BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# Understanding the −C−X1−X2−C− Motif in the Active Site of the Thioredoxin Superfamily: *E. coli* DsbA and Its Mutants as a Model System

Andrey Karshikoff,[†] Lennart Nilsson,*[,‡] and Nicolas Foloppe*[,§]

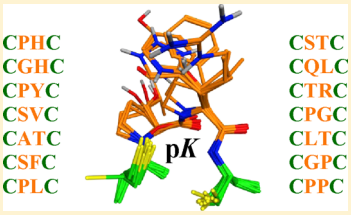
[†]Institute of Molecular Biology, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., bl. 21, Sofia 1113, Bulgaria

[‡]Department of Biosciences and Nutrition, Center for Biosciences, Karolinska Institutet, S-141 83 Huddinge, Sweden

[§]51 Natal Road, Cambridge CB1 3NY, U.K.

Supporting Information

**ABSTRACT:** *E. coli* DsbA is an intensively studied enzyme of the thioredoxin superfamily of thiol-disulfide oxidoreductases. DsbA catalyzes the disulfide bond formation and folding of proteins in the bacterial periplasm. DsbA and its mutants have highlighted the strong and puzzling influence of the −C−X1−X2−C− active site variants, found across the thioredoxin superfamily, on the ionization and redox properties of this site. However, the interpretation of these observations remains wanting, largely due to a dearth of structural information. Here, molecular dynamics simulations are used to provide extensive information on the structure and dynamics of reduced −C30−X31−X32−C33− motifs in wild type DsbA and 13 of its mutants. These simulations are combined with calculations of the p*K* of H32 and of the very low p*K* of the catalytic cysteine C30. In wild type DsbA, the titrations of C30 and H32 are shown to be coupled; the protonation states and dynamics of H32 are examined. The thiolate of C30 is stabilized by hydrogen bonds with the protein. Modulation of these hydrogen bonds by alteration of residue X32 has the greatest impact on the p*K* of C30, which rationalizes its higher p*K* in thioredoxin and tryparedoxin. Because of structural constrains, residue X31 has only an indirect and weak influence on the p*K* of C30. The dynamics of C30 is clearly related to its stabilizing interactions and p*K* value. Although relatively small differences between p*K*s were not reproduced in the calculations, the major trends are explained, adding new insights to our understanding of enzymes in this family.




The redox homeostasis of cells is highly regulated, and understanding the biochemical factors and interactions which control these redox properties is an area of intense research.[1,2] Equilibria between the thiol and disulfide forms of cysteine residues are major and essential components of the cellular redox environment. These equilibria are regulated by thiol-disulfide oxidoreductases. Many of these redox enzymes are structurally related to thioredoxin (Trx) and thus comprise the thioredoxin superfamily.[3,4] Most oxidoreductases of the thioredoxin superfamily contain the -C-X1-X2-C- motif in their active site, which is directly involved in the chemistry controlling the redox state, dithiol or disulfide, of their substrates. In addition to thioredoxin, enzymes in this superfamily include tryparedoxins,[5] glutaredoxins,[6] protein disulfide isomerase (PDI),[7] and DsbA.[8] These enzymes have markedly different reactivity, substrate selectivity, and redox potentials. Considerable efforts are directed at understanding the molecular determinants of these properties and why they vary across the thioredoxin superfamily.[3,9−11]

It is well recognized that the p*K* of the N-terminal cysteine of the −C−X1−X2−C− motif is critical to the chemistry and redox properties of the thiol-disulfide oxidoreductases.[3,10−12] This p*K* is lower than the p*K* of a free cysteine, especially in DsbA. Measured values of the p*K* of this N-terminal active site cysteine range from ∼7.2 in tryparedoxin[5] and ∼6.7 in Trx[13] to between 3.28 and 3.5 in *E. coli* DsbA.[8,10,11,14] The thiol p*K* is a key determinant of the chemistry of the sulfur atom, as a lowered p*K* affects both nucleophilicity and leaving group ability of the associated thiolate.[13,15−17] Thus, the p*K* of the N-terminal cysteine of the −C−X1−X2−C− motif is an important factor contributing to modulate the redox properties of the corresponding enzymes.[9,14,18,19] The factors which control this p*K* are debated.[3,9,20−22] They could involve electrostatic interactions with other charged side chains,[23] the influence of a helix macro dipole,[20,24] or direct hydrogen bonds which stabilize the thiolate sulfur.[22] Recent studies with glutaredoxins[22,25,26] and model systems[27] have highlighted the importance of direct hydrogen bonds to the thiolate. This suggests that residues X1 and X2 of −C−X1−X2−C− could both strongly influence the cysteine p*K* via direct contacts. Yet, the respective roles of X1 and X2 are poorly understood.

There is a wealth of experimental data pertaining to the p*K* of the active site N-terminal cysteine (C30) of *E. coli* DsbA and its mutants.[8,10,11,14] DsbA catalyzes the formation of protein disulfide bridges in the periplasm of gram-negative bacteria[2,28] and is one of the most investigated model systems to

**Received:** April 22, 2013
**Revised:** June 24, 2013
**Published:** July 23, 2013

understand the chemistry and function of thiol-disulfide oxidoreductases. The active site of wild type (WT) *E. coli* DsbA contains the redox-active −C30−P31−H32−C33− sequence. DsbA is a very oxidizing enzyme,[2,12,29] which is partially attributed to the unusually low p*K* of its cysteine C30.[11,12,14,18] A value of 3.5 was initially measured for the p*K* of C30,[8] essentially consistent with subsequent measurements.[10,11,14]

Crucially, the p*K* of C30 has also been determined for a good number of *E. coli* DsbA mutants,[11,14,30] providing a wealth of experimental data for theoretical analysis. Most of these mutants were obtained by changing residues −X31−X32− in the −C30−X31−X32−C33− active site motif, to mimic the active site of other enzymes of the Trx superfamily,[11] or by random mutagenesis.[14] These −X31−X32− mutants all increased the p*K* of Cys30 to various degrees, with some mutants resulting in striking upward p*K* shifts. The electrostatic contribution of E37 and E38 to the p*K* of C30 was also investigated by experimental mutagenesis, but these mutants shifted the p*K* of C30 by 0.5 pH units at most.[30] The −X31−X32− mutations are arguably the most interesting since they modify the core of the active site, provide models for other enzymes, and offer a systematic series to study the factors influencing the p*K* of C30. Also, a correlation was found between the redox potential and the p*K* of C30 for the −X31−X32− mutants of DsbA,[11,14] consistent with a comparison of wild type DsbA to thioredoxin.[18] Therefore, the −X31−X32− DsbA mutants are of clear interest for the molecular biology of redox thiol-disulfide regulation. Unfortunately, the dearth of structural information on the −X31−X32− mutants of DsbA has limited the interpretation of their properties. However, several experimental structures of reduced wild type DsbA are available, including X-ray[20] and NMR[31] structures. X-ray structures of three H32 mutants in the oxidized form have also been discussed.[32]

Several theoretical investigations addressed the ionization behavior of the active site of wild type and mutated DsbA. Pioneering electrostatic calculations exploited X-ray structures to yield important insights about the influence of the protein environment on the ionization of C30.[18,33,34] Taking the p*K* value of C30 of the wild type *E. coli* DsbA as the reference, the shifts of the ionization constant of this residue were evaluated for different mutants, after generation of multiple side chain rotamers for the -X1-X2- residues in a rigid protein framework.[9] Other calculations gave a p*K* of 3.7 for C30 in wild type *E. coli* DsbA[35] but did not analyze the physical factors determining the ionization behavior of C30.

The physical factors determining the ionization equilibria in proteins appear largely understood,[36−39] although the prediction of strongly shifted p*K* values, for instance for groups buried in the protein interior, still faces serious difficulties. Different computational approaches reproduce the experimental values with an average error of about 0.6 pH units.[40] The error in predicting p*K* values for buried sites is higher, about 1.0 pH units.[41−43] This seemingly large error corresponds to relatively small changes of the free energy of protonation/deprotonation, i.e., a 0.5 unit change in p*K* corresponds to a 1 kcal/mol free energy change. Therefore, accurate predictions of p*K* values in proteins remain challenging because they require the evaluation of a delicate balance between various factors, the result of which is small. Despite these limitations, current computational methods and force fields for calculation of charge−charge interactions and desolvation contributions (the main contributors to p*K* shifts) yield reasonable results, when based on sound structural information. Indeed, the computed p*K* values are very sensitive to the three-dimensional structure used to perform the calculations.[22,44−48] The p*K* shifts caused by side chain conformational transitions, and the corresponding change in solvent accessibility of the titratable groups, can result in p*K* shifts as large as 2 pH units.[49] Thus, a major avenue toward improving the accuracy of p*K* calculations is to address the influence of conformational flexibility of proteins.

Molecular dynamics (MD) simulation is a powerful method for investigating the structural flexibility and how it affects ionization equilibria. Simulations in explicit solvent can yield convincing dynamic structural models, including local adjustements resulting from one or a few mutations. Coupling p*K* calculations with MD gives in general a better agreement with the experimental data,[22,50−56] but there are also limitations. Insufficient conformational sampling would bias the resulting conformational ensemble toward the starting conformer. This, however, is a lesser issue with modern simulations over several tens of nanoseconds. Such simulation length tends to be adequate to sample side chains conformations in the active site of enzymes of the thioredoxin superfamily.[25,48,52] Another concern is the selection of fixed protein protonation state for the simulations. Indeed, p*K* calculations involve pH titrations, and the true conformational spaces corresponding to different pH regions may differ. This issue has been approached by constant pH MD simulations,[57] by introducing partial protonation[58] or by allowing the protonation state to fluctuate.[41,59] Replica exchange MD has also been used to better reflect the interrelationship between the protonation state and the conformational space.[60] Recently, we have employed another approach based on multiple pH regime MD simulation.[61] It produces MD trajectories with different fixed protonation states of the protein. This extends the structural sampling by exploring conformers expected to be populated at different pH. This approach was applied in the present work, to the key H32 in the active site of wild type DsbA (−C30−P31−H32−C33−) and its PDI-like mutant (−C30−G31−H32−C33−).

Here, we present MD simulations in explicit solvent of reduced wild type *E. coli* DsbA and 13 of its reduced active site mutants. It provides additional insights in the structure and electrostatics of the wild type active site, especially regarding the conformational dynamics and interactions of H32. It also offers novel structural information on the mutants which, to our knowledge, is otherwise not available. This is valuable since these mutants have played a key role in developing DsbA as a major model system for the thioredoxin superfamily. The structural models of the mutants are analyzed with a view to understand their influence on the p*K* of C30. It emerges that residue X31 does not significantly stabilize the thiolate, contrary to residue X32 which can hydrogen bond the thiolate via its backbone $N_{32}$−H, and sometimes via its side chain. The ability of various side chains at X32 to reach and hydrogen bond the thiolate is discussed. The simulated structures are used as the basis for p*K* calculations. The agreement between experimental and calculated p*K*s is sufficient to allow a detailed explanation of the low p*K* of Cys30 in wild type DsbA and an analysis of how the titrations of C30 and H32 are coupled. The calculations suggest an unexpected low p*K* for H32, possibly explaining older experimental data. For comparison, calculations based on a NMR structure of reduced DsbA are also presented. Specific insights are gained into the balance of


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

**Table 1. Overview of Simulated DsbA Systems, and Measured and Calculated p*K*s for C30**

| −C30−X31−X32−C33− sequence | MD length (ns) | $pK_{exp}$ [a] | $pK_{1/2}$ [b] | references for measured $pK_{exp}$ | comment |
|---|---|---|---|---|---|
| −C30−P31−H32−C33− | 240 | 3.28−3.5[c] | 3.73 ± 0.06 (1.17) | 8, 10, 11, 14 | *E. coli* wild type; was simulated with H32 charged (120 ns) or neutral (120 ns) |
| −C30−G31−H32−C33− | 244 | 3.71 | NA[d] | 11 | meant to mimick the active site of PDI; was simulated with H32 charged (132 ns) or neutral (112 ns) |
| −C30−P31−Y32−C33− | 120 | 3.75 | 5.74 ± 0.04 (0.57) | 11 | meant to mimick the active site of glutaredoxins |
| −C30−S31−V32−C33− | 120 | 4.23 | 5.75 ± 0.04 (0.55) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence |
| −C30−A31−T32−C33− | 120 | 4.34 | 5.87 ± 0.04 (0.60) | 11 | meant to mimick the active site of some yeast thioredoxin reductases |
| −C30−S31−F32−C33− | 120 | 4.34 | 5.59 ± 0.07 (0.59) | 14 | obtained by random mutagenesis of the −C30-X1−X2−C33− active site sequence |
| −C30−P31−L32−C33− | 120 | 4.42 | 5.79 ± 0.03 (0.54) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence |
| −C30−S31−T32−C33− | 120 | 4.45 | 5.62 ± 0.09 (0.61) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence |
| −C30−Q31−L32−C33− | 120 | 4.59 | 5.87 ± 0.04 (0.56) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence |
| −C30−T31−R32−C33− | 120 | 4.76 | 5.22 ± 0.08 (0.72) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence, but −C−T−R−C− is present in human thioredoxin reductase 1 |
| −C30−P31−G32−C33− | 120 | 4.85 | 5.46 ± 0.05 (0.60) | 11 | mutant initially designed to test the influence of the H32 side chain by comparison to the wild type |
| −C30−L31−T32−C33− | 120 | 4.86 | 5.85 ± 0.05 (0.54) | 14 | Obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence |
| −C30−G31−P32−C33− | 120 | 6.21 | 7.00 ± 0.01 (0.42) | 11 | meant to mimick the active site of a thioredoxin |
| −C30−P31−P32−C33− | 120 | 6.73 | 7.05 ± 0.02 (0.39) | 14 | obtained by random mutagenesis of the −C30−X1−X2−C33− active site sequence, but −C−P−P−C− is present in tryparedoxins. |

[a]Experimentally measured p*K*s. The systems are listed by increasing order of experimentally measured $pK_{exp}$ for C30. [b]Average calculated $pK_{1/2}$ for C30 ± standard error of the mean; the standard error of the mean was calculated based on sample means obtained on contiguous blocks of 10 ns each. The standard deviation is also provided in parentheses. This work uses the notation $pK_{1/2}$ to refer to the calculated ionization equilibria, and the notation $pK_{exp}$ to refers to experimental observations. [c]Range of measured $pK_{exp}$ values reported for C30 in wild type DsbA; the present work uses a reference value of 3.42 for the $pK_{exp}$ of C30 in wild type DsbA since it was obtained alongside several other $pK_{exp}$ values for the DsbA mutants.[14] [d]Not available.



dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

factors shifting the p$K$ of C30 in DsbA mutants. It is found that the dynamics of C30 is essential to explain the larger p$K$ differences across the series of mutants, with direct implications for thioredoxin and tryparedoxin.

## COMPUTATIONAL METHODS

**Preparation of Initial Structural Models.** The initial model for wild type (WT) DsbA was derived from the X-ray structure of reduced WT *E. coli* DsbA (PDB entry 1A2L, chain A). The missing atoms in the X-ray structure were built with the program MODELER.[62] This curated WT structure was used as the basis to build all of the 13 mutants by comparative (homology) modeling, again with the program MODELER. The mutants differ only by a change of the two −X31−X32− residues in the active site −C30−X31−X32−C33− sequence. With such minor changes, comparative modeling is expected to provide fairly reliable initial structural models for the mutants. For each mutant, 10 homology models were generated with MODELER, but little variation across models was observed. The model with the best score according to the MODELER objective function was selected. Visual inspection of the selected models confirmed that they were reasonably built. Then, the −C30−X31−X32−C33− sequence was energy-minimized while keeping the rest of the protein fixed, with the program CHARMM[63] and its all-atom protein force-field.[64,65] For the wild type (−C30−P31−H32−C33−) and the PDI-like mutant (−C30−G31−H32−C33−), H32 was prepared in both neutral (proton on $N_\delta$ of imidazole) and positively charged states. Altogether, this led to 16 initial structural models, which were then subjected to conformational sampling with molecular dynamics.

**Molecular Dynamics Simulations.** The same MD protocol was applied to the 16 systems, prepared as explained above. All MD simulations were performed with the program CHARMM and its all-atom protein force-field, a dielectric constant of 1.0, and atom-based nonbonded interactions truncated beyond 14 Å with force-shift.[66] Nonbonded lists were maintained to 16 Å and updated heuristically. There is ample evidence that the force-shift spherical cutoff used in the present work performs well, as well as the alternative PME schemes also in use. This is documented by detailed studies, which compared force-shift spherical cut-offs to PME and concluded that both schemes work equally well.[67−69] In addition, a recent study of a DNA oligomer in solution obtained very similar results with either CHARMM and the force-shift cutoff or AMBER and PME.[70] Also, the force-shift spherical cutoff was used across a series of studies on proteins in the Trx superfamily. These studies characterized the structure, dynamics, and p$K_a$'s of the ionized active-site cysteines, alongside detailed comparisons to experimental observables.[22,25,52,71,72] In every case, excellent agreement was found between simulations and experiment, even in predictive mode. This is a very strong validation of the force-shift spherical cutoff. Therefore, we continue to use this cutoff in the present work for reasons of consistency and to allow for comparisons across different proteins of the Trx superfamily.

C30 was treated as a thiolate in all systems, using the appropriate CHARMM nonbonded parameters.[22] All other titratable groups were assigned their conventional protonation state at pH 7, except His32 which was either neutral or positively charged. The initial orientation of the H32 side chain was such that the imidazole $N^{\delta}_{32}$−H pointed toward the sulfur of C30.

All simulations were performed in explicit water using the TIP3P model[73] with sodium counterions added to neutralize every system. Periodic boundary conditions were applied with a rhombohedric dodecahedron primary cell. The distance normal to two opposing faces of the dodecahedron was 80 Å, allowing for ample solvation of the protein. Covalent bonds involving hydrogen were constrained with SHAKE.[74]

First, the solvent was energy minimized keeping the protein fixed. In a second round of energy minimization, both the solvent and the protein were allowed to relax. Each system was then subjected to MD simulations, using the leapfrog integrator and a 0.002 ps time step. Heating to 300 K was performed in 30 ps, with the protein atoms harmonically constrained to their initial position with a force constant of 2.0 kcal/mol/Å$^2$. Each simulation was pursued at 300 K in the NVE ensemble for 20 ps, with the constraints on the protein kept. This was followed by unrestrained MD in the NVE ensemble during at least 120 ns for every system (Table 1).

**Structural Analyses.** We follow the standard recommendation for structural descriptions.[75] For torsion angles, the gauche minus ($g^-$) and gauche plus ($g^+$) ranges are defined as $-60° \pm 30°$ and $60° \pm 30°$, respectively.

Hydrogen bonds to the sulfur were identified using a distance cutoff of 4.0 Å between sulfur and heavy atom hydrogen bond donor D, combined with angular conditions S···H−D $\geq 120.0°$ and C−S···H $\geq 70.0°$, consistent with previous studies.[25,48,52]

**Ionization Equilibria.** The ionization equilibria of the individual titratable groups were calculated using the multiple hydrogen location approach.[76] This approach distinguishes alternative positions of the polar hydrogen atoms, such as binding to one or the other oxygen atoms of carboxyl groups, histidine tautomers, hydroxyl group rotamers. All these are treated as microscopic states of the individual functional groups. The output of the computations is the pH dependence of the probability for a given microscopic state to be occupied, $p_{i\alpha}$ (pH), where the index $i$ enumerates the ionizable groups in the protein, whereas $\alpha$ indicates its microscopic state (for instance protonated or deprotonated). That is, the approach gives the titration curves of the individual functional groups, rather than their p$K$ values. The method applied in the current work is as described in refs 47, 61, 76, and 77. The calculations for DsbA were performed in an initial grid box ($99 \times 99 \times 99$) with a grid size of 1.69 Å. The charge−charge interactions were calculated after focusing to a box with a grid size of 0.65 Å, whereas the calculations of the self-energies were performed after a third focusing on the geometrical centers of the individual ionizable groups to boxes with a grid size of 0.25 Å. The partial charges used in the p$K$ calculations, including those of the titratable functional groups, were taken from the CHARMM22 all-atom force field.[64] The van der Waals radii were taken from Rashin et al.[78] The solvent probe radius and the thickness of the ion exclusion layer were 1.4 Å and 2.0 Å, respectively. Relative dielectric constants of 78 for the solvent and 4 for the protein were used, together with an ionic strength of 0.12 M and a temperature of 298 K. The explicit treatment of the protein dynamics with the MD simulations prior to the p$K$ calculations could in principle call for a lower value of the protein dielectric constant, although the present value of 4 is already toward the lower end of the possible range. A value of 4 was retained since (i) it is empirically consistent with the other parameters involved in the p$K$ calculations, (ii) it yielded good agreement of the calculations with the well-established


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

measured p$K$ of wild type DsbA, and (iii) the interest in the present work resides more with the relative than with the absolute values of the p$K$ of C30 across the various active sites. The model p$K$ values were taken from ref 79. For cysteines, a model p$K$ of 8.22 was used.[80] For histidines, 3 model p$K$ values were used (7.1, 7.7, and 10.8) to take into account the different proton affinities, alternative protonation sites, and second deprotonation (which makes histidine negative), as previously described.[76]

**Coupling p$K$ Calculations with MD Simulation.** Protein conformers (snapshots) were collected each 100 ps during the 20 to 70 ns period of each MD simulation (501 snapshots per system). The populations of the microscopic states of the individual groups are obtained by averaging over the number of snapshots,

$$M\text{: } <p_{i\alpha}(\mathrm{pH})> = \frac{1}{M}\sum_{k=1}^{M} p_{i\alpha,k}(\mathrm{pH})$$

As discussed earlier,[61] one takes into account possible pH induced conformational changes by generating snapshots from MD runs with different initial protonation states of the protein that correspond to different pH values. Since this investigation is concerned with electrostatic interactions in the active site of DsbA, particular care was paid to the protonation state of H32 because it is in close vicinity of the catalytic C30 and is known to markedly stabilize the reduced form of *E. coli* DsbA.[32] Therefore, two sets of snapshots were used to calculate the titration curve of C30 in WT DsbA: set **SP** from the simulation with H32 protonated and set **SN** with H32 neutral. The final $p_{i\alpha}$ is obtained by weighting $<p_{i\alpha,k}(\mathrm{pH})>$ calculated from the two sets according to the deprotonation curve of H32:

$$p_{i\alpha}(\mathrm{pH}) = W(\mathrm{pH})<p_{i\alpha}(\mathrm{pH})>_{\mathrm{SP}} + (1 - W(\mathrm{pH}))<p_{i\alpha}(\mathrm{pH})>_{\mathrm{SN}} \quad (1)$$

If we consider the ionization of C30 only, the above general expression becomes

$$p_{(\mathrm{C30})\alpha}(\mathrm{pH}) = W(\mathrm{pH})<p_{(\mathrm{C30})\alpha}(\mathrm{pH})>_{\mathrm{SP}} + (1 - W(\mathrm{pH}))<p_{(\mathrm{C30})\alpha}(\mathrm{pH})>_{\mathrm{SN}} \quad (2)$$

where $\alpha$ denotes the deprotonated microscopic state of the thiol group and $p_{(\mathrm{C30}),\alpha}$ is the population of this microscopic state. The midpoint of $p_{(\mathrm{C30}),\alpha}(\mathrm{pH})$ is then considered as the p$K$ value of this group. In the above equation, the population $<p_{(\mathrm{C30})\alpha}(\mathrm{pH})>_{\mathrm{SP}}$ from set **SP** dominates at low pH, while $<p_{(\mathrm{C30})\alpha}(\mathrm{pH})>_{\mathrm{SN}}$ dominates at high pH.

Since the two sets **SP** and **SN** give different ionization curves (see Figure 5 and the associated discussion) the weight, $W(\mathrm{pH})$, used in the equation above is taken as the average of the corresponding deprotonation curves of H32:

$$W(\mathrm{pH}) = (<p_{(\mathrm{H32})\alpha}(\mathrm{pH})>_{\mathrm{SP}} + <p_{(\mathrm{H32})\alpha}(\mathrm{pH})>_{\mathrm{SN}})/2 \quad (3)$$

The microscopic state $\alpha$ in the above equation corresponds to the protonated imidazole of H32.

## RESULTS AND DISCUSSION

Table 1 presents the investigated DsbA systems. In this work, the experimentally measured p$K$ are denoted as $pK_{exp}$, while $pK_{1/2}$ refers to the calculated ionization equilibria. Several of the 14 investigated sequences were meant to model the active sites of other redox enzymes,[11] namely, PDI, glutaredoxins, thioredoxins, and some thioredoxin reductases. The −C30−P31−P32−C33− sequence produced by random mutagenesis[14] turned out to be biologically relevant since it is also present in tryparedoxins.[5] The $pK_{exp}$ of the N-terminal cysteine C40 in tryparedoxin is 7.2, close to its counterpart in the DsbA mutant ($pK_{exp}$ = 6.73). The −C30−T31−R32−C33− sequence, also initially generated by random mutagenesis, is present in human thioredoxin reductase 1.[81]

We first present an MD-based structural analysis of the reduced −C30−X31−X32−C33− motifs which, to our knowledge, offers the first opportunity to visualize the structure and dynamics of these DsbA mutants. It also provides qualitative insights into the factors stabilizing the thiolate. This is followed by a detailed quantitative analysis of the ionization properties of C30 and H32 in wild type DsbA, with electrostatic calculations based on the MD snapshots. For comparison, an equivalent analysis is performed using an NMR structure of reduced DsbA. Then, the quantitative analysis of the ionization properties of C30 is extended to the mutants of the −C30−X31−X32−C33− motif. This provides a context to discuss the factors controlling the p$K$ of C30, with implications for some key functional properties of DsbA and other enzymes represented by the mutants. It is also an opportunity to reflect on specific methodological aspects underpinning p$K$ calculations.

**Structure and Dynamics of the Reduced −C30−X31−X32−C33− Motifs in Wild Type DsbA and Its Mutants.** Figure 1 illustrates the structure and dynamics of the wild type −C30−P31−H32−C33− motif, with H32 protonated (lower pH) or neutral (higher pH). Exploring these two protonation states of H32 is relevant to the p$K$ calculations, in particular regarding the conformational dynamics of H32, as further elaborated in the next section. Figure 1 shows that the side chains of the cysteines of −C30−P31−H32−C33− adopt well defined conformations, with $\chi_1$ of C30 ($\chi_1^{30}$) and $\chi_1$ of C33 ($\chi_1^{33}$) populating almost exclusively the *trans* and $g^-$ orientations, respectively (Table S1 in Supporting Information). Figure 1 also shows that these cysteine conformations are key to the formation of hydrogen bonds stabilizing the thiolate of C30 (Table 2). The $\chi_1^{30}$ = *trans* conformation brings the sulfur of C30 ($S_{30}^{\gamma}$) to a position to accept hydrogen bonds from the backbone amide N−H groups of H32 and C33. In addition, $\chi_1^{33} = g^-$ for C33 allows the thiol of C33 to donate a hydrogen bond to the thiolate of C30. In the wild type sequence, the thiolate receives a fourth hydrogen bond from the imidazole side chain of H32. As would be expected, this fourth hydrogen bond is formed more frequently when H32 is protonated than when it is neutral (Figure 1 and Table 2). Several groups of the imidazole of H32 could *a priori* hydrogen bond the thiolate, namely, $N^{\delta}_{32}$−H, $N^{\varepsilon}_{32}$−H (protonated imidazole only), $C^{\delta}_{32}$−H, and $C^{\varepsilon}_{32}$−H. The $C^{\delta}_{32}$−H and $C^{\varepsilon}_{32}$−H groups were considered potential hydrogen bond donors since it is well documented that polarized C−H groups have some hydrogen bond donor character;[82,83] that includes aromatic C−H groups,[84,85] especially if they are polarized by electronegative heteroatoms like the nitrogen atoms of an imidazole moiety. Indeed, analysis of the simulations indicates that both $N^{\delta}_{32}$−H and $C^{\delta}_{32}$−H (but not $N^{\varepsilon}_{32}$−H nor $C^{\varepsilon}_{32}$−H) hydrogen bond the thiolate. From an X-ray crystallography point of view, both orientations of the imidazole ($N^{\delta}_{32}$−H or $C^{\delta}_{32}$−H hydrogen bond to the thiolate) are isomorphous with respect to the imidazole orientation (Figure S1, Supporting


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

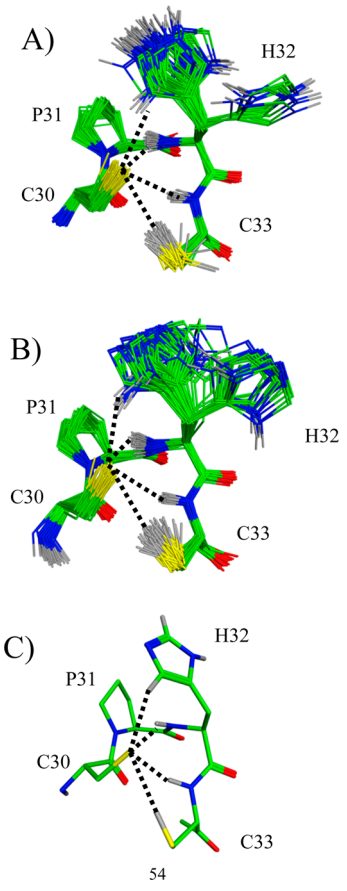



**Figure 1.** Structure and dynamics of the active site −C30−P31−H32−C33− motif in reduced wild type *E. coli* DsbA from MD simulations. The MD snapshots taken every nanosecond for −C30−P31−H32−C33− are shown with a protonated H32 (panel A) or a neutral H32 (panel B). Panel C illustrates a hydrogen bond between the thiolate of C30 and the $C^{\delta}$−H group of the H32 imidazole side chain. For clarity, only selected hydrogen atoms are shown (light gray). The hydrogen bonds are depicted as black dotted lines. Carbon atoms are green, nitrogen atoms blue, oxygen atoms red, and sulfur atoms yellow.

Information); therefore, it is likely that a $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond would not be incompatible with the X-ray data used to build the structure of reduced *E. coli* DsbA (PDB entry 1A2L). When the $C^{\delta}_{32}$−H hydrogen bonds the thiolate, the $N^{\delta}_{32}$−H is exposed to the bulk solvent and interacts with the explicit waters around the protein. The approximate frequency of interconversion between $N^{\delta}_{32}$−H or $C^{\delta}_{32}$−H as hydrogen bond donors to $S^{\gamma}_{30}$ can be inferred from plots of the distances $N^{\delta}_{32}$···$S^{\gamma}_{30}$ and $C^{\delta}_{32}$···$S^{\gamma}_{30}$ (Figure S3, Supporting Information). A distance of ∼3.5 Å between $S^{\gamma}_{30}$ and $N^{\delta}_{32}$ or $C^{\delta}_{32}$ corresponds to a hydrogen bond formed with $N^{\delta}_{32}$−H or $C^{\delta}_{32}$−H, respectively. As can be seen from Figure S3 (Supporting Information), there were frequent exchanges between $N^{\delta}_{32}$−H and $C^{\delta}_{32}$−H as the hydrogen bond donors to $S^{\gamma}_{30}$. This indicates that the observation of a $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond does not result from the simulations being accidentally trapped in this state, supporting further the notion of a $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond. Intriguingly, the simulations suggest that the hydrogen bond with $C^{\delta}_{32}$−H is formed more frequently than with $N^{\delta}_{32}$−H when the imidazole is protonated (Table 2). Thus, it would be interesting if the existence of the $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond could be probed by NMR, especially because the rate of occurrence of a $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond in the simulations might be force-field dependent. Yet, the stabilization of the thiolate by a C−H···S hydrogen bond would be reminiscent of the situation in glutaredoxins, where such an interaction has been proposed to involve the tyrosine or phenylalanine at X2 in −C−P−Y/F−C−,[25] typical of glutaredoxins. The aromatic $C^{\delta}$−H group of Y/F at X2 may donate a weak nonconventional hydrogen bond to the thiolate (Figure S2B, Supporting Information). Indeed, the measured p*K* of C30 in the DsbA mutant with a glutaredoxin-like sequence was 3.75, only marginally higher than the $pK_{exp}$ of ∼3.5 measured in WT DsbA (Table 1).

Overall, the structural analysis of wild type −C30−P31−H32−C33− supports the notion that its cysteines populate one overwhelmingly dominant conformation ($\chi_1^{30}$ = *trans*, $\chi_1^{33}$ = $g^-$) such that the thiolate is chelated by four frequently formed hydrogen bonds, essentially consistent with previous observations by X-ray crystallography.[20] This stabilization of the thiolate by local hydrogen bonds provides a strong qualitative explanation for the low p*K* of C30. In addition, the total count of hydrogen bonds to the thiolate in the wild type is larger than that with the mutants (Table 2 and Figure 2), consistent with C30 in the mutants having a p*K* always higher than that in the wild type.

The structural theme observed with the wild type is partly transferable to the mutants, although with some informative variations (Figure 2 and Figure S2, Supporting Information). Thus, C33 maintains its overwhelming preference for $\chi_1^{33} = g^-$ across all mutants (Table S1, Supporting Information). C30 also remains essentially in $\chi_1^{30}$ = *trans*, but this trend is weakened for the −C30−G31−P32−C33− (Trx-like) and −C30−P31−P32−C33− (tryparedoxin-like) mutants (Figure 3). The presence of P32 in these mutants removes the possibility of a $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond, thereby weakening the preference for $\chi_1^{30}$ = *trans*. This is consistent with $\chi_1^{30}$ populating the $g^-$ and "other" ranges more frequently in −C30−G31−P32−C33− and −C30−P31−P32−C33− than in the other sequences. Removal of the $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond explains qualitatively why the p*K* of C30 is clearly higher in −C30−G31−P32−C33− and −C30−P31−P32−C33− than those in other sequences (Table 1). Interestingly, the p*K* of C30 is 0.52 units higher in −C30−P31−P32−C33− than in −C30−G31−P32−C33− (Table 1); however, it cannot be rationalized by the removal of a further hydrogen bond between the thiolate and $N_{31}$−H. Indeed, a hydrogen bond between $N_{31}$−H and $S^{\gamma}_{30}$ is virtually never formed (Table 2 and Figure 2), even when X31 is not a proline. Nevertheless, in the simulations C30 populates $\chi_1^{30}$ = *trans* significantly less in −C30−P31−P32−C33− (73% of the time) than in −C30−G31−P32−C33− (96%), which in turn decreases the frequency of the $N_{33}$−H···$S_{30}{}^{\gamma}$ and $S^{\gamma}_{33}$−H···$S^{\gamma}_{30}$ hydrogen bonds (Figure 2 and Table 2). This subtle dynamical effect and its influence on thiolate stabilization could not be inferred from a straightforward examination of the sequences or from the initial homology models. It was only revealed when simulating the dynamics of the active sites and would presumably be difficult to detect experimentally. In sum, the structural models generated by the MD simulations (Figure 2) and the associated hydrogen bond frequencies to the thiolate (Table 2) offer a convincing qualitative explanation for the markedly higher p*K* of C30 in −C30−G31−P32−C33− and −C30−P31−P32−C33− than those in other sequences. Similar arguments are likely to be transferable to thioredoxins and probably rationalize


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

**Table 2. Hydrogen Bonds[a] between the Active Site Thiolate Sulfur and the −C30−X31−X32−C33− Motif in MD Simulations of DsbA**

| | backbone N−H hydrogen bond donors | | | side chain hydrogen bond donors | | | |
|---|---|---|---|---|---|---|---|
| sequence[b] | $N_{31}$−H | $N_{32}$−H | $N_{33}$−H | side chain of X31[d] | side chain of X32[d] | thiol $S^{\gamma}_{33}$−H of C33 | total[f] |
| −C30−PH−C33− H32 protonated | NA[c] | 97.2 | 88.6 | NA[c] | 44.5 ($C^{\delta}_{32}$−H)<br>8.0 ($N^{\delta}_{32}$−H)<br>0.0 ($N^{\varepsilon}_{32}$−H)<br>0.0 ($C^{\varepsilon}_{32}$−H) | 40.3 | 278.6 |
| −C30−PH−C33− H32 neutral | NA[c] | 94.3 | 83.9 | NA[c] | 16.9 ($C^{\delta}_{32}$−H)<br>16.8 ($N^{\delta}_{32}$−H)<br>0.0 ($C^{\varepsilon}_{32}$−H) | 44.9 | 258.8 |
| −C30−GH−C33− H32 protonated | 0.0 | 96.4 | 86.2 | NA[c] | 39.9 ($C^{\delta}_{32}$−H)<br>7.0 ($N^{\delta}_{32}$−H)<br>0.0 ($N^{\varepsilon}_{32}$−H)<br>0.0 ($C^{\varepsilon}_{32}$−H) | 39.0 | 268.5 |
| −C30−GH−C33− H32 neutral | 0.1 | 95.2 | 83.6 | NA[c] | 15.8 (H−$C^{\delta}_{32}$)<br>10.2 (H−$N^{\delta}_{32}$)<br>0.0 (H−$C^{\varepsilon}_{32}$) | 40.1 | 204.9 |
| −C30−PY−C33− | NA[c] | 92.5 | 86.9 | NA[c] | NA[c] | 47.2 | 226.6 |
| −C30−SV−C33− | 0.0 | 90.5 | 83.1 | 0.0 ($O^{\gamma}_{31}$−H) | NA[c] | 43.8 | 217.4 |
| −C30−AT−C33− | 0.0 | 84.4 | 88.9 | NA[c] | 3.0 ($O^{\gamma}_{32}$−H) | 48.5 | 224.8 |
| −C30−SF−C33− | 0.0 | 88.5 | 84.2 | 0.0 ($O^{\gamma}_{31}$−H) | NA[c] | 46.0 | 218.7 |
| −C30−PL−C33− | NA[c] | 92.0 | 87.5 | NA[c] | NA[c] | 45.4 | 224.9 |
| −C30−ST−C33− | 0.0 | 85.2 | 90.0 | 0.0 ($O^{\gamma}_{31}$−H) | 3.0 ($O^{\gamma}_{32}$−H) | 49.4 | 227.6 |
| −C30−QL−C33− | 0.0 | 84.9 | 84.2 | 0.0 ($N^{\varepsilon}_{31}$−H) | NA[c] | 44.7 | 213.8 |
| −C30−TR−C33− | 0.0 | 87.1 | 87.0 | 0.0 ($O^{\gamma}_{31}$−H) | 1.9 ($N^{\varepsilon}_{32}$−H)<br>1.7 ($N^{\eta 2}_{32}$−H)<br>0.0 ($N^{\eta 1}_{32}$−H) | 45.9 | 223.6 |
| −C30−PG−C33− | NA[c] | 97.1 | 90.1 | NA[c] | NA[c] | 40.9 | 228.1 |
| −C30−LT−C33− | 0.0 | 81.0 | 86.6 | NA[c] | 2.8 ($O^{\gamma}_{32}$−H) | 47.8 | 218.2 |
| −C30−GP−C33− | 0.0 | 0.0 | 47.6 | NA[c] | NA[c] | 50.3 | 97.9 |
| −C30−PP−C33− | NA[c] | NA[c] | 28.3 | NA[c] | NA[c] | 33.4 | 61.7 |

[a]Reported in percent of the time the hydrogen bonds are formed during the simulations. [b]The systems are presented by increasing experimental p$K_{exp}$ for C30 (see Table 1). [c]NA not applicable, i.e., when a hydrogen bond donor is not available. [d]The hydrogen-bonding groups are given in parentheses; each hydrogen-bonding group is listed individually when a side chain contains several. For the imidazole of H32, groups $C^{\delta}_{32}$−H and $C^{\varepsilon}_{32}$−H were considered as potential hydrogen bond donors. For the phenyl of Y32 in −C30−P31−Y32−C33−, the $C^{\delta}_{32}$−H was not included as hydrogen bond donor; see the main text and Figure S3B (Supporting Information) for a discussion of this point. [f]Total is the cumulated frequencies (in percent of the time) for formed hydrogen bonds and can thus be greater than 100%.

the relatively high p$K$ of 7.2 of the catalytic C40 in tryparedoxin.[5]

Thus, the count of hydrogen bonds to $S^{\gamma}_{30}$ separates rather well the active site sequences in three p$K$ groups (Figure 4), that is, (i) the lowest p$K$ values (wild type and PDI-like), (ii) the intermediate p$K$ values (from 3.71 to 4.86, most mutants), and (iii) the highest p$K$ values (p$K$ > 6.2, mutants −C30−G31−P32−C33− and −C30−P31−P32−C33−). However, a simple count of hydrogen bonds does not immediately rationalize the small p$K$ differences between mutants in the intermediate range of p$K$ values (Figure 4). Yet, one notes that the total frequency of hydrogen bonds formed by the thiolate is lower with −C30−G31−H32−C33− (model for PDI) than with the wild type −C30−P31−H32−C33−, despite both sequences having H32. PDI is an important enzyme involved in the formation and breakage of disulfide bonds in the endoplasmic reticulum of eukaryotic cells.[1] That the PDI sequence has the p$K$ of C30 closest to that of DsbA is consistent with both enzymes having the same four hydrogen bond donors to $S^{\gamma}_{30}$, including H32. This high similarity implies that the small p$K$ difference between the two enzymes must result from very subtle factors. There is no difference between the overall conformations of the cysteine side chains between DsbA and PDI (Table S1, Supporting Information). Nevertheless, hydrogen bonds to the thiolate are formed slightly less frequently in PDI than in DsbA (Table 2). It reflects small differences in the dynamics of the two active sites, i.e., the PDI site fluctuates more, with more frequent short-lived disruptions of hydrogen bonds to $S^{\gamma}_{30}$. This is in tune with P31 in DsbA being replaced by G31 in PDI since a glycine backbone is less constrained than that of a proline. No conformational transition was observed for the backbone torsions of G31; however, greater motions in the backbone of G31would suffice to make the PDI site looser than that in DsbA. This suggests another example of the fine-tuning of a p$K$ by the protein dynamics.

Another aspect which was expected to influence the p$K$ of C30 in the mutants is potentially the various degrees of hydrogen bonding of the thiolate by the −X31−X32− side chains. However, analysis of these putative hydrogen bonds shows that they are very rarely formed (Table 2, Figure 2, and Figure S2, Supporting Information), in contrast with H32 in the wild type and PDI sequences. Polar side chains at position X31 never hydrogen bond $S^{\gamma}_{30}$, simply as a geometrical consequence of the orientation of these side chains. This is


dx.doi.org/10.1021/bi400500e | Biochemistry 2013, 52, 5730−5745

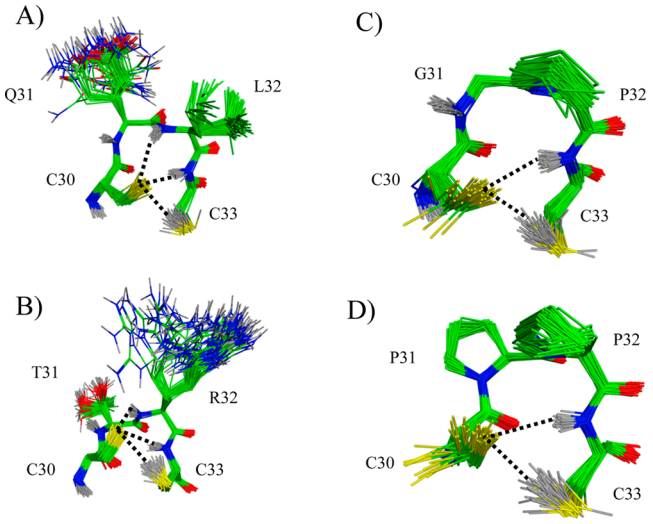



**Figure 2.** Structure and dynamics of the active site −C30−X31−X32−C33− motifs in selected DsbA mutants from MD simulations. The MD snapshots taken every nanosecond are shown for the mutants −C30−Q31−L32−C33− (panel A), −C30−T31−R32−C33− (panel B), −C30−G31−P32−C33− (panel C), and −C30−P31−P32−C33− (panel D). For clarity, only selected hydrogens are shown (light gray). The hydrogen bonds are depicted as black dotted lines. Carbon atoms are green, nitrogen atoms blue, oxygen atoms red, and sulfur atoms yellow.

illustrated with −C30−Q31−L32−C33− in Figure 2A. Q31 has a long flexible side chain with an amide hydrogen bond donor, but this side chain is projected away from C30 such that it cannot reach the thiolate. The same was observed when X31 was a serine (Figure S2, Supporting Information). Therefore, position X31 cannot be used to tune the thiolate p*K* by direct hydrogen bonding to $S^{\gamma}_{30}$ since both the side chain and the backbone of X31 are sterically precluded from forming such hydrogen bonds. However, as proposed above, position X31 can influence the p*K* indirectly by tuning the dynamics of the site.

At position X32, the mutants only explored polar side chains H32 (PDI-like), T32 and R32. T32 was present in −C30−A31−T32−C33−, −C30−S31−T32−C33−, and −C30−L31−T32−C33−, where the hydroxyl $O^{\gamma}_{32}$−H of T32 hydrogen bonded $S^{\gamma}_{30}$ ∼3% of the time. The rarity of these interactions is consistent with higher C30 p*K*s in these mutants than those in the wild type and PDI sequences. Another side chain of interest is R32 in −C30−T31−R32−C33− since this motif is present in human thioredoxin reductase 1 and because the role of charged side chains in stabilizing the thiolate in the thioredoxin superfamily has been an ongoing debate. Since R32 is positively charged, one might have expected that it would lower the p*K* of C30 relative to other mutants, but that is not the case. The measured $pK_{exp}$ of C30 is 4.76 in −C30−T31−R32−C33−, higher than in most mutants with a neutral side chain at position X32 (Table 1). Thus, the positive charge of R32 does not strongly influence the p*K* of C30. This can be understood

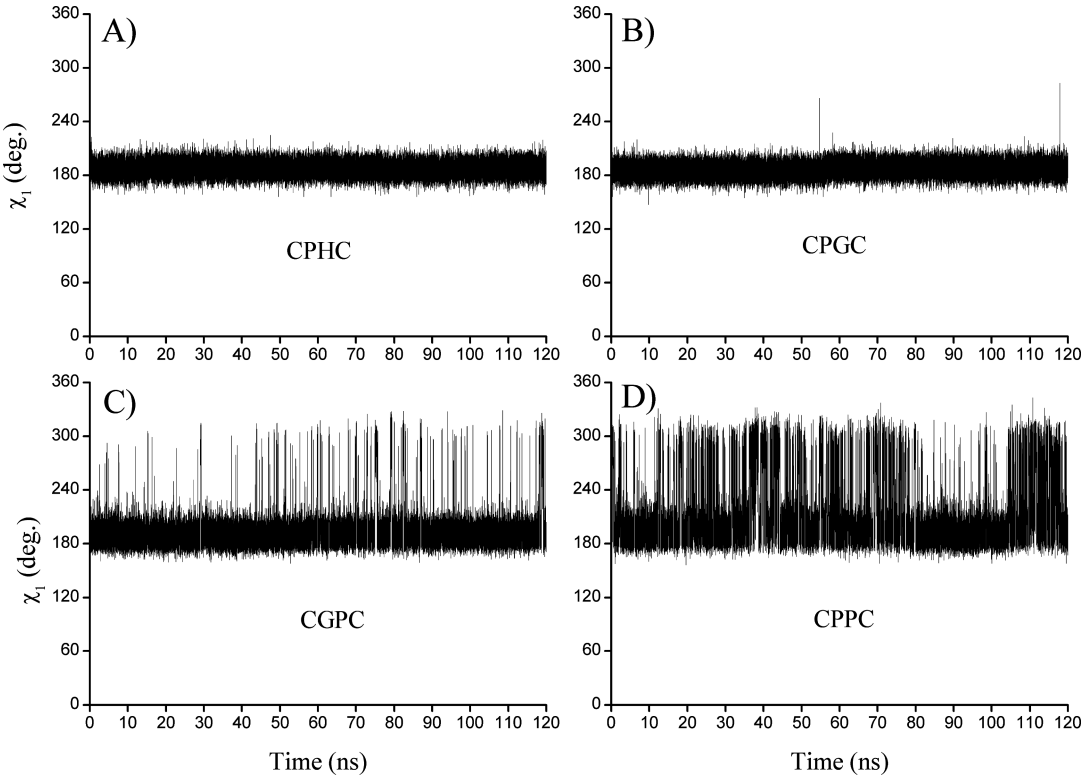



**Figure 3.** Influence of the −C−X1−X2−C− active site sequence on the dynamics of the catalytic cysteine C30. The side-chain $\chi_1$ torsion of C30 is plotted versus time, during the MD simulations of wild type DsbA (panel A, −C30−P31−H32−C33−), and the corresponding active site mutants −C30−P31−G32−C33− (panel B), −C30−G31−P32−C33− (panel C), and −C30−P31−P32−C33− (panel D). The $\chi_1$ torsion populates almost exclusively the *trans* orientation (180°) in the wild type −C−P−H−C− and mutant −C−P−G−C−. In contrast, the *trans* orientation is depleted in mutants −C−G−P−C− and −C−P−P−C− due to $\chi_1$ also populating the $g^-$ orientation (∼300°). The shift from *trans* to $g^-$ reflects the disruption of hydrogen bonds which stabilize the thiolate in the $\chi_1$ = *trans* orientation (Figures 1 and 2). This illustrates the subtle relationship between the structure and dynamics of the −C−X1−X2−C− sequence and the p*K* of its N-terminal cysteine.


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

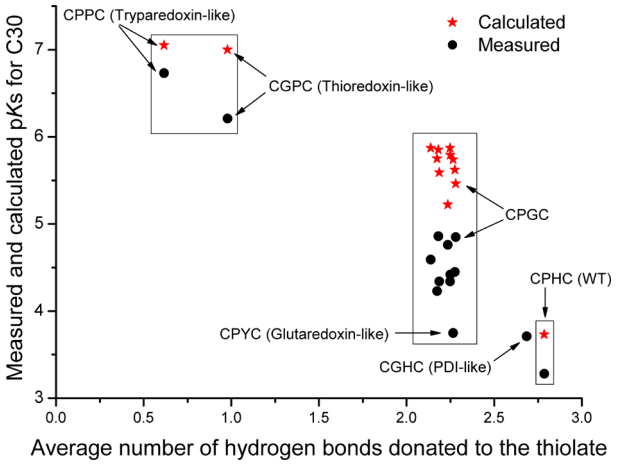



**Figure 4.** Average number of hydrogen bonds formed with the thiolate of C30 versus measured (black solid circles) and calculated (red stars) p*K*s of C30. The average number of hydrogen bonds formed with the thiolate was calculated based on the hydrogen bond donor groups listed in Table 2 and averaged over the length of the MD simulations. The mutants with the higher p*K*s for C30 are −C30−G31−P32−C33− and −C30−P31−P32−C33− (top left), with an average number of hydrogen bonds of 1.0 and 0.6, respectively. All other mutants are characterized by a higher number of hydrogen bonds (between 2.2 and 2.7) between C30 thiolate and the surrounding protein. The wild type DsbA (bottom right) has the highest average number of hydrogen bonds (2.8) and the most pronounced p*K* downshift. Overall, the plot highlights that a higher number of formed hydrogen bonds to the thiolate is associated with a lower p*K* of the catalytic cysteine C30. The glutaredoxin-like sequence would fit better with this trend if the $C^{\delta}_{32}$−H groups of the phenol of Y32 were included as hydrogen bond donors (see the main text and Figure S3B, Supporting Information). The calculated p*K* is not available for the −C−G−H−C− sequence.

qualitatively by visualizing MD snapshots of the −C30−T31−R32−C33− sequence (Figure 2B). The side chain of R32 is very mobile and spends most of the time pointing into the solvent, away from $S^{\gamma}_{30}$. In this orientation, the electrostatic interaction between the R32 guanidinium and the thiolate is mostly screened by the aqueous solvent separating these two charges. A direct hydrogen bond between these two groups was formed only ∼3% of the time. This is compatible with the p*K* of C30 in −C30−T31−R32−C33− not being particularly low and illustrates why long flexible charged side chains are not efficient at stabilizing the thiolate.

In contrast, a tyrosine or phenylanine at position X2 allows for a frequently formed $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ weakly polar interaction (Figure S2B, Supporting Information), which probably contributes to stabilize the thiolate,[25] consistent with the notably low p*K* of 3.71 for C30 in the glutaredoxin-like mutant (−C30−P31−Y32−C33−). In Table 2 and Figure 4, the $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ interaction was not included in the count of hydrogen bonds to the thiolate, possibly resulting in an artificially too low average frequency of hydrogen bonds to the thiolate for the glutaredoxin-like mutant. In Figure 4, including $C^{\delta}_{32}$−H···$S^{\gamma}_{30}$ in the hydrogen bond count would shift the glutaredoxin-like mutant closer to the PDI-like mutant and DsbA WT. This would stress further the correlation between a low p*K* of C30 and a high average frequency of hydrogen bonds to this thiolate. One might also speculate that by shielding the active site from solvent further, a bulky aromatic side chain at position 32 might enhance local electrostatic interactions.

Overall, the simulations of the various −C30−X31−X32−C33− motifs embedded in DsbA have yielded new insights regarding their structure and dynamics, which are relevant to a detailed understanding of their electrostatics. An examination of the cysteine side chain conformational dynamics and the related hydrogen bond patterns offers immediate insights regarding the extreme stabilization of the thiolate in the wild type and its lesser stabilization in the mutants. This supports the validity of the structural models derived by simulations. It is, however, important to keep in mind that the p*K* of C30 results from a balance of several factors, including solvation. Thus, explicit p*K* calculations are required to quantitatively approach the balance of these factors and the resulting p*K* values (see below).

**pH Related Conformational Flexibility of H32 in Wild Type DsbA.** A detailed treatment of the ionization properties of the active site of wild type DsbA has to consider the complication arising from two titratable groups in contact (C30 and H32) and hence the coupling between their titrations under changing pH. Ideally, one would have to treat the effect of the titration on the protonation states and resulting adjustments in the conformational flexibility of all titratable groups. However, pragmatically, only the protonation state of H32 was varied in the setup of the MD simulations to reflect pH differences. Two independent MD simulations were performed, with a protonated H32 and another with a neutral H32. These setups and MD trajectories are referred to as low pH (protonated H32, snapshot set **SP**) and high pH (neutral H32, set **SN**). In both pH regions, H32 adopts two groups of conformations, the populations of which are pH dependent (Figure 1). At low pH (**SP**), the conformers with the imidazole hydrogen bonded to the thiolate are significantly more populated than the conformers where the imidazole is more solvent exposed without hydrogen bonding the protein. At high pH (set **SN**), this second group of conformers (solvent exposed) is more populated, and visual inspection indicates that H32 is more flexible than at low pH (Figure 1). Hence, the MD simulations suggest that the charged form of H32 is less flexible and less accessible to the solvent than its neutral form. This illustrates that addition of a charge on a residue is not necessarily accompanied by a greater exposure of the charged moiety to the aqueous solvent, even when that would be sterically allowed. Thus, a direct contact between the protonated imidazole of H32 and the thiolate appears energetically more favorable than its extensive hydration. This echoes an earlier inference which suggested an electrostatic role for H32,[32] also consonant with some experimental observations on other systems.[86,87]

**Ionization Properties of C30 and H32 in Wild Type DsbA.** Figure 5 shows the impact of conformational sampling with the two H32 protonation states, **SP** and **SN**, on the calculated ionization properties of H32. As expected, the set **SN** results in a lower p*K* of H32, i.e., its neutral form is stabilized relative to the protonated form in **SP**. Conversely, the snapshot set **SP** stabilizes the protonated imidazole, consistent with more frequent contacts with the thiolate. Following the approach used in this study, we consider the average of the titration curves calculated over the two sets **SP** and **SN** as the titration curve of H32, which also determines the weight in eq 3 used to calculate the ionization of C30 (eq 2). The average calculated titration curves of H32 and C30 (Figure 5) do not obey the Henderson−Hasselbalch equation since they are characterized by two inflection points. The lower inflection points of both curves are at pH ∼3.40, whereas the upper ones are at pH 6.70


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

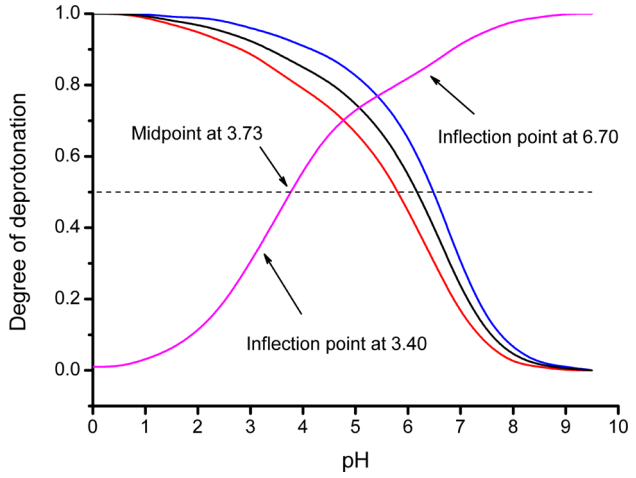



**Figure 5.** Calculated titration curves of C30 and H32 in wild type reduced *E. coli* DsbA. The curves for H32 are those obtained with MD snapshot sets **SP** (blue), **SN** (red), and the average over the two sets (black). For C30, only one titration curve (magenta) is shown, which is weighted according to eq 2. The state of half protonation is marked by the horizontal dotted line.

for C30 and at pH 6.47 for H32. In this way the two groups mirror their ionizations. Usually experimental observations of titration curves with two inflection points are interpreted as a superposition of two separate, noncoupled, titration curves; hence, two individual p*K* values are determined.[88] The alternative interpretation is that at least two closely situated titratable groups show the so-called mutual buffering effect.[38] That is, if the closely situated groups titrate within the same pH region, the protonation of one of them inhibits the protonation of the other(s) due to changing electrostatic influence and *vice versa*. As a result, the titration curves of such coupled groups have nonsigmoidal shapes. The conditions for this effect to occur and a quantitative description were presented earlier.[76] Such interpretation has so far relied on calculations on fixed three-dimensional protein structures, ignoring possible pH induced conformational responses. To our knowledge, the present study is the first one where cooperative ionization is theoretically described when pH induced conformational changes are taken into account. Thus, the theoretical ionization curves of C30 and H32 suggest that these titratable groups remain coupled despite the conformational relaxation that H32 undergoes. We refer to this effect as cooperative ionization. Since cooperative ionization results in a nonlinear dependence of p*K* on pH, it is appropriate (as stated above) to use p$K_{1/2}$, i.e., the pH at which the titratable group is half protonated, designated in Figure 5 as midpoint.

The calculated p$K_{1/2}$ values of H32 and C30 are 6.18 and 3.73, respectively. The latter is very close to the experimental value of 3.42[14] (see also Table 1). Both values are shifted down from their counterparts in model compounds. This is surprising for H32 since chemical intuition would suggest that the negative thiolate would shift the p$K_{1/2}$ of H32 to higher values. However, a detailed analysis of the calculations is required to understand the subtle interplay between C30 and H32 (see below). Interestingly, the calculated p$K_{1/2}$ of H32 echoes the experimental observation of a then "unknown group" titrating at p*K* ~6.7 which altered the reactivity of C30.[8] Such a marked influence on the reactivity of C30 suggests that this unknown group is in the vicinity of C30. The present analysis suggests that this group may well be H32. This possibility could be tested experimentally by studying again the reactivity of C30 after mutating H32 to a nontitratable residue other than proline.

Considering that H32 imidazole and C30 thiolate are close neighbors, one would expect that the charged form of these residues would stabilize each other. Therefore, at first glance, the extremely low p*K* value experimentally observed for C30 would reflect the electrostatic influence of the protonated H32 imidazole. According to the calculations, however, H32 is only partially protonated at the pH region where the deprotonation of C30 takes place. Hence, the stabilization of the C30 thiolate due to H32 reduces with increasing of pH. Thus, the very low p$K_{1/2}$ of C30 cannot be fully explained by a positive charge on H32. Instead, the low p$K_{1/2}$ of C30 originates mainly from the electrostatic influence of the partial (nontitratable) charges in its vicinity, including hydrogen-bonding dipoles. The section on Structure and Dynamics of the Reduced −C30−X31−X32−C33− Motifs above presented a qualitative interpretation of such effects in structural terms and descriptions of hydrogen bond counts between the thiolate and surrounding dipoles. The electrostatic calculations allow for dissecting and quantifying these contributions more precisely, with hydrogen-bonding groups being represented by permanent partial charges. As seen from Figure 6A, the p*K* shift resulting from the electrostatic interactions between C30 and the matrix of partial (non-

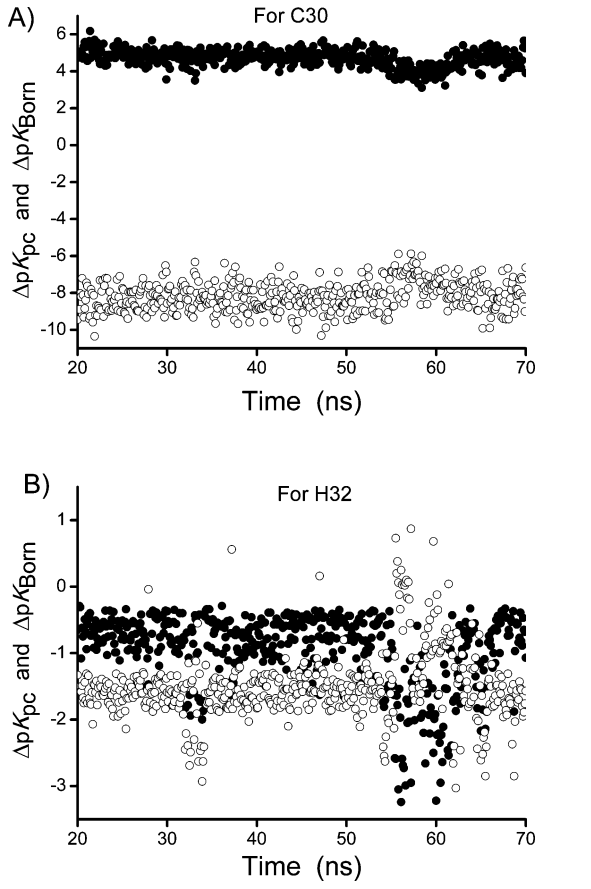



**Figure 6.** Calculated contributions to the shift of p*K* of specific residues due to their desolvation ($\Delta pK_{Born}$, filled circles) and interaction with the permanent (nontitratable) charges ($\Delta pK_{pc}$, open circles). These are shown for C30 (panel A) and H32 (panel B) in wild type DsbA (**SP** set), along the stretch of MD simulation used for p*K* calculations.


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

titratable) charges of the protein ($\Delta pK_{pc}$) is about −8.2, which is far from compensated by the desolvation contribution (on average $\Delta pK_{Born} \approx 4.7$). In this way, the intrinsic p$K$ (see for the definition of intrinsic p$K$ ref 89 or 76) of C30 is down-shifted to a value of 4.88.

H32 shares the same environment as C30; hence, the nontitratable partial charges are expected to shift the ionization equilibrium of H32 in a direction opposite to that of C30. Indeed, for H32 both $\Delta pK_{pc}$ and $\Delta pK_{Born}$ are negative (Figure 6B), thereby destabilizing the protonated form of H32. This results in a down-shift of the intrinsic p$K$ of H32 to a value of 4.42. This rationalizes the surprisingly low calculated $pK_{1/2}$ of H32. On the basis of these considerations, one can conclude that the dramatic reduction of $pK_{1/2}$ of C30 and the relatively moderate reduction of $pK_{1/2}$ of H32 are caused by their shared polar environment. The larger shift obtained for C30 is due to the burial of its side chain within an array of stabilizing hydrogen-bonding dipoles. This illustrates the interest of detailed p$K$ calculations which allow one to dissect the various physical contributions to a p$K$ shift.

**p$K$ of C30 Calculated Based on an NMR Structure of Wild Type DsbA.** Three-dimensional protein structures obtained by NMR spectroscopy provide information about the conformations that proteins may adopt in solution. In this context, a set of NMR conformers might be considered as an alternative to the MD sampling of conformations. Indeed, a previous study of three proteins found that p$K$ values averaged over NMR conformers were more accurate than those calculated from a single crystal structure.[90] Another study found that, in the regions where NMR and X-ray structures differed significantly, the p$K$ values calculated with the X-ray structures were in better agreement with the experiment.[91] Indeed, there is much debate about the accuracy of the details of protein NMR structures,[92] and it has been suggested that p$K$ calculations can help test the chemical relevance of NMR conformers.[48,52,93]

Thus, it is natural to investigate if calculations relying on an NMR structure of reduced wild type *E. coli* DsbA[31] (PDB entry 1A24, Figure S4 (Supporting Information)) can reproduce the low p$K$ value of C30. We have performed p$K$ calculations on each of the 20 NMR conformers, using the same protocol as that with the MD snapshots. The calculated titration curves are shown in Figure 7. The associated calculated $pK_{1/2}$ values for C30 cover a region from pH 5.7 and upward, with an average $pK_{1/2}$ of 7.3, well above the experimental p$K$ value of ∼3.5 (Table 1). This average $pK_{1/2}$ is only slightly lower than the standard p$K$ value of a free cysteine, reflecting weak stabilizing interactions between C30 and the protein. Indeed, none of the NMR conformers contains a hydrogen bond between C30 and H32, the two functional groups being separated by 5 to 10 Å. In addition, the C30 side chain points away from the neighboring helical backbone amide N−H groups, so those groups are too far to hydrogen bond $S^{\gamma}_{30}$ (Figure S4, Supporting Information). Consequently, the electrostatic stabilization of the thiolate is severely diminished. Thus, the NMR conformers are inconsistent with the low p$K$ of C30. In addition, C30 was modeled with a neutral thiol in the NMR structure, alongside a charged H32. This combination of protonation states corresponds to a pH region below the experimental p$K$ of C30, where the protein unfolds.[94] In sum, the present analysis indicates that the view given in Figure 1 provides a more realistic picture of the reduced active site of *E. coli* DsbA in solution since it is consistent with the low p$K$ of C30.

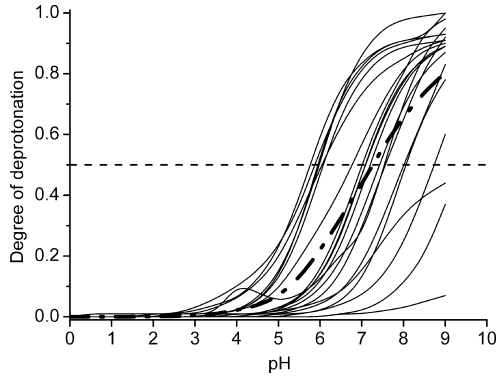



**Figure 7.** Titration curves of C30 calculated with an NMR structure of *E. coli* DsbA (PDB entry 1A24). There are 20 individual titration curves (thin solid black lines), corresponding to the 20 conformers of the NMR structure. The bold dashed-dot curve is the average over the 20 individual titration curves, and its intersection with the horizontal dashed line corresponds to the average calculated $pK_{1/2}$ = 7.3 for C30.

**Ionization of Cys30 in −C30−X31−X32−C33− Mutants of DsbA.** In most mutants (except −C30−G31−P32−C33− and −C30−P31−P32−C33−), the measured p$K$ of C30 is only slightly higher than that in the wild type, by 0.8 to 1.4 pH units (Table 1). This range of 0.6 pH units makes this set of measured $pK_{exp}$ values particularly challenging for p$K$ calculations.[21] Mutants −C30−G31−P32−C33− and −C30−P31−P32−C33− stand out with their comparatively higher measured C30 $pK_{exp}$ values, 6.21 and 6.73, respectively.

The p$K$ values of C30 in the mutants were also calculated and averaged based on MD snapshots and Poisson−Boltzmann theory (Table 1). The calculated p$K$s are systematically larger than their measured counterpart, by about one pH unit (Figure 4). Nevertheless, there is consistency between calculated and experimental p$K$s for some broad trends. First, all of the calculated p$K$s for C30 in the mutants are larger than that for C30 in the wild type. Second, the calculated p$K$s of C30 in the −C30−G31−P32−C33− and −C30−P31−P32−C33− mutants are distinctively higher than those in the other mutants, consistent with measurements. One notes that the calculated p$K$ differences for C30 between (i) the wild type and −C30−G31−P32−C33−, (ii) the wild type and −C30−P31−P32−C33−, and (iii) the wild type and the mutants with a p$K$ in the middle of the range (∼4.5) are much larger than the standard error of the mean (typically ∼0.05 p$K$ units, Table 1) for these calculated p$K$s. Thus, the largest p$K$ differences addressed in the following are not obscured by the magnitude of the fluctuations in the calculated p$K$s. A structural interpretation for the two higher p$K$s has already been suggested above. It proposes that the presence of a proline at position 32 removes the backbone $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond, thereby removing an interaction which stabilizes the thiolate in other sequences.

For the other mutants, with measured p$K$s from 3.71 to 4.86 (Table 1), there is no discernible agreement in the ranking of the calculated p$K$s compared to their measured counterparts (Table 1 and Figure 4). This is not surprising, considering the small differences between these p$K$s. Also, the corresponding measurements may not be free of experimental uncertainties. Still, the calculated p$K$s and their energetic components can lend themselves to interesting insights, for instance with the −C30−T31−R32−C33− mutant. It is remarkable that the p$K$ of C30 in −C30−T31−R32−C33− is up-shifted by more than 1 pH unit relative to that of the wild type, despite H32 being


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

substituted by another positively charged group. As already noted above, in the simulated structure of −C30−T31−R32−C33− the side chain of R32 points away from C30 and toward the solvent most of the time (Figure 2B), hydrating its guanidinium group. Since the influence of peripheral charged side chains on the thiolate has been much debated, it is of particular interest to compare the Coulombic interaction energies between the thiolate and the charged imidazole (wild type) and between the thiolate and the charged guanidinium (−C30−T31−R32−C33−). Figure 8A shows

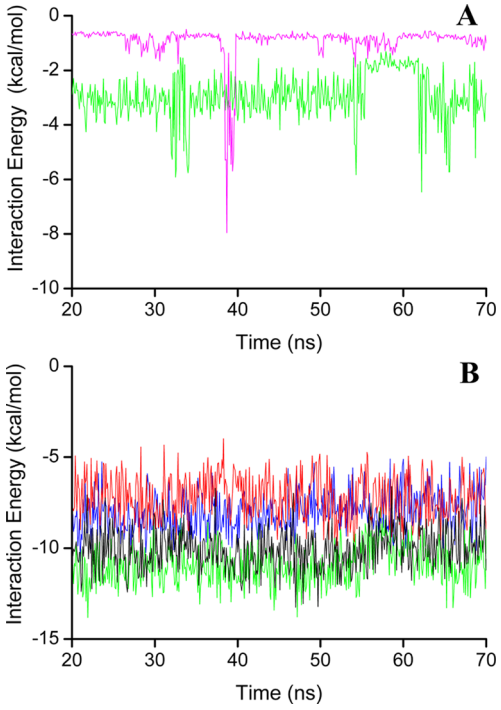



**Figure 8.** Electrostatic interaction energies (kcal/mol) calculated between the thiolate of C30 and selected groups. The time course of these energies is plotted for the MD snapshots used for the p*K* calculations (every 100 ps from 20 to 70 ns). Panel A shows the Coulombic interaction energy between the thiolate and the imidazole of H32 in WT DsbA (green) and between the thiolate and the guanidinium group of R32 in the −C30−T31−R32−C33− mutant (magenta). These energies were obtained with a positive net charge on both the imidazole and the guanidinium. Panel B shows the Coulombic interaction energy between the thiolate and the protein "permanent charges" used during the p*K* calculations, for WT DsbA (green), the −C30−P31−G32−C33− mutant (black), the −C30−G31−P32−C33− mutant (blue), and the −C30−P31−P32−C33− mutant (red). The "permanent charges" are the partial charges representing the covalent bond dipoles, including the backbone amide N−H dipoles (but excluding the +1 or −1 charges resulting from ionization upon changes in protonation state). The corresponding average values are given in the main text.

the time course of these interaction energies for the subset of snapshots used in the p*K* calculations. The thiolate−guanidinium interaction energy is weak most of the time (time average = −0.9 kcal/mol) and clearly less favorable than the thiolate−imidazole interaction energy (time average = −2.9 kcal/mol). The stabilization of the thiolate by R32 becomes much stronger only transiently at ∼40 ns (Figure 8A), when R32 hydrogen bonded the thiolate with an interaction energy up to −8 kcal/mol. However, such a hydrogen bond has only a weak statistical weight since it is a rare event. Thus, comparing the influences of the imidazole and guanidinium emphasizes again the importance of direct polar contacts to $S^{\gamma}_{30}$ for stabilization of the thiolate. Since the guanidinium of R32 hydrogen bonds the thiolate only rarely, the −C30−T31−R32−C33− mutant behaves somewhat like mutants with neutral side chains regarding the p*K* of C30.

Our theoretical interpretation developed with wild type *E. coli* DsbA, i.e., that the local protein polar environment dominates the p*K* down-shift of C30, is overall consistent with our analysis of the p*K* values of C30 for the mutants. This is apparent in the electrostatic interaction energies between the thiolate sulfur and the protein permanent charges obtained as part of the p*K* calculations. These charges are those which remain unchanged during the protonation/deprotonation of the titratable groups and comprise the partial charges characterizing the polar covalent groups, including the backbone amide N−H dipoles. Thus, the permanent charges largely represent the hydrogen-bonding interactions. The time course of the electrostatic interaction energies between the thiolate and the permanent charges is plotted in Figure 8B for the wild type and mutants −C30−P31−G32−C33−, −C30−G31−P32−C33−, and −C30−P31−P32−C33−. Mutant −C30−P31−G32−C33− is an interesting reference since it has no side chain at position X2 of −C−X1−X2−C− and therefore puts in perspective the stabilizing role of the side chain of H32 in the wild type. −C30−P31−G32−C33− is also a convenient control to interpret the thioredoxin-like −C30−G31−P32−C33− since they have the same composition of chemical groups but arranged differently around the thiolate. The time-averaged interaction energies between the thiolate and the permanent charges are −10.9, −10.1, −8.0, and −7.2 kcal/mol for −C30−P31−H32−C33−, −C30−P31−G32−C33−, −C30−G31−P32−C33−, and −C30−P31−P32−C33−, respectively. These interaction energies rank like the associated C30 $pK_{exp}$ values, highlighting the strong influence of the permanent charges (i.e., hydrogen bonds) on the thiolate stabilization.

Comparing −C30−P31−H32−C33− to −C30−P31−G32−C33− suggests that the H32 side chain stabilizes the thiolate by approximately −10.9 − (−10.1) = −0.8 kcal/mol via neutral polar contacts; H32 further stabilizes the thiolate via net charge−charge interactions (Figure 8A). Thus, one can essentially explain the higher p*K* of C30 in the mutants, relative to wild type, as resulting from the lack of the direct interaction between the H32 side chain and the thiolate. This is consistent with the wild type and PDI-like mutant having both H32 and the lowest measured p*K* of C30 among the mutants. The mutants, however, still have low p*K* values for C30 since most of them maintain the backbone $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond to the thiolate. The role of this interaction is reflected in the less favorable interaction energy between the thiolate and permanent charges in −C30−G31−P32−C33− (no $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond) relative to −C30−P31−G32−C33− ($N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond present), by approximately −8.0 − (−10.1) = 2.1 kcal/mol. This interaction energy is even less favorable (0.8 kcal/mol higher) in −C30−P31−P32−C33− than with −C30−P31−P32−C33−. This is consistent with the more frequent disruption of the $N_{33}$−H···$S_{30}{}^{\gamma}$ and $S^{\gamma}_{33}$−H···$S^{\gamma}_{30}$ hydrogen bonds in −C30−P31−P32−C33−, related to the greater flexibility of C30 in −C30−P31−P32−C33− than in −C30−G31−P32−C33− (see above and Figure 3). Altogether, the structural and energetic analyses point to a dominant role of direct hydrogen bonds to the sulfur for


dx.doi.org/10.1021/bi400500e | *Biochemistry* 2013, 52, 5730−5745

stabilization of its thiolate. This is consistent with an interpretation of the structure and biochemical properties of DsbA from N. *meningitidis*, where a threonine outside the −C−X1−X2−C− motif donates yet another hydrogen bond to the thiolate, resulting in a lower cysteine p*K* and a more oxidizing enzyme.[12]

In sum, even though the calculations did not completely reproduce the differences between measured p*K* values across the mutants, the results still give interpretable insights into the relative energetic contributions of specific interactions. The analysis of the mutants stresses that the p*K* of C30 is much more influenced by the residue at position X2 than by the residue at position X1.

## CONCLUSIONS

DsbA is an important enzyme in its own right but has also been a major model system to investigate the active sites of oxidoreductases across the Trx superfamily to understand how their −C−X1−X2−C− motifs influence the p*K* of the key catalytic cysteine. Thus, detailed experimental characterizations of wild type *E. coli* DsbA and its mutants have long been available. To complement and help interpret these experimental data, the present work embarked on an extensive computational analysis of these systems. This computational approach had several aims, including gaining additional structural information by MD simulations of wild type reduced DsbA and its mutants. For the mutants, the structural insights obtained here are new and comprehensive, and bridge a gap to provide a concrete context to study the electrostatics of their active sites. A detailed description of the salient structural and dynamical features of these sites gives specific insights into the interactions which stabilize the thiolate of C30 in the DsbA systems. This is accompanied by detailed p*K* calculations to provide additional physics based insights into the electrostatic factors at work. The p*K* calculations emphasized the importance of direct polar interactions with the thiolate, i.e., hydrogen bonds, as the dominant factor decreasing the p*K* of C30.

Specific new findings regarding wild type DsbA include the interaction between H32 and the thiolate via a C−H hydrogen bond suggested by the simulations, with both protonated and neutral H32. This might be probed by future NMR experiments. Generating structural models with both the neutral and protonated states of H32, to cover different pH ranges relevant to pH titrations, provided a firm foundation for the difficult treatment of the coupled titrations between C30 and H32. The calculated p*K* of C30 in wild type *E. coli* DsbA was found in agreement with its experimental counterpart. According to our calculations, this p*K* is primarily lowered by hydrogen bonds between (i) the thiolate and the backbone N−H group of H32, (ii) the thiolate and the backbone N−H group of C33, and (iii) the thiolate and the imidazole of H32. The thiol group of C33 also hydrogen bonds the thiolate in the MD simulations. For H32, our study yielded a calculated $pK_{1/2}$ = 6.18, suggesting that its imidazole is only partially protonated at pH 7. Although this appears counterintuitive, it can be rationalized by dissecting the factors which influence the calculated $pK_{1/2}$. Importantly, it echoes an early study of DsbA which uncovered the titration of a then unknown group at pH 6.7.[8] The present work suggests that H32 may be this group, which could be tested by performing the same experiments[8] with a DsbA mutant where H32 would be mutated. In addition, calculations of the coupling between the titrations of C30 and H32 anticipate departures from conventional sigmoidal titrations for C30 and H32. Experimental characterization of this effect would add to our understanding of DsbA and of coupled titrations in proteins in general.

The MD structural models of the DsbA mutants clarified important relationships between the structure and dynamics of their −C30−X31−X32−C33− active sites and the factors controlling the p*K* of C30. First, no hydrogen bond between residue X31 and the thiolate was uncovered during the simulations. The backbone N−H group of X31 is too far from the thiolate and does not point toward the thiolate. The side chain of X31 is projected away from the thiolate, into the solvent. Therefore, residue X31 does not significantly stabilize the thiolate. Instead, residue X32 can markedly stabilize the thiolate directly, via a backbone $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond and interaction between $S^{\gamma}_{30}$ and the side chain of X32. However, the only side chain at X32 found so far to markedly stabilize the thiolate is a histidine, for simple steric reasons. Other investigated polar side chains at X32 did not have the appropriate length or conformation to form long-lasting hydrogen bonds to the thiolate. Instead, they hydrogen bonded the thiolate only rarely and thus contributed little to lower the p*K* of C30. Threonine at X32 is too short to reach to the thiolate. Arginine at X32 is long and flexible enough to reach the thiolate, but the simulations show that R32 prefers to adopt another orientation, where its guanidinium points into the aqueous solvent. Thus, in the mutants X32 mainly lowers the p*K* of C30 via its backbone $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond.

When X32 is a proline, the $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond is removed, and the p*K* of C30 is distinctively higher, as evidenced by DsbA mutants −C30−G31−P32−C33− and −C30−P31−P32−C33−. This most likely explains the relatively high p*K*s of the catalytic cysteine in thioredoxins (−C−G−P−C− motif) and tryparedoxin (−C−P−P−C- motif). Removal of the $N_{32}$−H···$S^{\gamma}_{30}$ hydrogen bond destabilizes the $\chi_1^{30}$ = *trans* conformation of C30, in favor of the $\chi_1^{30} = g^-$ conformation. This increase in the $\chi_1^{30} = g^-$ population disrupts two other stabilizing hydrogen bonds between the thiolate and C33 (backbone $N_{33}$−H···$S^{\gamma}_{30}$ and thiol $S^{\gamma}_{33}$−H···$S^{\gamma}_{30}$), further increasing the p*K* of C30. Therefore, X32 is clearly more relevant than X31 in controlling the p*K* of the catalytic cysteine. The second residue critical to the stabilization of the p*K* is C33, also via hydrogen bonds. C33 was not varied in the present work, but its mutation to alanine in homologous enzymes is known to increase the p*K* of the catalytic cysteine. In sum, the present work proposes that the p*K* of C30 in wild type *E. coli* DsbA and mutants is primarily lowered by residues X32 and C33. The influence of X31 is much weaker but can be indirect and subtle, as illustrated when comparing −C30−P31−H32−C33− to the PDI-like sequence −C30−G31−H32−C33−. The P31 to G31 substitution influences the stiffness and dynamics of the site, resulting in slightly less frequent hydrogen bonds to the thiolate in the PDI-like sequence. The notions developed here on the respective roles of X31, X32, and C33 regarding thiolate stabilization should be largely transferable to other enzymes featuring −C−X1−X2−C− motifs.

The influence of protein structural details and their dynamics on the p*K* is only one of the factors hampering the very accurate calculation of p*K*s. Other factors touched upon in the present work include the need, in principle, to obtain structural ensembles at different pH regimes. Therefore, it is not surprising that the calculated p*K*s departed somewhat from their measured counterparts. Thus, accurate predictions of

small p*K* differences for these systems is still elusive. Yet, the Poisson−Boltzmann calculations provided valuable insights to dissect the various physical contributions to the p*K* shifts.

Although small p*K* differences are difficult to calculate accurately, it is encouraging that the p*K* calculations obtained the correct ranking between three main categories, i.e., (i) wild type DsbA (lowest C30 p*K*), (ii) mutants −C30−G31−P32−C33− and −C30−P31−P32−C33− (distinctively highest C30 p*K*s), and (iii) other −C30−X31−X32−C33− mutants (intermediate C30 p*K*s). Being able to discriminate between a p*K* of ∼3.5 and a p*K* of ∼6.0 is relevant to the thioredoxin superfamily since it includes enzymes with catalytic cysteines at the two extremes of this range and beyond.[26] Therefore, the p*K* calculation protocols employed here should prove helpful for future investigations of thiol-disulfide oxidoreductases.

## ■ ASSOCIATED CONTENT

### Supporting Information

List of structural descriptors for the reduced cysteines before they were subjected to MD simulations; a comparison of two isomorphous orientations of the imidazole of H32 which can both hydrogen bond C30 in the active site of wild-type DsbA; pictures of the structures of the active-sites (MD snapshots) for mutants −C30−G31−H32−C33−, −C30−P31−Y32−C33−, −C30−S31−V32−C33−, −C30−A31−T32−C33−, −C30−S31−F32−C33−, −C30−P31−L32−C33−, −C30−S31−T32−C33−, −C30−P31−G32−C33−, and −C30−L31−T32−C33−; a picture of the time series of the simulated $N^{\delta}_{32}\cdots S^{\gamma}_{30}$ and $C^{\delta}_{32}\cdots S^{\gamma}_{30}$ distances in wild-type DsbA; and a picture of the structure of the active site of the NMR structure (PDB entry 1A24) of *E. coli* DsbA. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION


### Corresponding Author

*(L.N.) Phone: +46-0-8-524-81099. E-mail: Lennart.Nilsson@ki.se. (N.F.) Phone: +44-0-1223-895-338. E-mail: nf_research@hotmail.com.



### Funding

A.K. thanks the Bulgarian National Research Fund (grant DRG-02/05). This work was also supported by the Swedish Research Council.

### Notes

The authors declare no competing financial interest.


## ■ ABBREVIATIONS USED

3D, three-dimensional; DsbA, disulfide bond protein A; *E. coli*, *Escherichia coli*; MD, molecular dynamics; NMR, nuclear magnetic resonance; PB, Poisson−Boltzmann; PDB, protein data bank; $pK_{1/2}$, calculated pH at which a titratable group is half protonated; $pK_{exp}$, experimentally measured p*K*; Trx, thioredoxin

## ■ REFERENCES


(1) Sevier, C. S., and Kaiser, C. A. (2006) Conservation and diversity of the cellular disulfide bond formation pathways. *Antioxid. Redox Signaling 8*, 797−811.

(2) Depuydt, M., Messens, J., and Collet, J.-F. (2011) How proteins form disulfide bonds. *Antioxid. Redox Signaling 15*, 49−66.

(3) Carvalho, A. P., Fernandes, P. A., and Ramos, M. J. (2006) Similarities and differences in the thioredoxin superfamily. *Prog. Biophys. Mol. Biol. 91*, 229−248.

(4) Martin, J. L. (1995) Thioredoxin - a fold for all reasons. *Structure 3*, 245−250.

(5) Reckenfelderbaumer, N., and Krauth-Siegel, L. R. L. (2002) Catalytic properties, thiol pK value, and redox potential of *Trypanosoma brucei* tryparedoxin. *J. Biol. Chem. 277*, 17548−17555.

(6) Holmgren, A. (1989) Thioredoxin and glutaredoxin. *J. Biol. Chem. 264*, 13963−13966.

(7) Kortemme, T., Darby, N. J., and Creighton, T. H. (1996) Electrostatic interactions in the active-site of the N-terminal thioredoxin-like domain of protein disulfide isomerase. *Biochemistry 35*, 14503−14511.

(8) Nelson, J. W., and Creighton, T. E. (1994) Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation *in vivo*. *Biochemistry 33*, 5974−5983.

(9) Moutevelis, E., and Warwicker, J. (2004) Prediction of pKa and redox properties in the thioredoxin superfamily. *Protein Sci. 13*, 2744−2752.

(10) Mössner, E., Huber-Wunderlich, M., and Glockshuber, R. (1998) Characterization of *Escherichia coli* thioredoxin variants mimicking the active-sites of other thiol/disulfide oxidoreductases. *Protein Sci. 7*, 1233−1244.

(11) Huber-Wunderlich, M., and Glockshuber, R. (1998) A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold. Des. 3*, 161−171.

(12) Lafaye, C., Iwema, T., Carpentier, P., Jullian-Binard, C., Kroll, J. S., Collet, J.-F., and Serre, L. (2009) Biochemical and structural study of the homologues of the thiol-disulfide oxidoreductase DsbA in *Neisseria meningitidis*. *J. Mol. Biol. 392*, 952−966.

(13) Kallis, G.-B., and Holmgren, A. (1980) Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-35 present in the reduced form of thioredoxin from *Escherichia coli*. *J. Biol. Chem. 255*, 10261−10265.

(14) Grauschopf, U., Winther, J. R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J. C. A. (1995) Why is DsbA such an oxidizing disulfide catalyst? *Cell 83*, 947−955.

(15) Szajewski, R. P., and Whitesides, G. M. (1980) Rate constants and equilibrium constants for thiol-disulfide interchange reactions involving oxidized glutathione. *J. Am. Chem. Soc. 102*, 2011−2026.

(16) Shakked, Z., Szajewski, R. P., and Whitesides, G. M. (1980) Rates of thiol-disulfide interchange reactions involving proteins and kinetic measurements of thiol $pK_a$ values. *Biochemistry 19*, 4156−4166.

(17) Winterbourn, C. C., and Metodiewa, D. (1999) Reactivity of biologically important thiol compounds with superoxide and hydrogen peroxide. *Free Radical Biol. Med. 27*, 322−328.

(18) Gane, P. J., Freedman, R. B., and Warwicker, J. (1995) A molecular model for the redox potential difference between thioredoxin and DsbA based on electrostatic calculations. *J. Mol. Biol. 249*, 376−387.

(19) Chivers, P. T., Prehoda, K. E., and Raines, R. T. (1997) The CXXC motif: a rheostat in the active site. *Biochemistry 36*, 4061−4066.

(20) Guddat, L. W., Bardwell, J. C. A., and Martin, J. (1998) Crystal structures of the reduced and oxidized DsbA: investigation of domain motion and thiolate stabilization. *Structure 6*, 757−767.

(21) Jao, S. C., Ospina, S. M. E., Berdis, A. J., Starke, D. W., Post, C. B., and Mieyal, J. J. (2006) Computational and mutational analysis of human glutaredoxin (thioltransferase): probing the molecular basis of the low $pK_a$ of cysteine 22 and its role in catalysis. *Biochemistry 45*, 4785−4796.

(22) Foloppe, N., Sagemark, J., Nordstrand, K., Berndt, K. D., and Nilsson, L. (2001) Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from *Escherichia coli*: comparison with functionally related proteins. *J. Mol. Biol. 310*, 449−470.

(23) Karala, A.-R., Lappi, A.-K., and Ruddock, L. W. (2010) A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. *J. Mol. Biol. 396*, 883−892.

(24) Wada, A. (1976) The α-helix as an electric macro-dipole. *Adv. Biophys. 9*, 1−63.

(25) Foloppe, N., and Nilsson, L. (2007) Stabilization of the catalytic thiolate in a mammalian glutaredoxin: structure, dynamics and electrostatics of reduced pig glutaredoxin and its mutants. *J. Mol. Biol. 372*, 798−816.
(26) Roos, G., Foloppe, N., and Messens, J. (2013) Understanding the $pK_a$ of redox cysteines: the key role of hydrogen bonding. *Antioxid. Redox Signaling 18*, 94−127.
(27) Roos, G., Loverix, S., and Geerlings, P. (2006) Origin of the $pK_a$ perturbation of Nterminal cysteine in alpha- and 3(10)-helices: a computational DFT study. *J. Phys. Chem. B 110*, 557−562.
(28) Bardwell, J. C. A., and Beckwith, J. (1993) The bonds that tie: catalyzed disulfide bond formation. *Cell 74*, 769−771.
(29) Wunderlich, M., and Glockshuber, R. (1993) Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci. 2*, 717−726.
(30) Hennecke, J., Spleiss, C., and Glockshuber, R. (1997) Influence of acidic residues and the kink in the active-site helix on the properties of the disulfide oxidoreductase DsbA. *J. Biol. Chem. 272*, 189−195.
(31) Schirra, H. J., Renner, C., Czisch, M., Huber-Wunderlich, M., Holak, T. A., and Glockshuber, R. (1998) Structure of reduced DsbA from *Escherichia coli* in solution. *Biochemistry 37*, 6263−6276.
(32) Guddat, L. W., Bardwell, J. C. A., Glockshuber, R., Huber-Wunderlich, M., Zander, T., and Martin, J. L. (1997) Stuctural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci. 6*, 1893−1900.
(33) Warwicker, J., and Gane, P. J. (1996) Calculation of Cys 30 ΔpK*a*'s and oxidising power for DsbA mutants. *FEBS Lett. 385*, 105−108.
(34) Warwicker, J. (1998) Modeling charge interactions and redox properties in DsbA. *J. Biol. Chem. 273*, 2501−2504.
(35) D'Ambrosio, K., Pedone, E., Langella, E., De Simone, G., Rossi, M., Pedone, C., and Bartolucci, S. (2006) A novel member of the protein disulfide oxidoreductase family from *Aeropyrum pernix* K1: structure, function and electrostatics. *J. Mol. Biol. 362*, 743−752.
(36) Bashford, D., and Karplus, M. (1991) Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation. *J. Phys. Chem. 95*, 9556−9561.
(37) Nielsen, J. E., and McCammon, J. A. (2003) Calculating pKa values in enzyme active site. *Protein Sci. 12*, 1894−1901.
(38) Yang, A.-S., Gunner, M. R., Sampogna, R., Sharp, K., and Honig, B. (1993) On the calculation of $pK_a$s in proteins. *Proteins 15*, 252−265.
(39) Rocchia, W., Alexov, E., and Honig, B. (2001) Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. *J. Phys. Chem. 105*, 6507−6514.
(40) Huang, R. B., Du, Q. S., Wang, C. H., Liao, S. M., and Chou, K. C. (2010) A fast and accurate method for predicting pK(a) of residues in proteins. *Protein Eng., Des. Sel. 23*, 35−42.
(41) Khandogin, J., and Brooks, C. L. (2006) Toward the accurate first-principles prediction of ionization equilibria in proteins. *Biochemistry 45*, 9363−9373.
(42) Mitra, R., Shyam, R., Mitra, I., Miteva, M. A., and Alexov, E. (2008) Calculating the protonation states of proteins and small molecules: Implications to ligand-receptor interactions. *Curr. Comput.-Aided Drug Des. 4*, 169−179.
(43) Aleksandrov, A., Polydorides, S., Archontis, G., and Simonson, T. (2010) Predicting the acid/base behavior of proteins: A constant-ph monte carlo approach with generalized born solvent. *J. Phys. Chem. B 114*, 10628−10633.
(44) Dillet, V., Dyson, H. J., and Bashford, D. (1998) Calculations of electrostatic interactions and $pk_a$s in the active site of *Escherichia coli* thioredoxin. *Biochemistry 37*, 10298−10306.
(45) Kieseritzky, G., and Knapp, E. W. (2008) Optimizing $pK_a$ computation in proteins with pH adapted conformations. *Proteins 71*, 1335−1348.
(46) Karshikoff, A., and Ladenstein, R. (2007) The Role of Electrostatic Interactions in the Stabilization of Proteins from Thermophiles, in *Protein Structures: Methods in Protein Structure and Stability Analysis* (Uversky, V. N., and Permyakov, E. A., Eds.) pp 71−109, Nova Science Publishers, New York.
(47) Bjelic, S., Wieninger, S., Jelesarov, I., and Karshikoff, A. (2008) Electrostatic contribution to the thermodynamic and kinetic stability of the homotrimeric coiled coil Lpp-56: a computational study. *Proteins 70*, 810−822.
(48) Foloppe, N., Vlamis-Gardikas, A., and Nilsson, L. (2012) The -Cys-X1-X2-Cys- motif of reduced glutaredoxins adopts a consensus structure that explains the low p*K*a of its catalytic cysteine. *Biochemistry 51*, 8189−8207.
(49) Koumanov, A., Karshikoff, A., Friis, E. P., and Borchert, T. V. (2001) Conformational averaging in pK calculations. Improvement and limitations in prediction of ionization properties of proteins. *J. Phys. Chem. B 105*, 9339−9344.
(50) Sandberg, L., and Edholm, O. (1997) $pK_a$ calculations along bacteriothodopsin molecular dynamics trajectory. *Biophys. Chem. 65*, 189−204.
(51) van Vlijmen, H. W. T., Schaefer, M., and Karplus, M. (1998) Improving the accuracy of protein $pK_a$ calculations - conformational averaging versus the average structure. *Proteins 33*, 145−158.
(52) Foloppe, N., and Nilsson, L. (2004) The Glutaredoxin -C-P-Y-C- motif: influence of peripheral residues. *Structure 12*, 289−300.
(53) Alexov, E. (2003) Role of the protein side-chain fluctuations on the strength of pair-wise electrostatic interactions: comparing experimental with computed $pK_a$s. *Proteins 50*, 94−103.
(54) Nielsen, J. E., and McCammon, J. A. (2003) On the evaluation and optimization of protein X-ray structures for pKa calculations. *Protein Sci. 12*, 313−326.
(55) Tarus, B., Straub, J. E., and Thirumalai, D. (2006) Dynamics of Asp23-Lys28 salt-bridge formation in $A\beta_{10-35}$ monomers. *J. Am. Chem. Soc. 128*, 16159−16168.
(56) Hajjar, E., Dejaegere, A., and Reuter, N. (2009) Challenges in pK(a) predictions for proteins: the case of Asp213 in human proteinase 3. *J. Phys. Chem. 113*, 11783−11792.
(57) Lee, M. S., Salsbury, F. R., and Brooks, C. L. (2004) Constant-pH molecular dynamics using continuous titration coordinates. *Proteins 56*, 738−752.
(58) Baptista, A. M., Martel, P. J., and Petersen, S. B. (1997) Simulation of protein conformational freedom as a function of pH: constant-pH molecular dynamics using implicit titration. *Proteins 27*, 523−544.
(59) Williams, S. L., de Oliveira, C. A. F., and McCammon, J. A. (2010) Coupling constant pH molecular dynamics with accelerated molecular dynamics. *J. Chem. Theory Comput. 6*, 560−5568.
(60) Dashti, D. S., Meng, Y., and Roitberg, A. E. (2012) pH-Replica exchange molecular dynamics in proteins using a discrete protonation method. *J. Phys. Chem. B 116*, 8805−8811.
(61) Nilsson, L., and Karshikoff, A. (2011) Multiple pH regime molecular dynamics simulation for p*K* calculations. *PLoS One 6*, e20116.
(62) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraint. *J. Mol. Biol. 234*, 779−815.
(63) Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem. 30*, 1545−1614.
(64) 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.

(65) MacKerell, A. D. J., Feig, M., and Brooks, I. C. L. (2004) Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc. 126*, 698−699.
(66) Steinbach, P. J., and Brooks, B. R. (1994) New spherical-cutoff methods for long-ranges forces in macromolecular simulation. *J. Comput. Chem. 15*, 667−683.
(67) Norberg, J., and Nilsson, L. (2000) On the truncation of long-range electrostatic interactions in DNA. *Biophys. J. 79*, 1537−1553.
(68) Beck, D. A. C., Armen, R. S., and Daggett, V. (2005) Cutoff size need not strongly influence molecular dynamics results for solvated polypeptides. *Biochemistry 44*, 609−616.
(69) Cheatham, T. E., and Kollman, P. A. (2000) Molecular dynamics simulation of nucleic acids. *Annu. Rev. Phys. Chem. 51*, 435−471.
(70) Heddi, B., Foloppe, N., Oguey, C., and Hartmann, B. (2008) Importance of accurate DNA structures in solution: the Jun-Fos model. *J. Mol. Biol. 382*, 956−970.
(71) Roos, G., Foloppe, N., Van Laer, K., Wyns, L., Nilsson, L., Geerlings, P., and Messens, J. (2009) How thioredoxin dissociates its mixed disulfide. *PLoS Comp. Biol. 5*, e1000461.
(72) Van Laer, K., Buts, L., Foloppe, N., Vertommen, D., Van Belle, K., Wahni, K., Roos, G., Nilsson, L., Mateos, L.M.., Rawat, M., van Nuland, N., and Messens, J. (2012) Mycoredoxin-1 is one of the missing links in the oxidative stress defence mechanism of Mycobacteria. *Mol. Microbiol. 86*, 787−804.
(73) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(74) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-integration of cartesian equations of motion of a system with constraints - molecular dynamics of N-alkanes. *J. Comput. Phys. 23*, 327−341.
(75) Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wuthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol. 280*, 933−952.
(76) Koumanov, A., Rüterjans, H., and Karshikoff, A. (2002) Continuum electrostatic analysis of irregular ionization and proton allocation in proteins. *Proteins 46*, 85−96.
(77) Koumanov, A., Benach, J., Atrian, S., Gonzàlez-Duarte, R., Karshikoff, A., and Ladenstein, R. (2003) The catalytic mechanism of Drosophila alcohol dehydrogenase: evidence for a proton relay modulated by the coupled ionization of the active site lysine/tyrosine pair and a NAD+ ribose OH switch. *Proteins 51*, 289−298.
(78) Rashin, A. A., Iofin, M., and Honig, B. (1986) Internal cavities and buried waters in globular proteins. *Biochemistry 25*, 3619−3625.
(79) Nozaki, Y., and Tanford, C. (1967) Examination of titration behaviour. *Methods Enzymol. 11*, 715−734.
(80) Bucholtz, K. M., Gareiss, P. C., Tajc, S. G., and Miller, B. L. (2004) Direct determination of thiol p$K_a$ by isothermal titration microcalorimetry. *J. Am. Chem. Soc. 126*, 10508−10509.
(81) Rundlöf, A.-K., and Arnér, E. S. J. (2004) Regulation of the mammalian selenoprotein thioredoxin reductase 1 in relation to cellular phenotype, growth, and signaling events. *Antioxid. Redox Signaling 6*, 41−52.
(82) Taylor, R., and Kennard, O. (1982) Crystallographic evidence for the existence of C-H.O, C-H.N, and C-H.Cl hydrogen bonds. *J. Am. Chem. Soc. 104*, 5063−5070.
(83) Desiraju, G. R., and Steiner, T. (1999) *IUCr Monographs on Crystallography: 9* Oxford Science Publications, Oxford, U.K..
(84) Hunter, C. A., and Sanders, J. K. M. (1990) The Nature of $\pi-\pi$ interactions. *J. Am. Chem. Soc. 112*, 5525−5534.
(85) Waters, M. L. (2002) Aromatic interactions in model systems. *Curr. Opin. Chem. Biol. 6*, 736−741.
(86) Isom, D. G., Castaneda, C. A., Velu, P. D., and Garcia-Moreno, B. (2010) Charges in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci. U.S.A. 107*, 16096−16100.
(87) Harms, M. J., Castaneda, C. A., Schlessman, J. L., Sue, G. R., Isom, D. G., Cannon, B. R., and Garcia-Moreno, B. (2009) The p*Ka* values of acidic and basic residues buried at the same internal location in a protein are governed by different factors. *J. Mol. Biol. 389*, 34−47.
(88) Spitzner, N., Lohr, F., Pfeiffer, S., Koumanov, A., Karshikoff, A., and Rüterjans, H. (2001) Ionization properties of titratable groups in ribonuclease $T_1$. I. p*Ka* values in the native state determined by two-dimensional heteronuclear NMR spectroscopy. *Eur. Biophys. J. 30*, 186−197.
(89) Tanford, C., and Kirkwood, J. G. (1957) Theory of titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc. 79*, 5333−5339.
(90) Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1996) The determination of p$K_a$s in proteins. *Biochemistry 35*, 7819−7833.
(91) Khare, D., Alexander, P., Antosiewich, J., Bryan, P., Gilson, M., and Orban, J. (1997) p$K_a$ measurements from nuclear magnetic resonance for B1 and B2 immunoglobin G-binding domain of protein G: comparison with calculated values for nuclear magnetic resonance and X-ray structures. *Biochemistry 36*, 3580−3589.
(92) Nabuurs, S. B., Nederveen, A. J., Vranken, W., Doreleijers, J. F., Bonvin, A. M. J. J., Vuister, G. W., Vriend, G., and Spronk, C. A. E. M. (2004) DRESS: a database of REfined solution NMR structures. *Proteins 55*, 483−486.
(93) Powers, N., and Jensen, J. H. (2006) Chemically accurate protein structures: validation of protein NMR structures by comparison of measured and predicted pKa values. *J. Biomol. NMR 35*, 39−51.
(94) Charbonnier, J. B., Belin, P., Moutiez, M., Stura, E. A., and Quemeneur, E. (1999) On the role of the cis-proline residue in the active site of DsbA. *Protein Sci. 8*, 96−105.