Role of the Variable Active Site Residues in the Function of Thioredoxin Family Oxidoreductases

ALEXANDRA T. P. CARVALHO,1 PEDRO A. FERNANDES,1 MARCEL SWART,2* JOOST N. P. VAN STRALEN,2 F. MATTHIAS BICKELHAUPT,2 MARIA J. RAMOS1

1REQUIMTE, Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre 687, P-4169-007 Porto, Portugal
2Department of Theoretical Chemistry and Amsterdam Center of Multiscale Modeling, Scheikundig Laboratorium der Vrije Universiteit, De Boelelaan 1083, NL-1081 HV Amsterdam, The Netherlands

Received 23 May 2008; Accepted 1 July 2008
DOI 10.1002/jcc.21086
Published online 8 September 2008 in Wiley InterScience (www.interscience.wiley.com).

Abstract: The enzymes of the thioredoxin family fulfill a wide range of physiological functions. Although they possess a similar CXYC active site motif, with identical environment and stereochemical properties, the redox potential and \( \mathrm{pK}_a \) of the cysteine pair varies widely across the family. As a consequence, each family member promotes oxidation or reduction reactions, or even isomerization reactions. The analysis of the three-dimensional structures gives no clues to identify the molecular source for the different active site properties. Therefore, we carried out a set of quantum mechanical calculations in active site models to gain more understanding on the elusive molecular-level origin of the differentiation of the properties across the family. The obtained results, together with earlier quantum mechanical calculations performed in our laboratories, gave rise to a consistent line of evidence, which points to the fact that both active site cysteines play an important role in the differentiation. In contrary to what was assumed, differentiation is not achieved through a different stabilization of the solvent exposed cysteine but, instead, through a fine tuning of the nucleophilicity of both active site cysteines. Reductant enzymes have both cysteine thiolates poorly stabilized, oxidant proteins have both cysteine thiolates highly stabilized, and isomerases have one thiolate (solvent exposed) poorly stabilized and the other (buried) thiolate highly stabilized. The feasibility of shifting the chemical equilibrium toward oxidation, reduction, or isomerization only through subtle electrostatic effects is quite unusual, and it relies on the inherent thermoneutrality of the catalytic steps carried out by a set of chemically equivalent entities all of which are cysteine thiolates. Such pattern of stabilization/destabilization, detected in our calculations is fully consistent with the observed physiological roles of this family of enzymes.


Key words: density functional calculations; disulfide; enzyme catalysis; molecular dynamics; reaction mechanisms

Introduction

The enzymes of the Thioredoxin family possess a high degree of structural similarity with at least one domain having the characteristic thioredoxin fold. However, they exhibit very different reducing potentials and hence fulfill very different functions, ranging from reductants [such as thioredoxins (Trx) and glutaredoxins (Grx)] to oxidants, such as disulfide bond formation A (DsbA), and even isomerases such as protein disulfide isomerase (PDI).1–5

All of these enzymes have a conserved stretch of sequence in the active site that contains a CXYC motif, where C stands for cysteine and XY are two variable residues. The two variable inner residues correspond to a glycine and a proline in Trx, to a proline and a tyrosine in Grx, to a glycine and a histidine in PDI, and to a proline and a histidine in DsbA (Table 1).

The diversity of functions of the family members is achieved by different, even though related, catalytic mechanisms. Scheme 1 gives an overview of the proposals for the most common

Additional Supporting Information may be found in the online version of this article.

*Present address: Institució Catalana de Recerca i Estudis Avançats (ICREA), E-08010 Barcelona, Spain, and Institut de Química Computacional, Universitat de Girona, Campus Montilivi, E-17071 Girona, Spain.

Correspondence to: F. M. Bickelhaupt or M. J. Ramos; e-mail: fm.bickelhaupt@few.vu.nl or mjramos@fc.up.pt

Contract/grant sponsor: Fundação para a Ciência e Tecnologia (FCT, Portugal); contract/grant number: SFRH/BD/17845/2004

Contract/grant sponsor: Netherlands Organization for Scientific Research (NWO-CW, NWO-NCF)

© 2008 Wiley Periodicals, Inc.
mechanisms: reduction by Trx, oxidation by DsbA, and isomerization by PDI.

The side chain of the solvent-exposed nucleophilic cysteine (C_{nuc}) must be deprotonated in all the mechanisms in which these proteins participate as reductants or isomerases. In such situations, the side chain thiolate (S_{bur}) performs a nucleophilic attack on the disulfide bond of the substrate (S_{S,S}) leading to the formation of the so-called mixed disulfide intermediate in which the enzyme and substrate are covalently bound via a new disulfide bridge. Trx represents the paradigm of the reductases. In these enzymes the buried cysteine thiol group of the originally reduced Trx is deprotonated. This initiates a second S_{S,S} step in which the C_{nuc} thiolate attacks the C_{nuc} sulfur atom involved in the disulfide bridge, causing the rupture of the latter and the release of the products, that is an oxidized enzyme and a reduced dithiol substrate (Scheme 1).

In the process of isomerization, the thiolate of the substrate that was released in the preceding step reacts intramolecularly with another substrate disulfide bond and initiates a cascade of intramolecular disulfide rearrangements. The conclusion of the substrate rearrangement happens when a substrate cysteine displaces the enzyme from the covalent complex forming the same substrate disulfide and reoxidizing them with different cysteine pairs (Scheme 1).

If the enzyme is an oxidant, the very unstable and reactive disulfide bond of the protein is attacked by a thiolate from a substrate (S_{S,S}), generating the mixed disulfide intermediate, and evolves later on to a reduced protein via nucleophilic attack of the same substrate on the mixed disulfide bond (S_{S,S}) (Scheme 1).

Glutaredoxins operate by a slightly different mechanism, because of the participation of glutathione, which resolves the mixed disulfide or reduces the enzyme disulfide. In either case, a second glutathione attacks the enzyme-glutathione mixed disulfide to yield diglutathione and regenerate the reduced enzyme.

The way the S_{bur} proton is lost is still not fully understood. It has been suggested that specific acidic groups in the vicinity of this cysteine (e.g., D26 in thioredoxin) are responsible for the deprotonation. The main problem with this hypothesis is that, in the case of D26, the carboxylate of this residue is more than 6 Å apart from S_{bur}. The results of a earlier mechanistic study on these enzymes led us to the alternative hypothesis of deprotonation by the anionic thiolate released from the substrate, after the attack on its disulfide in the first step of the reduction mechanism.

In earlier works we and others have studied the oxidation/reduction of disulfides. In our studies we have used two kinds of model compounds. The simpler model corresponded to a set of methyliothiol molecules and was intended to constitute a reference system for the intrinsic reactivity of the thiolates in thiol/disulfide exchange. The second system corresponded to a CGPC tetrapeptide model (the Trx active site sequence) and was aimed at illustrating the influence of the inner GP residues on the reactivity of the active site cysteines. We have noticed in that work that the GP motif influences the reactivity of C_{nuc} beyond what was a priori expectable. The results have shown that the S_{S,S} substitution reactions in vacuum proceeded via a triple-well potential energy surface (PES) with very symmetric SS interactions at characteristic stationary states in the methylthiol model, but very asymmetric SS interactions on the tetrapeptide model. The PES changed dramatically, on inclusion of a dielectric continuum solvent model (ε = 78.4) to a unimodal PES with a central barrier provided by a trisulfide transition state. Solvation managed to recover much of the symmetry that was lost in the tetrapeptide in vacuum.

Such observations led us to further investigate the influence of the inner XY residues in the reactivity profile of the C_{nuc}-XYC_{nuc} motif. These two residues are known to have a small influence on the pK_a of the family members, but their particular contribution to modulate the PES for the oxidation/reduction has never been possible to measure before. The obtained results allowed for a better understanding of the role of the inner residues of the active site motif in the reactivity of the thioredoxin family members. Taken together, the results gave us, for the first time, a glimpse into the complex and delicate machinery of this common Trx fold, which can regulate independently the pK_a and the nucleophilicity of the active sites, and promote such range of different reactions within an almost identical scaffold.

### Methods

All density functional calculations were carried out with the Amsterdam density functional (ADF) and the quantum regions interconnected by local descriptions (QUILD) programs. The energy of the stationary points was computed with the BP86 and OPBE functionals in combination with the TZ2P basis set. The latter is a large, uncontracted, triple-ζ set of Slater-type orbitals (STOs), containing two sets of polarization functions, namely 2p and 3d on H, and 3d and 4f on C, N, O, and S. The core shells of carbon (1s), nitrogen (1s), oxygen (1s), and sulfur (1s2s2p) were treated by the frozen-core approximation. An auxiliary set of s, p, d, f, and g STOs (fit functions) was used to fit the molecular density and to represent the Coulomb and exchange potentials accurately in each SCF cycle.

The geometry of the stationary points was optimized, using analytical gradient techniques, at the BP86/TZ2P level except for transition states, which were optimized at the BP86/DZP level. The DZP basis is an uncontracted set of Slater-type orbitals (STOs) of double-ζ quality, and has been augmented with...
one set of polarization functions, namely 2p on H and 3d on C, N, O, and S.

The choice of the DFT functionals was based on a recent study on the performance of DFT functionals for describing various $S_N2$ reactions. In this study, the OPBE functional was found to yield reaction profiles, in particular activation energies, in good agreement with high-level ab initio CCSD(T) benchmarks. The BP86 functional, on the other hand, is known to

Scheme 1. Catalytic mechanism for disulfide reduction and isomerization. The oxidation mechanism is similar to the reduction, but in the opposite direction.
perform better for describing hydrogen-bond distances and energies.\textsuperscript{35}

Vibrational analysis has confirmed the nature of the transition states through the number of imaginary frequencies. The character of the normal mode associated with the imaginary frequency of a transition state (i.e., the transition vector) was analyzed to ensure that the correct transition state was found.

Solvent effects have been estimated using the conductor-like screening model (COSMO),\textsuperscript{36,37} as implemented in the ADF program.\textsuperscript{38} We used a solvent-excluding surface with an effective radius for water of 1.9 Å, derived from the macroscopic density and molecular mass, and a relative dielectric constant of 78.4. The radii of the atoms were taken as MM3 radii,\textsuperscript{39} divided by 1.2, giving 1.350 Å for H, 1.700 Å for C, 1.608 Å for N, 1.517 Å for O, and 1.792 Å for S. These settings for COSMO comprise an ab initio approach to include solvent effects in the QM calculations.\textsuperscript{40}

The multipole derived charges (MDC) analysis\textsuperscript{41} uses the atomic multipoles (obtained from the fitted density) up to some level (quadrupole level in this work) and reconstructs these multipoles exactly (up to the same level) by distributing fractional point charges over all atoms. This is achieved by using Lagrange multipliers and a weight function to keep the multipoles local. Because the atomic multipoles are reconstructed up to the quadrupole level, the molecular multipoles are also represented up to this level.

Results

Once the protein substrate binds noncovalently to the hydrophobic surface of Trx with its oxidized disulfide positioned close to the Trx active site cysteines, the thiol-disulfide exchange reaction readily proceeds via a two step mechanism. In the first step, the thiolate of the C\textsubscript{nu}c attacks the substrate disulfide (S\textsubscript{C}2@S). The product of the reaction consists of a mixed disulfide with the enzyme and the substrate covalently bound (Scheme 1). Meanwhile, the pK\textsubscript{a} of C\textsubscript{bur} drops (partly due to the loss of the stabilizing interaction with the anionic C\textsubscript{nu}c) favoring its subsequent deprotonation. The identity of the base that deprotonates C\textsubscript{nu}c is unknown. In the second S\textsubscript{C}2@S step, the mixed disulfide is resolved by intramolecular attack of the thiolate of C\textsubscript{nu}c on C\textsubscript{nu}c yielding an oxidized enzyme and a reduced substrate (Scheme 1).

Although only the two active site cysteines seem to be directly involved in catalysis, the inner XY residues of the C\textsubscript{nu}c-XYC\textsubscript{bur} motif should play an important role. In fact, other family members with a similar active site (but different XY residues) have a completely different chemical behavior, acting as oxidants or as isomerases. To gain further understanding of the factors that control the active site reactivity, we decided to investigate the reaction mechanism using a C\textsubscript{nu}cGGC\textsubscript{nu}c mutant. Such mutations correspond to the total suppression of the influence of the side chains (but not the backbone dipoles) of the XY motif. These conceptual experiments are hard to perform experimentally because the mutation of two consecutive side chains into glycines generally leads to folding distortions making it impossible to isolate the two effects. However, such experiments can be made resorting to computer simulations, where the enzyme is kept in the original folding state, having no time to thermally progress to lower energy folding states, closer to the new native structure.

Mixed-Disulfide Formation in Vacuum

To study this catalytic cycle, we have used a model substrate constituted by two methyliod molecules oxidized to a disulfide; the thiol which is attacked by C\textsubscript{nu}c was named as central (S\textsubscript{c}) as it becomes the central atom in the hypothetical trisulphide ion, and the other thiol was named as leaving group (S\textsubscript{lg}) as it is the one which is eliminated from the disulfide.

The results have shown that this reaction in vacuum proceeds via a triple-well PES, as observed for the other models.\textsuperscript{13} Two different reactant complexes (IRC\textsubscript{1} and IRC\textsubscript{2}) were located with a feasible geometry in terms of the enzyme stereochemistry. In IRC\textsubscript{1} the thiolate of C\textsubscript{nu}c makes two hydrogen bridges with both methyl groups of the substrate. In IRC\textsubscript{2} (see Fig. 1) the thiolate of C\textsubscript{nu}c (S\textsubscript{nu}c) makes a hydrogen bond with the methyl group of the substrate. The difference in energy between the isolated reactants and reactant complexes (in kcal/mol) corresponds to −4.3 for IRC\textsubscript{1} and to −2.5 for IRC\textsubscript{2} at BP86/TZ2P. The IRC\textsubscript{1} structure is lower in energy than IRC\textsubscript{2}, but it is not directly connected to the transition state that leads toward the products of the catalytic cycle; such kind of minima is not very relevant for the purpose of evaluating the intrinsic reactivity of the CYC motifs in contrary to the ones involved in the reaction coordinate, they will be necessarily different within the framework of the protein structure.

Therefore, the system must progress from IRC\textsubscript{1} to IRC\textsubscript{2}. The transition state for this geometric rearrangement must be very shallow and not related to the chemistry of the enzyme, and therefore was not calculated. From IRC\textsubscript{2}, on, the reaction proceeds with the S\textsubscript{C}2 attack of S\textsubscript{nu}c on S\textsubscript{c} via a low-barrier transition state (ITS\textsubscript{1}, Fig. 1), which is just 1.40 kcal/mol above IRC\textsubscript{2}, to arrive to a transition complex (ITC, Fig. 2) located −1.6 kcal/mol below the reactants (see Table 2). The structure of the ITC intermediate shows two very asymmetric SS interactions, with a S\textsubscript{nu}c−S\textsubscript{c} distance of 3.41 Å (still too long to be considered a chemical bond) and the S\textsubscript{c}−S\textsubscript{lg} bond distance at 2.09 Å. A similar intermediate was also found in earlier calculations with the CGPC system.\textsuperscript{13} Consistently, an uneven charge distribution was observed, the negative charge being mostly concentrated in S\textsubscript{nu}c (−0.48 a.u., with a total charge of −0.96 a.u. in the tetrapeptide model), with the S\textsubscript{c}−S\textsubscript{lg} pair keeping its total charge very similar to the one obtained in a typical disulfide bond (total charge of the substrate of −0.04 a.u., compared with −0.07 a.u. at IRC\textsubscript{2}).

The asymmetry of this complex is clearly a consequence of the pattern of hydrogen bridges that the thiolate of C\textsubscript{nu}c makes with the protein backbone. Besides establishing a hydrogen bond with the thiol of the C\textsubscript{nu}c, the thiolate of C\textsubscript{nu}c engages in hydrogen bonds with the two amide groups of the GG backbone, and these interactions stabilize the charge concentration at S\textsubscript{nu}c, hampering the formation of the trisulphide anion, and lowering the intrinsic nucleophilicity of the thiolate. Such kind of effect was qualitatively expectable. What seems surprising is the extent to
which these thiolate-backbone hydrogen bonds can shift the
gometry of the transition complex toward the reactants side.

From this point, the reaction proceeds through another first-
order saddle point, ITS₂ (see Fig. 2), related to the breaking of
the bond between Sₙuc and S₉₄ and the formation of the Sₙuc—S₉₄
bond. This last bond is still elongated (2.25 Å) in ITS₂ and the
original Sₙuc—S₉₄ is already broken (3.85 Å). Looking at the geo-
metries of ITS₂ and ITC, it becomes clear that a stationary tri-

Figure 1. Selected structures from the first step in gas phase. Distances indicated in Å. Atomic
charges for the sulfur atoms are given in a.u. IRC₃, reactant complex; ITS₁, transition state. The thiol
and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn
in stick representation. The hydrogen bonds involving Sₙuc are also highlighted as black dashed lines.

Figure 2. Selected structures from the first step in gas phase. Distances are indicated in Å. Atomic
charges for the sulfur atoms are given in a.u. ITC, transition complex; ITS₂, transition state. The thiols
and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn
in stick representation. The hydrogen bonds involving Sₙuc are also highlighted as black dashed lines.
Table 2. Relative Energies (in kcal/mol) for the Stationary Points along the Potential Energy Surface for the Whole Catalytic Cycle, in the Gas Phase and in Water.

<table>
<thead>
<tr>
<th>Species</th>
<th>OPBE</th>
<th>BP86</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRC1</td>
<td>−1.09</td>
<td>−1.85</td>
</tr>
<tr>
<td>IRC2</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ITC</td>
<td>4.73</td>
<td>1.40</td>
</tr>
<tr>
<td>1TS</td>
<td>3.11</td>
<td>0.86</td>
</tr>
<tr>
<td>1TS₂</td>
<td>36.56</td>
<td>33.14</td>
</tr>
<tr>
<td>1PC₁</td>
<td>1.25</td>
<td>−3.53</td>
</tr>
<tr>
<td>1PC₂</td>
<td>4.45</td>
<td>4.11</td>
</tr>
<tr>
<td>2RC</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2TS₁</td>
<td>6.88</td>
<td>3.08</td>
</tr>
<tr>
<td>2TC</td>
<td>8.17</td>
<td>2.10</td>
</tr>
<tr>
<td>2TS₂</td>
<td>56.94</td>
<td>50.77</td>
</tr>
<tr>
<td>2PC</td>
<td>17.68</td>
<td>11.98</td>
</tr>
<tr>
<td>3RC₁</td>
<td>−2.62</td>
<td>−0.63</td>
</tr>
<tr>
<td>3RC₂</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3TS</td>
<td>24.52</td>
<td>20.70</td>
</tr>
<tr>
<td>3PC₁</td>
<td>0.95</td>
<td>−1.62</td>
</tr>
<tr>
<td>3PC₂</td>
<td>7.20</td>
<td>4.85</td>
</tr>
<tr>
<td>4RC</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4TS</td>
<td>43.86</td>
<td>31.92</td>
</tr>
<tr>
<td>4PC</td>
<td>15.94</td>
<td>5.88</td>
</tr>
</tbody>
</table>

The second step of the reduction mechanism consists of the nucleophilic attack of the thiolate of C_{bur} to the mixed disulfide bridge. This intramolecular reaction also proceeds via a triple-well PES featuring a pre- and a post-transition state (2TS₁ and 2TS₂) separated by a transition complex (2TC), just 2.10 kcal/mol above the reactants 2RC. Note that the reference for the energy is now 2RC, not 1RC₂ anymore, as at this stage a methythiol molecule has been eliminated from the system. To keep the energy of 2RC as the reference would imply the sum of the electronic energy of methythiol isolated in vacuum. As the molecule does not establish any interaction, the result would be an artificial increase in the energies of all the stationary points for the second reaction step, which would be translated upwards in the energy scale. Therefore, it is more correct to calculate the energy profile independently for each reaction step. The geometry of the first transition state (2TS₁, Fig. 4) is very similar to the first transition state of the first mechanistic step (1TS₁), exhibiting a very long distance between the nucleophile and the acceptor (3.77 Å, with a charge of −0.72 a.u. at S_{bur}), which is still engaged with S_c in a totally formed disulfide 2.10 Å.

As 2TS₁ decays to the 2TC intermediate (see Fig. 5) the attack distance (S_{snuc}−S_{bur}) is shortened to 3.05 Å and the S_{snuc}−S_c distance slightly increases to 2.16 Å. Such a structure, which is very similar to 1TS₁, still holds a very pronounced covalent character at the disulfide bond (S_{bur} charge of −0.66 a.u., compared with −0.73 a.u. at 2RC), with the attacking nucleophile kept in a noncovalent interaction with the slightly elongated disulfide. Note that such asymmetry is again a result of the anisotropic pattern of interactions involving the reacting species (stabilizing the negative charge at S_{bur} and decreasing its nucleophilicity). The equivalent mechanistic step in model compounds, where the environment is more isotropic, exhibits an almost symmetrical trisulphide anion in the corresponding stationary state. The energy of the barrier for this step amounts to 3.08 kcal/mol.

This asymmetrical intermediate progresses toward the final product through the 2TS₂ structure (see Fig. 5) with a barrier of 50.77 kcal/mol. The S_c group is already eliminated from the mixed disulfide at the transition stage (S_c−S_{snuc} distance of 4.20 Å), holding most of the negative charge of the system (−0.84 a.u.), and the final S_{snuc}−S_{bur} interaction is progressing toward a disulfide bond (2.40 Å). The products of the reaction (see Fig. 6) show an oxidized CGGC motif. Figures 5 and 6 illustrate the stationary points associated with the second stage of the oxidation reaction in vacuum.

**Mixed-Disulfide Formation in Water**

The active sites of enzymes are hydrophobic in general. It was shown in the past that in most cases the environment around the active site beyond the first-shell enzyme-substrate interactions could be successfully replaced by a dielectric continuum. The concept behind this approach is that most of the contribution for the activation and reaction energies comes from the reacting groups and from direct hydrogen bonds/interactions, and the remaining of the enzyme has a smaller, modulating role, in the
Figure 3. Products for the first step in gas phase. Distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. \( \text{IPC}_1 \) = product complex, before proton transfer from \( S_{\text{hag}} \) to \( S_{\text{hlg}} \). \( \text{IPC}_2 \) = product complex after proton transfer. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving \( S_{\text{hag}} (\text{IPC}_1) \) and \( S_{\text{hlg}} (\text{IPC}_2) \) are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Figure 4. Selected structures for the second part of the catalytic cycle in gas phase. Distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. \( \text{2RC} \), reactants complex; \( \text{2TS}_2 \), Transition state for the first step of the oxidation of the CGGC disulfide. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving \( S_{\text{hag}} \) are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]
active site chemistry. Even though the success of such an approach should not be taken as granted, there are no cases yet (as far as the authors are aware) in which it has been shown that the use of this modeling technique has changed the course of a catalytic mechanism.

In the particular case of the thioredoxin family, the active sites are partially accessible to the solvent. Therefore, a dielectric constant higher than the one typically used for buried active sites ($\varepsilon = 4$) should be used. Instead of trying to estimate a value for the dielectric constant that would mimic this “hybrid” environment we have opted for repeating the study in the opposite side of the spectrum, that is, in water phase. As we are dealing with active site models, not full enzymes, it is more insightful to follow the reaction mechanism in the two extreme situations (vacuum and water) and to understand the role of the environment, than to concentrate in reproducing the specific dielectric constant of one of the family members, at the cost of losing the global picture about the way the surroundings participate in the catalytic reactions. Therefore, the reactions taking place in the CGGC active site model were also studied within a dielectric continuum solvent (namely the COSMO model).

In aqueous solution, the profile of the PES for the first reaction (the formation of the mixed disulfide) changed from a triple-well PES to an essentially unimodal reaction. Interestingly, the same phenomenon was also seen in other related systems, as in the reactions between CH$_3$S$^-$ and CH$_3$SSCH$_3$, and between CGPC and CH$_3$SSCH$_3$.

![Figure 5](image1.png)  
**Figure 5.** Selected structures for the second part of the catalytic cycle in gas phase. Distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. 2TC, transition complex; 2TS$_2$, Transition state for the second step of the oxidation of the CGGC disulfide. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving S$_{nuc}$ (2TC, 2TS$_2$) and S$_{bur}$ (2TS$_2$) are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

![Figure 6](image2.png)  
**Figure 6.** Final products for the catalytic cycle in gas phase. Distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving S$_{nuc}$ are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]
Similar structures were found in water and in water for the reactant complexes, with the thiolate making one (3RC₂) and two (3RC₃) hydrogen bridges with the dimethylsulfide substrate; 3RC₂ is the relevant minimum that is associated with the reaction coordinate. The nucleophilic attack of S_nuc over the substrate disulfide progresses from 3RC₂ through a much more symmetric transition state (3TS), with a transition vector corresponding to the breaking of the bond between S_lg and S₀, and the formation of the bond between S_nuc and S₀ (see Fig. 7). The S_nuc−S_lg distance now corresponds to 2.47 Å and the S₀−S_lg distance corresponds to 2.58 Å. The saddle point 3TS now clearly corresponds to a trisulfide anion. The charge distribution is also much more symmetric and concentrated in the two terminal atoms of the trisulfide, with values of −0.60 a.u. in S_nuc and −0.77 a.u. in S_lg, consistent with the longer S₃S₃ distance compared with the one of S₃S₃.

The activation energy corresponds to 20.70 kcal/mol (Table 2). It is interesting to realize that the structure with a symmetric S_nuc−S₀−S_lg interaction does not exist as a stationary point in vacuum, but corresponds to a transition state in the presence of solvent. The difference relies on the capacity of the environment to stabilize the nascent thiolate (S_lg) to an extent where it can be compared with the stabilization of the more buried S_nuc. In the gas phase only S_nuc was stabilized, through hydrogen bonds to S_lg and to the backbone, leading to a greater asymmetry in the stationary points. These observations emphasize how much a simple environmental effect can influence the shape of this particular PES.

After the collapse of the transition state, the S_lg is eliminated and ends up involving in a strong ionic hydrogen bond with the thiol of C_bur (2.18 Å), in a perfect position for the necessary abstraction of the S_bur proton. In this way, the thiol eliminated from the substrate disulfide emerges again as the best candidate for proton abstraction to S_nuc (see Fig. 8). The same situation was noted in the gas phase, as well as with a CGPC model in an earlier work.¹³ These evidences, together with the absence of a suitable nucleophile in any of the family members, lead to the conclusion that the substrate is the most probable responsible species for the deprotonation and activation of S_nuc for the subsequent attack on S_nuc and oxidation of the active site. The mixed disulfide is located 1.62 kcal/mol below the reactants, and upon proton transfer from S_bur to S₀ (see corresponding structure in Fig. 8) it becomes located at 4.85 kcal/mol above the reactants (Table 2).

**Product Formation in Aqueous Solution**

This collapse of the mixed disulfide progresses through a unimodal PES in the condensed phase. The structure of the reactants (4RC) was found to be similar to those in the previous steps, with the buried thiolate establishing a long, weak interaction (6.1 Å) with the mixed disulfide, and carrying most of the negative charge of the system (−0.96 a.u.). The reaction proceeds to an almost symmetric transition state structure (4TS), which corresponds again to a trisulfide anion, characterized by an S_nuc−S_nuc distance of 2.66 Å and an S_nuc−S₀ distance of 2.67 Å (see Fig. 9).

These distances are longer and more symmetric than in the previous transition state (3TS). The charge distribution is also more symmetric, with values of −0.62 a.u. for S₀ and −0.60 a.u. for S_nuc. The barrier for this step amounts to 31.92 kcal/mol (Table 2). The product derived from this saddle point exhibits a fully formed intramolecular disulfide bond, corresponding to the oxidized form of the enzyme (see Fig. 10), with the eliminated S₀ carrying now the whole negative charge (−1.10 a.u.). The disulfide bond length corresponds to 2.07 Å. This structure is located 5.88 kcal/mol above the mixed disulfide (Table 2).

All the discussed structures are shown in Figure 6. Figure 11 shows the overall potential energy surface for the reactions in the gas phase and in water.

The picture emphasizes that the reactants, the mixed disulfide intermediate, and the products with the exception of the gas-phase products, all lie in close proximity in the energy scale. If the gas phase, the mixed disulfide is located 4.11 kcal/mol above the reactants and the product lies 7.13 kcal/mol above the mixed disulfide (11.98 kcal/mol above the reactants). The last value is the only one that deviates from thermoneutrality, probably due to the absence of the environment (either aqueous or proteic). In the more realistic aqueous environment the mixed disulfide is still at 4.85 kcal/mol above the reactants and the product lies 1.03 kcal/mol above the mixed disulfide (5.88 kcal/mol above the reactants).

The deviations from thermoneutrality probably arise from small differences between the interaction with the environment of the calculations (either water or gas phase) and the interactions with the physiological protein/water mixed environment.

As we will see subsequently, this intrinsic thermoneutrality has a fundamental importance in the wide range of functions played by the members of the Trx family.

The reaction barriers for the CGGC system are considerably higher than for the previous studied CGPC system. In the first step, in water, the 3RC₂ (reacting complex connected to the TS) is −1.85 above 3RC₁ in the present system. In the CGPC system 3RC₂ is 0.35 above 3RC₁. The transition state in the CGPC system is 12.96 kcal/mol above 3RC₂ compared with the 20.70 reported here (3TS). In the second step in water, the transition state in the CGPC system is 21.64 kcal/mol above 4RC compared with the 31.92 kcal/mol in the present system.

Regarding the distances in the CGPC system, the 3RC₂ and 4RC structures engage in one less hydrogen bridge with the thiolate, which makes this thiolate less stabilized and hence a better nucleophile. Accordingly, the distances between the thiolate and the attacked sulfur are larger in the CGGC system. In the transition states, the negative charge is spread over the three sulfurs and hence the amides do not establish or establish weaker hydrogen bonds with the sulfur atoms.

In conclusion, the substitution of a proline by a glycine (the introduction of an additional amide group) makes the reaction less favorable because the amide group stabilizes the anionic reactants and mixed disulfide more than it stabilizes the associated TSs.

**Discussion**

The thioredoxin family of enzymes catalyzes oxidation, reduction, and isomerization reactions apparently using the same cata-
Figure 7. Reactants and transition state for the first step the catalytic cycle in aqueous solvent. All distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving \( S_{\text{nuc}} \) are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Figure 8. Products for the first step the catalytic cycle in aqueous solvent, before (3PC₁) and after (3PC₂) proton transfer from \( S_{\text{nuc}} \) to \( S_{\text{lg}} \). All distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving \( S_{\text{lg}} \) are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]
lytic mechanism and a similar proteic scaffold. The molecular-level determinants for differentiation have remained unclear, elusive, and not understood so far, despite the great efforts to unravel them.

Taken together, the quantum-mechanical simulations on active-site models have highlighted an unusual characteristic associated with the reactivity of these enzymes: small details in the active site hydrogen-bonding network strongly influence the structure and energetics of the intermediates along the catalytic cycle. They influence in a decisive manner the reactivity of the participant nucleophiles, and the relative energy of the several intermediates, and even of the reactants and products.

In this way, the local electrostatic and hydrogen-bonding environment can tune the thermodynamics of the involved spe-

**Figure 9.** Reactants (4RC) and transition state (4TS) for the second step the catalytic cycle in aqueous solvent. All distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving S_{nur} are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

**Figure 10.** Products for the second step the catalytic cycle in aqueous solvent. All distances are indicated in Å. Atomic charges for the sulfur atoms are given in a.u. The thiols and backbone amine groups of the CGGC motif are depicted as spheres. All the other atoms are drawn in stick representation. The hydrogen bonds involving S_{c} are also highlighted as black dashed lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]
cies, favoring the overall equilibrium toward one side or the other of the cycle (i.e., toward oxidation or reduction).

Could this constitute the mechanism by which the Trx family members achieve their remarkable specificity? Is it reasonable to accept that small details on hydrogen bonding, or more generally, on the electrostatic potential around the active site can shift the equilibrium toward the reactants, or the products, or even to trap the cycle in a stabilized intermediate, promoting oxidation, reduction or isomerization? Such a mechanism is unprecedented in enzyme catalysis, as far as the authors know. Therefore, what makes this family of enzymes so particular? The answer lies in another very unusual characteristic of Trxs: all of the participants in the many chemical events are cysteines. Two cysteines in the active site of all the enzymes and two more cysteines whatever the substrate. In the latter, exceptions can occur even though the rare species that are not cysteines will still have thiolate groups as reacting groups. Therefore a very high degree of chemical similarity will be always present. Moreover, they all react in the same way, through nucleophilic addition of their side chain thiolates to a disulfide bond, in each single step of all the catalytic cycles. If the reader looks again to the figures above, he/she can see that the whole catalytic cycle is made of a series of identical \( S_N2 \) reactions, perpetuated by thiols with very similar chemical properties. It is perfectly obvious for the chemist that in this situation all reaction steps will be basically thermoneutral in an isotropic environment (as it happens in the model reaction involving only methylthiols, both in vacuum and in solvent\(^1\)). Any difference can only arise from the enzyme machinery. This is the reality of the Trx family, but it is also an extremely unusual situation in enzyme chemistry. In such context of “reactivity equilibrium” and thermoneutrality, it is perfectly understandable that the electrostatic potential created by the enzyme can determine the course of the reaction. Consequently, even though all the enzymes have a similar structure, their slight differences are enough to confer oxidizing, reducing, or isomerization properties.

Figure 11. Potential energy surfaces at BP86 for the whole catalytic cycles in gas phase (A and B) and water (C and D). Relevant distances are also depicted.
On the basis of the principle of thermodynamic control through local electrostatics around the cysteine thiols, it becomes possible for the first time to arrive at an understanding of the long-waited elusive properties that differentiate the enzyme members, which can be summarized as follows:

1. **Reductant enzymes** (as Trx) must possess a C_nuc not specifically stabilized, in order to act as an efficient nucleophile toward their substrate pairs of cysteines involved in a disulfide. They must also possess a C_bur not stabilized, in order to attack C_nuc and proceed to the reduction reaction toward the final stage, the formation of the enzyme disulfide. Interestingly, the pK_a of C_nuc (around 6.9–7.5) is absolutely perfect for this purpose: it is low enough to warrant deprotonation, but not lower than that, in order to keep a good nucleophilic character. The pK_a or nucleophilicity of C_bur has never been possible to measure, but the fact that Trx has a proline in the active site motif is certainly meaningful. Prolines usually have structural roles in proteins, due to the mechanical tension induced by the cyclization of their side chains with the amide backbone. However, the perfect superposition of all CXYC motifs in the family excludes such role for the active site proline in Trx. The incapacity of the active site proline to make a hydrogen bond to S_nuc is consistent with the absence of specific stabilization for this group and shall represent the real role of this active site motif.\(^{13}\)

2. **Oxidant enzymes** (such as DsbA) fall on the other side of the spectrum. They must possess an electrostatic environment that stabilizes both active site cysteines. C_bur is released upon attack of a substrate cysteine on the enzyme disulfide and its stabilization will promote this step. Remember that a small stabilization is enough, given the similarity between all the reaction groups. Unfortunately, it has not been possible to measure experimentally the nucleophilicity of the buried cysteines in any of the family members, so this point awaits definitive experimental confirmation. However, this is not the case for C_nuc in DsbA. The microenvironment model here proposed implies that C_nuc must also be highly stabilized, to be quantitatively released from the mixed disulfide in the second reaction step. It is impressive to realize that the experimental value of the pK_a for this cysteine (3.5) is consistently low, the lowest of the family, and by far the lowest of any of the substrate cysteines.

3. **Isomerases** (such as PDI), have a more obscure mechanism of action. According to the typical model of action for these enzymes, they must possess an electrostatic environment that stabilizes both active site cysteines. C_bur is released upon attack of a substrate cysteine on the enzyme disulfide and its stabilization will promote this step. Remember that a small stabilization is enough, given the similarity between all the reaction groups. Unfortunately, it has not been possible to measure experimentally the nucleophilicity of the buried cysteines in any of the family members, so this point awaits definitive experimental confirmation. However, this is not the case for C_nuc in DsbA. The microenvironment model here proposed implies that C_nuc must also be highly stabilized, to be quantitatively released from the mixed disulfide in the second reaction step. It is impressive to realize that the experimental value of the pK_a for this cysteine (3.5) is consistently low, the lowest of the family, and by far the lowest of any of the substrate cysteines.

\[\text{Scheme 2. Thiol-disulfide exchange reactions regulation across the thioredoxin family is promoted by the differential stabilization of the active site cysteines (C_nuc and C_bur). The red arrow represents reduction that is carried by thioredoxins and glutaredoxins. The blue arrow represents oxidation, which is performed by DsbA. And finally isomerisation (PDI) is represented in green. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]}\]
enzymes (illustrated in scheme 1), they must possess a moderately nucleophilic C
\textsubscript{nuc}, capable of attacking misformed disulfides and engaging in stable mixed disulfides. They should also constitute a good leaving group, to allow for the resolution of the mixed disulfide through intramolecular attack of another substrate thiol, generating the native bond. In such a situation, the enzyme will present a stable S\textsubscript{bur} to promote a stable mixed disulfide, and a moderately stable C\textsubscript{nuc}, to allow for attack and elimination as exposed above. Remarkably, the pK\textsubscript{a} of C\textsubscript{nuc} in PDI (4.5) is intermediate between the oxidant (Dsla) and reductant (Trx) values, in agreement with the proposed working scheme. Anyway, we must emphasize that the mechanism of isomerization is poorly known and probably it can occur through different pathways (e.g. cycles of reduction of misformed disulfides and oxidation in a different way). Therefore, an interpretation of the PDI catalytic power should be seen as tentative and definitive conclusions can only be obtained under a context of a better understanding of the (possible) multiple pathways for isomerization.

Scheme 2 summarizes the proposed general model for the action of the enzymes of this family, based on the inherent concept of thermoneutrality and specific stabilization of the different intervening cysteines.

Conclusion

In this work, we have calculated the PES for the oxidation of a di thiol unit of a CGGC tetrapeptide with concomitant reduction of the disulfide bond of an oxidized dimethylthiol molecule. This system corresponds to a model for the enzymatic oxidation of substrates by a Trx family enzyme. The two central residues were changed to glycines to understand the influence of the specific residues of the C\textsubscript{nuc}XYC\textsubscript{bur} on the properties of the active site of the family members. The results have shown that the backbone of the XY motif specifically stabilizes C\textsubscript{nuc} through hydrogen bonds between the peptidic amine and S\textsubscript{nuc}. The introduction of prolines in this conserved sequence (a common feature in the proteins of the family members will greatly influence the reactivity of the nucleophilic and buried cysteines. Taken together, the results obtained allowed to build up a consistent mechanism for the underlying molecular principles that enable this family of very similar enzymes to promote specifically oxidations, reductions or isomerizations, and to fulfill a wide variety of physiological roles.

References


Journal of Computational Chemistry DOI 10.1002/jcc