Glutathione Transferase: A First-Principles Study of the Active Site

Gian-Marco Rignanese,*,†,‡ Filippo De Angelis,§ Simone Melchionna,⊥,▽ and Alessandro De Vita‡,||

Contribution from the Unité de Physico-Chimie et de Physique des Matériaux, Université Catholique de Louvain, 1 Place Croix du Sud, B-1348 Louvain-la-Neuve, Belgium, Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), Ecublens, CH-1015 Lausanne, Switzerland, Dipartimento di Chimica e Centro di Studio CNR per il Calcolo Intensivo in Scienze Molecolari, Università di Perugia, I-06123 Perugia, Italy, INFM-Dipartimento di Fisica, Università di Roma "La Sapienza", P.le A. Moro 2, I-00185 Roma, Italy, and INFM and Dipartimento di Ingegneria dei Materiali, Università di Trieste, via A. Valerio 2, I-34127 Trieste, Italy

Received April 4, 2000. Revised Manuscript Received September 11, 2000

Abstract: We present a first-principles study of the interaction of glutathione (GSH) with the enzyme glutathione transferase (GST) and its Y6F mutant. By comparing a reduced model (5−19 atoms) of the interacting species with a larger model (127−131 atoms) including five amino acids of GST, we show that the protein environment effects must be taken into account to properly model the active site. We find that, in the case of the Tyr → Phe mutant, the experimental data on p*K* are reproduced, assuming that a water molecule interacts with the thiol group of GSH. Our results help to elucidate the role that Tyr and water may play as H-bond donors to the thiol group in the enzymatic reaction of GST.

I. Introduction

The preferred pathway for a chemical reaction involving large biomolecules is often determined by subtle effects of charge transfer, polarization, and atomic rearrangement in the active site region. In some cases, to model the reaction it is necessary to describe a fairly extended portion of the system at the fundamental quantum level. Modern density functional-based ab initio techniques are very appropriate tools for this, because of their efficiency and flexibility, whenever the approximations they involve can be assumed to be accurate enough for the problem under study. In the present article we report the study of glutathione (GSH), an ubiquitous tripeptide (γ-Glu-Cys-Gly) found in eukaryotic cells. GSH is implicated in many cellular functions which protect the living cell against toxicity and stress induced by environmental or endogenous chemical agents. These chemicals are deactivated by attaching the SH (thiol) group of glutathione to the hydrophilic moiety of the toxifying agent, thus rendering the electrophile harmless and ready to be removed from the organism. The properties of GSH and, in particular, the stability and structural features of its active thiol site upon variation of the local chemical environment are thus intensively studied, both experimentally[1−4] and theoretically.[5,6] Here, we investigate this problem using various computational schemes and perform first-principles computations on relatively large model systems. The accuracy of the approximations used (in representing the core ions, the electronic wave functions, and the electronic exchange and correlation effects) is checked on smaller systems previously studied in the literature. We also present the results obtained from an embedding scheme combining a quantum description of the active site and a semiempirical Hamiltonian.

GSH (Figure 1) is a substrate of a variety of enzymes such as glutathione transferase (GST), constituting an important class of enzymes that catalyze the process of deactivation of the electrophilic agents. GSTs are grouped starting mainly from the primary sequence into classes Alpha, Kappa, Mu, Pi, Sigma, Theta, Delta, and Zeta, forming homo- and heterodimers. Crystallographic and site-directed mutagenesis studies have shown that of all known mammalian GSTs, a number of amino acid residues of the GST enzyme are preserved and are presumably involved in the binding process.[7] In particular, a tyrosine residue located close to the N-terminal end of the enzyme (Tyr^7 of Pi class human GST or Tyr^9 and Tyr^6 in human Alpha and Mu GSTs, respectively) is believed to play a crucial role in catalysis. Before binding to the chemical agent, the GSH tripeptide is partially embedded in the glutathione transferase, its SH group facing the OH group of the tyrosine residue of GST (see Figure 2). Site-directed mutagenesis experiments have revealed that, by mutating the tyrosine residue into a phenylalanine (Phe^7) amino acid (Phe being almost identical to Tyr^7, except for a hydrogen atom saturating the carbon of the aromatic ring in lieu of the OH group of tyrosine, see Figure 3), the

* Corresponding author. Fax: (+32) 10-473359. E-mail: rignanese@pcpm.ucl.ac.be.

† Université Catholique de Louvain.

‡ Institut Romand de Recherche Numérique en Physique des Matériaux.

§ Università di Perugia.

⊥ Università di Roma "La Sapienza".

|| Università di Trieste.

▽ Present address: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1 EW, UK.

(1) Liu, S.; Zhang, P.; Ji, X.; Johnson, W. W.; Gilliland, G. L.; Armstrong, R. N. *J. Biol. Chem.* **1992**, *267*, 4296−4299.

(2) Ji, X.; Armstrong, R. N.; Gilliland, G. L. *Biochemistry* **1993**, *32*, 12949−12954.

(3) Allardyce, C. S.; McDonagh, P. D.; Lian, L.-Y.; Wolf, C. R.; Roberts, C. K. *Biochem. J.* **1999**, *343*, 525−531.

(4) Gustafsson, A.; Etahadieh, M.; Jemth, P.; Mannervik, B. *Biochemistry* **1999**, *38*, 16268−16275.

(5) Zheng, Y.-J.; Ornstein, R. L. *J. Biomol. Struct. Dyn.* **1996**, *14*, 231−233.

(6) Zheng, Y.-J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1997**, *119*, 1523−1528.

(7) Dirr, H. W.; Reinemer, P.; Huber, R. *Eur. J. Biochem.* **1994**, *220*, 645−661.

10.1021/ja001178o CCC: $19.00 © 2000 American Chemical Society
Published on Web 11/11/2000

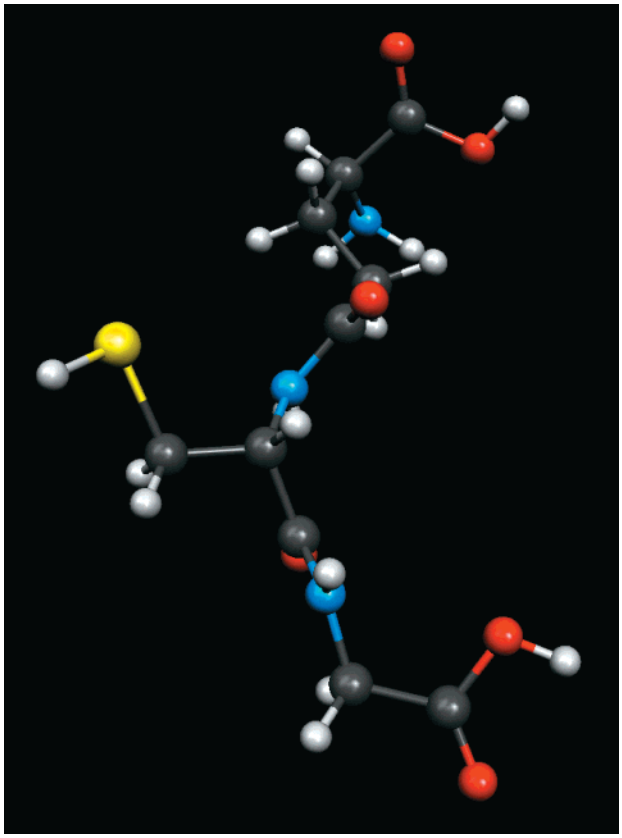

Figure 1. Ball-and-stick representation of glutathione as obtained from the crystallographic structure.[7] The C, N, O, S, and H atoms are represented in gray, blue, red, yellow, and white, respectively.

activity of the enzyme is lowered significantly, although it does not disappear completely.[1] In particular, it has been found that the p*K* is 9.0 for GSH in solution, 6.2 for the native complex (GST(Tyr)), and 7.8 for the Phe[7] mutant (GST(Phe)). Thus, the GSH group has a much larger tendency to deprotonate when bound to the native GST protein.

At first, these data have been interpreted in terms of a hydrogen bond formed between Tyr[7] and GSH (as illustrated by a dashed line in Figure 2), which lowers the p*K* of GSH in the binary complex and is the leading factor in the activation of the sulfur atom toward chemical bonding with external groups.[1] The low p*K* of the GSH−enzyme complex suggests that, at physiological pH, the native glutathione complex shows a deprotonated S atom. Upon mutation of Tyr into Phe, a water molecule may fill the empty space left by removal of the hydroxyl group, and, being close to the thiol, it may effectively alter the p*K* of GSH, thus determining the experimentally observed value.[6]

On the other hand, a different interpretation of the role of Tyr has been proposed on the basis of kinetic studies, according to which the OH group of Tyr is important only for the orientation of the SH group in the active site prior to deprotonation.[8] In the Theta and Delta class GSTs, where a serine residue replaces Tyr, the p*K* of GSH is again efficiently lowered to 6.6. This occurs despite the distance of the thiol from the carboxyl group of Ser being much larger than that from Tyr and despite the more acidic nature of Ser than Tyr.[9] In view of these observations, it has been proposed that the carboxyl group

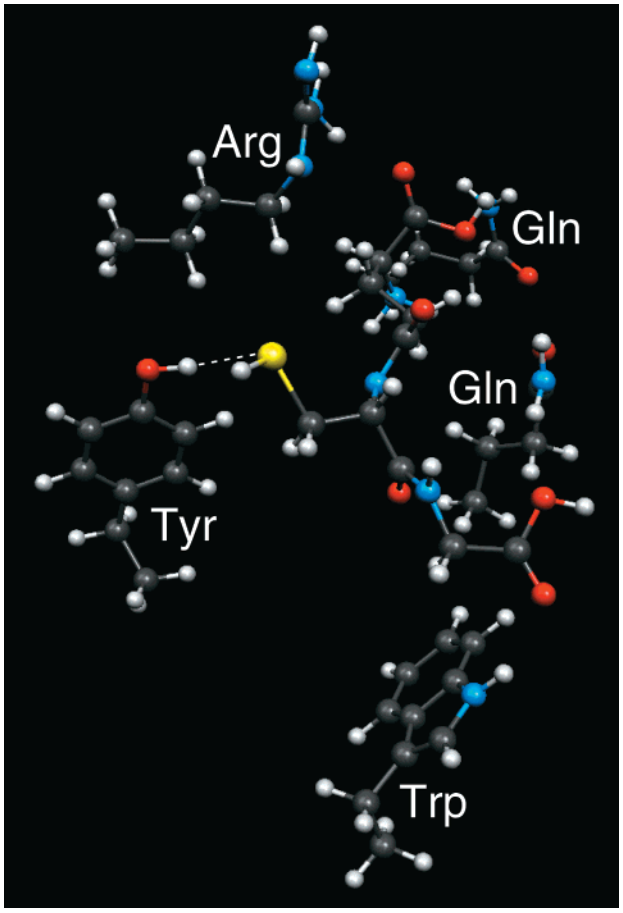

Figure 2. Ball-and-stick representation of our large-scale model for the native complex GST(Tyr) [LM 1]. It includes a complete representation of the GSH molecule (see Figure 1) and five blocking amino acids: Tyr[7], Arg[13], Trp[38], Gln[49], and Gln[62].

of GSH itself is mainly responsible for destabilizing the SH bond. The mechanism of proton extractions is thought to be due to a water molecule present in the active site (Figure 3) and able to shuttle the thiol proton out of the active site.[5] In this model, a water molecule is a H-bond acceptor for the thiol group and a H-bond donor for the carboxylate of Glu of GSH. This implies a structural rearrangement of GSH and GST during the enzymatic reaction, consistent with kinetic studies of Ricci and co-workers.[10,11] The change of p*K* by mutation Tyr → Phe is, in this case, attributed to a structural rearrangement involving GSH and/or GST, of a rather nonlocal character, eventually with the involvement of the C-terminal helix of GST, if present in the primary sequence.[3]

From inspection of the crystallographic data, it is apparent that GSH is partially exposed to water, whereas the thiol belongs to the region of peptide partially screened from the solvent. It is conceivable that a water molecule can be stuck in some sterically determined configuration in the proximity of the active site and may take part in the reaction process. Indeed, the crystal structure of Pi class GST exhibits two important water molecules in the region of the active site. A first water molecule is lost when GSH binds to the active site, and another is released when the GSH conjugate is bound.[12]

(8) Bjornestedt, R.; Stenberg, G.; Widersten, M.; Board, P. G.; Sinning, I.; Jones, T. A.; Mannervik, B. *J. Mol. Biol.* **1995**, *247*, 765−773.

(9) Caccuri, A. M.; Antonini, G., Nicotra, M.; Battistoni, A.; Lo Bello, M.; Board, P. G.; Parker, M. W.; Ricci, G. *J. Biol. Chem.* **1997**, *272*, 29681−29686.

(10) Ricci, G.; Caccuri, A. M.; Lo Bello, M.; Rosato, N.; Mei, G.; Nicotra, M.; Chiessi, E.; Mazzetti, A. P.; Federici, G. *J. Biol. Chem.* **1996**, *271*, 16187−16192.

(11) Caccuri, A. M.; Ascenzi, P.; Antonini, G.; Parker, M. W.; Oakley, A. J.; Chiessi, E.; Nuccetelli, M.; Battistoni, A.; Bellizia, A.; Ricci, G. *J. Biol. Chem.* **1996**, *271*, 16193−16198.

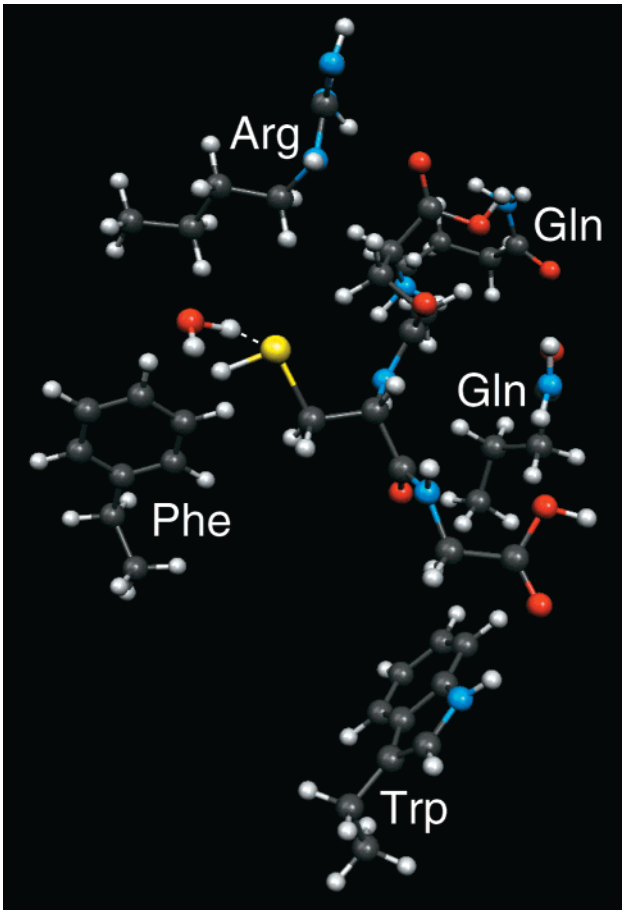

Figure 3. Ball-and-stick representation of our large-scale model for the mutant complex GST(Phe) with a water molecule near the active site [LM 5]. It consists of a complete representation of the GSH molecule (see Figure 1) and includes a water molecule and five blocking amino acids: Phe[7], Arg[13], Trp[38], Gln[49], and Gln[62].

The aim of the present work is to analyze the first proposed model, thus helping to elucidate the role of Tyr in the deprotonation of GSH and in the enzymatic reaction. The inclusion of a water molecule studied here is limited to the H-bond *donor* role that this molecule might play in the Tyr → Phe mutant. The spatial arrangement of the protons in the GSH−GST complex is a first important open question. So far, three different hypotheses have been proposed for the system obtained when one of the two hydrogen atoms of the active site (the thiol proton of GSH or that of the facing OH group of Tyr[7]) is removed. The first suggestion is that Tyr[7] retains its proton on the OH group and acts as a H-bond donor for glutathione, GS···HO−Tyr, thus stabilizing the S^- anion. In this configuration, the Tyr[7] group and a H-bond donor water molecule might have similar effects in lowering the deprotonation barrier. This is consistent with the experimental X-ray structure, although to our knowledge experimental data obtained from neutron diffraction or infrared spectroscopy are not presently available to draw a final conclusion about the proton location. Another possibility is that the Tyr[7] residue may lose its proton and act as a H-bond acceptor, GSH···O−Tyr, a configuration that assists the proton extraction from the thiol group without producing a substantial negative charge surplus on the S atom. Finally, a possible proton arrangement consists of a midway hydrogen bond, GS···H···O−Tyr. The H atom sits, in this case, at the center of a single-well potential with a short S−O distance instead of moving in a double-well H-bond with a large S−O distance.[13,14]

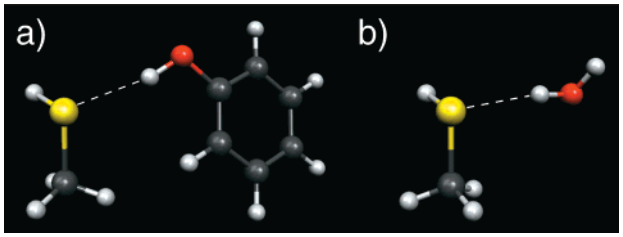

Figure 4. Ball-and-stick representations of our reduced models for (a) the native complex GST(Tyr) [RM 1] and (b) the mutant complex GST(Phe) with a water molecule near the active site [RM 5]. The GSH molecule is represented by a CH_3SH molecule, and the Tyr[7] residue is modeled by a phenol group, $(C_6H_5)OH$, whereas the Phe[7] residue is not taken into account explicitly.

These three models imply different atomic positions and charge arrangements at the enzyme active site, in particular as far as the sulfur atom is concerned. The role played by the S atom in the deprotonation mechanism, and therefore in the enzymatic process, may thus also differ, due to different interactions present in the electrostatic environment of the active site (e.g., in the presence of the electric dipole field generated by the helix α_1 of GST). A previous electronic structure investigation of the complex has considered a reduced representation of the GSH−GST active site at a MP2/MP4 level of accuracy.[6] In this work, GSH has been represented by a CH_3-SH molecule, while the Tyr[7] residue has been modeled more completely by a phenol $(C_6H_5)OH$ group (Figure 4). On the basis of this reduced model, it was concluded that a strong H-bond is formed between the two groups, in the GSH···Tyr configuration. An interesting result of this work is that, when the $(C_6H_5)OH$ molecule is substituted with a benzene ring to model the Phe[7] mutant, the experimental difference in p*K* values between the Tyr[7] and Phe[7] species is *not* recovered. The large discrepancy is speculated to be due to a water molecule stuck between the S atoms and Phe[7] ring, i.e., inside the volume once occupied by the OH group of Tyr[7]. This water molecule, forming a H-bond with the thiol group, lowers the p*K* of glutathione, so that the latter can still exhibit a deprotonated active site at pH 7. A different electronic structure investigation on a model system of Theta class GST has been recently been published.[15] Following a procedure similar to that used by Zheng and Ornstein, a reduced model for the thiolate environment has been used to investigate the deprotonation attitude of GSH. In the case of the Theta class GST, the role of Tyr is apparently played by a Ser hydroxyl. The authors, however, found that the proton stabilization cannot be explained simply by the GSH···Ser direct interaction and that the second shell environment needs to be considered.

The reduced models of the GSH−GST complex used in refs 6 and 15 are unsatisfactory from many points of view. First, due to their limited size, they cannot describe electronic structure rearrangement effects which may occur in the extended GSH−GST system once the proton is extracted. Second, the steric constraints imposed by the structure of the complete GSH molecule may play an important role in the energetics of the proton extraction. The size of the reduced models does not allow us to account for steric constraints when positioning the groups

(12) Parraga, A.; Garcia-Saez, I.; Walsh, S. B.; Tainer, T. J. *Biochem. J.* **1998**, *333*, 811−816.

(13) Kreevoy, M. M.; Liang, T. M. *J. Am. Chem. Soc.* **1980**, *102*, 3315−3322.

(14) Cleland, W. W. *Biochemistry* **1992**, *31*, 317−319.

(15) Flanagan, J. U.; King, W.; Parker, M. W.; Board, P. G.; Chelvanayagam, G. *Proteins* **2000**, *39*, 235−243.

and performing structural relaxations to the energy minimum. Finally, the computed energy differences (and consequently the p*K* values) do not include the effect of the electrostatic interaction between the environment and the active site atoms before and after extraction of the proton.

To investigate the importance of these effects, we perform density functional calculations for different representations of the system. We start by considering the same reduced model as in ref 6. In this case, we compare the results obtained using different GAUSSIAN basis sets and different approximations for the electronic exchange and correlation with those obtained with a plane-wave basis set, to assess the quality of the approximations used. Subsequently, we consider a large-scale model including a complete representation of the GSH molecule and five amino acids of the embedding GST enzyme (class Pi of the pig lung glutathione transferase[16]). With this model, we then study the Tyr^7 and Phe^7 mutant systems. In the case of mutation with Phe^7, for both the reduced and the large models, we analyze the system in the absence of solvent and after inclusion of a single water molecule in the active site region.

II. Technical Details

A. Structural Models. Our reduced model (RM) systems are the same as those described in ref 6. The GSH molecule is represented by a CH_3SH molecule, while the Tyr^7 residue is modeled by a phenol, $(C_6H_5)OH$, group (Figure 4). Our large-scale model (LM) consists of a complete representation of the GSH molecule (Figure 1) and includes five blocking amino acids—Tyr^7, Arg^{13}, Trp^{38}, Gln^{49}, and Gln^{62}—of class Pi of the pig lung GST.[16] These are the amino acids located inside an ellipsoidal region centered on the GSH center of mass, with its axes parallel and proportional in length to the eigenvectors of the molecular inertia tensor of the GSH molecule. The length of the largest axis of the ellipsoidal region is set to 10 Å. We have omitted other amino acids located farther away, because these are expected to have comparatively little effect on charge rearrangements upon proton extraction.[17] We believe that the five amino acids considered provide an extended representation of the GTS active site and electronic reservoir for possible charge rearrangements.

A ball-and-stick representation of this model is given in Figure 2. The amino acids are modeled by their side-chain residues only. The groups embedding the GSH group are free to move, except for the C_α atoms of each residue, which have fixed positions in space in order to inhibit uncontrolled drifting in the simulation cell. By imposing this constraint on the C_α, we allow each residue to slightly rototranslate in order to accommodate the GSH molecule while still taking into account the presence of the enzyme backbone, which is the less flexible part of a protein.[18] No position constraint is applied to the GSH molecule. The GSH initial structure is taken from the X-ray structure of the native complex, where the close-packing arrangement of atoms prevents the water molecules from approaching the active site. When considering the GSH−GST(Phe) system, we start from the same position as in the native Tyr^7 complex (which is the only available crystallographic structure), but we do not impose any blockage on the C_α atom of Phe^7. For both our reduced and large models, the whole system is first treated in vacuo: to study the GSH−GST complex, we do not include water molecules in the vicinity of the active site. When treating the Phe^7 mutant, we also investigate the possible presence of a water molecule placed at the active site, as illustrated in Figure 3. The water molecule is initially positioned in the region between GSH, Phe^7, and Arg^{13}. Its O atom is located 3.5 Å from the S atom of the GSH molecule, with one O−H bond pointing toward the S atom, and the other O−H bond pointing toward an empty region of the cell. The systems investigated are listed in Table 1.

Table 1. List of All the Systems Investigated in This Study[a]

no.	RM	LM
1	$CH_3SH\cdots(C_6H_5)OH$	$GSH\cdots Tyr^7$
2	$CH_3S^-\cdots(C_6H_5)OH$	$GS^-\cdots Tyr^7$
3	CH_3SH	$GSH\cdots Phe^7$
4	CH_3S^-	$GS^-\cdots Phe^7$
5	$CH_3SH\cdots H_2O$	$H_2O\cdots GSH\cdots Phe^7$
6	$CH_3S^-\cdots H_2O$	$H_2O\cdots GS^-\cdots Phe^7$

[a] We consider a reduced model (RM), in which the GSH molecule is represented by a CH_3SH molecule, and large-scale model (LM), which consists of a complete representation of the GSH molecule (see Figure 1) and includes five blocking amino acids—Tyr^7, Arg^{13}, Trp^{38}, Gln^{49}, and Gln^{62}—of class π of the pig lung GST. A ball-and-stick representation of LM 1, LM 5, RM 1, and RM 5 is given in Figures 2, 3, 4a, and 4b, respectively.

B. Theoretical Method. In our plane-wave calculations, the atomic coordinates are optimized to minimize the total energy using the Car−Parrinello method,[19,20] which provides the electronic structure as well as the forces that act on the ions. Only valence electrons are explicitly considered, and norm-conserving nonlocal pseudopotentials are used to account for the core−valence interactions.[21] We adopt periodic boundary conditions, keeping a minimum of 5 Å between repeated images to make any spurious interaction between the images negligible. The electronic wave functions at Γ-point of the supercell Brillouin zone and the electron density are expanded on a plane-wave basis set with kinetic energy cutoffs of 50 and 200 Ry, respectively. The exchange and correlation energies are evaluated in the generalized-gradient approximation of Perdew and Wang.[22]

In our GAUSSIAN calculations the geometry optimizations are performed without symmetry constraints using the default Berny[23] algorithm available in the Gaussian 98[24] package. We use the 6-311+G-(d,p) and the much larger 6-311++G(3df,2pd) basis sets, which are both derived from the original 6-311G basis set.[25] The energy differences computed with the two basis sets are found to differ by less than 2 kcal/mol. We also find that the results obtained with the 6-311++G(3df,2pd) basis set are systematically higher than those computed with the 6-311+G(d,p) basis. This can be interpreted in terms of an overstabilization of the protonated molecule with respect to the anion, presumably due to the large number of polarization functions. The exchange and correlation energies are evaluated using both the BPW91 functional,[22,26] whose correlation part is the same as for the plane-wave calculations, and the hybrid B3LYP functional.[27] The energy differences computed with these two functionals differ in all cases by less than 1 kcal/mol. For brevity, we report here only the results obtained with the 6-311+G(d,p) basis set and the BPW91 exchange-correlation

(16) Reinemer, P.; Dirr, H. W.; Ladenstein, R.; Schaffer, J.; Gallay, O.; Huber, R. *EMBO J.* **1991**, *10*, 1997−2005.

(17) An example of these is Glu^{95}, which forms a salt bridge with Arg^{13}. We note that, neglecting the presence of Glu^{95} in our calculations, we do not observe significant structural rearrangements of Arg^{13} with respect to its experimental structure.[16]

(18) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; Advances in Chemical Physics LXXI; John Wiley and Sons: New York, 1988.

(19) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.

(20) De Vita, A.; Canning, A.; Galli, G.; Gygi, F.; Mauri, F.; Car, R.; *Supercomput. Rev. (EPFL)* **1994**, *6*, 22−27.

(21) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993−2006.

(22) Perdew, J. P. In *Electronic Structure of Solids '91*; Ziesche, P., Eschrig, H., Eds.; Akademie Verlag: Berlin, 1991.

(23) Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996**, *17*, 49−56.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.2; Gaussian, Inc.: Pittsburgh, PA, 1998.

(25) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639−5648.

(26) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

functional. A complete study can be found elsewhere.[28] Frequency calculations have been performed on the RM 1−6 as well as on the phenol molecule in order to include thermal contributions to the energetics of the considered deprotonation reactions.

Finally, Oniom[29] calculations are performed on the LM systems, to compare results obtained by the fully ab initio treatment with those obtained by a combination of quantum mechanics and semiempirical methods. In this case, the RM 3−6 systems (cf. Table 1) are embedded in the environment of the five amino acids included in the LM systems. The geometries are fully optimized by adopting the 6-311+G(d,p) basis set with the BPW91 exchange-correlation functional for the quantum part and by means of the MNDO[30] semiempirical Hamiltonian for the residual part, using default cutoffs[29] for the separation between the quantum and semiempirical regions. The quantum part of the calculations is performed on exactly the same systems studied in the reduced model calculations (quantum systems 3−6). In the semiempirical part of the calculation, a C_α atom of the glutathione carbon chain is matched[29] to a hydrogen atom of the methyl group terminating the quantum system. Direct comparison can thus be made between the RM quantum calculations and the calculations using the Oniom model, and eventual differences can be discussed in terms of steric effects of the surrounding system, described at a semiempirical level, on the quantum system. Comparison with the LM results can also be made, to assess the usefulness of the approximation made in our Oniom calculations.

III. Results and Discussion

A. Deprotonation Energies. For the titration experiment, we consider the following deprotonation reactions:

$$\mathrm{GSH\cdots Tyr} \rightleftharpoons \mathrm{GS^-\cdots Tyr} + \mathrm{H^+} \quad (1)$$

$$\mathrm{GSH\cdots Phe} \rightleftharpoons \mathrm{GS^-\cdots Phe} + \mathrm{H^+} \quad (2)$$

$$\mathrm{GSH\cdots Phe} + \mathrm{H_2O} \rightleftharpoons \mathrm{H_2O\cdots GS^-\cdots Phe} + \mathrm{H^+} \quad (3)$$

$$\mathrm{H_2O\cdots GSH\cdots Phe} \rightleftharpoons \mathrm{H_2O\cdots GS^-\cdots Phe} + \mathrm{H^+} \quad (4)$$

where the "+" and "···" signs refer to the non-interacting and interacting species, respectively. Equations 1 and 2 refer to the ionization of the thiol moiety of the glutathione in the binary complex for the wild-type enzyme and the mutant, respectively. Equations 3 and 4 also refer to the deprotonation process in the presence of the Phe group. However, we here take into account the presence of a crystallographic water molecule which has moved close to the S atom, to form a hydrogen bond with it. In eq 3 we consider the case where this only happens once the thiol is deprotonated, whereas in eq 4, the water interacts with the S atom even before the deprotonation.

The deprotonation energies associated with the reactions 1−4 will be noted ΔE_w, $\Delta E_m{}^1$, $\Delta E_m{}^2$, and $\Delta E_m{}^3$, respectively. Their values can be calculated from the total energies[31] of systems 1−6, reported in Table 2, using the following expressions:

$$\Delta E_w = E_{tot}(2) - E_{tot}(1) \quad (5)$$

$$\Delta E_m{}^1 = E_{tot}(4) - E_{tot}(3) \quad (6)$$

$$\Delta E_m{}^2 = E_{tot}(6) - E_{tot}(3) - E_{tot}(H_2O) \quad (7)$$

$$\Delta E_m{}^3 = E_{tot}(6) - E_{tot}(5) \quad (8)$$

If we introduce the interaction energies,

$$E_{int}(\mathrm{GSH\cdots H_2O}) = E_{tot}(5) - E_{tot}(3) - E_{tot}(H_2O) \quad (9)$$

$$E_{int}(\mathrm{GS^-\cdots H_2O}) = E_{tot}(6) - E_{tot}(4) - E_{tot}(H_2O) \quad (10)$$

we can also write

$$\Delta E_m{}^2 = \Delta E_m{}^1 + E_{int}(\mathrm{GS^-\cdots H_2O}) \quad (11)$$

$$\Delta E_m{}^3 = \Delta E_m{}^2 - E_{int}(\mathrm{GSH\cdots H_2O}) \quad (12)$$

which isolate the contribution to the computed energies of the water molecule bonding to the active site after or before deprotonation.

In Table 3, we report the deprotonation energies for the RM as taken from ref 6 (where the energy differences and geometries were obtained at the MP2 and Hartree−Fock level of theory, respectively) and as computed with a GAUSSIAN and a plane-

Table 2. Total Electronic Energies for the RM and the LM, Computed with a GAUSSIAN Basis Set, and Obtained with a Plane-Wave Basis Set[31] [a]

	RM		LM	
no.	GAUSSIANs	plane-waves	GAUSSIANs	plane-waves
1	−746.253018161	−746.720195		−3175.022817
2	−745.710934731	−746.177035		−3174.506218
3	−438.728964414	−438.909239	−438.726253252	−3099.738688
4	−438.152221315	−438.336879	−438.150928375	−3099.203435
5	−515.183876877	−515.388230	−515.181041336	−3176.216659
6	−514.627323919	−514.834110	−514.625262051	−3175.696184

[a] The energies are expressed in atomic units. The energy of the water molecule is computed to be −76.4497248671 and −76.471039 au with GAUSSIAN and plane-wave basis sets, respectively.

Table 3. Deprotonation Energies[a] for the RM Taken from Ref 6, Computed with a GAUSSIAN Basis Set, and Obtained with a Plane-Wave Basis Set, for the Wild-Type Enzyme (ΔE_w), and for the Mutant in the Absence of Water ($\Delta E_m{}^1$), with a Water Molecule near the Sulfur Atom after Deprotonation ($\Delta E_m{}^2$), and in the Presence of a Water Molecule near the Sulfur Atom All along the Deprotonation Process ($\Delta E_m{}^3$)

	ref 6	GAUSSIANs	plane-waves
ΔE_w	345.3	340.2	340.8
$\Delta E_m{}^1$	368.1	361.9	359.2
$\Delta E_m{}^2$		346.0	342.7
$\Delta E_m{}^3$		349.2	347.7

[a] The energies are expressed in kilocalories per mole.

(28) De Angelis, F. Ph.D. thesis, Università di Perugia, 1999. A PDF file is available at the following URL: http://www.vitillaro.org/3d/filippo/gluta.pdf.

(29) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959−1967.

(30) Dewar, M.; Thiel, M. *J. Am. Chem. Soc.* **1977**, *100*, 4899.

(31) In our plane-wave calculations, pseudopotentials are used to account for the core−valence interactions. Therefore, the results of the calculations are sums of terms including atomic pseudoenergies. These contain an additive constant, which is fixed for every atomic species, and thus cancels out in computing energy differences. To recover the absolute energies which can be compared with those resulting from our GAUSSIAN calculations, we add and subtract the energies and pseudoenergies, respectively. These energies and pseudoenergies of the isolated atoms of each atomic species were obtained from the all-electron calculations performed to produce the pseudopotentials. In the following, the numbers between parentheses are the cutoff radii of the pseudo-wavefunctions, expressed in atomic units. For the C atom, the pseudopotential is generated using a $2s^2(1.47)\ 2p^2(1.47)$ configuration, leading to an energy and a pseudoenergy of −37.848098 and −5.386644 au, respectively. For the N atom, the pseudopotential is generated using a $2s^2(1.40)\ 2p^3(1.40)$ configuration, leading to an energy and a pseudoenergy of −54.545184 and −9.673056 au, respectively. For the O atom, the pseudopotential is generated using a $2s^2(1.45)\ 2p^4(1.45)$ configuration, leading to an energy and a pseudoenergy of −75.093368 and −15.756476 au, respectively. For the S atom, the pseudopotential is generated using a $3s^2(1.75)\ 3p^4(1.75)$ configuration, leading to an energy and a pseudoenergy of −398.260360 and −10.085795 au, respectively. For the H atom, the pseudopotential is generated using a $1s^1(1.20)$ configuration, leading to an energy and a pseudoenergy of −0.466859 and −0.466180 au, respectively.

Table 4. Selected Geometrical Parameters Computed with GAUSSIAN and Plane-Wave Basis Sets for RM 1−6[a]

	neutral system			charged system		
	RM no.	GAUSSIANs	plane-waves	RM no.	GAUSSIANs	plane-waves
d(C−S)	1	1.840	1.870	2	1.846	1.859
	3	1.838	1.871	4	1.848	1.874
	5	1.839	1.832	6	1.847	1.850
d(O−H)	1	0.984	1.001	2	1.077	1.099
	5	0.979	0.997	6	1.016	1.039
d(S···H)	1	2.392	2.311	2	1.918	1.888
	5	2.439	2.424	6	2.169	2.218
∠O−H−S	1	162.2	169.2	2	177.7	177.0
	5	167.2	168.4	6	168.9	171.0

[a] The bond lengths are expressed in angstroms. H refers to the hydrogen atom bonded to the oxygen atom, so that *d*(S···H) refers to the dashed line in Figure 4.

Table 5. Deprotonation Energies for the LM, Computed with a GAUSSIAN Basis Set and Obtained with a Plane-Wave Basis Set: for the Wild-Type Enzyme (ΔE_w), and for the Mutant in the Absence of Water (ΔE_m^1), with a Water Molecule near the Sulfur Atom after Deprotonation (ΔE_m^2), and in the Presence of a Water Molecule near the Sulfur Atom All along the Deprotonation Process (ΔE_m^3)[a]

	GAUSSIANs	plane-waves
ΔE_w		324.2 (−16.6)
ΔE_m^1	361.0 (−0.7)	335.9 (−23.3)
ΔE_m^2	345.6 (−0.4)	322.3 (−20.4)
ΔE_m^3	348.7 (−0.5)	326.6 (−21.1)

[a] The differences with respect to the corresponding RM are indicated in parentheses; a negative value means a lower deprotonation energy for the LM. The energies are expressed in kilocalories per mole.

wave basis set in the present study. The deprotonation energies ΔE_w and ΔE_m^1 from our GAUSSIAN calculations are lower than the MP2/6-31G* values taken from ref 6, with differences of 5.1 and 6.2 kcal/mol, respectively. This is not unreasonable since the present values are obtained with a larger basis set which includes diffuse functions and presumably leads to a higher stabilization of the anions. The deprotonation energies computed using a plane-wave basis set agree very well with those obtained with a GAUSSIAN basis set, with a maximum discrepancy of 3.3 kcal/mol for ΔE_m^2. As a check, we also computed the deprotonation energy ΔE_m^1 in the presence of a C_6H_6 molecule modeling the Phe[7] residue.[32] This value differs from the one obtained for the isolated CH_3SH molecule (RM 3−4) by less than 1 kcal/mol. This confirms that the Phe[7] group does not interact with the sulfur atom of the GSH molecule and hence does not need to be taken into account for the RM calculations.[6]

In Table 4, we report the values of selected geometrical parameters computed with the GAUSSIAN and the plane-wave basis sets. The agreement between these calculations is very good, considering that pseudopotentials are used in the plane-wave calculations. The data indicate that the hydrogen shared between the sulfur atom and the tyrosine OH is closer to the phenolic group, in agreement with the result of previous computational and experimental work.[1,6]

In Table 5, we present the deprotonation energies for the LM computed with a GAUSSIAN basis set (QM-MNDO Oniom calculation) and with a plane-wave basis set (fully ab initio description), as well as the differences between these values and those computed within the reduced model. A negative value for the difference means a lower deprotonation energy in the case of the large system. In the Oniom calculations, the effect of the environment is almost negligible, even if a small stabilization of the anion in the presence of a water molecule is achieved. Remarkably, the fully ab initio calculations give deprotonation energies which are significantly different from those computed for the reduced model. The average difference in energy is −20.3 kcal/mol.

We note that, when we adopt a QM-MNDO approach, the large model accounts for the steric effects of the cavity on the quantum part *only*. Indeed, the model does not allow for charge delocalization from the quantum region into the semiempirical region and vice versa, nor can it account for polarization effects of the protein cavity surrounding the GSH molecule in the enzyme. Whenever these effects become important, a fully ab initio treatment will produce results of higher quality than Oniom calculations. This is what happens in the present case, where the large differences observed between the results of the two models indicate that the environment plays a role which is not limited to imposing steric constraints on the relaxed geometries. We can rationalize this role in terms of the stabilization of the anion after deprotonation, due to rearrangements of the electronic structure throughout the protein cavity and polarization of the cavity itself. Indeed, after deprotonation, we observe significant charge rearrangements up to 7 Å from the S atom in the LM, while this only occurs within 4 Å in the much smaller RM.

Note also that, in the LM (which includes atoms up to 12 Å from the S atom), the charge rearrangements vanish rapidly and become negligible beyond 8 Å. This suggests that our LM is large enough to capture most of the effects of the charge rearrangements due to deprotonation, which is definitely not the case for the RM. The geometrical rearrangement of the active site points to the same kind of effect: the differences from the RM are larger for the fully ab initio than for the QM-MNDO approach. This is clearly seen from the results reported in Table 6, where we compare the geometrical parameters obtained for the LM with those obtained for the RM reported in Table 6. The geometrical parameters which show larger variations upon deprotonation are the *d*(S···H) distance and the ∠O−H−S angle, both in the Oniom and in the ab initio approach. However, a longer *d*(S···H) distance and a smaller value of the ∠O−H−S angle are obtained for the LM in the ab initio calculation, indicating that the S−H hydrogen is weaker when the protein environment effects are fully taken into account by a quantum description.

In Table 7, we compare the optimized geometrical parameters of the LM 1 with the available experimental data.[16] The agreement between the two sets of selected geometrical parameters is very good, suggesting that our LM representation allows for an accurate description of the spatial arrangement and relative orientation of atoms and groups in the enzyme active site.

B. Thermal Effects. To this point, our data on the energetics for hydrogen bond formation have referred to systems at zero temperature. However, the measured p*K* values are the ratio of concentrations of the dissociated species at constant atmospheric pressure and temperature of ∼300 K, so that the Gibbs free energy differences are required to properly estimate the absolute p*K* values. We estimate free energy differences by using entropies and enthalpies ($\Delta G = \Delta H - T\Delta S$) determined from the harmonic vibrational frequencies computed for the RM with the Gaussian 98 program. This is done through a standard statistical mechanics approach. In particular, enthalpies are determined using the following expression:

(32) It should be noted that the presence of the C_6H_6 molecule did not produce sizable steric effects.

Table 6. Selected Geometrical Parameters Computed with GAUSSIAN and Plane-Wave Basis Sets for LM 1−6[a]

	neutral system			charged system		
	LM no.	GAUSSIANs	plane-waves	LM no.	GAUSSIANs	plane-waves
d(C−S)	1		1.858 (−0.012)	2		1.859 (−0.001)
	3	1.854 (0.016)	1.876 (0.005)	4	1.884 (0.036)	1.848 (−0.026)
	5	1.855 (0.016)	1.877 (0.045)	6	1.879 (0.032)	1.854 (0.004)
d(O−H)	1		0.999 (−0.002)	2		1.053 (−0.046)
	5	0.979 (0.000)	0.992 (−0.005)	6	1.004 (−0.012)	1.019 (−0.020)
d(S···H)	1		2.423 (0.112)	2		2.053 (0.165)
	5	2.454 (0.015)	2.611 (0.187)	6	2.272 (0.103)	2.348 (0.130)
∠O−H−S	1		159.5 (−9.7)	2		169.6 (−7.4)
	5	159.5 (−7.7)	144.5 (−23.9)	6	174.0 (5.1)	169.7 (−1.3)

[a] The differences with respect to the corresponding RM are indicated in parentheses; a negative value means a decrease of the considered parameter for the LM. The bond lengths are expressed in angstroms. H refers to the hydrogen atom bonded to the oxygen atom, so that *d*(S···H) refers to the dashed lines in Figures 2 and 3.

Table 7. Comparison between Selected Experimental and Optimized Geometrical Parameters Computed with the Plane-Wave Basis Sets for LM 1[a]

	exptl	LM (plane-waves)
$d(C_{GSH}-O)$	4.391	4.383
d(S−O)	3.368	3.379
$\angle(C_{GSH}-O-C_{Tyr})$	22.3	23.9
$\angle(S-O-C_{Tyr})$	114.6	118.9

[a] The experimental values are taken from ref 16. The bond lengths are expressed in angstroms and bond angles in degrees. C_{GSH} refers to the carbon bound to the S atom in the glutathione molecule, while C_{Tyr} stands for the carbon bound to the oxygen in the Tyr^7 amino acid.

$$\Delta H = \Delta E + \Delta E_v^{\,0} + \Delta(\Delta E_v)^{298} + \Delta E_r^{\,298} + \Delta E_t^{\,298} + \Delta PV \quad (13)$$

Here, the first term represents the already computed deprotonation energy, the second term is the change in zero-point vibrational energy, and the third term is the change in vibrational energy when going from 0 to 298 K. The fourth, fifth, and sixth terms are due to variations of the rotational and translational energies and of the pressure volume work, respectively. All terms except the last can be evaluated with the Gaussian 98 program, using standard statistical thermodynamics.[33] The last term in the present case is $-RT$ (−0.59 kcal/mol at 298 K). We note that, for the plane-wave calculations on both RM and LM, we just add to the energy results the thermal corrections as obtained for the RM with the GAUSSIAN basis set. This may appear to be a crude approximation; however, we find that the computed thermal contributions to the Gibbs free energy differences are almost the same for all the considered RM deprotonation reactions, with average variations of about 1 kcal/mol. Therefore, when we consider differences of energy differences to estimate the p*K* values, the thermal contributions from different deprotonation processes almost cancel each other out.

The deprotonation free energies computed with a GAUSSIAN basis set for the RM and with a plane-wave basis set for both the RM and the LM are reported in Table 8. The variation of the deprotonation free energies from the wild-type enzyme to the mutant is related to the variation of the p*K* values measured experimentally by

$$\Delta\Delta G = \Delta G_w - \Delta G_m = 2.3RT\Delta pK \simeq -2.2 \text{ kcal/mol} \quad (14)$$

In Table 9, we report the computed variation of the free energies: $\Delta\Delta G^1$ in the absence of water, $\Delta\Delta G^2$ with a water molecule near the sulfur atom after deprotonation, and $\Delta\Delta G^3$ in the presence of a water molecule near the sulfur atom all along the process. The computed value of $\Delta\Delta G^1$ is −22.3 kcal/mol for the RM with the GAUSSIAN basis set, in good agreement with the value of −20.5 kcal/mol reported in ref 6. The corresponding result obtained with a plane-wave basis set does not differ much from this value, even if some lowering is observed. A better agreement with experiment is obtained on going to the large model representation of the system. However, the computed value of −12.3 kcal/mol is still about 10 kcal/mol higher than the experimental value in eq 14. If a water molecule is included in the simulated systems, the computed results get significantly closer to the experimental data. For the RM, the computed values for $\Delta\Delta G^2$ and $\Delta\Delta G^3$ are clearly lower than the $\Delta\Delta G^1$ value for the case without water, although still too high for both basis sets to match the experimental value. However, our most complete calculations on the LM yield −6.5 and −3.2 kcal/mol for the $\Delta\Delta G^2$ and $\Delta\Delta G^3$, respectively, in very good agreement with experiment.

Table 8. Deprotonation Gibbs Free Energies[a] for the RM and the LM, for the Wild-Type Enzyme (ΔG_w), and for the Mutant in the Absence of Water (ΔG_m^1), with a Water Molecule near the Sulfur Atom after Deprotonation (ΔG_m^2), and in the Presence of a Water Molecule near the Sulfur Atom All along the Deprotonation Process (ΔG_m^3)

	RM (GAUSSIANs)	RM (plane-waves)	LM (plane-waves)
ΔG_w	333.7	334.4	317.7
ΔG_m^1	356.0	353.3	330.0
ΔG_m^2	347.8	344.6	324.1
ΔG_m^3	343.5	342.0	320.9

[a] The energies are expressed in kilocalories per mole.

Table 9. Variation of the Deprotonation Gibbs Free Energies[a] between the Wild-Type Enzyme and the Mutant for the RM and the LM in the Absence of Water ($\Delta\Delta G^1$), with a Water Molecule near the Sulfur Atom after Deprotonation ($\Delta\Delta G^2$), and in the Presence of a Water Molecule near the Sulfur Atom All along the Deprotonation Process ($\Delta\Delta G^3$)

	RM (GAUSSIANs)	RM (plane-waves)	LM (plane-waves)
$\Delta\Delta G^1$	−22.3	−18.9	−12.3
$\Delta\Delta G^2$	−14.2	−10.3	−6.5
$\Delta\Delta G^3$	−9.8	−7.7	−3.2

[a] The energies are expressed in kilocalories per mole.

In summary, including the effects of the environment at the ab initio level using a large model representation of the enzyme cavity can improve the description of the interaction between the GSH's thiol group and possibly present chemical species interacting with it via hydrogen bonding. We point out that the computational models used here are still restricted since they do not extend to the entire enzyme, that we are neglecting solvation effects, and that we are extrapolating the thermal

(33) McQuarrie, D. A. *Statistical Thermodynamics*; Harper and Row: New York, 1973.

effects from the GAUSSIAN calculation on the reduced models. However, within these approximations, the experimentally measured differences in p*K* between the native and the mutant systems are only reproduced in our calculations by taking into account the presence of a water molecule near the sulfur atom.

IV. Conclusions

In this paper we have studied the interaction of glutathione (GSH) with the enzyme glutathione transferase (GST) by means of ab initio DFT calculations (Gaussian 98 and Car−Parrinello molecular dynamics). We have considered a reduced model of the interacting species, adopting CH_3SH as a model for GSH and C_6H_5OH as a model for the Tyr^7 residue of GST. To account for protein environment effects, we have also investigated a large model representation of the GSH−GST interacting system, including five blocking amino acids Tyr^7, Arg^{13}, Trp^{38}, Gln^{49}, and Gln^{62}, modeled by their side-chain residues. In both cases, we have taken into account the possible presence of a water molecule interacting with the thiol group of GSH.

Our results show that a fairly extended ab initio treatment of the GSH−GST complex is needed for this system, since our large model calculations predict proton extraction energies which are significantly different from those obtained by cruder representations. The present investigation also gives insight into the possible role of a water molecule in the catalytic activity of GSH. Indeed, the small p*K* variation on going from the native to the mutant species is not reproduced in our results on dry systems, although the difference between the two species of the Gibbs free energy variation upon deprotonation is much smaller in this work than in a previous study.[6] The experimental results are reproduced assuming that a water molecule plays a role in the activity of the GSH thiol group, namely by substituting the phenolic group of tyrosine in hydrogen bonding, as suggested in ref 6. This is in agreement with the observed presence of several crystallographic water molecules.[12]

Finally, we note that a model of water-assisted proton extraction is consistent with experiments and theoretical predictions where the water molecule is accommodated between the thiol group and the carboxylate of Glu of glutathione. For GST, the space made available upon the mutation $Tyr^7 \rightarrow$ Phe7 would then be filled by a water molecule, which partially restores the p*K* and allows for a residual catalytic activity of the mutated enzyme.

Acknowledgment. We thank Prof. G. Ricci for important discussions on the subject. Prof. A. Desideri is kindly acknowledged for suggesting the study of glutathione by ab initio methods and for important discussions in the early stages of this project. F.D. acknowledges the CNR and MURST for financial support. S.M. acknowledges the European Human Capital Mobility Program "Molecular dynamics and Monte Carlo simulations of quantum and classical systems" for a research fellowship.

JA001178O