The structure and dynamics of the Fe–CO bond in myoglobin

2003 J. Phys.: Condens. Matter 15 S1809

(http://iopscience.iop.org/0953-8984/15/18/314)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **15** (2003) S1809–S1822 PII: S0953-8984(03)54656-2

The structure and dynamics of the Fe–CO bond in myoglobin

Carme Rovira

Centre de Recerca en Química Teòrica, Barcelona Science Park, Josep Samitier 1-5, 08028 Barcelona, Spain

E-mail: crovira@pcb.ub.es

Received 11 October 2002
Published 28 April 2003
Online at stacks.iop.org/JPhysCM/15/S1809

Abstract

This paper is a review of our recent work on the structure and dynamics of the Fe–CO bond in carbonmonoxy myoglobin (MbCO), performed using density functional theory, Car–Parrinello molecular dynamics and hybrid quantum mechanics/molecular mechanics approaches. The results of these investigations have served to shed light onto one of the long standing questions in myoglobin research: whether the protein discriminates the CO ligand with respect to O_2 by distorting the FeCO bond. The calculations show that both in the gas phase and in the protein the Fe–CO bond is essentially linear and therefore exclude the hypothesis that the CO in MbCO is sterically hindered. In contrast, hydrogen bonding between the O_2 ligand and the His64 residue easily explains the protein discrimination for CO.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The binding of oxygen (O_2) and carbon monoxide (CO) to myoglobin and synthetic model compounds remains a topic of great ongoing interest [1]. CO is a ubiquitous toxin [2] because it is present in large amounts in cells and because its bond with the myoglobin iron atom is less dissociable than the bond of oxygen. It has been known since the 1970s that the affinity ratio CO/O_2 is lower in the protein than in synthetic model compounds by a factor of $\sim 10^2$ in the equilibrium constant for the ligand binding reaction (Springer *et al* [1]). In other words, the protein weakens the CO affinity with respect to that of O_2. Nevertheless, the origin of the protein discrimination for CO is still one of the most controversial issues in haem-protein research.

According to x-ray and neutron diffraction studies, the binding of CO to the haem active centre leads to a distorted FeCO unit. This distortion is generally described in terms of the tilt (δ) and bend (θ) angles, depicted in figure 1. Table 1 lists the different FeCO structures that have been reported for native MbCO since 1986 [3–9]. Two positions for the CO are given in

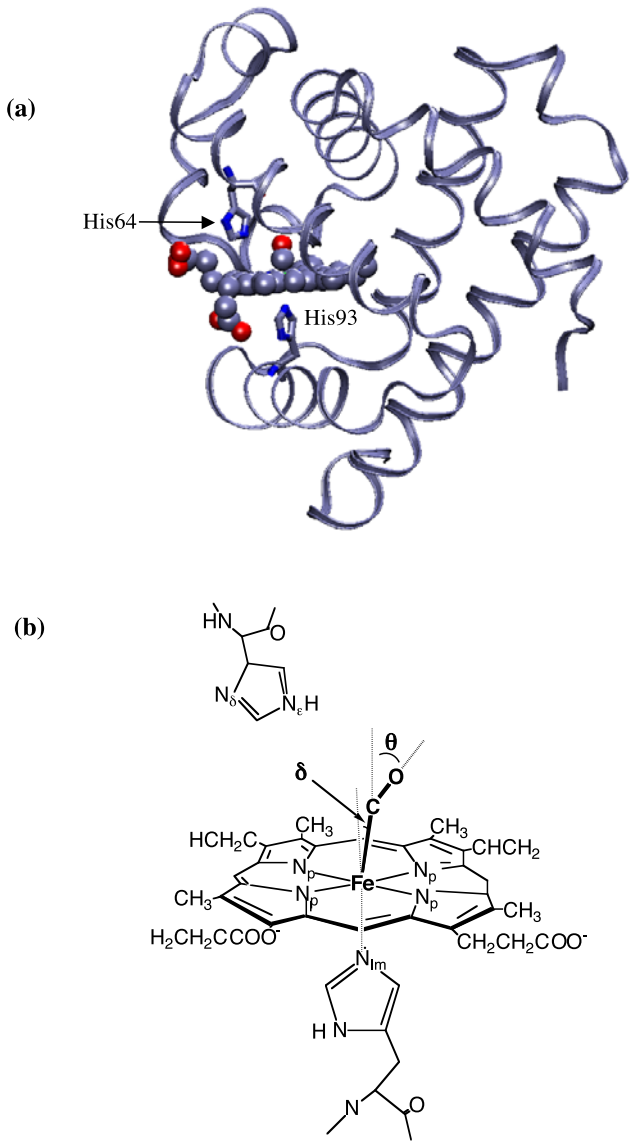


Figure 1. (a) Structure of MbCO (PDB entry 1BZR, at 1.15 Å resolution [7]). The haem active centre is highlighted (van der Waals spheres), as well as the proximal and distal histidines (His93 and His64, respectively, shown as sticks). (b) Tilt (δ) and bend (θ) angles commonly used to describe the FeCO distortion.

the first two entries of the table [3, 4] since the Fe–CO bond is affected by static disorder. As shown in table 1, most of the x-ray structures show large values of the tilt and bend angles. In particular, the two highest resolution structures (last two entries of the table) disagree in the degree of FeCO distortion. The magnitude of the Fe–C, Fe–N_p and Fe–N_{Im} distances show less variability among the different refinements, specially within the high resolution structures.

The results of the structural analyses of MbCO contrast with those on synthetic models of the protein (e.g. molecules of the the *picket fence* family) for which the FeCO fragment is linear.

Table 1. Structural information available (from x-ray diffraction and neutron scattering data) on the local structure around the iron atom in MbCO. Distances are given in angstroms and angles in degrees. Porphyrin nitrogens are denoted as N_p, while N_{Im} refer to the nitrogen atom of the axial imidazole which is coordinated to Fe. N_ε is one of the nitrogen atoms of the distal histidine (the one closest to CO, as shown in figure 1(b)).

Year	1986	1991	1994	1996	1999	1999	2002
Resolution	1.5 Å	1.8 Å	1.9 Å	2.0 Å	1.5 Å	1.2 Å	1.3 Å
Reference	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Structural parameter							
Bend (θ)	39/60	34/45	20	47.3	7.4	19	4.3
Tilt (δ)	2.7	13.2	3.9	>15	4.8	9	1.76
C–O	1.17/1.20	1.20	1.12	1.13	1.13	1.09	1.14
Fe–C	1.92	2.13	1.85	1.91	1.73	1.82	1.76
Fe–N_p	1.90–2.00	2.01	1.96	1.88–2.01	1.99–2.03	1.98	1.98–2.02
Fe–N_{Im}	2.19	2.20	2.30	2.27	2.11	2.06	2.06
$O\cdots N_\varepsilon$	3.96/2.66	3.48/2.60	2.81	7.44[a]	3.21	3.16/2.74/6.58[b]	2.81/3.41[b]

[a] This structure was solved at pH 4.0. At this pH the His64 residue is out of the haem pocket and thus far from the bound CO.

[b] The His64 residue was found to be disordered into more than one position.

Because of this, it has been often assumed that the Fe–CO bond in the protein is weaker than in the models and this argument has been used to explain the protein discrimination for CO [10]. However, several experiments do not support the so-called *steric interpretation*. For instance, kinetic and thermodynamic measurements in a variety of proteins have not found a clear relation between CO affinity and FeCO distortion [1] and spectroscopic studies have excluded the possibility of a large Fe–CO distortion in the protein [11]. On another hand, photoselection measurements have shown that the angle between the haem normal and the C–O dipole is less than 7° (Lim *et al* [1], Sage and Jee [12]). Therefore, even though experiments in synthetic models show that steric hindrance can control the binding affinity of CO and O_2 [13], the fact that the CO in myoglobin is sterically hindered is nowadays regarded with scepticism. To explain the protein discrimination for CO, other factors such as hydrogen bond or electrostatic interactions have been proposed by Springer *et al* [1], Ray *et al* [11] and Sage and Champion [14].

Theoretical calculations based on density functional theory (DFT) have been very useful to exclude the steric interpretation. Gas-phase calculations provide a precise knowledge of the intrinsic properties of the haem–ligand bonds [15]. For instance, Ghosh and Bocian [16] investigated the energetic cost of distorting the FeCO bond from linearity using a simplified model where the porphyrin is substituted by two amidinato ligands. This study demonstrated that small distortions of the FeCO moiety have a negligible cost in terms of energy. Very similar conclusions were reached by Rovira *et al* [17] using a larger model. Later on, Spiro and Kozlowski and Havlin *et al* [18] reinvestigated the FeCO bending and tilt potentials. The main findings of these static computations, along with the experimental information available, have been recently reviewed [19]. The dynamics of the FeCO bond at room temperature [20] completed the picture of the FeCO deformability. Very recently, the effect of the protein environment on the FeCO structure has also been analysed [21].

This paper summarizes our work in the modelling of the interaction of CO myoglobin using Car–Parrinello molecular dynamics and hybrid QM/MM approaches. The calculations provide additional support for the fact that the the Fe–CO bond in MbCO is linear. Moreover, it is shown that hydrogen bonding between the O_2 ligand and the His64 residue could explain the protein discrimination for CO.

2. Computational details

All calculations presented here are based on the DFT [22] within the local density approximation (LDA). The Kohn–Sham orbitals [23] are expanded in a plane wave (PW) basis set, with a kinetic energy cut-off of 70 Ryd. The Ceperley–Alder expressions for correlation and gradient corrections of the Becke–Perdew type are used [24], as well as *ab initio* pseudopotentials, generated using the Troullier–Martins scheme [25], including the non-linear core correction [26] for the iron atom. The Car–Parrinello method [27, 28], based on a combination of a molecular dynamics (MD) algorithm with electronic structure calculations by means of DFT, has been used with success in the study of different systems of biological interest [29]. Structure optimizations were performed with no constraints starting from non-symmetric structures. The convergence of the results with the energy cut-off in the PW expansion was investigated in a previous work [30].

MD 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 were enclosed in supercells of 16 Å $\times$ 16 Å $\times$ 20 Å, periodically repeated in space. They were allowed to evolve during 2 ps in order to achieve vibrational equilibration. The MD was performed for a total period of 15.5 ps. Hybrid quantum mechanics/molecular mechanics (QM/MM) calculations were done using the EGO-CPMD code [31], which is an interface between the classical MD code of Eichinger *et al* (named EGO_VIII [32]), based on the CHARMM force field [33], and the Car–Parrinello MD code written by Hutter (CPMD) [28]. The interface between the QM and MM regions is treated using the link atom approximation [31].

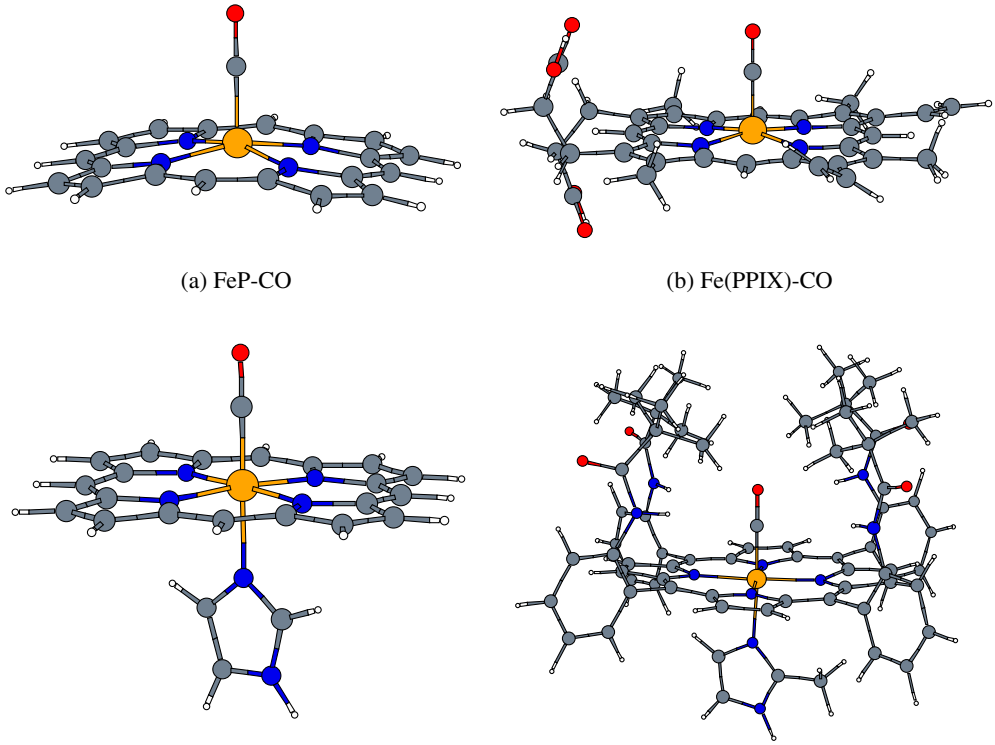
3. The Fe–CO bond in the gas phase

3.1. Structure

Several models have been used to investigate the properties of the Fe–CO bond in the gas phase. Two of the simplest ones are the five-coordinated iron–porphyrin–CO (FeP–CO) and the six-coordinated iron–porphyrin–imidazole–CO (FeP(Im)–CO), where the axial imidazole ligand mimics the effect of the proximal histidine aminoacid. Other authors have also used models based on substituting the iron–porphyrin by two amidinato ligands (i.e. a $(NH–CH{=}CH–CH{=}NH)^-$ ring structure) [16]. In order to quantify the effect of a crowded binding pocket, calculations have also been performed in larger models including the complete active centre of MbCO (Fe(PPIX)–CO, shown in figure 2, where PPIX is the protoporphyrin IX) and the *picket fence* biomimetic molecule Fe(T_{piv}PP)(2-meIm)–CO (T_{piv}PP = tetrapivalaminophenyl porphyrin, meIm = methylimidazole). The *picket fence* based molecules have been widely used as myoglobin models [14].

In agreement with experiments, the lowest energy spin state of the above systems is a closed shell singlet [14]. The corresponding optimized structures are shown in figure 2 and table 2 lists their most relevant structural parameters. As shown in figure 2, the five-coordinated FeP–CO and Fe(PPIX)–CO complexes are characterized by having a curved porphyrin. Disregarding the porphyrin non-planarity, the structure of the five-coordinated complexes is already quite similar to that of the six-coordinated complex FeP(Im)–CO.

The distortion of the porphyrin upon CO binding reinforces the bonding between the Fe(d_{z^2}) orbital and the $3\sigma_g$ orbital of the CO molecule (figure 3(a)). The symmetry reduction on going from a planar to a non-planar porphyrin allows the d_{z^2} orbital to mix with the p_z and s orbitals of iron. The resultant hybrid orbital, shown in figure 3(b), is stabilized further by this



(a) FeP-CO (b) Fe(PPIX)-CO

(c) FeP(Im)-CO (d) Fe(TpivPP)(2-meIm)-CO

Figure 2. Optimized structures for the myoglobin models investigated.

Table 2. Main parameters defining the optimized structures for the complexes investigated. Distances are given in Å, angles in degrees and energies in kcal mol^{-1}.

Structure	Fe–C	C–O	<Fe–C–O	Fe–N_p	Fe–N_{Im}
FeP–CO	1.69	1.17	180	1.99	—
FePPIX–CO	1.69	1.17	180	1.99	—
FeP(Im)–CO	1.72	1.17	180	2.02	2.10
Fe(T_{piv}PP)(2-meIm)–CO	1.72	1.17	180	2.01	2.11
Expt[a]	1.77	1.12	179	2.02	2.10

[a] Data corresponding to the Fe(TPP)(py)–CO complex [40].

mixing process (the antibonding interaction with the N_p ligands is reduced) and it becomes hybridized towards the missing CO ligand [34]. The overlap between this orbital and the $3\sigma_g$ orbital of the CO molecule is in this way enhanced, resulting in a more stable Fe–CO bond.

The calculations on large models evidence that the porphyrin substituents do not cause a significant change either on the Fe–CO bond or on the porphyrin structure: the porphyrin ring in FeP(Im)–CO is practically identical to that in Fe(T_{piv}PP)(1-meIm)–CO and the same occurs when comparing FeP–CO and Fe(PPIX)-CO. Only the Fe–N_ε axial bond is found to be longer for Fe(T_{piv}PP)(1-meIm)–CO compared to FeP(Im)–CO. Nevertheless, this is due to the steric interaction between the 2-me substituent of the imidazole and the porphyrin ring. In fact, additional calculations on the species obtained by removing the methyl group find that the Fe–N_ε distance is the same as the one found in FeP(Im)–CO. Thus, in spite of the structural

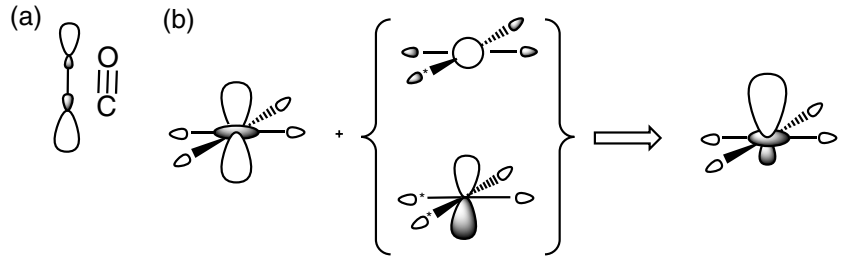


Figure 3. (a) $3\sigma_g$ orbital of the CO molecule. (b) Hybrid molecular orbital resulting from the combination of the three molecular orbitals depicted (see text). Although these orbitals have a main iron character, they also have a small contribution from the ligand orbitals.

complexity of the binding pocket, the local structure around the iron atom in Fe(T_{piv}PP)(1-meIm)–CO does not change with respect to the simplest iron–porphyrin models.

3.2. Binding energy

Binding energies with respect to dissociation of the axial ligands were computed as the difference between the energy of the complex and that of the isolated fragments at their optimized structures. The binding energy of the Fe–CO bond amounts to 26 kcal mol^{-1} for FeP–CO and it is much larger (35 kcal mol^{-1}) for FeP(Im)–CO. Therefore, the presence of the imidazole ligand stabilizes the Fe–CO bond. This can be understood in terms of the polarization of the d_{z^2} orbital which interacts with the $3\sigma_g$ orbital of the CO molecule [34]. Similarly, the binding energy of the Fe-Imidazole bond increases upon binding of CO (from 7 kcal mol^{-1} in the five-coordinated FeP(Im) to 12 kcal mol^{-1} for FeP(Im)–CO). It should be noted that a similar effect was found for oxygen complexes [35]. Therefore, one of the roles of the proximal histidine residue in myoglobin and haemoglobin is that of reinforcing the bond with the CO and O_2 ligands.

The influence of the porphyrin substituents in the energy of the Fe–CO bond appears to be very sensitive to the polarity of these substituents. For instance, in the case of Fe(PPIX)–CO, the energy of the Fe–CO bond does not change with respect to the unsubstituted complex FeP–CO. In contrast, a large enhancement of the binding energy is found for Fe(T_{piv}PP)(1-meIm)–CO with respect to FeP(Im)–CO (29 kcal mol^{-1}). Model calculations using point charges replacing the T_{piv}P substituents show that the energy increase can be attributed to electrostatic interactions of the CO with the polar T_{piv}P side chains. Nevertheless, this enhancement is likely to be overestimated in the calculation, since the strong electrostatic interaction of the T_{piv}P side chains with the CO ligand would decrease in the presence of a solvent, something which is not taken into account in the gas phase calculations. Nevertheless, the computed values are useful to identify trends along the series and to pinpoint the chemical groups responsible for a given change. In this respect, the results obtained show that although the structure of the FeCO fragment is insensitive to the presence of porphyrin substituents, the binding energy of the CO ligand is very sensitive to the polarity of these substituents (it does not change in the non-polar environment provided by the PPIX substituents, but it is significatively enhanced by the polar T_{piv}P substituents). In the context of the protein, these results suggests that the polarity of the haem pocket could be able to control the relative binding energy among the CO and O_2 ligands. Therefore, while steric hindrance probably has no relevance for the protein–ligand binding control, electrostatic effects are likely to play a major role in these processes.

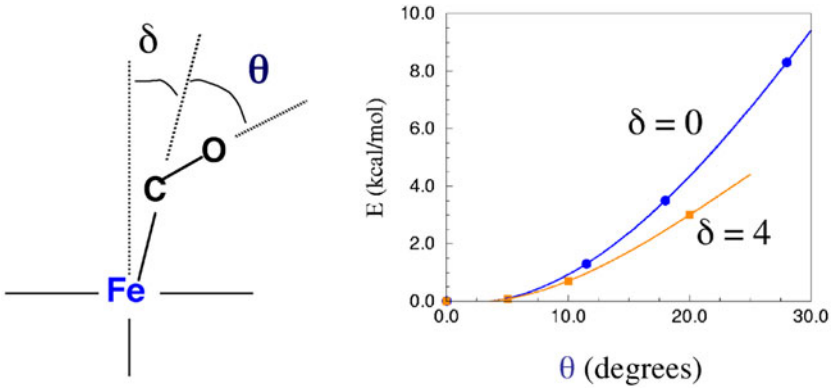


Figure 4. Energy cost of distorting the FeCO bond along the δ and bend θ coordinates.

3.3. Dynamics of the Fe–CO fragment

It is particularly interesting to quantify the energy involved in a distortion of the Fe–CO fragment, since this distortion has been traditionally invoked as responsible for the weakening of the Fe–CO bond in the protein [10]. To explain the reduced CO/O_2 affinity in the protein with respect to free haem in terms of a weaker Fe–CO bond, a decrease of the Fe–CO binding enthalpy by $\approx$2 kcal mol^{-1} would be required. This could be accomplished, for instance, by distorting the Fe–CO bond [10, 13].

Figure 4 shows the energy increase with respect to changes in the bend (θ) and tilt (δ) angles defining the Fe–CO structure. The calculations were performed by means of geometry optimization of the FeP(Im)–CO model, relaxing all degrees of freedom except the tilt and bend angles. It can be observed that small deformations (up to 7° in the bending angle, which are similar to the maximum distortions reported by photoselection and infrared crystallography measurements by Lim *et al* [1], Sage and Jee [12]) cost less than 0.5 kcal mol^{-1}. It is also apparent that Fe–C–O bending becomes easier when the bond is tilted (for a tilting of 4°, a bending by 10° requires only 0.7 kcal mol^{-1} while it requires 1 kcal mol^{-1} for the untilted bond). Similar results have been obtained by other groups at different levels of theory and using either the same or more simplified models [16, 18]. The results of figure 4 show that only distortions larger than $\sim$15° could be biologically relevant, as they would correspond to a weakening of the Fe–CO bond by $\sim$2 kcal mol^{-1}.

MD simulations at room temperature can provide a picture of the fluxionality of the Fe–CO bond. To this aim, a Car–Parrinello MD simulation was performed for the FeP(Im)–CO model for a total period of 15 ps. As a way to display the dynamical motion of the ligand, the projection of the C and O projections on the porphyrin plane was monitored along the simulation. Figure 5 shows the trajectories sampled by both the C and O atoms. The trajectories appear to be rather complex and concentrated around the iron atom, with the one of oxygen being more spread ($\approx$0.4 Å from the centre) than that of the carbon atom ($\approx$0.2 Å). Nevertheless, with respect to the size of the porphyrin core (Fe–N_p = 2.02 Å), the trajectories shown in figure 5 correspond to just a very small area over the iron atom (located at 0.0 in figure 5). Further analysis of the time evolution of the CO orientation (data not shown here) reveals that the projection of the C–O axis on the porphyrin plane visits all the porphyrin quadrants in a very short time ($\approx$0.5 ps). Therefore, the global picture that can be inferred from our simulation is that of an essentially upright FeCO unit, with the CO ligand undergoing a fast complex motion within a very small region around its equilibrium position. It should also be noted that, given the

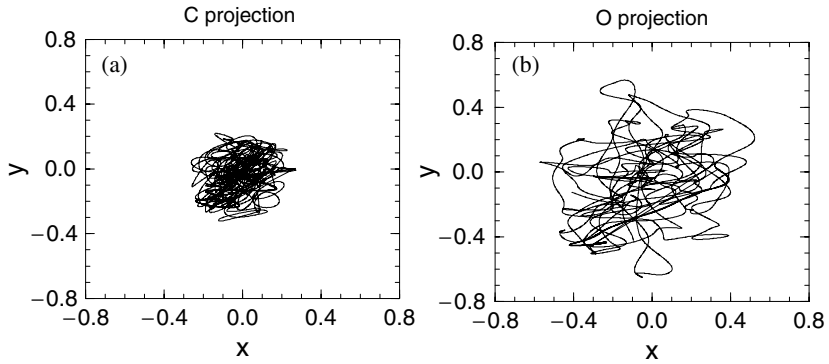


Figure 5. Configurational space sampled by the projection of the CO ligand on the porphyrin plane. The x, y axes are aligned with the Fe–N_p bonds. (a) Carbon atom; (b) oxygen atom. Values are given in angstroms.

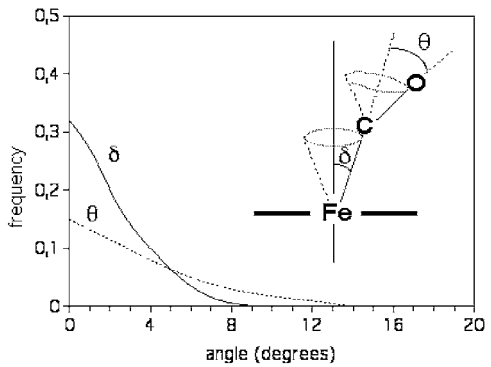


Figure 6. Frequency distribution of the δ and θ angles obtained from the CPMD simulation of the FeP(Im)–CO complex at 300 K. The distribution is not normalized.

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 fluxional CO molecule, which samples many different conformations with different probability in a short time.

Figure 6 shows the probability distribution of the δ and θ angles obtained from the simulation. It is apparent that small fluctuations of the tilt and bend angles ($\delta < 8°, \theta < 13°$) have a sizable probability of taking place, but larger deformations do not occur. Therefore, for an FeCO not perturbed by the environment, small δ–θ deviations can occur only due to the thermal motion of the ligand. Of course, these little deformations cannot have any relevance for the protein discrimination against CO.

In summary, the Fe–CO bond is perfectly linear in all models considered and only large deformations ($>15°$) should be of biological relevance. Therefore, unless the CO is severely distorted in the protein, the calculations exclude the steric interpretation as the origin of the protein discrimination for CO.

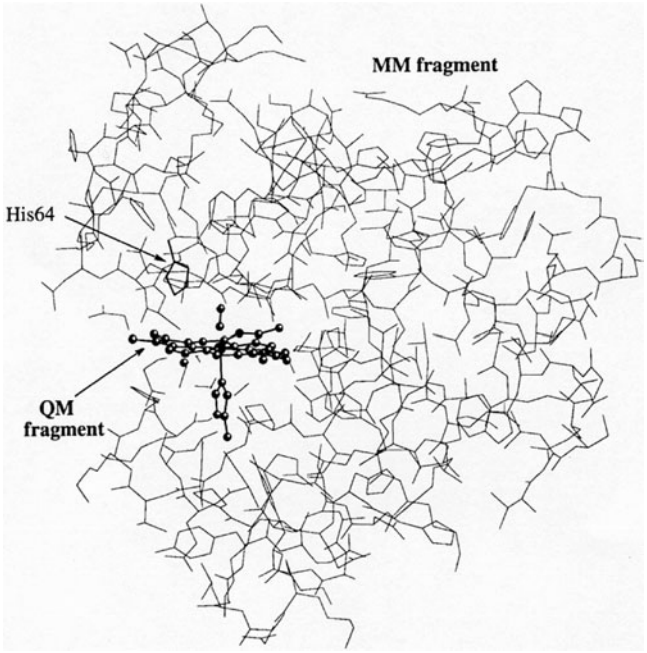


Figure 7. QM/MM partition chosen in the calculations including the protein environment.

4. The Fe–C–O structure in the protein

4.1. Structure

The modelling of the FeCO bond in MbCO was performed by means of the hybrid QM/MM approach described in section 2. Figure 7 depicts the type of QM–MM boundary used in the calculations. The CO ligand, the porphyrin and the axial imidazole were treated quantum mechanically (i.e. included in the QM region). The vinyl and propionate porphyrin substituents were not included in QM, since it had previously been found that they do not affect the properties of the Fe–ligand bonds (sections 3.1 and 3.2). On the other hand, it is crucial to consider the imidazole of the proximal His in QM, since the imidazole strengthens the Fe–CO bond (section 3.2). The use of this QM–MM partition ensures that the energy/spin/structure relations of the haem are well described. Four link atoms [31] are used to separate the porphyrin from the haem vinyl and propionate substituents, plus another link atom saturating the C–C bond of the His64 residue. In addition, the protein is enveloped in a 37 Å sphere of equilibrated TIP3P water molecules (see figure 8) in order to take into account solvation effects. The numbers of QM and MM atoms treated in the calculation are 63 and 20 000, respectively.

Before starting the calculations, it is important to chose an appropriate initial structure for the protein. While the x-ray structure of MbCO is a possible starting point, the fact that it corresponds to an average among the many different instantaneous protein conformations could lead to artifacts in the simulation. It is physically sounder to consider snapshots of classical MD simulations performed on the equilibrated protein. It is also desirable that, for consistency, these simulations are performed using the same force field as in the QM/MM calculations.

Classical simulations of MbCO using the CHARMM force field were done for different tautomerization states of the distal histidine residue (His64) [37]. These simulations showed

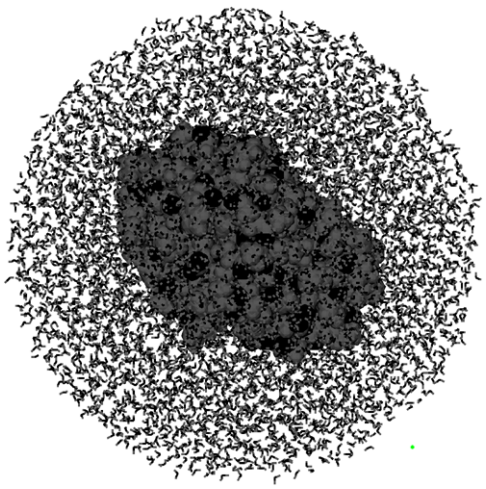


Figure 8. Protein solvated in the water shell, used in the QM/MM calculations.

that when His64 is protonated at N_δ (i.e., the N_δ tautomer) it often rotates so that it exposes either the N_δ–H bond or the unprotonated N_ε atom towards the CO. These two situations are depicted in figure 9. Two snapshots corresponding to these two extreme distal His conformations were taken, which are denoted as I and II. Another snapshot (III) from a simulation that started with the His64 protonated at N_ε (N_ε tautomer) was also considered. Since rotation of His64 did not occur in the timescale of the classical simulation (1 ns), it was artificially forced by inducing an 180° rotation around the His64 C–C bond (IV in figure 9). Finally, a fifth snapshot (V in figure 9) in which the distal His moved away from the CO was considered. This event occurred after 600 ps of the classical MD simulation. These protein configurations, denoted as I–V in figure 9, are representative of the dynamics of the haem pocket.

For each of these protein conformations, a structural relaxation was performed. Table 3 summarizes the results obtained for the structure around the Fe atom. As expected, the haem–ligand structure is not very sensitive to the protein conformation. In all cases we obtain a similar structure, which is not far from the gas phase values (last row in table 3). 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 (III). The maximum distortion observed for the FeCO angle is 3.9°. Therefore, the FeCO structure is not significantly influenced by the protein environment and the FeCO fragment can be considered as essentially linear. This is in contrast with the results of structural analyses (table 1) reporting a distorted FeCO moiety, but supports the conclusions of several other experimental studies that concluded that this distortion is marginal (Lim *et al* [1], Sage and Jee [12]).

4.2. Interaction of the ligand with the distal His residue

Having excluded the steric interpretation as the origin of the CO discrimination, it is interesting to analyse whether hydrogen bonding could explain such discrimination. To this aim, the energy of the interaction of the CO ligand with the distal histidine was analysed. In order to isolate the hydrogen bond interaction due to the His64 residue, the computations were done in the isolated haem–His93–CO$\cdots$His64 system for each of the protein conformations I–V.

The interaction energy of the ligand with the distal residue (last column in table 3) turns out to be very dependent on the conformation and protonation state of the distal His. The interaction between His64 and CO vanishes for configuration V. Configuration I leads to a

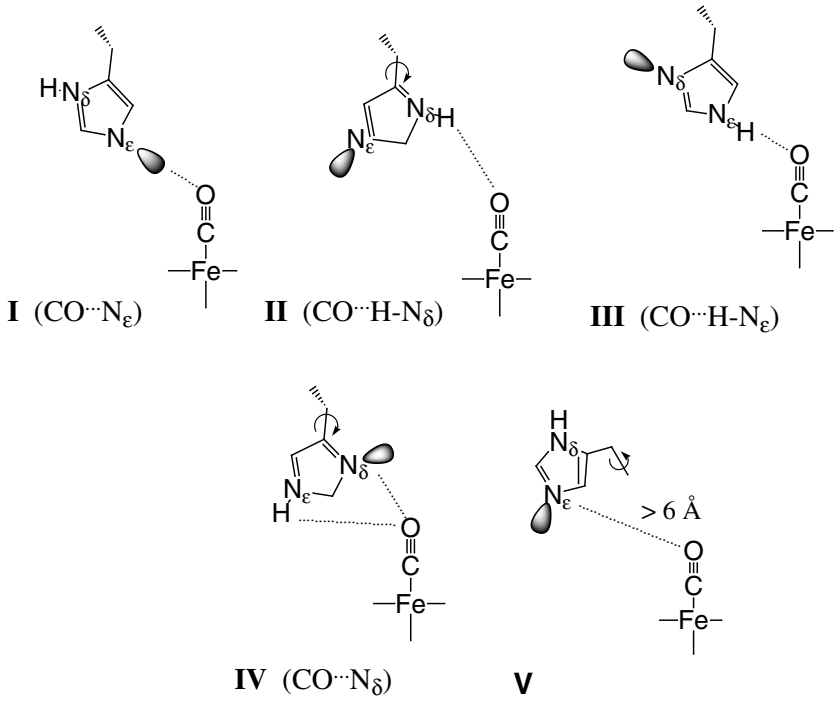


Figure 9. Orientation and tautomerization state of the His64 residue in each of the five different protein conformations considered (see text). Orientations I, II and V correspond to the N_δ tautomer of His64, while III and IV correspond to the N_ε tautomer (see text).

Table 3. Main parameters defining the optimized haem–CO structure of each protein conformation I–V. The last row corresponds to the results obtained for the FeP(Im)–CO isolated model (section 3.1).

Structure	Interaction type	O···X (Å)	C–O (Å)	Fe–C (Å)	<Fe–C–O (deg.)	Fe–N_p (Å)	$\Delta E_{O\cdots X}$ (kcal mol^{-1})
Exp.	CO···N_ε	4.0–2.60	1.09–1.21	1.73–2.21	120—172	2.01–2.06	+2.0
I	CO···N_ε	3.39	1.16	1.76	177.3	2.00–2.02	−2.5
II	CO···H-N_δ	3.47	1.16	1.75	179.3	1.98–2.03	−3.4
III	CO···H-N_ε	2.69	1.17	1.74	176.1	1.99–2.02	−0.9
IV	CO···N_δ	3.90	1.16	1.74	175.7	1.99–2.02	−0.1
	CO···H–C	2.18					
V	CO···N_ε	6.03	1.16	1.75	177.6	1.99–2.03	0.0
	CO···H–C	4.03					
FeP(Im)–CO	—	—	1.17	1.72	180.0	2.02	—

repulsive interaction (2 kcal mol^{-1}), while the interaction is favourable when the protonated nitrogen is close to the CO. The largest stabilization is found for the N_ε tautomer, as in this case the N_ε–H···OC interaction becomes geometrically more favoured than in the other cases. This is at variance from the common assumption [11] that only the binding of O_2 could be stabilized by interaction with the distal histidine. However, the calculations support recent resonance Raman measurements [38] that show spectroscopic evidence of a hydrogen bond between CO and His64.

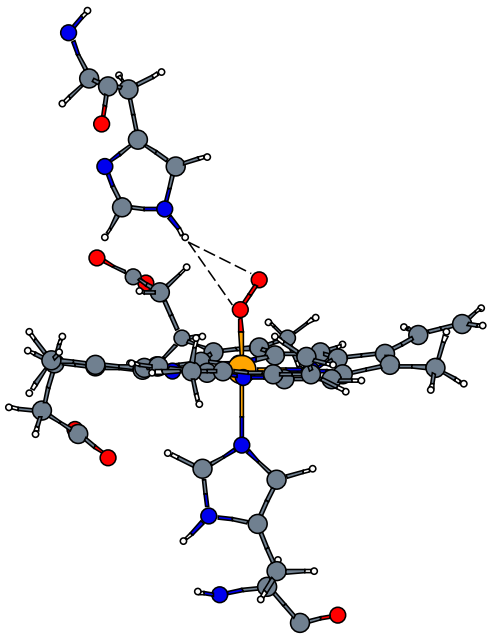


Figure 10. Orientation of the His64 residue with respect to the Fe–O_2 bond obtained for MbO_2.

In the case of MbO_2, there is experimental evidence that the His64 residue is protonated at N_ε [39]. Therefore, only protein conformation III was considered for the calculations. The corresponding orientation of the O_2 ligand obtained in the optimization is shown in figure 10. The computed hydrogen bond to O_2 is stronger than that of CO by 2.3 kcal mol^{-1}. This value should be regarded as a lower bound since no initial relaxation of the MbO_2 structure was performed and only one protein conformation was considered. Nevertheless, the difference obtained is sufficient to explain the protein discrimination for CO in terms of hydrogen bonding. On the other hand, other authors [41] have recently analysed the strength of this hydrogen bond using either a simplified model for HbO_2 or a model consisting on an iron–porphyrin interacting with an imidazole fixed at the x-ray position. The similarity of the results obtained seems to indicate that the strength of the His64$\cdots O_2$ interaction does not change with small displacements of the His64 residue. Therefore, while our results exclude steric effects as being relevant for the protein discrimination for CO, they support the hypothesis that the strong hydrogen bond interaction of the O_2 ligand with the His64 residue is at the basis of the protein discrimination for CO.

5. Summary and conclusions

In this study the structure and energetics of the binding of CO in myoglobin and synthetic models has been quantified. The binding of CO induces a significant curvature of the haem when binding to the iron atom, although the porphyrin planarity is restored by the binding of an imidazole ligand. The vinyl, methyl and propionic substituents of haem do not influence the structure and energy of the Fe–CO bond. Instead, the ligand binding energy is significantly enhanced by the presence of the polar porphyrin substituents of the *picket-fence* molecule.

Analysis of the energy increases upon bending (θ) and tilting (δ) the Fe–CO bond shows that only distortions larger than $\sim 15^\circ$ could be biologically relevant, as they would correspond

to a weakening of the Fe–CO bond by more than 2 kcal mol^{-1}. Small deformations have a negligible energy cost and can be available by thermal motion, as evidenced by Car–Parrinello molecular dynamics simulations at 300 K. The simulations show that the dynamics of the FeCO unit is characterized by a fast motion of the ligand around its equilibrium position, with a maximum distortion of $\theta \approx 13°$ and $\delta \approx 8°$.

Hybrid QM/MM calculations based on DFT combined with the CHARMM force field highlight the effect of the distal pocket conformation on the properties of the Fe–CO bond in MbCO. The calculations show that the local structure around the Fe atom is insensitive to the haem environment. An essentially linear FeCO bond is found for different distal pocket conformations, which leads to the conclusion that the haem–CO structure is quite robust and not influenced by the protein environment. Therefore, the results obtained rule out the steric interpretation to explain the origin of the protein discrimination for CO.

Unlike the structure, the strength of the CO$\cdots$His64 interaction appears to be very dependent on the protein conformation and, in particular, on the orientation and tautomerization state of His64. The CO ligand turns out to be substantially stabilized by interaction with the distal histidine residue, in contrast with the common assumption that such stabilization only occurs for O_2. Nevertheless, the strength of the CO$\cdots$His64 interaction is smaller than the one obtained for the oxygen ligand, with an energy difference of >2.3 kcal mol^{-1}. This gives support to the argument that myoglobin favours the binding of O_2 with respect to CO by reinforcing its hydrogen bond interaction with the His64 residue.

In summary, the calculations have quantified the interplay between the structure, energy and dynamics of the haem active centre and its interaction with the protein. This helps to understand several previously unclear aspects such as the precise structure of the Fe–CO bond, its intrinsic dynamics and the role of the proximal and distal histidines and its properties. The results obtained show that both in the gas phase and in the protein the Fe–CO bond is essentially linear and therefore exclude the hypothesis that the CO in MbCO is sterically hindered. In contrast, hydrogen bonding between the O_2 ligand and the His64 residue easily explains the protein discrimination for CO.

Acknowledgments

This work was supported by the CIRIT under project 2001SGR-00044. Computing support from the Max Planck Institute (Garching, Germany) and the CEPBA-IBM Research Institute of Barcelona (Spain) is acknowledged. CR thanks Jürg Hutter, Pietro Ballone, Karel Kunc, Mauro Boero, Enric Canadell and Roger Rousseau for many useful discussions, and specially Professor Michele Parrinello for his support and continuous interest throughout this work. The financial support of the 'Ramón y Cajal' programme of the Spanish Ministry of Science and Technology is also acknowledged.

References

[1] Springer B A, Sligar S G, Olson J S and Phillips G N 1994 *Chem. Rev.* **94** 699
Schlichting I, Berendzen J, Phillips G N and Sweet R M 1994 *Nature* **371** 808
Lim M, Jackson T A and Anfinrud P A 1995 *Science* **269** 962
Slebodnick C and Ibers J A 1997 *J. Biol. Inorg. Chem.* **2** 521
[2] Urbanetti J S 1981 *Prog. Clin. Biol. Res.* **51** 355
[3] Kuriyan J, Wilz S, Karplus M and Petsko G 1986 *J. Mol. Biol.* **192** 133
[4] Cheng X and Schoenborn B P 1991 *J. Mol. Biol.* **220** 381
[5] Schlichting I, Berendzen J, Phillips G N and Sweet R M 1994 *Nature* **371** 808
[6] Yang F and Phillips G N 1996 *J. Mol. Biol.* **256** 762

[7] Kachalova A N, Popov H and Bartunik D 1999 *Science* **284** 473
[8] Vojtechovsky J, Chu K, Berendzen J, Sweet R M and Schlichting I 1999 *Biophys. J.* **77** 2153
[9] Phillips G N and Romo T 2002 (PDB entry 1JW8) at press
[10] Stryer L 1995 *Biochemistry* 4th edn (New York: Freeman)
[11] Ray G B, Li X-Y, Ibers J A, Sessler J L and Spiro T G 1994 *J. Am. Chem. Soc.* **116** 162
Li X-Y and Spiro T G 1988 *J. Am. Chem. Soc.* **110** 6024
[12] Sage J T and Jee W 1997 *J. Mol. Biol.* **274** 21
[13] Collman J P 1997 *Inorg. Chem.* **36** 5145
[14] Sage J T and Champion P M 1996 Small substrate recognition in heme proteins *Compr. Supramol. Chem.* **5** 171–213
Olson J S and Phillips G N 1996 *J. Biol. Chem.* **271** 17593
Ostermann A, Waschipky R, Parak F G and Nienhaus G U 2000 *Nature* **404** 205
[15] Early studies using semiempirical and *ab initio* approaches were limited to using a fixed molecular structure and thus the structure and dynamics of the Fe–CO bond were not analysed.
See for instance
Case D A, Huynh B H and Karplus M 1979 *J. Am. Chem. Soc.* **101** 4433
Nakatsuji H, Tokita Y, Hasegawa J and Hada M 1996 *Chem. Phys. Lett.* **256** 220
[16] Ghosh A and Bocian D F 1996 *J. Phys. Chem.* **100** 6363
[17] Rovira C, Kunc K, Hutter J, Ballone P and Parrinello M 1997 *J. Phys. Chem.* A **101** 8914
[18] Havlin R H, Godbout N, Salzmann R, Wojdelski M, Arnold W, Schulz C E and Oldfield E 1998 *J. Am. Chem. Soc.* **120** 3144
Spiro T G and Kozlowski P M 1998 *J. Am. Chem. Soc.* **120** 4524
[19] Spiro T G and Kozlowski P M 2001 *Accounts Chem. Res.* **34** 137
[20] Rovira C and Parrinello M 2000 *Biophys. J.* **78** 93
[21] Rovira C, Schulze B, Eichinger M, Evanseck J and Parrinello M 2001 *Biophys. J.* **81** 435
[22] Hohenberg P and Kohn W 1964 *Phys. Rev.* B **136** 864
[23] Kohn W and Sham L J 1965 *Phys. Rev.* A **140** 1133
[24] Becke A D 1986 *J. Chem. Phys.* **84** 4524
Perdew J P 1986 *Phys. Rev.* B **33** 8822
[25] Troullier N and Martins J L 1991 *Phys. Rev.* B **43** 1993
[26] Louie S G, Froyen S and Cohen M L 1982 *Phys. Rev.* B **26** 1738
[27] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
Galli G and Parrinello M 1991 *Computer Simulation in Materials Science* ed V Pontikis and M Meyer (Dordrecht: Kluwer) and references therein
[28] CPMD 3.0h program, written by J Hutter (Max-Planck-Institut für Festkörperforschung, Stuttgart, 1998
[29] See for instance
Röthlisberger U, Carloni P, Doclo K and Parrinello M 2000 *J. Biol. Inorg. Chem.* **5** 236
Molteni C, Frank I and Parrinello M 1999 *J. Am. Chem. Soc.* **121** 12177
Alber F, Kuonen O, Scapozza L, Folkers G and Carloni P 1998 *Prot. Struct. Funct. Gen.* **31** 453
Carloni P, Sprik M and Andreoni W 2000 *J. Phys. Chem.* B **104** 823
Piana S, Sebastiani D, Carloni P and Parrinello M 2001 *J. Am. Chem. Soc.* **123** 8730
Sagnella D E, Laasonen K and Klein M L 1996 *Biophys. J.* **71** 1172
Hutter J, Carloni P and Parrinello M 1996 *J. Am. Chem. Soc.* **118** 7847
Rovira C, Carloni P and Parrinello M 1999 *J. Phys. Chem.* B **103** 7031
Rovira C, Kunc K and Parrinello M 2002 *Inorg. Chem.* at press
[30] Rovira C, Kunc K and Parrinello M 2001 *Inorg. Chem.* **40** 11
[31] Eichinger M, Tavan P, Hutter J and Parrinello M 1999 *J. Chem. Phys.* **110** 10452
[32] Eichinger M, Grubmüller H, Heller H and Tavan P 1997 *J. Comput. Chem.* **118** 1729
[33] Brooks B R *et al* 1983 *J. Comput. Chem.* **4** 187
[34] Albright T A, Burdett J K and Whangbo M-H 1985 *Orbital Interactions in Chemistry* (New York: Wiley)
[35] Rovira C and Parrinello M 2000 *Int. J. Quantum Chem.* **80** 1172
[36] Rovira C and Parrinello M 1999 *Chem. Eur. J.* **5** 250
[37] Schulze B and Evanseck J D 1999 *J. Am. Chem. Soc.* **121** 6444
[38] Unno M, Christian J F, Olson J S, Sage J T and Champion P M 1998 *J. Am. Chem. Soc.* **120** 2670
[39] Philips S E V and Schoenborn B P 1981 *Nature* **292** 81
[40] Peng S-M and J A Ibers 1976 *J. Am. Chem. Soc.* **98** 8032
[41] Sigfridson E and Ryde U 1999 *J. Biol. Inorg. Chem.* **4** 99
Scherlis D and Estrin D 2001 *J. Am. Chem. Soc.* **123** 8436