

**A DFT and ONIOM study of C–H hydroxylation catalyzed
by nitrobenzene 1,2-dioxygenase**

*Inacrist Geronimo and Piotr Paneth**

Institute of Applied Radiation Chemistry, Lodz University of Technology,

Żeromskiego 116, 90-924 Łódź, Poland

Preparation of nitrobenzene 1,2-dioxygenase-2-nitrotoluene (NBDO-2NT) complex and molecular dynamics (MD) simulation

The biological assembly of NBDO (with nitrobenzene substrate, PDB ID: 2BMQ),¹ an $\alpha_3\beta_3$ hetero-hexamer, was generated using MakeMultimer.² pKa values were determined using propKa 3.1.³⁻⁶ Asp and Glu were assigned as negative, and Arg and Lys as positive. The deprotonated form was used for Cys-79 and Cys-99, despite having pKa values greater than 7 since these are bonded to one of the Fe atoms of the Rieske cluster. Histidine residues, including those bound to one of the Fe atoms of the Rieske cluster (His-81 and His-102), were assigned as HIE (N_ϵ protonated). The exceptions are His-206 and His-211, which were assigned as HID (N_δ protonated), since these are bound to the Fe active site through the N_ϵ atom. Missing heavy and hydrogen atoms were added using the *tleap* module of AmberTools13.⁷

Docking of 2NT in the prepared enzyme-substrate complex was performed using Glide.⁸⁻¹¹ The receptor grid was generated with positional constraint of 2.5 Å radius around the carbon closest to the Fe center, and H-bond constraint defined between (a) nitro O atom of the substrate and amide H atom of Asn-258, and (b) nitro O atom and H atom of HOH-2401. The two water molecules coordinated to Fe in the active site (HOH-2227 and HOH-2401) were then replaced by a hydroperoxo ligand, with coordinates taken from the O₂ adduct of the naphthalene dioxygenase(NDO)-indole complex (PDB ID: 1O7N).¹² This was done after docking of the substrate since the latter is known to bind to the enzyme before O₂.¹³ The system was neutralized with 15 Na⁺ counterions and solvated in a periodic cubic box of TIP3P¹⁴ water, with a buffer distance of 12 Å between each wall and the closest atom in each direction.

MD simulations with the AMBER ff99SB force field,¹⁵ combined with newly developed parameters for the active site and Rieske cluster,¹⁶ were done using the AMBER 12 suite of programs.⁷ Periodic boundary conditions were applied using the Particle Mesh Ewald (PME) method,¹⁷ with a non-bonded cutoff of 10 Å. Energy minimization was done in two stages, where the solute was initially restrained with a harmonic force constant of 2.0 kcal•mol⁻¹Å⁻² to allow water and ions to relax, after which the entire system was minimized. The first stage was run for 1000 steps and the second for 2000 steps, with the first half of each stage done using the steepest descent method and the last half using the conjugate gradient algorithm. The system was then heated from 0 to 300 K for 50 ps, again with the solute restrained as in the first stage of minimization. Bonds involving hydrogen were constrained using the SHAKE algorithm¹⁸ and a 1 fs time step was used. The temperature was controlled using Langevin dynamics¹⁹ with a collision frequency of 1.0 ps⁻¹. NMR weight restraints were also used to linearly increase the temperature to avoid instabilities in the simulation. Finally, the system was equilibrated without restraints at constant pressure for 200 ps. Isotropic position scaling was used to maintain the pressure at 1 atm, with a relaxation time of 2 ps. The production phase was run for 2.5 ns using the same parameters as equilibration. Analysis of the trajectory was done using the *ptraj* module of AmberTools13.⁷

It was demonstrated in our recently published MD study of NBDO with nitrobenzene substrate that the enzyme maintains its stability throughout the simulation and that deviation from the initial crystal structure is not significant. Hydrogen bond interactions that may be important to catalytic activity were also identified.¹⁶ For the purpose of the present study, a shorter simulation of 2.5 ns was performed to relax the enzyme structure after docking of the substrate and addition of the hydroperoxo ligand prior to ONIOM calculations. Root-mean-square deviations (RMSD) of backbone and heavy atoms of the substrate pocket and the entire protein with respect to the crystal structure as a function of simulation time are shown in Fig. S1. The average structure for the last 2 ns was obtained and the calculated backbone and heavy atom RMSDs with respect to the crystal structure are 0.8 and 1.1 Å, which indicate

that the enzyme structure did not change significantly during the simulation. Initial structures for ONIOM calculations were taken from the trajectory within this time frame.

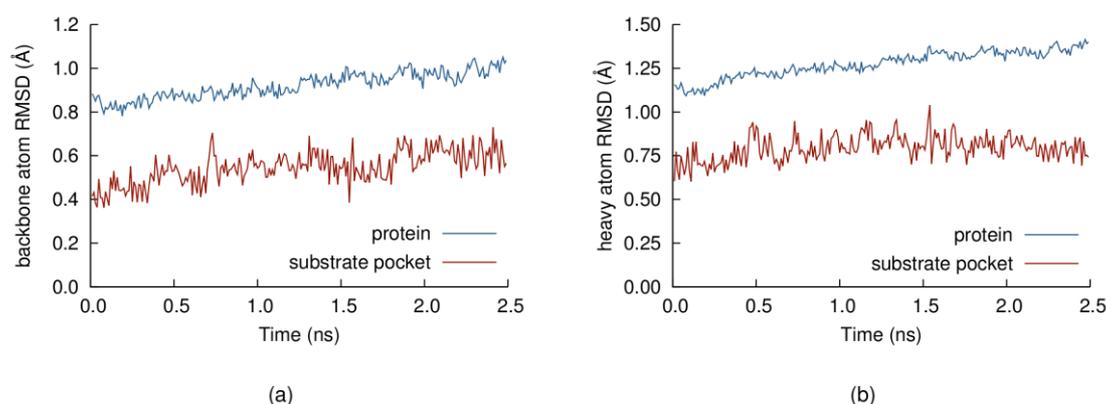


Fig. S1 (a) Backbone (CA, C, N) and (b) heavy atom root-mean-square deviation (RMSD) of the substrate pocket and the entire protein with respect to the crystal structure as a function of simulation time

Table S1 Hydrogen bond interactions formed with the active site region during the simulation reported as percent occupancy

acceptor ^a	donor ^a	% occupancy ^b
<i>His-206</i>		
His206:O	Thr210:H	99.60
Asp203:O	His206:HD1	96.39
His206:O	Trp209:H	87.95
Asp203:OD1	His206:H	46.18
Asp203:OD1	His206:HD1	44.58
Asp203:O	His206:H	30.92
Asp203:OD2	His206:H	14.86
Asp203:OD2	His206:HD1	12.45
<i>His-211</i>		
His211:O	Leu215:H	98.80
Val207:O	His211:H	73.90
HOH2338:O	His211:HD1	57.43
Asp359:OD1	His211:HD1	42.17
His211:O	Ala214:H	36.14
<i>Asp-360</i>		
Asp360:O	Asn363:H	99.20
Glu357:O	Asp360:H	90.76
Asp360:O	Met364:H	74.30
Asp360:OD2	HOH2339:H2	20.48
Asp360:OD2	HOH2339:H1	18.88
Asp360:OD1	HOH2334:H2	12.85
<i>HPO</i>		
HOH2334:O	HPO:H3	66.27
HOH2340:O	HPO:H3	32.13
HPO:O1	Asn199:HD21	31.33
<i>2NT</i>		
2NT:O1	Asn258:HD22	55.02
2NT:O2	Asn258:HD22	42.17

^a abbreviations: HPO: hydroperoxo ligand, HOH: crystallographic water, 2NT: 2-nitrotoluene

^b percent of time the hydrogen bond is formed over the trajectory

Hydrogen bond interactions formed by the active site region are summarized in Table S1. A more extensive analysis of hydrogen bonding, particularly on the interface of the active site and the adjacent Rieske cluster and the network of water molecules at the entrance of the active site, can be found in ref. 16. Most of the interactions listed in Table S1 are between backbone atoms given that the active site region is mostly hydrophobic. Sidechain interactions include those with Asp-203, Asn-199 and Asn-258. Hydrogen bonding with crystallographic waters (HOH) was also observed for the hydroperoxo ligand (HPO), His-211 and Asp-360.

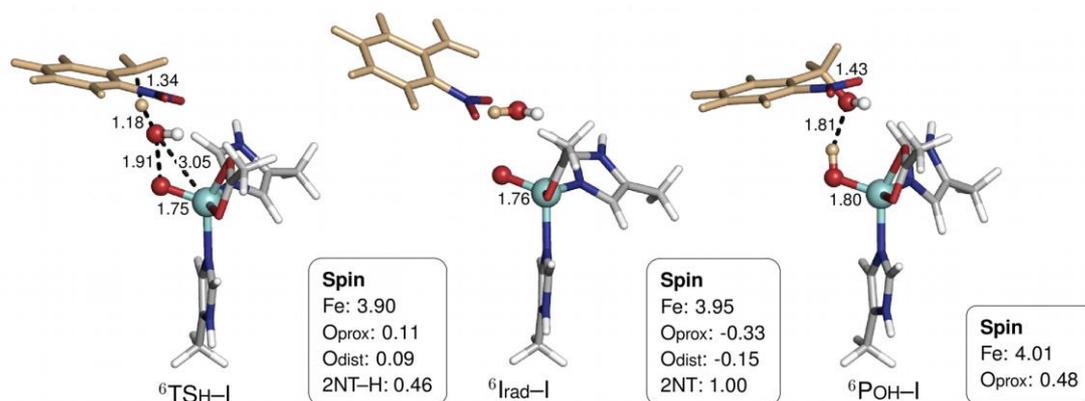


Fig. S2 Optimized geometries and spin populations for stationary points in mechanism I (S=5/2). Distances are given in Å

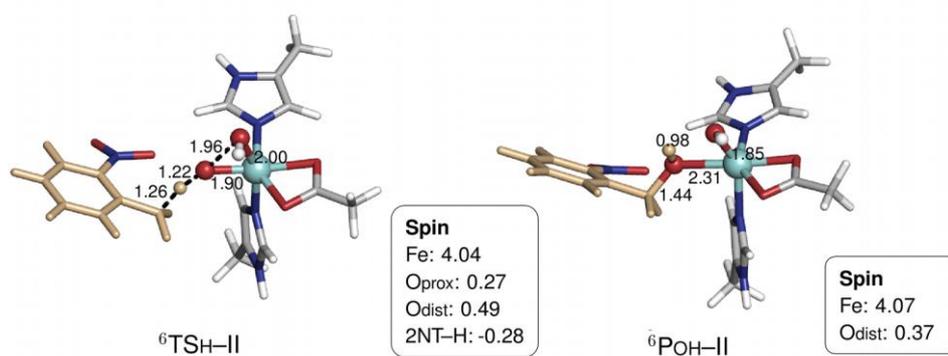


Fig. S3 Optimized geometries and spin populations for stationary points in mechanism II (S=5/2). Distances are given in Å

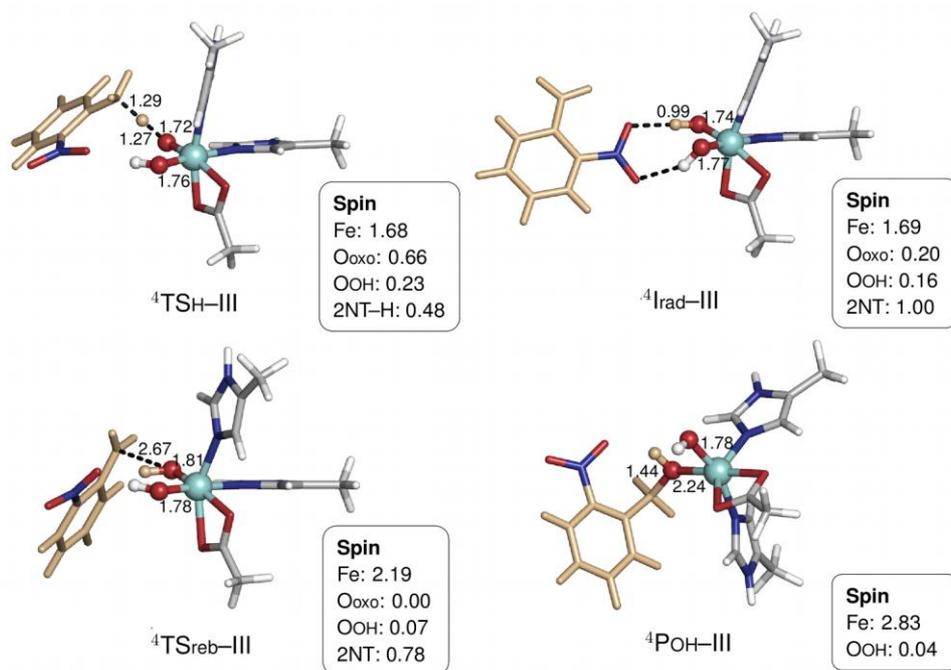


Fig. S4 Optimized geometries and spin populations for stationary points in mechanism III (S=3/2). Distances are given in Å

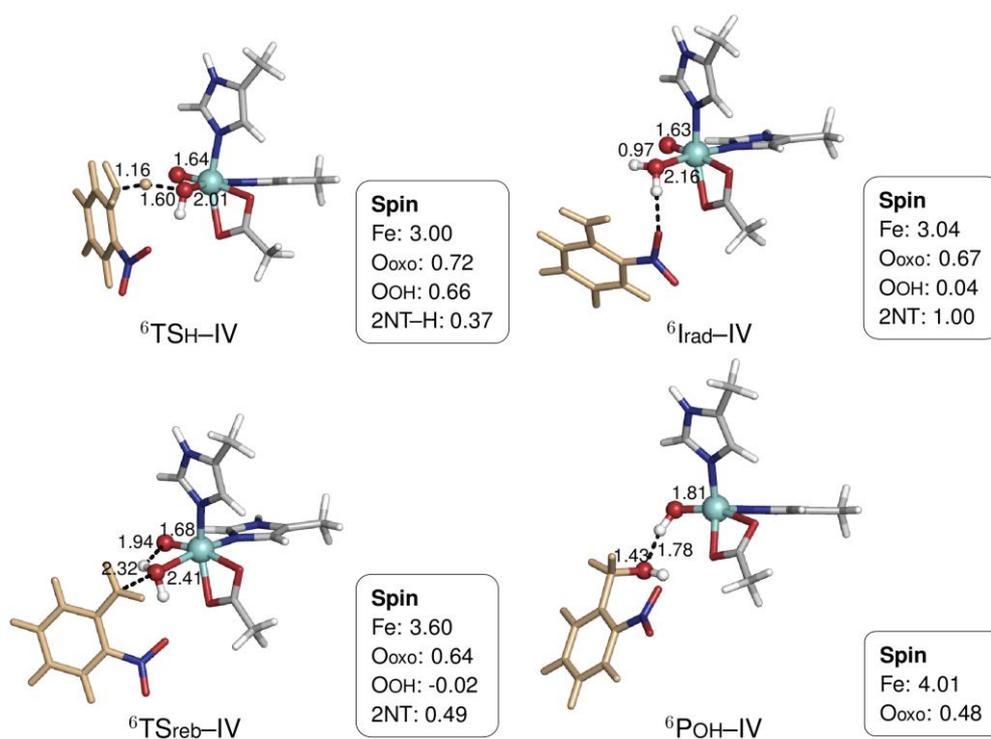


Fig. S5 Optimized geometries and spin populations for stationary points in mechanism IV (S=5/2). Distances are given in Å

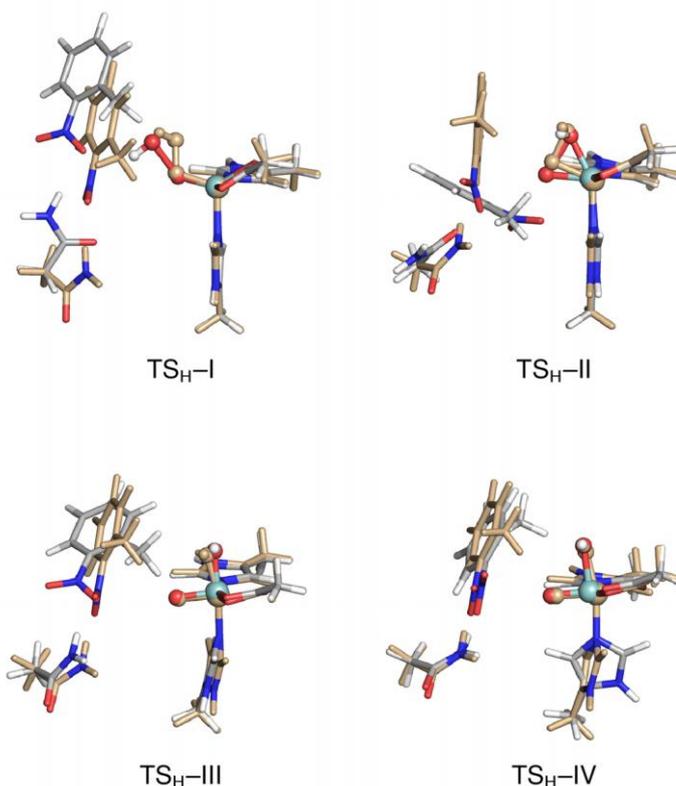


Fig. S6 Optimized M2 geometries for H abstraction transition state (grey) superimposed on reactant complex (beige), mechanisms I-IV

Table S2 Variation of O-O bond cleavage barrier (kcal/mol)^a with initial structure calculated using ONIOM-ME method

1 ^b	22.46
2	25.00
3	24.82
4	24.88
5	25.71
6	24.49

^a energies determined using LACV3P+* basis set with zero-point correction at LACVP* level

^b value reported in the main text

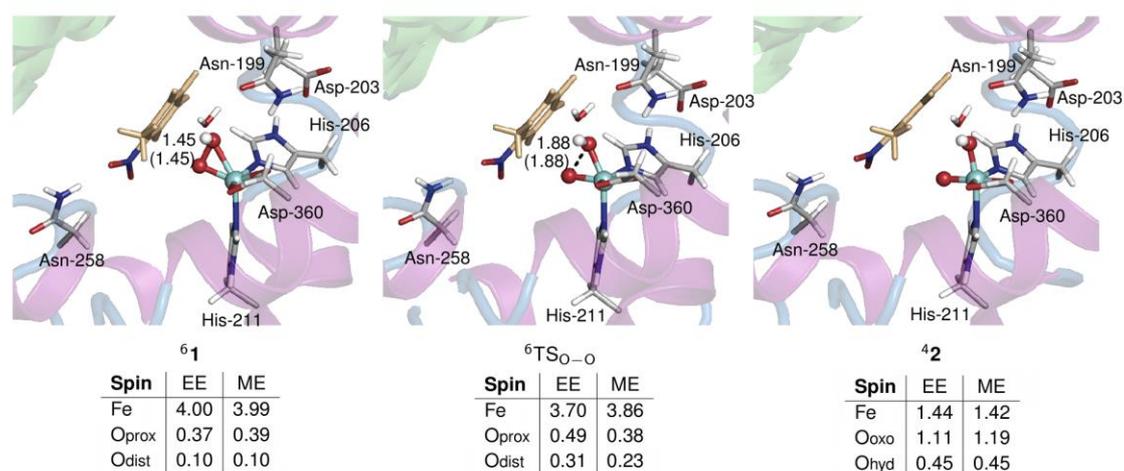


Fig. S7 Optimized geometries and spin populations of stationary points in O-O bond cleavage step calculated using mechanical (ME) and electronic (EE) embedding schemes. Distances (Å) in

parenthesis obtained using ME scheme. Residues and a water molecule forming hydrogen bonds with the QM region also shown

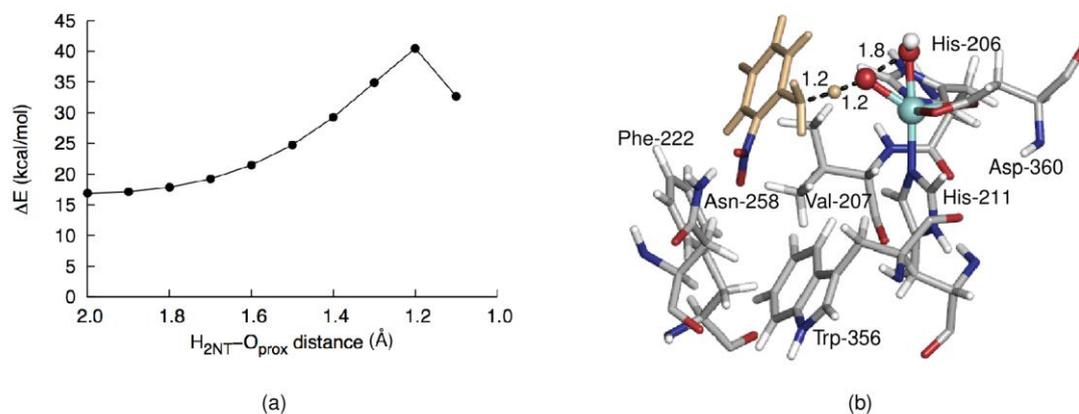


Fig. S8 (a) Potential energy scan along the H-O reaction coordinate, with O-O-H and O-H-C angles fixed at 170° and 175°, respectively; (b) Geometry at the highest point of the curve ($d_{\text{H-O}} = 1.2 \text{ \AA}$), which has an energy of 40.5 kcal/mol

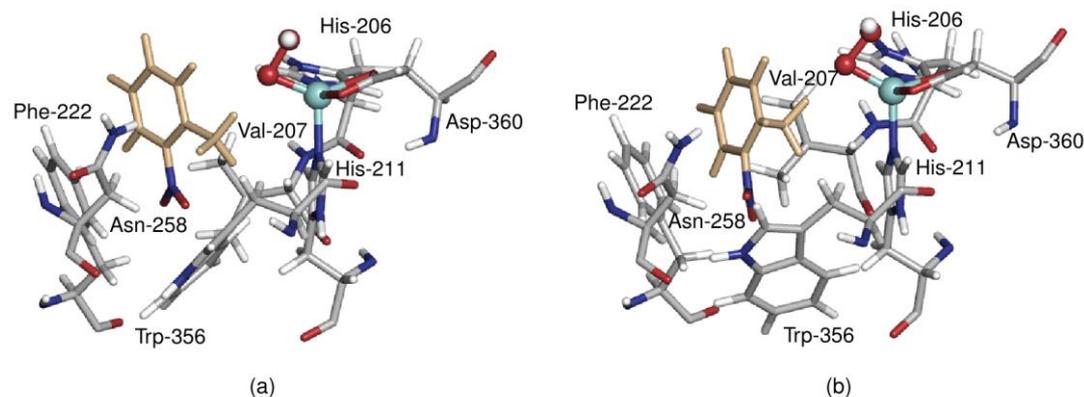


Fig. S9 Two other structures from the simulation at $d_{\text{H-O}} = 1.9 \text{ \AA}$, $\alpha_{\text{O-O-H}} = 170^\circ$ and $\alpha_{\text{O-H-C}} = 175^\circ$ showing steric hindrance from other residues that prevents rotation of the substrate to achieve maximum overlap between its π orbital and the O-O σ^* orbital of the hydroperoxy ligand

Table S3. H abstraction transition state (${}^4\text{TS}_\text{H}$) obtained using different initial structures. Calculations done using the ONIOM-ME method

	ΔE^\ddagger (kcal/mol) ^a	bond lengths (Å)		spin densities			
		C-H	O-H	Fe	O _{oxo}	O _{hyd}	2NT
1 ^b	7.24	1.23	1.39	1.51	0.85	0.21	0.49
2	9.09	1.24	1.36	1.60	0.66	0.14	0.65
3	8.33	1.23	1.38	1.50	0.87	0.24	0.44
4	8.61	1.24	1.37	1.52	0.84	0.21	0.46

^a energies determined using LACV3P+* basis set with zero-point correction at LACVP* level

^b value reported in the main text

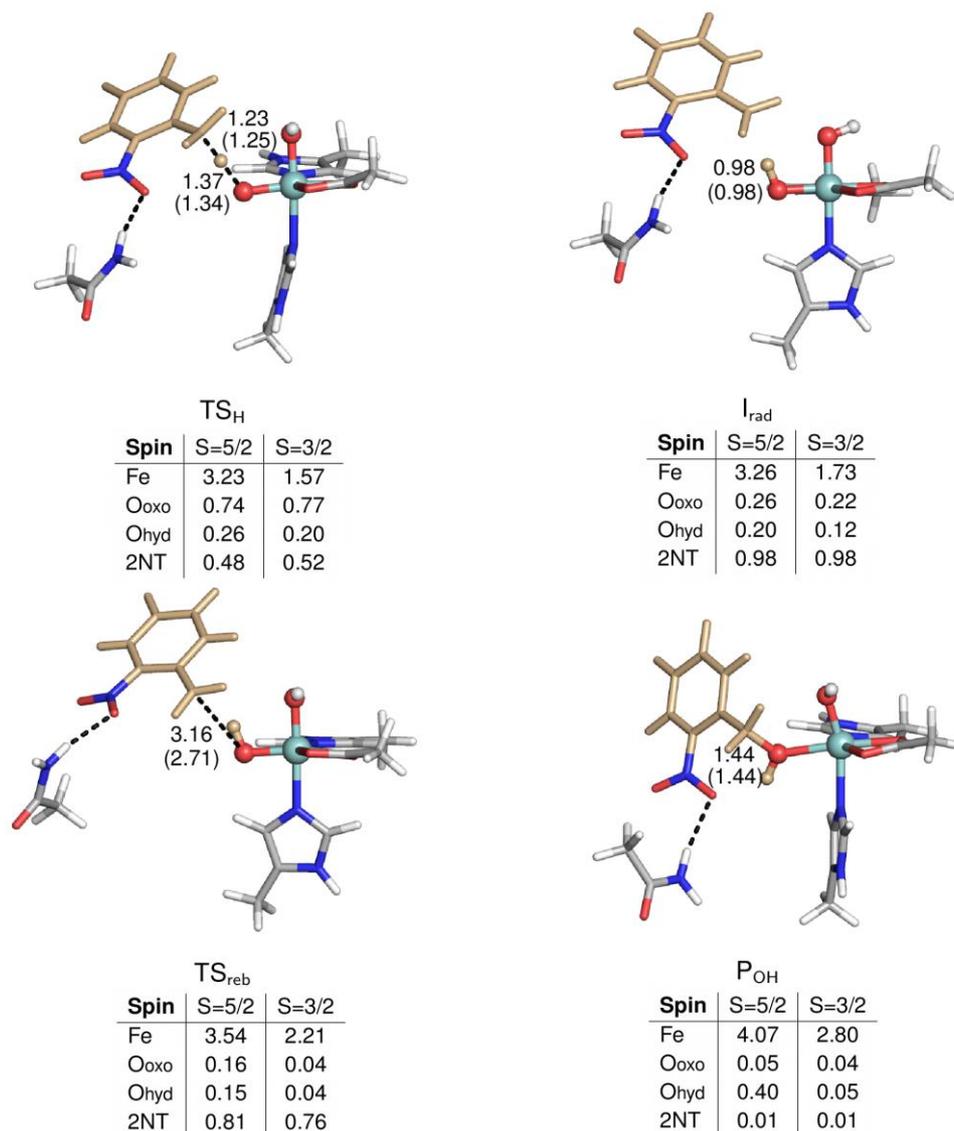


Fig. S10 Optimized M2 geometries and spin populations of stationary points in C–H hydroxylation step. Distances (Å) in parenthesis are for S=3/2

References

- 1 R. Friemann, M. M. Ivkovic-Jensen, D. J. Lessner, C.-L. Yu, D. T. Gibson, R. E. Parales, H. Eklund and S. Ramaswamy, *J. Mol. Biol.*, 2005, **348**, 1139–1151.
- 2 MakeMultimer.py, <http://watcut.uwaterloo.ca/tools/makemultimer/>, (accessed November 2012).
- 3 H. Li, A. D. Robertson and J. H. Jensen, *Proteins*, 2005, **61**, 704–721.
- 4 D. C. Bas, D. M. Rogers and J. H. Jensen, *Proteins*, 2008, **73**, 765–783.
- 5 M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
- 6 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284–2295.
- 7 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J.

- Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, T. L. V. Babin, S. Gusarov, A. Kovalenko and P. A. Kollman, AMBER 12, University of California, San Francisco, 2012.
- 8 Glide, version 5.6, Schrödinger, LLC, New York, NY, 2010.
 - 9 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, D. E. Shaw, M. Shelley, J. K. Perry, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.
 - 10 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.
 - 11 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177–6196.
 - 12 A. Karlsson, J. V. Parales, R. E. Parales, D. T. Gibson, H. Eklund and S. Ramaswamy, *Science*, 2003, **299**, 1039–1042.
 - 13 T. Ohta, S. Chakrabarty, J. D. Lipscomb and E. I. Solomon, *J. Am. Chem. Soc.*, 2008, **130**, 1601–1610.
 - 14 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
 - 15 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712–725.
 - 16 A. Pabis, I. Geronimo, D. M. York and P. Paneth, *J. Chem. Theory Comput.*, 2014, DOI: 10.1021/ct500205z.
 - 17 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
 - 18 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
 - 19 R. W. Pastor, B. R. Brooks and A. Szabo, *Mol. Phys.*, 1988, **65**, 1409–1419.