

# A combined density functional theory and molecular mechanics (QM/MM) study of FeCO vibrations in carbonmonoxy myoglobin

Marek Freindorf<sup>a,\*</sup>, Yihan Shao<sup>b</sup>, Shawn T. Brown<sup>b</sup>, Jing Kong<sup>b</sup>, Thomas R. Furlani<sup>a</sup>

<sup>a</sup> Center for Computational Research, State University of New York at Buffalo, 9 Norton Hall, Buffalo, NY 14260, USA

<sup>b</sup> Q-Chem Inc., 5001 Baum Boulevard, Suite 690, Pittsburgh, PA 15213, USA

Received 17 August 2005; in final form 24 November 2005

Available online 28 December 2005

## Abstract

Molecular oscillations of ligand motions at the active site of carbonmonoxy myoglobin have been calculated in a protein environment using the combined QM/MM approach. In these calculations, the active site was calculated at a quantum mechanical (QM) level of theory using the B3LYP/6-31 + G\* method, while the remaining protein was calculated at the molecular mechanical (MM) level utilizing the AMBER force field. The presence of a torsion mode and the second component of the bending vibration are proof of a bent CO geometry in the heme pocket.

© 2005 Elsevier B.V. All rights reserved.

The heme pocket of myoglobin controls the reactivity of the active site toward O<sub>2</sub> and CO adducts [1,2]. Collman et al. [3] proposed that the mechanism for the large difference in affinity between O<sub>2</sub> and CO is based on a steric interaction between the ligand adducted to the heme–Fe complex and a distal histidine in close contact with the ligand. In the protein-free heme–Fe–CO complex, a linear orientation of the Fe–C–O bond is preferred, while in myoglobin the CO ligand is bound to the heme–Fe complex in a bent conformation. There are several high-resolution X-ray crystallographic structures showing a range of displacements and orientations for the CO ligand in various crystal forms [4–12], and all of them show the heme–Fe–CO geometry to be non-linear. To date, spectroscopic data fail to provide any evidence in support of a non-linear Fe–C–O bond. In the middle-frequency spectral region there is a Fe–CO stretching mode, involving the Fe–C bond, which is experimentally observed [13] at 512 cm<sup>-1</sup>. However, there is dispute over the assignment of the Fe–C–O bending modes [14], which are expected to be non-degenerate in the protein due to steric hindrance with the heme pocket. One component of the bending

mode is assumed to be at 577 cm<sup>-1</sup> in the protein [15]. Unfortunately, there is no experimental confirmation of the second Fe–C–O bending mode, which in the gas phase is almost degenerate with the first component [16]. While there are also expected to be low-frequency Fe–CO bending and torsion modes, their IR and RR intensities are too weak to be observed. Assignment of these low-frequency modes is important because it provides insight into the geometry of the CO ligand bonded to the heme–Fe molecule in the protein. Based on the work presented here, we suggest that previously unassigned weak spectral features in the protein correspond to bending and torsion motions, thereby providing further proof of a bent CO geometry in the heme pocket.

In order to assign the low frequency spectral features in carbonmonoxy myoglobin to specific vibrational modes, we carried out a series of combined QM/MM calculations using a B3LYP/6-31 + G\*/AMBER potential [17]. The ligand oscillations of the active site in carbonmonoxy myoglobin have already been investigated in the literature by Rovira et al. [18], using a similar QM/MM approach, however the authors report only differences in ligand frequencies between the isolated system in the gas phase and a small QM model system of the protein active site. We report here results of more ambitious calculations which

\* Corresponding author. Fax: +1 716 645 6505.

E-mail address: [mfrein@buffalo.edu](mailto:mfrein@buffalo.edu) (M. Freindorf).

include additional vibrations along with the absolute value of the computed frequencies.

Details of the calculations are as follows. A near atomic resolution X-ray structure of Schlichting et al.[11], has the highest resolution (1.15 Å) published to date and accordingly was chosen as the starting point for our study. The 1A6G protein data bank contains three experimental conformers (denoted as A, B, and C) which differ primarily in the position of the distal histidine HIS<sub>64</sub>. In conformer B, the HIS<sub>64</sub> residue is slightly closer to CO than in the A conformer, while in conformer C, the HIS<sub>64</sub> residue is swung-out of the heme pocket. It is thought that conformers A and B correspond to different protonation states of HIS<sub>64</sub>, however there is still uncertainty in this assignment. The C conformer has been identified by a diffraction experiment at low pH, which indicates a double protonation state N<sub>δ<sub>e</sub></sub> of HIS<sub>64</sub>. Given the uncertainty associated with the protonation state of HIS<sub>64</sub> in the conformers, we calculated both protonation states N<sub>δ</sub> and N<sub>ε</sub> of HIS<sub>64</sub> in conformers A and B, named as: A-HIS<sub>64</sub>N<sub>δ</sub>H, A-HIS<sub>64</sub>N<sub>ε</sub>H, B-HIS<sub>64</sub>N<sub>δ</sub>H, B-HIS<sub>64</sub>N<sub>ε</sub>H. Conformer C has been represented by one protonation tautomer C-HIS<sub>64</sub>N<sub>δ<sub>e</sub></sub>H. Hydrogen atoms, which are missing in the experimental structure, were placed in their appropriate positions using the AMBER-8 program [19], and their positions were minimized at the molecular mechanical level, utilizing standard parameters for the heme-Fe-CO molecular complex [20] as implemented in the AMBER-8 program. After the energy minimization, each structure was divided into a QM region, which includes the heme-Fe-CO molecule and the proximal histidine HIS<sub>93</sub>, and a MM region, which includes the remaining protein. The chemical bond between the proximal histidine HIS<sub>93</sub> and a side chain of the protein was cut, and resulting free chemical valences were filled out by hydrogen atoms, according to the linking atom method [21,22]. We do not believe the presence of this covalent bond strongly perturbs the CO ligand oscillations because it is located more than 8Å from the CO ligand, and the CO ligand oscillations are mostly represented by local Fe-C-O modes. The QM region was simplified by eliminating side chains of the heme ring and a side chain of the HIS<sub>93</sub> residue, as was done in previous studies of this system [16,23]. From this starting point, a geometry optimization of the QM system was carried out with the MM system held fixed. The QM/MM calculations included external MM charges in the self consistent field (SCF) computations of the QM wave function, as well as a van der Waals interaction between the QM atoms and the MM atoms. The van der Waals interaction was approximated by a standard Lennard-Jones potential. The QM part of the combined QM/MM calculations was carried out at the B3LYP/6-31 + G\* level of theory, which has proven [24] to be very effective for computations of biological systems. The perturbing MM atoms were described by the AMBER-8 force field. The combined QM/MM calculations presented in our study were performed using the Q-Chem [25] program. The geometrical parameters of the heme-

Fe-CO active site in the protein and in the gas phase, which were determined in this study, are presented in Table 1. The torsion angle, denoted as  $\alpha$ , is defined as the N-Fe-C-O dihedral angle, where the N atom belongs to the heme ring. The bending angle, denoted as  $\beta$ , is the Fe-C-O angle, and the tilting angle, denoted as  $\gamma$ , is the N-Fe-C angle, where the N atom belongs to the proximal HIS<sub>93</sub> ring. According to our calculations, the optimal value for the torsion angle  $\alpha$  is the most sensitive geometric parameter with respect to the various protein conformers; varying from 0° to 100°. However its average value (for the conformers considered in this study) is 31° and is very close to experiment (36°). The optimal values of the other angles are close to the results of calculations reported by others [23,26] as well as experimental measurements, as indicated in Table 1. The bending angle ( $\beta$ ) which was calculated in our study (168°) is slightly smaller from the value of this angle calculated by Rovira [27] (177°) and Ryde et al. [23] (171°), using the similar QM/MM approach. The results of our present calculations are obtained using new values of Lennard-Jones parameters describing the interaction between QM and MM atoms. The new parameters used in our calculations were evaluated using the QM/MM approach, and they are slightly bigger (about 10–20%) than parameters used in combined QM/MM calculations by others, which were based on molecular mechanical force fields. Bigger Lennard-Jones parameters cause a stronger repulsive interaction between QM and MM atoms, which causes that the bending angle obtained in our study is slightly smaller than the angle calculated previously by others. We have also performed test calculations for the QM system in the protein having the CO ligand coordinated to the heme iron atom in a linear geometry. The energy difference between the QM system having the optimal bent CO geometry in the protein, and the QM system having the CO ligand bound to the heme molecule in a linear conformation in the protein, was found in our test calculations to have a value between 30 and 50 kcal/mol, depending on the protonation state of distal histidine and depending on the experimental protein conformer (A, B or C). This big CO bending energy which we have estimated in our test calculations in the protein, supports the initial explanation of protein discrimination (lowering affinity of CO binding to heme proteins) suggested many years ago by Collman et al.[3]. Therefore, this issue is a subject of our current investigations.

Once the geometry of the QM moiety was optimized, combined QM/MM calculations of the oscillation Hessian were performed, again with the MM portion of the protein held fixed. For comparison, we have also calculated the oscillation Hessian of the QM system in the gas phase (without the protein). The results of our calculations are shown in Table 2, and sketches of vibrational eigenvectors involving large amplitudes of the CO movement are shown in Fig. 1.

Three types of vibrational modes were investigated in the present study, namely (1) a low-frequency torsion

Table 1

The geometric parameters of the heme-Fe-CO molecule calculated in the protein and in the gas phase, which are compared with the X-ray experimental data

Geometrical parameter	Fe-C ( $r$ ) distance (Å)	C-O ( $d$ ) distance (Å)	N-Fe-C-O ( $\alpha$ ) torsion (°)	Fe-C-O ( $\beta$ ) bending (°)	N-Fe-C ( $\gamma$ ) tilting (°)
Gas-phase	1.78	1.15	–	180	180
<i>Protein</i>					
A-HIS <sub>64</sub> N <sub><math>\delta</math></sub> H	1.77	1.15	10	169	171
A-HIS <sub>64</sub> N <sub><math>\epsilon</math></sub> H	1.75	1.15	32	171	171
B-HIS <sub>64</sub> N <sub><math>\delta</math></sub> H	1.79	1.15	0	164	167
B-HIS <sub>64</sub> N <sub><math>\epsilon</math></sub> H	1.76	1.15	17	166	167
C-HIS <sub>64</sub> N <sub><math>\delta\epsilon</math></sub> H	1.77	1.15	100	168	171
Average value	1.76	1.15	31	168	170
Experiment <sup>a</sup>	1.82	1.09	36	171	172

<sup>a</sup> Ref. [11].

Table 2

Vibrational frequencies of the heme-Fe-CO complexes in the protein

Vibration (cm <sup>-1</sup> )	$\tau$	$\nu_1$	$\delta_a$	$\delta_b$	$\nu_2$
Gas-phase	–	483	547	557	1922
<i>Protein</i>					
A-HIS <sub>64</sub> N <sub><math>\delta</math></sub> H	100	502	552	596	1956
A-HIS <sub>64</sub> N <sub><math>\epsilon</math></sub> H	103	524	570	603	1926
B-HIS <sub>64</sub> N <sub><math>\delta</math></sub> H	126	470	548	582	1944
B-HIS <sub>64</sub> N <sub><math>\epsilon</math></sub> H	121	502	558	589	1903
C-HIS <sub>64</sub> N <sub><math>\delta\epsilon</math></sub> H	89	494	570	587	1915
Average value	108	498	560	591	1929
Experiment <sup>a</sup>		512	577	586	1944

The high-frequency vibration  $\nu_2$  obtained in our calculations, has been scaled by a factor 0.92, which reflects a common observation with DFT vibrational calculations for this mode [31]. The experimental value of 586 cm<sup>-1</sup> corresponds to an unassigned spectral peak displayed in Fig. 6 in a paper by Tsubaki et al. [13].

<sup>a</sup> Ref. [13].

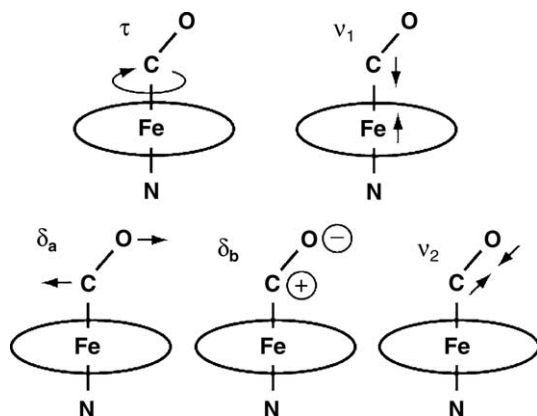


Fig. 1. Sketches of oscillation eigenvectors involving large CO amplitudes of the heme-Fe-CO complexes calculated in myoglobin.

vibration  $\tau$  of the Fe-CO bond, (2) middle-frequency vibrations  $\nu_1$ ,  $\delta_a$  and  $\delta_b$  involving stretching of the Fe-C bond and two bending modes of the Fe-C-O bond, and (3) a high-frequency oscillation  $\nu_2$  involving mostly stretching of the C-O bond. Table 2 shows the effect that the protein environment has on the vibrational frequencies of these

normal modes. There are vibrational modes between the low and middle-frequency region that also contribute to the CO movement, however the CO motion in these vibrations is strongly coupled with different heme ring motions, therefore it is difficult to categorize these vibrations as vibrations involving a large CO amplitude.

The low-frequency torsion mode,  $\tau$  (see Fig. 1) is not observed in the gas phase because of the linear alignment of the Fe-C-O bond. However in the protein, where the CO bond is bent, we computed a frequency of 108 cm<sup>-1</sup>, which is an arithmetic average value over all the conformers. In the middle-frequency spectral region, we observe quite good agreement between the calculated frequencies of the  $\nu_1$ ,  $\delta_a$  and  $\delta_b$  oscillations and experimental data. The average value of the frequencies of these vibrations, calculated in the protein, differ from experiment by less than 20 cm<sup>-1</sup>, whereas the gas phase calculations differ by 30 cm<sup>-1</sup> from experiment. Moreover, we observe in this spectral region that the calculated frequencies in the protein are systematically larger than the corresponding values in the gas phase. This observation agrees with the general expectation that an oscillating molecule should exhibit larger frequencies in the protein than in the gas phase due to additional interactions in the protein (mostly van der Waals) between the molecule and the protein environment.

In the high-frequency spectral region, the CO stretching vibrational mode  $\nu_2$  in the gas phase is calculated to be 1922 cm<sup>-1</sup>. In the protein, this vibration is calculated to be slightly higher (1929 cm<sup>-1</sup>), which is closer to the value observed experimentally (1944 cm<sup>-1</sup>). As was the case for the middle-frequency region, the calculated frequency of the high-frequency vibration  $\nu_2$  in the protein differs from experimental value by less than 20 cm<sup>-1</sup>. The RR experiment shows only one distinct band in the high-energetic spectral region, however the IR spectrum of this protein shows three bands [28] at this spectral range that are conventionally labeled A<sub>0</sub> (1967 cm<sup>-1</sup>), A<sub>1</sub> (1944 cm<sup>-1</sup>) and A<sub>2</sub> (1933 cm<sup>-1</sup>). While their IR relative intensity depends on a variety of external conditions, the intensity of the A<sub>1</sub> band remains the strongest, and it corresponds to the intensity maximum of the RR spectrum. There are several

models that seek to explain the origin of the three bands, however none of them seems to be generally accepted. Unfortunately, we are unable to elucidate the origin of these bands from the calculations carried out to date, however this subject is currently under further investigation.

In conclusion, the presence of the torsion mode  $\tau$  and the second component of the bending vibration  $\delta_b$  for CO bound to the heme complex in the protein are proof of a bent CO geometry in the protein. Our calculations also correctly reproduce the experimentally observed shift [29,30] in frequency of both stretching vibrations  $\nu_1$  and  $\nu_2$  with the change in the HIS<sub>64</sub> protonation state. We believe that this effect is due to the electrostatic interaction between the heme pocket and CO, and a full paper related to this subject is in preparation.

### Acknowledgements

The authors gratefully acknowledge support of NIH SBIR Grant 1 R43 GM065617.

### References

- [1] L. Stryer, *Biochemistry*, W.H. Freeman and Co., New York, 1988.
- [2] B.A. Springer, S.G. Sligar, J.S. Olson, G.N. Phillips Jr., *Chem. Rev.* 94 (1994) 699.
- [3] J.P. Collman, J.I. Brauman, T.R. Halbert, K.S. Suslick, *Proc. Natl. Acad. Sci. USA* 73 (1976) 3333.
- [4] J.C. Norvell, A.C. Nunes, B.P. Schoenborn, *Science* 190 (1975) 568.
- [5] J.C. Hanson, B.P. Schoenborn, *J. Mol. Biol.* 153 (1981) 117.
- [6] J. Kuriyan, S. Wilz, M. Karplus, G.A. Petsko, *J. Mol. Biol.* 192 (1986) 133.
- [7] X. Cheng, B. Schoenborn, *J. Mol. Biol.* 220 (1991) 381.
- [8] M.L. Quillin, R.M. Arduini, J.S. Olson, G.N. Phillips Jr., *J. Mol. Biol.* 234 (1993) 140.
- [9] F. Yang, G.N. Phillips Jr., *J. Mol. Biol.* 256 (1996) 762.
- [10] G.S. Kachalova, A.N. Popov, H.D. Bartunik, *Science* 284 (1999) 473.
- [11] J. Vojtechovsky, K. Chu, J. Berendzen, R.M. Sweet, I. Schlichting, *Biophys. J.* 77 (1999) 2153.
- [12] G.N. Phillips Jr., T. Romo, (PDB entry 1JW8), 2002, in press.
- [13] M. Tsubaki, R.B. Srivastava, N.T. Yu, *Biochemistry* 21 (1982) 1132.
- [14] S. Hirota, T. Ogura, K. Shizawa-Itoh, S. Yoshikawa, M. Nagai, T. Kitigawa, *J. Phys. Chem.* 98 (1994) 6652.
- [15] C. Rajani, J.R. Kincaid, *J. Am. Chem. Soc.* 120 (1998) 7278.
- [16] P.M. Kozlowski, K.M. Vogel, M.Z. Zgierski, T.G. Spiro, *J. Porphyrins Phthalocyanines* 5 (2001) 312.
- [17] M. Freindorf, Y. Shao, T.R. Furlani, J. Kong, *J. Comput. Chem.* 26 (2005) 1270.
- [18] C. Rovira, B. Schulze, M. Eichinger, J.D. Evanseckm, M. Parinello, *Biophys. J.* 81 (2001) 435.
- [19] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, R.J. Radmer, Y. Duan, J. Pitera, M. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, AMBER-8, University of California, San Francisco, 2004.
- [20] D.A. Giammona, Ph.D. thesis, University of California, Davis, 1984.
- [21] M.J. Field, P.A. Bash, M. Karplus, *J. Comput. Chem.* 11 (1990) 700.
- [22] U.C. Singh, P.A. Kollman, *J. Comput. Chem.* 7 (1986) 718.
- [23] E. Sigfridsson, U. Ryde, *J. Inorg. Biochem.* 91 (2002) 101.
- [24] P.R. Rablen, J.W. Lockman, W.L. Jorgensen, *J. Phys. Chem. A* 102 (1998) 3782.
- [25] J. Kong, C.A. White, A.I. Krylov, D. Sherrill, R.D. Adamson, T.R. Furlani, M.S. Lee, A.M. Lee, S.R. Gwaltney, T.R. Adams, C. Ochsenfeld, A.T.B. Gilbert, G.S. Kedziora, V.A. Rassolov, D.R. Maurice, N. Nair, Y. Shao, N.A. Besley, P.A. Maslen, J.P. Dombrowski, H. Daschel, W. Zhang, P.P. Korambath, J. Baker, E.F.C. Byrd, T.V. Voorhis, M. Oumi, S. Hirata, C.P. Hsu, N. Ishikawa, J. Florian, A. Warshel, B.G. Johnson, P.M.W. Gill, M. Head-Gordon, J.A. Pople, *J. Comput. Chem.* 21 (2000) 1532.
- [26] F. De Angelis, A. Jarzecki, R. Car, T.G. Spiro, *J. Phys. Chem. B* 109 (2005) 3065.
- [27] C. Rovira, *J. Phys.: Condes. Matter* 15 (2003) 1809.
- [28] J.O. Alben, D. Beece, S.F. Bowne, W. Doster, L. Eisenstein, H. Frauenfelder, D. Good, J.D. McDonald, M.C. Marden, P.P. Moh, L. Reinisch, A.H. Reynolds, E. Shyamsunder, K.T. Yue, *Proc. Natl. Acad. Sci. USA* 79 (1982) 3744.
- [29] B. Kushkuley, S.S. Stavrov, *Biophys. J.* 72 (1997) 899.
- [30] S. Frenzen, *J. Am. Chem. Soc.* 124 (2002) 13271.
- [31] G. Rauhut, P. Pulay, *J. Phys. Chem.* 99 (1995) 3093.