Published in final edited form as:

J Inorg Biochem. 2013 January ; 118C: 28–38. doi:10.1016/j.jinorgbio.2012.09.011.

Factors That Distort the Heme Structure in *Heme-Nitric Oxide/Oxygen-Binding (H-NOX) Protein Domains. A Theoretical Study*

Meng-Sheng Liao, Ming-Ju Huang, and John D. Watts*

Department of Chemistry, Jackson State University, Jackson, Mississippi 39217, USA

Abstract

DFT and dispersion-corrected DFT calculations were carried out to probe the factors that distort the heme structure in *Heme-Nitric oxide/Oxygen-binding (H-NOX) protein domains*. Various model systems that include heme, heme + surrounding residues, and heme + surrounding residues + additional protein environment were examined; the latter system was calculated with a quantum mechanics/molecular mechanics (QM/MM) method. The computations were extended to a myoglobin (Mb) protein, in which the heme structure is quite planar, in contrast to that in H-NOX. The natural tendency of the heme is to be planar. The strong structural distortion in H-NOX is mainly brought about by the intermolecular interactions between the whole heme molecule (heme ring plus its peripheral substituents) and the surrounding residues, among which the polar residues (Tyr140, Pro115, Met98) play major roles in distorting the heme structure. The two peripheral propionate substituents that are oriented on the same side of the heme plane can also make the molecule distort, but the distortion caused by this factor is not significant. In Mb, the surrounding residues considered are all nonpolar and do not cause a structural distortion. The different structural features of the heme macrocycle in the different proteins (H-NOX and Mb) are reproduced by the calculations. The dispersion correction is necessary, since it improves the calculated structures. The effects of the distortion on the binding affinity of the axial ligand to the heme were also examined.

Keywords

Heme; Structure; Distortion; H-NOX protein; Myoglobin; DFT calculations

1. Introduction

Hemoproteins play a key role in living organisms by transporting oxygen and other small molecules [1]. The same heme group, FePPIX(His) – iron protoporphyrin IX ligated with histidine, is present in these proteins but serves different purposes in each of them; in one it serves as oxygen carrier (hemoglobin, myoglobins), while in another it helps in oxygen metabolism (oxidases, peroxidases, catalases, and hydroxylases) and in electron transfer reactions (cytochromes). The presence of the protein environment is vital for these different functions, but is not fully understood. More recently, new functions of hemoproteins, e.g. sensing gas molecules and regulating transcription, have been identified [2–4]. Soluble

© 2012 Elsevier Inc. All rights reserved.

*Corresponding author (john.d.watts@jsums.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

guanylate cyclase (sGC) is a nitric oxide- (NO-) sensing hemoprotein, catalyzing the conversion of GTP to cGMP and thereby playing an important role in cGMP-mediated signaling pathways [5]. A most remarkable feature of sGC is that it selectively binds NO (and CO) but not O₂, in contrast to Mb (myoglobin) and Hb (hemoglobin). This ligand selectivity is very important for NO-specific signaling as sGC functions. Since there has been no structure determination for sGC so far, the reason why sGC achieves this “reverse” discrimination among the diatomic ligands is elusive.

Recently, a number of bacterial proteins (named H-NOX for *H*eme-*N*itric oxide/*O*Xygen-binding) with sequence homology to sGC have been identified and characterized [6–12]. Some of these bacterial proteins from eukaryotes and facultative aerobic prokaryotes also exclude O₂ while binding NO selectively, the same characteristic as sGC. These proteins may be called NO-selective H-NOXs. On the other hand, certain H-NOX proteins found in obligate anaerobic bacteria (prokaryotes) reversibly bind O₂ as well as NO and CO (similar to Mb/Hb). The latter proteins may be called O₂/NO binding H-NOXs. Marletta and co-workers [8] have recently reported the crystal structure of an O₂/NO binding H-NOX from the anaerobe *T. tengcongensis* (*Tt* H-NOX). A similar protein with structure determination was reported by Raman and co-workers [7]. One of the significant features of the O₂-bound structure of *Tt* H-NOX is the presence of a distal pocket hydrogen-bonding network, which includes Tyr140, Asn74, and Trp9. But these residues are absent in all of the domains of the studied NO-selective H-NOX. Since Tyr140 is involved in a short (2.74 Å) hydrogen bond to the bound O₂, this residue was suggested to be the key molecular factor for ligand discrimination in the H-NOX family [10].

Another notable feature of H-NOX(O₂) is the highly distorted structure of the heme skeleton in this kind of protein. This is in contrast to the heme in a myoglobin (MbO₂) [13], which possesses a nearly planar aromatic unit. Figure 1 illustrates the structure and residues surrounding the heme in an O₂ binding H-NOX domain [8] and in a wild-type MbO₂ [13]. The distortion from planarity may be traced to various sources, e.g. bulky axial ligands or peripheral substituents. At first we see in Figure 1 that the orientation of one propionate (-C₂H₄COOH) substituent at the periphery of the heme ring in H-NOX differs from that in Mb. In the former, the two propionates are located on the same side of the heme plane, while in the latter, the two propionates lie above and below the heme plane, respectively. The orientations of the propionates in each specific protein are actually relevant to the distributions of the residues surrounding the heme. In H-NOX, the Trp9 residue occupies a position that prevents one propionate from being located above the heme plane. It is the other way round for Mb, in which the His97 residue prevents one propionate group from locating below the heme plane.

Nonplanar conformations were suggested to be essential for biological functions and have been the subject of some studies [14–16]. As the first step in our theoretical studies of H-NOX proteins, we explore here factors that distort the heme structure in H-NOX protein domains. DFT (density functional theory) will be used in the calculations. It is desirable to use DFT for the efficient study of large biomolecules. However, the current DFT functionals are thought to be incapable of describing long-range nonlocal interactions such as dispersion properly [17–20]. Since noncovalent interactions between the heme moiety and the surrounding residues may play an important role in these biomolecules, another set of DFT calculations with a (semi-)empirical dispersion correction will also be performed. For the sake of comparison, the calculations will be extended to a myoglobin (Mb) protein, in which the heme structure is quite planar, in contrast to that in H-NOX.

2. Computational Details

2.1. Models

The systems that are used in the calculations are illustrated in Figure 2. The initial structure (Figure 2a) is based on the X-ray crystal data of H-NOX(O₂) (*code* 1U4H) [8]. From Figure 1a, there are eight (8) closest residues surrounding the heme active center; they are tryptophan 9 (Trp9), isoleucine 5 (Ile5), tyrosine 140 (Tyr140), asparagine 74 (Asn74), leucine 144 (Leu144), phenylalanine 78 (Phe78), proline 115 (Pro115), and selenomethionine 98 (Mse98), which have atoms falling within a ca. 10.0-Å radius of the Fe atom. In our actual calculations, the residues are somewhat simplified, as shown in Figure 2e. For example, the proximal histidine 102 (His102) is modeled as a 4-ethylimidazole (4-EtIm), which has been shown to be a well-simplified model for the histidine residue attached to FePPiX [21]. Furthermore, the axial ligand O₂ in Heme(O₂) is replaced with CO for computational convenience. Heme(CO) is closed-shell singlet, while Heme(O₂) has an open-shell ground state. A calculation on a large closed-shell system is much easier than that on the corresponding open-shell system, particularly in geometry optimizations. Unfortunately, no crystal structure of a *native* H-NOX(CO) protein is available. Comparing an MbO₂ (*code* 1A6M) [13] with an MbCO (*code* 1BZR) [22] (Section 3.5), the structure of the O₂ complex is in fact very similar to that of the CO complex (except for the Fe-O-O angle, which is ca. 120°). We believe that the structure of H-NOX(CO) is similar to the structure of H-NOX(O₂) too. Since peripheral substituents in the heme ring probably influence the structure of the heme, they are all included in the calculations here. Finally the systems are protonated (since no positions of hydrogen atoms are available in an X-ray crystal structure); the H atoms are omitted in the figures for clarity.

As mentioned in the Introduction, the aim of this work is to explore factors that distort the heme structure in the H-NOX protein domains. To do so, various systems have been examined; they include Heme(CO) (no residues surrounding the heme - Figure 2a), Heme(CO)+6Rs [Heme(CO) plus six distal residues above the heme plane - Figure 2c], Heme(CO)+2Rs [Heme(CO) plus two proximal residues below the heme plane - Figure 2d], Heme(CO)+8Rs [Heme(CO) plus all the eight residues surrounding the heme - Figure 2e], and Heme(CO)+8Rs + additional protein environment (APE). Heme(CO)+8Rs+APE represents the whole protein system; it was calculated using a quantum mechanics/molecular mechanics (QM/MM) method, where Heme(CO)+8Rs was treated at the QM level, while APE was accounted for at the MM level.

2.2. Computational Method

The calculations were carried out with the Amsterdam Density Functional (ADF) program package – ADF2010.02 [23,24]. The density functional used is BP, which contains Becke's 1988 gradient correction for exchange [25] plus Perdew's 1986 gradient correction for correlation [26]. This functional has been shown to give an excellent description of molecular structure in previous calculations [21]. As mentioned above, the current DFT functionals fail to describe dispersion interactions properly. In order to correct for this deficiency, one practical method is to add a (semi-)empirical correction of the form C_6R^{-6} to a density functional scheme to yield a DFT + E_{disp} model, in which the dispersion energy is calculated separately from the DFT calculation and simply added to the DFT energy [20], namely

$$E_{\text{tot}} = E_{\text{DFT}} + E_{\text{disp}}$$

Here E_{disp} is a sum of damped interatomic potentials that have the form C_6R^{-6} at long range.

$$E_{\text{disp}} = - \sum_n \sum_{i < j} s_n \frac{C_n^{ij}}{R_{ij}^n} f_{d,n}(R_{ij}), n=6, 8$$

where s_n is a global scaling factor, C_n^{ij} is the dispersion coefficient for atom pair ij and calculated from individual atomic C_n^i and C_n^j coefficients through a combination rule, and R_{ij} stands for an interatomic distance. The term $f_{d,n}(R_{ij})$ is a damping function, which is used to damp the R^{-n} term at short distances and to reduce the correlation effect on covalent bonds.

A large number of recent calculations (e.g. refs. [27–36]) performed with the DFT + E_{disp} technique for weakly bonded systems have been very promising; the new method is expected to yield improved quantitative results. The present ADF program includes three versions of Grimme's dispersion correction (DFT + E_{disp}) methods, labeled [37] as DFT-D1, -D2, and -D3. The -D3 version was developed very recently and contains many modifications and new features; it is claimed to have the advantages of higher accuracy, broader range of applicability, and less empiricism. Thus, we adopted the latest version DFT-D3 [37] in the present calculations.

The STO (Slater-type orbital) basis set used is the standard ADF-TZP, which is a triple- ζ plus one polarization function set. Frozen-core techniques [23] were used here to reduce the computational cost. To obtain accurate results, the valence set on Fe included sub-valence 3s and 3p shells. For Se, C, N, and O, ns and np were considered as valence shells. The other shells of lower energy, i.e. [Ne] for Fe, [Ar3d¹⁰] for Se, and [He] for C/N/O, were described as core and kept frozen.

Relativistic corrections for the valence electrons were calculated by the quasi-relativistic (QR) method [38].

In the QM/MM calculations, the initial structure of the system was taken from the Protein Data Bank (PDB *code* 1U4H). A utility program PDB2ADF [24] was used to transform a PDB file into an ADF input file for use with QM/MM calculations, where the protein system was divided into the QM region [Heme(CO)+8Rs] and MM region (APE). The MM calculations were performed with the AMBER95 force field [39]. The QM/MM covalent boundaries were capped by hydrogen link atoms and were treated by using the AddRemove coupling scheme developed by Swart [40]. The initial charges on each atom were assigned based on their type from the force field file. The present ADF program supports only the mechanical embedding (ME) scheme, which treats the electrostatic interaction between the QM and MM regions at the MM level. The solvating water molecules were not included in the calculations.

3. Results and Discussion

3.1. Heme(CO)

We start from the Heme(CO) moiety without the surrounding residues. The initial structure is based on the experimental data (*code* 1U4H) [8] and is seen to be very distorted (Figure 2a). However, the 4 equatorial, porphinato nitrogens almost form a plane, which is labeled here as the 4N plane; it stands for the mean plane defined by the 4 porphinato nitrogen atoms. The iron atom is essentially coplanar with the 4N plane. To give a quantitative description of the structural distortion of the molecule, we present the (vertical) distances

$R_{\text{plane}(4N)\cdots C_n}$ between the 4N plane and each of the peripheral carbon atoms (C_n). The C atoms with $n = 1 - 8$ are the peripheral β carbons, while those with $n = 9 - 12$ are the C atoms at the *meso* position (see Figure 1c). A positive value for $R_{\text{plane}(4N)\cdots C_n}$ means that the C atom lies above the 4N plane and a negative $R_{\text{plane}(4N)\cdots C_n}$ value indicates that the C atom lies below the 4N plane. The most positive and most negative values are indicated in bold in the tables; they represent the top and bottom C atoms respectively. The distance between the top and bottom C atoms [$\Delta(C_{\text{top}}-C_{\text{bot}}) = R_{\text{plane}(4N)\cdots C_{\text{top}}} - R_{\text{plane}(4N)\cdots C_{\text{bottom}}}$] in the vertical direction is taken as a measure of the structural distortion. We also calculated *rms* (root mean square) deviations from planarity for the C atoms at the β and *meso* positions, respectively. Since the *rms* values are relatively small [compared with $\Delta(C_{\text{top}}-C_{\text{bot}})$] and the change in the *rms* value from one system to another is not so obvious, $\Delta(C_{\text{top}}-C_{\text{bot}})$ probably gives a more realistic measure of the structural distortion. The experimental and optimized structural parameters for Heme(CO) are reported in Table 1. They include $R_{\text{Fe-N}(\text{eq})}$ (distance between Fe and the equatorial, porphinato nitrogen), $R_{\text{Ct}(4N)\cdots \text{Fe}}$ (distance between the center of the ring and Fe), $R_{\text{Fe-N}(\text{ax})}$ (axial Fe-N(ax) bond length), $R_{\text{Fe-C}(\text{CO})}$ (axial Fe-C(CO) bond length), $R_{\text{C-O}}$ (C-O bond length), $\angle \text{Fe-C-O}$ (Fe-C-O bond angle), as well as $R_{\text{plane}(4N)\cdots C_n}$. Since no experimental structural data are available that are related to the CO part, the experimental $R_{\text{Fe-C}(\text{CO})}$, $R_{\text{C-O}}$, and $\angle \text{Fe-C-O}$ values are set to estimated values on the basis of some relevant data in the literature. $R_{\text{Ct}(4N)\cdots \text{Fe}}$ in six-coordinate compounds is usually small, but not zero experimentally and may be different in different compounds. Hence, we have also presented the calculated $R_{\text{Ct}(4N)\cdots \text{Fe}}$ values for each system here.

First looking at the Fe-N(eq) and Fe-N(ax) bond lengths, as well as the $\text{Ct}(4N)\cdots \text{Fe}$ distance, there is good quantitative agreement between the calculation and experiment. The dispersion correction results in a shortening of axial $R_{\text{Fe-N}(\text{ax})}$ by ca. 0.04 Å and the DFT-D3 calculated $R_{\text{Ct}(4N)\cdots \text{Fe}}$ is somewhat more negative than that calculated with pure DFT. A negative $R_{\text{Ct}(4N)\cdots \text{Fe}}$ value denotes an out-of-plane displacement of Fe toward the CO ligand. The other structural parameters ($R_{\text{Fe-N}(\text{eq})}$, $R_{\text{Fe-C}(\text{CO})}$, $R_{\text{C-O}}$, $\angle \text{Fe-C-O}$) are affected little by the dispersion correction.

Examining the $R_{\text{plane}(4N)\cdots C_n}$ values, the experimental distances between the top and bottom C atoms, $\Delta(C_{\text{top}}-C_{\text{bot}})$, are as large as 1.26 Å for the β C atoms and 1.12 Å for the *meso* C atoms, indicating very large deviations of the C atoms from planarity. The major contributions to the deformation of the H-NOX domain heme group can be attributed to saddling, which involves the bending of pyrrole rings with respect to each other, and ruffling, described as a rotation of the pyrrole rings about the iron-nitrogen bonds. These two forms of deformation are illustrated in Figure 3. Upon geometry optimization with pure DFT, the heme macrocycle becomes quite planar, as is seen in Figure 2b. The $\Delta(C_{\text{top}}-C_{\text{bot}})$ values decrease to 0.34 and 0.25 Å respectively. But the DFT-D3 calculation gives slightly larger $\Delta(C_{\text{top}}-C_{\text{bot}})$ values than the DFT ones, indicating that the dispersion effects within the same heme group distort the structure slightly.

To further investigate effects of the peripheral substituents, additional calculations have been performed on FeP(His)(CO), where P is the simple porphine without any peripheral substituents. The initial structure for the geometry optimization on this complex is also based on the very distorted, experimental structure of the protein. All the calculated $R_{\text{plane}(4N)\cdots C_n}$ values turn out to be less than ± 0.1 Å even when the dispersion correction is made (see Table 1), which indicates that FeP in FeP(His)(CO) is almost planar after the geometry optimization, and the axial ligands do not cause structural distortion for the heme macrocycle ring. Simple metal porphines (MPs) including FeP are well known to be square planar, as established by experiment [41]. The calculated results are consistent with experiment.

3.2. Heme(CO)+6Rs

We now put the 6 distal residues above the DFT optimized Heme(CO) structure and optimize the structure of Heme(CO)+6Rs (see Figure 2c), where the initial positions of the residues are also based on the X-ray crystal determination [8] (the same is true for Heme(CO)+2Rs - Section 3.3 and Heme(CO)+8Rs - Section 3.4). A comparison between the experimental and optimized structural parameters is given in Table 2. For the external structure of Heme(CO), we give the distances between the Fe atom and the nearest neighbor C or N atoms from each residue. While the DFT optimized bond lengths are quite close to the experimental data within the internal structure of the heme, the situation is different for the external structure of the heme in the protein, where interactions between the heme and the surrounding residues are noncovalent. It is shown that the calculations with pure DFT greatly overestimate the intermolecular distances in the complex, thereby underestimating the interactions between the heme and the surrounding residues. According to the calculated $\Delta(C_{top}-C_{bot})$ values with DFT, the structure of the heme is only slightly more distorted in Heme(CO)+6Rs than in Heme(CO).

Since the dispersion interaction is attractive, the DFT-D3 method gives considerably shorter distances between the Fe atom and the residues than the DFT one. There is good agreement between the DFT-D3 calculated and experimental $R_{Fe \cdots Residue}$ values for most residues. Asn74 appears to be an exception, for which too short an $Fe \cdots Residue$ distance is obtained when the dispersion correction is made. Comparing Heme(CO)+6Rs to Heme(CO), the DFT-D3 calculations show that the $\Delta(C_{top}-C_{bot})$ values increase from 0.44 Å to 0.71 Å for the β C atoms and from 0.43 Å to 0.81 Å for the *meso* C atoms, which indicates that the structure of the heme is significantly more distorted in Heme(CO)+6Rs than in Heme(CO).

3.3. Heme(CO)+2Rs

We now consider another situation: Heme(CO)+2Rs, i.e. Heme(CO) (optimized with DFT) plus the two residues that lie below the heme plane. When the structure is optimized with pure DFT, the $Fe \cdots Residue$ distances are again overestimated greatly (see Table 2). Good agreement between the calculated and experimental $Fe \cdots Residue$ distances is obtained with the DFT-D3 method. According the calculations, the distortion of the heme structure is somewhat larger in Heme(CO)+2Rs than in Heme(CO)+6Rs, which indicates that the Mes98 and Pro115 residues are other important factors that distort the heme structure. The DFT-D3 calculated $\Delta(C_{top}-C_{bot})$ values in Heme(CO)+2Rs are as large as 0.89 Å for the β C atoms and 0.92 Å for the *meso* C atoms.

3.4. Heme(CO)+8Rs

Then we put all the 8 residues around the quite planar, DFT optimized Heme(CO). The results are presented in Table 2 as well. Although the DFT optimized $Fe \cdots Residue$ distances in Heme(CO)+8Rs are still significantly larger than the experimental ones, they are shorter than the corresponding values in Heme(CO)+6Rs or Heme(CO)+2Rs for most residues. That is, the interactions between the heme and the upper distal residues are enhanced by including the lower residues in the system, and vice versa for the interactions between the heme and the lower proximal residues. As a result, the DFT calculated $\Delta(C_{top}-C_{bot})$ values in Heme(CO)+8Rs are significantly larger than in Heme(CO)+6Rs and are also larger than in Heme(CO)+2Rs. Similarly, a significantly more distorted structure of heme is obtained when the dispersion correction is made. But we see that within the same DFT-D3 calculations, the structure of heme does not change much from Heme(CO)+2Rs to Heme(CO)+8Rs. The DFT-D3 calculated $\Delta(C_{top}-C_{bot})$ values are 0.96 Å for the β C atoms and 0.94 Å for the *meso* C atoms in Heme(CO)+8Rs, which are only slightly larger than the corresponding values in Heme(CO)+2Rs. A remarkable result here is that the conformation of the heme structure optimized with the surrounding residues is quite similar in shape to

that in the protein crystal structure. That is, the geometry optimization on Heme(CO)+8Rs also results in a combined saddling-ruffling structure for the heme molecule, though the initial structure of the Heme(CO) moiety in the geometry optimization is a 'planar' one. For several C atoms, there is even good quantitative agreement between the calculated and experimental $R_{\text{plane}(4N)\cdots Cn}$ values.

Section 3.1 has shown that isolated FeP(His)(CO) is planar. It would be of interest to see how the situation changes when we put residues around FeP(His)(CO). Two different initial structures of the FeP moiety were considered in the geometry optimizations. One is based on the very distorted, experimental structure of the protein; the other uses a completely planar FeP. Figure 4 illustrates the initial and optimized structures for FeP(His)(CO)+8Rs. The optimized structural parameters are presented in Table 3. It is shown that geometry optimizations with two different initial structures give final structures which are almost the same; they nearly coincide with each other. Since the final optimized structure does not depend on a choice of initial structure of the heme, the various structures optimized here can be considered to be reliable. The DFT calculated $\Delta(C_{\text{top}}-C_{\text{bot}})$ values in FeP(His)(CO)+8Rs are 0.56 Å for the β C atoms and 0.60 Å for the *meso* C atoms; they are not notably smaller than the corresponding values in Heme(CO)+8Rs (0.70 and 0.66 Å). When the dispersion correction is made, the calculated $\Delta(C_{\text{top}}-C_{\text{bot}})$ values in FeP(His)(CO)+8Rs are changed only slightly. This is different from the case of Heme(CO)+8Rs. In summary, a distortion of the heme surrounded by the residues would be smaller without peripheral substituents, indicating that the interaction between the peripheral substituents and the surrounding residues is one of factors that distort the heme structure.

3.5. Heme(CO)+8Rs+APE

For the quantum mechanical (QM) calculations on the above Heme(CO)+nRs systems, the residues above and below the heme plane are allowed to move freely in the geometry optimizations. Such an optimization procedure was justified in DFT calculations by De Angelis et al. [42] on a large Heme(AB) model system with 13 surrounding residues (AB = CO, O₂). In order to obtain more accurate results and to examine the effects of the additional protein environment (APE) on the structure of the heme, we performed quantum mechanics/molecular mechanics (QM/MM) calculations on the whole protein system Heme(CO)+8Rs+APE, where Heme(CO)+8Rs was treated at the QM level, while the APE was accounted for at the MM level. The initial structure of the Heme(CO) moiety in the geometry optimization is still the Heme(CO) structure optimized with DFT (i.e. it is planar), similar to the situations for the above Heme(CO)+nRs systems (for which the same planar, DFT optimized Heme(CO) structure was used as starting geometry). The DFT and DFT-D3 optimized structural parameters for Heme(CO)+8Rs+APE are presented in the last two columns of Table 2.

In the QM/MM calculations, the surrounding residues are then constrained to the protein backbone. In this case, the Fe $\cdots$ Residue distances optimized with pure DFT are less overestimated than those in Heme(CO)+8Rs. Correspondingly, the DFT calculated heme distortion in Heme(CO)+8Rs+APE is also more pronounced than in Heme(CO)+8Rs. The DFT results are improved by including the APE in the calculation through an MM method. Concerning the DFT-D3 results, there is no obvious trend in the Fe $\cdots$ Residue distances from Heme(CO)+8Rs to Heme(CO)+8Rs+APE; the DFT-D3 calculated $R_{\text{Fe}\cdots\text{Residue}}$ values in Heme(CO)+8Rs+APE may be larger (for Trp9 and Asn74) or smaller (for Leu144 and Pro115) than in Heme(CO)+8Rs; they change little for the other four residues. The DFT-D3 calculated $\Delta(C_{\text{top}}-C_{\text{bot}})$ value is larger in Heme(CO)+8Rs+APE (and in excellent agreement with experiment) than in Heme(CO)+8Rs for the β carbons, but the opposite is the case for the *meso* carbons. In summary, the DFT-D3 results with and without including APE can be considered to be comparable.

Figure 5 illustrates diagrams of overlay between the experimental (X-ray crystal) structure and optimized structures for the Heme(CO) moiety with and without including the surrounding residues in the calculation. In Figure 5a, the optimized structure (white line) of Heme(CO) without including residues is seen to be planar, deviating greatly from the experimental, distorted structure (black line). When the surrounding residues are included in the optimization, the calculated skeleton of the heme almost coincides with the experimental one (Figure 5b).

3.6. Heme(CO), Heme(CO)+8Rs, and Heme(CO)+8Rs+APE from Mb

For comparison, additional calculations have been carried out for Heme(CO), Heme(CO)+8Rs, and Heme(CO)+8Rs+APE that are from a myoglobin MbO₂ (*code* 1A6M) [corresponding H-NOX(O₂)], where the axial O₂ attached to the heme is replaced by CO for computational convenience. In fact, the crystal structure of *native* MbCO is also available (*code* 1BZR) [22], but does not show an obvious difference from that of MbO₂ (except for the Fe-O-O angle). Table 4 displays the experimental structural parameters in both MbO₂ and MbCO, which show that the structure of the O₂ complex is very similar to that of the CO complex, as mentioned in Section 2.1. As seen in Figure 1b, the heme macrocycle ring in MbAB (AB = O₂ or CO) is quite planar, in contrast to that in H-NOX(O₂) (Figure 1a). There are also eight residues surrounding the heme in Mb, but five are above the heme plane (histidine 64, valine 68, phenylalanine 43, isoleucine 107, leucine 29) and three lie below (histidine 97, leucine 89, isoleucine 99). The optimized structural parameters for Heme(CO), Heme(CO)+8Rs, and Heme(CO)+8Rs+APE from Mb are presented in Table 4, together with experimental data for comparison.

Examining the experimental $R_{\text{plane}(4N)\cdots C_n}$ values, two C atoms (C1 and C2, see Figure 1c) turn upwards rather greatly above the heme plane (this can be seen in Figure 1b); the deviations of the other C atoms from planarity are small. After the geometry of Heme(CO) is optimized, the upturning of these two C atoms is considerably reduced, and all the C atoms form a quite planar structure. The DFT and DFT-D3 calculated $\Delta(C_{\text{top}}-C_{\text{bot}})$ values here in Heme(CO) are 0.12 Å (for both β and *meso* C atoms) and 0.23 – 0.24 Å (for both β and *meso* C atoms), respectively, notably smaller than those in Heme(CO) from H-NOX. This indicates that without surrounding residues, the heme from Mb is somewhat more "planar" than that from H-NOX. As mentioned in Section 1, the two propionate groups at the periphery of the heme ring lie in the same direction in H-NOX, but in different directions in Mb. The calculations seem to indicate that propionate substituents in the same direction could distort the otherwise planar heme structure somewhat. In Mb, the two propionates lie above and below the heme plane respectively and so their effects on the structure conformation cancel each other and do not cause a structural distortion.

By putting the eight residues around the heme, a planar structure for the heme remains upon geometry optimization with DFT, and no obvious changes in the C-atom positions are found after the dispersion effects are taken into account. The QM/MM calculations on the whole protein system Heme(CO)+8Rs+APE do not notably change the $\Delta(C_{\text{top}}-C_{\text{bot}})$ values for both β and *meso* carbons. That is, the heme in Heme(CO)+8Rs+APE is still planar. The DFT-D3 calculated $R_{\text{plane}(4N)\cdots C1}$ and $R_{\text{plane}(4N)\cdots C2}$ in Heme(CO)+8Rs are ca. 0.13 Å, still significantly smaller than the experimental values (0.3 – 0.4 Å). The upturning of C1 and C2 from the heme plane in actual Mb probably results from the intermolecular interaction between the heme and the residue tyrosine 103 (Tyr103) (which is not displayed in Figure 1b). Tyr103 is a polar residue, lying sideways below the heme plane, the carbonyl group being relatively close to C1 and C2 and their bonded substituents (vinyl, methyl) [$R_{\text{Fe}\cdots\text{O}(\text{Tyr103})} = 9.04$ Å, $R_{\text{C1}\cdots\text{O}(\text{Tyr103})} = 5.01$ Å, $R_{\text{C2}\cdots\text{O}(\text{Tyr103})} = 5.90$ Å]. Examining the individual residues surrounding the heme in both H-NOX and Mb, we can say that the polar residues play a major role in distorting the heme structure. [Note that Tyr103 is treated only

at MM level in Heme(CO)+8Rs+APE.] The eight residues that surround the heme in Mb are exclusively nonpolar; none of them has strong effects distorting the heme structure. Figure 6 illustrates the experimental and optimized structures for Heme(CO)+8Rs from Mb, whose structural feature is also reproduced by the calculations.

3.7. Heme-CO Binding Energies

We also investigated the effects of the distortion on the binding affinity of the CO ligand to the heme. In doing so, we carried out another type of calculation by keeping the heme skeleton (which contains 4 N-atoms and 20 C-atoms) of the optimized Heme(CO) planar. This can be done with a constrained optimization method in ADF. Table 5 reports the calculated binding energies (E_{bind}) of CO to heme in different systems, together with available experimental data [43,44]. E_{bind} is defined as

$$-E_{\text{bind}}(\text{Heme-CO}) = E[\text{Heme(CO)+nRs}] - \{E(\text{Heme+nRs}) + E(\text{CO})\} \quad n=0, 8$$

Here $E[\text{Heme(CO)+nRs}]$, $E(\text{Heme+nRs})$, and $E(\text{CO})$ are the total energies of the indicated species, which are optimized independently.

The DFT calculated Heme-CO binding energy in Heme(CO) is 1.52 eV (1 eV = 23.06 kcal/mol = 96.5 kJ/mol), which is significantly larger than the experimental value (0.85 eV). Indeed the BP functional greatly overbinds for the CO ligand, but it gives an excellent description of molecular structure [21], as mentioned in Section 2.2. Similar results for E_{bind} were obtained with BP from calculations on FeP(His)-CO in the literature (1.52 eV from ref. [45] and 1.57 eV from ref. [46]). The dispersion contribution to $E_{\text{bind}}(\text{Heme-CO})$ is 0.14 eV. When the surrounding residues are included in the calculation, there are some changes in E_{bind} . The DFT calculated E_{bind} in Heme(CO)+8Rs becomes 1.56 eV, which is 0.04 eV larger than that in Heme(CO). In contrast, the DFT-D3 value of E_{bind} in Heme(CO)+8Rs is smaller by 0.07 eV as compared to that in Heme(CO). We should point out that the dispersion interaction between bound CO and the surrounding residues is always attractive. But according to the above definition for E_{bind} , the calculated 8Rs+Heme-CO binding energy is also related to the energy of *optimized* 8Rs+Heme. With the dispersion correction being made, the non-covalent interaction between the protein environment and the heme is enhanced when CO departs.

By keeping the skeleton of the heme planar in Heme(CO)+8Rs, the calculated Heme-CO binding energy would be 0.05 – 0.06 eV larger, indicating that the distortion decreases the axial ligand affinity to the heme.

We also evaluated the Heme-CO binding energy in Heme(CO)+8Rs from Mb. It is shown to be somewhat smaller (by 0.03 – 0.04 eV) than that in Heme(CO) from both DFT and DFT-D3 calculations. The calculated trend is in good agreement with the experimental one (0.05 eV). According to the DFT-D3 calculations, E_{bind} in Heme(CO)+8Rs from Mb is similar to that in Heme(CO)+8Rs from H-NOX, where planar heme is kept. Note that the heme in Heme(CO)+8Rs from Mb is planar.

4. Conclusions

The structure of the heme group in an H-NOX protein is highly distorted from planarity. The structural distortion is not an intrinsic property of the isolated molecule. Several factors are found to be responsible for the strong structural distortion in H-NOX. They include (1) the two peripheral propionate substituents that are oriented on the same side of the heme plane, which distort the structure somewhat but not significantly, (2) the interaction between the

heme and the upper distal residues, which distort the structure rather strongly, and (3) the interaction between the heme and the proximal residues below the heme plane (Pro115 and Mse98), which distort the structure more strongly than the upper distal residues. If we divide the whole heme molecule into two parts, namely the heme ring and its peripheral substituents, the distortion of the heme surrounded by the residues would be smaller without peripheral substituents. Therefore, the interaction between the peripheral substituents and the surrounding residues is a part of the contribution to the structural distortion. Among the various surrounding residues, the polar ones (Tyr140, Pro115, Mse98) play major roles in distorting the heme structure in H-NOX. In the case of Mb, the surrounding residues considered are all nonpolar and do not cause a structural distortion. The propionate substituents in the heme ring in Mb lie above and below the heme plane and do not make the heme distort as well. The different conformations of the heme structure in the different proteins (H-NOX and Mb) are well reproduced by the calculations.

The DFT results are improved by including the APE in the calculation through an MM method. This is because without constraints on the residues in Heme(CO)+8Rs in the geometry optimization, the DFT optimized heme...residue distances are greatly overestimated. When the dispersion correction is made in Heme(CO)+8Rs, there is a good agreement between the optimized and experimental structures, and the DFT-D3 results on Heme(CO)+8Rs without including APE are adequate and satisfactory.

We also investigated the effects of the distortion on the binding affinity of the axial ligand to the heme. The Heme-CO binding energy is smaller for the distorted macrocycle than for the planar one, indicating that the distortion decreases the axial ligand affinity to the heme.

Acknowledgments

This work was supported by Award Number SC1-HL096018 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health (NIH). The ADF calculations were run on a QuantumCube™ QS32-2800C computer from Parallel Quantum Solutions, LLC.

References

1. Antonini, E.; Rossi-Bernardi, L.; Chiancone, E. Hemoglobins. New York: Academic Press; 1981.
2. Lawson DM, Stevenson CE, Andrew CR, Eady RR. The EMBO J. 2000; 19:5661–5671.
3. Hao B, Isaza C, Arndt J, Soltis M, Chan MK. Biochemistry. 2002; 41:12952–12958. [PubMed: 12390021]
4. Karow DS, Pan D, Tran R, Pellicena P, Presley A, Mathies RA, Marletta MA. Biochemistry. 2004; 43:10203–10211. [PubMed: 15287748]
5. Toda N, Okamura T. Pharmacol. Rev. 2003; 55:271–324. [PubMed: 12773630]
6. Friedman J, Meharena YT, Wilks A, Poulos TL. J. Biol. Chem. 2007; 282:1066–1071. [PubMed: 17095508]
7. Nioche P, Berka V, Vipond J, Minton N, Tsai A-L, Raman CS. Science. 2004; 306:1550–1553. [PubMed: 15472039]
8. Pellicena P, Karow DS, Boon EM, Marletta MA, Kuriyan J. Proc. Natl. Acad. Sci. USA. 2004; 101:12854–12859. [PubMed: 15326296]
9. Boon EM, Marletta MA. Curr. Opin. Chem. Biol. 2005; 9:441–446. [PubMed: 16125437]
10. Boon EM, Marletta MA. J. Inorg. Biochem. 2005; 99:892–902. [PubMed: 15811506]
11. Boon EM, Huang SH, Marletta MA. Nature Chem. Biol. 2005; 1:53–59. [PubMed: 16407994]
12. Boon EM, Davis JH, Tran R, Karow DS, Huang SH, Pan D-H, Miazgowiec MM, Mathies RA, Marletta MA. J. Biol. Chem. 2006; 281:21892–21902. [PubMed: 16728401]
13. Vojtechovsky J, Chu K, Berendzen J, Sweet RM, Schlichting I. Biophys. J. 1999; 77:2153–2174. [PubMed: 10512835]

14. Shelnutt JA, Song X-J, Ma J-G, Jia SL, Jentzen W, Medforth CJ. Chem. Soc. Rev. 1998; 27:31–41.
15. Ravikanth M, Chandrashekar TK. Struct. Bonding. 1995; 82:105–188.
16. Haddad RE, Gazeau S, Pecaut J, Marchon J-C, Medforth CJ, Shelnutt JA. J. Am. Chem. Soc. 2003; 125:1253–1268. [PubMed: 12553827]
17. Jung Y, Head-Gordon M. Phys. Chem. Chem. Phys. 2006; 8:2831–2840. [PubMed: 16775638]
18. Sinnokrot MO, Sherrill CD. J. Phys. Chem. A. 2006; 110:10656–10668. [PubMed: 16970354]
19. Hesselmann A, Jansen G, Schütz M. J. Chem. Phys. 2005; 122 014103/1-17.
20. Grimme S. J. Comput. Chem. 2004; 25:1463–1473. [PubMed: 15224390]
21. Liao M-S, Huang M-J, Watts JD. J. Phys. Chem. A. 2010; 114:9554–9569. [PubMed: 20712371]
22. Kachalova GS, Popov AN, Bartunik HD. Science. 1999; 284:473–476. [PubMed: 10205052]
23. te Velde G, Bickelhaupt FM, van Gisbergen SJA, Fonseca-Guerra C, Baerends EJ, Snijders JG, Ziegler T. J. Comput. Chem. 2001; 22:931–967.
24. ADF2010.01, SCM, Theoretical Chemistry. Amsterdam, The Netherlands: Vrije Universiteit; 2010. see also <<http://www.scm.com>>.
25. Becke AD. Phys. Rev. A. 1988; 38:3098–3100. [PubMed: 9900728]
26. Perdew JP. Phys. Rev. B. 1986; 33:8822–8824.
27. Du Q-S, Liu P-J, Deng J. J. Chem. Theory Comput. 2007; 3:1665–1672.
28. Morgado CA, McNamara JP, Hillier IH, Burton NA, Vincent MA. J. Chem. Theory Comput. 2007; 3:1656–1664.
29. Jurecka P, Cerny J, Hobza P, Salahub DR. J. Comput. Chem. 2007; 28:555–569. [PubMed: 17186489]
30. Wu Q, Yang W-T. J. Chem. Phys. 2002; 116:515–524.
31. Zimmerli U, Parrinello M, Koumoutsakos P. J. Chem. Phys. 2004; 120:2693–2699. [PubMed: 15268413]
32. Elstner M, Hobza P, Suhai S, Kaxiras E. J. Chem. Phys. 2001; 114:5149–5155.
33. Ducere J-M, Cavallo L. J. Phys. Chem. B. 2007; 111:13124–13134. [PubMed: 17949028]
34. Thanthiriwatte KS, Hohenstein EG, Burns LA, Sherrill CD. J. Chem. Theory Comput. 2011; 7:88–96.
35. Marom N, Tkatchenko A, Scheffler M, Kronik L. J. Chem. Theory Comput. 2010; 6:81–90.
36. Barone V, Biczysko M, Pavone M. Chem. Phys. 2008; 346:247–256.
37. Grimme S, Antony J, Ehrlich S, Krieg H. J. Chem. Phys. 2010; 132 154104/1-19.
38. Ziegler T, Tschinke V, Baerends EJ, Snijders JG, Ravenek W. J. Phys. Chem. 1989; 93:3050–3056.
39. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. J. Am. Chem. Soc. 1995; 117:5179–5197.
40. Swart M. Int. J. Quantum Chem. 2003; 91:177–183.
41. Rosa A, Ricciardi G, Baerends EJ, van Gisbergen SJA. J. Phys. Chem. A. 2001; 105:3311–3327.
42. De Angelis F, Jarzecki AA, Car R, Spiro TG. J. Phys. Chem. B. 2005; 109:3065–3070. [PubMed: 16851321]
43. Radon M, Pierloot K. J. Phys. Chem. A. 2008; 112:11824–11832. [PubMed: 18942804]
44. Blomberg LM, Blomberg MRA, Siegbahn PEM. J. Inorg. Biochem. 2005; 99:949–958. [PubMed: 15811512]
45. Rovira C, Kunc K, Hutter J, Ballone P, Parrinello M. J. Phys. Chem. A. 1997; 101:8914–8925.
46. Strickland N, Harvey JN. J. Phys. Chem. B. 2007; 111:841–852. [PubMed: 17249828]

Highlights

- The natural tendency of the heme is to be planar.
- Several factors are responsible for the strong structural distortion in H-NOX.
- The polar residues play major roles in distorting the heme structure.
- The dispersion correction greatly improves the calculated structures.
- The different heme conformations in H-NOX and Mb are reproduced by the calculation.

Figure 1.

The structure and residues surrounding the heme in an O₂ binding H-NOX domain (a) and in a wild-type myoglobin (Mb) (b). (c) Protoporphyrin IX (FePPIX) of hemoproteins.

Figure 2.
Experimental and optimized structures for Heme(CO)+nRs ($n = 0, 2, 6, 8$) from H-NOX.

Figure 3.
Saddling and ruffling structures for heme macrocycle.

Figure 4.
Experimental and optimized structures for FeP(His)(CO)+8Rs from H-NOX.

(a) Heme(CO) is optimized without including surrounding residues

(b) Heme(CO) is optimized with including the surrounding residues

Figure 5.

Direct comparison of the experimental (X-ray crystal) structure (black line) with the optimized (DFT-D3) structures (white line) for the Heme(CO) moiety with and without including the surrounding residues. Hydrogen atoms are omitted for clarity.

Figure 6.
Experimental and optimized structures for Heme(CO)+8Rs from Mb.

Table 1

Experimental and optimized structural parameters for Heme(CO) and FeP(His)(CO) (distance R in Å, angle $\angle$ in degree)

	Heme(CO)			FeP(His)(CO)		
	Exptl ^a	DFT ^b	DFT-D3 ^c	DFT ^b	DFT-D3 ^c	DFT-D3 ^c
R _{Fe-N(ax)}	1.983	2.012	2.008	2.012	2.009	2.009
R _{C(4N)···Fe} ^d	-0.063	-0.029	-0.051	-0.035	-0.049	-0.049
R _{Fe-N(ax)}	2.068	2.059	2.022	2.058	2.022	2.022
R _{Fe-C(CO)}	(1.75) ^e	1.751	1.752	1.754	1.756	1.756
R _{C-O}	(1.17)	1.162	1.161	1.162	1.161	1.161
$\angle_{\text{Fe-C-O}}$	(180.0)	179.6	179.7	179.6	179.7	179.7
R _{plane(4N)···C1}	-0.218	-0.028	-0.134	0.040	-0.061	-0.061
R _{plane(4N)···C2}	0.123	0.049	0.016	0.073	0.016	0.016
R _{plane(4N)···C3}	0.493	0.267	0.238	0.080	0.071	0.071
R _{plane(4N)···C4}	0.211	0.182	0.104	0.048	-0.003	-0.003
R _{plane(4N)···C5}	-0.689	-0.048	-0.205	0.021	-0.062	-0.062
R _{plane(4N)···C6}	-0.357	0.035	-0.080	0.051	0.007	0.007
R _{plane(4N)···C7}	0.569	0.290	0.121	0.050	0.029	0.029
R _{plane(4N)···C8}	0.309	0.168	-0.093	0.018	-0.047	-0.047
$\Delta(\text{C}_{\text{top}}-\text{C}_{\text{bot}})$ (°)	1.26	0.34	0.44	0.06	0.13	0.13
rms (for β Cs)	0.41	0.17	0.14	0.05	0.04	0.04
R _{plane(4N)···C9}	0.682	0.151	0.190	0.076	0.098	0.098
R _{plane(4N)···C10}	-0.437	-0.045	-0.143	-0.008	-0.086	-0.086
R _{plane(4N)···C11}	0.299	0.185	0.199	0.061	0.088	0.088
R _{plane(4N)···C12}	-0.394	-0.060	-0.227	-0.011	-0.099	-0.099
$\Delta(\text{C}_{\text{top}}-\text{C}_{\text{bot}})$ (meso)	1.12	0.25	0.43	0.09	0.20	0.20
rms (for meso Cs)	0.47	0.13	0.19	0.05	0.09	0.09

^aX-ray crystal structural data on H-NOX(O₂) (code 1U4H). The R_{plane(4N)···C} values for the top and bottom C atoms are indicated in bold; the same is true for the other columns and tables.

^bThe initial structure for the geometry optimization with DFT is based on the experimental structure of the protein.

\$watermark-text

\$watermark-text

\$watermark-text

^cThe initial DFT optimized structure was adopted in the geometry optimization with DFT-D3; the same is true for the other tables (Tables 2 – 4).

^dA negative value denotes an out-of-plane displacement of Fe toward the CO ligand.

^eThe structural parameters in parentheses are those set to estimated values; the same is true for the other tables.

Table 2

Experimental and optimized structural parameters for Heme(CO)+6Rs, Heme(CO)+2Rs, Heme(CO)+8Rs, and Heme(CO)+8Rs+APE (APE = additional protein environment) from H-NOX (*code* 1U4H) (distance R in Å, angle $\angle$ in degree)

	Heme(CO)+6Rs			Heme(CO)+2Rs			Heme(CO)+8Rs			QM/MM ^b		
	Exptl	DFT ^a		DFT	DFT-D3		DFT	DFT-D3		DFT	DFT-D3	
		DFT	DFT-D3		DFT	DFT-D3		DFT	DFT-D3		DFT	DFT-D3
R _{Fe-N(eq)}	1.983	2.012	1.997	2.007	1.990	2.006	1.993	2.001	1.993			
R _{C1(dN)···Fe}	-0.063	-0.044	-0.058	-0.025	-0.061	-0.043	-0.062	-0.051	-0.067			
R _{Fe-N(ax)}	2.068	2.064	2.027	2.056	2.022	2.060	2.032	2.056	2.033			
R _{Fe-C(CO)}	(1.75)	1.736	1.713	1.752	1.753	1.732	1.712	1.729	1.716			
R _{CO}	(1.17)	1.170	1.173	1.162	1.162	1.172	1.175	1.171	1.175			
$\angle_{\text{Fe-C-O}}$	(180.0)	176.4	174.4	180.0	179.7	179.0	174.6	178.1	178.5			
R _{Fe···C (Ile5)}	5.979	7.090	6.403			6.666	5.639	6.241	5.604			
R _{Fe···N (Trp9)}	7.167	7.799	6.832			7.644	6.886	7.580	7.496			
R _{Fe···N (Asn74)}	6.799	8.127	6.172			8.131	6.581	7.234	7.116			
R _{Fe···C (Phe78)}	4.648	5.853	4.384			5.640	4.574	5.098	4.632			
R _{Fe···O (Tyr140)}	5.139	5.861	5.359			5.689	5.170	5.644	5.591			
R _{Fe···C (Leu144)}	4.743	6.918	4.821			6.582	6.097	4.867	4.597			
R _{Fe···Se (Mse98)}	5.668			6.129	5.551	6.128	5.568	6.028	5.541			
R _{Fe···C (Pro115)}	5.447			6.455	5.394	6.207	5.683	5.891	5.364			
R _{plane(4N)···C1}	-0.218	-0.091	-0.209	-0.086	-0.405	-0.187	-0.442	-0.195	-0.330			
R _{plane(4N)···C2}	0.123	0.021	0.073	0.149	-0.040	0.056	-0.084	0.073	-0.048			
R _{plane(4N)···C3}	0.493	0.209	0.286	0.230	0.165	0.208	0.153	0.240	0.392			
R _{plane(4N)···C4}	0.211	0.087	0.022	0.011	-0.161	-0.015	-0.093	-0.030	0.205			
R _{plane(4N)···C5}	-0.689	-0.146	-0.372	-0.064	-0.398	-0.228	-0.414	-0.308	-0.652			
R _{plane(4N)···C6}	-0.357	-0.007	-0.016	0.132	-0.108	0.023	-0.093	0.049	-0.406			
R _{plane(4N)···C7}	0.569	0.289	0.340	0.409	0.489	0.470	0.513	0.491	0.608			
R _{plane(4N)···C8}	0.309	0.121	-0.034	0.169	0.063	0.205	0.082	0.201	0.304			
$\Delta(\text{C}_{\text{top}}-\text{C}_{\text{bot}})(\beta)$	1.26	0.44	0.71	0.50	0.89	0.70	0.96	0.80	1.26			
<i>rms</i> (for β Cs)	0.41	0.15	0.22	0.19	0.28	0.22	0.29	0.25	0.41			

	Heme(CO)+6Rs		Heme(CO)+2Rs		Heme(CO)+8Rs		QM/MM ^b	
	Exptl	DFT ^a	DFT-D3	DFT	DFT-D3	DFT	DFT	DFT-D3
R _{plane(4N)···C9}	0.682	0.167	0.348	0.329	0.419	0.309	0.342	0.311
R _{plane(4N)···C10}	-0.437	-0.132	-0.366	-0.207	-0.428	-0.280	-0.329	-0.318
R _{plane(4N)···C11}	0.299	0.228	0.440	0.344	0.467	0.375	0.381	0.469
R _{plane(4N)···C12}	-0.394	-0.123	-0.359	-0.204	-0.452	-0.232	-0.298	-0.320
Δ(C _{10p} -C _{10a}) (<i>meso</i>)	1.12	0.36	0.81	0.55	0.92	0.66	0.71	0.79
<i>rms</i> (for <i>meso</i> Cs)	0.47	0.17	0.38	0.28	0.44	0.30	0.34	0.36

^aThe initial DFT optimized structure of Heme(CO) (nearly planar) (see Table 1) was adopted for the geometry optimizations with DFT on the various Heme(CO)+nRs systems as well as on Heme(CO)+8Rs+APE (APE = additional protein environment).

^bQM/MM calculations on Heme(CO)+8Rs(QM part)+APE(MM part)

Table 3

Experimental and optimized structural parameters for FeP(His)(CO)+8Rs from H-NOX (*code* 1U4H)
(distance R in Å, angle $\angle$ in degree)

	Exptl	DFT ^a	DFT ^b	DFT-D3
R _{Fe-N(cq)}	1.983	2.004	2.004	2.002
R _{Ct(4N)⋯Fe}	−0.063	−0.043	−0.044	−0.052
R _{Fe-N(ax)}	2.068	2.064	2.068	2.020
R _{Fe-C(CO)}	(1.75)	1.731	1.730	1.717
R _{C-O}	(1.17)	1.173	1.173	1.174
$\angle_{\text{Fe-C-O}}$	(180.0)	179.3	178.4	173.3
R _{Fe⋯C (Ile5)}	5.979	6.339	6.384	5.586
R _{Fe⋯N (Trp9)}	7.167	7.679	7.700	7.129
R _{Fe⋯N (Asn74)}	6.799	8.036	7.943	6.990
R _{Fe⋯C (Phe78)}	4.648	5.505	5.433	4.665
R _{Fe⋯O (Tyr140)}	5.139	5.695	5.687	5.321
R _{Fe⋯C (Leu144)}	4.743	6.109	6.011	5.946
R _{Fe⋯Se (Mse98)}	5.668	6.164	6.153	5.655
R _{Fe⋯C (Pro115)}	5.447	5.736	5.712	5.596
R _{plane(4N)⋯C1}	−0.218	−0.165	−0.174	−0.266
R _{plane(4N)⋯C2}	0.123	0.062	0.048	0.000
R _{plane(4N)⋯C3}	0.493	0.208	0.201	0.201
R _{plane(4N)⋯C4}	0.211	−0.029	−0.029	−0.049
R _{plane(4N)⋯C5}	−0.689	−0.218	−0.232	−0.202
R _{plane(4N)⋯C6}	−0.357	0.009	−0.011	0.039
R _{plane(4N)⋯C7}	0.569	0.330	0.331	0.303
R _{plane(4N)⋯C8}	0.309	0.121	0.127	0.065
$\Delta(\text{C}_{\text{top}}\text{--C}_{\text{bot}})$ (β)	1.26	0.55	0.56	0.57
<i>rms</i> (for β Cs)	0.41	0.18	0.18	0.18
R _{plane(4N)⋯C9}	0.682	0.322	0.310	0.338
R _{plane(4N)⋯C10}	−0.437	−0.283	−0.279	−0.298
R _{plane(4N)⋯C11}	0.299	0.303	0.293	0.321
R _{plane(4N)⋯C12}	−0.394	−0.220	−0.218	−0.297
$\Delta(\text{C}_{\text{top}}\text{--C}_{\text{bot}})$ (<i>meso</i>)	1.12	0.61	0.59	0.64
<i>rms</i> (for <i>meso</i> Cs)	0.47	0.28	0.28	0.31

^aThe initial structure for the geometry optimization is based on the distorted, experimental structure of the protein.

^bThe initial structure of the FeP moiety in the geometry optimization is set to be planar (see Figure 4).

Table 4

Experimental and optimized structural parameters for Heme(CO), Heme(CO)+8Rs, and Heme(CO)+8Rs+APE (APE = additional protein environment) from Mb (distance R in Å, angle $\angle$ in degree)

	Exptl (1A6M)	Exptl (1BZR)	Heme(CO)		Heme(CO)+8Rs		QM/MM ^b	
			DFT ^a	DFT-D3	DFT	DFT-D3	DFT	DFT-D3
R _{Fe-N(e)}	2.006	2.005	2.014	2.011	2.016	2.012	2.014	2.007
R _{C(dN)···Fe}	0.026	-0.016	-0.032	-0.048	-0.037	-0.055	-0.026	-0.041
R _{Fe-N(ax)}	2.064	2.112	2.065	2.018	2.061	2.021	2.057	2.034
R _{Fe-C(CO)}	(1.75)	1.731	1.751	1.753	1.737	1.739	1.750	1.741
R _{CO}	(1.17)	1.126	1.162	1.162	1.171	1.170	1.165	1.167
$\angle_{\text{Fe-C-O}}$	(180.0)	171.3	180.0	179.4	177.4	179.0	178.9	179.2
R _{Fe···N (His64)}	4.576	4.771			5.493	5.435	5.390	4.843
R _{Fe···C (Val68)}	4.702	4.791			6.438	5.340	6.221	5.637
R _{Fe···C (Phe43)}	5.228	5.223			6.223	5.332	6.090	5.277
R _{Fe···C (Ile107)}	6.549	6.546			7.222	5.441	6.778	5.629
R _{Fe···C (Leu29)}	7.380	7.219			7.676	6.473	7.556	7.250
R _{Fe···C (Leu89)}	5.192	5.259			6.623	5.737	6.676	5.358
R _{Fe···C (Ile99)}	5.770	5.913			7.189	5.589	6.646	5.547
R _{Fe···C (His97)}	5.084	5.213			6.355	6.184	6.290	6.146
R _{plane(4N)···C1}	0.395	0.359	0.045	0.027	0.152	0.125	0.200	0.202
R _{plane(4N)···C2}	0.347	0.316	0.097	0.127	0.102	0.125	0.105	-0.007
R _{plane(4N)···C3}	0.002	-0.003	0.017	-0.018	-0.091	-0.092	-0.121	-0.123
R _{plane(4N)···C4}	-0.026	0.004	0.004	-0.107	-0.049	-0.184	-0.087	-0.071
R _{plane(4N)···C5}	0.014	-0.082	0.120	-0.019	0.180	-0.084	0.196	0.220
R _{plane(4N)···C6}	-0.030	-0.085	0.119	0.058	0.145	0.057	0.198	0.099
R _{plane(4N)···C7}	0.165	0.150	0.088	0.020	-0.040	0.038	-0.007	-0.108
R _{plane(4N)···C8}	0.112	0.122	0.033	-0.052	-0.030	-0.005	-0.016	-0.035
$\Delta(\text{C}_{\text{top}}-\text{C}_{\text{bot}})$ (°)	0.43	0.44	0.12	0.23	0.27	0.31	0.32	0.34
rms (for β Cs)	0.20	0.19	0.08	0.07	0.11	0.10	0.14	0.13

	Exptl (1A6M)	Exptl (1BZR)	Heme(CO)		Heme(CO)+8Rs		QM/MM ^b	
			DFT ^a	DFT-D3	DFT	DFT-D3	DFT	DFT-D3
R _{plane(4N)···C9}	0.043	0.072	0.069	0.127	-0.029	0.045	0.007	-0.094
R _{plane(4N)···Cl0}	-0.008	-0.043	0.020	-0.110	0.067	-0.167	-0.002	0.051
R _{plane(4N)···Cl1}	0.061	0.057	0.084	0.098	0.000	0.127	0.099	-0.138
R _{plane(4N)···Cl2}	0.115	0.123	-0.035	-0.095	0.076	-0.016	0.059	0.154
Δ(C _{top} -C _{bot}) (<i>meso</i>)	0.12	0.17	0.12	0.24	0.11	0.29	0.10	0.29
<i>rms</i> (for <i>meso</i> Cs)	0.07	0.08	0.06	0.11	0.05	0.11	0.06	0.12

^aThe initial structure for the geometry optimization with DFT is based on the experimental structure of the protein.

^bQM/MM calculations on Heme(CO)+8Rs(QM part)+APE(MM part).

Table 5Calculated binding energies (E_{bind} in eV^a) of CO to heme in different systems

	DFT	DFT-D3	Exptl
Heme-CO	1.52	1.66	0.85 ^c
8Rs+Heme-CO (from H-NOX)	1.56	1.59	
8Rs+Heme-CO (planar heme) ^b (H-NOX)	1.61	1.65	
8Rs+Heme-CO (from Mb)	1.48	1.63	0.80 ^d

^a 1 eV = 23.06 kcal/mol = 96.5 kJ/mol.^b The planar skeleton of the heme (which contains 4 N-atoms and 20 C-atoms) for the optimized Heme(CO) is kept frozen in the geometry optimization on Heme(CO)+8Rs.^c Dissociation barrier for Mb, corrected for the absence of the protein environment (ref. [43]).^d Dissociation barrier for Mb (ref. [44]).