Published in final edited form as:

Biochemistry. 2017 December 26; 56(51): 6662–6676. doi:10.1021/acs.biochem.7b01002.

Effect of an Imposed Contact on Secondary Structure in the Denatured State of Yeast Iso-1-Cytochrome *c*

Travis A. Danielson^{†,||}, Jessica M. Stine^{†,||}, Tanveer A. Dar^{†,§}, Klara Briknarova^{†,‡}, and Bruce E. Bowler^{†,‡,*}

[†]Department of Chemistry and Biochemistry, University of Montana, Missoula, Montana 59812

[‡]Center for Biomolecular Structure and Dynamics, University of Montana, Missoula, Montana 59812

Abstract

There is considerable evidence that long-range interactions stabilize residual protein structure under denaturing conditions. However, evaluation of the effect of a specific contact on structure in the denatured state has been difficult. Iso-1-cytochrome *c* variants with a Lys54→His mutation form a particularly stable His-heme loop in the denatured state, suggestive of loop-induced residual structure. We have used multidimensional NMR methods to assign ¹H and ¹⁵N backbone amide and ¹³C backbone and side chain chemical shifts in the denatured state of iso-1-cytochrome *c* carrying the Lys54→His mutation in 3 and 6 M guanidine hydrochloride and at both pH 6.4, where the His54-heme loop is formed, and pH 3.6, where the His54-heme loop is broken. Using the secondary structure propensity score, with the 6 M guanidine hydrochloride chemical shift data as a random coil reference state for data collected in 3 M guanidine hydrochloride, we found residual helical structure in the denatured state for the 60's helix and the C-terminal helix, but not in the N-terminal helix in both the presence and absence of the His54-heme loop. Non-native helical structure is observed in two regions that form Ω-loops in the native state. There is more residual helical structure in the C-terminal helix at pH 6.4 when the loop is formed. Loop formation also appears to stabilize helical structure near His54, consistent with induction of helical structure observed when His-heme bonds form in heme-peptide model systems. The results are discussed in the context of the folding mechanism of cytochrome *c*.

TOC graphic

*Corresponding Author: bruce.bowler@umontana.edu.

Current address: Clinical Biochemistry, University of Kashmir, Hazratbal, Srinagar, Jammu & Kashmir, India, 190006

^{||}These authors contributed equally to this manuscript.

Supporting Information

Figure S1 shows secondary structure prediction by Agadir for the sequence of TM_T(-5)P/N52I/K54H iso-1-Cytc. Table S1 provides sequences of oligonucleotides used for site-directed mutagenesis. Tables S2 to S5 provide ¹H and ¹⁵N chemical shifts for the backbone amide NH and backbone and side chain ¹³C chemical shifts of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M and 6 M GdnHCl in the presence and absence of the His54-heme loop. Table S6 provides NMR acquisition and processing parameters.

ORCID

Bruce E. Bowler: 0000-0003-1543-2466

Notes

The authors declare no competing financial interest.

INTRODUCTION

The progression of events to a fully folded protein from a disordered state has been an area of considerable interest, and much progress has been made toward understanding this process.^{1, 2} The earliest events in protein folding likely involve the formation of simple loops. Considerable effort has focused on understanding the factors that control both the kinetics^{3–9} and thermodynamics^{10–18} of loop formation in disordered peptides and denatured proteins. While the bulk properties of protein denatured states are consistent with random coil behavior,^{19, 20} there is considerable evidence that significant residual structure can exist in the denatured states of proteins.^{21–25} Numerous NMR studies,^{26–54} studies on the thermodynamics of electrostatic interactions in the denatured state,^{55–62} Förster resonance energy transfer (FRET) studies,^{63–65} and some small angle X-ray scattering (SAXS)^{66–69} data support this conclusion. NMR studies of denatured proteins have been particularly fruitful. These studies have shown evidence for both native and non-native secondary structure in protein denatured states.^{22–26, 30, 35–42, 54, 70–73} NMR transverse relaxation experiments are sensitive to local chain stiffness providing evidence for hydrophobic clusters and residual secondary structure in protein denatured states as well as more flexible regions that may act as molecular hinges early in folding.^{38–42, 52, 71, 72, 74–78} Paramagnetic relaxation enhancement NMR experiments have further demonstrated long-range interactions in the denatured states of many proteins.^{43, 44, 46, 48, 49, 51, 53, 79–81} In a few cases, mutational analysis has demonstrated the importance of long-range contacts in stabilizing structure in the denatured state.^{76, 77, 82} However, it has been more difficult to study the effect of a specific well-defined contact on residual structure in the denatured state of a protein.

Two methods have been used extensively to impose defined contacts in the denatured state of a protein: engineered or naturally occurring disulfide bridges^{83–92} and histidine-heme loops.^{11, 14–18} In the case of disulfide bonds, the denaturant *m*-value for global unfolding is often decreased relative to the absence of the disulfide bond^{84, 87, 91} indicating that the disulfide bond induces a more compact denatured state.⁹³ Disulfide bonds have been used to probe folding pathways and can lead to remarkable acceleration of folding rates.^{87–89, 92} However, the effect on folding rates appears to depend on whether the disulfide bond encompasses a part of the protein which forms structure in the transition state.⁸⁶ In some cases, changes in the folding mechanism occur.^{89, 91} For a Cys85-Cys102 disulfide bond engineered into barnase, a significant dependence of disulfide loop stability in the denatured state on guanidine hydrochloride (GdnHCl) concentration was observed (*m*-value ~0.4 kcal/mol-M).⁸³ NMR data in this study were inconclusive regarding the nature of the residual

structure induced by disulfide loop formation. However, molecular dynamics simulations indicated the disulfide bond stabilized a natively like hairpin loop in the denatured state.⁸³ Notably, hydrophobic clusters observed in the denatured state of wild type barnase appeared to be disrupted by the Cys85-Cys102 disulfide bond.

For denatured state His-heme loop formation, GdnHCl *m*-values observed for the stability of the loops are typically near 0.2 kcal/mol-M for yeast iso-1-cytochrome *c* (iso-1-Cytc) and the four-helix bundle cytochrome *c*'.^{11, 16} These small *m*-values indicate that loop formation in the denatured state buries only a small amount of hydrophobic surface area for most single histidine variants of these proteins. However, in the case of a K54H variant of iso-1-Cytc, a relatively large GdnHCl *m*-value of 0.45 kcal/mol-M was observed¹⁶ indicating burial of a more significant amount of hydrophobic surface area upon His54-heme loop formation in the denatured state similar to that buried when the Cys85-Cys102 disulfide bond of barnase forms.⁸³ In addition, kinetic studies showed an unexpectedly slow rate of breakage of the very stable His54-heme loop, also consistent with stabilization of the closed loop by residual structure in the denatured state.¹⁵ While the His54-heme loop is a non-native contact, it presents an advantageous model system to characterize the effects of a specific well-defined long range interaction on a protein denatured state.

To investigate residual structure produced by the His54-heme loop formed by the K54H variant of iso-1-Cytc at residue level resolution, we assigned chemical shifts in this variant under denaturing conditions in the presence and absence of the His54-heme loop. We analyzed the observed chemical shifts relative to random coil shifts using the secondary structure propensity (SSP) score developed to assess residual structure in disordered proteins.⁹⁴ We observe significant residual secondary structure in the denatured state in 3 M GdnHCl in both the presence and absence of the His54-heme loop. As with barnase,⁸³ loop formation changes the distribution of residual structure.

EXPERIMENTAL PROCEDURES

Preparation of Iso-1-Cyc Variants

Preparation of variants started with the pRbs_BTR1 vector⁹⁵ carrying the coding sequence for a TM_T(-5)Q/N52I variant. TM stands for triple mutant and is shorthand for a variant with H26N, H33N and H39Q mutations, which eliminate His-heme misligation in the denatured state of iso-1-Cytc.^{18, 96} In addition to the T(-5)Q and N52I mutations, this variant also contains a C102S mutation to prevent disulfide dimerization during physical studies. The pRbs_BTR1 vector is a derivative of the pBTR1 vector,^{97, 98} which carries the iso-1-Cytc and heme lyase genes (*CYC1* and *CYC3*, respectively) from *Saccharomyces cerevisiae* allowing covalent attachment of heme to the CLQCH heme attachment sequence of iso-1-Cytc in the cytoplasm of *Escherichia coli*. The sequences of oligonucleotides used for mutagenesis are listed in Table S1 in the Supporting Information. Vector DNA for the expression of the TM_T(-5)P/N52I variant was prepared by the QuikChange protocol (Agilent Technologies) using the oligonucleotides Q(-5)P and Q(-5)P-r. In a similar manner, the vector for expression of the TM_T(-5)P/N52I/K54H variant was prepared by the QuikChange protocol using double-stranded pRbs_BTR1 DNA carrying the coding sequence for the TM_T(-5)P/N52I variant and the oligonucleotides K54H(I52) and

K54H(I52)-r. The entire coding regions of the TM_T(-5)P/N52I and TM_T(-5)P/N52I/K54H variants of iso-1-Cytc were verified by dideoxy sequencing (Murdock DNA Sequencing Facility, University of Montana).

Expression of Isotopically Labeled Iso-1-Cytc

Competent BL21-DE3 *E. coli* cells (EdgeBio) were transformed with the pRbs_BTR1 vector carrying the coding sequence of the TM_T(-5)P/N52I/K54H variant and plated on LB Agar plates containing 100 µg/mL ampicillin. Colonies on a plate were suspended in 3 mL of sterile LB medium, and the suspension was used to inoculate 1 L of minimal medium containing 100 µg/mL ampicillin. The 1 L of minimal medium consisted of the following components: 10 mL of 0.1 g/mL ¹⁵NH₄Cl, 1 mL of 1 M MgCl₂, 100 µL of 1 M CaCl₂, 1 mg thiamine, 5 mg FeSO₄•7H₂O, 1 mg biotin, 4 g ¹³C glucose, 16.8 mg δ-aminolevulinic acid, and 100 mL of a 10 × M9 salts solution at pH 7.4. The water, the M9 salts solution, and MgCl₂ solution were each autoclaved. The rest of the components were filter sterilized. The culture in minimal medium was grown at 37 °C with shaking at 125 rpm for 24 hrs. Cells were harvested using a Sorvall GS-3 rotor in a Sorvall RC-5C Plus floor model centrifuge running at 5000 rpm for 10 minutes. The supernatant was discarded and the resulting cell pellet was stored at -80 °C.

Expression of Iso-1-Cytc in Rich Media

Iso-1-Cytc variants were also expressed in Terrific Broth that contained 100 µg/mL ampicillin. The rest of the growth procedure was the same as for the preparation of isotopically labeled protein in the minimal medium.

Lysis and Crude Purification of Iso-1-Cytc

The cell pellet was removed from the -80 °C freezer and allowed to thaw in a 4 °C cold box and then returned to the -80 °C freezer. Two more rounds of freeze/thaw were performed. On the final thaw, the cells were resuspended in Lysis Buffer (100 mM Tris, pH 8.0, 500 mM NaCl, 1 mM EDTA, 2 mM PMSF plus a few crystals of DNase I and RNase A). Lysis buffer was used at 2 mL/g of cells to be lysed. The cell slurry was run through a French Press (SLM Aminco) 8-10 times to break the cells. This cell slurry was centrifuged twice, using a GSA rotor in a Sorvall RC-5C Plus floor model centrifuge at 10,000 rpm for 30 minutes, keeping the supernatant. Undesired proteins were precipitated by bringing the solution to 50% saturation with ammonium sulfate (31.4 g ammonium sulfate per 100 mL of lysate) and stirring 4-5 hours or overnight at 4 °C. The solution was centrifuged at 10,000 rpm for 30 minutes, discarding the pellet. The supernatant was transferred to 3500 molecular weight cut-off (MWCO) dialysis tubing and placed in 12.5 mM sodium phosphate, pH 7.2, 1 mM Na₂EDTA, 2 mM β-mercaptoethanol (β-ME) dialysis buffer and stirred gently overnight at 4 °C. The dialysis tubing was transferred to a second bottle of the dialysis buffer and stirred gently for 24 hours. The iso-1-Cytc solution was removed from the dialysis bags and centrifuged at 10,000 rpm for 30 minutes to remove any precipitate.

Purification of Iso-1-Cyt c by CM-Sepharose

Approximately 100 mL of CM-Sepharose resin was equilibrated in CM buffer A (50 mM sodium phosphate, pH 7.2, 1 mM EDTA, and 2 mM β -ME), and the dialyzed iso-1-Cyt c solution was added to the resin. After an hour of stirring at 4 °C, the resin was allowed to settle and the buffer was decanted off. The resin, now pink with iso-1-Cyt c protein, was resuspended in a small amount of CM Buffer A and packed into a column. Approximately 300-400 mL of CM Buffer A was run through the column as a wash. A 200 mL linear 0-0.8 M NaCl gradient was used to elute the protein from the CM-Sepharose by mixing CM Buffer A and CM Buffer B (CM Buffer A plus 0.8 M NaCl) in a gradient former. The eluted protein was concentrated by ultrafiltration and exchanged into HPLC Buffer A (50 mM sodium phosphate, pH 7.0) using Millipore Centricon and Centriprep 10,000 MWCO centrifugal filters and a benchtop Jouan CR4i centrifuge.

Purification of Iso-1-Cyt c by HPLC

Unlabeled protein was reduced with a small amount of sodium dithionite prior to HPLC, whereas ^{13}C and ^{15}N -labeled protein was oxidized with 5 mg of potassium ferricyanide per mg of crude protein prior to HPLC. Protein was loaded onto a BioRad UnoS6 cation exchange column and eluted with a gradient formed by mixing HPLC buffer A and HPLC Buffer B (50 mM sodium phosphate, 1 M NaCl, pH 7.0). The gradient used was: 0-7 minutes 0% B, 7-34 minutes increase linearly to 30% B, 34-40 minutes hold at 30% B, 40-43 minutes increase linearly to 100% B, 43-50 minutes hold at 100% B, 50-53 minutes decrease linearly to 0% B, 53-63 minutes hold at 0% B; flow rate of 3 mL/minute. The largest peak was collected and exchanged back to HPLC Buffer A and concentrated using the above mentioned Millipore centrifugal filters. The protein concentration was determined spectroscopically using a Beckman Coulter DU 800 Spectrophotometer at the isosbestic points 339, 526.5, and 541.75 nm of the iso-1-Cyt c spectrum as well as with the oxidized-state extinction coefficient at 550 nm.^{99, 100}

For the ^{13}C and ^{15}N -labeled protein, a mass spectrum was obtained with a Bruker microflex MALDI-TOF mass spectrometer to check the efficiency of labelling. The observed peak near 13252 Da is close to the expected mass of 13252.1 Da for uniformly ^{13}C - and ^{15}N -labeled protein.

Global Thermodynamic Stability Measurements

To measure global stability, GdnHCl denaturation was employed and monitored by circular dichroism (CD) spectroscopy. Protein in HPLC buffer A was oxidized with ferricyanide and then separated from ferricyanide and exchanged into CD Buffer (20 mM Tris, 40 mM NaCl, 1.0 mM NaEDTA, pH 7.0) by passing the protein solution through a Sephadex G-25 column equilibrated with CD Buffer. Concentration of the eluted protein was determined as described above. A volume of 850 μL of 4 μM native protein in CD Buffer was prepared and placed in a 10 $\times$ 4 mm quartz suprasil cell with space for a stir bar (Hellma part #: 119004F-10-40). A stock solution of 4 μM protein in approximately 6 M GdnHCl buffered with CD buffer was prepared from a 7.5 to 8.0 M GdnHCl solution buffered with CD buffer and the eluted protein stock in CD buffer. The GdnHCl titration was performed by titrating the 6 M GdnHCl stock solution into the solution in the cuvette with a Hamilton Microlab

500 titrator interfaced with an Applied Photophysics Chirascan CD spectrometer. A small stir bar was present in the cuvette and would stir at approximately 200 rpm for 100 seconds after addition of new denaturant stock to make certain mixing and equilibration were complete. The sample was allowed to settle without stirring for 10 seconds followed by data acquisition. Ellipticity was monitored at 222 nm, θ_{222} , diagnostic of α -helix, and 250 nm, θ_{250} , used as background, with 25 seconds of signal averaging at each wavelength. All data were acquired at 25 °C. $\theta_{222\text{corr}} (= \theta_{222} - \theta_{250})$ was plotted against GdnHCl concentration and fit to eq. 1,

$$\theta_{222\text{corr}} = \frac{\theta_N + \left[(\theta_D + m_D [\text{GdnHCl}]) \exp \left\{ \frac{m [\text{GdnHCl}] - \Delta G_u^{\circ'}(\text{H}_2\text{O})}{RT} \right\} \right]}{1 + \exp \left\{ \frac{m [\text{GdnHCl}] - \Delta G_u^{\circ'}(\text{H}_2\text{O})}{RT} \right\}} \quad (1)$$

where θ_N and θ_D are the ellipticities of the native and denatured protein, respectively, at 0 M GdnHCl, m_D is the denaturant dependence of the ellipticity of the denatured state, m is the GdnHCl concentration dependence of G_u , the free energy of unfolding, and $G_u^{\circ'}(\text{H}_2\text{O})$ is the free energy of unfolding extrapolated to 0 M GdnHCl. Reported parameters are the average and standard deviation of 3 to 4 separate titrations.

Denatured State Loop Formation

His-heme loop stability was measured in the denatured state (3 M GdnHCl, 5 mM Na_2HPO_4 and 15 mM NaCl) by pH titration. Loop equilibria were monitored at the heme Soret band (398 nm). Titrations were carried out at room temperature (22 ± 1 °C). A $3 \times \text{protein } 3 \times \text{buffer}$ stock solution (9 μM protein, 15 mM Na_2HPO_4 , and 45 mM NaCl) was made. The refractive index of an approximately 6 M GdnHCl solution was measured using water as background to determine the exact GdnHCl concentration.¹⁰¹

A 3 mL sample that contained 3 μM protein, $1 \times \text{buffer}$ and 3 M GdnHCl was prepared by combining appropriate amounts of $3 \times \text{protein } 3 \times \text{buffer}$ stock, 6 M GdnHCl solution, and water. The pH was measured with a Denver Instruments UB-10 pH meter. 60 μL of the sample were then transferred into a Beckman Coulter microcell cuvette, and an absorbance spectrum was acquired from 350 to 450 nm using a Beckman Coulter DU 800 spectrophotometer. To change the pH without changing the protein or GdnHCl concentration of the sample, $3 \times \text{protein } 3 \times \text{buffer}$ stock (20 μL), ~ 6 M GdnHCl (~ 30 μL), and HCl (~ 10 μL) in appropriate ratios, for a total volume of 60 μL , were added back to the sample and thoroughly mixed by pipetting. The volume of HCl added was constant, but the concentration of the acid was varied such that the pH of the sample decreased by ~ 0.2 pH units with each addition. After each pH adjustment, the pH and the absorption spectrum of the sample were measured as described above. At least 3 pH titrations were completed for each variant. The absorbance at 450 nm was used as a baseline and subtracted from the absorbance values at 350 to 450 nm. The adjusted absorbance at 398 nm, $A_{398\text{corr}} (= A_{398} - A_{450})$, was plotted versus pH and fit to a modified Henderson-Hasselbalch equation (eq. 2) to determine the apparent $\text{p}K_a$ of loop formation, $\text{p}K_a(\text{obs})$, and the number of protons, n_p , linked to the process.

$$A_{398\text{corr}} = \frac{A_{\text{LS}} + A_{\text{HS}} 10^{n\text{p}(pK_a(\text{obs}) - \text{pH})}}{1 + 10^{n\text{p}(pK_a(\text{obs}) - \text{pH})}} \quad (2)$$

In eq. 2, A_{LS} is $A_{398\text{corr}}$ when the loop is formed and the heme is in a low spin state, and A_{HS} is $A_{398\text{corr}}$ when the loop is broken and the heme is a high spin state, with water occupying the 6th coordination site.

Preparation of Iso-1-Cytc Samples for NMR Spectroscopy

For experiments in 3 M GdnHCl, 3 × protein 3 × buffer (~2100 μM oxidized ¹³C- and ¹⁵N-labeled TM_T(-5)P/N52I/K54H iso-1-Cytc, 15 mM Na₂HPO₄, and 45 mM NaCl) stock solution was made by passing the protein through a Sephadex G-25 column equilibrated in 3 × buffer (15 mM Na₂HPO₄, and 45 mM NaCl) followed by centrifugal ultrafiltration to adjust the protein concentration. This 3 × protein 3 × buffer stock was then combined with 6 M GdnHCl solution, D₂O and water to yield a sample containing ~700 μM protein, 1 × buffer (5 mM Na₂HPO₄, and 15 mM NaCl), 3 M GdnHCl and 10% D₂O. For experiments in 6 M GdnHCl, a 6 × protein 6 × buffer stock (~2400 μM oxidized ¹³C- and ¹⁵N-labeled TM_T(-5)P/N52I/K54H iso-1-Cytc, 30 mM Na₂HPO₄, and 90 mM NaCl), 8 M GdnHCl solution, D₂O and water were used to prepare a sample that contained ~400 μM protein, 1 × buffer, 6 M GdnHCl and 10% D₂O. The volume of the samples was 600 μL. The pH was measured with a Denver Instruments UB-10 pH meter. The procedure for changing the pH of the sample was very similar to the procedure described above for denatured state pH titrations. For example, to change the pH in the 3 M GdnHCl sample without changing the protein or GdnHCl concentration, a total of 12 μL was added to the sample: 3 × protein 3 × buffer stock (4 μL), ~6 M GdnHCl solution (~6 μL), D₂O (1.2 μL) and HCl (~.8 μL), and the sample was thoroughly mixed by pipetting. The volume of HCl added was constant, but the concentration of the acid was varied as needed to reach the desired pH value.

NMR Experiments

Chemical shifts of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl at pH 6.4 and pH 3.6 and in 6 M GdnHCl at pH 6.4 and pH 3.6 were assigned using 2D ¹H-¹⁵N HSQC¹⁰² and 3D HNCACB,¹⁰³ CBCA(CO)NH,¹⁰⁴ C(CO)NH,¹⁰⁵ HNCO,^{106, 107} and HN(CA)CO¹⁰⁸ spectra, which were acquired at all four solution conditions (Table S6). The experiments were performed at 25 °C using a 600 MHz Varian NMR System spectrometer equipped with a triple-resonance probe or a ¹³C-enhanced salt-tolerant cold probe. Standard BioPack (VnmrJ 2.2C software, Agilent) sensitivity enhanced pulse sequences that utilize gradients for coherence selection were employed. ¹H chemical shifts were referenced relative to internal DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), and ¹³C and ¹⁵N chemical shifts were referenced indirectly.^{109, 110} The data were processed with NMRPipe.¹¹¹ A squared cosine-bell window function and zero filling were applied in the ¹H and ¹³C dimensions, and linear prediction to double the number of data points, a squared cosine-bell window function and zero filling were applied in the ¹⁵N dimension prior to the respective Fourier transforms. Chemical shift assignments were made with the aid of Sparky (<http://www.cgl.ucsf.edu/home/sparky/>)¹¹² and CcpNmr Analysis version 2.2.2 software.¹¹³

Evaluation of Secondary Structure in the Denatured State

The $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts were used to estimate the secondary structure propensities in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc. The chemical shifts were re-referenced and the SSP score was calculated using eq. 3 as implemented in the SSP program, version 1.02.⁹⁴

$$SSP_i = \frac{\sum_{j=i-2}^{i+2} \sum_X \left\{ \begin{array}{l} \frac{|\Delta\delta X_j obs|}{\sigma_j \alpha} \text{ if } (\Delta\delta X_j obs)(\Delta\delta X_j \alpha) > 0 \\ - \frac{|\Delta\delta X_j obs|}{\sigma_j \beta} \text{ if } (\Delta\delta X_j obs)(\Delta\delta X_j \beta) > 0 \end{array} \right.}{\sum_{j=i-2}^{i+2} \sum_X \left\{ \begin{array}{l} \frac{|\Delta\delta X_j \alpha|}{\sigma_j \alpha} \text{ if } (\Delta\delta X_j obs)(\Delta\delta X_j \alpha) > 0 \\ + \frac{|\Delta\delta X_j \beta|}{\sigma_j \beta} \text{ if } (\Delta\delta X_j obs)(\Delta\delta X_j \beta) > 0 \end{array} \right.} \quad (3)$$

Eq. 3 yields the SSP score for a given residue i , where the summation over X is for all chemical shifts used, in our case $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$, $\delta X_j obs$ is the observed secondary chemical shift, $\delta X_j \alpha/\beta$ are the average secondary chemical shifts of fully formed secondary structure (α or β), and $\sigma_j \alpha/\beta$ are the standard deviations of $\delta X_j \alpha/\beta$. The random coil and secondary shifts and their standard deviations that are provided with the SSP program come from the RefDB database.¹¹⁴ A weighted average over a five-residue window is used to minimize the effects of chemical shifts that are poor measures of secondary structure, such as the chemical shifts of glycine residues. As the score reflects the fraction of secondary structure at a specific position, fully formed secondary structure would have an absolute value of 1. Regions of α and β structure yield positive and negative SSP scores, respectively.

RESULTS

Design of TM_T(-5)P/N52I/K54H Iso-1-Cytc

The goal of the current work is to determine the effect that formation of a simple loop has on the presence of ordered structure in an unfolded protein. In previous work, we have shown that a histidine at position 54 of iso-1-Cytc forms a particularly stable histidine-heme loop in the denatured state.^{15–18, 96} The denaturant dependence of the stability of the His54-heme loop in the denatured state is large compared to other histidine-heme loops suggesting induction of structure upon loop formation in the denatured state.^{16, 83} Thus, the His54-heme loop is well-suited as a model system for the effects of loop formation on residual structure in the denatured state.

Variants of iso-1-Cytc used to monitor histidine-heme loop formation in the denatured state contain a single histidine besides His18 from the CXXCH heme attachment motif. The histidine used for loop formation is engineered into a surface accessible site (Fig. 1). When iso-1-Cytc unfolds, the weak native state bond between Met80 and the heme iron breaks (Fig. 1). The engineered histidine forms a strong bond with the heme iron causing the formation of the denatured state histidine-heme loop. Thus, particularly stable histidine-heme loops can decrease the global stability of iso-1-Cytc by preferentially stabilizing the denatured state of iso-1-Cytc.¹¹⁵ The His54-heme loop strongly stabilizes the denatured

state of iso-1-Cytc.^{16–18} Thus, K54H variants of iso-1-Cytc are only marginally stable.^{17, 18} To ameliorate this problem, we use the mutation N52I, which is a global stabilizer of yeast iso-1-Cytc (see Fig. 1).¹¹⁶ Previous work shows that this mutation stabilizes the K54H variant by 2 kcal/mol,¹¹⁷ and should favor higher yields of isotopically-labeled protein for NMR experiments.

As in our previous work, the iso-1-Cytc variants used here are built on the TM variant (H26N, H33N and H39Q), in which all naturally occurring histidines are removed (except for His18, from the heme attachment sequence).¹⁸ Thus, only the engineered histidine (His54 in this work) can form a His-heme loop in the denatured state. Competition between the N-terminal amino group and histidine for heme binding introduces further complications.^{96, 118} When expressed in its native host, *S. cerevisiae*, the N-terminal amino group can be acetylated *in vivo* as a post-translational modification, preventing competition between the N-terminal amino group and histidine for heme binding in the denatured state.¹⁷ However, post-translational acetylation of the N-terminal amino group is not straightforward when producing isotopically-labeled protein in *E. coli*. The N-terminal α -amino group can be converted into an α -keto group via a transamination reaction^{118, 119} but the yield is too low for use with isotopically-labeled protein samples. Conversion of the terminal amino group to proline provides another strategy to reduce interference from the N-terminus of the protein. The pK_a of 10.47 ± 0.06 (ionic strength $\mu = 0.1$) for the secondary amine of proline^{120, 121} is 1.5 units higher than the pK_a of 8.97 ± 0.03 ($\mu = 0.1$) for α -amino group of threonine,¹²¹ the wild type residue at the N-terminus of iso-1-Cytc. Thus, N-terminal proline is expected to interact more weakly with the heme in the denatured state and interfere less with His-heme loop formation than a primary N-terminal α -amino group. TM_T(-5)P/N52I and TM_T(-5)P/N52I/K54H iso-1-Cytc variants were prepared for the work described here (see Fig. 1).

Stability of Iso-1-Cytc Variants

GdnHCl denaturation data for the TM_T(-5)P/N52I and TM_T(-5)P/N52I/K54H variants of iso-1-Cytc are shown in Figure 2. The TM_T(-5)P/N52I/K54H variant unfolds at considerably lower GdnHCl concentration compared to the TM_T(-5)P/N52I variant. At low concentrations of GdnHCl, we observe an increase in the negative ellipticity of iso-1-Cytc followed by a leveling off. This behavior is commonly observed for iso-1-Cytc.¹²² Previous work from the Tsong and Goto laboratories has shown that this effect is due to specific ion binding by Cl^- .^{123, 124} We typically exclude these data points from our data analysis and only fit the data after the native baseline has leveled off.^{125–128} Fits of the data in Figure 2 to a linear free energy relationship (eq. 1, Experimental Procedures) indicate that the K54H mutation destabilizes iso-1-Cytc by about 3 kcal/mol (Table 1). Comparison with GdnHCl denaturation data for the equivalent variants expressed in yeast with N-terminal acetylation, AcTMI52 and AcH54I52,¹¹⁷ shows that the *E. coli*-expressed variants with proline at the N-terminus have stabilities similar to their yeast-expressed counterparts.

In Figure 2, it is evident that the cooperative unfolding of the TM_T(-5)P/N52I/K54H variant is complete by 1.5 M GdnHCl. To be conservative, NMR studies were carried out at 3 M and

6 M GdnHCl to ensure that only residual structure in fully denatured TM_T(-5)P/N52I/K54H iso-1-Cytc is being monitored and not partially folded forms of the protein.

Loop Formation in the Denatured State

Met80 is a weak ligand for the heme of Cytc when the iron of the heme is in the ferric state.^{129, 130} When Cytc unfolds, Met80 (see Fig. 1) is replaced by histidines or lysines from the Cytc sequence,^{131–133} which have a higher affinity for ferric heme than methionine. For yeast iso-1-Cytc, we have engineered the protein so that a single histidine in the sequence can form a loop of defined size in the denatured state.^{16, 17} When pH is lowered, the histidine becomes protonated, and the loop is broken. Deprotonation of His54 and the coupled His-heme loop formation can be described by an apparent pK_a , $pK_a(\text{obs})$ (Fig. 3). The process is readily monitored at micromolar concentrations of iso-1-Cytc due to the shift in the intense Soret absorption band of the heme caused by the spin state change that results from replacing the strong field ligand, histidine, with the weak field ligand, water. Figure 4A shows pH titrations, which are used to evaluate loop formation equilibria in the denatured state, of the TM_T(-5)P/N52I and TM_T(-5)P/N52I/K54H variants of iso-1-Cytc in 3 M GdnHCl. The presence of His54 in TM_T(-5)P/N52I/K54H strongly shifts the $pK_a(\text{obs})$ to lower pH relative to TM_T(-5)P/N52I. Figure 4B shows spectra as a function of pH for TM_T(-5)P/N52I/K54H. A well-defined isosbestic point near 404 nm is observed from pH 7 to pH 3.7, consistent with two-state loop formation. Below this pH, the isosbestic point migrates to 398 nm, likely due to loss of His18 ligation. To minimize the effect of loss of His18 ligation on $pK_a(\text{obs})$ values, titrations are monitored at 398 nm.

The parameters from fitting the data in Figure 4 to eq. 2 (Experimental Procedures) are collected in Table 2. The proline secondary amino group shifts the $pK_a(\text{obs})$ upward by ~0.7 units relative to the primary α -amino group of the N-terminal threonine of the TM variant (Table 2). This shift is less than the ~1.5 unit shift expected based solely on the pK_a of the proline secondary amino group versus the primary α -amino group of threonine. Acetylating the N-terminus is more effective leading to $pK_a(\text{obs})$ values of 7.4 to 7.6.^{17, 118, 133} However, the shift is large enough to significantly decrease competition of the N-terminal group with His54 for binding to the heme in the denatured state. The similarity of the $pK_a(\text{obs})$ values for the TM_T(-5)P/N52I/K54H variant and the previously reported yeast-expressed Ach54I52 variant with an acetylated N-terminus (Table 2) is consistent with minimal competition of the proline secondary amino group with His54 binding in the denatured state of the TM_T(-5)P/N52I/K54H variant. Furthermore, $pK_a(\text{obs})$ for lysine binding to the heme of iso-1-Cytc in the denatured state shifts from 7.5 in the absence of a histidine to 9.0 in the presence of His54.¹³³ Thus, the proline secondary amino group should not displace His54 from the heme until significantly above pH 7 in denaturing solutions of GdnHCl. We note that the number of protons, n_p , linked to loop formation is ~1 in all cases (Table 2), as expected, because a histidine (or proline secondary amino group) must be deprotonated for the denatured state His-heme loop to form (Fig. 3).

Studies of the TM_T(-5)P/N52I/K54H Variant in the Denatured State by NMR Spectroscopy

The GdnHCl unfolding titration data for the TM_T(-5)P/N52I/K54H variant in Figure 2 show that it is fully in the denatured state above 1.5 M GdnHCl. The NMR studies described

here are carried out at 3 M and 6 M GdnHCl. Thus, the data report on the properties of the fully denatured state of the TM_T(-5)P/N52I/K54H variant at two different concentrations of GdnHCl.

From the denatured state pH titration data in Figure 4, it is evident that in 3 M GdnHCl the His54-heme loop of the TM_T(-5)P/N52I/K54H variant is essentially completely formed at pH 6.4 and almost completely broken at pH 3.6. Thus, NMR studies in the denatured state at pH 6.4 and pH 3.6 can provide for a comparison of secondary structure in the denatured state in the presence versus the absence of the His54-heme loop. It is important to evaluate quantitatively the degree to which the Pro(-5) secondary amino group competes with His54 for heme binding at pH 6.4 in the denatured state. The $pK_a(\text{obs})$ can be broken down into two terms, $pK_a(\text{LH}^+)$, which describes ionization of the ligand, and pK_{FeL} , which describes the affinity of the fully deprotonated ligand (His54 or Pro(-5) secondary amino group) for ferric heme, as given by eq. 4

$$pK_a(\text{obs}) = pK_a(\text{LH}^+) + pK_{\text{FeL}} \quad (4)$$

For the $pK_a(\text{LH}^+)$ of the secondary amino group of proline, pK_a values for the N-terminal amino group in short peptides provide an appropriate estimate. A pK_a value of 8.80 is reported for L-Pro-L-Leu-Gly-CONH₂ ($\mu = 0.1$, 25 °C).¹²¹ Using this value for $pK_a(\text{LH}^+)$, along with the $pK_a(\text{obs})$ in Table 2 for TM_T(-5)P/N52I, yields $pK_{\text{FeL}} = -2.13$ for N-terminal proline secondary amino group-heme binding in the denatured state in 3 M GdnHCl. Similarly, $pK_{\text{FeL}} = -2.04$ for His54-heme binding in the denatured state in 3 M GdnHCl using $pK_a(\text{obs})$ for TM_T(-5)P/N52I/K54H in Table 2 and $pK_a(\text{LH}^+) = 6.78$, an average value for histidine ionization in denaturing concentrations of GdnHCl extracted from His-heme loop formation data at eight different sequence positions from 3 – 6 M GdnHCl.¹¹ The pH dependence of the affinity of each ligand relative to ferric heme with water bound at low pH (Figure 3) is given by eq. 5.^{130, 133}

$$K_{\text{FeL}}(\text{H}^+) = \frac{K_{\text{FeL}}}{\left(1 + \frac{[\text{H}^+]}{K_a(\text{LH}^+)}\right)} \quad (5)$$

We can use eq. 5 to evaluate the partition function for heme ligand binding and the population of each heme-ligand pair¹³³ as a function of pH. For the denatured state of TM_T(-5)P/N52I/K54H in 3 M GdnHCl, we find that ~93% of TM_T(-5)P/N52I/K54H has water bound to the heme at pH 3.6, with the remainder having His54 as the ligand. At pH 6.4 in 3 M GdnHCl, ~96% of TM_T(-5)P/N52I/K54H has the His54-heme loop, whereas only ~3% forms the Pro(-5)-heme loop, with the remainder having water bound to the heme. Thus, interference from Pro(-5)-heme loop formation in the denatured state is minimal at pH 6.4 in 3 M GdnHCl.

The $pK_a(\text{obs})$ of AcH54I52 is 5.65 in 6 M GdnHCl,¹⁶ which leads to $pK_{\text{FeL}} = -1.13$. Thus, for the denatured state of TM_T(-5)P/N52I/K54H in 6 M GdnHCl, we estimate that ~99%

of TM_T(-5)P/N52I/K54H has water bound to the heme at pH 3.6, with the remainder having His54 as the ligand. At pH 6.4 in 6 M GdnHCl, ~72% of TM_T(-5)P/N52I/K54H has the His54-heme loop, whereas only ~10% forms the Pro(-5)-heme loop, leaving ~18% of the protein still having water bound to the heme. The strong GdnHCl dependence of the His54-heme loop lessens the proportion of the protein in the His54-heme conformer at pH 6.4, but it still remains the predominant form in solution.

We collected NMR data and assigned ^1H and ^{15}N backbone amide and ^{13}C backbone and side chain chemical shifts in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cyt c at pH 6.4 to report on residual secondary structure with the His54-heme loop in place and at pH 3.6 to report on residual secondary structure in the absence of this loop. Figure 5 shows 2D ^1H - ^{15}N HSQC spectra for the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cyt c at pH 6.4 and 3.6 in both 3 M and 6 M GdnHCl. The low chemical shift dispersion under all four denaturing conditions is typical for unfolded proteins. No partially unfolded forms are expected at 3 or 6 M GdnHCl based on the GdnHCl denaturation data in Figure 2. The crosspeaks in the ^1H - ^{15}N HSQC spectra are indeed sharp under all conditions, as expected for fully unfolded proteins. Another potential concern is aggregation. However, the sharp signals that we observe in the NMR spectra and the small-angle X-ray scattering (SAXS) data reported by the Doniach and Bilse laboratories^{68, 69} indicate that cytochrome c at millimolar concentrations does not aggregate in concentrated GdnHCl solutions. Also, the radius of gyration, R_g , of $30.3 \pm 0.1 \text{ \AA}$ ⁶⁸ observed in 3 – 5 M GdnHCl with SAXS measurements is in the range expected for an unfolded monomeric protein of the size of cytochrome c .¹³⁴

At pH 6.4, where the His54-heme loop is formed, backbone amide chemical shifts were assigned for 85 of 103 non-proline residues in 3 M GdnHCl and for 76 of 103 non-proline residues in 6 M GdnHCl (Tables S2 and S4 of the Supporting Information). Not surprisingly, assignments could not be made for residues located near the paramagnetic low-spin ferric heme. In 3 M GdnHCl, backbone amide chemical shifts of K11 through E21, which are adjacent to the CXXCH heme attachment site (C14 to H18), could not be assigned. In 6 M GdnHCl, the unassigned region adjacent to the heme attachment site is more extensive, with backbone amide chemical shifts missing for L9 to N26. The signals from amides whose chemical shifts could not be assigned in 6 M GdnHCl have low intensities in the HSQC spectrum in 3 M GdnHCl. The somewhat lower protein concentration in the 6 M GdnHCl samples is likely the primary reason assignments could not be made for residues 9 to 10 and 22 to 26. The other region where assignments could not be made was near His54, which is bound to the heme at pH 6.4. In 3 M GdnHCl, backbone amide chemical shifts were not assigned for D50 and I53 to N56. In 6 M GdnHCl, the region of incomplete assignments is again larger, with backbone amide chemical shifts missing for A51 to L58.

At pH 3.6, where the polypeptide chain of TM_T(-5)P/N52I/K54H iso-1-Cyt c is unconstrained, backbone amide chemical shifts were assigned for 83 of 103 non-proline residues in both 3 M and 6 M GdnHCl (Tables S3 and S5 of the Supporting Information). Backbone amide chemical shift assignments could not be made for residues T8 to N26 adjacent to the heme attachment site in 3 M GdnHCl and for L9 to V28 in 6 M GdnHCl. The

region with missing assignments near the heme attachment site is larger at pH 3.6 than at pH 6.4, possibly due to the paramagnetic ferric heme being high spin at low pH.

Interestingly, a signal from the backbone amide of F36 was not identified in the ^1H - ^{15}N HSQC spectra under any of the four denatured state conditions used for our NMR experiments. Signals from the neighboring residues including I35 and G37 were present in the HSQC spectra, although they are weak at pH 3.6 (Fig. 5). The correlations between the backbone amide of G37 and the $^{13}\text{C}'$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts of F36 were observable in the 3D triple resonance experiments under all conditions (Tables S2 – S5 of the Supporting Information). F36 is not near either the heme attachment site or His54.

DISCUSSION

SSP Score Analysis of the Denatured State of Iso-1-Cytc

To assess residual secondary structure in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc, we used the SSP score developed by Forman-Kay and coworkers with the re-referencing algorithm for $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts.⁹⁴ The re-referencing shifts are near -0.2 for all four experimental conditions. These shifts are somewhat large, given that all spectra were referenced to internal DSS,⁹⁴ and may reflect perturbations due to the paramagnetic heme. The SSP scores with the His54-heme loop formed at pH 6.4 in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc at 3 M and 6 M GdnHCl are shown in Figure 6A. Both positive (α -helix) and negative (β -sheet) SSP scores are observed. In general, the scores are more positive (or less negative) at 3 M versus 6 M GdnHCl. The profile of SSP scores is similar in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M and 6 M GdnHCl at pH 3.6 when the loop is broken (Figure 6B). There is evidence of residual helical structure corresponding to two of the three native α -helices of iso-1-Cytc, the 60's helix and the C-terminal helix, in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc both with and without the His54-heme loop formed. The magnitudes of the SSP scores, however, are large for a fully unfolded protein in high concentrations of denaturant⁷² and are more similar to the magnitudes of SSP scores observed for acid denatured proteins,^{45, 70–72} which typically retain more residual structure.

Correcting the SSP Score for the Paramagnetic Heme

A potential concern is the effect of the paramagnetic heme on chemical shifts for residues constrained to be near the heme in the denatured state either by the CLQCH heme attachment site or by the His54-heme bond at pH 6.4. To correct for this possibility, we used the SSP scores in 6 M GdnHCl as a reference state for the SSP scores for the denatured state in 3 M GdnHCl. Chemical shifts in 6 M GdnHCl have been used previously as a reference state to quantify residual secondary structure in the denatured state of apoflavodoxin under milder (3.4 M GdnHCl) denaturing conditions.¹³⁵ Using chemical shifts in 6 M GdnHCl also inherently corrects for the sequence dependence of random coil chemical shifts.¹³⁶ Because of the effects of GdnHCl on the stability of the His54-heme loop, this subtraction is not perfect at either pH, but should still eliminate the majority of the effects of the paramagnetic heme. Figure 7 shows SSP scores for the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl with 6 M GdnHCl as a reference state with the

His54-heme loop formed at pH 6.4 and in its absence at pH 3.6. The SSP scores in general fall between -0.1 and +0.1. Most SSP scores are positive indicative of residual helical structure on the order of 5-10% helix in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl at both pH 3.6 and 6.4.

Common Features of Residual Secondary Structure in the Denatured State in the Presence and Absence of the His54-heme loop

In general, the regions of residual helical structure in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl, in both the presence (pH 6.4) and the absence (pH 3.6) of the His54-heme loop, correspond to portions of iso-1-Cytc that are helical in the native state (Figure 7). However, there are also regions of non-native helical structure that are present in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl, in both the presence and the absence of the His54-heme loop. Regions of non-native secondary structure are common in denatured proteins.^{22, 26, 30, 35-42, 70-73}

There are three long helices in the native state of iso-1-Cytc. The presence or absence of residual structure in the denatured state of these sequence regions of TM_T(-5)P/N52I/K54H iso-1-Cytc is discussed first. The N-terminal helix has no consistent stretches of significant residual helical structure in the region between residues 2 and 7 that can be detected with a SSP score (Fig. 7). The large SSP scores at positions 6 and 7 (positive at pH 6.4 and negative at pH 3.6) may have considerable error because the NMR signals from these residues are very weak and thus the accuracy of the chemical shifts is likely lower. The lack of residual helical structure in this portion of the N-terminal helix in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl is consistent with the lower intrinsic helical propensity predicted by Agadir¹³⁷ for residues 2 – 7 of the N-terminal helix (Figure S1 of the Supporting Information). The 60's helix is populated to approximately 5 to 8 % in both the open and closed loop forms of the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl (Fig. 7). There is a region of non-native helix preceding the 60's helix (approximately residues 55 to 61), which is populated to a similar level. A peptide fragment corresponding to the C-terminal helix of horse cytochrome *c* has been shown to have significant helical content.¹³⁸ Consistent with this observation, there is also significant residual helical structure in the C-terminal helix, which is populated up to 5 to 7% (Fig. 7) in both the presence and the absence of the His54-heme loop in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl (Fig. 7).

Residual helical structure is also observed in the denatured state in 3 M GdnHCl outside of the sequence regions corresponding to the three long helices in the native state of TM_T(-5)P/N52I/K54H iso-1-Cytc. Starting at Pro76 in the middle of Ω -loop D (residues 70 to 85), there is non-native residual helical structure on the order of ~5% in both the presence and the absence of the His54-heme loop (Fig 7). There is also residual helical structure near residue 50, which begins near residue 45 and partially overlaps the short native 50's helix, which is formed by residues 49 to 55 in the native state of iso-1-Cytc. There is essentially no residual helical content in the C-terminal half of the native state 50's helix in the absence of loop formation (Fig. 7, pH 3.6; chemical shift data are not available for this region in the

presence of the His54-heme loop at pH 6.4). The low helical propensity of histidine¹³⁹ relative to the wild type lysine residue may be a factor in this case.

The absence of an observable signal from the backbone amide of F36 under any of the denaturing conditions tested is intriguing. It could be indicative of line broadening due to intermediate exchange through participation in a transient hydrophobic cluster or perhaps transient interactions with the heme. At pH 6.4, I35 and G37 give strong amide NH crosspeaks (Figure 5). At pH 3.6, the I35 and G37 crosspeaks are weaker (Figure 5, at 6 M GdnHCl, G37 is visible, but I35 is below the contour level; at 3 M GdnHCl, both crosspeaks are below the contour level). Thus, it appears that peak broadening is more pronounced when the loop is broken in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc and the polypeptide chain is less constrained. Roder's lab has shown based on NMR experiments that F36 and the preceding hydrophobic residue contribute significantly to nativelike structure that is present in an unfolding intermediate and also involves the N- and C-terminal helices and some residues in the 60's and 70's helices.¹⁴⁰ Our data indicate that F36 also may promote structure in the fully denatured state even in the presence of high concentrations of denaturant. Studies on heme peptide models show that Phe and Trp can make stabilizing contacts with the heme.^{141–143} In general, aromatic residues appear to be important for nucleating residual structure in the denatured state.^{10, 11, 13, 22, 27, 29, 76, 77}

Effect of the Imposed His54-heme Loop on Residual Structure in the Denatured State

The effect of the His54-heme loop on residual secondary structure in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl is relatively small (Fig. 7). Significant changes occur in the C-terminal helix. At pH 6.4 (loop formed) there is residual helical structure over much of the length of the C-terminal helix, whereas the C-terminal end of the helix has no significant residual structure at pH 3.6. This effect could be due to protonation of Glu103 which is expected to have a pK_a near 4.35 in concentrated GdnHCl solutions,¹⁴⁴ leading to loss of stabilizing salt bridge interactions with Lys99 and Lys100 ($i, i+3$ and $i, i+4$ helical contacts). The differences in residual helical structure in the presence (pH 6.4) and absence (pH 3.6) of the His54-heme loop at the C-terminal end of the C-terminal helix apparent in Figure 7 are not well predicted by Agadir (Figure S1 in the Supporting Information). The constraint imposed by loop formation may cause an increase in transient long-range interactions between the C-terminal helix and structure induced by the His-heme loop in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl, which stabilizes the C-terminal helix (see below).

The only other segment of TM_T(-5)P/N52I/K54H iso-1-Cytc that shows a significant enhancement in helical content upon loop formation is the region between residues 42 and 49, which immediately precedes His54. Given the proximity of these residues to the His54-heme contact, some caution should be exercised in interpreting these SSP scores. However, in 3 M GdnHCl at pH 6.4 in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc, the intensity of the backbone amide crosspeaks for the sequence preceding His54 shows periodicity suggestive of helical structure. His54 and I53 are not observed. I52 has a weak crosspeak, the A51 crosspeak is barely above the lower threshold of the spectrum shown in Figure 5 and D50 is not observed. In a helical structure, I52 would be furthest from the heme

and D50 would be expected to be nearest to the heme surface. Studies on heme peptide model systems show that when a His-containing peptide binds to heme there is induction of α -helix in the peptide.^{141–143, 145–147} Furthermore, the helical structure induced by His54-heme binding in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl may provide a docking site for the C-terminal helix leading to the observed increase in helical content for the C-terminal helix. In 6 M GdnHCl at pH 6.4, the amide NH crosspeaks are not observed from A51 to L58. The lower protein concentration in the 6 M GdnHCl experiments is likely a factor. However, the crosspeak for D50 is clearly visible, albeit weak (Figure 5), suggesting that the helical structure present in 3 M GdnHCl has been disrupted in 6 M GdnHCl.

Implications of Residual Structure in the Denatured State for the Folding Mechanism of Cytochrome c

Kinetic hydrogen-deuterium exchange (HX) data for horse Cytc, are consistent with the N- and C-terminal helices coming together to form the earliest folding intermediate, followed by the 60s helix and the Ω loop running from residues 19 – 36.^{148, 149} Our chemical shift data in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cytc in 3 M GdnHCl show that significant residual helical structure exists in both the 60s and C-terminal helices. Thus, these helices appear to have flickering structure that predisposes them to be involved in the early steps of the folding mechanism of cytochrome c. The portion of the N-terminal helix, which we can observe, does not contain significant residual helical structure. The helical content of isolated peptides corresponding to the N-terminal (3% helix), 60s (7% helix) and C-terminal (27% helix) helices of horse Cytc show that only the C-terminal helix is intrinsically stable.¹³⁸ In a compact molten globule at pD 2.2 and 1.5 M NaCl (A-state), the protection factors of the N-terminal, 60s and C-terminal helices are similar.¹⁵⁰ The same remains true at pD 2.2 and 0.02 M NaCl, where the acid state of horse Cytc is expanded to a pre-molten globule state, although the protection factors for the N-terminal, 60s and C-terminal helices are two orders of magnitude lower in the pre-molten globule state.^{151, 152} For the pre-molten globule form of horse Cytc, the protection factors suggest that the equilibrium constant for helix formation is near 10 for each of these helices, much higher than expected from the helical content of these helices in isolation. The similar protection factors suggest formation of modestly stable submolecular folding units via long-range interactions between these helices.^{151, 152} In 3 M GdnHCl, helical content decreases a further two orders of magnitude (5 – 10% helical content) in the 60s and C-terminal helices (Fig. 7). However, the similar helical content despite the significantly different intrinsic stabilities of isolated peptides corresponding to these helices¹³⁸ suggests the possibility of transient interactions between these helices even under the strong denaturing conditions of 3 M GdnHCl. The low helical content at the Gly83-Gly84 segment (Fig. 7), may act as a flexible molecular hinge that facilitates these long-range interactions in the denatured state as suggested for Gly-containing sequence segments in work from the Wright and Hosur labs.^{40–42, 52}

Our data also suggest that Phe36, which is part of the loop that folds in synche with the 60s helix, may be involved in a transient hydrophobic cluster in the denatured state involving the heme. Data from Roder's lab also indicates that Phe36 is involved in an equilibrium folding

intermediate.¹⁴⁰ The equilibrium HX data for both the molten globule and pre-molten globule states of horse Cyt_c show that amide NH of Phe36 is protected significantly from exchange.^{150, 151} Thus, Phe36 also may play an important role in establishing nativelike topological constraints, through a transient interaction with the heme (see Fig. 1) in the denatured state, providing for efficient folding.

CONCLUSION

The SSP scores observed in the denatured state of TM_T(-5)P/N52I/K54H iso-1-Cyt_c in 3 M GdnHCl at pH 6.4, when an imposed loop between His54 and the heme is present, and at pH 3.6, when the loop is absent, indicate that modest amounts of residual helical content (5 – 8%) remain in the 60's and C-terminal helices. Significant equilibrium HX protection factors also are observed for these helices in both the molten globule and per-molten globule states of horse Cyt_c. The portion of the N-terminal helix detectable in our data appears to be largely devoid of residual helical structure in both the presence and absence of the His54-heme loop. The similar residual helical content in the 60s and C-terminal helices in 3 M GdnHCl is consistent with data for both the pre-molten globule and molten globule states of horse Cyt_c.^{150, 151} Collectively, these observations suggest that the early formation of structure in these helices during folding are rooted in intrinsic topological biases that persist in the denatured state. In the presence of the His54-heme loop, there is an increase in residual helical structure that is evident at the C-terminus of the C-terminal helix. There may also be some enhancement in helical content for residues 42 – 49 preceding the His54-heme bond and in the 60's helix following it. Overall the observed effects of His54-heme loop formation are small, but similar in magnitude to the long-range changes in helical structure induced by mutations in the denatured state of acyl-coenzyme A binding protein.⁸²

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This research was supported by National Science Foundation grant, MCB-1412164 and NIH grant, R01GM074750, to B.E.B. The Bruker microflex MALDI-TOF mass spectrometer was purchased with Major Research Instrumentation Grant CHE-1039814 from the National Science Foundation. B.E.B. acknowledges support from CoBRE grant P20GM103546 from NIGMS. The NMR Core Facility at the University of Montana also received support from CoBRE grant P20GM103546 from NIGMS.

ABBREVIATIONS

iso-1-Cyt_c	iso-1-cytochrome <i>c</i>
CD	circular dichroism
GdnHCl	guanidine hydrochloride
G_u^0 (H₂O)	free energy of unfolding in the absence of denaturant

m rate of change of free energy of unfolding as a function of denaturant concentration

References

1. Daggett V, Fersht AR. Is there a unifying mechanism for protein folding? Trends Biochem Sci. 2003; 25:18–25.
2. Englander SW. Protein folding intermediates and pathways studied by hydrogen exchange. Annu Rev Biophys Biomol Struct. 2000; 29:213–238. [PubMed: 10940248]
3. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J. Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: testing a 1-dimensional diffusion model for peptide dynamics. J Phys Chem B. 2002; 106:11628–11640.
4. 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]
5. Hagen SJ, Hofrichter J, Szabo A, Eaton WA. Diffusion-limited contact formation in unfolded cytochrome *c*: estimating the maximum rate of protein folding. Proc Natl Acad Sci USA. 1996; 93:11615–11617. [PubMed: 8876184]
6. Jones CM, Henry ER, Hu Y, Chan CK, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA. Fast events in protein folding initiated by nanosecond laser photolysis. Proc Natl Acad Sci USA. 1993; 90:11860–11864. [PubMed: 8265638]
7. Fierz B, Satzger H, Root C, Gilch P, Zinth W, Kiefhaber T. Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. Proc Natl Acad Sci USA. 2007; 104:2163–2168. [PubMed: 17284588]
8. 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]
9. 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]
10. Khan MKA, Miller AL, Bowler BE. Tryptophan significantly stabilizes His-heme loops only when it is near a loop end. Biochemistry. 2012; 51:3586–3595. [PubMed: 22486179]
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. 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]
13. 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]
14. 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]
15. 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]
16. 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]
17. 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]
18. 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]
19. 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]
20. Tanford C. Protein denaturation. *Adv Protein Chem*. 1968; 23:121–282. [PubMed: 4882248]
21. Bowler BE. Residual structure in unfolded proteins. *Curr Opin Struct Biol*. 2012; 22:4–13. [PubMed: 21978577]
22. Bowler, BE. Globular proteins: characterization of the denatured state. In: Egelman, E., editor. *Comprehensive Biophysics*. Academic Press; Oxford; 2012. p. 72-114.
23. Mittag T, Forman-Kay JD. Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol*. 2007; 17:3–14. [PubMed: 17250999]
24. Dyson HJ, Wright PE. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv Protein Chem*. 2002; 62:311–340. [PubMed: 12418108]
25. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chem Rev*. 2004; 104:3607–3622. [PubMed: 15303830]
26. 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]
27. Marsh JA, Neale C, Jack FE, Choy WY, Lee AY, Crowhurst KA, Forman-Kay JD. Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J Mol Biol*. 2007; 367:1494–1510. [PubMed: 17320108]
28. 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]
29. 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]
30. 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]
31. Zhang O, Kay LE, Shortle D, Forman-Kay JD. Comprehensive NOE characterization of a partially folded large fragment of staphylococcal nuclease ¹³¹N, using NMR methods with improved resolution. *J Mol Biol*. 1997; 272:9–20. [PubMed: 9299333]
32. Zhang O, Forman-Kay JD. Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry*. 1995; 34:6784–6794. [PubMed: 7756310]
33. Mok YK, Kay CM, Kay LE, Forman-Kay J. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J Mol Biol*. 1999; 289:619–638. [PubMed: 10356333]
34. Neri D, Billeter M, Wider G, Wüthrich K. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*. 1992; 257:1559–1563. [PubMed: 1523410]
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. Bhavesh NS, Panchal SC, Mittal R, Hosur RV. NMR identification of local structural preferences in HIV-1 protease tethered heterodimer in 6 M guanidine hydrochloride. *FEBS Lett*. 2001; 509:218–224. [PubMed: 11741592]
37. Bhavesh NS, Sinha R, Mohan PMK, Hosur RV. NMR elucidation of early folding hierarchy in HIV-1 protease. *J Biol Chem*. 2003; 278:19980–19985. [PubMed: 12644464]
38. Bhavesh NS, Juneja J, Udgaonkar JB, Hosur RV. Native and nonnative conformational preferences in the urea-unfolded state of barstar. *Protein Sci*. 2004; 13:3085–3091. [PubMed: 15537753]
39. Chugh J, Sharma S, Hosur RV. Pockets of short-range transient order and restricted topological heterogeneity in the guanidine-denatured state ensemble of GED of dynamin. *Biochemistry*. 2007; 46:11819–11832. [PubMed: 17910478]
40. Kumar A, Srivastava S, Mishra RK, Mittal R, Hosur RV. Local structural preferences and dynamics restrictions in the urea-denatured state of SUMO-1: NMR characterization. *Biophys J*. 2006; 90:2498–2509. [PubMed: 16415059]

41. Kumar D, Chugh J, Sharma S, Hosur RV. Conserved structural and dynamics features in the denatured states of drosophila SUMO, human SUMO and ubiquitin proteins: Implications to sequence-folding paradigm. *Proteins: Struct Funct Bioinform.* 2009; 76:387–402.
42. Mishra P, Das S, Panicker L, Hosur MV, Sharma S, Hosur RV. NMR insights into folding and self-association of *Plasmodium falciparum* P₂. *PLoS One.* 2012; 7:e36279. [PubMed: 22567147]
43. Gillespie JR, Shortle D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol.* 1997; 268:158–169. [PubMed: 9149149]
44. Gillespie JR, Shortle D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J Mol Biol.* 1997; 268:179–184.
45. Rösner HI, Poulsen FM. Residue-specific description of non-native transient structures in the ensemble of acid-denatured structures of the all- β protein c-src SH3. *Biochemistry.* 2010; 49:3246–3253. [PubMed: 20218679]
46. 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]
47. Teilum K, Thormann T, Caterer NR, Poulsen HI, Jensen PH, Knudsen J, Kragelund BB, Poulsen FM. Different secondary structure elements as scaffolds for protein folding transition states of two homologous four-helix bundles. *Proteins: Struct Funct Bioinform.* 2005; 59:80–90.
48. 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]
49. Teilum K, Kragelund BB, Poulsen FM. Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. *J Mol Biol.* 2002; 324:349–357. [PubMed: 12441112]
50. Mohana-Borges R, Goto NK, Kroon GJA, Dyson HJ, Wright PE. Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J Mol Biol.* 2004; 340:1131–1142. [PubMed: 15236972]
51. 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]
52. Schwarzingner S, Wright PE, Dyson HJ. Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry.* 2002; 41:12681–12686. [PubMed: 12379110]
53. 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]
54. Yao J, Chung J, Eliezer D, Wright PE, Dyson HJ. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry.* 2001; 40:3561–3571. [PubMed: 11297422]
55. Cho JH, Sato S, Horng JC, Anil B, Raleigh DP. Electrostatic interactions in the denatured state ensemble: their effect upon protein folding and protein stability. *Arch Biochem Biophys.* 2008; 469:20–28. [PubMed: 17900519]
56. Cho JH, Raleigh DP. Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. *J Mol Biol.* 2005; 353:174–185. [PubMed: 16165156]
57. Cho JH, Sato S, Raleigh DP. Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J Mol Biol.* 2004; 338:827–837. [PubMed: 15099748]
58. 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]

59. 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]
60. Tan YJ, Oliveberg M, Davis B, Fersht AR. Perturbed pK_A -values in the denatured states of proteins. *J Mol Biol.* 1995; 254:980–992. [PubMed: 7500365]
61. Oliveberg M, Vuilleumier S, Fersht AR. Thermodynamic study of the acid denaturation of barnase and its dependence on ionic strength: evidence for residual electrostatic interactions in the acid/thermally denatured state. *Biochemistry.* 1994; 33:8826–8832. [PubMed: 8038174]
62. Oliveberg M, Arcus VL, Fersht AR. pK_A values of carboxyl groups in the native and denatured states of barnase: the pK_A values of the denatured state are on average 0.4 units lower than those of model compounds. *Biochemistry.* 34:9424–9433.
63. Pletneva EV, Gray HB, Winkler JR. Many faces of the unfolded state: conformational heterogeneity in denatured yeast cytochrome *c*. *J Mol Biol.* 2005; 345:855–867. [PubMed: 15588831]
64. Pletneva EV, Gray HB, Winkler JR. Snapshots of cytochrome *c* folding. *Proc Natl Acad Sci USA.* 2005; 102:18397–18402. [PubMed: 16344477]
65. Lee JC, Engman KC, Tezcan FA, Gray HB, Winkler JR. Structural features of cytochrome *c'* folding intermediates revealed by fluorescence energy-transfer kinetics. *Proc Natl Acad Sci USA.* 2002; 99:14778–14782. [PubMed: 12407175]
66. Choy WY, Mulder FAA, Crowhurst KA, Muhandiram DR, Millett IS, Doniach S, Forman-Kay JD, Kay LE. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J Mol Biol.* 2002; 316:101–112. [PubMed: 11829506]
67. Konuma T, Kimura T, Matsumoto S, Goto Y, Fujisawa T, Fersht AR, Takahashi S. Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase. *J Mol Biol.* 2011; 405:1284–1294. [PubMed: 21146541]
68. Segel DJ, Fink AL, Hodgson KO, Doniach S. Protein denaturation: a small angle X-ray scattering study of the ensemble of unfolded states of cytochrome *c*. *Biochemistry.* 1998; 37:12443–12451. [PubMed: 9730816]
69. Kathuria SV, Kayatekin C, Barrea R, Kondrashkina E, Graceffa R, Guo L, Nobrega RP, Chakravarthy S, Matthews CR, Irving TC, Bilsel O. Microsecond barrier-limited chain collapse observed by time-resolved FRET and SAXS. *J Mol Biol.* 2014; 426:1980–1994. [PubMed: 24607691]
70. Shan B, McClendon S, Rospigliosi C, Eliezer D, Raleigh DP. The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure. *J Am Chem Soc.* 2010; 132:4669–4677. [PubMed: 20225821]
71. Shan B, Eliezer D, Raleigh DP. The unfolded state of the C-terminal domain of the ribosomal protein L9 contains both native and non-native structure. *Biochemistry.* 2009; 48:4707–4719. [PubMed: 19301913]
72. Shan B, Bhattacharya S, Eliezer D, Raleigh DP. The low-pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry.* 2008; 47:9565–9573. [PubMed: 18707127]
73. Blanco FJ, Serrano L, Forman-Kay JD. High populations of non-native structures in the denatured state are compatible with the formation of the native folded state. *J Mol Biol.* 1998; 284:1153–1164. [PubMed: 9837733]
74. Wirmer J, Peti W, Schwalbe H. Motional properties of unfolded ubiquitin: a model for a random coil protein. *J Biomol NMR.* 2006; 35:175–186. [PubMed: 16865418]
75. Wirmer J, Schloerb C, Schwalbe H. Conformation and dynamics of nonnative states of proteins studied by NMR spectroscopy. *Protein Folding Handbook.* 2005; 2:737–808.
76. 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.

77. 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]
78. Schwalbe H, Fiebig KM, Buck M, Jones JA, Grimshaw SB, Spencer A, Glaser SJ, Smith LJ, Dobson CM. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry*. 1997; 36:8977–8991. [PubMed: 9220986]
79. Meng W, Lyle N, Luan B, Raleigh DP, Pappu RV. Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. *Proc Natl Acad Sci USA*. 2013; 110:2123–2128. [PubMed: 23341588]
80. Meng W, Luan B, Lyle N, Pappu RV, Raleigh DP. The denatured state ensemble contains significant local and long-range structure under native conditions: analysis of the N-terminal domain of ribosomal protein L9. *Biochemistry*. 2013; 52:2662–2671. [PubMed: 23480024]
81. Luan B, Lyle N, Pappu RV, Raleigh DP. Denatured state ensembles with the same radii of gyration can form significantly different long-range contacts. *Biochemistry*. 2014;39–47. [PubMed: 24280003]
82. Bruun SW, Ieřmantavius 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]
83. Clarke J, Hounslow AM, Bond CJ, Fersht AR, Daggett V. The effects of disulfide bonds on the denatured state of barnase. *Protein Sci*. 2000; 9:2394–2404. [PubMed: 11206061]
84. Betz SF, Pielak GJ. Introduction of a disulfide bond into cytochrome *c* stabilizes a compact denatured state. *Biochemistry*. 1992; 31:12337–12344. [PubMed: 1334426]
85. Pace CN, Grimsley GR, Thomson JA, Barnett BJ. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J Biol Chem*. 1988; 263:11820–11825. [PubMed: 2457027]
86. Clarke J, Fersht AR. Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. *Biochemistry*. 1993; 32:4322–4329. [PubMed: 8476861]
87. Parrini C, Bemporad F, Baroncelli A, Gianni S, Travaglini-Allocatelli C, Kohn JE, Ramazzotti M, Chiti F, Taddei N. The folding process of acylphosphatase from *Escherichia coli* is remarkably accelerated by the presence of a disulfide bond. *J Mol Biol*. 2008; 379:1107–1118. [PubMed: 18495159]
88. Mason JM, Gibbs N, Sessions RB, Clarke AR. The influence of intramolecular bridges on the dynamics of a protein folding reaction. *Biochemistry*. 2002; 41:12093–12099. [PubMed: 12356309]
89. Mason JM, Cliff MJ, Sessions RB, Clarke AR. Low energy pathways and non-native interactions: the influence of artificial disulfide bridges on the mechanism of folding. *J Biol Chem*. 2005; 280:40494–40499. [PubMed: 16216876]
90. Matsumura M, Matthews BW. Stabilization of functional proteins by introduction of multiple disulfide bonds. *Methods Enzymol*. 1991; 202:336–356. [PubMed: 1784181]
91. Feige MJ, Hagn F, Esser J, Kessler H, Buchner J. Influence of the internal disulfide bridge on the folding pathway of the C_L antibody domain. *J Mol Biol*. 2007; 365:1232–1244. [PubMed: 17112539]
92. Grantcharova VP, Riddle DS, Baker D. Long-range order in the src SH3 folding transition state. *Proc Natl Acad Sci USA*. 2000; 97:7084–7089. [PubMed: 10860975]
93. 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]
94. Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between α - and γ -synuclein: implications for fibrillation. *Protein Sci*. 2006; 15:2795–2804. [PubMed: 17088319]
95. Duncan MG, Williams MD, Bowler BE. Compressing the free energy range of substructure stabilities in iso-1-cytochrome *c*. *Protein Sci*. 2009; 18:1155–1164. [PubMed: 19472325]

96. 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]
97. 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]
98. 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]
99. Bowler BE, May K, Zaragoza T, York P, Dong A, Caughey WS. Destabilizing effects of replacing a surface lysine of cytochrome *c* with aromatic amino acids: implications for the denatured state. *Biochemistry.* 1993; 32:183–190. [PubMed: 8380333]
100. Margoliash E, Frohwirt N. Spectrum of horse-heart cytochrome *c*. *Biochem J.* 1959; 71:570–572. [PubMed: 13638266]
101. Nozaki Y. The preparation of guanidine hydrochloride. *Methods Enzymol.* 1972; 26:43–50. [PubMed: 4680720]
102. Kay LE, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc.* 1992; 114:10663–10665.
103. Wittekind M, Mueller L. HNCACB, a high-sensitivity 3D experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J Magn Reson B.* 1993; 101:201–205.
104. Grzesiek S, Bax A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc.* 1992; 114:6291–6293.
105. Grzesiek S, Anglister J, Bax A. Correlation of backbone amide and aliphatic side-chain resonances in $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of ^{13}C magnetization. *J Magn Reson B.* 1993; 101:114–119.
106. Ikura M, Kay LE, Bax A. A novel approach for sequential assignment of ^1H , ^{13}C , and ^{15}N spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry.* 1990; 29:4659–4667. [PubMed: 2372549]
107. Muhandrian DR, Kay LE. Gradient-enhanced triple resonance three-dimensional NMR experiments with improved sensitivity. *J Magn Reson B.* 1994; 103:203–216.
108. Clubb RT, Thanabal V, Wagner G. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ^1HN , ^{15}N , and $^{13}\text{C}'$ chemical shifts in ^{15}N - ^{13}C -labelled proteins. *J Magn Reson.* 1992; 97:213–217.
109. Wishart DS, Bigam CG, Yao J, Abilgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J Biomol NMR.* 1995; 6:135–140. [PubMed: 8589602]
110. Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K. Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J Biomol NMR.* 1998; 12:1–23. [PubMed: 9729785]
111. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995; 6:277–293. [PubMed: 8520220]
112. Goddard, TD., Kneller, DG. SPARKY. Vol. 3. University of California; San Francisco CA:
113. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct Funct Bioinform.* 2005; 59:687–696.
114. Zhang H, Neal S, Wishart DS. RefDB: A database of uniformly referenced protein chemical shifts. *J Biomol NMR.* 2003; 25:173–195. [PubMed: 12652131]
115. 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]
116. Hickey DR, Berghuis AM, Lafond G, Jaeger JA, Cardillo TS, McLendon D, Das G, Sherman F, Brayer GD, McLendon G. Enhanced thermodynamic stabilities of yeast iso-1-cytochromes *c*

with amino acid replacements at positions 52 and 102. *J Biol Chem.* 1991; 266:11686–11694. [PubMed: 1646814]

117. 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]
118. Hammack B, Godbole S, Bowler BE. Cytochrome *c* folding traps are not due solely to histidine-heme ligation: direct demonstration of a role for N-terminal amino group-heme ligation. *J Mol Biol.* 1998; 275:719–724. [PubMed: 9480763]
119. Dixon HBF, Fields R. Specific modification of NH₂-terminal residues by transamination. *Methods Enzymol.* 1972; 25:409–419. [PubMed: 23014421]
120. Clark RW, Youn H, Parks RB, Cherney MM, Roberts GP, Burstyn JN. Investigation of the role of the N-terminal proline, the distal heme ligand in the CO Sensor CoxA. *Biochemistry.* 2004; 43:14149–14160. [PubMed: 15518565]
121. Martell, AE., Smith, AB. Critical Stability Constants: Volume 6, Second Supplement. Springer Science + Business Media LLC; New York: 1989.
122. Hickey DR, McLendon G, Sherman F. Thermodynamic stabilities of the yeast iso-1-cytochromes *c* having amino acid substitutions for lysine 32. *J Biol Chem.* 1988; 263:18298–18305. [PubMed: 2848033]
123. Tsong TY. An acid induced conformational transition of denatured cytochrome *c* in urea and guanidine hydrochloride solution. *Biochemistry.* 1975; 7:1542–1547.
124. Hagihara Y, Tan Y, Goto Y. Comparison of the conformational stability of the molten globule and native states of horse cytochrome *c*. *J Mol Biol.* 1994; 237:336–348. [PubMed: 8145245]
125. Goldes ME, Jeakins-Cooley ME, McClelland LJ, Mou TC, Bowler BE. Disruption of a hydrogen bond network in human versus spider monkey cytochrome *c* affects heme crevice stability. *J Inorg Biochem.* 2016; 158:62–69. [PubMed: 26775610]
126. Bandi S, Bowler BE. Effect of an Ala81His mutation on the Met80 loop dynamics of iso-1-cytochrome *c*. *Biochemistry.* 2015; 54:1729–1742. [PubMed: 25671560]
127. McClelland LJ, Seagraves SM, Khan MKA, Cherney MM, Bandi S, Culbertson JE, Bowler BE. The response of Ω-loop D dynamics to truncation of trimethyllysine 72 of yeast iso-1-cytochrome *c* depends on the nature of loop deformation. *J Biol Inorg Chem.* 2015; 20:805–819. [PubMed: 25948392]
128. Cherney MM, Junior C, Bowler BE. Mutation of trimethyllysine-72 to alanine enhances His79-heme mediated dynamics of iso-1-cytochrome *c*. *Biochemistry.* 2013; 52:837–846. [PubMed: 23311346]
129. Xu Y, Mayne L, Englander SW. Evidence for an unfolding and refolding pathway in cytochrome *c*. *Nat Struct Biol.* 1998; 5:774–778. [PubMed: 9731770]
130. Adams, PA., Baldwin, DA., Marques, HM. The hemepeptides from cytochrome *c*: Preparation, physical and chemical properties, and their use as model compounds for the hemoproteins. In: Mauk, AG., Scott, RA., editors. *Cytochrome c: A Multidisciplinary Approach*. University Science Books; Sausalito, CA: 1996. p. 635-692.
131. Muthukrishnan K, Nall BT. Effective concentrations of amino acid side chains in an unfolded protein. *Biochemistry.* 1991; 30:4706–4710. [PubMed: 1851434]
132. Colón W, Wakem LP, Sherman F, Roder H. Identification of the predominant non-native histidine ligand in unfolded cytochrome *c*. *Biochemistry.* 1997; 36:12535–12541. [PubMed: 9376358]
133. Smith CR, Wandschneider E, Bowler BE. Effect of pH on the iso-1-cytochrome *c* denatured state: changing constraints due to heme ligation. *Biochemistry.* 2003; 42:2174–2184. [PubMed: 12590607]
134. Millett IS, Doniach S, Plaxco KW. Toward a taxonomy of the denatured state: small angle X-ray scattering studies of unfolded proteins. *Adv Protein Chem.* 2002; 62:241–262. [PubMed: 12418105]
135. Nabuurs SM, Westphal AH, van Mierlo CPM. Extensive formation of off-pathway species during folding of an α-β parallel protein is due to docking of (non)native structure elements in unfolded molecules. *J Am Chem Soc.* 2008; 130:16914–16920. [PubMed: 19053416]

136. Schwarzing S, Kroon GJA, Foss TR, Chung J, Wright PE, Dyson HJ. Sequence-Dependent Correction of Random Coil NMR Chemical Shifts. *J Am Chem Soc.* 2001; 123:2970–2978. [PubMed: 11457007]
137. Muñoz V, Serrano L. Elucidating the folding problem of α -helical peptides using empirical parameters III: temperature and pH dependence. *J Mol Biol.* 1994; 245:297–308.
138. Kuroda Y. Residual helical structure in the C-terminal fragment of cytochrome *c*. *Biochemistry.* 1993; 32:1219–1224. [PubMed: 8383525]
139. Rohl CA, Chakraborty A, Baldwin RL. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* 1996; 5:2623–2637. [PubMed: 8976571]
140. Latypov RF, Cheng H, Roder NA, Zhang J, Roder H. Structural characterization of an equilibrium unfolding intermediate in cytochrome *c*. *J Mol Biol.* 2006; 357:1009–1025. [PubMed: 16473367]
141. Cowley AB, Kennedy ML, Silchenko S, Lukat-Rodgers GS, Rodgers KR, Benson DR. Insight into heme potential redox control and functional aspects of six-coordinate ligand-sensing heme proteins from studies of synthetic heme proteins. *Inorg Chem.* 2006; 45:9985–10001. [PubMed: 17140194]
142. Williamson DA, Benson DR. Remarkable helix stabilization via edge-to-face tryptophan–porphyrin interactions in a peptide-sandwiched mesoheme. *Chem Commun.* 1998:961–962.
143. Liu D, Williamson DA, Kennedy ML, Williams TD, Benson DR. Aromatic side chain-porphyrin interactions in designed hemoproteins. *J Am Chem Soc.* 1999; 121:11798–11812.
144. Roxby R, Tanford C. Hydrogen ion titration curve of lysozyme in 6 M guanidine hydrochloride. *Biochemistry.* 1971; 10:3348–3352. [PubMed: 5165783]
145. Benson DR, Hart BR, Zhu X, Doughty MB. Design, synthesis, and circular dichroism investigation of a peptide-sandwiched mesoheme. *J Am Chem Soc.* 1997; 117:8502–8510.
146. Rosenblatt MM, Huffman DL, Wang X, Remmer HA, Suslick KS. Cyclic and hairpin peptide complexes of heme. *J Am Chem Soc.* 2002; 124:12394–12395. [PubMed: 12381164]
147. Huffman DL, Rosenblatt MM, Suslick KS. Synthetic heme-peptide complexes. *J Am Chem Soc.* 1998; 120:6183–6184.
148. Roder H, Elöve GA, Englander SW. Structural characterization of folding intermediates in cytochrome *c* by H-exchange labelling and proton NMR. *Nature.* 1988; 335:700–704. [PubMed: 2845279]
149. Hu W, Kan ZY, Mayne L, Englander SW. Cytochrome *c* folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc Natl Acad Sci USA.* 2016; 113:3809–3814. [PubMed: 26966231]
150. Jeng MF, Englander SW, Elöve GA, Wand AJ, Roder H. Structural description of acid-denatured cytochrome *c* by hydrogen exchange and 2D NMR. *Biochemistry.* 1990; 29:10433–10437. [PubMed: 2176867]
151. Jeng MF, Englander SW. Stable submolecular folding units in a non-compact form of cytochrome *c*. *J Mol Biol.* 1991; 221:1045–1061. [PubMed: 1658332]
152. Kuroda Y, Endo S, Kuniaki N, Wada A. Stability of α -helices in a molten globule state of cytochrome *c* by hydrogen-deuterium exchange and two-dimensional NMR spectroscopy. *J Mol Biol.* 1995; 247:682–688. [PubMed: 7723023]

Figure 1. Structure of iso-1-Cytc (PDB ID: 2YCC) showing the locations of residues substituted in this work. The residues at the substitution sites, Asn52, Lys54 and Thr(-5) are shown as stick models colored by element as are the heme and its ligands, Met80 and His18, and Phe36 which packs against the heme. Amino acid substitutions are indicated with the standard notation.

Figure 2. GdnHCl denaturation of the TM_T(-5)P/N52I (black) and TM_T(-5)P/N52I/K54H (red) variants of iso-1-Cytc. The data shown as open circles were not used in the fit (solid curves) to eq 1 (Experimental Procedures). All data were collected at 25 °C in 20 mM Tris (pH 7.0) 40 mM NaCl.

Figure 3.
Schematic representation of histidine-heme loop formation in the denatured state of TM_T(-5)P/N52I/K54H.

Figure 4.

Loop formation in the denatured state (3 M GdnHCl) of the TM_T(-5)P/N52I (black) and TM_T(-5)P/N52I/K54H (red) variants of iso-1-Cyt c. (A) Plots of $A_{398\text{corr}}$ versus pH. For TM_T(-5)P/N52I, the N-terminal secondary amino group of Pro(-5) forms the loop. For TM_T(-5)P/N52I/K54H, His 54 forms the loop. Solid black curves are fits to eq. 2 in Experimental Procedures. All data were collected at 22 ± 1 °C and in the presence of 5 mM Na_2HPO_4 and 15 mM NaCl. (B) The Soret band of the TM_T(-5)P/N52I/K54H variant at pH values ranging from 7.07 (peak near 407 nm, low spin heme) to pH 2.81 (peak near 396 nm, high spin heme). Absorbance at 450 nm was used as a baseline to correct for spectrometer drift during the titration experiments.

Author Manuscript

Figure 6.

SSP scores derived from $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ assignments at (A) pH 6.4 where the His54-heme loop is formed and at (B) pH 3.6, where the loop is broken. At both pH values, the SSP scores are shown in magenta for 3 M GdnHCl and in gray for 6 M GdnHCl. The position of helices in the native structure of iso-1-Cytc are shown schematically above the graph. Helices are represented as red cylinders and loops as blue lines.

Figure 7.

Plot of Δ SSP scores versus sequence position. The Δ SSP scores in the presence of the His54-heme loop (pH 6.4) are shown in red and in the absence of the loop (pH 3.6) are shown in gray. The position of helices in the native structure of iso-1-Cyt c are shown schematically above the graph. Helices are represented as red cylinders and loops as blue lines.

Table 1Thermodynamic Parameters for GdnHCl Denaturation of Iso-1-Cyt c Variants at 25 °C and pH 7

Variant	$G_u^{\circ'}(\text{H}_2\text{O})$ (kcal/mol)	m -value (kcal mol $^{-1}$ M $^{-1}$)	C_m^a (M)
TM_T(-5)P/N52I	7.07 ± 0.15	4.44 ± 0.12	1.59 ± 0.04
TM_T(-5)P/N52I/K54H	3.91 ± 0.15	3.93 ± 0.18	0.99 ± 0.01
AcTMI52 ^b	6.63 ± 0.23	3.82 ± 0.19	1.74 ± 0.06
AcH54I52 ^b	3.93 ± 0.02	3.96 ± 0.02	0.99 ± 0.01

^a C_m is the GdnHCl concentration at the midpoint of the unfolding titration.

^bParameters are taken from ref. ¹¹⁷. The AcTMI52 and AcH54I52 variants are identical to the TM_T(-5)P/N52I and TM_T(-5)P/N52I/K54H variants, respectively, except that instead of proline at position -5, they contain T(-5)S and K(-2)L mutations so that the N-terminus is acetylated when expressed in *S. cerevisiae*.

Table 2

Thermodynamic Parameters for Loop Formation in Iso-1-Cyt c Variants in the Denatured State (3 M GdnHCl) at 22 ± 1 °C

Variant	$pK_a(\text{obs})$	n_p
TM_T(-5)P/N52I	6.67 ± 0.01	1.02 ± 0.02
TM_T(-5)P/N52I/K54H	4.74 ± 0.04	0.78 ± 0.05
TM ^a	5.94 ± 0.08	0.9 ± 0.1
AcH54I52 ^b	4.80 ± 0.02	0.83 ± 0.02

^aThe TM variant was expressed in yeast and has threonine at the N-terminus; its N-terminal α -amino group is not acetylated. It also has the H26N, H33N and H39Q mutations. Parameters are from ref.¹⁸

^bParameters are taken from ref.¹⁶