Does the Active Site of Mammalian Glutathione Peroxidase (GPx) Contain Water Molecules? An ONIOM Study

Rajeev Prabhakar, Djamaladdin G. Musaev,* Ilja V. Khavrutskii, and Keiji Morokuma

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, Georgia 30322

Received: June 21, 2004; In Final Form: July 8, 2004

In the present theoretical study, the active site of mammalian glutathione peroxidase (GPx) has been refined by including protein/active site interactions using the two-layer ONIOM(QM:MM) method. In this study, a full model consisting of 3113 atoms (with 86 active site atoms in the quantum mechanical (QM) region and the rest in the molecular mechanics (MM) region) has been constructed on the basis of the 2.9 Å resolution X-ray structure of mammalian GPx. The "active site only" models give much larger root-mean-square (rms) deviations from the X-ray structure than the full model, indicating the importance of the protein environment on the structure of the active site. The still substantial rms errors of the optimized structures can be improved only when the full model complemented with two water molecules is considered, clearly indicating the existence of two missing water molecules at the active site of the mammalian GPx, which could be critical for the catalytic activity.

Glutathione peroxidase (GPx) constitutes a family of selenoproteins, which demonstrates a strong antioxidant activity and protects cell membranes and other cellular components against oxidative damage.[1,2] This enzyme reduces reactive oxygen species (ROS) like hydrogen peroxide and organic peroxides by utilizing glutathione as the reducing substrate[1]. Three different classes of Se-dependent GPx (cytoplasmic, plasma, and phospholipid hydroperoxide) have been known, while crystal structures of only bovine erythrocyte (intracellular enzyme) and human plasma (extracellular enzyme) GPx have been resolved at 2.0 and 2.9 Å resolutions, respectively.[3,4] The complete X-ray structure of only human plasma GPx (2.9 Å) is available in the literature.[4] The X-ray structures show that the enzyme is a tetramer, with two asymmetric units containing two dimers. Each of the dimers has two selenocysteine residues at their active sites (see Figure 1A). Although numerous experimental studies provide a significant amount of information regarding the structure and the catalytic activity of these important enzymes, still, several issues concerning the structure (existence of water molecules at the active site), the catalytic mechanism (nature of intermediates and transition states), and protein/active site interactions remain unresolved. These issues could be addressed by employing theoretical approaches in which these factors are examined separately.

The application of accurate quantum mechanical (QM) methods such as ab initio and density functional methods is simply not practical for large systems such as GPx. On the other hand, the use of less accurate methods (various molecular mechanics (MM) and semiempirical methods) could easily lead to wrong conclusions. Therefore, here, the use of hybrid methods, such as ONIOM,[5] is inevitable. However, the application of the ONIOM method on large systems such as GPx is not straightforward and requires special proficiency. The present communication represents the first application of the ONIOM approach on a system as large as the GPx model containing 3113 atoms in 196 amino acid residues. The prime objectives of this study are the elucidation of the structure of GPx and the role of protein/active site interactions for the refinement of the active site.

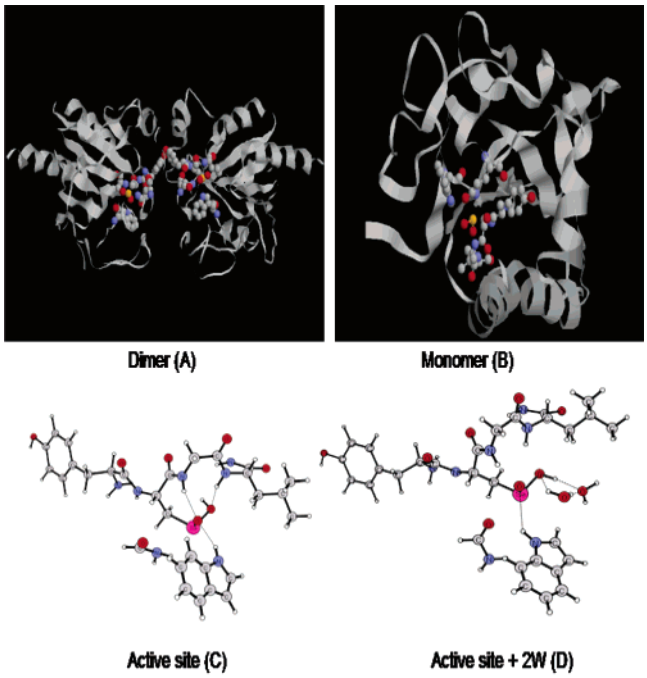

Figure 1. X-ray structures of (A) dimer, (B) monomer, (C) active site, and (D) optimized structure of the active site including two water molecules.

In this study, the two-layer ONIOM2(QM:MM) method has been employed, which treats the "active" part (where all catalytic processes take place) at the QM level and the remainder of the "real" system at the inexpensive MM level. In the present study, the entire monomer (system **I**, Figure 1B) is chosen as a real system, extracted from the dimeric X-ray structure[4] (Figure 1A) and treated at the MM level using the Amber force field.[6] This

* Corresponding author. E-mail: dmusaev@emory.edu. Phone: 1-404-727-2382.

10.1021/jp047303q CCC: $27.50
Published on Web 07/29/2004

TABLE 1: Calculated rms Deviations (in Å) for the Monomer and the Active Site

comparison	system	monomer	active site
X-ray−Amber	**I**	1.72	1.04
X-ray−ONIOM(HF:Amber)_ME	**I**	1.71	1.22
X-ray−ONIOM(B3LYP:Amber)_ME	**I**	1.71	0.97
X-ray−ONIOM(B3LYP:Amber)_EE	**I**	1.73	1.17
X-ray−ONIOM(B3LYP:Amber)_ME	**I(2W)**		0.79
X-ray−B3LYP	**I**, active site only		2.26
X-ray−B3LYP	**I(2W)**, active site only		1.48

choice of monomeric unit is justified by the fact that, in the crystal structure of the mammalian GPx (2.9 Å), the active site selenocysteine residues are well separated with a Se−Se distance of 23.2 Å. Hydrogen atoms not included in the PDB structure were added to the system containing 196 amino acid residues using the GaussView program.[7]

The active part (Figure 1C) of the real system is constructed by utilizing all the available experimental information. Since the selenocysteine residue[8] is experimentally suggested to play a critical role in the catalytic cycle,[3,4] it is included in the active part. Furthermore, in the X-ray structure,[4] Tyr48, Gly50, Leu51, Gln83, and Trp157 residues are shown to form a part of the cage around the selenocysteine residue; therefore, they are also included in the active part. The constructed system **I** consists of 3113 atoms, with 86 atoms in the QM region and the rest in the MM region. As discussed below, later, system **I** was extended by adding two water molecules in the active part (Figure 1D) and was then called system **I**(**2W**).

All the structures that belong to systems **I** and **I**(**2W**) are fully optimized at the Amber, ONIOM(HF/STO-3G:Amber)-_ME, and ONIOM(B3LYP/6-31G(d):Amber)_xx levels, where xx = ME or EE represents the mechanical embedding (ME) and the electronic embedding (EE)[9] schemes. To expound the surrounding protein effect, the geometries of the active parts of systems **I** and **I**(**2W**) without surrounding protein (Figure 1C and D) have been fully optimized at the B3LYP/6-31G(d) level ("active site only" system). All calculations were performed using the Gaussian 03 program package.[10]

All the results are given in Table 1, and comparisons have been made by the means of root-mean-square (rms) deviations between the X-ray and optimized structures using only non-hydrogen atoms. The rms deviations of 1.72, 1.71, and 1.71 Å between the X-ray structures and the Amber, ONIOM(HF/STO-3G:Amber)_ME, and ONIOM(B3LYP/6-31G(d):Amber)_ME structures, respectively, show that, irrespective of the method used, the calculated rms deviations remain the same and significant. The treatment of QM/MM interactions using the electronic embedding (ONIOM(B3LYP/6-31G(d):Amber)_EE) scheme gives an almost similar rms deviation of 1.73 Å.

A detailed analysis of the results shows that the largest deviations between the calculated and X-ray structures correspond to the residues positioned in the vicinity of the second monomer. Therefore, the inclusion of the second monomer into calculations should improve the calculated rms deviation. However, these types of large calculations were not performed in this paper.

The rms deviation of the active site atoms, crucial for the elucidation of the effect of the protein environment on the active site structure and in turn for the catalytic activity, is calculated to be 1.04 and 1.22 Å for the Amber and ONIOM(HF/STO-3G:Amber)_ME levels, respectively. The application of the higher-level ONIOM(B3LYP/6-31G(d):Amber)_ME method reduces it to 0.97 Å, which is still significant. Once again, the treatment of the QM/MM interactions using the electronic embedding (ONIOM(B3LYP/6-31G(d):Amber)_EE) scheme does not improve the result and gives a slightly larger rms deviation of 1.17 Å.

One of the major reasons for such a significant deviation between the calculated and X-ray structures could be the low resolution (2.9 Å) of the X-ray structure. It is well-known that such a low-resolution of X-ray structures neglects water molecules at the active site of the enzyme. It was found that the X-ray structure of the bovine erythrocyte GPx at a significantly higher resolution (2.0 Å) contains two water molecules at the active site. To corroborate the existence of water molecules at the active site of the mammalian GPx (2.9 Å), calculations have been performed including two water molecules at the active site of system **I** from the X-ray structure of the bovine erythrocyte GPx (2.0 Å). However, other possibilities suggesting the presence of either one or more than two water molecules cannot be ruled out. As shown in Table 1 (column 4, row 6), the inclusion of these two water molecules (system **I**(**2W**)) indeed reduces the rms deviation to 0.79 Å. This result suggests the existence of two bound water molecules at the active site of the mammalian GPx.

This conclusion is also supported by the active site only calculations, which give a very large rms deviation of 2.26 Å for the active site of system **I**. The inclusion of two crystal water molecules into the active site only system (the active site of system **I**(**2W**)) reduces this deviation to 1.48 Å, which clearly indicates the importance of these two water molecules at the active site. However, it has to be noted that even the rms deviation of 1.48 Å for the active site only calculation of system **I**(**2W**) is still much larger than 0.79 Å obtained including the protein environment. Thus, this result explicitly demonstrates the significance of the protein/active site interactions for the refinement of the active site structure and consequently for the enzyme activity. The results presented in this study clearly indicate that the computational methods used are less sensitive to the overall structure of the protein but have a much more pronounced effect on the active site structure of the protein. The effect of the protein environment on the catalytic mechanism of GPx is under study and will be discussed in the future.

In summary, the present study represents the first application of the ONIOM scheme to a system as large as mammalian GPx. The results obtained in this study clearly show the importance of the protein environment for the refinement of the active site of the enzyme. Without excluding other possibilities, this study suggests that the active site of the 2.9 Å resolution X-ray structure needs to be complemented by two water molecules, which could be crucial for the catalytic activity of the enzyme.

Acknowledgment. The present research is in part supported by a grant from the National Science Foundation (CHE-0209660). The authors acknowledge the Cherry L. Emerson Center of Emory University for the use of its resources, which is in part supported by a National Science Foundation grant (CHE-0079627) and an IBM Shared University Research Award.

Supporting Information Available: Supporting Information Tables S1 and S2 showing the Cartesian coordinates (in angstroms) of the active site structures of systems **I** and **I**(**2W**), respectively, and Supporting Information Tables S3 and S4 showing the partial RESP charges and Amber bonding parameters for the selenocysteine residue, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Mills, G. C. *J. Biol. Chem.* **1957**, *229*, 189−197.

(2) Flohè, L. In *Glutathione*; Dolphin, D., Avramovic, O., Poulson, R., Eds.; John Wiley & Sons: New York, 1989; pp 644−731.

(3) Epp, O.; Ladenstein, R.; Wendel, A. *Eur. J. Biochem.* **1983**, *133*, 51−69.

(4) Ren, B.; Huang, W.; Åkesson, B.; Ladenstein, R. *J. Mol. Biol.* **1997**, *268*, 869−885.

(5) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170−1179.

(6) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *16*, 1419−1432.

(7) 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.

(8) *GaussianView 3.0*; Gaussian Inc.: Pittsburgh, PA, 2003.

(9) The partial charges and bonding parameters required to describe selenocysteine residues have been obtained from B3LYP/6-31G(d) calculations on the 2-acetylamino-2-methylcarbamoyl-ethaneseleninic acid, using standard Amber RESP, and force constant estimation procedures.

(10) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.

(11) Frisch, M. J.; et al. *Gaussian 03*, revision B7; Gaussian Inc.: Pittsburgh, PA, 2003.