

The Fe-CO Bond Energy in Myoglobin: A QM/MM Study of the Effect of Tertiary Structure

Nikki Strickland, Adrian J. Mulholland and Jeremy N. Harvey

Centre for Computational Chemistry and School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.

Supplemental Information

Please note: Reference numbers with an "S" prefix are those listed in the supplemental information. All others refer to those in the main biophysical letters article.

STARTING COORDINATES

The 1.15 Å resolution refined crystal structures of sperm whale deoxy myoglobin (Mb) (14) (Protein Data Bank (PDB) entry 1a6g) and carbonmonoxy Mb (PDB entry 1a6n, with Asp122 changed to Asn122 to be consistent with 1a6g), with minor alternative conformations removed, were taken as the starting coordinates. Missing C-termini were added using the Deep View Swiss-Pdb Viewer (S1) and manipulated to the same orientation as the C-termini of the similar 1.15 Å resolution refined crystal structures of other sperm whale deoxyMb and MbCO structures (1bzp and 1bzs respectively (8)). Waters of crystallisation and sulfate ions were removed. The normal protonation state at pH 7 was assumed for all amino acids apart from histidines. Examination of the protein and solvent environment around the other histidines led to the assignment of histidines 12, 81, 82, 97 and 119 as the ϵ -tautomer (HIE), histidines 24, 48, 113 and 116 as the δ -tautomer (HID) and histidine 36 as being protonated on both the ϵ and δ nitrogens (HIS). For His64 we chose the HIE tautomer, in agreement with converging opinion (12, 13, S2-S10). However, some support for the His64 HID tautomer (8, S11) led us to run parallel calculations in this state, which are reported here but not commented on in the main article.

MOLECULAR DYNAMICS (MD) SIMULATIONS

Preparation of both systems and the resulting MD simulations were performed with the CHARMM v27b2 software package (16, S12) and a modified CHARMM all atom force-field (15), which included additional parameters for the deoxy and CO-heme groups (S13, S14). The electrostatic potential atomic charges used for deoxy-haem and CO-haem in the CHARMM force field were determined by optimisation of models, consisting of just the haem porphine, His93 and bound CO (when bound), in the Jaguar electronic structure package (18). The Jaguar calculations used to derive atomic charges used the standard B3LYP density functional method in conjunction with the standard Los Alamos ECP with Jaguar triple-zeta basis set (LACV3P) on Fe, the standard 6-311G* basis set on all N and the CO Ligand and the 6-31G* basis (with 5 spherical harmonic d polarization functions) on all other atoms. The solvent environment of the haem models was approximated with self-consistent reaction field method and Jaguar's Poisson-Boltzmann solver (S15, S16), with the outer dielectric constant of the solvent set to 4 and the radius of solvent probe molecule set to 2.60219. A nonbonding pair-list cut-off of 13 Å was used with a constant dielectric and switching function acting on the force between 8 and 12 Å throughout. The nonbonding pair-list was updated as necessary, using CHARMM's heuristic testing algorithm, and bonds to hydrogen atoms were constrained using the SHAKE algorithm (S17).

Hydrogen atoms were added and minimised on both rigid protein chains. Each protein was soaked in a cube of pre-equilibrated TIP3P water molecules and all TIP3P molecules < 2.7 Å and > 9 Å from the protein non-hydrogen atoms were removed. This created two systems of ~10000 atoms with a 9 Å layer of explicit TIP3P water molecules and a total charge of +2e. This approach allowed for complete solvation of the protein without solvating unnecessary volume, as may be the case when using periodic or stochastic boundary methods. The water was relaxed without any constraints/restraints around the fixed protein structure using 2*300 steps of the Steepest Descent (SD) minimisation algorithm, 1500 steps of the Adopted Basis Newton-Raphson (ABNR)

minimisation algorithm, heating to 300 K over 10 ps, equilibration for 25 ps then another series of minimisations. A 1 fs time step was used for all heating and equilibration during the system's preparation. The few waters which had 'boiled-off' from the outer layer were removed and all constraints and restraints were removed. The rest of the system was minimised with 2*250 SD steps then 2*1500 steps of ABNR. Initial velocities were assigned at 0 K according to a Gaussian distribution and the system was heated to 300 K over 50 ps, in 30 K temperature increments.

Velocities continued to be assigned according to a Gaussian distribution for the 50 ps early equilibration stage, then were reassigned every 100 fs to keep the temperature constant during the 300 ps late equilibration stage. At this point the fluctuations of Mb α -carbons were stable with respect to time and a 100 ps production run was carried out, also with a 1 fs time step. Coordinates were taken every 10 ps from the production run and waters which had 'boiled-off' were removed. The 10 trajectory coordinates and the solvated crystal structure of each Mb were minimised with 2*250 steps of SD minimisation the 2*1500 steps of ABNR minimisation. This resulted in 11 deoxyMb and 11 MbCO starting coordinates for QM/MM. The α -carbon root mean square deviations (rmsd) of the trajectory structures were stable with respect the crystal structures with time throughout the trajectories, with an average deviation of 0.61 Å and maximum deviation of 0.72 Å for the MbCO trajectory and an average and maximum of 0.63 Å and 0.71 Å respectively for the deoxyMb trajectory.

HYBRID QUANTUM MECHANICS/MOLECULAR MECHANICS (QM/MM) CALCULATIONS

Analogous pairs of deoxyMb and MbCO structures were created for each of the 22 starting coordinates; 11 R-state pairs were created by removing bound CO from MbCO and 11 T-state pairs were creating by adding bound CO to deoxyMb. Any water molecules present in deoxyMb's distal cavities were removed. The PDB coordinates were converted to a Tinker XYZ format and Mb's QM section was assigned as the heme porphine (heme group without side chains), the imidazole group of His93 and CO, if bound. All other Mb and solvent atoms were assigned to the MM region. The exclusion of the heme propionate and vinyl groups from the QM section has been shown to have little affect on the binding energy (S18) and has the advantage for us that the QM section can be directly compared to a protein-free gas-phase deoxy and CO-heme models one of us has previously studied (1). Separate QM and MM calculations are carried out and coordinated by our suite of UNIX shell scripts and Fortran 90 programs (QoMMMa (2)). This set of routines interacts with the unmodified Jaguar QM (18) and Tinker MM (19) codes to create, run and analyse the individual jobs whilst integrating the output, calculating the total QM/MM energy and gradients and optimising the QM section of Mb. The 9 covalent bonds joining the QM and MM atoms were modelled by a link atom method (S19), in a version similar to that described by Field et al (S20), where all terms containing only QM atoms were deleted in the MM energy expression.

In the QM calculation the MM atoms at the QM/MM division were substituted for capping hydrogen atoms, which were placed in the direction of the previous QM-MM bond and at a QM-H bond length defined in the CHARMM all-atom force field (15). The QM gradient on the capping atom was spread over both the QM atom it was bound to and the MM atom it replaced, based purely on the position of the capping atom. The capping atom was not explicitly optimised. The polarisation effect of the MM electronic environment on the QM region was represented by inserting an array of point charges corresponding to the MM atoms into the QM Hamiltonian. The charge on the MM link atoms and their closest MM neighbours were set to zero to avoid non-physical effects. Finally, possible steric interactions between QM and MM atoms were included as standard van der Waals radii in the MM force field. The Jaguar QM optimisations used a BFGS method with the standard B3LYP density functional method in conjunction with the standard Los Alamos ECP with Jaguar double-zeta basis set (LACVP) on Fe and the standard 6-31G basis on the rest of the QM section. 5 spherical harmonic d polarization functions were used on all atoms throughout. At the end of each QM/MM optimisation, single point calculations were performed on the QM section with a larger basis set, consisting of the standard Los Alamos ECP with Jaguar triple-zeta basis (LACV3P) on Fe, the standard 6-311G* basis on all N atoms and the CO ligand and the 6-31G* basis set on all other atoms. Restricted "wavefunctions" were used in all cases: RB3LYP for closed shell systems and ROB3LYP for open-shell systems.

At each of the QM geometries the MM region was fully optimised to within an RMS gradient of 0.1 kcal mol⁻¹ Å⁻¹ using the Tinker MM code (19) and the CHARMM (15) all-atom force field. The QM contribution to the

QM/MM energy gradient for the MM atoms was approximated by the gradient induced by a set of point charges positioned at the QM centres and corrected by the difference between this electric field contribution and the exact gradient term at the initial geometry. Mulliken charges were used, with the charge of the capping atom added to that of the QM atom to which it is bound. This approximation becomes exact as the stationary point is approached and many computations of the exact contribution of the QM region to QM/MM energy gradient for the MM atoms are saved.

The convergence of each deoxyMb and MbCO ‘pair’ to the same local substate was achieved by a series of iterations, similar to those described by Wirstam et al (17) in their studies of hemethyrin. QM/MM calculations were carried out separately on deoxyMb and MbCO to produce deoxyMb_1 and MbCO_1 respectively. The QM section of MbCO_1 was replaced by the QM section deoxyMb_1, to produce a hybrid of the MbCO_1 MM region with the deoxyMb_1 QM region. This hybrid was optimised to produce deoxyMb_2. Then the QM section of deoxyMb_2 was replaced by the QM section of MbCO_1 and the resulting structure was optimised to produce MbCO_2. Next the QM section of MbCO_2 was replaced by the QM section of deoxyMb_2 and optimised. This process, of replacing the QM section of MbCO_X with that of deoxyMb_X and optimising to produce deoxyMb_X+1, then replacing the QM section of deoxyMb_X+1 with the QM section of MbCO_X and optimising to produce MbCO_X+1, continued until the pair were in the same local minimum. Convergence was considered to be reached once the energy difference between MbCO_X and MbCO_X+1, as well as that between deoxyMb_X and deoxyMb_X+1, was less than 0.1 kcal/mol. The bond dissociation energy (BDE) was calculated as the difference in energy between MbCO and the sum of deoxyMb and CO(g) fragments, once MbCO and deoxyMb had converged to the same local minimum. A zero-point energy correction of -4.38 kcal/mol, as calculated on the QM-only section, was applied to each BDE.

RESULTS

TABLE S1 Zero-point energy corrected BDEs for 44 Mb conformations (kcal/mol). MM contribution to BDE in brackets

Starting point*	R-state BDE		T-state BDE	
	HID His64	HIE His64	HID His64	HIE His64
Crystal	12.33 (3.3)	10.68 (3.5)	7.03 (0.75)	6.36 (2.1)
1	11.28 (5.3)	9.67 (3.4)	5.57 (3.45)	7.34 (2.6)
2	10.78 (4.9)	10.38 (3.6)	3.39 (3.46)	7.85 (3.8)
3	11.87 (5.1)	8.38 (3.0)	1.59 (3.01)	7.61 (1.8)
4	9.28 (3.9)	10.82 (5.0)	1.66 (3.46)	3.01 (3.1)
5	11.87 (4.6)	10.01 (4.2)	4.23 (4.40)	6.23 (2.5)
6	9.96 (3.9)	8.44 (4.0)	3.51 (3.21)	6.58 (2.8)
7	10.53 (4.6)	8.56 (3.5)	4.58 (3.50)	9.03 (2.8)
8	10.37 (4.3)	8.11 (3.9)	4.63 (3.76)	8.56 (2.6)
9	11.92 (4.3)	9.40 (3.6)	5.59 (4.15)	5.47 (2.7)
10	12.06 (2.7)	9.17 (4.5)	3.61 (4.20)	7.37 (3.6)
Average	11.11 (4.3)	9.42 (3.8)	4.13 (3.40)	6.86 (2.8)

* All BDEs are calculated for Mb crystal or trajectory structures which have undergone MM then QM/MM optimisation. Structures 1 to 10 derive from trajectory snapshots taken every 10 ps from 10 ps onwards

TABLE S2 Key structural features for MbCO with R-state and T-state tertiary conformations (HIE His64 tautomer only)

Starting point*	Fe-C-O angle (°)		Tilt Angle (°) ^a		Fe distoop movement (Å) ^b		Distance between oxygen of bound CO and.....					
	R-state	T-state	R-state	T-state	R-state	T-state	Phe43 Cz 'H'		His64 Nε 'H'		Val68 Cγ2 'H' ^c	
	R-state	T-state	R-state	T-state	R-state	T-state	R-state	T-state	R-state	T-state	R-state	T-state
Crystal	174.37	172.93	4.95	5.67	-0.22	-0.27	2.52	2.53	3.09	2.45	2.56	2.53
1	176.06	174.21	3.13	3.14	-0.26	-0.33	2.65	2.51	2.38	1.89	2.61	2.57
2	176.79	175.13	1.97	2.09	-0.26	-0.30	2.82	2.55	2.03	1.89	2.59	2.48
3	174.37	176.50	3.33	1.41	-0.25	-0.32	2.58	2.56	2.22	1.91	2.55	2.53
4	176.83	174.91	1.95	2.88	-0.24	-0.44	3.26	2.52	3.09	2.31	2.63	2.43
5	174.21	175.59	3.48	2.34	-0.26	-0.34	3.27	2.53	2.05	2.07	2.61	2.49
6	174.12	176.24	3.40	1.41	-0.27	-0.34	3.48	2.60	2.24	1.94	2.67	2.61
7	173.55	176.98	4.64	1.02	-0.26	-0.31	3.18	2.53	2.69	1.94	2.61	2.59
8	175.42	174.70	3.05	2.14	-0.26	-0.33	3.34	2.58	2.71	1.98	2.58	2.59
9	175.73	171.70	3.06	5.51	-0.26	-0.34	3.26	2.83	3.02	1.87	2.56	2.55
10	175.14	171.48	3.36	5.12	-0.26	-0.32	3.57	2.87	2.52	1.96	2.53	2.56
Average	175.15	174.58	3.30	2.98	-0.25	-0.33	3.08	2.60	2.55	2.02	2.59	2.54

^aThe tilt angle is the angle between the Fe-C bond and the axis perpendicular to the plane of best fit through the four porphyrin nitrogens. ^bFe distoop movement is the difference between the iron distance out-of-plane for MbCO and deoxyMb, where the plane is the plane of best fit through the four porphyrin. ^cDistance between oxygen and bound CO and Val68 Cγ2 'H' refers to the distance with the nearest Cγ2 'H' to CO.

The effect of the MM electronic environment on the BDE was calculated by subtracting the single-point energy of each QM region without MM point charges from the same section with charges. The MM point charge stabilisation of the MbCO was subtracted from that of deoxyMb to reveal the charge contribution to the BDE (Table S3). A positive value indicates that the MM point charges stabilise MbCO more than they stabilise deoxyMb.

TABLE S3 Contribution of MM point charges to BDEs (kcal/mol)

Starting point*	R-state BDE		T-state BDE	
	HID His64	HIE His64	HID His64	HIE His64
Crystal	0.74	-0.87	1.43	-0.74
1	-0.34	-1.06	-0.58	1.74
2	0.08	0.28	-0.54	1.74
3	-0.08	-0.42	-0.55	1.29
4	0.23	-0.40	-0.60	-0.03
5	0.04	0.64	-1.32	1.24
6	-0.36	0.58	-0.76	1.52
7	-0.81	-0.36	-0.86	2.59
8	-0.35	-1.10	-1.03	1.47
9	0.63	-1.08	-1.17	1.72
10	0.34	-0.90	-1.64	1.22
Average	0.01	-0.43	-0.69	1.25

REFERENCES

- (S1) Guex, N., and Peitsch, M. C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*. 18:2714-23.
- (S2) Ray, G. B., Li, X. Y., Ibers, J. A., Sessler, J. L., and Spiro, T. G. 1994. How Far Can Proteins Bend the FeCO Unit - Distal Polar and Steric Effects in Heme-Proteins and Models. *J. Amer. Chem. Soc.* 116:162-76.
- (S3) Cui, Q., and Karplus, M. 2000. Molecular properties from combined QM/MM methods. I. Analytical second derivative and vibrational calculations. *J. Chem. Phys.* 112:1133-49.
- (S4) Kushkuley, B. and Stavrov, S. S. 1996. Theoretical study of the distal-side steric and electrostatic effects on the vibrational characteristics of the FeCO unit of the carbonylheme proteins and their models. *Biophys. J.* 70:1214-29.
- (S5) Kushkuley, B. and Stavrov, S. S. 1997. Theoretical study of the electrostatic and steric effects on the spectroscopic characteristics of the metal-ligand unit of heme proteins .2. C-O vibrational

- frequencies, O¹⁷ isotropic chemical shifts, and nuclear quadrupole coupling constants. *Biophys. J.* 72:899-912.
- (S6) Phillips, G. N., Teodoro, M. L., Li, T. S., Smith, B., and Olson, J. S. 1999. Bound CO is a molecular probe of electrostatic potential in the distal pocket of myoglobin. *J. Phys. Chem. B.* 103:8817-29.
- (S7) Straub, J. E., and Karplus, M. 1991. Molecular-Dynamics Study of the Photodissociation of Carbon-Monoxide from Myoglobin - Ligand Dynamics in the 1st 10 ps. *Chem. Phys.* 158:221-48.
- (S8) Unno, M., Christian, J. F., Olson, J. S., Sage, J. T., and Champion, P. M. 1998. Evidence for hydrogen bonding effects in the iron ligand vibrations of carbonmonoxy myoglobin. *J. Amer. Chem. Soc.* 120:2670-71.
- (S9) Nienhaus, K., Olson, J. S., Franzen, S., and Nienhaus, G. U. 2005. The origin of stark splitting in the initial photoproduct state of MbCO. *J. Amer. Chem. Soc.* 127:40-41.
- (S10) De Angelis, F., Jarzecki, A. A., Car, R., and Spiro, T. G. 2005. Quantum chemical evaluation of protein control over heme ligation: CO/O₂ discrimination in myoglobin. *J. Phys. Chem. B.* 109:3065-70.
- (S11) Schulze, B. G., and Evanseck, J. D. 1999. Cooperative role of Arg45 and His64 in the spectroscopic A(3) state of carbonmonoxy myoglobin: Molecular dynamics simulations, multivariate analysis, and quantum mechanical computations. *J. Amer. Chem. Soc.* 121:6444-54.
- (S12) MacKerell, A. D., Brooks, B. J., Brooks, C. L., Nilsson, L. I., Roux, B., Won, Y., and Karplus, M. in P.v.R. Schleyer, *e.a.* (Eds.), *The Encyclopedia of Computational Chemistry*. John Wiley & Sons, Chichester, 1998, p. 271-77.
- (S13) Li, H. Y., Elber, R., and Straub, J. E. 1993. Molecular-Dynamics Simulation of NO Recombination to Myoglobin Mutants. *J. Biol. Chem.* 268:17908-16.
- (S14) Meuwly, M., Becker, O. M., Stote, R., and Karplus, M. 2002. NO rebinding to myoglobin: a reactive molecular dynamics study. *Biophys. Chem.* 98:183-207.
- (S15) Tannor, D. J., Marten, B., Murphy, R., Friesner, R. A., Sitkoff, D., Nicholls, A., Ringnalda, M., Goddard, W. A., and Honig, B. 1994. Accurate First Principles Calculation of Molecular Charge-Distributions and Solvation Energies from Ab-Initio Quantum-Mechanics and Continuum Dielectric Theory. *J. Amer. Chem. Soc.* 116:11875-82.
- (S16) Marten, B., Kim, K., Cortis, C., Friesner, R. A., Murphy, R. B., Ringnalda, M. N., Sitkoff, D., and Honig, B. 1996. New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. *J. Phys. Chem.* 100:11775-88.
- (S17) Vangunsteren, W. F., and Berendsen, H. J. C. 1977. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* 34:1311-27.
- (S18) Rovira, C., Kunc, K., Hutter, J., Ballone, P., and Parrinello, M. 1997. Equilibrium geometries and electronic structure of iron- porphyrin complexes: A density functional study. *J. Phys. Chem. A.* 101:8914-25.
- (S19) Reuter, N., Dejaegere, A., Maigret, B., and Karplus, M. 2000. Frontier bonds in QM/MM methods: A comparison of different approaches. *J. Phys. Chem. A.* 104:1720-35.
- (S20) Field, M. J., Albe, M., Bret, C., Proust-De Martin F., and Thomas, A. 2000. The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comp. Chem.* 21:1088-100.