

An introduction to quantum chemical methods applied to drug design

Marco Stenta, Matteo Dal Peraro

Laboratory for Biomolecular Modeling, Institute of Bioengineering, School of Life Sciences, Swiss Federal Institute of Technology - EPF Lausanne CH-1015

TABLE OF CONTENTS

1. Abstract
2. Introduction: from drug discovery to drug design
3. Molecular modelling
4. Quantum chemistry-based methods
 - 4.1. Møller-Plesset perturbation theory
 - 4.2 Density Functional Theory (DFT)
 - 4.3. Dispersion force corrections in exchange-correlation functionals
 - 4.4. Semi-empirical methods
 - 4.5. Density Functional Tight-Binding
5. Molecular Mechanics approach
6. Hybrid Quantum Mechanics/Molecular Mechanics methods
7. Metalloenzymes: a challenge for QM methods
8. The Potential Energy Surface
9. The Free Energy Surface
10. Future perspectives
11. References

1. ABSTRACT

The advent of molecular medicine allowed identifying the malfunctioning of subcellular processes as the source of many diseases. Since then, drugs are not only discovered, but actually designed to fulfill a precise task. Modern computational techniques, based on molecular modeling, play a relevant role both in target identification and drug lead development. By flanking and integrating standard experimental techniques, modeling has proven itself as a powerful tool across the drug design process. The success of computational methods depends on a balance between cost (computation time) and accuracy. Thus, the integration of innovative theories and more powerful hardware architectures allows molecular modeling to be used as a reliable tool for rationalizing the results of experiments and accelerating the development of new drug design strategies. We present an overview of the most common quantum chemistry computational approaches, providing for each one a general theoretical introduction to highlight limitations and strong points. We then discuss recent developments in software and hardware resources, which have allowed state-of-the-art of computational quantum chemistry to be applied to drug development.

2. INTRODUCTION: FROM DRUG DISCOVERY TO DRUG DESIGN

The need to provide medicines can be traced back to the roots of humankind, when religious beliefs and common sense guided all attempts to cure disease. For many centuries during the *age of botanicals* (pre-1800s), medical care consisted of eating, drinking, or applying substances. Since the strategies for identifying new drugs were driven by empirical observation and intuition, each success was necessarily reliant on fortuitous accident. Pharmacopoeia regulating composition and dosage of the available drugs were introduced in Egypt as early as 16th Century BC (Ebers Papyrus, Edwin Smith Papyrus) and, later (5th Century BC) improved and refined in Greece (Corpus Hippocraticum); yet the first successful attempts to individuate illnesses by their symptoms only date back to the 19th Century. Since the cures were almost exclusively symptomatic, only a few authentic drugs were discovered in this period (quinine, cocaine, digitalis and, later, aspirin). The diffusion of the “microbial theory of diseases” at the turn of the last century was a key development in biology and medical science, paving the way for the advances of the 1920s and 1930s. The almost serendipitous discovery of

penicillin by Fleming provided the basis for the rational use of substances as specific drugs to cure specific illnesses. The *antibiotic era* began in the early 1940s, heralded by advances in synthetic organic and analytic chemistry, technical improvement of instruments and tools, a new multidisciplinary approach to studying disease, and an increased demand during World War II for penicillin derivatives and other antibiotics. Although the drug discovery process was becoming less serendipitous, the main approach in the 1950s and 1960s was still based on trial-and-error. Whenever a new active substance was discovered in nature, and its active principle purified, characterized and tested, entire libraries of similar compounds would then be synthesized, tested, and examined for higher activity levels and fewer side effects.

In the 1970s, cancer and heart diseases became the greatest targets for medical research. Drug discovery strategies based on massive and systematic scans of compound libraries were thus abandoned in favor of newer and more efficient approaches. To deal with such intrinsically multifactorial diseases, identifying specific targets appeared the only route to success(1). Recent advances in biochemistry, chemical biology, and biotechnology allowed scientists to look for drug targets at a subcellular level, which was desirable because many diseases were associated with the activity of nanoscale objects. The recent discoveries concerning DNA and protein structure dramatically changed landscape for those interested in finding new therapies. As soon as molecular and atomistic details about DNA replication and enzyme activity become available, the quest for molecules capable of interfering with those processes began in earnest. Drugs would not be discovered, but designed to exert an action on a specific target selected from thousands of potential targets in the organism's metabolic pathway. The atomistic structure of DNA and its complexes with proteins are now studied to identify molecules capable of interfering with replication (anticancer drugs). The active sites of enzymes are also studied in detail so that their activity can be better prevented by inhibitor molecules (antibacterial, anti-inflammatory drugs) (2, 3).

In this new era of drug development, computer sciences play many significant roles (4, 5). The most obvious is the classification, storage, and effective retrieval of the exponentially increasing amount of data coming from different research fields. Several drug development strategies are based on scanning databases of the properties of known compounds. Moreover, hardware improvements and the availability of new programming languages have allowed for some physical theories and methods to be employed in chemistry and biology. Computable models directly based on (or derived from) Quantum Mechanics (QM) are used to study the physical and chemical properties of molecules. These methods can provide insights at the molecular and atomistic level, which are usually difficult to obtain by experimental techniques. The reaction mechanism of enzymes, for instance, can be elucidated using these methods, providing a rational guide in the design of specific inhibitors. The binding affinity of drug candidates with proteins and DNA can also be tested

in silico, with considerable gain in both time and money. The increased reliability of QM-based methods means they are crucial drug development tools. The improved quality of the available molecular models and their widespread use in drug design are indeed contributing to a new era of molecular medicine.

3. MOLECULAR MODELLING

In recent decades, molecular modeling techniques have evolved with unexpected speed, becoming ever more reliable. As theoretical chemists developed new theories to describe reality, computational chemists were able to efficiently implement these models in parallel. Nowadays many experimental chemists, working in either organic or physical chemistry, can easily use modern software for both research and teaching (6).

The systems that can be considered by molecular modeling (6-8) range from small isolated molecules to large biological macromolecules (like proteins and DNA). Most molecular modeling studies involve three stages. In the first stage, the most appropriate level of theory must be chosen to describe the system in question. The balance between accuracy and speed must reflect the nature of the system as well as its size. The two most common computational approaches are based on quantum mechanics (QM) or molecular mechanics (MM). The former is based on an approximate solution of the Schrödinger equation, the latter on a classical description of the atoms following Newton's law. Both computational approaches provide the total energy for a given atomic configuration and can provide information about the system's thermodynamic properties. The second stage of a molecular modeling study is the actual calculation, using the chosen computational procedure to calculate the best geometrical arrangement (the one with the lowest energy), the reaction pathway, the behavior of the system as a function of time, and/or the value of relevant observables that are useful for rationalizing experimental data or predicting the behavior of the system in question. The third stage is the analysis of the results, which leads to the construction of an interpretative and predictive model.

4. QUANTUM CHEMISTRY-BASED METHODS

The so-called *ab initio* or first-principles methods are usually considered the most accurate and consistent (6) because they provide the best physical approximation of the system. The term *ab initio* implies that these methods, based on the laws of quantum mechanics, require the knowledge of just a few fundamental constants: the electron mass, the electron and nuclear charges, and the values of fundamental physical constants, such as Planck's constant and the speed of light.

The time-independent Schrödinger equation is used to find stationary, fixed states and is reported in Equation 1, where $\Psi(\vec{r})$ is the wavefunction describing the system, E is the energy of the system, and \hat{H} Hamiltonian operator. For a system of charged particles (electrons and

nuclei), \hat{H} can be written as in Equation 2, where \hat{T} in the kinetic energy operator, \hat{V} is the potential energy operator, m_a is the mass of the particle a , ∇_a^2 is the Laplacian operator for particle a , q_a and q_b are the charges of a and b , and \vec{r}_a and \vec{r}_b the positions of a and b .

Equation 1
Equation 2

$$\hat{H}\Psi = E\Psi$$

$$\hat{H} = \hat{T} + \hat{V} = \sum_a -\frac{\hbar^2}{2m_a} \nabla_a^2 + \sum_{a<b} \frac{q_a q_b}{|\vec{r}_a - \vec{r}_b|}$$

Since the ratio of electron-mass to proton-mass is about 1/1836, the electronic motion is much faster than the nuclear motion. Consequently, the electrons of a molecule can rapidly adjust to any change in the nuclear positions, and the energy of a molecule in its ground state can be considered a function of the nuclear coordinates only. This is the rationale behind the Born-Oppenheimer (BO) approximation (8, 9), which allows us to separate electronic and nuclear degrees of freedom. The total Hamiltonian can be split into its nuclear and electronic terms, as reported in Equations 3 and 4, where the indices μ, ν and i, j refer, respectively, to nuclei and electrons. Thus, the total Hamiltonian can be written in terms of one-, two- and zero-electron operators (see Equation 4).

Equation 3
Equation 4

$$\hat{H}_{nuc} = \sum_{\mu} -\frac{\hbar^2}{2m_{\mu}} \nabla_{\mu}^2$$

$$\hat{H}_{el} = \sum_i -\frac{\hbar^2}{2m_i} \nabla_i^2 - \sum_{\mu} \frac{Z_{\mu}}{|\vec{r}_{\mu} - \vec{r}_i|} + \sum_{i<j} \frac{\hbar^2}{|\vec{r}_i - \vec{r}_j|} + \sum_{\mu,\nu} \frac{Z_{\mu} Z_{\nu}}{|\vec{r}_{\mu} - \vec{r}_{\nu}|} = \sum_i \hat{h}_i + \sum_{i<j} \hat{g}_{ij} + \hat{h}_0$$

According to the Molecular Orbital (MO) approach, the molecular wavefunction can be written in terms of one-electron functions (spin-orbitals) $\psi_i(j)$ each defined as a product of a spatial part $\phi_i(j)$ (representing the “shape” of the orbital) and a spin function $\alpha_i(j)$ or $\beta_i(j)$, to identify the electron spin state (“up” or “down”). The i and j indices run on the spatial functions and the electrons, respectively. The anti-symmetric (due to the particle’s indistinguishability) electronic wave-function can be approximated as a product of spin-orbitals, in the form of a Slater Determinant (Equation 5). (10)

Equation 5

$$\Psi = (n!)^{-1/2} \begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) & \psi_2(1)\beta(1) & \cdots & \psi_{n/2}(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2) & \psi_2(2)\beta(2) & \cdots & \psi_{n/2}(2)\beta(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_1(n)\alpha(n) & \psi_1(n)\beta(n) & \psi_2(n)\beta(n) & \cdots & \psi_{n/2}(n)\beta(n) \end{vmatrix}$$

Each molecular orbital ϕ_i of a system can be expanded in terms of a set of N predefined one-electron functions $\chi_{\mu i}$ denoted as basis functions (basis set) according to Equation 6, where $c_{\mu i}$ are the molecular orbital expansion coefficients. When atomic orbitals (AO) are used as basis functions, this approach is often referred to as Linear Combination of Atomic Orbital (LCAO) approximation.

Equation 6

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu}$$

4.1. Hartree-Fock

For decades, the Hartree-Fock (11) (HF) method has been one of the most widely used computational approaches to determining the wavefunction within the molecular orbital model. The wavefunction is written in the form of a Slater determinant. The procedure is based on the use of the Hartree-Fock eigenvalues equations (Equation 7), where the Fock operator \hat{F} acts on the orbitals ϕ_i to give the same orbital function multiplied by a constant ϵ_i , which represents the energy of the orbital. The LCAO approximation can be applied to obtain Equation 8 and, after rearrangement, Equation 9.

Equation 7 $\hat{F}\phi_i = \epsilon_i \phi_i \quad (i=1,2,\dots,N)$

Equation 8 $\hat{F} \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} = \epsilon_i \sum_{\mu=1}^N c_{\mu i} \chi_{\mu}$

Equation 9 $\sum_{\mu=1}^N c_{\mu i} \hat{F} \chi_{\mu} = \epsilon_i \sum_{\mu=1}^N c_{\mu i} \chi_{\mu}$

Equation 10 $\mathbf{F}\mathbf{c} = \mathbf{\epsilon}\mathbf{c}$

The solution of the Hartree-Fock equations, even in the tensorial Roothaan-Hall formulation (Equation 10) (12), is not trivial, since the Fock operator is a function of the molecular orbitals, which themselves correspond to the solutions of the eigenvalues equations. The algorithm used to solve the Roothaan equations is an iterative approach known as Self Consistent Field (SCF) procedure. According to this algorithm, an initial guess for the Fock matrix is built using the molecular orbitals obtained using approximated procedures based on the Extended Huckel theory (6) or on semiempirical methods (6). Then, the Roothaan-Hall equations are solved to give atomic coefficients for each orbital and the relative energies and, consequently, the total energy. The procedure is repeated using the updated coefficients to build a new Fock matrix and a new value of energy is obtained. The new total energy value is compared to the previous one. The procedure stops when this energy difference is smaller than a chosen threshold.

In summary, the widespread HF-SCF approach relies on several approximations:

- 1) The Born-Oppenheimer approximation is inherently assumed. The true wavefunction should also be a function of the coordinates of each of the nuclei.
- 2) Typically, relativistic effects are completely neglected.
- 3) The representation of molecular orbitals is composed of a finite number of orthogonal one-electron wave-functions. The true wavefunction would need a complete (infinite) representation.
- 4) The energy eigenfunctions (wavefunctions) are assumed to be products of one-electron functions (Slater determinants). The effects of electron correlation, beyond that of exchange energy resulting from the anti-symmetry of the wavefunction, are neglected.

Nowadays more advanced approaches are preferred to HF, due to these intrinsic limitations. In fact, the missing electronic correlation effects of the HF-SCF method lead to large deviations with respect to experimental results. A number of approaches, usually denoted as post-Hartree-Fock methods, have been devised to include electron correlation in the multi-electron wave function, thus avoiding many of the HF pitfalls. One of these approaches, Møller-Plesset perturbation theory, (11, 13) treats correlation as a perturbation of the Fock operator (see below). Others methods expand the multi-electron wavefunction in terms of a linear combination of Slater determinants, such as Multi-Configurational Self Consistent Field (MC-SCF) (7, 8) Configuration Interaction (CI) (11) Complete active space SCF (CAS-SCF) (7, 8). Both MC-SCF and CAS-SCF are considered the reference methods for the study of processes involving multiple electron states. Photoinduced chemical reactions are an important application field, as exemplified by studies on vision processes initiated by rhodopsin light absorption (14) and by investigations into DNA light absorption processes (15). Another important field of applications is the study of reaction mechanisms involving low-lying excited states, a common example being the studies on cytochromes: ubiquitous and versatile catalysts involved in respiration, waste product degradation and drug excretion. (16, 17) The insights provided by QM methods (18) into cytochromes (e.g. reaction mechanism, substrate specificity) have been important for understanding the pharmacokinetics of many drugs.

4.2. Møller-Plesset perturbation theory

Møller-Plesset perturbation theory (MP) is one of the most popular post-Hartree-Fock *ab-initio* methods in quantum chemistry (the main idea was published in 1934 (13)). Rayleigh-Schrödinger perturbation theory (RS-PT) is used to improve the Hartree-Fock energy. MP2, MP3, MP4 acronyms are used to denote a truncation of the perturbation series up to the second, third, and fourth order, respectively. Although it is more computationally expensive than HF, the MP2 method is often used to treat systems where electronic correlation effects play a key role. For instance, weak interactions like dispersion (London) forces or π -stacking interactions between aromatic compounds can be properly described at the MP2 level. In several enzymes, substrate recognition and transition state stabilization are mediated by aromatic residues in the active

pocket (19), thus methods capable of accounting for such interactions can be useful in the search for inhibitors capable of displacing the natural substrate by establishing tighter interaction with the active site pocket.

4.3. Density Functional Theory (DFT)

DFT is somewhere between the *ab initio* and semiempirical approaches. According to its formalism, the basic quantity is not a many-body wavefunction, but the molecular electron density. The most common implementation of density functional theory is the Kohn-Sham (20) approach, where the intractable many-body problem of interacting electrons in a static nuclear electrostatic potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. This is composed of the nuclear potential and the effects of the interactions between the electrons, divided into exchange and correlation contributions. In Equation 11, the first term represents the electromagnetic interaction of the electron density with the external potential (it often corresponds to the Coulomb interaction between electrons and nuclei), the second represents the repulsion between the electron density and itself (Coulomb energy), the third term approximates the electronic kinetic energy (by expanding the density in a set of orbitals and computing the kinetic energy for the hypothetical system of non-interacting electrons), and the final term is the exchange-correlation functional which corrects the other terms.

Equation

11

$$E[\rho] = \int v(r)\rho(r)dr + J(\rho) + T_s(\rho) + E_{EX}(\rho)$$

The major problem within KS-DFT is the modeling of exchange and correlation interactions. The oldest and simplest approximation, called Local Density Approximation (LDA), was based on the Thomas-Fermi model of a uniform electron gas. In local exchange–correlation functionals, correlation and exchange interactions are functions which depend only on the scalar value of the electron density at a given point in space (Equation 12). A simple exchange functional (Slater) (21) is reported in Equation 13: it accounts for stabilization of high electron density correcting the electron-electron energy, which is overestimated by the Coulomb term. Combining this exchange functional with that proposed by Vosko, Wilk and Nusair leads to the SVWN functional (which is synonymous with LDA).

Equation

12

$$E_{EX}[\rho] = \int [\varepsilon_X\{\rho(r)\} + \varepsilon_C\{\rho(r)\}]\rho(r)dr$$

$$\text{Equation 13} \quad \varepsilon_X\{\rho(r)\} = -\frac{3}{2}\left(\frac{3}{\pi}\rho(r)\right)^{\frac{1}{3}}$$

Second generation exchange–correlation functionals include not only functions of the scalar density, but also functions of the gradient of the density. They are also referred to as “non-local” as the gradient introduces a certain degree of non-locality in the energy expression.

They should be more properly referred to as “gradient-corrected” (adopting the Generalized Gradient Approximation or GGA) because the exchange-correlation term is still an integral relating a defined energy contribution with the electron density of an infinitesimal volume. Since Becke’s proposal (22) of a gradient-corrected exchange functional, many attempts have been made to improve the reliability of the GGA DFT by developing correlation functionals with parameters obtained by fitting experimental data or generated to reproduce well-known physical principles. A popular GGA functional combines the Becke exchange expression with the Lee, Yang Parr correlation formulation(23), giving the commonly used BLYP functional. Other functional forms for both exchange and correlation expressions have been proposed and all their possible combinations constitute the large family of GGA DFT functionals. The next GGA functional generation, called meta-GGA, like the Tao, Perdew, Staroverov and Scuseria (TPSS),(24) include two more functions of the density, the Laplacian of the total density (or of the densities of spin-up and -down electrons) and the sum of the kinetic energy densities of the Kohn–Sham orbitals.

The observation that atomization energies (i.e. the energy required to break the molecule completely into its component atoms) are underestimated by HF method, and overestimated by both LDA and GGA methods, suggested some combined treatment to improve the results. Many sophisticated hybrid functionals have been developed. An early formulation (25) adopted a three-parameter functional where the exchange–correlation energy was expressed as a combination of the local exchange–correlation energy, the HF exchange energy, and the gradient corrections to the exchange and correlation energies as shown in Equation 14, where Becke exchange (22) and Perdew *et al.* Correlation (26) are used. Substituting the Perdew correlation with that of LYP (23) leads to the well-known B3LYP functional.

Equation

14

$$E_{XC} = E_{XC}^{LSDA} + a_0(E_X^{HF} - E_X^{LSDA}) + a_X \Delta E_X^{B88} + a_C \Delta E_C^{PW}$$

For studying large biomolecular systems, hybrid functionals like B3LYP are usually preferred to more accurate approaches like MPn or coupled cluster methods (6) because of a considerably lower computational cost, comparable to that of HF. Nevertheless, hybrid functionals have had little application for *ab initio* molecular dynamics: LDA and GGA functionals are used instead because of their lower computational cost.

The use of the DFT method has become very popular in the last two decades, including among experimental chemists. New functionals are constantly developed to improve accuracy and performance in dealing with particular classes of systems (gas clusters, metal surfaces, metal complexes, radical containing systems, unbound complexes, etc...). Due to the limited transferability of some of the new functionals, the choice of the correct functional for a given system becomes a major issue. However a large number of publications have reviewed the main application domain of the various

functionals (27-30), thus providing a valuable guide for the correct usage of the DFT method.

4.4. Dispersion force corrections in exchange-correlation functionals

DFT functionals generally provide a poor description of dispersion forces, which play a fundamental role in many aspects of chemistry and biochemistry. For instance, π -stacking interactions are often central to the correct folding of proteins, in determining packing of DNA base pairs, in enzymes’ substrate recognition and binding, etc. Current approximations of the gradient corrected XC functional prevent accurate description of π -stacking, since long-range electron correlation is not implicitly included in many functionals, which are thus unable to correctly reproduce the interactions between unbound chemical species.

Fortunately, several recent advances allow researchers to partially overcome such problems. The so-called DFT-D schemes (31, 32) can correct the DFT energy using a damped dispersion term that is able to reproduce the correct asymptotic form $-C_6R^{-6}$. Empirical corrections to DFT are a significant improvement over standard DFT functionals with no added computational cost. The functional form of the empirical dispersion term can be easily parameterized for use with any existing DFT functional, substantially increasing the domain of applications of GGA functionals. Alternatively, there is promise in the new multi-atom-centered expansion schemes introduced by Roethlisberger and co-workers for correcting the approximate XC functionals in DFT, in which the total electronic density of a system can be described as a sum of the atomic densities corrected for the interatomic interactions. This scheme (atom-centered potentials, ACPs) has proven particularly efficient when applied to the correction of dispersion forces in DFT (33-36). Dispersion-corrected atom-centered potentials (DCACPs) can reproduce the asymptotic R^{-6} behavior when a sufficiently complete basis set is used. A set of DCACPs is currently available for all the biologically relevant elements (37).

Other approaches involve reshaping the form of the exchange and correlation functionals, introducing new parameters taken from experiments or higher level computations, varying the amount of included HF exchange etc. A new family of functionals (MPW1B95, MPWB1K, M05, M05-2X, M06-2X, M08-HX and M08-SO) developed by Truhlar and co-workers (38-41) is proving particularly effective for studying biologically relevant systems (42, 43). A different approach is behind the family of functionals developed by Grimme (44). They comprise a mixture of DFT components and the MP2 correlation energy calculated with the DFT orbitals. Grimme referred to his functional as B2PLYP (45-48) (B88 exchange, 2 fitted parameters and perturbative mixture of MP2 and LYP); a version with improved performance (especially for weak interactions) is mPW2PLYP.(49) From the extensive calibration work, the new functionals appear to give better energetics and a narrower error distribution than B3LYP. Thus, the additional cost of the calculation of the MP2 energy may be well invested.

4.5. Semi-empirical methods

Semi-empirical (SE) methods such as AM1 (50), PM3 (51), PM6 (52) and RM1 (53) are based on HF formalism but contain several approximations and include some parameters from empirical data and some from higher level calculations. Because of their extremely low computational cost, SE methods have been used to study large molecules and are still used to investigate the dynamics of organic and bioorganic systems. Their general performance is affected by the approximations adopted to decrease the computational cost; still, for some selected systems they provide reliable results due to the particular parameterization. Nevertheless, the use of SE methods, even for large systems, is steadily decreasing as computational resources and code optimization progress allow the use of more reliable and transferable methods like DFT.

4.6. Density Functional Tight-Binding

DFTB is an alternative approach to the quantum chemical semi-empirical (6) methods. It corresponds to an approximate DFT scheme and is characterized by a computational speed similar to that of traditional semi-empirical quantum chemical methods (like AM1, and PM3), but requiring a smaller number of empirical parameters. The approximate DFTB method is derived from DFT by a second-order expansion of the DFT Kohn-Sham total energy with respect to charge density fluctuations. The zero-order approach is equivalent to a common standard non-self-consistent (Tight-Binding: TB) scheme, while at the second order, a transparent, parameter-free and readily calculable expression for generalized Hamiltonian matrix elements may be derived. DFTB can be seen as a tight binding method, parameterized from DFT. DFTB was augmented by a self-consistency treatment based on atomic charges in the so-called self-consistent charge density functional tight-binding (SCC-DFTB) method (54-57). The SCC-DFTB method has been applied to a large variety of problems in chemistry (58-60), physics (61, 62), materials science (54, 63), and biochemistry (64-70). During the past decade, it has been continually developed to improve its accuracy and transferability.

5. MOLECULAR MECHANICS APPROACH

Computational approaches based on Molecular Mechanics (MM) allow researchers to compute the energy and properties of large molecular systems. The total energy of a system is simply derived from the sum of different energetic contributions (additive force field). The MM Hamiltonian (6-8) thus comprises several terms, each one taking into account the contribution arising from the system's bonding (stretching, bending, and torsional) and non-bonding interactions (van der Waals, Coulombic) (see Equation 15). Both stretching and bending contributions (first and second terms in Equation 15) are expressed by simple harmonic potentials. The torsional contribution is described by a cosine series (third term) to account for multiple conformational minima. The classical 6-12 Lennard-Jones potential (6) and the Coulombic function have been used to describe, respectively, the van der Waals

and Coulombic non-bonding interactions (fourth and fifth terms of Equation 15). In the MM force field, the equilibrium bond lengths and angles (r_{eq} , ϑ_{eq}) and the spring constants (k_r , k_ϑ) are parameterized to reproduce experimental frequencies of specific sets of molecules, V_n is fitted to reproduce *ab initio* energies, A_{ij} and B_{ij} are obtained from *Monte Carlo*(6-8) simulations, and the atomic point charges (q_i , q_j) are derived from *ab initio* calculation and subsequent RESP (71) fitting (72) or by other procedures.

Equation

15

$$E_{tot} = \sum_{bonds} k_r (r - r_{eq})^2 + \sum_{bonds} k_\vartheta (\vartheta - \vartheta_{eq})^2 + \sum_{bonds} \frac{V_n}{2} [1 + \cos(n\varphi - \gamma)] + \sum_{i < j} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \sum_{i < j} \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}}$$

To obtain reliable values from MM calculations, a new atom definition has been adopted: all atoms in a molecule are classified as different atom types not only on the basis of the atomic number (as in QM approaches), but also according to their chemical environment (bonded atoms). This has led to the development of different parameters for aliphatic and aromatic carbon atoms, for carbonyl or alcoholic oxygen atoms, and so on. A particular MM Force Field (73) is defined on the basis of the adopted functional form for the energy expression and the specific values of the chosen parameters for the available atom types. Force fields like OPLS (74), AMBER (75-77), GROMOS (78) and CHARMM (79) are widely used to study standard biomolecular systems (protein, DNA, and RNA), while corrections and extensions are continually developed and improved to treat carbohydrates, lipids, enzyme cofactors, and other molecules of relevant biological interest, including drugs and ions. Particularly relevant in the field of drug design are both the Generalized AMBER Force Field (GAFF) (80) and CHARMM General Force Field (CgFF) (81). This is because they provide suitable parameters for drug-like organic molecules, which can be readily parameterized for virtual screening studies. MM potentials are in fact widely used in molecular docking protocols (82-85) devised to score the conformation of ligands within the active site of an enzyme. The small computational cost of these potentials allows for fast screening of large libraries of chemical compounds. Moreover, MM potentials, when coupled to advanced simulation techniques (such as free energy calculations), can be efficiently employed in the lead optimization step, which is aimed at optimizing the potency of drug candidates for specific targets (86-88).

6. HYBRID QUANTUM MECHANICS/MOLECULAR MECHANICS METHODS

Along the challenging pathway leading to the development of new drugs, the study of the structure and

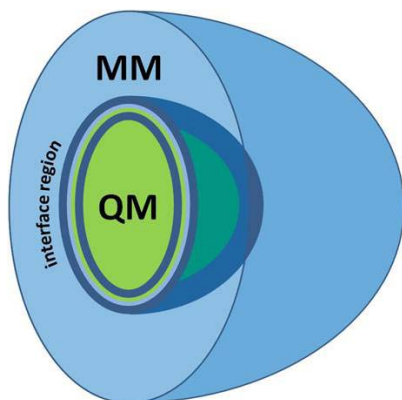


Figure 1. QM/MM partitioning scheme. Schematic representation of a system partitioned into two regions treated at the QM and MM level, respectively.

function of the targeted proteins is of paramount importance. While MM approaches are useful when investigating structural features of very large systems, they cannot be used to study reactive processes (when bond topology changes), as those take place at the enzymatic active site. In contrast, QM approaches, capable of dealing with both thermodynamics and kinetics of chemical reactions, cannot be used for large systems, like those of biological interest, because of their excessive computational demand. For many years, a partial solution involved building model systems comprising the sole reactive region and the nearest surrounding atoms. Using this approach, the reaction mechanism of several enzymatic reactions was elucidated, but the results had only a qualitative value since most of the environmental effects of the enzyme were overlooked and excluded from the model.

Recently, the most commonly used approach to overcoming such limitations is the combination of DFT (or other QM methods) with classical force-field-based molecular mechanics (MM) in the so-called quantum mechanics/molecular mechanics (QM/MM) schemes. The system is partitioned into two or more regions treated with different levels of theory (Figure 1). The chemically relevant part of the system may be described by QM, while the remaining portion is treated with faster MM methods. The choice of the QM region is not trivial as solvent molecules, ions and parts of protein/DNA may be directly involved in the process of interest. In addition, a proper treatment of the long-range electrostatic interactions between the QM and MM regions is crucial for obtaining a reliable approach. Originally proposed in 1976 by Warshel and Levitt, (89) many different QM/MM implementations have flourished in the last two decades, with such hybrid approaches successfully employed in both biological and material science fields (89-97). The gathered experience allowed researchers to define the most important issues that a QM/MM scheme should address to provide reliable results, with the treatment of the QM-MM boundary region being considered the most controversial (98-101). A more technical issue connected with the efficiency of QM/MM schemes is the actual implementation of the hybrid potential; several QM/MM implementations are constituted

by MM extensions to existing QM codes such as the ONIOM (102) implementation included in the GAUSSIAN (103) suite of programs, or the hybrid approaches included in CPMD (104) and CP2K (105). MM codes also feature interfaces to (or embed) QM algorithms (often semiempirical), as is the case with AMBER (68, 106, 107), CHARMM (108). A somewhat less efficient, but more flexible approach is provided by the “pipe” or “interface” codes, capable of interfacing existing and well-tested QM and MM software by means of light procedures for preparing the inputs for the chosen QM and MM programs, launching, and then combining the output of the employed codes. This approach is used in the COBRAMM (109-111), CHEMSHELL(112), COMQUM (113) and PUPIL (114) software, whose scope ranges from material science to drug discovery. QM/MM approaches have been successfully applied to the various phases of the drug design process (115-117), providing a better understanding of enzyme reaction mechanisms (118-122) and catalytic proficiency (109, 110). More recently, hybrid approaches have been integrated into computational protocols devised for docking (85, 123, 124) and computing the binding affinity of drugs (85, 125, 126), thus providing a useful tool for *in silico* screening of lead candidates (127-130).

7. METALLOENZYMES: A CHALLENGE FOR QM METHODS

Metal ions interact with cellular components and their role is of paramount importance e.g., in assuring structural stability, inducing conformational rearrangements, and conferring proper functionality in biological catalysis (metalloenzymes). Metalloenzymes are widespread proteins, ubiquitous in all life kingdoms, being involved in various biosynthetic processes. α -Carbonic anhydrases (CAs), matrix metalloproteinases (MMP), and angiotensin-converting enzyme (ACE), among others, are clinically exploited targets in the treatment or prevention of a variety of diseases such as congestive heart failure, hypertension, glaucoma, epilepsy, and cancer (131, 132). Other members of the family, like metallo- β -lactamases (MBL), are promising targets for developing new antibacterial drugs capable of avoiding bacterial resistance insurgence (133). Heme-group-containing enzymes are an important subgroup of the metalloenzyme family; cytochromes (16, 17), among others, have been investigated in detail (18, 134-136) to elucidate the degradation pathway of both endogenous metabolites and drug molecules, with the aim of rationally improving the pharmacokinetics of lead compounds. QM and MM methods, as well as the recent QM/MM hybrid techniques, have been used to study several metalloenzymes to better understand their catalytic features (125, 137-141) or to help in the lead optimization phase of structure-based drug design (85, 116, 117, 119, 126, 142-144). The investigation of metalloenzymes is a major challenge for computational methods because many metals can easily undergo spin state changes, allowing for different ligand arrangements. The polarizability in metal complexes is barely addressed by even the most sophisticated MM force fields. Several QM methods can be used to investigate metals' peculiar behavior, but at an increased computational cost. For this

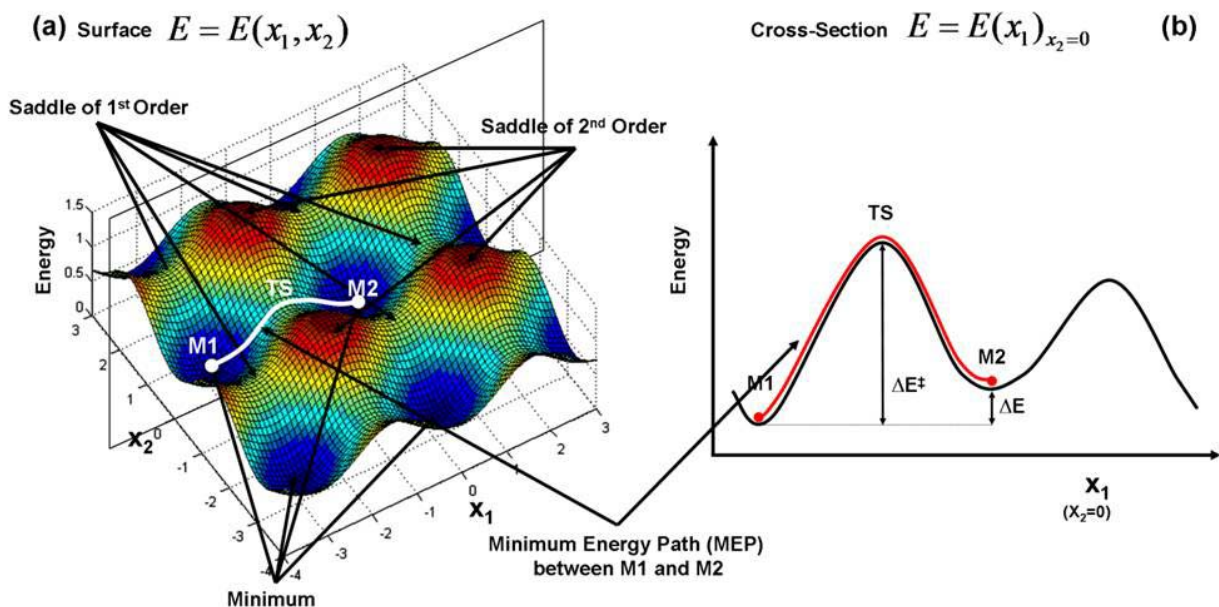


Figure 2. Potential Energy Surface representation. a) Graphical representation of a three-dimensional energy surface; b) bi-dimensional cross-section. The Minimum Energy Path (MEP) connecting the two critical points M1 and M2 is shown.

reason, many theoretical and computational chemists continue to focus on developing new techniques capable of capturing the elusive behavior of metalloenzymes. For these techniques to be used in the field of drug development, particular care must be devoted to the general usability of the methods and to the cost/accuracy balance.

8. THE POTENTIAL ENERGY SURFACE

The energy of a molecular system, computed at the QM, MM, or QM/MM level for a particular atomic arrangement, is often referred to as “single point energy”. Since the absolute energy has little physical meaning, it is common practice to compare the energy of two or more atomic configurations, since energy differences are related to physical observables. The Potential Energy Surface (PES) accounts for the dependency of the energy of a system with respect to its atomic configuration; it collects the potential energy of all the possible atomic configurations of a given system. In the Cartesian coordinate system $3N$ coordinates are required to describe a system comprising N atoms, since each atom is described by 3 coordinates that identify its position with respect to an arbitrary point (origin of the coordinate system). The internal coordinate system can be obtained by choosing $3N-6$ linearly independent coordinates (rotation and translation degrees of freedom are eliminated) that coincide with bond lengths, angles (plane angles) and dihedral angles (solid angles) between atoms (6, 8). The PES can be thus defined in terms of Cartesian or Internal coordinate systems.

Only a few points on the PES landscape have a chemical and physical significance. In particular, the points

with null first derivatives of the energy with respect to all the n coordinates are of great interest. Those characterized by all positive second derivatives can be either one of the many local minima or the unique global minimum of the energy, corresponding, respectively, to meta-stable nuclear configurations or to the most stable configuration. Points with null first derivatives and negative second derivatives with respect to $k < n$ coordinates are denoted as saddle point of index k . They have a chemical significance only if $k=1$, in which case they correspond to “transition structures”. An elementary reaction step is described as a transition from one equilibrium state (minimum) to a neighboring one *via* a single transition state. The reaction mechanism is given by the sequence of steps involved in a chemical process and corresponds on the PES to the Minimum Energy Path (MEP) connecting the two minima that represent reactants and products, respectively (Figure 2).

The localization of critical points on the PES is the main goal of many computational procedures aimed at studying the stability or reactivity of a system. This is because minima and first order saddles represent, respectively, stable nuclear arrangement and transition structures.

Since the PES complexity rapidly increases with the increasing number of coordinates, the search algorithm is crucial to locating critical points. Efficient algorithms for locating equilibrium and transition structures are now available in modern molecular software. These algorithms are based on the calculation of the first (gradient) and second derivatives (Hessian matrix) and allow researchers to perform a simultaneous optimization of the whole set of

coordinates. In general, the search algorithm is iterative and the geometry is gradually modified till the desired critical point is obtained. The fundamental equations for computing the coordinate variations at each step of the search procedure can be derived by assuming a quadratic shape of the PES (6-8). Should the surface be a real quadratic surface, the desired critical point will be obtained in a single step. In most cases, however, the surface is far from being quadratic, so a sequence of motions on the surface (optimization step) is usually required to locate the critical point. Because the calculation of the Hessian matrix is computationally expensive, approximate forms of the Newton-Raphson equations (involving approximated Hessian matrices) are usually employed. These methods are often referred to as “quasi-Newton” methods. The “steepest descent” method is far from being accurate but it is very fast and can be efficiently used to decide the first moves on a non-quadratic region of the PES, far away from the critical point. Then, in the vicinity of the critical point, the search algorithm can be switched to the Newton-Raphson method or to a more accurate quasi-Newton scheme. Examples of quasi-Newton methods are the widely used BSGF (145-148) scheme (Broyden-Fletcher-Goldfarb-Shanno), the L-BFGS (149, 150) scheme (Limited memory-BFGS), and the PSB (151) scheme (Powell-Symmetric-Broyden). Extensive exploration of a system’s configurational space is thus avoided by using an appropriate optimization algorithm, capable of locating critical points by computing the energy and gradient of a small number of atomic configurations.

9. THE FREE ENERGY SURFACE

The potential energy does not take into account the temperature of the system, thus it should not be related directly to physical observables like kinetic and equilibrium constants. Kinetics and thermodynamics of physical and chemical processes are directly related to free energy differences. In particular, the Gibbs free energy is defined when both pressure and temperature are constant, as in most physical systems under standard experimental conditions. The Gibbs free energy G is defined as the sum of an enthalpic term H , accounting for the potential contributions, and an entropic term $-TS$ (i.e. $G=H-TS$).

Computing the Free Energy Surface (FES) for a given system is an extremely difficult task, and several strategies based on either MD or Monte-Carlo techniques have been devised to compute free energy differences between states or free energy changes associated with processes. In all cases, the time required to explore and characterize the relevant portion of the FES is orders of magnitude greater than that required for PES exploration. This is because a large number of energy and gradient evaluations are required for statistical reasons. MD-based methods coupled to different levels of theory (e.g. HF, DFT, etc.) and different schemes (e.g. QM/MM) are valuable solutions for exploring the FES of a given system. Nevertheless, the time required to spontaneously explore the relevant portions of the FES is still out of reach for large systems of biological and pharmaceutical relevance, and for current software and hardware performance.

Enhanced sampling techniques are thus required to speed up calculations. Some of these procedures require the definition of one or more reaction coordinates so as to restrict the exploration of the phase space to a smaller, more interesting region chosen as representative of the process in question. Steered (SMD) and targeted (TMD) molecular dynamics (152, 153), umbrella sampling (US) (154-156) have been successfully used in the past to study both protein conformational changes and enzymatic reactions (6). More recent techniques like conformational flooding (CF) (157) and metadynamics (MMD) (158-160) do not require prior knowledge of the reaction coordinate components, allowing an unbiased investigation of the FES. In particular, the use of metadynamics schemes within the framework of QM/MM simulations have recently proven to be an effective protocol for investigating the reaction mechanism of potential drug target enzymes (161-163).

Within this framework, hybrid QM/MM (104, 105, 164) MD simulations based on the Car-Parrinello (CP) (165, 166) treatment of the electronic degrees of freedom has emerged as an efficient method for characterizing the reaction mechanism of several enzymes of pharmaceutical relevance (137, 138, 167-169). The CP scheme, as implemented in the CPMD code (165, 170), was originally developed for the study of condensed matter systems. In the past decade, it has frequently been used to treat isolated systems of biological relevance (171). In the DFT framework, the CP method allows researchers to describe the dynamics of the ground state, in which the electrons adiabatically follow the nuclear degrees of freedom. The advantage of this method over Born-Oppenheimer (BO) MD is that the wavefunction does not need to be optimized at every MD step, providing a considerable computational gain (for further details on the CP and BO methods, see the review by Marx and Hutter) (172). Recently, efficient optimization schemes (173) have been developed, which quickly converge the wavefunction of the system. This allows the use of BO-MD as alternative way of studying enzymatic catalysis in a QM/MM setup (e.g. see the implementation of BOMD in CP2K (166, 174) and the recent studies published using this scheme (161)).

10. FUTURE PERSPECTIVES

In recent decades, computational approaches have played an increasingly significant role in the field of drug development thanks to their improved reliability. In particular, QM and QM/MM-based protocols are today routinely used to investigate the reaction mechanism and catalytic proficiency of potential drug targets. CPU-demanding simulations are thus used to obtain structural and electronic information about reaction intermediates and transition states so that researchers can design analogue molecules that act as inhibitors. The ongoing effort of many research groups in academia and industry continues to improve the reliability and cost/efficiency ratio of QM-based approaches. QM/MM potentials have been designed to investigate biological systems without introducing unphysical truncations near the active site; thus, QM/MM hybrid schemes offer noticeable improvements over simple

QM treatment of small portions of the system, and have been used in recent years to study a large variety of biological systems. Researchers have thus been forced to test and improve the accuracy of QM methods and often to develop new theoretical approaches to cope with the diversity of chemical reactions and substrates encountered in biology. The need for a better description of metalloenzymes required accurate tests of old DFT functionals (28) and led to the development of novel strategies for studying transition metal complexes (175). Additionally, the increasing importance attributed to weak interactions in both complex stability and reactive processes is driving research into new QM methods capable of correctly describing noncovalent interactions at a reasonable computational cost.(43) Nonetheless, the high demand in terms of simulation time and hardware requirements remains a limiting factor for QM methods, especially in the field of drug design. Thus, many researchers have focused on developing new theories and better approximations to improve performance without impairing the accuracy of results. The RI approximation, (176) for example, is a recent breakthrough, which allows DFT and even MP2 to be efficiently used to study biological systems. While theoreticians develop more accurate methods and computer scientists implement these in efficient algorithms, developments in hardware are poised to herald a new era for QM-based methods. Since many available QM applications are so CPU-demanding (central processing unit), several strategies have been devised to efficiently share the load of a single calculation over a large number of CPUs. Algorithms were written to work in parallel on many processor units and wiring systems, which had been devised to allow for efficient communication between different nodes of a cluster. More recently, graphical processing units (GPU), formerly used only for graphical applications in professional design and video-gaming, have started to be used for general computing applications, benefiting from the evident advantages in terms of parallelization and hardware cost. Thus, QM algorithms have recently been rewritten using new programming languages designed for new-generation GPUs. Two-electron integral evaluation (177) and SCF procedure (178), both at the heart of many *ab initio* QM methods, have been shown to be faster on GPUs than on state-of-the-art CPUs by one to two orders of magnitude, opening the way for GPU-specific implementation of HF (179), DFT (180) and even RI-MP2 (181). These promising results are expected to steadily improve with the continuous progress in the GPU-based hardware and software. Soon, they will significantly affect the way we use QM-based methods to investigate biological systems and to rationally design more effective drugs.

11. REFERENCES

1. Markman, M. and Peereboom, D. M.: From serendipity to design: the evolution of drug development in oncology. *Cleve Clin J Med.*, 64(3), 155-63 (1997)
2. Anderson, A. C.: The Process of Structure-Based Drug Design. *Chem. Biol.*, 10(9), 787-797 (2003)
3. Colman, P. M.: Structure-based drug design. *Curr. Opin. Struct. Biol.*, 4(6), 868-874 (1994)
4. Guida, W. C.: Software for structure-based drug design. *Curr. Opin. Struct. Biol.*, 4(5), 777-781 (1994)
5. Jackson, R. C.: Update on computer-aided drug design. *Curr. Opin. Biotechnol.*, 6(6), 646-651 (1995)
6. Leach, A. R.: Molecular Modelling: Principles and Applications. Pearson Education EMA, UK (2001)
7. Cramer, C. J.: Essentials of Computational Chemistry: Theories and Models. (2004)
8. Jensen, F.: Introduction to Computational Chemistry. John Wiley & Sons, UK (1999)
9. Born, M. and Oppenheimer, R.: On the Quantum Theory of Molecules. *Ann. Physic.*, 84, 457-484 (1927)
10. Boys, S. F. *Proc. R. Soc. London*, 200, 542 (1950)
11. Szabo, A. and Ostlund, N. S.: Modern Quantum Chemistry. Dover Publications, INC, New York (1996)
12. Roothaan, C. C. J.: New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.*, 23(2), 69-89 (1951)
13. Möller, C. and Plesset, M. S.: Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.*, 46(7), 618-622 (1934)
14. Tomasello, G., Olaso-González, G., Altoè, P., Stenta, M., Serrano-Andrés, L., Merchán, M., Orlandi, G., Bottioni, A. and Garavelli, M.: Electrostatic Control of the Photoisomerization Efficiency and Optical Properties in Visual Pigments: On the Role of Counterion Quenching. *J. Am. Chem. Soc.*, 131(14), 5172-5186 (2009)
15. Conti, I., Altoe, P., Stenta, M., Garavelli, M. and Orlandi, G.: Adenine deactivation in DNA resolved at the CASPT2//CASSCF/AMBER level. *Phys. Chem. Chem. Phys.*, 12(19), 5016-5023 (2010)
16. Coon, M. J.: Cytochrome P450: Nature's Most Versatile Biological Catalyst. *Annu. Rev. Pharmacol. Toxicol.*, 45(1), 1-25 (2005)
17. Ow, Y.-L. P., Green, D. R., Hao, Z. and Mak, T. W.: Cytochrome c: functions beyond respiration. *Nat. Rev. Mol. Cell Biol.*, 9(7), 532-542 (2008)
18. Altun, A., Kumar, D., Neese, F. and Thiel, W.: Multireference Ab Initio Quantum Mechanics/Molecular Mechanics Study on Intermediates in the Catalytic Cycle of Cytochrome P450_{cam}. *J. Phys. Chem. A*, 112(50), 12904-12910 (2008)
19. Leach, A. G., Houk, K. N. and Raymond, J. L.: Theoretical Investigation of the Origins of Catalysis of a

- Retro-Diels-Alder Reaction by Antibody 10F11. *J. Org. Chem.*, 69(11), 3683-3692 (2004)
20. Kohn, W. and Sham, L. J.: Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A), A1133 (1965)
21. Slater, J. C.: A Simplification of the Hartree-Fock Method. *Phys. Rev.*, 81(3), 385 (1951)
22. Becke, A. D.: Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.*, 38(6), 3098-3100 (1988)
23. Lee, C., Yang, W. and Parr, R. G.: Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.*, 37(2), 785 (1988)
24. Tao, J., Perdew, J. P., Staroverov, V. N. and Scuseria, G. E.: Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.*, 91(14), 146401 (2003)
25. Becke, A. D.: Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7), 5648-5652 (1993)
26. Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J. and Fiolhais, C.: Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B: Condens. Matter Mater. Phys.*, 46(11), 6671 (1992)
27. Ziegler, T.: Approximate density functional theory as a practical tool in molecular energetics and dynamics. *Chem. Rev.*, 91(5), 651-667 (1991)
28. Harvey, J. N.: On the accuracy of density functional theory in transition metal chemistry. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 102, 203-226 (2006)
29. Bernardi, F., Bottoni, A. and Garavelli, M.: Exploring Organic Chemistry with DFT: Radical, Organo-metallic, and Bio-organic Applications. *Quant. Struct. - Act. Rel.*, 21(2), 128-148 (2002)
30. Geerlings, P., De Proft, F. and Langenaeker, W.: Conceptual density functional theory. *Chem. Rev.*, 103(5), 1793-1873 (2003)
31. Grimme, S.: Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.*, 25(12), 1463-1473 (2004)
32. Grimme, S.: Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27(15), 1787-1799 (2006)
33. von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U. and Sebastiani, D.: Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory. *Phys. Rev. Lett.*, 93(15), 153004 (2004)
34. Tapavicza, E., Lin, I. C., von Lilienfeld, O. A., Tavernelli, I., Coutinho-Neto, M. c. D. and Rothlisberger, U.: Weakly Bonded Complexes of Aliphatic and Aromatic Carbon Compounds Described with Dispersion Corrected Density Functional Theory. *J. Chem. Theory Comput.*, 3(5), 1673-1679 (2007)
35. Lin, I. C., von Lilienfeld, O. A., Coutinho-Neto, M. c. D., Tavernelli, I. and Rothlisberger, U.: Predicting Noncovalent Interactions between Aromatic Biomolecules with London-Dispersion-Corrected DFT. *J. Phys. Chem. B*, 111(51), 14346-14354 (2007)
36. Lin, I. C. and Rothlisberger, U.: Describing weak interactions of biomolecules with dispersion-corrected density functional theory. *Phys. Chem. Chem. Phys.*, 10(19), 2730-2734 (2008)
37. Lin, I. C., Coutinho-Neto, M. D., Felsenheimer, C., von Lilienfeld, O. A., Tavernelli, I. and Rothlisberger, U.: Library of dispersion-corrected atom-centered potentials for generalized gradient approximation functionals: Elements H, C, N, O, He, Ne, Ar, and Kr. *Phys. Rev. B: Condens. Matter Mater. Phys.*, 75(20), 205131 (2007)
38. Becke, A. D.: Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.*, 104(3), 1040-1046 (1996)
39. Adamo, C. and Barone, V.: Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. *J. Chem. Phys.*, 108(2), 664-675 (1998)
40. Zhao, Y. and Truhlar, D. G.: Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van der Waals Interactions. *J. Phys. Chem. A*, 108(33), 6908-6918 (2004)
41. Zhao, Y., Schultz, N. E. and Truhlar, D. G.: Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.*, 2(2), 364-382 (2006)
42. Zhao, Y. and Truhlar, D. G.: Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.*, 4(11), 1849-1868 (2008)
43. Zhao, Y. and Truhlar, D. G.: Density Functionals for Noncovalent Interaction Energies of Biological Importance. *J. Chem. Theory Comput.*, 3(1), 289-300 (2006)

44. Grimme, S.: Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.*, 124(3), 034108 (2006)
45. De Proft, F. and Geerlings, P.: Conceptual and Computational DFT in the Study of Aromaticity. *Chem. Rev.*, 101(5), 1451-1464 (2001)
46. Schwabe, T. and Grimme, S.: Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability. *Phys. Chem. Chem. Phys.*, 9(26), 3397-3406 (2007)
47. Grimme, S. and Neese, F.: Double-hybrid density functional theory for excited electronic states of molecules. *J. Chem. Phys.*, 127(15), 154116 (2007)
48. Neese, F., Schwabe, T. and Grimme, S.: Analytic derivatives for perturbatively corrected "double hybrid" density functionals: Theory, implementation, and applications. *J. Chem. Phys.*, 126(12), 124115 (2007)
49. Schwabe, T. and Grimme, S.: Towards chemical accuracy for the thermodynamics of large molecules: new hybrid density functionals including non-local correlation effects. *Phys. Chem. Chem. Phys.*, 8(38), 4398-4401 (2006)
50. Dewar, M. J. S., Zebisch, E. G., Healy, E. F. and 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.*, 107(13), 3902-3909 (1985)
51. Stewart, J. J. P.: Optimization of Parameters for Semi-Empirical Methods I-Method. *J. Comp. Chem.*, 10, 209-220 (1989)
52. Stewart, J.: Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.*, 13(12), 1173-1213 (2007)
53. Rocha, G. B., Freire, R. O., M. Simas, A. and Stewart, J. J. P.: RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J. Comput. Chem.*, 27(10), 1101-1111 (2006)
54. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S. and Seifert, G.: Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B: Condens. Matter Mater. Phys.*, 58(11), 7260 (1998)
55. Seifert, G.: Tight-Binding Density Functional Theory: An Approximate Kohn-Sham DFT Scheme. *J. Phys. Chem. A*, 111(26), 5609-5613 (2007)
56. Otte, N., Scholten, M. and Thiel, W.: Looking at Self-Consistent-Charge Density Functional Tight Binding from a Semiempirical Perspective. *J. Phys. Chem. A*, 111(26), 5751-5755 (2007)
57. Pu, J., Gao, J. and Truhlar, D. G.: Combining Self-Consistent-Charge Density-Functional Tight-Binding (SCC-DFTB) with Molecular Mechanics by the Generalized Hybrid Orbital (GHO) Method. *J. Phys. Chem. A*, 108(25), 5454-5463 (2004)
58. Cai, Z. L., Lopez, P., Reimers, J. R., Cui, Q. and Elstner, M.: Application of the Computationally Efficient Self-Consistent-Charge Density-Functional Tight-Binding Method to Magnesium-Containing Molecules. *J. Phys. Chem. A*, 111(26), 5743-5750 (2007)
59. Zheng, G., Witek, H. A., Bobadova-Parvanova, P., Irle, S., Musaev, D. G., Prabhakar, R., Morokuma, K., Lundberg, M., Elstner, M., Kohler, C. and Frauenheim, T.: Parameter Calibration of Transition-Metal Elements for the Spin-Polarized Self-Consistent-Charge Density-Functional Tight-Binding (DFTB) Method: Sc, Ti, Fe, Co, and Ni. *J. Chem. Theory Comput.*, 3(4), 1349-1367 (2007)
60. Heine, T., DosSantos, H. F., Patchkovskii, S. and Duarte, H. A.: Structure and Dynamics of β -Cyclodextrin in Aqueous Solution at the Density-Functional Tight Binding Level. *J. Phys. Chem. A*, 111(26), 5648-5654 (2007)
61. Hu, H., Lu, Z., Elstner, M., Hermans, J. and Yang, W.: Simulating Water with the Self-Consistent-Charge Density Functional Tight Binding Method: From Molecular Clusters to the Liquid State. *J. Phys. Chem. A*, 111(26), 5685-5691 (2007)
62. Reimers, J. R., Solomon, G. C., Gagliardi, A., Bilic, A., Hush, N. S., Frauenheim, T., DiCarlo, A. and Pecchia, A.: The Green's Function Density Functional Tight-Binding (gDFTB) Method for Molecular Electronic Conduction. *J. Phys. Chem. A*, 111(26), 5692-5702 (2007)
63. Frauenheim, T., Seifert, G., Elstner, M., Niehaus, T., hler, C., Amkreutz, M., Sternberg, M., Hajnal, Z., Carlo, A. D., Suhai, S. and ndor: Atomistic simulations of complex materials: ground-state and excited-state properties. *J. Phys.: Condens. Matter*, 14(11), 3015-3047 (2002)
64. Bhattacharyya, S., Stankovich, M. T., Truhlar, D. G. and Gao, J.: Combined Quantum Mechanical and Molecular Mechanical Simulations of One- and Two-Electron Reduction Potentials of Flavine Cofactor in Water, Medium-Chain Acyl-CoA Dehydrogenase, and Cholesterol Oxidase. *J. Phys. Chem. A*, 111(26), 5729-5742 (2007)
65. Kubar, T., Jurecka, P., Cerny, J., Rezac, J., Otyepka, M., Valdes, H. and Hobza, P.: Density-Functional, Density-Functional Tight-Binding, and Wave Function Calculations on Biomolecular Systems. *J. Phys. Chem. A*, 111(26), 5642-5647 (2007)
66. Riccardi, D. and Cui, Q.: pKa Analysis for the Zinc-Bound Water in Human Carbonic Anhydrase II: Benchmark for "Multiscale" QM/MM Simulations and Mechanistic Implications. *J. Phys. Chem. A*, 111(26), 5703-5711 (2007)

67. Riccardi, D., Li, G. and Cui, Q.: Importance of van der Waals Interactions in QM/MM Simulations. *J. Phys. Chem. B*, 108(20), 6467-6478 (2004)
68. Seabra, G. M., Walker, R. C., Elstner, M., Case, D. A. and Roitberg, A. E.: Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package. *J. Phys. Chem. A*, 111(26), 5655-5664 (2007)
69. Woodcock, H. L., Hodoscek, M. and Brooks, B. R.: Exploring SCC-DFTB Paths for Mapping QM/MM Reaction Mechanisms. *J. Phys. Chem. A*, 111(26), 5720-5728 (2007)
70. Xu, D., Guo, H. and Cui, Q.: Antibiotic Binding to DZinc β -Lactamase L1 from *Stenotrophomonas maltophilia*: SCC-DFTB/CHARMM and DFT Studies. *J. Phys. Chem. A*, 111(26), 5630-5636 (2007)
71. Bayly, C. I., Cieplak, P., Cornell, W. and Kollman, P. A.: A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, 97(40), 10269-10280 (1993)
72. Cieplak, P., Cornell, W. D., Bayly, C. and Kollman, P. A.: Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.*, 16(11), 1357-1377 (1995)
73. Jorgensen, W. L. and Tirado-Rives, J.: Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U. S. A.*, 102(19), 6665-6670 (2005)
74. Jorgensen, W. L., Maxwell, D. S. and Tirado-Rives, J.: Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, 118(45), 11225-11236 (1996)
75. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. and Kollman, P.: A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.*, 24(16), 1999-2012 (2003)
76. Ponder, J. W., Case, D. A. and Valerie, D.: Force Fields for Protein Simulations. In: *Advances in Protein Chemistry*. Academic Press, (2003)
77. Dudek, M. J. and Ponder, J. W.: Accurate modeling of the intramolecular electrostatic energy of proteins. *J. Comput. Chem.*, 16(7), 791-816 (1995)
78. Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Kräutler, V., Oostenbrink, C., Peter, C., Trzesniak, D. and van Gunsteren, W. F.: The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem.*, 26(16), 1719-1751 (2005)
79. MacKerell, A. D., Brooks, B., Brooks, C. L., Nilsson, L., Roux, B., Won, Y. and Karplus, M.: CHARMM: The Energy Function and Its Parameterization with an Overview of the Program. In: *The Encyclopedia of Computational Chemistry*. Ed P. v. R. S. e. al. John Wiley & Sons, Chichester (1998)
80. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. and Case, D. A.: Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9), 1157-1174 (2004)
81. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I. and Mackerell, A. D. J.: CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.*, 31(4), 671-690
82. Alonso, H., Bliznyuk, A. A. and Gready, J. E.: Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.*, 26(5), 531-568 (2006)
83. Beierlein, F., Lanig, H., Schrer, G., Horn, A. H. C. and Clark, T.: Quantum mechanical/molecular mechanical (QM/MM) docking: an evaluation for known test systems. *Mol. Phys.*, 101(15), 2469 - 2480 (2003)
84. Gervasio, F. L., Laio, A. and Parrinello, M.: Flexible Docking in Solution Using Metadynamics. *J. Am. Chem. Soc.*, 127(8), 2600-2607 (2005)
85. Khandelwal, A., Lukacova, V., Comez, D., Kroll, D. M., Raha, S. and Balaz, S.: A Combination of Docking, QM/MM Methods, and MD Simulation for Binding Affinity Estimation of Metalloprotein Ligands. *J. Med. Chem.*, 48(17), 5437-5447 (2005)
86. Jorgensen, W. L. and Thomas, L. L.: Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.*, 4(6), 869-876 (2008)
87. Zeevaart, J. G., Wang, L., Thakur, V. V., Leung, C. S., Tirado-Rives, J., Bailey, C. M., Domoal, R. A., Anderson, K. S. and Jorgensen, W. L.: Optimization of Azoles as Anti-Human Immunodeficiency Virus Agents Guided by Free-Energy Calculations. *J. Am. Chem. Soc.*, 130(29), 9492-9499 (2008)
88. Kim, J. T., Hamilton, A. D., Bailey, C. M., Domoal, R. A., Wang, L., Anderson, K. S. and Jorgensen, W. L.: FEP-Guided Selection of Bicyclic Heterocycles in Lead Optimization for Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *J. Am. Chem. Soc.*, 128(48), 15372-15373 (2006)
89. Warshel, A. and Levitt, M.: Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric

stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103(2), 227-249 (1976)

90. Sherwood, P.: Hybrid Quantum Mechanics/Molecular Mechanics Approaches. *NIC series*, 3, 285-305 (2000)

91. Gao, J. and Truhlar, D. G.: Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.*, 53(1), 467-505 (2002)

92. Vreven, T., Mennucci, B., da Silva, C. O., Morokuma, K. and Tomasi, J.: The ONIOM-PCM method: Combining the hybrid molecular orbital method and the polarizable continuum model for solvation. Application to the geometry and properties of a merocyanine in solution. *J. Chem. Phys.*, 115(1), 62-72 (2001)

93. Bakowies, D. and Thiel, W.: Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.*, 100(25), 10580-10594 (1996)

94. Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S. and Morokuma, K.: ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and $\text{Pt}(\text{P}(t\text{-Bu})_3)_2 + \text{H}_2$ Oxidative Addition. *J. Phys. Chem.*, 100(50), 19357-19363 (1996)

95. Maseras, F. and Morokuma, K.: IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.*, 16(9), 1170-1179 (1995)

96. Lin, H. and Truhlar, D. G.: QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.*, 117, 185-199 (2006)

97. Gao, J.: Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In: *Reviews in Computational Chemistry*. Ed K. B. Lipkowitz & D. B. Boyd. VHC Publishers, New York (1995)

98. Derat, E., Bouquant, J. and Humbel, S.: On the link atom distance in the ONIOM scheme. An harmonic approximation analysis. *J. Mol. Struct.*, 632(1-3), 61-69 (2003)

99. Ferre, N. and Olivucci, M.: The amide bond: pitfalls and drawbacks of the link atom scheme. *J. Mol. Struct.*, 632(1-3), 71-82 (2003)

100. Gao, J., Amara, P., Alhambra, C. and Field, M. J.: A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A*, 102(24), 4714-4721 (1998)

101. Pu, J., Gao, J. and Truhlar, D. G.: Generalized Hybrid Orbital (GHO) Method for Combining Ab Initio Hartree-Fock Wave Functions with Molecular Mechanics. *J. Phys. Chem. A*, 108(4), 632-650 (2004)

102. Vreven, T., Byun, K. S., Komaromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K. and Frisch, M. J.: Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *J. Chem. Theory Comput.*, 2(3), 815-826 (2006)

103. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C. and Pople, J. A.: Gaussian 03; Gaussian, Inc., Wallingford CT. In, (2004)

104. Laio, A., VandeVondele, J. and Rothlisberger, U.: A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.*, 116(16), 6941-6947 (2002)

105. Laino, T., Mohamed, F., Laio, A. and Parrinello, M.: An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.*, 1(6), 1176-1184 (2005)

106. Walker, R. C., Crowley, M. F. and Case, D. A.: The implementation of a fast and efficient hybrid QM/MM potential method within The Amber 9.0 sander module. *J. Comp. Chem.*, 29, 1019-1031 (2008)

107. Case, D. A., Darden, T. A., Cheatham, T. E. I., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A. E., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V. and Kollman, P. A.: AMBER 10. In, University of California: San Francisco (2008)

108. Woodcock, H. L., Hodoscek, M., Gilbert, A. T. B., Gill, P. M. W., Schaefer, H. F. and Brooks, B. R.: Interfacing Q-Chem and CHARMM to perform QM/MM reaction path calculations. *J. Comput. Chem.*, 28(9), 1485-1502 (2007)

109. Stenta, M., Calvaresi, M., Altoè, P., Spinelli, D., Garavelli, M., Galeazzi, R. and Bottoni, A.: Catalytic Mechanism of Diaminopimelate Epimerase: A QM/MM

- Investigation. *J. Chem. Theory Comput.*, 5(7), 1915-1930 (2009)
110. Stenta, M., Calvaresi, M., Altoè, P., Spinelli, D., Garavelli, M. and Bottoni, A.: The Catalytic Activity of Proline Racemase: A Quantum Mechanical/Molecular Mechanical Study. *J. Phys. Chem. B*, 112(4), 1057-1059 (2008)
111. Altoè, P., Stenta, M., Bottoni, A. and Garavelli, M.: A tunable QM/MM approach to chemical reactivity, structure and physico-chemical properties prediction. *Theor. Chem. Acc.*, 118(1), 219-240 (2007)
112. Sherwood, P., de Vries, A. H., Guest, M. F., Schreckenbach, G., Catlow, C. R. A., French, S. A., Sokol, A. A., Bromley, S. T., Thiel, W. and Turner, A. J.: QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.*, 632(1-3), 1-28 (2003)
113. Wang, Y., Shen, Y. and Ryde, U.: QM/MM study of the insertion of metal ion into protoporphyrin IX by ferrochelatase. *J. Inorg. Biochem.*, 103(12), 1680-1686 (2009)
114. Torras, J., de M. Seabra, G., Deumens, E., Trickey, S. B. and Roitberg, A. E.: A versatile AMBER-Gaussian QM/MM interface through PUPIL. *J. Comput. Chem.*, 29(10), 1564-1573 (2008)
115. Friesner, R. A.: Combined quantum and molecular mechanics (QM/MM). *Drug Discovery Today: Technologies*, 1(3), 253-260 (2004)
116. Gleeson, M. P. and Gleeson, D.: QM/MM Calculations in Drug Discovery: A Useful Method for Studying Binding Phenomena? *Journal of Chemical Information and Modeling*, 49(3), 670-677 (2009)
117. Gleeson, M. P. and Gleeson, D.: QM/MM As a Tool in Fragment Based Drug Discovery. A Cross-Docking, Rescoring Study of Kinase Inhibitors. *Journal of Chemical Information and Modeling*, 49(6), 1437-1448 (2009)
118. Mulholland, A. J.: Computational enzymology: modelling the mechanisms of biological catalysts. *Biochem. Soc. Trans.*, 036(1), 22-26 (2008)
119. Lodola, A., Woods, C. J., Mulholland, A. J., Ralph, A. W. and David, C. S.: Chapter 9 Applications and Advances of QM/MM Methods in Computational Enzymology. In: *Annual Reports in Computational Chemistry*. Elsevier, (2008)
120. Claeysens, F., Harvey, J. N., Manby, F. R., Mata, R. A., Mulholland, A. J., Ranaghan, K. E., Schütz, M., Thiel, S., Thiel, W. and Werner, H.-J.: High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chem. Int. Ed.*, 45(41), 6856-6859 (2006)
121. Rosta, E., Klahn, M. and Warshel, A.: Towards Accurate Ab Initio QM/MM Calculations of Free-Energy Profiles of Enzymatic Reactions. *J. Phys. Chem. B*, 110(6), 2934-2941 (2006)
122. Warshel, A.: Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu. Rev. Biophys. Biomol. Struct.*, 32(1), 425-443 (2003)
123. Sander, T., Liljefors, T. and Balle, T.: Prediction of the receptor conformation for iGluR2 agonist binding: QM/MM docking to an extensive conformational ensemble generated using normal mode analysis. *J. Mol. Graph. Model.*, 26(8), 1259-1268 (2008)
124. Cho, A. E., Guallar, V., Berne, B. J. and Friesner, R.: Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.*, 26(9), 915-931 (2005)
125. Kaukonen, M., Soderhjelm, P., Heimdal, J. and Ryde, U.: QM/MM-PBSA Method To Estimate Free Energies for Reactions in Proteins. *J. Phys. Chem. B*, 112(39), 12537-12548 (2008)
126. Khandelwal, A. and Balaz, S.: QM/MM linear response method distinguishes ligand affinities for closely related metalloproteins. *Proteins: Struct., Funct., Bioinf.*, 69(2), 326-339 (2007)
127. Jorgensen, W. L.: The Many Roles of Computation in Drug Discovery. *Science*, 303(5665), 1813-1818 (2004)
128. Acevedo, O. and Jorgensen, W. L.: Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic Reactions. *Acc. Chem. Res.* (2009)
129. Jorgensen, W. L.: Progress and Issues for Computationally-Guided Lead Discovery and Optimization. In: *Drug Design: Structure and Ligand-Based Approaches*. Ed K. Merz, C. Reynolds & D. Ringe. Cambridge University Press, Cambridge, UK (2009)
130. Jorgensen, W. L.: Efficient Drug Lead Discovery and Optimization. *Acc. Chem. Res.*, 42(6), 724-733 (2009)
131. Winum, J.-Y., Kohler, S., Scozzafava, A., Montero, J.-L. and Supuran, C. T.: Targeting Bacterial Metalloenzymes: A New Strategy for the Development of Anti-Infective Agents. *Anti-Infective Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry - Anti-Infective Agents)*, 7, 169-179 (2008)
132. White, R. J., Margolis, P. S., Trias, J. and Yuan, Z.: Targeting metalloenzymes: a strategy that works. *Curr. Opin. Pharmacol.*, 3(5), 502-507 (2003)
133. Kim, S.-K., Sims, C. L., Wozniak, S. E., Drude, S. H., Whitson, D. and W. Shaw, R.: Antibiotic Resistance in Bacteria: Novel Metalloenzyme Inhibitors. *Chemical Biology & Drug Design*, 74(4), 343-348 (2009)
134. Lüdemann, S. K., Lounnas, V. and Wade, R. C.: How do substrates enter and products exit the buried active site

of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.*, 303(5), 797-811 (2000)

135. Lüdemann, S. K., Lounnas, V. and Wade, R. C.: How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.*, 303(5), 813-830 (2000)

136. Winn, P. J., Lüdemann, S. K., Gauges, R., Lounnas, V. and Wade, R. C.: Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. U. S. A.*, 99(8), 5361-5366 (2002)

137. Dal Peraro, M., Llarrull, L. I., Rothlisberger, U., Vila, A. J. and Carloni, P.: Water-Assisted Reaction Mechanism of Monozinc β -Lactamases. *J. Am. Chem. Soc.*, 126(39), 12661-12668 (2004)

138. Dal Peraro, M., Vila, A. J., Carloni, P. and Klein, M. L.: Role of Zinc Content on the Catalytic Efficiency of B1 Metallo β -Lactamases. *J. Am. Chem. Soc.*, 129(10), 2808-2816 (2007)

139. Moran-Barrio, J., Gonzalez, J. M., Lisa, M. N., Costello, A. L., Dal Peraro, M., Carloni, P., Bennett, B., Tierney, D. L., Limansky, A. S., Viale, A. M. and Vila, A. J.: The Metallo-beta-lactamase GOB Is a Mono-Zn(II) Enzyme with a Novel Active Site. *J. Biol. Chem.*, 282(25), 18286-18293 (2007)

140. Ho, M.-H., Vivo, M. D., Peraro, M. D. and Klein, M. L.: Unraveling the Catalytic Pathway of Metalloenzyme Farnesyltransferase through QM/MM Computation. *J. Chem. Theory Comput.*, 5(6), 1657-1666 (2009)

141. Miscione, G., Stenta, M., Spinelli, D., Anders, E. and Bottoni, A.: New computational evidence for the catalytic mechanism of carbonic anhydrase. *Theor. Chem. Acc.*, 118(1), 193-201 (2007)

142. Irwin, J. J., Raushel, F. M. and Shoichet, B. K.: Virtual Screening against Metalloenzymes for Inhibitors and Substrates *Biochemistry*, 44(37), 12316-12328 (2005)

143. Hayik, S. A., Liao, N. and Merz, K. M.: A Combined QM/MM Poisson-Boltzmann Approach. *J. Chem. Theory Comput.* (2008)

144. Raha, K., Peters, M. B., Wang, B., Yu, N., Wollacott, A. M., Westerhoff, L. M. and Merz Jr, K. M.: The role of quantum mechanics in structure-based drug design. *Drug Discov. Today*, 12(17-18), 725-731 (2007)

145. Broyden, C. G.: The Convergence of a Class of Double-rank Minimization Algorithms. *J. Inst. Math. Appl.*, 6, 76-90 (1970)

146. Fletcher, R.: A New Approach to Variable Metric Algorithms. *Comput. J.*, 13, 317-322 (1970)

147. Goldfarb, D.: A Family of Variable Metric Updates Derived by Variational Means. *Math. Comput.*, 24, 23-26 (1970)

148. Shanno, D. F.: Conditioning of Quasi-Newton Methods for Function Minimization. *Math. Comput.*, 24, 647-656 (1970)

149. Nocedal, J.: Updating Quasi-Newton Matrices with Limited Storage. *Math. Comput.*, 35, 773-782 (1980)

150. Liu, D. C. and Nocedal, J.: On the Limited Memory Method for Large Scale Optimization. *Math. Prog. B*, 45(3), 503-528 (1989)

151. Powell, M. J. D.: A new algorithm for unconstrained optimization. Academic Press, New York (1970)

152. Gullingsrud, J. R., Braun, R. and Schulten, K.: Reconstructing Potentials of Mean Force through Time Series Analysis of Steered Molecular Dynamics Simulations. *J. Comput. Phys.*, 151(1), 190-211 (1999)

153. Jarzynski, C.: Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E*, 56(5), 5018 (1997)

154. Major, D. T. and Gao, J.: An Integrated Path Integral and Free-Energy Perturbation-Umbrella Sampling Method for Computing Kinetic Isotope Effects of Chemical Reactions in Solution and in Enzymes. *J. Chem. Theory Comput.*, 3(3), 949-960 (2007)

155. Marsili, S., Barducci, A., Chelli, R., Procacci, P. and Schettino, V.: Self-healing Umbrella Sampling: A Non-equilibrium Approach for Quantitative Free Energy Calculations. *J. Phys. Chem. B*, 110(29), 14011-14013 (2006)

156. Torrie, G. M. and Valleau, J. P.: Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.*, 23(2), 187-199 (1977)

157. Grubmüller, H.: Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E*, 52(3), 2893 (1995)

158. Barducci, A., Giovanni, B. and Parrinello, M.: Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.*, 100(2), 020603 (2008)

159. Ensing, B., Laio, A., Parrinello, M. and Klein, M. L.: A Recipe for the Computation of the Free Energy Barrier and the Lowest Free Energy Path of Concerted Reactions. *J. Phys. Chem. B*, 109(14), 6676-6687 (2005)

160. Laio, A. and Parrinello, M.: Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.*, 99(20), 12562-12566 (2002)
 161. Stanton, C. L., Kuo, I. F. W., Mundy, C. J., Laino, T. and Houk, K. N.: QM/MM Metadynamics Study of the Direct Decarboxylation Mechanism for Orotidine-5'-monophosphate Decarboxylase Using Two Different QM Regions: Acceleration Too Small To Explain Rate of Enzyme Catalysis. *J. Phys. Chem. B*, 111(43), 12573-12581 (2007)
 162. Petersen, L., Ard  vol, A., Rovira, C. and Reilly, P. J.: Mechanism of Cellulose Hydrolysis by Inverting GH8 Endoglucanases: A QM/MM Metadynamics Study. *J. Phys. Chem. B*, 113(20), 7331-7339 (2009)
 163. Bellucci, L., Laino, T., Tafi, A. and Botta, M.: Metadynamics Simulations of Enantioselective Acylation Give Insights into the Catalytic Mechanism of Burkholderia cepacia Lipase. *J. Chem. Theory Comput.*, 6(4), 1145-1156
 164. Laino, T., Mohamed, F., Laio, A. and Parrinello, M.: An Efficient Linear-Scaling Electrostatic Coupling for Treating Periodic Boundary Conditions in QM/MM Simulations. *J. Chem. Theory Comput.*, 2(5), 1370-1378 (2006)
 165. Car, R. and Parrinello, M.: Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.*, 55(22), 2471 (1985)
 166. K  hne, T. D., Krack, M., Mohamed, F. R. and Parrinello, M.: Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.*, 98(6), 066401 (2007)
 167. De Vivo, M., Dal Peraro, M. and Klein, M. L.: Phosphodiester Cleavage in Ribonuclease H Occurs via an Associative Two-Metal-Aided Catalytic Mechanism. *J. Am. Chem. Soc.*, 130(33), 10955-10962 (2008)
 168. Ho, M.-H., De Vivo, M., Dal Peraro, M. and Klein, M. L.: Understanding the Effect of Magnesium Ion Concentration on the Catalytic Activity of Ribonuclease H through Computation: Does a Third Metal Binding Site Modulate Endonuclease Activity? *J. Am. Chem. Soc.*, ASAP
 169. De Vivo, M., Ensing, B., DalPeraro, M., Gomez, G. A., Christianson, D. W. and Klein, M. L.: Proton Shuttles and Phosphatase Activity in Soluble Epoxide Hydrolase. *J. Am. Chem. Soc.*, 129(2), 387-394 (2007)
 170. Krack, M. and Parrinello, M.: All-electron ab-initio molecular dynamics. *Phys. Chem. Chem. Phys.*, 2(10), 2105-2112 (2000)
 171. Carloni, P., Rothlisberger, U. and Parrinello, M.: The Role and Perspective of Ab Initio Molecular Dynamics in the Study of Biological Systems. *Acc. Chem. Res.*, 35(6), 455-464 (2002)
 172. Marx, D. and Hutter, J.: Modern Methods and Algorithms of Quantum Chemistry. In: NIC Series. Ed J. Grotendorst. John von Neumann Institute for Computing, Julich (2000)
 173. VandeVondele, J. and Hutter, J.: An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.*, 118(10), 4365-4369 (2003)
 174. VandeVondele, J., Krack, M., Mohamed, F., Parrinello, M., Chassaing, T. and Hutter, J.: Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.*, 167(2), 103-128 (2005)
 175. Sorkin, A., Iron, M. A. and Truhlar, D. G.: Density Functional Theory in Transition-Metal Chemistry: Relative Energies of Low-Lying States of Iron Compounds and the Effect of Spatial Symmetry Breaking. *J. Chem. Theory Comput.*, 4(2), 307-315 (2008)
 176. Skylaris, C. K., Gagliardi, L., Handy, N. C., Ioannou, A. G., Spencer, S. and Willetts, A.: On the resolution of identity Coulomb energy approximation in density functional theory. *J. Mol. Struct. (THEOCHEM)*, 501-502, 229-239 (2000)
 177. Ufimtsev, I. S. and Martinez, T. J.: Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.*, 4(2), 222-231 (2008)
 178. Ufimtsev, I. S. and Martinez, T. J.: Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.*, 5(4), 1004-1015 (2009)
 179. Ufimtsev, I. S. and Martinez, T. J.: Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.*, 5(10), 2619-2628 (2009)
 180. Yasuda, K.: Accelerating Density Functional Calculations with Graphics Processing Unit. *J. Chem. Theory Comput.*, 4(8), 1230-1236 (2008)
 181. Vogt, L., Olivares-Amaya, R., Kermes, S., Shao, Y., Amador-Bedolla, C. and Aspuru-Guzik, A.: Accelerating Resolution-of-the-Identity Second-Order M  ller-Plesset Quantum Chemistry Calculations with Graphical Processing Units. *J. Phys. Chem. A*, 112(10), 2049-2057 (2008)
- Key Words:** Quantum Chemistry, Molecular Mechanics, QM/MM, Force Field, Drug Design, Enzyme Catalysis, DFT, CPMD, MP2, Review

Quantum chemical methods for drug design

Send correspondence to: Matteo Dal Peraro, EPFL SV
IB11 UPDALPE, AAB 017, Station 15, CH-1015,
Lausanne, Switzerland, Tel: 41216931861, Fax:
41216939665, E-mail: matteo.dalperaro@epfl.ch

<http://www.bioscience.org/current/volE3.htm>