obtained from silk fibroin, in which about 50% of the amino acid units are glycine (R = H) and over 20% are alanine (R = CH_3).

Of two spatial organizations of the peptide chains, helical and pleated, the α helix is a more common structure for the above reason.

The secondary structure of proteins is defined by hydrogen bonding

between peptide groups of the chain.

15.2.3. Tertiary Structure

So far we have considered the protein backbone regardless of the nature of their constituents. Different R groups of amino acids also affect the mode of spatial arrangement of proteins. The tertiary structure of proteins is their overall three-dimensional shape that arises from further folding of the polypeptide chain.

Disulfide linkages. Oxidative coupling of two SH groups of cysteine results in a disulfide bond (-S-S-) formation. This transformation can join two portions of a protein chain to form a loop or ring (Fig. 15.5), or can join two polypeptide chains together. The disulfide bond can be broken with reducing agents.

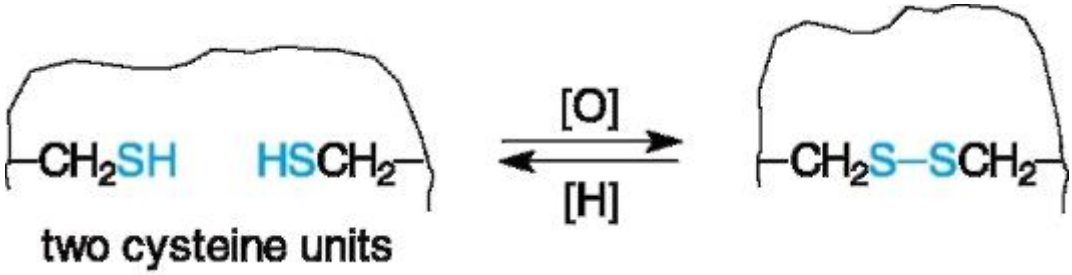


Figure 15.5. Disulfide bond formation and cleavage.

Dipolar interactions. Ionized amino and carboxyl groups, suitably disposed along the polypeptide chain, can give rise to electrostatic attraction between different segments of one chain or between separate chains (Fig. 15.6). The most ionogenic groups are the extra carboxyl groups of aspartic and glutamic acids that exist in a deprotonated form (-COO^-) at physiological pH's, whereas the basic functional groups of lysine and arginine are protonated in the same medium.

Figure 15.6. Electrostatic interaction

Figure 15.7. Hydrophobic interaction between ionized groups. between phenylalanine residues

Hydrophobic interactions. Attractive van der Waals' forces arise between nonpolar R groups of globular[1] proteins. These proteins tend to expose their polar (hydrophilic) groups to the aqueous environment and to dispose a maximum of nonpolar (hydrophobic) groups within the molecule (Fig. 15.7). As it follows from Table 15.1, phenylalanine and amino acids with alkyl R groups are responsible for hydrophobic interaction.

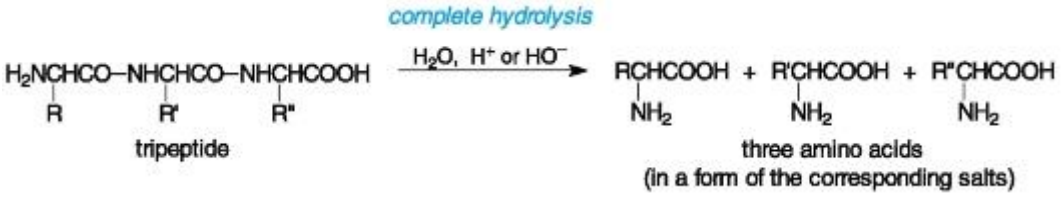
In concluding consideration of spatial organization of polypeptide chains, it should be noticed that the shape of the chain depends strongly on both amino acid content and amino acid sequence of a macromolecule.

[1] Proteins are classified as *globular* and *fibrous* ones according to their shapes. The former are of roughly spherical shape and are generally soluble in water. The most important representatives of these are enzymes. Fibrous proteins have a roughly linear shape and are water-soluble. They represent a constructive material of skin, hair, muscle, connective tissue, silk fibroin, etc.

15.2.4. Hydrolysis of Peptides and Proteins

Of all the chemical properties of peptides and proteins, in this section we shall only touch upon complete and partial hydrolyses. The former leads to amino acid components of the peptide, the latter - to a mixture of shortened peptides. Complete hydrolysis is a way to estimating amino acid content of peptides and proteins.

As has already been shown, amides undergo hydrolysis in an acidic or alkaline medium. Similarly, peptides can be hydrolyzed according to the general equation:



The alkaline hydrolysis is rarely applied in the peptide chemistry because of instability of some amino acids on heating with strong alkalis. The acid-catalyzed hydrolysis is preferably used[1] but it is not free of shortcomings too. They consist in the complete destroying tryptophan molecule (see Sec. 16.2.1) and the loss of information about asparagine and glutamine that are also hydrolyzed into the corresponding acids.

In living systems, proteins are hydrolyzed enzymically. That problem is beyond the scope of this book.

Additional Problems

15.6. Name the following amino acids by the substitutive IUPAC nomenclature:

(a) alanine;

(e) methionine;

(b) aspartic acid;

(f) phenylalanine;

(c) leucine;

(g) proline;

(d) lysine;

(h) tyrosine.

[1] Typical conditions for hydrolysis involve heating of a sample with 6 M HCl at 110 °C for 24 hours in the absence of oxygen.

15.7. Illustrate by equations the acid-base behaviour of the amino acids (a) - (c) in reactions with excess sodium hydroxide and of the amino acids (d) - (f) with excess hydrochloric acid:

(a) lysine;

(d) proline;

(b) aspartic acid;

(e) serine;

(c) cysteine;

(f) tyrosine.

15.8. Write equations for reactions of glutamine with:

(a) equimolar amount of hydrochloric acid;

(b) excess hydrochloric acid at room temperature;

(c) excess hydrochloric acid on heating.

Name the products obtained.

15.9. Complete the following equations:

(a) serine + excess NaOH →

(b) lysine + excess acetic anhydride →

(c) valine + ethanol/HCl → Name the products obtained.

15.10. Complete the following equations:

(a) glycine methyl ester + aqueous NaOH →

(b) lysine + methanol/HCl →

(c) tyrosine + HNO_3→

(d) valine + formaldehyde → Name the products obtained.

15.11. Predict appropriate structures of amino acids that on enzymic decarboxylation produce the following compounds:

(a) colamine (2-aminoethanol);

(b) 4-aminobutyric acid.

Write equations for these reactions.

15.12. Draw the structural formulas for each of the following peptides:

(a) cysteinyllysine;

(b) glycylproline;

(c) Asn-Ser-Met.

15.13. Write equations for acid-catalyzed (HCl) hydrolysis of the following peptides:

(a) phenylalanylvaline;

(b) Phe-Glu;

(c) Gln-Gly-Leu.

15.14. Write equations for base-catalyzed (NaOH) hydrolysis of the following peptides:

(a) glutamyltyrosine;

(b) Phe-Ile;

(c) His-Met-Asn.

15.15. Well-known sweetener *aspartame* (the trade name NutraSweet®) is the methyl ester of the dipeptide L-aspartyl-L-phenylalanine. Write the structural and the Fischer projection formulas for it.

Chapter 16. BIOLOGICALLY ACTIVE HETEROCYCLIC COMPOUNDS

Heterocycles form the largest group of organic compounds, and many have important biological properties. Most natural substances as well as synthetic and natural drugs contain heterocyclic fragments. More precisely, about two thirds of all organic compounds and over three quarters of drugs belong to heterocyclic compounds.

The main heteroatom that occurs in heterocyclic compounds is nitrogen, but there are also rings containing oxygen and sulfur.

In previous chapters we have encountered some compounds with the oxygen or nitrogen atom in a cycle. Formally, cyclic anhydrides, lactones, lactides, and lactams are heterocycles. But the mentioned compounds have the same chemistry as their open-chain counterparts: lactones and ordinary esters behave similarly, as well as lactams and ordinary amides, and so on. These cyclic compounds undergo

many reactions with ring opening that is not typical of «genuine» heterocycles, therefore they do not relate to heterocyclic compounds.

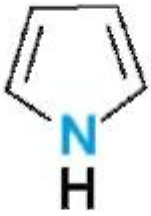
16.1. GENERAL CHARACTERISTICS OF HETEROCYCLIC SYSTEMS

16.1.1. Classification

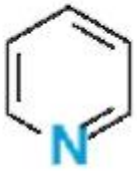
There is no unified classification of heterocyclic compounds because of their big variety (tens of thousands of types are known at present). They are usually classified in accordance with the following features of their skeleton.

According to the ring size. There are mainly three-, four-, five-, six-, and sevenmembered heterocycles. Out of them, fiveand six-membered are the most widespread.

aziridine



pyrrole



pyridine

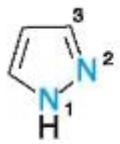
According to the heteroatom incorporated in a ring. As has already been said, the most important are nitrogen-, oxygen-, and sulfur-containing heterocycles.

furan

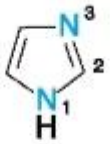
thiophene

pyrrolidine

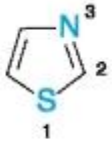
According to the number of heteroatoms and their mutual arrangement in a ring. The most common are compounds with one or two heteroatoms, but cycles with more heteroatoms are also known. Various combinations of heteroatoms are possible (for example, two nitrogens, nitrogen and oxygen, etc.), and heteroatoms can occupy the 1 and 2, 1 and 3, and 1 and 4 positions.



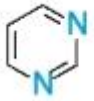
pyrazole



imidazole



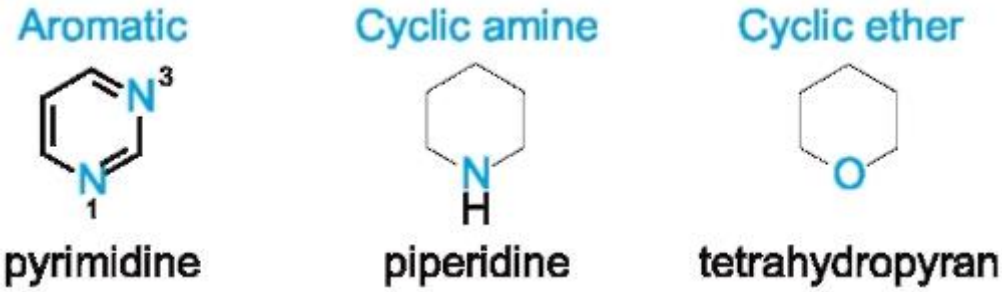
thiazole



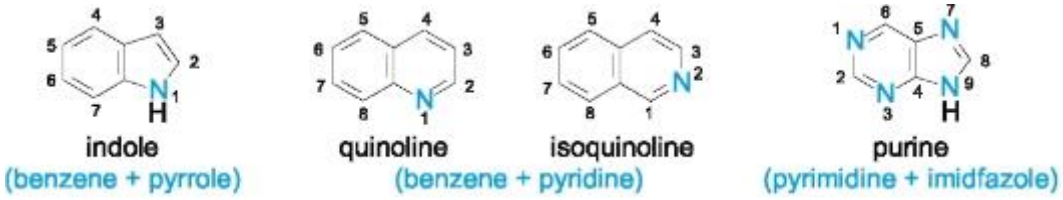
pyrimidine

piperazine

According to a degree of unsaturation of a ring. Heterocycles can be divided into two subgroups: aromatic and nonaromatic (fully or partly saturated). The former are much more important, therefore the main attention in this chapter will be devoted to them. Besides, fully saturated heterocycles are similar to their acyclic analogues, i. e. secondary amines, ethers, or sulfides.



According to the number of rings. In addition to monocyclic compounds, heterocycles with fused rings are well known too, carbocycles also being present in a fusedring system.

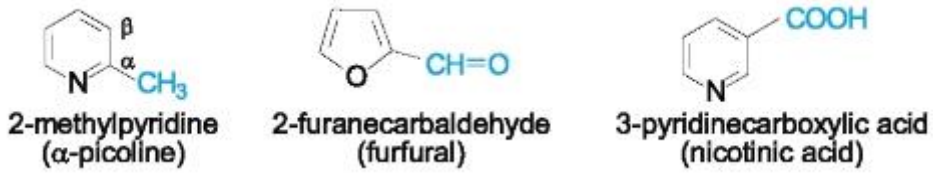


16.1.2. Nomenclature

Systematic nomenclature of heterocycles was developed at the end of the 19th century. It is known as*Hantzsch-Widman system* (see below). Nevertheless, several decades of trivial and semi-trivial names of heterocycles are widely used according to the IUPAC recommendations. Many examples were given in Sec. 16.1.1.

The position of a single heteroatom determines the numbering in monocyclic compound (but not in fused heterocycles). When the same heteroatom occurs twice or more, the numbering is chosen to give the lowest locants to the heteroatoms (see pyrimidine). In compounds with two nitrogens having a different configuration (see pyrazole and imidazole) the lowest locant gets the -NHfragment. When different heteroatoms are present, the priority of atoms in numbering is O > S > N (see thiazole).

The numbering in fused bicyclic compounds begins from the atom next to rings fusion in the direction to a heteroatom (see quinoline and isoquinoline). However, there are many exception to this rule, as, for example, in purine that retains historical numbering.
The general principles of substitutive nomenclature should be applied (Sec. 2.2) in systematic names of heterocycle derivatives, for example (trivial names are given in parentheses):



Hantzsch-Widman system. Monocyclic compounds containing one or more heteroatoms in a threeto ten-membered ring are named by combining the

appropriate *prefix* (or prefixes) that denotes a heteroatom with a *stem* that indicates a ring size and its unsaturation (Table 16.1).

For the most widespread heteroatoms the following prefixes are used: aza- for nitrogen, oxa- for oxygen, and thia- for sulfur (eliding terminal a where necessary for euphony). Thus, monocyclic heterocycles shown above will be named as follows: *oxole* (furan), *thiole* (thiophene), *azole* (pyrrol), *1,3-diazole* (imidazole), and *azine* (pyridine).

Table 16.1. The stems for designation of a ring size in the Hantzsch-Widman system

Number of atoms in the ring	Rings containing nitrogen		Rings containing no nitrogen	
	unsaturated*	saturated	unsaturated*	saturated
3	-irine	-iridine	-irene	-irane
4	-ete	-etidine	-ete	-etane
5	-ole	-olidine	-ole	-olane
6	-ine	**	-in	-ane
7	-epine	**	-epin	-epane
8	-ocine	**	-ocin	-ocane

* Corresponding to the maximum number of non-cumulative double bonds.

** Expressed by addition of the prefix perhydroto the name of the respective unsaturated compound.

16.2. GENERAL ASPECTS OF REACTIVITY OF AROMATIC HETEROCYCLES

Two heterocycles, pyridine and pyrrole, are of fundamental importance for a reason that will be considered in the next section.

16.2.1. Aromaticity of Pyridine and Pyrrole

Pyridine. This six-membered heterocycle has a structure similar to that of benzene, except that one CH fragment of the benzene ring is replaced by a nitrogen atom (Fig. 16.1, a).

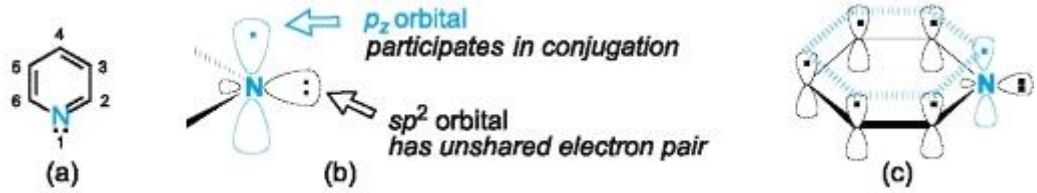


Figure 16.1. The pyridine molecule: the structure (a), the pyridine-type nitrogen atom (b), and the orbital picture (c). The orbital overlap below the ring and all hydrogen atoms are omitted for clarity.

Nitrogen in pyridine is sp^2-hybridized (as well as all carbons). The two hybrid orbitals of nitrogen, each with one electron, overlap sp^2 orbitals of the C-2 and C-6 atoms to form two σ bonds (these orbitals are not shown in Fig. 16.1, b). The remaining hybrid orbital of nitrogen possesses an unshared electron pair and does not form a bond. The unhybridized *p* orbital of nitrogen (with one electron) is

perpendicular to the plane of the ring and overlaps the *p* orbitals of carbons to form aromatic six π-electron cloud (Fig. 16.1, c). Otherwise, pyridine is very like benzene in its π-electron configuration, i. e. pyridine is*isoelectronic* to benzene. Thus, pyridine is in accord with all criteria of aromaticity:

- all atoms in the cycle are sp^2-hybridized,
- it represents a planar compound with cyclic conjugation,
- it obeys the Huckel's rule because it is a six π-electron system.

The conjugation energy in pyridine is sufficiently high (117 kJ/mol) but markedly less than that of benzene (151 kJ/mol). The difference is caused by the presence of the electronegative nitrogen atom that distorts complete uniformity of the π-electron cloud. This results in decreasing electronic density on the carbon atoms. Pyridine is referred to as *π-deficient* heterocycles for this reason.

Pyrrole. Five-membered heterocycle pyrrole contains all the carbons and the nitrogen atom in sp^2-hybridized state (Fig. 16.2, *a).*

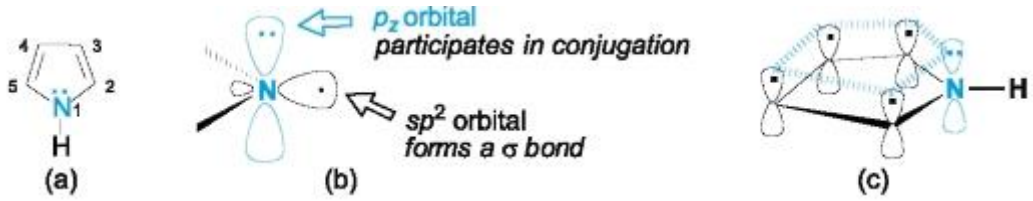


Figure 16.2. The pyrrole molecule: the structure (a), the pyrrole-type nitrogen atom (b), and the orbital picture (c). The orbital overlap below the ring and all C-H bonds are omitted for clarity.

All three hybrid orbitals on nitrogen form three σ bonds, two orbitals (not shown in Fig. 16.2, b) - with the atoms C-2 and C-5 and the third orbital - with hydrogen. The nitrogen lone pair of electrons occupies unhybridized *p* orbital and forms, together with four *p* orbitals on carbons, a delocalized π-electron cloud (Fig. 16.2, c). Pyrrole is therefore an aromatic compound.

In contrast to pyridine, pyrrole is said to be a *π-excessive* (or π-rich) heterocycle because *six π* electrons of the aromatic system spread over *five* atoms of the cycle. In other words, each *p* orbital contains more than one electron.
Finally, pay attention to different roles of the nitrogen atoms in pyridine and pyrrole. Nitrogen in pyridine contributes only one π electron to the aromatic sextet, while nitrogen in pyrrole contributes two *p* electrons (the lone pair) to the aromatic sextet.

Example 16.1. Prove that imidazole is an aromatic heterocycle. Draw its orbital picture, showing all *p*orbitals and all lone-pairs of electrons.

Solution. The atom N-1 is pyrrole-type nitrogen that has a lone pair of electrons on its unhybridized *p*orbital (Fig. 16.3). The atom N-3 is pyridine-type nitrogen with one electron on a *p* orbital.

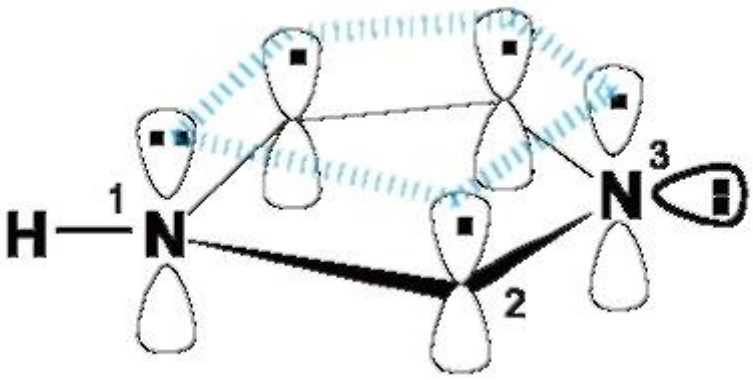


Figure 16.3. Orbital picture of imidazole. The orbital overlap below the ring and all C-H bonds are omitted for clarity.

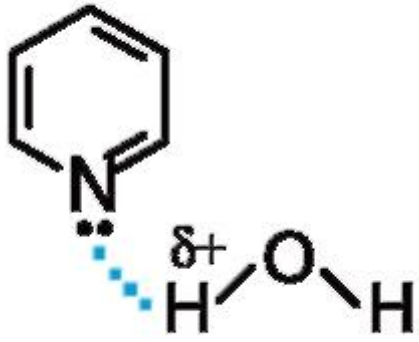
As a result of such difference in the nitrogen electron demand, pyridine and pyrrole behave in a different manner with respect to various chemical reagents.

Problem 16.1. Which of the following compounds represent aromatic heterocycles: (a) furan; (b) piperidine; (c) pyrimidine; (d) quinoline. Explain the reason for your choice.

16.2.2. Basicity and Acidity

In pyridine, the lone pair of electrons on the nitrogen atom is *not* a part of the aromatic π-electron system but occupies an sp^2 orbital in the ring plane (Fig. 16.1). Consequently, pyridine can donate this lone pair of electrons for accepting a proton and for hydrogen bonding with water. Unlike benzene, it is completely miscible with water.

Pyridine forms pyridinium salts in reaction with acids. It is a weakly basic compound, with pK_{BH^+} 5.2, which is comparable with the value of 4.6 for aniline. Pyridine is however much less basic than ammonia and aliphatic amines because the electronegativity of an sp^2-hybridized nitrogen is greater than that of an sp^3-hybridized nitrogen in ammonia and in aliphatic amines.



pyridinium chloride

In pyrrole on the contrary, the nitrogen lone pair of electrons is an essential part of the aromatic π system. Pyrrole is an extremely weak base; it has a pK_{BH^+} of -3.8, about 10^9 times weaker than pyridine. It can be protonated on carbon rather than on the nitrogen atom. Protonation of pyrrole would destroy the aromatic system, therefore it belongs to so-called *acidophobic* heterocycles. At the same time, pyrrole possesses a marked NH-acidity (pK_a 17.5 that is in the range of alcohol acidity). Thus, it can be deprotonated like alcohols in reactions with strong bases, for example, with sodium hydride:

pyrrole sodium pyrrolate

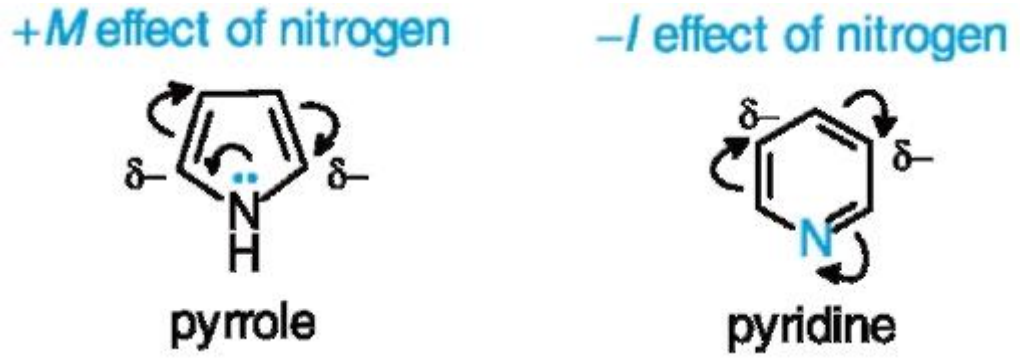
Note that both pyridinium and pyrrolate salts retain aromatic character.

Problem 16.2. Only one of two nitrogens in pyrazole is quite basic. Show which one it is, and explain the reason for your choice.

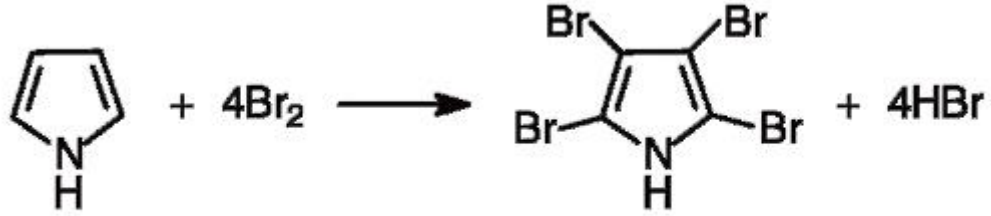
16.2.3. Substitution Reactions in Heterocycles

Electrophilic substitution is a typical reaction not only for aromatic hydrocarbons but also for aromatic heterocycles. The influence of a heteroatom on the reactivity of a heterocycle is similar to that of a substituent in the benzene ring.

In pyrrole, the donating effect of the nitrogen lone pair of electrons increases the electron density on the ring carbons, thus increasing reactivity of the heterocycle towards electrophiles. In pyridine, on the contrary, the pyridine-type nitrogen is an electron-withdrawing atom which decreases the electron density on the carbons (especially at the 2, 4, and 6 positions), thus making the heterocycle less reactive in electrophilic substitution reactions.

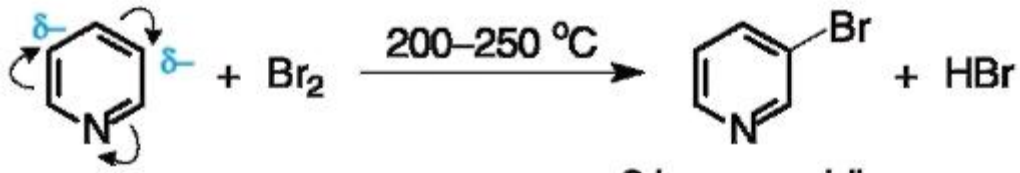


A difference in reactivity of pyrrole and pyridine is demonstrated in their bromination reactions. Pyrrole reacts readily with bromine at low temperature to yield the corresponding tetrabromo derivative:



2,3,4,5-tetrabromopyrrole

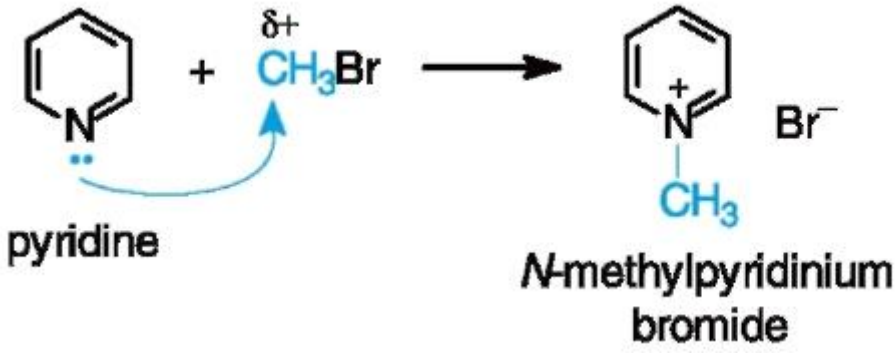
Bromination of pyridine can be carried out under drastic conditions in a low yield where a free-radical mechanism may operate. Electrophilic substitution nearly always takes place at the 3 position according to the bond polarization.



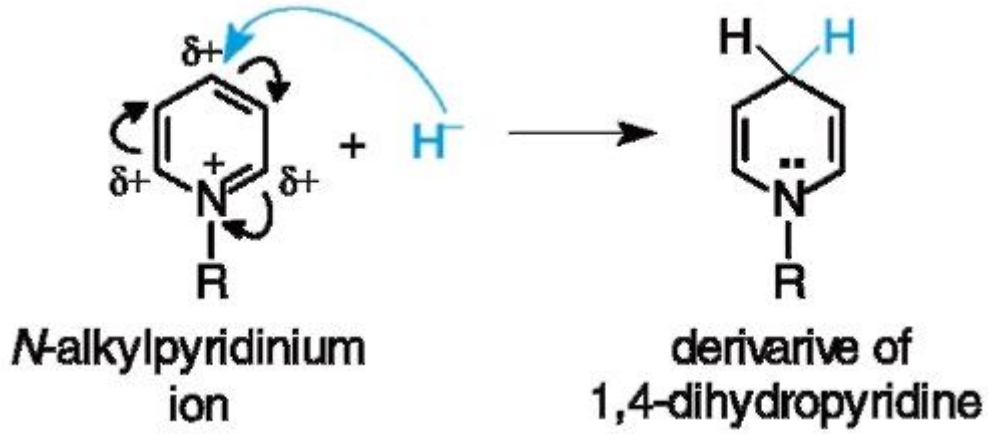
3-bromopyridine

Another factor decreasing the reactivity of pyridine in electrophilic substitution reactions is acid-base complexation between the basic nitrogen and the attacking electrophile, which may be often a proton. This results in the formation of a positive charge on the ring, further deactivating it. Thus, pyridine does not undergo FriedelCrafts alkylation and acylation.

Pyridine exhibits nucleophilic properties in the reaction with electrophiles. It behaves as tertiary amine and forms quaternary pyridinium salts when reacting with alkyl halides.



The aromatic ring of pyridinium salts is susceptible to nucleophilic attack owing to its high electron-deficiency. For example, a strongly nucleophilic hydride ion reacts with an alkylpyridinium salt by addition, producing non-aromatic 1,4-dihydropyridine derivative (attack of the 2 position is also possible):



The pyridinium ion is thus reduced, i. e. it acts as an oxidant. This ion is a part of the NAD+ molecule (for details see Sec. 17.4.2), one of the most important coenzymes in biological oxidations. So the above equation represents a simplified version of the biological reaction.

16.3. FIVE-MEMBERED RINGS WITH ONE NITROGEN

The pyrrole ring is a structural component of several biologically important compounds. It is often found as a fused-ring system indole. The latter is usually biosynthesized from the protein amino acidtryptophan. Indole itself and its 3-methyl derivative, skatole, are the decay products of proteins (both contribute to the odour of feces).

Tryptophan is decarboxylated *in vivo* to give tryptamine. Many compounds that contain the tryptamine skeleton have an effect on the brain and nervous system. For example, serotonin (5-hydroxytryptamine) is a neurotransmitter and vasoconstrictor active in the central nervous system. A disturbance in its metabolism leads to schizophrenia.

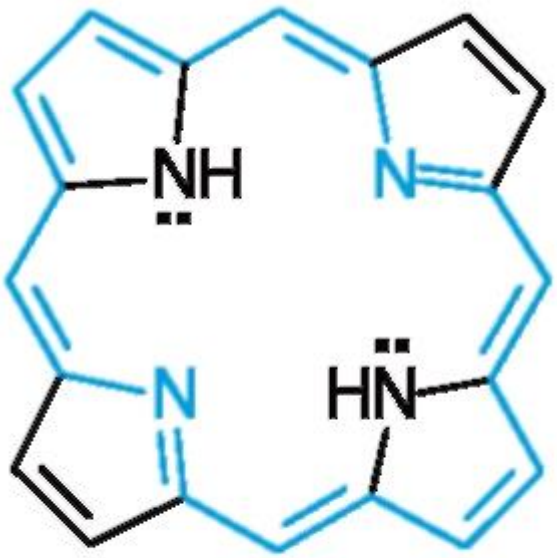
The structure of saturated pyrrole called pyrrolidine is a fragment of the wellknown alkaloid nicotine (Sec. 16.6.1).

The pyrrole ring is a structural component of many vitally important compounds called *porphyrins.* Their parent structure is porphine, a tetrapyrrolic macrocyclic system. Porphine represents a flat symmetrical molecule in which four pyrrole rings are linked by one-carbon bridges. It forms a conjugated system of eighteen π electrons shown in colour and is, therefore, aromatic.

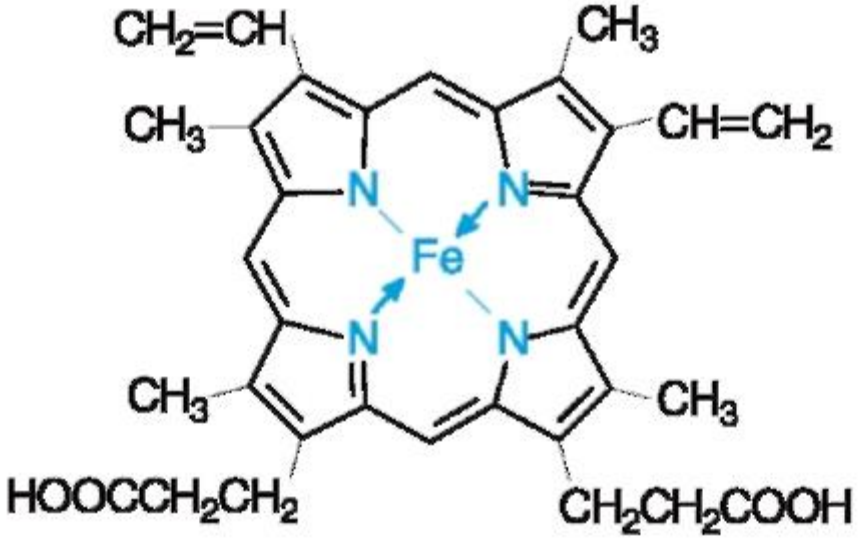
Note that 18 is the Huckel's number when $n = 4$. The formula of porphine represents one of the resonance contributing structures. Total amount of electrons in conjugation is 26, it also being the Huckel's number.

Aromaticity of porphine is confirmed by its high conjugation energy which amounts to 840 kJ/mol. All the porphyrins (substituted porphines with various side chains at the pyrrole rings) are exceptionally stable compounds that decompose at about 500 °C.

Porphyrins form metallic complexes in which two NH hydrogens are absent and each of the four nitrogens is bound to metal in the middle of the structure. The best known of these is heme, the iron(II)-porphyrin complex that imparts the red colour to blood. Heme is a constituent of the complex protein hemoglobin responsible for binding molecular oxygen in the process of respiration.



porphine



heme

Another example of tetrapyrrolic compounds is chlorophyll, the green plant pigment essential for photosynthesis. It represents a magnesium-porphyrin complex.

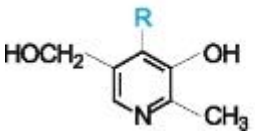
16.4. SIX-MEMBERED RINGS WITH ONE HETEROATOM

This group of heterocyclic compounds is the oldest one. Pyridine was discovered in the mid-19th century and its structure was established in 1869, shortly after the Kekule structure of benzene had been suggested.

16.4.1. Nitrogen-Containing Heterocycles

The main parent compounds of this type are pyridine and fused-ring heterocyclic systems of quinoline and isoquinoline. They are found in a small quantity in coal tar. Pyridine is a toxic liquid with unpleasant odour.

All these rings are constituents of many naturally occurring compounds and numerous drugs. Some physiologically important pyridine derivatives are related to vitamin B_6. This is a relatively simple pyridine derivative. The R group at C-4 may be a fragment of an alcohol, aldehyde, or amine. The vitamin functions as a coenzyme in the interconversions of oxo carboxylic acids and amino acids (Sec. 15.1.3).



vitamin B_6: R = CH_2OH, pyridoxin
R = CH=O, pyridoxal and its phosphate
R = CH_2NH_2, pyridoxamin

Problem 16.3. Which of the following statements corresponds to the structure and properties of pyridoxal?

(a) the unshared electrons of the nitrogen are involved in conjugation to form aromatic sextet;

(b) it contains the pyrrole ring system;

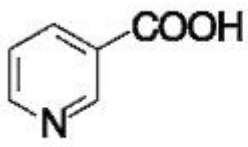
(c) it forms a sal t with hydrochloric acid;

(d) it forms a salt with potassium hydroxide, disposing both hydroxyl groups;

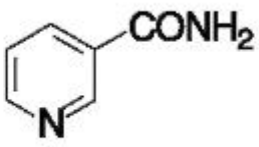
(e) it reacts with primary amines to give a Schiff's base.

Write equations for reactions, if they take place.

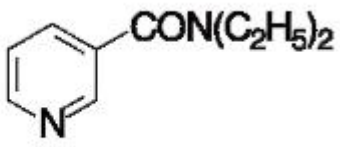
Two pyridine derivatives which relate to vitamin PP are nicotinic acid (Niacin) and its amide (Niacinamide). A substituted nicotinamide Cordiamine (Niacetamide) is used as a stimulator of the central nervous system.



nicotinic acid

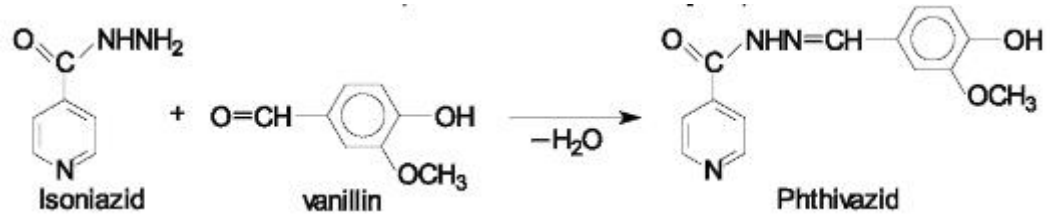


nicotinamide

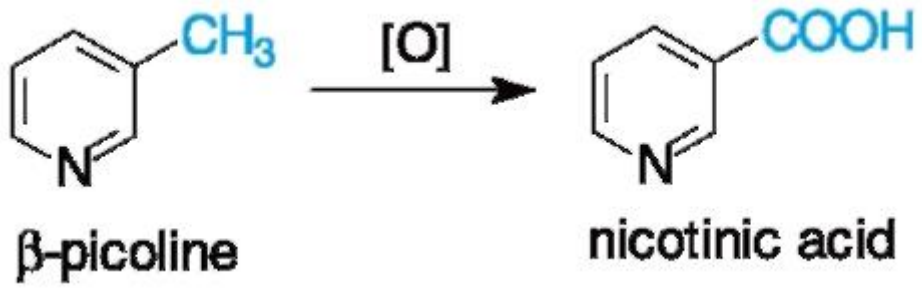


Cordiamine

Other examples of pyridine containing drugs are hydrazide derivatives of isonicotinic acid (4-pyridinecarboxylic acid). They are known as tuberculostatic drugs Isoniazid (Tubazid) and Phthivazid. The latter is prepared from Isoniazide by a familiar nucleophilic reaction with vanillin (an aromatic aldehyde) as shown below:



Nicotinic acid first obtained by oxidation of nicotine is now produced from available β-picoline (3-methylpyridine) and other 3-alkylpyridines also by oxidation, for example:

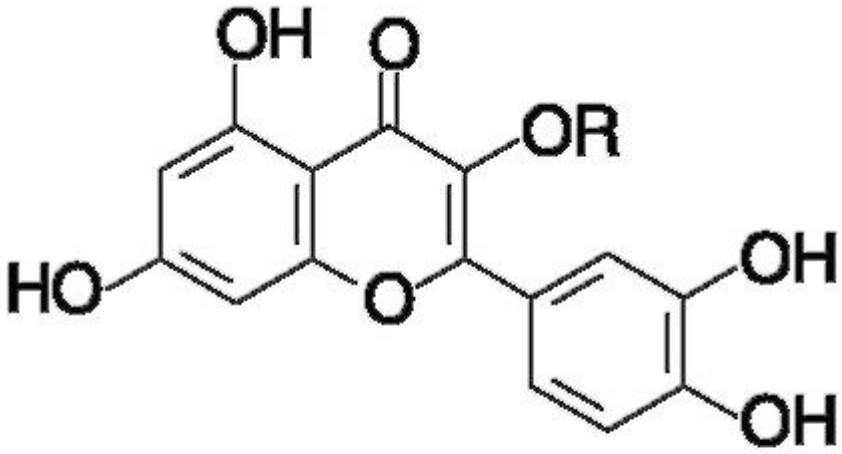


16.4.2. Oxygen-Containing Heterocycles

Compounds consisting of oxygen-containing heterocycles are widespread in nature; there are, first of all, cyclic forms of carbohydrates (Chapter 14). Unsaturated six-membered representatives - 2/ -pyran and 4/ -pyran[1] - are inherently unstable; moreover the former is at present unknown. Nevertheless, a group of the naturally occurring pyran derivatives called *flavonoids* has attracted attention of chemists and biochemists in the last decades. Basic structures of flavonoids are polycyclic compounds flavan and its 4-oxo derivative flavanone, which may also contain the C-2-C-3 double bond (not shown below):

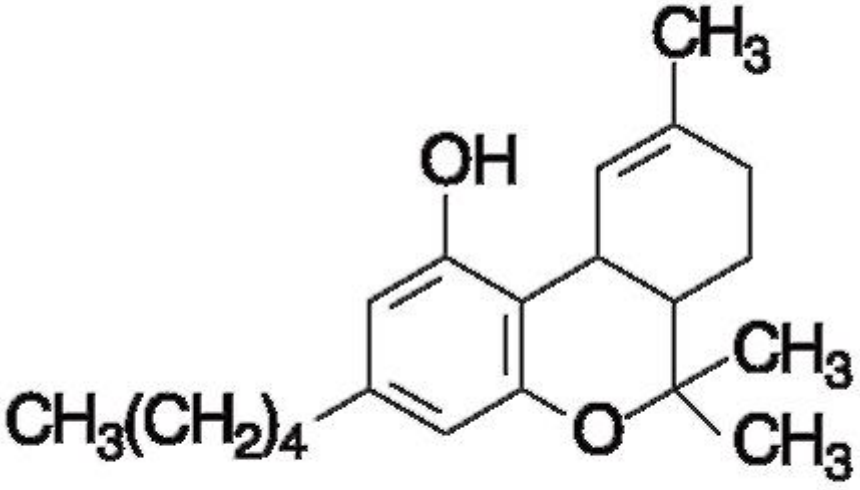
A characteristic feature of flavonoids is the presence of several hydroxyl groups (up to six) in all three rings, some of which are often bound to sugar units.

Flavonoids occur ubiquitously in plants and foods. It is estimated that the mean human intake of all flavonoids is about 100-150 mg per day. Many flavonoids possess multiple biological activities, including cardiovascular, anticarcinogenic, anti-inflammatory, immune-stimulating effects, some of them are used in medicine as, for example, quercetin (Quertin) and its sugar analogue rutin (Rutosid).



quercetin (R = H)

rutin (R = sugar unit)



tetrahydrocannabinol

It is known today that flavonoids take the protective role in the human body, saving the latter from destructive action of free radicals.

One compound containing a fused pyran ring system is of no benefit to humans; this istethahydrocannabinol, a psychotropic component of marijuana.

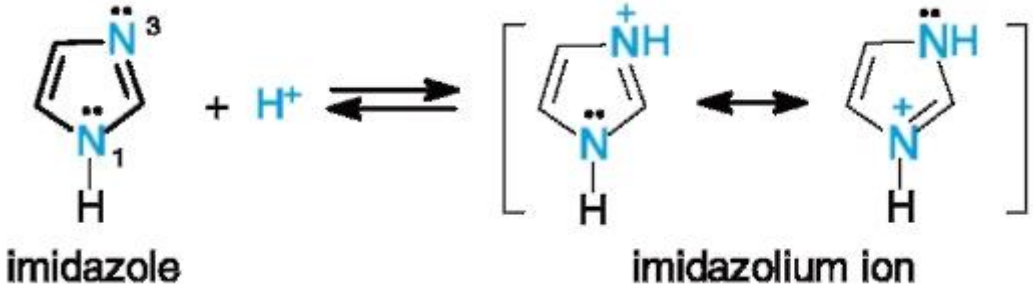
[1] The symbol H («indicated hydrogen») is used for accurate defining the position of the double bonds in compounds with the maximum number of non-cumulative double bonds.

16.5. RINGS WITH MORE THAN ONE HETEROATOM

Four important heterocyclic ring systems with more than one heteroatom are imidazole, pyrazole, pyrimidine, and purine. Other combinations of heteroatoms are known, of course, but we will be concerned with only nitrogen-containing aromatic heterocycles.

16.5.1. Imidazole and Pyrazole

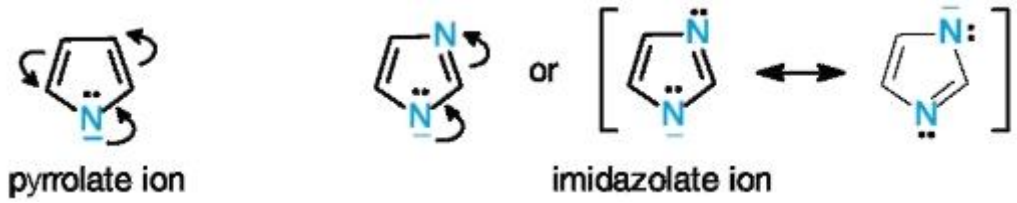
Imidazole is a little like pyrrole in the following way: the unshared electron pair on the N-1 atom is delocalized and is a part of the aromatic six π-electron system (Fig. 16.1). But the unshared electron pair on the N-3 atom is available for protonation. The $pK_{BH}+$ of imidazole is 7.0, so it is about 100 times more basic than pyridine and about 10^{11} times more basic than pyrrole. The positive charge in protonated imidazole can be delocalized over both nitrogens through resonance:



The similar consideration may be applied to pyrazole, which is an isomer of imidazole. The N-1 atom (pyrrole-type) represents the acidic site in both compounds whereas the C=N nitrogen (pyridine-type) is the basic site.

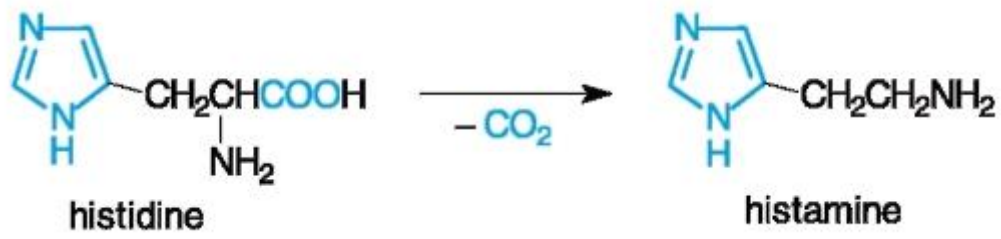
Example 16.2. As we have just seen, imidazole is much more basic than pyrrole. At the same time, imidazole is stronger as an acid too. Suggest an explanation for these facts.

Solution. It should be reemphasized that the best way for discussing acidity is to compare stability of conjugate bases, i. e. anions obtained after deprotonation of both heterocycles.

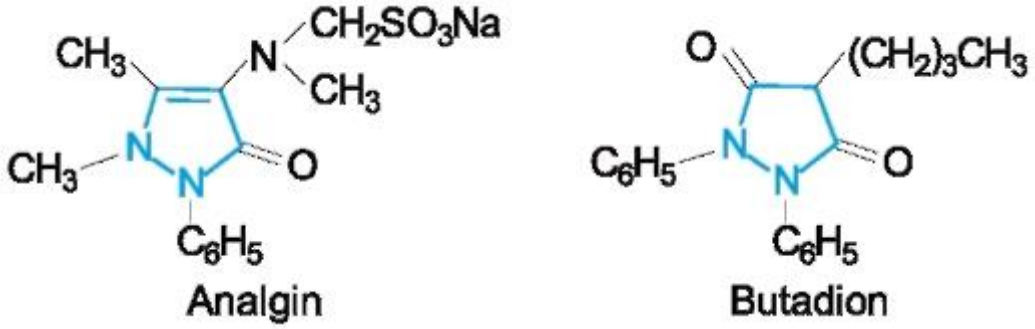


The imidazolate ion is better stabilized due to the presence of the second, electron-withdrawing nitrogen atom. In terms of the resonance theory, this anion represents two equal contributing structures. Indeed, the pK_a of imidazole is 14.2 and that of pyrrole is 17.5. Thus, imidazole is an amphoteric compound.

The imidazole unit is a part of the protein amino acid histidine. In some enzymic reactions, histidine situated on the active site of an enzyme can realize either acidic or basic catalysis; both are a consequence of amphoteric properties of imidazole. The imidazole derivative histamine that relates to biogenous amines is produced on decarboxylation of histidine in living systems.



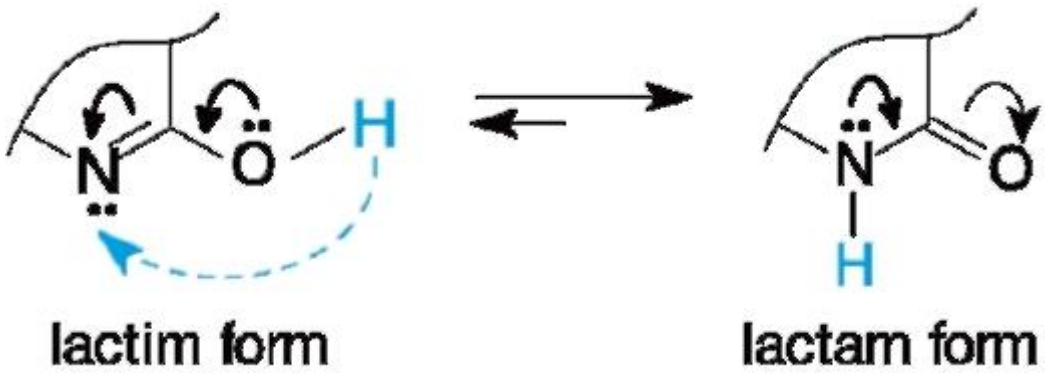
Pyrazole itself and its derivatives do not occur in nature, but the pyrazole skeleton is present in some analgesics and antipyretics, for example, in Analgin (Dipyrone) and Butadion (Phenylbutazone), that have many other synonyms.



16.5.2. Pyrimidine Derivatives

The most important pyrimidine derivatives are uracil, thymine, and cytosine called nucleic bases (or heterocyclic bases) since they are constituents of nucleic acids.

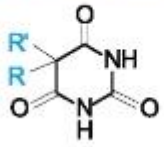
The nucleic bases are capable of existing in several tautomeric forms. A new kind of tautomerism arises for nitrogen-containing heterocycles that have the OH group attached to the C=N fragment of a ring system. This phenomenon is called *lactim-lactam tautomerism,* and tautomers are known as the *lactim form* and *lactam form.* The lactam tautomers are usually the predominant forms at equilibrium.



Problem 16.4. Draw the lactim tautomeric form for (a) uracil and (b) cytosine.

A group of compounds called *barbiturates*, whose usage ranges from mild sedatives to hypnotics and anesthetics, are also pyrimidine derivatives. Examples include Barbital (Veronal), the first synthetic soporific, Phenobarbital (Luminal), and Amobarbital (Amytal).

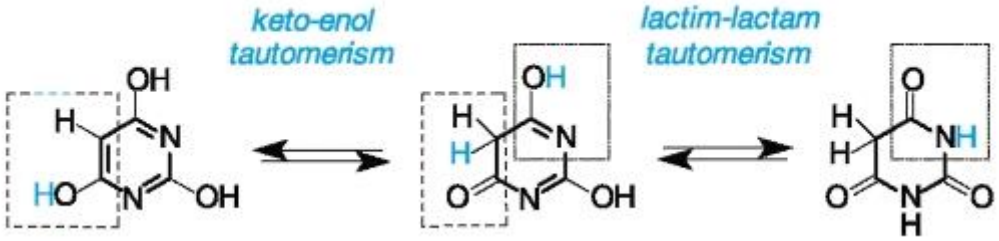
General formula of barbiturates



	R	R'
Barbituric acid	H	H
Barbital	C_2H_5	C_2H_5
Phenobarbital	C_2H_5	C_6H_5
Amobarbital	C_2H_5	$CH_2CH_2CH(CH_3)_2$

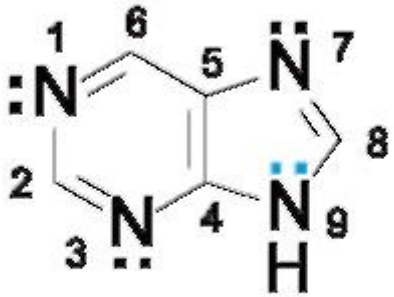
Barbituric acid exists in solution as a mixture of several tautomeric forms: the keto and enol forms (keto-enol tautomerism), and the lactim and lactam forms (not all the forms are given below):

TAUTOMERIC FORMS OF BARBITURIC ACID



16.5.3. Purine Derivatives

Purine represents a combination of two aromatic heterocycles imidazole and pyrimidine. Three of four nitrogens (N-1, N-3, and N-7) in the purine molecule are pyridine-type and the remaining (N-9) is a pyrrole-type nitrogen.

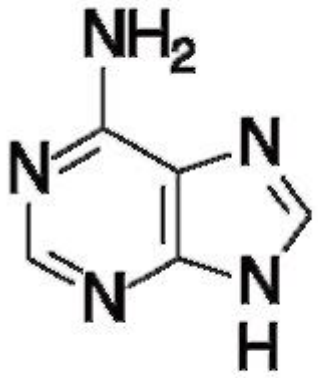


purine

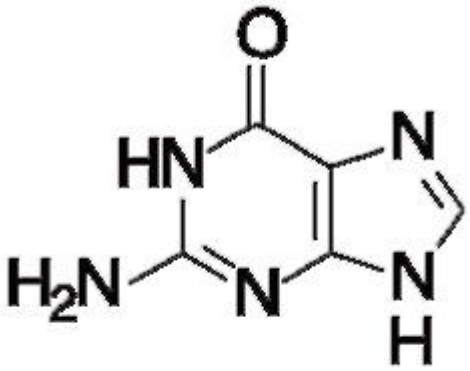
Purine is an aromatic compound because it has a cyclic system of conjugation with a *p* orbital on each atom and contains 10 π electrons (the Huckel's number). Each *p* orbital of five carbons and each of three pyridine-type nitrogens contributes one π electron to the aromatic system, and the pyrrole-type nitrogen contributes two *p* electrons.

Purine itself is not found in nature. Its two derivatives, adenine and guanine, represent purine components of nucleic acids in addition to three pyrimidine nucleic bases.

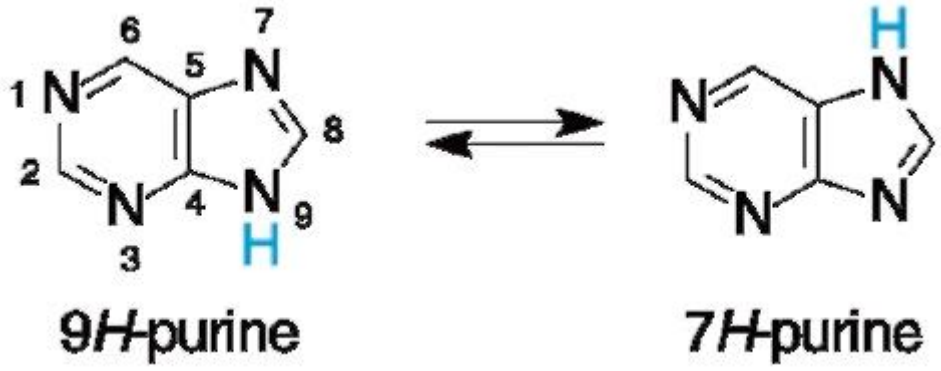
Purine and its derivatives are subjected to prototropic tautomerism caused by hydrogen migration between the positions N-7 and N-9.



adenine



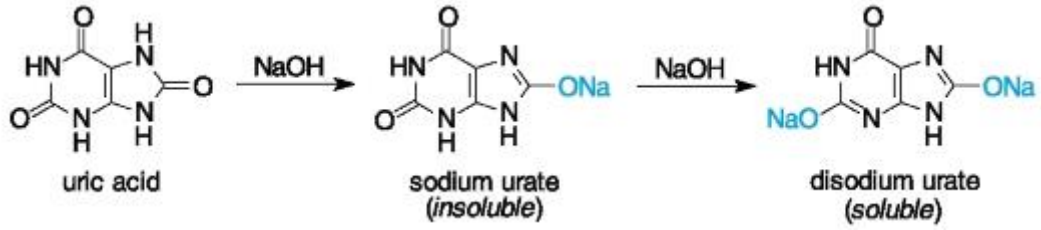
guanine



Crystalline purine is the 9/ tautomer as well as adenine and guanine incorporated into nucleic acids. Lactim-lactam tautomerism is possible for guanine.

Three hydroxylated purines are the products of nucleic acid metabolism. These are uric acid (2,6,8-trihydroxypurine), the final metabolite, xanthine (2,6-dihydroxypurine), and hypoxanthine (6-hydroxypurine), shown below in the most stable tautomeric forms:

Water-insoluble uric acid as a relatively strong acid (pK_{a1} 5.7, pK_{a2} 10.3) reacts with alkalis to form two series of salts called *urates,* for example:



Insoluble urates can be deposited in the joints and tendons as «stones» (or calculi) in some disorder in the human body.

Naturally occurring purine derivatives are N-methylated xanthines such as caffeine (present in coffee, tea, and cola beverages), theophylline (also present in coffee beans and tea leaves), and theobromine(found in cocoa).

Caffeine is known as a stimulator of the central nervous system; two other xanthines are also used in medicine. All the three compounds are sometimes assigned to alkaloids.

16.6. ALKALOIDS

Alkaloids are nitrogen-containing, mostly heterocyclic compounds that

produce striking physiological effects on animals.

Effects of alkaloids vary greatly from one compound to another. The term *alkaloid* originates from the fact that these substances are «alkali-like», i. e. they react with acids to form soluble salts. Moreover, alkaloids are present in plants as salts with organic acids (oxalic, malic, citric, and others).

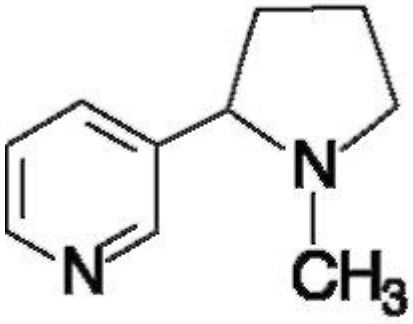
Common names are usually used for alkaloids because of the complexity of their structures. These names often reflect the botanical source of the substance. The alkaloid papaverine, for example, was isolated from the opium poppy, *Papaver somniferum;* the alkaloid cocaine - from *Erythroxylon coca,* etc. Sometimes the names of alkaloids are eccentric: the name of the another opium alkaloid morphine came from Morpheus, the Greek god of dreams; the name of the tobacco alkaloid nicotine came from J. Nicot, a French diplomat who brought tobacco seeds to France in 1560. Names of most alkaloids have the ending -ine that shows the amine nature of the compounds.

Over five thousand alkaloids are known today. The old classification of alkaloids used phylogenetic features, for example, tobacco alkaloids, coca alkaloids, and so on. The modern classification is based on the structure of a basic heterocyclic system though more than one heterocycle may be present.

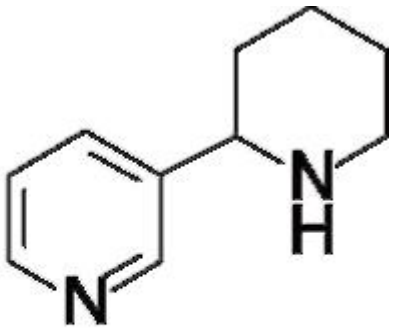
Only a few selected examples are considered in this section with minimal information on biological or pharmacological activity. Stereochemistry in all presented structures is omitted for simplicity but it should be remembered that almost all alkaloids are optically active compounds.

16.6.1. Pyridine Alkaloids

The major alkaloid of the tobacco leaves is nicotine, one of the simplest of all alkaloids, that contains, besides the pyridine ring, also the pyrrolidine ring. Its isomer anabasine found in the Central Asian plant*Anabasis aphylla* L. contains the piperidine ring.



nicotine

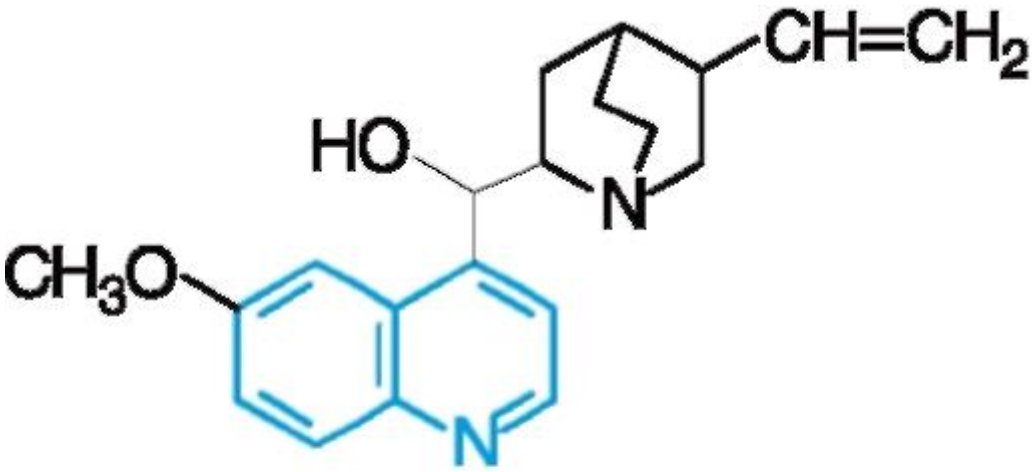


anabasine

The two alkaloids are toxic to humans (the effect of nicotine is well-known). It is interesting that anabasine as a hydrochloride salt is advised as the aid for breaking the habit of smoking[1]. Both compounds are used as agricultural insecticides and as raw materials for producing nicotinic acid by oxidation (similarly to the reaction described in Sec. 16.4.1).

Problem 16.5. Nicotine is present in tobacco as a salt with organic acids. Which of the two nitrogens of the nicotine molecule is more basic? Draw the structure of nicotine hydroxalate.

16.6.2. Quinoline and Isoquinoline Alkaloids

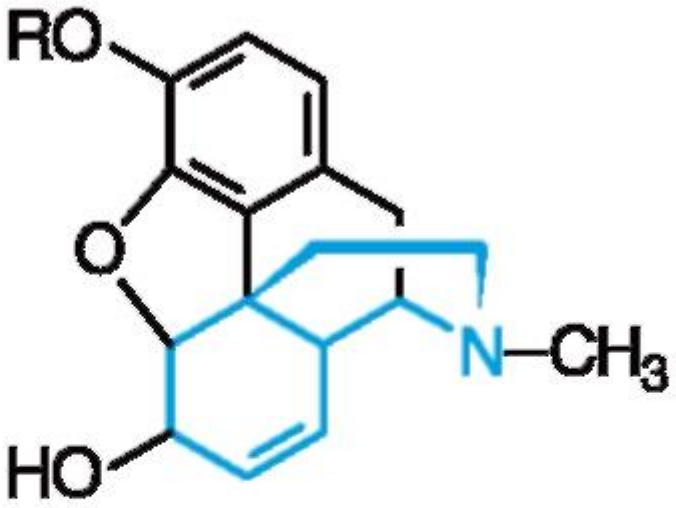


quinine

Quinine is the most known and the oldest alkaloid of the quinoline group. It contains an additional and unusual heterocyclic ring system - *quinuclidine* (non-coloured in the drawing). Quinine was isolated from cinchona bark in 1820 but the South American natives had used a decoction of the bark for centuries before this. For many long years quinine was a single remedy for malaria.

The group of isoquinoline alkaloids amounts to over 1000 substances which are located in about thirty plant families. We will be concerned with only several most known alkaloids of the opium poppy.

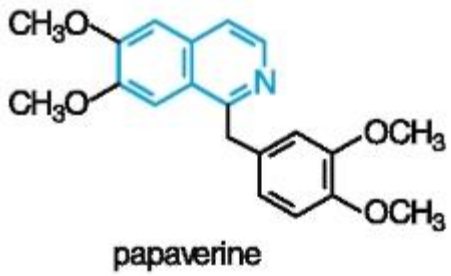
Morphine is the main component of opium and was first isolated from it in 1806. It was the first alkaloid obtained in a pure form, but its highly complicated structure was deduced only in 1920's and finally confirmed by independent masterly synthesis in 1952. Morphine remains one of the strongest analgesics though opium was used in ancient Egypt. Unfortunately, it has many disadvantages, and we will not discuss pharmacological and social aspects of morphine applications.



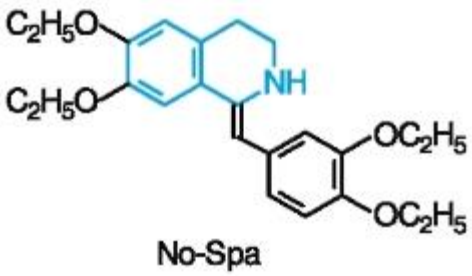
morphine (R = R' = H)
codeine (R = CH_3, R' = H)
heroin (R = R' = CH_3CO–)

The monomethyl ether of morphine, another component of opium, called *codeine* is useful as an anticough agent. The synthetic diacetyl derivative of morphine is the dolefully known narcotic heroin.

Papaverine mentioned above is a relatively simple isoquinoline derivative which is used as an antispasmodic medicine. A synthetic drug No-Spa (Drotaverine) has a similar structure but a more pronounced pharmacological effect.



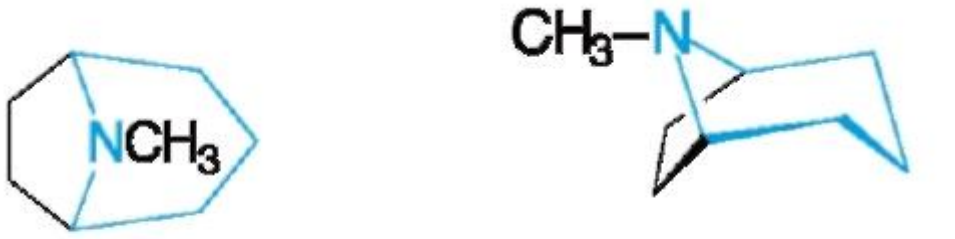
papaverine



No-Spa

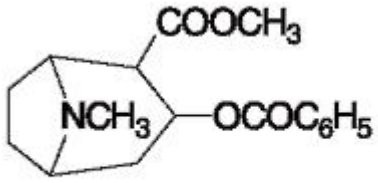
16.6.3. Tropane Alkaloids

Tropane is a saturated bicyclic system that consists of the pyrrolidine and piperidine rings (the latter is coloured):

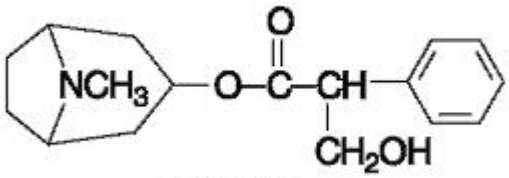


tropane and its most stable conformation

The oldest and the most known alkaloid in this group is cocaine, the first local anesthetic. Its administration is now limited because of toxicity and other side effects.



cocaine



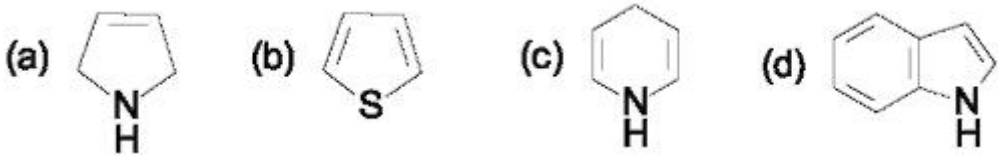
atropine

Atropine, the alkaloid of *Atropa belladonna* and other plants of *Solanaceae* family, is used in various spheres of medicine, particularly in ophthalmology to dilate the pupil of the eye.

Additional Problems

16.6. Are the heterocycles pyrimidine and imidazole that form a purine molecule aromatic themselves? Confirm your answer using the criteria of aromaticity.

16.7. Select aromatic compounds from the following ones. Classify the nitrogen atoms in the compounds (a), (c), and (d) as pyridineor pyrrole-type, *if any.*



16.8. Arrange ammonia, pyrrole, quinoline, and piperidine according to the increase in their basicity. Explain the reason of your arrangement.

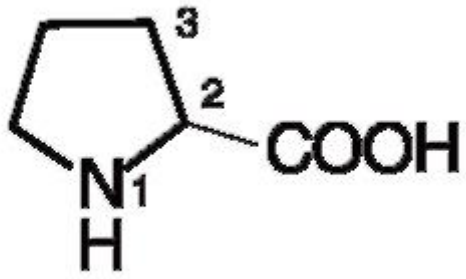
16.9. Imidazole is a crystalline compound that boils at 256 °C. Pyrimidine, whose molecule is larger than that of imidazole, has a significantly lower boiling point (124 °C). Give a reasonable explanation.

16.10. Which of the following statements corresponds to the structure and properties of the natural amino acid proline (see the drawing)?

(a) it is a carboxylic derivative of piperidine;

(b) it involves an aromatic pyrrole ring system;

(c) it forms a salt in the reaction with hydrobromic acid;



(d) it can exist as two stereoisomeric forms;

(e) it reacts with aqueous sodium hydroxide to form 4-hydroxyproline.

Write equations for all possible transformations.

16.11. The substance 2,5-dihydroxyfuran is much less stable than a tautomer, into which it is rapidly converted. Draw a structure of the tautomer and name it.

16.12. Suggest a synthetic pathway to Cordiamine, starting with 3-ethylpyridine. Give the IUPAC name for the product.

Find the most basic site in the Cordiamine molecule.

16.13. Barbital (5,5-diethyllbarbituric acid) is used in medicine as a soluble sodium salt called Barbital Sodium. Show acidic site (or sites) in the barbital molecule and draw a structure of the sodium salt.

16.14. Which of the following statements corresponds to the structure and properties of allopurinol (see the structure)?

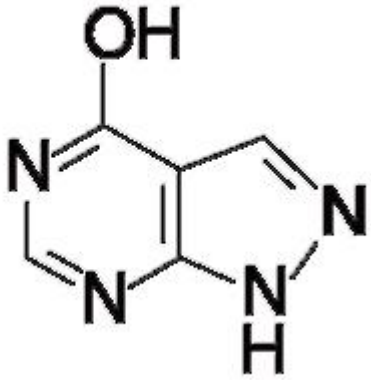
(a) the fused heterocyclic system fits criteria of aromaticity;

(b) it is a purine derivative;

(c) all nitrogens are sp^2-hybridized;

(d) it reacts with sodium hydroxide to give a salt;

(e) it can exist in a lactam tautomeric form in a solution.



Write equations for all possible transformations.

Nucleic acids, in addition to polysaccharides and proteins, represent the third type of biopolymers. They occur in all living systems, playing an exclusive role in the biosynthesis of proteins and in the transmission of hereditary characteristics. The nucleic acids are the chemical carriers of the genetic code, which prescribes the specific amino acid sequences in proteins.

Two types of nucleic acids, namely *ribonucleic acids* (RNA) and *deoxyribonucleic acids* (DNA)[1], differ in their structure and biological functions. In this chapter, the main attention will be paid to the structure and some chemical properties for better understanding biochemical behaviour of these biopolymers.

17.1. CONSTITUENTS OF NUCLEIC ACIDS

Just as polysaccharides are polymers made of monosaccharide units and as proteins are polymers made of amino acid units, nucleic acids are high molecular compounds that consist of building blocks called*nucleotides.* Nucleotides can be produced on enzyme-catalyzed hydrolysis of a nucleic acid (Fig. 17.1). The nucleotide, in its turn, can be hydrolyzed to yield a *nucleoside* and phosphoric acid. Each nucleoside can finally be cleaved into a sugar (pentose) and a pyrimidine or a purine nucleic base.

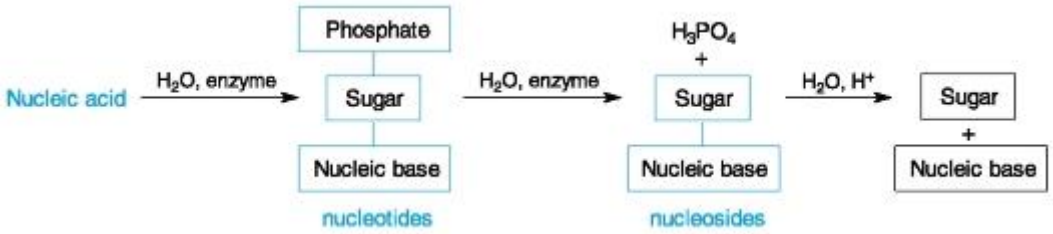
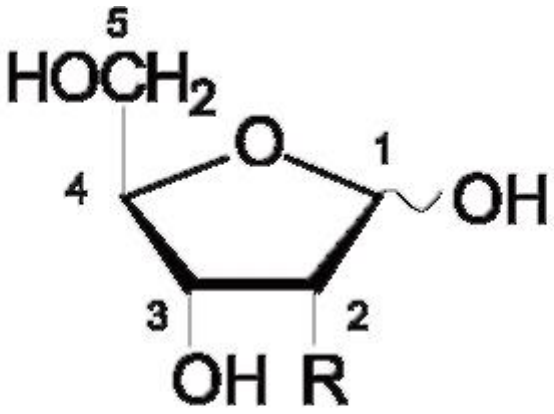


Figure 17.1. The principal scheme of splitting of nucleic acids. The rectangles designate constituents of a product.

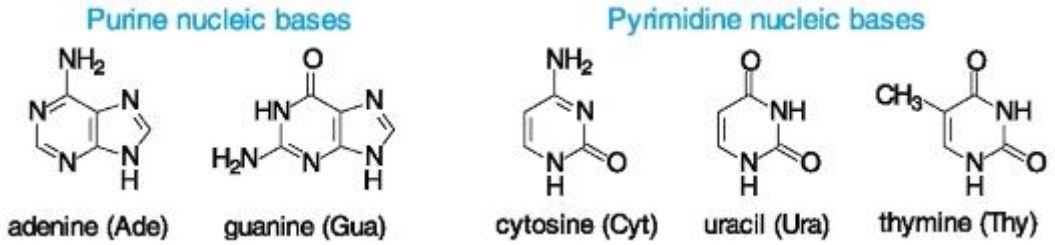
The sugar component of RNA is D-ribose, whereas DNA contains 2-deoxy-D-ribose, which lacks the hydroxyl group at C-2 (the names *ribose* and *deoxyribose* will be used hereafter). Both pentoses are in the furanose form.



D-ribose (R = OH)
2-deoxy-D-ribose (R = H)

17.1.1. Structure of Nucleosides and Nucleotides

Nucleic bases represent the hydroxy or(and) amino derivatives of purine or pyrimidine. The most common nucleic bases are adenine, guanine, and cytosine (from both RNA and DNA), uracil (from RNA), and thymine (from DNA). For convenience they are often abbreviated to their three initial letters as shown below. Thus each type of nucleic acids involves four nucleic bases, and DNA differs from RNA not only in its sugar content but also in one nucleic base (thymine instead of uracil, the former is the 5-methyl analogue of the latter).



Note the hydroxy derivatives are exclusively presented in the lactam tautomeric form (Sec. 16.5.2).

In nucleic acids, a nucleic base is connected to the anomeric carbon of the sugar by N-glycosidic bond to form a nucleoside.

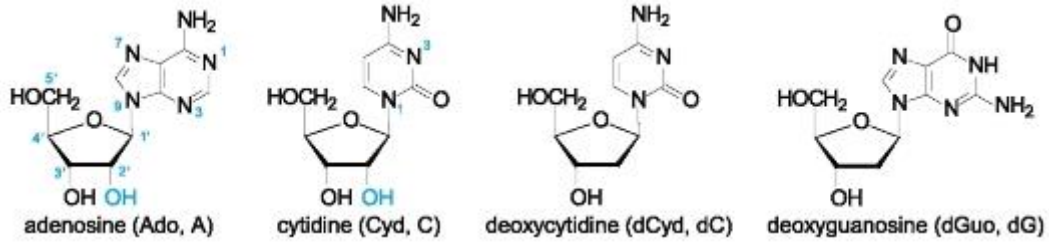
Nucleosides are N-glycosides of nucleic bases and D-ribose or 2-deoxy-D-ri

bose.

The pyrimidine bases are linked at the N-1 atom and the purine bases at the N-9 atom; the configuration of the glycosidic bond is β in all naturally occurring nucleosides.

Nucleosides are named by replacing the suffix -ine in the name of the corresponding nucleic base with the suffix -osine for the purine nucleosides or -idine for the pyrimidine nucleosides. Nucleosides containing deoxyribose get the prefix deoxy- to the name of the respective ribonucleoside. Nucleoside constituents are numbered independently like their components, except that primes are added to the numbers for the sugar unit.

For a long time, thymine nucleosides were named with some deviation from the above rule. Thus, the name «thymidine» was used for the deoxyriboside (since thymine is a component of DNA only) and the name «ribothymidine» meant the corresponding riboside. This exception was recently abandoned but the old names can be encountered in the literature.

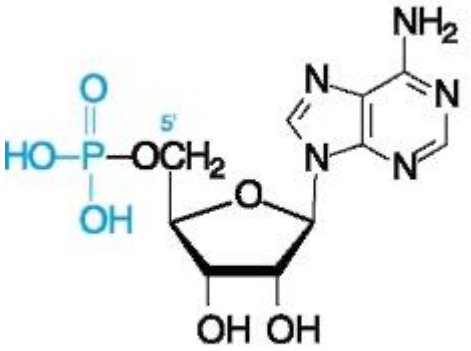


A three-letter abbreviations of nucleosides differ from those of nucleic bases in the last letter as shown in the examples above. One-letter symbols (A, C, G, T, U) are used for nucleoside *residues* in more complex structures. The letter d should be added in front to the abbreviation of deoxyribonucleosides.

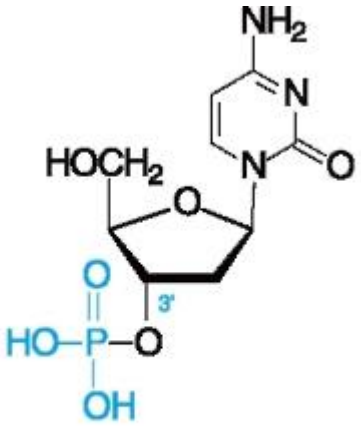
Polymeric chains of nucleic acids are composed of phosphorylated nucleosides. Such nucleoside derivatives are also components of nucleotide coenzymes (Sec. 17.4).

Nucleotides are phosphate esters of nucleosides.

Three ribonucleoside phosphates and two deoxyribonucleoside phosphates are possible according to the numbers of hydroxyl groups in the ribose unit. But only hydroxyls at C-3' and C-5' are esterified in nucleotide components of RNA. Here are two examples:



adenosine 5'-phosphate,
or 5'-adenylic acid



deoxycytidine 3'-phosphate
or 3'-deoxycytidylic acid

Nucleotides are named as the 3'- or 5'-phosphate esters of a nucleoside. Alternative names use the ending -ylic acid in place of -ine in the name of the parent nucleic base, as shown above. Nucleotide names are often abbreviated like corresponding nucleosides with additional small letter p, which is placed in front for 5'-phosphates and after the one-letter abbreviation for 3'-phosphates. In the biochemical literature, the abbreviation MP[1] stands for 5'-monophosphate, for example, AMP means adenosine 5'-phosphate. The complete names of nucleosides and nucleotides are given in Table 17.1.

Table 17.1. Names of nucleosides and nucleotides

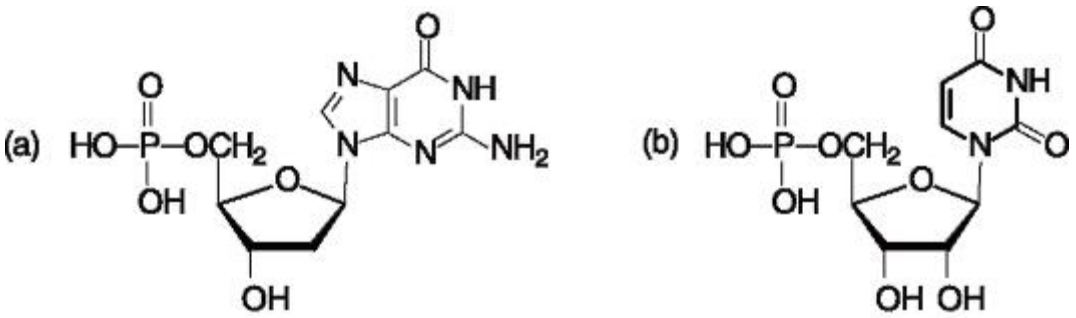
Nucleic base	*Source*	*Nucleoside*	*Nucleotide*		
			name	*abbreviations*	
Adenine	RNA	Adenosine	Adenosine 5'-phosphate	pA	AMP
	DNA	Deoxyadenosine	Deoxyadenosine 5'-phosphate	pdA	dAMP
Guanine	RNA	Guanosine	Guanosine 5'-phosphate	pG	GMP
	DNA	Deoxyguanosine	Deoxyguanosine 5'-phosphate	pdG	dGMP
Cytosine	RNA	Cytidine	Cytidine 5'-phosphate	pC	CMP
	DNA	Deoxycytidine	Deoxycytidine 5'-phosphate	pdC	dCMP
Uracil	RNA	Uridine	Uridine 5'-phosphate	pU	UMP
Thymine	DNA	Deoxythymidine	Deoxythymidine 5'-phosphate	pdT	dTMP

[1] As we will see later, some nucleotides are diphosphates (esters of diphosphoric acid) or triphosphates (esters of triphosphoric acid) that are abbreviated DP or TP, respectively.

Example 17.1. Draw the structures for (a) dGMP and (b) 5'-uridylic acid.

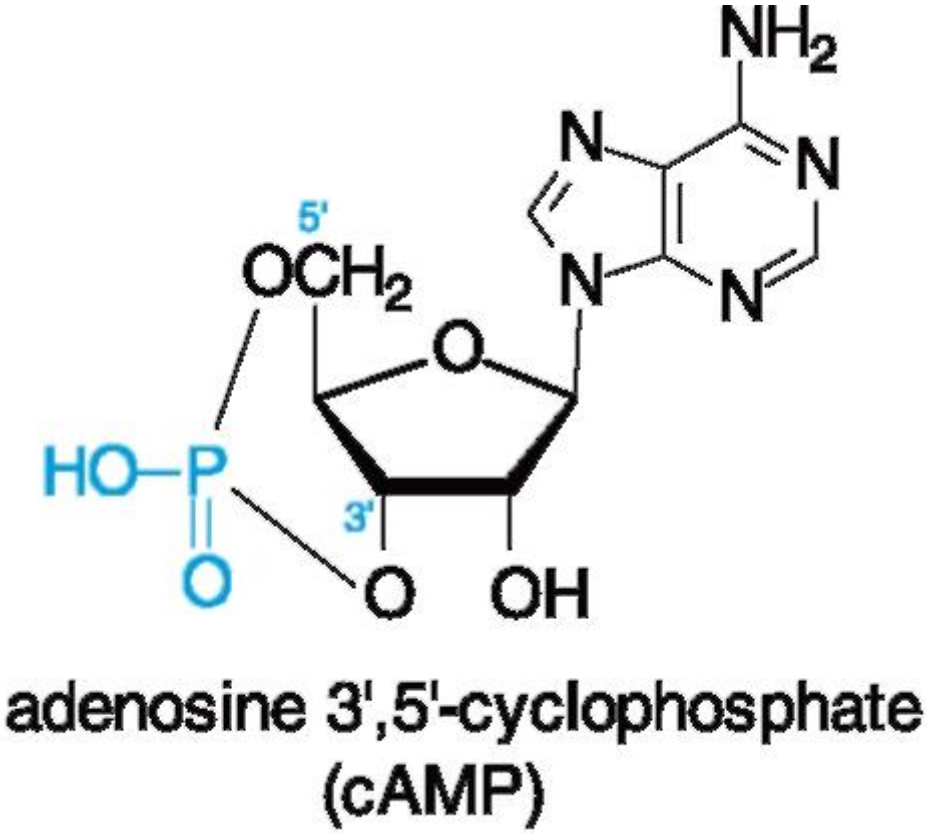
Solution. The letter *d* in the compound (a) indicates that this is a deoxynucleotide. The letter G which stands for the base is guanine, and the MP designates a monophosphate.

The compound (b) is 5'-phosphate of uridine. Thus, the structures are as follows:



Nucleosides can form cyclic phosphates, compounds in which two hydroxyls of the pentose (usually at C-3' and C-5') are esterified simultaneously by one molecule of phosphoric acid. Such cyclic phosphates are not, of course, components of nucleic acids but some of them are important in biological processes. For example, adenosine 3',5'-cyclophosphate is a mediator of a certain hormonal activity.

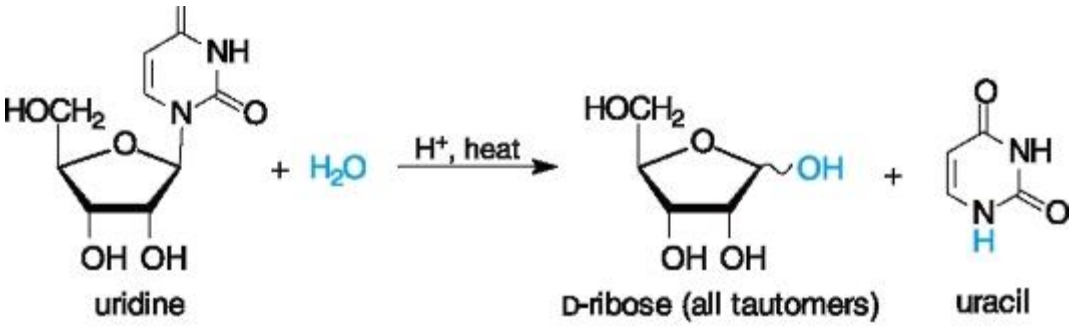
Problem 17.1. Draw the full structures of the following and give full names for the abbreviated compounds: (a) uridine; (b) dA; (c) pU.



adenosine 3',5'-cyclophosphate
(cAMP)

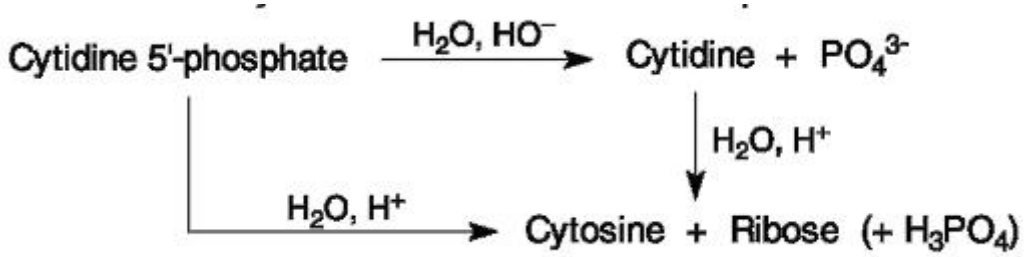
17.1.2. Some Chemical Properties of Nucleosides and Nucleotides

Hydrolysis. Like other glycosides, nucleosides are quite stable in an alkaline medium but they can be hydrolyzed in an acidic medium or by enzymes to the pentose sugar and the nucleic base, for example:o



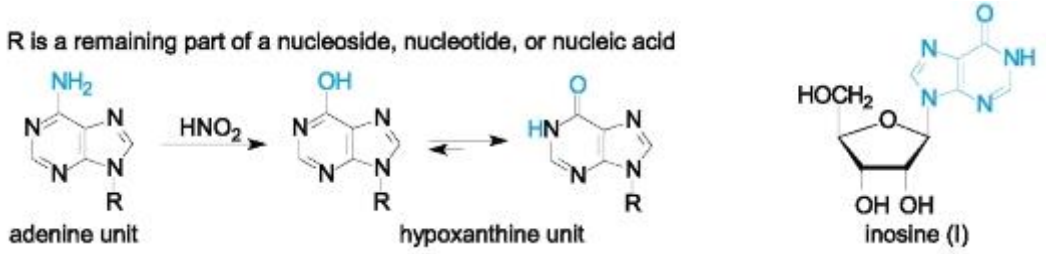
It should be noticed that N-glycosides are more resistant to the acid-catalyzed hydrolysis than ordinary O-glycosides and that deoxyribonucleosides undergo splitting much easier than ribonucleosides. Besides, deoxyribose is subjected to further transformation under acidic conditions and cannot be identified in hydrolysis products.

Complete acidic hydrolysis of nucleotides result in splitting both the phosphate ester bond and N-glycosidic bond (with the above mentioned limitation concerning deoxyribose). Nucleotides can also be hydrolyzed stepwise, first to a nucleoside and then to nucleoside components as shown above. The first step is performed under alkaline conditions that retain the N-glycosidic bond unaffected. These transformations are illustrated only with the names of compounds:



Problem 15.2. Write structural formulas for all compounds in the above scheme.

Modification of nucleic bases. Three nucleic bases - adenine, guanine, and cytosine - contain an amino group and, like amines, can react with nitrous acid (Sec. 8.4.2). When, for example, an adenine fragment is treated with nitrous acid, it is converted into a hypoxanthine derivative. Similarly, guanine is deaminated to give xanthine (2,6-dihydroxypurine), and cytosine is converted into uracil. By the way, hypoxanthine is a nucleic base of some intact RNA's; its nucleoside has the trivial name inosine(abbreviated I).



If deamination reaction takes place in living cells, it results in changing structure of nucleic acids thus affecting the genetic system. In other words, this reaction causes mutations, and nitrous acid is one of the most potent chemical mutagens.

Other known reactions that involve various functional groups of nucleic bases can be used for modification. These are alkylation, acylation, halogenation, oxidation, reactions with aldehydes and hydrazine. When applied to nucleic acids, these modifications help in determination of biopolymer structure and in genetic investigations.

17.2. PRIMARY STRUCTURE OF NUCLEIC ACIDS

Nucleotides link together in nucleic acids by forming a phosphate ester bond between the phosphate group of one nucleotide and the hydroxyl group of a sugar unit (ribose or deoxyribose) of another nucleotide. The overall structure of the nucleic acid is a macromolecule with a backbone of pentose components alternated with phosphate groups and with a nucleic base attached to each sugar unit (Fig. 17.2). The backbone of a DNA molecule differs from that of an RNA only in the absence of the OH group at C-2 of the ribose unit. The end of the polymeric chain that has a free hydroxyl at C-5' is called the *5' end*, and the other end (with a free OH at C-3') is called the *3 end.* Each end of the chain can be phosphorylated.

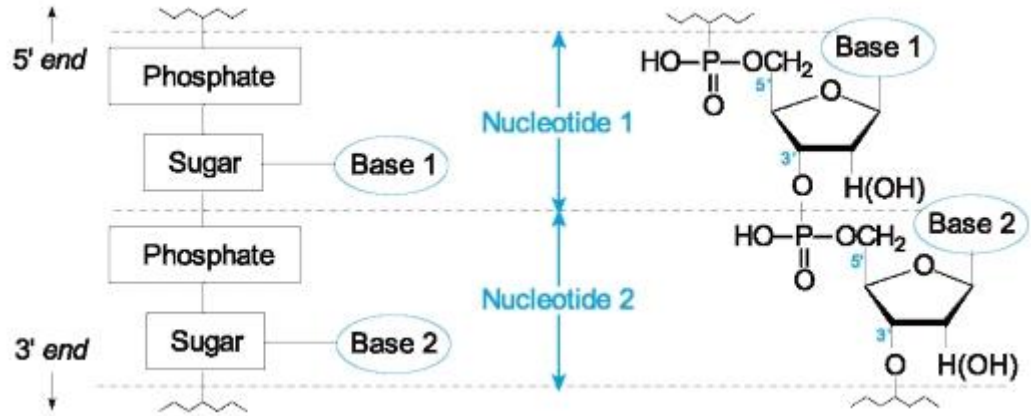
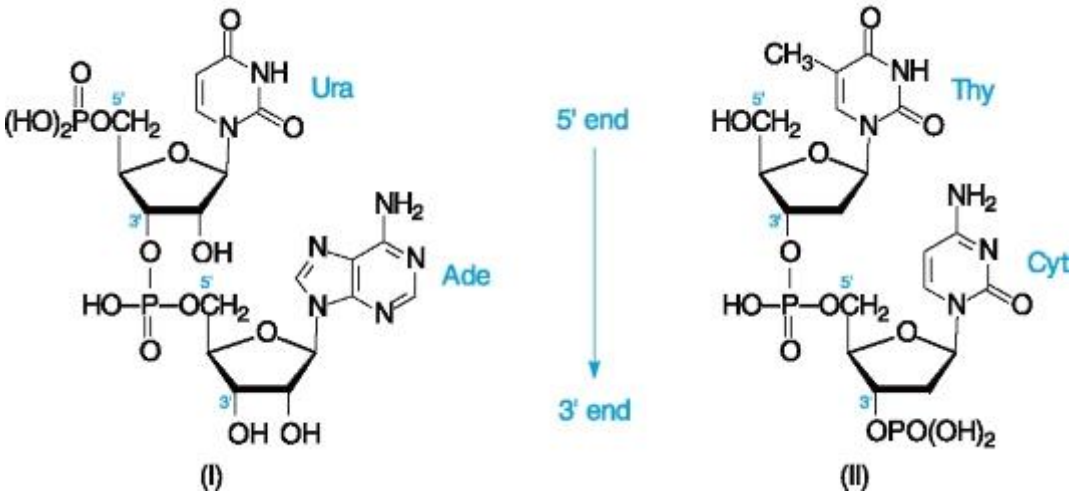


Figure 17.2. Generalized structure of nucleic acids.

Thus, RNA and DNA differ slightly in their backbones and nucleic base contents. Like proteins, each nucleic acid has a specific sequence of monomeric units in a polymeric chain.

A sequence in which nucleotides are bound in a chain is the primary structure of nucleic acids.

The nucleotide sequence in an oligoor polynucleotide is written by starting at the 5' end and identifying the nucleic bases in the direction of the 3' end. For convenience, all abbreviations given above may be used. Two examples illustrate this mode of representation. Thus, the dinucleotide (I) may be written in a short form as pUpA, and the dinucleotide (II) as d(TpCp).



For longer nucleotide chains, the letter *p* may be used only once designating phosphate group on either 5' or 3' end. The letter *d* may also be omitted when it is obvious that a nucleotide is a component of DNA (remember that thymine is present only in DNA). Thus, the above compound (I) may simply be written as pU-A, and the compound (II) as TpCp or T-Cp.

Problem 17.3. Draw the full structure of the trinucleotide U-C-Gp.

The problem of sequencing nucleic acids is much more difficult than that of sequencing proteins because of extremely high molecular mass (or long chain) of most nucleic acids. The shortest of nucleic acids is *transfer RNA* (tRNA)[1] that consists of about 70 to 95 nucleotide units, whereas the smallest DNA contains over 5,000 nucleotide units. The primary structures of hundreds of tRNA have become known since 1965. Further progress in nucleic acid sequencing has been striking.

The DNA sequencing is a powerful method applied in forensic medicine. Most of human DNA varies extremely from person to person and is characteristic of the individual (except for identical twins). The method known as *DNA profiling*, or DNA fingerprinting, is, indeed, similar to ordinary fingerprinting since it allows identifying a person. A paramount importance of the method is that it needs a negligible amount of the DNA sample (a few micrograms).

[1] A discussion of the genetic code is not the purpose of this book (this is the object of biochemistry).

17.3. SECONDARY STRUCTURE OF NUCLEIC ACIDS

The notion expressed in the heading is principally similar to that of proteins.

The secondary structure of nucleic acids is a spatial organization of a macromole cule that is defined by hydrogen bonding between nucleotide components of a chain (or chains).

Despite the fact that DNA molecules are much longer than that of RNA, the spatial arrangement of DNA was established first. A result of great importance in deducing the secondary structure of DNA was the finding that the ratio of adenine to thymine and of guanine to cytosine was very close to unity. On the basis of this fact and other physical evidence, J. Watson and F. Crick in 1953 made a revolutionary but now a classical proposal for the secondary structure of DNA supported shortly after by the X-ray data of M. Wilkins[1]. The most important features of the Watson-Crick model are the following:

- DNA consists of two polynucleotide chains (strands) coiled around a common axis in a *double helix;*
- the two strands are right-handed and run in opposite directions with regard to their 3' and 5' ends;
- the nucleic bases lie inside the helix and form *complementary pairs* by strong hydrogen bonds to each other; adenine is always paired with thymine, and guanine is always paired with cytosine;
- the helix is about 2.0 nm in diameter; the pitch of the helix is 3.4 nm and exactly ten nucleotide pairs are in each full turn.

Fig. 17.3 illustrates the above features the main of which is complementarity of the nucleic bases: A-T and G-C. This means when an adenine residue occurs in one strand, a thymine residue is opposite in the other strand. Likewise, whenever a guanine base occurs in one strand, a cytosine base appears opposite in the other strand. And vice versa. Thus the two strands of the double helix are complementary but not identical; they are said to be *anti-parallel.*

[1] The American biochemist James D. Watson, the English biochemist Francis H.C. Crick, and the English biophysicist Maurice H.F. Wilkins won the Nobel Prize for physiology and medicine (1962) for this discovery and establishing the hereditary role of DNA.

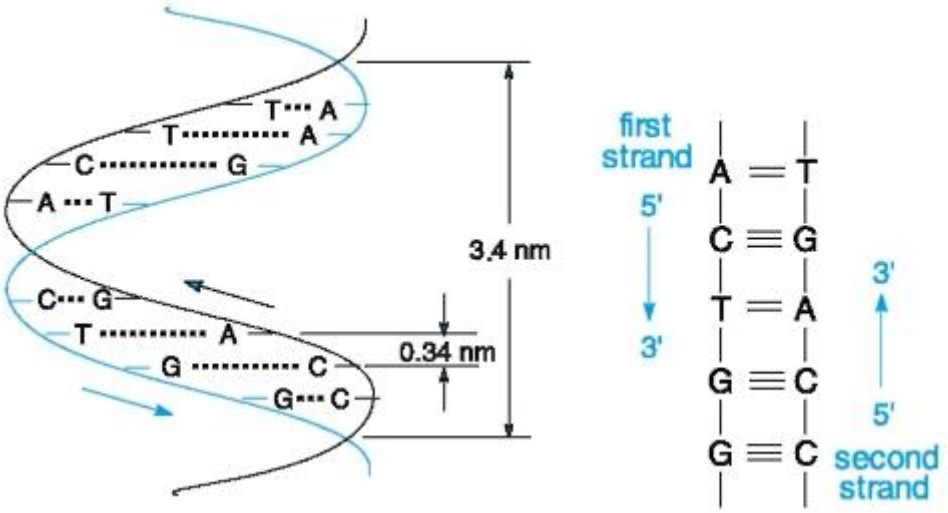


Figure 17.3. Schematic representations of the DNA double helix. The letters A, C, G, and T designate bases but not nucleosides. In the left drawing the dots indicate hydrogen bonds.

The A-T pair form two hydrogen bonds and the G-C pair is joined by three hydrogen bonds (Fig. 17.4). The geometry of both pairs is almost identical. Note that acidic sites in the hydrogen bonding are either an NH_2 group or NH fragment of a lactam form of thymine and guanine, while basic sites are either the pyridine-type nitrogen or the oxygen atom.

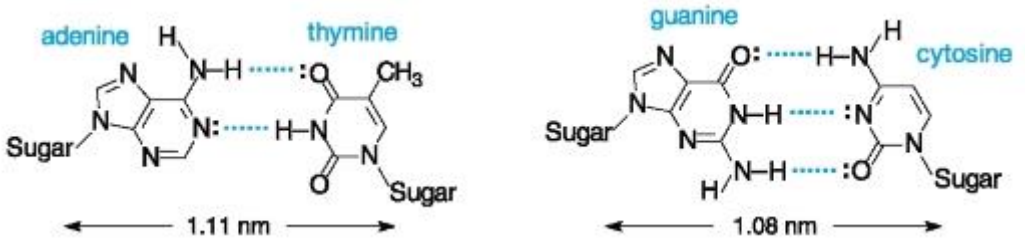


Figure 17.4. Hydrogen bonding in pairs A-T (two bonds) and C-G (three bonds).

Problem 17.4. What nucleotide sequence on one strand of DNA is complementary to the sequence ATACCTG (written from 5' to 3') on the other strand?

The principal difference in the structure of RNA is that they consist of only one polynucleotide chain. RNA has less rigid spatial arrangement when compared to that of DNA. Transfer RNA is the most studied of all RNA's because of its low molecular weight. The secondary structures of many tRNA's have been determined but it is too complicated to be presented here.

A typical tRNA has roughly the shape of a cloverleaf, as shown in Fig. 17.5 for tRNA that transports the amino acid alanine. The molecule is folded into several loops or stems by means of base pairing along the chain. Note that adenine forms the complementary pair A-U, which has the same geometry as the pair A-T in DNA.

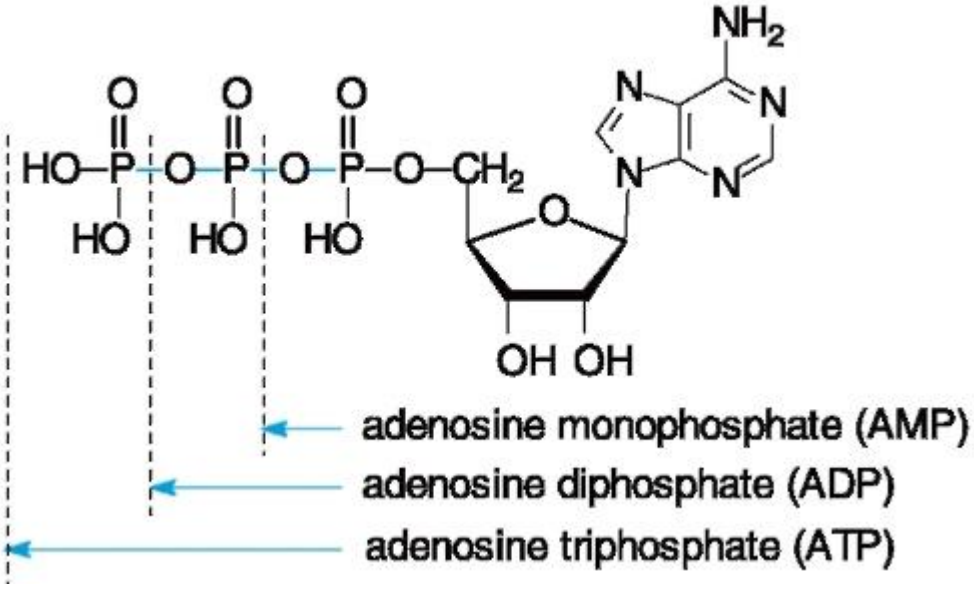
One stem of the molecule always terminates in the sequence CCA to which a specific amino acid is bonded through the free 3'-OH group of the terminal adenosine. Each tRNA also contains in one of the loops a sequence of three nucleotides (triplet) called an *anticodon.* The latter is highly important because it allows the tRNA to bind with the complementary nucleotide triplet of mRNA called a *codon.*Thus each codon (and anticodon) corresponds to only one amino acid constituting a genetic message.

17.4. NUCLEOSIDE PHOSPHATES IN BIOLOGICAL PROCESSES

Nucleotides are important not only because of their participation in biopolymers (RNA and DNA) but also because of other biological functions such as energy storage (ATP and ADP) and enzyme cofactors (NAD+, NADH, their phosphates, acetyl coenzyme A, and others).

17.4.1. Nucleoside Polyphosphates

In living cells nucleosides exist in several different phosphate forms. The most important are the 5'-phosphates of adenosine - AMP, ADP, and ATP, shown below[1]:



The bond P-O in AMP is a familiar ester bond, while the bonds P-O-P in ADP and ATP (coloured in the above structures) are phosphoric anhydride bonds. When the anhydride bond is hydrolyzed, a large amount of energy is released (compare the value given below with -12.5 kJ/mol for hydrolysis of the ester bond in aMp). For this reason, ATP and ADP are referred to as *high energy* compounds.

$$ATP + H_2O \longrightarrow ADP + H_3PO_4 ; \Delta G^\circ = -31.2 \text{ kJ/mol}$$

The biosynthesis of nucleoside triphosphates proceeds by phosphorylation of the low energy monoor diphosphate precursors. For example, ATP may be synthesized from ADP with either inorganic phosphate or another high energy triphosphate such as guanosine triphosphate (GTP). The energy liberated during oxidation of glucose and other substrates is used in the endothermic process.

$$ADP + H_3PO_4 + \text{energy} \longrightarrow ATP + H_2O$$

$$ADP + GTP \rightleftharpoons ATP + GDP$$

ATP can transfer its potential energy to various biochemically important compounds. For example, the initial step in the glycolysis process involves phosphorylation of glucose that results in the formation of glucose 6-phosphate.

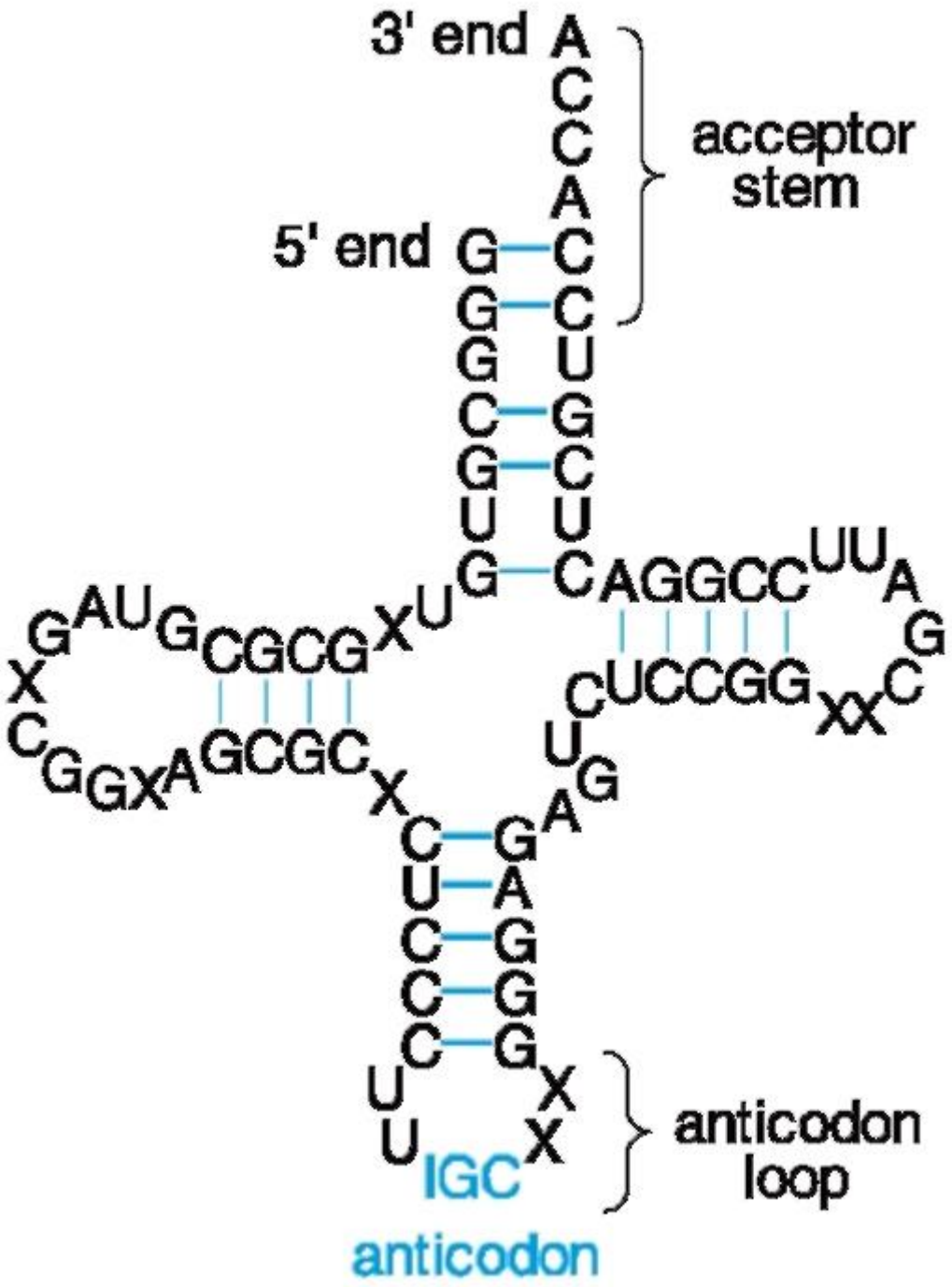
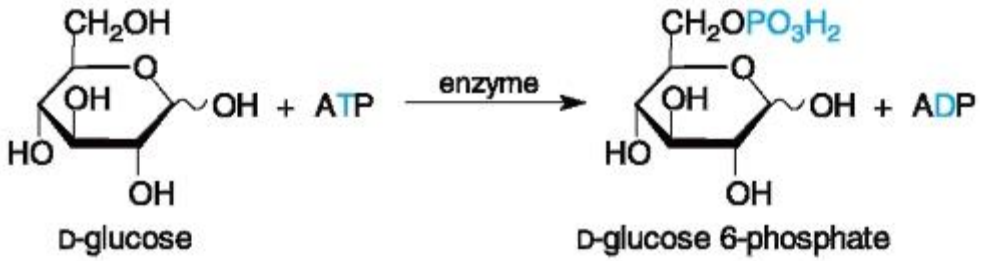
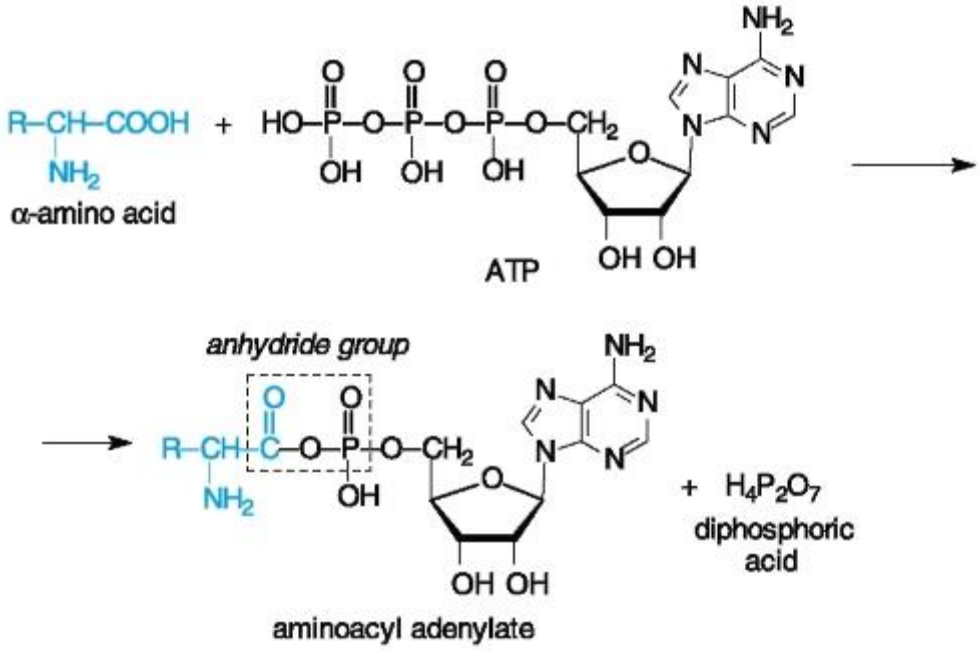


Figure 17.5. Structure of alanine tRNA from yeast. X's designate other nucleosides with unusual bases. Hydrogen bonds are shown by coloured dashes.

[1] In the biochemical literature, nucleoside phosphates, as well as other organic phosphates, are represented in an ionized form since they lose at least one hydroxylic proton at physiological pH (about 7).



A further example demonstrates the energy transfer from ATP to an α-amino acid in the form of an energy-rich anhydride bond. Phosphorylation of the amino acid results in the formation of a mixed anhydride called an *aminoacyl adenylate,* thus making the amino acid more reactive in protein biosynthesis.



The aminoacyl adenylate then reacts by acylation with the 3'-OH group at the 3'-terminal nucleotide of the specific tRNA (Fig. 17.4). Finally, the linked amino acid is transported to the ribosome where a peptide bond is synthesized.

Problem 17.5. Which structural features are present in the alanyl adenylate molecule?

(a) an amide bond;

(b) an ester bond;

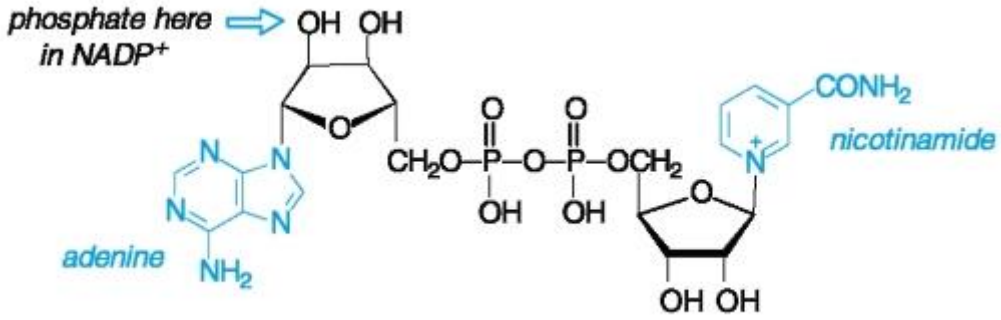
(c) an *N*-glycosidic bond;

(d) an anhydride group;

(e) a pyrimidine nucleic base.

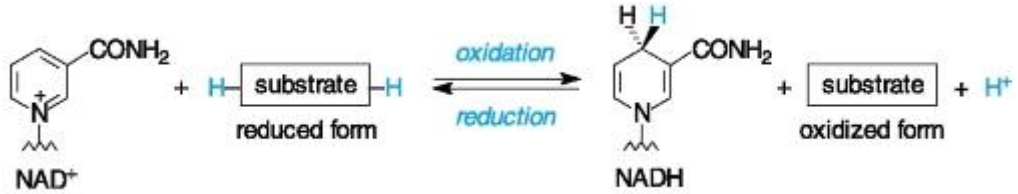
17.4.2. Nucleotide Coenzymes

Several coenzymes, that contain nucleotides in their structures, participate in various biochemical processes. Nicotinamide adenine dinucleotide (NAD+) and its phosphate (NADP+) are coenzymes involved in oxidation-reduction reactions.

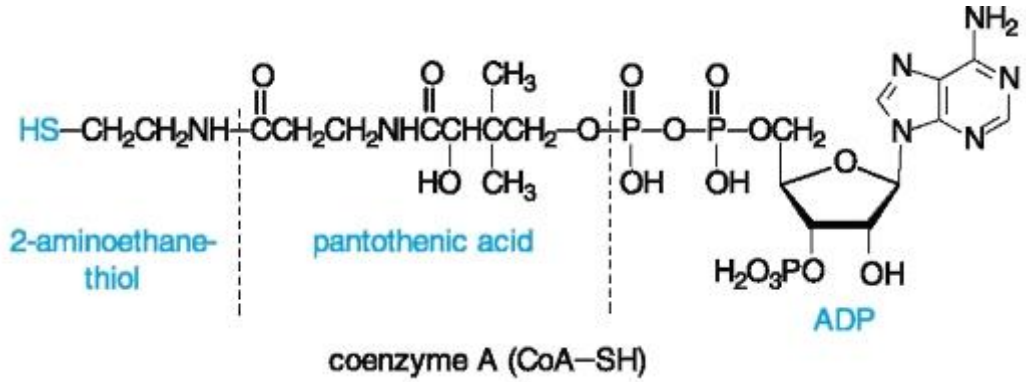


nicotinamide adenine dinucleotide (NAD^+)

The function of NAD^+ and $NADP^+$ as oxidants consists in formal abstraction of the hydride ion, H^-, from an organic substrate by the pyridinium ion of the nicotinamide part. In the reaction catalyzed by an enzyme *dehydrogenase*, both the coenzymes are converted to their reduced forms, NADH or NADPH, in which the pyridinium unit is reduced to 1,4-dihydropyridine. The reverse reaction occurs when NADH or NADPH reduces a substrate (only a changeable part of the coenzymes is shown below):

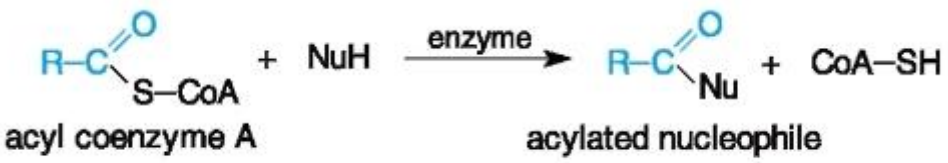


The most typical reactions of this kind such as oxidation of alcohols and hydroxy acids to corresponding carbonyl compounds, as well as reduction of a carbonyl group to hydroxyl group, were exemplified in Sec. 8.2.5 and 9.4.3.
Coenzyme A (CoA-SH or simply CoA) is another example of the complex nucleotide coenzymes. Its molecule consists of three parts: ADP (with additional phosphate group), linked by the ester bond to pantothenic acid, which, in turn, forms the amide bond to an amino thiol (see also Sec. 8.2.2).



coenzyme A (CoA–SH)

The important part of this rather complicated structure is the SH group that can be converted to thioester group to give an acylated derivative of the formula RC(O)SCoA. The latter reacts with many nucleophiles transferring an acyl group, as shown in the general equation:



The most important CoA derivative is acetyl coenzyme A, which serves as a universal acetyl-transfer agent in the cell (Sec. 9.3.3). It is also a donor of a twocarbon block in the aldol and Claisen condensations (Sec. 8.3.3 and 9.4).

Additional Problems

17.6. Pseudouridine is a minor (unusual) component of many RNA's. It represents a family of so-called C-glycosides in which ribose is β-linked to the C-5 atom of uracil. Draw the structural formula for pseudouridine.

17.7. Write an equation for hydrolysis of adenosine. Which conditions are necessary for this?

17.8. Write equations using the structural formulas both for the complete and the stepwise hydrolysis of adenosine 3'-phosphate.

17.9. Write the structure of a nucleoside, which will be formed on the treatment of cytosine with nitrous acid.

17.10. Show how uracil can form hydrogen bonds to an appropriate base in an RNA segment.

17.11. Draw the structural formula for phenylalanyl adenylate.

17.12. L-Malic acid is oxidized by NAD+ in the Krebs cycle. Write an equation for this reaction and name the product.

Type of bond*	Carbon hybridization	Bond length, pm	Bond energy, kJ/mol**
C–C	sp^3	154	350
C=C	sp^2	133	620
C≡C	sp	120	815
C–H	sp^3	109	415
=C–H	sp^2	108	435
Ar–H	sp^2	108	465
C–OH	sp^3	143	385
C–O–C	sp^3	143	335
Ar–OH	sp^2	136	435
C=O	sp^2	121	720
$C–NH_2$	sp^3	147	330
$Ar–NH_2$	sp^2	146	400
C–S	sp^3	182	265
C–F	sp^3	140	450
C–Cl	sp^3	177	340
C–Br	sp^3	194	270
C–I	sp^3	221	225
CO–H	sp^3	97	445
CO–OC	sp^3	146	165
CN–H	sp^3	101	390
CS–H	sp^3	133	345
H–H	–	74	435
H–OH	–	96	498
$H–NH_2$	–	101	425
H–F	–	92	565
H–Cl	–	127	431
H–Br	–	141	365
H–I	–	161	297
HO–OH	–	148	215
O=O***	–	121	494
F–F	–	142	158
Cl–Cl	–	199	240
Br–Br	–	228	192
I–I	–	267	150

* Unstated atoms are hydrogens or carbons.

** Average deviation up to ±10 kJ/mol over the range of a class.

*** About bonding in the O_2 molecule see Chapter 4.

Absolute configuration - the real three-dimensional structure of a chiral molecule. *See also* Relative configuration.

Acetals - the derivatives of a carbonyl compound containing two alkoxyl groups attached to the same carbon. The functional group is $>C(OR)_2$

Achiral - exhibiting no properties of «handedness». Achiral molecules have a plane or other elements of symmetry.

Acid anhydrides - derivatives of carboxylic acid that can be considered as a result of dehydration of two molecules of acids.

Acid halides - derivatives of carboxylic acids of the general formula RC(O)Hal.

Acidity constant, K_a- an evaluation of the strength of an acid HA that is considered as the equilibrium constant for the equation HA + H_2O ↔ H_3O+ + A". The acid strength is usually expressed as pK_a, which is -log K_a.

Activated complex - *see* Transition state.

Activating group - an electron-donating group, for example, a hydroxyl or amino group that increases the reactivity of aromatic rings towards electrophilic substitution reactions.

Activation energy - the difference in energy levels between reactants and transition state. It is energy that determines the rate of a reaction.

Acylation - introduction of an acyl group, RCO-, into a molecule. A hydroxyl or amino group is more often acylated to give an ester, RCOOR', or an amide, $RCONH_2$, respectively.

Aglycone - non-sugar portion of a glycoside.

Alcohols - organic compounds of the general formula ROH in which a hydroxyl group is attached to a saturated carbon.

Aldehydes - organic compounds of the general formula RCH=O.

Alditols - polyhydric alcohols resulting from reduction of the carbonyl group of a sugar.

Aldol condensation - an addition reaction between two molecules (identical or different) of carbonyl compounds resulting in the formation of an aldol.

Appendix 2. Glossary

Aldols - 1,3-hydroxy carbonyl compounds, products of an aldol condensation.

Aldonic acids - polyhydroxy carboxylic acids, non-classic sugars containing a carboxyl group in place of aldehyde group. Aliphatic - referring to nonaromatic series.

Alkaloids - naturally occurring compounds of basic character that produces striking physiological effects.

Alkylation - introduction of an alkyl group into a molecule. Amides - carboxylic acid derivatives of the general formula $RC(O)NH_2$. Amines - organic derivatives of ammonia that contain one or more organic groups instead of hydrogen atom(s). The general formulas of amines are RNH_2, R_2NH, and R_3N.

Amino sugars - non-classic sugars containing an amino group instead of an alcoholic hydroxyl group.

Androgens - male steroid sex hormones.

Angle strain - the resistance of a bond angle to deviation from its ideal tetrahedral value (about 109.5°). Angle strain is important in small-ring cycloalkanes.

Anomers - cyclic stereoisomers of sugars that differ only in configurations at the hemiacetal carbon (C-1 in aldoses and C-2 in ketoses).

anti Conformation - the arrangement around a C-C bond in which the two substituents occupy opposite positions in a Newman projection.

Antibonding orbital - an unfilled molecular orbital that is higher in energy than two initial atomic orbitals.

Anti-Markovnikov addition - an addition reaction of HX to an asymmetric unsaturated compound in which the H attaches to the less hydrogenated carbon, and the × attaches to the more hydrogenated carbon. *See also* Markovnikov's rule.

Aromatic compounds - compounds of specific reactivity that can be explained by electronic structure of cyclic conjugated systems. *See also* Aromaticity.
Aromaticity - special characteristics of cyclic conjugated systems that include unusual stability and a tendency to undergo substitution reactions rather than addition reactions. *See also* Huckel's rule.

Asymmetric centre - *see* Stereogenic centre.

Axial bond - a bond in chair cyclohexane that is parallel to the ring axis (or perpendicular to an average plane of the ring).

Ball-and-stick model - a molecular model that shows atoms as balls and covalent bonds as sticks.

Barbiturates - derivatives of barbituric acid substituted at position 5 with two (rarely with one) hydrocarbon group.

N-Bases - Brщnstedt bases whose nitrogen atom is responsible for protonation. O- and S-bases are similarly defined.

n-Bases - Brщnstedt bases that use nonshared electrons of a heteroatom for the proton acceptance.

π-Bases - Lewis bases that donate an electron pair of a double bond or of an aromatic ring to a Lewis acid, including a proton.

Basicity constant - an evaluation of the strength of a base B in terms of acidity of a conjugated acid BH^+. This constant, designated as $pK_{BH}+$, corresponds to the equilibrium $BH^+ + H_2O \leftrightarrow\pm B + H_3O^+$.

Bile acids - derivatives of hydrocarbon cholane containing a carboxyl group in a side chain at the ring D and several hydroxyl groups at other rings.

Bimolecular reaction - a reaction between two reagents.

Biopolymers - naturally occurring high-molecular compounds consisting of monomeric units such as amino acids, nucleotides, or monosaccharides. Biopolymers are proteins, nucleic acids, and polysaccharides.

Boat conformation - a conformation of cyclohexane that resembles the shape of a boat. It is less stable than a chair conformation because of a large number of eclipsing.
Bond angle - the angle formed between two adjacent bonds.

Bond dissociation energy - the amount of energy required for a homolytic cleavage of a covalent bond.

Bond length - the distance between the nuclei of two atoms bonded.

Bonding orbital - a molecular orbital that is lower in energy than two initial atomic orbitals.

Bronsted acids - neutral molecules or ions that donate proton to a base.

Bronsted bases - neutral molecules or ions that accept proton from a base.

Carbanions - species that contain a trivalent, negatively charged carbon atom, for example, R_2CH".

Carbocations - species that contain a trivalent, positively charged carbon atom, for example, R_3C^+.

Carbocyclic compounds - compounds that contain a ring (or rings) of carbon atoms only.

Carbohydrates - a big group of naturally occurring compounds that can be divided into monosaccharides, oligosaccharides, and polysaccharides (see *the definitions separately).*

Carbonyl compounds - a generalized name of aldehydes and ketones.

Carboxylic acids - organic compounds of the general formula RCOOH.

Cardiac glycosides - complex compounds of plant origin that represent glycosides of oligosaccharide and steroid alcohol.

Ceramides - derivatives of the amino alcohol sphingosine *N*-acylated with fatty acids.

Cerebrosides - glycolipids that constist of ceramide and monoosaccharide portions.

Chain reactions - reactions that involve a series of steps with each step producing a new reactive intermediate and thus propagating the next step.

Chair conformation - the lowest-energy conformation of cyclohexane that resembles the shape of a chair.

Chiral - exhibiting the property of «handedness». Chiral molecules do not have a plane and other elements of symmetry.
Chiral centre - *see* Stereogenic centre.

cis-trans Isomers - stereoisomers that differ in the positions of substituents relative to a double bond or the plane of a cyclic structure.

Coenzyme - a low-molecular compound, an assistant in enzymic reactions that functions as a carrier of atoms and groups.

Complementary bases - pairs on nucleic bases joined by hydrogen bonds. Complementarity stabilizes a structure of nucleic acids.

Condensation - a reaction with formation of a new C-C bond thus resulting in complication of carbon skeleton.

Condensed formulas - structural formulas in which dashes for the single bonds are omitted.

Configuration - the arrangement of atoms in space without regard to arrangements that differ only due to rotation about one or more single bonds.

Conformation - the three-dimensional shape of a molecule that arises only after rotation about single bonds.

Conjugate acid - the species (usually the cation) that results from protonation of a Bronsted acid.

Conjugate base - the species (usually the anion) that results from the removal of a proton from a Bronstedt acid.

Conjugation, - overlap of a *p* orbital of an atom adjacent to a double bond.

Conjugation, π,π - alternation of single and at least two multiple bonds in which *p* orbitals overlap.

Constitutional isomers - compounds that have the same molecular formula but differ in atoms connection.

Covalent bond - a bond formed by sharing even number of electrons between two nuclei.

Croton condensation - an addition reaction between two molecules (identical or different) of carbonyl compounds with a subsequent dehydration of an aldol.

Cyclo-oxo tautomerism - *see* Ring-chain tautomerism.

Dash formula - a representation of a molecule showing covalent bonds as dashes between atoms.

Deactivating group - an electron-withdrawing group, for example, a carboxyl, carbonyl, or nitro group that decreases the reactivity of aromatic rings towards electrophilic substitution reactions.

Decarboxylation - a reaction in which the carbon dioxide is liberated from a substrate.

Dehydration - a reaction that involves loss of the water molecule from a substrate.

Dehydrohalogenation - a reaction in which a hydrogen halide molecule is removed from a substrate.

Delocalization - a spreading out of electron density over a conjugated system.

Deoxy sugars - carbohydrates devoid of one or more hydroxyl groups.

Dextrorotatory - rotating the polarized light in a clockwise (right-handed) direction. Dextrorotatory compounds are designated as (+).

π-Diastereomers - configurational stereoisomers containing a π-bond.

σ-Diastereomers - a pair of stereoisomers that are not mirror images of one another. All their physical and chemical properties are different.

Dipolar ion - a neutral dipolar molecule that has a salt-like structure with positive and negative charges. For example, amino acids exist as dipolar ions, NH_3+-CHR-COO".

Disulfides - organic compounds of the general formula RSSR'.

DNA - deoxyribonucleic acid, the biopolymer consisting of nucleoside units linked together through phosphate-deoxyribose bonds. It is a carrier of genetic information of organisms.

Dreiding model - a molecular model that shows covalent bonds as sticks without representation of the atoms which are only implied.

Eclipsed conformation - the arrangement around a C-C bond in which the two substituents are in direct opposition to each other in a Newman projection.

Electronegativity - the ability of an atom to attract valence electrons and thus polarize a bond.

Electrophile - an electron-poor reagent (a neutral molecule or a cation) that accepts an electron pair from a nucleophile in a polar reaction.

Elimination - a removal small molecules (e. g. H_2O, NH_3, HCl) from a substrate.

Enantiomers - stereoisomers that are nonsuperposable mirror images of one another.

Enols - unstable unsaturated alcohol with a fragment C=C-OH.

Epimers - stereoisomers that differ in configuration at only one chiral centre.

Equatorial bond - a bond in chair cyclohexane that lies approximately in the rough plane of the ring.

Equilibrium constant, K_{eq}- a measure of the equilibrium position for the reaction A + B ↔C + D, that is expressed as $[C][D][A]^{-1}[B]^{-1}$.

Esterification - a formation of an ester in the reaction between a carboxylic acid and an alcohol.

Esters - carboxylic acid derivatives of the general formula RCOOR'.

Estrogens - female steroid sex hormones.

Ethers - compounds of the general formula R-O-R', where R and R' are alkyl or aryl groups.

Excited-state configuration - an electronic configuration of an atom in which an electron (or electrons) is promoted from a lower-energy orbital to a higher-energy level, for example, from 2*s* to 2*p* orbital in the carbon atom.

Fats - solid triacylglycerols usually of animal origin.

Fatty acids - long-chain carboxylic acids, constituents of lipids.

Fischer projection - a means of depicting the tetrahedral model of a molecule onto a plane.

Functional group - an atom or group of atoms in a molecule that determines chemical characteristics of a compound.

*gauche*Conformation - the arrangement around a C-C bond in which the two substituents lie 60° apart in a Newman projection.

Gangliosides - glycolipids that constitute of ceramide and oligosaccharide portions.
Glycolipids - lipids that consist of ceramide and monoor oligosaccharide portions.

Glycosides - cyclic acetals of a sugar in which hemiacetal OH group is replaced by the OR group.

N-Glycosides - nitrogen analogues of a glycoside in which the acetal oxygen is replaced by the nitrogen atom.

Glycosidic bond - a bond between anomeric carbon of monoor oligosaccharide and aglycone component OR.

Ground-state configuration - the lower energy, most stable electronic configuration of an atom.

Halogenation - introduction of a halogen atom(s) into a molecule by means of addition or substitution reaction.

Haworth formulas - representation of cyclic forms of sugars in the form of planar regular polygons (pentagon for furanoses and hexagon for pyranoses).

Hemiacetals - the derivative of a carbonyl compound containing a hydroxyl and an alkoxyl groups attached to the same carbon. The functional group is >C(OH)(OR^

Heterocyclic compounds - compounds containing in the cycle at least one heteroatom, an atom that is not carbon.

Heterofunctional compounds - compounds that involve two or more different functional groups.

Heterolytic bond cleavage - a kind of bond breaking in which a molecule A:B loses a fragment with both of the bonding electrons to produce the charged species B:" and A^+.

Heteropolysaccharides - polysaccharides consisting of different monosaccharide units.

Homolytic bond cleavage - a kind of bond breaking in which a molecule A: B splits in such a way that one electron is retained with each neutral fragment, A' and B'.

Homopolysaccharides - polysaccharides that built up from identical monosaccharide units.

Huckel's rule - a rule stating that a molecule with cyclic conjugation is aromatic if it contains $(4n + 2)$ πelectrons, where n is an integer.
Hybrid orbital - an orbital that is mathematically obtained by a combination of an *s* and *p* atomic orbitals.

Hydration - addition of water to an unsaturated compound.

Hydrogen bond - a weak attraction between a hydrogen atom bonded to an electronegative atom (the most often in groups OH or NH) and a lone-pair of electrons on another atom (usually O or N).

Hydrogenation - addition of hydrogen to an unsaturated compound to yield a saturated product.

Hydrolysis - splitting of a bond in a reaction with water.

Hydroperoxides - organic compounds of the general formula ROOH.

Inductive effect - the shifting of electrons in a bond in response to electronegativity of nearby atoms. The negative effect of a group is designated as -/, the positive effect - as +/.

Ionic bond - a bond between two ions of unlike charges due to the electrical attraction.

Isoelectric point - the pH value (designated p/) at which an amino acid or a protein exists in an electrically neutral dipolar form.

Isomers - compounds that have the same molecular formula but differ in their structures. *See also*Constitutional isomers *and* Stereoisomers.

Isoprene rule - according to this, a terpene molecule can be considered as a joining of isoprene (2-methyl-1,3-butadiene) unit connected mostly in a "head-totail" manner.

Isoprenoids - organic compounds whose framework consists of five-carbon isoprene units.

Keto-enol tautomerism - equilibrium between the keto and enol forms of a carbonyl compound, which consists in a proton migration from the α-CH acidic site to the oxygen atom of a carbonyl group.

Ketones - compounds of the general formula R-CO-R'.

Lactams - cyclic amides that can be considered as a result of the intramolecular cyclization of amino acids.

Lactim-lactam tautomerism - equilibrium between the lactam and lactim forms of 2-hydroxy substituted nitrogen-containing heterocycles.

Lactones - cyclic esters that can be considered as a result of the intramolecular cyclization of hydroxy carboxylic acids.

Leaving group - a group that is replaced in substitution reaction.

Levorotatory - rotating the polarized light in a counterclockwise (left-handed) direction. Lelorotatory compounds are designated as (-).

Lewis acids - compounds having a vacant orbital that can accept an electron pair from a base.

Lewis bases - compounds that donate an electron pair to a base.

Lipids - naturally occurring nonpolar substances that belong to many structural classes, including triacylglycerols, terpenoids, steroids, and others.

Locant - a numeral or a letter that shows a position of a substituent or a multiple bond in a parent structure.

Lone-pair electrons - a nonbonding pair of electrons that occupies a valence orbital.

Markovnikov's rule - this formulates the orientation of electrophilic addition reactions as follows: In the addition of HX to an asymmetric alkene, the H attaches to the more hydrogenated carbon, and the × attaches to the less hydrogenated carbon.

meso Compounds - optically inactive achiral stereoisomers containing chiral centres.

Mesomeric effect - the shifting of electron density caused by a substituent in conjugated system through *p*-orbital overlap.

Molecular formula - a formula that expresses the total numbers of each kind of atoms in a molecule.

Molecular orbital - an orbital that results from overlapping two or more atomic orbitals when bonds are formed.

Monofunctional compounds - compounds that have the only functional group.

Monosaccharides - polyhydroxy aldehydes or polyhydroxy ketones mainly in the cyclic hemiacetal form.
Mutarotation - the spontaneous change in optical rotation of the freshly prepared solution of a pure anomeric sugar.

Newman projection - a method for representing various conformations of a molecule when the C-C bond is viewed end-on.

Nomenclature - an arrangement of terms that describes complete structure of organic molecules.

Non-reducing sugars - sugars that do not reduce Ag+ ion (Tollens' reagent) or Cu^{2+} ion (Fehling's or Benedict's reagent). Nucleic acids - *see* DNA *and* RNA.

Nucleic base - purine and pyrimidine derivatives, constituents of nucleic acids.

Nucleophile - an electron-rich reagent (a neutral molecule or anion) that donates an electron pair to an electrophile in a polar reaction.

Nucleophilicity - ability to react with electrophilic carbon, i. e. affinity to carbon.

Nucleosides - nucleic acid constituents that represent an N-glycoside of a nucleic base and D-ribose or 2-deoxy-D-ribose.

Nucleotides - phosphate esters of a nucleoside. Nucleotides are the monomeric units of DNA and RNA.

Oils - liquid triacylglycerols usually isolated from plant sources.

Oligosaccharides - compounds that consist of several monosaccharide units linked together by a glycosidic bond between the anomeric carbon of one unit and a hydroxyl oxygen of another monosaccharide.

Optical isomers - *see* Enantiomers.

Optical activity - a physical characteristic of a substance attributed to rotation of plane-polarized light.

Orbital - a region of a space where the probability of finding an electron is large.

Parent name - a part of the full name of a compound from which a particular name is derived by prescribed rules.

Peptides - compounds composed of many (from 2 up to 100) amino acids connected through amide (peptide) bonds.
Peroxides - organic compounds of the general formula ROOR.

Phenols - compounds of the general formula ArOH in which a hydroxyl group is attached to an aromatic ring.

Phosphatides - derivatives of phosphatidic acids esterified on the OH group of phosphoric acid.

Phosphatidic acids - derivatives of glycerol-3-phosphate esterified on the OH groups of glycerol.

Phospholipids - lipids containing a phosphate residue.

Pi (π) bond - a covalent bond formed by sideways overlap of atomic orbitals.

Plane of symmetry - an imaginary plane that cuts an object in such a way that its two halves are reflections of each other.

Plane-polarized light - electromagnetic waves that vibrate in a single plane rather than in random planes as an ordinary light beam does.

Polar covalent bond - a chemical bond (single or double) between atoms that differ in their electronegativities, for example, C-N or C=O.

Polar reaction - a reaction that occurs between positive and negative charges (full or partial) on both reactants. This is usually a reaction between a nucleophile and an electrophile.

Polarity - the unsymmetrical distribution of electrons in a molecule that results from different electronegativities of atoms bonded.

Polarizability - the ability of electrons to respond to a changing electric field, as a result of its interaction with solvent or with polar reagents.

Polyfunctional compounds - compounds that involve two or more identical functional groups.

Polysaccharides - biopolymers built up of many monosaccharides linked together by glycosidic bonds.

Primary structure - a sequence of monomeric units in biopolymers (proteins, polysaccharides, and nucleic acids).

Proteins - biopolymers composed of many (over 100) amino acids connected through amide (peptide) bonds.

Pyridine-type nitrogen - the sp^2-hybridized nitrogen atom that has one electron on an unhybridized *p*orbital and lone pair of electrons on a hybridized orbital. The latter is available for protonation.

Pyrrole-type nitrogen - the sp^2-hybridized nitrogen atom that has lone-pair of electrons on an unhybridized *p* orbital and one electron on a hybridized orbital. The former is a part of the six-π-electron cloud.

Quinones - six-membered unsaturated 1,2- or 1,4-diketones.

Racemic mixture - an optically inactive mixture containing equal amounts of both enantiomers.

Radical - a species that has an unpaired electron in one of its orbitals. When used in nomenclature, this term denotes a part of a molecule that remains after removal of one or more hydrogen atoms from it.

Radical reaction - a reaction that involves symmetrical bond breaking and bond making with participation of a free radical reactant and intermediates.

Radicofunctional nomenclature - a method for constructing the name of a compound on basis a class name and radical(s) present.

Reaction energy diagram - a graphic representation of the reaction course in which potential energy is plotted as a function of reaction progress.

Reducing sugars - any sugar that can reduce Ag+ ion (Tollens' reagent) or Cu^{2+} ion (Fehling's or Benedict's reagent).

Regioselective reaction - a reaction in which one of the possible constitutional isomers is exclusively (or predominantly) produced rather than a mixture of products.

Relative configuration - a configuration of a chiral compound relative to D-(+)or L-(-)-glyceraldehyde that were chosen as configurational standards. *See also* Absolute configuration.

Resonance effect - *see* Mesomeric effect.

Resonance hybrid - a combination of contributing resonance forms for a compound that cannot be represented adequately by a single «ordinary» structure. Resonance structures exist only on paper.
Ring-chain tautomerism - interconversions of an open-chain and cyclic forms of monosaccharides in a solution.

Ring flip - the conversion of one chair conformation of cyclohexane into another chair conformation. This transforms axial substituents into equatorial ones, and vice versa.

RNA - ribonucleic acid, the biopolymer consisting of nucleoside units linked together through phosphate-ribose bonds.

Reconvention - a method for describing absolute configuration at stereogenic centre that is based on a priority of the four substituents attached to the centre.

eaponification - an old term used for alkaline hydrolysis of an ester to yield a carboxylic acid salt.

eaturated - referring to compounds that have only single C-C bonds and thus cannot undergo addition reactions.

eecondary structure - a higher level of biopolymer structure that involves three-dimensional organization of a macromolecule chain due to hydrogen bonding between various functional groups.

Shiff's bases - organic compounds of the general formula RCH=NR', the products of a reaction of aldehydes RCH=O and primary amines $R'NH_2$.

Sigma (σ) bond - a covalent bond formed by head-on overlap of atomic orbitals.

Skeletal formula - a representation of a molecule (mostly cyclic) that shows a molecular framework, omitting the symbols of CH fragments.

Soaps - mixtures of sodium or potassium salts of fatty acids produced on alkaline hydrolysis of triacylglycerols.

Space-filling model - a molecular model that indicates the space occupied by the atoms.

sp Orbital - a hybrid orbital derived from the combination of an s with a *p* atomic orbital. The two *sp*hybrid orbitals are oriented at an angle of 180° to each other.

sp^2Orbital - a hybrid orbital derived from the combination of an s with two *p* atomic orbitals. The three hybrid orbitals lie in a plane at angles of about 120° to each other.

sp^3Orbital - a hybrid orbital derived from the combination of an s with three *p* atomic orbitals. The four hybrid orbitals are directed at angles of about 109.5° to each other.

Specific rotation - the optical rotation of a sample adjusted to standardized conditions.

Sphingolipids - lipids containing an amino alcohol sphingosine.

Staggered conformation - the arrangement around a C-C bond in which any two substituents are at the maximum distance in a Newman projection.

Stereochemical nomenclature, d,l- an old nomenclature based on the assignment of a configuration relative to D- or L-glyceraldehyde as the reference compound. The D,L system is still used for amino acids and, especially, for carbohydrates.

Stereochemistry - the branch of chemistry that deals with the spatial aspects of molecules.

Stereogenic centre - an atom that has four different substituents attached and is therefore chiral.

Stereoisomers - isomers that have the same order of atoms attachment, but differ only in the arrangement of their atoms or groups in space.

Stereoselective reaction - a reaction in which one of the possible stereoisomers is exclusively (or predominantly) produced regardless of stereochemistry of a substrate.

Stereospecific reaction - a reaction in which a particular stereoisomer gives a specific stereoisomeric product.

Steric strain - a strain in a molecule when two groups are too close together.

Steroids - biologically active substances whose structure is based on the tetracyclic carbon skeleton of gonane.

Sterols - derivatives of hydrocarbon cholestane containing a hydroxyl group in the ring A.

Stick model - *see* Dreiding model.

Structural isomers - *see* Constitutional isomers.

Substituent - an atom or an atomic group that substitutes the atom H in an organic molecule.

Substitutive nomenclature - a method for constructing the name of a compound on principles of substitution in a parent structure.

Sulfides - organic compounds of the general formula RSR', where R and R' are alkyl or aryl groups, that are sulfur analogues of ethers.

Sulfonation - introduction of a sulfo group, SO_3H, into a molecule.

Sulfonic acids - organic compounds of the general formula RSO_3H.

Tautomerism - a dynamic isomerism, which consists in migration of some groups within a molecule and is accompanied by redistribution of electron density.

Tautomers - constitutional isomers that are interconvertable in a solution.

Terpeniods (terpenes) - lipids whose molecules are formally built up from isoprene units.

Tertiary structure - a higher level of biopolymer structure that involves an overall three-dimensional shape of a macromolecule due to hydrogen bonding and other types of interaction.

Thioesters - organic compounds of the general formula RC(O)SR', sulfur analogues of esters.

Thiols - compounds of the general formula RSH that are sulfur analogues of alcohols.

Torsional strain - a strain in a molecule resulting from electron repulsion between eclipsed bonds.

Transamination - enzymic interconversion of α-amino acids and α-oxo acids.

Transition state - a structure of an activated complex between reactants that corresponds maximum energy in the conversion of the reactants to products.

Triacylglycerols - ester of glycerol and long-chain fatty acids. Triacylglycerols are components of animal fats and vegetable oils.

Trivial names - historical names of compounds not reflecting their structure.

Unimolecular reaction - a reaction or its step that involves spontaneous transformation of the starting compound without participation of other reagents.
Unsaturated - referring to compounds that have multiple carbon-carbon bonds and thus can undergo addition reactions.

Uronic acids - polyhydroxy carboxylic acid, non-classic sugars resulting from oxidation of the CH_2OH group of a sugar to a carboxyl group.

Van der Waals' forces - the weak forces of intermolecular interaction arising without transfer of electrons. They are responsible for the attraction between molecules.

Waxes - esters of fatty acid and long-chain alcohols.

Zaitsev's rule - a rule stating that elimination reactions normally proceed with the formation of the more substituted alkene as a predominant
product. Zwitterion - *see* Dipolar ion.

Appendix 3. Answers to problems

TO THE STUDENT

As it was mentioned earlier, problem solving is, probably, the best key to studying chemistry. This work is very often like solving a puzzle; therefore never look at the

answer until you succeed (or failed) in your solution. The best way to use this Appendix is to compare your solutions and those given below, or to find explanations for the unsolved problems only after all attempts to solve them have failed.

Appendix 3 presents solutions to all the problems (in-text and additional problems). In many cases not only the answer is given, but also the explanation of the reasoning, even if it has not been questioned. Some problems have more than one solution, but only one is generally given. Thus, your answer may differ from the one presented here. The main thing is that your solution must be correct.

To Chapter 1

1.2. It is similar to methane tetrahedral structure, except for appreciable deviation from the value of 109.5°.

1.3. The CH_3 carbon is sp^3-hybridized and forms the H-C-H bond angles near 109°. Double-bond carbons are sp^2-hybridized and form the H-C-H and HC-C angles near 120°.

1.4. Both oxygen and nitrogen are sp^3-hybridized and have tetrahedral geometry. Three-dimensional structures are similar to those in Figures 1.10, *c* and 1.11, *b*, respectively.

1.5. $CH_3CH_2CH_2CH_2Br$, $CH_3CH_2CHBrCH_3$, $(CH_3)_2CHCH_2Br$, and $(CH_3)_3CBr$.

1.6. (a) CH_3NH_2; (b) $CH_3CH{=}CH_2$; (c) $CH_3CH_2CH_2Cl$; (d) $(CH_3)_3CH$; (e) $CH_3CH(OH)CH_3$. Two isomers (in total) are possible for compounds from (b) to (d), and three isomers - for the compound (e).

1.8. Sodium and oxygen form the ionic bond, $CH_3CH_2O^-\ Na^+$; other bonds in the molecule are covalent.

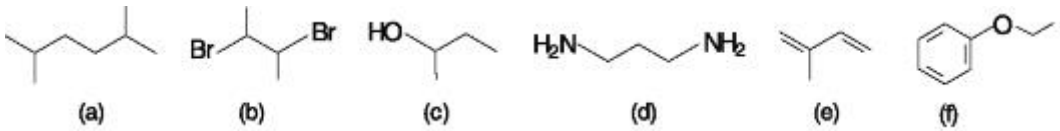
1.9. All bonds are covalent, except for $CaCl_2$.

1.10. All carbons are sp^2-hybridized in (a), (c), and (e); all carbons are sp^3-hybridized in (b) and (d). The triple-bond carbon in (f) is *sp*-hybridized, and two remaining sp^3-hybridized.

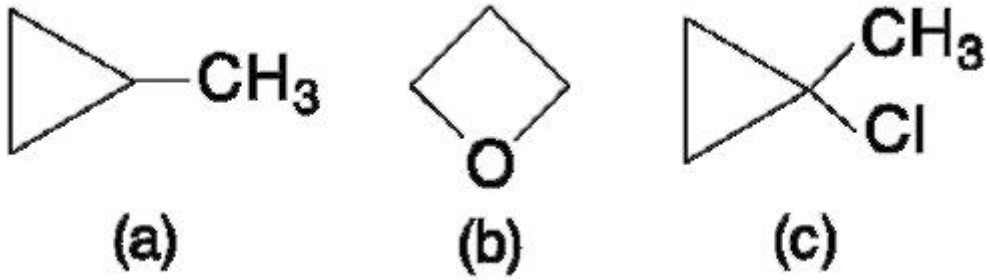
1.11. The compounds (b); (e) and (f).

1.12. (a) C4H8O2; (b) C4H5N; (c) C10H8; (d) C10H20O.

1.13. Only skeletal formulas shown:



1.15. Four cyclic constitutional isomers are possible for (c), and two isomers - for each of (a) and (b). The examples are:



To Chapter 2

2.1. *Hint.* Find in:

(a) four primary and two tertiary carbons;

(b) the same as in the compound (a);

(c) one primary, five secondary, one tertiary, and one quaternary carbons.

2.2. For answers see Problem 2.4.

2.3. Parent names are bold-faced:

(a) 2-methyl-1-propanamine (or isobutyramide);

(b) propoxyethylene (or propyl vinyl ether);

(c) ethylpropanedioic acid;

(d) 2,3-butanediol;

(e) phenylmethanol (or benzyl alcohol).

2.4. For answers see Problem 2.2.

2.5. Correct names are:

(a) 2-methylbutane;

(d) 4-aminobenzoic acid;

(b) tribromofluromethane;

(e) 1-amino-2-propanol;

(c) 2-chlorocyclohexanol;

(f) 3-methylaniline (or *m*-toluidine).

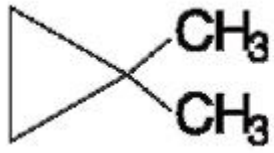
2.6. There are five isomers in all:

(a) methoxybenzene (or methyl phenyl ether);

(b) phenylmethanol (or benzyl alcohol);

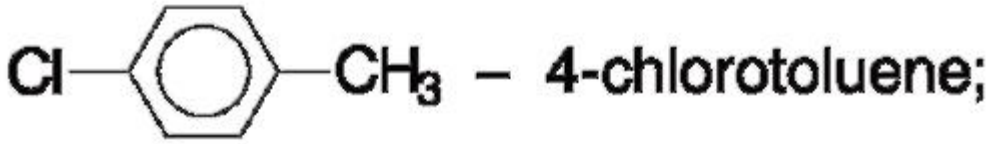
(c) three isomers *(ortho, meta,* and *para)* of methylphenol.

2.7. Variants:

(a)  – 1,1-dimethylcyclopropane;

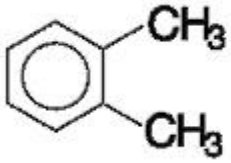
(b) $CH_3COCH_2CH_2CH_3$ - 2-pentanone;

(c) $CH_3CH_2CH_2COOH$ - butanoic acid (or butyric acid);

(d)  – 4-chlorotoluene;

(e) $H_2NCH_2CH_2OH$ - 2-aminoethanol.

2.8. Variants:

(a)  – 1,2-dimethylbenzene (or *o*-xylene);

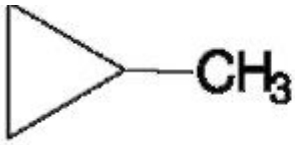
(b) $(CH_3)_2CHCHO$ - 2-methylpropanal;

(c) $(CH_3)_2CHNH_2$ - 2-propanamine (or isopropylamine);

(d) $HOCH_2CH_2OH$ - 1,2-ethanediol;

(e) $HOCH_2CH_2COOH$ - 3-hydroxypropanoic acid.

2.9. Variants (except for stereoisomers):

(a)  – methylcyclopropane,

$CH_3CH=CHCH_3$ - 2-butene, in all four isomers;

(b) $CH_2=CHOCH_2CH_3$ - ethoxyethylene, in all three isomers;

(c) $CH_3CH_2C(O)CH_2CH_3$ - 3-pentanone, only two isomers;

(d) $CH_3CH=CHCOOH$ - 2-butenoic acid, in all three isomers.

2.10. The name of the parent structure is bold-faced (italics). The full names are:

(a) 2,4-*pentadien*oic acid;

(b) 2,6-diaminohexanoic acid;

(c) 2,3-dihydroxy*propan*al;

(d) 2-isopropyl-5-methyl*cyclohexan*ol.

2.11. Start with the name of the parent structure (bold-faced, italics):

(a) 2-amino*ethane*sulfonic acid;

(b) 3-oxo*butan*oic acid;

(c) 2-amino-3-mercapto*propan*oic acid;

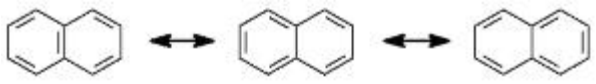
(d) 4-hydroxy-3-methoxy*benzaldehyde.*

To Chapter 3

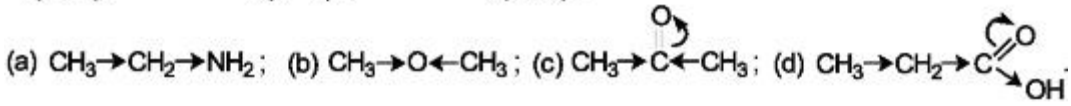
3.1. The compounds (a) and (c).

3.2. The conjugation in (b), no conjugation in (a), (c) and (d).

3.3. All the three resonance structures are:

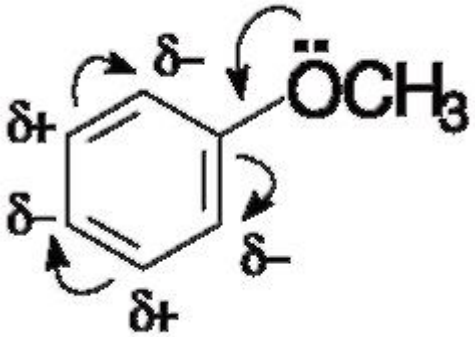


3.4.



An additional orbital overlap between oxygen of the hydroxyl group and C=O bond in (d) will be discussed later.

3.5. Electron density on all carbons of the benzene ring is increased due to the +M effect of the methoxyl group:



3.6. Only (b) and (e).

3.7. All, except for (d), are π,π-conjugated.

3.8. The compounds (a) and (d) are π,π-conjugated. The central carbon in the cumulated diene (c) is sp-hybridized and participates in formation of two π bonds that are mutually perpendicular oriented (remember *sp* hybridization of carbon). This does not result in orbital overlap of the two π bonds.

3.9. The compounds (a) and (d) are p,π-conjugated, the compound (c) is π,πconjugated.

3.10. Only (a) and (d) are conjugated: (a) is π,π-conjugated, (d) is p,π-conjugated.

3.11. All, except for (e).

3.12. All, except for (b). There are no aromatic compounds.

3.13. Electron-withdrawing: all, except for (f), which is neither donating nor withdrawing group.

3.14. Electron-donating: (a), (c), and (d); electron-withdrawing: (b), (e), and (f).
3.15. In (a), (c), and (f). The chlorine atom in (b) is electron-withdrawing substituent in spite of its $+M$effect.

3.16. In (b), (e), and (f). Both substituents in (c) are electron-withdrawing.

3.17. The stronger out of two electron-withdrawing substituent is a carboxyl group (which is the weaker?). There are two strong electron-donating groups in the compound (a). Donating property of an amino group (due to its strong $+M$ effect) is much more than those of a methyl group. Thus increasing order is the following: (d) < (e) < (c) < (f) << (b) < (a).

To Chapter 4

4.1. (a) Substitution reaction; (b) and (e) addition reactions; (c) elimination reaction; (d) oxidation reaction. Reactions (b) and (e) may also be classified as reduction reactions, but reaction (c) is never classified in organic chemistry as oxidation.

4.2. Reaction (1) is electrophilic substitution reaction (the $S_E$mechanism); reaction (2) is nucleophilic substitution reaction (the $S_N$mechanism).

4.3. (a) Elimination reaction; (b), (c), and (d) addition reactions; (e) oxidation reaction. Reactions (c) and (d) may also be classified as reduction reaction, and reaction (e) as addition reaction.

4.4. Electrophiles are (b) and (e), nucleophiles are (c) and (d). Methane is neither electrophile nor nucleophile.

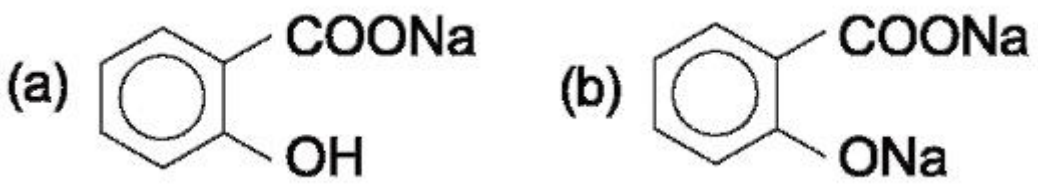
To Chapter 5

5.1. K_a44 corresponds to a negative value of pK_a(namely -1.6). Thus, HNO_3 is the strongest acid (its $pK_a$is the smallest); phenol is the weakest (its pK_a is the largest).

5.2. The iodide ion is much more stable because of its higher polarizability. For further explanation see the text.

5.3. (d) < (a) < (c) < (b). See the reaction in the text.

5.4. As a result of stronger acidity of the COOH group, the products are monosodium salt (a) and disodium salt (b), respectively:



5.5. (b) < (c) < (a) < (d). Any aliphatic amine is stronger as a base than ammonia.

5.6. (b) < (a) < (d) < (c).

p-Nitrophenol forms p-nitrophenoxide, p-$O_2NC_6H_4OM$+, in the reaction with a base MOH.

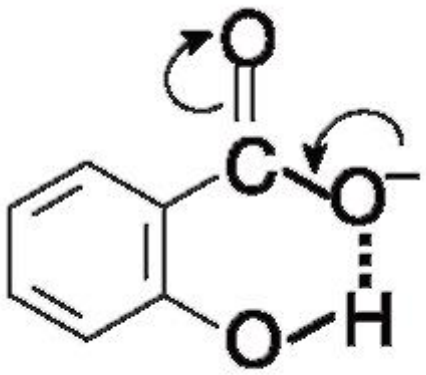
5.7. (d) < (b) < (a) < (c).

Cyclohexylamine reacts with an acid to form cyclohexylammonium salt, $C_6H_{11}NH_3^+X$.

5.8. *Method A.* Add excess aqueous NaOH to the mixture to extract phenol as a sodium salt. Acidify the alkaline solution to precipitate free phenol. (Don't use dilute solutions because of moderate solubility of phenol in water.) Aniline may be isolated from the benzene solution after evaporation of the solvent.*Method B.* Add excess HCl to the mixture to extract aniline as anilinium chloride. Complete a procedure without assistance by analogy with *Method A.*

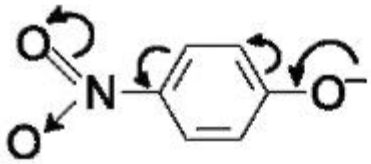
5.9. Use the procedure described in the Answer to Problem 5.8 *(Method* A). Only butanethiol will react with sodium hydroxide.

5.10. p-Hydroxybenzoic acid is weaker than benzoic acid due to the electron-donating effect of the hydroxyl group. On the contrary, a conjugate base of its *ortho* isomer (the salicylate ion) is stabilized better than the benzoate ion for the reason of intramolecular hydrogen bonding with the neighbouring hydroxyl group.

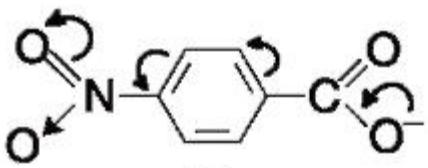


5.11. 4-Nitrophenol is a stronger acid than phenol, as well as 4-nitrobenzoic acid is stronger than benzoic acid itself. This is a result of electron-withdrawing effect of the NO_2 group that stabilizes the corresponding conjugate bases, phenoxide and benzoate ions, respectively.

However the influence of the NO_2 group is different in these nitro derivatives. In the nitrophenoxide ion (a), a negative charge is delocalized through a long chain of conjugation, whereas in the nitrobenzoate ion (b) the nitro group helps in charge delocalization only inductively, i. e. less effectively.



(a)



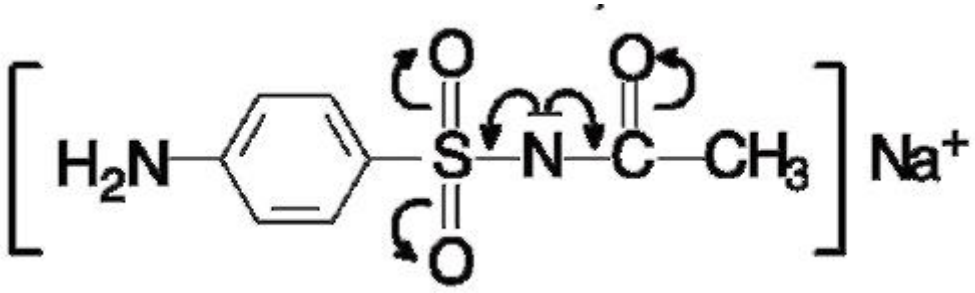
(b)

5.12. The most basic site is the nitrogen atom of the aliphatic part of the molecule. The next on basicity strength is the amino group attached to the benzene ring. Both

oxygen atoms are very weak basic sites, and the aromatic ring shows negligible basicity. Certainly, the aliphatic nitrogen participates in the salt formation.

5.13. The two most basic sites are the same as in the preceding example. Then basicity decreases in the following way: oxygen of the amido group > nitrogen of the amido group > the aromatic ring. Nitrogen of the aliphatic part of the molecule is protonated in the first place.

5.14. Significant acidity of the compound (higher, by the way, than that of phenol) is due to the presence of the fragment -SO_2NHCO-. A negative charge in the conjugate base is effectively delocalized through conjugation (see the structure of the salt).



The CHand NH-acidic sites of the groups $CH_3CO$and NH_2, respectively, are much weaker.

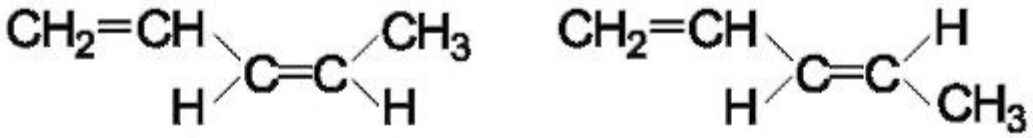
To Chapter 6

6.1. Both conformations are very similar to those of ethane, except for the replacement of one hydrogen with the methyl group.

6.2. The most stable is a chair conformation with equatorial position of the OH group.

6.3.

6.4. (b), (c), (e), and (f). Stereoisomers for the compound (f) are:

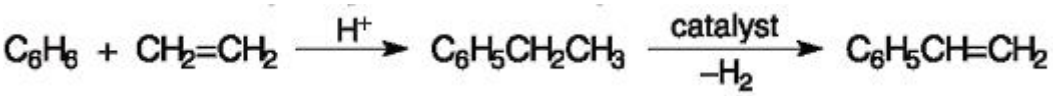


cis-1,3-pentadiene

trans-1,3-pentadiene

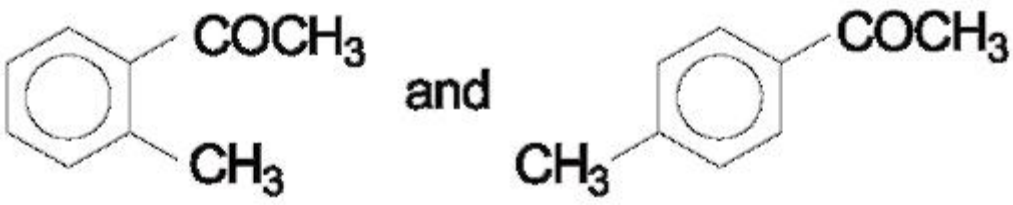
6.5. (a) 2-methyl-2-butanol; (b) 1-methylcyclohexanol.

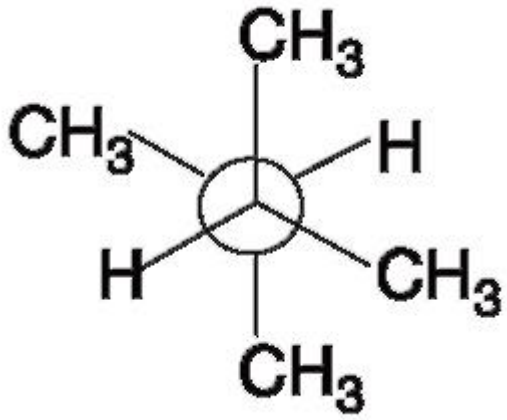
6.6. Ethylene is used for alkylation in the first step to produce ethylbenzene, which then will be dehydrogenated into styrene:



6.7. (a) < (d) < (b) < (c). Aniline is brominated with bromine-water to yield 2,4,6-tribromoaniline.

6.8. The products are: (a) a mixture of o- and p-chlorocumene (recall that cumene is isopropylbenzene); (b) a mixture of aromatic ketones, 2'-methyl- and 4'-methylacetophenone, respectively:





6.9.

6.10. The reaction can be performed only as free radical substitution; the product is bromocyclohexane.

6.11. (c) > (b) > (a). The compound (c) is more reactive than (b) because of two electron-donating substituents. The products are:

(a) 3-chloropropanoic acid (the anti-Markovnikov's product);

(b) 2-chlorobutane;

(c) isopropyl chloride (the Markovnikov's product).

6.12. Draw a tertiary carbocation in reaction (a) and two carbocations approximately equal in energy, which are formed in reaction (b). Therefore the products are:

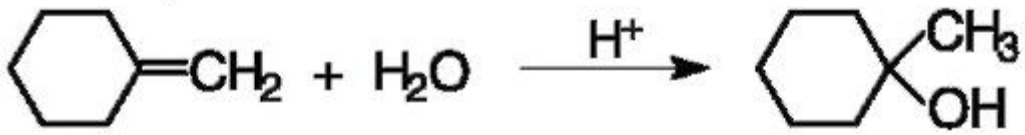
(a) 1-ethylcyclopentanol;

(b) a mixture of 2-bromopentane and 3-bromopentane.

6.13. Start with 3-methyl-1-butene.

6.14. Using the acid-catalyzed hydration reaction, start with: (a) 1-methylcyclohexene or (b) methylenecyclohexane (methylene is a divalent radical

>CH_2). For example:



6.15. 1-Chloro-2-butene (a major product at room temperature) and 3-chloro-1butene.

6.16. The products are:

(a) a mixture of *o*- and *p*-chlorotoluene (electrophilic substitution products);

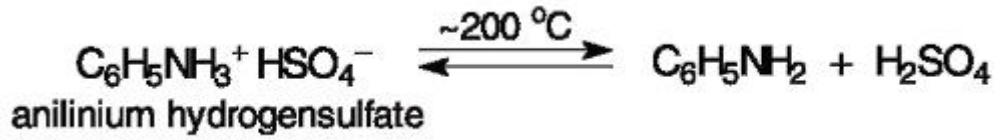
(b) bromophenylmethane, or benzyl bromide (a radical substitution product).

6.17. Use the following reaction sequences:

(a) bromination and separation of *p*-bromotoluene, then oxidation of the latter;

(b) the Friedel-Crafts alkylation, then nitration of toluene produced, and, finally, oxidation of m-nitrotoluene.

6.18. Aniline immediately forms a salt in the reaction with H_2SO_4 which is non-reactive because of a strong electron-withdrawing NH_3^+ substituent. Sulfonation proceeds only if this salt dissociates at high temperature with releasing a free base:



To Chapter 7

7.1. The products are: (a) benzyl alcohol; (b) $CH_3COOC_2H_5$ (ethyl acetate); (c) 1-butanethiol; d) tetramethyllammonium chloride.

7.2. (a) 3-Propen-2-ol; (b) 2-mercaptoethanol; (c) cyclohexanol; (d) 2-isopropylphenol; (e) 4-hydroxybenzoic acid. The compounds (a) and (c) are secondary alcohols; the compound (b) is a primary alcohol (and a primary thiol).

7.3. Both reactions are elimination reactions. According to the Zaytsev's rule, the main products are: (a) 2-methyl-2-butene; (b) 2-butene.

7.4. The products are:

(a) $CH_3CH_2CH_2CHO$, butanal (or butyraldehyde);

(b) $(CH_3)_2CHCHO$, 2-methylpropanal (or isobutyraldehyde);

(c) $CH_3CH_2COCH_3$, butanone;

(d) no reaction.

7.5. A bromide ion and, especially, an iodide ion are better nucleophiles than a chloride ion.

7.6. The product is anilinium hydrogensulfate, $C_6H_5NH_3$+ HSO_4

7.7. The products are:

(a) diethylmethylammonium iodide, $[(C_2H_5)_2NHCH_3]+ I^-$;

(b) tributylmethyammonuim iodide, $[(C_4H_9)_3NCH_3]+ I^-$;

(c) diethylmethylsulfonium iodide, $[(C_2H_5)_2SCH_3]+ I^-$.

7.8. The following reagents should be used in reactions:

(a) aqueous sodium hydroxide;

(b) alcoholic potassium hydroxide;

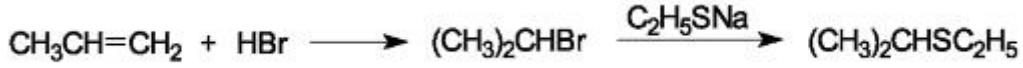
(c) excess methylamine and subsequent treatment of the product with an alkali.

7.9. The benzyl cation formed from benzyl chloride in nucleophilic substitution reactions is well stabilized through conjugation. Benzyl chloride can be hydrolyzed with boiling water, whereas alkyl chlorides are hydrolyzed only with alkalis.

7.10. Add bromine to allyl chloride, and subject the resulting 1,2-dibromo-3-chloropropane to alkaline hydrolysis.

7.11. Different resistance to hydrolysis of the two chlorides may be used. Chlorobenzene is quite stable on boiling with water, whereas benzyl chloride is hydrolyzed under the same conditions liberating a chloride ion. The latter may be detected in a usual manner, for example, by the reaction with Ag^+.

7.12. A possible synthetic pathway is:



7.13. Firstly, convert tert-butyl alcohol into tert-butyl bromide. Use then ethanol as a weak nucleophile in the final step:

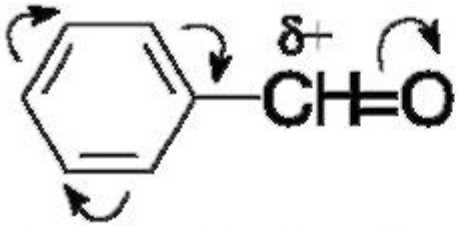
Explain why stronger nucleophilic reagents such as sodium ethoxide might not be used in this reaction.

7.14. Recall that electron-withdrawing substituents decrease basicity of a compound. This regards the second amino group of ethylenediamine that shows -/ effect with respect to the first one. Thus, the diamine is the weaker base.

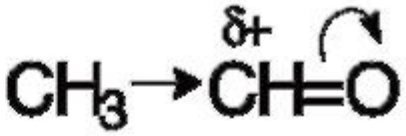
To Chapter 8

8.1. First, π,π conjugation of the carbonyl group and the benzene ring in benzaldehyde decreases δ+ on the carbonyl carbon. The donating inductive effect of the CH_3 group is less pronounced than the mesomeric effect of the benzene ring.

Second, the bulky phenyl group prevents the carbonyl carbon from nucleophilic



benzaldehyde

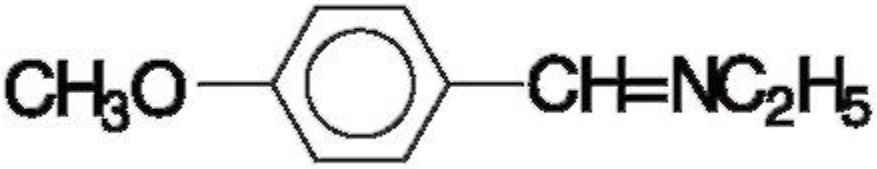


acetaldehyde

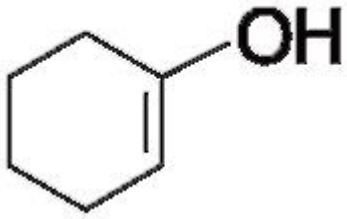
attack.

8.2. This is a familiar reaction of acetal formation from propanal and excess methanol in the presence of an acid catalyst.

8.3. The product is an imine, or a Schiff's base: p-Methoxylbenzaldehyde is less reactive because the methoxy group acts as an electron-donating substituent *(+M* effect) and decreases δ+ on the carbonyl



carbon.

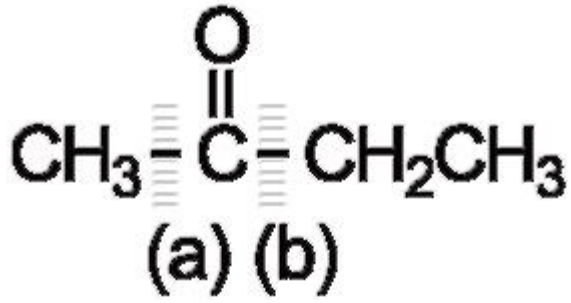


8.2. CH2=CHOH; (b) CH2=CH(OH)CH3; (c)

8.5. A catalyst converts an aldehyde molecule into a strong nucleophile (a carbanion) that attacks then the carbonyl carbon of the non-ionized aldehyde. The product is 3-hydroxy-2-ethylhexanal.

8.6. All, except for (b) and (d).

8.7. Cyclohexanone gives the same dicarboxylic acid whichever C-C bond next to the carbonyl group is cleaved. There are two sites of cleavage in butanone, (a) and (b). As a result, three different acids are produced: CH_3CH_2COOH + HCOOH



in the mode (a) and 2 moles of CH_3COOH in the mode (b).

8.8. The order of a reactivity increase is: (b) < (a) < (c). For explanation compare partial charges (δ+) on the carbonyl carbon in each compound.

8.9. Anhydrous protic acids (such as gaseous HCl) or Lewis acids are used as catalysts in reactions (a) and (b). The products are:

(a) $(CH_3)_2CHCH(OC_3H_7)_2$, an acetal;

(b) $C_6H_5CH(OH)OC_2H_5$, a hemiacetal;

(c) $(CH_3)_2C{=}NCH(CH_3)_2$, a substituted imine.

8.10. The products are:

(a) $C_6H_5CH_2CH{=}NNHC_6H_5$, a phenylhydrazone;

(b) $C_6H_5CH_2CH(OCH_3)_2$, an acetal;

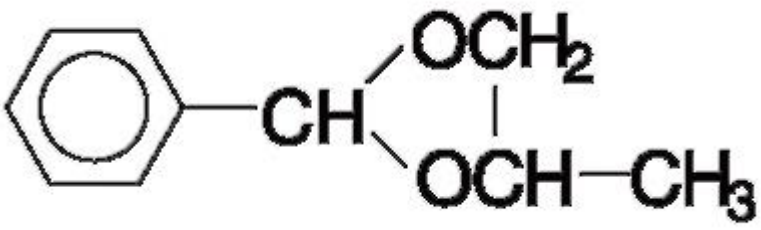
(c) $C_6H_5CH_2COOH$, a carboxylic acid.

8.11. Phenylhydrazine acts as a nucleophile by the $A_N$mechanism. The product is the phenylhydrazone, $CH_3CH_2CH{=}NNHC_6H_5$.

8.12. The acetal $(CH_3)_2CHCH_2CH(OC_2H_5)_2$, was subjected to acidic hydrolysis to give the mentioned products.

8.13. The product is a cyclic acetal

8.14. Only the compound (a).



8.15. The product is a hydroxy ketone $HOCH_2CH_2C(O)CH_3$ since only acetone possesses α-hydrogens.

8.16. The products are: (a) ethanol; (b) benzoic acid; (c) no reaction.

8.17. The products are: (a) cyclopentanol; (b) 1-butanol; (c) lactic acid.

8.18. The products are: (b) butyric (butanoic) acid; (c) propenoic (acrylic) acid; no reaction for (a) and (d).

To Chapter 9

9.1. (a) 3-Phenylpropenoic acid; (b) acetyl chloride; (c) chloroacetic acid; (d) methyl formate.

9.2. (c) > (a) > (d) > (b). The product is ammonium *p*-nitrobenzoate.

9.3. Consider a mechanism of nucleophilic substitution at sp^2-hybridized carbon. For equilibrium shifting apply one of Le Chatelier's principles, namely: i) use one of the reagents in excess; ii) remove any product formed (usually water can easier be removed).

9.4. Use the Fischer esterification for the one-step preparation of methyl salicylate.

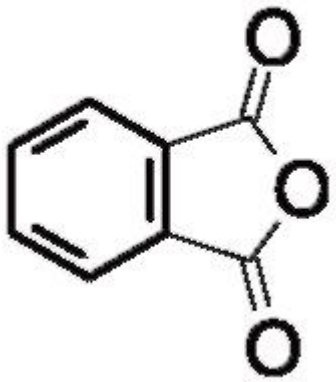
9.5. The products are:

(a) $CH_3COONa + CH_3OH$;

(b) $CH_3CONHCH_3 + CH_3COOH$;

(c) $CH_3COOK + C_2H_5NH_2$.

9.6. As it is said in the text, phthalic acid forms phthalic anhydride:



9.7. (a) $CH_3CH_2CH=CHCOOH$; (d) $CH_3CH_2CH_2CONHCH_3$;

(b) $C_2H_5OOC\text{-}COOC_2H_5$; (e) $C_6H_5CO\text{-}O\text{-}COC_6H_5$;

(c) $HCON(CH_3)_2$; (f) $(CH_3)_2CHC(O)Cl$.

9.8.

(a) 4-Phenyl-2-butenoic acid;

(b) isopropyl phenylacetate;

(c) dimethyl pentanedioate (or glutarate);

(d) acrylamide or propenamide;

(e) propanoic (or propionic) anhydride;

(f) bromoroacetyl chloride.

9.9. In contrast to carboxylic acids, their esters cannot form intermolecular hydrogen bonds.

9.10. The products are:

(a) $CH_3CH_2COOCH(CH_3)_2$ - isopropyl propionate (or propanoate);

(b) CH_3CH_2COONa - sodium propionate;

(c) $CH_3CH_2COONH_4$ - ammonium propionate;

(d) CH_3CH_2COOK - potassium propionate (not an esterification product).

9.11. The products are:

(a) $CH_3CH_2CH_2COONa$ - sodium butyrate, methanol, and CO_2;

(b) $CH_3CH_2CH_2CON(CH_3)_2$ - N,N-dimethylbutyramide and methanol;

(c) $CH_3CH_2CH_2COOC_2H_5$ - ethyl butyrate and methanol.

9.12. Only (a) and (d).

9.13. The products are: (a) ammonium benzoate; (b) 3-bromobenzoic acid. The latter is stronger than benzoic acid because bromine is an electron-withdrawing substituent.

9.14. The anti-Markovnikov addition product, $BrCH_2CH_2COOH$. This is a result of double bond polarization in acrylic acid due to -*M*-effect of the carboxyl group.

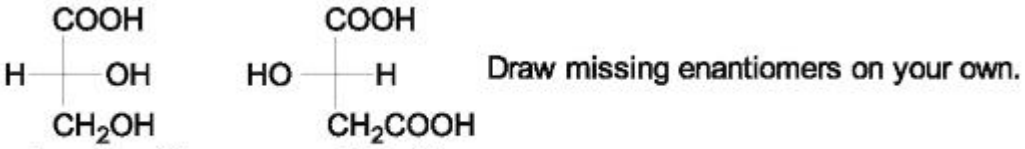
To Chapter 10

10.1. (a), (d), and (e).

10.2. (a) atom C-2; (b) atoms C-2 and C-3; (c) achiral compound; (d) atoms C-1 and C-2.

10.3. Find carbon having four different substituents in the adrenalin and dopa molecules. In the thalidomide molecule this is one of carbons in the six-membered heterocycle, which one?

10.4.



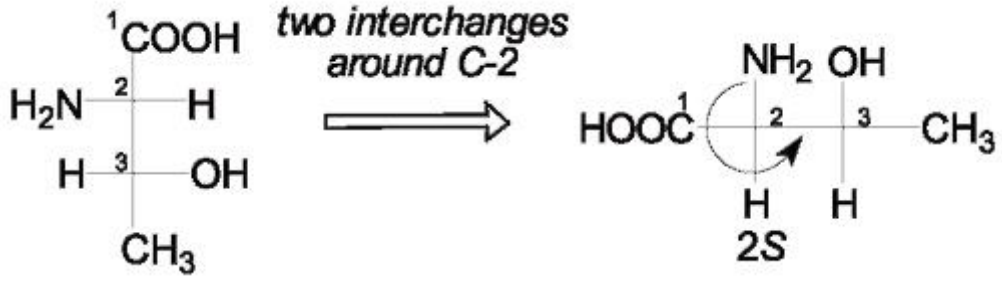
10.5.

(a) -Cl > $-CH_2OH$ > $-CF_3$ > -H;

(b) -OH > $-NH_2$ > - CH_2SH > -COOH (one S is preferred to three O);

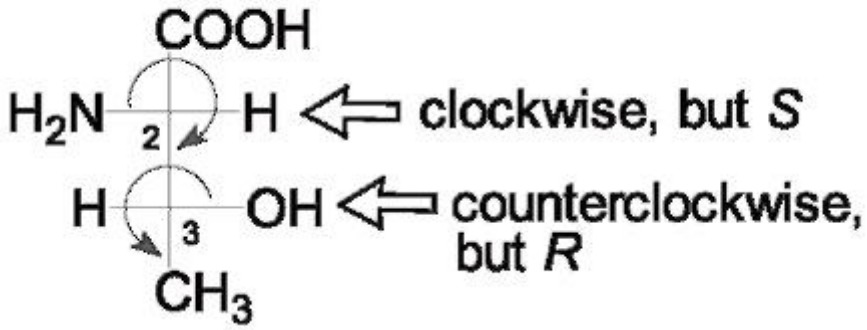
(c) -SH > $-OCH_3$ > $-NHCOCH_3$ > -CHO.

10.6. After two interchanges around C-2, it becomes clear that C-2 has the *S* configuration:



Repeat the same procedure at the next chiral centre to get the 3*R* configuration. Thus L-threonine has the 2S,3R configuration.

There is, however, a simpler way for the *R,S* assignment. A Fischer projection with a lowest priority group in a horizontal position (left or right) may be considered. But in this case the reverse directions should be applied, i. e. a clockwise priority decrease of the remaining groups signifies the *S*configuration, and *vice versa.* This method is especially useful for compounds with several chiral centres as it is exemplified in L-threonine (see the drawing).



10.7. Find two chiral centres in the compound (b) and only one in each remaining compound.

10.8. The compound (b) is achiral.

10.9. All, except for (a), (f), and (g). Two chiral centres are present in the latter but the molecule has a plane of symmetry; it is a *meso* compound.

10.10. The only interchange of any two substituents leads to inversion of configuration. Two interchanges do not alter the initial configuration. The Fischer projection may be turned by 180° in the plane of the paper (but not by 90°). (a) L-Alanine (2-aminopropanoic acid); (b) L-2-chlorobutyric acid; (c)L-glyceraldehyde.

10.11. For the compounds (b) and (c) only: (b) was assigned arbitrary to the D series; (c) is a racemic mixture, i. e. a mixture of the D- and L-enantiomers.

10.12. The *R* configuration, according to priority decreasing -OH > $-CH_2NHCH_3$ > > the benzene ring. The d,l system is helpless here.

10.13. For pairs (a) and (c) only:

(a) L-alanine is an optically active compound, whereas β-alanine is not; (c) (+)- and (-)-glyceraldehydes were assigned arbitrary to the D and L series, respectively.

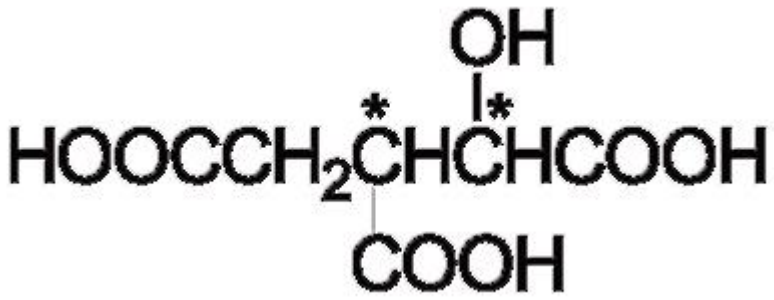
10.14. Draw the Fischer projections for:

(a) enantiomers of 2-butanol (there are no chiral ethers with the formula $C_4H_{10}O$);

(b) enantiomers of 2-bromopentane and 2-bromo-3-methylbutane.

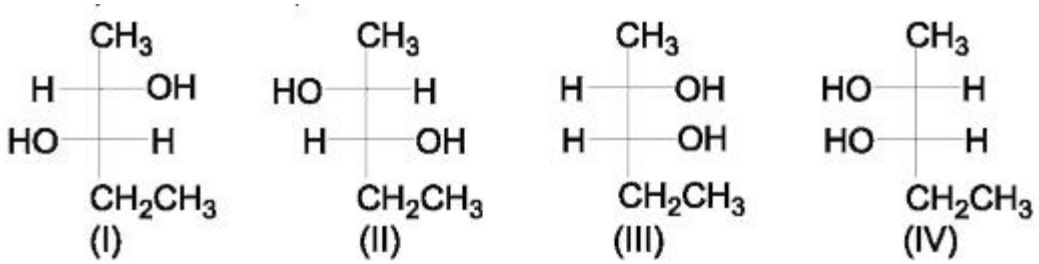
10.15. Draw the structure of citric acid and find two identical substituents at each carbon. Thus, citric acid is an achiral compound.

There are two chiral centres in a molecule of isocitric



acid (see folmula).

10.16.



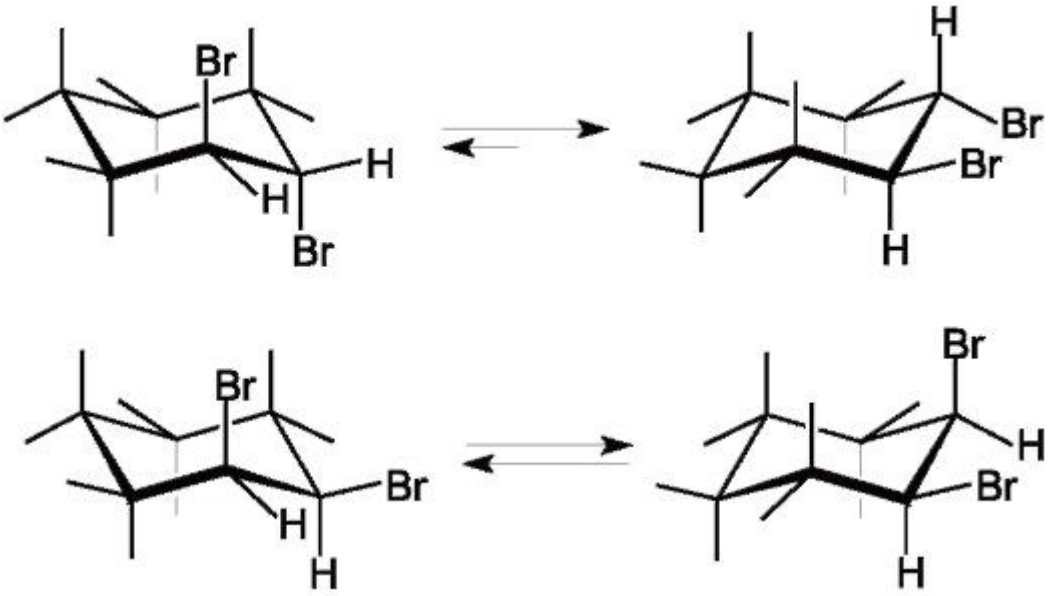
Enantiomers are the pairs (I) - (II) and (III) - (IV). Other pair combinations represent diastereomers. Note that neither (III) nor (IV) is a *meso* compound.

10.17. There are three chiral centres: C-1, C-3, and C-4. Eight stereoisomers are possible in all. The structure of (+)-menthol represents a mirror image of (-)-menthol. The two substances can be distinguished by polarimetry.

10.18. Only (a) as *cis-trans* isomers. Only (c), with a chiral atom C-3, as enantiomers. Draw the structures for the remaining *cis-trans* pairs on your own.

10.19. The *trans* isomer is more stable because in the most stable conformation both substituents are equatorial.

Both conformers of the *cis* isomer are of equal stability. Thus one bromine atom must be axial.



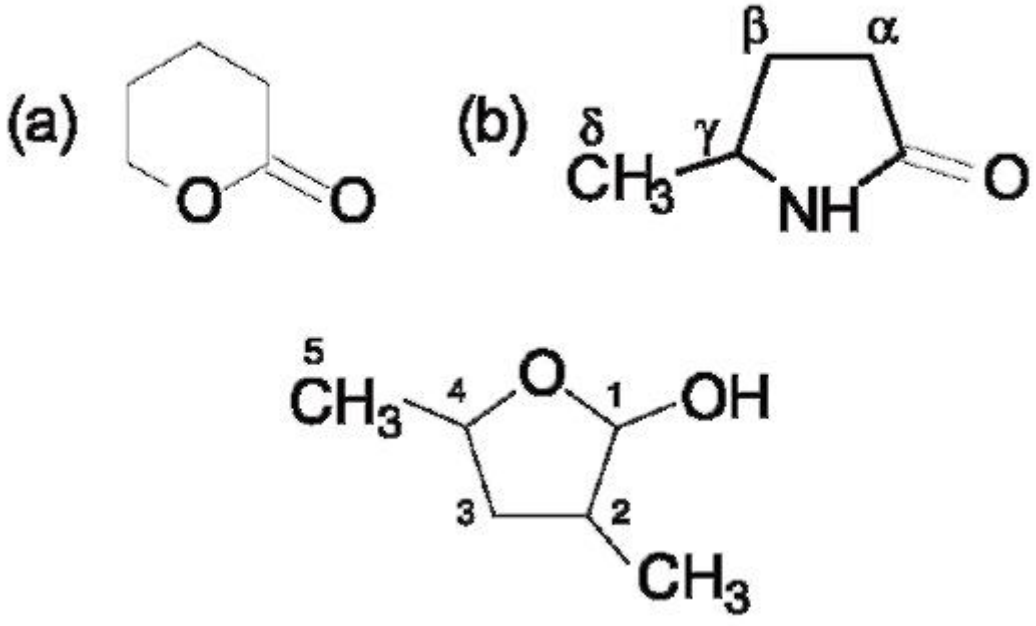
To Chapter 11

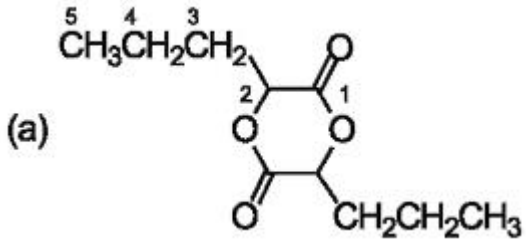
11.1. The compounds (a) and (c) are polyfunctional; the compound (b) is heterofuctional.

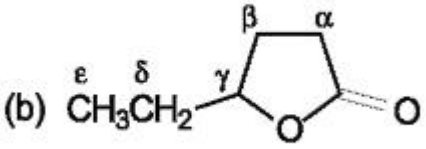
11.2. The lactone (a) is the product of cyclization of 5-hydroxyvaleric acid; the lactam (b) can be produced from 4-aminovaleric acid.

11.3. The structure of the product formed by intramolecular nucleophilic substitution is shown. It is predominant (about 88%) in the equilibrium mixture.

11.4. α-Hydroxy acids undergo intermolecular cyclization, whereas γ-hydroxy acids undergo intramolecular one. The most stable (sixor five-membered) cyclic ester is formed in each case, namely, the dilactide (a) from the former and the γ-lactone (b) from the latter (numbering in the parent hydroxy acids is shown below).



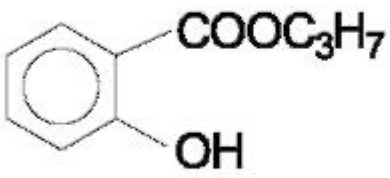




11.5. Only compounds (c) and (d).

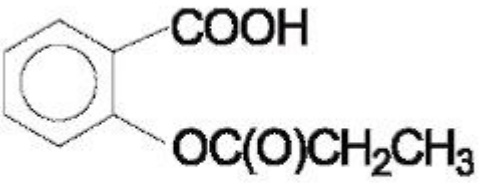
11.6. The compounds (d) and (f) are polyfunctional; the compounds (b) and (c) are heterofuctional.

11.7. The ester products are, for example (see structures). The former is produced by participation of the carboxyl group; the latter is formed by acylation of the phenolic hydroxyl group.



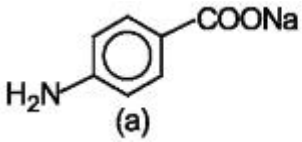
propyl salicylate

and

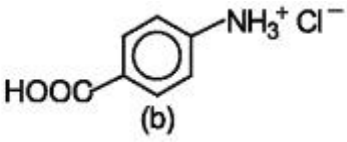


O-propanoylsalicylic acid

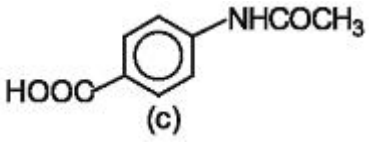
11.8. The products are:



(a)

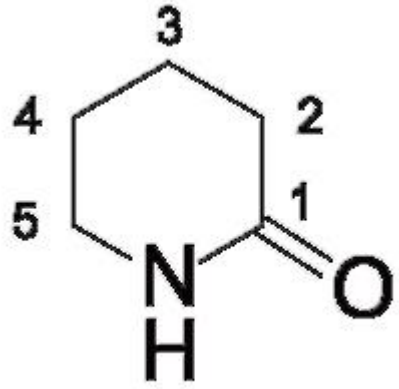


(b)



(c)

The compound (c) is formed by a nucleophilic acyl substitution mechanism. 11.9. A stable cyclic amide (see formula of δ-valerolactam) is formed only from δ- and γ-amino acids. β-Amino acids undergo elimination of ammonia to form 2-pentenoic acid.



11.10. The mentioned compounds can be formed on heating of 4-hydroxybutyric and 4-aminovaleric acids, respectively. Both the lactone and lactam can be hydrolyzed in both acidic and, better, alkaline media.

11.11. ε-Amino acids cannot form cyclic amides because of low stability of sevenmembered rings as compared with sixand five-membered ones. A linear polyamide $[-NH(CH_2)_5CO-]_n$, known as Nylon 66, is produced in this case.

11.12. Hydrogen bonding with water molecules prevents stabilization of the enol form of acetoacetic ester.

To Chapter 12

12.1. The products are:

(a) $C_{17}H_{33}COONa$, sodium oleate;

(b) $CH_3(CH_2)_7CHBr\text{-}CHBr(CH_2)_7COOH$, 9,10-dibromostearic acid;

(c) $CH_3(CH_2)_7CH(OH)\text{-}CH(OH)(CH_2)_7COONa$, sodium 9,10-dihydroxystearate.

12.2. The main fatty acid component of olive oil is a $C_{18:1}$ acid, whereas corn oil contains considerable amounts of polyunsaturated acids ($C_{18:2}$ and $C_{18:3}$).

12.3. The products are:

(b) $C_{17}H_{35}COOH$, stearic acid;

(c) $CH_3(CH_2)_4CHBrCHBrCH_2CHBrCHBr(CH_2)_7COOH$, 9,10,12,13-tetrabromostearic acid;

(d) $(C_{17}H_{33}COO)_2Mg$, magnesium oleate (a water-insoluble salt); (a) and (e) no reaction.

12.4. The products are: glycerol, phosphate, choline, and salts of oleic and stearic acids.

12.5. (a) and (e) are waxes; (b) is a fat; (c) is an oil.

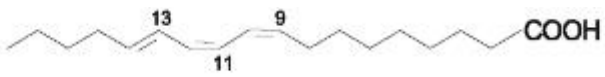
(a) $CH_3(CH_2)_{13}COOC_{16}H_{33}$

(b) $CH_2OCOC_{15}H_{31}$
$CHOCOC_{15}H_{31}$
$CH_2OCOC_{15}H_{31}$

(c) $CH_2OCOC_{13}H_{33}$
$CHOCOC_{17}H_{33}$
$CH_2OCOC_{17}H_{35}$

12.6. There is no logical connection. Margarinic acid occurs neither in margarine nor in vegetable oil that is subject to hydrogenation.

12.7. All the C=C double bonds in punicic acid are π,π-conjugated, one of them has the *E* configuration:



On the contrary, the double bonds in linoleic and linolenic acids are not conjugated and all of them have the *Z* configuration.

12.8. Use bromination reaction with measuring of bromine consumed. Bromine consumption for linoleic acid is near twice as much compared with that of oleic acid.

12.9. All, except for (a), can be hydrolyzed. Glycerol will be formed only from (b) and (d).

The hydrolysis products are:

(b) glycerol and three fatty acids (as salts);

(c) 1-*O*-pentadecylglycerol and two fatty acids (as salts);

(d) glycerol and phosphate.

12.10. Saponification of a wax gives rise to a long-chain alcohol and a carboxylate salt. The latter is a soap, by the way. The former may be treated with sulfuric acid to yield an alkyl hydrogensulfate $ROSO_3H$, which, after neutralization, leads to a salt, for example, $ROSO_3^-\ Na^+$.

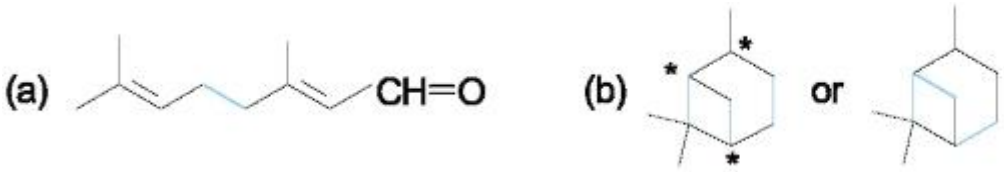
12.11. Alkaline hydrolysis of a triacylglycerol gives rise to water-soluble products: glycerol and salts of fatty acids. If a wax is subjected to saponification, one of the products (a long-chain alcohol) is insoluble.

12.12. Draw general formulas for each compound and divide each molecule into polar and nonpolar (long-chain) portions.

12.13. The products are: glycerol, phosphoric acid, colamine, and two fatty acids in a form, depending on pH.

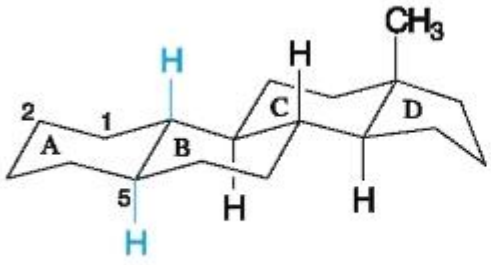
To Chapter 13

13.1. The joint positions are given in colour:

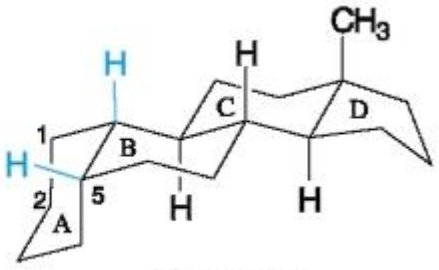


The *cis* and *trans* isomers are possible for the compound (a); the *trans* isomer is shown. Three chiral centres are in the compound (b) shown with asterisks.

13.2.



5α-estrane



5β-estrane

13.3. They have the same structure.

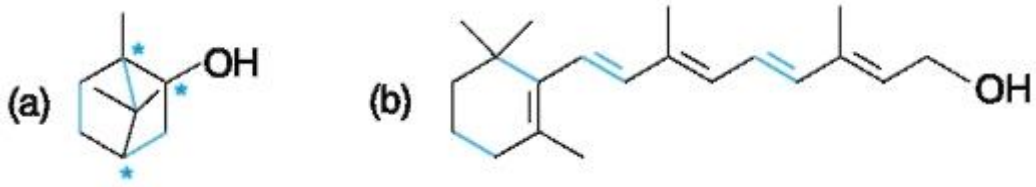
13.4. The products are:

(a) estradiol diacetate (both hydroxyl groups are esterified);

(b) a sodium salt (only at the phenolic OH group);

(c) 2,4-dibromo derivative (substitution at *ortho* positions to the phenolic hydroxyl group).

13.5. The joint positions are given in colour:



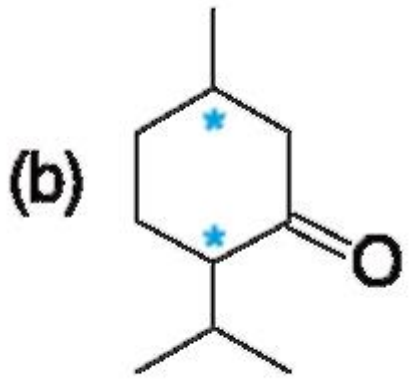
Three chiral centres are in the compound (a) shown with asterisks. Compound (b) represents the all-trans isomer. Various combinations for
the *cis* and *trans* configuration of the four double bonds in the chain are possible.

13.6.

(a) Retinal is an aldehyde therefore it gives the positive Fehling's and Tollens' tests, it forms a crystalline hydrazone derivative;

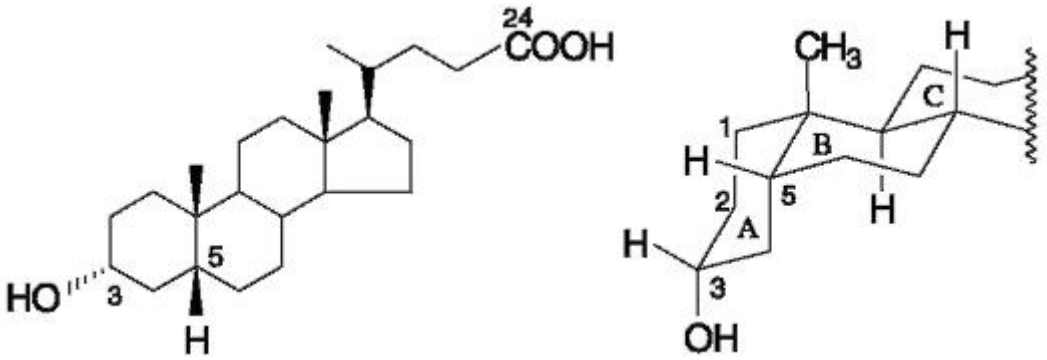
(b) α-pinene being unsaturated compound reacts with bromine water or permanganate by decoloration.

13.7. To distinguish between the compounds (a) and (b) use specific reactions of carbonyl compounds for menthone, for example, formation of crystalline hydrazone. Besides, menthol can be recognized by oxidation reaction with permanganate or chromic acid [recall changes in coloration after reduction of



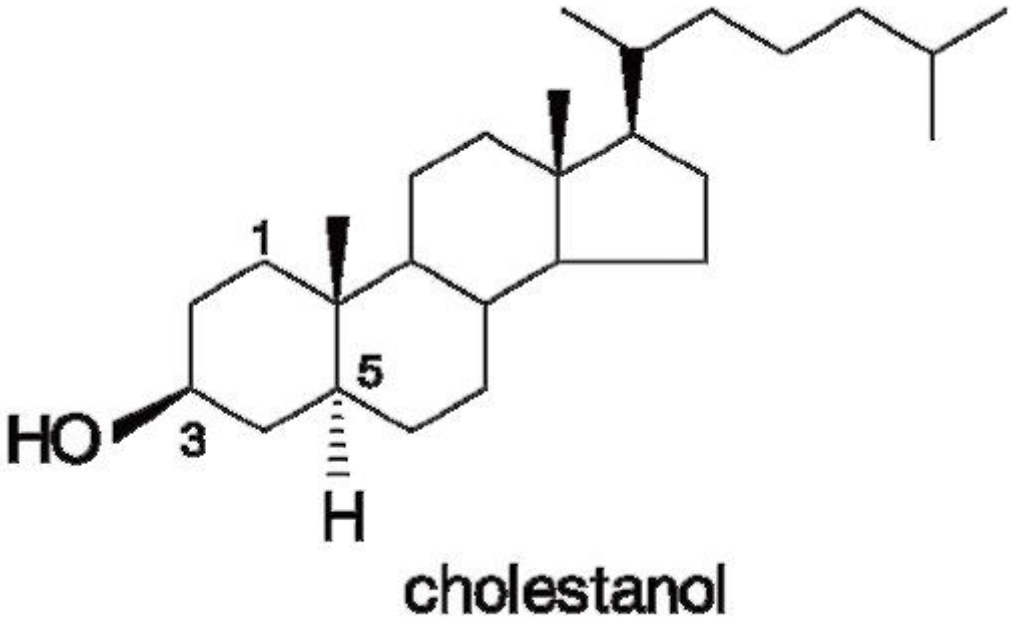
Mn(VII) and Cr(VI)]. Menthone is resistant to oxidation. Menthane is an optically inactive compound.

13.8. The rings A-B are *cis* fused; the rings B-C and C-D are *trans* fused. Looking at the conformational formula, define the position of the 3α-hydroxyl group.



13.9. Cholesterol is an unsaturated compound. Use qualitative tests for the detection of a C=C double bond, for example, reaction with bromine water or oxidation with potassium permanganate.

13.10. The products are:



cholestanol

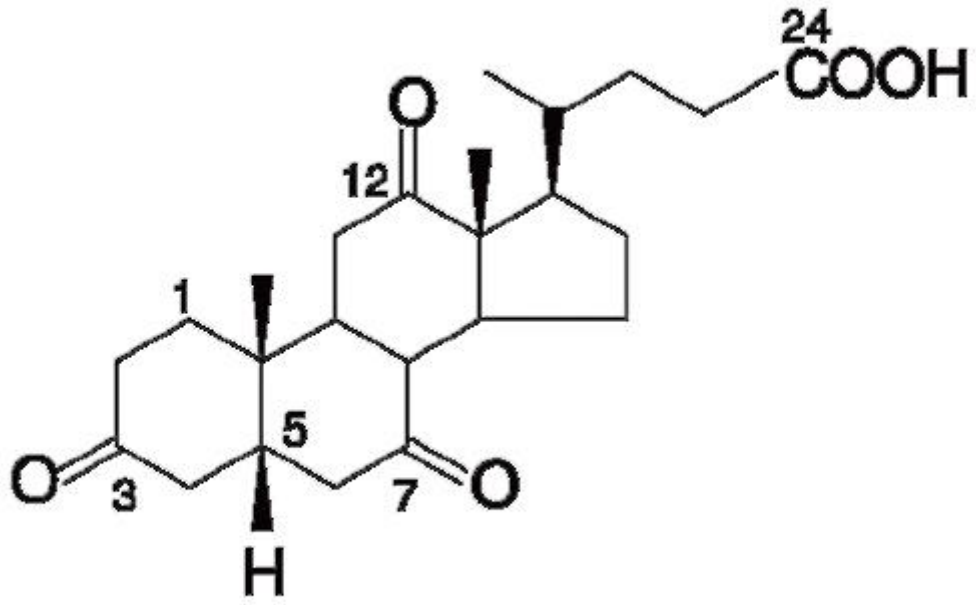
(b) a sodium salt (with participation of the phenolic hydroxyl group);

(c) dimethylacetal (with participation of the carbonyl group);

(d) phenylhydrazone (the same reactive site); (a) no reaction.

13.11. All the oxo groups will be reduced to hydroxyl groups, but three new chiral centres arise, C-3, C-7, and C-12. Thus, eight (2^3) stereoisomers are possible, only

one from which will be cholic acid (the 3α,7α,12α isomer).



dehydrocholic acid

To Chapter 14

14.1. The epimer at C-3 is d-allose, the epimer at C-4 is d-galactose, and the epimer at C-5 is an l-aldohexose (the OH at C-5 at the left). Since l-sugars are absent in Figure 13.1 you have to draw the denantiomer (mirror image) of the desired l-sugar and find it in the figure. It will be d-idose, consequently the epimer of d-glucose at C-5 is l-idose.

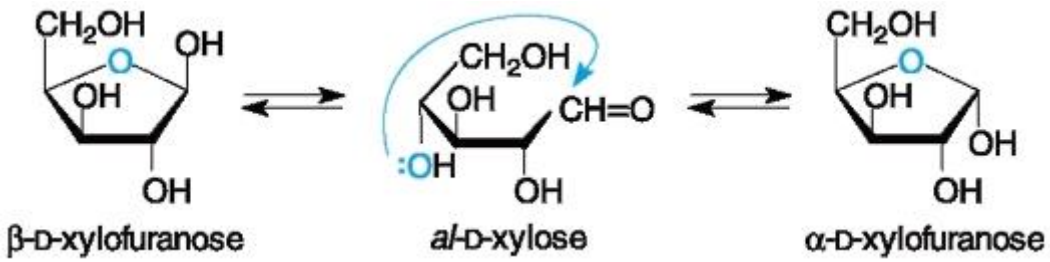
14.2. They are:

(a) diastereomers, epimers at C-2;

(b) diastereomers, not epimers because they differ in configuration at both C-2 and C-4;

(c) enantiomers.

14.3. The following scheme is half finished. Complete it by writing pyranose forms.

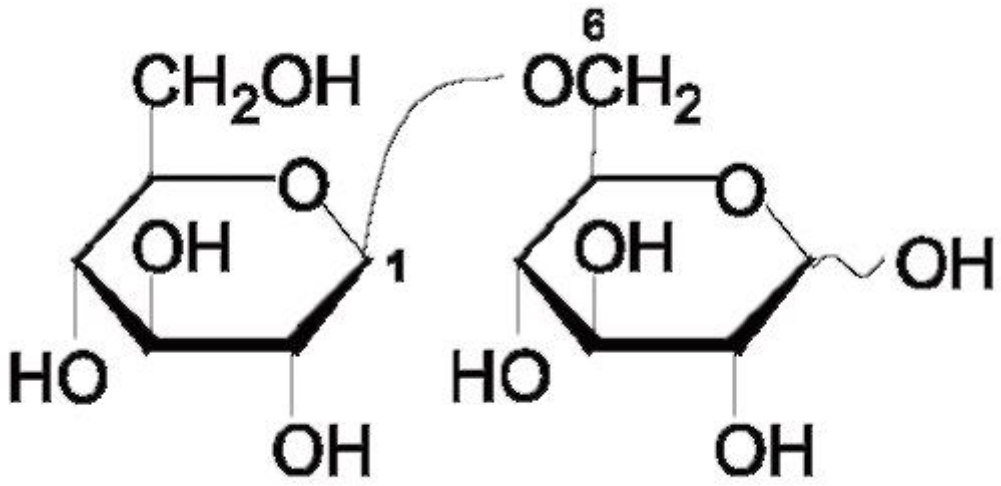


14.4. All, except for (a), since they contain a potential aldehyde group - hemiacetal hydroxyl (draw the structural formulas for each mentioned compound).

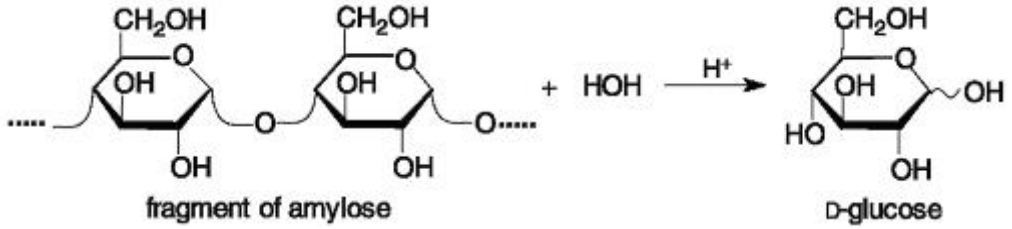
14.5. The tetrose d-erythrose (Fig. 14.2) or its enantiomer l-erythrose yield the same dicarboxylic acid.

14.6. Acid-catalyzed hydrolysis of glycosides yields a monosaccharide and an alcohol (or phenol).

14.7. Eight disaccharides in all. One of them, 6-O-(β-d-glucopyranosyl)-d-glucopyranose (common name gentiobiose), is represented. Draw, on your own, the structure for the second example, let α-d-Glcp-(13)-dGlcp.



14.8. Acid-catalyzed hydrolysis of amylose yields stepwise dextrins, maltose, and finally glucose (as a tautomeric mixture):



14.9. The xylitol molecule has a plane of symmetry that bisects the atom C-3, thus xylitiol is a *meso*compound.

14.10.

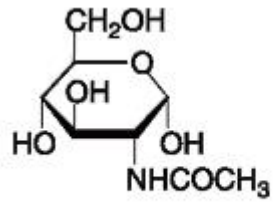
(a) Ribitol (neither d- nor l-; see comment at the answer to Problem 14.9);

(b) galactaric acid (also neither d- nor l-);

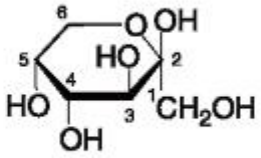
(c) phenyl α-d-galactopyranoside;

(d) 2-deoxy-β-d-ribofuranose 5-phosphate.

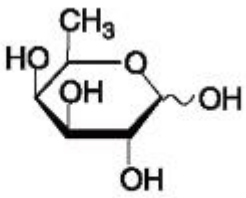
14.11. Only the Haworth formulas are shown:



(a) α-pyranose form

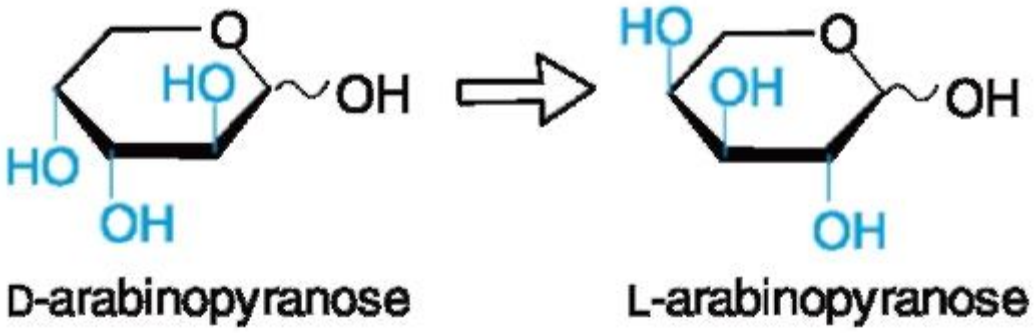


(b) β-pyranose form

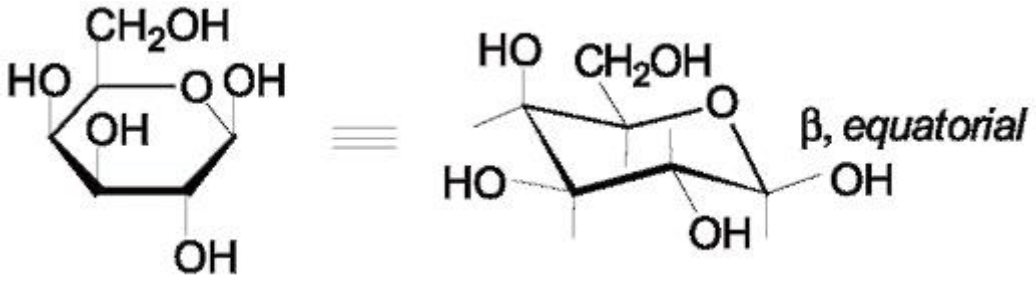


(c) both pyranose forms

(d) The simplest way to draw L-arabinopyranose is as follows: draw first its d enantiomer (Fig. 14.2), then invert *all* the chiral centers to make the initial l configuration again.



14.12. Equatorial position of the anomeric hydroxyl group is preferred, therefore the β anomer is more stable.



Any anomeric pair represents diastereomeric compounds that differ in all respects, namely, in physical, chemical, and biochemical properties.

14.13. Qualitative reactions are:

(a) reduction of the Tollens' and Fehling's (or Benedict's) reagents;

(b) reaction with copper(II) hydroxide.

14.14.

(a) An anomeric hydroxyl group only; the products are methyl α- and methyl β-D-mannopyranosides;

(b) all the hydroxyl groups. The product is the mixture of penta-O-acetyl-αand β-D-mannopyranose.

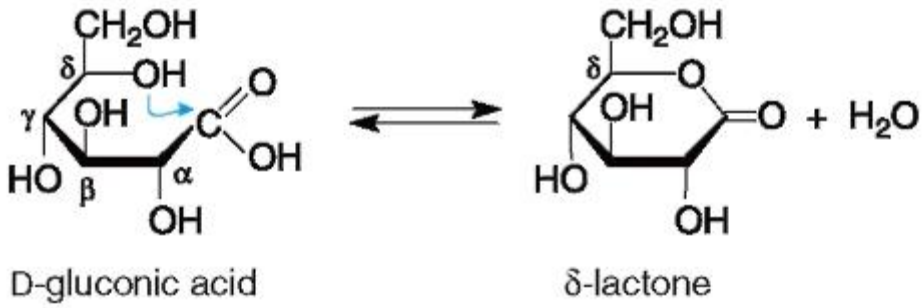
14.15. No product will be formed under stated conditions because of:

(I) alcoholic hydroxyl groups of monosaccharide do not react with an alcohol to form ethers;

(II) glycosides cannot be formed in the presence of water (hydrochloric acid contains at ~65% of water);

(III) even if anhydrous conditions would be used (i. e. gaseous HCl in methanol), not the β glycoside but an α,β mixture is formed.

14.16. D-Gluconic acid being hydroxy carboxylic acid undergoes cyclization to form γ- and δ-lactones.



D-gluconic acid δ-lactone

This equilibrium represents hydration-dehydration reaction but not tautomeric shift.

Draw, on your own, formation of a five-membered γ-lactone.

14.17. (b), (c), and (d).

14.18. β-d-Glucopyranose units linked by the (1 →3)-glycosidic bond.

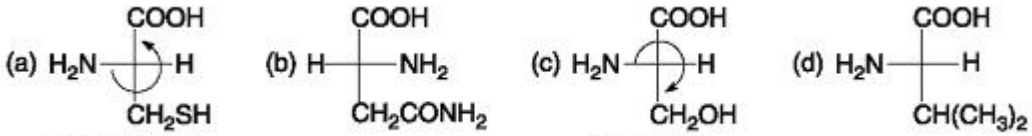
14.19. It is possible only for (a) since α-lactose mutarotates in a solution while sucrose does not.

14.20. Glycosidic bonds in cellulose are quite resistant to alkaline hydrolysis (recall properties of acetals). Cellulose acetate is an ester that will be hydrolyzed under alkaline conditions to give acetic acid and an alcohol, i. e. unsubstituted cellulose.

14.21. (c), (d), and (f).

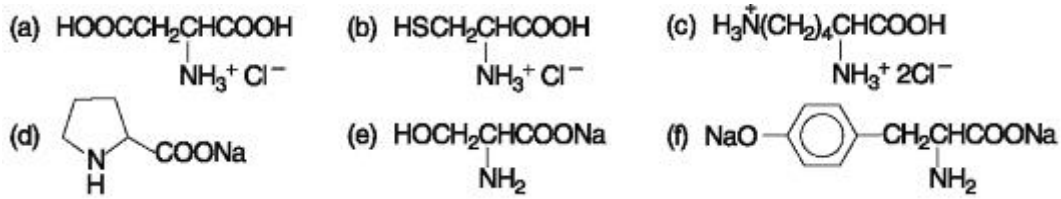
To Chapter 15

15.1.



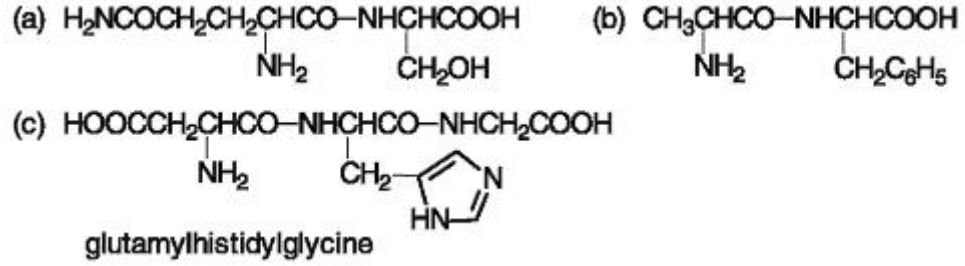
Explain, why clockwise decreasing of a priority results in the *S* configuration for the compound (c).

15.2. The products are:



15.3. The products are: (a) alanine methyl ester, hydrochloride; (b) N-hydroxymethylalanine; (c) lactic acid.

15.4. The products are glutamic acid and 2-oxopropanoic acid (pyruvic acid). 15.5.



15.6.

(a) 2-Aminopropanoic acid; (e) 2-amino-4-(methylthio)butanoic acid;

(b) aminobutanedioic acid; (f) 2-amino-3-phenylpropanoic acid;

(c) 2-amino-4-methylpentanoic acid; (g) 2-pyrrolidinecarboxylic acid;

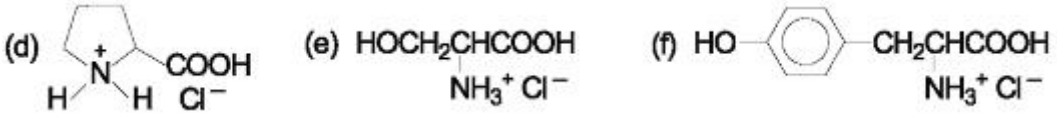
(d) 2,6-diaminohexanoic acid; (h) 2-amino-3-(4-hydroxyphenyl)-propanoic acid.

15.7. The products are:

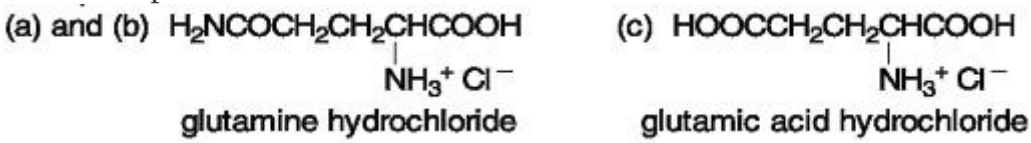
(a) lysine Na-salt;

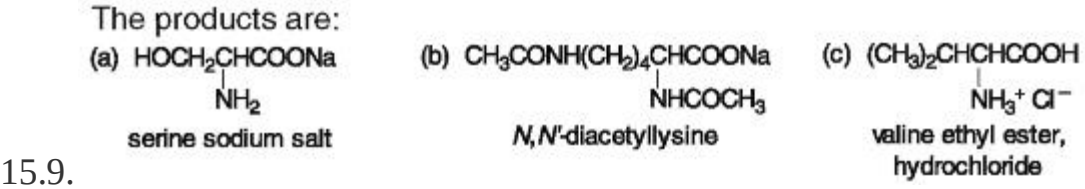
(b) aspartic acid disodium salt;

(c) cysteine *O*- and *S*-disodium salt;



15.8. The products are:



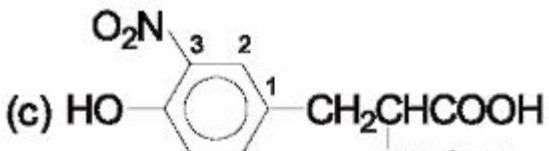


15.9.

15.10. The products are:

(a) glycine sodium salt;

(b) lysine methyl ester, dihydrochloride;



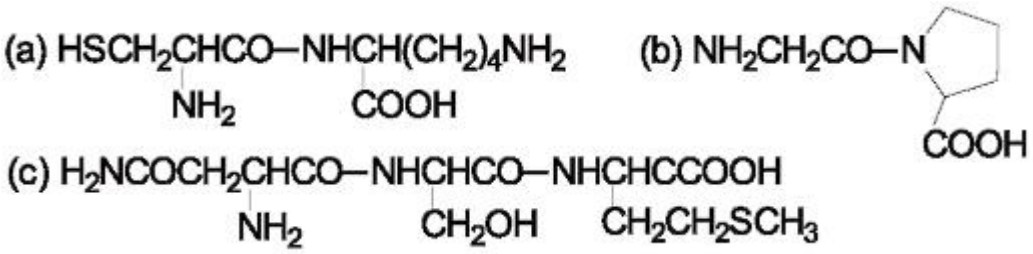
3-nitrotyrosine hydrochloride

N-(hydroxymethyl)valine

15.11. The compound

(a) can be produced from serine and the compound

(b) from glutamic acid.



15.12.

15.13. The products are (in the form of salts with hydrochloric acid):

(a) phenylalanine and valine;

(b) glutamic acid and phenylalanine;

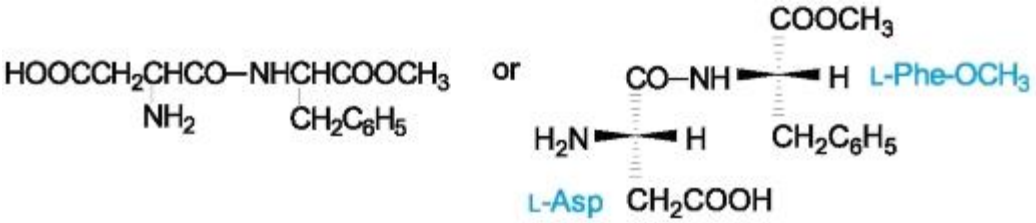
(c) glycine, leucine, and glutamic acid *(not*glutamine).

15.14. The products are (in the form of monoor disodium salts):

(a) glutamic acid and tyrosine;

(b) phenylalanine and isoleucine;

(c) histidine, methionine, and aspartic acid *(not*asparagine).



15.15.

To Chapter 16

16.1. All, except for (b). For the reasonable answer draw the orbital pictures and consider criteria of aromaticity.

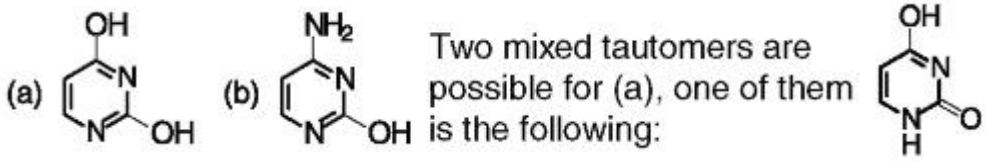
16.2. Consider electronic configuration of the two nitrogen atoms. The unshared electron pair on the N-1 atom is delocalized and is a part of the aromatic sixπ-electron system. Whereas the N-2 atom is the pyridine-type nitrogen whose lone pair of electrons can be protonated.

16.3. Only (c) and (e). Some comments:

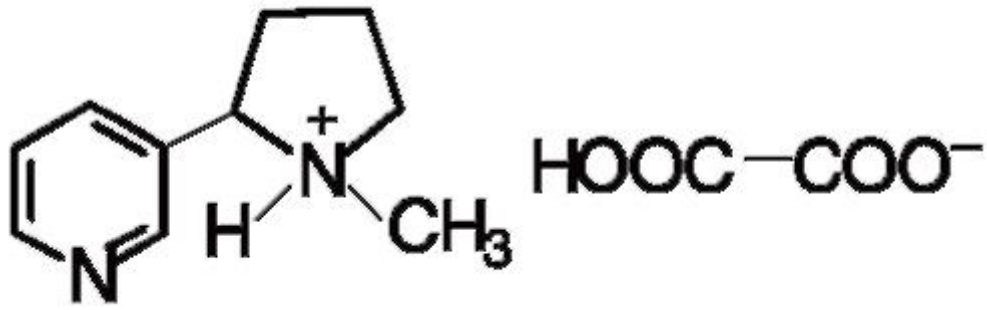
(d) only one hydroxyl group (that is attached to the ring) reacts with alkali;

(e) the structure of the Shiff's base in reaction with amine RNH_2 is the following:

16.4.



16.5. The nitrogen of the pyrrolidine ring is much more basic since pyrrolidine is the saturated tertiary amine. Its nitrogen atom will be protonated in the first place to form a salt:



16.6. See Example 16.1 and answer to Problem 16.1.

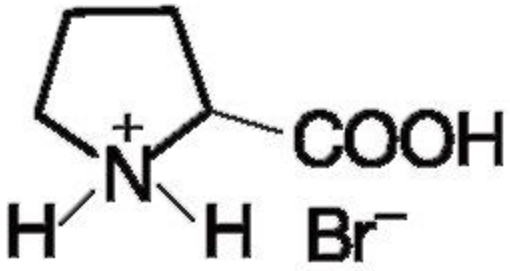
16.7. The compounds (b) and (d) are aromatic. Nitrogen in (a) is sp^3-hybridized (neither pyridinenor pyrrole-type). Nitrogen in (c) is sp^2-hybridized and participates in ρ,π conjugation (it is very similar to pyrrole-type), but the compound is not aromatic. Nitrogen in (d) is pyrrole-type.

16.8. Pyrrole < quinoline < ammonia < piperidine.

Piperidine is a typical secondary amine. Quinoline is approximately as basic as pyridine. For other explanations see the text.

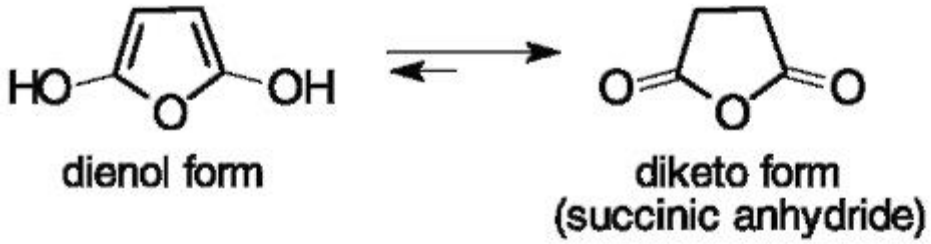
16.9. The difference can be accounted for by hydrogen bonding between the acidic (N-1) and basic (N-3) sites of amphoteric imidazole.

16.10. Only (c) and (d). The structure of a salt is as follows: The C-2 atom is a chiral centre. As regards to the item (e), proline reacts with NaOH but a sodium

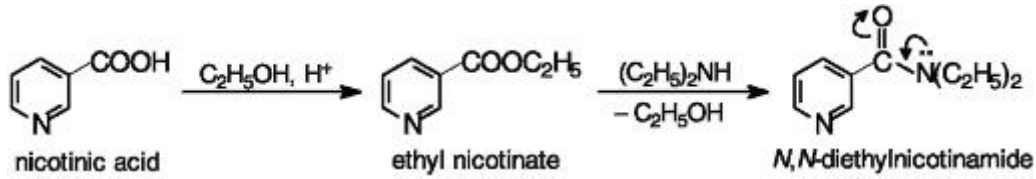


carboxylate salt is formed.

16.11. Keto-enol tautomerism is a reason for transformation:



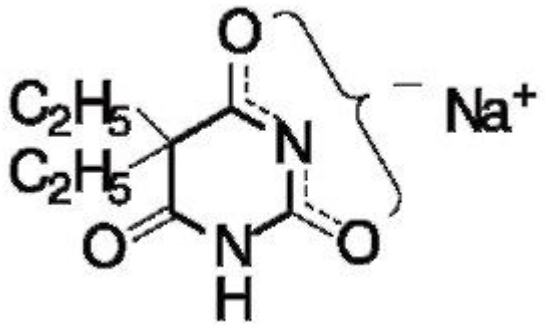
16.12. First use oxidation (as in the text) and then two successive nucleophilic substitution steps, namely, esterification of nicotinic acid followed by aminolysis of the ester:



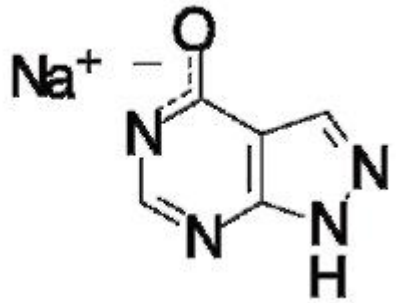
The nitrogen of the amido group is much weaker than the pyridine-type nitrogen due to conjugation of the lone pair of electrons with the C=O double bond.

16.13. Both NH groups represent acidic (not basic) sites. A possible structure for the salt is:

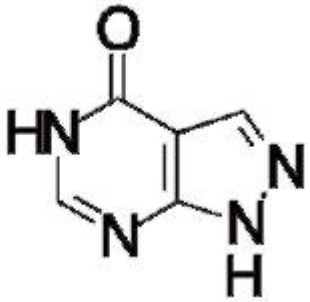
16.14. All, except for (b). The products of transformations are:



(d)



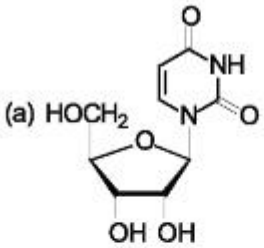
(e)



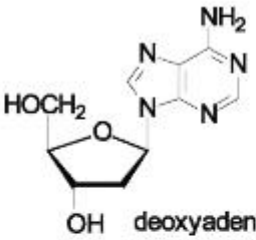
To Chapter 17

17.1.

(a)



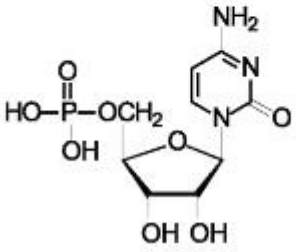
(b)



deoxyadenosine

(c) See Example 17.1 (b)

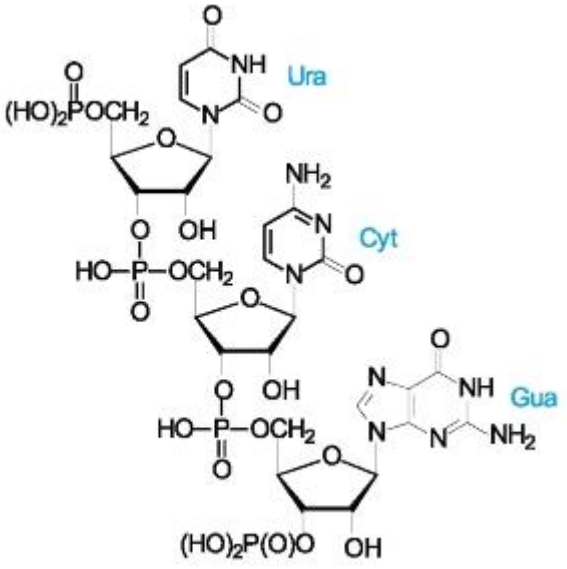
17.2.



cytidine 5'-phosphate

For the remaining formulas see the text.

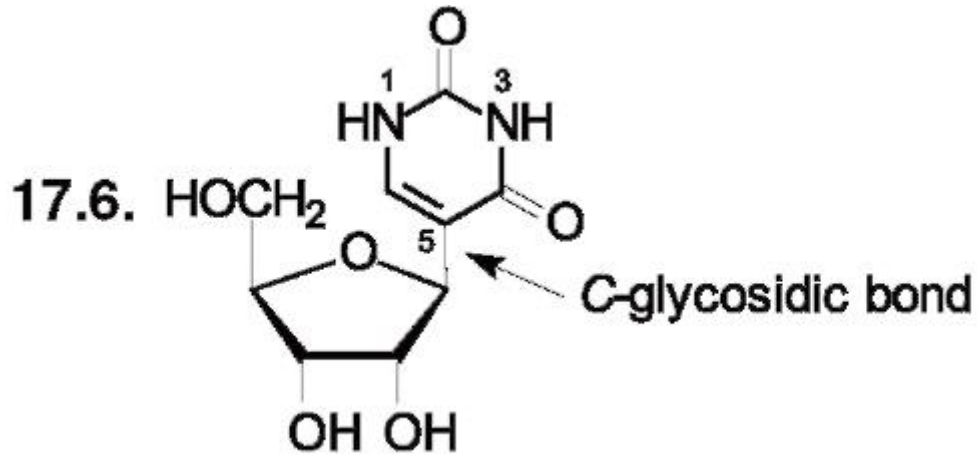
17.3.



17.4. CAGGTAT written from 5' to 3'.

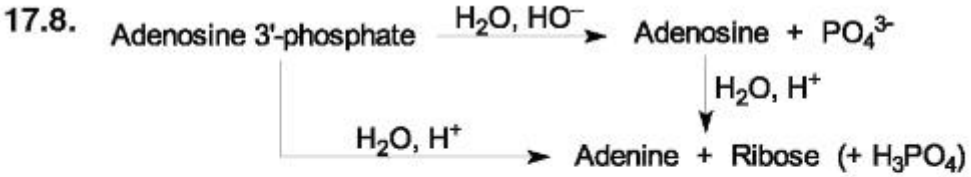
17.5. Draw the structural formula and find (b), (c), and (d).

17.6.



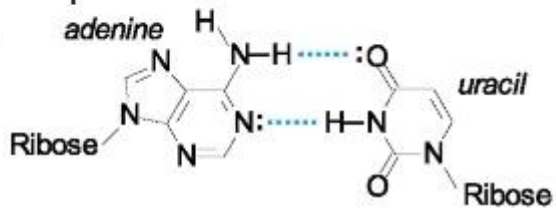
17.7. The products of the acidic hydrolysis are d-ribose and adenine.

17.8.



17.9. The product is uridine.

17.10.



17.11. Use the group $C_6H_5CH_2$- instead of R in the general formula given in Sec. 17.4.1.

17.12. Find the answer in Sec. 7.2.5.

References

Тюкавкина Н.А., Бауков Ю.И., Зурабян С.Э. Биоорганическая химия: учебник. - М.: ГЭОТАР-Медиа, 2012. - 412 с.

Руководство к лабораторным занятиям по биоорганической химии: учеб. пос. / под ред. Н.А. Тюкавкиной. - 4-е изд. - М.: Дрофа, 2008. - 318 с.

Hart H., Hart D.J., Craine L.E. Organic Chemistry (A Short Course). 10th edition. - Boston - NY: Houghton Mifflin Company, 1999. - 573 p.