CHAPTER 1
Introduction to Heme Proteins

1.1 Techniques Used to Study Protein Folding Pathways

Proteins are synthesized by the ribosome in a stepwise fashion, emerging as long, linear polypeptides that quickly fold to their native state. Characterization of the unfolded states of proteins is critical for fundamental understanding of protein folding pathways and, subsequently, is an integral part of experimental and theoretical studies.

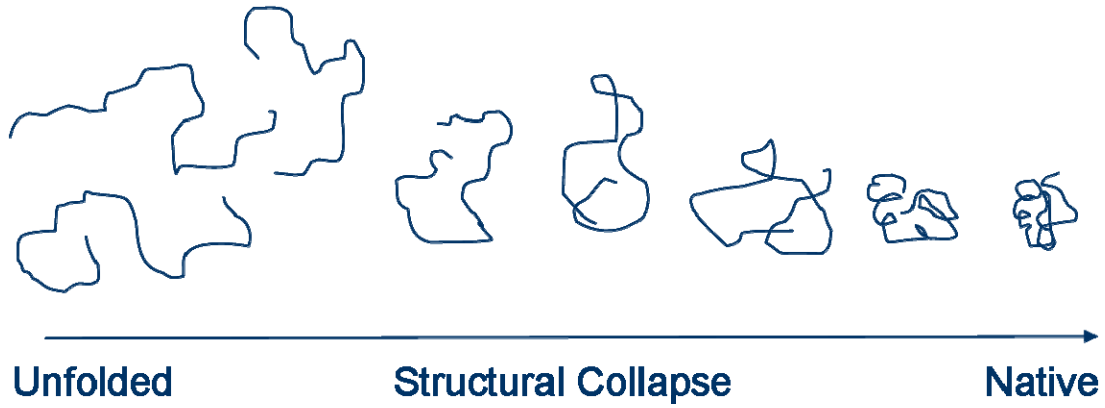

Figure 1.1. Protein folding reaction illustrating the progression of a protein from an extended structure to the native, functional structure.

A wide variety of experimental techniques have been employed to study proteins and their structure, the most common being Nuclear Magnetic Resonance (NMR), X-ray crystallography, circular dichroism (CD), Atomic Force Microscopy (AFM), fluorescence and absorbance.[1-8] While NMR and X-ray crystallography provide a great deal of information on protein structure, the research within this thesis investigates the unfolded states of heme proteins, which are better studied by spectroscopic measurements.

Metalloproteins, specifically the cytochromes, offer advantages for protein folding studies. In addition to being robust in a wide range of conditions, they also contain an optical probe, the iron-porphyrin, which is sensitive to protein structure and environment.[9] In addition, cofactors in metalloproteins introduce interesting features into the folding landscape. Cytochromes, like many metalloproteins, adopt discrete structures only when the heme cofactor is present.[11] Since the heme provides crucial contacts in the native protein, it is not surprising that this hydrophobic prosthetic group plays an important role in the folding mechanism. As such, the heme remains a useful probe for elucidating intramolecular changes during folding reactions.[12-15]

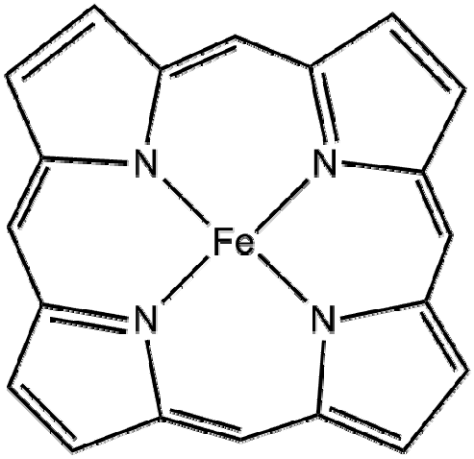

Figure 1.2. Structural framework of an iron phorphyrin. The porphyrin is similar between cytochromes, however the axial ligands are sequence and condition

The work in this thesis expands upon previous research in the Gray Group that has shown heterogenous unfolded populations in three alpha-helical proteins in the electron transport family of cytochromes: both horse heart[12,16,17] and *Saccharomyces cerevisiae* iso-1[10,15,18-21] cytochrome *c* (cyt *c*), *Rhodopseudomonas palustris* cytochrome *c*' (cyt *c*')[14,22-25], and the engineered cytochrome cb_{562} (cyt cb_{562})[26,27] from *E. coli* cytochrome b_{562}. While cyt *c* is partially helical, both cyt *c*' and cyt cb_{562} are classified as four-helix-bundle proteins. All three of these proteins are soluble, monomeric, contain covalently-bound porphyrin cofactors, and have a single tryptophan in their amino acid sequence. The electronic transitions of the porphyrin in each of these three proteins

convey information on the identity of ligands coordinated to the heme center, the exposure of the heme to the solvent, and conveniently, can act as an energy acceptor for fluorescence energy transfer (FET).[9,28] The single tryptophan is a natural fluorophore; its emission can be quenched by the heme in a distance-dependent manner, giving rise to information on the protein conformation.[29]

1.1.2. Absorbance Spectroscopy

Two absorption bands are most useful in cytochromes: the Soret band ($\pi \rightarrow \pi^*$),in the 400 nm region, and the Q band (also from porphyrin) found around 530 nm.[10] The absorbance maxima and molar absorptivities of these transitions are sensitive to the oxidation state of the metal center and the axial ligation, therefore also to different folding conditions (Figure 1.1) Absorbance is therefore a very useful measure of the environment of proteins and is utilized to probe small changes in structure and transient excited states. For instance, the Soret band of Fe(III)-cyt *c* shifts from 411 nm to 407 nm upon unfolding. An equilibrium denaturation curve can be constructed from this change in absorbance, providing information on the stability of the protein.

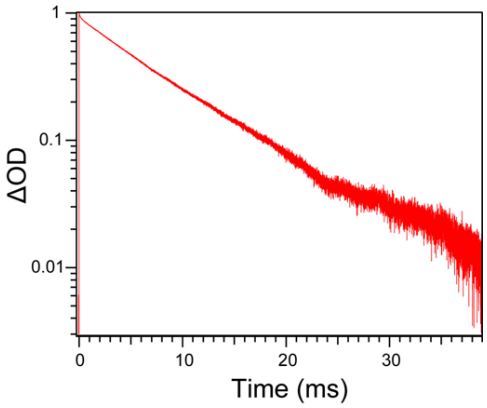

Figure 1.3. Long-lived triplet excited state of Zn-porphyrin.

Transient absorbance can also be used to measure excited states as a function of time, such as the triplet state lifetime of Zn-porphyrins described in Chapter 2.

1.1.3 Fluorescence Spectroscopy

Since native fluorophores such as tryptophan and tyrosine naturally occur in proteins, and it is relatively simple to introduce synthetic fluorophores to proteins,

fluorescence is a widely used probe of environment in biological systems. Tryptophan residues are very sensitive to environment and can be used to determine the general structure of proteins. For instance, the tryptophan emission maximum shifts to shorter wavelength upon entering a hydrophobic environment such as a membrane. In addition, it emission can be quenched through dipole-dipole interactions with a fluorescence acceptor[28].

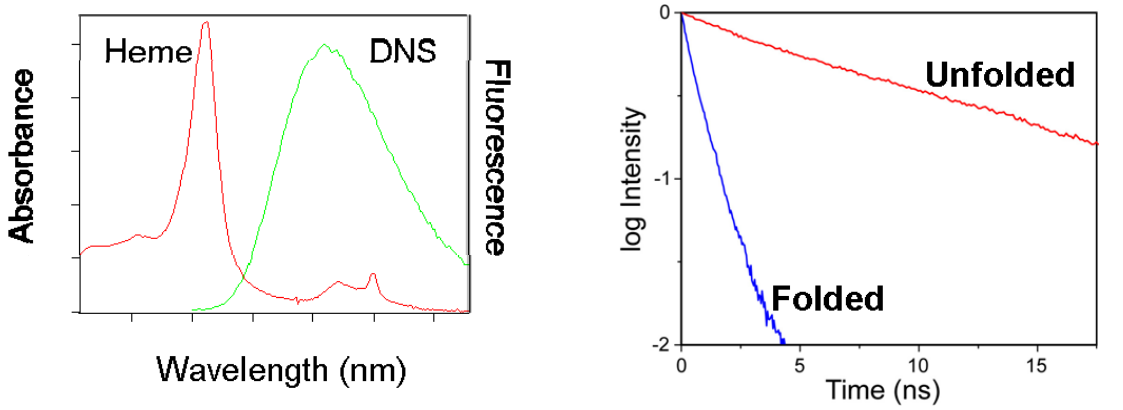

Figure 1.4. The spectral overlap (left) of the absorption spectrum of Cyt *c* (red) and emission of Dansyl model complex (green; at excitation wavelength 355 nm). The overlap of the ***D*** emission and ***A*** absorption results in distance-dependent fluorescence quenching (right).

The experimental techniques used to study the denatured states of proteins usually measure the average of all of the conformations in the ensemble and, therefore, reveal no details about the population distributions of these conformations. NMR, SAXS, and single molecule techniques have the capacity to resolve equilibrium population distributions, yet such experiments have limitations.[1-4] Unlike theory or single molecule experiments, most ensemble measurements cannot directly monitor the population distribution of conformations as a protein sample evolves from its unfolded to native state. Fluorescence energy transfer (FET) kinetics measurements provide a means of extracting these population distributions (distance distributions) directly from a bulk protein sample.[28] In folded Fe-cyt *c*, tryptophan fluorescence is quenched by the heme

due to overlap with the porphyrin absorption (Figure 1.2).[29] As this relationship is dependent on the distance between the donor and acceptor, the tryptophan fluorescence is quenched to a lesser degree in the extended structures of unfolded Fe-cyt *c*. Thus, monitoring the FET kinetics provides information on the structural conformation of the protein.

1.2 Protein Systems

1.2.2 Cytochrome *c*

Cytochrome *c* is a partially helical protein with a six-coordinate, low-spin hydrophobic porphyrin group that introduces complexity into its folding landscape. As mentioned above, the iron-porphyrin group of horse heart cytochrome *c* is a natural optical probe, providing structural information on the protein. When brighter fluorescence than the native W59 is desired, a yeast cytochrome c that contains a cysteine at the 102 position in the C-terminal alpha helix can be chemically modified with a fluorophore (note that cysteine residues can be introduced at desired positions through site-directed mutagenesis). While these two cyt *c* variants have nearly identical structures, the yeast-derived protein is less stable by approximately ΔG_f°=15 kJ mol^{-1}. The 60 % sequence overlap between the 108 residue *Saccharomyces cerevisiae* iso-1 homolog and the 105 residue horse variant is mostly in the heme region, with the differences primarily in the surface-exposed residues. Structures and sequences of the two cyt *c* homologs can be found in Figures 1.3 and 1.4, respectively.

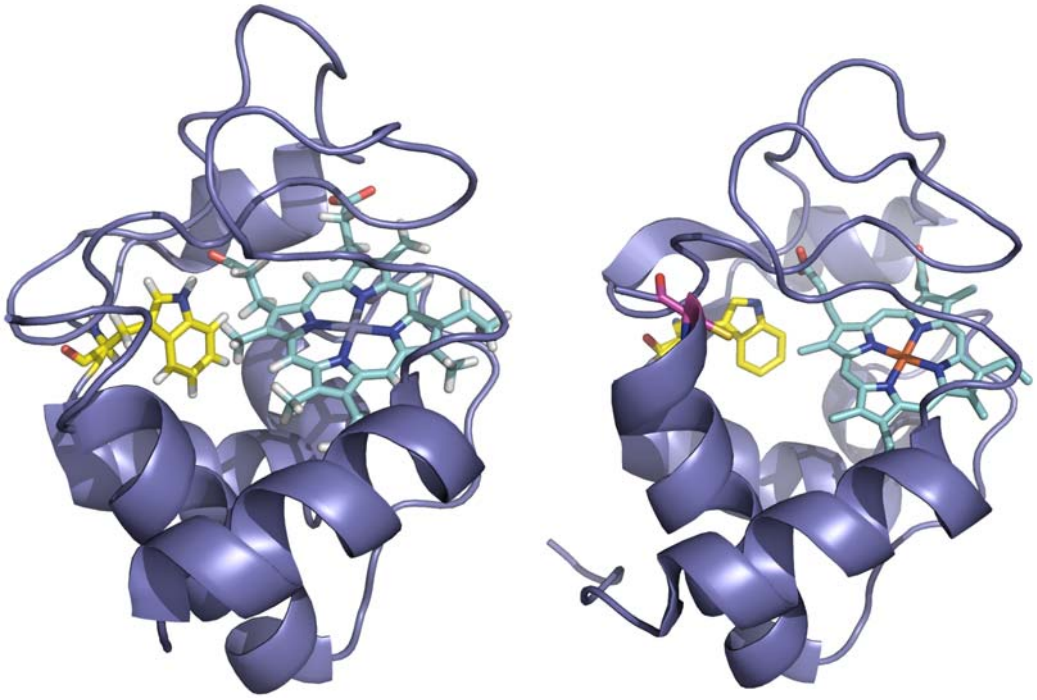

Figure 1.5. Crystal structures of horse heart cyt *c* (left) and *Saccharomyces cerevisiae* iso-1 cyt *c* (right). The heme group is highlighted in light blue; the native tryptophan residue is highlighted in yellow; the cysteine residue in yeast cyt *c* is highlighted in red. PDB: 1HRC (horse) and 1YCC (yeast).

```
HORSE:        GDVEKGKKIFVQKCAQCHTV  20
YEAST:  TEFKAGSAKKGATLFKTRCLQCHTV  25

  EKGGKHKTGPNLHGLFGRKTGQAPGFTYTD  50
  EKGGPHKVGPNLHGIFGRHSGQAEGYSYTD  55

  ANKNKGITWKEETLMEYLENPKKYIPGTKM  80
  ANIKKNVLWDENNMSEYLTNPXKYIPGTKM  85

  IFAGIKKKTEREDLIAYLKKATNE    104
  AFAGLKKEKDRNDLITYLKKACE     108
```

Figure 1.6. Sequences of horse heart cyt *c* (top) and *Saccharomyces cerevisiae* iso-1 cyt *c* (bottom). The conserved residues are highlighted in blue; the cysteines covalently bound to the porphyrin are highlighted in red; the axial ligands are highlighted in orange; the tryptophan residues are highlighted in green; the cysteine residue in yeast cyt *c* is highlighted in purple.

Since the porphyrin cofactor introduces complexity to cyt *c*, it is of great experimental interest. In order to study the heme through transient absorbance, the native iron metal center in h-cyt *c* can be substituted zinc, which has a long-lived triplet excited state that is sensitive to the local environment. It has been determined by both NMR and crystal structure that Zn-cyt *c* protein remains structurally equivalent to Fe(III)-cyt *c*.[2,3] Similarly, the stability of folded Zn-cyt *c* (ΔG_f° = -35 kJ/mol) tracks with Fe(III)-cyt *c* (ΔG_f° = -40 kJ/mol).

Former group members Julia Lyubovitsky, Jennifer Lee, and Kate Pletneva have extensively studied cyt *c* folding rates.[10,15,18-21] They found, for both Fe^{III} and Zn^{II}, rapid equilibrium between extended and nonnative compact structures, which then fold to the native conformation in milliseconds. The rate-limiting step of Fe(III)-cyt *c* folding is the substitution by Met80 for a non-native ligand, His33 (or H26, H33, H39 for y-cyt *c*)[10,30]. When the iron center of cyt *c* is replaced with zinc, the result is faster folding since the histidine misligation trap is eliminated.[25]

Previous studies have also shown heterogeneity in the unfolded states of cytochromes.[10,19] Since it is important to fully characterize the unfolded states in order to understand folding pathways, this thesis focuses on the local environment of the heme in Zn-cyt *c* and distance distributions from FET kinetics of DNS(C102)-cyt *c*. In order to examine the role of the protein and its hydrophobic character on the heme environment during folding, an enzymatically cleaved form of h-cyt *c*, *N*-acetyl-microperoxidase-8 (AcMP8), acts as a model of a fully solvent-exposed porphyrin group.[17,31]

1.2.3 Four-helix-bundles

A second class of protein, the four-helix bundle, raises an interesting issue of vastly different folding rates for structurally similar proteins.[26] Cytochrome *c*' from *Rhodopseudomonas palustris*, a 125-residue, high-spin, 5-coordinate heme protein, folds in a heterogenous nature, with some time scales (seconds) much greater than the folding rate of cyt *c*. Cytochrome cb_{562}, a 106 residue, low-spin, 6-coordinate heme protein with nearly identical in structure to cyt *c*', folds quickly, reaching the native state in less than a millisecond. Structures and sequences of cyt *c*' and cyt cb_{562} can be found in Figures 1.5 and 1.6, respectively.

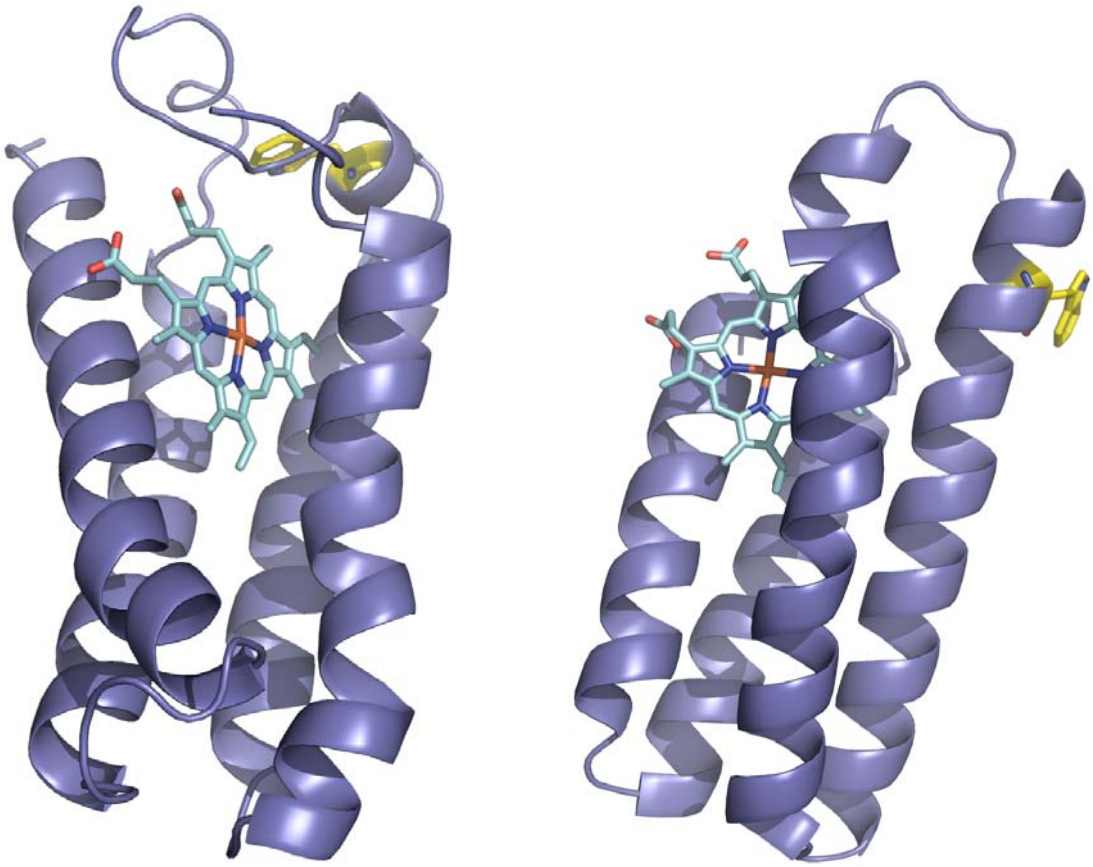

Figure 1.7. Crystal Structures of cyt *c*' (left) and cyt cb_{562} (right). The heme group is highlighted in light blue; the native tryptophan residue is highlighted in yellow. PDB: 1A7V (*c*') and 2BC5 (cb_{562}).

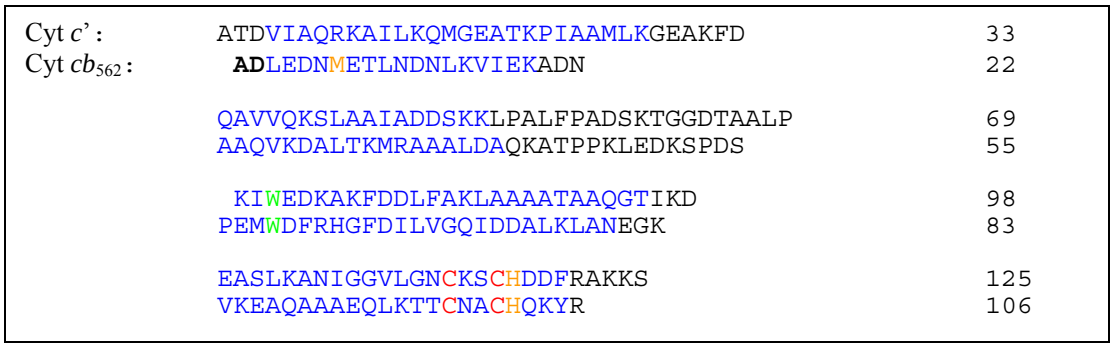

Figure 1.8. Sequences of cyt *c'* (top) and Cyt cb_{562} (bottom). The residues of the four helices that make up the bundle are highlighted in b; the cysteines covalently bound to the porphyrin are highlighted in red; the axial ligands are highlighted in orange; the tryptophan residues are highlighted in green.

1.3 Thesis Outline

The role of solvation in protein folding was investigated in Chapter 2. The triplet excited state of the Zn-porphyrin decays through coupling to the solvent, revealing a greater isotope effect for more solvent exposed conformations. Chapter 3 expands upon the isotope effect studies to reveal differences in the extended structures of temperature, guanidinium chloride-, and urea-denatured Zn- and Fe-substituted cyt *c*, cyt *c*', and cyt cb_{562} proteins. Thermally denatured DNS-cyt *c* is investigated in Chapter 4. Fluorophores developed for fluorescence and fluorescence energy transfer studies are reviewed in Chapter 5. Finally, Chapter 6 presents initial attempts to study protein folding under crowded conditions that better represent the macromolecularly crowded cell environment.

REFERENCES

(1) Sosnick, T.R.; Mayne, L.; Hiller, R.; Englander, S.W. The Barriers in Protein Folding. Structural Biology **1994**, *1*, 149-156.

(2) Dobson, C.M. Experimental Investigation of Protein Folding and Misfolding. *Methods* **2004**, *34*, 4-14.

(3) Krishna, M.M.G.; Hoang, L.; Lin, Y.; Englander, S.W. Hydrogen Exchange Methods to Study Protein Folding. *Methods* **2004**, *34*, 51-64.

(4) Zeeb, M.; Balbach, J. Protein Folding Studied by Real-time NMR Spectroscopy. *Methods* **2004**, *34*, 65-74.

(5) Rounsevell, R.; Forman, J.R.; Clarke, J. Atomic Force Microscopy: Mechanical Unfolding of Proteins. *Methods* **2004**, *34*, 100-111.

(6) Beck, D.A.C.; Daggett, V. Methods for Molecular Dynamics Simulations of Protein Folding/Unfolding in Solution. *Methods* **2004**, *34*, 112-120.

(7) Beechem, J. M.; Brand, L. Time-Resolved Fluorescence of Proteins, *Annual Review of Biochemistry* **1985**, *54*, 43-71.

(8) *Principles of fluorescence spectroscopy*, Lakowicz, J. R. Kluwer Academic/Plenum Publishers, New York, NY, 1999.

(9) Mines, G. A.; Winkler, J. R.; Gray, H. B. Spectroscopic Studies of Ferrocytochrome *c* Folding. In *Spectroscopic Methods in Bioinorganic Chemistry*; Solomon, E.;Hodgson, K., eds. American Chemical Society: Washington, D.C., 1997; Vol. 692, pp 188-211.

(10) Lyubovitsky, J. G. Thesis, California Institute of Technology, 2003.

(11) Feng, Y.; Wand, J.; Sligar, W. G. ^{1}H and ^{15}N NMR Resonance Assignments and Preliminary Structural Characterization of *Escherichia coli* Apocytochrome-b_{562}. *Biochemistry* **1991,** 30, 7711-7717.

(12) Telford, J. R.; Tezcan, F. A.; Gray, H. B.; Winkler, J. R. Role of Ligand Substitution in Ferrocytochrome *c* Folding. *Biochemistry* **1999,** 38, 1944-1949.

(13) Shastry, M. C. R.; Sauder, J. M.; Roder, H. Kinetic and Structural Analysis of Submillisecond Folding Events in Cytochrome *c*. *Acc. Chem. Res.* **1998,** 31, 717-725.

(14) Lee, J. C.; Gray, H. B.; Winkler, J. R. Cytochrome *c*' Folding Triggered by Electron Transfer: Fast and Slow Formation of Four-helix Bundles. *Proc. Natl. Acad. Sci. USA* **2001,** 98, 7760-7764.

(15) Lyubovitsky, J.G.; Gray, H.B.; Winkler, J.R. Mapping the Cytochrome *c* Folding Landscape. *J. Am. Chem. Soc.* **2002**, *124*, 5481-5485.

(16) Mines, G.A. Thesis, California Institute of Technology, 1997.

(17) Tezcan, F.A. Thesis, California Institute of Technology, 2001.

(18) Pletneva, E.V.; Gray, H.B.; Winkler, J.R. Snapshots of Cytochrome *c* Folding, *Proc. Natl. Acad. Sci U.S.A.* **2005**, *102*, 18397-18402.

(19) Pletneva, E.V.; Gray, H.B.; Winkler, J.R. Many Faces of the Unfolded State: Conformational Heterogeneity in Denatured Yeast Cytochrome *c*. *J. Mol. Bio.* **2005**, *345*, 855-867.

(20) Lyubovitsky, J.G.; Gray, H.B.; Winkler, J.R. Structural Features of the Cytochrome *c* Molten Globule Revealed by Fluorescence Energy Transfer Kinetics. *J. Am. Chem. Soc.* **2002**, *124*, 14840-14841.

(21) Lyubovitsky, J.G.; Winkler, J.R.; Gray, H.B. Real-time Folding Dynamics of S-cerevisiae iso-1 Cytochrome *c*. *J. Inorg. Biochem.* **2001**, *86*, 322-322.

(22) Kimura, T.; Lee, J.C.; Gray, H.B.; et al. Site-specific Collapse Dynamics Guide the Formation of the Ccytochrome c ' Four-helix Bundle, *Proc . Natl. Acad. Sci U.S.A.* **2007**, *104*, 117-122.

(23) Pletneva, E.V.; Zhao, Z.Q.; Kimura, T.; et al. Probing the Cytochrorne c ' Folding Landscape. *J. Inorg. Biochem.* **2007**, *101*, 1768-1775.

(24) Lee, J.C.; Engman, K.C.; Tezcan, F.A.; et al. Structural Features of Cytochrome c' Folding Intermediates Revealed by Fluorescence Energy-transfer Kinetics. *Proc. Natl. Acad. Sci U.S.A.* **2002**, *99*, 14778-14782.

(25) Lee, J.C. Thesis, California Institute of Technology, 2002.

(26) Faraone-Mennella, J.; Tezcan, F.A.; Gray, H.B.; et al. Stability and Folding Kinetics of Structurally Characterized Cytochrome c-b(562), *Biochemistry* **2006**, *45*, 10504-10511.

(27) Kimura, T.; Lee, J.C.; Gray, H.B.; et al. Folding Energy Landscape of Cytochrome *cb*(562), *Proc. Natl. Acad. Sci U.S.A.*. **2009**, *106*, 7834-7839.

(28) Arjara, G. Thesis, California Institute of Technology, 2007.

(29) Lamola, A.A. Electronic Energy Transfer in Solution: Theory and Applications. Chapter 2 p.17-132.

(30) Krishna, M.M.G.; Lin, Y.; Englander, S.W. Protein Misfolding: Optional Barriers, Misfolded Intermediates, and Pathway Heterogeneity. *J. Mol. Biol*. **2004**, *343*, 1095-1109.

(31) Kim, J.E.; Pribisko, M.A.; Gray, H.B.; Winkler, J.R. Zinc-porphyrin Solvation in Folded and Unfolded States of Zn-cytochrome *c*. *Inorg. Chem.* **2004**, *43*, 7953-7960.