9 SCIENTIFIC HIGHLIGHT OF THE MONTH

First principles simulations of hemeproteins: From the active center to the full protein

Carme Rovira[1] and Michele Parrinello [2]

[1] Departament de Química Física. Facultat de Química
Universitat de Barcelona.Martí i Franquès 1, 08028 Barcelona, Spain.
[2] CSCS- Swiss Center for Scientific Computing. ETH Zürich.
Via Cantonale, Galleria 2. CH-6928 Manno (TI), Switzerland.

Abstract

We present a computational study of the interplay between the structure, energy and dynamics of the active center of hemeproteins. Our calculations are based on density functional theory (DFT) combined with molecular dynamics (MD), within the Car- Parrinello scheme. Starting from the optimized structures of models for the active center, we quantify the trends in their ligand binding properties and their intrinsic dynamics at room temperature, helping to clarify experimental results for hemeproteins and biomimetics. The influence of the rest of the protein in the active center is investigated by means of hybrid QM/MM calculations based on density functional theory combined with a classical force field. It is shown that while the heme-CO bond in carbonmonoxy myoglobin (Mb-CO) is quite robust, both the CO stretch frequency and the strength of the CO···His64 interaction are very dependent on the orientation and tautomerization state of the distal histidine residue (His64). Further aspects of these interactions and its biological implications are discussed.

1. Introduction

Hemeproteins are proteins having a heme molecule as active center (Figure 1). They perform vital roles in cellular catalysis, transport and storage [1]. For example, hemoglobin (Hb) transports oxygen in the blood (from the lungs to the muscles) and transfers the oxygen to myoglobin (Mb) which stores it in muscles. The specificity of the protein continues to fascinate chemists, as it depends on many structural and dynamic features of not only the active center but also the surrounding polypeptide framework. These features are routinely studied by crystallographic methods, which traditionally represent a protein as a static molecule with a unique structure. However, today it is widely accepted that proteins are not static but complex dynamic systems which are characterized as "screaming and kicking" [2] and undergoing "proteinquakes" [3].

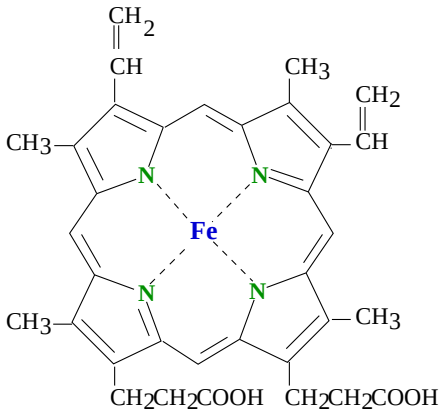


Figure 1. Molecular structure of heme b (also called protoporphyrin IX), the active center of myoglobin, hemoglobin and other hemeproteins. The substituent groups of the iron-porphyrin vary from one hemeprotein to another. In the case of myoglobin, they are two propionates, two vinyls and four methyl substituents.

Classical molecular dynamics simulations have revealed many dynamical features of these proteins, complementing the information obtained with structural techniques [4]. For instance, the most likely pathways for the ligand entry and exit from the active center of myoglobin have been identified with MD simulations [5]. However, subtle electronic/structural/spin changes (and chemical reactions) taking place at the active center cannot be described with the currently available force fields. Instead, these processes are traditionally studied by means of quantum chemistry methods using simplified models of the active center, often changed to make them as symmetric as possible. These type of computations are commonly performed at a fixed structure [6]. During the last few years, the fast development of efficient first principles methods based on Density Functional Theory (DFT) has allowed to model the reactivity of the active center beyond the "frozen structure approximation", thus capturing most of the chemistry of these systems [7]. In this respect, the first steps towards the modelling of the protein dynamics without relying on empirical parameters, have been performed [8]. In this article we will summarize our work in the modeling of myoglobin (Mb), a small globular hemeprotein that stores oxygen in muscles until it is released in the respiration process. Related works of us on hemoglobin

biomimetics, as well as other hemeproteins can be found in reference [9].

The structure of Mb is known from X-ray and neutron diffraction studies [4]. In addition, several features in the dynamics and electronic structure have been investigated by infrared, Raman, Mössbauer, and ESR spectroscopy [12]. Another important source of information is provided by experiments on synthetic models (biomimetics) such as the picket-fence-oxygen molecule and its derivatives [10], for which dynamic information in the femtosecond time scale is now available [11]. Despite all this studies, many essential aspects of this function, such as the way the protein controls the binding of ligands (O_2, CO and NO), the precise structure of the Fe- ligand bonds or the structure-spin-energy relations at the active center, are a topic of debate [1e].

Structural studies (neutron and X-ray diffraction [4]) show that the active center of myoglobin (the heme) is located on one side of the protein. The heme is attached to the polypeptide framework through a covalent bond between the nitrogen atom of one of the protein residues (His93, named as proximal histidine) and the iron atom. The opposite face of the iron-porphyrin (Figure 2) is free to bind oxygen. The oxygenated protein exhibits a bent end-on geometry of the Fe-O_2 fragment both in MbO_2 [14] and HbO_2 [13], as it is also found in isolated hemes [10]. Neutron and X-ray diffraction measurements have found that the O_2 is hydrogen-bonded with the N-H of His64 in MbO_2 and α-Hb [14] [13]. However, there is no evidence of a hydrogen bond in β-HbO_2, where free rotation of the ligand around its equilibrium position is expected [13]. Recent EPR measurements in cobalt substituted Hb have also found evidence of O_2 rotation [15]. In addition, the four-fold disorder found in the crystal structure of picket-fence-oxygen systems [10]. has been interpreted as a dynamic O_2 motion by both Mössbauer and NMR experiments [16].

The FeCO fragment is bent in all available X-ray and neutron diffraction analysis of carbonmonoxy myoglobin (MbCO). This distortion was early thought to be responsible for the well known protein discrimination against CO (the affinity ratio CO/O_2 is lower in the protein than in biomimetic systems). However, a surprising wide range of FeCO angles (7^o - 60^o) have been reported in different X-ray studies (two recent high resolution MbCO structures [4] still give discrepant Fe-C-O angles). Thus, the precise quantification of the heme-CO structure seems to require X-ray resolution higher than it is presently available. On the other hand, spectroscopic studies [17] have predicted just a small distortion (<FeCO angle $\geq 173^o$) which, according to recent theoretical investigations, has a negligible energetic cost [6g, 7e]. As a consequence, the functional role of the FeCO distortion is nowadays very questioned [18].

There is less structural information available for nitrogenmonoxy myoglobin (MbNO) [19] than for MbCO and MbO_2. Nevertheless, a few biomimetic systems with a NO ligand have been characterized. Thermodynamic measurements show that the NO ligand exhibits a unique effect upon binding to an iron-porphyrin derivative: it has a tendency to weaken the axial trans ligand bond [20]. This has been observed in MbNO and its biomimetics, as well as in other heme proteins such as Guanylate cyclase. In this case, the effect is so strong that the binding of NO to the heme induces the release of the trans axial histidine [20b].

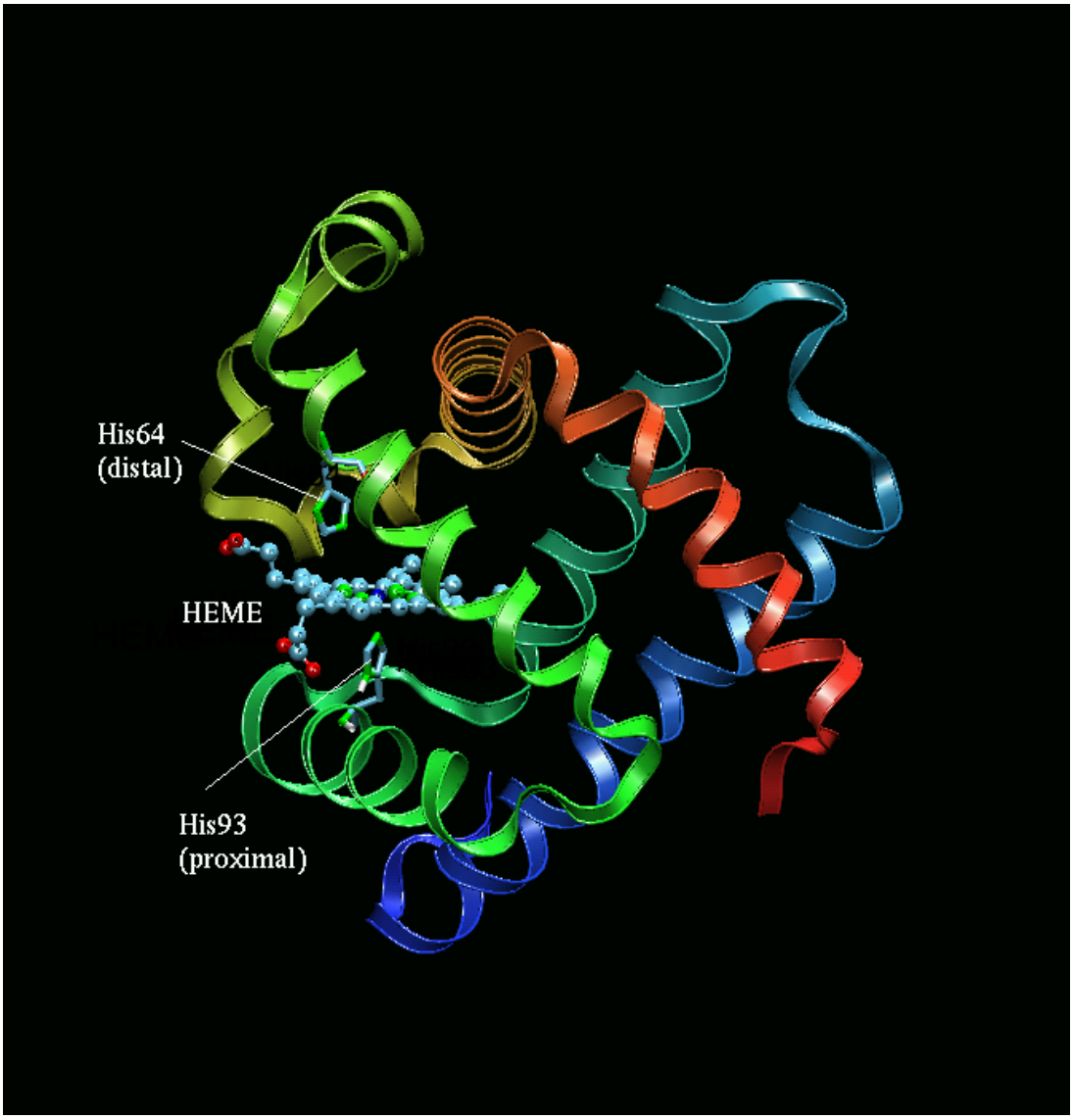


Figure 2. Structure of myoglobin (deoxy form). The heme active center is highlighted, as well as the proximal and distal histidines.

In order to understand all these issues, a precise knowledge of the intrinsic structural and dynamical properties of the heme-ligand bonds becomes of great interest. Theoretical studies could be very valuable in providing these data. It would also be extremely interesting to transcend a purely static point of view and examine fully the influence of thermal fluctuations. This is what motivated us to start a first principles study of the ligand binding properties of myoglobin. In a first step, we analyzed the interplay between the structure, energy and dynamics of the binding of O_2, CO and NO to the heme active center. The models we used for our study, shown in Figure 3, are of the type FeP-AB (FeP = iron- porphyrin, AB = CO, NO, O_2), Fe(Heme)-O_2 and FeP(Im)-AB (Im = imidazole). The imidazole molecule mimics the effect of the proximal histidine aminoacid (Figure 2). In a second step, we analyzed the interaction between the active center and the rest of the protein, by means of hybrid QM/MM simulations.

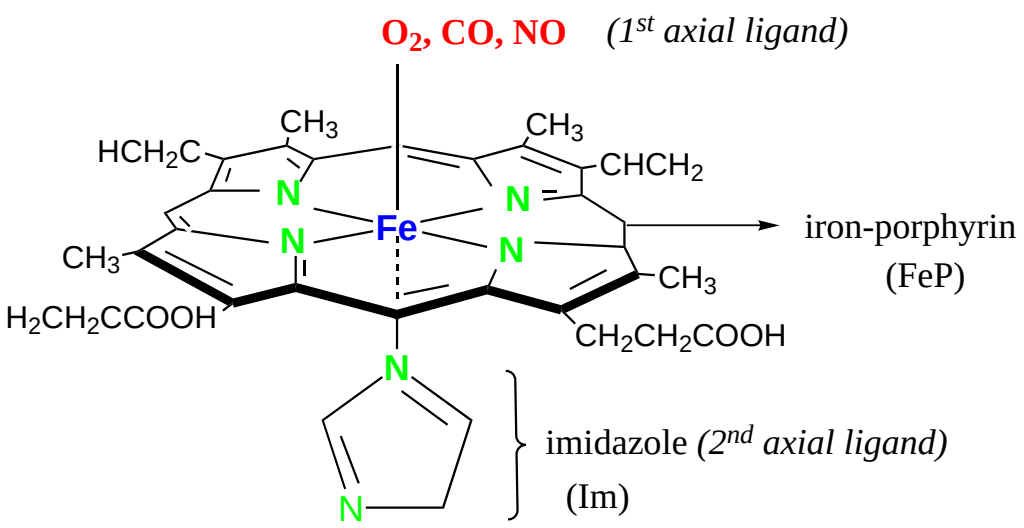


Figure 3. Active center models used in the calculations.

2. Myoglobin active center

2.1. Structure, energy and electronic state

The structures of iron-porphyrin (FeP), FeP-AB (AB = O_2, CO, NO) and FeP(Im)AB (Im = imidazole) were optimized, without symmetry constraints, for the lowest energy spin state of each system. This was found to be a triplet for FeP, singlet for FeP- CO, FeP-O_2, FeP(Im)CO and FeP(Im)-O_2, and doublet for FeP-NO and FeP(Im)-NO, which is in agreement with experiments. As shown in Figure 4 (top), all pentacoordinated FeP-AB complexes are characterized by having a curved porphyrin. This type of distortion reinforces the bonding between the Fe(d_{z2}) orbital and the $3\sigma_g$ orbital of the diatomic molecule (the d_{z2} orbital becomes hybridized by mixing some s and z character). The binding energy with respect to dissociation of the diatomic ligand amounts to 9 kcal/mol (FeP-O_2), 26 kcal/mol (FeP-CO) and 35 kcal/mol (FeP-NO). The enhanced binding of CO and NO compared to O_2 can be understood in terms of the variation of the Fe(d_{z2})- AB($3\sigma_g$) interaction. In the case of CO and NO, the $3\sigma_g$ orbital is more polarized towards the C and N atoms (the ones binding to Fe), while for O_2 it is shared among both oxygen atoms. It is also worth mentioning that the energy of these systems changes little upon rotation of the AB ligand with respect to the Fe-A bond (less than 2 kcal/mol), which indicates that a rotational motion of the ligand at room temperature is likely to occur. A side-on type of Fe-O_2 binding, where both oxygens are coordinated to the metal atom, was found to be unstable. The structure evolved towards the end-on global minimum in all attempts. Nevertheless, this type of structure has been unambiguously detected in the low temperature infrared (IR) and resonant Raman (RR) spectra on the co- condensation of irontetraphenyl-porphyrin (FeTPP) and O_2 [21]. The effect of the porphyrin tetraphenyl substituents, not present in the computed model, could be the reason of the stabilization of the side-on isomer. In fact, in our early study we found [9a] that the porphyrin tetrapivalaminophenyl substituents enhance substantially the O_2 binding to the picket-fence experimental model.

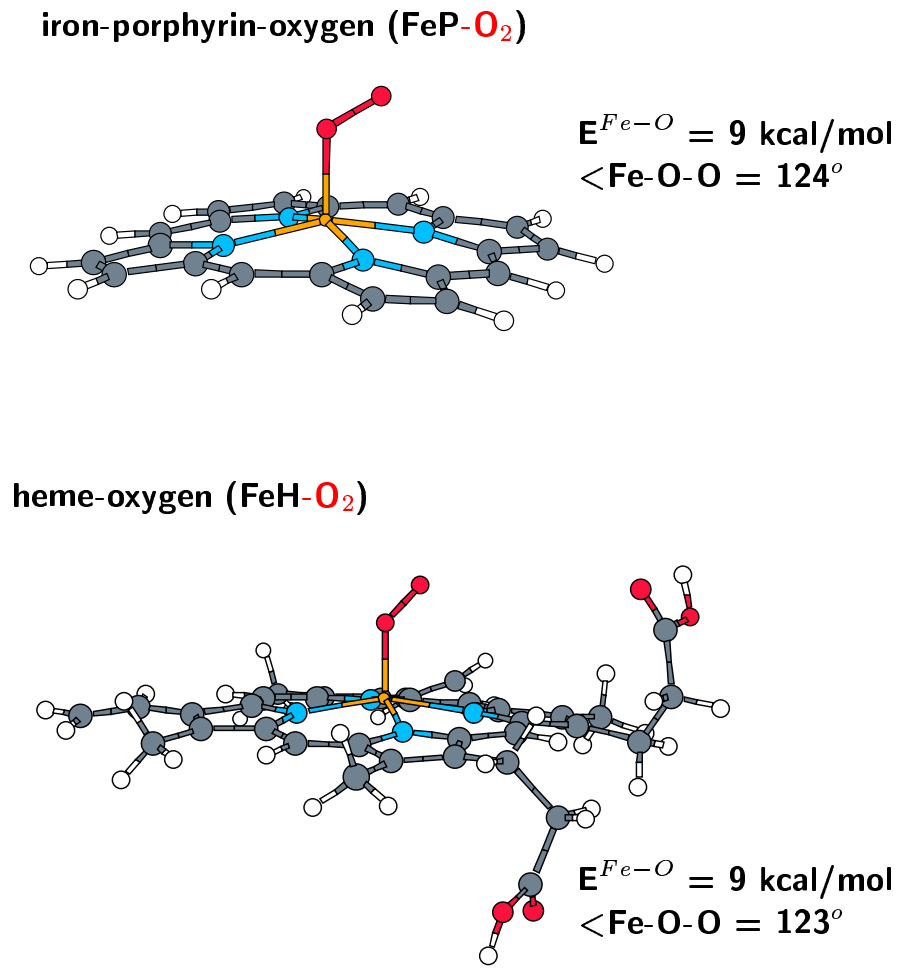


Figure 4. Top: Optimized structure of the iron-porphyrin-O_2 model. Bottom: Optimized structure of the same system considering the full heme b molecule FeH-O_2 (H = heme).

In order to relate these structure/energy changes with the properties of the protein, it is necessary to exclude any possible influence from the chemical groups which are closest to the iron-porphyrin (the porphyrin -$(CH_2)_3COOH$, -$CH{=}CH_2$, and -CH_3 substituents, as well as the proximal histidine residue). Thus, additional calculations including the porphyrin substituent groups (see Figure 4). Our calculations showed [9d] that the local structure of the FeO_2 bond in the FeH-O_2 complexes (H = heme b or protoporphyrin IX) is the same that in FeP-O_2. Most importantly, the binding energy of the ligand does not change. Similar results were obtained for the CO and NO complexes. Therefore, the porphyrin substituents present in the myoglobin active center do not change the main ligand binding properties of the iron-porphyrin.

In contrast, addition of an imidazole (Im) axial ligand (Figure 3) does introduce major changes on the binding of ligands. These changes are schematized in Figure 5. Variations in the internal porphyrin structure have not been detailed since these are minor. The only structural change due to Im is the lose of porphyrin out-of-planarity. However, the binding energy of the Fe-O_2

and Fe-CO bonds is substantially enhanced (left center in Figure 5). This can be understood in terms of the increase of the σ- donor character of the imidazole orbital that interacts with the Fe(d_{z2}) orbital. In contrast to CO and O_2, the energy of the Fe-NO bond shows practically no variation with the addition of the imidazole, even though a similar enhancement of the σ-donor character of the imidazole should be expected. However, in this case a competing effect appears, which is related to the occupation of the Fe(d_{z2}) orbital. This orbital (the HOMO one) is antibonding with respect to the Fe-NO interaction, and thus contributes to the weakening of the Fe-NO bond. The balance between both effects results in the insensitivity of the Fe-NO bond with respect to trans axial ligation.

Similar trends were found upon attaching the diatomic molecule to the FeP-Im system (right→center in Figure 5). The binding energy of the Fe-N_ϵ bond increases upon binding of CO or O_2, but changes little (it even weakens) upon NO binding. This type of changes are consistent with the well-known trans repulsive effect of the NO when binding to iron-porphyrin derivatives. Studies of reactions of heme models with imidazole, CO, O_2 and NO [20] (a) show that the addition of NO to an imidazole-bound iron- porphyrin weakens the Fe-Im bond, while the reverse is found for CO and O_2.

It is worth noting that the binding energy values obtained for O_2 are in agreement with thermodynamic measurements of oxygen binding to myoglobin and biomimetic heme models, for which Ho values in the range of 10-19 kcal/mol are reported [22]. The large imbalance between the binding energies of CO and O_2 in the gas phase underlines the fact that the protein environment plays a major role in modulating the relative binding between both ligands. Recent experiments on Fe$T_{pyr}PH_2$-NO have reported values a binding enthalpy of 29 kcal/mol [23], a value very similar to the one we computed for FeP-NO.

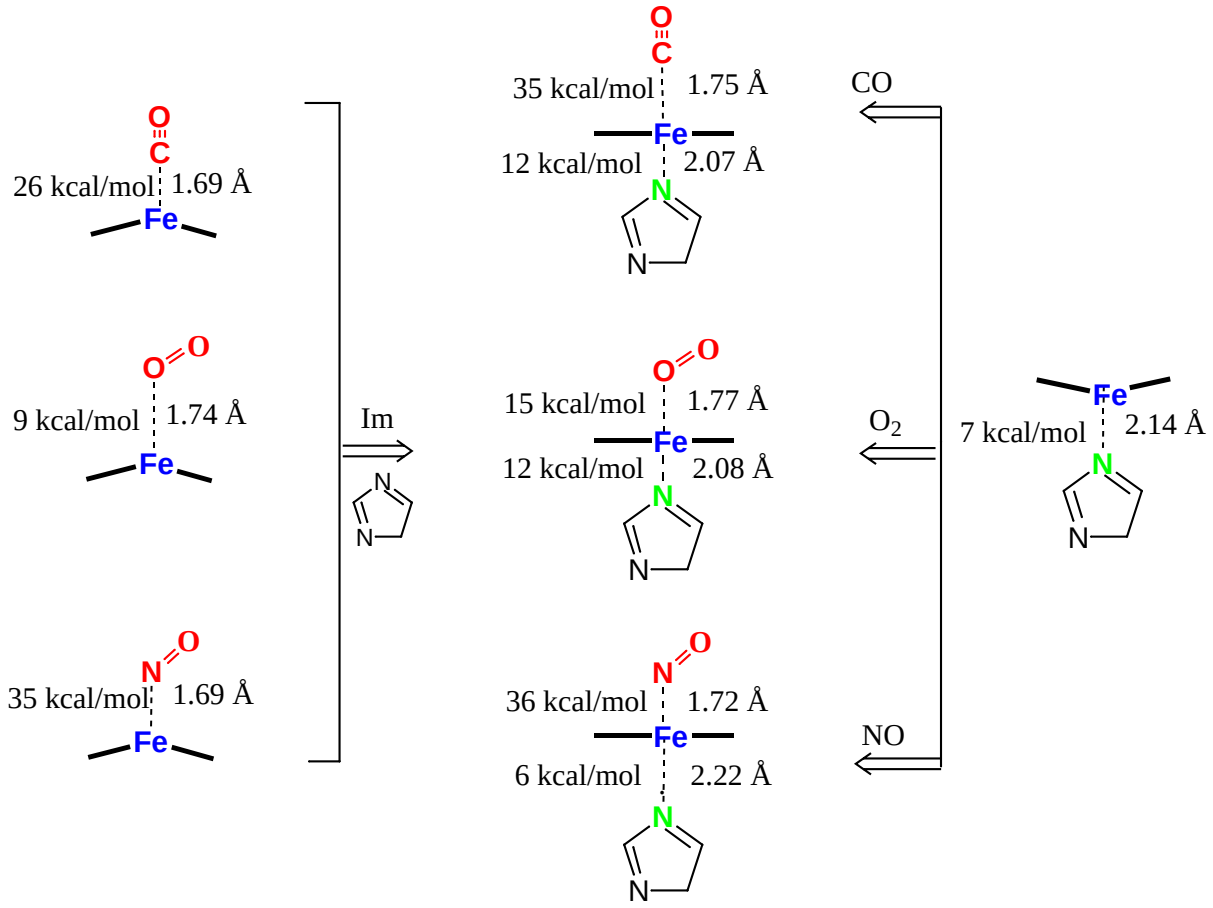


Figure 5. Scheme of the structure/energy changes upon addition of an imidazole axial ligand to the FeP(AB) systems (left to right) or addition of a diatomic AB molecule to the FeP(Im) system (right to left), with AB = CO, O_2, NO.

2.2. Heme-ligand dynamics

The dynamics of the CO ligand at room temperature was investigated in the FeP(Im)- CO model. Our simulations show that the CO ligand undergoes a fast motion around its equilibrium position: the projection of the oxygen atom on the porphyrin plane samples all porphyrin quadrants in less than 0.5 ps. An interesting property that can be extracted from the simulation is the allowed distortion of the Fe-C-O fragment. We quantified this distortion by using the tilt (δ) and bend (θ) angles, which have been often related to the protein discrimination against CO. Figure 6 shows the probability distribution of the δ and θ angles obtained from the dynamics. Small fluctuations ($\delta \leq 8^o$, $\theta \leq 13^o$) have a sizable probability to take place, but larger deformations do not occur. Therefore, for a FeCO not perturbed by the environment, small δ - θ deviations (similar to those reported by spectroscopic studies [17]) can occur just due to the thermal motion and not as a consequence of steric hindrance by the protein. This is consistent with the conclusions of recent DFT calculations [6g, 7e] which have predicted that small δ - θ variations do not have a significant energetic cost. It should also be noted that, given the complex motion of the ligand, the instantaneous structure of the FeCO unit cannot be easily

defined just in terms of the δ and θ angles; the problem should be best regarded as that of a highly dynamic FeCO moiety, sampling many different conformations with different probability in a short time.

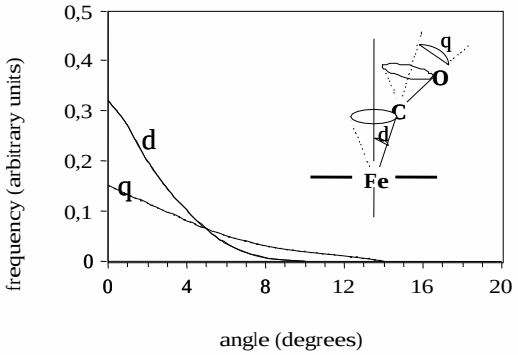
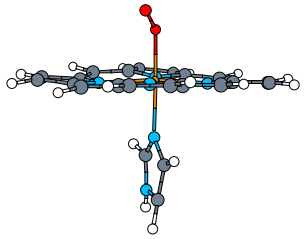
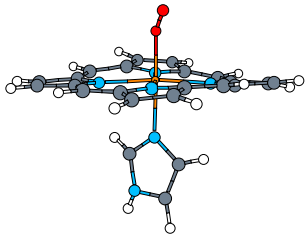


Figure 6. Frequency distribution corresponding to the two angles commonly used to describe the distortion of the FeCO fragment (θ = bending angle, δ = tilting angle).

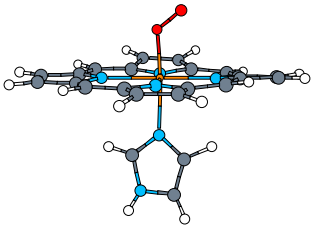
A MD simulation for the oxyheme model, FeP(Im)(O_2), was performed for a total time of 15.5 ps. The dynamics of the bent FeO_2 unit was found to be more complex than in the case of the CO system. Selected snapshots of this simulation are shown in Figure 7. During the first period of the simulation the O-O axis projection on the porphyrin plane lies on one of the porphyrin quadrants, although it undergoes large oscillations between the two closest Fe-N_p bonds (N_p = porphyrin nitrogen atom). However, after 2.2 ps the O_2 jumps over one Fe-N_p bond towards the next porphyrin quadrant. We observed five of these jumps during the whole simulation, with an average time interval of 4 - 6 ps. All transitions take place via rotation of O_2 around the Fe-O axis and involve a conformation with a more open Fe-O-O angle (124^o - 129^o) and the Fe-O bond slightly tilted (3-5^o) with respect to the heme perpendicular (i.e., the perpendicular direction with respect to the average plane defined by the four N_p atoms). This confirms the fact that the OO/Fe-N overlaping configuration is the transition state for the dynamic motion of O_2 between the porphyrin quadrants, something not obvious from experiments [24].



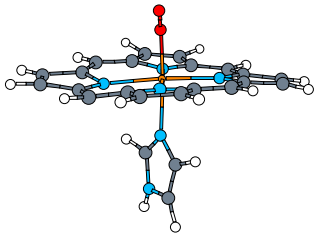
2.9 ps



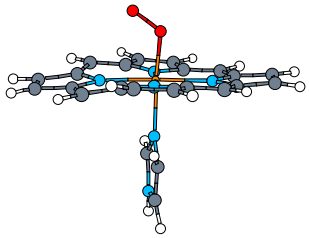
3.0 ps



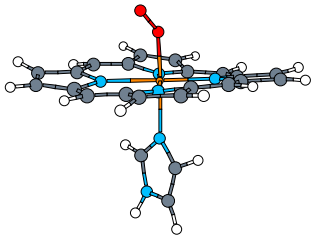
3.1 ps



7.7 ps



7.8 ps



10.4 ps

Figure 7. Snapshots of the dynamics of the $FeP(Im)O_2$ active center model.

The average structure of the FeO_2 fragment that we obtain from the simulation is very similar to the equilibrium structure of the FeP(Im)-O_2 system: Fe-O = 1.75 Å, <FeOO = 122^o, O-O = 1.30 Å. The average FeOO angle is though slightly more open (124^o) and the Fe-O distance is larger (1.86 Å) due to the anharmonicity of the corresponding vibrational modes. This structure is however very different from the experimental values of its closely related picket-fence myoglobin biomimetic: Fe-O = 1.75 Å, <FeOO = 129^o, 133^o; O-O = 1.15 Å, 1.17 Å. This could be due to the fact that X-ray do not measure distances and angles directly, but rather the positions of maximum probability, from which the other structural properties are deduced. We can prove this point from our MD trajectory. If we take the position of maximum probability and then extract distances and angles we find: Fe-O = 1.72 Å, <FeOO = 139^o, O-O = 1.19 Å. Although this is in better agreement with experiment, these values are irrealistic (the O-O distance, for instance, is shorter than the gas phase value of 1.23 Å). Therefore, although this reconciles theory and experiment, it also illustrates the risk of assigning a static structure to a highly dynamical moiety such as the FeO_2 unit. Similar considerations hold for the reported data on the proteins. In particular, the Fe-O-O angles reported by neutron and X-ray structures of oxymyoglobin [14] and oxyhemoglobin [13] are very discrepant (115^o for MbO_2, 153^o for α-Hb and 159^o for β-Hb) and not even sampled in our simulation. On the other hand, a recent X-ray structure [4g], reports a <FeOO angle very close to our computed value (122^o). It could be argued that the <FeOO angle in the protein is affected by H-bond to the His64 residue. However, it has been very recently shown [7e] that such a hydrogen bond does not modify significantly the structure of the FeO_2 unit.

Overall, our simulation reveals a highly anharmonic dynamics for the O_2 ligand: it undergoes large amplitude oscillations within one porphyrin quadrant and jumps from one to the other within 4-6 ps. This is consistent with the highly dynamic nature of O_2 bound to heme proposed by several experiments in proteins and biomimetics [10] [24] especially those which lack a hydrogen bond at the terminal oxygen. Ligand rotation in these models has been shown to occur by NMR experiments [24b,c] on the basis of the equivalence of the pyrrole proton resonances. Our results suggest that, for non-hydrogen bonded O_2, precise determination of the rate of rotation would require picosecond time resolution.

3. Interaction of the heme with the protein

The above results show that most of the properties of the Heme-Ligand bonds (Ligand = CO, O_2, NO), namely its spin state, structure and dynamics, are well reproduced by modeling only the active center. This leads to an obvious question, which is what is the role of the rest of the protein on the binding of ligands. The surrounding protein has in fact a major role, for instance, in controlling the entry and release of the ligands. Protein fluctuations open channels in the interior of the protein that allow the ligands to entry and reach the active center. Recent classical MD simulations have been able to identify some of these channels [5]. Another role, whose origin has not yet been elucidated, is the control of the binding affinity of different ligands. In fact, the values we compute for the binding energies have the right trends (NO > CO >> O_2) but not the right absolute values. The bond of the CO is far too strong to justify the

experimental CO/O_2 ratio in the equilibrium constants for the ligand binding reaction. This is however not surprising. It is known since the 70's that the relative binding CO/O_2 is controlled by the polypeptide framework [25]. A sensitive probe of this influence are the different CO stretch frequencies that appear in the IR spectrum of Mb-CO [26]. These peaks evidence that there are protein substates that interact differently with the ligand. Nevertheless, the relation between each CO stretch frequency and specific protein conformations has not yet been clarified [27].

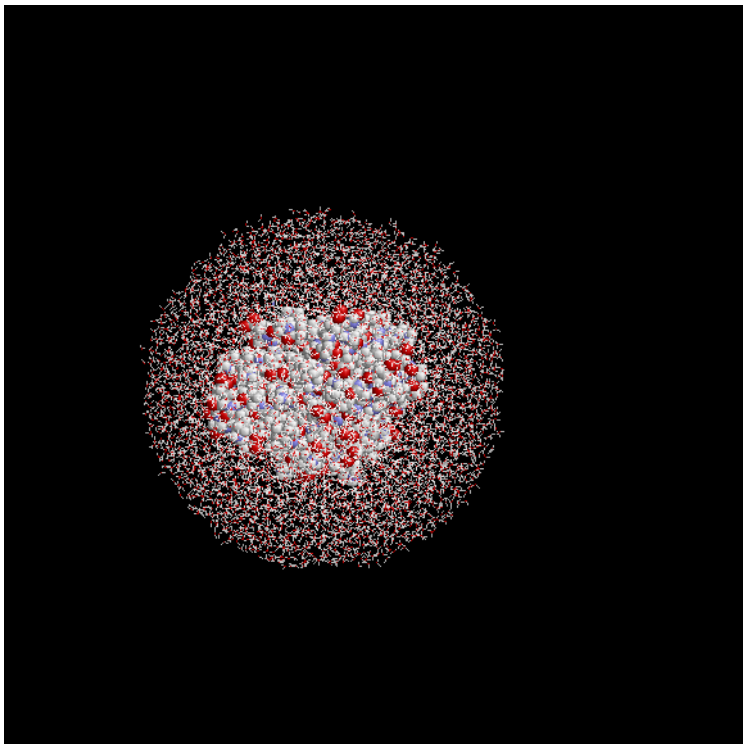
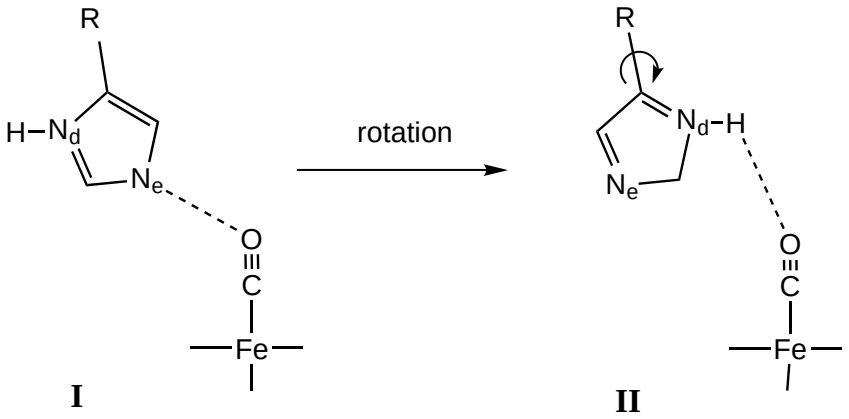


Figure 8. Protein embeeded in the water shell.

In an attempt to help the interpretation of the IR spectrum of Mb-CO, we decided to model the full protein at the QM/MM level and to evaluate the changes in the CO stretching frequency for different protein conformations. The QM/MM method we used combines the same first principles description of the active center that we have used up to now with a force-field treatment (using the CHARMM force field) of the rest of the protein. The QM- MM boundary is modeled with the use of link atoms (we refer to reference [28] for details on the QM/MM implementation). As QM region (the part that we will treat with DFT) we chose the CO ligand, the porphyrin, and the axial imidazole (see Figure 2). The porphyrin substituents were not included, since we previously found that they do not affect the properties of the Fe-ligand bonds. On the other hand including the Im of the proximal His (directly bonded to the heme) is crucial since it has an effect on strengthening the Fe-CO bond (section 2.1). With this QM-MM partition we can be confident that the energy/spin/structure relations of the heme will be well described. The protein was additionally enveloped in a 37 Å sphere of equilibrated TIP3P water molecules (see Figure 8) so as to take into account solvation effects. The number of QM and MM atoms treated in the calculation are 63 and 20000, respectively.

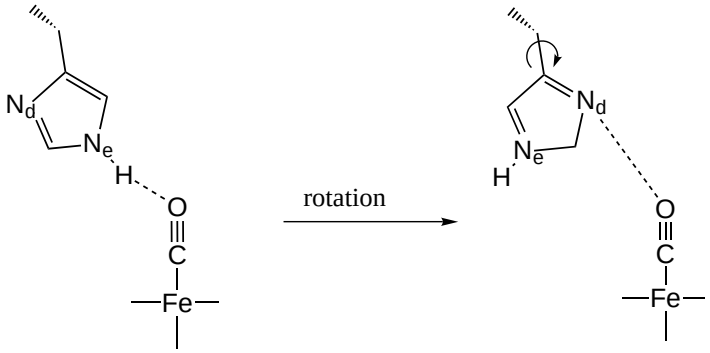
Before starting the calculations, it is crucial to decide what is the initial structure of the protein that we take for the calculation. Although one could just take the crystal structure directly, this is not very convenient since X-ray structures are an average among many different instantaneous protein conformations. Instead, it is physically more meaningful to consider snapshots of previous MD simulations done with the same force field used in the QM/MM calculations.

Classical simulations of MbCO using the CHARMM force field [29] were done considering different tautomerization states of the distal histidine residue (His64). These simulations evidenced that when His64 is protonated at N_δ (named N_δ-tautomer) it often rotates such that it can expose either the N_δ-H bond or the unprotonated N_ϵ atom towards the CO.



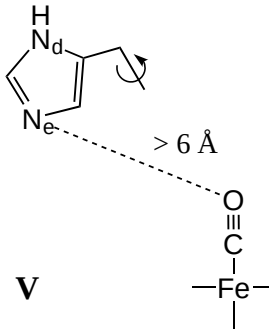
I **II**

Thus, we took snapshots corresponding to these two distal His conformations. We also took another snapshot (**III**) from a simulation started with the His64 proton placed on the N_ϵ atom (N_ϵ-tautomer). As rotation of His64 did not occur in the time scale of the classical simulation (1 ns) we forced it by inducing a 180 rotation around the C1-C2 bond of His64 (**IV**).



III **IV**

Finally, we took a fifth snapshot in which the distal His moved away from the CO (this occurred after 600 ps of the simulation).



V

These protein configurations, denoted **I**-**V**, are representative of the dynamics of the heme pocket. The results obtained in the structural relaxation for each of these protein conformations are summarized in the next table.

Table 1. Main parameters defining the optimized heme-CO structure of each protein conformation **I-V**. The last row corresponds to the results obtained for the heme-CO isolated model (section 2.1).

structure	interaction type	O···X	C-O	Fe-C	<FeCO	Fe-N_p
exp.	$CO\cdots N_\epsilon$	4.0-2.60	1.09-1.21	1.73-2.21	120 - 172	2.01-2.06
I	$CO\cdots N_\epsilon$	3.39	1.16	1.76	177.3	2.00-2.02
II	$CO\cdots N_\delta$	3.47	1.16	1.75	179.3	1.98-2.03
III	$CO\cdots H-N_\epsilon$	2.69	1.17	1.74	176.1	1.99-2.02
IV	$CO\cdots N_\delta$	3.90	1.16	1.74	175.7	1.99-2.02
	$CO\cdots H-C$	2.18				
V	$CO\cdots N_\epsilon$	6.03	1.16	1.75	177.6	1.99-2.03
	$CO\cdots H-C$	4.03				
heme-CO	–	-	1.17	1.72	180.0	2.02

As expected, the heme-ligand structure is not very sensitive to the protein conformation. All of them have similar Fe-ligand bonds (also similar from the gas phase values we obtained previously) and, most importantly: the Fe-CO angle is essentially linear in all cases, even when the proton of the distal His is close to the CO. This suggests that the FeCO fragment should be regarded as a linear bond. The fact that X-ray structures give a distorted FeCO bond is likely to be related to the limited resolution at region of the active center.

Concerning to the CO stretch frequency, we indeed obtain significant changes depending upon the protein conformation, as we can see in the following table.

Table 2. Shift of the C-O and Fe-C stretch frequencies with respect to the isolated heme-CO system for each of the protein conformations **I - V**. Hydrogen bond energies are also listed. Distances are given in Å, frequencies in cm^{-1} and energies in kcal/mol.

structure	interaction type	O···X	$\Delta\nu_{CO}$	$\Delta\nu_{FeC}$	$\Delta E_{O\ldots X}$
I	$CO\cdots N_\epsilon$	3.39	+14	-62	+2
II	$CO\cdots N_\delta$	3.47	-14	18	-2.5
III	$CO\cdots H-N_\epsilon$	2.69	-23	61	-3.4
IV	$CO\cdots N_\delta$	3.90	-4	10	-0.9
	$CO\cdots H-C$	2.18			
V	$CO\cdots N_\epsilon$	6.03	-1	10	-0.1
	$CO\cdots H-C$	4.03			
heme-CO	–	–	0	0	–

Whenever the proton of the distal His is close to the CO (**II**, **III**), the CO frequency decreases

(the shift is negative). Instead, the shift is positive when the unprotonated nitrogen comes close to CO (**I**). The largest downshift (-23 cm-1) is given by the CO···H-N interaction in the N_ϵ-tautomer (**III**). In this case, the possible N-H···OC hydrogen bond is geometrically more favored than in the other configurations. The configuration with the distal His far from the ligand (**V**) practically does not shift the CO frequency . This is in agreement with mutagenesis experiments showing an enhancement of the highest CO frequency peak in the IR spectrum when His64 is replaced for an apolar residue [30]. Configuration **IV** practically does not shift significantly the CO frequency either. This could be due to the opposite effect of both the protonated and the unprotonated N at intermediate distances. Our calculations also evidence the inverse correlation between the $\Delta\nu_{CO}$ and $\Delta\nu_{FeC}$ values ($\Delta\nu_{CO}$ increases as $\Delta\nu_{FeC}$ decreases). This general trend has been observed across a wide range of heme proteins and biomimetic systems [31].

The frequency changes of Table 2 can be rationalized in terms of variations in the Fe- CO back bonding (i.e. the interaction of the d-Fe levels with the empty σ^* CO orbitals). When a positive charge approaches the CO (such as the proton of His64), the orbitals are energetically stabilized. As they get closer in energy to the Fe-d orbitals, the back bonding increases and, as a consequence, the CO frequency decreases. Thus, a downshifted ν_{CO} is observed for the CO···H-N interactions (**II**, **III**). In contrast, a negative charge approaching the CO (such as a nitrogen lone pair) would decrease the back bonding and increase νCO. In agreement with this argument, an upshift of ν_{CO} is obtained for the arrangement I.

Similar changes in CO frequency have been observed experimentally. As mentioned above, the IR spectrum of carbonmonoxy myoglobin shows three main CO stretching bands. The origin of all these bands is still unclear but is often assumed that they correspond to different orientations and/or protonation state of the distal His [31]. The shifts in the CO frequency that we obtain give further support to this interpretation. Further details about this study including the interpretation of the IR spectrum in terms of our calculated shifts can be found in ref. [32].

As in the case of vibrational frequencies, the interaction energy of the ligand with the distal residue (last column in Table 2) is very dependent on the conformation and protonation state of the distal His. Configuration **I** lead to a repulsive interaction (2 kcal/mol), while the interaction is favorable when the protonated nitrogen is close to the CO. We find the largest stabilization for the N_ϵ tautomer, as in this case the H-bond is more favored. This is at variance from the common assumption [31a] that only O_2 could be hydrogen bonded to His64. However, our calculations support recent RR measurements [33] that show spectroscopic evidence of a hydrogen bond between CO and His64.

For the sake of comparison, we did a calculation replacing the CO by O_2 (just to estimate the strength of the analogous O_2···His64 interaction). It is commonly accepted that His64 is protonated at N_ϵ in MbO_2. Hence, we did the calculation on arrangement **III** where the proton of N_ϵ is pointing towards the ligand (Figure 9). The H-bond in that case amounts to about -5.1 kcal/mol. Therefore our calculations find the O_2 more stabilized by H-bond than the CO, although H-bond takes place in both cases.

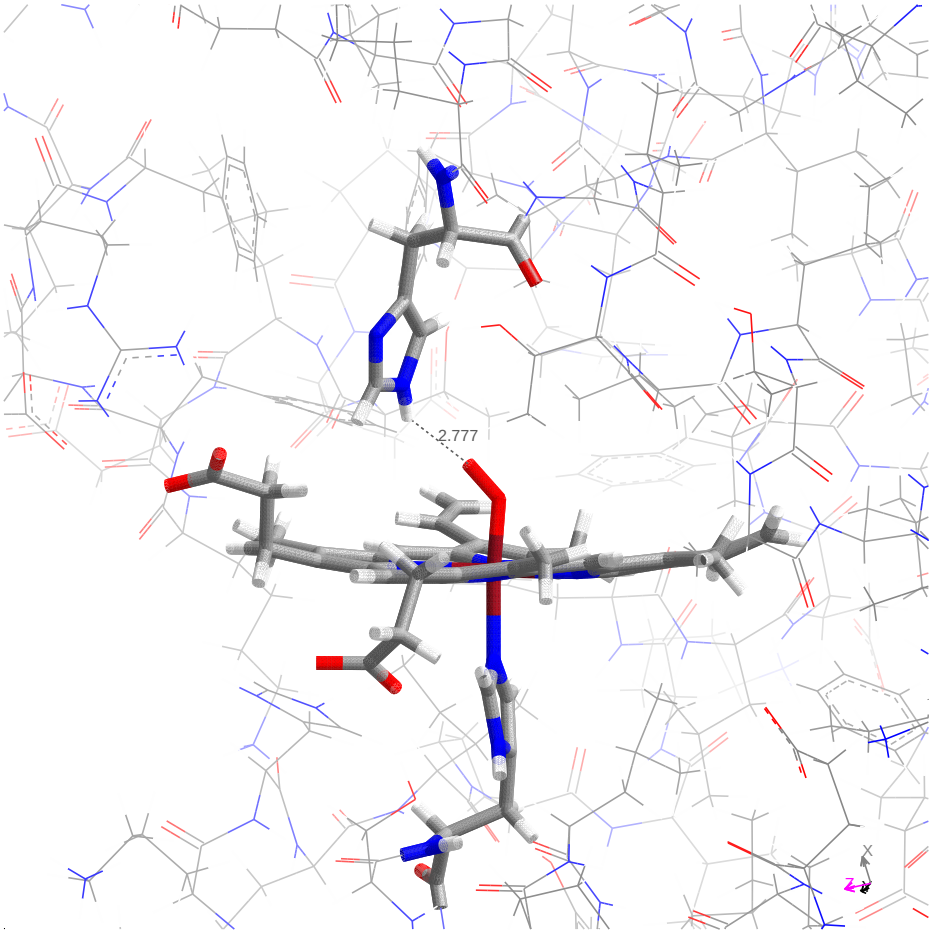


Figure 9. Optimized structure of oxymyoglobin (MbO_2) in the region around the heme active center.

4. Conclusions

In this study we have quantified the binding of CO, NO and O_2 to myoglobin. All three ligands induce a significant curvature of the heme active center when binding to the iron, although the porphyrin planarity is restored by the trans binding of an imidazole ligand. Unlike the imidazole axial ligand, the heme substituents do not influence the structure and the binding energy of the Fe-Ligand bonds (Ligand = CO, O_2, NO).

Significant differences in the binding properties of the three ligands are observed. *The Fe-O_2 bond is much weaker than Fe-CO and Fe-NO, and the binding angle increases on going from O_2 to CO.* Most of these changes can be traced back to differences in the electronic structure. In the case of the six-coordinated complexes, occupation of the Fe(d_{z2}) orbital in the NO complex leads to the elongation and weakening of the bond with the trans axial ligand, while for O_2 and CO the imidazole reinforces the iron-diatomic bond. Rotation of the oxygen around the Fe-O bond involves a small energy barrier (< 2 kcal/mol), which indicates that several rotational conformations could be available at room temperature. Indeed, our MD simulations show *the O_2 ligand undergoing large amplitude oscillations within one porphyrin quadrant and jumping to another quadrant in the picosecond time scale.* In contrast, just a fast motion of the ligand around its equilibrium position characterizes the dynamics of the FeCO unit, with a maximum distortion 13^o in the <FeCO angle.

Hybrid QM/MM calculations based on density functional theory combined with the CHARMM force field highlight the effect of the distal pocket conformation on the properties of the Fe-CO bond in MbCO and help the interpretation of the CO absorption bands in the IR spectrum. Our calculations, performed on selected snapshots from a classical MD trajectory, show that the local structure around the Fe atom practically recovers that of the gas-phase heme-CO system, where the FeCO is linear. Therefore, *the heme-CO structure seems to be quite robust and not influenced by the protein environment.* Instead, both the CO stretch frequency and the strength of the CO···His64 interaction appear to be very dependent on the orientation and tautomerization state of His64. This can be rationalized in terms of the changes in the Fe-CO back bonding when a positive/negative charged group approaches the CO ligand. Our calculations show that, *in contrast with the common assumption that only the binding of O_2 is stabilized by interaction with the distal histidine residue, a significantly stabilization also occurs for the CO ligand.* Nevertheless, the CO···His64 interaction is smaller than O_2···His64 one.

In summary, our calculations have quantified the interplay between the structure, energy and dynamics of the heme active center and its interaction with the protein. This helps to understand some of the unknown questions in hemeprotein research such as the precise structure of the Fe-ligand bonds, their intrinsic dynamics, the role of the proximal and distal histidines and the origin of the CO stretch bands of the IR spectrum of MbCO. Of course, many relevant biological processes, like ligand migration into the solvent and the folding of the proteins, occur in long time scales and thus cannot be treated currently by ab initio molecular dynamics (AIMD). However, we believe that the present study illustrates how AIMD simulations can provide useful hints in order to understand the mechanism of short time scale processes in proteins.

Acknowledgments

We thank the Garching Computer Center (Garching, Germany) for computing support. C. Rovira acknowledges the financial support of the "Incorporacin de doctores y tecnólogos" program of the Spanish Ministry of Science and Technology. C.R thanks useful discussions with many friends and colleagues during the years "in der Klinik": Jürg Hutter, Pietro Ballone, Karel Kunc, Mauro Boero, Roger Rousseau, Carla Molteni and Irmgard Frank.

Appendix: computational details

All calculations presented here are based on the density functional theory within the LDA and LSD approximations. The Kohn-Sham orbitals are expanded in a plane wave (PW) basis set, with a kinetic energy cutoff of 70 Ry. The Ceperley-Alder expression for correlation and gradient corrections of the Becke-Perdew type are used [34]. We employ ab initio pseudopotentials, generated using the Troullier-Martins scheme [35]. The following core-radii, in a.u., were used: 1.23 for the s, p atomic orbitals of carbon, 1.12 for s, p of N, 0.5 for the s of H, and 1.9, 2.0, 1.5, 1.97, respectively, for the s, p, d, f atomic orbitals of Fe. The non-linear core-correction [36] was used (with core- charge radius of 1.2 a.u.). We employed the Car-Parrinello molecular dynamics method [37] for optimization of the atomic structures. A successive use of quenching and annealing performed for about one picosecond was necessary to reach a final convergence of 10-5 and 5.10-4 a.u. for electronic and ionic gradients, respectively. Structure optimizations were performed with no constraints starting from non-symmetric structures. The convergence of our results with the energy cutoff in the PW expansion was investigated for an iron-porphyrin. The ordering of spin states was found to be insensitive to the PW cutoff value, and the energy differences changed only very slightly (within 0.5 kcal/mol) [9]e. Structural parameters were found to be even less sensitive than energy differences to the PW cutoff. Molecular dynamics simulations at room temperature were performed using a time step of 0.12 fs, with the fictitious mass of the Car-Parrinello Lagrangian set to 700 au. The deuterium mass for the hydrogen atoms was used. The systems, enclosed in supercells of 16 Å x 16 Å x 20 Å periodically repeated in space, were allowed to evolve during 2 ps in order to achieve vibrational equilibration. The MD was performed for a total period of 18 ps and 15.5 ps for FeP(Im)-CO and FeP(Im)-O_2, respectively. The hybrid QM/MM calculations have been done using the EGO-CPMD code [28]., which is an interface between the EGO classical code, based on the CHARMM force field, with the CPMD code of Stuttgart. Harmonic ligand stretch frequencies were computed from the diagonalization of the Hessian matrix obtained by numerical energy derivatives.

References

[1] See for instance: (a) M. F. Perutz, G. Fermi, B. Luisi, B. Shaanan, R. C. Liddington, R. C.; *Acc. Chem. Res.* **20**, 309, 1987. (b) B. A. Springer, S. G. Sligar, J. S. Olson, G. N.

Phillips,*Chem. Rev.* **94**, 669, 1994. (c) M. Momenteau, C. A. Reed, *Chem. Rev.***94**, 659, 1994. (d) J. S. Olson, G. N. Phillips, *J. Biol. Chem.***271**, 17593, 1996. (e) J. T. Sage, P. M. Champion, in "Small substrate recognition in heme proteins", *Comprehensive Supramolecular Chemistry*, vol 5, 171-213, 1996. (f) S. Hirota, L. Tiansheng, G. N. Phillips, J. S. Olson, M. Mukai, T. Kitagawa, *J. Am. Chem. Soc.*, **118**, 7845, 1996.

[2] (a) G. Weber. *Adv. Prot. Sci.*, **29**, 1, 1975. (b) H. Frauenfelder, S. G. Sligar, P. G. Wolynes. Science. **254**, 1598, 1991.

[3] A. Ansari, J. Berendsen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, R. D. Young. *Proc. Natl. Acad. Sci. USA*, **82**, 5000, 1985.

[4] (a) I. Schlichting, J. Berendzen, G. N. Phillips, R. M. Sweet, *Nature* **371**, 808, 1994. (b) F. Yang, G. N. Phillips *J. Mol. Biol.* **256**, 762, 1996. (c) J. Kuriyan, S. Wilz, M. Karplus, G. Petsko, *J. Mol. Biol.* **192**, 133, 1986. (d) X. Cheng, B. P. Schoenborn, *J. Mol. Biol.* **220**, 381, 1991. (e) F. Yang, G. N. Phillips, *J. Mol. Biol.* **256**, 762, 1996. (f) A. N. Kachalova, H. Popov, D. Bartunik, *Science* **284**, 473, 1999. (g) J. Vojtechovsky, K. Chu, J. Berendzen, R. M. Sweet, I. Schlichting, *Biophys. J.* **77**, 2153, 1999.

[5] B. Schulze, H. Grubmller, J. D. Evanseck. *J. Am. Chem. Soc.*, **122**, 8700, 2000.

[6] See for instance: (a) A. Dedieu, M.-M. Rohmer, A. Veillard, *Chem. Phys.* **97**, 10964, 1983. (b) A. Waleh, N. Ho, L. Chantranupong, G. H. Loew, *J. Am. Chem. Soc.* **111**, 2767, 1989. (c) W. D. Edwards, B. Weiner, M. C. Zerner, *J. Am. Chem. Soc.*, **108**, 2196, 1986. (d) P. Jewsbury, S. Yamamoto, T. Minato, M. Saito, T. Kitagawa, *J. Am. Chem. Soc.*, **116**, 11586, 1994. (e) H. Nakatsuji, J. Hasegawa, H. Ueda, M. Hada, *Chem. Phys. Lett.* **250**, 379, 1996 (f) R. H. Havlin, N. Godbout, R. Salzmann, M. Wojdelski, W. Arnold, C. E. Schulz, and E. Oldfield. *J. Am. Chem. Soc.* **120**, 3144, 1998. (g) Vangberg, T.; Bocian, D. F.; Ghosh, A. *J. Biol. Inorg. Chem.* **2**, 526, 1997.

[7] (a) T. G. Spiro, P. M. Kozlowsli, *J. Am. Chem. Soc.***120**, 4524, 1998. (b) F. Maseras, *N. J. Chem.***22**, 327, 1998. (c) C. Rovira, P. Ballone, M. Parrinello. *Chem. Phys. Lett.* **271**, 140, 1997. (d) C. Rovira, K. Kunc, J. Hutter, P. Ballone, M. Parrinello, *J. Phys. Chem. A* **101**, 8914, 1997. (e) E. Sigfridson, U. Ryde, *J. Biol. Inorg. Chem.* **4**, 99, 1999. (f) M. D. Segall, M. C. Payne, S. W. Ellis, G. T. Tucker, R. N. Boyes, *Phys. Rev. E* , 4618, 1998. (g) D. A. Scherlis, C. B. Cymeryng, D. A. Estrin, *Inorg. Chem.* **39**, 2352 (2000).

[8] (a) F.Alber, O.Kuonen, L.Scapozza, G.Folkers, P. Carloni, *Protens: Struc. Func. Gen.* **31**, 453, 1998. (b) D. E. Sagnella, K Laasonen, M. L. Klein, *Biophys. J.* **71**, 1172, 1996. (c) C. Rovira, M. Parrinello. *Biophys. J.* **78**, 93, 2000.

[9] (a) C. Rovira, M. Parrinello. *Chem. Eur. J.*, **5**, 250, 1999. (b) C. Rovira, P. Carloni, M. Parrinello. *J. Phys. Chem. B***103**, 7031, 1999. (c) M. Kaupp, C. Rovira, M. Parrinello. *J. Phys. Chem. B***104**, 5200, 2000. (d) C. Rovira, M. Parrinello. *Int. J. Quant. Chem.* **80**, 1172, 2000. (e) C. Rovira, K. Kunc, J. Hutter, M. Parrinello. *Inorg. Chem.* **40**, 11, 2001.

[10] (a) G. B. Jameson, G. A. Rodley, W. T. Robinson, R. R. Gagne, C. A. Reed, J. P. Collman, *Inorg. Chem.* **17**, 850, 1978. (b) B. Jameson, F. Molinaro, J. A. Ibers, J. P. Collman, J. I. Brauman, E. Rose, K. S. Suslick, *J. Am. Chem. Soc.* **102**, 3224, 1980.

[11] B. Steiger, J. S. Baskin, F. C. Anson, A. H. Zewail, *Angew. Chem. Int. Ed.* **39**, 257, 2000.

[12] A. X. Trautwein, E. Bill, E. L. Bominaar, H. Winkler, in *Bioinorganic Chemistry*, Vol. 78 (Springer-Verlag, Berlin, 1991), p. 1.

[13] B. Shaanan, *J. Mol. Biol.* **171**, 31, 1983.

[14] (a) S. E. V. Phillips, and B. P. Schoenborn, *Nature* **292**, 81, 1981. (b) S. E. V. Phillips, *J. Mol. Biol.* **142**, 531, 1980.

[15] J. H. Bowen, N. V. Shokhirev, A. M. Raitsimring, D. H. Buttlaire, F. A. Walker *J. Am. Chem. Soc.* **101**, 8683, 1997.

[16] (b) J. Mispelter, M. Momenteau, D. Lavalette, J.-M. Lhoste, *J. Am. Chem. Soc.* **105**, 5165, 1983. (c) E. Oldfield, H. C. Lee, C. Coretsopoulos, F. Adebodum, K. D. Park, S. Yang, J. Chung, B. Phillips, *J. Am. Chem. Soc.* **113**, 8680, 1991.

[17] (a) J. T. Sage, W. Jee, *J. Mol. Biol.* **274**, 21, 1997. (b) M. Lim, T. A. Jackson, P. A. Anfinrud, *Science* **269** 962, 1995.

[18] (a) C. Slebodnick, J. A. Ibers, *J. Biol. Inorg. Chem.* **2**, 521, 1997.

[19] (a) W. R. Scheidt, M. E. Frisse, *J. Am. Chem. Soc.* **97**, 17, 1975. (b) W. R. Scheidt, P. Piciulo, *J. Am. Chem. Soc.* **98**, 1913, 1976. (c) E. Allen Brucker, J. S. Olson, M. Ikeda-Saito, G. N. Phillips, *Prot. Str. Func. Gen.* **30**, 352, 1998.

[20] (a) T. G. Traylor, V. S. Sharma, *Biochemistry* **31**, 2847, 1992. (b) E. J. Rose, B. M. Hoffman, *J. Am. Chem. Soc.* **105**, 2866, 1983.

[21] (a) T. Watanabe, T. Ama, K. Nakamoto, *J. Phys. Chem.* **88**, 440, 1994. (b) L. M. Proniewicz, I. R. Paeng, K. Nakamoto, *J. Am. Chem. Soc.* **113**, 3294, 1991. (c) K. Nakamoto, *Coord. Chem. Rev.* **100**, 363, 1990.

[22] M. Momenteau, C. A. Reed, *Chem. Rev.* **94**, 659, 1994.

[23] O. Chem, S. Broh, A. Liechty, D. P. Ridge, *J. Am. Chem. Soc.* **121**, 11910, 1999.

[24] (a) K. Spartalian, G. Lang, J. P. Collman, R. R. Gagne, C. A. Reed, *J. Chem. Phys.* **63**, 5375, 1975. (b) J. Mispelter, M. Momenteau, D. Lavalette, J.-M. Lhoste, *J. Am. Chem. Soc.* **105**, 5165, 1983. (c) E. Oldfield, H. C. Lee, C. Coretsopoulos, F. Adebodum, K. D. Park, S. Yang, J. Chung, B. Phillips, *J. Am. Chem. Soc.* **113**, 8680, 1991.

[25] L. Stryer, *Biochemistry*, Fourth Edition, Freeman, New York, 1995.

[26] J. B. Johnson, D. C. Lamb, H. Frauenfelder, J. D. Mller, B. McMahon, G. U. Nienhaus, R. D. Young. *Biophys. J.* **71**, 1563, 1996.

[27] T. Li, M. L. Quillin, G. N. Phillips, J. S. Olson. *Biochemistry* **33**, 1433, 1994.

[28] M. Eichinger, P. Tavan, J. Hutter, M. Parrinello, *J. Chem. Phys.* **110**, 10452, 1999.

[29] Schulze, B. and J. D. Evanseck. *J. Am. Chem. Soc.* **121**, 6444, 1999..

[30] (a) D. P. Braunstein, K. Chu, K. D. Egeberg, H. Frauenfelder, J. R. Mourant, G. U. Nienhaus, P. Ormos, S. G. Sligar, B. A. Springer, R. D. Young, *Biophys. J.* **65**, 2447, 1993. (b) T. Li, M. L. Quillin, G. N. Phillips, J. S. Olson. *Biochemistry*, **33**, 1433, 1994.

[31] (a) G. B. Ray, X.-Y. Li, J. A. Ibers, J. L. Sessler, T. G. Spiro. *J. Am. Chem. Soc.* **116**, 162, 1994. (b) G. N. Phillips, M. L. Teodoro, T. LI, B. Smith, J. S. Olson. *J. Phys. Chem. B* **103**, 8817, 1999.

[32] C. Rovira, B. Schulze, M. Eichinger, J. D. Evanseck, M. Parrinello. *Biophys. J.* **81**, 435, 2001.

[33] M. Unno, J. F. Christian, J. S. Olson, J. T. Sage, P. M.Champion *J. Am. Chem. Soc.* **120**, 2670, 1998.

[34] (a) A. D. Becke, *J. Chem. Phys.* **84**, 4524, 1986. (b) J. P. Perdew, *Phys. Rev. B* **33**, 8822, 1986.

[35] N. Troullier, J. L. Martins, *Phys. Rev. B* **43**, 1993, 1991.

[36] S. G. Louie, S. Froyen, M. L. Cohen, *Phys. Rev. B* **26**, 1738, 1982.

[37] (a) Car, R; Parrinello M. *Phys. Rev. Lett.* **55**, 2471, 1985. (b) G. Galli, M. Parrinello, *Computer Simulation in Materials Science*, Edited by V. Pontikis and M. Meyer, Kluwer, Dordrecht, 1991, and references therein.