Published in final edited form as:
Biochemistry. 2012 May 1; 51(17): 3586–3595. doi:10.1021/bi300212a.

Tryptophan Stabilizes His-heme Loops in the Denatured State Only When It Is Near a Loop End[†]

Md. Khurshid A. Khan, **Abbigail L. Miller**, and **Bruce E. Bowler**[*]
Department of Chemistry and Biochemistry, Biochemistry Program and Center for Biomolecular Structure and Dynamics, University of Montana, Missoula, MT 59812

Abstract

We use a host-guest approach to evaluate the effect of Trp guest residues relative to Ala on the kinetics and thermodynamics of His-heme loop formation in the denatured state of iso-1-cytochrome *c* at 1.5, 3.0 and 6.0 M guanidine hydrochloride (GdnHCl). Trp guest residues are inserted in an alanine-rich segment placed after a unique His near the N-terminus of iso-1-cytochrome *c*. Trp guest residues are either four or ten residues from the His end of the 28-residue loop. We find the guest Trp stabilizes the His-heme loop at all GdnHCl concentrations when it is the 4th, but not the 10th, residue from the His end of the loop. Thus, residues near loop ends are most important in developing topological constraints in the denatured state that affect protein folding. In 1.5 M GdnHCl, loop stabilization is ~0.7 kcal/mol providing a thermodynamic rationale for the observation that Trp often mediates residual structure in the denatured state. Measurement of loop breakage rate constants, $k_{b,His}$, indicates that loop stabilization by the Trp guest residues occurs completely after the transition state for loop formation in 6.0 M GdnHCl. Under poorer solvent conditions, approximately half of the stabilization of the loop has developed in the transition state, consistent with contacts in the denatured state being energetically downhill and providing evidence for funneling even near the rim of the folding funnel.

It is long accepted that a protein which folds to a well-defined tertiary structure encodes its structural preference in its primary structure (1, 2). Furthermore, theoretical work indicates that foldable protein sequences have evolved to provide minimally frustrated energy landscapes (3, 4). Thus, primary sequence encodes not only structure, but an efficient folding mechanism. Given this context, it is somewhat surprising that denatured proteins have scaling properties consistent with a classic random coil (5). The resulting viewpoint that the denatured state ensemble (DSE)[1] is devoid of residual structure of substantive import has been reinforced by observations that polypeptides with simple (and unfoldable) sequences have random coil scaling properties (6–8). Theoretical studies, however, demonstrate that random coil scaling can be consistent with substantial residual structure

[†]This research was supported by NIH grant R01-GM074750 (B.E.B). Purchase of the Bruker microflex™ MALDI-ToF mass spectrometer was funded by NSF grant CHE-1039814.

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

Supporting Information. Representative equilibrium GdnHCl unfolding data and pH jump His-heme loop breakage kinetic data. Tables of $k_{f,His}$ and $k_{b,corr}$. This material is available free of charge via the Internet at http://pubs.acs.org.

[1]Abbreviations: DSE, denatured state ensemble; MD, molecular dynamics; GdnHCl, guanidine hydrochloride; iso-1-Cyt*c*, yeast iso-1-cytochrome *c*; ΔG_u^o (H_2O), free energy of unfolding in the absence of denaturant; *m*-value, slope of a plot of free energy of unfolding versus denaturant concentration; C_m, midpoint concentration for unfolding by guanidine hydrochloride; pK_a(obs), apparent pK_a for denatured state His-heme loop formation; pK_{loop}(His), pK for formation of a His-heme loop with a fully deprotonated histidine; ΔG_{loop}(His), free energy of formation of a His-heme loop with a fully deprotonated histidine; m_{eq}, slope of a plot of ΔG_{loop}(His) versus denaturant concentration; $k_{b,His}$, rate constant for His-heme loop breakage, $k_{f,His}$, rate constant for His-heme loop formation; $k_{b,corr}$, rate constant for His-heme loop breakage corrected for the viscosity of guanidine hydrochloride solutions; $m_b^{\ddagger}$, slope of a plot of RT*ln*($k_{b,corr}$) versus guanidine hydrochloride concentration; $\Delta G_b^{\ddagger}$, activation free energy for His-heme loop breakage, TS, transition state, δ_b, fraction of a contact that is broken in the transition state for His-heme loop breakage.

(9). We have recently shown with our His-heme loop formation method that the scaling properties of simple alanine-rich sequences and foldable sequences are similar (10). However, whereas the alanine-rich sequences adhere strictly to random coil scaling, the local variation in chain properties for a foldable heteropolymer is substantial (10, 11).

Considerable evidence has accrued, primarily from NMR structural methods and molecular dynamics (MD) simulations, that denatured states retain residual structure (12–14), both native and non-native in nature (15–24). Substantial data also show that residual structure can involve long-range tertiary interactions (17, 18, 25–32), even under strongly denaturing conditions (30). However, there are counter-examples where local clusters of residual structure have no detectable long-range interactions with other clusters (33, 34). Aromatic residues, particularly tryptophan (15, 16, 19, 22, 29, 30, 35–39), appear to promote residual structure in the DSE. Despite the plethora of structural data, thermodynamic data on the stability of residual structure in the DSE is scarce (12, 40), but important for understanding the role of residual structure in the DSE in specifying tertiary structure and limiting the search to find the native tertiary structure of a protein.

We have recently developed a host-guest variation of our His-heme loop formation method which has provided some of the first quantitative thermodynamic data on the tendency of different amino acids to stabilize residual structure in a denatured protein (41). Since the relative stability of simple loops will be important for defining the topology of a fold early in folding, quantifying the strength of residual structure interactions in the DSE and the distance over which these interactions operate is essential. The site of the guest residue in our initial host-guest study was five residues from one end of the loop (41). Thus, the primary source of the observed stabilization of the loop could be from interactions between residues near the loop ends, since the residues near the ends of the loop (histidine and heme in our system, see Figure 1) are held proximal to each other when the loop is formed. However, it is possible that the reduced conformational space in the closed loop form could promote longer-range interactions between residues within the loop even if they are not near the ends of the loop (42, 43).

In the present study, we apply our host-guest approach with a variant of yeast iso-1-cytochrome *c* (iso-1-Cyt*c*) that has a 12-residue alanine-rich sequence engineered into the N-terminus of the protein (the host variant, NH5A-2, see Figure 1). The NH5A-2 variant can form a 28-residue His-heme loop when the protein is in the denatured state. We use a set of guest variants to compare the effect on the stability of the 28-residue denatured state His-heme loop when an Ala residue is replaced with a guest Trp residue 4 positions from the His end of the loop (AW(4) variant, see Figure 1) with the effect on the stability of the 28-residue denatured state His-heme loop when an Ala residue is replaced with a guest Trp residue 10 positions from the His end of the loop (AW(10) variant, see Figure 1). We also introduce the two Trp guest residues simultaneously (AW(4,10) variant, see Figure 1).

When a disordered polypeptide chain forms a loop, the residues within the loop have less conformational freedom. This loss in conformational freedom can be viewed as an increase in the local concentration of the residues within the loop with respect to each other. The increase in the local concentration of the residues within the loop might be expected to promote interactions between residues within the loop through a mass action effect (Le Chatelier's principle) when the loop forms (42, 43). The effect of increased local concentration when the loop forms should be greater near the ends of the loop than near the center of the loop, since residues at the ends of the loops have more nearest neighbors after the loop forms (see Figure 1). Thus, depending on the sequence distance over which residue-residue interactions occur, the effect of replacing Ala with the more hydrophobic residue

Trp might be expected to be greater when it is placed near the end of the loop (AW(4) variant) versus further from the end of the loop (AW(10) variant).

Since solvent quality has significant effects on denatured state properties (44–46), we have made measurements of His-heme loop formation over the range 1.5 to 6.0 M guanidine hydrochloride (GdnHCl) to determine how solvent quality affects the thermodynamics and kinetics of residual structure stabilization by Trp relative to Ala in the denatured state. We find that the 28-residue denatured state His-heme loop is stabilized and that this stabilization increases with decreasing GdnHCl concentration when Trp guest is 4-residues, but not when it is 10-residues, from the loop end.

EXPERIMENTAL PROCEDURES

Preparation of Variants

The NH5A-2 host variant of iso-1-Cyt*c*, which contains a $(AAAAAK)_2$ insert between an engineered histidine at sequence position –2 (K(–2)H mutation; we use the sequence numbering convention of horse cytochrome *c*, thus, the first five sequence positions of iso-1-Cyt*c* are designated –5 to –1) and Ala(–1) was prepared as previously described (10). The AW(4), AW(10) and AW(4,10) variants which contain Ala→Trp substitutions in the $(AAAAAK)_2$ insert, as described in Results, were prepared using the QuikChange PCR-based site-directed mutagenesis protocol (Agilent Technologies), with the pBTR1 vector (47) containing the NH5A-2 variant as template. The sequence of the entire coding region of each variant was confirmed by dideoxy sequencing at the Murdock DNA Sequencing Facility (University of Montana). All variants contain H39Q, H33N and H26N mutations, in addition to the K(–2)H mutation relative to wild type yeast iso-1-Cyt*c*, so that His(–2) is the only histidine available to form a loop in the denatured state. All variants also contain a C102S mutation to prevent disulfide dimerization during physical studies.

The AW(4), AW(10) and AW(4,10) variants were expressed from *Escherichia coli* BL21(DE3) cells (EdgeBio) (48). Isolation and purification was as described previously (49, 50). These variants all gave multiple chromatographic peaks that required careful separation by HPLC on a BioRad UNO S column. We used the following gradient: 0% for 7 min; 0–30% in 8 min, 30–50% in 15 min (flow rate = 3 mL/min). The major peak was rechromatographed using the same gradient to ensure purity. The identity of the purified material was confirmed by MALDI-ToF mass spectrometry using a Bruker microflex™ mass spectrometer.

GdnHCl Denaturation

GdnHCl denaturation was carried out at pH 7 and 25 °C in 20 mM Tris, 40 mM NaCl, as previously described (49). Circular dichroism data at 222 nm as a function of GdnHCl concentration were fit to a linear free energy relationship using nonlinear least squares methods, as previously described (43). Due to the low stability of these variants, the length of the native state baseline was too short to reliably assess its denaturant dependence. Therefore, the native baseline was assumed to be invariant with denaturant concentration (43).

Equilibrium Loop Formation Measurements

Equilibrium loop formation was carried out at room temperature (22 ± 1 °C), as previously described (42, 49, 50), at 1.5, 3.0 and 6.0 M GdnHCl in the presence of 5 mM sodium phosphate dibasic, 15 mM NaCl, 1 mM EDTA. The integrity of each variant was checked by MALDI-ToF mass spectrometry before and after each experiment. No significant backbone cleavage occurred during any of the experiments reported here (50). Absorbance

data at 398 nm plotted against pH were fit to a modified form of the Henderson-Hasselbalch equation, as described previously (42, 51) giving the apparent pK_a, pK_a(obs), and the number of protons, n_p, associated with His-heme loop formation.

Kinetics of Loop Breakage

Loop breakage kinetics were measured by stopped-flow mixing methods (Applied Photophysics SX-20 spectrometer) monitored at 398 nm. Downward pH jumps were conducted at 1.5, 3.0 and 6.0 M GdnHCl at 25 °C. The starting solution contained 10 mM MES, 1.0 mM EDTA at pH 6.2, 6 μM protein and the desired concentration of GdnHCl. This solution was mixed 1:1 with 100 mM citrate containing 1.0 mM EDTA at either pH 3.5 or pH 3.0 and the same GdnHCl concentration. The actual ending pH values were measured directly from the collected effluent and are reported in Results. Data were fit to a single exponential rise to maximum equation (eq 1) where $A_{398}(t)$ is the absorbance at 398 nm as a

$$A_{398}(t)=\Delta A_{398}[1-\exp(-k_{obs}t)]+A^{\circ}_{398} \quad (1)$$

function of time, t, ΔA_{398} is the amplitude of the kinetic phase, k_{obs} is the observed rate constant and A°398 is the absorbance at 398 nm at time, t = 0. Double exponential fits were also attempted and residuals indicated a better fit with a low amplitude faster phase. However, the standard deviation of the larger rate constant was very large. The seemingly better fit of the double-exponential is probably due to fitting a pressure or flow artifact. Thus, we consider the single exponential fit to be adequate.

RESULTS

Design of Host-guest Variants

In our earlier study of the effects of aliphatic and aromatic amino acids on the stability of His-heme loops in the denatured state, we used a –HAAAAAK– host sequence engineered into the N-terminus of iso-1-Cyt*c* (41). Guest variants were prepared with the sequence –HAAAXAK– with X = L, F, Y and W. The use of an alanine-rich segment as the host-sequence was designed to afford a relatively neutral local sequence that would minimize local side chain interactions with the guest residue.

In the present study, our goal is to quantify the stability of sequentially short-range versus sequentially long-range interactions in the DSE. Thus, we use a longer Ala-rich host sequence, –AAAAAKAAAAAK–, inserted between an engineered histidine at position –2 and Ala(–1) of iso-1-Cyt*c* in our previously reported NH5A-2 variant (10). The host NH5A-2 variant and all of the guest variants form a 28-residue loop in the denatured state (Figure 1). We number the residues in the loop from 1 to 28 starting with His(–2), followed by the Ala-rich insert and then 15-residues of the natural sequence of iso-1-Cyt*c* ending with Cys14. For the purpose of loop numbering, we count the heme as part of the side chain of Cys14 since the heme is covalently attached to Cys14. In our previous host-guest study, we found that F, Y and W all stabilized the His-heme loop by 0.4–0.5 kcal/mol in 3 M GdnHCl (41). We use Trp as the guest residue in this study because structural studies on and simulations of the DSE show that Trp is often a component of residual structure (15, 16, 19, 22, 35–38). Trp has also been implicated in stabilizing long-range residual structure in the DSE (29, 30). Figure 1 shows the placement of the guest Trp residues within the host sequence at positions 4 and 10 of the loop. The AW(4) variant replaces the central Ala of the first five-alanine segment of the host sequence (labeled X_4 in Figure 1) with Trp, making the guest the 4th residue from the His(–2) end of the loop. The AW(10) variant replaces the central Ala of the second five-alanine segment of the host sequence (labeled X_{10} in Figure 1) with Trp, making the guest the 10th residue from the His(–2) end of the loop. Thus, the

guest Trp in the AW(10) variant tests whether a Trp distant from the loop ends can stabilize the denatured state loop relative to Ala due to interactions promoted by the reduced conformational space (higher local concentration) of the residues within the closed loop (42, 43). We have also prepared the AW(4,10) variant, which contains both Trp substitutions in the host sequence to test whether cooperative interactions between the two guests (relative to Ala) can stabilize the 28-residue denatured state His-heme loop.

We note that the lysines are placed at the end of each five-alanine segment in our host sequence to maintain aqueous solubility. The effect of charge on loop stability is potentially a concern. However, it has long been known that GdnHCl neutralizes electrostatic effects on protein stability (52) and it has also been shown that 750 mM NaCl is sufficient to disrupt electrostatic residual structure in the denatured state (53). Thus, electrostatics should have minimal effect on His-heme loop formation with the host-guest variants studied here.

His-heme Loop Formation

When histidine binds to heme in the denatured state of iso-1-Cyt*c*, loop formation is accompanied by release of a proton. Thus, a simple pH titration monitored at the heme Soret absorbance band, which is sensitive to the switch between the strong field histidine ligand and the weak field water ligand, is used to evaluate the His-heme loop equilibrium in the denatured state (see Figure 1). An apparent pK_a, pK_a(obs), is obtained, which is the sum of the histidine ionization equilibrium, pK_a(HisH$^+$), and the equilibrium for binding a deprotonated histidine to the heme, pK_{loop}(His) (eq 2). We have observed that

$$pK_a(obs)=pK_a(HisH^+)+pK_{loop}(His) \quad (2)$$

pK_a(HisH$^+$) = 6.6 ± 0.1 irrespective of sequence position or GdnHCl concentration (11, 42, 51). Thus, we treat pK_a(HisH$^+$) as a physical constant and obtain pK_{loop}(His) from pK_a(obs) with eq 2.

Loop formation equilibria in the denatured state are normally analyzed with eq 3, the Jacobson-Stockmayer equation (54), where ΔS_{loop} is the entropy of loop formation for a random

$$\Delta S_{loop}=-\nu_3 R\ln(n)+R\ln\left[\left(\frac{3}{2\pi C_n \ell^2}\right)^{\nu_3} V_i\right] \quad (3)$$

coil, n is the number of monomers in the loop, R is the gas constant, C_n is Flory's characteristic ratio for a sequence n residues in length (loop size for our purposes), ℓ is the distance between monomers, and V_i is the volume within which the two monomers must be constrained for a loop to form. In this host-guest study, all loops are the same size (n = 28). Thus, when comparing the stability of two loops, the first term of the Jacobson-Stockmayer equation cancels out, yielding eq 4 for the change in loop stability when a guest residue replaces alanine.

$$\Delta pK_{loop}(His)=pK_{loop}(His)_{guest}-pK_{loop}(His)_{host}=\nu_3\log(C_{n,guest}/C_{n,host}) \quad (4)$$

Since Flory's characteristic ratio is expected to be similar to alanine for all residues except proline and glycine (55), we assume that ΔpK_{loop}(His) is due to residual structure induced by the guest residue when the loop forms and not due to changes in C_n.

Stability of Variants

Global stability of the guest variants, AW(4), AW(10) and AW(4,10), was determined using GdnHCl as denaturant and circular dichroism at 222 nm as a probe of secondary structure (Figure S1 of the Supporting Information). The stability parameters obtained are shown in Table 1 along with those for the NH5A-2 variant for comparison. The stability of the AW(10) variant is similar to that of the NH5A-2 variant. By contrast, the AW(4) and AW(4,10) variants are less stable than the NH5A-2 variant, both in terms of the free energy of unfolding in the absence of denaturant, $\Delta G_u^o(H_2O)$, and the denaturation midpoint, C_m. The AW(4) and AW(4,10) variants both have a Trp near the histidine that binds to the heme in the denatured state, whereas the AW(10) variant does not. Thus, the Trp near His(–2) is likely responsible for the decreased stability of the AW(4) and AW(4,10) variants. The stability of all variants in Table 1 (see also Figure S1 of the Supporting Information) is low enough that all variants are completely denatured in 1.5, 3.0 and 6.0 M GdnHCl, the conditions used to measure denatured state His-heme loop formation in this study.

Equilibrium Denatured State His-heme Loop Formation

Denatured state His-heme loop formation was measured by pH titration monitored at 398 nm. Figure 2 compares data for formation of the 28-residue denatured state His-heme loop for the AW(4) and AW(10) variants obtained in 1.5 M GdnHCl. The midpoint of the titration, pK_a(obs), occurs at a significantly lower pH for the AW(4) variant, consistent with the 28-residue denatured state His-heme loop being more stable for this variant than for the AW(10) variant. Equilibrium parameters for loop formation for all variants at 1.5, 3.0 and 6.0 M GdnHCl are collected in Table 2. We note that in all cases, the number of protons released upon loop formation, n_p, is near 1, as expected for the one proton process outlined in Figure 1. It is also evident in Table 2 that pK_a(obs) for the AW(4) and AW(4,10) variants are similar and significantly less than pK_a(obs) for the host variant, NH5A-2, at all GdnHCl concentrations. On the other hand, the pK_a(obs) for the AW(10) variant is the same or slightly less than that of the NH5A-2 host variant at all GdnHCl concentrations. These results show that an Ala to Trp substitution near a loop end is more important for stabilizing a loop formed in the DSE since both AW(4) and AW(4,10) have a guest Trp that is 4 residues from the His(–2) loop end, whereas the guest Trp in the AW(10) variant is 10 residues from His(–2) loop end (see Figure 1).

Figure 3A shows plots of the free energy of loop formation for a fully deprotonated histidine, ΔG_{loop}(His), as a function of GdnHCl concentration for the host and guest variants. The slope of the dependence of ΔG_{loop}(His) on GdnHCl concentration, m_{eq}, is similar for the AW(10) and NH5A-2 variants (see also Table 3). The dependence of ΔG_{loop}(His) on GdnHCl concentration is significantly larger for the AW(4) and AW(4,10) variants (Figure 3A and Table 3). Thus, an Ala to Trp substitution 4 residues from the His(–2) end of the 28-residue denatured state His-heme loop increases m_{eq}, whereas an Ala to Trp substitution 10 residues from the His(–2) end of the 28-residue denatured state His-heme loop does not increase m_{eq}.

Kinetics of Denatured State His-heme Loop Formation

In previous work, we have shown that the kinetics of denatured state His-heme loop formation is consistent with a mechanism involving a rapid histidine deprotonation equilibrium followed by reversible binding of the histidine to the heme (56). This mechanism leads to the expression in eq 5 for the observed rate

$$k_{obs} = k_{b,His} + k_{f,His}\,(K_a(HisH^+)/([H^+] + K_a(HisH^+))) \quad (5)$$

constant for denatured state His-heme loop formation and breakage, k_{obs}, where $k_{b,His}$ is the rate constant for denatured state His-heme loop breakage, $k_{f,His}$ is the rate constant for denatured state His-heme loop formation and $K_a(HisH^+)$ is the acid dissociation constant for the histidine, His(–2), that forms the loop. Thus, pH jump experiments to pH's well below $K_a(HisH^+)$ ($pK_a(HisH^+)$ = 6.6 ± 0.1, irrespective of sequence position or GdnHCl concentration for iso-1-Cyt*c*, (42)) will yield $k_{b,His}$, since $k_{obs} \approx k_{b,His}$ under these conditions. We evaluated $k_{b,His}$ at 1.5, 3.0 and 6.0 M GdnHCl concentration using stopped-flow pH jump mixing from pH ~ 6 to pH ~ 3.5 and pH ~ 3.0 (Table 4 and Figure S2 of the Supporting Information). Within error, k_{obs} is indistinguishable at pH ~3.0 and pH ~ 3.5, indicating that the contribution of $k_{f,His}$ to k_{obs} is insignificant in the pH range 3 to 3.5, and thus provides a good measure of $k_{b,His}$.

Given that equilibrium denatured state His-heme loop formation becomes less favorable as GdnHCl concentration increases (Figure 3A, Table 2), $k_{b,His}$ is expected to increase with increasing GdnHCl concentration. In Table 4, k_{obs} increases from 1.5 to 3.0 M GdnHCl. However, it decreases at 6.0 GdnHCl. We have shown previously that $k_{b,His}$ is inversely proportional to the viscosity of GdnHCl solutions (11), as has been observed for the rate constant for protein unfolding (57–59). Thus, we have applied a viscosity correction to $k_{b,His}$ at 3.0 and 6.0 M GdnHCl, so that the rate constants at 3.0 and 6.0 M GdnHCl are corrected to solution viscosity at 1.5 M GdnHCl ($k_{b,corr}$, Table S2 in the Supporting Information). Plots of $RTln(k_{b,corr})$ versus GdnHCl concentration are provided in Figure 3B. After correction for solution viscosity, the slope of these plots should reflect the change in solvent exposed surface area in going from the closed-loop form to the transition state for loop breakage. The y-axes of the plots of ΔG_{loop}(His) versus GdnHCl concentration and of $RTln(k_{b,corr})$ versus GdnHCl concentration in Figure 3 both cover a range of 2 kcal/mol so that the GdnHCl dependence of the thermodynamic and kinetic data can be compared readily. It is obvious that the slopes of the plots of $RTln(k_{b,corr})$ versus GdnHCl concentration, $m_b^{\ddagger}$, in Figure 3B are less than those of the equilibrium plots in Figure 3A. Thus, m_{eq} is large compared to $m_b^{\ddagger}$ (Table 3). In analogy to the Tanford β-value used to evaluate the position of the transition state for protein folding (60), the ratio, $m_b^{\ddagger}/m_{eq} \approx 0.7$ to 0.8, indicates that the transition state for denatured state His-heme loop breakage lies near the closed-loop form. The smaller $k_{b,corr}$ values (and thus $RTln(k_{b,corr})$ values in Figure 3B) for the AW(4) and AW(4,10) variants relative to the NH5A-2 host variant indicate that the activation free energy for denatured state His-heme loop breakage, $\Delta G_b^{\ddagger}$, increases when the guest Trp is 4 residues from the His(–2) end of the loop. By contrast, $k_{b,corr}$ (and thus $RTln(k_{b,corr})$, see Figure 3B) is similar for the AW(10) guest variant and the NH5A-2 host variant. Thus, a guest Trp residue located 10 residues from the His(–2) end of the loop has little effect on $\Delta G_b^{\ddagger}$.

Discussion

Stabilization of Loops in the DSE Results from Interactions at Loop Ends

The qualitative effect of the guest residue on the stability of the 28-residue denatured state His-heme loop is evident from the plots in Figure 3A. We can quantitatively calculate the stabilization due to the guest residue, $\Delta\Delta G_{loop}$(His), as given by eq 6. The values for each guest variant are

$$\Delta\Delta G_{loop}(His) = \Delta G_{loop}(His)_{guest} - \Delta G_{loop}(His)_{host} \quad (6)$$

collected in Table 5. For the AW(10) variant, $\Delta\Delta G_{loop}$(His) is near zero at all GdnHCl concentrations indicating that a guest residue 10 residues distant from the loop end has almost no effect on denatured state His-heme loop stability relative to Ala. For the AW(4)

and AW(4,10) variants, $\Delta\Delta G_{loop}$(His) is of negative magnitude, indicating stabilization of the 28-residue denatured state His-heme loop. $\Delta\Delta G_{loop}$(His) increases from −0.25 kcal/mol in 6.0 M GdnHCl to − 0.6 to −0.7 kcal/mol at lower GdnHCl concentration. Both AW(4) and AW(4,10) have a guest Trp residue near the His(–2) end of the loop. Thus, a guest Trp that is 4 residues from the His(–2) end of the denatured state His-heme loop stabilizes the loop relative to Ala. Without additional experiments, the exact physical nature of the stabilizing interactions induced by the Trp guest residue four residues from the His(–2) end of the 28-residue denatured state loop is unclear. Possibilities could include dynamic hydrophobic interactions with the heme and with Leu9 or Phe10 (positions 23 and 24 of the 28-residue loop), which are all constrained to be closer to the Trp guest (position 4 of the 28-residue loop) when the loop forms (see Figure 1).

The effect of single-site variants on the stability of a compact DSE formed by the three-helix bundle protein, nucleophosmin C-terminal domain (61) in the presence of 0.5 M NaCl range from 0.5 to 1.7 kcal/mol. The 0.7 kcal/mol stabilization to the DSE caused by an Ala to Trp substitution four residues from the His(–2) end of the 28-residue loop in the denatured state at 1.5 M GdnHCl fall within this range. Shortle and others (62) have also observed similar effects of single-site mutations on denatured state thermodynamics. While the overall stability of these DSE interactions is small, the 0.7 kcal/mol observed at 22 °C for an Ala to Trp substitution four residues from the His(–2) end of the denatured state loop amounts to an ~3-fold change in the conformational distribution of the DSE. The maximum stabilization of the DSE of ~1.7 kcal/mol observed at 10 °C for nucleophosmin C-terminal domain amounts to an ~20-fold effect on the distribution of conformers in the DSE. These effects while small may be adequate to bias the conformational search to find the native state while minimizing the frustration of the folding landscape that would result from stronger interactions that would likely stabilize misfolds.

Experimental studies show that the solubility of the Trp side chain is considerably more responsive to increasing GdnHCl concentration than that of Ala (63, 64). Denaturant *m*-values for global unfolding are usually interpreted in terms of the change in solvent-accessible surface area, ΔSASA, linked to the conformational change (65). For urea denaturation, the dominant contribution to the *m*-value results from solvation of the peptide unit, not the side chains (66). However, a recent combined experimental and simulation study has shown that GdnHCl interacts much more strongly with aromatic than aliphatic model compounds (67), suggesting that aromatic side chains also contribute significantly to GdnHCl *m*-values. Thus, the increase in m_{eq} (Table 3) for equilibrium loop formation observed relative to the NH5A-2 variant for the AW(4) and AW(4,10) guest variants can be attributed to a change in the solvent-exposed surface area of the guest Trp near the His(–2) end of the 28-residue denatured state His-heme loop. Measurements of the difference in the GdnHCl concentration dependence of the solubility of an Ala versus a Trp side chain, indicate that the change in the m_{eq} should be ~0.19 kcal $mol^{-1}M^{-1}$ for transfer of the full side chain into GdnHCl solution (64). The change in m_{eq} observed here for the NH5A-2 and AW(10) variants versus the AW(4) and AW(4,10) variants indicates that about 50% of the solvent exposed surface area of the Trp side chain in the 4^{th} position from the end of the 28-residue denatured state His-heme loop is buried upon loop formation.

In our previous host-guest study, the guest Trp was at the 5^{th} position from the His(–2) end of the denatured state His-heme loop, yielding $\Delta\Delta G_{loop}$(His) = −0.44 ± 0.11 kcal/mol in 3.0 M GdnHCl. In this study, ΔG_{loop}(His) ~ −0.65 kcal/mol in 3.0 M GdnHCl for the AW(4) and AW(4,10) guest variants with a Trp at the 4^{th} position from the His(–2) end of the 28-residue denatured state His-heme loop. Thus, placing the guest residue one position closer to the His(–2) end of the loop increases the stability of the loop by ~0.2 kcal/mol in 3.0 M

GdnHCl, demonstrating that the stabilizing effect of the guest residue increases as the residue gets closer to the end of the denatured state His-heme loop.

Comparing $\Delta\Delta G_{loop}$(His) for the AW(4), AW(10) and AW(4,10) at 1.5, 3.0 and 6.0 M GdnHCl, the individual effects of the two Trp guest residues in AW(4,10) appear to be additive, within error. There is no indication of a cooperative interaction between the two Trp guest residues when the 28-residue denatured state His-heme loop forms with the AW(4,10) variant. Our work on a series of Ala-rich variants with loop sizes from 22 to 46 indicated that there is still steric strain in the 28-residue loop formed by the NH5A-2 variant (10). Thus, we cannot rule out the possibility that the lack of a cooperative interaction between the two Trp residues in the AW(4,10) variant results from inadequate flexibility in the 28-residue denatured state His-heme loop to allow the two Trp residues to interact.

Transition State for Loop Formation

The activation free energy for denatured state His-heme loop breakage, $\Delta G_b^{\ddagger}$, cannot be directly evaluated from our data. However, the effect of the guest residue on the relative height of the barrier for denatured state His-heme loop breakage can be evaluated from $k_{b,His}$ for the host and guest variants as given by eq 7 (41), where $k_{b,His(host)}$ is

$$\Delta\Delta G_b^{\ddagger} = RT\ln(k_{b,His(host)}/k_{b,His(guest)}) \quad (7)$$

the rate constant for denatured state His-heme loop breakage for the NH5A-2 variant at a given GdnHCl concentration and $k_{b,His(guest)}$ is the rate constant for denatured state His-heme loop breakage for one of the guest variants at the same GdnHCl concentration. The $\Delta\Delta G_b^{\ddagger}$ values collected in Table 5 show that the barrier for denatured state His-heme loop breakage increases somewhat as GdnHCl concentration decreases from 6.0 to 1.5 M.

Comparison of $\Delta\Delta G_{loop}$(His) and $\Delta\Delta G_b^{\ddagger}$ at 6.0 M GdnHCl in Table 5 shows that within error, $\Delta\Delta G_b^{\ddagger} = -\Delta\Delta G_{loop}$(His). The simplest interpretation of this observation is that, in 6.0 M GdnHCl, the Trp guest residue stabilizes the closed loop form of the 28-residue denatured state His-heme loop with no effect on the stability of either the transition state (TS) or the open loop form. In other words, the stabilizing interaction induced by the guest residue is not formed in either the open loop form or the TS for denatured state His-heme loop formation in 6.0 M GdnHCl. In analogy to φ values used in protein folding (60), we can define a similar parameter for denatured state loop formation to quantify the effect of a residue at a particular sequence position of the loop on structure stabilization in the TS versus the ground states for the loop equilibrium. For the loop equilibrium in the denatured state in the direction of loop breakage, we will refer to this as a δ_b value as defined in eq 8. For $\delta_b = 1$, the contact that stabilizes the closed-loop form of the

$$\delta_b = \frac{\Delta\Delta G_b^{\ddagger}}{-\Delta\Delta G_{loop}(His)} \quad (8)$$

denatured state loop in the guest variant breaks completely in the TS. For $\delta_b = 0$, the contact does not break on the way to the TS.

The δ_b values for all guest variants at all GdnHCl concentrations are listed in Table 5. We note that since $\Delta\Delta G_{loop}$(His) is small for the AW(10) variant, the errors in δ_b are large for this variant and thus δ_b for the AW(10) variant should be viewed with caution. In 6.0 M GdnHCl, $\delta_b = 1$, within error, for the AW(4) and AW(4,10) variants. Remembering that the *m*-values in Table 3 indicate that the TS for denatured state His-heme loop formation is near

the closed-loop form, one possible interpretation for $\delta_b = 1$ in 6.0 M GdnHCl is that the search to find the TS for loop formation from the open loop form is mostly entropic. In other words no significant polymer-polymer interactions occur in the TS for denatured state His-heme loop formation in 6.0 M GdnHCl. Such behavior might be expected for a polymer in a good solvent.

By contrast, $\delta_b = 0.5$ to 0.6 in 1.5 and 3.0 M GdnHCl. In analogy to analysis of φ values for protein folding (60), fractional δ_b could be interpreted as resulting from partial breakage of a contact in the TS or from parallel pathways with and without the contact broken in the TS. Either way, the conclusion remains the same. Under poorer solvent conditions, polymer-polymer contacts are now able to stabilize the TS for forming a simple loop in the DSE. In contrast to good solvent conditions (6.0 M GdnHCl), much of the increase in $\Delta\Delta G_{loop}$(His) at lower GdnHCl concentrations caused by a Trp guest residue near the end of the loop occurs in the TS. Thus, our host-guest model system shows that residual structure in the DSE near the rim of the folding funnel has funneling behavior in poorer solvents.

Implications for Specificity of Residual Structure in the DS

Studies on the DSE of a number of proteins show that aromatics, in particular Trp, are often involved in residual structure (15, 16, 19, 22, 29, 30, 35–38). In the case of Lysozyme (29, 30), Trp mediates long-range interactions in the DSE. Our data show that in lower GdnHCl concentrations, the Trp stabilizes residual structure in the DSE relative to Ala. Linear extrapolation of the data in Figure 3A to 0 M GdnHCl indicates that a single Trp, four residues from the end of a denatured state loop, can provide 0.9 to 1.1 kcal/mol of stabilization to residual structure in the DSE relative to Ala. Thus, in the absence of denaturant at 25 °C, a single Trp has a >5-fold greater tendency to form contacts in the DSE compared to Ala. Thus, it is not surprising that Trp is commonly observed as a component of residual structure in the DSE. Our data also show that the proximity of the side chain to the point of contact is important for stabilizing residual structure in the DSE. Thus, when two residues contact each other in the DSE, the persistence of that contact, and thus its ability to bias the conformational distribution of the DSE (and by extension the folding process), will depend on the sequence adjacent to the residues that make contact.

The evidence for transient long-range interactions in the DSE is substantial (25–31). Although, long-range interactions are not always observed (33, 34). Lysozyme (29, 30) and Im7 (33, 34) provide an interesting point of comparison. Lysozyme contains multiple Trp residues. Substitution of any of these tryptophans results in loss of loss of residual structure throughout the protein (29, 30). Im7 on the other hand contains only a single Trp and shows no evidence of mutation-induced long-range effects on residual structure in the DSE. The generality of this observation certainly requires further testing, but this comparison suggests an important role for Trp in mediating long-range interactions in the DSE. The significant stabilization of a His-heme loop we observe when Trp is near a contact interface and the denaturant dependence of this stabilization show that a single Trp can shift the population of species by >5-fold in the DSE. Our previous host-guest study (41), indicated that Phe and Tyr have similar stabilizing effects on His-heme loop formation in the DSE in 3.0 M GdnHCl. However, the denaturant concentration dependence of the solubility of the Trp side is significantly higher than for Phe and Tyr (64), particularly in urea solutions, the denaturant in the Lysozyme and Im7 studies (29, 30, 33, 34). Thus, the advantage of Trp as an agent to stabilize residual structure, particularly at long-range, should be substantial in the DSE in aqueous solution.

CONCLUSIONS

Using a host-guest approach, we have shown that residues proximal to the end of a loop are more important for stabilizing a loop in the denatured state than residues further from the end of the loop. Relative to Ala, a single Trp four residues from the histidine end of a loop can stabilize a 28-residue denatured state His-heme loop by 0.6 to 0.7 kcal/mol in 1.5 to 3.0 M GdnHCl. At a distance of ten residues from the histidine end of the loop, the ability of Trp relative to Ala to stabilize the 28-residue denatured state His-heme loop approaches 0 kcal/mol at all GdnHCl concentrations. Furthermore, under good solvent conditions, the guest Trp provides no stabilization of the TS for denatured state loop formation. However, as solvent conditions become poorer, the guest Trp is able to stabilize the TS for denatured state loop formation, demonstrating the energetically downhill nature of formation of simple loop contacts under solvent conditions that approach folding conditions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973; 181:223–230. [PubMed: 4124164]
2. Haber E, Anfinsen CB. Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. J Biol Chem. 1962; 237:1839–1844. [PubMed: 13903380]
3. Oliveberg M, Wolynes PG. The experimental survey of protein-folding energy landscapes. Q Rev Biophys. 2005; 38:245–288. [PubMed: 16780604]
4. Wolynes PG. Energy landscapes and solved protein-folding problems. Philos Trans R Soc London, Ser A. 2005; 363:453–464.
5. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, Hasan MZ, Pande VJ, Ruczinski I, Doniach S, Plaxco KW. Random coil behavior and the dimensions of chemically unfolded proteins. Proc Natl Acad Sci USA. 2004; 101:12491–12496. [PubMed: 15314214]
6. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T. Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. J Mol Biol. 2003; 332:265–274. [PubMed: 12946363]
7. Bieri O, Wirz J, Hellrung B, Schutkowski M, Drewello M, Kiefhaber T. The speed limit for protein folding measured by triplet-triplet energy transfer. Proc Natl Acad Sci USA. 1999; 96:9597–9601. [PubMed: 10449738]
8. Lapidus LJ, Eaton WA, Hofrichter J. Measuring the rate of intramolecular contact formation in polypeptides. Proc Natl Acad Sci USA. 2000; 97:7220–7225. [PubMed: 10860987]
9. Fitzkee NC, Rose GD. Reassessing random-coil statistics in unfolded proteins. Proc Natl Acad Sci USA. 2004; 101:12497–12502. [PubMed: 15314216]
10. 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]
11. Dar TA, Schaeffer RD, Daggett V, Bowler BE. Manifestations of native topology in the denatured state ensemble of *Rhodopseudomonas palustris* cytochrome c'. Biochemistry. 2011; 50:1029–1041. [PubMed: 21190388]
12. Bowler BE. Residual structure in unfolded proteins. Curr Opin Struct Biol. 2012; 22:4–13. [PubMed: 21978577]
13. Mittag T, Forman-Kay JD. Atomic-level characterization of disordered protein ensembles. Curr Opin Struct Biol. 2007; 17:3–14. [PubMed: 17250999]
14. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. Chem Rev. 2004; 104:3607–3622. [PubMed: 15303830]

15. 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]
16. Marsh JA, Neale C, Jack FE, Choy W-Y, 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]
17. 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]
18. 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]
19. Day R, Daggett V. Ensemble versus single-molecule protein unfolding. Proc Natl Acad Sci USA. 2005; 102:13445–13450. [PubMed: 16155127]
20. DeMarco ML, Alonso DOV, Daggett V. Diffusing and colliding: the atomic level folding/unfolding pathway of a small helical protein. J Mol Biol. 2004; 341:1109–1124. [PubMed: 15328620]
21. 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]
22. Wong K-B, 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]
23. Mayor U, Johnson CM, Daggett V, Fersht AR. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. Proc Natl Acad Sci USA. 2000; 97:13518–13522. [PubMed: 11087839]
24. 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]
25. 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]
26. Bruun SW, Iešmantavičius V, Danielsson J, Poulsen FM. Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein. Proc Natl Acad Sci USA. 2010; 107:13306–13311. [PubMed: 20624986]
27. Kortemme T, Kelly MJS, Kay LE, Forman-Kay J, Serrano L. Similarities between the spectrin SH3 domain denatured state and its folding transition state. J Mol Biol. 2000; 297:1217–1229. [PubMed: 10764585]
28. Felitsky DJ, Lietzow MA, Dyson HJ, Wright PE. Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. Proc Natl Acad Sci USA. 2008; 105:6278–6283. [PubMed: 18434548]
29. 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.
30. 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]
31. Fedyukina DV, Rajagopalan S, Sekhar A, Fulmer EC, Eun Y-J, Cavagnero S. Contributions of long-range interactions to the secondary structure of an unfolded globin. Biophys J. 2010; 99:L37–L39. [PubMed: 20816043]
32. Morrone A, McCully ME, Bryan PN, Brunori M, Daggett V, Gianni S, Travaglini-Allocatelli C. The denatured state dictates the topology of two proteins with almost identical sequence but different native structure and function. J Biol Chem. 2011; 286:3863–3872. [PubMed: 21118804]

33. Le Duff CS, Whittaker SB-M, Radford SE, Moore GR. Characterisation of the conformational properties of urea-unfolded Im7: implications for the early stages of protein folding. J Mol Biol. 2006; 364:824–835. [PubMed: 17045607]
34. Pashley CL, Morgan GJ, Kalverda AP, Thompson GS, Kleanthous C, Radford SE. Conformational properties of the unfolded state of Im7 in Nondenaturing Conditions. J Mol Biol. 2012; 416:300–318. [PubMed: 22226836]
35. 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]
36. 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]
37. 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]
38. 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]
39. 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.
40. Bowler BE. Thermodynamics of protein denatured states. Mol BioSyst. 2007; 3:88–99. [PubMed: 17245488]
41. Finnegan ML, Bowler BE. Propensities of aromatic amino acids versus leucine and proline to induce residual structure in the denatured-state ensemble of iso-1-cytochrome *c*. J Mol Biol. 2010; 403:495–504. [PubMed: 20850458]
42. 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]
43. 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]
44. Teufel DP, Johnson CM, Lum JK, Neuweiler H. Backbone driven collapse in unfolded protein chains. J Mol Biol. 2011; 409:250–262. [PubMed: 21497607]
45. Vitalis A, Wang X, Pappu RV. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J Mol Biol. 2008; 384:279–297. [PubMed: 18824003]
46. Ziv G, Thirumalai D, Haran G. Collapse transitions in proteins. Phys Chem Chem Phys. 2009; 11:83–93. [PubMed: 19081910]
47. 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]
48. Pollock WB, 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]
49. 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]
50. 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]
51. Rao KS, Tzul FO, Christian AK, Gordon TN, Bowler BE. Thermodynamics of loop formation in the denatured state of *Rhodopseudomonas palustris* cytochrome c': scaling exponents and the reconciliation problem. J Mol Biol. 2009; 392:1315–1325. [PubMed: 19647747]
52. Monera OD, Kay CM, Hodges RS. Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contribution of electrostatic interactions. Protein Sci. 1994; 3:1984–1991. [PubMed: 7703845]

53. 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]
54. Jacobson H, Stockmayer WH. Intramolecular reaction in polycondensations. I. The theory of linear systems. J Chem Phys. 1950; 18:1600–1606.
55. 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]
56. 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]
57. Jacob M, Geeves M, Holtermann G, Schmid FX. Diffusional barrier crossing in a two-state protein folding reaction. Nat Struct Mol Biol. 1999; 6:923–926.
58. Qiu L, Hagen SJ. Internal friction in the ultrafast folding of the tryptophan cage. Chem Phys. 2005; 312:327–333.
59. Qiu L, Hagen SJ. A limiting speed for protein folding at low solvent viscosity. J Am Chem Soc. 2004; 126:3398–3399. [PubMed: 15025447]
60. Fersht, A. Structure and Mechanism in Protein Science. W. H. Freeman and Company; New York: 1998.
61. Scaloni F, Federici L, Brunori M, Gianna S. Deciphering the folding transiton state structure and denatured state properties of nucleophosmin C-terminal domain. Proc Natl Acad Sci USA. 2010; 107:5447–5452. [PubMed: 20212148]
62. Shortle D. Staphylococcal nuclease: a showcase of *m*-value effects. Adv Protein Chem. 1995; 46:217–245. [PubMed: 7771319]
63. 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]
64. 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]
65. Myers JK, Pace CN, Scholtz JM. Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci. 1995; 4:2138–2148. [PubMed: 8535251]
66. Auton M, Holthauzen LMF, Bolen DW. Anatomy of energetic changes accompanying urea-induced protein denaturation. Proc Natl Acad Sci USA. 2007; 104:15317–15322. [PubMed: 17878304]
67. Mason PE, Dempsey CE, Neilson GW, Kline SR, Brady JW. Preferential interactions of guanidinium ions with aromatic groups over aliphatic groups. J Am Chem Soc. 2009; 131:16689–16696. [PubMed: 19874022]

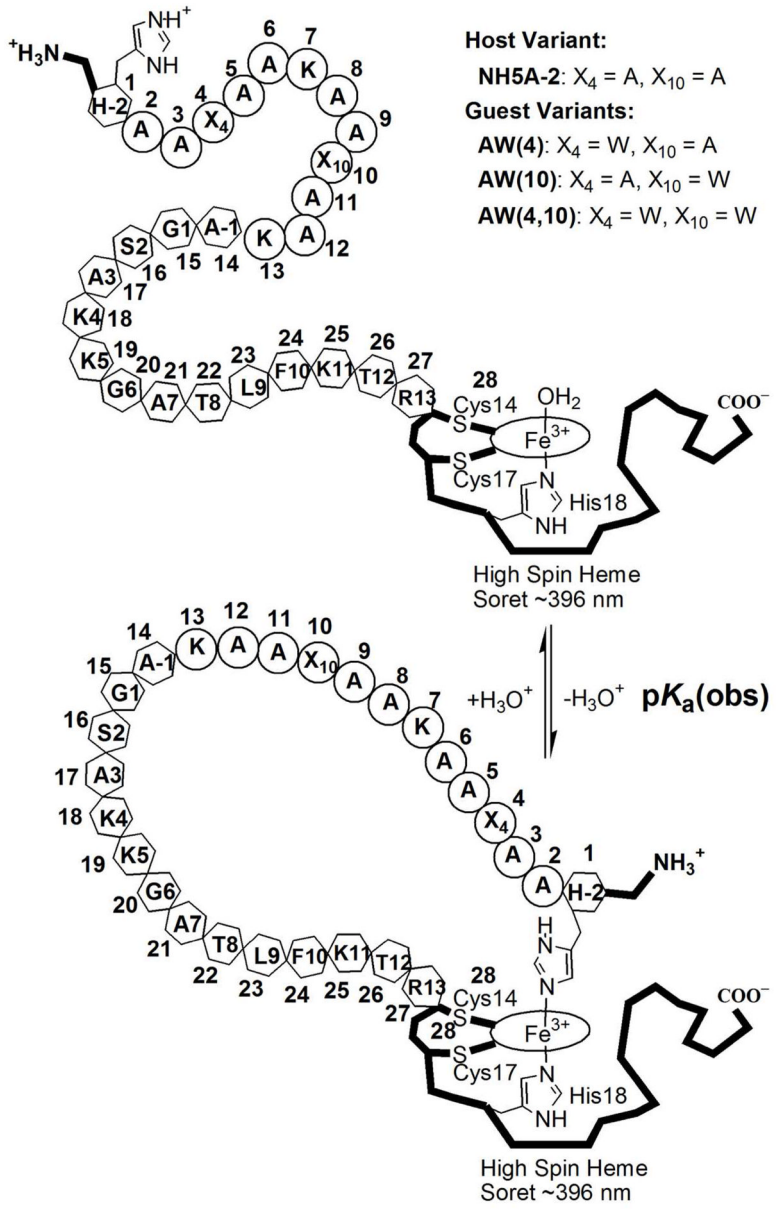

Figure 1.

Schematic representation of denatured state His-heme loop formation with the 28-residue loop containing a 12-residue alanine-rich host sequence. The residues of the alanine-rich insert between His(–2) and Ala(-1) of iso-1-Cyt*c* are represented by circles. The positions used for substitution of an Ala of the Ala-rich host sequence with a Trp are labeled X_4 and X_{10}, where the subscripted numbers represented the position of the amino acid monomer in the 28-residue denatured state loop. His(–2) derived from a K(–2)H mutation and the natural sequence running from Ala(–1) to Arg(13) at the C-terminal end of the loop are represent by hexagons. The sequence numbering for the natural sequence is based on the horse cytochrome *c* numbering convention applied to wild type (no Ala-rich insert) iso-1-Cyt*c*. Each member of the 28-residue loop is numbered 1 to 28 starting with His(–2) at the N-terminus of the denatured state loop and ending with Cys14 at the C-terminus of the denatured state loop. The iso-1-Cyt*c* polypeptide outside the residues involved in denatured state loop formation is represented schematically by thick black lines. The heme is

represented by an oval with an Fe^{3+} in the center. The His(–2) and His18 side chains are shown explicitly. The sulfur atoms of Cys14 and Cys17, which form thioether linkages to the heme, are also shown. The nomenclature for the variants is provided in the upper right corner of the scheme. For each variant the residue at position 4, X_4, and position 10, X_{10}, of the denatured state loop is given with the one letter abbreviations for Trp (W) and Ala (A).

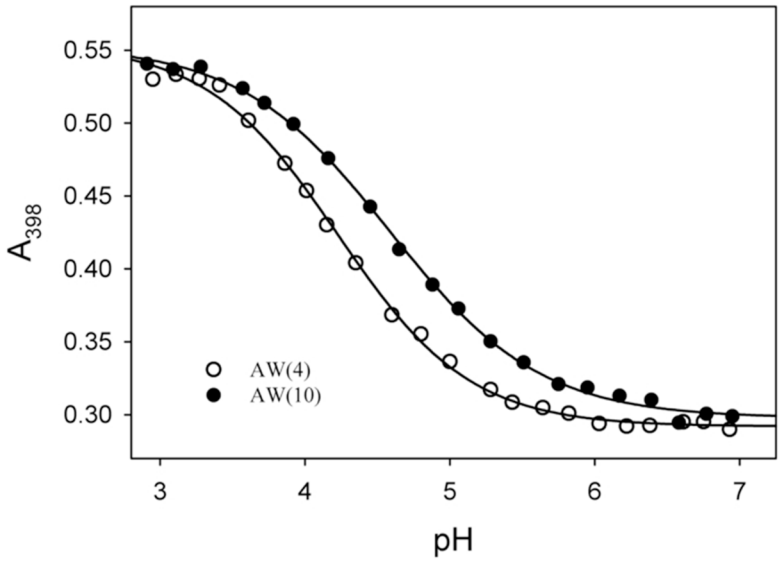

Figure 2.

Representative equilibrium denatured state His-heme loop formation titrations for the AW(4) and AW(10) guest variants. Data for the AW(4) (○) and AW(10) (●) variants were acquired at 3 μM protein concentration in the presence of 1.5 M GdnHCl at 22 ± 1 °C. Plots of A_{398} versus GdnHCl concentration were fit to a modified form of the Henderson-Hasselbalch equation that allows the number of protons, n_p, linked to denatured state His-heme loop formation to vary. The fits are shown as solid lines. Parameters from the fits are collected in Table 2.

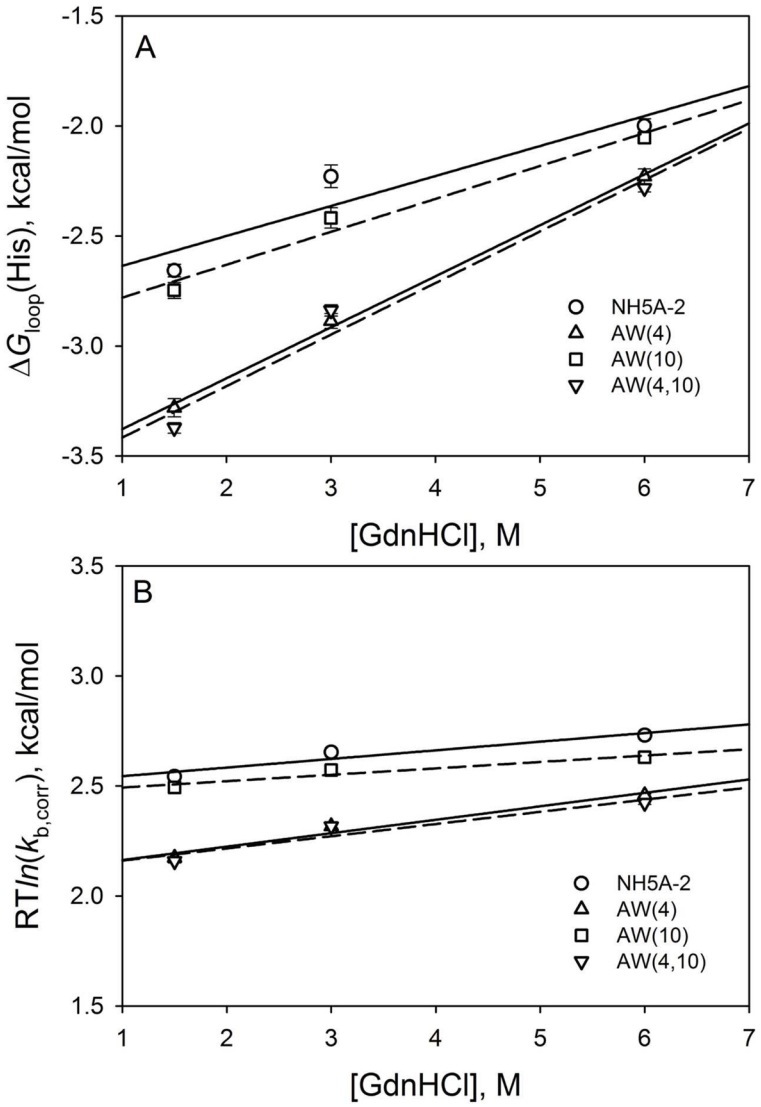

Figure 3.

Plots of (A) ΔG_{loop}(His) and (B) RT*ln*($k_{b,corr}$) as a function of GdnHCl concentration. ΔG_{loop}(His) was calculated as *ln*(10)RTpK_{loop}(His), where pK_{loop}(His) was derived from the pK_a(obs) data in Table 2 using eq 1 with pK_a(HisH$^+$) = 6.6 (42). In part B, the $k_{b,corr}$ values for the data at 3.0 and 6.0 M are from Table S2 in the Supporting Information and are corrected to the viscosity of 1.5 M GdnHCl, as described in the footnotes to Table S2 of the Supporting Information. We use the k_{obs} values obtained at pH 3.15 at 1.5 M GdnHCl for $k_{b,His}$ at 1.5 M GdnHCl. The solid (NH5A-2, AW(4)) and dashed (AW(10) and AW(4,10)) lines in parts A and B assume a linear relationship between free energy and GdnHCl concentration.

Table 1

Thermodynamic Parameters for GdnHCl Unfolding of Iso-1-Cyt*c* Variants at 25 °C and pH 7.[a]

Variant	$\Delta Gu^{o'}(H_2O)$ kcal/mol	*m*-value kcal $mol^{-1}M^{-1}$	C_m M
NH5A-2[b]	2.00 ± 0.41	4.32 ± 0.60	0.46 ± 0.04
AW(4)	1.00 ± 0.08	2.96 ± 0.23	0.34 ± 0.03
AW(10)	1.73 ± 0.05	3.53 ± 0.13	0.49 ± 0.02
AW(4,10)	1.06 ± 0.07	2.92 ± 0.15	0.36 ± 0.01

[a] Values are the averages and standard deviations of three independent trials.

[b] Data are from ref. (10).

Table 2

Thermodynamic Parameters for His-heme Loop Formation for Iso-1-Cyt*c* Variants at 22 ± 1 °C and in 1.5, 3.0 and 6.0 M GdnHCl.[a]

Variant	pK_a(obs)	n_p
	1.5 M GdnHCl	
NH5A-2	4.65 ± 0.02	1.00 ± 0.05
AW(4)	4.19 ± 0.03	0.96 ± 0.01
AW(10)	4.59 ± 0.03	0.83 ± 0.02
AW(4,10)	4.13 ± 0.02	0.97 ± 0.01
	3.0 M GdnHCl	
NH5A-2	4.97 ± 0.04[b]	1.05 ± 0.13[b]
AW(4)	4.48 ± 0.02	1.01 ± 0.05
AW(10)	4.83 ± 0.03	1.04 ± 0.02
AW(4,10)	4.52 ± 0.02	0.92 ± 0.06
	6.0 M GdnHCl	
NH5A-2	5.13 ± 0.02[b]	1.11 ± 0.01[b]
AW(4)	4.97 ± 0.02	1.01 ± 0.05
AW(10)	5.09 ± 0.01	1.04 ± 0.02
AW(4,10)	4.93 ± 0.01	0.92 ± 0.06

[a]Values are the average and standard deviation of three independent trials.

[b]Data are from *ref.* (10).

Table 3

Equilibrium and Kinetic Denaturant *m*-Values for Iso-1-Cyt*c* Variants.[a]

Variant	m_{eq}	$m_b^{\ddagger}$
NH5A-2	0.14 ± 0.05	0.04 ± 0.01
AW(4)	0.23 ± 0.01	0.06 ± 0.01
AW(10)	0.15 ± 0.02	0.03 ± 0.01
AW(4,10)	0.23 ± 0.04	0.06 ± 0.02

[a]Values obtained from the slopes of the linear fits to the data in Figure 3. Reported error is the standard error in the slope reported by SigmaPlot.

Table 4

Rate Constants for Loop Breakage at 25 °C in 1.5, 3.0 and 6.0 M GdnHCl for the Iso-1-Cyt*c* Variants.[a]

Variant	[GdnHCl], M	k_{obs}, s^{-1}	
		pH 3.15 ± 0.04[b]	pH 3.68 ± 0.05
NH5A-2	1.5	73.2 ± 0.7	73.5 ± 0.8
AW(4)		39.2 ± 0.2	42.9 ± 0.4
AW(10)		67.2 ± 0.8	66.8 ± 1.1
AW(4,10)		38.2 ± 0.6	43.6 ± 0.5
		pH 3.11 ± 0.05[b]	pH 3.59 ± 0.1
NH5A-2	3.0	79.9 ± 0.7[c]	77.9 ± 0.5[c]
AW(4)		45.3 ± 0.6	45.4 ± 0.5
AW(10)		69.7 ± 1.3	67.1 ± 1.6
AW(4,10)		45.3 ± 0.6	44.0 ± 0.5
		pH 3.06 ± 0.1[b]	pH 3.52 ± 0.14
NH5A-2	6.0	65.7 ± 2.0[c]	60.0 ± 1.4[c,d]
AW(4)		41.5 ± 0.4	40.7 ± 0.7
AW(10)		55.6 ± 0.7	52.2 ± 0.4
AW(4,10)		39.1 ± 0.7	38.6 ± 0.5

[a]Data are the average and standard deviation of at least six trials.

[b]The k_{obs} data from downward pH jump experiments at the lower pH value for each variant are equivalent to $k_{b,His}$ since the calculated contribution from k_f at these low pHs for the 28-residue loop is 2–4, 1–2 and 0.5–0.6 s^{-1} at 1.5, 3.0 and 6.0 M gdnHCl. The contribution of $k_{f,His}$ was calculated as $k_{f,His}/(1+10^{(6.6\text{-pH})})$ where pH is the final pH after stopped-flow mixing and $k_{f,His}$ is from Table S1 in the Supporting Information.

[c]Data are from *ref.* (10).

[d]Final measured pH for the downward jump to pH 3.5 was near pH 4 in this case.

Table 5

Mutation-induced $\Delta\Delta G_{loop}$(His), $\Delta\Delta G^{\ddagger}_{b}$, and δ_b for the Iso- 1-Cyt*c* Guest Variants at 25 °C.[a]

Variant	1.5 M GdnHCl	3.0 M GdnHCl	6.0 M GdnHCl
	$\Delta\Delta G_{loop}$(His), kcal/mol		
AW(4)	−0.62 ± 0.05	−0.66 ± 0.06	−0.23 ± 0.05
AW(10)	−0.09 ± 0.05	−0.19 ± 0.07	−0.05 ± 0.03
AW(4,10)	−0.72 ± 0.04	−0.61 ± 0.06	−0.28 ± 0.04
	$\Delta\Delta G_b^{\ddagger}$, kcal/mol		
AW(4)	0.37 ± 0.01	0.34 ± 0.01	0.27 ± 0.01
AW(10)	0.05 ± 0.01	0.08 ± 0.01	0.10 ± 0.02
AW(4,10)	0.39 ± 0.01	0.34 ± 0.01	0.31 ± 0.02
	δ_b		
AW(4)	0.59 ± 0.05	0.52 ± 0.05	1.2 ± 0.3
AW(10)	0.57 ± 0.31	0.43 ± 0.17	1.8 ± 1.2
AW(4,10)	0.54 ± 0.03	0.55 ± 0.06	1.1 ± 0.2

[a]Reported errors are from standard propagation of error.