1. Introduction

1.1. The Field of Organic Chemistry

1.1. The Field of Organic Chemistry
1.2. Origin of Natural Compounds
1.3. Classes of Natural Compounds
1.4. The Covalent Bond
1.5. Bond Dissociation Energy
1.6. Bond Polarization and Electronegativity
1.7. Dipole Moment
1.8. Intermolecular Interactions
1.9. *Van der Waals* Forces
1.10. Mechanisms of Chemical Reactions
1.11. Kinetics of Chemical Reactions
1.12. 'Stable' and Reactive Molecules
1.13. Molecular Shape
1.14. Isomerism
1.15. Functional Groups
1.16. Nomenclature of Organic Compounds

From its origins, organic chemistry has been the chemistry of living matter. As such, chemistry deals with the transformations of substances themselves, in which both the structures and the intrinsic properties of the constitutive molecules are changed.

The etymology of the name *chemistry*, as the science that developed from the medieval *alchemy* (*al* is simply the definite article 'the' in Arabic), is uncertain. Possibly, it derives from the Chinese word *jin*, *chin,* or *kim* for gold, the production of which by conversion of other metals was fervently attempted by all ancient civilizations.

The subject of organic chemistry is the description of the properties and reactions of derivatives of the sixth element of the periodic table: carbon. The fact that carbonic acid and its salts (carbonates) are traditionally assigned to the field of inorganic chemistry, whereas urea (carbonic acid diamide) and pentazole (a compound that does not contain any carbon atoms) are 'organic' compounds just illustrates the difficulties inherent in creating unequivocal definitions for concepts that in everyday usage have become common expressions.

However, what is the reason why the chemistry of one single element of the periodic table has been separated from the chemistry of all others? As a matter of fact, of the 50 million chemical substances registered before September 7, 2009 there are more derivatives of carbon than of all other elements together. The reasons for this great variety are the particular properties of this element, namely:

1) Carbon atoms form covalent bonds with two, three, or four other atoms, so that linear, planar, and three-dimensional structures, respectively, can be obtained when molecules are built up. Many of these structures are *isomeric* with one another (*cf. Sect. 1.14*), so a great variety of molecules can be obtained from a small number of atoms.
2) Carbon atoms form covalent bonds with one another, thus building up chains of any length. The sizes of molecules consisting of atoms of other elements (*e.g.*, boron, sulfur, or silicon)

that likewise form covalent bonds with themselves are considerably smaller.

3) Carbon atoms form strong, non-hydrolyzable (i.e., not cleaved by water) covalent bonds not only with hydrogen atoms but also with the so-called *heteroatoms* (*i.e.*, atoms of other elements such as nitrogen, oxygen, sulfur, and halogens). No marked preference for particular heteroatoms, like that of silicon for oxygen and fluorine, exists in the case of carbon.

4) Moreover, unlike most of the other elements of the periodic table, carbon atoms are able to form covalent double and triple bonds both with other C-atoms and with other elements such as oxygen, sulfur, and nitrogen. Because many of the resulting molecules can be isomers, the variety of possible structures is again increased.

It is therefore understandable that, out of all the chemical elements, carbon constitutes the basis for living matter. An important branch of organic chemistry, known as natural-products chemistry, deals with carbon compounds that are constituents of the different kinds of cells.

1.2. Origin of Natural Compounds

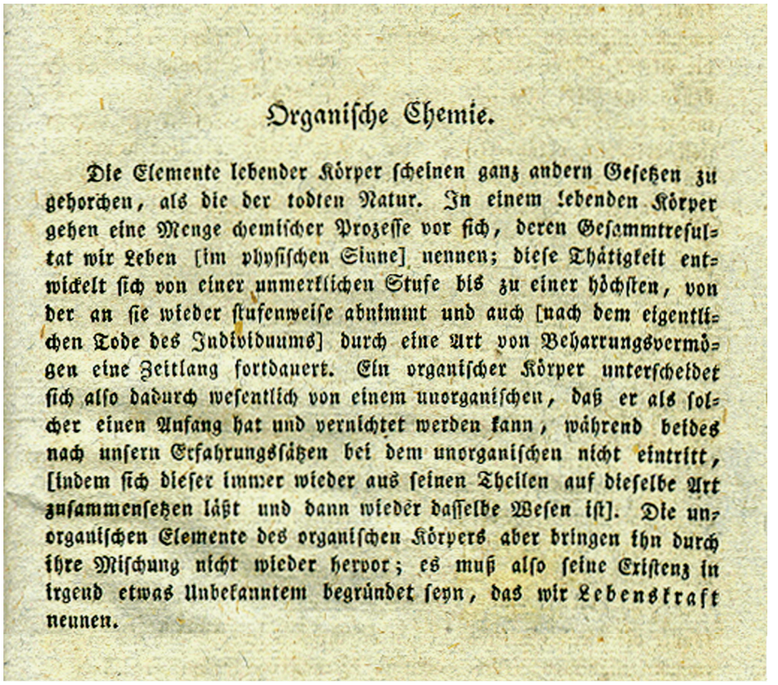
The concept of 'organic chemistry' goes back to *Jöns Jakob von Berzelius*. In his view, the formation of organic substances was the result of the activity of some inherent power that he called *vis vitalis* (*cf. Fig. 1.1*). His student *Friedrich Wöhler* showed in 1828 that urea (a typical 'organic' compound isolated from urine) can be obtained in the laboratory by heating ammonium cyanate (an inorganic substance):

$$(N{\equiv}C{-}O)^- \, NH_4^+ \rightarrow O{=}C(NH_2)_2$$



Although the structure of urea was elucidated only fifty years later (in 1879), by *Rudolf Leuckart* (*cf. Sect. 8.15*), *Wöhler*'s at the time sensational experiment is considered to have laid the foundations for organic chemical synthesis. Nevertheless, the conviction that the

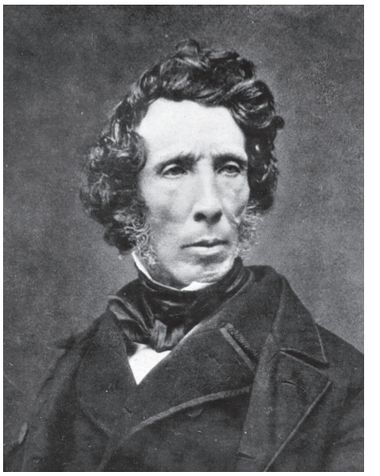
Jöns Jacob Baron Berzelius (1779–1848)
'A famous chemist, an unknown physician' became professor at the Surgical School in Stockholm in 1807, influencing the development of chemistry in Europe over more than five decades. His pioneering achievements include fundamental improvements of laboratory techniques, particularly in organic analysis, the establishment of the later superseded *dualistic* theory, and the use of chemical symbols based on the initial letters of the names of the elements, as well as the discovery of several at that time unknown elements (*e.g.*, silicon, zirconium, tantalum) and organic compounds (lactic acid from meat, starch, pyruvic acid, chlorophyll, *etc.*). (*Edgar Fahs Smith* Collection, with permission of the University of Pennsylvania Libraries, Philadelphia)



Organiſche Chemie.

Die Elemente lebender Körper ſcheinen ganz andern Geſetzen zu gehorchen, als die der todten Natur. In einem lebenden Körper gehen eine Menge chemiſcher Prozeſſe vor ſich, deren Geſammtreſultat wir Leben [im phyſiſchen Sinne] nennen; dieſe Thätigkeit entwickelt ſich von einer unmerklichen Stufe bis zu einer höchſten, von der an ſie wieder ſtufenweiſe abnimmt und auch [nach dem eigentlichen Tode des Individuums] durch eine Art von Beharrungsvermögen eine Zeitlang fortdauert. Ein organiſcher Körper unterſcheidet ſich alſo dadurch weſentlich von einem unorganiſchen, daß er als ſolcher einen Anfang hat und vernichtet werden kann, während beides nach unſern Erfahrungsſätzen bei dem unorganiſchen nicht eintritt, [indem ſich dieſer immer wieder aus ſeinen Theilen auf dieſelbe Art zuſammenſetzen läßt und dann wieder daſſelbe Weſen iſt]. Die unorganiſchen Elemente des organiſchen Körpers aber bringen ihn durch ihre Miſchung nicht wieder hervor; es muß alſo ſeine Exiſtenz in irgend etwas Unbekanntem begründet ſeyn, das wir Lebenskraft nennen.

Fig. 1.1. *Author's photograph of* Berzelius' *'Lehrbuch der Chemie'* ('Textbook of Organic Chemistry'; 2nd Vol., p. 1, by Druck und Verlag der J. B. Metzler'schen Buchhandlung, Stuttgart, 1832)



Friedrich Wöhler (1800 – 1882) took up the study of chemistry in 1823 with *Jöns Jacob Berzelius* in Stockholm. Upon his return to Germany, he began teaching chemistry in the municipal technical school in Berlin (1825 – 1831) and in the technical school in Kassel (1831 – 1836). In 1836, he was appointed professor of chemistry at the University of Göttingen (Germany). About the same time that he synthesized urea, he developed the first industrial method for the preparation of metallic aluminum. He also discovered calcium carbide, as well as several elements such as beryllium, yttrium, and crystalline silicon. *Wöhler* investigated the enzymatic breakdown of amygdalin, isolated cocaine for the first time, and carried out the synthesis of different organic compounds such as oxalic acid, benzoic acid (from benzaldehyde), benzoyl chloride, and benzamide. (*Edgar Fahs Smith* Collection, with permission of the University of Pennsylvania Libraries, Philadelphia)

(bio)synthesis of organic natural products must have been preceded by the appearance of living organisms remained unshakeable until the mid-20th century. It was not until 1953 that *Stanley L. Miller* succeeded in obtaining experimental evidence in support of *Oparin*'s hypothesis that organic compounds might have originated in the secondary earth's atmosphere believed to have existed 5,000 to 4,500 million years ago. In *Miller*'s experiments, mixtures of ammonia, methane, and water vapor, simulating the composition of a reducing early atmosphere (*cf. Sect. 2.16*), were subjected for several days to electrical discharges in a closed system at 80 °C (*cf. Fig. 1.2*), resulting in the observation of the formation of biologically important molecules such as amino acids (*e.g.*, glycine, alanine, serine, aspartic acid, *etc.*), as well as, among other compounds, lactic acid and succinic acid. A crucial role in the 'prebiotic' synthesis of organic molecules (*i.e.*, the synthesis likely to have taken place before the development of living organisms) was presumably played by hydrogen cyanide (HCN), which could have been formed in large quantities in a secondary earth's atmosphere under the influence of heat (volcanic eruptions), ultraviolet light (solar irradiation), or electrical discharges (lightning) through reactions of the type:

$$CH_4 + NH_3 \rightarrow HCN + 3\,H_2$$

Stanley Lloyd Miller (1930 – 2007)
carried out his famous experiments that demonstrated the formation of organic molecules under prebiotic conditions during his doctoral studies at the University of Chicago under Professor *Harold Clayton Urey* (1893 – 1981), the discoverer of the heavy hydrogen isotope known as deuterium. *Miller*'s later work as professor at the Department of Chemistry of the University of California at San Diego mainly concerned the formation of inclusion compounds, known as clathrates.

Alexander Ivanovich Oparin (1894 – 1980)
was professor of biochemistry at the University of Moscow (in 1929) and later (in 1946) director of the biochemical institute of the Russian Academy of Sciences in Moscow. Although best known for his contributions to studies of the origin of life, *Oparin* also devoted considerable efforts to enzymology and to the closely related subject of industrial biochemistry. In 1924, *Oparin* published his postulates on the origin of life that essentially agreed with the concepts independently formulated in 1928 by the British geneticist *John Burdon Sanderson Haldane* (1892 – 1964). The hypothesis that simple inorganic materials might have combined into complex organic molecules were experimentally corroborated by *S. L. Miller* in 1953.

On the one hand, HCN is the precursor of a series of reactive organic substances, such as prop-2-ynenitrile ($HC{\equiv}C{-}C{\equiv}N$), cyanamide ($H_2N{-}C{\equiv}N$), aminomalonitrile ($H_2N{-}CH(C{\equiv}N)_2$), *etc.*, which in the course of *chemical evolution* may have further reacted to yield more complex structures. It is noteworthy that similar organic compounds have been found in interstellar space. The largest organic molecule so far detected by radio telescope in a so-called dark nebula is 2,4,6,8,10-undeca-2,4,6,8,10-pentynenitrile:

$$H{-}C{\equiv}C{-}C{\equiv}C{-}C{\equiv}C{-}C{\equiv}C{-}C{\equiv}C{-}C{\equiv}N$$

On the other hand, even 'prebiotic' syntheses of biologically important molecules such as adenine – a constitutive element of more complex structures that are essential to primary metabolism, including DNA, RNA, ATP, NADH, FAD, *S*-adenosylmethionine, and coenzyme A (*cf. Fig. 10.80*) – can be reproduced in the laboratory.



Fig. 1.2. *Reconstruction of the apparatus used by* S. L. Miller *to simulate the conditions that presumably led to the synthesis of organic molecules in the primeval earth's atmosphere* (Author's photograph).

After *Miller*'s experiments, extensive studies of the prebiotic synthesis of nucleotides and their nonenzymatic polymerization to nucleic acids were carried out in different research groups, particularly that of *Leslie E. Orgel* at the *Salk* Institute for Biological Studies at San Diego, California (*cf. Sect. 10.20*).

Meanwhile, the assumption that the prebiotic earth's atmosphere was strongly reducing, as postulated by *Urey* and *Miller*, has been seriously called into question. The latest geochemical models suggest that the greater proportion of the hydrogen originally present on earth was lost during the planet's cooling phase, with CO_2 rather than CH_4 thus remaining as the main source of carbon. Under these conditions, the yields of amino acids formed by electrical discharges should have been considerably lower. However, it has been known for a long time that hydrogen is released on irradiation of aqueous solutions of Fe^{II} salts with UV light according to the following equation:

$$2\,Fe^{2+} + 7\,H_2O \rightarrow Fe_2O_3 + 4\,H_3O^+ + H_2$$

Leslie Eleazer Orgel (1927 – 2007)
started his career as a theoretical inorganic chemist with contributions to ligand field theory. In 1964, he was appointed Senior Fellow and Research Professor at the *Salk* Institute for Biological Studies, where he directed the Chemical Evolution Laboratory. He was also an adjunct professor in the Department of Chemistry and Biochemistry at the University of California, San Diego.

The thus formed hydrogen might have been the source for the synthesis of large quantities of CH_4, NH_3, HCN, H_2S, and other reactive intermediates that served as substrates for the abiotic synthesis of organic molecules.

An important research field of current interest in organic chemistry deals with the investigation of the interactions between low-molecular weight* molecules that lead to the formation of the so-called 'supramolecules'. Indeed, chemists have lately come to realize that some distinctive features of living matter, such as autonomous morphogenesis, are actually inherent molecular properties that fulfill the conditions for self-organization and autocatalytic replication* (*cf. Sect. 1.11*) as present on the cellular level.

Intuitively, 'life' is associated with motion and/or growth, although this kind of definition would not apply either to flowing water or to crystals in a saturated saline solution. In his essay '*Le Hasard et la nécessité*' ('Chance and Necessity', Paris 1970), *J. Monod* considers autonomous morphogenesis*, reproductive invariance*, and *teleonomy* (the tendency to have a purpose or project) as the essential properties of living matter. The concept of teleonomy (Gr.: *τέλος* (*télos*), end, purpose; and *νόμος* (*nómos*), law) was introduced into natural philosophy in 1958 by the US biologist *Colin Pittendrigh* as an antonym of *Aristotle*'s teleology (Gr.: *λόγος* (*lógos*), reasoning). It, therefore, replaces the teleological question *for what?* by the question *what advantage has a form, as developed in the course of evolution, to the continuance of a species?*, thus implying that the formation process can also be causally understood as a result of mutation and selection. Teleonomy is, therefore, an attribute of evolution (as a result from mutation and selection) rather than a property of living matter, because, at a very early stage of the world's evolution, as long as the primary population pressure was low as a result of an abundant supply of primary molecules in the environment, the appearance of living matter was presumably deprived of any causality. Moreover, because the ability of polynucleotide molecules to reproduce themselves, if necessary

Jacques Lucien Monod (1910 – 1976)
was appointed Professor of the Chemistry of Metabolism at the Sorbonne in 1959. In 1967, he became Professor at the Collège de France, and in 1971 he was appointed Director of the Institut *Pasteur* in Paris. Together with *François Jacob* he made fundamental contributions to the elucidation of how genes regulate the biosynthesis of enzymes. They thus discovered in 1960 that some particular DNA sequences, which they called *operons*, regulate gene expression in bacteria (*cf. Sect. 10.18*), and in 1961 they proposed the existence of *messenger* ribonucleic acid (mRNA), as a transmitter of genetic information from the DNA in the nucleus to the sites of protein biosynthesis in the cytoplasm. In 1965, *Monod*, together with *Jeffries Wyman* and *Jean-Pierre Changeux*, postulated the symmetry model (also known as the concerted model) to explain the kinetic properties of allosteric enzymes (*cf. Sect. 9.32*). *Monod* and *Jacob*, together with *André Lwolff* (1902 – 1994), shared the 1965 *Nobel* Prize for physiology or medicine.

Louis Pasteur (1822 – 1895)
His early investigations into the morphology of the tartaric acid crystals that separate in barrels containing fermenting wine led to the first resolution of enantiomers (*cf. Sect. 2.7*). In 1849, *Pasteur* became chairman of the chemistry department at the University of Strasbourg (France). Here, prompted by the problems of the French beverage industry at that time, he began studying alcoholic fermentation, discovering in the process that external microorganisms were involved and can be destroyed by careful heating, a procedure known today as *pasteurization*. In 1854, he was appointed Professor of Chemistry and Dean of the new science faculty at the University of Lille (France), and three years later (in 1857) he was named Director of Scientific Studies at the École Normale Supérieure in Paris. On behalf of the French government, *Pasteur* investigated the cause of pébrine, an epidemic disease of silkworms caused by protozoans of the species *Nosema bombycis* that in 1865 threatened to ruin the silk industry in southern France. In the course of these investigations, *Pasteur* established that fermentation, as well as contagious diseases and purulent infections such as sepsis, originate from microorganisms, thus establishing the fundamentals of microbiology and, particularly, of the special field of bacteriology. Beginning in 1881, *Pasteur* fell back on the work of the English surgeon *Edward Jenner* (1749 – 1823), who discovered in 1796 that inoculation for smallpox with matter from a patient affected with cowpox, a relatively harmless disease that could be contracted from cattle, can protect against any subsequent attack of smallpox. On the basis of this observation, Pasteur succeeded in the isolation of a vaccine (Sanskrit: *vaśa*, meaning cow) for rabies that he successfully tested in 1885. *Pasteur*'s pioneering work was honored with the foundation of the *Institut Pasteur* in Paris that he headed until his death.

with alteration of their base sequences (*cf. Sect. 10.18*), is an inherent property of these molecules, their teleonomy consists of their own reproduction. From the biological standpoint, therefore, the much-sought-after 'meaning of life' is found in reproduction in itself as the unavoidable necessity for the continuance of the species. On the other hand, autonomous morphogenesis and reproductive invariance are also characteristic features in the formation of crystals through aggregation of ions or molecules. However, in a living system, reproductive invariance cannot be absolute, because such a property would render the system itself incapable of evolving. Actually, in contrast with sporadic crystalline lattice defects, errors that necessarily occur during reproduction of living matter are transmitted to the descendants: that is, living matter displays a *hereditary variability* (*cf. Sect. 10.18*) that enables evolution to take place at all. Eventually, the evolution process decides whether or not the arisen 'error' is advantageous for survival. A further distinguishing property of living matter is its capability to transform exogenous substances into its own constituents (*i.e.*, to carry out *autogenous metabolism*). In point of fact, this latter property, together with *hereditary variability* must be regarded as the most characteristic attributes of living matter.

The characteristic properties of **living matter** are:

- *autonomous morphogenesis*
- imperfect *reproductive invariance* that allows
- *hereditary variability*, and
- ability to transform environmental substances into its own constituents, *i.e.*, to carry out *autogenous metabolism*.

On the basis of *Pasteur*'s experimental evidence, the German physician and founder of cellular pathology *Rudolf Virchow* (1821 – 1902) greatly contributed to the spread of the aphorism '*omnia cellula e cellula*', the axiom that every cell originates from a preexisting cell rather

than from amorphous material, thus overturning the belief in spontaneous generation, as formulated without any experimental evidence by *Aristotle* (384–322 BCE), that had prevailed until the end of the 19th century. Indeed, the question of the origin of the first living cell remains unanswered. From our current standpoint, therefore, chemical evolution developing from the origin of the earth (4,600 million years ago) until the beginning of biological evolution, believed to have been 3,500 million years ago, cannot be ruled out *a priori*.

1.3. Classes of Natural Compounds

A definition of *'nature'* is *per se* problematic. Because mankind itself is an integral part of nature, all substances that arise from metabolism associated with human somatic and intellectual activities are natural products. Of compounds of human origin, therefore, only those produced, either intentionally or unintentionally, through the *conscious* creative activity of the human intellect are to be regarded as artificial compounds.

Natural compounds are all chemical substances with structures not originally conceived by mankind. However, transformation products of artificial substances may be also natural compounds if they already pre-exist as such.

Natural compounds can be classified either according to their physiological *functions* or according to their *structures*. When the function of a particular natural compound is not specifically defined, there is understandably no relationship with its chemical structure, and so classification based on such a function is inadequate for systematic classification. This is the case for natural compounds categorized as pigments (*cf. Sect. 3.21*), vitamins (*cf. Sect. 3.22*), hormones (*cf. Sect. 7.4*), *etc.*

An appropriate classification of the majority of natural compounds is shown in *Fig. 1.4*. It is based on the *biogenesis* of the most important classes of natural compounds, which in most cases does correlate with their structural types. Accordingly, natural compounds are divided into *primary* and *secondary metabolites*. Primary metabolites are compounds that are found in all living organisms. They provide the fundamental elements of the phenomenon of 'life'. They include: the nucleic acids, as self-replicating carriers of genetic information, the proteins, as catalysts (enzymes) of all biochemical reactions, as well as the operational and structure-building elements in all organisms, the lipids, as constituents of compartments within the cell that enable the maintenance of metabolism, and the polysaccharides (glycans) that are the 'fuels' providing the cell with the necessary energy to carry out all endergonic biochemical reactions; several polysaccharides, like some proteins, are structure-building elements in organisms, particularly in bacteria and plants.

A plausible, though not established, pathway of the synthesis of a natural compound by a living organism is referred to as its **biogenesis**. If all intermediates have been identified experimentally, the process is called **biosynthesis**. This differentiation, however, is not always respected.

Most secondary metabolites are also indispensable to the maintenance of a metabolism. They are, however, not ubiquitous. A given secondary metabolite may be essential to some organisms but not to others. Frequently, they serve in the defense of an organism against

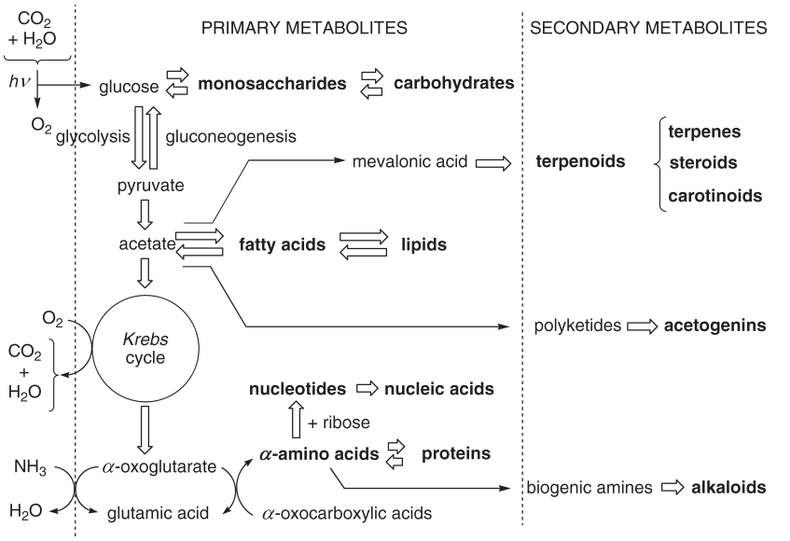


Fig. 1.3. *Main classes of natural compounds and their relation to primary cell metabolism*

its predators. Some examples are to be found in numerous poisons (toxins), hormones, pheromones (attractants), and scents (*cf. Sect. 3.20*).

1.4. The Covalent Bond

A **molecule** is a group of a limited number of atoms that are held together by covalent bonds.

Unlike metals and crystalline salts, which consist of three-dimensional arrangements either of atoms or of ions, the numbers of which are undefined, organic compounds are made up of *molecules*, the structures of which (*i.e.*, both their constitutions and their geometries) are determined by covalent bonds between the constitutive atoms. Covalent bonds are characterized both by their *lengths* and by their energies of *dissociation*, both of which are experimentally determinable. The distances between the nuclei of the atoms that make up a molecule can thus be determined with high accuracy by

The **forces** that hold together the atoms in a molecule are named **covalent bonds**.

The distance between the nuclei of two atoms linked together by a covalent bond is called the **bond length**.

Table. 1.1. *Nonpolar Covalent Atomic Radii of Common Elements.*

Atom	*r* [pm]	Atom	*r* [pm]
H	37	Cl	99
F	72	S	102
O	73	P	106
N	75	Br	114
C	77	I	133

X-ray diffraction in molecular crystals. From these values, the covalent radii of the individual atoms can easily be obtained (*cf. Table 1.1*). Atomic covalent radii are used both to estimate molecular bond lengths when these are not accessible experimentally and to construct molecular models (*cf. Sect. 1.13*). In both cases, however, it is assumed that atomic covalent radii are *additive*, independently of the structure of the molecule as a whole, a prerequisite that is only met within a certain tolerance, as a result of the polarization of covalent bonds connecting non-identical atoms (*cf. Sect. 1.6*).

The **covalent radius** is defined as half of the bond length of a covalent bond joining together two atoms of the same element in a molecule.

1.5. Bond Dissociation Energy

As mentioned above, any covalent bond has a characteristic energy of dissociation (also referred to as *bond strength*). The reciprocal relationship between the bond length (l) and the bond dissociation energy (D) may be represented as a potential-energy curve such as that shown in *Fig. 1.4*. This shows the dependence, in general, of the potential energy of any molecular system made up of two atoms; one of these (*e.g.*, a carbon atom) is considered to be located at the origin of a system of coordinates and the other (*e.g.*, a H-atom) is considered to move along the abscissa. If the distance between the two atoms is long enough (*i.e.*, at 'infinite' separation) there is no interaction between them at all, and the corresponding potential energy of the system is defined, by convention, as zero. On the other hand, the lowest limit to the internuclear distance is set by the electrostatic repulsion between the nuclei of the covalently bound atoms, as well as that between those of their core electrons that are not involved in the formation of the covalent bond, thus resulting in an equilibrium distance (l_0) characteristic of each covalent bond, which corresponds to the lowest point of the potential-energy curve. As a consequence, the cleavage of a bond can only take place with consumption of energy.

The energy required to cleave a covalent bond between two atoms homolytically by separating them from the bond distance to infinity is referred to as the **bond dissociation energy**.

The concept of the 'high-energy bond' in adenosine triphosphate (ATP), as suggested by *F. Lipmann* (*cf. Sect. 10.16*), might, therefore, serve as an illustration of which particular bond in a molecule releases more free energy* on hydrolysis*, but it is conceptually false, because the energy 'stored' in adenosine triphosphate is nothing more than the heat of reaction released when the following chemical reaction reaches its thermodynamic equilibrium, which under physiological conditions is displaced to the right (*cf. Sect. 10.16*):

$$ATP + H_2O \rightleftharpoons ADP + H_3PO_4 \quad (ADP = \text{adenosine diphosphate})$$

The above potential-energy curve can be approached through the equation proposed by *Philip McCord Morse* (1903–1985): $E(l) = D[1 - e^{-a(l-l_0)}]^2$, where a is a constant for the given molecule and $E(l)$ is the potential energy for an internuclear distance l, reckoned from the lowest point of the curve (*i.e.*, at $l = l_0$), rather than from

$E(l) = D\,(1 - e^{-a\,(l - l_0)})^2$
(a: a bond constant)

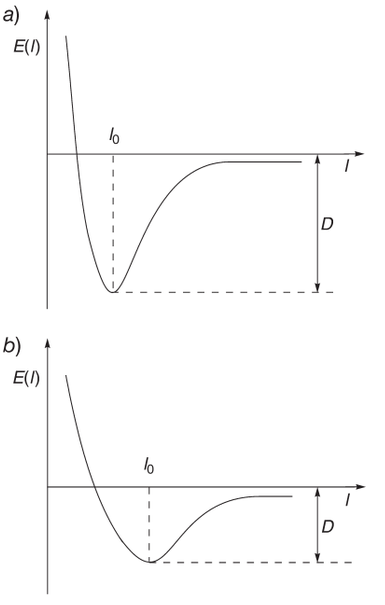


Fig. 1.4. *Potential-energy curve for a system consisting of two atoms that are linked together by* a) *a strong or* b) *a weak covalent bond*

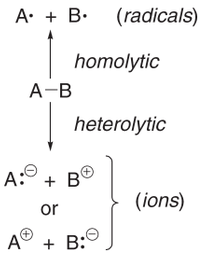


Fig. 1.5. Homolytic and heterolytic cleavage of a covalent bond

the abscissa (see above) as zero. Thus, as l approaches infinity, $E(l)$ becomes equal to D, the bond dissociation energy, as required by *Fig. 1.4*. However, as l approaches zero, the value or $E(l)$ remains finite, although it should increase to infinity. This weakness of the *Morse* function is not serious, because the conditions under which the internuclear distance approaches zero are of no practical significance. In fact, however, atoms that are linked together by covalent bonds oscillate about the equilibrium position defined by l_0. Therefore, as a mechanical model for a covalent bond, the reality is more like a coiled spring joining two balls together than a rod. Moreover, because the amplitude of this oscillation increases with rising temperature, the bond dissociation energy is temperature-dependent. Any covalent bond between two atoms can be represented by its own *Morse* curve. With a few exceptions, the bond dissociation energy increases when the bond length decreases; that is, the minimum of the curve flattens when l_0 increases (*cf. Fig. 1.4, b*).

Bond dissociation energies are given either in kilojoules ($1\ \text{kJ} = 10^3$ Joule) or in kilocalories ($1\ \text{kcal} = 10^3$ calories). *Table 1.2* summarizes the bond dissociation energies of covalent bonds between commonly occurring atoms in organic molecules. The bond dissociation energies defined in this way, though, are those required to cleave covalent bonds *homolytically*. In fact, however, a covalent bond may be cleaved either *homolytically* or *heterolytically* (*cf. Fig. 1.5*).

Free radicals are electrically uncharged molecules that contain odd numbers of electrons.

In the absence of solvation effects (*cf. Sect. 1.9*), homolytic bond cleavage, which leads to (uncharged) free radicals, requires less energy than its heterolytic counterpart, because the work required to overcome the *Coulomb* force of attraction between the ions resulting from the latter has to be supplied in addition to the bond dissocia-

Table. 1.2. *Bond Dissociation Energies* (in kilojoules per mole, *i.e.*, kJ mol^{-1}) *of Covalent Bonds between Atoms of the Same Element in Diatomic Molecules, As Well As between Tetragonal C-Atoms and Atoms of Other Common Elements.* Fourth column: lengths of the latter class of bonds.

X	D [kJ mol^{-1}]		l_{C-X} [pm]
	X–X	C–X	
H	434	415	111
D	439	447.6	110.6
F	158	488	135
O	498	360	143
N	945	308[a])	147
C	–	344	154
Cl	243	330	177
Br	193	288	194
I	151	216	214

[a]) Strongly dependent on the constitution of the given molecule.

tion energy. Heterolytic bond cleavage is, therefore, preferred in polar solvents because the solvation energy overcompensates the electrostatic interactions between the ions. In the gas phase, however, homolytic cleavage predominates. Probably, it also plays a more important role than assumed up to now in enzymatic reactions, because many of these reactions take place in hydrophobic 'pockets' within the enzyme.

The *energy content* of a molecule is always lower than that of the separated atoms altogether. If this were not the case, the molecule would spontaneously decompose into its constitutive elements. It results from this that the formation of a bond is always an *exothermic* process.

Chemical reactions that take place with evolution of energy (heat) are named *exothermic* or *exergonic*, depending on whether the released heat of reaction is the result of a decrease in the *enthalpy* (ΔH) or the *free energy* (ΔG), respectively, of the system involved. According to the thermodynamic definitions of ΔH ($= H_{\text{products}} - H_{\text{substrates}}$) and ΔG ($= G_{\text{products}} - G_{\text{substrates}}$), both quantities are < 0. Reactions for which $\Delta H > 0$ or $\Delta G > 0$ are called *endothermic* or *endergonic*, respectively.

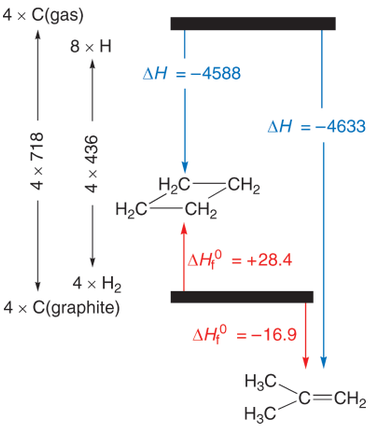


Fig. 1.6. *Formation of isobutene and cyclobutane from the constituent atoms in the gas phase, as well as from the corresponding elements in their standard states* (schematic; values in kJ mol^{-1})

The energy content of a molecule is the sum of the dissociation energies of the covalent bonds present in the molecule. It can be expressed in terms either of the molecule's *enthalpy of formation* (ΔH_f^0) or of its *atomization energy*. Obviously, in the case of a biatomic molecule, the atomization energy is identical with the bond dissociation energy. The difference between enthalpy of formation and atomization energy is illustrated in *Fig. 1.6* with the example of two compounds (isobutene and cyclobutane) the molecules of which contain the same number of identical atoms (*i.e.*, they are isomers; *cf. Sect. 1.14*). Most bond dissociation energies in organic molecules are indirectly determined from the experimentally easily available heats of combustion (ΔH_c^0) of the corresponding substances (*cf. Sect. 2.12*).

The **standard heat of formation** (ΔH_f^0) of a compound is the change in the heat content (ΔH) of the reacting system when 1 mole* of the substance is formed from its elements in their so-called standard states, *i.e.*, the state of aggregation (solid, liquid, or gas) in which they occur at 298.15 K and atmospheric pressure.

The **atomization energy** is the energy required to decompose a molecule into its constitutive atoms in the gas phase. The same quantity (with opposite sign) is the energy that is released when the same molecule is formed from the constitutive atoms in the gas phase.

Whereas the formation of a molecule from its constitutive atoms in the gas phase is always an exothermic process (*cf.* blue arrows in *Fig. 1.6*), the formation of an organic compound from the corresponding elements (*cf.* red arrows in *Fig. 1.6*) may be either exothermal (*i.e.*, $\Delta H_f^0 < 0$) or endothermal (*i.e.*, $\Delta H_f^0 > 0$), depending on whether the atomization energy of the molecule is higher or lower than the sum of the heat of sublimation of graphite (720 kJ gram-atom^{-1}) and the atomization energies of all other elements present in the molecule (*cf. Fig. 1.7*, left-hand side), their enthalpies of formation in the so-called standard state being zero, by definition, at all temperatures.

Following a suggestion by the Scottish chemist *Archibald Scott Couper* (*cf. Sect. 4.1*), covalent bonds in organic molecules are today represented by continuous or dotted lines between the symbols of the atoms that are linked together. As later postulated (1916) by *Gilbert Newton Lewis*, such a line symbolizes an electron pair possessing antiparallel spins (*i.e.*, with opposite angular momentum with reference to the sense of their translation in the space) and so is indifferent

Gilbert Newton Lewis (1875 – 1946) received his Ph.D. from Harvard University (Massachusetts) in 1899. After a short time (1905 – 1912) as professor of chemistry at Massachusetts Institute of Technology (MIT), he became professor of physical chemistry and dean of the College of Chemistry at the University of California, Berkeley in 1912. His main research field was chemical thermodynamics. In addition to his famous work on structural theory (1916) and his theory of acids* and bases* (1923), *Lewis* was the first to prepare deuterium oxide ('heavy water') in 1933 (*cf. Sect. 9.5*). In 1926, he coined the term 'photon' for the smallest unit of radiant energy.

as to whether each atom contributes to the formation of the covalent bond with one electron, or whether both electrons belong 'originally' to the same (donor) atom. In the latter case, the covalent bond is named a *dative* or *coordinative* bond; this is symbolized by an arrow pointing towards the acceptor atom (*cf. Fig. 6.1*).

In modern physics, the attractive force between atoms that are linked together is explained in terms of an interchange of bonding electrons between the atoms. This may be visualized in the form of two table tennis players who keep a certain distance from each other as the ball is hit to and fro during the play. The presence of a covalent bond consisting of two electrons (e_1^- and e_2^-) can thus be explained in terms of the simultaneous existence of two so-called resonance structures [Ae_1^- Be_2^- ⇌ Ae_2^- Be_1^-] of one and the same molecule (A – B; *cf. Sect. 3.13*). Other stabilizing forces between material particles are also understood as 'interchange forces' (*cf. Table 1.3*). The permanent interchange of (virtual) photons between two particles bearing opposite electric charges, for instance, generates the electromagnetic force. An atom can, therefore, be imagined as a material particle, the constituent parts of which (*i.e.*, the nucleus and electrons) interact through the emission and absorption of virtual photons. In a manner similar to that in which photons keep the electrons on their orbit around the atomic nucleus, other particles, so-called *gluons*, are responsible for the strong force that associate *quarks* with one another to make up nucleons (*i.e.*, protons* and neutrons).

Table. 1.3. *Fundamental Interchange Forces*

Interaction between ...	Interchanged particle	Spin
Electrical charges	Photon	1
Atoms	Electron	1/2
Quarks	Gluon	1

1.6. Bond Polarization and Electronegativity

A link between two atoms resulting from the permanent interchange of electrons between them both is referred to as a **covalent bond**.

Unlike bond lengths, bond dissociation energies cannot be calculated by summation of contributions from the individual atoms that are bound together. The reason for this is that covalent bonds linking atoms of different elements are *polarized*. Actually, **the more strongly polarized a covalent bond is, the stronger is it**, because in order to cleave such a bond not only the bond dissociation energy but also the *Coulomb* force of attraction between the electrical charges of the existing dipole must be overcome. Since the energy content of a molecule as a whole is the sum of the dissociation energies of the covalent bonds present in the molecule (*cf. Sect. 1.5*), the **ground state enthalpy of the molecule decreases as a result of bond polarization**.

If bond dissociation energies were additive, the bond dissociation energy of a bond linking two atoms of different elements would be approximately equal to the arithmetic mean value of the bond dissociation energies of the covalent bonds linking two atoms of the same elements. On this assumption, *Linus Pauling* realized that the experi-

Linus Carl Pauling (1901 – 1994)
was appointed full Professor at the California Institute of Technology in Pasadena in 1931. From 1963 to 1967, he was attached to the Center for the Study of Democratic Institutions at Santa Barbara, California, as a research professor; from 1967 to 1969 he was a professor of chemistry at the University of California at San Diego, and later on at Stanford University in California. In 1954, he was awarded the *Nobel* Prize for chemistry for his research into the nature of the chemical bond and its application to the elucidation of the structures of complex substances. Again in 1962, he received the *Nobel* Prize for peace in recognition of his efforts to ban the testing of nuclear weapons.

mentally determined bond dissociation energies are always higher than this mean value; *i.e.*,

$$D_{A-B} > 1/2\ (D_{A-A} + D_{B-B}),\ \text{or}\quad D_{A-B} = 1/2\ (D_{A-A} + D_{B-B}) + \Delta.$$

Because $\Delta > 0$, *Pauling* assumed that the difference between the experimentally determined and the calculated quantities is proportional to the square of the difference between the *electronegativities* of the atoms that are linked together, *i.e.*,

$$\Delta = 23\ (EN_A - EN_B)^2.$$

the proportionality constant thus being approximately equal to the conversion factor from electron volts to kilocalories, as energy units (*i.e.*, $1\ \text{eV} = 23.063\ \text{kcal}\,\text{mol}^{-1}$).

From the experimentally determined dissociation energies of covalent bonds between atoms of different elements, it is possible to draw up an electronegativity scale (*cf. Table 1.4*). Its significance is qualitative rather than quantitative, however.

As the 'size' of an atom depends on the balance between its (positive) nuclear charge and the repelling forces exerted mutually by the electrons, elements whose atoms are most compact are those with greatest tendency to attract electrons, *i.e.*, they are, in general, more electronegative.

Despite its name, atomic electronegativity, as defined above, is a thermodynamic rather than an electrical quantity that points to the relative attraction of an atom for the valence electrons *in a covalent bond*. An alternative, theoretically based definition of electronegativity was proposed by *Mulliken*. Accordingly, the electronegativity values (x) are equal to the arithmetic means of the ionization energies (I) and the electron affinities (EA) of the atoms that are linked together by a covalent bond, in their respective valence states, *i.e.*, $x = 1/2\ (I + EA)$. The electronegativity values calculated by this procedure are approximately propor-

Robert Sanderson Mulliken (1896 – 1986)
was professor of chemical physics at the University of Chicago (1931) and Florida State University in Tallahassee (1964). In 1966, his fundamental work on the development of quantum-mechanical methods to elucidate the nature of the chemical bond and the electronic structure of molecules was honored with the *Nobel* Prize for chemistry.

Table. 1.4. *Electronegativity Values for Some Elements, after* L. Pauling. The values quoted are based on the deviation from the geometric mean of the experimentally determined and calculated bond dissociation energies rather than on their arithmetic mean. Conceptually, however, the difference between the two calculation procedures is irrelevant.

Atom	Electronegativity (*EN*)	Atom	Electronegativity (*EN*)
H	2.1	Br	2.8
P	2.1	N	3.0
C	2.5	Cl	3.0
S	2.5	O	3.5
I	2.5	F	4.0

tional to the values given in *Table 1.4*, so x [*Mulliken*]/x [*Pauling*] ≈ 3.15. However, an advantage of *Pauling*'s scale is that it is based on experimentally easily accessible quantities, whereas both the ionization potentials and the electron affinities of atoms in the valence state, in which they occur in a molecule (*i.e.*, not as free atoms), are only accessible by theoretical methods.

The **oxidation number** of an atom in a molecule is determined by the difference between the number of valence electrons of the atom in question and the number of electrons that are assigned to it, as follows: *1*) the electron pair of each covalent bond between two atoms of different elements is assigned to the atom of higher electronegativity, and *2*) if the covalently bound atoms belong to the same element, one bonding electron is assigned to each of them.

The concept of electronegativity proved to be a heuristic principle in chemistry, because different molecular properties discussed in the following sections depend on it (*cf. Table 1.5*). Of far-reaching significance, although not considered in *Table 1.4*, is the fact that the electronegativity of an atom in a molecule depends on its oxidation number, *i.e.*, on the number of bonding electrons that may be assigned to it (*cf. Sect. 5.10*). Obviously, indeed, the **electronegativity of an atom increases when its oxidation number increases**. For this reason, the electronegativity of the carbon atom increases in the order:

$$C_{tetragonal} < C_{trigonal} < C_{digonal}$$

as theoretically justified in 1950 by *William E. Moffitt* (1925–1958).

As a consequence of the higher electronegativity of trigonal carbon atoms relative to tetragonal carbon atoms (*cf. Sect. 1.12*), a covalent bond linking a C=C group with an atom of electronegativity ≤ 2.5 is more strongly polarized and, therefore, stronger than a corresponding bond in an alkane (*cf. Sect. 3.5*).

The higher the difference in the electronegativities of two atoms linked together by a covalent bond, the higher is the bond's *ionic character* (*i.e.*, the *heterolytic* cleavage of such a bond is favored, provided that the medium is able to stabilize the resulting ions by solvation* (*cf. Sect. 1.9*).

This explains the higher acidities of hydrogen atoms that are bound to heteroatoms, relative to hydrogen atoms bound to C atoms, with the solvation of the H_3O^+ ion playing an important role.

Table. 1.5. *Molecular Properties Related to the Concept of Electronegativity*

- Oxidation number of the C-atom (*cf. Sect. 5.10*)
- Dipole moment (*cf. Sect. 1.7*)
- Hydrogen bonding (*cf. Sect. 5.5*)
- Acidity of covalently bound H-atoms (*cf. Sect. 5.6*)
- Nucleophilic substitution (*cf. Sect. 5.13*)

1.7. Dipole Moment

As a further consequence of the polarization of covalent bonds, the electrical center of the bonding electrons do not coincide with that of the nuclei, the latter being virtually at the mass center of the molecule. In other words, a covalent bond between two different atoms will be associated with a *dipole moment*; its value is given by the product of the electronic charge and the relative displacement (d) of the positive and negative electrical centers. The dipole moment is a vector; its direction coincides with the bond axis. In accordance with the suggestion of *Nevil Vincent Sidgwick* (1873–1952), the negative end, as indicated by the arrowhead, points to the atom of higher electronegativity (in physics, however, the sense of the dipole moment is defined inversely, from the negative to the positive charge).

The numerical magnitude of the dipole moment equals the product of the unit charge of electricity (*i.e.*, $q = 4.8029 \times 10^{-10}$ esu) and the distance (d) between the charge centers in Angstroms (Å):

$$|\mu| = q \times d$$

The unit defined in this way is called a *Debye* and is given the symbol D ($1\,\mathrm{D} = 3.33 \cdot 10^{-30}$ coulomb · m). However, because d is not known, the magnitude of the dipole moment is usually expressed as the product of the bond length (l) and a *partial* electric charge (δq), so that:

$$|\mu| = \delta q \times l,$$

and the corresponding bond polarization is represented as a vector between both partial charges, *i.e.*,



The dipole moments of most organic compounds lie in the range between 0 and 4 D (exceptionally up to 7 D). Dipole moments are not directly determinable by experimental methods. Usually, they can be obtained from measurements of the molar polarization (P), which is correlated with the dielectric constant (ε), the density (ρ), and the molecular weight (M) of the substance by the *Debye* equation:

$$P \equiv \frac{\varepsilon - 1}{\varepsilon + 1}\frac{M}{\rho} = \frac{4\pi N}{3}\left(\alpha + \frac{\vec{\mu}^2}{3k_\mathrm{b}T}\right)$$

where N being *Avogadro*'s number, k_B the *Boltzmann* constant, and T the absolute temperature given in kelvin.

The *Debye* equation shows two important aspects of shifting electrons within a covalent bond: the bond *polarization*, as manifested by

Peter Joseph Wilhelm Debye
(1884–1966)
After receiving his Ph.D. in physics from the University of Munich (1910), *Debye* taught physics at the universities of Utrecht (1912), Göttingen (1914), Zürich (1920), and Leipzig (1927) before becoming director of the Kaiser Wilhelm Institute for Theoretical Physics in Berlin (1935). In 1940, he was appointed professor at the University of Ithaca, N.Y., where he remained until he retired as chairman of the chemistry department in 1950. He was awarded the 1936 *Nobel* Prize for chemistry for his contributions to the elucidation of the molecular structure by means of dipole moment studies and X-ray diffraction analysis.

The lack of coincidence of the (positive) electrical center of two atoms that are linked together through a covalent bond with the (negative) electrical center of the bonding electrons is called **bond polarization**.

The ability of bonding electrons to be displaced along the bond axis in the presence of an external electrical field is called **bond polarizability**.

the dipole moment (μ), and the *polarizability* (α) of the molecule, as a temperature-independent parameter that is a measure of the ease with which the bond can be polarized, *i.e.*, how easily the bonding electrons can move along the bond axis, in the presence of an external electrical field. Actually, because of the bond polarizability, the dielectric constant of a substance that does not possess any permanent dipole moment is not zero (*cf. Table 5.2*).

Moreover, because an external electric field may be generated by ions in the medium, bond polarizability plays an important role in chemical reactions in the condensed phase. In general, the polarizability of a covalent bond increases when the distance between the bonding electrons and the nuclei (*i.e.*, the atomic radii) of the atoms that are linked together by the covalent bond increases.

For this reason, *C–H* bonds are much less polarizable than *C–C* bonds (*cf. Sect. 3.5*).

On the other hand, some apolar covalent bonds between atoms of equal electronegativity (*e.g.*, C–S or C–I) are strongly polarizable, thus being decisive for the reactivity of the whole molecule (*cf. Sect. 7.8*).

The dipole moment, as determined through the *Debye* equation, is the dipole moment of the molecule as a whole. Only in the case of a diatomic molecule is the dipole moment of the covalent bond between the two atoms identical with that of the molecule. The resultant dipole moment of a polyatomic molecule, however, is the *vector* sum of the individual dipole moments of the linkages, because each acts along the line connecting the centers of the atoms linked together by the associated covalent bond. For this reason, symmetrical molecules such as carbon tetrachloride (CCl_4), a methane derivative, in which the Cl-atoms occupy the corners of a regular tetrahedron (*cf. Sect. 2.6*), do not posses any dipole moment, although each individual C–Cl bond is strongly polarized (*cf. Table 1.6*). Some dipole moments of linkages have been estimated by comparison of the dipole moments of different compounds (*cf. Table 1.6*). However, as follows from *Fig. 1.7*, the dipole moments of dichloromethane (CH_2Cl_2) and chloroform ($CHCl_3$) are smaller than the corresponding values as calculated by vector addition of the C–Cl bond moment in methyl chloride (CH_3Cl). This is explained by the fact that the dipole moment of the C–Cl bond is not the same in all the above molecules. The

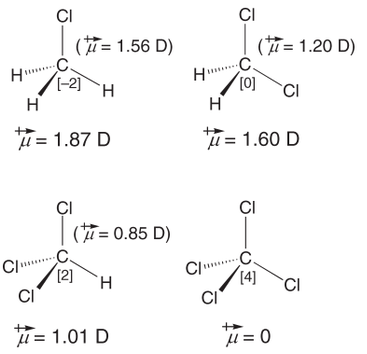


Fig. 1.7. *Dipole moments* (in Debye) *of chloromethanes.* For every compound, the bond moments of the C–Cl bonds as well as the oxidation number of the C-atom are given in round parentheses and square brackets, respectively.

Table. 1.6. *Dipole Moments of Some Covalent Bonds*

Bond	$\vec{\mu}$ [D]	Bond	$\vec{\mu}$ [D]
C–H	0.30	O–H	1.53
C–N	0.40	C–Cl	1.56
C–O	0.86	C=O	2.40
C=N	0.90	C≡N	3.60
N–H	1.31		

more Cl-atoms are bound to the C-atom, the higher its oxidation number becomes, thus increasing its electronegativity (*cf. Sect. 1.6*) and so the bond moments of the C–Cl bonds become smaller, as a result of the smaller difference in electronegativity between both atoms.

1.8. Intermolecular Interactions

As already mentioned in *Sect. 1.1*, an inherent property of living matter seems to be its striving towards the formation of more and more complex systems, even though such a process does not always satisfy the thermodynamic law of minimizing the free energy. Atoms thus combine to form molecules, molecules assemble to macromolecules (polymers), macromolecules aggregate to form 'supramolecules', microorganisms associate into symbionts, cells build up organs, individuals form societies, *etc*. On the molecular level, the forces that stabilize the aggregation of the constituent parts are known as weak interactions; they belong to three different types:

- *Van der Waals* forces (*cf. Sect. 1.9*),
- hydrogen bonding (*cf. Sect. 5.5*), and
- hydrophobic interactions (*cf. Sect. 9.16*).

Thus, whereas the assembly of atoms to build up molecules can be explained in terms of the formation of covalent bonds (*cf. Sect. 1.4*), the occurrence of more complex structures in macromolecules, as well as numerous phenomena that are observed in the condensed phase, can only be understood in terms of the effects of weak interactions. These may be both *intermolecular* (between molecules) and *intramolecular* (within the molecule). The dissociation energies of such weak interactions amount to only 4 to 25 kJ mol^{-1}, and so they are continually formed and interrupted either at room temperature or under physiological conditions. However, because large numbers of such interactions are usually present in one and the same macromolecule, they play important roles both in the shaping of the tertiary structures of proteins and nucleic acids, and in the stabilization of supramolecular structures such as membranes, ribosomes, the protein capsids of viruses, *etc*. Moreover, weak interactions are responsible for mutual recognition of molecules, such as in the interactions between enzymes and their substrates, between hormones and their receptors, *etc*.

1.9. *Van der Waals* Forces

Unlike in compounds such as table salt (Na^+Cl^-), the three-dimensional crystalline structure of which results from strong electrostatic interactions (known as *Coulomb* forces) between its constituent

Johannes Diderik van der Waals
(1837 – 1923)
After having finished elementary education in Leiden (the Netherlands), he became a schoolteacher. In 1864, he was appointed teacher at a secondary school at Deventer, and in 1866 he became director of a secondary school in the Hague. Although he had no knowledge of classical languages, and so was not allowed to take academic examinations, he continued studying at Leyden University in his spare time from 1862 to 1865, thus obtaining teaching certificates in mathematics and physics. In 1873, he obtained a doctorate degree, and in 1877 he was appointed professor of physics at the University of Amsterdam. In 1910, he was awarded the *Nobel* Prize for physics for his fundamental work on the equation of state for gases and liquids.

ions, the attractive forces between molecules are as a rule weak (*cf. Sect. 1.8*). Nevertheless, the fact that a large number of organic compounds exist as liquids or solids is the result of the so-called intermolecular *Van der Waals* forces. These are also electrostatic interactions, acting between dipoles that are either permanently present in the molecule or are generated for a short time by vibrations of the constitutive atoms about their equilibrium positions (*cf. Sect. 1.8*). As in the case of other electrostatic forces, the intensities of intermolecular *Van der Waals* forces in general increase when the distances between the interacting molecules decrease. Like atoms, however, molecules occupy volumes in space and, therefore, even in a crystal – as the most tightly packed of all possible spatial arrangements – they cannot come any closer to each other than a distance that is fixed by the repulsive interactions between the atoms that make up the molecule.

Actually, the distance between the molecules in a crystal is essentially determined by the atoms located on the molecules' surfaces, because such atoms belonging to different molecules 'touch' each other in the crystalline structure. As in the case of covalent bonds between atoms, the distances between the atoms that limit the approach of the molecules in a crystal (*i.e.*, the *Van der Waals* distances) can be experimentally measured by X-ray diffraction analysis, and from these the corresponding *Van der Waals* radii of the individual atoms can be determined (*cf. Table 1.7*). By this method, however, the *Van der Waals* radius of the C-atom has to be estimated from the radius of a methyl group (200 pm), because the C-atoms, unlike hydrogen or halogen atoms, are not located at the periphery of the molecule, and so do not experience direct contact in the crystal.

The shortest distance, in a molecular crystal, between atomic nuclei of the same element that do not belong to the same molecule is defined as the ***Van der Waals* distance**.

Half the *Van der Waals* distance is defined as the ***Van der Waals* radius** of an atom.

Van der Waals forces play a decisive role in the solubilities of substances in media in which chemical reactions take place. The electrostatic interactions between the dipole moments of the solvent molecules and those of the dissolved compound lead to *solvation* of the dissolved compound. This is an exergonic process (the free enthalpy decreases) that frequently stabilizes the reactive intermediates of chemical reactions (*cf. Sect. 3.8*).

Table. 1.7. Van der Waals *Radii of Some Atoms*

Atom	*r* [pm]	Atom	*r* [pm]
D	119.6	Cl	180
H	120	S	185
F	135	P	190
C	[a])	Br	195
O	140	I	215
N	150		

[a]) Not determinable experimentally.

1.10. Mechanisms of Chemical Reactions

A chemical reaction, as represented by a *chemical equation*, consists of the transformation of one or more compounds – called substrates or *starting materials* – into new compounds (*i.e.*, the reaction *products*), a process in which covalent bonds are broken and new ones are formed. A chemical equation that shows the structures both of the starting materials and of the products represents the *formal* course of the reaction. Accordingly, there are only five reaction types in organic chemistry:

- intramolecular rearrangement
- substitution
- addition
- elimination
- redox reaction

On the other hand, the mechanism of a chemical reaction describes the number of steps in which it takes place, the stereoselectivity of these steps (*cf. Sect. 3.6*), and which by-products may be formed. A (formal) substitution reaction, for example, may thus in principle take place by one of the following mechanisms:

- simultaneous cleavage of a covalent bond and formation of a new one (*cf. Sect. 5.13*),
- successive cleavage of a covalent bond and formation of a new one (*cf. Sect. 2.14* and 5.15),
- elimination and subsequent addition (*cf. Sect. 9.23*), and
- addition and subsequent elimination (*cf. Sect. 4.4*).

In older textbooks of organic chemistry, the atoms or functional groups involved in a chemical reaction were frequently encircled, thus indicating which structural modifications of the substrates take place as a result of the reaction. Although this symbolism may be still useful as a mnemonic aid (*cf. Fig. 1.8*), it does not take into consideration the fact that, in most cases, the transformation of the substrate of a chemical reaction into the products takes place through several *reaction steps* with formation of short-lived intermediates that eventually decompose to yield the reaction products (*cf. Sect. 1.12*).

The formation of, for example, an ester through a reaction between an alcohol and a carboxylic acid (*cf. Sect. 9.15*) involves the substitution of the alkyl residue of the alcohol for the H-atom in the OH group of the carboxylic acid, a process that may be *formally* interpreted as the release of a H_2O molecule with combination of the remaining atoms to form the ester molecule. *A priori*, however, there are two possible ways to release a H_2O molecule from the substrates (*cf. Fig. 1.9*), with elucidation of the *mechanism* of the reaction being the only way to differentiate between them.

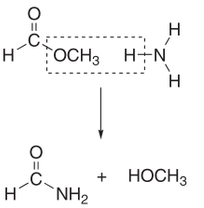


Fig. 1.8. *Chemical equation for the reaction between ammonia and methyl formate*

In the mechanistic sense, an **intermediate** is a molecule that, in a multistep process, acts both as the product of a reaction step and as the substrate of the next one.

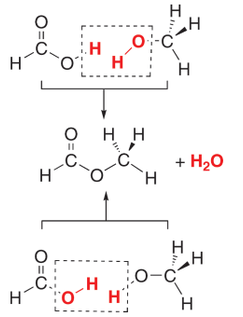


Fig. 1.9. *Formal synthesis of an ester through a reaction between the corresponding carboxylic acid and an alcohol through elimination of a H_2O molecule*

The elucidation of the mechanism of a chemical reaction assumes knowledge of the structures of *all* molecular species involved in the reaction, including those that cannot be either isolated or characterized by analytical methods, as well as the sequence of their formation in the course of the reaction. Actually, knowledge of the mechanisms of chemical reactions is a condition not only for understanding of the chemical reactivity of molecules but also for the design of the synthesis of either new or already known compounds, and the development of new reagents. Moreover, the mechanisms of most enzymatically catalyzed reactions can only be understood on the basis of analogous chemical processes (*cf. Table 1.8*).

Unfortunately, however, most of the mechanisms proposed for chemical reactions up to now are mere simplified *rationales* of the observed events, which indeed provide a basis for systematic subdivision of chemical reactions into a small number of reaction types, but often fail to predict the reactivity of a molecule on account of its structure. Nevertheless, the development of conceptual models relating to plausible mechanisms of chemical reactions is an important prerequisite for understanding of the connection between structures and reactivity of molecules. Actually, because the methodology of chemistry is basically empirical, the goal of being able to predict the structures of the products of a chemical reaction on the basis of knowledge of the structures of the reactants remains a long way off. Even so, conceptual models are of major heuristic significance, because they stimulate new experiments, thus keeping open the possibility of new discoveries.

Essentially, five experimental methods to elucidate the mechanism of a chemical reaction are available to the chemist. These are:

1) investigation of the reaction kinetics (if necessary in different solvents, and including the observation of isotopic effects),
2) either isolation of intermediates or analytical evidence of them,

Table. 1.8. *Mechanistically Analogous Chemical and Enzymatic Reactions*

Chemical reaction	Enzyme-catalyzed reaction	*cf. Sect.*
Aldol reaction	Biosynthesis of fructose 1,6-bisphosphate	*8.20*
Allylic rearrangement	Dimethylallyl pyrophosphate $\rightleftharpoons$ isopentenyl pyrophosphate	*3.9*
Benzoin condensation	2-Acetyllactate synthesis	*9.8*
Claisen condensation	Acetoacetyl coenzyme A biosynthesis	*9.23*
Claisen rearrangement	Chorismic acid $\rightarrow$ prephenic acid	*8.18*
Cope rearrangement	Germacrene A $\rightarrow$ β-elemene	*3.18*
Friedel–Crafts alkylation	Alkylation of 1,4-dihydroxy-2-naphthoic acid	*4.5*
Cationic polymerization	Biosynthesis of geranyl and farnesyl pyrophosphate	*3.19*
Keto–enol tautomerism	Glycerinaldehyde 3-phosphate $\rightleftharpoons$ dihydroxyacetone 1-phosphate	*8.16*
α-Ketol rearrangement	2-Acetyllactate $\rightarrow$ 3-hydroxy-2-oxovalerate	*9.8*
Meerwein–Ponndorf–Verley reduction	Acetaldehyde $\rightleftharpoons$ ethanol, among others	*8.11*
Phenol oxidation	Hypericin biosynthesis	*5.12*
S_N2 Reaction	Biosynthesis of adrenalin and phosphatidyl choline, among others	*7.8*
Wagner–Meerwein rearrangement	α-Terpineol $\rightarrow$ borneol, among others	*3.19*

3) execution of the same reaction with substrates that possess structures analogous to those of presumed intermediates,
4) substitution of the reactants with analogs of slightly altered structure (the most subtle modification being the replacement of appropriate atoms by their isotopes), and
5) ascertainment of the stereoselectivity of the reaction.

Examples of the application of the above methods to the investigation of reaction mechanisms of enzymatically catalyzed reactions are given in subsequent chapters. The most important concepts relating to investigation of reaction kinetics are explained in the following section.

1.11. Kinetics of Chemical Reactions

A first insight into the mechanism of a chemical reaction is given by determining the *reaction kinetics*, *i.e.*, the dependence of the reaction rate on the concentrations of the reactants.

The reaction velocity or reaction rate, which is dependent on the temperature, is defined as the modification of the concentrations (c) either of the reactants or of the reaction products per unit time (usually seconds); that is, reaction rate $= dc/dt$. The reaction rate is usually expressed, therefore, as moles per liter per second ($\text{M}^{-1}\,\text{s}^{-1}$).

In the simplest case, the reaction rate is directly proportional to the concentration of a single component in the reaction medium, *i.e.*,

$$dc/dt = -k_1 c$$

the minus sign meaning that the substrate concentration decreases during the reaction. A reaction with kinetics that obeys the above equation is called a *first-order* reaction. The proportionality factor k is called the *rate constant*; its reciprocal ($\tau = 1/k_1$) is the average lifetime of the reaction substrate.

Integration of the above differential equation yields:

$$\ln\frac{c}{c_0} = -k_1 t \text{ or } c = c_0\, \text{e}^{-kt}$$

c_0 being the substrate concentration at the beginning of the reaction (*i.e.*, $t = 0$). This is an exponential equation containing the variables c and t, as represented graphically in *Fig. 1.10*.

A corresponding expression for the increase in the product concentration during the reaction is obtained by integrating the following differential equation (*cf. Fig. 1.10*):

$$\frac{dc}{dt} = +k_1 c$$

Because the reaction rate depends on the concentrations of the reactants, it does not remain constant during the reaction. In general,

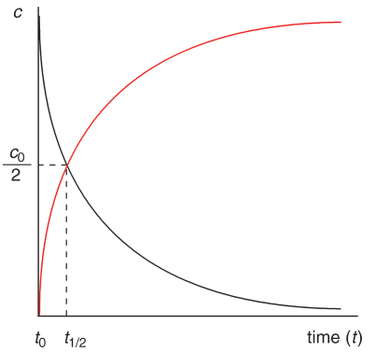


Fig. 1.10. *Changes in the substrate* (——) *and the product* (——) *concentrations as a function of time in a first-order reaction*

the reaction rate decreases either until an equilibrium is attained, at which the rates of the forward and the reverse reaction are equal (see below), or until the substrate of the reaction is exhausted, as in the case of an irreversible reaction. Therefore, in order to compare the reaction rates of different reactions, it is more appropriate to use the so-called half-life period ($\tau_{1/2}$) – the time for the concentration of the substrate to be reduced to *half* its initial value – rather than the corresponding rate constants. In the case of a first-order reaction, the half-life period is obtained from integration of the corresponding differential equation between the limits $t = 0$ (at the beginning of the reaction) and $t = (\tau_{1/2})$, so:

$$\ln \frac{c_0/2}{c_0} = -k_1(\tau_{1/2} - 0), \text{ and therefore:}$$

$$\ln 2 = k_1 \cdot \tau_{1/2}, \text{ or: } \tau_{1/2} = \frac{\ln 2}{k_1}$$

(ln 2 being the natural logarithm* of 2 = 0.69315)

Thus, in the case of reactions of the first order, the unit of the rate constant (k_1) is that of the reciprocal of time, quoted either as s^{-1} or as Hertz (Hz).

Exponential functions such as those discussed above do not only describe the kinetics of first-order reactions, including the decay of radioactive atoms, but they are also characteristic of a large number of natural macroscopic processes. Thus, in biology, both cell growth, in which a cell population doubles in size every generation through cell division, and the mortality rate within a population of living organisms depend on the sizes of the populations. Because changes in the rates of both events take place with the same sign, such processes are crucial for the evolution of the species in question.

In the case of a simple reaction in which two molecules of the same substrate combine to afford a single product molecule, according to the equation:

$$2\,A \rightarrow B$$

the reaction rate is proportional to the *square* of the substrate concentration:

$$\frac{dc}{dt} = -k_2 c^2$$

Such processes are reactions of the *second order*. Integration of the above differential equation yields:

$$\frac{1}{c} - \frac{1}{c_0} = k_2 t$$

c_0 being the substrate concentration at the beginning of the reaction: $t = 0$ (*cf. Fig. 1.11*).

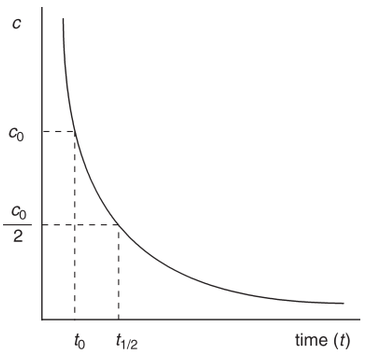


Fig. 1.11. *In a second-order reaction, the change in the substrate concentration, as a function of time, follows a hyperbolic equation*

Since the half-life period ($\tau_{1/2}$) of the reaction has elapsed when $c = c_0/2$, it results that:

$$\frac{2}{c_0} - \frac{1}{c_0} = k_2(\tau_{1/2} - 0)$$

and so the half-life period of a second-order reaction depends on the initial concentration of the substrate, *i.e.*,

$$\tau_{1/2} = \frac{1}{k_2 c_0}$$

the units of the rate constant (k_2) being $\text{mol}^{-1}\,\text{s}^{-1}$.

If more than one species are involved in the formation of the reaction product(s), the rate of the reaction depends, in general, on the concentration of each reacting species, as well as on the stoichiometric coefficients (m, n, p, ...) in the chemical equation:

$$\frac{dc}{dt} = -\,k\,[\text{A}]^m[\text{B}]^n[\text{C}]^p \ldots$$

The overall kinetic order of a reaction is equal to the sum of the exponents of the concentration terms of all reactant species in the rate expression, *i.e.*, $m + n + p + \ldots$

The **order of a reaction** is equal to the sum of the exponents of the reactant concentrations that determine the rate of the reaction.

Many chemical processes, however, take place as successions of *consecutive* reactions, rather than in a single reaction step. The simplest case arises when the consecutive processes are of the first order:

$$\text{A} \xrightarrow{k_\text{A}} \text{B} \xrightarrow{k_\text{B}} \text{C}$$

the reaction conditions being identical for all of the reaction steps. The overall reaction rate of such a multistep process is that of the slowest (*i.e.*, rate-determining) step. On the other hand, the rate of accumulation of B in such a system is equal to the difference in the rates of its formation from A and of its decomposition to C, so:

$$\frac{dc_\text{B}}{dt} = -\frac{dc_\text{A}}{dt} - \frac{dc_\text{C}}{dt}$$

Svante August Arrhenius (1859 – 1927)
submitted his doctoral thesis on the conductivity of electrolytes to the science faculty at Uppsala (Sweden) in 1884, where the fundamental importance of his publication was not well-understood by the faculty members. Later extensions of this very work earned *Arrhenius* the *Nobel* Prize for chemistry in 1903. He became professor of physics at the Stockolm Högskola in 1895. In 1906, he hypothesized that life might have spread from planet to planet by the transport of living spores through the universe; this theory, now known as panspermia, was later taken up and modified by the British astronomer and mathematician *Sir Fred Hoyle* (1915 – 2001) and others. *Arrhenius* also pointed out the existence of a 'greenhouse effect' through which small changes in the concentration of carbon dioxide in the atmosphere could considerably alter the average temperature of a planet.

The concentration of the substrate A, therefore, decreases steadily from the beginning, because it decomposes in an exponential manner, irrespective of the fate of its product B. Similarly, the concentration of C must increase steadily, and ultimately the value must become equal to the initial concentration of A. On the other hand, the amount of the intermediate B increases at first, but its rate of decomposition becomes greater as it accumulates, and so its concentration passes through a maximum, and finally it must become zero.

The rate of a chemical reaction is a *macroscopic* quantity that can normally be measured by simple analytical methods. Its quantitative dependence on *molecular* parameters that are essential to the mechanism of the reaction was established by *Arrhenius* in 1889, on the basis of the temperature dependence of the equilibrium constants of reversible reactions (*cf. Sect. 2.19*), which had been published five years earlier by *Van't Hoff* (*cf. Sect. 2.6*). The *Arrhenius* equation still today gives a sufficiently accurate relationship between the dependence of rate constants on the temperature and the so-called *activation energy* (E_a) of the reaction (*i.e.*, the lowest average energy that a molecule must possess for the reaction to be able to take place):

$$\frac{d \ln k}{dt} = \frac{E_a}{RT^2}$$

R being the gas constant* and T the absolute temperature*.

Integration of the *Arrhenius* equation yields:

$$\ln k = -\frac{E_a}{RT} + A \quad \text{or} \quad k = e^{-E_a/RT} \times e^A$$

Henry Eyring (1901 – 1981)
His most important work in the 1930s was his formulation of a general treatment of chemical reaction rates, now called transition-state theory. In 1938, he was appointed Professor of Chemistry at the University of Princeton in New Jersey, and in 1946 he became dean of the graduate school at the University of Utah in Salt Lake City, where he became involved in a great many other activities. Areas that had interested him at Princeton continued to maintain his curiosity, but many other areas involving molecular biology and highly applied matters also attracted his attention.

John Charles Polanyi (*1929)
Professor at the University of Toronto (Canada) from 1962, shared the 1986 *Nobel* Prize for chemistry with *Dudley Robert Herschbach* (*1932) and *Yuan Tseh Lee* (*1936) for their contributions relating to the dynamics of elementary chemical processes.

The constant of integration (A) is called the *frequency factor*: it represents the total frequency of encounters between the reactant molecules, irrespective of whether or not they possess sufficient energy to react.

The magnitude of the activation energy for a chemical reaction is considerably higher than the average kinetic energy of molecules at room temperature (*ca.* 2.5 kJ mol^{-1}). However, molecules are continually colliding with each other at rapid rates and exchanging kinetic energy. Therefore, at any given instant a few molecules are enough high in energy to forward the reaction.

As a result of the linear relationship between ln k and the reciprocal of the product $R \times T$, the parameters E_a and A can be determined from the inclination angle and the ordinate segment, respectively, of the straight line that is obtained by plotting ln k against $1/T$ in a two-dimensional Cartesian (*i.e.*, rectangular) coordinate system (*cf. Fig. 1.12*).

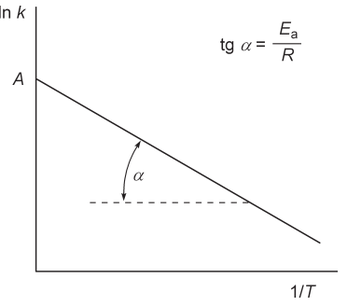


Fig. 1.12. *Determination of the activation energy* (E_a) *and the frequency factor* (A), *in the* Arrhenius *equation from the dependence of the reaction rate on the temperature*

An interpretation of the empirical parameters E_a and A is given in the context of the theory of absolute reaction rates formulated by *Eyring* and *Polanyi* in 1935. It is based on the concept of transition states of chemical reactions (*cf. Sect. 2.15*) and uses the methodology of statistical thermodynamics to determine the rate constant:

$$k = \frac{k_B T}{h} e^{-\Delta G/RT}$$

The *Eyring–Polanyi* equation thus contains, besides the free activation energy ($\Delta G^{\ddagger}$) of the reaction, exclusively fundamental physical constants, namely, the *Boltzmann* constant (k_B) and *Planck*'s constant (h).

In *Eyring*'s equation, the term $k_B T/h$ amounts to 6.2×10^{12} s^{-1} at room temperature (298.15 K). This value, which corresponds to the lifetime of the transition state, gives the approximate velocity of a reaction step, in the gas phase, for which both the activation enthalpy ($\Delta H^{\ddagger}$) and the activation entropy* ($\Delta S^{\ddagger}$) are equal to zero. In organic chemistry, the rate constants of most reactions that take place within one day lie in the range between 10^{-4} and 10^{-5} s^{-1}.

Because the energy content of the activated complex is without exception higher than that of the ground state, both $\Delta G^{\ddagger}$ and $\Delta H^{\ddagger}$ are positive quantities; $\Delta S^{\ddagger}$ may be either negative or positive, depending on whether translational, vibrational, or rotational degrees of freedom are lost or gained, respectively, in going from the ground state to the transition state.

In contrast to the *Arrhenius* equation, which is valid for any one rate constant, transition-state theory deals only with rate constants of individual reaction steps. As long as the *Arrhenius* equation is applied to the microscopic rate constant of a single reaction step, the parameters of the *Arrhenius* equation are related to those of the *Eyring's* equation by following equations:

$$A = \frac{k_B T}{h} e^{\Delta S^{\ddagger}/R} \quad \text{and} \quad E_a = \Delta H^{\ddagger} + \Delta nRT$$

Both equations are easily obtained from the definitions of the thermodynamic state-functions called internal energy (E), enthalpy (H), free energy (G), and entropy (S):

$$\Delta G = \Delta H - T\Delta S \quad \text{and} \quad \Delta H = \Delta E + \Delta nRT$$

(where R is the gas constant and T the absolute temperature of the concerned system).

For reactions in solution at constant pressure E_a and $\Delta H^{\ddagger}$ are related by:

$$E_a = \Delta H^{\ddagger} + (n - 1)\, RT$$

The number of molecules that are involved in the formation of the transition state of the rate-determining step of a chemical reaction is referred to as the reaction's **molecularity**.

n being the *molecularity* of the reaction step.

A reaction, in which a single molecule is involved in the formation of the transition state, is named *unimolecular*. Thus, in a unimolecular reaction ($n = 1$) the activation energy and the enthalpy of the transition state are equal ($E_a = \Delta H^{\ddagger}$). In a bimolecular reaction ($n = 2$) that takes place at room temperature ($T \approx 300$ K), the product RT is small; the difference between E_a and $\Delta H^{\ddagger}$ amounts to *ca.* 2.5 kJ mol^{-1}, a value that is within the limits of error of experimentally determined activation energies of chemical reactions.

The thermal decomposition of the *endo*-cyclopentadiene dimer, which takes place above 140 °C in the opposite sense to its formation from two cyclopentadiene molecules, may serve as an example of a unimolecular reaction (*cf. Fig. 1.13*). As determined by the *Arrhenius* equation, the reaction parameters are: $\Delta H^{\ddagger} = 138$ kJ mol^{-1} and $\Delta S^{\ddagger} = -8$ J mol^{-1} K^{-1}.

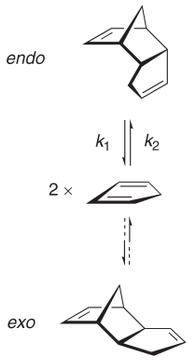


Fig. 1.13. *Reversible thermal dimerization of cyclopentadiene*

Moreover, because the reaction takes place in a single step, its rate depends only on the rate constant (k_1) and the substrate concentration, meaning that it is a reaction of the first order:

$$\frac{dc}{dt} = -k_1[\text{cyclopentadiene dimer}]$$

A chemical reaction in which the reactants are transformed into the products through a single transition state is called a **one-step reaction**.

As mentioned above, the decomposition of the cyclopentadiene dimer is a reversible reaction. Below 140 °C the dimer is present almost exclusively. The energy liberated by its formation is the result of the difference between the standard enthalpy of formation of the product and that of the reactants (*i.e.*, *two* cyclopentadiene molecules), so $\Delta H^0 = 188 - 2 \times 132 = -76$ kJ mol^{-1}.

The experimentally determined parameters for the dimerization of cyclopentadiene in the gas phase are: $\Delta H^{\ddagger} = 65$ kJ mol^{-1}and $\Delta S^{\ddagger} = -142$ J mol^{-1} K^{-1}. In relation to the activation entropy ($\Delta S^{\ddagger}$) of the decomposition of the cyclopentadiene dimer, the activation entropy of its formation is conspicuously high, because the number of mole-

cules present in the reaction volume decreases as a consequence of the formation of a 'compact' transition state from two cyclopentadiene molecules. On the other hand, the relatively low $\Delta H^{\ddagger}$ term for the dimerization of cyclopentadiene is characteristic of concerted reactions (*cf. Sect. 3.16*), in which bond formation accompanies bond breaking.

According to the stoichiometric equation of the reaction, the dimerization rate depends on the square of the concentration of the substrate, so it is a reaction of the second order:

$$\frac{dc}{dt} = -k_2[\text{cyclopentadiene}]^2$$

Although it may appear at first glance that the molecularity and the order of the reaction are synonymous, this redundancy is only true when particular conditions are met, as in the particular case of the dimerization of cyclopentadiene. The rate of a unimolecular reaction that takes place on a surface (*e.g.*, of an enzyme), for example, does not depend on the concentration of the substrate in the reaction medium. From the standpoint of kinetics, this reaction, the rate of which remains constant throughout all the reaction time, is a reaction of zero order (!), so:

$$\frac{dc}{dt} = -k$$

On the other hand, the order of a multistep reaction in which the unimolecular rate-determining step is not the first one may be greater than 1 (*cf. Sect. 5.15*). Moreover, bimolecular reactions that take place autocatalytically (see below), and those in which one reactant is present in a large excess (*e.g.*, as a solvent) are both reactions of *pseudo-first order*, because the concentration of one of the reactants remains practically unchanged during the reaction time.

The **rate-determining step** of a multistep process is the reaction step with highest free activation energy ($\Delta G^{\ddagger}$) (*cf. Sect. 3.8*).

As mentioned, the dimerization of cyclopentadiene is a reversible reaction. Therefore, at any temperature, cyclopentadiene and its dimer exist in thermodynamic equilibrium, so the composition of the reaction mixture is controlled thermodynamically. This means it is determined by the energy liberated by the reaction according to:

A chemical process is **thermodynamically controlled** if all its reaction steps are reversible.

$$\Delta G^0 = -RT \ln K$$

K being the equilibrium constant of the reaction, *i.e.*,

$$K = [\text{cyclopentadiene dimer}]/[\text{cyclopentadiene}]^2$$

Depending on its definition, the equilibrium constant (K) of a reversible reaction may be greater or smaller than 1. Except for acidity constants (*cf. Sect. 5.6*), K is commonly defined as the quotient of the concentrations of the components low in energy (in the numerator) divided by the concentrations of the components rich in energy (in the denominator), so that $K > 1$. Moreover, reaction equations are writ-

ten in such a way that the reaction takes place *exothermically* from the left to the right (*cf. Sect. 1.5*). This convention agrees with the second law of thermodynamics (*i.e.*, processes that take place spontaneously are irreversible). It must be pointed out, however, that the criterion of spontaneity for a reaction is not ΔG^0 but the free energy of the system as defined by: $\Delta G = \Delta G^0 + RT \ln [A]/[B]$; [A], [B], *etc*. being the concentrations of the reactants *outside* the equilibrium. Therefore, whether ΔG for a reaction is larger, smaller, or the same as ΔG^0 depends on the concentration of the reactants. Chemical reactions take place spontaneously if they are *exergonic*, which means if the free energy (ΔG) of the system decreases. Hence, the energy liberated by the reaction is equal to $-\Delta G^0$.

A thermodynamic equilibrium is a *dynamic* equilibrium; this means that, at any temperature, the rates both of the product formation and of the reverse reaction are equal:

$$k_2\,[\text{cyclopentadiene}]^2 = \frac{dc}{dt} = k_1\,[\text{cyclopentadiene dimer}]$$

$$\text{and hence: } K = k_2/k_1$$

Therefore, for any reversible reaction the equilibrium constant is equal to the ratio of the rate constants of the exothermic (exergonic) and the endothermic (endergonic) reaction.

In principle, any chemical reaction is reversible, meaning that the products are transformed back into the reactants under the same conditions as the reactants are converted into the products. This so-called principle of *microscopic reversibility* is a consequence of transition-state theory, according to which the same pathway that is traveled in the forward direction of a reaction will also be traveled in the reverse direction, because it affords the lowest energy barrier for either process.

A chemical reaction is virtually **irreversible** when the free energy of activation of the corresponding reverse reaction is higher than 128 kJ mol^{-1}.

In general, the more energy is liberated by a chemical reaction the lower is its activation enthalpy (*Bell–Evans–Polanyi* principle), and therefore, when the activation enthalpy of the reverse reaction is higher than 128 kJ mol^{-1} (*cf. Sect. 1.12*), the transformation of the products into the starting materials takes place so slowly that the reaction may be considered *irreversible*.

It must be pointed out, however, that the meaning of the term 'irreversible' is not the same in thermodynamics and in chemical kinetics. According to the second law of thermodynamics, an irreversible process (*e.g.*, the expansion of a gas into a vacuum) cannot be restored to its initial state without consuming energy. Likewise, the initial composition of a reaction mixture will not be spontaneously re-established when the thermodynamic equilibrium state is attained, although all reactions that lead to this equilibrium state are reversible.

The dimerization of cyclopentadiene, for instance, may afford two isomeric products that are known as *endo*- and *exo*-cyclopentadiene dimers (*cf. Fig. 1.13*). Because the mechanism of the formation of both isomers is the same, the experimentally observed *selectivity* of the reaction in favor of the *endo*-cyclopentadiene dimer, the energy

content of which is *higher* than that of the *exo*-isomer, can only be explained by the different rates of formation of the two isomers. Unlike reactions that lead to chemical equilibria, in which the products that are low in energy preponderate, most *irreversible* reactions are *kinetically* controlled, if the composition of the reaction mixture depends on the relative rates of formation of the different products.

A chemical reaction is **selective** when one reaction product out of several is formed preferentially.

This is the case in *competitive* processes, in which the reacting substance can be removed by two or more simultaneously occurring reactions, such as:

$$B \xleftarrow{k_B} A \xrightarrow{k_C} C$$

If it is assumed that both reactions are first order, the ratio of the concentrations of B and C at any instant is a constant equal to k_B/k_C, and independent of time. In most competitive processes, however, the substrate of a kinetically controlled chemical reaction consists of a mixture of isomers, each of which is transformed into a different product:

$$P_A \leftarrow A \rightleftharpoons B \rightarrow P_B$$

An example of this kind of process is given in *Fig. 3.37*. In such a case, the composition of the mixture of products depends only on the difference in the free energies of activation ($\Delta G^{\ddagger}$) of the transition states that lead to their formation, but not on the populations of the individual substrate isomers in the reaction mixture, provided that the energy barrier of the interconversion of the substrate isomers is low in relation to the activation energies of the corresponding reactions. Indeed, the composition of the overall product (P) from a kinetically controlled chemical reaction is determined by the ratio of the reaction rates with which the different products are formed:

$$P = \frac{d[P_A]}{dt} \Big/ \frac{d[P_B]}{dt} = \frac{[A] \cdot k_A}{[B] \cdot k_B}$$

Replacement of the rate constants by their corresponding expressions, as given by the *Eyring–Polanyi* equation, results in:

$$P = \frac{[A] \cdot k_B T/h \cdot e^{\left[\left(G_A^{\ddagger} - G_A\right)/RT\right]}}{[B] \cdot k_B T/h \cdot e^{\left[\left(G_B^{\ddagger} - G_B\right)/RT\right]}} = \frac{[A] \cdot e^{\left[\left(G_A^{\ddagger} - G_B^{\ddagger}\right)/RT\right]}}{[B] \cdot e^{\left[(G_A - G_B)/RT\right]}} = e^{\left(G_A^{\ddagger} - G_B^{\ddagger}\right)/RT}$$

The last expression results from: $[A]/[B] = K = e^{-\Delta G^0/RT} = e^{(G_A - G_B)/RT}$, K being the equilibrium constant of the reaction: $A \rightleftharpoons B$, according to the thermodynamic equation: $\Delta G^0 = -RT \ln K$. The value of P, therefore, depends exclusively on the difference between the free enthalpies of the transition states that lead to the products P_A and P_B, respectively. This is the so-called *Curtin–Hammet* principle.

As mentioned before, the rates of chemical reactions depend on the temperature. This temperature dependence can be easily evalu-

ated by dividing the *Eyring* equation of a reaction that is carried out at temperature T' by the corresponding equation of the same reaction taking place (at a slower rate) at temperature $T < T'$, so:

$$\ln (k'/k) = \ln (T'/T) + (\Delta H^{\ddagger}/R) (T' - T/T' \times T')$$

If it is assumed that $\Delta H^{\ddagger} \approx 118\ \mathrm{kJ\,mol^{-1}}$, an average value for a large number of chemical reactions with half-life periods amounting to 1 hour at $T = 373$ K (*i.e.*, 100 °C), the following expression is obtained for the ratio of the velocity constants on increasing the reaction temperature by $\Delta T = 10$ °C:

$$\ln (k'/k) = 118 \times 10/8.31 \times 373^2 \approx 1 \text{ and hence: } k'/k = 2.718$$

This is the known rule of thumb that states that an increase in the reaction temperature by *ca.* 10 °C results in a doubling of the reaction rate.

A **catalyst** reduces the free energy of activation of a chemical reaction by modifying the structure of the transition state; it does not, however, affect the position of equilibrium in a reversible reaction.

Furthermore, the rate of a chemical reaction depends on the effect of so-called *catalysts* (Gr.: καταλυσις (*katálysis*), meaning dissolution); these are components of the reaction mixture that modify (in particular, increase) the reaction rate without affecting the position of equilibrium in a reversible reaction. Because a catalyst, in general, is involved in the formation of the transition state, but is released unchanged chemically from the reaction products at the end of the reaction, its concentration in the reaction medium remains unaltered. Indeed, both the rate of formation of the products and the rate of the reverse reaction depend on the catalyst concentration:

$$\frac{dc}{dt} = k\ [\text{substrate(s)}]\,[\text{catalyst}]$$

Particularly interesting from the standpoint of biology and theory of evolution are *autocatalytic reactions*: chemical reactions in which the product catalyzes its own formation. Because the reaction rate increases when the concentration of the catalyst increases, an autocatalytic reaction is characterized by an initial period (so-called *period of induction*) during which the reaction process appears to be quiescent; thereafter, however, the reaction rate suddenly increases (*cf. Fig. 1.14*).

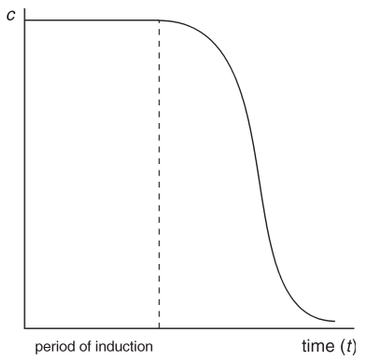


Fig. 1.14. *Autocatalytic reactions are characterized by the sigmoidal* (S-shaped) *dependence of reaction velocity on time*

The same behavior is observed in some enzymatically catalyzed reactions, in which the activity of the enzyme is enhanced by a change of conformation – a so-called *allosteric* modification (*cf. Sect. 9.31*) – that is induced by the substrate.

As already mentioned, study of the kinetics of a chemical reaction allows a preliminary insight into its mechanism; however, more detailed evidence of the real mechanism is only obtainable through more sophisticated methods.

1.12. 'Stable' and Reactive Molecules

There is hardly a more confusing concept in chemistry than that of 'stability'. In everyday language, stability is associated with unchangingness (Lat.: *stare*, meaning to stand, to be standing), thus ignoring the fact that economic systems as well as social order often survive due to their own powers of adaptation. In physics, a material system such as a pendulum is called stable if it returns to its initial state as soon as external forces (other than gravitation) are cancelled. Hence, the physical stability of a system is associated with its potential energy. Chemistry, on the other hand, deals with transformations of substances that only exceptionally revert completely to their initial states. It is, therefore, incorrect to speak about 'stability' in the physical (thermodynamic) sense, because this concept applies only to systems consisting of two or more substances that coexist in equilibrium, and, therefore, the energy liberated on establishing the equilibrium is a measure of the relative energy contents of both systems.

However, if the term 'stability' means the inertness of a substance towards a reaction, its energy content in relation to that of the reaction products does not play any role. Methane, for example, is a compound rich in energy in the customary sense, although it belongs to the class of saturated hydrocarbons that are characterized by their lack of reactivity. As a matter of fact, methane molecules have typically remained unchanged in their natural deposits for over hundreds of millions of years, as long as no oxygen has come into contact with them. In contrast, a mixture of methane and oxygen is a so-called *metastable** system, *i.e.*, it spontaneously transforms itself, liberating a considerable amount of energy (the heat of combustion), into a stable state consisting of CO_2 and H_2O, the moment the reaction is initiated (*cf. Sect. 2.14*). In chemistry, therefore, the concept of molecular 'stability' can be only defined in connection with the species that may be involved in a reaction, and it depends on the rates of the reactions that a particular substance may undergo with other compounds.

Individual molecules are *intrinsically kinetically* stable if they do not spontaneously decompose into smaller molecules. A definition of kinetic 'stability' can be formulated by assuming arbitrarily that the half-life period (*i.e.*, the time required for half the quantity of a substance to be consumed in a unimolecular process; *cf. Sect. 1.11*) amounts to 32 years (*cf. Table 1.9*).

In the customary sense, a molecule is intrinsically **stable** if the activation energy of its unimolecular decomposition amounts to more than 128 $kJ\,mol^{-1}$.

The thermodynamic 'stability' of a molecule, on the other hand, refers to its energy content relative to that of the free atoms from which the molecule is made (*cf. Sect. 1.5*). Therefore, comparison of the thermodynamic stabilities of molecules of different atomic compositions (*e.g.*, methane, nitrogen, water, carbon dioxide, *etc.*) is pointless, because they lack a common reference for their energy contents.

Table. 1.9. *Relationship between the Free Activation Energies* ($\Delta G^{\ddagger}$) *at 298.15 K* (= 25 °C) *of Reactions of the First Order and the Corresponding Kinetic Parameters.* For purposes of comparison, some physical processes are also included.

$\Delta G^{\ddagger}$ [kJ mol^{-1}]	k [s^{-1}]	$\tau_{1/2}$ [s]	Event
0	6.2×10^{12}	10^{-13}	Interval of an electron transition (10^{-15} to 10^{-16})
12	5×10^{10}	1.4×10^{-13}	Lifetimes of conformers (*cf. Sect. 2.19*)
42	3×10^{5}	2×10^{-6}	Lifetimes of excited electron configurations (10^{-6} to 10^{-9} s)
84	10^{-2}	51	Fast chemical reactions
128	7×10^{-10}	10^{9} (= 32 years)	Arbitrary limit of isolation of individual chemical species
168	3×10^{-17}	2×10^{16} (= 729×10^{6} years)	Estimated age of the earth 14.5×10^{16} s (= 4600 million years)

In general, an organic molecule made up of atoms belonging to the second row of the periodic table, containing fewer than eight electrons in their valence shells, is considered intrinsically unstable, and hence *reactive* (so-called *octet rule*).

Any atom or group of atoms that is attached to another atom is a **ligand** (Lat.: *ligare* meaning to bind) of the latter.

How many *ligands*, within a molecule, are bound to a particular atom is given by that atom's *coordination number*. Carbon atoms may display coordination numbers from 1 to 5 (*cf. Table 1.10*). Carbon atoms of coordination numbers 2, 3, and 4 are referred to as *digonal*, *trigonal*, and *tetragonal*, respectively (in the quantum-mechanical* theory of the covalent bond (*cf. Sect. 2.6*), the same atoms are called sp-, sp^2-, and sp^3-hybridized, resp.).

The **coordination number** of an atom in a molecule is the number of ligands that are bound to that particular atom.

The coordination number of an atom is not synonymous with its '*valence*'. Although this concept is rarely used in modern chemistry, it is sometimes helpful to bear in mind that in drawings of molecular structures of common organic molecules the C-atoms are always tetravalent (*i.e.*, they each contain $4 \times 2 = 8$ electrons in their valence shells, according to the octet rule) and are, therefore, always surrounded with four lines that symbolize covalent bonds. For most heteroatoms (*cf. Sect. 1.1*), the coordination number is usually lower than the valence number; valence electrons that are not used for bonding are called *nonbonding electrons*, or *lone-pair electrons.*

The **valence** number (or valency) of an atom is the number of electrons that this atom can either incorporate or release in order that its outer shell acquires the electron configuration of a noble gas (*i.e.*, eight electrons in the case of atoms belonging to the second row of the periodic table).

The term *valence* (Lat.: *valens*, meaning strong) was introduced in 1852 by *Percy Faraday Frankland* (1858 – 1946) to define the property of an atom to bind a certain number of other atoms. Later on (in 1919), based on *Lewis*' interpretation of chemical bonding, *Irving Langmuir* (1881 – 1957) introduced the name *covalent* (presumably from Lat.: *co-* = *cum*, meaning together, with) for chemical bonds between atoms, in contrast to electrostatic interactions between ions.

Molecules, however, include not only the submicroscopic components of substances that are considered 'stable', but also groups of atoms that exist for very short times as intermediates of chemical reactions (*cf. Table 1.10*). Of the compounds that a single C-atom may form with H-atoms, only CH_4 is in the customary sense stable (although metastable in the presence of oxygen). All other molecules represented in *Table 1.9* are too reactive to be isolated under conven-

Table. 1.10. *Currently Known Carbon Hydrides That Contain Only One C-Atom*

Number of coordination	Electrical charge in the molecule		
	− 1	0	+ 1
1	–	$\cdot\dot{C}H$ Methylydine	–
2	–	$:CH_2$ Methylene (a *carbene*)	–
3	$:CH_3^-$ methanide (a *carbanion*)	$\cdot CH_3$ methyl (a *radical*)	CH_3^+ methylium (a *carbenium* ion)
4	$\cdot CH_4^-$ (a radical anion)	CH_4 methane (an *alkane*)	$\cdot\, CH_4^+$ (a radical cation)
5	–		CH_5^+ Methanium (a *carbonium* ion)

tional conditions, either because they do not obey the octet rule (*i.e.*, $\cdot CH_4^-$, $:CH_2$, $\cdot CH_3$, $\cdot\dot{C}H$, CH_3^+, and $\cdot\, CH_4^+$), or because they are too strong as bases ($:C\dot{H}_3^-$) or acids (CH_5^+). However, as isolated molecules in the gas phase (*e.g.*, in interstellar space!) they may exist for quite long periods.

Carbon monoxide (CO) is the only diatomic carbon derivative in which the coordination number of the C-atom is equal to 1 that can be isolated under ordinary conditions. Carbon atoms of coordination number 2 occur in two 'stable' compounds, each containing only one C-atom: namely carbon dioxide (CO_2) and carbon disulfide (CS_2). Moreover, C-atoms may also have coordination numbers of 2 or 3 in compounds containing multiple C,C bonds, such as acetylene ($HC{\equiv}CH$) and ethylene ($H_2C{=}CH_2$), respectively.

Besides the reactants and the products, a number of short-lived intermediates, belonging to three different kinds of species (*cf. Table 1.11*), are involved in any chemical reaction. Moreover, depending on their constitutions, reactive intermediates belong to the seven different types listed in *Table 1.12*.

Table. 1.11. *Types of Molecules*

- 'Stable' molecules
- Reactive intermediates
- Transition states
- Excited states

Table. 1.12. *Reactive Intermediates in Organic Chemical Reactions*

- Free radicals
- Radical cations*
- Radical anions*
- Carbenes
- Carbenium ions
- Carbonium ions
- Carbanions*

Whereas the significance of *free radicals* in cell metabolism has been increasingly recognized recently, *carbenes* have yet to be considered as intermediates in enzymatic reactions. For this reason, neither the formation nor the reactivity of carbenes have been considered in this book, although they may play roles in enzymatic reactions that have previously been interpreted in terms of the involvement of carbenium ions as intermediates (*cf. Sect. 3.19*).

Both carbonium and carbenium ions are *carbocations* (in the nomenclature of organic compounds the suffix *-ium* indicates that the corresponding molecule is a positively charged ion). Before 1972, carbenium and carbonium ions were known as 'classical' and 'nonclassical' carbonium ions, respectively. Actually, a carbenium ion (*George A. Olah*, 1972) is the conjugate acid of a carbene. That is, a carbene is reversibly converted into a carbenium ion by reaction with a proton:

$$:CH_2 + H^+ \rightleftharpoons CH_3^+$$

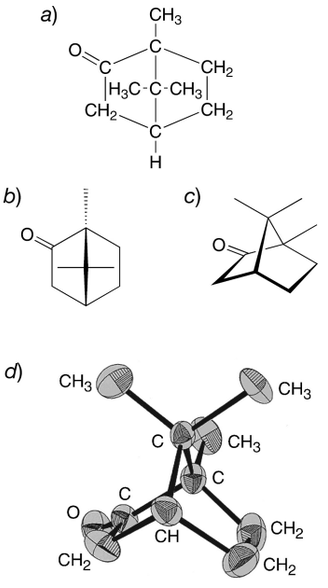


Fig. 1.15. *Different representations of the molecular structure of (+)-camphor.* In picture *d*, known as an ORTEP (Oak Ridge Thermal Ellipsoid Plot) drawing after the National Laboratory in Tennessee (USA), the boundaries within which the C-atoms and the O-atom vibrate are represented by ellipsoids (*cf. Sect. 1.5*). Modified from C. A. Bear & J. Trotter, *Acta Crystallogr., Sect. B* **1975**, *31*, 903–904. (Courtesy of Professor *James Trotter*, Chemistry Department of the University of British Columbia, Vancouver)

The **constitution** of a molecule describes the nature of the atoms and their sequence (*i.e.*, connectivity between them) in the molecule.

The **configuration** of a molecule describes the spatial arrangement of its constitutive atoms which distinguishes stereoisomers that cannot be interconverted by rotation about covalent bonds.

Carbonium ions, in which the coordination number of the C-atom is 5 (*cf. Table 1.10*), mainly occur as intermediates of molecular rearrangements (*cf. Sect. 5.21*). A compound of this kind, in which the C-atom is hypervalent, is possible because the electrons in the 1s shell participate in the bonding of the fifth ligand. Carbonium ions (*e.g.*, CH_5^+) have been detected both in the gas phase and – in solution – in the presence of so-called *superacids* (*i.e.*, mixtures of hydrofluoric acid (HF) or of fluorosulfonic acid (FSO_3H) either with SbF_5 or with another powerful *Lewis* acid), the acidities of which can be more than 10^{12} times higher than that of sulfuric acid (H_2SO_4).

Radical cations* are observed in the gas phase, for example, when organic molecules become ionized by a collision with electrons (*cf. Sect. 2.7*). On the other hand, radical anions, such as the *ketyl radical* (*cf. Sect. 8.11*), are frequently formed, in solution, as intermediates of reactions between organic molecules and alkali metals. Radical anions derived from alkyl halides thus serve as precursors of carbanions that are in general not accessible by deprotonation* of an alkane (*cf. Sect. 8.16*). According to the nomenclature of organic compounds, carbanions are characterized by the suffix *-ide* (*cf. Table 1.10*).

1.13. Molecular Shape

Although molecules may reach macroscopic dimensions (DNA molecules containing 4×10^6 base pairs are 1.36 mm long, for example), the vast majority of molecules consist of a few atoms up to a few hundred atoms and are, therefore, of submicroscopic size. Molecules are characterized both by their atomic compositions (*i.e.*, their molecular formulae) and their *structures* (*i.e.*, the connectivities between the constitutive atoms, including their spatial arrangements.

Any representation of a molecular structure should, therefore, show the positions of all atoms in the molecule. However, it would sometimes be confusing to represent every single atom by its chemical symbol (*cf. Fig. 1.15,a*). On the other hand, the molecular structure represented in *Fig. 1.15,a*, is incomplete, because it shows only the *constitution*, and not the *configuration*, of the molecule (*cf. Sect. 1.4*). For this reason, the molecular structures commonly used by chemists are mere ideograms that are often unintelligible to outsiders.

By convention, C- or H-atoms in a molecule are not represented by their chemical symbols, with the exception of H-atoms that are bound to heteroatoms. Carbon atoms are tacitly presumed to be located at the intersecting points of two or more bonds. Moreover, as many H-atoms as are required by the coordination number of each particular C-atom are imagined at the point that represents that C-atom, but they are not depicted in the molecular structure (*cf. Fig. 1.15,b*).

When a functional group that is involved in a chemical reaction needs to be emphasized, it is represented by the corresponding symbols of its constituent atoms. There are, however, no rules for this. It thus suffices in most cases, for example, to represent a methane molecule by its molecular formula (CH_4), and only rarely will the cor-

responding molecular structure be used for this purpose (*cf. Sect. 2.5*). On the other hand, it would be unsuitable to use a dot or a dash to represent a methane or an ethane molecule, respectively. Moreover, molecular structures must reproduce the spatial arrangements of the constitutive atoms. If the molecule, as a whole, is represented on a plane, there are two possibilities:

1) Bonds with ligands that are located above the plane are represented by heavy-type wedge bonds (*cf. Fig. 1.15*, *b*), and those with ligands that are situated below the plane by dashed bonds. These latter bonds may also be drawn wedge-shaped, pointing either towards the viewer or towards the ligand below the plane (*cf. Fig. 2.4*).
2) More appropriate, however, are perspective drawings that represent the real spatial structure of the molecule (*cf. Fig. 1.15*, *c*), as can be experimentally determined by X-ray diffraction analysis (*cf. Fig. 1.15*, *d*).

Furthermore, as explained in *Sect. 1.9*, use both of the *Van der Waals* radii and of the covalent radii of the constituent atoms, together with use of the bond angles, which can also be determined experimentally, enables the construction of molecular models that are usually indispensable for understanding of the three-dimensional shapes of molecules.

The use of three-dimensional molecular models can be traced back to *Le Bel* and *van't Hoff* in 1874 (*cf. Sect. 2.6*). The simplest of the molecular models are the so-called *ball-and-stick* models, the invention of which is attributed to *August von Hofmann* (*cf. Sect. 7.9*) around 1860. *Ball-and-stick* models, originally made of wood and assembled by connecting balls, representing atoms, by dowels, representing bonds, help visualization of the three-dimensional arrangement of the atoms in a molecule, but they are not to scale. On the other hand, the so-called *framework* models indicate correct bond distances and bond angles but do not show the atoms as such. Several inexpensive types of framework models, constructed from flexible plastic tubing and small metallic connecting pieces, are marketed for student use. A more sophisticated version of these models was developed by *André S. Dreiding*. A *Dreiding* model is constructed from stainless steel tubing of two different diameters, welded at a central point representing the atomic nucleus (*Fig. 1.16*). The bond lengths and angles precisely reproduce the average molecular dimensions, with 1 cm corresponding to 40 pm. To build the stereomodel, the units are assembled together by pushing the thinner tube of one unit into the larger tube of the other. A stop ensures that the tubes cannot be pushed together beyond the distance between the central points that corresponds to the internuclear bond distance in the molecule. In a molecular model constructed in this way, the free ends of each unit represent the nuclei of the H atoms. Although these models can be complemented by small semicircular plastic plates that can be rotated to describe

André Samuel Dreiding (1919 – 2013) taught chemistry from 1969 until his retirement in 1987 as a Professor at the University of Zurich. His contributions to the chemistry of natural products include the structure elucidation and synthesis of betalain pigments (*cf. Sect. 8.14*).

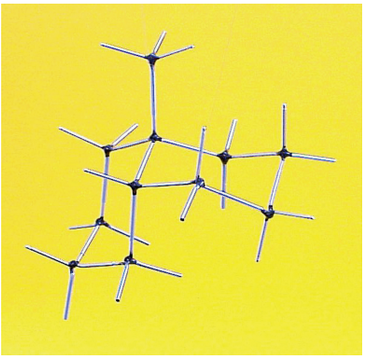


Fig. 1.16. Dreiding *model of cis-4a-methyldecalin* (*cf. Sect. 2.24*) (Author's photograph)

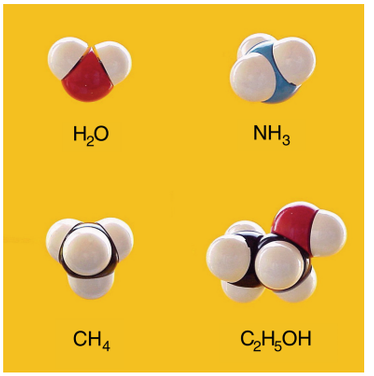


Fig. 1.17. *Space-filling molecular models of some simple molecules* (Author's photograph)

The **multiplicity** of a symmetry axis indicates how many times an object is brought into coincidence in the course of a complete rotation about this axis.

the *Van der Waals* volumes of the H-atoms, *Dreiding* models serve for the visualization of the three-dimensional structures of molecules rather than of the steric demands of their constituent atoms.

For this purpose, *space-filling* models, such as *Stuart–Briegleb* or *CPK*™ (*Corey–Pauling–Koltun*) molecular models are more appropriate (*cf. Fig. 1.17*). Indeed, such models show molecular volume at the expense of a clear representation of molecular geometry. The space-filling elements have the shape of spherical segments, the radii of which correspond to the *Van der Waals* radii of the chemical elements involved, and the distances between the centers of the sections, and the centers of the spheres correspond to the covalent radii. In addition, the dimensions of these elements are proportional to the atomic parameters, 100 pm on the molecular scale corresponding to 1.5 or 1.25 cm. Moreover, the different chemical elements are symbolized by different colors: white for H, black for C, red for O, blue for N, *etc.* (*cf. Fig. 1.17*).

Although molecular models of high-molecular-weight structures, such as polymers, proteins, or nucleic acids, can be constructed with space-filling molecular models (*cf. Fig. 1.18*), such structures are more appropriately represented nowadays by molecular graphics based on an appropriate electronic data processing (EDP) program (*cf. Fig. 1.19*). Such programs, however, use the same molecular parameters (*i.e.*, *Van der Waals* radii, bond lengths, and bond angles) as explained above. The computer-based pictures obtained in this way are chiefly used for *molecular modeling*, which enables interactions between substrate molecules, and proteins or enzymes to be simulated, thus assisting in the design of new chemotherapeutics.

A large number of molecular properties depend on the symmetries of the molecules' structures (*cf. Sect. 2.1*, as an example). Symmetrical objects, such as geometric figures, crystals, or molecules, may in general have axes and/or planes of symmetry that give rise to corresponding *symmetry operations* (*i.e.*, rotation about the axes and reflection in the planes, respectively). Both symmetry operations move the object, which is regarded as fixed at a point in space, into a position indistinguishable from (or equivalent to) the original position. Because symmetry operations meet the criteria for elements of a group, as defined in the branch of mathematics known as group theory, they belong, according to their number and geometrical properties, to so-called symmetry *point groups*. The number of different operations that can be performed in a group is called the order of the group. Axes of symmetry thus function as *symmetry operators*, represented in group theory as C_n, with n being the *multiplicity* or *order* of the axis. Planes of symmetry – denoted, in general, by σ – are differentiated by the subscripts v, h, and d, depending on their orientations with respect to the symmetry axes present in the molecule. A σ_v symmetry plane, therefore, contains a symmetry axis, a σ_h symmetry plane is perpendicular to a symmetry axis, and a σ_d symmetry plane bisects the angle between two symmetry axes.

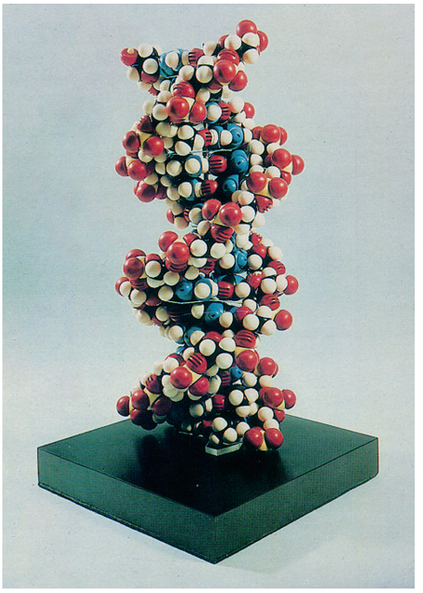


Fig. 1.18. *CPK Molecular model of a fragment of a B-DNA molecule.* (Courtesy of *Harvard Apparatus, Inc.*, Holliston, Massachusetts, USA)

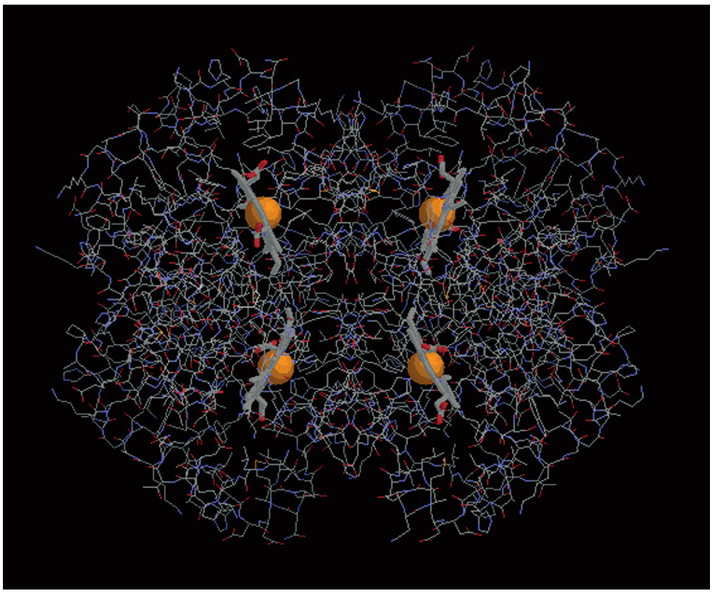


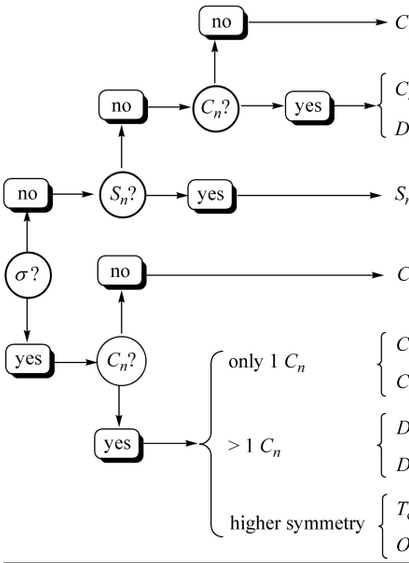
Fig. 1.19. *Three-dimensional structure of hemoglobin.* The four prosthetic groups (*i.e.*, the heme molecules) and the Fe^{II} ions bound to them are represented on a larger scale than the apoprotein (globin). (Courtesy of Professor *John J. Stezowski*, Chemistry Department at the University of Nebraska, Lincoln)

A molecule (or indeed any three-dimensional object) that does not possess *any* symmetry element is called *asymmetric*; it belongs to the point group C_1 (*cf. Sect. 2.8*). An object that possesses only symmetry axes but no planes of symmetry is called *dissymmetric*; if it possesses an axis of symmetry as the sole symmetry element, it belongs to the point group C_n. If more than one axis of symmetry is present, it belong to the point group D_n.

On the other hand, an object possessing a plane of symmetry but no axis of symmetry belongs to the point group C_s. Other C point groups contain, besides planes of symmetry, only axes of symmetry, whereas point groups containing more than one axis of symmetry are referred to as D groups. Point groups of higher symmetry (*e.g.*, T_d and O_h) are named according to the corresponding geometrical figures (*e.g.*, tetrahedron and octahedron, respectively). Most common organic molecules belong to the point groups specified in *Table 1.13*.

A particular category of symmetry groups symbolized by S_n contains objects that each possess only an *alternating axis of symmetry* (also called a rotation–reflection axis); these are brought into coincidence by rotation by $360^o/n$ about the axis and then reflected across a plane perpendicular to that axis (*cf. Sect. 2.8*).

Table. 1.13. *Flowchart for the Assignment of Organic Molecules to the Most Common Symmetry Point Groups*

Symmetry elements	Point group	Examples
	C_1 (asymmetric)	All molecules containig an asymmetric atom (*cf. Sect. 2.8*)
	C_n	1,2-Dichloroallene (C_2) (*cf. Sect. 3.27*)
	D_n (= C_n + n C_2)	Twist conformation of cyclohexane (D_2) (*cf. Sect. 2.22*)
	S_n	*anti* Conformation of (2*R*,3*S*)-2,3-dichlorobutane (S_2) (*cf. Sect. 2.11*)
	C_s	Chloroethylene
	C_{nv} (= C_n + n σ_v)	*cis*-1,2-Dichloroethylene (C_{2v}) (*cf. Sect. 3.2*)
	C_{nh} (= C_n + σ_h)	*trans*-1,2-Dichloroethylene (C_{2h}) (*cf. Sect. 3.2*)
	D_{nd} (= C_n + n C_2 + n σ_d)	Allene (D_{2d}) (*cf. Sect. 3.27*)
	D_{nh} (= C_n + n C_2 + σ_h)	Ethylene (D_{2h}), benzene (D_{6h}) (*cf. Sect. 4.1*)
	T_d (= 4 C_3 + 3 C_2 + 6 σ_d)	Methane, adamantane (*cf. Sect. 2.25*)
	O_h	(Octahedron)

1.14. Isomerism

Isomers are compounds that have identical molecular formulae.

Constitutional isomers are isomers that differ in the sequences of the atoms (*i.e.*, the connectivities between them) that constitute the molecules.

Stereoisomers are isomers of identical constitutions but differing in the spatial arrangements of their atoms.

Topoisomers are stereoisomers that arise from interlocking, intertwining, or threading of cyclic molecules.

Even molecules made up of only a few atoms can exist as isomeric structures (Gr.: *ίσος* (*ísos*) meaning equal, and *μέρος* (*méros*) meaning part). Particularly in the case of organic molecules, the number of possible structures in which the same constitutive atoms are linked together in a different way and/or with a different spatial arrangement increases enormously as the number of atoms in the molecule increases. Examples of the different types of isomeric organic molecules, as specified in *Table 1.14*, are given in the following sections. Unfortunately, it is not possible to achieve a rational classification of isomers on the basis of the presently existing definitions, because these definitions arose not systematically, but progressively, before all possible kinds of isomers were known. The classification shown in *Table 1.14* makes allowance for all known types of organic isomers; it has the didactic merit of being rigorously dichotomous (*i.e.*, each generic term leads to two specific terms that are mutually exclusive). However, a suitable term for the common feature of configurational isomers and topoisomers, not interconvertible by rotation about covalent bonds, is still lacking. Unlike *topomers* (Gr.: *τόπος* (*tópos*) meaning place), which can be interconverted without cleaving covalent bonds (*cf. Sect. 2.22*), *topoisomers* arise either from interlocking of cyclic molecules of sufficient size (*cf. Fig. 1.20*) or from intertwining of chemically independent subunits, thus forming single molecules.

Table. 1.14. *Dichotomous Classification of Isomers*

- **ISOMERS**
  - **Constitutional isomers**
    - Regioisomers (constitutional isomers that belong to the same class of compounds; *cf. Sect. 2.3*)
    - Chemoisomers (constitutional isomers that belong to different classes of compounds; *cf. Sect. 5.23*)
  - **Stereoisomers**
    - **Stereoisomers** which cannot be transformed into one another by rotation around covalent bonds
      - **Configurational isomers**
        - Enantiomers (*cf. Sect. 2.7*)
        - Diastereoisomers (*cf. Sect. 2.11*)
      - **Topological Isomers** (*cf. Fig. 1.21* and *Sect. 9.142*)
    - **Conformational isomers** (can be transformed into one another by rotation about covalent bonds)
      - σ-Conformers
        - Rotamers (*cf. Sect. 2.19*)
        - Atropisomers (*cf. Sect. 2.19*)
      - π-Conformers
        - (*Z*/*E*)-Isomers (*cf. Sect. 3.3*)
        - Axial-chiral allene (and cumulene) isomers (*cf. Sect. 3.27*)

For the sake of consistency, as imposed by the dichotomy of the classification shown in *Table 1.14*, (*Z*/*E*)-isomers (formerly known as *cis*/*trans* isomers) are classified in the group of conformational isomers as π-conformers, rather than as configurational isomers, as is usual in most textbooks on stereochemistry. Actually, however, (*Z*/*E*)-isomers are *not* configurational isomers, because they differ in the spatial arrangements of the ligands attached to two atoms that are linked together by a double bond, which can in principle rotate around the bond axis without separating the atoms. The reader will thus come to the conclusion that (*Z*/*E*)-isomers (*cf. Sect. 3.3*) have nothing in common with true diastereoisomers and hence with the cyclic *cis*/*trans* isomers dealt with in *Sect. 2.23*. Except for enantiomers, which are chiral according to their definition, and (*Z*/*E*)-isomers, which are achiral because they are planar, all other stereoisomers may be either chiral or achiral (*cf. Sect. 2.7*).



Fig. 1.20. *Cyclic and 'knot-like' cyclotripentacontane molecules* ($C_{53}H_{106}$) *are topoisomers of the same kind as are found in DNA molecules* (*cf. Fig. 9.143*) (Author's photograph)

1.15. Functional Groups

Despite some attempts made in the 1960s to use reaction mechanisms as a basis for structuring the contents of textbooks on organic chemistry, the concept of functional groups still proves extremely successful for systematizing study both of the structures and of the reactivities of organic molecules (*cf. Table 1.15*).

If a single molecule contains two functional groups sufficiently far apart from one another, the molecule should display the characteristic properties of both functional groups. The smaller the number of bonds between C-atoms that separate the two functional groups from one another, the greater is the effect that both functional groups exert in conjunction, and so they merge into a new functional group, the

A **functional group** is an atom or group of atoms in a molecule that is responsible, largely irrespective of the constitution of the molecule as a whole, for the characteristic physical and chemical properties of the molecule.

properties of which only partly correspond to those of the single components. An amido group, for example, although formally composed of an amino and a carbonyl group, has none of the basic properties characteristic of the amino group. The number of functional groups in organic chemistry is, therefore, quite large. The most important of these functional groups, as found in the molecules of natural compounds, are specified in *Table 1.15*. Moreover, organic compounds are classified on the basis of the (principal) functional group(s) present in their molecules.

1.16. Nomenclature of Organic Compounds

Organic compounds are named either by use of systematic nomenclature conventions or after their trivial (common) names, the etymologies of which are historically based. The huge number of known organic compounds calls for the use of systematic nomenclature, but a drawback of this is that the resulting names are often bewilderingly complex and so the (much shorter) trivial names are preferred. Internationally binding rules of systematic chemical nomenclature, first elaborated in 1892 in Geneva (Switzerland), have since 1947 been periodically revised and updated by the *IUPAC* (*International Union of Pure and Applied Chemistry*).

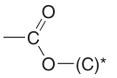
A compound name in the IUPAC nomenclature has three parts: a prefix, a parent name, and a suffix. The prefix specifies the position(s) and type(s) of the *substituent(s)* in the molecule, the parent name refers to the carbon skeleton of the molecule (as a rule an alkane), and the suffix identifies the functional group or class of compound to which the molecule belongs (*cf. Table 1.15*). The naming of an organic compound begins with the numbering of the parent hydrocarbon (*cf. Sect. 2.4*); the positions of the functional groups on the molecule are indicated through the numbers (so-called *locants*) of the C-atoms that carry the corresponding substituent. A systematic name contains *only one* suffix that corresponds to the *principal* functional group in the molecule. If more than one functional group is present, the others are indicated as prefixes. Which functional group is the principal one is decided by the following priority sequence:

Carboxylic acid > ester > amide > nitrile > aldehyde > ketone > alcohol > amine > imine > ether.

Any atom or group of atoms that replaces a H-atom in a molecule is called a **substituent** (Lat.: *substituere*, meaning to put in place of). An atom or molecule is said to be *substituted*, when one or more of the H atoms originally attached to it has been replaced by another ligand.

The location of a substituent with respect to the main functional group in a molecule is sometimes given with a Greek letter as the locant (α for the C-atom that is bound to the functional group, β for the next one, *etc.*), but with the use of ω for the *last* position in the C-chain, regardless of the chain's length. In modern nomenclature, however, such locants are used to indicate the position of a substituent with

Table. 1.15. *Most Important Functional Groups in Organic Chemistry*

Functional group	Class of compound	Prefix	Suffix	
—OH	Alcohols		hydroxy-	-ol
—O—(C)*	Ethers		oxy-	ether
	Amines		amino-	-amine
	Alkenes		–	-ene
	Aldehydes		formyl	-al
	Ketones		oxo-	-one
	Carboxylic acids		carboxy-	-oic acid
	Esters (carboxylic esters)		[alk]oxycarbonyl	-oate
	Amides (carboxamides)		carbamoyl-	-amide
	Imines		imino-	-imine
—C≡C—	Alkynes		–	-yne
—C≡N	Nitriles		cyano-	-nitrile

(C)* symbolizes any (digonal, trigonal, or tetragonal) C atom.

respect to a functional group in general, rather than as a constituent part of a systematic name.