
# FACULTY OF ENGINEERING AND TECHNOLOGY

## Department of Biotechnology

# Introduction to the concepts of molecular modeling

abc

Molecular modeling is a general term that covers a wide range of molecular graphics and computational chemistry techniques used to build, display, manipulate, simulate, and analyze molecular structures and to calculate properties of these structures.

Molecular modeling is used in several different research areas, and therefore the term does not have a rigid definition. To a chemical physicist, molecular modeling might imply performing a high quality quantum mechanical calculation using a supercomputer on a structure with 4 or 5 atoms; to an organic chemist, molecular modeling might mean displaying and modifying a candidate drug molecule on a desktop computer. The criterion for a successful modeling experiment should not be how accurately the calculations are performed, but whether they are useful in rationalizing the behavior of the molecule or in enhancing the creativity of the chemist in the design of novel compounds.

This section gives a brief introduction to techniques that can be used to model molecules of interest to the biopharmaceutical industry.

XYZ

abc

computational chemistry. Molecular graphics is the core of a modeling system, providing for the visualization of molecular structures and their properties. Molecular graphics provides the ability to display structures in a variety of styles (from simple line displays to solid renderings known as CPKs) and color schemes, with visual aids such as depth-cueing, and the ability to move the structures interactively in three dimensions (3D). Simple tools for manipulating structures, such as for modifying torsion angles and calculating geometry, are frequently included under the molecular graphics banner.Visualization of molecular properties is an extremely important aspect of molecular modeling. The properties might be calculated using a computational chemistry program and visualized as 3D contours, along with the associated structures. While manipulation of structures is usually interactive, the calculation of properties may require significant computer time. Calculations are usually run in the background rather than interactively, leaving the modeling system free for interactive work. The graphics part of the

XYZ

abc

exist for building and modifying structures.Small molecule structures can be built in 3D by joining basic building blocks from a fragment library and then modifying atom and bond types (this functionality is provided by the Builder module; see Basic Applications 1 and 2). Macromolecule structures such as proteins and nucleic acids, consisting of large numbers of specific units, can be built by specifying the sequence of the units and the conformation in which the units should be joined. When necessary, individual units (amino acids or nucleotides) may be modified or replaced in the same manner that small molecule structures are modified (see the Biopolymer module documentation).

An alternative to building in 3D is to sketch a structure freehand in 2D and then convert the sketch into 3D. This approach has the advantage of speed and the ease with which complex structures can be built; only element types, connectivity, and (where necessary) stereochemistry need to be defined. This capability is provided by the **MolBuilder** toolbox of the Builder module,

# XYZ

abc

abc

the energy of the structure is brought to a minimum. The structure with the lowest energy is considered to have the most stable arrangement and, by definition, the optimum geometry. Minimization generally results in a modeled structure with a close resemblance to a real physical structure. The ability to compute the energy of a structure is a necessary part of the minimization process and is an extremely important aspect of a modeling system.Molecular mechanics techniques take a classical approach to calculating the energy of a structure. The molecule is treated essentially as a set of charged point masses (the atoms) which are coupled together with springs. The total energy of a structure is calculated using an analytical function that sums the individual energy terms. At its simplest level, the function includes bond stretching, valence angle bending, torsion, and nonbond interaction terms, which associate an energetic penalty to the structure based upon deviations from an idealized geometry. For instance, the bond term is a summation over all bonds in the structure, in which the energy of each bond is

# XYZ

abc

abc

can be either attractive or repulsive depending on the sign of the charges involved. The number of nonbonded atom pairs rises with the square of the number of atoms, and therefore the nonbond interaction term dominates the computer time required for energy evaluation in calculations on larger structures. For this reason, it is normal to limit the nonbond energy evaluation to atom pairs within a certain cutoff distance, based on the assumption that atom pairs separated by a larger distance make a negligible contribution to the total energy.

The idealized reference values of bond lengths and angles, plus the force constants and the van der Waals radii and associated constants required to calculate the nonbond interactions, are stored in a file that is referred to when the energy calculation is run. The combination of these parameters with the functional forms of the individual energy terms is known as a *forcefield*. It is important to appreciate that the reference values of bonds and angles cannot be based on element type alone. For example, a carbon-carbon single bond is longer

# XYZ

abc

XYZ

abc

Molecular mechanics enables the energy of a structure to be evaluated quickly and may be applied to structures of a size up to and including large proteins. Energy calculations have a range of applications in molecular modeling. They can be used in conformational analysis to evaluate the relative stability of different conformers (see below) and to predict the equilibrium geometry of a structure. They can also be used to evaluate the energy of two or more interacting molecules, such as when docking a substrate into an enzyme active site (Basic Application 3 illustrates docking in the formic acid dimer using the Docking module.)

As described above, molecular mechanics energy calculations are an integral part of energy minimizations. From the energy functions it is possible to evaluate the forces acting on the atoms. Minimization uses information on the atomic forces to adjust atomic coordinates in an iterative manner to bring the structure to a minimum energy conformation.

# XYZ

abc

abc

motion.Starting with the molecular mechanics energy description of the structure as described above, the forces acting on the atoms can be evaluated. Since the masses of the atoms are known, Newton's second law of motion (force = mass x acceleration) may be used to compute the accelerations, and thus the velocities, of the atoms. The accelerations and velocities may then be used to calculate new positions for the atoms over a short time step (around 1 femtosecond, where 1 fs is equal to $10^{-15}$ s), thus moving each atom to a new position in space. This process iterates many thousands of times, generating a series of conformations of the structure known as a trajectory. A simulation is frequently run for many tens of picoseconds (1 ps is equal to $10^{-12}$ s).

The velocities of the atoms are related directly to the temperature at which the simulation is run. A simulation run at 300 K provides information on structural fluctuations that occur around the starting conformation, perhaps to illustrate which parts of a molecule are most flexible, and also can provide

# XYZ

abc

XYZ

abc

are separated by a certain distance. This information can be added to a simulation in the form of *restraints*. During a simulation a force can be applied to an atom pair to restrain it to the specified separation. Restraints are important because they provide a certain amount of control over a simulation. (Basic Applications 10 and 11 illustrate the use of restraints in a dynamics simulation.)

Minimizations and molecular dynamics simulations may be performed in Insight II using the Discover modules. (For more information, refer to the Forcefield-Based Simulations documentation.)

Using molecular dynamics simulations it is easy to generate a vast amount of data. The problem is in analyzing the data. The simulation is usually analyzed first in a *qualitative* way by replaying the simulation as a movie, a process known as animation. A simulation can be analyzed *quantitatively* by defining properties of interest and then graphing those properties against each other. For example, graphs that illustrate how the geometry or energy of the structure varies

# XYZ

abc

**Quantum Mechanics**

The electronic structure of a molecule is of prime importance in determining its properties. When a drug molecule and receptor interact, each "sees" the other as a blob of electron density that is held together by the positive charges of the atomic nuclei. Although molecular mechanics calculations are extremely useful, they consider essentially only the position of the nuclei and therefore cannot fully represent chemical reality. Molecular mechanics provides no information on electronic structure and, furthermore, cannot be used when the molecule is not in its ground state or when covalent bonds are being broken or formed. Some properties can be derived only with quantum mechanics calculations.Starting with a specified nuclear geometry, quantum mechanics calculations solve the Schrödinger equation for this arrangement of electrons and nuclei. This yields both the energy of the molecule and the associated wave function from which electronic properties, such as electron density, can be calculated.

The energy of a structure calculated via quantum mechanics can be used in conformational searches, in the same way that the molecular mechanics energy is used. Quantum mechanics calculations can also be used for energy minimization. However, quantum mechanics calculations typically consume a far greater amount of computer resource than molecular mechanics calculations and are therefore generally limited to small molecules, whereas molecular mechanics can be applied to structures up to the size of large proteins. Molecular mechanics and quantum mechanics should thus be viewed as complementary techniques. For instance, conformational energy calculations for a peptide are best carried out using molecular mechanics. However, molecular mechanics is generally ineffective for handling conjugated systems, while quantum mechanics, in calculating electronic structure, takes account of conjugation automatically and is therefore recommended for optimizing the structure of a small molecule containing conjugated systems.

The wave function can be used to calculate a range of chemical properties, which can be used in structure-activity studies. These include electrostatic potential, electron density, dipole moment, and the energies and positions of frontier orbitals. As with the analysis of a molecular dynamics calculation, molecular graphics is essential for visualizing these properties. Quantum mechanics calculations are also used frequently to derive atom-centered partial charges (although the term *charge* itself does not have a strict quantum mechanical definition). Charges have a wide range of applications in modeling and are used in the calculation of electrostatic energies in molecular mechanics calculations and in computing electrostatic potentials.

# MCQs

The most widely used quantum mechanics packages are the public-domain programs AMPAC and MOPAC, to which an interface is provided in the Ampac/Mopac module. These programs utilize the molecular orbital formalism in solving the Schrödinger equation, and are therefore known as molecular orbital programs

RAMA
UNIVERSITY
UTTAR PRADESH
KANPUR

RAMA
UNIVERSITY
UTTAR PRADESH
KANPUR

RAMA
UNIVERSITY
UTTAR PRADESH
KANPUR