

Chapter 3

Biophysical Characterization of Cytochrome *cb*₅₆₂

3.1 Introduction

To prevent heme dissociation during refolding, the porphyrin vinyl groups of cytochrome *cb*₅₆₂'s heme were covalently attached to two mutated cysteine residues, R98C and Y101C, through the conserved CXXCH cytochrome *c*-type motif [37]. The heme environment is not significantly perturbed from cytochrome *b*₅₆₂, and the average RMSD for α -carbons is 0.42 Å [41, 37]. An additional mutation, K59W, aids in expression and can be used as a fluorescent probe. Cytochrome *cb*₅₆₂ folds reversibly and has enhanced stability in denaturant unfolding conditions (folding free energy change of -42 kJ/mol at 0 M Gdn) when compared to cytochrome *b*₅₆₂ (-30 kJ/mol) [37].

In this chapter, we characterize native and unfolded cytochrome *cb*₅₆₂. We monitor how the heme environment changes with denaturation and pH changes by UV-visible absorption spectroscopy. We measure the stability of the protein to denaturation, and we then investigate the unfolded state ensemble conformations

by trFRET. Residues have been selected throughout the protein for labeling with a dansyl (Dns) fluorophore in single-cysteine variants.

3.2 Protein Purification and Characterization

Cytochrome *cb*₅₆₂ was purified and the labeled protein was separated by ion-exchange chromatography on an FPLC system. A representative chromatogram is shown in Figure 3.1. At pH 4.5, Dns-labeled protein is less cationic than unlabeled protein, whereas Ru-labeled protein is significantly more cationic than unlabeled protein. The elution conductivity varies with mutations.

The folded structure and purity were confirmed by UV-visible adsorption spectroscopy, and the mass was verified by electrospray ionization mass spectrometry (ESI-MS). Representative mass spectra for the Dns-labeled (Chapters 3 and 5) and Ru-labeled (Chapters 4 and 9) proteins are shown in Figures 3.2 and 3.3.

3.3 Stability of Labeled Mutants to Denaturation

Solvent-exposed residues that form no intramolecular bonds were selected for photosensitizer labeling to minimize structural perturbations. To confirm that the mutations and Dns labels do not destabilize the protein, denaturation curves were obtained by circular dichroism (CD) and UV-visible spectroscopy. The relative populations of folded and unfolded protein when the protein is denatured at various concentrations of guanidine hydrochloride (Gdn) were measured.

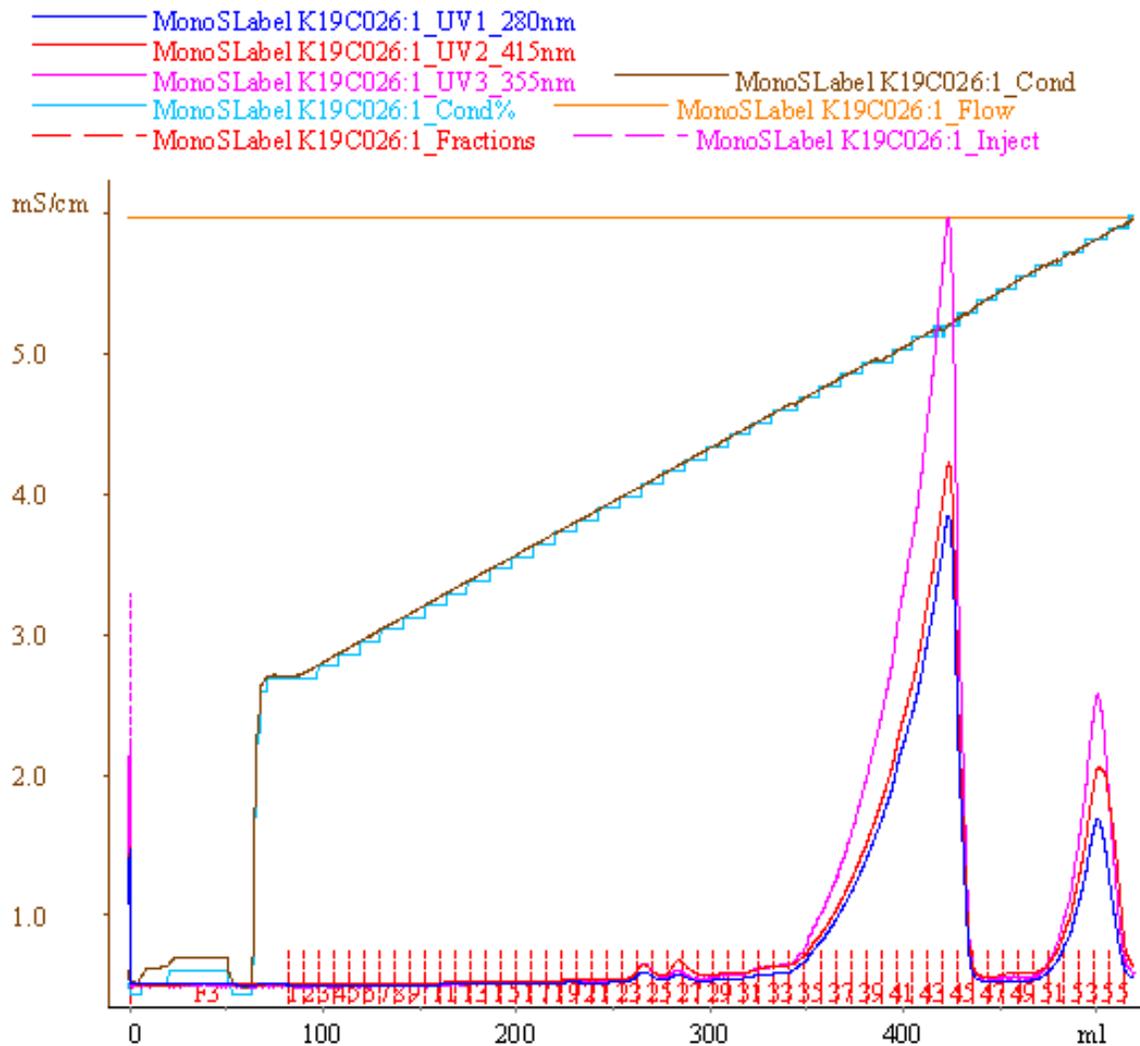


Figure 3.1: Representative FPLC chromatogram. Separation of Dns-labeled (first peak) and unlabeled (second peak) cytochrome cb_{562} (variant K19C) by cation-exchange chromatography (pH 4.5).

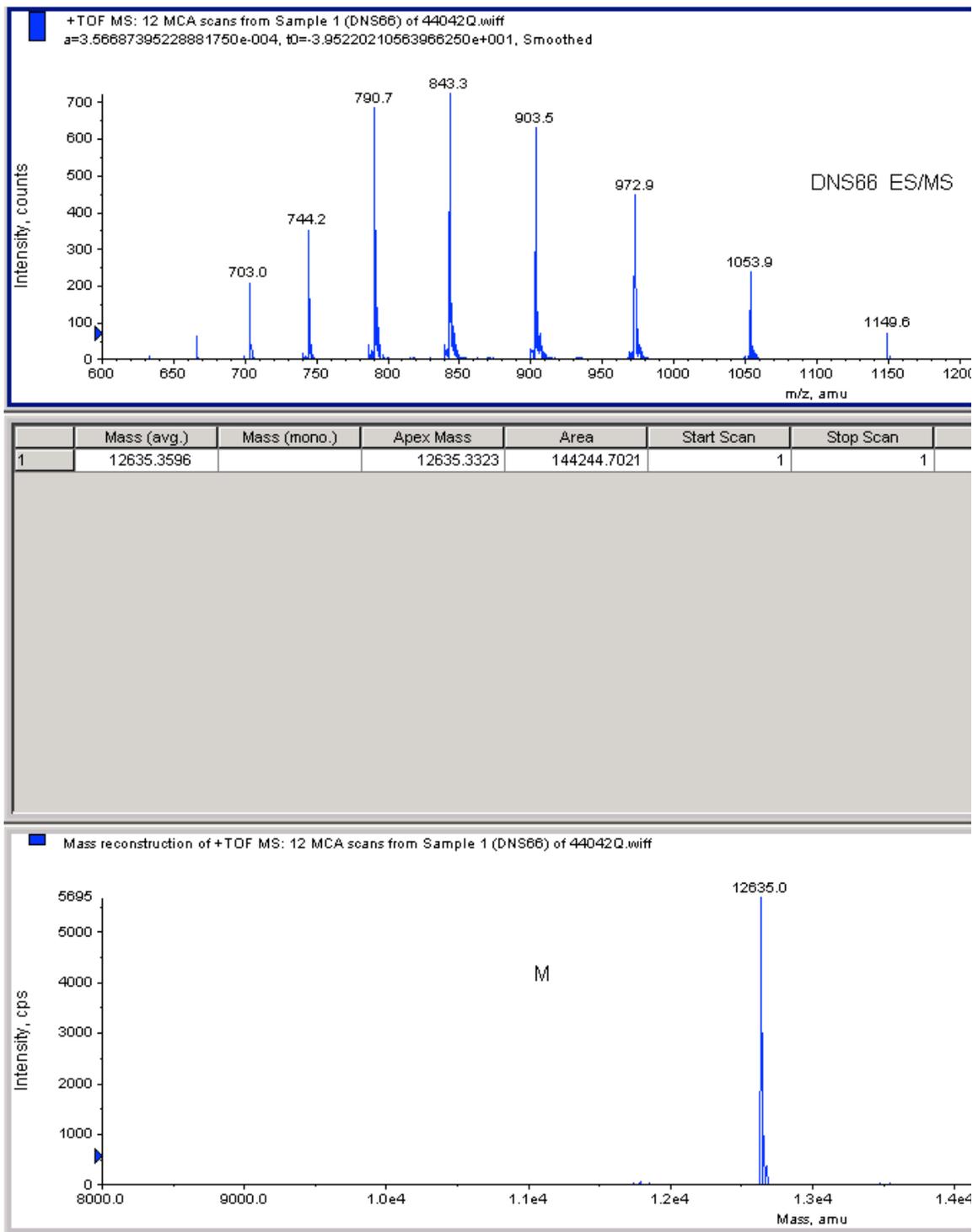


Figure 3.2: Representative electrospray ionization mass spectrum of Dns-labeled cytochrome cb_{562} (variant D66C).

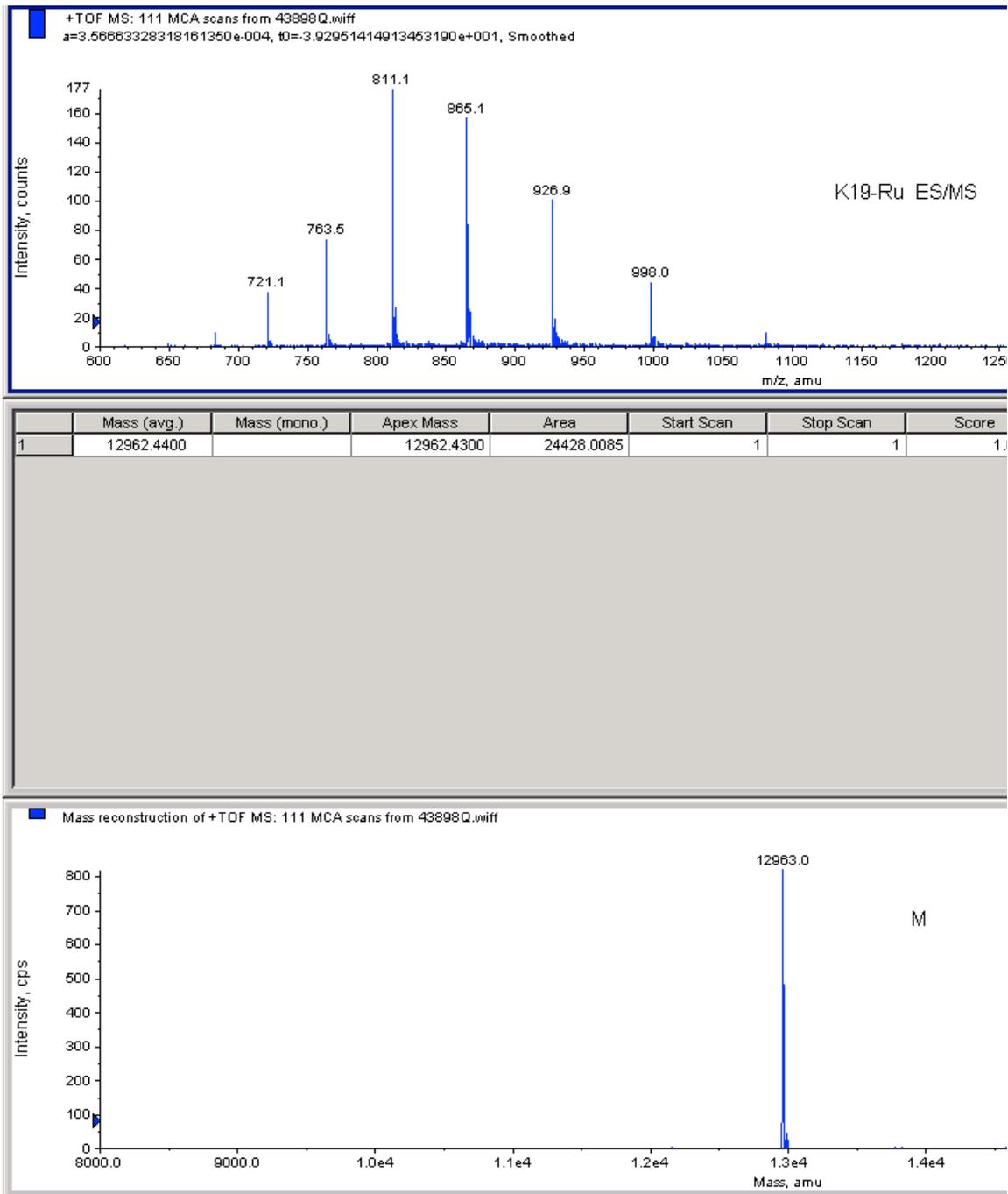


Figure 3.3: Representative electrospray ionization mass spectrum of Ru-labeled cytochrome cb_{562} (variant K19C).

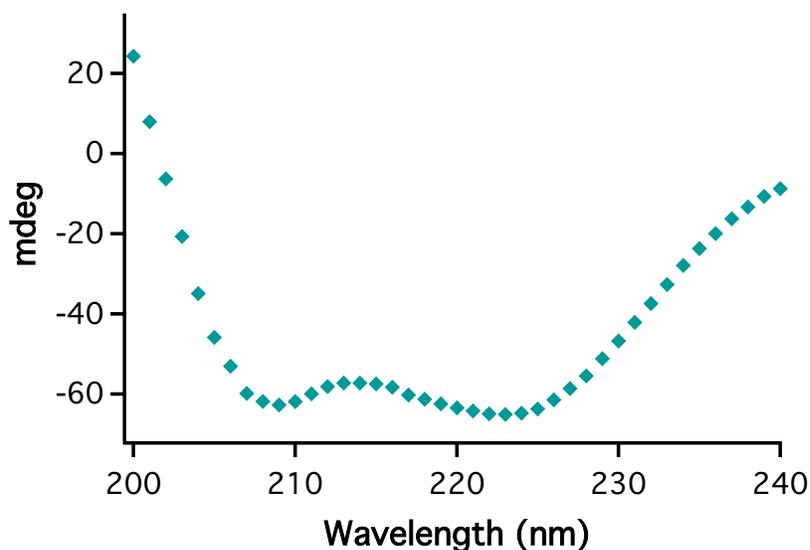


Figure 3.4: Representative circular dichroism spectrum of Dns-labeled cytochrome *cb*₅₆₂ (variant E92C).

CD spectroscopy provides information about the secondary structure and helical content of a protein. The circular dichroism spectrum of cytochrome *cb*₅₆₂ is shown in Figure 3.4. The two minima around 222 and 209 nm are indicative of α -helical structure. As the protein unfolds, the signal at 222 nm decreases until it reaches baseline (Figure 3.5).

UV-visible absorption spectroscopy probes the heme environment. As the protein unfolds, the heme Soret absorbance maximum red shifts. In the unfolding curve in Figure 3.6, the ratio of absorbance at the wavelengths that correspond to the maxima for the unfolded and folded protein are plotted.

The steep transition in the denaturation curves is indicative of a cooperative transition (Figure 3.7). The midpoint concentration of denaturant ($[Gdn]_{1/2}$) and a term representing the change in solvent accessibility upon unfolding (m) are

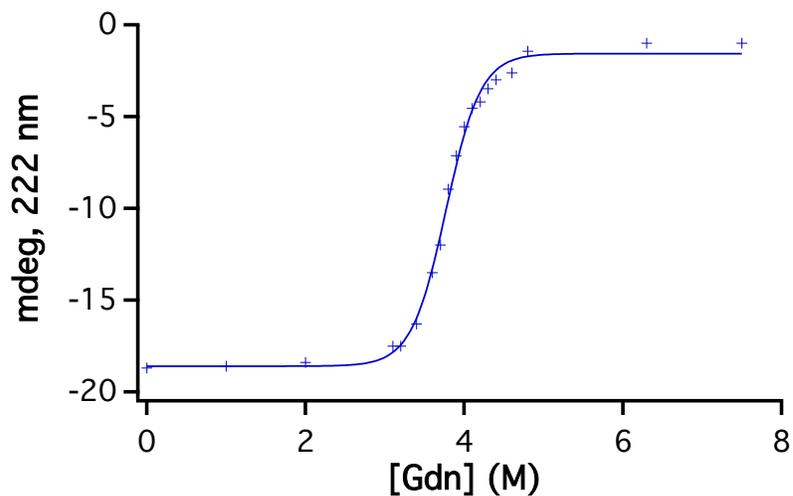


Figure 3.5: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} probed by CD spectroscopy at pH 5 (variant E92C).

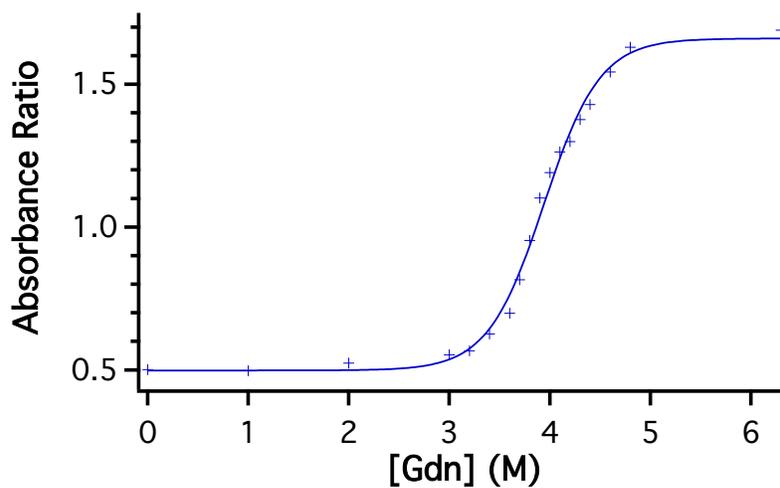


Figure 3.6: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} probed by UV-visible absorption spectroscopy at pH 5 (variant E92C). The ratio of absorbance at 402 nm to 415 nm is plotted.

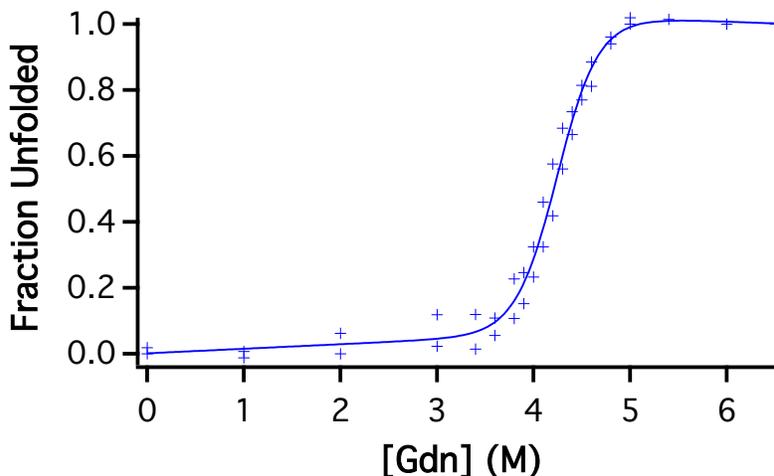


Figure 3.7: Chemical denaturation curve of Dns-labeled cytochrome cb_{562} (variant K19C, CD and UV-visible data).

Table 3.1: Denaturation parameters for Dns-labeled cytochrome cb_{562} variants.

Variant	$[\text{Gdn}]_{1/2}$ (M)	m (kJ/mol/M)	$-\Delta G$ (kJ/mol)
WT [36]	4.2	10	42
Dns19	4.2	11	46
Dns66	4.4	10	44
Dns92	3.8/3.9	12/9	40

extracted by fitting denaturant unfolding curves to a two-state model [59]. The free energy of folding at 0 M Gdn is the product of $[\text{Gdn}]_{1/2}$ and $-m$. $[\text{Gdn}]_{1/2}$ values for the labeled variants (Table 3.1) are comparable to the wild-type protein (4.2 M) [36], suggesting minimal perturbation of the structure and stability from incorporation of the cysteine mutation and Dns fluorophore. The Dns92 variant is slightly destabilized.

We have also measured the stability of cytochrome cb_{562} to thermal denaturation by CD spectroscopy (Figure 3.8). Cytochrome cb_{562} is exceptionally stable and contains significant secondary structure at 90°C. The partial unfolding that is

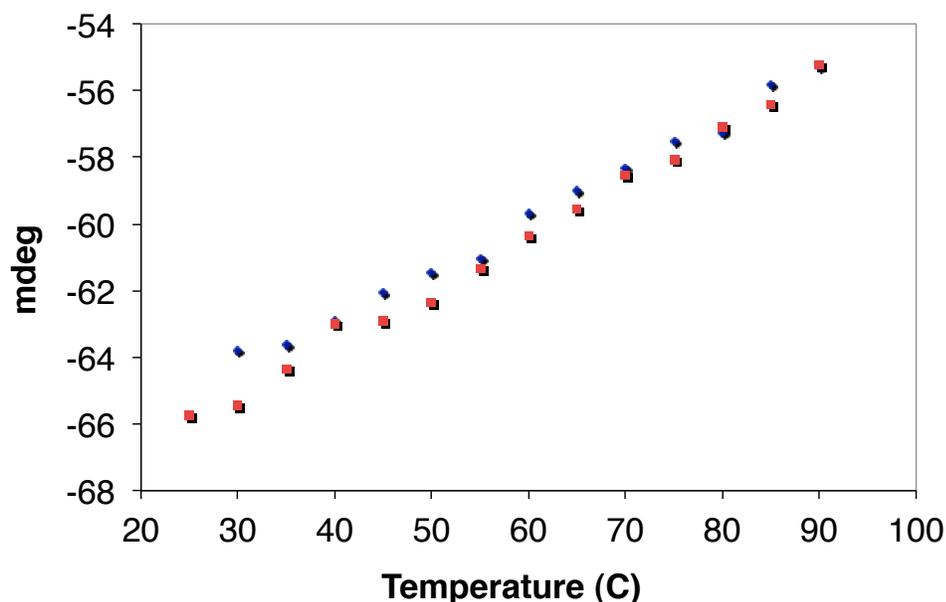


Figure 3.8: Temperature denaturation of Dns92-cytochrome cb_{562} . Unfolding is reversible; unfolding and refolding curves are shown in blue and red, respectively.

observed is not cooperative. Thus, heat-induced denaturation would be ineffective for studying the refolding of cytochrome cb_{562} .

3.4 Steady-State trFRET Measurements

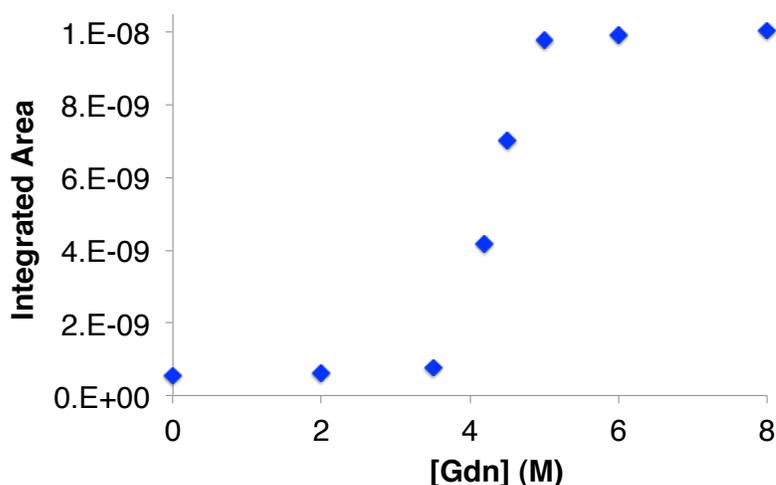
We employed trFRET to characterize the conformational populations of Dns-labeled cytochrome cb_{562} variants upon denaturation with Gdn. The Gdn-dependence of the normalized integrated fluorescence intensity (Figure 3.9), comparable to the CD and absorption denaturation curves, is indicative of cooperative unfolding.

Distance distributions were extracted from fluorescence decays (Figures 3.10–3.12). At 0 M Gdn, Dns–heme distances are consistent with the crystal structure distances in the folded wild-type protein (Table 3.2); the most probable distances

Table 3.2: Donor-acceptor distances in angstroms for Dns-labeled cytochrome cb_{562} variants.

Variant	Distance C_{γ} -Fe*	$r_{\text{mode}}^{\dagger}$	$r_{\text{mean}}^{\dagger}$	Random coil distance ‡	$r_{\text{mode}, 6 \text{ M Gdn}}^{\dagger}$
19	19.8	22	23	96	≥ 50
66	14.9	19	21	62	≥ 50
92	16.3	20	20	28	22

*Measured from the crystal structure (PDB 2BC5).

 † Extracted from the Dns-heme distance distributions at pH 4. ‡ Predicted by calculation [60].**Figure 3.9:** Denaturant-dependence of integrated normalized Dns fluorescence decays for Dns19-cytochrome cb_{562} .

extracted from the fluorescence decays (r_{mode}) are slightly longer (2–4 Å) than the C_{γ} -Fe distances in the crystal structure, likely owing to the length of linker between the Dns fluorophore and the C_{γ} atom. Denaturation with Gdn slowed the fluorescence decay in all Dns variants, consistent with an increase in the average Dns-heme distance. Dns fluorescence decays could not be fit to single exponential functions as distance components exist with various probabilities, consistent with the presence of multiple conformations.

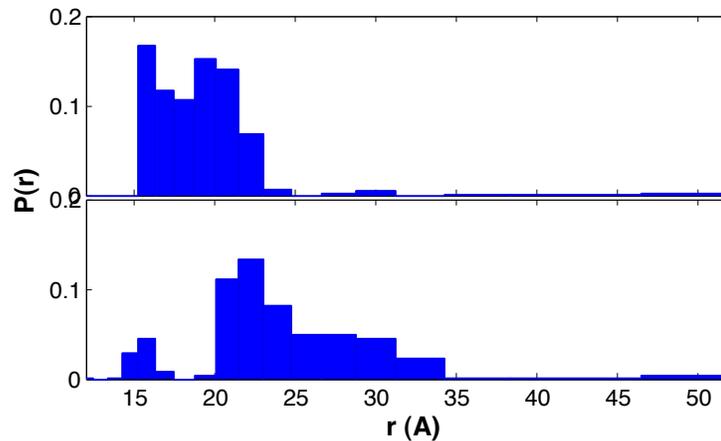


Figure 3.10: Dns92–heme distance distributions for native (upper) and denatured (8 M Gdn, lower) cytochrome *cb*₅₆₂ at pH 5 by trFRET.

A random coil is characterized by a statistical distribution of conformations that lack well-defined structure. The most probable Dns–heme distances for the extended states roughly correspond to the calculated random coil distances (Table 3.2). The extended conformations of Dns92 (Figure 3.10) are distributed around this random coil distance. The predicted random coil distances for mutants Dns19 and Dns66 are beyond the resolution of the measurement (> 52 Å); however, notable populations at ≥ 50 Å are consistent with the predicted distances (Figures 3.11–3.12). Significantly, the extended structures are more populated for mutant K19C, which has a longer calculated random coil distance. In contrast, the unfolded conformations of cytochrome *c'* do not resemble a random coil; trFRET measurements of Trp32–heme distances are similar to Trp72–heme distances, despite the significantly greater distance between Trp32 and the heme in the polypeptide chain [28].

Dns92, located on helix 4, probes formation of the secondary structure of

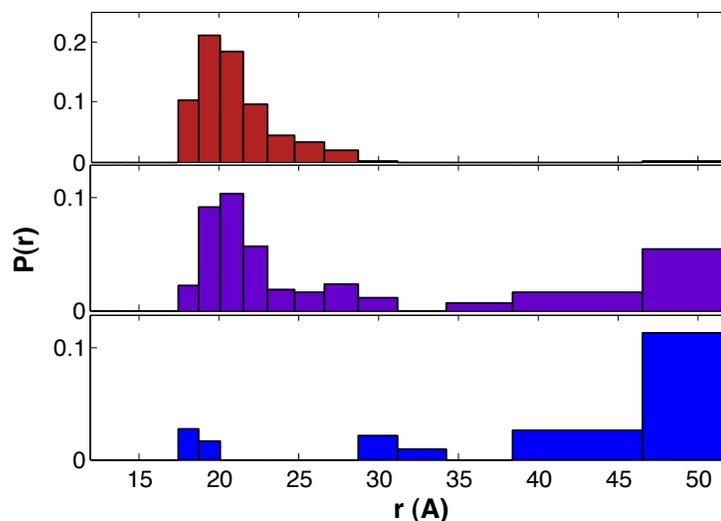


Figure 3.11: Dns66–heme distance distributions of cytochrome cb_{562} in 0 M Gdn (red), 4.2 M Gdn (violet), and 6 M Gdn (blue) at pH 4 by trFRET.

this helix. Residue 92 is located near the heme on the polypeptide chain, so only modest broadening of the D-A distribution is observed upon denaturation (Figure 3.10). Despite the overlap of states in the unfolded ensemble, there is an appreciable occupation of extended states (>25 Å). The compact unfolded conformations could consist of bent non-helical polypeptide conformations that bring the Dns fluorophore and heme close together, as well as native-like contacts; heme misligation may increase the occupation of this compact population (see section 3.5).

Variant Dns66 probes tertiary contacts between helix 3 and the heme, attached to helix 4 (Figure 3.11). The unfolded state ensemble consists predominantly of extended structures (94% at ≥ 30 Å, 86% at ≥ 40 Å). At the denaturation midpoint ([Gdn] = 4.2 M), the protein ensemble equally resembles the folded and unfolded protein with 47% extended conformations (> 30 Å).

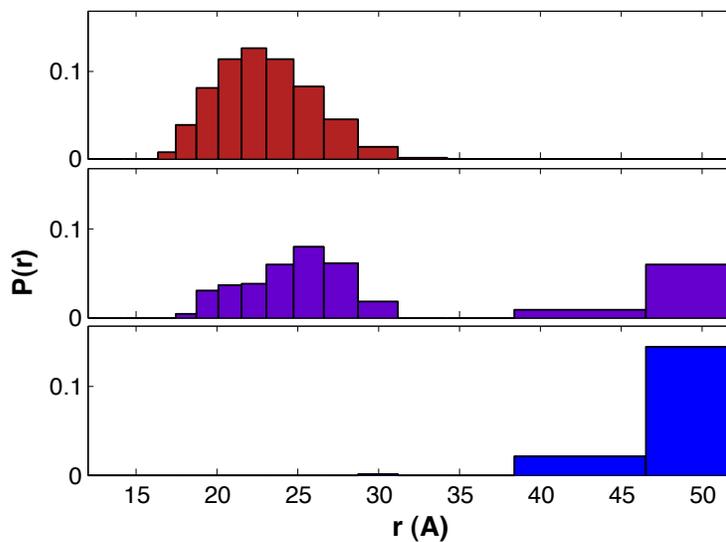


Figure 3.12: Dns19–heme distance distributions of cytochrome cb_{562} in 0 M Gdn (red), 4.2 M Gdn (violet), and 6 M Gdn (blue) at pH 4 by trFRET.

Variant Dns19 probes tertiary contacts between helix 1 and the heme (Figure 3.12). The unfolded state ensemble is populated exclusively with extended structures. At the denaturation midpoint, two populations with distributions that resemble the native and unfolded states are observed.

3.5 Heme Environment¹

The heme coordination of cytochrome *cb*₅₆₂ is shown in Figure 3.13. When the protein is folded, the heme is axially ligated to the sidechains of Met7 and His102. Ferricytochrome (Fe³⁺) *cb*₅₆₂ is unfolded by Gdn with a denaturation midpoint of 4.2 M Gdn. Changes in the heme environment with different solvent conditions were probed by UV-visible spectroscopy (Figure 3.14).

Ferricytochrome *cb*₅₆₂ has a Soret maximum of 415 nm. When unfolded in 6 M Gdn at pH 4, the Soret maximum blue shifts to 400 nm as the heme transitions from low spin to high spin, suggesting that the Met7 sulfur ligation is replaced by a water molecule [37]. As the pH of the unfolded protein is increased, the Soret peak red shifts ($\lambda_{\text{max}} = 412$ nm). The shift and pH range are consistent with heme ligation by a histidine imidazole. This observation suggests that in denatured cyt *cb*₅₆₂, the nonnative histidine ligand (His63) coordinates to the heme at pH ≥ 5 .

Conformational populations of Dns66-cytochrome *cb*₅₆₂ were determined by trFRET measurements at variable Gdn concentrations and pH (Figures 3.15). Extended conformations with >50 -Å distances comprise a significant population of denatured proteins (6 M Gdn). At pH 4, there are few compact states. With increasing pH, however, the proportion of the compact population increases, consistent with a decrease in the configurational freedom and the number of states available to the unfolded protein upon formation of a 36-residue loop by His63 ligation to the heme.

¹Adapted from Bouley Ford, N.; Shin, D.W.; Gray, H.; Winkler, J. "Tertiary Contact Dynamics in Unfolded Cytochrome *cb*₅₆₂" *Submitted*.

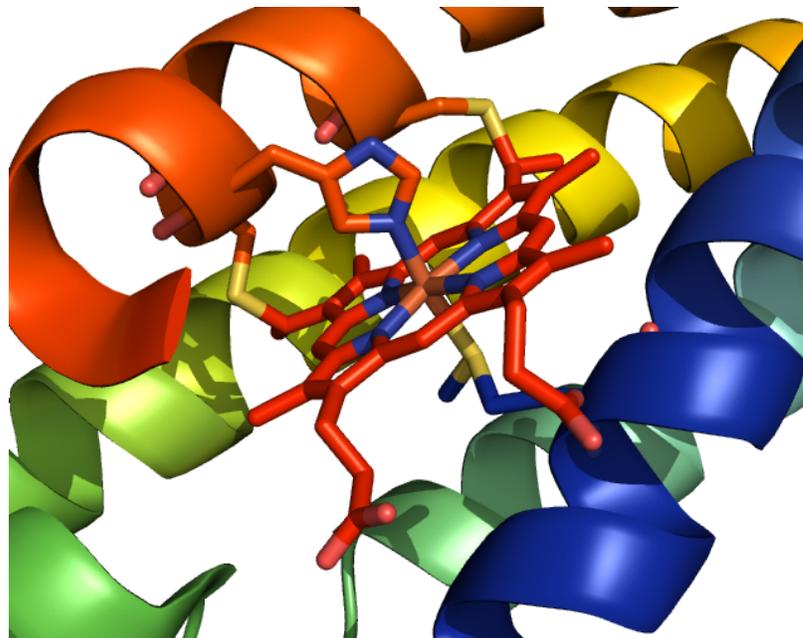


Figure 3.13: Heme environment of cytochrome cb_{562} (PDB 2BC5). The heme (red) is attached covalently to the fourth helix (orange) and ligated to Met 7 on the first helix (blue) and His102 on the fourth helix.

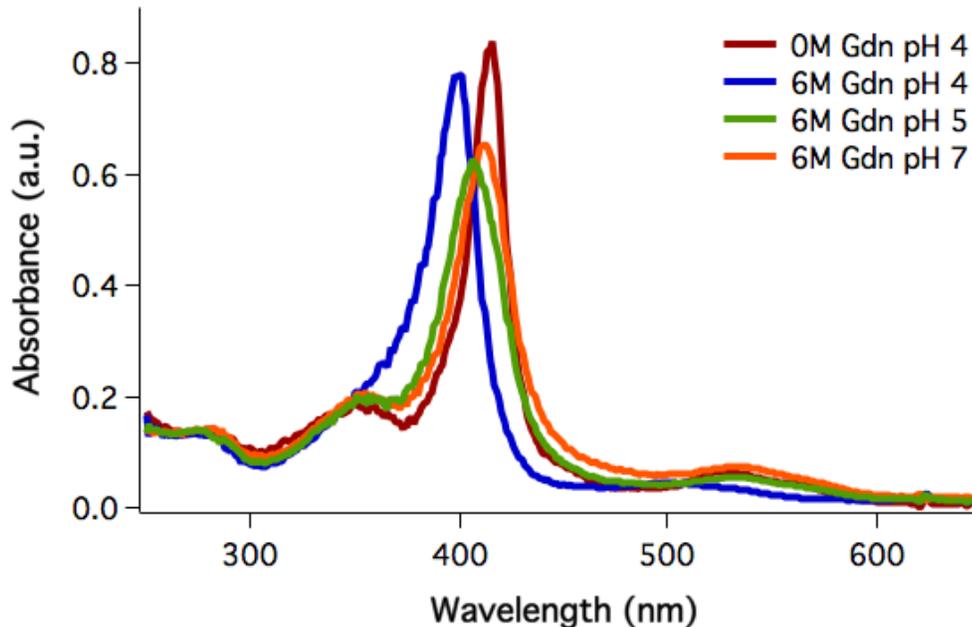


Figure 3.14: UV-visible spectra of cytochrome cb_{562} , variant K32C. The heme is ligated by Met7 and His102 when folded ($\lambda_{max} = 415$ nm). When the ferricytochrome is denatured by 6 M Gdn, Met7 is replaced by water or His63 ($\lambda_{max} = 400$ nm at pH 4, 406 nm at pH 5, 412 nm at pH 7).

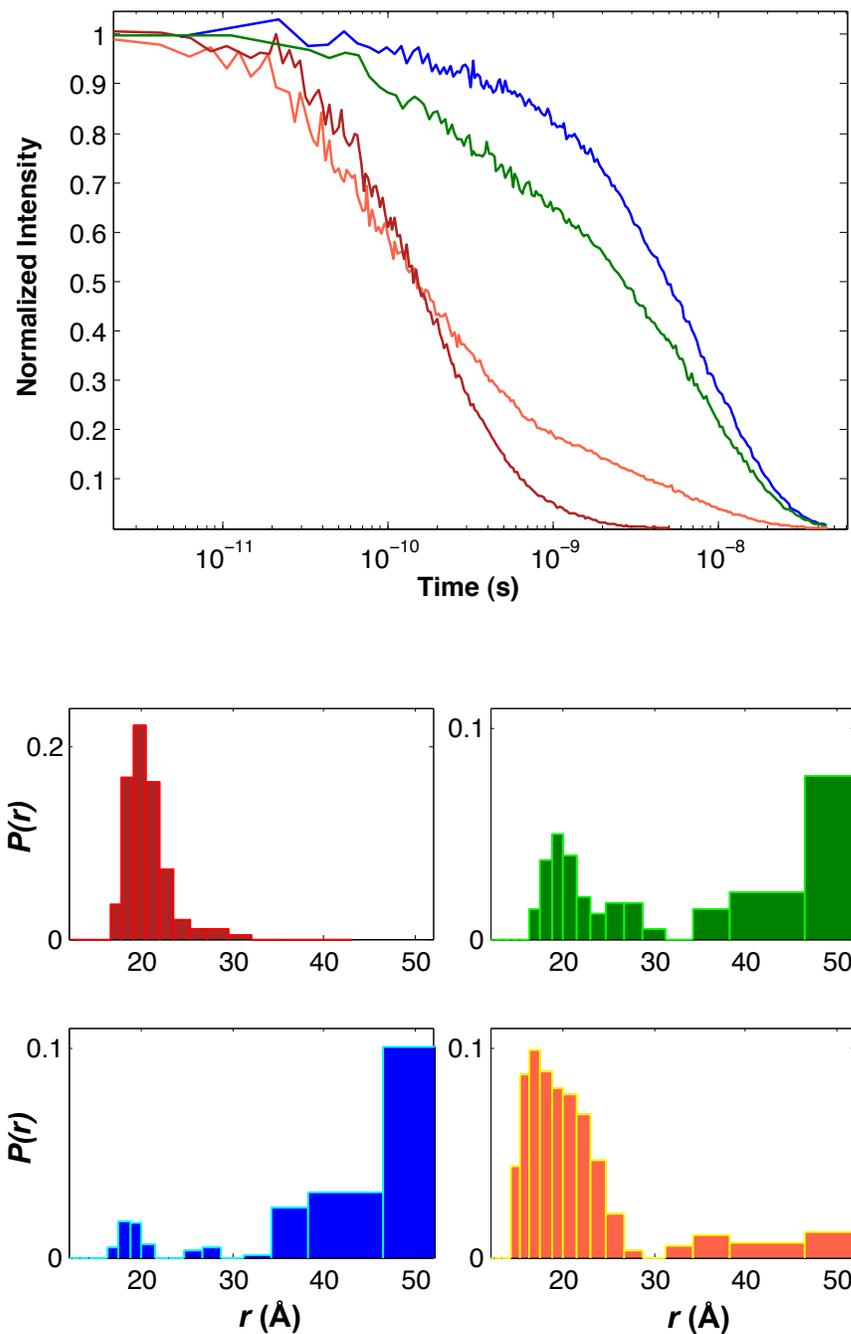


Figure 3.15: Fluorescence decays and corresponding distance distributions of Dns66-cytochrome cb_{562} : folded at pH 4 (red) and unfolded in 6 M Gdn at pH 4 (blue), pH 5 (green), and pH 7 (orange).

3.6 Conclusions

Cytochrome cb_{562} is highly stable to chemical (Gdn) and thermal denaturation, exhibiting incomplete unfolding at 90°C. The protein exhibits cooperative unfolding in equilibrium conditions ($D_{50\%} = 4.2$ M Gdn) by circular dichroism, UV-visible absorbance, and fluorescence spectroscopy.

The properties of the unfolded state ensemble strongly influence the folding mechanism. At pH 4, the unfolded state, probed by trFRET with a Dns fluorophore attached to K19C, D66C, or E92C, resembles a random coil. However, pH studies of unfolded cytochrome cb_{562} by UV-visible absorption spectroscopy indicate that the heme likely coordinates to the nonnative ligand His63 at pH > 4. Misligated conformations, probed by trFRET measurements of Dns66, are significantly more compact.

3.7 Acknowledgments

Parts of this work were carried out in collaboration with Caltech undergraduates Yuehan Huang and Dong Woo Shin.

Chapter 4

Tertiary Contact Dynamics in Unfolded Cytochrome *cb*₅₆₂ by Electron Transfer¹

4.1 Introduction

When studying the dynamics of unfolded proteins, frustration manifests as internal friction, slowing intrachain diffusion in regions of the polypeptide. As formation of tertiary contacts is a necessary step in folding, the speed limit is set by the conformational dynamics in the unfolded state [14, 15, 16, 17]. Statistical mechanical models suggest that the rates of tertiary contact formation in polymers are determined by the intrachain diffusion coefficient and chain length [61, 62]. Approximate analytical solutions to the Smoluchowski equation for idealized proteins predict that contact rate constants will show a power-law distance dependence on n ($n^{-1.5}$), where n is the number of peptide bonds between contacts [61]. A model that takes into account the loop size formation probability predicts a

¹Adapted from Bouley Ford, N.; Shin, D.W.; Gray, H.; Winkler, J. "Tertiary Contact Dynamics in Unfolded Cytochrome *cb*₅₆₂" *Submitted*.

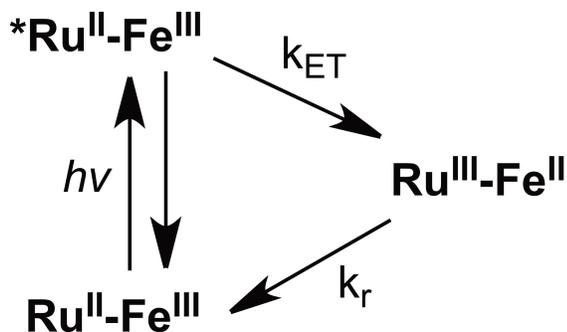


Figure 4.1: Luminescence of photoexcited *Ru^{II} is quenched by electron transfer upon contact with the ferriheme.

maximum rate at $n = 10$ with an $n^{-3.2}$ dependence for larger loops and slower rates for small loops, owing to increased stiffness [62]. Analysis of contact quenching experiments probing intrachain diffusion in unstructured polypeptides revealed a power-law dependence for large loops with an n -independent asymptote for small loops [63, 64].

In this chapter, we describe the contact dynamics of denatured cytochrome cb_{562} , a minimally frustrated protein. We examine the formation of transient contacts between the heme and $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ complexes (Figure 1.7) that have been covalently attached at various positions in the protein chain. Rates of intrachain diffusion in chemically denatured cytochrome cb_{562} molecules have been obtained from analysis of time-resolved luminescence measurements, thereby allowing us to assess the degree of frustration in the energy landscape of this four-helix bundle. We also have found that forming one tertiary contact can accelerate formation of subsequent contacts.

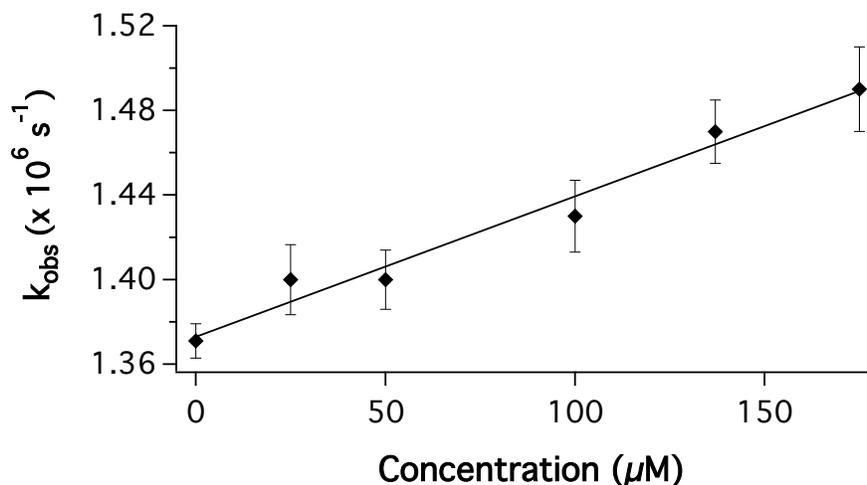


Figure 4.2: Stern-Volmer plot for the quenching of photoexcited $\text{Ru}(\text{bpy})_3^{2+}$ by cytochrome cb_{562} (Fe^{3+}) in 6 M Gdn at pH 4. The data fit to a k_Q of $6.6(\pm 0.5) \times 10^8 \text{ M}^{-1}\text{s}^{-1}$.

4.2 Contact Quenching

Following photoexcitation, $^*[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ luminesces with an unquenched lifetime of $\sim 1 \mu\text{s}$ ($k_{\text{label}} = 1.10(\pm 0.04) \times 10^6 \text{ s}^{-1}$ at pH 4–7 with 6–8M Gdn). Contact with a ferriheme quenches the $^*\text{Ru}^{\text{II}}$ state via electron transfer (Figure 4.1). We have performed quenching studies to determine the rate constant for the bimolecular electron transfer reaction using $\text{Ru}(\text{bpy})_3^{2+}$ as a model for the Ru-label $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$. Quenching was examined at several protein concentrations in 6 M Gdn (Figure 4.2). We measure a quenching rate constant (k_Q) of $6.6(\pm 0.5) \times 10^8 \text{ M}^{-1}\text{s}^{-1}$. Correcting for the viscosity of 6 M Gdn [65] gives $k_Q = 1.2 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, consistent with a reaction near the diffusion-controlled limit.

```

ADLEDNMETLNDNLKVIEKADNAAQVKDAL 30
TKMRAAALDAQKATPPKLEDKSPDSPEMWD 60
FRHGFDILVGQIDDALKLANEGKVKEAQA 90
AEQLKTTCNACHQKYR 106

```

Figure 4.3: Sequence of cytochrome cb_{562} . Cysteines by which the porphyrin is attached covalently to the protein backbone are shown in red. Heme ligands are shown in orange with the non-native ligand His63 italicized. Residues that are mutated to cysteine in single-mutation variants are shown in blue.

4.3 Ruthenium-Labeled Variants

Residues labeled with $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ in single-cysteine variants are highlighted in the sequence of cytochrome cb_{562} in Figure 4.3. The circular dichroism (Figure 4.4) and heme Soret absorption spectra of labeled proteins are virtually identical with those of the unlabeled wild-type, confirming that the mutations and Ru-labeling do not disrupt the protein secondary structure or heme environment.

4.4 Transient Contact Formation Rates

The formation of transient loops in the polypeptide chain of cytochrome cb_{562} was probed by monitoring electron transfer between the heme and $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ complexes attached at various positions on the polypeptide chain. The porphyrin vinyl groups of the heme are covalently attached to the protein backbone of this 106-residue protein via two mutant cysteine residues, R98C and Y101C, in a CXXCH cyt *c*-type motif [37]. Data were collected at pH 4 to inhibit His63 ligation of the heme and allow the denatured proteins a full range of motion.

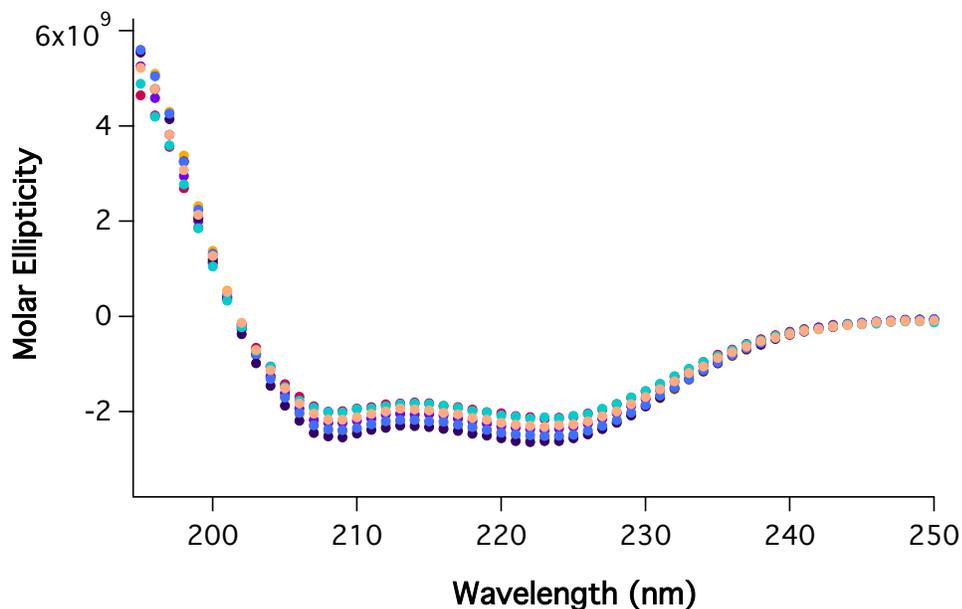


Figure 4.4: Circular dichroism spectra of cytochrome cb_{562} Ru-variants at pH 7: Ru19 (yellow), Ru32 (red), Ru51 (purple), Ru66 (deep purple), Ru77 (blue), Ru83 (teal), and Ru92 (coral).

As bimolecular electron transfer is diffusion-limited, the rate constant for the labeled proteins ($k_{\text{ET}} = k_{\text{obs}} - k_{\text{label}}$) corresponds to the specific rate of transient contact formation.

Luminescence decays for Ru-variants in 6 M Gdn and the corresponding $1/k_{\text{ET}}$ values are shown in Figures 4.5 and 4.6. The degree of quenching correlates with the efficiency of the electron transfer reaction. The luminescence decays exhibit single-exponential kinetics, consistent with interconversion among protein conformations that is rapid compared to luminescence decay. Luminescence decay rate constants (k_{obs}) for the variants range from $1.53(\pm 0.03) \times 10^6 \text{ s}^{-1}$ for Ru51 to $9.4(\pm 0.6) \times 10^6 \text{ s}^{-1}$ for Ru92 (Table 4.1). Luminescence decays for Ru19 and Ru32 with the greatest Ru-heme distance are not quenched. For these variants, electron

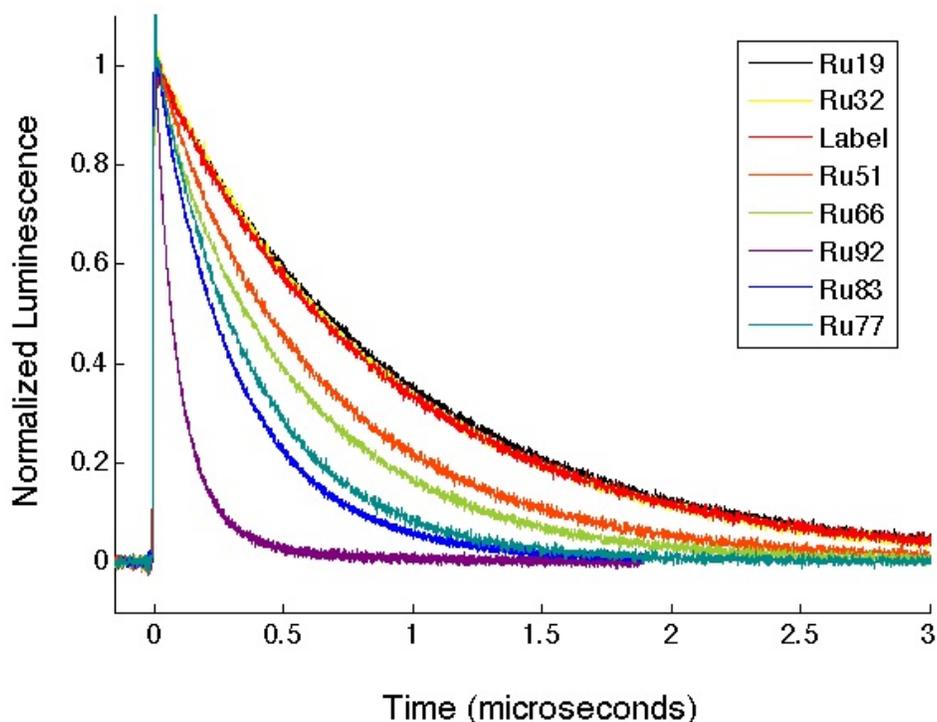


Figure 4.5: Luminescence decays of $^*\text{Ru}^{\text{II}}$ in the Ru-labeled cytochrome cb_{562} variants and the unquenched free label at pH 4 in 6 M Gdn.

transfer is not competitive with excited state deactivation, indicative of a contact time greater than $20 \mu\text{s}$ (the standard deviation for k_{ET}^{-1}).

We observe a Gdn-dependence of contact formation rate constants (k_{ET}), as expected for a diffusion-limited process. For example, k_{ET} for Ru77 is $1.37(\pm 0.07) \times 10^6 \text{ s}^{-1}$ at 6 M Gdn and $9.1(\pm 0.5) \times 10^5 \text{ s}^{-1}$ at 8 M Gdn. A viscosity dependence also has been observed for intrachain diffusion in denatured cytochrome c and in Gly-Ser-repeat polypeptides [15, 64]. Prior work on end-to-end contact kinetics in Gly-Ser-repeat polypeptides found a linear dependence of $\ln(k)$ on $[\text{Gdn}]$ [23]. If we assume similar behavior in cytochrome cb_{562} , then our two data points are consistent with a slope of $-\Delta \ln(k) / \Delta [\text{Gdn}] = 0.2 \text{ M}^{-1}$, comparable to the values

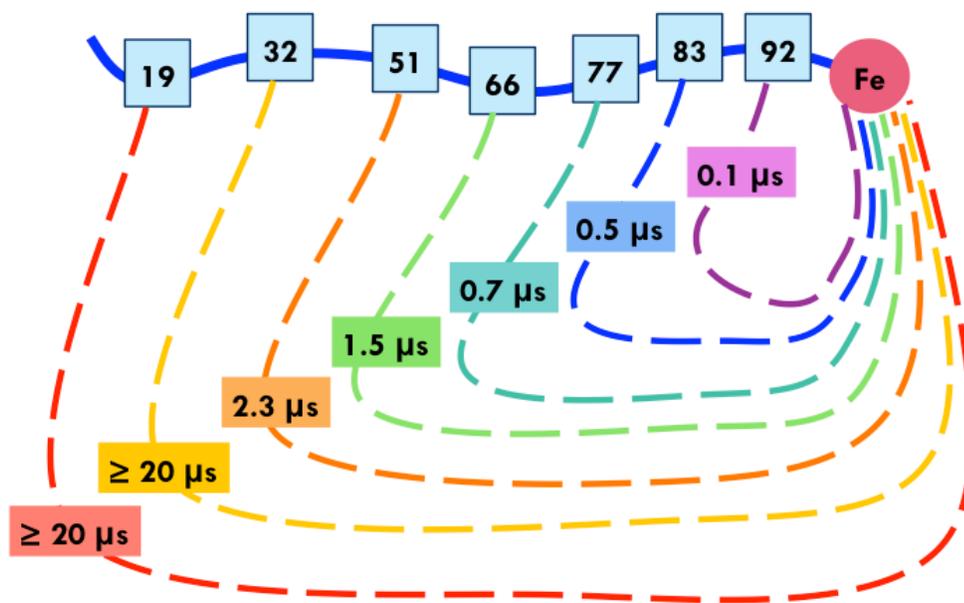


Figure 4.6: Schematic of the contact formation time constants for denatured Ru-cytochrome cb_{562} variants, which show a distance dependence.

Table 4.1: Fitted rate constants for luminescence data collected at pH 4 in 6 M Gdn and the corresponding contact formation rate constants.

Variant	Distance (n)	k_{obs} (s^{-1})	k_{ET} (s^{-1})
Free label	N/A	$1.09(\pm 0.03) \times 10^6$	N/A
Ru19	79	$1.06(\pm 0.02) \times 10^6$	Not observed
Ru32	66	$1.14(\pm 0.04) \times 10^6$	Not observed
Ru51	47	$1.53(\pm 0.03) \times 10^6$	$4.4(\pm 0.4) \times 10^5$
Ru66	32	$1.75(\pm 0.04) \times 10^6$	$6.6(\pm 0.5) \times 10^5$
Ru77	21	$2.46(\pm 0.07) \times 10^6$	$1.37(\pm 0.07) \times 10^6$
Ru83	15	$3.1(\pm 0.2) \times 10^6$	$2.0(\pm 0.2) \times 10^6$
Ru92	6	$9.4(\pm 0.6) \times 10^6$	$8.3(\pm 0.6) \times 10^6$

extracted from Gly-Ser-repeat polypeptides (0.19–0.23 M⁻¹) [23]. Extrapolating to 0 M Gdn solution suggests that the Ru77-heme contact rate constant would be $5 \times 10^6 \text{ s}^{-1}$ in buffer solution without Gdn. Applying the same extrapolation to our largest measured rate constant (Ru92, 10^7 s^{-1} in 6 M Gdn) leads to an estimated contact time of ~ 30 -ns for Ru92-heme in water.

The contact formation rate constants display a power-law dependence on n , the number of peptide bonds between the Ru-label and the heme (Figure 4.7). A fit to the function $k_{\text{ET}} = A \cdot n^y$ gives the following parameters: $A = 1.20(\pm 0.09) \times 10^8 \text{ s}^{-1}$ and $y = -1.49 \pm 0.04$. The random polymer model by Szabo, Schulten, and Schulten predicts this trend ($k \propto n^{-1.5}$) for purely entropy-controlled intrachain diffusion in ideal, freely-jointed Gaussian chains [61]. In contrast, end-to-end contact measurements of unstructured Gly-Ser-repeat polypeptides of variable length in water and 8 M Gdn exhibit power-law slopes of -1.7 and -1.8, respectively, which are consistent with the Gaussian-chain model when excluded-volume effects are included [64]. Importantly, our data indicate that cytochrome cb_{562} behaves as a random coil when denatured and that excluded volume corrections to our intrachain contact rates are not needed.

The adherence of the data ($n \geq 6$) to the random polymer model provides strong support for the proposal that cytochrome cb_{562} (helix 2/3 loop to the heme) is minimally frustrated. Deviations from the power-law function would be expected for frustrated regions that exhibit increased internal friction. And indeed, deviations were found for α -synuclein, the intrinsically disordered polypeptide

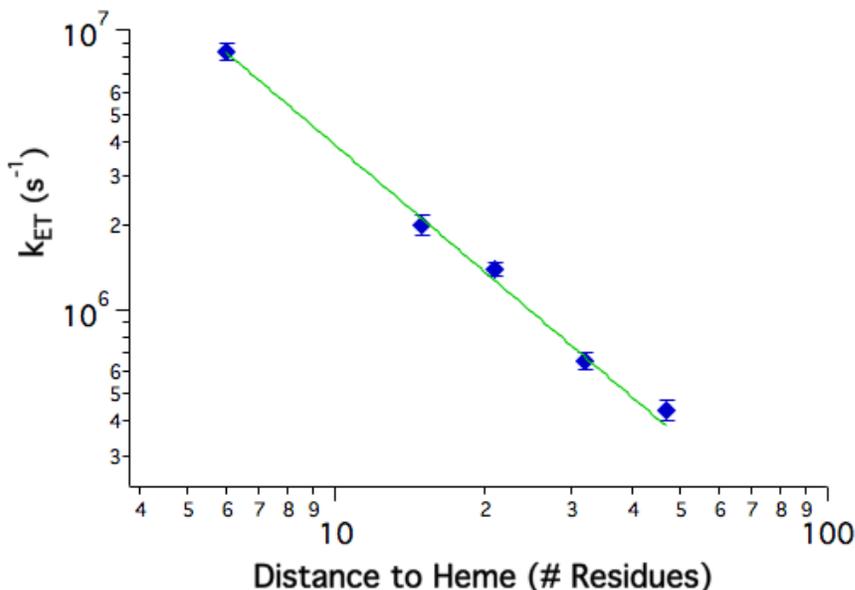


Figure 4.7: Log/log plot of contact formation rates versus number of peptide bonds between transient intramolecular contacts (Table 4.1). Ru-variants are denatured in 6 M Gdn at pH 4. The curve fits to a power function with a slope of -1.49 ± 0.04 .

implicated in Parkinson's disease [19]. It will be of special interest to compare the unfolded dynamics of cytochrome cb_{562} to those of cytochrome c' , as the latter protein displays probe-dependent folding kinetics [28].

Our contact rate constants for the denatured cytochrome cb_{562} polypeptide are substantially smaller than those reported for Gly-Ser-repeat polypeptides in Gdn, a trend that also was found for the contact dynamics of α -synuclein in the absence of denaturants [64, 18]. For comparison, contact formation for Ru77 ($n=21$) in 8 M Gdn is five times slower than for the repeat polypeptides in 8 M Gdn [64]. The different probes of contact events (i.e., triplet-triplet energy transfer versus electron transfer) may contribute to the discrepancy. If the *Ru to heme electron transfer reaction is not strictly diffusion controlled, the observed contact times could be

systematically greater than those found using energy-transfer quenching.

Alternatively, the intrachain diffusion coefficient in cytochrome cb_{562} may be smaller than that for repeat polypeptides. Although Gdn viscosity effects, as we have noted, are similar for cytochrome cb_{562} and repeat polypeptides, the increased conformational freedom of glycine as compared to other amino acids could contribute to a somewhat increased diffusion coefficient in the synthetic peptides. The presence of glycine can only partially explain the difference however; polyserines diffuse two times slower than Gly-Ser polypeptides [64].

Dynamical drag introduced by segments of the polypeptide chain exterior to the loop also likely contributes to the slower contact rates found in the 106-residue protein [66]. In synthetic polypeptides, interior-to-end contact rates decrease by as much as 2.5-fold compared to end-to-end contacts [67]. In proteins, interior contact dynamics would be more affected by drag from the external protein chain than contacts toward the ends of the chain. This drag could account for the differences in the power-law dependence for interior loop formation in cytochrome cb_{562} ($y = -1.5$) compared to that found for end-to-end contact formation ($y = -1.8$) in synthetic peptides.

4.5 Dynamics within a Constrained Loop

At $\text{pH} > 4$, Ru66 quenching kinetics probe contact formation within a 36-residue loop formed by the His63 misligation to the heme (Figure 4.8). Under these conditions, Ru66 cytochrome cb_{562} exhibits biexponential luminescence decay

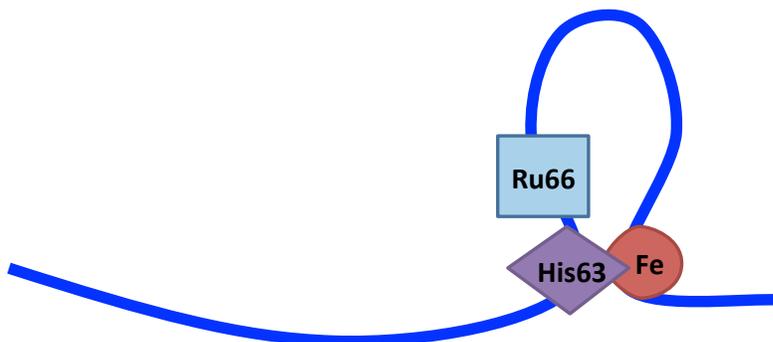


Figure 4.8: Schematic of long-lived loop formed by ferriheme misligation with His63 at pH > 4.

Table 4.2: pH dependence of contact formation rate constants for Ru66 in 6 M Gdn.

pH	First k_{ET} (s^{-1})	Second k_{ET} (s^{-1})
4	6.6×10^5 (100 %)	N/A
5	6×10^5 (60 %)	1.2×10^7 (40 %)
7	3×10^5 (17 %)	1.0×10^7 (83 %)

kinetics (Figure 4.9), consistent with two populations that are not exchanging on the microsecond timescale. The UV-visible spectrum of cytochrome cb_{562} suggests that there is an increase in the degree of His63 heme misligation from pH 5 to 7 as the Soret band further red shifts, concurrently with an increase in the compact conformational population observed in trFET measurements (see section 3.5). Over this pH range, the contribution of the second rate constant to the biexponential fit increases from 40% to 83% (Table 4.2 and Figure 4.10). It is likely, therefore, that the additional compact population is attributable to heme-misligated structures.

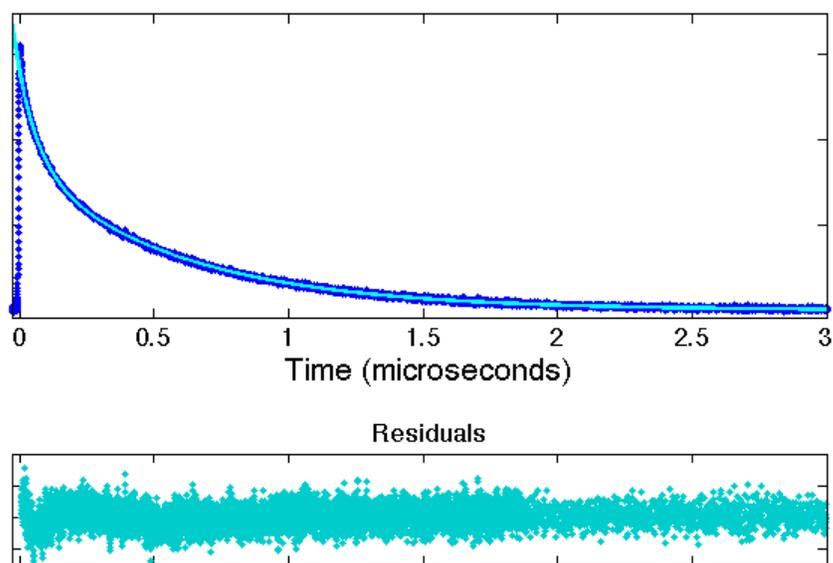


Figure 4.9: Biexponential fit of luminescence decay of Ru66 at pH 5.

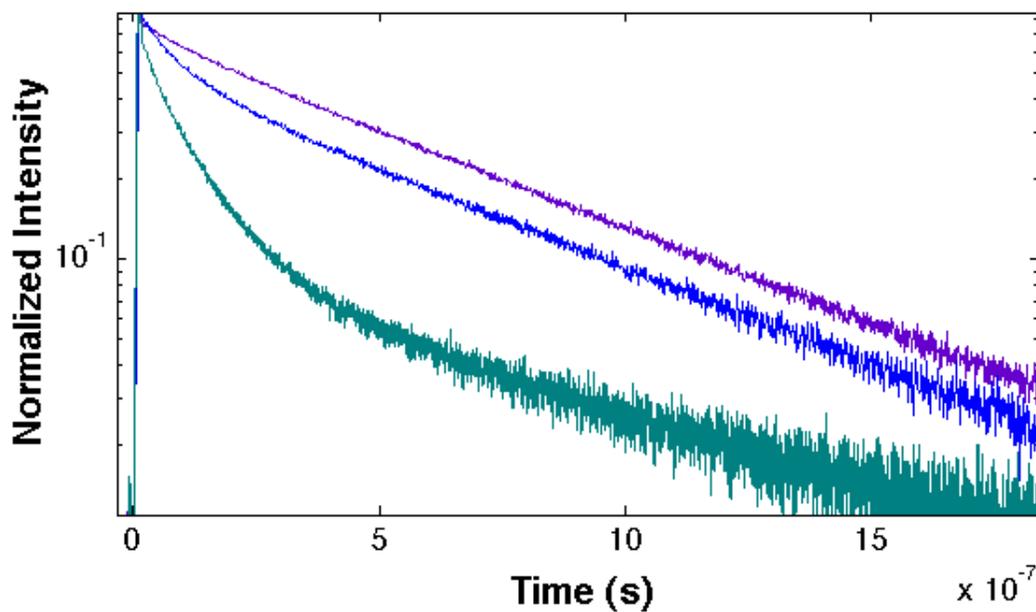


Figure 4.10: Luminescence decays of Ru66 at pH 4 (purple), pH 5 (blue), and pH 7 (teal). At pH 5 and 7, biexponential quenching by two distinct populations, the freely diffusing and misligated forms, is observed.

The 36-residue static loop formed by misligation serves as a model for early formation of tertiary contacts in protein folding. Ru66-heme contacts form more rapidly for misligated populations ($k_{\text{ET}} \approx 10^7 \text{ s}^{-1}$) than in the random coil ($k_{\text{ET}} \approx 10^{5.8} \text{ s}^{-1}$). The heme-His63 contact acts as a nucleation site that facilitates interactions with nearby Ru66, with formation of this new contact approaching the folding speed limit. In contrast to Ru66, two variants between His63 and the heme, Ru77 and Ru83, do not show enhanced electron transfer in the misligated population; single exponential luminescence fits at pH 7 correspond to pH 4 fits ($k_{\text{ET}} = 1.4 \times 10^6 \text{ s}^{-1}$ for Ru77 at both pH 4 and 7; $k_{\text{ET}} = 2.0 \times 10^6 \text{ s}^{-1}$ and $8.4 \times 10^5 \text{ s}^{-1}$ for Ru83 at pH 4 and 7, respectively). Although the His63-heme contact speeds formation of a new contact with nearby Ru66, it does not assist residues further away in the loop. This behavior is consistent with a cooperative folding process. If a native tertiary contact is formed early in the folding pathway, it can limit the conformational search for other tertiary contacts by effectively decreasing n for nearby residues.

We observe a divergence from the power-law behavior when we include the misligated populations of Ru66 (Figure 4.11). Contact formation may be slowed due to decreased access to the heme when ligated by a histidine side chain instead of a water molecule; this rate constant could be a lower limit for loops of comparable size. Kiefhaber and coworkers found that the increased rigidity of smaller loops produces a limiting contact rate constant (k_0): $k = (1/k_0 + 1/(A^*n^y))^{-1}$ [64]. The influence of the rate-limiting factor was less pronounced in Gdn than

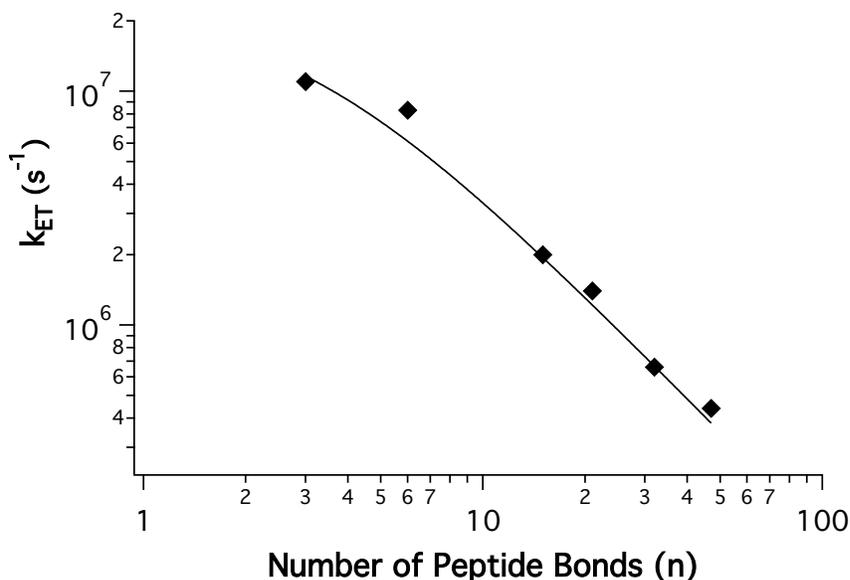


Figure 4.11: Fit of contact rate constants to $k = (1/k_0 + 1/(A*n^y))^{-1}$ with the following constraints: $A = 1.2 \times 10^8 \text{ s}^{-1}$ and $y = -1.5$. Rate constants reach an upper limit of $k_0 = 2.3 \times 10^7 \text{ s}^{-1}$. Additional short-distance data points are needed to determine if the function accurately describes the data set.

in water (A and k_0 converging), possibly owing to decreased chain stiffness in Gdn [64, 68]. Including Ru66 (pH 7) data in a fit to the extended equation with constrained A and y parameters gives $k_0 = 2.3 \times 10^7 \text{ s}^{-1}$, or an upper time constant of 43 ns for close contacts (Figure 4.11).

4.6 Conclusions

The upper limit for the kinetics of the folding reaction will be governed by the rate at which intrachain contacts are formed. We measure a limiting time constant of $0.1 \mu\text{s}$ in 6 M Gdn, extrapolated to an estimated 30 ns in water. We observe a power-law dependence of contact formation rate constants on n , the length of the polypeptide chain between the contacts. Adherence of our data to a slope of -1.5, consistent with theoretical models for ideal polypeptides, demonstrates that cytochrome cb_{562} is minimally frustrated (loops 3–4). Contacts that stabilize the native structure are thus strongly favored during folding. Early formation of these native contacts can aid the protein in forming additional contacts, as we have observed, by effectively decreasing n . These events are expected to increase in occurrence as the folding reaction proceeds, in accord with a funneled energy landscape [8, 11], to the native state.

4.7 Acknowledgments

This work was carried out in collaboration with Dong Woo Shin, a Caltech student. We would like to thank the following former and current group members for the synthesis of the Ru label: Lionel Cheruzel, Jeffrey Warren, Maraia Ener, and Katja Luxem.

Chapter 5

Resolving the Fast Folding of Cytochrome cb_{562} with Microfluidic Mixing

5.1 Introduction

Cytochrome cb_{562} belongs to a unique family of cytochromes with similar topologies but divergent folding kinetics. We would like to fully characterize the folding mechanisms of cytochromes cb_{562} and c' and isolate the factors causing the differences. In this chapter, we resolve the folding kinetics of ferricytochrome cb_{562} by trFRET and continuous flow mixing.

Previous folding studies of ferrocycytochrome cb_{562} were initiated by electron injection and monitored by transient absorption spectroscopy [36]. Met7 heme coordination occurs within 10 μs , and the absorption spectrum resembles the folded state after 100 μs . This experimental result suggests that persistent heme ligation aids in the folding process. Folding of the ferricytochrome is not expected to be as rapid as the ferrocycytochrome since Met7 is more strongly coordinated to

the heme in the reduced protein.

Folding of ferricytochrome cb_{562} was investigated by stopped-flow mixing [37]. When monitored by Trp59 fluorescence, more than half of the signal is quenched during the 5 ms mixing dead time. The remaining amplitude decays with a denaturant-dependent rate constant that extrapolates to $4.2 \times 10^2 \text{ s}^{-1}$ at 0 M Gdn. Absorption-monitored kinetics are biphasic and include the fast phase observed with fluorescence and an additional denaturant-independent slow phase ($k_{\text{obs}} \approx 5 \text{ s}^{-1}$). This slow phase likely involves optimization of the native heme environment, including the Met7 ligation.

Microfluidic continuous flow mixing (150- μs dead time) allows for the resolution of the early folding of cytochrome cb_{562} [29]. Trp59(helix3)-heme distance distributions were probed throughout folding by trFRET. Helices 3-4 fold with a rate constant of 240-260 s^{-1} . Interestingly, cytochrome c' also exhibited apparent two-state folding when Trp32-heme and Trp72-heme distances were probed, but differing rates [28]. This probe dependence of the rate constants are indicative of a sequential folding mechanism where the region around the helix 1/2 loop forms native contacts more rapidly ($k = 170 \text{ s}^{-1}$) than the helix 2/3 loop ($k = 15 \text{ s}^{-1}$).

In this chapter, we investigate the fast folding of cytochrome cb_{562} with probes that report on helix 1-heme and helix 3-heme folding. Dansyl (Dns) fluorophores are attached to K19C or D66C in single-cysteine variants (Figure 5.1). By comparing the rate constants for Dns19- and Dns66-cytochrome cb_{562} , we learn whether the protein folds with a concerted or sequential mechanism.

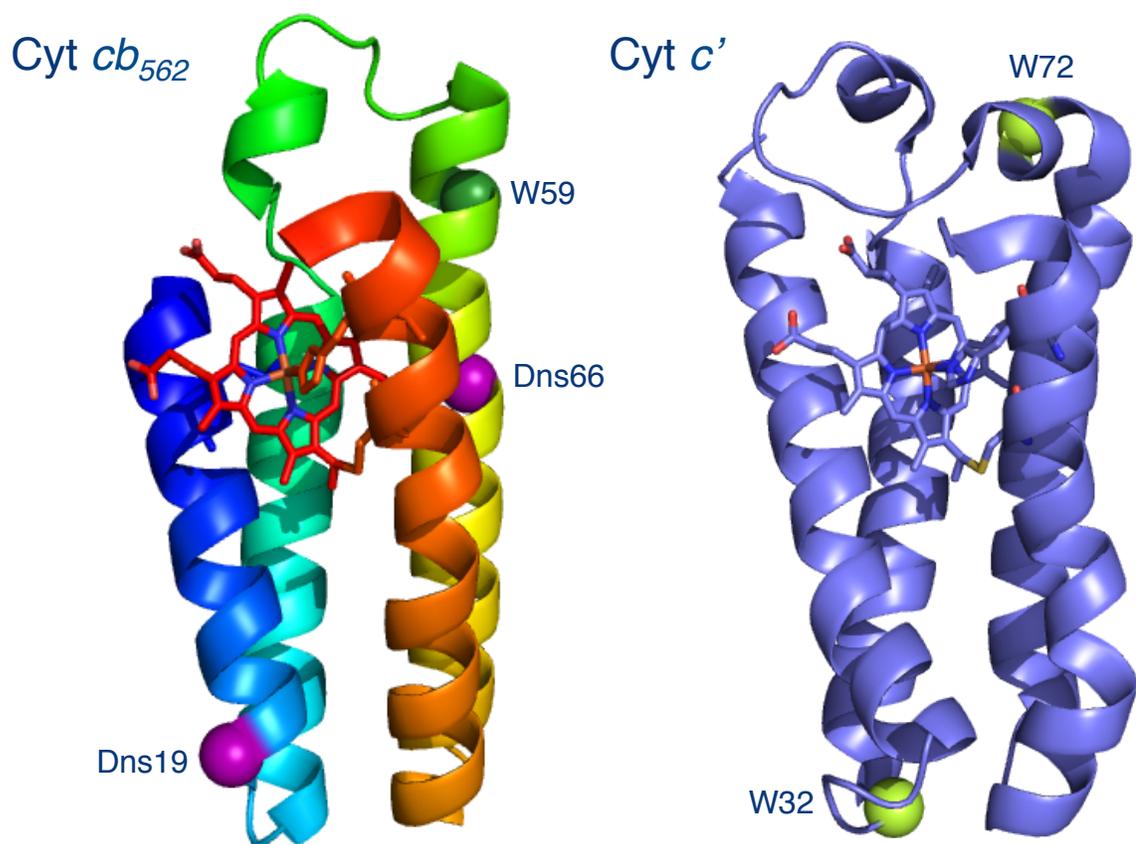


Figure 5.1: Cytochrome cb_{562} with helix I to helix IV/heme colored from blue to red (PDB 2BC5) and cytochrome c' (PDB 1MQV). Dansyl-labeling positions and tryptophan residues are highlighted as purple and green spheres, respectively. Folding is measured as a function of fluorophore to heme distance.

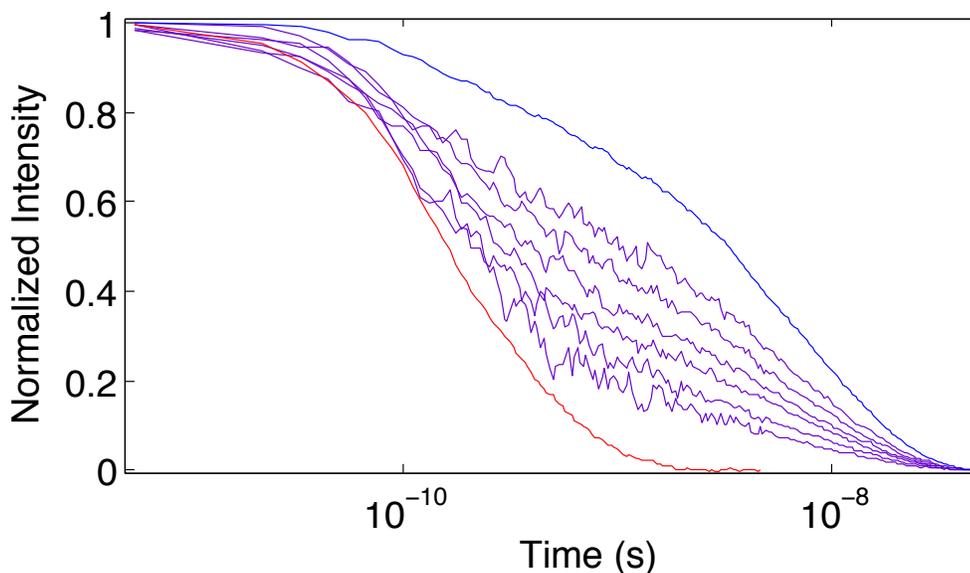


Figure 5.2: Dns66 fluorescence decays: unfolded (blue), during the folding reaction (purple), and folded (red). Dns66-cytochrome cb_{562} folding was triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 4.1, ambient temperature).

5.2 Dansyl-D66C: Folding of Helices 3-4

We triggered refolding using a continuous flow mixer to dilute denatured protein ($[Gdn]= 6$ M) with folding buffer ($[Gdn]= 0$ M) on a sub-millisecond timescale ($[Gdn]_{\text{final}}= 1$ M) and monitored the reaction progress using trFRET. Measurements were performed at pH 4 to prevent heme misligation by His63. The Dns66 fluorescence decay rate increases as the protein folds, indicating more efficient energy transfer to the heme (Figure 5.2). $P(r_{\text{DA}})$ distributions for Dns66 were extracted from fitting the fluorescence decay curves (Figure 5.3). Populations of intermediate and extended states decrease over time as additional compact structures are formed.

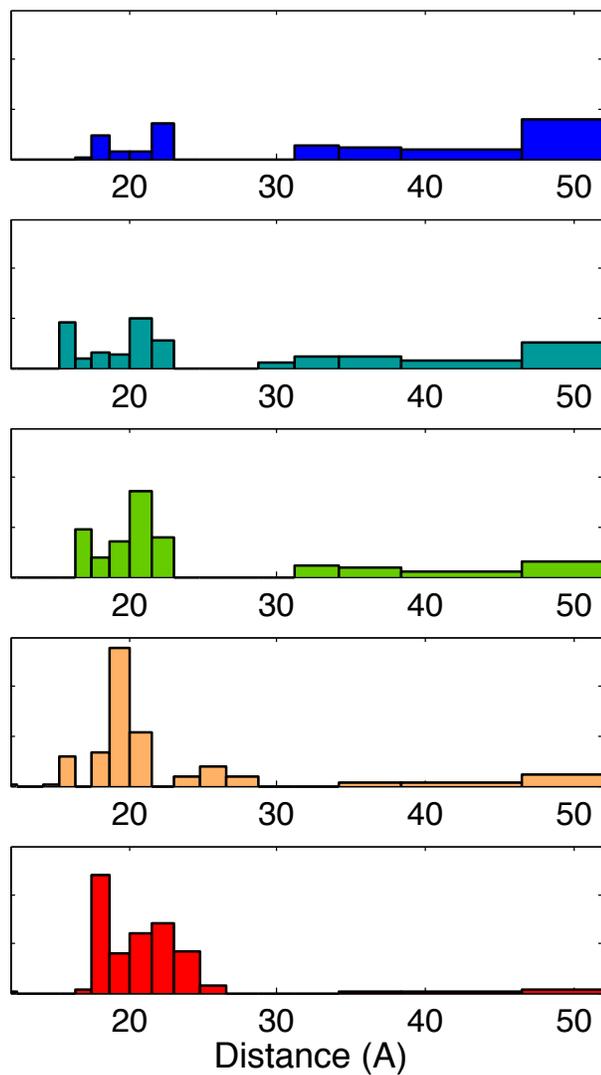


Figure 5.3: Conformational changes of cytochrome cb_{562} during folding, probed by Dns66-heme distances: unfolded (blue), after 1.5 ms (teal), after 4 ms (green), after 6 ms (orange), and folded (red).

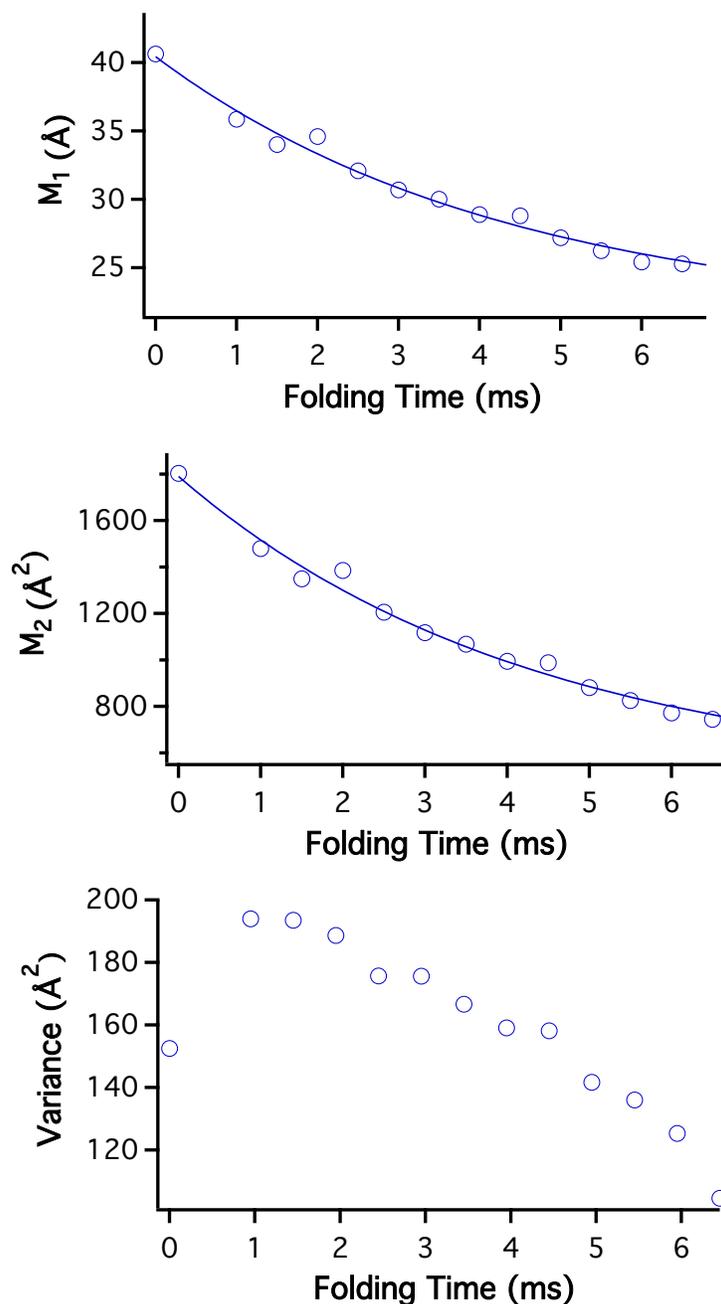


Figure 5.4: Analysis of the $P(r_{DA})$ fits of the folding kinetics of Dns66-cytochrome cb_{562} . (A) Mean distance between Dns66 and the heme (M_1) as a function of folding time. The solid line is a monoexponential fit with a rate constant of 230 s^{-1} . The infinite-time value of M_1 for the folded protein (22 \AA) was included in the fit. (B) Time course of the second moment (M_2). The solid line is a monoexponential fit with a rate constant of 230 s^{-1} . The infinite-time value of M_2 for the folded protein (500 \AA^2) was included in the fit. (C) Time course of the variance. The variance for the folded protein is 33 \AA^2 . For a two-state folding process, the variance should be biphasic with a growth rate constant equal to $2k_{\text{obs}}$ and a decay equal to k_{obs} [29].

The kinetics of $P(r_{\text{DA}})$ evolution were evaluated by moment analysis (see Methods 2.5). For a two-state process, the time course of the first (M_1) and second (M_2) moments will be exponential, with a rate constant corresponding to k_{obs} for the reaction [29]. We measure a k_{obs} of 230 s^{-1} (Figure 5.4A/B). The folded M_1 and M_2 were included in the fits; the trendlines agree well with the folding data and extrapolate to the moments of the native protein. The time course of the variance (V) in a two-state transition is biphasic, with a growth rate constant equal to $2k_{\text{obs}}$, and a decay equal to k_{obs} [29]; two phases are observed, consistent with a two-state process (Figure 5.4C). These results indicate that helix 3 forms native contacts with helix 4 by a two-state transition from the unfolded state ensemble to the native state with a rate constant of 230 s^{-1} .

5.3 Dansyl-K19C: Folding of Helices 1-4

Dns19 results parallel Dns66 findings. $P(r_{\text{DA}})$ distributions for the refolding of Dns19 (Figure 5.6) were extracted from fitting the fluorescence decay curves (Figure 5.5). An increase in energy transfer efficiency is observed throughout folding, consistent with the compaction of protein conformations. The unfolded state ensemble of Dns19 consists almost exclusively of extended structures, which gradually contract to intermediate then compact states during folding.

The kinetics of $P(r_{\text{DA}})$ evolution were evaluated by moment analysis. We fit the time courses of M_1 and M_2 to k_{obs} of 220 and 230 s^{-1} , respectively (Figure 5.7A/B). The folded M_1 and M_2 were included in the fits; the fits are consistent with the

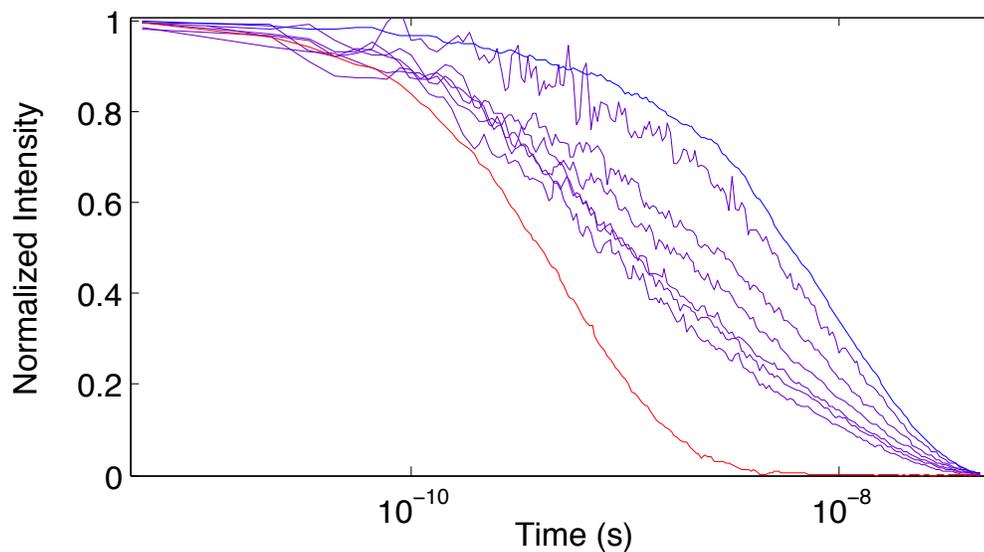


Figure 5.5: Dns19 fluorescence decays: unfolded (blue), during the folding reaction (purple), and folded (red). Dns19-cytochrome cb_{562} folding was triggered by Gdn jump from 6 to 1 M in a continuous flow mixer (pH 4.1, ambient temperature).

moments of the native protein and partially folded ensembles. Two phases are observed for V , as predicted for a two-state process (Figure 5.7C). These results indicate that the protein structure between the helix 1/2 loop and the heme folds by a two-state transition with a rate constant of 220–230 s^{-1} .

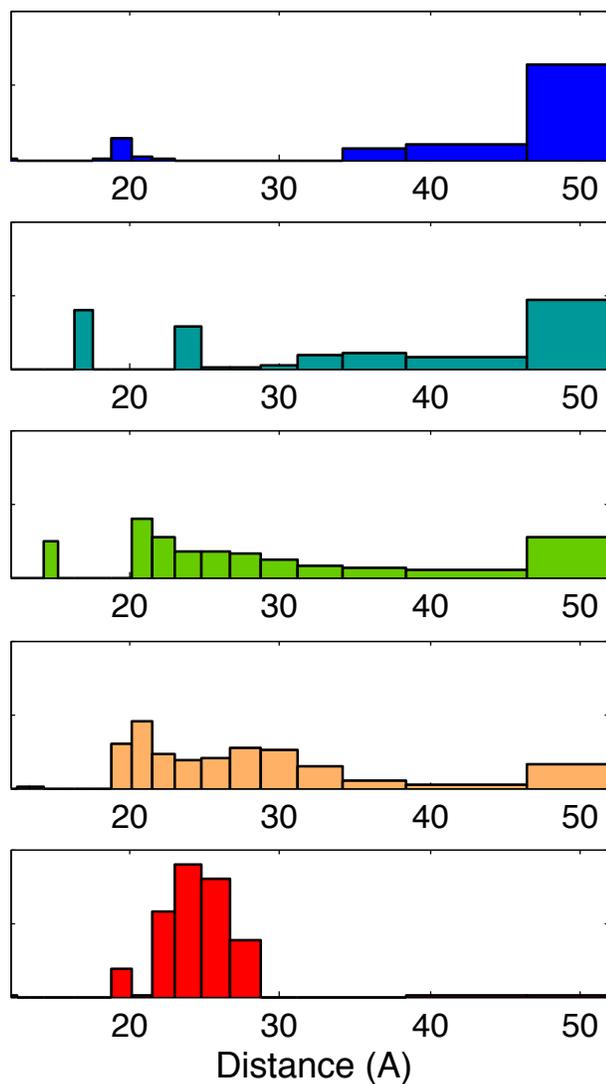


Figure 5.6: Conformational changes of cytochrome cb_{562} during folding, probed by Dns19-heme distances: unfolded (blue), after 1.3 ms (teal), after 3.3 ms (green), after 5.3 ms (orange), and folded (red).

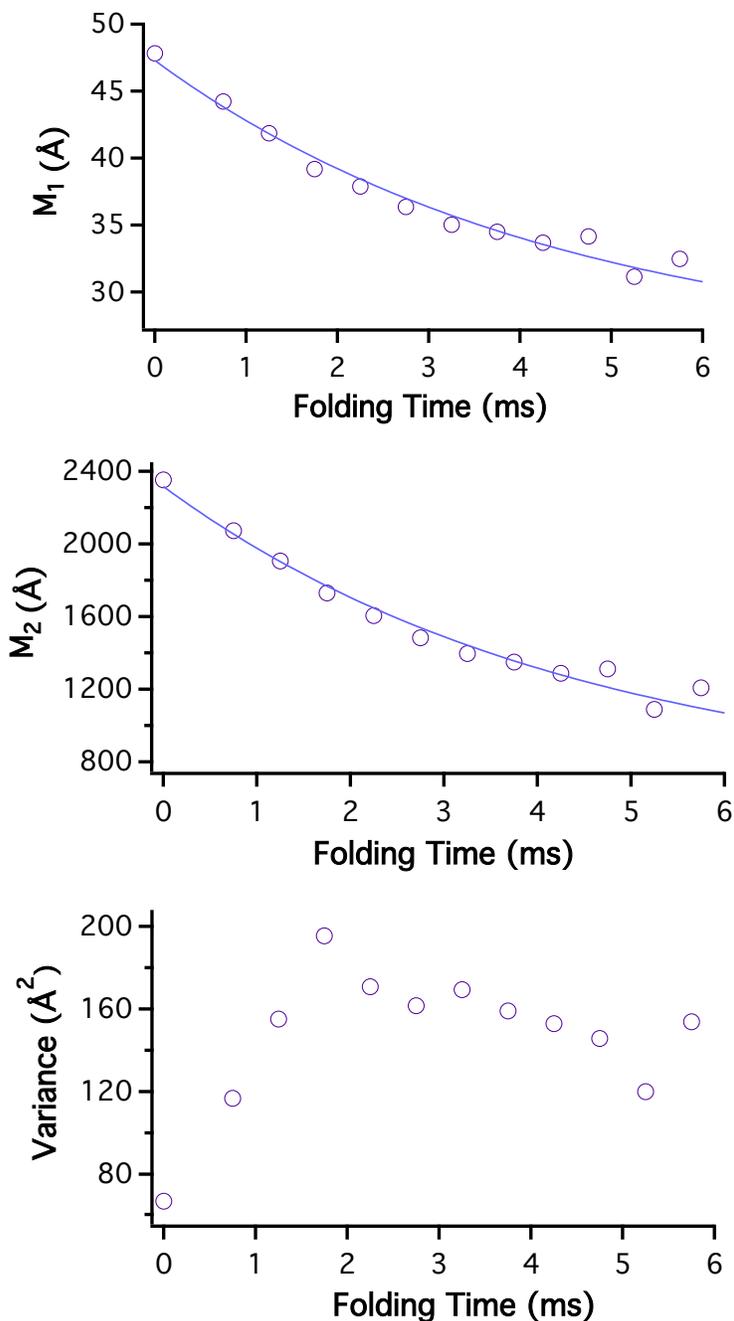


Figure 5.7: Analysis of the $P(r_{DA})$ fits of the folding kinetics of Dns19-cytochrome cb_{562} . (A) Mean distance between Dns19 and the heme (M_1) as a function of folding time. The solid line is a monoexponential fit with a rate constant of 220 s^{-1} . The infinite-time value of M_1 for the folded protein (25 Å) was included in the fit. (B) Time course of the second moment (M_2). The solid line is a monoexponential fit with a rate constant of 230 s^{-1} . The infinite-time value of M_2 for the folded protein (620 Å^2) was included in the fit. (C) Time course of the variance. The variance for the folded protein is 11 Å^2 . For a two-state folding process, the variance should be biphasic with a growth rate constant equal to $2k_{\text{obs}}$ and a decay equal to k_{obs} [29].

Cyt cb_{562} : A D [L E D N M E T L N D N L K V I E K] A D N 22
 Cyt c' : A T D [V I A Q R K A I L K Q M G E A T K P I A A M L K] G E A K F D 33
 [A A Q V K D A L T K M R A A A L D A] Q K A T P P K L E D K S P D S 55
 [Q A V V Q K S L A A I A D D S K K] L P A L F P A D S K T G G D T A A L P 69
 [P E M W D F R H G F D I L V G Q I D D A L K L A N] E G K 83
 [K I W E D K A K F D D L F A K L A A A T A A Q G T] I K D 98
 [V K E A Q A A A E Q L K T T C N A C H Q K Y] R 106
 [E A S L K A N I G G V L G N C K S C H D D F] R A K K S 125

Figure 5.8: Sequences of cytochromes cb_{562} and c' with the helices separated by brackets. Hydrophobic residues are highlighted in purple. Dansyl-labeling sites are colored blue. Tryptophan residues (and Phe32 of the F32W/W72F cytochrome c' variant) are shown in green. Cysteine residues binding the heme and axial ligands are shown in red and orange, respectively.

5.4 Folding Mechanism of Cytochrome cb_{562}

The amino acid sequences of cytochromes cb_{562} and c' (Figure 5.8) give some insight into the divergence of folding kinetics. Cytochrome c' has a more hydrophobic sequence than cytochrome cb_{562} and shows evidence of a hydrophobic collapse, with one of the refolding populations more compact than the native state [28]. Hydrophobic collapse is not observed for cytochrome cb_{562} . Helix 1 folds on a similar timescale in both proteins (4 ms), but helices 2–3 fold significantly slower in cytochrome c' (70 ms) than in cytochrome cb_{562} (4 ms). Collapse of hydrophobic patches in these regions could slow formation of helices and native intermolecular contacts [72, 73].

Frustration calculations [74] for native contacts in these proteins are consistent with our kinetics data. Both cytochromes cb_{562} and c' show a large degree of frustration between helix 1 and helix 2 (Figure 5.9). However, cytochrome c' has

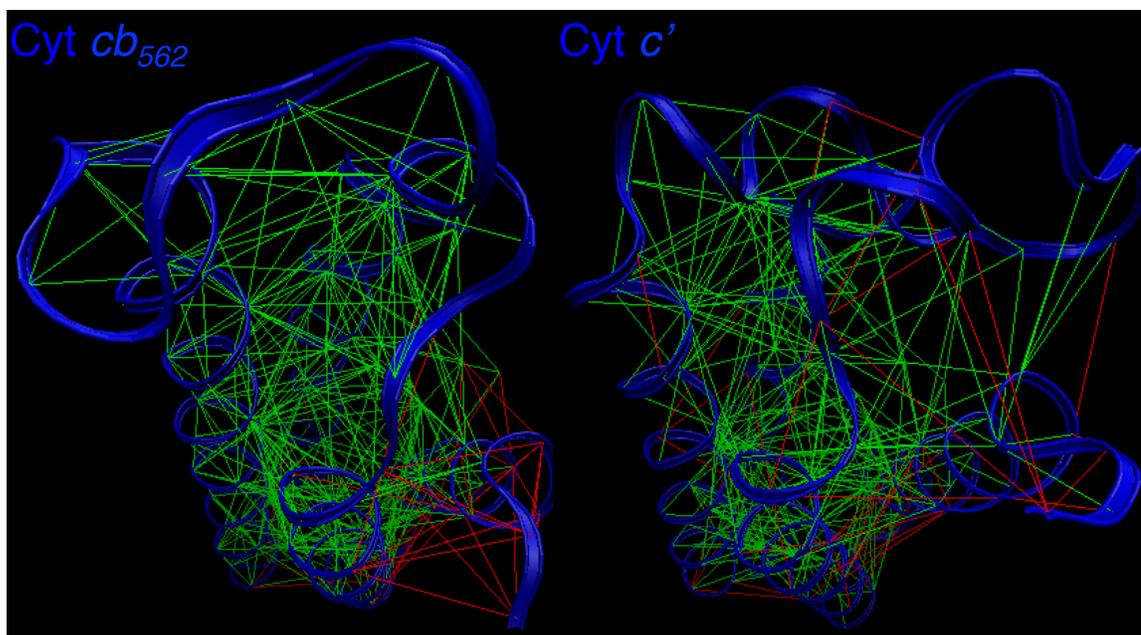


Figure 5.9: Top-down view of cytochromes cb_{562} (PDB 2BC5) and c' (PDB 1MQV) with the loop region in the foreground and the N-terminus in the bottom right corner. Frustrated contacts and minimally frustrated contacts are highlighted in red and green, respectively. The methodology is described in reference [74].

additional frustrated contacts within the helix 2/3 loop region and between the loop and the N-terminus. This indicates that these contacts are not energetically optimized, and nonnative contacts may be more favorable, which would slow folding. Cytochrome cb_{562} folding in this region is expected to be faster than for cytochrome c' .

Folding kinetics of cytochrome cb_{562} variants Dns19, Dns66, and Trp59 are similar, indicative of a concerted folding mechanism. Our results provide strong support for cooperative, two-state folding. These fluorophores probe the majority of the protein, between the loop end of helix 1 and the heme, close to the C-terminus. The N-terminus could fold at a different rate than Dns19; however, the fast-phase rate constant observed by both heme absorption and Trp59 fluorescence

after the stopped flow mixing dead time suggests that the heme environment (including interactions with helix 1) forms during macromolecular assembly [37]. The modeled frustration in helix 1, thus, does not slow folding. An intermediate with helix 1 unfolded was observed after the rate-limiting transition state for engineered variants of, or "redesigned," apocytochrome b_{562} [69, 70, 71]. This hidden intermediate could also be involved in cytochrome cb_{562} folding, but may not be relevant to the heme-containing protein if the energy landscape is affected by the presence of the heme or the altered sequence.

Our experiments of helix 3 folding (Dns66) were carried out at pH 4 to avoid heme misligation in the unfolded protein (Chapters 3-4). Kinetics measurements for Trp59, nearby on helix 3, were carried out at pH 5 [29], where His63 is ligated to the heme in about 40% of the molecules. The presence of this contact does not significantly impact the folding rate; Trp59 probed folding is only marginally faster ($k = 240\text{--}260\text{ s}^{-1}$) than Dns66 folding ($k = 230\text{ s}^{-1}$). This robustness of the rate suggests that the contact does not divert molecules along a slower folding pathway. His63 is relatively close to the heme in the native protein (14 Å) and thus may not inhibit, and may even aid, native contact formation. The folding kinetics may also be unaffected if the conversion between the misligated and fully extended forms is fast. (The conversion time is unknown at 1 M Gdn, but greater than several microseconds at 6 M Gdn.) The energy landscape theory predicts malleable folding pathways in minimally frustrated proteins; therefore, subsequent folding events may not be affected by the early conformational

constraint.

5.5 Conclusions

Cytochrome cb_{562} exhibits cooperative two-state folding on the millisecond timescale, in contrast to cytochrome c' , which exhibits slower folding in the highly frustrated helix 2/3 loop. Folding kinetics studies reveal that the region between the helix 1/2 loop and the heme folds at the same rate as the region between helix 3 and the heme in a concerted folding event. The rate of helix 3 folding is unaffected by His63 ligation of the heme in the unfolded state. We have previously shown that helices 3–4 of cytochrome cb_{562} are minimally frustrated (Chapter 4); these additional studies suggest that the majority of the protein (helix 1/2 loop to the heme) is minimally frustrated.

5.6 Acknowledgments

The frustration calculations were performed by Patrick Weinkam while in Peter Wolynes' group at the University of California, San Diego.