Supporting Information

Paramagnetic NMR Relaxation and Molecular Mechanics Studies of Chloroperoxidase-Indole Complex: Insights into the Mechanism of Chloroperoxidase-Catalyzed Regioselective Oxidation of Indole.

Rui Zhang, Qinghao He, David Chatfield, Xiaotang Wang[*]

Department of Chemistry & Biochemistry, Florida International University, Miami, FL,33199

1. The parameterization of the heme thiolate

Parameters for the heme ligated with cysteine in the P450-cam system for use with the GROMOS force field were developed by Helms et al. in 1995 [(1)]. Unfortunately, the spin state of heme in P450-cam system is low spin (S=1/2), while the spin state of heme in CPO's resting state is high spin (S=5/2). Therefore we determined the heme-thiolate parameters appropriate for CPO's resting state.

In this work, the bonded parameters for the heme ligated with cysteine were mainly obtained from geometry optimization and the hessian matrix using Seminario's method [(2)] For bonded parameters, the heme structure was taken from a high-resolution X-ray crystal coordinates (PDB code: 2CIW, resolution at 1.15 Å) and ligated with methylthiolate (representing cysteine). The side chains of the heme were omitted to simplify the calculation, since the experimental geometries of porphyrin models can be reproduced to a high accuracy without simple side chains (e.g., alkyl and vinyl groups) [(3)]. Hydrogen atoms were added manually using the package MOLDEN[(4)], version 5.0. *Ab initio* DFT geometry optimizations at the spin-unrestricted TPSS [(5,

[6] level was performed by ORCA (version 2.9) [7] with both the def2-SVP [8] and def2-SVP/J [9] basis sets on all atoms. The TPSS functional was used because of its high accuracy and fast calculation speed for reproducing the experimental geometry [3]. The def2-SVP/J basis set is an auxiliary basis that, when used with the resolution-of-the-identity (RI) approach, speeds up the calculations significantly [9]. The vibrational frequencies were calculated by two-sided numerical differentiation of analytical gradients. The hessian matrix was analyzed by Seminario's method [2] to derive the force constants for bond stretching and angle bending. While the dihedral terms can be fit to vibrational frequencies by Seminario's method, the harmonic approximation is particularly poor for the torsional degrees of the freedom [10]. Instead, the terms were derived from a relaxed potential energy scan (PES) [11]. A series of constrained geometry optimizations was performed by freezing different N-Fe-S-C dihedral angles at each step, while the other coordinates were relaxed. The optimized structures were then energy minimized with GROMACS [12] (Groningen Machine for Chemical Simulations), version 4.5.5, using a conjugate gradient algorithm converged to a gradient of less than 5 kcal/mol/Å in which the dihedral terms were set to zero. The difference in energy between the two scans (QM and MM) was then fit to a trigonometric expansion to derive the dihedral constant. Our *ab initio* calculation indicates a total barrier of 1.50 kJ/mol, and our molecular simulation shows a non-bond energy contribution of 0.28 kJ/mol to the dihedral barrier. The bond stretching, angle and dihedral-angle parameters are listed in Table S1, and the methods for deriving these constants are shown below.

For the non-bonded parameters, van der Waals parameters for the heme were taken directly from the GROMOS53a6 force field. The atomic partial charges of the heme system were obtained by the CHELPG (CHarges from Electrostatic Potentials using a Grid based method) approach [13] as implemented in ORCA. The heme structure from the X-ray crystal coordinates

(PDB code: 2CPO) was optimized at the same level of theory with the def2-SVPD [14] basis set. For this purpose, the side chains of heme were attached to the structure. The charges of equivalent atoms were symmetrized with the constraint that the total charge on the heme-thiolate system be -2 (see Figure S1).

Table S1. Bonded parameters of ferric heme.

Groups	Equilibrium Values	Force constant
Bond	Bond Length (Å)	(kJ/(mol·nm^2))
Fe-N	2.10	4.62×10^5
Fe-S	2.30	6.35×10^5
N-C(heme)	1.38	6.18×10^5
Angle	Bending Angle (°)	(kJ/mol)
N-Fe-N	87.1	293.38
Fe-N-C	126.1	656.75
Fe-S-CH_3	105.4	216.21
S-Fe-N	103.0	265.06
Dihedral Angle	Multiplicity (n)	(kJ/mol)
N-Fe-S-CH_3	4(φ_0=0)	0.1525

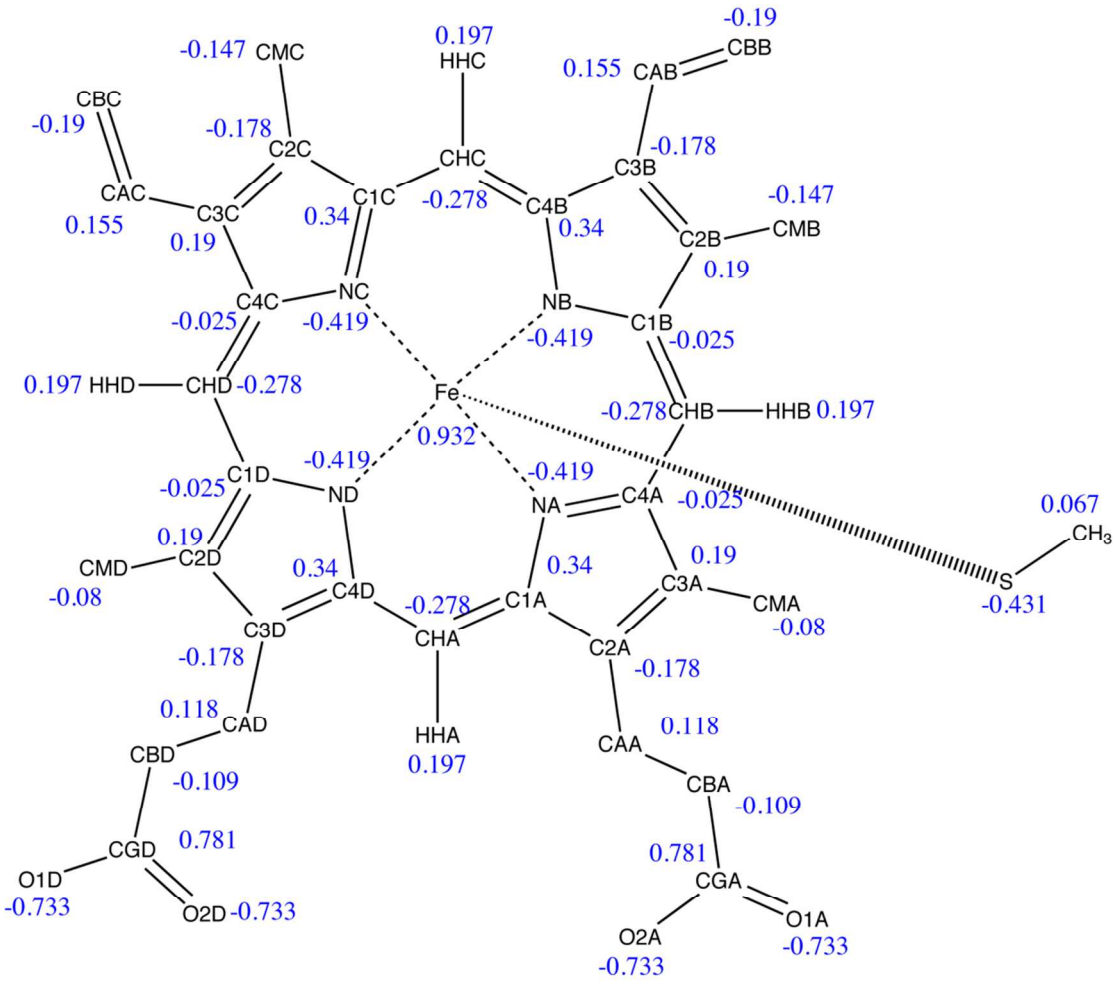

Figure S1. The structure of heme-thiolate with GROMOS atom types and the CHELPG charges (shown in blue).

2. Distances between protons of indole and the heme iron obtained from NMR relaxation at pH 5.0 and from restrained MD simulation.

Table S2. The dissociation constant (K_D) of CPO-indole complex and the distances (r) between protons of indole and heme iron derived from proton relaxation time at pH 5.0 and from restrained MD simulation.

position	δ (ppm)	T_{1b} (s)	r (Å)	K_D (mM)	MD r (Å)[a]
2-H	7.31	6.0×10^{-5}	4.1	23	3.0
3-H	6.50	4.1×10^{-4}	5.6	23	3.9
4-H	7.61	7.3×10^{-3}	9.1	22	6.5
5-H	7.05	1.5×10^{-2}	10.2	26	8.6
6-H	7.14	1.8×10^{-2}	10.5	21	9.1
7-H	7.44	9.1×10^{-3}	9.4	22	7.7

[a]The MD distances (r) were averaged ovedr the last 50 ps of the simulations.

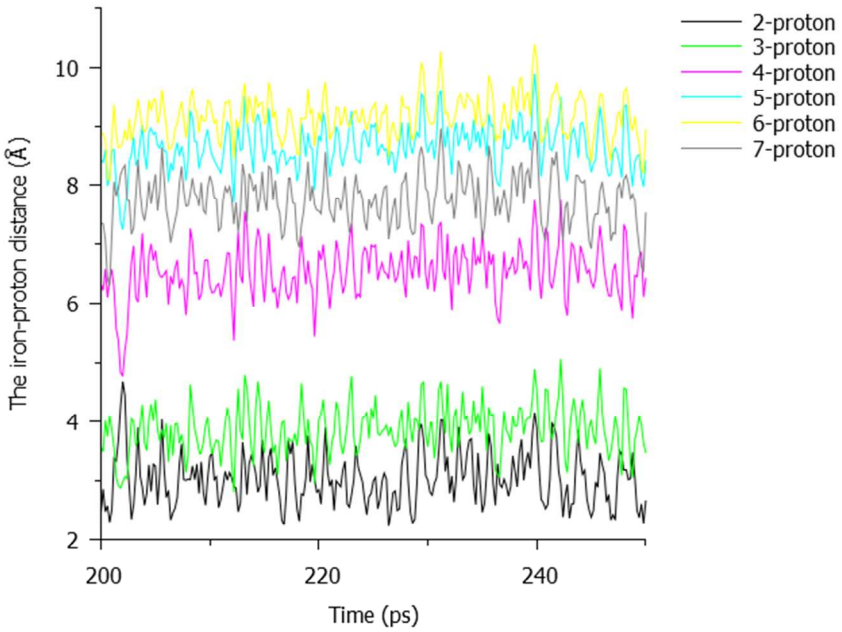

Figure S2. The plot of the MD distances (Å) between protons of indole and the heme iron as a function of time (ps).

3. Root mean-square deviations (RMSDs) and root mean-square fluctuations (RMSFs) relative to the minimized crystal structure of CPO calculated for the active-site residues during the last 50 ps of MD simulation.

Table S3. The RMSD (Å) and RMSF (Å) of the active residues averaged over the last 50 ps of simulation.

Residues	RMSD (Å)	RMSF (Å)
Ile 68	0.2	0.2
Leu 70	0.04	0.05
Asn 74	0.06	0.09
Phe 103	0.1	0.1
Ile 179	0.07	0.09
Val 182	0.3	0.2
Glu 183	0.8	0.2
Phe 186	0.1	0.1

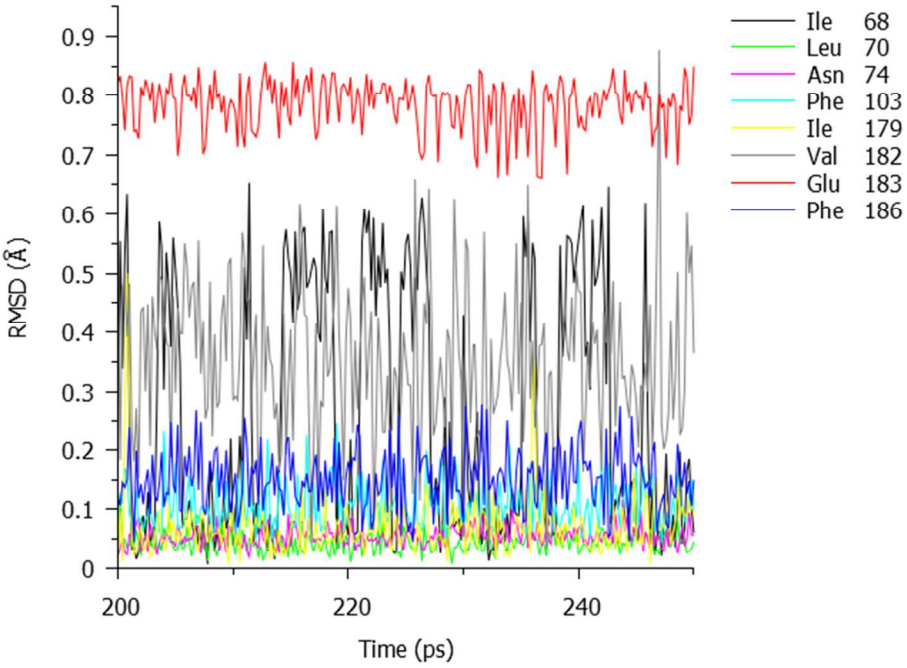

Figure S3. The RMSD (Å) of residues in the active site of CPO with respect to the starting structure during the last 50 ps of the simulation.

4. Mean potential energy calculated by GROMACS during the last 50 ps of simulated annealing.

Table S4. The calculated energy (kJ/mol) during the last 50 ps of the simulation

Residues	Average Energy[a,b]	
	Coul	LJ
Val 67	-2.1±0.3	-2.7±0.3
Ile 68	-0.4±0.03	-2.4±0.3
Leu 70	1.1±0.1	-7.2±0.1
Ala 71	1.7±0.2	-4.0±0.3
Asn 74	-1.2±0.3	-2.8±0.1
Phe103	-3.1±0.2	-10.4±0.1
Ile 179	-1.3±0.1	-8.0±0.2
Val 182	0.02±0.06	-8.8±0.1
Glu 183	-25.7±1.1	-5.6±0.2
Phe 186	-1.4±0.1	-6.6±0.2
Ala 267	-0.9±0.1	-1.6±0.2

[a]LJ, Lennard-Jones potential energy; Coul, coulomb potential energy (rcoulomb=1.2 nm, rvdw=1.4 nm). [b]Statistics over 50001 steps (150.0 through 250.0 ps).

REFERENCES

1. Helms, V., and Wade, R. C. (1995) Thermodynamics of water mediating protein-ligand interactions in cytochrome P450cam: a molecular dynamics study, *Biophys. J. 69*, 810-824.

2. Seminario, J. M. (1996) Calculation of intramolecular force fields from second-derivative tensors, *Int. J. Quantum Chem. 60*, 1271-1277.

3. Rydberg, P., and Olsen, L. (2009) The accuracy of geometries for iron porphyrin complexes from density functional theory, *J. Phys. Chem. A 113*, 11949-11953.

4. Schaftenaar, G., and Noordik, J. H. (2000) Molden: a pre- and post-processing program for molecular and electronic structures, *J. Comput. Aided. Mol. Des. 14*, 123-134.

5. Perdew, J. P., and Wang, Y. (1992) Accurate and simple analytic representation of the electron-gas correlation energy, *Phys. Rev. B Condens. Matter. 45*, 13244-13249.

6. Tao, J., Perdew, J. P., Staroverov, V. N., and Scuseria, G. E. (2003) Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids, *Phys. Rev. Lett. 91*, 146401.

7. Neese, F., and Wennmohs, F. *ORCA- an ab initio, DFT and semiempirical SCF-MO package*.

8. Weigend, F., and Ahlrichs, R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *Phys. Chem. Chem. Phys. 7*, 3297-3305.

9. Weigend, F. (2006) Accurate Coulomb-fitting basis sets for H to Rn, *Phys. Chem. Chem. Phys. 8*, 1057-1065.

10. Burger, S. K., Lacasse, M., Verstraelen, T., Drewry, J., Gunning, P., and Ayers, P. W. (2012) Automated parametrization of AMBER force field terms from vibrational analysis with a focus on functionalizing dinuclear Zinc (II) scaffolds, *J. Chem. Theory Comput. 8*, 554-562.

11. Autenrieth, F., Tajkhorshid, E., Baudry, J., and Luthey-Schulten, Z. (2004) Classical force field parameters for the heme prosthetic group of cytochrome c, *J. Comput. Chem. 25*, 1613-1622.

12. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, *J. Chem. Theory Comput. 4*, 435-447.

13. Breneman, C. M., and Wiberg, K. B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis, *J. Comput. Chem.*, 361-373.

14. Rappoport, D., and Furche, F. (2010) Property-optimized gaussian basis sets for molecular response calculations, *J. Chem. Phys. 133*, 134105.