Biophysical Chemistry 149 (2010) 1–11

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Review

QM/MM methods: Looking inside heme proteins biochemisty

Victor Guallar *, Frank H. Wallrapp

ICREA Research Professor, Life Science Department, Barcelona Supercomputing Center, Jordi Girona, 29, 08034 Barcelona, Spain

ARTICLE INFO

Article history:
Received 11 January 2010
Received in revised form 15 March 2010
Accepted 16 March 2010
Available online 20 March 2010

ABSTRACT

Mixed quantum mechanics/molecular mechanics methods offer a valuable computational tool for understanding biochemical events. When combined with conformational sampling techniques, they allow for an exhaustive exploration of the enzymatic mechanism. Heme proteins are ubiquitous and essential for every organism. In this review we summarize our efforts towards the understanding of heme biochemistry. We present: 1) results on ligand migration on globins coupled to the ligand binding event, 2) results on the localization of the spin density in compound I of cytochromes and peroxidases, 3) novel methodologies for mapping the electron transfer pathways and 4) novel data on Tryptophan 2,3-dioxygenase. For this enzyme our results strongly indicate that the distal oxygen will end up on the C3 indole carbon, whereas the proximal oxygen will end up in the C2 position. Interestingly, the process involves the formation of an epoxide and a heme ferryl intermediate. The overall energy profile indicates an energy barrier of approximately 18 kcal/mol and an exothermic driving force of almost 80 kcal/mol.

Contents

1. Introduction

The computational modeling, both at the software (methods) and hardware development, has evolved considerably since the early

* Corresponding author.
E-mail address: victor.guallar@bsc.es (V. Guallar).

0301-4622/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2010.03.010

work in the 70s and 80s. Computational predictions are becoming significantly more accurate in predicting enzymatic mechanisms and rates, binding energies, docked structures, complex network pathways, gene annotation, etc. As a consequence, the level of confidence towards computational modeling is growing among scientists, and many experimental laboratories work in close collaboration with computational modelers. Furthermore, some experimentalists routinely use a computational modeling software themselves to analyze their results and obtain a detailed view of the mechanism in atomic resolution. In the near future one could expect that computational techniques would determine, at the initial stages of research projects, many aspects of the experimental work. Simulations will give a level of confidence to certain experiments, reducing cost and labor.

Understanding biochemical mechanisms at atomic and electronic detail is of crucial importance in industrial catalysis, in biomedical research, etc. Theoretical modeling can involve three main steps from a general point of view: 1) building the model, 2) projecting the conformational sampling and 3) mapping the enzymatic chemical process.

The first step involves the construction of a 3-dimensional structure including all atoms needed to properly characterize the biological system. In enzymatic systems this often requires modeling the entire protein or catalytic domain. An alternative is to use a reduced model of the active site, including only dozens of atoms held together by means of harmonic constraints. The small size of the reduced system allows for a full quantum mechanical (QM) treatment of the model. While this has been shown to be a valid approach in many systems [1], the development of mixed quantum mechanics and classical mechanics (QM/MM) techniques has driven many theoreticians towards the description of the entire enzymatic system [2–6]. The computational cost of QM/MM techniques is similar to that of QM methods alone. Furthermore, by using the correct protein constraints, QM/MM methods might reduce significantly the number of theoretical assays. QM methods based on reduced models, however, are still quite useful to track the protein effects and perform comparison studies. When X-ray or NMR 3-dimensional coordinates are not available, it is also possible to use closely related structures to produce a model using comparative techniques, generally called homology modeling. However, when studying a chemical process, small differences in the active site model may produce significant changes in the potential energy surface. Thus, homology modeling in biochemistry should be restricted to systems with a very high level of sequence identity in the active site. Finally, the initial structure will often require extensive manual editing, as a comprehensive check on protonation states, hydrogen bond network, possible crystal contact artifacts, etc. is mandatory.

Once the model is chosen, it is generally not enough to perform the study on a single set of atomic coordinates. The crystal structure (respectively the homology model) may describe an intermediate of and enzymatic reaction with no specific interest for the researcher. Therefore, it might be needed to add some cofactor, protons or electrons to model the active state properly. As a consequence, these additions might result in conformational changes that need to be described. Furthermore, it is also possible that the enzymatic mechanism is enhanced by means of some low frequency collective motion of the system. To account for this conformational sampling effectively, one should discard the electronic degrees of freedom (due to computational expense) and model the system by means of molecular mechanics (MM), using a classical force field. There exist several common techniques such as molecular dynamics [7–10], Monte Carlo [11–13], protein structure prediction algorithms [14,15], robotic algorithms [16], etc. It is possible to speed up the sampling by using models with lesser atomic resolution. Several of these so-called coarse grained models have been introduced, where an entire residue is described by a reduced number of beads [17,18]. Additionally, normal modes from a reduced system, where only the alpha carbon from each protein residue is included in a connectivity network matrix, have also been applied to obtain the low frequency conformational motion of the system [19]. However, performing conformational sampling different from an all atom representation may introduce too many changes in the active site.

The last step is to study the electronic states involved in the chemical process. Although it is possible to model a bond breaking event by parameterization of the classical force fields introducing, for example, coupling between different harmonic or Morse potentials [20], an accurate calculation will require the use of quantum chemistry calculations. As mentioned above, QM/MM methods allow the adequate description of large biological units, such as an enzyme. They join together a quantum and a classical representation of different sectors of a complex condensed phase system. The reactive region of the active site can be treated with a robust ab initio QM methodology, whereas the remainder of the protein can be modeled at the MM level, providing the appropriate structural constraints and electrostatic and van der Waals interactions with the reactive core.

The methodology used in the analysis of the chemical process would depend on the QM Hamiltonian chosen to describe the process. When using less expensive semiempirical QM methods (or small QM regions and basis sets), we can perform short molecular dynamic trajectories, build potential of mean force or free energy profiles for the chemical process [21–23]. If a more expensive Hamiltonian is used, for example at the Hartree–Fock (HF) or density functional (DFT) level of theory, the study will usually narrow down to one or several reaction coordinates. Recently, Thiel et al. applied QM/MM methods in studies on NMR chemical shifts [24] as well as on the stability of redox eliectronmers [25]. Further recent applications include excitation energies of rhodopsin [26] and on the reactive geometry of the HIV-1 protease [27]. Next to the QM/MM method, there exist approaches of utilizing divide-and-conquer algorithms within the QM calculations, firstly introduced by Yang et al [28], these approaches are able to compute large QM regions by iteratively calculating only fragments until convergence [29–32]. Application of these methods are MP2 calculations of a synthetic protein [33] or structural refinement of the photosystem II [31], for example. Very interesting is the combination of accelerated dynamical sampling with DFT based methods, which has started to be used in mapping biochemical events with high accuracy [34]. These methods, however, involve large computational costs and it should be referred to only when necessary. Many times a simpler reaction coordinate will allow for a sufficient qualitative description of the process [2].

Of particular interest are heme proteins. These proteins are ubiquitous and essential for every organism [35–40]. The complexity of the iron containing porphyrin, together with the different active site environments, gives to this group of metalloproteins multiple functions such as oxygen transport, oxidative catalysis, electron transport, etc. Theoretical methods offer a very valuable tool to study the electronic state of the metal center, to map the electronic delocalization and the energy barriers for the catalytic chemical reaction. Recent QM/MM studies focus on the structural and electronic state of compound I of cytochrome P450cam [41–45], of nitric oxide synthase[46], as well as cytochrome c peroxidase and ascorbate peroxidase [42,47], for example. Godfrey et al. published a study on specifying the electronic state of taurine/α-ketoglutarate dioxygenase and predicting the presence of low reaction barriers within the catalytic cycle [48]. Furthermore, Crespo et al. applied QM/MM on identifying the NO detoxification mechanism of oxy-truncated hemoglobin N [49].

Within this context we would like to present recent results discovered by our group. We have studied multiple heme systems including globins, cytochromes and peroxidases, and pioneered the combination of protein structure prediction techniques with QM/MM methods. In this review we summarize some of our results and introduce some novel data on tryptophan 2,3-dioxygenase (TDO).

2. Methods

In the following section we summarize the methods employed for the studies reviewed below. We only give a general description of the most common methods used in the different systems, as it is possible to find more details on particular implementations in the corresponding publications.

2.1. Ligand and protein dynamic exploration

To couple ligand and protein dynamics and to perform the protein conformational sampling we have developed our own method. The protein energy landscape exploration (PELE) algorithm combines a steered stochastic approach with protein structure prediction methods, capable of projecting the migration dynamics of ligands in proteins. PELE's heuristic algorithm is based on consecutive iteration of three main steps:

1) Ligand and protein perturbation. Ligand perturbation involves a translation and rotation of the center of mass of the ligand. Protein perturbation involves alpha carbon displacement following an anisotropic network model approach (ANM) [18].
2) Side chain sampling. The algorithm proceeds by placing all side chains within a given distance from the ligand (user defined) using the algorithms designed by Prof. Jacobson at UCSF [14].
3) Energy minimization. The last step in every move involves the minimization of a region including, at least, all residues local to the atoms involved in steps 1 and 2.

These three steps compose of a move, which is accepted (defining a new minima) or rejected based on a Metropolis criterion for a given temperature. Typically, a simulation involves several processors running multiple steps and sharing information towards addressing a common task. The task might include minimization of global energy or ligand binding energy, driving the ligand far from a present or fixed point, etc. The method and initial tests have been published showing the capabilities for millisecond landscape exploration [13,50].

2.2. QM/MM

All QM/MM calculations were performed with the Qsite program [51]. Most applications used the DFT B3LYP functional [52,53] and 6-31G* basis set (including a lacv3p pseudopotential for metals) [54]. Starting conformations for QM/MM analysis were obtained after a short MD equilibration of the system, aiming mostly for the explicit water equilibration, and by removing all water molecules beyond 10 Å from the protein surface. All geometry optimizations were then performed by constraining the outermost 5 Å of the water layer. We should emphasize here that in most calculations we aim only for a qualitative view. Recent QM/MM studies indicate that in order to obtain quantitative results it is necessary to use large QM regions [55,56]. Additionally, when dealing with iron spin states, the basis set and DFT functional used do not allow for a quantitative analysis neither. Thus, care must be taken when applying QM/MM techniques. When multiple close electronic states may exist, some basis set and DFT analysis should be conducted even to reach a qualitative view.

QM/MM methods add also an additional level of complexity when compared with QM calculations. The addition of the MM atom electrostatic terms in the one electron Hamiltonian makes it more difficult to obtain converged energy profiles, requiring often multiple runs of the reaction coordinate back and forth. They also require a preliminary study in order to properly position the QM/MM boundary region. Another shortcoming, when compared with QM techniques, is the necessity of a proper description of the enzymatic environment. Small changes in the vicinity of the QM region can have a large effect on the energy profiles for the chemical reaction. These changes might be the result of an inaccurate initial crystal structure, from the wrong interpolation of missing atoms, from the incorrect assignment of protonation states, etc.

2.3. QM

DFT QM calculations on model systems were performed with the Jaguar program [57], and the same functionals and basis sets as in the QM/MM analysis. Gas phase second-order Moller Plesset (MP2) QM calculations were performed using the Gaussian03 program [58].

2.4. MD

MD analysis was performed for model equilibration and, in some systems, for analysis of conformational states. MD was obtained with NAMD version 2.6 [8]. Proteins were usually solvated in a 0.5 M NaCl solution in periodic boxes. Systems were then equilibrated by a minimization followed by 10 ps of MD in the NVT ensemble followed by 0.5 ns each in the NPT and NVT ensembles with a step size of 1 fs. If necessary, production runs usually consisted of ~5 ns with a 2 fs step size in the NVT ensemble using the particle mesh Ewald method to treat long-range electrostatics [59].

2.5. Model setup

Model setup routinely includes the crystal (or NMR structure) inspection followed by water soaking and system equilibration. We first protonate the system, correct the prosthetic group and manually inspect all histidines, glutamines and asparagines, aiming for a correct hydrogen bond network. Next we check ionic groups for possible obvious changes in protonation states. All water molecules present in the crystal structures are always included in the models. Furthermore, if some water molecule may act in the chemical process it is included in the quantum region. We proceed by inspecting possible crystal contact artifacts (in X-ray structures). This step involves mainly surface side chains and protruding loops. The system is then placed in a box of pre-equilibrated waters including ions. We further equilibrate the waters constraining all C-alpha positions (unless some crystal contact artifact was detected) and reduce the final system by removing all water molecules beyond 10 Å from the protein surface.

3. Review of previous results

3.1. Globin studies

We have performed various biophysical and biochemical studies on globin proteins, including myoglobin, hemoglobin and truncated hemoglobins, results of which we are showing in the following section.

3.1.1. Myoglobin carbon monoxide binding and migration

By means of QM/MM methods we first studied the carbon monoxide (CO) deligation event in myoglobin (Mb) [60]. We found that there is a transition from a six coordinated singlet ligated species into a quintet five coordinated heme. In a previous work, using time dependent DFT and a diatomic in molecules semiempirical approach, we observed multiple crossing of quintet and triplet excited states along the carbon monoxide deligation coordinate [61]. The QM/MM energy profiles also indicate a very close proximity of these two spin states. The quintet species highly correlates with the high-resolution crystallographic deligated structure. We also applied protein structure prediction methods along the deligation reaction coordinate aiming to reproduce and understand the overall conformational changes associated with the ligand bond breaking. Our results solved an apparent disagreement between two sets of high-resolution crystal structures for MbCO and deoxyMb. The main effect observed after CO dissociation is a concerted rotation of the E and F helices, which

hold the heme like a clamshell. The rotation, a response to deligation forces, pushes the F helix away from the heme because of the iron spin conversion, allowing the E helix to collapse toward the heme as nonbonded contacts on the distal side are relieved. Additional helix and loop conformational changes stem from these primary events. In particular the tight packing around the active site is used to propagate motion ~25 away from the active site center.

Myoglobin has also been studied in our group with PELE, the algorithm we developed for studying ligand migration in proteins [13]. In our early studies, where protein flexibility was reduced to geometry optimization following the initial perturbation step, minimization (in current versions normal modes are used), most of the protein exit pathways involved the crossing of the CD loop. In this path we find a cavity containing three of the six Phe residues present in Mb. This result was consistent with mutation studies by Scott et al. where mutation of these Phe residues largely affected the entry and escape kinetics of the ligand [62]. Mutation of all residues in the vicinity of the other cavities (Xe1, Xe4, etc.) did not show any significant effect on the ligand kinetics (see, for example, Fig. 5 in [62]). Our results indicate that the Phe residues participate actively in the migration path. Performing quantum chemical (MP2 with cc-pvqz basis set) calculations on a CO molecule placed on top of a benzene ring, we obtained ~1.5 kcal/mol energy stabilization. Furthermore, mutational studies by Schotte et al. have shown a direct crystallographic evidence of a CO molecule being stacked with a Phe residue, the L29F mutated residue [63]. Thus, it appears as the Phe residues can guide and stabilize diatomic molecules with π electronic density next to the most flexible region (CD loop) in the system. Additionally, the different stacking energies might differentiate the migration dynamics of different ligands. The importance of the CD region on the globin heme-binding family has been corroborated by Alberti et al. [64]. Using Gaussian network model analysis the authors showed the existence of conserved low frequency collective modes involving the CD helixes and the loop connecting them.

The current PELE version is able to explicitly add a backbone perturbation by means of an alpha carbon displacement. We have performed additional runs with this new perturbation scheme, which has been shown to give good results in other globin systems (see below TrHbO, for example). Fig. 1 indicates the escape pathways observed when running 40 trajectories (each trajectory consisting of 48 CPU hours in 12 processors). Here, half of the pathways still prefer the CD loop area as an escape path. However, further pathways appear involving a larger ligand migration, which can be seen in Fig. 1. These pathways are in good agreement with molecular dynamics studies [65].

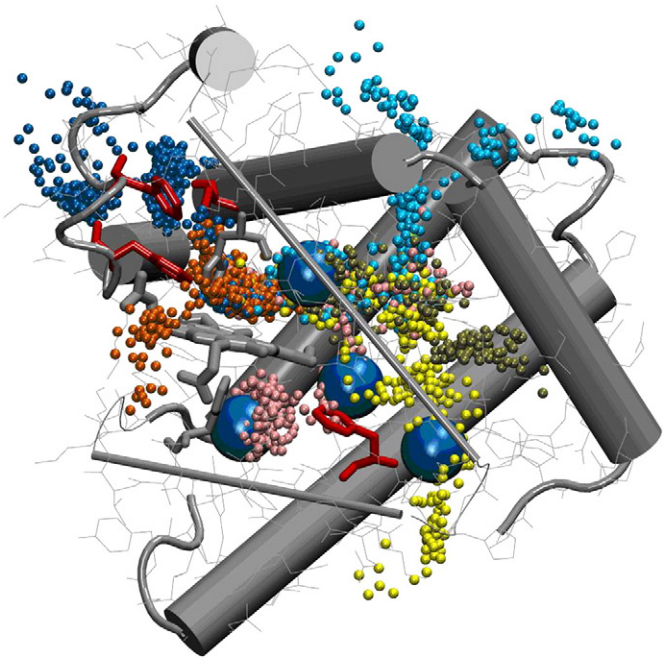

Fig. 1. Multiple exit pathways found in Mb with PELE. Carbon monoxide migration pathways are shown with small spheres in different colors. Big blue spheres depict the xenon cavities. The heme group and proximal histidine are shown in grey (stick representation). Phenylalanines are shown in red (stick representation).

3.1.2. Hemoglobin T/R states binding energies

By combining protein structure prediction algorithms with QM/MM methods we were able to obtain an atomic description for the carbon monoxide binding mechanism in both the deoxy T and the oxy R states of human Hemoglobin [66]. To prepare the system, we performed ligand migration in the T state and ligand dissociation in the R state. Conformational sampling in the active site unbound models was produced by means of PELE. After the conformational clustering, QM/MM methods were used to obtain carbon monoxide binding energies. The results clearly indicate the decrease of affinity by ~7 kcal/mol in the T state, the first ab initio theoretical prediction of the dependence of binding affinity on allostery. By mixing quantum mechanical methods and molecular mechanics, the simulation allows for a description of the atomic mechanism responsible for the change in binding energy. Approximately 30–40% of this difference between the relaxed and tense protein quaternary structures result from the protein strain induced in the heme and its ligands, especially in one of the porphyrine pyrrole rings. As expected from the Perutz models [67], significant energy differences come from the proximal histidine strain. Interestingly, about half of the energy difference arises from protein contacts, involving mainly those residues responsible for locking the quaternary changes at the "hinge" region of the $\alpha1\beta2$ quaternary interface, for example contacts involving the FG corner. Small fluctuations in these contact regions are propagated by means of the highly evolved packing in the quaternary system.

3.1.3. Truncated hemoglobin

TrHbO is a globin protein from *Mycobacterium tuberculosis* with unclear physiological functions and antibacterial target interest. The G8Trp residue (a triptophan residue in the 8th position of the G helix), which is highly conserved, is believed to play a critical role in ligand binding and stabilization [68,69]. The mechanism by which this residue tunes the protein-ligand interactions was modeled with PELE and QM/MM methods [50]. The study was complemented with mutational experimental studies by Prof. Syun Ru Yeh, at Albert Einstein School of Medicine.

Applying QM/MM methods we studied the carbon monoxide deligation from the heme group for the WT and the $G8_{WF}$ mutant species. The energy profiles clearly indicated the importance of the G8 Trp in the active site binding energies. We observed that the tryptophan sterically hinders the ligand deligation. Furthermore, we found that the G8 Trp is important in anchoring the CD1 Tyr and E11 Leu side chain groups, allowing the stabilization of the heme-bound ligand via H-bonds donated from the G8 Trp and the CD1 Tyr.

After deligation, the ligand dynamics in the protein frame was studied with PELE. As illustrated in Fig. 2 top panel, the dissociated ligand immediately migrates to an active site cavity above the heme iron center, denoted as docking site A. Subsequently, the CO finds a quick transition to an additional active site cavity, denoted as docking site B, which is located at 5.2 Å away from the heme iron center. The ligand escapes the active site populating the C cavity and bypassing the G8 Trp. This mechanism is shared by both the WT and the mutant. The two bottom panels indicate the energy profile for this ligand

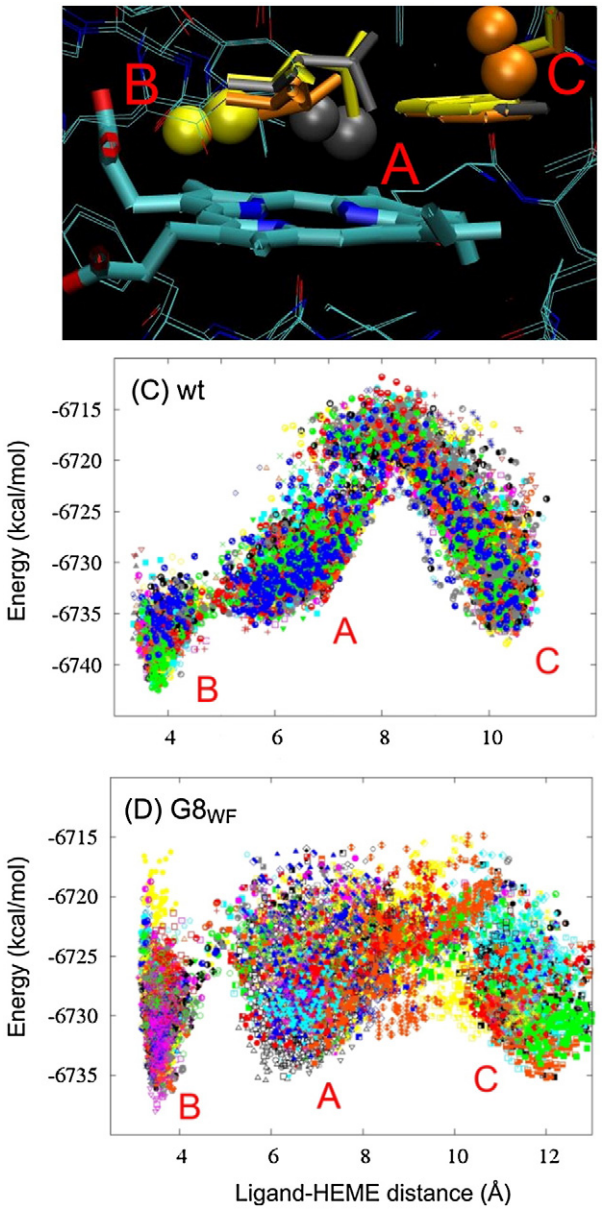

Fig. 2. Top panel: representative snapshots following CO dissociation. The heme, E11 Leu and G8 Trp are shown in a stick presentation. Energy landscape associated with CO migration in the wild type (middle panel) and $G8_{WF}$ mutant (bottom panel).

migration. Each point with different color represents a minimum found by a different processor in PELE. It is clearly seen how the G8F mutation largely facilitates the escape of the ligand from the active site, by reducing the energy barrier for the A to C transition. The energy profiles are almost quantitatively equal to the kinetic observations by our coworkers [50].

3.2. Cytochrome and peroxidase studies

In addition to the research on globin proteins, we have performed several studies on cytochromes and peroxidases. Besides ligand migration in Cytochrome P450, these involve mainly QM/MM studies of the distribution of the spin density on different reactive compounds and on mutants. We have also performed electron transfer calculations by means of a new algorithm, the QM/MM e-Pathway [70].

3.2.1. Camphor migration in cytochrome P450

The camphor binding pathway in cytochrome P450 has been the subject of intensive study and therefore it was chosen as one of our initial benchmark studies for PELE. Recent crystallographic structures trapped two different synthetic molecular wires, indicating the importance of the motion of the F and G helix in the substrate entry pathway [71]. We performed 10 different exploration paths with different Metropolis temperatures ranging from 300–2000 K. All paths, even at 2000 K, resulted in camphor leaving the active site closely following the crystallographic wire ligand [13]. The right panel in Fig. 3 indicates the superposition of several camphor snapshots (green) along the exit pathway with one of the crystallographic carbon wires (red stick presentation). On the left panel in the same figure. we can observe the motion of helixes F and G along the camphor exit pathway. It demonstrates that the ligand diffusion is mainly gated by the movements of several phenylalanine residues (Phe87, Phe98 and Phe193, shown in yellow). The important role of these Phe residues agrees with recent observations using molecular dynamics techniques by Wade et al. [72] and Shaik et al. [73].

3.2.2. Compound I

Of particular interest in heme biochemistry is Compound I (Cpd I), a highly reactive oxyferril intermediate species known to be a major player in oxidative catalysis in cytochromes and peroxidases [74–76]. Cpd I is an iron(IV) oxyferryl species with two unpaired electrons in the iron-oxo moiety and a third unpaired electron found in a porphyrin or protein radical. Here, small differences in the active site can introduce significant changes in the spin localization. Cytochrome c peroxidase (CcP), ascorbate peroxidase (APX) and the bifunctional catalase–peroxidase (mKatG) enzymes constitute a good example of this diversity. These enzymes have similar active sites but the third unpaired electron locates into a proximal tryptophan residue in CcP (Trp191), into the porphyrin ring in APX and into a distal tryptophan in mKatG (Trp107) [77]. Due to its high reactive properties, characterization of Cpd I has been particularly elusive [78]. Computational modeling allows the study of the electron delocalization by means of solving the time independent Schrödinger equation.

Modeling of Cpd I spin density for the three enzymes is shown in Fig. 4. For each system, the quantum region included the heme group, the axial histidine (covalent link of the heme group to the protein), and the distal and axial tryptophan residues. As expected, two unpaired electrons are located in the d and p orbitals of the iron-oxo moiety. The main difference between the systems is the localization

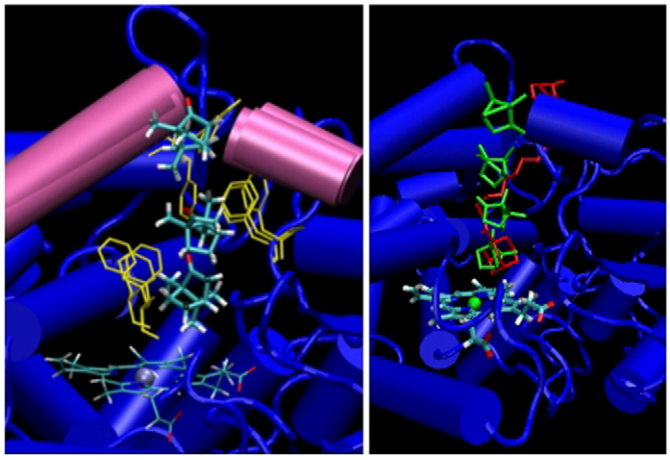

Fig. 3. Camphor migration in cytochrome P450cam obtained with PELE.

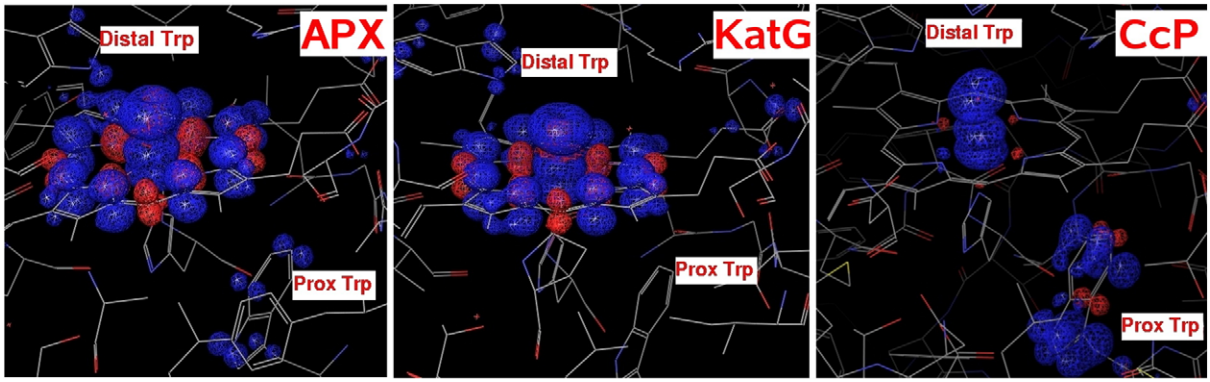

Fig. 4. Spin densities obtained for the CcP, APX and mKatG active sites.

nature of the third unpaired electron. Within the CcP system the unpaired electron is localized at the proximal Trp. In mKatG the third unpaired electron is mostly localized in the porphyrin group but with a notable contribution in the distal tryptophan. In the APX enzyme the unpaired electron is more delocalized, with a main contribution in the porphyrin ring but with a small component in the proximal Trp and a minimal one in the distal Trp. The theoretical spin densities obtained largely agree with the experimental knowledge of the spin distribution for these three systems [77,79].

Interestingly, for all systems the third unpaired electron shows some delocalization of spin density into the heme propionate lone pair orbitals. In previous studies, motivated by the experimental work of Barrows et al. [79], we could transfer almost 0.4 electron units from the porphyrin moiety into the tryptophan by altering the electrostatic screening of the APX's propionates [80]. Further analysis of the role of the propionates in peroxidases indicated that this carboxylic group could play an active role in the transport of electrons in/out the heme group. We named this mechanism “the propionate e-pathway”.

Another example of spin density plasticity was obtained in our collaborative work with Prof. Emma Raven on cytochrome c peroxidase [81]. We studied the evolution of spin densities on cytochrome c peroxidase wild type and the W191F mutant. The theoretical results indicate a shift in spin density from Trp191 to the porphyrin center and a distal tryptophan, Trp51 upon mutation. The radical component in Trp51 agrees nicely with the partial formation of a covalent link between Trp51 and heme observed by our coworkers.

3.2.3. Tracking electron migration pathways

In order to track electron migration, we have proposed a new algorithm called the QM/MM e-Pathway [70]. The underlying idea is that regions contained in the MM part cannot allocate electrons. Hence, if an unpaired electron is present in the system it must necessarily occupy the quantum region. Thus, the method strategy is based on modifying the QM region following the evolution of the spin density. For example, if an electron transfer pathway between two distant redox centers is studied, the methodology consists of the following procedure. As the initial setup, the parameters for the donor and the acceptor in the neutral state are derived, meaning both sites are either oxidized in the case of excess electron transfer or reduced in the case of hole transfer. In both cases the electron (hole) has left one site but has not yet arrived at the other. The parameterization consists of a QM/MM energy minimization of the donor and acceptor (both in their respective oxidation state) from which we extract the geometry and atomic charges from the electrostatic potential. This parameterization is important because in the next step of the procedure both residues are going to be included only in the classical region. Thus, there is going to be no electronic description of the donor and acceptor. Instead, the focus is on the “transfer region” in-between, the region that now contains the electron (hole). After this setup an iterative process is started where initially the entire transfer region is included in the QM region of the QM/MM calculation. The first acceptor of the electron is found by performing a single point calculation of the system where one electron is added (or subtracted if a hole transfer process is studied) to the QM region; now in a doublet spin state. Proceeding to the next iteration, the previously identified residue is turned into a classical residue by excluding it from the quantum region. In doing so, the method does not allow for an electronic description of it and thus, the electron (hole) needs to find its next host. The procedure is repeated until the identified residues connect to a direct pathway between the donor and acceptor.

3.2.4. P450cam–Pdx electron transfer

Although the mechanism of interaction and electron transfer for the redox couple of cytochrome P450 camphor with putidaredoxin (P450cam–Pdx) has been under investigation for over 30 years, the exact mechanism and electron transfer pathway has not been fully understood, yet [82–84]. Applying the above-explained procedure, we reported the first ab initio quantum chemistry description of the electron migration between the two active sites [85]. Starting from the two separate crystal structures of P450cam and Pdx, we used the docking program pyDock [86] as well as the protein modeling program PELE [13] to generate a few complex structures having iron–iron distances between 12.7 and 15.7 Å and being very similar to the structures proposed in the literature [87–89]. The results of mapping the excess electron transfer pathway between the Fe_2S_2 cluster of Pdx and the heme of P450cam clearly identify residue Asp38 of Pdx and Arg112 of P450cam as the key residues within the electron transfer pathway. Here, the electron is believed to migrate from the iron–sulfur cluster to the carbonyl group of Asp38, then jump onto P450cam into the side chain of Arg112 and continue by moving to Cys357, being six coordinated to the iron of the heme. A further possible pathway goes through Asp38 before directly jumping from Arg112 onto the heme through one of the propionate being in very close vicinity of 2.8 to 5.1 Å in the four examined complex conformations.

3.2.5. CcP–Cyt c electron transfer

We also investigated the electron transfer pathway of another heme electron transfer, namely cytochrome c peroxidase with cytochrome c (CcP–Cyt c), based on the crystal structure of the cross linked complex provided by Guo et al. [90]. Here, an electron is known to travel from the heme in Cyt c to the heme in CcP being more than 25 Å away from each other. The study also applies the QM/MM e-Pathway approach by having both hemes parameterized in their oxidized states and only having the transfer region under the QM description [70]. When studying the electron injection, the method iteratively identifies residues Phe82 and Cys81 of Cyt c and residues

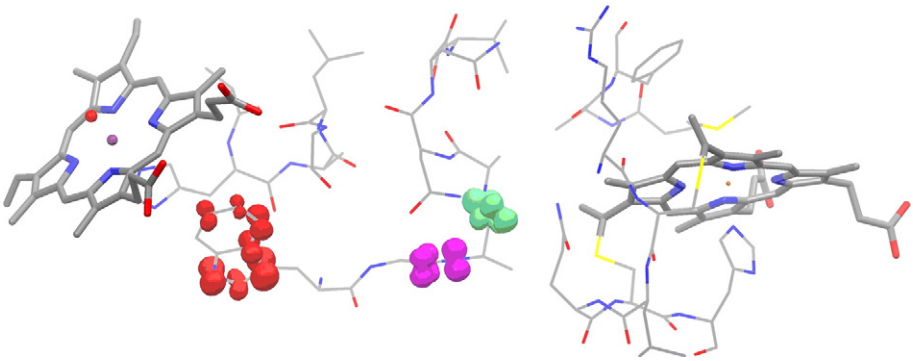

Fig. 5. Main ET pathway obtained with the QM/MM e-Pathway method in CcP–Cyt c.

Asn196, Ala176 and His175 of CcP as the key amino acids involved in the electron transfer. The electron hole procedure, which is expected to be active in this system after Cpd I formation in CcP, identifies Trp191, Gly192 and Ala194 on CcP as the main pathway, Fig. 5, in agreement with a previously proposed pathway [91].

3.2.6. Ascorbate peroxidase

A similar approach, involving switching on/off residues in the quantum region, was applied to study the electron transfer pathway between the ascorbate peroxidase and its substrate ascorbate [92]. In Fig. 6 it is possible to observe how adding/removing the ascorbate ligand to the quantum region points to the heme propionate groups as the active components of electron transfer pathway between the ligand and the iron center. An analogous approach and results were obtained for the cytochrome c peroxidase from Pseudomonas aeruginosa [92]. This is one of the simplest di-heme systems with a high potential electron transfer heme and a low potential catalytic heme.

Shown above, the results of applying QM/MM e-Pathway identify several residues which define one electron transfer pathway, which can be extended to various pathways if we proceed with further iterations. The method provides a robust electron affinity ranking of the residues (respectively molecular orbitals) at the *ab initio* level of theory in the transfer region. Once these residues are found, however, it is necessary to address whether they will act as true intermediates, known as sequential hopping mechanism, or will only assist the direct electron transfer pathway between the donor and acceptor, known as bridge-mediated superexchange mechanism. To address this issue it is necessary to perform electron transfer rate calculations.

We have already started working in the computation of the electronic coupling elements and rate constants in small peptides [93]. Working with Trp-(Pro)n-Trp ($n = 1$ to 6) polypeptides, we have applied semiempirical INDO/S and *ab initio* Hartree–Fock together with the generalized Mulliken–Hush method (GMH) [94,95]. We demonstrated that the coupling values strongly fluctuate throughout the peptide dynamics and the mechanism of electron transfer is affected by the presence of solvent through restriction of the conformational space. Our results emphasized the impact of structural fluctuations on the electronic coupling and hence also on protein electron transfer, being well established in the literature [96–102]. Future studies will combine QM/MM e-Pathway with electronic coupling calculations in the protein complexes studied above.

4. Novel results on tryptophan 2,3-dioxygenase

QM/MM reaction mechanism in tryptophan 2,3-dioxygenase Heme. Tryptophan 2,3-dioxygenase (TDO) is a heme enzyme that catalyses oxidation of L-tryptophan to N-formyl-kynurenine,1-3, in a mechanism that involves binding of dioxygen to ferrous iron, shown in Fig. 7 [103]. Following recent studies by Raven et al. we have focused on the direct electrophilic addition to dioxygen [104]. The possibility of the indole proton abstraction using an active site base has also been investigated.

Starting with the crystal structure 1NGK (chain L), protons were added and all Asp, Glu, Lys and Arg residues were modeled in their ionic state. After visual inspection, histidines 75, 181 and 196 were protonated at both epsilon and delta positions, and histidine 45 only at the epsilon position (the remainder adopting the default delta protonation). The water molecule (HOH 2) at the axial heme position was used to model the dioxygen molecule. The system was solvated and equilibrated at 300 K by means of a 2 ns molecular dynamics using the Desmond package [105]. A layer of 10 Å water molecules was kept for the QM/MM simulations, following the procedure described earlier [106]. All QM/MM calculations were performed with the Qsite program [51], using the unrestricted DFT B3LYP level

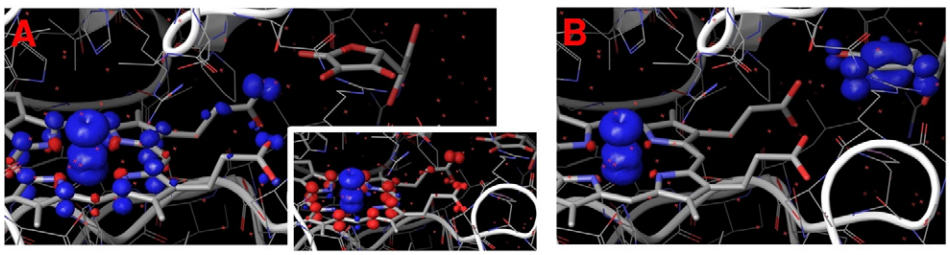

Fig. 6. Spin density for putative active species, compound I, in ascorbate peroxidase. In panel A the quantum region includes the heme group, the oxo ligand and the proximal histidine. In panel B the quantum region also includes the substrate ascorbate. Main pictures in each panel show the quartet spin state; the inset on the right bottom corner in panel A shows the doublet spin state.

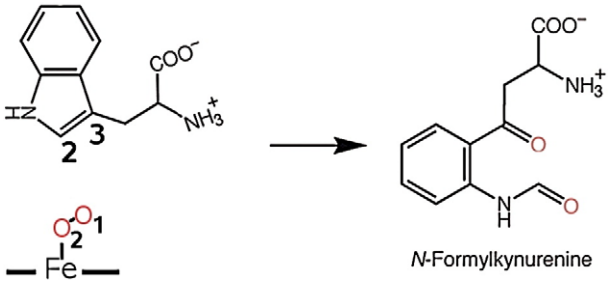

Fig. 7. L-tryptophan to N-formyl-kynurenine biochemical process in TDO.

of theory and 6-31G* basis set (including the lacv3p pseudo potential for the iron center). Unless noted otherwise, the quantum region included the heme group, the axial histidine and dioxygen, and the tryptophan ligand.

The QM/MM geometry optimization of the reactants, namely the ferrous dioxygen species, was first carried out in the singlet and in the triplet spin states. As observed previously by Morokuma et al., the singlet state is ~9 kcal/mol more stable than triplet state [107]. Furthermore, the Mulliken population analysis reflects a ferric-superoxide nature, where one spin resides on the Fe atom and the other one delocalizes over the dioxygen axial ligand. Based on previous studies on this enzyme we first modeled the Trp proton abstraction by His55 [104,108]. In the initial crystal structure, this histidine is in close contact with the tryptophan ligand. Moreover the orientation for a hydrogen abstraction is quite favorable, with a hydrogen–nitrogen distance of only 3.2 Å. We built several reaction coordinates for the proton abstraction with different quantum regions, including and excluding the heme group. All our attempts, however, did not produce any stable product for the abstraction. For all different quantum regions the energy profile increases monotonically to ~30 kcal/mol. Thus, the active site does not seem to be optimized for the indole proton abstraction. This is in agreement with the chemistry of indoles, being well documented and not favorable to the loss of the indole proton. Recent studies by Raven et al, mutating His55 and also using a methyl substitution on N1 of tryptophan, also ruled out the possibility of the hydrogen abstraction mechanism [104].

We proceeded by studying the direct electrophile attack of the indole by the dioxygen axial ligand. In indoles the electrophile becomes preferentially attached at the third position. Thus, we first studied the addition of the distal oxygen to C3 (see Fig. 7). The blue squares in Fig. 8 indicate the energy profile for this bond formation reaction coordinate. The potential is quite repulsive, with an associated barrier of ~39 kcal/mol and an endothermicity of ~37.5 kcal/mol. A free geometry optimization produces a stable product. The addition of the zero point energy along the C–O reaction coordinate, however, might turn the potential energy into a monotonically increasing energy profile. The distortions in the Trp and in the porphyrin rings in this product species indicate that the steric repulsion is a major component of this large endothermic profile. The red rhombus in Fig. 8 indicate the energy profile for the addition of the distal oxygen into C2. Contrary to the attack to C3, we observed a more stable product and reduced energy barrier of ~13.5 kcal/mol. The inset in Fig. 8 shows the geometry for the oxygen addition products after the attack in C2. The new C2–O distance is 1.46 Å, the O–O distance 1.45 Å and the Fe–O distance is 1.81 Å. We name this intermediate as Int1 in the following.

Int1 is equivalent to the first intermediate shown by Morokuma et al. in a recent gas phase reduced model [107]. The gas phase study, however, did not allow to model correctly the protein constraints and could not distinguish properly between C2 and C3. The addition reaction takes place as a radical, and a diradical ferric intermediate is obtained. Starting from this intermediate, we proceed by enlarging the Fe–O and the O–O distances shown in Fig. 8. Enlarging the Fe–O distances quickly increases the energy by ~25 kcal/mol. Increasing the O–O distance, however, produces a very stable intermediate, Int2, where the oxygen forms an epoxide cyclic ether with C2 and C3. Interestingly, we obtain a heme ferryl intermediate in a dioxygenase enzyme. When writing this report, it has appeared an experimental and computational study confirming the existence of this intermediate [109]. The energy barrier for this epoxide formation is ~8 kcal/mol starting from Int1 and ~15/kcal/mol starting from the reactants. Interestingly, Int2 is ~13 kcal/mol more stable than Int1, which translates in being ~6 kcal/mol more stable than the reactants. The B panel in Fig. 8 shows the nature of this epoxide intermediate. The C2–O and C3–O distances are 1.46 and 1.49, respectively.

Proceeding from Int2, we built a reaction coordinate for the attack of the proximal oxygen, O2, on the C2, the closest tryptophan carbon at 2.75 Å, as seen in panel B in Fig. 8. The energy profile for this reaction coordinate indicates a barrier of ~16 kcal/mol with an exothermicity of ~33 kcal/mol to give a new intermediate, Int3. As

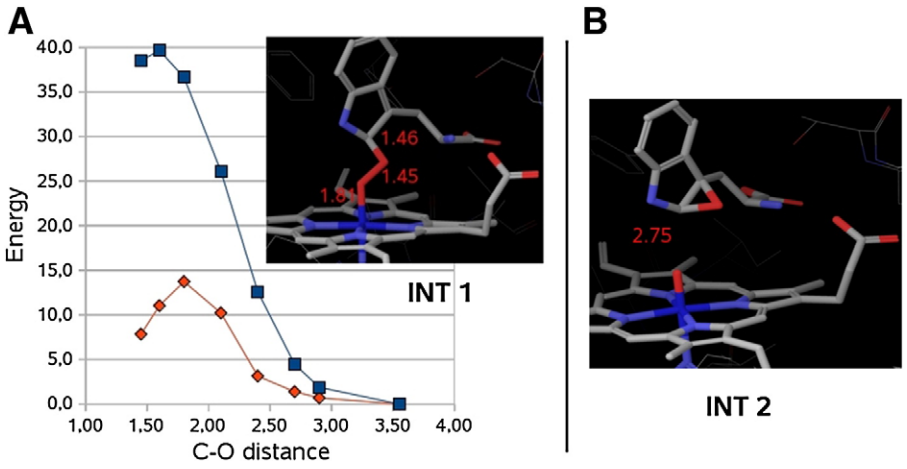

Fig. 8. Panel A: Reaction energy profile for the C2–O1 (blue) and C3–O1 (red) attack. Inset in panel A shows the Int1 intermediate. Panel B: Int2 intermediate.

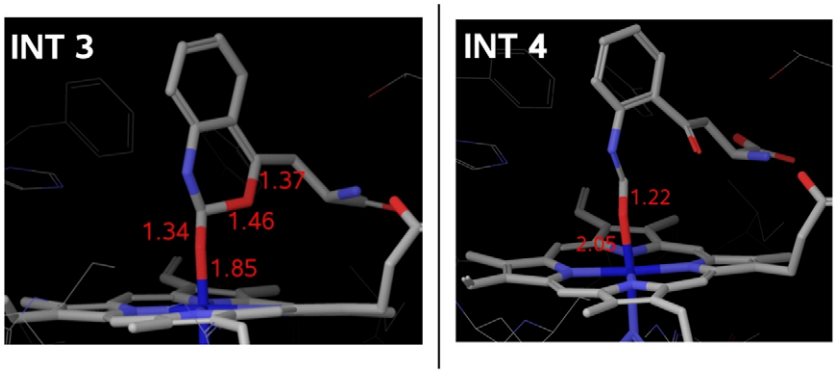

Fig. 9. Int3 (left) and Int4 (right) intermediates. Main distances shown in angstroms.

seen in Fig. 9, in Int3 the epoxide ring "opens" into a 6 member hetero ring with a 1.85 Å Fe–O2 bond. O1 is now forming a stronger bond with C3 than with C2, having bond lengths of 1.37 Å respectively 1.46 Å. Although Int3 is significantly more stable than any previous minimum, it only constitutes a weak intermediate. It sits in a very shallow valley and small vibrations on the C2–O1 mode would bring it spontaneously to a much deeper minimum. When enlarging this distance by only 0.1 Å, associated with a ~1.5 kcal/mol energy increase, the system proceeds to Int4, an intermediate which is ~75 kcal/mol more stable than the reactants. Thus, at this level of theory, the addition of zero point energy along the C2–O1 coordinate might turn Int3 into a non-stationary species in the potential energy. As seen in Fig. 9, Int4 clearly shows the breakage of the C2–O1 bond, the elongation of the Fe–O1 bond and a strengthening of both C2–O1 and C3–O2 bonds to ~1.22 Å.

The final step in the biochemical cycle is now the deligation of the N-formyl-kynurenine product. This step involves now only ~6 kcal/mol and produces an exothermic reaction of ~85 kcal/mol in total. The overall energy profile for the process is shown in Fig. 10. This final step involves a singlet to triplet (or quintet) transition, as seen, for example, in carbon monoxide deligation in myoglobin [60].

These results strongly indicate the initial attack of the distal oxygen (O1) on C2. However O1 will end up in C3 after the epoxide formation and opening, whereas the proximal oxygen (O2) will then end up in C2. Interestingly, the process involves the formation of a heme ferryl intermediate. The overall energy profile indicates an energy barrier of approximately 18 kcal/mol and an exothermic driving force of almost 80 kcal/mol.

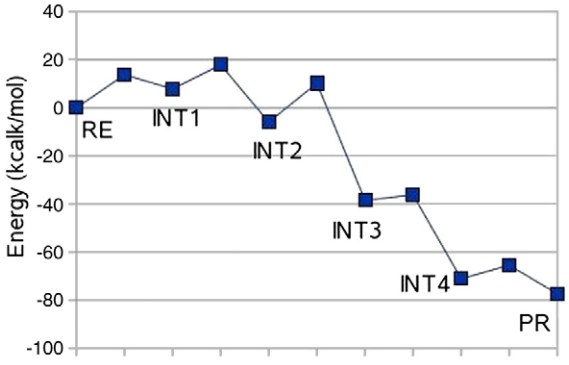

Fig. 10. Overall energy profile for the oxidation of L-tryptophan to N-formyl-kynurenine,1-3. RE and PR indicate reactants and products.

5. Conclusions

We underlined here the possibilities of mixed quantum mechanics and molecular mechanics methods in addressing the enzymatic mechanisms in heme proteins. When combined with protein structure prediction techniques (or other sampling techniques), the coupling of the biochemical and biophysical processes in the enzyme can be studied.

The studies presented here include a wide diversity of enzymes with multiple specific findings. However we can draw some overall conclusions:

1. Distal residues regulating the ligand binding and biochemical process are of great importance. Small changes in these residues introduce large changes in specificity and in the energy barriers associated with the biochemical event or the ligand migration.
2. The importance of these distal residues is mostly based on steric packing. This tight and highly specific packing is also used to propagate information across the enzyme.
3. Phenylalanine residues appear as one of the main gate control in ligand diffusion. Interestingly, they use their electronic properties to attract diatomic molecules with a π bond.
4. DFT methods offer a very valuable tool for mapping differences in spin density and electron transfer processes in heme proteins.

Points 1 and 4 are in line with previous QM/MM studies, many of them summarized in this review. The combination of protein structure prediction techniques with QM/MM methods, however, has allowed the atomic and electronic detailed observation of new mechanisms like the ones described in points 2 and 3.

Acknowledgments

Computational resources were provided by the Barcelona Supercomputing Center. Work was supported by startup funds from the Barcelona Supercomputer Center and through the Spanish Ministry of Education and Science through the project CTQ2007-62122/BQU. We further want to thank all coworkers in the reviewed studies.

References

[1] P.E.M. Siegbahn, M.R.A. Blomberg, Transition-metal systems in biochemistry studied by high-accuracy quantum chemical methods, Chemical Reviews 100 (2000) 421–437.
[2] R.A. Friesner, V. Guallar, Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis, Annual Review of Physical Chemistry 56 (2005) 347–389.
[3] A. Warshel, A. Bromberg, Oxidation of 4A, 4B-dihydrophenanthrenes. 3. A theoretical study of large kinetic isotope effect of deuterium in initiatio step of therma reaction, Journal of Chemical Physics 52 (1970) 1262–1269.

[4] J.L. Gao, D.G. Truhlar, Quantum mechanical methods for enzyme kinetics, Annual Review of Physical Chemistry 53 (2002) 467–505.
[5] P. Sherwood, A.H. de Vries, M.F. Guest, G. Schreckenbach, C.R.A. Catlow, S.A. French, A.A. Sokol, S.T. Bromley, W. Thiel, A.J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S.C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, C. Lennartz, QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis, Journal of Molecular Structure: THEOCHEM 632 (2003) 1–28.
[6] H.M. Senn, W. Thiel, QM/MM Methods for Biological Systems, Springer, Berlin, 2007.
[7] B. Brooks, M. Karplus, Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor, Proceedings of the National Academy of Sciences 80 (1983) 6571–6575.
[8] J.L. Klepeis, K. Lindorff-Larsen, R.O. Dror, D.E. Shaw, Long-timescale molecular dynamics simulations of protein structure and function, Current Opinion in Structural Biology 19 (2009) 120–127.
[9] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, Journal of Computational Chemistry 26 (2005) 1781–1802.
[10] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, Journal of Chemical Theory and Computation 4 (2008) 435–447.
[11] Z.Q. Li, H.A. Scheraga, Monte Carlo-minimization approach to the multiple-minima problem in protein folding, Proceedings of the National Academy of Sciences 84 (1987) 2633–2636.
[12] J.P. Ulmschneider, W.L. Jorgensen, Monte Carlo backbone sampling for polypeptides with variable bond angles and dihedral angles using concerted rotations and a Gaussian bias, Journal of Chemical Physics 118 (2003) 4261–4271.
[13] D.A. Evans, D.J. Wales, Free energy landscapes of model peptides and proteins, Journal of Chemical Physics 118 (2003) 3891–3897.
[14] K.W. Borrelli, A. Vitalis, R. Alcantara, V. Guallar, PELE: Protein energy landscape exploration, A novel Monte Carlo based technique, Journal of Chemical Theory and Computation 1 (2005) 1304–1311.
[15] M.P. Jacobson, D.L. Pincus, C.S. Rapp, J.F. Tyler, B. Honig, D.E. Shaw, R.A. Friesner, A hierarchical approach to all-atom protein loop prediction, Proteins 55 (2004) 351–367.
[16] G. Verbitsky, R. Nussinov, H. Wolfson, Flexible structural comparison allowing hinge-bending, swiveling motions, Proteins: Structure, Function, and Genetics 34 (1999) 232–254.
[17] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, S.J. Marrink, The MARTINI coarse-grained force field: extension to proteins, Journal of Chemical Theory and Computation 4 (2008) 819–834.
[18] A. Liwo, P. Arlukowicz, C. Czaplewski, S. Oldziej, J. Pillardy, H.A. Scheraga, A method for optimizing potential-energy functions by a hierarchical design of the potential-energy landscape: application to the UNRES force field, Proceedings of the National Academy of Sciences 99 (2002) 1937–1942.
[19] I. Bahar, A.R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, Folding & Design 2 (1997) 173–181.
[20] A. Warshel, R.M. Weiss, An empirical valence bond approach for comparing reactions in solutions and in enzymes, Journal of the American Chemical Society 102 (1980) 6218–6226.
[21] J.J. Ruiz-Pernia, E. Silla, I. Tunon, S. Marti, V. Moliner, Hybrid QM/MM potentials of mean force with interpolated corrections, Journal of Physical Chemistry B 108 (2004) 8427–8433.
[22] H. Hu, Z.Y. Lu, W.T. Yang, QM/MM minimum free-energy path: methodology and application to triosephosphate isomerase, Journal of Chemical Theory and Computation 3 (2007) 390–406.
[23] M. Garcia-Viloca, C. Alhambra, D.G. Truhlar, J.L. Gao, Hydride transfer catalyzed by xylose isomerase: mechanism and quantum effects, Journal of Computational Chemistry 24 (2003) 177–190.
[24] K.R. Geethalakshmi, M.P. Waller, W. Thiel, M. Bühl, 51V NMR chemical shifts calculated from QM/MM models of peroxo forms of vanadium haloperoxidases, Journal of Physical Chemistry B 113 (2009) 4456–4465.
[25] J.C. Schoneboom, S. Cohen, H. Lin, S. Shaik, W. Thiel, Quantum mechanical/molecular mechanical investigation of the mechanism of C-H hydroxylation of camphor by cytochrome P450(cam): theory supports a two-state rebound mechanism, Journal of the American Chemical Society 126 (2004) 4017–4034.
[26] M. Wanko, M. Hoffmann, T. Frauenheim, M. Elstner, Computational Photochemistry of Retinal proteins, Journal of Computer-Aided Molecular Design 20 (2006) 511–518.
[27] V. Carnevale, S. Raugei, M. Neri, S. Pantano, C. Micheletti, P. Carloni, Multi-scale modeling of HIV-1 proteins, Journal of Molecular Structure: THEOCHEM 898 (2009) 97–105.
[28] W. Yang, Direct calculation of electron density in density-functional theory, Physical Review Letters 66 (1991) 1438.
[29] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Fragment molecular orbital method: an approximate computational method for large molecules, Chemical Physics Letters 313 (1999) 701–706.
[30] Y. Kiyota, J.-Y. Hasegawa, K. Fujimoto, B. Swerts, H. Nakatsuji, A multicore QM/MM approach for the geometry optimization of chromophore aggregate in protein, Journal of Computational Chemistry 30 (2009) 1351–1359.
[31] E. Sproviero, M. Newcomer, J. Gascón, E. Batista, G. Brudvig, V. Batista, The MoD-QM/MM methodology for structural refinement of photosystem II and other biological macromolecules, Photosynthesis Research 102 (2009) 455–470.
[32] X. He, K.M. Merz, Divide and conquer Hartree–Fock calculations on proteins, Journal of Chemical Theory and Computation 6 (2010) 405–411.
[33] D.G. Fedorov, K. Ishimura, T. Ishida, K. Kitaura, P. Pulay, S. Nagase, Accuracy of the three-body fragment molecular orbital method applied to Moeller–Plesset perturbation theory, Journal of Computational Chemistry 28 (2007) 1476–1484.
[34] B. Ensing, M. De Vivo, Z.W. Liu, P. Moore, M.L. Klein, Metadynamics as a tool for exploring free energy landscapes of chemical reactions, Accounts of Chemical Research 39 (2006) 73–81.
[35] C.J. Reddy, B.R. Gibney, Heme proteins assemblies, Chemical Reviews 104 (2004) 617–649.
[36] M. Sono, M.P. Roach, E.D. Coulter, J.H. Dawson, Heme-containing oxygenases, Chemical Reviews 96 (1996) 2841–2887.
[37] P.J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B.R. Cowen, D. Ehrenstein, H. Frauenfelder, J.B. Johnson, D.C. Lamb, S. Luck, J.R. Mourant, G.U. Nienhaus, P. Ormos, R. Philipp, A.H. Xie, R.D. Young, Ligand-binding to heme-proteins – connection between dynamics and function, Biochemistry 30 (1991) 3988–4001.
[38] J.H. Dawson, Probing structure–function relations in heme–containing oxygenases and peroxidases, Science 240 (1988) 433–439.
[39] S.L. Mayo, W.R. Ellis, R.J. Crutchley, H.B. Gray, Long-range electron transfer in heme proteins, Science 233 (1986) 948–952.
[40] J. Wyman Jr, M.L. Anson, T.E. John, Heme Proteins, Academic Press, New York, 1948.
[41] C.M. Bathelt, J. Zurek, A.J. Mulholland, J.N. Harvey, Electronic structure of compound I in human isoforms of cytochrome P450 from QM/MM modeling, Journal of the American Chemical Society 127 (2005) 12900–12908.
[42] J.N. Harvey, C.M. Bathelt, A.J. Mulholland, QM/MM modeling of compound I active species in cytochrome P450, cytochrome *c* peroxidase, and ascorbate peroxidase, Journal of Computational Chemistry 27 (2006) 1352–1362.
[43] J. Zheng, D. Wang, W. Thiel, S. Shaik, QM/MM study of mechanisms for compound I formation in the catalytic cycle of cytochrome P450cam, Journal of the American Chemical Society 128 (2006) 13204–13215.
[44] A. Altun, S. Shaik, W. Thiel, What is the active species of cytochrome P450 during camphor hydroxylation? QM/MM studies of different electronic states of compound I and of reduced and oxidized iron–oxo intermediates, Journal of the American Chemical Society 129 (2007) 8978–8987.
[45] A. Altun, V. Guallar, R.A. Friesner, S. Shaik, W. Thiel, The effect of heme environment on the hydrogen abstraction reaction of camphor in P450(cam) catalysis: a QM/MM study, Journal of the American Chemical Society 128 (2006) 3924–3925.
[46] K.-B. Cho, E. Derat, S. Shaik, Compound I of nitric oxide synthase: the active site protonation state, Journal of the American Chemical Society 129 (2007) 3182–3188.
[47] S.P. deVisser, What affects the quartet–doublet energy splitting in peroxidase enzymes? Journal of Physical Chemistry (2005).
[48] E. Godfrey, C.S. Porro, S.P. de Visser, Comparative quantum mechanics/molecular mechanics (QM/MM) and density functional theory calculations on the oxo–iron species of taurine/α-ketoglutarate dioxygenase, Journal of Physical Chemistry 112 (2008) 2464–2468.
[49] A. Crespo, M.A. Marti, S.G. Kalko, A. Morreale, M. Orozco, J.L. Gelpi, F.J. Luque, D.A. Estrin, Theoretical study of the truncated hemoglobin HbN: exploring the molecular basis of the NO detoxification mechanism, Journal of the American Chemical Society 127 (2005) 4433–4444.
[50] V. Guallar, C. Lu, K. Borrelli, T. Egawa, S.-R. Yeh, Ligand migration in the truncated hemoglobin-II from mycobacterium tuberculosis: the role of G8 tryptophan, Journal of Biological Chemistry 284 (2009) 3106–3116.
[51] Schrödinger Inc., QSite 5.5, 2007, Portland.
[52] A.D. Becke, Density-functional thermochemistry 3. The role of exact exchange, Journal of Chemical Physics 98 (1993) 5648–5652.
[53] C. Lee, W. Yang, R.G. Parr, Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density, Physical Reviews B 37 (1988) 785–789.
[54] P.J. Hay, W.R. Wadt, Abinitio effective core potentials for molecular calculations – potentials for the transition-metal atoms Sc to Hg, Journal of Chemical Physics 82 (1985) 270–283.
[55] I. Solt, P. Kulhanek, I. Simon, S. Winfield, M.C. Payne, G. Csanyi, M. Fuxreiter, Evaluating boundary dependent errors in QM/MM simulations, Journal of Physical Chemistry B 113 (2009) 5728–5735.
[56] C.V. Sumowski, C. Ochsenfeld, A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems, Journal of Physical Chemistry 113 (2009) 11734–11741.
[57] Schrödinger Inc., Jaguar 7.6, 2000, Portland.
[58] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A. , T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Y. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian 03, Revision A.1, 2003, Pittsburgh PA.
[59] T. Darden, D. York, L. Pedersen, Particle mesh Ewald - an $N\log(N)$ method for Ewald sums in large systems, Journal of Chemical Physics 98 (1993) 10089–10092.
[60] V. Guallar, A. Jarzecki, R.A. Friesner, T.G. Spiro, Modeling of ligation-induced helix/loop displacements in myoglobin; toward an understanding of hemoglobin allostery, Journal of the American Chemical Society 128 (2006) 5427–5435.

[61] C.J. Margulis, V. Guallar, E. Sim, R.A. Friesner, B.J. Berne, A new semiempirical approach to study ground and excited states of metal complexes in biological systems, Journal of Physical Chemistry B 106 (2002) 8038–8046.
[62] E.E. Scott, Q.H. Gibson, J.S. Olson, Mapping the pathways for O-2 entry into and exit from myoglobin, Journal of Biological Chemistry 276 (2001) 5177–5188.
[63] D. Bourgeois, B. Vallone, F. Schotte, A. Arcovito, A.E. Miele, G. Sciara, M. Wulff, P. Anfinrud, M. Brunori, Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography, Proceedings of the National Academy of Sciences 100 (2003) 8704–8709.
[64] S. Maguid, S. Fernandez-Alberti, L. Ferrelli, J. Echave, Exploring the common dynamics of homologous proteins application to the globin family, Biophysical Journal 89 (2005) 3–13.
[65] J.Z. Ruscio, D. Kumar, M. Shukla, M.G. Prisant, T.M. Murali, A.V. Onufriev, Atomic level computational identification of ligand migration pathways between solvent and binding site in myoglobin, Proceedings of the National Academy of Sciences 105 (2008) 9204–9209.
[66] R.E. Alcantara, C. Xu, T.G. Spiro, V. Guallar, A quantum-chemical picture of hemoglobin affinity, Proceedings of the National Academy of Sciences 104 (2007) 18451–18455.
[67] M.F. Perutz, A.J. Wilkinson, M. Paoli, G.G. Dodson, The stereochemical mechanism of the cooperative effects in hemoglobin revisited, Annual Review of Biophysics and Biomolecular Structure 27 (1998) 1–34.
[68] C. Lu, T. Egawa, M. Mukai, R.K. Poole, S.R. Yeh, Hemoglobins from *Mycobacterium tuberculosis* and *Campylobacter jejuni*: a comparative study with resonance Raman spectroscopy, Methods in Enzymology 437 (2008) 255–286.
[69] H. Ouellet, M. Milani, M. LaBarre, M. Bolognesi, M. Couture, M. Guertin, The roles of Tyr(CD1) and Trp(G8) in *Mycobacterium tuberculosis* truncated hemoglobin O in ligand binding and on the heme distal site architecture, Biochemistry 46 (2007) 11440–11450.
[70] V. Guallar, F. Wallrapp, Mapping protein electron transfer pathways with QM/MM methods, Journal of The Royal Society Interface 5 (2008) 233–239.
[71] A.-M.A. Hays, A.R. Dunn, R. Chiu, H.B. Gray, C.D. Stout, D.B. Goodin, Conformational states of cytochrome P450cam revealed by trapping of synthetic molecular wires, Journal of Molecular Biology 344 (2004) 455–469.
[72] S.K. Ludemann, V. Lounnas, R.C. Wade, 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, Journal of Molecular Biology 303 (2000) 797–811.
[73] D. Fishelovitch, S. Shaik, H.J. Wolfson, R. Nussinov, Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism, Journal of Physical Chemistry B 113 (2009) 13018–13025.
[74] T. Egawa, H. Shimada, Y. Ishimura, Evidence for compound I formation in the reaction of cytochrome-P450cam with M-Chloroperbenzoic acid, Biochemical and Biophysical Research Communications 201 (1994) 1464–1469.
[75] L.A. Fishel, M.F. Farnum, J.M. Mauro, M.A. Miller, J. Kraut, Y. Liu, X.L. Tan, C.P. Scholes, Compound I radical in site-directed mutants of cytochrome c peroxidase as probed by electron paramagnetic resonance and electron-nuclear double resonance, Biochemistry 30 (1991) 1986–1996.
[76] M. Wirstam, M.R.A. Blomberg, P.E.M. Siegbahn, Reaction mechanism of compound I formation in heme peroxidases: a density functional theory study, Journal of the American Chemical Society 121 (1999) 10178–10185.
[77] Z. Pipirou, A.R. Bottrill, C.M. Metcalfe, S.C. Mistry, S.K. Badyal, B.J. Rawlings, E.L. Raven, Autocatalytic formation of a covalent link between tryptophan 41 and the heme in ascorbate peroxidase, Biochemistry 46 (2007) 2174–2180.
[78] P. Jones, H.B. Dunford, The mechanism of compound I formation revisited, Journal of Inorganic Biochemistry 99 (2005) 2292–2298.
[79] T.P. Barrows, T.L. Poulos, Role of electrostatics and salt bridges in stabilizing the compound I radical in ascorbate peroxidase, Biochemistry 44 (2005) 14062–14068.
[80] V. Guallar, B. Olsen, The role of the heme propionates in heme biochemistry, Journal of Inorganic Biochemistry 100 (2006) 755–760.
[81] Z. Pipirou, V. Guallar, J. Basran, C.L. Metcalfe, E.J. Murphy, A.R. Bottrill, S.C. Mistry, E.L. Raven, Peroxide-dependent formation of a covalent link between Trp51 and the heme in cytochrome *c* peroxidase, Biochemistry 48 (2009) 3593–3599.
[82] H. Shimada, S. Nagano, H. Hori, Y. Ishimura, Putidaredoxin-cytochrome P450cam interaction, Journal of Inorganic Biochemistry 83 (2001) 255–260.
[83] R.C. Wade, D. Motiejunas, K. Schleinkofer, Sudarko, P. J. Winn, A. Banerjee, A. Kariakin and C. Jung, Multiple molecular recognition mechanisms. Cytochrome P450–a case study, Biochimica et Biophysica Acta 1754 (2005) 239–244.
[84] K. Harada, K. Sakurai, K. Ikemura, T. Ogura, S. Hirota, H. Shimada, T. Hayashi, Evaluation of the functional role of the heme-6-propionate side chain in cytochrome P450cam, Journal of the American Chemical Society 19 (2007) 19.
[85] F. Wallrapp, D. Masone, V. Guallar, Electron transfer in the P450cam/PDX complex. the QM/MM e-pathway, Journal of Physical Chemistry 112 (2008) 12989–12994.
[86] T.M.-K. Cheng, T.L. Blundell, J. Fernandez-Recio, pyDock: Electrostatics and desolvation for effective scoring of rigid–body protein–protein docking, Proteins 68 (2007) 503–515.
[87] T.C. Pochapsky, T.A. Lyons, S. Kazanis, T. Arakaki, G. Ratnaswamy, A structure–based model for cytochrome P450cam–putidaredoxin interactions, Biochimie 78 (1996) 723–733.
[88] A.E. Roitberg, M.J. Holden, M.P. Mayhew, I.V. Kurnikov, D.N. Beratan, V.L. Vilker, Binding and electron transfer between putidaredoxin and cytochrome P450cam. Theory and experiments, Journal of the American Chemical Society 120 (1998) 8927–8932.
[89] A. Karyakin, D. Motiejunas, R.C. Wade, C. Jung, FTIR studies of the redox partner interaction in cytochrome P450: the Pdx–P450cam couple, Biochimica et Biophysica Acta (BBA) – General Subjects 1770 (2007) 420–431.
[90] M. Guo, B. Bhaskar, H. Li, T.P. Barrows, T.L. Poulos, Crystal structure and characterization of a cytochrome *c* peroxidasecytochrome *c* sitespecific cross-link, Proceedings of the National Academy of Sciences 101 (2004) 5940–5945.
[91] H. Pelletier, J. Kraut, Crystal structure of a complex between electron transfer partners, cytochrome *c* peroxidase and cytochrome *c*, Science 258 (1992) 1748–1755.
[92] V. Guallar, Heme electron transfer in peroxidases: the propionate e-pathway, Journal of Physical Chemistry B 112 (2008) 13460–13464.
[93] F. Wallrapp, A. Voityuk, V. Guallar, Solvent effects on donor–acceptor couplings in peptides. a combined QM and MD study, Journal of Chemical Theory and Computation 5 (2009) 3312–3320.
[94] R.J. Cave, M.D. Newton, Calculation of electronic coupling matrix elements for ground and excited state electron transfer reactions: comparison of the generalized Mulliken–Hush and block diagonalization methods, Journal of Chemical Physics 106 (1997) 9213–9226.
[95] R.J. Cave, M.D. Newton, Generalization of the Mulliken–Hush treatment for the calculation of electron transfer matrix elements, Chemical Physics Letters 249 (1996) 15–19.
[96] I.A. Balabin, D.N. Beratan, S.S. Skourtis, Persistence of structure over fluctuations in biological electron–transfer reactions, Physical Review Letters 101 (2008) 158102–158104.
[97] S.S. Skourtis, I.A. Balabin, T. Kawatsu, D.N. Beratan, Protein dynamics and electron transfer: electronic decoherence and non–condon effects, Proceedings of the National Academy of Sciences 102 (2005) 3552–3557.
[98] J. Wolfgang, S.M. Risser, S. Priyadarshy, D.N. Beratan, Secondary structure conformations and long range electronic interactions in oligopeptides, Journal of Physical Chemistry B 101 (1997) 2986–2991.
[99] T. Kawatsu, T. Kakitani, T. Yamato, Destructive interference in the electron tunneling through protein media, Journal of Physical Chemistry B 106 (2002) 11356–11366.
[100] N.E. Miller, M.C. Wander, R.J. Cave, A theoretical study of the electronic coupling element for electron transfer in water, Journal of Physical Chemistry 103 (1999) 1084–1093.
[101] E.W. Castner, D. Kennedy, R.J. Cave, Solvent as electron donor: donor/acceptor electronic coupling is a dynamical variable, Journal of Physical Chemistry 104 (2000) 2869–2885.
[102] S.S. Skourtis, G. Archontis, Q. Xie, Electron transfer through fluctuating bridges: on the validity of the superexchange mechanism and time-dependent tunneling matrix elements, Journal of Chemical Physics 115 (2001) 9444–9462.
[103] Y. Zhang, S.A. Kang, T. Mukherjee, S. Bale, B.R. Crane, T.P. Begley, S.E. Ealick, Crystal structure and mechanism of tryptophan 2, 3-dioxygenase, a heme enzyme involved in tryptophan catabolism and in quinolinate biosynthesis(,), Biochemistry 46 (2007) 145–155.
[104] N. Chauhan, S.J. Thackray, S.A. Rafice, G. Eaton, M. Lee, I. Efimov, J. Basran, P.R. Jenkins, C.G. Mowat, S.K. Chapman, E.L. Raven, Reassessment of the reaction mechanism in the heme dioxygenases, Journal of the American Chemical Society 131 (2009) 4186–4187.
[105] N.D.E. Shaw Research, Desmond Molecular Dynamics System, version 2.X, 2008, York, NY.
[106] V. Guallar, K.W. Borrelli, A binding mechanism in protein–nucleotide interactions: implication for U1A RNA binding, Proceedings of the National Academy of Sciences 102 (2005) 3954–3959.
[107] L.W. Chung, X. Li, H. Sugimoto, Y. Shiro, K. Morokuma, Density functional theory study on a missing piece in understanding of heme chemistry: the reaction mechanism for indoleamine 2, 3–dioxygenase and tryptophan 2, 3–dioxygenase, Journal of the American Chemical Society 130 (2008) 12299–12309.
[108] D. Batabyal, S.R. Yeh, Substrate–protein interaction in human tryptophan dioxygenase: the critical role of H76, Journal of the American Chemical Society 131 (2009) 3260–3270.
[109] A. Lewis-Ballester, D. Batabyal, T. Egawa, C.Y. Lu, Y. Lin, M.A. Marti, L. Capece, D.A. Estrin, S.R. Yeh, Evidence for a ferryl intermediate in a heme-based dioxygenase, Proceedings of the National Academy of Sciences 106 (2009) 17371–17376.