

# Computational Chemistry Methods

Subjects: Others | Physics, Atomic, Molecular & Chemical

Contributor: Albert Rimola, Piero Ugliengo

The main objective of computational chemistry is to solve chemical problems by simulating chemical systems (molecular, biological, materials) in order to provide reliable, accurate and comprehensive information at an atomic level. To this end, there are two main methodological families: those based on quantum chemical methods and those based on molecular mechanics. The former are methods in which the electrons are explicitly accounted for, while in the latter their presence is hidden in the force field.

Keywords: computational chemistry ; quantum chemistry ; density functional theory (DFT) ; materials modelling ; periodic surfaces ; cluster models ; astrochemistry ; interstellar grains ; silicates ; water ice

---

## 1. Quantum Chemical (QC) Methods

QC methods, also called electronic structure, first-principles or ab initio methods, calculate how electrons and nuclei interact by solving the time-independent electronic Schrödinger equation in the Born–Oppenheimer approximation. QC can be classified into two groups based on either wave function methods or the density functional theory (DFT).

### 1.1. Wave Function-Based Methods

These are methods that, by solving the electron Schrödinger equation, calculate the explicitly correlated electronic wave functions. The simplest wave function-based method is the Hartree–Fock (HF), in which the multielectron wave function is approximated to a single Slater determinant (the mathematical expressions for wave functions in quantum mechanics). This approximation, however, leads to the main pitfall of HF: it neglects the instantaneous electron correlation. That is, it considers that one electron moves in an averaged field due to the remaining electrons, when in reality the motion of the electrons is correlated (the motion of one electron depends on the instantaneous, mutual interaction with the other electrons). The lack of instantaneous electron correlation introduces an error in the HF wave function and energy, seriously affecting the prediction of the kinetic barriers or the description of London dispersion forces.

Despite this inaccuracy, the HF solution can be systematically improved. One way is through the many body perturbation theory, culminated in the Møller–Plesset (MPn) methods <sup>[1]</sup>, or by expressing the wave function as a linear combination of Slater determinants, i.e., the HF one plus those representing single, double, triple, etc., excitations, giving rise to the configuration interaction (CI) methods <sup>[2]</sup>. The coupled cluster (CC) theory also introduces excited state configurations to the wave function but adopting an exponential expansion (through the cluster operator) on the HF one <sup>[3]</sup>. The most popular CC derivation is the CCSD(T) method <sup>[4]</sup>, in which single (S) and double (D) excitations are included through the cluster operator and triple (T) excitations by using the perturbation theory. CCSD(T) extensively includes electron correlation and, in combination with largely extended basis sets, is considered the “gold standard” in QC <sup>[5]</sup>.

Despite the improved quality of the post-HF methods, they are considerably computationally expensive so that their applicability is hampered by the size of the systems. CCSD(T) calculations for more than 25–30 atoms are unfeasible so that, in practice, these calculations in surface modelling are mainly used for calibrating more approximate (but cheaper) methods to ensure their suitability in the description of the electronic structure of the simulated surfaces.

### 1.2. Methods Based on the Density Functional Theory (DFT)

The DFT foundations lie on the mathematical formulations developed by Kohn and Sham in 1965 <sup>[6]</sup>. The central theorem states that the ground state energy of a non-degenerate electronic system is unequivocally defined by the total electron density  $\rho(r)$  through a universal and exact mathematical functional (a functional takes a function and defines a single number from it, like in a definite integral)—that is, the energy of a system can be expressed as a functional  $E$  of  $\rho(r)$ , i.e.,  $E[\rho(r)]$ . This represents an enormous advantage with respect to the wave function-based methods because, in DFT, the  $\rho(r)$  for a system of  $N$  electrons only depends on 3 spatial coordinates, while the corresponding wave function depends on  $3N$  spatial and  $N$  spin variables. Because of this, DFT methods are even cheaper than HF (according to the choice of the

functional  $E[\rho(r)]$ ) while including a significant fraction of the missing instantaneous electron correlation. The disadvantage of DFT is that although the existence of the universal functional relating the electron density to the energy of a system can be mathematically demonstrated, the exact form of  $E[\rho(r)]$  is still unknown. The search for functionals connecting these two quantities, known as exchange-correlation functional  $E_{xc}[\rho(r)]$ , leads to the existence of a wide variety of DFT methods [7].

The simplest approach for  $E_{xc}[\rho(r)]$  is the Local Density Approximation (LDA), which implicitly assumes that  $E_{xc}[\rho(r)]$  does only depend on the functional expression on the  $\rho(r)$  value at the local position  $r$ . LDA is improved by the Generalized Gradient Approximation (GGA), in which the  $E_{xc}[\rho(r)]$  depends not only on the density  $\rho(r)$  but also on the gradient of the density,  $\nabla\rho(r)$ , this way accounting for the spatial varying electron density present in any chemical system. Meta-GGA functionals are a new class of methods in which higher order density gradients or the kinetic energy density are accounted for. Combination of conventional GGA  $E_{xc}[\rho(r)]$  with HF-like electron exchange functional  $E_x(\text{HF})$  gives rise to the DFT hybrid functionals. For these cases, since the weight factor for each component in the definition of the  $E_{xc}[\rho(r)]$  cannot be assigned from first-principles, a certain degree of empiricism is used, for instance by fitting the coefficients to experimental data or to post-HF wave-function-based results. Finally, hybrid-meta GGA methods apply a meta-GGA similar concept to conventional hybrid functionals so that these methods depend on the electron density and its gradient, a percentage of exact exchange and the kinetic energy density.

As mentioned, a wide number of DFT methods have been developed. However, only some of them are most commonly used since they provide a reasonable description of the electronic structure and related properties of broad different molecular and solid-state systems. Usual GGA methods are PBE [8], OPBE [9][10], PW91 [11][12] and BLYP [13][14]. Conventional hybrid functionals are B3LYP [14][15], BHLYP [14][16], PBE0 [17][18] and wB97X [19]. The most used meta-GGA and hybrid-meta-GGA are functionals developed by Truhlar and coworkers, particularly those of the M06 family: the M06L meta-GGA [20] and the M06 and M06-2X hybrid-meta-GGA [21].

Among the main drawbacks of the DFT methods (specially the oldest ones) is that they do not cope with long-range non-covalent (i.e., dispersion or London) interactions. Although this is not particularly important in the modelling of conventional surfaces (i.e., metals, oxides, etc.), it gains relevance when a proper simulation of the elementary steps taking place at the surfaces (i.e., adsorption, diffusion, reaction) is aimed. A pragmatic solution to account for dispersion interactions in DFT methods was proposed by Grimme, in which an a posteriori correction term based on an atom–atom additive London-type empirical potential (D) is introduced to supplement the DFT energy [22]. Different correction terms have been sequentially proposed and improved (D [23], D2 [24], D3 [25] and D4 [26]). The main advantage of the DFT-D scheme is that it maintains the original accuracy of the pure DFT method for short-range interactions while including the dispersion interactions at a negligible computational cost.

### 1.3. Semiempirical Methods

Despite the enormous cost-effective advantage of the DFT methods with respect to the wave function ones, DFT simulations are also limited by the size of the systems up to hundreds of atoms or even thousands of atoms with high performance computing allowing massively parallel calculations. This is an important aspect in surface modelling, because, in most of the cases, realistic extended systems are inevitably large (e.g., amorphous surfaces), reaching the limits of the DFT applicability.

Semiempirical methods are faster alternative approaches to overcome the size limitations of DFT. These methods are derived from the pure QC methods but with different approximations (e.g., omission of the bielectronic integrals) while including some empirical parameters to mitigate the errors due to the approximations and to account for electron correlation effects. Due to these simplifications and parametrizations, semiempirical methods are faster than their ab initio counterparts, permitting the treatment of large chemical systems, impossible to deal with the pure QC methods. However, accuracy of the semiempirical methods strongly depends on whether the parametrization is suitable for the specific case under study. That is, results can be very wrong if the simulated system does not fit with the training set used to parametrize the method.

Semiempirical methods can be classified into several groups: (i) by Pople [27][28][29][30]—CNDO/2, INDO and NDDO—in which the parametrization is mostly based on HF results; (ii) by Dewar—AM1 [31] and MNDO [32]—which are NDDO-derived methods parametrized with experimental fittings; (iii) the PMx ( $x = 3, 6$  and  $7$ ) family developed by Stewart [33][34] [35], in which more experimental parameters are introduced in NDDO; and (iv) by the tight binding approaches (e.g., SCC-DFTB [36][37]), which are the semiempirical methods derived from DFT. Very recently, Grimme and coworkers developed the GFN-xTB family [38][39][40], a set of tight binding-based methods that model covalent as well as H-bond and dispersion interactions with improved accuracy, parametrized for almost the entire periodic table of elements ( $Z \leq 86$ ).

## 1.4. Basis Set

Basis sets are a linear combination of basis functions used to build molecular orbitals, which in turn approximately represent the electronic wave function. Thus, the use of basis sets in electronic structure calculations is compulsory. The quality of the calculation and the accuracy of the derived information strongly depend on the completeness of the chosen basis set.

In addition to the Slater-type orbitals (STO), two types of basis sets are usually used to simulate molecular and periodic systems: Gaussian-type orbitals (GTOs) and plane waves (PWs). GTOs are localized functions centered on the atoms. They are commonly used in molecular (i.e., finite) calculations because they obey the typical radial–angular decomposition and exhibit the spatial and symmetry properties of atomic orbitals. PWs are periodic functions in nature, i.e., they are not localized functions but uniformly diffuse in space, so that their use requires the adoption of periodic boundary conditions (in which an infinite system is approximated by repeating a unit cell, see below). Because of that, they are widely used in the simulation of periodic solid-state systems.

The main advantage of GTOs with respect to PWs is that the number of basis functions exclusively depends on the number of atoms of the system, while the number of PWs, since they fill the 3D space, depends on the unit cell size. Advantages of PWs with respect to GTOs are that their quality is controlled by a unique parameter (the cut-off kinetic energy) and that they do not suffer from basis set superposition errors (BSSE), which can be dramatically critical for adsorption purposes. BSSE is an artefact when using truncated GTO basis sets, in which there is a system composed of two fragments: one fragment uses the basis set of the other and vice versa, this way leading to an overestimation of the total energy and, therefore, affecting the optimized geometries. BSSE is usually corrected by the counterpoise method <sup>[41]</sup>, with schemes to correct energies (CP) and geometries (gCP). In relation to that, it is worth mentioning the recent development of the HF-3c method <sup>[42]</sup>. HF-3c is based on HF using the minimal MINIX basis set in which 3 correction (3c) terms are applied (upon inclusion of empirical parameters) to include London dispersions (D3 scheme), to alleviate BSSE (gCP scheme), and to correct for short-range effects caused by basis set incompleteness (i.e., to remove the systematic overestimation of bond lengths for electronegative elements when employing small basis sets).

## 2. Classical Molecular Mechanic (MM) and Molecular Dynamic (MD) Simulations

Methods based on classical MM (also called force field methods) ignore the electrons and their motion. A chemical system is described by a “ball and spring” model, in which atoms are represented by balls of different sizes and softness and the bonds by springs of different stiffness. The energy of the system is calculated as a function of the nuclear positions only. To this end, force fields (FFs), a set of interatomic potentials (IPs) encompassing energy functions and parameters, are used to define the molecular mechanics energy measuring the degree of mechanical strain within the system. As electrons are not explicitly accounted for in the classical MM, very large systems (thousands of atoms) can be simulated to predict conformational flexibility and relative stability (e.g., protein folding).

FFs contain energy functions that include bonded and non-bonded terms to represent the intra- (i.e., covalent) and inter- (i.e., non-covalent) molecular forces within the system. In essence, the various contributions present in the IPs should model: (i) the interaction between pairs of bonded atoms (covalent bond stretching), (ii) the energy required for bending an angle formed by three atoms (angle bending), (iii) the energy change associated with a bond rotation (torsional or dihedral), and iv) the non-bonded pairwise interactions, usually using a Coulomb potential for electrostatic interactions and a Lennard-Jones potential for van der Waals interactions.

The energy functions contain a set of parameters (e.g., equilibrium values for bond lengths, angles, dihedrals and atomic charges) that define the different types of atoms, chemical bonds, angles torsions, non-bonded interactions and other terms. The FF parametrization (i.e., the numerical assignment of the parameters) was historically derived from experimental data, but nowadays has been replaced by QC calculations.

Due to the cheap cost, MM methods are adopted to simulate the dynamics of the nuclear motion within a molecule through molecular dynamic (MD) simulations. In MD, successive configurations of the system are generated by integrating the Newton's laws of motion. The result is a trajectory that specifies how the positions and velocities of the nuclei vary with time, in which at each nuclear position, the molecular mechanics energy and the nuclear forces of the system are evaluated. Thus, MD simulations are a good choice to study the evolution in time–space phase of the atomic positions subject to the internal forces of chemical nature, while the kinetic energy associated with the nuclear motion provides the temperature of the system. Classical MD simulations can reach time-scales of hundreds of picoseconds to microseconds, which allows the ergodicity of the system to be reached.

MD simulations can also be carried out by moving the nuclei within the electronic field generated by the corresponding electronic wave function or, more conveniently, by the electron density usually computed within the DFT. For these cases, the electrons are treated quantum mechanically, while the nuclei, within the Born–Oppenheimer approximation, are treated as classical particles so that their dynamics can be followed by the integration of the Newton equations, like in the MM method. Hence, these simulations are called *ab initio* molecular dynamics (AIMD). In AIMD simulations, although being computationally more expensive than pure classical MD, the introduction of the electronic structure of the system allows the study of bond breaking/formation as the result of internal exchange of energy. Thanks to the huge improvement of the power of high-performance computers, AIMD can nowadays reach tens to hundreds of picoseconds.

To overcome the high cost and limited time evolution intrinsically linked to AIMD, the relatively new methodologies of machine learning (ML) can be adopted. ML is becoming an explosive field, not only in problems such as image recognition, language translation or to play the GO game but also to tackle physical chemistry problems. The present review cannot expand to include the multitude of new contributions by ML in the present topic due to space constraints and limited authors experience in the ML field. In essence, neural network (NN) approaches can be used to model the potential energy surface for almost any kind of systems through a specific learning based on *ab initio* data run on a limited set of representative cases. The interested reader can refer to this perspective article <sup>[43]</sup>. An impressive example of the success of ML in greatly expanding the applicability of AIMD quality calculation up to 100 million atoms at the rate of 1ns/day was recently made possible by the combination of the Deep Potential Molecular Dynamics approach using a highly optimized code (GPU DeePMD-kit) on the Summit supercomputer (see Reference <sup>[44]</sup>).

---

## References

1. Møller, C.; Plesset, M.S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* 1934, 46, 618–622.
2. Siegbahn, P.E.M. The Configuration Interaction Method. In *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*; Roos, B.O., Ed.; Springer: Berlin/Heidelberg, Germany, 1992; pp. 255–293.
3. Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ur sell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* 1966, 45, 4256–4266.
4. Raghavachari, K.; Trucks, G.W.; Pople, J.A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* 1989, 157, 479–483.
5. Sherrill, C.D. Frontiers in electronic structure theory. *J. Chem. Phys.* 2010, 132, 110902.
6. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* 1965, 140, A1133–A1138.
7. Sousa, S.F.; Fernandes, P.A.; Ramos, M.J. General Performance of Density Functionals. *J. Phys. Chem. A* 2007, 111, 10439–10452.
8. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* 1996, 77, 3865–3868.
9. Hoe, W.-M.; Cohen, A.J.; Handy, N.C. Assessment of a new local exchange functional OPTX. *Chem. Phys. Lett.* 2001, 341, 319–328.
10. Handy, N.C.; Cohen, A.J. Left-right correlation energy. *Mol. Phys.* 2001, 99, 403–412.
11. Perdew, J.P.; Chevary, J.A.; Vosko, S.H.; Jackson, K.A.; Pederson, M.R.; Singh, D.J.; Fiolhais, C. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* 1992, 46, 6671–6687.
12. Perdew, J.P.; Burke, K.; Wang, Y. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B* 1996, 54, 16533–16539.
13. Becke, A.D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 1988, 38, 3098–3100.
14. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 1988, 37, 785–789.
15. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 1993, 98, 5648–5652.
16. Becke, A.D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* 1993, 98, 1372–1377.
17. Perdew, J.P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* 1996, 105, 9982–9985.

18. Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* 1999, 110, 6158–6170.
19. Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* 2008, 128, 084106.
20. Zhao, Y.; Truhlar, D.G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* 2006, 125, 194101.
21. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* 2008, 120, 215–241.
22. Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* 2011, 1, 211–228.
23. Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* 2004, 25, 1463–1473.
24. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* 2006, 27, 1787–1799.
25. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 2010, 132, 154104.
26. Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* 2017, 147, 034112.
27. Pople, J.; Beveridge, D. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, NY, USA, 1970.
28. Pople, J.A.; Segal, G.A. *Approximate Self-Consistent Molecular Orbital Theory. II. Calculations with Complete Neglect of Differential Overlap*. *J. Chem. Phys.* 1965, 43, S136–S151.
29. Pople, J.A.; Segal, G.A. *Approximate Self-Consistent Molecular Orbital Theory. III. CNDO Results for AB<sub>2</sub> and AB<sub>3</sub> Systems*. *J. Chem. Phys.* 1966, 44, 3289–3296.
30. Pople, J.A.; Beveridge, D.L.; Dobosh, P.A. *Approximate Self-Consistent Molecular-Orbital Theory. V. Intermediate Neglect of Differential Overlap*. *J. Chem. Phys.* 1967, 47, 2026–2033.
31. Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 1985, 107, 3902–3909.
32. Dewar, M.J.S.; Thiel, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.* 1977, 99, 4899–4907.
33. Stewart, J.J.P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* 1989, 10, 209–220.
34. Stewart, J.J.P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* 2007, 13, 1173–1213.
35. Stewart, J.J.P. Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Model.* 2013, 19, 1–32.
36. Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B* 1995, 51, 12947–12957.
37. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* 1998, 58, 7260–7268.
38. Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86). *J. Chem. Theory Comput.* 2017, 13, 1989–2009.
39. Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* 2019, 15, 1652–1671.
40. Pracht, P.; Caldeweyher, E.; Ehlert, S.; Grimme, S. A Robust Non-Self-Consistent Tight-Binding Quantum Chemistry Method for large Molecules. *ChemRxiv* 2019.
41. Boys, S.F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* 1970, 19, 553–566.

42. Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comput. Chem.* 2013, 34, 1672–1685.
43. Ceriotti, M. Unsupervised machine learning in atomistic simulations, between predictions and understanding. *J. Chem. Phys.* 2019, 150, 150901.
44. Jia, W.; Wang, H.; Chen, M.; Lu, D.; Lin, L.; Car, R.; E, W.; Zhang, L. Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Atlanta, GA, USA, 9–19 November 2020; Article 5.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/16030>