

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/274401212>

# Dissection of the Water Cavity of Yeast Thioredoxin 1: The Effect of a Hydrophobic Residue in the Cavity

ARTICLE in **BIOCHEMISTRY** · APRIL 2015

Impact Factor: 3.02 · DOI: 10.1021/acs.biochem.5b00082 · Source: PubMed

---

CITATIONS

2

---

READS

68

## 5 AUTHORS, INCLUDING:



[Francisco Gomes-Neto](#)

Fundaão Oswaldo Cruz

8 PUBLICATIONS 49 CITATIONS

[SEE PROFILE](#)



[Ana Paula Valente](#)

Federal University of Rio de Janeiro

117 PUBLICATIONS 2,327 CITATIONS

[SEE PROFILE](#)



[Fabio C L Almeida](#)

Federal University of Rio de Janeiro

112 PUBLICATIONS 1,329 CITATIONS

[SEE PROFILE](#)

## Dissection of the Water Cavity of Yeast Thioredoxin 1: The Effect of a Hydrophobic Residue in the Cavity

Anwar Iqbal,<sup>†</sup> Francisco Gomes-Neto,<sup>§</sup> Catarina Akiko Myiamoto,<sup>||</sup> Ana Paula Valente,<sup>‡</sup> and Fabio C. L. Almeida<sup>\*,†</sup>

<sup>†</sup>Institute of Medical Biochemistry, National Center of Nuclear Magnetic Resonance Jiri Jonas, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

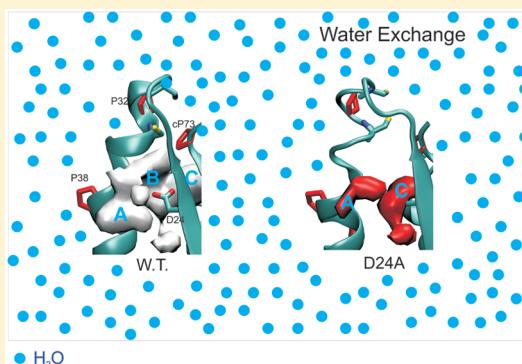
<sup>‡</sup>Center of Structural Biology and Bioimaging (CENABIO), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>§</sup>Laboratory of Toxicology, Instituto Oswaldo Cruz, Fiocruz, 21045-900 Rio de Janeiro, Brazil

<sup>||</sup>Faculdades Integradas de Três Lagoas-AEMS, Três Lagoas, Brazil

### Supporting Information

**ABSTRACT:** The water cavity of yeast thioredoxin 1 (*y*Trx1) is an ancestral, conserved structural element that is poorly understood. We recently demonstrated that the water cavity is involved in the complex protein dynamics that are responsible for the catalytically relevant event of coupling hydration, proton exchange, and motion at the interacting loops. Its main feature is the presence of the conserved polar residue, Asp24, which is buried in a hydrophobic cavity. Here, we evaluated the role of the solvation of Asp24 as the main element that is responsible for the formation of the water cavity in thioredoxins. We showed that the substitution of Asp24 with a hydrophobic residue (D24A) was not sufficient to completely close the cavity. The dynamics of the D24A mutant of *y*Trx1 at multiple time scales revealed that the D24A mutant presents motions at different time scales near the active site, interaction loops, and water cavity, revealing the existence of a smaller dissected cavity. Molecular dynamics simulation, along with experimental molecular dynamics, allowed a detailed description of the water cavity in wild-type *y*Trx1 and D24A. The cavity connects the interacting loops, the central  $\beta$ -sheet, and  $\alpha$ -helices 2 and 4. It is formed by three contiguous lobes, which we call lobes A–C. Lobe A is hydrophilic and the most superficial. It is formed primarily by the conserved Lys54. Lobe B is the central lobe formed by the catalytically important residues Cys33 and Asp24, which are strategically positioned. Lobe C is the most hydrophobic and is formed by the conserved *cis*-Pro73. The central lobe B is closed upon introduction of the D24A mutation, revealing that independent forces other than solvation of Asp24 maintain lobes A and C in the open configuration. These data allow us to better understand the properties of this enzyme.



Thioredoxins (Trx) are small proteins that are present in all organisms from Archaea to the higher eukaryotes and are essential in many life forms.<sup>1</sup> They display disulfide reductase activity, which is essential for maintaining the oxidative–reductive potential of the cells. Thioredoxin reductase recycles the reduced Trxs in the cell by using the reductive potential of NADPH.<sup>2–4</sup> Trx binds to several targets in the cell, thereby participating in multiple regulatory events. For example, Trx is involved in apoptosis (Ask1), tumor suppression (TxNIP), control of DNA synthesis and gene expression (ribonucleotide reductase, NF- $\kappa$ B), control of oxidative events (peroxiredoxins), and nitric oxide (NO) signaling. In yeast, Trx also controls several processes, including 3'-phosphoadenylsulfate (PAPS) reductase activity.<sup>2</sup>

The conformation of thioredoxins has been conserved for at least 4 billion years of evolution.<sup>5</sup> The water cavity is one of the main conserved structural elements.<sup>6–8</sup> It is our goal to describe the structural properties of the water cavity. The central residue

of the water cavity is the largely hydrated, conserved, buried aspartic acid 24 (*y*Trx1).<sup>9</sup> During the 1980s and 1990s, the acceptance of a proton by this aspartic acid in the reduction of the protein target was proposed as one of the mechanisms for the activation of the internal cysteine.<sup>7,10–13</sup> We showed a complementary role for Asp24 via coupling hydration and conformational motion of the water cavity and interacting loops.<sup>9</sup> The conformational equilibrium of thioredoxins is essential for understanding the catalytic cycle. The water cavity is the central structural element for *y*Trx1 dynamics.

The water cavity is underneath the interacting loops and is protected by  $\alpha$ -helix 2 and the central  $\beta$ -sheet. Motion at the water cavity affects the interacting loops and vice versa. The D24N mutation suppresses part of the conformational motion

**Received:** January 28, 2015

**Revised:** March 30, 2015

**Published:** April 1, 2015



at the water cavity.<sup>9</sup> Hydration studies and molecular dynamics simulations showed that Asn24 is more exposed to water than Asp24 is and that the water cavity is smaller in D24N. The pK<sub>a</sub> values of the buried aspartic acid and cysteine have been determined to co-titratae at approximately 7.6, which is near the optimal catalytic activity of thioredoxin. The internal negatively charged residue and the water cavity are stabilized by the presence of a conserved lysine ( $\beta$ 3 strand) in the vicinity. The long lysine side chain creates a positively charged environment near the buried aspartic acid and extending up to the protein surface.<sup>12</sup>

Our group determined the solution structure of the reduced *Saccharomyces cerevisiae* Trx1<sup>14,15</sup> (yTrx1) and Trx2<sup>16</sup> (yTrx2). The crystallographic oxidized yTrx1 and yTrx2 structures are also available.<sup>17–19</sup> The structure of the active site of yTrx1 in the reduced state<sup>15</sup> is characterized by the exposed Cys30 and the internal Cys33, which faces the water cavity. The buried aspartic acid is Asp24, and the conserved lysine is Lys54. The interacting surface is formed by three loop regions: residues 56–59, 70–73, and 90–93 ( $\beta$ 3– $\alpha$ 3 loop,  $\alpha$ 3– $\beta$ 4 loop, and  $\beta$ 5– $\alpha$ 4 loop, respectively). Pro73 is the *cis*-proline in the  $\alpha$ 3– $\beta$ 4 loop, which is important for the catalytic activity of Trxs and precludes metal binding.<sup>20</sup> The water cavity is underneath the interacting loops.<sup>15,21</sup>

The conformational free energy is the driving force for Trx activity, in which the oxidized conformation is more stable than the reduced conformation. The protonation state of the buried aspartate finely tunes the stability of the protein. For activity in *Escherichia coli* thioredoxin, the protonation of Asp26 stabilizes the protein, with an increase of the  $\Delta G^\circ$  for unfolding from 8.6 to 10.6 kcal/mol. In a similar manner, the D26A mutation stabilizes the protein, with a  $\Delta G^\circ$  of unfolding of 13.2 kcal/mol.<sup>12</sup>

The total energy of a protein is a balance between the conformational free energy and the solvation energy. For thioredoxins, the opening–closing equilibrium of the water cavity changes the overall energy of the system, and the desolvation of the buried aspartic acid may contribute significantly to the overall free energy. However, the question of how the free energy of solvation of the buried polar group affects the dynamics and stability of Trx and, specifically, the water cavity remains unanswered. The solvation free energy is high for the aspartate residue ( $-82.57/-80.65$  kcal/mol) and smaller for the asparagine side chain ( $-17.47/-9.70$  kcal/mol).<sup>22</sup> The solvation free energy of the protonated aspartate is similar to that of the asparagine. Our early work suggested that the D24N mutation facilitated the desolvation that is necessary for closing the inner part of the water cavity, thereby quenching part of the dynamics of the protein.<sup>9</sup>

In the work presented here, we measured the effect of a neutral hydrophobic residue in the center of the cavity (Ala24). Specifically, we evaluated the dynamics of the D24A mutant and the role of residue Asp24 in the formation of the water cavity and coupling of dynamics and molecular recognition in thioredoxins. We showed that Asp24 is fundamental for the integrity of the water cavity and positioning of the catalytically relevant residues. Additionally, there are other elements of the cavity that are conserved and important for the constitution of the cavity. The dynamics of the D24A mutant at multiple time scales, as supported by molecular dynamics simulations, revealed key residues for the conformational equilibrium of the cavity. The cavity is located between the central  $\beta$ -sheet and  $\alpha$ -helices 2 and 4. It is formed by three contiguous lobes, which

we call lobes A–C. Lobe A is the most superficial lobe and the most hydrophilic. It is formed by the conserved side chain Lys54. Lobe B is the central lobe and is formed by catalytic residues Cys33 and Asp24. Lobe C is the most hydrophobic and is formed by the conserved *cis*-Pro73. The central lobe B is closed in the D24A mutant, whereas lobes A and C remain open, thus allowing us to study and understand their properties.

## METHODS

### Expression and Purification of Yeast Thioredoxin 1.

Plasmid pET17-b containing the sequence of yTrx1 and mutant D24A were used for transformation in *E. coli* Bl21(DE3) cells. Cells were grown in M9 medium enriched with <sup>15</sup>NH<sub>4</sub>Cl as the only source of nitrogen. Isopropyl  $\beta$ -D-1-thiogalactopyranoside (1 mM) was added once the cells grew to an OD<sub>600</sub> of 0.7. Following centrifugation at 6000g for 30 min, cells were resuspended in 50 mM Tris-HCl buffer (pH 7) at 4 °C and disrupted by sonication. After a second bout of centrifugation, the protein supernatant was purified with an ion exchange Q-Sepharose column (GE Lifescience) with a salt gradient and by size exclusion on Sephacryl 26/60 resin (GE Lifescience) with 20 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7).

**NMR Experiments.** NMR spectra were recorded on a 600 MHz Bruker Avance III instrument at 25 °C. The NMR spectra were processed using NMRPipe, version 2.3,<sup>35</sup> and analyzed using CCPNNMR Analysis, version (analysis 2.4.0).<sup>23</sup> All samples of yTrx1 and the D24A mutant were prepared at a concentration of 200  $\mu$ M with 20 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7) and a H<sub>2</sub>O/D<sub>2</sub>O mixture (90/10) in the oxidized and reduced state. The protein was reduced by the addition of deuterated dithiothreitol (DTT) (30 mM).

**Relaxation Experiments.** The relaxation experiments were performed on a 600 MHz Bruker Avance III instrument at 25 °C. The relaxation experiments were used to determine the <sup>15</sup>N  $R_1$ , <sup>15</sup>N  $R_2$ , and <sup>1</sup>H–<sup>15</sup>N heteronuclear nuclear Overhauser effect (NOE) with 1024 and 96 complex points in  $F_2$  (<sup>1</sup>H) and  $F_1$  (<sup>15</sup>N), respectively.  $R_1$  and  $R_2$  were measured with delays of 20, 50, 70, 100 (three times), 150, 300, 500, 1000, and 1500 ms and 47, 136, 78.56 (three times), 109.984, 141.408, 172.832, 204.256, 235.68, 267.104, and 298.528 ms, respectively. The errors in the peak intensities were calculated from the  $R_1$  and  $R_2$  experiments with replicate delay times to calculate the standard deviation. The <sup>1</sup>H–<sup>15</sup>N NOE was acquired with or without proton saturation for 5 s. The  $R_1$ ,  $R_2$ , and NOE values were obtained using CCPNNMR Analysis.<sup>23</sup> The <sup>1</sup>H–<sup>15</sup>N heteronuclear NOEs were determined from the ratio of peak intensities with and without the saturation of amide protons. The relaxation parameters were fit in accordance with the Lipari–Szabo formalism<sup>24,25</sup> to obtain the intramolecular dynamics for each residue. We used TENSOR2<sup>26</sup> to define a motional model with an axially symmetric diffusion model in the presence of an anisotropic tensor. These calculations were performed using the data from the residues whose  $R_2/R_1$  ratios were within one standard deviation of the average  $R_2/R_1$  ratio. The deposited structures of reduced [Protein Data Bank (PDB) entry 2I9H<sup>15</sup>] and oxidized (PDB entry 3F3Q<sup>18</sup>) yTrx1 were used to calculate the diffusion tensors.

**Dipolar Interaction with Water.** The three-dimensional (3D) NOESY-HSQC and 3D ROESY-HSQC spectra were used to obtain the NOE at the laboratory frame and the rotating frame (ROE), respectively, of residues of wild-type [<sup>15</sup>N]Trx1 or mutant D24N [<sup>15</sup>N]Trx1 with water. These data were acquired with 1024, 32, and 64 complex points in  $F_3$  (<sup>1</sup>H),

$F_2$  ( $^{15}\text{N}$ ), and  $F_1$  ( $^1\text{H}$ ), respectively, using 16 scans and a mixing time of 20 ms.

**Mapping of Structural Change.** The  $^1\text{H}-^{15}\text{N}$  HSQC spectra were recorded at 600 MHz using  $^{15}\text{N}$ -labeled wild-type yTrx1 and mutant D24A. All experiments were conducted at 25 °C in 20 mM  $\text{NaH}_2\text{PO}_4$  (pH 7) and 10%  $\text{D}_2\text{O}$ . The chemical shifts perturbations (CSPs) were calculated using the equation  $\Delta\delta = [(\Delta^{\text{HN}})^2 + (\Delta^{\text{N}}/10)^2]^{1/2}$ , where  $\Delta^{\text{HN}}$  and  $\Delta^{\text{N}}$  are the differences between chemical shifts, in parts per million, observed for the hydrogen and nitrogen nuclei, respectively.

**Molecular Dynamics (MD) Simulations.** We performed two sets of MD simulations: (i) wild-type yTrx1 (reduced and oxidized states) and (ii) D24A mutant (reduced and oxidized states). We used GROMACS version 4.5.4<sup>27</sup> to build the models and for energy minimization and all MD simulations.

As initial coordinates for the reduced state (wild type), we used the NMR solution structure of PDB entry 2I9H,<sup>15</sup> and for the oxidized state (wild type), we used the crystal structure of PDB entry 3F3Q<sup>18</sup> without the crystallographic water molecules. The ionization state of the side chains and the N- and C-termini was fixed using PROPKA setting the pH to 7.0.<sup>28</sup>

The D24A mutant models (reduced and oxidized forms) were built using the molecular visualization program VMD-XPLOR<sup>29</sup> by applying the module “mutate residue”.

The Amber99sb force field was used to set up the molecular topologies.<sup>30</sup> Water was modeled using the TIP3P model.<sup>31</sup> The proteins were inserted into the center of a cubic box (8 nm side) with a total volume of 512 nm<sup>3</sup>. Solvent was added totaling 16543 water molecules. To neutralize the total charge of the system, three molecules of water were replaced with sodium ions in the models representing the wild-type protein (reduced and oxidized states); in the models representing D24A mutants, two molecules of water were replaced with sodium ions (reduced and oxidized states).

The initial models (protein, water, and ions) were subjected to a three-step energy minimization: (i) position-restrained steepest descents applied on heavy atoms (harmonic constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup>), (ii) steepest descents, and (iii) conjugate gradient to correct bad contacts and geometry deviations. The models were equilibrated allowing solvation and temperature stabilization throughout three sequential short MD simulation steps.

The MD simulations were performed in the *NPT* ensemble, which consists of an ensemble keeping the number of particles ( $N$ ), system pressure ( $P$ ), and temperature ( $T$ ) constant and conserved. The pressure was controlled by coupling the system to the isotropic Berendsen pressure bath (at 1 atm with a time constant of 1.0 ps). The temperature was controlled by coupling the system to the Velocity Rescale Thermostat (at 298 K with a time constant of 0.1 ps), working independently for protein, ions, and solvent. The LINCS<sup>32</sup> and SETTLE<sup>33</sup> algorithms were used to constrain the bond lengths of protein and solvent, respectively. The long-range electrostatics were modeled using the particle mesh Ewald method (PME),<sup>34</sup> and short-range repulsive/attractive dispersion interactions were modeled using the Lennard-Jones potential, both with a 1.0 nm cutoff for the real-space calculation.<sup>35</sup> The MD integration time step used was 2 fs for recording coordinates every 10 ps and recording energies every 0.5 ps.

In the first step, we used the GROMACS simulated annealing protocol to perform a linear increase in temperature, from 0 to 298 K over 2 ns, applying a positional restraint potential on the protein heavy atoms (1000 kJ mol<sup>-1</sup> nm<sup>-2</sup>),

thereby allowing solvation of the protein by the water molecules. The second step was the stepwise reduction of the position restraint potential from 1000 to 0 kJ mol<sup>-1</sup> nm<sup>-2</sup> in a sequence of 26 MD simulations of 200 ps each (1000, 800, 600, 400, 300, 250, 200, 175, 150, 125, 100, 75, 50, 25, 23, 20, 18, 16, 14, 12, 10, 8, 6, 4, 2, and 0 kJ mol<sup>-1</sup> nm<sup>-2</sup>). In the third step, the system was equilibrated in a 1 ns MD simulation where the rmsd stabilized at  $1 \pm 0.2$  Å for backbone and  $2 \pm 0.2$  Å for side chains atoms. The production phase consisted of the simulation of all models for a total time of 200 ns.

**MD Simulation Analysis.** The generated trajectories were analyzed using GROMACS version 4.5.4. All images were generated using VMD.<sup>29</sup>

**The Radial Distribution Function (RDF) of Water Molecules as a Function of Asp/Ala N Atom.** The probability of finding a water molecule around residue 24 (Asp or Ala) was calculated from the radial distribution function between the amidic nitrogen atom (N) and the oxygen atom of water molecules using the software g\_rdf included in the GROMACS application suite.

**Quantification of Water Molecules Surrounding the Asp/Ala N Atom.** To determine the number of water molecules around the amidic nitrogen (N) of amino acid Asp/Ala24 over the MD simulation time, we used the g\_mean software included in the GROMACS application suite. The first step was to calculate the fluctuation of the distances between the N atom and the oxygen of the water molecules (OW atom) during the MD simulation time. Finally, we computed only the water molecules present in the distance cutoff established in the radial distribution analysis that was previously described.

The number of water molecules (inside the distance cutoff) was expressed as a function of the MD simulation time (picoseconds).

**Hydrogen Bond Analysis.** The identification of hydrogen bonds between the amino acid residues and the water molecules was performed with a geometrical criterion. When the triplets of donor, hydrogen, and acceptor atoms are at a distance of <0.35 nm and at an angle of <30°, these atoms are considered hydrogen-bonded (ON state).<sup>36</sup> Following the MD simulation trajectory, a condition from the geometrical criterion for the triplet of donor, hydrogen, and acceptor atoms turns this hydrogen bond to the OFF or ON state.

A hydrogen bond existence map was generated for all hydrogen bonds between the protein atoms and the water, where each hydrogen bond detected was identified by a unique (and sequential) numerical index and plotted as a function of simulation time. A hydrogen bond in the ON state is represented by a red pixel, and a bond in the OFF state is represented by a white pixel.

To evaluate the time during which a hydrogen bond is into ON state, the number of red pixels was computed and expressed as a function of a percentage of the total frames in the simulation. To express a change in the behavior in the hydrogen bond profile, we chose only the most prevalent hydrogen bond (the longest-lived) for each amino acid residue and plotted the data as a function of residue number.

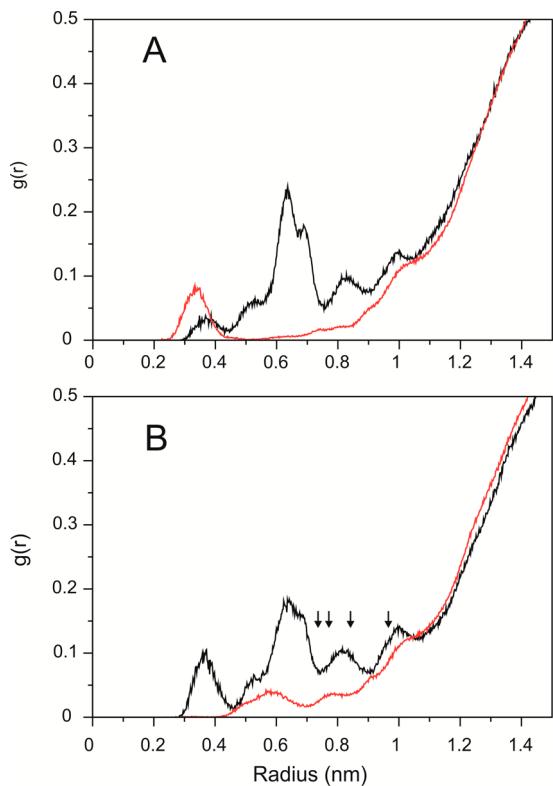
**Cavity Detection in Trx1.** Cavity detection was performed using the Fpocket software based on the detection of  $\alpha$ -spheres from the protein structure over all coordinates calculated during the MD simulation (20000 structures per MD simulation).<sup>37</sup> When  $\alpha$ -spheres clustered together, cavities were identified, thus allowing the calculation of properties for

this feature. All images were generated using VMD-Xplor,<sup>29</sup> and the cavities were represented as solid isosurfaces.

## RESULTS

To obtain information about the state of hydration of the water cavity, we ran molecular dynamics (MD) simulations of the D24A mutant and compared the results with those of wild-type  $\gamma$ Trx1 for both oxidized and reduced states. The starting points for simulations were the solution structure of the reduced state<sup>21</sup> and the crystal structure of the oxidized state.<sup>18</sup> The mutant was constructed *in silico* and was stable throughout the simulation.

The radial distribution function [ $g(r)$ ] (Figure 1) of water relative to the residue 24 amide nitrogen atom reflects the water



**Figure 1.** Radial distribution function of water molecules relative to the amide nitrogen of Ala24 (red) in the D24A mutant and Asp24 (black) in wild-type  $\gamma$ Trx1. Plot A is the radial distribution function of the reduced state and plot B that of the oxidized state. The arrows indicate the distance relative to the nitrogen of Asp24 of the four nearest crystallographic waters.

density around this atom. The water cavity of wild-type  $\gamma$ Trx1 forms clusters of water molecules along the MD trajectory. On average, the cavity is filled with one to five water molecules at the same time.<sup>9</sup> Densities at distances of  $>9$  Å reflect the bulk water outside the cavity. We observed five high-density peaks for wild-type  $\gamma$ Trx1 for both oxidized and reduced states (approximately 3.8, 5.2, 6.2, 7.0, and 8.2 Å). These sites are the five water pockets inside the cavity, which are similar for the reduced and oxidized states (Figure 1). The arrows in Figure 1 show the four nearest crystallographic water molecules found in the oxidized state. These are the only internal molecules in the crystal structure. The fourth water molecule is already at the surface. The radial distribution function shows a good density of water at the distances expected for the crystallographic water

molecules. It means that we reasonably frequently find water at the same position as in the crystal during the MD simulation.

For the D24A mutant, the density of water decreased as a whole, suggesting a partial closing of the cavity. There is water inside the cavity, with occupancy in three sites for the reduced state (approximately 3.2, 7.5, and 8.0 Å) and three sites for the oxidized state (approximately 5.0, 6.0, and 7.9 Å). There is only one major density site for reduced D24A (3.2 Å) and one major site for oxidized D24A (5.0 Å). The density observed at larger distances, up to 7 Å, is considerably lower for the D24A mutant. This can be explained by the decrease in the size of the cavity with one preserved site at 3.2 Å, which exhibits high density. The other sites are preserved but with a significantly lower probability of occupancy compared with that of wild-type  $\gamma$ Trx1.

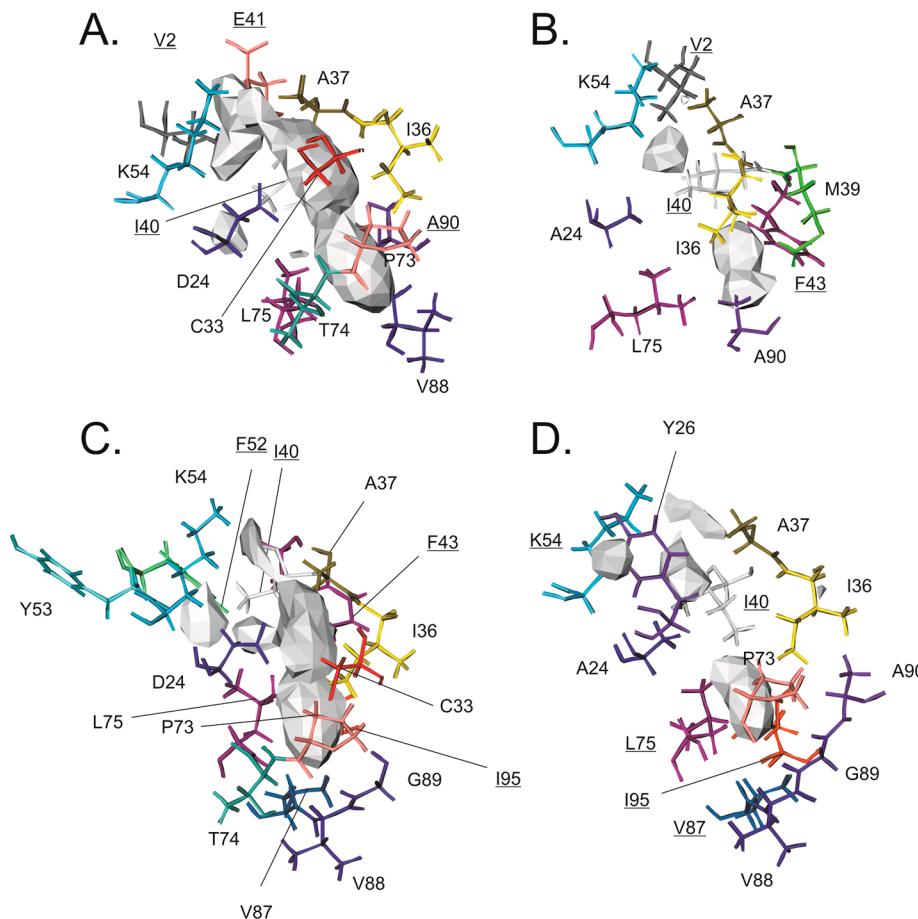
To obtain information about the location of water, we used two approaches: detection of the cavity with MD simulation and analysis of the water hydrogen bonding pattern. These data were further validated with chemical shift perturbations elicited by the mutation and confrontation with the protein dynamics obtained from <sup>15</sup>N nuclear spin relaxation experiments.

We detected the presence of cavities in the  $\gamma$ Trx1 structures using 20000 structures extracted from the MD simulations (Figure 2). The detection is based on an ensemble of  $\alpha$ -spheres fit along the MD trajectory, which allows us to locate intermediately sized  $\alpha$ -spheres that are typical of cavities.<sup>37</sup> This method was used previously to detect cavities of  $\gamma$ Trx1 and the D24N mutant.<sup>9</sup> Figure 2 shows the detected cavities that are representative of the MD trajectories of both oxidized and reduced states. Figure S1 of the Supporting Information shows the localization of the cavities at the  $\gamma$ Trx1 structure and the superposition of the wild-type and D24A mutant cavities. Figure 2 describes in detail the residues that are facing the cavity, and Figure S1 of the Supporting Information gives a reference of the position of the cavity, showing that upon mutation, the cavity became smaller, truncated, and delocalized.

Cavity detection agrees with the occupancy of water near residue 24 observed from the  $g(r)$  analysis. For wild-type  $\gamma$ Trx1, both oxidized and reduced, the water cavity was located underneath the interacting loops and between the central  $\beta$ -sheet and  $\alpha$ -helices 2 and 4. The effect of the mutation was truncation of the water cavity, which was split in half (Figure 2). This observation led to the conclusion that there is a central lobe that depends on the presence of Asp24 and two other lobes that are independent of Asp24.

For the sake of clarity, we classified the water cavity into three distinct lobes, which we call A–C. Table 1 shows the residues facing each lobe and summarizes the findings in Figure 2. Note that there are residues that are at the interface between two lobes. Lobe B is the central lobe, which contains the catalytic residues Cys33 and Asp24. In both the reduced and oxidized states, lobe B is the largest; in both states, particularly in the reduced state, Cys33 is strategically positioned in the top of Asp24 throughout the MD simulation.

Lobe A is localized between  $\alpha$ -helix 2 and  $\beta$ -strand 1. It is the most hydrophilic because of the presence of Lys54, which allows water exchange. This residue is largely conserved in evolution and is important for the stabilization of thioredoxins. The Lys54 side chain is facing the cavity and allows the interface between the water cavity and the protein surface. The other residues are hydrophobic, including V2 and I40. F52 is facing lobe A in the oxidized state, whereas Glu41 is facing lobe A in reduced  $\gamma$ Trx1.



**Figure 2.** Water cavities of wild-type yTrx1 and the D24A mutant. Panels A and C show the cavity of wild-type yTrx1<sup>9</sup> in the reduced and oxidized forms, respectively (cavity surface colored gray). Panels B and D show the cavity of the D24A mutant of the reduced and oxidized forms, respectively. To facilitate the analysis of the figure, the residues are color-coded according to the residue positions. The orientation of all cavities is similar, with the top view of the protein and the interacting loops (not being shown) at the top. For wild-type yTrx1, Cys33 and Pro73 are at the top of the cavity. Residues that are in the back of the cavity are underlined. All residues participating in the cavities are shown. We used Fpocket<sup>37</sup> (mdpocket) to compute the cavities over the 200 ns MD simulation. To compute the isosurface, we used the same isovalue of 3 for the four simulations.

**Table 1. Description of the Amino Acid Residues That Are Facing the Water Cavity of Wild-Type, D24A, and D24N yTrx1 in the Reduced and Oxidized States<sup>a</sup>**

cavity	reduced			oxidized		
	WT	D24A	D24N	WT	D24A	D24N
lobe A	V2, A37, E41, K54, I40	V2, A37, K54	Y26, C33, K34, M35, A37, K54	I40, F52, Y53, K54	Y26, K54, C33	V2, Y26, A37, E41, K54
lobe B	D24, C33, I36, A37, I40	closed cavity lobe: A24, I36, I40	V2, N24, I36, I40	D24, C33, I36, A37, I40	closed cavity lobe: A24, C33, I36, A37, I40	smaller lobe: N24, Y26, I36, I40
lobe C	I36, P73, T74, L75, V88, A90	I36, M39, F43, L75, A90	closed cavity lobe: V22, L75, F52	I36, P73, T74, L75, V87, V88, G89, I95	I36, P73, L75, V87, V88, G89, A90, I95	smaller lobe: P73, L75

<sup>a</sup>The water cavity was classified into three lobes (A–C). The table shows the residues facing each of lobes. Note that there are residues that are at two lobes because they are at the interface between the two.

Lobe C is located in the space near  $\alpha$ -helix 2, the  $\beta$ 5– $\alpha$ 4 loop, and  $\alpha$ -helix 4. It is the most hydrophobic lobe and is formed exclusively by hydrophobic residues. The fact that this lobe could be detected in the D24A mutant reveals that the cavity can be formed from an imperfect hydrophobic packing that is conserved in evolution and is independent of Asp24 solvation. The most hydrophilic residue is Pro73, which is at the protein surface. Pro73 is also largely conserved in evolution. In addition, lobe C is near the protein surface on the opposite side of Lys4.

There are three conserved prolines surrounding the water cavity: Pro32, *cis*-Pro73, and Pro38 (Figure S1 of the Supporting Information). The relative hydrophilicity of the proline residues could be a key aspect in the formation of the water cavity. The polarity of Lys4 in lobe A and Pro73 in lobe C is most likely very important in controlling the water exchange with the bulk water. The central lobe B is the most protected in the center.

The main consequence of the D24A mutation was the closing of the central lobe (Figure 2 and Table 1). It was

replaced with a hydrophobic well-packed core formed by residues Ile40, Ile36, and Ala24. Interestingly, lobes A and C are still present in both oxidized and reduced states. As mentioned above, the presence of lobes A and C means that the hydrophobic mispacking of lobe C, the hydrophilicity of Pro73, and the hydrophilicity of lobe A are essential for the formation of the water cavity. It also means that lobes A and C are independent of the solvation of Asp24.

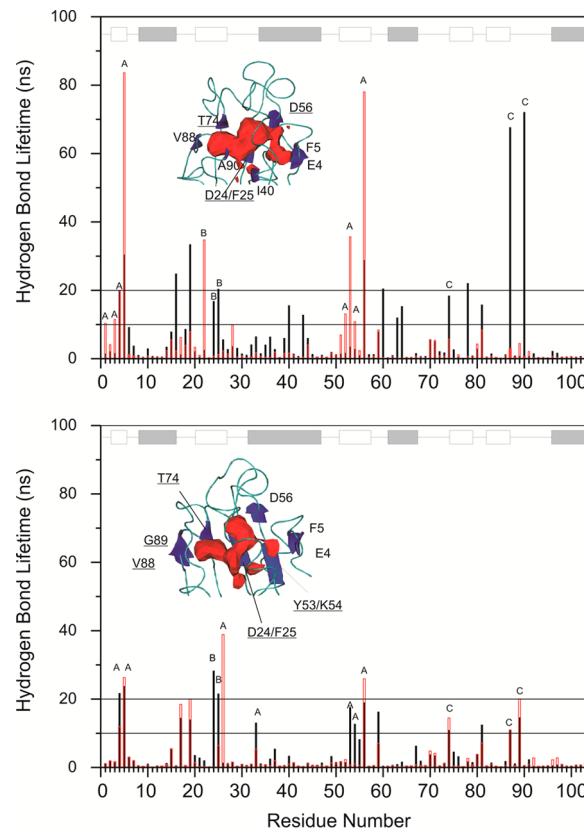
For comparison, we plotted the cavity of the D24N mutant, studied in a previous work<sup>9</sup> (Table 1 and Figure S2 of the Supporting Information). Different from the case for the D24A mutant, for D24N, the three lobes of the cavity are present, including the central lobe B. The D24N cavity is smaller than the cavity of the wild type. The hydrophilic lobe A is the most pronounced in D24N, as expected from the previous conclusion that the D24N cavity is more superficial and permeable to water.<sup>9</sup>

We also aimed to assess long-lived hydrogen bonding with water to obtain information about the presence of water at the described cavities. Thus, we detected the most prevalent hydrogen bond with water for each amide of the protein (Figure 3). This analysis was the most suitable way of detecting long-lived hydrogen bonding. The description of the longest-lived water hydrogen bond tells us that at least one water molecule stays at a given site for a certain time. Each bar in Figure 3 represents the water molecule that had the longest lifetime for each amide during the 200 ns MD simulation. We observed important differences when we compared wild-type and D24A *y*Trx1.

For the reduced state *y*Trx1, we observed that the following residues in the water cavity display a hydrogen bonding lifetime of >10 ns (present for more than 5% of the MD simulation time): Glu4, Phe5, Asp24, Phe25, Asp56, Ile40, Thr74, Val88, and Ala90. The level of water hydrogen bonding with residues participating in lobes B and C decreased substantially with the D24A mutation. It is worth noting that there was a substantial decrease in the water hydrogen bond lifetime at the amides participating in the hydrophobic lobe C. In the reduced state of D24A, lobes A and C of the cavity are present, but there is only water with a very low residence time.

Similar behavior was observed for the oxidized state of *y*Trx1. Most of the amides with water hydrogen bond residence times of >10 ns are in the proximity of the water cavity. Figure 3 shows the water cavity and the residues with a water hydrogen bonding time of >10 ns, which were Glu4, Phe5, Asp24, Phe25, Cys33, Tyr53, Lys54, Asp56, Thr74, Val88, Gly89, and Ala90. In an MD simulation of D24A, the mutation led to a significant decrease in the level of water hydrogen bonding with residues in lobe B, with the exception of Tyr26, which is actually located near lobe C. Lobes A and C remain hydrated, as observed by the hydrogen bonding pattern of the residues participating in the cavity.

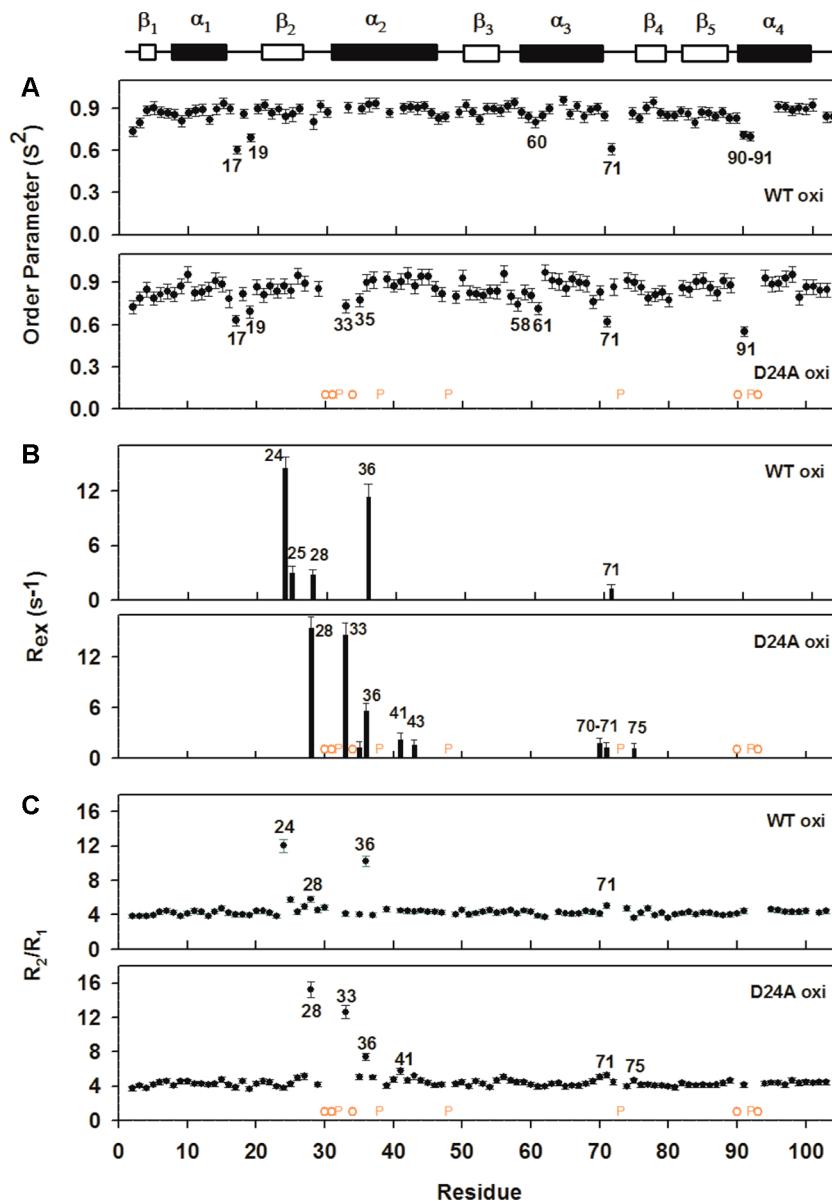
A detailed understanding of protein motion is informative with regard to the function of a protein, provided that the protein dynamics describe the equilibrium among the thermally accessible conformational states of a protein. Proteins visit high-energy conformational states to comply with their function.<sup>38</sup> We used <sup>15</sup>N nuclear relaxation parameters of the D24A mutant of *y*Trx1 to analyze the dynamic properties for the reduced and oxidized states and confront the theoretical data obtained from the MD simulations. A concentration of 200  $\mu$ M was used for all experiments for reduced and oxidized states because we found that wild-type protein at a higher concentration (>200



**Figure 3.** Hydrogen bond of water with protein amide. The graphics show the lifetime of the longest-lived water hydrogen bond (most prevalent) for the amino acid residue. The count was performed during a 200 ns MD simulation for each amino acid. The top panel represents data for the reduced state, and the bottom panel represents data for the oxidized state. The letters above the bars indicate whether the residue resides in lobe A (A), lobe B (B), or lobe C (C) of the water cavity. Bars with no label are from residues that are not part of the water cavity. The red bars refer to data for the D24A mutant, and the black bars refer to data for wild-type *y*Trx1. The insets are the ribbon representations of the *y*Trx1 structure. The insets show the residues that are part of the cavity with a total water hydrogen bonding time of >10 ns. Some water hydrogen bonds lived for more than 30% of the MD simulations: Phe5 and Asp56 for the wild type and Val88 and Ala90 for the D24A mutant all for the reduced state. Note also that most of the longest-lived water hydrogen bonds are at the main water cavity.

$\mu$ M) displayed changes in chemical shifts, relaxation parameters, and slow dynamics. The previously assigned amide resonances<sup>14,21</sup> were used for analysis. Residues Cys30, Gly31, Lys34, Ala90, and Ala93 were absent in the <sup>15</sup>N HSQC spectra because of extensive line broadening in the D24A mutant and for the wild-type protein.<sup>9</sup> We measured longitudinal relaxation rates ( $R_1$ ), transverse relaxation rates ( $R_2$ ), and <sup>15</sup>N/<sup>1</sup>H heteronuclear NOEs (Figures S3 and S4 of the Supporting Information).

The order parameter ( $S^2$ ) and conformational exchange ( $R_{ex}$ ) based on the Lipari–Szabo model-free approach<sup>24,25</sup> for reduced and oxidized wild-type *y*Trx1 (for comparison, see ref 9) and the D24A mutant (Figures 4 and 5) show that the D24A mutation changed the motions of residues facing the bulk water and water cavity. In the oxidized form, wild-type *y*Trx1 showed milli- to microsecond motions ( $R_{ex} > 0$ ) in residues facing the water cavity, i.e., Asp24 and Ile36. All of the

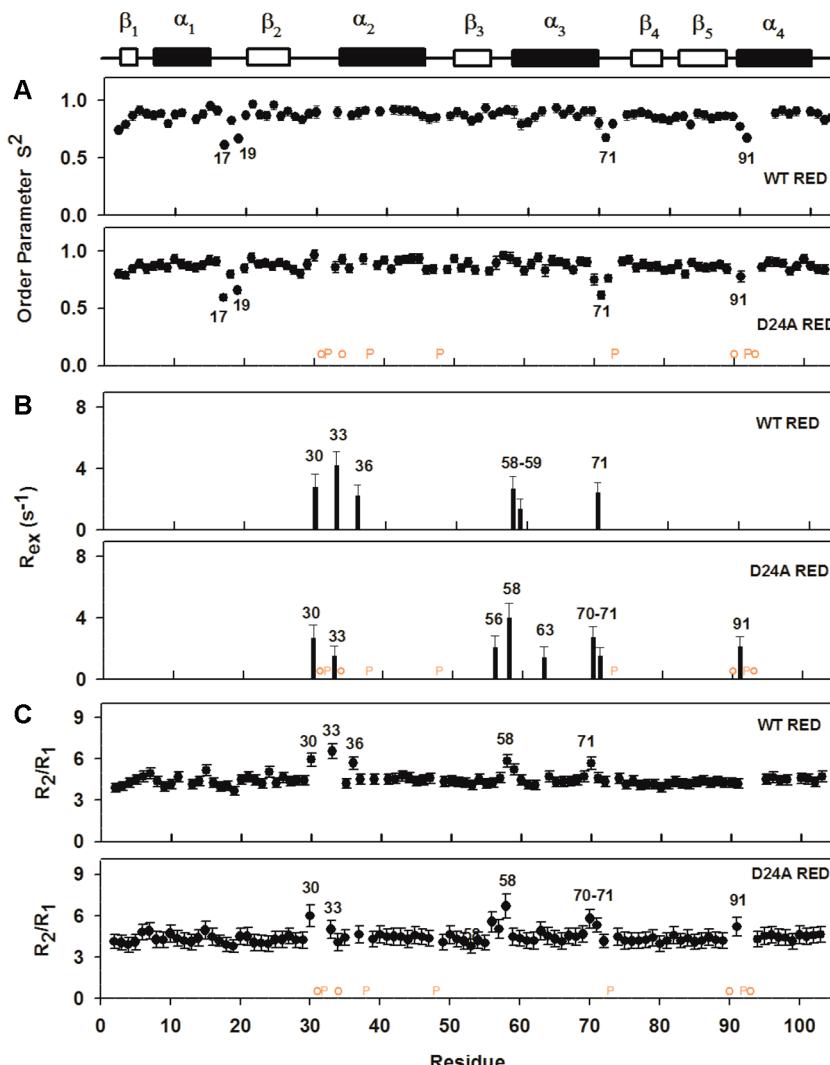


**Figure 4.** Dynamics based on the Lipari–Szabo model-free formalism of oxidized yTrx1 and mutant D24A. (A) Order parameter ( $S^2$ ) for oxidized wild-type yTrx1 (top) and its D24A mutant. (B) Rate of residues in conformational exchange ( $R_{ex}$ ) as a function of residue number of wild-type yTrx1 (top) and its D24A mutant (bottom). (C)  $R_2/R_1$  ratio as a function of residue number of wild-type yTrx1 (top) and its D24A mutant (bottom). An empty circle represents resonances absent from the HSQC spectrum due to line broadening, and “P” represents proline. Both yTrx1 and D24A were at 200  $\mu$ M and 298 K in 20 mM phosphate buffer (pH 7.0) with 30 mM perdeuterated DTT.

other residues with an  $R_{ex}$  of  $>0$  were in the interacting loops. For the D24A mutant, the motion of residue Ala24 was completely quenched ( $R_{ex} = 0$ ), and conformational exchange was observed for other residues facing the water cavity (Cys33 and Ile36). The quenching of Ala24 corroborates the closing of lobe B of the water cavity. The motion observed for Ile36 and Cys33 is very likely, because the Ile36 side chain is facing lobe C of the cavity and Cys33 is at the upper side of the cavity, near the interacting loops. The increase in the  $R_{ex}$  of Thr28 is noteworthy; it demonstrates the change of motion caused by the mutation in the upper region of the water cavity.

We also analyzed the  $R_2/R_1$  ratio for wild-type and mutant proteins (Figures 4C and 5C). An increased value of  $R_2/R_1$  was found in various residues, thus confirming their slow exchange behavior independent of any fitting or physical models.

Thermal flexibility, which is the fast motion that occurs on the pico- to nanosecond time scale, occurs in the same regions for both the oxidized and reduced states. Similar behavior was observed for the D24A mutant, with the same loops showing thermal flexibility compared with wild-type yTrx1. The most noticeable change was the small decrease in the order parameters of residues Cys33 and Met35 involved in the active site loop for oxidized D24A. The fast dynamics remain very similar to those of the wild-type yTrx1 protein and the D24N mutant,<sup>9</sup> as revealed by the unchanged order parameter. We found thermal flexibility for the N-terminal residues [e.g., Gln17 and Lys19 (which are the part of the  $\alpha_1-\beta_2$  loop), residue Ala71 ( $\alpha_3-\beta_4$  loop), and Ala90 and Asn91 ( $\beta_5-\alpha_4$  loop)]. The  $\alpha_3-\beta_4$  and  $\beta_5-\alpha_4$  loops are known to participate in interactions with other targets,<sup>39,40</sup> whereas the  $\alpha_1-\beta_2$  loop is on the opposite side of the interaction surface.



**Figure 5.** Dynamics based on the Lipari–Szabo model-free formalism of reduced yTrx1 and its D24A mutant. (A) Order parameter ( $S^2$ ) for reduced wild-type yTrx1 (top) and its D24A mutant. (B) Rate of residues in conformational exchange ( $R_{ex}$ ) as a function of residue number of wild-type yTrx1 (top) and its D24A mutant (bottom). (C)  $R_2/R_1$  ratio as a function of residue number of wild-type yTrx1 (top) and its D24A mutant (bottom). An empty circle represents resonances absent from the HSQC spectrum due to line broadening, and “P” represents proline. Both yTrx1 and D24A were at 200  $\mu$ M and 298 K in 20 mM phosphate buffer (pH 7.0) with 30 mM perdeuterated DTT.

Similar behavior was observed for yTrx1 in the reduced state. We observed motions in many residues involving the interacting loops (Cys30, Asp56, Asp89, Ser70, Ala71, and Asn91). Interestingly, Val63 is in conformational exchange exclusively with D24A. This residue holds  $\alpha$ -helix 3 ( $\alpha_3$ ) by interacting with the hydrophobic core at the opposite side of the  $\beta$ -sheet compared with the side of the water cavity.

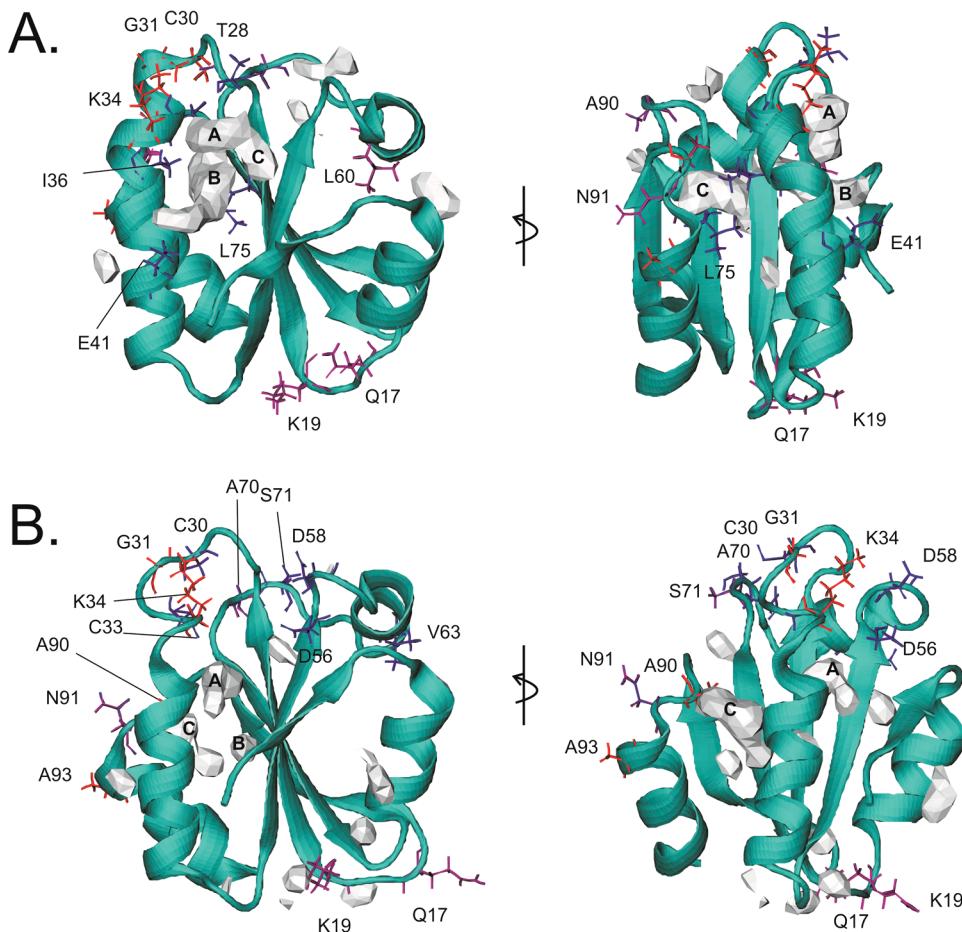
It is important to mention the strong correlation between the protein dynamics and the proximity of the residue to the protein cavities. Figure 6 shows the water cavity colored gray, using an isovalue of 2, which lowers the threshold for finding cavities. We detected the main water cavity, which was described above, and other small cavities that are more superficial but important for explaining the protein dynamics.

The dynamics of several residues could be explained by the proximity to the water cavity. For both the oxidized and reduced states, almost all of the residues in conformational exchange are near the main water cavity. This is especially clear for the oxidized state. A residue such as Glu41, which apparently is not as close to the cavity, is actually near an

extension of lobe B of the water cavity and can be observed only at an isovalue of 2 (Figure 6). The same happens for the reduced state; however, in the reduced state, there are residues in conformational exchange that are indirectly in contact with the water cavity, but always very close. This is true for some residues in the interaction loops.

Leu60 (oxidized) and Val63 (reduced) are far from the main cavity but near a small cavity below helix 3. The same is true for the residues in thermal dynamics Gln17 and Lys19, which are near a small cavity near this loop (Figure 6). Residues Ser71 and Asn91 are near the main cavity.

It is well-known from the analysis of the structures available that the behavior of  $\alpha$ -helix 2 is different in the oxidized and reduced states<sup>16,18,19,21</sup> (i.e., PDB entries 2HI9 and 3F3Q). The reduced state exhibits a comparatively shorter  $\alpha$ -helix 2 (ranging from 33 to 46 residues), whereas in the oxidized state,  $\alpha$ -helix 2 ranges from 30 to 46 residues. Figure 6 shows the residues in conformational exchange in the structure for reduced and oxidized states. Note that many residues in conformational exchange belong to  $\alpha$ -helix 2 and that the



**Figure 6.** Representation of the dynamics of D24A yTrx1. Panel A shows the results for the oxidized protein, and panel B shows the results for the reduced protein. A ribbon representation of the structure of D24A yTrx1 is colored cyan. The water cavity is represented by the gray volumes. The residues showing dynamics are shown as sticks colored blue for milli- to microsecond dynamics, red for the extreme broadened by intermediate exchange, and purple for pico- to nanosecond dynamics. The residues that undergo milli- to microsecond dynamics are Thr28, Cys33, Ile36, Glu41, Ser71, and Leu75 for the oxidized state and Cys30, Cys33, Asp56, Asp58, Val63, Ala70, and Asn91 for the reduced state. The residues in the broadened by intermediate exchange are Cys30, Gly31, Lys34, Ala90, and Ala93 for the oxidized protein and Gly31, Lys34, Ala90, and Ala93 for the reduced protein. The residues with thermal motion (pico- to nanosecond dynamics) are Gln17, Lys19, Ser71, and Asn91 for both reduced and oxidized proteins. An isovalue of 2 was used to compute the water cavity.

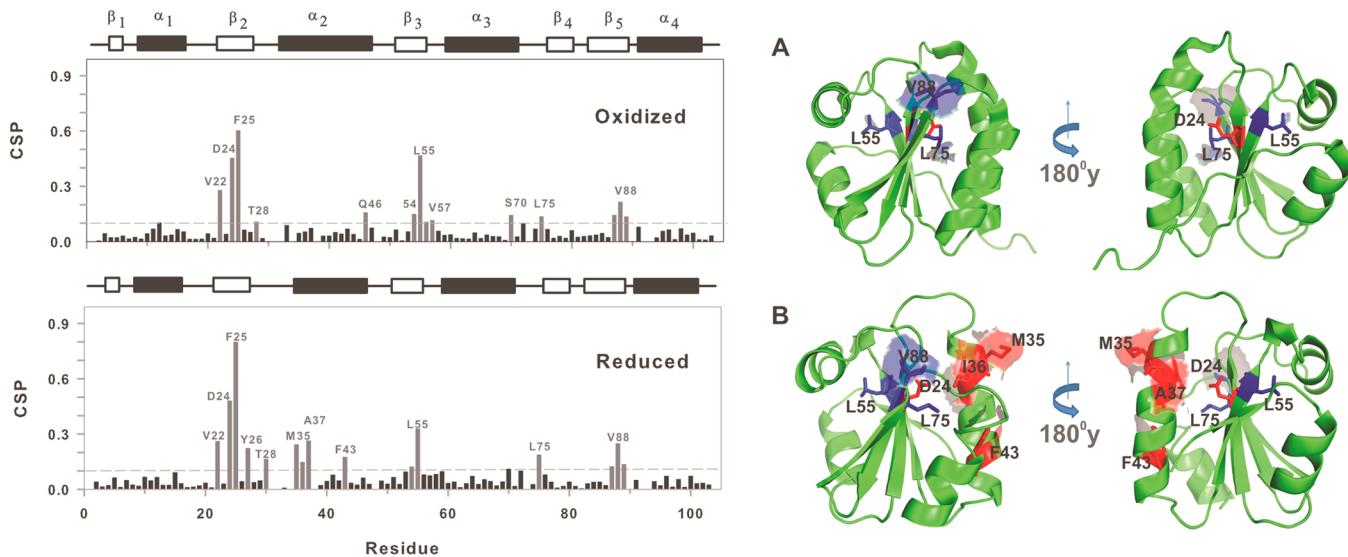
interacting loops always exhibit conformational exchange. Cys33 and Ile36 are part of  $\alpha$ -helix 2 for both reduced and oxidized states, whereas Cys30 is the N-terminal cap in the oxidized state. Asp24 is protected from direct contact of solvent with  $\alpha$ -helix 2, and Ile36 is directly facing Asp24.

We probed the effects of the mutation by examining the chemical shift perturbation (CSP) elicited by the D24A (Figure 7). CSP reflects the changes in the chemical environment observed for each nucleus upon mutation. It is a valuable tool for informing how the conformational change is transmitted along the protein chain, mostly through tertiary hydrophobic contacts. We considered changes near the mutation site (according to amino acid sequence) to be local effects of the mutation. Interestingly, nonlocal changes were observed mostly for hydrophobic residues. Residues Leu55 and Val88 showed high CSPs for both the reduced and oxidized states. These are key hydrophobic residues that are facing the opposite side of the central  $\beta$ -sheet relative to the side where the water cavity is located. This is the hydrophobic core, which is rearranged upon mutation. It is worth noting that Leu55 is adjacent to Lys54, which is the conserved positively charged residue that faces the water cavity side and is near Asp/Ala24. Another key residue is

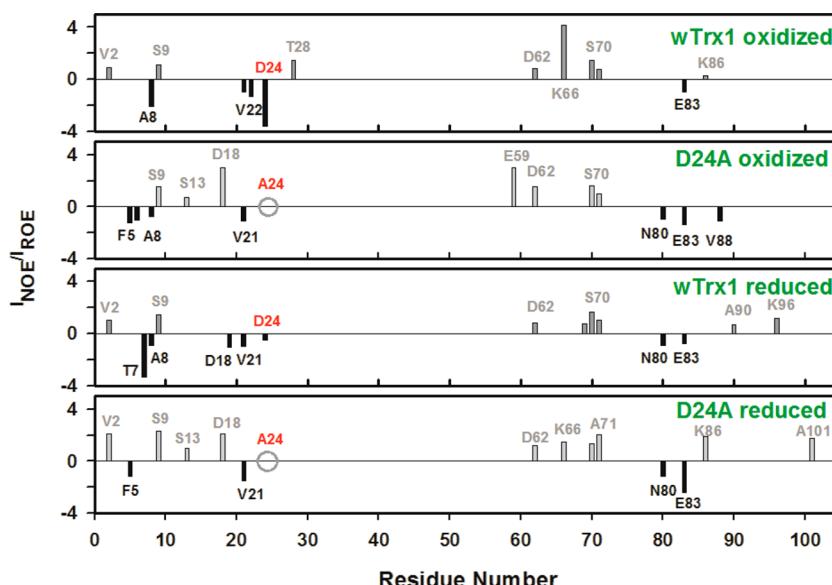
Leu75, which is on the same side of the water cavity. These observations are clear indications that a rearrangement of the hydrophobic packing was induced by the D24A mutation. The same was found for the D24N mutant.<sup>9</sup>

We observed changes that are exclusive to the reduced state. There are prominent CSPs at  $\alpha$ -helix 2, including residues Met35, Ile36, Ala37, and Phe43. Note again that these are the hydrophobic residues that are rearranged by the mutation. This helix is more sensitive to reorientation in the reduced state. These residues are most likely the hydrophobic interaction that serves as the scaffold of the water cavity. Leu75 has the same role.

To understand the effects of hydration on the water cavity, we directly measured the dipolar coupling between water and protein hydrogen atoms by NMR. The method, which was used for yTrx1 in a previous publication,<sup>9,41,42</sup> consisted of measuring nuclear Overhauser correlations in the laboratory frame (NOE) and in the rotating frame (ROE). A negative NOE/ROE intensity ratio reveals interactions between the water and the protein residue. Because nonlocal water molecules may also contribute to the negative ratio,<sup>43</sup> observation of tightly bound water molecules is not in itself



**Figure 7.** Chemical shift perturbation (CSP) induced by the D24A mutation in oxidized and reduced states. The plot shows CSP as a function of residue number. On the bottom, the ribbon and the surface representation of the yTrx1 structures demonstrate CSP changes by mutation, where panels A and B represent oxidized and reduced yTrx1, respectively. Note that many residues close to the mutation site showed chemical shift changes. The most apparent chemical shift changes upon mutation (CSP > 0.1) are shown above the dotted line and are colored red in structures A and B, based on PDB entries 2I9H and 3F3Q, respectively.



**Figure 8.** Plot of the NOE/ROE ratio as a function of residue number. The NOE corresponds to the laboratory frame, and the ROE corresponds to the rotating frame; both were measured with a mixing time of 20 ms. The nuclear spin dipolar interactions of water and HN are shown by negative values. The NMR data resemble those of the wild-type crystal structure of PDB entry 3F3Q, considering their hydration state.

conclusive, e.g., for Asp24. However, it is conclusive for thioredoxins at position 24 because it is well-known that a tightly bound water is present in the crystal structures for the reduced and oxidized state, hydrogen-bonded to the carboxyl side chain of Asp24.<sup>44–46</sup> For the reasons noted above, we limit our analysis to the solvation of this site. Figure 8 shows the absence of NOE/ROE connecting water and the Ala24 amide hydrogen, whereas a negative NOE/ROE ratio is present for wild-type yTrx1.<sup>9</sup> This result suggests either that there is an absence of water and the cavity is closed or that water is present at fast exchange. The micro- to millisecond dynamics observed for the residues facing the water cavity point to the second explanation because the cavity still exists for D24A. For the D24N mutant, we showed that there is an increase in the level

of exposure of Asn24 to water, fast exchange with the bulk water, and the reduced size of the cavity.<sup>9</sup>

The effect of the D26A mutation on *E. coli* Trx was previously reported, with a drastic drop of the disulfide reductase activity (~10-fold)<sup>7</sup> and the stabilization of the reduced state by 4.6 kcal/mol.<sup>12</sup> To evaluate if the same is true for yTrx1, we measured its activity and stabilization and the effect of the D24A mutation.

For the disulfide reductase yTrx1 activity, we used two substrates, insulin and L-cystine (Figure S5 of the Supporting Information). As observed for *E. coli* Trx, we found that the activity toward insulin is higher than the activity toward L-cystine by a factor of 1.7. The mutation led to an ~15-fold

decrease toward insulin as a substrate and ~5-fold toward L-cystine as a substrate.

For the stabilization, we used FoldX,<sup>47</sup> which can predict with good reliability the protein stability. The FoldX stabilization free energy ( $\Delta\Delta G_u = \Delta G_u^{\text{mutant}} - \Delta G_u^{\text{Trx1}}$ ) is presented in Table S1 of the Supporting Information. The D24A mutant in the reduced state stabilized the yTrx1 by 2.6 kcal/mol. We also observed stabilization in the oxidized state of D24A and for the D24N mutant in the range of 0.6–1 kcal/mol.

## ■ DISCUSSION

The concepts of minimal frustration and energy landscape allow an understanding of the energetics of protein folding, in the presence of few folding intermediates. The principle of minimal frustration states that the evolution shaped the amino acid content of proteins and its interactions to favor the native structure, minimizing non-native contact that leads to kinetic traps. The energy landscape is smoother than would be expected by chance, biasing the protein folding toward the native state and just a few high-energy states.

Some degree of frustration remains, as can be observed in the presence of local minima in the energy landscape of proteins (intermediate and excited states).<sup>48,49</sup> Recently, a correlation between protein dynamics and frustration level has been found.<sup>50</sup> Regions involved in protein dynamics display higher frustration levels, particularly those in conformational exchange, displaying dynamics on the micro- to millisecond time scales.<sup>50</sup> Excited conformational states were also favored by evolution in the shaping of the energy landscape. High-energy states that are useful for biological function were conserved by evolution. In fact, allosteric and catalytic binding sites can have higher degrees of local frustration.<sup>51</sup> For thioredoxins, the water cavity must also be conserved by evolution and the high degree of frustration of the residues at the cavity might be reflected by its protein dynamics. Here we showed that the dynamics of almost all residues could be explained by the localization within or in the proximity of the water cavity (Figure 6). The absence of internal stable contacts between the residues facing a cavity can produce an increased level of frustration, and thus, internal hydration can also induce increased levels of frustration because of the exchange between the internal water and bulk water.

Frustration levels can be calculated by analyzing how mutations and conformational changes shift the energetics of a protein. We used the Frustratrometer<sup>52</sup> (<http://www.frustratometer.tk/>), a tool to localize the level of frustration for each interaction. The highly mutational frustrated interaction is depicted in Figure S6 of the Supporting Information. The mutational frustration level consists of calculation of the energy of each interaction in the native protein and comparison of it with the statistics of energies calculated for mutations at each native location. If mutations consistently destabilize a native interaction, then this interaction is minimally frustrated. On the other hand, highly frustrated interactions are stabilized by changing the native location through mutations. Figure S6 of the Supporting Information shows that the region of the water cavity of yTrx1 is the most mutationally frustrated region in the protein, for both the reduced and oxidized states. We observed a better correlation between the dynamics and mutational frustration levels than between configurational frustration levels.<sup>52</sup>

We observed a good correlation between the residues that are undergoing protein dynamics (on multiple time scales) and

highly frustrated contacts. Figure S6 of the Supporting Information shows that both the number of highly frustrated residues and the counts of frustrated interactions per residue decrease considerably upon introduction of the D24A mutation. For reduced yTrx1, we observed residues Asp24, Tyr26, Asp51, Lys54, Asp56, and Pro73 with more than four counts per residue. Among these, all but Asp51 are located in the water cavity. For oxidized yTrx1, we observed residues Asp24, Lys54, and Pro73 with more than four counts per residue. All of them are located in the water cavity. It is also important to mention that for the D24A mutant there is no residue showing more than four counts per residue, corroborating the partial closing of the cavity.

In all situations studied here, all the residues undergoing protein dynamics on multiple time scales are involved directly in a highly frustrated contact or in the proximity. All residues of reduced yTrx1 that undergo protein dynamics are involved directly in or in the proximity of frustrated contacts, such as Gln17, Lys19, Cys30, Gly31, Cys33, Lys34, Ile36, Asp58, Glu59, Ser71, Ala90, Asn91, and Ala93. The same is true with oxidized yTrx1, involving residues Gln17, Lys19, Asp24, Thr28, Cys30, Gly31, Lys34, Ile36, Asp58, Glu59, Ser71, Ala90, and Ala93.

In spite of the good correlation, we also observed that there are several residues participating in mutationally frustrated contacts that are not undergoing protein dynamics. We believe that to make a more accurate correlation of protein dynamics and frustration level it is necessary to analyze an ensemble of structures, possibly generated by MD simulation, rather than a unique structure as presented here.

yTrx1 shows dynamics on multiple time scales. The residues that undergo milli- to microsecond dynamics are Thr28, Cys33, Ile36, Glu41, Ser71, and Leu75 for the oxidized state and Cys30, Cys33, Asp56, Asp58, Val63, Ala70, and Asn91 for the reduced state. The residues broadened by intermediate exchange are Cys30, Gly31, Lys34, Ala90, and Ala93 for the oxidized protein and Gly31, Lys34, Ala90, and Ala93 for the reduced protein. The residues with thermal motion (pico- to nanosecond dynamics) are Gln17, Lys19, Ser71, and Asn91 for both reduced and oxidized protein.

The MD simulation corroborates the existence of a water cavity in the presence of the Asp to Ala mutation. The detection of resident water and the proximity to the distorted cavity were mapped. It is quite evident now that the water cavity couples the slow motions of the buried hydrophobic core residues and other consecutive loops. The decrease in the number of hydrogen bonds and solvent exposure and the experimental dynamics support the description of the water cavity presented here.

Evolution plays with a multitude of forces.<sup>53</sup> This is the rule for understanding the formation of the water cavity in thioredoxins. We demonstrated that the water cavity of yTrx1 is a result of the increased frustration level near the catalytic site of thioredoxins. Several conserved residues act in the formation of the water cavity. These included residues located at  $\alpha$ -helices 2 and 4, the central  $\beta$ -sheet, and the interacting loops. Of particular importance are several residues and factors. (i) The first is Asp24, for which a high solvation free energy increases the rate of hydration of lobe B. Its solvation free energy is controlled by the state of ionization. (ii) The second is Lys54, at the surface of the hydrophilic lobe A, which confers hydrophilicity for lobe A of the cavity. It is at the cavity-solvation shell interface, being probably very important for the

water and consequently proton exchange. (iii) The third is *cis*-Pro73, the most hydrophilic residue in the hydrophobic lobe C. The *cis* conformation warrants that the side chain face the cavity. The hydrophilicity of this *cis*-Pro73 is probably important for water and proton exchange. (iv) The fourth is the presence of a cage of prolines surrounding the water cavity (Pro32, Pro38, and *cis*-Pro73). The prolines provide rigidity and at the same time the hydrophilicity necessary for the formation of the water cavity. (v) The fifth is hydrophobic mismatch that must be conserved in evolution. The existence of lobes A and C even when lobe B is closed by the D24A mutation is very relevant, because it reveals that the three lobes are independent and that there are multiple paths for water exchange. (vi) Helices 2 and 4 are key secondary structure elements for the water cavity.<sup>54</sup> They contain key residues in conformational exchange.

For thioredoxins, the opening–closing equilibrium of the water cavity changes the overall energy of the system. This work provides insights into the mechanism of coupling the internal hydration and conformation exchange (free energy) in thioredoxins and the residues that are participating in that process. The conformational equilibrium probably involves several degrees of internal hydration, with the less hydrated conformations (more closed) being the more stable. For *E. coli* thioredoxin, the protonation of Asp26 stabilizes the protein ( $\Delta\Delta G^\circ = -2$  kcal/mol). The D26A mutation stabilizes as well with a  $\Delta\Delta G^\circ$  of  $-4.6$  kcal/mol.<sup>12</sup> The FoldX free energy was calculated for yTrx1 and mutants D24A and D24N (Table S1 of the Supporting Information). Both mutants perturbed the cavity and are more stable than the wild type.

Thioredoxin catalytic activity demands that internal water molecules mediate proton transfer and molecular recognition of the target protein. The water cavity controls thioredoxin catalysis, pH dependence, and molecular recognition of target proteins. Our work suggested multiple ways for water and proton to exchange in yTrx1. Binding to cellular targets may block in different manners the path of water to the cavity and consequently induces less hydrated closed conformations. The coupling of the internal hydration and conformation exchange in thioredoxins provides a fine-tuning of the frustration level and consequently of the energy level.

## ■ ASSOCIATED CONTENT

### Supporting Information

Several supporting figures and a table. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Av. Carlos Chagas Filho, 373, CCS, Annex CNRMN, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil 21941-902. E-mail: falmeida@cnrmn.bioqmed.ufrj.br. Phone: +5521-31042326.

### Author Contributions

A.I. and F.G.-N. contributed equally to this work.

### Funding

We thank the following funding agencies for funding: FAPERJ, CNPq, CAPES, and the program Science Without Borders-CNPq.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We are thankful to Adolfo Moraes, Viviane de Paula, and Anderson Pinheiro for carefully reading the manuscript and presenting their suggestions. We are also thankful to the National Center of Nuclear Magnetic Resonance (CNRMN) for the NMR time. We are thankful to Dr. Ernesto Roman for the fruitful discussion on protein frustration.

## ■ REFERENCES

- (1) Buchanan, B. B., Holmgren, A., Jacquot, J.-P., and Scheibe, R. (2012) Fifty years in the thioredoxin field and a bountiful harvest. *Biochim. Biophys. Acta* 1820, 1822–1829.
- (2) Arnér, E. S., and Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.* 267, 6102–6109.
- (3) Oliveira, M. A., Discola, K. F., Alves, S. V., Medrano, F. J., Guimarães, B. G., and Netto, L. E. S. (2010) Insights into the specificity of thioredoxin reductase-thioredoxin interactions. A structural and functional investigation of the yeast thioredoxin system. *Biochemistry* 49, 3317–3326.
- (4) Potamitou, A., Holmgren, A., and Vlamis-Gardikas, A. (2002) Protein levels of *Escherichia coli* thioredoxins and glutaredoxins and their relation to null mutants, growth phase, and function. *J. Biol. Chem.* 277, 18561–18567.
- (5) Ingles-Prieto, A., Ibarra-Molero, B., Delgado-Delgado, A., Perez-Jimenez, R., Fernandez, J. M., Gaucher, E. A., Sanchez-Ruiz, J. M., and Gavira, J. A. (2013) Conservation of Protein Structure over Four Billion Years. *Structure* 21, 1–8.
- (6) LeMaster, D. M., Springer, P. A., and Unkefer, C. J. (1997) The role of the buried aspartate of *Escherichia coli* thioredoxin in the activation of the mixed disulfide intermediate. *J. Biol. Chem.* 272, 29998–30001.
- (7) Dyson, H. J., Jeng, M. F., Tenant, L. L., Slaby, I., Lindell, M., Cui, D. S., Kuprin, S., and Holmgren, A. (1997) Effects of buried charged groups on cysteine thiol ionization and reactivity in *Escherichia coli* thioredoxin: Structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry* 36, 2622–2636.
- (8) Ji, C., Mei, Y., and Zhang, J. Z. H. (2008) Developing polarized protein-specific charges for protein dynamics: MD free energy calculation of pK<sub>a</sub> shifts for Asp26/Asp20 in thioredoxin. *Biophys. J.* 95, 1080–1088.
- (9) Cruzeiro-Silva, C., Gomes-Neto, F., Machado, L. E. S. F., Miyamoto, C. A., Pinheiro, A. S., Correa-Pereira, N., de Magalhães, M. T. Q., Valente, A. P., and Almeida, F. C. L. (2014) Hydration and conformational equilibrium in yeast thioredoxin 1: Implication for H<sup>+</sup> exchange. *Biochemistry* 53, 2890–2902.
- (10) Wilson, N. A., Barbar, E., Fuchs, J. A., and Woodward, C. (1995) Aspartic Acid 26 in Reduced *Escherichia coli* Thioredoxin Has a pK<sub>a</sub> >9. *Biochemistry* 34, 8931–8939.
- (11) Jeng, M., and Dyson, H. J. (1996) Direct Measurement of the Aspartic Acid 26 pK<sub>a</sub> for Reduced *Escherichia coli*. *Biochemistry* 35, 1–6.
- (12) Langsetmo, K., Fuchs, J. A., and Woodward, C. (1991) The conserved, buried aspartic acid in oxidized *Escherichia coli* thioredoxin has a pK<sub>a</sub> of 7.5. Its titration produces a related shift in global stability. *Biochemistry* 30, 7603–7609.
- (13) Setterdahl, A. T., Chivers, P. T., Hirasawa, M., Lemaire, S. D., Keryer, E., Miginiac-Maslow, M., Kim, S.-K., Mason, J., Jacquot, J.-P., Longbine, C. C., de Lamotte-Guery, F., and Knaff, D. B. (2003) Effect of pH on the oxidation-reduction properties of thioredoxins. *Biochemistry* 42, 14877–14884.
- (14) Pinheiro, A. S., Amorim, G. C., Netto, L. E. S., Valente, A. P., and Almeida, F. C. L. (2006) <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N resonance assignments for the reduced forms of thioredoxin 1 and 2 from *S. cerevisiae*. *J. Biomol. NMR* 36 (Suppl. 1), 35.
- (15) Pinheiro, A. S., Amorim, G. C., Netto, L. E. S., Almeida, F. C. L., and Valente, A. P. (2008) NMR solution structure of the reduced form

- of thioredoxin 1 from *Saccharomyces cerevisiae*. *Proteins: Struct., Funct., Bioinf.* 70, 584–587.
- (16) Amorim, G. C., Pinheiro, A. S., Netto, L. E. S., Valente, A. P., and Almeida, F. C. L. (2007) NMR solution structure of the reduced form of thioredoxin 2 from *Saccharomyces cerevisiae*. *J. Biomol. NMR* 38, 99–104.
- (17) Zhang, Y., Bao, R., Zhou, C.-Z., and Chen, Y. (2008) Expression, purification, crystallization and preliminary X-ray diffraction analysis of thioredoxin Trx1 from *Saccharomyces cerevisiae*. *Acta Crystallogr. F* 64, 323–325.
- (18) Bao, R., Zhang, Y., Lou, X., Zhou, C.-Z., and Chen, Y. (2009) Structural and kinetic analysis of *Saccharomyces cerevisiae* thioredoxin Trx1: Implications for the catalytic mechanism of GSSG reduced by the thioredoxin system. *Biochim. Biophys. Acta* 1794, 1218–1223.
- (19) Bao, R., Chen, Y., Tang, Y., Janin, J., and Zhou, C. (2007) Crystal Structure of the Yeast Cytoplasmic Thioredoxin Trx2. *Bioinformatics* 249, 246–249.
- (20) Su, D., Berndt, C., Fomenko, D. E., Holmgren, A., and Gladyshev, V. N. (2007) A conserved cis-proline precludes metal binding by the active site thiolates in members of the thioredoxin family of proteins. *Biochemistry* 46, 6903–6910.
- (21) Pinheiro, A. S., Amorim, G. C., Eduardo, L., Netto, S., Almeida, F. C. L., and Valente, A. P. (2007) NMR solution structure of the reduced form of thioredoxin 1 from *Saccharomyces cerevisiae*. *Proteins* 70, 584–587.
- (22) Dixit, S. B., Bhasin, R., Rajasekaran, E., and Jayaram, B. (1997) Solvation thermodynamics of amino acids. Assessment of the electrostatic contribution and force-field dependence. *J. Chem. Soc., Faraday Trans.* 93, 1105–1113.
- (23) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* 59, 687–696.
- (24) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* 104, 4546–4559.
- (25) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *J. Am. Chem. Soc.* 104, 4559–4570.
- (26) Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* 16, 23–28.
- (27) Berendsen, H. J. C., Vandervelde, D., and Vandrunen, R. (1995) Gromacs: A Message-Passing Parallel Molecular-Dynamics Implementation. *Comput. Phys. Commun.* 91, 43–56.
- (28) Li, H., Robertson, A. D., and Jensen, J. H. (2005) Very fast empirical prediction and rationalization of protein pK<sub>a</sub> values. *Proteins: Struct., Funct., Bioinf.* 61, 704–721.
- (29) Schwieters, C. D., and Clore, G. M. (2001) The VMD-XPLOR visualization package for NMR structure refinement. *J. Magn. Reson.* 149, 239–244.
- (30) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65, 712–725.
- (31) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102, 3586–3616.
- (32) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18, 1463–1472.
- (33) Miyamoto, S., and Kollman, P. A. (1992) Settle: An Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* 13, 952–962.
- (34) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 103, 8577–8593.
- (35) Jorgensen, W. L., and Tiradore, J. (1988) The Opls Potential Functions for Proteins: Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* 110, 1657–1666.
- (36) Luzar, A. (2000) Resolving the hydrogen bond dynamics conundrum. *J. Chem. Phys.* 113, 10663–10675.
- (37) Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009) Fpocket: An open source platform for ligand pocket detection. *BMC Bioinf.* 10, 168–177.
- (38) Valente, A. P., Miyamoto, C. A., and Almeida, F. C. L. (2006) Implications of protein conformational diversity for binding and development of new biological active compounds. *Curr. Med. Chem.* 13, 3697–3703.
- (39) Cheng, Z., Zhang, J., Ballou, D. P., and Williams, C. H. (2011) Reactivity of thioredoxin as a protein thiol-disulfide oxidoreductase. *Chem. Rev.* 111, 5768–5783.
- (40) Holmgren, A., and Lu, J. (2010) Thioredoxin and thioredoxin reductase: Current research with special reference to human disease. *Biochem. Biophys. Res. Commun.* 396, 120–124.
- (41) Otting, G., Liepinsh, E., and Wüthrich, K. (1991) Protein hydration in aqueous solution. *Science* 254, 974–980.
- (42) Otting, G., Liepinsh, E., Farmer, B. T., and Wüthrich, K. (1991) Protein hydration studied with homonuclear 3D <sup>1</sup>H NMR experiments. *J. Biomol. NMR* 1, 209–215.
- (43) Modig, K., Liepinsh, E., Otting, G., and Halle, B. (2004) Dynamics of protein and peptide hydration. *J. Am. Chem. Soc.* 126, 102–114.
- (44) Weichsel, A., Kem, M., and Montfort, W. R. (2010) Crystal structure of human thioredoxin revealing an unraveled helix and exposed S-nitrosation site. *Protein Sci.* 19, 1801–1806.
- (45) Campos-Acevedo, A. A., Garcia-Orozco, K. D., Sotelo-Mundo, R. R., and Rudiño-Piñera, E. (2013) Expression, purification, crystallization and X-ray crystallographic studies of different redox states of the active site of thioredoxin 1 from the whiteleg shrimp *Litopenaeus vannamei*. *Acta Crystallogr. F* 69, 488–493.
- (46) Bao, R., Zhang, Y., Zhou, C.-Z., and Chen, Y. (2009) Structural and mechanistic analyses of yeast mitochondrial thioredoxin Trx3 reveal putative function of its additional cysteine residues. *Biochim. Biophys. Acta* 1794, 716–721.
- (47) Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* 320, 369–87.
- (48) Leopold, P. E., Montal, M., and Onuchic, J. N. (1992) Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. U.S.A.* 89, 8721–8725.
- (49) Bryngelson, J. D., Onuchic, J. N., Soccia, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* 21, 167–195.
- (50) Whittaker, S. B.-M., Clayden, N. J., and Moore, G. R. (2011) NMR characterisation of the relationship between frustration and the excited state of Im7. *J. Mol. Biol.* 414, 511–529.
- (51) Ferreiro, D. U., Komives, E. A., and Wolynes, P. G. (2014) Frustration in biomolecules. *Q. Rev. Biophys.*, 1–79.
- (52) Jenik, M., Parra, R. G., Radusky, L. G., Turjanski, A., Wolynes, P. G., and Ferreiro, D. U. (2012) Protein frustratometer: A tool to localize energetic frustration in protein molecules. *Nucleic Acids Res.* 40, 348–351.
- (53) Gianni, S., Camilloni, C., Giri, R., Toto, A., Bonetti, D., Morrone, A., Sormanni, P., Brunori, M., and Vendruscolo, M. (2014) Understanding the frustration arising from the competition between function, misfolding, and aggregation in a globular protein. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14141–14146.

- (54) Vazquez, D. S., Sánchez, I. E., Garrote, A., Sica, M. P., and Santos, J. (2015) The *E. coli* thioredoxin folding mechanism: The key role of the C-terminal helix. *Biochim. Biophys. Acta* 1854, 127–137.