Published in final edited form as:

J Mol Biol. 2007 August 17; 371(3): 577–584.

Sequence Composition Effects on Denatured State Loop Formation in Iso-1-cytochrome c Variants: Polyalanine versus Polyglycine Inserts

Franco O. Tzul, Eydiejo Kurchan, and Bruce E. Bowler*

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

Abstract

Protein folding is dependent on the formation and persistence of simple loops during the earliest events of the folding process. Ease of loop formation and persistence is believed to be dependent on the steric properties of the residues involved in loop formation. We have investigated this conformational factor in the denatured state of iso-1-cytochrome *c* using a five alanine insert in front of a unique histidine in the N-terminal region of the protein. The alanines have then been progressively substituted with sterically less-constrained glycine residues. Guanidine-HCl unfolding shows that all variants have a free energy of unfolding of ~2 kcal/mol. The low stability of these variants is well accounted for by stabilization of the denatured state by histidine-heme loop formation. The stability of the 22-residue histidine-heme loop has been measured in 3 M guanidine hydrochloride for all variants. Surprisingly, relative to alanine, glycine has only a very modest effect on equilibrium loop stability. Thus, the greater flexibility that glycine confers on the main chain provides no advantage in terms of the persistence of simple loops early in folding. The underlying basis for the similar behavior of loops with polyalanine versus polyglycine inserts is discussed in terms of the current knowledge of the structure and loop formation kinetics of glycine-versus alanine-rich peptides.

Keywords

Denatured States; Protein Folding; Loop Formation; Sequence Composition; Side Chain Sterics

Early work on protein folding assumed that the denatured state of a protein is devoid of any persistent structure. The protein polypeptide chain was considered to be a random coil having minimal structural periodicity.¹ Nowadays, it is widely accepted that protein denatured states can contain structures of intermediate complexity known as residual structure.² These structures have been shown to persist even under relatively high denaturing conditions^{3–10} and are likely an important factor in protein folding.^{8,9} Such residual structure can significantly impact the stability of the denatured state (for a recent review see ref. ¹¹). In particular, studies on the pH dependence of the stability of the N-terminal domain of ribosomal protein L9 (NTL9)^{12–14} and staphylococcal nuclease,¹⁵ as well as investigations into the effects of surface electrostatics on the stability of RNase T1, RNase Sa and T4 lysozyme^{16–18} show that ionic interactions can stabilize protein denatured states by 1 to 4 kcal/mol. These lingering structures

* To whom correspondence should be addressed. Phone (406) 243-6114. Fax (406) 243-4227. E-mail: 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.

in a protein's denatured state can limit the conformational search or "walk across conformational space" towards a thermodynamic minimum (bottom of folding funnel). In so doing, these non-random structures greatly increase protein folding efficiency,¹⁹ thereby satisfying Levinthal's Paradox. Unfortunately, the question of how this increased efficiency is accomplished is still not fully understood.

In order to determine the mechanism by which residual structure increases protein folding efficiency, the formation and persistence of such structure must first be assessed. Statistically it has been shown that simple loops are much easier to form as compared to more complex β turns, β sheets and α helices.²⁰ Loop formation is necessary in nucleation of more complex structures such as α helix secondary structure^{21,22} and relatively large loops have been implicated in the folding mechanism of some proteins.^{23–25} Hence, "loop closure" can be considered a basic step in protein folding reactions, which sets a "speed limit" for folding.^{20, 26,27}

Very primitive loop structures are the basic starting point in residual structure formation. Thus, investigations into loop formation and their development into residual structure are vital. To this end, it is essential to understand the fundamental conformational constraints that act on a disordered polypeptide chain. This investigation focuses on a single 22 residue loop formed from the N-terminal side of iso-1-cytochrome *c* and acts as a simple test of important factors in polymer theory. Specifically, we investigate flexibility of the main chain due to steric bulk of the R-group and the effect that less conformationally restrictive amino acid residues would have on polypeptide compactness and thus the equilibrium stability of loop formation. A more flexible chain is more compact¹ and loop formation might be expected to be more favorable in a more compact polymer. However, Robinson & Sauer²⁸ suggested that equilibrium stability of loops or linkers between two subunits of a protein is primarily dependent on composition and residue ratios rather than sequence. In particular, a balance between adequate versus too much flexibility was essential for optimal stability. This investigation offers further insights into this notion.

Mutant design

Mutants were designed with the triple mutant (TM) variant of yeast iso-1-cytochrome *c* as the template.^{29,30} This TM variant has the mutations H26N, H33N and H39Q which remove all histidines that can participate in denatured state His-heme ligation. Since the TM variant does not form His-heme loops in the denatured state, it also acts as a control to assess the effects of such loops on global stability. We have found in previous work that the N-terminal amino group competes with His-heme loop formation in the denatured state of iso-1-cytochrome *c*^{31,32} and must be blocked through *in vivo* N-terminal acetylation in yeast to allow His-heme affinity in the denatured state to be measured accurately.³² We have also studied variants having the mutation K(-2)H (the tuna numbering system is used whereby the first 5 amino acids in yeast are numbered -5 to -1) that can make a 16 residue His-heme loop in the denatured state.^{33,34} The histidine at position -2 in these variants has a high affinity for the heme due its location on the N-terminal side of the site of heme attachment to iso-1-cytochrome *c*.³³ Thus, homopolymeric sequence insertions have been made on the N-terminal side of the site of heme attachment to make N-terminal amino group competition with His-heme binding in the denatured state insignificant. This is important for the current work since the N-terminus of cytochrome *c* is not acetylated in *Escherichia coli* which is being used to express the variant proteins described here.

We initially inserted the sequence KAAAAA in between F(-3) and K(-2) of the TM variant to produce the variant NK5A. Alanines were used in this insert because they have one of the simplest side chains and on the Ramachandran plot most amino acids are found to be similar

to alanine. Lysine was added to this insert to maintain solubility and prevent aggregation. The insertion is in a disordered region of the protein, as judged from the high thermal factors for residues -5 to -1 in the crystal structure of yeast iso-1-cytochrome *c*.³⁵ Therefore, mutations in this region are not expected to have significant effects on the overall structure or stability of the native fold. The inserted lysine in this cassette was then mutated to a histidine residue to produce the NH5A variant which is capable of forming a 22 residue His-heme loop in the denatured state (the loop begins with Cys 14, which forms a thioether bond to a heme vinyl substituent and ends at the engineered histidine which binds to the Fe³⁺ in the heme). Figure 1 is a schematic representation of how this loop would form from the N-terminal side of the heme.

To test the effect of chain flexibility, the alanines in the NH5A insert were progressively replaced with glycines to produce variants Gly1 to Gly5 as depicted in Figure 1. Sequential replacement of Ala with Gly could have been accomplished in many ways. We arbitrarily chose to replace Ala with Gly outward from the center of the five Ala insert. We note that there are no prolines in any of the His-heme loops which could modulate loop formation through *cis/trans* proline isomerization.^{26,27} The loop also contains two additional glycines (Gly 1 and 6) and two β -substituted amino acids (Thr 8 and 12). Our previous work indicates that changes in β -substituted amino acid content have modest effects on loop equilibria.³³ Including glycines 1 and 6, the overall glycine content in the 22 residue loop goes from 9% to 32% as the NH5A variant is progressively converted to the Gly5 variant, substantially increasing the flexibility of the polypeptide segment involved in loop formation.

Stability of Variants

The global stability of all variants was measured by guanidine hydrochloride (gdnHCl) denaturation monitored by circular dichroism spectroscopy (see Supplementary Data). Immediately obvious from Table 1 is the fact that all the mutants capable of forming the 22 residue His-heme loop in the denatured state have approximately half the free energy of unfolding, $\Delta G_u^0(\text{H}_2\text{O})$, of the TM and NK5A variants. Coupled with the denatured state loop formation results discussed later, this difference in $\Delta G_u^0(\text{H}_2\text{O})$ strongly suggests that the His-heme loop formed under denaturing conditions stabilizes the denatured state relative to the fully folded form (see below). No significant differences in the stabilities of the NH5A and Gly1 to Gly5 variants are observed. The denaturation midpoints, C_m , are all less than 0.5 which indicates that these proteins are fully unfolded in 3 M gdnHCl, the conditions for the pH titrations used to determine denatured state loop stability.

Effects of sequence composition on loop formation in the denatured state

One of the advantages of using His-heme ligation as a probe is that a direct correlation can be made between the apparent pK_a , $pK_a(\text{obs})$, of titrating the histidine off the heme and the strength of that ligand interaction. Typically, the apparent pK_a is lower than the pK_a of free histidine because deprotonation of histidine is coupled to favorable formation of a bond between Fe³⁺ in the heme and the histidine imidazole side chain. The stability of the His-heme bond is dependent on three main factors. First, chain stiffness can decrease the stability if the polypeptide chain forming the loop is too short or too sterically restrictive, thereby hindering effective His-heme molecular orbital overlap.^{32,33} These hindrances can be relieved to increase the bond strength by either increasing loop size or decreasing residue bulk. Second, loop entropy decreases the His-heme bond strength as the loop becomes larger. Too long a loop increases the conformational space for the histidine to search, thereby decreasing the probability of making an actual His-heme bond. Third, residual structure formation increases the His-heme bond strength due to the presence of additional stabilizing contributions in the loop, if the structure is induced by loop formation. The summed stability of these contributors

is conferred onto the His-heme bond. Lower pK_a 's for the His-heme ligation are a direct indication that the interaction is stronger and thus requires more acid to break the bond by protonating the imidazole ring of the histidine. Using this probe of relative denatured state loop stability for the NH5A and GlyX (where X = 1 to 5) variants gives us direct information on the effect that sequence composition and thus main chain sterics has on loop formation in the denatured state.

Loop breakage is detected at 398 nm which is the wavelength to which the low spin Fe^{3+} -heme Soret band shifts when a high spin Fe^{3+} -heme forms due to the weak field water ligand being bound to the Fe^{3+} instead of the previously bound strong field histidine ligand (See Fig. 1). A typical pH titration in 3 M gdnHCl is shown in Figure 2 along with a plot of spectra as a function of pH. For the mutants in this series of experiments, the pH range of 7 to 3 was more than adequate to define the apparent pK_a . In particular, upper and lower baselines are well-formed allowing for straightforward extraction of $pK_a(obs)$ from fits of the data to the Henderson-Hasselbalch equation. The inset in Figure 2 shows spectra as a function of pH for the titration in Figure 2 and is typical for titrations with these variants. There are two isosbestic points at 403 nm (above pH 3.5) and 398 nm (below pH 3.5). This second isosbestic point is possibly due to His 18 titrating off the heme. Since all experiments monitor loop breakage at 398 nm, the apparent pK_a 's for loop breakage obtained at 398 nm should be insensitive to breakage of the His 18-heme bond. Table 2 summarizes the $pK_a(obs)$ results for the six different His-heme loop forming variants, as well as for the NK5A and TM variants. This table also shows the experimental values of n , the number of protons involved in loop breakage. All variants yield $n \sim 1$, which is in agreement with the scheme for His-heme loop equilibria in Figure 1.

Previous experiments done in our lab with the TM variant gave $pK_a(obs) = 5.94$ in 3 M gdnHCl for N-terminal amino group ligation to the heme.^{30,31} The NK5A variant yields a $pK_a(obs)$ of 6.22 in 3 M gdnHCl, an increase of ~ 0.3 units relative to the TM variant (Table 2). This increase is consistent with the N-terminal amino group being 6 amino acid residues further from the heme.

Replacing the lysine in NK5A with a histidine (NH5A) causes a drastic drop in the $pK_a(obs)$ to ~ 4.6 , consistent with the high affinity of histidines for the heme observed for histidines on the N-terminal side of the site of heme attachment in iso-1-cytochrome *c*.³³ Thus, interference from the N-terminal amino group in assessing $pK_a(obs)$ for histidine-heme ligation with the variants studied here should be minor. Another comparison in this regard is instructive. Previous studies gave $pK_a(obs) = 4.20 \pm 0.04$ for the variant AcH(-2).³³ This variant has an acetylated N-terminus and a histidine at position -2. The H(-2)I52 variant also has a histidine at position -2 (I52 indicates a stabilizing N52I mutation).³⁴ The H(-2)I52 variant is expressed in bacteria and thus its N-terminus is not acetylated. Its $pK_a(obs)$ reported here is 4.22 ± 0.09 . The fact that these variants' $pK_a(obs)$ values are identical within error demonstrates that the N-terminal amino group does not perturb $pK_a(obs)$ values for His-heme ligation from the N-terminal side of the heme.

The $pK_a(obs)$ values in Table 2 also show that His-heme ligation is stronger for the 16 residue loop formed by the H(-2)I52 variant than for the longer 22 residue loop formed by the NH5A variant. The increase in the $pK_a(obs)$ of ~ 0.4 units for His-heme ligation in 3 M gdnHCl caused by the AAAA⁵AK insert is similar to the ~ 0.3 unit increase for N-terminal amino group-heme ligation for the NK5A versus the TM variant. Thus, loop affinity scales similarly with loop size for the two types of denatured state loop formation.

The $pK_a(obs)$ results for the GlyX variants in Table 2 are surprising. The expected trend of progressively decreasing $pK_a(obs)$ values with increasing chain flexibility was not observed. The initial data set collected showed a decrease in $pK_a(obs)$ values – though subtle – for the

series of variants, NH5A, Gly1 and Gly2. However, instead of continuing this trend, the $pK_a(\text{obs})$ values increased for the series of variants, Gly3, Gly4 and Gly5. The differences in the $pK_a(\text{obs})$ values were very small. Therefore, to evaluate the error in the data due to calibrating the pH meter, a second independent data set of $pK_a(\text{obs})$ values was collected (Table 2). Figure 3 shows the plot of both data sets and even though the differences are small, the trend is still the same. There is a subtle decrease in the $pK_a(\text{obs})$ reaching a minimum for the Gly2 or Gly3 variant. This result is in accord with the chain being more flexible and thus more compact upon addition of less sterically restrictive glycines versus alanines. At higher glycine content there is a slight increase in the $pK_a(\text{obs})$ values. The greater flexibility presumably leads to a greater loss in conformational space when the loop forms for inserts with three or more of the 5 alanines replaced with glycine. These observations are in accord with the results on linker composition in Arc repressor by Robinson & Sauer.²⁸ In that work it was shown that an adequate percentage of glycine residues in the linker was very useful in bringing the two subunits together. However, having too many glycines in the linker was counterproductive.

Perhaps what is most surprising is the observation that the $pK_a(\text{obs})$ is identical for the NH5A and the Gly5 variants. The increase in overall glycine content from 9% (NH5A) to 32% (Gly5) in the 22 residues between the histidine which forms the loop and the heme should significantly decrease the average end-to-end separation of the histidine and the heme. The Jacobson-Stockmayer equation^{39,40} (eqn. 1) is typically used to predict loop formation equilibria,

$$\Delta S_{\text{loop}} = -v_3 R \ln(n) + R \ln \left[\left(\frac{3}{2\pi C_n l^2} \right)^{v_3} V_i \right] \quad (1)$$

where n is the number of monomers in the loop, R is the gas constant, C_n is Flory's characteristic ratio, which is sensitive to chain flexibility,⁴¹ l is the distance between monomers, V_i is the approach volume of the atoms involved in loop formation and v_3 is the scaling exponent for loop formation. Using the Jacobson-Stockmayer equation it can be shown that $\Delta pK_a(\text{obs})$ for the Gly5 variant versus the NH5A variant is given by eqn. 2. If we use $v_3 = 2.1$ (average of theoretical

$$\Delta pK_a(\text{obs}) = pK_a(\text{obs})_{\text{Gly5}} - pK_a(\text{obs})_{\text{NH5A}} = \left(\frac{-v_3}{\ln(10)} \right) \left[\ln \left(\frac{C_{n,\text{NH5A}}}{C_{n,\text{Gly5}}} \right) \right] \quad (2)$$

values for a random coil with excluded volume),^{11,42} then set the Flory characteristic ratios for the NH5A variant (9% Gly, 22 residue loop) to 5.5 and for the Gly5 variant (32% Gly, 22 residue loop) to 3.5,⁴¹ the expected ΔpK_a would be -0.41. Our experimental results clearly are counter to this prediction, the question then is why.

The peptide systems used to monitor the kinetics of loop formation allow much more freedom in the stereochemistry of loop formation than is true for our histidine-heme loop formation system. Histidine-heme binding requires axial docking of the histidine ligand to the iron in the heme and this constraint would be expected to seriously affect formation of short loops. Since our set of variants forms loops from the N-terminal side of the heme (see Fig. 1),³³ the heme excluded volume effect, which causes a decrease in the favorability of loop formation for loops <35 residues formed from the C-terminal side of the heme,³² is not a factor. Also, the favorability of loop formation is higher for the 16 residue versus the 22 residue N-terminal histidine-heme loop (Table 2), indicating that chain stiffness is not a dominant factor at this loop length. Thus, for the set of 22 residue loops studied here heme excluded volume and chain stiffness are unlikely to influence the observed result. The clustering of glycines near the histidine which forms the loop could also influence loop formation. However, the observations of Robinson & Sauer²⁸ argue against this possibility. Their work shows that the overall %Gly

composition of the linker in single chain Arc repressors, and not its specific sequence, is the dominant factor controlling stability and folding kinetics.

The assumption implicit in the Jacobson-Stockmayer treatment of loop formation is that the entropy decrease can be approximated by the required volume reduction to form a loop given a Gaussian distribution of conformations (see eqn. 1). Thus, a more compact chain should form loops more readily. However, this approximation may not account well for the larger density of conformations in a more flexible polypeptide chain with high glycine content. Comparison of rates of loop formation for polyserine versus poly(GlySer) peptides show that the glycine containing sequences have rates of loop formation that are faster by a factor of 2 for loop sizes from 3 to 12.⁴³ The $\log(k)$ versus $\log(n)$ curves for both are decreasing linearly by $n = 12$, and while data are only available for poly(GlySer) for larger loop sizes the factor of 2 difference looks to persist to larger loop sizes based on extrapolation of the polyserine data. Krieger et al.⁴³ noted that the enhanced rate of loop formation for poly(GlySer) sequences is smaller than might be expected given the greater flexibility and shorter end-to-end distance expected for poly(GlySer) versus polyserine peptides, likely as a result of the greater conformational space available to glycine containing sequences. In a more recent study,²⁷ a single glycine was embedded into polyserine sequences capable of forming loops of 4 to 12 residues progressively dropping glycine content from ~25% to ~8%. Rates of loop formation for this set of peptides move smoothly from the faster poly(GlySer) to the slower polyserine rate versus loop length curve. Thus, we expect that rates of loop formation should increase smoothly going from the NH5A (9% Gly) to the Gly5 (32% Gly) variant. The invariance in the equilibrium $pK_a(\text{obs})$ values from Table 2 suggests that the rate of loop breakage increases in nearly equal measure with the rate of loop formation. Thus, the greater flexibility speeds loop breakage, as well, which presumably reflects the greater number of ways the histidine can move away from the heme after the Fe^{3+} -histidine bond breaks.

Recent studies on polyglycine sequences using residual dipolar NMR and small-angle X-ray scattering methods have provided evidence that polyglycine segments prefer elongated conformations that are in fact more extended than polyalanine.⁴⁴ Stiffening of the backbone by hydration provides a possible explanation. The low sensitivity of $pK_a(\text{obs})$ to replacement of alanine with glycine could also result from polyglycine sequences being more elongated than previously thought. Kinetics experiments on loop formation and breakage are currently underway with these NH5A and GlyX variants to determine whether the insensitivity of $pK_a(\text{obs})$ to %Gly in the insert is due to compensating effects of both faster loop formation and breakage as opposed to polyglycine sequences being more elongated in aqueous solution than previously thought.

Effects of denatured state loop formation on overall protein stability

It is becoming increasingly evident that the denatured state of a protein can have an important impact on protein stability.^{8,9,11–18,45} In previous work, we have noted correlations between denatured state loop stability and global protein stability.^{32,33,37} All variants capable of forming a His-heme loop in the denatured state are strongly destabilized relative to the NK5A variant (Table 1) which forms the much less favorable N-terminal amino group-heme loop in the denatured state. How much of this destabilization can be accounted for by differences in denatured state loop stability? The stabilities of the NH5A and GlyX variants are all within error the same, with an average stability of 1.98 kcal/mol, which is 2.23 kcal/mol less stable than the NK5A variant (Table 1). At pH 7, where global stability is measured, both types of loops are fully or almost fully formed. Thus, the difference in denatured state stability due to loop formation can be evaluated as $\ln(10)RT\Delta pK_a(\text{obs})$. The average $pK_a(\text{obs})$ for the NH5A and GlyX variants is 4.56 (Table 2). Thus, relative to the less stable loop formed by the NK5A variant, the stabilization of the denatured state for the NH5A and GlyX variants is 2.24 kcal/

mol. Thus, denatured state loop stability completely accounts for the decrease in $\Delta G_u^0(\text{H}_2\text{O})$ for the NH5A and GlyX variants relative to the NK5A variant. While denatured state His-heme loop formation is specific to heme proteins, denatured state metal ligation may be a general modulator of metalloprotein stability that must be considered.

Conclusion

Experimentation on protein denatured states can still be considered a “frontier” area of research. There are many factors that influence protein folding, some of which are still not well-understood. Characterization of the stabilizing interactions and structures in the denatured state are considerably more difficult to assign than in the native state. The results of this paper clearly indicate that side chain sterics has a modest effect on equilibrium loop formation for the set of 22 residue loops studied here. Surprisingly the all glycine insert (Gly5) has essentially the same $pK_a(\text{obs})$ as the all alanine insert (NH5A). This observation may result from polyglycine being more elongated than previously suspected. Alternately, high glycine content and the accompanying increased flexibility may speed up the rate of loop formation but at the expense of persistence. A number of studies suggest that rapid conformational equilibration occurs during protein folding.^{46–48} Thus, both speed of contact formation and the persistence of a contact may be important in selecting which structures form during folding. Kinetics investigations are underway to provide more insight into this issue.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by NIH Grant GM074750 (B.E.B.).

References

1. Tanford C. Protein denaturation. *Adv Protein Chem* 1968;23:121–282. [PubMed: 4882248]
2. Fleming, PJ.; Rose, GD. Conformational properties of unfolded proteins. In: Buchner, J.; Kiefhaber, T., editors. *Protein Folding Handbook: Part I*. Wiley-VCH; Weinheim, Germany: 2005. p. 710-736.
3. Garcia P, Serrano L, Durand D, Rico M, Bruix M. NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: Implications for protein folding initiation. *Protein Sci* 2001;10:1100–1112. [PubMed: 11369848]
4. 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 non-native protein. *Science* 2002;295:1719–1722. [PubMed: 11872841]
5. Bond CJ, Wong K, Clarke J, Fersht AR, Daggett V. Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway. *Proc Natl Acad Sci USA* 1997;94:13409–13413. [PubMed: 9391038]
6. Shortle D, Ackerman MS. Persistence of native-like topology in a denatured protein in 8 M urea. *Science* 2001;293:487–489. [PubMed: 11463915]
7. Kazmirski SL, Wong K, Freund MV, Tan Y, Fersht AR, Daggett V. Protein folding from a highly disordered denatured state: The folding pathway of chymotrypsin 2 inhibitor at atomic resolution. *Proc Natl Acad Sci USA* 2001;98:4349–4354. [PubMed: 11274353]
8. Dill K, Shortle D. Denatured states of proteins. *Annu Rev Biochem* 1991;60:795–825. [PubMed: 1883209]
9. Shortle D. Structural analysis of non-native states of proteins by NMR methods. *Curr Opin Struct Biol* 1996;6:24–30. [PubMed: 8696969]
10. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiagarajan P, Sosnick TR, Hasan MZ, Pande VS, 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]
11. Bowler BE. Thermodynamics of protein denatured states. *Mol BioSyst* 2007;3:88–99. [PubMed: 17245488]
  12. Cho J, Sato S, Raleigh D. Thermodynamics and kinetics 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]
  13. Cho J, Raleigh D. 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]
  14. Anil B, Craig-Shapiro R, Raleigh DP. Design of a hyperstable protein by rational consideration of unfolded state interactions. *J Am Chem Soc* 2006;128:3144–3145. [PubMed: 16522085]
  15. Whitten ST, Garcia-Moreno EB. pH dependence of the stability of staphylococcal nuclease: Evidence of substantial electrostatic interactions in the denatured state. *Biochemistry* 2000;39:14292–14304. [PubMed: 11087378]
  16. Grimsley GR, Shaw KL, Fee LR, Alston RW, Huyghues-Despointes BM, Thurlkill RL, Scholtz JM, Pace CN. Increasing protein stability by altering long-range coulombic interactions. *Protein Sci* 1999;8:1843–1849. [PubMed: 10493585]
  17. 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]
  18. 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]
  19. Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat Struct Biol* 1997;4:10–19. [PubMed: 8989315]
  20. Eaton WA, Munoz V, Thompson PA, Henry ER, Hofrichter J. Kinetics and dynamics of loops, α -helices, β -hairpins, and fast folding proteins. *Acc Chem Res* 1998;31:745–753.
  21. Chan HS, Dill KA. Intrachain loops in polymers: Effects of excluded volume. *J Chem Phys* 1989;90:492–509.
  22. Weikl TR, Dill KA. Folding rates and low-entropy-loss routes of two-state proteins. *J Mol Biol* 2003;329:585–598. [PubMed: 12767836]
  23. Fersht AR. Transition-state structure as a unifying basis in protein folding mechanisms: Contact order, chain topology, stability and the extended nucleus mechanism. *Proc Natl Acad Sci USA* 2000;97:1525–1529. [PubMed: 10677494]
  24. Roder H, Elöve GA, Englander SW. Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and protein NMR. *Nature* 1988;335:700–704. [PubMed: 2845279]
  25. Krishna MMG, Englander SW. The N-terminal and C-terminal motif in protein folding and function. *Proc Natl Acad Sci USA* 2005;102:1053–1058. [PubMed: 15657118]
  26. Pierce MM, Nall BT. Coupled kinetic traps in cytochrome *c* folding: His-heme misligation and proline isomerization. *J Mol Biol* 2000;298:955–969. [PubMed: 10801361]
  27. 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]
  28. Robinson CR, Sauer RT. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc Natl Acad Sci USA* 1998;95:5929–5934. [PubMed: 9600894]
  29. Godbole S, Hammack B, Bowler BE. Measuring denatured state energetics: Deviations from random coil behavior and implications for the folding of iso-1-cytochrome *c*. *J Mol Biol* 2000;296:217–228. [PubMed: 10656828]
  30. 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]
  31. Hammack BN, 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]

32. 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]
33. 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]
34. 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]
35. Berghuis AM, Brayer GD. Oxidation state-dependent conformational changes in cytochrome *c*. *J Mol Biol* 1992;223:959–976. [PubMed: 1311391]
36. Deng WPD, Nickoloff JA. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal Biochem* 1992;200:81–88. [PubMed: 1595905]
37. Wandschneider E, Hammack BN, Bowler BE. Evaluation of cooperative interactions between substructures of iso-1-cytochrome *c* using double mutant cycles. *Biochemistry* 2003;42:10659–10666. [PubMed: 12962490]
38. Kristinsson R, Bowler BE. Communication of stabilizing energy between substructures of a protein. *Biochemistry* 2005;44:2349–2359. [PubMed: 15709747]
39. Cantor, CR.; Schimmel, PR. *Biophysical Chemistry, Part III: The Behavior of Biological Molecules*. W. H. Freeman and Co.; San Francisco: 1980. p. 1205-1209.
40. Flory, PJ. *Statistical Mechanics of Chain Molecules*. John Wiley & Sons, Inc.; New York: 1969. p. 392-396.
41. Miller WG, Brant DA, Flory PJ. Random coil configurations of polypeptide copolymers. *J Mol Biol* 1967;23:67–80.
42. Chan HS, Dill KA. The effect of internal constraints on the configurations of chain molecules. *J Phys Chem* 1990;92:3118–3135.
43. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T. Dynamics of unfolded polypeptide chains as models for the earliest steps in protein folding. *J Mol Biol* 2003;332:265–274. [PubMed: 12946363]
44. Onishi S, Kamikubo H, Onitsuka M, Kataoka M, Shortle D. Conformational preference of polyglycine in solution to elongated structure. *J Am Chem Soc* 2006;128:16338–16344. [PubMed: 17165789]
45. Anil B, Song B, Tang Y, Raleigh DP. Exploiting the right side of the Ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability. *J Am Chem Soc* 2004;126:13194–13195. [PubMed: 15479052]
46. Yeh S, Rousseau DL. Folding intermediates in cytochrome *c*. *Nat Struct Biol* 1998;5:222–228. [PubMed: 9501916]
47. Lee JC, Chang I, Gray HB, Winkler JR. The cytochrome *c* folding landscape revealed by electron-transfer kinetics. *J Mol Biol* 2002;320:159–164. [PubMed: 12079375]
48. Lyubovitsky JG, Gray HB, Winkler JR. Mapping the cytochrome *c* folding landscape. *J Am Chem Soc* 2002;124:5481–5485. [PubMed: 11996590]

Figure 1.

Schematic representation of denatured state His-heme ligation from the N-terminal side of the heme of iso-1-cytochrome *c*. Ala/Gly inserts made following the histidine inserted next to Phe (-3) for the NH5A and GlyX variants are shown. Lys(-2) terminates each insert sequence. Mutations were made using the unique restriction site elimination method³⁶ as previously described³⁷ (for details on protein expression and purification see Supplementary Data).

Figure 2.

A representative pH titration curve monitoring His-heme ligation in 3 M gdnHCl , 5 mM sodium phosphate, 15 mM NaCl . Data were acquired at room temperature, $22 \pm 1^\circ\text{C}$, using a Beckman DU 640 spectrophotometer following previously discussed titration procedures.^{29–33} Data were fitted to a modified form of the Henderson-Hasselbalch equation, allowing extraction of both the $\text{pK}_a(\text{obs})$ and the number of protons coupled to loop formation.^{29–33} Titrations were done in sets of three starting from pH 7 down to about pH 3 in intervals of approximately 0.2 pH units. The inset shows absorbance spectra from 350 to 450 nm as a function of pH over the pH range 7 to 3 (bottom to top at 398 nm) in approximate increments of 0.2 pH units.

Figure 3.

Graph of the two independent data sets showing pH titration reproducibility due to error in calibrating the pH meter. The errors for each $pK_a(\text{obs})$ are the standard deviation for three pH titrations done sequentially for each variant in each data set. The “%Gly in insert” axis indicates the percentage of the 5 alanines in the insert in the NH5A variant which has been replaced with glycines. Thus, the NH5A variant has 0 “%Gly in insert” and Gly5 variant has 100 “%Gly in insert”.

Table 1 Thermodynamic parameters for gdnHCl induced unfolding of iso-1-cytochrome *c* variants at pH 7 and 25 °C.

Iso-1-cytochrome <i>c</i> variant	$\Delta G_u^0(H_2O)$ (kcal/mol) ^a	<i>m</i> -value (kcal/(mol*M)) ^a	<i>C_m</i> (M) ^a
TM ^b	4.02 ± 0.17	3.99 ± 0.17	1.008 ± 0.0003
NK5A	4.21 ± 0.16	4.11 ± 0.01	1.03 ± 0.04
NH5A	2.00 ± 0.08	4.52 ± 0.18	0.44 ± 0.03
Gly1	1.87 ± 0.28	4.36 ± 0.63	0.43 ± 0.02
Gly2	2.05 ± 0.16	4.42 ± 0.19	0.46 ± 0.03
Gly3	1.97 ± 0.19	4.42 ± 0.67	0.45 ± 0.04
Gly4	2.04 ± 0.17	4.72 ± 0.29	0.43 ± 0.03
Gly5	1.92 ± 0.21	5.06 ± 0.84	0.38 ± 0.03

Average values obtained from circular dichroism experiments monitoring alpha helical content of variants at 222 nm. Ellipticity at 222 nm versus [gdnHCl] data were obtained with an Applied Photophysics PiStar 180 Spectrometer coupled to a Hamilton Microlab 500 titrator as previously described.³⁸ The buffer present during gdnHCl titrations was 20 mM Tris, 40 mM NaCl, pH 7. Ellipticity data at 250 nm, acquired simultaneously, were subtracted as background at each [gdnHCl]. Corrected data as a function of [gdnHCl] (see Supplementary Data) were fit to a linear free energy relationship to obtain $\Delta G_u^0(H_2O)$ and *m*, as described previously.³² Two-state folding is assumed in these fits.

^a Average values reported are based on 3 separate experiments and MALDI-TOF analysis was done before each run to assess the degree of fragmentation (see Supplementary Data). Fragmentation greater than 15% of the major peak required a restart of the entire experimental attempt (ie. starting from reisolation and repurification of the protein). The errors reported are standard deviations.

^b Data values obtained from Ref. 30.

Table 2

Thermodynamic parameters for denatured state loop formation at 22 °C in 3 M gdnHCl for variants of iso-1-cytochrome *c*.

Iso-1-cytochrome <i>c</i> Variant	$pK_a(obs)^a$		n-value, ^{ab}	
	Data Set 1	Data Set 2	Data Set 1	Data Set 2
TM ^c	5.94 ± 0.08		0.9 ± 0.1	
H(-2)I52	4.22 ± 0.09		1.11 ± 0.07	
NK5A ^d	6.22 ± 0.06		0.88 ± 0.12	
NH5A	4.64 ± 0.02	4.59 ± 0.04	1.06 ± 0.03	0.99 ± 0.05
Gly1	4.54 ± 0.02	4.56 ± 0.02	0.97 ± 0.02	1.02 ± 0.06
Gly2	4.45 ± 0.03	4.53 ± 0.02	1.03 ± 0.02	0.98 ± 0.04
Gly3	4.51 ± 0.02	4.53 ± 0.02	1.11 ± 0.12	0.99 ± 0.02
Gly4	4.58 ± 0.03	4.58 ± 0.01	1.08 ± 0.04	1.05 ± 0.01
Gly5	4.59 ± 0.02	4.60 ± 0.02	1.04 ± 0.03	1.13 ± 0.10

^a Average and standard deviation reported are based on 3 titration experiments carried out in sequence. MALDI-TOF analysis was done before each run to ensure minimal fragmentation (see Supplementary Data). Fragmentation greater than 15% of the m/z peak for the intact protein required a restart of the entire experiment (ie. reisolation and repurification of the protein). Parameters were obtained as described in the caption to Figure 2.

^b Proton transfer numbers (n-values) for each mutant are expected to be 1 (see Figure 1).

^c Data values obtained from Ref. 30.

^d This mutant does not show any fragmentation.