J|A|C|S
COMMUNICATIONS

Published on Web 12/18/2001

Effects of the Protein Environment on the Structure and Energetics of Active Sites of Metalloenzymes. ONIOM Study of Methane Monooxygenase and Ribonucleotide Reductase

Maricel Torrent, Thom Vreven,‡ Djamaladdin G. Musaev, and Keiji Morokuma*

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

Ödön Farkas§ and H. Bernhard Schlegel

Department of Chemistry, Wayne State University, Detroit, Michigan 48202

Received July 11, 2001

Most theoretical studies to date on transition-metal complexes in biological systems have been performed with quantum chemical methods for models consisting of up to 50 atoms.[1] With such limited models, the question remains whether inclusion of the surrounding protein in the calculation may change the results. Recent development of hybrid methods, such as the QM/MM (quantum mechanics/molecular mechanics) method[2] as well as the more general ONIOM (our own *n*-layered integrated molecular orbital and molecular mechanics) method,[3] has made larger molecular systems within reach of accurate calculations.[4]

Two members of the non-heme family, methane monooxygenase and ribonucleotide reductase, possess very similar active sites in the units called MMOH and R2, respectively.[5] X-ray structures of both the oxidized forms, $MMOH_{ox}$ and $R2_{met}$ (resting state), and the reduced forms, $MMOH_{red}$ and $R2_{red}$ (active state), are now available.[6] ONIOM studies of these systems are expected to shed new insights beyond model studies performed to date.[7]

We have recently developed new efficient ONIOM2(QM:MM) and ONIOM3(QM:QM:MM) codes, implemented in the development version of Gaussian,[8] and present here its first application to the models of metalloenzymes in which a large number of protein residues are added in the calculation. The main novel features of our code are the double iteration geometry optimization scheme for ONIOM(QM:MM) calculations, including freezing a selection of the MM atoms,[9] and the inclusion of the QM-MM interaction at either the classical MM level (mechanical embedding) or the QM level (electronic embedding).[10]

The left section of Table 1 gives the results of density functional (DFT) geometry optimization at the B3LYP/lanl2dz level of the "active site only" models of these four forms of the enzymes. These model active sites contain 39−46 atoms (depending on the species, in column 2), consisting of two Fe centers and the first shell ligands of four formates, two imidazoles, and a few oxo, hydroxo, and/or aquo groups (see Figure S1 in Supporting Information).[11] When one superposes the B3LYP-optimized "active site only" structures on the experimental structures, as shown in Figure 1 A for $R2_{met}$, there are substantial differences in geometry of the active site. The spatial orientations of rings and bonds are causing these differences; in particular, the imidazole rings (pointing forward in Figure 1) rotate away from the experimental direction.

The difference between the B3LYP-optimized structure and the experiment has been quantitized as the root-mean-square (RMS) and maximum deviations by using Tinker v3.8.[12,13] Only the mass-weighted coordinates of non-hydrogen atoms of the active site were included. As can be seen from Table 1, RMS deviations of non-hydrogen atoms are in the range 0.623−1.029 Å. The maximum deviations range from 1.10 to 2.62 Å. The agreement is qualitatively good, suggesting that the "active site only" model is a reasonable starting point for theoretical study of these metalloenzymes.[14]

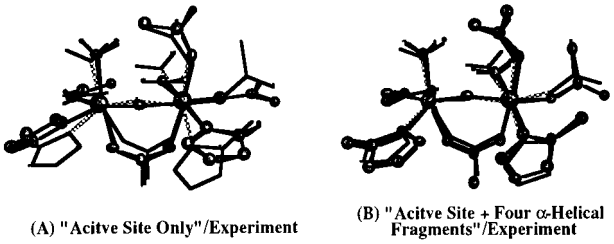

Figure 1. Superposition of (A) the B3LYP-optimized and experimental (thin wireframe drawing) structures and (B) the ONIOM2-optimized and experimental structures of the active site of $R2_{met}$.

Next we studied the "active site + four α-helical fragments" models of these enzymes, using ONIOM2 combining B3LYP/lanl2dz and Amber.[15] We included 62 residues within the Fe-centered spheres of radii of 15 and 18 Å for MMOH and R2, respectively, containing only the central parts of helices B, C, E, and F.[16] Hydrogen atoms were added to the selected fragments from the PDB structures with use of GaussView,[17] followed by further Amber force field minimization; this was used as the starting structure for the ONIOM optimizations. The total number of atoms included is shown in column 5 of Table 1, which shows that the B3LYP part is the same to the "active site only" model. Preliminary full ONIOM2 geometry optimization gave unrealistic expansions of the protein structure, most likely due to the lack of the remainder of the enzyme. Therefore, the structure was optimized keeping all main-chain atoms of the considered helical fragments frozen at the experimental position.[18] The results are shown in section B of Table 1.

The most notable in comparison of the "active site + four α-helical fragments" results with the "active site only" results above is that the new optimized structures agree with the experiment much more closely than the preceding results. RMS and maximum deviations are reduced from 0.62−1.02 to 0.34−0.48 Å, and from 1.01−2.62 to 0.51−0.81 Å, respectively (in column 8). Figure 1B shows this very clearly; now all the active-site atoms are in close vicinity of the experimental position. Inclusion of the aspartates that are hydrogen bonded to imidazoles eliminates the ring rotation. One can conclude that the local structure of active site of these metalloenzymes is in part controlled by the protein environment.

‡ Current address: Gaussian, Inc., 140 Washington Ave., North Haven, CT 06473.
§ Department of Organic Chemistry, Eötvös Loránd University of Budapest, H-1518 Budapest, 112 P.O. Box 32, Hungary.

10.1021/ja016589z CCC: $22.00

Table 1. Total Energies (au), Deformation Energies (kcal/mol), and Deviations from Experiment (Å) for Active Site Models of R2 and MMOH

	(A) "active site only" model B3LYP			(B) "active site + four α-helical fragments" model						
				ONIOM2(B3LYP:Amber)[b]					ONIOM3 (B3LYP:HF:Amber)[b]	
enzyme	no. of atoms	energy of the active site	RMS/max dev (Å)	# of atoms	energy of the active site	deformation	RMS/max dev (Å)	RMS/max dev (Å)[c]	no. of atoms	RMS/max dev (Å)
$R2_{met}$	43	−1684.177591	0.623/1.099	43:1034	−1684.145609	20.1	0.335/0.519	0.367/0.620	43:88:1034	0.230/0.359
$R2_{red}$	39	−1532.507052	0.641/1.308	39:1030	−1532.488915	11.4	0.344/0.747	0.393/0.993		
$MMOH_{ox}$	46	−1760.645165	1.029/1.940	46:987	−1760.570193	47.0	0.478/0.641	0.501/0.934		
$MMOH_{red}$	42	−1609.000745	1.021/2.619	42:983	−1608.946283	34.2	0.358/0.809	0.336/0.569		

[a] The mass-weighted root-mean-square deviations of non-H atoms of the active site from the experiment. See text for details. [b] Geometry optimization performed under the mechanical embedding with all main-chain atoms frozen at the experimental geometry (electronic embedding results are similar). Number of atoms includes both DFT (in parentheses) and Amber atoms. See text for the definition of deformation energy. [c] The results of geometry optimization with only the eight terminal C_α atoms frozen.

We also performed geometry optimizations with only the eight terminal C_α atoms of the four helical fragments fixed at the experimental position. The resultant RMS and maximum deviations (column 9, Table 1) are 0.34−0.50 Å and 0.57−0.99 Å, only slightly larger than those with the frozen main-chain atoms. This suggests that once these four helical chains in the immediate vicinity are considered, the active site structure is self-maintained. We also performed optimization for $R2_{met}$ using the three-layer ONIOM3-(B3LYP: HF/STO-3G: Amber) method with all the main-chain atoms fixed. We used the atomic partition (43:88:1034), i.e., in addition to the ONIOM2 above, the intermediate model system, consisting of 43 active site atoms and 45 side-chain and C_α atoms of the first coordination shell,[19] was treated at the Hartree−Fock/STO-3G level. The resultant RMS and maximum deviations (last column, Table 1) are 0.230 and 0.359 Å, respectively, compared to 0.335 and 0.519 Å above. This indicates clearly that the electronic effects of the protein residues, neglected in the ONIOM2-(B3LYP: Amber) treatment, significantly improve the agreement and suggests the usefulness of the ONIOM3 method.

Concerning the B3LYP energies of the active site in Table 1, that of the "active site + four α-helical fragment" structure (column 6) is always higher than that of the "active site only" structure (column 3), since the latter fully optimized "gas phase" structure is deformed and destabilized by the effect of the protein environment by the amount shown as "deformation" (in column 7, in kcal/mol). The amount of deformation energy is substantial, suggesting that the active site is highly strained. This strain is crucial for the reactivity; the strained configuration acts as a device for energy storage (which will be used in later steps of the cycle).

The flexibility of the active site was already assessed in a previous study.[14] Here, the most intriguing finding in the deformation energy is that for both enzymes the destabilization is larger in the oxidized form than in the reduced form (by 8.7 kcal/mol for R2 and 12.8 kcal/mol for MMOH). This indicates that the thermodynamics of the active part in going from the resting to the active state is made more exothermic by the protein environment. Since the protein models used in the present calculations for the resting and active states are stereochemically different, one cannot draw conclusions whether the entire reaction step has become energetically more favorable by the protein environment. This point will be discussed in detail in the future.

In summary, the present work illustrates the importance of including the protein environment beyond the core of the active site in these enzymes. Indeed, the calculations have demonstrated that the interaction of active site with protein highly strains the former, and that these strain energies are different for different forms of the enzymes. These findings indicate the crucial role of the protein-active site interaction in the enzymatic activities. The usefulness of the ONIOM approaches was also demonstrated.

Acknowledgment. The authors acknowledge grants (CHE-9627775 to K.M. and CHE-9874005 to H.B.S.) from the National Science Foundation, grants from Gaussian, Inc., and generous support of computing time from NCSA.

Supporting Information Available: Figure S1 showing the selected atomic frames to model the four active sites (full model in DFT and QM region in ONIOM); a list of the added Amber parameters is also included (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Torrent, M.; Musaev, D. G.; Basch, H.; Morokuma, K. *J. Comput. Chem.* **2001**, *23*, 59−76.
(2) (a) Warshel, A.; Levitt, M. *J. Mol. Biol*. **1976**, *103*, 227−249. (b) Singh, U. C.; Kollman, P. A. *J. Comput. Chem*. **1986**, *7*, 718−730. (c) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem*. **1990**, *11*, 700−733. (d) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 7, pp 119−185. (e) Mordasini, T. Z.; Thiel, W. *Chimia* **1998**, *58*, 288−291.
(3) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem*. **1995**, *16*, 1170−1179. (b) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem*. **1996**, *100*, 19357. (c) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. THEOCHEM* **1999**, *461*−*462*, 1. (d) Vreven, T.; Morokuma, K. *J. Comput. Chem*. **2000**, *16*, 1419−1432.
(4) In contrast to QM/MM schemes, ONIOM can combine QM and QM, as well as QM and MM, in multiple layers.
(5) Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S.-K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y.-S.; Zhou, J. *Chem. Rev*. **2000**, *100*, 235−349 and references therein.
(6) (a) Whittington, D. A.; Lippard, S. J. *J. Am. Chem. Soc*. **2001**, *123*, 827−838. (b) Logan, D. T.; Su, X. D.; Åberg, A.; Regnström, K.; Hajdu, J.; Eklund, H.; Nordlund, P. *Structure* **1996**, *4*, 1053−1064. (c) Nordlund, P.; Eklund, H. *J. Mol. Biol*. **1993**, *232*, 123−164.
(7) (a) Basch, H.; Mogi, K.; Musaev, D. G.; Morokuma, K. *J. Am. Chem. Soc.* **1999**, *121*, 7249−7256. (b) Basch, H.; Musaev, D. G.; Mogi, K.; Morokuma, K. *J. Phys. Chem. A* **2001**, *105*, 3615−3622. (c) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2001**, *6*, 27−45. (d) Dunietz, B. D.; Beachy, M. D.; Cao, Y.; Whittington, D. A.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc*. **2000,** *122*, 2828−2839. (e) Yoshizawa, K. *J. Inorg. Biochem*. **2000**, *78*, 23−34.
(8) *Gaussian 99*, Development Version (Revision B.09+), Frisch, M. J.; et al. Gaussian, Inc.: Pittsburgh, PA, 2000.
(9) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* Submitted for publication.
(10) Vreven, T.; Komáromi, I.; Dapprich, S.; Byun, K. S.; Montgomery, J. A., Jr.; Morokuma, K.; Frisch, M. J. In preparation.
(11) Ferromagnetic coupling was assumed in all cases. Spin multiplicity = 11 (oxidized forms) and 9 (reduced forms).
(12) http://dasher.wustl.edu/tinker.
(13) (a) Kearsley, S. J. *J. Comput. Chem*. **1990**, *11*, 1187−1192. (b) Diamond, R. *Acta Crystallogr. A* **1988**, *44*, 211−216. (c) McLachlan, A. D. *Acta Crystallogr. A* **1982**, *38*, 871−873. (d) Nyburg, S. C. *Acta Crystallogr. B* **1974**, *30*, 251−253.
(14) Torrent, M.; Musaev, D. G.; Morokuma, K. *J. Phys. Chem. B* **2001**, *105*, 322−327.
(15) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(16) Selected protein fragments: (1) For MMO, αB 106−120, αC 140−154, αE 203−218, and αF 234−249. (2) For R2: αB 76−90, αC 111−125, αE 198−210, and αF 227−245. Protonation states for individual residues were chosen according to p*K* values and experimental pH (7). New terminal residues were turned into N_t (cationic) and C_t (anionic) residues.
(17) http://www.gaussian.com/gvbroc.htm.
(18) Added Amber parameters for link atoms are given in the Supporting Information. Charges for the peptide fragments are taken from ww-w.amber.ucsf.edu/amber/ff94.
(19) D84, E115, E118, E204, E238, and H238.

JA016589Z