ANNALS OF PHARMACY AND PHARMACEUTICAL SCIENCES
Volume 4 | Issue 1&2 | April & October, 2013 | 26-35

A REVIEW

Molecular mechanics and dynamics

HEMANT U. CHIKHALE, POORVASHREE P. JOSHI, AMIT G. NERKAR AND SANJAY D. SAWANT

ABSTRACT

Computational chemistry is a branch of chemistry that uses principles of computer science to assist in solving chemical problems. It uses the results of theoretical chemistry, incorporated into efficient computer programs, to calculate the structures and properties of molecules and solids. It is widely used in the design of new drugs and materials. Because of the complexity of biological systems, computer methods have become increasingly important in the life sciences. With faster and more powerful computers larger and more complex systems may be explored using computer modelling or computer simulations. Various simulation techniques such as molecular mechanics (MM) and molecular dynamics (MD), Monte Carlo, and Brownian dynamics, as well as hybrids of these methods, have emerged as simplifications of the exact quantum mechanical description for large molecules.

Key words : Molecular mechanics (MM), Molecular dynamics (MD), Monte carlo (MC)

How to cite this paper : Chikhale, Hemant U., Joshi, Poorvashree P., Nerkar, Amit G. and Sawant, Sanjay D. (2013). Molecular mechanics and dynamics. *Ann. Pharm. & Pharm. Sci.*, **4** (1&2) : 26-35.

Article chronicle : Received : *15.09.2013;* **Accepted :** *27.09.2013*

INTRODUCTION

A model of the real world is constructed for both measurable and unmeasurable properties are computed, and are compared with experimentally determined properties. This comparison validates or invalidates the model that is used. In the former case the model may be used to study relationships between model parameters and assumptions or to predict unknown or unmeasurable quantities. Since chemistry concerns the study of properties of substances or molecular systems in terms of atoms, the basic challenge facing computational chemistry is to describe or even predict.

– The structure and stability of a molecular system,

– The (free) energy of different states of a molecular system,

– Reaction processes within molecular systems. (Wilfred, EVG, *et al.*[1]). Molecular mechanics (force field) calculation is the most commonly used type of calculation in computational medicinal chemistry, and a large number of different force fields have been developed over the years Molecular mechanics is the application of classical mechanics to molecules. All of the equations and associated parameters used to calculate each energy term are collectively called the “force field”. Different force fields have been developed for different molecular types (e.g. small organic molecules vs. large biomolecules). Since molecular mechanics does not deal directly with electrons and orbitals, we cannot study chemical reactions or predict the reactivity of the molecules studied with this technique.

$$\dot{g}_{steric} Z \dot{g}_{bonds} K_r(r J r_{eq})^2 H \dot{g}_{angles} K^{-}(^{-} J ^{-}_{eq})^2 H \dot{g}_{dihedrals} \frac{V_n}{2}[1 H \cos(\text{¦} J \text{¥})] \quad (1)$$

MEMBERS OF THE RESEARCH FORUM

Address for correspondence :
HEMANT U. CHIKHALE, Post Graduate Department of Pharmaceutical and Medicinal Chemistry, STES’s Smt. Kashibai Navale College of Pharmacy, Kondhwa, PUNE (M.S.) INDIA

Coopted auhors :
POORVASHREE P. JOSHI, AMIT G. NERKAR AND SANJAY D. SAWANT, Post Graduate Department of Pharmaceutical and Medicinal Chemistry, STES’s Smt. Kashibai Navale College of Pharmacy, Kondhwa, PUNE (M.S.) INDIA

Applications :

In computational terms, molecular mechanics is the least expensive (fastest) method. It is especially well suited for providing excellent structural parameters in terms of bond distances, angles, etc., for the most stable conformation of a molecule. This so-called "geometry optimization" is often used as the first step before a calculation of another type is performed. This is done to insure that the molecule is in its' lowest energy state so that calculated results can be compared to those done experimentally. Since molecular mechanics is computationally inexpensive, it is often the only method available for use with large molecules, especially those of biochemical interest such as proteins. (Holtje *et al.*,[2])

Basic principles of molecular mechanics :

Empirical force-field methodology is based on classical mechanics and on the fundamental assumption that the total "steric" energy of a structure can be expressed as a sum of contributions from many interaction types. Another important assumption is that the force-field and its parameters, which have been determined from a set of molecules, are used to study other molecules. Molecular mechanics is one of the computational techniques used by chemists to study molecular systems. Basically molecular mechanics was developed to describe molecular structures and properties in as practical a manner as possible. Here practical implies that results should be accurate, available within a reasonable amount of computational time and be applicable to a large variety of systems. As a result of these requirements quantum mechanical first principle methods and semi empirical methods are not always suited (or even feasible). The ranges of applicability of molecular mechanics (force field) techniques include:

- Molecules containing thousands of atoms.
- Organics, oligonucleotides, peptides, and saccharides (metallo-organics and inorganics in some cases).
- Vacuum, implicit, or explicit solvent environments.
- Ground state only.
- Thermodynamic and kinetic properties.

Essentially molecular mechanics is an attempt to provide a reliable recipe for obtaining the potential energy surface for the movement of atoms within molecules. Molecular mechanics is characterized by high computational speed in comparison to first principle or semi empirical methods. This arises from the simple mathematical formulation - mainly simple arithmetic expressions rather than complex integral evaluations. As a result of the speed the technique allows for its use in procedures such as molecular dynamics, conformational energy searching, and docking, which require large numbers of energy evaluation.

What is a force field :

Very briefly, a force field is a mathematical function which returns the energy of a system as a function of the conformation of the system. But a better idea may be obtained by considering the situation physically. Consider a molecule as a collection of atoms held together by elastic forces. Now the forces can be written in terms of potential energy functions of various structural features such as bond lengths, bond angle, non bonded interactions etc. Hence force fields are also sometimes referred to as potentials. Thus the energy, E, of a molecule in a force field arises from the deviations from the ideal structural features and so can be written approximately as

$$E = Es + Eb + E\mu + Enb + ... \quad (2)$$

Here, E is termed as the steric energy. Es is the energy for bond stretching, Eb is the energy for bond angle bending, Eù is the torsional energy due to twisting about bonds and Enb is the energy for non bonded interactions. If there are other mechanisms affecting energy such as electrostatic repulsions or hydrogen bonding then these may be included in E by adding appropriate terms into the above expression. It should be kept in mind that E is a measure of the intra-molecular strain relative to a hypothetical situation. Thus E on its own has no physical meaning. Many different kinds of force fields have been developed over the years. Some include additional energy terms that describe other kinds of deformations. Some force-fields account for coupling between bending and stretching in adjacent bonds in order to improve the accuracy of the mechanical model.

Design of a force field :

The design of force fields for molecular mechanics is guided by the following principles:

- Nuclei and electrons are lumped into atom-like particles.
- Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).
- Interactions are based on springs and classical potentials.
- Interactions must be pre assigned to specific sets of atoms.
- Interactions determine the spatial distribution of atom-like particles and their energies.

Mechanical effect :

Stretching & bending :

Considering the idea of a molecule to be a collection of masses connected by springs. Thus by applying Hookes Law we can evaluate the energy required to stretch and bend bonds from their ideal values. Thus Es and Eb may be expressed as :

$$E_s = \sum_{i=1}^{N} \frac{K_i^s}{2}(l_i - l_i^0)^2 \quad (3)$$

$$E_b = \sum_{i<j}^{M} \frac{K_{ij}^b}{2}(\bar{}_{ij} - {}^{-0}_{ij})^2 \quad (4)$$

where, N is the total number of bonds and M is the total number of bond angles in the molecule. k_i^s and K_{ij}^b are are the force constants for stretching and bending, respectively. Li and θij are the actual bond lengths and bond angles. Finally l^0i and θ^0ij are ideal bond lengths and bond angles. Unique k_i^s and θ^0ij and values are assigned to each pair of bonded atoms based on their types (e.g. C-C, C-H, O-C, etc.). Similarly K_{ij}^b and è^{0}ij parameters for angle bending are assigned to each bonded triplet of atoms based on their types (e.g. C-C-C, C-O-C, C-C-H, etc.). The formulation above is only a first approximation. There are various factors which can be taken into account to improve the accuracy for these include noting that bond stretching requires more energy than bond bending and so for a molecule being deformed most of the distortion should occur in the bond angles rather than bond lengths. Another point to consider is that Hookes Law overestimates the energy required to achieve large distortions. Another aspect is that as a bond angle gets compressed the two associated bond lengths become longer.

Torsion :

Now we consider the form of the Eω term. The energy due to torsion is usually expressed in terms of a Fourier series,

$$E\mu = p½[V1(1+\cos\mu)+V2(1+\cos 2\mu F+V3\,(1+\cos 3\mu)+...] \quad (5)$$

Where, the sum is over all unique sequences of bonded atoms. In general the series is truncated at the third term, V_1, V_2 and V_3 being chosen so that the resultant conformation agree well with experiment for a given group of molecules.

Non bonded interactions :

The final term contributing to E is the energy from pair wise non bonded Interactions. Such interactions are modeled by London dispersive forces (for the attraction) and Van der Waals forces (for the repulsion). Some of the common potential functions implementing the above are the Lennard Jones and Buckingham potentials :

$$V_{LJ} = \frac{A}{r^{12}} - \frac{B}{r^6} \quad (6)$$

$$V_{Buck} = A' \exp\frac{B'}{r} - \frac{C}{r^6} \quad (7)$$

Usually the parameters for the non bonded energy terms are obtained by measuring non bonded contact distances in crystalline hydrocarbons, diamond, and graphite and Van der Waals contact data for rare gas atoms. Parameters for other atoms are then obtained by extrapolation or interpolation. One major assumption is that the potential derived from intermolecular interactions can accurately reproduce intermolecular interactions. In addition the interactions are considered to be pair wise additive.

Electrostatic Interactions :

An important non bonded energy term that is always taken into account is the electrostatic interactions. Typically the electrostatic interaction dominates the total energy of a system by a full magnitude. The electrostatic contribution is modeled using a Coulombic potential :

$$E_{Coul} = \sum_{i,j-1} \frac{q_i q_j}{r_{ij}} \quad (8)$$

The electrostatic energy is a function of the charge on the non-bonded atom- s, their inter atomic distance, and a molecular dielectric expression that accounts for the decrease of electrostatic interaction due to the environment (such as by solvent or the molecule itself). A linearly varying distance dependent dielectric (*i.e.* 1=r) is sometimes used to account for the increase in environmental bulk as the separation distance between interacting atoms increases. The accuracy of the electrostatic term depends on the correct assignment of charges to individual atoms. There are two main problems in charge assignment :

- Experimental data of charges for molecules at atom resolutions does not exist.
- Molecular mechanics assumes that nuclei are isolated entities.

The fact that nuclei share common electrons is not considered. Thus exact localization of electrons by considering atoms as fixed point charges is not correct Charge assignment methodology differs from force field to force field and is more art than science. Summarizing we see that the terms describing energy changes from bond length, bond angle and torsions are well understood and can be accurately included in the overall energy expression. The most influential term, the electrostatic term, however is not fully understood. Hence the variation in results from different force fields can be attributed, to a large extent, to the electrostatic term. Of course the nature of the parameterization generating the various force constants and ideal lengths and angles will also affect the applicability of a force field. In general once must consider with which group of molecules or systems a given force field has been parameterized – keeping this fact in mind one may then use the force field on an unknown but similar system. Generality is still a problem with force fields, though with the development of the Universal Force Field (UFF) an attempt has been made to develop a generalized force field applicable to a large portion of the periodic table and not be restricted to particular groupings of atoms such as proteins, nucleic acids etc. (Guha *et.al.*[3]) Molecular mechanics methods are several orders of magnitude faster than Quantum mechanics methods, and for problems where MM methods are well defined, the accuracy is good.

Chemical effect :

Molecular mechanics is a model for chemical behavior. Simple mechanical effect, such as stretching, bending, and torsion supplemented with an electrostatic scheme, is not sufficient alone to reproduce some well- known structural or spectroscopic phenomena such as the electro negativity, anomeric, and Bohmann effect. These effect and other phenomena might be classified as chemical effect.

Electronegativity effect :

An examination of chemical-bonding effect indicates that the length of a particular bond is highly dependent on the electronegativity of the attached substituents. For example, the $C_{sp3} - C_{sp3}$ in ethane has a bond length of 1.534 A^0. What happen if hydrogen is replaced by an electronegative atom? In ethyl chloride, the $C_{sp3} - C_{sp3}$ bond shrink to 1.528 A^0. In molecular orbital term, the above described bond shrinkage can be explained. When an electronegative atom, such as chlorine replaces hydrogen, there is more *P* character present on the carbon forming the C-Cl bond, which is a result of the tendency for chlorine to attract electron relative to carbon. In turn, because of there must be conservation of S and P orbital's, the hybrid orbital of the same carbon bearing the electronegative atom must have more S character when forming $C_{sp3} - C_{sp3}$ bond. This loss of P- character leads to shorter observed bond length. The effect depends on the electronegative atom itself (e.g. F, Cl, Br or other) and how many atoms are present. This is referred to as primary electronegativity effect but the effect is also observed for bonds not directly attached to the electronegative center such as the C_2-C_3 bond of propyl chloride this through-bond effect is quickly diminished, but is pronounced enough to be considered structurally important and is classified as the secondary electronegative effect. Both a primary and secondary electronegativity effect is taken into account in MM3 and MM4. The program examines the connectivity of every bond. If an electronegative atom present, the default natural bond length l is shortened by a specified amount $a l_e^{ij}$.

$$a l^{ij}_{e-primary} = a l^{ij}_{e-sub1} + 0.62\, a l^{ij}_{e-sub2} + 0.62^2\, a l^{ij}_{e-sub3} + ... \quad (9)$$

$$D l^{ij}_{e-secondary} = 0.40\, D l^{ij}_{e-primary} \quad (10)$$

The shortening depends on the electronegativity of the atom and the number of electronegative atoms included, according to equation (8), (9) for the primary and secondary electronegativity effect.

Anomeric effect :

For dimethoxymethane and related system of the type CH_3-O-CH_2-X there is an unusual conformational preference that can not be rationalized by steric argument alone. It turn out that the gauche- gauche conformation is the most stable. This conformational preference falls into a general category of stereo-electronic effect. The generalized anomeric effect, as it is sometime referred to in example, is best known in cyclic carbohydrate compounds and derivatives. Various arguments about the origin of the anomeric effect have been suggested, ranging from unfavorable dipole-dipole interaction to n-σ* MO interaction, most likely, both effects are operating simultaneously to various degrees.

Bohmann effect :

The length of C-H bonds adjacent to atoms bearing at least one lone pair of electrons (e.g. amine, alcohol, ethers, and fluorides) is a function of the orientation of the C-H bond relative to the lone pair of electrons. The difference in C-H bands has been observed as low as 2700 $cm^{-1.}$ When a C-H bond adjacent to the nitrogen in an amine, for example, is located anti to the lone pair of electrons, resonance argument similar to those used in the foregoing to des cribes the anomeric effect can be involved.

Parameterization :

All Semiemperical contains parameter. They either replace integrals that are calculated analytically or they are part of empirical formulas that describes the chemical bonding, usually in the two-center one electron part. These parametric formulas are designed to compensate for the neglect of a large part of the inter atomic, three and four center term that have to be taken in account.

Classification of parameter :

The parameter discussed here can be classified into two groups

- Experimentally derived fixed parameters.
- Adjustable parameter.

All quantum chemical methods, even if they are considered as derived from first principle method make use of at least the first group of parameters. Examples are the atomic masses and atomic heats of formation.

The experimentally derived parameters used in semiempirical methods are the orbital energies of valence and inner orbital. The adjustable parameter can be further classified into atomic and bond parameter. The most important adjustable atomic parameter are the exponent of the atomic basis function. In some methods (INDO/S MSINDO, and AM1), different values of exponents are used for the evaluation of intra- and interatomic integrals. Other methods use atomic parameters (Slater - Condon factors) for the one-center terms U or use this integral directly as adjustable parameter. For the two-center one electron integrals, all methods use atomic parameters, either called β (MINDO/3, MNDO, AM1, PM3 and SAM1) or K (MSINDO, with a slightly different functionality) Bond parameter αAB that depend on the atomic number of two

atoms appear in some semiempirical methods. They are either used in the correction term fAB of the internuclear repulsion or in the parameterized function for Hμσ. (Bowen *et al.*,[4], Bulintick *et al.*,[5])

Algorithms :

The algorithms fall into three broads categories.

- Programs that suggest a bioactive conformation given the set of matching point.
- Programs that are devoted to discovering both the matching points and the proposed bioactive conformation.
- The programs that not only determine the matching points and the proposed bioactive conformation but also performs a 3D QSAR analysis on the resulting overlays.

The one of a particular problem depends on the properties of the compounds as well as the software available.

Specialized methods to proposed a bioactive conformation given the points to match :

The active analog approach and receptor :

If one has enough SAR information to proposed the points to match, the problem becomes simply selecting the proposed bioactive conformation(s) of each compound. One approach to this problem is to perform rigid rotation of all rotatable bonds in each molecule, tabulating which distance between the atoms or points of interest are occupied by a conformation of reasonable energy. Proposed bioactive conformations thus correspond to conformations that occupy the distance bins occupied by all other molecules of the set.

Ensemble distance geometry :

Distance geometry is a conformational search strategy in which the algorithm starts with a matrix of allowed interpoint (usually all interatomic) distances. It then randomly select points within the allowed ranges and refines these distance bounds so that they obey the triangle inequality.

Ensemble distance geometry uses distance geometry with three enhancements.

- All molecules that one wishes to examine for a pharmacophores are included in one distance matrix.
- The distance matrix has been enhanced such that the lower bounds of all intermolecular distance are set to zero, thus allowing the molecules to superimpose.
- The upper bound of the distance between the matching points is also set to zero or small number. The result is that the algorithm produces a set of overlapped conformations ready for optimization.

The problem with distance geometry is that, because the distances for any solution are set at random, many duplicates of some conformations are generated. This requires intervention to get a unique set of conformation. The advantage of ensemble distance geometry over rigid rotation is that ring conformations, bond length, and bond angle are not frozen.

Boltzmann jump :

Monte Carlo sampling of conformational space is another popular method to generate conformation of interest. It generates a new structure from the previous one by making random changes to one or more torsion angles. If the new structure is lower in energy than the previous ones, it is kept. If it is higher in energy than the pervious one by ΔE then it is retained with a probability equal to the Boltzmann factor $e^{\Delta E/RT}$, where R is the gas constant and T the temperature. This procedure allows the search to overcome torsional barriers. In practice, one minimizes the conformations when they differ substantially from the privous one.

Chemometrics methods :

A principal component analysis of the distance between five proposed pharmacophore points in all low-energy conformer of active HMG-CoA reductase inhibitors reveled that only three of these distances was necessary to explain 85.5% of the variances of the 10 distances. Subsequent cluster analysis of the conformer using these three distances produced one cluster of conformation s that includes four of the five active compounds and only two of the six inactive compounds. The conformations and distances that correspond to this cluster were selected as the bioactive conformation and

Programme	No.of molecules	Conformations	Method of of search	Scorings	Ligand features	Target features
AUTOFIT[69]	2	Precalculated	Exhaustive	RMS	Functional points,user controlled	Functional points,user controlled
COMPASS[163]	2-32	Precalculated (ex.Used X-ray structure)	Clique detection of interpoints distances.	Distance tolerance	All atoms, properties assigened by the programs	Not used
ChemDBS-3D[49]	2-32	Precalculated 2,3 And 4 points distance keys	Intersection of 3 or 4-center pharmacophores distance keys	Distance tolerance	H-bond donars and acceptor , positive atoms, aromatic ring.	Not Used

pharmacophore distance respectively. The lacks of reactivity of the two inactive compounds that have the pharmacophore distance are explained by difference in electrostatics potential.

Specialized methods to detect corresponding points :

Clique detection :

The earliest methods to detect the feature that can overlap in 3D are an extension of work that clique detection methods efficiently identify the maximum common 3D substructure in a set of molecule. Clique detection algorithm use as input a matrix of interpoint distance for each 3D object, a conformation of molecule in the case of pharmacophore mapping the algorithm examine this matrix to identify those set of distances that are common to all 3D objects and for which all distances between these points are included in the set. The early work used each heavy atom as a point in the graph. The resulting solution is the maximum common 3D substructure of the molecule.

The utility of clique detection for pharmacophore mapping required two changes.

- Considering multiple conformations.
- Defining the points to be generalized recognition features.

By incrementing the tolerance at which distances matched, and/or the conformations included. The disadvantage of clique detection methods is that typically one compound is selected as a reference and other compounds are fit to it. If a poor choice of reference molecule has been made, then the results can be confusing or useless.

Genetic algorithm :

A genetic algorithm performs its search by analogy to biological evolution. Possible solution is represented as alleles in a chromosome per molecule. The genetic operators of mutation and crossover operate to optimize some fitness (scoring) function for the whole set of individuals. The scoring function of the genetic algorithm is based on combination of

- The number and the similarity of the feature tat have been overlaid;
- The volume integrals of the overly and
- The van der Waal energy of the molecular conformation defined by the torsion angles encoded in the chromosome. Other programmes use different chromosomes and fitness function.

Partial least squares :

One may use a 3D QSAR study to select the bioactive conformation of the molecules. In multiway PLS one provides the potency of each molecule and sets of CoMFA descriptors for each potential alignment of the compounds with each conformation of the most potent compounds. The multiway PLS identifies the overlay and conformation of the reference compounds that best explains the biological potency of the compounds. This overlay was then used to generate a traditional CoMFA, which was shown to have excellent predictive ability on an external test set.

In related methods, 4D QSAR, thousand of conformer of each molecule are first generated by molecular dynamics sampling. For each particular pharmacophore chosen by the investigator, the alignments of each molecule is described by the fraction of time that its conformer occupy each particular type and location of points in a lattice that correlated with biological activity. This analysis is repeated for all potential superpositions. The model with the most predictive power is selected as that which most accurately describes the pharmacophore.

Constrained conformational search :

One can modify the conformational search used in the active analog approach to also examine possible matching points. This report also included a detailed analysis of how to rank possible pharmacophores using 3D database searching.

Statistical classification of activities of molecules for pharmacophore identification :

Recursive partitioning is a data-mining that uses statistical test to identify descriptor of objects that separates one class from another; in our context it would use molecular descriptor to identify those that separates activities from in actives. It works in tree fashion: one variable is used to make the first split, then all molecules in each subset are further independently divided, one descriptor at a time until either the compounds are completely classified correctly or further splits will not meet some statistical significance criterion.

In SCAMPI, the molecular descriptors are related to pharmacophore distance. Combined with the recursive partitioning is a constrained conformational search strategy, an enhancement of the idea in active analog approach and RECEPTOR. Slightly overlapping distance are set for pair of potential pharmacophoric atoms and most predictive distance. Pharmacophore pair is used as basis for first split for all molecules that contains this distance; the conformational search is repeated using this constrain and new sets for all 3-point pharmacophore. This process continues until no more pharmacophore point can be added /until a five point pharmacophore is found.

Gibbs sampling of pharmacophore bitmaps :

A convenient way to summarized the potential pharmacophore of molecule is to generates key in bitmap or fingerprint that describes the presence or absence of particular geometric feature such as distance between two points of a particular character or the angle between 3-points an interesting new approach each conformer of a molecule populates a sets of key of predetermined length. The result is that each molecule

can be associated with a different length bitmaps. The modification is required that the matches are of bits associated with only one conformer after bit sting alignment is discovered the corresponding conformation are identify by a clique detection algorithm an advantage of the Gibbs sampling is that it will detect two binding modes if that is necessary to explain data.

EGSITE 3D-Model of binding sives :

Early studies to derive an explicit model used Xoronoi polyhedron of different size shape and properties to quantitatively explain the observed SAR, EGSITE is an outgrowth of the approach, the algorithm start by placing all molecule into an uniform macromolecular binding site, then incrementally partitions the space into regions of different character and fits the ligands into complementary region by changing both its orientation in space and its conformation.

Methods to used pharmacophore feature in QSAR :

One may choose to evaluate or use pharmacophore for alignments in more traditional 3D-QSAR such as CoMFA, several program make specific use of pharmacophore feature in a 3D-QSAR using QSAR to direct the search for pharmacophore, the hydrogen module of catalyst searches first for pharmacophore present in the most potent ligands but not in less potent compounds. It then uses simulated annealing on all compounds to identify one or more equation that fit potency to the presence or absence of pharmacophore feature including a weight and distance dependence for the quality of the fit. The newer program phase score potential pharmacophore by the RMS fit, the deviation from ideal of the cosine of hydrogen bond angle, the volume overlap, and the energy of the conformations. It includes a penalty if not all molecule fit the hypothesis. Its technique can detect the situation where two binding modes are needed to explain the data. It then uses PLS on subfields generated around the molecule to produce a quantitative model. In the electron conformational methods the final QSAR is based on the presence or absence of the discovered pharmacophoric groups, corrected for the accessibility of the groups to potential interaction sites on the target biomolecules. (Theil W.[6] Comprehensive medicinal chemistry[7], Computer Aided Drug Design, Jonathan S. Mason. Ed.)

Molecular dynamics :

Molecular dynamics (MD) emerged as one of the first simulation methods from the pioneering applications to the dynamics of liquids by Alder and Wainwright and by Rahman in the late 1950s and early 1960s. Due to the revolutionary advances in computer technology and algorithmic improvements, MD has subsequently become a valuable tool in many areas of physics and chemistry. Since the 1970s MD has been used widely to study the structure and dynamics of macromolecules, such as proteins or nucleic acids**.** There are two main families of MD methods, which can be distinguished according to the model (and the resulting mathematical formalism) chosen to represent a physical system. In the 'classical' mechanics approach to MD simulations molecules are treated as classical objects, resembling very much the 'ball and stick' model. Atoms correspond to soft balls and elastic sticks correspond to bonds. The laws of classical mechanics define the dynamic of the system. The 'quantum' or 'first-principles' MD simulations, which started in the 1980s with the seminal work of Car and Parinello, take explicitly into account the quantum nature of the chemical bond. The electron density function for the valence electrons that determine bonding in the system is computed using quantum equations, whereas the dynamics of ions (nuclei with their inner electrons) is followed classically.[7]

Why Classical Molecular Dynamics?

- Includes temperature effects.
- Able to treat comparatively large systems for a relatively long time.
- Help interpret experiments, and to provide alternative interpretations.
- Test theories.
- Obtain detailed molecular level information.
- Structure *and* dynamics (compared to QC, MM, and MC . . .).
- Improve our understanding of nature through model-building. (Meller, J., *et al.*[8])

In the broadest sense, molecular dynamics is concerned with molecular motion. Motion is inherent to all chemical processes. Simple vibrations, like bond stretching and angle bending, give rise to IR spectra. Chemical reactions, hormone-receptor binding, and other complex processes are associated with many kinds of intra- and intermolecular motions. (Daniel S.[9]).

Simulation technique in molecular dynamic :

The most problematic consideration in the field of molecular simulation is that of limited available computational power. Simplification of the system or of the performed calculations artificially increases the available computational power; less computer time is needed for each simulation step, leading to greater sampling of the studied system. Various simulation techniques such as molecular mechanics (MM) and molecular dynamics (MD), Monte Carlo, and Brownian dynamics, as well as hybrids of these methods, have emerged as simplifications of the exact quantum mechanical description for large molecules.

A simulation of molecular dynamics is an important challenge of today. Biologists started recently to study the behavior of large molecular systems consisting of several

thousands of atoms with the help of parallel computers. The computational complexity of such systems increases considerably with the molecule size. Part of molecular dynamic application is to simulate the behavior of multiple-particles systems via numerical simulation techniques. The principle is to compute time series of positions and velocities of the atoms by integrating Newton's equation of motion. At each step of molecular dynamic application, the forces between atoms have to be computed. The computation of non-bonded forces, the Van der Waals and electrostatic (Coulomb) forces between charged atoms is the most consuming portion for a typical simulation. (Bernard *et al.*,[10]) By modeling the motions of atoms within a molecular system, molecular dynamics (MD) simulations can serve as a computational "microscope" onto phenomena that are difficult to observe experimentally. Such simulations hold great promise in biochemistry and molecular biology, where they allow functional observation of proteins, nucleic acids, membranes, and other building blocks of the cell. Unfortunately, many of the events of greatest biological and pharmaceutical interest take place on time scales that are still beyond the reach of MD simulations on modern computers. An MD simulation of such an event might involve tens of thousands of atoms, representing one or more biological macromolecules surrounded by a solvent environment consisting of solvated ions (in some cases) and a large number of water molecules. Because the vibrational frequencies of these atoms typically limit each simulation time step to a few femtoseconds, simulations of more than a microsecond have thus far proven infeasible on systems with more than about ten thousand atoms. It seems clear that longer simulations will require the application of massive computational parallelism. The scalability of MD codes, however, has historically been limited by formidable inter-processor communication requirements associated with the exchange of atomic positions and inter-atomic force data. A number of established MD codes, including CHARMM, AMBER, GROMACS, and NAMD are widely used in the research community, each supporting somewhat different features and targeting somewhat different goals. Of these, NAMD is regarded as the most scalable and the most efficient on highly parallel runs, although GROMACS typically achieves superior performance on single-processor runs. IBM's recently developed Blue Matter MD code was designed to scale up to the full 128K-processor size of Blue Gene/L. Further advances in the parallel execution of MD simulations, however, could have important implications for both scientific research and the development of novel pharmaceutical compounds. In an MD simulation, the positions and velocities of particles corresponding to atoms evolve according to the laws of classical physics. Each time step of an MD simulation involves

– Computing forces on each particle (*force computation*) and
– Using these forces to compute updated positions and velocities on each particle by numerically integrating Newton's laws of motion (*integration*). Most of the computational load lies in the force computation. (Kevin *et al.*,[11]).

Molecular dynamics simulation :

In the Molecular Dynamics (MD) method a trajectory (configurations as a function of time) of the molecular system is generated by simultaneous integration of Newton's equations of motion (10) and (11) for all the atoms in the system.

$$d^2ri(t)/dt^2 = mi;^1 Fi \quad (11)$$

$$Fi = - 6 V(r,, \ldots . rN)/ari \quad (12)$$

The force on atom i is denoted by 6 and time is denoted by *t.* MD simulation requires calculation of the gradient of the potential energy V(r), which therefore must be a differentiable function of the atomic coordinates ri. The integration of Equation (11) is performed in small time steps *t*, typically 1-10 for molecular systems. Static equilibrium quantities can be obtained by averaging over the trajectory, which must be of sufficient length to form a representative ensemble of the state of the system. In addition, dynamic information can be extracted. Another asset of MD simulation is that non-equilibrium properties can be efficiently studied by keeping the system in a steady non-equilibrium state. Viewed as a technique to search configuration space, the power of MD lies in the fact that the kinetic energy present in the system allows it to surmount energy barriers that are of order of *k*, Tper degree of freedom. By raising the temperature *T* a larger part of conformation space can be searched, as has been shown by DiNola *et al.,* 1881 who generated a series of different conformations of the hormone somatostatin by applying MD at T = 600 K and at 1200 K. At the elevated temperature the total energy and potential energy are monitored for conspicuous fluctuations which may signal a possibly significant conformational change. When minima in the total energy occur, the system is cooled down and equilibrated at temperature (300 K). In this way different conformations with comparable free energy were obtained. We note however that the search at elevated temperature favors selection of higher entropy conformations. Searching conformation space by MD is expected to be efficient for molecules up to about 100 atoms. For larger molecules, which may and are likely to show a particular topological fold, MD methods will not be able to generate major topological rearrangements. Even when the barriers separating two topologically different low energy regions of conformation space are of the order of k, T, the time needed for traversing them may be much too long to be covered in a MD simulation of 10-100 ps.

Monte carlo simulation :

The monte carlo (MC) simulation procedure by which a (canonical) ensemble is produced consists of the following steps. Given a starting configuration r, a new configuration is generated by random displacement of one (or more) atoms :

$$\mathbf{T_{s+1} = r_s + ar} \quad (13)$$

The random displacements Δr should be such that in the limit of a large number of successive displacements the available Cartesian space of all atoms is uniformly sampled. This does not mean that the actual sampling must be carried out in Cartesian space. It can be done, e.g., in internal coordinate space (r, θ, ω), but since the equivalent volume element is r^2 sinθ drdθdω, the sampling in internal coordinate space must be non-uniform in order to produce a uniform sampling in terms of Cartesian coordinates.

The newly generated configuration r_{s+1} is either accepted or rejected on the basis of an energy criterion involving the change $\Delta E = V(r_{s+1})–V(r_s)$ of the potential energy with respect to the previous Configuaration. The new configuration is accepted when ΔE < 0, or if ΔE > 0 when e-A'''kBT > R, where R is a random number taken from a uniform distribution over the interval (0, 1). Upon acceptance, the new configuration is counted and used as a starting point for the next random displacement. If the criteria are not met, the new configuration r_{s+1} is rejected. This implies that the previous configuration r_s is counted again and used as a starting point for another random displacement. It is relatively easy to understand that this procedure will generate a Boltzmann ensemble. We consider two configuarations r1 and r2 with energies El = V(r,) < V(r,) = E2. The probability of stepping from configuration r2 to r1 equals 1, the reverse step has a probability exp. (– (E, - E,/k, T). when the populations p_2 and p_1 of.

The two configuration are in equilibrium, one has detailed balanced conditions (14) or (15).

$$\mathbf{p_2.1 = P_1 e^{-(E_2-E_1)} K_3 T} \quad (14)$$

$$\frac{p_1}{p_2} Z \frac{e^{-E_1/K_3.T}}{e^{-E_2/K_3.T}} \quad (15)$$

Each configuration occurs with a probability proportional to its Boltzmann factor J. The advantage of this (Metropolis) Monte Carlo or Boltzmann sampling over random sampling is that most sampled configurations are relevant (low energy), while with random sampling much computational effort is likely to be spent on irrelevant (high energy) configurations. In order to obtain high computational efficiency, one would like to combine a large (random) step size with a high acceptance ratio. This is possible when applying MC techniques to simulate simple atomic or molecular liquids. However, for complex systems involving many covalently bound atoms, a reasonable acceptance ratio can only be obtained for very small step size. This is due to the fact that a random displacement will inevitably generate very high bond energy of the bonds of the displaced atom. This makes MC methods rather inefficient for (macro) molecular systems.

Application of simulations :

- Understanding in Terms of Atomic Properties.
- Interpretation of Biochemical Data on Repressor-Operator Binding by MD Computer Simulation.
- Interpretation of Biophysical Data on membrane Properties by MD Computer Simulation.
- Determination of Spatial Molecular Structure on the Basis of 2D-NMR, X-Ray or Neutron Diffraction Data.
- Prediction of Free Energy Changes by MD Simulation.
- Prediction of Structural Changes by MD Simulation.

Simulation program available :

AMBER :

- Molecular modelling suite written specifically for simulating biological molecules.
- Can perform many types of calculation including:
- Structure optimization, molecular dynamics and vibrational analysis.
- Includes a wide range of tools for analyzing the results and setting up the calculations.

GAUSSAIN :

- Able to perform molecular mechanics calculations on single molecules.
- Includes molecular dynamics and structure optimization.

GROMACS :

- GROMACS help in obtaining high throughput. In all cases, the efficient parallelization, and neighbor searching, and all the other optimized algorithm.
- Efforts to build an API including all facets of MD simulations will allow users to focus on applications rather than coding. (David *et al.*,[12]).

Conclusion :

Computer simulations have altered the interplay between experiment and theory. The essence of the simulation is the use of the computer to model a physical system. Calculations implied by a mathematical model are carried out by the machine and the results are interpreted in terms of physical properties. Since computer simulation deals with models it may be classified as a theoretical method, On the other hand, physical quantities can (in a sense) is measured on a computer, justifying the term 'computer experiment'. The crucial advantage of simulations

is the ability to expand the horizon of the complexity that separates 'solvable' from 'unsolvable'. Basic physical theories applicable to biologically important phenomena, such as quantum, classical and statistical mechanics, lead to equations that cannot be solved analytically, except for a few special cases. The quantum Schrodinger equation for any atom but hydrogen (or any molecule) or the classical Newton's equations of motion for a system on the more practical side, computer experiments can be used to discover and design new molecules. Testing properties of a molecule using computer modelling is faster and less expensive than synthesizing and characterizing it in a real experiment. Drug design by computer is commonly used in the pharmaceutical industry.

REFERENCES

1. **Wilfred, E.V.G. and Berendsen, H.J.C.** (1990). Computer simulation of molecular dynamics:Methodology, applications and perspectives in chemistry. *Angen. Chem. lnt. Ed. Engl.,* **29** (9) : 992-1023.
2. **Holtje, H.D. and Wolfgang, S.** Molecular modeling basic principles and application, computational tools for geometry optimization. Second edition, Wiley-VCH GmbH & Co. KGaA. 15-16.
3. **Guha, R. and Sardar, R.** (2001). Force fields in molecular mechanics.
4. **Bowen, J.P. and Liang, G.** New vista in molecular mechanics Practical application of computer aided drug design, Charifson PS. editor, Marcel Dekker INC. 515-519.
5. **Bulintick, P. and Winter, H.D.** (2009). Computational medicinal chemistry for drug discovery, First Indian edition, CRC Press Taylor and Francis group. London. New York. 37-45.
6. **Theil, W.** Modern Methods and Algorithms of Quantum Chemistry, NIC Series, **3** : 261-283.
7. Comprehensive medicinal chemistry.-II, Computer Aided Drug Design. Vol. 4, Jonathan S. Mason. Editor, Elsevier publisher. 121, 128-134.
9. **Meller, J. and Nicholas, C.** Molecular Dynamics., New York, USA.
10. **Daniel, S.** (2007). Molecular Dynamics, Uppsala Multidisciplinary Center for Advanced Computational Science. Spring.
11. **Bernard, P.E. and Namath, L.** Large scale simulation of parallel molecular dynamic. INRIA Lorraine, Villers les Nancy, FRANCE.
12. **Kevin, J. and Bowers, K.** Algorithms for Molecular Dynamics Simulations on Commodity Clusters *D.E.* Shaw Research, LLC, New York, 10036, USA
13. **David, V.S., Berk, H.** (2011). GROMACS–the road ahead, John Wiley & Sons, WIREs *Comp. Mol. Sci.,* 1.