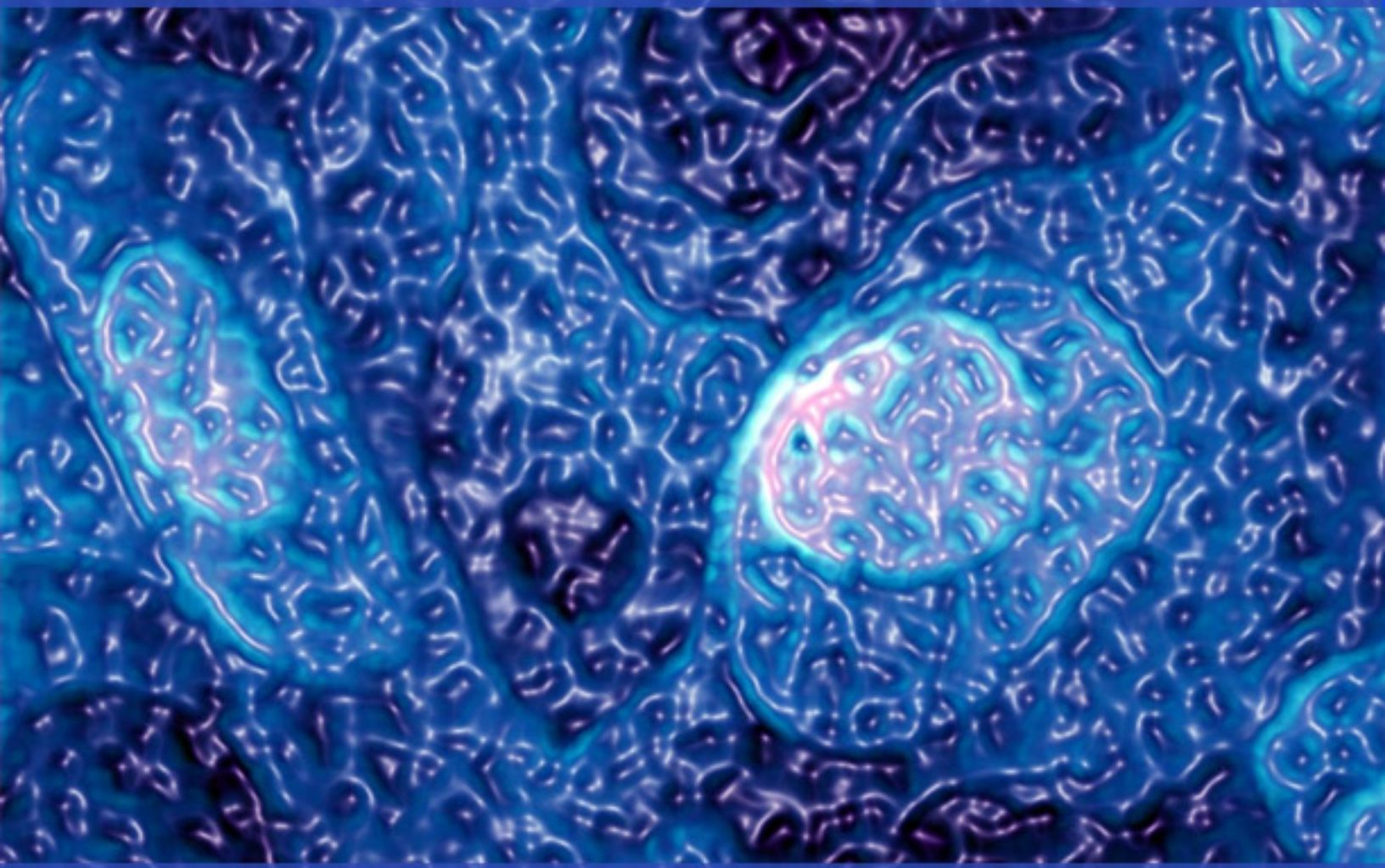
Molecular Engineering



Nohemi Altman

First Edition, 2012

ISBN 978-81-323-3085-1

Published by:
Research World
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Introduction

Molecular engineering is any means of manufacturing molecules. It may be used to create, on an extremely small scale, most typically one at a time, new molecules which may not exist in nature, or be stable beyond a very narrow range of conditions.

Today this is an extremely difficult process, requiring manual manipulation of molecules using such devices as a scanning tunneling microscope. Eventually it is expected to exploit life-like self-replicating 'helper molecules' that are themselves engineered. Thus the field can be seen as a precision form of chemical engineering that includes protein engineering, the creation of protein molecules, a process that occurs naturally in biochemistry, e.g., prion reproduction. However, it provides far more control than genetic modification of an existing genome, which must rely strictly on existing biochemistry to express genes as proteins, and has little power to produce any non-proteins.

Molecular engineering is an important part of pharmaceutical research and materials science.

Emergence of scanning tunneling microscopes and picosecond-burst lasers in the 1990s, plus discovery of new carbon nanotube applications to motivate mass production of these custom molecules, drove the field forward to commercial reality in the 2000s.

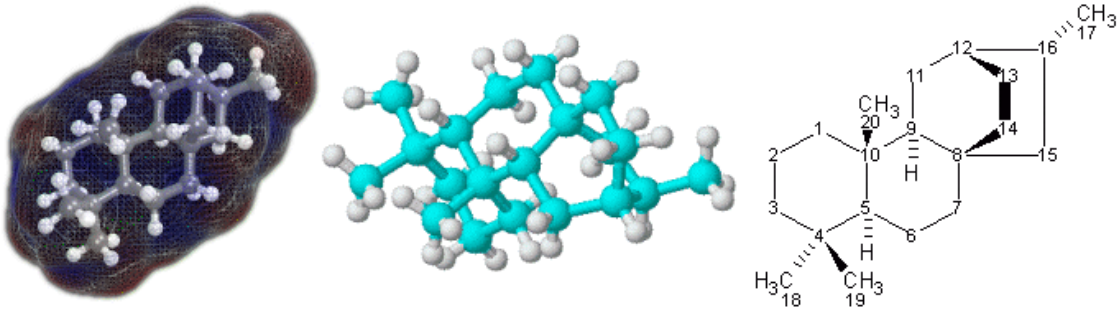
As it matures, it is seeming to converge with mechanical engineering, since the molecules being designed often resemble small machines. A general theory of molecular mechanosynthesis to parallel that of photosynthesis and chemosynthesis (both used by living things) is the ultimate goal of the field. This may lead to a molecular assembler, according to some, such as K. Eric Drexler, Ralph Merkle, and Robert Freitas, and of the potential for integrating vast numbers of assemblers into a kg-scale nanofactory.

Molecular engineering is sometimes called generically "nanotechnology", in reference to the nanometre scale at which its basic processes must operate. That term is considered to be vague, however, due to misappropriation of the word in association with other

techniques, such as X-ray lithography, that are not used to create new free-floating ions or molecules.

Chapter 1

Molecule



3D (left and center) and 2D (right) representations of the terpenoid molecule atisane

A **molecule** is an electrically neutral group of at least two atoms held together by covalent chemical bonds. Molecules are distinguished from ions by their electrical charge. However, in quantum physics, organic chemistry, and biochemistry, the term *molecule* is often used less strictly and applied to polyatomic ions.

In the kinetic theory of gases, the term *molecule* is often used for any gaseous particle regardless of its composition. According to this definition noble gas atoms are considered molecules despite the fact that they are composed of a single non-bonded atom.

A molecule may consist of atoms of a single chemical element, as with oxygen (O_2), or of different elements, as with water (H_2O). Atoms and complexes connected by non-covalent bonds such as hydrogen bonds or ionic bonds are generally not considered single molecules.

Molecules as components of matter are common in organic substances (and therefore biochemistry). They also make up most of the oceans and atmosphere. A large number of familiar solid substances, however, including most of the minerals that make up the crust, mantle, and core of the Earth itself, contain many chemical bonds, but are *not* made of

identifiable molecules. No typical molecule can be defined for ionic crystals (salts) and covalent crystals (network solids), although these are often composed of repeating unit cells that extend either in a plane (such as in graphene) or three-dimensionally (such as in diamond or sodium chloride). The theme of repeated unit-cellular-structure also holds for most condensed phases with metallic bonding. In glasses (solids that exist in a vitreous disordered state), atoms may also be held together by chemical bonds without any definable molecule, but also without any of the regularity of repeating units that characterises crystals.

Molecular science

The science of molecules is called *molecular chemistry* or *molecular physics*, depending on the focus. Molecular chemistry deals with the laws governing the interaction between molecules that results in the formation and breakage of chemical bonds, while molecular physics deals with the laws governing their structure and properties. In practice, however, this distinction is vague. In molecular sciences, a molecule consists of a stable system (bound state) comprising two or more atoms. Polyatomic ions may sometimes be usefully thought of as electrically charged molecules. The term *unstable molecule* is used for very reactive species, i.e., short-lived assemblies (resonances) of electrons and nuclei, such as radicals, molecular ions, Rydberg molecules, transition states, van der Waals complexes, or systems of colliding atoms as in Bose-Einstein condensate

History and etymology

According to Merriam-Webster and the Online Etymology Dictionary, the word "molecule" derives from the Latin "moles" or small unit of mass.

- **Molecule** (1794) – "extremely minute particle," from Fr. *molécule* (1678), from Mod.L. *molecula*, dim. of L. moles "mass, barrier". A vague meaning at first; the vogue for the word (used until late 18th century only in Latin form) can be traced to the philosophy of Descartes.

Although the existence of molecules has been accepted by many chemists since the early 19th century as a result of Dalton's laws of Definite and Multiple Proportions (1803–1808) and Avogadro's law (1811), there was some resistance among positivists and physicists such as Mach, Boltzmann, Maxwell, and Gibbs, who saw molecules merely as convenient mathematical constructs. The work of Perrin on Brownian motion (1911) is considered to be the final proof of the existence of molecules.

The definition of the molecule has evolved as knowledge of the structure of molecules has increased. Earlier definitions were less precise, defining molecules as the smallest particles of pure chemical substances that still retain their composition and chemical properties. This definition often breaks down since many substances in ordinary experience, such as rocks, salts, and metals, are composed of large networks of chemically bonded atoms or ions, but are not made of discrete molecules.

Molecular size

Most molecules are far too small to be seen with the naked eye, but there are exceptions. DNA, a macromolecule, can reach macroscopic sizes, as can molecules of many polymers. The smallest molecule is the diatomic hydrogen (H_2), with a length of 0.74 Å. Molecules commonly used as building blocks for organic synthesis have a dimension of a few Å to several dozen Å. Single molecules cannot usually be observed by light (as noted above), but small molecules and even the outlines of individual atoms may be traced in some circumstances by use of an atomic force microscope. Some of the largest molecules are macromolecules or supermolecules.

Radius

Effective molecular radius is the size a molecule displays in solution. The table of permselectivity for different substances contains examples.

Molecular formula

A compound's empirical formula is the **simplest** integer ratio of the chemical elements that constitute it. For example, water is always composed of a 2:1 ratio of hydrogen to oxygen atoms, and ethyl alcohol or ethanol is always composed of carbon, hydrogen, and oxygen in a 2:6:1 ratio. However, this does not determine the kind of molecule uniquely – dimethyl ether has the same ratios as ethanol, for instance. Molecules with the same atoms in different arrangements are called isomers. Also carbohydrates, for example, have the same ratio (carbon:hydrogen:oxygen = 1:2:1) (and thus the same empirical formula) but different total numbers of atoms in the molecule.

The molecular formula reflects the exact number of atoms that compose the molecule and so characterizes different molecules. However different isomers can have the same atomic composition while being different molecules.

The empirical formula is often the same as the molecular formula but not always. For example the molecule acetylene has molecular formula C_2H_2, but the simplest integer ratio of elements is CH.

The molecular mass can be calculated from the chemical formula and is expressed in conventional atomic mass units equal to 1/12 of the mass of a neutral carbon-12 (^{12}C isotope) atom. For network solids, the term formula unit is used in stoichiometric calculations.

Molecular geometry

Molecules have fixed equilibrium geometries—bond lengths and angles— about which they continuously oscillate through vibrational and rotational motions. A pure substance is composed of molecules with the same average geometrical structure. The chemical formula and the structure of a molecule are the two important factors that determine its

properties, particularly its reactivity. Isomers share a chemical formula but normally have very different properties because of their different structures. Stereoisomers, a particular type of isomers, may have very similar physico-chemical properties and at the same time different biochemical activities.

Molecular spectroscopy

Molecular spectroscopy deals with the response (spectrum) of molecules interacting with probing signals of known energy (or frequency, according to Planck's formula). Molecules have quantized energy levels that can be analyzed by detecting the molecule's energy exchange through absorbance or emission. Spectroscopy does not generally refer to diffraction studies where particles such as neutrons, electrons, or high energy X-rays interact with a regular arrangement of molecules (as in a crystal).

Theoretical aspects

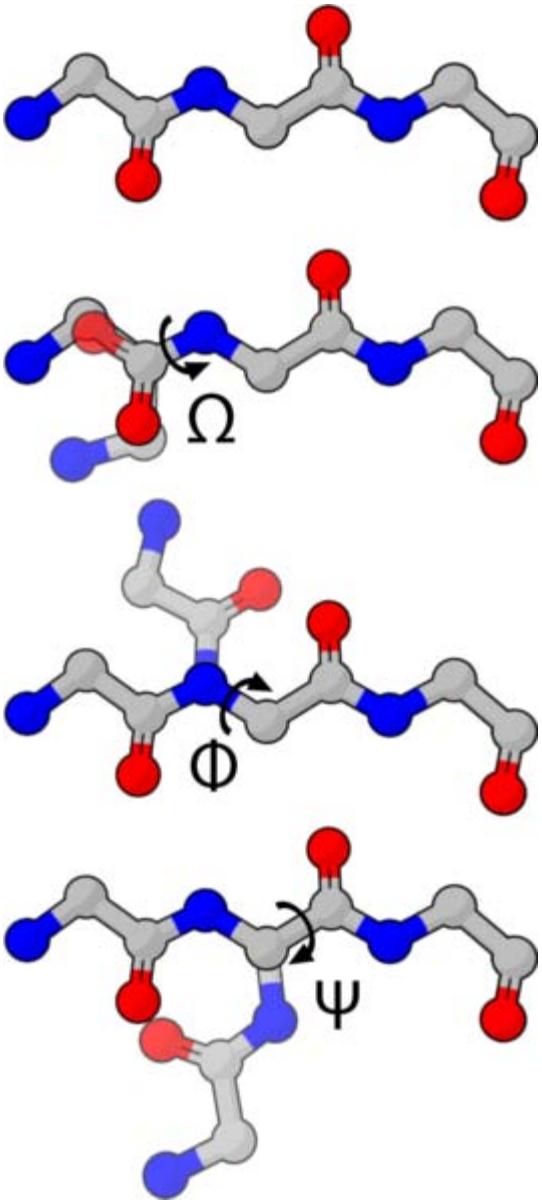
The study of molecules by molecular physics and theoretical chemistry is largely based on quantum mechanics and is essential for the understanding of the chemical bond. The simplest of molecules is the hydrogen molecule-ion, H_2^+, and the simplest of all the chemical bonds is the one-electron bond. H_2^+ is composed of two positively charged protons and one negatively charged electron, which means that the Schrödinger equation for the system can be solved more easily due to the lack of electron–electron repulsion. With the development of fast digital computers, approximate solutions for more complicated molecules became possible and are one of the main aspects of computational chemistry.

When trying to define rigorously whether an arrangement of atoms is "sufficiently stable" to be considered a molecule, IUPAC suggests that it "must correspond to a depression on the potential energy surface that is deep enough to confine at least one vibrational state". This definition does not depend on the nature of the interaction between the atoms, but only on the strength of the interaction. In fact, it includes weakly bound species that would not traditionally be considered molecules, such as the helium dimer, He_2, which has one vibrational bound state and is so loosely bound that it is only likely to be observed at very low temperatures.

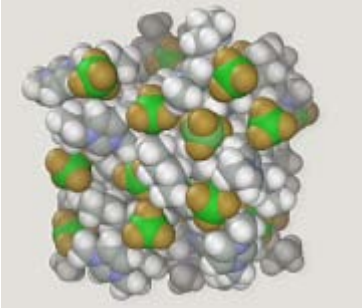
Chapter 2

Molecular Modelling and Molecular Design Software

Molecular modelling



The backbone dihedral angles are included in the molecular model of a protein.



Modelling of ionic liquid

Molecular modelling encompasses all theoretical methods and computational techniques used to model or mimic the behaviour of molecules. The techniques are used in the fields of computational chemistry, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies. The simplest calculations can be performed by hand, but inevitably computers are required to perform molecular modelling of any reasonably sized system. The common feature of molecular modelling techniques is the atomistic level description of the molecular systems; the lowest level of information is individual atoms (or a small group of atoms). This is in contrast to quantum chemistry (also known as electronic structure calculations) where electrons are considered explicitly. The benefit of molecular modelling is that it reduces the complexity of the system, allowing many more particles (atoms) to be considered during simulations.

Molecular mechanics

Molecular mechanics is one aspect of molecular modelling, as it refers to the use of classical mechanics/Newtonian mechanics to describe the physical basis behind the models. Molecular models typically describe atoms (nucleus and electrons collectively) as point charges with an associated mass. The interactions between neighbouring atoms are described by spring-like interactions (representing chemical bonds) and van der Waals forces. The Lennard-Jones potential is commonly used to describe van der Waals forces. The electrostatic interactions are computed based on Coulomb's law. Atoms are assigned coordinates in Cartesian space or in internal coordinates, and can also be assigned velocities in dynamical simulations. The atomic velocities are related to the temperature of the system, a macroscopic quantity. The collective mathematical expression is known as a potential function and is related to the system internal energy (U), a thermodynamic quantity equal to the sum of potential and kinetic energies. Methods which minimize the potential energy are known as energy minimization techniques (e.g., steepest descent and conjugate gradient), while methods that model the behaviour of the system with propagation of time are known as molecular dynamics.

$$E = E_{\text{bonds}} + E_{\text{angle}} + E_{\text{dihedral}} + E_{\text{non-bonded}}$$
$$E_{\text{non-bonded}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}$$

This function, referred to as a potential function, computes the molecular potential energy as a sum of energy terms that describe the deviation of bond lengths, bond angles and torsion angles away from equilibrium values, plus terms for non-bonded pairs of atoms describing van der Waals and electrostatic interactions. The set of parameters consisting of equilibrium bond lengths, bond angles, partial charge values, force constants and van der Waals parameters are collectively known as a force field. Different implementations of molecular mechanics use different mathematical expressions and different parameters for the potential function. The common force fields in use today have been developed by using high level quantum calculations and/or fitting to experimental data. The technique known as energy minimization is used to find positions of zero gradient for all atoms, in other words, a local energy minimum. Lower energy states are more stable and are commonly investigated because of their role in chemical and biological processes. A molecular dynamics simulation, on the other hand, computes the behaviour of a system as a function of time. It involves solving Newton's laws of motion, principally the second law, $\mathbf{F} = m\mathbf{a}$. Integration of Newton's laws of motion, using different integration algorithms, leads to atomic trajectories in space and time. The force on an atom is defined as the negative gradient of the potential energy function. The energy minimization technique is useful for obtaining a static picture for comparing between states of similar systems, while molecular dynamics provides information about the dynamic processes with the intrinsic inclusion of temperature effects.

Variables

Molecules can be modelled either in vacuum or in the presence of a solvent such as water. Simulations of systems in vacuum are referred to as *gas-phase* simulations, while those that include the presence of solvent molecules are referred to as *explicit solvent* simulations. In another type of simulation, the effect of solvent is estimated using an empirical mathematical expression; these are known as *implicit solvation* simulations.

Applications

Molecular modelling methods are now routinely used to investigate the structure, dynamics, surface properties and thermodynamics of inorganic, biological and polymeric systems. The types of biological activity that have been investigated using molecular modelling include protein folding, enzyme catalysis, protein stability, conformational changes associated with biomolecular function, and molecular recognition of proteins, DNA, and membrane complexes.

Molecular design software

Molecular design software is a software for molecular modeling, distinctive property of which is the presence of the special support for developing the molecular models.

In contrast to the usual molecular modeling programs such as the molecular dynamics and quantum chemistry programs, such software directly supports the aspects related to the construction of molecular models:

- Molecular graphics
- interactive molecular drawing and conformational editing
- building of polymeric molecules, crystals and solvated systems
- partial charges development
- geometry optimization
- support for the different aspects of Force Field development
- *etc*.

Comparative table of packages covering the major aspects of molecular design

3D - Molecular Graphics, **Mouse** - drawing molecule by mouse, **Poly** - polymer building, **DNA** - Nucleic acid building, **Pept** - Peptide building, **Cryst** - crystal building, **Solv** - solvent addition, **Q** - partial charges, **Dock** - docking, **Min** - optimization, **MM** - Molecular mechanics, **QM** - Quantum mechanics. **FF** - Support for Force Field development. **QSAR** - 2D, 3D and Group QSAR.

	3D	Mouse	Poly	DNA	Pept	Cryst	Solv	Q	Dock	Min	MM	QM	FF	QSAR	Homepage	Comments
AMBER				+	+			+		+	+		+		ambermd.org	Classical molecular modeling program
ArgusLab	+	+			+			+	+	+	+	+			Planaria Software	A molecular modeling, graphics, and drug design program
Ascalaph Designer	+	+	+	+	+	+	+	+		+	+	+	+		Agile Molecule	common molecular modeling suite
Avogadro (software)	+	+			+	+		+		+	+		+		OpenMolecules.net	Extensible, free, open source molecular editor

BALL / BALLView	+	+			+	+			+	+	+	+	+		ball-project.org	open source, C++ molecular modelling and visualization tool with scripting interface
BOSS								+		+	+	+	+		Yale University	OPLS inventor
DOCK								+	+	+	+				University of California	DOCK algorithm
Firefly (PC GAMESS)								+		+		+	+		Moscow State University	ab initio and DFT computational chemistry program
FoldX								+		+	+		+		CRG	A force field for energy calculations and protein design
Maestro	+	+		+	+	+	+	+	+	+	+	+	+	+	Schrodinger	A molecular modeling, visualization, and drug design program
Materials Studio	+	+		+			+	+	+	+	+	+	+	+	Accelrys	software environment
MedeA	+			+			+				+	+	+		Materials Design	software environment for inorganic materials science
MOE	+	+		+	+	+	+	+	+	+	+	+	+	+	Chemical Computing Group	Molecular Operating Environment
VLifeMDS	+	+							+	+	+	+	+	+	+	VLife Sciences Technologies Pvt. Ltd.	VLife Molecular Design Suite
NAB					+						+	+			Rutgers University	molecular manipulation language for nucleic acids
PCMODEL	+			+	+	+				+	+	+	+		Serena Software	common molecular modeling tool
SPARTAN	+	+			+	+			+	+	+	+	+		Wavefunction	molecular modeling tool with molecular mechanics and quantum chemical engines
StruMM 3D	+	+		+	+	+	+	+	+	+	+	+			Exorga Software	molecular modeling tool

(STR3DI32)										
TINKER				+		+	+	+	Washington University	freeware, tools for protein design
VEGA	+	+	+	+	+	+			Università degli Studi di Milano	a bridge between most of the molecular software packages

Chapter 3

Molecular Graphics

Molecular graphics (**MG**) is the discipline and philosophy of studying molecules and their properties through graphical representation. IUPAC limits the definition to representations on a "graphical display device". Ever since Dalton's atoms and Kekule's benzene, there has been a rich history of hand-drawn atoms and molecules, and these representations have had an important influence on modern molecular graphics. Here we concentrates on the use of computers to create molecular graphics. Note, however, that many molecular graphics programs and systems have close coupling between the graphics and editing commands or calculations such as in molecular modelling.

Relation to molecular models

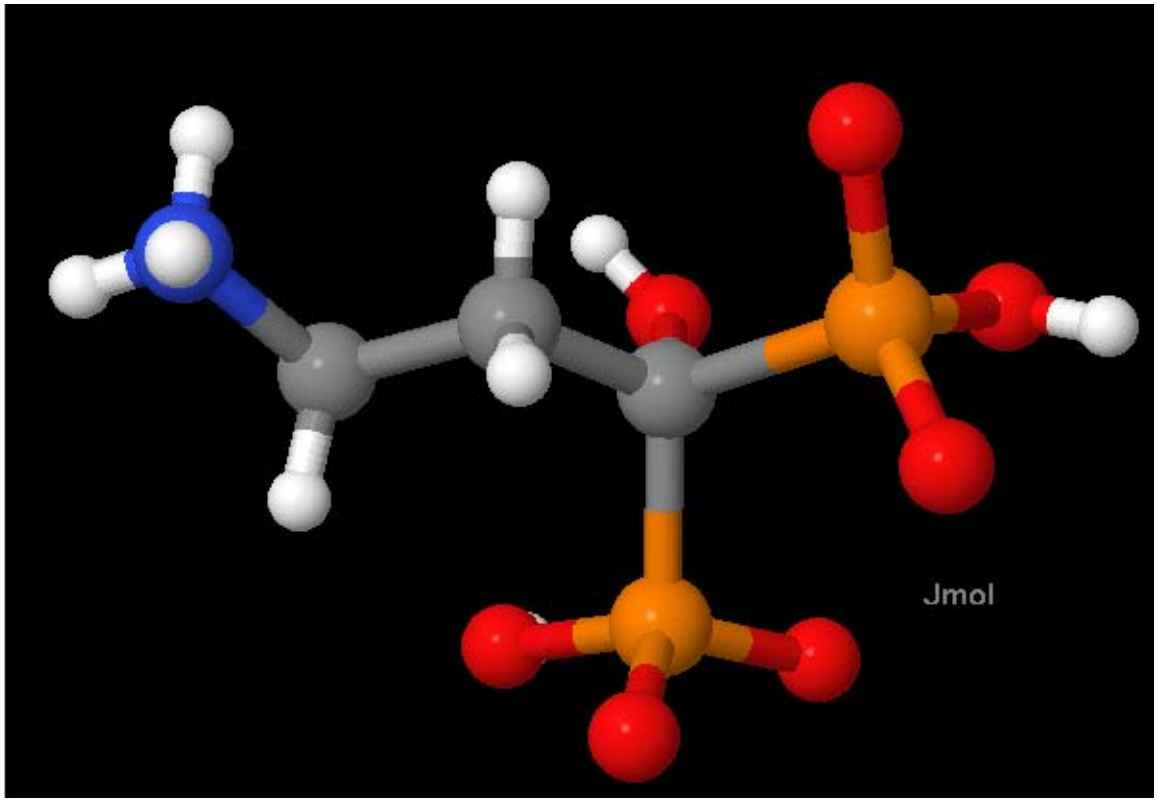


Fig. 1. Key: Hydrogen = white, carbon = grey, nitrogen = blue, oxygen = red, and phosphorus = orange.

There has been a long tradition of creating molecular models from physical materials. Perhaps the best known is Crick and Watson's model of DNA built from rods and planar sheets, but the most widely used approach is to represent all atoms and bonds explicitly using the "ball and stick" approach. This can demonstrate a wide range of properties, such as shape, relative size, and flexibility. Many chemistry courses expect that students will have access to ball and stick models. One goal of mainstream molecular graphics has been to represent the "ball and stick" model as realistically as possible and to couple this with calculations of molecular properties.

Figure 1 shows a small molecule ($NH_3CH_2CH_2C(OH)(PO_3H)(PO_3H)$-), as drawn by the Jmol program. It is important to realize that the colors and shapes are purely a convention, as individual atoms are not colored, nor do they have hard surfaces. Bonds between atoms are also not rod-shaped.

Comparison of physical models with molecular graphics

Physical models and computer models have partially complementary strengths and weaknesses. Physical models can be used by those without access to a computer and now can be made cheaply out of plastic materials. Their tactile and visual aspects cannot be easily reproduced by computers (although haptic devices have occasionally been built). On a computer screen, the flexibility of molecules is also difficult to appreciate; illustrating the pseudorotation of cyclohexane is a good example of the value of mechanical models.

However, it is difficult to build large physical molecules, and all-atom physical models of even simple proteins could take weeks or months to build. Moreover, physical models are not robust and they decay over time. Molecular graphics is particularly valuable for representing global and local properties of molecules, such as electrostatic potential. Graphics can also be animated to represent molecular processes and chemical reactions, a feat that is not easy to reproduce physically.

History

Initially the rendering was on early Cathode ray tube screens or through plotters drawing on paper. Molecular structures have always been an attractive choice for developing new computer graphics tools, since the input data are easy to create and the results are usually highly appealing. The first example of MG was a display of a protein molecule (Project MAC, 1966) by Cyrus Levinthal and Robert Langridge. Among the milestones in high-performance MG was the work of Nelson Max in "realistic" rendering of macromolecules using reflecting spheres.

By about 1980 many laboratories both in academia and industry had recognized the power of the computer to analyse and predict the properties of molecules, especially in materials science and the pharmaceutical industry. The discipline was often called "molecular graphics" and in 1982 a group of academics and industrialists in the UK set up the Molecular Graphics Society (MGS). Initially much of the technology concentrated

either on high-performance 3D graphics, including interactive rotation or 3D rendering of atoms as spheres (sometimes with radiosity). During the 1980s a number of programs for calculating molecular properties (such as molecular dynamics and quantum mechanics) became available and the term "molecular graphics" often included these. As a result the MGS has now changed its name to the Molecular Graphics and Modelling Society (MGMS).

The requirements of macromolecular crystallography also drove MG because the traditional techniques of physical model-building could not scale. Alwyn Jones' FRODO program (and later "O") were developed to overlay the molecular electron density determined from X-ray crystallography and the hypothetical molecular structure.

Art, science and technology in molecular graphics

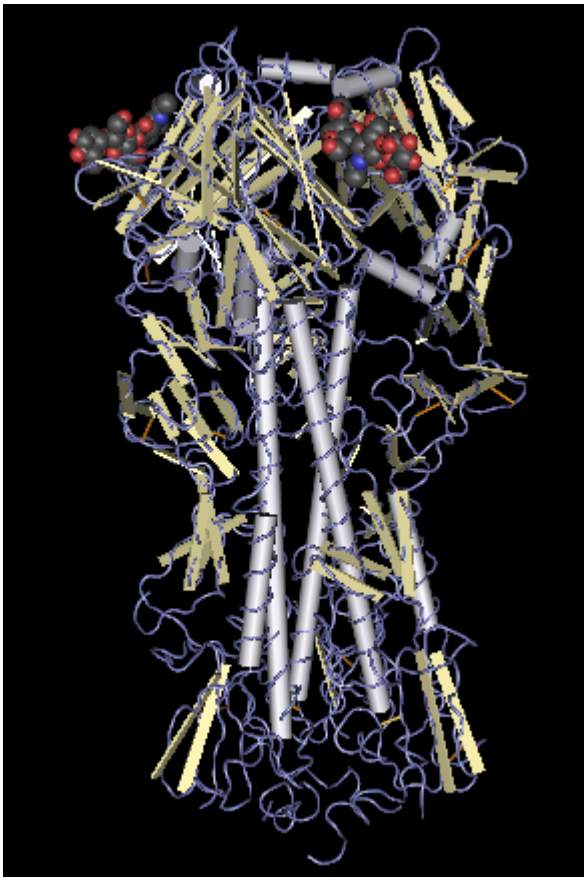


Fig. 2. Image of hemagglutinin with alpha helices depicted as cylinders and the rest of the chain as silver coils. The individual protein atoms (several thousand) have been hidden. All of the non-hydrogen atoms in the two ligands (presumably sialic acid) have been shown near the top of the diagram. Key: Carbon = grey, oxygen = red, nitrogen = blue.

Both computer technology and graphic arts have contributed to molecular graphics. The development of structural biology in the 1950s led to a requirement to display molecules with thousands of atoms. The existing computer technology was limited in power, and in

any case a naive depiction of all atoms left viewers overwhelmed. Most systems therefore used conventions where information was implicit or stylistic. Two vectors meeting at a point implied an atom or (in macromolecules) a complete residue (10-20 atoms).

The macromolecular approach was popularized by Dickerson and Geis' presentation of proteins and the graphic work of Jane Richardson through high-quality hand-drawn diagrams such as the "ribbon" representation. In this they strove to capture the intrinsic 'meaning' of the molecule. This search for the "messages in the molecule" has always accompanied the increasing power of computer graphics processing. Typically the depiction would concentrate on specific areas of the molecule (such as the active site) and this might have different colours or more detail in the number of explicit atoms or the type of depiction (e.g., spheres for atoms).

In some cases the limitations of technology have led to serendipitous methods for rendering. Most early graphics devices used vector graphics, which meant that rendering spheres and surfaces was impossible. Michael Connolly's program "MS" calculated points on the surface-accessible surface of a molecule, and the points were rendered as dots with good visibility using the new vector graphics technology, such as the Evans and Sutherland PS300 series. Thin sections ("slabs") through the structural display showed very clearly the complementarity of the surfaces for molecules binding to active sites, and the "Connolly surface" became a universal metaphor.

The relationship between the art and science of molecular graphics is shown in the exhibitions sponsored by the Molecular Graphics Society. Some exhibits are created with molecular graphics programs alone, while others are collages, or involve physical materials. An example from Mike Hann (1994), inspired by Magritte's painting *Ceci n'est pas une pipe*, uses an image of a salmeterol molecule.

"*Ceci n'est pas une molecule*," writes Mike Hann, "serves to remind us that all of the graphics images presented here are not molecules, not even pictures of molecules, but pictures of icons which we believe represent some aspects of the molecule's properties."

Space-filling models

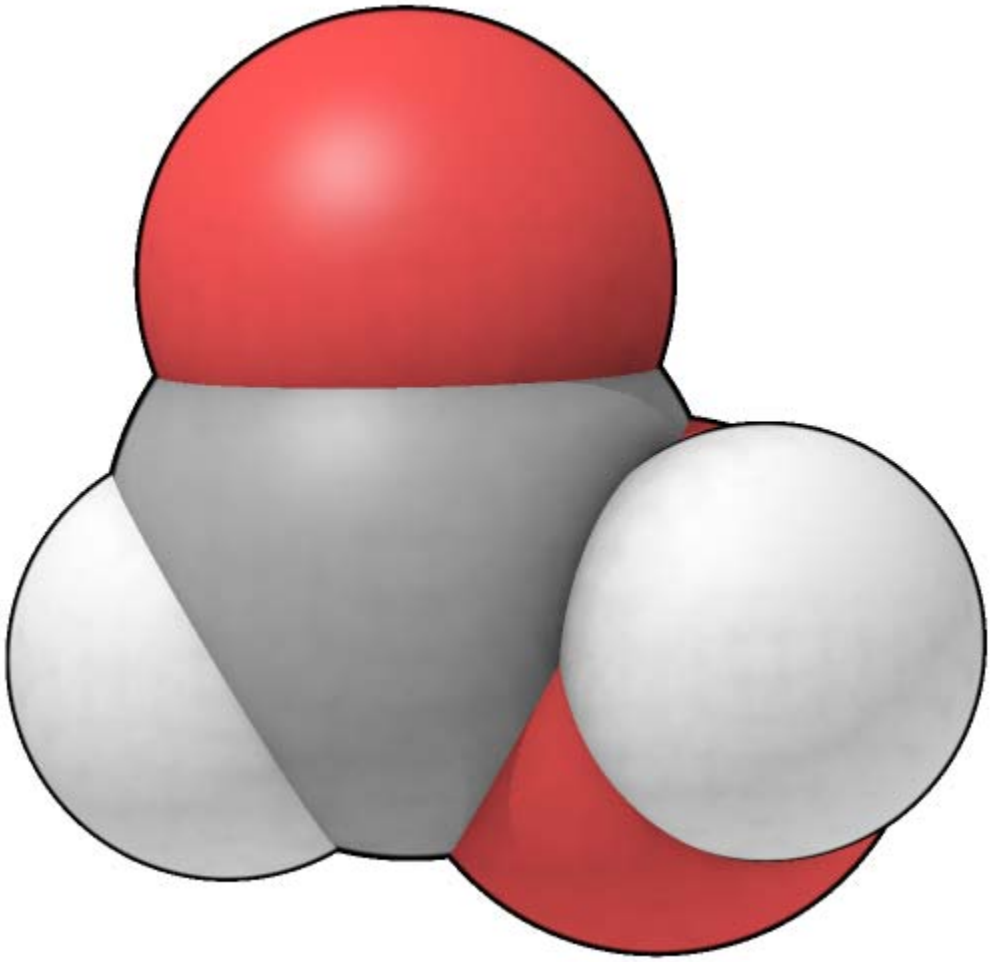


Fig. 4. Space-filling model of formic acid. Key: Hydrogen = white, carbon = black, oxygen = red.

Fig. 4 is a "space-filling" representation of formic acid, where atoms are drawn to suggest the amount of space they occupy. This is necessarily an icon: in the quantum mechanical representation of molecules, there are only (positively charged) nuclei and a "cloud" of negative electrons. The electron cloud defines an approximate size for the molecule, though there can be no single precise definition of size. For many years the size of atoms has been approximated by mechanical models (CPK), where the atoms have been represented by plastic spheres whose radius (van der Waals radius) describes a sphere within which "most" of the electron density can be found. These spheres could be clicked together to show the steric aspects of the molecule rather than the positions of the nuclei. Fig. 4 shows the intricacy required to make sure that all spheres intersect correctly, and also demonstrates a reflective model.

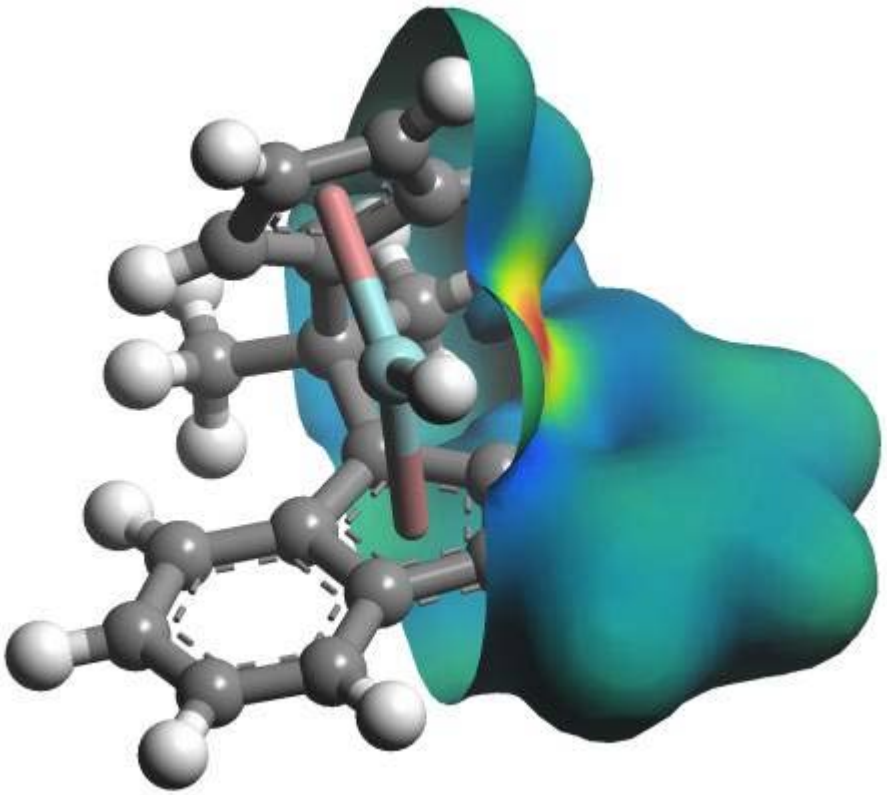


Fig. 5. A molecule (zirconocene) where part (left) is rendered as ball-and-stick and part (right) as an isosurface.

Since the atomic radii (e.g. in Fig. 4) are only slightly less than the distance between bonded atoms, the iconic spheres intersect, and in the CPK models, this was achieved by planar truncations along the bonding directions, the section being circular. When raster graphics became affordable, one of the common approaches was to replicate CPK models *in silico*. It is relatively straightforward to calculate the circles of intersection, but more complex to represent a model with hidden surface removal. A useful side product is that a conventional value for the molecular volume can be calculated.

The use of spheres is often for convenience, being limited both by graphics libraries and the additional effort required to compute complete electronic density or other space-filling quantities. It is now relatively common to see images of surfaces that have been colored to show quantities such as electrostatic potential. Common surfaces in molecular visualization include solvent-accessible ("Lee-Richards") surfaces, solvent-excluded ("Connolly") surfaces, and isosurfaces. The isosurface in Fig. 5 appears to show the electrostatic potential, with blue colors being negative and red/yellow (near the metal) positive (there is no absolute convention of coloring, and red/positive, blue/negative are often reversed). Opaque isosurfaces do not allow the atoms to be seen and identified and it is not easy to deduce them. Because of this, isosurfaces are often drawn with a degree of transparency.

Technology

Early interactive molecular computer graphics systems were vector graphics machines, which used stroke-writing vector monitors, sometimes even oscilloscopes. The electron beam does not sweep left-and-right as in a raster display. The display hardware followed a sequential list of digital drawing instructions (the display list), directly drawing at an angle one stroke for each molecular bond. When the list was complete, drawing would begin again from the top of the list, so if the list was long (a large number of molecular bonds), the display would flicker heavily. Later vector displays could rotate complex structures with smooth motion, since the orientation of all of the coordinates in the display list could be changed by loading just a few numbers into rotation registers in the display unit, and the display unit would multiply all coordinates in the display list by the contents of these registers as the picture was drawn.

The early black-and white vector displays could somewhat distinguish for example a molecular structure from its surrounding electron density map for crystallographic structure solution work by drawing the molecule brighter than the map. Color display makes them easier to tell apart. During the 1970s two-color stroke-writing Penetron tubes were available, but not used in molecular computer graphics systems. In about 1980 Evans & Sutherland made the first practical full-color vector displays for molecular graphics, typically attached to an E&S PS-300 display. This early color tube was expensive, because its was originally engineered to withstand the shaking of a flight-simulator motion base.

Color raster graphics display of molecular models began around 1978 as seen in this paper by Porter on spherical shading of atomic models. Early raster molecular graphics systems displayed static images that could take around a minute to generate. Dynamically rotating color raster molecular display phased in during 1982-1985 with the introduction of the Ikonas programmable raster display.

Molecular graphics has always pushed the limits of display technology, and has seen a number of cycles of integration and separation of compute-host and display. Early systems like Project MAC were bespoke and unique, but in the 1970s the MMS-X and similar systems used (relatively) low-cost terminals, such as the Tektronix 4014 series, often over dial-up lines to multi-user hosts. The devices could only display static pictures but were able to evangelize MG. In the late 1970s, it was possible for departments (such as crystallography) to afford their own hosts (e.g., PDP-11) and to attach a display (such as Evans & Sutherland's MPS) directly to the bus. The display list was kept on the host, and interactivity was good since updates were rapidly reflected in the display—at the cost of reducing most machines to a single-user system.

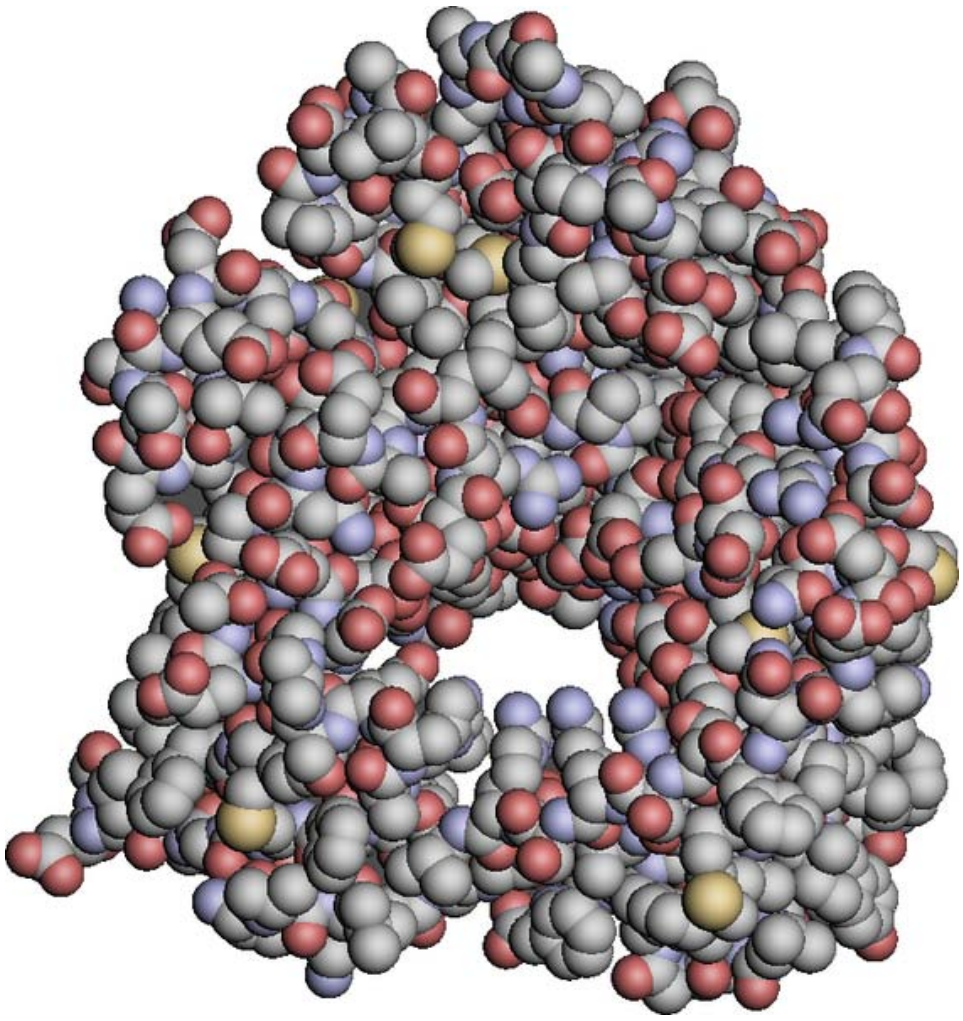
In the early 1980s, Evans & Sutherland (E&S) decoupled their PS300 display, which contained its own display information transformable through a dataflow architecture. Complex graphical objects could be downloaded over a serial line (e.g. 9600 baud) and then manipulated without impact on the host. The architecture was excellent for high performance display but very inconvenient for domain-specific calculations, such as

electron-density fitting and energy calculations. Many crystallographers and modellers spent arduous months trying to fit such activities into this architecture.

The benefits for MG were considerable, but by the later 1980s, UNIX workstations such as Sun-3 with raster graphics (initially at a resolution of 256 by 256) had started to appear. Computer-assisted drug design in particular required raster graphics for the display of computed properties such as atomic charge and electrostatic potential. Although E&S had a high-end range of raster graphics (primarily aimed at the aerospace industry) they failed to respond to the low-end market challenge where single users, rather than engineering departments, bought workstations. As a result the market for MG displays passed to Silicon Graphics, coupled with the development of minisupercomputers (e.g., CONVEX and Alliant) which were affordable for well-supported MG laboratories. Silicon Graphics provided a graphics language, IrisGL, which was easier to use and more productive than the PS300 architecture. Commercial companies (e.g., Biosym, Polygen/MSI) ported their code to Silicon Graphics, and by the early 1990s, this was the "industry standard". Dial boxes were often used as control devices.

Stereoscopic displays were developed based on liquid crystal polarized spectacles, and while this had been very expensive on the PS300, it now became a commodity item. A common alternative was to add a polarizable screen to the front of the display and to provide viewers with extremely cheap spectacles with orthogonal polarization for separate eyes. With projectors such as Barco, it was possible to project stereoscopic display onto special silvered screens and supply an audience of hundreds with spectacles. In this way molecular graphics became universally known within large sectors of chemical and biochemical science, especially in the pharmaceutical industry. Because the backgrounds of many displays were black by default, it was common for modelling sessions and lectures to be held with almost all lighting turned off.

In the last decade almost all of this technology has become commoditized. IrisGL evolved to OpenGL so that molecular graphics can be run on any machine. In 1992, Roger Sayle released his RasMol program into the public domain. RasMol contained a very high-performance molecular renderer that ran on Unix/X Window, and Sayle later ported this to the Windows and Macintosh platforms. The Richardsons developed kinemages and the Mage software, which was also multi-platform. By specifying the chemical MIME type, molecular models could be served over the Internet, so that for the first time MG could be distributed at zero cost regardless of platform. In 1995, Birkbeck College's crystallography department used this to run "Principles of Protein Structure", the first multimedia course on the Internet, which reached 100 to 200 scientists.



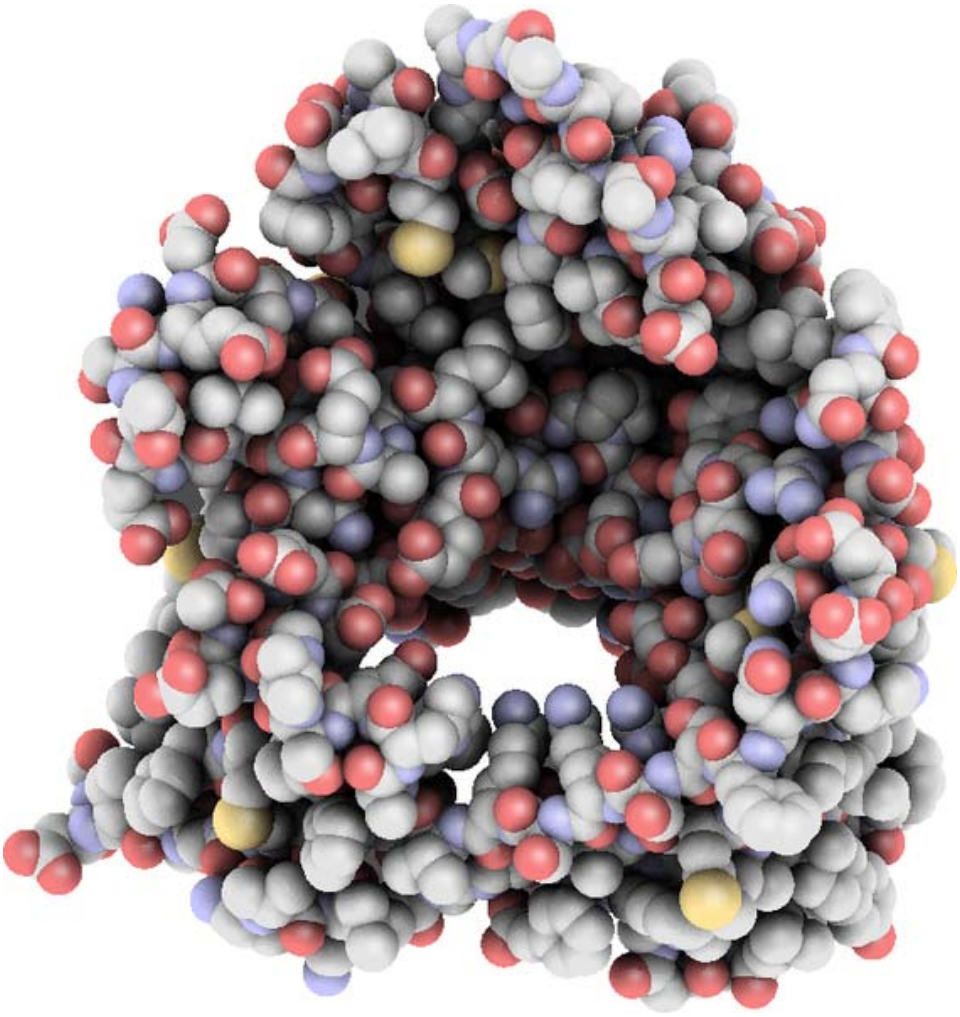


Fig. 6. A molecule of Porin (protein) shown without ambient occlusion (above) and with (below). Advanced rendering effects can improve the comprehension of the 3D shape of a molecule.

MG continues to see innovation that balances technology and art, and currently zero-cost or open source programs such as PyMOL and Jmol have very wide use and acceptance.

Recently the wide spread diffusion of advanced graphics hardware, has improved the rendering capabilities of the visualization tools. The capabilities of current shading languages allow the inclusion of advanced graphic effects (like ambient occlusion, cast shadows and non-photorealistic rendering techniques) in the interactive visualization of molecules. These graphic effects, beside being eye candy, can improve the comprehension of the three dimensional shapes of the molecules. An example of the effects that can be achieved exploiting recent graphics hardware can be seen in the simple open source visualization system QuteMol.

Algorithms

Reference frames

Drawing molecules requires a transformation between molecular coordinates (usually, but not always, in Angstrom units) and the screen. Because many molecules are chiral it is essential that the handedness of the system (almost always right-handed) is preserved. In molecular graphics the origin (0, 0) is usually at the lower left, while in many computer systems the origin is at top left. If the z-coordinate is out of the screen (towards the viewer) the molecule will be referred to right-handed axes, while the screen display will be left-handed.

Molecular transformations normally require:

- scaling of the display (but not the molecule).
- translations of the molecule and objects on the screen.
- rotations about points and lines.

Conformational changes (e.g. rotations about bonds) require rotation of one part of the molecule relative to another. The programmer must decide whether a transformation on the screen reflects a change of view or a change in the molecule or its reference frame.

Simple

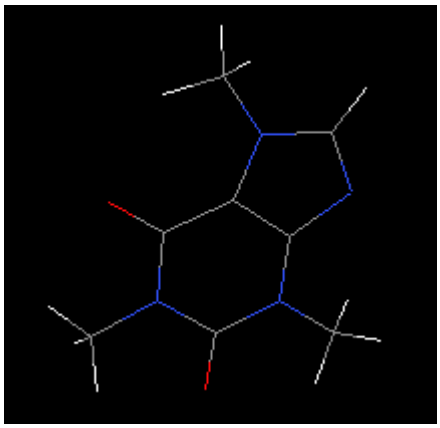


Fig. 7. Stick model of caffeine drawn in Jmol.

In early displays only vectors could be drawn e.g. (Fig. 7) which are easy to draw because no rendering or hidden surface removal is required.

On vector machines the lines would be smooth but on raster devices Bresenham's algorithm is used (note the "jaggies" on some of the bonds, which can be largely removed with antialiasing software.)

Atoms can be drawn as circles, but these should be sorted so that those with the largest z-coordinates (nearest the screen) are drawn last. Although imperfect, this often gives a reasonably attractive display. Other simple tricks which do not include hidden surface algorithms are:

- colouring each end of a bond with the same colour as the atom to which it is attached (Fig. 7).
- drawing less than the whole length of the bond (e.g. 10%-90%) to simulate the bond sticking out of a circle.
- adding a small offset white circle within the circle for an atom to simulate reflection.

Typical pseudocode for creating Fig. 7 (to fit the molecule exactly to the screen):

```
// assume:
// atoms with x, y, z coordinates (Angstrom) and elementSymbol
// bonds with pointers/references to atoms at ends
// table of colours for elementTypes
// find limits of molecule in molecule coordinates as xMin, yMin, xMax,
yMax
scale = min(xScreenMax/(xMax-xMin), yScreenMax/(yMax-yMin))
xOffset = -xMin * scale; yOffset = -yMin * scale
for (bond in $bonds) {
  atom0 = bond.getAtom(0)
  atom1 = bond.getAtom(1)
  x0 = xOffset+atom0.getX()*scale; y0 = yOffset+atom0.getY()*scale //
(1)
  x1 = xOffset+atom1.getX()*scale; y1 = yOffset+atom1.getY()*scale //
(2)
  x1 = atom1.getX();  y1 = atom1.getY()
  xMid = (x0 + x1) /2;  yMid = (y0 + y1) /2;
  colour0 = ColourTable.getColour(atom0.getSymbol())
  drawLine (colour0, x0, y0, xMid, yMid)
  colour1 = ColourTable.getColour(atom1.getSymbol())
  drawLine (colour1, x1, y1, xMid, yMid)
}
```

Note that this assumes the origin is in the bottom left corner of the screen, with Y up the screen. Many graphics systems have the origin at the top left, with Y down the screen. In this case the lines (1) and (2) should have the y coordinate generation as:

```
 y0 = yScreenMax -(yOffset+atom0.getY()*scale) // (1)
 y1 = yScreenMax -(yOffset+atom1.getY()*scale) // (2)
```

Changes of this sort change the handedness of the axes so it is easy to reverse the chirality of the displayed molecule unless care is taken.

Advanced

For greater realism and better comprehension of the 3D structure of a molecule many computer graphics algorithms can be used. For many years molecular graphics has

stressed the capabilities of graphics hardware and has required hardware-specific approaches. With the increasing power of machines on the desktop, portability is more important and programs such as Jmol have advanced algorithms that do not rely on hardware. On the other hand recent graphics hardware is able to interactively render very complex molecule shapes with a quality that would not be possible with standard software techniques.

Chronology

Formative Era: 1960-1976

The formative years of molecular graphics can be summarized by

- Hard-copy molecular graphic in heavy, routine use by chemists (ORTEP, PLUTO, etc.);
- Interactive molecular graphics prototype systems producing little or no publishable chemistry.

The problems for interactive molecular graphics were

- The small address space of laboratory computers (12-bit and 16-bit) limited the size of programs and data that could be easily handled to below the size of most real problems;
- Interactive computer graphics displays were only just fast enough to display simple molecules, below the size of most real problems.

Developer(s)	Approximate date	Technology	Comments
Crystallographers	< 1960	Hand-drawn	Crystal structures, with hidden atom and bond removal. Often clinographic projections.
Johnson, Motherwell	ca 1970	Pen plotter	ORTEP, PLUTO. Very widely deployed for publishing crystal structures.
Cyrus Levinthal, Bob Langridge, Ward, Stots	1966	Project MAC display system, two-degree of freedom, spring-return velocity joystick for rotating the image.	First protein display on screen. System for interactively building protein structures.
Barry	1969	Link 300 computer with a dual trace oscilloscope display.	Interactive molecular structure viewing system. Early examples of dynamic rotation, intensity depth·cueing, and side-by-

			side stereo. Early use of the small angle approximations (a = sin a, 1 = cos a) to speed up graphical rotation calculations.
Ortony	1971	Designed a stereo viewer (British patent appl. 13844/70) for molecular computer graphics.	Horizontal two-way (half-silvered) mirror combines images drawn on the upper and lower halves of a CRT. Crossed polarizers isolate the images to each eye.
Ortony	1971	Light pen, knob.	Interactive molecular structure viewing system. Select bond by turning another knob until desired bond lights up in sequence, a technique later used on the MMS-4 system below, or by picking with the light pen. Points in space are specified with a 3-D ”bug” under dynamic control.
Barry, Graesser, Marshall	1971	CHEMAST: LINK 300 computer driving an oscilloscope. Two-axis joystick, similar to one used later by GRIP-75 (below).	Interactive molecular structure viewing system. Structures dynamically rotated using the joystick.
Tountas and Katz	1971	Adage AGT/50 display	Interactive molecular structure viewing system. Mathematics of nested rotation and for laboratory-space rotation.
Perkins, Piper, Tattam, White	1971	Honeywell DDP 516 computer, EAL TR48 analog computer, Lanelec oscilloscope, 7 linear potentiometers. Stereo.	Interactive molecular structure viewing system.
Wright	1972	GRIP-71 at UNC-CH: IBM System/360 Model 40 time-shared computer, IBM 2250 display, buttons, light pen, keyboard.	Discrete manipulation and energy relaxation of protein structures. Program code became the foundation of the GRIP-75 system below.

Barry and North	1972	Oxford Univ.: Ferranti Argus 500 computer, Ferranti model 30 display, keyboard, track ball, one knob. Stereo.	Prototype large-molecule crystallographic structure solution system. Track ball rotates a bond, knob brightens the molecule vs. electron density map.
North, Ford, Watson	Early 1970s	Leeds Univ.: DEC PDP·11/40 computer, Hewlett-Packard display. 16 knobs, keyboard, spring-return joystick. Stereo.	Prototype large-molecule crystallographic structure solution system. Six knobs rotate and translate a small molecule.
Barry, Bosshard, Ellis, Marshall, Fritch, Jacobi	1974	MMS-4: Washington Univ. at St. Louis, Link 300 computer and an LDS-1 / Link 300 display, custom display modules. Rotation joystick, knobs. Stereo.	Prototype large-molecule crystallographic structure solution system. Select bond to rotate by turning another knob until desired bond lights up in sequence.
Cohen and Feldmann	1974	DEC PDP-10 computer, Adage display, push buttons, keyboard, knobs	Prototype large-molecule crystallographic structure solution system.
Stellman	1975	Princeton: PDP-10 computer, LDS-1 display, knobs	Prototype large-molecule crystallographic structure solution system. Electron density map not shown; instead an "H Factor" figure of merit is updated as the molecular structure is manipulated.
Collins, Cotton, Hazen, Meyer, Morimoto	1975	CRYSNET , Texas A&M Univ. DEC PDP-11/40 computer, Vector General Series 3 display, knobs, keyboard. Stereo.	Prototype large-molecule crystallographic structure solution system. Variety of viewing modes: rocking, spinning, and several stereo display modes.
Cornelius and Kraut	1976 (approx.)	Univ, of Calif. at San Diego: DEC PDP-11/40 emulator (CalData 135), Evans and Sutherland Picture System display, keyboard, 6 knobs. Stereo.	Prototype large-molecule crystallographic structure solution system.

(Yale Univ.)	1976 (approx.)	PIGS: DEC PDP-11/70 computer, Evans and Sutherland Picture System 2 display, data tablet, knobs.	Prototype large-molecule crystallographic structure solution system. The tablet was used for most interactions.
Feldmann and Porter	1976	NIH: DEC PDP—11/70 computer. Evans and Sutherland Picture System 2 display, knobs. Stereo.	Interactive molecular structure viewing system. Intended to display interactively molecular data from the AMSOM – Atlas of Macromolecular Structure on Microfiche.
Rosenberger et al.	1976	MMS-X: Washington Univ. at St. Louis, TI 980B computer, Hewlett-Packard 1321A display, Beehive video terminal, custom display modules, pair of 3-D spring-return joysticks, knobs.	Prototype large-molecule crystallographic structure solution system. Successor to the MMS-4 system above. The 3-D spring-return joysticks either translate and rotate the molecular structure for viewing or a molecular substructure for fitting, mode controlled by a toggle switch.

A few paragraphs each on many of the display and crystallographic systems above are in a review chapter in Lipscomb.

MMS-4 and MMS-X

In a noteworthy attempt to overcome the low speed of graphics displays of the time took place at Washington University in St. Louis. Dave Barry's group attempted to leapfrog the state of the art in graphics displays by making custom display hardware to display images complex enough for large-molecule crystallographic structure solution, fitting molecules to their electron-density maps. The MMS-4 (table above) display modules were slow and expensive, so a second generation of modules was produced for the MMS-X (table above) system. The 16-bit computer initially driving the display was a limitation that may have delayed productive crystallography until the years of the "Mature Era" below.

Mature Era: 1977-present

Developer(s)	Approximate date	Technology	Comments
Britton, Lipscomb, Pique, Wright, Brooks	1977	GRIP-75 at UNC-CH: Time-shared IBM System/360 Model 75 computer, DEC PDP 11/45 computer, Vector	First large-molecule crystallographic structure solution.

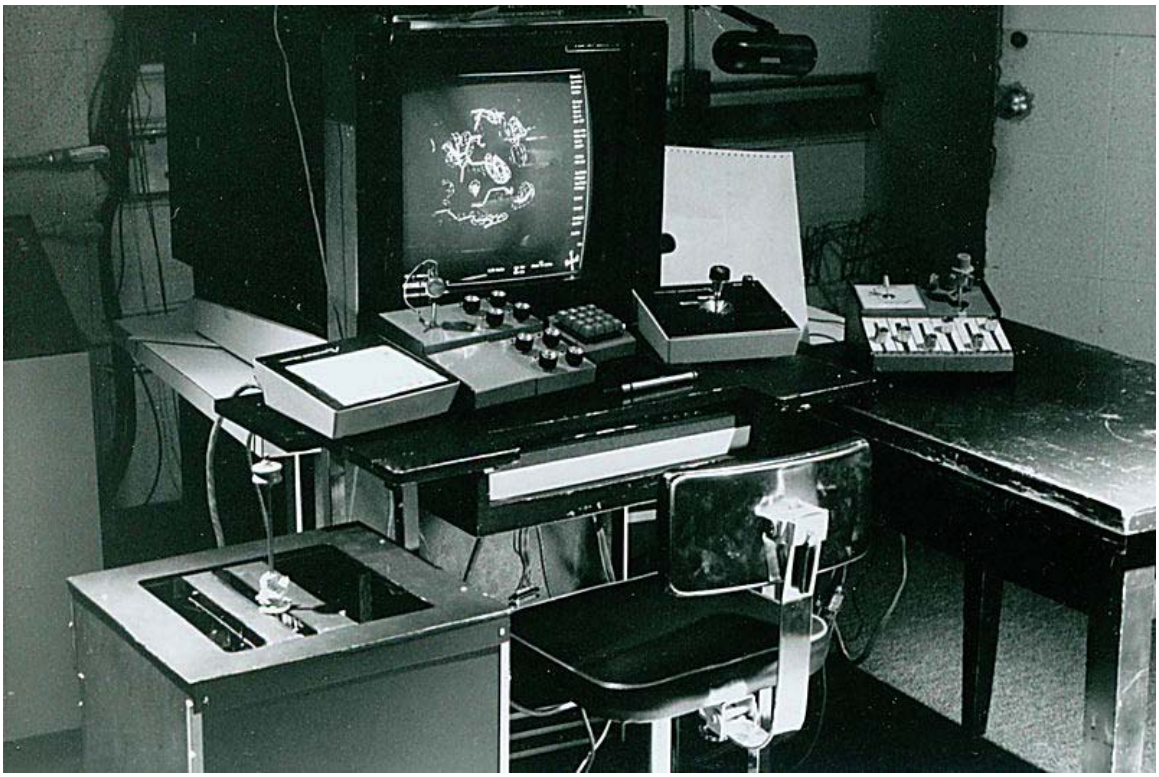
		General Series 3 display, 3-D movement box from A.M. Noll and 3-D spring return joystick for substructure manipulation, Measurement Systems nested joystick, knobs, sliders, buttons, keyboard, light pen.	
Jones	1978	FRODO and RING Max Plank Inst., Germany, RING: DEC PDP-11/40 and Siemens 4004 computers, Vector General 3404 display, 6 knobs.	Large-molecule crystallographic structure solution. FRODO may have run on a DEC VAX-780 as a follow-on to RING.
Diamond	1978	Bilder Cambridge, England, DEC PDP-11/50 computer, Evans and Sutherland Picture System display, tablet.	Large-molecule crystallographic structure solution. All input is by data tablet. Molecular structures built on-line with ideal geometry. Later passes stretch bonds with idealization.
Langridge, White, Marshall	Late 1970s	Departmental systems (PDP-11, Tektronix displays or DEC-VT11, e.g. MMS-X)	Mixture of commodity computing with early displays.
Davies, Hubbard	Mid-1980s	CHEM-X, HYDRA	Laboratory systems with multicolor, raster and vector devices (Sigmex, PS300).
Biosym, Tripos, Polygen	Mid-1980s	PS300 and lower cost dumb terminals (VT200, SIGMEX)	Commercial integrated modelling and display packages.
Silicon Graphics, Sun	Late 1980s	IRIS GL (UNIX) workstations	Commodity-priced single-user workstations with stereoscopic display.
EMBL - WHAT IF	1989, 2000	Machine independent	Nearly free, multifunctional, still fully supported, many free servers based on it
Sayle, Richardson	1992, 1993	RasMol, Kinemage	Platform-independent MG.
MDL (van Vliet,	1995–1998	Chime	proprietary C++ ; free

Maffett, Adler, Holt)			browser plugin for Mac (OS9) and PCs
ChemAxon	1998-	MarvinSketch & MarvinView. MarvinSpace (2005)	proprietary Java applet or stand-alone application.
Community efforts	2000-	Jmol, PyMol, Avogadro, PDB	Open-source Java applet or stand-alone application.
NOCH	2002-	NOC	Powerful and open source code molecular structure explorer
LION Bioscience / EMBL	2004-	SRS 3D	Free, open-source system based on Java3D. Integrates 3D structures with sequence and feature data (domains, SNPs, etc.).
San Diego Supercomputer Center	2006-	Sirius	Free for academic/non-profit institutions
Weizmann Institute of Science - Community efforts	2008-	Proteopedia	Collaborative

Productive use of interactive molecular graphics became possible in the Mature Era largely by two hardware advances:

- 24-bit and 32-bit computers became affordable to university departments, increasing the size of programs and data that could be easily handled to the size of many real problems. Early examples include GRIP-75, which used in part a time-shared IBM System/360 Model 75, and FRODO, which used a DEC VAX-780.
- Interactive computer graphics displays became fast enough to easily display complex molecules and electron-density maps. The first such displays widely used were the Vector General Series 3 and the Evans and Sutherland Picture System 2, MultiPicure System, and PS-300

First large-molecule crystallographic computer graphics structure solutions



GRIP-75 molecular graphics system console, Univ. of N. Carolina, Chapel Hill

Before computer graphics could do the job, mechanical methods were used to fit large molecules to their electron density maps. Contour circles around high electron density were drawn on large plastic sheets. Bingo chips were placed on the plastic sheets where atoms were interpreted to be in the case of the solution of carboxypeptidase A. This was superseded by the Richards Box in which an adjustable brass Kendrew molecular model was placed front of a 2-way mirror, behind which were plastic sheets of the electron density map. This optically superimposed the molecular model and the electron density map. The model was moved to within the contour lines of the superimposed map. Then, atomic coordinates were recorded using a plumb bob and a meter stick. These mechanical methods are no longer used.

The first large molecule whose atomic structure was *partly* determined on a molecular computer graphics system was Transfer RNA by Sung-Hou Kim's team in 1976. They used the GRIP-75 computer graphics system after initial fitting on a mechanical Richards Box.

The first large molecule whose atomic structure was *entirely* determined on a molecular computer graphics system was neurotoxin A from venom of the Philippines sea snake, by Tsernoglou, Petsko, and Tu, using GRIP-75, with a statement of being first in 1977.

The Richardson group, using GRIP-75, published partial atomic structure results of the protein superoxide dismutase in 1977.

GRIP-75 was gradually superseded by systems that crystallographers could have in their own labs (FRODO, RING, Builder, MMS-X, etc.). There was only one GRIP-75 system built, and crystallographers had to travel to use it, and after solving about two dozen molecular structures it was retired.

All of these early systems were sometimes called "Electronic Richards Boxes" because they worked likewise. Crystallographers manually moved parts of the molecule (using joysticks and knobs) to within the contour lines of the electronically superimposed map.

Nowadays fitting of the molecular structure to the electron density map is largely automated by algorithms with computer graphics a guide to the process. An example is the XtalView XFit program.

Chapter 4

Molecular Machine

A **molecular machine**, or nanomachine, is defined as a discrete number of molecular components which perform mechanical-like movements (output) in response to specific stimuli (input). More generally, the expression is often applied to molecules that simply mimic functions which occur at the macroscopic level. The term is also common in nanotechnology, and a number of highly complex molecular machines have been proposed towards the goal of constructing a molecular assembler. Molecular machines can be divided into two broad categories: synthetic and biological.

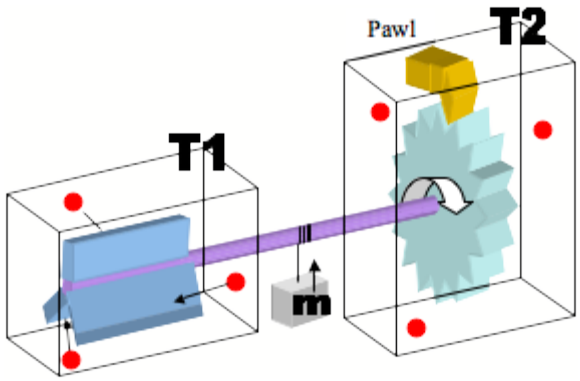
Molecular systems capable of shifting a chemical or mechanical process away from equilibrium represent a potentially important branch of chemistry and nanotechnology. As the gradient generated from this process is able to perform useful work, by definition, these types of systems are examples of molecular machinery.

Historical Insight and Studies

There are two thought experiments that form the historical basis for molecular machines: Maxwell's demon and Feynman's Ratchet (or Brownian ratchet). Maxwell's Demon is well described elsewhere, and a slightly different interpretation of Richard Feynman's ratchet is given here.

Imagine a very small system (seen below) of two paddles or gears connected by a rigid axle and that it is possible to keep these two paddles at two different temperatures. One of the gears (at T2) has a pawl that is rectifying the system motion, and therefore, the axle can only move in a clockwise rotation, and in doing so, it could lift a weight (m) upward upon ratcheting. Now imagine if the paddle in box T1 was in a much hotter environment than the gear in box T2; it would be expected that the kinetic energy of the gas molecules (red circles) hitting the paddle in T1 would be much higher than the gas molecules hitting the gear at T2. Therefore, with lower kinetic energy of the gases in T2, there would be very little resistance from the molecules on colliding with the gear in the statistically

opposite direction. Further, the ratcheting would allow for directionality, and slowly over time, the axle would rotate and ratchet, lifting the weight (m).



Schematic figure of Feynman's Ratchet

As described, this system may seem like a perpetual motion machine; however, the key ingredient is the heat gradient within the system. This ratchet does not threaten the second law of thermodynamics, because this temperature gradient must be maintained by some external means. Brownian motion of the gas particles provides the power to the machine, and the temperature gradient allows the machine to drive the system cyclically away from equilibrium. In Feynman's ratchet, random Brownian motion is not fought against, but instead, harnessed and rectified. Unfortunately, temperature gradients cannot be maintained over molecular scale distances because of molecular vibration redistributing the energy to other parts of the molecule. Furthermore, despite Feynman's machine doing useful work in lifting the mass, using Brownian motion to power a molecular level machine does not provides any insight on how that power (or potential energy of the lifted weight, m) can be used to perform nanoscale tasks.

Modern Insights and Studies

Unlike macroscopic motion, molecular systems are constantly undergoing significant dynamic motions subject to the laws of Brownian mechanics (or Brownian motion), and as such, harnessing molecular motion is a far more difficult process. At the macroscopic level, many machines operate in the gas phase, and often, air resistance is neglected, as it is insignificant, but analogously for a molecular system in a Brownian environment, molecular motion is similar "to walking in a hurricane, or swimming in molasses." The phenomenon of Brownian motion (observed by Robert Brown (botanist), 1827) was later explained by Albert Einstein in 1905. Einstein found that Brownian motion is a consequence of scale and not the nature of the surroundings. As long as thermal energy is applied to a molecule, it will undergo Brownian motion with the kinetic energy appropriate to that temperature. Therefore, like Feynman's strategy, when designing a molecular machine, it seems sensible to utilize Brownian motion rather than attempt to fight against it.

Like macroscopic machines, molecular machines typically have movable parts. However, while everyday macroscopic machines may provide inspiration for molecular machines, it is misleading to draw analogies between their design strategy; the dynamics of large and small length scales are simply too different. Harnessing Brownian motion and making molecular level machines is regulated by the second law of thermodynamics, with its often counter-intuitive consequences, and as such, we need another inspiration.

Although it is a challenging process to harness Brownian motion, nature has provided us with several blueprints for molecular motion performing useful work. Nature has created many useful structures for compartmentalizing molecular systems, hence creating distinct non-equilibrium distributions; the cell membrane is an excellent example. Lipophilic barriers make use of a number of different mechanisms to power motion from one compartment to another.

Examples of molecular machines

From a synthetic perspective, there are two important types of molecular machines: molecular switches (or shuttles) and molecular motors. The major difference between the two systems is that a switch influences a system as a function of state, whereas a motor influences a system as function of trajectory. A switch (or shuttle) may appear to undergo translational motion, but returning a switch to its original position undoes any mechanical effect and liberates energy to the system. Furthermore, switches cannot use chemical energy to repetitively and progressively drive a system away from equilibrium where a motor can.

Synthetic

A wide variety of rather simple molecular machines have been synthesized by chemists. They can consist of a single molecule; however, they are often constructed for mechanically-interlocked molecular architectures, such as rotaxanes and catenanes.

- Molecular motors are molecules that are capable of unidirectional rotation motion powered by external energy input. A number of molecular machines have been synthesized powered by light or reaction with other molecules.

- A molecular propeller is a molecule that can propel fluids when rotated, due to its special shape that is designed in analogy to macroscopic propellers. It has several molecular-scale blades attached at a certain pitch angle around the circumference of a nanoscale shaft.

- A molecular switch is a molecule that can be reversibly shifted between two or more stable states. The molecules may be shifted between the states in response to changes in e.g. pH, light, temperature, an electrical current, microenvironment, or the presence of a ligand.

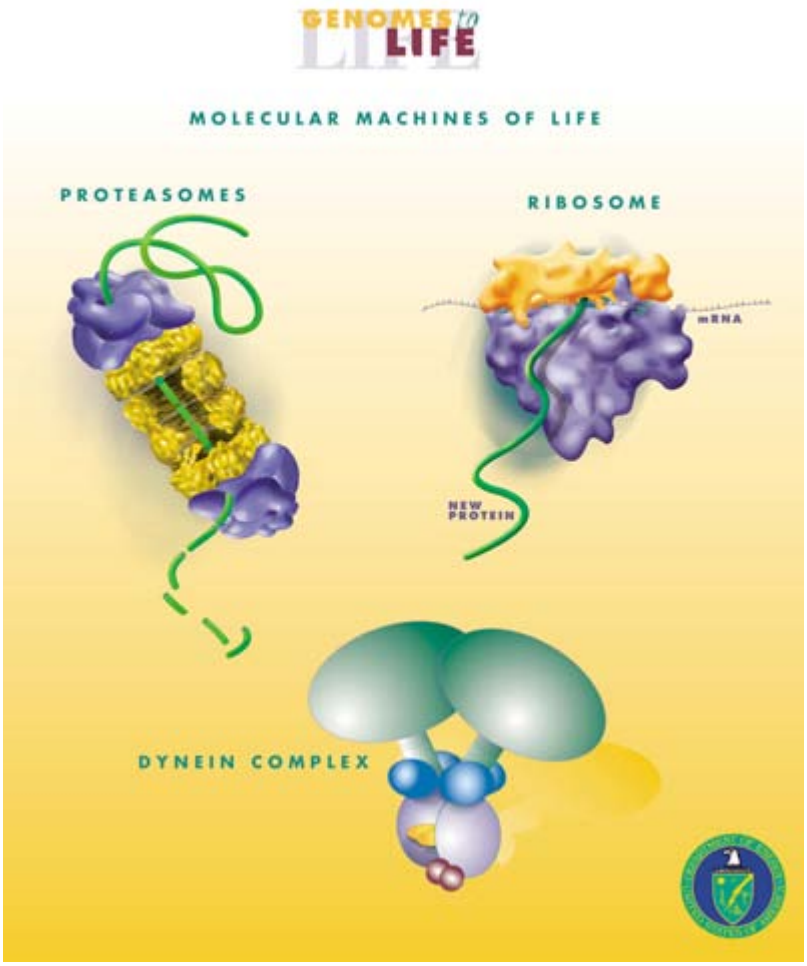
- A molecular shuttle is a molecule capable of shuttling molecules or ions from one location to another. A common molecular shuttle consists of a rotaxane where the macrocycle can move between two sites or stations along the dumbbell backbone.

- Molecular tweezers are host molecules capable of holding items between its two arms. The open cavity of the molecular tweezers binds items using non-covalent bonding including hydrogen bonding, metal coordination, hydrophobic forces, van der Waals forces, π-π interactions, and/or electrostatic effects. Examples of molecular tweezers have been reported that are constructed from DNA and are considered DNA machines.

- A molecular sensor is a molecule that interacts with an analyte to produce a detectable change. Molecular sensors combine molecular recognition with some form of reporter, so the presence of the item can be observed.

- A molecular logic gate is a molecule that performs a logical operation on one or more logic inputs and produces a single logic output. Unlike a molecular sensor, the molecular logic gate will only output when a particular combination of inputs are present.

Biological



Some biological molecular machines

The most complex molecular machines are found within cells. These include motor proteins, such as myosin, which is responsible for muscle contraction, kinesin, which moves cargo inside cells away from the nucleus along microtubules, and dynein, which produces the axonemal beating of motile cilia and flagella. These proteins and their nanoscale dynamics are far more complex than any molecular machines that have yet been artificially constructed.

The detailed mechanism of ciliary motility has been described by Satir in a 2008 review article. A high-level-abstraction summary is that, "[i]n effect, the [motile cilium] is a nanomachine composed of perhaps over 600 proteins in molecular complexes, many of which also function independently as nanomachines."

Theoretical

The construction of more complex molecular machines is an active area of theoretical research. A number of molecules, such as molecular propellers, have been designed, although experimental studies of these molecules are inhibited by the lack of methods to construct these molecules. These complex molecular machines form the basis of areas of nanotechnology, including molecular assembler.

Chapter 5

Molecular Assembler

A **molecular assembler** is a "proposed device able to guide chemical reactions by positioning reactive molecules with atomic precision", as defined by K. Eric Drexler. Some biological molecules such as ribosomes fit this definition, since they receive instructions from messenger RNA and then assemble specific sequences of amino acids to construct protein molecules. However, the term "molecular assembler" usually refers to theoretical human-made devices.

Beginning in 2007, the British Engineering and Physical Sciences Research Council funds development of ribosome-like molecular assemblers. Clearly, molecular assemblers are possible in this limited sense. A technology roadmap project, led by the Battelle Memorial Institute and hosted by several U.S. National Laboratories has explored a range of atomically precise fabrication technologies, including both early-generation and longer-term prospects for programmable molecular assembly; the report was released in December, 2007.

Likewise, the term "molecular assembler" has been used in science fiction and popular culture to refer to a wide range of fantastic atom-manipulating nanomachines, many of which may be physically impossible in reality. Much of the controversy regarding "molecular assemblers" results from the confusion in the use of the name for both technical concepts and popular fantasies. In 1992, Drexler introduced the related but better-understood term "molecular manufacturing," which he defined as the programmed "chemical synthesis of complex structures by mechanically positioning reactive molecules, not by manipulating individual atoms."

Here we, mostly discusses "molecular assemblers" in the popular sense. These include hypothetical machines that manipulate individual atoms and machines with organism-like self-replicating abilities, mobility, ability to consume food, and so forth. These are quite different from devices that merely (as defined above) "guide chemical reactions by positioning reactive molecules with atomic precision".

Because synthetic molecular assemblers have never been constructed and because of the confusion regarding the meaning of the term, there has been much controversy as to whether "molecular assemblers" are possible or simply science fiction. Confusion and controversy also stem from their classification as nanotechnology, which is an active area of laboratory research which has already been applied to the production of real products; however, there had been, until recently, no research efforts into the actual construction of "molecular assemblers". A primary criticism of the computational research into products of advanced "molecular assemblers" is that the structures investigated are impossible to make today.

Nanofactories

A **nanofactory** is a proposed system in which nanomachines (resembling molecular assemblers, or industrial robot arms) would combine reactive molecules via mechanosynthesis to build larger atomically precise parts. These, in turn, would be assembled by positioning mechanisms of assorted sizes to build macroscopic (visible) but still atomically-precise products.

A typical nanofactory would fit in a desktop box, in the vision of K. Eric Drexler published in *Nanosystems: Molecular Machinery, Manufacturing and Computation* (1992), a notable work of "exploratory engineering". During the last decade, others have extended the nanofactory concept, including an analysis of nanofactory convergent assembly by Ralph Merkle, a systems design of a replicating nanofactory architecture by J. Storrs Hall, Forrest Bishop's "Universal Assembler", the patented exponential assembly process by Zyvex, and a top-level systems design for a 'primitive nanofactory' by Chris Phoenix (Director of Research at the Center for Responsible Nanotechnology). All of these nanofactory designs (and more) are summarized in Chapter 4 of *Kinematic Self-Replicating Machines* (2004) by Robert Freitas and Ralph Merkle. The Nanofactory Collaboration, founded by Robert Freitas and Ralph Merkle in 2000, is a focused ongoing effort involving 23 researchers from 10 organizations and 4 countries that is developing a practical research agenda specifically aimed at positionally-controlled diamond mechanosynthesis and diamondoid nanofactory development.

In 2005, a computer-animated short film of the nanofactory concept was produced by John Burch, in collaboration with Drexler. Such visions have been the subject of much debate, on several intellectual levels. No one has discovered an insurmountable problem with the underlying theories and no one has proved that the theories can be translated into practice. However, the debate continues, with some of it being summarized in the Molecular nanotechnology article.

If nanofactories could be built, severe disruption to the world economy would be one of many possible negative impacts, though it could be argued that this disruption would have little negative effect if everyone had such nanofactories. Great benefits also would be anticipated. Various works of science fiction have explored these and similar concepts. The potential for such devices was part of the mandate of a major UK study led by mechanical engineering professor Dame Ann Dowling. The report is now complete.

Self-replication

"Molecular assemblers" have been confused with self-replicating machines. To produce a practical quantity of a desired product, the nanoscale size of a typical science fiction universal molecular assembler requires an extremely large number of such devices. However, a single such theoretical molecular assembler might be programmed to self-replicate, constructing many copies of itself. This would allow an exponential rate of production. Then after sufficient quantities of the molecular assemblers were available, they would then be re-programmed for production of the desired product. However, if self-replication of molecular assemblers were not restrained then it might lead to competition with naturally occurring organisms. This has been called ecophagy or the grey goo problem.

One method to building molecular assemblers is to mimic evolutionary processes employed by biological systems. Biological evolution proceeds by random variation combined with culling of the less-successful variants and reproduction of the more-successful variants. Production of complex molecular assemblers might be evolved from simpler systems since "A complex system that works is invariably found to have evolved from a simple system that worked. . . . A complex system designed from scratch never works and can not be patched up to make it work. You have to start over, beginning with a system that works." However, most published safety guidelines include "recommendations against developing ... replicator designs which permit surviving mutation or undergoing evolution".

Most assembler designs keep the "source code" external to the physical assembler. At each step of a manufacturing process, that step is read from an ordinary computer file and "broadcast" to all the assemblers. If any assembler gets out of range of that computer, or when the link between that computer and the assemblers is broken, or when that computer is unplugged, the assemblers stop replicating. Such a "broadcast architecture" is one of the safety features recommended by the "Foresight Guidelines on Molecular Nanotechnology", and a map of the 137-dimensional replicator design space recently published by Freitas and Merkle provides numerous practical methods by which replicators can be safely controlled by good design.

Drexler and Smalley debate

One of the most outspoken critics of some concepts of "molecular assemblers" was Professor Richard Smalley (1943-2005) who won the Nobel prize for his contributions to the field of nanotechnology. Smalley believed that such assemblers were not physically possible and introduced scientific objections to them. His two principal technical objections were termed the "fat fingers problem" and the "sticky fingers problem". He believed these would exclude the possibility of "molecular assemblers" that worked by precision picking and placing of individual atoms. Drexler and coworkers responded to these two issues in a 2001 publication.

Smalley also believed that Drexler's speculations about apocalyptic dangers of self-replicating machines that have been equated with "molecular assemblers" would threaten the public support for development of nanotechnology. To address the debate between Drexler and Smalley regarding molecular assemblers *Chemical & Engineering News* published a point-counterpoint consisting of an exchange of letters that addressed the issues.

Regulation

Speculation on the power of systems that have been called "molecular assemblers" has sparked a wider political discussion on the implication of nanotechnology. This is in part due to the fact that nanotechnology is a very broad term and could include "molecular assemblers." Discussion of the possible implications of fantastic molecular assemblers has prompted calls for regulation of current and future nanotechnology. There are very real concerns with the potential health and ecological impact of nanotechnology that is being integrated in manufactured products. Greenpeace for instance commissioned a report concerning nanotechnology in which they express concern into the toxicity of nanomaterials that have been introduced in the environment. However, it makes only passing references to "assembler" technology. The UK Royal Society and UK Royal Academy of Engineering also commissioned a report entitled "Nanoscience and nanotechnologies: opportunities and uncertainties" regarding the larger social and ecological implications on nanotechnology. This report does not discuss the threat posed by potential so-called "molecular assemblers."

Formal scientific review

In 2006, U.S. National Academy of Sciences released the report of a study of molecular manufacturing as part of a longer report, *A Matter of Size: Triennial Review of the National Nanotechnology Initiative* The study committee reviewed the technical content of *Nanosystems*, and in its conclusion states that no current theoretical analysis can be considered definitive regarding several questions of potential system performance, and that optimal paths for implementing high-performance systems cannot be predicted with confidence. It recommends experimental research to advance knowledge in this area:

> "Although theoretical calculations can be made today, the eventually attainable range of chemical reaction cycles, error rates, speed of operation, and thermodynamic efficiencies of such bottom-up manufacturing systems cannot be reliably predicted at this time. Thus, the eventually attainable perfection and complexity of manufactured products, while they can be calculated in theory, cannot be predicted with confidence. Finally, the optimum research paths that might lead to systems which greatly exceed the thermodynamic efficiencies and other capabilities of biological systems cannot be reliably predicted at this time. Research funding that is based on the ability of investigators to produce experimental demonstrations that link to abstract models and guide long-term vision is most appropriate to achieve this goal."

Grey goo

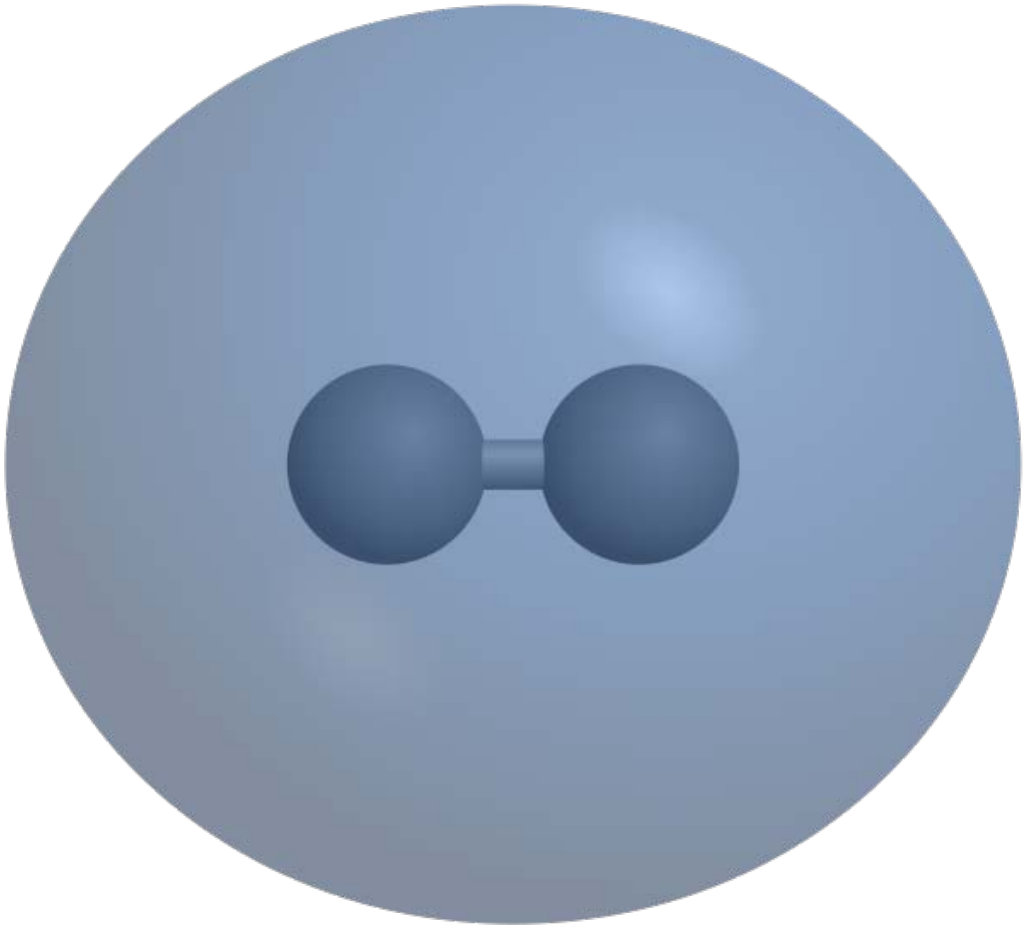
One potential scenario that has been envisioned is out-of-control self-replicating molecular assemblers in the form of grey goo which consumes carbon to continue its replication. If unchecked such mechanical replication could potentially consume whole ecoregions or the whole Earth (ecophagy), or it could simply outcompete natural lifeforms for necessary resources such as carbon, ATP, or UV light (which some nanomotor examples run on). It is worth noting that the ecophagy and 'grey goo' scenarios, like synthetic molecular assemblers, are based upon still-theoretical technologies that have not yet been demonstrated experimentally.

Chapter 6

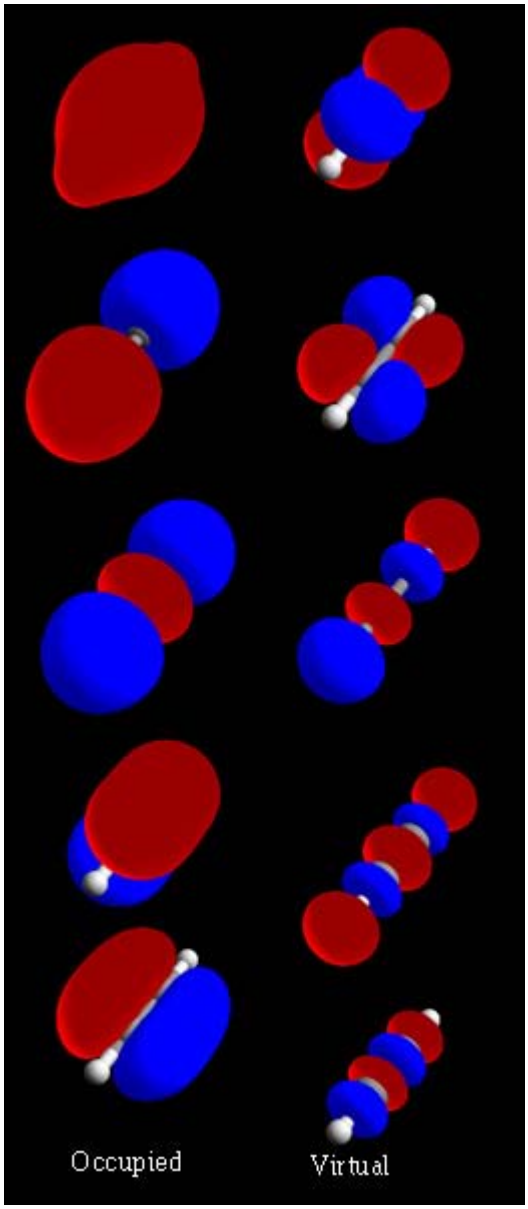
Molecular Orbital and Molecular Orbital Theory

Molecular orbital

In chemistry, a **molecular orbital** (or **MO**) is a mathematical function describing the wave-like behavior of an electron in a molecule. This function can be used to calculate chemical and physical properties such as the probability of finding an electron in any specific region. The term "orbital" was first used in English by Robert S. Mulliken as the English translation of Schrödinger's 'Eigenfunktion'. It has since been equated with the "region" generated with the function. Molecular orbitals are usually constructed by combining atomic orbitals or hybrid orbitals from each atom of the molecule, or other molecular orbitals from groups of atoms. They can be quantitatively calculated using the Hartree-Fock or Self-Consistent Field (SCF) methods.



H_2 1sσ bonding molecular orbital



Complete acetylene (H–C≡C–H) molecular orbital set

Overview

A molecular orbital (MO) can specify the electron configuration of a molecule: the spatial distribution and energy of one (or one pair of) electron(s). Most commonly an MO is represented as a linear combination of atomic orbitals (the LCAO-MO method), especially in qualitative or very approximate usage. They are invaluable in providing a simple model of bonding in molecules, understood through molecular orbital theory.

Most present-day methods in computational chemistry begin by calculating the MOs of the system. A molecular orbital describes the behavior of one electron in the electric field generated by the nuclei and some average distribution of the other electrons. In the case

of two electrons occupying the same orbital, the Pauli principle demands that they have opposite spin. Necessarily this is an approximation, and highly accurate descriptions of the molecular electronic wave function do not have orbitals.

Qualitative discussion

For an imprecise, but qualitatively useful, discussion of the molecular structure, the molecular orbitals can be obtained from the "Linear combination of atomic orbitals molecular orbital method" ansatz. Here, the molecular orbitals are expressed as linear combinations of atomic orbitals.

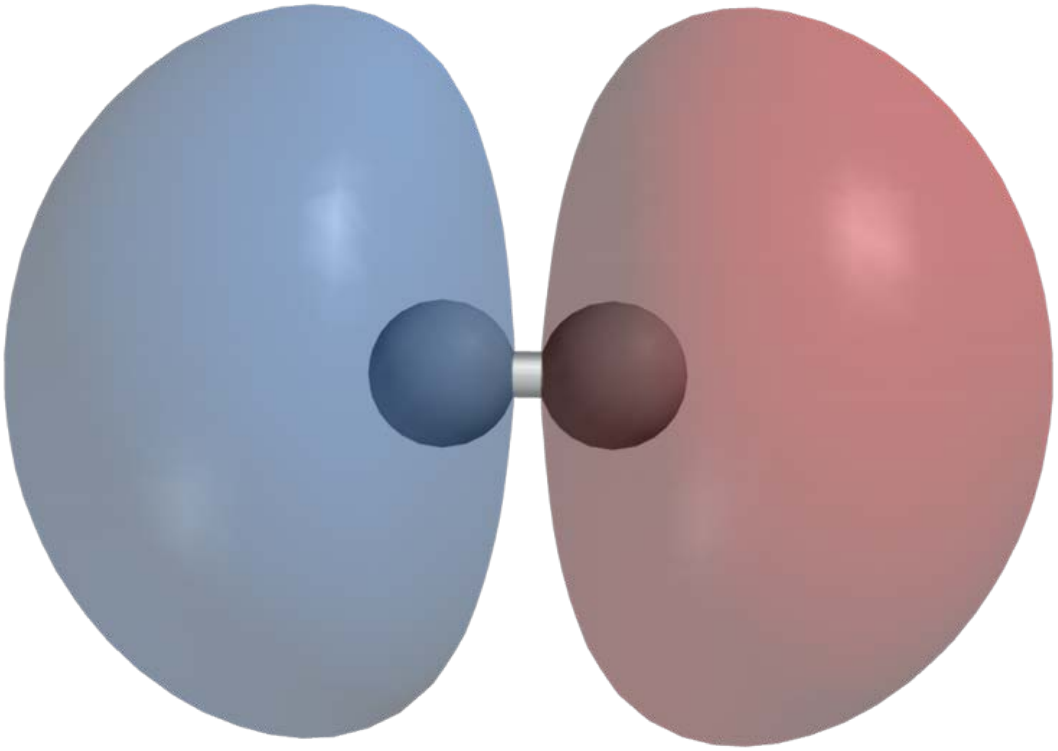
Molecular orbitals were first introduced by Friedrich Hund and Robert S. Mulliken in 1927 and 1928. The linear combination of atomic orbitals or "LCAO" approximation for molecular orbitals was introduced in 1929 by Sir John Lennard-Jones.. His ground-breaking paper showed how to derive the electronic structure of the fluorine and oxygen molecules from quantum principles. This qualitative approach to molecular orbital theory is part of the start of modern quantum chemistry.

Some properties:

- The number of molecular orbitals is equal to the number of atomic orbitals included in the linear expansion,
- If the molecule has some symmetry, the degenerate atomic orbitals (with the same atomic energy) are grouped in linear combinations (called **symmetry adapted atomic orbitals (SO)**) which belong to the representation of the symmetry group, so the wave functions that describe the group are known as **symmetry-adapted linear combinations** (**SALC**).
- The number of molecular orbitals belonging to one group representation is equal to the number of symmetry-adapted atomic orbitals belonging to this representation,
- Within a particular representation, the symmetry-adapted atomic orbitals mix more if their atomic energy levels are closer.

Examples

H_2



H_2 1sσ* antibonding molecular orbital

As a simple example consider the hydrogen molecule, H_2, with the two atoms labelled H' and H". The lowest-energy atomic orbitals, 1s' and 1s", do not transform according to the symmetries of the molecule. However, the following symmetry adapted atomic orbitals do:

1s' - 1s"	Antisymmetric combination: negated by reflection, unchanged by other operations
1s' + 1s"	Symmetric combination: unchanged by all symmetry operations

The symmetric combination (called a bonding orbital) is lower in energy than the basis orbitals, and the antisymmetric combination (called an antibonding orbital) is higher. Because the H_2 molecule has two electrons, they can both go in the bonding orbital, making the system lower in energy (and hence more stable) than two free hydrogen atoms. This is called a covalent bond. The *bond order* is equal to the number of bonding electrons minus the number of antibonding electrons, divided by 2. In this example there

are 2 electrons in the bonding orbital and none in the antibonding orbital; the bond order is 1, and there is a single bond between the two hydrogen atoms.

He_2

On the other hand, consider the hypothetical molecule of He_2 with the atoms labelled He' and He". Again, the lowest-energy atomic orbitals, 1s' and 1s", do not transform according to the symmetries of the molecule, while the following symmetry adapted atomic orbitals do:

1s' - 1s"	Antisymmetric combination: negated by reflection, unchanged by other operations
1s' + 1s"	Symmetric combination: unchanged by all symmetry operations

Similar to the molecule H_2, the symmetric combination (called a bonding orbital) is lower in energy than the basis orbitals, and the antisymmetric combination (called an antibonding orbital) is higher. However, in its neutral ground state, each helium atom contains two electrons in its 1s orbital, combining for a total of four electrons. Two electrons fill the lower energy bonding orbital, while the remaining two fill the higher energy antibonding orbital. Thus, the resulting electron density around the molecule does not support the formation of a bond between the two atoms (called a sigma bond); therefore, the molecule does not exist. Another way of looking at it is that there are two bonding electrons and two antibonding electrons; therefore, the bond order is 0 and no bond exists.

Noble gases

Considering a hypothetical molecule of He_2, since the basis set of atomic orbitals is the same as in the case of H_2, we find that both the bonding and antibonding orbitals are filled, so there is no energy advantage to the pair. HeH would have a slight energy advantage, but not as much as H_2 + 2 He, so the molecule exists only a short while. In general, we find that atoms such as He that have completely full energy shells rarely bond with other atoms. Except for short-lived Van der Waals complexes, there are very few noble gas compounds known.

Ionic bonds

When the energy difference between the atomic orbitals of two atoms is quite large, one atom's orbitals contribute almost entirely to the bonding orbitals, and the other's almost entirely to the antibonding orbitals. Thus, the situation is effectively that some electrons have been transferred from one atom to the other. This is called an (mostly) ionic bond.

MO diagrams

For more complicated molecules, the wave mechanics approach loses utility in a qualitative understanding of bonding (although is still necessary for a quantitative

approach). The qualitative approach of MO uses a molecular orbital diagram. In this type of diagram, the molecular orbitals are represented by horizontal lines; the higher a line, the higher the energy of the orbital, and degenerate orbitals are placed on the same level with a space between them. Then, the electrons to be placed in the molecular orbitals are slotted in one by one, keeping in mind the Pauli exclusion principle and Hund's rule of maximum multiplicity (only 2 electrons, having opposite spins, per orbital; have as many unpaired electrons on one energy level as possible before starting to pair them).

HOMO and LUMO

The highest occupied molecular orbital and lowest unoccupied molecular orbital are often referred to as the HOMO and LUMO, respectively. The difference of the energies of the HOMO and LUMO, termed the band gap, can sometimes serve as a measure of the excitability of the molecule: the smaller the energy, the more easily it will be excited.

More quantitative approach

To obtain quantitative values for the molecular energy levels, one needs to have molecular orbitals which are such that the configuration interaction (CI) expansion converges fast towards the full CI limit. The most common method to obtain such functions is the Hartree–Fock method which expresses the molecular orbitals as eigenfunctions of the Fock operator. One usually solves this problem by expanding the molecular orbitals as linear combinations of gaussian functions centered on the atomic nuclei. The equation for the coefficients of these linear combinations is a generalized eigenvalue equation known as the Roothaan equations which are in fact a particular representation of the Hartree-Fock equation.

Simple accounts often suggest that experimental molecular orbital energies can be obtained by the methods of ultra-violet photoelectron spectroscopy for valence orbitals and X-ray photoelectron spectroscopy for core orbitals. This however is incorrect as these experiments measure the ionization energy, the difference in energy between the molecule and one of the ions resulting from the removal of one electron. Ionization energies are linked approximately to orbital energies by Koopmans' theorem. While the agreement between these two values can be close for some molecules, it can be very poor in other cases.

Molecular orbital theory

In chemistry, **molecular orbital (MO) theory** is a method for determining molecular structure in which electrons are not assigned to individual bonds between atoms, but are treated as moving under the influence of the nuclei in the whole molecule. In this theory, each molecule has a set of molecular orbitals, in which it is assumed that the molecular orbital wave function ψ_f may be written as a simple weighted sum of the n constituent atomic orbitals χ_i, according to the following equation:

$$\psi_j = \sum_{i=1}^{n} c_{ij} \chi_i$$

The c_{ij} coefficients may be determined numerically by substitution of this equation into the Schrödinger equation and application of the variational principle. This method is called the linear combination of atomic orbitals (LCAO) approximation and is used in computational chemistry. An additional unitary transformation can be applied on the system to accelerate the convergence in some computational schemes. Molecular orbital theory was seen as a competitor to valence bond theory in the 1930s, before it was realized that the two methods are closely related and that when extended they become equivalent.

History

Molecular orbital theory was developed, in the years after valence bond theory had been established (1927), primarily through the efforts of Friedrich Hund, Robert Mulliken, John C. Slater, and John Lennard-Jones. MO theory was originally called the Hund-Mulliken theory. The word *orbital* was introduced by Mulliken in 1932. By 1933, the molecular orbital theory had become accepted as a valid and useful theory. According to German physicist and physical chemist Erich Hückel, the first quantitative use of molecular orbital theory was the 1929 paper of Lennard-Jones. The first accurate calculation of a molecular orbital wavefunction was that made by Charles Coulson in 1938 on the hydrogen molecule. By 1950, molecular orbitals were completely defined as eigenfunctions (wave functions) of the self-consistent field Hamiltonian and it was at this point that molecular orbital theory became fully rigorous and consistent. This rigorous approach is known as the Hartree–Fock method for molecules although it had its origins in calculations on atoms. In calculations on molecules, the molecular orbitals are expanded in terms of an atomic orbital basis set, leading to the Roothaan equations. This led to the development of many ab initio quantum chemistry methods. In parallel, molecular orbital theory was applied in a more approximate manner using some empirically derived parameters in methods now known as semi-empirical quantum chemistry methods.

Overview

Molecular orbital (MO) theory uses a linear combination of atomic orbitals (LCAO) to represent molecular orbitals involving the whole molecule. These are often divided into bonding orbitals, anti-bonding orbitals, and non-bonding orbitals. A molecular orbital is merely a Schrödinger orbital which includes several, but often only two nuclei. If this orbital is of type in which the electron(s) in the orbital have a higher probability of being *between* nuclei than elsewhere, the orbital will be a bonding orbital, and will tend to hold the nuclei together. If the electrons tend to be present in a molecular orbital in which they spend more time elsewhere than between the nuclei, the orbital will function as an anti-bonding orbital and will actually weaken the bond. Electrons in non-bonding orbitals tend to be in deep orbitals (nearly atomic orbitals) associated almost entirely with one nucleus

or the other, and thus they spend equal time between nuclei or not. These electrons neither contribute nor detract from bond strength.

Molecular orbitals are further divided according to the types of atomic orbitals combining to form a bond. These orbitals are results of electron-nucleus interactions that are caused by the fundamental force of electromagnetism. Chemical substances will form a bond if their orbitals become lower in energy when they interact with each other. Different chemical bonds are distinguished that differ by electron cloud shape and by energy levels.

MO theory provides a global, delocalized perspective on chemical bonding. For example, in the MO theory for hypervalent molecules it is unnecessary to invoke a major role for d-orbitals, whereas valence bond theory normally uses hybridization with d-orbitals to explain hypervalency. In MO theory, *any* electron in a molecule may be found *anywhere* in the molecule, since quantum conditions allow electrons to travel under the influence of an arbitrarily large number of nuclei, so long as permitted by certain quantum rules. Although in MO theory *some* molecular orbitals may hold electrons which are more localized between specific pairs of molecular atoms, *other* orbitals may hold electrons which are spread more uniformly over the molecule. Thus, overall, bonding (and electrons) are far more delocalized (spread out) in MO theory, than is implied in valence bond (VB) theory. This makes MO theory more useful for the description of extended systems.

An example is that in the MO picture of benzene, composed of a hexagonal ring of 6 carbon atoms. In this molecule, 24 of the 30 total valence bonding electrons are located in 12 σ (sigma) bonding orbitals which are mostly located between pairs of atoms (C-C or C-H), similar to the valence bond picture. However, in benzene the remaining 6 bonding electrons are located in 3 π (pi) molecular bonding orbitals that are delocalized around the ring. Two are in an MO which has equal contributions from all 6 atoms. The other two orbitals have vertical nodes at right angles to each other. As in the VB theory, all of these 6 delocalized pi electrons reside in a larger space which exists above and below the ring plane. All carbon-carbon bonds in benzene are chemically equivalent. In MO theory this is a direct consequence of the fact that the 3 molecular pi orbitals form a combination which evenly spreads the extra 6 electrons over 6 carbon atoms.

In molecules such as methane, the 8 valence electrons are found in 4 MOs that are spread out over all 5 atoms. However, it is possible to approximate the MOs with 4 localized orbitals similar in shape to sp^3 hybrid orbitals predicted by VB theory. This is often adequate for σ (sigma) bonds, but it is not possible for the π (pi) orbitals. However, the delocalized MO picture is more appropriate for ionization and spectroscopic predictions. Upon ionization of methane, a single electron is taken from the MO which surrounds the whole molecule, weakening all 4 bonds equally. VB theory would predict that one electron is removed for an sp^3 orbital, resulting in the need for resonance between four valence bond structures, each of which has a one-electron bond.

As in benzene, in substances such as beta carotene, chlorophyll or heme, some electrons the π (pi) orbitals are spread out in molecular orbitals over long distances in a molecule,

giving rise to light absorption in lower energies (visible colors), a fact which is observed. This and other spectroscopic data for molecules are better explained in MO theory, with an emphasis on electronic states associated with multicenter orbitals, including mixing of orbitals premised on principles of orbital symmetry matching. The same MO principles also more naturally explain some electrical phenomena, such as high electrical conductivity in the planar direction of the hexagonal atomic sheets that exist in graphite. In MO theory, "resonance" (a mixing and blending of VB bond states) is a natural consequence of symmetry. For example, in graphite, as in benzene, it is not necessary to invoke the sp^2 hybridization and resonance of VB theory, in order to explain electrical conduction. Instead, MO theory simply recognizes that some electrons in the graphite atomic sheets are completely delocalized over arbitrary distances, and reside in very large *molecular orbitals* that cover an entire graphite sheet, and some electrons are thus as free to move and conduct electricity *in the sheet plane,* as if they resided in a metal.

Chapter 7

Molecular Model

A **molecular model** is a physical model that represents molecules and their processes. The creation of mathematical models of molecular properties and behaviour is **molecular modelling**, and their graphical depiction is **molecular graphics**, but these topics are closely linked and each uses techniques from the others. In this, "molecular model" will primarily refer to systems containing more than one atom and where nuclear structure is neglected. The electronic structure is often also omitted or represented in a highly simplified way.

Overview

Physical models of atomistic systems have played an important role in understanding chemistry and generating and testing hypotheses. Most commonly there is an explicit representation of atoms, though other approaches such as soap films and other continuous media have been useful. There are several motivations for creating physical models:

- as pedagogic tools for students or those unfamiliar with atomistic structures;
- as objects to generate or test theories (e.g., the structure of DNA);
- as analogue computers (e.g., for measuring distances and angles in flexible systems);
- as aesthetically pleasing objects on the boundary of art and science.

The construction of physical models is often a creative act, and many bespoke examples have been carefully created in the workshops of science departments. There is a very wide range of approaches to physical modelling, and this lists only the most common or historically important. The main strategies are:

- bespoke construction of a single model;
- use of common materials (plasticine, matchsticks) or children's toys (Tinkertoy(TM), Meccano, Lego, etc.);

- re-use of generic components in kits (ca. 1930s to present).

Models encompass a wide range of degrees of precision and engineering: some models such as J.D. Bernal's water are conceptual, while the macromodels of Pauling and Crick and Watson were created with much greater precision.

Molecular models have inspired molecular graphics, initially in textbooks and research articles and more recently on computers. Molecular graphics has replaced some functions of physical molecular models, but physical kits continue to be very popular and are sold in large numbers. Their unique strengths include:

- cheapness and portability;
- immediate tactile and visual messages;
- easy interactivity for many processes (e.g., conformational analysis and pseudorotation).

History

In the 1600s, Johannes Kepler speculated on the symmetry of snowflakes and also on the close packing of spherical objects such as fruit (this problem remained unsolved until very recently). The symmetrical arrangement of closely packed spheres informed theories of molecular structure in the late 1800s, and many theories of crystallography and solid state inorganic structure used collections of equal and unequal spheres to simulate packing and predict structure.

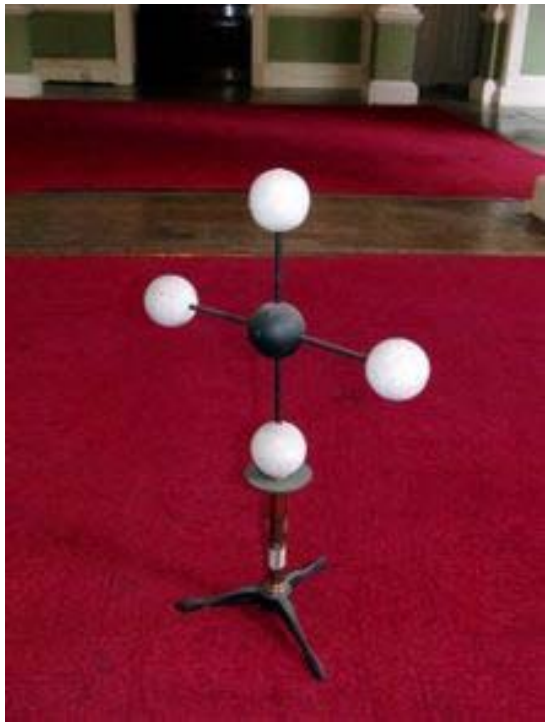


Fig. 1. Hofmann's model for methane.

John Dalton represented compounds as aggregations of circular atoms, and although Loschmidt did not create physical models, his diagrams based on circles are two-dimensional analogues of later models. Hofmann is credited with the first physical molecular model around 1860 (Fig. 1). Note how the size of the carbon appears smaller than the hydrogen. The importance of stereochemistry was not then recognised and the model is essentially topological (it should be a 3-dimensional tetrahedron).

J.H. van 't Hoff and J. le Bel introduced the concept of chemistry in space—stereochemistry in three dimensions. van 't Hoff built tetrahedral molecules representing the three-dimensional properties of carbon.

Models based on spheres

Robert Hooke proposed a relationship between crystals and the packing of spheres . R. Haüy argued that the structures of crystals involved regular lattices of repeating units with shapes similar to the macroscopic crystal. Barlow, who jointly developed the theories of space groups, proposed models of crystals based on sphere packings (ca. 1890).

Fig. 2. Sodium chloride (NaCl) lattice, showing close-packed spheres representing a face-centered cubic AB lattice similar to that of NaCl and most other alkali halides. In this

model the spheres are equal sizes whereas more "realistic" models would have different radii for cations and anions.

The binary compounds sodium chloride (NaCl) and caesium chloride (CsCl) have cubic structures but have different space groups. This can be rationalised in terms of close packing of spheres of different sizes. For example, NaCl can be described as close-packed chloride ions (in a face-centered cubic lattice) with sodium ions in the octahedral holes. After the development of X-ray crystallography as a tool for determining crystal structures, many laboratories built models based on spheres. With the development of plastic or polystyrene balls it is now easy to create such models.

Models based on ball-and-stick

The concept of the chemical bond as a direct link between atoms can be modelled by linking balls (atoms) with sticks/rods (bonds). This has been extremely popular and is still widely used today. Initially atoms were made of spherical wooden balls with specially drilled holes for rods. Thus carbon can be represented as a sphere with four holes at the tetrahedral angles $\cos^{-1}(-1/3) \approx 109.47°$.

A problem with rigid bonds and holes is that systems with arbitrary angles could not be built. This can be overcome with flexible bonds, originally helical springs but now usually plastic. This also allows double and triple bonds to be approximated by multiple single bonds (Fig. 3).

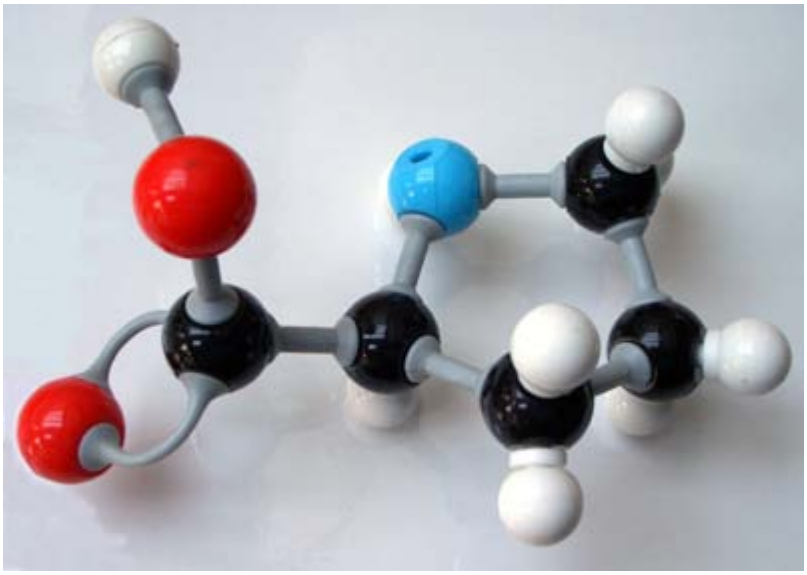


Fig 3. A modern plastic ball and stick model. The molecule shown is proline.

Figure 3 represents a ball-and-stick model of proline. The balls have colours: **black** represents carbon (C); **red**, oxygen (O); **blue**, nitrogen (N); and white, hydrogen (H). Each ball is drilled with as many holes as its conventional valence (C: 4; N: 3; O: 2; H: 1) directed towards the vertices of a tetrahedron. Single bonds are represented by (fairly)

rigid grey rods. Double and triple bonds use two longer flexible bonds which restrict rotation and support conventional cis/trans stereochemistry.

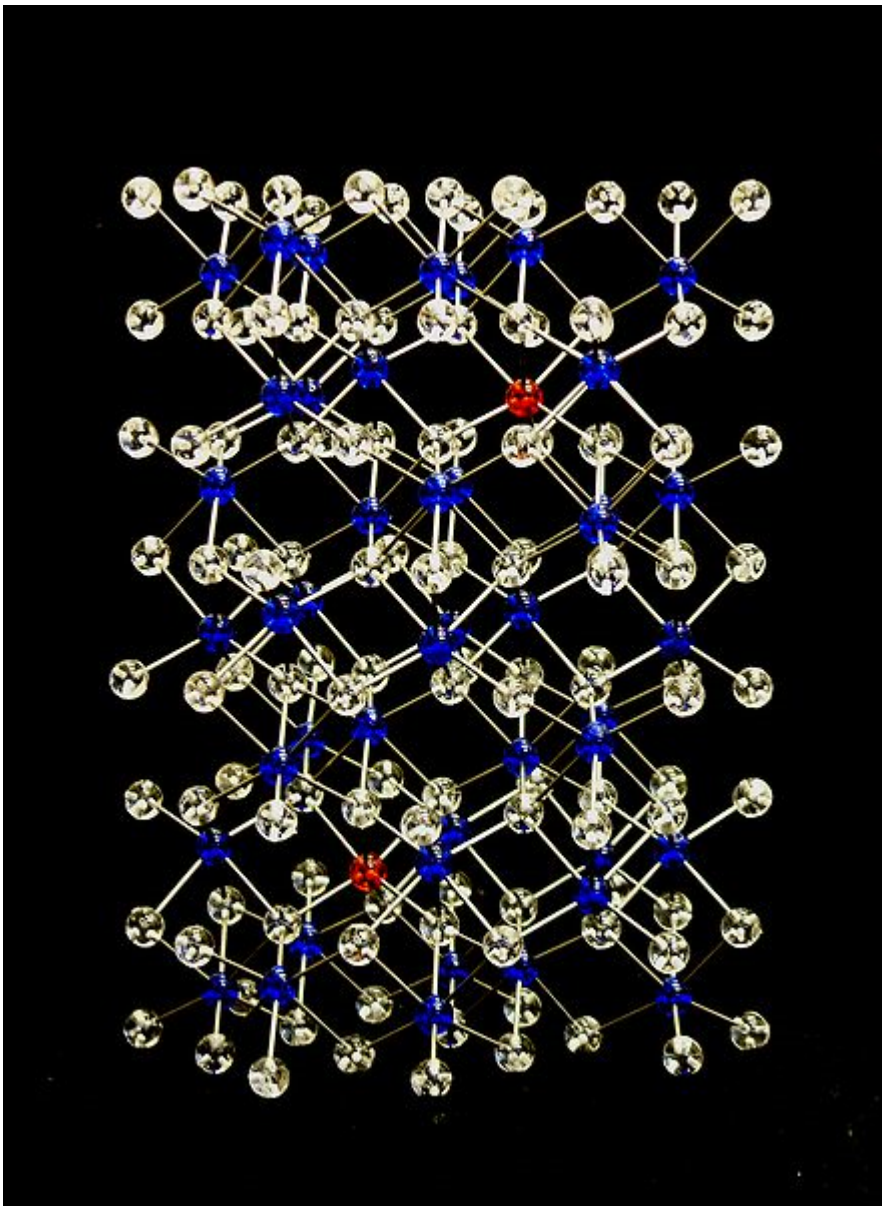


Fig. 4. Beevers ball and stick model of ruby (Cr-doped corundum) made with acrylic balls and stainless steel rods.

However, most molecules require holes at other angles and specialist companies manufacture kits and bespoke models. One of the earlier companies was Woosters at Bottisham, Cambridgeshire, UK. Besides tetrahedral, trigonal and octahedral holes, there were all-purpose balls with 24 holes. These models allowed rotation about the single rod bonds, which could be both an advantage (showing molecular flexibility) and a disadvantage (models are floppy). The approximate scale was 5 cm per ångström (0.5 m/nm or 500,000,000:1), but was not consistent over all elements.

Arnold Beevers in Edinburgh (now operating as Miramodus) created small models using PMMA balls and stainless steel rods. By using individually drilled balls with precise bond angles and bond lengths in these models, large crystal structures to be accurately created, but with light and rigid form. Figure 4 shows a unit cell of ruby in this style.

Skeletal models

Crick and Watson's DNA model and the protein-building kits of Kendrew were among the first skeletal models. These were based on atomic components where the valences were represented by rods; the atoms were points at the intersections. Bonds were created by linking components with tubular connectors with locking screws.

Andre Dreiding introduced a molecular modelling kit (ca. 1975) which dispensed with the connectors. A given atom would have solid and hollow valence spikes. The solid rods clicked into the tubes forming a bond, usually with free rotation. These were and are very widely used in organic chemistry departments and were made so accurately that interatomic measurements could be made by ruler.

More recently, inexpensive plastic models (such as Orbit) use a similar principle. A small plastic sphere has protuberances onto which plastic tubes can be fitted. The flexibility of the plastic means that distorted geometries can be made.

Polyhedral models

Many inorganic solids consist of atoms surrounded by a coordination sphere of electronegative atoms (e.g. PO_4 tetrahedra, TiO_6 octahedra). Structures can be modelled by gluing together polyhedra made of paper or plastic.

Composite models

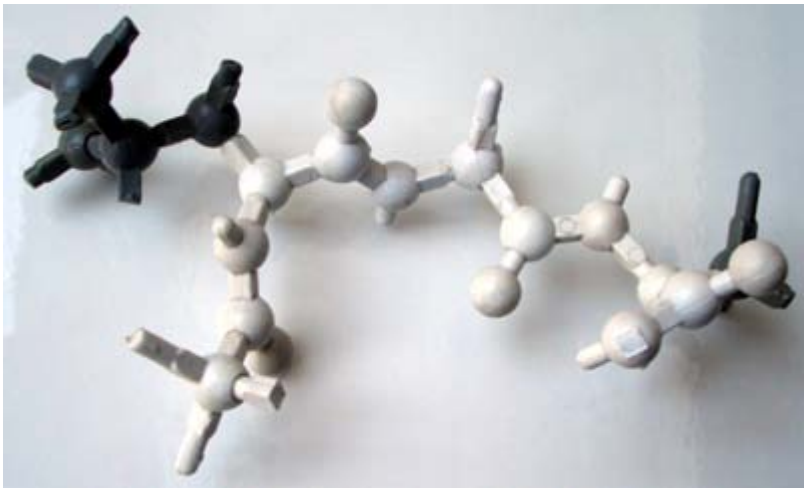


Fig. 5. A Nicholson model, showing a short part of protein backbone (white) with side chains (grey). Note the snipped stubs representing hydrogen atoms.

A good example of composite models is the Nicholson approach, widely used from the late 1970s for building models of biological macromolecules. The components are primarily amino acids and nucleic acids with preformed residues representing groups of atoms. Many of these atoms are directly moulded into the template, and fit together by pushing plastic stubs into small holes. The plastic grips well and makes bonds difficult to rotate, so that arbitrary torsion angles can be set and retain their value. The conformations of the backbone and side chains are determined by pre-computing the torsion angles and then adjusting the model with a protractor.

The plastic is white and can be painted to distinguish between O and N atoms. Hydrogen atoms are normally implicit and modelled by snipping off the spokes. A model of a typical protein with approximately 300 residues could take a month to build. It was common for laboratories to build a model for each protein solved. By 2005, so many protein structures were being determined that relatively few models were made.

Computer-based models

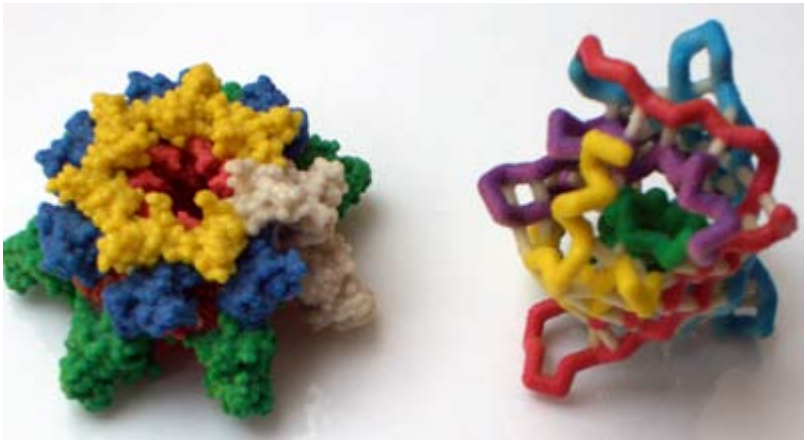


Fig. 6. Integrated protein models.

With the development of computer-based physical modelling, it is now possible to create complete single-piece models by feeding the coordinates of a surface into the computer. Figure 6 shows models of anthrax toxin, left (at a scale of approximately 20 Å/cm or 1:5,000,000) and green fluorescent protein, right (5 cm high, at a scale of about 4 Å/cm or 1:25,000,000) from 3D Molecular Design. Models are made of plaster or starch, using a rapid prototyping process.

It has also recently become possible to create accurate molecular models inside glass blocks using a technique known as subsurface laser engraving. The image at right (Fig. 7) shows the 3D structure of an *E. coli* protein (DNA polymerase beta-subunit, PDB code 1MMI) etched inside a block of glass by British company Luminorum Ltd.

Fig. 7. Protein model in glass.

Common colors

Some of the most common colors used in molecular models are as follows:

Hydrogen	White
Alkali Metals	Violet
Alkaline-Earth Metals	Dark Green
Boron, Most Transition Metals	Peach/Salmon
Carbon	Black
Nitrogen	Dark Blue
Oxygen	Red
Fluorine, Chlorine	Green
Bromine	Dark Red
Iodine	Dark Violet
Noble Gases	Cyan
Phosphorus	Orange
Sulfur	Yellow
Titanium	Gray

Chronology

This table is an incomplete chronology of events where physical molecular models provided major scientific insights.

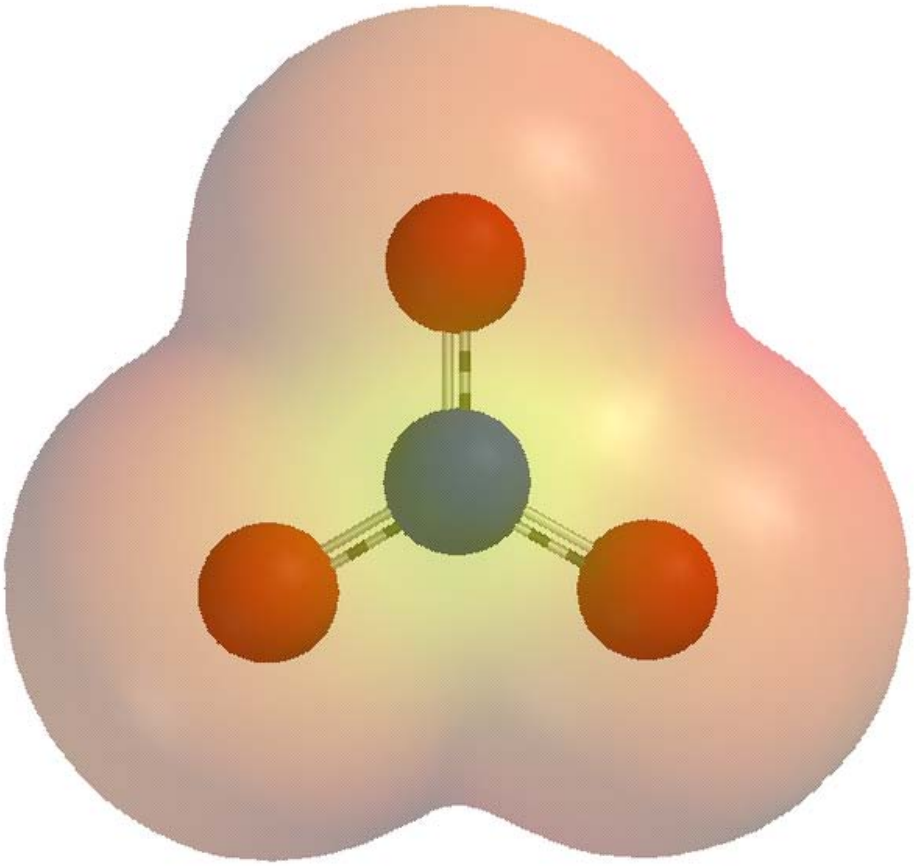
developer(s)	**approximate date**	**technology**	**comments**

Kepler			sphere packing, symmetry of snowflakes.
Loschmidt		2-D graphics	representation of atoms and bonds by touching circles
Hofmann		ball-and-stick	first recognisable physical molecular model
van't Hoff		paper?	representation of atoms as tetrahedra supported the development of stereochemistry
Bernal		Plasticine and spokes	model of liquid water
Corey, Pauling, Koltun (CPK coloring)		Space filling models of alpha-helix, etc.	Pauling's "Nature of the Chemical Bond" covered all aspects of molecular structure and influenced many aspects of models
Crick and Watson		spikes, flat templates and connectors with screws	model of DNA
Molecular graphics	ca 1960	display on computer screens	complements rather than replaces physical models

Chapter 8

Polyatomic Ion and Molecular Hamiltonian

Polyatomic ion



An electrostatic potential map of the nitrate ion (NO_3^-). Areas coloured red are lower in energy than areas coloured yellow

A **polyatomic ion**, also known as a **molecular ion**, is a charged species (ion) composed of two or more atoms covalently bonded or of a metal complex that can be considered as acting as a single unit in the context of acid and base chemistry or in the formation of salts. The prefix "poly-" means "many," in Greek, but even ions of two atoms are commonly referred to as polyatomic. In older literature, a polyatomic ion is also referred to as a **radical**, and less commonly, as a **radical group**. In contemporary usage, the term radical refers to free radicals which are (not necessarily charged) species with an unpaired electron.

For example, a hydroxide ion is made of one oxygen atom and one hydrogen atom: its chemical formula is $(OH)^-$. It has a charge of −1. An ammonium ion is made up of one nitrogen atom and four hydrogen atoms: its chemical formula is $(NH_4)^+$. It has charge of +1.

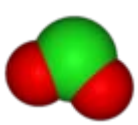
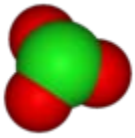
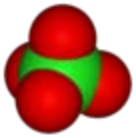
A polyatomic ion can often be considered as the conjugate acid or conjugate base of a neutral molecule. For example the sulfate anion, SO_4^{2-}, is derived from H_2SO_4 which can be regarded as $SO_3 + H_2O$.

Nomenclature

There are two "rules" that can be used for learning the nomenclature of polyatomic ions. First, when the prefix bi- is added to a name, a hydrogen is added to the ion's formula and its charge is increased by 1, the latter being a consequence of the hydrogen ion carrying a +1 charge. An alternate to the bi- prefix is to use the word hydrogen in its place: the anion derived from $H^+ + CO_3^{2-}$, HCO_3^- can be called either bicarbonate or hydrogen carbonate.

Note that many of the common polyatomic anions are conjugate bases of acids derived from the oxides of non-metallic elements. For example the sulfate anion, SO_4^{2-}, is derived from H_2SO_4 which can be regarded as $SO_3 + H_2O$.

The second rule looks at the number of oxygens in an ion. Consider the chlorine oxoanion family:

oxidation state	−1	+1	+3	+5	+7
anion name	chloride	hypochlorite	chlorite	chlorate	perchlorate
formula	Cl^-	ClO^-	ClO_2^-	ClO_3^-	ClO_4^-
structure					

First, think of the -ate ion as being the "base" name, in which case the addition of a per- prefix adds an oxygen. Changing the -ate suffix to -ite will reduce the oxygens by one, and keeping the suffix -ite and adding the prefix hypo- reduces the number of oxygens by

two. In all situations, the charge is not affected. The naming pattern follows within many different oxyanion series based on a standard root for that particular series. The -ite has one less oxygen than the -ate, but different -ate anions might have different numbers of oxygen atoms.

These rules will not work with all polyatomic ions, but they do work with the most common ones (sulfate, phosphate, nitrate, chlorate).

Examples of common polyatomic ions

The following tables give examples of commonly-encountered polyatomic ions. Only a few representatives are given, as the number of polyatomic ions encountered in practice is very large.

Anions	
Acetate (ethanoate)	CH_3COO^- or $C_2H_3O_2^-$
Benzoate	$C_6H_5COO^-$ or $C_7H_5O_2^-$
Bicarbonate (hydrogen carbonate)	HCO_3^-
Carbonate	CO_3^{2-}
Cyanide	CN^-
Hydroxide	OH^-
Nitrate	NO_3^-
Phosphate	PO_4^{3-}
Sulfate	SO_4^{2-}

Cations	
Ammonium	NH_4^+
Hydronium	H_3O^+
Mercury(I)	Hg_2^{2+}
Tropylium	$C_7H_7^+$

Molecular Hamiltonian

In atomic, molecular, and optical physics as well as in quantum chemistry, **molecular Hamiltonian** is the name given to the Hamiltonian representing the energy of the electrons and nuclei in a molecule. This Hermitian operator and the associated Schrödinger equation play a central role in computational chemistry and physics for computing properties of molecules and aggregates of molecules, such as thermal conductivity, specific heat, electrical conductivity, optical, and magnetic properties, and reactivity.

The bricks of the molecule are the nuclei, characterized by their atomic number, Z, while the electrons, which have negative elementary charge, q, are its mortar. The charge of a nucleus is Zq. Electrons and nuclei are, to a very good approximation, point charges and point masses. The molecular Hamiltonian is a sum of several terms: its major terms are the kinetic energies of the electrons and the Coulomb (electrostatic) interactions between the two kinds of charged particles. The Hamiltonian that contains only the kinetic energies of electrons and nuclei, and the Coulomb interactions between them, is known as the **Coulomb Hamiltonian**. From it are missing a number of small terms, most of which are due to electronic and nuclear spin.

Although it is generally assumed that the solution of the time-independent Schrödinger equation associated with the Coulomb Hamiltonian will predict most properties of the molecule, including its shape (three-dimensional structure), calculations based on the full Coulomb Hamiltonian are very rare. The main reason is that its Schrödinger equation is very difficult to solve. Applications are restricted to small systems like the hydrogen molecule.

Almost all calculations of molecular wavefunctions are based on the separation of the Coulomb Hamiltonian first devised by Born and Oppenheimer. The nuclear kinetic energy terms are omitted from the Coulomb Hamiltonian and one considers the remaining Hamiltonian as a Hamiltonian of electrons only. The stationary nuclei enter the problem only as generators of an electric potential in which the electrons move in a quantum mechanical way. Within this framework the molecular Hamiltonian has been simplified to the so-called **clamped nucleus Hamiltonian**, also called **electronic Hamiltonian**, that acts only on functions of the electronic coordinates.

Once the Schrödinger equation of the clamped nucleus Hamiltonian has been solved for a sufficient number of constellations of the nuclei, an appropriate eigenvalue (usually the lowest) can be seen as a function of the nuclear coordinates, which leads to a potential energy surface. In practical calculations the surface is usually fitted in terms of some analytic functions. In the second step of the Born-Oppenheimer approximation the part of the full Coulomb Hamiltonian that depends on the electrons is replaced by the potential energy surface. This converts the total molecular Hamiltonian into another Hamiltonian that acts only on the nuclear coordinates. In the case of a breakdown of the Born-Oppenheimer approximation—which occurs when energies of different electronic states are close—the neighboring potential energy surfaces are needed.

The nuclear motion Schrödinger equation can be solved in a space-fixed (laboratory) frame, but then the translational and rotational (external) energies are not accounted for. Only the (internal) atomic vibrations enter the problem. Further, for molecules larger than triatomic ones, it is quite common to introduce the harmonic approximation, which approximates the potential energy surface as a quadratic function of the atomic displacements. This gives the **harmonic nuclear motion Hamiltonian**. Making the harmonic approximation, we can convert the Hamiltonian into a sum of uncoupled one-dimensional harmonic oscillator Hamiltonians. The one-dimensional harmonic oscillator is one of the few systems that allows an exact solution of the Schrödinger equation.

Alternatively, the nuclear motion (rovibrational) Schrödinger equation can be solved in a special frame (an Eckart frame) that rotates and translates with the molecule. Formulated with respect to this body-fixed frame the Hamiltonian accounts for rotation, translation and vibration of the nuclei. Since Watson introduced in 1968 an important simplification to this Hamiltonian, it is often referred to as **Watson's nuclear motion Hamiltonian**, but it is also known as the **Eckart Hamiltonian**.

Coulomb Hamiltonian

The algebraic form of many observables—i.e., Hermitian operators representing observable quantities—is obtained by the following quantization rules:

- Write the classical form of the observable in Hamilton form (as a function of momenta **p** and positions **q**). Both vectors are expressed with respect to an arbitrary inertial frame, usually referred to as *laboratory-frame* or *space-fixed frame*.
- Replace **p** by $-i\hbar\nabla$ and interpret **q** as a multiplicative operator. Here ∇ is the nabla operator, a vector operator consisting of first derivatives. The well-known commutation relations for the **p** and **q** operators follow directly from the differentiation rules.

Classically the electrons and nuclei in a molecule have kinetic energy of the form $p^2/(2m)$ and interact via Coulomb interactions, which are inversely proportional to the distance r_{ij} between particle *i* and *j*.

$$r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j| = \sqrt{(\mathbf{r}_i - \mathbf{r}_j)\cdot(\mathbf{r}_i - \mathbf{r}_j)} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}.$$

In this expression $\mathbf{r}_i$ stands for the coordinate vector of any particle (electron or nucleus). But from here on we will reserve capital **R** to represent the nuclear coordinate, and lower case **r** for the electrons of the system. The coordinates can be taken to be expressed with respect to any Cartesian frame centered anywhere in space, because distance, being an inner product, is invariant under rotation of the frame and, being the norm of a difference vector, distance is invariant under translation of the frame as well.

By quantizing the classical energy in Hamilton form one obtains the a molecular Hamilton operator that is often referred to as the **Coulomb Hamiltonian**. This Hamiltonian is a sum of 5 terms. They are

1. The kinetic energy operators for each nucleus in the system;
2. The kinetic energy operators for each electron in the system;
3. The potential energy between the electrons and nuclei - the total electron-nucleus Coulombic attraction in the system;
4. The potential energy arising from Coulombic electron-electron repulsions
5. The potential energy arising from Coulombic nuclei-nuclei repulsions - also known as the nuclear repulsion energy.

1. $$\hat{T}_n = -\sum_i \frac{\hbar^2}{2M_i}\nabla^2(\mathbf{R}_i)$$
2. $$\hat{T}_e = -\sum_i \frac{\hbar^2}{2m_e}\nabla^2(\mathbf{r}_i)$$
3. $$\hat{U}_{en} = -\sum_i \sum_j \frac{Z_i e^2}{4\pi\epsilon_0 \left|\mathbf{R}_i - \mathbf{r}_j\right|}$$
4. $$\hat{U}_{ee} = \frac{1}{2}\sum_i \sum_{j\neq i} \frac{e^2}{4\pi\epsilon_0 \left|\mathbf{r}_i - \mathbf{r}_j\right|} = \sum_i \sum_{j>i} \frac{e^2}{4\pi\epsilon_0 \left|\mathbf{r}_i - \mathbf{r}_j\right|}$$
5. $$\hat{U}_{nn} = \frac{1}{2}\sum_i \sum_{j\neq i} \frac{Z_i Z_j e^2}{4\pi\epsilon_0 \left|\mathbf{R}_i - \mathbf{R}_j\right|} = \sum_i \sum_{j>i} \frac{Z_i Z_j e^2}{4\pi\epsilon_0 \left|\mathbf{R}_i - \mathbf{R}_j\right|}.$$

Here M_i is the mass of nucleus i, Z_i is the atomic number of nucleus i, and m_e is the mass of the electron. The Laplace operator of particle i is : $\nabla^2(\mathbf{r}_i) \equiv \boldsymbol{\nabla}(\mathbf{r}_i)\cdot\boldsymbol{\nabla}(\mathbf{r}_i) = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$. Since the kinetic energy operator is an inner product, it is invariant under rotation of the Cartesian frame with respect to which x_i, y_i, and z_i are expressed. The kinetic energy operator, however, is not invariant under translation (choice of origin of the frame).

Small terms

In the 1920s much spectroscopic evidence made it clear that the Coulomb Hamiltonian is missing certain terms. Especially for molecules containing heavier atoms, these terms, although much smaller than kinetic and Coulomb energies, are nonnegligible. These spectroscopic observations led to the introduction of a new degree of freedom for electrons and nuclei, namely spin. This empirical concept was given a theoretical basis by Paul Dirac when he introduced a relativistically correct (Lorentz covariant) form of the one-particle Schrödinger equation. The Dirac equation predicts that spin and spatial

motion of a particle interact via spin-orbit coupling. In analogy spin-other-orbit coupling was introduced. The fact that particle spin has some of the characteristics of a magnetic dipole led to spin-spin coupling. Further terms without a classical counterpart are the Fermi-contact term (interaction of electronic density on a finite size nucleus with the nucleus), and nuclear quadrupole coupling (interaction of a nuclear quadrupole with the gradient of an electric field due to the electrons). Finally a parity violating term predicted by the Standard Model must be mentioned. Although it is an extremely small interaction, it has attracted a fair amount of attention in the scientific literature because it gives different energies for the enantiomers in chiral molecules.

The Schrödinger equation of the Coulomb Hamiltonian

The Coulomb Hamiltonian has a continuous spectrum due to the center of mass motion of the molecule in homogeneous space. In classical mechanics it is easy to separate off the center of mass (COM) motion of a system of point masses. Classically the motion of the COM is uncoupled from the other motions. The COM moves uniformly (i.e., with constant velocity) through space as if it were a point particle with mass equal to the sum M_{tot} of the masses of all the particles.

In quantum mechanics a free particle has as state function a plane wave function, which is a non-square-integrable function of well-defined momentum. The kinetic energy of this particle covers $\mathbb{R}^+$. The position of the COM is uniformly probable everywhere, in agreement with the Heisenberg uncertainty principle.

Quantum mechanically the proper separation of the COM motion is much more cumbersome than in classical mechanics. By introducing the coordinate vector $\mathbf{X}$ of the center of mass as three of the degrees of freedom of the system and eliminating the coordinate vector of one (arbitrary) particle, so that the number of degrees of freedom stays the same, one obtains by a linear transformation a new set of coordinates $\mathbf{t}_i$. These coordinates are linear combinations of the old coordinates of *all* particles (nuclei *and* electrons). By applying the chain rule one can show that

$$H = -\frac{\hbar^2}{2M_{\text{tot}}}\nabla^2(\mathbf{X}) + H' \quad \text{with} \quad H' = -\sum_{i=1}^{N_{\text{tot}}-1}\frac{\hbar^2}{\mu_{ii}}\nabla^2(\mathbf{t}_i) - \sum_{i,j=1}^{N_{\text{tot}}-1}\frac{\hbar^2}{\mu_{ij}}\boldsymbol{\nabla}(\mathbf{t}_i)\cdot\boldsymbol{\nabla}(\mathbf{t}_j) + V(\mathbf{t}).$$

The first term of H is the kinetic energy of the COM motion, which can be treated separately since H' does not depend on $\mathbf{X}$. As just stated, its eigenstates are plane waves. The constants $1/\mu_{ij}$ are positive and are linear combinations of all the inverse masses $1/m_i$. They are generalized reduced masses. The potential $V(\mathbf{t})$ consists of the Coulomb terms expressed in the new coordinates. The first term of H' has the usual appearance of a kinetic energy operator. The second term is known as the **mass polarization** term. The translationally invariant Hamiltonian H' can be shown to be self-adjoint and to be bounded from below. That is, its lowest eigenvalue is real and finite. Although H' is necessarily invariant under permutations of identical particles (since H and the COM kinetic energy are invariant), its invariance is not manifest.

Not many actual molecular applications of H' exist, see, however, the seminal work on the hydrogen molecule for an early application. In the great majority of computations of molecular wavefunctions the electronic problem is solved with the clamped nucleus Hamiltonian arising in the first step of the Born-Oppenheimer approximation.

Clamped nucleus Hamiltonian

The clamped nucleus Hamiltonian describes the energy of the electrons in the electrostatic field of the nuclei, where the nuclei are assumed to be stationary with respect to an inertial frame. The form of the electronic Hamiltonian is

$$\hat{H}_{\mathrm{el}} = \hat{T}_e + \hat{U}_{en} + \hat{U}_{ee} + \hat{U}_{nn}.$$

The coordinates of electrons and nuclei are expressed with respect to a frame that moves with the nuclei, so that the nuclei are at rest with respect to this frame. The frame stays parallel to a space-fixed frame. It is an inertial frame because the nuclei are assumed not to be accelerated by external forces or torques. The origin of the frame is arbitrary, it is usually positioned on a central nucleus or in the nuclear center of mass. Sometimes it is stated that the nuclei are "at rest in a space-fixed frame". This statement implies that the nuclei are viewed as classical particles, because a quantum mechanical particle cannot be at rest. (It would mean that it had simultaneously zero momentum and well-defined position, which contradicts Heisenberg's uncertainty principle).

Since the nuclear positions are constants, the electronic kinetic energy operator is invariant under translation over any nuclear vector. The Coulomb potential, depending on difference vectors, is invariant as well. In the description of atomic orbitals and the computation of integrals over atomic orbitals this invariance is used by equipping all atoms in the molecule with their own localized frames parallel to the space-fixed frame.

As explained in the Born-Oppenheimer approximation, a sufficient number of solutions of the Schrödinger equation of H_{el} leads to a potential energy surface (PES) $V(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N)$. It is assumed that the functional dependence of V on its coordinates is such that

$$V(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N) = V(\mathbf{R}'_1, \mathbf{R}'_2, \ldots, \mathbf{R}'_N)$$

for

$$\mathbf{R}'_i = \mathbf{R}_i + \mathbf{t} \ \text{(translation)} \ \text{and} \ \mathbf{R}'_i = \mathbf{R}_i + \frac{\Delta\phi}{|\mathbf{s}|}(\mathbf{s} \times \mathbf{R}_i) \ \text{(infinitesimal rotation)},$$

where **t** and **s** are arbitrary vectors and $\Delta\varphi$ is an infinitesimal angle, $\Delta\varphi >> \Delta\varphi^2$. This invariance condition on the PES is automatically fulfilled when the PES is expressed in terms of differences of, and angles between, the $\mathbf{R}_i$, which is usually the case.

Harmonic nuclear motion Hamiltonian

In the remaining part we assume that the molecule is semi-rigid. In the second step of the BO approximation the nuclear kinetic energy T_n is reintroduced and the Schrödinger equation with Hamiltonian

$$\hat{H}_{\mathrm{nuc}} = -\frac{\hbar^2}{2}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{1}{M_i}\frac{\partial^2}{\partial R_{i\alpha}^2} + V(\mathbf{R}_1,\ldots,\mathbf{R}_N)$$

is considered. One would like to recognize in its solution: the motion of the nuclear center of mass (3 degrees of freedom), the overall rotation of the molecule (3 degrees of freedom), and the nuclear vibrations. In general, this is not possible with the given nuclear kinetic energy, because it does not separate explicitly the 6 external degrees of freedom (overall translation and rotation) from the 3*N*-6 internal degrees of freedom. In fact, the kinetic energy operator here is defined with respect to a space-fixed (SF) frame. If we were to move the origin of the SF frame to the nuclear center of mass, then, by application of the chain rule, nuclear mass polarization terms would appear. It is customary to ignore these terms altogether and we will follow this custom.

In order to achieve a separation we must distinguish internal and external coordinates, to which end Eckart introduced conditions to be satisfied by the coordinates. We will show how these conditions arise in a natural way from a harmonic analysis in mass-weighted Cartesian coordinates.

In order to simplify the expression for the kinetic energy we introduce mass-weighted displacement coordinates

$$\boldsymbol{\rho}_i \equiv \sqrt{M_i}(\mathbf{R}_i - \mathbf{R}_i^0).$$

Since

$$\frac{\partial}{\partial \rho_{i\alpha}} = \frac{\partial}{\sqrt{M_i}(\partial R_{i\alpha} - \partial R_{i\alpha}^0)} = \frac{1}{\sqrt{M_i}}\frac{\partial}{\partial R_{i\alpha}},$$

the kinetic energy operator becomes,

$$T = -\frac{\hbar^2}{2}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\frac{\partial^2}{\partial \rho_{i\alpha}^2}.$$

If we make a Taylor expansion of *V* around the equilibrium geometry,

$$V = V_0 + \sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left(\frac{\partial V}{\partial \rho_{i\alpha}}\right)_0 \rho_{i\alpha} + \frac{1}{2}\sum_{i,j=1}^{N}\sum_{\alpha,\beta=1}^{3}\left(\frac{\partial^2 V}{\partial \rho_{i\alpha}\partial \rho_{j\beta}}\right)_0 \rho_{i\alpha}\rho_{j\beta} + \cdots,$$

and truncate after three terms (the so-called harmonic approximation), we can describe V with only the third term. The term V_0 can be absorbed in the energy (gives a new zero of energy). The second term is vanishing because of the equilibrium condition. The remaining term contains the Hessian matrix **F** of V, which is symmetric and may be diagonalized with an orthogonal $3N$ x $3N$ matrix with constant elements:

$$\mathbf{QFQ}^{\mathrm{T}} = \mathbf{\Phi} \quad \text{with} \quad \mathbf{\Phi} = \mathrm{diag}(f_1, \ldots, f_{3N-6}, 0, \ldots, 0).$$

It can be shown from the invariance of V under rotation and translation that six of the eigenvectors of **F** (last six rows of **Q**) have eigenvalue zero (are zero-frequency modes). They span the *external space*. The first $3N$-6 rows of **Q** are—for molecules in their ground state—eigenvectors with non-zero eigenvalue; they are the internal coordinates and form an orthonormal basis for a $3N$ - 6 dimensional subspace of the nuclear configuration space $\mathbf{R}^{3N}$, the *internal space*. The zero-frequency eigenvectors are orthogonal to the eigenvectors of non-zero frequency. It can be shown that these orthogonalities are in fact the Eckart conditions. The kinetic energy expressed in the internal coordinates is the internal (vibrational) kinetic energy.

With the introduction of normal coordinates

$$q_t \equiv \sum_{i=1}^{N}\sum_{\alpha=1}^{3} Q_{t,i\alpha}\rho_{i\alpha},$$

the vibrational (internal) part of the Hamiltonian for the nuclear motion becomes in the *harmonic approximation*

$$\hat{H}_{\mathrm{nuc}} \approx \frac{1}{2}\sum_{t=1}^{3N-6}\left[-\hbar^2\frac{\partial^2}{\partial q_t^2} + f_t q_t^2\right].$$

The corresponding Schrödinger equation is easily solved, it factorizes into $3N$-6 equations for one-dimensional harmonic oscillators. The main effort in this approximate solution of the nuclear motion Schrödinger equation is the computation of the Hessian **F** of V and its diagonalization.

This approximation to the nuclear motion problem, described in $3N$ mass-weighted Cartesian coordinates, became standard in quantum chemistry, since the days (1980s-1990s) that algorithms for accurate computations of the Hessian **F** became available. Apart from the harmonic approximation, it has as a further deficiency that the external (rotational and translational) motions of the molecule are not accounted for. They are

accounted for in a rovibrational Hamiltonian that sometimes is called *Watson's Hamiltonian*.

Watson's nuclear motion Hamiltonian

In order to obtain a Hamiltonian for external (translation and rotation) motions coupled to the internal (vibrational) motions, it is common to return at this point to classical mechanics and to formulate the classical kinetic energy corresponding to these motions of the nuclei. Classically it is easy to separate the translational—center of mass—motion from the other motions. However, the separation of the rotational from the vibrational motion is more difficult and is not completely possible. This ro-vibrational separation was first achieved by Eckart in 1935 by imposing by what is now known as Eckart conditions. Since the problem is described in a frame (an "Eckart" frame) that rotates with the molecule, and hence is a non-inertial frame, energies associated with the fictitious forces: centrifugal and Coriolis force appear in the kinetic energy.

In general, the classical kinetic energy T defines the metric tensor $\mathbf{g} = (g_{ij})$ associated with the curvilinear coordinates $\mathbf{s} = (s_i)$ through

$$2T = \sum_{ij} g_{ij}\dot{s}_i\dot{s}_j$$

.

The quantization step is the transformation of this classical kinetic energy into a quantum mechanical operator. It is common to follow Podolsky by writing down the Laplace-Beltrami operator in the same (generalized, curvilinear) coordinates **s** as used for the classical form. The equation for this operator requires the inverse of the metric tensor **g** and its determinant. Multiplication of the Laplace-Beltrami operator by $-\hbar^2$ gives the required quantum mechanical kinetic energy operator. When we apply this recipe to Cartesian coordinates, which have unit metric, the same kinetic energy is obtained as by application of the quantization rules.

The nuclear motion Hamiltonian was obtained by Wilson and Howard in 1936, who followed this procedure, and further refined by Darling and Dennison in 1940. It remained the standard until 1968, when Watson was able to simplify it drastically by commuting through the derivatives the determinant of the metric tensor. We will give the ro-vibrational Hamiltonian obtained by Watson, which often is referred to as the **Watson Hamiltonian**. Before we do this we must mention that a derivation of this Hamiltonian is also possible by starting from the Laplace operator in Cartesian form, application of coordinate transformations, and use of the chain rule. The Watson Hamiltonian, describing all motions of the N nuclei, is

$$\hat{H} = -\frac{\hbar^2}{2M_{\text{tot}}}\sum_{\alpha=1}^{3}\frac{\partial^2}{\partial X_\alpha^2} + \frac{1}{2}\sum_{\alpha,\beta=1}^{3}\mu_{\alpha\beta}(\mathcal{P}_\alpha - \Pi_\alpha)(\mathcal{P}_\beta - \Pi_\beta) + U - \frac{\hbar^2}{2}\sum_{s=1}^{3N-6}\frac{\partial^2}{\partial q_s^2} + V.$$

The first term is the center of mass term

$$\mathbf{X} \equiv \frac{1}{M_{\mathrm{tot}}} \sum_{i=1}^{N} M_i \mathbf{R}_i \quad \text{with} \quad M_{\mathrm{tot}} \equiv \sum_{i=1}^{N} M_i.$$

The second term is the rotational term akin to the kinetic energy of the rigid rotor. Here $\mathcal{P}_\alpha$is the α component of the body-fixed *rigid rotor angular momentum operator*. The operator $\Pi_\alpha$is a component of an operator known as the *vibrational angular momentum operator* (although it does *not* satisfy angular momentum commutation relations),

$$\Pi_\alpha = -i\hbar \sum_{s,t=1}^{3N-6} \zeta_{st}^{\alpha}\, q_s \frac{\partial}{\partial q_t}$$

with the *Coriolis coupling constant*:

$$\zeta_{st}^{\alpha} = \sum_{i=1}^{N} \sum_{\beta,\gamma=1}^{3} \epsilon_{\alpha\beta\gamma} Q_{s,i\beta}\, Q_{t,i\gamma} \quad \text{and} \quad \alpha = 1,2,3.$$

Here $\varepsilon_{\alpha\beta\gamma}$ is the Levi-Civita symbol. The terms quadratic in the $\mathcal{P}_\alpha$are centrifugal terms, those bilinear in $\mathcal{P}_\alpha$and $\Pi_\beta$are Coriolis terms. The quantities $Q_{s,i\gamma}$ are the components of the normal coordinates introduced above. Alternatively, normal coordinates may be obtained by application of Wilson's GF method. The 3 x 3 symmetric matrix $\boldsymbol{\mu}$is called the *effective reciprocal inertia tensor*. If all q_s were zero (rigid molecule) the Eckart frame would coincide with a principal axes frame and $\boldsymbol{\mu}$would be diagonal, with the equilibrium reciprocal moments of inertia on the diagonal. If all q_s would be zero, only the kinetic energies of translation and rigid rotation would survive.

The potential-like term *U* is the *Watson term*:

$$U = -\frac{1}{8} \sum_{\alpha=1}^{3} \mu_{\alpha\alpha}$$

proportional to the trace of the effective reciprocal inertia tensor.

The fourth term in the Watson Hamiltonian is the kinetic energy associated with the vibrations of the atoms (nuclei) expressed in normal coordinates q_s, which as stated above, are given in terms of nuclear displacements $\rho_{i\alpha}$ by

$$q_s = \sum_{i=1}^{N} \sum_{\alpha=1}^{3} Q_{s,i\alpha} \rho_{i\alpha} \quad \text{for} \quad s = 1, \ldots, 3N-6.$$

Finally V is the unexpanded potential energy by definition depending on internal coordinates only. In the harmonic approximation it takes the form

$$V \approx \frac{1}{2} \sum_{s=1}^{3N-6} f_s q_s^2.$$

Chapter 9

Molecular Dynamics

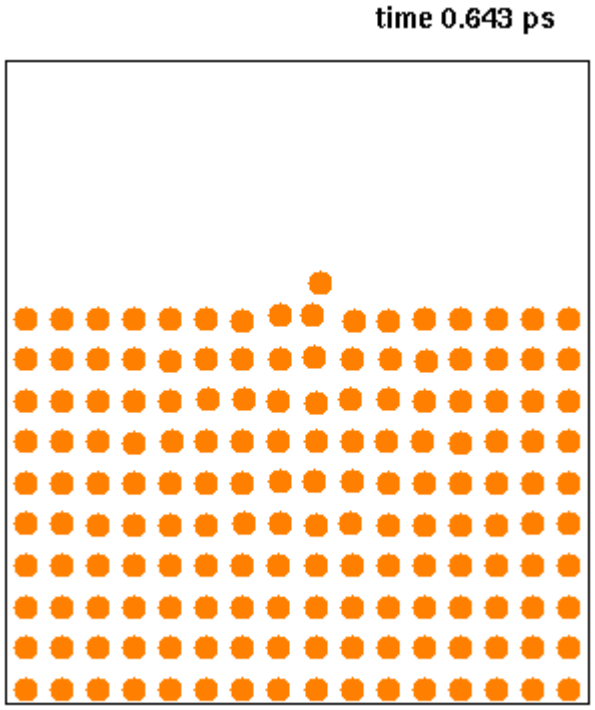
Molecular dynamics (**MD**) is computer simulation of physical movements by atoms and molecules.

Molecular dynamics simulation is frequently used in the study of proteins and biomolecules, as well as in materials science. It is tempting, though not entirely accurate, to describe the technique as a "virtual microscope" with high temporal and spatial resolution. Whereas it is possible to take "still snapshots" of crystal structures and probe features of the motion of molecules through NMR, no current experimental technique allows access to all the time scales of motion with atomic resolution. Richard Feynman once said that "If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms." Molecular dynamics lets scientists peer into the motion of individual atoms in a way which is not possible in laboratory experiments.

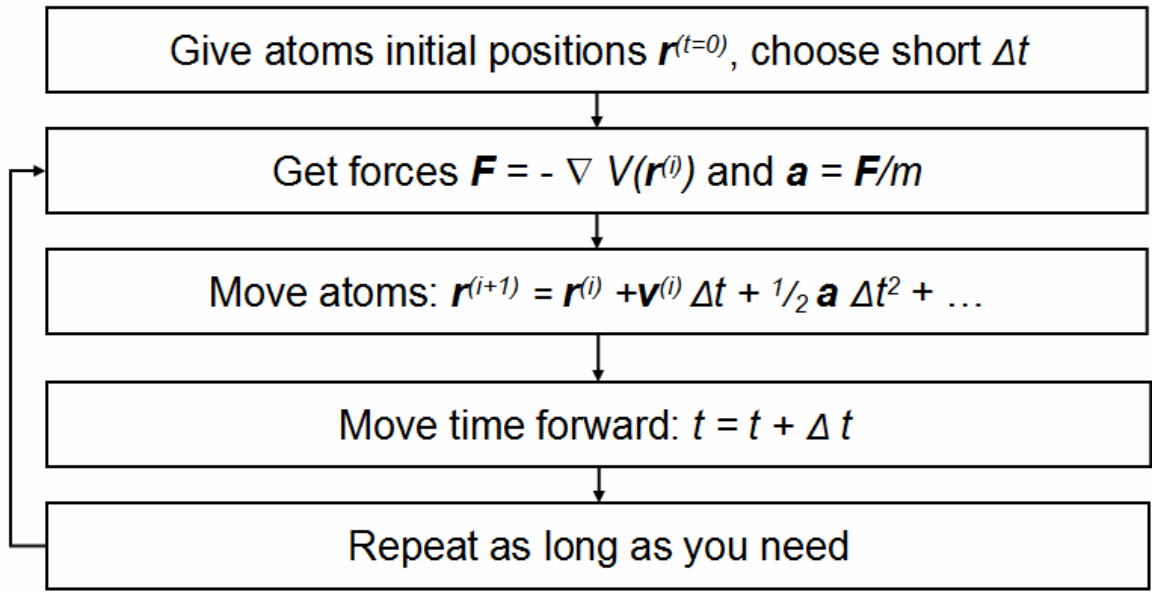
Molecular dynamics is a specialized discipline of molecular modeling and computer simulation based on statistical mechanics; the main justification of the MD method is that statistical ensemble averages are equal to time averages of the system, known as the ergodic hypothesis. MD has also been termed "statistical mechanics by numbers" and "Laplace's vision of Newtonian mechanics" of predicting the future by animating nature's forces and allowing insight into molecular motion on an atomic scale. However, long MD simulations are mathematically ill-conditioned, generating cumulative errors in numerical integration that can be minimized with proper selection of algorithms and parameters, but not eliminated entirely. Furthermore, current potential energy functions (also called force-fields) are, in many cases, not sufficiently accurate to reproduce the dynamics of molecular systems, so the much more computationally demanding Ab Initio Molecular Dynamics method must be used. Nevertheless, molecular dynamics techniques allow detailed time and space resolution into representative behavior in phase space for carefully selected systems.

Before it became possible to simulate molecular dynamics with computers, some undertook the hard work of trying it with physical models such as macroscopic spheres. The idea was to arrange them to replicate the properties of a liquid. J.D. Bernal said, in 1962: "... I took a number of rubber balls and stuck them together with rods of a selection of different lengths ranging from 2.75 to 4 inches. I tried to do this in the first place as casually as possible, working in my own office, being interrupted every five minutes or so and not remembering what I had done before the interruption." Fortunately, now computers keep track of bonds during a simulation.

Because molecular systems generally consist of a vast number of particles, it is in general impossible to find the properties of such complex systems analytically. When the number of particles interacting is higher than two, the result is chaotic motion. MD simulation circumvents the analytical intractability by using numerical methods. It represents an interface between laboratory experiments and theory, and can be understood as a "virtual experiment". MD probes the relationship between molecular structure, movement and function. Molecular dynamics is a multidisciplinary method. Its laws and theories stem from mathematics, physics, and chemistry, and it employs algorithms from computer science and information theory. It was originally conceived within theoretical physics in the late 1950s and early 1960s , but is applied today mostly in materials science and the modeling of biomolecules.



Example of a molecular dynamics simulation in a simple system: deposition of a single Cu atom on a Cu (001) surface. Each circle illustrates the position of a single atom; note that the actual atomic interactions used in current simulations are more complex than those of 2-dimensional hard spheres.



Highly simplified description of the molecular dynamics simulation algorithm. The simulation proceeds iteratively by alternatively calculating forces and solving the equations of motion based on the accelerations obtained from the new forces. In practise, almost all MD codes use much more complicated versions of the algorithm, including two steps (predictor and corrector) in solving the equations of motion and many additional steps for e.g. temperature and pressure control, analysis and output.

Areas of Application

There is a significant difference between the focus and methods used by chemists and physicists, and this is reflected in differences in the jargon used by the different fields. In chemistry and biophysics, the interaction between the particles is either described by a "force field" (**classical MD**), a quantum chemical model, or a mix between the two. These terms are not used in physics, where the interactions are usually described by the name of the theory or approximation being used and called the potential energy, or just the "potential".

Beginning in theoretical physics, the method of MD gained popularity in materials science and since the 1970s also in biochemistry and biophysics. In chemistry, MD serves as an important tool in protein structure determination and refinement using experimental tools such as X-ray crystallography and NMR. It has also been applied with limited success as a method of refining protein structure predictions. In physics, MD is used to examine the dynamics of atomic-level phenomena that cannot be observed directly, such as thin film growth and ion-subplantation. It is also used to examine the physical properties of nanotechnological devices that have not or cannot yet be created.

In applied mathematics and theoretical physics, molecular dynamics is a part of the research realm of dynamical systems, ergodic theory and statistical mechanics in general. The concepts of energy conservation and molecular entropy come from thermodynamics.

Some techniques to calculate conformational entropy such as principal components analysis come from information theory. Mathematical techniques such as the transfer operator become applicable when MD is seen as a Markov chain. Also, there is a large community of mathematicians working on volume preserving, symplectic integrators for more computationally efficient MD simulations.

MD can also be seen as a special case of the discrete element method (DEM) in which the particles have spherical shape (e.g. with the size of their van der Waals radii.) Some authors in the DEM community employ the term MD rather loosely, even when their simulations do not model actual molecules.

Design Constraints

Design of a molecular dynamics simulation should account for the available computational power. Simulation size (n=number of particles), timestep and total time duration must be selected so that the calculation can finish within a reasonable time period. However, the simulations should be long enough to be relevant to the time scales of the natural processes being studied. To make statistically valid conclusions from the simulations, the time span simulated should match the kinetics of the natural process. Otherwise, it is analogous to making conclusions about how a human walks from less than one footstep. Most scientific publications about the dynamics of proteins and DNA use data from simulations spanning nanoseconds (1E-9 s) to microseconds (1E-6 s). To obtain these simulations, several CPU-days to CPU-years are needed. Parallel algorithms allow the load to be distributed among CPUs; an example is the spatial or force decomposition algorithm .

During a classical MD simulation, the most CPU intensive task is the evaluation of the potential (force field) as a function of the particles' internal coordinates. Within that energy evaluation, the most expensive one is the non-bonded or non-covalent part. In Big O notation, common molecular dynamics simulations scale by $O(n^2)$ if all pair-wise electrostatic and van der Waals interactions must be accounted for explicitly. This computational cost can be reduced by employing electrostatics methods such as Particle Mesh Ewald ($O(n\log(n))$), P3M or good spherical cutoff techniques ($O(n)$).

Another factor that impacts total CPU time required by a simulation is the size of the integration timestep. This is the time length between evaluations of the potential. The timestep must be chosen small enough to avoid discretization errors (i.e. smaller than the fastest vibrational frequency in the system). Typical timesteps for classical MD are in the order of 1 femtosecond (1E-15 s). This value may be extended by using algorithms such as SHAKE, which fix the vibrations of the fastest atoms (e.g. hydrogens) into place. Multiple time scale methods have also been developed, which allow for extended times between updates of slower long-range forces.

For simulating molecules in a solvent, a choice should be made between explicit solvent and implicit solvent. Explicit solvent particles (such as the TIP3P, SPC/E and SPC-f water models) must be calculated expensively by the force field, while implicit solvents

use a mean-field approach. Using an explicit solvent is computationally expensive, requiring inclusion of roughly ten times more particles in the simulation. But the granularity and viscosity of explicit solvent is essential to reproduce certain properties of the solute molecules. This is especially important to reproduce kinetics.

In all kinds of molecular dynamics simulations, the simulation box size must be large enough to avoid boundary condition artifacts. Boundary conditions are often treated by choosing fixed values at the edges (which may cause artifacts), or by employing periodic boundary conditions in which one side of the simulation loops back to the opposite side, mimicking a bulk phase.

Microcanonical ensemble (NVE)

In the **microcanonical**, or **NVE** ensemble, the system is isolated from changes in moles (N), volume (V) and energy (E). It corresponds to an adiabatic process with no heat exchange. A microcanonical molecular dynamics trajectory may be seen as an exchange of potential and kinetic energy, with total energy being conserved. For a system of N particles with coordinates X and velocities V, the following pair of first order differential equations may be written in Newton's notation as

$$F(X) = -\nabla U(X) = M\dot{V}(t)$$
$$V(t) = \dot{X}(t).$$

The potential energy function $U(X)$ of the system is a function of the particle coordinates X. It is referred to simply as the "potential" in Physics, or the "force field" in Chemistry. The first equation comes from Newton's laws; the force F acting on each particle in the system can be calculated as the negative gradient of $U(X)$.

For every timestep, each particle's position X and velocity V may be integrated with a symplectic method such as Verlet. The time evolution of X and V is called a trajectory. Given the initial positions (e.g. from theoretical knowledge) and velocities (e.g. randomized Gaussian), we can calculate all future (or past) positions and velocities.

One frequent source of confusion is the meaning of temperature in MD. Commonly we have experience with macroscopic temperatures, which involve a huge number of particles. But temperature is a statistical quantity. If there is a large enough number of atoms, statistical temperature can be estimated from the *instantaneous temperature*, which is found by equating the kinetic energy of the system to $nk_BT/2$ where n is the number of degrees of freedom of the system.

A temperature-related phenomenon arises due to the small number of atoms that are used in MD simulations. For example, consider simulating the growth of a copper film starting with a substrate containing 500 atoms and a deposition energy of 100 eV. In the real world, the 100 eV from the deposited atom would rapidly be transported through and shared among a large number of atoms (10^{10} or more) with no big change in temperature. When there are only 500 atoms, however, the substrate is almost immediately vaporized

by the deposition. Something similar happens in biophysical simulations. The temperature of the system in NVE is naturally raised when macromolecules such as proteins undergo exothermic conformational changes and binding.

Canonical ensemble (NVT)

In the canonical ensemble, moles (N), volume (V) and temperature (T) are conserved. It is also sometimes called constant temperature molecular dynamics (CTMD). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat.

A variety of thermostat methods is available to add and remove energy from the boundaries of an MD system in a more or less realistic way, approximating the canonical ensemble. Popular techniques to control temperature include velocity rescaling, the Nosé-Hoover thermostat, Nosé-Hoover chains, the Berendsen thermostat and Langevin dynamics. Note that the Berendsen thermostat might introduce the flying ice cube effect, which leads to unphysical translations and rotations of the simulated system.

It is not trivial to obtain a canonical distribution of conformations and velocities using these algorithms. How this depends on system size, thermostat choice, thermostat parameters, time step and integrator is the subject of many articles in the field.

Isothermal-Isobaric (NPT) ensemble

In the isothermal-isobaric ensemble, moles (N), pressure (P) and temperature (T) are conserved. In addition to a thermostat, a barostat is needed. It corresponds most closely to laboratory conditions with a flask open to ambient temperature and pressure.

In the simulation of biological membranes, isotropic pressure control is not appropriate. For lipid bilayers, pressure control occurs under constant membrane area (NPAT) or constant surface tension "gamma" (NPγT).

Generalized ensembles

The replica exchange method is a generalized ensemble. It was originally created to deal with the slow dynamics of disordered spin systems. It is also called parallel tempering. The replica exchange MD (REMD) formulation tries to overcome the multiple-minima problem by exchanging the temperature of non-interacting replicas of the system running at several temperatures.

Potentials in MD simulations

A molecular dynamics simulation requires the definition of a potential function, or a description of the terms by which the particles in the simulation will interact. In chemistry and biology this is usually referred to as a force field. Potentials may be defined at many levels of physical accuracy; those most commonly used in chemistry are based on molecular mechanics and embody a classical treatment of particle-particle

interactions that can reproduce structural and conformational changes but usually cannot reproduce chemical reactions.

The reduction from a fully quantum description to a classical potential entails two main approximations. The first one is the Born-Oppenheimer approximation, which states that the dynamics of electrons is so fast that they can be considered to react instantaneously to the motion of their nuclei. As a consequence, they may be treated separately. The second one treats the nuclei, which are much heavier than electrons, as point particles that follow classical Newtonian dynamics. In classical molecular dynamics the effect of the electrons is approximated as a single potential energy surface, usually representing the ground state.

When finer levels of detail are required, potentials based on quantum mechanics are used; some techniques attempt to create hybrid classical/quantum potentials where the bulk of the system is treated classically but a small region is treated as a quantum system, usually undergoing a chemical transformation.

Empirical potentials

Empirical potentials used in chemistry are frequently called force fields, while those used in materials physics are called just empirical or analytical potentials.

Most force fields in chemistry are empirical and consist of a summation of bonded forces associated with chemical bonds, bond angles, and bond dihedrals, and non-bonded forces associated with van der Waals forces and electrostatic charge. Empirical potentials represent quantum-mechanical effects in a limited way through ad-hoc functional approximations. These potentials contain free parameters such as atomic charge, van der Waals parameters reflecting estimates of atomic radius, and equilibrium bond length, angle, and dihedral; these are obtained by fitting against detailed electronic calculations (quantum chemical simulations) or experimental physical properties such as elastic constants, lattice parameters and spectroscopic measurements.

Because of the non-local nature of non-bonded interactions, they involve at least weak interactions between all particles in the system. Its calculation is normally the bottleneck in the speed of MD simulations. To lower the computational cost, force fields employ numerical approximations such as shifted cutoff radii, reaction field algorithms, particle mesh Ewald summation, or the newer Particle-Particle Particle Mesh (P3M).

Chemistry force fields commonly employ preset bonding arrangements (an exception being *ab-initio* dynamics), and thus are unable to model the process of chemical bond breaking and reactions explicitly. On the other hand, many of the potentials used in physics, such as those based on the bond order formalism can describe several different coordinations of a system and bond breaking. Examples of such potentials include the Brenner potential for hydrocarbons and its further developments for the C-Si-H and C-O-H systems. The ReaxFF potential can be considered a fully reactive hybrid between bond order potentials and chemistry force fields.

Pair potentials vs. many-body potentials

The potential functions representing the non-bonded energy are formulated as a sum over interactions between the particles of the system. The simplest choice, employed in many popular force fields, is the "pair potential", in which the total potential energy can be calculated from the sum of energy contributions between pairs of atoms. An example of such a pair potential is the non-bonded Lennard-Jones potential (also known as the 6-12 potential), used for calculating van der Waals forces.

$$U(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$

Another example is the Born (ionic) model of the ionic lattice. The first term in the next equation is Coulomb's law for a pair of ions, the second term is the short-range repulsion explained by Pauli's exclusion principle and the final term is the dispersion interaction term. Usually, a simulation only includes the dipolar term, although sometimes the quadrupolar term is included as well.

$$U_{ij}(r_{ij}) = \sum \frac{z_i z_j}{4\pi\epsilon_0} \frac{1}{r_{ij}} + \sum A_l \exp \frac{-r_{ij}}{p_l} + \sum C_l r_{ij}^{-n_j} + \cdots$$

In many-body potentials, the potential energy includes the effects of three or more particles interacting with each other. In simulations with pairwise potentials, global interactions in the system also exist, but they occur only through pairwise terms. In many-body potentials, the potential energy cannot be found by a sum over pairs of atoms, as these interactions are calculated explicitly as a combination of higher-order terms. In the statistical view, the dependency between the variables cannot in general be expressed using only pairwise products of the degrees of freedom. For example, the Tersoff potential, which was originally used to simulate carbon, silicon and germanium and has since been used for a wide range of other materials, involves a sum over groups of three atoms, with the angles between the atoms being an important factor in the potential. Other examples are the embedded-atom method (EAM) and the Tight-Binding Second Moment Approximation (TBSMA) potentials, where the electron density of states in the region of an atom is calculated from a sum of contributions from surrounding atoms, and the potential energy contribution is then a function of this sum.

Semi-empirical potentials

Semi-empirical potentials make use of the matrix representation from quantum mechanics. However, the values of the matrix elements are found through empirical formulae that estimate the degree of overlap of specific atomic orbitals. The matrix is then diagonalized to determine the occupancy of the different atomic orbitals, and empirical formulae are used once again to determine the energy contributions of the orbitals.

There are a wide variety of semi-empirical potentials, known as tight-binding potentials, which vary according to the atoms being modeled.

Polarizable potentials

Most classical force fields implicitly include the effect of polarizability, e.g. by scaling up the partial charges obtained from quantum chemical calculations. These partial charges are stationary with respect to the mass of the atom. But molecular dynamics simulations can explicitly model polarizability with the introduction of induced dipoles through different methods, such as Drude particles or fluctuating charges. This allows for a dynamic redistribution of charge between atoms which responds to the local chemical environment.

For many years, polarizable MD simulations have been touted as the next generation. For homogenous liquids such as water, increased accuracy has been achieved through the inclusion of polarizability. Some promising results have also been achieved for proteins. However, it is still uncertain how to best approximate polarizability in a simulation.

Ab-initio methods

In classical molecular dynamics, a single potential energy surface (usually the ground state) is represented in the force field. This is a consequence of the Born-Oppenheimer approximation. In excited states, chemical reactions or a more accurate representation is needed, electronic behavior can be obtained from first principles by using a quantum mechanical method, such as Density Functional Theory. This is known as Ab Initio Molecular Dynamics (AIMD). Due to the cost of treating the electronic degrees of freedom, the computational cost of this simulations is much higher than classical molecular dynamics. This implies that AIMD is limited to smaller systems and shorter periods of time.

Ab-initio quantum-mechanical methods may be used to calculate the potential energy of a system on the fly, as needed for conformations in a trajectory. This calculation is usually made in the close neighborhood of the reaction coordinate. Although various approximations may be used, these are based on theoretical considerations, not on empirical fitting. *Ab-Initio* calculations produce a vast amount of information that is not available from empirical methods, such as density of electronic states or other electronic properties. A significant advantage of using *ab-initio* methods is the ability to study reactions that involve breaking or formation of covalent bonds, which correspond to multiple electronic states.

A popular software for *ab-initio* molecular dynamics is the Car-Parrinello Molecular Dynamics (CPMD) package based on the density functional theory.

Hybrid QM/MM

QM (quantum-mechanical) methods are very powerful. However, they are computationally expensive, while the MM (classical or molecular mechanics) methods are fast but suffer from several limitations (require extensive parameterization; energy estimates obtained are not very accurate; cannot be used to simulate reactions where covalent bonds are broken/formed; and are limited in their abilities for providing accurate details regarding the chemical environment). A new class of method has emerged that combines the good points of QM (accuracy) and MM (speed) calculations. These methods are known as mixed or hybrid quantum-mechanical and molecular mechanics methods (hybrid QM/MM). The methodology for such techniques was introduced by Warshel and coworkers. In the recent years have been pioneered by several groups including: Arieh Warshel (University of Southern California), Weitao Yang (Duke University), Sharon Hammes-Schiffer (The Pennsylvania State University), Donald Truhlar and Jiali Gao (University of Minnesota) and Kenneth Merz (University of Florida).

The most important advantage of hybrid QM/MM methods is the speed. The cost of doing classical molecular dynamics (MM) in the most straightforward case scales $O(n^2)$, where n is the number of atoms in the system. This is mainly due to electrostatic interactions term (every particle interacts with every other particle). However, use of cutoff radius, periodic pair-list updates and more recently the variations of the particle-mesh Ewald's (PME) method has reduced this between O(n) to $O(n^2)$. In other words, if a system with twice as many atoms is simulated then it would take between two to four times as much computing power. On the other hand the simplest *ab-initio* calculations typically scale $O(n^3)$ or worse (Restricted Hartree-Fock calculations have been suggested to scale $\sim O(n^{2.7})$). To overcome the limitation, a small part of the system is treated quantum-mechanically (typically active-site of an enzyme) and the remaining system is treated classically.

In more sophisticated implementations, QM/MM methods exist to treat both light nuclei susceptible to quantum effects (such as hydrogens) and electronic states. This allows generation of hydrogen wave-functions (similar to electronic wave-functions). This methodology has been useful in investigating phenomena such as hydrogen tunneling. One example where QM/MM methods have provided new discoveries is the calculation of hydride transfer in the enzyme liver alcohol dehydrogenase. In this case, tunneling is important for the hydrogen, as it determines the reaction rate.

Coarse-graining and reduced representations

At the other end of the detail scale are coarse-grained and lattice models. Instead of explicitly representing every atom of the system, one uses "pseudo-atoms" to represent groups of atoms. MD simulations on very large systems may require such large computer resources that they cannot easily be studied by traditional all-atom methods. Similarly, simulations of processes on long timescales (beyond about 1 microsecond) are prohibitively expensive, because they require so many timesteps. In these cases, one can

sometimes tackle the problem by using reduced representations, which are also called coarse-grained models.

Examples for coarse graining (CG) methods are discontinuous molecular dynamics (CG-DMD) and Go-models. Coarse-graining is done sometimes taking larger pseudo-atoms. Such united atom approximations have been used in MD simulations of biological membranes. The aliphatic tails of lipids are represented by a few pseudo-atoms by gathering 2 to 4 methylene groups into each pseudo-atom.

The parameterization of these very coarse-grained models must be done empirically, by matching the behavior of the model to appropriate experimental data or all-atom simulations. Ideally, these parameters should account for both enthalpic and entropic contributions to free energy in an implicit way. When coarse-graining is done at higher levels, the accuracy of the dynamic description may be less reliable. But very coarse-grained models have been used successfully to examine a wide range of questions in structural biology.

Examples of applications of coarse-graining in biophysics:

- protein folding studies are often carried out using a single (or a few) pseudo-atoms per amino acid;
- DNA supercoiling has been investigated using 1-3 pseudo-atoms per basepair, and at even lower resolution;
- Packaging of double-helical DNA into bacteriophage has been investigated with models where one pseudo-atom represents one turn (about 10 basepairs) of the double helix;
- RNA structure in the ribosome and other large systems has been modeled with one pseudo-atom per nucleotide.

The simplest form of coarse-graining is the "united atom" (sometimes called "extended atom") and was used in most early MD simulations of proteins, lipids and nucleic acids. For example, instead of treating all four atoms of a CH_3 methyl group explicitly (or all three atoms of CH_2 methylene group), one represents the whole group with a single pseudo-atom. This pseudo-atom must, of course, be properly parameterized so that its van der Waals interactions with other groups have the proper distance-dependence. Similar considerations apply to the bonds, angles, and torsions in which the pseudo-atom participates. In this kind of united atom representation, one typically eliminates all explicit hydrogen atoms except those that have the capability to participate in hydrogen bonds ("polar hydrogens"). An example of this is the Charmm 19 force-field.

The polar hydrogens are usually retained in the model, because proper treatment of hydrogen bonds requires a reasonably accurate description of the directionality and the electrostatic interactions between the donor and acceptor groups. A hydroxyl group, for example, can be both a hydrogen bond donor and a hydrogen bond acceptor, and it would be impossible to treat this with a single OH pseudo-atom. Note that about half the atoms

in a protein or nucleic acid are nonpolar hydrogens, so the use of united atoms can provide a substantial savings in computer time.

Examples of applications

Molecular dynamics is used in many fields of science.

- First macromolecular MD simulation published (1977, Size: 500 atoms, Simulation Time: 9.2 ps=0.0092 ns, Program: CHARMM precursor) Protein: Bovine Pancreatic Trypsine Inhibitor. This is one of the best studied proteins in terms of folding and kinetics. Its simulation published in Nature magazine paved the way for understanding protein motion as essential in function and not just accessory.

- MD is the standard method to treat collision cascades in the heat spike regime, i.e. the effects that energetic neutron and ion irradiation have on solids an solid surfaces.

The following two biophysical examples are not run-of-the-mill MD simulations. They illustrate notable efforts to produce simulations of a system of very large size (a complete virus) and very long simulation times (500 microseconds):

- MD simulation of the complete satellite tobacco mosaic virus (**STMV**) (2006, Size: 1 million atoms, Simulation time: 50 ns, program: NAMD) This virus is a small, icosahedral plant virus which worsens the symptoms of infection by Tobacco Mosaic Virus (TMV). Molecular dynamics simulations were used to probe the mechanisms of viral assembly. The entire STMV particle consists of 60 identical copies of a single protein that make up the viral capsid (coating), and a 1063 nucleotide single stranded RNA genome. One key finding is that the capsid is very unstable when there is no RNA inside. The simulation would take a single 2006 desktop computer around 35 years to complete. It was thus done in many processors in parallel with continuous communication between them.

- Folding Simulations of the Villin Headpiece in All-Atom Detail (2006, Size: 20,000 atoms; Simulation time: 500 μs = 500,000 ns, Program: folding@home) This simulation was run in 200,000 CPU's of participating personal computers around the world. These computers had the folding@home program installed, a large-scale distributed computing effort coordinated by Vijay Pande at Stanford University. The kinetic properties of the Villin Headpiece protein were probed by using many independent, short trajectories run by CPU's without continuous real-time communication. One technique employed was the Pfold value analysis, which measures the probability of folding before unfolding of a specific starting conformation. Pfold gives information about transition state structures and an ordering of conformations along the folding pathway. Each trajectory in a Pfold calculation can be relatively short, but many independent trajectories are needed.