Published in final edited form as:

J Mol Biol. 2010 November 5; 403(4): 495–504. doi:10.1016/j.jmb.2010.09.004.

Propensities of Aromatic Amino Acids versus Leucine and Proline to Induce Residual Structure in the Denatured State Ensemble of Iso-1-cytochrome *c*

Michaela L. Finnegan and Bruce E. Bowler*

Department of Chemistry and Biochemistry, Biochemistry Program and Center for Biomolecular Structure & Dynamics, University of Montana, Missoula, Montana, 59812, USA

Abstract

Histidine-heme loop formation in the denatured state of a protein is a sensitive means to probe for residual structure under unfolding conditions. In this study, we use a host-guest approach to investigate the relative tendencies of different amino acids to promote residual structure under denaturing conditions. The host for this work is a 6 amino acid insert of five alanines followed by a lysine engineered immediately following a unique histidine near the N-terminus of yeast iso-1-cytochrome *c*. We substitute the 4th alanine in this sequence, HAAAXAK, with X = Trp, Phe, Tyr and Leu. The effects of proline are tested with substitutions at positions 1 and 5 in the insert, HPAAAAK and HAAAAPK, respectively. Thermodynamic studies on His-heme loop formation in 3 M guanidine hydrochloride reveal significant stabilization of residual structure by aromatic amino acids, particularly, Trp and Phe, and minimal stabilization of residual structure by Leu. Prolines disfavor His-heme loop formation slightly, presumably due to enhanced chain stiffness. Kinetic studies reveal that much of the change in His-heme loop stability for the aromatic amino acids is caused by a slowing of the rate of His-heme loop breakage, indicating that residual structure is preferentially stabilized in the closed-loop form of the denatured state.

Keywords

Denatured States; Protein Folding; Loop Formation; Residual Structure; Chain Stiffness

There is strong evidence that non-random structure persists in disordered and denatured proteins.^{1,2} The Non-random structure is often local in nature,³ with persistent secondary structure particularly common.^{4,5} However, recent work, especially with the paramagnetic relaxation enhancement^{2,6-8} and transverse relaxation NMR methods,⁹⁻¹² has shown that non-random structure can be stabilized by long-range tertiary interactions. Residual structure is most prominent in the absence of denaturing agents^{13,14} and under weaker denaturing conditions.^{5,6} In many instances, though, significant residual structure persists in the denatured state in the presence of 3 M guanidine hydrochloride (gdnHCl)^{5,6,15} or even in 8 M urea.^{3,11}

There has been significant progress in quantitative evaluation of the free energy associated with electrostatic interactions in the denatured state ensemble (DSE).¹⁶⁻²² Electrostatic

* To whom correspondence should be addressed. Phone (406) 243-6114. Fax (406) 243-4227. bruce.bowler@umontana.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

interactions have been shown to stabilize the DSE by 1 to 4 kcal/mol. Thus, electrostatic effects in the DSE can strongly influence the overall stability of a protein. Mutational studies on surface-exposed sites for several proteins show reverse hydrophobic effects indicating that hydrophobic interactions are important in the DSE of proteins, as well.²³⁻²⁸ Mutation of charged residues to hydrophobic residues has been shown to decrease the heat capacity change, ΔC_p , for thermal unfolding consistent with formation of hydrophobic clusters in the DSE.^{29,30} Simulation and experiment on the drkN SH3 domain, barnase and hen egg white lysozyme have led to the conclusion that aromatic residues, in particular, promote hydrophobic clusters in the DSE.^{10,11,14,31-37} However, simulation and experiment on chymotrypsin inhibitor 2 and α -lactalbumin show that larger aliphatic side chains are also sufficient to nucleate hydrophobic clusters in the DSE.^{9,38} The various hydrophobicity scales rank the relative hydrophobicities of aliphatics and aromatics differently.³⁹⁻⁴² Thus, a quantitative assessment of the relative tendencies of these two types of amino acid side chains to promote residual structure under a particular set of denaturing conditions would be useful. Similarly, a comparison of the relative impact of electrostatic versus hydrophobic interactions on the energetics of the DSE would be valuable.

To quantify the relative free energies of hydrophobic residual structure interactions induced by aromatic and aliphatic residues in the DSE, we apply our denatured state histidine-heme loop formation assay.^{23,43-45} In particular, we substitute single aromatic or aliphatic residues into a 22-residue loop formed in the denatured state of yeast iso-1-cytochrome *c* (iso-1-Cytc) and measure the effect on loop stability. For comparison, we have also measured the effect of insertion of a single proline into the loop to evaluate the relative importance of backbone versus side-chain effects on the DSE.

Design of variants

We use a host-guest approach to evaluate the effects of proline, the aromatic amino acids, and the large aliphatic amino acid, leucine, on His-heme loop formation in the DSE of iso-1-Cytc. We have used leucine to represent large aliphatic residues because it is not β -substituted, like valine and isoleucine, and thus its effect on main chain sterics will be more similar to the aromatic side chains.⁴⁶ The host sequence we have used to examine the effects of these amino acids on denatured state His-heme loop formation is an AAAA⁴⁷AK insert between His(-2) and Ala(-1) of iso-1-Cytc⁴⁷ (Fig. 1; horse numbering is used, thus the 5 amino acid N-terminal extension of iso-1-Cytc is numbered -5 to -1). The placement of the insert was chosen so as not to disrupt the N-terminal helix and thereby maintain the integrity of the protein as a whole. Since iso-1-Cytc is a *c*-type cytochrome (heme is covalently attached to the polypeptide chain through a CXXCH heme attachment motif, Cys14, Cys17 and His18 in iso-1-Cytc, see Figs. 1 and 2), the polypeptide chain remains attached to the heme when the protein is unfolded by denaturants such as gdnHCl. Thus, for the iso-1-Cytc variant containing the AAAA⁴⁷AK host sequence (NH5A variant), the stability of a 22-residues His-heme loop is readily measured under denaturing conditions via a simply pH titration, yielding an apparent pK_a , $pK_a(\text{obs})$ (Fig. 2).

To compare the relative abilities of aliphatic and aromatic residues to stabilize residual structure in the DSE, we replace the 4th alanine in the insert (Fig. 1). The inserted residue replaces an alanine and can be viewed as the addition of an aromatic group or two methyl groups, as in the case of leucine. The inserted residue is also in a relatively neutral environment, surrounded by alanines on either side. Furthermore, the guest amino acid is 4 residues from His(-2), thus, it will not be held immediately adjacent to the heme when the His(-2)-heme loop forms. These design features of the host-guest approach should provide for a good general estimate of the relative tendency of aromatic and aliphatic amino acids to promote residual structure compared to an alanine at the same sequence position. When the

22-residue His(-2)-heme loop forms under denaturing conditions, the volume within which the aliphatic or aromatic residue guest is constrained will decrease significantly. Therefore, we expect an increased potential for hydrophobic residual structure upon loop formation due to a mass action effect.⁴⁵

In the case of the proline residues, we have replaced the 1st and 5th alanines in the insert. The interest in proline next to the histidine stems from our previous studies on the effects of glycine on His-heme loop formation which showed that the backbone flexibility of the residue next to the histidine which forms the loop has a significant effect on the persistence (breakage rate) of the loop. Glycine next to the loop forming histidine significantly increases the rate of loop breakage.⁴⁸ Thus, we would predict that a proline in the same position would slow loop breakage. The proline at the 5th position is meant to test the effect of proline on the overall stiffness of the loop. We have also prepared a variant with both a Trp in place of the 4th alanine and a proline in place of the 5th alanine. This variant will examine how chain stiffness and hydrophobicity interact in modulating the propensity to form primitive contacts early in folding. The complete set of variants used in this study is outlined in Fig. 2 along with the abbreviation used for each variant.

Theoretical treatment of loop formation

When a His-heme loop forms in the denatured state of a *c*-type cytochrome such as iso-1-Cytc, a proton is released (Fig. 2). Thus, an apparent pK_a , $pK_a(\text{obs})$, can be extracted from the midpoint of the pH-dependent loop formation equilibrium. Since this equilibrium is essentially a competition between the heme and a proton for binding to the histidine, a lower $pK_a(\text{obs})$ indicates a more stable His-heme loop. This equilibrium can be treated as a stepwise process involving deprotonation of histidine followed by loop formation (Eq 1)

$$pK_a(\text{obs}) = pK_a(\text{HisH}^+) + pK_{\text{loop}}(\text{His}) \quad (1)$$

where $pK_a(\text{HisH}^+)$ is the acid ionization constant of His(-2) in this case and $pK_{\text{loop}}(\text{His})$ reflects the equilibrium constant for binding of a deprotonated histidine to the heme to form a loop in the denatured state.

For a random coil polypeptide, the Jacobson-Stockmayer equation⁴⁹ is used to predict the probability of loop formation (Eq 2). In Eq 2, v_3 is the scaling exponent for loop formation in a 3-dimensional polymer, R is the gas constant, n is the number of residues in the loop, C_n is

$$\Delta S_{\text{loop}} = -v_3 R \ln(n) + R \ln \left[\left(\frac{3}{2\pi C_n l^2} \right)^{v_3} V_i \right] \quad (2)$$

Flory's characteristic ratio for two monomers forming a loop including n residues and V_i is the volume within which the two monomers must be constrained for a loop to form. Since all His-heme loops have the same length in this study ($n = 22$), we can write Eq 3 for the difference in the $pK_a(\text{obs})$, $\Delta pK_a(\text{obs})$, for the host sequence and a variant containing a guest residue within the

$$\Delta pK_a(\text{obs}) = pK_a(\text{obs})_{\text{guest}} - pK_a(\text{obs})_{\text{host}} = \nu_3 \log \left(\frac{C_{n,\text{guest}}}{C_{n,\text{host}}} \right) \quad (3)$$

AAAAAK host sequence. To a first approximation, Flory's characteristic ratio, which reflects chain stiffness, will be similar to alanine for all amino acids except glycine, proline and the β -substituted amino acids. Thus, for the W4, F4, Y4 and L4 variants, changes in $pK_a(\text{obs})$ relative to the NH5A variant can be attributed to residual structure induced by loop formation rather than changes in C_n . Proline severely restricts backbone flexibility relative to alanine,⁵⁰ thus, a change in C_n is expected to dominate for the P1 and P5 variants.

Stability of variants

Variants were prepared using the pBTR1 vector^{51,52} carrying the NH5A variant as template for standard PCR-based mutagenesis. Proteins were expressed in BL21 DE3 *Escherichia coli* cells, as previously described.^{47,48,52} These variants all have a free N-terminal amino group which can bind to the heme under denaturing conditions.⁵³ However, histidines on the N-terminal side of the site of heme attachment have high enough affinity for the heme^{47,48,54} that binding of the N-terminal amino group to the heme does not interfere with His-heme binding in the denatured state.

As a prerequisite to measurements of His-heme loop formation in the denatured state, we measured the global stability of each variant using gdnHCl denaturation monitored by circular dichroism spectroscopy at 222 nm (Fig. 3).^{44,55} Table 1 compiles the stability data for all variants. The denaturation midpoints, C_m , vary between 0.4 and 0.5 M gdnHCl. Given the instability of these variants and the curvature in the native baseline (Fig. 3), the errors in evaluating the gdnHCl m -value are larger than usual (up to 0.6 kcal mol⁻¹ M⁻¹, Table 1). As a result, the $\Delta G_u^0(\text{H}_2\text{O})$ values for all variants are mostly the same within error. Stability data for the NK5A (AAAAAK insert with the native Lys instead of His at position -2) is included for comparison with the variants containing the unique histidine at position -2. The differences in stability, $\Delta G_u^0(\text{H}_2\text{O}) \sim 4$ kcal mol⁻¹ for the NK5A variant versus $\Delta G_u^0(\text{H}_2\text{O}) \sim 2$ kcal mol⁻¹ for the variants with His at position -2, indicates that the formation of the 22-residue His-heme loop stabilizes the denatured state relative to the native state.⁴⁷ The low C_m values for all His(-2)-containing variants demonstrate that all are fully unfolded in 3 M gdnHCl (see also Fig. 3), the conditions used for studies on denatured state His-heme loop formation.

Equilibrium His-heme loop formation

His-heme loop stabilities for each variant were measured under denaturing conditions (3 M gdnHCl) by increasing proton concentration in steps of ~ 0.2 pH units from pH 7 to pH 2. Within this range the unique histidine, a strong field ligand, is titrated off the heme-group and is replaced by water, a weak field ligand. This change is monitored at 398 nm where the greatest change in absorbance occurs as the Fe⁺³-heme Soret band shifts due to the spin state change of the heme (Fig. S1 in Supporting Information). Figure 4 shows the titration curves for three of the variants, NH5A, P1 and F4. The titration curve for the P1 variant (pink) is shifted to slightly higher pH relative to the NH5A variant (black), whereas the titration curve for the F4 variant (green) is shifted to lower pH relative to the NH5A variant. Plots of absorbance data from pH 7 to pH 3 show a single isosbestic point, indicating that equilibrium loop formation in 3 M gdnHCl is a two-state process (Fig. S1 in Supporting Information). The $pK_a(\text{obs})$ values for all variants assuming a two-state equilibrium are collected in Table 2. We note for all variants that the number of protons released upon loop formation, n_p , is near 1 (Table 2) as expected from the reaction scheme in Fig. 2.

For both variants with a single proline guest residue in the AAAAAK host, $pK_a(\text{obs})$ increases slightly relative to the NH5A variant. This increase in the $pK_a(\text{obs})$ corresponds to disfavoring loop formation ($\Delta\Delta G_{\text{loop}}(\text{His})$ in Table 2) by about 0.15 kcal mol⁻¹ for the P1 and P5 variants relative to the NH5A variant. This observation is consistent with the expected effects of proline on chain stiffness. Use of Eq 3, with $v_3 = 2$,⁵⁶ indicates that Flory's characteristic ratio increases by 10-15% when a single proline is introduced into a 22 residue loop.

All of the variants containing aromatic guest residues (W4, W4P5, F4, Y4) show a significant decrease in $pK_a(\text{obs})$ indicating that loop formation is favored by substitution with an aromatic amino acid. In 3 M gdnHCl, the observed decreases in $pK_a(\text{obs})$ correspond to a stabilization of the His-heme loop by 0.35 to 0.5 kcal mol⁻¹ relative to the NH5A variant ($\Delta\Delta G_{\text{loop}}(\text{His})$ in Table 2). The stabilizing effect is largest for Phe and least for Tyr, with Trp in between in its effect. By contrast, the $pK_a(\text{obs})$ for the L4 variant indicates that, within error, replacing a single alanine with leucine has no effect on the stability of the 22-residue His-heme loop in 3 M gdnHCl.

We have also prepared the W4P5 variant, which has both aromatic and proline guest residues, to test how these two types of substitutions interact to affect the conformational properties of the DSE. Within error the $pK_a(\text{obs})$ and $\Delta\Delta G_{\text{loop}}(\text{His})$ values are the same for the W4 and the W4P5 variants in 3 M gdnHCl.

Kinetics of His-heme loop formation

In previous work, we have shown that the kinetics of His-heme loop formation in the denatured state is consistent with a mechanism in which the histidine undergoes a rapid deprotonation equilibrium followed by binding of the deprotonated histidine to the heme.⁵⁷ For this kinetic mechanism, the pH dependence of the observed rate constant for loop formation and breakage, k_{obs} is given by Eq 4, where k_b is the rate constant for loop breakage, k_f is the rate

$$k_{\text{obs}} = k_b + k_f (K_a(\text{HisH}^+) / ([\text{H}^+] + K_a(\text{HisH}^+))) \quad (4)$$

constant for loop formation and $K_a(\text{HisH}^+)$ is the acid dissociation constant for the histidine which forms the loop. Since $K_a(\text{HisH}^+) = 6.6 \pm 0.1$ for iso-1-CytC irrespective of sequence position or gdnHCl concentration,⁴⁵ k_b will equal k_{obs} for pH jump mixing from pH values where the His-heme loop is formed to pH values in the range 3 to 4, where the loop is broken (Table S1 and Fig. S2 in Supporting Information). In Table 3, we report the k_b values obtained from downward pH jumps to pH 3.4. Since loop formation occurs within the dead time of our stopped-flow instrument, we take advantage of the observation that equilibrium loop formation is two-state to obtain k_f . The k_f values obtained from k_b and $K_{\text{loop}}(\text{His})$ (derived using Eq 1 and the data in Table 2) are reported in Table 3.

The rate constants for loop formation for the variants containing Pro are significantly smaller than for the NH5A variant. This result indicates that the additional stiffness conferred by proline significantly impacts the conformational properties of the main chain, causing the root mean square end-to-end distance to be longer and thus, slowing loop formation. Our equilibrium measurements indicated that C_n increases by 10-15% for the P1 and P5 variants relative to the NH5A variant, consistent with the slower rates of loop formation for the P1 and P5 variants. For the P5 variant, k_b is identical to k_b for the NH5A variant. Thus, for prolines away from loop ends, the effects of chain stiffness on loop stability are due entirely to a slowing of the rate of loop formation. The decrease in k_f is

most pronounced when proline is immediately adjacent to the histidine that forms the loop (P1 versus P5 in Table 3). Thus, the effects of proline on backbone stiffness are most prominent for residues near the proline. Similarly, substitution of a single proline in the center of a set of polyserine peptides slows rates of end-to-end loop formation most for the shortest peptides,⁵⁸ also consistent with the effect of proline being most significant on nearby residues.

By contrast, the effect of the guest residue on k_f is modest within error for variants with aromatic and aliphatic guests. However, it is evident in Table 3, that k_b is significantly smaller for breakage of the 22-residue His-heme loop when the loop contains an aromatic amino acid. This observation suggests that compared to alanine, the primary effect of aromatic residues is to stabilize the closed loop form relative to the transition state for loop breakage. When leucine is introduced into the loop there is also a decrease in k_b , although it is less than for the aromatic amino acids.

Using the k_b for the host sequence (NH5A variant) as a reference point, the relative increase in the activation free energy for loop breakage, $\Delta\Delta G_b^\ddagger$ (Table 3), is largest for the aromatic guest residues. We attribute the positive values of $\Delta\Delta G_b^\ddagger$ to stabilization of the closed loop form caused by residual structure induced by the aromatic side chains upon loop formation. For the F4 variant, the residual structure induced by Phe relative to Ala in the closed loop form is >0.4 kcal/mol (Table 3), accounting for essentially all of the increase in the equilibrium stability of the loop (compare to $\Delta\Delta G_{\text{loop}}(\text{His})$ in Table 2). Given that our measurements were made in 3 M gdnHCl, the stabilization due to hydrophobic residual structure is substantial. We cannot rule out the possibility that the aromatic guest residues destabilize residual structure in both the open loop form and the transition state relative to the closed loop form, however, this explanation seems less likely.

When proline is introduced into the 22-residue loop there is no effect on k_b when the proline is five residues from the histidine that forms the loop (P5 variant). On the other hand, when the proline is next to the histidine that forms the loop (P1 variant), a significant slowing in the rate of loop breakage is observed (compare also $\Delta\Delta G_b^\ddagger$ in Table 3). This observation is analogous to the results of our previous work on glycine substitutions in the AAAA^{AK} insert of the NH5A variant. When a sterically flexible glycine is next to the histidine that forms the loop, k_b increases by about 20%.⁴⁸ Proline clearly has the opposite effect. Thus, the flexibility of the residue next to the loop forming histidine has a significant effect on His-heme loop persistence in the denatured state. We note that $\Delta\Delta G_b^\ddagger$ is smaller for the W4P5 variant relative to the W4 variant. While a small effect, the data suggest that having Pro next to Trp interferes with optimal stabilization of residual structure by Trp. Again, local chain sterics appear to have an important effect on the behavior of the DSE.

Importance of aromatic residues for residual structure in the DSE

We have used a host-guest approach to compare the tendency of aromatic versus aliphatic residues to promote hydrophobic residual structure in the DSE. Both kinetic and thermodynamic data ($\Delta\Delta G_{\text{loop}}(\text{His})$ in Table 2 and $\Delta\Delta G_b^\ddagger$ in Table 3) show that leucine has a weaker tendency to induce residual structure in 3 M gdnHCl compared to the aromatic amino acids. Thus, our data indicate that aromatics are more effective than aliphatics at inducing residual structure in the DSE. Given that the transfer free energies of aromatics from water into gdnHCl solution are larger and more sensitive to gdnHCl concentration than for leucine,^{15,42} the relative effectiveness of aromatics at inducing residual structure in the DSE compared to aliphatics may be somewhat underestimated by our $pK_a(\text{obs})$ measurements in 3 M gdnHCl.

The absolute magnitudes of $\Delta\Delta G_{\text{loop}}(\text{His})$ and $\Delta\Delta G_{\text{b}}^{\ddagger}$ are comparable indicating that for the aromatic guest residues most of the stabilization is due to residual structure in the closed loop form. It is interesting to note that for the aromatic guests, $\Delta\Delta G_{\text{b}}^{\ddagger}$ is uniformly smaller in absolute magnitude than $\Delta\Delta G_{\text{loop}}(\text{His})$. This observation suggests that the aromatic guests also stabilize the transition state (TS) for loop formation relative to the open loop form. However, the activation free energy for loop formation, $\Delta\Delta G_{\text{f}}^{\ddagger} (= \Delta\Delta G_{\text{loop}}(\text{His}) + \Delta\Delta G_{\text{b}}^{\ddagger})$, is only significant relative to the error for the Y4 ($\Delta\Delta G_{\text{f}}^{\ddagger} = -0.14 \pm 0.06 \text{ kcal mol}^{-1}$) and F4 ($\Delta\Delta G_{\text{f}}^{\ddagger} = -0.10 \pm 0.06 \text{ kcal mol}^{-1}$) variants. Thus, it appears that some of the residual structure that stabilizes the His-heme loops may form in the TS, but that most forms after the TS for loop formation.

Considerable research has demonstrated that aromatic residues can participate in a broader array of interactions in the native state structures of proteins than aliphatic amino acids.⁵⁹ The electronic quadrupole moment of aromatic rings allows them to engage in weakly polar interactions that are considerably stronger than the London forces that predominate when aliphatic residues pack together in the core of a folded protein. The intrinsic difference between aliphatic and aromatic residues is evident from studies on the thermodynamics of hydration of small molecule analogs of aliphatic versus aromatic amino acids.⁶⁰ While hydration of aromatics is favorable and that of aliphatics is unfavorable, the two classes have similar “hydrophobicities” based on liquid to water transfer free energies because interactions between aromatic molecules are stronger than interactions between aliphatic molecules in the liquid state. Weakly polar interactions between aromatic side chains provide stabilization on the order of 1 to 2.5 kcal/mol,^{61,62} and are known to affect the distribution of angles between aromatic rings in proteins.^{62,63} Introduction of an aromatic-aromatic interaction into xylanase, for example, has been shown to significantly stabilize the protein.⁶⁴ The cation- π interaction between histidine and aromatic residues provides another illustration of the importance of the weakly polar interactions available to aromatics in peptides and proteins. Studies on alanine-based peptides⁶⁵ show that HisH^+ interacts with Phe but not with a cyclohexane amino acid to stabilize α -helix structure. The cation- π interaction between His18 and Trp94 of barnase imparts a stabilization of 0.8 to 1 kcal/mol to the protein.⁶⁶ This stabilization is lost when Trp94 is mutated to Leu. A broad survey of side chain interactions in crystallographically characterized proteins shows that, except for Glu and Asp, most amino acid side chains show preferential association with aromatic residues.^{62,67} In contrast, aliphatic residues are only preferentially associated with other aliphatic residues. Recent work on intrinsically-disordered proteins has also shown that the aromatic amino acids are the most structure forming residues in proteins.⁶⁸ The greater structure-forming propensity of aromatic amino acids over aliphatic amino acids likely arises from the broad array of weakly polar interactions available to aromatic residues, while aliphatic residues can only participate in weaker London force interactions. Our present work shows that the greater ability of aromatic amino acids to stabilize structure relative to aliphatic amino acids persists in the DSE, even in 3 M gdnHCl, providing stabilization that is less than in folded proteins but of the same order of magnitude. Thus, our thermodynamics results explain why aromatic residues are prevalent in hydrophobic clusters in the DSE.^{10,11,14,31-37}

Relative importance of hydrophobic versus electrostatic residual structure in the DSE

Electrostatic interactions have been shown to stabilize the DSE by 1 to 4 kcal mol⁻¹.¹⁶⁻²¹ The stabilization of a His-heme loop by 0.3 to 0.5 kcal mol⁻¹ that we observe here for an aromatic guest residue in a 22-residue loop, while significant, is modest by comparison. However, electrostatic interactions in the DSE likely involve multiple residues, whereas in our case we are measuring the impact of introducing only a single aromatic residue relative

to alanine. Presumably introduction of multiple aromatic residues into an alanine background would lead to larger stabilization of the DSE.

Importance of the present work for protein folding

The present work adds to a growing body of evidence that the DSE is not a thermodynamically featureless part of the energy landscape for protein folding.^{16,19} More importantly, it shows that “hydrophobic” amino acids, particularly aromatics can contribute significantly to stabilizing interactions that bias the DSE. In terms of the mechanism of protein folding, the role of small loops is often emphasized.^{69,70} However, our work on denatured state His-heme loop formation suggests that long-range interactions can be important early in folding. In earlier work on loop formation in the denatured state of iso-1-Cytc, we observed that the gdnHCl *m*-value for a 37 residue loop formed by a histidine at position 54 was $0.46 \pm 0.06 \text{ kcal mol}^{-1} \text{ M}^{-1}$. The gdnHCl *m*-value decreased progressively to $0.15 \pm 0.05 \text{ kcal mol}^{-1} \text{ M}^{-1}$ for an 83 residue His-heme loop in the denatured state of iso-1-Cytc. This result is consistent with the volume constriction imposed by loop formation favoring residual structure by a mass action effect. The significant stabilization of a 22-residue loop in 3 M gdnHCl observed here for single aromatic guest residues also suggests that relatively large loops can also play a role in the folding mechanism of a protein.

Summary

Using a host-guest approach, we have shown that aromatic residues have a stronger effect on loop stability in comparison to the aliphatic residue, leucine. Previous work on the factors that stabilize native state structure in proteins suggests that the stronger stabilizing effect of aromatic residues in the DSE likely arises from the weakly polar interactions available to aromatic but not aliphatic residues. The residual structure induced by aromatic amino acids primarily affects the persistence of the loop, indicating that it results from the volume constriction imposed by loop formation. The fact that the effect is significant for a 22-residue loop suggests that larger loops may play a more prominent role in the mechanism of protein folding than previously thought. By contrast, proline has little effect on loop persistence, unless it is immediately adjacent to the residue involved in loop formation. The primary effect of proline is to slow loop formation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by award number GM074750 (B.E.B.) and American Recovery and Reinvestment Act supplement GM074750-04S1 from the National Institute of General Medical Sciences. M.L.F acknowledges support from The University of Montana and the MT NSF EPSCoR grant EPS-0701906.

References

1. Mittag T, Forman-Kay JD. Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol.* 2007; 17:3–14. [PubMed: 17250999]
2. Eliezer D. Biophysical characterization of intrinsically disordered proteins. *Curr Opin Struct Biol.* 2009; 19:23–30. [PubMed: 19162471]
3. Neri D, Billeter M, Wider G, Wuthrich K. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science.* 1992; 257:1559–1563. [PubMed: 1523410]

4. Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SMV, Alonso DOV, Daggett V, Fersht AR. The complete folding pathway of a protein from nanoseconds to microseconds. *Nature*. 2003; 421:863–867. [PubMed: 12594518]
5. Nabuurs SM, Westphal AH, van Mierlo CPM. Extensive formation of off-pathway species during folding of an α - β parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J Am Chem Soc*. 2008; 130:16914–16920. [PubMed: 19053416]
6. Kristjansdottir S, Lindorff-Larsen K, Fieber W, Dobson CM, Vendruscolo M, Poulsen FM. Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *J Mol Biol*. 2005; 347:1053–1062. [PubMed: 15784263]
7. Lindorff-Larsen K, Kristjansdottir S, Teilum K, Fieber W, Dobson CM, Poulsen FM, Vendruscolo M. Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J Am Chem Soc*. 2004; 126:3291–3299. [PubMed: 15012160]
8. Lietzow MA, Jamin M, Dyson HJ, Wright PE. Mapping long-range contacts in a highly unfolded protein. *J Mol Biol*. 2002; 322:655–662. [PubMed: 12270702]
9. Wirmer J, Berk H, Ugolini R, Redfield C, Schwalbe H. Characterization of the unfolded state of bovine α -lactalbumin and comparison with unfolded states of homologous proteins. *Protein Sci*. 2006; 15:1397–1407. [PubMed: 16731974]
10. Wirmer J, Schloerb C, Klein-Seetharaman J, Hirano R, Ueda T, Imoto T, Schwalbe H. Protein interactions: modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew Chem Int Ed*. 2004; 43:5780–5785.
11. Klein-Seetharaman J, Oikawa M, Grimshaw SB, Wirmer J, Duchardt E, Ueda T, Imoto T, Smith LJ, Dobson CM, Schwalbe H. Long-range interactions within a nonnative protein. *Science*. 2002; 295:1719–1722. [PubMed: 11872841]
12. Schwarzsinger S, Wright PE, Dyson HJ. Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry*. 2002; 41:12681–12686. [PubMed: 12379110]
13. Shan B, Bhattacharya S, Eliezer D, Raleigh DP. The low-pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry*. 2008; 47:9565–9573. [PubMed: 18707127]
14. Zhang O, Forman-Kay JD. NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry*. 1997; 36:3959–3970. [PubMed: 9092826]
15. O'Brien EP, Ziv G, Haran G, Brooks BR, Thirumalai D. Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc Natl Acad Sci USA*. 2008; 105:13403–13408. [PubMed: 18757747]
16. Bowler BE. Thermodynamics of protein denatured states. *Mol BioSyst*. 2007; 3:88–99. [PubMed: 17245488]
17. Trefethen JM, Pace CN, Scholtz JM, Brems DN. Charge-charge interactions in the denatured state influence the folding kinetics of ribonuclease Sa. *Protein Sci*. 2005; 14:1934–1938. [PubMed: 15937282]
18. Pace CN, Alston RW, Shaw KL. Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci*. 2000; 9:1395–1398. [PubMed: 10933506]
19. Cho JH, Raleigh DP. Experimental characterization of the denatured state ensemble of proteins. *Methods in Molecular Biology* (Totowa, NJ, United States). 2009; 490:339–351.
20. Cho JH, Raleigh DP. Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. *J Mol Biol*. 2005; 353:174–185. [PubMed: 16165156]
21. Cho JH, Sato S, Raleigh DP. Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J Mol Biol*. 2004; 338:827–837. [PubMed: 15099748]
22. Kuhlman B, Luisi DL, Young P, Raleigh DP. pK_a values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions. *Biochemistry*. 1999; 38:4896–4903. [PubMed: 10200179]

23. Bowler, BE. Thermodynamic approaches to understanding protein denatured states. In: Creamer, TP., editor. *Unfolded Proteins: From Denatured to Intrinsically Disordered*. Nova Science Publishers; Hauppauge, NY: 2008. p. 23-50.
24. Pakula AA, Sauer RT. Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface. *Nature*. 1990; 344:363–364. [PubMed: 2314475]
25. Herrmann L, Bowler BE, Dong A, Caughey WS. The effects of hydrophilic to hydrophobic surface mutations on the denatured state of iso-1-cytochrome *c*: investigation of aliphatic residues. *Biochemistry*. 1995; 34:3040–3047. [PubMed: 7893716]
26. Bowler BE, May K, Zaragoza T, York P, Dong A, Caughey WS. Destabilizing effects of replacing a surface lysine of cytochrome *c* with aromatic amino acids: implications for the denatured state. *Biochemistry*. 1993; 32:183–190. [PubMed: 8380333]
27. Krowarsch D, Otlewski J. Amino-acid substitutions at the fully exposed P1 site of bovine pancreatic trypsin inhibitor affect its stability. *Protein Sci*. 2001; 10:715–724. [PubMed: 11274462]
28. Grimsley GR, Shaw KL, Fee LR, Alston RW, Huyghues-Despointes BMP, Thurlkill RL, Scholtz JM, Pace CN. Increasing protein stability by altering long-range coulombic interactions. *Protein Sci*. 1999; 8:1843–1849. [PubMed: 10493585]
29. Robic S, Guzman-Casado M, Sanchez-Ruiz JM, Marqusee S. Role of residual structure in the unfolded state of a thermophilic protein. *Proc Natl Acad Sci USA*. 2003; 100:11345–11349. [PubMed: 14504401]
30. Fu H, Grimsley G, Scholtz JM, Pace CN. Increasing protein stability: importance of ΔC_p and the denatured state. *Protein Sci*. 2010; 19:1044–1052. [PubMed: 20340133]
31. Marsh JA, Forman-Kay JD. Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *J Mol Biol*. 2009; 391:359–374. [PubMed: 19501099]
32. Marsh JA, Neale C, Jack FE, Choy WY, Lee AY, Crowhurst KA, Forman-Kay JD. Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J Mol Biol*. 2007; 367:1494–1510. [PubMed: 17320108]
33. Crowhurst KA, Forman-Kay JD. Aromatic and methyl NOEs highlight hydrophobic clustering in the unfolded state of an SH3 domain. *Biochemistry*. 2003; 42:8687–8695. [PubMed: 12873128]
34. Crowhurst KA, Tollinger M, Forman-Kay JD. Cooperative interactions and a non-native buried Trp in the unfolded state of an SH3 domain. *J Mol Biol*. 2002; 322:163–178. [PubMed: 12215422]
35. Bezsonova I, Evanics F, Marsh JA, Forman-Kay JD, Prosser RS. Oxygen as a paramagnetic probe of clustering and solvent exposure in folded and unfolded states of an SH3 domain. *J Am Chem Soc*. 2007; 129:1826–1835. [PubMed: 17253684]
36. Wong KB, Clarke J, Bond CJ, Neira JL, Freund SMV, Fersht AR, Daggett V. Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *J Mol Biol*. 2000; 296:1257–1282. [PubMed: 10698632]
37. Day R, Daggett V. Ensemble versus single-molecule protein unfolding. *Proc Natl Acad Sci USA*. 2005; 102:13445–13450. [PubMed: 16155127]
38. Kazmirski SL, Wong KB, Freund SMV, Tan YJ, Fersht AR, Daggett V. Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution. *Proc Natl Acad Sci USA*. 2001; 98:4349–4354. [PubMed: 11274353]
39. Fauchère JL, Pliška V. Hydrophobic parameters π of amino acid side chains from the partitioning of N-acetyl-amino acid amides. *Eur J Med Chem - Chim Ther*. 1983; 18:369–375.
40. Wolfenden R, Andersson L, Cullis PM, Southgate CCB. Affinities of amino acid side chains for solvent water. *Biochemistry*. 1981; 20:849–855. [PubMed: 7213619]
41. Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol*. 1984; 179:125–142. [PubMed: 6502707]
42. Nozaki Y, Tanford C. The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. *J Biol Chem*. 1970; 245:1648–1652. [PubMed: 5438355]
43. Godbole S, Bowler BE. A histidine variant of yeast iso-1-cytochrome *c* that strongly affects the energetics of the denatured state. *J Mol Biol*. 1997; 268:816–821. [PubMed: 9180374]

44. Hammack BN, Smith CR, Bowler BE. Denatured state thermodynamics: residual structure, chain stiffness and scaling factors. *J Mol Biol.* 2001; 311:1091–1104. [PubMed: 11531342]
45. Wandschneider E, Bowler BE. Conformational properties of the iso-1-cytochrome *c* denatured state: dependence on guanidine hydrochloride concentration. *J Mol Biol.* 2004; 339:185–197. [PubMed: 15123430]
46. Beck DAC, Alonso DOV, Inoyama D, Daggett V. The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proc Natl Acad Sci USA.* 2008; 105:12259–12264. [PubMed: 18713857]
47. Tzul FO, Kurchan E, Bowler BE. Sequence composition effects on denatured state loop formation in iso-1-cytochrome *c* variants: polyalanine versus polyglycine inserts. *J Mol Biol.* 2007; 371:577–584. [PubMed: 17583729]
48. Tzul FO, Bowler BE. Importance of contact persistence in denatured state loop formation: kinetic insights into sequence effects on nucleation early in folding. *J Mol Biol.* 2009; 390:124–134. [PubMed: 19426739]
49. Jacobson H, Stockmayer WH. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J Chem Phys.* 1950; 18:1600–1606.
50. Schimmel PR, Flory PJ. Conformational energy and configurational statistics of poly-L-proline. *Proc Natl Acad Sci USA.* 1967; 58:52–59. [PubMed: 5231619]
51. Rosell FI, Mauk AG. Spectroscopic properties of a mitochondrial cytochrome *c* with a single thioether bond to the heme prosthetic group. *Biochemistry.* 2002; 41:7811–7818. [PubMed: 12056913]
52. Pollock WBR, Rosell FI, Twitchett MB, Dumont ME, Mauk AG. Bacterial expression of a mitochondrial cytochrome *c*: trimethylation of Lys72 in yeast iso-1-cytochrome *c* and the alkaline conformational transition. *Biochemistry.* 1998; 37:6124–6131. [PubMed: 9558351]
53. Hammack B, Godbole S, Bowler BE. Cytochrome *c* folding traps are not due solely to histidine-heme ligation: direct demonstration of a role for N-terminal amino group-heme ligation. *J Mol Biol.* 1998; 275:719–724. [PubMed: 9480763]
54. Smith CR, Mateljevic N, Bowler BE. Effects of topology and excluded volume on protein denatured state conformational properties. *Biochemistry.* 2002; 41:10173–10181. [PubMed: 12146983]
55. Kristinsson R, Bowler BE. Communication of stabilizing energy between substructures of a protein. *Biochemistry.* 2005; 44:2349–2359. [PubMed: 15709747]
56. Tzul FO, Bowler BE. Denatured states of low complexity polypeptide sequences differ dramatically from those of foldable sequences. *Proc Natl Acad Sci USA.* 2010; 107:11364–11369. [PubMed: 20534569]
57. Kurchan E, Roder H, Bowler BE. Kinetics of loop formation and breakage in the denatured state of iso-1-cytochrome *c*. *J Mol Biol.* 2005; 353:730–743. [PubMed: 16185706]
58. Krieger F, Möglich A, Kiefhaber T. Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc.* 2005; 127:3346–3352. [PubMed: 15755151]
59. Burley SK, Petsko GA. Weakly polar interactions in proteins. *Adv Protein Chem.* 1988; 39:125–189. [PubMed: 3072867]
60. Makhatadze GI, Privalov PL. Energetics of interactions of aromatic hydrocarbons with water. *Biophys Chem.* 1994; 50:285–291. [PubMed: 8011949]
61. Burley SK, Petsko GA. Aromatic-aromatic interaction: a mechanism to stabilize proteins. *Science.* 1985; 229:23–28. [PubMed: 3892686]
62. Chakrabarti P, Bhattacharyya R. Geometry of nonbonded interactions involving planar groups in proteins. *Prog Biophys Mol Biol.* 2007; 95:83–137. [PubMed: 17629549]
63. Mitchell JBO, Laskowski RA, Thornton JM. Non-randomness in side-chain packing: the distribution of interplanar angles. *Proteins: Struct Funct Genet.* 1997; 29:370–380. [PubMed: 9365991]
64. Georis J, Esteves FDL, Lamotte-Brasseur J, Bougnet V, Devreese B, Giannotta F, Granier B, Frere JM. An additional aromatic interaction improves the thermostability and thermophilicity of a

- mesophilic family 11 xylanase: structural basis and molecular study. *Protein Sci.* 2000; 9:466–475. [PubMed: 10752608]
65. Armstrong KM, Fairman R, Baldwin RL. The ($i, i + 4$) Phe-His interaction studied in an alanine-based α -helix. *J Mol Biol.* 1993; 230:284–291. [PubMed: 8450542]
66. Loewenthal R, Sancho J, Fersht AR. Histidine-aromatic interactions in barnase. Elevation of histidine pK_a and contribution to protein stability. *J Mol Biol.* 1992; 224:759–770. [PubMed: 1569555]
67. Saha RP, Bahadur RP, Chakrabarti P. Interresidue contacts in proteins and protein-protein interfaces and their use in characterizing the homodimeric interface. *J Proteome Res.* 2005; 4:1600–1609. [PubMed: 16212412]
68. Campen A, Williams RM, Brown CJ, Meng J, Uversky VN, Dunker AK. TOP-IDP-Scale: a new amino acid scale measuring the propensity for intrinsic disorder. *Protein Peptide Lett.* 2008; 15:956–963.
69. Weikl TR, Dill KA. Folding rates and low-entropy-loss routes of two-state proteins. *J Mol Biol.* 2003; 329:585–598. [PubMed: 12767836]
70. Dill KA, Ozkan SB, Shell MS, Weikl TR. The protein folding problem. *Annu Rev Biophys.* 2008; 37:289–316. [PubMed: 18573083]

Fig. 1. Ribbon diagram of the structure of yeast iso-1-cytochrome *c* (pdb code: 2ycc with the K(-2)H mutation introduced with the mutate function of HyperChem) showing the site of the AAAAAK insert between His(-2) and Ala(-1) (shown as stick models colored by element). The guest amino acids are shown below the alanines in the AAAAAK insert that they replace. The heme cofactor is shown in blue with the iron in red. The heme ligands, Met80 and His18, and Cys14 and Cys17 which covalently attach the heme to the polypeptide chain are shown as stick models colored by element.

Fig. 2. Schematic of His-heme loop formation in the denatured state of iso-1-Cytc. The scheme illustrates the placement of the inserts used in these experiments, along with the abbreviation used for each variant and the sequence of the insert in each variant.

Fig. 3.

Plot of ellipticity at 222 nm, θ_{222} , versus gdnHCl concentration for the L4 variant at 25 °C. Data are shown as open and closed circles. Ellipticity of gdnHCl induced unfolding was monitored at 222 nm to measure the decrease in α helical content. Data were also collected at 250 nm and subtracted as background. Measurements were taken using an Applied Photophysics Chirascan CD Spectrometer linked to a Hamilton Microlab 500 titrator unit. Protein at 4 μ M concentration in $\sim$ 6 M gdnHCl containing 20 mM Tris, 40 mM NaCl, 1 mM EDTA at pH 7 was titrated into 4 μ M protein in 20 mM Tris, 40 mM NaCl, 1 mM EDTA at pH 7 as described previously.⁵⁵ The solid curve is a non-linear least squares fit to a linear free energy relationship assuming two-state folding. Due to the downward curvature in the early part of the native baseline, these points (open circles) are not included in the fit. To avoid bias in the native baseline due to this curvature, we assume that the native baseline is independent of gdnHCl concentration.⁴⁴ Thermodynamic parameters from the fit are collected in Table 1.

Fig. 4.

Comparison of equilibrium denatured state His-heme loop formation for the NH5A (black), P1 (pink), and F4 (green) variants in 3 M gdnHCl at 20 ± 1 °C. The right axis corresponds to the NH5A data. The plots for F4 and P1 correspond to the left axis. pH titrations in 3 M gdnHCl were carried out as described previously.⁴⁵ Wavelength scans from 350 nm to 450 nm were taken using a Beckman DU 800 spectrophotometer over the course of the titration from about pH 7 down to about pH 2 in steps of ~ 0.2 pH units. Protein concentration was 3 μ M and the buffer present was 5 mM sodium phosphate, 15 mM NaCl, 1 mM EDTA. Absorbance data at 398 nm (wavelength at which extinction coefficient of heme Soret band changes most significantly when the strong field His is replaced with a weak field H_2O) versus pH were fit to a modified version of the Henderson-Hasselbalch equation (solid lines) which allows the number of protons, n_p , linked to the process to vary. Parameters from the fits for all variants are collected in Table 2.

Table 1Stability parameters for iso-1-cytochrome *c* variants at 25 °C

Variant	$\Delta G_u^o(\text{H}_2\text{O})$ kcal mol ⁻¹	<i>m</i> -value kcal mol ⁻¹ M ⁻¹	<i>C_m</i> M
NK5A ^a	4.21 ± 0.16	4.11 ± 0.01	1.03 ± 0.04
NH5A ^a	2.00 ± 0.08	4.52 ± 0.18	0.44 ± 0.03
P1	2.60 ± 0.13	5.14 ± 0.19	0.507 ± 0.006
P5	2.22 ± 0.14	4.40 ± 0.42	0.51 ± 0.04
W4	2.10 ± 0.17	5.50 ± 0.41	0.381 ± 0.006
W4P5	2.10 ± 0.15	5.36 ± 0.45	0.392 ± 0.008
L4	2.11 ± 0.24	4.95 ± 0.54	0.427 ± 0.004
Y4	1.81 ± 0.19	4.42 ± 0.21	0.41 ± 0.03
F4	1.59 ± 0.07	3.86 ± 0.30	0.41 ± 0.01

Parameters are the average and standard deviation of three independent trials.

^aParameters for the NH5A and NK5A variants are from ref. 47.

Table 2

Equilibrium denatured state His-heme loop formation for iso-1-cytochrome *c* variants at 20 ± 1 °C in 3 M gdnHCl

Variant	$pK_a(\text{obs})$	n_p	$\Delta\Delta G_{\text{loop}}(\text{His})^a$ kcal mol ⁻¹
NH5A ^b	4.62 ± 0.04	1.06 ± 0.04	-
P1	4.75 ± 0.05	1.08 ± 0.30	0.18 ± 0.08
P5	4.71 ± 0.02	1.00 ± 0.07	0.13 ± 0.06
W4	4.29 ± 0.08	0.97 ± 0.07	-0.44 ± 0.11
W4P5	4.36 ± 0.08	1.02 ± 0.12	-0.36 ± 0.12
L4	4.57 ± 0.05	0.92 ± 0.09	-0.07 ± 0.08
Y4	4.35 ± 0.01	1.14 ± 0.06	-0.36 ± 0.06
F4	4.24 ± 0.01	1.22 ± 0.02	-0.52 ± 0.06

Parameters are the average and standard deviation of three trials.

^a Calculated as $\Delta\Delta G_{\text{loop}}(\text{His}) = \ln(10)RT[pK_a(\text{obs})_{\text{guest}} - pK_a(\text{obs})_{\text{NH5A}}]$. Error is from standard propagation of the errors in $pK_a(\text{obs})$.

^b Parameters for the NH5A variant are from ref. 47.

Table 3Kinetics of denatured state His-heme loop formation in 3 M gdnHCl at 25 °C for iso-1-cytochrome *c* variants

Variant	k_b^a s ⁻¹	k_f^b s ⁻¹	$\Delta\Delta G_b^{\ddagger c}$ kcal mol ⁻¹
NH5A ^d	93.0 ± 0.2	8.9 ± 0.8 × 10 ³	-
P1	80.3 ± 0.7	5.7 ± 0.6 × 10 ³	0.087 ± 0.005
P5	94 ± 3	7.3 ± 0.5 × 10 ³	-0.01 ± 0.02
W4	52.7 ± 0.6	1.1 ± 0.2 × 10 ⁴	0.336 ± 0.007
W4P5	57.9 ± 0.6	1.0 ± 0.2 × 10 ⁴	0.281 ± 0.006
L4	74.2 ± 0.4	8.0 ± 1.5 × 10 ³	0.134 ± 0.001
Y4	64.4 ± 0.2	1.15 ± 0.04 × 10 ⁴	0.217 ± 0.002
F4	45.9 ± 0.4	1.06 ± 0.03 × 10 ⁴	0.418 ± 0.005

^aValues for k_b are the average and standard deviation of k_{obs} for eight trials obtained at pH 3.4.

^bLoop formation rates constants, k_f , are too large to be measured directly by our Applied Photophysics SX-20 stopped-flow instrument. Thus, k_f was evaluated as $k_f = k_b \times K_{loop}(His) = k_b \times 10^{-pK_{loop}(His)}$, where $pK_{loop}(His)$ was evaluated with Eq 1 using $pK_a(HisH^+) = 6.6$.⁴⁵

^c $\Delta\Delta G_b^{\ddagger} = RT \ln(k_b, NH5A / k_b, guest)$, error is from standard propagation of the error in k_b .

^dData from ref. 48.