BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Mapping the Active Site Helix-to-Strand Conversion of CxxxxC Peroxiredoxin Q Enzymes

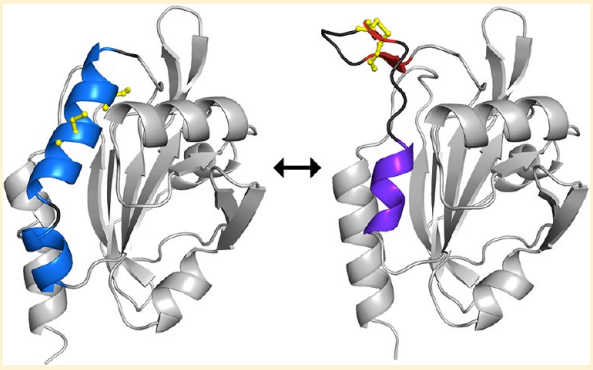
Arden Perkins,[†] Michael C. Gretes,[†] Kimberly J. Nelson,[‡] Leslie B. Poole,[‡] and P. Andrew Karplus*,[†]

[†]Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331, United States

[‡]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, United States

Supporting Information

ABSTRACT: Peroxiredoxins (Prx) make up a family of enzymes that reduce peroxides using a peroxidatic cysteine residue; among these, members of the PrxQ subfamily are proposed to be the most ancestral-like yet are among the least characterized. In many PrxQ enzymes, a second "resolving" cysteine is located five residues downstream from the peroxidatic Cys, and these residues form a disulfide during the catalytic cycle. Here, we describe three hyperthermophilic PrxQ crystal structures originally determined by the RIKEN structural genomics group. We reprocessed the diffraction data and conducted further refinement to yield models with R_{free} values lowered by 2.3−7.2% and resolution extended by 0.2−0.3 Å, making one, at 1.4 Å, one of the best resolved peroxiredoxins to date. Comparisons of two matched thiol and disulfide forms reveal that the active site conformational change required for disulfide formation involves a transition of ∼20 residues from a pair of α-helices to a β-hairpin and 3_{10}-helix. Each conformation has ∼10 residues with a high level of disorder providing slack that allows the dramatic shift, and the two conformations are anchored to the protein core by distinct nonpolar side chains that fill three hydrophobic pockets. Sequence conservation patterns confirm the importance of these and a few additional residues for function. From a broader perspective, this study raises the provocative question of how to make use of the valuable information in the Protein Data Bank generated by structural genomics projects but not described in the literature, perhaps remaining unrecognized and certainly underutilized.



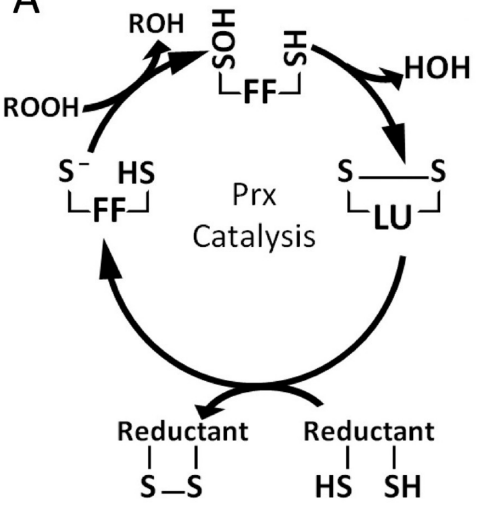
Aerobic organisms must cope with reactive oxygen species (ROS) that can damage DNA, proteins, and lipids, potentially causing loss of membrane integrity and cell death.[1−4] Meeting this challenge are a variety of antioxidant defenses, including enzymes of the ubiquitous and highly expressed peroxiredoxin (Prx) family, which specialize in reducing hydrogen peroxide and organic peroxides and can also reduce peroxynitrite.[5−8] Studies of Prxs have provided insight into their catalytic mechanism[9] and their critical functions within the cell, including possible roles in eukaryotes in cell signaling and tumor suppression.[10,11] Prxs are also potential drug targets,[5] as for instance, it has been demonstrated that Prx knockout strains of the causative agent of malaria, *Plasmodium falciparum*, cope poorly with oxidative stress.[12,13]

Prxs are divided by sequence similarity into six subfamilies, Prx1, Prx6, Prx5, Tpx, AhpE, and BCP-PrxQ,[1,14] all of which share a globular tertiary structure, with a conserved core of five α-helices and seven β-strands.[1,7] Their enzymatic activity requires an active site peroxidatic Cys residue (designated as C_P or S_PH for the thiol), located within a strictly conserved PXXX(T/S)XXC_P motif.[14] The catalytic cycle involves three chemical steps (Figure 1A). First, the S_P^- thiolate attacks a peroxide substrate in an S_N2 reaction to form sulfenic acid ($R\text{-}S_POH$).[9] Second, to recycle the C_P, and to prevent over-oxidation from a second peroxide substrate, the subset called 2-Cys Prxs possess a second Cys residue (designated C_R for resolving) that reacts with the sulfenic acid to form a disulfide bond.[15] The disulfide is then reduced by an external agent, such as thioredoxin, to regenerate the substrate-ready form.[1,5,9,15]

Importantly, disulfide formation in all Prxs requires a local unfolding of the active site. The substrate-ready conformation with the C_P thiolate at the bottom of an active site pocket is called fully folded (FF), and the disulfide form is called locally unfolded (LU).[1] As reviewed by Hall et al.,[1] each type of 2-Cys Prx has unique structural features that effectively stabilize discrete FF and LU conformations, and these have been described for the Prx1, Tpx, and Prx5 subfamilies as well as the BCP-PrxQ subgroup that has the C_R residue in helix α3. The one 2-Cys Prx subgroup not yet characterized in this way is the BCP-PrxQ subgroup that has C_R in helix α2, with just four residues separating it and the C_P. It is this common but little characterized[1,5,16] subgroup that is the subject of this work.

Received: July 28, 2012
Revised: August 24, 2012
Published: August 28, 2012

A



B

Type	Number	Structure	PDB	Form	Ref.
$C_R^{\alpha 2}$	511	*Ap*PrxQ	2CX3	LU	SG
		*Ap*PrxQ	2CX4	LU/FF	SG
		*St*PrxQ	2YWN	LU	SG
		*Sc*PrxQ	2A4V	FF	*(20)*
		*Ss*PrxQ4	3HJP	FF	*(18)*
		*Ss*PrxQ1	3DRN	FF	*(19)*
$C_R^{\alpha 3}$	74	*Xf*PrxQ	3IXR	FF	*(33)*
		*Xc*PrxQ	3GKM	FF	*(30)*
		*Xc*PrxQ	3GKK	LU	*(30)*
		*Xc*PrxQ	3GKN	LU/FF	*(30)*
C_R^{0}	387	None	--	--	--

Figure 1. Prx catalytic cycle and structural knowledge of PrxQ enzymes. (A) Catalytic cycle of 2-Cys Prxs indicating the protein conformation (FF or LU) dominant at each redox state. (B) Listed for each PrxQ subtype are the number of sequences identified in a 2008 database search[14,42] and the crystal structures determined and conformations seen. Gray denotes the structures in this study, with "SG" standing for structural genomics. The structures include *Ap*PrxQ (*Aeropyrum pernix* PrxQ), *St*PrxQ (*Sulfolobus tokodaii* PrxQ), *Sc*PrxQ (*Saccharomyces cerevisiae* PrxQ), *Ss*PrxQ4 (*Sulfolobus solfataricus* PrxQ4), *Ss*PrxQ1 (*S. solfataricus* PrxQ1), *Xf*PrxQ (*Xylella fastidiosa* PrxQ), and *Xc*PrxQ (*Xanthomonas campestris* PrxQ).

Table 1. Data Collection and Refinement Statistics for New and Previously Deposited Models

Data Collection	*St*PrxQ-LU[a]	*Ap*PrxQ-FF/LU[b]	*Ap*PrxQ-LU[c]
space group	$C222_1$	$P4_122$	$P6_422$
unit cell *a*, *b*, *c* (Å)	69.34, 78.65, 61.96	132.39, 132.39, 106.53	127.00, 127.00, 104.87
resolution (Å)	50.00−1.40 (1.42−1.40)[d]	50.00−2.00 (2.05−2.00)[e]	50.00−2.30 (2.34−2.30)[f]
completeness	99.6 (42.0)	95.5 (93.5)	99.8 (99.5)
no. of unique reflections	29551	63966	22182
multiplicity	5.8 (2.2)	12.1 (6.7)	13.0 (8.1)
$\langle I/\sigma \rangle$	29.4 (1.8)	38.0 (0.6)	20.7 (0.4)
$CC_{1/2}$	0.999 (0.809)	0.999 (0.236)	1.000 (0.162)
Refinement	**2YWN/polished**	**2CX4/polished[g]**	**2CX3/polished[h]**
resolution (Å)	1.6/1.4	2.3/2.0	2.6/2.3
R (%)	20.0/12.0	20.3/19.5	20.6/18.6
R_{free} (%)	22.1/14.9	26.3/23.4	24.7/22.4
no. of molecules in the asymmetric unit	1/1	8/4	4/2
no. of protein residues	150/151	1284/642	641/321
no. of water molecules	155/305	386/272	196/69
total no. of non-hydrogen atoms	1361/1573	10743/5694	5364/2714
rmsd for lengths (Å)	0.005/0.008	0.007/0.01	0.009/0.010
rmsd for angles (deg)	1.2/1.1	1.1/1.1	1.2/1.15
Ramachandran plot (%)			
preferred	98.6/99.3	84.7/96.7	89.1/94.3
allowed	1.4/0.7	11.7/3.1	9.5/4.8
outliers	0.0/0.0	3.6/0.2[i]	1.4/1.0[j]
$\langle B \rangle$ factor (Å^2)[k]	iso/aniso	iso/TLS	iso/TLS
main chain	14/15	49/55	65/77
side chains and waters	19/21	49/66	65/84
new PDB entry	−/4G2E	−/4GQC	−/4GQF

[a]Crystallization conditions as reported in the PDB header: 0.2 M ammonium acetate, 0.1 M HEPES, 25% PEG 3350, pH 7.5, vapor diffusion, sitting drop, 293 K. [b]Crystallization conditions as reported in the PDB header: ammonium sulfate, MES, TRIS, sodium chloride, dithiothreitol, pH 7, vapor diffusion, sitting drop, 285 K. [c]Crystallization conditions as reported in the PDB header: ammonium sulfate, 2-propanol, TRIS sodium chloride, dithiothreitol, pH 8, vapor diffusion, hanging drop, 293 K. [d]Resolution cutoff previously 1.6 Å with R_{meas} in the highest-resolution bin being 12%; R_{meas} at 1.4 Å is 51%. [e]Resolution cutoff previously 2.30 Å with R_{meas} being ∼120%; R_{meas} at 2.0 Å is 350%. [f]Resolution cutoff previously 2.60 Å with R_{meas} being ∼160%; R_{meas} at 2.3 Å is 480%. [g]The polished version also contained 30 sulfate molecules. [h]The polished version also contained nine sulfate molecules. [i]Gly3, Thr50, and Glu52 in chain A that all have weak density. [j]Ala91 in chain D with φ and ψ values of −73.4° and 26.7°, respectively, is very close to an allowed region. [k]Individual atomic *B* factors were refined anisotropically (aniso), isotropically (iso), or isotropically plus using TLS with one group per monomer (TLS).

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

Because the name BCP (bacterioferritin comigratory protein) is a historical holdover based on gel migration, we favor simply using the more informative PrxQ name for this subfamily and its members. This both clearly identifies them as Prxs and matches our recently proposed nomenclature for Prxs from parasites.[17] Here, we will use PrxQ, even for proteins that have been previously termed "BCP". For instance, *Escherichia coli* BCP will be termed *E. coli* PrxQ.

The PrxQ group includes monomers and dimers[18] and has been proposed to represent the subfamily most like the ancestral Prx.[1] Among PrxQs, the C_R position may vary, being absent (~38%) or being located in either α2 (~55%) or α3 (~7%).[1] These subtypes are here designated C_R^0, $C_R^{\alpha 2}$, and $C_R^{\alpha 3}$. *E.* coli PrxQ (a $C_R^{\alpha 2}$ enzyme) is the founding and best studied member of the subfamily[16] but has yet to be crystallized. To date, seven PrxQ structures have been reported in the literature (Figure 1B), including both FF and LU conformations of a $C_R^{\alpha 3}$-type PrxQ, but only the FF conformation of a $C_R^{\alpha 2}$-type PrxQ. Among these structures, only one, *Sulfolobus solfataricus* PrxQ4 (*Ss*PrxQ4), is a dimer.[18]

Fortuitously, although the $C_R^{\alpha 2}$ LU conformation is missing among the described structures, it actually has been determined in three crystal structures of dimeric PrxQs from the hyperthermophilic aerobes *Sulfolobus tokodaii* (*St*PrxQ) and *Aeropyrum pernix* (*Ap*PrxQ). The sequence of *St*PrxQ is 54% identical with that of *Ap*PrxQ and 78% identical with that of *Ss*PrxQ4, the previously described $C_R^{\alpha 2}$ PrxQ in the FF conformation. The three structures, one for *St*PrxQ and two for *Ap*PrxQ, were deposited more than five years ago by the RIKEN Structural Genomics Group, and have been mentioned in a few papers,[1,4,16,18−20] but have never been carefully interpreted and described in a primary publication.

Here, we report such an analysis of the *St*PrxQ and *Ap*PrxQ structures. With support from scientists at RIKEN, we have reprocessed the diffraction data and improved the previously deposited structures by additional rounds of polishing refinement. We provide a description of these structures and a detailed analysis of the conformational changes associated with local unfolding in the $C_R^{\alpha 2}$ PrxQ enzymes.

EXPERIMENTAL PROCEDURES

Polishing Refinements of Protein Data Bank (PDB) Entries 2YWN, 2CX3, and 2CX4. Original diffraction data for the structures of *St*PrxQ (PDB entry 2YWN), *Ap*PrxQ-LU (PDB entry 2CX3), and *Ap*PrxQ-FF/LU (PDB entry 2CX4) were obtained from the RIKEN Structural Genomics/Proteomics Initiative and processed and scaled using XDS.[21] The resolution cutoffs were based on the new criteria of $CC_{1/2} > 0$ with a *P* value of <0.001,[22] except that for the *St*PrxQ data set the high-resolution cutoff of 1.4 Å was limited by completeness (dropping below 50%) rather than signal strength (Table 1). We found using Pointless[23] that the data for 2CX4 could be processed in space group $P4_122$ and showed no evidence of twinning, contrasting with the previous assignment (according to the PDB file header) of space group $P4_1$ with a twinning fraction of 0.5, and twin operator *k*, *h*, −*l*. Similarly, the data for 2CX3, previously handled as space group $P6_4$, processed well in space group $P6_422$. No change was made to the space group of the data of 2YWN. To generate initial electron density maps for the structures with new space groups, molecular replacement was conducted with Phaser[24] using as search models four and two chains from the deposited 2CX4 and 2CX3 models, respectively.

Refinements were conducted using BUSTER[25] and/or PHENIX[26] without the use of noncrystallographic symmetry restraints, and with manual rebuilding conducted in Coot.[27] Ordered waters were added manually and automatically to difference map peaks of >3.3ρ_{rms} also having >1.0ρ_{rms} density in the $2F_o - F_c$ map and at least one hydrogen bonding neighbor within 2.4−3.5 Å. In the final models, waters were sorted and renumbered in order of decreasing electron density in the final $2F_o - F_c$ maps. During the final refinement rounds, geometry restraint weights were adjusted to minimize R_{free}, and Molprobity[28] was used to find steric problems. Each refinement is briefly described below; the refinement statistics, as well as those of the original PDB entries, are listed in Table 1.

For *St*PrxQ, minimization of the 2YWN coordinates at 1.6 Å resolution by BUSTER led to *R* and R_{free} values of 19.0 and 22.5%, respectively, and an initial difference map with most peaks being unmodeled water sites. Little density was present for the backbones of residues 44−49, so these were left unmodeled as in the original PDB entry, even though there was weak density at the expected position of the C_P−C_R disulfide (inferred from a structural overlay of deposited 2CX4 and 2CX3). A few rounds of refinement brought *R* and R_{free} to 15.7 and 18.2%, respectively, and the resolution was extended to 1.5 Å and then to 1.4 Å. The higher resolution revealed an alternate conformation of the catalytic Arg118. At this stage, refinement with PHENIX with riding hydrogens and individual anisotropic *B* factors decreased *R* and R_{free} to 12.7 and 15.8%, respectively. Electron density near Arg118, modeled as waters 17, 45, and 52, has some continuity and may be an acetate. The highest remaining difference peak, at 6.1ρ_{rms}, is the unmodeled C_P−C_R disulfide, and the final *R* and R_{free} values are 12.0 and 14.9%, respectively (Table 1).

*Ap*PrxQ-LU (originally PDB entry 2CX3) contains two half-dimers in the asymmetric unit, both in the LU conformation with a C_P−C_R disulfide. Initial BUSTER minimization at 2.6 Å of the molecular replacement solution (with no waters) resulted in *R* and R_{free} values of 21.1 and 24.4%, respectively. Manual changes included mostly adjustments of some side chain rotamers and the addition of many waters. TLS refinement using one group per monomer lowered the R_{free} by 3.3%. Extension of the resolution to 2.3 Å allowed the modeling of nine sulfate molecules as well as two glycerol molecules. The final *R* and R_{free} values were 18.6 and 22.4%, respectively (Table 1).

*Ap*PrxQ-FF/LU (originally PDB entry 2CX4) has two dimers in the asymmetric unit, each containing one subunit in the FF conformation and one in the LU conformation with a C_P−C_R disulfide. Initial BUSTER minimization at 2.3 Å resolution of the molecular replacement solution led to *R* and R_{free} values of 21.3 and 24.5%, respectively. The main manual changes included removal of a Cys80−Cys80′ (the prime denotes the second chain of a dimer) intermolecular disulfide bond, and the addition of many waters. As refinement progressed, 30 sulfates were added with occupancies ranging from 0.3 to 0.9. TLS refinement using one group per monomer decreased the R_{free} ~1.0%. Extending the resolution to 2.0 Å allowed identification of three glycerol molecules and further water sites. During the final stages, an oxidized dithiothreitol was modeled into a large oblate electron density peak in the active sites of the FF chains. During refinement, residual negative difference density at the DTT sulfur atoms existed, which we attributed to restraints tethering their *B* factors to be close to those of the much better ordered hydroxyls; this was

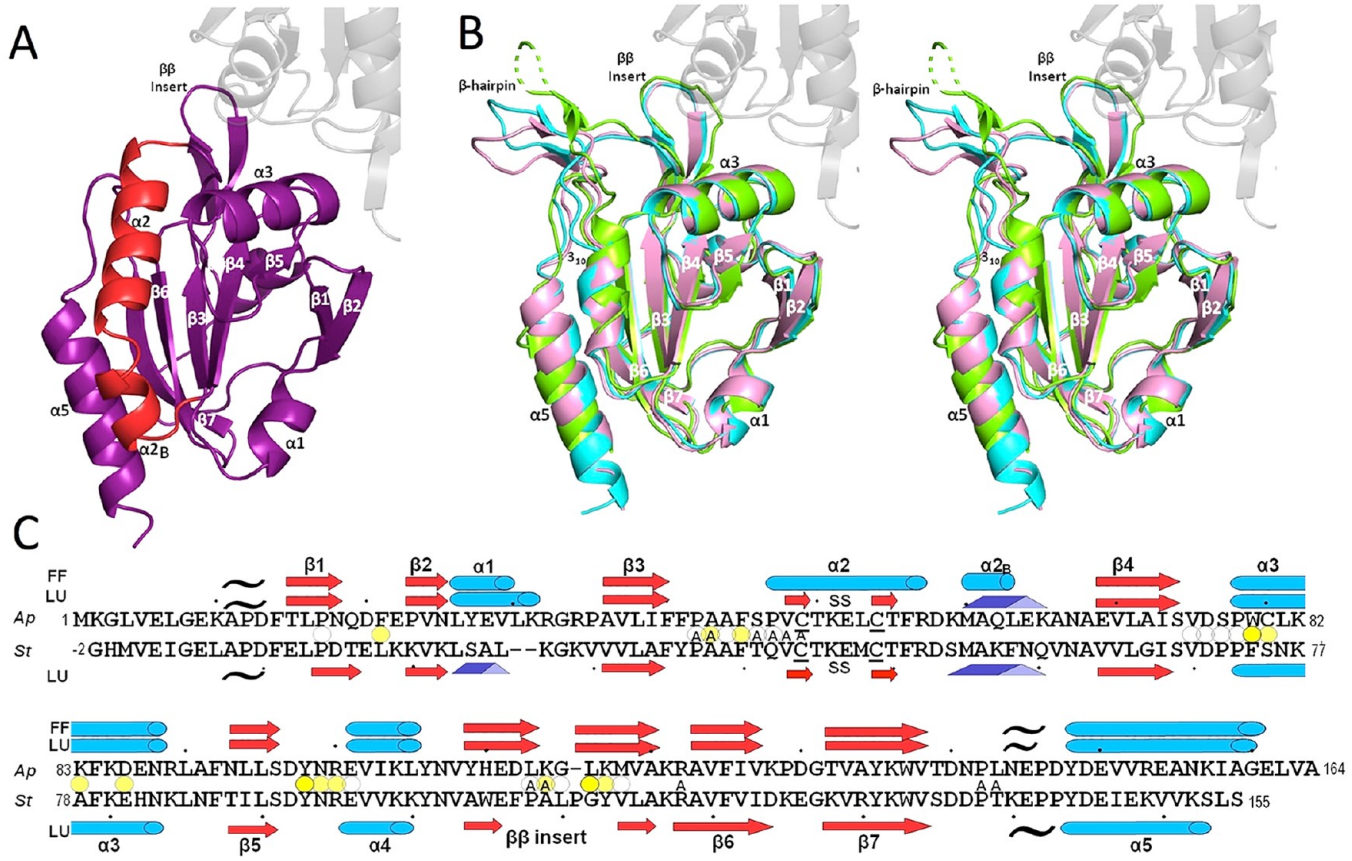


Figure 2. Overall structures of *Ap*PrxQ and *St*PrxQ. (A) The FF subunit of the *Ap*PrxQ-FF/LU structure is shown with the main segment involved in the conformational transition (residues 46−67) colored red. The universally conserved Prx secondary structures are labeled (except α4, which is behind β5). (B) Stereoview of superimposed representative LU chains from *Ap*PrxQ-FF/LU (violet), *Ap*PrxQ-LU (cyan), and *St*PrxQ (green). The unmodeled part of the *St*PrxQ β-hairpin is shown as a dashed line. The area of transition highlighted in panel A becomes a β-hairpin and 3_{10}-helix (both labeled). For perspective, here and in panel A, the second subunit of the dimer is shown for *Ap*PrxQ-LU (ghost gray). (C) Structure-based sequence alignment of *Ap*PrxQ and *St*PrxQ indicating placements of α-helices (cyan cylinders), β-strands (red arrows and pale red arrows for the inferred *St*PrxQ β-hairpin), 3_{10}-helices (blue triangular rods), and P_{II} spirals (black tildes). Ellipsoids mark residues with surface areas buried at the dimer interface of >100 Å^2 (dark yellow), 25−100 Å^2 (medium yellow), and <25 Å^2 (white). *Ap*PrxQ residues having surface area buried by the modeled DTT inhibitor are noted with A for active site.

addressed by setting the occupancy for the sulfur atoms to 0.7, which led to a clean difference map in the area. Also, compared to those of 2CX4, the chains were renamed so that the two asymmetric dimers are composed of chains AC and BD with chains A and B adopting the LU conformation and chains C and D adopting the FF conformation. The final R and R_{free} values were 19.5 and 23.4%, respectively (Table 1).

RESULTS AND DISCUSSION

As expected, the PrxQ chains are all compact domains with the Prx fold and associate to form A-type dimers (further discussed below). In the combined asymmetric units of the three crystal structures, a total of seven chains were modeled, including five with LU conformations and two with FF conformations. *St*PrxQ (PDB entry 4G2E), at 1.4 Å resolution, is the highest-resolution structure, actually the second highest resolution of any Prx structure, and its single chain (a half-dimer) in the asymmetric unit adopts the LU conformation. In this structure, the active site loop (residues 44−49) has little density and has not been modeled. The better diffracting of the *Ap*PrxQ crystal forms, at 2.0 Å resolution, has two dual-conformation (FF/LU paired) dimers in the asymmetric unit, and thus, we call it *Ap*PrxQ-FF/LU (PDB entry 4GQC). Terminal residues 1−3 and 164 are not modeled in the LU chains, and residues 1, 2, and 164 are not modeled in the FF chains. The other *Ap*PrxQ crystal form, at 2.3 Å, has two LU half-dimers in the asymmetric unit and will be called *Ap*PrxQ-LU (PDB entry 4GQF). These chains are missing residues 1 (chain A) and 1 and 2 (chain B). The noncrystallographic symmetry-related chains overlay within ∼0.2−0.3 Å, and between the two *Ap*PrxQ crystal forms, the LU chains excluding residues 45−55 (see below) agree within ∼0.4 Å. Comparisons of the *St*PrxQ (LU) and *Ap*PrxQ LU structures give a C_α rmsd of ∼1.2−1.6 Å, depending on the chain compared. A DALI[29] search of the PDB shows that *Ss*PrxQ4 (PDB entry 3HJP) is the closest other structure to both *Ap*PrxQ and *St*PrxQ; the levels of sequence identity are 46 and 78%, respectively.

Compared with the previously deposited structures, the reprocessing of the diffraction data and individualized refinements has led to improved models for all three PrxQ crystal structures. The resolutions were extended by 0.2−0.3 Å, and refinement lowered the overall R_{free} values by 2−7%, even using the higher-resolution limit (Table 1). The further refinement changed some key aspects of the structure, including removing interchain Cys80−Cys80′ disulfides from the *Ap*PrxQ-FF/LU structure, adding a dithiothreitol (DTT) ligand to the active

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

sites of the FF chains, identifying alternate conformations for residues important in catalysis, and identifying many sulfate ions. The additional major change was correcting the space groups of the two *Ap*PrxQ crystal forms (see Experimental Procedures). In our subsequent discussion of PrxQ structures, *Ap*PrxQ numbering will be used unless otherwise specified.

Overall Structures of *Ap*PrxQ and *St*PrxQ. *Fully Folded ApPrxQ.* The FF structures (*Ap*PrxQ chains C and D) have the standard Prx secondary structures of five α-helices and seven β-strands (Figure 2A), with the C_P-containing helix α2 sitting in a cradle with β-strands β4, β3, β6, and β7 providing a base and helices α3 and α5 on the two sides.[1,15] Notable variations on the standard Prx fold are as follows: helix α2 has a short bent extension we designate $\alpha2_B$, and both the *Ap*PrxQ and *St*PrxQ structures have an additional β-hairpin between α4 and β6, also present in *Ss*PrxQ4.[18] Also, *St*PrxQ substitutes a 3_{10}-helix in place of α1, as was observed in *Ss*PrxQ4[18] (Figure 2B,C).

In terms of the active site, the FF conformation tends to be well-conserved among Prxs,[1,9] and *Ap*PrxQ is no exception with the C_P (Cys49 in the initial turn of α2) along with Arg122 and Ser46 at the bottom of the active site pocket (Figure 3). As was

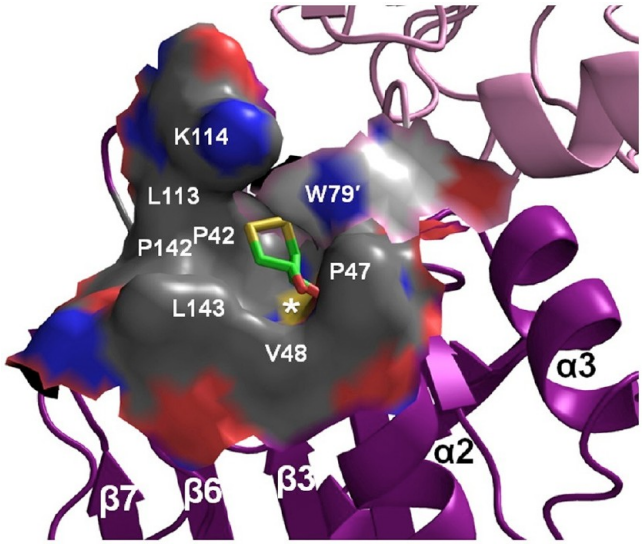


Figure 3. *Ap*PrxQ FF active site hydrophobic collar. Shown is the active site molecular surface annotated with residue numbers and showing the bound DTT (sticks). The surface is colored by element, distinguishing carbons from the main monomer (dark gray) or the other subunit (light gray), nitrogens (blue), oxygens (red), and sulfurs (yellow). The peroxidatic Cys thiolate surface is marked by a white asterisk. The catalytic Arg122 contributes the blue nitrogen surface left of the asterisk, and barely visible red surface right of the asterisk is from the Ser46 hydroxyl.

seen for Tpx[15] and Prx5 subfamily enzymes, the upper part of the *Ap*PrxQ active site pocket is lined with a collar of nonpolar side chains (Figure 3) and seems to be well-suited to preferentially reducing organic peroxides, with Trp79′ from the other subunit of the dimer contributing to this collar, suggesting it is part of the complete active site. Indeed, persistent pancake-like electron density was observed (Figure S1a of the Supporting Information) that we eventually modeled as an oxidized dithiothreitol (DTT). Although the density was not definitive on its own, the interpretation was made because DTT was present in the crystallization buffer and the density was a good match for how DTT was seen bound to the FF conformation of human PrxV (PDB entry 3MNG)[9] with its two hydroxyls mimicking a peroxide substrate (Figure S1b of the Supporting Information). We rationalize that the rather weak density for the disulfide of DTT occurs because it has much more variation in position than the much better localized hydroxyls.

DTT is the first substrate-mimicking ligand to be identified in a FF $C_R^{\alpha2}$ subtype (FF *Ss*PrxQ4 was modeled with a chloride ion in the active site)[18] and also represents the only such ligand for a dimeric PrxQ. Examples of ligands in monomeric PrxQs include a formate in FF *Xc*PrxQ (PDB entry 3GKM)[30] and a citrate molecule in *Ss*PrxQ1,[19] though their exact positions overlay only loosely with the expected H_2O_2 binding site.[9] None of the monomeric PrxQs have a hydrophobic collar as pronounced as that of the dimeric forms but have active sites that are much shallower and more polar. Whereas it was proposed for the Tpx enzymes that the active site hydrophobic collar lends specificity for organic peroxides,[1] no PrxQ enzymes have yet been shown to have such a preference.[16,31−33] Indeed, the monomeric *E. coli* PrxQ has essentially equivalent k_{cat}/K_m values for both hydrogen peroxide and the bulky organic cumene hydroperoxide, although it is less active with *tert*-butyl hydroperoxide.[16] Nevertheless, because Trp79′ makes a substantial contribution of nonpolar surface area to the $C_R^{\alpha2}$ active site, it could allow dimer formation to influence substrate specificity and catalytic activity.

Locally Unfolded Conformation of $C_R^{\alpha2}$ PrxQ Enzymes. The $C_R^{\alpha2}$ LU conformation is seen in four *Ap*PrxQ chains (two from each crystal form) and the single *St*PrxQ chain. Compared with the FF conformation, the only substantive backbone changes involve residues 46−68, with residues 46−56 forming a β-hairpin extending out from the protein core and residues 59−65 forming a 3_{10}-helix (Figure 2). A striking feature is that the direction at which the hairpin extends varies among the independent LU chains (Figure 2B), even between the *Ap*PrxQ chains, in which case the variation cannot be due to sequence differences. At their tips, the hairpins span a range of ∼15 Å, though this is an estimate, because density was so weak for the terminal residues of the *St*PrxQ hairpin that they were not modeled.

Despite the variation in position, the internal conformations of the hairpin are nearly identical (Figure 4). The β-hairpin is internally stabilized by three β-strand hydrogen bonds and the disulfide bridge, as well as hydrogen bonds at the tip of the hairpin between the Thr50 side chain hydroxyl and the backbone nitrogens of Glu52 and Leu53. Additionally, the C_P backbone nitrogen exhibits a hydrogen bond with the carbonyl of Leu143 that is only present in *Ap*PrxQ-FF/LU, providing the single external stabilizing interaction for the hairpin position. Because the turn is highly mobile, with *B* factors of >80 Å^2 in the *Ap*PrxQ-FF/LU structure (see Figure 6A), the conformational details must be considered tentative.

Sulfate Binding. An additional noteworthy feature of the *Ap*PrxQ structures, which were crystallized using ammonium sulfate as a precipitant, is the presence of many ordered sulfate molecules. Although ammonium sulfate is a common crystallizing agent, such extensive localization of sulfates does not generally occur. Interestingly, only one sulfate site, found near His110, was present in both crystal forms and both conformations, implying that most are not discrete well-defined sulfate binding sites. It has been suggested in studies of halophiles that their proteins have evolved to be stable at the high concentrations of sodium partly through an increase in the

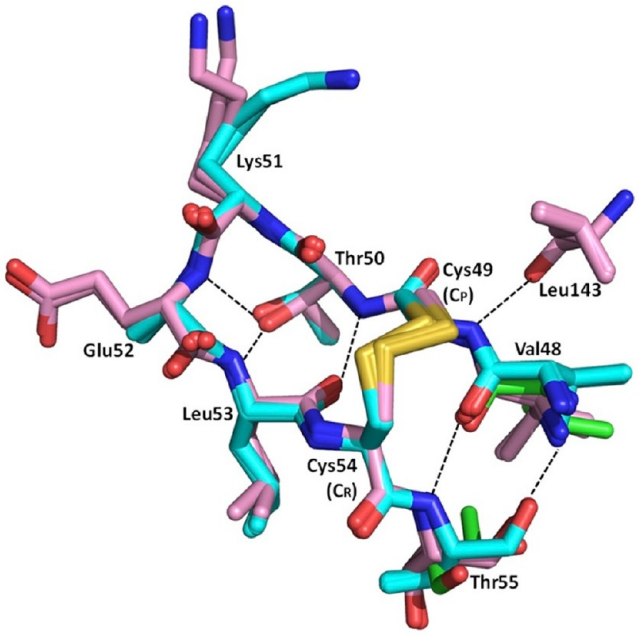


Figure 4. Consistent internal conformation of the *Ap*PrxQ and *St*PrxQ β-hairpins. The overlaid β-hairpins of all five LU chains are shown: two from *Ap*PrxQ-FF/LU (violet), two from *Ap*PrxQ-LU (cyan), and one from *St*PrxQ (green). Residues are labeled, and H-bonds (dashed lines) internal to the hairpin and to the carbonyl of residue 142 are shown. As modeled, the conformation is a type I-like turn with φ and ψ values of approximately −70° and −35° for Lys51 and φ and ψ values of approximately −130° and −30° for Glu52.

number of surface acidic residues that coordinate the positively charged sodium ions.[34] Similarly, because the environment of the hyperthermophile *A. pernix* contains high concentrations of sulfate species, especially thiosulfate,[35] the organism's proteins may have evolved to take advantage of interactions with these solutes to achieve thermostability.

Transition between FF and LU Conformations of $C_R^{\alpha 2}$ PrxQs. A naïve prediction might be that formation of a disulfide bond between Cys residues separated by just four amino acids might involve only modest structural changes, as is true, for example, in the active site of glutathione reductase.[36] However, an overlay of the FF and LU *Ap*PrxQ structures shows that the transition between FF and LU conformations involves very large (4−17 Å) backbone shifts of residues 47−67 and smaller (0.5−2 Å) shifts of residues 80−90 and 140−160 (Figure 5A). The substantial rearrangement of residues 47−67 involves the unfolding of helix α2 to make a β-hairpin and the extension of helix $\alpha 2_B$ to become a 3_{10}-helix (Figure 5B); the smaller shifts are adaptations that accommodate the changes in residues 47−67. In this striking transition, to allow the 11 residues in the β-hairpin to stick out like a tongue from the protein surface, it can be seen that helix $\alpha 2_B$ provides excess slack that can be spooled out in the form of a 3_{10}-helix that is more tightly associated with the protein core to cover ∼8.5 Å of the gap that would otherwise have been opened (Figure 5B).

The ability of $\alpha 2_B$ to undergo a transition to a 3_{10}-helix is made possible by key nonpolar side chains that can interact with three hydrophobic pockets in the protein core (Figure 5B,C). In FF *Ap*PrxQ, the anchoring side chains are Leu53, Phe56, and Leu63 and are replaced in the LU form by Met60, Leu63, and Ala66, respectively. Because the side chain positions do not match perfectly, the core residues lining the pocket (such as Ile39 of β3, Phe84 of α3, and Tyr148 of α5) adjust, explaining the smaller backbone shifts in residues 80−90 and 140−160. Also, the side chains of Glu145 and Asp58 aid the transition by making N-cap hydrogen bonds with the backbone NH groups of residues 59 and 60 in the first turn of the 3_{10}-helix. The movements near residue 140 also reflect shifts in the backbone carbonyl interactions with Arg122 and the gain of the Leu143 carbonyl H-bond with the β-hairpin (Figure 4). The small shifts in α3 are partly related to the movement of Arg57, which in the FF form links α2 with the $\alpha 3-\beta 5$ loop by H-bonding with the Asn88 carbonyl. The transition is also helped by the relatively smooth, nonpolar faces of β-strands 3−5, with few "sticky" polar interactions, as was similarly described for Tpx subfamily enzymes.[15]

Because of the high degree of sequence similarity between FF *Ss*PrxQ4 and LU *St*PrxQ (with only two differences in the 21-residue transition segment, *St*PrxQ Gln42 → Ser and Gln60 → Glu), this pair provides a second view of the FF ↔ LU conformational transition. Other than differences near residues 25 and 110 that are not correlated to the transition, this comparison also shows large backbone shifts of residues 47−67 and smaller shifts near residues 85−90 and 140−150 (Figure 5A, green line). For this pair, all of the features described above are present, including the two 3_{10}-helix N-capping interactions and the existence of three key hydrophobic anchor positions, which in this case involve FF Met53, Phe56, and Phe63 being replaced by LU Met60, Phe63, and Val66, respectively.

Equally dramatic are changes in chain disorder that are associated with the FF ↔ LU transition of residues 47−67 (Figure 6). In the FF active site, residues 47−57, corresponding to helix α2, are as ordered as the protein core but are much less ordered when adopting the LU β-hairpin. In contrast, $\alpha 2_B$ residues 58−67 are rather disordered in the FF form but become much more ordered in LU as the single weak anchoring provided by residue 63 is replaced by strong anchoring of residues 60, 63, and 66 (Figure 5C). The high *B* factors of $\alpha 2_B$ in the FF form can be understood in that the whole helix is tethered to the protein core by only one nonpolar side chain (position 63) and that tethering involves a larger side chain (Leu in *Ap*PrxQ and Phe in *St*PrxQ) sticking only its tip into a small pocket that is sized more optimally for an Ala (as in *Ap*PrxQ) or a Val (as in *St*PrxQ).

An interesting question for understanding the Prx FF ↔ LU transition is whether there is always an ongoing rapid interchange between the two conformations or whether the transition occurs only when triggered by a redox change in the active site C_P residue. Protein crystallography cannot provide data to answer this question, but NMR can. Fortuitously, an extensive NMR study of a monomeric $C_R^{\alpha 2}$ PrxQ, *Arabidopsis thaliana* PrxQ (*At*PrxQ)[37] with a sequence ∼40−50% identical to those of *Ap*PrxQ and *St*PrxQ, provides further insight. The dynamics of both the reduced (dithiol) and oxidized (disulfide) forms were studied. Of the 140 non-proline residues in the protein, 132 could be assigned in the disulfide form, but only 63 assigned in the reduced form. The unassigned residues suffered from extensive line broadening due to conformational exchange dynamics, meaning that by this measure the FF conformation is more disordered than the LU conformation. The regions involved were in four segments that correlate reasonably well with the regions that undergo conformational and/or environmental change in the FF ↔ LU transition (Figure 5A). Among the residues without large C_α shifts, residues 35−45 pack against $\alpha 2_B$ and residues 110−125 form part of the active site

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

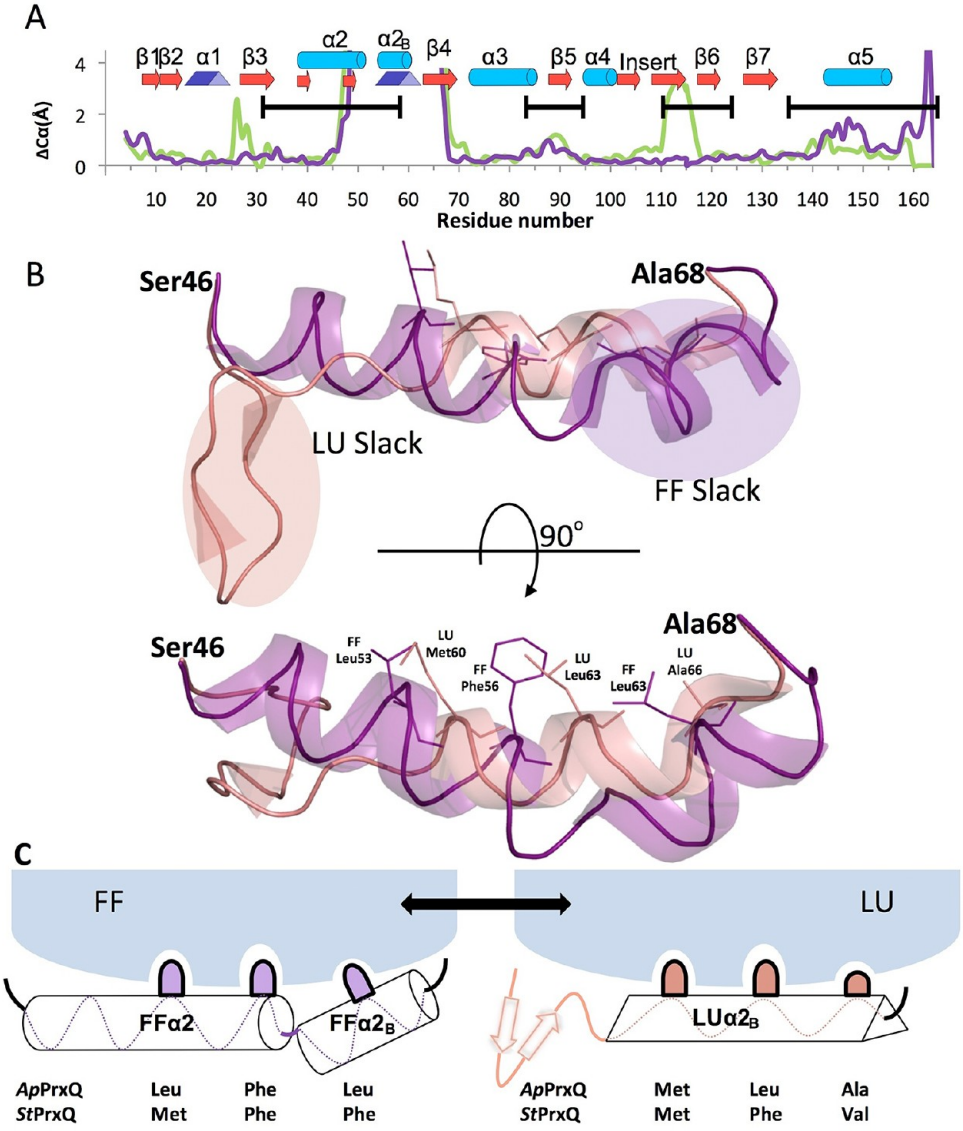


Figure 5. Essential structural features of the PrxQ $C_R^{\alpha 2}$ FF ↔ LU transition. (A) C_α shifts between the FF and LU conformations are given for the *Ap*PrxQ pair (purple) and the *St*PrxQ/*Ss*PrxQ4 pair (green). *Ap*PrxQ numbering is used, and <4 Å shifts are shown in the inset. Black bars indicate the regions undergoing intermediate exchange in the *At*PrxQ NMR study.[37] Peaks near residues 25 and 110 in the *St*PrxQ/*Ss*PrxQ4 pair are not related to the FF ↔ LU transition but are due to a one-residue insertion in *St*PrxQ4 and the influence of crystal contacts, respectively. (B) Orthogonal views of the backbone ribbon for residues 46−68 in the FF (purple) and LU (salmon) conformations. Side chains for the hydrophobic anchors are shown as sticks (labeled in the bottom view), and regions in each conformation not well anchored and providing slack that extends as part of the conformational change are indicated in the top view. (C) Schematic model of how the hydrophobic anchors shift interactions with the protein core during the conformational change.

collar, so in both cases, their environment becomes much more solvent-exposed in the LU conformation.

These data indicate that even though fully reduced PrxQ molecules are largely populating the FF conformations, they are also constantly sampling the LU conformation, according to model-free analysis, at a rate of ~1650 times per second.[37] For Prx1 subfamily enzymes,[38] proteins with a disrupted A-type interface had a less stabilized and less restricted FF conformation,[39] implying that dimeric PrxQ proteins may have slightly more restricted conformational dynamics than the monomer of the NMR study. That the disulfide form is on the whole less dynamic, despite the floppy β-hairpin (which in fact consists of the main unassigned residues in this form), can be understood in that the formation of the disulfide "locks" the protein into the LU conformation by removing its ability to sample the FF conformation until the disulfide is reduced. These studies support the view that local unfolding of even the reduced form of the protein occurs sufficiently frequently to support catalysis, and that formation of S_POH does not trigger the conformational change but simply allows the chemistry of disulfide formation to trap the protein in the LU conformation.

The facile local unfolding of the reduced protein implies that the energy barrier between conformations is small, with neither state being highly stabilized. We speculate that whereas the FF conformation has many interactions stabilizing it (e.g., much buried surface and salt bridges spanning from Lys59 to Asp149 and from Lys51 to Glu45), $\alpha 2_B$ is relatively poorly stabilized. In contrast, in the LU conformation, the $\alpha 2_B$ residues are interacting well with the protein, but the $\alpha 2$ residues (in the β-hairpin) are poorly stabilized. Together, this leads to neither

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

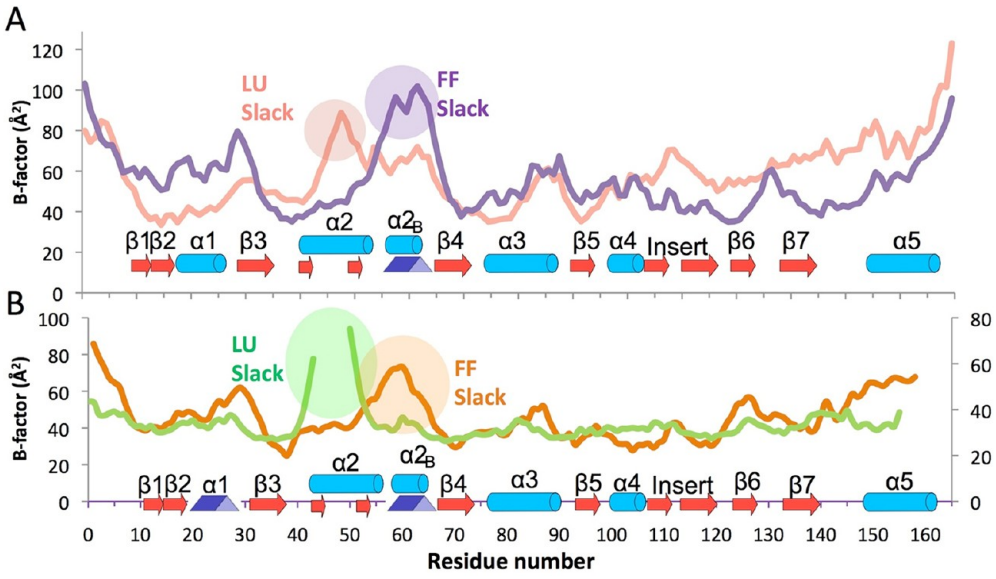


Figure 6. Changes in disorder associated with the FF ↔ LU transition. (A) *B* factors of the FF (purple) and LU (salmon) chains from the *Ap*PrxQ-FF/LU crystal form. Secondary structure along the chain is indicated. (B) Same as panel A, but for the *Ss*PrxQ4 FF chain (orange) and the *St*PrxQ LU chain (green) using *St*PrxQ numbering. The left *y*-axis is the *St*PrxQ *B* factor and the right *y*-axis the *Ss*PrxQ4 *B* factor. In *St*PrxQ, α1 is a 3_{10}-helix (as shown), but in *Ss*PrxQ4, it is an α-helix as it is in most other Prxs. This change is not related to the FF ↔ LU transition but is due to a one-residue insertion in *Ss*PrxQ4.

state being highly stable and levels the energy landscape, aiding the process of constant transition between conformations.

Dimer Interface. As previously mentioned, all chains we have analyzed are A-type dimers, with interfaces involving ~30 residues from strands β1 and β2, and the loops prior to α2−α4 (Figure 7A).[12,38] These interacting regions group well into the defined sequence regions numbered 0−4 that were used for other Prxs.[12,38] As was noted in the description of *Ss*PrxQ4,[18] the *Ap*PrxQ and *St*PrxQ interfaces involve residues from a $\beta\beta$ insertion (here designated region 4) absent in the monomeric *At*PrxQ[37] and *Sc*PrxQ.[20] However, the monomeric *Xc*PrxQ and *Xf*PrxQ have such an insertion (though it is extended and contains a higher proportion of residues with polar side chains), meaning this insertion cannot be strictly used as a marker for dimeric PrxQs.

Despite the dramatic shift between conformations FF and LU, we find that the dimer interface remains virtually identical among the FF/LU, LU/LU, and FF/FF (from *Ss*PrxQ4) dimers (Figure 7A). As all are essentially equivalent, we focus our description on the *Ap*PrxQ interface, noting differences where needed. The dimerization interface buries ~2000 Å^2 of surface (1050 and 915 Å^2 per chain in *Ap*PrxQ and *St*PrxQ, respectively). Most of the buried area is nonpolar with the four most buried positions being Phe45 (70−75 Å^2), Trp79 (150 Å^2), Tyr98 (120 Å^2), and Leu116 (105−110 Å^2), one each from regions 1−4, respectively (Figure 7B). Leu116 of region 4 is part of the $\beta\beta$ insertion, and its role is conserved between the PrxQ dimers as a tight packing interaction into a nonpolar pocket formed by the side chains of Tyr98′, Trp79′, Pro78′, and Phe21′. The insertion is also stabilized by variable hydrogen bonds depending on sequence differences. An important point is that the close Cys80−Cys80′ pair (unique to *Ap*PrxQ) packs together but does not form a disulfide (Figure 7C).

In terms of well-buried polar interactions, there is one water site (Water81) at the dimer interface conserved among *Ap*PrxQ, *St*PrxQ, and *Ss*PrxQ4. It is among the more highly ordered water sites in the structures (the waters are ordered based on decreasing electron density). Water81 makes bridging H-bonds among the Ser77 hydroxyl, the Asp76′ carboxylate, and the Ala44′ backbone oxygen (Figure 7C). Additional polar interactions that are not in common occur near the periphery of the interfaces, such as the symmetric intermolecular salt bridges in *Ap*PrxQ between Lys83 and Glu87′ and between Glu87 and Lys83′.

As was noted above, the interfaces of *Ap*PrxQ and *St*PrxQ contribute to the active site (Figures 2A and 3). In fact, in addition to Trp79′, six other residues interacting with the DTT ligand (Figure 2C) are also involved in the dimer interface. An interesting question is how the FF/LU mixed dimers came to be stabilized in the crystal form. The systematic heterogeneity seems to be related to crystal packing interactions: in the FF/LU crystal form, it is impossible for the FF chains to be LU because the LU disulfide loop would collide with Lys31 of the adjacent FF chain. We hypothesize that crystals are nucleated by the FF/LU heterodimers that form as the DTT present becomes oxidized, and then the crystal can grow as more protein oxidizes. Once an FF/LU dimer is incorporated into the crystal lattice, the crystal packing interactions and the DTT binding hinder the LU transition of the FF subunit, so it cannot be oxidized via disulfide exchange. As long as some reduced DTT remains, disulfide exchange can occur to allow remaining soluble dimers to adopt the crystallizable FF/LU form. We emphasize that the existence of the FF/LU dimers should not be taken as evidence of half-of-the-sites reactivity or another more general allosteric relationship.

As a final thought about the relevance of the dimer interface, a characterization of *Ss*PrxQs[31] found that the dimeric *Ss*PrxQ4 was substantially more heat stable, displayed optimal activity levels at higher temperatures, and retained a higher relative activity when incubated at 95 °C, compared to the monomeric *Ss*PrxQ1 and SsPrxQ3. *Ss*PrxQ4 was also found to be exceptionally resistant to chemical denaturation, retaining 50% activity after 30 min in 6 M urea.[31] It may be that the

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

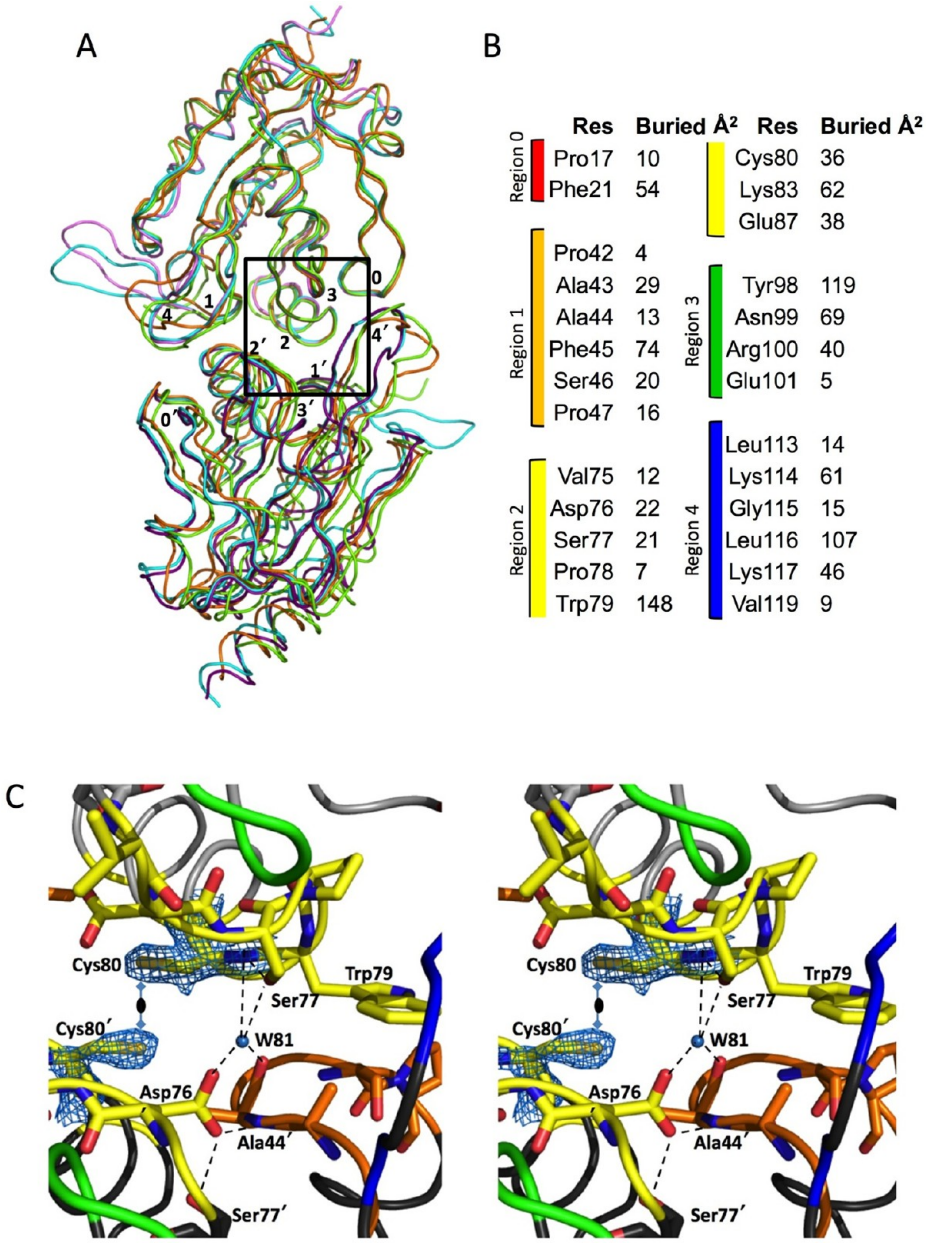


Figure 7. PrxQ dimer interface that is unperturbed during the catalytic cycle. (A) *Ap*PrxQ-LU/LU (cyan/cyan), *Ap*PrxQ-FF/LU (purple/pink), *St*PrxQ-LU/LU (green/green), and *Ss*PrxQ4-FF/FF (gold/gold) dimers overlaid based on the top subunit. (B) Residues with >3 Å^2 of surface buried at the *Ap*PrxQ-LU/LU dimer interface grouped into sequence regions 0−4[38] and colored to correspond with panel C. (C) Close-up (stereo) of interactions of the two-fold *Ap*PrxQ-FF/LU interface. The coordination of conserved buried Water81 is shown, as well the primary conformation and $2F_o - F_c$ electron density at $2.0\rho_{rms}$ for Cys80 and Cys80′, which do not form a disulfide (the sulfur atoms are separated by a distance of ∼3.5 Å).

more stable dimeric forms of PrxQs would allow these hyperthermophiles to temporarily endure a larger range of temperature fluctuation by providing protection from reactive oxygen species at temperatures that would denature the organism's other PrxQ enzymes.

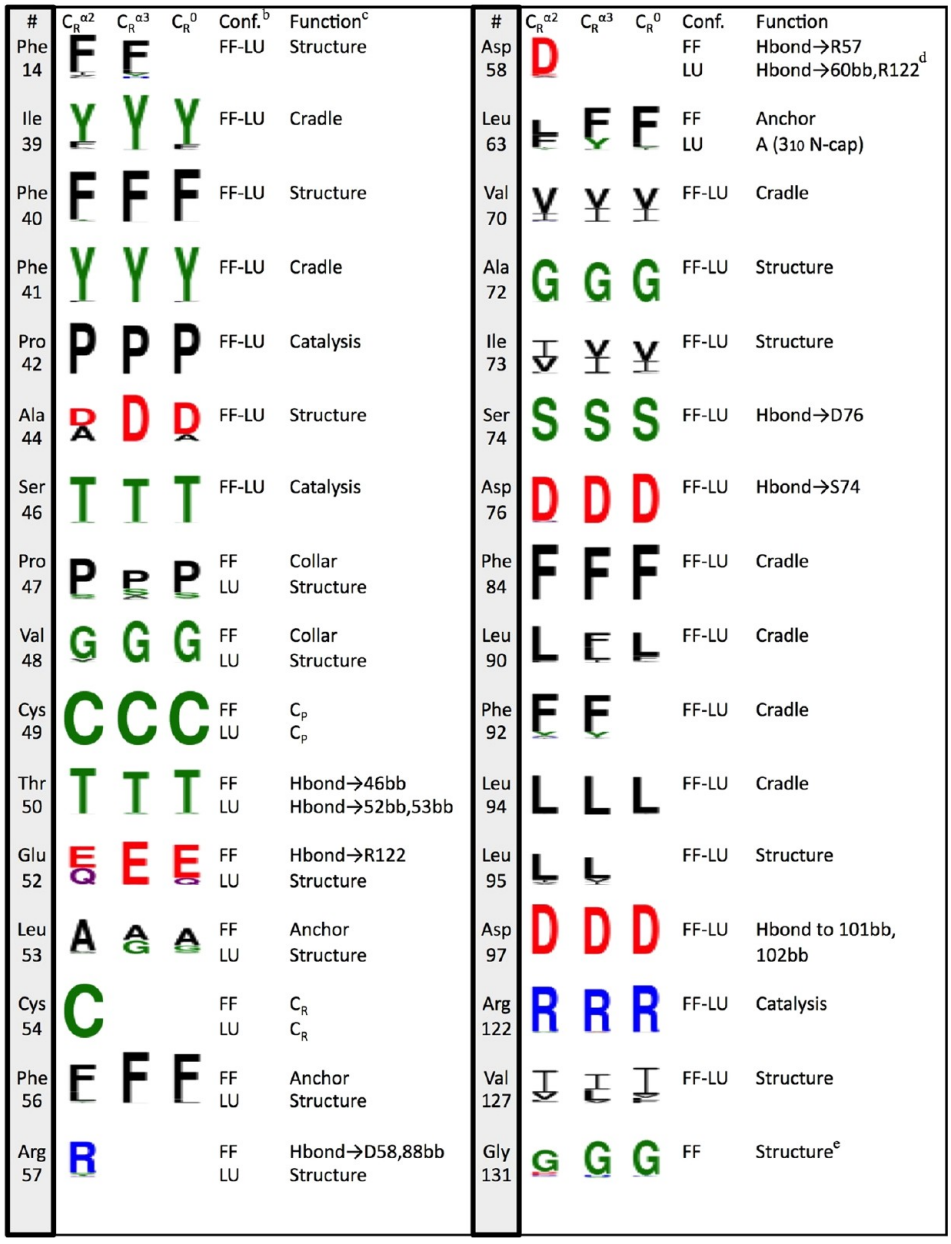
Sequence Conservation among $C_R^{\alpha 2}$, $C_R^{\alpha 3}$, and C_R^{0} PrxQs. The PrxQ subfamily contains a surprising diversity in the mechanism by which the sulfenic acid form is resolved, containing $C_R^{\alpha 2}$, $C_R^{\alpha 3}$, and C_R^{0} Prxs. Whereas the $C_R^{\alpha 2}$ PrxQs undergo the drastic unraveling of α2 documented here, $C_R^{\alpha 3}$ PrxQ conformations change little between redox states, requiring only the inward bending of the α2−α3 loop and the flipping of the C_P and C_R side chains to form the disulfide.[30]

To see what sequence features are conserved with the $C_R^{\alpha 2}$ subtype, conservation patterns were assessed for 511 $C_R^{\alpha 2}$, 74 $C_R^{\alpha 3}$, and 387 C_R^{0} PrxQ sequences, and the roles of each highly conserved position were identified (Table 2 and Figure 8). By also assessing conservation at these positions among the other subtypes, we are able to comment on which patterns appear to be unique to the $C_R^{\alpha 2}$ subtype and which are conserved in PrxQs in general. We consider three tiers of highly conserved residues: ≥98% identical, ≥90% identical, or ≥90% identical between two residue types, all for residues present in at least 95% of the sequences.

The highly conserved residues are mostly centered around the FF active site, especially positions 38−58, which

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

Table 2. Conservation of PrxQ $C_R^{\alpha 2}$ Positions[a]



[a]The sequence relative to *Ap*PrxQ is given. PrxQ sequences identified as previously described were edited to remove sequences that were incomplete or did not align properly at the absolutely conserved residues (Cys49 and Arg122), and subdivided based on the location of the C_R. Conservation of positions among subgroups is expressed as a standard SeqLogo of 511 $C_R^{\alpha 2}$ sequences, 74 $C_R^{\alpha 3}$ sequences, and 387 C_R^{0} sequences. Those positions that exhibited conservation patterns of ≥90 or ≥90% between two residues are shown, with blanks indicating a lack of conservation for that subtype. The overall size of each SeqLogo image corresponds to the number of total sequences that aligned at this position, and the relative height of the single-letter residue abbreviation corresponds to the percentage of conservation, with a minimal sequence alignment cutoff of 95%. For example, position Cys49 represents 100% of sequences aligning and 100% conservation of a Cys. The data were visualized using WebLogo 3.3 (http://weblogo.threeplusone.com/).[43] [b]The conformation(s) in which the function occurs: fully folded (FF), locally unfolded (LU), or both (FF-LU). [c]General function and important interactions are defined as follows: catalysis, involved in catalysis and/or substrate positioning; Collar, part of the hydrophobic collar of the active site; Anchor, one of the hydrophobic anchors involved in the change in conformation; Cradle, packs against $\alpha 2$, $\alpha 2_B$, or the active site; Hbond, makes important structural interactions through hydrogen bonding, with the residue specified if the bond is to another side chain or "bb" if to a backbone atom; Structure, maintains general protein structure. [d]The distance between Asp58 and Arg122 is variable and in some chains is too great for formation of a salt bridge. [e]Adopts Gly-specific Φ and ψ values of 90° and 0°, respectively.

encompasses the end of $\beta 3$, the loop before helix $\alpha 2$ (the C_P loop), and most of $\alpha 2$. This can be interpreted as a conservation of redox function, given that these proteins have an identical mode and site for catalysis, making the structure and stability of the $\alpha 2$ region paramount. Closely related to this are the side chains that line the cradle in which helices $\alpha 2$ and $\alpha 2_B$ lie, and this accounts for most of the other residues that have a similar level of conservation across all three PrxQ subtypes. In contrast to what was seen for the Tpx subfamily of Prxs,[15] residues at the dimer interface are not well-conserved. This is not surprising given that the $C_R^{\alpha 2}$ subtype includes both dimers and monomers. In fact, only five dimer interface

dx.doi.org/10.1021/bi301017s | *Biochemistry* 2012, 51, 7638−7650

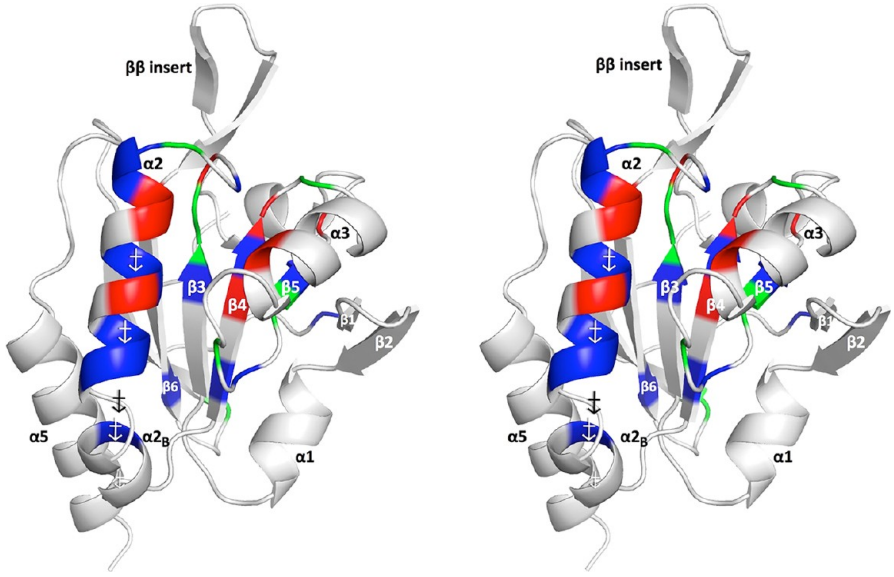


Figure 8. Positions highly conserved in $C_R^{\alpha 2}$ PrxQ enzymes. Stereoview backbone ribbon of the FF subunit from the *Ap*PrxQ-FF/LU structure colored to highlight conservation, distinguishing positions that are ≥98% identical (red), ≥90% identical (green), and ≥90% identical between two amino acids (blue). The remaining residues are colored gray. The five anchor residues are marked by anchor symbols.

positions are highly conserved (Pro42, Ser46, Pro47, Phe56, and Asp76), and none of these have more than 25 Å^2 of surface buried at the interface (Figure 7B), consistent with the conservation of these positions being related to functions distinct from dimerization. The *ββ* insertion (positions 113–117) common to all known PrxQ dimers is present in only ~30 of the 511 sequences, suggesting $C_R^{\alpha 2}$ PrxQs are predominantly monomers (or form dimers of a different type), and even that number may be an overestimate because, as mentioned above, some monomers contain this insertion.

Particularly interesting are positions with conservation patterns unique to the $C_R^{\alpha 2}$ subtype. These include the completely unique Cys54 (C_R), Arg57, and Asp58, and the somewhat unique Leu53, Phe56, and Leu63. Each of these positions has been mentioned above as having specific roles in the FF ↔ LU transition (Table 2). The conservation of these positions provides further evidence that their specific roles are rather important for facilitating the transition. That there are no further residues with a distinct $C_R^{\alpha 2}$-specific conservation pattern supports the idea that we have a handle on the important structure–function features of the $C_R^{\alpha 2}$ PrxQ forms.

Challenge and Opportunity of Unpublished Structural Genomics Results. Structural genomics groups have contributed greatly to our knowledge of protein structures. Already by 2006 their work accounted for ~20% of the total number of structures added to the PDB.[40] However, less well-documented is what fraction of these structures have never received the individual attention and detail-oriented analyses needed for a primary literature report. Our analyses here of the three PrxQ structures illustrate not only that such unpublished structures can contain a wealth of information useful to a field (e.g., ref 41) but also that they can have real shortcomings, errors and incomplete interpretations, that detract from their reliability. This is of course a significant problem when, as for the previously deposited *Ap*PrxQ and *St*PrxQ entries, the unvetted structures get used "as is" by others for structure-based sequence alignments,[1,14,18,19,30] assigning residue functionality,[14] threading,[37] protein–protein docking studies,[18] homology modeling,[18] overlays,[30] and other structure-based comparisons.[1,30] For unpublished structural genomics structures to provide the maximal benefit, they must be carefully analyzed and described, and how this will be done is an important unresolved issue that raises novel ethical and policy questions that need to be explicitly addressed by the structural biology community.

■ ASSOCIATED CONTENT

Supporting Information

Evidence of the modeled DTT (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (541) 737-3200. E-mail: karplusp@science.oregonstate.edu.

Funding

This study was supported by National Institutes of Health Grant RO1 GM050389 to L.B.P. and P.A.K.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

For conducting this study, we are grateful to Professors S. Yokoyama and S. Kuramitsu of RIKEN Systems and Structural Biology Center who provided the original diffraction data for the structures in this study and whose support allayed our concerns about the appropriateness of describing unpublished work conducted by others.

■ ABBREVIATIONS

Prx, peroxiredoxin; C_P, peroxidatic cysteine; C_R, resolving cysteine; FF, fully folded; LU, locally unfolded; BCP, bacterioferritin comigratory protein; $C_R^{\alpha 2}$, PrxQ enzymes with C_R located in *α*2; $C_R^{\alpha 3}$, PrxQ enzymes with C_R located in *α*3; C_R^{0}, PrxQ enzymes lacking C_R; *Ss*PrxQ1, *S. solfataricus* PrxQ1; *Ss*PrxQ4, *S. solfataricus* PrxQ4; *Sc*PrxQ, *Saccharomyces cerevisiae* PrxQ; *Xf* PrxQ, *X. fastidiosa* PrxQ; *Xc*PrxQ, *Xa. campestris* PrxQ;

*At*PrxQ, *Ar. thaliana* PrxQ; *Ap*PrxQ, *A. pernix* PrxQ; *St*PrxQ, *S. tokodaii* PrxQ; PDB, Protein Data Bank; DTT, 1,4-dithiothreitol; NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation.

REFERENCES

(1) Hall, A., Nelson, K., Poole, L. B., and Karplus, P. A. (2011) Structure-based Insights into the Catalytic Power and Conformational Dexterity of Peroxiredoxins. *Antioxid. Redox Signaling 15*, 795−815.

(2) Winterbourn, C. C. (2008) Reconciling the chemistry and biology of reactive oxygen species. *Nat. Chem. Biol. 4*, 278−286.

(3) Niki, E., Yamamoto, Y., Komuro, E., and Sato, K. (1991) Membrane Damage Due to Lipid Oxidation. *J. Clin. Nutr. 53*, 201S−205S.

(4) Nakamura, T., Kado, Y., Yamaguchi, T., Matsumura, H., Ishikawa, K., and Inoue, T. (2010) Crystal Structure of Peroxiredoxin from *Aeropyrum Pernix* K1 Complexed with Its Substrate, Hydrogen Peroxide. *J. Biochem. 147*, 109−115.

(5) Karplus, P. A., and Hall, A. (2007) Structural Survey of the Peroxiredoxins. *Peroxiredoxin Systems Subcell. Biochem. 44*, 41−60.

(6) Hofmann, B., Hecht, H. J., and Flohé, L. (2002) Peroxiredoxins. *Biol. Chem. 383*, 347−364.

(7) Knoops, B., Loumaye, E., and Eecken, V. (2007) Evolution of the Peroxiredoxins. *Peroxiredoxin Systems Subcell. Biochem. 44*, 27−40.

(8) Trujillo, M., Ferrer-Sueta, G., Thomson, L., Flohé, L., and Radi, R. (2007) Kinetics of Peroxiredoxins and their Role in the Decomposition of Peroxynitrite. *Peroxiredoxin Systems Subcell. Biochem. 44*, 83−113.

(9) Hall, A., Parsonage, D., Poole, L. B., and Karplus, P. A. (2010) Structural Evidence that Peroxiredoxin Catalytic Power Is Based on Transition-State Stabilization. *J. Mol. Biol. 402*, 194−209.

(10) Wood, Z. A., Poole, L. B., and Karplus, P. A. (2003) Peroxiredoxin Evolution and the Regulation of Hydrogen Peroxide Signaling. *Science 300*, 650−653.

(11) Hall, A., Karplus, P. A., and Poole, L. B. (2009) Typical 2-Cys peroxiredoxins: Structures, mechanisms and functions. *FEBS J. 276*, 2469−2477.

(12) Sarma, G. N., Nickel, C., Rahlfs, S., Fischer, M., Becker, K., and Karplus, P. A. (2005) Crystal Structure of a Novel *Plasmodium falciparum* 1-Cys Peroxiredoxin. *J. Mol. Biol. 346*, 1021−1034.

(13) Krnajski, Z., Gilberger, T. W., Walter, R. D., Cowman, A. F., and Müller, S. (2002) Thioredoxin Reductase Is Essential for the Survival of *Plasmodium falciparum* Erythrocytic Stages. *J. Biol. Chem. 277*, 25970−25975.

(14) Nelson, K. J., Knutson, S. T., Soito, L., Klomsiri, C., Poole, L. B., and Fetrow, J. S. (2010) Analysis of the peroxiredoxin family: Using active-site structure and sequence information for global classification and residue analysis. *Proteins: Struct., Funct., Bioinf. 79*, 947−964.

(15) Hall, A., Sankaran, B., Poole, L. B., and Karplus, P. A. (2009) Structural Changes Common to Catalysis in the Tpx Peroxiredoxin Subfamily. *J. Mol. Biol. 393*, 867−881.

(16) Reeves, S. A., Parsonage, D., Nelson, K. J., and Poole, L. B. (2011) Kinetic and Thermodynamic Features Reveal That *Escherichia coli* BCP Is an Unusually Versatile Peroxiredoxin. *Biochemistry 50*, 8970−8981.

(17) Gretes, M. C., Poole, L. B., and Karplus, P. A. (2012) Peroxiredoxins in Parasites. *Antioxid. Redox Signaling 17*, 608−633.

(18) Limauro, D., D'Ambrosio, K., Langella, E., De Simone, G., Galdi, I., Pedone, C., Pedone, E., and Bartolucci, S. (2010) Exploring the catalytic mechanism of the first dimeric Bcp: Functional, structural and docking analyses of Bcp4 from *Sulfolobus solfataricus*. *Biochimie 92*, 1435−1444.

(19) D'Ambrosio, K., Limauro, D., Pedone, E., Galdi, I., Pedone, C., Bartolucci, S., and De Simone, G. (2009) Insights into the catalytic mechanism of the Bcp family: Functional and structural analysis of Bcp1 from *Sulfolobus solfataricus*. *Proteins: Struct., Funct., Bioinf. 76*, 995−1006.

(20) Choi, J., Choi, S., Chon, J. K., Choi, J., Cha, M. K., Kim, I. H., and Shin, W. (2005) Crystal structure of the C107S/C112S mutant of yeast nuclear 2-Cys peroxiredoxin. *Proteins: Struct., Funct., Bioinf. 61*, 1146−1149.

(21) Kabsch, W. (2010) XDS. *Acta Crystallogr. D66*, 125−132.

(22) Karplus, P. A., and Diederichs, K. (2012) Linking Crystallographic Model and Data Quality. *Science 336*, 1030−1033.

(23) Evans, P. (2005) Scaling and assessment of data quality. *Acta Crystallogr. D62*, 72−82.

(24) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.

(25) Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2012) Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D68*, 368−380.

(26) Afonine, P. V., Grosse-Kunstleve, R. W., Adams, P. D., Lunin, V. Y., and Urzhumtsev, A. (2007) On macromolecular refinement at subatomic resolution with interatomic scatterers. *Acta Crystallogr. D63*, 1194−1197.

(27) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.

(28) Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res. 32*, W615−W619.

(29) Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res. 38*, W545−W549.

(30) Liao, S. J., Yang, C. Y., Chin, K. H., Wang, A. H. J., and Chou, S. H. (2009) Insights into the Alkyl Peroxide Reduction Pathway of *Xanthomonas campestris* Bacterioferritin Comigratory Protein from the Trapped Intermediate−Ligand Complex Structures. *J. Mol. Biol. 390*, 951−966.

(31) Limauro, D., Pedone, E., Galdi, I., and Bartolucci, S. (2008) Peroxiredoxins as cellular guardians in *Sulfolobus solfataricus*: Characterization of Bcp1, Bcp3 and Bcp4. *FEBS J. 275*, 2067−2077.

(32) Wakita, M., Masuda, S., Motohashi, K., Hisabori, T., Ohto, H., and Takamiya, K. (2007) The Significance of Type II and PrxQ Peroxiredoxins for Antioxidative Stress Response in the Purple Bacterium *Rhodobacter sphaeroides*. *J. Biol. Chem. 282*, 27792−27801.

(33) Horta, B. B., Oliveira, M. A. D., Discola, K. F., Cussiol, J. R. R., and Netto, L. E. S. (2010) Structural and Biochemical Characterization of Peroxiredoxin Qβ from *Xylella fastidiosa* Catalytic Mechanism and High Reactivity. *J. Biol. Chem. 285*, 16051−16065.

(34) Madern, D., Ebel, C., and Zaccai, G. (2000) Halophilic adaptation of enzymes. *Extremophiles 4*, 91−98.

(35) Sako, Y., Norimichi, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T., and Maruyama, T. (1996) *Aeropyrum pernix*, gen. nov., sp. nov., a Novel Aerobic Hyperthermophilic Archaeon Growing at Temperatures up to 100 °C. *Int. J. Syst. Bacteriol. 46*, 1070−1077.

(36) Berkholz, D. S., Faber, H. R., Savvides, S. N., and Karplus, P. A. (2008) Catalytic Cycle of Human Glutathione Reductase Near 1 Å Resolution. *J. Mol. Biol. 382*, 371−384.

(37) Ådén, J., Wallgren, M., Storm, P., Weise, C. F., Christiansen, A., Schroder, W. P., Funk, C., and Wolf-Watz, M. (2011) Extraordinary μs−ms backbone dynamics in *Arabidopsis thaliana* peroxiredoxin Q. *Biochim. Biophys. Acta 1814*, 1880−1890.

(38) Wood, Z. A., Poole, L. B., Hantgan, R. R., and Karplus, P. A. (2002) Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry 41*, 5493−5504.

(39) Parsonage, D., Youngblood, D. S., Sarma, N., Wood, Z. A., Karplus, P. A., and Poole, L. B. (2005) Analysis of the Link between Enzymatic Activity and Oligomeric State in AhpC, a Bacterial Peroxiredoxin. *Biochemistry 44*, 10583−10592.

(40) Chandonia, J. M., and Brenner, S. E. (2006) The Impact of Structural Genomics: Expectations and Outcomes. *Science 311*, 347−351.

(41) Knaus, T., Eger, E., Koop, J., Stipsits, S., Kinsland, C. L., Eali, S. E., and Macheroux, P. (2012) Reverse Structural Genomics: An

Unusual Flavin-Binding Site in a Putative Protease from *Bacteroides thetaiotaomicron*. *J. Biol. Chem. 287*, 27490−27498.
(42) Soito, L., Williamson, C., Knutson, S. T., Fetrow, J. S., Poole, L. B., and Nelson, K. J. (2010) PREX: PeroxiRedoxin classification indEX, a database of subfamily assignments across the diverse peroxiredoxin family. *Nucleic Acids Res. 39*, D332−D337.
(43) Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) Weblogo: A sequence logo generator. *Genome Res. 14*, 1188−1190.