10 SCIENTIFIC HIGHLIGHT OF THE MONTH: "Ab Initio Modeling of Biological Systems"

Ab Initio Modeling of Biological Systems

Simone Raugei and Paolo Carloni

International School for Advanced Studies (SISSA/ISAS)

and

INFM - DEMOCRITOS MOdeling Center for Research In aTOmistic Simulation

Via Beirut 2-4, 34014 - Trieste, Italy

e-mail: raugei@sissa.it

carloni@sissa.it

Abstract

Biological systems are particularly challenging for ab initio quantum mechanical methods, yet first principles calculations can now tackle problems of great current biological interest, that cannot be solved by a different approaches. Here we outline the state-of-the art of ab initio biological modeling by presenting a brief survey of new trends in the development of algorithms as well as few representative applications.

1 Introduction

One of the goals of modern biology is the understanding of biological phenomena at the molecular level, which involves the study of the structure of biomolecules and their functions. A detailed atomistic investigation of a biological system requires often the knowledge of its electronic structure. For instance, enzymatic reactions are bond-forming and bond-breaking phenomena, that require a quantum mechanical description [1] Another example is represented by photoreceptors (such as rhodopsin [2, 3, 4]), which involve excited states and the interaction between the biomolecules and the electromagnetic field. QM-based approach may be also of invaluable help to correctly describe polarization effects (*e.g.* in ion channels [5]) as well as metallo-proteins, where very subtle chemical phenomena (such as the fact that the metal ion ligand bond has a partially covalent nature) play an important role. Finally, QM calculations allows for comparison with a variety of spectroscopic data, such as IR [6, 7, 8, 9, 10], Raman [11, 12], and NMR [11, 13], which can be obtained from the electronic structure calculations without additional assumptions.

The study of biologically active molecules from first principles, such as proteins and DNA, however poses its own unique set of problems. Biological systems present an extremely high degree of complexity. First

of all, proteins may contain several hundreds of amino acids, *i.e.* thousands of atoms. Secondly, biological processes occur in aqueous solution and span millisecond or even seconds time scales. In addition, very often solvent molecules have an active role and have to be explicitly considered. Furthermore, dynamical effects both on short and long time scales are extremely important and must be taken into account in biological modeling. Therefore, the overall accuracy and predictive capability of computational models for biological systems are limited by (i) the accuracy to which relevant phase space regions are sampled, (ii) the degree to which the microscopic system on the computer reflects the typically macroscopic system in nature, and, in addition, (iii) the accuracy to which interaction forces are described.

Despite the explosive growth of computer power over the past two decades has led to the development of large-scale simulation techniques, direct application of first principle approaches to the study of biomolecules, although important, is still limited for the size of the model system that can be studied [14, 15, 16]. Few noticeable exceptions are represented by very recent studies on DNA [17, 18]. A successful strategy is represented by the coupling of multi-nanosecond classical molecular dynamics (MD) simulations [19, 20] based on empirical force fields to hybrid quantum mechanical/molecular mechanics (QM/MM) [21] optimizations and/or MD simulations. In a typical protocol classical MD is used to sample the relevant conformational space of the system and then QM/MM calculations are performed on representative configurations, where the region of interest, *e.g.*, the active site of an enzyme, is treated at QM level. This approach has been widely used in many studies [1, 3, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31].

Here we shall summarize some of the most common techniques used in the ab initio modeling of biological systems, namely density functional theory and the so-called hybrid QM/MM methods (section 2). Subsequently, we illustrate the power of the methodology for the modeling of biological systems by a survey of selected very recent applications from our and other groups. This survey is clearly very limited in scope and cannot cover important contributions from several active groups. For a more exhaustive reviews, see, e.g. Refs. [1, 32, 33]. First, we shall discuss fully ab initio calculations on a solvated DNA filament, which represent some of the biggest calculations ever performed on a biological system [17, 18] (section 3.1). Then, we will show the power of the combined use of classical and hybrid QM/MM simulations to the study of enzymatic catalysis, taking as illustrative example the decarboxylation of orotidine 5′-monophosphate by orotidine 5′-monophosphate decarboxylase [28] (section 3.2.1). Finally, we illustrate an extension to this approach to the excited-state dynamics of the retinal protonated Schiff base in rhodopsin [3] (section 3.2.2). Then we conclude with some final remarks (section 4).

2 Computational tools for atomistic biomodeling

2.1 Quantum mechanical methods

Because of the large number of atoms that must be explicitly taken into account ($> 50\text{-}80$) when dealing, for instance, with an active site of an enzyme, it is prohibitive the use of post-Hartree-Fock methods [14, 15, 16]. For this reason the most common method applied to biological systems is density functional theory (DFT), which achieves good accuracy and efficiency in large scale systems. A detailed description of DFT can be found in Ref. [34].

Large-scale DFT calculations on biological system are often performed with a plane wave (PW) basis set. Although localized basis sets require fewer basis functions per atom than PW, in a Kohn-Sham scheme the computational cost scales linearly with the number of PW, whereas with a localized basis set the cost scales as the cube of the number of basis functions. PW offer further advantages: (i) convergence can be easily monitored increasing the number of PW (which depends on only one parameter, *i.e.* the cut-off energy); (ii) forces in a PW calculation are not affected by Pulay terms; (iii) a PW basis set is completely unbiased and, therefore, no basis superposition error is present. Calculations with PW can

practically deal only with valence electrons and pseudo potentials for the core electron must be used. The most common pseudopotentials are the norm-conserving pseudopotentials of Troullier and Martins [35], and Goedecker, Teter and Hutter [36], and the ultra soft non-norm-conserving pseudopentials due to Vanderbilt [37].

A promising valid alternative to pure PW calculations is represented by hybrid Gaussian and plane waves (GPW) method [38, 39]. The GPW method provides an efficient way to treat the Hartree (Coulomb) energy and the orthogonalization of the wave functions. These terms are notoriously not scaling linearly with system size, and they dominate the computational cost for larger systems in standard electronic structure calculations. The GPW method uses an atom-center Gaussian-type basis set to describe the wave functions, but uses an auxiliary plane waves basis (PW) set to describe the density. With a density represented as PW on a regular grid, the efficiency of Fast Fourier Transform can be exploited to solve the Poisson equation and to obtain the Hartree energy in a time scale that scales linearly with the system size. Some applications based on this method have appeared in the literature [40, 41, 42, 43, 44].

In a DFT calculation a crucial role is played by the choice of the exchange and correlation functional. Although the simple local density approximation (LDA) is extremely good for studying solids, it does not give reasonably good bonding energy and lengths and other molecular properties. Therefore, when dealing with biological systems, it is necessary to go beyond LDA and use a gradient corrected functional. One the the most common functionals used in biomodeling is the one that uses the Becke exchange [45] and the Lee, Parr and Yang correlation [46], or BLYP. BLYP functional gives a quite accurate description of hydrogen-bonded systems [47, 48, 49, 50, 51]. Another very used functional is the hybrid B3LYP functional [52], which combines Slater, Hartree-Fock and Becke [45] exchange with the correlation terms due to Lee, Parr and Yang [46]. This functional has shown to give results in some cases comparable to MP2 calculations, although it is prohibitively expensive in a PW scheme. When in the system are present transition metals, it has been shown that the BLYP functional fails in most cases and the Becke and Perdew (BP) functional is a better alternative [45, 53].

The majority of the studies of chemical reactivity carried out so far using DFT seem to indicate that the reaction barriers heights are generally underestimated (see for example Refs. [30, 54, 55, 56, 57]) by many of the current implementation of DFT. Proton transfer and S_N2 reactions are particularly challenging for DFT. This has been reported [58, 59, 60] to be the case for the very popular BLYP exchange and correlation functional, or its improved modification B3LYP, on the basis of a comparison with both MP2 calculations and experimental results [61, 62]. The major failure have been found in proton transfer [54, 55] and S_N2 reactions [58, 56] (exceptions exist, see, *e.g.* Ref. [63]). In many circumstances this deficiency is to be attributed to the failing of many generalized gradient functionals in describing three-centers, two-electron chemical bonds, as has been thoroughly discussed by others (see for example Ref. [64]). Therefore, in order to apply the DFT method to the studies of chemical reaction mechanisms the question of the accuracy of the calculated barrier heights is of primary importance. The validity of the functional adopted must always be controlled against more accurate approaches or the experimental data.

An additional comment is in order about the fact that current gradient-corrected density functionals are unable to describe dispersive interactions [65, 66, 67, 68], which is a key interaction for biological systems. Therefore, if van der Waals interactions are crucial in the system being studied they must be recovered in the calculation. As DFT methods include (at least partially) correlation effects in the distance range where density overlap between two interacting fragments is appreciable, the simplest way to recover dispersion interactions at long range is the introduction of empirical damped dispersion terms of the form $f(R)\, C_6 R^{-6}$, where f is a damping function [69, 70, 71]. This approach has already been applied in the framework of ab initio Car-Parrinello simulations [72, 70]. More recently, in a pseudopotential calculation, it has been proposed to construct an effective potential consisting of optimized nonlocal higher angular momentum dependent terms for all of the atoms in the system in order to compensate the absence of dispersion forces in a generalized gradient approximation functional [73, 74]. This approach

has given remarkably good results for many "pathological" systems [73, 75] and represents a promising method to be applied to the study of biological systems.

2.2 Hybrid quantum mechanical/molecular mechanics methods

Chemical and biochemical processes usually take place in complex heterogeneous condensed phase environments consisting of thousands of atoms. The simplest approach to overcome the limitation on the size of the system is to isolate the part of the system that is the most relevant for the function of the system itself. Usually this requires the use of geometrical constraints to impose that the original (local) structure is preserved during structural relaxation and/or dynamics. One has always to check the reliability of the model chosen enlarging systematically the system and ensuring that the results of the calculation are converged with respect to the size. This approach may require the use of very large quantum models, which include elements of the biological system not important for the function itself but that are crucial for its mechanical stability. In this scheme, long-range effects may be incorporated by including the environment electrostatic field [76] and the solvation may be taken into account by a dielectric continuum [77].

A better solution that is often used is to employ an hybrid QM/MM approach [21] in which the whole system is partitioned into a localized chemically active region, treated at quantum mechanical level, and the remaining part of the system, treated with empirical force fields. Several partitioning schemes exist [21, 78, 79, 80]. The major difficulty lies in achieving a good description of the interface region, where covalently bonds may be cut. The most common method used is the link atom approach, where the valence of QM atoms at the interface is saturated by a parameterized monovalent pseudoatom(s) or by simple hydrogen atom(s) (the so-called capping technique).

Crucial is also the treatment of the long-range electrostatic interactions. In this regard, among the possible protocols, a recent and efficient scheme is the fully Hamiltonian approach due to Rothlisberger and coworkers [79, 80, 81]. The MM subsystem is split in three regions centered on the QM part. For computational efficiency reasons, only the electrostatic interaction between the QM system and a first shell of atoms around it is explicitly taken into account. A modified Coulomb interaction is introduced in order to have the correct interaction properties and avoid an unphysical escape of electronic density to the MM atoms (the so-called spill-out effect) [79]. The electrostatic interaction between the QM region and an intermediate MM region is calculated by point charge-point charge electrostatic potential where the point charges for the atoms in the quantum region are obtained by a RESP procedure [80]. The interaction between the QM part and the more distant MM atoms are expressed in terms of multipole moments of the QM charge density and classical point charges.

The scheme due to Rothlisberger and coworkers has already been applied to many biochemical processes in the framework of the ab-initio Car-Parrinello molecular mechanics [3, 25, 26, 27, 28, 29, 30, 31]. In Section 3.2 two applications of this approach are presented [28, 3]. Other valuable QM/MM approaches are not discussed here [21].

We conclude remarking that also in a QM/MM scheme the reliability of the QM model chosen must be always checked.

2.3 Molecular dynamics

Molecular dynamics (MD) simulation based on the description of interatomic interactions via empirical force fields is a well consolidated used tool for the study of biomolecular systems. Currently, it is possible to study thousands of atoms on time scales that can reach reach the 100-ns range. Its success resides in the development of realistic force fields.

However, there are many areas in which the use of effective potentials may be not appropriate, and more sophisticated and accurate approaches are required. An alternative is offered by ab initio molecular dynamics (AIMD). AIMD in the Born-Oppheneimer approximation can be efficiently performed using the scheme proposed by Car and Parrinello (CP) [82, 83], where the electronic degrees of freedom (represented by one electron wave functions) are treated as fictitious classical variables, using a extended Lagrangian approach similar to those used in classical (with empirical force fields) MD to study thermodynamic ensembles other the the microcanonical ensemble [19, 20]. It is always possible to choose a generalized mass associated to the fictitious electronic degrees of freedom in a such a way that the latter adjust instantaneously to changes in the nuclear coordinates and the resulting dynamics is adiabatic. The method is extensively discussed in Ref. [83] (some important remarks on the effects of the fictitious electronic mass are given in Refs. [42, 84, 85]). CP MD has been intensively applied in materials science and in chemical problems. The advantage of the method lies essentially in the quality of the inter-atomic potential at all of the phase space points. Since the method allows an exact consideration of the anharmonic effects at finite temperature, a detailed study of the energy redistribution among the degrees of freedom as a function of time and of the polarization effects along a reaction pathway is feasible.

Limitations arising from the problems outlined in the introduction are clearly present. Given the computational cost and the space and time restrictions of using an ab initio approach it is not possible to explore the phase space of a biological system. It is necessary therefore to make hypotheses on the action mechanisms and/or structures to be considered. This is somehow a limitation in the predicting power of first principle methodologies applied to biochemical processes. As already mentioned, one successful strategy to alleviate this drawback is to couple multi-nanosecond classical molecular dynamics to AIMD. In practice the conformational space of the system is sampled via standard molecular dynamics simulations and then, on representative configurations, a AIMD/MM calculation is performed.

Finally, many chemical process involve bond breaking and forming, which often require the crossing of a relatively large activation barrier. Hence, a method to sample the relevant reaction coordinate must be employed. An huge effort has been mounted to develop sampling techniques of reaction paths in the condensed phase, as is shown by the impressive number of papers devoted to this subject [86, 87, 88, 89, 90, 91]. To date, there is not any really efficient way to explore the phase space of a reactive system in the condensed phase, which can be used routinely in ab initio simulations, that does not rely at least on an initial guess of a reaction coordinate. From a computational point of view, one of the best techniques to be applied in CPMD simulations is the *Blue Moon Ensemble* constrained molecular dynamics [91, 92], in which the system is forced to move along a suitable coordinate. Recently, multiple steering molecular dynamics [86] has also been applied in a CP framework [28, 93]. This approach allows one to work out the free-energy profile along selected reactive routes via non equilibrium MD [86]. To date, the most powerful technique to find reactive pathways is probably represented by the so-called metadynamics of hills method, which was introduced few years ago by Laio and Parrinello [87, 94]. The method is based on a coarse-grained history-dependent dynamics (metadynamics) that is able to explore the free energy in the space defined by a manifold of collective coordinates S_α that characterize the reaction process. At each metadynamics step the system evolution is guided by the combined action of the thermodynamic force (which would trap the system in the free-energy wells) and a history-dependent force, which disfavors configurations in S_α space that have already been visited. The history-dependent potential, is constructed as a sum of Gaussians centered on each value of the S_α already explored during the dynamics. This approach has been successfully applied to the study of the first steps of the oxidative damage of DNA via radical cation formation [27].

3 Applications

Because of the large number of ab initio applications to biological systems already presented in the literature, it is clearly impossible to review all of the work appeared so far. Here we report only few representative examples, which give the flavor of what it can be done nowadays exploiting the best computational resources and the state-of-the-art computational methodologies.

3.1 Electronic structure DNA

3.1.1 Electronic structure of Wet DNA

The computational study of nucleic acids and other biopolymers in laboratory-realizable conditions has until now focused mainly on their structural properties, while the nature of the electronic structure has received far less attention. Yet the electronic states play an important role in determining the interatomic forces, as they lead to electronic polarization effects and many-body forces and provide an accurate description of the dielectric properties. Furthermore, they are crucial in phenomena like radiation-induced damage [95]. More recently, there has been great interest in the nature of the electronic structure of DNA because of its potential applications to nanotechnology [96, 97, 98, 99, 100, 101]. Only recently has an effort been made to study periodically infinite double strands with high quality ab initio calculations by Gervasio *et al.* [27].
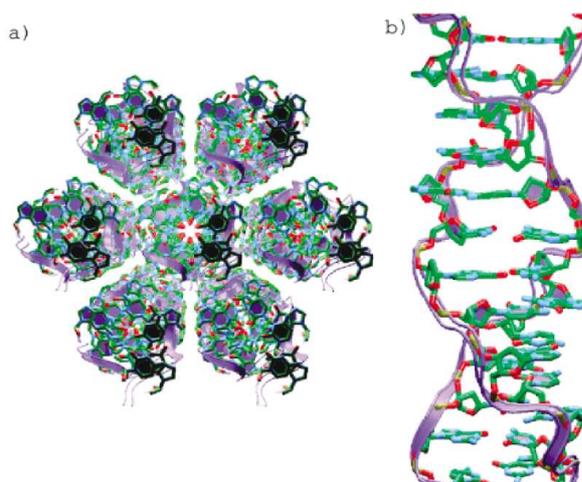

Figure 1: View of the three-dimensional structure of the G:C dodecamer [102] (a) along, (b) orthogonal to the **c** (z) axis. Water molecules, counterions and hydrogens have been removed for clarity. The sugar-phosphate backbone is represented as ribbons.

The system studied is a fully hydrated double strand self-complementary DNA (polyd(GpCp) [102], see Fig. 1). This is an infinitely repeated biopolymer, which in the unit cell contains twelve guanine:cytosine (G:C) pairs in the Z conformation. Polyd(GpCp) adopts the Z conformation only under conditions of high ionic strength and thus is rarely observed in nature. This choice, however, was made since in this particular crystal structure the DNA is continuous across crystal boundaries and because the Z conformation exhibits less thermal fluctuations than either A or B forms. [103] The system included other than the solvation water, all of the counterions, which are crucial for the stability of the DNA [104]. The model contained 654 heavy atoms and 540 hydrogen atoms. Geometry relaxation and subsequent electronic structure calculations explicitly taken into account 3,960 valence electrons (for sodium semicore

states were also considered).

The calculations revealed that twelve quasi-degenerate states are positioned at the top of the valence band. These states have a π character and are mostly localized on the G nucleobases. Each individual state of this manifold is spread over several G bases (Fig. 2). The states immediately below the top of the valence band were assign to states localized also on G. The first C localized state is at 0.78 eV below the top.
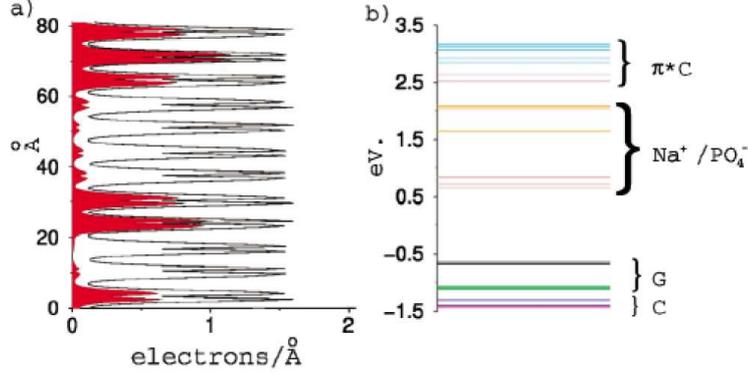

Figure 2: (a) Electronic charge density $\rho_e(z)$ along the z axis integrated over the x and y directions. The integrals are performed only in regions surrounding the G basis. In red the $\rho_e(z)$ of the top state is shown. The black line give the total $\rho_e(z)$ of the manifold. (b) Schematic level diagram around the Fermi level. The Fermi level positioned in the middle of the gap has been chosen as the zero of energy.

The DFT gap between empty and occupied states is particularly small, being only 1.28 eV. This reflects the nature of the state at the bottom of the conduction band, which is a charge transfer state where one electron has been moved outside the helix mostly on the Na^+ counterions and on the PO_4^- groups. The effect of water molecules on this state and on the value of the gap was assessed repeating the calculation by removing the water molecules but otherwise leaving the geometry of the DNA and counterions unchanged. The gap is much reduced. This reveals the electrostatic nature of the charge transfer states and the fundamental role of water in shielding the DNA from the electrostatic field of counterions.

3.1.2 Charge localization in a DNA filament

Charge transfer in DNA is currently the subject of intense theoretical and experimental investigation [105, 106]. This is due both to a possible use of DNA as a component in nanoelectronic and electrochemical devices [107] and to the fundamental role of conductivity in the oxidative damage of DNA [27, 108]. Recent experiments have provided contradictory results, ranging from a highly conducting wire [109] and a proximity induced superconductor [110] to a semiconductor [111] or an insulator[112, 113]. Wet DNA has been shown to be a charge carrier when its length is shorter than $\simeq 20$ Å [114]. While DNA helices longer than $\simeq 40$ Å or in dry conditions were generally found to be insulators or high-bandgap semiconductors [115].

In a very recent study Gervasio *et al.* [18] by using fully ab initio calculations examined the effect of charge defect and the stabilization mechanism of the hole localization [116, 117, 118, 119, 120].

On the basis of this investigation, the authors excluded that in polyd(GpCp) charge localization is directly due to helix distortions. On the other hand evidence for a proton-coupled charge transfer mechanism was found. In fact, when the proton to the G:C base pair where the spin density has a maximum was
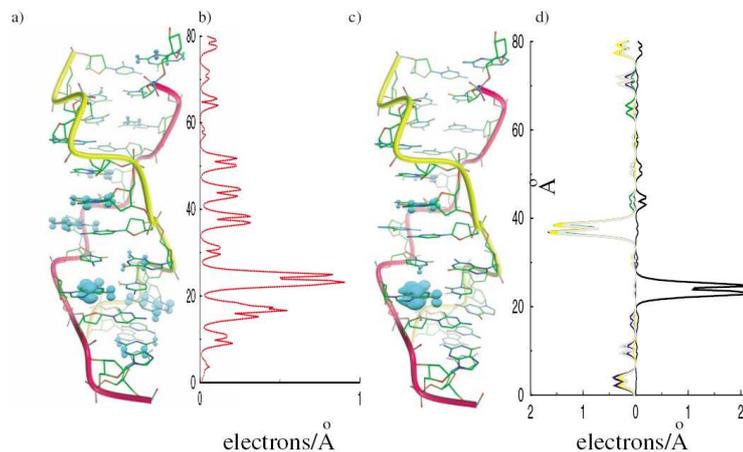
Figure 3: View of the structure of the polyd(GpCp) and of the spin density isosurface (in cyan) associated with the radical cation state a) before and c) after localization of the spin defect due to proton transfer from G to C. The isosurfaces represented have a value of 10^{-3} electrons $\mathring{A}^{-3}$. b) and d) Electronic spin density $\rho_s(z)$ projected along the z axis. b) the $\rho_s(z)$ corresponding to a). d) right, the $\rho_e(z)$ corresponding to the status depicted in c). d) left, four projections of the $\rho_e(z)$ corresponding to the shift of the proton from a different G each taken every 48 fs.

displaced, a remarkable localization of the hole on G was noticed. This lead to a situation in which the unpaired spin (the hole) is on G, yielding a $(G_{-H})^{\cdot}$ state, while H^+ moves to C, which becomes protonated (Fig. 3c and d, black line). This important result is fully consistent with a wealth of experimental data on solvated DNA [117, 119, 121].

3.2 QM/MM studies on enzymatic function

3.2.1 Ground state destabilization vs transition state stabilization

Several experimental, but somehow controversial, evidences indicate that certain enzymes can work destabilizing the ground state (GSD) of the substrate/enzyme complex [122, 123] instead of stabilizing the transition state [124], for example exerting an electrostatic or a geometrical frustration on the substrate [123]. This possibility has given rise to a long standing debate. The structure and the biological role of the few enzymes for which a ground state destabilization has been proposed are very different from each other. Therefore, to proof/disproof a ground state destabilization is necessary to analyze singularly a representative number of enzymes.

One of the enzyme taken as reference for the transition state versus ground state destabilization controversy is orotidine 5′-monophosphate decarboxylase (ODCase), which catalyzes the decarboxylation of orotidine 5′-monophosphate (Omp) to a major precursor in the formation of pyrimidine nucleotides, uridine 5′-monophosphate [122]. ODCase represents a crucial enzyme in the de-novo synthesis of DNA bases and represents one of the most proficient enzymes, enhancing the rate of spontaneous substrate decarboxylation by more than 17 orders of magnitude. Its remarkable catalytic power is entirely dependent on non-covalent binding forces and does not involve metals or other cofactors [125]. This is a very unusual characteristic among decarboxylating enzymes [122].

The remarkable proficiency of the enzyme has been suggested to be caused by GSD on the basis of X-Ray crystallography experiments [126, 127], kinetic ^{15}N and ^{13}C isotope effects [128], and free energy and

binding energy calculations [126]. It is proposed that the electrostatic repulsion between the substrate and the nearby Asp70 carboxylate would drive the decarboxylation. This repulsion would be counterbalanced by the favorable binding of the phosphate tail [122, 126, 129].

On the contrary, other studies, based on EVB free energy perturbation simulations and binding energies affinities and classical MD simulations [130], invoke an ordinary transition state stabilization mechanism.
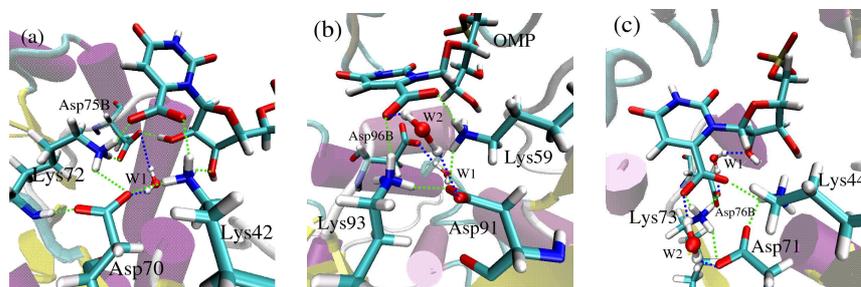

Figure 4: Snapshot of ODCase/Omp complex ground state. (a) ODCase form *Methane bacterium thermoautotrophicum*; (b) ODCase form yeast *Saccharomyces cerevisiae*; (c) ODCase form *Escherichia Coli*.

Very recently long time scale classical MD and hybrid CP/MM simulations were employed to help clarify the issue [28] The lack of structural (experimental) information on the ground state of ODCase/Omp complex was overcome by a careful construction of the model and the analysis of three different strains of the enzyme. It was found that the ODCase/substrate complex is characterized by a very stable charged network Omp-Lys-Asp-Lys-Asp, which with a direct decarboxylation driven by a ground state destabilization (Fig. 4).

On the contrary, a direct decarboxylation induced by an electrostatic stabilization of the transition state is consistent with this study. The calculated activation free energy (Fig. 5) for the direct decarboxylation with the formation of a C6 carboanionic intermediate, yields an overall rate enhancement by the enzyme ($k_{cat}/k_{wat} = 3.5 \cdot 10^{16}$) in agreement with experiments ($k_{cat}/k_{wat} = 1.7 \cdot 10^{17}$).

The decarboxylation is accompanied by the movement of a fully conserved lysine residue toward the developing negative charge at C6 position. The calculated trend of the electrostatic interaction energy suggests an extremely important role in the catalysis for Lys72, which is able to stabilize the developing negative charge at C6, and therefore the TS or, equivalently, the intermediate, and eventually provides the proton that saturates the valence of the C6 atom of Ump$^-$. Of course, water cannot stabilize the developing C6 charge (and therefore the intermediate) equally well.

This picture confirms also for ODCase the validity of the concept of preorganization energy [124], in which the electrostatic field generated by the enzyme active site is already organized in such a way to be complementary to the transition state charge distribution. No evidences of GSD was found as (i) the geometry of the substrate in enzyme cavity resembles very much that the most stable conformer in aqueous solution; (ii) the H-bond network accommodating the substrate is remarkably stable. In addition to Lys72 (MTBO numbering), [126, 124, 130] also Lys42 turned out to interact with carboxyl moiety of the substrate, which strongly contributes to shield the Omp carboxylate from the forming Asp70. This feature is common to all of the ODCase/Omp complexes investigated. Finally, this study clearly shows the importance of both the construction of a reliable enzyme/substrate complex and the proper sampling of the configurational space of the complex itself.
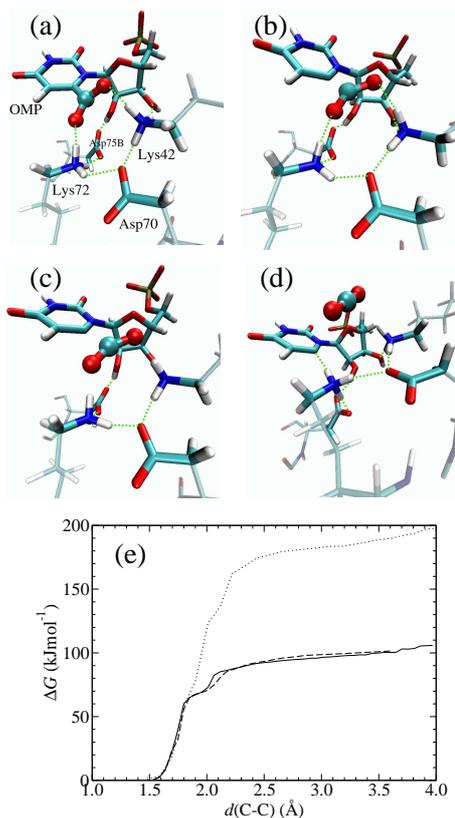
Figure 5: Decarboxylation of orotidine 5′-monophosphate. (a) Reactants; (b) breaking of the C–C bond; (c) breaking of the $CO_2 \cdots (Lys)_2$ H-bond; (d) Intermediate; (e) free energy profiles. In (e) solid and dotted lines refer to the direct decarboxylation in the enzyme cavity and in aqueous solution, respectively (the dashed line is the free energy profile calculated averaging over only one trajectory, see Ref. [28]). The residues included in the quantum sub system are evidenced: the charged network Lys-Asp-Lys-Asp and the entire substrate (Omp).

3.2.2 Photoreaction in rhodopsin

Rhodopsin proteins consist of a bundle of seven α-helices, which hold retinal chromophore in its protonated Schiff base (RPSB). The primary event in the process of vision is the absorption of a photon by the retinal. Light absorption by rhodopsin leads to conformational change in the chromophore, which is followed by a complex signal transduction pathway leading to the stimulation of the optic nerve.

The 11-cis to all-trans isomerization of the RPSB in the protein environment is ultrafast (200 fs) [131] and very efficient (quantum yield 0.65), [132] in contrast to the same photoreaction in solution. This efficiency is very intriguing because of the steric confinement of the RPSB to a small binding pocket that should hamper the large movements required to adopt an all-trans conformation. Experimental evidence reveals that bathorhodopsin, the first thermally equilibrated intermediate in the signaling cascade, exhibits a strained all-trans RPSB and stores about 134 kJ/mol of the photon energy [133].

Recently Rothlisberger and coworkers [3] have studied the conformational change of retinal on absorption of a proton with hybryd QM/MM MD methodologies, which consider the chromophore at ab-initio level, while take into account the heterogeneity and complexity of both the protein and the membrane environment by a classical force field. For the description of the first excited singlet state (S_1) the restricted open-shell Kohn-Sham (ROKS) algorithm was used [134]. This provides a good compromise between accuracy and computational efficiency.
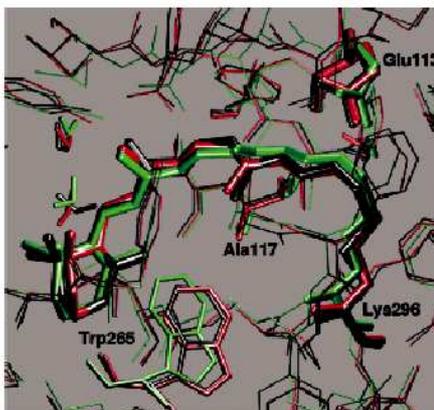
Figure 6: Superposition of the chromophore structure in the protein binding pocket in the dark state (black), at the S1 f S0 transition (red), and after 500 fs of relaxation in the isomerized state (green).

Starting from configurations sampled in a dark-state simulation, 23 excited-state QM/MM trajectories of about 100 fs each were calculated. The excited-state configuration of the RPSB is characterized by the well-known inversion of the bond length pattern. In S_1, especially the bonds C_9-C_{10} (1.44 Å), C_{11}-C_{12} (1.43 Å), and C_{13}-C_{14} (1.43 Å) are elongated, thus lowering the barrier toward isomerization. Whereas the electronic structure would be unselective toward the rotation of any of these double bonds, the protein environment favors C_{11}-C_{12} bond isomerization by steric strain. In fact, the dihedral angles from C_7 to C_{11} and from C_{12} to N deviate in S_1, similar to S_0, only by at most 15° from a perfect trans conformation. A very small barrier for the rotation around the pre-twisted dihedral angle C_{10}-C_{11}-C_{12}-C_{13} was found. This dihedral rotates to -100° and no recrossing to the dark state -35° was observed. When released to the ground state after 90 fs, the RPSB evolves toward a highly twisted all-trans structure. Remarkably, in the excited state, no atom moves more than 0.8 Å, that is, only 0.3 Åmore than the maximal thermal displacement in the dark state. After a 500 fs ground state relaxation, it was found that only the methyl groups C_{19} and C_{20} move further away from their position in the dark state. The strain is propagated through the carbon chain. The unexpected small difference between the primary photoproduct and the dark-state structure is also evident in Fig. 6.

Therefore, the 11-cis to all-trans isomerization is possible within the binding pocket with a minor atomic rearrangement, which produces a highly strained chromophore. The estimated energy stored in the system after isomerization is 117 kJ/mol, which is in good agreement with the experimentally determined energy storage in bathorhodopsin (134 kJ/mol). In the all-trans conformer, the energy stored in the internal degrees of freedom of the RPSB (+75 kJ/mol) and the van der Waals (steric) interaction energy between the RPSB and the protein (+10 kcal/mol) increase substantially, while the electrostatic interaction energy remains unaffected (< 0.4 kJ/mol difference).

In conclusion, according to this study, the initial step of vision can be viewed as the compression of a molecular spring that can then release its strain by altering the protein environment in a highly specific manner.

4 Conclusions

Biological systems are particularly challenging for ab initio quantum mechanical methods. Nevertheless, first principles calculations can be used to attack problems of great current biological interest, that cannot

be solved by a different approaches. In this paper we have described the state-of-the-art in ab initio (density functional theory) modeling of biological systems. This survey is clearly very limited in scope and cannot cover all of the (many) aspects of ab initio investigations of biologically relevant systems. We have tried to critically analyze the most commonly used methodologies. The most serious limitations encountered when studying a biological system are essentially (i) the accuracy to which relevant phase space regions are sampled, (ii) the degree to which the microscopic system on the computer reflects the typically macroscopic system in nature, and, in addition, (iii) the accuracy to which interaction forces are described. The first two limitations can be overcome combining ab initio calculations with more accurate sampling methodologies such as classical molecular dynamics and efficient sampling techniques of rare events. For what concern the last limitation, results must always be checked against more accurate schemes. Clearly this can be done only for model systems, which capture all the main chemical aspects of the system being studied.

Acknowledgment

References

[1] P. Carloni, U. Roethlisberger, and M. Parrinello. The role and perpective of ab initio molecular dynamics in the study of biological systems. *Acc. Chem. Res.*, 35:455–464, 2002.

[2] C. Molteni, I. Frank, and M. Parrinello. An excited state density functional theory study of the rhodopsin chromophore. *J. Am. Chem. Soc.*, 121:12177–12183, 1999.

[3] U. F. Rohrig, L. Guidoni, A. Laio, I. Frank, and U. Rothlisberger. A molecular spring for vision. *J. Am. Chem. Soc.*, 126:15328–15329, 204.

[4] R. Rousseau, V. Kleinschmidt, U. W. Schmitt, and D. Marx. Modeling protonated water networks in bacteriorhodopsin. *Phys. Chem. Chem. Phys.*, 6:1848–1859, 2004.

[5] L. Guidoni and P. Carloni. Potassium permeation through the kcsa channel: a density functional study. *Biochim. Biophys. Acta*, 2002.

[6] P. L. Silvestrelli, M. Bernasconi, and M. Parrinello. Ab initio infrared spectrum of liquid water. *Chem. Phys. Lett.*, 277:478–482, 1997.

[7] M. Bernasconi, P. L. Silvestrelli, and M. Parrinello. Ab-initio absorption study of the hydrogen-bond symmetrization in ice. *Phys. Rev. Lett.*, 81:1235–1239, 1998.

[8] Z. W. Zhu and M. E. Tuckerman. Ab initio molecular dynamics investigation of the concentration dependence of charged defect transport in basic solutions via calculation of the infrared spectrum. *J. Phys. Chem. B*, 106:8009–8018, 2002.

[9] M. P Gaigeot and M. Sprik. Ab initio molecular dynamics computation of the infrared spectrum of aqueous uracil. *J. Phys. Chem. B*, 107:10344–10358, 2003.

[10] M. Pagliai, G. Cardini, R. Righini, and V. Schettino. Hydrogen bond dynamics in liquid methanol. *J. Chem. Phys.*, 119:6655–6662, 2003.

[11] A. Putrino, D. Sebastiani, and M. Parrinello. Generalized variational density functional perturbation theory. *J. Chem. Phys.*, 113:7102–7109, 2002.

[12] A. Putrino and M. Parrinello. Anharmonic raman spectra in high-pressure ice from ab initio simulations. *Phys. Rev. Lett.*, 88:176401–1–176401–4, 2002.

[13] S. Piana, D. Sebastiani, P. Carloni, and M. Parrinello. Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site. *J. Am. Chem. Soc.*, 123:8730–8737, 2001.

[14] P. Hobza and J. Sponer. Structure, energetics, and dynamics of the nucleic acid base pairs: Nonempirical ab initio calculations. *Chem. Rev.*, 99:3247–3276, 1999.

[15] P. E. M. Siegbahn and M. R. A. Blomberg. Transition-metal systems in biochemistry studied by high-accuracy quantum chemical methods. *Chem. Rev.*, 100:421–437, 2000.

[16] F. Himo and P. E. M. Siegbahn. Quantum chemical studies of radical-containing enzymes. *Chem. Rev.*, 103:2421–2456, 2003.

[17] F. L. Gervasio, P. Carloni, and M. Parrinello. Electronic Structure of Wet DNA. *Phys. Rev. Lett.*, 89:108102–1–4, 2002.

[18] F. L. Gervasio, A. Laio, M. Parrinello, and M. Boero. Charge localization in dna fibers. *Phys. Rev. Lett.*, 94:158103, 2005.

[19] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids.* Claredon Press, Oxford, 1987.

[20] D. Frenkel and B. Smith. *Understanding molecular dynamics.* Accademic Press, 1996.

[21] P. Sherwood. Hybrid quantum mechanics/molecular mechanics approaches. In J. Grotendorst, editor, *Modern Methods and Algorithms of Quantum Chemistry Proceedings*, volume 3 of *NIC*, pages 285–305, Jülich, 2000. John von Neumann Institute fo Computing.

[22] M. Garcia-Viloca, J. Gao, M. Karplus, and D. G. Truhlar. How enzymes work: Analysis by modern rate theory and computer simulations. *Science*, 303:186–195, 2004.

[23] G. Nadig, L. C. van Zant, S. L. Dixon, and Jr. K. M. Merz. Charge-transfer interactions in macromolecular systems: a new view of the protein/water interface. *J. Am. Chem. Soc.*, 120:5593–5594, 1998.

[24] S. Piana, P. Carloni, and M. Parrinello. Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease . *J. Mol. Biol.*, 319:567–583, 2002.

[25] M. Sulpizi, A. Laio, J. VandeNondele, A. Cattaneo, and U. Rothlisberger. Reaction mechanism of caspases: Insights from qm/mm car-parrinello simulations. *Proteins*, 52:212–224, 2003.

[26] A. Magistrato, W. F. DeGrado, A. Laio, U. Rothlisberger, J. VandeVondele, and M. L. Klein. Characterization of the dizinc analogue of the synthetic diiron protein df1 using ab initio and hybrid quantum/classical molecular dynamics simulations. *J. Phys. Chem. B*, pages 4182–4188, 2003.

[27] F. L. Gervasio, A. Laio, M. Iannuzzi, and M. Parrinello. Influence of dna structure on the reactivity of the guanine radical cation. *Chem. Eur. J.*, 10:4846–4852, 2004.

[28] S. Raugei, M. Cascella, and P. CArloni. A proficient enzyme: Insights on the mechanism of orotidine monophosphate decarboxylase from computer simulations. *J. Am. Chem. Soc.*, 126:15730–15737, 2004.

[29] M. Dal Peraro, L. I. Llarrull, U. Rothlisberger, A. J. Vila, and P. Carloni. Water-assisted reaction mechanism of monozinc beta-lactamases. *J. Am. Chem. Soc.*, 126:12661–12668, 2004.

[30] S. Piana, D. Bucher, P. Carloni, and U. Rothlisberger. Reaction mechanism ofhiv-1 protease by hybrid car-parrinello/classical md simulations. *J. Phys. Chem. B*, 108:11139–11149, 2004.

[31] M. Cascella, C. Micheletti, U. Rothlisberger, and P. Carloni. Evolutionarily conserved functional mechanics across pepsin-like and retroviral aspartic proteases. *J. Am. Chem. Soc.*, 127:3734–3742, 2005.

[32] P. Carloni and U. Rothlisberger. Simulations of enzymatic systems: Perspectives from car-parrinello molecular dynamics simulations. In L. Eriksson, editor, *Theoretical Biochemistry - Processes and Properties of Biological Systems*, page 215. Elsevier Science, 2001.

[33] U. Rothlisberger. 15 years of car-parrinello simulations in physics, chemistry. In J. Leszczynski, editor, *Biology in Computational Chemistry: Reviews of Current Trends*, volume 6. World Scientific, 2001.

[34] R.G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, Oxford, 1989.

[35] N. Troullier and J. L. Martins. Efficient pseudopotential for plane-wave calculations. *Phys. Rev. B*, 43(3):1993–2006, 1991.

[36] S. Goedecker, M. Teter, and J. Hutter. Separable dual-space gassian pseudopotential. *Phys. Rev. B*, 54(3):1703–1710, 1996.

[37] D. Vanderbilt. Soft self-consistent pseudopotential in a generalized eigenvalue formalism. *Phys. Rev. B*, 41(11):7892–7895, 1990.

[38] G. Lippert, J. Hutter, and M. Parrinello. A hybrid gaussian and plane wave density functional scheme. *Mol. Phys.*, 92:477–486, 1997.

[39] J. VandeVondele, M. Krack, M. Parrinello F. Mohamed, T. Chassaing, and J. Hutter. Quickstep: fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comp. Phys. Comm.*, 2004. submitted.

[40] M. Krack, A. Gambirasio, and M. Parrinello. Ab initio x-ray scattering of liquid water. *J. Chem. Phys.*, 117:9409–9412, 2002.

[41] V. Buch, F. Mohamed, M. Krack, J. Sadlej, J. P. Devlin, and M. Parrinello M. Solvation states of hcl in mixed ether:acid crystals: A computational study. *J. Chem. Phys.*, 121:12135–12138, 2004.

[42] I. F. W. Kuo, C. J. Mundy, M. J. McGrath, J. I. Siepmann, J. VandeVondele, M. Sprik, J. Hutter, B. Chen, M. L. Klein, F. Mohamed, M. Krack, and M. Parrinello. Liquid water from first principles: Investigation of different sampling approaches. *J. Phys. Chem. B*, 108:12990–12998, 2004.

[43] J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik, and M. Parrinello. The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water. *J. Chem. Phys.*, 122:014515–1–014515–10, 2005.

[44] A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello. Assessing the accuracy of metadynamics. *J. Phys. Chem. B*, 109:6714–6721, 2005.

[45] A. D. Becke. Density-functional exchange-energy approximation with correct aymptotic behaviour. *Phys. Rev. A*, 38(6):3098–3100, 1988.

[46] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the correlation energy. *Phys. Rev. B*, 37(2):785–789, 1988.

[47] M. Sprik, J. Hutter, and M. Parrinello. *Ab initio* molecular dynamics simulation of liquid water: comparison of three gradient-corrected density functionals. *J. Chem. Phys.*, 105(3):1142–1152, 1996.

[48] P. L. Silvestrelli and M. Parrinello. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.*, 82:3308–3311, 1999.

[49] P. L. Silvestrelli and M. Parrinello. Structural, electronic, and bonding properties of liquid water from first principles. *J. Chem. Phys.*, 111:3572–3580, 1990.

[50] S. Raugei and M. L. Klein. Nuclear quantum effects and hydrogen bonding in liquids. *J. Am. Chem. Soc.*, 125:8992–8993, 2003.

[51] B. Chen, I. Ivanov, M. L. Klein, and M. Parrinello. Hydrogen bonding in water. *Phys. Rev. Lett.*, 91:21550311–2155034, 2003.

[52] A. D. Becke. Density-funtional thermochemistry. III. the role of exact exchange. *J. Phys. Chem.*, 98(7):5648–5652, 1993.

[53] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomegeneous elctron gas. *Phys. Rev. B*, 33(12):8822–8824, 1986.

[54] Z. Latajka, Y. Bouteiller, and S. Scheiner. Critical assessment of density functional theory methods for study of proton transfer processes. $(HFH)^-$. *Chem. Phys. Lett.*, 234:159–164, 1995.

[55] J. Backer, J. Andzelm, M. Muir, and P. R. Taylor. $OH + H_2 \rightarrow H_2O + H$. The importance of 'exact echange' in density functional theory. *Chem. Phys. Lett.*, 237:53–60, 1995.

[56] S. Parthiban, G. de Oliveira, and J. M. L. Martin. Benchmark ab initio energy profiles for the gas-phase S_N2 reactions $Y^- + CH_3 \rightarrow CH_3Y + X^-$ (X Y = F, Cl, Br). *J. Phys. Chem. B*, 105:895–904, 2001.

[57] S. Raugei, G. Cardini, and V. Schettino. Microsolvation effect on chemical reactivity: the case of $Cl^- + CH_3Br$ S_N2 reaction. *J. Chem. Phys.*, 114:4089–4098, 2001.

[58] S. Raugei and M. L. Klein. *Ab initio* Molecular Dynamics Investigation of the Formyl Cation in the SbF_5/HF Superacid. *J. Phys. Chem. B*, 105:8212–8219, 2001.

[59] E.J. Meijer and M. Sprik. Ab Initio Molecular Dynamics Study of the Reaction of Water with Formaldehyde in Sulfuric Acid Solution. *J. Am. Chem. Soc.*, 120:6345–6355, 1998.

[60] E.J. Meijer and M. Sprik. A density functional study of the addition of water to SO_3 in the gas phase and in acqueous solution. *J. Phys. Chem. A*, 102:2893–2898, 1998.

[61] M. N. Glukhovtsev, A. Pross, and L. Radom. Gas-phase identity S_N2 reactions of halide anions with methyl halides: a high-level computational study. *J. Am. Chem. Soc.*, 117:2024–2032, 1995.

[62] M. N. Glukhovtsev, R.D. Bach, A. Pross, and L. Radom. The performance of B3-LYP density functional theory in describing S_N2 reactions at satured carbon. *Chem. Phys. Lett.*, 260:558–564, 1996.

[63] S. Raugei and M. L. Klein. On the quantum nature of an excess proton in liquid hydrogen fluoride. *ChemPhysChem*, 5:1569–1576, 2004.

[64] O. V. Gritsenko, B. Ensing, P. R. T. Schipper, and E. J. Baerends. Comparison of the accurate Kohn-Sham solution with the generalized gradient approximations (GGAs) for the S_N2 reaction $F^-+CH_3F\rightarrow FCH_3+F^-$: a qualitative rule to predict success of failure of GGAs. *J. Phys. Chem*, A104:8558–8565, 2000.

[65] S. Kristyán and P. Pulay. Can (semi)local density-functional theory account for the london dispersion forces. *Chem. Phys. Lett.*, 229:175, 1994.

[66] E. J. Meijer and M. Sprik. A density-functional study of the intermolecular interactions of benzene. *J. Chem. Phys.*, 105(19):8684–8689, 1996.

[67] W. Kohn, Y. Meir, and D. E. Makarov. van der Waals energies in density functional theory. *Phys. Rev. Lett.*, 80(19):4153–4156, 1998.

[68] S. Raugei, G. Cardini, and V. Schettino. Intermolecular interactions in the $N_2 - N_2$ dimer. *Mol. Phys.*, 95:477–480, 1998.

[69] X. Wu, M. C. Vargas, S. Nayak, V. Lotrich, and G. Scoles G. Towards extending the applicability of density functional theory to weakly bound systems. *J. Chem. Phys.*, 115:8748–8757, 2001.

[70] U. Zimmerli, M. Parrinello, and P. Koumoutsakos. Dispersion corrections to density functionals for water aromatic interactions. *J. Chem. Phys.*, 120:2693–2699, 2004.

[71] S. Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *J. Comp. Chem.*, 25:1463–1473, 2004.

[72] M. Bernasconi, M. Parrinello, G.L. Chiarotti, P. Focher, and E. Tosatti. Anisotropic *a*-c:h from compression of polyacetylene. *Phys. Rev. Lett.*, 76:2081–2083, 1996.

[73] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. Optimization of effective atom centered potentials for london dispersion forces in density functional theory. *Phys. Rev. Lett.*, 93:153004, 2004.

[74] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. Variational optimization of effective atom centered potentials for molecular properties. *J. Chem. Phys.*, 122:014113, 2005.

[75] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. Performance of optimized atom-centered potentials for weakly bonded systems using density functional theory. *Phys. Rev. B*, 71:195119, 2005.

[76] F. Alber and P. Carloni. Alber f, carloni p ab initio molecular dynamics studies on hiv-1 reverse transcriptase triphosphate binding site: Implications for nucleoside-analog drug resistance. *Protein Sci.*, 9:2535–2546, 2000.

[77] J. Tomasi and M. Persico. Molecular-interactions in solution - an overview of methods based on continuous distributions of the solvent. *Che. Rev.*, 94:2027–2094, 1994.

[78] M. Eichinger, P. Tavan, J. Hutter, , and M. Parrinello. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.*, 110:10452–10467, 1999.

[79] A. Laio, J. VandeVondele, and U. Rothlisberger. A hamiltonian electrostatic coupling scheme for hybrid car-parrinello molecular dynamics simulations. *J. Chem. Phys.*, 116:6941–6947, 2002.

[80] A. Laio, J. VandeVondele, and U. Rothlisberger. D-RESP: Dynamically generated electrostatic potential derived charges from quantum mechanics/molecular mechanics simulations. *J. Phys. Chem. B*, 106:7300–7307, 2002.

[81] M.C. Colombo, L. Guidoni, A. Laio, A. Magistrato, P. Maurer, S. Piana, U. Röhrig, K. Spiegel, M. Sulpizi, J. VandeVondele, M.Zumstein, , and U. Rothlisberger. Hybrid qm/mm car-parrinello simulations of catalytic and enzymatic reactions. *Chimia*, 56:13–19, 2002.

[82] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471–2474, 1985.

[83] D. Marx and J. Hutter. Ab intio molecular dynamics: theory and implementation. In J. Grotendorst, editor, *Modern Methods and algorithms of quantum chemistry*, volume 1 of *NIC series*, pages 301–349, Jülich, 2000. John von Neumann Institute for Computing.

[84] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. *J. Chem. Phys.*, 120:300–311, 2004.

[85] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. ii. *J. Chem. Phys.*, 121:5400–5409, 2004.

[86] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.

[87] A. Laio and M. Parrinello. Escaping free-energy minima. *PNAS*, 99:12562–12566, 2002.

[88] D. Passerone and M. Parrinello. Action-derived molecular dynamics in the study of rare events. *Phys. Rev. Lett.*, 87:108302–, 2001.

[89] R. Elber. Reaction path studies of biological molecules. In R. Elber, editor, *Recent developments in theoretical studies of proteins*. World Scientific, 1996.

[90] C. Dellago, P. G. Bolhuis, and D. Chandler. On the calculation of rate constants in the transition path ensemble. *J. Chem. Phys.*, 110:6617–6624, 1999.

[91] M. Sprik and G. Ciccotti. Free energy from constrained molecular dynamics. *J. Chem. Phys.*, 109:7737–7744, 1998.

[92] P. E. Blöch, H. M. Senn, and A. Togni. Molecular reaction modelling from ab-initio molecular dynamics. In D. G. Truhlar and K. Morokuma, editors, *Transition state modelling for catalysis*, ACS Symposium Series, pages 89–99, Washington, DC, 1999. American Chemical Society.

[93] M. Cascella, S. Raugei, and P. Carloni. Formamide hydrolysis investigated by multiple steering ab-initio molecular dynamics. *J. Phys. Chem. B*, 108:369–375, 2004.

[94] M. Iannuzzi, A. Laio, and M. Parrinello. Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. *Phys. Rev. Lett.*, 90:238302, 2003.

[95] D. T. Odom and J. K. Barton. Long-Range Oxidative Damage in DNA/RNA Duplexes. *Biochemistry*, 40:8727–8737, 2001.

[96] C. A. Mirkin, R. L. Letsinger, R. C. Mucic, and J. J. Storhoff. A DNA-based method for rationally assembling nanoparticles into macroscopic materials . *Nature*, 382:607–609, 1996.

[97] J. J. Storhoff. Programmed Materials Synthesis with DNA. *Chem. Rev.*, 99:1849–1862, 1999.

[98] A. P. Alivisatos, K. P. Johnson, T. E. Wilson, C. J. Loveth, M. P. Bruchez, and P. G. Schulz. Organization of 'nanocrystal molecules' using DNA. *Nature*, 382:609–611, 1996.

[99] Y. Okata, T. Kobayashi, K. Tanaka, and M. Shimomura. Anisotropic Electric Conductivity in an Aligned DNA Cast Film . *J. Am. Chem. Soc.*, 120:6165–6166, 1998.

[100] E. Braun, Y. Eichen, U. Silvan, and G. Ben-Yoseph. DNA-templated assembly and electrode attachment of a conducting silver wire . *Nature*, 391:775–778, 1998.

[101] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394:539–544, 1998.

[102] C. Ban, B. Ramakrishnan, and M. Sundaralingam. Crystal structure of the self-complementary 5'-purine start decamer d(GCGCGCGCGC) in the Z-DNA conformation. I. *Biophys. J.*, 71:1215–1221, 1996.

[103] K. K. Irikura, B. Tidor, B. R. Brooks, and M. Karplus. Transition from b to z dna: contribution of internal fluctuations to the configurational entropy difference. *Science*, 229:571–572, 1985.

[104] D. L. Beveridge and K. J. McConnell. Nucleic acids: theory and computer simulation. *Current Opin. Struct. Biol.*, 10:182–196, 2000.

[105] G. B. Schuster, editor. *Longe-Range Charge Transfer in DNA I*, volume 236 of *Topics in Current Chemistry*. Springer-Verlag, Heidelberg, 2004.

[106] G. B. Schuster, editor. *Longe-Range Charge Transfer in DNA II*, volume 237 of *Topics in Current Chemistry*. Springer-Verlag, Heidelberg, 2004.

[107] S. O. Kelley, N. M. Jackson, M. G. Hill, and J. K. Barton. Long-Range Electron Transfer through DNA Films. *Angew. Chem. Int. Ed.*, 38:941–945, 1999.

[108] T. Douki, J. L. Ravanat, D. Angelov, J. R. Wagner, and J. Cadet. Effects of duplex stability on charge-transfer efficiency within dna. *Top. Curr. Chem.*, 236:1–25, 2004.

[109] R. Holmlin, P. Dandliker, and J. Barton. Charge transfer through the dna base stack. *Angew. Chem. Int. Ed.*, 36:2714–2730, 1997.

[110] A. Yu. Kasumov, M. Kociak, S. Gueron, B. Reulet, V. T. Volkov, D. V. Klinov, and H. Bouchiat. Proximity-induced superconductivity in dna. *Science*, 291:280–282, 2001.

[111] D. Porath, A. Bezryadin, S. deVries, and C. Dekker. Direct measurements of electrical transport through dna molecules. *Nature*, 403:635–638, 2000.

[112] P. J. de Pablo, F. Moreno-Herrero, J. Colchero, J. Gómez Herrero, P. Herrero, A. M. Baró, P. Ordejón, J. M. Soler, and E. Artacho. Absence of dc-Conductivity in *lambda*-DNA. *Phys. Rev. Lett.*, 85:4992–4995, 2000.

[113] C. Gomez-Navarro, F. Moreno-Herrero, P. J. de Pablo, P. J. Colchero, J. Gmez Herrero, and A. M. Bar. Contactless experiments on individual dna molecules show no evidence for molecular wire behavior. *Proc. Natl. Acad. Sci. USA*, 99:8484, 2002.

[114] D. Porath, G. Cuniberti, and R. Di Felice. Charge Transport in DNA-based Devices. In G. Schuster, editor, *Long Range Charge Transfer in DNA*, volume 237, pages 183–228, Heidelberg, 2004. Springer-Verlag.

[115] R. G. Endres, D. L. Cox, and R. R. P. Singh. Colloquium: The quest for high-conductance dna. *Rev. Mod. Phys.*, 76:195–214, 2004.

[116] R. Bruinsma, G. Grüner, M. R. D'Orsogna, and J. Rudnik. Fluctuation-facilitated charge migration along dna. *Phys. Rev. Lett.*, 85:4393–4396, 2000.

[117] B. Giese and S. Wessely. The significance of proton migration during hole hopping through DNA . *Chem. Commun.*, 20:2108–2109, 2001.

[118] S. Steenken. Electron transfer in dna? competition by ultra-fast proton transfer? *Biol. Chem.*, 378:1293–1297, 1997.

[119] V. Shafirovich, A. Dourandin, and N. E. Geacintov. Proton-coupled electron-transfer reactions at a distance in dna duplexes: Kinetic deuterium isotope effect. *J. Phys. Chem. B*, 105:8431–8435, 2001.

[120] R. N. Barnett, C. L. Cleveland, A. Joy, U. Landman, and G. Schuster. Charge Migration in DNA: Ion-Gated Transport. *Science*, 294:567–571, 2001.

[121] O. Schiemann, N. J. Turro, and J. K. Barton. Epr detection of guanine radicals in a dna duplex under biological conditions: Selective base oxidation by ru(phen)2dppz3+ using the flash-quench technique. *J. Phys. Chem. B*, 104:7214–7220, 2000.

[122] B. G. Miller and R. Wolfenden. Catalytic proficiency: the unusual case of OMP decarboxylase. *Annu. Rev. Biochem.*, 71:847–885, 2002.

[123] W. P. Jencks. Binding energy, specificity, and enzymic catalysis: the Circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.*, 43:219–410, 1975.

[124] A. Warshel, M. Strajbl, J. Villà, and J. Florian. Remarkable rate enhancement of orotidine $5'$-monosphosphate decarboxylase is due to transition-state stabilization rather than to ground-state destabilization. *Biochemistry*, 39:14728–14738, 2000.

[125] A. Radzicka and R. Wolfenden. A proficient enzyme. *Science*, 267:90–93, 1995.

[126] N. Wu, Y. Mo, J. gao, and E. F. Pai. Electrostatic stress in catalysis: structure and mechanism of the enzyme orotidine monophosphate decarboxylase. *PNAS*, 97:2017–2022, 2000.

[127] T. C. Appleby, C. Kinsland, T. P. Begley, and S. E. Ealick. The crystal structure and mechanism of orotidine $5'$-monosphosphate decarboxylase. *PNAS*, 97:2005–2010, 2000.

[128] M. A. Rishavy and W. W. Cleland. Determination of the mechanism of orotidine $5'$-monosphosphate decarboxylase by isotope effect. *Biochemistry*, 39:4569–4574, 2000.

[129] T. P Begley, T. C. Appleby, and S. E. Ealick. The stuctural basis for the remarkable catalytic proficiency of orotidine $5'$-monophosphate decarboxylase. *Curr. Opin. Struct. Biol.*, 10:711–718, 2000.

[130] S. Hur and T. C. Bruice. Molecular dynamic study of orotidine $5'$-monophosphate decarboxylase in ground and in intermediate state: a role of the 203-218 loop dynamics. *PNAS*, 99:9668–9673, 2002.

[131] R. W. Schoenlein, L. A. Peteanu, R. A. Mathies, and C. V. Shank. The 1st step in vision - femtosecond isomerization of rhodopsin. *Science*, 254:412–415, 1991.

[132] J. E. Kim, M. J. Tauber, and R. A. Mathies. Wavelength dependent cis-trans isomerization in vision. *Biochemistry*, 40:13774–13778, 2001.

[133] R. R. Birge and B. W. Vought. Energetics of rhodopsin photobleaching: Photocalorimetric studies of energy storage in early and later intermediates. *Methods Enzymol.*, 315:143–163, 2000.

[134] I. Frank, J. Hutter, D. Marx, and M. Parrinello. Molecular dynamics in low-spin excited states. *J. Chem. Phys.*, 108:4060–4069, 1998.