J|A|C|S
ARTICLES
Published on Web 07/26/2002

Quantum Mechanical/Molecular Mechanical Free Energy Simulations of the Glutathione *S*-Transferase (M1-1) Reaction with Phenanthrene 9,10-Oxide

Lars Ridder,*,† Ivonne M. C. M. Rietjens,‡ Jacques Vervoort,§ and Adrian J. Mulholland*,†

Contribution from the School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom, Division of Toxicology, Wageningen University, Tuinlaan 5, 6703 HE Wageningen, The Netherlands, and Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

Received January 18, 2002

Abstract: Glutathione *S*-transferases (GSTs) play an important role in the detoxification of xenobiotics in mammals. They catalyze the conjugation of glutathione to a wide range of electrophilic compounds. Phenanthrene 9,10-oxide is a model substrate for GSTs, representing an important group of epoxide substrates. In the present study, combined quantum mechanical/molecular mechanical (QM/MM) simulations of the conjugation of glutathione to phenanthrene 9,10-oxide, catalyzed by the M1-1 isoenzyme from rat, have been carried out to obtain insight into details of the reaction mechanism and the role of solvent present in the highly solvent accessible active site. Reaction-specific AM1 parameters for sulfur have been developed to obtain an accurate modeling of the reaction, and QM/MM solvent interactions in the model have been calibrated. Free energy profiles for the formation of two diastereomeric products were obtained from molecular dynamics simulations of the enzyme, using umbrella sampling and weighted histogram analysis techniques. The barriers (20 kcal/mol) are in good agreement with the overall experimental rate constant and with the formation of equal amounts of the two diastereomeric products, as experimentally observed. Along the reaction pathway, desolvation of the thiolate sulfur of glutathione is observed, in agreement with solvent isotope experiments, as well as increased solvation of the epoxide oxygen of phenanthrene 9,10-oxide, illustrating an important stabilizing role for active site solvent molecules. Important active site interactions have been identified and analyzed. The catalytic effect of Tyr115 through a direct hydrogen bond with the epoxide oxygen of the substrate, which was proposed on the basis of the crystal structure of the (9*S*,10*S*) product complex, is supported by the simulations. The indirect interaction through a mediating water molecule, observed in the crystal structure of the (9*R*,10*R*) product complex, cannot be confirmed to play a role in the conjugation step. A selection of mutations is modeled. The Asn8Asp mutation, representing one of the differences between the M1-1 and M2-2 isoenzymes, is identified as a possible factor contributing to the difference in the ratio of product formation by these two isoenzymes. The QM/MM reaction pathway simulations provide new and detailed insight into the reaction mechanism of this important class of detoxifying enzymes and illustrate the potential of QM/MM modeling to complement experimental data on enzyme reaction mechanisms.

Introduction

Glutathione *S*-transferases (GSTs) are a family of enzymes that catalyze the conjugation of the tripeptide glutathione (γ-Glu-Cys-Gly) to a wide range of compounds. This reaction is an important step in the detoxification of a large variety of xenobiotic (e.g., carcinogenic) compounds as well as drugs (including anticancer drugs) in mammals and other organisms.[1] The conjugation of glutathione makes these compounds more soluble, which facilitates their excretion via the kidney or the bile as mercapturic acids. Understanding of GST activity, therefore, provides important insight into human health implications of toxins and drugs. An important group of substrates for class μ and π GST enzymes consists of epoxides of polycyclic aromatic hydrocarbons, primarily formed in the body by cytochrome P450-catalyzed biotransformation reactions.[2,3] The conjugation of glutathione to these substrates proceeds via a nucleophilic attack of the thiolate sulfur on either of the epoxide carbons, which results in cleavage of the epoxide ring. Phenanthrene 9,10-oxide has become a model substrate for studying the effect of active site polymorphism on the regio- and

* Authors to whom correspondence should be addressed. L.R.: (tel) +44 117 9546863; (e-mail) l.ridder@bristol.ac.uk. A.J.M.: (tel) +44 117 928 9097; (e-mail) adrian.mulholland@bristol.ac.uk.
† University of Bristol.
‡ Division of Toxicology, Wageningen University.
§ Laboratory of Biochemistry, Wageningen University.

(1) Salinas, A. E.; Wong, M. G. *Curr. Med. Chem.* **1999**, *6*, 279−309.
(2) Cobb, D.; Boehlert, C.; Lewis, D.; Armstrong, R. N. *Biochemistry* **1983**, *22*, 805−812.
(3) Armstrong, R. N. *Chem. Res. Toxicol.* **1997**, *10*, 2−18.

10.1021/ja0256360 CCC: $22.00

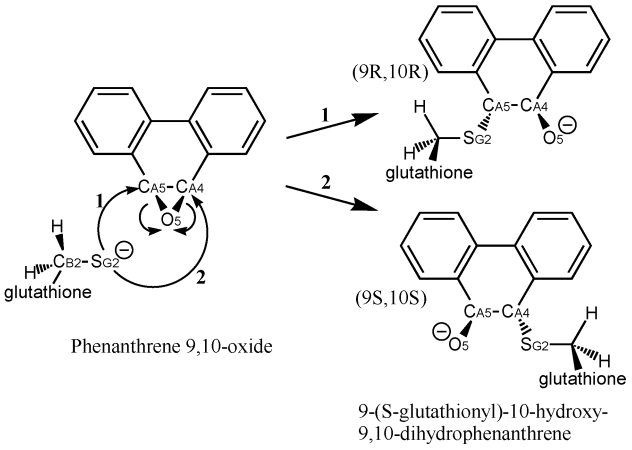

Figure 1. Conjugation of glutathione to phenanthrene 9,10-oxide by glutathione *S*-transferase, leading to two diastereomeric products. Note that although the products are formed by the attack on two distinguishable substrate carbon atoms, the products are diastereomers because of the symmetry of the phenanthrene 9,10-oxide substrate. Atom labels according to PDB entry 2GST. This labeling is different from the numbering used in the standard chemical nomenclature: the 9- and 10-positions in phenanthrene 9,10-oxide refer to the CA4 and CA5 atoms in our model. Note that, in (9*R*,10*R*)-9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene, the 9-position is represented by CA5 and the 10-position by CA4, whereas in the (9*S*,10*S*) diastereomer, this is the other way around.

stereospecificity of the addition of GSH to epoxides (Figure 1). For example, the M1-1 and M2-2 isoenzymes from rat show a significantly different stereospecificity toward this substrate. (M1-1 and M2-2 from rat, named according to the nomenclature for human cytosolic GSTs,[4] were formerly designated isoenzymes 3-3 and 4-4, respectively.) M1-1 gives approximately equal amounts of both diastereomeric products, whereas M2-2 produces essentially only the (9*S*,10*S*) diastereomer of 9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene.[2,5] Crystal structures of the M1-1 isoenzyme from rat in complex with each of the two diastereomeric products have been reported.[6] The reasons for the difference in specificity between the isoenzymes, however, are still not fully understood.

A number of other aspects of the GST-catalyzed conjugation of glutathione to epoxides remain to be clarified. An interesting feature of the reaction is the desolvation of the thiolate moiety of glutathione along the reaction path, which was indicated by solvent isotope effects.[7,8] Zheng and Ornstein[9] performed model calculations on the conjugation reaction of glutathione to 1-chloro-2,4-dinitrobenzene, which indeed suggest a strong solvation effect (modeled by a dielectric continuum) on the energetics of the reaction. However, these calculations do not provide a highly defined picture of the desolvation process, as the solvent was not modeled explicitly, and also did not include the active site environment provided by the enzyme. Another aspect that requires further investigation arises from a marked difference between the crystal structures with the two diastereomeric products of the reaction with phenanthrene 9,10-oxide.[6] In the crystal structure with the (9*S*,10*S*) diastereomer bound to the enzyme, a direct hydrogen bond between the 10-hydroxyl moiety of the product and Tyr115 is observed. The presence of this hydrogen bond supports the proposal that Tyr115 plays an important catalytic role by providing "electrophilic assistance" to the ring opening reaction. However, in the crystal structure with the (9*R*,10*R*) diastereomer, a water molecule mediates the hydrogen bonding to Tyr115.[6] This raises the question of what the exact role of Tyr115 in the conjugation reaction is, and to what extent the crystal structures represent the reactive active site.

In the present study, a combined quantum mechanical/molecular mechanical (QM/MM) method[10] has been applied to study the reaction of GST with phenanthrene 9,10-oxide and to address the issues mentioned above. In the QM/MM approach, a small part of a system, such as the reacting groups in the active site of an enzyme, is treated at a quantum chemical level, whereas the large surrounding region is modeled using molecular (classical) mechanics. This approach allows simulations of reactions in large systems such as proteins. QM/MM simulations of enzyme-catalyzed reactions can provide useful insight into enzyme mechanisms at a level of detail often beyond that attainable from experiment. It is therefore becoming an increasingly important tool to complement experimental biochemistry.[11−16]

The combination of certain aspects of the mechanism of GSTs makes accurate simulation of the reaction a challenge. The reaction involves a nucleophilic attack by a thiolate moiety, which is a demanding reaction from a quantum chemical point of view. Sulfur, as a third-row element, shows complex chemistry and its accurate description imposes some requirements on ab initio theoretical treatments, such as inclusion of electron correlation and the use of reasonably large basis sets. This is especially true for the present reaction in which the formal charge on the sulfur changes from negative to neutral. The calculations performed by Zheng and Ornstein on the conjugation reaction of glutathione to 1-chloro-2,4-dinitrobenzene illustrate that the outcomes are very sensitive to the level of quantum chemical theory applied.[9] In the present study, a semiempirical (AM1) treatment of the QM region is applied, in which the parameters for sulfur are calibrated to reproduce high-level ab initio results. The use of semiempirical molecular orbital theory as an adjustable potential (rather than a predictive method) was proposed by Truhlar and co-workers[17,18] and has recently been shown to be very successful in simulations of condensed-phase reactions.[19,20] It combines accuracy with

(4) Mannervik, B.; Awasthi, Y. C.; Board, P. G.; Hayes, J. D.; Diilio, C.; Ketterer, B.; Listowsky, I.; Morgenstern, R.; Muramatsu, M.; Pearson, W. R.; Pickett, C. B.; Sato, K.; Widersten, M.; Wolf, C. R. *Biochem. J.* **1992**, *282*, 305−306.
(5) Zhang, P. H.; Liu, S. X.; Shan, S. O.; Ji, X. H.; Gilliland, G. L.; Armstrong, R. N. *Biochemistry* **1992**, *31*, 10185−10193.
(6) Ji, X. H.; Johnson, W. W.; Sesay, M. A.; Dickert, L.; Prasad, S. M.; Ammon, H. L.; Armstrong, R. N.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 1043−1052.
(7) Huskey, S. E. W.; Huskey, W. P.; Lu, A. Y. H. *J. Am. Chem. Soc.* **1991**, *113*, 2283−2290.
(8) Chen, W. J.; Graminski, G. F.; Armstrong, R. N. *Biochemistry* **1988**, *27*, 647−654.
(9) Zheng, Y. J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1997**, *119*, 648−655.
(10) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(11) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826−5832.
(12) Hartsough, D. S.; Merz, K. M. *J. Phys. Chem.* **1995**, *99*, 11266−11275.
(13) Harrison, M. J.; Burton, N. A.; Hillier, I. H. *J. Am. Chem. Soc.* **1997**, *119*, 12285−12291.
(14) Ridder, L.; Mulholland, A. J.; Vervoort, J.; Rietjens, I. M. C. M. *J. Am. Chem. Soc.* **1998**, *120*, 7641−7642.
(15) Mulholland, A. J.; Lyne, P. D.; Karplus, M. *J. Am. Chem. Soc.* **2000**, *122*, 534−535.
(16) Ridder, L.; Mulholland, A. J.; Rietjens, I. M. C. M.; Vervoort, J. *J. Am. Chem. Soc.* **2000**, *122*, 8728−8738.
(17) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618−4627.
(18) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *233*, 231−236.
(19) Bash, P. A.; Ho, L. L.; MacKerell, A. D.; Levine, D.; Hallstrom, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3698−3703.

computational efficiency. The latter is important when dealing with the other challenges of simulating the reactions catalyzed by GSTs, which arise from the very large solvent-exposed active site. Its aspecificity, illustrated by the ability to bind a large variety of substrates, implies that the actual enzyme−substrate complex is flexible and may comprise an ensemble of different conformations of protein and ligand as well as the presence of many possible water configurations within the active site. Furthermore, the energetics of the reaction are expected to be strongly influenced by solvation effects, as mentioned above. It is not sufficient to represent such a system by a single conformation, as the energy profiles would be highly dependent on the protein, reactant, and solvent configuration. This means that a simple adiabatic mapping approach, which has previously provided useful insight into some enzymes[14,16,21,22] (e.g., those with buried, tightly ordered active sites) is not applicable here. In the present study, therefore, the reaction was studied by free energy calculations based on QM/MM molecular dynamics umbrella sampling, which accounts for the flexibility of the reactants and the enzyme active site as well as for the explicit solvation in the active site of GST. The results are in good agreement with experiment and provide new insight into the detailed catalytic mechanism of this important enzyme reaction.

Methods

Simulation System. The starting models for the present simulations were based on the two crystal structures of the M1-1 isoenzyme of rat liver glutathione *S*-transferase in complex with (two diastereomers of) 9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene[6] (PDB codes: 2GST and 3GST), the two possible products of the reaction. The position of the glutathionyl part of the ligand in these structures is identical to the glutathione ligand (GSH) in the crystal structure of the enzyme−glutathione complex (PDB code: 6GST), indicating that these structures are good starting points for modeling the reaction.

It is well established that the thiol group of GSH deprotonates upon binding in the active site of GST.[23,24] Thus, the reactions modeled in the present study are the nucleophilic attack of the anionic thiolate moiety of GS^- on either carbon of the epoxide ring. In the initial product of this reaction, the resulting hydroxylate oxygen is anionic (it is assumed that ring opening is not strictly concerted with protonation of the oxygen). On the other hand, the crystal structures used to build the models are likely to represent the complexes with the neutral (protonated) products of the reaction. To test these assumptions, each crystal structure (each containing a different diastereomer of the product) was used to build two models (four models in total), one of the neutral product (protonated at O5) and one of the anionic product (deprotonated at O5). Construction of and simulations with the four models were carried out with the CHARMM[25] program as follows.

Hydrogen atoms were built on the basis of CHARMM internal coordinates.[26] Hydrogen atoms on crystal water molecules were placed using the HBUILD routine[27] in CHARMM. The simulations were focused on the active site of the α subunit, using a stochastic boundary approach:[28] the simulation system contained a spherical selection of residues centered around the reacting atoms in the α subunit (i.e., the mass-weighted center of all 9-(*S*-mercaptomethyl)-10-hydroxy-9,10-dihydrophenanthrene atoms). All protein residues and crystal waters with at least one atom within 18 Å of the center were selected. This system was then solvated by superimposing an 18-Å-radius sphere of water and deleting any added water molecule of which the oxygen was within 2.6-Å distance of another non-hydrogen atom. After 10 ps of Langevin dynamics (300 K) for all water molecules (all other atoms fixed), the solvation procedure was repeated. Finally, 100 steps of steepest descent minimization (of water molecules only) were performed. After solvation, the models with the (9*S*,10*S*) product consisted of 118 residues of subunit A, 21 residues of subunit B, the (9*S*,10*S*)-9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene ligand (bound to the active site of subunit A), and 250 water molecules (including 103 selected crystal waters). The models with the (9*R*,10*R*) product consisted of 122 residues of subunit A, 26 residues of subunit B, the (9*R*,10*R*)-9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene ligand (bound to the active site of subunit A), and 242 water molecules (including 97 crystal waters). Further calculations were performed using a combined QM/MM potential. The QM region, i.e., the part of the ligand including the cysteine side chain and the conjugated phenanthrene, as well as the side chain of Tyr6, was treated with a modified semiempirical AM1 method (see below). Tyr6 was included in the QM region because it has been proposed to play an essential role in stabilizing the thiolate sulfur in the reactant state.[23] To describe the two bonds crossing the QM/MM boundary, link atoms[10] were placed between the Cα and Cβ of the cysteinyl side chain of glutathionyl and between the Cβ and Cγ of the Tyr6 side chain. The link atoms are QM hydrogens without classical van der Waals or bonded force field terms. They do, however, interact with all MM charges, where the charges on the link atom host groups have been set to zero (this is the HQ-type link atom,[29] which has been found to perform better than the original QQ-type link atom[10]). Although link atoms are an approximate approach to QM/MM partitioning of covalently bonded systems, they have been found to give good results for a number of simulations of enzyme reactions.[14,16,21,29,30] The overall charge on the QM system is 0 or −1, depending on the protonation state of the product. Other atoms were treated molecular mechanically, using the CHARMM22 all-hydrogen force field.[26] Throughout, a 4-Å buffer zone was defined as all atoms further than 14 Å away from the center of the sphere, in which the nonsolvent heavy atoms were harmonically restrained to their crystal coordinates with force constants based on model average *B* factors.[21,28] These restraints were scaled (linearly and in four steps) from zero at 14 Å from the center of the system, to a maximum at 18 Å. During dynamics simulations, Langevin dynamics were applied for the buffer region (updated every 50 steps), using friction coefficients of 250 ps^{-1} for non-hydrogen protein atoms and 62 ps^{-1} on water oxygen atoms.[28] A deformable boundary potential[31] was applied on the water oxygens. SHAKE[32] was applied to fix all MM bonds involving hydrogen atoms, and a 1-fs time step was used. Nonbonded interactions were calculated within a 13-Å cutoff distance. MM/MM electrostatic interactions were scaled down using an atom-based SHIFting function, whereas QM/MM electrostatics were scaled down by a group-based SWITching function between 8 and 13 Å (the QM atoms are divided into two groups: those of Tyr6 are in one group, and the 9-(*S*-mercaptomethyl)-10-hydroxy-9,10-dihydrophenanthrene atoms are in

(20) Lau, E. Y.; Kahn, K.; Bash, P. A.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9937−9942.
(21) Mulholland, A. J.; Richards, W. G. *Proteins* **1997**, *27*, 9−25.
(22) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.
(23) Liu, S. X.; Zhang, P. H.; Ji, X. H.; Johnson, W. W.; Gilliland, G. L.; Armstrong, R. N. *J. Biol. Chem.* **1992**, *267*, 4296−4299.
(24) Graminski, G. F.; Kubo, Y.; Armstrong, R. N. *Biochemistry* **1989**, *28*, 3562−3568.
(25) CHARMM. Available from the CHARMM Development Project, Harvard University, Cambridge, MA.
(26) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(27) Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148−156.
(28) Brooks, C. L.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(29) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(30) Cunningham, M. A.; Ho, L. L.; Nguyen, D. T.; Gillilan, R. E.; Bash, P. A. *Biochemistry* **1997**, *36*, 4800−4816.
(31) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(32) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

the other). The nonbonded list was updated either heuristically (for minimizations) or every 25 steps (for dynamics simulations). Simulations were carried out with CHARMM version 27b2.[25]

Simulations of the Product Complexes. To test the QM/MM model with respect to its ability to maintain the correct structure of the QM region in the active site, as well as to investigate the influence of the protonation state of the product on the QM structure, all four models of the product complexes were used for QM/MM molecular dynamics simulations. After 200 steps of steepest descent (SD) minimization, the product complexes were equilibrated for 15 ps. Then, 20-ps simulations were performed, the average QM structures of which were compared to the crystal structures coordinates.

Free Energy Simulations of the Conjugation Reaction. As the reaction is assumed to initially result in the anionic product, the two models containing the deprotonated products were used for free energy simulations. The initial geometries of both QM/MM models were first optimized toward the reactant state as follows. In each case, harmonic restraints were applied to the two C−O distances corresponding to the epoxide bonds in the reactant state. A total of 500 steps of steepest descent minimization were used to convert the geometry of the QM atoms into the epoxide/GS$^-$ conformation (i.e., the reactant state). Subsequently, 200 steps of SD minimization were performed without the harmonic restraints on the epoxide ring, yielding a stable epoxide/GS$^-$ structure. This minimized structure of the reactant state was equilibrated by 15 ps of stochastic boundary molecular dynamics. At this point, for reasons described in the Results and Discussion section, the equilibrated structure obtained for the model based on the X-ray structure with the (9*S*,10*S*) ligand (2GST) was selected as the starting point for subsequent free energy calculations based on umbrella sampling.[33] Starting from this equilibrated structure of the reactants, two series of simulations were performed, in which two different reaction coordinates, leading to the two diastereomeric products, were applied. These reaction coordinates were defined as follows: $r_1 = d(\mathrm{CA5} - \mathrm{O5}) - d(\mathrm{SG2} - \mathrm{CA5})$, which results in the (9*R*,10*R*) diastereomer and $r_2 = d(\mathrm{CA4} - \mathrm{O5}) - d(\mathrm{SG2} - \mathrm{CA4})$, which results in the (9*S*,10*S*) diastereomer (labeling as in Figure 1). In each series, one of the two reaction coordinates was harmonically restrained to a range of values (at every 0.1 Å) along the reaction coordinate. A force constant of 200 kcal mol^{-1} Å^{-2} was used, which resulted in simulations sampling sufficiently overlapping regions of the reaction coordinate. Each simulation consisted of 10 ps of equilibration followed by 20 ps of sampling dynamics. Each subsequent equilibration is started from the 1-ps point of the adjacent equilibration run. The reaction coordinate statistics of the various simulations were combined by means of a weighted histogram analysis method.[34,35]

Reaction-Specific Parametrization of Sulfur. The QM region was treated with the semiempirical AM1 method.[36] Although AM1 is a reasonably good model for many organic molecules, sometimes better than Hartree−Fock ab initio treatments, the reliability and accuracy of AM1 may vary significantly for different systems. A recent development for good-quality condensed-phase simulations is to reparametrize MNDO-type methods, such as AM1, to reproduce experimental or high-level ab initio data for the specific chemical reaction of interest.[17−19] For the reaction of GST, sulfur is the most important element to focus on, where the accuracy of AM1 is concerned. Sulfur is a difficult element in the context of semiempirical MO theory, partly because of the many different chemical configurations in which it occurs (including three different valence states).[37] In the relatively large QM system of our model system, the other elements, H, C, and O, are present in various chemical environments; i.e., C and H are present in an aliphatic as well as an aromatic configuration, and O is present in both the tyrosine side chain and the epoxide ring of the substrate. It was therefore considered appropriate to use the standard parameters (with a more general validity) for these elements, whereas significant improvement was expected from a reaction-specific optimization of the parameters for sulfur (which occurs only once in the present QM system). This approach is valid since in standard AM1, the parameters for C, H, N, and O[36] were developed prior to, and independently from those for sulfur,[37] and it has the advantage of retaining standard (general) AM1 treatment for most of the QM system under investigation.

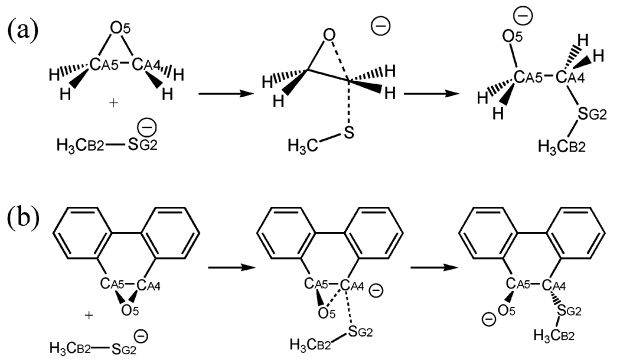

Figure 2. Small (a) and large (b) gas-phase model systems used for testing and reparametrization of AM1. Labels are chosen to correspond to atom labels in the crystal structure PDB entry 2GST.

Two gas-phase models for the reacting system were examined for validation (and optimization) of AM1. A small model was defined as the reaction between methylthiolate and epoxyethane (Figure 2a). A larger model was defined as the reaction between methylthiolate and phenanthrene 9,10-oxide, which corresponds to the QM region of the QM/MM model without the side chain of Tyr6 (Figure 2b). The reactants, transition states, and products of both model reactions were optimized in the gas phase at the AM1 as well as HF/6-31G* and B3LYP/6-31+G** levels, using MOPAC-93[61] and Gaussian-98,[38] respectively. MP2/6-31+G** optimizations were performed for the small model only. The small model was subsequently used to optimize the AM1 parameters for sulfur to produce results closer to those obtained at the higher levels.

The parameters of MNDO-type methods can be optimized efficiently using genetic algorithms.[18,19] The algorithm used in the present study is similar to that described by Bash et al.[19] and implemented as a modification to the libGA code by Corcoran and Wainwright.[39] A pool of 300 chromosomes, each with a length of 10 variables (AM1 parameters for sulfur) was generated from a random Gaussian distribution with an SD of 3% around the standard AM1 values.[37] The chromosomes are represented as arrays of real numbers, as opposed to a binary representation used in the original genetic algorithms.[40] Uniform crossover and mutation probabilities of 0.7 and 0.5 were applied, where mutated values were taken from a random Gaussian distribution with an SD of 1% around the current value. A rank-biased steady-state population replacement was used.

(33) Proust-De Martin, F.; Dumas, R.; Field, M. J. *J. Am. Chem. Soc.* **2000**, *122*, 7688−7697.
(34) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(35) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450−1462.
(36) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(37) Dewar, M. J. S.; Yuan, Y. C. *Inorg. Chem.* **1990**, *29*, 3881−3890.
(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A. J.; 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.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*, Revision A.6; Gaussian, Inc., Pittsburgh, PA, 1998.
(39) Corcoran, A. L.; Wainwright, R. L. In *Proceedings of the 1993 ACM/SIGAPP Symposium on Applied Computing*; Indianapolis: IN, 1993. (ftp://www.aic.nrl.navy.mil/pub/galist/src/libga100.tar. Z).
(40) Goldberg, D. E. *Genetic algorithms in search, optimization and machine learning*; Addison-Wesley: Reading, MA, 1989.

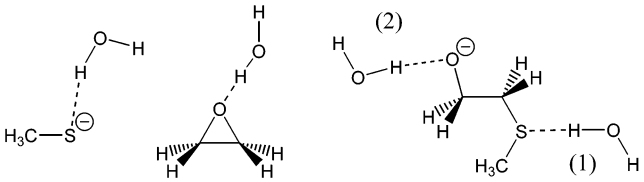

Figure 3. Solute−solvent complexes used for calibration of the thiolate sulfur and epoxide oxygen Lennard-Jones parameters.

The "fitness" of a set of sulfur parameters (a chromosome) was defined as the sum of (weighted) errors in the resulting AM1 energies (enthalpies of formation) for the methylthiolate (×2), the transition state for its conjugation to epoxyethane (×2), and the resulting product (×1) of the small gas-phase model reaction optimized in MOPAC-93. The numbers in parentheses are the weighting factors used in the fitness function, which were applied to ensure that a good result for the barrier was obtained with priority to a good result for the reaction energy. The target enthalpy for methylthiolate was the experimental heat of formation, i.e. −14.3 kcal/mol.[41] The target enthalpies for the transition state and the product were derived by adding the relative enthalpies with respect to the reactants calculated at the higher levels to the AM1 enthalpy for epoxyethane (−8.96 kcal/mol) plus the experimental value for methylthiolate. (The use of the AM1 ΔH_f for epoxyethane is appropriate as only *S* parameters were to be optimized.) Thus, the target enthalpies (of formation) used for the transition state and product were −20 and −40 kcal/mol, respectively.

It is important to note that the optimization of the sulfur parameters was based on the smaller model system only (Figure 2a) and that the larger model (Figure 2b) was used as an independent test of how well the modified (reaction-specific) version of AM1 (denoted as AM1-SRP) performs for the reaction of interest.

Calibration of QM/MM Interactions. In the QM/MM model used,[10] QM/MM interactions involve a classical (MM) Lennard-Jones term. Often, Lennard-Jones parameters optimized for MM models[26] are used for the QM atoms. These parameters, although they often provide reasonable values, are not necessarily optimal for QM/MM models. Better results can be found by optimizing the Lennard-Jones parameters of the QM atoms to reproduce interaction energies and geometries of small model complexes.[10,19] As with the development of nonbonded parameters for purely MM calculations with the CHARMM22 parameter set,[26] typically complexes with a single water molecule are chosen. This type of approach was followed here and is consistent with the CHARMM22 MM parameter set used for the system.

A key aspect of the present simulations is to investigate the effect of the solvent in the active site on the energetics of the reaction. The main effects are expected from changes in solvation of the thiolate sulfur and the epoxide oxygen along the reaction coordinate. Therefore, it was important to test and calibrate the QM/MM interactions for these two atoms. Figure 3 presents the model complexes used for the calibration. Geometries and interaction energies were calculated using HF/6-31G*, which is the standard level of theory used in the development of MM parameters,[26] B3LYP/6-31+G** and MP2/6-31+G**. These methods have been found to give good results for hydrogen-bonded complexes.[19,26,42,43] The MM parameters for the QM oxygen and sulfur atoms were manually fitted to reproduce interaction energies and geometries of the complexes calculated at the higher levels of theory.

Approximate Modeling of Mutations. An important goal in connecting simulations of enzyme-catalyzed reactions to experimental investigations is to reproduce and interpret the results of site-directed mutagenesis experiments. In principle, the effects of mutations can be studied by free energy perturbation (FEP) calculations using the QM/MM method, with the mutation modeled in the MM system. For example, the effects on the reactants and a model of the transition state could be studied, as in FEP calculations using MM models of enzymic transition states and substrates.[44,45] Such calculations are, however, computationally demanding at the QM/MM level. We have therefore applied a more approximate, but computationally efficient procedure to examine the effects of mutations on the reaction.

Four mutations were selected, as discussed in the Results and Discussion section, which do not involve an increase in the number of atoms, and for which the structural changes are expected to be small. The effects of these mutations were modeled as the average change in QM/MM electrostatic interaction energy upon reassigning charges to the MM atoms of the mutated side chains, with the averaging carried out on the wild-type structure. In the case of the Tyr115Phe mutation, for example, a zero charge was assigned to the hydroxyl hydrogen of Tyr115 and the charges of the other atoms were changed to the standard values of those of a Phe residue in the CHARMM22 force field[26] (the hydroxyl oxygen is given the charge of a phenyl hydrogen). For five simulations, corresponding to the reactants and the products and transition states of both diastereomeric reactions, the average energetic effect of the "mutation" on the QM energy was determined for 100 structures, taken from the trajectory every 0.2 ps. This procedure does not include steric and structural changes, but these are expected to be relatively small in the cases under consideration. The approach is similar in spirit to simple interaction energy analyses, which have proved useful in identifying important residues and their function in earlier QM/MM studies of enzyme reactions.[11,16,21,30,46]

Table 1. Enthalpies[a] (kcal/mol) of the Transition States (ΔH_{TS}) and Products (ΔH_{Prod}) of Two (Small and Large) Model Reactions in Gas Phase (Shown in Figure 2), Relative to the (Separate) Reactants, Calculated at Different Levels of Theory

enthalpies,[a] kcal/mol	ΔH_{TS} (small)	ΔH_{prod} (small)	ΔH_{TS} (large)	ΔH_{prod} (large)
HF/6-31G*	19.65	−0.18	19.18	−3.16
B3LYP/6-31+G**[b]	3.23	−13.74	0.22	−16.45
MP2/6-31+G**	3.71	−18.91	7.60[c]	−21.22[c]
AM1	2.96	−26.75	4.20	−30.61
AM1-SRP	3.88	−18.14	5.54	−21.99

[a] Energies obtained for HF, B3LYP and MP2 calculations were converted to enthalpies by applying the unscaled thermal corrections derived at the HF/6-31G* level. [b] The B3LYP method is known to sometimes underestimate barriers.[47] [c] These values were obtained by a single-point LMP2/6-31+G** calculation on the B3LYP/6-31+G** geometry, using Jaguar 4.0.[48]

Results and Discussion

Calibration of the Model. Table 1 shows the results of the testing and optimization of AM1 for the two gas-phase model systems (Figure 2). The results from the HF/6-31G* calculations differ significantly from the results obtained at the B3LYP/6-31+G** and MP2/6-31+G** levels of theory. This indicates that, for reliable results, electron correlation must be included in the ab initio calculations on the present model systems. For the small model system, B3LYP/6-31+G** and MP2/6-31+G** give similar results for the energy difference between

(41) Bartmess, J. E. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 2001. (http://webbook.nist.gov).
(42) Pan, Y. P.; McAllister, M. A. *Theochem−J. Mol. Struct.* **1998**, *427*, 221−227.
(43) Gilli, G.; Gilli, P. *J. Mol. Struct.* **2000**, *552*, 1−15.
(44) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564−568.
(45) Mulholland, A. J.; Grant, G. H.; Richards, W. G. *Protein Eng.* **1993**, *6*, 133−147.
(46) Ridder, L.; Mulholland, A. J.; Rietjens, I. M. C. M.; Vervoort, J. *J. Mol. Graphics* **1999**, *17*, 163−175.
(47) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936−2941.
(48) *Jaguar 4.0*, Schrödinger Inc., 1998.

Table 2. Comparison of the Geometries (Å and deg) for the Two (Small and Large) Model Systems (Figure 2) Obtained with B3LYP/6-31+G**, MP2/6-31+G**, AM1, and AM1-SRP

	B3LYP/ 6-31+G**	MP2/ 6-31+G**	AM1	AM1-SRP
		Methanethiolate		
CB2−SG2	1.85	1.83	1.73	1.78
		TS Small Model (Figure 2a)		
CB2−SG2	1.84	1.82	1.74	1.78
SG2−CA4	2.64	2.50	2.31	2.30
CB2−SG2−CA4	99.2	96.3	114.5	108.3
SG2−CA4−CA5	113.4	112.6	116.2	115.6
		Adduct Small Model (Figure 2a)		
CB2−SG2	1.84	1.81	1.76	1.80
SG2−CA4	1.89	1.85	1.78	1.89
CB2−SG2−CA4	104.0	103.6	114.5	103.8
SG2-CA4-CA5	120.3	120.0	116.2	116.8
		TS Large Model (Figure 2b)		
CB2−SG2	1.84		1.74	1.78
SG2−CA4	2.68		2.33	2.33
CB2−SG2−CA4	94.1		109.5	109.5
SG2−CA4−CA5	109.4		115.9	115.8
		Adduct Large Model (Figure 2b)		
CB2−SG2	1.83		1.76	1.81
SG2−CA4	1.89		1.81	1.93
CB2−SG2−CA4	100.2		102.4	100.6
SG2−CA4−CA5	115.9		113.9	113.1

[a] Atom labels were chosen to correspond with atom labels in the crystal structure PDB entry 2GST.

the separate reactants and the transition state (i.e., the "barrier" ΔH_{TS}) but somewhat different results with respect to the exothermicity of the reaction (ΔH_{prod}). AM1 gives a similar value for the barrier but overestimates the exothermicity by at least 8 kcal/mol in comparison to MP2/6-31+G**. This overestimation of the exothermicity by AM1 is possibly due to an underestimation of the stability of the anionic reactant, methylthiolate.

Table 1 also presents the results for AM1-SRP with the specific reaction parameters for sulfur (available in the Supporting Information, Table S1), which were obtained using the genetic algorithm described in the Methods section. The AM1-SRP results are significantly improved over AM1 with respect to exothermicity. The reaction energetics with the new parameters are now very close to the MP2/6-31+G** results and represent the best possible result that could be obtained by optimizing the parameters for sulfur only. With these modified parameters, the exothermicity calculated with the larger model is improved over standard AM1. With respect to the barriers, some uncertainty remains because of the disagreement between the B3LYP/6-31+G** and LMP2/6-31+G**//B3LYP/6-31+G** energies. The AM1 and AM1-SRP barriers are between those obtained at the higher levels. Thus, they seem to provide reasonable values, however, with an uncertainty of a few kilocalories per mole.

Table 2 shows that, although the AM1 parameter optimization for sulfur was set up to improve the energetic properties and not the geometrical properties (i.e., geometrical properties were not included in the parameter optimization), the geometries around the sulfur atom in the calculated structures are, nevertheless, improved relative to standard AM1. This confirms that the sulfur parameters obtained provide a better description of

Table 3. Solvent Interaction Energies (kcal/mol), Calculated as the Energy Difference between the Solvent Complexes Presented in Figure 3 and the Isolated (Water and Reactant/Product) Molecules[a]

	methanethiolate	oxirane[b]	product 1	product 2
HF/6-31G*	−13.80	−5.75	−6.35	−22.40
B3LYP/6-31+G**[c]	−15.38	−6.15	−6.95	−22.02
MP2/6-31+G**[c]	−15.93	−7.48	−7.16	−22.74
AM1	−11.98	−3.17	−3.94	−14.82[d]
AM1-SRP	−11.85	−3.17	−4.06	−14.67[d]
AM1-SRP/MM	−15.92	−4.33	−7.74	−15.09
AM1-SRP/MM-SRP	−15.41	−5.03	−7.34	−18.08

[a] MM indicates the standard CHARMM force field;[26] MM-SRP refers to optimized VDW parameters for the sulfur and oxygen (O5) atoms in the QM region. [b] High-level QCISD(T)/6-311++G(2d,p) calculations with counterpoise correction for basis set superposition error (BSSE) are reported to yield −5.4 kcal/mol.[49] [c] Calculated counterpoise (CP) corrections indicate that the basis set superposition error is small with the basis set used. (B3LYP/6-31+G**, <0.7 kcal/mol; MP2/6-31+G**, <2.7 kcal/mol). [d] AM1 and AM1-SRP predict wrong (bifurcated) geometries for these complexes. In contrast, AM1-SRP/MM and AM1-SRP/MM-SRP geometries are correct.

the specific reaction of interest than the standard AM1 parameters.[37]

The results of the testing and optimization of the nonbonded QM/MM interactions of the thiolate sulfur and the epoxide oxygen with water (Figure 3) are summarized in Tables 3 and 4. The higher levels of theory (HF/6-31G*, B3LYP/6-31+G**, MP2/6-31+G**) agree on the interaction energies to within 2 kcal/mol. The hydrogen bond distances are systematically shorter by the electron-correlated methods than by HF. In pure QM calculations, AM1 and AM1-SRP underestimate the hydrogen bond interaction energies in comparison to the higher level results. Also, the geometries for the O···H−O hydrogen bonds are poor. These observations are in line with previous studies showing that AM1 generally underestimates hydrogen bond energies and may predict poor (e.g., bifurcated) geometries for hydrogen bonds from water molecules[36,50−52] It is encouraging to observe that the QM/MM interaction energies, as well as the QM/MM geometries, are much better than AM1, even when standard MM VDW parameters are used for the QM atoms. Similar improvements of QM/MM interactions over pure semiempirical QM results have been found by others.[10]

By adjusting the nonbonded parameters of the thiolate sulfur and the epoxide oxygen (parameters available in the Supporting Information, Table S2), further improvement of the QM/MM interactions (the improved Lennard-Jones parameters are designated MM-SRP in Tables 3 and 4) was achieved. Although significantly improved, the stabilization by the hydrogen bond involving the hydroxylate O5 oxygen of the product (column 5 in Table 3) could not be fully recovered by fitting the VDW parameters, without yielding unrealistically short hydrogen bonds. This impossibility to obtain a perfect fit for the Lennard-Jones parameters might be due to an intrinsic limitation of the current QM/MM methods, which neglect polarization of the MM atoms and which employ fixed Lennard-Jones parameters for QM atoms that may change their chemical nature during a reaction. Nevertheless, the reparametrization has

(49) Caminati, W.; Moreschini, P.; Rossi, I.; Favero, P. G. *J. Am. Chem. Soc.* **1998**, *120*, 11144−11148.
(50) Rzepa, H. S.; Yi, M. Y. *J. Chem. Soc., Perkin Trans. 2* **1990**, 943−951.
(51) Zheng, Y. J.; Merz, K. M. *J. Comput. Chem.* **1992**, *13*, 1151−1169.
(52) Mulholland, A. J.; Richards, W. G. *J. Phys. Chem. B* **1998**, *102*, 6635−6646.

Table 4. Hydrogen Bond Geometries (Å) in the Solvent Complexes Presented in Figure 3, Characterized by the Distances to Both the (Bonding) Hydrogen and Oxygen Atoms of the Water Molecule

	methanethiolate		oxirane		product (1)		product (2)	
	S–HW	S–OW	O–HW[a]	O–OW	S–HW	S–OW	O–HW	O–OW
HF/6-31G*	2.49	3.40	2.03	2.90	2.73	3.58	1.69	2.66
B3LYP/6-31+G**	2.24	3.23	1.89	2.82	2.49	3.44	1.52	2.56
MP2/6-31+G**	2.21	3.22	1.90	2.80	2.51	3.45	1.52	2.56
AM1-SRP/MM	2.24	3.13	2.01	2.97	2.47	3.34	1.92	2.77
AM1-SRP/MM-SRP	2.29	3.18	1.84	2.79	2.53	3.40	1.65	2.46

[a] Experimental value, 1.92 Å.[49]

resulted in improved QM/MM interactions with a residual maximum error of 4 kcal/mol in the QM/MM interactions involving the hydroxylate oxygen of the product. The latter may cause the exothermicity of the reaction in the full QM/MM model to be somewhat underestimated.

It should be noted that, in the full QM/MM model, the important hydrogen bond between Tyr6 and the thiolate sulfur of glutathione is treated entirely AM1-SRP (as both the side chain of Tyr6 and the thiolate group of glutathione are included in the QM region). Test calculations indicate that the strength of this hydrogen bond is somewhat underestimated by AM1-SRP, very similar to the results obtained for hydrogen bonds from water (Table 3, column 2) and in line with previous studies of hydrogen bonds by AM1.[36,51,52] However, in line with the results in Table 3 (columns 2 and 4), this underestimation (compared to, for example, B3LYP/6-31+G**) is likely to be similar in the reactants and the products. Thus, as this hydrogen bond is maintained throughout all simulations (see paragraph on the hydrogen bond analysis), the underestimation in hydrogen bond stabilization is similar along the entire reaction coordinate and should therefore not affect the calculated free energy profiles significantly.

Free Energy Profiles. An important difference in the two X-ray structures, which provide the initial coordinates for the QM/MM simulations, is observed in the hydrogen bonding around the O5 hydroxyl moiety. In the X-ray structure with the (9*S*,10*S*) ligand (2GST), a clear hydrogen-bonding interaction between O5 of the ligand and the hydroxyl group of Tyr115 is present,[6] which was preserved during equilibration dynamics of the reactant state of this model. However, in the other X-ray structure with the diastereomeric (9*R*,10*R*) ligand (3GST), no such direct hydrogen bond is observed. Instead, a bound water is observed that appears to connect the O5 and Tyr115 through a hydrogen-bonding network.[6] During the first 2 ps of equilibration dynamics for this second starting model (after the structure was minimized toward the reactant state as described in Methods), this water left its interconnecting position and a direct hydrogen bond between the O5 and Tyr115 was restored. This observation suggests that a configuration with a direct hydrogen bond between O5 and Tyr115 is a better representation of the reactive (enzyme−substrate complex) active site. Therefore, all subsequent calculations described hereafter were based on the first model, built from the crystal structure with the (9*S*,10*S*) diastereomer, in which this hydrogen bond is present from the beginning. Starting from this equilibrated model, two series of 20-ps simulations were performed, to sample each of the two reaction coordinates leading to the two diastereomeric products. The free energy profiles obtained on the basis of these simulations are presented in Figure 4. Very similar profiles, with almost identical barriers, were obtained when the analysis was done using data from the first 10 ps of the various simulations only or using data from the last 10 ps. This indicates that the individual simulations were long enough to obtain converged statistics.

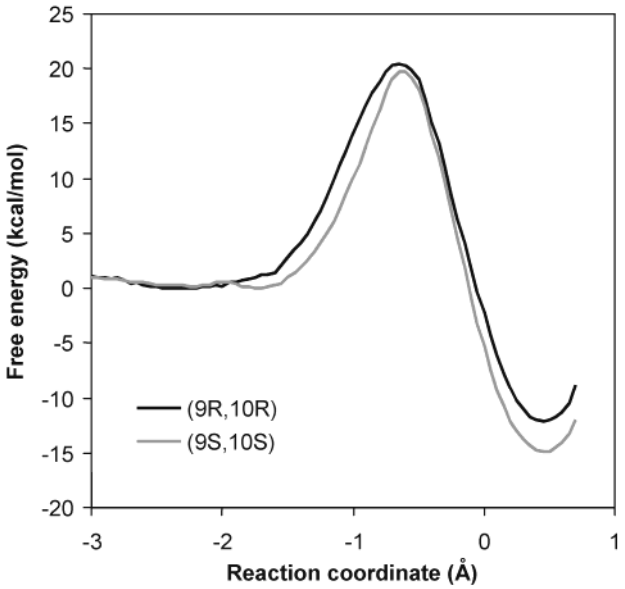

Figure 4. Free energy profiles for the reactions leading to the two diastereomeric products with (9*S*,10*S*) and (9*R*,10*R*) chirality.

The free energy barriers for the two reactions are very similar, which is in agreement with the experimental observation that the two products are formed in similar amounts.[5] Also, the value for the free energy barrier of ∼20 kcal/mol compares favorably with the experimental catalytic rate constant of 0.4 s^{-1} at 298 K,[53] which corresponds to an experimental activation energy of ∼18 kcal/mol. The difference of 2 kcal/mol between the calculated and experimental value is within the uncertainty of the simulation system, for example, due to some uncertainty in the gas-phase barrier by the AM1-SRP method as described above (Table 1, column 4). The maximums in the free energy profiles correspond to a reaction coordinate value of ∼−0.6 Å. Figure 5 shows the average structures of the two simulations with the reaction coordinates restrained to −0.6 Å. These structures can be interpreted as approximate transition states on the free energy surface as defined by the present QM/MM model.

The free energy profiles indicate that the simulated reaction step is exergonic by 10−15 kcal/mol. On the basis of Tables 1 and 3, the model uncertainty with respect to this exergonicity can be estimated. Table 1 indicates that the AM1-SRP model

(53) Johnson, W. W.; Liu, S. X.; Ji, X. H.; Gilliland, G. L.; Armstrong, R. N. *J. Biol. Chem.* **1993**, *268*, 11508−11511.

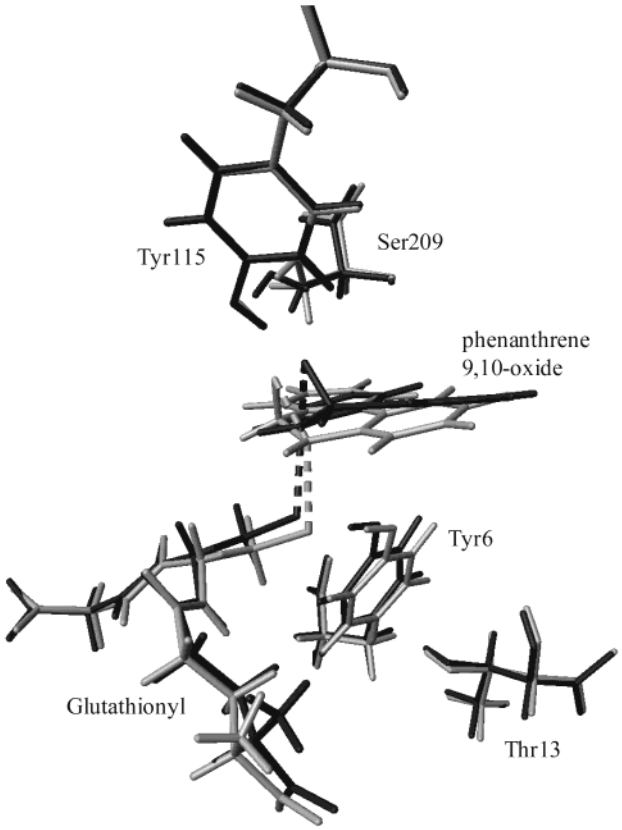

Figure 5. Overlay of the average structures during two (20-ps) simulations with the reaction coordinates (leading to the two diastereomeric products) restrained to −0.6 Å, representing approximate transition states on the QM/MM free energy surface. The transition-state structure leading to the (9*S*,10*S*) diastereomer is colored lighter than the transition-state structure leading to the (9*R*,10*R*) diastereomer.

employed is very close to LMP2/6-31+G** with respect to exothermicity of the reaction, but the difference with B3LYP/6-31+G** leaves the possibility of an overestimation of the exothermicity by a few kilocalories per mole. On the other hand, Table 3 indicates that the stabilization of the hydroxylate moiety of the product, modeled through QM/MM interactions, may be underestimated, also by a few kilocalories per mole. Thus, the exergonicities indicated by the free energy profiles (Figure 4) may contain errors of in the order of at most 5 kcal/mol, which may partly cancel out. These errors do not affect the conclusion that the simulated reaction is exergonic in nature.

The agreement between the calculated activation barriers and the experimental rate constant, and the reasonable exothermicity of the simulated reaction step, support the mechanism underlying the present model. Although our calculations cannot rule out a concerted protonation of the O5 oxygen of the product (e.g., by Tyr115), the calculations show that a strictly concerted protonation is not required to make the epoxide ring opening exergonic in nature. However, at physiological conditions, as the active site is highly solvent accessible and the pK_a of the product is well above 7 (typically for a secondary alcohol, pK_a > 10), protonation may occur readily after formation of the product in its anionic form. In general, free energy barriers to proton transfers are much lower than the barrier calculated for the epoxide ring opening.[54] Especially proton transfers along preformed hydrogen bonds are very fast.[54] Thus, a proton could initially come from Tyr115, which, in turn, could be protonated from solution. (Formally, the proton comes from the thiol group of glutathione: the thiol is deprotonated upon binding to active site, due to lowering of its pK_a through active site stabilization of the thiolate.)

(54) Guthrie, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 12886−12890.

Table 5. Root-Mean-Square Deviations (RMSD, in Å) for the QM Atoms, between the Crystal Structures of the Product Complexes and the Average Coordinates during Several Simulations of the Same Product[a]

RMSD	from initial models		from free energy pathway	
diastereomer	anionic	neutral	anionic	neutral
9*S*,10*S*	1.41 (0.74)	0.64 (0.62)	1.42 (0.82)	0.60 (0.76)
9*R*,10*R*	1.71[b] (0.59)	0.31 (0.56)	2.51[c]	2.17[c]

[a] Values for all unconstrained atoms (i.e., <14 Å from center) are given in parentheses. [b] This simulation crashed after 10 ps of equilibration due to problems in SCF convergence in the QM structure. The value is for the average structure over the 5−10-ps equilibration dynamics. [c] Note that these RMSDs are between simulations based on the 2GST crystal structure and the 3GST crystal coordinates. These crystal structures, however, overlay well enough (e.g., the RMSD for the CA backbone atoms between 2GST and 3GST is 0.14 Å) for the RMSD analysis to be made for the QM atoms.

Structural Comparison between the Simulations and the Crystal Structures. Simulations of four different product complexes were performed as described in the Methods section, starting from the crystal structure coordinates. Table 5 presents the root-mean-square differences (RMSD) for the QM atoms between the crystal structures and average structures of various simulations of the product complexes. For both crystal structures, the initial geometries of the QM atoms are maintained best by the models containing the neutral products (protonated on O5). This is indicated by small RSMD values, i.e., <0.7 Å, similar in magnitude to RMSD values for all unconstrained atoms in the model (given in parentheses in Table 5). The simulations of the deprotonated products result in significantly larger deviations in the average structures of the QM atoms (>1.4 Å). This suggests that the crystal structures indeed represent the neutral (protonated products), rather than the anionic intermediate forms of the products formed initially upon glutathione conjugation.

As a further test, Table 5 also shows RMSD values for another four simulations, which correspond to the end points of the free energy pathways. The RMSD value obtained from the end point simulation of the free energy pathway toward the (9*S*,10*S*) diastereomer (which involves the deprotonated product) is of the same magnitude as the value obtained from the simulation starting from the initial model (with the crystal structure coordinates). When the product in this end point simulation was protonated and further simulation (10-ps equilibration and 20-ps simulation) was performed, the average structure became more "crystal structure-like" as indicated by the decrease in the RMSD value. This further supports the suggestion that the crystal structure represents the neutral product, formed upon protonation of the initial anionic product. The RMSD value from the end point simulation of the free energy pathway towards the (9*R*,10*R*) diastereomer is significantly larger, which is due to a distinct difference in the active site configuration. In the crystal structure (3GST), a water molecule forms a hydrogen-bonding chain between Tyr115 and the O5 hydroxyl moiety of the product. During the pathway simulations, however, this water bridging interaction was not

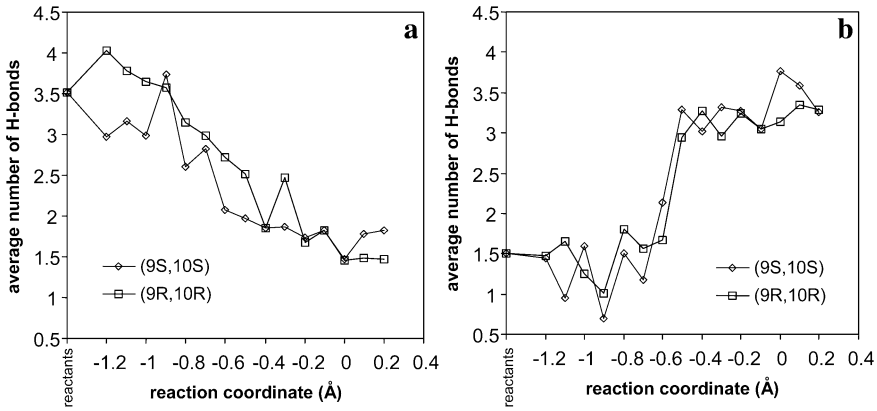

Figure 6. Average number of hydrogen bonds with (a) the thiolate sulfur and (b) the oxirane oxygen, plotted for the various simulations along the reaction coordinate.

observed. Instead, a direct hydrogen bond was present between Tyr115 and O5. After protonation of O5 in the product simulation, this direct hydrogen bond was maintained during the 15-ps equilibration and 20-ps subsequent dynamics.

Active Site Interactions. Several active site interactions, which are apparent from the crystal structure,[6] are conserved during the dynamics simulations. First, the interaction between the thiolate sulfur and Tyr6 is present throughout all simulations. Several theoretical studies[55−57] have suggested a role for Tyr6 in stabilizing the reactive, deprotonated form of glutathione in the active site. This stabilization of deprotonated glutathione is experimentally indicated by its low pK_a of 6.2 in the native enzyme, compared to 7.8 in the Tyr6Phe mutant and 9.0 in solution.[55,58]

A second important interaction, already mentioned above, is between the substrate oxygen and Tyr115. On the basis of the crystal structure of GST in complex with the (9*S*,10*S*) diastereomer of the product, it was proposed that Tyr115 could play a role in catalysis by providing "electrophilic assistance" in the addition of glutathione to phenanthrene 9,10-oxide, through a direct hydrogen bond to the oxirane oxygen in the transition state.[6] This was supported by a 100-fold decrease in activity of the Tyr115Phe mutant toward this substrate.[53] A direct hydrogen bond is observed throughout the present simulations. As is clear from the results of the modeling of the mutation to phenylalanine (see below), this hydrogen bond interaction becomes stronger as the oxygen becomes increasingly anionic during the reaction, thereby stabilizing the transition states and products relative to the reactants.

A number of "second-sphere" interactions have been identified in the crystal structures of GST M1-1.[6] Most remarkably, an on-face hydrogen bond was proposed between Thr13 and the π-electron density of the aromatic ring of Tyr6.[6,55,58] This interaction has been proposed to lower the proton affinity of Tyr6 by stabilizing the electron density of the π system. This decreased proton affinity would result in a stronger hydrogen bond to the thiolate group of the enzyme-bound glutathione and a further decrease of the pK_a of glutathione within the active site. pH-dependence studies of the Thr13Ala and Thr13Val mutants indicate that the on-face hydrogen bond may contribute to the lowering of the pK_a of enzyme-bound GSH by ~0.7 log unit.[55,58] This effect was qualitatively reproduced by ab initio calculations on small models.[55] The on-face hydrogen bond interaction is conserved well in the present simulations, modeled as a QM/MM interaction.

Two other previously identified second-sphere interactions, between the backbone N−H moiety of Leu12 and the hydroxyl moiety of Tyr6[58] and between Ser209 and Tyr115,[6] are also observed in the present simulations (Ser209 being the donor). The latter hydrogen bond, however, was frequently replaced by a direct interaction between Ser209 and the oxirane oxygen. This may reflect the likely possibility that the protonation state during the simulated reaction is different from the protonation state in the crystal structure as discussed above. The direct interaction will be stronger with the negatively charged hydroxylate oxygen in the initial deprotonated form of the product than with the neutral (protonated) product likely to be present in the crystal structure. Thus, the simulations suggest that Ser209 could play a more important role in catalysis (through a direct hydrogen bond) than the crystal structures suggest (i.e., through a second-sphere interaction).

Hydrogen Bond Analysis. To gain insight into the effect of desolvation of the glutathione thiolate along the reaction coordinate, the various dynamics trajectories along the reaction coordinate were analyzed for hydrogen bonds between the thiolate sulfur and nearby solvent and protein residues. For each simulation, 100 structures, taken from the dynamics trajectory every 0.2 ps, were analyzed. For analysis purposes, a hydrogen bond to sulfur was defined as a hydrogen to sulfur distance of less than 2.8 Å with a donor−H···S angle greater than 90°.[59,60] The results of the analysis were not highly sensitive to the precise choice of cutoff parameters. In Figure 6a, the average number of hydrogen bonds, using this definition, is plotted as a function of the reaction coordinate. The average number of hydrogen bonds to sulfur clearly decreases, from about 3.5 to 1.5, as the reaction proceeds, indicating a "desolvation" of the thiolate. Further analysis showed that, as mentioned above, the hydrogen bond with Tyr6 persists throughout all simulations. All other hydrogen bonds to the thiolate sulfur, and the changes in the average number, involve water molecules. A similar

(55) Liu, S. X.; Ji, X. H.; Gilliland, G. L.; Stevens, W. J.; Armstrong, R. N. *J. Am. Chem. Soc.* **1993**, *115*, 7910−7911.
(56) Zheng, Y. J.; Ornstein, R. L. **1997**, *119*, 1523−1528.
(57) Rignanese, G. M.; De Angelis, P.; Melchionna, S.; De Vita, A. *J. Am. Chem. Soc.* **2000**, *122*, 11963−11970.
(58) Xiao, G. Y.; Liu, S. X.; Ji, X. H.; Johnson, W. W.; Chen, J. H.; Parsons, J. F.; Stevens, W. J.; Gilliland, G. L.; Armstrong, R. N. *Biochemistry* **1996**, *35*, 4753−4765.
(59) Hadfield, A. T.; Mulholland, A. J. *Int. J. Quantum Chem.* **1999**, *73*, 137−146.
(60) Braatz, J. A.; Paulsen, M. D.; Ornstein, R. L. *J. Biomol. Struct. Dyn.* **1992**, *9*, 935−949.
(61) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−115.

Table 6. Effects (in kcal/mol) of Four Selected Mutations on the Barriers and Reaction Energies Obtained from Average Changes in the QM Energy (QMEL Term) over 100 Snapshots Taken from the Simulations of the Reactants, the Transition State ($r = -0.6$ Å), and the Products[a]

	(9*S*,10*S*) diastereomer		(9*R*,10*R*) diastereomer	
mutant	$\Delta\Delta E_{act}$	$\Delta\Delta E_{reaction}$	$\Delta\Delta E_{act}$	$\Delta\Delta E_{reaction}$
Asn8Asp	1.6	2.33	4.99	3.12
Thr13Ala	0.75	0.45	−0.11	0.42
Tyr115Phe	4.44	8.78	8.03	15.97
Ser209Ala	3.69	6.15	3.79	1.95

[a] Data available in the Supporting Information, Table S3.

analysis was performed for hydrogen bond interactions with the epoxide oxygen. Here, hydrogens closer than 2.5 Å from the epoxide oxygen, where the donor—H···O angle is greater than 90° were considered to form hydrogen bonds.[59] The results, shown in Figure 6b, suggest a clear change in solvation as the system crosses the barrier ($r = -0.6$ Å). At this point, the average number of hydrogen bonds changes from ∼1.5 to more than 3. Further analysis showed that, during the dynamics, the epoxide oxygen sometimes exchanged hydrogen bonds from solvent with hydrogen bonds from the active residues Tyr115 and Ser209 and the other way around.

These results should be interpreted with care. The plotted points along the *x*-axis are based on the various simulations that in fact cover overlapping regions of the reaction coordinate. Also, the simulations are too short to obtain highly accurate statistics for hydrogen bonding. Therefore, Figure 6 provides only a qualitative picture of the (de)solvation along the reaction coordinate. Nevertheless, the figure provides a clear indication of changes in solvation during the reaction. The thiolate is well solvated initially. Along the reaction coordinate, sulfur becomes less ionic and the stabilizing interactions with solvent molecules become weaker (Table 3, columns 2 and 4). Also, the solvent accessibility of the sulfur atom becomes smaller as the reaction proceeds. For the epoxide oxygen (O5), an opposite effect is observed: the oxygen becomes more ionic as the reaction proceeds and the stabilizing interactions with hydrogen bond donating groups increase in number and strength (Table 3, columns 3 and 5). These solvation effects will have a significant influence on the reaction energetics.

Effects of Active Site Mutations. Four mutations were selected for analysis of their effects on the energetics of the reaction. The Tyr115Phe mutation has been reported to result in a decrease in activity toward the phenanthrene 9,10-oxide substrate by 2 orders of magnitude, without changing the stereoselectivity of the reaction.[53] The Thr13Ala mutation has been shown to result in a rise in pK_a of the enzyme-bound glutathione by 0.7 unit.[55,58] Two other mutations, Ser209Ala and Asn8Asp (as well as Thr13Ala), represent differences between the isoenzymes. The M1-1 and M2-2 isoenzymes catalyze the conjugation of glutathione to phenanthrene 9,10-oxide with significant differences in efficiency and stereoselectivity.[2,5] The selected mutations could be modeled by the method described in the Methods section. The results are presented in Table 6.

The Tyr115Phe mutation results in an increase in the barriers of the modeled reactions. This can be explained by the removal of the hydrogen-bonding hydroxyl group, which, in the wild-type enzyme, stabilizes the O5 oxirane oxygen atom increasingly as the reaction proceeds. This result agrees with experiment, which shows a decrease in the catalytic rate constant in the Tyr115Phe mutant,[53] and with the proposal that Tyr115 assists the reaction.[6] The 100-fold decrease in catalytic activity in the mutant corresponds to an increase in the activation barrier of ∼3 kcal/mol. The present (approximate) results suggest an increase of 4−8 kcal/mol (Table 6). A possible explanation for this overestimation could be that in the actual mutant additional hydrogen-bonding interactions between the oxirane oxygen and water may partially compensate for the absence of the interaction with a tyrosine. Such possible additional interactions with water are not accounted for in the present modeling of this mutation.

The results for the Ser209Ala mutation indicate a relatively strong stabilizing effect of this residue, which has previously been proposed to be involved via a second-sphere interaction, i.e., by hydrogen bonding to Tyr115.[6] In the present dynamics simulations, however, Ser209 often interacts directly with the oxirane oxygen, especially toward the product state (in which the oxygen has a formal negative charge). The energetic effect is illustrated by comparing the simulations of two diastereomeric products. In the simulation of the (9*S*,10*S*) product, the direct interaction between Ser209 and the oxirane oxygen plays an important role and is present in a significant fraction of the snapshots analyzed, whereas in the simulation of the (9*R*,10*R*) product, it is practically absent. This explains the smaller effect (by 4 kcal/mol) of mutating Ser209 to Ala in the simulation of the (9*R*,10*R*) product, compared to the simulation of the (9*S*,-10*S*) product.

The results for the Thr13Ala mutation indicate a stabilizing effect of the on-face hydrogen bond between Thr13 (MM) and Tyr6 (QM) on the QM system (see Supporting Information). However, the effects of mutating the Thr13 to Ala on the barriers and reaction energies are too small, compared to the accuracy of the analysis (see Supporting Information), to provide significant insight into the role of the on-face hydrogen bond in terms of catalysis of the conjugation step. This second-sphere interaction, propagated through the aromatic system of Tyr6, probably requires higher level QM treatment to properly represent its catalytic effect.

A (longer range) electrostatic effect can be expected from the mutation of Asn8 present in the M1-1 isoenzyme to Asp8 present in the M2-2 isoenzyme. In the present simulations, the mutation increases the barrier of the reaction toward the (9*R*,-10*R*) product significantly more than the barrier toward the (9*S*,-10*S*) product (by a few kcal/mol). This is in good agreement with the fact that M2-2 isoenzyme preferentially forms the (9*S*,-10*S*) product, whereas the M1-1 isoenzyme forms both products in approximately equal amounts.[5] It is important to mention that the Asn8 residue is at the surface of the protein. Its electrostatic effect is therefore likely to be lowered through solvent shielding and possible interactions with counterions. Nevertheless, the present results suggest that the differently charged residues at the 8-position could be one of the factors causing the different ratios of the diastereomeric products formed by both isoenzymes.

Conclusions

The approach employed in the present study, which combines a fast and accurate AM1-SRP treatment with QM/MM dynamic free energy simulations, allows accurate calculations of many enzyme-catalyzed reactions. This approach successfully models

the conjugation of glutathione to phenanthrene 9,10-oxide, catalyzed by glutathione *S*-transferase M1-1. A selection of critical parameters in the QM/MM model, i.e., the AM1 parameters for sulfur and the QM/MM nonbonded interaction parameters for sulfur and oxygen, have been optimized for the conjugation reaction to give results of an accuracy comparable to ab initio level methods including electron correlation. The free energy profiles obtained compare favorably with the experimental rate constant for the overall reaction and with the observed ratio of diastereomeric products formed. This supports the reaction modeled in the present simulations as the rate-limiting step in the overall reaction cycle, as well as supporting the mechanism underlying the present model. The simulations indicate that the conjugation of deprotonated glutathione to the substrate, resulting in an anionic product with a charged hydroxylate oxygen (O5), is exergonic by 10−15 kcal/mol and therefore does not require protonation of O5 in a strictly concerted process. However, protonation of O5 can be expected to occur readily after epoxide ring opening. Structural comparison of various simulations of the product state with the crystal structures suggests that the latter represent complexes with the neutral (protonated) products.

The present simulations yield insight into the role of water in the highly solvent accessible active site of glutathione *S*-transferase. They provide a detailed picture of the desolvation of the thiolate sulfur, involving a decrease in solvent hydrogen bonds, and solvation of the oxirane oxygen, through an increase in hydrogen bonds, along the reaction pathway. Analysis of the simulations and modeling of several active site mutations provides insight into the catalytic role of several active site interactions and allows the study of the variation in catalytic specificity of different isoenzymes. The catalytic role of a direct hydrogen bond between the epoxide oxygen and Tyr115 in the reaction is supported by the observation that this interaction is maintained throughout the present simulations. This interaction stabilizes the epoxide oxygen, which becomes more anionic as the reaction proceeds. Modeling of the Tyr115Phe mutation shows an increase in the barrier for conjugation, in agreement with experiments showing a decrease in catalytic activity of this mutant.[53] The results also suggest that an indirect interaction with Tyr115 mediated by a bound water molecule, present in one of the X-ray structures (3GST), may not be relevant to the reactant complex and the rate-limiting conjugation step. Ser209 may not only be involved through a "second-sphere" hydrogen bond to Tyr115 but may also directly interact with the epoxide oxygen changing to a negatively charged hydroxylate oxygen along the reaction coordinate. In line with this possibility, the modeled Ser209Ala mutation results in an increase in barrier, which is indicative of a catalytic role of Ser209. More detailed simulations would be required to further establish the exact role of Ser209 in catalysis.

Finally, the results obtained for the Asn8Asp mutation, representing one of the differences between the M1-1 and M2-2 isoenzymes, suggest a differential effect of this mutation on the reactions leading to the two diastereomeric products. A more significant increase in the barrier toward the (9*R*,10*R*) diastereomer, compared to the barrier toward the other product, is in line with the change in product ratio between M1-1 and M2-2 isoenzymes. Thus, the mutation at position 8 is identified as a possible factor causing the difference in stereospecificity toward of the diastereomeric products between the M1-1 and M2-2 isoenzymes.

Acknowledgment. This research has been supported by a Marie Curie Fellowship of the European Community program Quality of Life within the fifth framework under Contract QLRI-CT-1999-51244. A.J.M. is an EPSRC Advanced Research Fellow. We acknowledge the reviewer for bringing the paper by Caminati et al.[49] to our attention.

Supporting Information Available: Three tables listing AM1-SRP parameters for sulfur, modified van der Waals parameters for sulfur and oxygen and average effects of modeled mutations on the QM energies. Coordinates of optimized structures and complexes presented in Figures 2 and 3 (results presented in Tables 1−4). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0256360