

Chapter 1

Background

1.1 Our Approach

We are interested in dynamic processes in proteins that occur over a broad range of timescales: protein folding, intrachain diffusion in unfolded proteins, and electron tunneling through proteins. Small-scale motions such as side-chain rotations in proteins occur in picoseconds, whereas large-scale motions such as intrachain diffusion occur in microseconds. The entire folding reaction for a protein can take microseconds to seconds, depending on the protein.

We probe these dynamics by attaching a fluorescent photosensitizer to the protein. We investigate the process on a suitable laser system, which must excite the sample significantly faster than the timescale of the process. For example, we monitor electron transfer events that occur on a microsecond timescale with a nanosecond laser and a long-lived (microseconds) luminescent ruthenium complex. Conformational rearrangements in unfolded proteins, which occur on a similar timescale, are also studied with this system.

We measure intramolecular distances in proteins using a picosecond laser

system and an organic fluorophore with an excited state lifetime of tens of nanoseconds. Since large-scale motions occur significantly slower than the nanosecond regime, we capture a "snapshot" of the ensemble of protein conformations. We can follow the folding process by collecting snapshots of protein ensembles over milliseconds; this is carried out by coupling the picosecond laser system to a continuous flow mixer to trigger the folding reaction.

1.2 Protein Dynamics

1.2.1 Protein Folding

Protein folding is an intriguing biological problem with significant societal impacts. Protein misfolding and aggregation have been implicated in several diseases, including Alzheimer's, Parkinson's, and prion diseases [1]. A rigorous understanding of protein folding mechanisms would enable development of new pharmaceuticals to treat or prevent these diseases. For example, a drug developed by Genentech for prevention of Alzheimer's disease is currently in trial [2]. The drug is designed to target plaques of the protein beta amyloid and their oligomeric precursors that are thought to cause neurodegeneration.

Protein folding is also, in and of itself, an interesting biophysical problem. Proteins are heterogeneous polymers with unique primary amino acid sequences that dictate the secondary and tertiary structure, or the local and global intramolecular contacts, respectively. Many proteins fold *in vitro*, or outside of the cell,

from an ensemble of unfolded states to a single native state without the help of chaperones. Scientists are attempting to understand this process from multiple angles—through structure prediction from amino acid sequences [3, 4], folding kinetics and equilibrium measurements [5], computational modeling of protein dynamics [6], and unfolded state dynamics models and experiments.

The energy landscape theory [7, 8, 9, 10], developed by Wolynes, Bryngelson, and Onuchic, is a conceptual framework for the energetics of protein folding. In contrast to a simple free energy diagram with a discrete reaction pathway, multiple folding pathways are available for a protein to reach its global minimum, the native (folded) state. The principle of minimal frustration [11] describes how proteins with numerous accessible states often can fold efficiently to the native state, and why some proteins do not.

In minimally frustrated proteins, contacts that stabilize the native structure are strongly favored during folding, so that molecules appear to be funneled to the native state. Each newly formed native contact reduces the conformational search for the native structure. Frustration in the energy landscape will appear as kinetic traps, or energy barriers that can slow folding. A rough landscape, in which non-native contacts are energetically competitive with native contacts, can lead to deviations from simple first-order kinetics. Intermediates arise when there is a build-up of conformations due to a large energy barrier. If region(s) of the protein are more frustrated than others, a sequential folding mechanism with region-specific folding kinetics can result. It is, thus, necessary to probe multiple

regions of the protein for a global mechanism.

Although predicting folding mechanisms is very challenging, some important factors guiding folding have been identified [12, 13, 3]. These factors include topology (protein fold), the number of amino acids between tertiary contacts, chain length, and properties of amino acid sidechains, e.g., hydrophobicity. It is difficult to determine the impact of each parameter as they are interdependent. Therefore, we will compare proteins to other members of their family that share a common fold.

1.2.2 Intrachain Diffusion

We investigate tertiary contact formation rates in unfolded cytochrome *cb*₅₆₂. As formation of tertiary contacts is a necessary step for folding, the folding speed limit is set by the conformational dynamics in the unfolded state [14, 15, 16, 17]. When studying unfolded dynamics, frustration manifests as internal friction, slowing intrachain diffusion in regions of the polypeptide.

The Gray Group previously employed electron transfer quenching to estimate contact formation rates in denatured cytochrome *c* [15] and in α -synuclein [18, 19], the polypeptide implicated in Parkinson's disease. We have now extended our work on contact dynamics to include denatured cytochrome *cb*₅₆₂. Adiabatic electron transfer requires a close-contact encounter complex. If the electron transfer reaction is diffusion-limited, then the electron transfer rate constant corresponds to contact formation. Photosensitizers are used to probe the kinetics

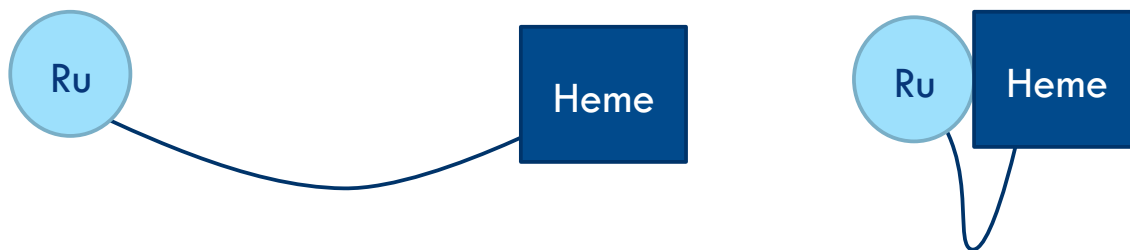


Figure 1.1: Contact quenching to probe intrachain diffusion. We examine formation of transient contacts between the heme and $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ complexes that have been covalently attached at various positions in unfolded cytochrome cb_{562} .

of intrachain diffusion (Figure 1.1).

1.2.3 Electron Tunneling

Electron transfer in metalloproteins is an important phenomenon in biology that occurs by quantum mechanical tunneling. Electron tunneling through protein medium occurs in nanoseconds to microseconds when there is a high driving force. We are interested in the factors that determine the electron transfer rate, especially the tunneling distance and the types of bonds in the pathway. We probe the kinetics of electron tunneling in folded cytochrome cb_{562} from a ruthenium photosensitizer, attached to various positions in the protein, to the heme using time-resolved spectroscopy (Figure 1.2).

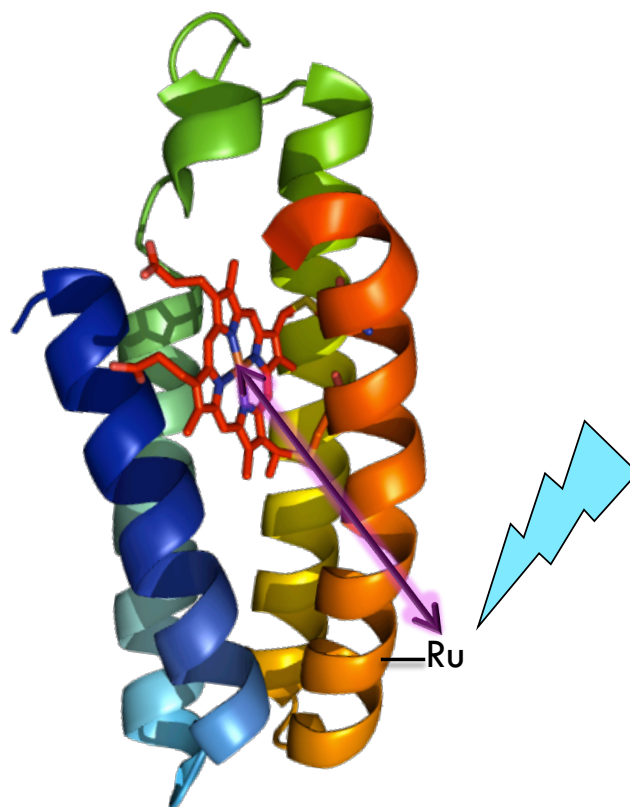


Figure 1.2: We measure ruthenium-heme electron tunneling rates in cytochrome *cb*₅₆₂ following photoexcitation of $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ complexes that have been covalently attached at various positions in the protein.

1.3 Techniques to Probe Protein Dynamics

1.3.1 Time-Resolved FRET

Fluorescence energy transfer, or FRET, provides information about the distance, or the distribution of distances for an ensemble between two fluorophores. We use this technique to measure intramolecular distances in proteins and characterize heterogeneous ensembles of conformations, providing a "snapshot" of the different populations present at a point in time.

FRET is a nonradiative transfer of energy between two fluorophores that can occur when there is a spectral overlap between donor fluorescence and acceptor absorption. The efficiency of this process is described by Förster's model of dipole-dipole energy transfer: $E = R_0^6 / (R_0^6 + r_{DA}^6)$, where r_{DA} is the distance between the fluorophores and R_0 is the Förster distance, dependent on the spectral overlap and the dipole orientations [20].

In time-resolved FRET (trFRET), energy transfer is observed as a decrease in the lifetime of the donor fluorescence. The fluorescence decay is fit to an exponential function with each decay rate constant k corresponding to a donor-acceptor (D-A) distance (Figure 1.3). A single-exponential decay would suggest a homogeneous solution, whereas a multiexponential decay would indicate that there is a heterogeneous ensemble of molecules in different conformations.

Typically, proteins must be modified to incorporate fluorescent probes for FRET studies. (The fluorescent amino acid tryptophan can also serve as a probe.)

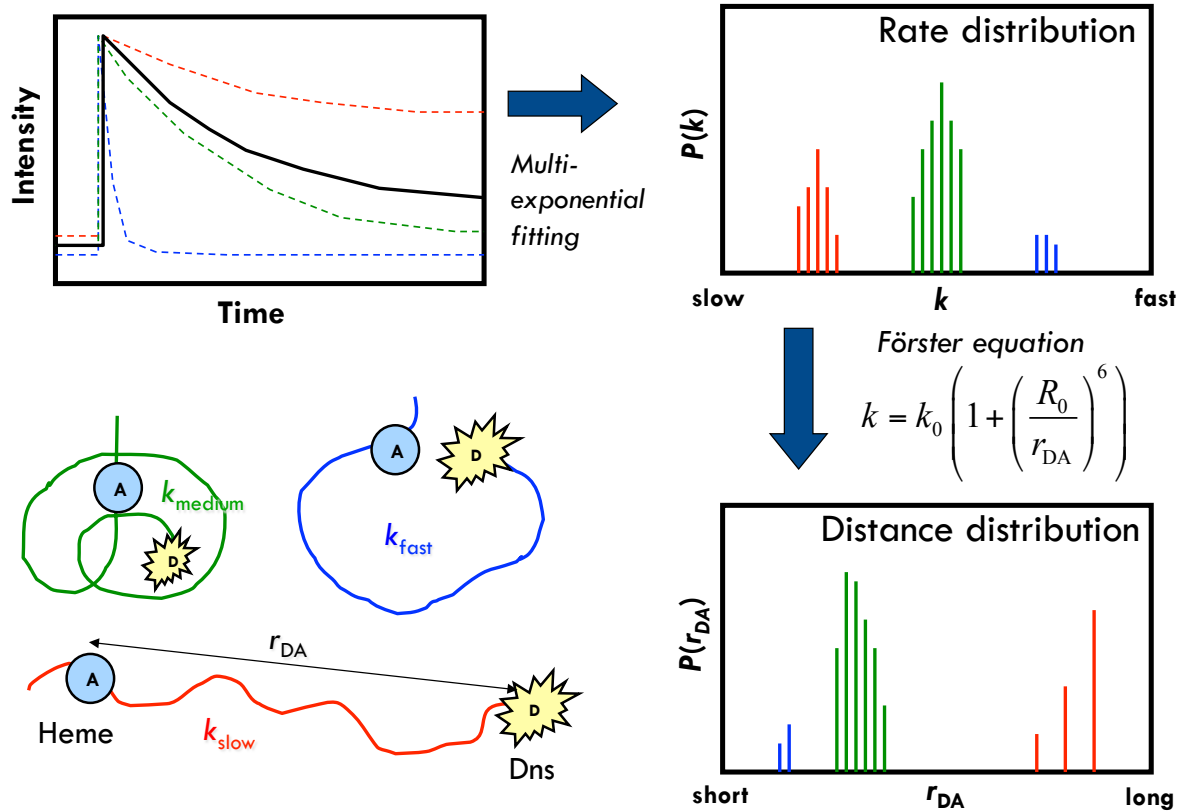


Figure 1.3: Time-resolved fluorescence energy transfer: Illustration of data analysis and results for a solution containing compact, intermediate, and extended protein conformations. The three populations contribute to the multiexponential luminescence decay, shown in black. Image courtesy of Seiji Yamada.

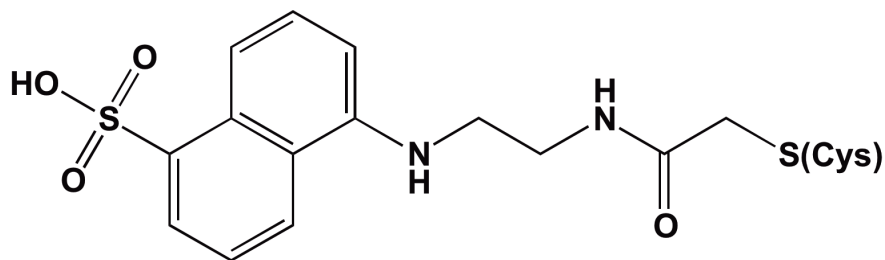


Figure 1.4: Dansyl fluorophore used in FRET studies.

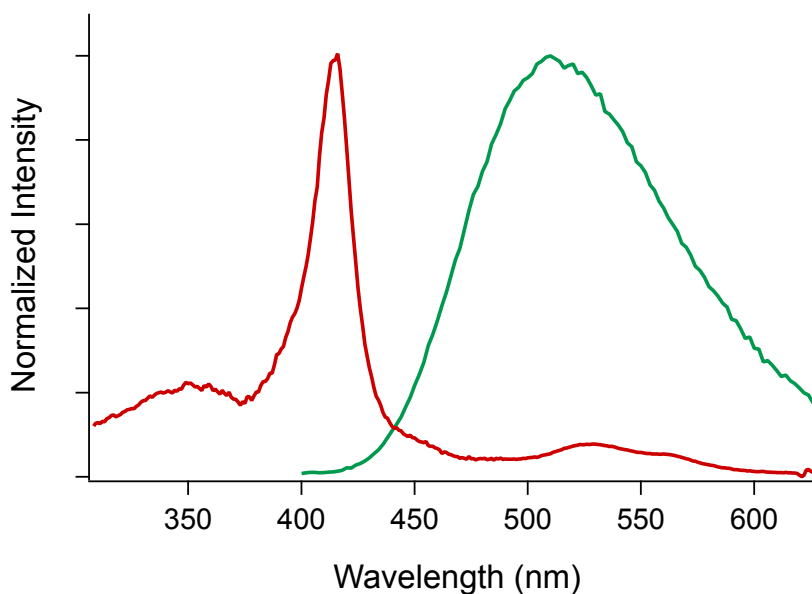


Figure 1.5: Spectral overlap of dansyl fluorescence (green, 355 nm excitation) and cytochrome *cb*₅₆₂ heme absorbance (red).

We utilize the small organic fluorophore dansyl, attached covalently to cysteine residues, as the donor (Figure 1.4). The heme in the protein is the acceptor in our system. The dansyl-heme spectral overlap is shown in Figure 1.5. The Dns-heme pair has a Förster critical length (R_0) of 39 Å [21], which is optimum for measuring distances from 12–60 Å.

1.3.2 Continuous Flow Mixing to Trigger Folding

Protein folding is triggered by a change from conditions that disfavor folding, e.g., high temperature, acidic pH, or denaturants, to conditions that favor folding. Chemical denaturants, including guanidine hydrochloride (Gdn) and urea, are proposed to denature proteins by interacting with and disrupting the hydrogen bonding of the protein backbone and/or altering the intermolecular bonding of water [22, 23]. To trigger folding, denatured protein is mixed rapidly with buffer to decrease the denaturant concentration. In stopped flow mixing, the two solutions are mixed into a cell with high pressure. The "dead time," or time needed for complete sample mixing, of these instruments ($\sim 1\text{--}3$ ms [24]) does not allow for resolution of fast folding processes. Newer models with smaller volumes and pneumatic drives reach dead times of 0.5–1 ms.

In contrast, continuous flow mixing allows for the resolution of submillisecond folding, including burst phases and early intermediates [25]. In continuous flow mixing, two solutions are rapidly and continuously mixed together in a microfluidic channel [26, 27]. The Gray Group has utilized continuous flow mixing to reveal the early phases of the folding of four-helix bundle cytochromes cb_{562} and c' [28, 29]. In this thesis, we resolve the fast formation of native cytochrome cb_{562} and the folding intermediate of cytochrome c_{552} by ultrafast continuous flow mixing.

We couple the mixer (150 μs dead time) with a picosecond streak camera to simultaneously monitor multiple mixer positions, corresponding to time points in

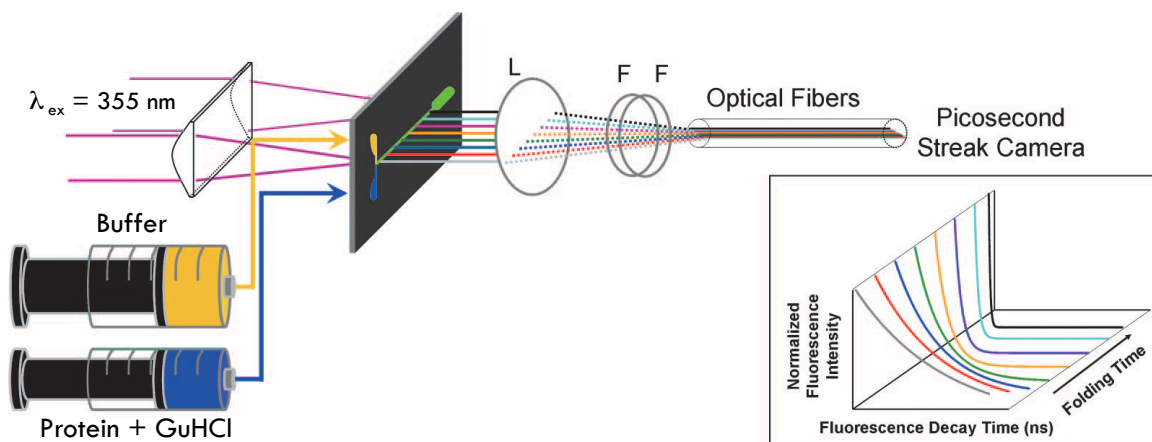


Figure 1.6: A continuous flow mixer is coupled to trFRET measurements. Folding is initiated upon mixing (green) of the unfolded protein (blue) and buffer (yellow). As the solution travels down the microfluidic channel, the protein has more time to fold; each position corresponds to a time point in the folding reaction. The excitation beam is focused on the channel, and fluorescence decays are collected with a picosecond streak camera. The detector allows for simultaneous data collection for multiple mixer positions (folding times). Modified with permission from Kimura et al. [28].

the folding reaction (Figure 1.6). The collected fluorescence decays (trFRET data) provide information about the changing distributions of protein conformations during folding, including the heterogeneity of structural ensembles. Measurement of a single donor-acceptor pair reveals the timescale for the formation of the structure in the corresponding regions of the protein. Fitting the time course of the mean donor-acceptor distance to an exponential function elucidates whether there is two-state folding with a single energetic barrier or if there are kinetic intermediates. Comparing the rate constants for two donor-acceptor pairs indicates whether the regions fold simultaneously or sequentially. Insights about folding mechanisms, whether there is a cooperative folding mechanism, a single sequential pathway, or multiple pathways, can be gleaned from this method.

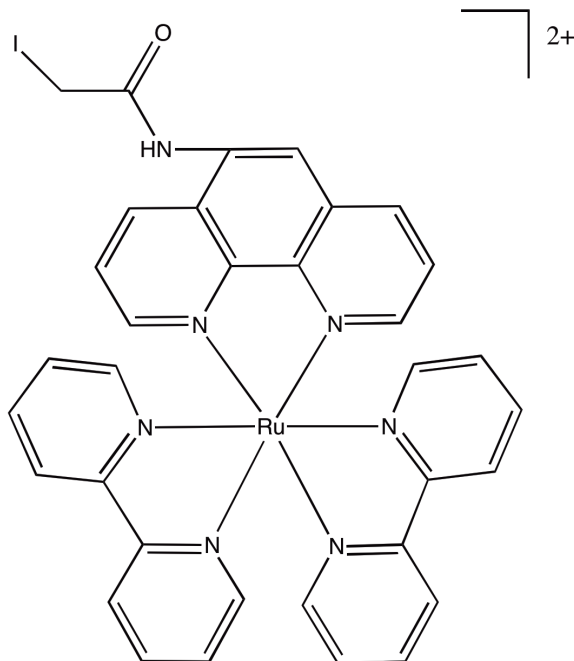


Figure 1.7: Ruthenium photosensitizer: $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$.

1.3.3 Photochemical Triggers of Electron Transfer

Long-lived excited state metal complexes enable us to study electron transfer reactions on the microsecond timescale. The Gray Group has developed ruthenium complexes to probe electron transfer reactions in proteins [30]. We attach the complex $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$ (Figure 1.7) to proteins site-specifically at mutated cysteines [31]. Electron transfer to the protein's heme occurs via adiabatic contact quenching or nonadiabatic electron tunneling. This reaction has a high driving force and is essentially activationless [32]. We photoexcite the Ru label in the MLCT absorption band (Figure 1.8) and observe electron transfer by monitoring the quenching of its luminescence or the absorption of transient species (Ru^{III} and Fe^{II} , Figure 1.9).

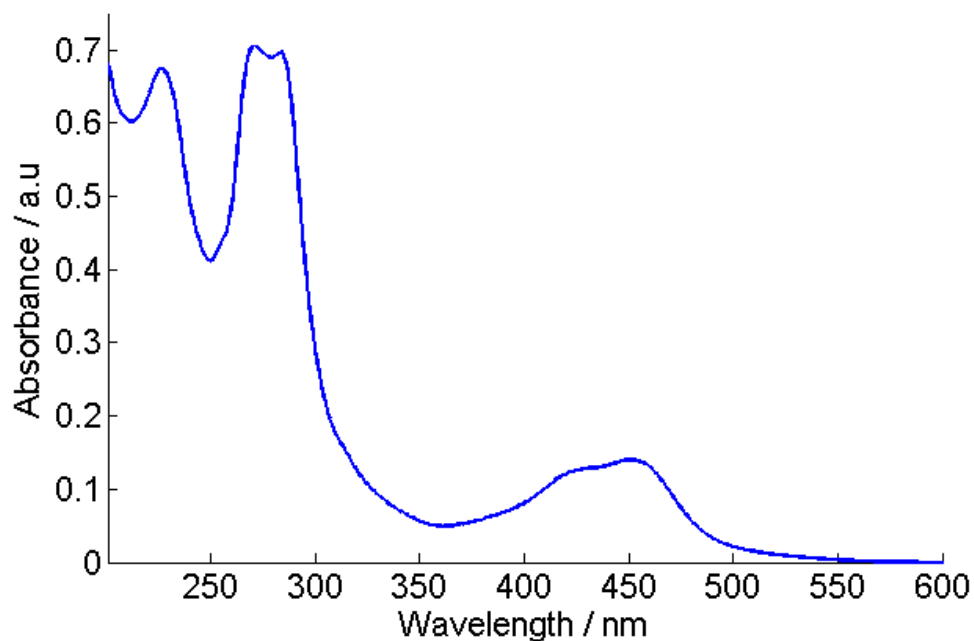


Figure 1.8: Absorption spectrum of $[\text{Ru}(\text{bpy})_2(\text{IA-phen})]^{2+}$. The label is photoexcited at 480 nm where there is minimal overlap with the heme Soret absorbance.

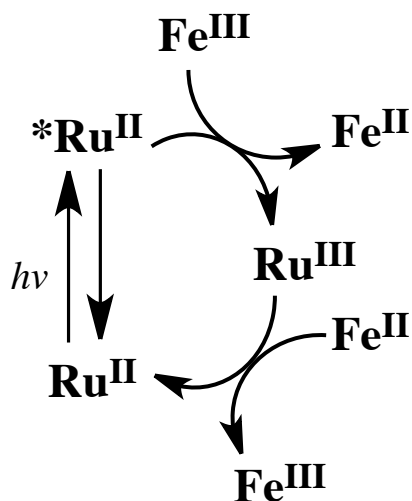


Figure 1.9: Ruthenium photochemistry: Triggering electron transfer to heme Fe^{III} .

1.4 Proteins of Interest

1.4.1 Cytochrome cb_{562} , a Four-Helix Bundle Cytochrome

The Gray Group studies the folding mechanisms of a family of four-helix bundle cytochromes, including cytochromes c' , c_{556} , and b_{562} [33, 34, 35, 36, 37, 38, 28, 29, 39, 40]. While protein families typically have fairly conserved sequences and topologies, the members of this family are unique in that they have almost superimposable structures (Figure 1.10) [41, 42, 43] but unusually low sequence homology of only 10-20% [44]. Similar topology leads one to predict that the proteins should have similar folding kinetics; however these the four-helix bundle cytochromes have widely divergent folding kinetics. This rare characteristic creates a unique opportunity to study the relationship between amino acid sequences and folding pathways in proteins with similar topologies.

The folding kinetics of cytochrome c' are highly diverse, with time constants ranging from milliseconds to seconds [28, 38]. Probe-dependent folding kinetics and heme transient absorption spectra reveal deviations from two-state folding behavior and a sequential folding mechanism. Ferricytochrome c' has heterogeneous conformations during refolding, including a population that is more compact than the native state, suggestive of hydrophobic collapse during folding [28]. The large degree of frustration is partially attributed to proline isomerizations [38]. Folding of ferrocycytochrome c' is further slowed by methionine misligations of the heme [34].

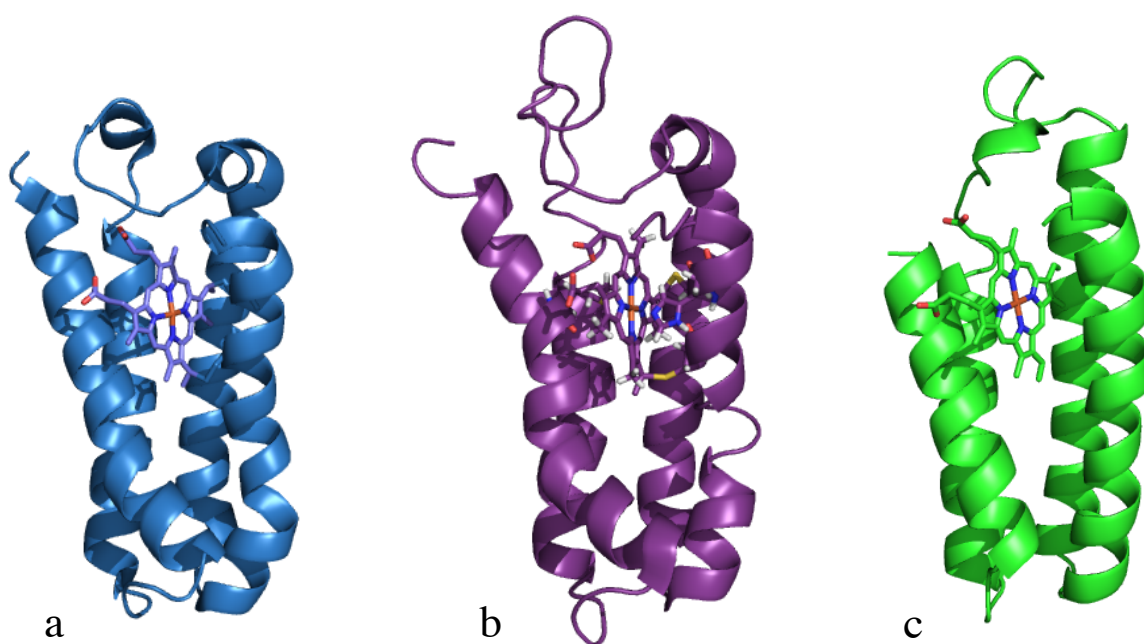


Figure 1.10: Structures of four-helix bundle cytochromes. Structures cytochrome c' and cytochrome b_{562} have a backbone atom RMSD of 3.4 Å, but only 15% sequence homology. Despite their conserved topology, the four-helix bundle cytochromes have highly divergent folding kinetics. (a) Cytochrome c' from *R. palustris* (PDB 1A7V). (b) Cytochrome c_{556} from *R. palustris* (PDB 1S05). (c) Cytochrome b_{562} from *E. coli* (PDB 256B).

On the other hand, cytochrome b_{562} folds the fastest among the four-helix bundle cytochromes with simple first-order folding kinetics [33]. Folding of ferrocycytochrome b_{562} was triggered by electron injection to the oxidized iron center in slightly denaturing conditions, taking advantage of stability differences between the ferric and ferrous oxidation states. Folded protein was formed on a submillisecond timescale; however, folding rate determination was limited by irreversible heme dissociation. Furthermore, evidence suggests that the measured rate constant may only represent the folding event for a small population of a native-like conformation present in low concentrations of denaturant [33, 45].

To overcome the complication of heme dissociation during refolding, cytochrome b_{562} with its heme bound solely by heme-ligand interactions was modified to cytochrome cb_{562} with additional *c*-type covalent heme linkages to the protein backbone [37]. Cytochromes c' and cb_{562} now share a common topology and similar heme environment so that the influence of the primary sequences on folding kinetics can be isolated from these factors. Cytochrome cb_{562} has been shown by Trp59-heme trFRET studies to fold on the millisecond timescale with simple first-order folding kinetics [29]; we further characterize cytochrome cb_{562} 's folding mechanism in this thesis.

Cytochrome cb_{562} is an excellent model α -helical protein. It is very stable and refolds *in vitro*. Due to its high degree of symmetry, it has been utilized as a building block for macromolecular protein assemblies [46]. In addition, it is thought to be an electron transfer protein, although its specific function is

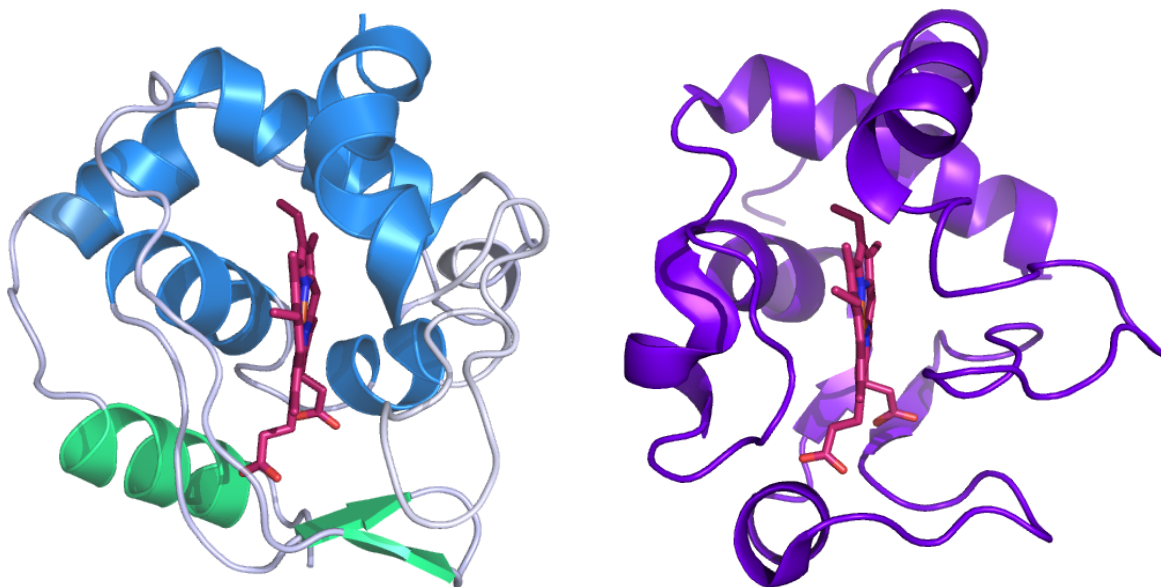


Figure 1.11: Structures of class-I c-type cytochromes c_{552} from *T. thermophilus* (blue/green, PDB 3VNW) and yeast *c* (violet, PDB 1YCC). The N- and C-terminal helices cross each other (apexes). The structural features in green are distinct to cytochrome c_{552} .

unknown. We have shown that Zn-cytochrome b_{562} is photoconductive [47], and in this thesis, we report electron tunneling rates for Ru-modified cytochrome cb_{562} .

1.4.2 Cytochrome c_{552} , a Class I C-Type Cytochrome

Cytochrome c_{552} is a small soluble, thermostable protein containing a single heme bound covalently via two cysteines and axially ligated by His15 and Met69. Cytochrome c_{552} is found in the hyperthermophile *Thermus thermophilus* with an optimal growth temperature of $\sim 65^{\circ}\text{C}$ that was discovered in a Japanese hot spring. It is the natural electron donor for cytochrome ba_3 oxidase during the terminal step of respiration [48].

Cytochrome c_{552} belongs to the class-I c-type cytochrome family, which includes

yeast cytochrome *c*. Cytochromes in this family share the common structural feature where the N- and C-terminal helices cross each other (Figure 1.11). However, cytochrome *c*₅₅₂ has the distinct structural components of a β -sheet around the heme propionate and extra helices in the C-terminal region (Figure 1.11, green region). We are interested in the role of these distinct structural features and their possible correlation to the thermostability of the protein.

Brunori and coworkers examined the folding pathway of this protein by native tryptophan (Trp91) fluorescence kinetics and revealed that there is an on-pathway folding intermediate [49]. In this work, Trp91 probed structures near the crossed helices of cytochrome *c*₅₅₂, but did not report on the heme β -sheet or the C-terminal helices. We would like to further characterize the folding intermediate structurally.

1.5 Thesis Overview

We gain insight into cytochrome cb_{562} 's dynamics and folding mechanism in Chapters 3–5. In Chapter 3, we spectroscopically characterize cytochrome cb_{562} 's structural conformations, heme environment, and stability to chemical denaturants. In Chapter 4, we probe the unfolded dynamics and intrachain diffusion of cytochrome cb_{562} by electron transfer. In Chapter 5, we investigate the folding kinetics of cytochrome cb_{562} by trFRET and microfluidic mixing.

We explore the folding of cytochrome c_{552} in Chapters 6–7. We characterize an equilibrium unfolding intermediate of cytochrome c_{552} by trFRET and improve our understanding of the physical basis for its thermostability in Chapter 6. In Chapter 7, we resolve cytochrome c_{552} 's folding intermediate by trFRET and microfluidic mixing.

The findings from the protein folding studies are summarized in Chapter 8. Finally, we investigate the effects of structural changes on electron tunneling rates in cytochrome cb_{562} (Chapter 9).