

Preferred Hydrogen-Bonding Partners of Cysteine: Implications for Regulating Cys Functions

Karine Mazmanian,^{†,‡,§} Karen Sargsyan,[†] Cédric Grauffel,[†] Todor Dudev,^{||} and Carmay Lim^{*,†,⊥}

[†]Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan

[‡]Chemical Biology and Molecular Biophysics Program, Taiwan International Graduate Program, Academia Sinica, Taipei 11529, Taiwan

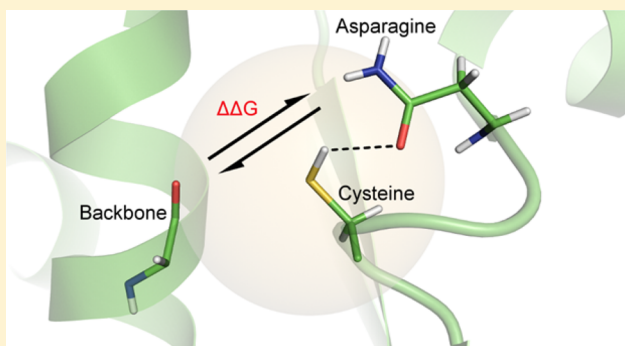
[§]Institute of Biochemical Sciences, National Taiwan University, Taipei 10617, Taiwan

^{||}Faculty of Chemistry and Pharmacy, Sofia University, Sofia 1164, Bulgaria

[⊥]Department of Chemistry, National Tsing Hua University, Hsinchu 300, Taiwan

S Supporting Information

ABSTRACT: The hydrogen-bonding interactions of cysteine, which can serve as a hydrogen-bond donor and/or acceptor, play a central role in cysteine's diverse functional roles in proteins. They affect the balance between the neutral thiol (SH) or thiolate (S[−]) and the charge distribution in the rate-limiting transition state of a reaction. Despite their importance, no study has determined the preferred hydrogen-bonding partners of cysteine serving as a hydrogen-bond donor or acceptor. By computing the free energy for displacing a peptide backbone hydrogen-bonded to cysteine with amino acid side chains in various protein environments, we have evaluated how the strength of the hydrogen bond to the cysteine thiol/thiolate depends on its hydrogen-bonding partner and its local environment. The predicted hydrogen-bonding partners preferred by cysteine are consistent with the hydrogen-bonding interactions made by cysteines in 9138 nonredundant X-ray structures. Our results suggest a mechanism to regulate the reactivity of cysteines and a strategy to design drugs based on the hydrogen-bonding preference of cysteine.



INTRODUCTION

Although cysteine (Cys) is the second least abundant amino acid (aa) residue (only ~1.9% of all aa residues in proteins), it nevertheless plays diverse functional roles in proteins (see Chart 1).¹ Cysteine can play a catalytic role by acting as (i) a potent nucleophile in enzymes such as cysteine proteases and ubiquitin ligases/deubiquitinases, (ii) an acid/base in abstracting or donating a proton to or from a substrate during the catalytic reaction, and (iii) a transition-state stabilizer (e.g., in stabilizing an oxyanion hole). It can also play a catalytic role by activating a water molecule, cofactor, or substrate or by participating in redox reactions catalyzed by enzymes in the thiol oxidoreductase family.^{2–4} Apart from a catalytic role, Cys can play a structural role to stabilize the protein by forming disulfide bridges⁵ or by binding metal ions such as Fe, Zn, Cd, and Hg to form mono/polynuclear clusters.⁶ In addition to its catalytic and structural roles, Cys can serve a regulatory role by acting as a switch to regulate the activity of proteins such as transcription factors, protein kinases, and metabolic enzymes.^{4,7}

The wide variety of Cys functional roles can be attributed to the physicochemical properties and rich chemistry of its side chain.⁶ The Cys chemical reactivity in a protein at physiological pH is dictated by its solvent accessibility and its hydrogen-

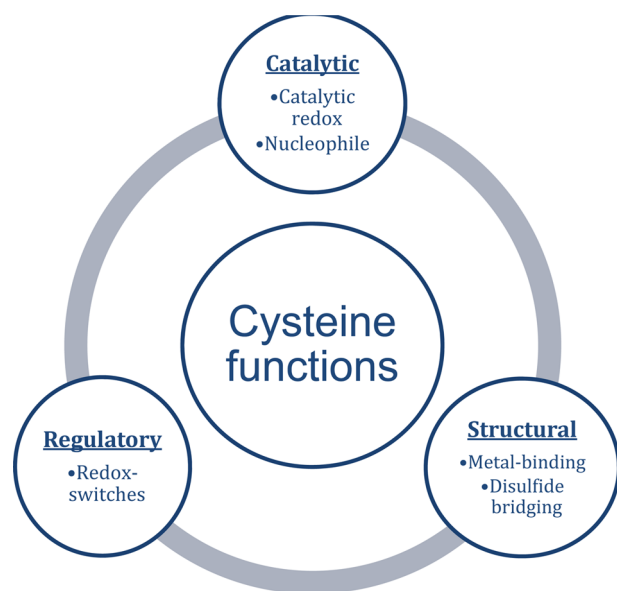
bonding interactions, which can tip the balance between the neutral thiol (SH) or thiolate (S[−]) to favor the protonated or deprotonated state^{8,9} and can stabilize charges in the rate-limiting transition state of a reaction.^{10,11} Among the 20 aa residues, Cys is found to be the least solvent-exposed residue in proteins.¹ It can serve as a hydrogen bond (HB) donor when protonated as well as a HB acceptor in both protonated and deprotonated states.^{4,8–10,12} Thus, elucidating the factors governing the hydrogen-bonding interactions with the Cys thiol or thiolate is crucial for understanding and predicting its reactivity.

Compared to studies on HBs to CH, NH, and OH, there are fewer studies on HBs to SH or S[−] despite their ubiquitous presence in proteins. Most of these studies have focused on HBs involving S formed by small molecules (e.g., hydrogen sulfide) to elucidate the directionality, nature, and strength of such HBs.^{13–18} The few studies on HBs to Cys in proteins have revealed the following:

Received: August 11, 2016

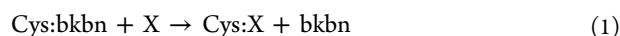
Revised: September 16, 2016

Published: September 16, 2016

Chart 1. Diagram Summarizing the Different Functional Roles of Cys in Proteins

- (1) Cys is found more often as a HB donor than as a HB acceptor in protein X-ray structures by a ratio of 5:1.¹⁹
- (2) Compared to HBs to oxygen, those to the larger and more polarizable sulfur are longer^{19,20} with a mean H...S(acceptor) distance of 2.80 ± 0.26 Å.
- (3) The Cys side chain in an α helix frequently forms HBs with the backbone oxygen, stabilizing the helix.²⁰
- (4) Intermolecular HBs formed between the Cys(S) and the backbone N can lower the pK_a of a N-terminal Cys in a helix,²¹ whereas intramolecular HBs formed between the Cys(SH) and the backbone O can lower the pK_a of active-site Cys.²²
- (5) HBs to a Zn-bound Cys(S^-) in Zn finger proteins reduce the negative charge on all Zn-bound Cys(S^-) atoms, even those with no hydrogen-bonding interactions.¹⁰

Although hydrogen-bonding interactions of the Cys side chain clearly depend on its partner, little attention has been paid to cysteine's preference for its HB donor or acceptor partner. Hence, it is not known (i) if there is a preference among the aa residues or peptide backbone to accept/donate a HB from/to protonated Cys(SH) or deprotonated Cys(S^-), (ii) which residues are preferred, and (iii) how the protein environment of the Cys would modulate this preference. To address this, we have focused on metal-free unmodified cysteines in this study and have examined the competition among the various groups in a protein to accept a HB from neutral Cys(SH) or donate a HB to protonated Cys(SH) or deprotonated Cys(S^-). The outcome of this competition was assessed by the free energy, ΔG^ϵ , for displacing the peptide backbone (denoted as bkbn) that is hydrogen-bonded to Cys(SH) or Cys(S^-) with an aa side chain X in a protein environment characterized by an effective dielectric constant ϵ ; i.e.,



In eq 1, X is one of the possible hydrogen-bonding partners of Cys, viz., the side chains of Ser/Thr, Tyr, Asn/Gln, His, Arg, or Lys. The different protein environments of the Cys, which may be buried or solvent-exposed, were taken into account by

computing the ΔG^ϵ for ϵ ranging from 4 to 80 according to eq 2:

$$\Delta G^\epsilon = \Delta G^1 + \Delta \Delta G_{\text{solv}}^\epsilon \quad (2)$$

In eq 2, ΔG^1 is the gas-phase free energy for eq 1, while $\Delta \Delta G_{\text{solv}}^\epsilon$ is the solvation free energy difference between the products and reactants for a given ϵ . A negative ΔG^ϵ implies that Cys prefers to interact with the aa side chain X rather than the peptide backbone in an environment characterized by ϵ . The results were then compared with trends from a statistical survey of protein X-ray structures. They reveal that cysteines from protein X-ray structures usually form HBs in accord with the results from the computed free energies. The trends obtained help to understand the reactivity and diverse functional roles of cysteine.

METHODS

Models Used. The aa residue side chains in their neutral or ionized form at physiological pH were modeled by the model compounds, as summarized in Table 1. All models were built using GaussView 5.²³

Table 1. Compounds Modeling Amino Acid Side Chains and Peptide Backbone Groups

residue	model compound	chemical formula
backbone	N-methylacetamide	$\text{CH}_3\text{CONHCH}_3$
Arg ⁺	methylguanidinium	$\text{C}_2\text{H}_8\text{N}_3^+$
Asn/Gln	acetamide	CH_3CONH_2
Asp ⁻ /Glu ⁻	acetate	CH_3COO^-
Cys	methanethiol	CH_3SH
Cys ⁻	methanethiolate	CH_3S^-
His ⁰	imidazole	$(\text{CH})_3\text{N}(\text{NH})$
His ⁺	imidazolium	$(\text{CH})_3(\text{NH})_2^+$
Lys ⁺	methylammonium	CH_3NH_3^+
Met	dimethyl sulfide	$(\text{CH}_3)_2\text{S}$
Ser/Thr	methanol	CH_3OH
Trp	3-methylindole	$\text{C}_9\text{H}_9\text{N}$
Tyr	phenol	$\text{C}_6\text{H}_6\text{O}$

Calibrating the Geometry Optimization Method. To obtain reliable geometries of the hydrogen-bonded complexes involving sulfur, the experimentally determined structure of CH_3SH ²⁴ was used to calibrate the geometry using M06 functionals (M06, M06-L, M06-HF, M06-2X)²⁵ in combination with different Pople's basis sets containing both polarization and diffuse functions. The four M06 functionals were chosen, as they account for the dispersion contribution and have been recommended for noncovalent interactions.²⁶ The results in Supporting Table S1 show that the M06-2X functional with the 6-311++G(d,p) efficiently reproduced the experimentally observed CH_3SH geometry. The computed C–S (1.82 Å) and S–H (1.34 Å) bond distances are identical to the measured ones, while the H–S–C and H–C–H angles are within 1.2° of the experimental values (96.5 and 109.8°, respectively). In previous studies, the M06-2X functional was found to reproduce the relative energies of Cys conformers computed using more accurate CBS-QB3, G3B3, G4MP2, and G4 composite methods.²⁷ Out of 64 tested density functionals, M06-2X was found to be the most accurate one for predicting pK_a values of aa residues and proton transfer between different residues.^{28,29} Thus, the M062X/6-311++G(d,p) method was used to compute the geometries and the vibrational frequencies

Table 2. Root-Mean-Square Deviations (RMSDs) of the D⋯A and H⋯A Distances (Å) and the D–H⋯A Angles (deg) in the M062X/6-311++G(d,p) Optimized Structures from Those in the Corresponding CSD Structure

CSD ID (HB type)	RMSD (Å) (D⋯A and H⋯A)	RMSD (deg) (D–H⋯A)
WUJWAZ (S–H⋯O)	0.05	0.4
RONVEV (O–H⋯S; S–H⋯S)	0.04	5.5
ADABIQ (N–H⋯S)	0.05	0.9
YIRGEM (S–H⋯N)	0.10	1.6
BABCOW (O–H⋯S; N–H⋯S)	0.11	4.8

Table 3. Comparison between Computed and Experimental Solvation Free Energies (in kcal/mol) and pK_a Values^a

molecule AH	$\Delta G_{\text{solv}}^{\text{80}}$ (AH)	$\Delta G_{\text{solv}}^{\text{80}}$ (A)	pK _a
acetic acid CH ₃ COOH	−6.7 (−6.7 ^b)	−76.3	4.9 (4.75)
imidazolium (CH) ₃ N(NH)	−62.6	−9.9 (−10.2 ^c)	7.2 (7.1)
phenol C ₆ H ₆ O	−6.4 (−6.6 ^b)	−70.1	10.2 (10.0)
methanethiol CH ₃ SH	−1.0 (−1.2 ^b)	−73.5	10.7 (10.4)
methylammonium CH ₃ NH ₃ ⁺	−73.4	−4.4 (−4.6 ^b)	10.9 (10.6)
methylguanidinium C ₂ H ₈ N ₃ ⁺	−56.7	−11.1 (−11.2 ^d)	13.6 (13.4)
methanol CH ₃ OH	−5.0 (−5.1 ^b)	−96.0	15.8 (15.5)
N-methylacetamide CH ₃ CONHCH ₃	−9.7 (−10.0 ^e)		
acetamide CH ₃ CONH ₂	−9.6 (−9.7 ^f)		
dimethyl sulfide (CH ₃) ₂ S	−1.4 (−1.5 ^g)		
3-methylindole C ₉ H ₁₀ N	−5.9 (−5.9 ^g)		

^aExperimental values, where available, are in parentheses. The pK_a values were computed using the experimental hydration free energy of the proton (−264.0 kcal/mol).³⁵ ^bFrom Kelly, 2005.⁴² ^cFrom Lim et al., 1991.³⁹ ^dFrom Vorobyov et al., 2008.⁴³ ^eFrom Wolfenden, 1978.⁴⁴ ^fFrom Wolfenden, 1981.⁴⁵ ^gFrom Sitkoff et al., 1994.⁴⁶

of the Cys complexes using the Gaussian 09³⁰ program. No imaginary frequencies were found in any of the fully optimized structures.

Next, we tested the performance of the M062X/6-311++G(d,p) method in reproducing the experimental geometries of Cambridge Structural Database (CSD)³¹ structures with intramolecular or intermolecular sulfur-centered HBs. Their fully optimized structures are shown in Supporting Figure S1 along with the donor (D)⋯acceptor (A) and H⋯A distances as well as D–H⋯A angles. The root-mean-square deviations (RMSDs) of the D⋯A and H⋯A distances as well as the D–H⋯A angles in the M062X/6-311++G(d,p) optimized structures from those in the corresponding CSD structures (Table 2) show that they are generally comparable to the respective differences (∼0.05 Å and 4.3°) between the X-ray and neutron diffraction structures (BABCOW01) of the same molecule.³² Thus, the M062X/6-311++G(d,p) method could model HBs involving sulfur with different binding partners.

Calibrating Gas-Phase Free Energy Calculations. The vibrational frequencies were scaled by an empirical factor of 0.979³³ and used to compute the change in the thermal energy including zero-point energy, ΔE_{th} , and entropy ΔS for eq 1. The differences in the electronic energy ΔE_{el} , ΔE_{th} , ΔPV (work term), and ΔS between the products and reactants in eq 1 were used to compute the gas-phase free energy ΔG^1 at a temperature T of 298.15 K using the Gaussian 09³⁰ program according to eq 3:

$$\Delta G^1 = \Delta E_{\text{el}} + \Delta E_{\text{th}} + \Delta PV - T\Delta S \quad (3)$$

To choose an optimal method for computing ΔG^1 , the experimental ΔG^1 for deprotonating CH₃SH (351.3 ± 0.4 kcal/mol at 298.15 K)³⁴ was employed, as deprotonation and hydrogen-bonding interactions of the Cys model are interconnected.²⁷ Thus, the ΔG^1 for deprotonating CH₃SH at 298.15 K was computed using various density functionals and

basis sets probed in previous studies²⁸ and the ΔG^1 of the proton (−6.28 kcal/mol³⁵). The computed CH₃SH deprotonation free energies in Supporting Table S2 show that three density functionals (B-971, B-98, and PBE1-PBE) along with the 6-31+G(2d,2p) basis set could reproduce the experimental ΔG^1 for deprotonating CH₃SH to within 0.1 kcal/mol. Since the B-98/6-31+G(2d,2p) method had been shown to perform well in computing the energies of hydrogen-bonding and weak noncovalent interactions,^{36,37} it was chosen to compute the ΔG^1 of eq 1. Basis set superposition error was estimated using the counterpoise method and added to ΔG^1 to yield the final energies reported herein. We also verified that the B-98/6-31+G(2d,2p) method yielded a very similar ΔG^1 for eq 1 as the B-971 and PBE1-PBE functionals with the same basis set (see Supporting Table S3).

Calibrating the Solvation Free Energy Calculations.

The solvation free energy $\Delta G_{\text{solv}}^{\text{e}}$ was estimated by solving Poisson's equation using finite difference methods^{38,39} with the MEAD (Macroscopic Electrostatics with Atomic Detail) program,⁴⁰ as described in previous works.⁴¹ The continuum dielectric calculations employed M062X/6-311++G(d,p) optimized geometries and natural bond orbital atomic charges as well as effective solute radii R_{eff} (Supporting Table S4) that have been adjusted to reproduce the experimental solvation free energies and/or pK_a of the compounds in Table 3.

Database of Nonredundant Cys Structures. To verify the trends from the computed free energies, we retrieved <2.5 Å X-ray structures of Cys-containing proteins from the Protein Data Bank (PDB).⁴⁷ Representative structures (sharing <30% sequence identity) were selected according to the algorithm described in <http://www.rcsb.org/pdb/statistics/clusterStatistics.do>. This resulted in 9138 nonredundant structures containing 52,698 metal-free, unmodified Cys residues that are *not* involved in disulfide bond formation, as described in the PDB files.

Using 19 protein structures determined by neutron diffraction, Zhou et al.¹⁹ had assessed the hydrogen positions predicted by the program REDUCE⁴⁸ and found an average root-mean-square error of 0.22 Å. Hence, the REDUCE program was used to add hydrogen atoms to each of the 9138 nonredundant structures. Only heavy atoms with *B*-factors <40 Å² were analyzed for HBs, as defined below.

HB Criteria. HBs formed by a Cys were assessed using geometric criteria established from meta-analyses of previous studies including statistical analyses of S-containing HBs in CSD and PDB X-ray structures.^{19,49} A HB from the donor Cys SH group to a heavy atom acceptor A was defined by $S\cdots A < 4.3$ Å, $H\cdots A < 3.0$ Å, $S-H\cdots A > 90^\circ$, and $H(D)\cdots A-AA$ or $H\cdots A-bisector > 90^\circ$ (Figure 1). On the other hand, a HB to the

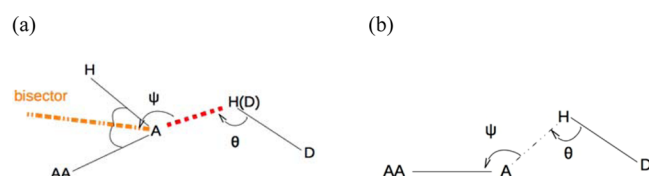


Figure 1. HBs between acceptor (A) and donor (D) heavy atoms. AA is the acceptor antecedent heavy atom. (a) Acceptor is covalently bonded to a hydrogen, e.g., SH(Cys), OH(Ser), or another heavy atom, e.g., N(His) or S(Met). (b) Acceptor has no hydrogen, e.g., S^- , sp^2 O, or O^- .

acceptor Cys SH or S^- atom from a heavy atom donor D was defined by $D\cdots S < 4.1$ Å, $H\cdots S < 3.2$ Å, $D-H\cdots S > 90^\circ$, and $H\cdots S-AA$ or $H\cdots S^H-bisector > 80^\circ$. Note that, if only the heavy atom–heavy atom distance ($S\cdots A < 4.3$ Å or $D\cdots S < 4.1$ Å) criterion was met but the other geometric requirements were not, a HB was not considered to be present in the X-ray structure.

RESULTS

Preferred HB Acceptors of Cysteine Thiol. The His imidazole N, Ser/Thr/Tyr hydroxyl O, the peptide backbone/Asn/Gln carbonyl O, the Asp/Glu carboxylate O, and the Cys/Met S can all accept a HB from protonated Cys⁰ thiol. To assess which of these aa residues prefer to accept HBs from the Cys H(S), their hydrogen-bonded complexes with methyl thiol were optimized at the M062X/6-311++G(d,p) level (see Supporting Table S5 for the HB parameters). During optimization of methyl thiol hydrogen-bonded to acetate, proton transfer occurred forming $CH_3S^-\cdots CH_3COOH$. To assess if such proton transfer might occur in the protein, the $CH_3SH\cdots CH_3COO^-$ complex was optimized using the conductor-like polarizable continuum model (CPCM) in the Gaussian 09³⁰ program with a dielectric constant ϵ ranging from 4 to 80. The results show that the proton remained on CH_3SH with the sulfur 3.24 Å from the nearest carboxylate O. Hence, the $CH_3SH\cdots CH_3COO^-$ hydrogen-bonded complex was reoptimized with the $S\cdots O^-$ hydrogen-bonding distance constrained to 3.24 Å. The resulting structure and the other fully optimized geometries (Figure 2) were used to compute the free energies ΔG^ϵ ($\epsilon = 1-80$) for replacing the peptide backbone with various HB acceptor side chains (Supporting Table S6).

In the absence of the protein, the SH group strongly prefers to donate a HB to the Asp⁻/Glu⁻ carboxylate ($\Delta G^1 = -11.8$ kcal/mol, Table 4), forming the shortest and most linear HB

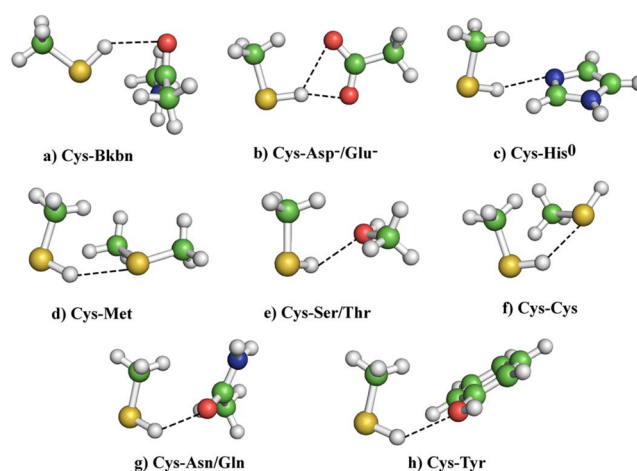


Figure 2. Fully optimized hydrogen-bonded complexes of neutral Cys⁰ as a HB donor, except the $CH_3SH\cdots CH_3COO^-$ complex (Figure 2b), which was optimized with the $S-H\cdots O^-$ distance constrained to 3.24 Å. Carbon in green, nitrogen in blue, oxygen in red, sulfur in yellow, and hydrogen in off-white.

Table 4. Computed Free Energies (ΔG^ϵ in kcal/mol) for $CysSH:bkbk + X \rightarrow CysSH:X + bkbk$, where X is a HB Acceptor, in a Medium with an Effective Dielectric Constant ϵ^a

HB acceptor	ΔG^1	ΔG^4	ΔG^{30}	ΔG^{80}
acetate (Asp ⁻ /Glu ⁻) ^b	-11.8	-3.0	-0.6	-0.3
imidazole (His ⁰)	-1.8	-1.7	-1.9	-1.9
dimethylsulfide (Met)	-1.1	-1.8	-2.0	-2.0
methanol (Ser/Thr)	-0.3	-0.1	0.2	0.2
methylthiol (Cys)	-0.2	-0.8	-1.0	-1.0
acetamide (Asn/Gln)	0.5	0.8	0.9	0.9
phenol (Tyr)	0.7	0.5	0.6	0.7

^a $\Delta G^\epsilon < -2.5$ kcal/mol are highlighted in bold. ^bThe geometry of the $CH_3SH\cdots CH_3COO^-$ hydrogen-bonded complex was constrained-optimized at the M062X/6-311++G(d,p) level with the $S\cdots O^-$ hydrogen-bonding distance fixed at 3.24 Å.

compared with the other aa side chains (Supporting Table S5). The anionic Asp⁻/Glu⁻ strongly polarizes the S–H bond: the NBO charge on S in the $CH_3SH\cdots CH_3COO^-$ complex ($-0.19e$) is significantly more negative than that in free Cys⁰ ($-0.04e$). In contrast, the neutral aa side chains do not significantly polarize the S–H bond, as they increase the S negative charge in free Cys⁰ by only 0.01–0.04e (see Supporting Figure S2).

In the presence of the protein, the preference for Asp⁻/Glu⁻ is significantly dampened because the solvation free energy of the bulky anionic $CH_3SH\cdots CH_3COO^-$ complex cannot compensate for the larger desolvation penalty of CH_3COO^- . Hence, the Cys thiol favors the Asp⁻/Glu⁻ carboxylate over the backbone carbonyl group only when it is quite buried ($\Delta G^4 = -3$ kcal/mol). However, when Cys⁰ is partially/fully solvent exposed, it is not very discriminatory toward its HB acceptors: the ΔG^ϵ values differ by no more than 2 kcal/mol.

Preferred HB Donors of Cysteine Thiol. In addition to serving as a HB donor, Cys⁰ can accept HBs from the backbone/His/Trp NH, Asn/Gln NH₂, Ser/Thr/Tyr OH, the Cys SH, the Lys ammonium, and the Arg amino/imino groups. Interestingly, the HB parameters in Supporting Table S5 show that Cys⁰ formed shorter and more linear HBs with the

positively charged His⁺, Lys⁺, and Arg⁺ (H...S < 2.5 Å, N–H...S > 150°) than with neutral HB donors (H...S > 2.5 Å, N–H...S < 150°). Furthermore, Arg⁺ donates two hydrogen bonds to CysSH, unlike the other residues. The fully optimized structures of the hydrogen-bonded complexes with Cys⁰ as a HB acceptor in Figure 3 were used to compute the ΔG^e free energies for replacing the peptide backbone with various HB donor side chains (Supporting Table S7).

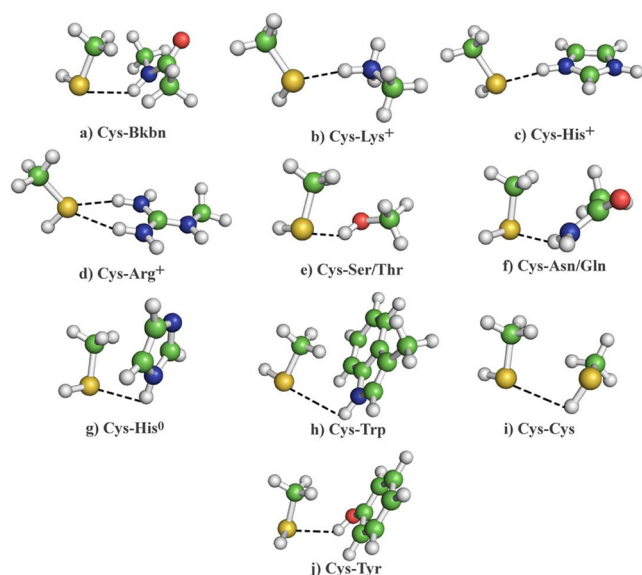


Figure 3. Optimized hydrogen-bonded complexes of neutral Cys⁰ as a HB acceptor. The coloring scheme is the same as that in Figure 2.

The Cys⁰ thiol is more discriminatory as a HB acceptor than as a HB donor and prefers to interact with aa side chains rather than the peptide backbone (ΔG^e all negative, Table 5).

Table 5. Computed Free Energies (ΔG^e in kcal/mol) for CysSH:bkbk + X → CysSH:X + bkbk, Where the Thiol Is a HB Acceptor, in a Medium with an Effective Dielectric Constant ϵ^a

HB donor	ΔG^1	ΔG^4	ΔG^{30}	ΔG^{80}
methyllumonium (Lys ⁺)	−16.9	−7.3	−4.9	−4.6
imidazolium (His ⁺)	−13.6	−7.4	−5.6	−5.4
methylguanidinium (Arg ⁺)	−12.4	−7.7	−6.4	−6.3
methanol (Ser/Thr)	−3.5	−3.6	−3.6	−3.6
acetamide (Asn/Gln)	−2.2	−2.3	−2.2	−2.2
imidazole (His ⁰)	−1.9	−1.4	−1.3	−1.4
3-methylindole (Trp)	−1.8	−1.6	−1.3	−1.3
methanethiol (Cys)	−1.6	−2.0	−2.0	−2.1
phenol (Tyr)	−1.2	−1.2	−1.0	−1.0

^a $\Delta G^e < -2.5$ kcal/mol are highlighted in bold.

Whereas the Cys⁰ SH shows no strong preference for any HB acceptors except for Asp[−]/Glu[−] when buried, its S strongly prefers to accept a HB from positively charged aa side chains (His⁺, Lys⁺, or Arg⁺) and to a lesser extent from the Ser/Thr hydroxyl group, regardless of its solvent exposure. However, increasing solvent exposure of Cys⁰ attenuates the HB strength of the charged donors whose desolvation penalties exceed the solvation free energies of their hydrogen-bonded complexes. Nevertheless, Cys⁰ still prefers charged HB donors to the backbone amide even when it is solvent exposed.

Preferred HB Donors of Deprotonated Cysteine.

Unlike Cys⁰, the Cys thiolate (Cys[−]) can only serve as a HB acceptor. In the fully optimized structures of Cys[−] hydrogen-bonded to Lys⁺/His⁺, the thiolate became protonated (Figure 2b,c). Proton transfer from Lys⁺/His⁺ to Cys[−] was also observed in the CPCM-optimized hydrogen-bonded structures except when $\epsilon = 80$. Hence, a Cys[−]...Lys⁺/His⁺ structure was obtained by constraining the S[−]...N distance to the value obtained using CPCM with $\epsilon = 80$. The resulting Cys[−]...Lys⁺/His⁺ structure and the other fully optimized geometries in Figure 4 were used to compute the ΔG^e (Supporting Table S8).

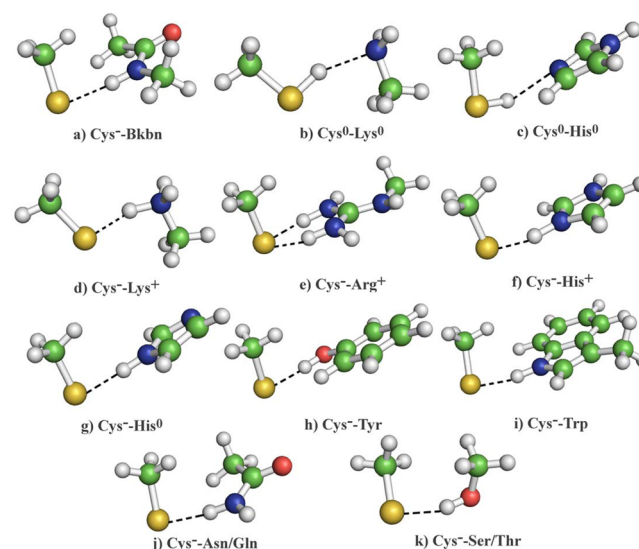


Figure 4. Optimized hydrogen-bonded complexes of Cys[−] as HB acceptor. The coloring scheme is the same as that in Figure 2.

Like the Cys⁰ thiol as a HB acceptor, Cys[−] is also discriminatory toward its HB donors, but the preference for charged HB donors depends on its relative solvent exposure: When the thiolate group is relatively buried, it strongly prefers positively charged aa side chains (Lys⁺, Arg⁺, and His⁺) as HB donors (negative ΔG^4 , Table 6). When it is partially solvent exposed, it prefers Lys⁺ and His⁺ only if it could accept a proton from these side chains forming Cys⁰...Lys⁰/His⁰ HBs (negative ΔG^{30} , Table 6). On the other hand, when the thiolate group is fully solvent exposed, positively charged side chains become disfavored compared with the neutral peptide amide group because they form neutral Cys[−]...Lys⁺/His⁺/Arg⁺ complexes, whose solvation free energy gain cannot outweigh the cost of desolvating the cationic aa side chains (see Table 3).

Unlike the Cys⁰ thiol, Cys[−] prefers the His⁰ imidazole among the neutral HB donors and disfavors the Ser/Thr side chain, as compared to the backbone amide, especially when buried (positive ΔG^4 , Table 6). When Cys is protonated, the size and flexibility of its HB donor seem to define HB complex formation: In comparison with the backbone amide group, the smaller and less rigid Ser/Thr hydroxyl group can better overlap with the sulfur lone pair, forming a stronger HB with Cys⁰, as evidenced by the shorter O...S (3.33 Å) and H...S (2.51 Å) distances and the larger O–H...S angle (143°) compared to those formed by the backbone amide (N...S = 3.47 Å, H...S = 2.78 Å, and N–H...S angle = 126°). However, when Cys is deprotonated, charge(S[−])...dipole electrostatic

Table 6. Computed Free Energies (ΔG^e in kcal/mol) for $\text{CysS}^-:\text{bkbn} + \text{X} \rightarrow \text{CysS}^-\text{:X} + \text{bkbn}$, Where the Cys^0 Thiolate Is a HB Acceptor, in a Medium with an Effective Dielectric Constant ϵ^a

HB donor	ΔG^1	ΔG^4	ΔG^{30}	ΔG^{80}
methylammonium (Lys^+ transfers proton to Cys^-)	-122.4^b	-31.8^b	-3.4^b	-0.5^b
imidazolium (His^+ transfers proton to Cys^-)	-109.0^b	-29.7^b	-5.6^b	-3.2^b
methylammonium (Lys^+) ^c	-100.6^c	-19.7^c	4.3^c	6.7^c
methylguanidinium (Arg^+)	-90.1	-19.2	2.1	4.4
imidazolium (His^+) ^c	-88.1^c	-17.1^c	3.7^c	5.8^c
imidazole (His)	-3.6	-3.4	-3.0	-2.9
phenol (Tyr)	-1.7	-1.4	-1.2	-1.2
3-methylindole (Trp)	0.1	1.6	1.6	1.6
acetamide (Asn/Gln)	2.6	2.7	3.0	3.0
methanol (Ser/Thr)	7.8	2.8	1.2	1.1

^a $\Delta G^e < -2.5$ kcal/mol are highlighted in bold, whereas $\Delta G^e > 2.5$ kcal/mol are highlighted in italic. ^bBased on geometries optimized at the M062X/6-311++G(d,p) level in the gas phase where CysS^- becomes protonated. ^cThe geometry of the hydrogen-bonded complex was constrained-optimized at the M062X/6-311++G(d,p) level with the $\text{CysS}^-\cdots\text{H}(\text{N})$ distance constrained to 3.06 and 3.00 Å for Lys^+ and His^+ , respectively, to prevent proton transfer to CysS^- .

interactions become dominant in a low dielectric environment. As the dipole moment of the backbone amide or His^0 imidazole (~ 3.9 D) is roughly twice that of the Ser/Thr hydroxyl group (~ 1.9 D), a buried Cys^- prefers to accept a HB from the backbone amide or His^0 imidazole rather than the Ser/Thr side chain.

Preferred HB Partners of Cys from PDB Structures. To verify the preferred hydrogen-bonding partners of Cys predicted by the free energies ΔG^e , we analyzed the hydrogen-bonding interactions made by Cys in 9138 non-redundant X-ray structures in the PDB containing 52,698 metal-free Cys residues (see Methods). Since eq 1 represents a competition between the backbone and aa side chains to form HBs with Cys, we first identified all competing backbone and side chain donor/acceptor heavy atoms within 4.3 Å from the Cys S. Given the Cys protonation state assigned by the REDUCE program, each potential side chain X listed in Tables 4–6 was evaluated to see if it formed a HB with Cys according to geometric criteria (see Methods) and, if it did, whether its rival backbone acceptor/donor formed a HB or not. In the competition between side chain and backbone acceptor/donor to form HBs with Cys, let $N_{\text{X,W}}$ denote the number of side chain X winners (i.e., the number of side chain X that formed a HB with Cys but not the competing backbone) and let $N_{\text{X,L}}$ denote the number of side chain X losers (i.e., the number of side chain X that did not form a HB with Cys but the backbone did). Since hydrogen atoms are not seen in the X-ray structures and were added using the REDUCE program (see Methods), we attempted to reduce errors in the PDB analysis by using sufficient samples for statistical analyses. Thus, we computed the % fraction of side chain X winners ($f_{\text{X,W}}$) only when the total number of X competitors, $N_{\text{X}} = N_{\text{X,W}} + N_{\text{X,L}}$, exceeded 100. N_{X} was found to be >100 only for the side chains listed in Table 7, but it was zero or <100 for the other side chains.

The results in Table 7 show the % fraction of side chain X winners correlated with the respective ΔG^e ($\epsilon = 4\text{--}80$) in Tables 4 and 6. The $f_{\text{X,W}}$ of 59% for Cys as HB donor indicates that Cys has a weak preference to form a HB with another Cys

Table 7. Correlation between % Fraction of Side Chain X Winners ($f_{\text{X,W}}$) and ΔG^e (in kcal/mol) for $\text{Cys}:\text{bkbn} + \text{X} \rightarrow \text{Cys}:\text{X} + \text{bkbn}$ in a Medium with an Effective Dielectric Constant ϵ

HB type	X	N_{X}	$f_{\text{X,W}}$	ΔG^4	ΔG^{80}
$\text{CysSH}\cdots\text{S}$	Cys	708	59.4%	-0.8	-1.0
$\text{CysSH}\cdots\text{O}$	Ser/Thr	103	49.5%	-0.1	0.2
$\text{CysS}^-\cdots\text{H}$	Ser/Thr	103	46.4%	2.8	1.1

rather than the peptide backbone, consistent with the small negative ΔG^e (~ -1 kcal/mol) for Cys in Table 4. The $f_{\text{X,W}}$ of $\sim 50\%$ for Ser/Thr to accept a HB from protonated CysSH indicates that Cys^0 has no preference to donate a HB to the Ser/Thr hydroxyl or the backbone carbonyl group, in accord with the near zero ΔG^e for Ser/Thr in Table 4. In contrast, the $f_{\text{X,W}}$ for Ser/Thr to donate a HB to deprotonated Cys^- is $\sim 46\%$, in line with the small positive ΔG^e for Ser/Thr in Table 6.

DISCUSSION

The ability of Cys to form HBs has been challenged for a long time due to the low electronegativity of sulfur. Experimental and theoretical evidence^{15,50,51} now indicates that Cys can be a HB donor and/or acceptor and its hydrogen-bonding interactions influence its reactivity and thus biological functions.^{8,10,11,52} However, no study (to our knowledge) has determined the preferred hydrogen-bonding partners of Cys serving as a HB donor or acceptor. Here, we have elucidated how the strength of the HB to the Cys thiol/thiolate depends on its hydrogen-bonding partner and its local environment as well as the physical basis for the observed preference (see below).

Physical Basis for the Selectivity of Cys Hydrogen-Bonding Partners. The preferred hydrogen-bonding partner of Cys depends largely on the local protein environment (effective dielectric constant ϵ), which dictates the nature of the interaction stabilizing the HB. A buried Cys is quite selective toward its hydrogen-bonding partner, preferring charged aa side chains to neutral ones or the peptide backbone. For buried Cys residues, the nature of the electrostatic interactions dictates its preferred charged hydrogen-bonding partners: When the Cys is protonated, its weak S–H bond is polarized by HB interactions, inducing a dipole; hence, its HB strength depends on induced-dipole(S–H) \cdots charge/dipole interactions.¹⁷ On the other hand, when the Cys is deprotonated, its HB strength depends on charge(S^-) \cdots charge/dipole interactions.¹⁷ As induced-dipole(S–H)/charge(S^-) \cdots charge interactions are more favorable than induced-dipole(S–H)/charge(S^-) \cdots dipole interactions, a buried Cys prefers negatively charged $\text{Asp}^-/\text{Glu}^-$ as a HB acceptor (Table 4) and positively charged Lys^+ , His^+ , or Arg^+ as HB donors (Tables 5 and 6).

Solvent exposure of the Cys weakens its hydrogen-bonding partner selectivity. A solvent-exposed Cys is selective toward HB donors but not HB acceptors. Protonated Cys^0 still favors positively charged Lys^+ , His^+ , or Arg^+ as HB donors, but deprotonated Cys^- favors neutral imidazole instead of positively charged aa side chains. Thus, the Cys^- thiolate is stabilized by basic charged residues in a “dry” environment but by neutral aa residues when solvent exposed. Unlike buried Cys^0 , which favors $\text{Asp}^-/\text{Glu}^-$ as a HB acceptor, solvent-exposed Cys^0 exhibits no strong preference toward HB acceptors ($|\Delta G^{80}| \leq 2.0$ kcal/mol, Table 4).

Cysteine discerns its HB donor partners better than its HB acceptor ones: the $|\Delta G^{\circ}|$ free energies for HB acceptors are smaller (Table 4) than those for HB donors (Tables 5 and 6). Cysteine is relatively nonselective toward its HB acceptors because its weak S–H bond enables sharing of its H(S) with any HB acceptor. On the other hand, Cys is much more selective toward its HB donor because of its S lone pair orientations and the nature of the electrostatic interactions and solvent effects, as discussed above.

Comparison with Experiment and Previous Studies.

The above finding that Cys is less discriminatory as a HB donor than as a HB acceptor appears to be consistent with the empirical observation that Cys serves as a HB donor more often than as a HB acceptor in protein X-ray structures¹⁹ (see the Introduction). The calculations can generally reproduce (within experimental error) the average HB interaction geometries of Cys seen in protein X-ray structures¹⁹ (see Supporting Table S5). They predict that, in a solvent-inaccessible environment, a $\text{Cys}^0 \cdots \text{Lys}^0$ HB might be more stable than a $\text{Cys}^- \cdots \text{Lys}^+$ HB, whereas the latter is more stable than the former in an aqueous medium. Interestingly, this prediction is supported by IR spectroscopy,⁵³ which shows that, in the absence of water, both $\text{Cys}^0 \cdots \text{Lys}^0$ and $\text{Cys}^- \cdots \text{Lys}^+$ structures were found, but hydration tipped the equilibrium in favor of the $\text{Cys}^- \cdots \text{Lys}^+$ salt-bridged structure.

The outcome of the competition between the backbone and α side chains to form HBs with Cys predicted by the ΔG° (eq 1) appears consistent with that deduced from 9138 non-redundant X-ray structures (see Table 7). Also consistent with the positive ΔG° computed for Lys^+ , His^+ , or Arg^+ in Table 6, basic charged residues are rarely observed near nucleophilic Cys^- in the solvent-exposed active sites of thioredoxin-fold enzymes; instead, these solvent-exposed reactive thiolates usually form HBs with neutral polar groups, especially backbones.^{8,52} In line with the results in Table 6 showing that CysS^- forms a strong HB with protonated His^+ when buried or neutral His^0 when exposed, the nucleophilic Cys30 in the disulfide-binding protein A forms a strong HB with a conserved His32⁵⁴ that accounts for its unusually depressed pK_a of ~ 3.5 .⁵⁵

Implications for Regulating Cys Functions by Modulating HBs. Our results, revealing how the strength of a HB to Cys varies with the physicochemical properties of its hydrogen-bonding partner, suggest a mechanism to regulate the reactivity of cysteines: Changing the cysteine's hydrogen-bonding partner and dielectric environment would alter the HB strength and thus the degree of thiolate stabilization, which in turn would affect the cysteine's nucleophilicity/reactivity. The cysteine's hydrogen-bonding partner and dielectric environment could be altered via protein conformational changes that position different hydrogen-bonding partners and limit/enable water access. The cysteine's hydrogen-bonding partner could also be altered via mutagenesis during evolution or in vitro/in vivo experiments. For example, mutations of the wild-type His32 in the disulfide-binding protein A, which forms a strong HB with the nucleophilic Cys30, to Gly increased the pK_a of catalytic Cys30 from 3.5 to 4.85, leading to a decrease in the enzyme's redox potential.⁵⁶

Our results also suggest a strategy to design drugs on the basis of the HB preferences of Cys: If the drug target is an enzyme requiring a catalytic Cys, a drug candidate could be designed to form HBs with a catalytic Cys^- to lower/abolish its intrinsic nucleophilicity, thus deactivating the enzyme. In fact,

there are some noncovalent inhibitors that rely on HBs to an active site Cys.⁵⁷ For example, the pharmacophore ketone group of the series of hydroxymethyl ketone inhibitors forms HBs to Cys25 in cysteine protease cruzain from protozoan parasite *Trypanosoma cruzi*.⁵⁷ As another example, the protonated amine nitrogen of a potent benzylamine inhibitor of caspase-1 is positioned within hydrogen-bonding distance to the active-site Cys.⁵⁸ The drug target can also be a viral protein where HBs to Cys are important for maintaining the integrity of the protein structure. This is exemplified by the viral DNA packaging P22 portal,⁵⁰ where HBs to Cys have been shown to play an important role in protein complex assembly; in this case, a drug candidate could be designed to compete with the HB acceptors/donors of the "structural" Cys.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b08109.

Density functional and basis set calibration data for geometries (Table S1) and gas-phase free energies (Table S2); comparison of the three best-performing functionals in computing gas-phase free energies (Table S3); calibrated effective solute radii for the amino acid model compounds (Table S4); hydrogen-bond parameters of optimized Cys hydrogen-bonded complexes and comparison with available experimental data (Table S5); thermochemical data for calculated free energies at $T = 298.15$ K (Tables S6–S8); superimposition of M062X/6-311++G(d,p)-optimized and X-ray structures of compounds from Cambridge Structural Database structures (Figure S1); NBO sulfur charges of the free and hydrogen-bonded Cys complexes (Figure S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: carmay@gate.sinica.edu.tw. Phone: +886-2-26523031 (Office); +886-2-27899043 (Lab).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Yu-Ming Lee for helpful comments. This work was supported by funds from Academia Sinica and MOST (Grant No. 98-2113-M-001-011), Taiwan.

■ REFERENCES

- (1) Marino, S. M.; Gladyshev, V. N. Cysteine Function Governs Its Conservation and Degeneration and Restricts Its Utilization on Protein Surfaces. *J. Mol. Biol.* **2010**, *404*, 902–916.
- (2) Harrison, M. J.; Burton, N. A.; Hillier, I. H. Catalytic Mechanism of the Enzyme Papain: Predictions with a Hybrid Quantum Mechanical/Molecular Mechanical Potential. *J. Am. Chem. Soc.* **1997**, *119*, 12285–12291.
- (3) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* **2002**, *324*, 105–121.
- (4) Pace, N. J.; Weerapana, E. Diverse functional roles of reactive cysteines. *ACS Chem. Biol.* **2013**, *8*, 283–296.
- (5) Hogg, P. J. Disulfide bonds as switches for protein function. *Trends Biochem. Sci.* **2003**, *28*, 210–214.

- (6) Giles, N. M.; Watts, A. B.; Giles, G. I.; Fry, F. H.; Littlechild, J. A.; Jacob, C. Metal and redox modulation of cysteine protein function. *Chem. Biol.* **2003**, *10*, 677–693.
- (7) Barford, D. The role of cysteine residues as redox-sensitive regulatory switches. *Curr. Opin. Struct. Biol.* **2004**, *14*, 679–686.
- (8) Roos, G.; Foloppe, N.; Messens, J. Understanding the pKa of Redox Cysteines: The Key Role of Hydrogen Bonding. *Antioxid. Redox Signaling* **2013**, *18*, 94–127.
- (9) Marino, S. M. Protein flexibility and cysteine reactivity: influence of mobility on the H-bond network and effects on pKa prediction. *Protein J.* **2014**, *33*, 323–336.
- (10) Lee, Y.-M.; Lim, C. Factors Controlling the Reactivity of Zinc Finger Cores. *J. Am. Chem. Soc.* **2011**, *133*, 8691–8703.
- (11) Ferrer-Sueta, G.; Manta, B.; Botti, H.; Radi, R.; Trujillo, M.; Denicola, A. Factors Affecting Protein Thiol Reactivity and Specificity in Peroxide Reduction. *Chem. Res. Toxicol.* **2011**, *24*, 434–450.
- (12) Desiraju, G. R. A Bond by Any Other Name. *Angew. Chem., Int. Ed.* **2011**, *50*, 52–59.
- (13) Platts, J. A.; Howard, S. T.; Bracke, B. R. F. Directionality of Hydrogen Bonds to Sulfur and Oxygen. *J. Am. Chem. Soc.* **1996**, *118*, 2726–2733.
- (14) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. Ab initio study of hydrogen-bonded complexes of small organic molecules with water. *J. Phys. Chem. A* **1998**, *102*, 3782–3797.
- (15) Wennmohs, F.; Staemmler, V.; Schindler, M. Theoretical investigation of weak hydrogen bonds to sulfur. *J. Chem. Phys.* **2003**, *119*, 3208–3218.
- (16) Melandri, S. “Union is strength”: How weak hydrogen bonds become stronger. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13901–13911.
- (17) Biswal, H. S.; Bhattacharyya, S.; Bhattacharjee, A.; Wategaonkar, S. Nature and Strength of Sulfur Centered Hydrogen Bonds: Laser Spectroscopic Investigations in Gas Phase and Quantum Chemical Calculations. *Int. Rev. Phys. Chem.* **2015**, *34*, 99–160.
- (18) Antonijević, I. S.; Janjić, G. V.; Milčić, M. K.; Zarić, S. D. Preferred Geometries and Energies of Sulfur–Sulfur Interactions in Crystal Structures. *Cryst. Growth Des.* **2016**, *16*, 632–639.
- (19) Zhou, P.; Tian, F.; Lv, F.; Shang, Z. Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins: Struct., Funct., Genet.* **2009**, *76*, 151–163.
- (20) Gregoret, L. M.; Rader, S. D.; Fletterick, R. J.; Cohen, F. E. Hydrogen bonds involving sulfur atoms in proteins. *Proteins: Struct., Funct., Genet.* **1991**, *9*, 99–107.
- (21) Roos, G.; Loverix, S.; Geerlings, P. Origin of the pKa perturbation of N-terminal cysteine in alpha- and 3(10)-helices: a computational DFT study. *J. Phys. Chem. B* **2006**, *110*, 557–562.
- (22) Chakrabarti, P.; Pal, D. An electrophile-nucleophile interaction in metalloprotein structures. *Protein Sci.* **1997**, *6*, 851–859.
- (23) Dennington, R.; Keith, T.; Millam, J. *GaussView*, version 5; Semichem Inc.: Shawnee Mission, KS, 2009.
- (24) Lide, D. R. Structure of Free Molecules in the Gas Phase. In *CRC Handbook of Chemistry and Physics*, 93rd ed. (2012–2013); Haynes, W. M., Lide, D. R., Bruno, T. J., Eds.; CRC Press: Boca Raton, FL, 2012; pp 19–47.
- (25) Zhao, Y.; Truhlar, D. G. Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.
- (26) Walker, M.; Harvey, A. J.; Sen, A.; Dessent, C. E. Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions. *J. Phys. Chem. A* **2013**, *117*, 12590–12600.
- (27) Riffet, V.; Frisona, G.; Bouchoux, G. Acid–base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met). *Phys. Chem. Chem. Phys.* **2011**, *13*, 18561–18580.
- (28) Brás, N. F.; Perez, M. A.; Fernandes, P. A.; Silva, P. J.; Ramos, M. J. Accuracy of Density Functionals in the Prediction of Electronic Proton Affinities of Amino Acid Side Chains. *J. Chem. Theory Comput.* **2011**, *7*, 3898–3908.
- (29) Ugur, I.; Marion, A.; Parant, S.; Jensen, J. H.; Monard, G. Rationalization of the pKa values of alcohols and thiols using atomic charge descriptors and its application to the prediction of amino acid pKa's. *J. Chem. Inf. Model.* **2014**, *54*, 2200–2213.
- (30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
- (31) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 171–179.
- (32) Van Roey, P.; Kerr, K. A. The structure of 2,6-dimethylpiperidinium dithiosalicylate. X-ray diffraction study at 200 K and neutron diffraction study at 20 K. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1981**, *37*, 1679–1685.
- (33) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. Computational Thermochemistry: Scale Factor Databases and Scale Factors for Vibrational Frequencies Obtained from Electronic Model Chemistries. *J. Chem. Theory Comput.* **2010**, *6*, 2872–2887.
- (34) Ervin, K. M.; Nickel, A. A.; Lanorio, J. G.; Ghale, S. B. Anchoring the Gas-Phase Acidity Scale from Hydrogen Sulfide to Pyrrole. Experimental Bond Dissociation Energies of Nitromethane, Ethanethiol, and Cyclopentadiene. *J. Phys. Chem. A* **2015**, *119*, 7169–7179.
- (35) Alongi, K. S.; Shields, G. C.; Wheeler, R. A. Theoretical Calculations of Acid Dissociation Constants: A Review Article. *Annu. Rep. Comput. Chem.* **2010**, *6*, 113–138.
- (36) Boese, A. D.; Handy, N. C. New exchange-correlation density functionals: The role of the kinetic-energy density. *J. Chem. Phys.* **2002**, *116*, 9559–9569.
- (37) Zhao, Y.; Pu, J.; Lynch, B. J.; Truhlar, D. G. Tests of second-generation and third-generation density functionals for thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2004**, *6*, 673–676.
- (38) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **1988**, *9*, 327–335.
- (39) Lim, C.; Bashford, D.; Karplus, M. Absolute pKa Calculations with continuum dielectric methods. *J. Phys. Chem.* **1991**, *95*, 5610–5620.
- (40) Bashford, D. An object-oriented programming suite for electrostatic effects in biological molecules An experience report on the MEAD project. In *Scientific Computing in Object-Oriented Parallel Environments*; Ishikawa, Y., Oldehoeft, R., Reyniers, J. W., Tholburn, M., Eds.; Springer: Berlin, Heidelberg, 1997; Vol. 1343, pp 233–240.
- (41) Dudev, T.; Lim, C. A DFT/CDM study of metal-carboxylate interactions in metalloproteins: Factors governing the maximum number of metal-bound carboxylates. *J. Am. Chem. Soc.* **2006**, *128*, 1553–1561.
- (42) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute–Water Clusters. *J. Chem. Theory Comput.* **2005**, *1*, 1133–1152.
- (43) Vorobyov, I.; Li, L.; Allen, T. W. Assessing atomistic and coarse-grained force fields for protein-lipid interactions: The formidable challenge of an ionizable side chain in a membrane. *J. Phys. Chem. B* **2008**, *112*, 9588–9602.
- (44) Wolfenden, R. Interaction of a peptide bond with solvent water-vapor phase analysis. *Biochemistry* **1978**, *17*, 201–204.
- (45) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. Affinities of Amino Acid Side Chains for Solvent Water. *Biochemistry* **1981**, *20*, 849–855.
- (46) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (47) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (48) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **1999**, *285*, 1735–1747.

- (49) Baker, E. N.; Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179.
- (50) Rodríguez-Casado, A.; Thomas, G. J. J. Structural Roles of Subunit Cysteines in the Folding and Assembly of the DNA Packaging Machine (Portal) of Bacteriophage P22. *Biochemistry* **2003**, *42*, 3437–3445.
- (51) Biswal, H. S.; Shirhatti, P. R.; Wategaonkar, S. O-H ··· O versus O-H ··· S Hydrogen Bonding. 2. Alcohols and Thiols as Hydrogen Bond Acceptors. *J. Phys. Chem. A* **2010**, *114*, 6944–6955.
- (52) Salsbury, F. R. J.; Knutson, S. T.; Poole, L. B.; Fetrow, J. S. Functional site profiling and electrostatic analysis of cysteines modifiable to cysteine sulfenic acid. *Protein Sci.* **2008**, *17*, 299–312.
- (53) Kristof, W.; Zundel, G. Proton transfer and polarizability of hydrogen bonds formed between cysteine and lysine residues. *Biopolymers* **1982**, *21*, 25–42.
- (54) Karshikoff, A. N.; Nilsson, L.; Foloppe, N. 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. *Biochemistry* **2013**, *52*, 5730–5745.
- (55) Guddat, L. W.; Bardwell, J. C.; Glockshuber, R.; Huber-Wunderlich, M.; Zander, T.; Martin, J. L. Structural analysis of three His32 mutants of DsbA: Support for an electrostatic role of His32 in DsbA stability. *Protein Sci.* **1997**, *6*, 1893–1900.
- (56) Huber-Wunderlich, M.; Glockshuber, R. A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Folding Des.* **1998**, *3*, 161–171.
- (57) Huang, L.; Brinen, L. S.; Ellman, J. A. Crystal Structures of Reversible Ketone-Based Inhibitors of the Cysteine Protease Cruzain. *Bioorg. Med. Chem.* **2003**, *11*, 21–29.
- (58) Löser, R.; Abbenante, G.; Madala, P. K.; Halili, M.; Le, G. T.; Fairlie, D. P. Noncovalent Tripeptidyl Benzyl- and Cyclohexyl-Amine Inhibitors of the Cysteine Protease Caspase-1. *J. Med. Chem.* **2010**, *53*, 2651–2655.