HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2022 October 19.

Published in final edited form as:

Biochemistry. 2021 October 19; 60(41): 3071–3085. doi:10.1021/acs.biochem.1c00400.

Denatured State Conformational Biases in Three-Helix Bundles Containing Divergent Sequences Localize Near Turns and Helix Capping Residues

Moses J. Leavens[†,‡], **Lisa E. Spang**[†,‡], **Melisa M. Cherney**[†,‡], **Bruce E. Bowler**[†,‡,*]

[†]Department of Chemistry & Biochemistry, University of Montana, Missoula, Montana 59812, United States

[‡]Center for Biomolecular Structure & Dynamics, University of Montana, Missoula, Montana 59812, United States

Abstract

Rhodopseudomonas palustris cytochrome c', a four-helix bundle, and the second ubiquitin-associated domain, UBA(2), a three-helix bundle from the human homolog of yeast Rad23, HHR23A, deviate from random coil behavior under denaturing conditions in a fold-specific manner. The random coil deviations in each of these folds occur near interhelical turns and loops in their tertiary structures. Here, we examine an additional three-helix bundle with an identical fold to UBA(2), but a highly divergent sequence, the first ubiquitin-associated domain, UBA(1), of HHR23A. We use histidine-heme loop formation methods, employing eight single histidine variants, to probe for denatured state conformational bias of a UBA(1) domain fused to the N-terminus of iso-1-cytochrome *c* (iso-1-Cyt*c*). Guanidine hydrochloride (GuHCl) denaturation shows that the iso-1-Cyt*c* domain unfolds first, followed the UBA(1) domain. Denatured state (4 M and 6 M GuHCl) histidine-heme loop formation studies show that as the size of the histidine-heme loop increases, loop stability decreases, as expected for the Jacobson-Stockmayer relationship. However, loops formed with His35, His31 and His15, of UBA(1), are 0.6 – 1.1 kcal/mol more stable than expected from the Jacobson-Stockmayer relationship, confirming the importance of deviations of the denatured state from random coil behavior near interhelical turns of helical domains, for facilitating folding to the correct topology. For UBA(1) and UBA(2),

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

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Tables S1 and S2 provide sequences for oligonucleotides used for cloning and mutagenesis. Table S3 contains MALDI-TOF data for the UBA(1) – iso-1-Cyt*c* variants. Table S4 contains loop breakage rate constants for the UBA(1) domain fused to iso-1-Cyt*c* in 6 M and 4 M GuHCl at pH 3.0 and 3.5. Table S5 provides loop formation rate constants. Table S6 provides *m*-values for the thermodynamics and kinetics of loop formation. Figure S1 shows a typical kinetic trace for loop breakage with a UBA(1) – iso-1-Cyt*c* variant. Figure S2 is a plot of pK_{loop}(His) versus Log(N) at 6 M GuHCl for UBA(1) compared to UBA(2). Figure S3 is a plot of Log(k_f) versus Log(N) at 4 M and 6 M GuHCl for UBA(1) compared to data for homopolymeric sequences of Ala and Gln. Figures S4 and S5 are hydropathy plots for UBA(1) – iso-1-Cyt*c* and UBA(2) – iso-1-Cyt*c*, respectively. Figures S6 and S7 show ILV clusters for the UBA(2) domain. Figure S8 is a plot of k_b versus HpC for UBA(2) – iso-1-Cyt*c*. Figures S9 and S10 contain plots of local NCPR versus sequence position and k_b versus local NCPR for UBA(1) – iso-1-Cyt*c* and UBA(2) – iso-1-Cyt*c*, respectively. The Supporting Information is available free of charge on the ACS Publications website.

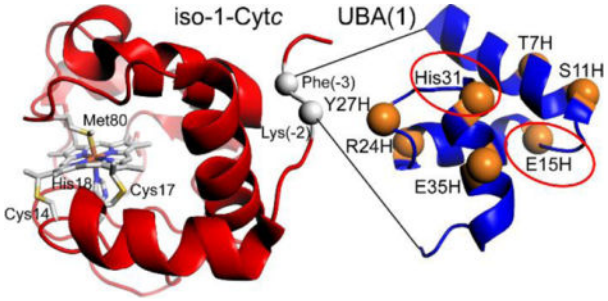
Accession Codes

UniProt ID: yeast iso-1-cytochrome *c*, P00044; UV excision repair protein RAD23 homolog A, P54725.

The authors declare no competing financial interest.

hydrophobic clusters on either side of the turns partially explain deviations from random coil behavior, however, helix capping also appears to be important.

Graphical Abstract



INTRODUCTION

One conclusion of Levinthal's paradox is that efficient folding requires a biased search.[1] Early estimates of the magnitude of necessary bias indicated conformational biases on the order of 2kT would be sufficient to allow a protein to fold on a biological time scale.[2] Experimental measurements of conformational biases resulting from residual structure in the denatured state ensemble (DSE) indicate that they are on this order of magnitude.[3–10] However, numerous observations have indicated that denatured proteins possess bulk dimensions compatible with random coil behavior,[11] based on the relationship between the radius of gyration, R_g, and the number of residues N in proteins under denaturing conditions ($R_g \alpha N^{\nu}$, where ν is the scaling exponent). In particular, values of ν near 0.6, as expected for a random coil are observed for large sets of chemically denatured proteins.[12,13] Despite the prevalence of bulk properties consistent with random coil behavior there are many examples of proteins that contain significant residual structure.[14–26] Moreover, the interactions in the DSE are long range and involve hydrophobic stabilization, indicating that these interactions may be critical for an efficient conformational search to find the native state of a protein chain.[27]

Small helical bundles are well-studied as models for protein folding.[28] Persistent helical bias is readily studied with peptide fragments,[29,30] and identified by deviation from random coil chemicals shifts under denaturing conditions.[31–33] The DSE of acyl-Coenzyme A binding protein (ACBP), a four-helix bundle, has been particularly well-studied.[34–39] ACBP possesses long range contacts which manifest its native state topology. Work on the three-helix bundle B fragment from protein A, shows formation of native-like secondary structure in the DSE, where a C-terminal hairpin is thought to be the region of local collapse in its DSE.[40,41] Moreover, in a minimally topologically frustrated thermodynamic and kinetic model of the B fragment from protein A, the highest ϕ values are located in turn regions, suggesting these regions are required in the transition state for folding.[42] Our lab also has shown that thermodynamically significant biases occur near turn sequences of helical bundles.[4,17,43,44]

Recent work on the DSE has focused on the denaturant dependence of R_g as measured by small-angle X-ray scattering (SAXS) versus end-to-end distances, R_{ee}, as measured by fluorescence resonance energy transfer (FRET).[45–55] In general, both SAXS and FRET agree that $\nu \sim 0.6$ in 6 M guanidine hydrochloride (GuHCl) and 8 M urea and that there is a gradual decrease in ν as denaturant concentration is lowered, with more pronounced effects occurring below ~2 – 3 M denaturant concentration. The significance of DSE compaction for protein folding is still debated,[45,46,52,53] however, distribution of charge[56] and clustering of hydrophobic groups[46] along the primary structure of a protein have emerged as possible factors that control the dimensions of the DSE and may impact the efficiency of protein folding. In addition, these factors may affect protein aggregation, an important factor in many human diseases.[57]

Our lab developed a method to evaluate the thermodynamics and kinetics of His-heme loop formation in the DSE.[8,58–60] The advantage of this method is that it can interrogate the properties of the DSE along the length of the primary structure of a protein. The method is adept at detecting local sequence-specific variations in the properties of the DSE that result from proteins being heteropolymers composed of a set of twenty amino acids with disparate properties. Thus, in principle, the effects of local clustering of hydrophobic groups or charge distribution should be detectable.

We and others have shown that loop formation with sequences of amino acids that are of low-complexity conforms well to simple polymer models.[61–68] Thus, we use the Jacobson-Stockmayer (JS) equation (eq 1),[69] which makes the assumption that loop formation for a

$$\Delta S_{\text{loop}} = -\nu_3 \mathrm{R} ln(\mathrm{N}) + \mathrm{R} ln\left(\left(3/2\pi C_{\mathrm{n}} l^2\right)^{\upsilon 3} \mathrm{V_i}\right) \quad (1)$$

random coil is solely limited by entropy, to provide a reference state to evaluate deviations of foldable heteropolymeric sequences from random coil behavior. In eq 1, C_n is Flory's characteristic ratio, V_i is the approach volume the monomers must be contained within for a loop to form, l is the distance between the monomers in the polymer, R is the gas constant, ν_3 is the loop formation scaling exponent, and N is the number of residues in the loop. Loop formation requires the ends of a loop to search in three-dimensions to make contact, so, the scaling exponent for the loop size dependence of N, ν_3, is three-fold larger than ν for the dependence of R_g on N. Thus, ν_3 is 1.8 for a random coil when its excluded volume is accounted for.[70] However, simulations indicate that excluded volume effects could increase ν_3 to 2.1 – 2.4.[71,72]

Initially, the His-heme loop formation method was limited to proteins containing a covalently attached heme (c-type heme).[4,44,59,60] We later showed that the method could be used to evaluate the conformational properties of low complexity homopolymeric sequences inserted at the N-terminus of yeast iso-1-cytochrome *c* (iso-1-Cyt*c*),[64–66] which demonstrated that His-heme loop formation with foldable proteins differed markedly from loop formation with low complexity sequences.[66] Recently, we showed that the method could be applied to the DSE of UBA(2), a three-helix bundle from the human homolog of yeast Rad23A (HHR23A),[73] by fusing it to the N-terminus of iso-1-Cyt*c*,[43] demonstrating

that the thermodynamic and kinetic properties of the DSE of any fold can be interrogated with the His-heme loop formation method.

In our previous report on the kinetics and thermodynamics of His-heme loop formation in the DSE of UBA(2), a scaling exponent consistent with a random coil with excluded volume was observed, as was true for our studies of cytochrome c', Cytc', a four-helix bundle.[4,44] While we observed that loop formation occurs with lower probability as the His-heme loop increases in size, a subset of His-heme loops formed under denaturing conditions by these three-helix and four-helix bundles are more favorable than expected. The kinetics of loop breakage for this same subset of His-heme loops in the denatured state is also slower than expected for a random coil.[66] The portions of primary structure containing this thermodynamic and kinetic bias in the DSE, localize near interhelical turns or loops in the tertiary structures of Cytc' and UBA(2),[4,43,44] suggesting that residues near interhelical turns form unusually stable contacts in the DSE in these folds. Furthermore, molecular dynamics (MD) simulations of Cytc' under denaturing conditions revealed that the persistence of turns in this four-helix bundle is stabilized by hydrophobic side chains, particularly aromatic side chains.[4] These results indicate that DSE interactions may be critical for establishing native state topology.

Here, we apply the denatured state His-heme loop formation method to examine the intrinsic DSE conformational properties of the first Ubiquitin-associated domain, UBA(1), from HHR23A,[74] using a UBA(1) – iso-1-Cyt*c* fusion protein (Figure 1). UBA(1) and UBA(2) have identical three-helix bundle native state topologies,[74] but only ~20% amino acid sequence identity. The His-heme loop equilibria of the UBA(1) – iso-1-Cytc fusion protein yield a larger degree of scatter with respect to the JS equation than observed with the UBA(2) – iso-1-Cyt*c* fusion protein. As in our previous work,[4,17,43,44] strong deviations from the properties of a random coil occur along the primary structure near both interhelical turns in the native structure of UBA(1). Denatured state His-heme loop breakage kinetics data show that rate constants for loop breakage, k_b, for histidines probing these segments of the primary structure of UBA(1), are small compared to other His-heme loops. Moreover, we find that the magnitude of k_b for His-heme loops for both UBA(1) and UBA(2) show a correlation with the degree of hydrophobic clustering along the primary structure, consistent with the recent proposal that clustering of hydrophobic groups affects the compactness of the DSE.[46] However, the outliers to this correlation form the most stable His-heme loops and are involved in helix capping, leading to the novel conclusion that helix capping also stabilizes denatured state residual structure.

EXPERIMENTAL PROCEDURES

Cloning the UBA(1) gene into the pRbs_BTR1fuse Plasmid.

The first ubiquitin associated domain, UBA(1), which comprises residues 162 – 204 of HHR23A, was inserted between Phe(-3) and Lys(-2) of the yeast iso-1-Cyt*c* domain using the unique *Eco*RI and *Ngo*MIV restriction sites of the pRbs_BTR1fuse vector[43]. The pGEX-2T plasmid containing the UBA(1) gene was a gift from Juli Feigon.[74] The gene was amplified by PCR with primers that put *Eco*RI and *Ngo*MIV restriction sites, respectively, at the 5' and 3' ends of the gene (Table S1). After purification (QIAquick PCR

purification kit, Qiagen), the PCR product was cloned into pRbs_BTR1fuse using the unique *Eco*RI and *Ngo*MIV restriction sites of the vector, as described previously.[43] The resulting pRbs_BTR1(UBA1_Cc) plasmid was sequenced (Genomics Core Facility, University of Montana) confirming the presence of the fusion protein gene.

Site-Directed Mutagenesis to Prepare Single Histidine Variants of the Fusion Protein.

The 43 residue UBA(1) domain of the fusion protein has a histidine at position 31. A pseudo-wild type (pWT) UBA(1) – iso-1-Cyt*c* variant with a H31N substitution was created to provide a template for preparation of single histidine variants. pRbs_BTR1(UBA1_Cc) with the pWT gene was used to prepare T7H, S11H, E15H, R24H, Y27H, and E35H variants of UBA(1) – iso-1-Cyt*c*. PCR-based site-directed mutagenesis was accomplished with the QuikChange Lightning kit (Agilent) employing the primers in Table S2. All single histidine sites have high solvent accessibility as assessed with the GetArea algorithm (Sealy Center for Structural Biology, University of Texas Medical Branch, http://curie.utmb.edu/getarea.html).[76] The wild type His31 construct also was used for His-heme loop formation studies. A tyrosine to glutamine mutation at position 27 was prepared in the His31 background to examine the role of a local nonpolar side chain on the DSE properties near the second turn of UBA(1). Each variant was sequenced (Eurofins Genomics, Louisville, Kentucky) to confirm its identity.

Preparation of UBA(1) – iso-1-Cyt*c* Variants.

Expression and purification of pWT and single-histidine variants of UBA(1) – iso-1-Cyt*c* used previously described methods.[43] In brief, the pRbs_BTR1(UBA1_Cc) plasmid containing the desired UBA(1) – iso-1-Cyt*c* variant was transformed into competent phage T1 resistant BL21(DE3) *Escherichia coli* cells (New England Biolabs) following the manufacturer's recommended method. After suspending transformed cells from L-ampicillin plates into 0.5 mL of sterile 2xYT media, the cell suspension was used to inoculate 1 L 2xYT cultures in 2.8 L Fernbach flasks containing 100 mg of ampicillin. Cultures were incubated at 37 °C for 30 hours in a shaker incubator set to shake at 150 rpm. After harvesting (Sorvall Lynx 6000, F12 rotor, 5000 rpm. 10 minutes), cell pellets were preserved in a −80 °C or used immediately.

Extraction and purification of protein followed previously described protocols.[43,77–80] In summary, thawed cells were dispersed in 50 mM Tris, 500 mM NaCl, 1 mM EDTA, pH 8.0, 2.5 mM PMSF and lysed with a Qsonica Q700 sonicator. Lysate was cleared (Sorvall Lynx 6000, F14 rotor, 10,00 rpm, 30 minutes), brought to 8% ammonium sulfate and equilibrated overnight at 4 °C with stirring. After removal of precipitated protein impurities (Sorvall Lynx 6000, F14 rotor, 10,000 rpm, 30 minutes), dialysis of the supernatant against MilliQ water (two changes) containing 1 mM Na_2EDTA and 1 mM β-mercaptoethanol (βME) was used to remove the ammonium sulfate. The dialysate was brought to pH 6.0 by addition of 50 mM sodium phosphate monobasic and batch loaded onto CM-sepharose resin (100 mL, GE Healthcare Life Sciences). After loading the resin into a glass column, it was washed with MilliQ water followed by elution of UBA(1) – iso-1-Cyt*c* with a linear gradient (200 mL) from 0 to 0.8 M NaCl with 50 mM sodium phosphate pH 6.0, 1 mM Na_2EDTA, 2 mM βME as the buffer. Centrifuge ultrafiltration was used to transfer purified

UBA(1) – iso-1-Cyt*c* into 25 mM sodium phosphate, pH 6.0, 1 mM Na_2EDTA. If not used immediately, UBA(1) – iso-1-Cyt*c* variants were stored at –80 °C after flash freezing with liquid N_2.

Final purification of UBA(1) – iso-1-Cyt*c* variants involved high resolution cation exchange chromatography (5.0 mL HiTrap SP HP column, ÄKTAprime plus chromatography system, GE Healthcare Life Sciences). After oxidation with $K_3[Fe(CN)_6]$, Sephadex G25 (GE Healthcare Life Sciences, superfine grade) chromatography was used to exchange variants into the desired buffer and to remove excess ferricyanide.

Purity was assessed by SDS-PAGE. The molecular weight of each variant was measured by matrix-assisted laser-desorption time-of-flight (MALDI-TOF) mass spectrometry (Bruker microflex mass spectrometer, Table S3). UBA(1) – iso-1-Cyt*c* variants were characterized by MALDI-TOF mass spectrometry immediately before and after each experiment to ensure that the fusion protein variants had not undergone cleavage.

Guanidine Hydrochloride Denaturation of UBA(1) – iso-1-Cyt*c* Variants.

All UBA(1) – iso-1-Cyt*c* variants were oxidized with $K_2[Fe(CN)_6]$. Sephadex G25 chromatography was used to remove excess $K_2[Fe(CN)_6]$ and exchange the variants into CD buffer (20 mM Tris, 40 mM NaCl, 1 mM Na_2EDTA, pH 7.0), as described above. Protein concentration and degree of oxidation were evaluated using the 339, 526.5 and 541.8 nm isosbestic points and the redox sensitive 550 nm absorbance band.[81] The concentration of the ~6 M GuHCl stock containing CD buffer was evaluated from refractive index measurements.[82] A Hamilton Microlab 500 Titrator was used to titrate a given UBA(1) – iso-1-Cyt*c* variant (4 μM) in ~ 6 M GuHCl containing CD buffer into a 4 μM solution of the variant in CD buffer. The titrator was controlled by the software of our Applied Photophysics Chirascan CD Spectrophotometer, which measured ellipticity at 222 nm (α-helix) and 250 nm (background) after each step of the titration. For each variant, a plot of $\theta_{222corr}$ (θ_{222nm} - θ_{250nm}) versus GuHCl concentration was fit to eq 2,

$$\theta_{222corr} = \theta_D + m_D[GuHCl] + \frac{\theta_N + m_N[GuHCl] - \theta_D - m_D[GuHCl] + \left[(\theta_I - \theta_D - m_D[GuHCl])e^{\frac{m_{NI}[GuHCl] - \Delta G_{NI}^{\circ\prime}(H2O)}{RT}}\right]}{\left[1 + \left(e^{\frac{m_{NI}[GuHCl] - \Delta G_{NI}^{\circ\prime}(H2O)}{RT}}\right)\left(1 + e^{\frac{m_{ID}[GuHCl] - \Delta G_{ID}^{\circ\prime}(H2O)}{RT}}\right)\right]} \quad (2)$$

Eq 2 describes 3-state unfolding and assumes a linear free energy relationship between the free energy of unfolding, ΔG_u, and the concentration of GuHCl.[83,84] In eq 2, θ_N and m_N, and, θ_D and m_D, are the intercepts and slopes of the native and denatured state baselines, respectively. θ_I is the intermediate state baseline. The parameters m_{NI} and m_{ID}, respectively, are the rates of change of the free energy of the native to intermediate (ΔG_{NI}) and intermediate to denatured state (ΔG_{ID}) transitions as a function of GuHCl concentration. $\Delta G_{NI}^{o'}(H_2O)$ and $\Delta G_{ID}^{o'}(H_2O)$ are, respectively, free energies for the native to intermediate and intermediate to denatured state transitions at 0 M GuHCl. Fits assumed $m_N = 0$ because in general the native baselines for UBA(1) – iso-1-Cyt*c* variants were too short to fit m_N reliably. Each variant was subjected to three independent trials.

Denatured State His-Heme Loop Formation Titrations.

A Beckman DU 800 UV-Vis spectrophotometer was used to monitor pH titrations at room temperature (22 ± 1 °C) in the denatured state. Protein concentration was 3 μM. Titrations at 4 M and 6 M GuHCl used a 15 mM NaCl, 5 mM Na_2HPO_4, 1 mM Na_2EDTA buffer, and previously reported methods.[43,85] Absorbance at 398 nm, A_{398}, was used to follow His-heme loop formation. Absorbance at 450 nm, A_{450}, was used as a background wavelength. The apparent pK_a, pK_a(obs), for His-heme loop formation was extracted from fits of $A_{398corr}$ ($A_{398} - A_{450}$) versus pH data to eq 3. In eq 3 n is the proton linkage number for His-heme loop formation, A_{LS} is $A_{398corr}$ at high pH (low spin heme

$$A_{398corr} = \frac{A_{LS} + A_{HS} \times 10^{n[pK_a(obs) - pH]}}{1 + 10^{n[pK_a(obs) - pH]}} \quad (3)$$

with histidine as the ligand) and A_{HS} is $A_{398corr}$ at low pH (high-spin heme with H_2O displacing histidine).

His-heme Loop Breakage Kinetics.

pH jump mixing employed an Applied Photophysics SX20 stopped-flow apparatus to measure the kinetics of His-heme loop breakage at 25 °C. Previously described methods were used.[43] All UBA(1) – iso-1-Cyt*c* variants were oxidized, transferred into MOPS buffer (10 mM MOPS, 40 mM NaCl, 2 mM Na_2EDTA, pH 6.8) and their concentrations evaluated, as described above. Refractive index measurements were used to evaluate the exact concentrations of 6 M or 8 M GuHCl stocks containing MOPS buffer, as described above. Solutions of 4 M or 6 M GuHCl containing MOPS buffer and 6 μM protein were prepared. They were mixed 1:1 with GuHCl at 4 M or 6 M buffered at pH 3 or pH 3.5 (100 mM citrate) using the SX20 stopped-flow instrument. After mixing, the protein concentration was 3 μM. For the pWT variant, the starting pH was 8 (20 mM Tris, 2 mM Na_2EDTA, pH 8.0). The pH after mixing was assessed by measuring pH after 1:1 manual mixing of the GuHCl citrate and GuHCl MOPS buffer solutions. The dead time of our stopped-flow was found to be 2 ms by reduction of dichlorophenolindophenol with L-ascorbic acid.[86] Before fitting kinetic data, 0.002 s was added to each time point. A single exponential equation was utilized for fitting kinetic data.

RESULTS

GuHCl Denaturation of UBA(1) – iso-1-Cyt*c* Variants.

GuHCl denaturation followed by CD was used to evaluate the global stability of all UBA(1) – iso-1-Cyt*c* variants at pH 7 and 25 °C. Figure 2 shows a plot of $\theta_{222corr}$ versus GuHCl concentration for each of three single histidine variants (T7H, R24H, and His31) and for the pWT variant. The data are consistent with 3-state equilibrium unfolding for all variants of UBA(1) – iso-1-Cyt*c* as previously observed with UBA(2) – iso-1-Cyt*c* variants.[43] Parameters from fits to a 3-state equilibrium unfolding model (eq 2, Experimental Procedures) can be found in Table 1. We assigned the thermodynamic parameters from the initial stage of unfolding to the iso-1-Cyt*c* domain ($\Delta G_{NI}^{o'}(H_2O)$ and m_{NI}) and those of the second step to the UBA(1) domain ($\Delta G_{ID}^{o'}(H_2O)$ and m_{ID}) based on previously

published GuHCl denaturation experiments on the individual domains.[87,88] In previous work on UBA(2)-Cyt*c*,[43] the Cyt*c* domain also unfolded first followed by the UBA(2) domain.

Iso-1-Cyt*c* with all histidines removed except His18 has an unfolding midpoint near 1 M GuHCl and an *m*-value between 3.8 and 4.0 kcal/mol-M,[88] close to the m_{NI} values in Table 1. Previous work on the UBA(2) domain reports an *m*-value near 1.2 kcal/mol^{-1}M^{-1},[87] close to the m_{ID} values in Table 1. As observed for UBA(2) – iso-1-Cyt*c*,[43] the *m*-value for the N→I transition, m_{NI}, assigned to unfolding of the iso-1-Cyt*c* domain increases significantly when the UBA(1) domain has surface histidines compared to m_{NI} for the pWT variant. Each single histidine substitution decreases the midpoint for the native to intermediate transition, C_{mNI}, by 0.3 – 0.6 M. However, the free energy for the native to intermediate transition at 0 M GuHCl, $\Delta G_{NI}^{o'}(H_2O)$, mostly shows little change because of the increase in m_{NI}.

Many of the single histidine variants of UBA(1) – iso-1-Cyt*c* show little to no effect on the UBA(1) domain stability. For the pWT variant, the midpoint for the transition between the intermediate and denatured states, C_{mID}, is 2.3 M. For the E35H, His31/Y27Q and Y27H variants, C_{mID} is within error of this value. For the remaining variants, C_{mID} decreases significantly. The GuHCl unfolding *m*-values for the UBA(1) domain, m_{ID}, remain more or less constant for most variants. However, m_{ID} decreases for the His31 and T7H variants compared to pWT. As a result, $\Delta G_{ID}^{o'}(H_2O)$ decreases significantly for all variants except for the E35H, His31/Y27Q and Y27H variants. The degree of population of the denatured state for all variants at both 4 M GuHCl and 6 M GuHCl (conditions of His-heme loop formation titrations) was evaluated using the parameters in Table 1. In all cases, the denatured state is at least 95% populated at both GuHCl concentrations.

Thermodynamics of Denatured State His-Heme Loop Formation for UBA(1) – iso-1-Cyt*c* Variants.

The His-heme loop formation method was developed as a means to probe the thermodynamics of the denatured state along the length of a protein. Here, we apply the method to UBA(1), a three-helix bundle. The His-heme methodology requires a heme and a single engineered histidine within the fold under study, to form a His-heme loop of a particular size. As with our previous work on the UBA(2) – iso-1-Cyt*c*,[43] we fused UBA(1) to the N-terminus of iso-1-Cyt*c* (Figure 1) to provide a heme. Under denaturing conditions only an engineered histidine in UBA(1) is available, because all the native histidines of iso-1-Cyt*c* have been replaced. Cys14 is the nearest site of heme attachment to the UBA(1) domain in the fusion protein (see Figures 1 and 3). Thus, 16 residues from the primary structure of iso-1-Cyt*c* are contained within each loop formed by a histidine from the UBA(1) domain. Chain stiffness becomes a factor for loop sizes less than 16.[89] Therefore, for comparing the denatured state properties of UBA(1) with those of a random, insertion between Lys(-2) and Phe(-3) is ideal. Our lab has used this insertion site for previous work on denatured state loop formation with homopolymeric amino acid sequences,[3,64–66] providing for straightforward comparison to these data.

Under denaturing conditions, Met80 is a weak ligand for Fe^{3+} that is readily displaced by water,[90] permitting binding of the unique histidine (for pWT, lysine[91]) to the heme. Because the histidine side chain can be protonated, a simple denatured state pH titration allows the

relative stabilities of a set of His-heme loops to be compared through the apparent pK_a, pK_a(obs) obtained from the titration (Figure 3).

The pH titration data for the pWT, T7H, R24H and His31 UBA(1) – iso-1-Cyt*c* variants in 6 M GuHCl are provided in Figure 4. The His-heme loops are fully formed at pH 7 for all variants except T7H, yielding a low spin state for the heme with the peak of the heme Soret band near 408 nm. As pH decreases, and water, a weak field ligand, displaces the strong field ligand (His, Lys) from UBA(1), the heme undergoes a low spin to a high spin transition. Fitting the DSE pH titration data to the Henderson-Hasselbach equation (eq 3, modified to account for proton linkage number, *n*, see Experimental Procedures), allows determination of pK_a(obs) (Table 2), and thus the relative stability of each His-heme loop. The pWT variant has pK_a(obs) near 7.3 (6 M GuHCl) and 6.7 (4 M GuHCl), which sets the upper limit for the pK_a(obs) values that can be measured under denaturing conditions. For single histidine variants, the E35H, His31 and E15H variants have the lowest pK_a(obs) values at both 4 and 6 M GuHCl indicating that these variants form the most stable His-heme loops in the DSE of UBA(1). The data in Table 2 show pK_a(obs) increases with increasing loop size at both 4 and 6 M GuHCl. However, the E15H variant is a notable exception. But, in general under both denaturing conditions, longer His-heme loops are less stable consistent with the JS equation (eq 2). Because the histidine must lose one H^+ to form the His-heme loop (Figure 3), *n* is expected to be close to one, as observed (Table 2).

The equilibrium for histidine-heme loop formation occurs in two steps as denoted by eq 4. First, the histidine ionizes, pK_a($HisH^+$), then the deprotonated histidine binds to the heme, pK_{loop}(His).

$$pK_a(\text{obs}) = pK_a\left(\text{HisH}^+\right) + pK_{loop}(\text{His}) \tag{4}$$

For iso-1-Cyt*c*, pK_a($HisH^+$) is 6.6 for a range of loop sizes and GuHCl concentrations.[85] Thus, the pK_{loop}(His) values in Table 2 were obtained by subtracting 6.6 from pK_a(obs) for each variant. For pK_{loop}(His), a larger negative magnitude corresponds to a His-heme loop that is more stable.

His-Heme Loop Breakage Kinetics in the DSE of UBA(1) – iso-1-Cyt*c*.

His-heme loop breakage and formation kinetics follow a kinetic model involving rapid deprotonation of histidine and subsequent binding of the histidine to the heme.[92] Eq 5 describes the dependence of the observed rate constant, k_{obs},[92] on pH. In eq 5, K_a($HisH^+$) is the acid dissociation constant, and k_f and k_b are rate constants for loop formation and breakage, respectively.

$$k_{obs} = k_b + k_f\left(\frac{K_a\left(\text{HisH}^+\right)}{K_a\left(\text{HisH}^+\right) + \left[\text{H}^+\right]}\right) \tag{5}$$

When pH<<pK_a($HisH^+$), $k_{obs} \approx k_b$. Because pK_a($HisH^+$) is near 6.6, histidine-heme loop breakage kinetics at 4 M and 6 M GuHCl were evaluated by pH jump down to pH 3.0 and 3.5. Histidine-heme loop breakage kinetic traces are consistent with single exponential

kinetics (Figure S1). The k_{obs} values obtained at pH 3.5 and 3.0 are similar for all UBA(1) – iso-1-Cyt*c* variants, confirming that $k_{obs} \approx k_b$ under these conditions (Table S4).

The values of k_b at 6 M GuHCl for the His-heme loops studied for the UBA(1) and UBA(2) domains fused to iso-1-Cyt*c* are plotted versus loop size in Figure 5. The k_b values for His-heme loops containing poly(Ala) inserted between Phe(-2) and Lys(-3) of iso-1-Cyt*c*[66] are also provided in Figure 5. As previously observed with the UBA(2) – iso-1-Cyt*c* fusion protein, the magnitude of k_b varies irregularly versus loop size. By contrast, for the polyalanine-containing His-heme loops (poly(Ala)), which behave as random coils in strong denaturant solution,[66] the magnitude of k_b is largest for small His-heme loops but gradually plateaus to values near 60 s^{-1} at longer loop sizes. The irregular behavior of k_b values as a function of loop size for UBA – iso-1-Cyt*c* fusion proteins indicates that regions of the primary structure of UBA domains are prone to residual or non-random structure in the DSE. In 6 M GuHCl, unusually persistent His-heme loops in 6 M GuHCl are observed for the E27H (~46 s^{-1}) variant of UBA(2), and the His31 (~29 s^{-1}), His31/Y27Q (~29 s^{-1}) and His15 (~25 s^{-1}) variants of UBA(1). These variants also have His-heme loops with high equilibrium stability (Table 2).

DISCUSSION

Thermodynamics of Fusion Protein Denaturant Unfolding.

We previously fused the UBA(2) domain to the N-terminus of iso-1-Cyt*c*, and upon introducing single surface-exposed histidines into UBA(2) discovered that they decreased the C_{mNI} for unfolding of the iso-1-Cyt*c* domain.[43] A similar effect occurs when UBA(1) is fused to iso-1-Cyt*c*. Notably, the C_{mID} for UBA(1) unfolding is ~1.4 M greater than the C_{mNI} for iso-1-Cyt*c* unfolding for most single histidine fusion protein variants (Table 1). Both the UBA(2)[43] and the UBA(1) domains of the respective fusion proteins remain folded after the iso-1-Cyt*c* domain has unfolded. Therefore, folded UBA domains with a surface histidine appear to be able to unfold the iso-1-Cyt*c* domain by binding to the heme. As the length of the denatured state His-heme loop produced by the appended UBA(1) domain increases, the stability of that loop generally decreases. Thus, as His-heme loop size increases, iso-1-Cyt*c* domain stability might be expected to increase. However, because the UBA(1) domain is not fully unfolded when the iso-1-Cyt*c* domain unfolds, a simple correlation of this sort is unlikely and is not observed (Table 1). The high local concentration of an independently folded domain linked to iso-1-Cyt*c* is sufficient to decrease the C_{mNI} of the iso-1-Cyt*c* domain.

For pWT, which does not contain a surface histidine, the C_{mNI} value of 0.9 M for the iso-1Cyt*c* domain is the same as C_m reported for an isolated iso-1-Cyt*c* domain (~1.0 M).[88] The linked UBA(1) domain significantly lowers the *m*-value indicating that interactions between the two domains affect the cooperativity of unfolding of iso-1-Cyt*c*. The single histidine variants of the UBA(1) domain have C_{mNI} values less than pWT (0.23 to 0.55 M, Figure 2 and Table 1). Interestingly, disrupting the weak heme-Met80 bond with a histidine on the surface of an attached folded domain also increases the cooperativity of unfolding (*m*-value increases), so, that the magnitudes of $\Delta G_{NI}^{o'}(H_2O)$ for the variants are similar to

pWT despite the decrease in C_{mNI}. Thus, introduction of a solvent accessible histidine into the UBA(1) domain catalyzes unfolding of the iso-1-Cyt*c* domain.

For wild type UBA(1) which contains His31, the unfolding data were not consistent with 3-state unfolding (see Table 1). The native histidine at position 31 in UBA(1) destabilizes the iso-1-Cyt*c* domain to such an extent that it is completely unfolded at 0 M GuHCl (see Figure 2). The iso-1-Cyt*c* domain of His31/Y27Q variant is also destabilized, but, the unfolding transition of the iso-1-Cyt*c* domain is observable, indicating that Tyr27 stabilizes the binding of the native state of UBA(1) to the iso-1-Cyt*c* heme. Thus, both the placement of the surface histidine within the UBA(1) domain and the residues of the native state of UBA(1) that interact non-covalently with the heme affect the degree to which the linked UBA domain destabilizes iso-1-Cyt*c*.

As observed for UBA(2) – iso-1-Cyt*c*,[43] the impact of single histidines on the UBA(1) domain stability is variable. The C_m for the pWT UBA(1) variant is 2.3 M. Several of the single histidine variants (E35H, Y27H, and T7H) have C_m values within error of pWT. However, the His31, R24H, E15H and S11H variants have C_m values between 0.6 M and 1.4 M. As a result, these latter variants all show a significant decrease in $\Delta G_{ID}^{\circ'}(H_2O)$ of the UBA(1) domain relative to pWT. Moreover, because the values of $\Delta G_{ID}^{o'}(H_2O)$ and m_{ID} for the E35H, His31/Y27Q and Y27H variants are similar to those of pWT, binding to the heme does not necessarily affect the UBA(1) domain stability. For His31 and T7H, m_{ID} decreases, implying binding of a histidine from UBA(1) to the iso-1-Cyt*c* heme can impact the structure of UBA(1). The E15H, R24H and S11H variants have significantly decreased C_{mID} and $\Delta G_{ID}^{\circ'}(H_2O)$ compared to pWT, but m_{ID} remains unchanged, indicating that UBA(1) domains can remain folded but have a decrease in stability caused by the histidine substitution.

Comparison of Denatured State His-Heme Loop Formation for UBA(1) versus UBA(2).

As discussed in the Introduction, denatured state loop stability versus loop size can be compared to the expectations for a random coil using the JS relationship (eq 2).[69] Eq 6 assumes the free energy of loop formation, $\Delta G_{loop}(His)$, is completely entropic.

$$\Delta G_{loop}(His) = ln(10)RTpK_{loop}(His) = -T\Delta S_{loop} \quad (6)$$

Substituting ΔS_{loop} as expressed by eq 1 into eq 6 yields eq 7,

$$pK_{loop}(His) = pK_{loop}(His)_{ref} + \nu_3 log(N) \quad (7)$$

where $pK_{loop}(His)_{ref}$ is $pK_{loop}(His)$ for a loop nominally containing 1 residue and ν_3 is the loop formation scaling exponent. Thus, $pK_{loop}(His)$ is proportional to the logarithm of the loop size, N, with a slope of ν_3. The magnitude of ν_3 ranges from 1.8 to 2.4, for a random coil polymer with excluded volume.[70–72] Of course, $\Delta G_{loop}(His)$ is not entirely entropic because formation of a His-Fe^{3+} bond will have an enthalpic component. However, the enthalpy of His-Fe^{3+} bond formation should not vary significantly between variants and is incorporated into $pK_{loop}(His)_{ref}$. Thus, deviations from the linearity expected from eq 7 represent non-random interactions in the denatured state.

Figure 6 shows the dependence of pK_{loop}(His) on log(N) for UBA domain data obtained in 4 M GuHCl. Figure 6 also includes, previously published data for histidine-heme loop formation with iso-1-Cyt$c$[66] containing variable-length insertions of poly(Ala) at the N-terminus. The poly(Ala) data adhere closely to eq 7. For the UBA domains, the scatter about the best fit line to eq 7 is significant, indicating that segments of primary structure in the UBA domains have non-random behavior. The previously reported pK_{loop}(His) data for the UBA(2) had only a single His-heme loop (E27H variant) that deviated significantly from the best fit line to eq 7 (solid red line, Figure 6), allowing a reasonable fit to eq 7. The slope of the best fit line for UBA(2) is similar to those observed for the poly(Ala) sequences. At 4 M GuHCl, pK_{loop}(His) for the E27H variant of UBA(2) is 0.28 units more negative than expected from the best fit line to eq 7 at 4 M GuHCl. For UBA(1), three of seven variants have unusually strong His-heme loops in 4 M GuHCl (Figure 6). The remaining four pK_{loop}(His) data points for UBA(1) fall close to the best fit line for UBA(2), suggesting that the best fit line for the UBA(2) data provides a good approximation for segments of UBA(1) that approach random coil behavior.

To test the robustness of this approximation, we fit the UBA(2) data to eq 7 without the E27H data point (red dotted line, Figure 6). We also did a combined fit to eq. 7 of the four UBA(1) data points that lie close to the best fit UBA(2) line and the UBA(2) data excluding the E27H data point (blue dotted line, Figure 6). Both data sets yielded best fit lines similar to the fit to the complete UBA(2) data set, but with better correlation coefficients. Similar results were obtained for the pK_{loop}(His) data of the UBA domains at 6 M GuHCl (Figure S2) indicating that the E27H data point for UBA(2) and the E15H, His31, His31/Y27Q and E35H data points for UBA(1) are outliers with respect to random coil behavior. Therefore, we have used the best fit lines to eq 7 to the combined UBA(1) and UBA(2) data excluding these 5 data points, to evaluate the deviation of the unusually stable His-heme loops of UBA(1) from random coil behavior. At 4 M and 6 M GuHCl, the E35H, His31 and E15H variants deviate from this best fit line by –0.47 to –0.86 pK units. All three loops are more stabilizing than expected, suggesting that the local sequence is prone to residual structure. Converting these deviations to free energy units [$\Delta\Delta G_{loop}$(His) = ln(10)RT$_{\Delta}pK_{loop}$(His)] for the 6 M GuHCl data, the E35H and His31 variants form loops that are 0.74 ± 0.01 kcal/mol and 0.63 ± 0.01 kcal/mol more stable than expected, respectively, and the E15H loop is 1.10 ± 0.07 kcal/mol more stable than expected. At 4 M GuHCl, the E35H loop is 0.65 ± 0.07 kcal/mol more stable than expected, and the His31 and E15H loops are 0.9 ± 0.1 kcal/mol and 0.95 ± 0.05 kcal/mol more stable than expected, respectively.

In previous work with UBA(2)[43] and the four-helix bundle Cytc',[4] we have observed that unusually stable loops are formed by histidines adjacent to turns in the native state structure. Molecular dynamics simulations of the DSE of Cytc' indicated that the topologies of the turns were retained through transient hydrophobic clusters. For UBA(2), the unusually stable loop (E27H, stabilized by 0.41 ± 0.03 kcal/mol at 4 M GuHCl and 0.5 ± 0.1 kcal/mol at 6 M GuHCl relative to the combined UBA(1)/UBA(2) random coil reference state) was adjacent to the helix 2/helix 3 turn. In the case of UBA(1), unusually stable denatured state loops occur near both the helix 1/helix 2 turn (E15H) and the helix 2/helix3 turn (His31). The other stable loop (E35H) is one turn into helix 3. For Cytc', which has four long (20 – 24 residue) helices, stable loops adjacent to turns were also up to a turn into the helices.

Thus, stable loops may reflect hydrophobic residues near turns that can induce dynamic hydrophobic clusters that bias the denatured state ensemble toward the native topology.

Relationship of Residual Structure to Loop Breakage Kinetics.

The values of k_b plotted as a function of loop size (Figure 5) vary irregularly. Our previous results with the UBA(2) domain,[43] found that the magnitude of k_b was small for the stable loop formed by the E27H variant. For UBA(1), we observe similar behavior for the stable His-heme loops formed by the E35H, His31 and E15H variants. These three His-heme loops have the smallest k_b values at both 4 M and 6 M GuHCl (Figure 5, Figure S2, Table S4). These findings are consistent with these regions of the primary structure of UBA(1) stabilizing persistent structure in the DSE, compared to the primary structure of other portions of UBA(1).

Our His-heme loop formation study of the four-helix bundle Cytc', showed that His-heme loops with small k_b values were associated with histidines adjacent to the helix1/helix 2 turn and at the ends of the Ω-loop that joins helix 2 to helix 3.[4] MD simulations of the DSE of Cytc' showed that hydrophobic (particularly aromatic) residues make persistent contacts which favor the topology of the turns. Numerous studies demonstrate that aromatic residues cause stabilization of residual structure and promote long range interactions in the DSE.[14,15,93–97] Because our previous work showed that aromatic residues near the histidine involved in the His-heme loop can stabilize the loop by up to 0.7 kcal/mol, we added a Y27Q substitution to the His31 variant. This substitution neither changes pK_{loop}(His) (Table 2), nor k_b for loop breakage (Table S4) significantly. Thus, Tyr27 does not stabilize residual structure near the second interhelical turn of UBA(1).

Our previous work showed that stabilizations of ~0.3 kcal/mol and ~0.7 kcal/mol were achieved at 3 M and 6M GuHCl, respectively, when an Ala→Trp replacement was made three residues from the histidine involved in loop formation.[3] Similarly, when an alanine was replaced with Trp, Tyr or Phe, four residues from the histidine involved in loop formation, stabilizations of 0.4 – 0.5 kcal/mol were observed in 3 M GuHCl. By contrast, a Leu substitution caused minimal loop stabilization.[5] In all cases, the aromatic residue that stabilized the loop was within the His-heme loop, whereas Tyr27 is outside the loop formed by His31. However, the R24H variant which forms a loop that includes Tyr27 is not unusually stable (Figure 6) nor is its k_b for loop breakage unusually small (Figure 5). Tyr 36 is within the loops formed by His31 and His35 (E35H variant) and may contribute to the stability of these loops, perhaps via dynamic hydrophobic clusters in the DSE, which could stabilize the topology of the helix 2/helix 3 turn.

Scaling Properties of UBA(1).

Because of the degree of scatter in the dependence of pK_{loop}(His) on Log(N) for UBA(1) (Figure 6), we used the best fit line to the JS relationship (eq 7) of the data for UBA(1) and UBA(2), excluding the five data points that exhibit non-random behavior in the DSE, to estimate the stabilization of His-heme loops relative to the expectations for random coil behavior. These fits yield $\nu_3 = 2.2 \pm 0.3$ at 4 M GuHCl and $\nu_3 = 2.0 \pm 0.4$ at 6 M GuHCl. Direct fits of the plots of pK_{loop}(His) versus Log(N) for the UBA(1) domain yield scaling

exponents, ν_3, of 3.3 ± 1.1 and 2.7 ± 1.2 at 4 and 6 M GuHCl, respectively. The errors in these values because of the scatter in the data provide a poor estimate of the scaling behavior of UBA(1). We have observed in our previous studies that much of the non-random behavior of His-heme loop formation results from k_b. Plots of the logarithm of the rate constant for loop formation, k_f (Table S5), versus loop size, N, for the UBA(1) domain show less scatter (Figure S3) and yield values for ν_3 of 2.5 ± 0.4 at 4 M GuHCl and 1.7 ± 0.4 at 6 M GuHCl for the UBA(1) data. These values are within the bounds expected for a random coil with excluded volume (1.8 – 2.4).[70–72] A number of studies,[98,99] including our work on His-heme loop formation,[4,43,92] have shown that the presence of local and long-range residual structure is compatible with random coil scaling properties.

Data in Figure 6 and Figure S3 show that the magnitudes pK_{loop}(His) and Log(k_f) for the UBA(1) data at 4 M GuHCl, respectively, are similar to those for poly(Ala) segments measured in 6 M GuHCl. This observation indicates that the denatured state of UBA(1) is more expanded than that of poly(Ala). In Figure S3, Log(k_f) for poly(Gln) segments at 6 M GuHCl also are plotted and have values that are similar to those for UBA(1) in 6 M GdnHCl. The Flory characteristic ratio, C_n, a measure of chain stiffness, is larger for poly(Gln) segments than for poly(Ala) segments,[64] suggesting that the DSE of UBA(1) has relatively high chain stiffness. We have shown that mutations expected to decrease the helical propensity of the third helix of Cyt*c*′ lead to more negative values of pK_{loop}(His) and larger values of k_f. Thus, residual helical structure in the denatured state can enhance chain stiffness producing an apparently more expanded DSE. The three-helix bundle engrailed homeodomain, has an expanded denatured state that results from residual helical structure.[100,101] The same may be the case for the HHR23A UBA domains.

Rationale for differences in patterns of His-heme loop persistence.

Persistent loops were observed near both turns of UBA(1). However, for both the three-helix bundle UBA(2) and the four-helix bundle Cyt*c*′, His-heme loops containing persistent residual structure are not observed for histidine probes near every interhelical turn, indicating that the DSE does not need to be biased near all interhelical turns to provide for efficient formation of the native state structure. Figure 7 shows plots, for both UBA domains, of normalized Kyte-Doolittle hydrophobicity[102] and hybrid Miyazawa-Jernigan/Hopp-Woods (MJHW) hydrophobicity, a scale recently introduced to identify hydrophobic clusters.[46] Overall, the average hydrophobicity of UBA(1) is slightly higher than that of UBA(2). However, the degree of hydrophobic clustering, as measured by HpC (the area under the portions of the MJHW plots above 0, normalized for sequence length),[46] is much higher for UBA(2) (0.088) than for UBA(1) (0.051) because of a prominent hydrophobic cluster in helix 2 of UBA(2). Figure S4 shows that the Kyte-Doolittle hydrophobicity is more or less constant for the single His variants of UBA(1) – iso-1-Cyt*c*. However, HpC determined with the MJHW hydrophobicity scale, is more significantly affected by the single-histidine substitutions, particularly for helix 2. For UBA(2) – iso-1-Cyt*c*, the effects of the histidine substitutions are minor for both hydrophobicity scales (Figure S5).

For UBA(1) – iso-1-Cyt*c*, the HpC of 0.51 for the WT (His31) protein is near the lower end of the distribution expected for a randomized sequence.[46] For UBA(2) – iso-1-Cyt*c*, the HpC

for pWT is near the upper end of the expected distribution.[46] There are also clear differences in the patterns of hydrophobicity between the two UBA domains. UBA(1) has peaks in hydrophobicity for all three helices, whereas UBA(2) only has peaks in hydrophobicity for helices 2 and 3. UBA(2) has much stronger hydrophobicity in helix 2 than UBA(1), the difference being more pronounced for the MJHW scale.

In all cases, the most stable His-heme loops sit between two hydrophobic peaks, suggesting that such positions are optimal for His-heme loop formation to sense hydrophobic regions on either side of a turn that could promote turn formation in the DSE. For UBA(1), His15 is midway between the helix 1 and helix 2 hydrophobic peaks and His31 (and to a lesser extent His35) is midway between the helix 2 and helix 3 hydrophobic peaks. For UBA(2), His27 is midway between the helix 2 and helix 3 hydrophobic peaks. The lack of a strong hydrophobic peak for helix 1 of UBA(2) may explain why single His variants near turn 1 of UBA(2) do not form persistent His-heme loops.

While the most stable loops appear to have hydrophobic clusters on either side, the observation that the histidines that form the most stable loops are in polar regions of the primary sequence is somewhat unsatisfying. This property is not surprising given that turns can be identified from minima in hydropathy plots.[103] So, we also considered a third measure of hydrophobicity, clustering of Ile, Leu and Val side chains. ILV clusters have recently been suggested to stabilize partially unfolded states of proteins.[104] Similar clusters of ILV residues might be expected to stabilize residual structure in denatured states. Using the ProteinTools web server,[105] we generated ILV clusters for UBA(1) and UBA(2). UBA(1) has one large ILV cluster of seven residues that encompasses the entire hydrophobic core, whereas UBA(2) has two ILV clusters one containing 6 residues and a smaller one with two residues (Figure S6). Figure 8 shows the ILV cluster of UBA(1) as it frames each turn of UBA(1). The packing of the ILV cluster around both turns is good, indicating that the ILV cluster could act to stabilize the loops formed by the E15H, His31 and E35H variants. For UBA(2) (Figure S7), the ILV clusters are not as tightly pack around turn 1 and are relatively distant from turn 2. Thus, the unusual stability of the His-heme loop formed by the E27H variant of UBA(2) cannot be explain by a nearby ILV cluster.

Further inspection of the structure of the UBA domains around the sites of the histidine variants which form highly stable loops show that they are often at sites involved in helix capping.[106] The position of the E15H variant is the N-cap residue of helix 2 of UBA(1) (see Figure 8A). His 31 of UBA(1) is next to the N-cap residue (Pro30) of helix 3. For UBA(2), the Glu27 (E27H variant) provides part of the C-cap of helix 2 (Figure S7B). Hydrophobic interactions are usually considered a predominant cause of residual structure in denatured states of proteins. Analysis of the ILV clusters of the UBA domains indicate that hydrophobic interactions cannot explain all cases of residual structure detected by our His-heme loop formation method. For three of the four His-heme loops (UBA(1), E15H, His31; UBA(2), E27H) which are stabilized by residual structure in the denatured state, helix capping interactions may explain the observed residual structure. The His-heme method may be particularly adept at detecting this type of residual structure because it is known that binding of histidine-containing peptides to hemes can induce helical structure.[107–109]

Detection of DSE compactness with His-heme loop breakage rates.

There has been considerable interest in the factors that affect the compactness of the DSE of foldable proteins and of intrinsically disordered proteins.[45–55] When the DSE is constrained by a His-heme loop, the tendency of the sequence to become more compact should be accentuated. If the DSE near the His that forms the loop is prone to compaction, loop breakage might be expected to slow leading to a smaller magnitude for k_b. The presence of hydrophobic clusters, as measured by HpC,[46] has been proposed to correlate well with compaction of the DSE of a protein. Figure 9 shows a plot of k_b versus HpC for the single His variants of UBA(1). The set of UBA(1) – iso-1-Cyt*c* variants cover a reasonable range of HpC values allowing for a good test of whether k_b correlates with HpC. The E15H and His31 (and His31/Y27Q) variants, which place histidines near both turns of the UBA(1) domain, appear to be outliers. However, the remaining variants have k_b values that correlate well with HpC at both 6 M GuHCl (R = 0.96) and 4 M GuHCl (R = 0.89). This result suggests that variants with pK_{loop}(His) values adhering closely to the JS equation (Figure 6) have k_b values that decrease in response to compaction induced by increased hydrophobic clustering. Because the histidine that forms the loop causes the change in HpC, k_b is sensing the local tendency for hydrophobic clustering. The E35H variant has the largest change in HpC, suggesting that its low k_b is not caused by its location near a turn, but is because of the high local HpC. For the E15H, His31 and His31/Y27Q variants, the large deviation from the k_b versus HpC correlation provides further support for a role of helix capping in stabilizing residual structure in the denatured state.

Figure S8 shows the correlation of k_b with HpC for the UBA(2) – iso-1-Cyt*c* single His variants. The E27H variant, which forms a persistent His-heme loop and is involved in helix-capping in the native state of UBA(2), is an outlier, demonstrating again the importance of helix capping for promoting residual structure in the denatured state. As for UBA(1), the variants with His-heme loops that adhere closely to the JS equation also have k_b values that correlate with HpC. The range of HpC is smaller for this set of variants and the correlation coefficients are somewhat smaller. The HpC values of UBA(2) – iso-1-Cyt*c* are larger compared to UBA(1) – iso-1-Cyt*c*, yet the range of k_b values is similar. Therefore, it appears that HpC is not the only factor that affects the magnitude of k_b. SAXS and FRET do not generally observe strong effects on the dimension (R_g, R_{ee}) of the DSE above 3 M GuHCl. The ability of His-heme loop formation to detect the conformational properties of a protein sequence at higher GuHCl concentrations likely results from the restriction that the His-heme bond imposes on the DSE ensemble.

Do electrostatics affect His-heme loop formation?

There has been considerable interest in the effects of electrostatics on the thermodynamics of the DSE,[8,9,110–112] and more recently the impact of charge distribution, as measured by the parameter κ,[56,113] on the compactness of disordered states of a protein. The His-heme loop formation measurements presented here were carried out in 4 M and 6 M GuHCl. High salt conditions have been shown to screen electrostatic interactions in the DSE.[111] As expected, no significant correlation of k_b with κ, for the sequence region containing the UBA(1) domain (N-terminus to the heme attachment site) is observed.

However, when a His-heme loop is formed, the denatured state is constrained and it is possible that electrostatic interactions between the segment of the UBA domain near the histidine forming the loop and the heme, may not be fully screened by high concentrations of GuHCl. An effect of electrostatics on the rate of loop breakage might result. The intrinsic pK_a of heme propionates is near 5.[114] At pH 3.0 – 3.5, where loop breakage is measured, the net charge on the heme is +1. Therefore, as the local net charge per residue (NCPR) near the histidine forming the loop becomes more positive, k_b is expected to increase because of the increasingly repulsive interactions. Figures S9 and S10 show plots of local NCPR versus sequence position and of k_b versus local NCPR for UBA(1) – iso-1-Cyt*c* and our previously reported data for UBA(2) – iso-1-Cyt*c*.[43] If the data for histidines adjacent to turn sequences are omitted from the correlation, the correlation between k_b and local NCPR goes in the expected direction.

Thus, it appears that for His-heme loop formation, both HpC and local NCPR affect k_b for His-heme loop breakage because of local chain compaction caused by hydrophobic clustering and local charge interactions with the heme, respectively. However for histidines near turn sequences, k_b appears to be dominated by helix-capping effects.

CONCLUSIONS

Similar to the UBA(2) domain from the C-terminus of HHR23A, we observe global properties of the internal UBA(1) domain from HHR23A under denaturing conditions (4 and 6 M GuHCl) consistent with a random coil with excluded volume (ν_3 = 1.8 – 2.4). Extensive scatter about the best fit line to the JS relationship shows that histidine at positions 15, 31 and 35 near turns between the helices form persistent His-heme loops. This result is consistent with our observations for Cyt*c*′ and UBA(2), that persistent denatured state His-heme loops occur for histidines at or near interhelical turns in their tertiary structures.[4,43] Therefore, in agreement with the four-helix bundle Cyt*c*′ and the three-helix bundle UBA(2), denatured state conformational biases in helical bundles localize to sequences at the interface between turns and helices, which may help govern the gross topology of the DSE of α-folds. Plots of k_b versus HpC indicate that HpC clusters affect the stability of His-heme loops in the denatured state. However, ILV cluster analysis of UBA(1) and UBA(2) show that while residual structure for His-heme loops formed near turns 1 and 2 of UBA(1) could be stabilized by ILV clusters, there is no significant ILV cluster near turn 2 of UBA(2) to account for the stable loop formed by the E27H variant of UBA(2). The sequence positions of the E27H variant of UBA(2) and the E15H and His31 variants of UBA(1) are all involved in helix capping leading to the novel conclusion that helix capping also contributes to residual structure in protein denatured states. This view is reinforced by the observation that these variants are all outliers in the correlation between k_b and HpC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

We thank George Rose for pointing out the correlation between helix capping and the positions of histidines that form unusually stable loops in the denatured state.

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 and is part of the University of Montana Mass Spectrometry Core Facility which was supported by COBRE phase II grant (P20 GM103546) to the Center for Biomolecular Structure and Dynamics. M.J.L. acknowledges the Sloan Indigenous Graduate Partnership of the Alfred P. Sloan Foundation, the National Science Foundation (DEB 0614406) and NSF EPSCoR Track-1 EPS-1101342 (INSTEP 3) for graduate education support. Lisa Spang was supported by an NSF REU grant aimed at providing Native American students with summer research opportunities (CHE-1359440).

ABBREVIATIONS

iso-1-Cyt*c*	iso-1-cytochrome *c*
CD	circular dichroism
GuHCl	guanidine hydrochloride
UBA(1)	human HHR23A ubiquitin-associated domain 1
UBA(1) – iso-1-Cyt*c*	fusion protein with UBA(1) inserted near the N-terminus of iso-1-Cyt*c*
pWT	UBA(1)-iso-1-Cyt*c* carrying a H31N mutation in the UBA(1) domain

REFERENCES

(1). Levinthal C (1969) How to fold graciously, In Mössbaur Spectroscopy in Biological Systems (DeBrunner JTP, and Munck E, Eds.), pp 22–24, University of Illinois Press, Allerton House, Monticello, Illinois.

(2). Zwanzig R, Szabo A, and Bagchi B (1992) Levinthal's paradox, Proc. Natl. Acad. Sci. U.S.A 89, 20–22, 10.1073/pnas.89.1.20 [PubMed: 1729690]

(3). Khan MKA, Miller AL, and Bowler BE (2012) Tryptophan significantly stabilizes His-heme loops only when it is near a loop end, Biochemistry 51, 3586–3595, 10.1021/bi300212a. [PubMed: 22486179]

(4). Dar TA, Schaeffer RD, Daggett V, and Bowler BE (2011) Manifestations of native topology in the denatured state ensemble of *Rhodopseudomonas palustris* cytochrome c', Biochemistry 50, 1029–1041, 10.1021/bi101551h. [PubMed: 21190388]

(5). Finnegan ML, and Bowler BE (2010) 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 403, 495–504, 10.1016/j.jmb.2010.09.004. [PubMed: 20850458]

(6). Trefethen JM, Pace CN, Scholtz JM, and Brems DN (2005) Charge-charge interactions in the denatured state influence the folding kinetics of ribonuclease Sa, Protein Sci. 14, 1934–1938, 10.1110/ps.051401905. [PubMed: 15937282]

(7). Cho J-H, and Raleigh DP (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins, J. Mol. Biol 353, 174–185, 10.1016/j.jmb.2005.08.019. [PubMed: 16165156]

(8). Bowler BE (2007) Thermodynamics of protein denatured states, Mol. BioSyst 3, 88–99, 10.1039/B611895J. [PubMed: 17245488]

(9). Cho J-H, Sato S, and Raleigh DP (2004) 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 338, 827–837, 10.1016/j.jmb.2004.02.073. [PubMed: 15099748]

(10). Cho J-H, Meng W, Sato S, Kim EY, Schindelin H, and Raleigh DP (2014) Energetically significant networks of coupled interactions within an unfolded protein, Proc. Natl. Acad. Sci. U.S.A 111, 12079–12084, 10.1073/pnas.1402054111. [PubMed: 25099351]

(11). McCarney ER, Kohn JE, and Plaxco KW (2005) Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins, Crit. Rev. Biochem. Mol. Biol 40, 181–189, 10.1080/10409230591008143. [PubMed: 16126485]

(12). 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, and Plaxco KW (2004) Random coil behavior and the dimensions of chemically unfolded proteins, Proc. Natl. Acad. Sci. U.S.A 101, 12491–12496, 10.1073/pnas.0403643101. [PubMed: 15314214]

(13). Tanford C (1968) Protein denaturation, Adv. Protein Chem 23, 121–282, 10.1016/S0065-3233(08)60401-5. [PubMed: 4882248]

(14). Klein-Seetharaman J, Oikawa M, Grimshaw SB, Wirmer J, Duchardt E, Ueda T, Imoto T, Smith LJ, Dobson CM, and Schwalbe H (2002) Long-range interactions within a nonnative protein, Science 295, 1719–1722, 10.1126/science.1067680. [PubMed: 11872841]

(15). Wirmer J, Schloerb C, Klein-Seetharaman J, Hirano R, Ueda T, Imoto T, and Schwalbe H (2004) Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations, Angew. Chem. Int. Ed 43, 5780–5785, 10.1002/anie.200460907.

(16). Shortle D, and Ackerman MS (2001) Persistence of native-like topology in a denatured protein in 8 M urea, Science 293, 487–489, 10.1126/science.1060438. [PubMed: 11463915]

(17). Bowler BE (2012) Residual structure in unfolded proteins, Curr. Opin. Struct. Biol 22, 4–13, 10.1016/j.sbi.2011.09.002. [PubMed: 21978577]

(18). Bowler BE (2012) Globular proteins: characterization of the denatured state, In Comprehensive Biophysics (Egelman E, Ed.), pp 72–114, Academic Press, Oxford, 10.1016/B978-0-12-374920-8.00305-2.

(19). Mittag T, and Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles, Curr. Opin. Struct. Biol 17, 3–14, 10.1016/j.sbi.2007.01.009. [PubMed: 17250999]

(20). Marsh JA, Baker JMR, Tollinger M, and Forman-Kay JD (2008) Calculation of Residual Dipolar Couplings from Disordered State Ensembles Using Local Alignment, J. Am. Chem. Soc 130, 7804–7805, 10.1021/ja802220c. [PubMed: 18512919]

(21). Marsh JA, Neale C, Jack FE, Choy W-Y, Lee AY, Crowhurst KA, and Forman-Kay JD (2007) Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure, J. Mol. Biol 367, 1494–1510, 10.1016/j.jmb.2007.01.038. [PubMed: 17320108]

(22). Gillespie JR, and Shortle D (1997) 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 268, 179–184, 10.1006/jmbi.1997.0953.

(23). Gillespie JR, and Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels, J. Mol. Biol 268, 158–169, 10.1006/jmbi.1997.0954. [PubMed: 9149149]

(24). Schwalbe H, Fiebig KM, Buck M, Jones JA, Grimshaw SB, Spencer A, Glaser SJ, Smith LJ, and Dobson CM (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea, Biochemistry 36, 8977–8991, 10.1021/bi970049q. [PubMed: 9220986]

(25). Neri D, Billeter M, Wider G, and Wüthrich K (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor, Science 257, 1559–1563, 10.1126/science.1523410. [PubMed: 1523410]

(26). Zhang O, and Forman-Kay JD (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions, Biochemistry 36, 3959–3970, 10.1021/bi9627626. [PubMed: 9092826]

(27). Dill KA, and Shortle D (1991) Denatured states of proteins, Annu. Rev. Biochem 60, 795–825, 10.1146/annurev.bi.60.070191.004051. [PubMed: 1883209]

(28). Dyer RB (2007) Ultrafast and downhill folding, Curr. Opin. Struct. Biol 17, 38–57, 10.1016/j.sbi.2007.01.001. [PubMed: 17223539]

(29). Tang Y, Rigotti DJ, Fairman R, and Raleigh DP (2004) Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: the villin headpiece subdomain, Biochemistry 43, 3264–3272, 10.1021/bi035652p. [PubMed: 15023077]

(30). Luisi DL, Wu W-J, and Raleigh DP (1999) Conformational analysis of a set of peptides corresponding to the entire primary sequence of the N-terminal domain of the ribosomal protein L9: evidence for stable native-like secondary structure in the unfolded state, J. Mol. Biol 287, 395–407, 10.1006/jmbi.1999.2595. [PubMed: 10080901]

(31). Eliezer D (2007) Characterizing residual structure in disordered protein states using nuclear magnetic resonance, Methods in Molecular Biology (Totowa, NJ, United States) 350, 49–67, 10.1385/1-59745-189-4:49. [PubMed: 16957317]

(32). Dyson HJ, and Wright PE (2004) Unfolded proteins and protein folding studied by NMR, Chem. Rev 104, 3607–3622, 10.1021/cr030403s. [PubMed: 15303830]

(33). Marsh JA, Singh VK, Jia Z, and Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation, Protein Sci. 15, 2795–2804, 10.1110/ps.062465306. [PubMed: 17088319]

(34). Ozenne V, Noel JK, Heidarsson PO, Brander S, Poulsen FM, Jensen MR, Kragelund BB, Blackledge M, and Danielsson J (2014) Exploring the minimally frustrated energy landscape of unfolded ACBP, J. Mol. Biol 426, 722–734, 10.1016/j.jmb.2013.10.031. [PubMed: 24211721]

(35). Modig K, Jürgensen VW, Lindorff-Larsen K, Fieber W, Bohr HG, and Poulsen FM (2007) Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis, FEBS Lett. 581, 4965–4971, 10.1016/j.febslet.2007.09.027. [PubMed: 17910956]

(36). Teilum K, Kragelund BB, and Poulsen FM (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, J. Mol. Biol 324, 349–357, 10.1016/S0022-2836(02)01039-2. [PubMed: 12441112]

(37). Lindorff-Larsen K, Kristjansdottir S, Teilum K, Fieber W, Dobson CM, Poulsen FM, and Vendruscolo M (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein, J. Am. Chem. Soc 126, 3291–3299, 10.1021/ja039250g. [PubMed: 15012160]

(38). Kristjansdottir S, Lindorff-Larsen K, Fieber W, Dobson CM, Vendruscolo M, and Poulsen FM (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies, J. Mol. Biol 347, 1053–1062, 10.1016/j.jmb.2005.01.009. [PubMed: 15784263]

(39). Bruun SW, Iešmantavičius V, Danielsson J, and Poulsen FM (2010) Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein, Proc. Natl. Acad. Sci. U.S.A 107, 13306–13311, 10.1073/pnas.1003004107. [PubMed: 20624986]

(40). Bai Y, Karimi A, Dyson HJ, and Wright PE (1997) Absence of a stable intermediate on the folding pathway of protein A, Protein Sci. 6, 1449–1457, 10.1002/pro.5560060709. [PubMed: 9232646]

(41). Myers JK, and Oas TG (2001) Preorganized secondary structure as an important determinant of fast protein folding, Nat. Struct. Biol 8, 552–558, 10.1038/88626. [PubMed: 11373626]

(42). Shea J-E, Onuchic JN, and Brooks CL (1999) Exploring the origins of topological frustration: design of a minimally frustrated model of fragment B of protein A, Proc. Natl. Acad. Sci. U.S.A 96, 12512–12517, 10.1073/pnas.96.22.12512. [PubMed: 10535953]

(43). Leavens MJ, Cherney MM, Finnegan ML, and Bowler BE (2018) Probing denatured state conformational bias in a three-helix bundle, UBA(2), using a cytochrome c fusion protein, Biochemistry 57, 1711–1721, 10.1021/acs.biochem.8b00015. [PubMed: 29480716]

(44). Rao KS, Tzul FO, Christian AK, Gordon TN, and Bowler BE (2009) Thermodynamics of loop formation in the denatured state of *Rhodopseudomonas palustris* cytochrome c': scaling exponents and the reconciliation problem, J. Mol. Biol 392, 1315–1325, 10.1016/j.jmb.2009.07.074. [PubMed: 19647747]

(45). Clark PL, Plaxco KW, and Sosnick TR (2020) Water as a good solvent for unfolded proteins: folding and collapse are fundamentally different, J. Mol. Biol 432, 2882–2889, 10.1016/j.jmb.2020.01.031. [PubMed: 32044346]

(46). Bowman MA, Riback JA, Rodriguez A, Guo H, Li J, Sosnick TR, and Clark PL (2020) Properties of protein unfolded states suggest broad selection for expanded conformational ensembles, Proc. Natl. Acad. Sci. U.S.A 117, 23356–23364, 10.1073/pnas.2003773117. [PubMed: 32879005]

(47). Riback JA, Bowman MA, Zmyslowski A, Knoverek CR, Jumper J, Kaye EB, Freed KF, Clark PL, and Sosnick TR (2018) Response to comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water", Science 361, eaar7949, 10.1126/science.aar7949. [PubMed: 30166460]

(48). Riback JA, Bowman MA, Zmyslowski AM, Knoverek CR, Jumper John M., Hinshaw JR, Kaye, Freed KF, Clark PL, and Sosnick TR (2017) Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water, Science 358, 238–241, 10.1126/science.aan5774. [PubMed: 29026044]

(49). Borgia A, Zheng W, Buholzer K, Borgia MB, Schüler A, Hofmann H, Soranno A, Nettels D, Gast K, Grishaev A, Best RB, and Schuler B (2016) Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods, J. Am. Chem. Soc 138, 11714–11726, 10.1021/jacs.6b05917. [PubMed: 27583570]

(50). Hofmann H, Soranno A, Borgia A, Gast K, Nettels D, and Schuler B (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy, Proc. Natl. Acad. Sci. U.S.A 109, 16155–16160, 10.1073/pnas.1207719109. [PubMed: 22984159]

(51). Best RB, Zheng W, Borgia A, Buholzer K, Borgia MB, Hofmann H, Soranno A, Nettels D, Gast K, Grishaev A, and Schuler B (2018) Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water", Science 361, eaar7101, 10.1126/science.aar7101. [PubMed: 30166459]

(52). Best RB (2020) Emerging consensus on the collapse of unfolded and intrinsically disordered proteins in water, Curr. Opin. Struct. Biol 60, 27–38, 10.1016/j.sbi.2019.10.009. [PubMed: 31805437]

(53). Peran I, Holehouse AS, Carricoa IS, Pappu RV, Bilsel O, and Raleigh DP (2019) Unfolded states under folding conditions accommodate sequence-specific conformational preferences with random coil-like dimensions, Proc. Natl. Acad. Sci. U.S.A 116, 12301–12310, 10.1073/pnas.1818206116. [PubMed: 31167941]

(54). Fuertes G, Banterle N, Ruff KM, Chowdhury A, Pappu RV, Svergun DI, and Lemke EA (2018) Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water", Science 361, eaau8230, 10.1126/science.aau8230. [PubMed: 30166461]

(55). Fuertes G, Banterle N, Ruff KM, Chowdhury A, Mercadante D, Koehler C, Kachala M, Estrada Girona G, Milles S, Mishra A, Onck PR, Gräter F, Esteban-Martín S, Pappu RV, Svergun DI, and Lemke EA (2017) Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements, Proc. Natl. Acad. Sci. U.S.A 114, E6342–E6351, 10.1073/pnas.1704692114. [PubMed: 28716919]

(56). Das RK, and Pappu RV (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues, Proc. Natl. Acad. Sci. U.S.A 110, 13392–13397. [PubMed: 23901099]

(57). Ke PC, Zhou R, Serpell LC, Riek R, Knowles TPJ, Lashuel HA, Gazit E, Hamley IW, Davis TP, Fändrich M, Otzen DE, Chapman MR, Dobson CM, Eisenberg DS, and Mezzenga R (2020) Half a century of amyloids: past, present and future, Chem. Soc. Rev 49, 5473–5509, 10.1039/C9CS00199A. [PubMed: 32632432]

(58). Bowler BE (2008) Thermodynamic approaches to understanding protein denatured states, In Unfolded Proteins: From Denatured to Intrinsically Disordered (Creamer TP, Ed.), pp 23–50, Nova Science Publishers, Hauppage, NY.

(59). Hammack BN, Smith CR, and Bowler BE (2001) Denatured state thermodynamics: residual structure, chain stiffness and scaling factors, J. Mol. Biol 311, 1091–1104, 10.1006/jmbi.2001.4909. [PubMed: 11531342]

(60). Godbole S, and Bowler BE (1997) A histidine variant of yeast iso-1-cytochrome *c* that strongly affects the energetics of the denatured state, J. Mol. Biol 268, 816–821, 10.1006/jmbi.1997.0999. [PubMed: 9180374]

(61). Krieger F, Fierz B, Bieri O, Drewello M, and Kiefhaber T (2003) Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding, J. Mol. Biol 332, 265–274, 10.1016/S0022-2836(03)00892-1. [PubMed: 12946363]

(62). Fierz B (2005) Dynamics of unfolded polypeptide chains, In Protein Folding Handbook (Buchner J, and Kiefhaber T, Eds.), pp 809–855, 10.1002/9783527619498.ch22.

(63). Bieri O, Wirz J, Hellrung B, Schutkowski M, Drewello M, and Kiefhaber T (1999) The speed limit for protein folding measured by triplet-triplet energy transfer, Proc. Natl. Acad. Sci. U.S.A 96, 9597–9601, 10.1073/pnas.96.17.9597. [PubMed: 10449738]

(64). Khan MKA, and Bowler BE (2012) Conformational properties of polyglutamine sequences in guanidine hydrochloride solutions, Biophys. J 103, 1989–1999, 10.1016/j.bpj.2012.09.041. [PubMed: 23199927]

(65). Finnegan ML, and Bowler BE (2012) Scaling properties of glycine-rich sequences in guanidine hydrochloride solutions, Biophys. J 102, 1969–1978, 10.1016/j.bpj.2012.03.049. [PubMed: 22768954]

(66). Tzul FO, and Bowler BE (2010) Denatured states of low complexity polypeptide sequences differ dramatically from those of foldable sequences, Proc. Natl. Acad. Sci. U.S.A 107, 11364–11369, 10.1073/pnas.1004572107. [PubMed: 20534569]

(67). Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, and Hofrichter J (2002) 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 106, 11628–11640, 10.1021/jp020829v.

(68). Lapidus LJ, Eaton WA, and Hofrichter J (2000) Measuring the rate of intramolecular contact formation in polypeptides, Proc. Natl. Acad. Sci. U.S.A 97, 7220–7225, 10.1073/pnas.97.13.7220. [PubMed: 10860987]

(69). Jacobson H, and Stockmayer WH (1950) Intramolecular reaction in polycondensations. I. The theory of linear systems, J. Chem. Phys 18, 1600–1606, 10.1063/1.1747547.

(70). de Gennes P-G (1979) Scaling Concepts in Polymer Physics, Cornell University Press, Ithaca, NY.

(71). Chan HS, and Dill KA (1990) The effect of internal constraints on the configurations of chain molecules, J. Chem. Phys 92, 3118–3135, 10.1063/1.458605.

(72). Redner S (1980) Distribution functions in the interior of polymer chains, J. Phys. A: Math. Gen 13, 3525–3541.

(73). Dieckmann T, Withers-Ward ES, Jarosinski MA, Liu C-F, Chen ISY, and Feigon J (1998) Structure of a human DNA repair protein UBA domain that interacts with HIV-1 Vpr, Nat. Struct. Biol 5, 1042–1047, 10.1038/4220. [PubMed: 9846873]

(74). Mueller TD, and Feigon J (2002) Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions, J. Mol. Biol 319, 1243–1255, 10.1016/S0022-2836(02)00302-9. [PubMed: 12079361]

(75). Berghuis AM, and Brayer GD (1992) Oxidation state-dependent conformational changes in cytochrome *c*, J. Mol. Biol 223, 959–976, 10.1016/0022-2836(92)90255-I. [PubMed: 1311391]

(76). Fraczkiewicz R, and Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, J. Comput. Chem 19, 319–333, 10.1002/(SICI)1096-987X(199802)19:3<319::AIDJCC6>3.0.CO;2-W.

(77). Cherney MM, Junior C, and Bowler BE (2013) Mutation of trimethyllysine-72 to alanine enhances His79-heme mediated dynamics of iso-1-cytochrome *c*, Biochemistry 52, 837–846, 10.1021/bi301599g. [PubMed: 23311346]

(78). Redzic JS, and Bowler BE (2005) Role of hydrogen bond networks and dynamics in positive and negative cooperative stabilization of a protein, Biochemistry 44, 2900–2908, 10.1021/bi048218b. [PubMed: 15723532]

(79). Wandschneider E, Hammack BN, and Bowler BE (2003) Evaluation of cooperative interactions between substructures of iso-1-cytochrome *c* using double mutant cycles, Biochemistry 42, 10659–10666, 10.1021/bi034958t. [PubMed: 12962490]

(80). Goldes ME, Jeakins-Cooley ME, McClelland LJ, Mou T-C, and Bowler BE (2016) Disruption of a hydrogen bond network in human versus spider monkey cytochrome *c* affects heme crevice stability, J. Inorg. Biochem 158, 62–69, 10.1016/j.jinorgbio.2015.12.025. [PubMed: 26775610]

(81). Margoliash E, and Frohwirt N (1959) Spectrum of horse-heart cytochrome *c*, Biochem. J 71, 570–572, 10.1042/bj0710570. [PubMed: 13638266]

(82). Nozaki Y (1972) The preparation of guanidine hydrochloride, Methods Enzymol. 26, 43–50, 10.1016/S0076-6879(72)26005-0. [PubMed: 4680720]

(83). Pace CN, Shirley BA, and Thomson JA (1989) Measuring the conformational stability of a protein, In Protein structure: a practical approach (Creighton TE, Ed.), pp 311–330, IRL Press at Oxford University Press, New York.

(84). Schellman JA (1978) Solvent denaturation, Biopolymers 17, 1305–1322, 10.1002/bip.1978.360170515.

(85). Wandschneider E, and Bowler BE (2004) Conformational properties of the iso-1-cytochrome *c* denatured state: dependence on guanidine hydrochloride concentration, J. Mol. Biol 339, 185–197, 10.1016/S0022-2836(04)00331-6. [PubMed: 15123430]

(86). Tonomura B, Nakatani H, Ohnishi M, Yamaguchi-Ito J, and Hiromi K (1978) Test reactions for a stopped-flow apparatus. Reduction of 2,6-dichlorophenolindophenol and potassium ferricyanide by L-ascorbic acid, Anal. Biochem 84, 370–383, 10.1016/0003-2697(78)90054-4. [PubMed: 626384]

(87). Anil B, Song B, Tang Y, and Raleigh DP (2004) Exploiting the right side of the Ramachandran plot: substitution of glycines by D-alanine can significantly increase protein stability, J. Am. Chem. Soc 126, 13194–13195, 10.1021/ja047119i. [PubMed: 15479052]

(88). Duncan MG, Williams MD, and Bowler BE (2009) Compressing the free energy range of substructure stabilities in iso-1-cytochrome *c*, Protein Sci. 18, 1155–1164, 10.1002/pro.120. [PubMed: 19472325]

(89). Smith CR, Mateljevic N, and Bowler BE (2002) Effects of topology and excluded volume on protein denatured state conformational properties, Biochemistry 41, 10173–10181, 10.1021/bi0259249. [PubMed: 12146983]

(90). Xu Y, Mayne L, and Englander SW (1998) Evidence for an unfolding and refolding pathway in cytochrome *c*, Nat. Struct. Biol 5, 774–778, 10.1038/1810. [PubMed: 9731770]

(91). Smith CR, Wandschneider E, and Bowler BE (2003) Effect of pH on the iso-1-cytochrome *c* denatured state: changing constraints due to heme ligation, Biochemistry 42, 2174–2184, 10.1021/bi026827i. [PubMed: 12590607]

(92). Kurchan E, Roder H, and Bowler BE (2005) Kinetics of loop formation and breakage in the denatured state of iso-1-cytochrome *c*, J. Mol. Biol 353, 730–743, 10.1016/j.jmb.2005.08.034. [PubMed: 16185706]

(93). Crowhurst KA, and Forman-Kay JD (2003) Aromatic and methyl NOEs highlight hydrophobic clustering in the unfolded state of an SH3 domain, Biochemistry 42, 8687–8695, 10.1021/bi034601p. [PubMed: 12873128]

(94). Crowhurst KA, Tollinger M, and Forman-Kay JD (2002) Cooperative interactions and a non-native buried Trp in the unfolded state of an SH3 domain, J. Mol. Biol 322, 163–178, 10.1016/S0022-2836(02)00741-6. [PubMed: 12215422]

(95). Marsh JA, and Forman-Kay JD (2009) Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints, J. Mol. Biol 391, 359–374, 10.1016/j.jmb.2009.06.001. [PubMed: 19501099]

(96). Day R, and Daggett V (2005) Ensemble versus single-molecule protein unfolding, Proc. Natl. Acad. Sci. U.S.A 102, 13445–13450, 10.1073/pnas.0501773102. [PubMed: 16155127]

(97). Wong K-B, Clarke J, Bond CJ, Neira JL, Freund SMV, Fersht AR, and Daggett V (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding, J. Mol. Biol 296, 1257–1282, 10.1006/jmbi.2000.3523. [PubMed: 10698632]

(98). Fitzkee NC, and Rose GD (2004) Reassessing random-coil statistics in unfolded proteins, Proc. Natl. Acad. Sci. U.S.A 101, 12497–12502, 10.1073/pnas.0404236101. [PubMed: 15314216]

(99). Meng W, Lyle N, Luan B, Raleigh DP, and Pappu RV (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure, Proc. Natl. Acad. Sci. U.S.A 110, 2123–2128, 10.1073/pnas.1216979110. [PubMed: 23341588]

(100). DeMarco ML, Alonso DOV, and Daggett V (2004) Diffusing and colliding: the atomic level folding/unfolding pathway of a small helical protein, J. Mol. Biol 341, 1109–1124, 10.1016/j.jmb.2004.06.074. [PubMed: 15328620]

(101). Gianni S, Guydosh NR, Khan F, Caldas TD, Mayor U, White GWN, DeMarco ML, Daggett V, and Fersht AR (2003) Unifying features in protein folding mechanisms, Proc. Natl. Acad. Sci. U.S.A 100, 13286–13291, 10.1073/pnas.1835776100. [PubMed: 14595026]

(102). Kyte J, and Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein, J. Mol. Biol 157, 105–132, 10.1016/0022-2836(82)90515-0. [PubMed: 7108955]

(103). Rose GD (1978) Prediction of chain turns in globular proteins on a hydrophobic basis, Nature 272, 586–590, 10.1038/272586a0. [PubMed: 643051]

(104). Kathuria SV, Chan YH, Nobrega RP, Özen A, and Matthews CR (2016) Clusters of isoleucine, leucine, and valine side chains define cores of stability in high-energy states of globular proteins: sequence determinants of structure and stability, Protein Sci. 25, 662–675, 10.1002/pro.2860. [PubMed: 26660714]

(105). Ferruz N, Schmidt S, and Höcker B (2021) ProteinTools: a toolkit to analyze protein structures, Nucleic Acids Res. 49, W559–W566, 10.1093/nar/gkab375. [PubMed: 34019657]

(106). Aurora R, and Rose GD (1998) Helix capping, Protein Sci. 7, 21–38, 10.1002/pro.5560070103. [PubMed: 9514257]

(107). Cowley AB, Kennedy ML, Silchenko S, Lukat-Rodgers GS, Rodgers KR, and Benson DR (2006) Insight into heme potential redox control and functional aspects of six-coordinate ligand-sensing heme proteins from studies of synthetic heme proteins, Inorg. Chem 45, 9985–10001, 10.1021/ic052205k. [PubMed: 17140194]

(108). Williamson DA, and Benson DR (1998) Remarkable helix stabilization via edge-to-face tryptophan–porphyrin interactions in a peptide-sandwiched mesoheme, Chem. Commun, 961–962, 10.1039/A708368H.

(109). Huffman DL, Rosenblatt MM, and Suslick KS (1998) Synthetic heme-peptide complexes, J. Am. Chem. Soc 120, 6183–6184, 10.1021/ja9704545.

(110). Cho J-H, and Raleigh DP (2006) Electrostatic interactions in the denatured state and in the transition state for protein folding: effects of denatured state interactions on the analysis of transition state structure, J. Mol. Biol 359, 1437–1446, 10.1016/j.jmb.2006.04.038. [PubMed: 16787780]

(111). Kuhlman B, Luisi DL, Young P, and Raleigh DP (1999) 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 38, 4896–4903, 10.1021/bi982931h. [PubMed: 10200179]

(112). Swint-Kruse L, and Robertson AD (1995) Hydrogen bonds and the pH dependence of ovomucoid third domain stability, Biochemistry 34, 4724–4732, 10.1021/bi00014a029. [PubMed: 7718578]

(113). Holehouse AS, Das RK, Ahad JN, Richardson MOG, and Pappu RV (2017) CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins, Biophys. J 112, 16–21, 10.1016/j.bpj.2016.11.3200. [PubMed: 28076807]

(114). Pielak GJ, Auld DS, Betz SF, Hilgen-Willis SE, and Garcia LL (1996) Nuclear magnetic resonance studies of class I cytochromes c, In Cytochrome c: A Multidisciplinary Approach (Scott RA, and Mauk AG, Eds.), pp 203–284, University Science Books, Sausalito, CA.

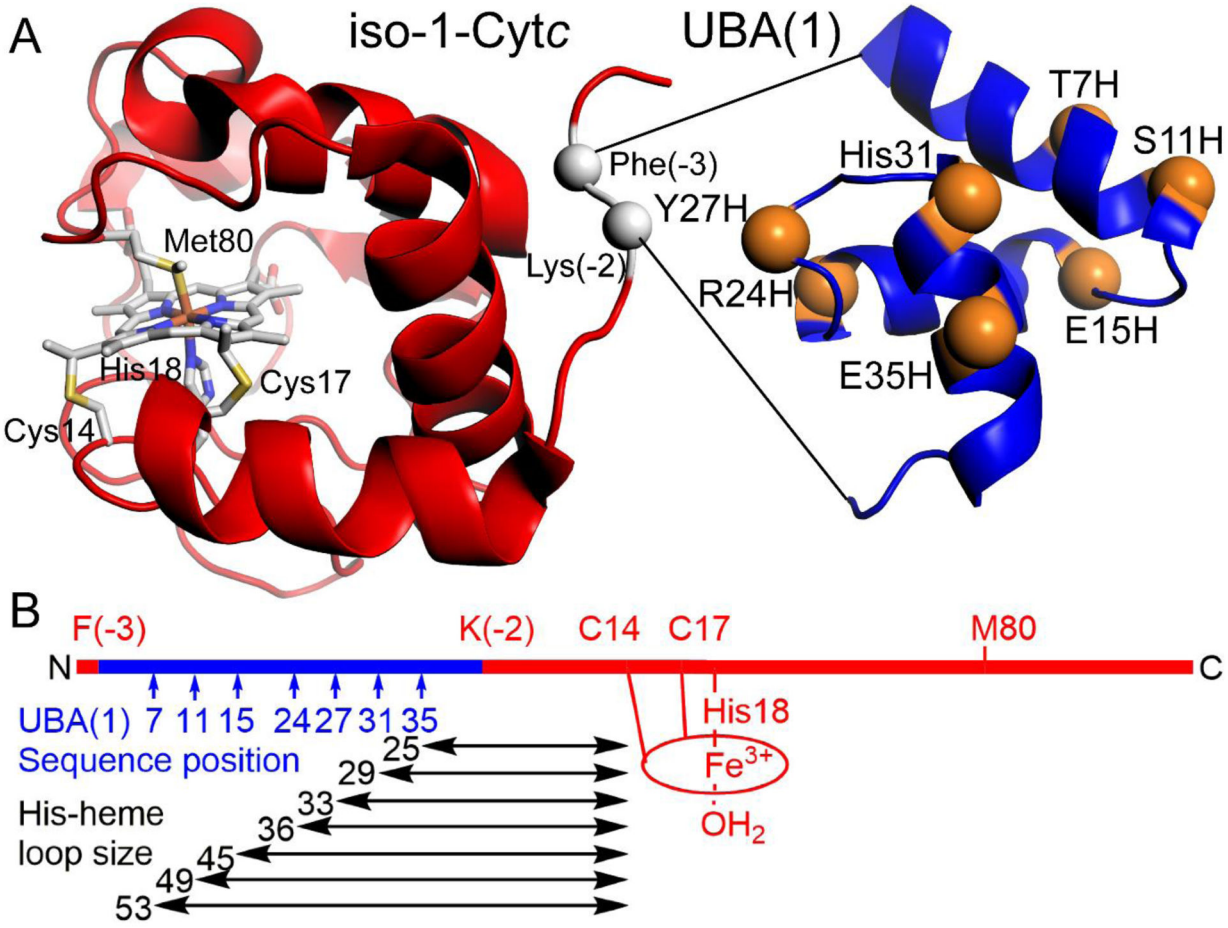


Figure 1.

(A) Yeast iso-1 cytochrome *c* (red, PDB ID: 2YCC)[75] UBA(1) (blue, PDB: 1IFY)[74] fusion protein. UBA(1) is inserted between Phe(-3) and Lys(-2) (white spheres) of iso-1-Cyt*c*. Single histidine substitution sites in UBA(1) are shown with orange spheres. Stick models are used to represent the heme and its points of attachment to iso-1-Cyt*c* at Cys14, Cys17, His18 and Met80. His31 was mutated to Asn for the other single His variants. A Y27Q variant was prepared in the presence of His31. (B) Schematic representation showing how His-heme loop size varies in the denatured state for single histidine variants at positions 7, 11, 15, 24, 27, 31 or 35 in the sequence of UBA(1) when the histidine displaces the water in the 6th coordination site of the heme of the fusion protein.

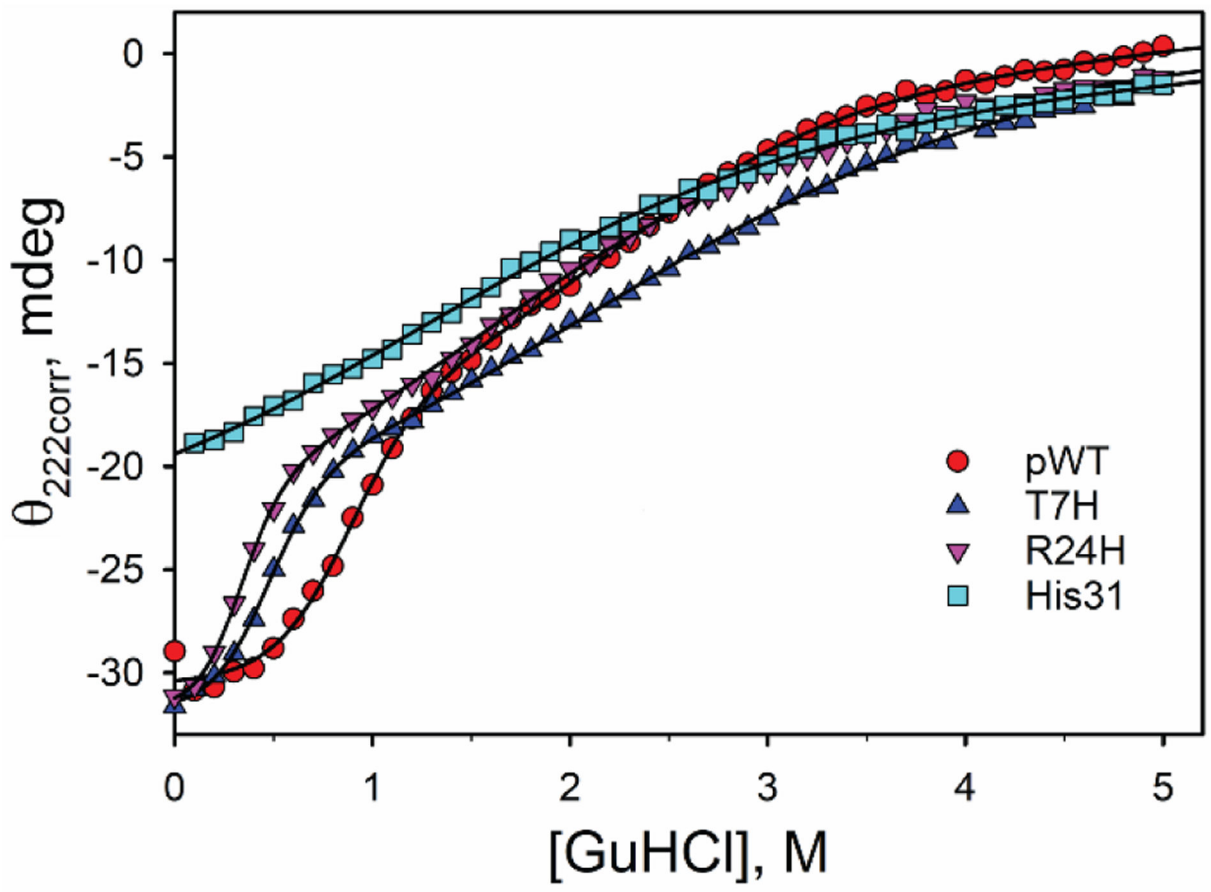


Figure 2.

GuHCl denaturation curves at pH and 25 °C for variants of UBA(1) – iso-1-Cyt*c*, pWT (red circles), T7H (blue triangles), R24H (pink inverted triangles), and His31 (cyan squares). Plots show the dependence of corrected ellipticity at 222 nm, $\theta_{222corr}$ on the concentration of GuHCl. The curves (black) show the fits to eq 2 (Experimental Procedures) except for the His31 (WT) protein, which is fit to a two-state transition. Parameters from the fits can be found in Table 1.

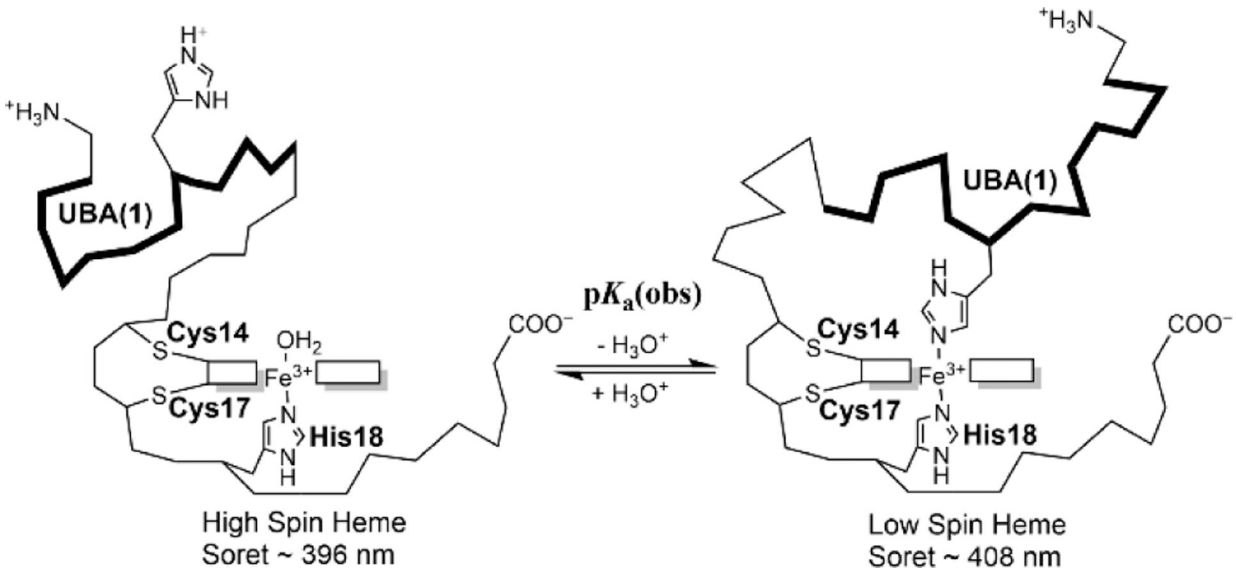


Figure 3.

Loop formation for UBA(1) – iso-1-Cyt*c* shown schematically. At low pH (left) the His-heme loop is broken. At high pH (right), the His-heme loop is formed.

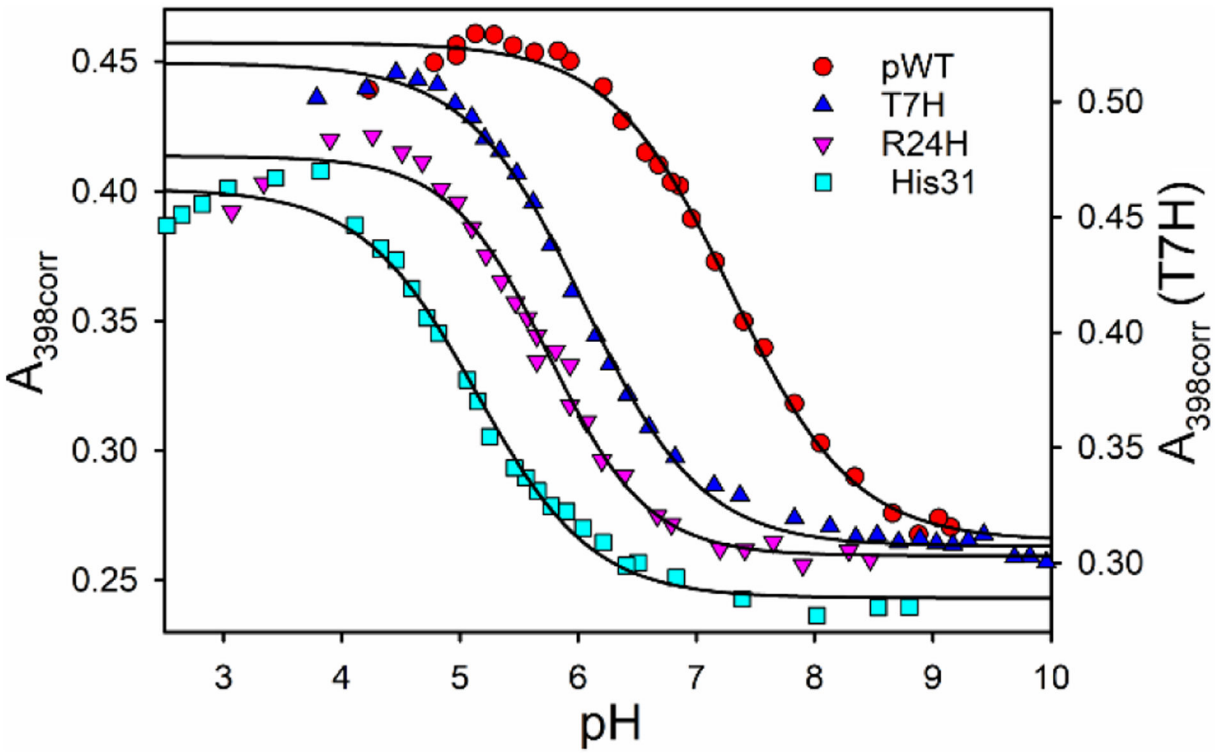


Figure 4.

Denatured state (6 M GuHCl) pH titration data for selected UBA(1) – iso-1-Cyt*c* variants. The dependence of $A_{398corr}$ on pH is shown for the pWT (red circles), T7H (blue triangles), R24H (pink inverted triangles), and His31 (cyan squares) variants. The curves (black) show the fits to eq 3 (Experimental Procedures).

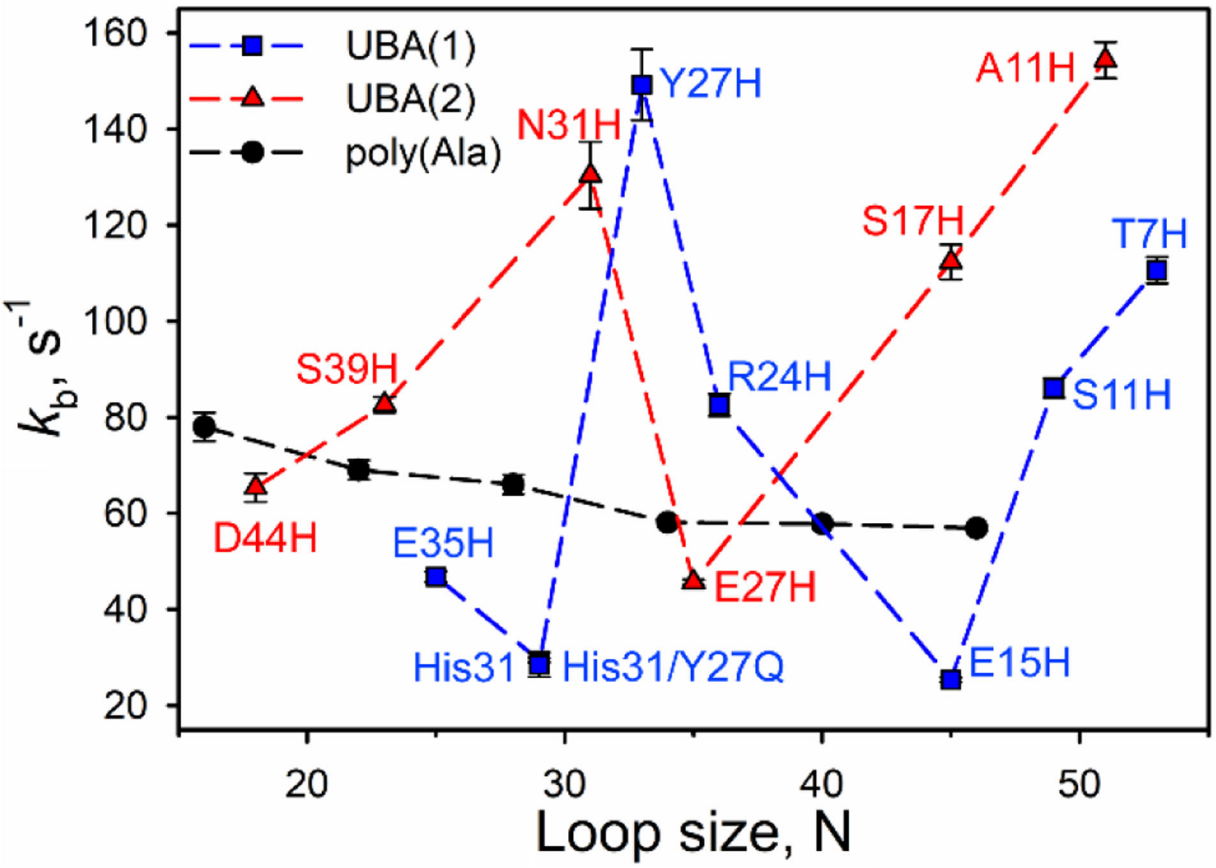


Figure 5.

Values of k_b at 25 °C and 6 M GuHCl plotted against loop size, N, for UBA(1) – iso-1-Cyt*c* (blue squares). Data at 6 M GuHCl for UBA(2) – iso-1-Cyt*c* (red triangles),[43] and alanine inserts $(KAAAAA)_n$[66] (n = 1 – 5, black circles) between Phe(-3) and Lys(-2) of iso-1-Cyt*c* are included for comparison.

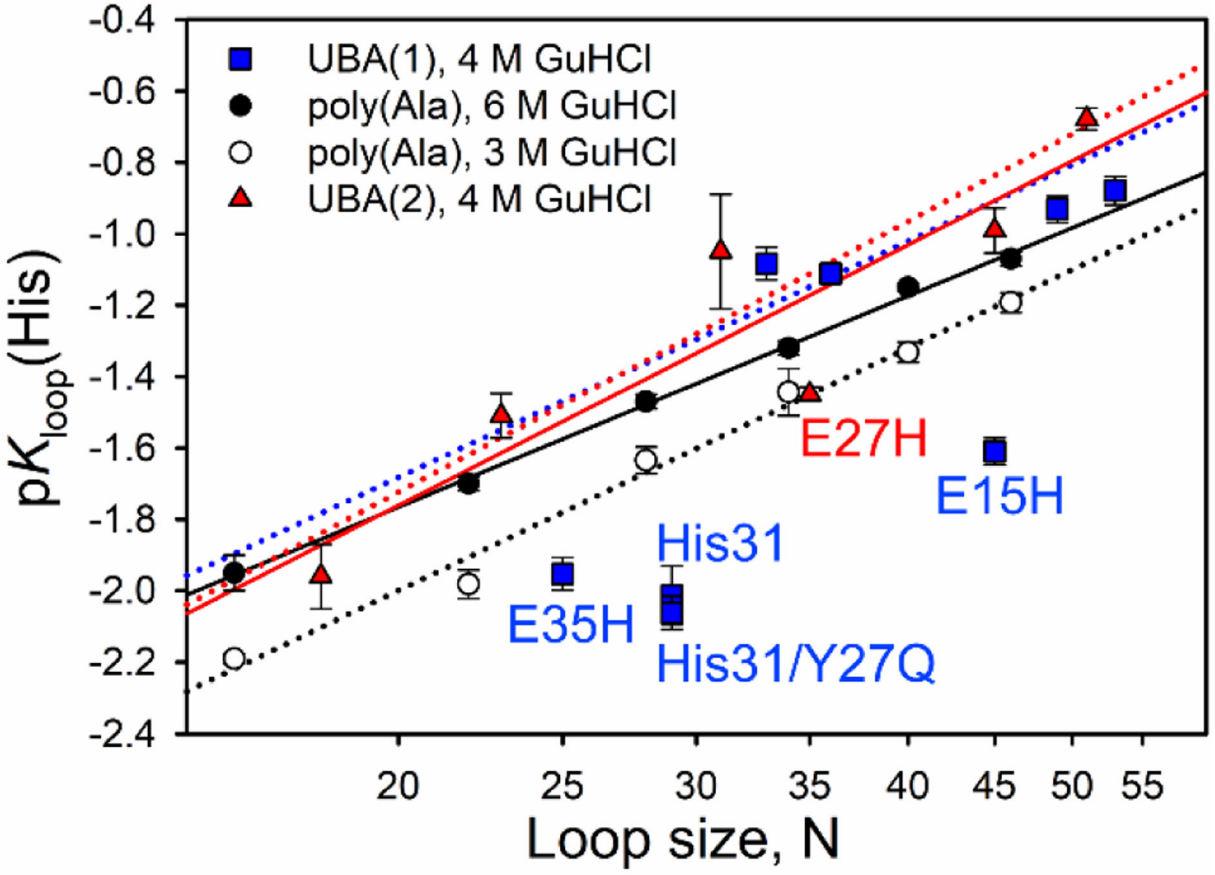


Figure 6.

Loop stability, pK_{loop}(His), as a function of loop size (logarithmic scale) under denaturing conditions (4 M Gu HCl) for the three-helix bundles UBA(1) (blue squares) and UBA(2)[43] (red triangles) inserted at the N-terminus of iso-1-Cyt*c*. pK_{loop}(His) data at 4 M (open circles) and 6 M (black circles) GuHCl for homopolymeric segments of alanine, $(KAAAAA)_n$,[66] inserted into iso-1-Cyt*c* between Phe(-3) and Lys(-2) are displayed for comparison. The UBA(1) and UBA(2) variants with slow loop breakage rates (See Figure 5) are labeled. Fits to eq 7 are shown for the UBA(2) data (red line, R = 0.912), the poly(Ala) data at 4 M (dotted black line) and 6 M (black line) GuHCl. Scaling exponents ν_3 obtained from the fits to eq 7 are 2.4 ± 0.5, 2.3 ± 0.1 and 1.97 ± 0.05, respectively. The dotted red line shows the effect of fitting the UBA(2) data without the E27H data point (ν_3 = 2.5 ± 0.4, R = 0.959). The blue dotted line is a fit to the combined UBA(1) and UBA(2) data with neither the E15H, His31, His31/Y27Q and E35H data points nor the E27H UBA(2) data point included (ν_3 = 2.2 ± 0.3, R = 0.938).

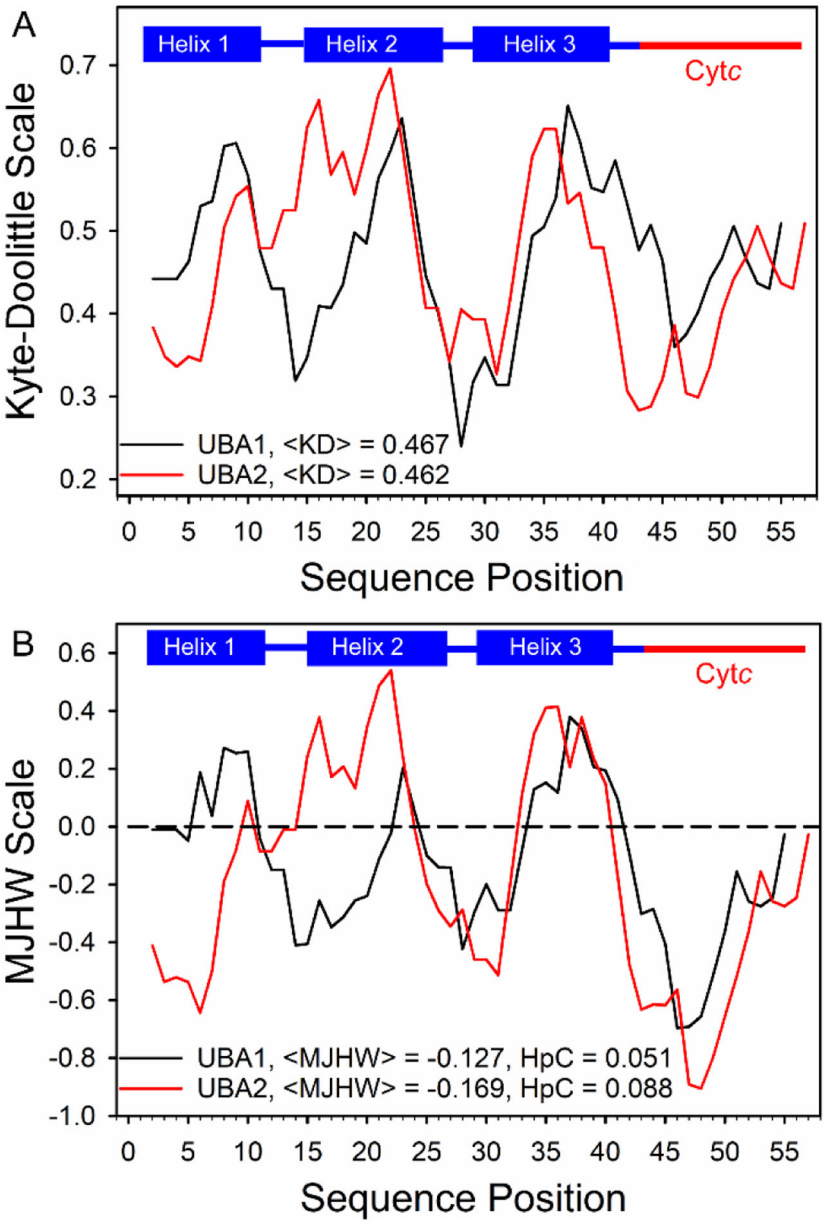


Figure 7.

Plot of hydrophobicity versus sequence position for WT (His31) UBA(1) – iso-1-Cyt*c* (black line) versus pWT (C26A) UBA(2) – iso-1-Cyt*c* for the (A) normalized Kyte-Doolittle and (B) hybrid Miyazawa-Jernigan/Hopp-Woods (MJHW) scales. Hydrophobicity was calculated using a nine-residue window. Reported average values of Kyte-Doolittle hydrophobicity, <KD>, and MJHW hydrophobicity, <MJHW>, and for HpC are for the regions between the N-terminus and Cys14 of the fusion proteins (i.e., the portion of the protein involve in loop formation). The sequence numbering corresponds to that used for the UBA(1) domain in Figure 1. Hydrophobic clusters used to calculate HpC correspond to regions of the MJHW plots with values above zero (dashed line). The secondary structure of the UBA domains in the fusion proteins is shown above the plots. Blue rectangles

correspond to helices and blue lines correspond to the turns. The portion of the sequence derived from Cyt*c* is labeled and denoted with a red line.

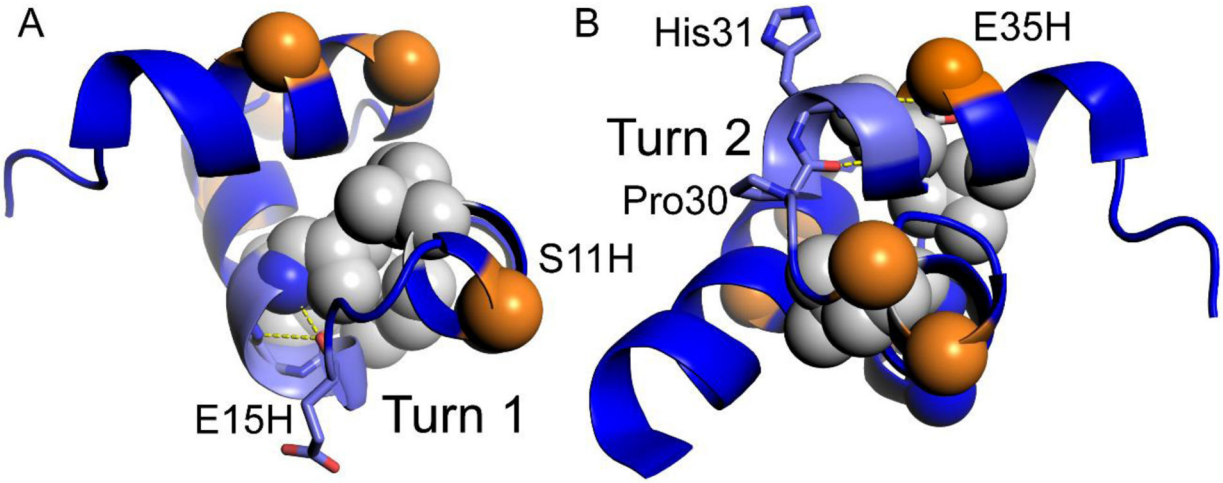


Figure 8.

Portions of the ILV cluster of UBA(1) that pack around each turn of UBA(1). (A) Turn 1 of UBA(1). The hydrogen bonds from the carbonyl of the N-cap residue Glu15 to amide NH atoms of Val19 and Arg18 are shown as yellow dashed lines. (B) Turn 2 of UBA(1). The hydrogen bonds from the carbonyl of the N-cap residue Pro30 to the amide NH of Val34 and from His31 to the amide NH of Glu35 are shown as yellow dashed lines. The positions of the histidine variants are shown with orange spheres.

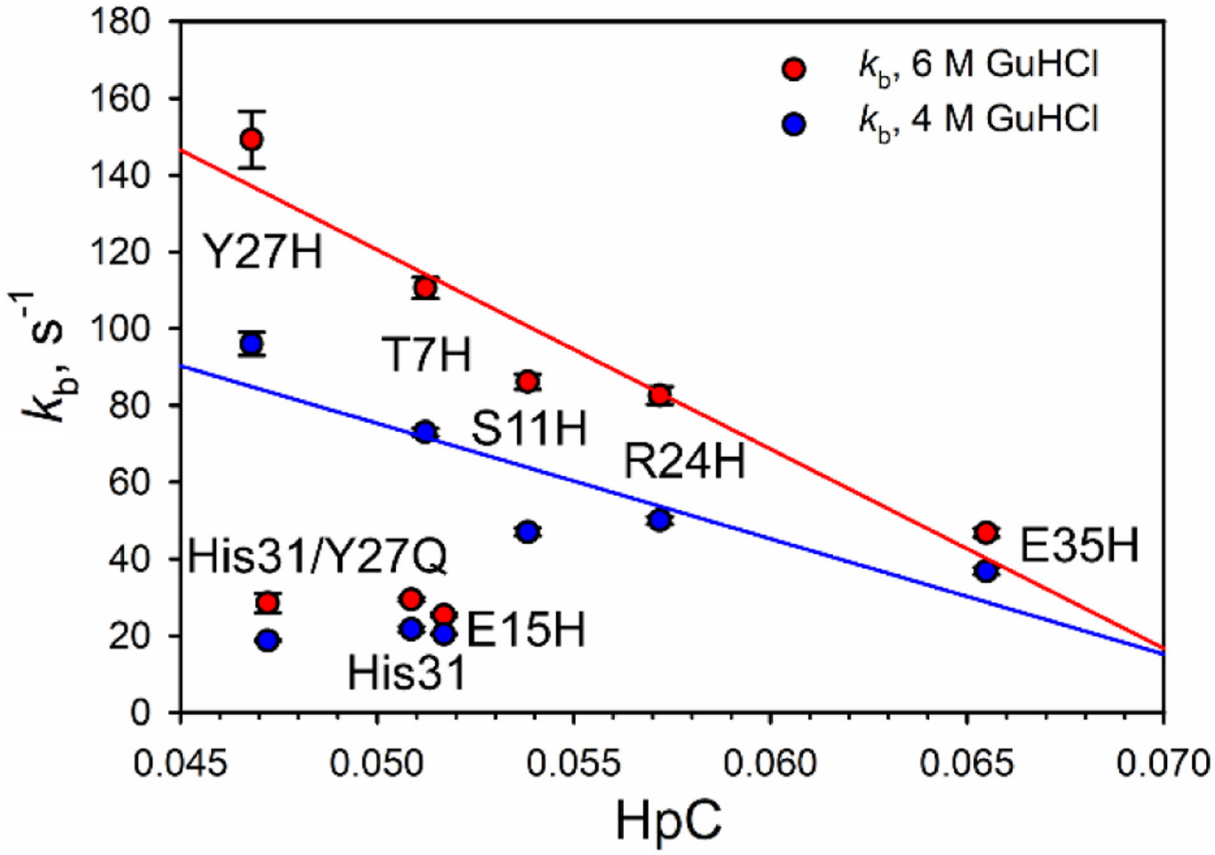


Figure 9.

Plot of k_b for His-heme loop breakage versus the degree of hydrophobic clustering, HpC, of the portion of the UBA(1) – iso-1-Cyt*c* N-terminal to the point of heme attachment. The solid red line shows the correlation of k_b with HpC for loop breakage in 6 M GuHCl (R = 0.96). The solid blue line is the same correlation at 4 M GuHCl (R = 0.89). The correlations do not include the His variants adjacent to turn 1 (E15H) and turn 2 (His31 and His31/Y27Q).

Table 1.

GuHCl Denaturation Parameters for UBA(1) – iso-1-Cyt*c* variants at pH 7.0 and 25 °C.

variant[a]	m_{NI}, kcal mol^{-1} M^{-1}	$\Delta G_{NI}^{\circ\prime}$ (H_2O), kcal mol^{-1}	C_{mNI}, M	m_{ID}, kcal mol^{-1} M^{-1}	$\Delta G_{ID}^{\circ\prime}$ (H_2O), kcal mol^{-1}	C_{mID}, M
pWT	2.6 ± 0.4	2.3 ± 0.4	0.86 ± 0.02	0.9 ± 0.09	2.1 ± 0.2	2.3 ± 0.4
E35H (25)	6 ± 1	2.4 ± 0.6	0.40 ± 0.03	0.83 ± 0.05	1.8 ± 0.1	2.17 ± 0.05
His31 (29)[b]	-	-	-	0.52 ± 0.06	0.63 ± 0.06	1.2 ± 0.1
His31/Y27Q (29)	4.4 ± 0.7	1.1 ± 0.5	0.23 ± 0.07	0.86 ± 0.01	2.19 ± 0.04	2.54 ± 0.03
Y27H (33)	5 ± 1	3.0 ± 0.5	0.55 ± 0.01	0.83 ± 0.07	1.8 ± 0.2	2.14 ± 0.04
R24H (36)	4.8 ± 0.5	1.7 ± 0.2	0.36 ± 0.01	0.7 ± 0.2	1.0 ± 0.3	1.40 ± 0.05
E15H (45)	4.1 ± 0.7	2.0 ± 0.9	0.5 ± 0.1	1.0 ± 0.2	0.6 ± 0.1	0.67 ± 0.03
S11H (49)[c]	4.8 ± 0.5	2.3 ± 0.2	0.48 ± 0.01	0.78 ± 0.04	1.00 ± 0.04	1.3 ± 0.1
T7H (53)	4.8 ± 0.7	2.3 ± 0.3	0.49 ± 0.04	0.4 ± 0.2	0.7 ± 0.4	1.8 ± 0.3

[a] The number in brackets is the His-heme loop size formed in the denatured state.

[b] His31 could only be fit to two-state model.

[c] To fit the S11H data, θ_I was constrained using values observed for the R24H variant as a guide.

Table 2.

Stability and Linkage Parameters for Loop Formation with UBA(1) – iso-1-Cyt*c* Variants at 22 ± 1°C in 6 M and 4 M GuHCl

Variant	Loop size	pK_a(obs)	*n*	pK_{loop}(His)
		6 M GuHCl		
pWT	-	7.31 ± 0.04	0.78 ± 0.06	-
E35H	25	4.94 ± 0.01	1.09 ± 0.06	–1.66 ± 0.01
His31 (WT)	29	5.15 ± 0.01	0.89 ± 0.06	–1.45 ± 0.01
His31/Y27Q	29	5.08 ± 0.02	1.07 ± 0.01	–1.52 ± 0.02
Y27H	33	6.02 ± 0.02	1.01 ± 0.06	–0.58 ± 0.02
R24H	36	5.75 ± 0.01	1.06 ± 0.03	–0.85 ± 0.01
E15H	45	5.19 ± 0.05	0.83 ± 0.05	–1.41 ± 0.05
S11H	49	5.98 ± 0.04	0.94 ± 0.04	–0.63 ± 0.04
T7H	53	6.03 ± 0.02	0.88 ± 0.07	–0.57 ± 0.02
		4 M GuHCl		
pWT	-	6.69 ± 0.05	0.80 ± 0.04	-
E35H	25	4.65 ± 0.05	1.08 ± 0.08	–1.95 ± 0.05
His31 (WT)	29	4.59 ± 0.08	0.95 ± 0.13	–2.01 ± 0.08
His31/Y27Q	29	4.54 ± 0.05	0.90 ± 0.05	–2.06 ± 0.05
Y27H	33	5.52 ± 0.05	1.15 ± 0.06	–1.08 ± 0.05
R24H	36	5.49 ± 0.03	0.99 ± 0.04	–1.11 ± 0.03
E15H	45	4.99 ± 0.04	0.89 ± 0.01	–1.61 ± 0.04
S11H	49	5.67 ± 0.04	0.96 ± 0.05	–0.93 ± 0.04
T7H	53	5.72 ± 0.04	1.10 ± 0.05	–0.88 ± 0.04