Mapping Heme Protein Folding Landscapes

Thesis by

Jennifer Chen Lee

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2002

(Defended May 28, 2002)

Acknowledgements

It is difficult not to be sappy after these wonderful five years. So, here it is... First, I want to thank my family, my grandparents and especially, my mother, for their unconditional love and support. (Even though most of the time, they are not quite sure of what I am doing.) I cannot be more indebted to my mother who has made many sacrifices so that I can have the opportunity to pursue my dreams (even if it is to become a scientist).

What can I say about my beloved advisors? There are not enough words in the world (nor am I eloquent enough, which Jay will agree) to describe how I feel about Harry and Jay. Harry is not only a world renowned scientist, but he is *the most caring, generous, and absolutely, the funniest advisor of all*! In his own right, Jay is also unbelievably supportive (always answering questions with new experiments) and fun to be around! He has taught me everything from "doing good science" to where not to go for car repairs. Unfortunately, to his dismay, I am still struggling with grammar. (After all, I am not a native speaker.) Throughout my graduate career, both Harry and Jay have been there for me everyday. Most of the time, I think Jay's office belongs to me. (It's my lunch room!) I am always sitting on his couch. Seriously, without Harry and Jay my life would not have been the same. I can't imagine what graduate school would have been like with any other advisors (or with just one). I will desperately miss our morning coffee breaks together. Who will be the next coffee girl?

In the last five years, I have been lucky to have worked with so many great people. I thank Pernilla (the first Swedish woman) for getting me started in lab.

Everything I know about molecular biology, I have learned from Michele, who is the best mentor. I will never forget working with I-Jy, the loudest and most energetic Chinese woman I know. I am honored to have had the opportunity to work with Lisa, a superb undergraduate researcher. Lastly, I have to thank Cissi (the second Swedish woman) for deciding that California winters were better than Sweden's. I will never forget our "shared" lunches and times in the laser lab.

I would also like to thank Corinna for being my best friend. She is the most loyal and dependable friend that anyone can hope for. Despite my own twisted sense of humor, she always can laugh with/at me. I don't think that I could have endured these years without her being beside me (cooking/eating, drinking, and of course, shopping!). Berkeley is lucky to get to experience her. Go BEARS!

The GRAY group has been the best! Akif has always been up for anything (coffee breaks, advice, and sometimes arguments). We have had some fun times hanging out. I would also like to acknowledge other past Gray group members, Ivan, Jon, Liz, and Adrian for always being so helpful. I would like to thank the current group members and especially my office roommate, Kristine, for putting up with me.

My graduate career is rich with some of the best experiences of my life, one of which is meeting my future husband, Seth. (Harry is a fabulous matchmaker!) Ever since I have met Seth, my life is full of laughter and free from worries (most of the time). He is a truly patient (hard-of-hearing) and loving

(forgiving) man for suffering through my proposal exam and thesis writing (and other craziness). I cannot imagine my life without him!

Lastly, I would like to thank my cat, Tiger. Although he may not always had the choice, he is always there when I come home (ever since the 7th grade). As he lies next to me now, I will never forget the support (hairball) and the love (drool) he has given me.

Abstract

The folding dynamics of two four-helix bundles, Fe^{II}-cytochrome b_{562} (*E. coli*) and Fe^{II}-cytochrome *c'* (*R. palustris*), have been investigated: in each case, folding was triggered by electron transfer (ET) to the unfolded Fe^{III}-protein and probed by transient absorption spectroscopy. Although these two proteins are topologically homologous, they display strikingly different folding kinetics: for Fe^{II}-cyt b_{562}, the extrapolated time constant in the absence of denaturant (≈ 5 μs) is near the predicted value for intrachain diffusion; in contrast, Fe^{II}-cyt *c'* folding is heterogeneous over a ms-s time range.

Nonnative methionine ligands in Fe^{II}-cyt *c'* contribute to the complex folding kinetics. Removal of either Met-15 or Met-25 leads to simpler kinetics than found in the pseudo-wild-type (Q1A) protein. The amplitude of the intermediate folding phase (10^2-10^1 s^{-1}) is diminished, suggesting that formation of near-native contacts may facilitate folding.

Cyt *c'* has a wild-type Trp-72 residue that can be exploited in fluorescence energy transfer (FET) kinetics measurements. Also, we have crystallographically characterized (1.78 Å) a Trp-mutant, Q1A/F32W/W72F, in order to probe distances between position-32 and the heme during folding. Analysis of FET kinetics from Trp-72 and Trp-32 shows that there is structural heterogeneity in the denatured protein. Even at high guanidine hydrochloride (GuHCl) concentrations, 50% of the polypeptides adopt compact conformations (25-35 Å) in cyt *c'*. A burst phase (≤ 5 ms) is revealed when stopped-flow triggered refolding is probed by Trp fluorescence intensity: measurements on the Q1A

protein show that 75% of the Trp-72 fluorescence (83% for Trp-32, Q1A/F32W/W72F) is quenched in the mixing deadtime, suggesting that most of the polypeptides have collapsed.

Electronically excited Zn-porphyrin in GuHCl-denatured Zn-substituted cytochrome *c* (Zn-cyt *c*) reduces $Ru(NH_3)_6^{3+}$ about ten times faster than when embedded in the folded protein. Measurements of ET kinetics during folding reveal a burst intermediate in which one-third of the ensemble has a protected Zn-porphyrin and slow ET kinetics; the remaining fraction exhibits fast ET characteristic of a solvent-exposed cofactor. Importantly, our FET (cyt *c*') and ET (Zn-cyt *c*) experiments show that collapsed nonnative structures are not substantially more stable than extended polypeptide conformations.

Table of Contents

List of Tables and Figures

Chapter 1

Protein folding triggered by electron transfer

Introduction

Newly synthesized polypeptide chains must fold into discrete, three-dimensional structures to function correctly. Misfolding can lead to disease and even death in organisms (1). Folding in the cell occurs in a highly crowded macromolecular environment (~ 300 mg/mL) and vectorially from N- to C-terminus (2). To ensure that proteins reach their native states and avoid potentially toxic aggregates, cells have evolved a complex machinery of molecular chaperones that assists the folding of nascent polypeptides and rescues proteins from stress-induced denaturation (3).

In vitro folding experiments, using chemically denatured proteins, have demonstrated that even in the absence of cellular components, folding of small soluble proteins is spontaneous (4). Therefore, the information for formation of native structures is encoded in the primary amino-acid sequence. Since it is impossible for unfolded polypeptides to randomly sample all possible conformations during the folding process (Levinthal's paradox), it has been suggested that proteins must fold via well-defined pathways (5). Recently, the protein folding problem has been described in terms of an energy landscape in configurational space with a deep minimum (a funnel) at the folded structure (6, 7). Understanding the mechanism of how secondary and tertiary structures of proteins are formed from nonnative conformations is a continuing challenge facing experimentalists and theorists.

Elucidation of key events in protein folding calls for investigations on timescales that range from picoseconds to minutes (Figure 1.1). The fastest

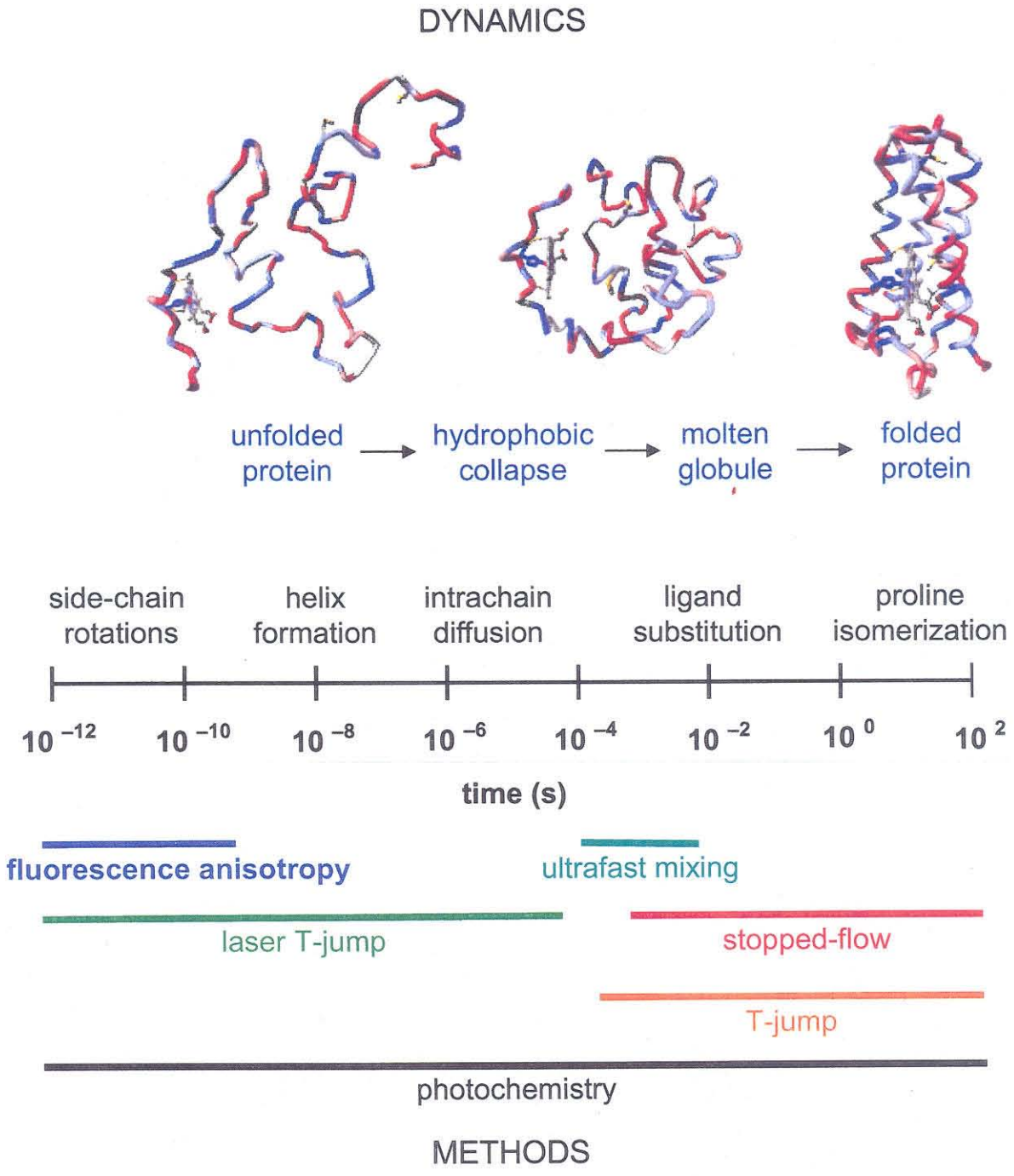


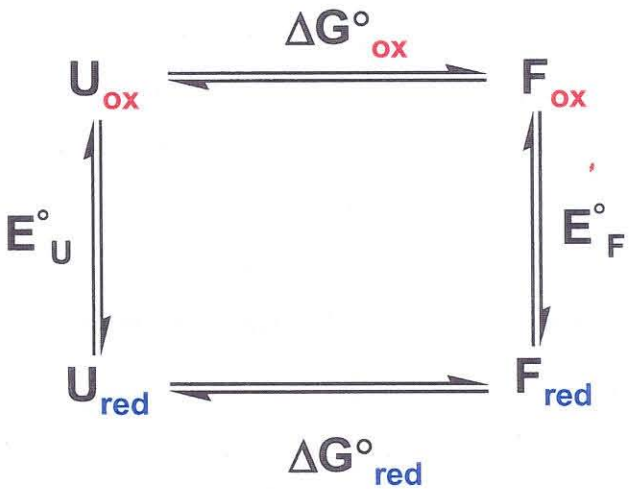
Figure 1.1. Time regimes and experimental methods employed for the study of protein dynamics.

nuclear motions, rotations about single bonds, occur on the picosecond timescale, whereas helical secondary structures are formed in hundreds of nanoseconds (8–11). Tertiary contact formation, involving large scale backbone motions, occurs in the microsecond to millisecond time range (12–14). In contrast, misfolded or trapped structures (*e.g.*, proline isomers) can take several seconds to minutes to fold correctly (15).

Formation of partially folded structures, the so-called "burst" intermediates observed during a stopped-flow mixing deadtime, has been the most challenging to study because it occurs on a submillisecond timescale (12, 13, 16–19). Identification and characterization of these compact structures are critical in understanding the role of hydrophobic collapse during protein folding. Development of rapid folding triggers (*e.g.*, ultrafast mixers and photochemical methods) is necessary for studies on these fast timescales (20–22).

Redox-coupled folding

The basic requirement for experimental investigations of protein folding dynamics is some means of triggering the folding process. We have developed a new method for initiating protein folding using electron-transfer (ET) chemistry (23-26). ET triggered folding is based on the observation that the formal potentials for redox cofactors in the hydrophobic interiors of folded proteins often are shifted dramatically from their aqueous solution values (27). A simple thermodynamic cycle can be drawn connecting an oxidized (P_{ox}) and reduced



$$\Delta\Delta G^{\circ}_{red-ox} = -nF(\Delta E^{\circ}_{F-U})$$

Figure 1.2. Thermodynamic cycle illustrating the relationship between the formal potentials of redox cofactors and the free energy of folding.

(P_{red}) protein in both folded and unfolded states (Figure 1.2). If the active-site reduction potentials are different for the folded and unfolded proteins ($\Delta E^{o}_{f} = E^{o}_{F} - E^{o}_{U}$), then the folding free energies of the oxidized and reduced proteins will differ by a comparable amount ($\Delta\Delta G^{o}_{f} = \Delta G^{o}_{f, ox} - \Delta G^{o}_{f, red}$). The shift in the reduction potential reflects the differential stabilities of the oxidized and reduced folded proteins. If the difference between the reduction potentials is sufficiently large and their folding free energies depend similarly on denaturant concentration, it is possible to find denaturant conditions where one oxidation state of the protein is fully folded while the other is unfolded (Figure 1.3). Most heme proteins have $\Delta E^{o}_{f} > 0$; thus, in the denaturant concentration region that lies between the two unfolding transitions, electron injection into the oxidized protein will lead to the formation of the reduced folded state. One very attractive feature of this technique is that the initiation of the folding reaction is only limited by the ET time.

Photochemical Triggers

There is a vast array of photoactive complexes that can rapidly inject electrons into proteins on timescales as short as a few hundred nanoseconds. Electronically excited $Ru(bpy)_3^{2+}$ ($^*Ru(bpy)_3^{2+}$; bpy= 2,2'-bipyridine) is a powerful reductant (E^{o} ($Ru^{3+}/^{*2+}$) = –0.85 V vs NHE), and its 600 ns decay time makes it an excellent reagent for reducing unfolded proteins on the microsecond timescale (23). Subsequent charge recombination regenerates initial reagents within a few milliseconds. This reversibility permits extensive signal averaging

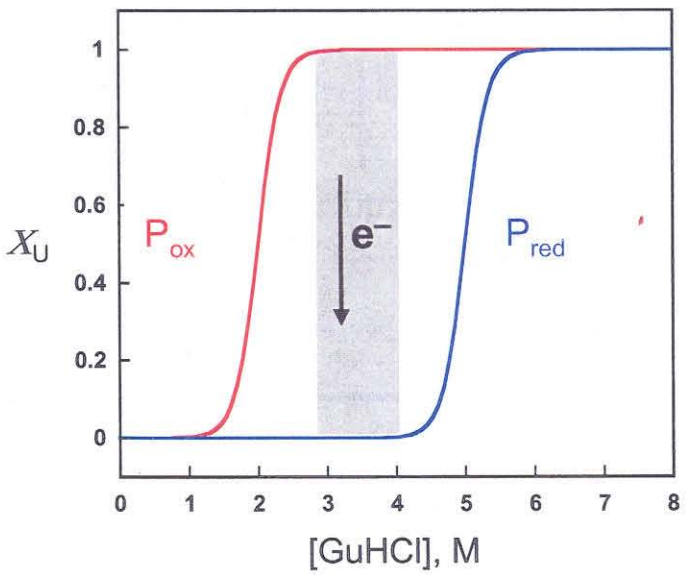


Figure 1.3. Representative denaturant titration curves for oxidized (P_{ox}) and reduced (P_{red}) forms of a redox-active protein with $\Delta E^\circ_f > 0$ eV (X_U is the fraction of unfolded protein). The shaded region represents denaturant concentrations in which ≥ 99% of the oxidized protein is unfolded and ≥ 99% of the reduced protein is folded. Under these conditions, rapid electron injection into the unfolded oxidized protein will initiate a folding reaction.

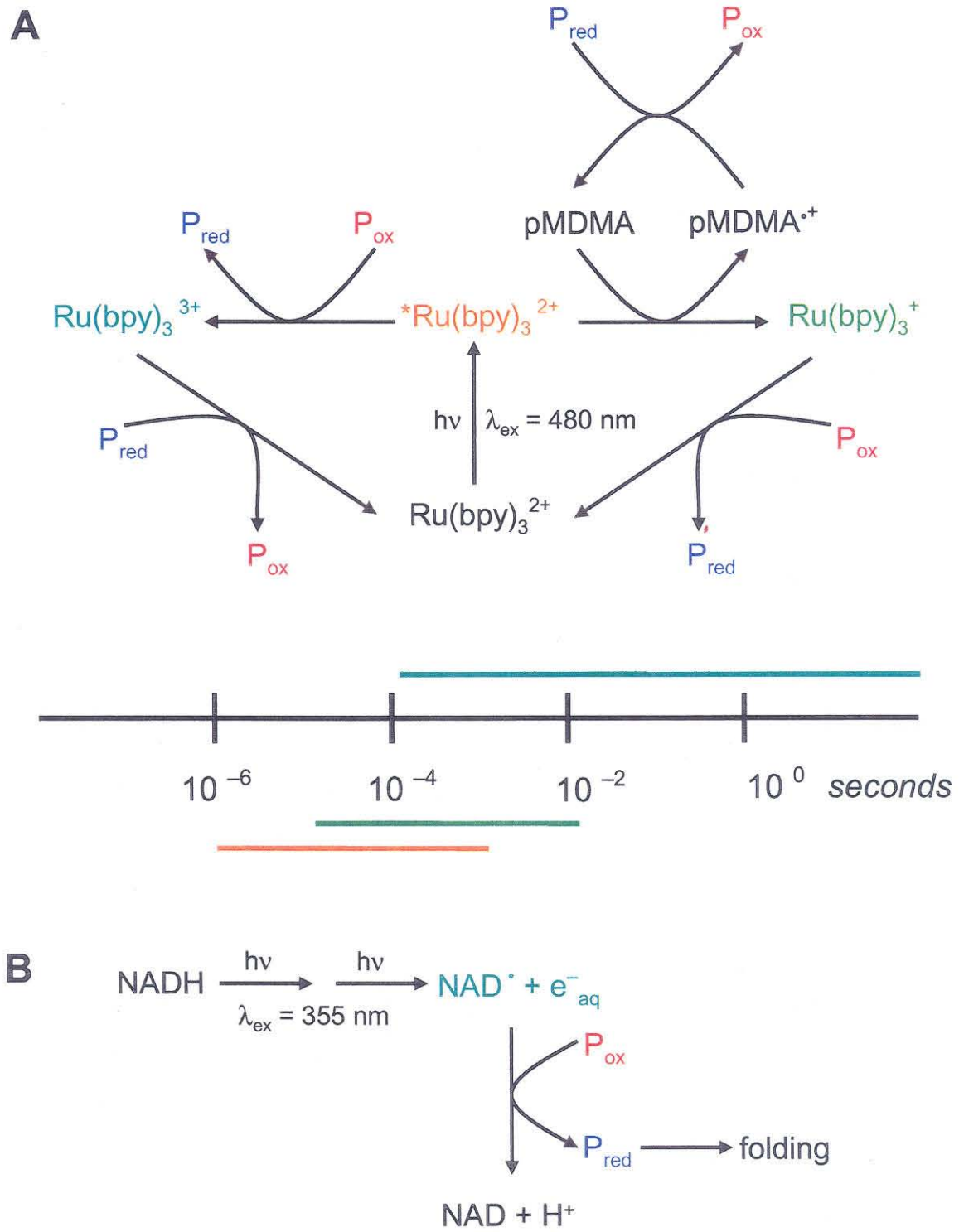


Figure 1.4. Reversible (A) and irreversible (B) reaction schemes for photochemical ET-triggered protein folding (P_{ox} and P_{red} are oxidized and reduced proteins). Time regimes for different photochemical sensitizers are also illustrated.

but limits the folding observation window to the 1–1000 μs range (Figure 1.4A). The yield of electron injection into an unfolded oxidized protein is often limited. However, reductive quenching with *p*-methoxy-N,N'-dimethylaniline (pMDMA) permits a tenfold increase in injection yield with the formation of $Ru(bpy)_3^+$ (Figure 1.4A), and with only a modest sacrifice in injection timescale (~10 μs) (28). Complete folding of a protein can require hundreds of milliseconds to many seconds. For measurements on these longer timescales, two-photon (355 nm) excitation of NADH produces reagents ($NAD^•$ and e^-_{aq}) that can reduce unfolded heme proteins (~100 μM) in about 100 μs (Figure 1.4B) (25). The combination of $^*Ru(bpy)_3^{2+}$, $^*Ru(bpy)_3^{2+}$/pMDMA, and $NAD^•/e^-_{aq}$ (NADH) permits investigations of 1-μs to > 1-s folding events in heme proteins.

Conclusion

Oxidation-state-dependent folding free energies are general properties of redox-active proteins. Photochemical ET methods allow investigations of metalloprotein folding that would be difficult to accomplish by other means, in particular, by providing the time resolution required to reveal the earliest steps in the folding process. In parallel with the exploitation of existing probes of folding, we must develop new methods that report not just on reaction progress but on ensemble heterogeneity as well.

References

1. Dobson, C. M. (2001) *Phil. Trans. Roy. Soc. B* **356,** 133-145.

2. Frydman, J. (2001) *Annu. Rev. Biochem.* **70,** 613-647.

3. Hartle, F. U. and Hartle-Hayer, M. (2002) *Science* **295,** 1852-1858.

4. Anfinsen, C. B. (1973) *Science* **181,** 223-230.

5. Levinthal, C. (1968) *J. Chim. Phys.* **65,** 44.

6. Bryngelson, J. D., Onuchic, J. N. and Wolynes, P. G. (1995) *Proteins: Struct. Func. Gen.* **21,** 167-195.

7. Onuchic, J. N., Lutheyschulten, Z. and Wolynes, P. G. (1997) *Annu. Rev. Phys. Chem.* **48,** 545-600.

8. Chen, L. X.-Q., Petrich, J. W., Fleming, G. R. and Perico, A. (1987) *Chem. Phys. Lett.* **139,** 55-61.

9. Beals, J. M., Haas, E., Krausz, S. and Scheraga, H. A. (1991) *Biochemistry* **30,** 7680-7692.

10. Woodruff, W. H., Dyer, R. B., Callender, R. H., Paige, K. and Causgrove, T. (1994) *Biophys. J.* **66,** A397.

11. Phillips, C. M., Mizutani, Y. and Hochstrasser, R. M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7292-7296.

12. Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. and Nagamura, T. (1987) *FEBS Lett.* **221,** 115-118.

13. Jones, B. E., Beechem, J. M. and Matthews, C. R. (1995) *Biochemistry* **34,** 1867-1877.

14. Nölting, B., Golbik, R. and Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10668-10672.

15. Schmid, F. X., Mayr, L. M., Mücke, M. and Schönbrunner, E. R. (1993) *Adv. Protein Chem.* **44,** 25-66.

16. Gutin, A. M., Abkevich, V. I. and Shakhnovich, E. I. (1995) *Biochemistry* **34,** 3066-3076.

17. Ptitsyn, O. B. (1995) *Adv. Protein Chem.* **47,** 83-229.

18. Ptitsyn, O. B. (1995) *Curr. Opin. Struct. Biol.* **5,** 79-84.

19. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E. and Razgulyaev, O. I. (1990) *FEBS Lett.* **262,** 20-24.

20. Eaton, W. A., Thompson, P. A., Chan, C. K., Hagen, S. J. and Hofrichter, J. (1996) *Structure* **4,** 1133-1139.

21. Roder, H. and Shastry, M. C. R. (1999) *Curr. Opin. Struc. Biol.* **9,** 620-626.

22. Shastry, M. C. R., Luck, S. D. and Roder, H. (1998) *Biophys. J.* **74,** 2714-2721.

23. Pascher, T., Chesick, J. P., Winkler, J. R. and Gray, H. B. (1996) *Science* **271,** 1558-1560.

24. Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1997) *J. Am. Chem. Soc.* **119,** 9562-9563.

25. Telford, J. R., Tezcan, F. A., Gray, H. B. and Winkler, J. R. (1999) *Biochemistry* **38,** 1944-1949.

26. Telford, J. R., Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1998) *Acc. Chem. Res.* **31,** 755-763.

27. Bixler, J., Bakker, G. and McLendon, G. (1992) *J. Am. Chem. Soc.* **114,** 6938-6939.

28. Wittung-Stafshede, P., Lee, J. C., Winkler, J. R. and Gray, H. B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 6587-6590.

Chapter 2

Cytochrome b_{562} folding triggered by electron transfer: approaching the speed limit for formation of a four-helix bundle[†]

[†]Adapted from Wittung-Stafshede, P.; Lee, J. C.; Winkler, J. R.; Gray, H. B.; *Proc. Natl. Acad. Sci. USA* **1999, 96, 6587-6590**.

Introduction

Spontaneous refolding of proteins occurs on timescales that range from microseconds to hours (1–10). Cytochrome b_{562} (cyt b_{562}), a soluble 106-residue heme protein from *Escherichia coli* (11–14), folds relatively rapidly (9, 15). The protein is comprised of four antiparallel α-helices that wrap into a left-handed bundle; the heme is inserted among the helices, axially ligated by Met-7 and His-102 (Figure 2.1) (16). The heme reduction potential in the folded state [189 mV vs. NHE (pH 7) (17)] is ≈330 mV higher than in the unfolded form [–150 mV (18, 19)], indicating that the driving force for folding is greater for the reduced than for the oxidized protein (9, 15, 20, 21). Guanidine hydrochloride (GuHCl) reversibly unfolds the oxidized protein with a midpoint at 1.8 M (pH 7). Unfolding the reduced protein requires a GuHCl concentration > 5 M and, owing to heme dissociation, is irreversible. Under suitable denaturing conditions (2 M ≤ [GuHCl] ≤ 3 M), electron transfer to unfolded Fe^{III}–cyt b_{562} initiates folding of the reduced protein. Earlier, it was reported that the native low-spin Fe^{II} protein forms ≈1 ms after reduction of unfolded Fe^{III}–cyt b_{562} (15).

We have investigated the dependence of the Fe^{II}–cyt b_{562} folding kinetics on GuHCl concentration. Extrapolation of the observed rate constants to zero denaturant concentration suggests that the folding of this four-helix bundle in water approaches the theoretical speed limit for intrachain diffusion of the polypeptide.

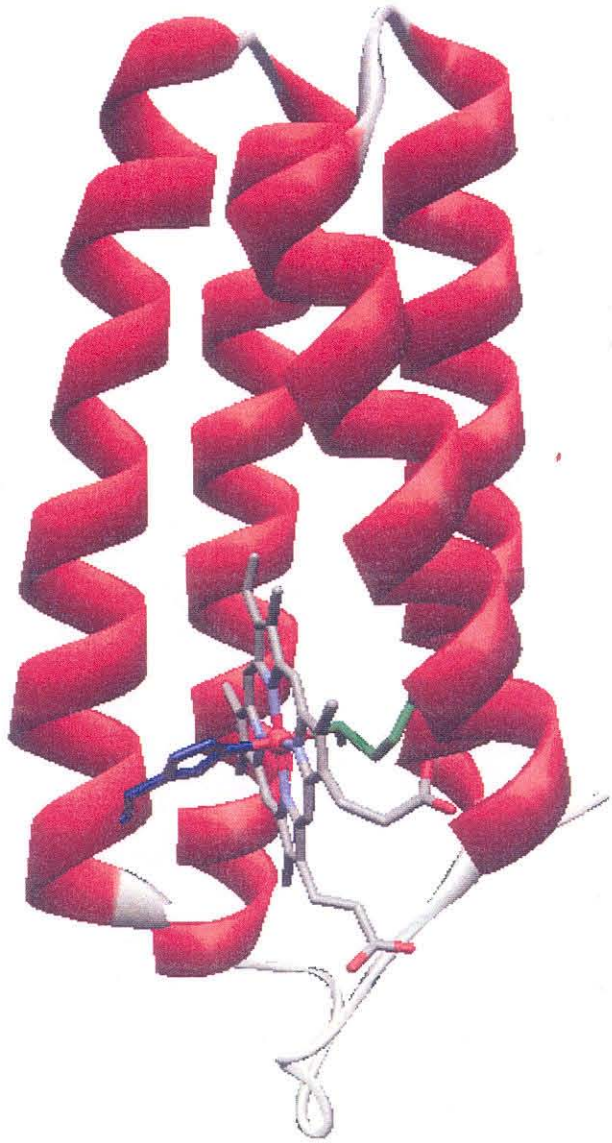


Figure 2.1. Ribbon structure of cytochrome b_{562} from *E. coli* (PDB 256B). The heme axial ligands, Met-7 (green) and His-102 (blue), are highlighted.

Materials and Methods

Guanidine hydrochloride (Sigma, ultrapure grade), tris(2, 2'-bipyridine)-ruthenium(II) chloride (Strem, Newburyport, MA), and 1,4-dihydronicotinamide adenine dinucleotide (NADH, Sigma) were used as received. *Para*-methoxydimethyl-aniline (pMDMA) was synthesized according to a published procedure (22). Plasmid (pNS207), containing the cytochrome b562 gene (provided by S. G. Sligar, University of Illinois), was expressed in *E. coli* (strain BL21).

Cyt b_{562} was isolated and purified according to a procedure given by Farrow (23). The purity of proteins used in all experiments was assessed from the 418:280 nm absorbance (A_{418}/A_{280} = 5.6–6.0). Protein concentrations were determined by using the extinction coefficients: ε_{418} = 0.117 mM^{-1} cm^{-1} [Fe^{III}] and ε_{426} = 0.180 mM^{-1} cm^{-1} [Fe^{II}] (12, 24). Unfolding curves for Fe^{III}– and Fe^{II}–cyt b_{562} were generated from absorption and far-UV circular dichroism spectra according to standard procedures (25). In experiments with the Fe^{II} protein, a slight excess of sodium dithionite was added to keep the unfolded protein in the reduced state. GuHCl stock solutions were adjusted to pH 7.0–7.1 by using 1.0 M NaOH. The concentration of GuHCl was determined through measurements of the solution refractive index (Milton Roy Abbe-3L Refractometer) (25).

Steady-state absorption spectra were measured on a Hewlett–Packard 8452 diode array spectrophotometer. Circular dichroism spectra were obtained by using an Aviv 62ADS spectropolarimeter (Aviv Associates, Lakewood, NJ). Transient absorption kinetics measurements were made as described (26). The

excitation source was a XeCl excimer-pumped dye laser (480 nm, 25 ns, 1–4 mJ) in experiments with $Ru(bpy)_3^{2+}$ and the third harmonic (355 nm) of a Q-switched Nd:YAG laser in experiments using NADH as the photoreductant. Samples for folding kinetics measurements were held in 1 mm sealed cuvettes (0.8–1.5 mL) and contained the following: cyt b_{562} (15–100 μM); either $Ru(bpy)_3^{2+}$ (1.0–1.5 equivalent) and pMDMA (≈10 mM) or NADH (2.0–2.5 equivalent); and GuHCl (2–3 M, pH 6.6). Samples were deoxygenated by repeated evacuation/Ar-fill cycles on a Schlenk line. At each GuHCl concentration, the reported rate constant is an average of several kinetics traces.

The yield of folded reduced protein (Φ_f), relative to unfolded reduced protein generated by laser excitation, was calculated from initial and final absorption changes at each GuHCl concentration. The extinction coefficient for the unfolded reduced protein was determined by using a reductive flash quench procedure. Laser-excited $Ru(bpy)_3^{2+}$ was quenched with pMDMA to generate $Ru(bpy)_3^{+}$ and $pMDMA^{\bullet+}$. The concentrations of the charge separated products were determined from their molar difference spectra. The photochemically generated $Ru(bpy)_3^{+}$ reduces unfolded Fe^{III}–cyt b_{562} in 10 μs. On complete reduction of the heme, the transient spectrum contains contributions from $pMDMA^{\bullet+}$ and unfolded Fe^{II}–cyt b_{562} at known concentrations. Extinction coefficients for unfolded Fe^{II}–cyt b_{562} then were determined from the transient kinetics at various wavelengths by subtracting the contribution from $pMDMA^{\bullet+}$. The observed rate constant (k_{obsd}) was interpreted as the sum of a folding rate constant (k_f) and a dissociation rate constant (k_{diss}):

$$k_{obsd} = k_f + k_{diss}$$

$$k_f = k_{obsd}\,\Phi_f$$

Folding rate constants were extracted from observed rates by using the spectroscopically determined folding yields.

RESULTS

Equilibrium Unfolding. The absorption spectra of the oxidized and reduced states of folded cyt b_{562} show Soret-region features that are characteristic of low-spin d^5 and d^6 hemes (19). On GuHCl denaturation of Fe^{III}–cyt b_{562}, the Soret absorption blue-shifts, broadens and loses intensity at the maximum; these changes in the spectrum reflect the formation of a high-spin d^5 heme (Figure 2.2). GuHCl titrations probed by changes in absorbance or circular dichroism spectra indicate that the unfolding midpoint for Fe^{III}–cyt b_{562} (1.8 M) is much lower than that for reduced protein (5.7 M) (Figure 2.3). The unfolded oxidized protein spontaneously refolds on dilution of GuHCl, confirming that the heme stays bound to the polypeptide for protein concentrations > 10 μM (27). Unfolded Fe^{II}–cyt b_{562}, however, does not readily refold on dilution of the denaturant; heme dissociation is likely responsible for this behavior. Between 2 and 3 M GuHCl (pH 7), a significant amount of folded Fe^{II}–cyt b_{562} is observed on addition of sodium dithionite to solutions of unfolded Fe^{III}–cyt b_{562} (2.33 M GuHCl, 50% yield; 3.02 M, 30%). At higher GuHCl concentrations, heme dissociation from the peptide appears to be faster than protein folding.

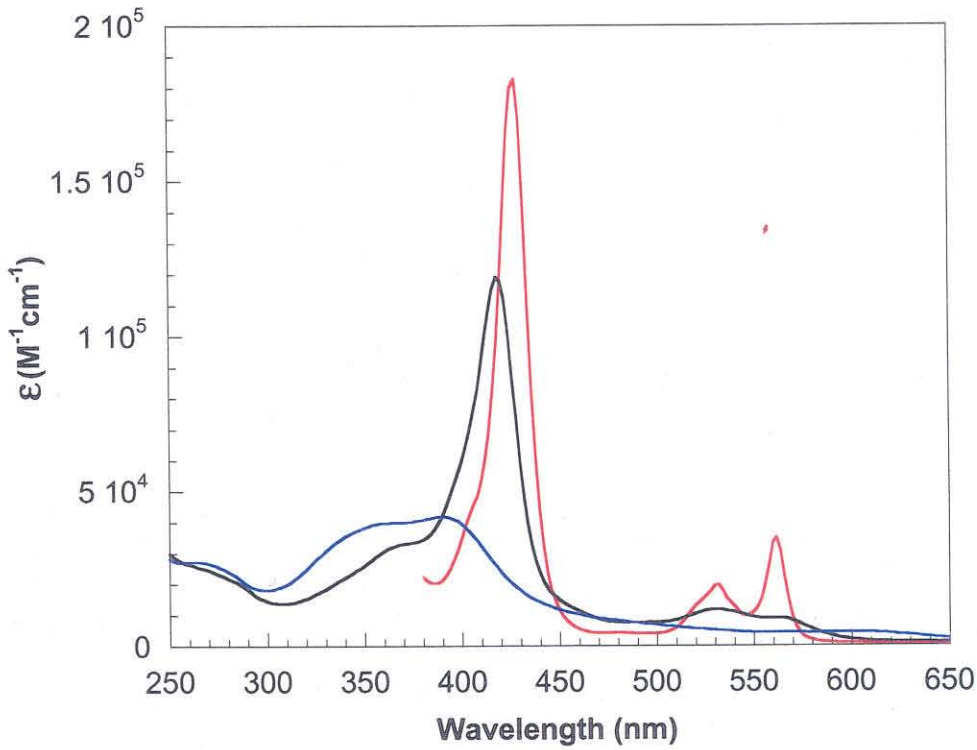


Figure 2.2. Absorption spectra of folded Fe^{III}-cyt b_{562} (black), Fe^{II}-cyt b_{562} (red), and unfolded Fe^{III}-cyt b_{562} (blue).

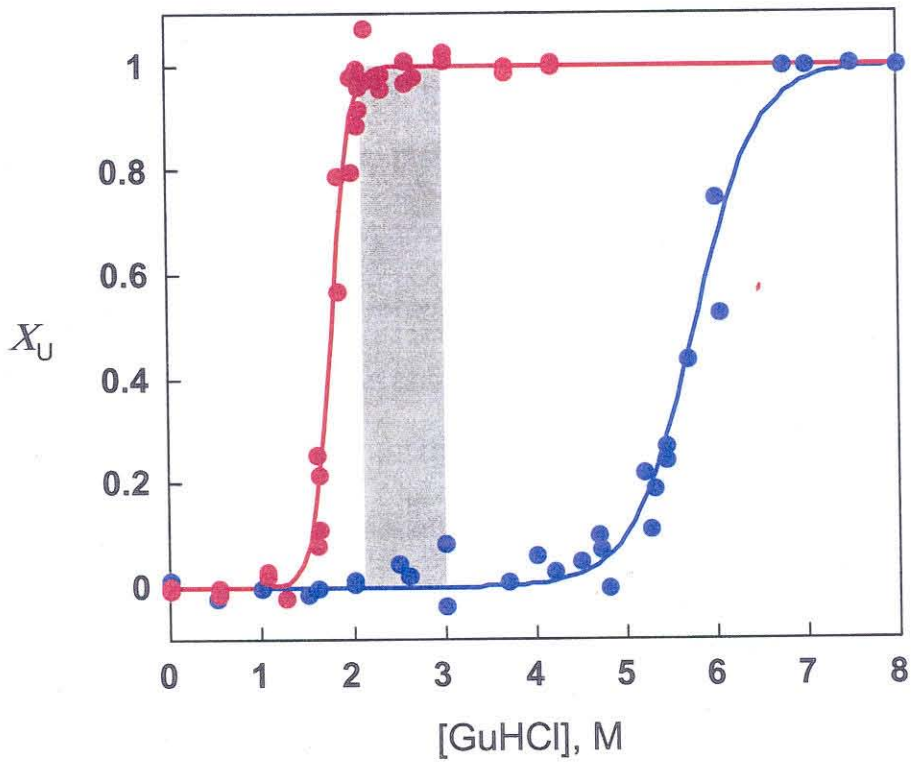


Figure 2.3. Denaturation curves for Fe^{III}- (red) and Fe^{II}-cyt b_{562} (blue) generated from Soret absorption and CD spectral data. The shaded region indicates the GuHCl concentrations where ET triggered folding experiments were performed.

Fe^{II}–cyt b_{562} Folding. Measurements of Fe^{II}–cyt b_{562} kinetics were, therefore, confined to the 2–3 M GuHCl concentration range. The folding of Fe^{II}–cyt b_{562} was initiated by laser excitation (355 nm) of NADH to produce two powerful reductants, a solvated electron and the NAD radical (15, 28, 29). The initial transient spectrum (≈100 μs) reflects formation of reduced unfolded protein. The transient spectrum recorded 10 ms after excitation (2.5 M GuHCl) agrees well with steady-state absorption differences between unfolded Fe^{III}–cyt b_{562} and folded Fe^{II}–cyt b_{562} (15) indicating that folding has occurred on this timescale. Kinetics were monitored by Soret absorption changes (400–430 nm) that indicate the formation of a low-spin ferroheme. The folding reactions were slower at higher GuHCl concentrations; the observed first-order rate constants (k_{obsd}) range from 3.3(7) x 10^3 to 7.4(9) x 10^3 s^{-1} (2–3 M GuHCl) (Figure 2.4). Specific rate constants for formation of low-spin folded cyt b_{562}, extracted from the observed rate constants, vary from 1.5(2) x 10^3 (2.2 M GuHCl) to 3.0(9) x 10^2 s^{-1} (3.0 M GuHCl) (Figure 2.5). (The rate constants for the competing heme-dissociation process range from 3 to 7 x10^3 s^{-1} between 2 and 3 M GuHCl.) Although a relatively narrow range of GuHCl concentrations was explored, there appears to be a linear relationship between the logarithm of the folding rate and the GuHCl concentration. Extrapolation to 0 M GuHCl suggests a time constant of ≈5 μs for folding Fe^{II}–cyt b_{562} in water.

Discussion

A recent molecular-dynamics simulation carried out on an unfolded 36-

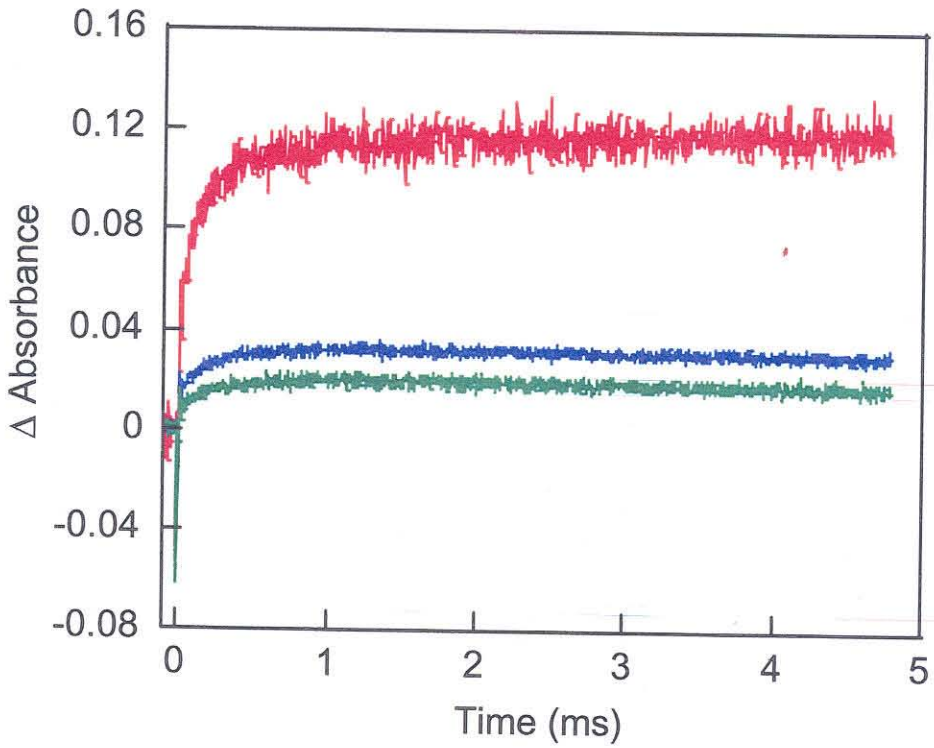


Figure 2.4. Folding kinetics of Fe^{II}-cyt b_{562} observed at 425 nm with varying concentrations of GuHCl (2.43 M (red); 2.60 M (blue); 2.83 M GuHCl (green)).

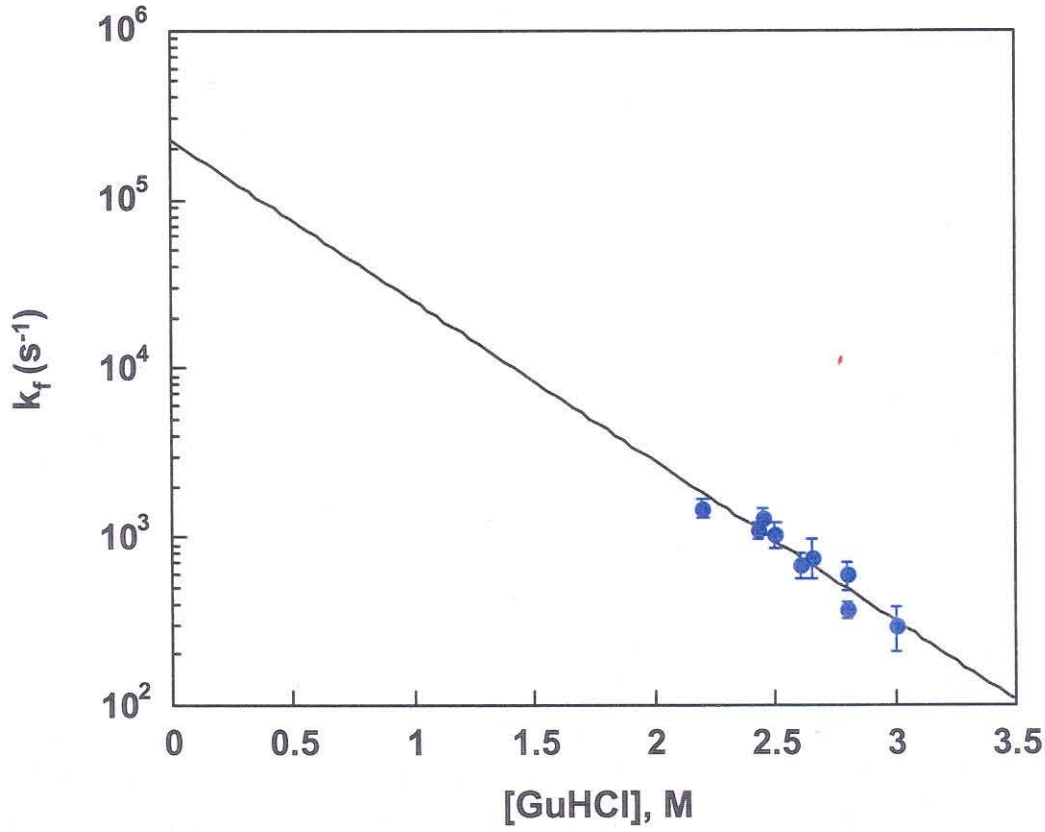


Figure 2.5. Dependence of Fe^{II}-cyt b_{562} folding rates on GuHCl concentration (pH 6.6). Observed rate constants, folding rate constants ([GuHCl], yield of folded protein): 3.3(7) x 10^3, 1.5(2) x 10^3 s^{-1} (2.20 M, 45%); 3.4(1) x 10^3, 1.1(1) x 10^3 s^{-1} (2.43 M, 34%); 6.0(7) x 10^3, 1.3(2) x 10^3 s^{-1}(2.45 M, 22%); 3.2(6) x 10^3, 1.1(2) x 10^3 s^{-1} (2.50 M, 33%); 5.4(2) x 10^3, 6.8(2) x 10^2 s^{-1} (2.60 M, 13%); 9.5(3) x 10^3, 7.6(2) x 10^2 s^{-1} (2.65 M, 8%); 5.5(6) x 10^3, 6.0(1) x 10^2 s^{-1} (2.80 M, 6%); 6.3(2) x 10^3, 3.7(4) x 10^2 s^{-1} (2.80 M, 9%); 7.4(9) x 10^3, 3.0(9) x 10^2 s^{-1} (3.00 M, 4%).

residue protein suggests that partly folded structures can occur within tens of nanoseconds (30). Experimental studies on apomyoglobin (134 residues) indicate that some secondary structure forms in the first 10 μs of folding (31, 32). Polymer diffusion models have been used to estimate that the minimum time constant for folding a small, single-domain protein to its native structure is on the order of 1 μs (8, 33). More recent estimates suggest a lower limit of 0.2 ms for protein folding (34). Our extrapolated folding time for Fe^{II}-cyt b_{562} approaches this limit.

Another four-helix bundle, the acyl CoA binding protein (ACBP), folds more slowly than Fe^{II}–cyt b_{562} [700 s^{-1}, 20°C (35, 36)]. ACBP has a so-called up-down-down-up four-helix bundle structure held together by hydrophobic junctions, where the helices interact pairwise or in groups of three (37). By contrast, the cyt b_{562} bundle is an up-down-up-down arrangement, and the four helices are almost perfectly parallel to each other (16). ACBP is less stable than Fe^{II}–cyt b_{562} [ΔG_f = –29 (35, 36) and ≈ –43 kJ mol^{-1} (15), respectively], a property that may affect the folding kinetics, because, in general, the tertiary folds of these two proteins are similar. Although reduced horse heart cytochrome *c* is much more stable [ΔG_f = –74 kJ mol^{-1} (38)] than Fe^{II}–cyt b_{562}, it folds much more slowly [estimated folding time of 100 ms in aqueous solution (20, 28, 38)]. Part of the difference is undoubtedly attributable to ligand substitution kinetics that retard the folding of cytochrome *c* (28). But even at reduced pH, where the heme in unfolded cytochrome *c* is not misligated, protein folding is still slower than for cyt b_{562} (5).

Why does folding of Fe^{II}–cyt b_{562} occur near the maximum speed limit when the folding of other helical proteins is much slower? Theoretical analysis indicates that symmetrical structures will have relatively smooth energy landscapes for folding (1, 39), and the up-down-up-down four-helix-bundle structure of cytochrome b_{562} is one of the most symmetrical of protein architectures. It is interesting to note that an inverse correlation between folding rates and contact order (the average relative sequence separation of native contacts) has been observed for a set of single-domain proteins that exhibit two-state folding (40). Although two-state folding has not yet been confirmed in cyt b_{562}, the contact order of this protein is extremely low,~7.5% (K. W. Plaxco, personal communication), which is consistent with its very local, all helical structure (16). This low contact order corresponds to a predicted folding time of $\approx$ 10 μs, which is in good agreement with our extrapolated value of 5 μs. The longer folding times of the helical proteins cytochrome *c* and ACBP (20, 28, 35, 36, 38) are consistent with the predictions of the contact-order model (contact orders: cyt *c*, 11.2%; ACBP, 14%) (40).

The fast folding observed for cyt b_{562} may be a consequence of the competing heme-dissociation step. Unfolded cyt b_{562} is expected to be an ensemble of randomly configured, heme-bound polypeptides. The kinetics for conversion of this ensemble into a homogeneous population of folded proteins could be highly heterogeneous. Some configurations in the unfolded ensemble could fold quite rapidly whereas others might require much longer times (41). This division of the unfolded ensemble into fast and slow folding populations can

have important implications when a second process competes with folding. In Fe^{II}–cyt b_{562}, it is possible that only fast folding is observed because the heme dissociates from the slow-folding configurations. Because the free heme has a relatively weak spectroscopic signature, its presence would be difficult to detect amid a population of folded reduced protein. If heme dissociation does select against slow folding configurations in Fe^{II}–cyt b_{562}, then we expect that four-helix bundle *c*-type heme proteins [e.g., cytochrome *c*' (11, 42)] would require much longer times to form fully folded ensembles.

References

1. Onuchic, J. N., Lutheyschulten, Z. and Wolynes, P. G. (1997) *Annu. Rev. Phys. Chem.* **48,** 545-600.
2. Shakhnovich, E. I. (1997) *Curr. Opin. Struct. Biol.* **7,** 29-40.
3. Dobson, C. M., Sali, A. and Karplus, M. (1998) *Angew. Chem., Int. Ed. Eng.* **37,** 868-893.
4. Dill, K. A. and Chan, H. S. (1997) *Nature Struct. Biol.* **4,** 10-19.
5. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X. and Bai, Y. (1998) *Acc. Chem. Res.* **31,** 737-744.
6. Shastry, M. C. R., Sauder, J. M. and Roder, H. (1998) *Acc. Chem. Res.* **31,** 717-725.
7. Yeh, S.-R., Han, S. and Rousseau, D. L. (1998) *Acc. Chem. Res.* **31,** 727-736.

8. Eaton, W. A., Muñoz, V., Thompson, P. A., Henry, E. R. and Hofrichter, J. (1998) *Acc. Chem. Res.* **31,** 745-753.

9. Telford, J. R., Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1998) *Acc. Chem. Res.* **31,** 755-763.

10. Dodge, R. W. and Scheraga, H. A. (1996) *Biochemistry* **35,** 1548-1559.

11. Moore, G. R. and Pettigrew, G. W. (1990) *Cytochromes c: Evolutionary, Structural, and Physicochemical Aspects* (Springer-Verlag, New York).

12. Itagaki, E. and Hager, L. P. (1966) *J. Biol. Chem.* **241,** 3687-3695.

13. Itagaki, E., Palmer, G. and Hager, L. P. (1967) *J. Biol. Chem.* **242,** 2272-2277.

14. Nikkila, H., Gennis, R. B. and Sligar, S. G. (1991) *Eur. J. Biochem.* **202,** 309-313.

15. Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1997) *J. Am. Chem. Soc.* **119,** 9562-9563.

16. Hamada, K., Bethge, P. H. and Mathews, F. S. (1995) *J. Mol. Biol.* **247,** 947-962.

17. Barker, P. D., Butler, J. L., de Oliveira, P., Hill, H. A. O. and Hunt, N. I. (1996) *Inorg. Chim. Acta* **252,** 71-77.

18. Bixler, J., Bakker, G. and McLendon, G. (1992) *J. Am. Chem. Soc.* **114,** 6938-6939.

19. Tezcan, F. A., Winkler, J. R. and Gray, H. B. (1998) *J. Am. Chem. Soc.* **120,** 13383-13388.

20. Pascher, T., Chesick, J. P., Winkler, J. R. and Gray, H. B. (1996) *Science* **271,** 1558-1560.

21. Winkler, J. R., Wittung-Stafshede, P., Leckner, J., Malmström, B. G. and Gray, H. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 4246-4249.

22. Sekiya, M. and Leonard, N. J. (1968) *J. Org. Chem.* **33,** 318-312.

23. Farrow, N. A. (1999) in *Chemistry* (California Institute of Technology, Pasadena).

24. Moore, G. R., Williams, R. J. P., Peterson, J., Thomson, A. J. and Mathews, F. S. (1985) *Biochim. Biophys. Acta* **829,** 83-96.

25. Pace, N. C., Shirley, B. A. and Thomson, J. A. (1990) in *Protein Structure: A Practical Approach*, ed. Creighton, T. F. (IRL Press, Oxford), pp. 311-330.

26. Stowell, M. H. B., Larsen, R. W., Winkler, J. R., Rees, D. C. and Chan, S. I. (1993) *J. Phys. Chem.* **97,** 3054-3057.

27. Robinson, C. R., Liu, Y., Thomson, J. A., Sturtevant, J. M. and Sligar, S. G. (1997) *Biochemistry* **36,** 16141-16146.

28. Telford, J. R., Tezcan, F. A., Gray, H. B. and Winkler, J. R. (1999) *Biochemistry* **38,** 1944-1949.

29. Orii, Y. (1993) *Biochemistry* **32,** 11910-11914.

30. Duan, Y., Wang, L. and Kollman, P. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9897-9902.

31. Gruebele, M., Sabelko, J., Ballew, R. and Ervin, J. (1998) *Acc. Chem. Res.* **31,** 699-707.

32. Dyer, R. B., Gai, F., Woodruff, W. H., Gilmanshin, R. and Callender, R. H. (1998) *Acc. Chem. Res.* **31,** 709-716.

33. Hagen, S. J., Hofrichter, J., Szabo, A. and Eaton, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11615-11617.

34. Thirumalai, D. (1999) *J. Phys. Chem. B* **103,** 608-610.

35. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. and Poulsen, F. M. (1995) *Biochemistry* **34,** 7217-7224.

36. Kragelund, B. B., Højrup, P., Jensen, M. S., Schjerling, C. K., Juul, E., Knudsen, J. and Poulsen, F. M. (1996) *J. Mol. Biol.* **256,** 187-200.

37. Andersen, K. V. and Poulsen, F. (1992) *J. Mol. Biol.* **226,** 1131-1141.

38. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. and Gray, H. B. (1996) *Chem. Biol.* **3,** 491-497.

39. Wolynes, P. G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14249-14255.

40. Plaxco, K. W., Simons, K. T. and Baker, D. (1998) *J. Mol. Biol.* **277,** 985-994.

41. Guo, Z. and Thirumalai, D. (1996) *J. Mol. Biol.* **263,** 323-343.

42. Bertini, I., Dikiy, A., Luchinat, C., Macinai, R. and Viezzoli, M. S. (1998) *Inorg. Chem.* **37,** 4814-4821.

Chapter 3

Cytochrome *c*' folding triggered by electron transfer: fast and slow formation of four-helix bundles[†]

[†]Adapted from Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 7760-7764.

Introduction

Energy-landscape descriptions of protein folding emphasize the roles of free energy and energy-surface ruggedness as key determinants of folding dynamics (1–3). The many correlations of folding times with free energy support this point of view, but it is difficult to reconcile the model with the seemingly simple kinetics that often are observed. Analyses of experimental data point to a relationship between the folding kinetics and the topology of the native protein structure (4–8). Proteins that, on average, have a large number of long-range contacts in the folded state tend to fold more slowly than those with a preponderance of short-range contacts. The questions remain: is it the depth and roughness of the funnel, the topology of the folded state, or some combination of the two that determines the folding mechanism?

Short-range contacts outnumber long-range interactions in helical bundles; if topology is the key, then these proteins should fold rapidly (4, 5). Highly helical ferrocytochrome b_{562} (Fe^{II}–cyt b_{562}) (9, 10), acyl-CoA binding protein (ACBP) (11–14), the E colicin binding immunity proteins Im7 and Im9 (15, 16), and the N-terminal domain of phage λ repressor (17) all fold on the millisecond timescale. These observations appear to support the notion that folded-state structural topology is of central importance in folding dynamics and that helical bundle structures are inherently fast folding.

We have shown previously that Fe^{II}–cyt b_{562} folding can be triggered by electron transfer (ET), a method that takes advantage of the greater stabilities of reduced heme proteins in native states (9, 10, 18, 19). Although folded Fe^{II}–cyt

b_{562} was observed within milliseconds after reduction of the unfolded oxidized protein, no more than half of the reduced protein successfully developed native structure (10). Rapid heme dissociation from the polypeptide (k_{diss} ~2–7 x 10^{3} s^{-1}) limited the yield of the folding reaction. The heme-loss step selects fast-folding conformations from the unfolded ensemble; if there are slow-folding components, they cannot be detected. Under these circumstances, the observed kinetics reflect heme-dissociation dynamics rather than folding.

Cytochrome *c*' (cyt *c*') from the photosynthetic bacterium *Rhodopseudomonas palustris* is a monomeric, soluble, 125-residue, four-helix-bundle heme protein. Importantly, the porphyrin is bound to the polypeptide with two thioether links near the C terminus (Cys-113 and Cys-116) (20–22). Although sharing just 19% sequence identity and 40% similarity (23), cyt *c*' and cyt b_{562} have quite similar folds (1.6-Å rms deviation in α-carbon position) (Figure 3.1) (24, 25). Cyt b_{562} has a six-coordinate, low-spin heme with Met-7 and His-102 axial ligands [$(^{H102}N)\{^{Por}N_4Fe^{III}\}(^{M7}S)^{+}$] and a reduction potential of 180 mV vs. NHE (24, 26). Cyt *c*' has a high-spin, five-coordinate heme, axially ligated by His-117 [$(^{H117}N)\{^{Por}N_4Fe^{III}\}^{+}$] and a reduction potential of 100 mV (27–32). In cyt *c*', the side-chain of Leu-12 fills the space occupied by a sixth ligand in cyt b_{562} (Figure 3.1 *Inset*) (25); movement of this bulky group is necessary for ligand binding.

The $Fe^{III/II}$ reduction potential is high enough to permit ET triggering of Fe^{II}–cyt *c*' folding in the 2.0–2.9 M guanidine hydrochloride (GuHCl) concentration range. Fe^{II}–cyt *c*' folding is highly heterogeneous, spanning a time

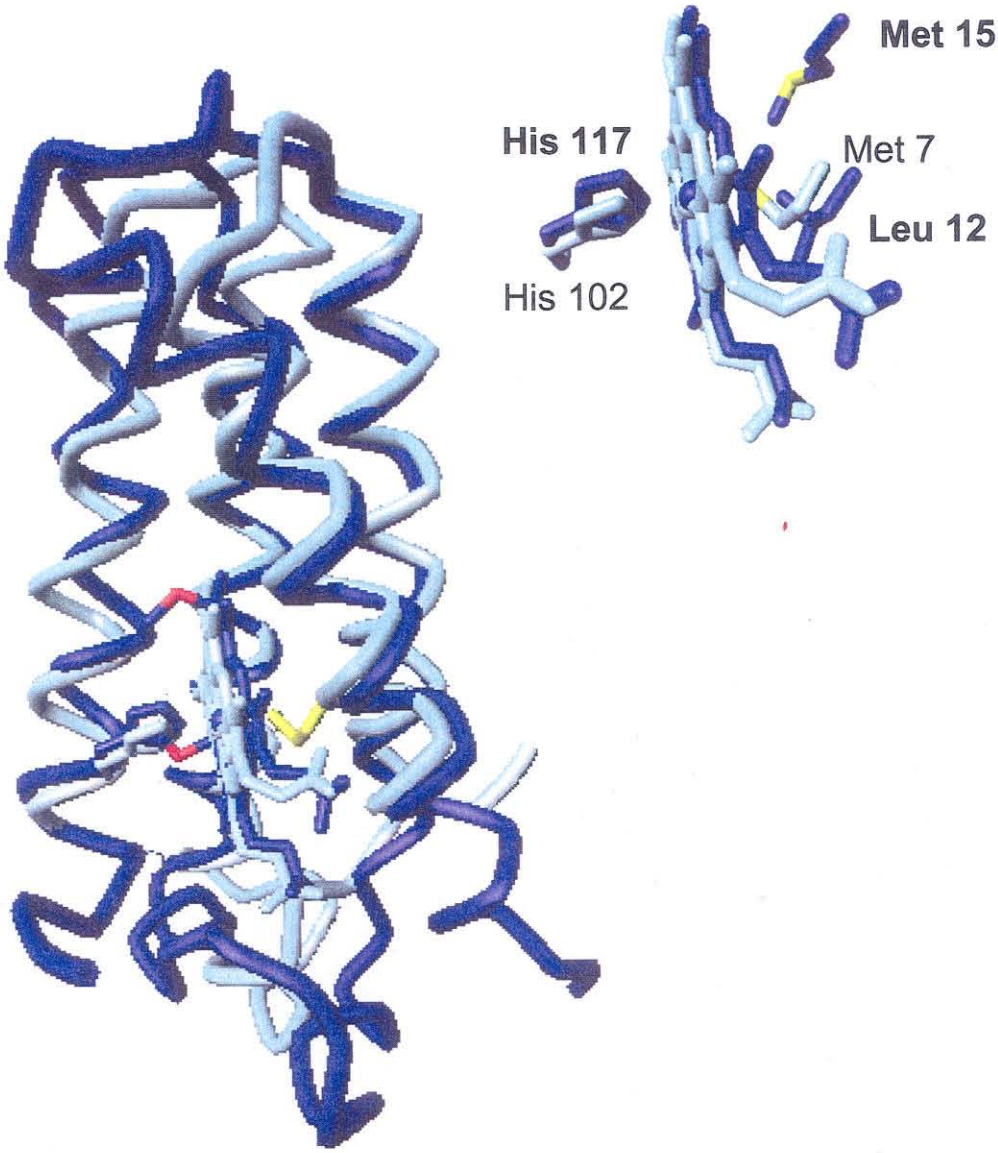


Figure 3.1. Comparison of heme environments. In cyt b_{562} (cyan), the heme iron is axially ligated to His-102 and Met-7, whereas in cyt *c'* (blue), the heme has only one axial ligand, His-117 with the side chain of Leu-12 at the other axial site. The nearest methionine residue (Met-15) in cyt *c'* is also shown. Backbone atoms of four α-helices of cyt *c'* (PDB 1A7V) are superimposed on the corresponding atoms in cyt b_{562} (PDB 256B) with a calculated rms deviation of 1.6 Å .

range from milliseconds to several seconds. Importantly, the folding kinetics of Fe^{III}–cyt *c*' and $Fe^{II}(CO)$–cyt *c*' are less complex, indicating that nonnative ligand traps may be responsible for the complexity of Fe^{II}–cyt *c*' folding.

Materials and Methods

GuHCl (Sigma, ultrapure grade), tris(2,2'-bipyridine)ruthenium(II) chloride ($[Ru(bpy)_3]Cl_2$, Strem), and NADH (Sigma) were used as received.

Spectroscopic measurements were made as follows: absorption, Hewlett-Packard 8452 diode array spectrophotometer; CD, Aviv 62ADS spectropolarimeter; and luminescence, Hitachi F-4500 spectrofluorimeter (λ_{ex} = 290 nm; λ_{obsd} = 300–500 nm). Mass spectral analyses were done at the Protein/Peptide Microanalytical Laboratory in the Beckman Institute at Caltech.

R. palustris strain 37 (ATCC 17007) cyt *c*' was isolated and purified by using minor modifications of published procedures (31, 33). Cyt *c*' was chromatographed either on a CM Sepharose column equilibrated with 5 mM sodium phosphate and 10 mM NaCl (pH 7.0) in the presence of ferricyanide and eluted with the charging buffer, or on a column of hydroxyapatite (type I, Bio-Rad) equilibrated with 3 mM sodium phosphate and 15 mM NaCl (pH 7.0) in the presence of mercaptoethanol and eluted with a linear gradient from 10 mM to 100 mM sodium phosphate buffer (15 mM NaCl, pH 7.0). Proteins used in all experiments had purity indices ($A_{280\ nm}/A_{398\ nm}$, oxidized) between 0.22 and 0.25 (literature value, 0.21) (33). Although only a single fraction was isolated chromatographically, mass spectra revealed two components in the protein

sample. The mass of one component matched that expected on the basis of the amino acid sequence (13,748 daltons); the other had a mass 17 atomic mass units lower than predicted. From further protein sequencing analysis, it was revealed that the N-terminus undergoes a post-translational modification where Gln1 cyclizes and forms pyroglutamic acid (losing 17 amu) (see Appendix A). Protein concentrations were obtained from UV-visible absorption measurements (Fe^{III}–cyt *c*', $\varepsilon_{398\ nm}$ = 85 $mM^{-1}cm^{-1}$; Fe^{II}–cyt *c*', $\varepsilon_{426\ nm}$ = 99 $mM^{-1}cm^{-1}$) (34). GuHCl concentrations were extracted from refractive index determinations (Abbe-3L refractometer, Milton Roy, Rochester, NY) (35).

Denaturation curves for Fe^{II}–cyt *c*' were determined in the presence of excess sodium dithionite. Transient absorption spectra and kinetics were measured by using an apparatus described previously (36). The excitation source was the third harmonic of a Q-switched Nd:YAG laser (355 nm) for electron photoinjection from NADH, and a Nd:YAG pumped optical parametric oscillator (480 nm) for injection from $Ru(bpy)_3^{2+}$ and for CO photolysis experiments. A 16-bit, 20 MS/s digital oscilloscope (CompuScope 1602, Gage Applied Science, Inc.) was used to record the complete kinetics traces from 10^{-3} to 10^{1} s. Samples for kinetics measurements were deoxygenated by repeated evacuation/argon-fill cycles on a Schlenk line. Typical samples contained cyt *c*' (100–200 μM), either $Ru(bpy)_3^{2+}$ (1.0–1.5 equivalents) or NADH (2.0–2.5 equivalents), and GuHCl (1.8–3.5 M) in 100 mM sodium phosphate buffer (pH 7). The pH and index of refraction were measured after laser experiments. Reduced protein was stirred under 1 atm of CO for at least 30 min after deoxygenation to

prepare Fe^{II}(CO)–cyt *c*'.

Measured kinetics traces were logarithmically compressed (100 points per time decade), normalized [$I(t) - I(t_\infty)/I(t_0) - I(t_\infty)$], and fit by using a non-negative linear least-squares minimization algorithm to project the folding kinetics onto a basis of exponential decay functions with rate constants in the range 10^5 to 10^{-2} s^{-1} (10 rate constants per decade, spaced logarithmically).

The folding kinetics of Fe^{III}–cyt *c*' were measured by using a BioLogic SFM-4S stopped-flow mixer, coupled via optical fiber to a monochromator fitted with a five-stage photomultiplier tube. Probe light was provided by a 200-W Hg/Xe arc lamp. The PMT output was recorded with the 16-bit digital oscilloscope. Kinetics traces were measured at 372 nm (formation of folded protein, $\Delta\varepsilon = 28$ $mM^{-1}cm^{-1}$) and 400 nm (disappearance of the unfolded protein, $\Delta\varepsilon = -44$ mM^{-1} cm^{-1}) for a time period up to 25 s, at a sampling rate of 2 or 5 kS/s. In all refolding experiments, unfolded Fe^{III}–cyt *c*' (3.3 M GuHCl) was either diluted 6- or 12-fold with a combination of Hepes buffer ([Hepes] = 0.1 M, pH 7) and Hepes/GuHCl ([Hepes] = 0.1 M, [GuHCl] = 3.3 M, pH 7), giving final protein concentrations of 1.5 μM and varying GuHCl concentrations (0.6–1.80 M). Kinetics traces were fit to single exponential functions.

Results

Equilibrium Unfolding. GuHCl titration curves generated from far-UV CD, Trp-72 fluorescence, and Soret absorbance spectra show that reduced cyt *c*' is more stable toward unfolding ($[GuHCl]_{½}$= 2.9(1) M) than oxidized cyt *c*' ($[GuHCl]_{½}$=

1.9(1) M) (Figure 3.2). The stability of folded cyt *c'*, judged by the folding free-energy change extrapolated to zero denaturant ($\Delta G_f°$), is measurably lower than that of cyt b_{562} [Fe(III): $\Delta G_f°$(cyt *c'*) = –33(3) kJ mol^{-1}, $\Delta G_f°$(cyt b_{562}) = –42(3) kJ mol^{-1}; Fe(II): $\Delta G_f°$(cyt *c'*) = –40(4) kJ mol^{-1}, $\Delta G_f°$(cyt b_{562}) ~–78 kJ mol^{-1}]. The difference in folding free energies between Fe^{III}– and Fe^{II}–cyt *c'* ($\Delta\Delta G_f°$ =7 kJ mol^{-1} ≈70 meV) is less than half that predicted on the basis of the reduction potential of the folded protein (100 mV vs. NHE) and the potential expected for a heme in aqueous solution (–150 mV) (37). The titrations suggest that the potential of the unfolded protein is ≈ 0 V, possibly because a second axial ligand binds to the heme when the protein unfolds. Although the difference in folding free energies is smaller than anticipated, it is large enough to permit ET triggering of Fe^{II}–cyt *c'* folding between 2 and 2.9 M GuHCl.

Ligand Binding to Unfolded Cyt c'. Addition of GuHCl to solutions of Fe^{III}–cyt *c'* narrows the Soret absorption band (λ_{max} = 400 nm, ε = 133 mM^{-1}•cm^{-1}) (Figure 3.3). The absorption profile is analogous to that of the M80A mutant of cytochrome *c* (λ_{max} = 400 nm, ε = 164 mM^{-1}•cm^{-1},pH = 4.5) (38) and is consistent with the formation of a six-coordinate, high-spin heme $(^{H117}N)\{^{Por}N_4Fe^{III}\}(OH_2)^+$. The absorption spectrum in alkaline solution (pH > 10, [GuHCl] ≥ 3.5 M) is typical of a low-spin Fe^{III} heme (λ_{max} = 406 nm, ε = 100 mM^{-1} •cm^{-1}) (Figure 3.4) (38); in this case, it is possible that a lysine residue binds to form $(^{H117}N)\{^{Por}N_4Fe^{III}\}(^{K}N)^+$. Unlike cytochrome c, cyt *c'* has no His residues

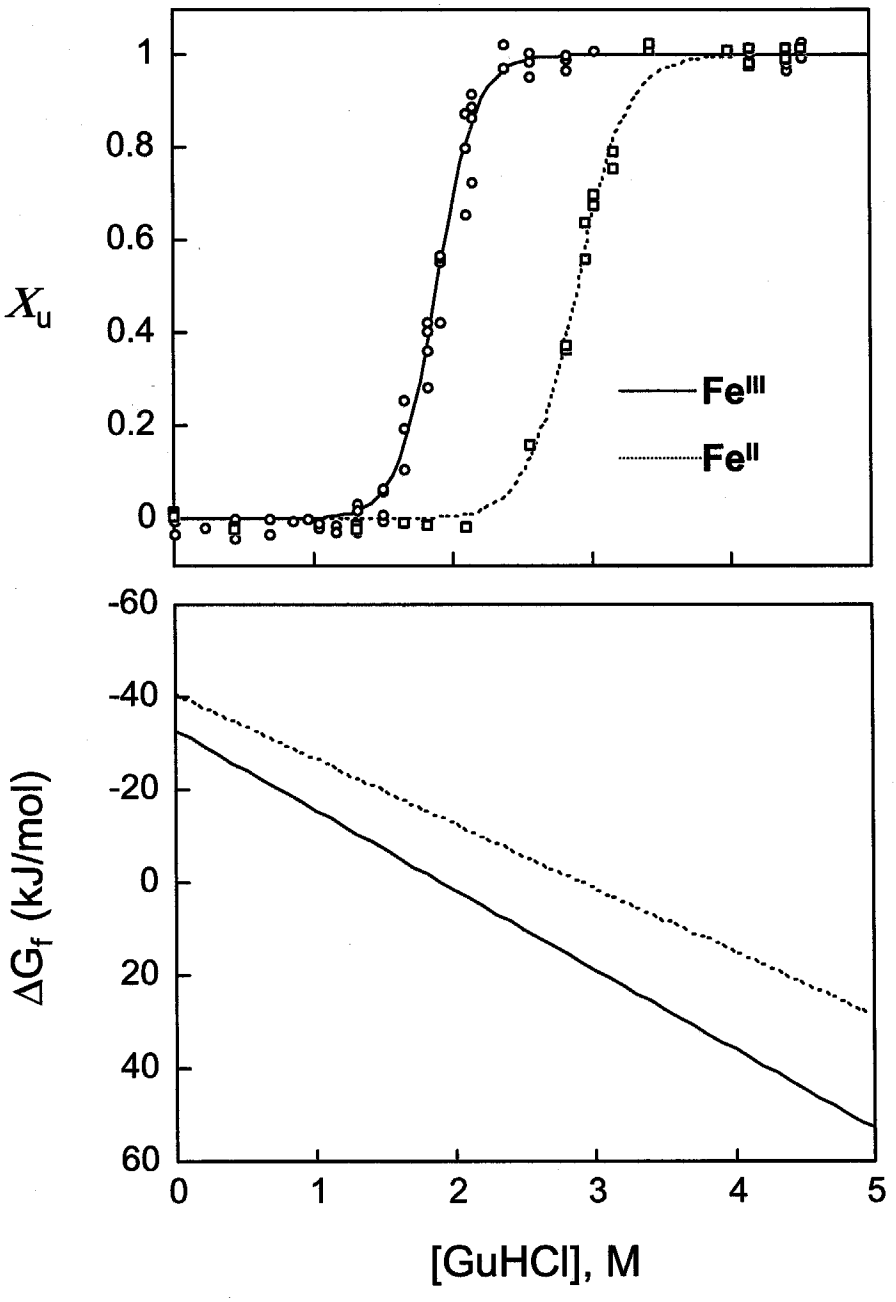


Figure 3.2. GuHCl titration curves generated from far UV-CD, Trp-72 fluorescence, and Soret absorbance (X_u is the unfolded protein fraction).

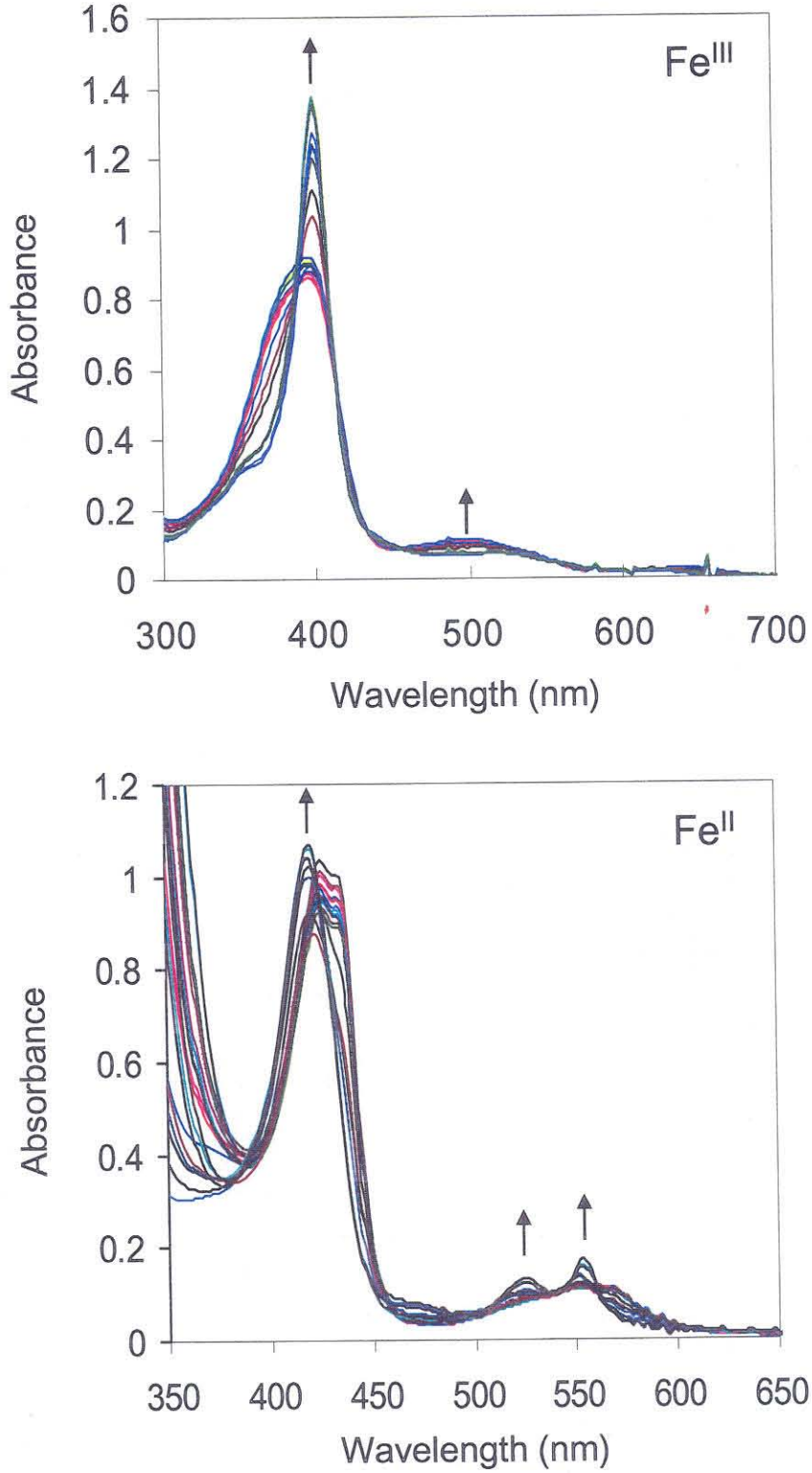


Figure 3.3. Changes in the absorption spectra of oxidized (top) and reduced (bottom) cyt *c'* upon GuHCl unfolding.

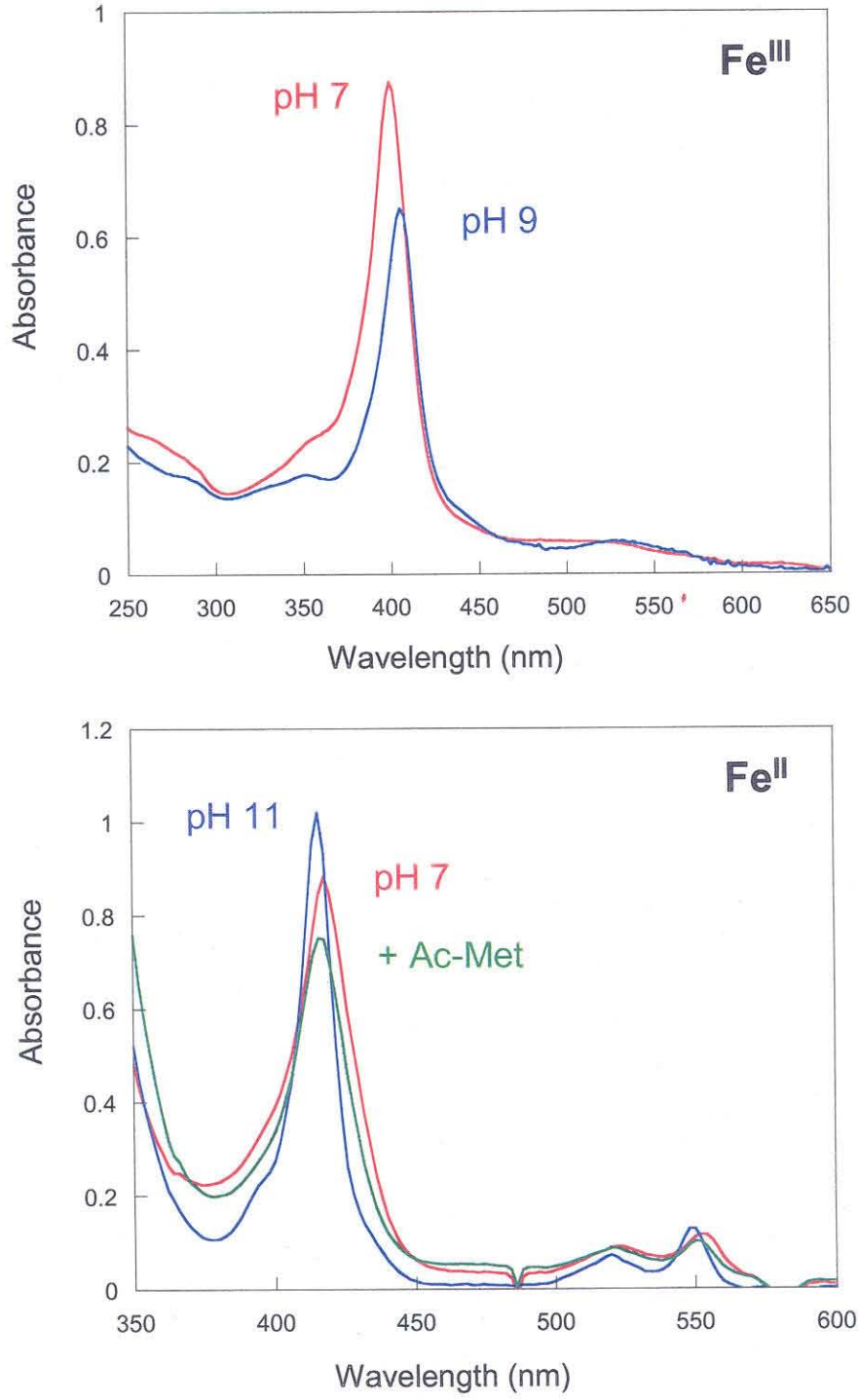


Figure 3.4. Absorption spectra of unfolded oxidized (top) and reduced (bottom) cyt *c*'. In both oxidation states, the unfolded protein binds Lys residues in alkaline solutions. Addition of N-acetyl-L-methionine (30 mM) to Fe^{II}-cyt *c*' causes little spectral perturbation.

(other than the native His-117 ligand) available to bind to the Fe^{III} heme in the unfolded protein.

GuHCl denaturation blue shifts the Soret absorption of Fe^{II}–cyt *c'* (folded: λ_{max} = 426, 435 nm; unfolded: λ_{max} = 420 nm), and two sharp Q-bands appear (λ_{max} = 524, 552 nm) (Figure 3.3). These spectroscopic features are consistent with a six- coordinate, low-spin heme. The absence of strongly shifted peaks in the ^1H NMR spectrum of unfolded Fe^{II}–cyt *c'* also signals the presence of a diamagnetic heme. The absorption spectrum of the unfolded Fe^{II} protein is independent of pH in the range 4–10, although changes in Soret absorption at higher pH levels are consistent with lysine binding (λ_{max} = 416, 520, 550 nm) (Figure 3.4). Addition of N-acetyl-L-methionine to unfolded Fe^{II}– cyt *c'* causes little perturbation of the spectrum, suggesting methionine sulfur coordination [(^{H117}N){$^{Por}N_4Fe^{II}$}(^{M}S), Met-15 and/or Met-25] in unfolded Fe^{II}–cyt *c'* (Figure 3.4). Methionine binding in unfolded Fe^{II}–cyt *c'* would raise the reduction potential and could explain the small difference between the folding free-energy changes in Fe^{II}– and Fe^{III}–cyt *c'*.

Misligated Met residues must dissociate from the heme before cyt *c'* can adopt its native folded structure. Indeed, the folding dynamics of Fe^{II}–cyt *c'* could reflect the kinetics of ligand substitutions. We have used two methods to examine the kinetics of ^{M}S binding to Fe^{II} in unfolded cyt *c'*. Laser-excited $Ru(bpy)_3^{2+}$ reduces unfolded Fe^{III}–cyt *c'* ([GuHCl] = 4.4 M) in less than a microsecond. After the rapid [1.4(1) x 10^6 s^{-1}] change in absorbance corresponding to heme reduction, there is no substantial variation in the

spectrum of the reduced protein until charge recombination regenerates unfolded Fe^{III}–cyt *c'*. The transient difference spectrum recorded 20 μs after the laser flash reveals a mixture of high- and low-spin hemes, suggesting that a nonnative sixth ligand rapidly binds to the heme in the unfolded reduced protein. Ligand binding dynamics also were examined after photodissociation of CO from unfolded Fe^{II}(CO)-cyt *c'* (4.5 M GuHCl, pH 7). The transient absorption difference spectrum measured immediately after CO dissociation (*t* = 100 ns) exhibits a broad feature in the Q-band region (530–570 nm) corresponding to a five-coordinate heme (^{H117}N){ $^{Por}N_4Fe^{II}$}. Subsequent absorption changes arise from the formation of a low-spin heme [k_{obs} = 5(2) x 10^4 s^{-1}] before CO rebinding (Figure 3.5). Both experiments indicate that ^{M}S can bind to the heme in unfolded Fe^{II}–cyt *c'* on the microsecond timescale (39–42).

Fe^{II}-cyt c' Folding. Fe^{II}–cyt *c'* folding was initiated by rapid electron injection (≈100 μs) into unfolded oxidized protein ([GuHCl] = 2.02–2.54 M) after two-photon laser excitation of NADH (43, 44). Under these conditions, heme reduction is slower than binding of the nonnative sixth ligand. The progress of the folding reaction was monitored by heme absorption in the Soret and Q-band regions from 10^{-4} to 1 s after excitation. The observed kinetics are highly heterogeneous. A small fraction (≈20%) of the population forms a high-spin heme species in about a millisecond. Complete formation of the fully folded ensemble requires several seconds (Figure 3.6). Rate constants for Fe^{II}–cyt *c'* span a range from 10^3 to 10^{-1} s^{-1} revealing fast-folding (7.0 x 10^3 s^{-1}, 8%; 5.7 x

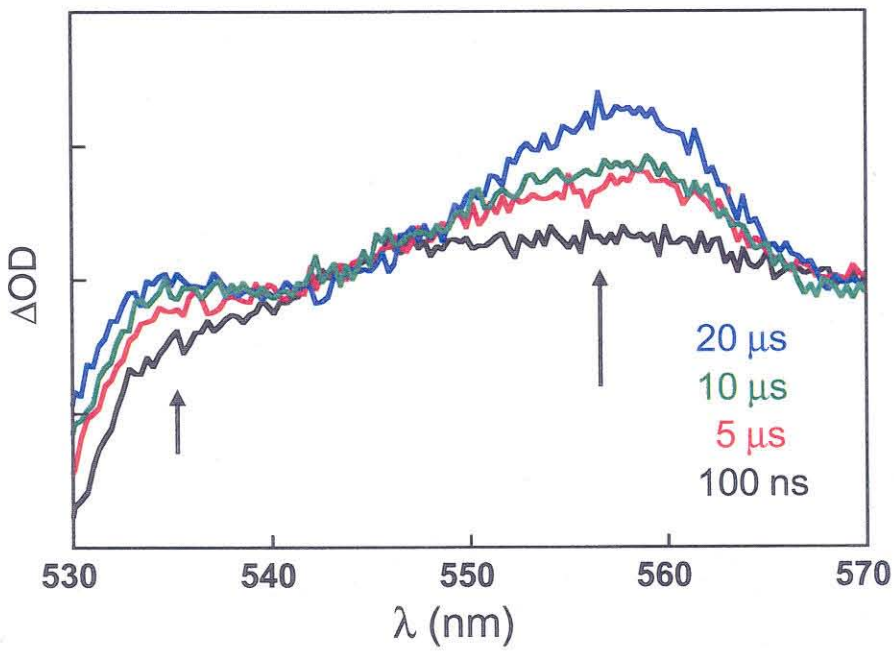


Figure 3.5. Transient absorption spectra following CO photodissociation from Fe^{II}(CO)-cyt *c'*.

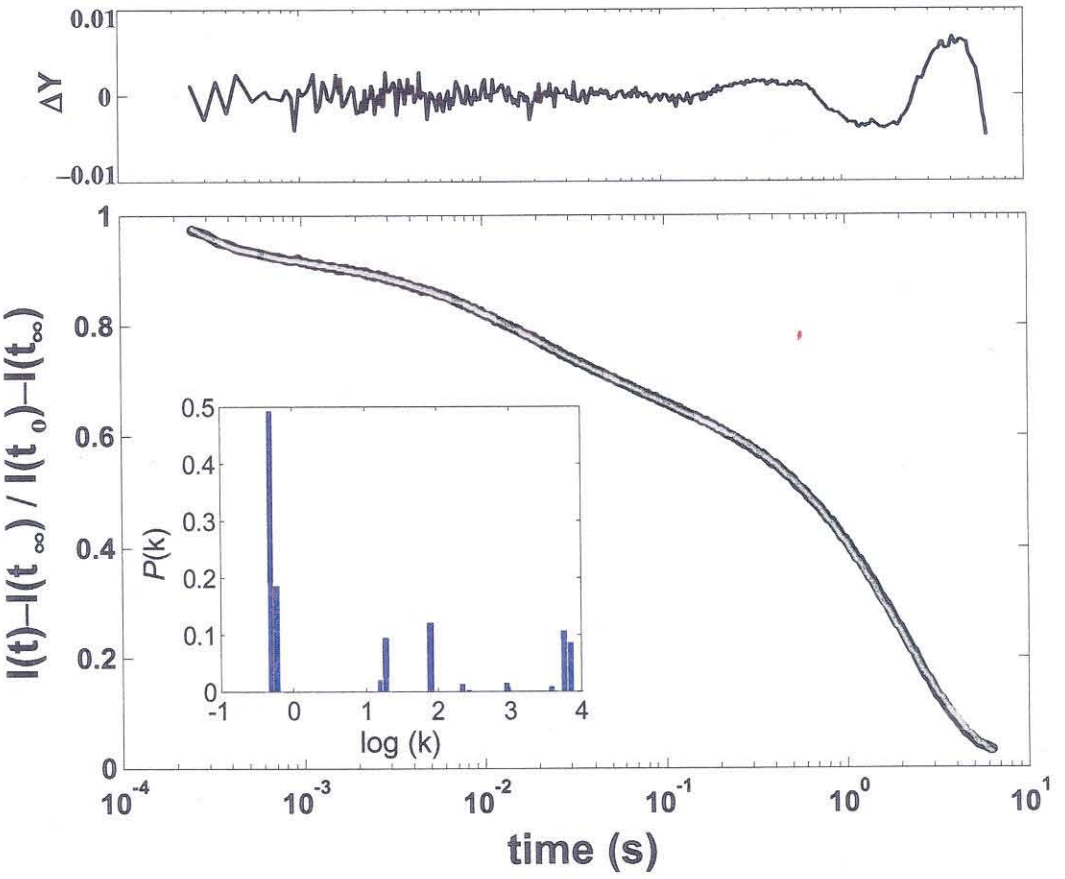


Figure 3.6. Normalized folding kinetics (observed at 440 nm, black line) fit to 80 rate constants spanning logarithmically 10^5 to 10^{-2} s^{-1} (gray line) by using a non-negative least-squares algorithm. (*Upper*) Residuals from the fit. (*Inset*) The projected population of rate constants.

10^3 s^{-1}, 9%), intermediate-folding (9.0 x10^2 to 1.5 x 10^1 s^{-1}, 24%), and slow-folding (5.9 x 10^{-1} s^{-1}, 16%; 4.8 x 10^{-1} s^{-1} 43%) components in the protein ensemble (Figure 3.6 *Inset*). Although the relative populations and the number of different rate constants vary slightly for kinetics measured at different wavelengths, heterogeneous behavior is always observed. On very long time scales, oxidation of the reduced protein leads to some uncertainty in the extracted rate constants (although the samples were carefully deoxygenated, trace amounts of oxygen were always present). The transient difference spectrum recorded 100 μs after electron injection is characteristic of a mixture of five-coordinate, high-spin and six-coordinate, low-spin ferrohemes, suggesting that the slower folding populations are misligated. The difference spectra measured at 1 and 500 ms have been fit to a combination of [(Fe^{II}-folded) – (Fe^{III}-unfolded)] and [(Fe^{II}-unfolded) – (Fe^{III}-unfolded)] molar difference spectra. The resulting folded and unfolded populations (1 ms, 25% folded, 75% unfolded; 500 ms,52% folded, 48% unfolded) are consistent with the measured amplitude changes in the single wavelength kinetics.

Fe^{II}-cyt c' Folding in the Presence of CO. To determine whether misligated heme traps are responsible for the slow Fe^{II}–cyt *c'* folding, we examined ET-triggered folding of Fe^{II}–cyt *c'* in the presence of CO. In the unfolded protein (3.32 M GuHCl, pH 7), CO binding occurs within a millisecond [1.0(4) x10^3 s^{-1}; 1 atm CO], whereas the folded protein binds CO much more slowly [1.4(4) s^{-1}]. At lower denaturant concentration ([GuHCl] = 2.1 M, pH 7) in the presence of CO,

reduction of unfolded Fe^{III} –cyt *c'* is followed by two kinetics phases: a 1.4(6) x 10^3 s^{-1} step corresponding to CO binding, and a 2.8(9) s^{-1} component due to Fe^{II}(CO)–cyt *c'* folding (Figure 3.7). The 1 ms transient difference spectrum can be modeled by two absorbing species: the major component is unfolded Fe^{II}(CO)–cyt *c'*, and the minor fraction is five-coordinate, folded Fe^{II}–cyt *c'*. Subsequent changes in absorbance are associated with the conversion of unfolded Fe^{II}(CO)– to folded Fe^{II}(CO)–cyt *c'*.

Fe^{III}–cyt c' Folding. Nonnative methionine-heme ligation is disfavored in unfolded Fe^{III}–cyt *c'* (37), and the axial heme coordination sites are occupied by water and His-117. The stopped-flow triggered folding of Fe^{III}–cyt *c'* is adequately described by a single kinetics phase with a rate constant that decreases exponentially with increasing denaturant concentration (Figure 3.8). The folding rate constant [1.0(2) s^{-1}; [GuHCl] = 0.95 M; ΔG_f = –16 kJ mol^{-1}] is comparable to that of Fe^{II}(CO)–cyt *c'* and of the slower folding population of Fe^{II}–cyt *c'* ([GuHCl] = 2.16 M; ΔG_f = –10 kJ mol^{-1}). The extrapolated rate constant for folding Fe^{III}–cyt *c'* in the absence of denaturant is 2.5 x 10^2 s^{-1} ($\Delta G_f°$ = –33(3) kJ mol^{-1}).

Discussion

It is reasonable to expect that topologically homologous proteins would fold at similar rates, yet cytochromes b_{562} and *c'* display quite disparate folding kinetics. Apparently, there are factors beyond structural topology that must be considered to explain the folding kinetics of these four-helix bundles. Three key

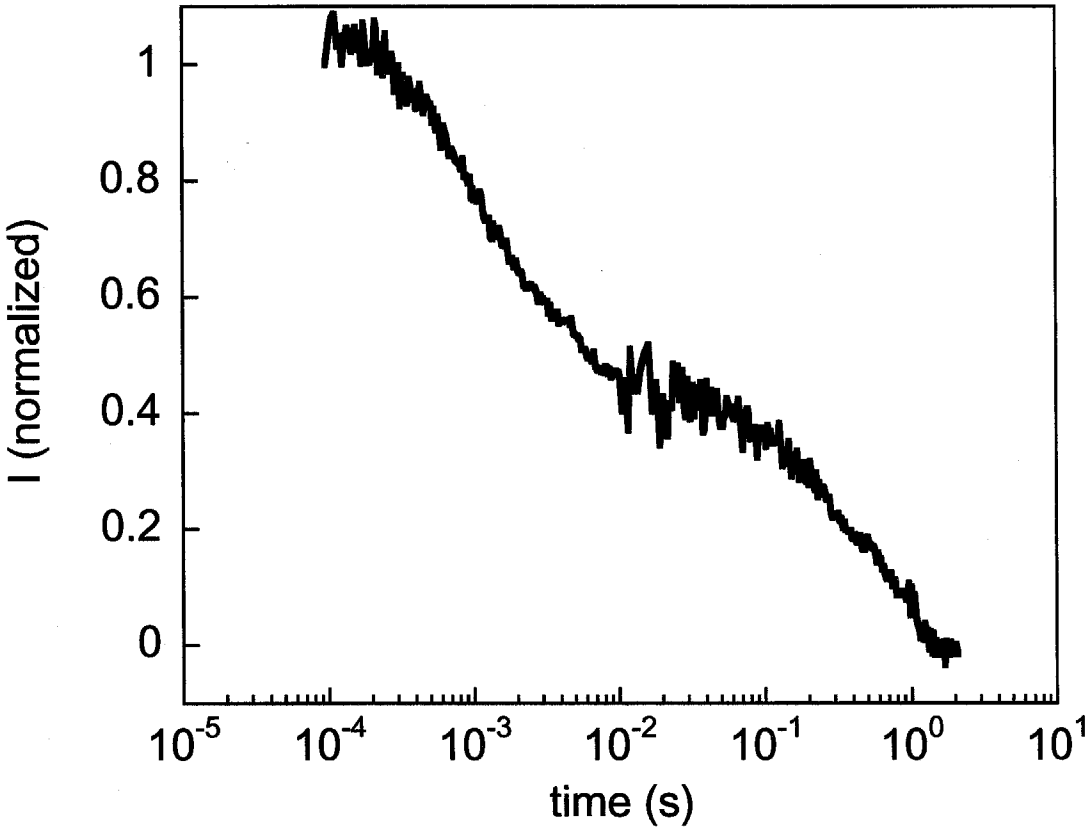


Figure 3.7. Normalized folding kinetics (observed at 416 nm) of Fe^{II}-cyt *c*' in the presence of CO (2.1 M GuHCl, pH 7, 1 atm CO).

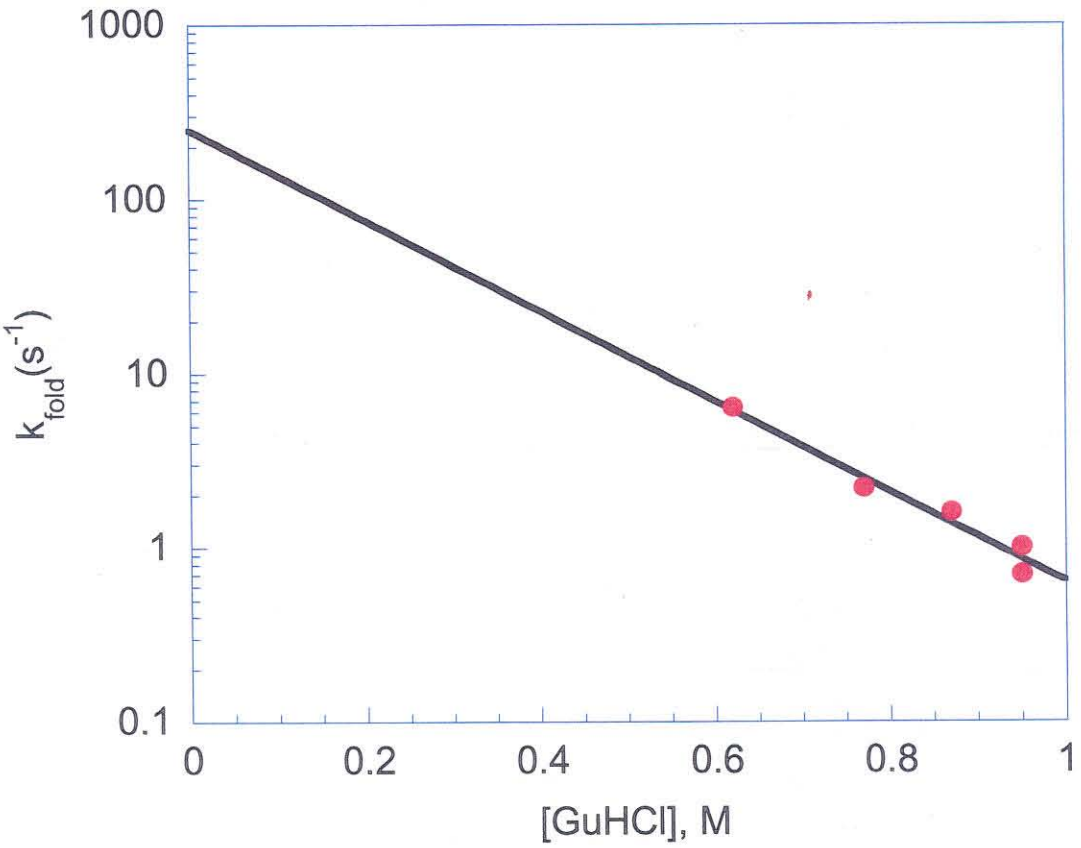


Figure 3.8. Dependence of Fe^{III}-cyt *c'* folding rates on GuHCl concentration.

features of cyt b_{562} and *c'* are likely to contribute to the differences in folding kinetics: covalent attachment of the porphyrin to the polypeptide, nonnative heme–ligand traps, and folding driving force.

The rapid dissociation of the heme from unfolded Fe^{II}–cyt b_{562} limits the timescale for observation of folding kinetics. Because the porphyrin is covalently attached in Fe–cyt *c'*, the folding kinetics cover a far wider time range and are considerably more complex. A fraction of the unfolded Fe^{II}–cyt *c'* ensemble refolds rapidly, but several seconds are required for the entire sample to fold. Proline isomerization is known to inhibit protein folding (45) but is unlikely to be responsible for the slow Fe–cyt *c'* kinetics. Although there are four proline residues in the cyt *c'* sequence, the observed folding rates depend on denaturant concentration and are substantially faster than typical proline isomerizations (45). It is interesting to note that the fast-folding fraction of Fe^{II}–cyt *c'* is roughly comparable to the yield of folded Fe^{II}–cyt b_{562}. It is possible, then, that if heme dissociation could be inhibited, Fe^{II}–cyt b_{562} would display slower and more complex folding kinetics.

The refolding of Fe^{II}–cyt *c'* can be described qualitatively by a kinetic partitioning mechanism (46, 47). At the instant that folding is initiated, a fraction of the denatured proteins will have adopted conformations that can smoothly and rapidly fold to the native structure. The fast-folding population of Fe^{II}–cyt *c'* would correspond to this group. Folding of the remaining fraction is frustrated by transient trapping in local minima on the folding energy landscape. Escape from these misfolded structures is an activated process that slows formation of the

native structure. The partition factor that determines the balance between fast- and slow-folding populations depends on the primary sequence as well as the refolding conditions (46–51). Simulated folding kinetics of model three-helix-bundle proteins (48, 49) point to a central role for the folding free-energy gap in determining the partitioning between fast- and slow-track folding. The heterogeneous folding kinetics of Fe^{II}–cyt *c*' are strikingly similar to the simulated folding kinetics of a small-gap three-helix bundle (49).

Nonnative methionine ligands in unfolded Fe^{II}–cyt *c*' contribute to the heterogeneity of the folding kinetics. In addition to the fast-folding population, there are intermediate ($\approx 10^2\ s^{-1}$) and slow-folding ($< 1\ s^{-1}$) components. Iron–sulfur bond dissociation is not rate limiting, because ferroheme–methionine ligand exchange is faster than either folding phase (40). In the presence of CO, the nonnative Met ligands are displaced from the ferroheme, and folding is dominated by a slow phase with a time constant of ≈350 ms. Similarly, Fe^{III}–cyt *c*' folding is not complicated by heme–ligand binding, and a single 1 s phase predominates. These observations suggest that, in the absence of misligation, cyt *c*' requires 500 ms to 1 s to adopt a folded structure. The presence of an intermediate ($\approx 10^2\ s^{-1}$) folding phase in Fe^{II}–cyt c' implies that nonnative methionine ligation can, in some instances, facilitate refolding. The protein has two Met residues that can bind to the ferroheme when the protein unfolds. It is possible that the intermediate phase arises from protein populations with $(^{117}N)\{^{Por}N_4Fe^{II}\}(^{M15}S)$ heme coordination, and the slow phase corresponds to populations with $(^{H117}N)\{^{Por}N_4Fe^{II}\}(^{M25}S)$ ligation (Figure 3.9). Met-15 is in the

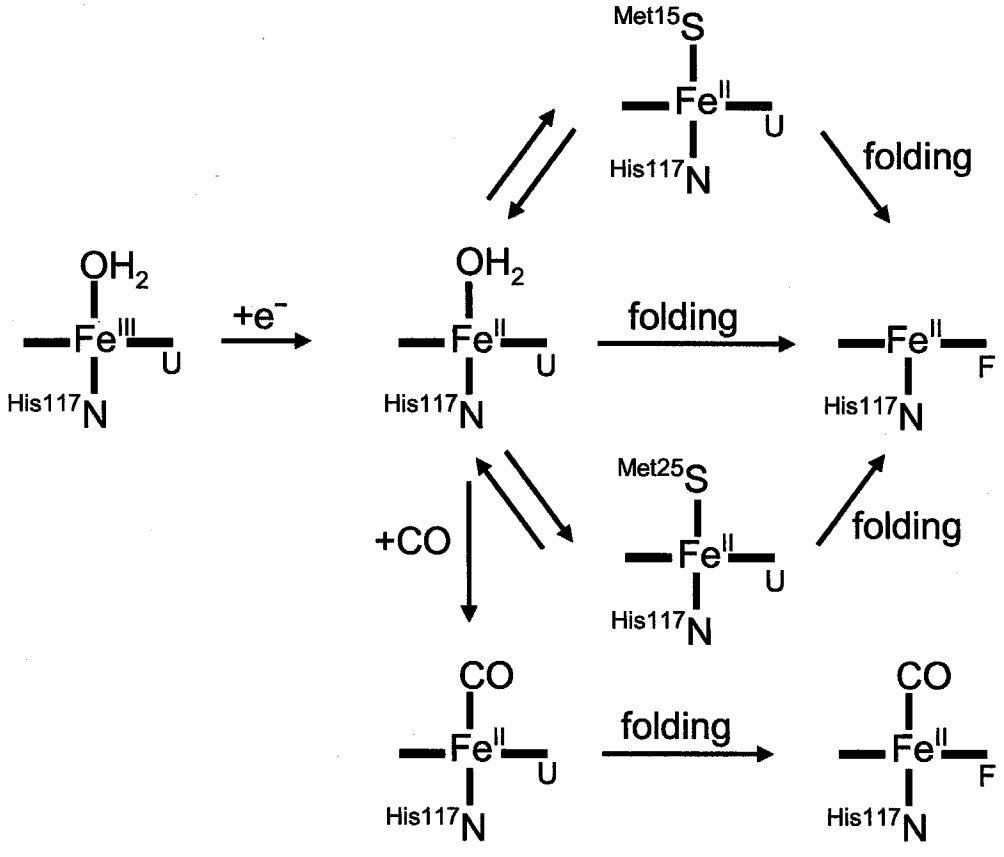


Figure 3.9. Schematic representation of heme coordination environment after electron-transfer triggered folding of Fe^{III}-cyt *c'* in the presence and absence of CO. The subscripts F and U refer to folded and unfolded polypeptides.

distal heme pocket in the folded protein, so an intermediate with $(^{H117}N)\{^{Por}N_4Fe^{II}\}(^{M15}S)$ coordination might fold more readily than a $(^{H117}N)\{^{Por}N_4Fe^{II}\}(^{M25}S)$ form.

The Fe–cyt *c*' folding rates are compared in Figure 3.10 to those of four structurally related proteins (11, 15, 19). The four-helix bundles Im7 and Im9 have 60% sequence identity and nearly identical three-dimensional structures, yet Im9 folds more than 10 times slower than Im7 (15). Curiously, Im9 displays the hallmarks of two-state folding, whereas intermediates are implicated in the faster Im7 folding process (15, 16). The folding of ACBP has been interpreted in terms of a two-state model (12, 13), although subsequent quenched-flow hydrogen exchange studies revealed that an intermediate is formed before the rate-limiting step (14). Among helical-bundle structures, there is a wide variation in measured folding rates as functions of $-\Delta G_f$ (Figure 3.10). It is clear that, even in the absence of misligation, cyt *c*' is the slowest folding protein in this structural class. The difference is even greater when account is taken of the fact that the Im7, Im9, and ACBP folding kinetics were measured at reduced temperatures. Cyt *c*' is substantially larger than the other helical bundles (Im7, 85 residues; Im9, 85 residues; ACBP, 86 residues; and cyt *c*', 126 residues), and its size may be in part responsible for its slower folding. On a free-energy basis, the Fe–cyt *c*' folding kinetics more closely resemble those of cyt *c* (52) than those of the nonheme helical-bundle proteins. Nonnative ligand binding clearly can perturb heme protein folding; it is likely that noncovalent, nonnative heme–polypeptide contacts represent additional sources of frustration.

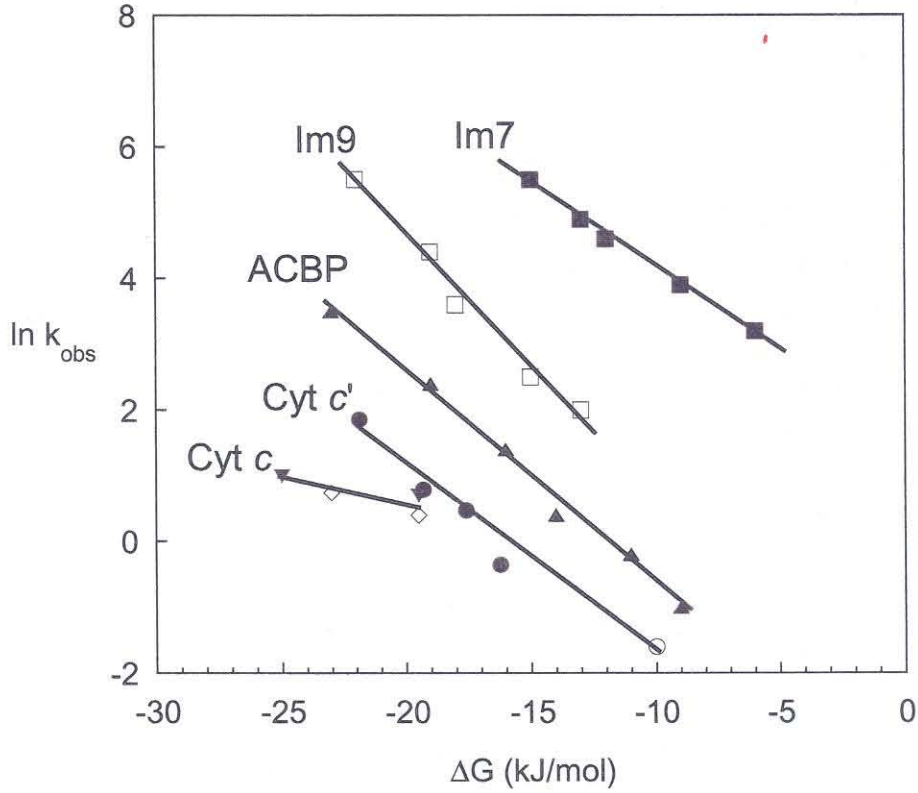


Figure 3.10. Measured folding rates of helical proteins as functions of folding driving force (ΔG_f). Im7 (85 aa), pH 7, 10 °C (■) (15); Im9 (85 aa), pH 7, 10 °C (□) (15); bovine ACBP (86 aa), pH 5.3, 5 °C (♦) (12); Fe^{II}-cyt *c'* (126 aa), pH 7, 22 °C, slowest phase (O); Fe^{III}-cyt *c'*, pH 7, 22 °C (●); equine Fe^{II}-cyt *c* (104 aa), pH 7, 22 °C (◊) (52); yeast Fe^{II}-cyt *c* (108 aa), pH 7, 22 °C (♦) (52).

References

1. Bryngelson, J. D., Onuchic, J. N. and Wolynes, P. G. (1995) *Proteins: Struct. Func. Gen.* **21,** 167-195.
2. Onuchic, J. N., Lutheyschulten, Z. and Wolynes, P. G. (1997) *Annu. Rev. Phys. Chem.* **48,** 545-600.
3. Dill, K. A. and Chan, H. S. (1997) *Nature Struct. Biol.* **4,** 10-19.
4. Plaxco, K. W., Simons, K. T. and Baker, D. (1998) *J. Mol. Biol.* **277,** 985-994.
5. Plaxco, K. W., Simons, K. T., Ruczinski, I. and Baker, D. (2000) *Biochemistry* **39,** 11177-11183.
6. Dinner, A. R. and Karplus, M. (2001) *Nature Struct. Biol.* **8,** 21-22.
7. Debe, D. A., Carlson, M. J. and Goddard III, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 2596-2601.
8. Muñoz, V. and Eaton, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11311-11316.
9. Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1997) *J. Am. Chem. Soc.* **119,** 9562-9563.
10. Wittung-Stafshede, P., Lee, J. C., Winkler, J. R. and Gray, H. B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 6587-6590.
11. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. and Poulsen, F. M. (1995) *Biochemistry* **34,** 7217-7224.
12. Kragelund, B. B., Højrup, P., Jensen, M. S., Schjerling, C. K., Juul, E., Knudsen, J. and Poulsen, F. M. (1996) *J. Mol. Biol.* **256,** 187-200.

13. Kragelund, B. B., Osmark, P., Neergaard, T. B., Schiødt, J., Kristiansen, K., Knudsen, J., Poulsen and M., F. (1999) *Nature Struct. Biol.* **6,** 594-601.

14. Teilum, K., Kragelund, B. B., Knudsen, J. and Poulsen, F. (2000) *J. Mol. Biol.* **301,** 1307-1314.

15. Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C. and Radford, S. E. (1999) *J. Mol. Biol.* **286,** 1597-1608.

16. Capaldi, A. P., Shastry, M. C. R., Kleanthous, C., Roder, H. and Radford, S. E. (2001) *Nature Struct. Biol.* **8,** 68-72.

17. Huang, G. S. and Oas, T. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6878-6882.

18. Pascher, T., Chesick, J. P., Winkler, J. R. and Gray, H. B. (1996) *Science* **271,** 1558-1560.

19. Telford, J. R., Wittung-Stafshede, P., Gray, H. B. and Winkler, J. R. (1998) *Acc. Chem. Res.* **31,** 755-763.

20. Meyer, T. E. and Kamen, M. D. (1982) *Adv. Protein Chem.* **35,** 105-212.

21. Moore, G. R., McClune, G. J., Clayden, N. J., Williams, R. J. P., Alsaadi, B. M., Angstrom, J., Ambler, R. P., van Beeumen, J., Tempst, P., Bartsch, R. G., Meyer, T. E. and Kamen, M. D. (1982) *Eur. J. Biochem.* **123,** 73-80.

22. Moore, G. R. and Pettigrew, G. W. (1990) *Cytochromes c: Evolutionary, Structural, and Physicochemical Aspects* (Springer-Verlag, New York).

23. Ambler, R. P., Bartsch, R. G., Daniel, M., Kamen, M. D., McLellan, L., Meyer, T. E. and Beeumen, J. V. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6854-6857.

24. Hamada, K., Bethge, P. H. and Mathews, F. S. (1995) *J. Mol. Biol.* **247,** 947-962.

25. Shibata, N., Iba, S., Misaki, S., Meyer, T. E., Bartsch, R. G., Cusanovich, M. A., Morimoto, Y., Higuchi, Y. and Yasuoka, N. (1998) *J. Mol. Biol.* **284,** 751-760.

26. Moore, G. R., Williams, R. J. P., Peterson, J., Thomson, A. J. and Mathews, F. S. (1985) *Biochim. Biophys. Acta* **829,** 83-96.

27. Klerk, H. D., Bartsch, R. G. and Kamen, M. D. (1965) *Biochim. Biophys. Acta* **97,** 275-280.

28. Strekas, T. C. and Spiro, T. G. (1974) *Biochim. Biophys. Acta* **351,** 237-245.

29. Fujii, S., Yoshimura, T., Kamada, H., Yamaguchi, K., Suzuki, S., Shidara, S. and Takakuwa, S. (1995) *BBA Prot. Struct. Molec. Enzym.* **1251,** 161-169.

30. Clark, K., Dugad, L. B., Bartsch, R. G., Cusanovich, M. A. and Mar, G. N. L. (1996) *J. Am. Chem. Soc.* **118,** 4654-4664.

31. Jackson, J. T., Mar, G. N. L. and Bartsch, R. G. (1983) *J. Biol. Chem.* **258,** 1799-1805.

32. Barakat, R. and Strekas, T. C. (1982) *Biochim. Biophys. Acta* **679,** 393-399.

33. Bartsch, R. G. (1971) *Meth. Enzymol.* **23,** 344-363.

34. Bartsch, R. G. (1978) in *The Photosynthetic Bacteria*, eds. Clayton, R. K. and Sistrom, W. R. (Plenum Press, New York), pp. 249-279.

35. Nozaki, Y. (1972) *Meth. Enzymol.* **26,** 43-50.

36. Stowell, M. H. B., Larsen, R. W., Winkler, J. R., Rees, D. C. and Chan, S. I. (1993) *J. Phys. Chem.* **97,** 3054-3057.

37. Tezcan, F. A., Winkler, J. R. and Gray, H. B. (1998) *J. Am. Chem. Soc.* **120,** 13383-13388.

38. Bren, K. L. and Gray, H. B. (1993) *J. Am. Chem. Soc.* **115,** 10382-10383.

39. Hagen, S. J., Hofrichter, J., Szabo, A. and Eaton, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11615-11617.

40. Hagen, S. J., Hofrichter, J. and Eaton, W. A. (1997) *J. Phys. Chem. B* **101,** 2352-2365.

41. Lapidus, L. J., Eaton, W. A. and Hofrichter, J. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 7220-7225.

42. Hagen, S. J., Carswell, C. W. and Sjolander, E. M. (2001) *J. Mol. Biol.* **305,** 1161-1171.

43. Telford, J. R., Tezcan, F. A., Gray, H. B. and Winkler, J. R. (1999) *Biochemistry* **38,** 1944-1949.

44. Orii, Y. (1993) *Biochemistry* **32,** 11910-11914.

45. Schmid, F. X., Mayr, L. M., Mücke, M. and Schönbrunner, E. R. (1993) *Adv. Protein Chem.* **44,** 25-66.

46. Thirumalai, D., Klimov, D. K. and Woodson, S. A. (1997) *Theor. Chem. Acc.* **96,** 14-22.

47. Thirumalai, D. and Klimov, D. K. (1999) *Curr. Opin. Struct. Biol.* **9,** 197-207.

48. Zhou, Y. and Karplus, M. (1999) *J. Mol. Biol.* **293,** 917-951.

49. Zhou, Y. and Karplus, M. (1999) *Nature* **401,** 400-403.

50. Shea, J.-E., Onuchic, J. N. and Brooks III, C. L. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 12512-12517.

51. Shea, J.-E., Onuchic, J. N. and Brooks III, C. L. (2000) *J. Chem. Phys.* **113,** 7663-7671.

52. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. and Gray, H. B. (1996) *Chem. Biol.* **3,** 491-497.

Chapter 4

Does methionine misligation facilitate cytochrome *c*' folding?[†]

[†]Expression and purification of Q1A/M15L/M25A, Q1A/E17C, and dansyl-labeled Q1A/E17C cyt *c*' proteins were done in collaboration with Lisa Van Hoozer. Fluorescence lifetime measurements were done in collaboration with K. Cecilia Engman.

Introduction

Rapid electron-transfer triggering (≈ 100 μs) of ferrocytochrome *c*' (Fe^{II}–cyt *c*') folding reveals complex kinetics spanning times from milliseconds to tens of seconds (1). Fe^{II}–cyt *c*' folding can be described by a partitioning mechanism in which some denatured structures are poised to fold quickly while other unfolded molecules fall into traps (*e.g.*, misligation) on the folding landscape (2, 3). In contrast, folding of Fe^{III}–cyt *c*' and Fe^{II}(CO)–cyt *c*' exhibit just one kinetics phase, suggesting that heme misligation is responsible for the heterogeneous folding behavior (1). The absorption spectrum and reduction potential of unfolded Fe^{II}–cyt *c*' (extracted from unfolding curves) suggest that Met residues (Met-15 and/or Met-25) bind to the ferroheme in the denatured protein (Figure 4.1). Unlike cytochrome *c* folding, where the rate-limiting step is ligand dissociation (4–7), elimination of misligated traps in cyt *c*' apparently does not affect the overall folding rate (≈ seconds). Instead, the observation of an intermediate-folding population (≈ 100 ms) suggests that formation of near-native contacts (Met-15 ligation) could speed up cyt *c*' folding.

Cyt *c*' has been cloned and expressed in *E. coli*, allowing for site-directed mutagenesis of the methionine residues (8). Single (M15L and M25A) and double (M15L/M25A) methionine mutants have been prepared and characterized. By studying these mutant proteins, we should be able to determine whether the two methionine residues are responsible for the different folding populations. In addition, a Q1A/E17C mutant has been prepared and labeled with a fluorescent dansyl-molecule, 1,5-IAEDANS (5-((((2-

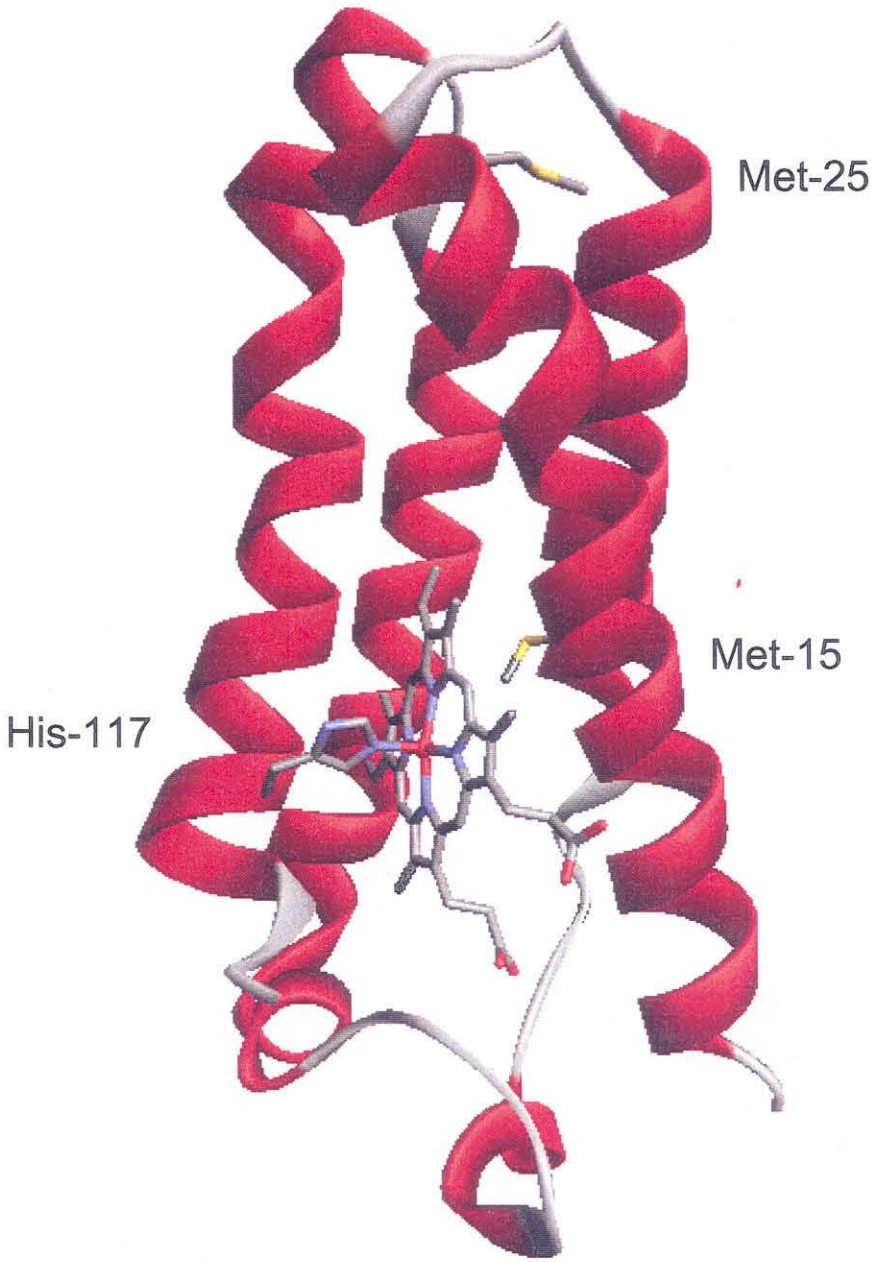


Figure 4.1. Crystal structure of *R. palustris* cyt *c'* with the methinione residues highlighted. In the folded protein, Met-15 sits above the heme pocket whereas Met-25 is much more distant from the heme in the first helix.

iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid). Fluorescence energy transfer (FET) from the dansyl fluorophore to the heme moiety will allow us to monitor the polypeptide backbone movements associated with methionine binding. Measurements of FET kinetics could indicate whether either of the methionine residues, Met-15 or Met-25, is bound in the reduced unfolded state.

Materials and Methods

Guanidine hydrochloride (GuHCl, Sigma, ultrapure grade) and NADH (Sigma) were used as received. *R. palustris* cyt *c'* gene was cloned (plasmid cytcp1) and the protein was heterologously expressed (with plasmid pEC86, kindly provided by Dr. Linda Thöny-Meyer) in *E. coli* (strain BL21(DE3) (9)). Pseudo-wild-type (Q1A) and mutant proteins (Q1A/M15L, Q1A/M25A, Q1A/M15L/M25A, and Q1A/E17C) were generated by site-directed mutagenesis using a Quickchange kit (Stratagene). Detailed procedures are described in Appendix A. DNA sequencing was performed by the Caltech DNA Sequencing Core Facility.

Cyt *c'* proteins were expressed and purified with minor modifications of procedures described in Appendix A. Proteins were extracted by complete cell lysis and were purified by cationic exchange chromatography using first a CM Sepharose Fast Flow (Pharmacia) batch column (eluted with 10 mM sodium phosphate and 50-100 mM NaCl, pH 6) followed by a CM Sepharose CL-6B (Sigma) column (eluted with a linear salt gradient from 0 to 100 mM NaCl in 10 mM sodium phosphate buffer, pH 6) at 4 °C. Q1A/E17C cyt *c'* (50 μM in 20 mM

sodium phosphate buffer, pH 7), after treating with DTT and desalted with a gel-filtration column (PD-10, Pharmacia), was reacted with a cysteine-selective labeling reagent, 1,5-IAEDANS (10x molar excess, dissolved in minimal dimethyl sulfoxide (<1% of total solution)) in the dark for 3.5 h (room temperature). Dansyl-labeled protein was repeatedly chromatographed on a Mono S 10/10 column (Pharmacia) to separate from unmodified cyt *c*' (Figure 4.2). Mass spectral (Figure 4.3) and protein sequence (Figure 4.4) analyses were conducted by the Protein/Peptide Microanalytical Laboratory in the Beckman Institute at Caltech. Purity of protein samples was assessed by SDS-PAGE gel electrophoresis and absorbance ratio ($\lambda_{280\ nm}$:$\lambda_{398\ nm}$) between 0.19–0.21 (with the exception of the dansyl-labeled protein; $\lambda_{280\ nm}$:$\lambda_{398\ nm}$ = 0.22–0.23).

Absorption and CD spectra were measured on a Hewlett-Packard 8452 diode array spectrophotometer and Aviv 62ADS spectropolarimeter. Denaturation curves were generated from heme absorption and CD spectral changes according to standard procedures (10). GuHCl concentrations were determined by refractive index measurements (11).

Transient absorption spectra and kinetics were measured by using an apparatus described previously (12). The excitation source was the third harmonic of a Q-switched Nd:YAG laser (355 nm) for electron photoinjection from NADH. A 16-bit, 20 MS/s digital oscilloscope (CompuScope 1602, Gage Applied) was used to record the complete kinetics traces from 10^{-3} to 10^{1} s (sampled at 10 kS/s). Samples for kinetics measurements were deoxygenated by repeated evacuation/argon-fill cycles on a Schlenk line. Typical samples

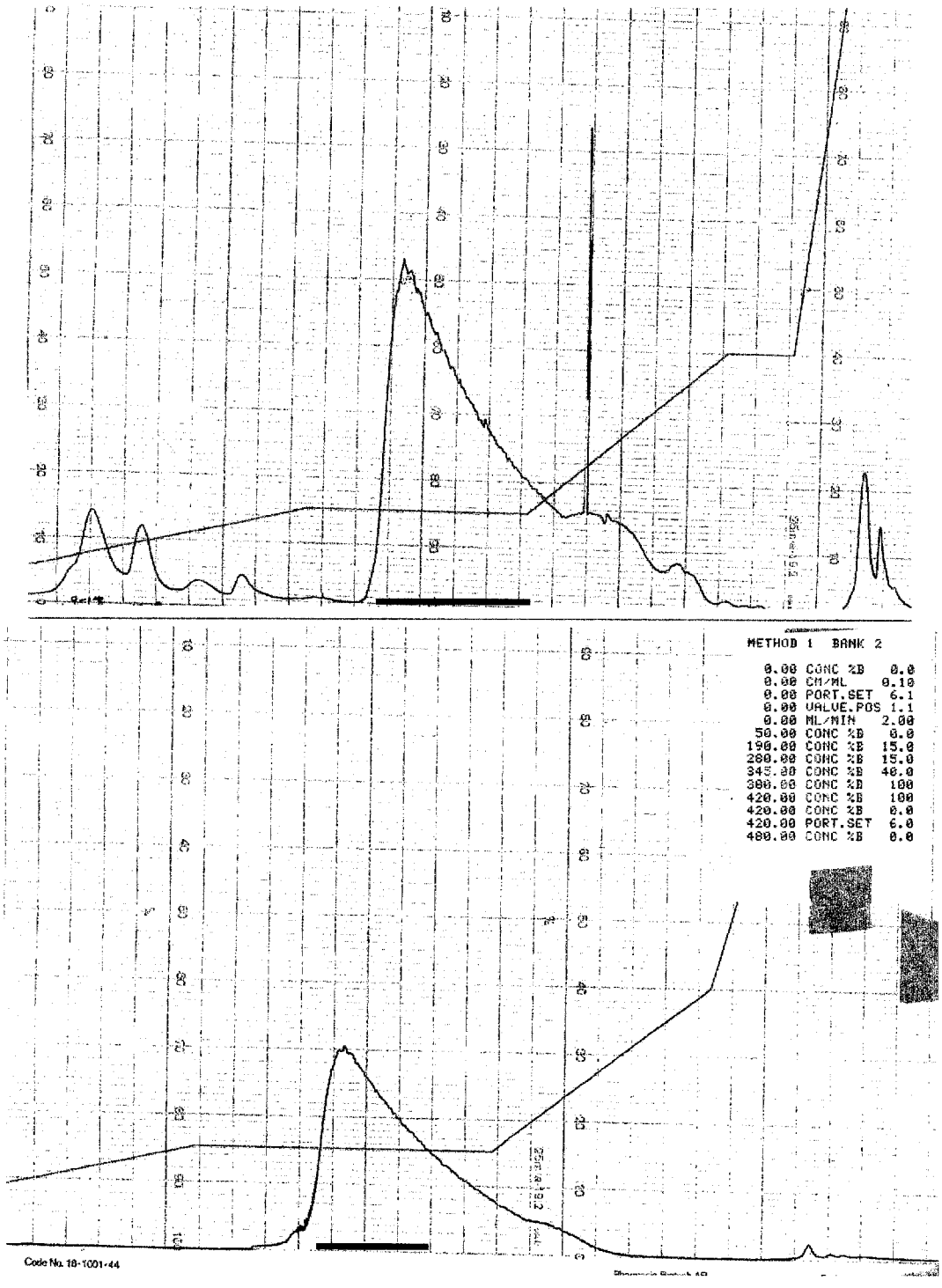
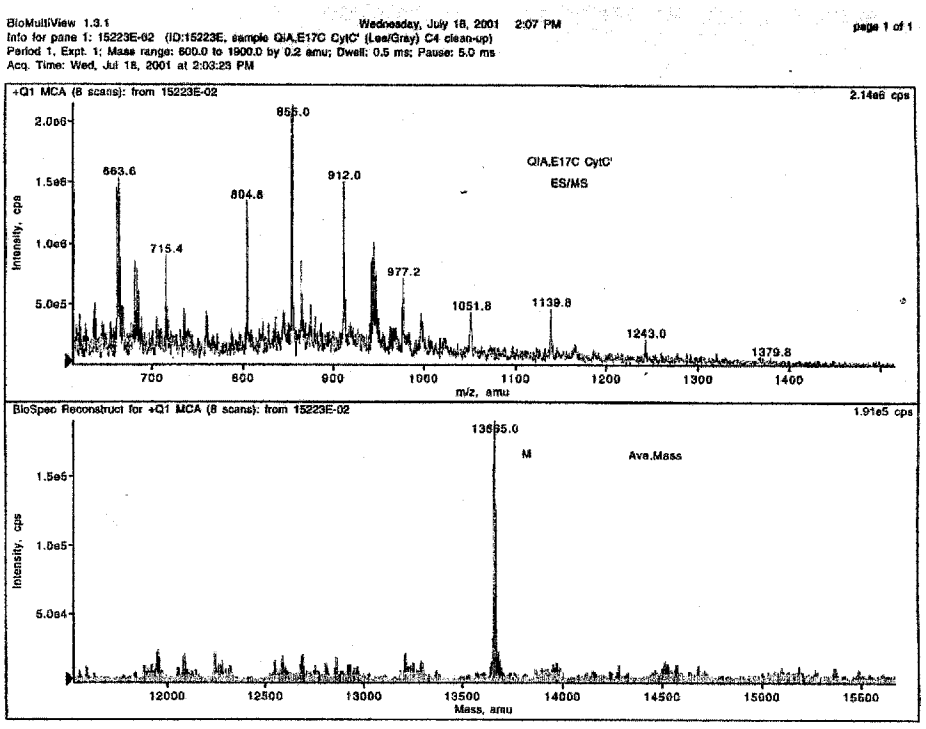


Figure 4.2. FPLC purifications of 1,5-IAEDANS-labeled Q1A/E17C cyt *c*'. The black line (top panel) highlights fractions pooled and repurified. The black line (bottom panel) indicates fractions pooled for experiments.

A



B

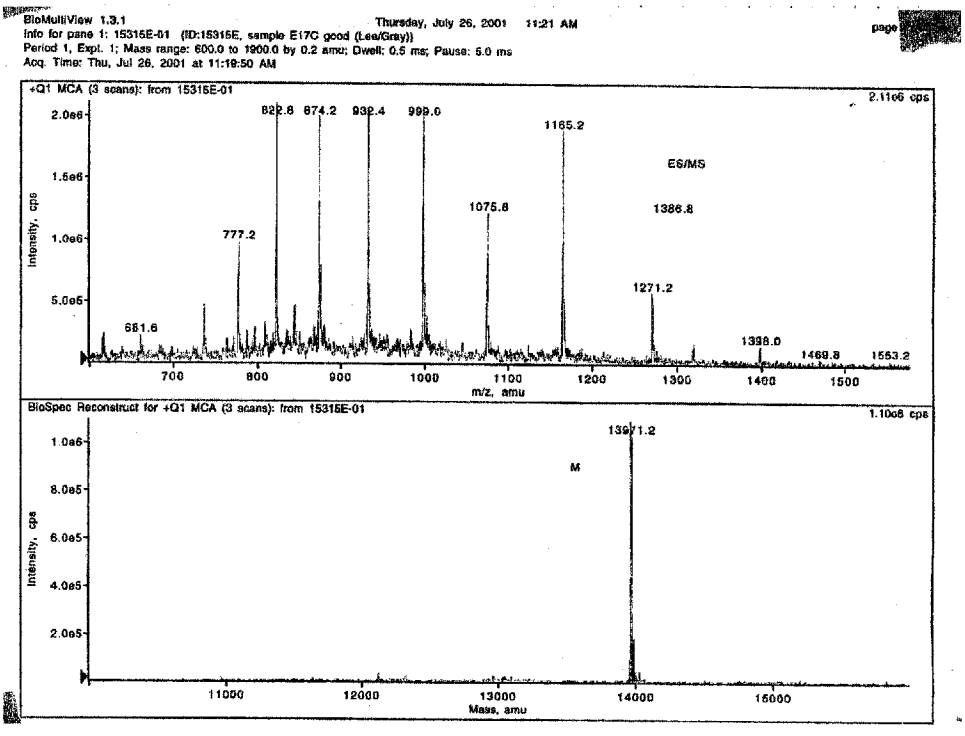


Figure 4.3. Electrospray mass spectra of Q1A/E17C (A) and 1,5-IAEDANS-labeled Q1A/E17C (B) cyt *c*'.

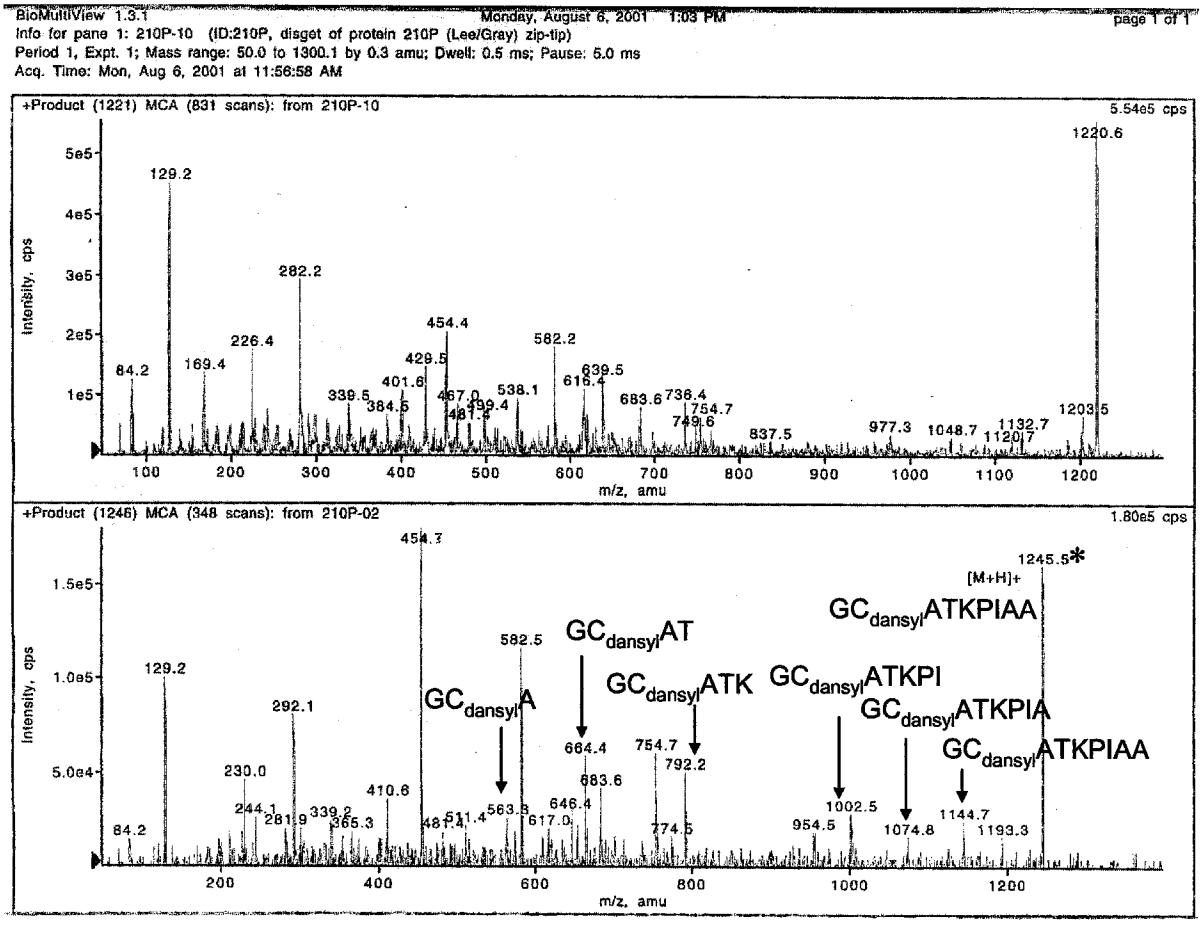


Figure 4.4. MS/MS fragmentation pattern of CNBr-cleaved peptide (residues 16-24, 1245.5 amu) of 1,5-IAEDANS-labeled Q1A/E17C cyt *c*', confirming the position of the fluorophore.

contained cyt *c*' (100 μM), NADH (200 μM), and GuHCl (1.6–2.2 M) in 100 mM sodium phosphate buffer (pH 7).

Measured kinetics traces were logarithmically compressed (100 points per time decade), normalized $[I(t) - I(t_\infty)/I(t_0) - I(t_\infty)]$, and fit by using a nonnegative linear least-squares minimization algorithm, LSQNONNEG, in MATLAB (The Math Works, Inc.) to project the folding kinetics onto a basis of exponential decay functions with rate constants ranging from 10^5 to 10^{-2} s^{-1} (10 rate constants per decade, spaced logarithmically).

Fluorescence lifetime measurements were measured by using the third harmonic of a regeneratively amplified femtosecond Ti:sapphire laser (Spectra-Physics) at 290 nm as the excitation source and a picosecond streak camera (Hamamatsu) as the detector (optical configurations are illustrated in Chapter 5). A cutoff filter was used to observe emission wavelengths above 475 nm. FET kinetics were fit by using a LSQNONNEG algorithm to give probability distributions of rate constants, $P(k)$, in the range 10^{11} to 10^7 s^{-1} (rate constants were spaced with a resolution ratio (k_{i+1}/k_i) of 1.2).

Results

Equilibrium unfolding. GuHCl titration curves show that the stabilities of both oxidized ($[GuHCl]_{½}$ = 1.9(1) M; $\Delta G_f°$ = –33(5) kJ mol^{-1}) and reduced ($[GuHCl]_{½}$ = 2.9(1) M; $\Delta G_f°$ = –38(5) kJ mol^{-1}) forms of recombinant Q1A cyt *c*' are nearly identical with that of the corresponding wild-type protein (Fe^{III}: $[GuHCl]_{½}$ = 1.9(1) M; $\Delta G_f°$ = –33(3) kJ mol^{-1} and Fe^{II}: $[GuHCl]_{½}$ = 2.9(1) M; $\Delta G_f°$ = –40(4) kJ mol^{-1})

isolated from *R. palustris* (Figure 4.5) (1). In contrast, the single methionine mutants of cyt *c'* are measurably less stable than Q1A cyt *c'* (Table 4.1). In particular, the folding free-energy change extrapolated to zero denaturant ($\Delta G_f°$) in the Q1A/M25A mutant is ≈10 kJ mol^{-1} lower than that of pseudo-wild-type Q1A cyt *c'*. Although the methionine mutants are destabilized compared to the wild-type protein, the differences in folding free energies between Fe^{III}– and Fe^{II}–cyt *c'* proteins ($\Delta\Delta G_f°$ =7(2) kJ mol^{-1} ≈70(20) meV) are similar to that found for Q1A cyt *c'*, suggesting that a methionine residue is ligated to the Fe^{II}–center in the unfolded mutant proteins. This conclusion is based on the estimated reduction potential (~0 mV) of unfolded cyt *c'*, which is close to that measured for a solvent-exposed heme with His and Met ligands (Ac-MP11, proteolytic fragment from cyt *c*, E° = –67 mV vs NHE) (13). The reduction potential for a His-bound heme in aqueous solution is substantially lower (E° = –134 mV vs NHE) (13). Although the GuHCl unfolding midpoints have shifted, there is still a [GuHCl] window in

Protein	**Fe^{III}**		**Fe^{II}**	
	$[GuHCl]_{½}$, M	$-\Delta G_f°$, kJ mol^{-1}	$[GuHCl]_{½}$, M	$-\Delta G_f°$, kJ mol^{-1}
WT	1.9(1)	33(3)	2.9(1)	40(4)
Q1A	1.9(1)	33(5)	2.9(1)	38(4)
Q1A/M15L	1.7(1)	28(5)	2.8(1)	37(1)
Q1A/M25A	1.1(1)	21(1)	2.1(1)	28(1)

Table 4.1. Folding thermodynamics values extracted from two-state fits of unfolding titration curves.

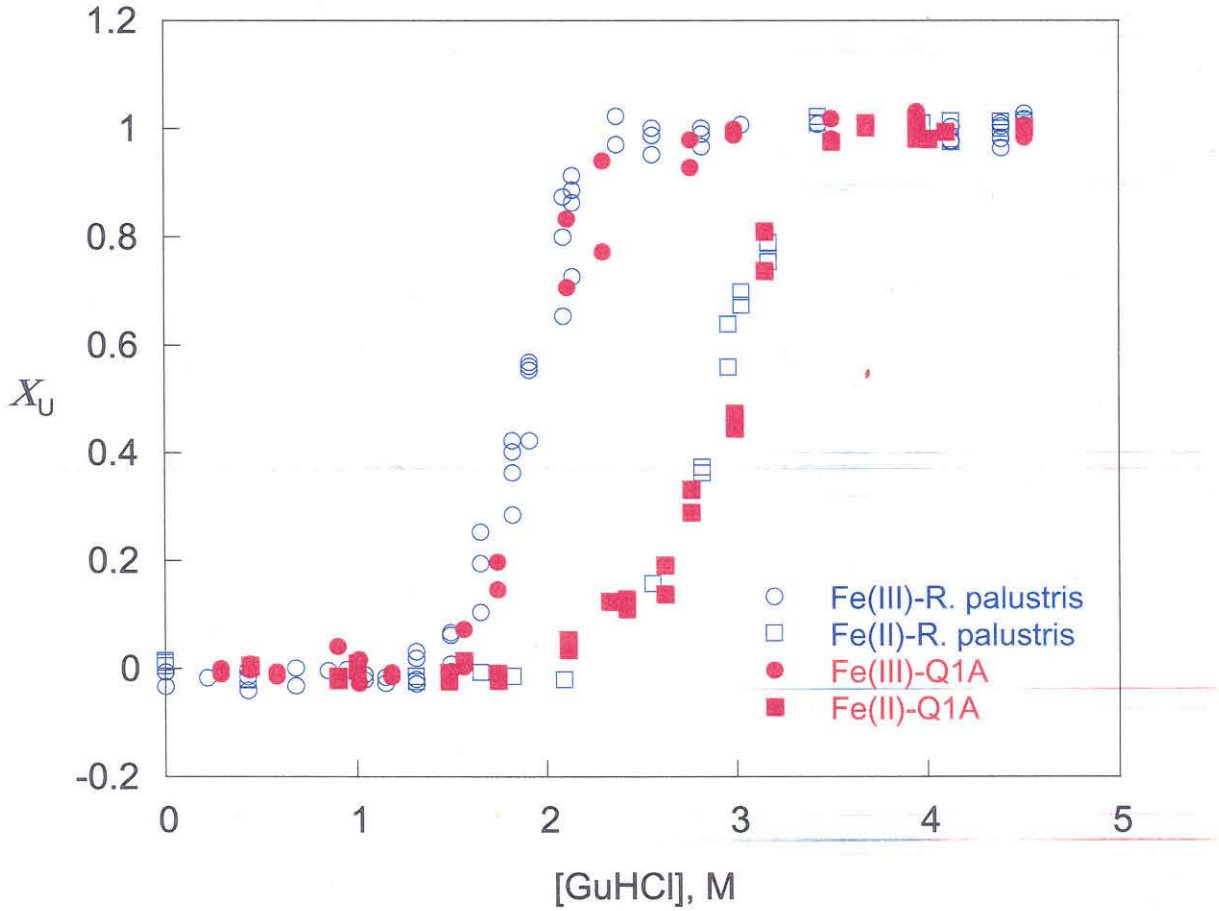


Figure 4.5. Equilibrium GuHCl unfolding curves for oxidized (circles) and reduced (squares) cyt *c'* (*R. palustris,* blue) and Q1A cyt *c'* (expressed in *E. coli*, red).

which ET triggering of Fe^{II}–cyt *c'* folding is possible(Q1A/M15L: 1.8 to 2.7 M GuHCl ; Q1A/M25A: 1.2 to 2.1 M GuHCl) (Figure 4.6). Unfortunately, the double Met-mutant (Q1A/M15L/M25A) did not express in high levels (< 1mg/ L) and has been difficult to purify. Consequently, folding studies were limited to the single Met mutants (Q1A/M15L and Q1A/M25A).

Ligand binding in Met-cyt c' mutants. Removal of just one methionine residue did not change the cyt *c'* absorption or CD spectrum, indicating that the heme coordination environments and secondary structures of the folded mutant proteins are unchanged from those of the native protein (Figure 4.7). Upon GuHCl unfolding, the Soret bands of the oxidized Met-mutants (Q1A/M15L and Q1A/M25A) behave similarly to that of the Q1A protein: the bands narrow and increase in intensity (Figure 4.8). Unfolding the single Met mutants of Fe^{II}–cyt *c'* blue shifts the Soret absorption band and two sharper Q-bands appear, consistent with the formation of low-spin, six-coordinate species (Figure 4.9). These spectral changes are reminiscent of those observed for the wild-type protein. However, in the triple mutant (without any methionine residues, Q1A/M15L/M25A), the absorption changes upon denaturation of the reduced state are modest: the Soret band is slightly broadened, red-shifted, and less intense; and the two Q-bands are also broadened and exhibit smaller extinction coefficients (Figure 4.10). Although unlikely at pH 7, it is also possible that there are other nonnative ligands (*e.g.*, lysine residues or the N-terminus (14)) that are coordinated to the Fe^{II}–center in unfolded cyt *c'*. In addition, water could be the

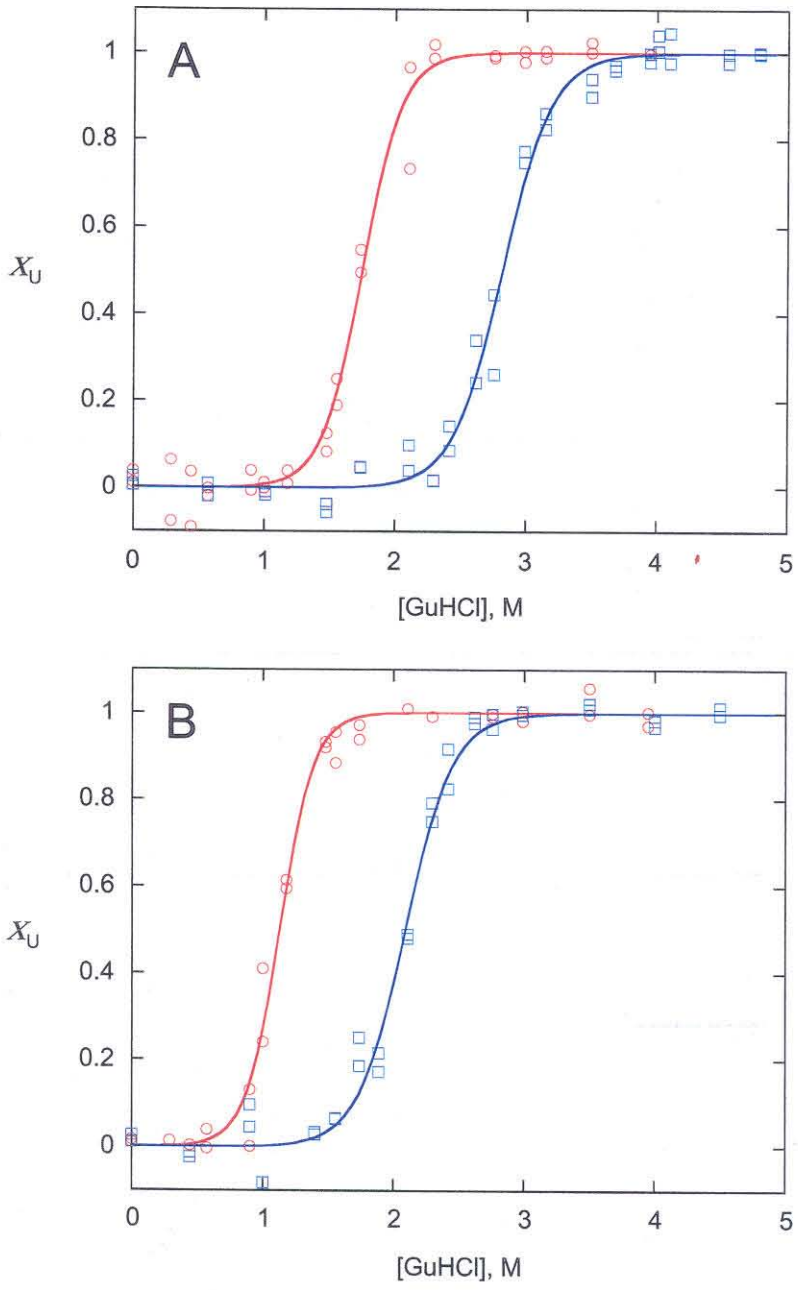


Figure 4.6. Denaturation curves for Fe^{III}- and Fe^{II}-cyt *c'* methionine mutants. (A) Q1A/M15L; (B) Q1A/M25A.

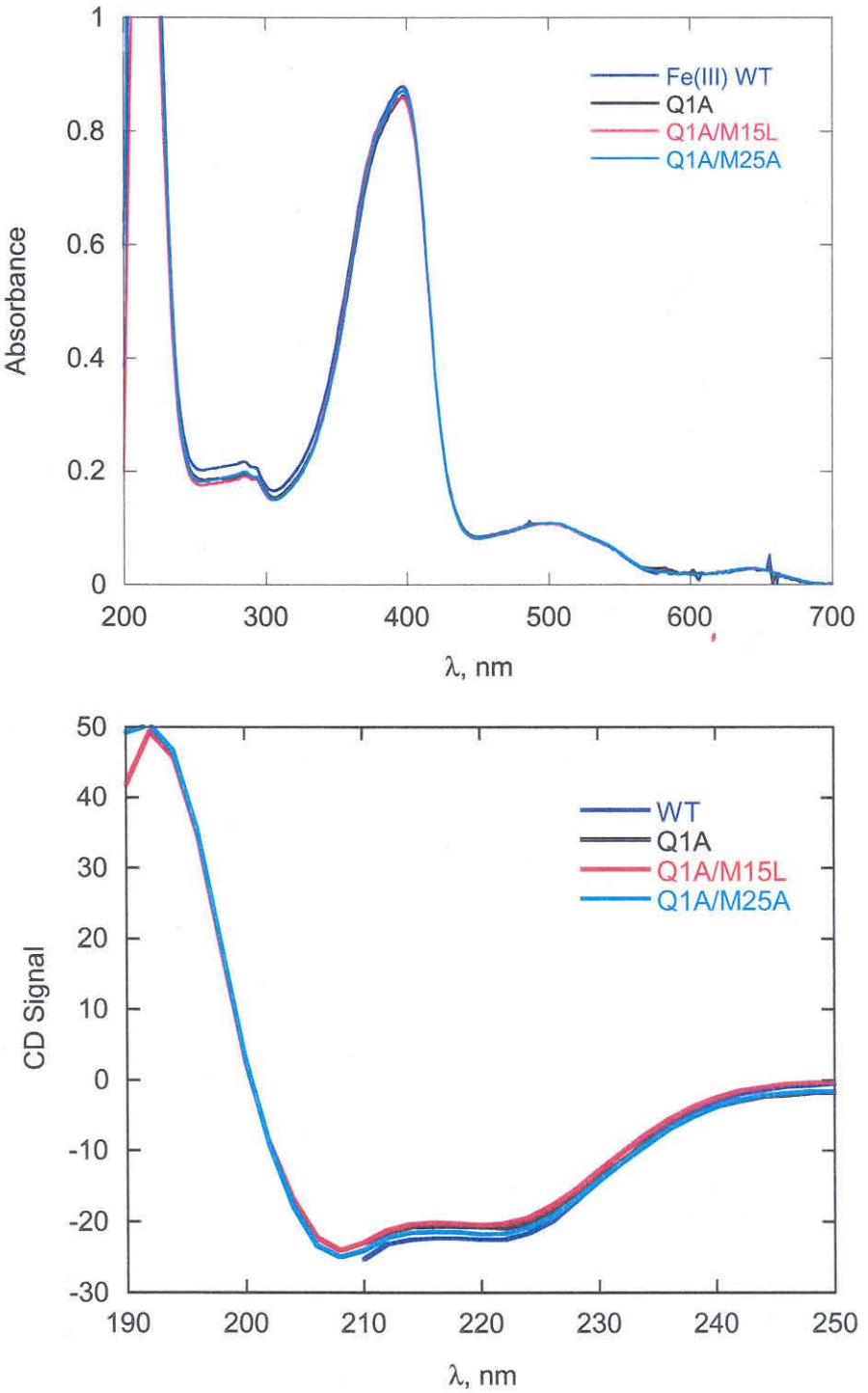


Figure 4.7. UV-Vis absorption and far-UV CD spectra of folded oxidized cyt *c*' proteins (*R. palustris* (blue); Q1A (black); Q1A/M15L (magenta); Q1A/M25A (cyan)).

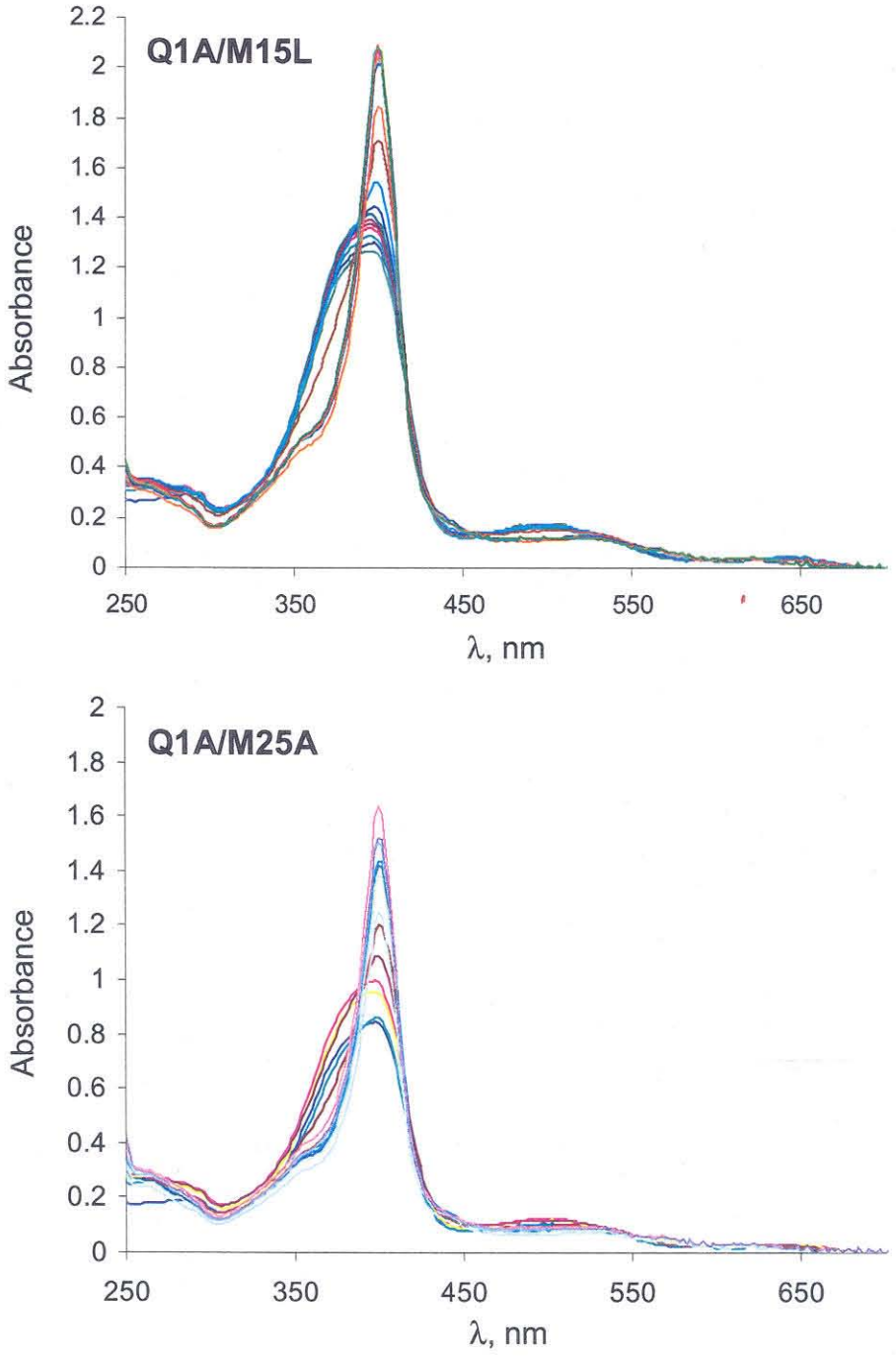


Figure 4.8. Changes in UV-Vis absorption spectra upon unfolding of oxidized single Met cyt *c*' proteins (Q1A/M15L (top panel); Q1A/M25A (bottom panel)).

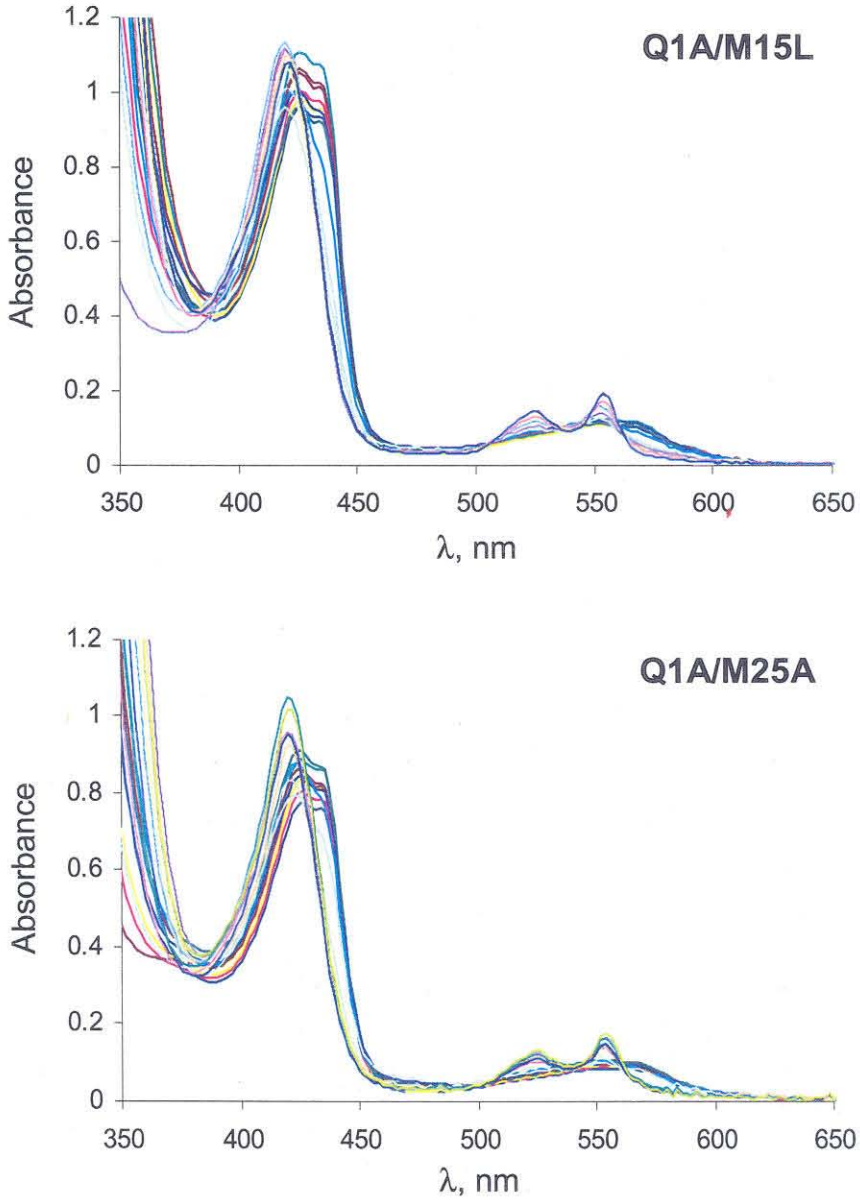


Figure 4.9. Changes in UV-Vis absorption spectra upon unfolding of reduced single Met cyt *c*' proteins (Q1A/M15L (top panel); Q1A/M25A (bottom panel)).

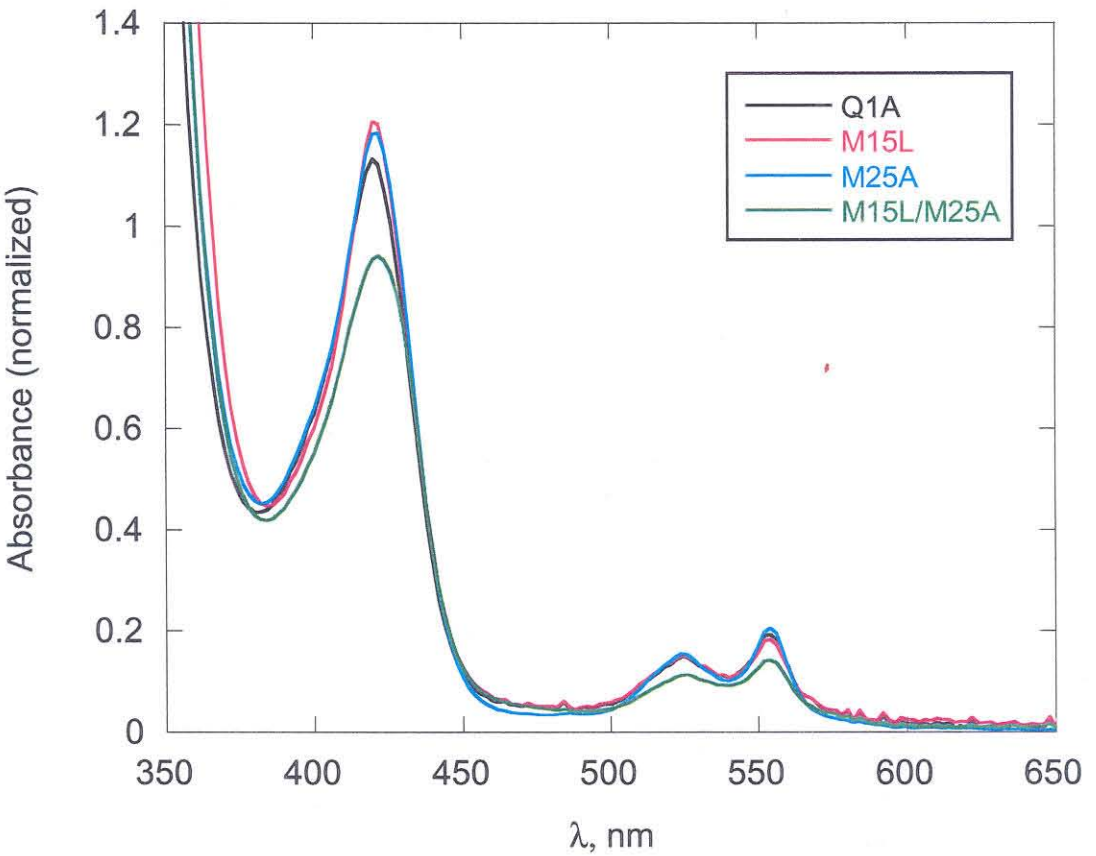


Figure 4.10. Comparison of absorption spectra of unfolded reduced cyt *c*' mutants (Q1A (black); Q1A/M15L (magenta); Q1A/M25A (cyan); Q1A/M15L/M25A (green), in 4.5 M GuHCl, pH 7).

sixth ligand in some fraction of unfolded Q1A/M15L/M25A Fe^{II}–cyt *c*' molecules. Similar coordination modes have been proposed in the case of the heme octapeptide from cyt *c* (MP8): the absorption spectrum of Fe^{II}-AcMP8 is consistent with an equilibrium mixture of high- and low-spin hemes, where the high-spin form is five-coordinate and the low-spin species is six-coordinate with H_2O as the sixth ligand (15). The differences between the absorption spectrum of unfolded Q1A/M15L/M25A cyt *c*' and those of the wild-type and single Met mutants are similar to those observed when Fe^{II}-AcMP8 is titrated with methionine (16).

Spectroscopic studies of unfolded methionine mutants gave ambiguous results. It is not clear whether methionine residues are coordinated in the reduced unfolded protein. In order to address this issue, we labeled Cys-17 in a mutant (Q1A/E17C) with a dansyl fluorophore. If either Met-15 or Met-25 is bound in the reduced unfolded state, the dansyl group should be near the heme and fluorescence will be quenched by energy transfer. If no nonnative residues on the first helix are ligated to the Fe^{II}–center, the dansyl fluorescence decay kinetics should be similar to that of the unfolded (and unligated) Fe^{III}–protein.

In both oxidation states of the folded protein, the dansyl excited state is short-lived, owing to rapid energy transfer to the heme. The decay kinetics for the Fe^{III}–and Fe^{II} proteins are virtually identical (Figure 4.11). The kinetics are biphasic with a dominant fast (75%, ~3×10^9 s^{-1}) population and a rather substantial population of slow (25%, ~10^8 s^{-1}) components. The latter may arise from contamination with unfolded polypeptides or unbound dye (Figure 4.12). An

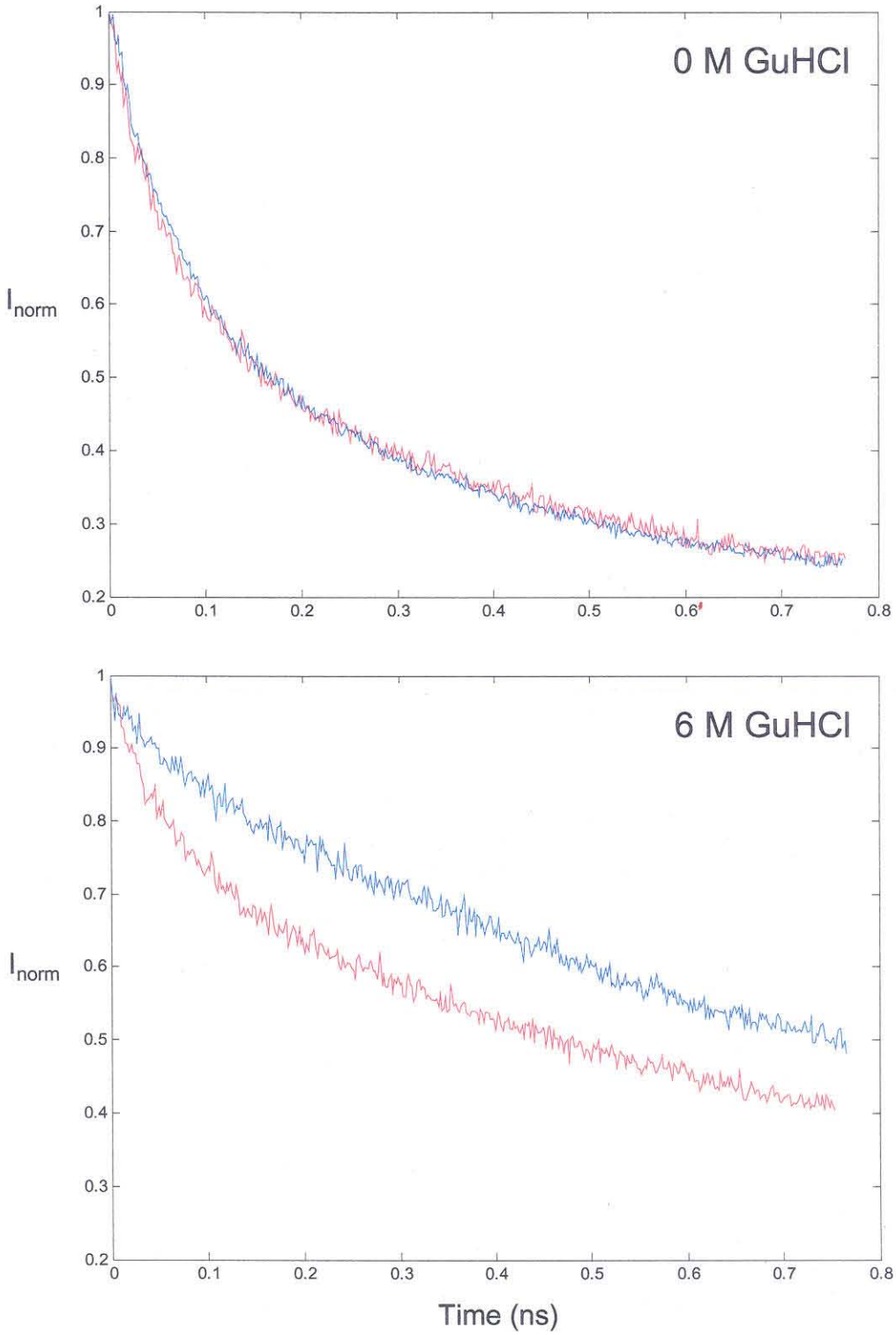


Figure 4.11. FET kinetics of folded (top panel) and unfolded (bottom panel) Fe^{III}- (blue line) and Fe^{II}- (red line) dansyl-labeled Q1A/E17C cyt *c*' (λ_{ex} = 290 nm; $\lambda_{obsd} \geq$ 475 nm).

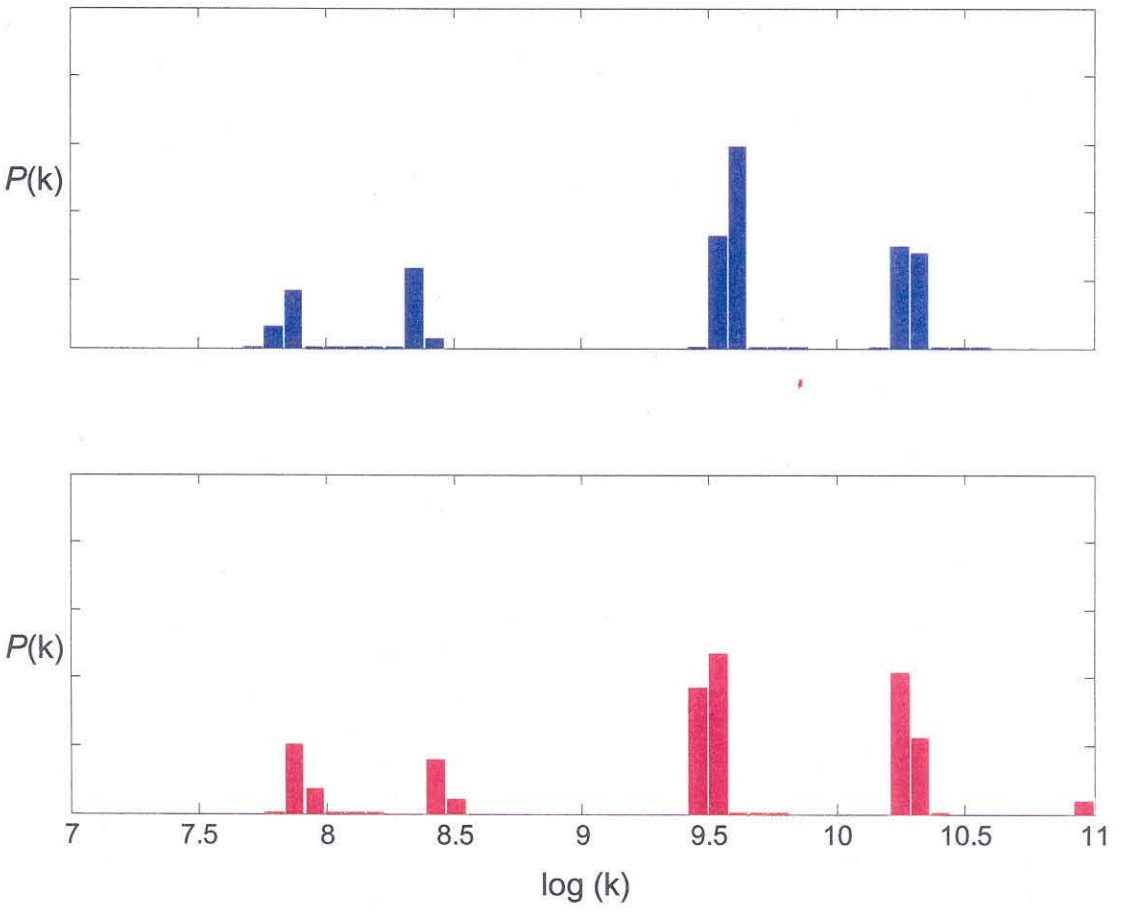


Figure 4.12. Probability distributions of FET rate constants extracted from LSQNONNEG fits of folded dansyl-labeled Q1A/E17C cyt *c'* in 100 mM sodium phosphate buffer, pH 7(Fe^{III} (blue); Fe^{II} (red)).

intriguing alternative possibility is that these nanosecond timescale snapshots reveal large amplitude fluctuations in the protein structure.

Addition of GuHCl to oxidized and reduced cyt *c'* leads to a significant increase in dansyl fluorescence. The fraction of slow-decay components increases (Fe^{III}: 60%; Fe^{II}: 40%), but, in unfolded Fe^{II}–cyt *c'*, there is a noticeable population of fast-decay components, suggesting that residue 17 is closer to the heme in the reduced state (Figure 4.13). It is important to note that in both oxidation states there are decay components attributable to collapsed structures ($\sim 10^9\ s^{-1}$), even at very high denaturant concentration (6 M). Thus, although the CD spectrum is consistent with complete disruption of secondary structure at high [GuHCl], it is apparent that there are significant populations of collapsed conformations in the denatured protein.

Single Met Fe^{II}–cyt c' folding. We have compared the kinetics of Fe^{II}–cyt *c'* folding in the Q1A/M15L and Q1A/M25A mutants. The GuHCl concentrations were adjusted to afford a constant folding free energy (Q1A: [GuHCl] = 2.2 M, $\Delta G_f° = -9$ kJ mol $^{-1}$; Q1A/M15L: [GuHCl] = 2.2 M, $\Delta G_f° = -8$ kJ mol $^{-1}$; Q1A/M25A: [GuHCl] = 1.6 M, $\Delta G_f° = -6$ kJ mol $^{-1}$). Two-photon laser excitation of NADH was used to reduce the unfolded oxidized proteins (17). The progress of the folding reaction was monitored by transient Soret absorption spectroscopy from 10^{-4} to 10^1 s after excitation. The kinetics for Q1A, Q1A/M15L, and Q1A/M25A cyt *c'* are shown in Figure 4.14. The folding kinetics of pseudo-wild-type Q1A cyt *c'* are heterogeneous, similar to the behavior of native *R. palustris*

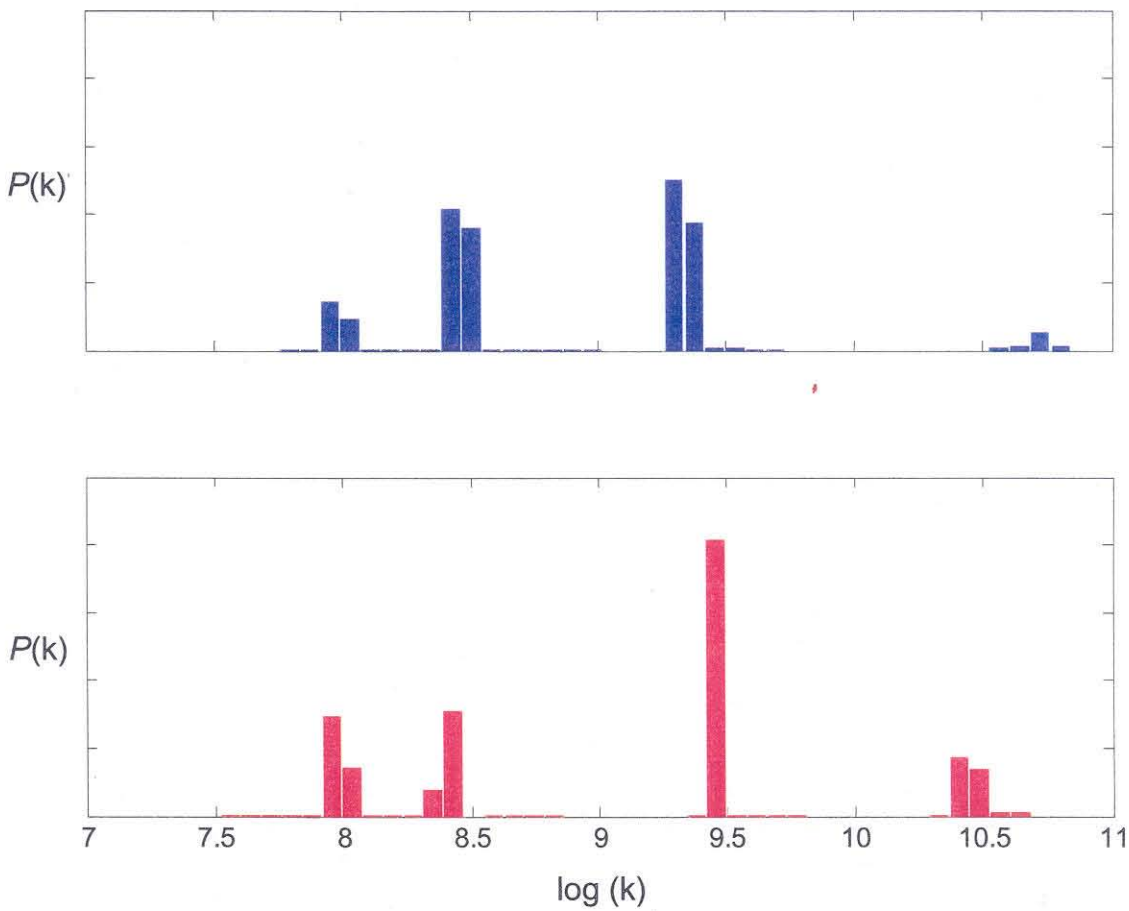


Figure 4.13. Probability distributions of FET rate constants extracted from LSQNONNEG fits of unfolded dansyl-labeled Q1A/E17C cyt *c'* in 6 M GuHCl and 100 mM sodium phosphate buffer, pH 7(Fe^{III} (blue); Fe^{II} (red)).

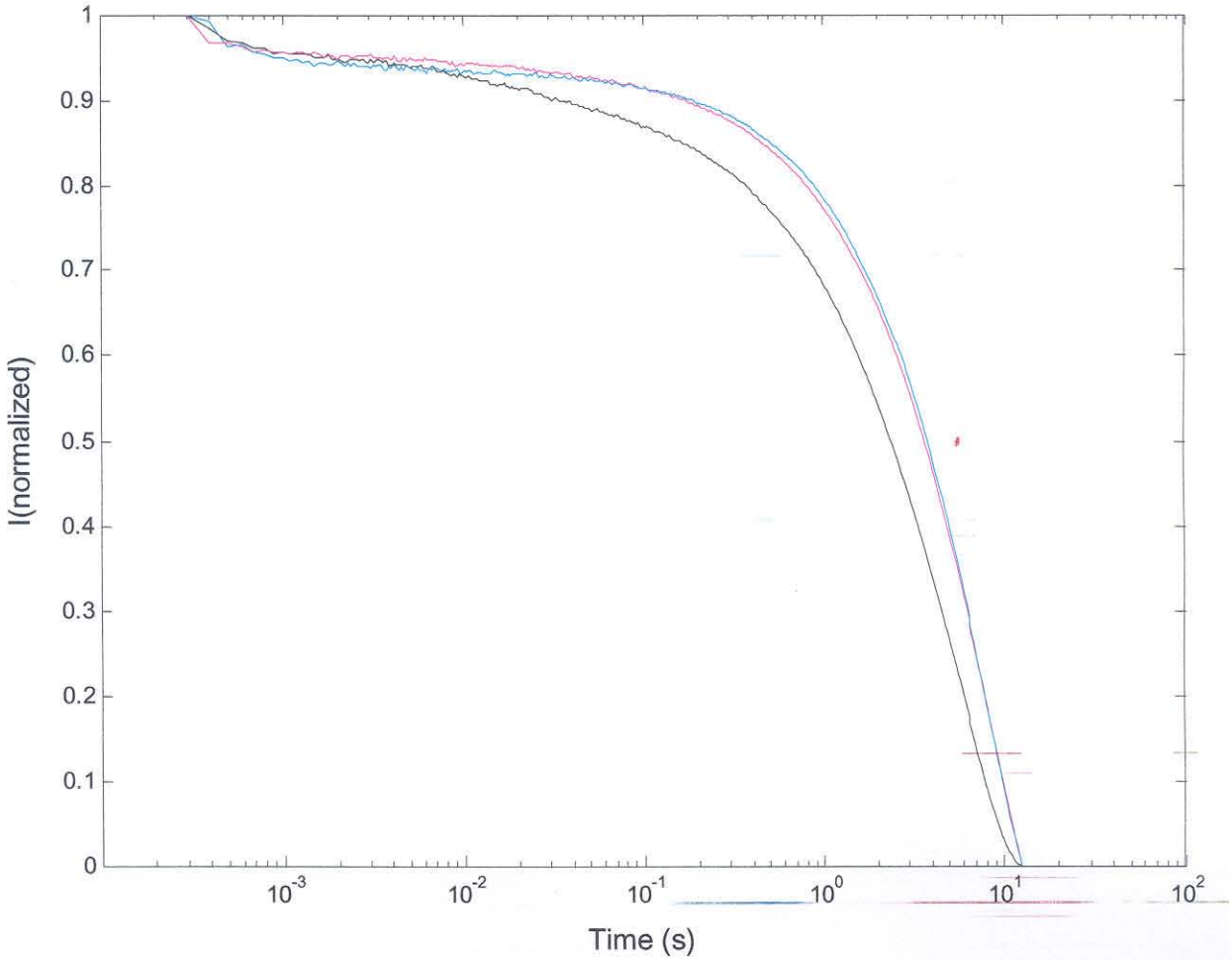


Figure 4.14 . Normalized folding kinetics (observed at 440 nm in 0.1 M sodium phosphate buffer, pH 7) at comparable folding free energies of Fe^{II}–Q1A cyt *c*' (black line; 2.2 M GuHCl, $\Delta G_f^\circ = -9$ kJ mol $^{-1}$), Fe^{II}–Q1A/M15L cyt *c*' (magenta line; 2.2 M GuHCl, $\Delta G_f^\circ = -8$ kJ mol $^{-1}$), and Fe^{II}–Q1A/M25A cyt *c*' (cyan line; 1.6 M GuHCl, $\Delta G_f^\circ = -6$ kJ mol $^{-1}$).

cyt *c*' (1). There are fast- (~10^4 s^{-1}), intermediate- (~10^2 – 10^1 s^{-1}), and slow-folding (~10^{-1} s^{-1}) populations (Figure 4.15). One striking feature of these kinetics is that there are substantially more (>20%) of the slow-folding populations than found for the wild-type protein. We have attributed this difference to the lower concentration of dissolved oxygen in the samples. Reduced unfolded cyt *c*' is readily oxidized by O_2, a reaction that interrupts protein folding.

Removal of one of the methionine residues does not have a substantial effect on either the very fast- or slow-folding components. Rather, the intermediate-folding group is most influenced by the methionine mutations. Indeed, the non-negative least-squares fits reveal that these Met residues contribute to the heterogeneity of cyt *c*' folding (Figure 4.15). One interpretation of the kinetics is that Met-ligated structures facilitate cyt *c*' folding; the removal of Met residues eliminates the intermediate folding rates. In view of the fact that Fe^{III}–cyt *c*' and Fe^{II}(CO)–cyt *c*' fold on the 10^0-10^1 s timescale, it seems unlikely that ligation by either Met residue retards folding. It is important to note also that the methionine-misligated structures are minor components of the cyt *c*' folding ensemble. Cyt *c*' folding is dominated by processes that occur on a seconds timescale.

Discussion

Cytochrome folding is complicated and often retarded by heme misligation (18-20). Ligand substitution processes at the heme strongly influence peptide

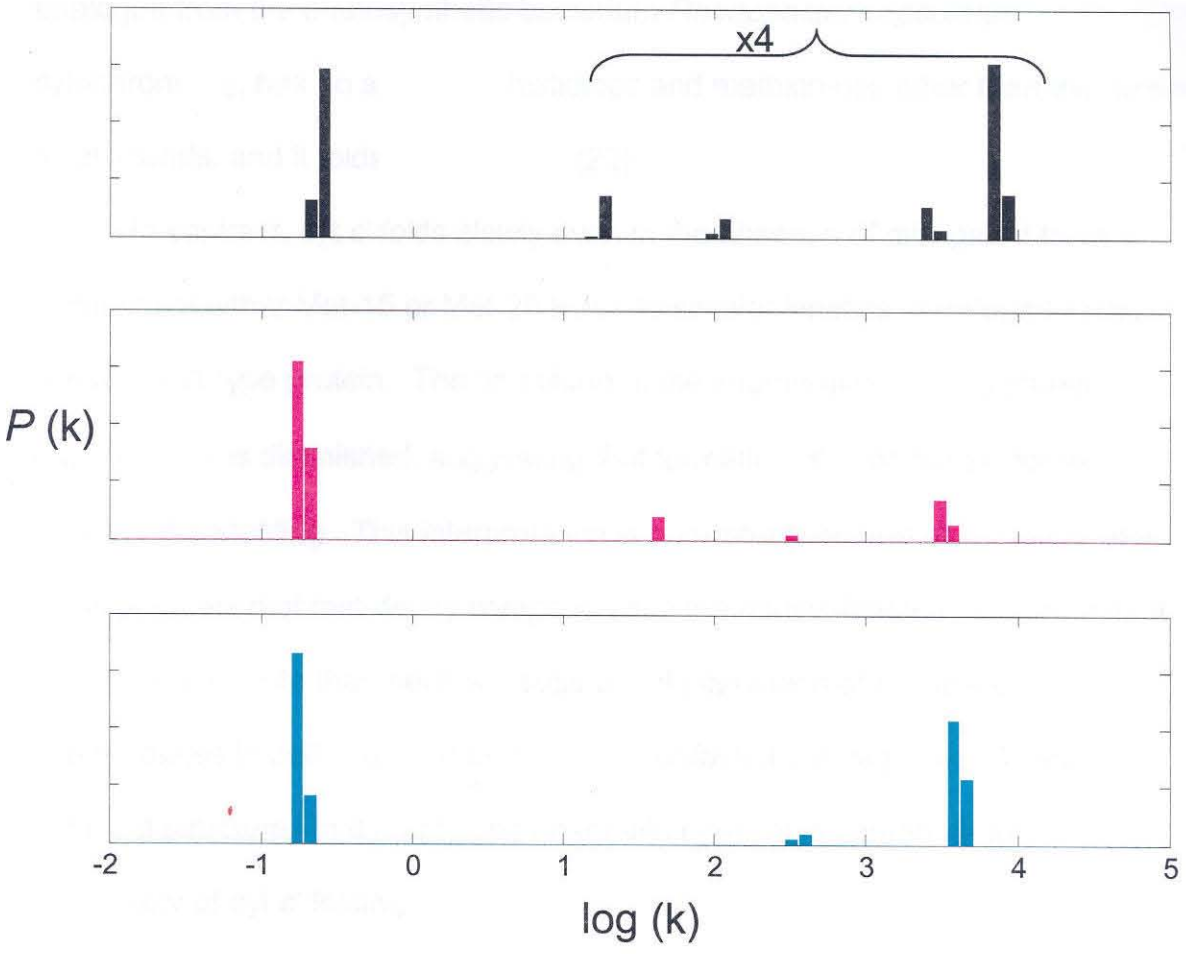


Figure 4.15. Probability distributions of folding rate constants extracted from LSQNONNEG fits of Q1A (black), Q1A/M15L (magenta), and Q1A/M25A (cyan) Fe^{II}–cyt *c*' folding kinetics. The amplitudes of rate constants from the range 10^4 to 10^1 s^{-1} have been magnified (x4) to illustrate differences in the mutants in that region.

backbone dynamics. During the folding of equine cytochrome *c* at neutral pH, correction of His-misligation requires hundreds of milliseconds. When His-33 is removed by site-directed mutagenesis or displaced from the heme by high concentrations of imidazole, folding is accelerated (4-7, 21). A structural analogue from the photosynthetic bacterium *Rhodobacter capsulatus*, cytochrome c_2, has no additional histidines and methionines other than the native axial ligands, and it folds quite rapidly (22).

In contrast, cyt *c'* folds slowly even in the absence of misligated traps. Removal of either Met-15 or Met-25 leads to simpler kinetics than found in the pseudo-wild-type protein. The amplitude of the intermediate folding phase (10^2–10^1 s^{-1}) is diminished, suggesting that formation of near-native contacts may facilitate folding. This interpretation is consistent with the observation of a small population of fast-decay components in the dansyl-labeled Fe^{II}–protein. It is important to note that there is a substantial population of collapsed intermediates in both oxidized and reduced unfolded polypeptides. These compact structures in the unfolded ensemble could also contribute to the complexity of cyt *c'* folding.

The role of Met-misligated structures in cyt *c'* folding remains a mystery. The available data suggest that misligated hemes are not responsible for slow folding. Indeed, measurements on the single Met mutants point to accelerated folding as a result of forming near-native contacts in the denatured Fe-Met protein. Favorable metal/amino-acid ligation could reduce the size of the conformational space for an unfolded polypeptide and lower the barrier to folding.

But, because Met residues only bind to a fraction of the ferrohemes in cyt *c'*, these intermediates constitute a small population (~10%) of the molecules in ET-triggered folding. Folding of the remaining fraction may be frustrated by transient trapping in nonnative compact structures. From our work, it is clear that cyt *c'* folding is controlled by polypeptide dynamics rather than by Met-Fe substitution processes.

References

1. Lee, J. C., Gray, H. B. and Winkler, J. R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7760-7764.
2. Thirumalai, D., Klimov, D. K. and Woodson, S. A. (1997) *Theor. Chem. Acc.* **96,** 14-22.
3. Thirumalai, D. and Klimov, D. K. (1999) *Curr. Opin. Struct. Biol.* **9,** 197-207.
4. Elöve, G., Bhuyan, A. K. and Roder, H. (1994) *Biochemistry* **33,** 6925-6935.
5. Sosnick, T. R., Mayne, L. and Englander, S. W. (1996) *Proteins: Struct. Func. Gen.* **24,** 413-426.
6. Yeh, S.-R., Takahashi, S., Fan, B. and Rousseau, D. (1997) *Nature Struct. Biol.* **4,** 51-56.
7. Colón, W., Wakem, L. P., Sherman, F. and Roder, H. (1997) *Biochemistry* **36,** 12535-12541.

8. McGuirl, M. M., Lee, J. C., Lyubovitsky, J. G.,Thanyakoop, C., Winkler, J. R.and Gray, H. B., manuscript in preparation.

9. Arslan, E., Schulz, H., Zufferey, R., Künzler, P. and Thöny-Meyer, L. (1998) *Biochem. Biophys. Res. Commun.* **251,** 744-747.

10. Pace, N. C., Shirley, B. A. and Thomson, J. A. (1990) in *Protein Structure: A Practical Approach*, ed. Creighton, T. F. (IRL Press, Oxford), pp. 311-330.

11. Nozaki, Y. (1972) *Meth. Enzymol.* **26,** 43-50.

12. Stowell, M. H. B., Larsen, R. W., Winkler, J. R., Rees, D. C. and Chan, S. I. (1993) *J. Phys. Chem.* **97,** 3054-3057.

13. Battistuzzi, G. B., M.; Cowan, J. A.; Ranieri, A.; Sola, M. (2002) *J. Am. Chem. Soc.***124**, 5315-5324.

14. Hammock, B., Godbole, S. and Bowler, B. E. (1998) *J. Mol. Biol.* **275,** 719-724.

15. Tezcan, F. A. (2001) in *Chemistry* (California Institute of Technology, Pasadena).

16. Tezcan, F. A., Winkler, J. R. and Gray, H. B. (1998) *J. Am. Chem. Soc.* **120,** 13383-13388.

17. Orii, Y. (1993) *Biochemistry* **32,** 11910-11914.

18. Babul, J. and Stellwagen, E. (1972) *Biochemistry* **11,** 1195-1200.

19. Muthukrishnan, K. and Nall, B. T. (1991) *Biochemistry* **30,** 4706-4710.

20. Godbole, S. and Bowler, B. E. (1997) *J. Mol. Biol.* **268,** 816-821.

21. Telford, J. R., Tezcan, F. A., Gray, H. B. and Winkler, J. R. (1999) *Biochemistry* **38,** 1944-1949.

22. Sauder, J. M., McKenzie, N. E. and Roder, H. (1996) *Biochemistry* **35,** 16852-16862.

Chapter 5

Cytochrome *c*' folding probed by fluorescence energy-transfer kinetics[†]

[†]Work on Q1A and Q1A/F32W/W72F cyt *c*' folding was done in collaboration with K. Cecilia Engman. Crystal structure of Q1A/F32W/W72F cyt *c*' was determined by Dr. F. Akif Tezcan.

Introduction

Recent studies employing electron- (1) and energy-transfer (2) kinetics have revealed structural heterogeneity in cytochrome *c* (cyt *c*) folding that is not apparent when other spectroscopic probes are used. Fluorescence energy-transfer (FET) kinetics measured with dansyl-labeled cyt *c* show that two rapidly equilibrating populations of collapsed (average dansyl-heme distance, r_{avg} = 27 Å) and extended (r_{avg} > 50 Å) structures are populated throughout the folding process (2). Similarly, measurements of electron-transfer (ET) kinetics during Zn-cyt *c* folding reveal a burst intermediate in which one-third of the ensemble has a protected Zn-porphyrin and slow ET kinetics; the remaining fraction shows fast ET characteristic of a solvent-exposed redox cofactor (1). These findings from FET and ET experiments suggest that compact nonnative structures are not substantially more stable than fully extended peptides.

Shakhnovich and coworkers have examined the possibility that global hydrophobic collapse is not an obligatory step in protein folding (3). By modifying the inter-residue potential of a lattice-model protein, two limiting folding regimes could be examined: the so-called attractive case in which a collapsed intermediate forms, and the repulsive case in which folding proceeds without global collapse. When correlated with the primary amino-acid sequence and overall stability, it appears that the more hydrophobic and less optimized sequences tend to have a higher propensity to form compact structures, whereas the less hydrophobic and strongly optimized sequences lead to repulsive interactions that disfavor compaction (3).

The surprising observation of rapidly equilibrating extended and compact nonnative cyt *c* molecules demands further investigations of the energetics and dynamics of folding intermediates. One possibility that should be explored is that nonnative compact conformations have highly hydrated cores, as recent theoretical studies support the notion that significant numbers of water molecules can be trapped during formation of collapsed structures (4, 5). Of course, such hydration may not be general, and the equilibrium between collapsed and extended conformations could be a peculiar property of cyt *c*. It is reasonable to suggest that different protein folds and sequences will modulate the balance between compact and extended structures. Indeed, the relative energetics of collapsed intermediates and extended polypeptides can have important kinetics consequences. FET kinetics studies of other cytochromes are necessary to determine if this is a common protein folding characteristic and how the equilibrium between compact and extended structures affects the overall folding dynamics.

The slow and highly heterogeneous folding kinetics of ferrocytochrome *c*' (cyt *c*') from *R. palustris* are unusual for a four-helix bundle (6). Many helical bundles are fast-folding, owing to their highly symmetrical structures and preponderance of short-range contacts (7–9). Our investigations of cyt *c*' thus far, which have employed primarily ET triggering and exclusively heme absorption probes, have indicated that either Met-15 or Met-25 binds to the unfolded ferroheme (6). Further, the mutagenesis results have suggested that the faster folding phases may arise from polypeptides containing Met-ligation

(Chapter 4). In the absence of misligation, Fe^{III}–cyt *c*' refolding can be described adequately by a single kinetics phase on a seconds timescale. Our measurements have revealed the timescale required for proper heme coordination, but have provided little information about the polypeptide dynamics. In order to address this issue, we have extended these studies to include use of FET probes to gain insight into polypeptide dynamics that accompany Fe^{III}–cyt *c*' folding. (Recall that Fe^{III}–cyt *c*' does not contain a protein-derived sixth ligand.) Cyt *c*' has a wild-type Trp72 residue that serves as a natural fluorophore for FET kinetics measurements (Figure 5.1) (10). In addition, we have prepared and crystallographically characterized a position-32 mutant, Q1A/F32W/W72F, in order to probe distances between another site on the polypeptide chain (Trp-32) and the heme during folding.

Materials and Methods

Guanidine hydrochloride (GuHCl, Sigma, ultrapure grade) and L-tryptophan (Sigma) were used as received. The *R. palustris* cyt *c*' gene was cloned (plasmid cytcp1) and the protein was heterologously expressed (with plasmid pEC86, kindly provided by Dr. Linda Thöny-Meyer (11)) in *E. coli* (strain BL21(DE3)). Mutant protein (Q1A/F32W/W72F) was generated by site-directed mutagenesis using a Quickchange kit (Stratagene). DNA sequencing was performed by the Caltech DNA Sequencing Core Facility.

Q1A/F32W/W72F cyt *c*' was expressed and purified according to procedures described in Appendix A. Mass spectral and protein sequence

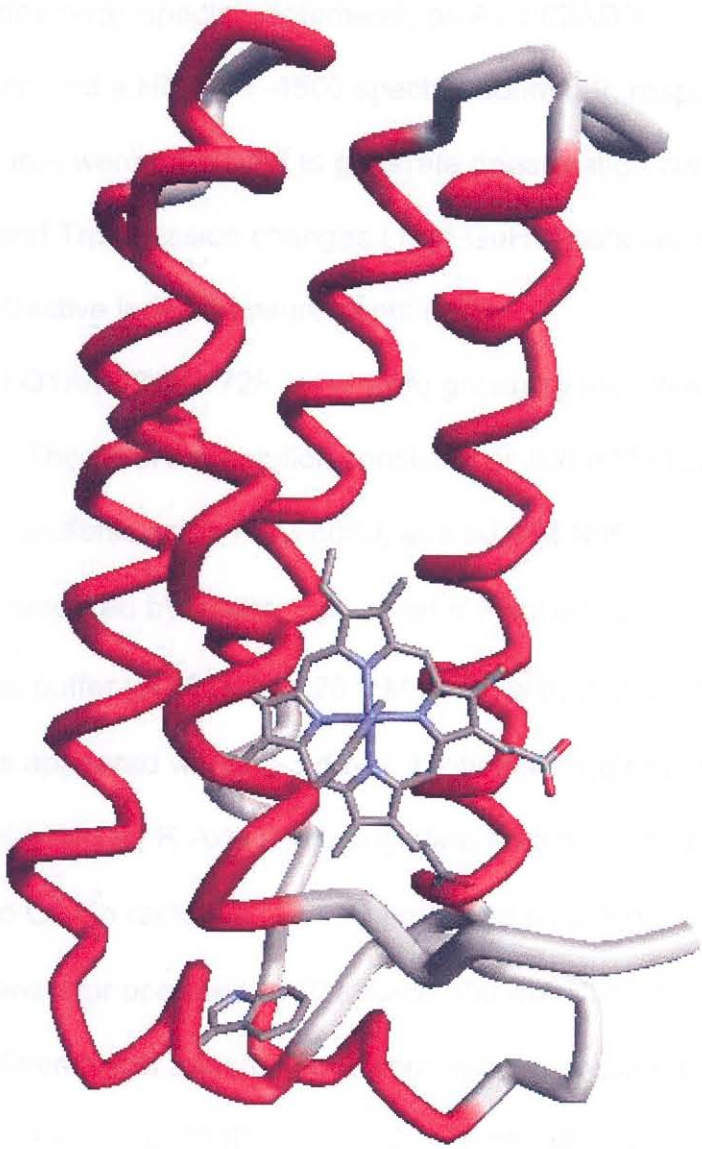


Figure 5.1. Structure of *R. palustris* cyt *c'*. Trp-72 is highlighted.

analyses were conducted by the Protein/Peptide Microanalytical Laboratory in the Beckman Institute at Caltech (results are given in Appendix A).

Absorption, CD, and luminescence spectra were measured on a Hewlett-Packard 8452 diode array spectrophotometer, an Aviv 62ADS spectropolarimeter, and a Hitachi F-4500 spectrofluorimeter, respectively. Standard procedures were employed to generate denaturation curves from heme absorption, CD, and Trp emission changes (12). GuHCl concentrations were determined by refractive index measurements (13).

Crystals of Q1A/F32W/W72F cyt *c'* were grown by the sitting-drop vapor diffusion method. The reservoir solution consisted of 100 mM MES (2-[N-morpholino]ethanesulfonic acid), PEG 6000, and 50 mM NaCl (pH 6). The protein drop was prepared by mixing 2 μL of cyt *c'* solution (850 μM in 10 mM sodium phosphate buffer (pH 6.0) and 28 mM NaCl) with 2 μL of the reservoir solution. Crystals appeared within 2-3 days. X-ray diffraction data were collected at 100 K using an R-Axis II imaging plate area detector and monochromatized Cu-Kα radiation (1.54 Å) produced by a Rigaku RU 200 rotating anode generator operated at 50 kV and 100 mA. The structure was determined by refinement of a model from isomorphous crystals (space group $P2_12_12_1$) of *R. palustris* cyt *c'* (PDB:1A7V) (10) against diffraction data using DENZO (Table 5.1) (14). Rigid-body, simulated-annealing, positional and thermal refinement with CNS (15), amid rounds of manual rebuilding, and water placement with XFIT (16), produced the final model. The Ramachandran plot was calculated with PROCHECK (17).

Statistic	Value
Residues	125 + 1 Heme (last 2 residues deleted)
Waters	262
Unit cell dimensions (Å)	33.83 x 62.09 x 113.42
No. of monomers/ unit cell	2
Symmetry group	$P2_12_12_1$
Resolution (Å)	1.78
X-ray wavelength (Å)	1.54
No. of measurements	150461
No. of unique reflections	23685
Completeness (%)	99.7 (99.0)
$\langle I/\sigma I \rangle^{\dagger}$	24.8 (2.0)
R_{sym} (%)	6.2 (49)
R (%)	21.7 (30.7)
Free R (%)	26.6 (38.3)
Rms Bnd (Å)	0.017
Rms Ang (°)	1.79
Average temperature factor (Å)	
All protein	34.91
Main chain atoms	31.29
Side chain atoms	37.95
Heme atoms	26.98
Water molecules	41.67
Ramachandran plot (%)	
Residues in most favored regions	92.6
Residues in additional allowed regions	7.0
Residues in generously allowed regions	0.5
Residues in disallowed regions	0.0

Table 5.1. X-ray diffraction data collection and refinement statistics for Q1A/F32W/W72F cyt *c*'. †intensity signal to noise ratio. $R_{sym} = \Sigma\Sigma_j |I_j - \langle I \rangle| / \Sigma\Sigma_j |I_j|$. $R = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$ for all reflections. Free R calculated against 8% of the reflections removed at random.

Trp fluorescence decay kinetics at varying concentrations of GuHCl were measured by using the third harmonic of a regeneratively amplified femtosecond Ti:sapphire laser at 290 nm as the excitation source and a picosecond streak

camera (Hamamatsu C5680) as the detector (Figure 5.2). An interference filter was used to select Trp emission (λ = 355±5 nm). The kinetics data were fit using a MATLAB (The Mathworks, Inc.) algorithm (LSQNONNEG) that minimizes the sum of the squared deviations (χ^2) between observed and calculated values of *I(t)*, subject to a nonnegativity constraint. LSQNONNEG produces the narrowest *P(k)* distributions with relatively few nonzero components. A simple coordinate transformation recasts the *P(k)* results obtained by LSQNONNEG fitting as probability distributions over *r* (Equation 5.1) (18, 19). The value of J for the Trp-heme donor-acceptor pair was determined on the basis of several measurements of both Trp fluorescence and cyt *c'* absorption, at varying concentrations between 5 to 120 μM (Figure 5.3). The value of τ_D (3.5×10^{-9} s) was obtained from luminescence decay measurements with L-tryptophan.

$$r_o = (8.8 \times 10^{23} \kappa^2 n^{-4} \Phi_D J)^{1/6}$$

$$J = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 \, d\lambda$$

$$k_{FET} = \frac{1}{\tau_D}\left(1 + \left(\frac{r_o}{r}\right)^6\right)$$

Equation 5.1. Förster energy-transfer equations and parameters (r_o, critical distance, κ, orientation factor, n, index of refraction, Φ_D, quantum yield of donor, J, overlap integral between donor emission and acceptor absorption spectra, k_{FET}, rate of energy transfer).

The folding kinetics of Q1A and Q1A/F32W/W72F Fe^{III}–cyt *c'* were measured by using a BioLogic SFM-4S stopped-flow mixer, coupled via optical fiber to a monochromator fitted with a five-stage photomultiplier tube. For both

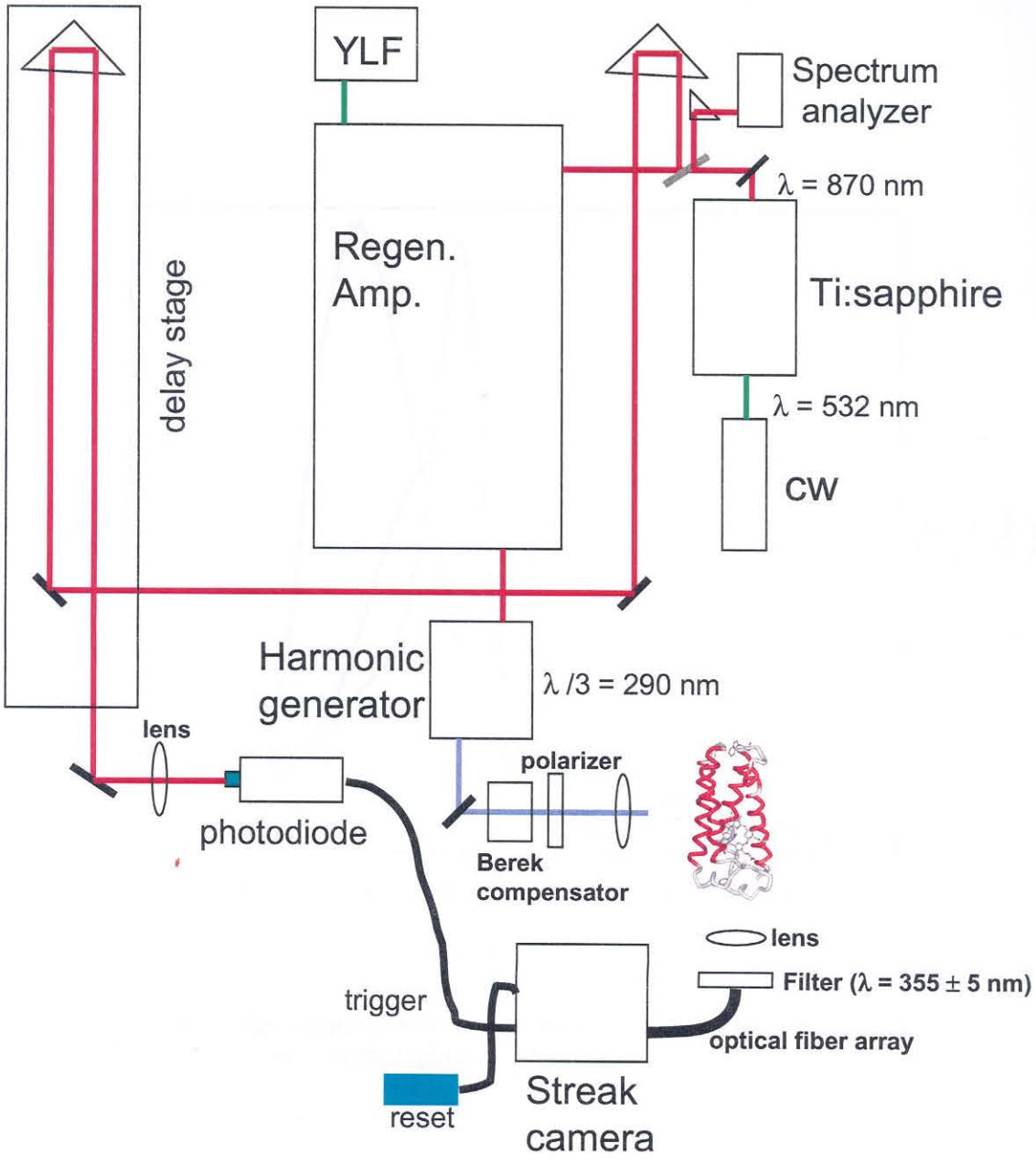


Figure 5.2. Illustration of the femtosecond Ti-sapphire laser/picosecond streak camera setup for fluorescence lifetime measurements (not drawn to scale).

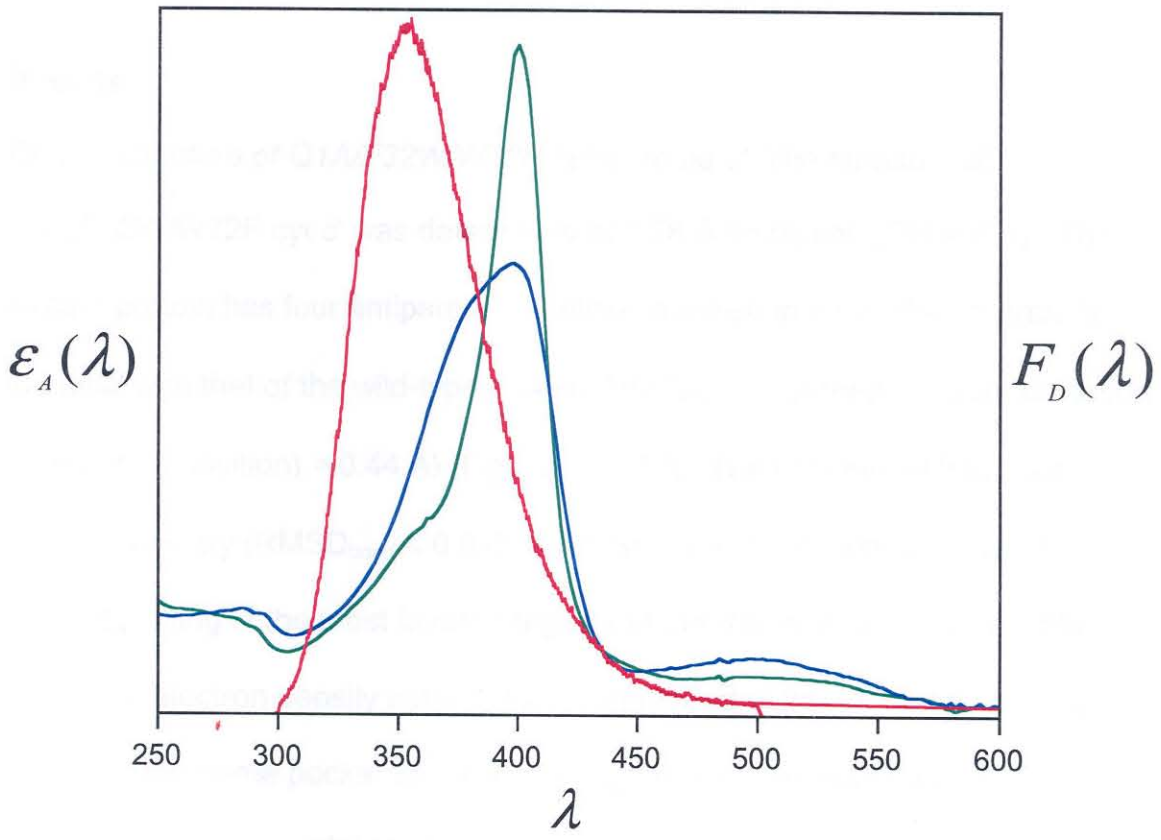


Figure 5.3. The spectral overlap of Trp emission (red) and heme absorption of folded (blue) and unfolded (green) Q1A cyt *c*'.

absorption- and fluorescence-monitored folding experiments, the probe light was provided by a 200 W Hg/Xe arc lamp (absorption: $\lambda_{probe} \geq 350$ nm; emission: 260 nm$\leq \lambda_{ex} \leq$ 355 nm). The PMT output was recorded with a 16-bit, 20 MS/s digital oscilloscope (CompuScope 1602, Gage Applied Science, Inc.). The stopped-flow mixing deadtime under these conditions was $\approx$ 5 ms.

Results

Crystal structure of Q1A/F32W/W72F cytochrome c'. The structure of Q1A/F32W/W72F cyt *c'* was determined at 1.78 Å resolution (Table 5.1). The mutant protein has four antiparallel α-helices bundled in a fold that is virtually identical with that of the wild-type protein ($RMSD_{C\alpha}$ (root-mean-square deviation in α-carbon position) = 0.44 Å) (Figure 5.4) (10). The final model has good stereochemistry ($RMSD_{bond} \leq 0.017$ Å, $RMSD_{angle} \leq 1.79°$) with most residues (92.6 %) falling in the most favored regions of the Ramachandran plot (Table 5.1). The electron density maps ($2F_o$-F_c; contoured at 2σ) of the two mutation sites and the heme pocket are shown in Figure 5.5. The heme iron is five-coordinate and the metal-to-His-117 distance is 2.00 Å. The vacant sixth coordination site is surrounded by Leu-12, Met-15, Leu-85, three Phe residues (Phe-55, Phe-78, and Phe-82), and the methylene chain of Arg-8 (Figure 5.6). There are minor perturbations in the hydrophobic heme pocket of the mutant protein, particularly in the Phe-82 and Leu-85 side chains. The replacement of Trp-72 with Phe-72 affects the hydrophobic interactions with helix-2 (Figure 5.4).

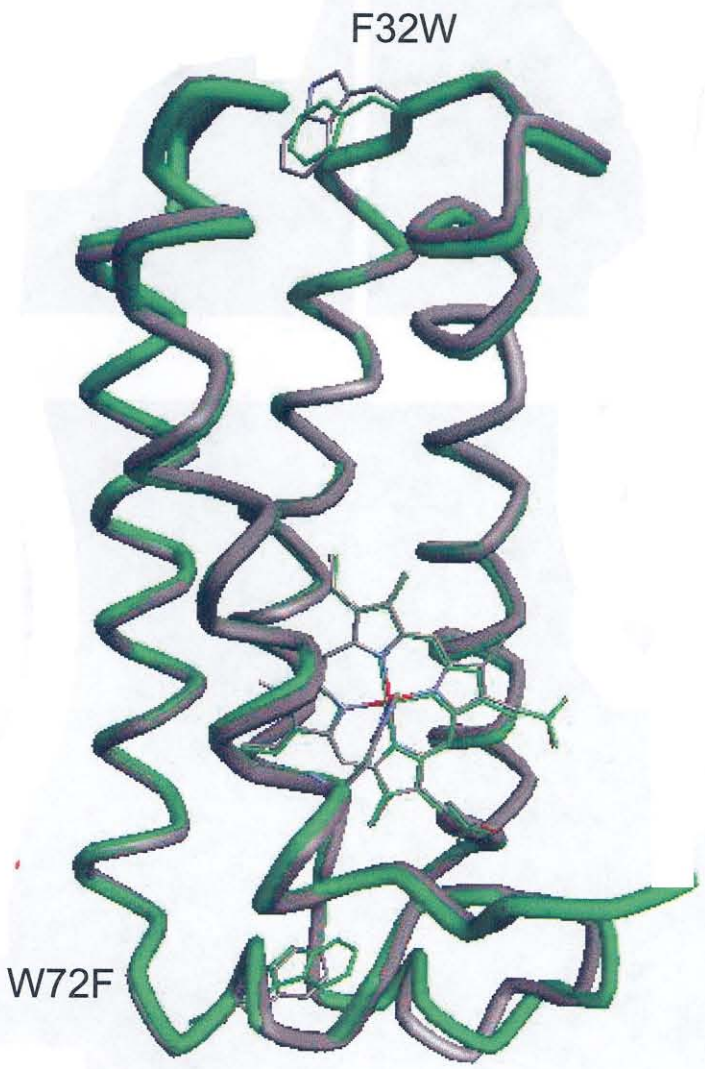


Figure 5.4. Structures of two cytochromes *c*'. Backbone atoms of the *R. palustris* protein (PDB 1A7V, green) are superimposed on the corresponding atoms in the Q1A/F32W/W72F mutant (gray)with a calculated RMSD of 0.44 Å. The mutated side chains are also shown.

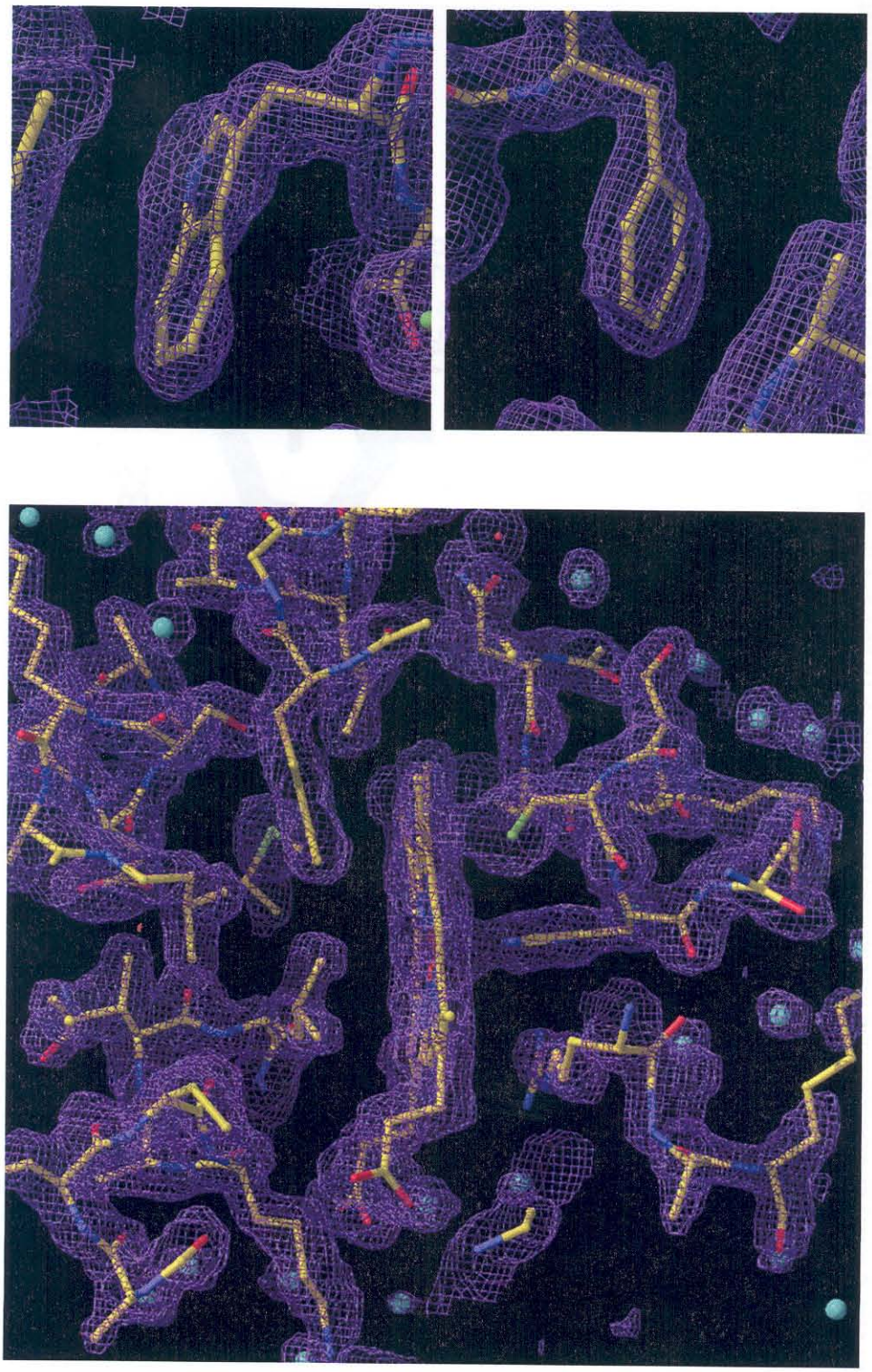


Figure 5.5. The $2F_o–F_c$ electron density maps (contour 2σ) of the mutation sites (F32W and F72F) and the heme pocket of Q1A/F32W/W72F cyt *c*'.

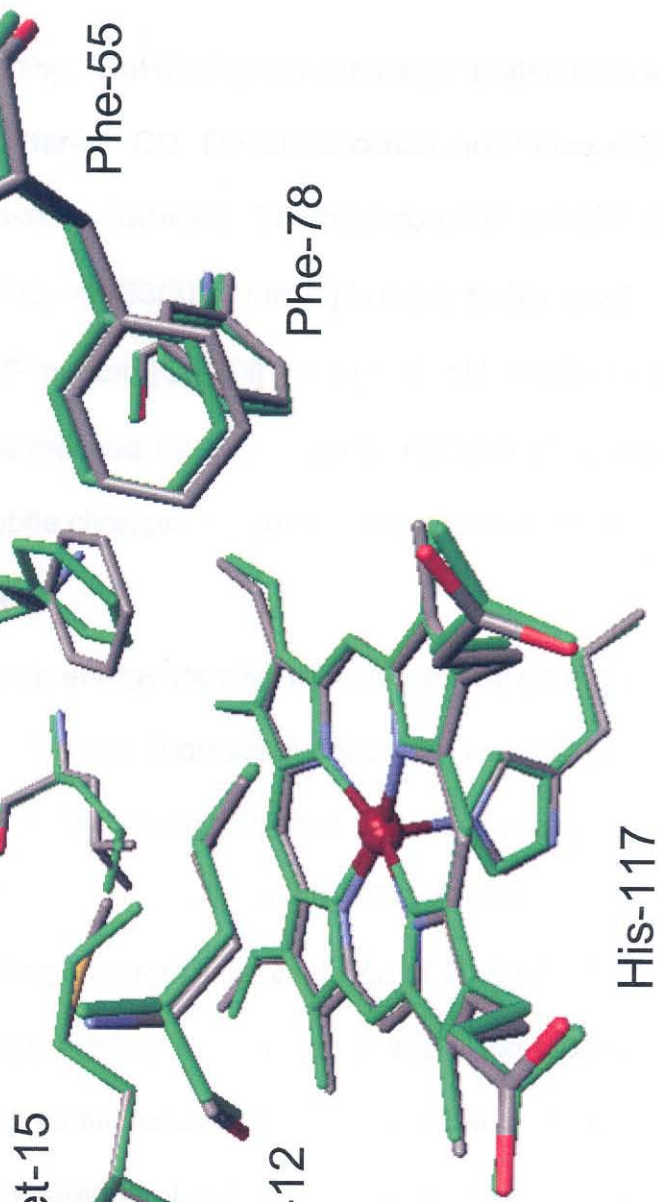


Figure 5.6. Overlay of residues in the hydrophobic heme pockets of *R. palustris* (green, PDB 1A7V) and Q1A/F32W/W72F (gray) cyt *c'*.

The introduction of Trp-32 at the top of the bundle does not seem to perturb the surrounding structure.

Equilibrium unfolding. GuHCl titration curves generated from ensemble-averaged probes (far-UV CD, Trp fluorescence, and heme absorption) are well described by two-state equilibria. The pseudo-wild-type Q1A protein ($[GuHCl]_{½}$ = 1.9(1) M; ΔG_f^o (H_2O) = –33(3) kJ mol $^{-1}$) is more stable toward unfolding than the Q1A/F32W/W72F mutant ($[GuHCl]_{½}$ = 1.(1) M; ΔG_f^o (H_2O) = –16(2) kJ mol $^{-1}$) (Figure 5.7). The reduced stability of the Q1A/F32W/W72F mutant most likely is attributable to subtle changes in hydrophobic interactions accompanying the Phe ↔ Trp exchange.

Dipole-dipole energy-transfer rates depend upon the inverse sixth power of the distance between a fluorescent donor (**D**) and acceptor (**A**). A **DA** pair is characterized by a critical length (r_o) that defines the range of distances (r) that can be probed ($0.3r_o < r < 1.5r_o$) (18, 19). The critical length (r_o) of the Trp-heme pair is 34 Å, yielding a useful distance range of 10-50 Å. The Trp-72 fluorescence decay kinetics in the folded pseudo-wild-type (Q1A) protein can be fit to a single exponential function with a decay time of 41 ps, corresponding to a Trp-72 to heme separation of ~15 Å. This value is in good agreement with the crystallographically determined distance. Fluorescence decay kinetics in the folded mutant (Q1A/F32W/W72F) are consistent with a distribution of Trp-32 to heme distances centered at 20 Å with a full width at half maximum (fwhm) of 2 Å, similar to the distance (~ 25 Å) found in the crystal structure. The shorter

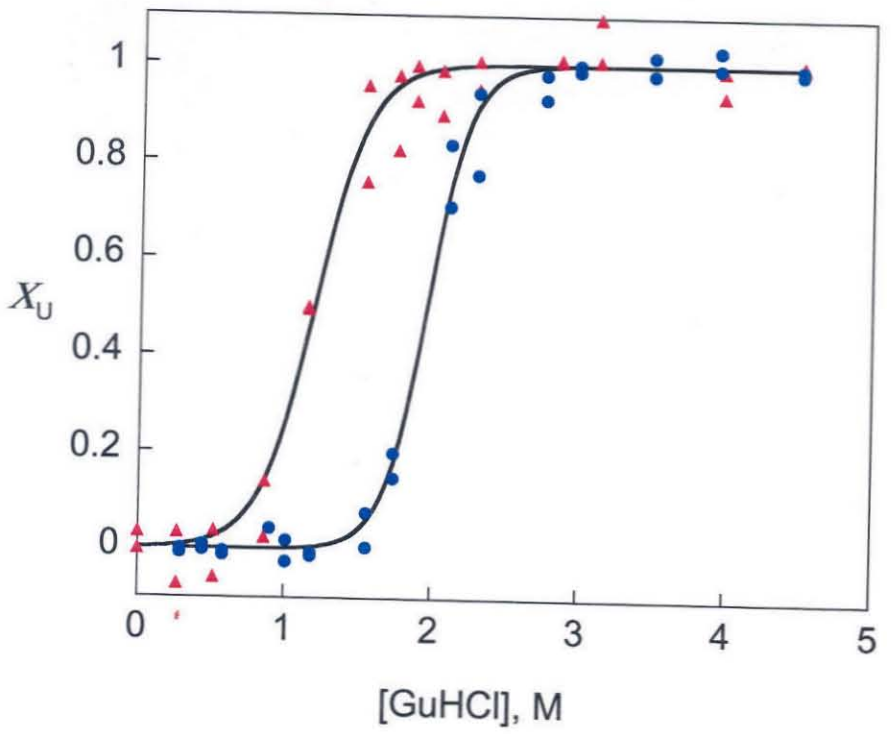


Figure 5.7. Denaturation curves for Q1A (●) and Q1A/F32W/W72F (▲) cyt *c*' determined by heme absorption, CD, and Trp fluorescence.

distance determined from the FET kinetics is likely due to favorable orientation of the Trp and heme transition dipoles in the folded protein.

More complex behavior is apparent with the addition of GuHCl (Figure 5.8). As GuHCl is added to the protein solution, the Trp-72 and Trp-32 decay kinetics become biphasic: the appearance of slow decay times (~3.5 ns) indicates the presence of unfolded polypeptides. Additionally, the fast-decay components shift to slower rates with increasing [GuHCl], suggesting a loosening of the native structure. The FET kinetics data obtained in the presence of denaturant can be interpreted in terms of a heterogeneous collection of compact and extended conformations (Figure 5.9). The mean **DA** distance moves to larger values as [GuHCl] increases. At the GuHCl midpoints (Q1A: 1.9 M; Q1A/F32W/W72F: 1.1 M), a substantial population of polypeptides with intermediate **DA** distances (25–30 Å) is revealed. Finally, at GuHCl concentrations above the midpoint, the small-r distribution loses all amplitude in favor of extended (r > 45 Å) and compact conformations (with mean distances of ~24 and ~32 Å). Despite having the Trp residues separated by 40 residues, the distance distributions extracted from the two different proteins are very similar. The variations in the **DA** distance distributions are consistent with cooperative unfolding of cyt *c'*, but the FET kinetics reveal structural heterogeneity that is not apparent from ensemble-averaged spectroscopic probes.

Fe^{III}–cytochrome c' folding. The native heme environment of Fe^{III}–cyt *c'* forms in a single kinetics phase (~ 1 s time constant) following stopped-flow dilution of

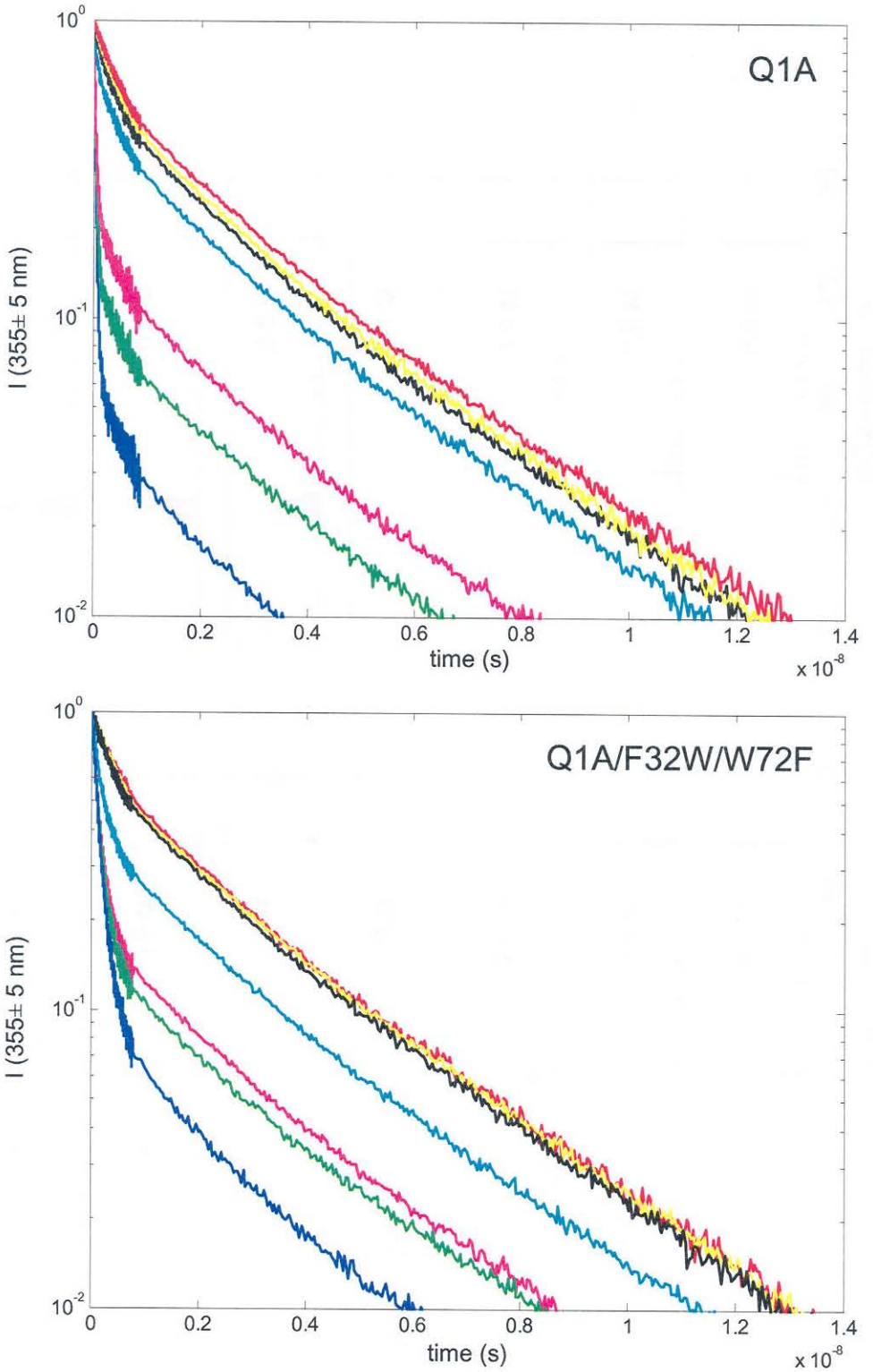


Figure 5.8. Trp fluorescence decay kinetics for Q1A (upper) and Q1A/F32W/W72F (lower) cyt *c'* as a function of [GuHCl] (Q1A: 0.0, 0.9, 1.5, 1.9, 2.0, 2.5, and 4.0 M GuHCl; Q1A/F32W/W72F: 0.0, 0.6, 0.8, 1.1, 1.4, 1.9, and 2.5 M GuHCl, pH 7).

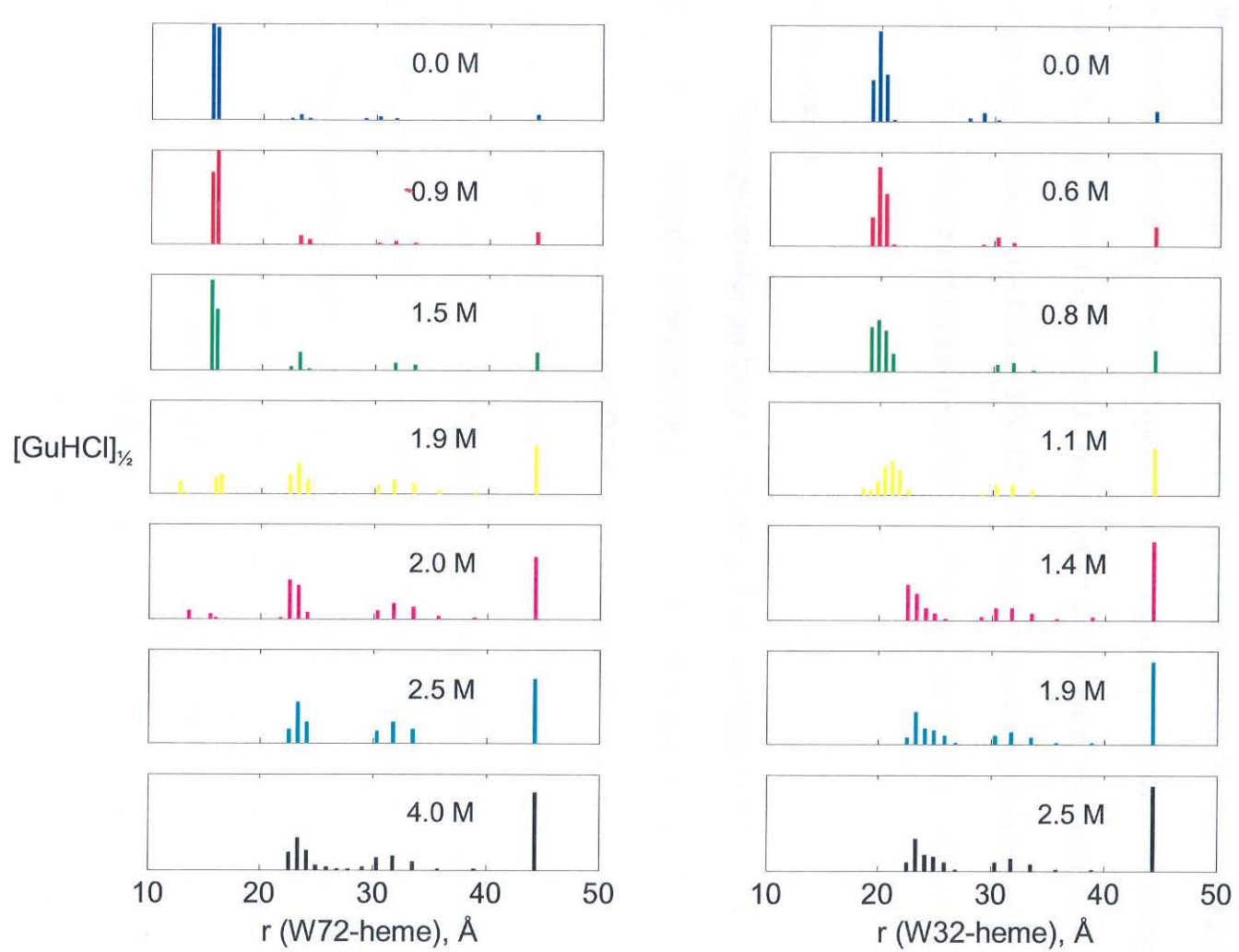


Figure 5.9. Trp-heme distance distributions extracted from FET kinetics obtained for Q1A (left panel) and Q1A/F32W/W72/F (right panel) cyt *c*' with variations in [GuHCl] ($[GuHCl]_{½}$ is the unfolding midpoint).

GuHCl (6). A burst phase is revealed when refolding is probed by Trp fluorescence intensity. The majority (~ 75%) of the Trp-72 fluorescence is quenched in the mixing deadtime (~ 5 ms), while the remaining fraction takes a few seconds (Figure 5.10). Similar behavior is observed for the Trp-32 mutant. At comparable folding free energies, the folding kinetics measured using heme absorption spectroscopy occurs on the order of seconds. Fluorescence measurements on the Q1A/F32W/W72F mutant show an increase in the amplitude of the burst phase (~83 %) and folding rate (~20 s^{-1}) (Figure 5.11).

Discussion

Energy-transfer partners incorporated into a single polypeptide can provide site-specific conformational detail. Measurements of FET kinetics can give probability distributions of **DA** distances. Indeed, analysis of Trp-heme FET kinetics from two distinct residues (Trp-72 (native) and Trp-32 (mutant)) in unfolded cyt *c'* reveals underlying structural heterogeneity in the denatured protein. Even at GuHCl concentrations well beyond the unfolding transition, a substantial fraction of the polypeptides (~ 50%) adopt compact conformations (25-35 Å) in both pseudo-wild-type and mutant cyt *c'* (Figure 5.8). In contrast, the fraction (~10 %) of compact structures (29 Å) in unfolded dansyl-labeled yeast cyt *c* (DNS(Cys102)-cyt *c*) is significantly smaller (2). This dramatic difference in the relative composition of collapsed and extended structures in denatured cyts *c'* and *c* must be due to differences in their primary amino-acid sequences.

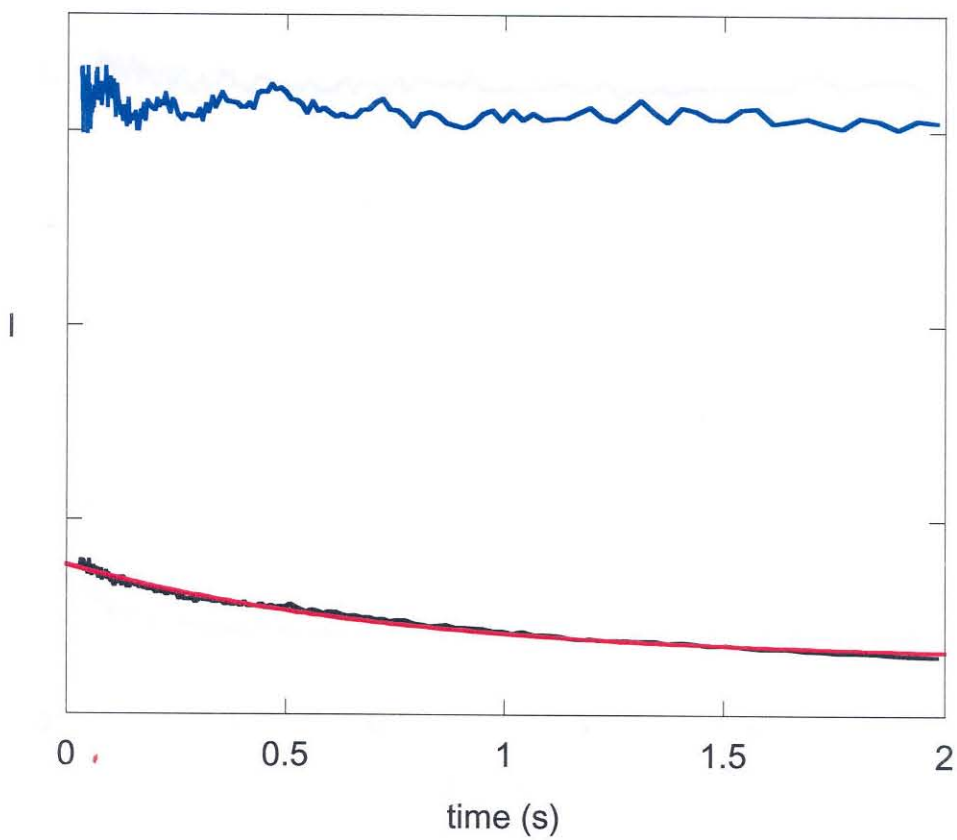


Figure 5.10. Folding kinetics of Q1A cyt *c*' measured by Trp-72 fluorescence intensity ([GuHCl] = 1.2 M; [cyt *c*'] = 10 μM). The data (black curve; logarithmically compressed to contain100 points per decay) can be adequately fit by a single exponential decay (red curve, k_f = 1.4 s^{-1}). The fluorescence intensity of the unfolded protein is represented by the blue curve.

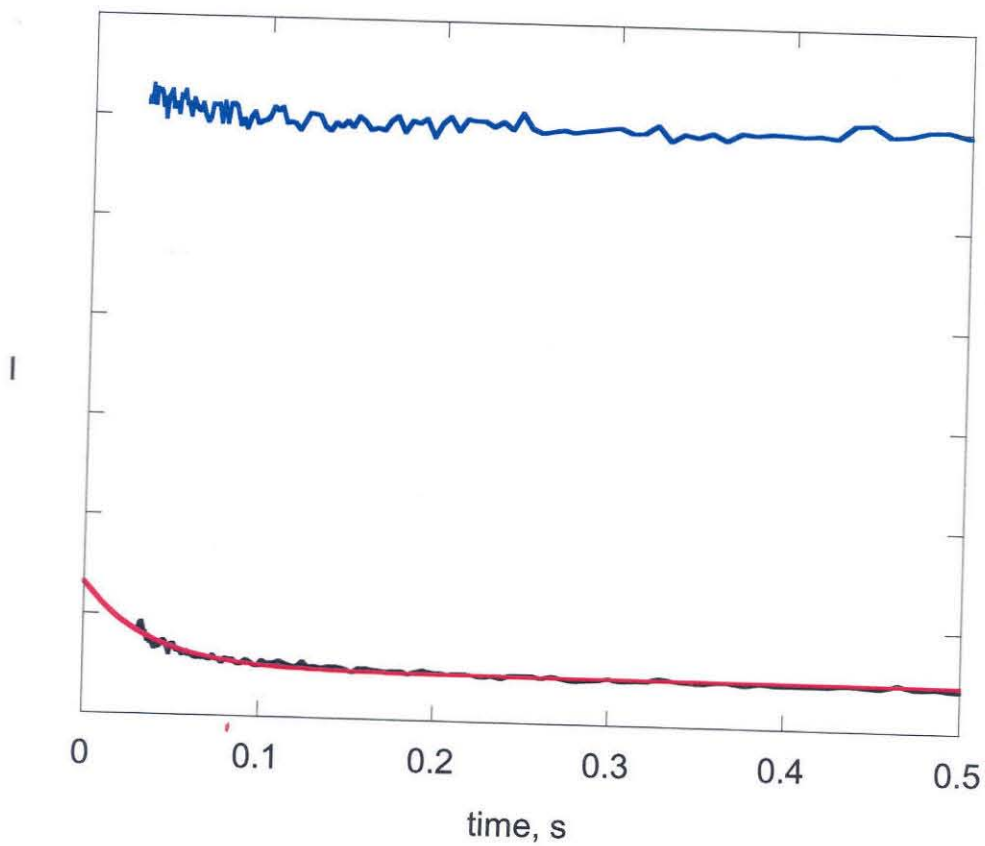


Figure 5.11. Folding kinetics of Q1A/F32W/W72F cyt *c*' measured by Trp-32 fluorescence intensity ([GuHCl] = 0.4 M; [cyt *c*'] = 10 μM). The data (black curve, logarithmically compressed to contain100 points per decay) can be adequately fit by a single exponential decay (red curve, k_f = 27 s^{-1}). The fluorescence intensity of the unfolded protein is represented by the blue curve.

Sequence comparison of the two proteins shows that cyt *c'* has a substantially higher fraction of nonpolar residues (54.4 %) than cyt *c* (49.1 %) (Table 5.2). The cyt *c'* sequence contains a large number of small residues (*e.g.*, 25 alanines), while cyt *c* has many charged and basic residues(*e.g.*, 16 lysines). Using Eisenberg's normalized consensus hydrophobicity scale (Figure 5.12) (20), the calculated average residue hydrophobicity for cyt *c'* is significantly greater

		Cyt *c'*		**Cyt *c***	
Property	Residues	Number	Mole %	Number	Mole %
Tiny	A, C, G, S, T	47	37.6	34	31.5
Small	A, C, D, G, N , P, S, T, V	69	55.2	52	48.1
Aliphatic	I, L, V	21	16.8	15	13.9
Aromatic	F, H, W, Y	7	5.6	14	13.0
Nonpolar	A, C, F, G, I, L, M, P, V, W, Y	68	54.4	53	49.1
Polar	D, E, H, K, N, Q, R, S, T	57	45.6	55	50.9
Charged	D, E, H, K, R	37	29.6	34	31.5
Basic	H, K, R	21	16.8	23	21.3
Acidic	D, E	16	12.8	11	10.2

Table 5.2 Amino-acid compositions of cyts *c'* and *c*.

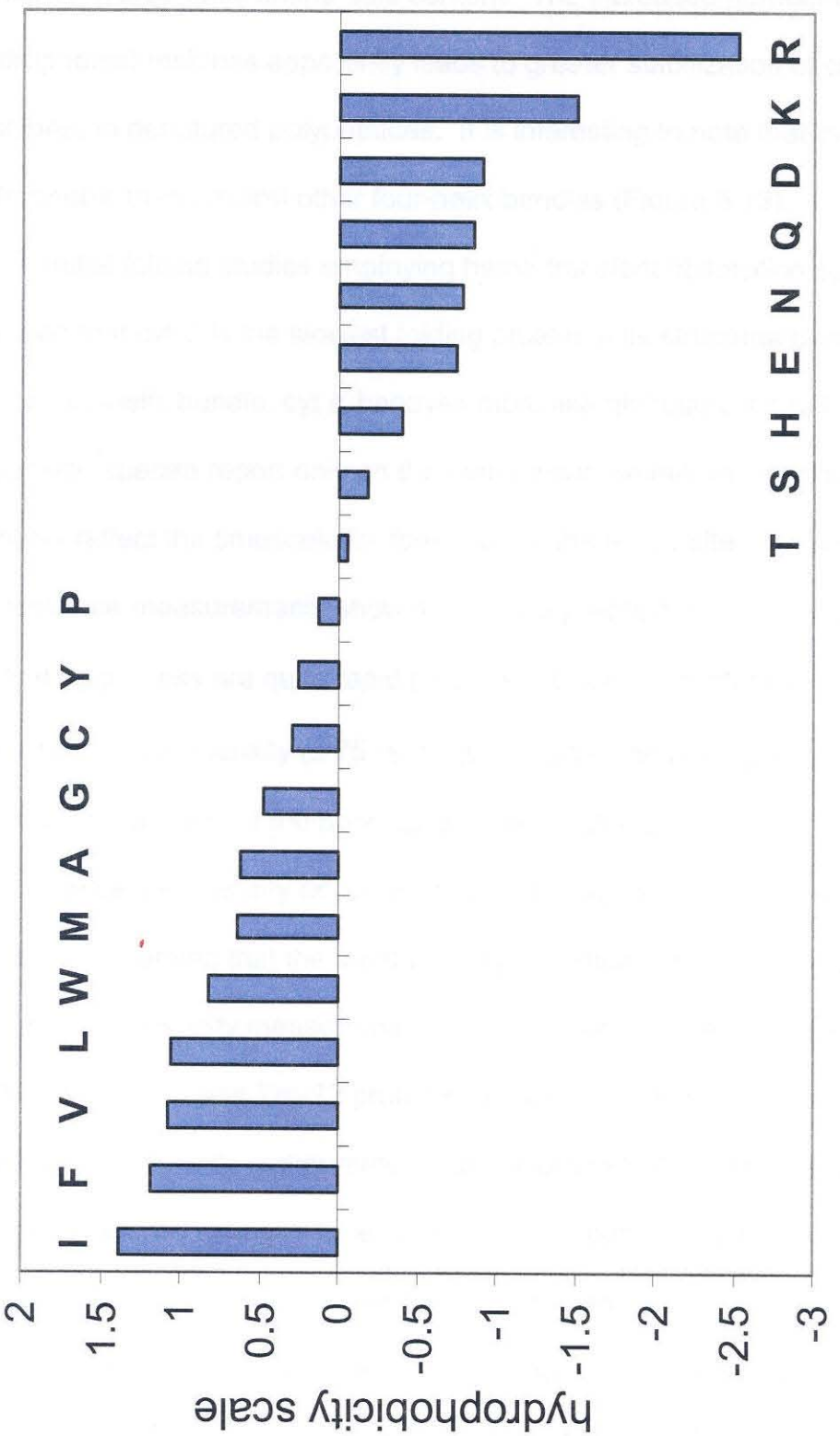


Figure 5.12. Eisenberg's hydrophobicity scale for individual amino acids. The values are normalized so that the mean value of the hydrophobicities is zero and the standard deviation is unity.

than for cyt *c*. Compact denatured structures may be prevalent in cyt *c'*, owing to its higher hydrophobic amino-acid content. The increased number of nonpolar (hydrophobic) residues apparently leads to greater stabilization of compact structures in denatured polypeptides. It is interesting to note that cyt *c'* is more hydrophobic than several other four-helix bundles (Figure 5.13).

Initial folding studies employing heme transient absorption spectroscopy revealed that cyt *c'* is the slowest folding protein in its structural class (6). Rather than a four-helix bundle, cyt *c'* behaves more like globular cyt *c* (21). Since absorption spectra report only on the heme environment, the measured kinetics primarily reflect the timescale for formation of the active site. In contrast, Trp fluorescence measurements show that the polypeptide dynamics associated with the folding process are quite rapid ($\leq$ 5 ms). Upon dilution of GuHCl, most of the Trp fluorescence intensity ($\leq$ 75 %) is quenched in the mixing deadtime, suggesting that most of the polypeptides have collapsed. Further reductions in Trp fluorescence intensity occur on the order of hundreds of milliseconds to seconds. Assuming that the burst phase population is homogeneous, the Trp fluorescence intensity measurements suggest average **DA** distances of 26 and 28 Å in the Trp-32 and Trp-72 proteins, respectively. It is not possible from these fluorescence intensity measurements to decipher whether there is just one collapsed intermediate or a heterogeneous distribution of compact (folded and/or nonnative) and extended structures. Measurements of FET kinetics during folding are necessary to provide the probability distributions of **DA** distances. The results for DNS(Cys102)-cyt *c* indicate that the latter scenario is more likely;

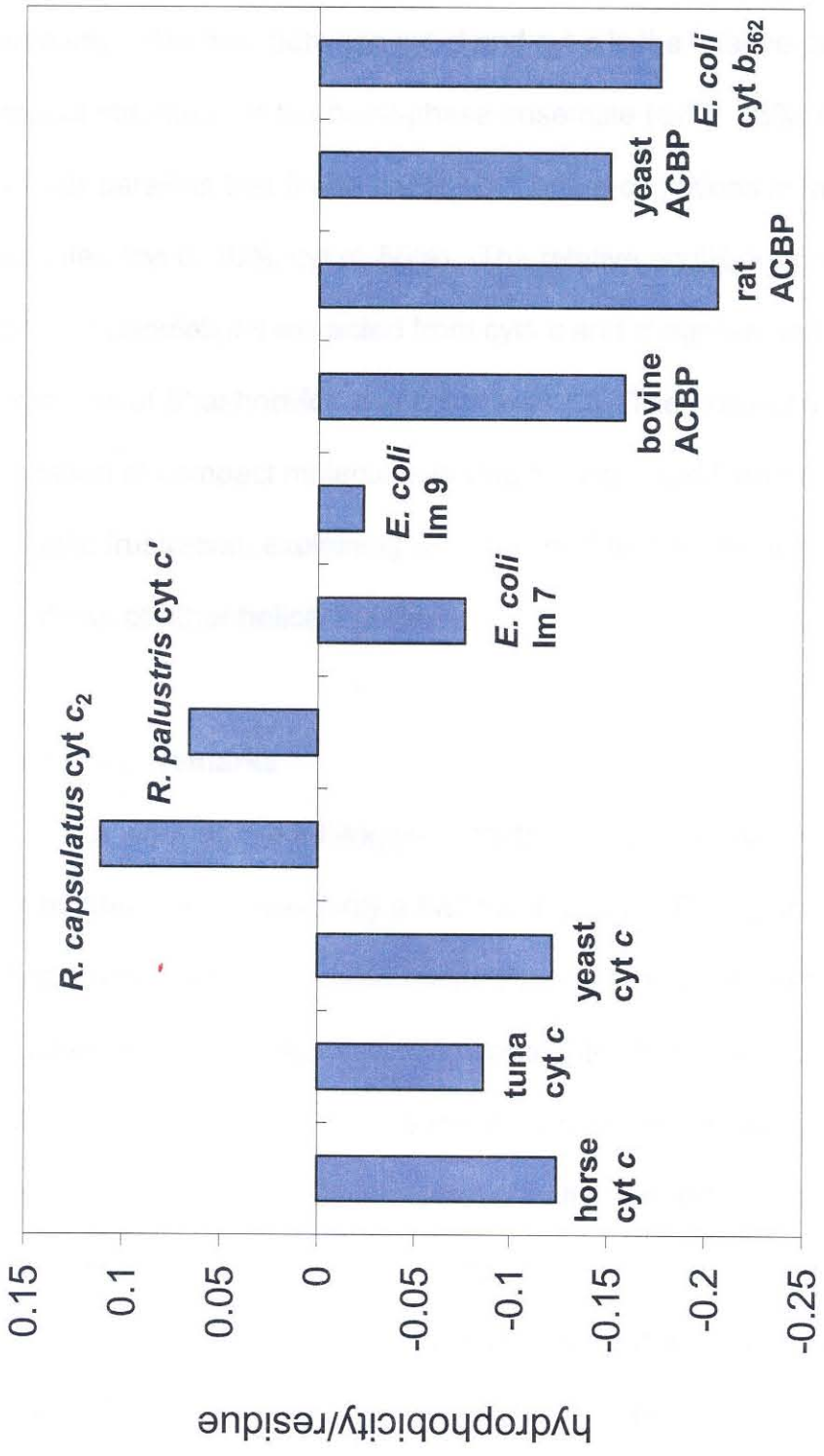


Figure 5.13. Comparisons of calculated average hydrophobicity values for cytochrome *c* and four-helix bundles (Im 7(E colicin binding immunity protein), Im 9, ACBP (Acyl-coenzyme A binding protein), cyt *c'*, and cyt b_{562}) .

that is, as soon as folding is initiated, there is an equilibrium between compact (~ 75-80%) and extended (~20-25 %) molecules. If this interpretation is correct, then a key difference between cyt *c'* and cyt *c* is the relative population of compact structures in the burst-phase ensemble (cyt *c*, 40%; cyt *c'*, 75%). This behavior parallels that found under equilibrium conditions in the denatured molecules (cyt *c*, 10%; cyt *c'*, 50%). The relative equilibrium of collapsed and extended populations extracted from cyts *c* and *c'* agrees well with the predictions of Shakhnovich and coworkers (3). The presence of a large population of compact molecules during folding could lead to topological and energetic frustration, explaining why the cyt *c'* folding dynamics are so different from those of other helical bundles.

Concluding Remarks

Our work on the ET-triggered folding of cyt *c'* shows clearly that a four-helix bundle is not necessarily a fast-folding motif. Fe^{II}-cyt *c'* exhibits complex folding kinetics with slow components (Chapter 3). An important finding is that nonnative methionine ligand binding contributes to the heterogeneous folding kinetics. Interestingly, unlike His-misligation which inhibits cyt *c* folding, Met-ligation appears to facilitate formation of native Fe^{II}-cyt *c'* (Chapter 4). Even in the absence of misligation (Fe^{III}-cyt *c'*), the folding dynamics are not simple. Fluorescence (Trp)-monitored folding reveals that a substantial compaction of the polypeptide occurs in milliseconds, but several seconds are required for the heme to adopt its native environment. The dominantly hydrophobic sequence of

this four-helix bundle appears to favor the formation of nonnative compact structures; an intriguing suggestion is that, following rapid hydrophobic collapse, the slow step involves "squeezing" water out of the heme pocket (4, 5).

References

1. Lee, J. C., Chang, I.-J., Gray, H. B. and Winkler, J. R. (2002) *J. Mol. Biol.*, in press.
2. Lyubovitsky, J. G., Gray, H. B. and Winkler, J. R. (2002) *J. Am. Chem. Soc.***124,** 5481-5485.
3. Gutin, A. M., Abkevich, V. I. and Shakhnovich, E. I. (1995) *Biochemistry* **34,** 3066-3076.
4. Sheinerman, F. B. and Brooks, C. L. (1998) *J. Mol. Biol.* **278,** 439-456.
5. Cheung, M. S., García, A. E. and Onuchic, J. N. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 685-690.
6. Lee, J. C., Gray, H. B. and Winkler, J. R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7760-7764.
7. Plaxco, K. W., Simons, K. T. and Baker, D. (1998) *J. Mol. Biol.* **277,** 985-994.
8. Plaxco, K. W., Simons, K. T., Ruczinski, I. and Baker, D. (2000) *Biochemistry* **39,** 11177-11183.
9. Wolynes, P. G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14249-14255.

10. Shibata, N., Iba, S., Misaki, S., Meyer, T. E., Bartsch, R. G., Cusanovich, M. A., Morimoto, Y., Higuchi, Y. and Yasuoka, N. (1998) *J. Mol. Biol.* **284,** 751-760.

11. Arslan, E., Schulz, H., Zufferey, R., Künzler, P. and Thöny-Meyer, L. (1998) *Biochem. Biophys. Res. Commun.* **251,** 744-747.

12. Pace, N. C., Shirley, B. A. and Thomson, J. A. (1990) in *Protein Structure: A Practical Approach*, ed. Creighton, T. F. (IRL Press, Oxford), pp. 311-330.

13. Nozaki, Y. (1972) *Methods Enzymol.* **26,** 43-50.

14. Otwinowski, Z. M., W. (1997) *Methods Enzymol.* **276,** 307-326.

15. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998) *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **54,** 905-921.

16. McRee, D. (1992) *J. Mol. Graphics* **10,** 44-46.

17. Laskowski, R. A., McArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283-291.

18. Förster, T. (1948) *Ann. Phys. (Leipzig)* **2,** 55-75.

19. Wu, P. and Brand, L. (1994) *Anal. Biochem.* **218,** 1-13.

20. Eisenberg, D., Schwarz, E., Komaromy, M. and Wall R. (1984) *J. Mol. Biol.* **179,** 125-142.

21. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. and Gray, H. B. (1996) *Chem. Biol.* **3,** 491-497.

Chapter 6

The cytochrome *c* folding landscape revealed by electron-transfer kinetics[†]

[†]Adapted from Lee, J. C.; Chang, I.-J.; Gray, H. B.; Winkler, J. R. *J. Mol. Biol.*, in press. Preparation and purification of Zn-cyt *c* were done by Dr. I-Jy Chang. Folding experiments were done in collaboration with Dr. I-Jy Chang.

Introduction

Great strides have been made recently in understanding exactly how a polypeptide self-assembles into the native structure of a protein (1–10). Energy landscape models, in particular, have provided a conceptual framework that has aided both the design and interpretation of folding experiments (11, 12). The configurational entropy of a polypeptide, represented by the lateral dimension on a funnel-shaped folding energy landscape, is minimized at the native conformation. This unfavorable configurational entropy change is balanced by a decrease in enthalpy due to native-contact formation, and an increase in solvent entropy (1). It is the large change in configurational entropy that can introduce mechanistic (and, in turn, kinetics) complexity. Without detailed maps of the folding energy landscape, it is uncertain at what stage the conformational heterogeneity of the unfolded ensemble is lost. This issue is difficult to resolve experimentally; folding probes must report not just the average progress of the unfolded ensemble toward the folded state, but on the ensemble heterogeneity as well. Indeed, few spectroscopic methods can even distinguish between single and bimodal distributions of molecules in a partially folded ensemble (Figure 6.1).

The reactivity of a polypeptide can be an extremely sensitive indicator of structural heterogeneity. With a carefully selected probe reaction, a bimodal distribution of protein conformations would exhibit biphasic kinetics, whereas a single-mode distribution would react in a single phase. The key requirement is that the probe reaction be fast compared to the timescale of folding ($<10^{-4}$ to $>10^{1}$ s). Electron-transfer (ET) reactions could be excellent folding probes

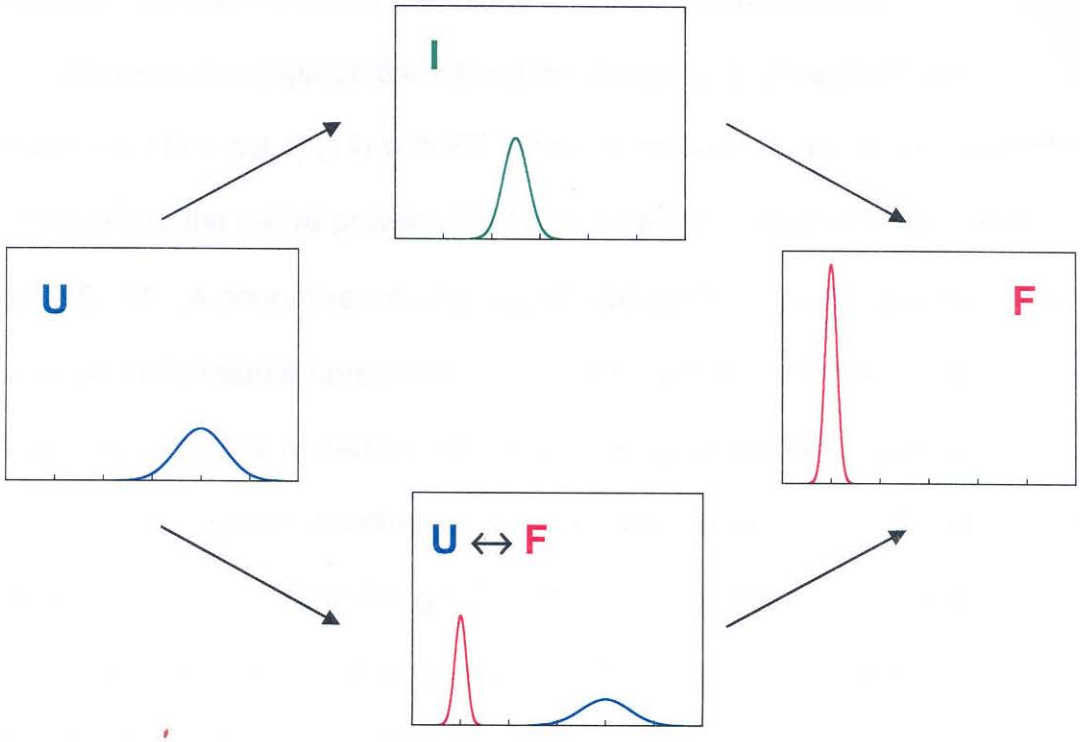


Figure 6.1. Schematic representation of polypeptide ensembles before, during, and after a folding process. A two-state conversion of the unfolded (**U**) into the folded (**F**) ensemble would proceed along the lower path. Population of folding intermediates (**I**) in represented by the upper path. Measurements of ET kinetics can be used to distinguish an ensemble composed solely of **I** from a bimodal distribution of **U** and **F**.

because rates at high driving forces are determined by the distance and medium separating the two redox partners (13). Buried redox centers in proteins often exchange electrons rather slowly with reagents in solution. Unfolding will greatly increase the accessibility of a redox cofactor and can lead to much faster ET (Figure 6.2). Compact intermediates might be expected to exhibit ET rates somewhere between those of the folded and unfolded molecules.

We have investigated the folding landscape of Zn(II)-substituted cytochrome *c* (Zn–cyt *c*) (14) with ET reactivity probes. Zn-cyt *c* is structurally homologous to the native protein (Fe–cyt *c*), which has been studied in great detail (15–18). A nonnative axial ligand (His-26 or His-33) replaces Met 80 at neutral pH in denatured horse heart Fe–cyt *c* (19); the rate-limiting folding step in this case is correction of heme misligation. Owing to weaker binding and faster substitution at the sixth coordination site, replacement of Fe with Zn will eliminate axial ligand traps during refolding (20). The key advantage of the Zn-substituted protein is the availability of a long-lived ($t \sim 15$ ms), powerfully reducing ($E^\circ = -0.8$ V *vs.* NHE) triplet excited state (*Zn–cyt *c*), which can be prepared in 90% yield (21) with 580 nm laser excitation (22).

Materials and Methods

Zn–cyt *c* was prepared from the horse heart protein (Sigma) according to previously published procedures (14). Guanidine hydrochloride (UltrapureGrade, Sigma) and ruthenium(III) hexaammine (Strem) were used without further purification. GuHCl unfolding curves were generated from CD, absorption, and

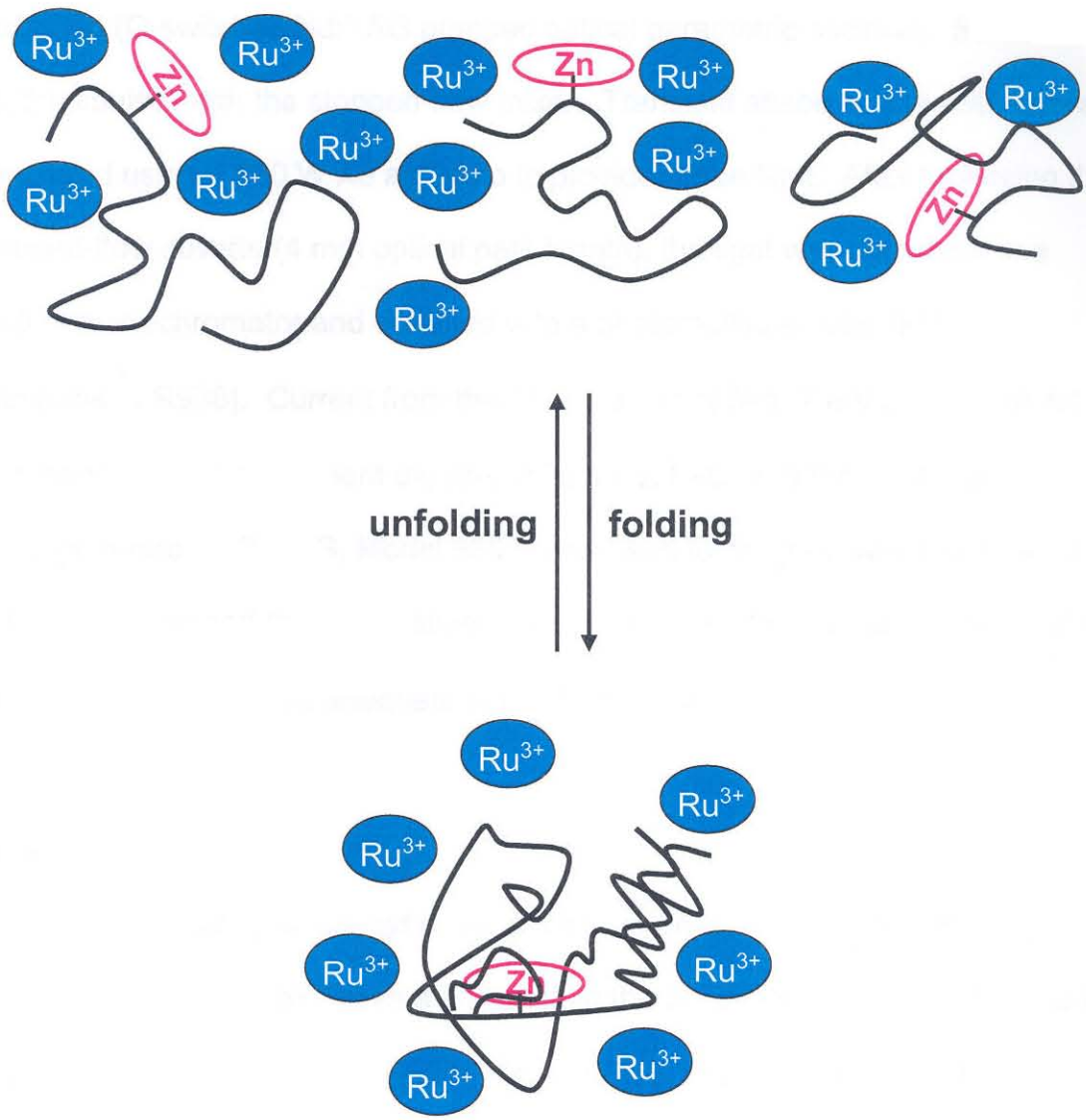


Figure 6.2. Cartoon illustrating the differences in solvent accessibility of the Zn-porphyrin in unfolded and folded polypeptides.

ET kinetics according to standard procedures (23). Stopped-flow triggered folding kinetics were measured using a Biologic SFM-4 stopped-flow mixer. ET kinetics were measured during Zn–cyt *c* folding by synchronizing pulsed laser excitation (Q-switched Nd:YAG pumped optical parametric oscillator, 580 nm, 10 ns, 5mJ/pulse) with the stopped-flow mixer. Transient absorption at 450 nm was measured using a 150 W Xe arc lamp to provide probe light. After traversing the stopped-flow cuvette (4 mm optical path length), the light was dispersed in a 0.25 m monochromator and detected with a photomultiplier tube (PMT, Hamamatsu R938). Current from the PMT was amplified (7 mV/μA, 100 MHz) and recorded with a transient digitizer (500 Ms/s, LeCroy 9354A). A gate and delay generator (EGandG, Model 9650) was used to program variable delay time between mixing and laser excitation of Zn–cyt *c*. Data from 10 laser shots ([Zn–cyt *c*] = 10 μM) provided adequate signal-to-noise levels.

Results

Equilibrium unfolding of Zn–cyt c. Addition of guanidine hydrochloride (GuHCl) to solutions of Zn–cyt *c* produces a blue shift in the Soret absorption band (folded, λ_{max} = 426 nm; unfolded, λ_{max} = 418 nm), giving a species with a spectrum similar to that of Zn(II)-substituted myoglobin (λ_{max} = 414 nm) (Figure 6.3) (24). GuHCl unfolding curves generated from Soret absorption and far-UV CD spectra show that the stability of folded Zn-cyt *c* is comparable to that of the Fe(III) form (Fe(III): ΔG_f^o = –40(1) kJ mol^{-1}, $[GuHCl]_{½}$ = 2.8(1) M; Zn(II): ΔG_f^o = –35(2) kJ mol^{-1}, $[GuHCl]_{½}$ = 2.9(1) M) (Figure 6.4) (25, 26). In contrast to Fe–cyt *c*, the

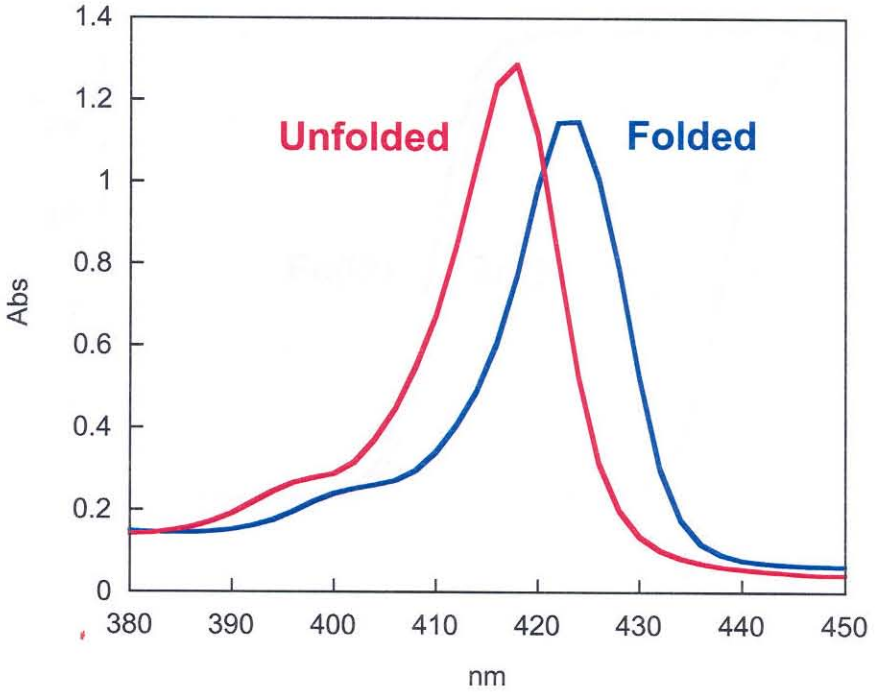


Figure 6.3. Absorption spectra of folded (in 100 mM sodium phosphate buffer, pH 7) and unfolded (in 4.5 M GuHCl, pH 7) Zn-cyt *c*.

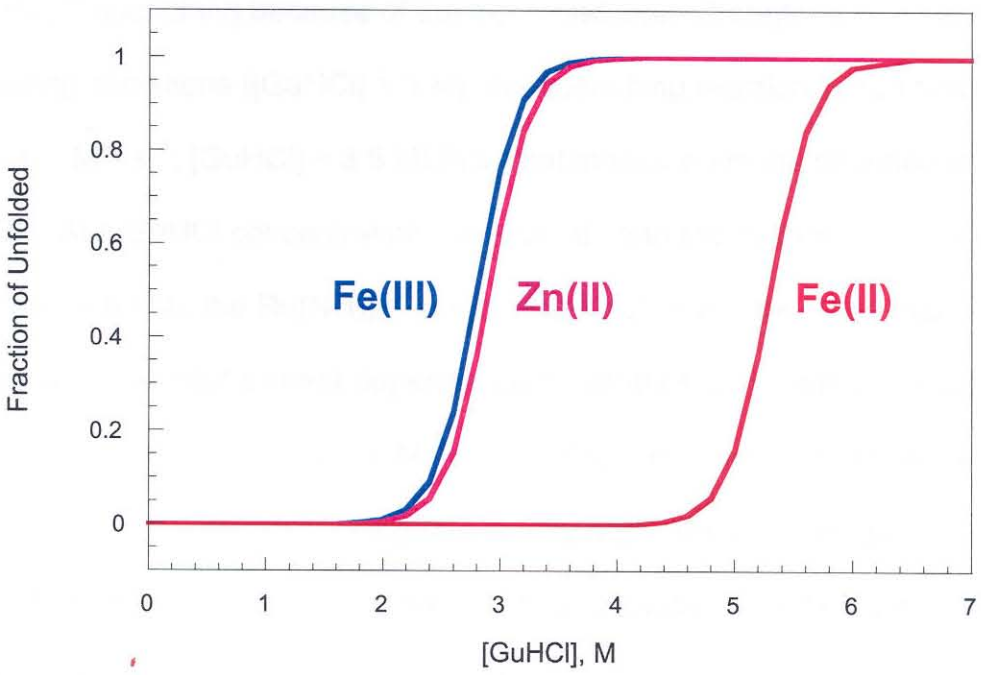


Figure 6.4. Equilibrium unfolding curves of Fe^{III}-, Zn^{II}-, and Fe^{II}-cyt *c* generated from absorption and CD data.

Zn center in the unfolded protein is not ligated by any peptide side chain other than the native His-18 (20).

Triplet-excited Zn–cyt *c* is quenched by $Ru(NH_3)_6^{3+}$, producing Zn–cyt $c^{\bullet+}$ and $Ru(NH_3)_6^{2+}$ (Figure 6.5, *inset*). The reported quenching rate constant measured under native conditions is $1.4 \times 10^7\ M^{-1}\ s^{-1}$ (22). Moderate concentrations of GuHCl (<1.0 M) do not unfold the protein, but accelerate $Ru(NH_3)_6^{3+}$ quenching because of the increased ionic strength (27). Under denaturing conditions ([GuHCl] > 3 M), the quenching reaction is 100 times faster ($1.4 \times 10^9\ M^{-1}\ s^{-1}$, [GuHCl] = 3.5 M) than that measured in the absence of GuHCl. At a GuHCl concentration corresponding to the midpoint of Zn–cyt *c* unfolding (2.85 M), the $Ru(NH_3)_6^{3+}$ quenching kinetics are biexponential. Both rate constants exhibit a linear dependence on $[Ru(NH_3)_6^{3+}]$, giving respective values of 7.2×10^7 and $1.0 \times 10^9\ M^{-1}s^{-1}$ for folded and unfolded ensembles (Figure 6.5). The latter quenching rate is very high, owing to the greater accessibility of the Zn–porphyrin in the unfolded protein. The *Zn–cyt *c* ET kinetics are generally consistent with a two-state unfolding process. The unfolding isotherm generated from ET kinetics exhibits a transition midpoint at 2.8(1) M GuHCl, in good agreement with those obtained from far-UV CD and heme absorption measurements (Figure 6.6).

Zn–cyt c folding kinetics. Changes in Zn–cyt *c* Soret absorption (418 and 426 nm) after stopped-flow dilution of denaturant (initial [GuHCl] = 3.4 M; final [GuHCl] = 1.20-2.60 M) were examined: the transient absorption kinetics are

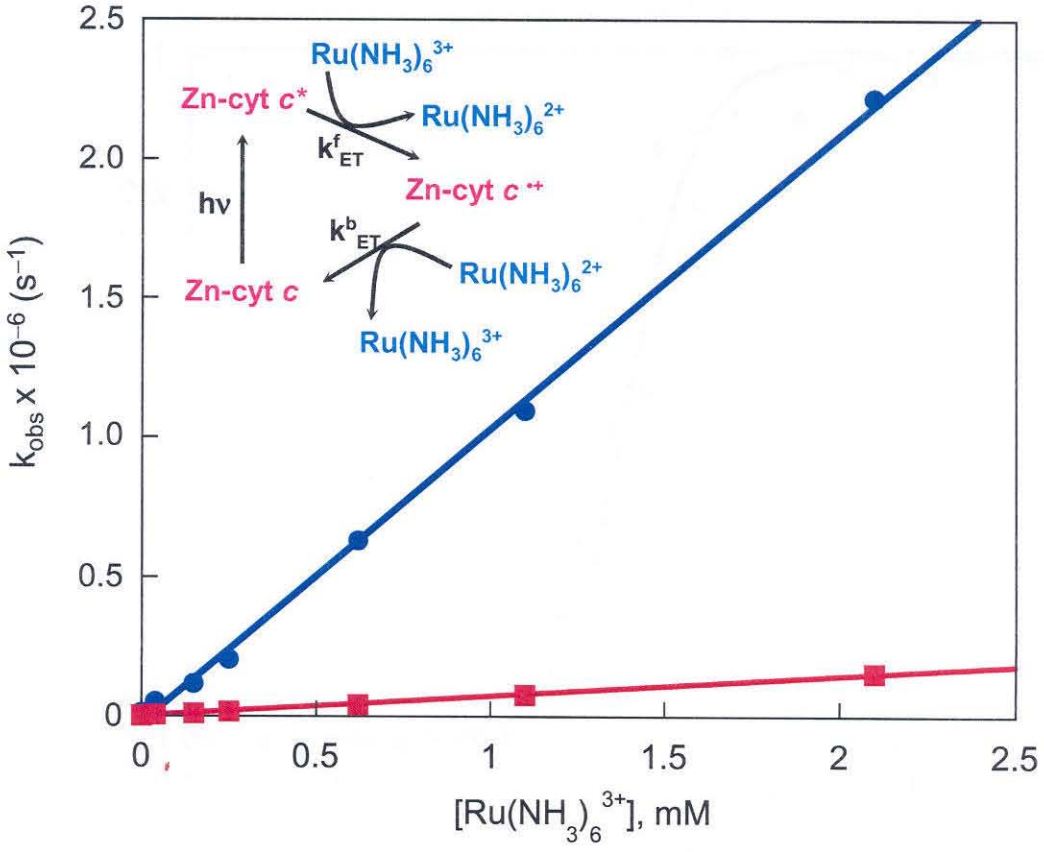


Figure 6.5. ET quenching of *Zn-cyt *c* by $Ru(NH_3)_6^{3+}$ (inset: reaction scheme) in deoxygenated GuHCl solution (2.85 M, 100 mM sodium phosphate, pH 7). The bimolecular ET reaction is about 10-fold slower for the folded (7.2×10^7 M^{-1} s^{-1}, ■) than the unfolded (1×10^9 M^{-1} s^{-1}, ●) protein.

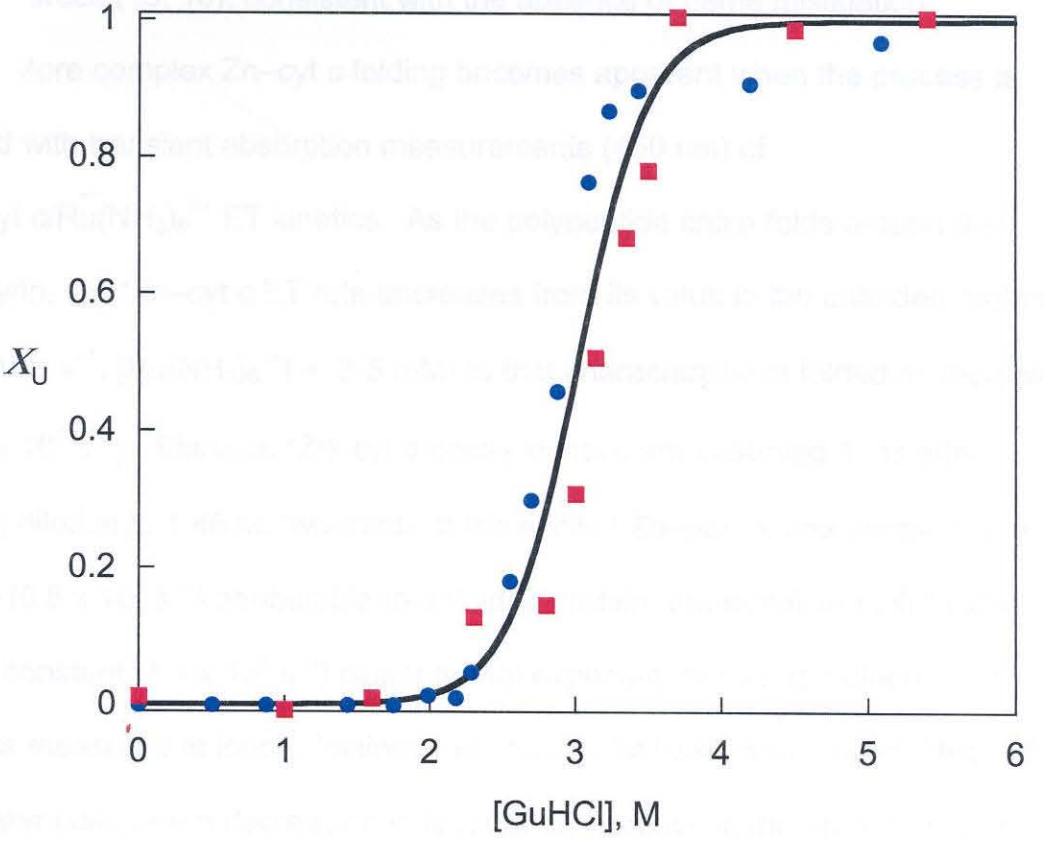


Figure 6.6. Denaturation curve for Zn-cyt *c* from Soret absorption (■) and ET rate (●).

exponential functions and the observed rate constants depend linearly on denaturant concentrations, decreasing from 1.3(2) x 10^2 s^{-1} at 1.20 M GuHCl to 1.1(1) x 10^1 s^{-1} at 2.60 M GuHCl (Figure 6.7). The extrapolated time constant for refolding in the absence of denaturant is about a millisecond. The Zn–cyt *c* folding rate is about 10 times higher than that of the Fe(III) protein at comparable driving forces (15, 16), consistent with the absence of heme misligation.

More complex Zn–cyt *c* folding becomes apparent when the process is probed with transient absorption measurements (450 nm) of *Zn–cyt *c*/$Ru(NH_3)_6^{3+}$ ET kinetics. As the polypeptide chain folds around the porphyrin, the *Zn–cyt *c* ET rate decreases from its value in the unfolded protein (≈7 x 10^6 s^{-1}, $[Ru(NH_3)_6^{3+}]$ = 3-5 mM) to that characteristic of folded molecules (≈2.5 x 10^5 s^{-1}). Biphasic *Zn–cyt *c* decay kinetics are observed 1 ms after GuHCl dilution to 1.46 M: two-thirds of the excited Zn–porphyrins decay in a fast phase (6.8 x 10^6 s^{-1}) attributable to unfolded protein; the remaining third exhibits a rate constant (4.1 x 10^5 s^{-1}) closer to that expected for folded molecules. ET kinetics measured at longer folding times remain biphasic with the amplitude of the faster component decreasing in favor of an increase in the amplitude of the slow component (Figure 6.8).

Both *Zn–cyt *c* decay rates decrease by about a factor of 2 (1 ms, 6.8 x 10^6, 4.1 x 10^5 s^{-1}; 50 ms, 3.2 x 10^6, 2.4 x 10^5 s^{-1}) as the folding reaction proceeds (Figure 6.9). This variation in quenching efficiency reflects a gradual collapse of polypeptide structures during folding. For the slower phase, the two-fold reduction in rate constant is consistent with the conversion of compact,

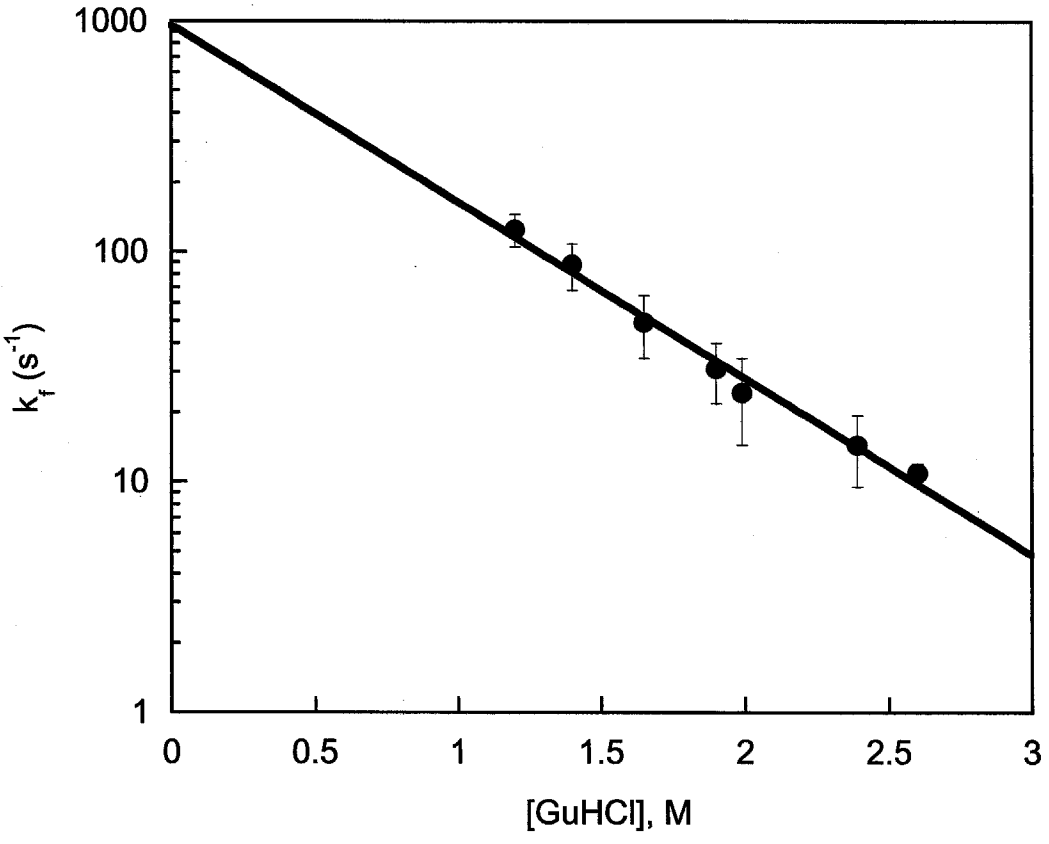


Figure 6.7. Dependence of Zn-cyt *c* folding rates on GuHCl concentration.

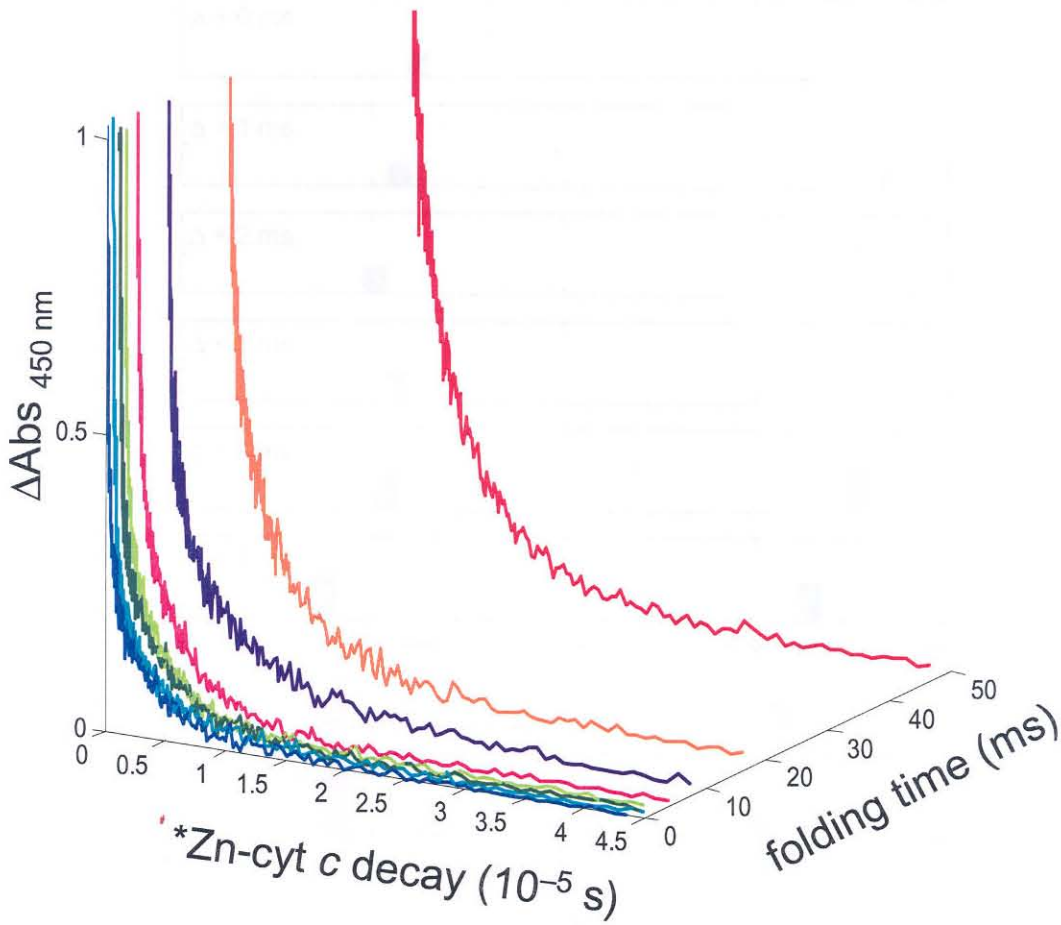


Figure 6.8. Stopped-flow triggered folding of Zn-cyt *c* ([GuHCl] = 1.46 M; [Zn-cyt *c*] = 10 μM; $[Ru(NH_3)_6]^{3+}$ = 4 mM; pH 7) monitored by laser transient absorption measurements of *Zn-cyt *c* decay (450 nm, signal normalized to an initial value of 1.0) at different delay times after mixing (0, 1, 2, 3, 5, 10, 20, and 50 ms).

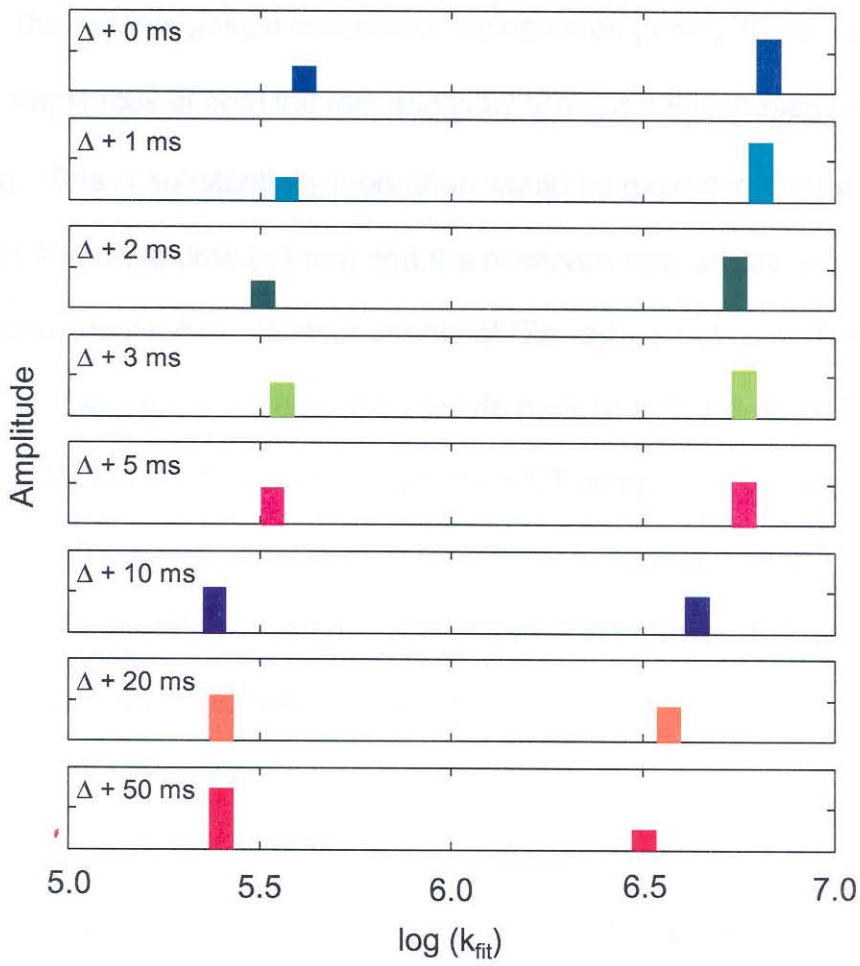


Figure 6.9. Rate constants and relative amplitudes from biexponentical fits of *Zn-cyt c decay data measured during folding. The observed rate constants decrease by about a factor of two as folding proceeds.

nonnative conformations to fully folded protein molecules.

The amplitudes of the two *Zn–cyt *c* decay phases vary exponentially with folding time and the rate constant (33 s^{-1}, [GuHCl] = 1.46 M) is in reasonable agreement with that measured by Soret absorption spectroscopy. It is noteworthy that, at the earliest measured folding times (1 ms), there are significant amplitudes in *both* the fast and slow *Zn–cyt *c* ET phases (≈67% fast; ≈33% slow). This is substantially more than would be expected on the basis of the stopped-flow dead-time (~1 ms) and the observed rate constants extracted from the signal amplitudes. Measurements of *Zn–cyt *c* kinetics at different GuHCl concentrations consistently extrapolate back to a "burst-phase" ensemble (15, 16, 28, 29) with a 2:1 ratio of fast and slow ET components (Figure 6.10): these results demonstrate that the burst ensemble is heterogeneous; molecules in one-third of the protein population have compact structures, and ones in the remaining fraction have exposed Zn–porphyrins.

Discussion

It is apparent that there is underlying complexity in Zn–cyt *c* folding. The fraction of the burst ensemble (~1/3) with slow *Zn–cyt *c* decay kinetics could be fully folded protein or an ensemble of compact nonnative structures. The former possibility would be an example of "fast-track" folding (30-32), where about a third of the unfolded Zn–cyt *c* molecules adopt conformations that can refold very quickly. The remaining protein molecules have relatively exposed porphryin groups, and fold on a substantially longer timescale. Alternatively, the 2:1 ratio of

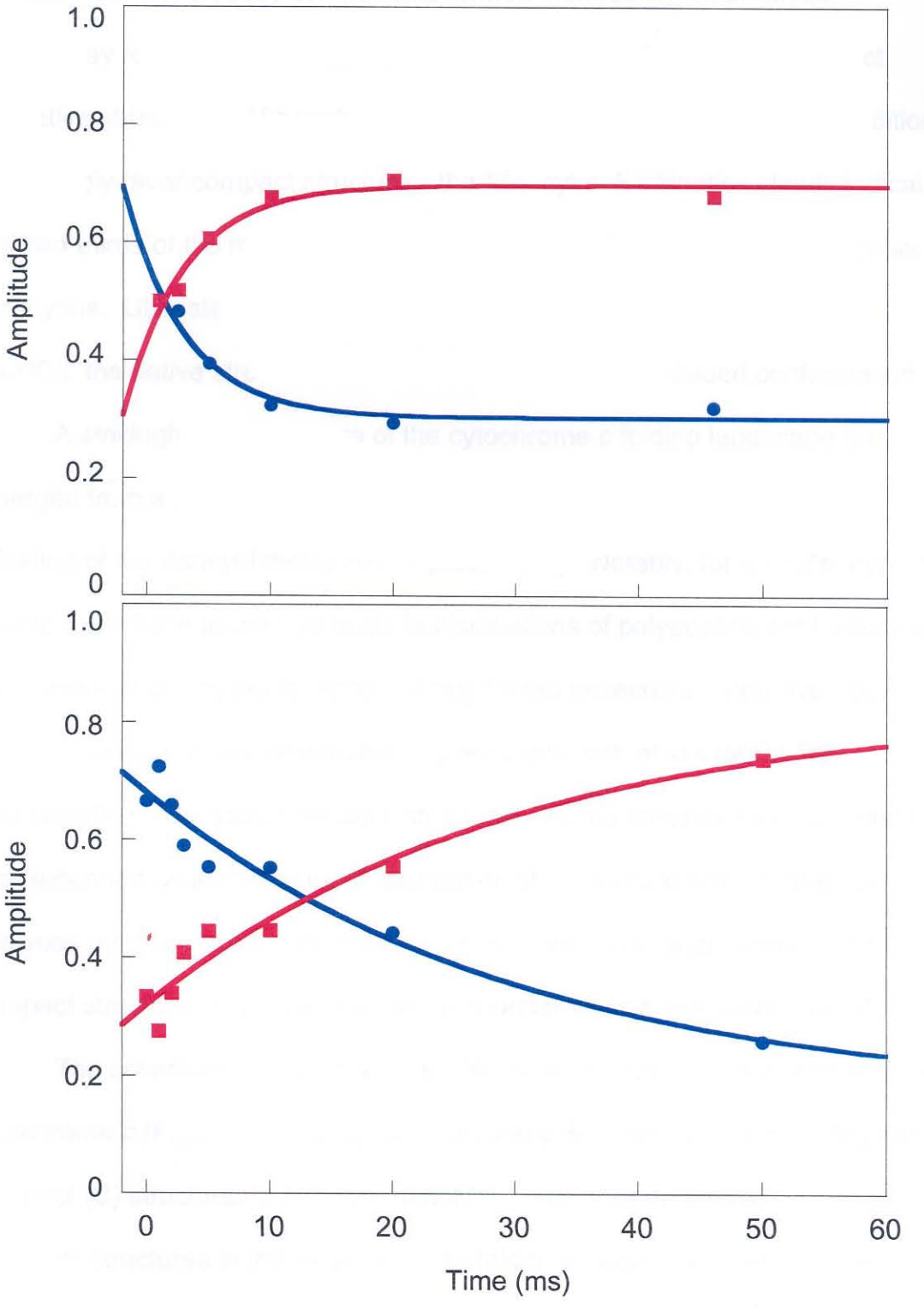


Figure 6.10. Amplitudes of fast and slow decay constants extracted from biexponential fits to *Zn-cyt *c* decay kinetics as functions of time after the initiation of protein folding. Top panel: final [GuHCl] = 1.10 M; folding rate constant, $k_f = 200\ s^{-1}$. Bottom panel: final [GuHCl] = 1.46 M; $k_f = 33\ s^{-1}$.

fast:slow *Zn–cyt *c* decay components formed immediately after dilution of GuHCl may reflect a shift in the equilibrium between unfolded and compact nonnative structures. Although it is often assumed that native solvent conditions will strongly favor compact structures, the *Zn–cyt *c* ET kinetics clearly indicate that two-thirds of the molecules in the burst intermediates have highly exposed porphyrins. Ultimately, the entire protein ensemble folds because, at low [GuHCl], the native structure is much more stable than unfolded conformations.

A strikingly similar picture of the cytochrome *c* folding landscape has emerged from an analysis of fluorescence energy-transfer kinetics during refolding of the dansyl-labeled native protein (33). Notably, for both Zn–cyt *c* and Fe–cyt *c*, we have found that multiple distributions of polypeptide conformations are populated during the formation of fully folded molecules. Moreover, our kinetics measurements demonstrate quite clearly that, at low denaturant concentrations, compact nonnative structures are not substantially more stable than extended conformations. These observations accord with theoretical analyses by Shakhovich and coworkers that have shown that extended and compact structures could be in equilibrium under folding conditions (34, 35).

The schematic folding-energy profile derived from our measurements on cytochrome *c* (Figure 6.11) suggests that nearly degenerate extended (**E**) and compact (**C**) structures equilibrate rapidly (<1 ms) after denaturant is diluted. Compact structures in the Fe protein can fold by crossing an energy barrier that has a substantial ligand-substitution component. This barrier is lowered in the Zn protein because there is no misligation (20). Owing to topological frustration, not

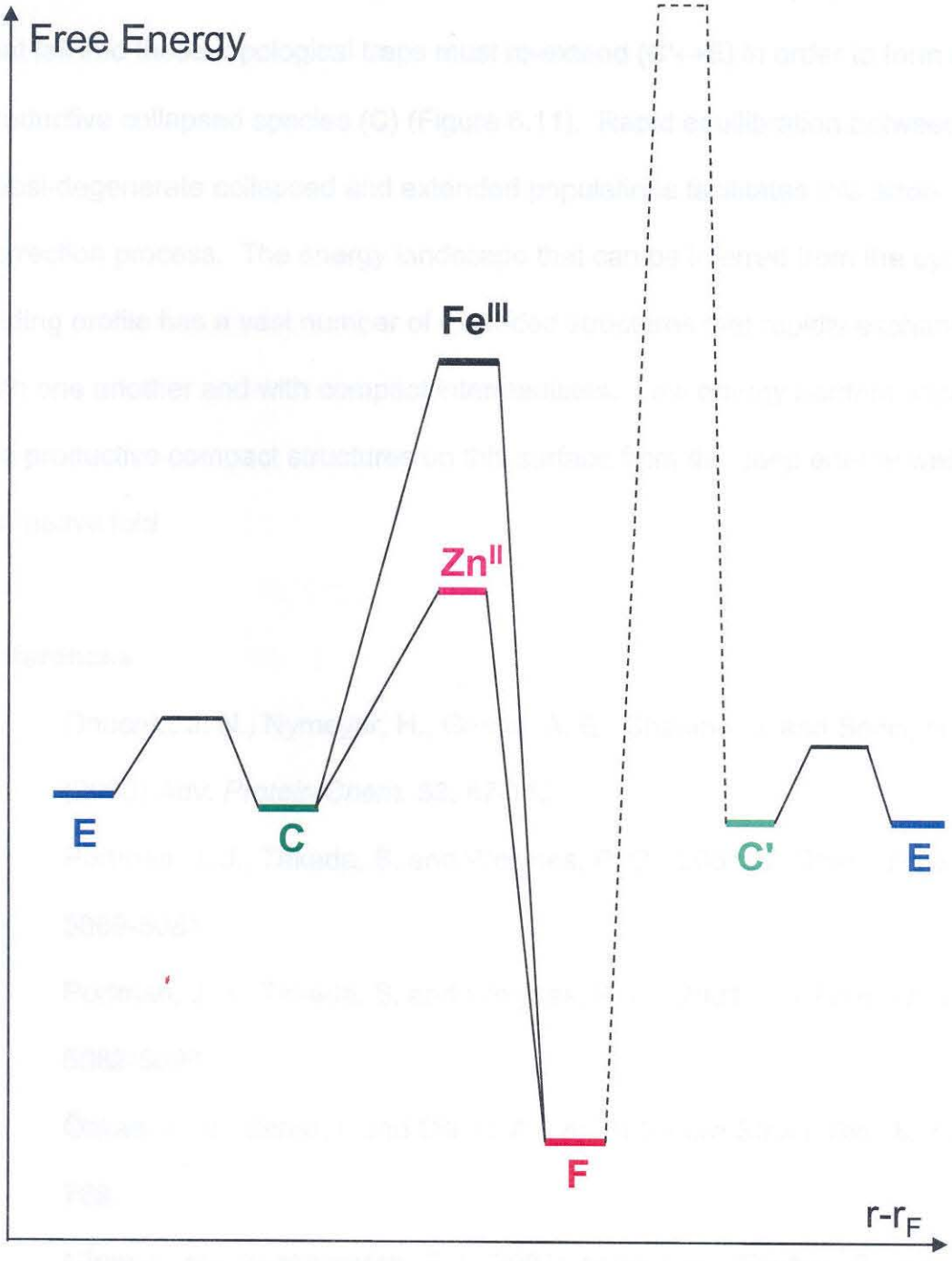


Figure 6.11. Qualitative folding energy profile for cytochrome *c* as a function of the deviation of the protein radius (r) from its value in the native fold (r_F). Extended (E) and compact (C) structures are nearly degenerate and equilibrate on a submillisecond timescale. Compact structures fold in the Fe protein by crossing an energy barrier with a sizable ligand-substitution component. There is no misligation for folding from the compact structures of the Zn protein. Topologically trapped compact structures (C') must extend and collapse to a productive structure (C) in order to reach the native conformation.

all collapsed conformations (**C'**) can proceed to folded product (31). Molecules that fall into these topological traps must re-extend (**C'↔E**) in order to form a productive collapsed species (**C**) (Figure 6.11). Rapid equilibration between quasi-degenerate collapsed and extended populations facilitates this error-correction process. The energy landscape that can be inferred from the cyt *c* folding profile has a vast number of extended structures that rapidly exchange with one another and with compact intermediates. Low energy barriers separate the productive compact structures on this surface from the deep energy well of the native fold.

References

1. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J. and Socci, N. D. (2000) *Adv. Protein Chem.* **53,** 87-152.
2. Portman, J. J., Takada, S. and Wolynes, P. G. (2001) *J. Chem. Phys.* **114,** 5069-5081.
3. Portman, J. J., Takada, S. and Wolynes, P. G. (2001) *J. Chem. Phys.* **114,** 5082-5096.
4. Özkan, S. B., Bahar, I. and Dill, K. A. (2001) *Nature Struct. Biol.* **8,** 765-769.
5. Mirny, L. and Shakhnovich, E. I. (2001) *Annu. Rev. Biophys. Biomol. Struct.* **30,** 361-396.
6. Kazmirski, S. L., Wong, K. B., Freund, S. M. V., Tan, Y. J., Fersht, A. R. and Daggett, V. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4349-4354.

7. Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M. V., Fersht, A. R. and Daggett, V. (2000) *J. Mol. Biol.* **296,** 1257-1282.

8. Eaton, W. A., Munoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R. and Hofrichter, J. (2000) *Annu. Rev. Biophys. Biomol. Struct.* **29,** 327-359.

9. Kuwata, K., Shastry, R., Cheng, H., Hishino, M., Batt, C. A., Goto, Y. and Roder, H. (2001) *Nature Struct. Biol.* **8,** 151-155.

10. Jager, M., Nguyen, H., Crane, J. C., Kelly, J. W. and Gruebele, M. (2001) *J. Mol. Biol.* **311,** 373-393.

11. Bryngelson, J. D., Onuchic, J. N. and Wolynes, P. G. (1995) *Proteins: Struct. Func. Gen.* **21,** 167-195.

12. Dill, K. A. and Chan, H. S. (1997) *Nature Struct. Biol.* **4,** 10-19.

13. Gray, H. B. and Winkler, J. R. (1996) *Annu. Rev. Biochem.* **65,** 537-561.

14. Vanderkooi, J. M., Landesberg, R., Hayden, G. and Owen, C. S. (1977) *Eur. J. Biochem.* **81,** 339-347.

15. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X. and Bai, Y. (1998) *Acc. Chem. Res.* **31,** 737-744.

16. Shastry, M. C. R., Sauder, J. M. and Roder, H. (1998) *Acc. Chem. Res.* **31,** 717-725.

17. Hagen, S. J., Hofrichter, J. and Eaton, W. A. (1997) *J. Phys. Chem. B* **101,** 2352-2365.

18. Pierce, M. M. and Nall, B. T. (1997) *Prot. Sci.* **6,** 618-627.

19. Colón, W., Wakem, L. P., Sherman, F. and Roder, H. (1997) *Biochemistry* **36,** 12535-12541.

20. Scheidt, W. R., Kastner, M. E. and Hatano, K. (1978) *Inorg. Chem.* **17,** 706-710.

21. Sudha, B. P., Dixit, N., Moy, V. T. and Vanderkooi, J. M. (1984) *Biochemistry* **23,** 2103-2107.

22. Elias, H., Chou, M. H. and Winkler, J. R. (1988) *J. Am. Chem. Soc.* **110,** 429-434.

23. Pace, N. C., Shirley, B. A. and Thomson, J. A. (1990) in *Protein Structure: A Practical Approach*, ed. Creighton, T. F. (IRL Press, Oxford), pp. 311-330.

24. Cowan, J. A. and Gray, H. B. (1989) *Inorg. Chem.* **28,** 2074-2078.

25. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. and Gray, H. B. (1996) *Chem. Biol.* **3,** 491-497.

26. Tezcan, F. A., Crane, B. R., Winkler, J. R. and Gray, H. B. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5002-5006.

27. Wherland, S. and Gray, H. B. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2950-2954.

28. Akiyama, S., Takahashi, S., Ishimori, K. and Morishima, I. (2000) *Nature Struct. Biol.* **7,** 514-520.

29. Sauder, J. M. and Roder, H. (1998) *Folding and Design* **3,** 293-301.

30. Zhou, Y. and Karplus, M. (1999) *Nature* **401,** 400-403.

31. Thirumalai, D., Klimov, D. K. and Woodson, S. A. (1997) *Theor. Chem. Acc.* **96,** 14-22.

32. Mirny, L. A., Abkevich, V. and Shakhnovich, E. I. (1996) *Folding and Design* **1,** 103-116.

33. Lyubovitsky, J. G., Gray, H. B. and Winkler, J. R. (2002) *J. Am. Chem. Soc.* **124,** 5481-5485.

34. Finkelstein, A. V. and Shakhnovich, E. I. (1989) *Biopolymers* **28,** 1681-1694.

35. Gutin, A. M., Abkevich, V. I. and Shakhnovich, E. I. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1282-1286.

Appendix A

E. coli Expression and Purification of *R. palustris* cytochrome *c*′[†]

[†]The cytochrome *c*′ gene was cloned by Dr. Michele M. McGuirl. Site-directed mutagenesis, protein expression, and purification protocols were developed in collaboration with Lisa Van Hoozer and K. Cecilia Engman. Site-directed mutagenesis of L12H, M15H, and F32W/W72F proteins was performed by K. Cecilia Engman.

In order to prepare *R. palustris* cytochrome *c'* in large quantities, the cyt *c'* gene was cloned and co-expressed in *E. coli* with the cyt *c* maturation (*ccm*) gene cassette developed by Thöny-Meyer and coworkers [1, 2]. With this heterologous expression system, we are able to produce wild-type and mutant cyt *c'* on the tens- of-milligrams/liter-of-media scale.

Site-Directed Mutagenesis and Protein Expression

The DNA sequence of cyt *c'* (Figure A.1) was cloned into the Nde*I* and Bam*HI* sites of pET20b(+) (Novagen). The resulting plasmid (pETcp1) was electrochemically co-transformed with pEC86 [1] into *E. coli* strain BL21(DE3) and grown in TB media (1 L media = 12 g bacto-tryptone; 24 g yeast extract; 4 mL glycerol; 100 mL of potassium phophate buffer (23.1 g KH_2PO_4/ 125.4 g K_2HPO_4 in 1 L of H_2O); 900 mL of H_2O) in the presence of 100 mg/L ampicillin and 30 mg/L chloramphenicol. To improve the expression yield, a rare Arg codon in the periplasmic leader sequence was mutated from AGA to CGT. With this silent mutation, the expression level was increased to ≈10 mg/ L media from 200μg/L media. From mass spectral and protein sequence analyses (performed by the Protein/Peptide Microanalytical Laboratory at Caltech) of the wild-type protein expressed in *E. coli*, it was revealed that the amino-terminal Gln residue was post-translationally modified to pyroglutamic acid (Figure A.2). To circumvent this problem, the N-terminus was mutated to Ala (Q1A). Q1A is designated as the pseudo-wild-type protein for further DNA manipulation and protein study.

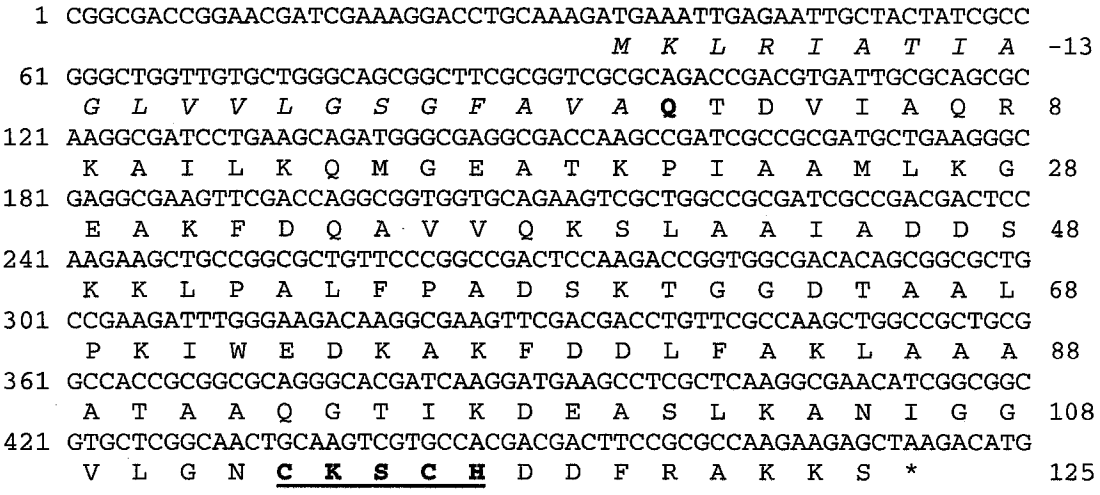


Figure A.1. DNA sequences and deduced protein sequences for *R. palustris* cyt *c'*. Periplasmic leader sequences are in italics with the first amino acid of the mature protein in bold. Heme binding motif (CXXCH) is in bold and underlined. The amino acids are numbered with reference to the mature protein sequence.

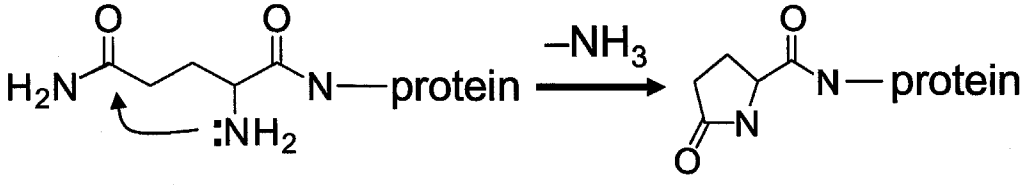


Figure A.2. N-terminal modification of wild-type cytochrome *c'*. The Gln residue cyclizes and forms pyroglutamic acid.

All site-directed mutagenesis was performed using a Quickchange kit (Stratagene) with slight modification of the protocols. The primers are listed below (5' to 3', the mutation site is underlined):

M15L

m15lcpf	ATC CTG AAG CAG CTG GGC GAG GCG ACC
m15lcpr	GGT CGC CTC GCC CAG CTG CTT CAG GAT

M25A

m25acpf	CCG ATC GCC GCG GCG CTG AAG GGC GAG
m25acpr	CTC GCC CTT CAG CGC CGC GGC GAT CGG

E17C

e17ccpf	CTG AAG CAG ATG GGC TGC GCG ACC AAG CCG ATC GCC
e17ccpr	GGC GAT CGG CTT GGT CGC GCA GCC CAT CTG CTT CAG

L12H

l12hcpf	CGC AAG GCG STC CAC AAG CAG ATG GGC
l12hcpr	GCC CAT CTG CTT GTG GAT CGC CTT GCG

M15H

m15hcpf	CG ATC CTG AAG CAG CAC GGC GAG GCG ACC
m15hcpr	GGT CGC CTC GCC GTG CTG CTT CAG GAT CG

F32W

f32wcpf	G GGC GAG GCG AAG TGG GAC CAG GCG GTG GTG
f32wcpr	CAC CAC CGC CTG GTC CCA CTT CGC CTC GCC C

W72F

w72fcpf	G CTG CCG AAG ATT TTC GAA GAC AAG GCG
w72cpr	CGC CTT GTC TTC GAA AAT CTT CGG CAG C

The PCR mixture used is as follows:

2 μL 10x buffer
25 ng double-stranded DNA (Q1A)
50 ng primer (forward)
50 ng primer (backward)
0.2 μL 10mM dNTPs
H_2O to total volume of 20μL
0.4 μLof PfuTurbo (DNA polymerase)

PCR was performed in a DNA thermocycler (MJ Research) with the programmed sequence listed below:

Step 1	1 cycle	95 °C	30 s
Step 2	16 cycles	95 °C	30 s
		55 °C	60 s
		68 °C	8 min 24 s

Following digestion, the desired PCR product was transformed into an *E. coli* cell line (XL-1 Blue, Stratagene) for sequencing. The digestion and transformation protocols are listed below:

<u>digestion</u>

1) add 1 μL DpnI (20U/μL)

2) incubate at 37 °C for 1 h

3) heat for 20 min at 80 °C

4) store at -20 °C or transform immediately

<u>transformation</u>

1) pre-chill PCR tubes

2) set thermocycler to 42 °C

3) gently defrost cells from -70 °C freezer (keeping them on ice)

4) mix 25 μL cells with 1 μL DNA (from PCR or 5 mL miniprep)

5) let tube sit at least for 5 min on ice

6) apply heat pulse for 40 s at 42 °C

7) let tube sit for 2 min on ice

8) add 200 μL of NZY$^+$ broth (RT) and transfer to 14 mL Falcon round bottom tubes

9) shake at 37 °C for 1 h

10) plate* (20 μL and 200 μL) and grow overnight at 37 °C on LB plates

*for pETcp plasmid only use 100 μg/mL of ampicillin and for co-transformation of pETcp and pEC86 use 100 μg/mL of ampicilin and 30 μg/mL of chloramphenicol

Then, a single colony is picked and grown in LB media (10 g bactotryptone; 5 g yeast extract; 10 g NaCl in 1 L water) overnight (≈18 h). Extraction of DNA is achieved by using a QIAprep Spin Miniprep kit (Qiagen). The resulting DNA is sequenced to confirm the desired mutations (performed by the Caltech DNA Sequencing Core Facility). Finally, the plasmid is co-transformed with pEC86 into a BL21(DE3) cell line following procedures described above. Due to the heterologous expression, there is significant colonial variation in protein expression. An initial screen of the colonies (single colonies were picked and grown overnight in 1-mL of LB or TB media) is used to select for protein expression. Colonies are selected by the pink color of the cell pellet.

Protein is prepared by first growing pre-cultures (10-15 mL TB media) for 12 h; prior to innoculation (1:100) and growth of 1 to 1.5-L cultures (16-20 h), the

cells were spun down and rinsed with fresh media. Induction with IPTG does not seem to dramatically improve the yield of expression.

Protein Purification

Early preparations of cyt *c'* (Q1A, Q1A/M15L, Q1A/M25A, and Q1A/M15L/M25A) were purified from total cell lysis. It is also possible to extract the protein (Q1A, Q1A/F32W/W72F, Q1A/L12H, and Q1A/L15H) using osmotic shock. Although complete cell lysis allows the extraction of more protein, in certain mutants it proved to be more difficult to obtain pure samples. The following protocol was used for cell lysis:

1) resuspend cell pellet in cell lysis buffer (80 mL/g wet cell pellet, 20% sucrose solution containing 1mM EDTA and 30 mM Tris, pH 8)

2) add 10 mg of lysoyme, 3.5 mg of DNase, and 3.5 mg of RNase per g of wet cell paste and stir at 4 °C for 1-2 h

3) centrifuge at 8000 rpm for 25-30 min and collect supernatant (yellowish-red solution)

4) adjust the pH to 6 with1N HCl and dilute solution with cold water to obtain an ionic strength of 5-7 mM prior to column chromatograph

Osmotic shock procedures are described below:

1) resuspend fresh cell pellet in sucrose solution (40-50 mL/g wet cell pellet, 20% sucrose, 1 mM EDTA, and 30 mM Tris, pH 8)

2) centrifuge at 5000 rpm for 10 min and discard supernatant

3) resuspend cell pellet in cold buffer solution (5 mM $MgSO_4$ and 2 mM sodium phosphate, 30-40 mL/g wet cell paste), let sit on ice for 5-10 min

4) centrifuge at 8500 rpm for 20-25 min and collect supernatant

5) adjust the crude protein solution to pH 6 with1N HCl and dilute solution with cold water to obtain an ionic strength of 5-7 mM prior to column chromatography After cell lysis/osmotic shock, the crude protein solution is loaded onto a CM Sepharose Fast Flow column (Pharmacia) equilibrated with 5 mM sodium phosphate buffer, pH 6. The column was washed with at least 10x column volume of 5 or 10 mM sodium phosphate and cyt *c'* was eluted with 10 mM sodium phosphate and 50 mM NaCl, pH 6. Protein was further purified on a Mono S 10/10 column (Pharmacia) with an FPLC system (linear gradient from 0-100 mM NaCl in 10 mM sodium phosphate buffer, pH 6) (Figure A.3 (Q1A/F32W/W72F), A.4 (Q1A/L12H), A.5(Q1A/M15H)). Identity and purity of the protein were checked with electrospray mass spectrometry and SDS gel electrophoresis (Table A.1). Protein digestion and sequencing also were performed on the Q1A/F32W/W72F mutant to ensure that the correct mutations were introduced (Figure A.6).

cytochrome *c'*	calculated mass	measured mass
wild type	13,747	13,746
wild type (pyroglutamic acid)	13,732	13,733
Q1A	13,690	13,691
Q1A/L12H	13,714	13,716
Q1A/M15L	13,671	13,670
Q1A/M15H	13,696	13,697
Q1A/E17C	13,663	13,665
Q1A/M25A	13,629	13,630
Q1A/M15L/M25A	13,611	13,613
Q1A/F32W/W72F	13,690	13,691

Table A.1. Mass spectral data for all proteins expressed in *E. coli*.

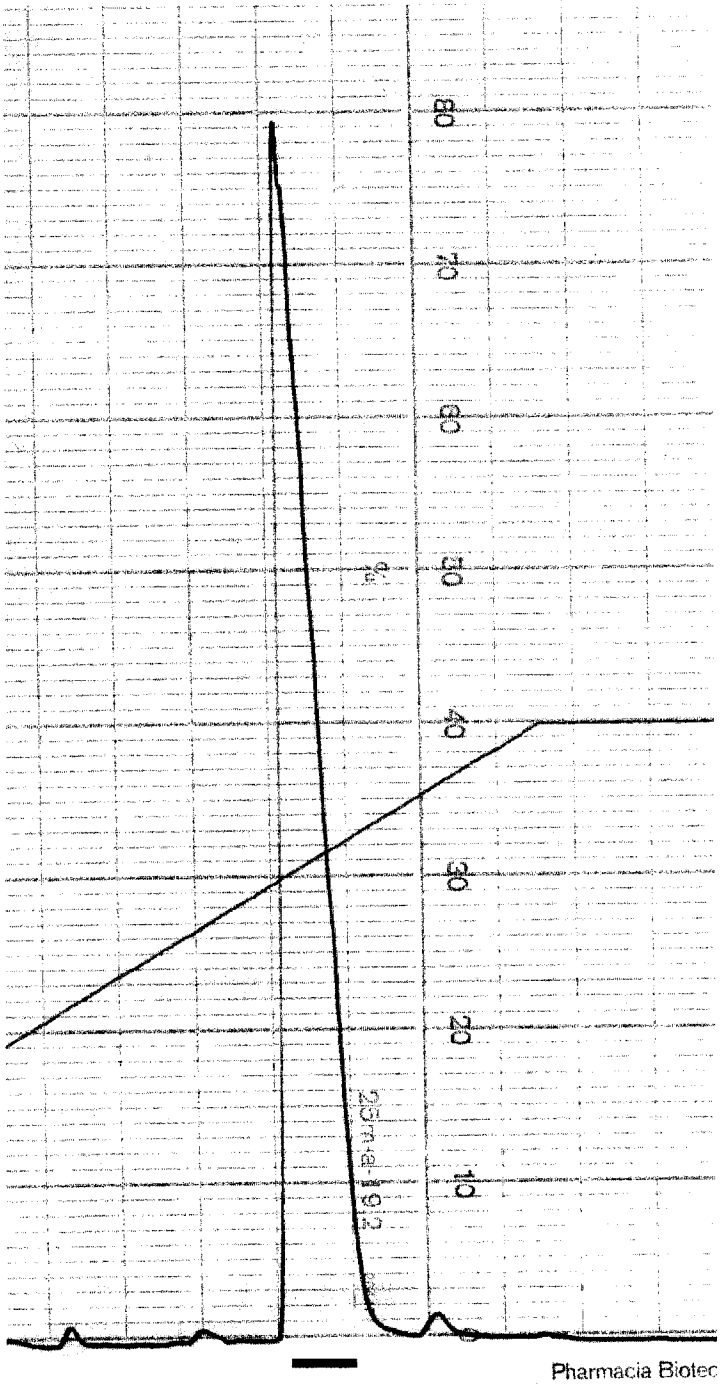


Figure A.3. FPLC chromatogram of Q1A/F32W/W72F cyt *c'*.

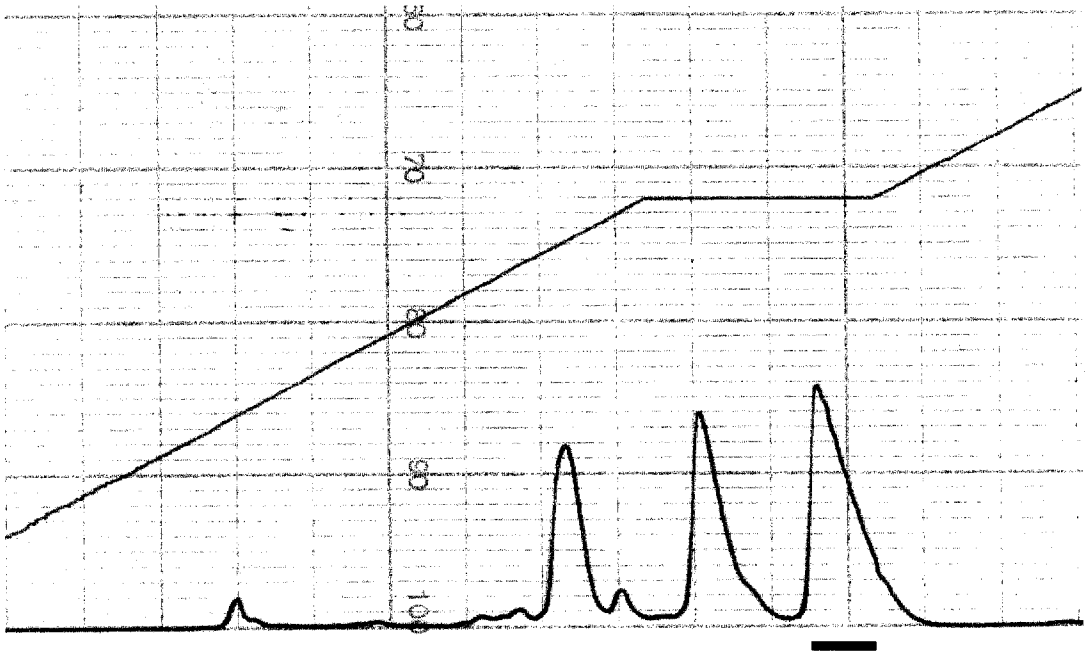


Figure A.4. FPLC chromatogram of Q1A/L12H cyt *c'*.

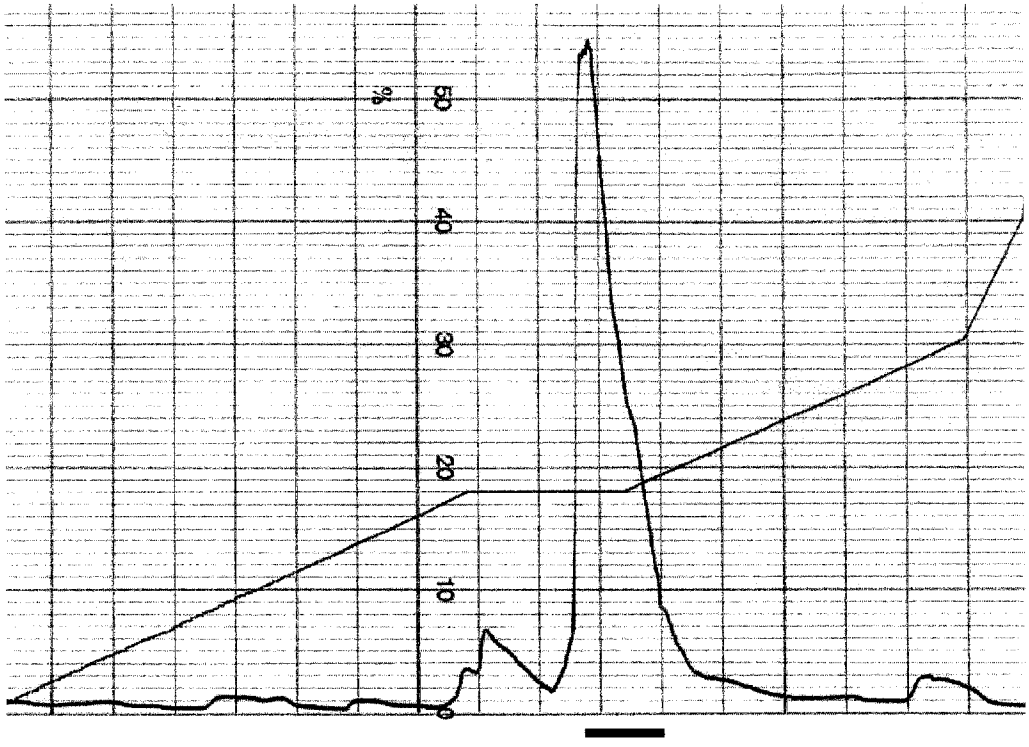


Figure A.5. FPLC chromatogram of Q1A/M15H cyt *c'*.

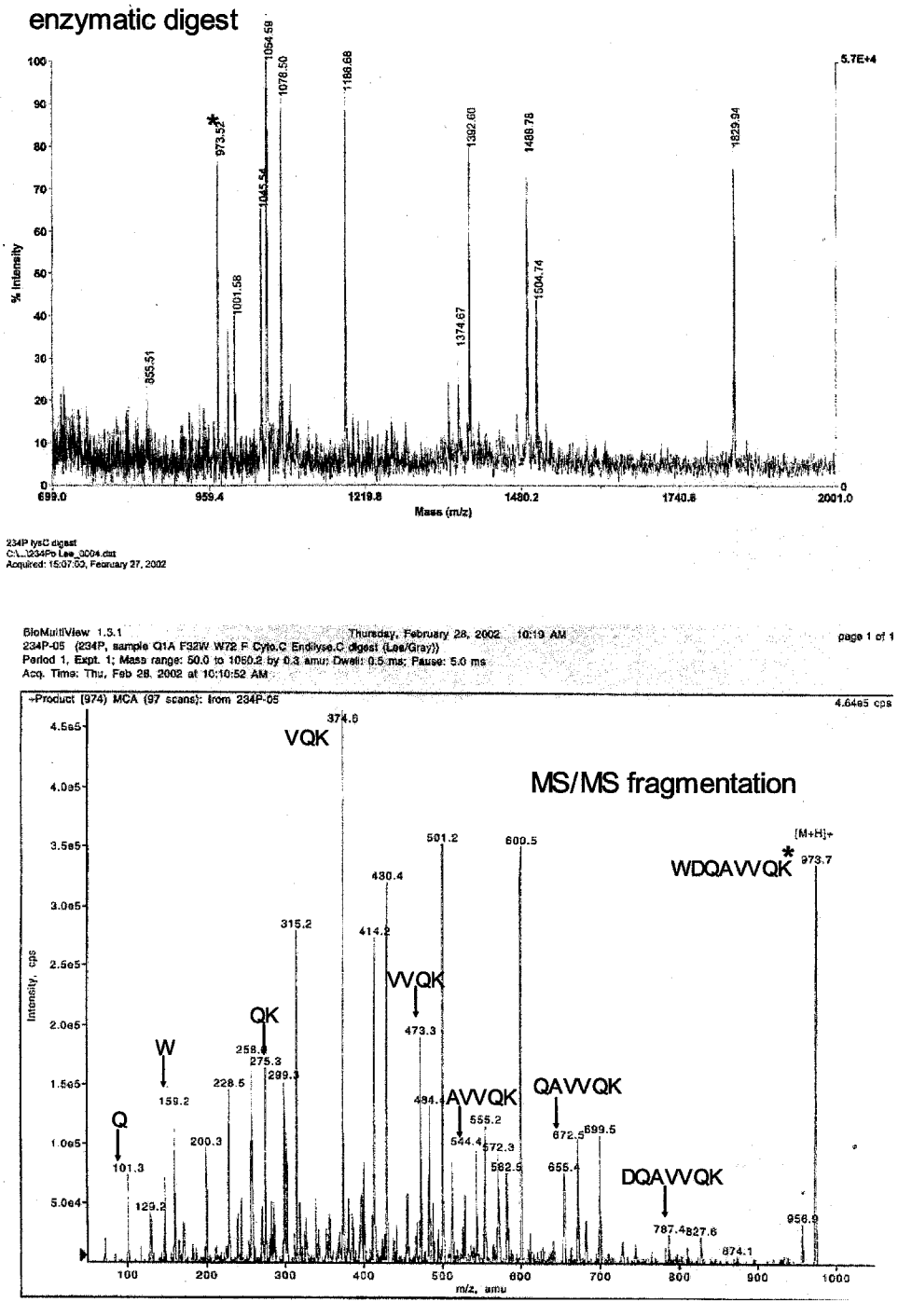


Figure A.6. Mass spectral analyses of enzymatic digest (endolys-C) of Q1A/F32W/W72F cyt *c'*. Electrospray mass spectrum of the digest (top) and ms/ms fragmentation of the peptide W_{32}-K_{39} (bottom).