n→π* Interactions Modulate the Properties of Cysteine Residues and Disulfide Bonds in Proteins

Henry R. Kilgore and Ronald T. Raines*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Supporting Information

ABSTRACT: Noncovalent interactions are ubiquitous in biology, taking on roles that include stabilizing the conformation of and assembling biomolecules, and providing an optimal environment for enzymatic catalysis. Here, we describe a noncovalent interaction that engages the sulfur atoms of cysteine residues and disulfide bonds in proteins—their donation of electron density into an antibonding orbital of proximal amide carbonyl groups. This n→π* interaction tunes the reactivity of the CXXC motif, which is the critical feature of thioredoxin and other enzymes involved in redox homeostasis. In particular, an n→π* interaction lowers the pKₐ value of the N-terminal cysteine residue of the motif, which is the nucleophile that initiates catalysis. In addition, the interplay between disulfide n→π* interactions and C5 hydrogen bonds leads to hyperstable β-strands. Finally, n→π* interactions stabilize vicinal disulfide bonds, which are naturally diverse in function. These previously unappreciated n→π* interactions are strong and underlie the ability of cysteine residues and disulfide bonds to engage in the structure and function of proteins.

INTRODUCTION

The cysteine residues of proteins have unique attributes. Their thiolhydryl groups not only manifest potent nucleophilicity, but also undergo a facile oxidation reaction to generate disulfide bonds. The descendant cystines are active components of catalytic, oxidation–reduction, and signal transduction pathways, and have distinct physicochemical properties.

Approximately 20% of human proteins are predicted to contain a disulfide bond. Although prevalent, the two sulfur atoms of disulfide bonds are not known to engage with other functional groups in proteins. The unique attributes of disulfide bonds and their component sulfur atoms enticed us to consider their electronic structure in detail.

In a disulfide bond, one lone pair of each sulfur atom resides in a nondegenerate s-type orbital (nₛ; Figure 1A), and the other resides in a nondegenerate p-type orbital (nₚ; Figure 1B). We envisioned that these four lone pairs could interact with nearby carbonyl groups. In particular, donation of lone-pair electron density into the π* orbital of an adjacent carbonyl group could lead to an n→π* interaction (Figure 1C and D). The shape and higher energy of nₛ orbitals confers larger contributions relative to those of nₚ orbitals. The existence of such an interaction would underlie an aspect of disulfide bonds that is now unappreciated.

Herein, we use computational methods and bioinformatic analyses to provide evidence that n→π* interactions that originate from sulfur play important roles in the structure and function of proteins. The effects arise from the tuning of the thermodynamic stability of the disulfide bonds, thios, and thiolates of cysteine residues. We find these effects to be especially important in the reactivity of the CXXC motifs in enzymic active sites, interplay with the C5-hydrogen bonds of β-strands, and polarization of electron density in vicinal disulfide bonds.

RESULTS AND DISCUSSION

Protein structures are stabilized by a web of interplaying noncovalent interactions. This web overpowers entropy only barely, as the free energy difference between the folded and unfolded states is merely 5–15 kcal/mol. We examined three
interaction of $S_i$ dimensional structure. The results revealed a chain of interactions that stabilize the oxidized state of the CXXC motif. Moreover, $S_i$ cysteine residue in a CXXC motif typically has a $pK_a$ value of 8.7. In contrast, the N-terminal cysteine residue in a CXXC motif typically has a $pK_a$ value below physiological pH and is thus highly nucleophilic. The origin of this anomalous acidity has been unclear, despite extensive investigation.

CXXC motifs often reside at the N-terminus of an $\alpha$-helix. In that context, the sulfur atom ($S_i$) of only the N-terminal cysteine residue is exposed to solvent. Solvent-accessible surface area calculations on the crystal structures of oxidized and reduced states of thioredoxin and thioredoxin-2 show that the C-terminal cysteine is completely inaccessible regardless of redox state (Figure S1 of the Supporting Information, SI). Moreover, $S_i$ of the N-terminal cysteine residue experiences an increase of $\sim$6-fold in solvent-accessible surface area upon reduction of the active-site disulfide bond. Accordingly, we focused our attention on $S_i$, which is the linchpin of the CXXC motif.

We began by performing Natural Bond Orbital (NBO) second-order perturbation theory calculations on 7 different proteins with an oxidized CXXC motif and a known three-dimensional structure. The results revealed a chain of $n\rightarrow\pi^*$ interactions that stabilize the oxidized state of the CXXC motif (Figure 2A, Table S1). Foremost in this network is the interaction of $S_i$ and the $C_i=O_i$ carbonyl group. Specifically, lone-pair electron density is donated from this sulfur atom into the $\pi^*$ orbital of the carbonyl group, generating a strong $n\rightarrow\pi^*$ interaction in the oxidized, thiol, and thiolate states (Figures 2B–D; Table S2). The chain is propagated by the formation of a $C_i=O_i\cdots O_{i+1} n\rightarrow\pi^*$ interaction (Figure 2E; Tables S1 and S2), and then a $C_{i+1}=O_{i+1}\cdots O_{i+2} n\rightarrow\pi^*$ interaction (Figure 2F; Tables S1 and S2). This chain of $n\rightarrow\pi^*$ interactions was apparent in all 7 proteins examined and appears to be a ubiquitous feature of CXXC motifs.

Next, we examined oxidized CXXC motifs with known crystal structures and reduction potentials. We found that stronger $n\rightarrow\pi^*$ interactions correlate with lower reduction potentials, that is, more stable disulfide bonds (Figure 3). The effect here is not major, given that $100\text{ mV}$ corresponds to $2.3\text{ kcal/mol}$. Nonetheless, the electron-donation that arises from disulfide $n\rightarrow\pi^*$ interactions is likely to increase the electrophilicity of a disulfide bond and thereby enhance its reactivity in thiol–disulfide interchange reactions.

To understand how the chain of $n\rightarrow\pi^*$ interactions within CXXC motifs might be leveraged to perform biochemical functions, we examined well-characterized thioredoxins in more detail. In a CXXC motif, $S_i$ has three relevant states: disulfide, thiol, and thiolate. Conversion between these states does not induce substantial conformational changes (Figures S1 and S2). The major change incurred upon reduction of the disulfide bond is in the $\chi_1$ dihedral angle (that is, $N_C\cdots C=P\cdots S_i$), which rotates toward the solvent (Figure S2). In the descendant thiol and thiolate, $S_i$ forms a hydrogen bond with water rather than with $S_i\cdots \cdots H$ or another enzymic functional group. Inspection of both of these three states reveals that all are stabilized by a $S_i\cdots C_i=O_i n\rightarrow\pi^*$ interaction (Figure 4; Tables S1 and S2).

Figure 2. Network of $n\rightarrow\pi^*$ interactions within the CXXC motif. (A) Electron donation in the oxidized state. (B) $S_i\cdots C_i=O_i n\rightarrow\pi^*$ interaction in the oxidized state. (C) $S_{i+1}\cdots C_{i+2}=O_i n\rightarrow\pi^*$ interaction in the thiol state. (D) $S_{i+1}\cdots C_{i+2}=O_i n\rightarrow\pi^*$ interaction in the thiolate state. (E) $C_i=O_i\cdots C_{i+1}=O_{i+1} n\rightarrow\pi^*$ interaction in the thiolate state. (F) $C_{i+1}=O_{i+1}\cdots C_{i+2}=O_{i+2} n\rightarrow\pi^*$ interaction in the thiolate state. Structures are from PDB entries 1ert and 1eru.11

Figure 3. Graph of the relationship between calculated $E_{n\rightarrow\pi^*}$ values and measured $E^{\circ}$ values for CXXC motifs: *Escherichia coli* DsbA (black; PDB entry 1a2) and three variants of *Staphylococcus aureus* thioredoxin (blue; PDB entries 2o7k, 2o8s, and 2o8t).
Moreover, the Siγ···CπO in → π* interaction tends to be stronger than the CπO···Ci+1πOi+1 or CπOi+1···Ci+2πOi+2 interaction. A critical step in catalysis by thioredoxin is deprotonation of Siγ to form the nucleophilic thiolate.12a We find that the Siγ···CπO in → π* interaction in the thiolate state is much greater than that in the thiol state (Figure 4). This difference is likely to make a significant contribution to the diminished pKₐ of the N-terminal cysteine residue in CXXC motifs. The extant explanation for this low thiol pKₐ value relies on a presumed macrodipole of the α-helix.17 The dipole of an α-helix18 has not been well-replicated in model systems.19 Moreover, slightly downstream to many CXXC motifs is a proline residue, which induces a kink in the α-helix.20 Such a kink would interrupt the projection of the electric field along the helical axis. Notably, calculations of this thiol pKₐ have yielded values that are much greater than those observed by experiment,14,17,21 consistent with n→π* interactions being absent from the Hamiltonians employed in typical calculations.

Interplay of Disulfide n→π* Interactions with C5 Hydrogen Bonds. A C5 hydrogen bond is an intrinsic feature of β-strands, arising from the overlap of an nπ*-type carbonyl lone pair with the σ* orbital of an adjacent amide N–H bond (Figure 5A).22 Because a large fraction of disulfide bonds in β-strands participate in highly stabilizing n→π* interactions, we sought to examine the interplay between a C5 hydrogen bond and a disulfide n→π* interaction (Figure 5B). To do so, we examined a disulfide bond that originates from a β-strand (Figure 5C).

A disulfide n→π* interaction from Siγ into a carbonyl group polarizes the electron density of the carbonyl group toward its oxygen (Figure 5B). The ensuing increase in electron density could result in a stronger C5 hydrogen bond. We performed

![Figure 4](image1.png)

**Figure 4.** Graph showing calculated En→π* values (in kcal/mol) within the CXXC motifs of *Homo sapiens* thioredoxin and thioredoxin-2, and *Drosophila melanogaster* thioredoxin. Data are listed in Table S2.

![Figure 5](image2.png)

**Figure 5.** Interplay between a disulfide n→π* interaction and C5 hydrogen bond in a β-strand. (A) Natural bond orbitals showing a disulfide n→π* interaction. (B) Network of natural bond orbitals in which the n→π* interaction from panel A enhances an n→σ* interaction (that is, a C5 hydrogen bond) within the half-cystine residue. (C) Image of a model disulfide bond. (D) Scan of the dihedral angle ξ (which is defined in the inset of panel E) in the presence of a disulfide n→π* interaction of En→π* = 1.65 kcal/mol; data are listed in Table S3. (E) Scan of the dihedral angle ξ in the absence of a disulfide n→π* interaction; data are listed in Table S4. The structure in panels A–C is from PDB entry 4gn2 (Table S5) and was used in the calculations of panels D and E.
relaxed scan calculations of the dihedral angle $\xi$ (that is, $H_1^\alpha \cdots C_n^\alpha \cdots N_\beta \cdots H_2$). Each step of these calculations was then subjected to NBO calculations to deconvolute the stabilizing interactions.\textsuperscript{23} Specifically, donation of $n_\sigma$ electron density leads to $\Delta E_{n_\sigma \pi^*} = 0.30$ kcal/mol at the maximum, an increase of 42% over that in the absence of a $n \rightarrow \pi^*$ interaction. Moreover, the maximal $E_{n_\sigma \pi^*}$ is achieved at a dihedral angle $\xi$ that is lower by 15°. In essence, the disulfide $n \rightarrow \pi^*$ interaction increases the polarization of the acceptor carbonyl group, resulting in an increase in the energy of an associated $C_5\pi$ hydrogen bond. This interplay between disulfide and $C_5\pi$ hydrogen bonds bears resemblance to systems in which hydrogen bonds can provide an electron density that the disulfide bond, thereby creating an electropositive and hydrophobic surface (Figure S3), especially in the trans-up/down conformations (Figure S3A,B).

**CONCLUSIONS**

The role of $n \rightarrow \pi^*$ interactions in protein structure and function became apparent in the early 2000s.\textsuperscript{26} Our data expose new terrain in this landscape: the mixing of sulfur $n_\sigma$ and $n_\pi$ orbitals with proximal carbonyl groups can provide an exceptionally strong $n \rightarrow \pi^*$ interaction that enhances the stability of host secondary structures. In general, the stabilization of oxidized, thiol, or thiolate states through interactions extends to protein function. In particular, the thermodynamic stability of the CXXC motif, which is the centerpiece of redox homeostasis, is underpinned by $n \rightarrow \pi^*$ interactions. Finally, we note that the enhanced ability of selenium to donate an $n \rightarrow \pi^*$ interaction\textsuperscript{29} suggests that the effects that we observe with cysteine residues could be amplified with selenocysteine.

**EXPERIMENTAL METHODS**

Calculations. All quantum mechanical calculations were performed with Gaussian 09, revision E.01\textsuperscript{31} at the M062x/6-311+g(2df,p) level of theory. Energies (i.e., $E_{n_\sigma \pi^*}$ and $E_{n \pi^*}$) were calculated by second-order perturbation theory analysis of optimized structures as implemented with NBO 6.0\textsuperscript{12} in Gaussian 09, revision.
The atomic coordinates of CXXC motifs were extracted from the PDB files of parent enzymes. The Cα atoms (and thus the side chains) were fixed while other main-chain atoms were allowed to optimize. Optimized structures were consistent with those from molecular dynamics and QM/MM calculations. One-dimensional scan calculations were performed by increasing the dihedral angle θ in 1°-steps and allowing the structure to optimize.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b09701.

Tables S1–S5, S6–S18 (atomic coordinates of CXXC motifs), S19–S24 (atomic coordinates of vicinal disulfide bonds), and Figures S1–S3 (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

*trraines@mit.edu*

**ORCID**

Ronald T. Raines: 0000-0001-7164-1719

**Funding**

This work was supported by Grant R01 GM044783 (NIH). Calculations made use of the Molecular Graphics and Computational Facility at the University of California, Berkeley, which was supported by Grant S10 OD023532 (NIH).

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank Dr. Emily R. Garnett for pointing us toward the disulfide bonds of human chorionic gonadotropin.

**REFERENCES**


active site proline determines the reducing power of *Staphylococcus aureus* thioredoxin. *J. Mol. Biol.* 2007, 368, 800–811.


