



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 4, April 2018

Computer Simulation in Solving Chemical Problems of Chemistry

Dr. Anil Kumar

Dept. of Chemistry, D.A.V. (PG) College, Dehradun, India

ABSTRACT: Computational chemistry is a branch of chemistry that uses computer simulation to assist in solving chemical problems. It uses methods of theoretical chemistry, incorporated into computer programs, to calculate the structures and properties of molecules, groups of molecules, and solids. It is essential because, apart from relatively recent results concerning the hydrogen molecular ion (dihydrogen cation, see references therein for more details), the quantum many-body problem cannot be solved analytically, much less in closed form. While computational results normally complement the information obtained by chemical experiments, it can in some cases predict hitherto unobserved chemical phenomena. It is widely used in the design of new drugs and materials.

Examples of such properties are structure (i.e., the expected positions of the constituent atoms), absolute and relative (interaction) energies, electronic charge density distributions, dipoles and higher multipole moments, vibrational frequencies, reactivity, or other spectroscopic quantities, and cross sections for collision with other particles.

KEYWORDS: computational chemistry, experiments, multipole, vibrational, spectroscopic, collision

I. INTRODUCTION

The methods used cover both static and dynamic situations. In all cases, the computer time and other resources (such as memory and disk space) increase quickly with the size of the system being studied. That system can be a molecule, a group of molecules, or a solid. Computational chemistry methods range from very approximate to highly accurate; the latter is usually feasible for small systems only. Ab initio methods are based entirely on quantum mechanics and basic physical constants. Other methods are called empirical or semi-empirical because they use additional empirical parameters.

Both ab initio and semi-empirical approaches involve approximations. These range from simplified forms of the first-principles equations that are easier or faster to solve, to approximations limiting the size of the system (for example, periodic boundary conditions), to fundamental approximations to the underlying equations that are required to achieve any solution to them at all. For example, most ab initio calculations make the Born–Oppenheimer approximation, which greatly simplifies the underlying Schrödinger equation by assuming that the nuclei remain in place during the calculation. In principle, ab initio methods eventually converge to the exact solution of the underlying equations as the number of approximations is reduced. In practice, however, it is impossible to eliminate all approximations, and residual error inevitably remains. The goal of computational chemistry is to minimize this residual error while keeping the calculations tractable.

In some cases, the details of electronic structure are less important than the long-time phase space behavior of molecules. This is the case in conformational studies of proteins and protein-ligand binding thermodynamics. Classical approximations to the potential energy surface are used, typically with molecular mechanics force fields, as they are computationally less intensive than electronic calculations, to enable longer simulations of molecular dynamics. Furthermore, cheminformatics uses even more empirical (and computationally cheaper) methods like machine learning based on physicochemical properties. One typical problem in cheminformatics is to predict the binding affinity of drug molecules to a given target. Other problems include predicting binding specificity, off-target effects, toxicity, and pharmacokinetic properties.



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 4, April 2018

Building on the founding discoveries and theories in the history of quantum mechanics, the first theoretical calculations in chemistry were those of Walter Heitler and Fritz London in 1927, using valence bond theory. The books that were influential in the early development of computational quantum chemistry include Linus Pauling and E. Bright Wilson's 1935 Introduction to Quantum Mechanics – with Applications to Chemistry, Eyring, Walter and Kimball's 1944 Quantum Chemistry, Heitler's 1945 Elementary Wave Mechanics – with Applications to Quantum Chemistry, and later Coulson's 1952 textbook Valence, each of which served as primary references for chemists in the decades to follow.

With the development of efficient computer technology in the 1940s, the solutions of elaborate wave equations for complex atomic systems began to be a realizable objective. In the early 1950s, the first semi-empirical atomic orbital calculations were performed. Theoretical chemists became extensive users of the early digital computers. One major advance came with the 1951 paper in Reviews of Modern Physics by Clemens C. J. Roothaan in 1951, largely on the "LCAO MO" approach (Linear Combination of Atomic Orbitals Molecular Orbitals), for many years the second-most cited paper in that journal. A very detailed account of such use in the United Kingdom is given by Smith and Sutcliffe.^[1] The first ab initio Hartree–Fock method calculations on diatomic molecules were performed in 1956 at MIT, using a basis set of Slater orbitals. For diatomic molecules, a systematic study using a minimum basis set and the first calculation with a larger basis set were published by Ransil and Nesbet respectively in 1960.^[2] The first polyatomic calculations using Gaussian orbitals were performed in the late 1950s. The first configuration interaction calculations were performed in Cambridge on the EDSAC computer in the 1950s using Gaussian orbitals by Boys and coworkers.^[3] By 1971, when a bibliography of ab initio calculations was published,^[4] the largest molecules included were naphthalene and azulene.^{[5][6]} Abstracts of many earlier developments in ab initio theory have been published by Schaefer.^[7]

In 1964, Hückel method calculations (using a simple linear combination of atomic orbitals (LCAO) method to determine electron energies of molecular orbitals of π electrons in conjugated hydrocarbon systems) of molecules, ranging in complexity from butadiene and benzene to ovalene, were generated on computers at Berkeley and Oxford.^[8] These empirical methods were replaced in the 1960s by semi-empirical methods such as CNDO.^[9]

In the early 1970s, efficient ab initio computer programs such as ATMOL, Gaussian, IBMOL, and POLYAYTOM, began to be used to speed ab initio calculations of molecular orbitals. Of these four programs, only Gaussian, now vastly expanded, is still in use, but many other programs are now in use. At the same time, the methods of molecular mechanics, such as MM2 force field, were developed, primarily by Norman Allinger.^[10]

One of the first mentions of the term computational chemistry can be found in the 1970 book Computers and Their Role in the Physical Sciences by Sidney Fernbach and Abraham Haskell Taub, where they state "It seems, therefore, that 'computational chemistry' can finally be more and more of a reality."^[11] During the 1970s, widely different methods began to be seen as part of a new emerging discipline of computational chemistry.^[12] The Journal of Computational Chemistry was first published in 1980.

Computational chemistry has featured in several Nobel Prize awards, most notably in 1998 and 2013. Walter Kohn, "for his development of the density-functional theory", and John Pople, "for his development of computational methods in quantum chemistry", received the 1998 Nobel Prize in Chemistry.^[13] Martin Karplus, Michael Levitt and Arieh Warshel received the 2013 Nobel Prize in Chemistry for "the development of multiscale models for complex chemical systems".^[14]

II. DISCUSSION

The term theoretical chemistry may be defined as a mathematical description of chemistry, whereas computational chemistry is usually used when a mathematical method is sufficiently well developed that it can be automated for implementation on a computer. In theoretical chemistry, chemists, physicists, and mathematicians develop algorithms and computer programs to predict atomic and molecular properties and reaction paths for chemical reactions. Computational chemists, in contrast, may simply apply existing computer programs and methodologies to specific chemical questions.



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 4, April 2018

Computational chemistry has two different aspects:

- Computational studies, used to find a starting point for a laboratory synthesis or to assist in understanding experimental data, such as the position and source of spectroscopic peaks.
- Computational studies, used to predict the possibility of so far entirely unknown molecules or to explore reaction mechanisms not readily studied via experiments.

Thus, computational chemistry can assist the experimental chemist or it can challenge the experimental chemist to find entirely new chemical objects.

Several major areas may be distinguished within computational chemistry:

- The prediction of the molecular structure of molecules by the use of the simulation of forces, or more accurate quantum chemical methods, to find stationary points on the energy surface as the position of the nuclei is varied.
- Storing and searching for data on chemical entities
- Identifying correlations between chemical structures and properties (see quantitative structure–property relationship (QSPR) and quantitative structure–activity relationship (QSAR)).
- Computational approaches to help in the efficient synthesis of compounds.
- Computational approaches to design molecules that interact in specific ways with other molecules (e.g. drug design and catalysis).

Computational chemistry is not an exact description of real-life chemistry, as our mathematical models of the physical laws of nature can only provide us with an approximation. However, the majority of chemical phenomena can be described to a certain degree in a qualitative or approximate quantitative computational scheme.

Molecules consist of nuclei and electrons, so the methods of quantum mechanics apply. Computational chemists often attempt to solve the non-relativistic Schrödinger equation, with relativistic corrections added, although some progress has been made in solving the fully relativistic Dirac equation. In principle, it is possible to solve the Schrödinger equation in either its time-dependent or time-independent form, as appropriate for the problem in hand; in practice, this is not possible except for very small systems. Therefore, a great number of approximate methods strive to achieve the best trade-off between accuracy and computational cost.

Accuracy can always be improved with greater computational cost. Significant errors can present themselves in ab initio models comprising many electrons, due to the computational cost of full relativistic-inclusive methods. This complicates the study of molecules interacting with high atomic mass unit atoms, such as transitional metals and their catalytic properties. Present algorithms in computational chemistry can routinely calculate the properties of small molecules that contain up to about 40 electrons with errors for energies less than a few kJ/mol. For geometries, bond lengths can be predicted within a few picometers and bond angles within 0.5 degrees. The treatment of larger molecules that contain a few dozen atoms is computationally tractable by more approximate methods such as density functional theory (DFT).

There is some dispute within the field whether or not the latter methods are sufficient to describe complex chemical reactions, such as those in biochemistry. Large molecules can be studied by semi-empirical approximate methods. Even larger molecules are treated by classical mechanics methods that use what are called molecular mechanics (MM). In QM-MM methods, small parts of large complexes are treated quantum mechanically (QM), and the remainder is treated approximately (MM).

III. RESULTS

One molecular formula can represent more than one molecular isomer: a set of isomers. Each isomer is a local minimum on the energy surface (called the potential energy surface) created from the total energy (i.e., the electronic



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 4, April 2018

energy, plus the repulsion energy between the nuclei) as a function of the coordinates of all the nuclei. A stationary point is a geometry such that the derivative of the energy with respect to all displacements of the nuclei is zero. A local (energy) minimum is a stationary point where all such displacements lead to an increase in energy. The local minimum that is lowest is called the global minimum and corresponds to the most stable isomer. If there is one particular coordinate change that leads to a decrease in the total energy in both directions, the stationary point is a transition structure and the coordinate is the reaction coordinate. This process of determining stationary points is called geometry optimization.

The determination of molecular structure by geometry optimization became routine only after efficient methods for calculating the first derivatives of the energy with respect to all atomic coordinates became available. Evaluation of the related second derivatives allows the prediction of vibrational frequencies if harmonic motion is estimated. More importantly, it allows for the characterization of stationary points. The frequencies are related to the eigenvalues of the Hessian matrix, which contains second derivatives. If the eigenvalues are all positive, then the frequencies are all real and the stationary point is a local minimum. If one eigenvalue is negative (i.e., an imaginary frequency), then the stationary point is a transition structure. If more than one eigenvalue is negative, then the stationary point is a more complex one and is usually of little interest. When one of these is found, it is necessary to move the search away from it if the experimenter is looking solely for local minima and transition structures.

The total energy is determined by approximate solutions of the time-dependent Schrödinger equation, usually with no relativistic terms included, and by making use of the Born–Oppenheimer approximation, which allows for the separation of electronic and nuclear motions, thereby simplifying the Schrödinger equation. This leads to the evaluation of the total energy as a sum of the electronic energy at fixed nuclei positions and the repulsion energy of the nuclei. A notable exception is certain approaches called direct quantum chemistry, which treat electrons and nuclei on a common footing. Density functional methods and semi-empirical methods are variants of the major theme. For very large systems, the relative total energies can be compared using molecular mechanics. The ways of determining the total energy to predict molecular structures are:

The programs used in computational chemistry are based on many different quantum-chemical methods that solve the molecular Schrödinger equation associated with the molecular Hamiltonian. Methods that do not include any empirical or semi-empirical parameters in their equations – being derived directly from theoretical principles, with no inclusion of experimental data – are called *ab initio* methods. This does not imply that the solution is an exact one; they are all approximate quantum mechanical calculations. It means that a particular approximation is rigorously defined on first principles (quantum theory) and then solved within an error margin that is qualitatively known beforehand. If numerical iterative methods must be used, the aim is to iterate until full machine accuracy is obtained (the best that is possible with a finite word length on the computer, and within the mathematical and/or physical approximations made).

The simplest type of *ab initio* electronic structure calculation is the Hartree–Fock method (HF), an extension of molecular orbital theory, in which the correlated electron–electron repulsion is not specifically taken into account; only its average effect is included in the calculation. As the basis set size is increased, the energy and wave function tend towards a limit called the Hartree–Fock limit. Many types of calculations (termed post-Hartree–Fock methods) begin with a Hartree–Fock calculation and subsequently correct for electron–electron repulsion, referred to also as electronic correlation. As these methods are pushed to the limit, they approach the exact solution of the non-relativistic Schrödinger equation. To obtain exact agreement with the experiment, it is necessary to include relativistic and spin orbit terms, both of which are far more important for heavy atoms. In all of these approaches, along with a choice of method, it is necessary to choose a basis set. This is a set of functions, usually centered on the different atoms in the molecule, which are used to expand the molecular orbitals with the linear combination of atomic orbitals (LCAO) molecular orbital method ansatz. *Ab initio* methods need to define a level of theory (the method) and a basis set.

The Hartree–Fock wave function is a single configuration or determinant. In some cases, particularly for bond-breaking processes, this is inadequate, and several configurations must be used. Here, the coefficients of the configurations, and of the basis functions, are optimized together.



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 4, April 2018

The total molecular energy can be evaluated as a function of the molecular geometry; in other words, the potential energy surface. Such a surface can be used for reaction dynamics. The stationary points of the surface lead to predictions of different isomers and the transition structures for conversion between isomers, but these can be determined without full knowledge of the complete surface.

A particularly important objective, called computational thermochemistry, is to calculate thermochemical quantities such as the enthalpy of formation to chemical accuracy. Chemical accuracy is the accuracy required to make realistic chemical predictions and is generally considered to be 1 kcal/mol or 4 kJ/mol. To reach that accuracy in an economic way it is necessary to use a series of post-Hartree–Fock methods and combine the results. These methods are called quantum chemistry composite methods.

Density functional methods

Density functional theory (DFT) methods are often considered to be ab initio methods for determining the molecular electronic structure, even though many of the most common functionals use parameters derived from empirical data, or from more complex calculations. In DFT, the total energy is expressed in terms of the total one-electron density rather than the wave function. In this type of calculation, there is an approximate Hamiltonian and an approximate expression for the total electron density. DFT methods can be very accurate for little computational cost. Some methods combine the density functional exchange functional with the Hartree–Fock exchange term and are termed hybrid functional methods.

Semi-empirical methods

Semi-empirical quantum chemistry methods are based on the Hartree–Fock method formalism, but make many approximations and obtain some parameters from empirical data. They were very important in computational chemistry from the 60s to the 90s, especially for treating large molecules where the full Hartree–Fock method without the approximations were too costly. The use of empirical parameters appears to allow some inclusion of correlation effects into the methods.

Primitive semi-empirical methods were designed even before, where the two-electron part of the Hamiltonian is not explicitly included. For π -electron systems, this was the Hückel method proposed by Erich Hückel, and for all valence electron systems, the extended Hückel method proposed by Roald Hoffmann. Sometimes, Hückel methods are referred to as "completely empirical" because they do not derive from a Hamiltonian.^[15] Yet, the term "empirical methods", or "empirical force fields" is usually used to describe Molecular Mechanics.^[16]

Molecular mechanics

In many cases, large molecular systems can be modeled successfully while avoiding quantum mechanical calculations entirely. Molecular mechanics simulations, for example, use one classical expression for the energy of a compound, for instance, the harmonic oscillator. All constants appearing in the equations must be obtained beforehand from experimental data or ab initio calculations.

The database of compounds used for parameterization, i.e., the resulting set of parameters and functions is called the force field, is crucial to the success of molecular mechanics calculations. A force field parameterized against a specific class of molecules, for instance, proteins, would be expected to only have any relevance when describing other molecules of the same class.

These methods can be applied to proteins and other large biological molecules, and allow studies of the approach and interaction (docking) of potential drug molecules.^{[17][18]}

Methods for solids

Computational chemical methods can be applied to solid-state physics problems. The electronic structure of a crystal is in general described by a band structure, which defines the energies of electron orbitals for each point in the Brillouin zone. Ab initio and semi-empirical calculations yield orbital energies; therefore, they can be applied to band structure



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 4, April 2018

calculations. Since it is time-consuming to calculate the energy for a molecule, it is even more time-consuming to calculate them for the entire list of points in the Brillouin zone.

Chemical dynamics

Once the electronic and nuclear variables are separated (within the Born–Oppenheimer representation), in the time-dependent approach, the wave packet corresponding to the nuclear degrees of freedom is propagated via the time evolution operator (physics) associated to the time-dependent Schrödinger equation (for the full molecular Hamiltonian). In the complementary energy-dependent approach, the time-independent Schrödinger equation is solved using the scattering theory formalism. The potential representing the interatomic interaction is given by the potential energy surfaces. In general, the potential energy surfaces are coupled via the vibronic coupling terms.

The most popular methods for propagating the wave packet associated to the molecular geometry are:

- the split operator technique,
- the Chebyshev (real) polynomial,
- the multi-configuration time-dependent Hartree method (MCTDH),
- the semiclassical method.

IV. CONCLUSIONS

Molecular dynamics

Molecular dynamics (MD) use either quantum mechanics, molecular mechanics or a mixture of both to calculate forces which are then used to solve Newton's laws of motion to examine the time-dependent behavior of systems. The result of a molecular dynamics simulation is a trajectory that describes how the position and velocity of particles varies with time. The phase point of a system described by the positions and momenta of all its particles on a previous time point will determine the next phase point in time by integrating over Newton's laws of motion.

Monte Carlo

Monte Carlo (MC) generates configurations of a system by making random changes to the positions of its particles, together with their orientations and conformations where appropriate. It is a random sampling method, which makes use of the so-called importance sampling. Importance sampling methods are able to generate low energy states, as this enables properties to be calculated accurately. The potential energy of each configuration of the system can be calculated, together with the values of other properties, from the positions of the atoms.^[19]

Quantum mechanics/Molecular mechanics (QM/MM)

QM/MM is a hybrid method that attempts to combine the accuracy of quantum mechanics with the speed of molecular mechanics. It is useful for simulating very large molecules such as enzymes.

The atoms in molecules (QTAIM) model of Richard Bader was developed to effectively link the quantum mechanical model of a molecule, as an electronic wavefunction, to chemically useful concepts such as atoms in molecules, functional groups, bonding, the theory of Lewis pairs, and the valence bond model. Bader has demonstrated that these empirically useful chemistry concepts can be related to the topology of the observable charge density distribution, whether measured or calculated from a quantum mechanical wavefunction. QTAIM analysis of molecular wavefunctions is implemented, for example, in the AIMAll software package.

Many self-sufficient computational chemistry software packages exist. Some include many methods covering a wide range, while others concentrate on a very specific range or even on one method. Details of most of them can be found in:



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 4, April 2018

- Biomolecular modelling programs: proteins, nucleic acid.
- Molecular mechanics programs.
- Quantum chemistry and solid state-physics software supporting several methods.
- Molecular design software
- Semi-empirical programs.
- Valence bond programs.

REFERENCES

1. Smith, S. J.; Sutcliffe, B. T. (1997). "The development of Computational Chemistry in the United Kingdom". *Reviews in Computational Chemistry*. 10: 271–316.
2. ^ Schaefer, Henry F. III (1972). *The electronic structure of atoms and molecules*. Reading, Massachusetts: Addison-Wesley Publishing Co. p. 146.
3. ^ Boys, S. F.; Cook, G. B.; Reeves, C. M.; Shavitt, I. (1956). "Automatic fundamental calculations of molecular structure". *Nature*. 178 (2): 1207. Bibcode:1956Natur.178.1207B. doi:10.1038/1781207a0. S2CID 4218995.
4. ^ Richards, W. G.; Walker, T. E. H.; Hinkley R. K. (1971). *A bibliography of ab initio molecular wave functions*. Oxford: Clarendon Press.
5. ^ Preuss, H. (1968). "DasSCF-MO-P(LCGO)-Verfahren und seine Varianten". *International Journal of Quantum Chemistry*. 2 (5): 651. Bibcode:1968IJQC....2..651P. doi:10.1002/qua.560020506.
6. ^ Buenker, R. J.; Peyerimhoff, S. D. (1969). "Ab initio SCF calculations for azulene and naphthalene". *Chemical Physics Letters*. 3 (1): 37. Bibcode:1969CPL.....3...37B. doi:10.1016/0009-2614(69)80014-X.
7. ^ Schaefer, Henry F. III (1984). *Quantum Chemistry*. Oxford: Clarendon Press.
8. ^ Streitwieser, A.; Brauman, J. I.; Coulson, C. A. (1965). *Supplementary Tables of Molecular Orbital Calculations*. Oxford: Pergamon Press.
9. ^ Pople, John A.; Beveridge, David L. (1970). *Approximate Molecular Orbital Theory*. New York: McGraw Hill.
10. ^ Allinger, Norman (1977). "Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms". *Journal of the American Chemical Society*. 99 (25): 8127–8134. doi:10.1021/ja00467a001.
11. ^ Fernbach, Sidney; Taub, Abraham Haskell (1970). *Computers and Their Role in the Physical Sciences*. Routledge. ISBN 978-0-677-14030-8.
12. ^ "vol 1, preface". *Reviews in Computational Chemistry*. 1990. doi:10.1002/9780470125786. ISBN 9780470125786.^[permanent dead link]
13. ^ "The Nobel Prize in Chemistry 1998".
14. ^ "The Nobel Prize in Chemistry 2013" (Press release). Royal Swedish Academy of Sciences. October 9, 2013. Retrieved October 9, 2013.
15. ^ Counts, Richard W. (1987-07-01). "Strategies I". *Journal of Computer-Aided Molecular Design*. 1 (2): 177–178. Bibcode:1987JCAMD...1..177C. doi:10.1007/bf01676961. ISSN 0920-654X. PMID 3504968. S2CID 40429116.
16. ^ Dinur, Uri; Hagler, Arnold T. (1991). Lipkowitz, Kenny B.; Boyd, Donald B. (eds.). *Reviews in Computational Chemistry*. John Wiley & Sons, Inc. pp. 99–164. doi:10.1002/9780470125793.ch4. ISBN 9780470125793.
17. ^ Rubenstein, Lester A.; Zauhar, Randy J.; Lanzara, Richard G. (2006). "Molecular dynamics of a biophysical model for β_2 -adrenergic and G protein-coupled receptor activation" (PDF). *Journal of Molecular Graphics and Modelling*. 25 (4): 396–409. doi:10.1016/j.jmgm.2006.02.008. PMID 16574446. Archived (PDF) from the original on 2008-02-27.
18. ^ Rubenstein, Lester A.; Lanzara, Richard G. (1998). "Activation of G protein-coupled receptors entails cysteine modulation of agonist binding" (PDF). *Journal of Molecular Structure: THEOCHEM*. 430: 57–71. doi:10.1016/S0166-1280(98)90217-2. Archived (PDF) from the original on 2004-05-30.
19. ^ Allen, M. P. (1987). *Computer simulation of liquids*. D. J. Tildesley. Oxford [England]: Clarendon Press. ISBN 0-19-855375-7. OCLC 15132676.