Free Energy Calculations on Disulfide Bridges Reduction in Proteins by Combining ab Initio and Molecular Mechanics Methods

Catalina David and Mironel Enescu*

Laboratoire de Chimie Physique et Rayonnement, UMR CEA E4, University of Franche-Comte, 16 route de Gray, 25030 Besancon, France

Received: October 29, 2009; Revised Manuscript Received: January 19, 2010

Free energy profiles were calculated for the reduction of the four disulfide bridges in lysozyme by tris(2-carboxyethyl)phosphine (TCEP). The computational method combines high-precision density functional theory (DFT) calculations performed on the core of the reactant system with classical mechanical free energy evaluations based on the sampling of the configuration space of reaction environment. The predicted reaction energy barriers are in satisfactory agreement with experimental data, proving that the present method provides a reliable description of the mechanism of reaction. The role of the protein environment in this mechanism is further emphasized by analyzing the different contributions to the free energy profiles. It is shown that the protein environment affects the reaction by three factors: polarizability, steric hindrance of the reactant site, and S−S bridge distortion due to structural constraints. The corresponding effects are quantitatively evaluated, and the results are discussed in connection with the current two-step reaction model for the reduction of S−S bridges in proteins.

Introduction

Disulfide bridges (S−S) play a major role in the structural stability and folding pathways of proteins.[1] Consequently, their reduction is a common and efficient method in studying conformational properties of proteins.[2] Among the reducing agents used for cleaving S−S bridges, tris(2-carboxyethyl)phosphine (TCEP) is particularly interesting because in this case the reaction (Scheme 1) is practically irreversible.[3] The rate-determining step is the substitution of one S atom in the S−S bridge by the P atom,[4] and it is followed by the hydrolysis of the intermediate product.

The reactivity of a S−S bridge is clearly dependent on the reaction environment.[5,6] As for the most reactions occurring in protein environment, the main factors governing this reactivity are expected to be the accessibility of the reactive groups and the polarizability of their environment. Conformational opening events (local unfolding) are supposed to play an important role in the accessibility of the S−S bridge by the reducing agent.[7] Another factor which is specific to this reaction is the protein unfolding subsequent to the S−S bridge breaking that contributes to the stabilization of the reaction product.[8] The role of the protein structure in the reactivity of S−S bridges was directly demonstrated in the case of the protein disulfide isomerase (PDI) that is a catalyst of disulfide formation in the eukaryotic endoplasmatic reticulum.[9]

Lysozyme is an excellent model for studying the mechanism of S−S bridge reduction. Among its four S−S bridges, one is solvent exposed while the three others are more or less buried. The experimental kinetics of lysozyme reduction by TCEP has been recently reported.[10] Its heterogeneity strongly suggests that the four S−S bridges have different reaction rate constants. The aim of the present paper is to highlight the role of the protein environment in the reactivity of the four S−S bridges in lysozyme by analyzing the corresponding theoretical free energy curves.

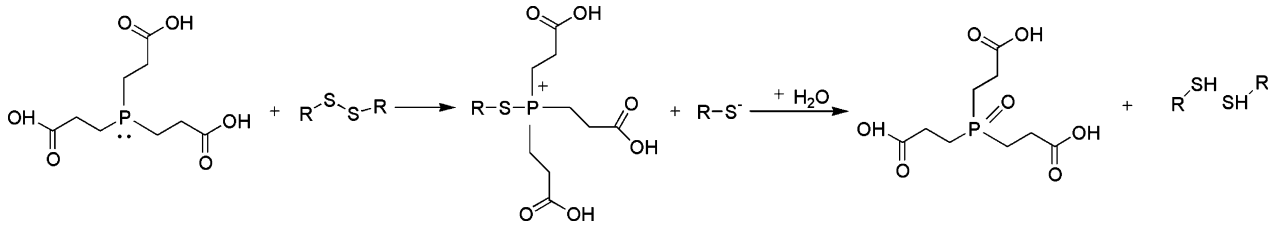
Molecular modeling is becoming a favorite tool in analyzing the mechanisms of chemical reactions in protein environment. The quantum mechanics/molecular mechanics (QM/MM) hybrid methods[11,12] provide accurate values for the energy of the reactant system that include the environment contribution. However, this energy is very sensitive to the configuration of the whole system which can present significant variations in the case of large systems. Hence, the potential energy profiles obtained by simply driving the reaction coordinate could be less relevant for the reaction kinetics in this case. Instead, free energy calculations by sampling the configuration space are necessary. The QM/MM sampling was successfully applied in the study of some enzymatic reactions. In this kind of calculations, a semiempirical method is used in treating the QM subsystem (model system)[13−16] because of the very high computational cost of the QM/MM sampling with an ab initio QM method. However, in many cases, the semiempirical treatment of the reaction is not satisfactory. An alternative approach is based on the ab initio QM/MM (QM(ai)/MM) sampling of Warshel and co-workers[17,18] in which the sampling of the configuration space is performed using an empirical reference potential (usually the empirical valence bond potential, EVB). Then the free energy changes between the system described by the reference potential and by the density functional theory are calculated with the free energy perturbation method. Along the same line, it was shown that the ab initio QM/MM simulation can be accelerated by using an average field approach for describing bulk solvent interactions.[19,20]

A good compromise between accuracy and computational cost is realized by the quantum mechanics free energy method (QM-FE) where the sampling is limited to the MM subspace while preserving the ab initio treatment of the QM subsystem. In the original formulation of this method proposed by Jorgensen and co-workers,[21,22] a reaction trajectory is generated by performing

* Corresponding author. E-mail: mironel.enescu@univ-fcomte.fr. Fax: (+33) 3 81 66 65 22. Tel.: (+33) 3 81 66 65 21.

10.1021/jp910340t
Published on Web 02/04/2010

SCHEME 1: Disulfide Bridge Reduction by TCEP



gas-phase QM calculations on a model system. Structures extracted on that trajectory are mapped into protein (or solvated), and the environment contribution to the free energy is calculated by performing molecular dynamics simulations at the MM level with frozen QM coordinates. The main limitation of this first formulation of the QM-FE method is that the constraints imposed on the QM system by the protein environment are not taken into account when calculating the quantum energy E_{QM}. Moreover, the method is applicable unless the reaction trajectory in the protein environment is close enough to that in the gas phase. Both difficulties are overcome if, following Yang and co-workers, the reaction trajectory is determined by performing QM/MM optimizations with electronic embedding that allows partial atomic charges variation.[23−25] Although in the absence of conformation averaging the values of the total QM/MM energy on that trajectory have very limited relevance, the QM part of this energy is expected to be more stable giving rise to a reliable QM reaction trajectory. The free energy variation along that trajectory due to the interaction between the model system and its environment is then calculated by sampling the MM subspace. A drawback of this improved QM-FE method arises from the possible convergence failure of the QM/MM optimization with electronic embedding.[12] Indeed, the standard QM/MM optimization algorithms for large size molecular systems are based on an iterative procedure: the QM and the MM subsystems are alternatively optimized and alternatively frozen until a simultaneous convergence is achieved.[26] In the case of strong QM/MM coupling the convergence is not guaranteed.

Given the molecular size of TCEP, reaction 1 should occur at the protein surface which means that the reaction environment is highly polarizable. Moreover, the product of the first reaction step is very polar, and thus the coupling between reactants and environment is strong. It is not surprising that ab initio studies on model systems[4] strongly suggest that this first reaction step is not possible in the gas phase. Consequently, the QM-FE original scheme is not applicable in the present case. On the other hand, our QM/MM test optimizations of the reaction product using electronic embedding failed because of the strong QM/MM coupling. An alternative solution is to determine the starting reaction trajectory by ab initio calculations on the model system in aqueous solution using the polarizable continuum model (PCM[27]), as already suggested for enzymatic reaction by Kollman et al.[28] This choice is appropriate for studying reaction 1 since the polarizability of the reaction environment is of the same order of magnitude as that of the pure solvent. The presently reported QM-FE calculations on the S−S bridges reduction in lysozyme are based on this approach. The resulting QM-PCM optimized structures of the model system were reoptimized in the protein environment at the QM/MM level with frozen reaction coordinate and electronic embedding. This procedure allowed us to take into account the environment constraints when calculating the QM energy. It provides mechanical and electronic coupling between model system and its environment, although the coupling is static. The environment contribution to the profile of the total free energy was then calculated by performing MM molecular dynamic simulations. Usually, in the QM-FE method this contribution is calculated using the free energy perturbation method.[23,24] We propose here an alternative approach based on the thermodynamic integration method (TDI).[29]

Method

Theory. In the subtractive scheme of the hybrid QM/MM method, the potential energy of a molecular system is written as follows:[26]

$$E(X_{R\text{-}M}, X_M) = E_{M;QM}(X_M) + E_{R;MM}(X_{R\text{-}M}, X_M) - E_{M;MM}(X_M) \quad (1)$$

Here R and M denote the real and the model system, respectively, X_M is the ensemble of nuclear coordinates of the model system, $X_{R\text{-}M}$ is the ensemble of nuclear coordinates of the environment, $E_{M;QM}$ is the quantum energy (electronic energy plus nuclear repulsion energy) of the model system, $E_{M;MM}$ is the MM potential energy of the model system, and $E_{R;MM}$ is the MM potential energy of the real system. Let ξ be the reaction coordinate supposed to be dependent on X_M only. The partial partition function related to ξ is

$$Z(\xi) = \int \exp[-\beta E(X_{R\text{-}M}, X'_M, \xi)]\, dX_{R\text{-}M}\, dX'_M \quad (2)$$

where X'_M denotes the ensemble of coordinates of the model system excepting ξ and $\beta = 1/k_BT$ (T is the temperature and k_B the Bolzmann constant). In writing eq 2, kinetics contributions to Z and some multiplying factors were omitted because of their cancellation when free energy differences are calculated. Correspondingly, the partial free energy of the system is given by

$$F(\xi) = -\frac{1}{\beta}\ln Z(\xi) \quad (3)$$

Obviously, the energy appearing in eq 2 is to be calculated using partial atomic charges that are dependent on ξ. The main approximation in the QM-FE method is to neglect the fluctuations of $E_{M;QM}(X'_M,\xi)$ around its mean value on the reaction trajectory:

$$E_{M;QM}(X'_M, \xi) \cong \langle E_{M;QM}(X'_M, \xi)\rangle_\xi \quad (4)$$

Here $\langle\rangle_\xi$ designates averaging over a canonical ensemble generated with fixed ξ. In practice, this mean value is approximated by the quantum energy of the model system resulting from a QM/MM optimization with fixed ξ:

$$\langle E_{M;QM}(X'_M, \xi)\rangle_\xi \cong E_{M;QM}(X_{M;min}(\xi)) \quad (5)$$

$X_{M;min}(\xi)$ are the coordinates of the model system that minimize the QM/MM energy for a given ξ. The partial free energy of the system becomes[23]

$$F(\xi) = E_{M;QM}(X_{M;min}(\xi)) + F_{env}(\xi) \quad (6)$$

where

$$F_{env}(\xi) = -\frac{1}{\beta}\ln(\int \exp[-\beta(E_{R;MM}(X_{R\text{-}M}, X'_M, \xi) - E_{M;MM}(X'_M,\xi))]\, dX_{R\text{-}M}\, dX'_M) \quad (7)$$

The variation of F_{env} when passing between two points, i and $i + 1$, on the reaction trajectory is

$$\Delta F_{env}(i \rightarrow i+1) = F_{env}(\xi_{i+1}) - F_{env}(\xi_i) \quad (8)$$

In the standard formulation of the QM-FE method, $F_{env}(\xi_i)$ is calculated by fixing X_M at $X_{M;min}(\xi_i)$. The free energy variation $\Delta F_{env}(i \rightarrow i+1)$ is thus obtained by perturbing $X_{M;min}(\xi_i)$ with $X_{M;min}(\xi_{i+1})$.[15] However, the subtractive QM/MM scheme offers an alternative to this approach: instead of fixing X_M, one can replace $E_{M;MM}$ in eq 7 by its mean value:

$$E_{M;MM}(X'_M, \xi_i) \cong \langle E_{M;MM}(X'_M, \xi_i)\rangle_{\xi_i} \quad (9)$$

One notes that this approximation is consistent with that assumed in eq 4 for the $E_{M;QM}$ energy. Now, the partial free energy of the system is

$$F(\xi_i) = E_{M;QM}(X_{M;min}(\xi_i)) - \langle E_{M;MM}\rangle_{\xi_i} + F_{MM}(\xi_i) \quad (10)$$

Here $F_{MM}(\xi_i)$ is simply the partial free energy of the real system calculated at the MM level:

$$F_{MM}(\xi_i) = -\frac{1}{\beta}\ln(\int \exp[-\beta(E_{R;MM}(X_{R\text{-}M}, X'_M, \xi_i)]\, dX_{R\text{-}M}\, dX'_M) \quad (11)$$

In the current application of the TDI method to free energy profile calculations, the free energy variation between the states ξ_1 and ξ_2 is expressed as[30,31]

$$\Delta F(\xi_1 \rightarrow \xi_2) \approx \int_{\xi_1}^{\xi_2} \langle F_c\rangle_\xi\, d\xi \quad (12)$$

where F_c is the force of the constraint and $\langle\ \rangle$ designates averaging over canonical ensemble (in writing eq 12 it was assumed that the metric-tensor correction[32] is negligible[31]). When chemical changes are produced, the evaluation of F_c requires molecular dynamic simulation with variable partial atomic charges. We can avoid this complication by using molecular simulations performed on mixed states. Mixing the initial and the final state of the transition is possible if the constraint on ξ is replaced by a harmonic restraint. The modified partial free energy is defined as follows:

$$\bar{F}_{MM}(\xi_i) = -\frac{1}{\beta}\ln(\int \exp[-\beta(\bar{E}_{RMM}(X_{R\text{-}M}, X'_M, \xi;\xi_i)] \times dX_{R\text{-}M}\, dX'_M\, d\xi) \quad (13)$$

where $\bar{E}_{R;MM}$ is the energy of the restrained system:

$$\bar{E}_{R;MM}(X_{R\text{-}M}, X'_M, \xi;\xi_i) = E_{R;MM}(X_{R\text{-}M}, X'_M, \xi) + \frac{k}{2}(\xi - \xi_i)^2 \quad (14)$$

Here k is the force constant of the harmonic restraint. Obviously, if k is strong enough, the variation of the partial free energy of the restrained system upon the displacement from ξ_i to ξ_{i+1} is very close to that of the constrained system. This may be seen by noticing that in the limit of high k the integral with respect to ξ on the right-hand side of eq 13 converges to the Dirac functional (up to a multiplying factor). A detailed analysis of this approximation is given in the Supporting Information. Hence

$$\Delta\bar{F}_{MM}(i \rightarrow i+1) \cong \Delta F_{MM}(i \rightarrow i+1) \quad (15)$$

Now, the free energy variation is calculated by defining a hybrid energy that mixes the i and $i + 1$ states:

$$\bar{E}(\lambda) = (1 - \lambda)\bar{E}_{R;MM}(\xi_i) + \lambda\bar{E}_{R;MM}(\xi_{i+1}) \quad (16)$$

For simplicity, in writing eq 16, the dependence of $\bar{E}_{R;MM}$ on $X_{R\text{-}M}$, X'_M, and ξ was omitted. Formally, the free energy variation between the states i and $i + 1$ is

$$\Delta\bar{F}_{MM}(i \rightarrow i+1) = \int_0^1 \frac{\partial\bar{F}(\lambda)}{\partial\lambda}\, d\lambda \quad (17)$$

According to eqs 2, 3, and 16

$$\frac{\partial\bar{F}(\lambda)}{\partial\lambda} = \langle\bar{E}_{R;MM}(\xi_{i+1}) - \bar{E}_{R;MM}(\xi_i)\rangle_\lambda \quad (18)$$

The average in eq 18 is performed over a canonical ensemble generated with $\bar{E}(\lambda)$ as system energy. The integral in the right-hand side of eq 18 can be evaluated numerically by choosing several values between 0 and 1 for the mixing parameter λ.

Computational Method. The reaction trajectory was determined by performing QM-PCM optimizations on the solvated model system (Figure 1). The calculations were performed with Gaussian 03 program[33] at the B3LYP/6-311+G(d,p) level of theory. The driving coordinate in this series of partial optimizations was the distance between the P atom of TCEP and the nearest S atom of the S−S bridge. As expected, the reaction coordinate is a combination between the S−P and the S−S distances (d(S−P) and d(S−S), respectively). In the subsequent QM/MM and MM calculations, the reaction coordinate was frozen by fixing the two distances at their related values. Several structures on the reaction trajectory were extracted for which partial atomic charges were calculated by using the natural population analysis (NPA) of Reed et al.[34] These charges are directly related to the electronic density and were found to faithfully follow the progress of the chemical reaction. Electrostatic potential fitting (ESP) charges calculated using the Merz−Singh−Kollman scheme[35] were also tested, but they were

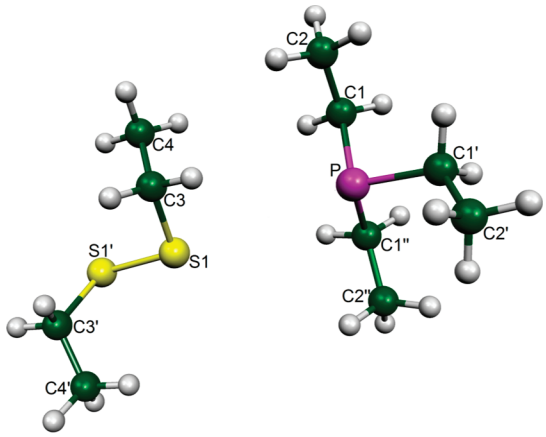


Figure 1. Model system for the disulfide bridge reduction by TCEP used in QM-PCM and QM/MM calculations.

found to be not stable enough in the present case (see Supporting Information). On the other hand, the NPA and ESP dipole moments calculated for the model system do not differ significantly (see Supporting Information). Thus, in the present case, using NPA charges for the model system and ESP charges for the MM system in the QM/MM calculations is not inconsistent.

The structures of the model system extracted from the QM-PCM reaction trajectory were transposed in the protein environment. The starting geometry of lysozyme was the X-ray diffraction structure of hen egg white lysozyme (HEWL).[36] In order to make possible a comparison with the experimental data on the reaction kinetics already reported, the protonation state of amino acid residues was that one stable at pH 3 while for TCEP was chosen the neutral form (Scheme 1). The MM calculations were performed with the Amber 9 simulation package.[37] The partial atomic charges of TCEP and the eight cysteine residues were calculated by the NPA method while for the rest of the system the charges of the Amber force field were used. The MD simulations were performed at constant temperature (300 K) and pressure (1 atm) using a 50 × 60 × 60 TIP3P[38] solvent box with periodic boundary conditions. Cl^- counterions were added in order to preserve the neutrality of the solution. All covalent bonds involving H atoms were constrained by the SHAKE algorithm while the time step for the integration of the equations of motion was 2 fs. The cutoff of the no bonded interactions was set at 13.5 Å. The trajectory production was preceded by a heating and an equilibration period of 20 ps each.

The QM/MM optimizations with frozen reaction coordinate (more specifically, with frozen d(S−P) and d(S−S)) were performed with Gaussian 03 program based on the ONIOM algorithm.[26,39] The structures were first optimized with mechanical embedding, then reoptimized with electronic embedding. The electronic embedding calculations were performed by scaling the electrostatic interactions with a factor of 0.6 in order to avoid over polarization of the model system. The present scaling simulates the partial disorientation of the solvent molecules due to their dynamics. The model system was that shown in Figure 1. The geometry of the real (extended) system was obtained from that used in the MM simulations by deleting the solvent molecules whose distance with respect to the closest protein or TCEP atom was greater than 10 Å. The periodic boundary conditions were not maintained in this kind of calculations. The real system was treated using the Amber force field as implemented in Gaussian 03 while the QM system was

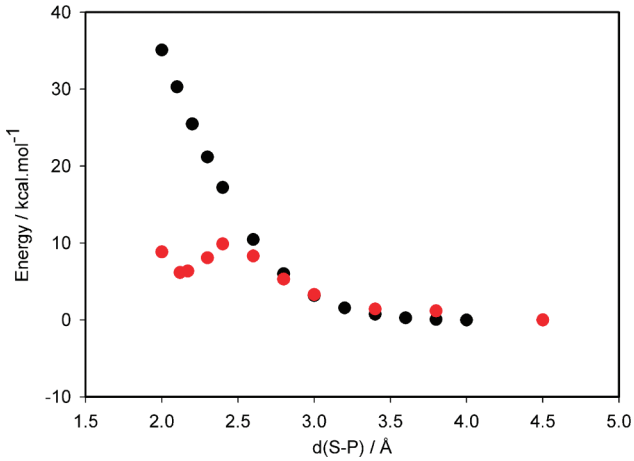


Figure 2. QM calculations on the model system: potential energy curve calculated in vacuo (black) and PCM free energy curve calculated in aqueous solution (red).

treated at the B3LYP/6-311+G(d,p) level of theory. The optimized structures were further used to calculate the quantum energy $E_{M;QM}(X_{M;min}(\xi_i))$ in eq 10, according to the formula

$$E_{M;QM}(X_{M;min}(\xi_i)) = \langle \Psi | \hat{H} | \Psi \rangle - \sum_{\alpha \in M, \beta \in R\text{-}M} \frac{q_\alpha q_\alpha}{r_{\alpha\beta}} \quad (19)$$

where $\langle \Psi | \hat{H} | \Psi \rangle$ is the QM energy in the external point charge field (electronic embedding) and q_α and q_β are the partial atomic charges of the model system and, respectively, of its environment.

The free energy variation $\Delta \bar{F}_{MM}(i \rightarrow i+1)$ was determined by the TDI method. Each configuration i on the reaction trajectory was defined by fixing the partial atomic charges of the model system at $q_M(\xi_i)$ and by using harmonic restraints on d(S−P) and d(S−S) at $d(S-P)_i$ and $d(S-S)_i$, respectively. Practically, the restrains were imposed by creating a S−S bond and a S−P bond for which the lengths were adjusted for each configuration i. The force constant for these bonds was set at 250 kcal $Å^{-2}$. The average in the eq 18 was based on simulation trajectories of 0.5 ns each generated with the hybrid energy given by eq 16.

The remaining term in eq 10, the mean energy $\langle E_{M;MM} \rangle_{\xi_i}$, was evaluated by performing similar molecular dynamic simulations with the mixing parameter λ set at 0.

Results

QM Calculations on the Model System. Constrained optimizations performed for different S−P distances at the B3LYP/6-311+G(d,p) level of theory on the model system in Figure 1 indicate that in vacuo the S−S bridge is stable in the presence of TCEP. Indeed, the potential energy of the system increases continuously upon decreasing the S−P distance (Figure 2, black points). The S−S bond length increases from 2.1 Å (unperturbed S−S bridge) to 2.6 Å (d(S−P) = 2.0 Å), but it is not broken even for shorter distances between the two reactants.

A different evolution is predicted in aqueous solution: the QM-PCM free energy of the system presents a maximum at d(S−P) = 2.4 Å and a local minimum at d(S−P) = 2.12 Å (Figure 2, red points). The free energy profile in Figure 2 does not include the thermal and entropy contributions of the solute. The system geometry obtained by constrained optimization at

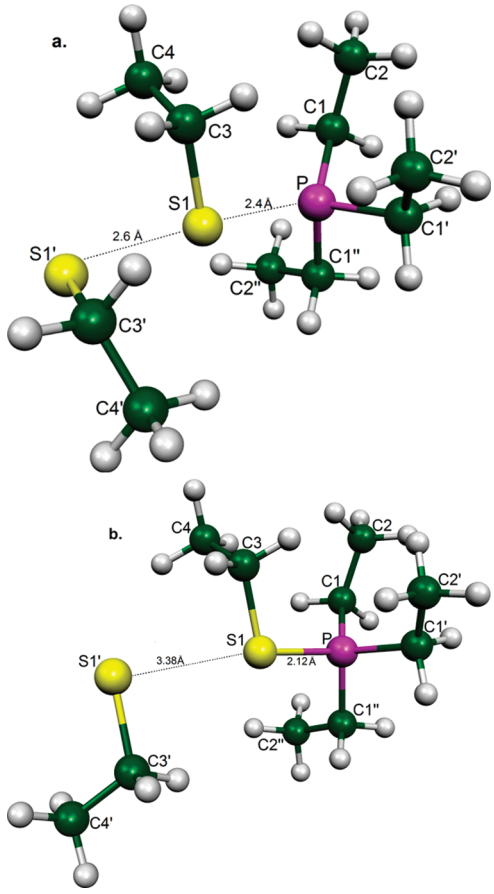


Figure 3. QM-PCM calculation on the model system in aqueous solution: (a) geometry of the transition state TS and (b) geometry of the complex of intermediate products IP.

d(S−P) = 2.4 Å is very close to that of a transition state (TS): this was confirmed by performing an unconstrained TS reoptimization starting from this point. The structure thus obtained (Figure 3a) possesses an imaginary frequency and the corresponding normal mode is a combination of d(S−P) and d(S−S). The reaction products are the S−P bond intermediate in Scheme 1 and the thiolate. Their complex (IP) is presented in Figure 3b. The predicted intermediate products are not very stable with respect to the free reactants since the energy barrier for the backward reaction is only 4 kcal mol^{-1}.

The present result is in agreement with that reported by Dmitrenko et al.[4] which performed QM-COSMO optimizations on a similar model system, but of smaller size. We note, however, that the method they used allowed them to identify a transition state only in the presence of two explicit water molecules interacting with their model system. This confirm the important role that the environment plays in this reaction.

Structures extracted from the reaction trajectory in Figure 2 (red points) were subsequently used in the QM-FE calculations.

QM-FE Calculations on the Reduction of GSSG by TCEP. QM-FE modeling of GSSG reduction by TCEP provides a very useful term of comparison for the reduction of S−S bridges in proteins. It also gives occasion to a comparison between the QM-PCM and QM-FE methods since the S−S bridge environment in GSSG is not very different from that in the model system. The QM-FE free energy profile calculated for the GSSG reduction is presented in Figure 4. The interaction between the two reactants was followed up to a S−P distance of 7.0 Å where the free energy variation with respect to the reaction coordinate becomes negligible. Consequently, the origin of free energy in

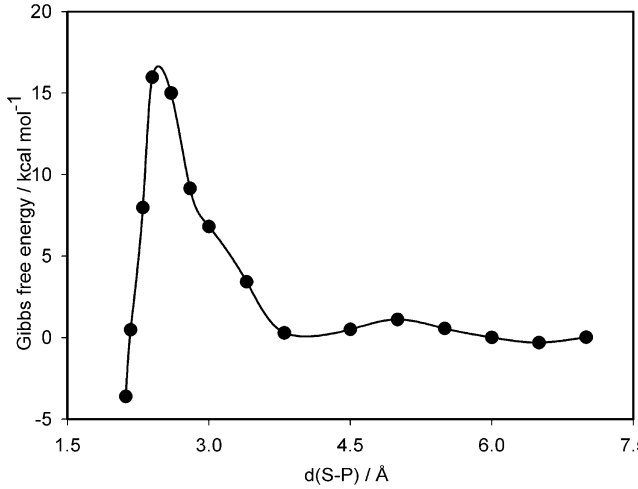


Figure 4. Gibbs free energy profile calculated for the reduction of GSSG by TCEP using the QM-FE method.

Figure 4 was chosen at d(S−P) = 7.0 Å. The predicted reaction energy barrier is about 5 kcal mol^{-1} higher as compared to that calculated for the model system using the QM-PCM method. However, the most significant difference between the two results is observed for the stability of the IP complex: in the QM-FE calculation IP is more stable with respect to the reactants by about 3.6 kcal mol^{-1}. This result is in agreement with the experimental evidence that the rate limiting step in the S−S bridge reduction by TCEP is the first reaction step,[3] thus implying that the intermediate products are relatively stable.

As expected, the difference between the products stabilities predicted by the two computational methods is due to a different evaluation of the interaction energy between the model system and its environment. In the QM-FE evaluation, the products−environment interaction energy is lower by 62 kcal mol^{-1} with respect to that calculated for the reactants while this difference is only 42 kcal mol^{-1} in the QM-PCM approach. On the other hand, the difference between the two calculated energy barriers has a different origin. Indeed, for a given ξ, the QM/MM reoptimized structures of the model system in GSSG were slightly different with respect to the starting QM-PCM optimized structures. The corresponding quantum energies ($E_{M;QM}$) were also slightly different. The main contribution to the difference between the two energy barriers comes from this quantum energy. Obviously, the changes in the model system geometry, and hence in the quantum energy, are also due to the environment interaction as described by the QM/MM calculation method.

It will be shown below that the QM-FE energy barrier calculated for the reduction of GSSG by TCEP is in good agreement with the experimental value. It appears that, even in the case of GSSG, the QM-FE method takes advantage over the QM-PCM method of a more realistic treatment of the reaction environment.

QM-FE Calculations on the Reduction of the Four S−S Bridges in Lysozyme. Gibbs free energy profiles calculated for the reduction of S−S bridges in protein environment (Figure 5) show significant differences with respect to that obtained for GSSG. The reaction energy barrier varies between 12.8 and 24.5 kcal mol^{-1} while the minimum value corresponds to the cys6−cys127 bridge. According to the solvent accessibility evaluations performed with the NACCESS[40] program, this bridge is also the only one that is solvent exposed. In fact, the shape of the curves in Figure 5 gives a direct indication about

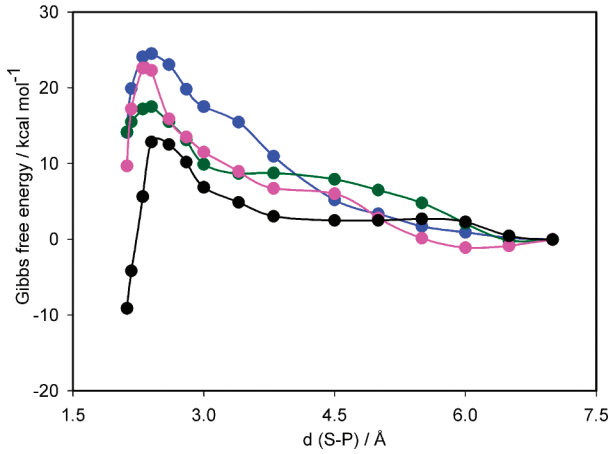


Figure 5. Gibbs free energy profiles calculated for the reduction of the four S−S bridges in lysozyme by TCEP using the QM-FE method: cys6−cys127 (black), cys30−cys115 (blue), cys64−cys80 (green), and cys76−cys94 (pink).

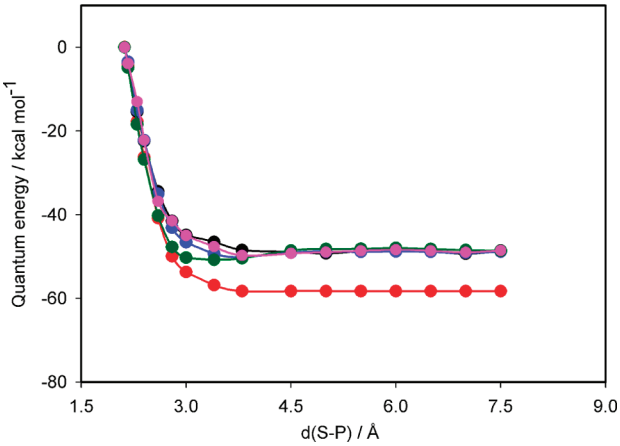


Figure 6. Quantum energy (E_{QM}) contribution to the Gibbs free energy profile calculated for cys6−cys127 (black), cys30−cys115 (blue), cys64−cys80 (green), cys76−cys94 (pink), and GSSG (red).

the accessibility of S−S bridges by TCEP: the free energy rising upon the decrease of d(S−P) from 7.0 to 4.0 Å is clearly related to a steric hindrance. This rising is present for all the bridges except the bridge cys6−cys127, a result that is consistent with the static accessibilities provided by NACCESS.

A more explicit description of the protein environment effects is obtained when the two contributions to the Gibbs free energy in eq 6 are separately analyzed. The internal quantum energy $E_{M;QM}$ calculated for GSSG and the four S−S bridges of lysozyme is represented in Figure 6 as a function of d(S−P). For each curve, the $E_{M;QM}$ energy of the IP complex was chosen as reference. As expected from the in vacuo calculations on the model system, the internal quantum energy of the separated reactants is significantly lower than that of the IP complex. The energy difference between the two states (separated reactants and IP, respectively) is more important for GSSG, −58 kcal mol^{-1}, while for the four S−S bridges in lysozyme it is only −48 kcal mol^{-1}. We attribute this dissimilarity between GSSG and protein to a destabilization of the S−S bridges in protein induced by structural constraints. The bridge cleavage partially relaxes this distortion that still affects the products, as indicated by the absolute values of the internal quantum energy of the IP complexes (not shown). The S−S bridge distortion in protein has a nonnegligible effect on the energy barrier of reaction also. A comparison between the $E_{M;QM}$ contributions to this barrier

TABLE 1: Internal Quantum Mechanics and Environment Interaction Contributions to the Reaction Energy Barriers[a]

disulfide bridge	ΔE_{QM}/kcal mol^{-1}	ΔF_{env}/kcal mol^{-1}
GSSG	32.0	–16.0
cys6–cys127	26.1	–13.3
cys30–cys115	26.6	–2.1
cys64–cys80	21.7	–4.2
cys76–cys94	26.4	–4.1

[a] Differences ΔE_{QM} and ΔF_{env} were calculated between d(S−P) = 2.4 Å and d(S−P) = 7.0 Å, respectively.

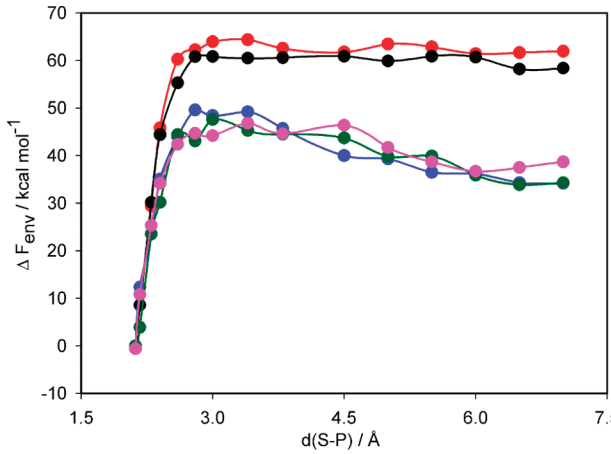


Figure 7. Contribution of the environment interaction (F_{env}) to the Gibbs free energy profile calculated for cys6−cys127 (black), cys30−cys115 (blue), cys64−cys80 (green), cys76−cys94 (pink), and GSSG (red).

in GSSG and protein is given in Table 1. The four S−S bridges in lysozyme have $E_{M;QM}$ contributions to the reaction energy barrier smaller by 5.4−10.3 kcal mol^{-1} with respect to GSSG. A remarkably small contribution is observed for the cys64−cys80 bridge, indicating that in this case the bridge distortion relaxes sooner along the reaction trajectory.

The second contribution to the Gibbs free energy to be analyzed is the environment contribution F_{env}. The related curves calculated for GSSG and the four S−S bridges of lysozyme are presented in Figure 7. For each curve, the free energy origin corresponds to the IP complex. It is not surprising that the solvent exposed bridge cys6−cys127 and GSSG have very similar environment contributions to the free energy profile. An interesting similarity is also observed between the three remaining lysozyme S−S bridges. Going further into details, one notices that the main factors characterizing the reaction environment are the polarizability and the S−S bridge accessibility. Their effects along the reaction trajectory are difficult to separate. However one can attribute the step decrease of ΔF_{env} upon the transition TS → IP to the solvation of the reactant system showing an increasing polarity during this transition. This solvation is more effective by about 15 kcal mol^{-1} for the solvent exposed bridges GSSG and cys6−cys127 with respect to the three other bridges. On the other hand, the ΔF_{env} rising observed for these later bridges upon the decrease of d(S−P) from 7.0 to 4.0 Å can be attributed to the steric hindrance. Indeed, the weak and almost constant reactant polarity in this d(S−P) range could not explain a ΔF_{env} variation of about 10−15 kcal mol^{-1}.

Another important point to be considered is the ΔF_{env} contribution to the reaction energy barrier. This contribution is generally negative thus favorable to the reaction (Table 1). The contribution is very important for GSSG and cys6−cys127, but rather modest for the other bridges.

Discussion

The viability of the present theoretical results may be directly checked since experimental data on GSSG and lysozyme reduction by TCEP are now available. A rate constant of 15.2 M^{-1} s^{-1} was recently reported for the GSSG reduction while the experimental kinetics of the reduction of S−S bridges in lysozyme were found to be heterogeneous and were fitted using two rate constants, 0.45 and 0.052 M^{-1} s^{-1}, respectively.[10] Based on these values, one can evaluate the corresponding reaction energy barriers by using the Eyring's formula of the reaction rate constant in the transition state theory (TST) approximation.[41] This gives an energy barrier of 16.5 kcal mol^{-1} for the reduction of GSSG and two energy barriers, 18.0 and 19.5 kcal mol^{-1}, respectively, for the lysozyme reduction. At first sight, the theoretical energy barriers reported in the present paper might not be directly compared with these experimentally derived values. Indeed, the theoretical barriers were obtained from the variation of the partial Gibbs free energy. The real energy barrier of reaction includes also the contribution of the translation and rotational motions to the dissociation Gibbs free energy of the complex of reactants. In fact, it was already shown that in aqueous solutions at standard concentration (1 mol dm^{-3}) the median value of this contribution is zero with a standard deviation of $5RT$.[42] Hence the theoretical energy barriers derived from free energy profiles may be assimilated within a good approximation to the real reaction energy barriers. Under these considerations, one can state that the energy barrier calculated for the reduction of GSSG, that is, 16.0 kcal mol^{-1}, is in good agreement with the experimentally derived value.

In the case of lysozyme S−S bridges, the agreement between calculated and experimental energy barriers is variable. The predicted energy barrier for the bridge cys6−cys127, 12.8 kcal mol^{-1}, is lower by about 5.2 kcal mol^{-1} with respect to the lowest experimental energy barrier for lysozyme reduction. That calculated for the bridge cys30−cys115, 24.5 kcal mol^{-1}, exceeds the highest experimental value by about 5 kcal mol^{-1}. The energy barriers calculated for the two other S−S bridges, cys64−cys80 and cys76−cys94, are 17.5 and 22.3 kcal mol^{-1}, respectively. This is close enough to the interval of the experimental values.

One may prudently conclude that the present computation method when applied to the S−S bridge reduction gives at least a good qualitative description, even in the case of a complex reaction environment such as the protein structure. Hence a re-examination of the reaction model in the light of these new results becomes possible. Currently, it is admitted that the reduction of S−S bridges in protein occurs in two steps: in the first step, the access to the bridge is opened by a local protein unfolding, and in the second, the exposed bridge is reduced by the reducing agent.[7] In this model, the Gibbs free energy of local unfolding is calculated as the difference between the energy barrier of reduction of a S−S bridge in protein environment and that of a fully solvent exposed bridge in a model system such as GSSG.[8,43] According to the theoretical analysis presented above, the reaction energy barrier in protein is dependent not only on the S−S bridge accessibility but also on the environment polarizability as well as on the S−S bridge distortion. Interestingly, even when the bridge is in contact with the TCEP after the opening of the reaction channel, the environment polarizability of an initially buried S−S bridge is not that of a solvent exposed S−S bridge. This can be seen by comparing the free energy variations between the states TS and IP for different reactant systems, as shown in Figure 7. The present result indicates that the local unfolding has a limited amplitude and

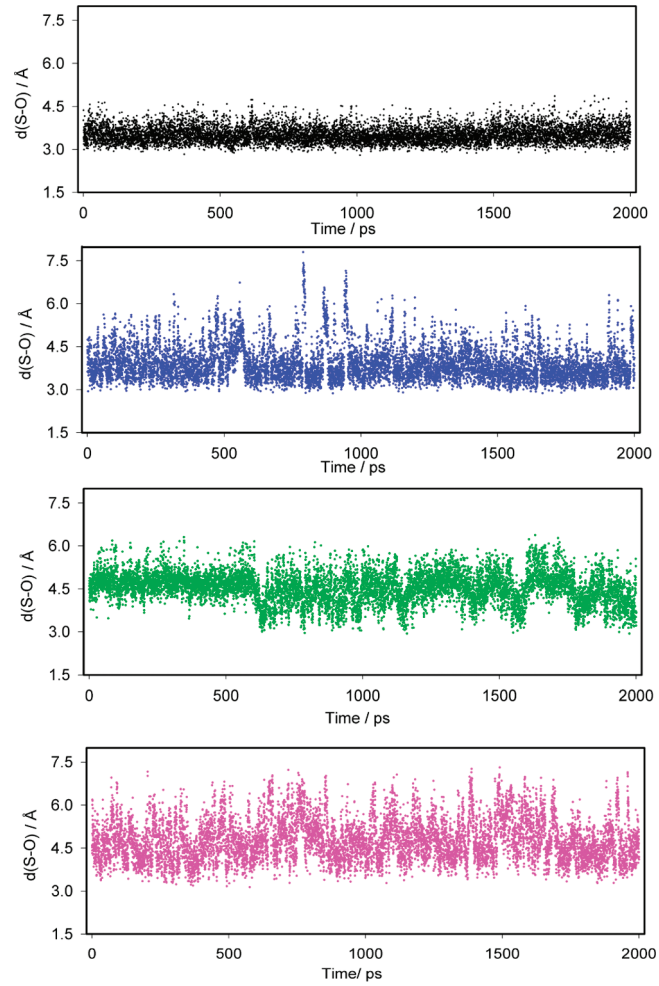


Figure 8. Distribution of distances between the most exposed S atom of a bridge and the closest water molecule calculated along a dynamic simulation trajectory for cys6−cys127 (black), cys30−cys115 (blue), cys64−cys80 (green), and cys76−cys94 (pink).

the environment of a buried bridge does not become equivalent to that of a naturally exposed bridge. Hence, the contribution of the local unfolding to the reaction energy barrier could not be obtained by simply subtracting the energy barrier of reduction of a fully exposed S−S bridge. Alternatively, the curves in Figure 7 suggest that, in the case of the three buried or partially buried S−S bridges in lysozyme, the steric hindrance contribution to this energy barrier is situated between 10 and 15 kcal mol^{-1}. However, one cannot deduce from a static reaction parameter that is the free energy profiles whether this hindrance is over taken by local unfolding or by diffusion.

We further investigated the dynamics of the S−S bridges environment in lysozyme by monitoring the distance between the most exposed S atom of a bridge and the closest water molecule, along a dynamic simulation trajectory of 2 ns (Figure 8). The distribution of distances obtained for the bridge cys6−cys127 is by far the least dispersed and it is centered on about 3.4 Å. This distribution can be considered as a reference for a solvent-exposed S−S bridge. For the bridges cys30−cys115 and cys64−cys80, one can identify short segments of trajectory of several tens of picoseconds where the S−O distances are confined in the interval 3.0−4.5 Å, thus closely resembling to the distance distribution of a solvent-exposed bridge. On the other hand, in the case of the bridge cys76−cys94, the S−O distances are always superior to 3.3 Å and it is hard to identify similar segments of trajectory. These results appear to be

consistent with the hypothesis of local unfolding at least in the case of the bridges cys30−cys115 and cys64−cys80. More analyses are necessary in order to clarify the dynamic aspect of the reaction.

In conclusion, the present work provides reliable qualitative and quantitative results contributing to a better understanding of the mechanism of the S−S bridges reduction in protein environment. The proposed reaction pattern may be enhanced by extending the present analysis to other S−S bridge containing proteins.

Supporting Information Available: Evaluation of the harmonic restraints approximation used in calculating the partial free energy variation; Cartesian coordinates for the transition state TS and the intermediate products complex IP as well as the corresponding Gaussian 03 input files used in geometry optimization; ESP and NPA partial atomic charges for the main atoms of the model system; comparison between the dipole moments of model system calculated with the NPA and the ESP charges; example of AMBER input for molecular dynamics simulation used in TDI calculations; examples of Gaussian input file for QM/MM optimizations with mechanical and electronic embedding. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Vaz, D. C.; Rodriguez, J. R.; Sebald, W.; Dobson, C. M.; Brito, M. M. *Protein Sci.* **2006**, *15*, 33.

(2) Wendemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. *Biochemistry* **2000**, *39*, 4207.

(3) Cline, D. J.; Redding, S. E.; Brohawn, S. G.; Psathas, J. N.; Schneider, J. P.; Thorpe, C. *Biochemistry* **2004**, *43*, 15195.

(4) Dmitrenko, O.; Thorpe, C.; Bach, R. D. *J. Org. Chem.* **2007**, *72*, 8298.

(5) Goldberg, D. P.; Bekeart, L. S.; Lahern, D. A.; Zhou, J. D. *Biochemistry* **1993**, *32*, 2835.

(6) Wilkinson, B.; Xiao, R.; Gilbert, H. F. *J. Biol. Chem.* **2005**, *280*, 11483.

(7) Feng, Z.; Butler, M. C.; Alom, S. L.; Loh, S. H. *J. Mol. Biol.* **2001**, *314*, 153.

(8) David, C.; Foley, S.; Mavon, C.; Enescu, M. *Biopolymers* **2008**, *89*, 623.

(9) Wilkinson, B.; Xiao, R.; Gilbert, H. F. *J. Biol. Chem.* **2005**, *280*, 11483.

(10) David, C.; Foley, S.; Enescu, M. *Phys. Chem. Chem. Phys.* **2009**, *14*, 2532.

(11) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

(12) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185.

(13) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296.

(14) Cummins, P. L.; Rostov, I. V.; Gready, J. E. *J. Chem. Theory Comput.* **2007**, *3*, 1203.

(15) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J. Chem. Theory Comput.* **2006**, *2*, 452.

(16) Li, G.; Zhang, X.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 8643.

(17) Strajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333.

(18) Rosta, E.; Klähn, M.; Warshel, A. *J. Phys. Chem. B* **2006**, *110*, 2934.

(19) Sanchez, M. L.; Martin, M. E.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Comput. Chem.* **2000**, *21*, 705.

(20) Rosta, E.; Haranczyk, M.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **2008**, *112*, 5680.

(21) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154.

(22) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *184*, 2974.

(23) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483.

(24) Kaukonen, M.; Söderhjelm, P.; Heimdal, J.; Ryde, U. *J. Chem. Theory Comput.* **2008**, *4*, 985.

(25) Hu, H.; Lu, Z.; Yang, W. *J. Chem. Theory Comput.* **2007**, *3*, 390.

(26) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125.

(27) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43.

(28) Kollman, P. A.; Kuhn, B.; Donini, O.; Perakyla, M.; Stanton, R.; Bakowies, D. *Acc. Chem. Res.* **2001**, *34*, 72.

(29) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.

(30) Mülders, T.; Krüger, P.; Swegat, W.; Schlitter, J. *J. Chem. Phys.* **1995**, *104*, 4869.

(31) Senn, H. M.; Thiel, S.; Thiel, W. *J. Chem. Theory Comput.* **2005**, *1*, 494.

(32) den Otter, W. K. *J. Chem. Phys.* **2000**, 7283.

(33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; 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. Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.04*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(34) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.

(35) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(36) Jakoncic, J.; Aslantas, M.; Honkimaki, V.; Dimichel, M.; Stojanoff, V. *J. Appl. Crystallogr.* **2006**, *39*, 831.

(37) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, 2006.

(38) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. P. *J. Chem. Phys.* **1983**, *79*, 926.

(39) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654.

(40) Hubbard, S. J.; Thornton, J. M. *NACCESS, computer program*; Department Biochemistry and Molecular Biology, University College, London, 1993.

(41) Cardey, B.; Enescu, M. *ChemPhysChem* **2005**, *6*, 1175.

(42) Yu, Y. B.; Privalov, P. L.; Hodges, R. S. *Biophys. J.* **2001**, *81*, 1632.

(43) Li, Y.-J.; Rothwarf, D. M.; Scheraga, H. A. *Nat. Struct. Biol.* **1995**, *2*, 489.

JP910340T