



Publicly Accessible Penn Dissertations

1-1-2013

The Physical Chemistry Underlying the Assembly and Midpoint Potential Control in a Series of Designed Protein-Maquettes

Lee Andrew Solomon

University of Pennsylvania, solomoan@mail.med.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>



Part of the [Biochemistry Commons](#), and the [Biophysics Commons](#)

Recommended Citation

Solomon, Lee Andrew, "The Physical Chemistry Underlying the Assembly and Midpoint Potential Control in a Series of Designed Protein-Maquettes" (2013). *Publicly Accessible Penn Dissertations*. 802.
<http://repository.upenn.edu/edissertations/802>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/802>
For more information, please contact libraryrepository@pobox.upenn.edu.

The Physical Chemistry Underlying the Assembly and Midpoint Potential Control in a Series of Designed Protein-Maquettes

Abstract

In nature, oxidoreductase proteins are responsible for many enzymatic processes critical to life. These proteins often rely on the presence of non-proteinaceous cofactors to take part in the enzymatic function. The most common, central to my thesis, is heme B. Depending on the protein environment, this cofactor can take part in functions as diverse as electron transfer (cytochromes), oxygen transport (hemoglobins), oxygen reduction (oxidases), carbon-hydroxylation (oxygenases), and superoxide production (NADH oxidase).

In natural oxidoreductases, determination of the course and rates of heme-protein association, what barriers are encountered, what affinity is achieved, and what are the oxidation-reduction potentials, is critical for understanding the rules of assembly and function of the different activities performed. In the growing field of research attempting to make man-made oxidoreductases, the same understanding is required for progress to be made toward construction of novel enzymes. However, this understanding is still out of reach in natural oxidoreductases because of the immense complexity of natural proteins, while for man-made designs progress has only recently reached a point where an in-depth systematic study can be contemplated.

My thesis states: Simple non-natural proteins (maquettes) designed from first principles to ligate heme, can be used to uncover the factors derived from the oligomeric and structural state of related maquette and also derived from porphyrin variants of heme B, that govern rates of incorporation and ligation of heme B into a maquette. Maquettes are ideal platforms to demonstrate what aspects of a protein govern heme redox potentials, a key parameter underlying the diversity of hemoprotein functions.

The findings from my work provide the first views of heme and maquette assembly: spontaneous, rapid and with high affinity association. They also provide a foundation for understanding what controls redox potentials of the heme and perspective on this control. The work offers insight into similar processes in natural oxidoreductases, but the concepts and principles uncovered in this thesis will be vital in the development of novel functions applied in man-made applications in vitro and in vivo.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Biochemistry & Molecular Biophysics

First Advisor

Peter L. Dutton

Keywords

Assembly, Heme, Maquettes, Midpoint, Porphyrins, Proteins

Subject Categories

Biochemistry | Biophysics

THE PHYSICAL CHEMISTRY UNDERLYING THE
ASSEMBLY AND MIDPOINT POTENTIAL CONTROL IN
A SERIES OF DESIGNED PROTEIN-MAQUETTES

Lee A. Solomon

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

P. Leslie Dutton Ph.D., FRS

Eldridge Reeves Johnson Professor of Biochemistry and Biophysics and Director of
the Johnson Foundation for Molecular Biophysics

Graduate Group Chairperson

Kathryn M. Ferguson Ph.D.

Associate Professor of Physiology

Dissertation Committee:

Sergei Vinogradov Ph.D., Associate Professor of Biochemistry and Biophysics

Ian Blair Ph.D., A.N. Richards Professor of Pharmacology

Ivan Dmochowski Ph.D., Associate Professor of Chemistry

Greg Van Duyne Ph.D., Jacob Gershon-Cohen Professor of Biochemistry and Biophysics

Kim Sharp Ph.D., Associate Professor of Biochemistry and Biophysics

Jose Cerda Ph.D., Assistant Professor of Chemistry and Chemical Biology at St. Joseph's
University

THE PHYSICAL CHEMISTRY UNDERLYING THE
ASSEMBLY AND MIDPOINT POTENTIAL CONTROL IN
A SERIES OF DESIGNED PROTEIN-MAQUETTES

© 2013

Lee Andrew Solomon

Dedication

To everyone who supported me, both friends and family. This could not have been done without you

Acknowledgements

I have a lot of people to thank due to their significant assistance to this work, scientifically, mentally and emotionally.

First off is my professor, P. Leslie Dutton. The insight he has given me over the years has been crucial to my development as a scientist and person. From the first conversation we had to the latest I seem to always learn something new. Thanks to his encouragement I have grown and thanks to what I have learned from him I will continue to do so. In my future career, wherever it may take me, I will remember what I have learned here, and hope that I can inspire future scientists the way Les has inspired me.

There is no Mathematica program powerful enough to calculate how much Christopher C. Moser has helped me. From his technical help with figures and with equipment to his broader help with crafting ideas and everything else, his input into my development cannot be understated. He has shown me the fun and benefits of a Friday afternoon experiment, which will surely be something I carry on, no matter where I end up. Bohdana Discher has been a constant source of help and good cheer in the lab. Even when her office moved, her presence in the lab could be felt, and I was happy to know she was there to help and provide feedback to what I was doing, or to let me know I was not alone in being frustrated with equipment malfunctions.

It's hard for anyone to start in a lab and that stayed true for me. I was fortunate however to have Ross Anderson around to help me in those early days. Though I started working under him as a rotation student, he became a valued friend. To this day his friendship and advice are still present in my life and I could not be more grateful. After he left I was equally as fortunate to meet Goutham Kodali. Goutham has been a fantastic

influence, both with how he thinks and his approach to science. I hope that well after we're gone our collaborations can continue. Joe Mitala was helpful developing the WAGWA peptide I used, and providing other support. Mike Englander, Geetha Goparaju and Doug Marshall also provided much-needed assistance and advice in the different stages of my tenure here in the lab.

I have been incredibly fortunate to have some wonderful graduate students to have interacted with over the years. Sarah Chobot was always a staunch ally. From serving as my CAPSA deputy to editing this very thesis, she has helped me immensely. I cannot thank her enough. Tammer Farid was my first bay-mate and a good friend, accompanying me around the world, and always willing to hang out and give me some good advice on running and cars. Bruce Lichtenstein helped me before I even came to the lab as my TA in the first days in Philadelphia. His breadth of knowledge never ceases to impress me, and, though he was tough, I think his questioning made me think harder about my ideas, making them better.

Molly Sheehan joined the lab after me and we quickly became good friends. She was a great help both developing ideas and being there to talk to when I needed a break. Bryan Fry was, from the start, a fun and smart influence on the lab. Happy to talk about science or star trek, his time was always appreciated. Chris Bialas, Nathan Ennist, Zhenyu Zhao, Josh Mancini, and Jessica Siedlecki all joined the lab after me, and I know they will go far in life, and I wish them all the luck in the world.

I would also like to thank my undergraduate advisor Dr. Gerald Koudelka. He set me on my current path and made me realize science can be more than classes, cultivating my passion early on.

My thesis committee: Sergei Vinogradov; Ian Blair; Greg Van Duyne; Robin Hochstrasser, Ivan Dmochowski, Kim Sharp, and Jose Cerda, were also helpful providing valued feedback and direction over the course of my studies here.

However none of this would have occurred if I did not have a strong support system. From my earliest days my Mother, Father and Brother has supported my interests, never wanting anything more than to see me happy. I will never be able to thank them enough. My girlfriend Michelle always supported me at the best and worst of times. I am the luckiest man in the world to be with her, and I'm thankful she will be with me in the future. Finally, my friends, I could not be more fortunate to have such great people in my life. Their contributions to this, though mainly by helping me stay sane, are incalculable. I am immensely grateful for them all.

Abstract

THE PHYSICAL CHEMISTRY UNDERLYING THE ASSEMBLY AND MIDPOINT POTENTIAL CONTROL IN A SERIES OF DESIGNED PROTEIN-MAQUETTES

Lee A. Solomon

P. Leslie Dutton

In nature, oxidoreductase proteins are responsible for many enzymatic processes critical to life. These proteins often rely on the presence of non-proteinaceous cofactors to take part in the enzymatic function. The most common, central to my thesis, is heme B. Depending on the protein environment, this cofactor can take part in functions as diverse as electron transfer (cytochromes), oxygen transport (hemoglobins), oxygen reduction (oxidases), carbon-hydroxylation (oxygenases), and superoxide production (NADH oxidase).

In natural oxidoreductases, determination of the course and rates of heme-protein association, what barriers are encountered, what affinity is achieved, and what are the oxidation-reduction potentials, is critical for understanding the rules of assembly and function of the different activities performed. In the growing field of research attempting to make man-made oxidoreductases, the same understanding is required for progress to be made toward construction of novel enzymes. However, this understanding is still out of reach in natural oxidoreductases because of the immense complexity of natural proteins, while for man-made designs progress has only recently reached a point where an in-depth systematic study can be contemplated.

My thesis states: Simple non-natural proteins (maquettes) designed from first principles to ligate heme, can be used to uncover the factors derived from the oligomeric and structural state of related maquette and also derived from porphyrin variants of heme B, that govern rates of incorporation and ligation of heme B into a maquette. Maquettes are ideal platforms to demonstrate what aspects of a protein govern heme redox potentials, a key parameter underlying the diversity of hemoprotein functions.

The findings from my work provide the first views of heme and maquette assembly: spontaneous, rapid and with high affinity association. They also provide a foundation for understanding what controls redox potentials of the heme and perspective on this control. The work offers insight into similar processes in natural oxidoreductases, but the concepts and principles uncovered in this thesis will be vital in the development of novel functions applied in man-made applications *in vitro* and *in vivo*.

Table of Contents

Dedication.....	iii
Acknowledgements.....	iv
Abstract.....	vii
List of Topics.....	ix
List of Tables.....	xiii
List of Figures.....	xiv
Chapter 1: Introduction.....	1
1.1: The Heme Cofactor.....	1
1.1.1: Heme In Nature.....	1
1.1.2: Alternative Heme Variants.....	2
1.1.3: Alternative Tetrapyrrole Cofactors In Nature.....	4
1.1.4: Heme Regulation In The Cell.....	5
1.1.5: Heme-Protein Assembly.....	7
1.1.6: Oxidation Reduction Midpoint Potential.....	10
1.2: Practical Approaches To The Study Of Protein Assembly And E_m Control.....	13
1.2.1: Evolutionary Complexity Hindering Study Of Assembly And E_m Control.....	13
1.2.2: Protein Design As A Tool For Understanding Natural Proteins.....	15
1.2.3: Measuring the E_m Value.....	18
1.3: Goals Of This Thesis.....	21
1.6: References.....	22
Chapter 2: Maquette Development And History.....	26
2.1: What Is A Maquette?.....	26
2.1.1: History Of Maquettes.....	27
2.2: The Current Generation Of Maquettes And Their Attributes.....	29
2.3: Potential of Maquettes Past The Scope Of This Work.....	31
2.4: References.....	32
Chapter 3: The Protein Influence On Holoprotein Assembly Kinetics.....	34

3.1: Introduction.....	34
3.2: Kinetics Of Holoprotein Assembly.....	34
3.2.1: <i>Proteins Used For This Study.....</i>	34
3.2.2: <i>Kinetic Characterization Of The Proteins Used.....</i>	37
3.2.3: <i>Temperature Dependent Kinetic Analysis.....</i>	42
3.2.4: <i>Analysis Of The Assembly Kinetics.....</i>	45
3.3: Proof Of A Partitioned Heme State.....	47
3.4: Conclusions.....	51
3.5: References.....	53
Chapter 4: The Cofactor Influence On Holoprotein.....	54
Assembly Kinetics	
4.1: Introduction.....	54
4.2: Results.....	54
4.2.1: <i>Fe-Porphyrin Assembly And The Effect Of Substituents.....</i>	54
4.2.2: <i>Effect Of The Metal On Assembly.....</i>	63
4.2.3: <i>Effect Of The Protein On Porphyrin Aggregates.....</i>	64
4.2.4: <i>Effect Of Porphyrin Skeleton On Assembly.....</i>	64
4.2.5: <i>Test Of Solubility With A Synthetic Porphyrins.....</i>	65
4.3: Discussion.....	67
4.3.1: <i>Porphyrin Control Over The Rate Of Assembly.....</i>	67
4.3.2: <i>Natural Relevance.....</i>	68
4.3.3: <i>Conclusions.....</i>	69
4.4: References.....	69
Chapter 5: Midpoint Potential Control In A Series Of Maquettes.....	71
5.1: Introduction.....	71
5.2: Results.....	73
5.2.1: <i>Environmental Changes.....</i>	73
5.2.2: <i>Ligation Changes.....</i>	78
5.2.3: <i>Using Cofactors To Shift The Potential.....</i>	79
5.2.4: <i>Using Alternative Porphyrins.....</i>	82
5.3: Discussion.....	85
5.3.1: <i>Midpoint Potential Changes Compared To Previous Maquette Work</i>	85

5.3.2: Magnitude Of The Potential Changes In A New Generation Of Maquettes..	87
5.3.3: Conclusions.....	87
5.4: References.....	88
Chapter 6: Functions Possible In Maquettes When The Assembly Is Understood	90
6.1: Introduction.....	90
6.2: Results.....	90
6.2.1: Electron Transfer.....	90
6.2.2: Porphyrin Sequestration.....	94
6.2.3: Oxygen Binding.....	95
6.2.4: Light Harvesting	96
6.3: Discussion.....	97
6.4: References.....	99
Chapter 7: Conclusions.....	101
7.1: References.....	105
Appendix 1: Materials and Methods.....	106
A.1: Materials.....	106
A.2: Porphyrins.....	106
A.3 Proteins.....	107
A.4: Partition coefficients.....	108
A.5: Rates.....	109
A.6: Spectra.....	110
A6.1: Fluorescence Spectra.....	110
A.7: K_d Titrations.....	111
A.8: E_m titrations.....	111
A.8.1: Spectral-Electrochemistry Titrations.....	113
A.9: Greater Than Millisecond rate values.....	113
A.10: Intra-Maquette Electron Transfer.....	114
A.11: Low Temperature Electron Transfer.....	114
A.12: Oxygen binding.....	115
A.13: Oxygen Binding And Decay Rates.....	115
A.14: References.....	116
Appendix 2: Sequence Appendix.....	117
A2.2: References.....	121

Appendix 3: Spectral Appendix.....	123
A3.1: Heme.....	123
A3.2: Alternate Fe-porphyrin Kinetics.....	124
A3.3: Alternate Fe-porphyrin Redox Titrations.....	133
A3.4: Zn-Centered Chlorins And Porphyrins.....	137
A3.5: References.....	137

List of Tables

Chapter 3

Table 3.1: Descriptions of Proteins In Chapter 3.....	36
Table 3.2: Rate Constants and Pseudo Thermodynamic Parameters of Assembly.....	44

Chapter 4

Table 4.1: Characterization of Fe-Porphyrin Binding to the Single Chain Maquette.....	57
---	----

List of Figures

Chapter 1

Figure 1.1: The Heme Cofactor.....	2
Figure 1.2: Mitochondrial Respiratory Chain.....	3
Figure 1.3: ISD Heme Transport System.....	6
Figure 1.4: Cytochrome b_6 Assembly.....	8
Figure 1.5: Reduction Oxidation Potentials of Natural Proteins.....	10
Figure 1.6: Evolutionary Complexity.....	14
Figure 1.7: Spectrophotometric Cuvette for Potentiometric Titrations.....	21

Chapter 2

Figure 2.1: Binary Patterning.....	26
Figure 2.2: Thermal Denaturation of Maquette Proteins.....	30
Figure 2.3: NMR Characterization of Single Chain Maquette Protein.....	31
Figure 2.4: Heme Affinity of the Single Chain Protein.....	32

Chapter 3

Figure 3.1: Affinity Titrations of Heme B for the Homodimer and Structured Homotetramer Proteins.....	37
Figure 3.2: Rates versus Concentration Plots for Heme B and the Six Proteins Used to Describe Protein Assembly.....	38
Figure 3.3: Burst Phase of the Homodimer Protein	39
Figure 3.4: Heme Binding Site Knockouts.....	41
Figure 3.5: Four Second Traces of Heme Assembly.....	42
Figure 3.6: Eyring Plot Describing Heme Assembly.....	43
Figure 3.7: Demonstration of Heme Partitioning.....	48
Figure 3.8: Zn-Protoporphyrin IX Binding to His-less Single Chain.....	49
Figure 3.9: Fluorescence Titration of Heme Binding.....	50

Scheme 3.1: Reaction Scheme of Heme-Protein Assembly.....	35
---	----

Chapter 4

Figure 4.1: Selection of Fe-porphyrins.....	55
Figure 4.2: Assembly as a Function of Time for the Fe-Porphyrins and the Single Chain Protein.....	56
Figure 4.3: Log-P Values Compared to Assembly Rates.....	58
Figure 4.4: Description of the Amphiphilic Hypothesis.....	59
Figure 4.5: Affinity Titrations of a selection of Fe-Porphyrins.....	60
Figure 4.6: Ferric versus Ferrous Heme Assembly.....	61
Figure 4.7: Comparison of Fe and Zn Centers of Protoporphyrin IX.....	62
Figure 4.8: Zn-Protoporphyrin IX Aggregate Being Broken Up By the Single Chain Proteins.....	63
Figure 4.9: Comparison of Chlorin and Porphyrin Assembly.....	64
Figure 4.10: Comparison of Soluble and Insoluble Porphyrin Assembly.....	66

Chapter 5

Figure 5.1: Charge Patterning of Helices.....	72
Figure 5.2: Thermal Denaturation of Different Charge Patterned 4- α -Helix Bundle Proteins.....	73
Figure 5.3: Affinity of Charge Patterned Proteins for Heme B.....	74
Figure 5.4: Potentiometric Titrations of Charge Patterned Proteins.....	75
Figure 5.5: Potentiometric Titration of Heme; Different Domains In a Single Protein.....	76
Figure 5.6: Potentiometric Titrations of Heme B; Attempts at Charge Patterning.....	77
Figure 5.7: Potentiometric Titrations of Different Ligation Combinations to Heme B.....	79
Figure 5.8: Spectra of 5-Coordinate Heme B.....	80
Figure 5.9: Pymol Models of Heme Placement in Maquettes.....	81
Figure 5.10: Offset Heme Characterization.....	82
Figure 5.11: Midpoint Values of Various Alternative Porphyrins in the Single Chain Protein.....	84
Figure 5.12: Summary of Maquette Midpoint Values and Their Comparison to Nature.....	86

Chapter 6

Figure 6.1: Electron Transfer Between Two Maquettes.....	91
Figure 6.2: Low Temperature Electron Transfer Between Two Maquettes.....	93
Figure 6.3: Oxygen Binding to the Single Chain Maquette.....	94
Figure 6.4: Oxyferrous Decay With Different Ligations to Heme.....	95
Figure 6.5: Light Harvesting In a Maquette.....	97

Appendix 3: Spectral Appendix

Figure A3.1: Heme B Spectra In Various States and Solvents.....	123
Figure A3.2: Heme In Cytochrome b_{562}	123
Figure A3.3: Heme B Kinetic Spectra.....	124
Figure A3.4: Mesoporphyrin Kinetic Spectra.....	125
Figure A3.5: Deuteroporphyrin Kinetic Spectra.....	126
Figure A3.6: IsoHematoporphyrin Kinetic Spectra.....	127
Figure A3.7: Diacetyl Deuteroporphyrin Kinetic Spectra.....	128
Figure A3.8: Tetracarboxyphenylporphyrin Kinetic Spectra.....	129
Figure A3.9: 2,6-Dinitrileporphyrin Kinetic Spectra.....	130
Figure A3.10: Etioporphyrin Kinetic Spectra.....	131
Figure A3.11: Protoporphyrin IX Dimethyl Ester Kinetic Spectra.....	132
Figure A3.12: Heme A Kinetic Spectra.....	133
Figure A3.13: Mesoporphyrin Redox Spectra and Nerst Curve.....	134
Figure A3.14: Deuteroporphyrin Redox Spectra and Nerst Curve.....	134
Figure A3.15: IsoHematoporphyrin Redox Spectra and Nerst Curve.....	135
Figure A3.16: Diacetyl Deuteroporphyrin Redox Spectra and Nerst Curve.....	135
Figure A3.17: Tetracarboxyphenylporphyrin Redox Spectra and Nerst Curve.....	136
Figure A3.18: Heme A Redox Spectra and Nerst Curve.....	136
Figure A3.19: Zn-Centered Tetrapyrrole Kinetic Spectra.....	137

Chapter 1: Introduction

1.1: The Heme Cofactor

1.1.1: Heme In Nature

The large family of oxidoreductase proteins often employs a wide variety of cofactors capable of taking part in enzymatic reactions. Organic, inorganic, and in some cases amino acids themselves are used to setup electron transfer pathways to catalyze these functions. Making up a large number of oxidoreductase proteins are ones that contain a specific cofactor named heme B (Fe-protoporphyrin IX) (Figure 1.1-B).¹ This cofactor is an iron-centered tetrapyrrole used for a number of functions including oxygenases (heme oxygenase),² cytochrome P450s,³ electron transfer (cytochrome c),⁴ small molecule binding (myoglobin, soluble guanylyl cyclase),⁵ peroxidases (Cytochrome c peroxidase),⁶ and quinone oxidation and reduction (redox) reactions (cytochrome *bc*₁).⁷ As such, it is critical we have a deep understanding of these proteins as they are important to cellular viability, and have effects that extend into medical disorders such as ophthalmoplegia, and cardiomyopathy.⁸

Of the many reactions hemoproteins carry out, the respiratory electron transport chain reactions are amongst the most prevalent throughout nature (Figure 1.2).^{9,10} These reactions couple electron transfer to proton transport across a membrane generating a large electrochemical potential ($\Delta\mu_{H^+}$), which is then used to power ATP synthesis. Heme plays a vital role in this process being present in, and integral to the function of, respiratory pathway proteins Complex II, III, and IV (Figure 1.2).¹¹ In these proteins it is used to shuttle single electrons reduce or oxidize quinone cofactors and reduce O₂ to water.¹⁰ Heme is

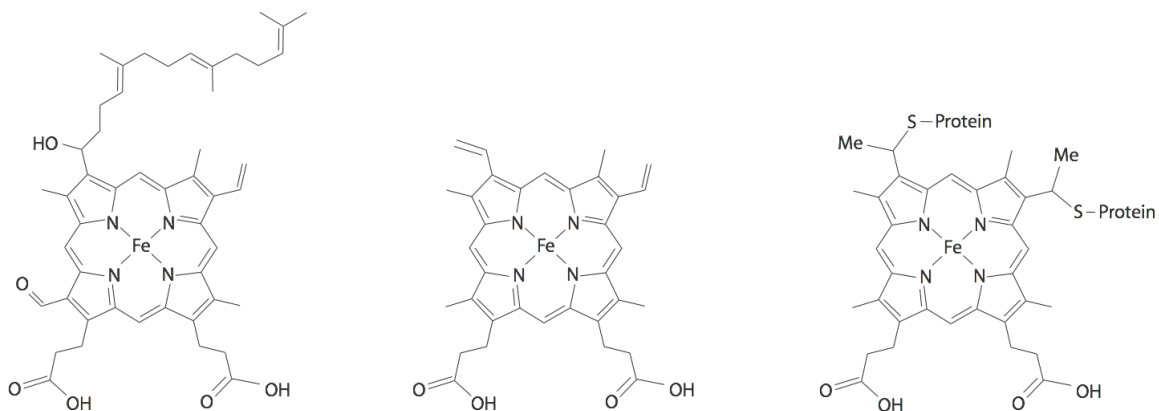


Figure 1.1: From Left to right: Heme A, Heme B and Heme C. These are the three most common naturally occurring heme cofactors, used in a wide variety of functions

also present in photosynthetic proteins, most notably cytochrome *b₆f* and Photosystem II.^{9,12,13} Again, in these proteins heme is central to electron and proton transfer reactions generating a proton gradient for use in ATP synthesis.¹⁴

Due to the prevalence of the heme cofactor, and its importance to cellular viability, the interactions it has with proteins need to be understood. My thesis describes the physical chemistry underlying assembly of this cofactor with protein and the effect of the protein environment on the redox chemistry it takes part in.

1.1.2: Alternative Heme Variants

Though heme B is the most widely used porphyrin it is not the only one utilized in nature. There are a wide variety of tetrapyrroles that display diverse functions. Heme B itself is often modified so that its functional range can be expanded. Heme A is an example of how heme B is modified in nature to change its function. This cofactor is present in cytochrome oxidase in mammals.^{15,16} This variant has the same tetrapyrrole ring as heme B however its substituents are modified (Figure 1.1-Left). Both heme A and heme O have, at

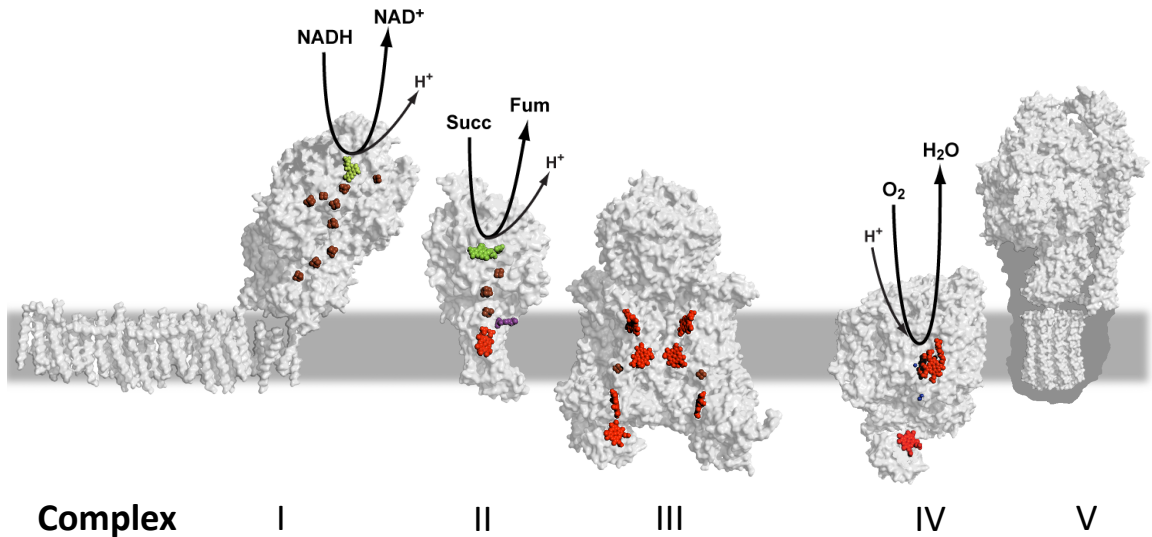


Figure 1.2: Mitochondrial respiratory electron transport chain. A proton motive force ($\Delta\mu\text{H}^+$) is generated by proton pumping, the energy for which is obtained through electron transfer. Heme cofactors are represented in red and present throughout a majority of these proteins.

the C3 position, a farnesyl tail, a long hydrophobic chain that is used to anchor the heme into a large, membrane bound protein.¹⁷ This is in place of a vinyl group present in heme B. Heme A also has, at the C18 position, a formyl group in place of a methyl group. This electron withdrawing modification substantially raises the redox midpoint potential to make it a potent oxidant. While it is functionally beneficial for these modifications to be made, these modifications have inherent complications. Proteins that utilize heme A require a large hydrophobic cavity specifically for the farnesyl tail. In nature, as a result of these issues, this cofactor is constructed and inserted into cytochrome oxidase through a complicated system of protein chaperones. Heme A is controlled by the COX system. It is synthesized by COX10 and COX15, and then chaperoned into the appropriate site in the protein.¹⁷ Due to its higher oxidation potential, this porphyrin is used for oxygen reduction seen in cytochrome oxidase and requires extensive modifications far different and more complex than a standard heme B protein.

Heme C is another heme B variant. This variant uses the C3 and C8 vinyl groups as points of covalent attachment to a protein that contains a CXXCH motif (Figure 1.1-Right). This simple system allows for many different ligands and environments without the worry of losing the heme to solution. This variant requires chemical reactions for its assembly, and multiple proteins in the cytochrome-c maturation (CCM) system have evolved to chaperone c-heme cofactors into place, and aid in the covalent linkage to the protein.⁴ With this attachment motif proteins retain heme despite not being optimized for a high affinity. The attachment motif greatly expands the range and abilities of many hemoproteins, though the complex system of proteins interfere with the study of how proteins affect the assembly of hemoproteins.

1.1.3: Alternative Tetrapyrrole Cofactors In Nature

In addition to heme B and its variants there are also many types of porphyrins that have both alternate metals and structures. A common tetrapyrrole alternative to heme is the chlorin, the basis of chlorophyll pigments central to photosynthesis.¹⁸ Chlorophyll is well-known for its extensive modifications to its ring structure and a Mg metal in the center.¹⁹ The metal in the center allows for this pigment to effectively absorb light and transfer this energy through a series of cofactors and proteins eventually leading to its conversion into chemical energy that can be stored by plants.¹⁸

Using the same starting material as heme B, these pigments must all be synthesized and assembled in the cell on an as needed basis without excess released. Like heme A chlorophyll has a large phytyl tail maintaining them in hydrophobic environments. The various assembly processes have been well documented from the initial signaling to final

cofactor insertion. The fact that heme B precursors are the starting material only increases the importance of understanding porphyrin interactions with natural proteins, as they lead to further understanding of the interactions made between proteins and other modified variants of this cofactor.

Photosynthetic and respiratory systems are vital throughout all kingdoms of life, and all have the same heme cofactor.¹⁰ However, despite this importance, tetrapyrrole cofactors are toxic if it is not in an appropriate protein environment.¹¹ Because of this problem, their concentration is highly regulated (both acquisition and removal), which will be described below.

1.1.4: Heme Regulation In The Cell

Many organisms employ systems of acquisition and transport in order to get heme B into the cell. The ISD (iron-responsive surface determinant) proteins in *Staphylococcus aureus* are one such example (Figure 1.3).²⁰ These proteins are membrane transporters that bind hemoproteins in the exterior environment and transport them across the membrane. Once heme is in the cell many other proteins of this system dismantle the holoprotein and shuttle the heme to either an apoprotein or to heme oxygenase for degradation. Another example is the hemopexin system, seen in many organisms such as *Haemophilus influenzae*.^{21,22} As in the ISD systems, this protein is secreted by the cell and binds free heme in the extracellular environment with one of the highest known affinities, having K_d values estimated in the sub-picomolar range. When it binds heme the complex then binds to a cellular surface protein called HxuC (Hemopexin Utilization), which mediates its entry into the cell, followed by interaction with the HxuA and HxuB proteins that mediate heme release from hemopexin for uptake by cellular machinery.

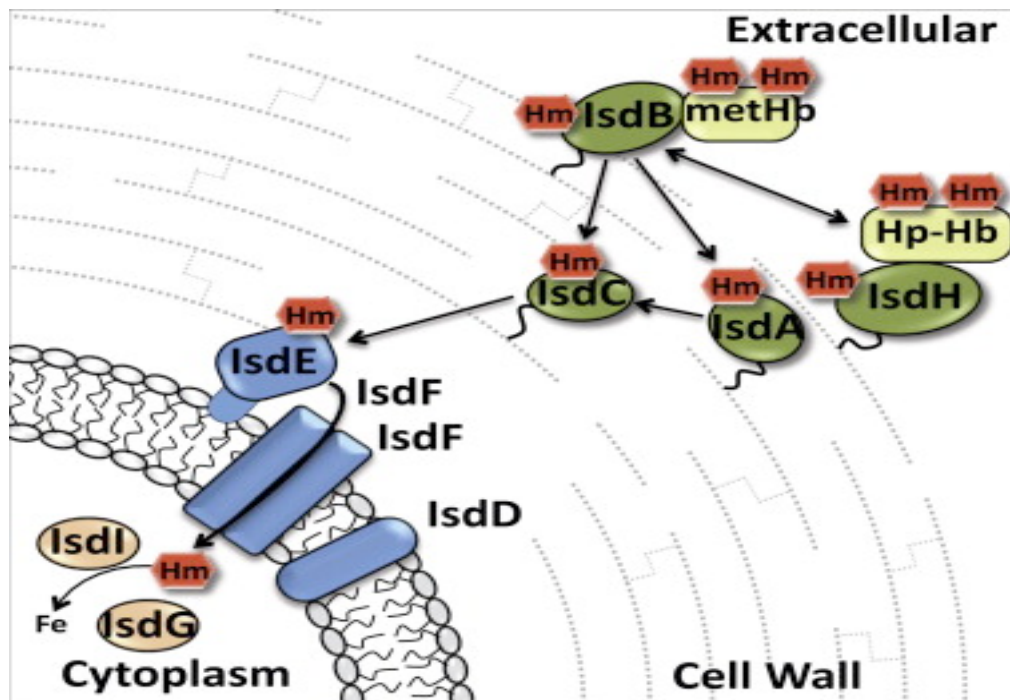


Fig. 1.3: Schematic representation of the Isd system heme transport components. Heme transport and iron liberation is accomplished by the coordinated effort of nine Isd proteins. IsdA, IsdB, IsdC, and IsdH (green) are covalently anchored to the cell wall. IsdE and IsdF (blue) are the binding protein and permease components of an ABC transporter, respectively. IsdE is shown in the heme-bound state prior to complexation with IsdF for transport. IsdD (blue) is a membrane protein of unknown function. IsdG and IsdI are cytoplasmic heme-degrading enzymes. The ligand preferences for each member are illustrated as Hm (heme), metHb (methemoglobin) and Hp (Haptoglobin-hemoglobin). For simplicity, IsdB and IsdH are shown interacting with only one protein ligand, but in fact, IsdB also binds Hp-Hb and IsdH also binds metHb. The predominant heme transfer path in the Isd system is represented by arrows. Figure and legend reproduced from: Jason C. Grigg, Georgia Ukpabi, Catherine F.M. Gaudin, and Michael E.P. Murphy in the paper "Structural biology of heme binding in the *Staphylococcus aureus* ISD system" *Journal of Inorganic Biochemistry*, Volume 104, Issue 3, 2010. 341-348.¹²

Once in the cell the concentration of heme must be carefully maintained, as free heme is cytotoxic. It has been shown to cause a multitude of problems ranging from aggregation to the production of reactive oxygen species (ROS).^{13,23-25} The acquisition of

heme must be kept in check by the degradation or sequestration of heme. Degradation is predominantly carried out by the Heme-oxygenase protein.² This protein consumes a molecule of NADH to convert heme into biliverdin, which is eventually converted to glucouronate and can be easily exported as waste. Cellular proteins such as ferritin can bind up the free Fe left over from this process. Though this method is the most common, certain organisms have been shown to circumvent it.²⁶

1.1.5: Heme-Protein Assembly

After heme has been acquired and transported into the cell, holoproteins must assemble and attain a functional state. Little is known about spontaneous *in vivo* protein assembly. It is difficult to monitor these processes due to the many factors brought on by the complicated nature of a living cell. It is not possible, with the current level of technology, to monitor a single protein such as myoglobin from the initial translation steps through heme assembly. Furthermore, genetic mutations required to either halt the processes or attach tags to monitor the protein throughout various stages may interfere with the function, inhibiting the protein's normal activities and interactions.

Assembly can be monitored in complex systems, however. If a protein requires multiple chaperones to aid in folding and heme association, knocking out the chaperones pauses the process at a specific step. The most relevant work examining heme proteins is not with soluble hemoproteins, but membrane bound ones. Cytochrome *b₆f* was examined *in vivo* with the aim of determining its assembly process. Kuras, Wollman et. al. were able to elucidate a genetic pathway that regulated the binding of heme B to this complex (Figure 1.4).²⁷ In the *C. reinhardtii* organism, there are nuclear genes called CCB genes

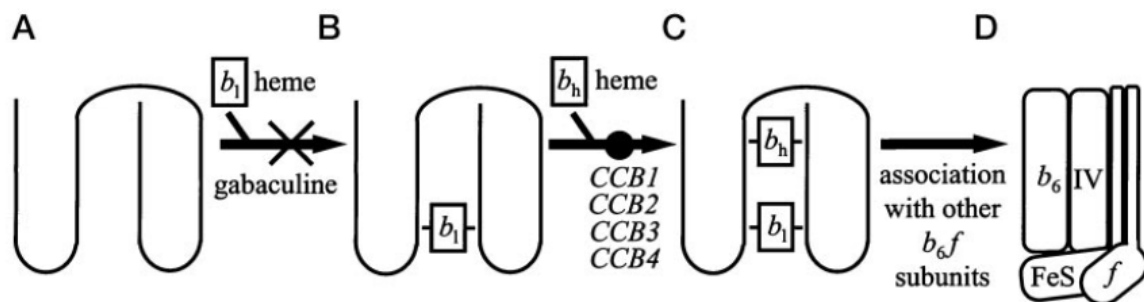


Figure 1.4: Schematic pathway of the conversion of apocytochrome to holo-cytochrome b_6 and patterns after urea/SDS-PAGE. A, membrane integration occurs even in absence of heme association. B, there is formation of a b_1 heme-dependent intermediate, which can be prevented by heme depletion (gabaculine treatment). C, Ccb1–Ccb4 encoded nuclear factors are necessary for the production of the holo-form showing both b_h heme binding and b_1 binding. D, holo-cytochrome b_6 accumulates in a protease-resistant form, upon association with the other b_6f subunits. Figure and legend reproduced from: Richard Kuras, Catherine de Vitry, Yves Choquet, Jacqueline Girard-Bascou, Duane Culler, Sylvie Buschlen, Sabeeha Merchant, and Francis-Andre Wollman, “Molecular Genetic Identification of a Pathway for Heme Binding to Cytochrome b_6 ” *Journal of Biological Chemistry*, Vol 272, No. 51, 1997, pp. 32427-32435.²⁰

(construction cytochrome B) that are known to control *c*-type assembly, however they appear to be implicated in the *b*-type assembly with the b_6f complex. When these proteins are rendered non-functional, the cells appear to have the same phenotype as the b_6f complex with its heme-binding site removed. Similarly, in yeast, the bc_1 complex requires the Cbp protein system to assemble.²¹

It is unclear what the mechanism of this process is in both cases. The work done has only been at the genetic level as more direct biochemical observation is not possible. Based on structural and in vitro data, the proteins are thought to assemble with heme without any assistance. Proteins are needed to deliver heme but not directly insert it. Currently it is thought that the heme is sequestered until the protein is able to accept it, though it is unknown where.¹¹

Prokaryotic systems have likewise been examined though not much more is known. One clear difference here is that heme assembly is thought to be a spontaneous process requiring no exterior proteins. Palombo and Daley show have worked out the process of the, cytochrome *bo*₃ protein of *E. coli* (complex IV of the electron transport chain) providing a mechanism for how this protein comes together.²⁸ Subunit I must associate with II and III before heme can bind. *In vitro* experiments show that in the apo form subunit I binds heme with low affinity, and when these three subunits are together the affinity increases significantly. Heme-binding and additional conformational changes allows subunit II to associate making a fully functional protein. It is still unknown, however, where the heme originates and the mechanism by which it is delivered to the proteins.

In vitro experiments have been done examining the affinity of heme to natural proteins. The tetrapyrrole ring of heme is itself very hydrophobic and providing a watertight pocket minimizes interactions with the aqueous solvent.²⁹⁻³¹ In addition to providing a hydrophobic pocket, the protein provides coordination bond(s) to the central metal. Typically this is a histidine residue, although cysteine, and methionine are known to ligate as well.^{32,33} These two factors are prevalent throughout many types of folds in nature. Globin and hemopexin folds are amongst the best known and both share a large amount of hydrophobicity in their binding pockets. Interestingly many hydrophobic residues serve an additional purpose. Aromatic residues such as tryptophan and phenylalanine, in addition to generating a water-tight pocket, also provide pi-stacking interactions increasing the affinity through this added stabilization.

The structure of the heme-binding cavity is critical to the protein not just for function but in many cases folding as well. Myoglobin relies on heme for its folding process leading to a final rigid structure. The importance of heme ligation is evidenced by the fact that the PDB has no published X-ray or NMR structures of myoglobin without heme or a heme-analogue present.^{34,35} Cytochrome *b*₅₆₂ and *b*₅ for example, have unfolded sections that gain a large amount of rigidity upon heme binding.^{36,37}

1.1.6: Oxidation Reduction Midpoint Potential

Natural Heme Proteins

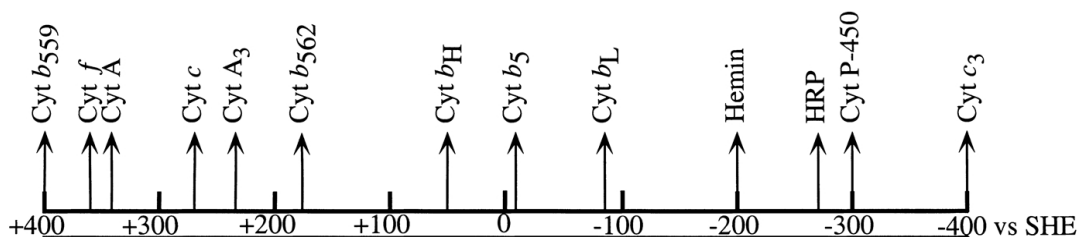


Figure 1.5: The reduction midpoint potential of free hemin, FePPIXCl, in aqueous solution is shown for comparison. Reduction midpoint potential values correspond to pH 7-8 range. References for redox activity values, Figure, and Legend adapted from: Julia M. Shifman, Brian R. Gibney, R. Eryl Sharp, and P. Leslie Dutton “Heme Redox Control in de Novo Designed Four- α -Helix Bundle Proteins” *Biochemistry*, Volume 39, 2000, 14813-21.⁴¹

The many different oxidoreductase functions share a common aspect: Electron transfer. Monooxygenase proteins reduce oxygen in a controlled manner and cytochrome *c* peroxidase oxidizes cytochrome *c*. Depending on environmental factors the heme cofactor can either accept or donate electrons with a substrate.³⁸ This is described by a thermodynamic parameter called the oxidation-reduction midpoint potential (E_m). More specifically, this value is the potential of the environment where the oxidized and reduced

components of the half-cell reaction are equal. When two redox-active molecules are mixed together, electrons are spontaneously exchanged from the molecule with the lower E_m to the one with the higher value. This is often exploited in nature; changing the environment around individual cofactors creates a series of increasing E_m values and a pathway for electrons to flow, like those seen in respiration and photosynthesis from above.
7,39,40

Nature has many ways of shifting the midpoint potential of not only heme B, but the wide variety of cofactors seen in nature stabilizing either the oxidized or reduced state (Figure 1.5). The most common heme variant, heme B, has an inherent midpoint potential of -200mV (Hemin in Figure 1.5) in water but in natural oxidoreductases this is expanded to a very wide range, as low as -400mV in Cyt c_3 and up to +400mV in cytochrome b_{559} .⁴¹ One way is to attach different substituents to alter the electron density of the ring. Various changes to its tetrapyrrole ring results in different heme types such as heme A, seen in cytochrome c oxidase (E_m values of +200-350mV),⁴² and Heme- d_1 in Nitrite reductase (E_m of +234mV).⁴³ In natural systems, any modification to the tetrapyrrole needs to be done in parallel to co-evolving the binding sites in order to accommodate the modified cofactors. It is this coevolution of the protein that complicates the study of the cofactors. Making a certain change to one amino acid can have unforeseen effects shifting the potential not by direct interactions but rather by weakening the binding site or changing the water content of the local environment.

Having different ligands to the heme Fe also changes its propensity for being oxidized or reduced.^{44,45} Histidine and methionine have been observed ligating heme B and C^{46,47} and both have a significant effect on the E_m . Methionine ligates more strongly to

ferrous iron as opposed to ferric, due to it being softer according to Hard Soft Acid Base theory.⁴⁸ Sulfur increasing the E_m is not always the case, however. Thiolate ligation bonds, with their negative charge stabilize the ferric Fe, thereby suppressing the midpoint reduction potential.

The axial iron ligands can also factor into the midpoint potential by changing the spin state of the Fe at the heme center, a common trend seen in proteins such as cytochrome P450.^{3,49} In the low-spin state Fe (III) is more stable, and the heme exhibits a lower midpoint potential. However, in the high-spin 5-coordinate state the midpoint potential increases. These midpoint shifts are important to the function of the protein. Cytochrome P450 is a well-known oxygenase that generates a potent oxidant capable of generating radicals. In P450 enzymes the redox potential is part of the catalytic mechanism. In its resting state, the heme is in a low spin 6-coordinate state. When substrate binds, a water ligand is removed and the heme converts to a 5-coordinate state with a corresponding increase in the reduction potential. This shift is integral to the mechanism as it allows for a controlled reduction of the heme and the beginning stages of the catalytic cycle.

The environment however, is another common way of modulating this value. The midpoint potential can be highly affected by water exposure as ferric heme is stabilized by hydrogen bonds in water with a E_m value of -200mV (Figure 1.5).⁴¹ Water access is not the only environmental factor affecting the heme. Many amino acid residues are charged, either positively or negatively, at neutral pH. Having these in the vicinity of the heme, or even coordinated to the Fe atom can cause significant shifts in the midpoint potential.⁵⁰ Having a lysine residue near the heme can cause a drastic increase in the midpoint

potential by destabilizing the oxidized Fe (III) state through coulombic interactions. Alternatively, having a glutamate nearby, with a corresponding negative charge, will have the opposite effect of lowering the E_m by stabilizing the Fe (III) state. These types of amino acids are commonly used in natural proteins to cause shifts in the potential.⁵¹⁻⁵³

Natural proteins often take advantage of these redox differences to control function. However, natural proteins often use a combination of these effects to modulate the potential. This complicates the study of how much one aspect of the environment affects the heme when in a protein. A single amino acid may have unseen or masked effects when mutated, or may affect the environment from a substantial distance away from the cofactor. This, along with the initial heme assembly, have evolved over millennia, building in a significant tolerance to external environmental factors and genetic manipulation. These multiple utilities make for both a fascinating yet complex system whose understanding remains thus far incomplete. Simplified systems are needed that can determine how the individual contributions affect the assembly and environmental factors of oxidoreductase functions.

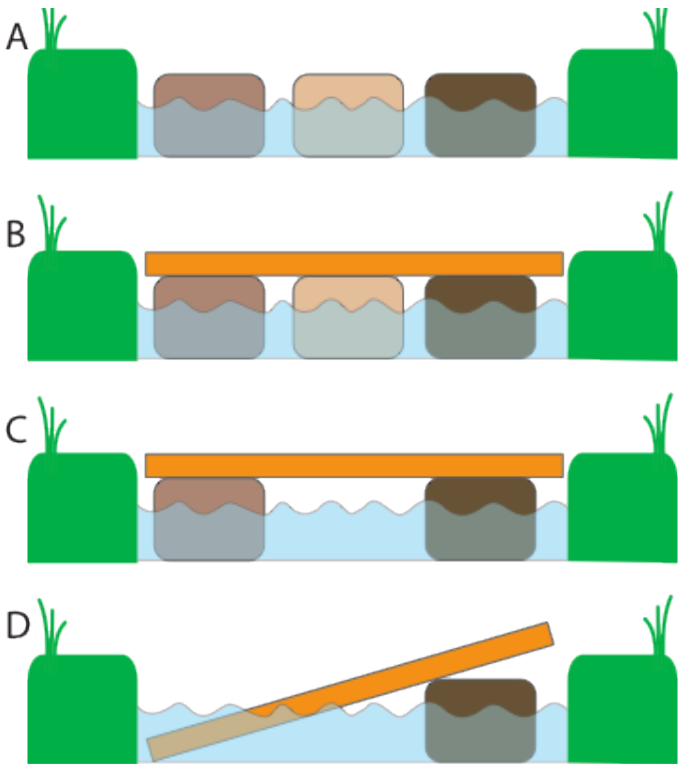
1.2: Practical Approaches To The Study of Protein Assembly And E_m Control

1.2.1: Evolutionary Complexity Hindering Study Of Assembly And E_m Control

Determining the function and understanding the natural role of an oxidoreductase protein is complicated at every step, starting from the earliest genetic stages. The protein is commonly removed from its natural context, and in doing so there may be functions removed from the enzyme.

Most laboratories study the natural proteins directly in order to understand how the systems operate. They may mutate various residues and look for a change in the structure or function. However, by doing this both the structure and function may be adversely affected in unforeseen ways. These changes may be direct and predicted, but can also have effects from sites remote to the active site. These effects could be from structural perturbations or gross misfolding that changes other parameters of the protein (cofactor midpoint potential for example). Unintended consequences arise from the various pressures of evolution, which may have developed many complex and redundant features in natural proteins. These redundant mutations make the protein robust; however, they can inhibit their study by adding unknown layers of complexity. For example, a series of mutations can end up compensating for one another functionally so if an important amino acid is removed others can counterbalance it. Figure 1.6 depicts the relationship of

Figure 1.6: Depiction of the complexity brought on by evolution. When a function is present, such as crossing a bridge (A), a nonfunctional piece can be added (B). If this nonfunctional piece is rendered important (C) it cannot be removed again. This is a prime example of Muller's ratchet. Alternatively, if (B) is naturally found and studied it will be unclear which is providing function, the orange plank or the beige stone. Removal of either still allows for crossing of the river, the intended function of the bridge.



redundancy in a simple cartoon where the function is crossing a river.⁵⁴ At first, it is possible to simply walk across on the stones to get from one side to the other (Fig. 1.6A). A mutation is then made wherein a plank is added which has little or no functional effect (1.6B). At this point, removing either the plank or the middle stone would not impair the function of getting across; if one were studying this crossing removing either of the two individually would not clearly demonstrate how these pieces contribute to getting across.

This figure depicts another problem as well, one of evolved complexity. In 1.6A the function is simple, as it is in 1.6B as well, though slightly improved. However, when a mutation is made as in 1.6C, the middle stone is removed making the plank critical to the function. Removing any piece would now completely wipe out the ability to cross the river. In natural proteins these types of mutations can occur greatly complicating the study of a natural function by adding complexity that obfuscates a clear picture of what is going on. These are two simple examples that describe the inherent problems of directly studying natural proteins. There are many unforeseen changes that can be made originating from evolution over the many millennia these proteins have existed.

1.2.2: Protein Design As A Tool for Understanding Natural Proteins

With protein design, the drawbacks of evolution are avoided as functions are engineered from basic chemical principles using amino acids. Functions can be directly studied and evolutionary complexity can be ignored as it is not part of the design process. However, there is not one single method of protein design, and different laboratories utilize varying approaches.

Despite the complexity of natural proteins, many laboratories utilize a method where minor mutations are made to well studied naturally occurring protein structures with

the aim of changing their function. This method makes changes as small as point mutations in the hope that it alters the activity without affecting other attributes not directly related to function. This is beneficial as it allows for a fully functional binding and/or active site to be formed, with the literature providing guidance on which mutations can be tolerated. The change in function slightly circumvents the evolutionary complexity because the protein was not originally intended to perform this function so compensatory and redundant mutations have less of an effect. In addition, as the function is novel to this structure, many basic characterizations must be performed from scratch. One specific example of this is seen in myoglobin. Matsui, Morishima et. al. have made point mutations to both histidine residues in myoglobin changing its function from an oxygen binding protein to one with peroxidase activity.⁵⁵ Simply changing the electronic structure on the heme through the mutation of the ligating-his residue into a cys residue significantly changes the function. This has also been seen in the work of Yeung, Lu et. al..⁵⁶ They also used myoglobin as a starting point, and by making mutations in the heme pocket they were able to convert the function from an oxygen binding protein to a nitric oxide reductase.

Directed evolution is another approach made use of in the field of protein design.^{57,58} With this method, a gene is randomized by methods such as error-prone PCR or mutator strains of bacteria. Many different versions are made creating a large library of genes, with some libraries getting as large as 10^{15} different mutations. These mutants are put through a high-throughput screen and proteins that have attained the desired target are examined to determine what residues they now have and how they factor into the desired function. This method works backward from the previous methodology, it does not involve a consideration of what steps are required before mutagenesis is done. Instead it keeps

randomly evolving a gene such that it becomes more and more functional, and then the sequence and structure are examined to determine what amino acids are involved, and how the protein operates. This method has been very successful at a variety of structural functional and topological changes, modifying proteins to have new quaternary structures, functions, and midpoint potentials.⁵⁹ This method is primarily limited by the size of the library, which in turn is limited by the assay size. As assays get more sophisticated the potential for this field will grow. This technique has incredible potential, able to start with even the simplest proteins, in some cases using only seven amino acids, and come out with complex functions.

A third common method used is computational design.^{60,61} In this method a protein is modeled with the aid of various computational tools, energetic minimization and Rosetta are common examples. In enzymatic designs, the proteins are constructed based around stabilizing the transition state or binding a specific cofactor.^{62,63} The computer is able to generate energy-minimized structures that would perform this action thereby generating a new design. There are two methods a computer can use to do this. First, the program can scan the library of known structures to find an appropriate scaffold and modify it as mentioned above. This has many of the advantages as site directed mutagenesis, but with the advantage of having a computer minimize the energies involved, thereby optimizing the intended function. The second method of design is to have the computer itself build a structure de novo. This is obviously much more computationally taxing as a variety of factors must be taken into account and energy minimization steps must be done frequently. As a consequence, these two methods are often combined. A

computer will scan the PDB and use various pieces of specific proteins to compile the best structure it can compute.

A fourth design approach makes use of simplified structures constructed from chemical principles. Named the maquette approach, it is much like the site-directed mutagenesis, but with one key difference, it starts with a simple scaffold protein that has no natural counterpart, not a natural protein. Due to its simplicity, every amino acid can be structurally and functionally accounted for. When changes are made to this protein the gain and or loss of a function can be readily assessed.⁶⁴ This approach works as follows: starting off with a simple scaffold protein whose sole function is to fold, and knowing the function desired, minor modifications to the scaffold can be made with the aim of building in the goal. As these mutations are made in discrete steps and the scaffold is well defined, each mutation's effect can be directly characterized. If unintended consequences arise from a change, we can either go back to the previous design or continue on noting how the protein was affected in the current step. There is no mimicry; rather changes are chosen based on the underlying chemistry of the intended function. Cofactor-binding sites from natural proteins are not imported directly; rather they are examined and the chemical bonds made to heme or other cofactors are engineered into the new constructs. For example heme ligates to histidine residues in a hydrophobic core. The myoglobin-binding site sequence was not imported directly or mimicked, rather, to the hydrophobic core already part of the protein, two his residues were incorporated, achieving a sub-micromolar K_d value.⁶⁵

1.2.3: Measuring The E_m Value

The diversity of heme functions is strongly influenced by the midpoint potential. As such, the ability to describe how the protein environment is able to control this parameter is amongst the most important aspects of understanding and reproducing the variety of oxidoreductase functions. In both natural and synthetic systems a clear means of obtaining this value and being able to observe how mutations have affected it is crucial. The following section describes a straightforward method applying chemical concepts to the study of protein electrochemistry.

Below is a description how electron transfer is measured and the underlying concepts. A more in depth discussion of these concepts can be found in the paper by P. Leslie Dutton.⁶⁶ The midpoint potential is a thermodynamic parameter describing electron transfer, as such it can be directly related to Gibbs free energy through the following equation:

$$\Delta G = -zF\Delta E$$

where ΔG is the Gibbs free energy, z is the number of electrons involved in the redox half cell, F is Faraday's constant (96,493 J/V), and ΔE is the difference in reduction potential between the two half-cell components of the reaction. This is easily coupled to the other free energy equation:

$$\Delta G = \Delta G^\circ - RT \ln(K_{eq})$$

where R is the gas constant, T is the temperature in Kelvin and K_{eq} is the equilibrium constant between reactants A and B. In this experiment, the equilibrium is between the oxidized and reduced species. The ratio of [reduced]/[oxidized]. ΔG° is the free energy

under standard unit conditions. Combining these two equations yields the following Nernst equation:

$$E_{\text{cell}} = \Delta E^{\ominus} + \frac{RT}{zF} \ln \frac{[\text{A-Red}][\text{B-Ox}]}{[\text{A-Ox}][\text{B-Red}]}$$

where $\Delta E^{\ominus} = (E^{\ominus}_{\text{A}} - E^{\ominus}_{\text{B}})$ is the difference in standard potential between A and B. The point where the standard reduction potential E_{cell} is equal to the ΔE^{\ominus} is the midpoint potential of the redox pair being measured.

There is no absolute value for a E_{m} value so a scale must be defined. In the literature the standard hydrogen electrode (SHE) scale is commonly used. This scale is based on the equilibrium between H_2 and $2\text{H}^+ + 2\text{e}^-$. This can take the place of component A in the above equation reducing it to the following:

$$E_{\text{h}} = E_{\text{B}}^{\ominus} + \frac{RT}{zF} \ln \frac{[\text{B-Red}]}{[\text{B-Ox}]}$$

Where E_{h} is the potential of the environment measured with the SHE. Given this, at equilibrium between A-Red and A-Ox, E_{h} is equal to E^{\ominus}_{A} which is defined as the midpoint potential for the B-component half cell.

Redox reactions are not always independent of pH and the $[\text{H}^+]$ must be factored into the equation above. In the case where the electron loss is coupled to proton loss



The Nernst equation can be rewritten as such

$$E_{\text{h}} = E_{\text{B}}^{\ominus} + \frac{RT}{zF} \ln \frac{[\text{Red-H}]}{[\text{Ox}][\text{H}^+]}$$

Which can be further expanded to

$$E_h = E_B^\ominus + \frac{RT}{zF} \ln \frac{[\text{Red-H}]}{[\text{Ox}]} - (0.06 \times \text{pH})$$

From this it is clear that the midpoint will change as a function of pH, 60mV for every pH unit under normal conditions. This can be a useful diagnostic tool, uncovering potential structural or electrochemical changes the protein undergoes as a function of redox state or pH. This relationship, because it is based in equilibrium, must be reversible. If re-oxidation

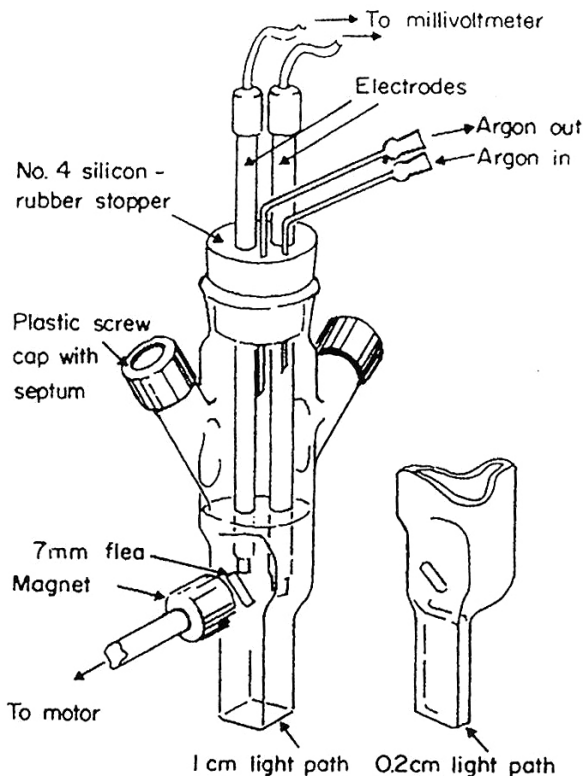


Figure 1.7: Spectrophotometric cuvette for redox potentiometric titrations at ambient temperatures. Figure reproduced from Dutton P.L. *Methods of Enzymology*. Vol. 54, 1978, pp 411.³⁷

of a reduced cofactor does not share the same potential at the same conditions it is indicative of some perturbation as well.³⁷

The method of Dutton et. al. is the most straightforward method of observing this value. By using a specially designed cuvette (Figure 1.7) we are able to monitor both the UV/Vis spectra while at the same time monitoring the potential of the environment. The UV/Vis spectra allow for a direct observation of the concentration of each redox state, which can be coupled to the electrode reading. The use of spectra also make clear any changes to the system not associated with reduction such as free cofactor, or other

side redox reactions.

This method will be used throughout this work, and all numbers are relative to the Standard Hydrogen electrode (unless otherwise noted).

1.3: Goals Of This Thesis

The work detailed in this thesis deals with the effect of the protein on the assembly and control over the heme E_m , two aspects that play a significant role in function. I aim to uncover the physical chemical principles that underlie both of these processes and clearly demonstrate them in a simple protein system.

Based on the goals of this thesis, the maquette approach is the best suited to answer the questions compared to other design procedures. The benefit of maquettes comes from the simple scaffold unburdened by evolutionary complexity. The physical chemistry of a certain process is more clearly demonstrated as a single change does not have allosteric or long-range effects that cannot be explained. The simple scaffold has does not have any overlapping amino acid functions or irreducible complexity seen in a natural proteins. Each amino acid can be directly traced back to the chemical principles underlying their role in the sequence. Adding to that simplicity is the rigorous characterization these maquette proteins are subject to for every mutation.

Through the efforts of this thesis, a better understanding of oxidoreductases becomes clear. The physical chemistry that underlies the assembly kinetics of these proteins is investigated along with the midpoint control. These are two critical aspects of oxidoreductase function that are poorly understood in natural proteins.

1.4: References

- (1) Smith, L. J.; Kahraman, A.; Thornton, J. M. *Proteins* **2010**, *78*, 2349.
- (2) de Montellano, P. R. O. *Curr Opin Chem Biol* **2000**, *4*, 221.
- (3) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem Rev* **2005**, *105*, 2253.
- (4) Kranz, R. G.; Richard-Fogal, C.; Taylor, J. S.; Frawley, E. R. *Microbiol Mol Biol R* **2009**, *73*, 510.
- (5) Makino, R.; Park, S.; Obayashi, E.; Iizuka, T.; Hori, H.; Shiro, Y. *J Biol Chem* **2011**, *286*.
- (6) Alves, T.; Besson, S.; Pereira, A. S.; Pettigrew, G. W.; Moura, J. J. G.; Moura, I. *J Inorg Biochem* **2001**, *86*, 122.
- (7) Daldal, F. *Bba-Bioenergetics* **2004**, *1658*, 63.
- (8) Diaz, F.; Kotarsky, H.; Fellman, V.; Moraes, C. T. *Semin Fetal Neonatal Med* **2011**, *16*, 197.
- (9) Baniulis, D.; Yamashita, E.; Zhang, H.; Hasan, S. S.; Cramer, W. A. *Photochem Photobiol* **2008**, *84*, 1349.
- (10) Mulkidjanian, A. Y. *Bba-Bioenergetics* **2005**, *1709*, 5.
- (11) Kim, H. J.; Khalimonchuk, O.; Smith, P. M.; Winge, D. R. *Bba-Mol Cell Res* **2012**, *1823*, 1604.
- (12) Nelson, N.; Yocum, C. F. *Annu Rev Plant Biol* **2006**, *57*, 521.
- (13) Kumar, S.; Bandyopadhyay, U. *Toxicol Lett* **2005**, *157*, 175.
- (14) Buchachenko, A. L.; Kuznetsov, D. A.; Breslavskaya, N. N. *Chem Rev* **2012**, *112*, 2042.
- (15) Soto, I. C.; Fontanesi, F.; Liu, J. J.; Barrientos, A. *Bba-Bioenergetics* **2012**, *1817*, 883.
- (16) Carr, H. S.; Winge, D. R. *Accounts Chem Res* **2003**, *36*, 309.
- (17) Wang, Z. H.; Wang, Y. X.; Hegg, E. L. *J Biol Chem* **2009**, *284*, 839.
- (18) Minagawa, J.; Takahashi, Y. *Photosynth Res* **2004**, *82*, 241.
- (19) Beale, S. I.; Cornejo, J. *Arch Biochem Biophys* **1983**, *227*, 279.
- (20) Grigg, J. C.; Ukpabi, G.; Gaudin, C. F. M.; Murphy, M. E. P. *J Inorg Biochem* **2010**, *104*, 341.
- (21) Fournier, C.; Smith, A.; Delepelaire, P. *Mol Microbiol* **2011**, *80*, 133.
- (22) Tolosano, E.; Altruda, F. *DNA Cell Biol* **2002**, *21*, 297.
- (23) Weissbuch, I.; Leiserowitz, L. *Chem Rev* **2008**, *108*, 4899.
- (24) Costa, C. A.; Trivelato, G. C.; Pinto, A. M. P.; Bechara, E. J. H. *Clin Chem* **1997**, *43*, 1196.
- (25) Brennan, M. J. W.; Cantrill, R. C.; Kramer, S. *Int J Biochem* **1980**, *12*, 833.
- (26) Paiva-Silva, G. O.; Cruz-Oliveira, C.; Nakayasu, E. S.; Maya-Monteiro, C. M.; Dunkov, B. C.; Masuda, H.; Almeida, I. C.; Oliveira, P. L. *P Natl Acad Sci USA* **2006**, *103*, 8030.
- (27) Kuras, R.; de Vitry, C.; Choquet, Y.; Girard-Bascou, J.; Culler, D.; Buschlen, S.; Merchant, S.; Wollman, F. A. *J Biol Chem* **1997**, *272*, 32427.
- (28) Palombo, I.; Daley, D. O. *Febs Lett* **2012**, *586*, 4197.
- (29) Argos, P.; Rossmann, M. G. *Acta Crystallogr A* **1975**, *31*, S55.
- (30) Argos, P.; Rossmann, M. G. *Biochemistry-U S* **1979**, *18*, 4951.
- (31) Rossmann, M. G.; Argos, P. *J Biol Chem* **1975**, *250*, 7525.
- (32) Aldag, C.; Gromov, I. A.; Garcia-Rubio, I.; von Koenig, K.; Schlichting, I.;

- Jaun, B.; Hilvert, D. *P Natl Acad Sci USA* **2009**, *106*, 5481.
- (33) Zhang, P.; Malolepsza, E.; Straub, J. E. *J Phys Chem B* **2012**, *116*, 6980.
- (34) Culbertson, D. S.; Olson, J. S. *Biochemistry-U.S.* **2010**, *49*, 6052.
- (35) Culbertson, D. S. Doctoral, Rice University, 2010.
- (36) Sturtevant, J. M.; Robinson, C. R.; Liu, Y. F.; Thomson, J. A.; Sligar, S. G. *Biochemistry-U.S.* **1997**, *36*, 16141.
- (37) Davis, R. B.; Lecomte, J. T. J. *Biopolymers* **2008**, *90*, 556.
- (38) David L. Nelson, M. M. C. *Lehninger Principles of Biochemistry*; Third Edition ed.; Worth Publishers: New York, NY, 2000.
- (39) Swierczek, M.; Cieluch, E.; Sarewicz, M.; Borek, A.; Moser, C. C.; Dutton, P. L.; Osyczka, A. *Science* **2010**, *329*, 451.
- (40) Rajagukguk, S.; Bryant, J.; Merbitz-Zahradnik, T.; Rajagukguk, R.; Pielak, G.; Trumpower, B.; Durham, B.; Millett, F. *Biophys J* **2004**, *86*, 470A.
- (41) Shifman, J. M.; Gibney, B. R.; Sharp, R. E.; Dutton, P. L. *Biochemistry-U.S.* **2000**, *39*, 14813.
- (42) Hendler, R. W.; Sidhu, G. S.; Pardhasaradhi, K. *Biophys J* **1990**, *58*, 957.
- (43) Doi, M.; Shioi, Y.; Morita, M.; Takamiya, K. *Eur J Biochem* **1989**, *184*, 521.
- (44) Michel, L. V.; Ye, T.; Bowman, S. E. J.; Levin, B. D.; Hahn, M. A.; Russell, B. S.; Elliott, S. J.; Bren, K. L. *Biochemistry-U.S.* **2007**, *46*, 11753.
- (45) Olsson, M. H.; Ryde, U. *J Biol Inorg Chem* **1999**, *4*, 654.
- (46) Feng, Y. Q.; Wand, A. J.; Sligar, S. G. *Biochemistry-U.S.* **1991**, *30*, 7711.
- (47) Stevens, J. M.; Uchida, T.; Daltrop, O.; Ferguson, S. J. *Biochem Soc T* **2005**, *33*, 792.
- (48) Jungwirth, U.; Kowol, C. R.; Keppler, B. K.; Hartinger, C. G.; Berger, W.; Heffeter, P. *Antioxid Redox Signal* **2011**, *15*, 1085.
- (49) Sligar, S. G.; Cinti, D. L.; Gibson, G. G.; Schenkman, J. B. *Biochem Biophys Res Commun* **1979**, *90*, 925.
- (50) Zhou, H. X. *J Biol Inorg Chem* **1997**, *2*, 109.
- (51) Shifman, J. M.; Moser, C. C.; Kalsbeck, W. A.; Bocian, D. F.; Dutton, P. L. *Biochemistry-U.S.* **1998**, *37*, 16815.
- (52) Mao, J. J.; Hauser, K.; Gunner, M. R. *Biochemistry-U.S.* **2003**, *42*, 9829.
- (53) Reddi, A. R.; Reedy, C. J.; Mui, S.; Gibney, B. R. *Biochemistry-U.S.* **2007**, *46*, 291.
- (54) Dutton, P. L.; Moser, C. C. *Faraday Discussions* **2011**, *148*, 443.
- (55) Matsui, T.; Nagano, S.; Ishimori, K.; Watanabe, Y.; Morishima, I. *Biochemistry-U.S.* **1996**, *35*, 13118.
- (56) Yeung, N.; Lin, Y. W.; Gao, Y. G.; Zhao, X.; Russell, B. S.; Lei, L.; Miner, K. D.; Robinson, H.; Lu, Y. *Nature* **2009**, *462*, 1079.
- (57) Hecht, M. H.; Das, A.; Go, A.; Bradley, L. H.; Wei, Y. N. *Protein Sci* **2004**, *13*, 1711.
- (58) Jackel, C.; Kast, P.; Hilvert, D. *Annu Rev Biophys* **2008**, *37*, 153.
- (59) Moffet, D. A.; Foley, J.; Hecht, M. H. *Biophys Chem* **2003**, *105*, 231.
- (60) Tian, P. *Chem Soc Rev* **2010**, *39*, 2071.
- (61) Samish, I.; MacDermaid, C. M.; Perez-Aguilar, J. M.; Saven, J. G. *Annual Review of Physical Chemistry, Vol 62* **2011**, *62*, 129.
- (62) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364.

- (63) Fleishman, S. J.; Baker, D. *Cell* **2012**, *149*, 262.
- (64) Lichtenstein, B. R.; Farid, T. A.; Kodali, G.; Solomon, L. A.; Anderson, J. L.; Sheehan, M. M.; Ennist, N. M.; Fry, B. A.; Chobot, S. E.; Bialas, C.; Mancini, J. A.; Armstrong, C. T.; Zhao, Z.; Esipova, T. V.; Snell, D.; Vinogradov, S. A.; Discher, B. M.; Moser, C. C.; Dutton, P. L. *Biochem Soc Trans* **2012**, *40*, 561.
- (65) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; Degrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425.
- (66) Dutton, P. L. *Methods Enzymol* **1978**, *54*, 411.

Chapter 2: Maquette Development and History

2.1: What Is A Maquette?

As mentioned in the introduction we aim to use maquettes to learn about a protein fold or function by recreating it from the most basic chemical principles in a simple scaffold. This can only be done if the scaffold itself is well understood such that the role of each amino acid is characterized. The simple scaffold we use in this work is a 4-helix bundle, initially developed by Regan and Degrado in 1989.¹

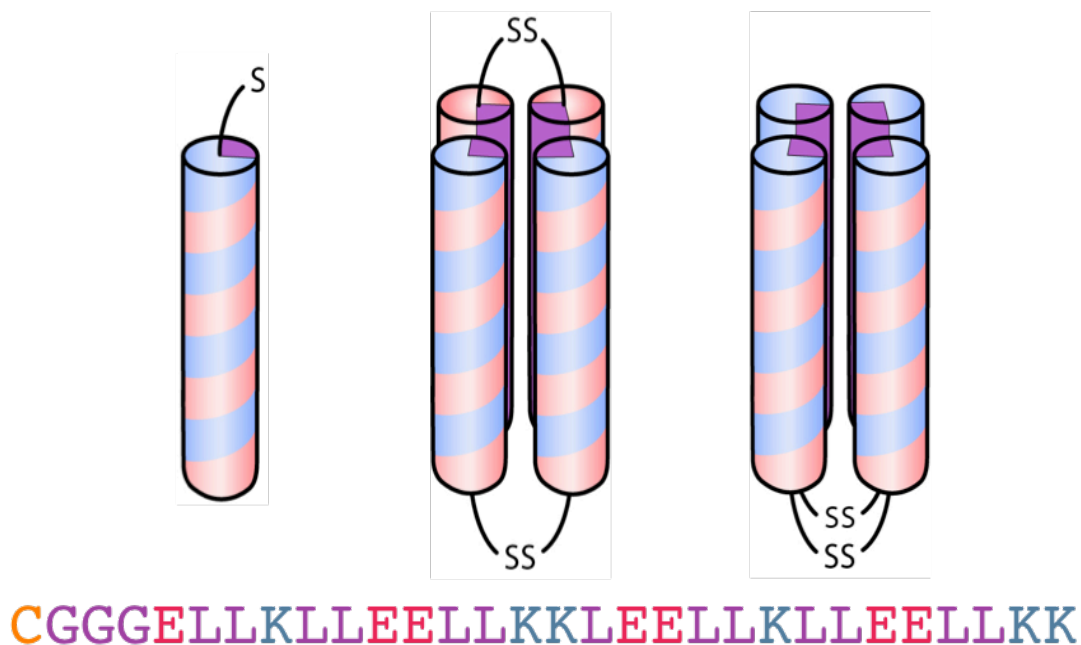


Figure 2.1: Demonstration of binary patterning in a simple diagrammatic cartoon. Blue color represents positively charged residues, red negatively charged and purple hydrophobic. Arranging these residues, based on their properties, as shown allows for the spontaneous formation of a four alpha-helix bundle in solution. The nonpolar residues are buried, and the charged residues form salt bridges along the outside of the helices. Shown here is a single helix (left), that can oligomerize and assume two different topologies: syn (middle), and anti (right).

These bundles, from the earliest Degrado variant to the most recent positive variant, all have some basic attributes in common. First is the structure. All of the maquettes are 4- α -helix bundles. They achieve this state by binary patterning, the positioning of their

positive, negative, and hydrophobic charges to maximize helical propensity and drive the formation of the bundle architecture seen all throughout maquettes in literature (Figure 2.1). From the original three amino-acid helix (Sequence Appendix) the sequence has been altered many times which will be touched on below, however a more thorough review can be found in Lichtenstein et. al..²

2.1.1: History Of Maquettes

The first maquette created was a basic 4- α -helix bundle introduced by DeGrado and Regan. This was both the original and the simplest protein seen being comprised of only 3 amino acid types. Its design was solely aimed at assembling into a tetrameric four-helix bundle with no further biologically relevant function. Its simplicity lead to it becoming the scaffold for future designs.

In order to increase the functional abilities of the protein carefully considered mutations were introduced. Most importantly histidine residues were added buried in the protein core in order for heme to bind and cysteines were added in order for the tetramer to link functionally creating a dimer of dimers. This protein, named H10H24 was able to bind four heme B cofactors and all of them had discernable midpoint potentials and K_d values.³ This protein was later changed into H10A24, replacing one of the histidine residues per helix with alanines residues.⁴ With these mutations each dimer unit can only bind one heme, making electrochemical and affinity characterization less complicated.

The H10H24 protein was subjected to structural studies as well.⁵ Huang et. al. were able to use this protein as a scaffold to generate a protein from which a crystal structure was obtained. This provided structural information, useful for both the current designs and as a guide for future variants (BB, Sequence Appendix).

In the crystal structure the protein dimers were stably in an anti conformation. This was exploited to create a new variant. No longer four single helices, this new protein was dimer of two helix-loop-helix motifs. This structural change revealed many possibilities, as the symmetry of the protein was now reduced from four identical helices to two identical helix-loop-helix motifs.^{6,7} Heme-binding site knockouts were possible and the sequence could be better dispersed for NMR. This protein was further modified to include a disulfide bond covalently linking the dimers together. This formed a structure colloquially referred to as a candelabra, noted for its resemblance to the tool used to hold candles. With this protein, previous members of the laboratory were able to observe an oxyferrous state, the first example in a designed protein.⁸

This candelabra was converted into a true monomer by linking the two-dimer units with a peptide chain comprised of glycine residues. This new protein, named the "Single Chain" has the same sequence and orientation of helices as the candelabra however there is less conformational freedom at the candle end. This new change also makes the symmetry even lower, as the entire sequence can be expressed as a single gene, whereas in the candelabra it was expressed as two dimers that were disulfide-linked.⁹

Though this appears to be a simple change converting a dimer into a monomer, it has much more significant effects. This change reduces the symmetry of the protein, a major hindrance to new functions. Binding sites can be designed that incorporate dipoles and other uneven attributes not possible when the amino acids on either side are identical. These changes have lead to the creation of single histidine binding sites, which can discriminate between Fe and Zn tetrapyrrole cofactors. This confers site specificity in maquettes and can also be used to incorporate different ligations not possible before.

2.2: The Current Generation Of Maquettes And Their Attributes

The single chain monomer shows a high degree of thermal stability and tolerance to sequence changes, both in the core and the surface.

The high α -helical content of these proteins makes circular dichroism a useful tool for determining the stability of these maquettes.^{9,10} Figure 2.2 shows the temperature melts for the single chain protein as well as variants whose sequences have been altered both in the core, and on the surface. The changes allow for a high level of control over the T_m of the protein. Core packing interactions can be changed to increase the stability of the apo state, as can surface charges to generate more salt bridges. NMR can give information on the tertiary structure of the protein in many ways complimentary to CD. It shows stability changes in the form of increased tertiary structure. The single chain for example gains a significant amount of structure upon heme addition, depicted as shifts and clarity of NMR peaks (Figure 2.3). This structure is brought on by the tethering effect of the heme when it forms coordination bonds to the histidine residues on each helix. These coordination bonds also reduce the ability of the helices to rotate, further increasing their structural rigidity.

An important feature of all maquettes, from the earliest up to the current single chain, is heme binding. Previous work by Gibney et. al. has shown where in the sequence histidine residues can best be placed to maximize the heme affinity.⁴ The A-position of a standard helical wheel is important as it allows for the heme to be buried in a hydrophobic core. This has been maintained throughout all heme-binding maquette variants, which has lead to sub-nanomolar affinities (Figure 2.4). This is an important result, ensuring that not only are the proteins stable when mutated, but heme binding can be readily monitored and adverse mutations can be easily identified.

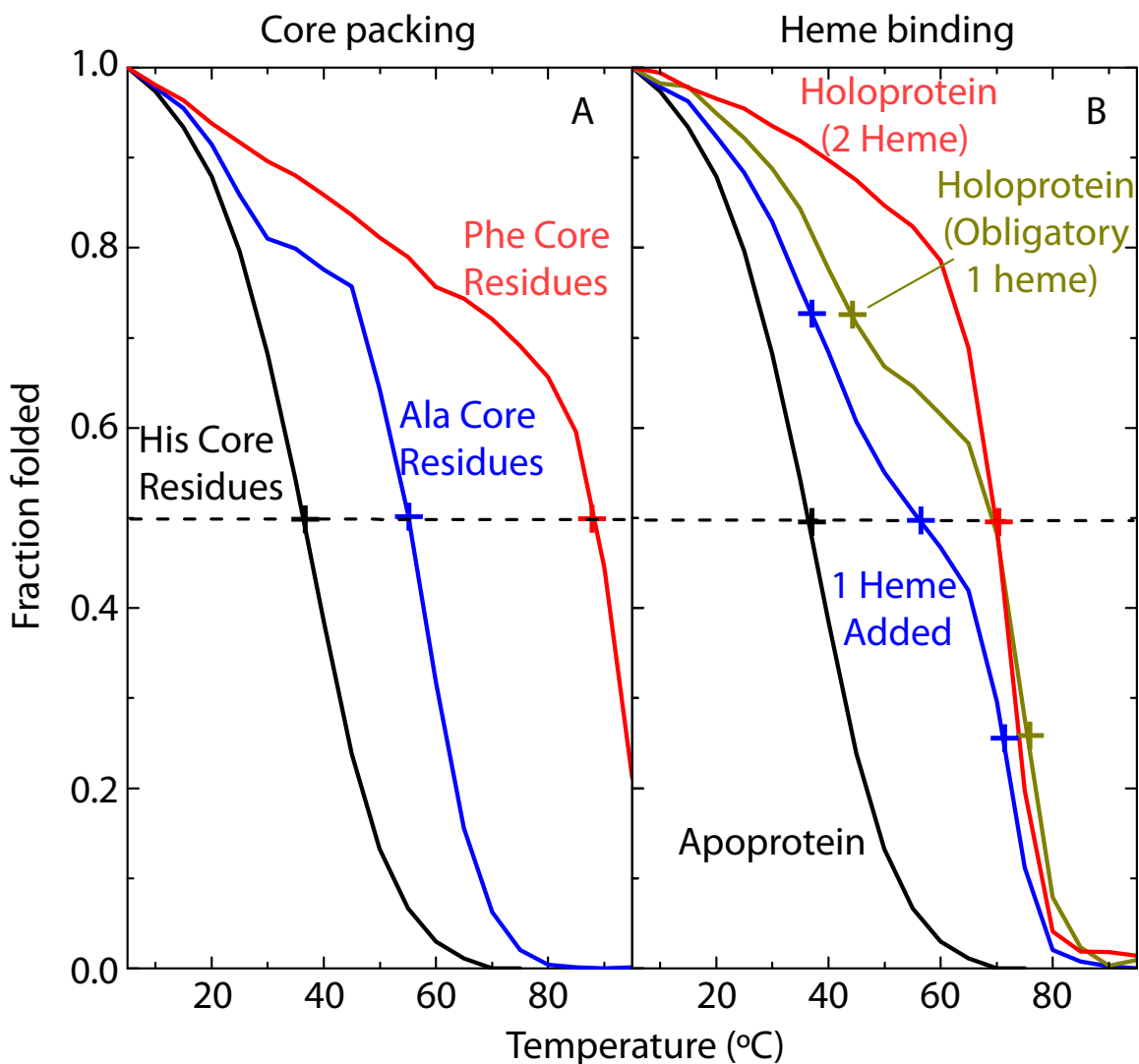


Figure 2.2: Thermal denaturation of α -helical content monitored as a loss of circular dichroism at 222 nm. **(A)** T_m of the apo maquette increases from 37 to 55 and 95 $^{\circ}\text{C}$ as the Histidines (black) are replaced by Alanines (blue) and Phenylalanines (red), respectively. **(B)** The impact of heme binding on the thermal stability of the helices is shown. Addition of one heme to Single chain reveals two transitions at 38 and 72 $^{\circ}\text{C}$ (blue). Similarly, binding one heme to a single-heme binding maquette has two separate transitions at 39 and 75 $^{\circ}\text{C}$. Upon adding 2 hemes to the single chain, there is a single T_m transition at 72 $^{\circ}\text{C}$, a 35 degrees increase compared to the single chain apo maquette. All CD experiments were performed with 20 μM protein in 20 mM CHES at pH 9 with 150 mM KCl. The CD signal amplitude is normalized to 1 at 5 $^{\circ}\text{C}$. Figure and Legend reproduced from Farid, Kodali, Solomon et. al.. Nature Chem. Biol. 2013: Submitted⁹

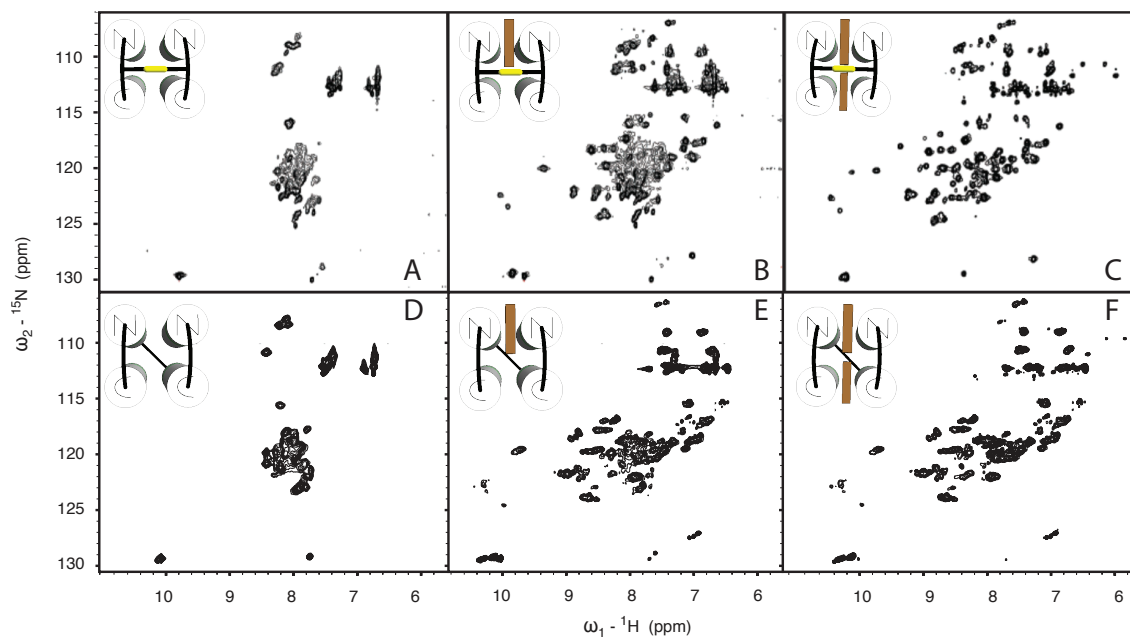


Figure 2.3: 750 MHz ^{15}N -HSQC showing changes in NMR spectral dispersion upon heme binding. Similar results are seen for the homo-dimer candelabra (top) and the single chain version Single chain (bottom). In the absence of cofactor, NMR resonances are relatively undispersed (left column). Addition of one equivalent of Fe(III) protoporphyrin IX (center column) induces partial dispersion, indicative of stable structure. Addition of a second equivalent of heme (right column) induces further structuring. NMR of candelabra was performed with 350 μM protein in 25 mM KH_2PO_4 , pH 6.6 at 18 $^\circ\text{C}$. NMR of Single chain was performed with 200 μM protein in 50 mM KH_2PO_4 , 50 mM KCl, pH 7.9, at 25 $^\circ\text{C}$. Data from Farid, Kodali, Solomon et. al. Nature Chem. Biol. 2013: Submitted⁹

2.3: Potential Of Maquettes Past The Scope Of This Work

Though not part of the scope of this work, it bears mentioning that the maquette platform can be used to study more than just heme B cofactors. It has in the past been shown to allow for c-type heme maturation, covalently attaching a heme to the protein allowing for massive changes in the protein to not affect where the heme is or whether or not it is bound. Maquettes have also been used to study non-heme organic cofactors such as Flavins. Sharp and Farid have both shown flavin attachment coupled to light activated

electron transfer in a maquette platform.^{9,11} Quinone has also been attached covalently to the maquette.¹² The work of Lichtenstein et al have developed a amino acid cofactor which can be easily incorporated into a maquette backbone with simple Merrifield chemistry. This has advantages over the cofactor work relying on a disulfide bond, as the attachment to the protein is not susceptible to redox changes.

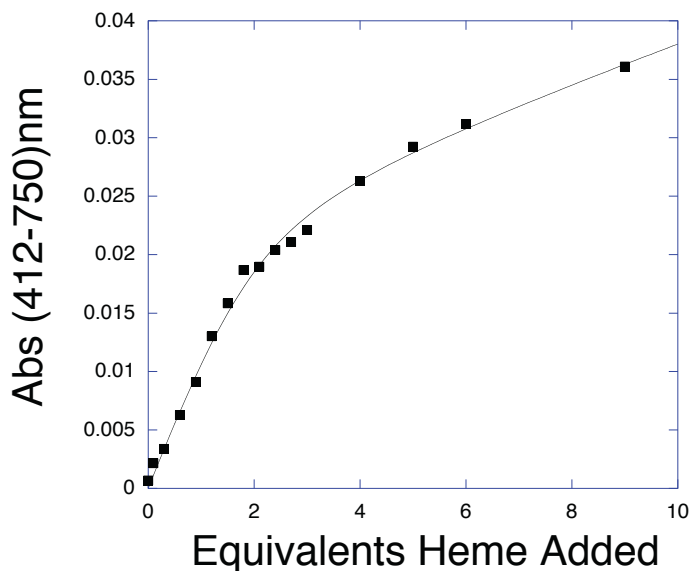


Figure 2.4: K_d titration data for the single chain protein (at 6nM), This experiment was carried out at room temperature in 20mM CHES 150mM KCl buffer pH 9. K_d value calculated as less than 2nM, listed in Table 3.1.

These various cofactor attachments have allowed for a lot of functions to be studied in maquettes. Flavin coupled with heme has been shown by sharp and Farid to be capable of light induced electron transfer. The benefit to doing this in maquettes is that they allow for the control over positions in the protein. Distances can be adjusted to control specific electron transfer rates based on established guidelines. Oxygen binding has also been described in this system. Maquettes, like certain natural proteins in the globin family, are able to exclude water from the protein core such that an oxyferrous state is stabilized.⁹

2.4: References

- (1) Regan, L.; Degrado, W. F. *Science* **1988**, *241*, 976.
- (2) Lichtenstein, B. R.; Farid, T. A.; Kodali, G.; Solomon, L. A.; Anderson, J. L. R.; Sheehan, M. M.; Ennist, N. M.; Fry, B. A.; Chobot, S. E.; Bialas, C.; Mancini, J. A.; Armstrong, C. T.; Zhao, Z. Y.; Esipova, T. V.; Snell, D.; Vinogradov, S. A.; Discher, B. M.; Moser, C. C.; Dutton, P. L. *Biochem Soc T* **2012**, *40*, 561.

- (3) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; Degrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425.
- (4) Gibney, B. R.; Dutton, P. L. *Protein Sci* **1999**, *8*, 1888.
- (5) Huang, S. S.; Gibney, B. R.; Stayrook, S. E.; Dutton, P. L.; Lewis, M. *J Mol Biol* **2003**, *326*, 1219.
- (6) Huang, S. S.; Koder, R. L.; Lewis, M.; Wand, A. J.; Dutton, P. L. *P Natl Acad Sci USA* **2004**, *101*, 5536.
- (7) Koder, R. L.; Valentine, K. G.; Cerda, J.; Noy, D.; Smith, K. M.; Wand, A. J.; Dutton, P. L. *J Am Chem Soc* **2006**, *128*, 14450.
- (8) Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.
- (9) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton1 In *Nat Chem Biol* 2013; Vol. Submitted.
- (10) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry-Us* **1998**, *37*, 4635.
- (11) Sharp, R. E.; Moser, C. C.; Rabanal, F.; Dutton, P. L. *P Natl Acad Sci USA* **1998**, *95*, 10465.
- (12) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun* **2009**, 168.

Chapter 3: The Protein Influence On Holoprotein Assembly Kinetics

3.1: Introduction

The main goal of this section is to uncover the physical chemistry of heme and protein assembly. Throughout the literature, the thermodynamics have been investigated significantly more than the kinetics.¹ Despite how much is known about the thermodynamics of heme-protein assembly, the kinetics generally remain poorly understood. Hargrove et. al. explored the effect of mutating various residues in the heme cavity of myoglobin looking for a change in binding rate of heme B, demonstrating how individual residues do not contribute significantly to this process.² Various groups have looked at the binding rates of proteins such as cyt *b*₅ and cytochrome *c* peroxidase but still no formal set of engineering guidelines have emerged.^{3,4}

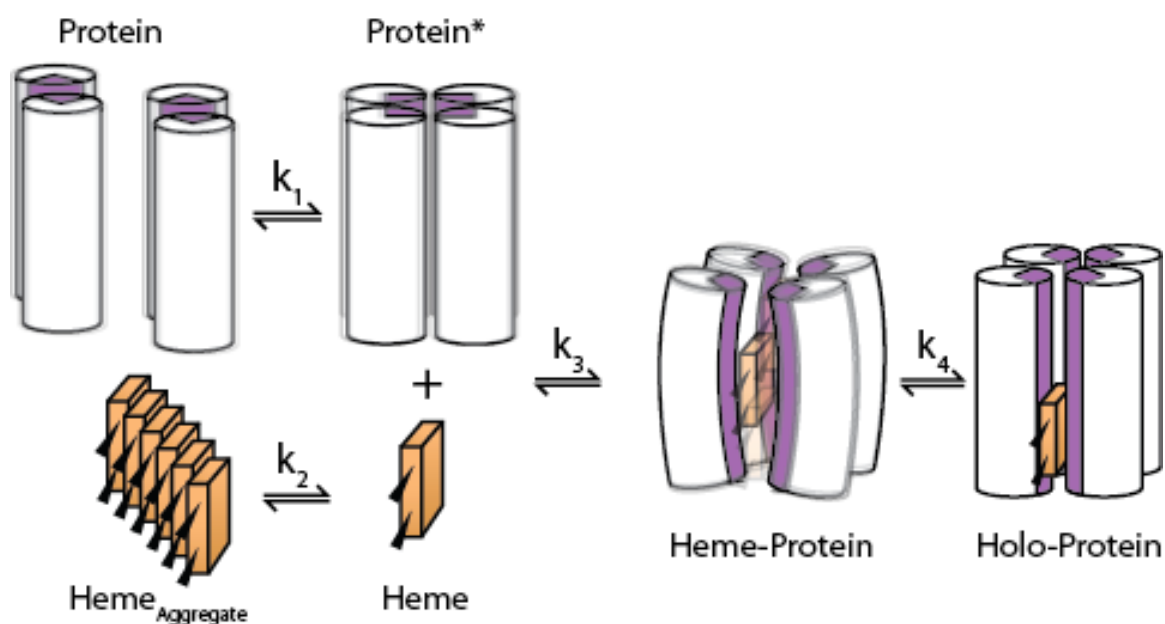
In this chapter, I present the study of heme-protein assembly kinetics using protein maquettes and the natural protein Cytochrome *b*₅₆₂. These proteins are all 4- α -helix bundles with nanomolar heme B affinities. The low K_d values allow me to focus squarely on the assembly mechanism, and characterize the steps involved in holoprotein construction.

This work will illustrate how assembly kinetics can be incorporated into maquette designs. The barriers are clearly shown, and ways to circumvent them are explained. When these concepts are applied, a designed protein capable of an assembly reaction that rivals natural proteins can be generated.

3.2: Kinetics Of Holoprotein Assembly

3.2.1: Proteins Used For This Study

Each protein in this paper is limited at a different stage of the process depending on its structural characteristics (Scheme 3.1, Table 3.1). k_1 presents protein pre-assembly, depicted as the coming together of two dimers, though more generally referring to the protein accessing a structure suitable for heme-binding. The third rate, k_3 , is the partitioning step. This rate constant represents the speed at which a cofactor moves from the solvent to the protein core. k_4 , the final rate, is the ligation step, representing the cofactor finding the proper residue and forming a coordination bond as well as the protein structurally conforming to the presence of heme. k_2 depicts the cofactor deaggregation. This step pertains to heme or other cofactors being in a soluble, monomeric and non-aggregated state freely available to associate with protein, this is the topic of chapter 4.



Scheme 3.1: Reaction scheme for the protein Tetrapyrrole binding reaction. k_1 demonstrates the protein forming an appropriate binding site. Concurrently k_2 illustrates the the cofactor also becoming able to assemble. k_3 then depicts the rate of partitioning forming an unligated complex of the cofactor and protein, where the heme is in the core but not bound. Finally, this complex finds its final conformation at a rate of k_4 leading to the final protein.

	Structured HomoTetramer	Unstructured HomoTetramer	Homodimer	Tethered Homodimer	Single Chain	Cytochrome <i>b</i> -562
Structured in Apo	Yes	No	No	No	No	No*
Structured in Holo	No	No	Yes	Yes	Yes	Yes
Kd (nM)	2	15	NA	<1	<1	6
Limiting Rate	k_3	k_1	k_1	k_2	k_2	k_2

Table 3.1: Protein topologies, structural states, heme B affinities, and limiting rates of the proteins used in this chapter. The topologies are depicted as cartoons of the various designs. Helices are depicted as cylinders and loops as curves. Curves with “SS” in the middle denote a disulfide bond linking loop sections together. Structure in apo and holo states is based on NMR data: Structured Homotetramer and Unstructured homotetramer,⁵ Tethered Homodimer,⁶ Single Chain,⁷ and *b*₅₆₂.⁸ Heme B affinities for Unstructured Homotetramer, Tethered Homodimer and *b*₅₆₂ were obtained from the literature.¹⁰⁻¹² Single Chain, Structured Homotetramer and Homodimer are from this work.

K_d values for the structured homotetramer, and homodimer had to first be obtained as the rest have been obtained previously from the literature.⁹⁻¹¹ Figure 3.1 shows these experiments. These experiments were performed as described in the methods section and were fitted as described in the methods section. All of these affinities are in the same general nanomolar range. The homodimer is an exception, it is unable to generate a K_d value, as, at the concentrations required for this assay, the protein is not dimerized in solution and thus not in a structurally associated state.

To explore the homodimer’s lack of apparent binding, a reverse titration was done. In this experiment, instead of heme being added to a fixed amount of protein, the homodimer protein was added to a fixed amount of heme. Figure 3.1B shows this

experiment. At low concentrations, there appears to be little binding, which begins to increase in slope, achieving a maximum at $2\mu\text{M}$. Below $2\mu\text{M}$ the protein is unable to hold its dimer form dissociating into solution, whereas above $2\mu\text{M}$ it is able to form a sufficient binding site for the heme to assemble with.

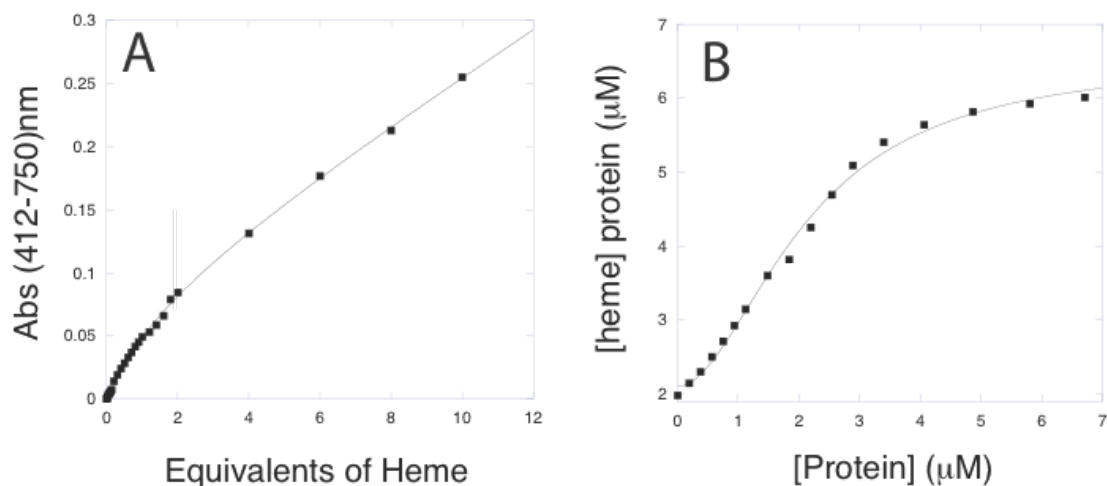


Figure 3.1: K_d titration data for (A) the structured homotetramer (at 50nM), and (B) the homodimer ($6\mu\text{M}$ heme). All experiments were carried out at room temperature in 20mM CHES 150mM KCl buffer pH 9. Structured homotetramer titration was performed using a 10cm path length cuvette. The homodimer was not carried out in the same conditions as it is unable to bind heme at nM concentrations. For this protein, the heme concentration was set at $6\mu\text{M}$, and the protein was added in aliquots as depicted in (B). K_d value for structured homotetramer is calculated as $<2\text{nM}$, listed in Table 3.1.

3.2.2: Kinetic Characterization Of The Proteins Used

The first step in examining this kinetic mechanism was to determine the order of the reaction. This was done by rapidly mixing a range of concentrations of protein with excess heme or an excess of the single chain protein with a range of concentrations of heme. In all of the following assembly experiments assembled proteins were monitored by the shift in absorbance at the Soret band (Figure A3.1-3). As shown in figure 3.2 all but two of these

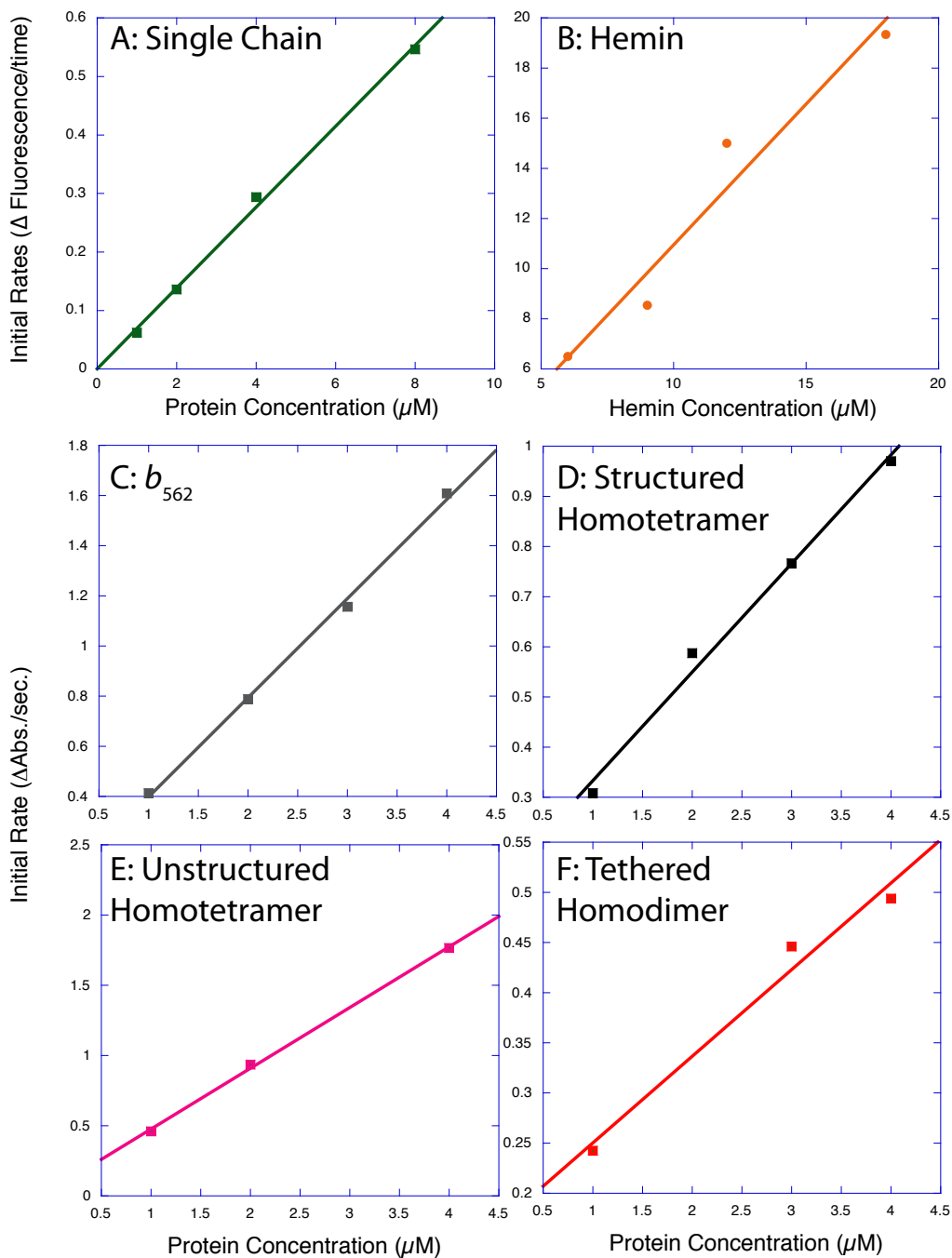


Figure 3.2: Plots of rates versus concentration of the single chain protein (A), Cyt-b562 (C), Structured homotetramer (D), Unstructured homotetramer (E), and the Tethered homodimer (F). The rates were taken from the initial slope of the reaction (first 20msec.). All experiments were carried out at 20°C in 20mM CHES 150mM KCl buffer pH 9, heme at 20 μ M. (B) Rates of heme vs concentration were carried out with excess single chain protein, 30 μ M, in 20mM CHES 150mM KCl buffer pH 9 at 20°C. Rates for the single chain and heme were obtained on a fluorescence stopped flow monitoring the change of emissions at 350nm in the first 20 msec.

proteins follow standard second order behavior (1st order with respect to both reactants) increasing in rate commensurate with concentration. Panel B shows that the heme itself is also first order in this reaction, increasing in rate commensurate with concentration.

The tethered-homodimer and the homodimer require further explanation. The homodimer appears to have two distinct phases depending on the concentration (Figure 3.3). In the first twenty milliseconds of the reaction there is a noticeable burst phase for the three and four micromolar concentrations. This is most likely brought on by the tendency of the protein to be partially assembled in solution at these concentrations whereas below these concentrations, it behaves in a fashion more resembling a monomeric structure. There is a slight burst phase in the two-micromolar trace however no noticeable one in the one-micromolar trace. Once the burst phase is over (approximately 20msec.) the kinetics appear to once again have standard first order kinetics within the error of the

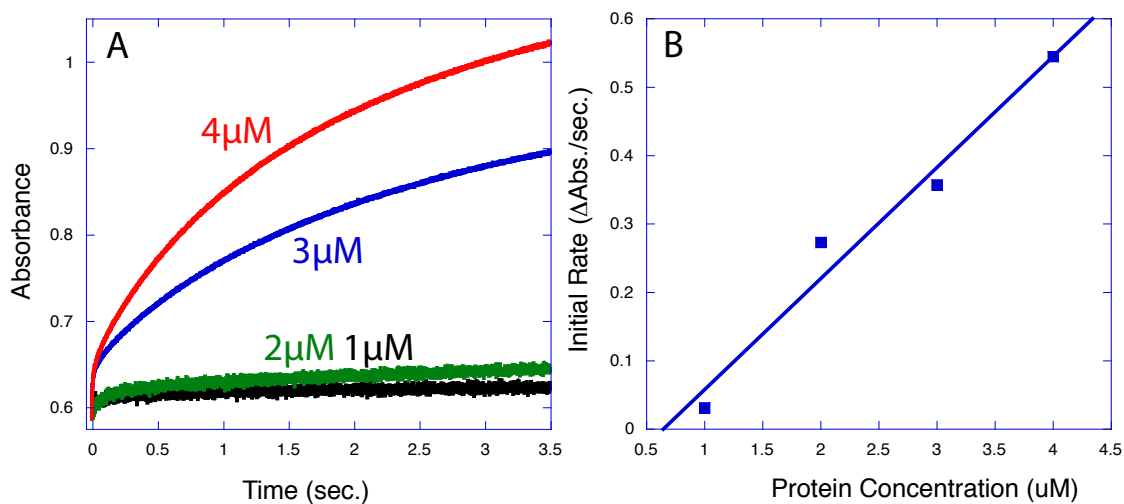


Figure 3.3: (A) Absorbance at 412nm for the homodimer protein over 3.5 seconds tracking the assembly, and (B) rates as a function of concentration of protein. At the initial 20 milliseconds of the reaction there is a short lived burst phase in the 2, 3, and 4 μM traces but not in the 1 μM sample. Based on the affinity data this is due to the lack of associated monomers in solution. The rates in (B) follow first order kinetics with respect to the protein above the 1 μM sample. All data was collected at 20°C in 20mM CHES 150mM KCl buffer pH 9. Heme at 20 μM

measurements (Figure 3.3, Panel B). The tethered homodimer in contrast does not have a burst phase. Its kinetics are not first order with respect to the protein but appear to be $\frac{1}{2}$ order, with the total reaction being 1.5 order overall.

I next had to determine any differences in the rates of different binding sites. Of the six proteins tested, the homodimer, tethered homodimer, and the single chain have two heme-binding sites. Both the unstructured and structured homotetramer proteins have two sites but one cannot be removed as any change to the sequence is bound to the four-fold symmetry of these maquettes. Cyt-*b*₅₆₂ only has one binding site. The single-chain protein and tethered homodimer were compared to each site knockout. In these proteins, either the 7 or 42 histidine was removed (Sequence Appendix). In the single-chain protein the removal of residues 7 and 42 yield one mono-his heme-binding sites per bundle. Mono-histidine sites have such a low affinity they can be ignored in this experiment. Figure 3.4 shows the effect of these knockouts on the binding kinetics. In the single-chain the rates between H7A (200msec) and H42A (159msec) differ by approximately 40msec, sufficiently close to treat them as the same. The tethered homodimer mutant is not as simple. The standard variant and the H42 site knockout appear to have the same rate (halftime of 250msec) however, the H7 site knockout is approximately 3 times slower (halftime 750msec). What this means is the H7 site potentially requires a pre-assembly step, one that the H42 sites does not. This pre-assembly is achieved upon heme binding to the H42 site thus explaining why the variant with both sites achieves the same kinetics as the H7-knockout. Moving forward, the effect of knocking out a binding site can be ignored as when one site is not present the rate does not appear to be affected compared to both sites intact.

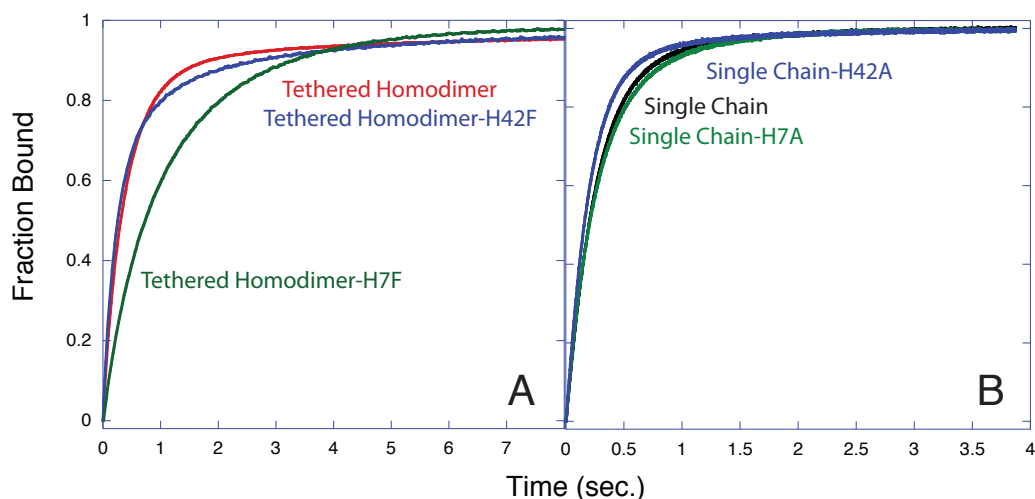


Figure 3.4: Rates for (A) the Tethered homodimer and its site knockout mutants. The H42F and wild-type variants share similar rates, but the H7F variant binds heme much slower. This is likely due to a mild pre-assembly reaction not present in H42F or the “wild-type variant”. (B) The single chain protein and its individual site knockouts, unlike the tethered homodimer, all share the same kinetics with halftimes differing by 40 msec.. All data was collected at 20°C in 20mM CHES 150mM KCl buffer pH 9.

Finally, I had to determine the differences in binding rate. The six variants and heme B were rapidly mixed at 30°C. Due to the bis-histidine ligation, heme bound to protein has an absorbance at 412nm far to the red of free heme (395nm). Figure 3.5 shows each protein has different kinetics spanning a wide range of rates. The heme-binding of the three fastest proteins, single-chain, *b*₅₆₂, and tethered homodimer all had halftimes within 500 msec (rate constants of $3.04 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $3.04 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, and $1.73 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ respectively, Table 2). This contrasts with the three other proteins: homodimer, structured homotetramer, and unstructured homotetramer, which had much slower assembly rate constants ($3.24 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$, $7.3 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$, and $2.68 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ respectively Table 3.2). From this data one piece of

information becomes immediately clear: proteins that are obligatory monomeric in solution are able to assemble faster than ones in which the four-helix bundle architecture is

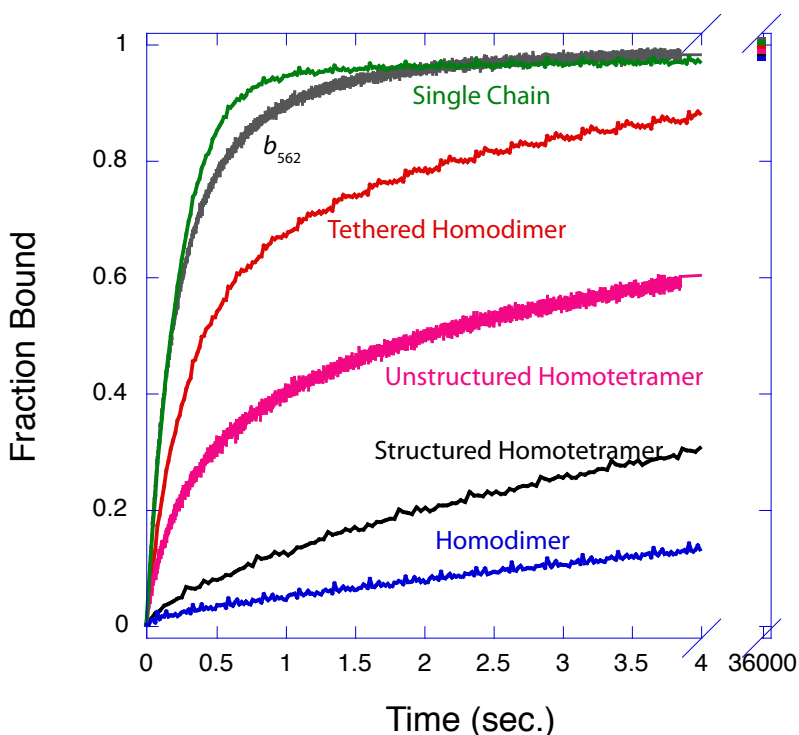


Figure 3.5: Comparison of the binding rates of heme and the various proteins used. The proteins that are monomeric in solution achieve the fastest rate (single chain, tethered homodimer, and b_{562}) whereas removal of the monomeric character substantially slows this process (homodimer, unstructured homotetramer, and structured homotetramer). All data was collected at 30 degrees Celsius in 20mM CHES 150mM KCl buffer pH 9. All proteins are at $5\mu\text{M}$, Heme is at $10\mu\text{M}$

composed of separate pieces. This could be due to either having the dimers come together in solution or having to rearrange them into a suitable form for assembly. The specific barriers and reasons for the unique kinetics were not clear from this data; another approach was needed.

3.2.3: Temperature Dependent Kinetic Analysis

Temperature dependent kinetics were used to identify the specific barriers hindering assembly. Figure 3.6 shows what is known as an Eyring plot.¹² This analysis uses the rate constants at various temperatures to explore the transition state. The underlying theory treats the transition state and the ground state reactants as being in equilibrium. This mathematical expression generates pseudo free energy from the equilibrium, which can be

further broken down into pseudo enthalpic and entropic components. These give a

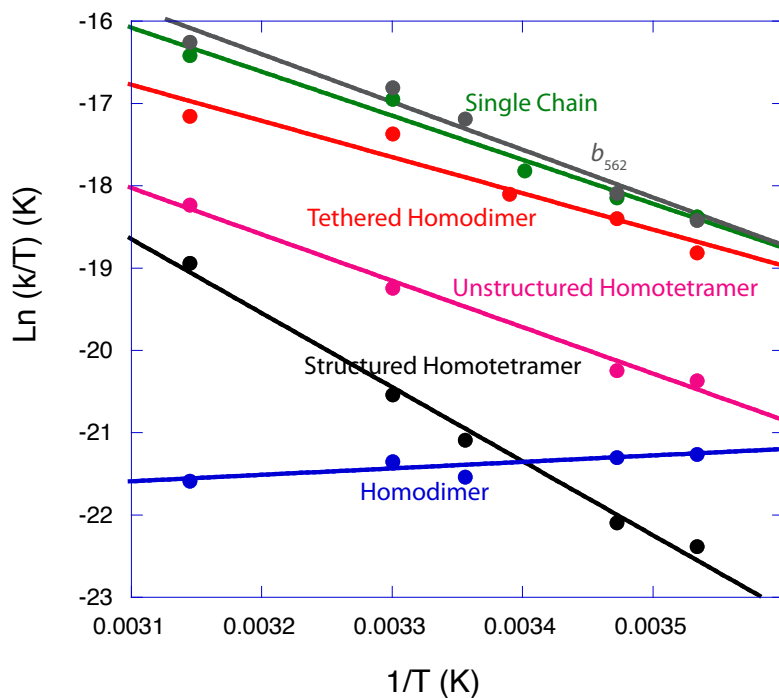


Figure 3.6: Eyring plot of the rates as a function of temperature for the homodimer (blue), structured homotetramer (black), unstructured homotetramer (pink), tethered homodimer (red), single chain (green) and Cyt b_{562} . All rates were taken from initial slopes of 412nm absorbance converted into concentration of holoprotein at various temperatures.

valuable insight into how the transition state is energetically limited and can indicate how the reaction is limited generally.

The Eyring analysis clearly demonstrated the barriers for each protein.

In the case of the structured homotetramer, a protein known to have a rigid apo structure,¹³ heme assembly is the most sensitive to

temperature. This behavior is due to its high transition state enthalpy (17.9 kcal/mol), brought on by its tightly packed apo-core (Table 3.2, Figure 3.6).

This is contrasted with the unstructured homo-tetramer. The sequence of these two proteins differs by four residues in each helix (I6L, F14L, H24A, and M31L; Sequence Appendix). These changes cause its structure to be unstructured on NMR timescales in both apo and holo states (Table 3.1). This dramatically increases the heme assembly rate by lowering the transition state (TS) enthalpy to 11 kcal/mol (Table 3.2). The transition state

entropy is higher (48.3 cal/mol-K), though does not impose a barrier as intense as the enthalpy.

The homodimer assembly is very slow and appears to be independent of temperature, indicative of an entropically limited transition state, has a ΔH^\ddagger of -1559 cal/mol and ΔS^\ddagger of 94 cal/mol-K (Table 3.2). Previous reports have often attributed this to active site rearrangements and other protein folding reactions, which are possible here as well.¹⁴⁻¹⁶ However, the data do not make clear what exactly is happening.

The tethered-homodimer and homodimer are identical in sequence and function, except for a cysteine residue replacing a serine in the loop region (Sequence Appendix, Table 3.1). This cysteine creates a disulfide bond between the two dimers, which makes the protein a monomer in solution. This seemingly minor change has a major effect on the rate decreasing the entropy to -53.44 cal/mol-K and increasing the enthalpy 8.73 kcal/mol (Table 3.2). These changes greatly increase the rate over the temperature range monitored and, since the entropy is no longer the major barrier, reinstates the temperature dependence seen with the other maquettes (Figure 3.6).

The single-chain protein shares the majority of its sequence with the tethered homodimer protein. The exception being an amino acid loop converting two dimers into a monomer (Sequence Appendix). Compared to the homodimer, this loop decreases the

Table 3.2: Summary of rates and transition state parameters of the heme binding reaction of the proteins used in this chapter.

Protein	k at 30°C (303K) (M ⁻¹ s ⁻¹)	ΔH^\ddagger (cal M ⁻¹)	ΔS^\ddagger (cal M ⁻¹ K ⁻¹)
Structured Homotetramer	7.3x10 ³	17.87x10 ³	-28.85
Unstructured Homotetramer	2.68x10 ⁴	11.20x10 ³	-48.3
Homodimer	3.24x10 ³	-1.56x10 ³	-94.94
Tethered Homodimer	1.73x10 ⁵	8.73x10 ³	-53.44
Single Chain	3.04x10 ⁵	10.62x10 ³	-46.22
<i>b</i> ₅₆₂	3.04x10 ⁵	11.53x10 ³	-42.89

entropic barrier to -46.22 cal/mol-K, and increases the enthalpy to 10.6 kcal/mol (Table 3.2). These parameters combine to yield the fastest reaction rate constant of all the proteins in this work.

Cytochrome b_{562} is unlike the other 5 designed maquettes in the sense that it is a natural protein present in *E. coli* thought to be involved in electron transfer.¹⁷ This protein binds heme at a rate apparently equal to the SC maquette, both having comparable transition state parameters and rate constants (Table 2).

3.2.4: Analysis Of The Assembly Kinetics

The above data uncover differences in the rates, and kinetic limitations of the 6 protein variants, and describes where in scheme 3.1 they are limited. Generally, the two major limiting steps for the protein seem to be k_1 and k_3 , no protein appears limited at the ligation step k_4 . Rate k_1 can be described clearly by the data in figure 3.5, as the assembly rates vary significantly depending on the oligomeric state of the protein, seen here with the monomeric proteins assembling with the highest rate (Figure 3.5). Single-chain b_{562} and the tethered homodimer bind heme at the highest rate. These three proteins are all monomeric in solution. The other three proteins are all dimeric in solution, and bind heme at a much slower rate. For example, the homodimer protein does not have any designed interactions to promote monomeric character and is therefore unable to bind heme at an appreciable rate. The addition of either a disulfide bond or amino acid loop connection drastically improves the assembly rate, facilitating the generation of a heme-binding pocket by either keeping the individual helices of the protein close to one another, or limiting its ability to rearrange in solution. Rearrangements have been seen before in maquettes, in the work on the H10S24 protein of Grosset et. al.,¹⁸ This protein had the ability to assume either a syn

or anti conformation depending on the redox state of the heme. This entropic barrier would obviously slow the assembly rate as the protein would have to conformationally rearrange in order to form the binding site for heme. Having a linker, stopping the protein from rearranging, would lower this barrier and increase the rate as seen with the proteins here.

This behavior is similar to that of some natural proteins; for instance cytochrome *b*₅ consists of two domains, one of which binds heme.¹⁹ When it is in its apo form the heme-binding domain is unstructured, only folding into its native state upon heme binding. The other domain, which acts analogously to our third loop or disulfide bond linker, holds it together. If this other domain were not present the heme-binding loops would be unable to conform to heme, as they would not be held near each other rendering the protein useless.

The core packing of the protein is the other major barrier to the assembly rate. This enthalpic barrier is the result of having to break apart interior van der Waals interactions and water shell hydrogen bonds. An example of this is the structured homotetramer. This protein is known to have a rigid apo state brought on by core packing interactions and hydrogen bonds between histidine residues.⁹ In order for a heme to bind, these must first be broken which requires a significant amount of thermal energy slowing the reaction. If these interactions are removed, and the apo structure is destabilized as with the unstructured homotetramer protein, the rate increases substantially. These two proteins share a high level of sequence overlap (85% homology) differing only in their core packing, yet the assembly rate constants differ by a factor of 3.7 (Table 2, Sequence Appendix). This barrier goes from being limited at k_3 to being limited at k_1 .

Based on the similarity in rate over the temperature range shown between single-chain and *b*₅₆₂, and their unrelated sequences it seems reasonable to suggest that the rate-limiting step is likely to be the availability of porphyrin in solution (k_2 in scheme 3.1).

Heme as a cofactor when free in solution is well known to form soluble oligomers.²⁰ However, the protein is only able to bind monomers of heme, therefore the assembly is most likely limited by the monomer availability for these two proteins.

The variants with the fastest rate have in common a molten globular apo state, that upon heme binding, changes to a rigidly structured holo state. It seems likely that the apo state makes available many energetically equal microstates for the heme to bind as opposed to one specific binding ready conformation. The proteins also have monomeric character. Being in a single structure helps to lower the ΔS^\ddagger by limiting its ability to assume other topologies as well as keep pieces from dissociating.

3.3: Proof Of A Partitioned Heme State

Of the various states in the mechanism discussed above (Scheme 3.1), the only one that cannot be directly monitored in the course of a normal reaction is the partitioned state. In order to prove its existence we used a protein that resembles the single-chain, except for four mutations changing the histidine residues to phenylalanine, the single-chain-F (Sequence Appendix). This allows for all of the steps to occur except for histidine ligation.

In figure 3.7 we see that, upon combining the histidine-less protein with heme cofactor in solution, there is a change in the spectra. The shoulder on the free heme spectra decreases while the peak at 400nm sharpens and increases a small amount. The spectra appear to be shifting to resemble more of heme in DMSO, a solvent where it is well known to not aggregate (Spectral Appendix: Figure A3.1). This data indicates that the heme in the protein core is monomeric and able to remain in that state.

To further observe this state and confirm its existence the absorbance was monitored again. However, in this experiment, rather than heme, which has a minor

spectral shift in the UV-Vis absorbance range, Zn-protoporphyrin IX (ZnPPIX) was used as it

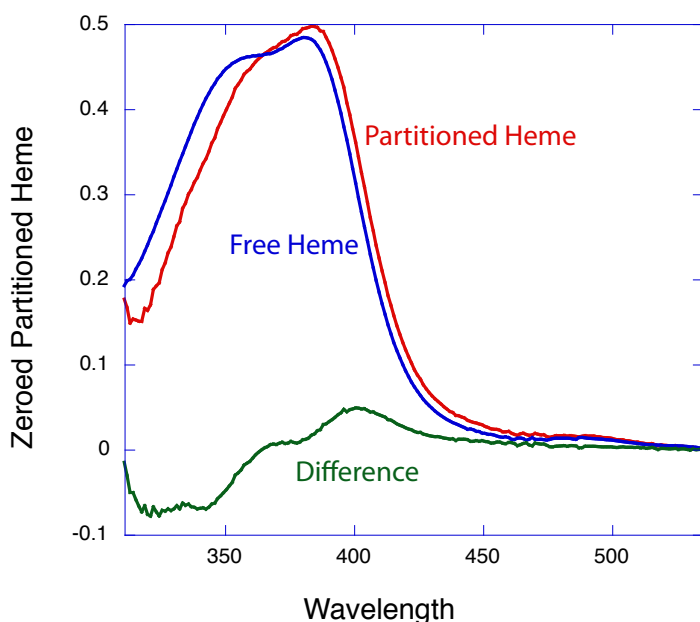


Figure 3.7: Heme has a spectral shift when mixed with a protein that has no histidine residues (red) compared to free heme in solution (blue). The green trace depicts the difference spectrum of the other two (protein - free). Work done in CHES buffer pH 9 at room temperature

has a larger difference between its aggregate and partitioned states. In figure 3.8, we see a shift in the Soret band of the ZnPPIX over the course of 90 minutes after addition to an aqueous solution of histidine-less protein. This slow change is consistent with the breaking up of the aggregates as the porphyrin partitions into the protein core. It is not clear from this data whether the protein is directly interacting with the

porphyrin-aggregate, acting like a deaggregate, or just associating with the small amount of free porphyrin in solution and shifting the equilibrium. However, without the ability to ligate, the hydrophobic interactions appear strong enough to allow for some interactions between the two species that causes a spectral shift consistent with an increase in monomer concentration.

To once more prove that the protein was interacting with the porphyrin and absorbance changes were not artifactual, I carried out a fluorescence titration on single chain, single chain-F, and a small peptide with the sequence WAGWA (Figure 3.9). WAGWA was chosen for its solubility. This experiment monitored a drop in the tryptophan

fluorescence signal due to the presence of heme quenching it. This only occurs when heme is near a tryptophan residue taking part in an energy transfer reaction. In this figure, we

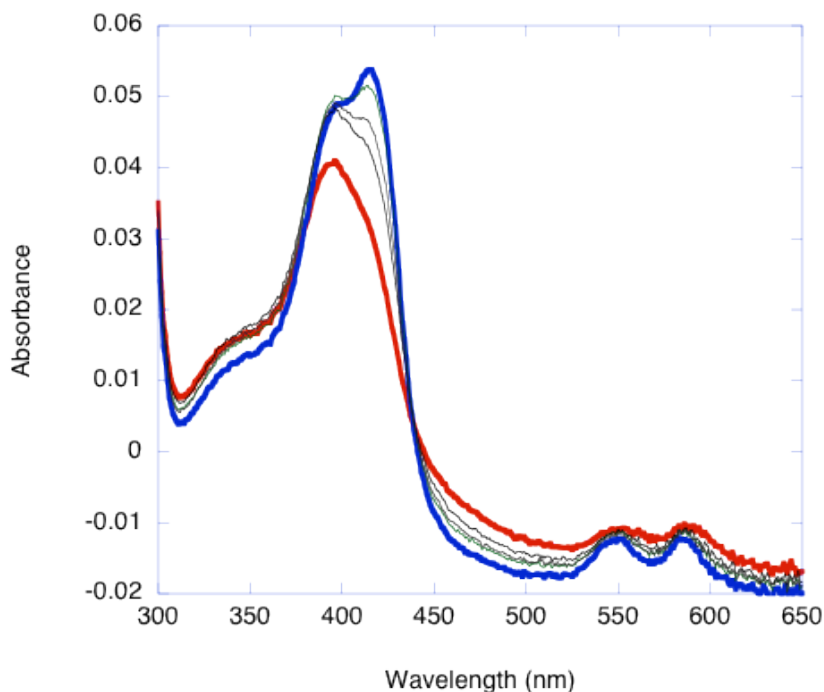


Figure 3.8: ZnPIX associating with the single-chain histidine-less protein. There is a noticeable spectral shift between the starting spectra (red) and the spectra after 90 minutes (blue). Data collected in 20mM CHES 150mM KCl buffer pH 9 at room temperature.

see an expected drop in tryptophan fluorescence as heme is added to the single-chain protein. The sharp drop in signal persists until the heme binding sites are saturated at which point there is no signal change. The WAGWA peptide has no core into which to partition nor does it

have any sort of binding site for the heme to associate. The decrease in fluorescence in this case is due to the background absorbance of light by the heme and any collision based quenching that occurs in solution. The third trace, single chain-F, has a quenching profile that is in between the two, indicating that it has more heme associated with it than the WAGWA but not as much as the ligation-able single chain. This can be explained by the heme partitioning into the protein core of single chain-F without being held there tightly, thus dissociating.

The spectral changes in concert with the fluorescence titration demonstrate the presence of a state between that of the heme-ligated state and that of the heme-unbound state. The spectra begin to

assume characteristics similar to that of heme in DMSO, which is in agreement with the literature as non-aggregated spectra. These states form too quickly to have their kinetics traced in the single chain, tethered homodimer, and

*b*₅₆₂ proteins. In order for this state to be observed it requires a protein that cannot ligate to heme with any residues. Structured homotetramer is unable to retain its

structure without the histidine residues in the core, as they are responsible for maintaining its structure in the apo state and the holo state therefore observing a heme-partitioned state is impossible. This state is also not observable in the homodimer protein as it requires the

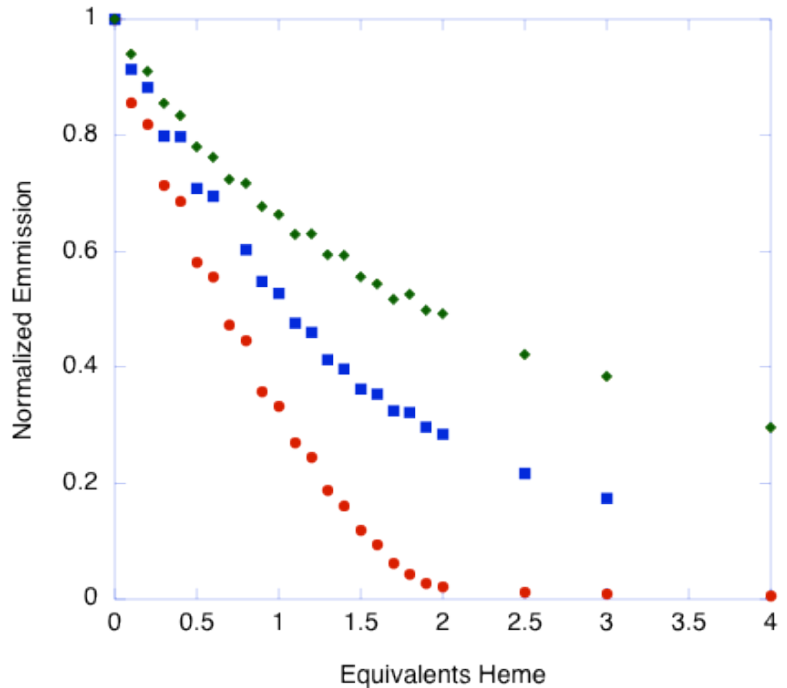


Figure 3.9: Fluorescence titration shows the quenching of tryptophan when heme is added. The single chain protein (red circles) loses all of its fluorescence intensity at 2 equivalents of heme, while the free WAGWA peptide (green diamonds) loses only half at this point. The His-less protein (blue squares) loses an intermediate amount indicating that, despite its low affinity, heme is partitioned in or held near the Trp residues. Data collected in 20mM CHES 150mM KCl pH 9 buffer at room temperature. Excitation: 280nm, Emission at 350nm. Emissions normalized to the signal at 0 eq. heme added

heme as a structural piece to form the binding site. Without the histidine residues this protein cannot dimerize much less form a cavity for the heme to partition into.

3.4: Conclusions

The different temperature response behaviors can be used to probe which aspect of the reaction is rate limiting for each protein and make clear what alterations can be made to speed up the assembly rate. Using the data above as the basis of a model, it is possible to compare one protein's heme binding rate over a range of temperatures and uncover the limiting step of assembly. This will be a useful tool for protein designers going forward, helping them to explain why a particular design would have a slower assembly rate than another.

For a maximally fast rate two features must be at the core of the design: Monomeric character and molten globularity of the apo-form. The monomeric character significantly lowers the entropy of the transition state, which manifests as either diffusional assembly of dimers or topological rearrangements. This is most clearly seen in k_1 and has been observed in previous maquette designs.

A molten globular apo-core can considerably increase the rate by lowering the enthalpy. This concept reduces the ΔH^\ddagger by removing any van der Waals interactions holding the protein core together. With these removed, heme insertion into the core is easier, as described by k_3 . Together these two parameters are the most pertinent to protein designs aiming to control the kinetics of the assembly reaction.

The concept of a molten globular core has been discussed previously.²¹ Many protein design methodologies energetically minimize the structure they are generating in the presence of a specific cofactor such that it has little to no motion at its lowest energetic

state. This makes the holoprotein very rigid with the hope that it tightly binds its cofactor. This can, however, have the problem of obstructing binding by shifting the barrier to k_3 as the core too rigid to allow heme to enter.

The existence of a heme-partitioned state fills out the last place of the scheme. If a protein were able to bind cofactors without the need for ligating amino acids, different metals could be inserted into the tetrapyrrole skeleton and the redox properties would be significantly expanded.

While being a boon for protein design, this work also sheds light on the natural process of heme assembly. Many proteins are unstructured in their apo forms and the addition of heme induces a rigid holoprotein. It is possible this is an evolved trait aiming to optimize heme-proteins assembly. This is not clear from any of the above data, as maquettes do not resemble natural proteins in any way, but the underlying physical chemistry applies to any four- α -helix bundle. Cytochrome b_{562} follows these principles, and it is likely that others would as well if examined more closely.

We have developed, in this chapter, a complete understanding of the mechanism and rate of the heme assembly in protein maquettes. This work builds on previous studies of Hargrove et. al. that stated specific residues are not important to the binding rate.² Here we agree with that work and expand on it; we demonstrate how the gross overall conformation and topology of the protein are critical, being the points at which the transition state is most affected. These concepts appear to match natural proteins such as b_{562} , as this protein for example, like the single chain, undergoes a structural stabilization upon the apo to holo transition. More generally, this approach is applicable for natural and designed proteins. It can uncover the course of binding and remove barriers in the way.

3.5: References

- (1) Argos, P.; Rossmann, M. G. *Acta Crystallogr A* **1975**, *31*, S55.
- (2) Hargrove, M. S.; Barrick, D.; Olson, J. S. *Biochemistry-U.S.* **1996**, *35*, 11293.
- (3) Hunter, C. L.; Lloyd, E.; Eltis, L. D.; Rafferty, S. P.; Lee, H.; Smith, M.; Mauk, A. G. *Biochemistry-U.S.* **1997**, *36*, 1010.
- (4) Asakura, T.; Kobayash.T; Chance, B. *J Biol Chem* **1974**, *249*, 1799.
- (5) Gibney, B. R.; Huang S. S.; Skalicky J. J.; Fuentes E. J.; Wand A. J.; Dutton, P. L. *Biochemistry* **2001**, *40*, 10550.
- (6) Koder, R. L.; Valentine, K. G.; Cerda, J.; Noy, D.; Smith, K. M.; Wand, A. J.; Dutton, P. L. *JAm Chem Soc* **2006**, *128*, 14450.
- (7) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton1 In *Nat Chem Biol* 2013; Vol. Submitted.
- (8) Feng, Y. Q.; Wand, A. J.; Sligar, S. G. *Biochemistry-U.S.* **1991**, *30*, 7711.
- (9) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry-U.S.* **1998**, *37*, 4635.
- (10) Sturtevant, J. M.; Robinson, C. R.; Liu, Y. F.; Thomson, J. A.; Sligar, S. G. *Biochemistry Us* **1997**, *36*, 16141.
- (11) Zhang, L.; Anderson, J. L. R.; Ahmed, I.; Norman, J. A.; Negron, C.; Mutter, A. C.; Dutton, P. L.; Koder, R. L. *Biochemistry-U.S.* **2011**, *50*, 10254.
- (12) Houston, P. L. *Chemical Kinetics and Reaction Dynamics*; First ed.; Dover Publications Inc.: Mineola, NY, 2001.
- (13) Huang, S. S.; Gibney, B. R.; Stayrook, S. E.; Dutton, P. L.; Lewis, M. *J Mol Biol* **2003**, *326*, 1219.
- (14) Borgmann, U.; Moon, T. W.; Laidler, K. J. *Biochemistry-U.S.* **1974**, *13*, 5152.
- (15) Hewson, W. D.; Dunford, H. B. *Can J Chem* **1975**, *53*, 1928.
- (16) Eftink, M. R.; Anusiem, A. C.; Biltonen, R. L. *Biochemistry-U.S.* **1983**, *22*, 3884.
- (17) Itagaki, E.; Hager, L. P. *Biochem Bioph Res Co* **1968**, *32*, 1013.
- (18) Grosset, A. M.; Gibney, B. R.; Rabanal, F.; Moser, C. C.; Dutton, P. L. *Biochemistry Us* **2001**, *40*, 5474.
- (19) Davis, R. B.; Lecomte, J. T. J. *Biopolymers* **2008**, *90*, 556.
- (20) Kuzelova, K.; Mrhalova, M.; Hrkal, Z. *Bba-Gen Subjects* **1997**, *1336*, 497.
- (21) Huang, S. S.; Koder, R. L.; Lewis, M.; Wand, A. J.; Dutton, P. L. *P Natl Acad Sci USA* **2004**, *101*, 5536.

Chapter 4: The Cofactor Influence On Holoprotein Assembly Kinetics

4.1: Introduction

In this chapter the assembly of a holoprotein is examined with the effect of the cofactor being the main focus. The previous chapter only addressed the cofactor itself to show that it is first order with respect to the reaction kinetics. Being first order, it therefore plays an integral part in the assembly and its contributions need to be addressed.

The only previous work that examined the kinetics of maquette-porphyrin binding was performed by Gibney et. al..¹ Investigating the difference in the binding of heme B and heme A in a designed 4-helix bundle, they observed heme A assembles at a slower rate. This was attributed to the farnesyl tail inhibiting requiring burial in the core. Other groups have examined the effect of adding alternative hemes to globin proteins however, the underlying physical chemistry was not addressed.^{2 3} The effect of porphyrin aggregation on the assembly rate was previously addressed, however this was only with heme B, other cofactors with differing structures were not considered.⁴ Using Hemopexin they were able to describe the effect of aggregation on the assembly rate in a variety of environmental conditions. The work in this chapter aims to build upon that previous work, describing how varying structures of tetrapyrrole cofactors affect the rate.

In this chapter, we fill in the final piece of Scheme 3.1. I aim to clarify how chemical aspects of the porphyrin control the assembly rate, and how it can be managed using a variety of natural and synthetic porphyrins structurally related to heme B.

4.2: Results

4.2.1: Fe-Porphyrin Assembly And The Effect Of Substituents

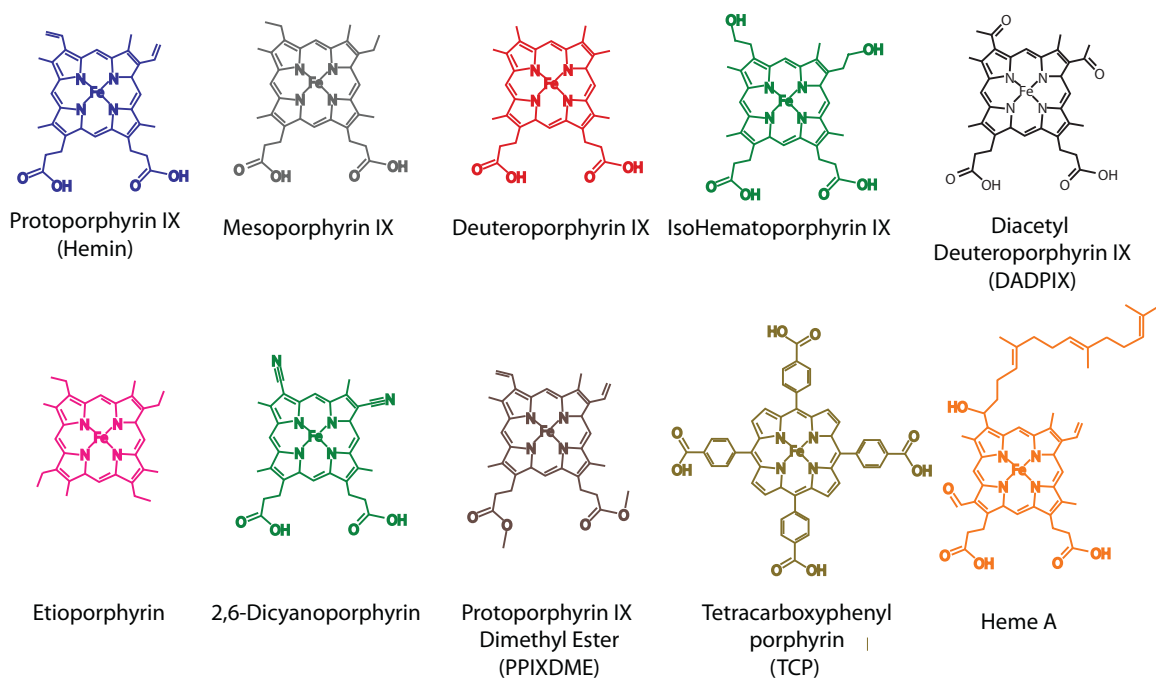


Figure 4.1: The Library of porphyrins used in this chapter. These porphyrins were chosen based on their diverse structures and substituent characteristics

In order to elucidate the effect the porphyrin has on holoprotein assembly we selected a series of porphyrins with different attributes to compare (Figure 4.1). The porphyrins selected all are standard tetrapyrroles with an iron center, however their ring substituents differ significantly as depicted in Figure 4.1. These different porphyrins were chosen based on their variety of substituents and, based on common chemical knowledge of water interactions, can be broadly divided up into three classes: (1) Hydrophobic porphyrins (Etioporphyrin, Protoporphyrin IX Dimethyl Ester,) contain side-chains that do not have strong polar interactions with water; (2) Hydrophilic porphyrins (IsoHematoporphyrin, Tetracarboxyphenyl Porphyrin) have side-chains that are polar or charged, therefore able to hydrogen bond to water; (3) Amphiphilic porphyrins (Protoporphyrin IX, Mesoporphyrin IX, Deuteroporphyrin IX, 2,6-Dicyanoporphyrin,

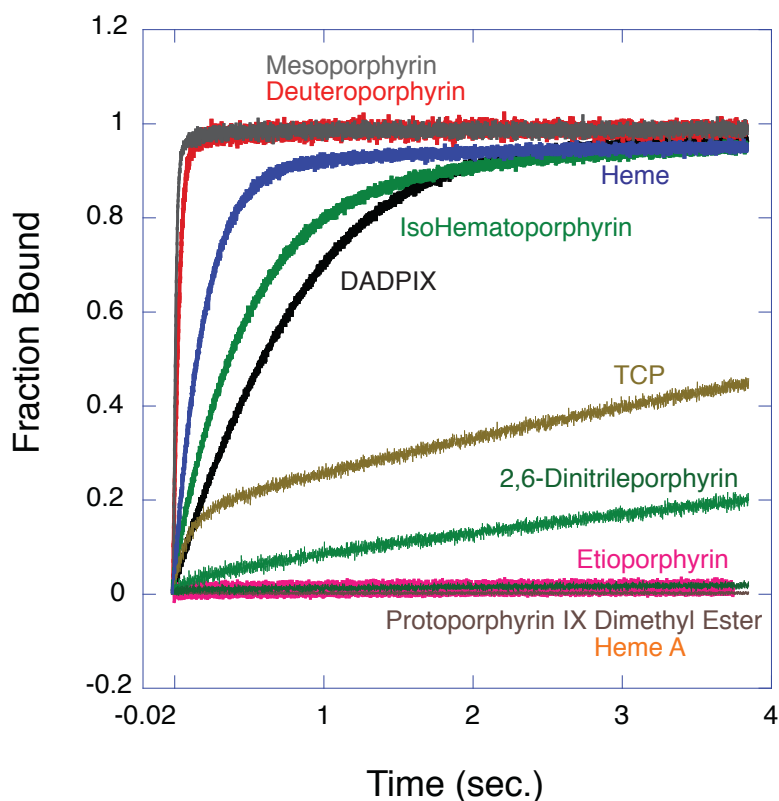


Figure 4.2: Assembly traces of the porphyrins from figure 1 (5 μ M) and the single chain protein (24 μ M). Deuteroporphyrin (red), Mesoporphyrin (grey), Heme (blue), Iso Hematoporphyrin (forest green), Diacetyl Deuteroporphyrin (black), Tetracarboxyphenyl porphyrin (gold), 2,6-Dinitrile porphyrin (light green), Etioporphyrin (pink), Protoporphyrin IX Dimethyl ester (brown), and Heme A (Orange). All data collected in 20mM CHES 150mM KCl buffer pH 9 at 25°C .

single-chain protein mentioned previously (rate constants listed in table 4.1). What emerges from the data is a general trend of the amphiphilic proteins with small substituents binding with faster rates than the hydrophobic or hydrophilic porphyrins. Hydrophobic porphyrins have the hardest time binding even with substituents as small as ethyl groups, as is the case with Etioporphyrin.

Diacetyl Deuteroporphyrin IX, and Heme a), which have both hydrophobic and hydrophilic side-chains. All of these are shown in Figure 4.1.

The various porphyrins were the rapidly mixed with the single chain proteins from the previous chapter and the binding rate was measured as the change in the Soret band of the porphyrin as a function of time (Soret bands A3.1-A3.12). Figure 4.2 shows this series of porphyrins binding to the

Porphyrin	Rate constant M⁻¹ s⁻¹	Log-P (CHES pH 9)	K_d Values (nM)	E_m (mV)
Hemin	82068	0.46254	<2	-261
Mesoporphyrin	484079	0.55195	<10	-319
Deuteroporphyrin	176639	0.65704	80	-268
IsoHematoporphyrin	58649	0.1196	1120	-236
DADPIX	30371	0.12379	<10	-48
Etioporphyrin	3796	1.4218	--	NA
FePPIX-DME	3140	0.85306	--	NA
Tetracarboxyphenyl porphyrin	12557	0.02769	--	-320
2,6-Dintrile porphyrin	7118	0.036413	--	*+150
Heme a	683	0.93846	--	+145

Table 4.1: Rate Constants, Partition Coefficients, K_d values (where possible), and E_m values for the Fe-porphyrins in this work.

There appeared to be no direct relation between the rates and the assigned classes of hydrophobicity. I next quantified the solubility to better classify the porphyrins, and again investigate a relationship with the rates. This was measured in Partition coefficients (log-P values) between octanol and 20mM CHES 150mM KCl buffer pH 9. The porphyrins were mixed in a 50/50 octanol/buffer mixture and allowed to incubate for one hour at room temperature. After this incubation the concentration in each phase were determined (Table 4.1). Figure 4.3 shows the rate constants plotted against the log P values. In this figure there appears to be two trends. Mesoporphyrin is at the apex and from there if a porphyrin becomes either more or less hydrophobic the rate appears to drop. This can be explained by an amphiphilic-centered hypothesis. Having to bury a polar group inside a hydrophobic protein core is energetically difficult causing the rate to slow, and conversely having to get insoluble porphyrins in solution is also difficult as those porphyrins have a propensity to aggregate. The ideal cofactor is one that has both hydrophobic character to aid in burial and hydrophilic character to increase its solubility. Figure 4.4 demonstrates

this with heme, the standard cofactor in this work. The heme vinyl groups contribute hydrophobic character and are easily buried in the protein core whereas the propionate groups contribute solubility, helping the heme to remain in aqueous solvent and not

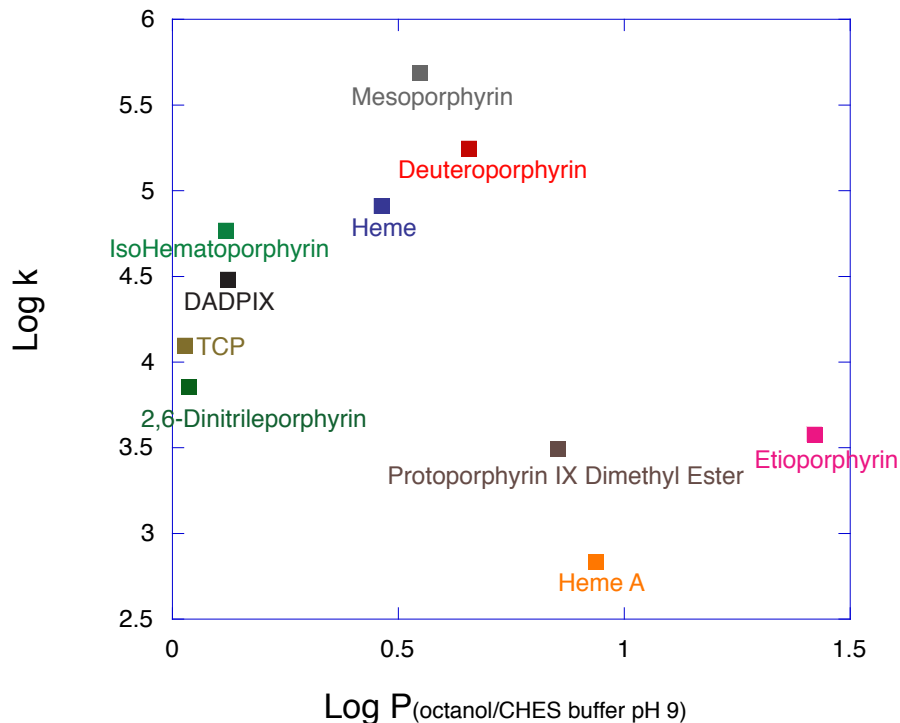


Figure 4.3: Log of rate constant (k) as a function of the partition coefficient (\log - P) in 20mM CHES 150mM KCl buffer pH 9. Rates were taken from the previous graph. Log- P values were determined between CHES buffer and octanol at room temperature, and allowed to equilibrate for approximately one hour. With Mesoporphyrin appearing to bind the fastest, becoming more hydrophobic or hydrophilic appears to slow the assembly rate.

aggregate. This hypothesis helps to explain why Mesoporphyrin binds with the highest rate. It has an amphiphilic character leading to a low propensity to aggregate and ease of burial in the protein core. This correlation between rate and solubility can be

used as a rough guide for porphyrin selection as this relationship appears to play some factor in the binding mechanism.

In agreement with the kinetics, the thermodynamics indicate a need for a clear separation of hydrophilic and hydrophobic faces pH. This work will be detailed in Kodali

et. al. (in preparation), showing that tetrapyrroles with the highest affinities have a clear hydrophobic face and a clear hydrophilic face.

While the plots show a correlation between log-P values and binding rate it may

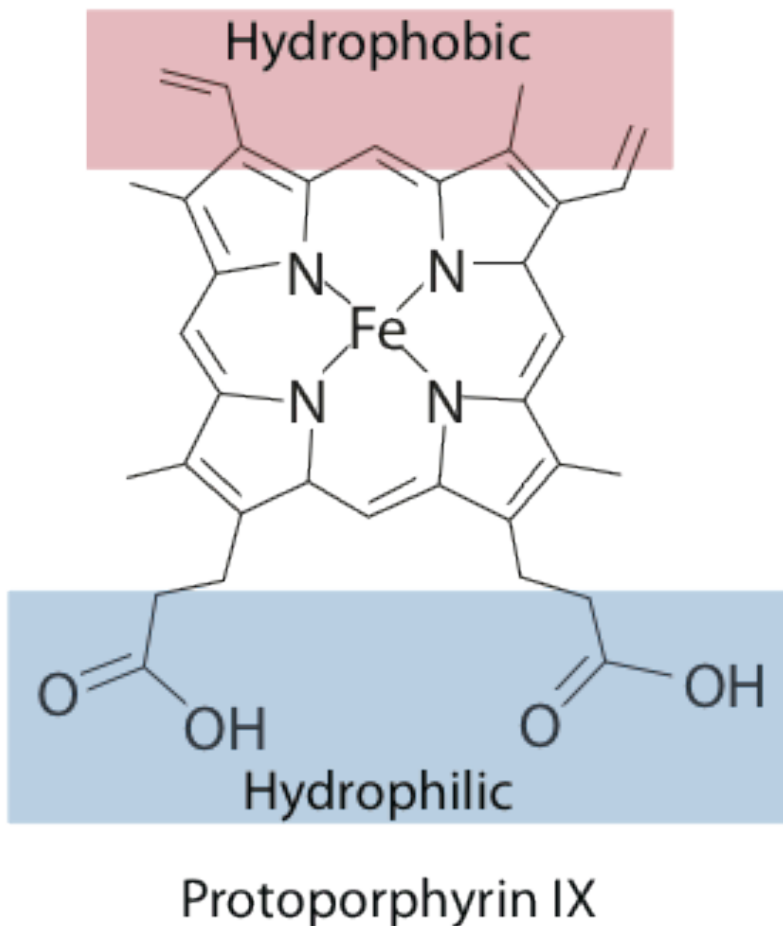


Figure 4.4: Illustrating the amphiphilic hypothesis in heme. The pink shaded area denotes the hydrophobic face while the blue shaded area denotes the hydrophilic face. It is essential that a porphyrin has both to maximize its solubility with interactions with the protein.

not be as straightforward.

In addition to vast differences in solubility, these porphyrins also have a vast difference in size.

Etioporphyrin and heme A are both slow binding porphyrins. Though both have different log-P values

they also have vastly different sizes of their ring substituents. Heme A has a

large bulky farnesyl tail, and Etioporphyrin has, in

that same C8 position, a far less bulky ethyl group.

Despite both being unable

to interact with water, they have different interactions with the protein. The bulk of the farnesyl tail hinders the heme from getting buried in the core, both because it is insoluble and because a large portion of the protein interior must conform to a bulky substituent. This is an obvious hindrance to this process, one not encountered with the smaller substituents

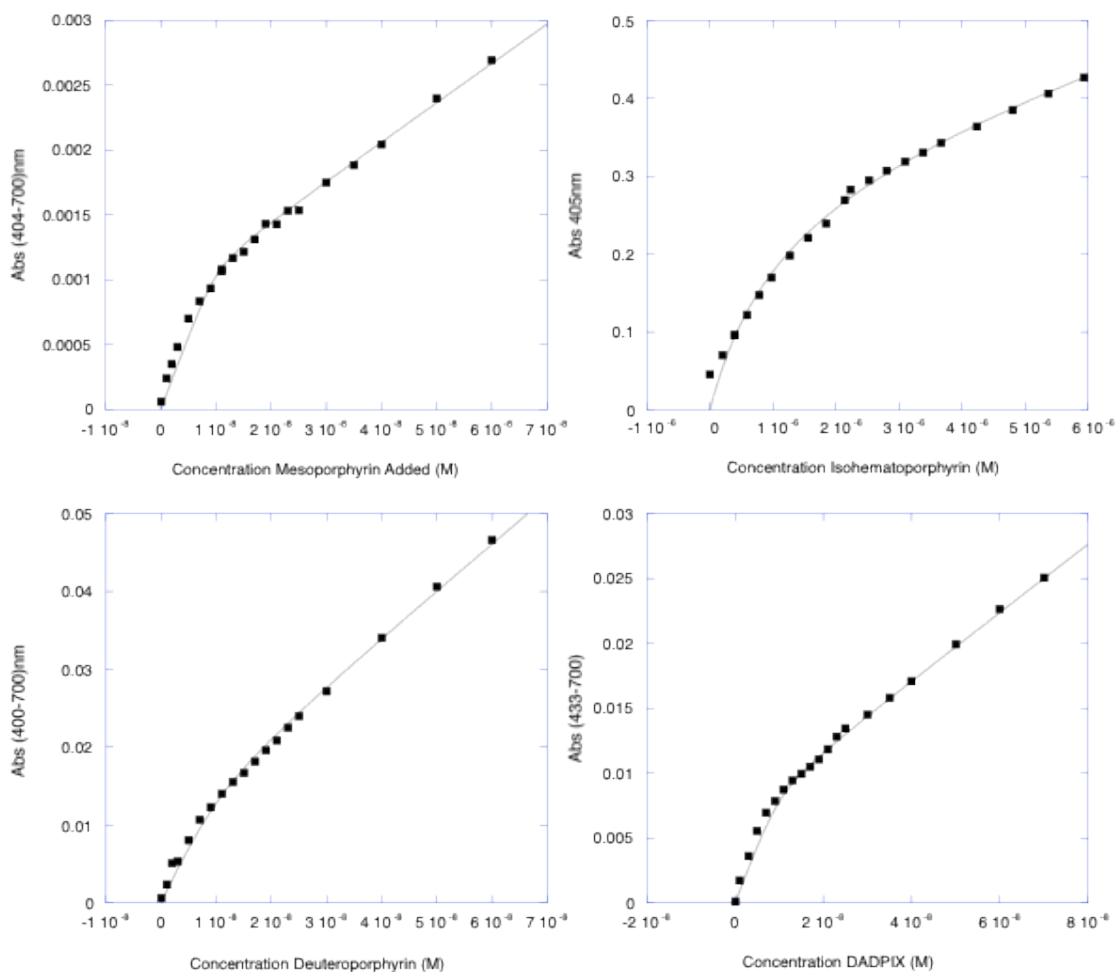


Figure 4.5: Affinity titrations of a selection of porphyrins to the single-chain protein at 10nM. The Iso Hematoporphyrin trace was done with 2.9 μ M protein. All titrations were done in 20mM CHES 150mM KCl buffer pH 9 at room temperature.

on Etioporphyrin. The data here do not provide enough information to make clear whether solubility or size is the main barrier,

Fe-TCP is another example of size being a possible hindrance to assembly. This cofactor has large carboxy-phenyl groups at its four meso-positions, and burying these in the hydrophobic core incites two problems. First, there is a large amount of steric bulk,

requiring the protein to attain a new structure before binding, similar to heme A. The second is the burial of charged groups. Sequestering these in a core devoid of water would require a large amount of energy. This energy would manifest in a kinetic barrier slowing assembly.

In an attempt to determine a relationship between the porphyrin affinity by protein and the rate the K_d values were measured for the soluble porphyrins. Figure 4.5 depicts the

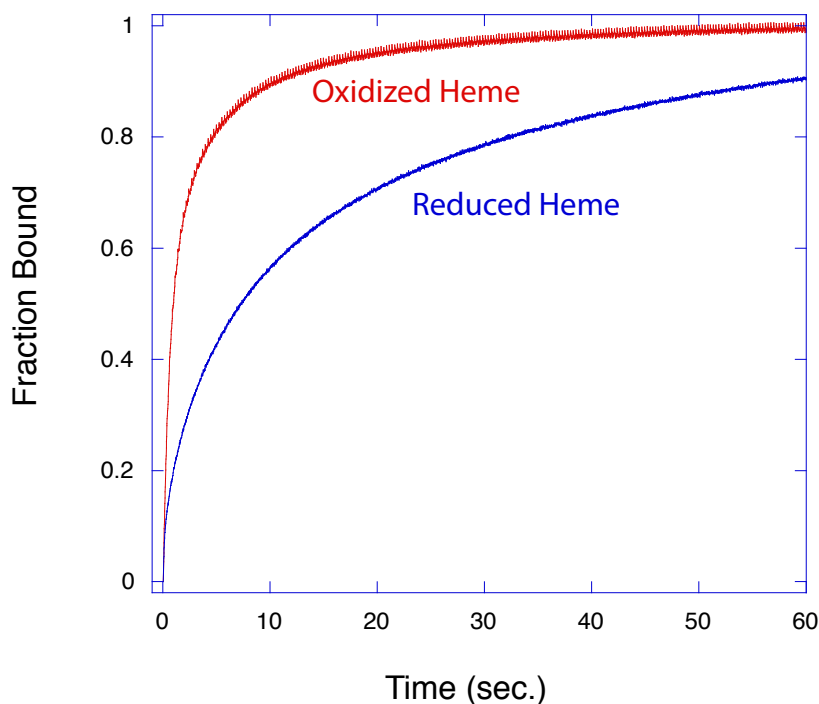


Figure 4.6: Comparison of ferric (red) and ferrous (blue) heme binding to the tethered homodimer protein. Protein at $10\mu\text{M}$ and porphyrin at $5\mu\text{M}$. Data collected at 25°C in 250mM Borate 150mM KCl buffer pH 9. Oxidized heme data is the change in 412nm over time, and Reduced heme is change in 427nm over time.

redox states. To test this effect, ferric and ferrous heme were rapidly mixed with the tethered homodimer protein and the Soret band was monitored over time (Figure 4.6). In this experiment assembled proteins were monitored by the shift in absorbance at the Soret

affinities for the four porphyrins Deuteroporphyrin, Mesoporphyrin, DADPIX, and Iso Hematoporphyrin. The K_d values are listed in table 4.1. From the data here, no discernible connection was made between the affinity and rate.

I next wanted to test any effect that resulted from a change in

band (Figure A3.1). The ferrous heme assembles with protein at a slower rate than the ferric state. This can be explained using the idea of solubility from above. The oxidized porphyrin's Fe-center has a +1 charge due to the Ferric iron being in a +3 state and the porphyrin ring having a -2 charge. When this Fe is reduced to a plus +2 state the net charge of the tetrapyrrole ring becomes neutral and the compound would have a higher propensity to aggregate in water causing the rate to decrease. This is not the only explanation, it is also possible interactions with ligands are affected. The electronic environment of heme in the

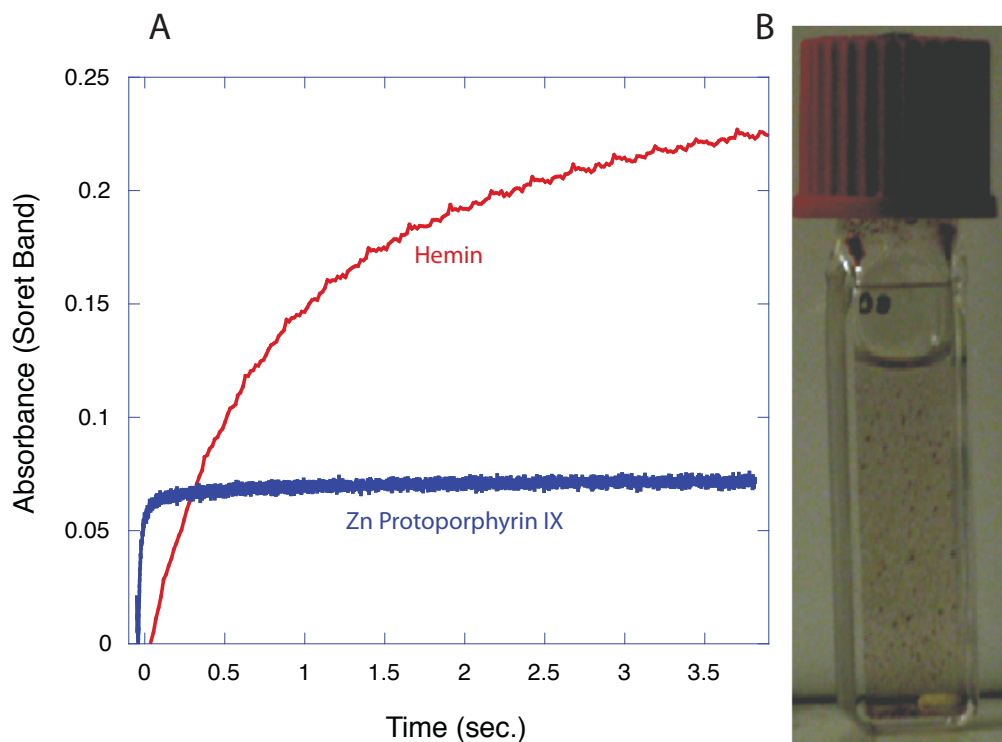


Figure 4.7: Comparison of ZnPPiX and heme assembling with the tethered homodimer protein (A). Over the course of 4 seconds Zn appears to bind quickly compared to the tethered homodimer. However the stoichiometry is not consistent. This can be explained by looking at the stock of ZnPPiX (B). It is in an aggregated state and precipitating out of solution. All data was collected at 20°C in 250mM Borate 150mM KCl pH 9.

oxidized state may bind more tightly due to coulombic interactions having a higher affinity for the lone pairs of histidine.

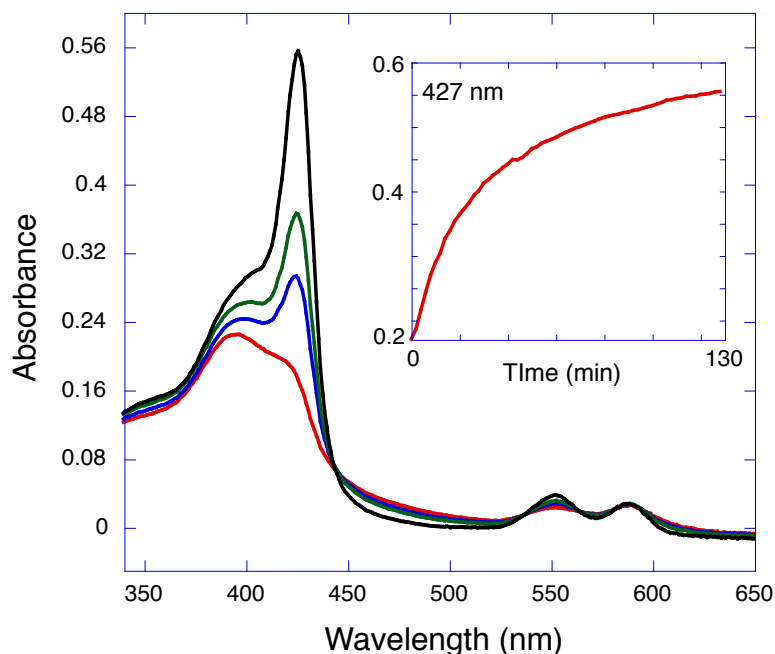


Figure 4.8: ZnPPIX aggregate (prepared by leaving ZnPPIX in buffer overnight) being broken up and assembling with the single chain protein. Protein at $5\mu\text{M}$ and porphyrin at an unknown quantity due it being in an aggregated state. Inset depicts the change in the Soret band (427nm) as a function of time. Experiment performed at room temperature in 20mM CHES 150mM KCl buffer pH 9.

4.2.2: Effect Of The Metal On Assembly

Having looked at different Fe-centered porphyrins, I next investigated the effect of the metal center to determine if it has an effect on the rate. I mixed the tethered-homodimer protein with heme and Zn protoporphyrin IX (ZnPPIX). These two porphyrins differ only by the metal at the center. In this experiment

assembled proteins were monitored by the shift in absorbance at the Soret band (Figure 4.8). Figure 4.7-A shows the traces of these two porphyrins over four seconds. Though the Zn appears to assemble faster than the Fe, the final stoichiometry is much lower. This is due to the aggregation of ZnPPIX that can be plainly seen in Figure 4.7-B. In this case, the solubility of ZnPPIX is a clear barrier to the assembly rate, significantly slowing it compared to iron. Based on this data, the metal appears to have a substantial effect, it

being the only difference between these two porphyrins. Heme has been shown to aggregate in solution slightly however, not to the degree of ZnPPIX.⁴

4.2.3: Effect Of The Protein On Porphyrin Aggregates

Despite what is shown in figure 4.7 it is possible to achieve assembly and rescue a cofactor from an aggregated state, however the rate in this scenario is greatly slowed. Figure 4.8 shows a protein removing ZnPPIX from an aggregate state. In this experiment, protein was added to a pre-aggregated sample of ZnPPIX and the absorbance at 427nm was monitored, the Soret band of protein-bound ZnPPIX. Over the course of 1.5 hours, binding appears to be limited by the de-aggregation rate yet again. The protein appears capable of a deaggregase activity breaking up aggregates of porphyrin and sequestering it.

4.2.4: Effect Of Porphyrin Skeleton On Assembly

The other aspect of the tetrapyrrole structure to be investigated for rate differences is the ring. To

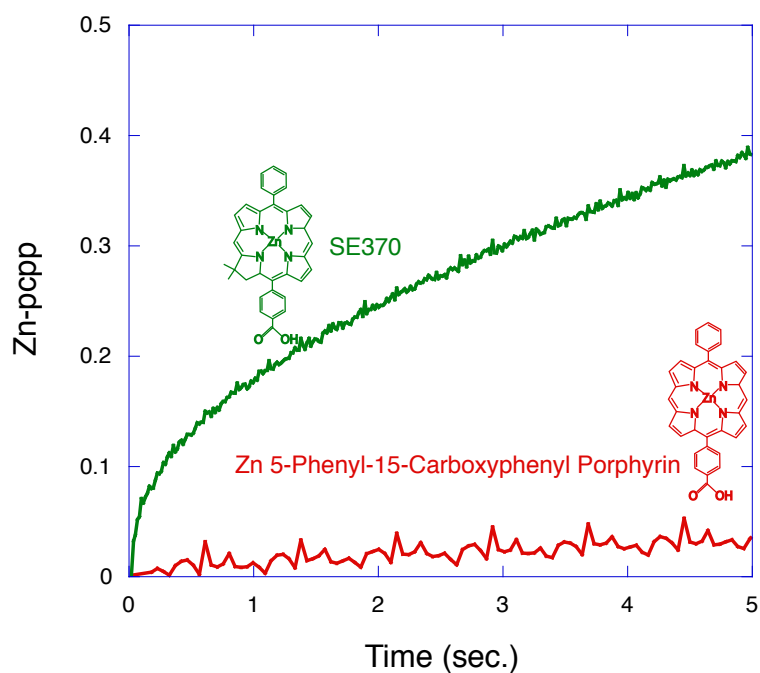


Figure 4.9: Comparison of the assembly rates between a similarly structured Zn centered chlorin (green) and porphyrin (red). The saturation of a pyrrol ring appears to increase the rate despite substituents and metal being the same between the two. Data was collected at 20°C in 20mM CHES 150mM KCl buffer pH 9.

test this we used a pair of synthesized tetrapyrrole cofactors SE370, a chlorin type, and Zn-5-phenyl-15-carboxyphenyl Porphyrin (ZnPCPP), given to us by the Jonathan Lindsey and Sergei Vinogradov labs respectively. These two cofactors differ in structure by the saturation of a single pyrrole ring and the addition of methyl groups (Figure 4.9). In this experiment assembled proteins were monitored by the shift in absorbance at the Soret band (Figure A3.19). The change in the ring structure has an obvious effect; the rate increasing when it is saturated. It is likely that any differences stem from the methyl groups or puckering of the tetrapyrrole ring breaking up aggregates or self-assembled oligomers in solution, providing monomeric chlorins for the protein to assemble with.

4.2.5: Test Of Solubility With A Synthetic Porphyrin

The Vinogradov group synthesized a new porphyrin taking into account both the amphiphilic requirements that allowed me to test the amphiphilic theory directly with a Zn-centered porphyrin. This porphyrin is the ZnPCPP modified to have a newkome dendrimer (containing 3 carboxylic acid groups) for increased solubility. The binding rate for this was compared to ZnPCPP from before. Figure 4.10 shows the comparison between these two rates. In this experiment assembled proteins were monitored by the shift in absorbance at the Soret band (Figure A3.19). The newkome-porphyrin binds to the protein within milliseconds of being exposed in a stopped flow, and majorly in the dead time of the instrument. As such no rate was fitted for this porphyrin. This synthesized porphyrin has a zinc center, as it's primary design goal was to end up as a light harvesting and electron transfer porphyrin.⁶ It is compared to a porphyrin that is structurally the same except that it does not have the newkome dendrimer. This porphyrin does not appear to bind at a comparable rate; its lack of a solubilizing group hindering this reaction. No rate

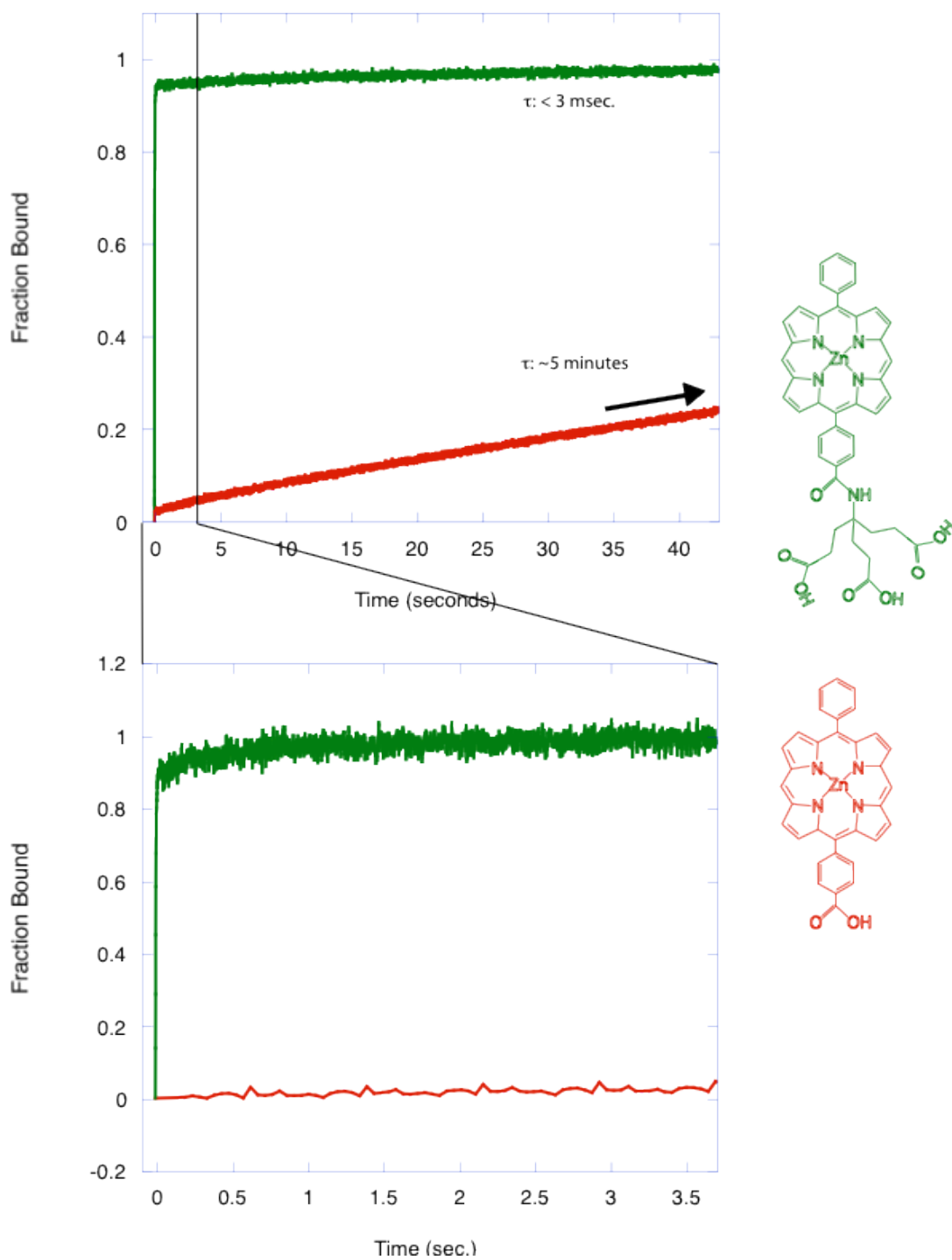


Figure 4.10: Assembly kinetics of an insoluble porphyrin (red), and the same one but with a chemical modification aimed at solublizing it (green). The soluble group, called a newkome dendrimer, greatly increases the rate of assembly. All rates collected in 20mM CHES 150mM KCl buffer pH 9 at 25°C. Protein at 24 μ M and porphyrin at 5 μ M.

was fitted for this, as it appears to not bind. Though no rates could be determined, this clearly shows the effect of increasing the solubility of a porphyrin on the assembly rate. Increasing from not binding at all to binding within the dead-time of the machine.

4.3: Discussion

4.3.1: Porphyrin Control Over The Rate Of Assembly

The data above describe the obstacles involved with selecting a tetrapyrrole to assemble with a 4-helix bundle protein. More generally this data describes the kinetics of assembly from the perspective of the cofactor, any tetrapyrrole-binding protein could be affected by these factors.

Figure 4.3 highlights a major hindrance to cofactor binding. Increasing the cofactor-binding rate depends heavily on increasing the cofactor's amphiphilicity, measured here by partition coefficients. The best example of this is observed in the differences in assembly rate between heme B and Fe-PPIX-DME incorporation. The only difference between these two cofactors is the methylation of the propionates, the overall structure is otherwise unperturbed, but simply removing the negative charges on the cofactor eliminates its solubility. The partition coefficient of heme B (0.46 Octanol/pH 9 aqueous buffer) increases drastically (0.85 Octanol/pH 9) when in a PPIX-DME form, and the rate is similarly affected, decreasing from $8.2 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ down to $3.1 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$.

Heme can similarly be compared to the 2,6-Dinitrileporphyrin. These two cofactors have the same structure save for the vinyl groups at C3 and C8 being converted to nitrile groups. The nitrile groups lower the log-P value from 0.46 to 0.036 and the rate from $8.2 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ to $7.1 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$, an order of magnitude change in rate. In this case, the increased solubility hinders the rate by requiring polar groups to be buried. The structure of

the porphyrin is not significantly changed but the assembly is greatly affected. Together these two examples point to the log-P value as being highly correlated with the assembly rate. One can think of these as a guide to cofactor selection and a method of diagnosing potential barriers to binding.

Solubility, while appearing to play a role in this process, is not necessarily the main factor. Large cofactors, like heme A and Fe-TCP, are likely limited by substituent size as well. Bulky sidechains would hinder assembly as the protein must adopt a conformation that can accommodate these large chemical groups. This is in contrast to the small groups on Mesoporphyrin or Deuteroporphyrin consisting of ethyl and methyl groups respectively. These do not require significant conformational changes, the holoprotein can easily assemble.

Changes to the tetrapyrrole skeleton have noticeable effects on the assembly rate. Chlorins have a saturated pyrrole ring as a signature part of their structure. This greatly changes their photo-physical and other inherent properties. In this context, it may speed up their assembly reaction with the protein by destabilizing their aggregated state compared to a porphyrin with similar substituents.

4.3.2: Natural Relevance

Nature has a variety of protein systems designed to control the assembly of holo-proteins and circumvent issues with cofactor aggregation. In many cases, tetrapyrroles are stored in proteins, such as HasA, which stores heme B.⁷ There are cytotoxic consequences brought on by free cofactor; it is known to aggregate,⁸ can generate an insoluble hemozoin crystal,⁹ and also has been shown to catalyze Fenton chemistry releasing reactive oxygen species (ROS).¹⁰ Aggregated heme also induces cellular pro-inflammatory-signaling

pathways. Aggregates of bacteriochlorophylls do not necessarily trigger Fenton chemistry because they lack iron, but their photophysical properties are changed substantially from the free cofactor, hindering them from taking part in light capture reactions.¹¹ In either case, loss of a prosthetic group is a major issue to protein function and cell viability.

Maquettes *in vitro* do not have proteins to limit aggregates, and a change in solubility can impede assembly. However, maquettes do have an advantage over natural proteins in this regard. Their ability to bind both natural-modified cofactors and synthetic tetrapyrroles allows for changes in solubility to be made with little to no effect on the affinity or electrochemical properties.

4.3.3: Conclusions

The work in this chapter provides a rudimentary guide to tetrapyrrole cofactor selection when dealing with proteins. Though the work here uses exclusively 4-helix bundles proteins, the concepts are equally applicable as they pertain to the cofactor. The protein was chosen as it is the least likely to have a rate limiting step that would be slower than that of the de-aggregation or solubility issues here. The rate-limiting step regarding the cofactor in assembly does not change.

4.4: References

- (1) Gibney, B. R.; Isogai, Y.; Rabanal, F.; Reddy, K. S.; Grosset, A. M.; Moser, C. C.; Dutton, P. L. *Biochemistry-Us* **2000**, *39*, 11041.
- (2) Gibson, Q. H.; Antonini, E. *Biochem J* **1960**, *77*, 328.
- (3) Tamura, M.; Woodrow, G. V., 3rd; Yonetani, T. *Biochim Biophys Acta* **1973**, *317*, 34.
- (4) Kuzelova, K.; Mrhalova, M.; Hrkal, Z. *Bba-Gen Subjects* **1997**, *1336*, 497.
- (5) Albert Leo, C. H., David Elkins *Chem Rev* **1971**, *71*, 525.
- (6) Mohanty, S. K.; Subuddhi, U.; Baskaran, S.; Mishra, A. K. *Photochem Photobiol Sci* **2007**, *6*, 1164.
- (7) Yukl, E. T.; Jepkorir, G.; Alontaga, A. Y.; Pautsch, L.; Rodriguez, J. C.; Rivera, M.; Moenne-Loccoz, P. *Biochemistry-Us* **2010**, *49*, 6646.

- (8) Asher, C.; de Villiers, K. A.; Egan, T. J. *Inorg Chem* **2009**, *48*, 7994.
- (9) Weissbuch, I.; Leiserowitz, L. *Chem Rev* **2008**, *108*, 4899.
- (10) Prousek, J. *Pure and Applied Chemistry* **2007**, *79*, 2325.
- (11) de Paula, J. C.; Robblee, J. H.; Pasternack, R. F. *Biophys J* **1995**, *68*, 335.

Chapter 5: Midpoint Potential Control In A Series Of Maquettes

5.1: Introduction

In natural systems many oxidoreductase functions are carried out with a heme B cofactor. These range from electron transfer to oxygen binding and alkane hydroxylation.¹⁻³ One major aspect of the heme that allows it to have such diverse roles is its midpoint potential (E_m).⁴ This thermodynamic parameter is a measure of how reducing the heme is, and its control is crucial to the protein properly achieving its intended function.

For many years maquettes have been used to study how the protein exerts control over this parameter. Brian Gibney and Julia Shifman studied a previous generation of maquettes related to the unstructured homotetramer protein of chapter 3. In the H10A24 protein they were able to increase the heme B midpoint potential from approximately -200mV up to -100mV using environmental factors, and charge effects.⁵⁻⁸ This 2.3 kcal/mol change is impressive but still far from the range seen in natural systems, which can reach upward of +400mV or 9.22 kcal/mol.⁵

In this chapter, I study how holoproteins modulate the midpoint potential in a new generation of designed proteins, more malleable and asymmetric than the ones used previously, a consequence of them being monomeric.⁹ This asymmetry allows for the creation of separate environments within the same protein and more control over ligation states. Through well known potentiometric methods,⁴ I was able to show how various changes to the protein and heme affect the E_m , shedding light on this process in natural systems.

From a design perspective, this chapter demonstrates the most basic principles underlying construction of a synthetic oxidoreductase protein. From selection of cofactor to

placement and midpoint potential, the principles in this paper can be used to tailor maquettes to be able to perform any redox function.

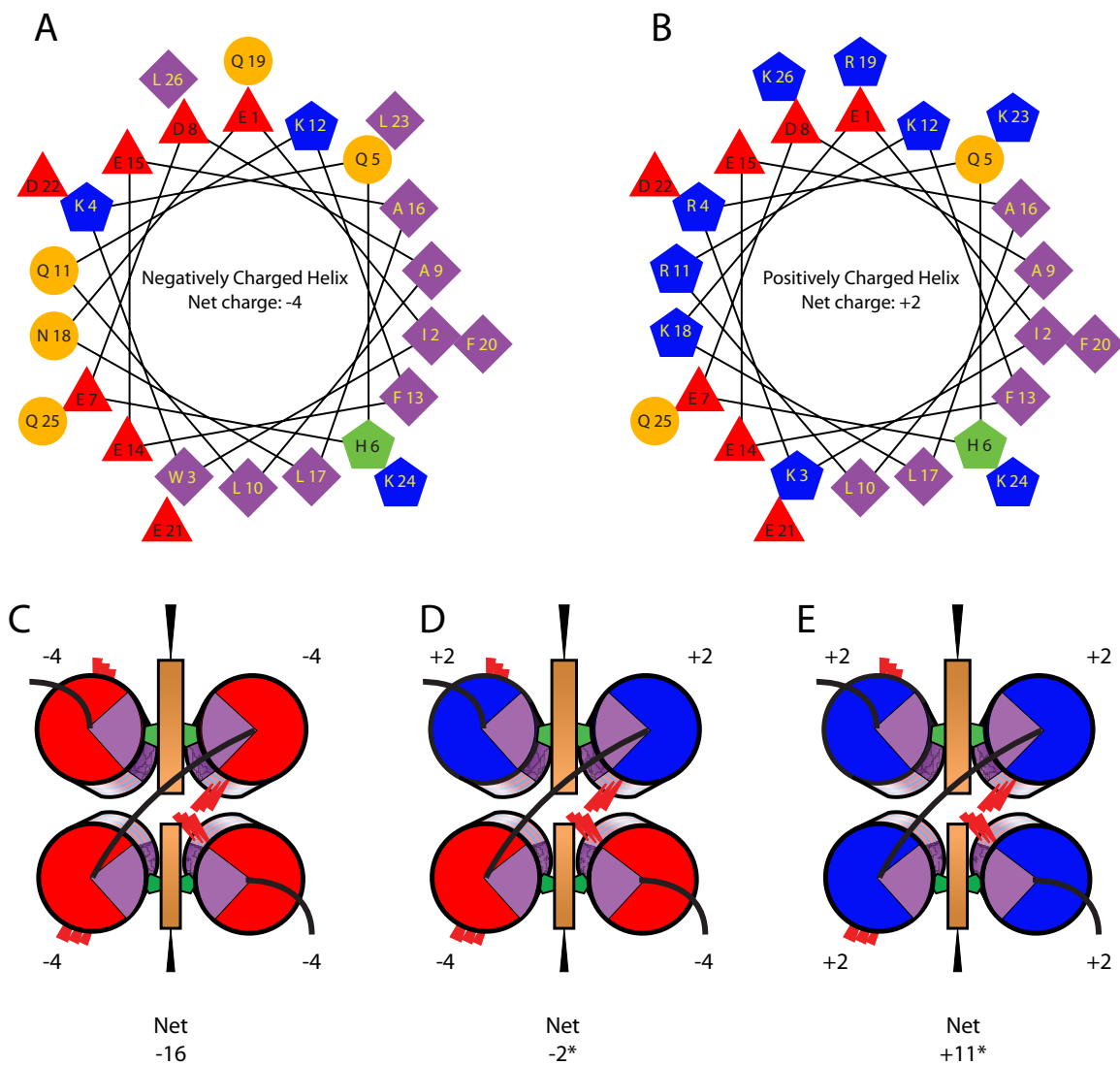


Figure 5.1: Helical wheels of the single chain protein (A), and a positive variant of that design (B). Heme binding His residues are colored green, positively charged residues are colored blue, negatively charged residues are red, polar-uncharged residues are in yellow, and hydrophobic residues are in purple. These helices can be connected in various ways as shown below: All negative (C), mixed half positive and half negative (D), and all positive (E). The various combinations of helices afford the protein different sets of attributes that relate to the charge imposed, the hydrogen bonds formed and the burial of certain amino acids.

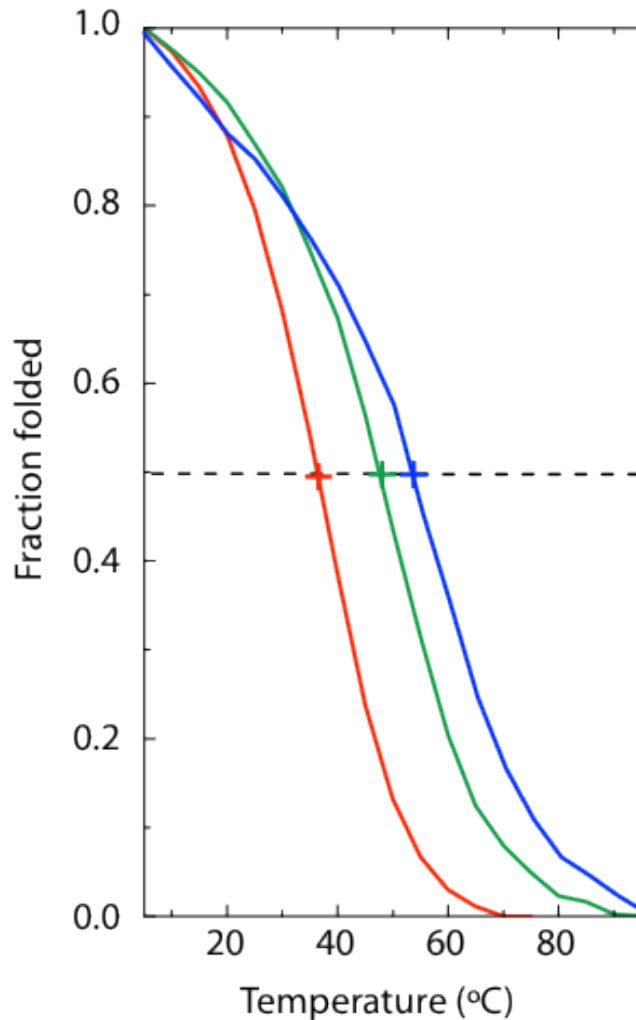


Figure 5.2: Melts of apo-state of the single-chain (red), the binary single chain (green), and the positive mutant (blue). Using helices from figure 5.1, three distinct designs were generated, all with varying T_m values, Single chain: 37°C, Binary: 48°C, and Positive: 55°C. These different T_m values are brought on by differences in charge patterning along the surface of the protein generating coulombic pairing between amino acids. Data for the single chain and binary were collected by Goutham Kodali. All three melts were performed in 20mM CHES 150mM buffer pH 9. All protein concentrations were approximately 5 μ M

5.2: Results

5.2.1: Environmental Changes

The first set of changes aimed at effecting the midpoint potential is altering the general charge patterning of surface residues. Figure 5.1 (Panels A and B) shows a comparison of two individual helices that have different net charges along the outside of the helix. These have been combined to form three 4- α -helix bundles with different charge properties (Figure 5.1 C-E). Starting with the single chain protein from previous chapters (Figure 5.1-C), whose net charge is -16; the charge pattern along the outside was altered such that the net charges become -2 (Figure 5.1-D) and +11 (Figure 5.1-E).

To test for the any change in

stability, circular dichroism (CD) temperature melts were performed (Figure 5.2). The positive mutant, thanks to an increase in charge pairing along the helical surface is rendered more thermostable by 0.6 kcal/mol (a T_m increase from 37°C to 55°C). The binary mutant has a T_m of 48°C, which is between the positive and negative charged variants, as expected. It is unclear from these data whether or not there are multiple transitions overlaid. NMR could not be done on the positive mutant due to aggregation issues above 30µM.

The K_d value was also examined to ensure heme affinity would not complicate measurements. Figure 5.3 depicts the binding titration of the positive mutant compared to the single chain variant from before. There is a noticeable increase in this parameter, going

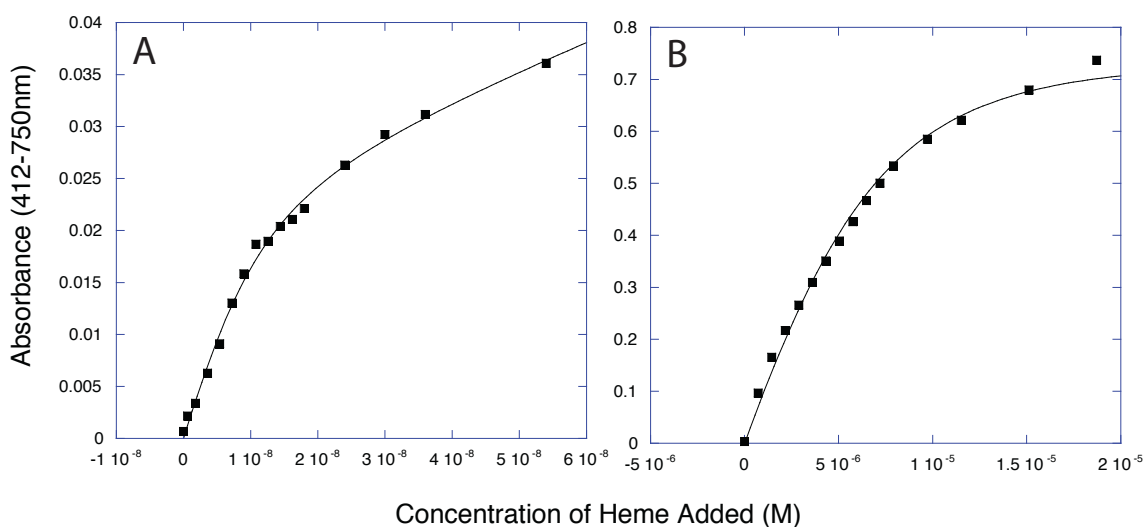


Figure 5.3: Comparison of the single chain protein (A) and the positive variant (B) binding affinities. Protein was at 6nM in (A) and 5µM in (B). The addition of positive charges causes a severe lowering of the affinity, driving it up from sub-nanomolar to micromolar. Titrations were done in 20mM CHES 150mM KCl buffer pH 9 at room temperature. The concentration of heme (in Molar) is depicted on the x-axis. Note the difference in order of magnitude between the two proteins

from sub-nanomolar up to 1.5 μ M. Though this sequence change resulted in a loss of affinity it was not severe enough to prohibit further studies.

This charge patterning has a significant effect altering the midpoint potential. In the single chain variant the E_m value is -290mV, which starts with a charge of -16. In the binary mutant this E_m is split, accounting for the multiple environments, to -224mV and -151mV. In the positive variant, the split is once again removed and the higher

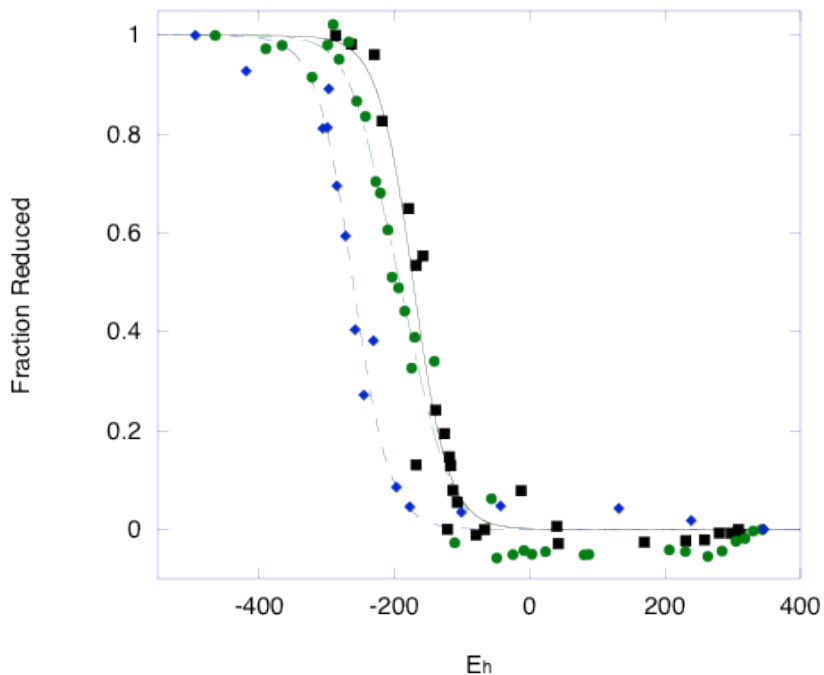


Figure 5.4: Altering the charges along the outside of the protein can allow for a significant change in the midpoint potential of a heme B cofactor. Blue Diamonds: -16 charged single chain protein with one heme, midpoint potential of -260mV. Black Squares: +11 charged single chain protein with two heme cofactors, both potentials: -151mV. Green Circles: Binary charged single chain maquette with two heme cofactors. This shows two potentials at -151mV and -225mV, a substantial shift. Titrations were done at room temperature in Phosphate buffer pH 8. All proteins at 25 μ M. Mediators are listed in the Materials and Methods section. Titrations carried out at 22 $^{\circ}$ C in 100mM PO₄ 150mM KCl pH 9

potential remains, -151mV (Figure 5.4). The positive charged and negative charged proteins were fit to a single 1- E_m curve, while the binary mutant was fit to a 2- E_m curve (see materials and methods). This is a 3.22kcal/mol difference between the -16 and +11 charge patterns of the single chain, and positive variant maquette (140mV E_m difference). This large

split between the two comes to approximately 5mV/charge, a value that does not match 11-14mV/charge seen in the literature.¹⁰ However, the binary mutant has two different midpoint potential values, and based on the sequence similarity, it is reasonable to assume that the two positive helices in the binary mutant contain the heme with a -151mV E_m value. In this case, half of one protein is modified, the change in charge is only 12, which gives an 11.6 mV/charge change (0.26 kcal/mol-charge), well in line with previous studies.

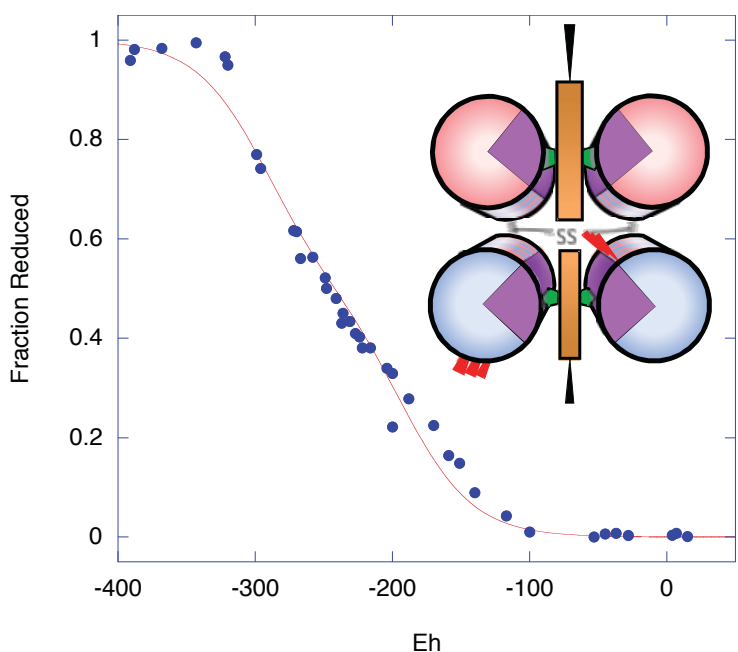


Figure 5.5: Redox Titration of the F56HA3 protein. The independence of helices allows for a modular control over the midpoint potential as well as functions. One helix-pair has had glutamates replaced with alanines causing a 100mV midpoint potential split (-290mV and -190mV), see sequence appendix. Mediators are listed in the Materials and Methods section. Titrations carried out at 22°C in 20mM CHES 150mM KCl pH 9. Holoprotein at 25µM.

There also appears to be a crossover effect from the positive helices raising the E_m of the heme bound to the negative helices from -290mV up to -224mV ($\Delta E_m = 66\text{mV}$, 1.5 kcal/mol). This is most likely due to buried lysine residues in the core mildly destabilizing the nearby positive ferric heme B.

This experiment clearly shows the maquettes ability to create E_m differentiation between separate heme sites. This site

independence can be seen with another example of charge patterning. The homo-dimeric protein F56HA3 has two domains, one of which has glutamates buried near the heme and

the other has alanines in place of them (Sequence Appendix). The effect of this differences is observed through a midpoint potential shift from -290mV to -190mV, a difference in the two E_m values is 100mV

(Figure 5.5), and is similar to that seen in the H10A24 protein.⁵ This difference in potential arises from placing negatively charged Glu residues near the heme in one domain. This stabilizes the ferric heme by 2.3kcal/mol through columbic interactions. With these residues removed the charge stabilization is also removed, leading to an

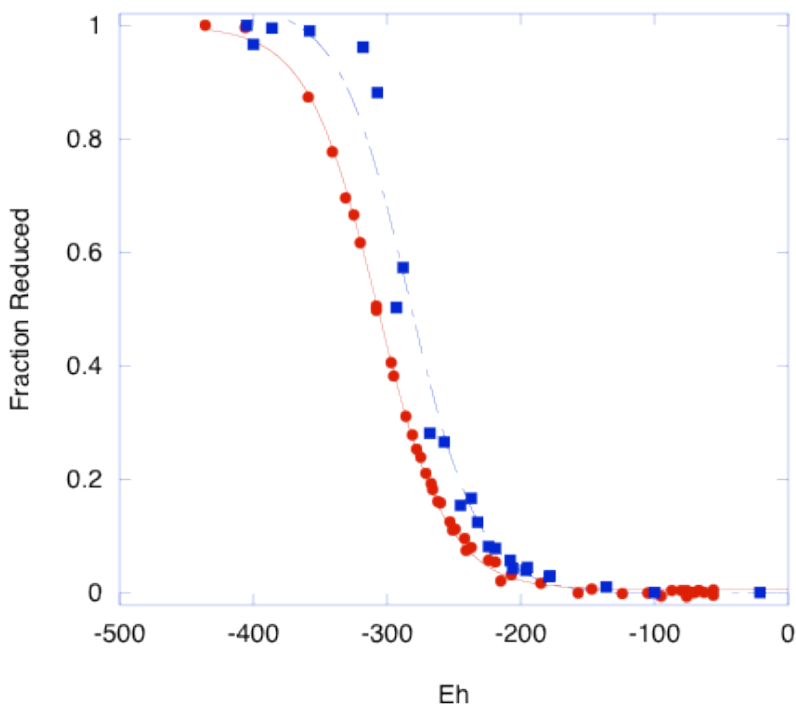


Figure 5.6: Redox titrations of the mutants E8A (red) and E57A (blue). Each mutant shows similar E_m values of -283 and -307 respectively. Further proof that shifting the midpoint of heme in maquettes requires more significant mutations. Mediators are listed in the Materials and Methods section. Titrations carried out at 22°C in 20mM CHES 150mM KCl pH 9

increase in the midpoint potential. Here, the curve was fit to a 2- E_m Nernst curve fit (see materials and methods).

Changing the midpoint via charge patterning is not a trivial matter of placing charges near the heme. Farid et al have shown that making a single tryptophan to lysine mutation near the heme is not sufficient to shift the potential significantly.⁹ This change resulted in an approximately 10mV increase. Though in agreement with literature values,

this change is not outside of the error of the potentiometric method used. More significant charge modifications are needed to make an impact on the potential.

Further proof that solitary charge modifications are not sufficient to increase the potential was needed to confirm the observation of Farid et. al.. The proteins E8A and E57A were chosen (Sequence Appendix). Potentiometric analysis does not show a substantial E_m shift compared to the original tethered homodimer, all E_m values are approximately -280mV (Figure 5.6). The mutation of two glutamate residues is not enough to cause a shift comparable to that seen in the F56HA3 mutant. Here, the curve was fit to a 1- E_m Nernst curve fit (see materials and methods).

5.2.2: Ligation Changes

Different porphyrin ligations are common in natural proteins, and often shift the potential significantly.¹¹ Figure 5.7 demonstrates the change in E_m when a bis-his ligation is mutated to his-cys. This change causes the E_m to change from -260mV to -280mV. Though slight, this change also opens up the possibilities for other functions as the cys residue is used as a ligand in many P450 enzymes.^{1,12} Going one step further and changing the ligation to cys-ala has a more profound effect on the midpoint potential, raising it from -260 to approximately -165mV, (95mV difference). This ligation change also yields spectra that correspond to a 5-coordinate state typically seen in P450s (Figure 5.8). Here, the three curves were fit to a 2- E_m Nernst curve fit (see materials and methods).

Ligation changes are not as robust as environmental changes. While cysteine residues are tolerated, both with a distal his residue or without, other ligations tend to not have a comparable affinity. Methionine-histidine ligation pairs, like those seen in the natural protein b_{562} , do not bind heme in maquettes.¹³ This lack of affinity is also seen in

natural systems.

Changing the Met residue of b_{562} to a Histidine increases the affinity by at least 100 fold.¹¹

Single his binding sites also do not have the requisite affinity for these studies.

In maquettes, a single heme site is unable to bind heme with a measureable K_d value. Efforts to improve the binding site have also met with failure. Single

cys residue sites have been successful as these are thought to be aided by the thiolate group generating an ionic bond between the negatively charged sulfur and the positively charged heme-Fe to improve the affinity.

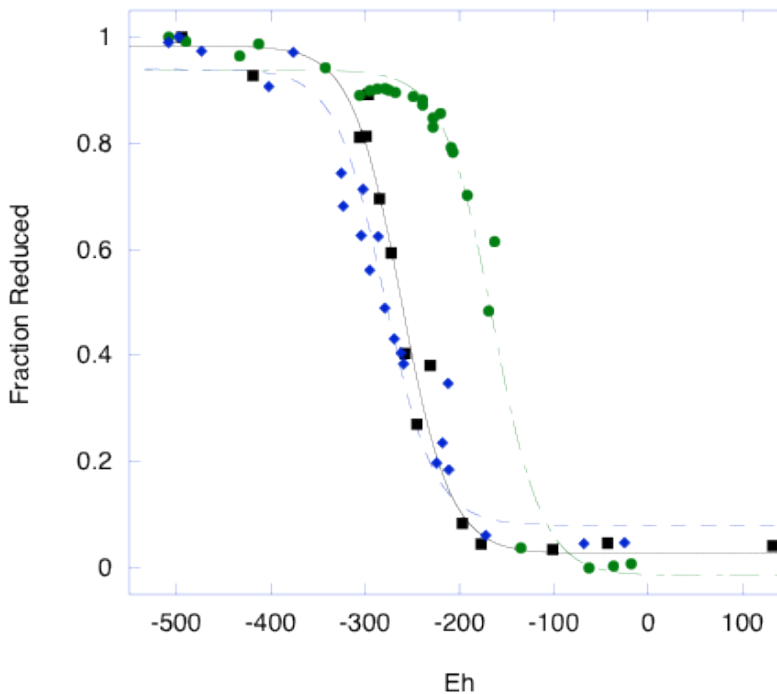


Figure 5.7: Redox titration of bis-his (black), his-cys (blue), and cys-ala (green) ligations in the single chain protein. All done in Phosphate buffer pH 8 at room temperature with 25 μ M protein. The his-cys ligation scheme does not have a substantial effect on the midpoint potential, lowering it from -260mV to -280mV. The cys-ala ligation however, has a large effect raising the E_m to -165, a 95mV increase. Mediators are listed in the Materials and Methods section. Titrations carried out at 22 $^{\circ}$ C in 100mM PO₄ 150mM KCl pH 9

5.2.3: Using Cofactors To Shift The Potential

In order to explore other methods to change the midpoint potential of heme inside a protein I moved another heme-cofactor to within van der Waals distance so that the two

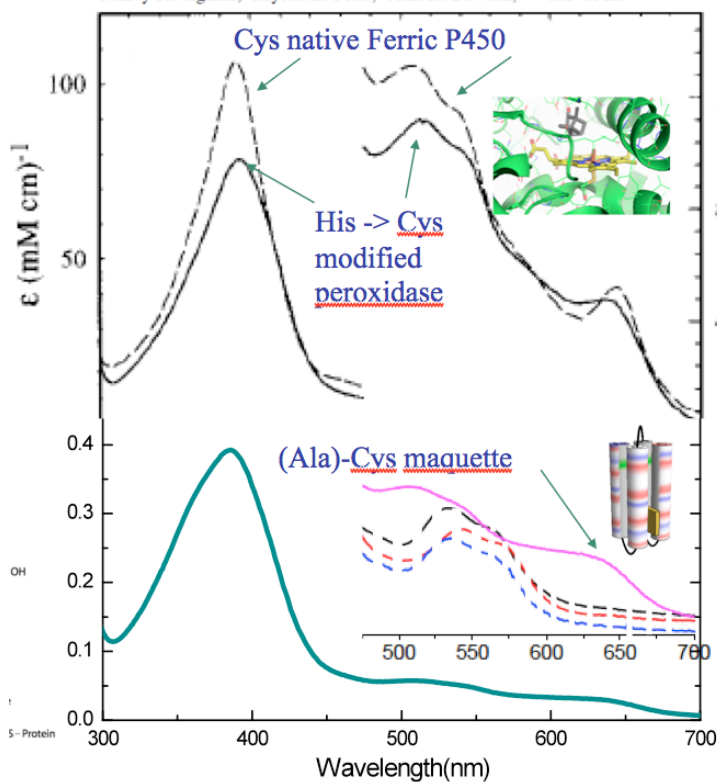


Figure 5.8: The Cys-Ala maquette bound to ferric heme shows UV/Vis spectra equivalent to ferric cytochrome P450 in the 5-coordinate state. Figure provided by Goutham Kodali.

heme-cofactors would be electronically coupled. This involved moving the ligating histidine amino acids up one heptad in the tethered homodimer protein. The first attempt to achieve two hemes in contact was to move the second heme adjacent with the first (Figure 5.9-Adjacent). This was successful when done in the H10H24 maquette, as seen in the work of Robertson et. al.^{7,14} The adjacent-tethered homodimer was not able to bind heme sufficiently as the decrease in conformational freedom does not provide enough room. To address this, the next variant moved the heme down one heptad such that the two cofactors were now "Offset" to one another (Figure 5.9-Offset). This comes at a cost of a higher K_d value, in the micromolar range (Figure 5.10). The weakened K_d is thought to be from crowding and steric clashes of 2 heme cofactors buried in the core. In previous variants such as H10A24 this is not a problem. The reason for this has to do with the way the heme binds and the protein's oligomeric state. H10A24 is a homodimer and the heme binds not between two separate dimers, but two helices of one dimer (Table 3.1). This alleviates crowding by allowing the two separate pieces to mildly dissociate from one

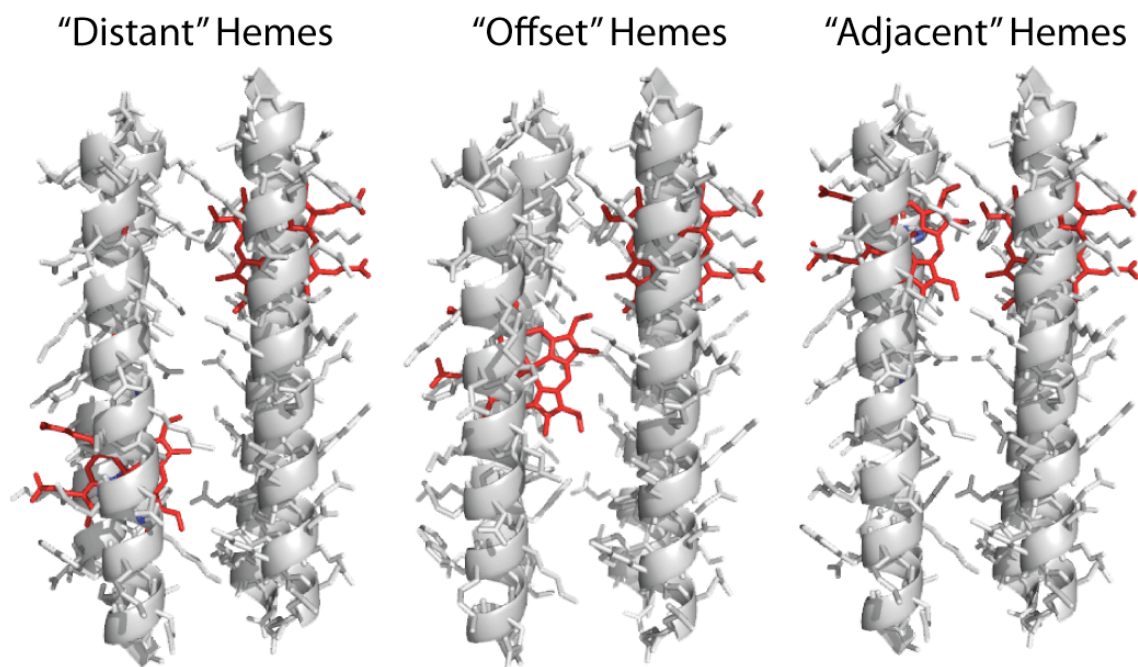


Figure 5.9: Pymol models of the three possible placements of heme in the tethered homodimer maquette. The original model had both heme cofactors at opposite ends of the protein (distant). Offset and Adjacent were attempted however, only the offset variant was able to accommodate two heme cofactors in one protein.

another generating more room to accommodate two cofactors. In the candelabra, as it is linked at the bottom, that is not an option. The lack of conformational freedom brought on by the disulfide bond restricts the binding sites. The offset protein also falls victim to crowding but not enough to eliminate binding. Attempts to recreate this in the single chain protein were fruitless. Due to the extra loop, this protein has even less conformational freedom at either end than the tethered homodimer and heme could not bind 2 heme cofactors in offset positions. Attempts were made to make more room in the binding site by removing bulky amino acids in the core, however, heme binding did not improve. Steric clashes brought on by the lack of conformational freedom appear to dominate when it comes to alternate heme placement.

Despite these various affinity issues, Figure 5.10 shows redox titration of the offset-candelabra mutant with E_m values of -280mV and -180mV, a split of 100mV. This effect

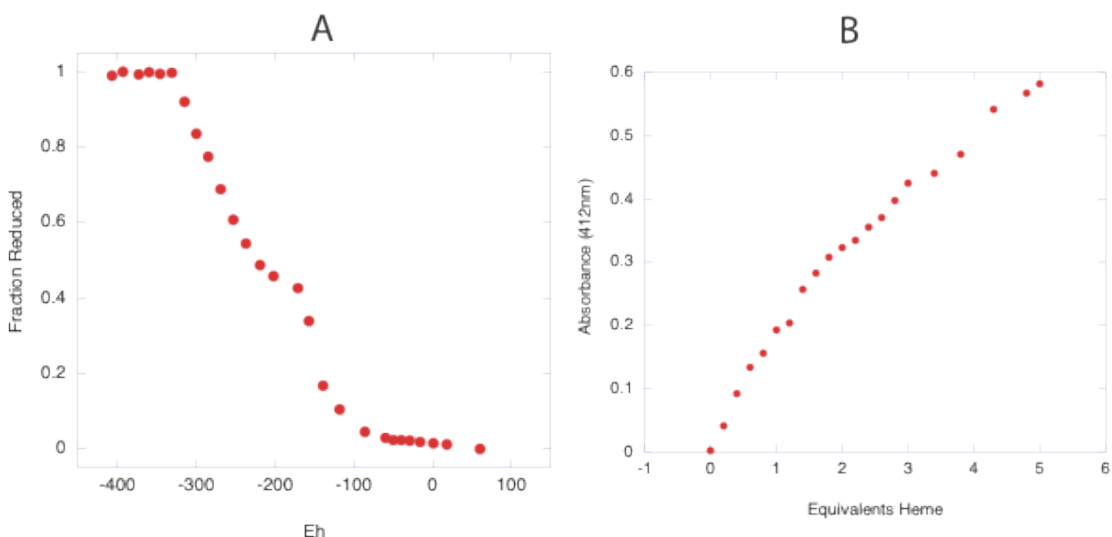


Figure 5.10: Redox titration (A) of the offset mutant with two heme cofactors bound. The proximity of the two cofactors causes a split in potentials of approximately 100mV (-290, and -180). Titration was performed at room temperature in 250mM Borate 150mM KCl buffer pH 9 with 25 μ M protein. Mediators are listed in the Materials and Methods section. (B), Kd titration of the Offset mutant. The proximity of the two heme cofactors, though possible, lowers the affinity to micromolar level. Titration done in Borate buffer pH 9 at room temperature.

stems from the positive charge on one heme affecting the other one nearby changing the ΔG of reduction by 2.3 kcal/mol. The affinity for the reduced heme is not as high as for the oxidized, and when the heme is reduced it dissociates from the protein, as such, only the first reductive direction is depicted. Here, the offset data was fit to a 2- E_m Nernst curve fit while the homodimer data was fit with a 1- E_m Nernst curve (see materials and methods).

5.2.4: Using Alternative Porphyrins

The other method investigated to change the midpoint potential of the heme was to simply change the heme itself to a different Fe-centered porphyrin. Here we used the array

of synthetic porphyrins from the previous chapter and obtained E_m values for the ones that bind sufficiently (Etioporphyrin and protoporphyrin IX dimethyl ester were not included due to difficulties in binding). Figure 5.11 show a large range of potentials spanning over 350mV, all with the single chain protein and similar K_d values (Table 4.1). Here, all of the porphyrin data were fit to $1-E_m$ Nernst curves (see materials and methods).

As shown above, one heme has a potential of -260mV in the single chain protein. Using the porphyrins I can expand the range of potentials in maquettes not only in the more oxidizing direction but also the more reducing direction (Table 4.1 Figure 5.11). Mesoporphyrin despite differing by a minor structural change has a 60mV shift downward. This saturation of the vinyl groups of heme appears to alter the electron density on the tetrapyrrole ring, stabilizing the Ferric (III) state. This is not the case with Deuteroporphyrin whose midpoint is -270. The structural difference is minor however the potential difference is more significant. It is possible that the conversion of the two ethyl groups to methyl, removing two CH_3 groups, allows more water near the heme changing the interactions with the protein core increasing the potential by solvating the porphyrin. Fe-Tetracarboxyphenyl porphyrin (TCP) has electron donating groups in the form of four carboxyphenyl substituents that drive the potential down 40mV.

The rest of this series of porphyrins change the potential in the positive direction (Table 4.1, Figure 5.11). IsoHematoporphyrin has a potential of -236mV. This is most likely caused by not the hydroxyl groups withdrawing electron density but rather them increasing the water concentration in the core. Hydroxyl groups are polar as they have hydrogen bond acceptors from the oxygen lone pairs and donors from the hydrogen. These could help stabilize water inside the core of the protein by bonding with them and allowing them to interact with the tetrapyrrole. Heme A has an E_m of -145mV most directly from an

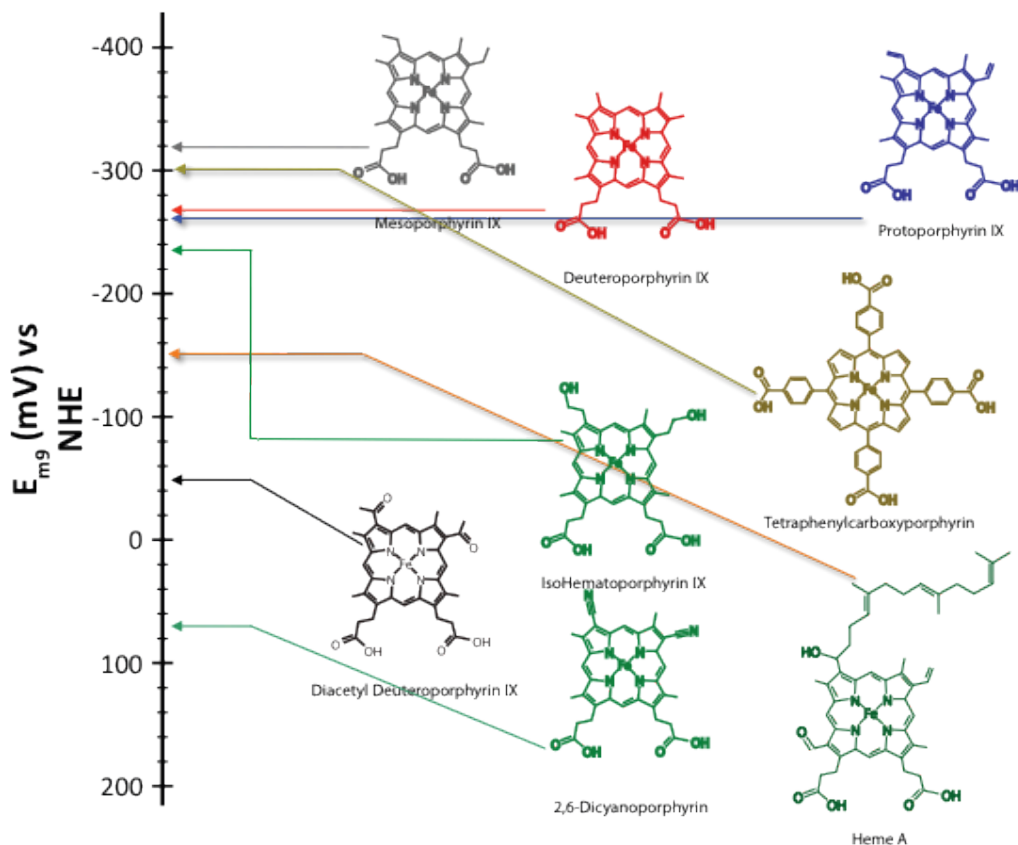


Figure 5.11: Reduction potentials of the Fe-porphyrins able to bind to the single chain protein. All titrations done at 25 μ M except heme A, which was done at an unknown concentration due to poor affinity. All titrations done in 20mM CHES 150mM KCl buffer pH 9 at room temperature. Mediators are listed in the Materials and Methods section. Data summarized in Table 4.1.

electronegative formyl group directly attached to the porphyrin ring. The addition of another electron withdrawing group, as seen in DADPIX, causes the porphyrin to have a potential of -48mV. This porphyrin has the two vinyl groups of heme B replaced with highly electronegative acetyl groups. Compared to one formyl group in heme A, the two here have a similar effect, raising the potential approximately 100mV/carbonyl group.

Past work from the lab performed by Sarah Chobot and Bruce Lichtenstein has added another porphyrin to the list.¹⁵ The 2,6-Dinitriroporphyrin E_m value was obtained in

the candelabra protein as opposed to the single chain, which is not thought to be a factor considering the similarity of sequence and E_m values of heme B in either protein. The replacement of the two vinyl groups with nitrile ones places two incredibly electron withdrawing groups directly attached to the heme causing the most massive shift seen in this work, up to +150mV (Table 4.1, Figure 5.11).

As expected the porphyrins with the more electron withdrawing groups, such as Diacetyl Deuteroporphyrin IX with its two acetyl groups directly connected to the porphyrin ring, have higher potentials compared to those with electron donating or electron neutral groups such as Iso Hematoporphyrin which has hydroxyl groups two methylene groups away from the porphyrin ring, only marginally interacting with the electronics.

5.3: Discussion

5.3.1: Midpoint Potential Changes Compared To Previous Maquette Work

In this chapter I attempted to alter the midpoint potential in a new series of maquette proteins. Only through changing the porphyrin was I able to get out of the range achieved by Gibney and Shifman in their earlier work. One common drawback that all maquettes have compared to natural proteins is the burial of the heme away from water. Natural proteins often bury the heme cofactor away from solvent. This takes cofactors from the high dielectric of water to the much lower dielectric of the protein core where charges are not stable. In maquettes the heme is not so buried, and it is unknown how dry the core is. As such, the dielectric inside the protein may be higher and the positive ferric-heme cofactor more stable.

Changing the potential of heme to match nature is still out of reach in maquettes. Nature has evolved hemoproteins incorporating many of the chemical concepts seen

above to shift the potential an impressive range along with burial of the cofactor away from water (Figure 5.12). Having established the magnitude of each modification, extending the E_m range appears to be a matter of combining them together. As seen in Figure 1.2, many proteins in the respiratory chain remove the heme from water, sequestering it in the membrane, and combine that with alternative ligations, other cofactor placement or

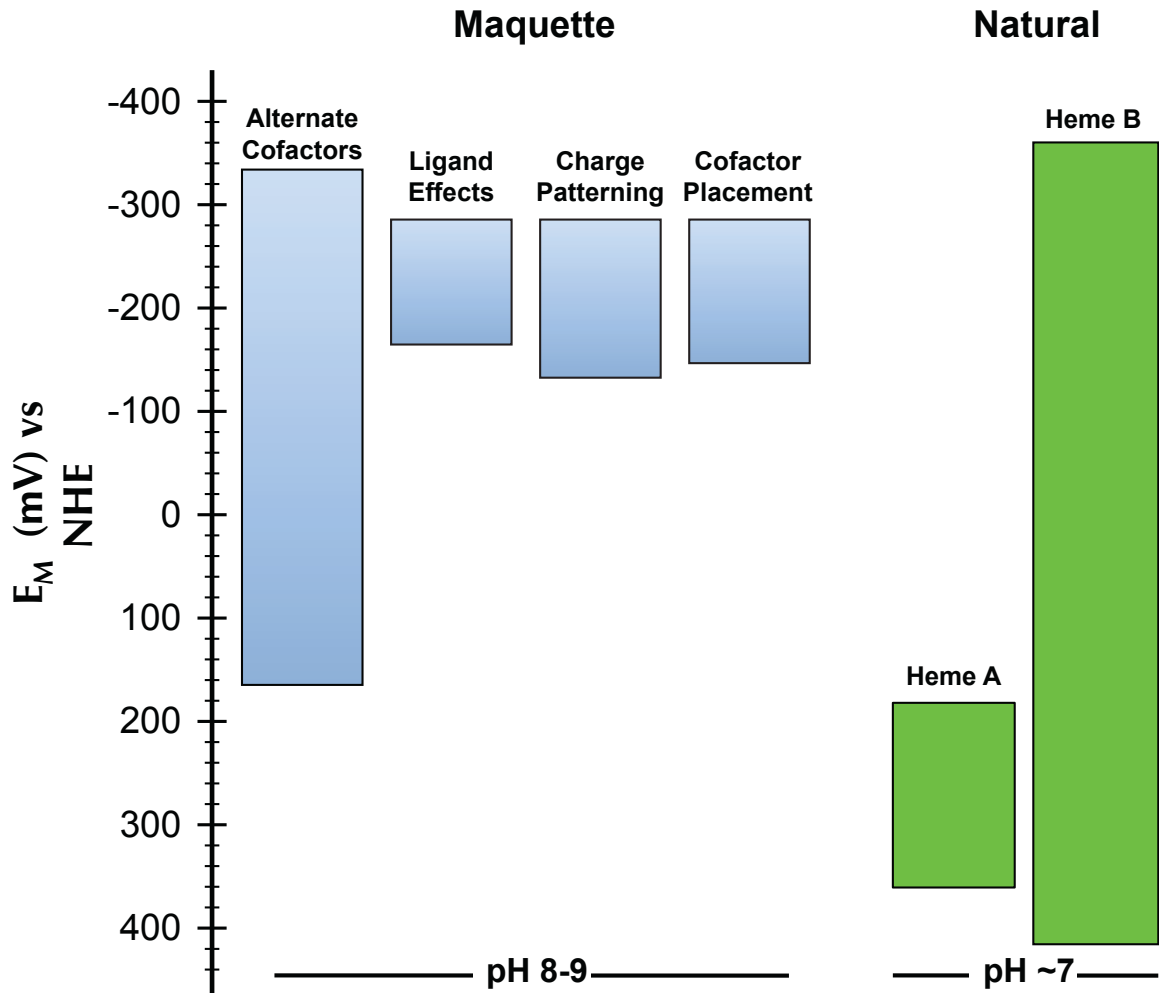


Figure 5.12: Summary of the ranges of midpoint potential values seen in maquettes, and in natural proteins. Maquette values are from this work, and the natural proteins are adapted from Shifman et. al.⁵

environmental effects to shift the potential by ~400mV.

Despite me being unable to achieve midpoint potentials out of the range of Gibney, Shifman et. al., I was able to show a shift the potential in novel ways. Due to the symmetry imposed by the earlier maquettes separate environments were impossible. In this current generation, splitting the potential of two heme cofactors by 100mV in the same protein 13Å away from one another. In addition to that, switching the ligands from bis-histidine to Cys-Ala and His-Met was not possible before either. These not only can shift the E_m value but can lead to more diversity in function.

5.3.2: Magnitude Of The Potential Changes In A New Generation Of Maquettes

Changes in E_m of 361mV were seen upon changing the porphyrin, going from -311mV with Mesoporphyrin up to 50mV with the 2,6-Dinitrileporphyrin in the same protein (Table 4.1). Adding electron-withdrawing groups to the porphyrin greatly adjusts the redox potential, an effect seen in nature between groups like heme B and heme A.¹⁶

A major advantage with this system is the affinities of maquettes for the various porphyrins. DADPIX and Mesoporphyrin, with potentials at -48mV and -319mV respectively, both have sub-nanomolar affinities with the single chain porphyrin (Table 4.1).

Changing the protein sequence to alter the charge environment is another way to shift the potential. However, this appears to have an upper limit of 150mV, as further charge modifications cause the protein to aggregate. Despite that limitation, a major advantage is the ability to include two separate charge domains in a single protein.⁹

5.3.3: Conclusions

Protein modularity is an important aspect of the maquette platform. Being able to generate two separate environments, with two separate E_m values, allows one to control the

direction and speed of electron transfer in a protein system. Having cofactors at various distances is the first step to reproducing electron transport chains similar to the ones seen in nature, creating a pathway of cofactors with increasing E_m values for electrons to follow.

The range of potentials, and different ways of achieving them, make maquettes the best way to study of protein electron-transport chains as well as the invention of novel oxidoreductase functions in a protein scaffold. Maquettes are capable of reproducing natural electron transfer reactions both the driving force (E_m), as well as the distances between cofactors. From these parameters, according the established Moser-Dutton guidelines of electron transfer,^{20,21,22} the rates can be easily controlled in this system, to the point where maquettes can artificially replicate the events of nature.

The work in this chapter will be invaluable when maquettes are to be transitioned into living cells. Not only will they have to assemble with cofactors, but they will have to regulate the electrochemistry of the cofactors as well. Using what has been shown here, this transition will be straightforward, and maquettes will be able to both incorporate into natural electron transfer processes as well as form their own.

5.4: References

- (1) de Montellano, P. R. O. *Chem Rev* **2010**, *110*, 932.
- (2) Daldal, F. *Bba-Bioenergetics* **2004**, *1658*, 63.
- (3) Tamura, M.; Woodrow, G. V., 3rd; Yonetani, T. *Biochim Biophys Acta* **1973**, *317*, 34.
- (4) Dutton, P. L. *Methods Enzymol* **1978**, *54*, 411.
- (5) Shifman, J. M.; Gibney, B. R.; Sharp, R. E.; Dutton, P. L. *Biochemistry-Us* **2000**, *39*, 14813.
- (6) Shifman, J. M.; Moser, C. C.; Kalsbeck, W. A.; Bocian, D. F.; Dutton, P. L. *Biochemistry-Us* **1998**, *37*, 16815.
- (7) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry-Us* **1998**, *37*, 4635.
- (8) Gibney, B. R.; Huang, S. S.; Skalicky, J. J.; Fuentes, E. J.; Wand, A. J.; Dutton, P. L. *Biochemistry-Us* **2001**, *40*, 10550.

- (9) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton¹ In *Nat Chem Biol* 2013; Vol. Submitted.
- (10) Caffrey, M. S.; Cusanovich, M. A. *Arch Biochem Biophys* **1991**, 285, 227.
- (11) Hay, S.; Wydrzynski, T. *Biochemistry-Us* **2005**, 44, 431.
- (12) Sligar, S. G.; Cinti, D. L.; Gibson, G. G.; Schenkman, J. B. *Biochem Biophys Res Commun* **1979**, 90, 925.
- (13) Itagaki, E.; Hager, L. P. *Biochem Bioph Res Co* **1968**, 32, 1013.
- (14) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; Degrado, W. F.; Dutton, P. L. *Nature* **1994**, 368, 425.
- (15) Hokanson, S. C. PhD, University of Pennsylvania, 2012.
- (16) Gibney, B. R.; Isogai, Y.; Rabanal, F.; Reddy, K. S.; Grosset, A. M.; Moser, C. C.; Dutton, P. L. *Biochemistry-Us* **2000**, 39, 11041.
- (17) Osyczka, A.; Moser, C. C.; Daldal, F.; Dutton, P. L. *Nature* **2004**, 427, 607.
- (18) Nelson, N.; Yocum, C. F. *Annu Rev Plant Biol* **2006**, 57, 521.
- (19) Kim, H. J.; Khalimonchuk, O.; Smith, P. M.; Winge, D. R. *Bba-Mol Cell Res* **2012**, 1823, 1604.
- (20) Carr, H. S.; Winge, D. R. *Accounts Chem Res* **2003**, 36, 309.
- (21) Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. *Nature* **1992**, 355, 796.
- (22) Page, C. C.; Moser, C. C.; Dutton, P. L. *Current Opinion in Chemical Biology* **2003**, 7, 551.

Chapter 6: Functions Possible In Maquettes When The Assembly Is Understood

6.1: Introduction

The main focus of protein design is to learn about protein function through designing a series of natural and engineered activities. How specifically to do this differs by approach, but all approaches aim to generate enzymatic reactions, and learn how proteins have evolved to be the catalysts they are.

As discussed in previous chapters, assembly and environment are key issues underlying any function. The past chapters of this thesis have investigated these aspects in protein maquettes describing how the protein itself contributes to making a functional holo-protein. In this section we describe a series of functions carried out by these proteins using tetrapyrrole cofactors, the designs of which were aided by our newfound knowledge of protein biophysical chemistry.

6.2: Results

6.2.1: *Electron Transfer*

The first function I attempted to recreate is a simple demonstration of electron transfer between two maquettes, reproducing the many electron transfer reactions common to oxidoreductase proteins.¹ I assembled two separate proteins using what I have learned about modulating midpoint potential. The first protein is heme B bound to the single-chain protein. This had a measured midpoint potential of -260mV. This was mixed with another single-chain protein containing Diacetyl Deuteroporphyrin IX (DADPIX) that has an E_m value of -50mV. Together, these two proteins generate a 210mV driving force between the

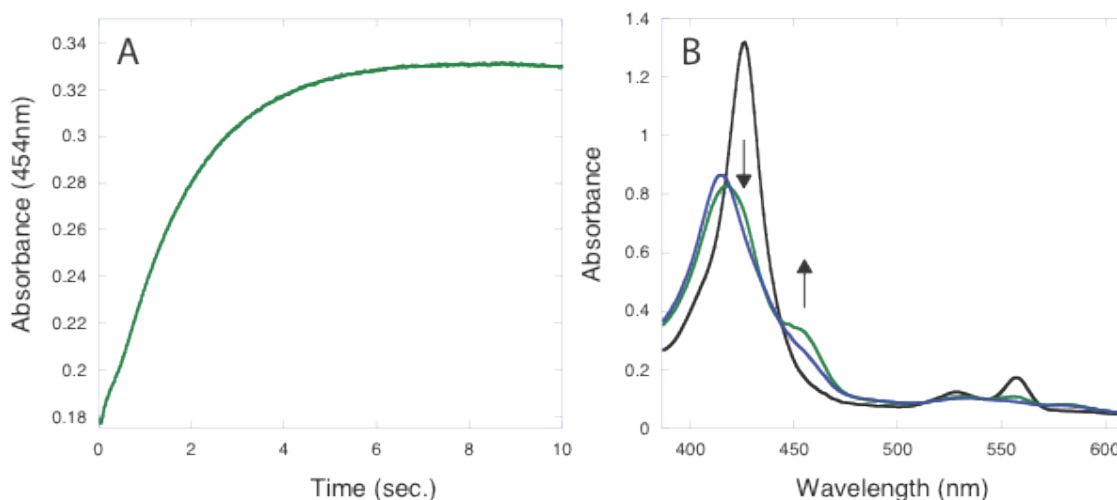


Figure 6.1: Demonstration of electron transfer between two maquettes using different cofactors. Here showing 5 μ M heme in the single chain negative protein transferring an electron to 5 μ M DADPIX in the single chain protein. Panel (A) depicts the increase of absorbance at 454nm, the Soret band of reduced DADPIX. The spectra of which are shown in (B) Spectra collected in a stopped-flow apparatus at 20 $^{\circ}$ C in CHES buffer pH 9.

pair, or a ΔG of -4.84 kcal/mol (Figure 5.11, Table 4.1). Figure 6.1A shows a spectral trace at 454nm, the reduced DADPIX soret peak, as a function of time. This peak appears while the reduced heme peak at 427nm decays (trace not shown). Figure 6.1-B shows UV/Vis spectra at three separate time points throughout this experiment. The arrows are highlighting the growth of the reduced DADPIX. These experiments were done at 15 $^{\circ}$ C in CHES buffer pH 9, and the absorbance at 314nm, the absorbance maximum of dithionite, was monitored to assure no excess reductant was present.

These two proteins exchange electrons at a rate of approximately $5.8 \times 10^{-7} \text{ M s}^{-1}$, a rate 3.6 times faster the calculated diffusion limited rate of $1.62 \times 10^{-7} \text{ M s}^{-1}$. The calculated diffusion-limited rate constant was determined through the following equation²

$$k_d = 4\pi R (D_{\text{Heme}} + D_{\text{DADPIX}})$$

Where k_d is the diffusion limited rate constant, R is the critical distance at which a reaction will take place, and D is the diffusion coefficient which can be determined from the following equation:²

$$D = \frac{k_b T}{6 \pi \eta a}$$

Where k_b is Boltzmann's constant, T is the temperature in Kelvin, η is the viscosity of the solution (here assumed to be 1.002 MPa•sec, the value of water at 20°C), and a is the radius of the molecule at which a reaction will take place (here assumed to be 10Å).²

The two diffusion constants (D) were both assumed to be the same as the difference between the two samples is a partially buried porphyrin, and the protein used was the same in both samples. This value was calculated to be $4.289 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$, which gives a k_d of $1.078 \times 10^{-17} \text{ m}^3 \text{ s}^{-1}$. Multiplying by Avogadro's number and accounting for m^3 into liters converts this value into a rate constant of $6.493 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$. Multiplying this value by the concentration of each protein species (5µM each) gives the final rate of $1.62 \times 10^{-7} \text{ M s}^{-1}$. The experimental rate was obtained from the slope of a linear regression fit to the first 1.5 seconds of the reaction (Figure 6.1A Inset).

Despite the slight discrepancy, these two rates are very close to one another, indicating that the reaction is likely diffusion limited. The calculated rate however made some assumptions that could account for the discrepancy. First, the calculation assumed that the viscosity of the solution at 20°C was that of water (1.002 mPa sec), not water with salt and pH 9 buffer, which would increase the viscosity. Furthermore, the viscosity was not altered, a common test of diffusion limited reactions. The other assumption, that 13Å was

the distance at which the two proteins had to be from one another before they would transfer electrons (assuming a reorganization energy of 0.7). It may be farther considering

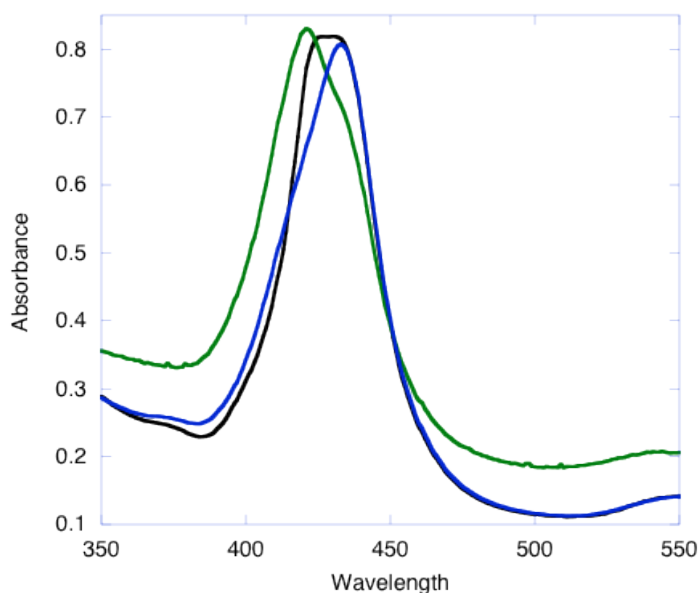


Figure 6.2: Electron transfer at low temperatures between heme and DADPIX in the single chain protein. The system was setup with reduced holoprotein bound to CO mixed with oxidized DADPIX-holoprotein (black trace). The setup was flashed to remove the CO from the heme and allow electron transfer (blue trace), and then allowed to fully oxidize (green trace). Each holoprotein at 5 μ M and data collected at -10 $^{\circ}$ C in CHES buffer pH 9 with 30% v/v ethylene glycol.

the high driving force for this reaction (-4.84 kcal/mol).

As a control to provide further support that this reaction is occurring, and it is not an artifact of excess reductant, the reaction was also done with a CO trapped heme reduced state, or a carboxyferrous state. Doing this allowed for a scrubbing of dithionite by bubbling through oxygen gas, which can be replaced with an inert gas when the experiment begins (Ar was

used in this experiment). Upon a light flash from a projection lamp, CO is photolysed from the heme and the electrons are free to be transferred. This is seen in figure 6.2, where upon the addition of light, the heme appears to go oxidized while the DADPIX appears to not only go reduced, but also bind CO. This method is limited in the sense that kinetics could not be obtained.

Inter-protein electron transfer has been done many times before with natural proteins, but has not yet been examined with two synthetic proteins. This data shows not

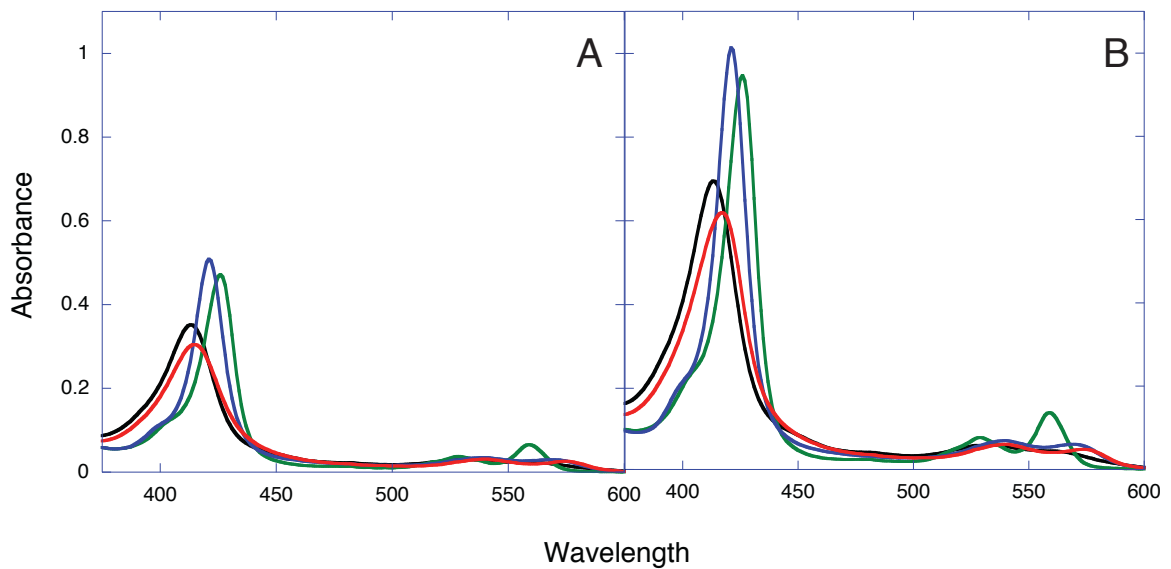


Figure 6.3: The Single chain protein is able to generate an oxyferrous state independently at each heme binding site. Here one (A) and two (B) heme containing proteins are depicted going from oxidized (black) to reduced (green) then to carboxyferrous (blue) and finally, the oxyferrous (red). Experiments done at -15°C in CHES buffer pH 9 with 30% v/v ethylene glycol. $4.5\mu\text{M}$ protein was used for these experiments.

only electron transfer but has evidence of gaseous ligand transfer as well. This is the first part of a system of oxidoreductase proteins that can generate an electron transport chain and possibly shuttle substrates along the same gradient.

6.2.2: Porphyrin Sequestration

An equally simple function is akin to that of the Histidine rich protein (HRP) of *Plasmodium falciparum*.^{3,4} This protein's sole known function is to bind and sequester free heme present in a red blood cell (RBC) when the parasite is present due to the toxic effects of free heme that have been discussed.^{5,6} This protein's function is amongst the first steps in hemozoin formation.⁷ This simple function has been shown to be possible throughout this thesis as the sequestration process of the heme is rapid, discussed in chapters 2 and 3.

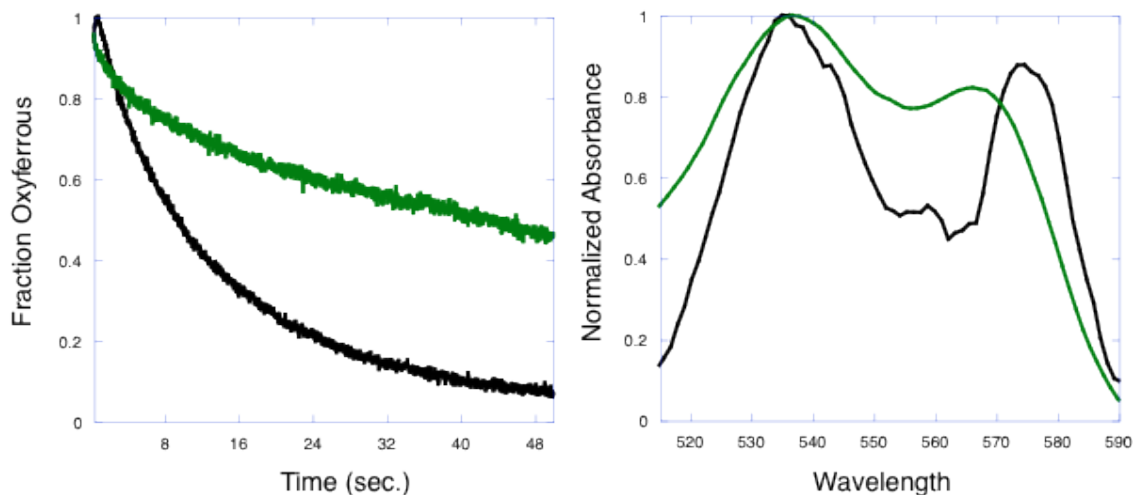


Figure 6.4: Oxyferrous decay with either the his-cys (green) single chain or the bis-his (black) single chain protein. Despite no change in E_m , the oxyferrous lifetime is extended approximately four times its original length. Data collected in CHES buffer pH 9 at 15°C. Protein at 10 μ M. Traces shown are the decay at 576nm (green) and 574nm (black).

HRP, like the maquettes is a helical histidine rich protein patterned such that many histidines are buried in a hydrophobic patch greatly increasing the affinity of the protein for heme.

6.2.3: Oxygen Binding

Another function we were able to achieve that is slightly more complex is oxygen binding.^{8,9} The previous work in this lab has established this state before, uncovering the engineering guidelines needed to form an oxyferrous state,¹⁰ and generating an oxyferrous state with a half-life of approximately 2.25 seconds. This oxyferrous state takes advantage of the modular nature of heme maquettes. In both sites it forms and decays at the same rate, independent of the other heme occupancy (Figure 6.3).

Using our knowledge of protein assembly, we were able to increase the oxyferrous state half-life to approximately 10.2 seconds with the use of an alternate ligation scheme. Replacing one of the histidine residues with a cysteine allows for a large increase in the lifetime brought about by a few possible factors (Figure 6.4). The cysteine, being a smaller residue than the histidine, leaves more room in the binding pocket better able to accommodate the oxygen. The different geometry can also result in a more stable hydrogen bond between the cysteine and oxygen. It is possible that this better oriented bond is able to hold the oxygen in place relative to histidine due to the geometry and steric factors. Extending the lifetime of the oxyferrous state is a big step moving forward. It is the first step toward more complex chemistry in the maquette interior.

6.2.4: Light Harvesting

The final function being discussed is a rudimentary light harvesting reaction. This is a primary step in photosynthesis, and is performed with chromophores capable of absorbing photons and transferring the energy to other cofactors that induce more complex chemistry.¹¹ In maquettes, I was able to reproduce the initial steps of this process. This work follows from that of Dror Noy, however that work chose cofactors empirically, not with considerations of their binding rates or affinities.¹² I ligated Zn Mesoporphyrin to a modified single-chain protein (H7H112: Sequence Appendix). When this was done we were able to generate a FRET pair between that pigment and tryptophan. Figure 6.5 shows four different excitation spectra, where we are exciting the chromophore pairs over the range shown and collecting the emissions at 630nm, the fluorescence maximum of Zn-mesoporphyrin. When the cofactor is excited in this wavelength range there is no fluorescence, regardless of the solvent used (DMSO or aqueous CHES Buffer). The protein

has some fluorescence emission in this range, as it contains tryptophan, and it is increased

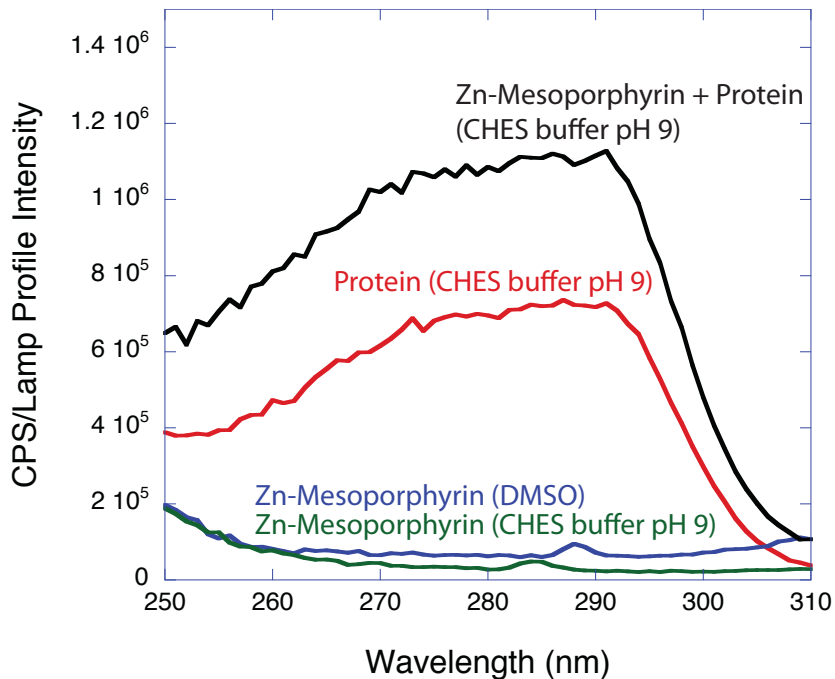


Figure 6.5: FRET between tryptophan and Zn-Mesoporphyrin in the single chain protein. If the chromophore is excited between 250 and 310nm there is no emission at 630nm regardless of the solvent being CHES (blue) or DMSO (green). The protein at 5 μ M has a mild emission at this wavelength (red). When the protein is excited with a chromophore bound it is able to transfer energy over and the total emission at 630nm is increased.

significantly when mixed with the pigment. The increased emission comes from tryptophan indole ring absorbing the energy from light and transferring it to the Zn-mesoporphyrin via FRET. With an R^0 value of 30Å this is a very efficient reaction. What is shown here are the first steps in light harvesting function akin to the

LHC complex in plants and photosynthetic bacteria. In the future, this could potentially be combined with the functions above to generate an entire system wherein a photon can be absorbed and transferred to another chromophore photolysing a CO molecule and opening it up for some other function.

6.3: Discussion

The work shown in this chapter details some of the various functions achieved in a maquette platform. Beginning with the simple heme binding and slowly working up to more and more complex functions we show the immense versatility of the maquette platform and approach working in knowledge of the most basic physical-chemical principles to further our designs.

The simplest functions of all the ones discussed, the one that underlies all the others, is heme binding. This function is well described in the previous chapters but is important in nature as well. In addition to the HRP protein in *Plasmodium falciparum*,³ there are many other proteins whose sole purpose is to bind and sequester heme.¹³ The hemopexin and ISD proteins, as mentioned in the introduction, are excreted proteins with possibly the strongest known affinities for heme.^{13,14} Due to the low K_d values, it is feasible to think that maquettes would be ideal substitutions for this natural function. Their assembly mechanism has been discussed here thoroughly.

In addition to their function as heme transporting proteins, these maquettes can also be exploited for their electron transfer properties. Maquettes can be tailored to adopt a wide range of potentials, through environmental effects or alternate porphyrins. This simple function has a vast amount of potential for future use. Biology uses vast electron transport chains in two of its most fundamental processes: ATP synthesis and Photosynthesis.¹⁵ Maquettes with varying potentials thereby generating an artificial electron transport. By demonstrating a relatively simple example of this complex process, maquette show their potential for higher order functions, which can be combined with the other functions to mimic photosynthesis or respiration activities.

The Dutton lab has also gone a long way toward generating membrane bound proteins capable of electron transfer.¹⁶ These can be readily incorporated into these

electron transport schemes thereby increasing the functional versatility of this system and allowing the generation of a proton gradient across a membrane. Maquettes are also capable of incorporating quinine amino acids,¹⁷ with this, the entire function of the *bc₁* protein can be recreated in a much more simple system.

This chapter does more than demonstrate simple functions in a maquette system it demonstrates the potential of the maquette system when the underlying principles are known. In a simple scaffold there are a myriad of possibilities that can be attained, and the more that is known about the underlying engineering required to assemble a specific combination of cofactors with protein, the more complex the functions can become. Using this work, the transition from the test tube to the cell is not far off. When the holomaquettes can assemble spontaneously, a cell can be fed a specific cofactor. This can lead to novel functions being carried out by living systems.

6.4: References

- (1) Moser, C. C.; Page, C. C.; Farid, R.; Dutton, P. L. *J Bioenerg Biomembr* **1995**, *27*, 263.
- (2) Houston, P. L. *Chemical Kinetics and Reaction Dynamics*; First ed.; Dover Publications Inc.: Mineola, NY, 2001.
- (3) Marletta, M. A.; Schneider, E. L. *Biochemistry-Us* **2005**, *44*, 979.
- (4) Pagola, S.; Stephens, P. W.; Bohle, D. S.; Kosar, A. D.; Madsen, S. K. *Nature* **2000**, *404*, 307.
- (5) Hao, K.; Hanawa, H.; Ding, L. M.; Ota, Y.; Yoshida, K.; Toba, K.; Ogura, M.; Ito, H.; Kodama, M.; Aizawa, Y. *Mol Immunol* **2011**, *48*, 1191.
- (6) Yukl, E. T.; Jepkorir, G.; Alontaga, A. Y.; Pautsch, L.; Rodriguez, J. C.; Rivera, M.; Moenne-Loccoz, P. *Biochemistry-Us* **2010**, *49*, 6646.
- (7) Weissbuch, I.; Leiserowitz, L. *Chem Rev* **2008**, *108*, 4899.
- (8) Tamura, M.; Woodrow, G. V., 3rd; Yonetani, T. *Biochim Biophys Acta* **1973**, *317*, 34.
- (9) Zhang, L.; Anderson, J. L. R.; Ahmed, I.; Norman, J. A.; Negron, C.; Mutter, A. C.; Dutton, P. L.; Koder, R. L. *Biochemistry-Us* **2011**, *50*, 10254.
- (10) Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.
- (11) Minagawa, J.; Takahashi, Y. *Photosynth Res* **2004**, *82*, 241.
- (12) Noy, D.; Moser, C. C.; Dutton, P. L. *Bba-Bioenergetics* **2006**, *1757*, 90.

- (13) Grigg, J. C.; Ukpabi, G.; Gaudin, C. F. M.; Murphy, M. E. P. *J Inorg Biochem* **2010**, *104*, 341.
- (14) Tolosano, E.; Altruda, F. *DNA Cell Biol* **2002**, *21*, 297.
- (15) David L. Nelson, M. M. C. *Lehninger Principles of Biochemistry*; Third Edition ed.; Worth Publishers: New York, NY, 2000.
- (16) Ye, S. X.; Discher, B. M.; Strzalka, J.; Xu, T.; Wu, S. P.; Noy, D.; Kuzmenko, I.; Gog, T.; Therien, M. J.; Dutton, P. L.; Blasie, J. K. *Nano Lett* **2005**, *5*, 1658.
- (17) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun* **2009**, 168.

Chapter 7: Conclusions

This thesis explains the way the maquette approach works. We define operational benchmarks from the underlying physical chemistry, and using that we are able to incorporate novel functions into the maquette framework. Understanding how the protein affects the assembly and reactivity of various cofactors is more important now as the new generation of maquettes, and functions becomes more complex.

This ground up approach starts with the most basic aspect of function: assembly. Up to now, the lab has taken advantage of the simple system and spontaneous assembly. Optimizing this process is key for a multitude of reasons. If maquettes are to be inserted into cells they must be able to compete with the natural and present proteins when it comes to binding cofactors. Alternatively, if cofactors are to be added to cells they must be able to insert themselves into the machinery present. Maquettes must be capable of quickly assembling in a test tube so that their assembly *in vivo* can be optimized and devoid of side reactions. Assembly is also relevant for industrial applications. It is not profitable for a company to have a holoprotein that spontaneously assembles over the course of weeks or requires chaperones. A much more practical option is one where the protein can spontaneously assemble on the seconds timescale and is robust enough to be transported to distant locations, synthetic blood is one example. If this could be developed rapidly and transported long distances it could provide medical aid to people in need all over the world.

Chapters 3 is an investigation of that and how assembly is modified through changes to the protein sequence. This has uncovered different types of kinetic limitations enforced through the protein's gross structural characteristics, not by individual amino

acids. Both entropic and enthalpic barriers were uncovered that have been seen before. We have shown that proteins must be monomeric in solution as to avoid problems linked to preassembly. These problems are mainly caused by diffusion and protein rearrangement and are defined as an entropic barrier. We also observed that a densely packed core is also a drawback, this one being enthalpic in nature. This knowledge allows for future designs to be enhanced by being able to pull out what is kinetically inhibiting them. The Eyring plot of figure 3.6 can be thought of as a diagnostic tool, one that can be utilized when there are problems binding. Once the barrier has been uncovered, it is a simple matter of either making the core less stable or adding a structural element.

Chapter 4 deals with the other half of this reaction: The cofactor. In this section we see various different factors that can impede its binding. Though not necessarily due to solubility, the rate of assembly is highly correlated with the log-P value, providing a rough guide to aid in cofactor selection. This is not the only mitigating factor, the structure of the tetrapyrrole ring itself, and the metal center also have noticeable effects on the assembly rate. These principles could be a boon to future designers. If assembly is slow, but the protein appears optimized, the cofactor may not be suitable for rapid assembly. Looking at its log-P value can test this hypothesis, which can be modified, for example with a new dendrimer increasing the rate.

It is hoped that with this knowledge the future designs will not be limited with assembly steps before they are ready to assume their function.

Though assembly is important, making sure the protein sequence will have a desired effect on the reactivity is equally important. All throughout nature there are hemoproteins with a wide range of potentials.¹ Bringing maquettes into the cell and having them take part in biology means they will need to have an equally diverse range of

potentials. The maquette approach also allows a high level of control over distances and the scaffold's high tolerance to mutations help to create a docking site for other proteins. This could lead to the beginning of artificial electron transport chains.

Chapter 5 investigates the electrochemistry thoroughly, and we examine four areas where the maquette can be applied such that a heme B cofactor has a wide range of potentials. The most intense range came from changing the cofactor itself. Attaching electron-withdrawing groups to the porphyrin directly drove the potential up approximately 370mV. Other methods of changing the environment, or ligations also had a significant effect, shifting the E_m by as many as 140mV.

As in nature, the maquette approach is not bound to one set of potential shifting mutations. Various ones can be mixed together to generate high potential heme cofactors spanning the range of natural enzymes. This work will be integral to these designs, laying the groundwork for the future of raised heme midpoint potentials. If maquettes are to be introduced into living systems, it is these principles that will influence their sequence.

By addressing the underlying physical chemistry we are able to engineer functions into a maquette system. Being able to both bind a series of cofactors and modulate their potentials was the first step in creating the first pair of maquettes capable of electron transfer. Though a small first step, this concept is key to the beginning of larger arrays that can eventually mimic the electron transport properties of the respiration enzymes in nature. In fact progress has already been made by Fry, Discher et. al. putting maquettes into membranes and transporting electrons across it.² Coupled with the progress made with Flavins and quinones the maquette approach can these pieces together and generate a completely synthetic proton motive force.

Light harvesting is another area where the maquette approach could have a major impact. Aided with the work here describing cofactor selection and protein requirements, a light harvesting and light-energy transfer maquette was produced. This maquette, also demonstrates a proof of principle concept that this complex function is possible on a much smaller and simpler scale. With knowledge of chromophore attributes, expanding the range of light that can be used is the next step to generating complex light activated charge separation and from there complex, high oxidative chemistry, like that seen in oxygen generation. Goutham Kodali has already begun adding a wide array of chromophores extending the range of captured solar light from what is seen in nature.

Though simple, heme binding and sequestration plays a vital role in cellular viability. As mentioned previously, free heme is toxic and free light harvesting tetrapyrroles also can be potentially hazardous to the cell.^{3,4} The function of binding thus plays a vital role keeping ROS levels low. Indeed many organisms have complex systems that do nothing but simply bind heme and sequester it until it is needed later.⁵ In those terms the maquettes are already poised to play a role in the cell, protecting from ROS and aggregated porphyrins. The single chain maquette is only limited in its heme binding by the cofactor. In the cell free heme is kept to a minimum, and therefore heme aggregation is a minor issue. The assembly would be limited by diffusion alone.

Finally this work has also improved the function of oxygen binding. By adding a cysteine residue to the heme-binding site the oxyferrous lifetime was extended approximately 5 times its original. This lifetime-extension has many important applications like synthetic blood as mentioned above. This work establishing increased E_m values could also have a substantial impact in this function as well, and Molly Sheehan is working toward that end.

This thesis uncovers the physical chemistry key to many diverse oxidoreductase functions. We have described the binding kinetics filling out Scheme 3.1. The barriers can be clearly uncovered and the protein can be modified to avoid them. The cofactor effects have also been characterized. We have now a deeper understanding of holoprotein assembly, and can better modify designs in the future to reflect that. We have also expanded the range of electrochemical potentials in maquettes, and have paved a clear path toward generating artificial electron transport chains. Future maquette work will be able to expand the functional repertoire in these systems better explaining natural enzymes as well as putting synthetic ones to use for the benefit of society.

7.1: References

- (1) Shifman, J. M.; Gibney, B. R.; Sharp, R. E.; Dutton, P. L. *Biochemistry-US* 2000, 39, 14813.
- (2) Ye, S. X.; Discher, B. M.; Strzalka, J.; Xu, T.; Wu, S. P.; Noy, D.; Kuzmenko, I.; Gog, T.; Therien, M. J.; Dutton, P. L.; Blasie, J. K. *Nano Lett* 2005, 5, 1658.
- (3) Hao, K.; Hanawa, H.; Ding, L. M.; Ota, Y.; Yoshida, K.; Toba, K.; Ogura, M.; Ito, H.; Kodama, M.; Aizawa, Y. *Mol Immunol* 2011, 48, 1191.
- (4) de Paula, J. C.; Robblee, J. H.; Pasternack, R. F. *Biophys J* 1995, 68, 335.
- (5) Yukl, E. T.; Jepkorir, G.; Alontaga, A. Y.; Pautsch, L.; Rodriguez, J. C.; Rivera, M.; Moenne-Loccoz, P. *Biochemistry-US* 2010, 49, 6646.

Appendix 1: Materials And Methods

A.1: Materials

All materials were purchased from Sigma-Aldrich unless otherwise noted. Solvents were purchased from Fisher Scientific.

A.2: Porphyrins

All porphyrins were purchased from Frontier Scientific unless otherwise noted. They were dissolved in DMSO and their concentration was determined by mass with a couple of exceptions. Heme B, Fe-mesoporphyrin, and Fe-Deuteroporphyrin had their concentrations measured by the pyridine hemochrome assay, utilizing the extinction coefficient at 556nm as described here.¹ Of a solution of 40% pyridine with 100mM NaOH, 997uL was placed into a cuvette. To this 1uL of a heme stock was added as well as 2uL of a $K_3Fe(CN)_6$ to ensure the sample was fully oxidized. After full oxidation was confirmed, a spatula tip containing sodium dithionite was added to the cuvette. The mixture was given 10 seconds to reduce, and a spectrum was taken. This was done with at least 4 different volumes varying only the heme and pyridine-mixture volumes, not the $K_3Fe(CN)_6$. The absorbance of 556nm – 540nm was monitored and plotted as a function of heme volume added. This was fit to a line with a slope of $(34.7 \times [\text{heme}])$ and a intercept set to zero. This line is a simple use of beer's law coupling the absorbance to the concentration. The value of specific extinction coefficients can be found in the spectral appendix.

Heme-a was purified and its mass was quantified as described here.² Bruce R. Lichtenstein synthesized 2,6-Dinitrile porphyrin. Its synthesis protocols can be found in his PhD thesis. Its concentration was determined by weight and its purity is unknown.³

A.3 Proteins

The proteins used in this paper were obtained as follows. The gene for the protein was ordered from DNA2.0. The construct contained a His-tag, a linker, and then the protein of interest. No fusion was used. Upon delivery of the gene the DNA was resuspended in TE buffer and transformed into both DH5 α and BL-21 strain E. coli cells. For expression, the BL-21 cells were grown in TB media until the OD600 was at 0.6 AU then over-expression was induced by addition of IPTG to a concentration 1mM. The cells over-expressed the protein for 4.5 hours and then were spun down to a pellet. The pellet was then resuspended in lysis buffer (300mM NaCl, 50mM NaH₂PO₄, 10mM Imidazole, pH 7) and sonicated at an amplitude of 90 (5 times, 20 seconds each time). The lysate was spun down for 1 hour at 20,000g. The supernatant was then run through a GE-Histrap column. The eluate was TEV cleaved and rerun on the column. The flowthrough of this second column was dialyzed into CHES buffer (20mM CHES, 150mM KCl, pH 9).⁴

Its concentration was determined by the absorbance at 280nm using an extinction coefficient of 22,500 M⁻¹ cm⁻¹. The extinction coefficient was determined with the ProtParam software on the ExPASy website (<http://web.expasy.org/protparam/>).

Where the proteins could not be ordered as is and mutagenesis had to be done primers were ordered from Integrated DNA Technologies (IDT) and dissolved to a level of 10uM. Using this stock, 1 μ L was added to a solution containing the plasmid (obtained from either DNA2.0 or from a previous lab member). 1 μ L of this primer was mixed with 5ng of

template DNA, polymerase mix to 1X (2X master mix: 4 μ L 5X Phusion buffer, 0.5 μ L 10mM dNTP, 0.2 μ L Phusion Hot Start, 0.2 μ L 50mM MgCl₂, 5.1 μ L Milli-Q water), and Milli-Q water to obtain a 10 μ L final volume. This was subject to a PCR cycle of 98°C for 30 seconds, followed by 98°C for 10 seconds, Annealing temperature for 30 seconds, 72°C for the extension time (repeat for 15 cycles), followed by a final extension step at 72°C. When this was completed, the reaction was held at 4°C until needed. This DNA was then transformed as discussed above and the protein was purified as discussed pending sequencing.

Certain proteins were unable to be made in *E. coli*. These were prepared at a 0.1mmol scale on a CEM Liberty microwave peptide synthesizer using standard FMOC/tBu protection protocols.⁵ Amino acids were purchased from Nova Biochem. The side chain protecting groups were as follows: Cys (trt), Lys (Boc), His (Boc), and Asp/Glu (OtBu). After synthesis the protein was cleaved from the resin by incubating it with a mixture of trifluoroacetic acid (TFA), ethanedithiol, Anisole and thioanisole in a 9: 0.2 : 0.5 : 0.3 ratio for two and a half hours protected from light. The crude product then had excess reagent removed *in vacuo* and was subsequently precipitated with methyl,t-butyl ether. The peptide was then purified by reverse phase HPLC using a C18 column and a linear gradient of Acetonitrile with 0.1% (v/v) TFA in water with 0.1% (v/v) TFA. The products were verified with MALDI Mass spectroscopy using a 2-(4¹-Hydroxybenzeneazo)benzoic acid (HABA) matrix.

A.4: Partition Coefficients

Partition coefficients were determined by dissolving the chromophore in 1-octanol and taking a spectrum to make sure there was no aggregation.⁶ This was then mixed with

an equal volume water, vortexed, and allowed to sit for one hour at room temperature. After this hour the spectra of the octanol layer was taken again. The solet absorbance was taken again and subtracted from the pre-mixed solet, this value is counted as the concentration of porphyrin in the water layer. The two values were analyzed in the following way:

$$\text{Log } [\text{Porphyrin}]_{\text{octanol layer}} / [\text{Porphyrin}]_{\text{water layer}}$$

This value is known as the Log-P value, or partition coefficient. The concentration of porphyrin in the aqueous layer was calculated as the concentration in Octanol before minus the concentration in octanol after mixing and incubation. This is described below

$$[\text{Porphyrin}]_{\text{aqueous layer}} = [\text{Porphyrin}]_{\text{octanol before mixing}} - [\text{Porphyrin}]_{\text{octanol after mixing and incubation}}$$

A.5: Rates

Millisecond scale time measurements were made on an OLIS RSM 1000 stopped flow spectrophotometer. The tetrapyrrole and protein were added to separate syringes and shot together monitoring the absorbance (typically from 386nm to 611nm unless otherwise noted). The apparatus had a Fischer-Scientific IsoTemp 3031 water bath attached, allowing for temperature control. The experiments in this thesis were determined at variable degrees Celsius, noted in the specific experiments. Multiple spectroscopic species were determined using SVD analysis included in the OLIS Globalworks software. Individual wavelengths were identified by eye for further kinetic analysis.

The samples had to be loaded onto the machine already in buffer. Solutions of DMSO were not used for these experiments.

Data Fitting for the Eyring plots is as follows: The rates were determined by first fitting the initial slopes from 0 up to 20msec. The slope was then divided by an extinction

coefficient of 110,000 M⁻¹ cm⁻¹ to obtain the value in the proper units. The rates were then fit to the following equation to obtain the pseudo-thermodynamic parameters of the transition state obtained from Transition State Theory:⁷

$$k = \kappa (k_B T/h) e^{(\Delta S^\ddagger/R)} e^{(-\Delta H^\ddagger/RT)}$$

Where k is the initial rate for the reaction. k_B is the boltzman constant, and h is plank's constant. T is the temperature in Kelvin. R is the gas constant in units of cal K⁻¹ mol⁻¹. ΔS^\ddagger and ΔH^\ddagger are the entropy and enthalpy differences between the ground reactant state and the transition state respectively. κ is transmission coefficient, a term used to describe the amount of reactions that go to the transtition state and proceed back to reactants, here it is assumed to be one, indicating that all transition states proceed to a product and do not dissociate back into reactants.

A.6: Spectra

Protein solutions were prepared in CHES buffer. Binding was monitored by UV/Vis Soret absorbance on a Varian Cary-50 spectrophotometer at room temperature in a 1 cm path quartz cuvette. Secondary structure was monitored by CD spectroscopy (Aviv Model 410) with a 1 mm path quartz cuvette. Thermal denaturation monitored the ellipticity at 222 nm every 5°C from 5°C to 95°C after 15 minutes of equilibration. Melting temperatures were calculated using a Boltzmann equation with one term for each observed transition.

A6.1: Fluorescence Spectra

Fluorescent data was collected on Florolog Jobin-Yvon fluorometer. A standard temperature of 25°C was held constant with a Fischer-Scientific IsoTemp 3031 water bath.

A 1cm path-length cuvette, graded for fluorescent work was used. Experiments were performed in CHES buffer pH 9.

A.7: K_d Titrations

Affinity titration experiments were performed with low concentration of protein typically in the nanomolar range. Porphyrin was added in 0.1 equivalents and a spectrum was taken. These were fit to the tight binding equation:⁴

$$A = H_{\text{tot}} \epsilon_{\text{free}} + (\epsilon_{\text{bound}} - \epsilon_{\text{free}}) \frac{(K_D + P_{\text{tot}} + H_{\text{tot}} - \sqrt{(K_D + P_{\text{tot}} + H_{\text{tot}})^2 - 4 P_{\text{tot}} H_{\text{tot}}})}{2}$$

where A is the absorbance at the Soret band, H_{tot} is the total heme concentration, P_{tot} is the total protein concentration, K_d is the affinity constant, ϵ_{bound} is the extinction coefficient of bound porphyrin, and ϵ_{free} is the extinction coefficient of free porphyrin. These were performed in either a 1cm or 10cm pathlength cuvette. The cuvette volumes of each are 3mL or 30mL respectively.

A.8: E_m Titrations

Redox titrations were performed in combination with UV/Vis monitoring, adapted from Dutton et. al.^{8,9} Samples of typically 5-25 μ M were monitored electrochemically by a calomel electrode purchased from Radiometer analytical. The change from oxidized to reduced was monitored by the change in absorbance at the q-band region of the bis-his ligated porphyrin. These titrations were done anaerobically with a constant stream of Ar being blown over the top of the sample. The E_h was modulated by 1-3 μ L injections of a

freshly prepared sodium dithionite solution when trying to lower the potential or a potassium ferricyanide solution when trying to increase the potential. The following redox mediators were used: Anthra-quinone-2-sulfonate (20 μ M), Benzly Viologen (10 μ M), Methyl Viologen (10 μ M), Sulfanilamide (10 μ M), Indigo trisulfonate (10 μ M), Phenazine (10 μ M), Pyocyanin (10 μ M), and Hydroxy-Napthquinone (10 μ M).

The apparatus used was a specialized redox-cuvette designed by Dutton et. al. and is described here.⁸

The data was analyzed the same as Gibney et al 1998 Biochemistry. The absorbance value of the reduced peak (corrected for baseline shifts) was plotted against voltage and fit to the Nerst equation described in the introduction and displayed below:¹⁰

$$E_{\text{cell}} = E_{\text{cell}}^{\ominus} - \frac{RT}{zF} \ln \frac{[\text{Red}]}{[\text{Ox}]}$$

Where E_{cell} is the reduction potential, $E_{\text{cell}}^{\ominus}$ is the standard potential, R is the gas constant, T is the temperature in Kelvin, z is the number of moles of electrons involved in the half cell reaction, F is Faraday's constant (96485 C mol⁻¹), [Red] is the concentration of reduced components, and [Ox] is the concentration of oxidized components.

This equation is converted to the following when used for fitting:

$$\text{Fraction Reduced} = \frac{1}{10^{\frac{E_h - E_m}{[RT/nF]}} + 1}$$

When two midpoint potentials are present, the equation can be expanded to account for that.

$$\text{Fraction Reduced} = \frac{0.5}{10^{\frac{E_h - E_{m1}}{RT/nF}} + 1} + \frac{0.5}{10^{\frac{E_h - E_{m2}}{RT/nF}} + 1}$$

Where the fraction reduced is a combination of the sum of the concentrations of each species.

A.8.1: Spectral-Electrochemistry Titrations

Spectroelectrochemical titrations were carried out with a platinum working electrode, gold counter wire, and a Ag/AgCl reference electrode. A CH Instruments (Austin Texas USA) Electrochemical workstation was used with the CH Instruments interface program version 9.07. In a cuvette with a 0.1cm path-length, 200 μ M single chain protein with tetracarboxyphenylporphyrin was added, dissolved in 20mM CHES 150mM KCl buffer pH 9. Mediators used were: Anthra-quinone-2-sulfonate (100 μ M), Benzly Viologen (50 μ M), Methyl Viologen (50 μ M), Sulfanilamide (50 μ M), Indigo trisulfonate (50 μ M), Phenazine (50 μ M), Pyocyanin (50 μ M), and Hydroxy-Naphthquinone (50 μ M). The electrodes were placed in the solution and the cuvette was sealed with parafilm. A potential was set by the computer and the current was monitored. When the current reached zero, a spectra was taken. This was repeated in 20mV steps from -450mV up to -150mV. This was done multiple times (2 oxidative, and 2 reductive) to see if there was any hysteresis. Afterward, the absorbance value of 535nm-517nm was plotted against voltage and fit to the Nernst equation as above yielding an E_m value.

A.9: Greater Than Millisecond Rate Values

Zinc-protoporphyrin IX (ZnPPIX) deaggregation experiments were performed on a Varian Cary-500 spectrophotometer. ZnPPIX was dissolved in aqueous buffer at a concentration of 24 μ M and allowed to crash out of solution overnight. This solution of aggregated chromophore was inverted four-five times to equally disperse the aggregates and 500 μ L was removed and added to a cuvette. To this cuvette 500 μ L of a 6 μ M protein was added and a spectrum was taken every 5 minutes for two hours.

A.10: Intra-maquette Electron Transfer

Electron transfer was first done on a OLIS-stopped flow apparatus as described above. The temperature was controlled with a Fischer-Scientific IsoTemp 3031 water bath. In each syringe was added the single chain protein with either DADPIX or heme B. These were degassed in the syringe for approximately one hour. To the degassed heme syringe, dithionite was added and monitored at 314nm to ensure no excess reductant was present, and the heme was monitored at 425nm to ensure it was reduced. If too much reductant was present, FeCN was added to remove it. When it was reduced, and no excess reductant could be observed, the stopped flow was fired mixing the samples together. Electron transfer was monitored in the Soret region.

A.11: Low Temperature Electron Transfer

For the low temperature electron transfer a 3mL cuvette was used. 1.5mL of 5 μ M heme in the single chain protein (CHES buffer pH 9, 30% v/v ethylene glycol) was added to this protein and degassed under Ar flow for 1 hour. This was reduced until a peak at 314nm was detected. The temperature was lowered to -10°C. At this point CO gas was bubbled through until no reduced heme could be detected, as all had been converted to

CO-ferrous heme. Oxygen was then bubbled through until no dithionite could be detected at 314nm. At this point, 1.5mL (CHES buffer pH 9, 30% v/v ethylene glycol) of 5 μ M DADPIX-single chain protein was added. A spectrum was taken and the entire sample was degassed with Ar gas for 1.5 hours. At this point, a projection lamp was turned on for approximately ten seconds to flash off the CO and a spectrum was immediately taken to observe any electron transfer. The flashes and spectra were repeated as necessary until no more change was detected.

A.12: Oxygen Binding

Oxygen binding was performed on an Varian Cary 50 as described above. 5 μ M protein with one or two heme cofactors (CHES buffer pH 9, 30% v/v ethylene glycol) was added to a 3mL cuvette and degassed under Ar flow for 1 hour. Following this, dithionite was added until the protein appeared fully reduced and no evidence of oxidized protein could be detected. At this point CO gas was bubbled through until all the reduced protein appeared converted to CO-ferrous. When the reduced heme was bound to CO, oxygen gas was bubbled through to the point that no dithionite could be detected, and then for 0.5 hours more. A projection lamp was then used to flash off the CO in the presence of oxygen. Spectra were taken until no change could be detected.¹¹

A.13: Oxygen Binding And Decay Rates

Oxygen binding was performed on an OLIS stopped flow as described above. In one syringe was a solution of CHES buffer pH 9 saturated with oxygen. Saturation was achieved by bubbling O₂ gas through for approximately 1 hour. In the other syringe was 5 μ M protein was added and reduced with dithionite until the protein appeared fully

reduced. The absorbance at 314nm was monitored to ensure no excess reductant was present. These two syringe contents were shot against each other at 15°C and the absorbance between 386 and 611nm was collected. The data was analyzed by monitoring the absorbance over time at 547nm for bis-his oxyferrous, and 576nm for his-cys oxyferrous. The lifetimes were determined by fitting the decay to simple first order exponential kinetics.¹¹

A.14: References

- (1) Berry, E. A.; Trumpower, B. L. *Anal Biochem* **1987**, *161*, 1.
- (2) Takemori, S.; King, T. E. *J Biol Chem* **1965**, *240*, 504.
- (3) Lichtenstein, B. R. Doctoral, University of Pennsylvania, 2010.
- (4) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton¹ In *Nat Chem Biol* 2013; Vol. Submitted.
- (5) Bodansky, M. *Peptide Chemistry: A Practical Approach*; 2nd ed.; Springer-Verlag: New York, 1993.
- (6) Albert Leo, C. H., David Elkins *Chem Rev* **1971**, *71*, 525.
- (7) Houston, P. L. *Chemical Kinetics and Reaction Dynamics*; First ed.; Dover Publications Inc.: Mineola, NY, 2001.
- (8) Dutton, P. L. *Methods Enzymol* **1978**, *54*, 411.
- (9) Shifman, J. M.; Gibney, B. R.; Sharp, R. E.; Dutton, P. L. *Biochemistry-U.S* **2000**, *39*, 14813.
- (10) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry-U.S* **1998**, *37*, 4635.
- (11) Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.

Appendix 2: Sequence Appendix

The following sequences were used throughout this thesis. This section will provide the sequence, appropriate references, and relevant background information.

H10A24 or Unstructured Homotetramer:

CGGGELWKLHEELLKKFEELLKLAEERLKKL

This sequence was originally developed by Brian Gibney and published in various papers.¹⁻³ It binds two *b*-type hemes per four helices and has split midpoint potentials at -230mV and -148mV. This was chosen to be included in this work as it has been thoroughly characterized structurally and electrochemically, providing a maquette example of a protein with a split potential stemming from heme proximity. Though one helix, the N-terminal Cysteine residue allows it to form a dimerized helix-loop-helix motif

BB or the Structured Homotetramer:

CGGGEIWKLHEEFLKKFEELLKLHEERLKKM

First developed and published by Steve Huang, this protein was chosen due to its stable apo structure and low-resolution holo-state.^{4,5} I measured the K_d value as well as the binding rates. No midpoint titration data of this protein is included in this work. Like H10A24, this protein is a single helix that forms a helix-loop-helix motif through its N-terminal cys residue

C38S or the Homodimer

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

First developed by Ronald Koder, this protein was chosen for this paper due to its resemblance to the tethered-homodimer without the tether. I attempted to measure the K_d value however at the concentrations required the protein was not dimerized sufficiently.

HP7 or Tethered Homodimer

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGCGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

Also developed by Ronald Koder, this protein was used as the first example of an oxyferrous state in an artificial framework.^{6,7} It was included in this work as it is a dimeric protein that functions as a monomer making it a key example for binding kinetics. Its low and previously established midpoint potential and simple sequence provided a scaffold for probing the effects of modifications on the E_m value. It also was the basis for the oxygen binding work attempted here.

BT6 or the Single Chain

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

Designed as a thesis project by Tammer Farid and Bruce Lichtenstein, this protein is based on HP7 however with the disulfide removed and a peptide linker in its place.⁸ This was chosen as the final step of this sequence family in rate studies. It has the least amount of symmetry allowing for single amino acid mutations and was subsequently used as a basis for further designs

Cytochrome b_{562}

ADLEDNMETLNDNLKVIKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSEPMKDF
RHGFDILVGQIDDALKLANEGKVKEAQAEEQLKTTRNAYHQKYR

A single chain four-helix bundle protein found in *E. coli* thought to be used in electron transfer reactions.⁹⁻¹¹ This protein was chosen, as it is a well-known and studied heme-

binding protein, to provide a comparison of natural proteins to maquettes. It has a 6nM K_d value and a +189mV E_m value.¹²

BT6-F or SC-F

GEIWKQFEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQFEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQFEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQFEDALQKFEDALNQFEDLKQL

This protein was designed by Tammer Farid and used as a non-heme-binding control for various studies here.

GLBin-Heme

GEIWKQHEDALQKFEEALNQFEDLKQLGGSGKGS GG
EIKRQHEDALRKFEALKRFEDKKQKGGSGSGSGG
EIWKQHEDALQKFEEALNQFEDLKQLGGSGKGS GG
EIKRQHEDALRKFEALKRFEDKKQK

Goutham Kodali and myself designed this protein. I used it as a test of charge patterning on the E_m values of heme in maquettes.

Positive mutant (POS-mut)

GEIKRQHEDALRKFEALKRFEDKKQKGGSGKGS GG
EIWKRHEDALRKFEALKRFEDKKQKGGSGKGS GG
EIWKRHEDALRKFEALKRFEDKKQKGGSGKGS GG
EIKQRHEDALRKFEALKRFEDKKQK

I designed this protein as a tool to further study the effect of charge patterning on the heme E_m value.

F56HA3

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGCGSGG
EIWKQFADALQKFADALNQHADLKQL

This protein, a variant of HP7, was used in this work to show the effect of charge patterning on E_m values. Instead of adding positive charges, this simply has negative charges removed.

HP7A6

GEIWKQHADALQKFADALNQFADLKQLGGSGCGSGG
EIWKQHADALQKFADALNQFADLKQL

This protein was designed for the purposes of E shifts and gaseous ligand binding control.

E8A

GEIWKQHADALQKFEDALNQFEDLKQLGGSGCGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

Designed to generate a midpoint potential shift albeit unsuccessfully. Appears to have a similar affinity for heme as HP7

E57A

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGCGSGG
EIWKQHEDALQKFEDALNQFADLKQL

Like E8A, this protein was designed to generate a E_m shift, though also unsuccessful.

Appears to have a similar affinity for heme as HP7

GLBin-K59,112L

GEIWKQHEDALQKFEEALNQFEDLKQLGGSGKGS GG
EIKRQHEDALRKFEELKRFEDLKQKGGSGSGSGG
EIWKQHEDALQKFEEALNQFEDLKQLGGSGKGS GG
EIKRQHEDALRKFEELKRFEDLKQK

This protein designed by Goutham Kodali was used to elucidate the change of E_m and K_d by the change in a buried Lysine.

BT6-CYP

GEIWKQCEDALQKFEDALNQFEDLKQLGGSGSGSGG

EIWKQAEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQL

This protein was designed by Goutham Kodali and used to show the effect of ligation changes on the E_m of a heme.

BT6-HisCys

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQCEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQL

This protein was designed by Goutham Kodali and used to show the effect of ligation changes on the E_m of a heme.

HP7-Offset

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGCGSGG
EIWKQFEDALQKHEDALNQFEDLKQL

This protein was designed by me and used to move the heme around inside the bundle to generate a split midpoint potential

BT6-H7-112

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQAEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

This protein was used for the light capture reactions. It was chosen over BT6 proper for its lack of ability to bind heme B as there are no bis-his binding sites.

A2.2: References

- (1) Gibney, B. R.; Huang, S. S.; Skalicky, J. J.; Fuentes, E. J.; Wand, A. J.; Dutton, P. L. *Biochemistry-U.S.* **2001**, *40*, 10550.
- (2) Gibney, B. R.; Isogai, Y.; Rabanal, F.; Reddy, K. S.; Grosset, A. M.; Moser, C. C.; Dutton, P. L. *Biochemistry-U.S.* **2000**, *39*, 11041.
- (3) Shifman, J. M.; Gibney, B. R.; Sharp, R. E.; Dutton, P. L. *Biochemistry-U.S.* **2000**, *39*, 14813.
- (4) Huang, S. S.; Gibney, B. R.; Stayrook, S. E.; Dutton, P. L.; Lewis, M. *J Mol Biol* **2003**, *326*, 1219.
- (5) Huang, S. S.; Koder, R. L.; Lewis, M.; Wand, A. J.; Dutton, P. L. *P Natl Acad Sci USA* **2004**, *101*, 5536.
- (6) Koder, R. L.; Valentine, K. G.; Cerda, J.; Noy, D.; Smith, K. M.; Wand, A. J.; Dutton, P. L. *J Am Chem Soc* **2006**, *128*, 14450.
- (7) Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.
- (8) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton1 In *Nat Chem Biol* 2013; Vol. Submitted.
- (9) Feng, Y. Q.; Sligar, S. G. *Biochemistry-U.S.* **1991**, *30*, 10150.
- (10) Feng, Y. Q.; Wand, A. J.; Sligar, S. G. *Biochemistry-U.S.* **1991**, *30*, 7711.
- (11) Itagaki, E.; Hager, L. P. *Biochem Bioph Res Co* **1968**, *32*, 1013.
- (12) Sturtevant, J. M.; Robinson, C. R.; Liu, Y. F.; Thomson, J. A.; Sligar, S. G. *Biochemistry-U.S.* **1997**, *36*, 16141.

Appendix 3: Spectral Appendix

A3.1: Heme

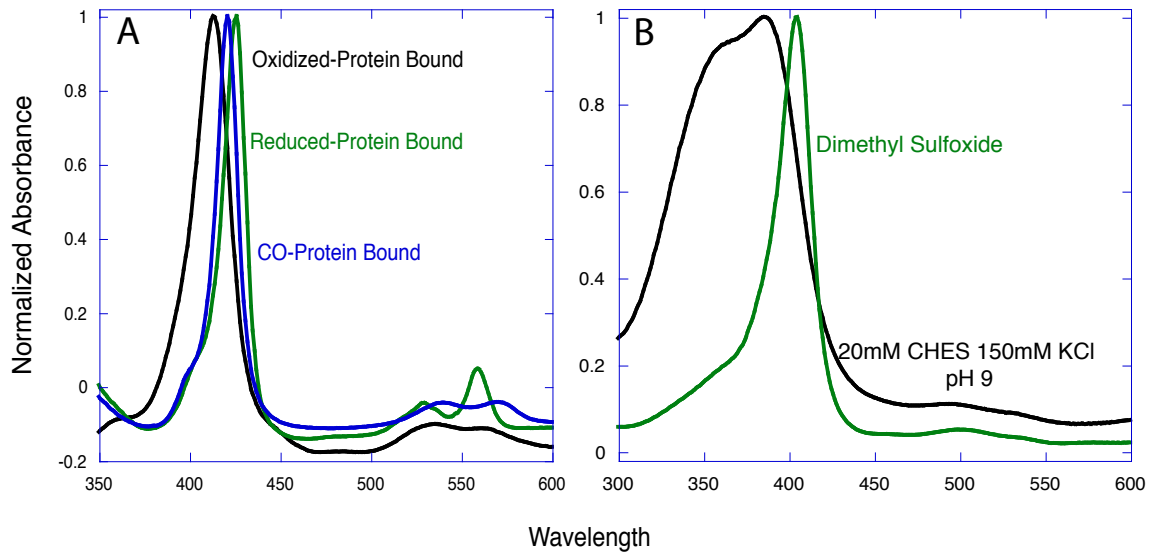


Figure A3.1: (A) Heme bound to protein in its oxidized (black), reduced (green) and carbon-monoxide (CO) bound state (blue). (B) Heme in Dimethyl Sulfoxide (DMSO) and in 20mM CHES 150mM KCl buffer. Wavelength in nm. Absorbance Normalized to Soret peak.

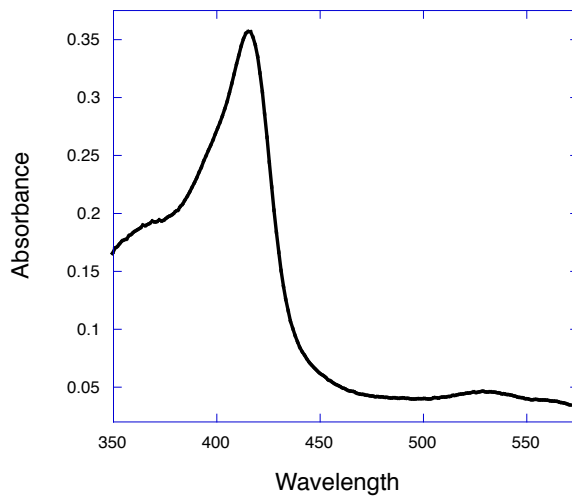


Figure A3.2: Spectra of Heme at 3.5 μM in the b562 protein

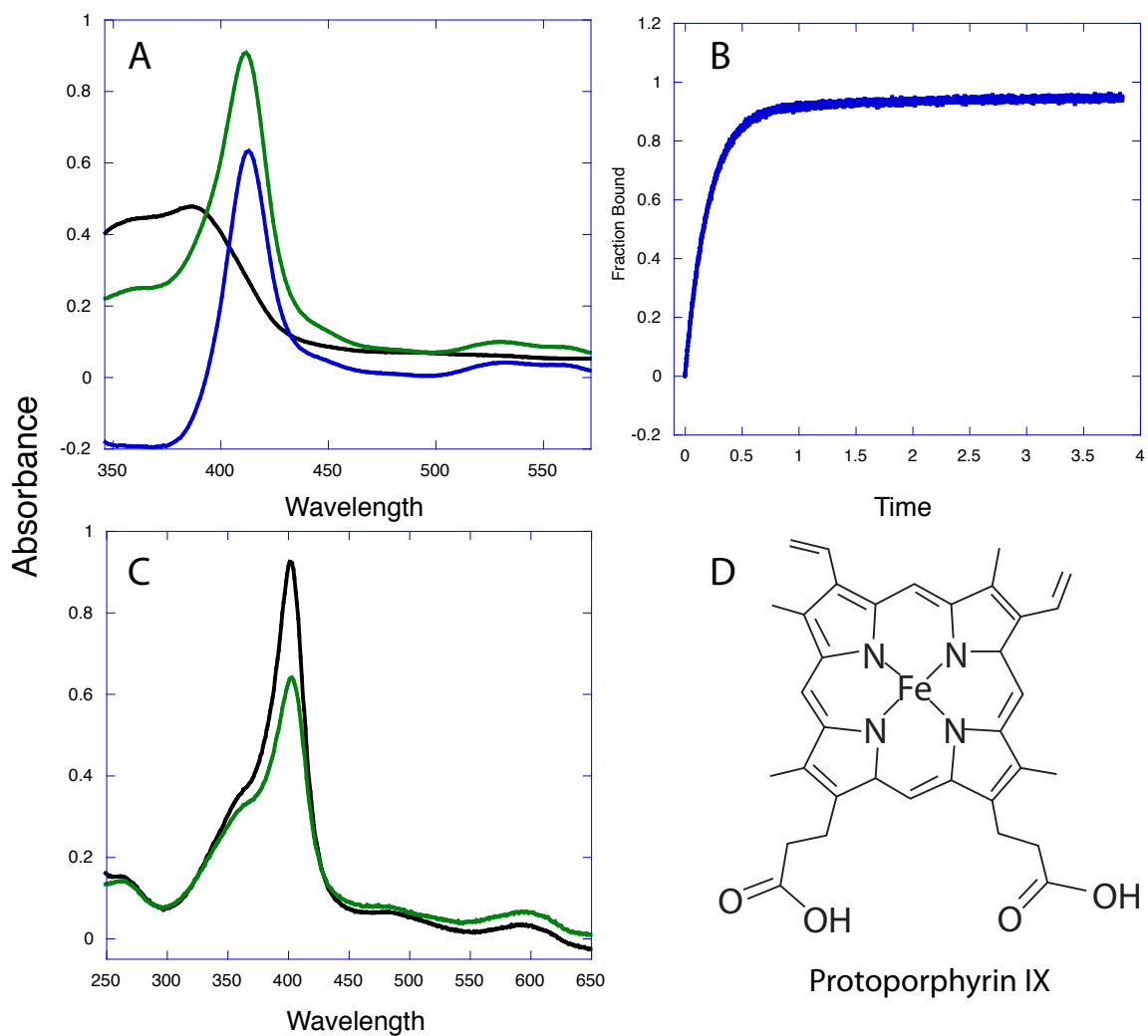


Figure A3.3: (A) Free heme B in solution (black), heme bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of heme B assembling with the single chain protein. (C) Heme B in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Heme B. All wavelengths are in nm. The extinction coefficient at 412nm is $109,700 \text{ M}^{-1}\text{cm}^{-1}$.¹

A3.2: Alternate Fe-porphyrin Kinetics

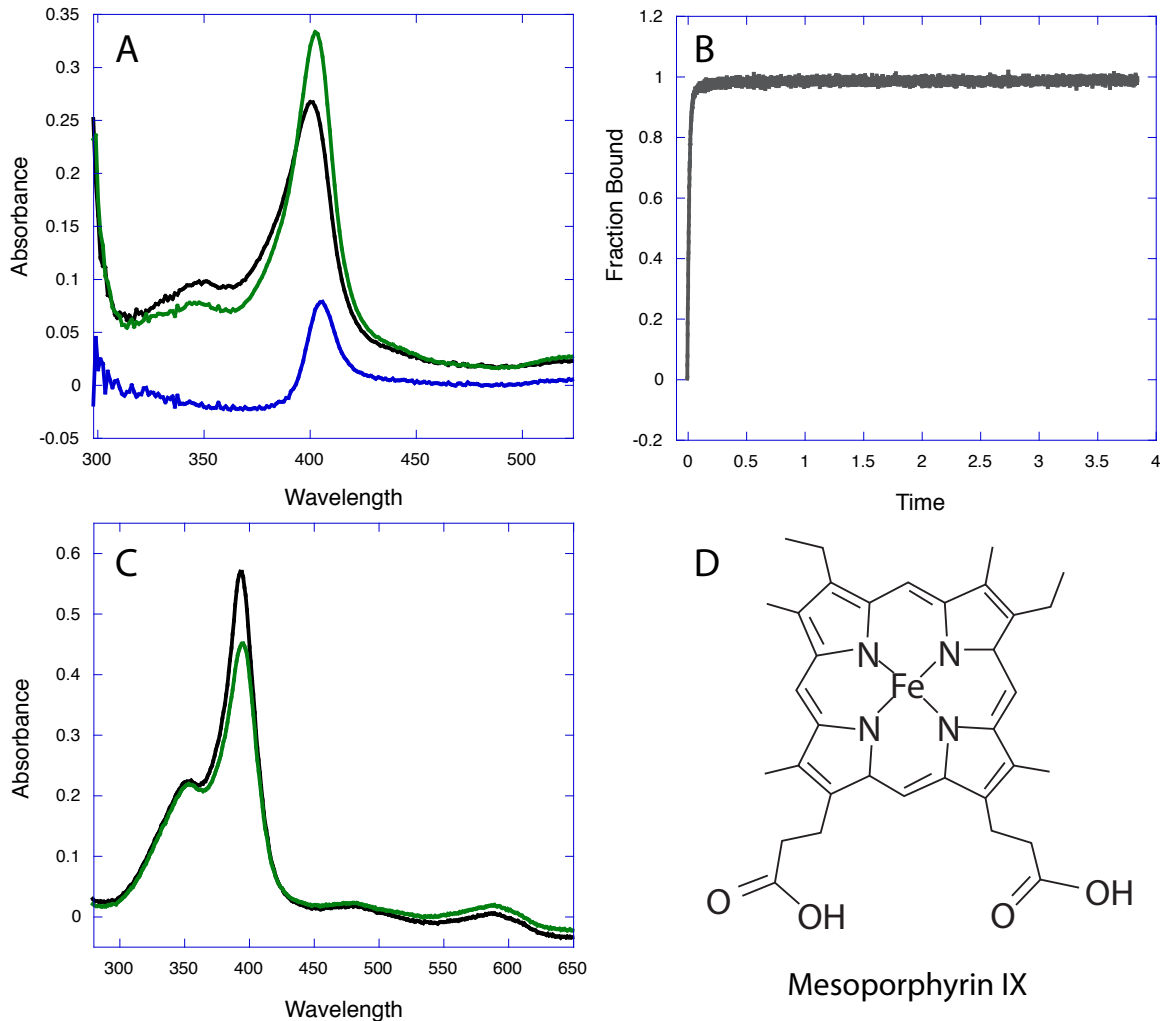


Figure A3.4: (A) Free Mesoporphyrin in solution (black), Mesoporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Mesoporphyrin assembling with the single chain protein. (C) Mesoporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Mesoporphyrin. All wavelenghts are in nm. The extinction coefficient at 405nm is $115,000 \text{ M}^{-1}\text{cm}^{-1}$.²

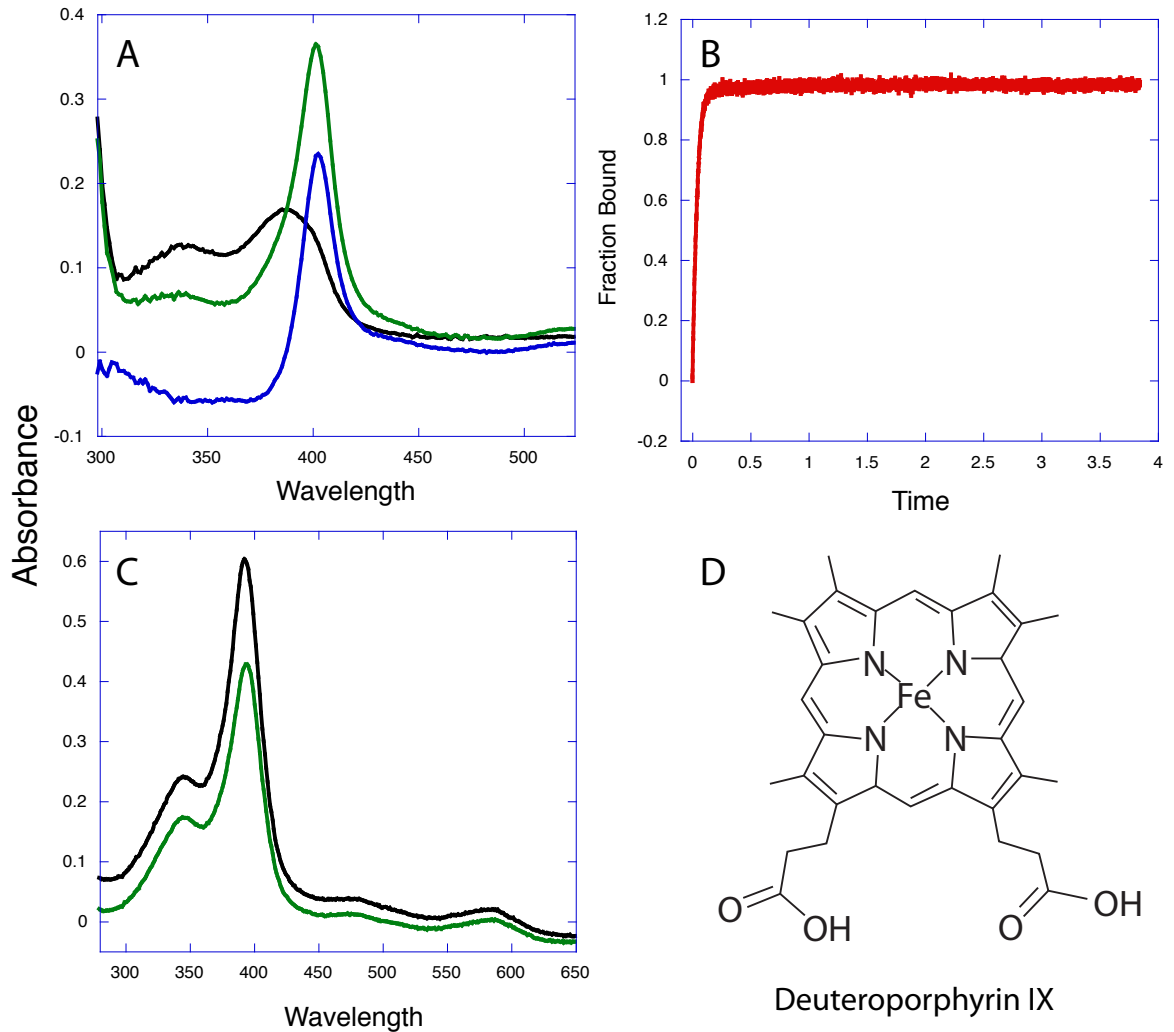


Figure A3.5: (A) Free Deuteroporphyrin in solution (black), Deuteroporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Deuteroporphyrin assembling with the single chain protein. (C) Deuteroporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Deuteroporphyrin. All wavelengths are in nm. The extinction coefficient at 403nm is $122,000 \text{ M}^{-1}\text{cm}^{-1}$.²

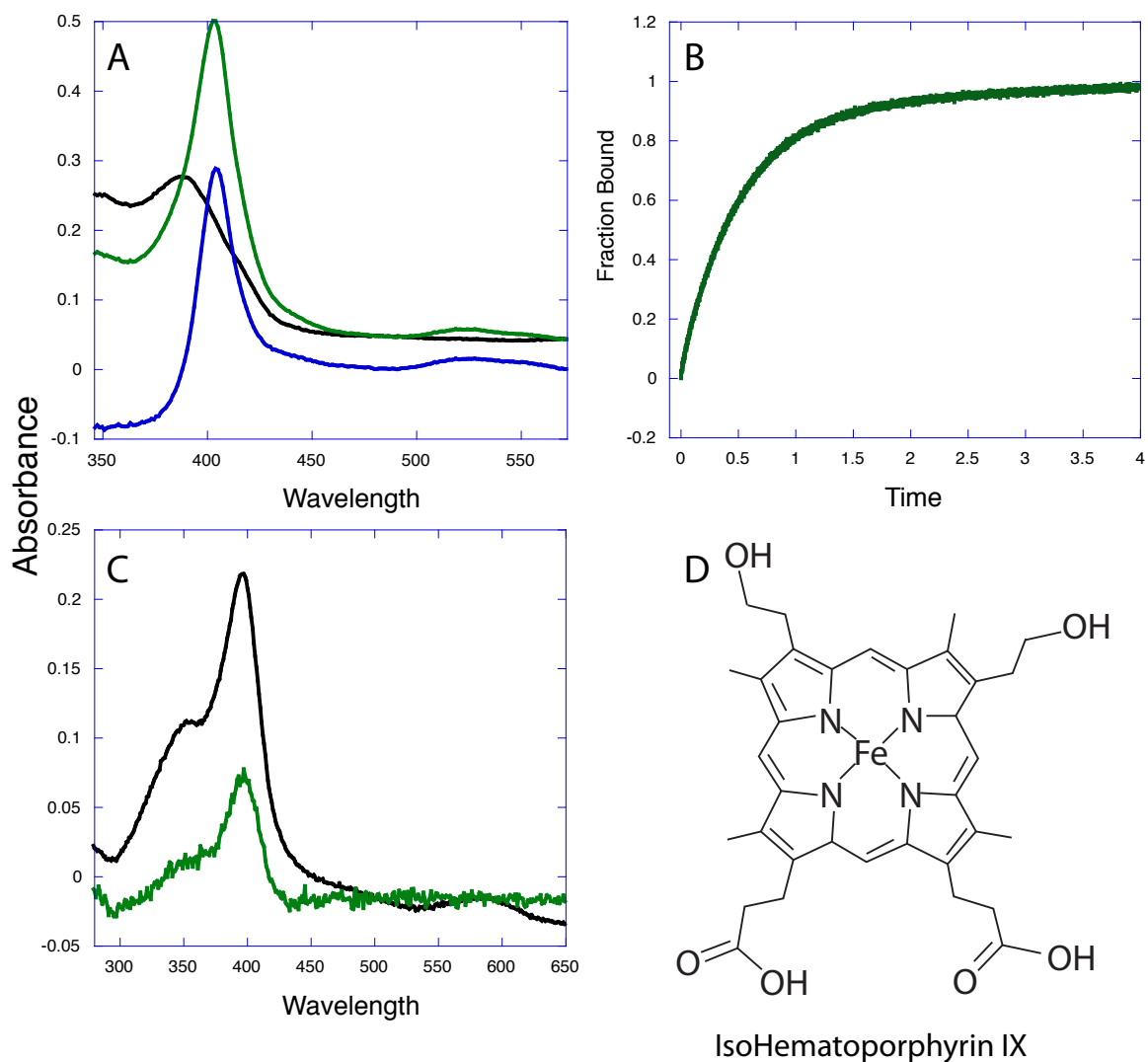


Figure A3.6: (A) Free Iso Hematoporphyrin in solution (black), Iso Hematoporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Iso Hematoporphyrin assembling with the single chain protein. (C) Iso Hematoporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Iso Hematoporphyrin. All wavelengths are in nm.

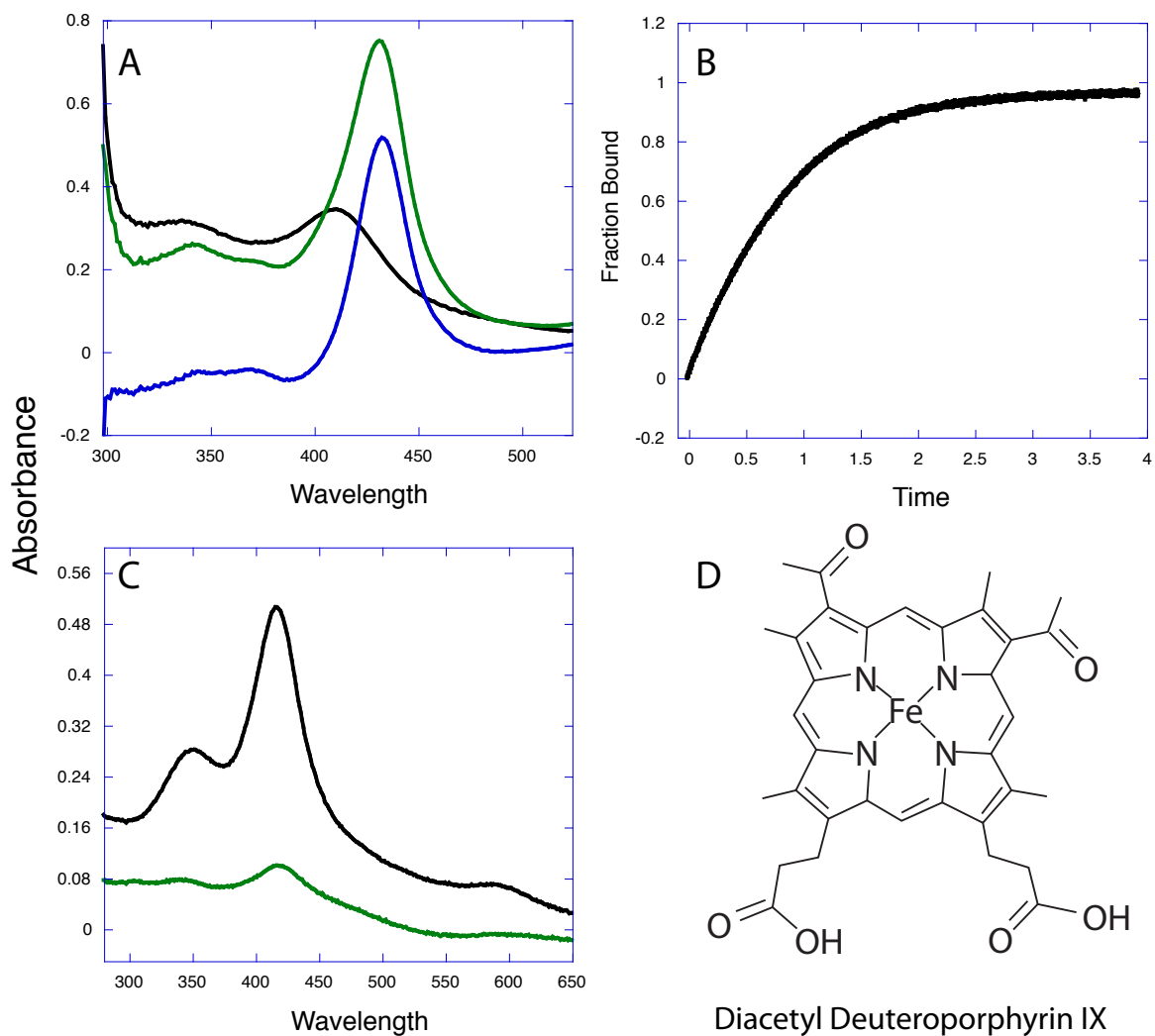


Figure A3.7: (A) Free Diacetyl Deuteroporphyrin in solution (black), Diacetyl Deuteroporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Diacetyl Deuteroporphyrin assembling with the single chain protein. (C) Diacetyl Deuteroporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Diacetyl Deuteroporphyrin. All wavelengths are in nm. The extinction coefficient at 426nm is $89,000 \text{ M}^{-1}\text{cm}^{-1}$.³

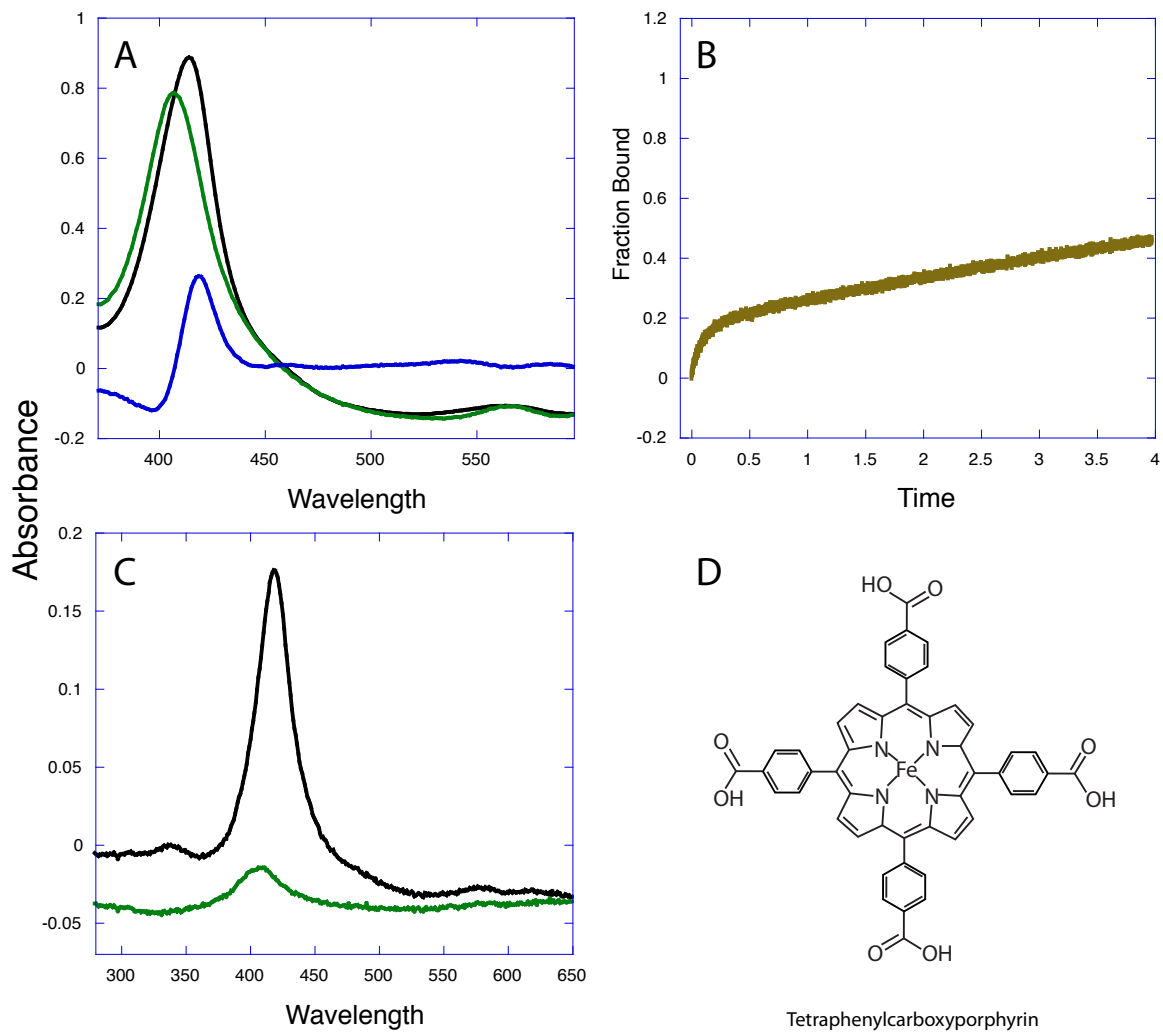


Figure A3.8: (A) Free Tetracarboxyphenylporphyrin in solution (black), Tetracarboxyphenylporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Tetracarboxyphenylporphyrin assembling with the single chain protein. (C) Tetracarboxyphenylporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Tetracarboxyphenylporphyrin. All wavelenghts are in nm.

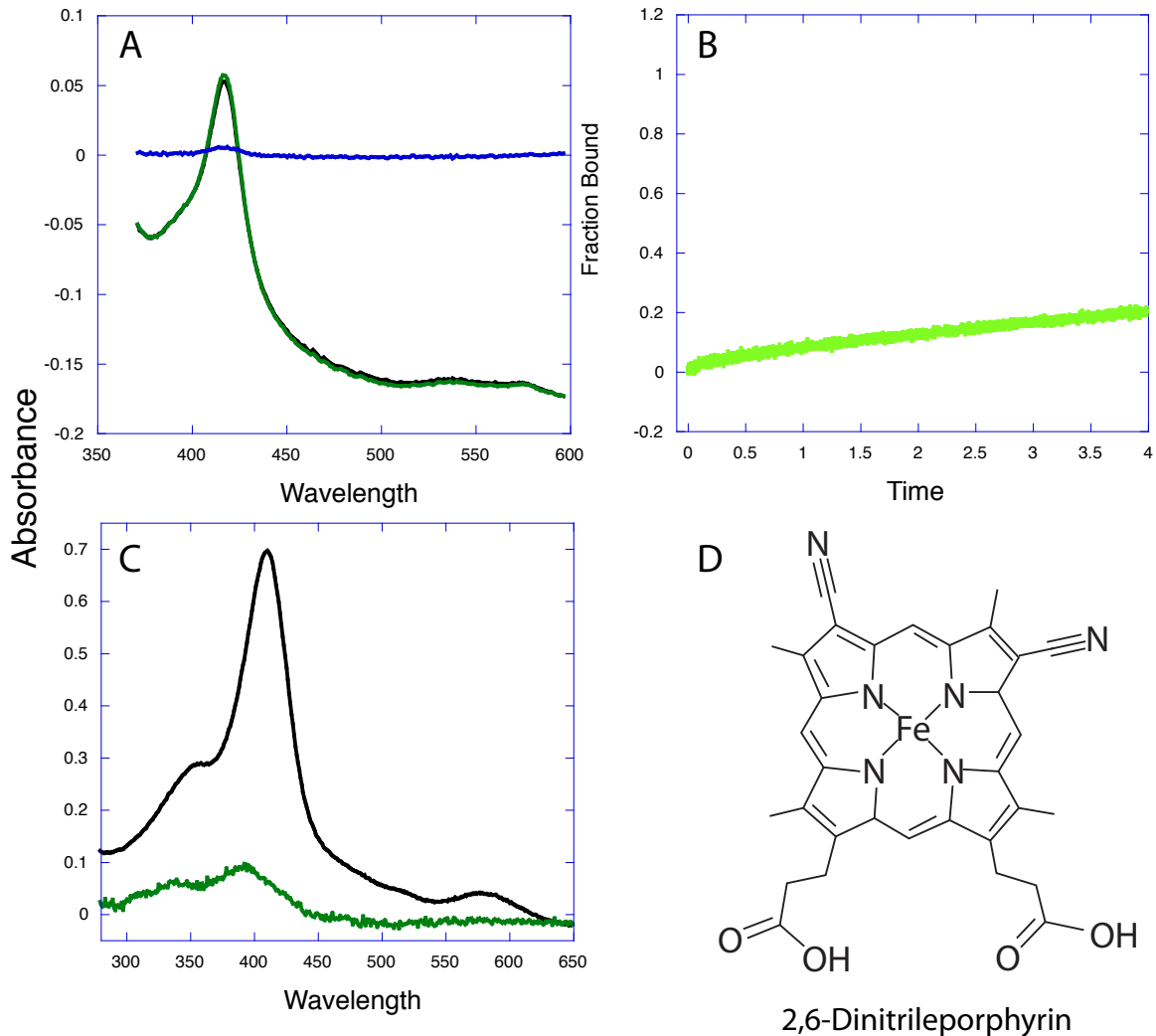


Figure A3.9: (A) Free 2,6-Dinitrileporphyrin in solution (black), 2,6-Dinitrileporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of 2,6-Dinitrileporphyrin assembling with the single chain protein. (C) 2,6-Dinitrileporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of 2,6-Dinitrileporphyrin. All wavelengths are in nm.

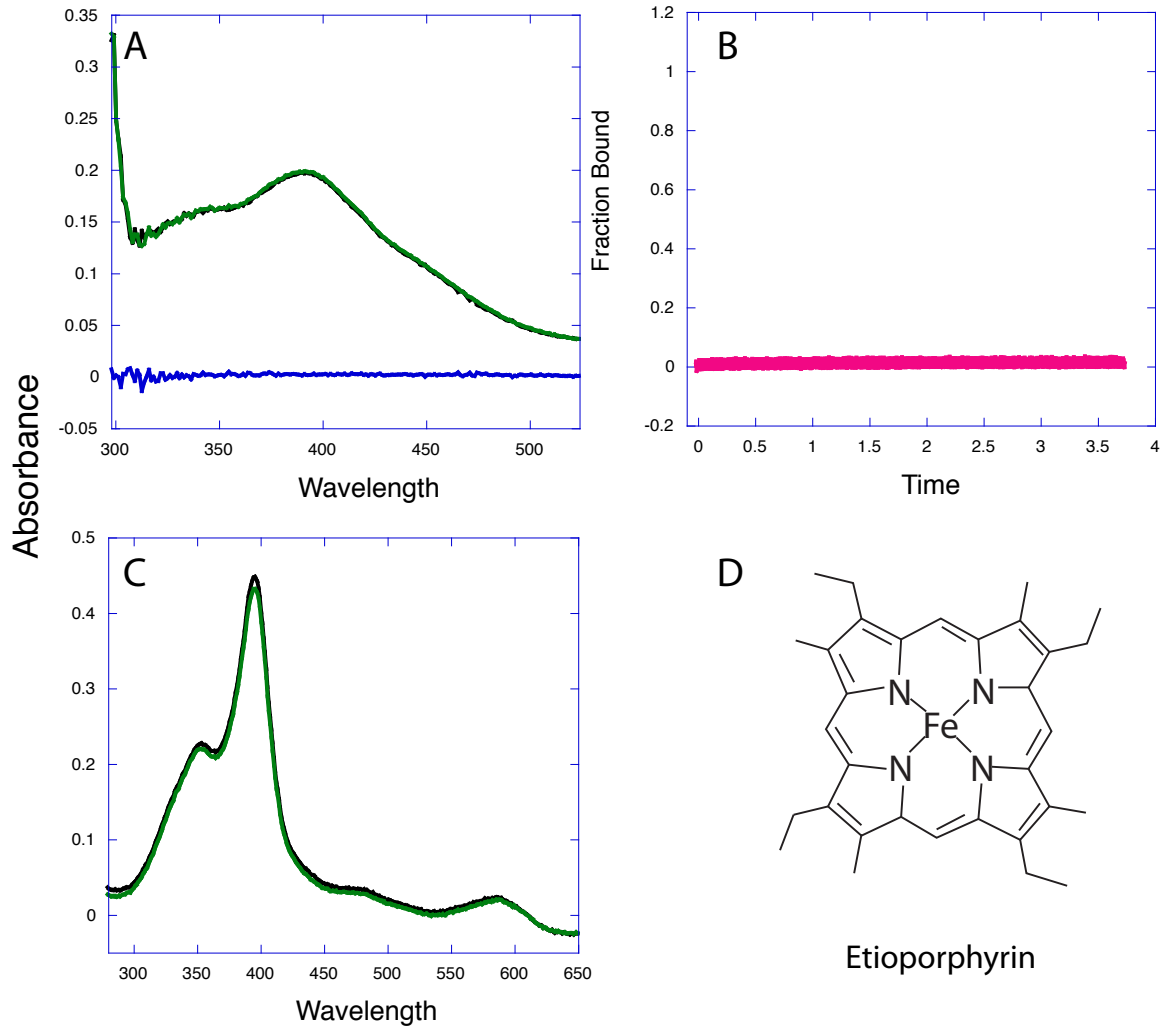


Figure A3.10: (A) Free Etioporphyrin in solution (black), Etioporphyrin bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Etioporphyrin assembling with the single chain protein. (C) Etioporphyrin in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Etioporphyrin. All wavelengths are in nm. The extinction coefficient at 395nm is $98,000 \text{ M}^{-1}\text{cm}^{-1}$.⁴

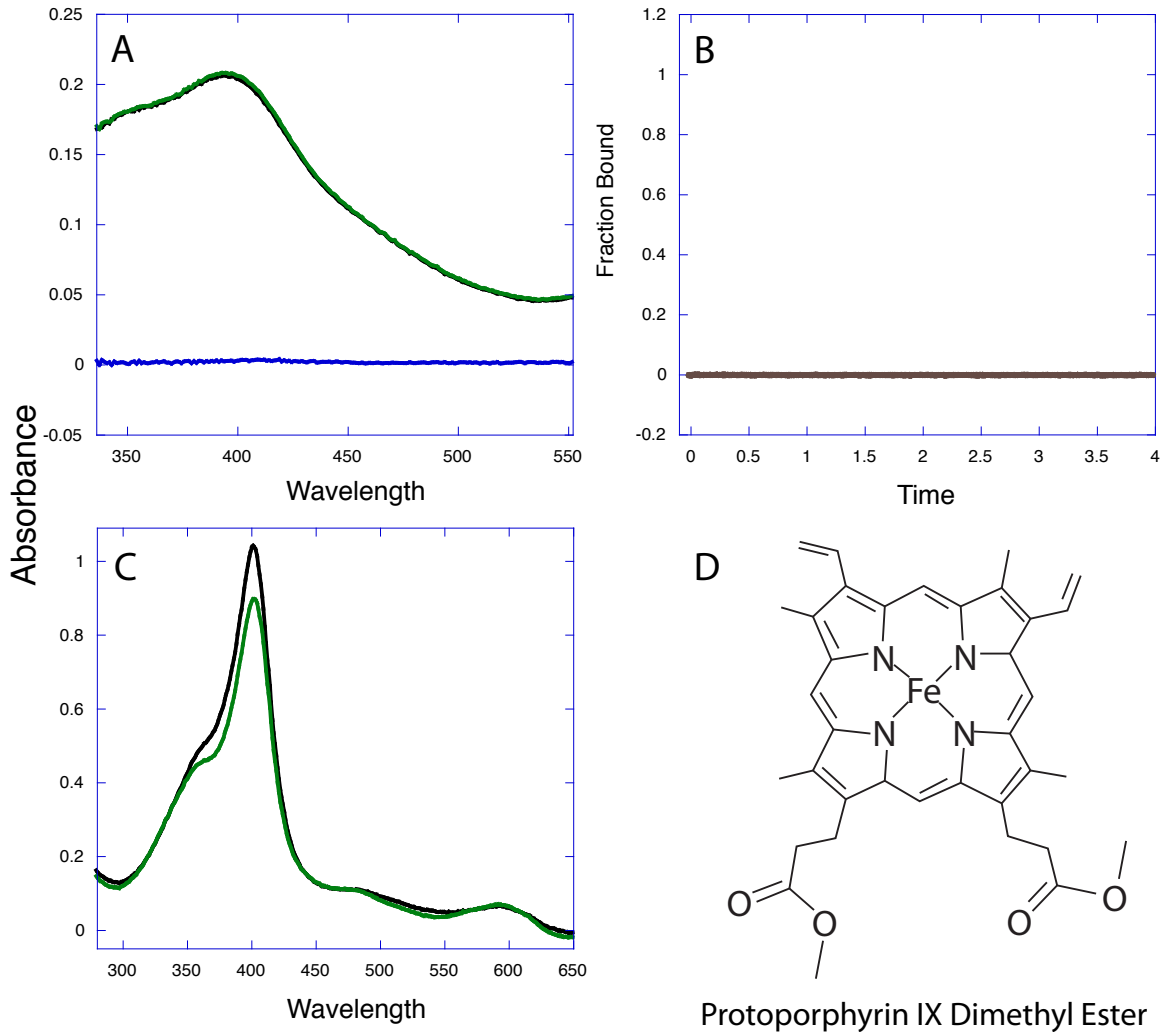


Figure A3.11: (A) Free Protoporphyrin IX Dimethyl Ester in solution (black), Protoporphyrin IX Dimethyl Ester bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Protoporphyrin IX Dimethyl Ester assembling with the single chain protein. (C) Protoporphyrin IX Dimethyl Ester in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Protoporphyrin IX Dimethyl Ester. All wavelengths are in nm. The extinction coefficient was assumed to be equal to that of Heme B.

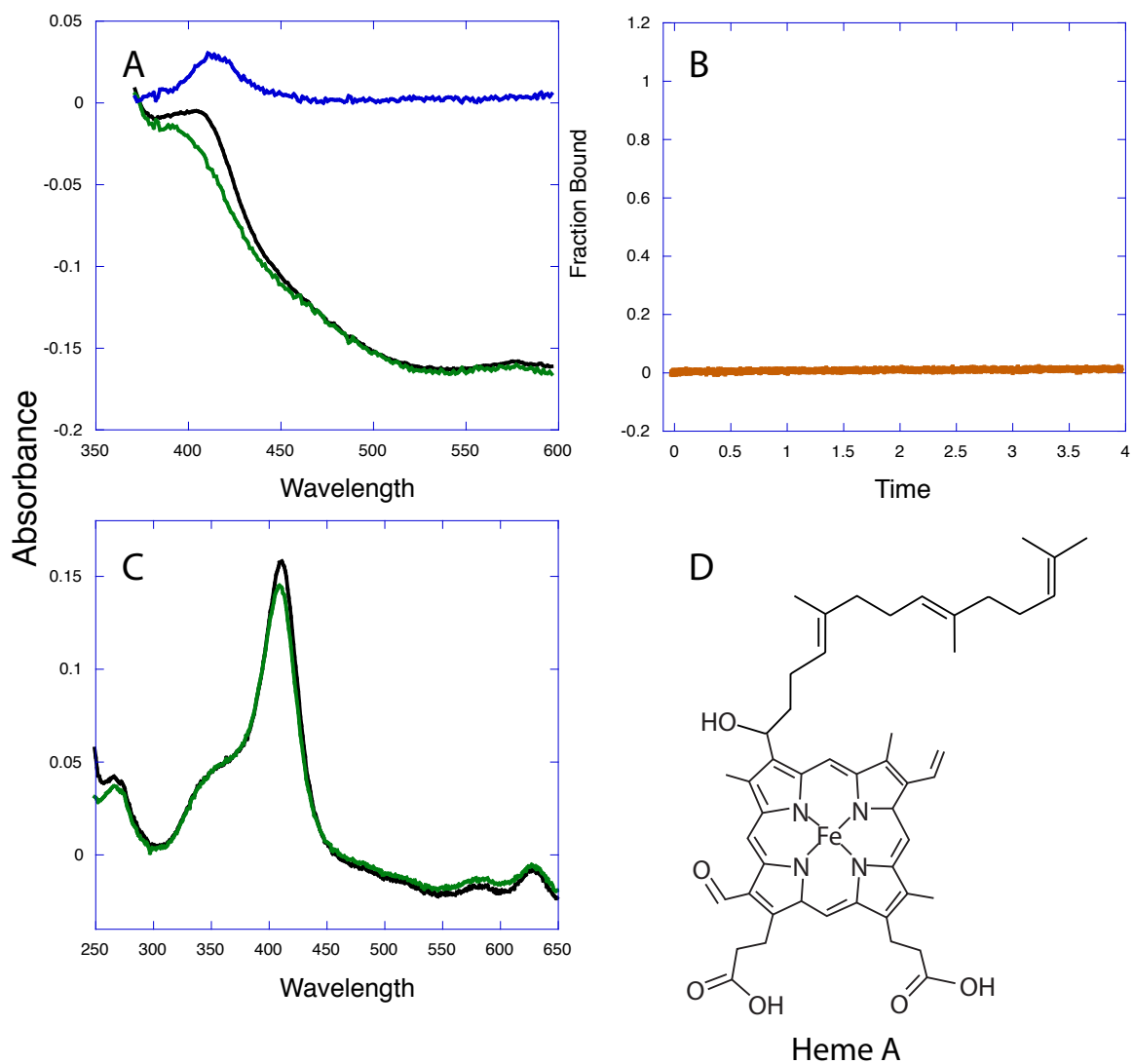


Figure A3.12: (A) Free Heme A in solution (black), Heme A bound to protein (green), and the difference spectrum (blue). (B) 4 second trace of Heme A assembling with the single chain protein. (C) Heme A in octanol (black) and octanol after mixing with 20mM CHES 150mM KCl buffer pH 9. (D) Chemical structure of Heme A. All wavelengths are in nm. The extinction coefficient at 420nm is $131,000 \text{ M}^{-1}\text{cm}^{-1}$.⁵

A3.3: Alternate Fe-porphyrin Redox Titrations

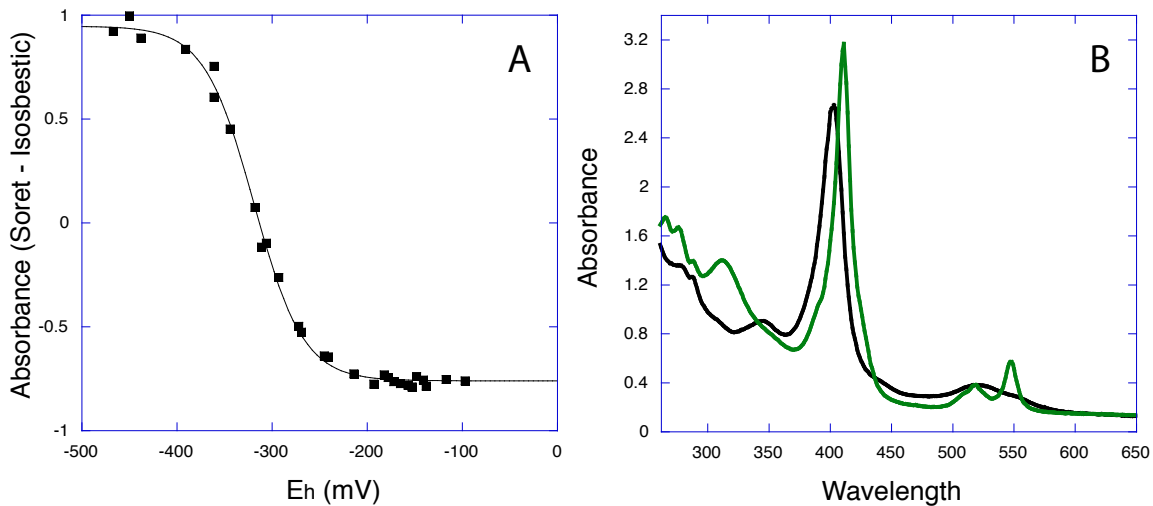


Figure A3.13: (A) Mesoporphyrin redox titration and (B) Oxidized (black) and reduced (green) spectra. Soret at 411nm and Isosbestic point at 404nm

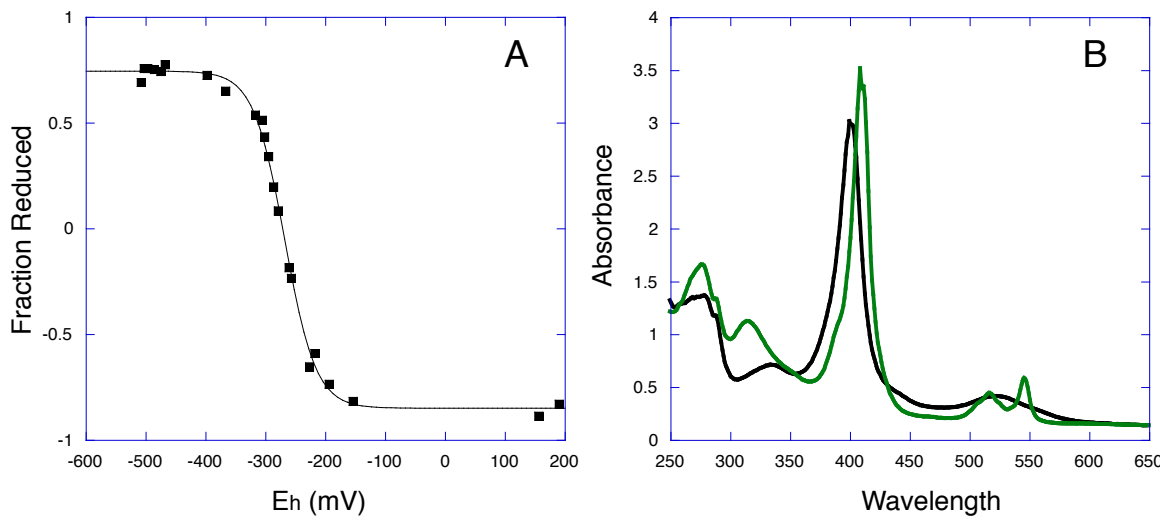


Figure A3.14: (A) Deuteroporphyrin redox titration and (B) Oxidized (black) and reduced (green) spectra. Fraction reduced calculated from ratio of 546nm/531nm

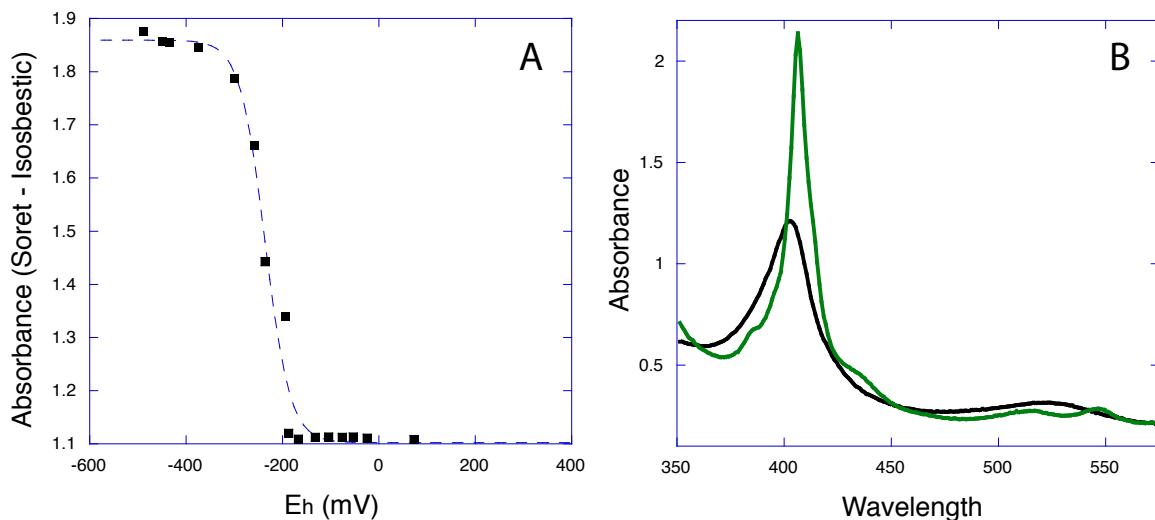


Figure A3.15: (A) IsoHematooporphyrin redox titration and (B) Oxidized (black) and reduced (green) spectra. Soret at 406nm and Isosbestic point at 452nm

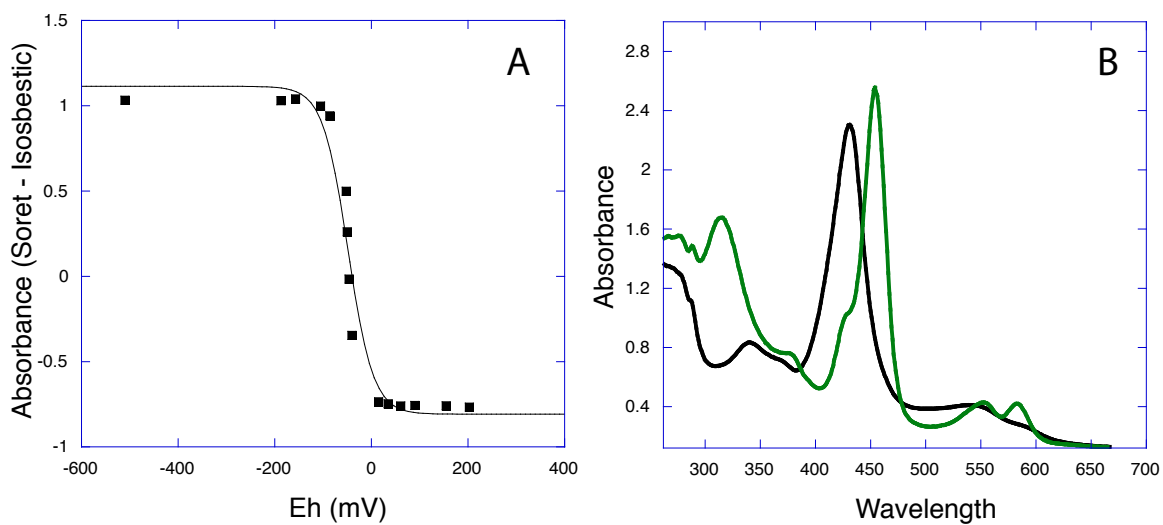


Figure A3.16: (A) Diacetyl Deuteroporphyrin redox titration and (B) Oxidized (black) and reduced (green) spectra. Soret at 454nm and Isosbestic point at 443nm

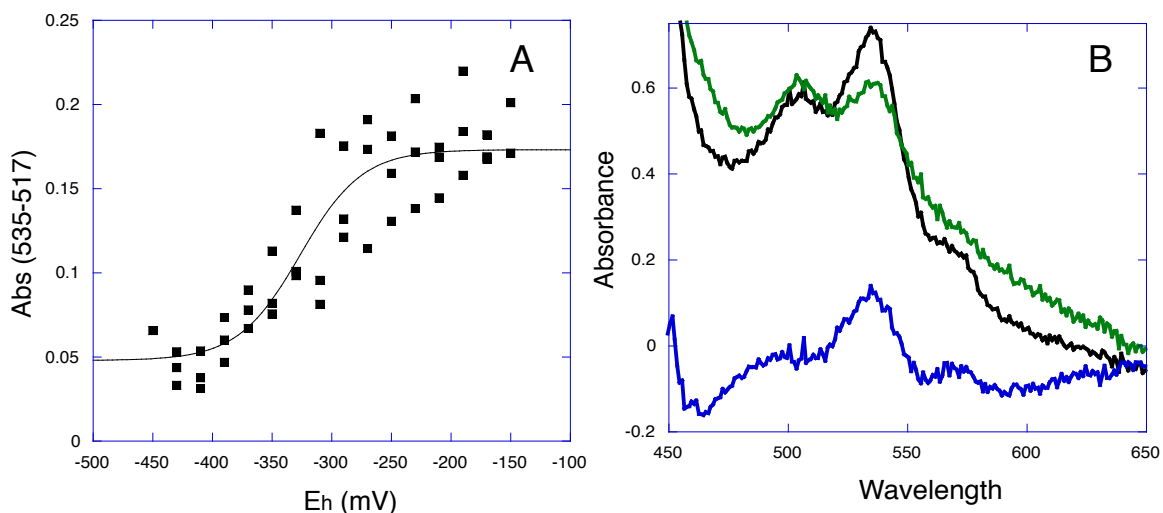


Figure A3.17: (A) Tetracarboxyphenylporphyrin Spectral-Electrochemical titration and (B) Oxidized (black) and reduced (green) spectra. Difference spectrum is in blue

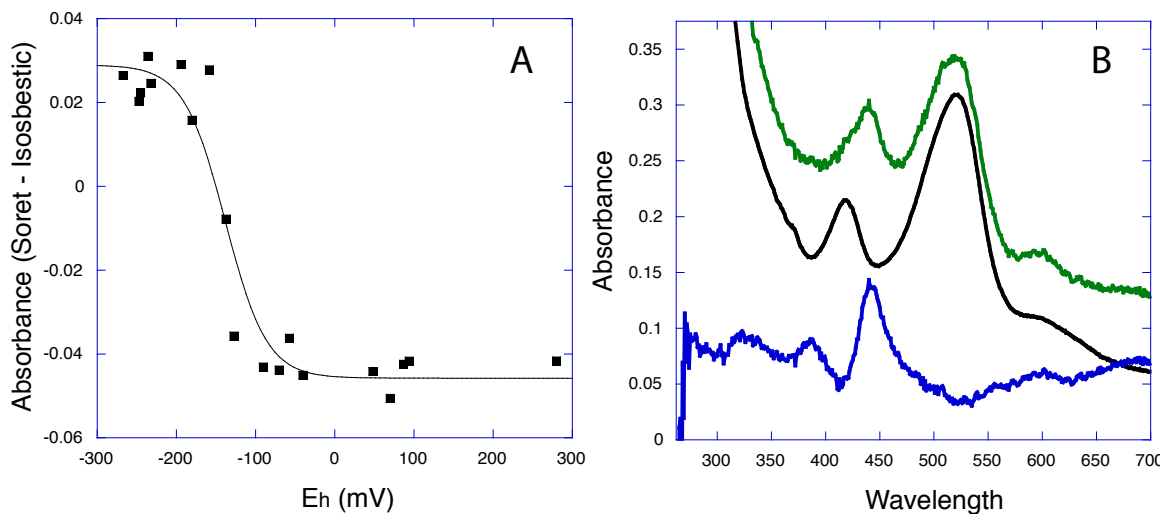


Figure A3.18: (A) Heme A redox titration and (B) Oxidized (black) and reduced (green) spectra. A difference spectrum (blue) was added for clarity. Soret at 443nm and Isosbestic point at 414nm

A3.4: Zn-Centered Chlorins And Porphyrins

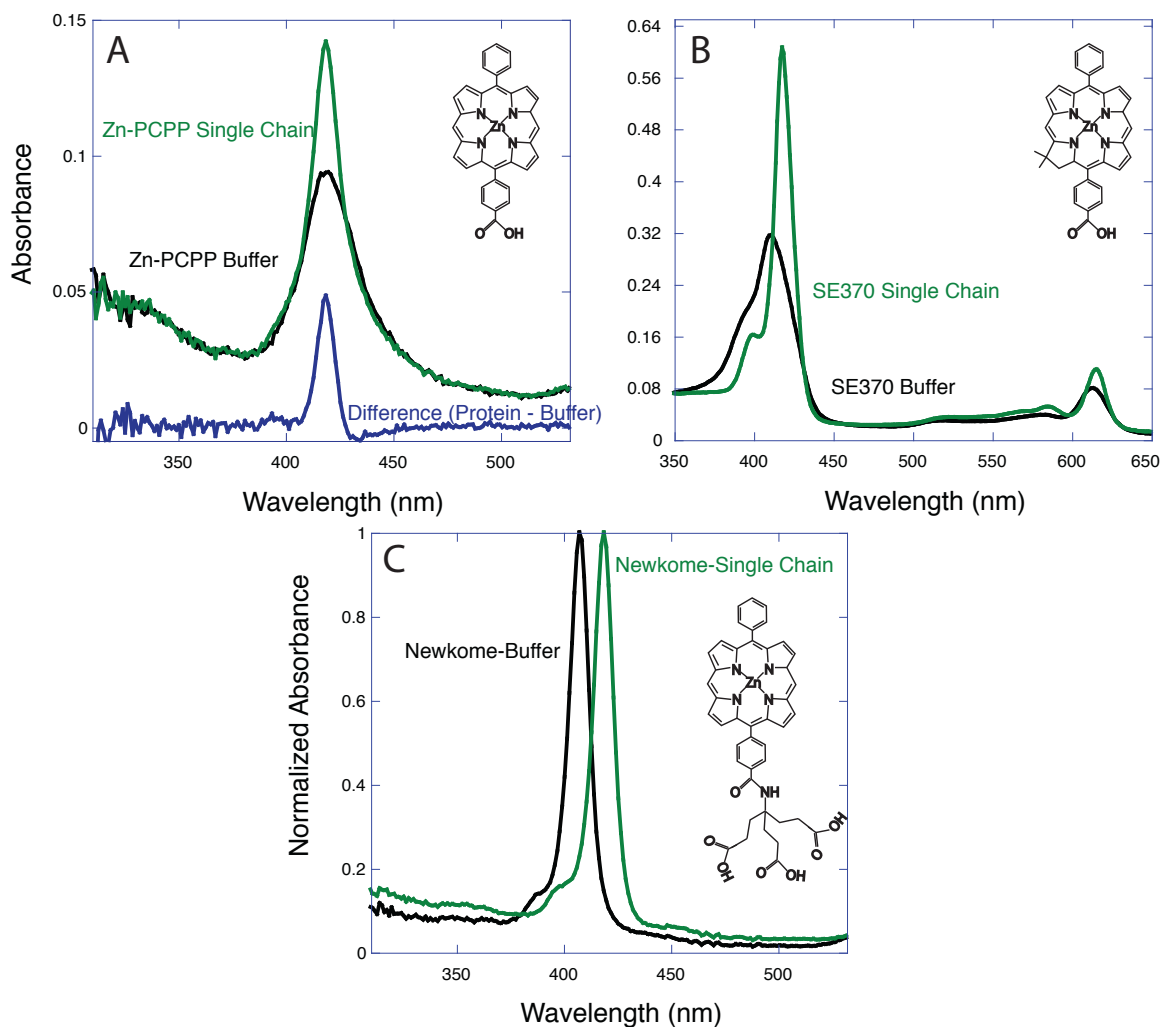


Figure A3.19: (A) Spectra of Zn-5-Phenyl-15-Carboxyphenyl Porphyrin (Zn-PCPP) in buffer (black), and bound to the Single Chain protein (green). A difference spectra has been added to better show the wavelength used for assembly studies (blue). This chlorin has been donated by the laboratory of Jonathan Lindsey at North Carolina State university. (B) Spectra of SE370 in aqueous buffer (black) and bound to the single chain protein (green). (C) Newkome Solubilized Zn-PCPP in aqueous buffer (black) and bound to the single chain protein (green). All spectra are in 20mM CHES 150mM KCl buffer. Wavelength in nm. Absorbance Normalized to Soret peak, where normalized.

A3.5: References

(1) Tammer A. Farid, G. K., Lee A. Solomon, Bruce R. Lichtenstein, Molly M. Sheehan, Bryan A. Fry, Chris Bialas, Nathan M. Ennist, Jessica A. Siedlecki, Zhenyu Zhao, Matthew A. Stetz, Kathleen G. Valentine, J. L. Ross Anderson, Bohdana M. Discher, A. Joshua Wand, Christopher C. Moser and P. Leslie Dutton¹ In *Nat Chem Biol* 2013; Vol. Submitted.

(2) Morgan, W. T.; Mullereb.U *J Biol Chem* **1972**, 247, 7181.

(3) Zhuang, J. Y.; Amoroso, J. H.; Kinloch, R.; Dawson, J. H.; Baldwin, M. J.; Gibney, B. R. *Inorg Chem* **2006**, 45, 4685.

(4) Reading, N. S.; Aust, S. D. *Arch Biochem Biophys* **1998**, 359, 291.

(5) Zhuang, J. Y.; Reddi, A. R.; Wang, Z. H.; Khodaverdian, B.; Hegg, E. L.; Gibney, B. R. *Biochemistry-U* **2006**, 45, 12530.