



Publicly Accessible Penn Dissertations


Spring 5-17-2010

A New Approach to Understanding Biological Control of Quinone Electrochemistry

Bruce R. Lichtenstein

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

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

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

Recommended Citation

Lichtenstein, Bruce R., "A New Approach to Understanding Biological Control of Quinone Electrochemistry" (2010). *Publicly Accessible Penn Dissertations*. 180.
<http://repository.upenn.edu/edissertations/180>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/180>
For more information, please contact libraryrepository@pobox.upenn.edu.

A New Approach to Understanding Biological Control of Quinone Electrochemistry

Abstract

Oxidoreductases play pivotal roles in energy capturing and converting processes of life. During these processes, quinones shuttle protons and reducing equivalents between membrane-bound oxidoreductases that generate the proton motive force during oxidative phosphorylation and photophosphorylation. A key mechanistic feature of these oxidoreductases is their ability to tune the reduction potentials of the hydroquinone, semiquinone and oxidized states of their substrate quinones. This level of control allows for maximization of conversion efficiency between the energy of the quinone reducing equivalents and the proton motive force, and prevents side reactions that may be fatal to cells. A half-century of experimental study and computational modeling of the respiratory and photosynthetic complexes has revealed little information on how this mechanistic control is accomplished. To obtain mechanistic insights into the control process, it is necessary to eliminate the biological complexity intrinsic to natural quinone oxidoreductases and create experimental systems that are simplified maquettes of quinone active sites. In this work, development of a naphthoquinone amino acid (Naq), modeled after vitamin K, allowed the creation of a range of quinone peptide maquettes designed to address uncertain mechanistic details of biological quinone control. In a simple heptamer, Naq acquires properties of quinone cofactors found in the three distinct classes of active sites of membrane oxidoreductases under different experimental conditions. Study of Naq in a lanthanide ion binding EF hand peptide allowed observation of the effect of a structural transition from coil to alpha-helix on the aqueous midpoint potential of Naq and measurement of the rate of electron transfer between reduced and oxidized Naq. Naq was also incorporated into a structured miniprotein based upon the TrpCage using a combination of the SCADS computational approach and iterative redesign by hand, creating a simple scaffold for evaluating effects of changing the local environment on Naq. Finally, using expressed protein ligation, Naq was incorporated into a single chain heme-binding maquette. Studies using this multi-cofactor protein to explore electron transfer reactions to and from Naq like those critical to respiration and photosynthesis are underway.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Biochemistry & Molecular Biophysics

First Advisor

P. Leslie Dutton

Keywords

Quinone, Naq, Maquette, Protein Design, Biophysics, Electrochemistry

Subject Categories

Biochemistry | Biophysics | Materials Chemistry | Organic Chemistry

A NEW APPROACH TO UNDERSTANDING BIOLOGICAL CONTROL
OF QUINONE ELECTROCHEMISTRY

Bruce R. Lichtenstein

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

2010

Supervisor of Dissertation:

Graduate Group Chair:

P. Leslie Dutton, Ph.D., Eldridge Reeves
Johnson Professor of Biochemistry and
Biophysics and Director of the Johnson
Foundation for Molecular Biophysics

Kathryn M. Ferguson, Ph.D., Associate
Professor of Physiology

Dissertation Committee:

William F. DeGrado, Ph.D., George W. Raiziss Professor of Biochemistry and Biophysics

Ivan J. Dmochowski, Ph.D., Assistant Professor of Chemistry

S. Walter Englander, Ph.D., Jacob Gershon-Cohen Professor of Medical Science and
Professor of Biochemistry and Biophysics

Feng Gai, Ph.D., Associate Professor of Chemistry

Jeffery G. Saven, Ph.D., Associate Professor of Chemistry

A. Joshua Wand, Ph.D., Benjamin Rush Professor of Biochemistry and Biophysics

A NEW APPROACH TO UNDERSTANDING
BIOLOGICAL CONTROL OF QUINONE
ELECTROCHEMISTRY

© 2010

Bruce Robert Lichtenstein

Dedicated to the memory of
Professor Ahamindra Jain who
taught us to approach questions
with our entire toolbox

Acknowledgements

The work described in this dissertation could not have been accomplished without the support and intellectual contributions of a large number of individuals. The text itself was read and edited by Vonni Moorman, who in all likelihood suffers nightmares from my butchering of the English language and grammar. Without her copyediting and not so subtle rallying cries of ‘get it done’ and ‘you can do this,’ I am unsure this beast would have been readable let alone complete in a timely fashion. Vonni’s contributions extend beyond the final phases of the dissertation, as her scientific efforts were integral in the NMR measurements described in Chapter 3. For all of her work and friendship, I am deeply indebted and grateful. Scientific contributions from José Cerda, including all electrochemical characterization described, and Tammer Farid and Ross Anderson for their work in establishing the utility and importance of the BTX family of proteins were integral to this work and I owe each of them more than a few beers, preferably in Brisol, UK where Ross is now a lecturer. I also cannot forget Seung-gu Kang who tirelessly pushed *SCADS* to do things with Naq that I wouldn’t have imagined possible even when I was somewhat remiss in updating him on the progress of our NaqCage project. Ron Koder’s input at the beginning stages of this project date from the day he offered me a job in Les Dutton’s laboratory while I was still at Swarthmore suffering through his seminar course. I would not have worked in Les’s lab had this introduction not been made, and Ron’s interests and intellectual contributions to the maquette work definitely shaded the early stages of the work described in this dissertation. My work and friendship with Ron is likely to continue indefinitely as he is a fount of ideas with powerful magnets under his control.

Ultimately this project could not have been complete without the efforts of my advisor Les and Chris Moser in supporting my wandering in this forest of my own creation and finding ways to keep me paid and the project funded. Their advise and encouragement over the years has proven helpful in furnishing this disseration. I am grateful for the years of support I have had in the Dutton laboratory and they were critical to each and every one of them.

Finally there are those without whom I could not have completed this project whose contributions were less tangible but no less appreciated. Anne Jones and Volker Benkert were wonderful friends during Anne's tenure as a post-doc in Les's lab. Scientific discussions with Anne and social-political discussion with Volker rounded out those years of my life while they were in Philadelphia. They made me a part of their family for those brief years, and I cannot begin to repay this kindness. Ross Anderson and Rachel Moffat were also wonderful friends (however Scottish) both encouraging me to get out of my apartment and exposing me to more theater produced by Swarthmore alums than I thought reasonably possible. I owe almost the entirety of my social existence to the two of them while they were here. I must thank Joe Kielec for providing decoration for my overly messy desk and while I strongly disagree with his allegiance to the ninjas, he fought admirably against the democratic pirate tide. Bohdana Discher has been in Les's lab since before I arrived and has tolerated many many late afternoon discussions with me about various topics. Her orthogonal perspective on the work described in this dissertation always proved enlightening, and I am in her debt. Fellow graduate students in the Dutton Lab, Tammer Farid, Sarah Chobot Hokanson and Lee Solomon, amongst others, provided not only good friendships over the years, but provided constant reminders that I should have been writing when instead I

wanted to play some more in lab. It has also been great to work in the lab with post-docs José Cerda, Ron Koder, Audria Stubna and Doug Marshall all of whom approach the subject of science from very different perspectives but with much passion and with whom conversations helped shape my days. I cannot have survived graduate school without the friendships of these and many other individuals. To my friends from former lives Seeta, Robyn, Dan, Imo, Marah, Peter and Brianne, your support over the years has been invaluable.

I of course would not be here at all if it weren't for my family from my parents who reproduced all those thirty years ago to my siblings who provided some of the rich tapestry of my childhood and more recent appreciation of almost all things Lichtenstein – notably not the kidney stones. I cannot say enough about the passion and intellectual curiosity fostered in that household, and even though it seems that my life has been guided by those eager to get out of my way and let me experience life as I am to experience it, the support I was provided by my family is incalculable.

Lastly, I would like to thank my Dissertation Committee (Josh Wand, Bill DeGrado, Walter Englander, Feng Gai, Ivan Dmochowski, Jeff Saven) for reading and providing constructive criticism of this dissertation.

ABSTRACT

A NEW APPROACH TO UNDERSTANDING BIOLOGICAL CONTROL OF QUINONE

ELECTROCHEMISTRY

Bruce R. Lichtenstein

Advisor: P. Leslie Dutton

Oxidoreductases play pivotal roles in energy capturing and converting processes of life. During these processes, quinones shuttle protons and reducing equivalents between membrane-bound oxidoreductases that generate the proton motive force during oxidative phosphorylation and photophosphorylation. A key mechanistic feature of these oxidoreductases is their ability to tune the reduction potentials of the hydroquinone, semiquinone and oxidized states of their substrate quinones. This level of control allows for maximization of conversion efficiency between the energy of the quinone reducing equivalents and the proton motive force, and prevents side reactions that may be fatal to cells. A half-century of experimental study and computational modeling of the respiratory and photosynthetic complexes has revealed little information on how this mechanistic control is accomplished. To obtain mechanistic insights into the control process, it is necessary to eliminate the biological complexity intrinsic to natural quinone oxidoreductases and create experimental systems that are simplified maquettes of quinone active sites. In this work, development of a naphthoquinone amino acid (Naq), modeled after vitamin K, allowed the creation of a range of quinone peptide maquettes designed to address uncertain mechanistic details of biological quinone control. In a simple heptamer, Naq acquires properties of quinone cofactors found in the three distinct classes of active sites of membrane oxidoreductases under different experimental conditions. Study of Naq in a

lanthanide ion binding EF hand peptide allowed observation of the effect of a structural transition from coil to alpha-helix on the aqueous midpoint potential of Naq and measurement of the rate of electron transfer between reduced and oxidized Naq. Naq was also incorporated into a structured miniprotein based upon the TrpCage using a combination of the *SCADS* computational approach and iterative redesign by hand, creating a simple scaffold for evaluating effects of changing the local environment on Naq. Finally, using expressed protein ligation, Naq was incorporated into a single chain heme-binding maquette. Studies using this multi-cofactor protein to explore electron transfer reactions to and from Naq like those critical to respiration and photosynthesis are underway.

Table of Contents

| | |
|---|------|
| Acknowledgements | iv |
| Abstract | vii |
| List of Tables | xiii |
| List of Figures | xiv |
| Chapter 1: Introduction | 1 |
| 1.1 Quinones | 1 |
| 1.1.1 <i>Quinone Electrochemistry</i> | 4 |
| 1.2 Biological Control of Quinone Electrochemistry | 5 |
| 1.3 References | 8 |
| Chapter 2: Design, Development and Exploration of a New Naphthoquinone Cofactor | 11 |
| 2.1 Introduction | 11 |
| 2.2 Design of a New Quinone Cofactor | 16 |
| 2.3 Synthesis of Naq, the Naphthoquinone Amino Acid | 20 |
| 2.3.1 <i>Side Chain Protecting Groups</i> | 21 |
| 2.3.2 <i>Completion of the synthesis of Naq</i> | 28 |
| 2.4 Initial Electrochemical Studies of Naq in an Ultra-Short Peptide | 32 |
| 2.5 HeptaNaq | 34 |
| 2.5.1 <i>HeptaNaq(OMe/OMe) Spectroscopy and Fluorescence</i> | 35 |
| 2.5.2 <i>HeptaNaq(OMe/OMe) Deprotection</i> | 36 |
| 2.5.3 <i>HeptaNaq Aqueous Spectroscopy and Potentiometry</i> | 38 |
| 2.5.4 <i>HeptaNaq Aprotic Cyclic Voltammetry</i> | 41 |
| 2.6 Conclusion | 43 |
| 2.7 References | 45 |
| Chapter 3: Secondary Structure Effects on the Electrochemistry of the Naphthoquinone Amino Acid, Naq | 47 |
| 3.1 Introduction | 47 |
| 3.2 P2Naq Thermodynamics | 53 |
| 3.3 Synthesis, Deprotection and Initial Characterization of P2Naq | 56 |
| 3.3.1 <i>Deprotection/Activation of P2Naq(OMe/OMe)</i> | 57 |
| 3.3.2 <i>Spectroscopy of P2Naq</i> | 61 |
| 3.3.3 <i>Borohydride Oxidase Activity of P2Naq</i> | 62 |

| | |
|--|-----|
| 3.4 Synthesis of Nitrogen Labeled Peptides for Site-Specific Helicity by NMR | 62 |
| 3.4.1 <i>Synthesis of ¹⁵N labeled Fmoc-Naq(OMe/OMe)-OH</i> | 63 |
| 3.4.2 <i>Synthesis of ¹⁵N labeled P2Naq(OMe/OMe), Random Coil Control and Helical Control</i> | 64 |
| 3.5 Circular Dichroism of P2Naq, RCCNaq, and P2NaqExt | 66 |
| 3.5.1 <i>Circular Dichroism of P2Naq_{ox} and La³⁺- Bound P2Naq_{ox}</i> | 67 |
| 3.5.2 <i>Circular Dichroism of P2Naq_{red} and La³⁺- Bound P2Naq_{red}</i> | 68 |
| 3.5.3 <i>Circular Dichroism of RCCNaq_{ox} and RCCNaq_{red}</i> | 70 |
| 3.5.4 <i>Circular Dichroism of P2NaqExt_{ox} and P2NaqExt_{red}</i> | 71 |
| 3.5.5 <i>Naq specific CD signals in P2Naq and P2NaqExt</i> | 73 |
| 3.6 Measurement of Lanthanide Ion Binding Affinity of P2Naq_{ox} and P2Naq_{red} | 75 |
| 3.6.1 <i>Lanthanide Binding of P2Naq_{ox}</i> | 77 |
| 3.6.2 <i>Lanthanide Binding of P2Naq_{red}</i> | 77 |
| 3.6.3 <i>Completing the Thermodynamic Box</i> | 82 |
| 3.7 Backbone Chemical Shift Determination for P2Naq, RCCNaq, and P2NaqExt | 83 |
| 3.7.1 <i>Chemical Shifts</i> | 86 |
| 3.7.2 <i>Quantitative J-Correlation and Measures of ϕ</i> | 88 |
| 3.7.3 <i>Walking the Backbone with a ¹⁵N Filtered NOESY</i> | 91 |
| 3.7.4 <i>Measuring Electron Transfer Rates Between Lanthanide-Bound P2Naq_{ox} and P2Naq_{red}</i> | 92 |
| 3.8 Measurement of the Change in Midpoint of Naq from Coil to Helix | 96 |
| 3.9 Conclusion | 99 |
| 3.10 References | 101 |

Chapter 4: Differentiating Oxidized and Reduced Naq in Designed

| | |
|--|-----|
| Protein | 103 |
| 4.1 Introduction | 103 |
| 4.2 SCADS Redesign of ^{wt}TrpCage into Redox State Specific NaqCages | 107 |
| 4.2.1 <i>Optimizing ^{wt}TrpCage Scaffold</i> | 108 |
| 4.2.2 <i>Incorporating Naq into SCADS Amino Acid Library</i> | 108 |
| 4.2.3 <i>Identification of Mutation Sites</i> | 111 |
| 4.3 Synthesis and Redesign of ^{ox}NaqCage | 114 |
| 4.3.1 <i>Synthesis and Deprotection of ^{ox}NaqCage Target Sequences</i> | 115 |
| 4.3.2 <i>Iterative Redesign of ^{ox}NaqCage</i> | 116 |
| 4.4 NMR of ^{ox}NaqCage 7c | 121 |
| 4.5 Conclusion | 126 |
| 4.6 References | 128 |

Chapter 5: Design and Properties of a Single-Chain Four-Helix Bundle

| | |
|---------------------------|-----|
| Maquette | 130 |
| 5.1 Introduction | 130 |
| 5.2 Design of BT3 | 136 |
| 5.2.1 <i>N-Cap design</i> | 137 |

| | |
|---|-----|
| 5.2.2 Cross Bundle Loop Design | 138 |
| 5.2.3 Addition of an N-Terminal Methionine | 139 |
| 5.2.3 Breaking the Symmetry of HP7 | 139 |
| 5.3 Confirmation of General Properties of Heme Binding Maquette BT3 | 140 |
| 5.3.1 Heme Binding | 141 |
| 5.3.2 Molecular Oxygen Binding | 142 |
| 5.3.3 Redesign | 142 |
| 5.4 cBT3 | 144 |
| 5.4.1 In vivo c-Type Maturation | 145 |
| 5.4.2 Potentiometry and Molecular Oxygen Binding | 146 |
| 5.5 Conclusion | 147 |
| 5.6 References | 148 |
| | |
| Chapter 6: Incorporation and Deprotection of Naq in a Four-Helix Bundle Maquette | 149 |
| 6.1 Introduction | 149 |
| 6.2 Building a Maquette Competent Expressed Protein Ligation Vector | 151 |
| 6.2.1 Constructing a BT3-Intein Expression System | 154 |
| 6.2.2 Chemical Ligation of BTX ³ with Synthetic Helices | 156 |
| 6.2.3 Deprotecting Naq(OMe/OMe) in a Four Helix Bundle | 158 |
| 6.3 Reductive Deprotection of Naq in a Four-Helix Bundle | 159 |
| 6.3.1 Synthesizing Fmoc-Naq(OProp/OProp)-OH | 160 |
| 6.3.2 Incorporation of Naq(OProp/OProp) into Single Helix BT0.4 | 161 |
| 6.3.3 Deprotection of Naq(OProp/OProp) in BT0.4 | 162 |
| 6.3.4 Incorporation and Deprotection of Naq(OProp/OProp) in a Four Helix Bundle | 164 |
| 6.4 Future Directions | 166 |
| 6.4 References | 168 |
| | |
| Chapter 7: Conclusion | 170 |
| 7.1 References | 174 |
| | |
| Appendix A: Experimental Methods | 175 |
| A.1 Introduction | 175 |
| A.2 General Methods | 175 |
| A.3 Naq Synthetic Methods | 177 |
| A.3.1 Methyl 1-hydroxy-4-methoxy-2-naphthoate (1) | 177 |
| A.3.2 Methyl 1,4-dimethoxy-2-naphthoate (2) | 178 |
| A.3.3 1,4-Dimethoxy-2-hydroxymethylnaphthalene | 179 |
| A.3.4 1,4-Dimethoxy-2-bromomethylnaphthalene (3) | 180 |
| A.3.5 N-(Diphenylmethylene)-L-1,4-dimethoxy-2-naphthalanine tertbutyl ester (4) | 178 |
| A.3.6 L-1,4-dimethoxy-2-naphthalanine hydrochloride (6) | 182 |
| A.3.7 Fmoc-Naq(OMe/OMe)-OH (7) | 183 |

| | |
|--|-----|
| A.4 Naq Deprotection Methods | 184 |
| A.4.1 Deprotection of heptaNaq(OMe/OMe) with Cerium Ammonium Nitrate | 184 |
| A.4.2 Deprotection of P2Naq(OMe/OMe) with Dichlorodicyanoquinone (DDQ) | 184 |
| A.4.3 Deprotection of Naq(OProp/OProp) in BT6.4 | 185 |
| A.5 Uncommon Characterization Methods | 185 |
| A.5.1 Determination of Enantiomeric Excess of the Synthesis of Naq | 185 |
| A.5.2 Determination of Molar Absorptivity of heptaNaq | 186 |
| A.5.3 Spectroelectrochemistry | 186 |
| A.5.4 New Spectroelectrochemical Cell | 188 |
| A.5.5 Cyclic Voltammetry of heptaNaq in DMF | 189 |
| A.6 Chemical Ligation Methods | 189 |
| A.7 Iron(III) 2,4-Dicyanodeuteroporphyrin IX Synthetic Methods | 190 |
| A.7.1 Protoporphyrin IX dimethylester | 191 |
| A.7.2 2,4-Diformylprotoporphyrin IX dimethyl ester | 192 |
| A.7.3 2,4-Dicyanoporphyrin IX dimethyl ester | 193 |
| A.7.4 Iron(III) 2,4-Dicyanoporphyrin IX dimethyl ester | 194 |
| A.7.5 Iron(III) 2,4-Dicyanoporphyrin IX | 195 |
| A.8 References | 196 |

List of Tables

Chapter 3. Secondary Structure Effects on the Electrochemistry of the Naphthoquinone Amino Acid, Naq

| | |
|---|----|
| Table 3.1. Primary sequence of peptides under study | 65 |
| Table 3.2. Chemical shifts of backbone atoms of ^{15}N labeled Ala ₁₃ , Naq ₁₄ and Ala ₁₅ in P2Naq, RCCNaq and P2NaqExt | 86 |
| Table 3.3. Backbone $J_{\text{H}\alpha\text{H}\text{N}}$ scalar coupling constants and ϕ | 90 |
| Table 3.4 Computed fractional helicities. | 96 |

Chapter 4: Differentiating Oxidized and Reduced Naq in Designed Protein

| | |
|---|-----|
| Table 4.1. Sequence for the $^{\text{wt}}$ TrpCage and SCADS target sequences for $^{\text{ox}}$ NaqCage and $^{\text{red}}$ NaqCage miniproteins | 114 |
| Table 4.2. Sequences for the $^{\text{wt}}$ TrpCage, optimized $^{\text{mut}}$ TrpCage, initial SCADS target sequence 1 for $^{\text{ox}}$ NaqCage, and sequences of iteratively redesigned $^{\text{ox}}$ NaqCage | 115 |
| Table 4.3. Backbone assignments | 125 |

Chapter 5: Design and Properties of a Single-Chain Four-Helix Bundle Maquette

| | |
|---|-----|
| Table 5.1. Sequences of maquettes discussed in text showing evolution of design | 134 |
|---|-----|

Chapter 6: Incorporation and Deprotection of Naq in a Four-Helix Bundle Maquette

| | |
|--|-----|
| Table 6.1. Sequences of maquettes discussed in text | 153 |
| Table 6.2. Manipulation of the expressed BT3 ³ intein fusion protein to obtain soluble expression | 154 |

List of Figures

Chapter 1: Introduction

| | |
|---|---|
| <i>Figure 1.1. Genetic para- and ortho-quinones showing two electron, two-proton redox chemistry common at physiological pH</i> | 2 |
| <i>Figure 1.2. Natural quinone cofactors</i> | 3 |
| <i>Figure 1.3. The nine redox states of quinones shown with 2-methyl-1,4-naphthoquinone's redox couples</i> | 4 |
| <i>Figure 1.4. Natural quinone oxidoreductases classes</i> | 5 |

Chapter 2: Design, Development and Exploration of a New Naphthoquinone Cofactor

Figures

| | |
|---|----|
| <i>Figure 2.1. Naturally occurring quinone cofactors derived from amino acids</i> | 13 |
| <i>Figure 2.2. Membrane soluble para-quinone cofactors used in respiration and photosynthesis</i> | 17 |
| <i>Figure 2.3. Possible backbone connectivity regiomers of the naphthoquinone amino acid</i> | 18 |
| <i>Figure 2.4. 1,4-Dihydroxy-2-bromomethylnaphthalene</i> | 21 |
| <i>Figure 2.5. Final desired structure of Fmoc-Naq(OMe/OMe)-OH</i> | 28 |
| <i>Figure 2.6. Cyclic Voltammetry of Ac-Naq-OMe</i> | 33 |
| <i>Figure 2.7. Cartoon of oxidized heptaNaq in poly-proline II conformation</i> | 34 |
| <i>Figure 2.8. Average molar absorbance spectrum of heptaNaq(OMe/OMe)</i> | 35 |
| <i>Figure 2.9. Absorbance spectra of 1,4-naphthoquinone and heptaNaq</i> | 37 |
| <i>Figure 2.10. Reduced and oxidized spectra of heptaNaq in the presence of mediators</i> | 39 |
| <i>Figure 2.11. Oxidation-reduction chemistry of heptaNaq</i> | 40 |
| <i>Figure 2.12. Cyclic voltammetry of heptaNaq in DMF</i> | 42 |
| <i>Figure 2.13. Absorption spectrum and structure of 2-hydroxy-3-methyl-1,4-naphthoquinone</i> | 44 |

Schemes

| | |
|---|----|
| <i>Scheme 2.1. Covalent modification of peptide-incorporated cysteine by 1,4-naphthoquinone</i> | 15 |
| <i>Scheme 2.2. Generic electrophilic aromatic amino acid synthesis</i> | 20 |
| <i>Scheme 2.3. Synthesis of a naphthoquinone derived aryl-alkyl bromide</i> | 22 |
| <i>Scheme 2.4. Loss of naphthoquinone protecting groups upon oxidation</i> | 23 |
| <i>Scheme 2.5. Example synthesis of a protected hydroxymethyl naphthoquinone</i> | 25 |
| <i>Scheme 2.6. Incomplete protection of 1,4-dihydroxynaphthoic acid</i> | 25 |
| <i>Scheme 2.7. Progress towards a reductive route to a stably protected naphthoquinone aryl alkyl bromide</i> | 26 |

| | |
|--|----|
| <i>Scheme 2.8. Successful generation of naphthoquinone aryl-alkyl bromide desired for amino acid synthesis</i> | 27 |
| <i>Scheme 2.9. Creation of C^α-C^β bond in naphthoquinone amino acid</i> | 29 |
| <i>Scheme 2.10. Acidic deprotection of naphthoquinone amino acid backbone in methanol</i> | 29 |
| <i>Scheme 2.11. Final synthetic steps towards Fmoc-Naq(OMe/OMe)-OH.</i> | 30 |
| <i>Scheme 2.12. Synthesis of Ac-Naq-OMe</i> | 32 |

Chapter 3. Secondary Structure Effects on the Electrochemistry of the Naphthoquinone Amino Acid, Naq

Figures

| | |
|--|----|
| <i>Figure 3.1. Likely helix rotamers of Naq</i> | 48 |
| <i>Figure 3.2. Representative solution structure of P2A</i> | 49 |
| <i>Figure 3.3. The lanthanide binding coupled electrochemistry of P2Naq</i> | 53 |
| <i>Figure 3.4. Titration of P2Naq(OMe/OMe) with substoichiometric aliquots of PIFA</i> | 57 |
| <i>Figure 3.5. Molar absorptivity of P2Naq in 8 M GdnHCl</i> | 61 |
| <i>Figure 3.6. Borohydride reduction of P2Naq in the presence of oxygen</i> | 62 |
| <i>Figure 3.7. CD spectra of apo and holo P2Naq_{ox}</i> | 68 |
| <i>Figure 3.8. CD spectra of apo and holo P2Naq_{red}</i> | 69 |
| <i>Figure 3.9. CD spectra of RCCNaq_{ox} and RCCNaq_{red}</i> | 70 |
| <i>Figure 3.10. CD spectra of P2NaqExt_{ox} and P2NaqExt_{red}</i> | 71 |
| <i>Figure 3.11. Naq chromophore specific molar ellipticities in P2Naq</i> | 74 |
| <i>Figure 3.12. Simultaneous fit of five identical lanthanide ion titrations of P2Naq_{ox}</i> | 77 |
| <i>Figure 3.13. P2A titration under reducing conditions</i> | 79 |
| <i>Figure 3.14. Spectroelectrochemical titration of both apo and holo P2Naq</i> | 82 |
| <i>Figure 3.15. Backbone ¹⁵N labeling of the amino acid sequence A₁₃ Q₁₄ A₁₅</i> | 83 |
| <i>Figure 3.16. Graphical representation of the H^α chemical shift data in Table 3.2</i> | 87 |
| <i>Figure 3.17. Karplus curve</i> | 88 |
| <i>Figure 3.18. ¹⁵N-filtered HNHA of RCCNaq_{ox}</i> | 89 |
| <i>Figure 3.19. ¹⁵N-filtered NOESY spectra of lanthanide bound P2Naq_{ox} and P2NaqExt_{ox}</i> | 91 |
| <i>Figure 3.20. Chemical exchange in mix redox state sample of holo P2Naq</i> | 92 |
| <i>Figure 3.21. Fractional helicities calculated using CD and NMR data</i> | 97 |

Schemes

| | |
|--|----|
| <i>Scheme 3.1. Backbone cyclization of Naq when treated with TMSI</i> | 59 |
| <i>Scheme 3.2. Synthesis route to ¹⁵N labeled intermediate needed to make backbone labeled Naq.</i> | 64 |
| <i>Scheme 3.3. The major route of equilibration between quinones and hydroquinones in aqueous media</i> | 94 |

Chapter 4: Differentiating Oxidized and Reduced Naq in Designed Protein

Figures

| | |
|--|-----|
| <i>Figure 4.1. The effect of progressively and independently designing proteins to stabilize only one of two different redox states of Naq</i> | 104 |
| <i>Figure 4.2. A representative NMR structure of the TrpCage</i> | 106 |
| <i>Figure 4.3. Optimized scaffold for SCADS</i> | 109 |
| <i>Figure 4.4. Potential energy surface of side chain dihedral angles of Ac-Naq_{ox}-NHMe and Ac-Naq_{red}-NHMe</i> | 111 |
| <i>Figure 4.5. Identification of SCADS mutation site</i> | 113 |
| <i>Figure 4.6. CD spectral data of the iteratively redesigned ^{ox}NaqCages</i> | 116 |
| <i>Figure 4.7. H^α region of natural abundance ¹³C-HSQC and ¹H TOCSY spectra of ^{ox}NaqCage 7c</i> | 123 |

Chapter 5: Design and Properties of a Single-Chain Four-Helix Bundle

Maquette

Figures

| | |
|--|-----|
| <i>Figure 5.1. Crystal Structure of apo H10H24-L6I-L13F-L31M</i> | 132 |
| <i>Figure 5.2. Protein topologies of multichain heme-binding maquettes HP1 and HP7, and single chain heme-binding maquette BT6</i> | 135 |
| <i>Figure 5.3. Iron(III) protoporphyrin IX (PPIX) endpoint titrations of BT3</i> | 141 |
| <i>Figure 5.4. Formation of oxy-ferrous state in hemin bound BT3 in a stopped-flow experiment</i> | 142 |
| <i>Figure 5.5. 'Two-heme' binding restored in the BT6 single chain heme-binding maquette</i> | 143 |
| <i>Figure 5.6. Spectra of ferric Heme A, Heme B and Heme C bound to protein maquettes</i> | 145 |
| <i>Figure 5.7. Potentiometry of cBT3</i> | 146 |
| <i>Figure 5.8. Formation of oxy-ferrous cBT3 in a stopped-flow experiment</i> | 147 |

Chapter 6: Incorporation and Deprotection of Naq in a Four-Helix Bundle Maquette

Figures

| | |
|---|-----|
| <i>Figure 6.1. Diagram of possible intein joint sites in BT3</i> | 151 |
| <i>Figure 6.2. SDS-Page protein gel of BT6³ intein construct from expression to formation of BT6³-mensa thioester</i> | 156 |
| <i>Figure 6.3. UV/vis spectrum of BT5.3 incorporating Naq(OMe/OMe)</i> | 157 |
| <i>Figure 6.4. Energetically minimized structure of HP7 based upon the crystal structure of H10H24-L6I-L13F-L31M</i> | 161 |

| | |
|---|-----|
| <i>Figure 6.5. Spectrum of BT6.4 after reductive deprotection of Naq(OProp/OProp)</i> | 164 |
| <i>Figure 6.6. Formation of carboxy-ferrous complex of heme bound BT6.4 upon treatment with borohydride</i> | 166 |
| <i>Figure 6.7. Schematic of CO-flash photolysis of a single heme bound BT6.4-like maquette incorporating an activated Naq</i> | 167 |
| Schemes | |
| <i>Scheme 6.1. Native chemical ligation between two peptide fragments</i> | 150 |
| <i>Scheme 6.2. Synthesis of Fmoc-Naq(OProp/OProp)-OH</i> | 160 |
| Appendix A: Experimental Methods | |
| Figures | |
| <i>Figure A.1. Schematic of spectroelectrochemical cell built by José Cerda</i> | 186 |
| <i>Figure A.2. New spectroelectrochemical cell.</i> | 188 |

Chapter 1: Introduction

1.1 Quinones

All known life depends upon the generation of a chemical gradient across a membrane. Chemiosmosis of protons across membranes in every known cellular life form is used to drive the production of ATP and all of the cellular functions dependent upon it. Prior to the chemiosmotic hypothesis, it was generally believed that photosynthesis and respiration produced a high-energy intermediate leading to ATP synthesis. When Mitchell originally theorized the possibility that proton translocation across a membrane could effect the same action as the proposed intermediate, there was no clear mechanism or species by which the so-called proton motive force could be generated. Eventually however, it became clear that the lipophilic quinone cofactors which were known to be involved in electron transport as substrates for various respiratory complexes were both the proton and electron carrier driving a large part of chemiosmosis.³⁻⁵

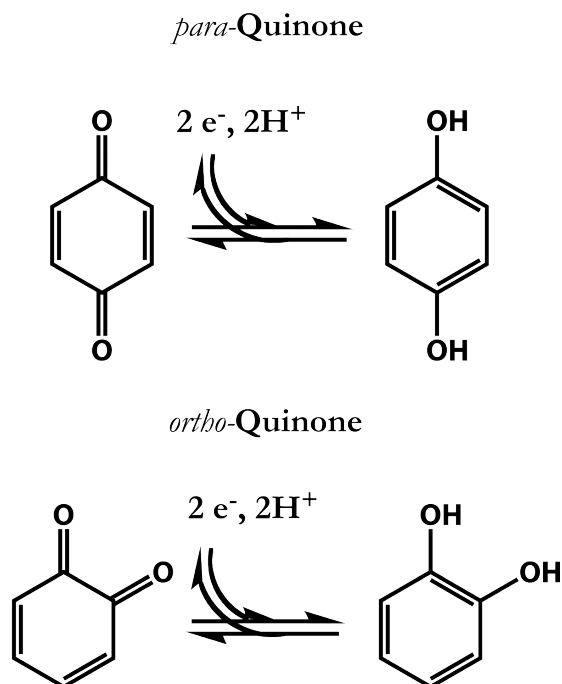


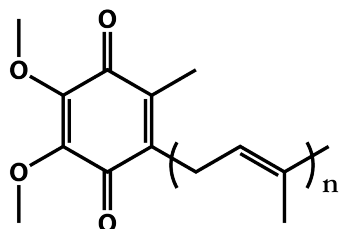
Figure 1.1. Genetic *para*- and *ortho*-quinones showing two electron, two-proton redox chemistry common at physiological pH.

The quinones are a diverse class of biological redox molecules implicated in everything from the electron transfer pathways of photosynthesis and respiration to catalysis of critical reactions in the cross linking of interstitial tissue and alternative metabolic pathways⁶⁻⁸ to antibacterial and antigenic quinones used in a broad array of biological defense mechanisms.^{9,10} This expansive utility comes from modifications around a quinone core; loosely defined as a cyclic, fully conjugated dione which can

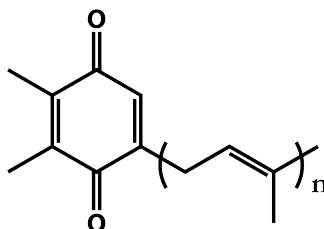
alternate between aromatic and non-aromatic forms in a two-electron (and often, two proton) reaction. Differing only in the placement of the ketone/hydroxyl substituents, the *para*-quinone and *ortho*-quinone isomers, Figure 1.1, are used in very different manners in nature. *para*-Quinones, like those found in the membranes of mitochondria and chloroplasts, Figure 1.2, are principally involved in electron transport in the generation of the proton motive force that ultimately drives ATP synthesis. Vitamin K, a naphthoquinone, is a notable exception as a critical cofactor in the Vitamin K dependent carboxylases.^{6,7,11} As will be discussed in more detail in Chapter 2, the common *ortho*-quinone cofactors derived from amino acids *in situ* in the quinoprotein family, Figure 1.2, also act directly on substrates including glucose,^{12,13} methanol,¹⁴ and lysine.¹⁵ The reasons behind the functional segregation

of *para*- and *ortho*-quinones in nature is not clear, but it might indicate a general need for *ortho*-positioned ketones to allow quinones to effect substrate level redox transformations.

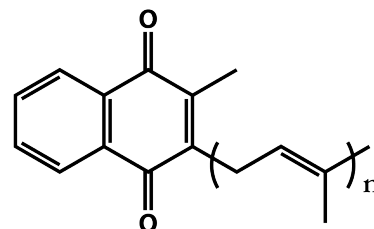
para-Quinones



Ubiquinone

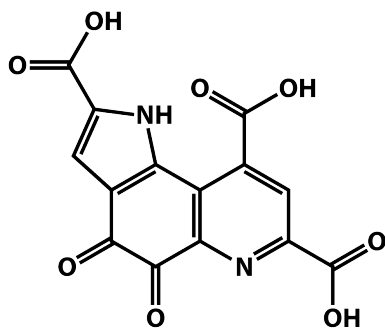


Plastoquinone

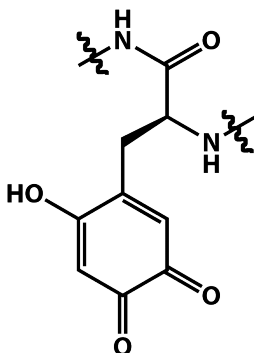


Vitamin K

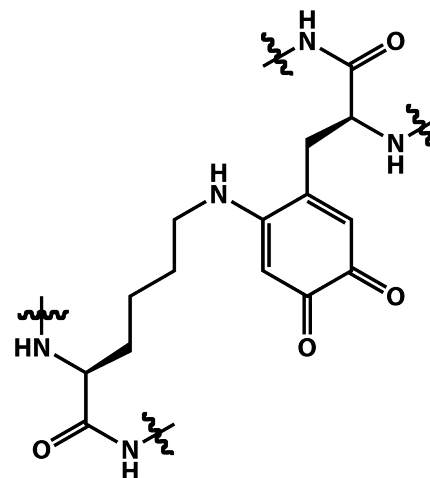
ortho-Quinones



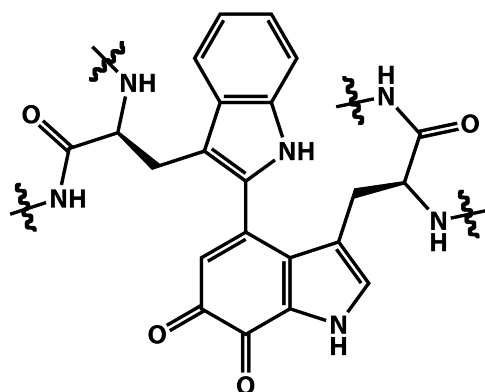
PQQ



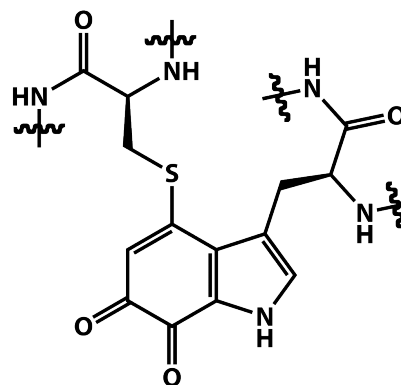
TPQ



LTQ



TTQ



CTQ

Figure 1.2. Natural quinone cofactors. PQQ, pyrroloquinoline quinone; TPQ, topaquinone; LTQ, lysine tyrosylquinone; TTQ, tryptophan tryptophylquinone; CTQ, cysteine tryptophylquinone.

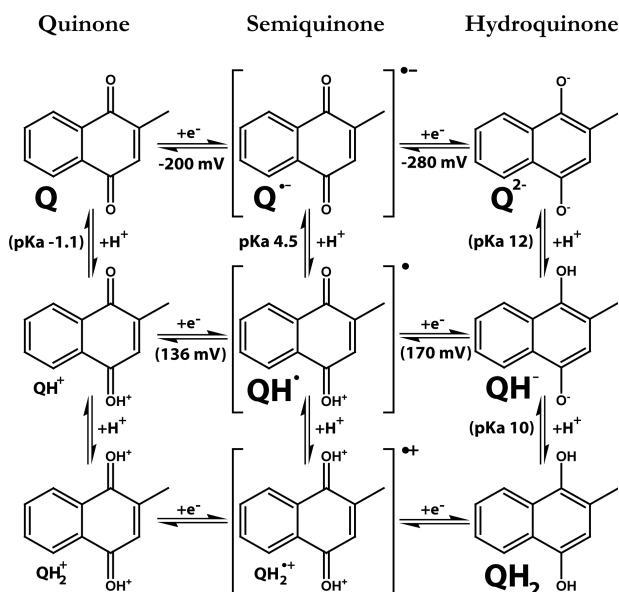


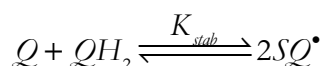
Figure 1.3. The nine redox states of quinones shown with 2-methyl-1,4-naphthoquinone's redox couples.² Values in parentheses are educated guesses or calculated from educated guesses. All known natural proteins make use of the states indicated in bold.

1.1.1 Quinone Electrochemistry

Quinones are almost universally capable of being reduced by two electrons and concomitantly acquiring two protons. This generates nine possible redox states, Figure 1.3. In solution at physiological pH, most quinones are reduced from the fully oxidized quinone form to the fully reduced, fully protonated hydroquinone

in a concerted two-electron, two-

proton reaction.¹⁶ Deviations from this behavior, however, can be observed if the typically unstable single electron semiquinone state is somehow stabilized. The stability of a semiquinone is a measure of the likelihood that a quinone and a hydroquinone colliding in solution will react to produce two semiquinones



where Q, QH₂ and SQ[•] are the fully oxidized, fully reduced and singly reduced species, respectively, without consideration of their protonation state. The energetic gauge of a semiquinone's stability is simply the free energy of this reaction

$$\Delta G = -RT \ln(K_{stab}) = -FE_m(Q \rightarrow SQ^{\bullet}) + FE_m(SQ^{\bullet} \rightarrow QH_2)$$

where F is Faraday's constant, $E_m(Q \rightarrow SQ^{\cdot-})$ and $E_m(SQ^{\cdot-} \rightarrow QH_2)$ are the midpoint potentials for the half-cell reactions indicated in the parentheses, and K_{stab} is the equilibrium constant for this reaction

$$K_{stab} = \frac{[SQ^{\cdot-}]^2}{[QH_2][Q]}$$

K_{stab} is a measure of the stability of the semiquinone; the greater the K_{stab} the more semiquinone will be present at a given pH and potential. At neutral pH in aqueous solutions, most quinones have a very low semiquinone stability ($K_{stab} < 10^{-3}$) thus become reduced in a concerted manner.¹⁷

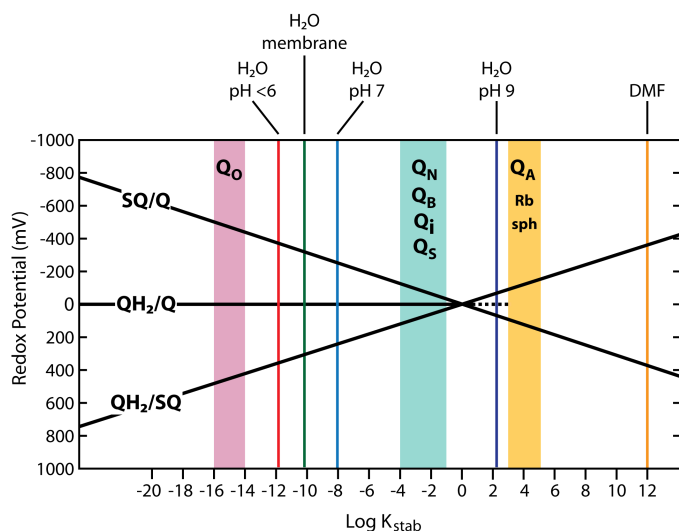


Figure 1.4. Natural quinone oxidoreductases classes (orange, teal, and purple) in relation to semiquinone stability ($\log(K_{stab})$) of quinones in active sites. Also shown (black lines) is the relation of redox potential of the semiquinone redox couples to $\log(K_{stab})$. Observed $\log(K_{stab})$ in DMF and water at various pHs are indicated. Figure after Ding *et al.*¹

1.2 Biological Control of Quinone Electrochemistry

While most quinones are reduced in a concerted two-electron, two-proton reaction at physiological conditions, natural quinone oxidoreductases are not limited to alternating between the quinone and fully reduced, fully protonated hydroquinone state at

their active sites. Through an array of poorly understood mechanisms, the active sites of quinone oxidoreductases selectively stabilize certain of the nine redox states seen in Figure

1.3. The active sites can be broken down into three broad classes, Figure 1.4. Quinones in Q_o -type sites, like that in b_c , show semiquinone stabilities similar to those seen in water and are typically found with multiple cofactors nearby to allow concerted uptake or release of electrons and protons.¹⁸⁻²¹ Quinones in Q_i -type sites, also found in b_c , have semiquinone stabilities much closer to parity and thus can form the single electron state on the way to being fully reduced. Unlike the Q_o -type sites, there is no need for multiple nearby cofactors since electrons can go onto or off of a quinone in the site one at a time. Finally, quinones in Q_A -type sites in bacterial photosynthetic reaction centers demonstrate highly stabilized semiquinones, but cannot be doubly reduced under most circumstance. Quinones in these sites are typically involved in the first electron transfer steps after photo-oxidation of water and thus need to have very low potentials and only transiently hold electrons.^{22,23}

While the differentiation of the active sites of quinone oxidoreductases demonstrate a 1 V (23 kcal mol⁻¹) difference in the stability of the semiquinone, there is little known about the actual mechanisms of control. Recent theoretical treatments²⁴⁻²⁶ have added to the understanding of the interplay between protein environment and active site quinones with particular focus on pKa and electrostatic effects, but have yet to clarify a great deal of the mechanistic ambiguity. Most experimental approaches have focused on mutational studies of the readily modified and studied bacterial reaction centers that, while providing a rough sketch of factors involved in controlling protonation and electrostatic drivers of quinone properties, have not been particularly well suited to giving a broad view, especially of the Q_o -type sites.^{22,23} In contrast, most of what is known about how quinones are likely to operate in natural systems comes from aqueous solution studies performed thirty years ago.^{16,27,28} In addition, studies of quinone properties in organic solvents have expanded understanding of

the role of structural modifications on their electrochemistry²⁹⁻³¹ as well as revealing a great deal about the effects of hydrogen bonding and electrostatic interactions on quinone redox properties and electron transfer rates.³²⁻³⁵ Despite the fact that these studies have increased understanding of physical-chemical attributes of quinones, they are of little immediate utility in determining the manner in which biological control is manifest as quinone properties change radically between protic and aprotic mediums.

The nature of the divergent activities of quinones in the active sites of membrane oxidoreductases remains an open question. Further, the interplay of quinone electrochemical properties and active site architecture in producing productive vectorial electron transfer reactions is not fully understood.²¹ Due to the experimental limitations of mutagenic studies on complex multisubunit proteins, it is not clear whether continued systematic study of the respiratory and photosynthetic complexes will yield any additional insights into the mechanisms of control. Rather, a new approach may be needed that carries the advantages of the small molecule studies in water and organic solvents, but maintains the structural and architectural diversity of peptide polymers. The non-natural protein maquettes were originally developed to explore the mechanisms of control of the redox properties of heme binding sites in membrane oxidoreductases.³⁶ Removed from the complex overlapping functions of the primary sequence of amino acids in any naturally evolved protein, the maquettes have been used to simplify the study of a growing set of cofactors in a protein environment,³⁶⁻⁵³ including quinones as will be discussed in detail in Chapter 2.⁵⁴⁻⁵⁷ Maquettes are not limited to incorporating single cofactors,⁴⁹ and provide an unprecedented opportunity to construct electron transfer active sites that emulate the function of natural oxidoreductase active sites. In seeking to understand the natural engineering principles

guiding biological mechanisms of quinone electrochemical control and oxidoreductase quinone active site design, it is clear that the development of a quinone maquette family would greatly expand the cache of tools that can be brought to bear on this pursuit.

1.3 References

- (1) Ding, H.; Moser, C. C.; Robertson, D. E.; Tokito, M. K.; Daldal, F.; Dutton, P. L. *Biochemistry-U.S.* **1995**, *34*, 15979.
- (2) *Physical Chemistry of Semiquinones*; Swallow, A. J., Ed.; Academic Press: New York, 1982.
- (3) Mitchell, P. *FEBS Lett.* **1975**, *56*, 1.
- (4) Mitchell, P. *FEBS Lett.* **1975**, *59*, 137.
- (5) Mitchell, P. *J. Theor. Biol.* **1976**, *62*, 327.
- (6) Rishavy, M. A.; Pudota, B. N.; Hallgren, K. W.; Qian, W.; Yakubenko, A. V.; Song, J. H.; Runge, K. W.; Berkner, K. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13732.
- (7) Dowd, P.; Hershline, R.; Ham, S. W.; Naganathan, S. *Science* **1995**, *269*, 1684.
- (8) Duine, J. A. *J. Biosci. Bioeng.* **1999**, *88*, 231.
- (9) Yamashita, M.; Kaneko, M.; Tokuda, H.; Nishimura, K.; Kumeda, Y.; Iida, A. *Bioorganic & Medicinal Chemistry* **2009**, *17*, 6286.
- (10) Kim, H. W.; Lee, C. H.; Lee, H. S. *Food Science and Biotechnology* **2009**, *18*, 755.
- (11) Suttie, J. W. *Annu. Rev. Biochem.* **1985**, *54*, 459.
- (12) Yuhashi, N.; Tomiyama, M.; Okuda, J.; Igarashi, S.; Ikebukuro, K.; Sode, K. *Biosens. Bioelectron.* **2005**, *20*, 2145.
- (13) Matsushita, K.; Shinagawa, E.; Inoue, T.; Adachi, O.; Ameyama, M. *FEMS Microbiol. Lett.* **1986**, *37*, 141.
- (14) Davidson, V. L. *Arch. Biochem. Biophys.* **2004**, *428*, 32.
- (15) Csizsar, K. *Prog. Nucleic Acid Res. Mol. Biol.* **2001**, *70*, 1.
- (16) Rich, P. R.; Bendall, D. S. *FEBS Lett.* **1979**, *105*, 189.
- (17) Clark, W. M. *Oxidation/Reduction Potentials of Organic Systems*; The Williams and Wilkins Co.: Baltimore, 1960.
- (18) Chance, B. **1952**.
- (19) Erecinska, M.; Chance, B.; Wilson, D. F.; Dutton, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 50.
- (20) Dutton, P. L.; Moser, C. C.; Sled, V. D.; Daldal, F.; Ohnishi, T. *Biochim. Biophys. Acta* **1998**, *1364*, 245.
- (21) Osyczka, A.; Moser, C. C.; Daldal, F.; Dutton, P. L. *Nature* **2004**, *427*, 607.
- (22) Rinyu, L.; Martin, E. W.; Takahashi, E.; Maroti, P.; Wraight, C. A. *BBA-Bioenergetics* **2004**, *1655*, 93.
- (23) Takahashi, E.; Wells, T. A.; Wraight, C. A. *Biochemistry-U.S.* **2001**, *40*, 1020.
- (24) Gunner, M. R. *Photosynth Res* **2008**, *97*, 1.
- (25) Gunner, M. R.; Madeo, J.; Zhu, Z. *J Bioenerg Biomembr* **2008**, *40*, 509.
- (26) Zhu, Z. Y.; Gunner, M. R. *Biochemistry-U.S.* **2005**, *44*, 82.

- (27) Rich, P. R. *FEBS Lett* **1981**, *130*, 173.
- (28) Rich, P. R. *Biochimica Et Biophysica Acta* **1981**, *637*, 28.
- (29) Prince, R. C.; Dutton, P. L.; Bruce, J. M. *Febs Letters* **1983**, *160*, 273.
- (30) Prince, R. C.; LloydWilliams, P.; Bruce, J. M.; Dutton, P. L. *Methods in Enzymology* **1986**, *125*, 109.
- (31) Ashnagar, A.; Bruce, J. M.; Dutton, P. L.; Prince, R. C. *Biochimica Et Biophysica Acta* **1984**, *801*, 351.
- (32) Gupta, N.; Linschitz, H. J. *Am. Chem. Soc.* **1997**, *119*, 6384.
- (33) Wain, A. J.; Wildgoose, G. G.; Heald, C. G. R.; Li, J.; Jones, T. G. J.; Compton, R. G. J. *Phys. Chem. B* **2005**, *109*, 3971.
- (34) Okamoto, K.; Ohkubo, K.; Kadish, K. M.; Fukuzumi, S. *Journal of Physical Chemistry A* **2004**, *108*, 10405.
- (35) Yuasa, J.; Yamada, S.; Fukuzumi, S. *Chemistry-a European Journal* **2008**, *14*, 1866.
- (36) 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.
- (37) Dai, Q. H.; Tommos, C.; Fuentes, E. J.; Blomberg, M. R.; Dutton, P. L.; Wand, A. J. *Am. Chem. Soc.* **2002**, *124*, 10952.
- (38) Sharp, R. E.; Moser, C. C.; Rabanal, F.; Dutton, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 10465.
- (39) Kennedy, M. L.; Gibney, B. R. *J. Am. Chem. Soc.* **2002**, *124*, 6826.
- (40) Noy, D.; Discher, B.; Lear, J. D.; Ashur, I.; Scherz, A.; Dutton, P. L. *Biophys. J.* **2003**, *84*, 176A.
- (41) Razeghifard, A. R.; Wydrzynski, T. *Biochemistry-U.S.* **2003**, *42*, 1024.
- (42) Summa, C. M.; Rosenblatt, M. M.; Hong, J. K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Bio.* **2002**, *321*, 923.
- (43) Calhoun, J. R.; Kono, H.; Lahr, S.; Wang, W.; DeGrado, W. F.; Saven, J. G. *J. Mol Biol* **2003**, *334*, 1101.
- (44) Magistrato, A.; DeGrado, W. F.; Laio, A.; Rothlisberger, U.; VandeVondele, J.; Klein, M. L. *J. Phys. Chem. B.* **2003**, *107*, 4182.
- (45) Maglio, O.; Nistri, F.; Pavone, V.; Lombardi, A.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3772.
- (46) Nanda, V.; Rosenblatt, M. M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P. L.; Saven, J. G.; Degrado, W. F. *J Am Chem Soc* **2005**, *127*, 5804.
- (47) Bender, G. M.; Lehmann, A.; Zou, H.; Cheng, H.; Fry, H. C.; Engel, D.; Therien, M. J.; Blasie, J. K.; Roder, H.; Saven, J. G.; DeGrado, W. F. *J Am Chem Soc* **2007**, *129*, 10732.
- (48) Ishida, M.; Dohmae, N.; Shiro, Y.; Oku, T.; Iizuka, T.; Isogai, Y. *Biochemistry-U.S.* **2004**, *43*, 9823.
- (49) 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.
- (50) Wei, Y. N.; Kim, S.; Fela, D.; Baum, J.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13270.
- (51) Laplaza, C. E.; Holm, R. H. *J. Biol. Inorg. Chem.* **2002**, *7*, 451.
- (52) Schnepf, R.; Horth, P.; Bill, E.; Wieghardt, K.; Hildebrandt, P.; Haehnel, W. *J. Am. Chem. Soc.* **2001**, *123*, 2186.

- (53) Welch, J. T.; Kearney, W. R.; Franklin, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3725.
- (54) Hay, S.; Westerlund, K.; Tommos, C. J. *Phys. Chem. B* **2007**, *111*, 3488.
- (55) Li, W.; Heinze, J.; Haehnel, W. J. *Am. Chem. Soc.* **2005**, *127*, 6140.
- (56) Li, W. W.; Hellwig, P.; Ritter, M.; Haehnel, W. *Chem. Eur. J.* **2006**, *12*, 7236.
- (57) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun (Camb)* **2009**, 168.

Chapter 2: Design, Development and Exploration of a New Naphthoquinone Cofactor

2.1 Introduction

It is somewhat remarkable that no generic, well-defined motif for quinone binding exists across the many various quinone oxidoreductases associated with metabolism, respiration and photosynthesis. While a weak motif was identified that was mildly predictive of Qi-type sites,² it is apparent that the conserved amino acids in the preserved sequences are mostly important for increased semiquinone stability rather than direct and generic binding of the quinone head group or common phytyl tails which offer very little in the way of specific contacts. It is interesting to consider that specificity might not actually be biologically essential in energy converting protein complexes, and in fact it may have been selected against in organisms capable of using various quinone cofactors under different metabolic conditions.

This lack of a well-defined quinone-binding motif precludes the adoption of a conserved sequence to create a quinone-binding site in a maquette. However, nature has

circumvented the need for binding specificity by making use of covalent quinone cofactors generated *in situ* from amino acids in the primary sequence of proteins. These so called quinoproteins are a relatively new class of alternative oxidoreductases that perform a variety of biological substrate level redox functions normally associated with enzymes that use flavin and nicotinamide cofactors, for instance, in amine oxidases and alcohol dehydrogenases, amongst others. Discovered in 1979 with the identification of the PQQ soluble cofactor, a variety of post-translationally generated quinone cofactors have been described, Figure 2.1.³⁻⁷ The mechanism of maturation of the cofactor generically involves oxidation and simultaneous covalent modification of either tyrosine or tryptophan.⁸ However, the detailed mechanisms for cofactor generation vary widely. The major biosynthetic route to tyrosine-derived cofactors appears to involve direct oxidation of the source aromatic residue with molecular oxygen in an autocatalytic copper-dependent manner.⁸ The mechanisms for formation of tryptophan-derived cofactors are seemingly more complex, requiring the involvement of heme-containing maturation oxidoreductases for formation of the TTQ cofactor⁹ and a variety of chaperons for formation of the CTQ cofactor.¹⁰

While the inclusion of a fully self-maturing quinone cofactor would obviously be an ideal approach to developing a quinone-containing maquette, it would require mimicry of the natural quinoproteins. The extent of the mimicry would prevent full control and thus understanding of the effects of the cofactor's environment on the quinone redox chemistry, a crucial piece of knowledge still broadly lacking. Rather, the inspiration gained from nature's use of covalent quinone-amino acids suggests another possibility: the use of a natural or synthetic quinone amino acid in the primary sequence of the protein maquette. In fact, one such natural amino acid exists, L-3,4-dihydroxyphenylalanine (DOPA), Figure 2.1,

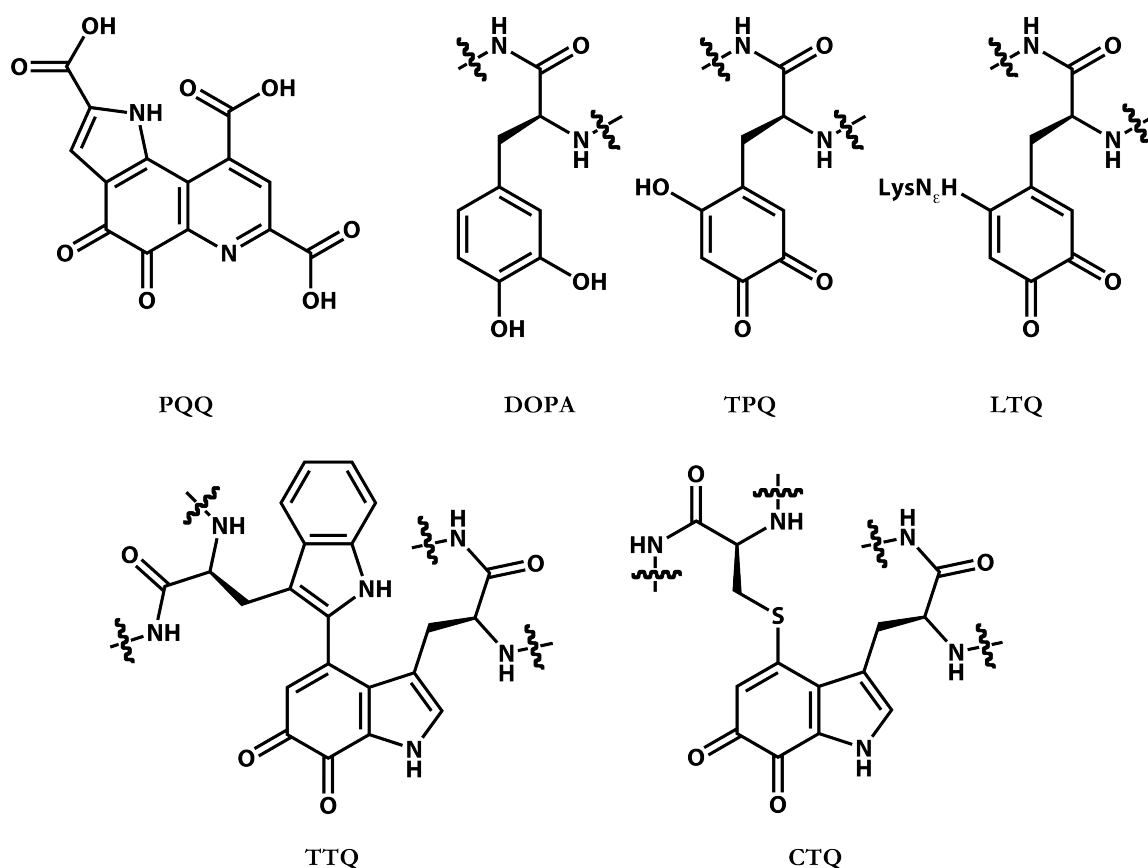


Figure 2.1. Naturally occurring quinone cofactors derived from amino acids. PQQ, pyrroloquinoline quinone; TPQ, topaquinone; LTQ, lysine tyrosylquinone; TTQ, tryptophan tryptophylquinone; CTQ, cysteine tryptophylquinone.

commonly used as a treatment for Parkinson's disease, and is, importantly, a likely intermediate in the spontaneous formation of the TPQ and LTQ cofactors in both natural⁸ and fortuitously generated quinoproteins.¹¹

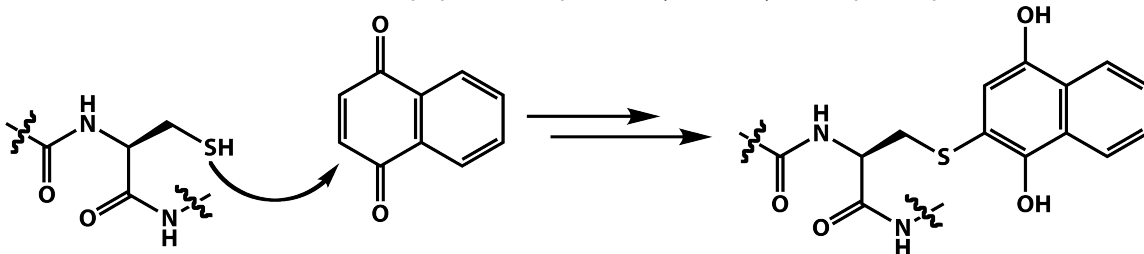
In work concurrent with that reported here, the Schultz group evolved a tRNA-aminoacyl tRNA-synthetase pair for *in vivo* incorporation DOPA into sperm whale myoglobin revealing limitedly explored changes in the redox chemistry of the DOPA amino acid and the heme.¹² The Stubbe group subsequently reported the semisynthesis of ribonucleotide reductase (RNR) $\beta 2$ -subunit with a DOPA amino acid incorporated into a synthetic fragment in the place of a tyrosine suspected of playing a critical role in the

prevalent radical hopping mechanism.¹³ Unsurprisingly, they observed that stabilization of a radical at the tyrosine position quenched nucleotide reduction. In addition, the formation of the DOPA radical (with concomitant reduction of the crucial diferric-tyrosyl radical cofactor) was dependent upon the presence of an allosteric effector and/or substrate in the full $\alpha 2/\beta 2$ complex, and occurred with kinetics similar to that observed in the intein generated wild type protein. The maintenance of a ligand response and competent kinetics suggests that the DOPA radical (and presumably the wild type tyrosine radical) is located along the radical propagation pathway. Stubbe and Schultz then collaborated on the *in vivo* incorporation of higher potential DOPA analogs at other sites in RNR providing insights into a potential hydrogen abstraction mechanism as a basis for catalysis at the nucleotide reduction active site,¹⁴ and other groups followed.¹⁵⁻¹⁸ While DOPA provides an important precedent for both *in vivo* and semisynthetic incorporation of a redox active quinone amino acid, its redox potential (~ 570 mV vs. NHE at pH 7.0)¹³ places it outside of the useful biological range for most energy conserving electron transfer reactions. It is also revealing that later groups used DOPA to allow for site specific redox activated cross linking, a creative use of the DOPA quinone form's intrinsic reactivity towards nucleophiles. This high reactivity in the oxidized form limits the useful lifetime of proteins in which it is incorporated and further reduces DOPA's appeal as a candidate for incorporation into quinone maquettes.

The electrophilicity of quinones arises from the presence of multiple Michael centers around the quinone ring in addition to the slight tendency for Schiff-base formation at the quinone carbonyls. In quinones with unsubstituted Michael centers, there is a well-known reaction that can occur with sufficiently nucleophilic atoms, such as the sulfur of cysteine,

Scheme 2.1, which results in a covalent bond and preservation of quinone redox function. This reaction provides the basis for another approach to covalently binding quinones to proteins. Recent efforts, simultaneous to the work presented here, have extended this technique to natural¹⁹ and synthetic^{20,21} proteins by reacting a cysteine with quinones having at least one unsubstituted Michael center. This technique is robust and allows for the incorporation of a diverse array of quinone head groups at a single cysteine site. The application of this approach has produced a rudimentary model of early mechanistic stages of photosynthesis in cytochrome *b562*, notably a four-helix bundle, demonstrating light induced electron transfer between a zinc chlorin and a quinone,¹⁹ as well as probative studies on the effects of a protein environment on the incorporated quinone.^{21,22}

Scheme 2.1. Covalent modification of peptide-incorporated cysteine by 1,4-naphthoquinone



The ease of use and expansive opportunities for quinone head group substitutions makes the cysteine substitution technique quite attractive from a design standpoint, but it is not without limitations. The need for a covalent linkage between the quinone head group and cysteine gives the conjugate amino acid a great deal of rotational flexibility due to the presence of three bonds between the backbone and quinone. The subsequent lack of a well-defined and small set of rotamers may explain the loss of apparent singular structure in a four-helix bundle modified by a quinone from the work of Haehnel.²⁰ In addition, although Wydrzynski reports only a small change in midpoint potential (25 mV) of the ubiquinone-0

head group upon binding and being buried in cytochrome *b562* compared to solution, Haehnel and Tommos observe massive changes in redox potentials (both groups, at least +110 mV)^{21,22} and apparent pKas (-3.4 pH units in Tommos' laboratory).²² Intriguingly, in both the Haehnel and Tommos proteins there are nearby (i->i+4) positively charged residues (lysine or arginine) in the primary sequence suggesting that perhaps there are cation- π interactions between the quinone headgroups and the local positively charged residue side chains. However, the values regularly measured for cation- π interactions (-0.9 kcal/mol between a tryptophan and lysine residue in a closely related protein to the one used in Tommos' work)²³ are not sufficiently energetic to fully account for the changes observed (0.9 kcal/mol is 20 mV for a two electron center). Further, the data for small molecule models of the cysteine ligation products show either small²⁰ or large²² variations in potentials from the free quinones. This unanticipated and inexplicable irregularity in the electrochemical properties of the covalently attached quinones are incredibly interesting, but suggest that this approach might not be conducive to constructing advanced quinone maquettes where the redox chemistry of the cofactors is somewhat predictable.

2.2 Design of a New Quinone Cofactor

The work described above provides a general framework for the desired features of a quinone cofactor whose incorporation into maquettes will allow systematic study of the effect of protein environment on quinone chemistry. In the need to create functional models of the critical reactions of photosynthesis and respiration, the primary consideration is the redox chemistry of the quinone cofactor. The midpoint potentials of the *ortho*-quinones, like DOPA, are generally much higher than their *para*-quinone equivalents. Indeed, in the low

potential reactions of respiration and photosynthesis, none of the quinone cofactors are *ortho*-quinones. In addition to affording redox chemistry closer to that observed in the major natural quinone oxidoreductases, developing a *para*-quinone cofactor has the distinct advantage of implicitly increasing the chemical stability of the oxidized form, thereby avoiding the cross-linking chemistry demonstrated by DOPA, while providing a functional group whose redox chemistry is more in line with natural systems.

Nevertheless, *para*-quinones are quite reactive at their unsubstituted Michael centers, so it is important that as many of the electrophilic sites around the new cofactor's quinone ring are substituted to reduce its reactivity. In the quinones used in respiration and photosynthesis, both the naphthoquinone core of menaquinone (Vitamin K) and the benzoquinone core of ubiquinone, Figure 2.2, have all of their Michael centers substituted. Lacking the methoxy substitutions of ubiquinone, the naphthoquinone head group of menaquinone is more robust, and in principle, provides an easier synthetic target making it the obvious choice for the functionality of the new quinone cofactor.

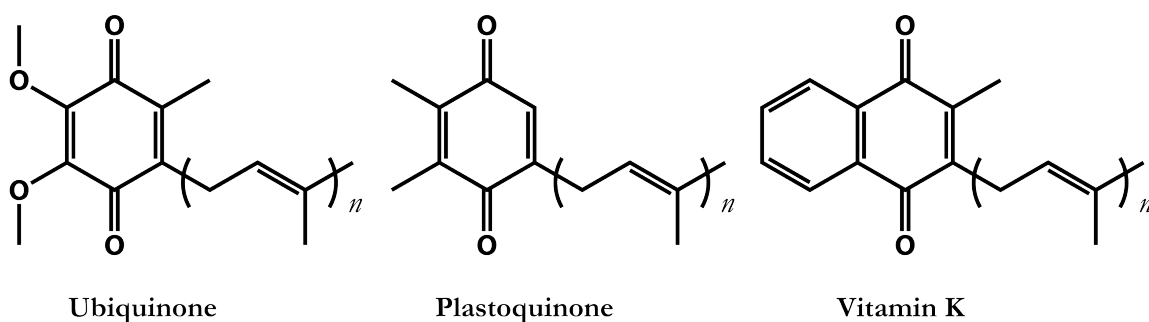


Figure 2.2. Membrane soluble *para*-quinone cofactors used in respiration and photosynthesis

As previously noted, the Haehnel group has developed a four-helix bundle containing a naphthoquinone-modified cysteine,²⁰ but this attachment scheme provides limited control over the position of the quinone head group and lacks consistent electrochemical behavior.¹⁹⁻²² It seems more prudent to create a naphthoquinone cofactor

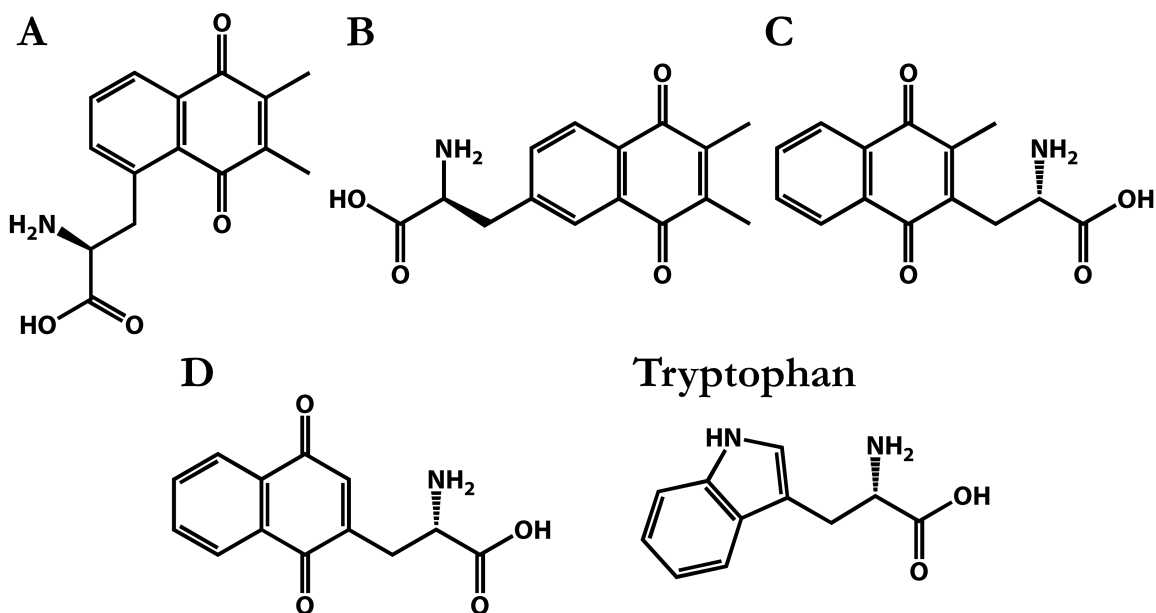


Figure 2.3. Possible backbone connectivity regioisomers of the naphthoquinone amino acid (A-C) and the final design (D); see text for details.

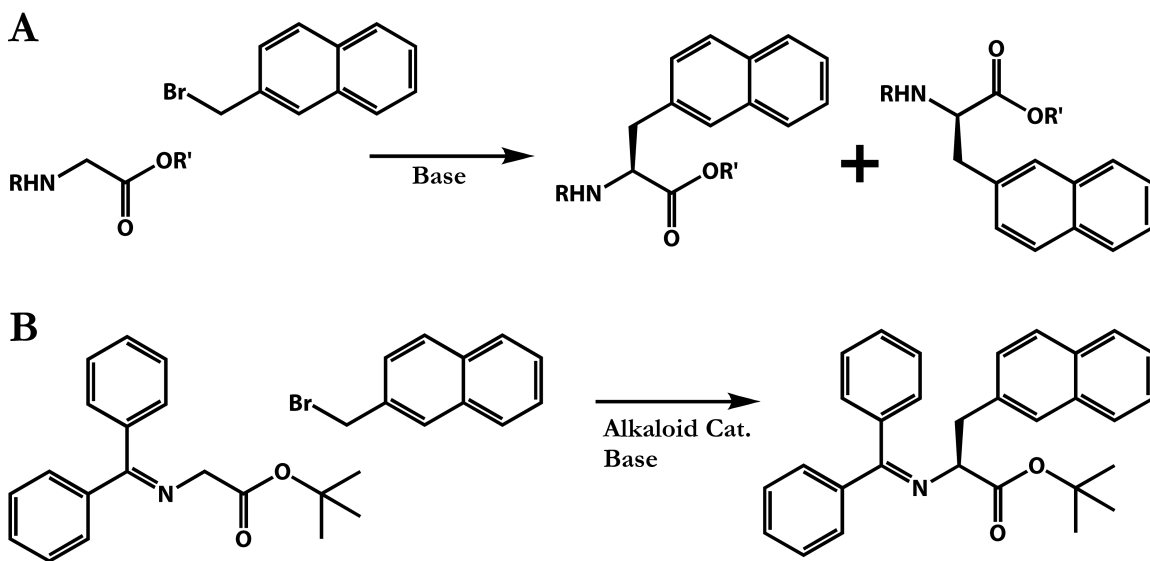
that is attached to the peptide backbone with a single bridging methylene, much like tryptophan or tyrosine. The comparison of the double ringed naphthoquinone side chain to tryptophan is apt, and of the three possible naphthoquinone side chains with methyl substituted Michael centers, Figure 2.3A-C, the only side chain which closely approximates the indole ring of tryptophan's steric bulk is the one in which the backbone is bridged by a methylene to the quinone side of the naphthalene ring, Figure 2.3C. While the proximity of the naphthoquinone core to the backbone might prove somewhat disruptive of secondary structure, its likely benefits are much larger, giving us the opportunity to place the quinone headgroup with better knowledge of its location relative to the remainder of the protein as compared to the quinone-modified cysteine cofactors. However, in order to partially address the possible disruption of secondary structure by reducing the steric volume of the naphthoquinone amino acid near the backbone and bringing its overall geometry closer to

that of tryptophan, sacrificing some amount of chemical stability proves necessary. Unblocking the non-methylene bridged Michael center accomplishes this goal, Figure 2.3D.

The design of the naphthoquinone amino acid, abbreviated Naq or **Q**, takes into consideration the likely structural constraints of a protein environment. Trimming the side chain to give Naq a steric bulk similar to that of tryptophan in theory provides enough leeway so that the naphthoquinone amino acid can be incorporated into virtually any protein element which tryptophan can be incorporated. The fact that the target design for Naq is modeled on a standard aromatic α -amino acid provides it with substantial conformational rigidity lacking in the cysteine-bridged quinone cofactors, and actually approximates the main chain side of the natural quinone cofactor amino acid TPQ, Figure 2.1. Finally, the pursuit of a naphthoquinone amino acid provides the best opportunity for developing a cofactor whose redox activity and general reactivity fall within the same range as those of natural quinone cofactors. This is necessary for fully emulating quinone-dependent respiration and photosynthetic reactions in new quinone maquettes.

The useful monomeric form of the naphthoquinone amino acid, Naq, does not resemble the image in Figure 2.3D. To fully realize the incorporation of Naq into maquettes, it is necessary to create a backbone-protected version to be used in peptide synthesis. Additionally, the primary peptide synthesis chemistry used in the Dutton laboratory is based upon the base labile 9-fluorenylmethylloxycarbonyl (Fmoc) amine-protecting group so the final desired form of Naq has the main chain amine Fmoc-protected and the carboxylate free to be activated for peptide chain extension.

Scheme 2.2. Generic electrophilic aromatic amino acid synthesis (A). The O'Donnell asymmetric amino acid synthesis (B)



2.3 Synthesis of Naq, the Naphthoquinone Amino Acid

The synthesis of Naq and characterization of heptaNaq found in the remaining sections of this chapter has been published already by Lichtenstein *et al.*²⁴

The naphthoquinone amino acid, Naq, in some ways resembles a standard aromatic amino acid, like tryptophan, tyrosine or phenylalanine. This is especially true in its reduced form where Naq is nearly isosteric to the artificial amino acid L-2-naphthalanine. The chemical syntheses of natural and artificial aromatic α -amino acids typically take advantage of the high electrophilicity of aryl-alkyl halides to form the bond between the final amino acid's α -carbon and β -carbon, derived from the formerly halogenated aryl-alkyl methylene, Scheme 2.2A. The new bond creates a chiral center at the α -carbon, and unless the reaction is highly asymmetric, the mixture of enantiomers would need to be separated to yield only the desired L-amino acid. Though synthesis literature abounds with amino acid synthesis examples dating back to Strecker in the mid-19th century, perhaps the most common

currently used highly enantioselective amino acid synthesis is based upon work by O'Donnell *et al.*²⁵ Reaction of a bulky glycine *tert*-butyl ester, whose α -protons are rendered more acidic by the modification of the glycine amine to a benzophenone imine, with an electrophile in the presence of a base and an alkaloid chiral phase transfer catalyst yields the electrophile substituted amino acid derivative in high enantiomeric excess, Scheme 2.2B. In most of the literature using this technique, the free amino acid is released under aqueous acidic conditions; though unreported, it is in principle also possible to remove both the benzophenone imine and *tert*-butyl ester under nucleophilic basic conditions. The simplicity and scalability of the O'Donnell approach made it an attractive synthetic transformation for the centerpiece of the synthesis of Naq.

2.3.1 Side Chain Protecting Groups

It was clear that to make use of the O'Donnell amino acid synthesis the desired naphthoquinone side chain needed to be in an aryl-alkyl halogenated form such as 1,4-dihydroxy-2-bromomethylnaphthalene, Figure 2.4, a compound very unlikely to be stable due to direct conjugation of the 1-position hydroxyl with a hypothetical cation formed from elimination of the bromide at the 2-position methylene. Indeed, the synthesis of the

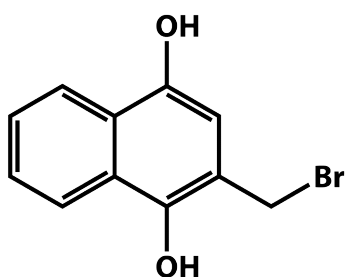


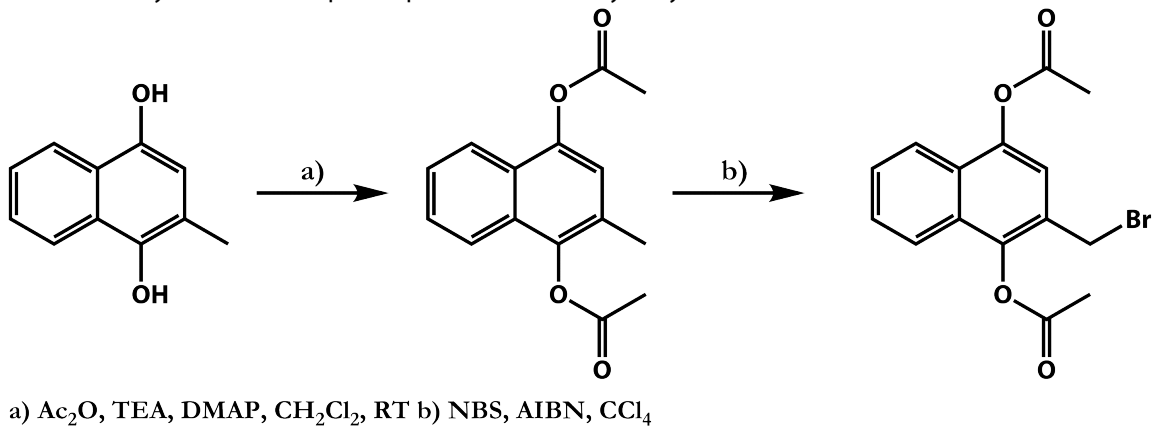
Figure 2.4. 1,4-Dihydroxy-2-bromomethylnaphthalene

naphthoquinone amino acid needed to be designed around maintenance of the quinone functionality through various reaction conditions the cofactor was likely to be exposed. In designing Naq, the focus was on

making sure the cofactor had the best possible chance at demonstrating redox activity over a useful pH and electrochemical range in aqueous solution. Little, if any, consideration was made as to how to best effect the final design and the stability of the monomeric cofactor in non-aqueous environments most often employed during synthetic transformations.

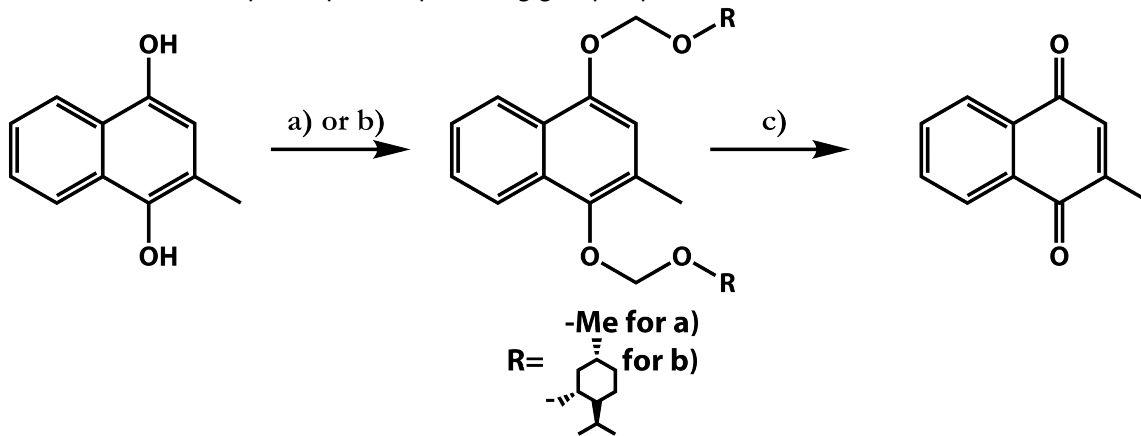
Quinones are notoriously reactive. From uncontrolled polymerization to the electrophilicity highlighted in earlier sections of this chapter, the best possible synthetic design to avoid these issues would take into consideration the fact that the reduced form is more stable in general, and that if redox cycling between the reduced and oxidized forms can be limited or stopped, the likelihood of successful synthetic transformations is far higher. The most straightforward method to obtain a naphthoquinone aryl-alkyl bromide that meets the above criteria is to acylate reduced 1,4-methylnaphthoquinone with acetic anhydride (Ac_2O) followed by bromination under radical conditions, Scheme 2.3.²⁶ Unfortunately, the acetyl groups are readily hydrolyzed, and, in considering further transformations, especially during peptide synthesis, the reactivity of acyl-aryl alcohols towards generic nucleophiles suggested that the naphthoquinone would not survive the synthesis. Instead, protecting groups that are more robust are needed.

Scheme 2.3. Synthesis of a naphthoquinone derived aryl-alkyl bromide.



Initial approaches to generating a protected naphthoquinone side chain capable of being transformed into an aryl-alkyl bromide focused on acid labile protecting groups that would remain attached during both the remainder of the amino acid synthesis and peptide synthesis. Deprotection of the naphthoquinone amino acid side chain could then be effected under the common acidic trifluoroacetic acid/scavengers conditions used in Fmoc/tBu peptide synthesis. It was hoped that basic conditions could be used in removing the benzophenone imine and *tert*-butyl group needed in the asymmetric O'Donnell synthesis.²⁵ Early attempts to form the *tert*-butyl ethers, commonly used to protect the tyrosine hydroxyl in peptide synthesis, proved fruitless and so more readily installed acid labile groups were pursued. To this end the acid labile methoxymethyl and menthoxymethyl ethers of 1,4-dihydroxy-2-methylnaphthalene were generated under basic conditions with chloromethyl methyl ether and (-)-chloromethyl menthyl ether, respectively, Scheme 2.4. It was predicted that the chiral menthoxymethyl ether could ease separation of the L- and D- amino acids potentially formed during the amino acid synthesis reaction by merit of creation of non-equivalent diastereomers. Unfortunately, all attempts to brominate the 2-methyl under

Scheme 2.4. Loss of naphthoquinone protecting groups upon oxidation.

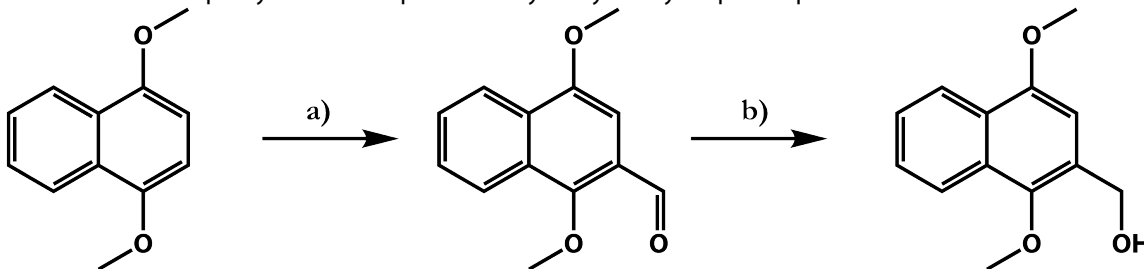


a) MOMCl, NaOH(aq), H₂O, CH₂Cl₂, RT b) (-)-Chloromethylmentylether, NaOH(aq), H₂O, CH₂Cl₂, RT c) NBS, AIBN, CCl₄

radical conditions with N-bromosuccinamide (NBS) failed to form the desired aryl-alkyl ether, and resulted in both loss of the protecting groups and re-oxidation of the quinone, Scheme 2.4. This was due in large part, no doubt, to the oxidative conditions commonly used for the radical brominations. This observation led to the discovery of a critical literature report in which the use of moderate hypervalent iodide(III) oxidants allowed the deprotection/activation of quinones from the methyl ether protected forms.²⁷

The abject failure of the radical bromination to produce useful products for continued transformations led to a rethinking in the general approach to the aryl-alkyl bromide. Instead of using a radical-based bromination, it is possible to convert aryl-alkyl alcohols to their respective bromides under non-oxidative conditions using phosphorus tribromide, or equivalent reagent. There are several methods to obtain the aryl-hydroxymethyl needed for this conversion. There is a literature method in which a methyl ether protected hydroquinone is reacted with dimethylformamide in the presence of the dehydrating agent phosphorous oxychloride to yield the formylated quinone, Scheme 2.5.²⁸ Subsequent reduction affords the desired alcohol. However, this transformation actually involves four steps starting with the highly reactive 1,4-naphthoquinone, and it seemed more straightforward to start with the natural product, 1,4-dihydroxy-2-naphthoic acid which can be transformed into the same alcohol in fewer steps and is naturally pre-reduced and air stable.

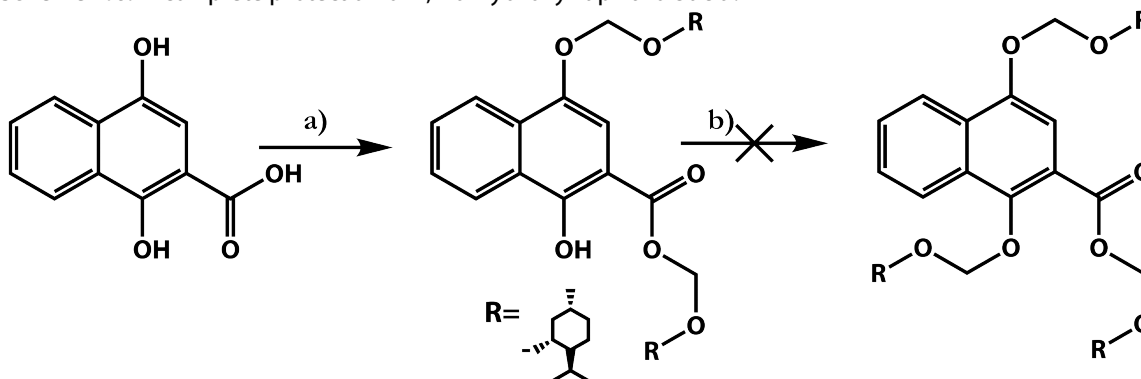
Scheme 2.5. Example synthesis of a protected hydroxymethyl naphthoquinone.



a) DMF-POCl₃, 80 °C b) NaBH₄, MeOH, RT

Early protecting schemes of the 1,4-dihydroxy-2-naphthoic acid proved difficult as the proton on the 1-hydroxyl group has a very high pK_a due to potential bridging to the carboxylate carbonyl, which is made higher upon deprotonation of the carboxylate. Under basic conditions, only the carboxylate and the 4-hydroxyl group became protected with the menthoxymethyl group, Scheme 2.6. Further protection of the 1-hydroxyl with chloromethyl menthyl ether, using sodium hydride or butyl lithium bases proved inconsistent. It was believed that by converting the carboxylate to an ester prior to attempting protection of the hydroxyls, the effective pK_a of the 1-hydroxyl could be reduced and full protection would occur more readily. Surprisingly, under acidic conditions in methanol, the obtained product

Scheme 2.6. Incomplete protection of 1,4-dihydroxynaphthoic acid.

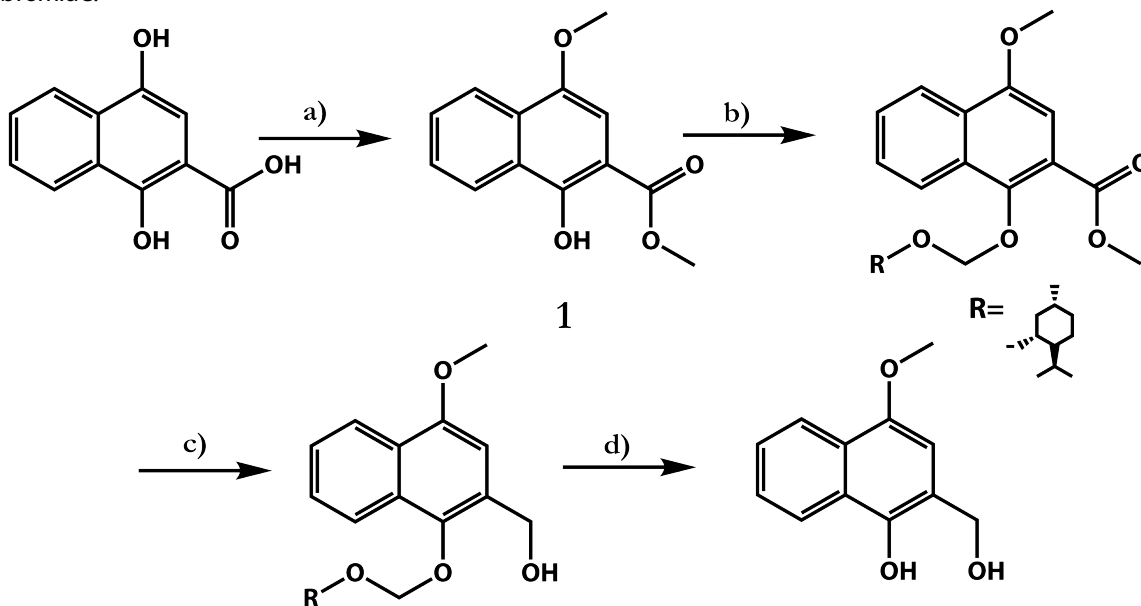


a) (-)-Chloromethylmentylether, NaH, DMF b) (-)-Chloromethylmentylether, NaH or Butyllithium, DMF

was not the ester, but rather the doubly methylated methyl 1-hydroxy-4-methoxy-2-naphthoate **1**, Scheme 2.7, in 93% yield. The addition of two methyl groups under the conditions used was completely unanticipated, and does not appear in the modern synthetic literature at all. A single test of the generality of the reaction was performed using ethanol, which proved effective.

Having partially accomplished the goal of deactivating the carboxylate, the free hydroxyl in doubly protected methyl naphthoate **1** was protected as the mentoxymethyl ether under basic conditions, Scheme 2.7. Reduction with lithium aluminum hydride afforded the desired alcohol, Scheme 2.7. However, upon treatment of the alcohol with phosphorous tribromide in carbon tetrachloride to generate the aryl-alkyl bromide, the only product observed was the alcohol in which the mentoxymethyl ether was deprotected, Scheme 2.7, with no sign of successful bromination. This surprising result, though similar in

Scheme 2.7. Progress towards a reductive route to a stably protected naphthoquinone aryl alkyl bromide.

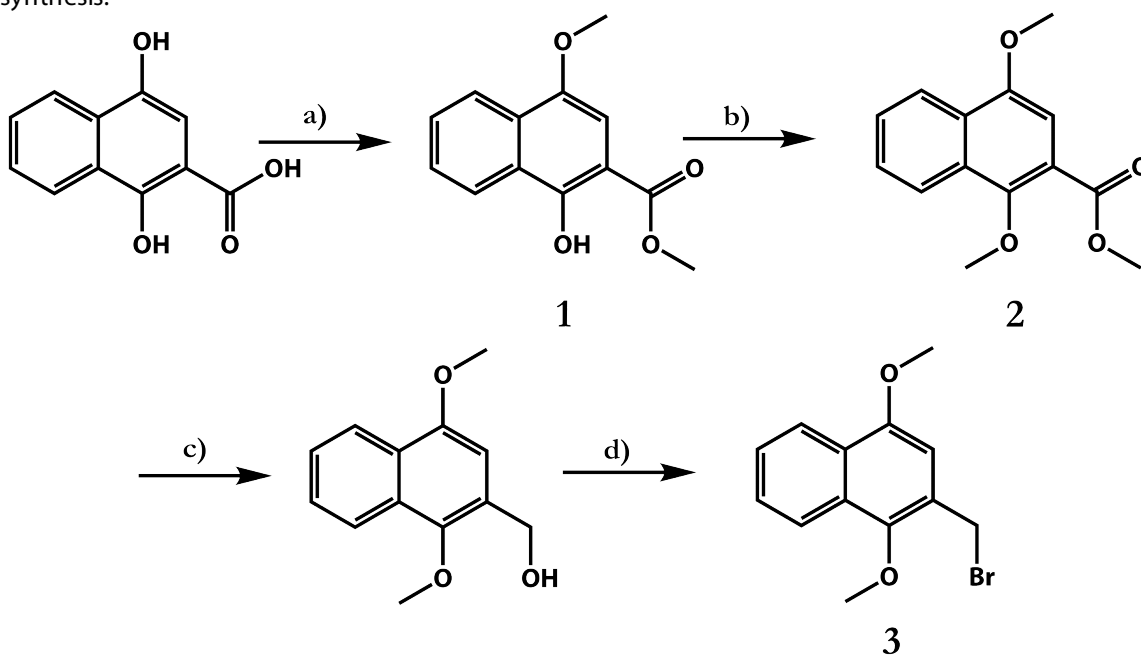


a) H_2SO_4 , MeOH, Reflux, 93.4% b) (-)-Chloromethylmentylether, NaH, DMF c) LiAlH_4 , THF, 0 °C
d) PBr_3 , CH_2Cl_2

effect, was not entirely reminiscent of the prior deprotections observed upon attempted bromination, in that the naphthoquinone remained reduced.

At this point, with seemingly little to gain from the chiral menthoxy methyl or achiral methoxy methyl protecting groups, especially in light of the previous observation that quinone methyl ether deprotection can occur under mildly oxidizing conditions, it was decided to alkylate the 1-hydroxyl with methyl iodide, and continue the synthesis. This decision bears significant weight on the remainder of this work as will be seen. Effecting the desired transformation of the methyl-naphthoate **1** to methyl-1,4-dimethoxy-2-naphthoate **2** afforded the desired triply methylated product in high yield (90%), Scheme 2.8. Subsequent reduction with lithium aluminum hydride, and bromination using phosphorous tribromide yielded the desired aryl-alkyl bromide 1,4-dimethoxy-2-bromomethyl-naphthalene **3** needed

Scheme 2.8. Successful generation of naphthoquinone aryl-alkyl bromide desired for amino acid synthesis.



a) H_2SO_4 , MeOH , Reflux, 93.4% b) CH_3I , K_2CO_3 , DMF , Reflux, 90% c) LiAlH_4 , THF , 0°C , quant.
d) PBr_3 , CCl_4 , 0°C , 95.8%

for the O'Donnell amino acid synthesis in nearly quantitative yield.²⁵

The transformations in four steps from 1,4-dihydroxy-2-naphthoic acid to 1,4-dimethoxy-2-bromomethyl-naphthalene **3** are high yielding, easily purified and readily scale to forty grams, the largest reaction size attempted during this work. It is apparent that the two steps towards the triply methylated methyl 1,4-dimethoxyl-2-naphthoate **2** allows for different and orthogonal protecting groups on either of the hydroquinone hydroxyls. However, attempts to silylate or otherwise install non-alkyl orthogonal protecting groups on the 1-hydroxyl failed to produce appreciable quantities of the desired products and these synthetic directions were abandoned. Given the success of installing the methyl protecting groups, the final desired form of Naq, ready for peptide synthesis can now be drawn, Figure 2.5.

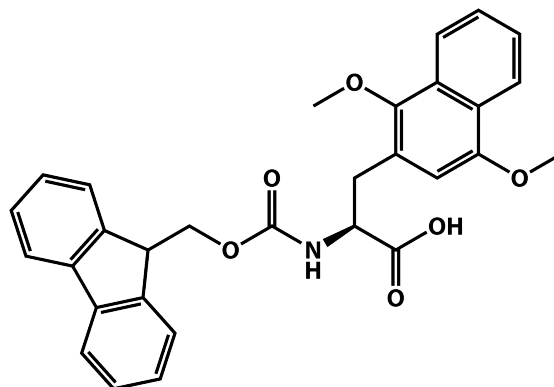
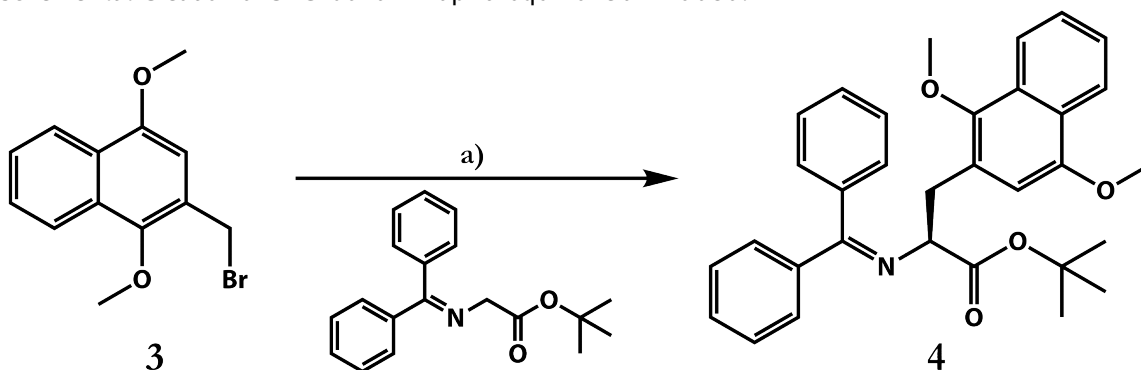


Figure 2.5. Final desired structure of Fmoc-Naq(OMe/OMe)-OH.

2.3.2 Completion of the synthesis of Naq

With the 1,4-dimethoxy-2-bromomethyl-naphthalene **3** aryl-alkyl bromide necessary for the O'Donnell amino acid synthesis in hand, it was combined with *N*-(diphenylmethylene)glycine *tert*-butyl ester in the presence of cesium hydroxide and the third-generation asymmetric cinchonidium phase transfer catalyst,²⁹ at -20 °C, Scheme 2.9. The product, *L*-*N*-(diphenylmethylene)-1,4-dimethoxy-2-naphthalanine *tert*-butyl ester **4**, was afforded in an 86% yield with a 93% enantiomeric excess, as determined by chiral

Scheme 2.9. Creation of C^α-C^β bond in naphthoquinone amino acid.

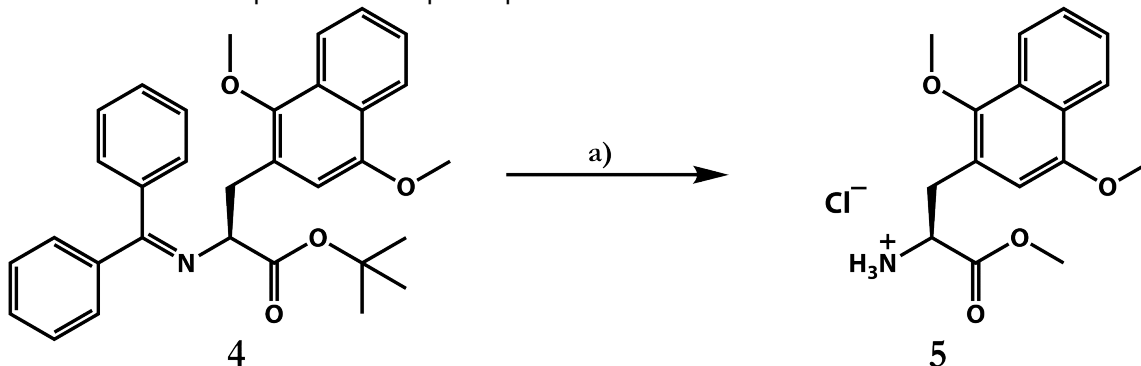


a) CsOH, 3rd Generation Cinchonidium PTC, Tol, -20 °C, 86%, 93% ee

HPLC of the free amino acid -- meaning approximately 4% of the product was the undesired D-enantiomer. Given the high stereochemical enrichment afforded by the O'Donnell amino acid synthesis, no further separation of the enantiomers was deemed necessary.

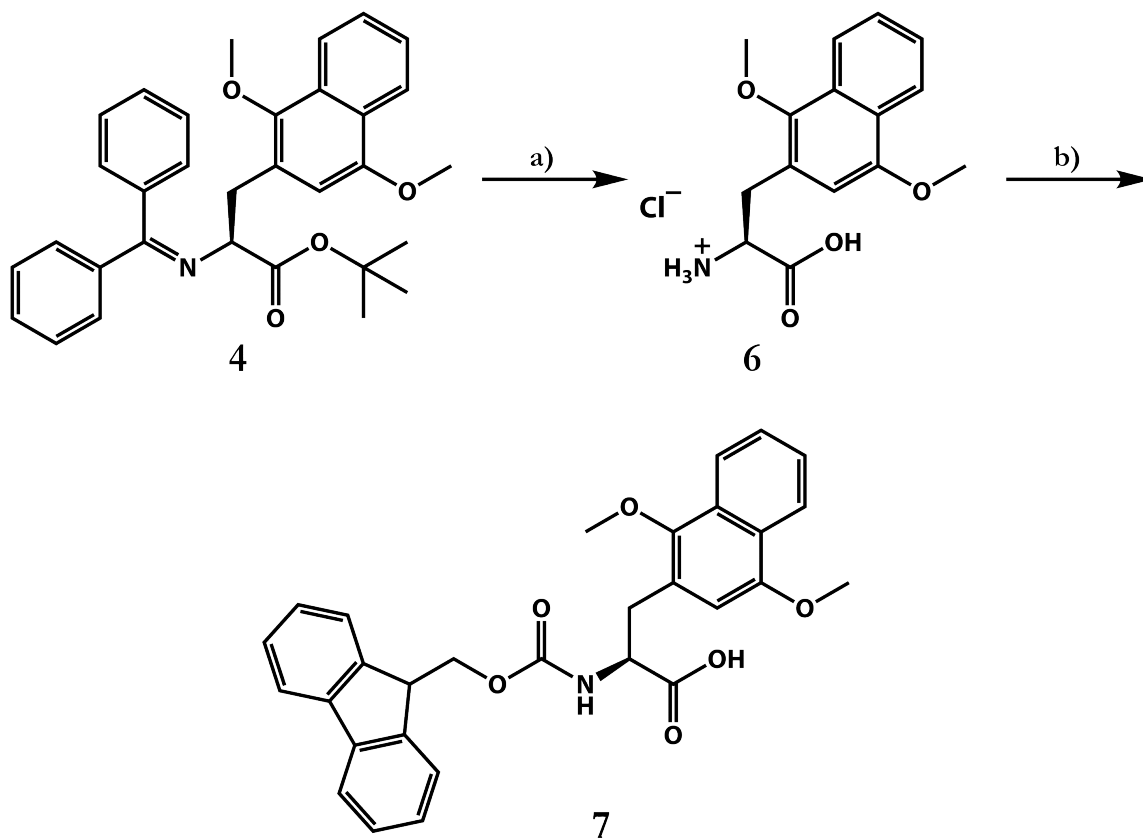
The protected amino acid 4 was committed to standard aqueous acidic conditions used in removing the diphenylmethylene and *tert*-butyl moiety protecting the backbone, rapidly resulting in a black tar. Under suspicions that the acidic water was somehow releasing the quinone functionality, followed by uncontrolled polymerization, 4 was treated with acidic

Scheme 2.10. Acidic deprotection of naphthoquinone amino acid backbone in methanol.



a) SOCl₂, MeOH, Reflux

Scheme 2.11. Final synthetic steps towards Fmoc-Naq(OMe/OMe)-OH.



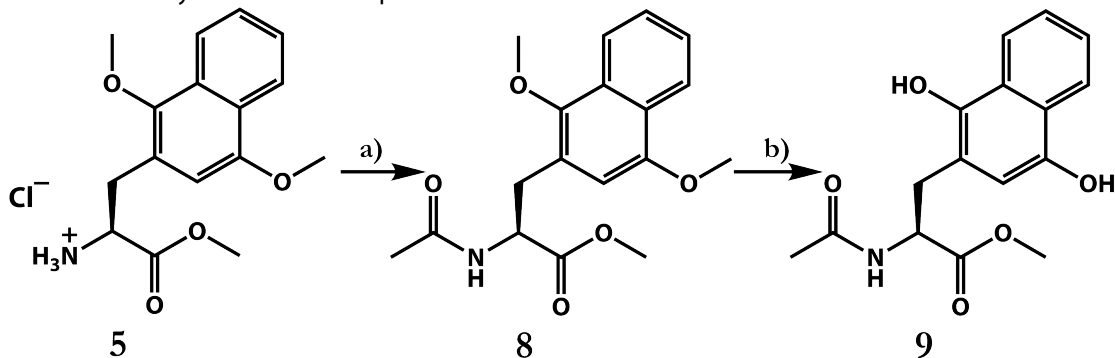
a) TFA, EDT, 98.2% b) Fmoc-OSu, NaHCO₃, DMF, 60 °C, 85%

methanol generated with thionyl chloride, to afford L-1,4-dimethoxy-2-naphthalanine methyl ester hydrochloride **5**, Scheme 2.10. While the methyl ester could easily be converted into the free amino acid, it seemed apparent that anhydrous acidic conditions could afford the desired free amino acid without the need to purify and deprotect the methyl ester. Using trifluoroacetic acid (TFA) with ethanedithiol (EDT) as a scavenger for the diphenylmethylene and *tert*-butyl cations likely to be released, protected amino acid **4** was converted to the free side chain protected naphthoquinone amino acid L-1,4-dimethoxy-2-naphthalanine hydrochloride salt **6**, Scheme 2.11, nearly quantitatively. This reaction darkens quickly if the scavenger is added after the TFA, suggesting either modification of the amino

acid by the released protecting groups or, possibly more likely, an oxidation dependent process which is suppressed by the moderately reducing EDT. Protection of the amine of **6** with Fmoc-OSu in the presence of stoichiometric bicarbonate gives the 9-fluorenylmethyloxycarbonyl (Fmoc) derived L-Fmoc-Naq(OMe/OMe)-OH **7** desired for Fmoc peptide synthesis, in an 85% yield, Scheme 2.11. Intriguingly, if the crude reaction mixture is allowed to sit, the 9-methylfluorene byproduct polymerizes.

The complete synthesis of L-Fmoc-Naq(OMe/OMe)-OH takes seven steps from the readily available natural product 1,4-dihydroxy-2-naphthoic acid. The key reaction creating the asymmetric α -carbon- β -carbon bond proceeds in high yield, with high enantiomeric excess thereby allowing the remainder of the synthesis to proceed without need for separation of the L- and D- enantiomers. While the methyl protecting groups proved quite robust in the chemical synthesis, their use in a synthetic peptide requires a separate deprotection step from the remainder of the peptide side chains. This feature has its advantages in that peptides containing the protected Naq will be much longer lived than peptides in which Naq was already deprotected. This simplifies peptide chemistry and allows peptides incorporating a protected Naq to be treated as any peptide can, without much consideration for the stability of Naq under various conditions, something that cannot be done with activated quinones like DOPA. It is important to note that like DOPA, attempts were made with a collaborator, Ryan Mehl, to integrate Naq into an orthogonal tRNA/tRNA synthetase pair for expression of proteins incorporating Naq in their primary sequence. The free amino acid proved somewhat toxic, and no viable hits were obtained after three attempts, so this aspect of the project was abandoned.

Scheme 2.12. Synthesis of Ac-Naq-OMe.



a) Ac_2O , NaOAc , DMAP , 94% b) PIFA , MeCN , H_2O ; reductive workup

2.4 Initial Electrochemical Studies of Naq in an Ultra-Short Peptide

The free amino acid methyl ester **5**, Scheme 2.12, can be readily converted into the peptide Ac-Naq(OMe/OMe)-OMe, **8**, with under standard acetylation conditions with 4-dimethylamino pyridine serving as a catalyst. As mentioned in the previous section, methyl ether protected quinone can be activated by using a hypervalent iodide(III) species like phenyliodine bis(trifluoroacetate) (PIFA) in nearly quantitative yields,²⁷ and so **8** was treated with PIFA in a water acetonitrile mixture, giving the hydroquinone product **9** after a reductive workup, Scheme 2.12.

Cyclic voltammetry studies of a short Naq containing peptide **9**, Figure 2.6, showed very strongly scan-rate dependent waves characteristic of the common ecec proton-dependent electrochemical mechanisms observed in cyclic voltammetry of quinones in aqueous solution.^{30,31} At the slowest scan rate, 1 mV s^{-1} , the half-wave potential, $E_{1/2}$, is -58 mV vs. NHE, close to the anticipated midpoint of 2-methylnaphthoquinone at pH 8.6 of approximately -110 mV vs. NHE. Besides obvious differences in local environment between 2-methylnaphthoquinone and Ac-Naq-OMe, the observed difference in midpoint potential could be do the quasi-reversibility of the quinone electrochemistry in the cyclic voltammetry

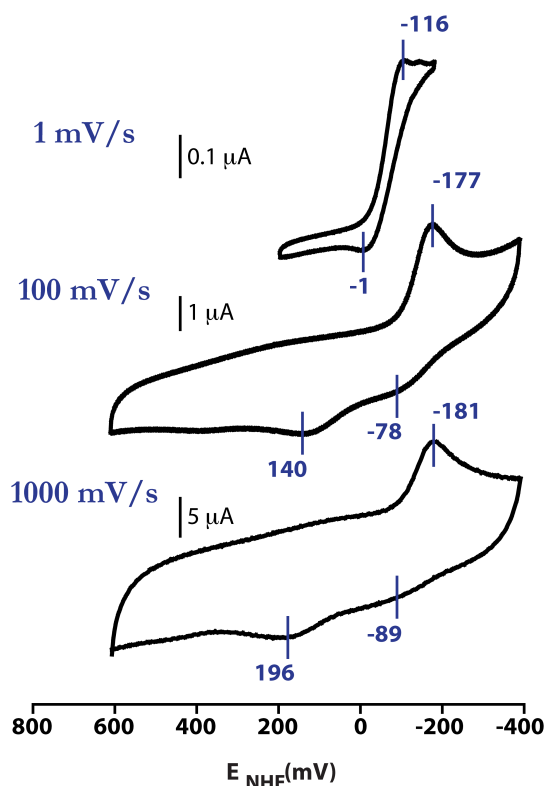


Figure 2.6. Cyclic Voltammetry of Ac-Naq-OMe in 25 mM Na₂PO₄, 250 mM KCl, pH 8.6 at various scan rates revealing complex ecec chemistry expected of quinones. Peak potentials are versus NHE, corrected from the Ag/AgCl reference used in the experiment. The working and counter electrodes were Pt.

studies. The non-ideality of the electrochemistry is such that even at the slow scan rate of 1 mV s⁻¹, the peak separation is 115 mV, far from the ideal diffusion limited 29 mV predicted for a two electron species at room temperature. Additionally, characteristic of a non-reversible electrochemical process, there is an observed scan rate dependence of the measured half-wave potential.

While the cyclic voltammetry of **9**

confirmed that Naq was indeed electrochemically active and displayed the standard suite of proton dependent quinone electron transfer chemistries when

deprotected using the hypervalent iodide(III) species PIFA, subsequent attempts to use the hypervalent iodide species with peptides of greater sequence diversity, incorporating the redox active amino acid tyrosine, proved thoroughly impossible as will be discussed in greater length in Chapter 3. Since little was known about how Naq would interact within a peptide containing diverse amino acid functionalities ranging from nucleophilic terminal amines and lysines, to redox active amino acids like tyrosine, tryptophan, cysteine and methionine, it became necessary to create a slightly longer model peptide with limited

functionality that would serve as model system for defining the range of activities likely to be observed in functionally diverse Naq containing peptide maquettes.

2.5 HeptaNaq

To ease reference to Naq in peptide sequences, it was given the single letter code **Q**, which is the archaic Greek letter koppa. With the sequence of PDP-Naq(**Q**)-PDQ-NH₂ and conceived as an unambiguous Naq containing short peptide, Figure 2.7, heptaNaq was designed around water solubility, containing both highly water soluble prolines and aspartates. In addition, the prolines provided structural rigidity thought likely to ease interpretation of electrochemical behaviors. The sequence acronym PDQ belies its original

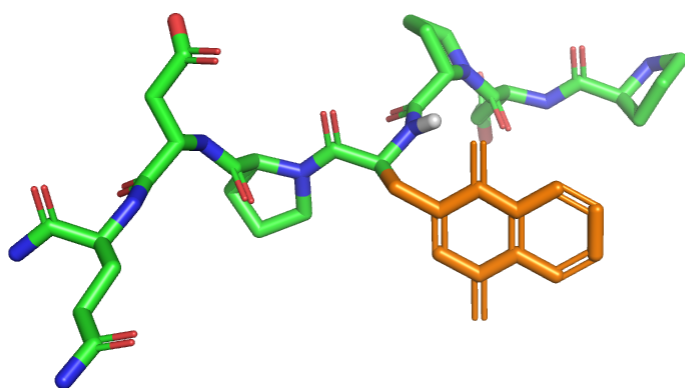


Figure 2.7. Cartoon of oxidized heptaNaq in poly-proline II conformation

purpose as a quick proving ground for deprotection strategies in peptides of greater complexity than Ac-Naq-OMe **9** as highly unanticipated problems arose in the course of attempting to deprotect Naq in such environments. The

problems yielded to neither re-synthesis of Naq using different protecting groups briefly alluded to before (the observations discussed in Chapter 6, notwithstanding) nor variegation of deprotection strategies, which will be discussed at greater length in Chapter 3.

HeptaNaq was synthesized using standard Fmoc-solid phase peptide synthesis procedures and protocols and purified as a single peak on HPLC. There was no evidence of

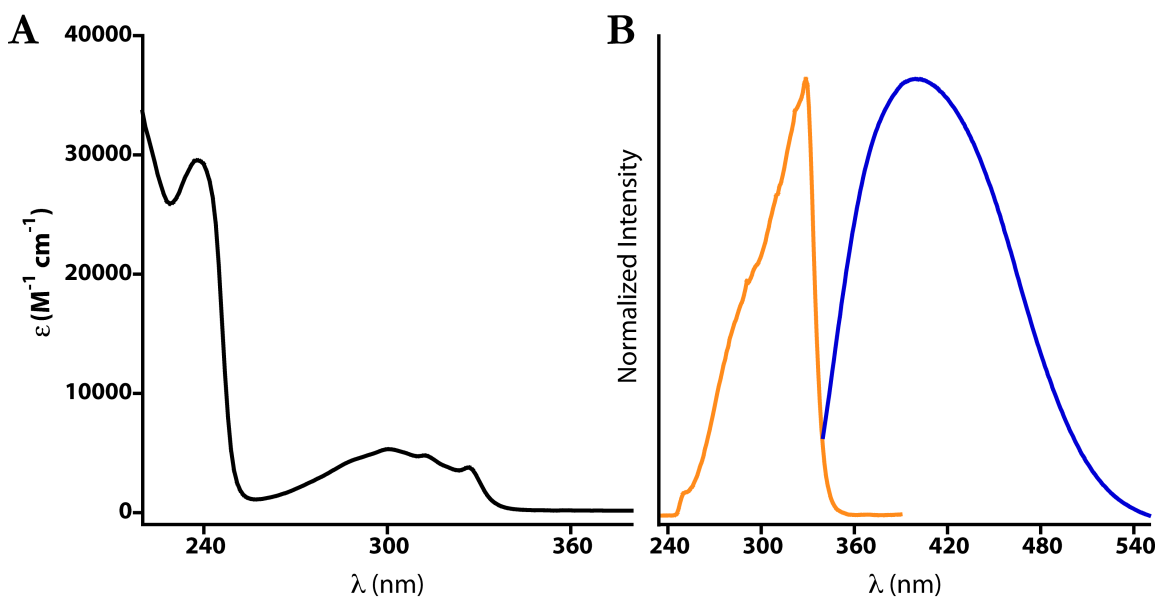


Figure 2.8. Average molar absorbance spectrum in 100 mM NaPi, pH 7.4 (A) and normalized fluorescence excitation (orange, measured at 400 nm) and emission (blue, measured at 329 nm) in 20 mM Pipes, 100 mM KCl, pH 6.57 (B) of heptaNaq(OMe/OMe).

lack of incorporation of Naq into the sequence or of degradation of the amino acid during peptide cleavage in Reagent R (TFA/thioanisole/EDT/anisole, 95:5:3:2), which like the cocktail used to deprotect Naq's backbone during its synthesis contains EDT, suggesting that Fmoc-Naq(OMe/OMe)-OH behaves appreciably similar to other natural and non-natural Fmoc amino acids used in peptide synthesis, an important feature needed for Fmoc-Naq(OMe/OMe)-OH's continued use.

2.5.1 HeptaNaq(OMe/OMe) Spectroscopy and Fluorescence

In the course of studying the deprotection chemistry of Naq(OMe/OMe), it became necessary to measure the molar absorptivity of heptaNaq to perform reactions with stoichiometric amounts of deprotecting reagents. Average molar absorptivities were measured using total Kjeldahl nitrogen to determine heptaNaq concentration in solutions

with known spectra, Figure 2.8A. The molar absorptivity at 301 nm, determined via linear regression fitting of the Kjeldahl and absorbance data as opposed to simply averaging, $5065 \text{ M}^{-1} \text{ cm}^{-1}$, is used as a standard value for determining protected Naq-containing peptide concentrations since no common amino acids have significant absorbances at that wavelength and the spectrum is not appreciably pH dependent.

HeptaNaq(OMe/OMe), and indeed, seemingly all peptides that contain protected Naq, has a remarkably bright blue fluorescence observable by eye when taking standard UV/*vis* spectra, Figure 2.8B.

2.5.2 HeptaNaq(OMe/OMe) Deprotection

Hypervalent iodide(III) species, such as PIFA, which were successfully used to deprotect Naq in the peptide Ac-Naq-OMe, react through a covalent intermediate with the species they are oxidizing. Though never attempted, heptaNaq would likely be efficiently deprotected by PIFA; however, the covalent intermediate biases the reaction of the hypervalent iodide(III) species towards reacting with species that have less steric bulk, and in peptides containing both Naq(OMe/OMe) and tyrosine, the formation of the covalent intermediate heavily favors oxidation of tyrosine over deprotecting Naq(OMe/OMe). Nevertheless, use of other non-covalent oxidants might deprotect Naq(OMe/OMe) first, especially if the redox potential of the reduced protected Naq is lower than the potential of the tyrosine.

After numerous deprotection attempts with other methodologies, some of which will be discussed in Chapter 3, a range of single electron oxidants were used to attempt to deprotect heptaNaq(OMe/OMe); only one, with stoichiometric quantities of cerium

ammonium nitrate (CAN), was able to completely oxidize heptaNaq(OMe/OMe). The nominal midpoint potential of CAN is 1.7 V vs. NHE and is likely far more oxidizing than either tyrosine or Naq(OMe/OMe) are reducing, thereby requiring addition of almost exactly two equivalents per protected Naq to be selective. Equal concentrations of heptaNaq(OMe/OMe) and heptaTyr, where the Naq has been mutated to a tyrosine, were reacted with two Naq(OMe/OMe) equivalents of CAN activating Naq with no major tyrosine oxidation. This selectivity is enhanced at low pH as the redox chemistry of tyrosine is pH dependent and the methyl-protected Naq is not anticipated to have a proton dependent couple. As will be discussed in Chapter 3, another oxidant dichlorodicyanoquinone (DDQ) that has a lower potential than CAN and is an obligate two-electron redox center, like Naq, is far more selective; nevertheless, heptaNaq is activated quantitatively by treatment with stoichiometric amounts of cerium ammonium nitrate.

2.5.3 HeptaNaq Aqueous Spectroscopy and Potentiometry

The aqueous spectra of 1,4-naphthoquinone and 2-methyl-1,4-naphthoquinone only differ by the appearance of absorption bands at 262 nm and 268 nm in the latter, attributed to absorption changes associated with the quinoid ring, Figure 2.9A.³² Activated oxidized heptaNaq (heptaNaq_{ox}) has a UV/*vis* spectrum that lies between the two species, although it is on first glance more similar to 1,4-naphthoquinone than to the more structurally similar 2-methyl-1,4-naphthoquinone, Figure 2.9B. The suppression of the bands at 262 nm and 268 nm in heptaNaq may be due to increased structural rigidity of the methylene off of the quinone in Naq reducing some of the electronic effects of the added methyl in desymmetrizing the cofactor as seen in 2-methyl-1,4-naphthoquinone. Regardless, heptaNaq's UV/*vis* spectrum is highly characteristic of naphthoquinones, with the broad moderate band at 239 nm imparting a slight orange tinge to concentrated solutions.

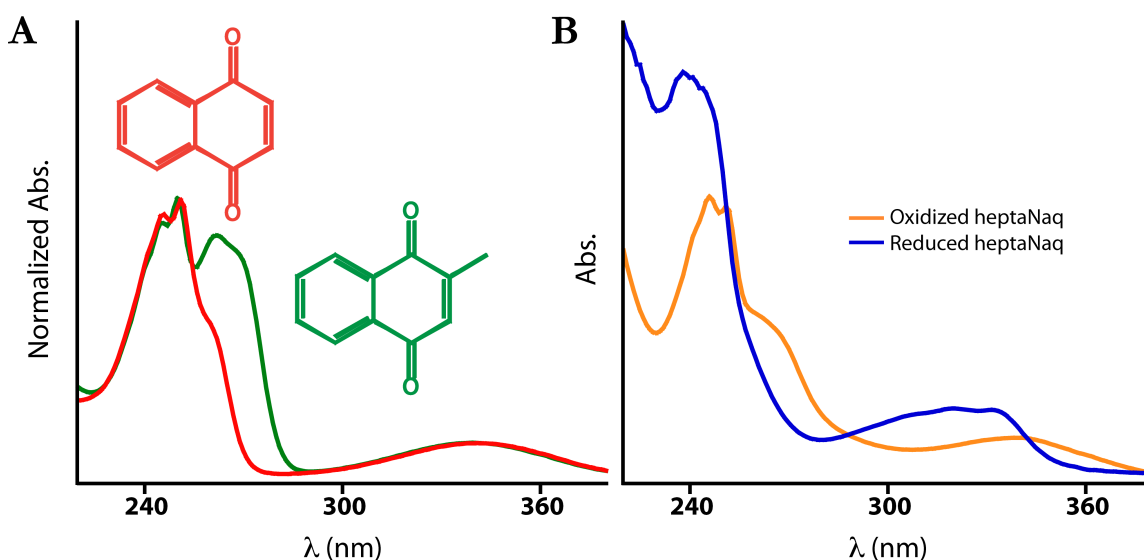


Figure 2.9. Normalized absorbance spectra of 1,4-naphthoquinone (red) and 2-methyl-1,4-naphthoquinone (green) in 25 mM KPi, 100 mM KCl, pH 8.0 (A) and oxidized (orange) and borohydride reduced (blue) heptaNaq in 20 mM Pipes, 100 mM KCl, pH 6.57.

Similarly, the reduced spectrum, Figure 2.9B, generated by addition of a slight excess of borohydride, is also as expected, demonstrating a significant red shift to the absorbance maximum at 320 nm when compared to the maxima in protected heptaNaq(OMe/OMe) at 301 nm. Excluding the work using a ubiquinone (UQ-0) head group,^{19,21} it should be noted that although the various groups that have used cysteine-modified quinones have reported spectra of peptides incorporating the quinones, none have been so unambiguously the correct species, perhaps indicating chemical modification of the quinones beyond the ligation chemistry or showing evidence of aggregation phenomena that could be significantly altering the quinone environment.²⁰⁻²²

With the spectroscopy of the reduced and oxidized heptaNaq well established, and with the observation that the erratic electrochemical results observed in the quinone-modified cysteine peptides were all determined using decidedly non-equilibrium cyclic voltammetry, heptaNaq's aqueous electrochemistry was determined by

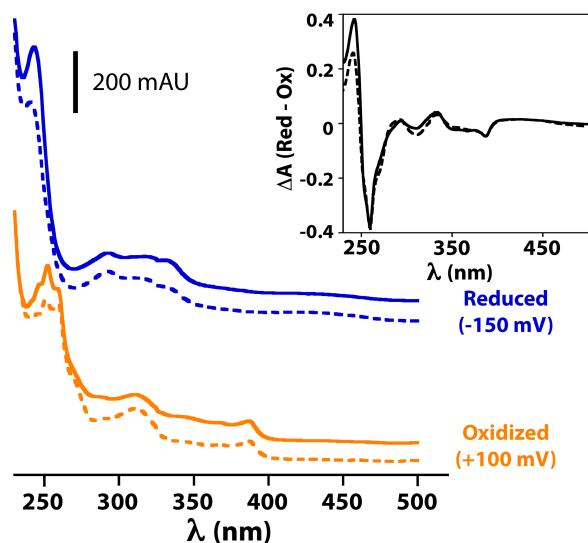


Figure 2.10. Reduced and oxidized spectra of heptaNaq (----) and 2-methyl-1,4-naphthoquinone in the presence of redox mediators. Inset: Redox difference spectra of the two species.

spectroelectrochemistry using a home built spectroelectrochemistry cell developed by José Cerda. Initial studies of heptaNaq without redox mediators and with an unmodified, large surface area, gold working electrode produced somewhat non-ideal results. Aggregation of the heptaNaq on unmodified gold foil spoiled attempts at reversing the direction of the titration since the

aggregates tended to dislodge during the course of the experiment. Without mediation, heptaNaq would demonstrate classic two-electron two-proton spectral changes but have an apparent one-electron Nernst response, yet ultimately displaying the same two-electron midpoint as determined under more ideal conditions. This is to be expected to some degree, as there is no physical meaning of a set point redox potential when there are few redox active charge carriers buffering it, which, in the absence of mediators beyond heptaNaq, would be the case at the edges of the heptaNaq Nernst curve. However, when the gold working electrode was modified with a short mercaptoalkylalcohol, and well-spaced mediators were used, heptaNaq demonstrated ideal redox behavior.

The reduced, oxidized and redox difference spectra observed during the course of the spectroelectrochemistry titrations matched that of the model compound 2-methyl-1,4-naphthoquinone under identical conditions, Figure 2.10. The observed maxima at 243 nm, and isosbestic point at 252 nm were identical to the values observed for the oxidized and

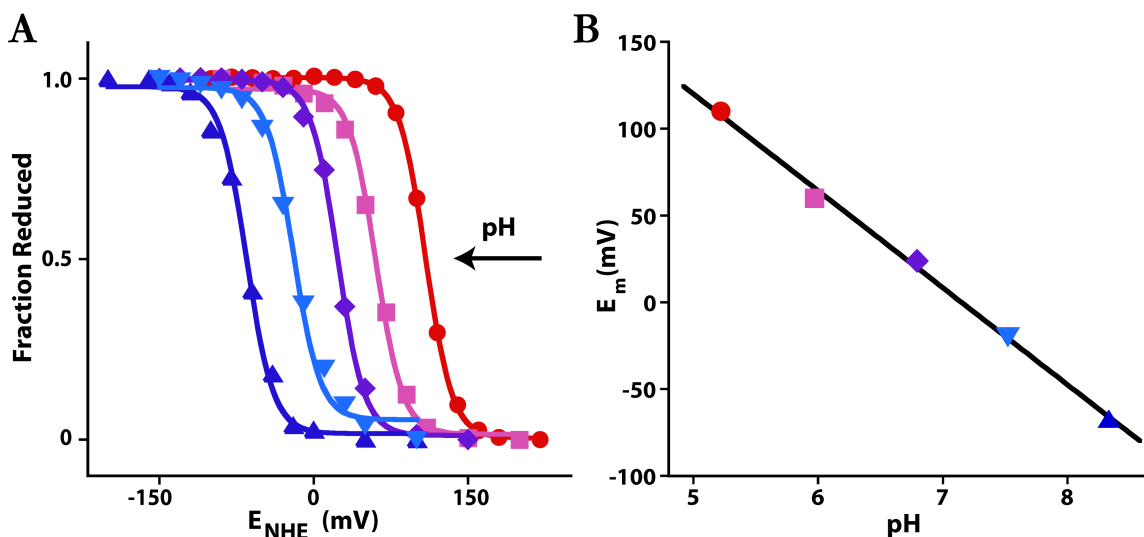


Figure 2.11. Oxidation-reduction chemistry of heptaNaq. Spectroelectrochemical data fit to the two-electron Nernst equation (A) from which the values of E_m were obtained. The pH dependence of E_m is drawn with line for a two-electron two-proton coupled redox center (slope -58 mV/pH at 20 °C).

borohydride-reduced species, Figure 2.9B. The pH dependence of the redox midpoint potential, E_m , of heptaNaq was determined spectroelectrochemically. Titrations were followed by measuring the difference between the absorbance at the maxima and isosbestic wavelengths as a function of redox potential, Figure 2.11. HeptaNaq's oxidation-reduction exhibits classic two-electron reversible Nernst behavior ($n = 2.0$) over the pH range 5.2-8.3 at 20 °C. The pH dependence of heptaNaq matches the expected two-electron, two-proton coupled oxidation-reduction of quinones over this pH range., Figure 2.11.³³ The E_m of heptaNaq at pH 7.5 is -20 mV, a value comparable to that of the Naq analog 2-methyl-1,4-naphthoquinone (E_m -47 mV at pH 7.5). The modest difference is similar, albeit smaller, to that observed in the cyclic voltammetry of Ac-Naq-OMe. Environmental effects caused by the peptide or more specific hydrogen bond interactions of the Naq side chain with heptaNaq's backbone can explain the subtle difference between the model compound and heptaNaq.

Overall the redox chemistry of heptaNaq is ideal, showing neither a significant change in the pKa of the hydroquinone hydroxyls as observed by the cysteine coupled quinone in the work of Hay *et al.*,²² nor deviating from ideal behavior of known quinone cofactors in the pH range explored. In addition, the aqueous solution equilibrium potentiometry confirmed that Naq, as designed, maintains a sufficiently low redox potential for use in maquettes built to emulate the low potential reactions of respiration and photosynthesis. However, while confirmation of aqueous electrochemical behavior is important for establishing the broad utility of Naq in soluble peptides, most of the low potential chemistry of quinones occurs in environments with low dielectrics and proton limitation within membranes and membrane protein complexes.

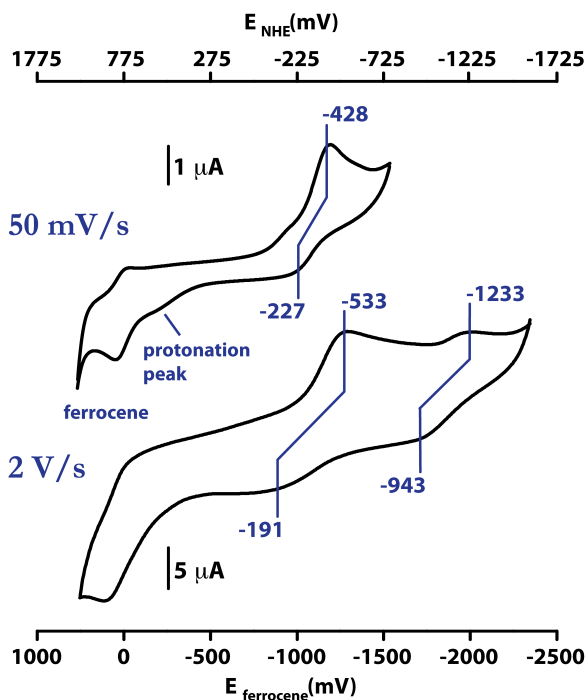


Figure 2.12. Cyclic voltammetry of heptaNaq in DMF. Slow (50 mV s^{-1}) and fast (2 V s^{-1}) scan rates redox couples with different proton dependencies. Peak potentials are presented relative to NHE using the ferrocene couple (766 mV versus NHE) as an internal standard.¹

2.5.4 HeptaNaq Aprotic Cyclic

Voltammetry

It is not surprising that heptaNaq displays high solubility in dimethylformamide (DMF), an aprotic solvent whose dielectric constant is 38.3 D falling in-between the average cited dielectric of between $2.5\text{--}4 \text{ D}$ for proteins,³⁴ which are of course more heterogeneous than a solution model would provide, and 80 D for water. Ideally suited to approximate the behavior of some of the slightly watery active sites of membrane

quinone oxidoreductases, solution cyclic voltammetry in DMF does not suffer from disruptive IR drops due to high solvent resistance as is observed in low dielectric solvents like benzene³⁵ making observation of proton-limited heptaNaq electrochemistry readily accessible.

With cyclic voltammetry using slow scan rates in DMF, heptaNaq oxidation-reduction peaks display a clear scan rate dependency not unlike those observed in the cyclic voltammetry of Ac-Naq-OMe in aqueous buffer. The peaks probably originated with restricted protonation by protons associated with the peptide itself as HPLC purified heptaNaq is likely to have at least the N-terminal amine and the two aspartates protonated due to the presence of trifluoroacetic acid (TFA) in the elution solvent (some mixture of

water and acetonitrile with 0.1% TFA). Indeed, at 50 mV s⁻¹ a “protonation peak” is observed indicative of the availability of protons in the DMF solution, Figure 2.12. Further, although there is a significant depression of the half-wave potential, $E_{1/2}$ -328 vs. NHE, compared to the aqueous midpoint, a complex ece(c) electrochemical mechanisms is evident which is not observed from studies with 2-methyl-1,4-naphthoquinone under identical conditions. However, unlike with the aqueous solution cyclic voltammetry of Ac-Naq-OMe, with increasing scan rates both anodic and cathodic peaks split into two one-electron steps and approach scan rate independence characteristic of kinetic bypass of protonation and development of fully reversible redox chemistry. Thus at fast scan rates in the range of 2 V s⁻¹ the two proton uncoupled half-wave potentials are -362 mV and -1088 mV vs. NHE. Like the previously reported redox couples for aqueous cyclic voltammetry and equilibrium spectroelectrochemistry, these values are modestly positive of the values of 2-methyl-1,4-naphthoquinone in anhydrous DMF (-408 mV and -1142 mV vs. NHE). The difference can once again be explained by interactions with the peptide backbone, or the possible presence of minute amounts of residual water in the experimental setup.³⁶

2.6 Conclusion

The aprotic cyclic voltammetry of heptaNaq matches very closely that of the model compound 2-methyl-1,4-naphthoquinone, suggesting, with the aqueous spectroelectrochemical behavior, that the overall design of Naq as a cofactor capable of robustly emulating natural quinone cofactors in respiratory and photosynthetic membrane oxidoreductases succeeded. In addition, it is clear that considering the range of redox potentials demonstrated by heptaNaq in both protic and aprotic mediums Naq may be a

important first step towards creating a family of electron transfer maquettes capable of demonstrating very low and moderate electrochemical behaviors not unlike the membrane complexes of photosynthesis. Indeed, since DMF has a significantly higher dielectric constant than is predicted for the average protein environment and there is a clear presence of protons in the ‘aprotic’ cyclic voltammetry of heptaNaq, the numbers obtained for the half-wave potentials represent an upper limit for the real value in the lower dielectric environment of a protein core. Nevertheless, even if the DMF values are accurate, it is

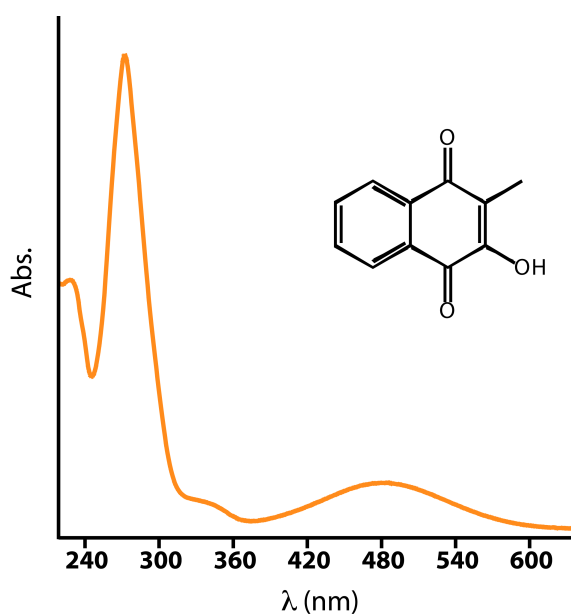


Figure 2.13. Absorption spectrum and structure of 2-hydroxy-3-methyl-1,4-naphthoquinone.

apparent that Naq may serve as a shuttle for low potential electrons capable of reducing hydrogen or other similarly stable high-energy fuels.

With the clear spectroscopy of freshly prepared Naq established, it is possible to observe a slow degradation of the cofactor on the days time scale at room temperature. The likely end product is an analog of 2-hydroxy-3-methyl-1,4-

naphthoquinone, a species that has a significantly different absorbance spectrum than Naq, Figure 2.13, with a notable fairly intense relatively red absorption band at 480 nm. Unfortunately, there is no good method of avoiding the hydroxylation degradation product as three distinct routes exist for its formation, namely direct nucleophilic modification of Naq by hydroxide or water, reaction of oxidized Naq with peroxide followed by a slow rearrangement (a reaction which was used to produce the reference compound in Figure

2.13), or reaction of reduced Naq with molecular oxygen reminiscent of the formation of natural quinoprotein cofactors. Stocks of activated Naq containing peptides are seemingly stable indefinitely at -80 °C, and last for at least six months at -20 °C; stocks of protected Naq(OMe/OMe) containing peptides are indefinitely stable under most reasonable conditions. Fortunately, the distinct spectroscopy of the degradation product allows a quick check for cofactor purity.

While the studies of heptaNaq confirmed adherence to design principles of the new quinone cofactor Naq, it should be clear that the choice of methyl protecting groups for the Naq side chain almost prevented a broad utility of the naphthoquinone amino acid. Ultimately, the methyl protecting groups allowed the completion of much of this work, though their deprotection was far less straightforward than initially believed. In the following chapter a more thorough discussion on the attempts to remove the methyl ether protecting groups will be presented as it amounted to a major stumbling block. Finally, in Chapter 6, a new, seemingly completely orthogonal protecting scheme will be presented.

2.7 References

- (1) Prince, R. C.; Dutton, P. L.; Bruce, J. M. *Febs Letters* **1983**, *160*, 273.
- (2) Fisher, N.; Rich, P. R. *J Mol Biol* **2000**, *296*, 1153.
- (3) Duine, J. A. *J. Biosci. Bioeng.* **1999**, *88*, 231.
- (4) Rinaldi, A. C.; Rescigno, A.; Rinaldi, A.; Sanjust, E. *Bioorg. Chem.* **1999**, *27*, 253.
- (5) Davidson, V. L. *Arch. Biochem. Biophys.* **2004**, *428*, 32.
- (6) Murakami, Y.; Tachi, Y.; Itoh, S. *Eur. J. Org. Chem.* **2004**, 3074.
- (7) Datta, S.; Mori, Y.; Takagi, K.; Kawaguchi, K.; Chen, Z. W.; Okajima, T.; Kuroda, S.; Ikeda, T.; Kano, K.; Tanizawa, K.; Mathews, F. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14268.
- (8) Schwartz, B.; Klinman, J. P. *Vitam. Horm.* **2001**, *61*, 219.
- (9) Wilmot, C. M.; Davidson, V. L. *Curr Opin Chem Biol* **2009**, *13*, 462.
- (10) Ono, K.; Okajima, T.; Tani, M.; Kuroda, S.; Sun, D.; Davidson, V. L.; Tanizawa, K. *Journal of Biological Chemistry* **2006**, *281*, 13672.

- (11) Aberg, A.; Ormo, M.; Nordlund, P.; Sjoberg, B. M. *Biochemistry-Us* **1993**, *32*, 9845.
- (12) Alfonta, L.; Zhang, Z.; Uryu, S.; Loo, J. A.; Schultz, P. G. *J Am Chem Soc* **2003**, *125*, 14662.
- (13) Seyedsayamdost, M. R.; Stubbe, J. *J. Am. Chem. Soc.* **2006**, *128*, 2522.
- (14) Seyedsayamdost, M. R.; Xie, J.; Chan, C. T.; Schultz, P. G.; Stubbe, J. *J Am Chem Soc* **2007**, *129*, 15060.
- (15) Burdine, L.; Gillette, T. G.; Lin, H. J.; Kodadek, T. *J. Am. Chem. Soc.* **2004**, *126*, 11442.
- (16) Burdine, L.; Kodadek, T. *Chem Biol* **2004**, *11*, 593.
- (17) Umanah, G. K.; Son, C.; Ding, F.; Naider, F.; Becker, J. M. *Biochemistry-Us* **2009**, *48*, 2033.
- (18) Umeda, A.; Thibodeaux, G. N.; Zhu, J.; Lee, Y.; Zhang, Z. *J. Chembiochem* **2009**, *10*, 1302.
- (19) Hay, S.; Wallace, B. B.; Smith, T. A.; Ghiggino, K. P.; Wydrzynski, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17675.
- (20) Li, W.; Heinz, J.; Haehnel, W. *J. Am. Chem. Soc.* **2005**, *127*, 6140.
- (21) Li, W. W.; Hellwig, P.; Ritter, M.; Haehnel, W. *Chem. Eur. J.* **2006**, *12*, 7236.
- (22) Hay, S.; Westerlund, K.; Tommos, C. *J. Phys. Chem. B* **2007**, *111*, 3488.
- (23) Berry, B. W.; Elvekrog, M. M.; Tommos, C. *J Am Chem Soc* **2007**, *129*, 5308.
- (24) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun (Camb)* **2009**, 168.
- (25) O'Donnell, M.; Bennett, W.; Wu, S. *J. Am. Chem. Soc.* **1989**, *111*, 2353.
- (26) Berk, S. C.; Yeh, M. C. P.; Jeong, N.; Knochel, P. *Organometallics* **1990**, *9*, 3053.
- (27) Tohma, H.; Morioka, H.; Harayama, Y.; Hashizume, M.; Kita, Y. *Tet. Lett.* **2001**, *42*, 6899.
- (28) Bulbule, V. J.; Koranne, P. S.; Munot, Y. S.; Borate, H. B.; Deshpande, V. H. *Synthetic Communications* **2003**, *33*, 587.
- (29) Corey, E. J., Noe, M. C. *Org. Synth.* **2003**, *80*, 38.
- (30) Rich, P. R. *Biochim. Biophys. Acta-Bioenerg.* **2004**, *1658*, 165.
- (31) Quan, M.; Sanchez, D.; Wasylkiw, M. F.; Smith, D. K. *J Am Chem Soc* **2007**, *129*, 12847.
- (32) Singh, I.; Ogata, R. T.; Moore, R. E.; Chang, C. W. J.; Scheuer, P. J. *Tetrahedron* **1968**, *24*, 6053.
- (33) Clark, W. M. *Oxidation/Reduction Potentials of Organic Systems*; The Williams and Wilkins Co.: Baltimore, 1960.
- (34) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097.
- (35) Cerda, J. F.; Koder, R. L.; Lichtenstein, B. R.; Moser, C. M.; Miller, A. F.; Dutton, P. L. *Org Biomol Chem* **2008**, *6*, 2204.
- (36) Prince, R. C.; LloydWilliams, P.; Bruce, J. M.; Dutton, P. L. *Methods in Enzymology* **1986**, *125*, 109.

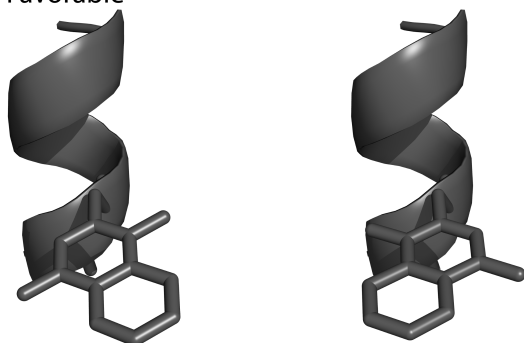
Chapter 3: Secondary Structure Effects on the Electrochemistry of the Naphthoquinone Amino Acid, Naq

3.1 Introduction

In the development of the naphthoquinone amino acid Naq the advantages of creating a quinone α -amino acid that could be installed in a peptide with limited conformational flexibility outweighed concern of any disruption of that peptide's secondary structure (Chapter 2). Care was taken to minimize the likely disruption caused by excessive steric bulk near the β -carbon by sacrificing some amount of chemical stability, but the overall design took into little account the need for side chain torsion angles (rotamers) that were compatible with the major secondary structural elements. Hand modeling of Naq into idealized and natural alpha helices suggested that while two of the four likely helical χ_1/χ_2 rotamers would be highly disruptive to the backbone due to steric occlusion with the 1-carbonyl (or 1-hydroxyl), the other two rotamers should be tolerated, Figure 3.1. However, this simple analysis ignores the possibility that Naq could stabilize local non-helical conformations to such an extent that helix formation would be unlikely, in essence, reducing

Naq's alpha helix propensity. For the remainder of the chapter the terms alpha helix and helix are used interchangeably. Much literature (see Pace and Scholtz's creation of a 'universal' alpha-helix propensity scale based upon prior measurements² and references therein), has focused on the establishment of helix propensity scales using an array of different approaches from early statistical analysis of structures of proteins³ to experimental systems with synthetic host-guest peptides. The experimental techniques used to ascertain the helical propensity of natural and artificial amino acids run the gamut from applying

Sterically Favorable



Sterically Hindered

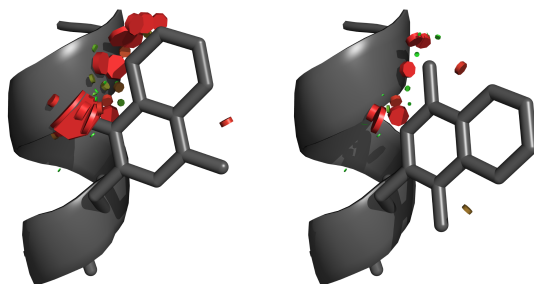


Figure 3.1. Likely helix rotamers of Naq. Steric conflicts are shown as red cylinders as generated by Pymol.

Makhatadze group took advantage of a family of short peptides that demonstrated very efficient metal dependent helix formation.⁸⁻¹⁰ This peptide family is based upon the calcium binding loop III of native rat testis calmodulin.¹⁰ The 12-residue peptide, P1 (sequence: Ac-

modified Lifson-Roig or Zimm-Bragg theories in conjunction with two-state folding models of monomeric helices^{4,6} to measurements of relative affinities during binding events to self⁷ or a metal^{8,9} that result in helix stabilization or formation of helices, respectively.

In one of the more recent contributions to the body of literature on the measurement of helical propensity, the

DKDGDGYISAAE-NH₂), contains a single point mutation that changed the native binding selectivity from calcium to lanthanide ions, with a binding affinity for the latter of $K_d=9.4 \mu\text{M}$.⁸ As with calcium induction of helix formation in EF-hand proteins, in the presence of saturating lanthanide ion, the final three residues of P1 form a rigid helix due to the terminal glutamic acid involvement in ligating the metal. Extension of this helix with an additional four residues (AAAQ) produced the peptide P2; the structure of which at 2 °C confirmed the helicity of the final seven or eight residues (the terminal glutamine in P2 may be unstructured), Figure 3.2. Lopez *et al.* further developed the peptide as a system for measuring the enthalpy and entropy of helix formation while taking advantage of the

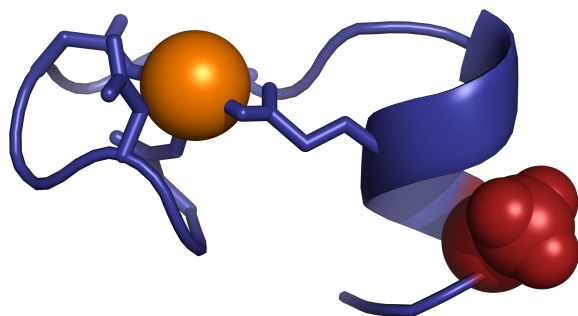


Figure 3.2. Representative solution structure of P2A; lanthanide ion in orange, guest site alanine in red.

lanthanide induced changes in CD signal to evaluate the actual change in helicity upon ligand binding.⁹ This formed the basis for use of the peptide P2 as a host system for measuring the detailed thermodynamics of

helix formation using several different amino acids at position 14, the so called P2X peptides (sequence: Ac-DKDGDGYISAAEAXAQ-NH₂, where **X** is a guest amino acid under study and in the lanthanide bound form points away from the peptide ‘core,’ Figure 3.2).⁸

Using isothermal titration calorimetry (ITC) and fluorescence titrations to follow the lanthanide binding response of the sole tyrosine residue, Richardson *et al.* evaluated the enthalpy and free energy of lanthanide binding for the P2X family (where X= A, F, G, V, L, I, J, B, N, Q, S, T -- and an observed binding range from 2.4 μM for P2A to 6.5 μM for

P2G).⁸ The enthalpy observations surprisingly revealed three distinct classes of amino acids, which are separated based upon high, intermediate, and low Δh_{α} ; it had been previously believed that the enthalpy of formation of a helix was driven by non-amino acid specific interactions like the backbone hydrogen bonds and van der Waals contacts. The classes follow two empirical rules, namely, if the amino acid is β -branched or can hydrogen bond to the backbone it belongs in class two. Non- β -branched aliphatic residues make up class one, and glycine is the sole member of class three. There does not appear to be a single explanation for the observed classes of amino acids. Hydration and the nature of the unfolded state are surely components, but the clear-cut observation of the significance of varying amino acid-specific enthalpic effects on helix formation is an important discovery in itself. More importantly, however, for the purposes of this work, was the observation that the ‘site-specific’ free energy ($\Delta\Delta G$) values obtained from the P2X peptides and the compiled propensity scale created by Pace and Scholtz² matched. This realization proved the utility of P2X as a straightforward thermodynamic host-guest system for accessing the helical propensity of artificial amino acids, including Naq.

In the work of the Makhatadze group, their treatment of the thermodynamic data played a critical role in the success of the P2X peptide as a system for measuring the energies of helical propensity and so it will be briefly summarized here with modification. The fluorescence titrations of the P2X peptides showed tighter binding affinities for the lanthanide ion than that of the lanthanide binding loop peptide P1. The energy difference between the P2X and P1 is the effective energy of helix formation in the last three or four residues of P2X

$$\Delta G(P2X) = -RT \ln \left(\frac{K_d(P2X)}{K_d(P1)} \right) \quad \text{Eq. 1}$$

where $K_d(P2X)$ is the measured dissociation constant of lanthanide ion for P2X, and $K_d(P1)$ is the measured dissociation constant for P1, 9.4 μM . Since the P2X peptides demonstrated some amount of helicity in the absence of lanthanide, this energy is corrected to account for the changes in helicity, assuming the lanthanide bound form is fully helical

$$\Delta G_o(P2X) = \frac{\Delta G(P2X)}{1 - F_\alpha^-(P2X)} \quad \text{Eq. 2}$$

where $F_\alpha^-(P2X)$ is the fraction helical in the unbound peptide derived from ellipticities of experimental and reference states

$$F_\alpha^- = \frac{[\Theta(P2X)]_c - [\Theta^-(P2X)]}{[\Theta(P2X)]_c - [\Theta(P2X)]_b} \quad \text{Eq. 3}$$

The value for $[\Theta(P2X)]_b$ represents the ellipticity of a fully helical control, in this case the lanthanide bound form which is assumed to be >95% helical, and $[\Theta^-(P2X)]$ is the measured ellipticity for the lanthanide unbound form. $[\Theta(P2X)]_c$ is

$$[\Theta(P2X)]_c = \frac{13 \cdot [\Theta^-(P1)] + 4 \cdot [\Theta(RCC)]}{17} \quad \text{Eq. 4}$$

where $[\Theta^-(P1)]$ is the ellipticity of P1 in the absence of lanthanide, and $[\Theta(RCC)]$ is the ellipticity of a random coil control peptide designated P6N5 in Makhatadze's group that represents the final four residues in P2X. The weights (13, 4 and 17) arise from the circular dichroism formalism of counting backbone peptide units including the N-terminal acetyl and C-terminal amide. $[\Theta(P2X)]_c$ can be thought to represent a mean molar residue ellipticity

spectrum of an idealized random coil form of the P2X peptide and thus $F_{\alpha}^{-}(P2X)$ is an accurate measure of the fraction of helix in the lanthanide unbound form assuming the lanthanide bound form actually represents $[\Theta(P2X)]_b$. In the Makhatadze group's treatment of their P2X titration data, it was assumed that the globally responsive circular dichroism spectrum giving the fractional helicity in the unbound form adequately represented the local properties of the amino acid in the guest site, an assertion they defended by matching the measured fractional helicity with that calculated using AGADIR, a program based upon empirical derived helix formation energies refined using observed fractional helicities obtained from CD spectra.^{8,11} In any case, it is clear that if the lanthanide bound form does not represent the true measure of full helicity then it is necessary to find an alternative measure of $[\Theta(P2X)]_b$ and the equation for $\Delta G_o(P2X)$ must be modified

$$\Delta G_o(P2X) = \frac{\Delta G(P2X)}{F_{\alpha}^{+}(P2X) - F_{\alpha}^{-}(P2X)} \quad \text{Eq. 5}$$

where $F_{\alpha}^{+}(P2X)$ is the fractional helicity of the lanthanum bound form. Considerations such as these were used by the Bierzynski group¹² in explaining slight disparities being their observations of the enthalpy of helix formation and those of the Makhatadze group in the same peptide family.⁹

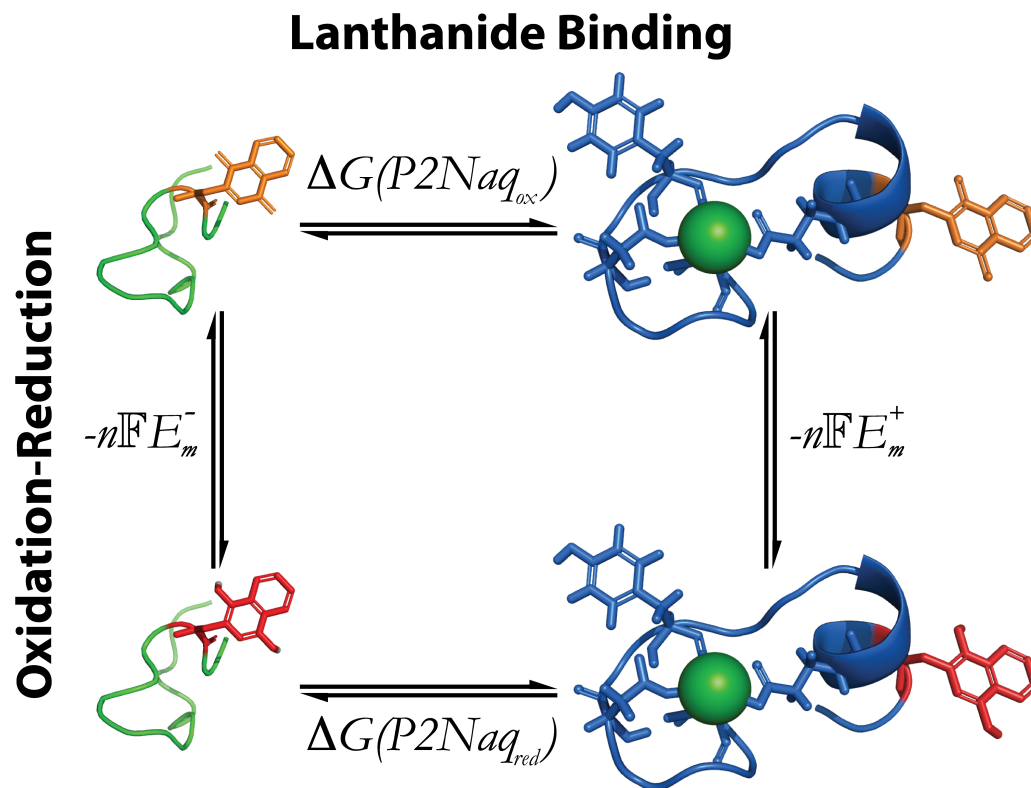


Figure 3.3. The lanthanide binding coupled electrochemistry of P2Naq

3.2 P2Naq Thermodynamics

The discussion thus far has ignored that Naq has two stable thermodynamic states at neutral pH. In essence, P2Naq, the P2X peptide with Naq substituted into the guest site, actually allows the measurement of the helical propensity of both forms of the amino acid, and there is reason to believe they may be different as their abilities to hydrogen bond the backbone are distinct. The binding of lanthanum to reduced and oxidized P2Naq (P2Naq_{red} and P2Naq_{ox}, respectively) is thermodynamically coupled to the midpoint potentials of the bound and unbound forms of P2Naq, generating a thermodynamic box, Figure 3.3. The P2Naq system actually allows the measurement of the change in redox potential associated with a change in secondary structure of Naq. In an ideal system, where the lanthanide bound

form of the peptide is fully helical and the unbound form has no helical character, the measure of this value is straightforward

$$\Delta\Delta G_{\alpha c} = -n\mathbb{F}\Delta E_{m,\alpha c} = -n\mathbb{F}E_m^+ + n\mathbb{F}E_m^- = \Delta G(P2Naq_{red}) - \Delta G(P2Naq_{ox}) \text{ Eq. 6}$$

where $\Delta\Delta G_{\alpha c}$ is the free energy change from coil to helix, n is the number of electrons associated with Naq reduction ($n = 2$), \mathbb{F} is Faraday's constant ($23.06 \text{ kcal mol}^{-1} \text{ V}^{-1}$), $\Delta E_{m,\alpha c}$ is the redox midpoint change of Naq going from coil to helix, E_m^+ and E_m^- represent the lanthanide bound and unbound forms assumed to represent fully helical and fully coil, respectively, and $\Delta G(P2Naq_{ox})$ and $\Delta G(P2Naq_{red})$ represent the lanthanide binding energies of the oxidized and reduced form of P2Naq corrected for the binding energy of P1 to lanthanide as in **Eq. 1**. Unfortunately, the system is not ideal and the need to correct the thermodynamic values to account for fractional helicity would prevent direct use of measured redox midpoints. Combining fractional helicities with experimentally determined midpoints makes little thermodynamic sense since midpoint potentials represent the energy of a half-cell and are meaningless without a coupled redox reaction. Nevertheless, assuming a two-state model, the expected change in midpoint potential for the coil to helix transition can be calculated via binding affinity differences

$$\Delta\Delta G_{\alpha c} = -n\mathbb{F}\Delta E_{m,\alpha c} = \frac{\Delta G(P2Naq)_{red}}{F_{\alpha}^+(P2X)_{red} - F_{\alpha}^-(P2X)_{red}} - \frac{\Delta G(P2Naq)_{ox}}{F_{\alpha}^+(P2X)_{ox} - F_{\alpha}^-(P2X)_{ox}} \text{ Eq. 7}$$

where $F_{\alpha}^+(P2X)_{ox}$ and $F_{\alpha}^-(P2X)_{ox}$ represent the fractional helicity in the oxidized lanthanide bound and unbound forms, respectively, and $F_{\alpha}^+(P2X)_{red}$ and $F_{\alpha}^-(P2X)_{red}$ represent the fractional helicity in the reduced lanthanide bound and unbound forms, respectively. $\Delta\Delta G_{\alpha c}$ and $\Delta E_{m,\alpha c}$ calculated as such serve as global measures of helical

propensity for both forms of Naq. These values are not directly comparable to the values of helical propensity, which are generally standardized using helix formation energies of peptides with alanine in the guest site; yet it is an effective measure of the energetic coupling of the coil to helix transition to electrochemical changes in Naq and should represent an energy of disruption or stabilization of a helix that would occur on changing Naq's redox state. The closest observation of a change in an independently measurable physical property with natural amino acids would be the observed rotamer dependent fluorescence lifetimes of tryptophan and tyrosine.¹³⁻¹⁵

Without making a large number of assumptions, there is perhaps no good way to adequately measure the lanthanide bound fraction helicity, $F_{\alpha}^{+}(P2X)_{ox}$, and similar values site-specifically with circular dichroism, with single site aromatic amino acid dichroism being a possible and noteworthy exception. Model states and thermal unfolding assumptions are perhaps good enough for a global estimate, but to determine site-specific probabilities another technique is needed. The chemical shifts of the $^1\text{H}^{\alpha}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, and ^{13}CO atoms in amino acids have long been known to be responsive to secondary structure.^{16,17} With selective labeling, it is therefore possible to obtain an unbiased measure of helical content in a site-specific manner by observing the corresponding chemical shifts of the amino acid in question

$$F_{\alpha} = \frac{\delta_{obs} - \delta_c}{\delta_h - \delta_c} \quad \text{Eq. 8}$$

where δ_{obs} is the observed chemical shift for the atom in question in the experimental peptide, and δ_h and δ_c are the reference chemical shifts for helix and coil form of the same atom, respectively. As with the circular dichroism based fraction helicities, the choice in helix

and coil controls is vitally important. Fortunately, under most circumstances the long distance effects on chemical shifts in NMR are minimal, and the local environment can be approximated or even copied by directly repeating the local sequence of the experimental system into controls designed to take on the helix and coil geometries. This technique, for instance, has been used to great effect by the Waters' lab in their thermodynamic studies of cation- π interactions in β -hairpins.¹⁸

3.3 Synthesis, Deprotection and Initial Characterization of P2Naq

As previously mentioned, P2Naq is based upon the P2X peptide family used by the Makhataдзе group to measure the enthalpic contributions of various amino acids at a guest site in the formation of a helix.⁸ Its synthetic sequence is Ac-DKDGDGYISAAEA \mathbf{Q} (OMe/OMe)AQ-NH₂, where \mathbf{Q} once again represents Naq. Fmoc-solid phase peptide synthesis was once again straightforward; an increased yield was obtained by synthesizing the isoleucine-serine sequence into the growing peptide chain using a Fmoc-protected pseudo-proline derivative, where the serine residue side chain hydroxyl is reversibly bound to its own amine with an acid labile dimethylacetal group thereby reducing the likelihood of premature termination of chain extension. Additionally, when synthesized on a CEM Liberty microwave synthesizer, it was necessary to switch from the common piperidine Fmoc deprotection reagent to the less basic piperazine to circumvent succinimide formation between Asp₅ and Gly₆, increasing the overall yield of the peptide substantially. Peptide identity was confirmed via MALDI-TOF-MS. It is of note that purified and lyophilized P2Naq(OMe/OMe) is more or less completely insoluble in any solution besides at least 4 M GdnHCl or neat TFA, or mixtures of the two (high concentrations of urea is

possibly effective, but it was not tested). When added to aqueous solutions that do not contain high denaturant concentrations, the peptide forms hard, millimeter scale gel-like aggregates that are transparent with no apparent scattering from the bulk solution. They can, however, be recovered quickly using neat TFA or slowly adding >4 M GdnHCl solutions. In addition to P2Naq(OMe/OMe), peptides P1 and P2A, as described above, were synthesized under the same conditions.

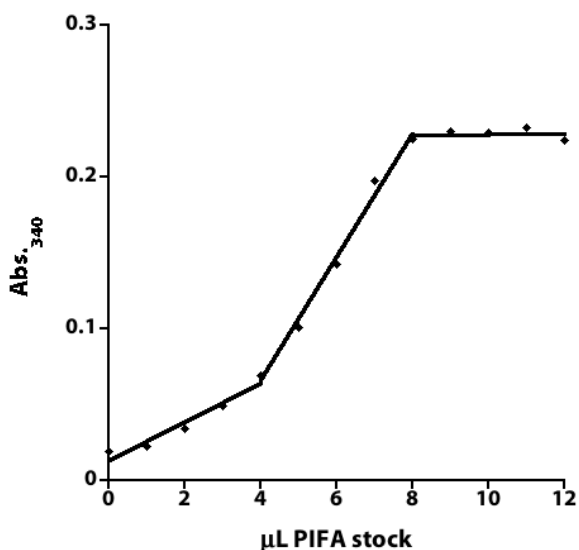


Figure 3.4. Titration of P2Naq(OMe/OMe) with substoichiometric aliquots of PIFA. One equivalent is added before oxidation of Naq(OMe/OMe) is evident as observed at 340 nm. The lines are drawn to guide the eye.

3.3.1 Deprotection/Activation of

P2Naq(OMe/OMe)

Despite the success of using the hypervalent iodide(III) oxidant PIFA to activate Naq in Ac-Naq(OMe/OMe)-OMe, all attempts to similarly deprotect P2Naq(OMe/OMe) failed. Indeed, while initial experiments were performed with blind excess of the PIFA reagent resulting in indistinct

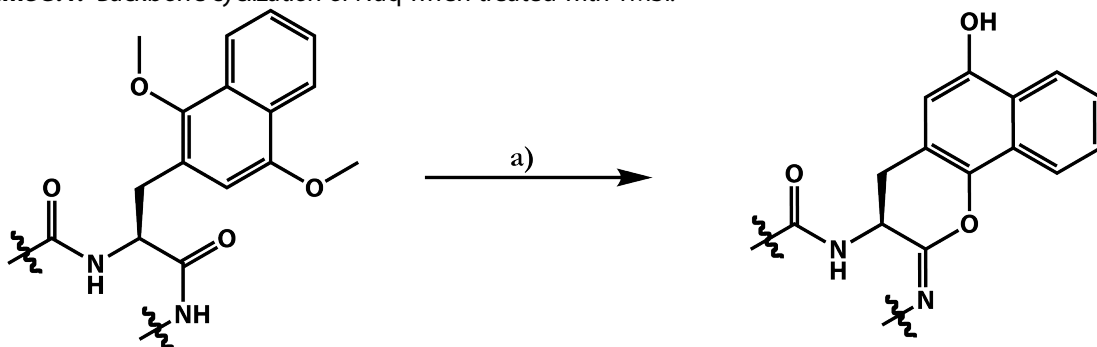
products, none of which had an obvious, non-aggregate naphthoquinone spectrum, even careful titrations of PIFA into a solution of P2Naq(OMe/OMe) showed an unexpected one equivalent lag of in the formation of an activated Naq product, Figure 3.4. The best explanation was that another species in solution, presumably the Tyr₇ of P2Naq(OMe/OMe), was reacting significantly faster with PIFA than Naq(OMe/OMe) since

oxidation is effected through a covalent adduct between the oxidant and its target and tyrosine is less sterically occluded than Naq(OMe/OMe).

Since it was unknown at the time whether the observed phenomena was going to be true for all oxidants commonly used to activate methyl-ether protected quinones suggesting that the tyrosine is far more reductive than Naq(OMe/OMe) is, other common methods of removing methyl-ether protecting groups were explored, all of which can be described as push-pull nucleophilic displacement reagents, where an oxophilic Lewis acid activates the methoxy sufficiently to allow a nucleophile to displace the methyl group freeing the hydroxyl. The vast majority of the reagents attempted, from BF₃-methylsulfide or AlCl₃/EDT in CH₂Cl₂, both of which only partially deprotected Naq(OMe/OMe) even though they were in vast excess, to TMSBr/EDT/thioanisole/m-cresol, which showed no appreciable deprotection, were completely unsatisfactory. There were two exceptions to this observation. In the case of triflic acid/thioanisole in TFA, there was clear removal of the methyl protecting groups from Naq(OMe/OMe), but it was met with a mixture of backbone hydrolysis products around Naq; as if Naq, once deprotected, was involved in the pseudo-catalytic hydrolysis of peptide bonds before and after itself. On the other hand, using TMSI/EDT/thioanisole in TFA to deprotect P2Naq(OMe/OMe) gave a single predominant product with a very clear reduced spectrum not unlike that of the protected Naq(OMe/OMe). Mass spectroscopy, both MALDI-TOF-MS and ESI-MS, revealed a dehydration reaction had occurred, but this was not all together unexpected given the number of amides in the peptide. However, the product was not redox active as purified, and only displayed redox activity after being accidentally exposed to a highly acidic aqueous solution. Given the observations, it is believed that the dehydration product was actually a

cyclization of reduced Naq to the peptide backbone, Scheme 3.1. The heptaNaq project described in Chapter 2 was conceived to untangle the various reactions that were complicating the selective deprotection of Naq in P2Naq. Subsequent attempts to further fine-tune the reactivity of the push-pull systems proved difficult; the use of TMS-triflate in TFA, for instance, generated both presumptive backbone cyclization and hydrolysis products, while using B(OTFA)₃ in TFA in a large excess seemingly hydrolyzed the backbone of an extended helix from the N- and C- terminals and, though interpretation of the MALDI-TOF spectra of the HPLC purified products was difficult, did not appear to deprotect Naq(OMe/OMe).

Scheme 3.1. Backbone cyclization of Naq when treated with TMSI.



a) TMSI, EDT, Thioanisole, TFA

The success of using stoichiometric amounts of cerium ammonium nitrate (CAN) to deprotect heptaNaq(OMe/OMe) in the presence of heptaTyr (Chapter 2), provided the basis for successfully deprotecting P2Naq(OMe/OMe). Treatment of P2Naq(OMe/OMe) in 8 M GdnHCl with 100 mM TFA, see discussion above concerning the solubility of P2Naq(OMe/OMe), with 2-3 equivalents of CAN at room temperature for twenty minutes yields deprotected P2Naq_{ox} in a ~60% overall HPLC yield; the remaining 40% is made up of ~20% unactivated P2Naq(OMe/OMe) and ~20% unidentified peptide byproducts. Above 3

equivalents and the reaction specificity for Naq(OMe/OMe) declined rapidly. It was observed that the unidentified byproduct percentage tracked with the amount of unconverted starting material depending upon how quickly the CAN was added to the solution, which flashes brown after CAN addition, suggesting that a slightly less reactive oxidant might be effective in increasing the yield and specificity.

Having exhausted the readily available or easily produced single electron moderate oxidants and found none to be sufficient, it was discovered that there is a commonly used obligate two-electron quinone, dichlorodicyanoquinone (DDQ), in organic synthesis whose redox potential was sufficiently above that of the single electron oxidants attempted. Unfortunately, DDQ is not long lived in aqueous solutions, but it does dissolve in acetonitrile (MeCN) and is stable in TFA; thus, reaction of P2Naq(OMe/OMe) in neat TFA with precisely one equivalent of DDQ afforded the deprotected P2Naq_{ox} in ~80% yield after purification by HPLC. The yield was increased to be effectively quantitative when the reaction is kept at 0 °C for at least 30 minutes. Slight excess DDQ does not appear to oxidize the peptide further, but rather, appears to form a covalent adduct to the peptide and makes purification significantly more difficult. The reaction of Naq(OMe/OMe) with DDQ is notably slower than the reaction with CAN and stays brown for tens of minutes before becoming slightly yellow, indicating long lived formation of semiquinones or charge transfer complex between DDQ and Naq. Some amount of the selectivity can probably be explained by the much reduced effective redox potential of DDQ compared to CAN, but in addition, it seems apparent that the two-electron nature of DDQ oxidation may actually bias reaction towards Naq(OMe/OMe) over Tyr₇, which is significantly less likely to be doubly oxidized.

3.3.2 Spectroscopy of P2Naq

Like protected heptaNaq(OMe/OMe), it became necessary in the course of study to determine the molar absorptivity of P2Naq_{ox} to make possible the determination of concentrations of experimental solutions. The Kjeldahl total nitrogen of solutions of P2Naq(OMe/OMe) with known spectra were determined and the molar absorptivity spectrum calculated, Figure 3.5. As with the spectroscopy of heptaNaq (Chapter 2), the

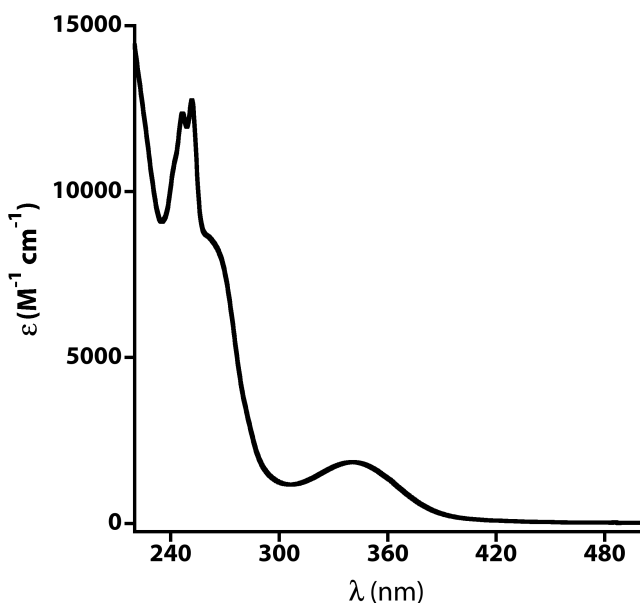


Figure 3.5. Molar absorptivity of P2Naq in 8 M GdnHCl, 100 mM TFA. See text for details.

spectrum is does not deviate from the spectroscopy of known naphthoquinones, with features lying between those of 2-methyl-1,4-naphthoquinone and 1,4-naphthoquinone. As anticipated, there is only slight evidence that the tyrosine chromophore is present.

The molar absorptivity at 252 nm is 12700 M⁻¹ cm⁻¹, and at 340 nm is 1840 M⁻¹ cm⁻¹. The band at 340 nm is

sufficient to measure the concentration of most peptides containing activated Naq since it does not overlap with any natural amino acids. In situations where the peptide concentration is too low to measure the 340 nm band reliably, the tyrosine absorbance at 252 nm is 316 M⁻¹ cm⁻¹,¹⁹ which represents an insignificant 2.4% adjustment so the value 12700 M⁻¹ cm⁻¹ can be used without modification.

3.3.3 Borohydride Oxidase Activity of P2Naq

Like most natural quinones and the peptide heptaNaq, Naq in P2Naq_{ox} can be reversibly reduced by borohydride; oxidation of P2Naq_{red} by molecular oxygen reverses the process effectively making P2Naq a catalyst for the oxygen dependent oxidation of borohydride, Figure 3.6. Depending upon the steric occlusion of the Naq cofactor, the reaction is likely to be slightly irreversible as covalent

modification of reduced Naq by molecular oxygen is a possible path given the biologically important reaction of reduced naphthoquinone vitamin K with molecular oxygen in the cross linking of interstitial tissue.^{20,21}

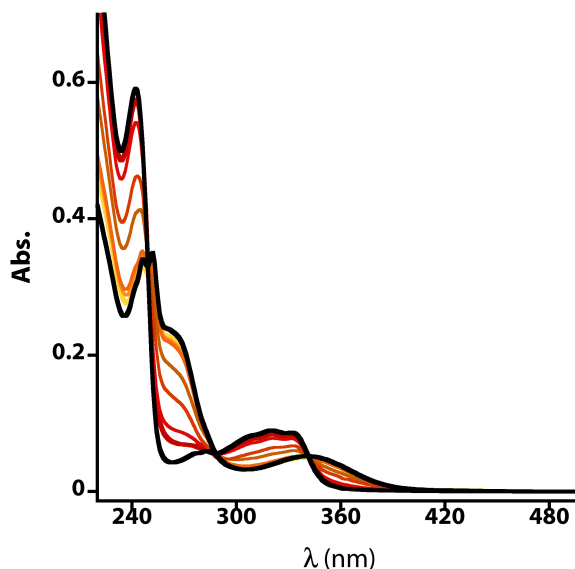


Figure 3.6. Borohydride reduction of 28 μM P2Naq in 20 mM Cacodylate, 100 mM KCl, pH 6.9, followed by reoxidation of P2Naq_{red} by O₂. Black spectra are of initial oxidized and reduced samples. Red to yellow spectra were taken at various time points after reduction.

3.4 Synthesis of Nitrogen Labeled Peptides for Site-Specific Helicity by NMR

As described in the Section 3.2 on P2Naq thermodynamics, it is possible to use chemical shift measurements to ascertain the fraction helicity in a site-specific manner. Unfortunately, while highly desirable for a full description of the conformation of Naq in any peptide, a full labeling the backbone and C ^{β} is not possible in any cost effective manner, rather a choice must be made between ¹⁵N labeling the amine or ¹³C labeling the C ^{α} , both of which provide convenient handles with which to determine the chemical shifts of ¹³C ^{α} and

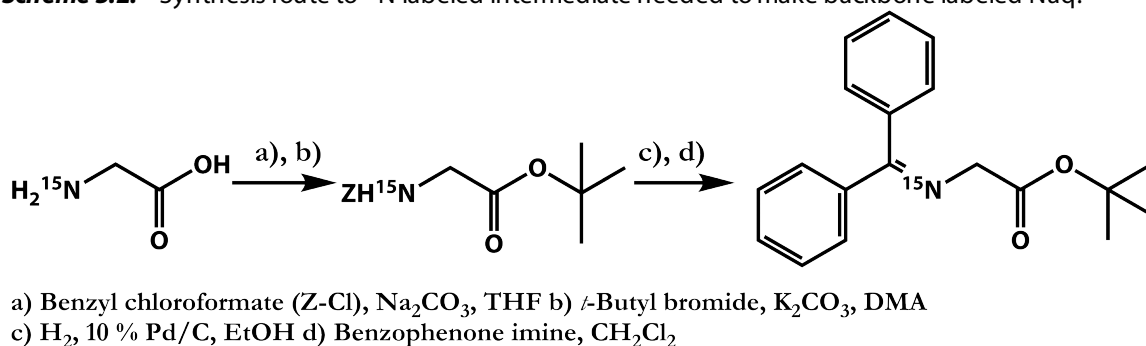
$^1\text{H}^\alpha$ through selective pulse sequences. These chemical shift values can then be used calculate F_α , the fractional helicity. It is also clear that focusing on Naq alone in the guest site of P2Naq might not provide enough information about the general conformation of the peptide around Naq to give a clear gauge of what is actually happening in Naq's local environment. Therefore, it seemed prudent to label the preceding and following residues as well, both alanines in P2Naq. The importance of tracking the identity of the amino acids as their chemical shifts change from experimental peptide to experimental peptide cannot be overstated. While intuitive tracking of the chemical shifts is possible in the absence of clear connectivity data, ^{15}N labeling the amine provides the best experimental route to not depending upon inference in that it is possible to use a nitrogen filtered HNCA experiment to get backbone connectivities in generic peptides by revealing cross peaks not only between the amide proton and the C^α of a residue but also between the amide proton and the C^α of the $i-1$ residue. Additionally, with ^{15}N labeling, a nitrogen filtered NOESY allows walking the backbone amides in helices. Pulse sequences that provide the same information do not readily exist for peptides incorporating $^{13}\text{C}^\alpha$ labeling alone. ^{15}N labeling also provides an opportunity to use the nitrogen filtered HNHA for measuring J-coupling between the amide proton and $^1\text{H}^\alpha$ and, thus, the ϕ , CO-N- C^α -CO, dihedral angle giving another physical parameter about the nature of the peptide backbone in the environment around Naq.

3.4.1 Synthesis of ^{15}N labeled Fmoc-Naq(OMe/OMe)-OH

In synthesizing either ^{15}N or $^{13}\text{C}^\alpha$ labeled Naq, no major changes in the synthetic route need to be considered. The ease of synthesis of the ^{15}N or $^{13}\text{C}^\alpha$ labeled backbone

precursor *N*-(diphenylmethylene)glycine *tert*-butyl ester from ^{15}N or $^{13}\text{C}^\alpha$ labeled glycine allows direct incorporation of the labels into the synthesis as it currently exists, Scheme 3.2.²² Thus ^{15}N Fmoc-Naq(OMe/OMe)-OH was synthesized by incorporating the ^{15}N -*N*-(diphenylmethylene)glycine *tert*-butyl ester into the synthesis reported²³ and described in Chapter 2.

Scheme 3.2.²² Synthesis route to ^{15}N labeled intermediate needed to make backbone labeled Naq.



3.4.2 Synthesis of ^{15}N labeled P2Naq(OMe/OMe), Random Coil Control and Helical Control

As described above, the sequence of ^{15}N labeled Ala₁₃, Naq₁₄ and Ala₁₅ were incorporated into P2Naq as well as controls for random coil and helical conformation; the sequences of the three peptides are shown in Table 3.1. ^{15}N labeled P2Naq (sequence Ac-DKDGDGYISAAEA**QAQ**-NH₂, where the emboldened amino acids are ^{15}N labeled) was synthesized and deprotected with DDQ as described above. Unlike the Makhatazde random coil control P6N6, which had very little sequence similarity to the terminal sequence of P2A,⁹ the random coil control (RCCNaq) was designed to be a nearly one-for-one duplication of the terminal six amino acids (sequence Ac-AEA**QAE**-NH₂). The mutation of the terminal glutamine residue to a glutamate was made to increase solubility of the resulting peptide and decrease the likelihood of any short helix formation as there should be an *i*-

>i+4 side chain charge-charge repulsion. Deprotection of RCCNaq(OMe/OMe) with DDQ as described proceeded smoothly producing a peptide with a very clean characteristic oxidized Naq spectrum.

Table 3.1. Primary sequence of peptides under study. Amino acids in bold are to be labeled for NMR studies; the numbering scheme for those amino acids is preserved from the longer peptides in the shorter RCCNaq.

| Peptide | Sequence |
|----------|---|
| P2Naq | Ac-DKDGDGYISAAEA ₁₃ Q ₁₄ A ₁₅ Q-NH ₂ |
| RCCNaq | Ac-AEA ₁₃ Q ₁₄ A ₁₅ E-NH ₂ |
| P2NaqExt | Ac-DKDGDGYISAAEA ₁₃ Q ₁₄ A ₁₅ QAAAAEAAAAEAAAAE-NH ₂ |

Richardson *et al.* assumed that the lanthanide ion bound P2X peptides were fully helical,⁸ arguing that given the strength of helix formation in the P2A peptide upon lanthanide binding as evidenced by both NMR and CD experiments, the P2X peptides should have similarly efficient structural transitions. Unfortunately, with the concern that Naq in both its reduced and oxidized forms might significantly cap helix formation through hydrogen bonds between the backbone amides and side chain hydroxyls and carbonyls, respectively, this assumption appeared illegitimate. In designing a better representation of the helical conformation, sequence identity was preserved beyond the terminal six amino acids as in the random coil peptide. Indeed, the entire sequence of P2Naq was copied wholesale and appended to 15 residues of one of Baldwin's 'unusually stable' monomeric alanine-based alpha-helical peptides.²⁴ Although only sixteen residues in length, the Baldwin peptides displayed upwards of 80% helicity under ideal conditions (low temperature, high ionic strength). The best designs had a pentameric repeating unit of four alanines and a single charged residue. This design assured that the helices did not aggregate in solution since the charged residues would be almost evenly dispersed around the helix (every 140°,

assuming 100° per residue). Since P2Naq is highly negatively charged without a ligand, the Baldwin helix chosen to append for the helical control peptide was similarly charged giving the sequence of the helical control termed P2NaqExt as Ac-DKDGDGYISAAEA**QA**QAAAAEAAAAEAAAAE-NH₂. With the added extended helical sequence, it was hoped that this peptide would be sufficient to induce Naq into a helical conformation, if possible. P2NaqExt was readily synthesized with acetic anhydride capping after each amino acid and deprotected with DDQ. Again, spectroscopy revealed a clear unambiguous Naq signal.

3.5 Circular Dichroism of P2Naq, RCCNaq, and P2NaqExt

The P2X family of proteins demonstrates clear lanthanide ion dependent circular dichroism (CD) changes in the amide region associated with helix formation. This fact was used by Richardson *et al.* to measure fractional helicity in lanthanide free P2X to correct the measured ‘helix formation’ energies for calculation of the site-specific enthalpy and free energy contributions to helix formation for the guest amino acid according to **Eq. 2**. In addition to providing one measure of alpha-helical content in the lanthanum unbound state, circular dichroism provides a convenient method for confirming gross properties of P2Naq, RCCNaq and P2NaqExt.

While it is certainly clear that the CD spectroscopy of the oxidized peptides should be straightforward, the CD spectroscopy of the reduced peptides presents somewhat more of a challenge. As will be discussed in greater detail in Section 3.6, the choice in reductants is not obvious. The standard reductant dithionite is known to extensively modify quinones, and indeed, there is evidence that Naq becomes sulfonated in the presence of dithionite.

Although as noted above, borohydride or related borane compounds function reasonably well, it is all but impossible use sufficient quantities of excess borohydride in high concentration samples to assure maintenance of complete reduction. In the work on developing methods for measuring the binding constant of P2Naq_{red} to lanthanide ion, it was discovered that hydrogen gas (H₂) can be used in association with platinum nanoparticles to reduce both P2Naq and the benzyl viologen redox dye, which serves as both a redox buffer and a convenient visible measure of the redox state of P2Naq since it is significantly more reducing than Naq is likely to be. One drawback of using benzyl viologen is that it has a high molar absorptivity in the far-UV region used to observe protein amide CD signals, so there is a small decrease in the quality of the spectra for the reduced samples because of this. All samples were measured on an Aviv 410 Circular Dichroism Spectrometer in triplicate at 25 °C from 200 nm to 260 nm. The wavelength was limited to this range because of technical challenges of acquiring data below 200 nm.

3.5.1 Circular Dichroism of P2Naq_{ox} and La³⁺- Bound P2Naq_{ox}

The circular dichroism spectrum of P2Naq_{ox} was measured in the far-UV at a variety of concentrations averaged, Figure 3.7A. The spectrum appears to be a mixture of a small amount of helix in a random coil background; not qualitatively different than that reported for P2A.¹⁰ There was no observed concentration dependence on the mean residue molar ellipticity of P2Naq_{ox} at 217 nm, -1930 mdeg cm² dmol⁻¹, Figure 3.7. The averaged CD spectrum of lanthanide ion bound P2Naq_{ox} shows a dramatic shift towards a slightly more alpha helical spectrum with a significant increase in the mean residue molar ellipticity at 217 nm to -6800 mdeg cm² dmol⁻¹, Figure 3.7A. However, when this mean residue molar

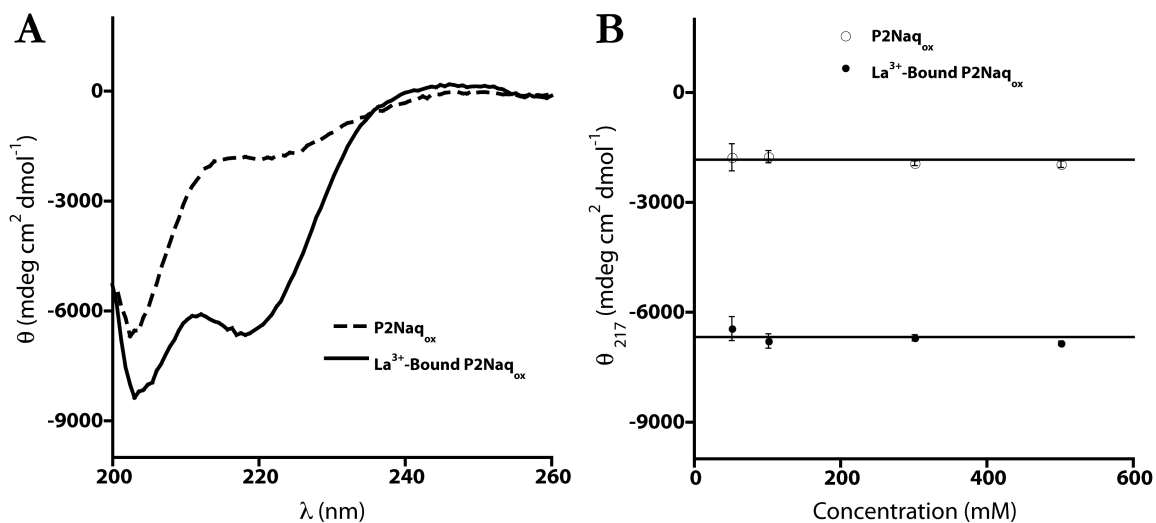


Figure 3.7. CD spectra of apo and holo P2Naq_{ox} (A). Concentration dependence of CD signal at 217 nm for apo and holo P2Naq_{ox} shown with experimental errors and lines drawn through the average (B). CD spectra were taken at 25 °C in 20 mM Cacodylate, 100 mM KCl, pH 6.9.

ellipticity is compared to the reported values for the P2A peptide, approximately -11000 mdeg cm² dmol⁻¹,^{9,12} it is evident that the helix is not forming entirely upon lanthanide ion binding. Qualitatively, when compared to the spectrum of lanthanide bound P2Naq_{ox}, the spectrum of lanthanide bound P2A^{9,12} appears to have significantly less of the strong random coil-associated signal near 200 nm. These observations lend credence to the idea that lanthanide bound P2Naq might not be fully helical. No concentration dependence of the mean residue molar ellipticity of lanthanide bound P2Naq_{ox} was found, Figure 3.7B.

3.5.2 Circular Dichroism of P2Naq_{red} and La³⁺- Bound P2Naq_{red}

As above, CD spectra were measured of the ligand free P2Naq_{red} and lanthanide ion bound P2Naq_{red}. In the absence of lanthanide the CD spectrum of P2Naq_{red} appears similar to that of unbound P2Naq_{ox}, but with a notable decrease in the signal at 217 nm, Figure 3.8A, -1200 mdeg cm² dmol⁻¹. While there may be a slight concentration dependence at 217 nm, the values measured for the mean residue molar ellipticity of P2Naq_{red} are within error

of each other, Figure 3.8B. Upon being saturated with lanthanide ion, the CD spectrum of P2Naq_{red} changes radically as with P2Naq_{ox}, Figure 3.8A. Similar to P2Naq_{ox} the mean molar residue ellipticity at 217 nm is -6920 mdeg cm² dmol⁻¹ suggesting that although the C-terminals of both peptides may not becoming entirely helical the same fraction of helix is forming. With the exception of an anomalously low measured mean residue molar ellipticity for the sample at 50 μM, there does not appear to be a concentration dependence of the CD signal for lanthanide bound P2Naq_{red}, Figure 3.8B. If investigated closely, there is an apparent positive signal present at ~245 nm in the lanthanide bound P2Naq_{red} CD spectrum that is not present in the spectrum of lanthanide bound P2Naq_{ox} or in the spectrum of ligand free P2Naq_{red}. The correlation of this observed band's maxima with the approximate absorption maxima of the reduced Naq did not go unnoticed.

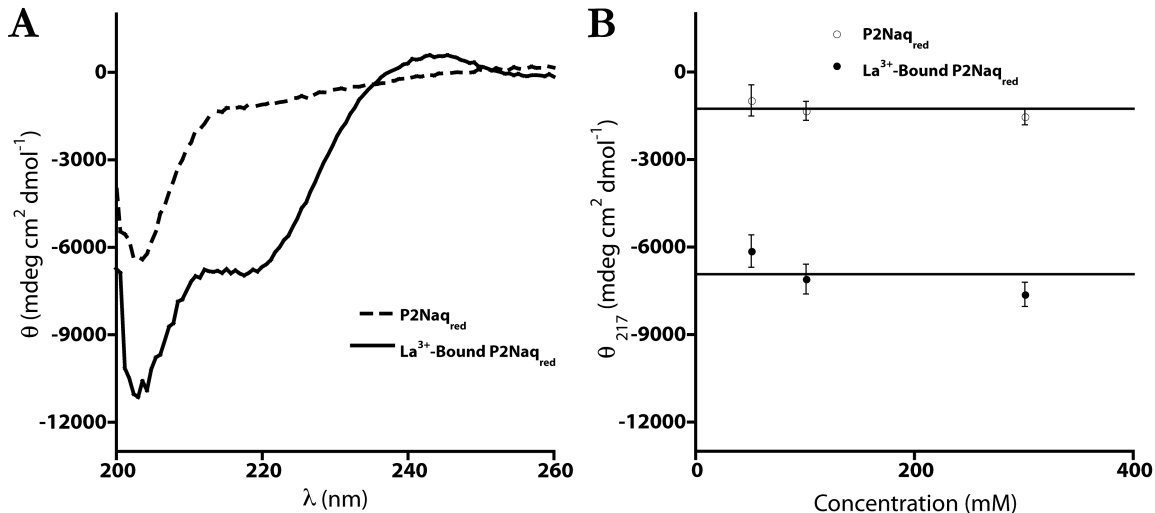


Figure 3.8. CD spectra of apo and holo P2Naq_{red} (A). Concentration dependence of CD signal at 217 nm for apo and holo P2Naq_{red} shown with experimental errors and lines drawn through the average (B). CD spectra were taken at 25 °C in 20 mM Cacodylate, 100 mM KCl, pH 6.9.

3.5.3 Circular Dichroism of RCCNaq_{ox} and RCCNaq_{red}

The CD spectrum of RCCNaq_{ox} confirms its design as a random coil standard, revealing no clear conformational bias, Figure 3.9A; and while the signals were smaller than the noise especially at the lower concentrations of peptide, there was no clear concentration dependence of the CD signal at 217 nm, Figure 3.9B. On the other hand, the CD spectrum of RCCNaq_{red} is even more featureless, Figure 3.9A; and increased noise is apparent in the plotting the concentration dependence, Figure 3.9B. With such small signals, it appears that the addition of benzyl-viologen to the sample, coupled with the higher molar absorptivity of the reduced Naq cofactor and the presence of amides in the protective polymer of the colloidal platinum, may have significantly decreased sample quality; yet the differences between the two samples at wavelengths associated with coil-to-helix transition in the P2Naq peptides are small (at 217 nm, ~ 370 mdeg cm² dmol⁻¹, which is approximately the average error for the two samples).

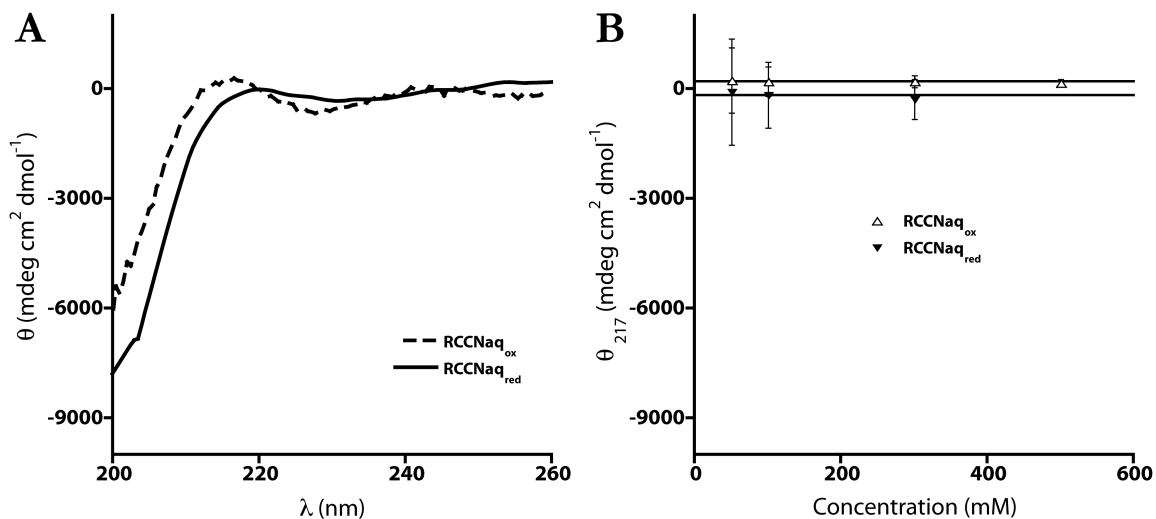


Figure 3.9. CD spectra of RCCNaq_{ox} and RCCNaq_{red} (A). Concentration dependence of CD signal at 217 nm for RCCNaq_{ox} and RCCNaq_{red} shown with experimental errors and lines drawn through the average (B). CD spectra were taken at 25 °C in 20 mM Cacodylate, 100 mM KCl, pH 6.9.

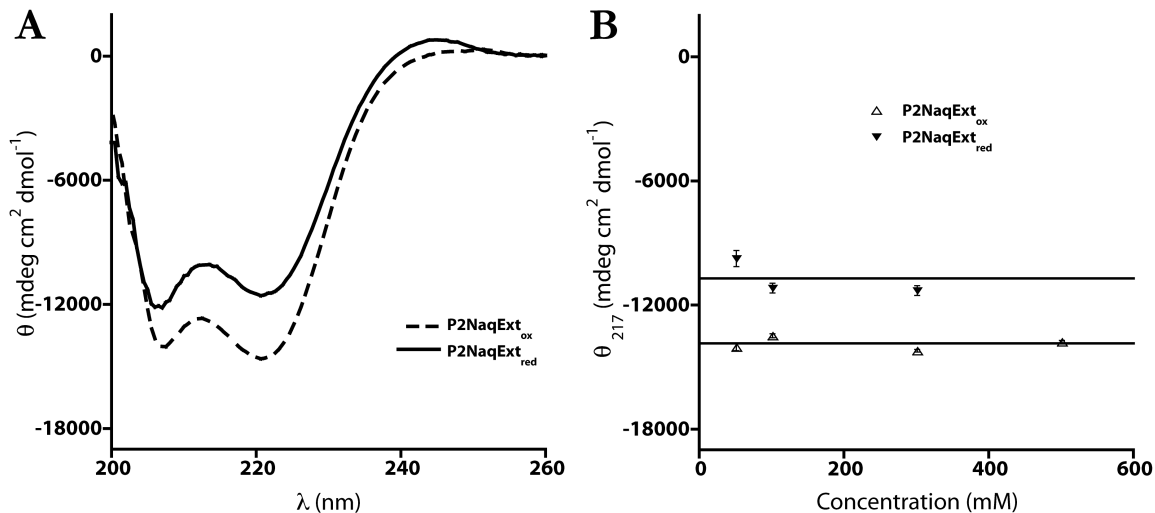


Figure 3.10. CD spectra of P2NaqExt_{ox} and P2NaqExt_{red} (A). Concentration dependence of CD signal at 217 nm for P2NaqExt_{ox} and P2NaqExt_{red} shown with experimental errors and lines drawn through the average (B). CD spectra were taken at 25 °C in 20 mM Cacodylate, 100 mM KCl, pH 6.9.

3.5.4 Circular Dichroism of P2NaqExt_{ox} and P2NaqExt_{red}

In designing the helical control peptide P2NaqExt, a length of a free helix stable in solution was appended to the P2Naq sequence. It is clear that to effect the desired result of forcing Naq into a helical conformation, this peptide, which presumably still binds lanthanide, should have its physical properties measured in saturating lanthanide ion; thus in all studies with P2NaqExt, saturating lanthanide is present. The CD spectra of P2NaqExt_{ox} and P2NaqExt_{red} can be seen in Figure 3.10A. The two peptides are clearly helical. Modifying **Eq. 4** we can obtain an anticipated value for the mean residue molar ellipticity for a P2NaqExt assuming it is simply a combined species

$$[\Theta^+(P2NaqExt)] = \frac{17 \cdot [\Theta(Baldwin)] + 17 \cdot [\Theta^+(P2X)]}{32} \quad \text{Eq. 9}$$

where $[\Theta^+(P2NaqExt)]$ is the predicted ellipticity of P2NaqExt, $[\Theta(Baldwin)]$ is the ellipticity of one of the Baldwin²⁴ peptides related to the sequence added to P2Naq, and

$[\Theta^+(P2X)]$ is either $[\Theta^+(P2Naq)]$ or $[\Theta^+(P2A)]$, the ellipticities of the P2Naq or the idealized ellipticities of P2A, respectively. The scaling numbers come from the backbone peptide lengths of the individual species. This calculation allows us to gauge how successfully the addition of the Baldwin peptide to P2Naq increased the helicity of joint product. For the purposes of this calculation, it is assumed that due to the relative flatness of the P2X peptide's mean residue molar ellipticity from 217 nm to 222 nm the value at 217 nm is equal to the value at 222 nm. The mean residue molar ellipticity of Baldwin's peptide is taken as $-11000 \text{ mdeg cm}^2 \text{ dmol}^{-1}$,²⁴ and that of P2A is taken as $-11000 \text{ mdeg cm}^2 \text{ dmol}^{-1}$.^{9,12} In the case of P2NaqExt_{ox}, using the ellipticity of lanthanide bound P2Naq_{ox} for $[\Theta^+(P2X)]$, $[\Theta(P2NaqExt)]$ is predicted to be $\sim -9500 \text{ mdeg cm}^2 \text{ dmol}^{-1}$, in comparison to the observed $-14500 \text{ mdeg cm}^2 \text{ dmol}^{-1}$. In the case of P2NaqExt_{red}, using ellipticity of lanthanide bound P2Naq_{red} for $[\Theta^+(P2X)]$, $[\Theta(P2NaqExt)]$ is predicted to be $-9520 \text{ mdeg cm}^2 \text{ dmol}^{-1}$, in comparison to the observed $-11500 \text{ mdeg cm}^2 \text{ dmol}^{-1}$. For both redox states of P2NaqExt, the predicted value assuming $[\Theta^+(P2X)]$ is the ellipticity of lanthanide bound P2A is $-11700 \text{ mdeg cm}^2 \text{ dmol}^{-1}$. The observed ellipticities compared to the predicted ellipticities strongly support the conclusion that the P2NaqExt control is helical. Using P2A as a proxy for a fully helical P2Naq unit, both P2NaqExt_{ox} and P2NaqExt_{red} compare favorably suggesting that in the absence of two helix nucleation events, the amino acids around Naq are, in fact, helical. It is not surprising that P2NaqExt_{ox} proves significantly more helical than predicted since without the need to overcome the barrier for or probability of nucleation, the helix would be likely to extend much further. However, it is surprising that P2NaqExt_{red} is not significantly different than that predicted in the idealized case, suggesting something else might be going on.

The cause of the observed significant difference between the oxidized and reduced forms of P2NaqExt is not immediately obvious. It may be possible that due to the experimental setup for the reduced peptides, the concentrations were significantly off, and this difference is amplified in the P2NaqExt experimental results. In addition, benzyl viologen, the dye added as a redox indicator in the reduced samples, is a dication in its oxidized form and was observed to cause the precipitation of P2NaqExt when both species were in high (~ 1 mM) concentrations, suggesting the presence of intermolecular interactions that might affect CD signals. Finally, it is entirely possible that reduced Naq completely disrupts or caps any helix formed nearby; however, NMR results of P2NaqExt_{red} discussed below will show that this is not likely to be the case and that the significant difference in the CD signal between the oxidized and reduced P2NaqExt is probably due to experimental parameters that were not ideal.

As with the samples of P2Naq and RCCNaq, there is no observed concentration dependence of the CD signal with either P2NaqExt_{ox} or P2NaqExt_{red}, Figure 3.10B. It is of note, that once again, in the reduced 50 μ M sample there is an anomalously low measured mean residue molar ellipticity at 217 nm.

3.5.5 Naq specific CD signals in P2Naq and P2NaqExt

The observation of a slightly different CD spectrum in the 240 nm-260 nm region for lanthanide bound P2Naq_{ox} and P2Naq_{red}, Figure 3.11B, was further confirmed with P2NaqExt_{ox} and P2NaqExt_{red}, Figure 3.11C. In the helical control peptides, the spectral features are more clearly those of the oxidized and reduced Naq. On closer inspection of the P2Naq_{ox} and P2Naq_{red} samples, only P2Naq_{ox} shows any sort of identifiable spectral features,

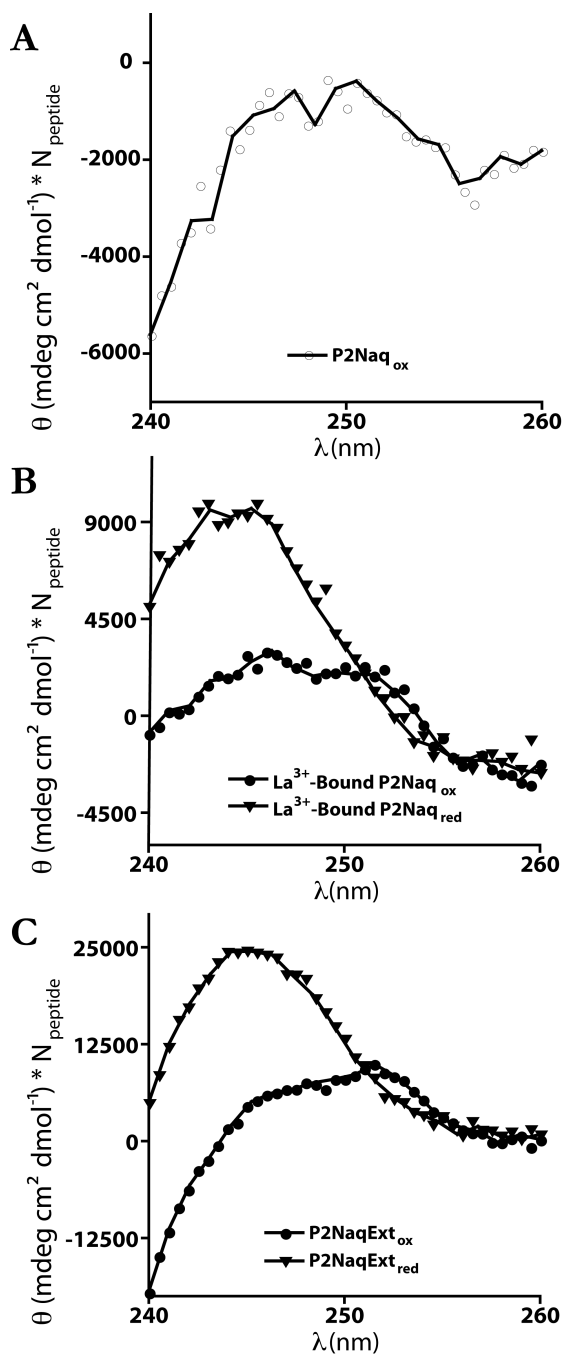


Figure 3.11. Naq chromophore specific molar ellipticities in P2Naq_{ox} (A), and both redox states of holo P2Naq (B) and P2NaqExt (C). N_{peptide} is the number of backbone peptide units. Data fit with cubic spline. See text for details.

Figure 3.11A. The CD signals of the RCCNaq random coil peptides do not show any discernible spectral feature relating to either form of the amino acid. The CD signal of aromatic amino acids in the near-UV is known to be responsive to local protein environment like the net-dipole moment, which affect the symmetry of the chromophore's environment and thus the CD signal through changes in the chromophore's polarizability. The apparent tracking of Naq's chromophore CD signal with helicity in these peptides, with no apparent response in random coil peptides and the strongest response in the helical control peptides, can serve as an additional approximate 'site-specific' measure of the fraction helicity of Naq in the apo and holo, reduced and oxidized P2Naq peptides. By using non-residue averaged redox difference CD spectra to remove the effects of the remainder of the peptide backbone from the

analysis, the observed changes in this region of the CD spectrum can be tied directly to the

Naq chromophore. Use of the relevant calculated redox difference ellipticities in **Eq. 3** allows an average measure of the fraction helicity of Naq in the reduced and oxidized peptides in question. While this method might suffer somewhat from low signal strength, this is not anticipated to prevent the use of the approach as an approximation of the real value as a method to confirm the results obtained as described above by CD or NMR. Additionally, non-identical CD signals of the reduced and oxidized peptides due to structural differences are not expected to be particularly strong in the relevant region between 240-260 nm, especially since neither amides nor tyrosine absorb appreciably in this region. Finally, while it is not clear if the Naq chromophore's CD signal is *linearly* responsive to the helicity of its environment or how this might occur, a possible explanation for the observed response is that it is dependent on the strength of the local helix dipole which should scale with length of the helix.

3.6 Measurement of Lanthanide Ion Binding Affinity of P2Naq_{ox} and P2Naq_{red}

The CD response of the P2Naq peptides to saturating lanthanide confirmed that ligand binding induced helix formation. In previous reports with the P2X peptides,⁸ binding affinities were measured using changes in Tyr₇'s fluorescence intensity. Initial attempts to use fluorescence to measure P2Naq_{ox}'s lanthanide binding affinity failed, possibly due to charge transfer between Tyr₇ and Naq₁₄ suppressing tyrosine's natural fluorescence. This left circular dichroism as the only spectroscopic technique capable of being used to measure lanthanide binding. The CD titrations were significantly eased using an automated titration apparatus, which eliminated the significant variability associated with manual measurements. Data collected at 25 °C was fit to the tight binding equation

$$CD_{signal} = \epsilon_{apo} [P2Naq]_{tot} + (\epsilon_{holo} - \epsilon_{apo}) \frac{-A - \sqrt{A^2 - 4[P2Naq]_{tot} [La^{3+}]_{tot}}}{2} \quad \text{Eq. 10}$$

where CD_{signal} is the measured CD signal at a chosen wavelength, ϵ_{apo} and ϵ_{holo} are uncorrected ellipticities of the ligand free and bound peptides, A is $(K_d + [La^{3+}]_{tot} + [P2Naq]_{tot})$, K_d is the dissociation constant, $[La^{3+}]_{tot}$ is the total concentration of lanthanide ion, and $[P2Naq]_{tot}$ is the total concentration of P2Naq. Lanthanum chloride stocks were made by dissolving solid $LaCl_3$ in the titration buffer to approximately 10 mM, making aliquots and snap freezing. The stocks were kept at -20 °C to prevent precipitation of lanthanum carbonate formed from lanthanide ion and dissolved carbon dioxide. Stock concentration was determined using a spectrophotometric end point titration of a lanthanide ion stock/xylene orange solution with a calibrated EDTA solution.²⁵ The EDTA solution was calibrated with a standardized zinc solution with xylene orange. This calibration scheme varies markedly from the method used by Richardson *et al.*,⁸ who made their lanthanide ion stocks by mass, a technique that proved unreliable for the titrations. In the titrations, the peptide concentration was always approximately 5 μ M, as determined by UV/*vis* spectroscopy using the molar absorptivities determined as described above.

3.6.1 Lanthanide Binding of P2Naq_{ox}

The lanthanide titration of P2Naq_{ox} observed with CD was relatively straightforward. Five identical titrations were simultaneously fit to the tight binding Eq. 10 at 217 nm after their initial CD signal was zeroed due to significant variance, Figure 3.12. The dilution effect was accounted for by discreetly fitting concentrations of P2Naq_{ox}, giving a K_d of $3.2 \mu\text{M} \pm 0.4 \mu\text{M}$. This K_d is the tightest non-alanine

dissociation constant measured for the P2X peptide family.⁸

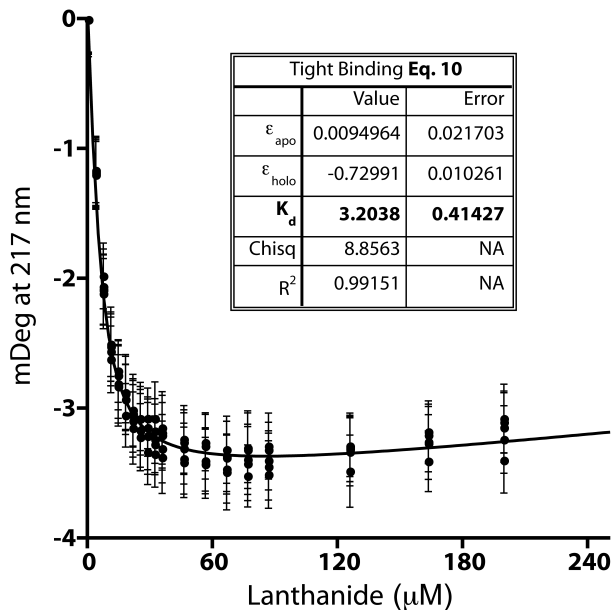


Figure 3.12. Simultaneous fit of five identical lanthanide ion titrations of P2Naq_{ox} with errors in 20 mM Cacodylate, 100 mM KCl, pH 6.9. Data was fit to the tight binding equation Eq. 10. See text for details.

3.6.2 Lanthanide Binding of P2Naq_{red}

Measuring lanthanide binding to P2Naq_{red} proved problematic. It was determined early that large excesses (>100 eq) of the water stable ammonia borane (NH_3BH_3) complex could reduce Naq at neutral pH without the concomitant production of hydrogen gas seen with borohydride reagents. Unfortunately, it appeared that the product borate, from the oxidation of the borane, bound lanthanide very tightly; causing the titrations to have both a lag before binding to P2Naq_{red} began and, since there was likely continued oxidation of the

borane during the course of the titration, the effective K_d measured was not accurate given that the exchangeable lanthanide was not equal to the total lanthanide added.

Both the common catalytic oxygen scrubbing system of glucose oxidase/glucose/catalase and the less common xanthine oxidase/xanthine/methyl viologen interfered with lanthanum binding. Spectroscopic considerations made it clear that by adding sufficient quantities of any of the available, non-metal binding, redox mediators to reduce P2Naq and maintain a scrubbed sample, the absorbance in 200-260 nm range would far exceed that needed for good CD measurements. The need for a spectroscopically silent reductant led to consideration of the use of hydrogen gas, H_2 . The only literature report known to employ hydrogen gas directly to reduce a redox active protein used a natural hydrogenase to effect reduction.²⁶ While an attempt was made to find sources of natural hydrogenases, a brief literature search made it clear that although seemingly never attempted before, the use of platinum nanoparticles could facilitate reduction of P2Naq with H_2 . It was observed that at a high, but unbuffered, pH, platinum nanoparticles could reduce methyl viologen in the presence of hydrogen.²⁷ Since the viologen redox dyes do not have pH dependent redox couples, a quick back of the envelope calculation suggested that the platinum nanoparticles could conceivably reduce the higher potential benzyl viologen at the titration pH 6.9 (the potential of hydrogen gas at pH 6.9 is -414 mV; benzyl viologen is -311 mV). Nanoparticles with a 40 kD polyvinylpyrrolidinone protective polymer in a 40:1 monomer/platinum ratio generated under standard conditions²⁸ were capable of mediating H_2 reduction of a benzyl viologen solution at pH 6.9 as indicated by the appearance of the blue-purple viologen radical in less than 30 s. To ensure there was no interference of

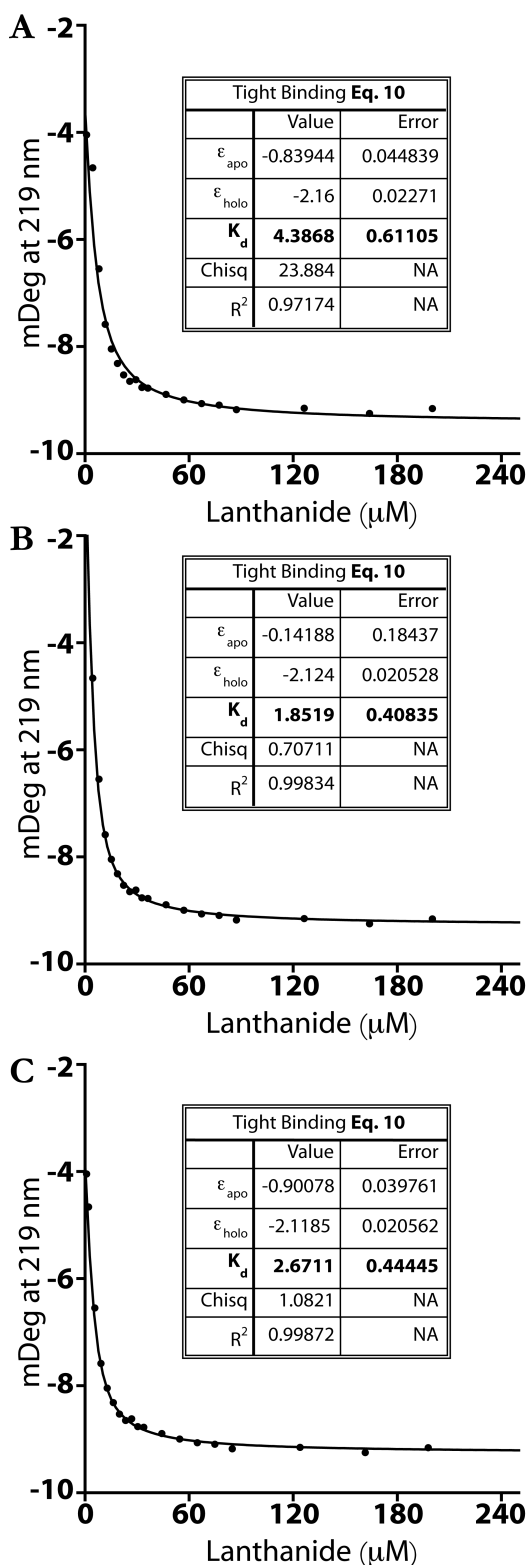


Figure 3.13. P2A titration under reducing conditions; fit with **Eq. 10** as measured (A), ignoring first point (B), and making a linear adjustment to the data (C). See text for details.

lanthanide binding by free platinum ions, the colloidal platinum was dialyzed extensively in pure water.

In the reduced titrations, platinum nanoparticles were kept at a platinum atom concentration within an order of magnitude below the concentrations of the peptides being studied. Additionally, benzyl viologen was 50 μM , a concentration not anticipated to interfere with CD signals but enough that it might serve as a redox buffer. Under these conditions, in the absence of positive pressure hydrogen, P2A, the reference peptide from the work of Richardson *et al.*,⁸ bound lanthanide ions with the same affinity, within error, as that reported, 2.4 μM .⁸ However, under positive pressure hydrogen gas, with the sample reduced as observed by the blue-purple of the benzyl viologen radical, the P2A titration gave a much higher K_d , 4.4 μM , Figure 3.13A.

While this type of variance might be

anticipated from one titration to another, the lack of this significant variability in the P2Naq_{ox} titration data gave pause. It was also clear that the difference between the experimental value and the reported value was solely due to what appeared to be an incomplete addition of the first aliquot of lanthanide. This seeming incomplete addition of the first aliquot was also observed under similar conditions for titrations of P2Naq_{red}. In some of these titrations, there was no observable change in CD signal after the initial 5 μ L aliquot. Changing the concentration of the platinum nanoparticles did not change the stoichiometry of the apparent pause, and since the pause was observed in both P2Naq_{red} and P2A under similar conditions, it was not a peptide specific problem. It is possible that when under positive pressure, even at the lowest over pressure setting possible, there is significant compression of the fluid line (of a bubble, for instance) from the titration apparatus to the sample cuvette. If this were the case or if there was an exogenous lanthanide binder in solution, the data could be corrected by assuming linearity in the concentration dependence of the CD signal change for the first two responsive aliquots. The lanthanide concentrations for all aliquots could then be corrected by subtracting the concentration of lanthanide that was apparently not added in the initial partial aliquots. The other possibility is that the lanthanide solution is being partially added to the cuvette during manipulation of the titrant line in setting up the titration. The correction for this would simply be to ignore the first, 'without-lanthanide' point. This second possible cause for the observed delay in response is unlikely to be the case since there were titrations in which the pause lasted for two aliquots, thus suggesting that greater than 5 μ L of the lanthanide solution was added before the titration began. 5 μ L represents the volume contained in 2 cm of the steel transfer tubing and it is highly improbable that much compression occurred during manipulation. In

addition, when fitting the data using the assumption that the first point is to be ignored, the fit predicts an experimental change in CD signal for P2Naq_{red} (and P2A) that is much larger than the signal change observed when the titrant tube is added to the sample after the first CD signal is recorded. In any event, correcting the P2A titration as described by assuming the first point is in error, K_d $1.8 \mu\text{M} \pm 0.4 \mu\text{M}$ as seen in Figure 3.13B, and by assuming the first aliquot is partially consumed, K_d $2.7 \mu\text{M} \pm 0.4 \mu\text{M}$ as seen in Figure 3.13C, reveals that the fitted K_d 's become significantly closer to the reported value. The assumption that the first aliquot is only partially added matches the expected value more closely, and because of the reasons outlined above this method was chosen for correcting the titration data observed with P2Naq_{red}.

Unlike P2Naq_{ox}, five identical experiments were not preformed with P2Naq_{red}; however, five experiments, two in duplicate, were fit as with **Eq. 10**; no correction for dilution was needed since the lanthanum titrant solution had P2Naq_{red} at the same concentration as the sample. This procedural adjustment was made to ensure a good endpoint was reached, as opposed to the trailing upward endpoint of the P2Naq_{ox} titrations, and so that the sample would remain reduced while the titrant was being added. Both the titrant and the sample contained benzyl viologen as a redox indicator and buffer to ensure maintenance of a reduced sample during the experiment. The five values for the measured lanthanide binding K_d of P2Naq_{red} were in the range between $4.5 \mu\text{M}$ and $4.6 \mu\text{M}$ with an average of $4.6 \mu\text{M} \pm 0.1 \mu\text{M}$, where the reported error is propagated from the error associated with each measurement ($\sigma=0.04 \mu\text{M}$). If the method of correcting for the apparent pause in the titration data were in error, it is unclear how such disparate titrations (varying by the concentrations of peptide, lanthanide in the titrant solution and colloidal

platinum, wavelengths of observation and methods for degassing the solutions), would produce nearly identical measures of binding affinity.

3.6.3 Completing the Thermodynamic Box

With the measurement of the two lanthanide ion binding affinities corresponding to $\Delta G(P2Naq_{ox})$ and $\Delta G(P2Naq_{red})$ complete, a measure of the difference between the midpoint potentials of the lanthanide bound and unbound P2Naq can be made.

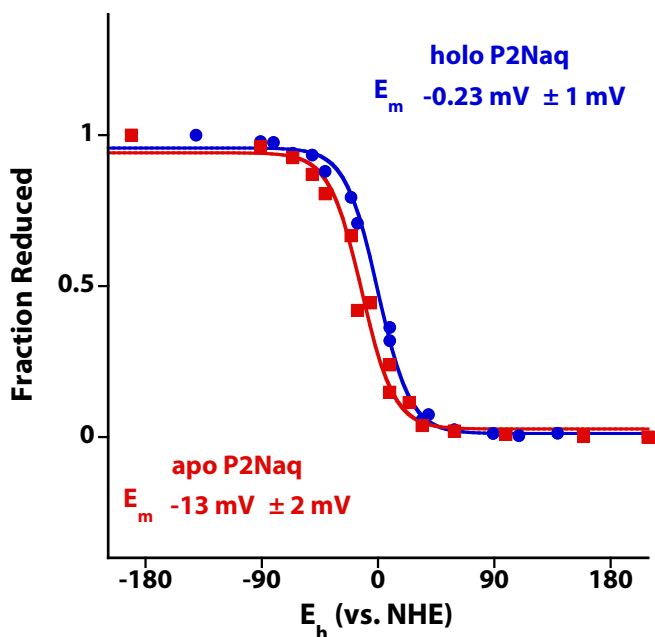


Figure 3.14. Spectroelectrochemical titration of both apo (red) and holo (blue) P2Naq in 20 mM Cacodylate, 100 mM KCl, pH 6.9 fit to the two electron Nernst equation. The change in measured midpoint is of the opposite sign and of a larger magnitude than measured with the lanthanide titrations.

$\Delta G(P2Naq_{ox})$, as defined by **Eq. 1** and using the reported value of $K_d(P1)$ of $9.4 \mu\text{M}$, is $-0.64 \text{ kcal mol}^{-1}$; $\Delta G(P2Naq_{red})$ is $-0.42 \text{ kcal mol}^{-1}$. Using **Eq. 6** this gives a ΔG_{+-} , the change in free energy between the bound and unbound P2Naq, of $0.22 \text{ kcal mol}^{-1}$ and a $\Delta E_{m,+-}$, the change

in P2Naq's midpoint with and without lanthanum, of -5 mV . A change in midpoint of 5 mV is not

reliably measurable by any electrochemical technique available in the Dutton Laboratory.

Early redox titrations of P2Naq with and without lanthanide indicated as much with the measured potentials for lanthanum bound and unbound forms orbiting each other, Figure 3.14 (José Cerda, personal communication). The electrochemistry is very much similar to

that of heptaNaq showing clear two-electron Nernst behavior in the presence of mediators for both lanthanide bound, $E_m -0.23 \text{ mV} \pm 1 \text{ mV}$, and unbound forms, $E_m -13 \text{ mV} \pm 2 \text{ mV}$. The $\Delta E_{m,+}$ is 13.2 mV, which is of the opposite sign and of a larger magnitude than the lanthanide binding data would suggest. It is illuminating to consider that this measured difference in energy would predict a difference in lanthanide affinities between P2Naq_{red} and P2Naq_{ox} on the same order as that seen between alanine and glycine in P2X, the largest difference measured by Richardson *et al.*⁸ In the attempt to better measure the redox difference between the lanthanum bound and unbound forms, a new spectroelectrochemical cell is being developed which will be described in a section in Appendix A.

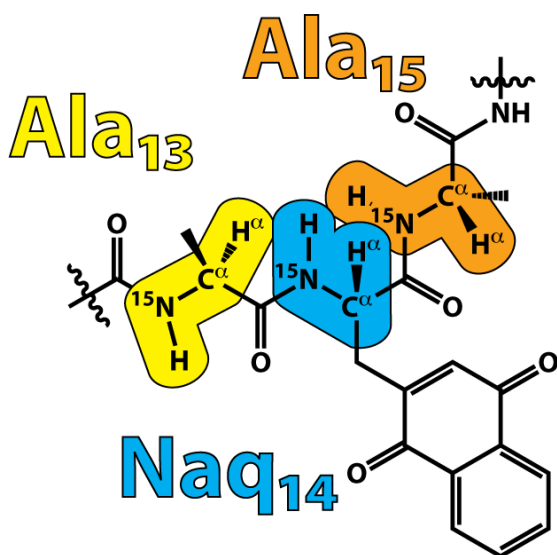


Figure 3.15. Backbone ^{15}N labeling of the amino acid sequence $\text{A}_{13}\text{Q}_{14}\text{A}_{15}$. The amide proton, ^{15}N amide nitrogen, C^α and H^α are indicated for clarity. NMR data was collected to observe the chemical shifts of these four nuclei.

3.7 Backbone Chemical Shift

Determination for P2Naq, RCCNaq, and P2NaqExt

The nitrogen labeled peptides P2Naq, RCCNaq, and P2NaqExt were created to allow the unambiguous measurement of site-specific helicities of Naq and the two adjacent alanines to Naq, Figure 3.15. In all three peptides, the $\text{A}_{13}\text{Q}_{14}\text{A}_{15}$ (the numbering is preserved from

P2Naq and P2NaqExt, even in the case of RCCNaq) amino acid sequence was uniformly ^{15}N labeled to provide for a full description of the local conformation of Naq in the various forms of the test peptide P2Naq. Circular dichroism as discussed above confirmed the

success of the designs of the control peptides RCCNaq and P2NaqExt as models of random coil and helical conformations, respectively. In addition, the lack of concentration dependence of the CD signals for all of the peptides in both reduced and oxidized forms further suggests that any NMR results observed with peptides in the low mM range should be applicable to the thermodynamics measured with peptides in the low μM range. All NMR data was collected with Veronica Moorman in the laboratory of A. Joshua Wand.

In the initial design and selection of ^{15}N labeling, it was thought that three experiments (the ^{15}N filtered HNHA, NOESY and HNCA) would sufficiently describe the backbone chemical shifts, $^{13}\text{C}^\alpha$ and H^α , the dihedral angle ϕ via J-coupling and the connectivity between the side chains; ensuring identity of compared nuclei was preserved across the various samples. For practical reasons the length of data collection was a concern given the slow hydroxylation of oxidized Naq in solution. In trial experiments, it was evident that the HNCA was not going to give sufficient signal strength to allow full data collection in a reasonable (<12 hours) amount of time. To obtain $^{13}\text{C}^\alpha$ chemical shifts, a ^{13}C -HSQC was needed, and a ^{15}N -HSQC was also added to the experiment list. In all, ^{13}C -HSQC, ^{15}N -HSQC, and ^{15}N filtered HNHA and NOESY spectra were taken of each peptide, excluding the P2NaqExt peptides, which were not observed with ^{13}C -HSQC due to the supposition that the alanine chemical shifts for the residues adjacent to Naq would be indistinguishable from the background of 16 hypothetically helical alanines. In addition, no $^{13}\text{C}^\alpha$ chemical shift data for Naq was obtained since in the ^{13}C -HSQC the chemical shift of Naq's H^α overlapped with the H_2O proton peak, even in a D_2O sample. ^{13}C labeling is a potential route to obtaining these values, and given the lack of success in obtaining data from the HNCA experiment, ^{13}C labeling might have been more advantageous. Nevertheless, full H^α

data was obtained for all of the peptides; tracking the identities of the nuclei and their amino acids by inference proved slightly easier than expected since the chemical shift behaviors of the two alanines across the peptides were very different. The NOESY spectra were not entirely useful for identifying amino acids as NOE peaks only appeared in some samples, notably only in the lanthanide containing P2NaqExt and P2Naq samples. It is also noteworthy that due to degeneracy of the amide proton chemical shifts in some of the samples, J-coupling values and hence values for ϕ are unreliable, but this is a minor point.

All data was collected at 25 °C with temperature calibrated to methanol, ^{15}N , ^{13}C , and ^1H chemical shifts were standardized to the ^1H chemical shift of DSS by IUPAC recommendations.²⁹ Concentrations of the peptides ranged from 1 mM-2.6 mM; samples were adjusted to pH 6.9 ± 0.1 as measured by a micro pH electrode. Samples which contained lanthanide had approximately 10 mM LaCl_3 added prior to pH calibration. Measurements were performed on a Bruker 500 MHz Advance II Spectrophotometer equipped with a cold probe. All of the reduced samples had 11 μM colloidal platinum prepared as described above and in most cases, 100 μM benzyl viologen. Samples to be reduced were stirred in a small pear shaped round bottom under a fast flow of H_2 for at least 15 minutes before colloidal platinum was added. Reduction took ~ 5 minutes as indicated by the solutions developing a blue-purple color characteristic of the benzyl-viologen radical. After reduction was observed the samples were quickly cannulated into a clean NMR tube that had been purged with H_2 through a rubber septum. Attempts to conventionally flame seal the NMR tubes were unsuccessful, so after sample transfer the rubber septum of the NMR tube was dipped in a quick setting epoxy and allowed to harden for an hour before NMR experiments were undertaken. In all cases, the epoxy sealed septum held for the

course of the experiments, with very little evidence of oxidation of the benzyl viologen (loss of blue-purple color) at the top of the NMR sample. Intriguingly, a sample prepared in this way oscillated in the apparent deepness of the blue-purple color over the course of months; a phenomenon that cannot be explained by oxygen leakage into the tube nor a simple two component reaction. It is likely, given the complex redox active nature of the cacodylate buffer that a complex oscillatory reaction has accidentally been established in this sample.

Table 3.2. Chemical shifts of backbone atoms of ^{15}N labeled Ala₁₃, Naq₁₄ and Ala₁₅ in P2Naq, RCCNaq and P2NaqExt under a variety of conditions. Chemical shifts are in ppm and referenced to DSS. The $^1\text{H}^\alpha$ and $^{13}\text{C}^\alpha$ chemical shifts in bold were used in calculating site-specific fraction helicity see text for details. Nuclei have same naming convention as in Figure 3.15.

| Peptide | Ala ₁₃ | | | | Naq ₁₄ | | | Ala ₁₅ | | | |
|----------------------------|-------------------|--------------|-----------------|------|-------------------|-----------------|------|-------------------|--------------|-----------------|------|
| | H^α | C^α | ^{15}N | H | H^α | ^{15}N | H | H^α | C^α | ^{15}N | H |
| <i>Ligand Free</i> | | | | | | | | | | | |
| P2Naq _{ox} | 4.08 | 50.61 | 123.2 | 7.97 | 4.65 | 117.0 | 8.05 | 4.24 | 50.22 | 124.5 | 8.11 |
| P2Naq _{red} | 4.09 | 50.67 | 123.8 | 8.06 | 4.59 | 118.1 | 8.09 | 4.20 | 50.19 | 124.6 | 7.81 |
| <i>Lanthanide Bound</i> | | | | | | | | | | | |
| P2Naq _{ox} | 3.99 | 51.26 | 122.5 | 8.25 | 4.60 | 115.8 | 7.77 | 4.27 | 50.38 | 122.7 | 7.74 |
| P2Naq _{red} | 4.00 | 51.16 | 122.3 | 8.06 | 4.55 | 118.0 | 7.90 | 4.19 | 50.26 | 123.1 | 7.74 |
| <i>Random Coil Control</i> | | | | | | | | | | | |
| RCCNaq _{ox} | 4.15 | 49.85 | 123.9 | 7.99 | 4.69 | 117.8 | 8.20 | 4.27 | 49.83 | 125.3 | 8.29 |
| RCCNaq _{red} | 4.15 | 50.11 | 124.5 | 8.11 | 4.64 | 118.7 | 8.12 | 4.21 | 49.76 | 125.5 | 7.86 |
| <i>Helical Control</i> | | | | | | | | | | | |
| P2NaqExt _{ox} | 3.98 | - | 124.3 | 8.75 | 4.46 | 117.6 | 8.20 | 4.09 | - | 123.3 | 7.95 |
| P2NaqExt _{red} | 4.04 | - | 124.0 | 8.62 | 4.46 | 119.7 | 8.22 | 4.06 | - | 121.9 | 7.91 |

3.7.1 Chemical Shifts

The measured chemical shifts for the various backbone atoms indicated in Figure 3.15 can be found in Table 3.2. The chemical shifts of the $^{13}\text{C}^\alpha$ of Naq in all peptides and in alanines of the P2NaqExt peptides are absent for the reasons described above. The chemical

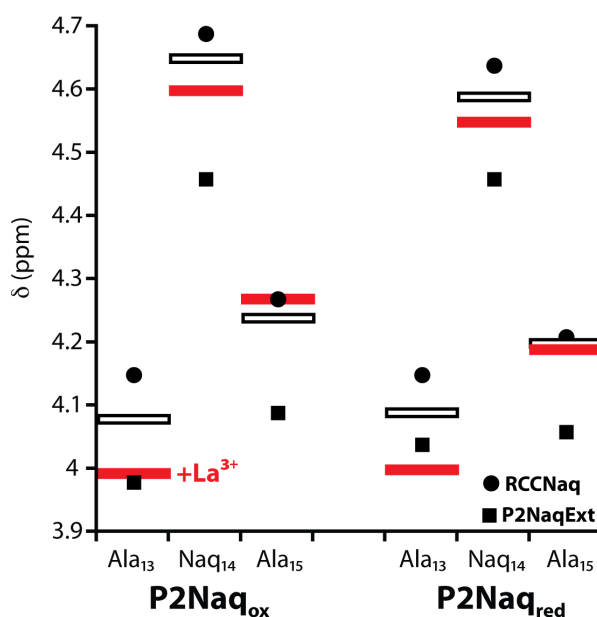


Figure 3.16. Graphical representation of the H^α chemical shift data in Table 3.2. P2Naq δ_{H^α} (outlined boxes) and lanthanide-bound P2Naq δ_{H^α} (red boxes). Redox dependent chemical shifts for random coil (circle) and helical controls (box) are shown with their corresponding atom in P2Naq. See text for details.

shift values for $^1H^\alpha$ of Naq₁₄ show some trends, Figure 3.16, with the values of the experimental P2Naq peptides for the most part falling in between the values obtained for RCCNaq and P2NaqExt. On the other hand, while there are significant differences between the RCCNaq and P2NaqExt chemical values for the $^1H^\alpha$ of Ala₁₅ in both redox states, Figure 3.16, the chemicals shift for these atoms and the corresponding $^{13}C^\alpha$ s in the experimental P2Naq

peptides all lie very close to the random coil peptide (towards higher ppm values in Figure 3.16, e.g.), giving more evidence that as suggested by CD, the helix is not fully formed in P2Naq even in the presence of lanthanide. This contrasts nicely with the chemical shifts of Ala₁₃, Figure 3.16, which like Naq₁₄, has values of $^1H^\alpha$ that are well dispersed and show clear variance between the apo and holo P2Naq peptides. In addition, the $^{13}C^\alpha$ for Ala₁₃ values have a much larger range than those of Ala₁₅ suggesting the clear conformational change implicated by the $^1H^\alpha$ chemical shift changes is real. The chemical shift values suggest quite strongly that upon lanthanide binding in the P2Naq peptides Ala₁₃ becomes significantly helical, Naq₁₄ becomes slightly more helical, and Ala₁₅ does not change. This further suggests that Naq₁₄ is a distinct ‘helix breaker,’ and that there is possibly not enough energy in helix

formation from the last two residues to drive it into a fully helical conformation. These results will be discussed more quantitatively when considered with the results of the CD studies in the next section.

3.7.2 Quantitative J-Correlation and Measures of ϕ

J-coupling is a direct consequence of nuclei being bound together as magnetization is transfer from nuclei to bonded nuclei through the intervening electrons that make up the bond(s) between them. The efficiency of this transfer is highly dependent on the relative orientation of the bonds, which when part of a three-bond system describes a dihedral angle according to the Karplus equation, Figure 3.17

$$J(\theta) = A \cos^2(\theta) + B \cos(\theta) + C \quad \text{Eq. 11}$$

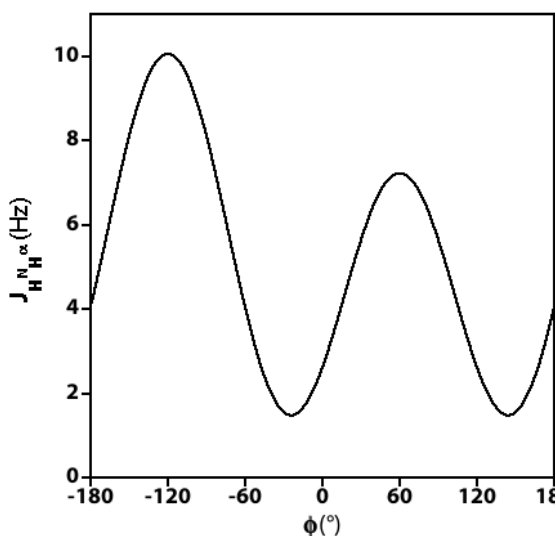


Figure 3.17. Karplus curve described by Eq. 11 calibrated to ϕ ($\phi = \theta + 60^\circ$). The empirically derived values A, 7.09; B, -1.42; C, 1.55 are from reference¹ and were used in the calculation of ϕ from experimentally determined $J_{H\alpha HN}$.

where $J(\theta)$ is the measured J-coupling between bonded nuclei, θ is the dihedral angle between the first and third bond defined physically, and A, B and C are empirically derived parameters dependent on the identities of the bonds in the dihedral angle.¹ In the case of the backbone dihedral ϕ , $\phi = \theta + 60$ due to convention. J-coupling is most commonly manifest as a splitting of peaks in 1D ^1H NMR. Measurement of splitting parameters in the 1D or 2D NMR

spectra of proteins and peptides is made more difficult due to the large amount of overlap between nearly chemically identical nuclei. Quantitative measurement in experiments like the HNHA take advantage of dephasing effects associated with J-coupling dependent magnetization transfer observed in the cross and diagonal peaks

$$\frac{S_{cross}}{S_{diagonal}} = -\tan^2(2\pi J_{H^\alpha H^N} \zeta) \quad \text{Eq. 12}$$

where S_{cross} and $S_{diagonal}$ are the peak volumes of the amide correlated cross ($^1\text{H}^\alpha$ -amide ^1H) and (amide ^1H) diagonal peaks, respectively, $J_{H^\alpha H^N}$ is the measured J-coupling between the $^1\text{H}^\alpha$ and amide $^1\text{H}^N$, and ζ is a pulse delay during which J-coupling dephasing occurs.³⁰ An

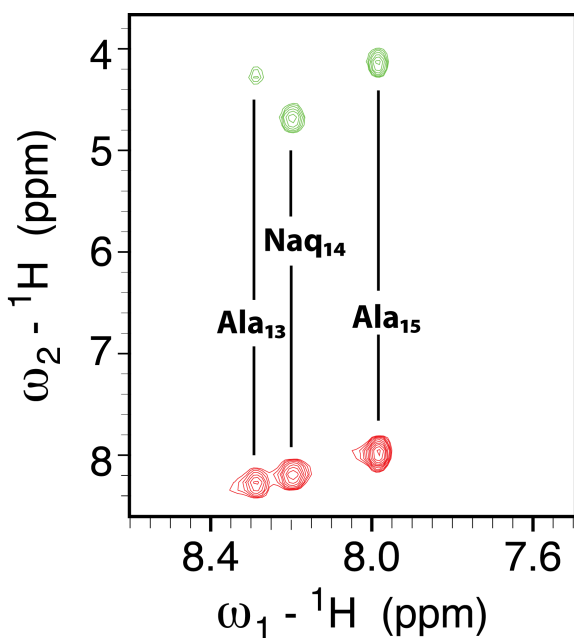


Figure 3.18. ^{15}N -filtered HNHA of RCCNaq_{ox} showing characteristic peaks with positive intensity along the amide proton diagonal (red) and negative cross peaks correlated to the amide proton and H^α (green). Measurement of the relative volumes of the peaks gives $J_{H^\alpha H^N}$ according to **Eq. 12**. The J-coupling values can then be used to calculate the backbone dihedral ϕ . Data was collected as described in the text with $\zeta=13.05$ and 128 transients in the indirect dimension.

example of a nitrogen filtered HNHA spectrum from this experimental set is seen in Figure 3.18.

While **Eq. 11** describes four potential dihedral angles, three of the angles described are unlikely since they are either positive or very near to -180° . To clarify analysis these values were ignored, Table 3.3. Excluding values of ϕ that were determined from nearly degenerate amide protons, lightened in the table, there is an overall trend towards more positive ϕ values as

peptides become more helical. Comparing the ϕ for all P2NaqExt and RCCNaq, the reported values of ϕ for the helical control are closest to the idealized helical value of $\sim -58^\circ$ for an alpha helix. Other trends are harder to distinguish given a lack of credible data, and the measured angles in the reduced P2NaqExt peptide are significantly higher than in the oxidized peptide. This may have been a result of an accidental excess of benzyl viologen radical in this particular sample, with the radical causing differential relaxation of the transitions that lead to the cross and diagonal peaks.³² In any event, the HNHA J-coupling experiments are not sufficient to identify the conformation of the backbone since lacking the ϕ angle prevents full attribution. Indeed, the measured ϕ angles for the RCCNaq peptides could very well be found in a naturally occurring helix.

Table 3.3. The backbone $J_{H\alpha HN}$ scalar coupling constants as measured using volumes of cross and diagonal peaks from HNHA experiments with $\zeta=13.05$ ms as proscribed by Eq. 12. Volumes were measured by Gaussian integration in Sparky.³¹ ϕ was subsequently calculated using Eq. 11 with values of A, B and C from reference¹ (A 7.09; B -1.42; C 1.55). Values in grey were obtained from degenerate diagonal peaks and are unreliable, but included for completeness.

| Peptide | Ala ₁₃ | | Naq ₁₄ | | Ala ₁₅ | |
|----------------------------|-----------------------|--------|-----------------------|--------|-----------------------|--------|
| | $J_{H\alpha HN}$ (Hz) | ϕ | $J_{H\alpha HN}$ (Hz) | ϕ | $J_{H\alpha HN}$ (Hz) | ϕ |
| <i>Ligand Free</i> | | | | | | |
| P2Naq _{ox} | 5.63 | -71.7° | 6.87 | -80.5° | 6.30 | -76.4° |
| P2Naq _{red} | 5.00 | -67.2° | 7.16 | -82.6° | 5.99 | -74.2° |
| <i>Lanthanide Bound</i> | | | | | | |
| P2Naq _{ox} | 4.63 | -64.5° | 6.18 | -75.6° | 4.83 | -66.0° |
| P2Naq _{red} | 5.46 | -70.5° | 6.23 | -75.9° | 5.34 | -69.6° |
| <i>Random Coil Control</i> | | | | | | |
| RCCNaq _{ox} | 5.74 | -72.5° | 7.39 | -84.4° | 5.97 | -74.1° |
| RCCNaq _{red} | 5.20 | -68.6° | 5.97 | -74.1° | 6.05 | -74.6° |
| <i>Helical Control</i> | | | | | | |
| P2NaqExt _{ox} | 4.49 | -63.5° | 4.98 | -67.0° | 4.31 | -62.1° |
| P2NaqExt _{red} | 5.20 | -68.6° | 5.59 | -71.4° | 4.64 | -64.5° |

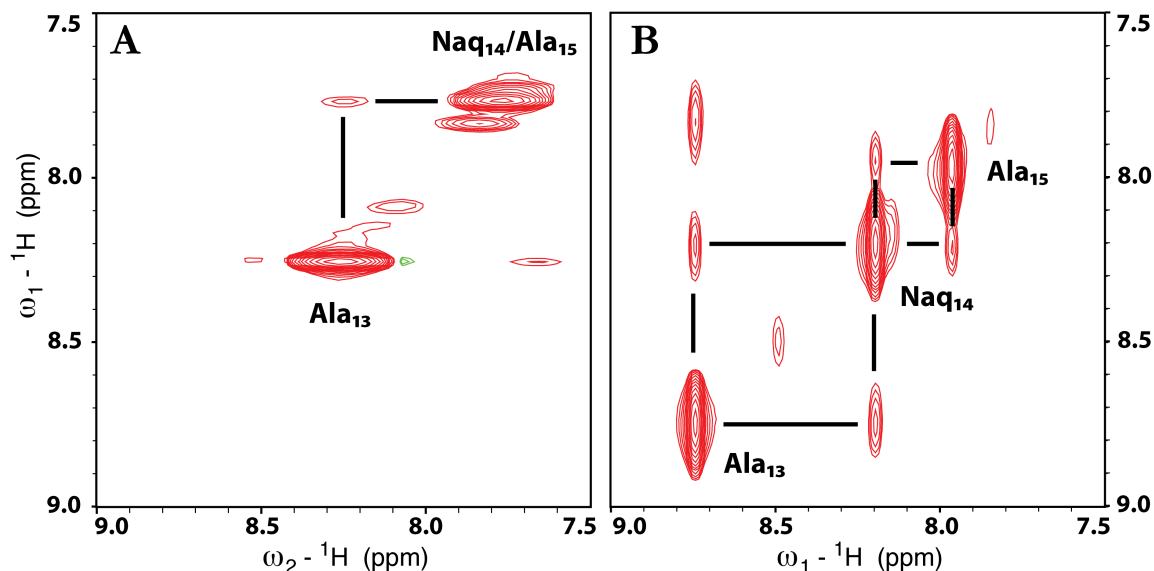


Figure 3.19. ^{15}N -filtered NOESY spectra of lanthanide bound P2Naq_{ox} (A) and P2NaqExt_{ox} (B). Black bars overlaying spectra track NOE cross peak correlations. The correlation observed in lanthanide bound P2Naq_{ox} suggests some presence of helicity; however, the slight correlation compares unfavorably to the clear backbone walk from the bottom left peak of (B) corresponding to Ala₁₃ to Naq₁₄ and finally to Ala₁₅. Data was collected as described in the text with 128 transients collected in the indirect dimension.

3.7.3 Walking the Backbone with a ^{15}N Filtered NOESY

In proteins with well-defined alpha helices, it is possible to walk the backbone from amide $i \rightarrow i+1$. This is true because of the close proximity in space of the amide protons of adjacent residues allowing efficient NOE magnetization transfer. The NOESY spectra of the P2Naq and RCCNaq peptides were mostly devoid of such NOE signals, Figure 3.19A, even in the lanthanide bound P2Naq samples. However, in the P2NaqExt helical control samples, it became possible to observe a very clear backbone-amide to backbone-amide in the NOESY spectra, Figure 3.19B, adding further evidence of helicity of these controls, and the lack of full helicity of lanthanide bound and unbound P2Naq.

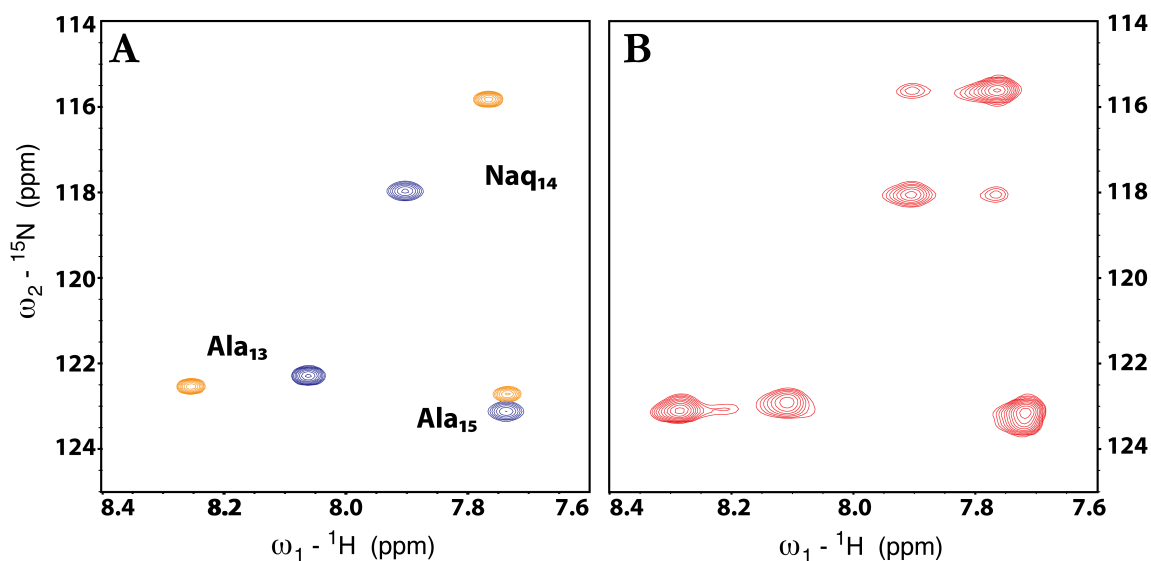


Figure 3.20. Chemical exchange in mix redox state sample of holo P2Naq. ^{15}N -HSQC overlay of pure lanthanide bound P2Naq_{ox} (orange) and P2Naq_{red} (blue) (A); and mixed lanthanide bound P2Naq_{ox} and P2Naq_{red} in an approximate 60:40 ratio. The appearance of two additional Naq₁₄ ‘cross peaks’ in the mixed sample along with observations of chemical exchange in the sample’s NOESY spectrum suggests that electron transfer was occurring between P2Naq_{ox} and P2Naq_{red}. The chemical shifts in (B) are uncorrected. See text for details.

3.7.4 Measuring Electron Transfer Rates Between Lanthanide-Bound P2Naq_{ox} and P2Naq_{red}

While most samples were prepared fully oxidized or fully reduced, two partially reduced lanthanide bound P2Naq samples were prepared accidentally. The NOESY of both samples indicated some amount of slow chemical exchange was occurring between the oxidized and reduced forms of the peptide. However, this process was more clearly observed in the ^{15}N -HSQC spectrum, Figure 3.20B. In comparing to the ^{15}N -HSQC of pure oxidized and pure reduced samples, combined in Figure 3.20A, it is clear that in the mixed sample two additional ‘cross peaks’ arise at the intersection of the reduced amide proton and oxidized amide nitrogen chemical shifts, and vice versa. It appears that chemical exchange is manifest in the proton dimension alone based upon the fact that the peaks begin to merge only parallel to the proton axis; the asymmetry of the exchanging peaks is further evidence that the exchange process is fairly slow, since in the absence of exchange, peaks should be

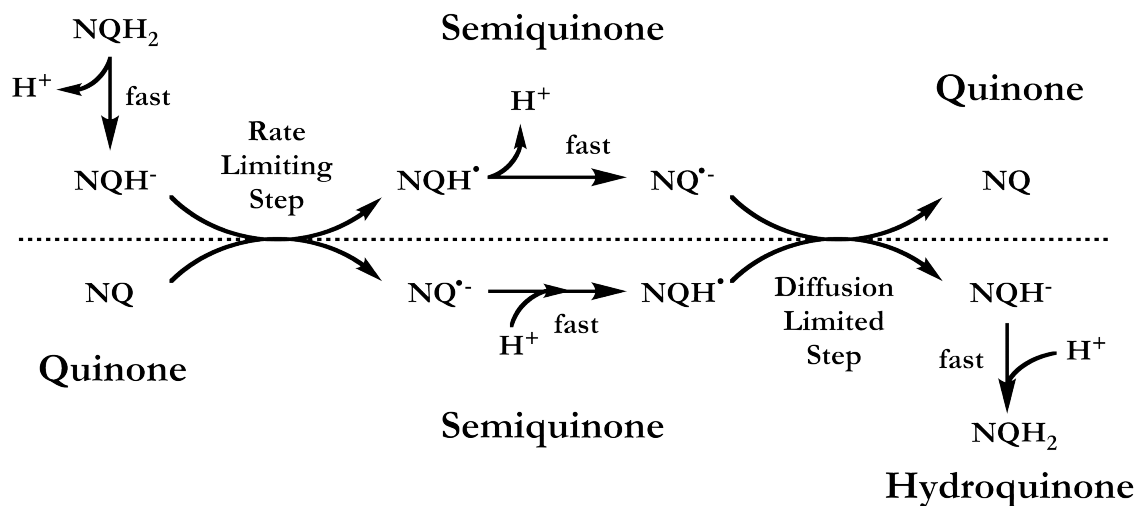
found only at the pure oxidized and reduced positions. It may be that exchange is manifest in the nitrogen dimension as well, but the overall exchange is so slow that the length of the ^{15}N chemical shift evolution period in the pulse sequence is insufficient to fully observe line broadening associated with the exchange process. The rate of the slow chemical exchange process can be measured by

$$k = \pi(h_{ex} - h_o) \quad \text{Eq. 13}$$

where k is the apparent rate constant, h_{ex} and h_o are the peak widths at half max (FWHM) of the exchanging peak and a non-exchanging peak, respectively. Using the program Sparky³¹ peak widths were measured in both the ^1H and ^{15}N dimension for the four peaks of Naq observed for the exchange process and four peaks of significantly lower intensity seemingly unassociated with either the predominant P2Naq species or the exchange process. Averages of the observed peak widths were used in the calculation of k as described in **Eq. 13**. The value of k measured in the proton dimension is 14 s^{-1} ; that for the nitrogen dimension was 2.5 s^{-1} . The range for the FWHM values for the ^{15}N dimension for both the exchanging peaks and non-exchanging peaks overlapped significantly so the value is quite likely zero.

Scheme 3.3. After Rich,³³ the major route of equilibration between quinones and hydroquinones in aqueous media at neutral to acidic pH, where proton transfer is faster than electron transfer. NQH_2 and NQH^- are the doubly protonated and singly protonated forms of the naphthohydroquinone, respectively, NQH^\bullet and $\text{NQ}^{\bullet-}$ are the singly and fully deprotonated forms of the naphthosemiquinone, respectively, and NQ is naphthoquinone.

Hydroquinone



A major route of the equilibration between reduced and oxidized quinones in aqueous solution reaction is well established,³³ Scheme 3.3; the rate limiting step is the bimolecular reaction between a deprotonated hydroquinone and a quinone forming a pair of unstable semiquinones which after proton exchange react with each other in a diffusion controlled manner producing a new quinone and hydroquinone. Rich³³ determined the rate of electron transfer between quinones of known midpoint and semiquinone potential differences, which provide a good benchmark to compare the observed electron transfer rate in the mixed P2Naq NMR sample. It was observed that the logarithm of the rate scaled linearly with the driving force for the rate-limiting step creating the pair of unstable semiquinones

$$\Delta G_o = E_m(QH^- / QH^\bullet) - E_m(Q^{\bullet-} / Q) \quad \text{Eq. 14}$$

where ΔG_o is the driving force for the rate limiting step, and $E_m(QH^- / QH^\bullet)$ and $E_m(Q^{\bullet-} / Q)$ are the midpoint potentials for the single electron oxidation of singly deprotonated hydroquinone and the single electron reduction of quinone to the anionic semiquinone, respectively. The rate calculated from the observed line broadening is a first order rate, which must be converted to an effective second order rate by dividing by the concentration of the interacting species. The sample from which the rate was calculated was approximately 1.6 mM total lanthanide bound P2Naq, and by NMR, approximately 40% was reduced. Using the lesser concentration species as the entity which is getting reduced, since in effect the 20% excess oxidized P2Naq is not involved in the reaction, 0.64 mM of each species was reacting in solution giving a second order rate constant of $2.2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ using the value of the exchange rate obtained from line broadening in the proton dimension. Using the values for $E_m(QH^- / QH^\bullet)$ and $E_m(Q^{\bullet-} / Q)$ for 2-methyl-1,4-naphthoquinone calculated from reference³⁴ and assuming approximate the pKa, 12 of the hydroquinone dianion, ΔG_o is -370 mV, which according to the studies by Rich³³ corresponds to a rate of approximately $10^5 \text{ M}^{-1} \text{ s}^{-1}$, which is reasonably close to the measured second order rate constant given the need to estimate the pKa and the likelihood that fewer collisions are productive when the quinone is attached to a peptide chain. It should be noted that this might be the first time that a self-quinone redox reaction was observed, since NMR is the only technique suited for its measurement.

3.8 Measurement of the Change in Midpoint of Naq from Coil to Helix

All of the data presented thus far was collected with the goal of measuring $\Delta E_{m,ac}$, the change in Naq's midpoint potential from coil to helix; a value that would represent the difference in helical propensity of the two forms of the amino acid in an ideal system. Yet P2Naq is not ideal. The fraction helicity from the CD data calculated with the assumption that RCCNaq is a random and lanthanide bound P2Naq is fully helical using **Eq. 3**, Figure 3.21A, is compared to that of the NMR data using **Eq. 8** with the same assumptions, Figure 3.21B, or assuming that P2NaqExt is fully helical at Naq₁₄, Figure 3.21C. The results are also tabulated in Table 3.4, and additionally compared to the results of calculations considering the observed Naq chromophore CD signature as discussed in Section 3.5.5.

It is apparent that the assumption of full helicity for lanthanide bound P2Naq is erroneous. While this was somewhat clear from the discussion of the CD data, it is made especially clear when considering that using the same assumptions with the NMR data, in

Table 3.4. Computed fractional helicities.

| Peptide | Ala ₁₃ | | Naq ₁₄ | | | | Ala ₁₅ | |
|-------------------------|-------------------|-------------------|---------------------|-------------------|-------------------------|-------------------|--------------------|--------------------|
| | NMR _{CD} | NMR | CD | NMR _{CD} | NMR | CD _{Naq} | NMR _{CD} | NMR |
| <i>Ligand Free</i> | | | | | | | | |
| P2Naq _{ox} | 0.48 ^a | 0.40 ^b | 0.22 ^c | 0.53 ^a | 0.20^b | 0.08 ^d | -5.40 ^a | 0.13 ^b |
| P2Naq _{red} | 0.48 ^a | 0.57 ^b | 0.06 ^c | 0.60 ^a | 0.29^b | | 0.71 ^a | 0.09 ^b |
| <i>Lanthanide Bound</i> | | | | | | | | |
| P2Naq _{ox} | 1.00 ^a | 0.94 ^b | 1.00 ^{a,c} | | 0.38^b | 0.36 ^d | 1.00 ^a | -0.01 ^b |
| P2Naq _{red} | | 1.35 ^b | | | 0.49^b | | | 0.16 ^b |

^aNMR_{CD}: Fraction helicities according to **Eq. 8** assuming holo P2Naq are fully helical; averaging values obtained for ¹³C^α and ¹H^α if available. ^bNMR: Fraction helicities according to Eq. 8 assuming P2NaqExt are fully helical; averaging values obtained for ¹³C^α and ¹H^α if available. ^cCD: Fraction helicities according to **Eq. 3**; assuming holo P2Naq are fully helical. ^dCD_{Naq}: Fraction helicities according to **Eq. 3** using the assumptions discussed in Section 3.5.5. Values in bold are used to calculate coil to helix midpoint potential shift.

effect ignoring the P2NaqExt control, Figure 3.21B, wildly disparate results are observed for Ala₁₅. In addition, under these assumptions with the NMR ¹H^α chemical shifts, the projected percent helicity of Naq₁₄ in apo P2Naq is well above the values reported for any amino acids in the studies of the Makhatazde group,⁸ and seemingly significantly different (NMR 53% versus 22% CD for P2Naq_{ox}, and 60% NMR versus 6% CD for P2Naq_{red}) than the values obtained by the CD data itself. This later disparity strikes at the core argument made by Richardson *et al.* that the measured CD helicities are good proxies for the helicity at the guest site in the P2X peptides. This is clearly not a universal rule. In addition, since lanthanide bound P2Naq in both redox states does not become fully helical, it is impossible to ascertain the helical propensity of Naq in both redox state with respect to alanine in P2A; such a comparison would ignore the effects of helix breaking and capping that are at work in the P2Naq peptides.

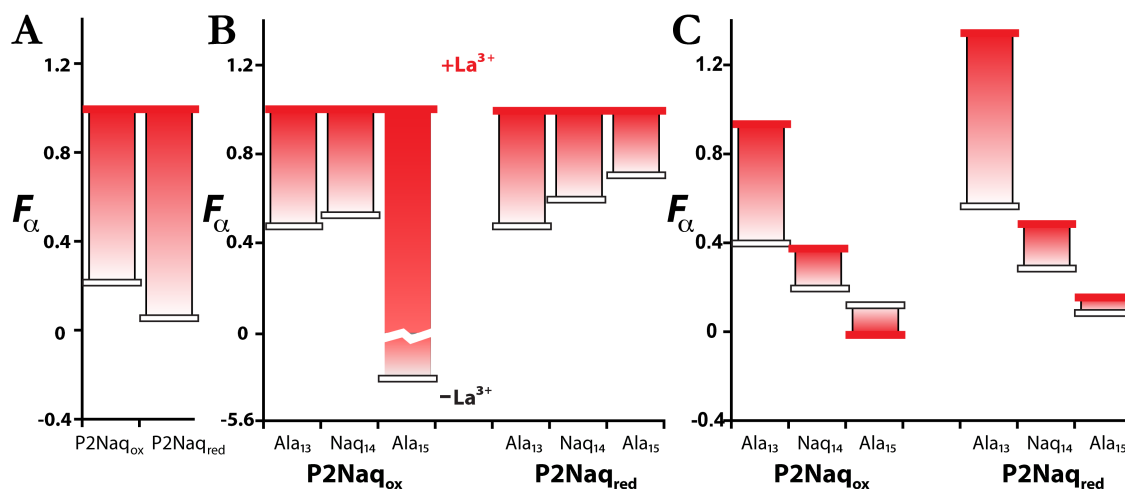


Figure 3.21. Fractional helicities calculated using CD and NMR data as described in the text. Gradients are to guide the eye between the fractional helicities of apo (outlined white boxes) and lanthanide bound P2Naq (filled red boxes). In (A) and (B), lanthanide bound P2Naq are assumed fully helical; in (C) the NMR helical control P2NaqExt is taken as the helical reference state.

On the other hand, while the data set is not complete due to the absence of the fractional helicity from the $^{13}\text{C}^\alpha$ values, with the assumption that the P2NaqExt are fully helical, the $^1\text{H}^\alpha$ numbers provide a clearer description of what is happening in the peptides. It appears that there is limited if any helicity in Ala₁₅ even with saturating lanthanide in the P2Naq peptides. Ala₁₃ behaves as expected, going from approximately 50% helical in both redox states of P2Naq to fully helical upon saturating lanthanide. The exceptionally high 134% helicity for Ala₁₃ the lanthanide bound P2Naq_{red} appears to be due to a slightly higher than anticipated Ala₁₃ $^1\text{H}^\alpha$ chemical shift in P2NaqExt_{red} as is evident in Figure 3.16. A small adjustment of the chemical shift by 0.03 ppm, the width of the peak in the HNHA spectrum, brings the value down to a more reasonable 107% helicity. Naq₁₄ in both oxidized and reduced P2Naq increases in total fractional helicity upon lanthanide binding, while not becoming fully helical. The redox average values for fraction helicity obtained by NMR in both the lanthanide unbound (25%) and bound forms (43%), are close to the values calculated by using the redox difference CD signal of the Naq chromophore at 243 nm (8% and 36%). The better match to the lanthanide bound peptides might be attributable to better signal-to-noise in significantly helical samples. Nevertheless, it is a nice confirmation that the NMR results are somewhat correlative to more global measures of helical content, since the CD signal of the chromophore is reporting on much larger scale environmental effects.

Finally, using the fractional helicities determined by $^1\text{H}^\alpha$ NMR using RCCNaq as a random coil control and P2NaqExt as a helical control, a measure of $\Delta E_{m,ac}$ can be made using the lanthanide binding affinities corrected by the lanthanide affinity of P1 as described above and **Eq. 7**. Without any further assumptions, $\Delta E_{m,ac}$ is measured to be -16 mV. Assuming that the NMR fractional helicities can be better represented by averages between

the values of the two redox states, $\Delta E_{m,ac}$ is -13 mV. These are not large numbers, roughly 0.7 kcal mol⁻¹, which alone is not likely to destabilize a long helix upon a redox change (to some degree evidenced by the P2NaqExt control). However, it is possible that by placing multiple Naq's along a helix, it might be possible to create a redox sensitive conformational switch. The switch may be highly cooperative depending upon whether nearby Naq's affect each other's midpoint potentials. Without such cross talk, it is anticipated that by reducing one Naq and partially destabilizing the helix, the remainder will follow more readily.

3.9 Conclusion

A fundamental assumption in this analysis has been that in P2Naq the only major difference in structure in the two different redox states is associated with Naq, that is, that all of the other amino acids adopt the same conformation and interact with each other in the same way. While this may be true for amino acids prior to Naq, namely the preceding alanine and the entire lanthanide ion binding motif represented by P1, given both the NMR observations of Ala₁₃ and the fact that the peptides both still bind lanthanide ions in an expected range, the residues following Naq are clearly unstructured and may be adopting different unstructured states. However, it is not anticipated that the last two residues of P2Naq will vary from either their random coil state, suggested to be true by the NMR of Ala₁₅, or, perhaps more significantly, between reduced and oxidized P2Naq thus allowing their energetic contributions to be balanced upon subtracting the corrected free energies of binding.

The final result of this work is a measure of the change in midpoint potential of Naq on going from a coil state to a helical state. The determined $\Delta E_{m,ac}$ is not large,

approximately -16 mV or 0.7 kcal mol⁻¹, a pittance compared to the general stability of most proteins (on the order of 7 kcal mol⁻¹ for example, for the holo four-helix bundle cytochrome *b562*).³⁵ It is difficult to say exactly what the source of the 16 mV difference is. The value of $\Delta E_{m,oc}$ is approximately on the order of a hydrogen bond, and while it is certainly likely that the ability for the side chain to hydrogen bond in the helical state is reduced in comparison to the random coil state, it is difficult to ascertain how this might effect differentiation between the two states. The change in midpoint might also arise from other sources, like the presence of a helix dipole perturbing, however incrementally, the pKa of the reduced amino acid, or perhaps it is even more directly related to the helical rotamer(s) of Naq and the ability of the methylene protons to hyperconjugate into the ring system. Regardless, it is clear that in short helices or near the end of long helices Naq is quite likely a helix breaker in both redox states. While this certainly limits the range of structural environments Naq can be placed, it might be possible to design around the helix-breaking tendency by introducing competitive hydrogen bonding partners that selectively stabilize Naq's helical rotamer(s). It may be possible to do this in a catholic manner, using side chain amides of asparagine and glutamine or the hydroxyls of threonine and serine to serve as hydrogen bond donors and acceptors for the carbonyls of the oxidized form or the hydroxyls of the reduced form, respectively. It is important to realize, however, that in the P2NaqExt peptides, without any designed interaction, Naq was integrated into a stretch of helix which had all of the hallmarks of a natural helix, ranging from the CD spectrum to the observation of backbone amide to backbone amide NOE's to the measure of ϕ , which in the case of P2NaqExt_{ox} was nearly ideal for the stretch of AQA. However, in the future care will certainly need to be taken on incorporating Naq near the end of helices.

3.10 References

- (1) Hu, J. S.; Bax, A. J. *Am. Chem. Soc.* **1997**, *119*, 6360.
- (2) Pace, C. N.; Scholtz, J. M. *Biophys J* **1998**, *75*, 422.
- (3) Chou, P. Y.; Fasman, G. D. *Biochemistry-U.S.* **1974**, *13*, 211.
- (4) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Biochemistry-U.S.* **1997**, *36*, 10923.
- (5) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Proc Natl Acad Sci U S A* **1997**, *94*, 2833.
- (6) Park, S. H.; Shalongo, W.; Stellwagen, E. *Biochemistry-U.S.* **1993**, *32*, 7048.
- (7) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646.
- (8) Richardson, J. M.; Lopez, M. M.; Makhatadze, G. I. *Proc Natl Acad Sci U S A* **2005**, *102*, 1413.
- (9) Lopez, M. M.; Chin, D. H.; Baldwin, R. L.; Makhatadze, G. I. *Proc Natl Acad Sci U S A* **2002**, *99*, 1298.
- (10) Siedlecka, M.; Goch, G.; Ejchart, A.; Sticht, H.; Bierzynski, A. *Proc Natl Acad Sci U S A* **1999**, *96*, 903.
- (11) Munoz, V.; Serrano, L. *Nat Struct Biol* **1994**, *1*, 399.
- (12) Goch, G.; Maciejczyk, M.; Oleszczuk, M.; Stachowiak, D.; Malicka, J.; Bierzynski, A. *Biochemistry-U.S.* **2003**, *42*, 6840.
- (13) Ross, J. B.; Wyssbrod, H. R.; Porter, R. A.; Schwartz, G. P.; Michaels, C. A.; Laws, W. R. *Biochemistry-U.S.* **1992**, *31*, 1585.
- (14) Ross, J. B.; Laws, W. R.; Buku, A.; Sutherland, J. C.; Wyssbrod, H. R. *Biochemistry-U.S.* **1986**, *25*, 607.
- (15) Laws, W. R.; Ross, J. B.; Wyssbrod, H. R.; Beechem, J. M.; Brand, L.; Sutherland, J. C. *Biochemistry-U.S.* **1986**, *25*, 599.
- (16) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135.
- (17) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry-U.S.* **1992**, *31*, 1647.
- (18) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2006**, *128*, 13586.
- (19) Du, H.; Fuh, R. C. A.; Li, J. Z.; Corkan, L. A.; Lindsey, J. S. *Photochem. Photobiol.* **1998**, *68*, 141.
- (20) Rishavy, M. A.; Pudota, B. N.; Hallgren, K. W.; Qian, W.; Yakubenko, A. V.; Song, J. H.; Runge, K. W.; Berkner, K. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13732.
- (21) Dowd, P.; Hershline, R.; Ham, S. W.; Naganathan, S. *Science* **1995**, *269*, 1684.
- (22) Siebum, A. H. G.; Woo, W. S.; Lugtenburg, J. *Eur. J. Org. Chem.* **2003**, 4664.
- (23) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun (Camb)* **2009**, 168.
- (24) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 5286.
- (25) Agrawal, Y. K.; Thomaskutty, P. T. *Analyst* **1987**, *112*, 1015.
- (26) Fonseca, B. M.; Saraiva, I. H.; Paquete, C. M.; Soares, C. M.; Pacheco, I.; Salueiro, C. A.; Louro, R. O. *J. Biol. Inorg. Chem.* **2009**, *14*, 375.
- (27) Ershov, B. G. *Russ. Chem. Bull.* **2001**, *50*, 626.
- (28) Chen, C. W.; Tano, D.; Akashi, M. *J. Colloid Interface Sci.* **2000**, *225*, 349.
- (29) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wuthrich, K. *J. Biomol NMR* **1998**, *12*, 1.

- (30) Vuister, G. W.; Bax, A. J. *Am. Chem. Soc.* **1993**, *115*, 7772.
- (31) Goddard, T. D.; Kneller, D. G. *University of California, San Francisco.*
- (32) de Alba, E.; Tjandra, N. *J Biomol NMR* **2006**, *35*, 1.
- (33) Rich, P. R. *Biochimica Et Biophysica Acta* **1981**, *637*, 28.
- (34) *Physical Chemistry of Semiquinones*; Swallow, A. J., Ed.; Academic Press: New York, 1982.
- (35) Kimura, T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc Natl Acad Sci U S A* **2009**, *106*, 7834.

Chapter 4: Differentiating Oxidized and Reduced Naq in Designed Protein

4.1 Introduction

The projects described in Chapters 2 and 3 establish Naq as a well-defined new quinone cofactor demonstrating the full range of quinone electrochemical properties observed in natural system with real, but perhaps limited, perturbation of short helices. An observed moderate change in midpoint potential upon going from a random coil state to a helix is tantalizing evidence that Naq's environment may have larger measurable electrochemical effects. Indeed, it may be possible to design two structurally identical peptides that separately tune the environment around Naq to selectively stabilize the neutral oxidized (quinone) or neutral reduced (hydroquinone) redox states. The measured midpoint potentials of Naq in peptides designed with such differentiation in mind should separate according to the redox state that was stabilized in the design process. So in considering a design that stabilizes the introduction of an oxidized Naq cofactor, the midpoint potential should become more negative compared to an undersigned reference state, in essence,

making it more energetically costly to introduce the reduced ‘mutation.’ The opposite should hold true for a peptide designed to stabilize the hydroquinone for the same reasons. In effect, by driving the design of the peptides towards stabilizing particular redox states there should be a measurable difference in their respective midpoint potentials, Figure 4.1. The likely order of this difference is unknown, but given that the two forms of the amino acid are

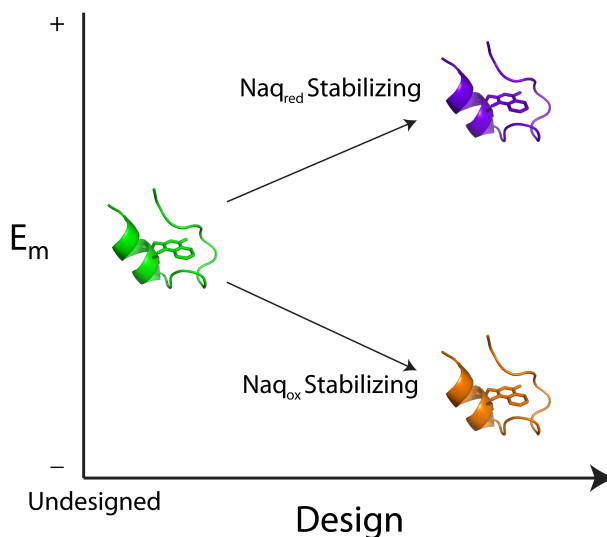


Figure 4.1. The effect of progressively and independently designing proteins to stabilize only one of two different redox states of Naq. As designs separately become more stable through specific interactions with Naq in its different redox states, a difference in midpoint potential between the peptides will develop. This cannot progress indefinitely; rather, there is an unknown limit where the design scaffold reaches its maximal differential stabilization of the two redox states. This is true even if negative design is introduced to destabilize the opposing redox state. In nature, however, evolution is not limited to considering only two forms of a quinone. At least six quinone states are at play in natural systems, each with their own evolving axis of stabilization.

neutral under most pH conditions, perturbation is likely to be manifest by either the significant differences in hydrogen bonding geometry preferred by the two forms of Naq’s side chain, or via more generic and possibly less energetic interactions affected by the difference in electron density of the two redox states. The largest differentiating electrostatic effects might be driven by cation-pi interactions,^{1,2} but contacts such as these would also include pi-pi interactions as well as subtle changes in van der Waals (vdW) forces due to polarizability changes between the two redox states. It

should be clear that if the design process works, the observed differentiation will provide new information with respect to the extent and ability of natural systems to modify the electrochemistry of generic quinone cofactors.^{3,4}

The typical approach towards maquette design in the Dutton laboratory is to start with a candidate sequence that is likely to adopt a particular structural fold and iterate over several generations until the desired properties are obtained.⁵ While this approach has succeeded primarily because it is an evidence based design process,⁶ it is not likely to provide enough structural resolution and global consideration of the energy landscape of a peptide to allow the development of two structurally similar peptides designed to selectively stabilize different redox states of Naq. On the other hand, a variety of computational protein design approaches^{7,8} exist that might be able to address this question. The choice of the Saven laboratory's statistical computationally assisted design strategy (*SCADS*) is obvious since, while it requires the input of initial backbone coordinates for a target design, it is not exclusively dependent upon natural structures to generate pair-wise interaction energies between amino acids, is fast and it generates a testable number of possible designs which can be further optimized in an iterative computational approach not unlike the hand-designed maquette approach. In addition, as will be discussed below, *SCADS* has been used in the redesign of a short miniprotein which is an ideal starting point for a structural scaffold incorporating Naq.⁹

SCADS uses an informational entropy approach to generate site-specific and amino acid-specific probabilities of stabilizing mutations.¹⁰ The sequence entropy derived directly from these probabilities is constrained using mean-field energy terms that are dependent upon the amino acid probabilities, allowed rotamers, pair-wise interactions with nearby residues and backbone geometry. Numerically maximizing the sequence entropy subject to these constraints using Lagrange multipliers gives a list of site-specific probabilities for the ensemble of amino acids considered. This approach allows sequence optimization from

single sites and a handful of amino acids to an entire protein considering all amino acids. In addition, non-bonding energies are pre-calculated in a pair-wise manner using the AMBER force field so the algorithm is not computationally taxing. One advantage that this approach allows is that the energetic constraints can implicitly incorporate the energies of an ensemble of unfolded states.

One of the concerns with any design process is blindly destabilizing the protein target through a build up of subtly destabilizing mutations. This is particularly true when the target protein is small enough to be produced on a peptide synthesizer, as is the case with any simple design incorporating the Naq cofactor. Fortunately, SCADS has been used to stabilize the smallest known independently folding protein.⁹ At 20 amino acids, the Andersen group's ^wTrpCage is structured, cooperatively folds and consists of an N-terminal helix, 3₁₀ helical loop, followed by a proline rich C-terminal tail, Figure 4.2.¹¹ It folds around a centrally buried tryptophan, which additionally makes a stabilizing hydrogen bond to the amide backbone of a C-terminal Arg₁₆, thereby 'stapling' the structure together.^{11,12} In mutagenesis

N₁L₂Y₃I₄Q₅W₆L₇K₈D₉G₁₀G₁₁P₁₂S₁₃S₁₄G₁₅R₁₆P₁₇P₁₈P₁₉S₂₀

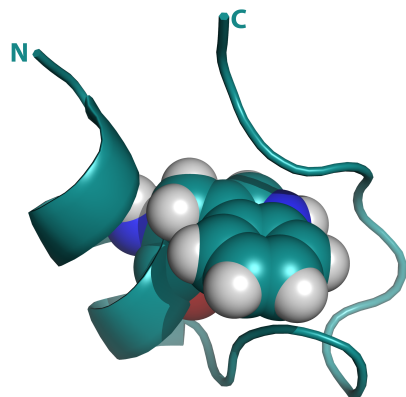


Figure 4.2. A representative NMR structure of the TrpCage. Primary sequence shown at top for reference; buried ^wTrp₆ is shown in spheres.

studies, the ^wTrp₆ was found to be essential; significantly, in mutants that showed significant 'cage' structure, the only other amino acids tolerated at position 6 were naphthalanine residues, which are structural homologs to Naq, suggesting that burial of the large planar hydrophobic ring system of tryptophan drives protein folding.¹³ In addition, there is evidence for stabilizing salt

bridges between ^{wt}Asp₉ and ^{wt}Arg₁₆, and the N-terminal amine and C-terminal carboxylate.^{11,12} A second helix stabilizing interaction exists between ^{wt}Asp₉ and ^{wt}Gln₅.^{12,14} In a subsequent paper published after the initiation of this work, the Andersen group used alanine scanning mutagenesis to optimize the stability of the ^{wt}TrpCage fold finding an approximately 14 °C increase in melting temperature in a protein that had two alanine mutations (L2A and I4A), and an N1D capping mutation in the N-terminal helix.¹² The stability of the helix and the van der Waals interaction of ^{wt}Trp₆ with ^{wt}Pro₁₉ proved above all else to be essential for the mini-protein to both cooperatively fold and achieve the TrpCage structure.

Given the ^{wt}TrpCage's small size, which eases synthesis and structural studies, the relatively high contact order of the buried tryptophan, partial tolerance of insertion of the Naq homolog 2-naphthalanine at the site of the tryptophan and the successful use of *SCADS* in redesigning the protein, it is the natural choice for a structural framework for separately stabilizing reduced and oxidized Naq in a peptide. In addition, since the Naq₆ site would likely be partially buried, it would provide an opportunity to observe how changes in protein environmental hydrophobicity affect Naq's redox properties, an environmental constraint that was not adjustable in heptaNaq (Chapter 2) or the P2Naq (Chapter 3) peptides. Selectively redesigning the ^{wt}TrpCage to incorporate either Naq_{ox} or Naq_{red} should produce so-called NaqCages with different electrochemical midpoints.

4.2 SCADS Redesign of ^{wt}TrpCage into Redox State Specific NaqCages

Initial attempts to generate usable *SCADS* output by the author were unsuccessful and as the collaboration developed with the Saven laboratory, Seung-gu Kang took the

initiative to generate the designs discussed below. The work described in this section is that of Seung-gu Kang, a graduate student in Jeff Saven's laboratory.

4.2.1 Optimizing ^{wt}TrpCage Scaffold

The NMR structure of the ^{wt}TrpCage¹¹ consists of the top 38 models that matched the NMR constraints. However, *SCADS* requires the input of a single backbone pdb structure for mutagenic optimization. The individual structure which showed the lowest overall z-value (deviation from the mean) for the structural parameters (ϕ , ψ , χ_1 , χ_2) defining the local geometry of Trp_g and which visually confirmed the known stabilization ^{wt}Asp₉-^{wt}Arg₁₆ salt bridge was intact was selected for use as a starting backbone (model 34 in 1L2Y.pdb). Unfavorable van der Waals contacts in the chosen structure were eliminated by minimizing the energy of the structure using the AMBER potential parameters with the secondary structures constrained.

4.2.2 Incorporating Naq into SCADS Amino Acid Library

To make good functional use of the Saven group's design algorithm *SCADS*, it is necessary to supply both forms of the non-natural naphthoquinone amino acid, Naq, structural details that readily exist for natural amino acids including calculating the partial charges on the atoms for AMBER force field calculations, generation of rotamer families, one-body backbone interactions, and pre-calculating two-body interactions with the rotamers of other amino acids. This process begins with the calculation of the structure of the amino acid using geometry optimization in Gaussian03 with the HF/6-31G* quantum

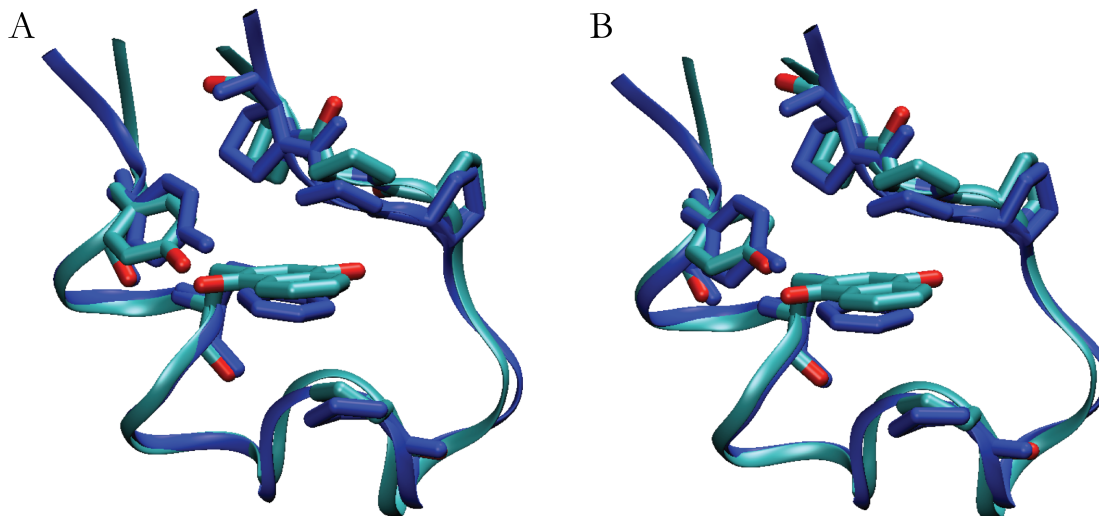


Figure 4.3. Optimized scaffold for *SCADS*. Optimized geometry for $^{\text{ox}}\text{NaqCage}$ backbone overlaid upon TrpCage model 34 from NMR structure (r.m.s.d 0.880 Å from wildtype) (A). Optimized geometry for $^{\text{red}}\text{NaqCage}$ backbone overlaid upon TrpCage model 34 from NMR structure (r.m.s.d 0.289 Å from wildtype) (B). Images created by Seung-gu Kang. See text for details.

basis sets. van der Waals parameters were obtained from the AMBER force field. Partial charges were calculated by fitting the Gaussian generated surface electrostatic potential with restrained electrostatic potential (RESP) charge fitting using the RED II tool. This process generates values for point charges at the sites of atoms, which are needed for AMBER force field calculations. For Naq_{red} , since the hydroxyl protons on the side chain can adopt several conformations, the side chain protons were discretely modeled and the partial charges assigned to the atoms in Naq_{red} were Boltzmann weighted averages of the atomic charges obtained from each possible conformation.

The values derived for the partial charges of both redox states of Naq were used in defining the scaffold to be given to *SCADS*. Separately Naq_{ox} and Naq_{red} were substituted into the $^{\text{wt}}\text{Trp}_6$ site of the optimized $^{\text{wt}}\text{TrpCage}$ described above. In both cases, the geometry of the backbone was further optimized using AMBER to accommodate the Naq side chain. The end structures were verified by ϕ - ψ analysis, with all but $^{\text{wt}}\text{Gly}_{15}$ in the favored regions

of the Ramachandran plot; it is notable that ^{wt}Gly₁₅ in the wild type TrpCage can be substituted with D-Ala in a strongly stabilizing mutation.¹⁵ R.m.s deviations from the backbone of the ^{wt}TrpCage starting structure were computed: 0.880 Å between the NaqCage_{ox} and the ^{wt}TrpCage, and 0.289 Å between NaqCage_{red} and the ^{wt}TrpCage, Figure 4.3. Several interesting features were observed in these geometry optimized initial states. In the case of the NaqCage_{ox}, though not observed, the Naq side chain carbonyls can possibly hydrogen bond to the backbone amide proton of ^{wt}Leu₇ (a likely helix breaking interaction) and the amide proton of a residue at position 18 (in the case of the starting structure, a ^{wt}Pro₁₈ which does not have an amide). In the case of NaqCage_{red}, due to the hydroxyls of the side chain being hydrogen bond donors and acceptors, possible backbone interactions exist with the carbonyl of ^{wt}Tyr₃ and the amide of ^{wt}Leu₇ (both helix breaking), and the carbonyl of ^{wt}Arg₁₆, a possibly stabilizing long distance buried polar contact.

Although it is possible to force the Naq side chains to take the geometry proscribed in the optimized scaffolds seen in Figure 4.3, to further expand the capabilities of *SCADS* with respect to Naq side chain flexibility, rotamer families were generated. Using the backbone geometry defined by the optimization procedure, a rotamer (χ_1, χ_2 in five degree increments) potential energy surface of Ac-Naq-NHMe in both redox was determined, Figure 4.4. Rotamer libraries were built from rotamers within 1.0 kT above the local minima, giving 38 rotamers for Naq_{ox} and 22 rotamers for Naq_{red}. While this procedure is slightly biasing towards the backbone geometry calculated in the optimization procedure, without an actual structure of Naq in a helix, this is the only method available to provide Naq with side chain flexibility. It is notable that the rotamer adopted by ^{wt}Trp₆ in the wild type protein is near one of the local minima for both redox states of Naq. Finally, one-body interactions of

the rotamers in the library with the backbone were calculated, as were the two-body interactions between these rotamers and the rotamers of natural amino acids from Roland Dunbrack's library.¹⁶

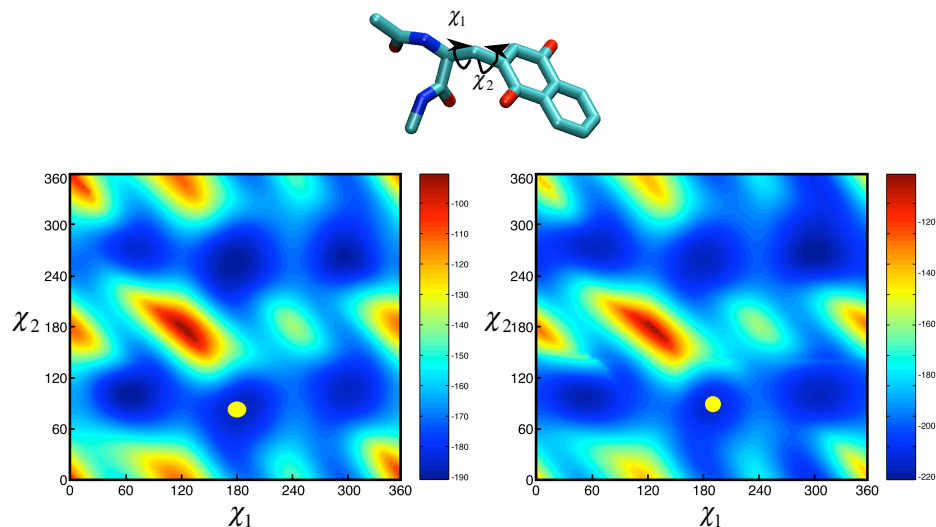


Figure 4.4. Potential energy surface of side chain dihedral angles χ_1 and χ_2 of Ac-Naq_{ox}-NHMe (left) and Ac-Naq_{red}-NHMe (right). The image at top is of Ac-Naq-NHMe. The yellow circle in either potential energy surface indicates the local minima in which the ^wTrp₆ conformation resides in model 34 of the TrpCage structure. Images and figure layout created by Seung-gu Kang. See text for details.

4.2.3 Identification of Mutation Sites

SCADS was run incorporating Naq_{ox} and Naq_{red} as described above. The ^{mut}Naq₆ site was left unvaried but allowed to adopt rotamers as determined above, generating probabilities for each rotamer as a natural result of the *SCADS* algorithm. All other 19 sites were individually allowed to mutate to all natural amino acids except Cys. This procedure allowed the calculation of site specific frustration values defined as $\ln(w_i^{mp} / w_i^{wt})$, where w_i^{mp} is the site specific *SCADS* generated probability of finding the most probable amino acid at site *i*, and w_i^{wt} is the same for the ‘wild-type’ amino acid. This procedure was performed at two ‘temperatures’ represented by $\beta = 1/kT$ values of 0.5 mol kcal⁻¹ and 2 mol kcal⁻¹.

Frustration values so defined represent a measure of the comparative stabilization of mutations made against a wild type background. To gauge the selective effects of mutations on the stability of one redox state versus the other an average, site-specific, relative entropy was calculated

$$\bar{S}_{\text{rel},i} = \frac{1}{2} \sum_{\alpha=1}^{19} (w_i^{\text{ox}}(\alpha) - w_i^{\text{red}}(\alpha)) \ln(w_i^{\text{ox}}(\alpha) / w_i^{\text{red}}(\alpha))$$

where $w_i^{\text{ox}}(\alpha)$ and $w_i^{\text{red}}(\alpha)$ are the probabilities of finding amino acid α at site i in the Naq_{ox} stabilizing design ($^{\text{ox}}\text{NaqCage}$) or Naq_{red} stabilizing design ($^{\text{red}}\text{NaqCage}$), respectively. The $\bar{S}_{\text{rel},i}$ value gives a measure of the importance of a specific site with regards to stabilizing one form of the amino acid versus the other, thereby allowing a selection of mutants based upon both the broad stabilization of the fold, high frustration, low $\bar{S}_{\text{rel},i}$, and redox-specific interactions, moderate to high frustration, high $\bar{S}_{\text{rel},i}$. Selection of globally stabilizing mutations and redox state differentiating mutations can then be made.

The top eight frustration sites were selected for inspection. This was more than simply a practical measure, as beyond about eight sites conservative mutations began to dominate suggesting that no further useful information was to be derived from analysis. Of the eight sites, five were eliminated from consideration. Mutations at $^{\text{wt}}\text{Lys}_8$, $^{\text{wt}}\text{Ser}_{13}$, $^{\text{wt}}\text{Gly}_{15}$ and $^{\text{wt}}\text{Pro}_{17}$ were eliminated since they were highly solvent accessible and therefore had little interaction with $^{\text{mut}}\text{Naq}_6$. In addition, $^{\text{wt}}\text{Gln}_5$ and $^{\text{wt}}\text{Ser}_{13}$ were targeted for conservative mutations at moderate temperature. While $^{\text{wt}}\text{Gln}_5$ was suggested to mutate to either Asn or Asp at moderate or low temperature, respectively, this site was left un-mutated given its role in the QXXXD helix stabilization motif. Finally, with $^{\text{wt}}\text{Lys}_8$ it was clear that mutation might

locally stabilize the helix but long distance interactions with other mutant sites, I4D for instance, appeared to be probable, something the single site analysis would miss.

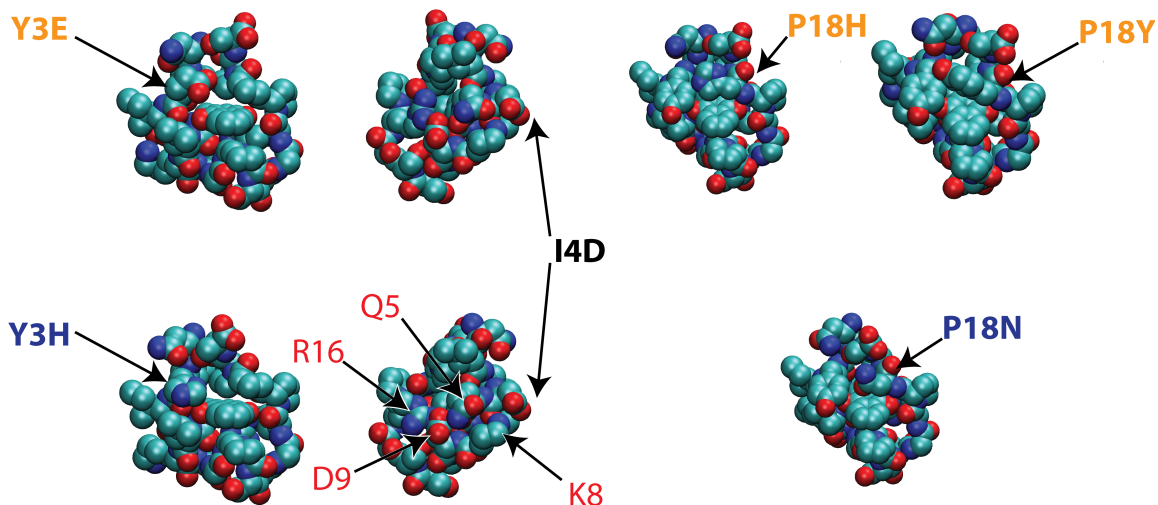


Figure 4.5. Identification of SCADS mutation sites. Orange (top) sites correspond to mutations selected for stabilization of $^{\text{ox}}\text{NaqCage}$. Blue (bottom) sites correspond to mutations selected for stabilization of $^{\text{red}}\text{NaqCage}$. Black (middle) site corresponds to identical mutation suggested to stabilize the fold of both peptides. Red sites (bottom-center) correspond to sites known for their stabilizing or potentially stabilizing interactions in the $^{\text{wt}}\text{TrpCage}$. Structures were generated by Seung-gu Kang as output from SCADS.

Three sites were ultimately chosen for mutation: the aforementioned $^{\text{wt}}\text{Ile}_4$ which is highly frustrated in both redox states, and $^{\text{wt}}\text{Tyr}_3$ and $^{\text{wt}}\text{Pro}_{18}$ which are frustrated but also have a significant $\bar{S}_{\text{rel},i}$, Figure 4.5. Mutation of Pro_{18} is absolutely required due to significant steric contact between $^{\text{mut}}\text{Naq}_6$ and the rigid $^{\text{wt}}\text{Pro}_{18}$ potentially disrupting stabilizing interactions required for maintenance of the $^{\text{wt}}\text{TrpCage}$ fold; in both redesigned proteins site 18 showed the highest $\bar{S}_{\text{rel},i}$. The Naq_{ox} stabilizing design incorporates either a $^{\text{ox}}\text{Tyr}_{18}$ or $^{\text{ox}}\text{His}_{18}$. In both cases the residues make good aromatic vdW contact with the oxidized $^{\text{mut}}\text{Naq}_6$. The suggestion of a mutation of the $^{\text{wt}}\text{Tyr}_3$ to $^{\text{ox}}\text{Glu}_3$ is somewhat surprising given the conservation of the $^{\text{wt}}\text{Tyr}_3$ in the mutants generated by Andersen's lab and its pivotal role in establishing the fold.^{11,12,15} However, long-range electrostatic interactions appear to be

Table 4.1. Sequence for the ^wTrpCage and *SCADS* target sequences for ^{ox}NaqCage and ^{red}NaqCage miniproteins. **Q** is the one letter code for Naq. Sites of mutation are indicated with subscript site numbering. GN is generational number.

| Peptide | GN | Sequence |
|------------------------|----|---|
| ^w TrpCage | 1 | NLY ₃ I ₄ QW ₆ LKDGGPSSGRPP ₁₈ PS |
| ^{ox} NaqCage | | NLE ₃ D ₄ Q Q ₆ LKDGGPSSGRPY ₁₈ PS |
| | | NLE ₃ D ₄ Q Q ₆ LKDGGPSSGRPH ₁₈ PS |
| ^{red} NaqCage | | NLE ₃ D ₄ Q Q ₆ LKDGGPSSGRPN ₁₈ PS |

driving *SCADS*'s selection of this mutation. The Naq_{red} stabilizing design incorporates an ^{red}Asn₁₈ residue that makes good vdW contacts and electrostatic interactions with the hydrophobic core of the redesigned protein. The mutation at ^wTyr₃ is the only mutation seemingly taking advantage of differential hydrogen bonding abilities of the two forms of the amino acid with a ^{red}His₃ mutation capable of hydrogen bonding to the hydroxyls on the reduced ^{mut}Naq₆ residue. One common mutation was the mentioned I4D, which makes good stabilizing electrostatic intrahelical contacts with ^wLys₈. Table 4.1 summarizes the suggested mutations from *SCADS* output.

4.3 Synthesis and Redesign of ^{ox}NaqCage

With three mutant sequences identified as potentially uniquely stabilizing Naq_{ox} and Naq_{red} in the ^wTrpCage fold, it was decided to initially focus on testing the ^{ox}NaqCage mutants; in particular, the mutant **GN 1**, Table 4.1, which had a ^{ox}Tyr₁₈ mutation and showed excellent hydrophobic burial of the Naq side chain. There was also some concern, though unfounded, that the ^{ox}His₁₈ mutation might covalently modify Naq. In addition, at the time of this early work there was no convenient way to keep the NaqCages reduced during the course of a CD or NMR experiment, precluding rapid testing of design sequences. Peptide concentrations were determined by absorbance spectra using the molar absorptivity of the oxidized Naq cofactor determined in Chapter 3.

4.3.1 Synthesis and Deprotection of ^{ox}NaqCage Target Sequences

All ^{ox}NaqCage peptides discussed in this section were synthesized using standard Fmoc-SPPS on Fmoc-Serine Wang resin on a CEM synthesizer ensuring that the C-terminal was a carboxylate and not the more typical amides. This was essential since there is evidence that a stabilizing interaction between the N-terminal amine and C-terminal carboxylate. Fmoc deblock was effected using piperazine instead of the more typical piperidine to avoid succinimide formation in the DG sequence of some of the peptides. Deprotection of the natural amino acids using a standard TFA/scavenger cocktail mix and Naq(OMe/OMe) using the DDQ procedure described in Chapter 3 proceeded without issue. As with all other peptides discussed in the preceding chapters, the UV/*vis* spectrum of the deprotected Naq containing peptides were unambiguously naphthoquinone. All peptide identities were confirmed by MALDI-MS.

Table 4.2. Sequences for the ^wTrpCage, optimized ^{mut}TrpCage, initial SCADS target sequence **1** for ^{ox}NaqCage, and sequences of iteratively redesigned ^{ox}NaqCage. **Q** is the one letter code for Naq. Sites of mutation are indicated with subscript site numbering. Lower case q is the residue dGln GN is generational number. GN **7c** was selected on the basis of CD spectral analysis as a NMR structure target.

| Peptide | GN | Sequence |
|------------------------|-----------|---|
| ^w TrpCage | | N ₁ L ₂ Y ₃ I ₄ QW ₆ L ₇ KD ₉ G ₁₀ GPS ₁₃ SGRPP ₁₈ PS |
| ^{mut} TrpCage | | D ₁ A ₂ Y ₃ A ₄ QW ₆ L ₇ KD ₉ G ₁₀ GPS ₁₃ SGRPP ₁₈ PS |
| ^{ox} NaqCage | 1 | N ₁ L ₂ E ₃ D ₄ Q Q ₆ L ₇ KD ₉ G ₁₀ GPS ₁₃ SGRPY ₁₈ PS |
| | 2 | D ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ G ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 3a | N ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ G ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 3b | N ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ G ₁₀ GPA ₁₃ SGRPH ₁₈ PS |
| | 4 | N ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 5 | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ G ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 6a | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 6b | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ L ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPH ₁₈ PS |
| | 7a | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ T ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 7b | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ N ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 7c | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ Q ₇ KD ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |
| | 8 | T ₋₃ P ₋₂ E ₋₁ Q ₁ A ₂ E ₃ E ₄ Q Q ₆ Q ₇ KE ₉ q ₁₀ GPA ₁₃ SGRPY ₁₈ PS |

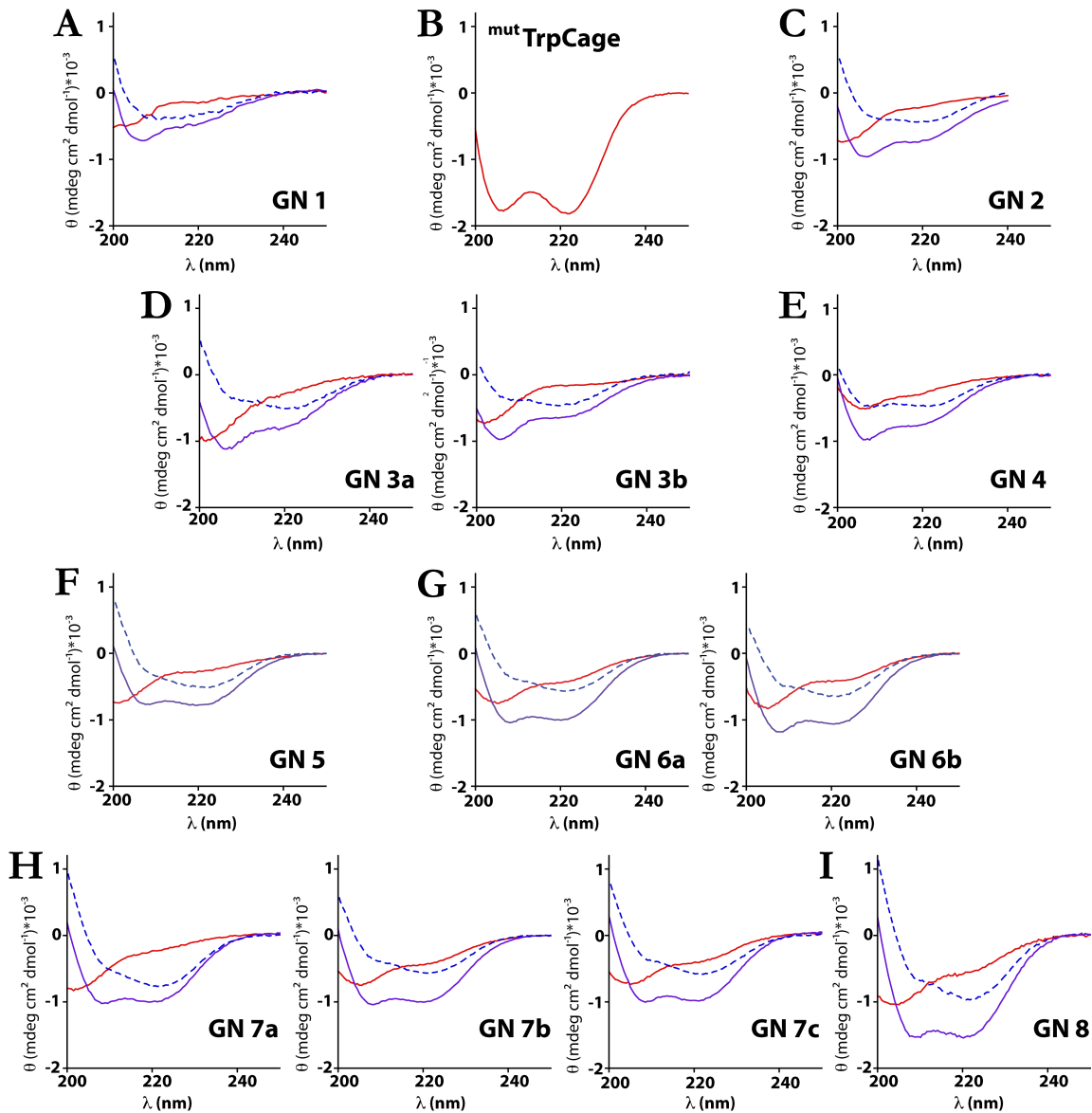


Figure 4.6. CD spectral data of the iteratively redesigned ^{ox}NaqCages and the ^{mut}TrpCage, the primary sequence of which most of the mutant sequences are based. CD spectra were collected at 2 °C in 50 mM K₂PO₄, pH 7 buffer. Samples were equilibrated at 2 °C for 30 min prior to acquisition of data.

4.3.2 Iterative Redesign of ^{ox}NaqCage

The CD spectrum, taken at 2 °C, of target 1, Figure 4.6A, in the absence of TFE is characteristic of a random coil peptide, with perhaps a small amount of alpha-helical character. It compares unfavorably to the CD spectrum of ^{opt}TrpCage (mutant 10b, in the

Andersen naming scheme),¹² Figure 4.6B, which shows a more classical alpha-helix CD response. The strength of the ⁹⁹TrpCage CD signal cannot be compared directly since there is evidence that because the chromophore's CD signal at 226 nm becomes negative and intensifies in helices, the magnitude of the 222 nm band associated with helices projects higher helical content than is actually in the structure.¹¹ In the presence of 30% trifluoroethanol (TFE), a secondary structure stabilizing reagent which acts to lower the barrier to desolvation of the peptide backbone,¹⁷ the CD spectrum appears to indicate the protein is more helical, but still has significant random coil character. The difference CD spectrum shows that TFE did effect a coil-to-helix transition, but it was a small change. This is perhaps not surprising given the results discussed in Chapter 3 that suggest that Naq in both redox states is significantly helix destabilizing, but these results were actually the first indication that Naq might disrupt helices.

The lack of helical CD spectrum was a sign that the fold was not forming properly, and rather than focus on obtaining melting and thermodynamic data for what little fold might have been present, it was decided that it might be better to attempt to get to a peptide that appeared significantly folded (as measured by an alpha-helical CD response) in the presence or absence of TFE. Obtaining an NMR structure of this peptide would then guide further design efforts, including recommitting the obtained structure back into *SCADS*.

While this initial work was underway, Andersen's group published another paper on the TrpCage detailing the results of alanine scanning the protein sequence and identifying several stabilizing mutations that increased the melting point of the mutant protein by 14 °C.¹² After some discussion with Seung-gu Kang, it was decided that it would not likely change the *SCADS* results if the three most stabilizing mutations (N1D, L2A and S13A)

identified by this paper.¹² were incorporated into the target sequence. Additionally, the I4D mutation that *SCADS* recommended was once again changed to ^{ox}Glu₄ since *SCADS* had also recommended this mutation to some degree and it would eliminate the potentially helix breaking residue Asp from the N-terminal of the helix, Table 4.2, GN **2**. The CD spectra of this second generational ^{ox}NaqCage design, Figure 4.6C, varied significantly from the preceding GN **1**, indicating that stabilization of the N-terminal helix effected similar responses in the ^{ox}NaqCage as in the ^{mutr}TrpCage.

The two GN **3** mutants (**3a** and **3b**), Table 4.2 and Figure 4.6D, reverted the N1D mutation since it was apparent that the different capping strength of the Asp and Asn residues was not significant. Mutant **3b** with a ^{ox}His₁₈ residue further tested whether the choice of the P18Y (**3a**) mutation over the P18H was at fault for the non-ideal CD spectral response. Neither mutant was much of an improvement over previous generations, and further confirmed that something specifically related to Naq could be disrupting the fold. In addition, it appeared that the ^{ox}Tyr₁₈ mutant was slightly more helical even in the CD sample lacking TFE, suggesting that the choice of target **1** as a starting point was not an issue.

The significant increase in melting temperature of the mutationally optimized ^{opr}TrpCage over the original ^{wt}TrpCage paled in comparison to the ΔT_m caused by introducing a dGln (lower case q) at the ^{wt}Gly₁₀ site (Vikas Nanda, personal communication). It was believed that this mutation would not only be tolerated, since achiral ^{wt}Gly₁₀ adopts a right handed backbone conformation, but actively stabilize the short 3₁₀-helix by effectively capping it. This confirms the later findings by Andersen's group that by introducing dAla at both ^{wt}Gly₁₀ and ^{wt}Gly₁₅, they could get a ^{mutr}TrpCage that would only partially melt at 83 °C. Thinking that perhaps by stabilizing the fold outside of the helix might encourage the

protein to become structured, a dGln residue was introduced into site 10, produced GN 4, Table 4.2. The CD spectrum of GN 4 without TFE showed the largest deviation from the previous generations, appearing significantly less random coil, but not comparably helical, Figure 4.6E. In the presence of TFE, the CD spectrum looks appreciably identical to all previous generations of the ^{ox}NaqCage under similar conditions. It would appear, then, that the method of stabilization found with the dGln₁₀ residue in the ^{wr}TrpCage is somewhat independent of helix formation, presumably acting mostly on the 3₁₀-helix. In directing further efforts, it became clear that the dGln residue was possibly important for the overall topology of the protein; in addition, coming at the end of the helix it would guarantee that in TFE the alpha-helix would not continue unabated through the 3₁₀-helix since it would be incompatible with right handed alpha-helix propagation.

The need to stabilize the alpha helix of the ^{ox}NaqCage mutants, drove the design of GN 5, Table 4.2. The replacement of the N-terminal Asn₁ residue with the strong helix capping motif TPEQ,¹⁸ afforded a slight increase in the amount of helix in the absence of TFE, Figure 4.6F. However, in the presence of TFE, the spectrum takes on a much more notable alpha-helical character, suggesting that significantly more stabilization of the helix has occurred. It is particularly telling that in the with and without TFE difference CD spectrum, the apparent change occurs mostly at the alpha-helix associated 222 nm band, suggesting that elimination of the random coil, whose CD signal overlaps with the 208 nm helical band, is more complete than in the previous mutant generations.

Combining the successful TPEQ helix cap and the dGln₁₀ mutation allowed the comparison of the site 18 mutants (P18Y (**6a**) and P18H (**6b**)) in a more structured context, GN 6 in Table 4.2. Both peptides demonstrated increased helicity in their CD spectra with

and without TFE compared to previous generations, Figure 4.6G. Once again it appeared that the P18H mutant was slightly less helical, so for the remainder of the mutants, the P18Y mutation was kept.

Hand modeling of the ^{ox}NaqCage using the published structure of the ^{wT}TrpCage and editing in Naq_{ox} at site 6 suggested that if Naq was disrupting the helical fold, it could be doing so by hydrogen bonding to the backbone of the helix and prematurely capping it, an effect that would likely be minimized in the presence of TFE because of improved competition between solvent hydrogen bonding contacts and disruptive peptide hydrogen bond contacts. Further modeling suggested that incorporating a hydrogen bond donor at site 7 (^{wT}Leu₇), a hydrogen bond between Naq's side chain 1-position carbonyl could be effectively competed off and the helix stabilized further. Three mutant GN 7, Table 4.2, were synthesized with hydrogen bond donating residues at site 7 in the background of GN **6a**. GN **7a**, with a L7T mutation, and GN **7b**, with a L7N, mutation were highly disruptive of whatever helical content was observed in the absence of TFE, Figure 4.6H compared to Figure 4.6H. In the presence of TFE both mutants behaved identically to the previous generation, suggesting that instead of interacting with Naq's carbonyl, these two helix disrupting residues were breaking an insipient helix near Naq; perhaps also suggesting that there might be other, unknown, structural perturbations occurring in the protein. On the other hand GN **7c**, with a L7Q mutation did not behave appreciably different than the previous generation of peptides under either condition in CD suggesting at the very least that the mutation was not particularly disruptive to the overall structure.

Finally, *SCADS* also suggested that a functionally conservative mutation (D9E) of the Asp₉ residue involved in a critical salt bridge with Arg₁₆ might help stabilize the fold by

allowing the salt bridge to occur over longer distances that might be necessary with Naq incorporated at site 6. Mutating Asp₉ to Glu₉ in the background of GN **7c** afforded GN **8**, Table 4.2. Qualitatively, the CD spectra of **8** in the absence and presence of TFE look appreciably identical to that of **7c**, though perhaps **7c** is a bit more convincingly helical in the absence of TFE, Figure 4.6I. The nearly 1.5 fold increase in signal intensity of both the TFE-free and TFE containing sample cannot be explained by any structural consideration, and is likely due to the fact that the measured sample concentration was probably wrong. Regardless, the D9E mutation does not appear to significantly change the CD properties of the parent peptide **7c**.

Qualitatively, the two best peptides from this iterative design process are **6b** and **7c**, and their CD spectra are nearly identical in the absence and presence of TFE. Given the possible hydrogen bond contact between the Gln₇ residue of **7c**, this peptide was selected for further study using NMR.

4.4 NMR of ¹⁵NaqCage 7c

Without any tertiary structural data provided by CD directed mutagenesis, it was prudent to select the best candidate mutant from the iterative redesign cycle to determine, if possible, an NMR structure. This has two advantages. First, while gross changes in CD spectra were observed over the course of the iterative redesign cycle which tracked with an intuitive understanding of the anticipated effects of the mutations, CD gives no real tertiary structural information, and the stabilization of the helix afforded by the iterative redesign could be significantly disruptive to the fold as a whole, however unlikely. Second, if a structure is obtained by collection of relevant NMR data, this structure can guide both hand

and *SCADS* redesign efforts. This would aid the process immensely since it would allow *SCADS* to mutate a backbone with a structure based upon NMR obtained physical constraints as opposed to relying on the AMBER force-field minimization that was performed in generating the original ^{ox}NaqCage target backbone. The NMR determined backbone could be used for both forms of the amino acid and thus aid the overall progress of the project towards maximal E_m separation between ^{ox}NaqCage and ^{red}NaqCage.

A set of ¹H NMR data sufficient to determine the structure of **7c** was collected and processed at CCNY with Ronald Koder, Hsin Wang and Lei Zhang; a graduate student in Ronald Koder's laboratory. Data was collected on a 600 MHz Varian Inova spectrophotometer equipped with a cold probe. In particular, 2D TOCSY, 2D NOESY, natural abundance ¹³C-HSQC, and 3D NOESY-TOCSY data were collected at 2 °C. An 80 ms DIPSY spin lock was used for the TOCSY experiments, and NOESY data were collected at 100 ms, 150 ms, and 200 ms mixing times. Natural abundance ¹³C-HSQC data was collected with signal enhancement. The 3D TOCSY-NOESY was collected with an 80 ms spin lock and 100 ms NOE mixing time. The samples were at 1.2 mM peptide, in a 50 mM K₂PO₄ buffer, pH 7, containing 10% D₂O and 33% deuterated TFE. pH was uncorrected after addition to peptide. Solvent suppression was accomplished with a presat on water. Spectra were standardized to the residual TFE resonance at 3.919 ppm.

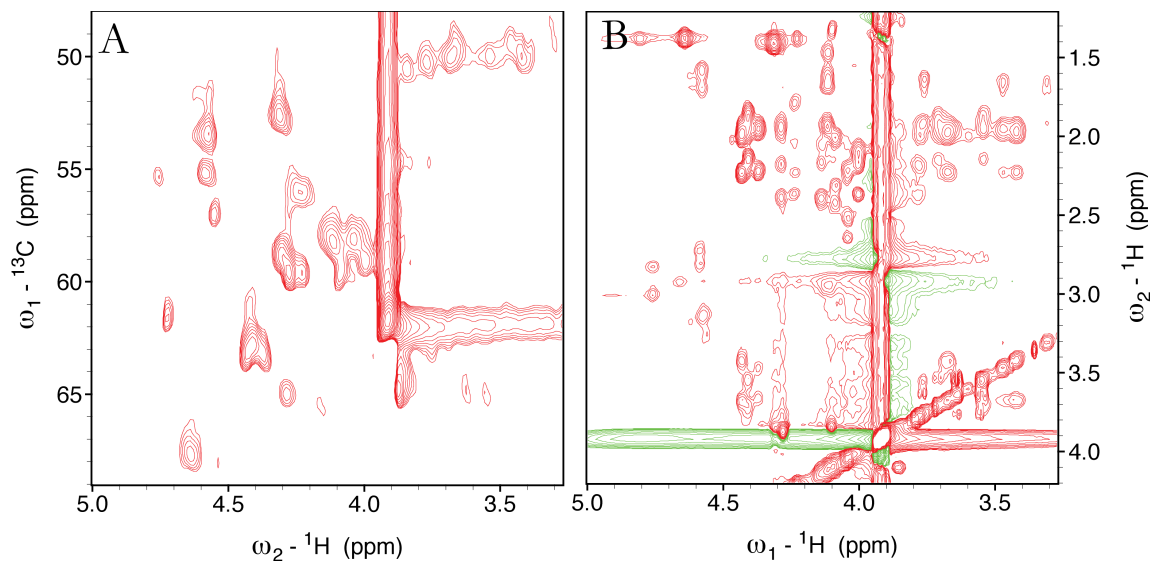


Figure 4.7. H^α region of natural abundance ^{13}C -HSQC (A) and ^1H TOCSY (B) spectra of $^\alpha\text{NaqCage 7c}$. It is possible to count nearly the correct number of $H^\alpha\text{-C}^\alpha$ correlations in the ^{13}C -HSQC. The TOCSY spectrum shows decent dispersion of the Glx residue spin systems, but suffers, like the ^{13}C -HSQC and NOESY (not shown), from a large streak corresponding to the residual protons on the deuterated TFE used in the sample. Positive peaks are in red, negative in green. See text for details.

On first glance the NMR data looked incredibly promising. The natural abundance ^{13}C -HSQC had 20 of the expected 23 $H^\alpha\text{-C}^\alpha$ correlations, Figure 4.7A. This chemical dispersion suggested that the various, and chemically identical amino acids were in different structural environments which would not be true if the protein was either unfolded or particularly dynamic. The TOCSY, which reveals covalently bonded spin systems through scalar (J) coupling, also appeared to be particularly clean given the number of repetitive residues (7 Glx, 4 Pro, 2 Ser, 2 Ala, 2 Gly). In particular, it was possible to count all seven Glx spin systems in the H^α correlated part of the spectrum, Figure 4.7B, and identify the spin systems of every residue except for one proline and one glycine. The aromatic spin systems of Tyr₁₈ and Naq₆ overlapped significantly but were clearly observed in the TOCSY. However, using the NOESY to walk the backbone proved to be quite limited. While the

missing proline's side chain became apparent in the NOESY, unfortunately, there is an unknown TOCSY spin system on top of the region of amide-amide correlations that would ordinarily be used for tracing the backbone of a helix. While some such correlations are observed, there are just few enough that the helix of the ^{15}N NaqCage cannot be unambiguously pieced together. For instance, Ala₂, which is surrounded by two stretches (two, and three residues) of Glx residues, would in the best case show identical helical amide proton-amide proton NOEs in both directions. This sequence homogeneity is complicated by the fact that there are a number of degenerate amide proton chemical shifts. The NOESY data, unfortunately, does not allow long runs of assignments, and thus it is somewhat difficult to identify long (sequence) range NOEs; none had been identified when assignments were stopped.

The 3D NOESY-TOCSY was acquired to address some of the degeneracy seen in the 2D NMR; although its value was somewhat limited due to the need to keep experiment time under 5 days, which is about the length of time it takes Naq_{ox} to get modified appreciably enough to see peaks growing in even at 2 °C. New peaks observed in the 3D NOESY-TOCSY merely confirmed what was known from the 2Ds. Finally, the need for 33% TFE removes a significant portion of all of the spectra from analysis and introduces streaks of NOE peaks that could either be due to NOE contact to TFE or the few H^α protons that lie close to its chemical shift. It is unclear whether the missing glycine resonances are coincident with the TFE or presat water peak. Use of a different solvent suppression pulse sequence may have significantly reduced the size of the residual TFE proton peaks. In particular, the WET pulse sequence uses shaped gradients to target specific

proton chemical shift regions of the spectra thus allowing multiple solvent peaks to be suppressed, but it was not used at the time of time of data collection.

All told, while 20 out of 23 spin systems could be clearly identified, only the backbone protons of 16 residues could be assigned, Table 4.3. It is somewhat telling that the majority of the residues without definite backbone assignments are the Glx residues of the helix, where assignment was slowed by the lack of unambiguous amide proton-amide proton NOEs. The remaining Gly and Ser resonances might be coincident with the TFE peak. It is possible, with continued effort, that the NMR data collected would be sufficient to complete the assignments. The well-dispersed spin systems are evidence enough for this, but without long distance NOE contacts the structure is not solvable. With the current level of assignments complete and without any obvious long range NOEs observable, it may not be possible to obtain a structure without recollection of data with better solvent suppression and higher resolution in the indirect dimensions.

Table 4.3. Backbone assignments. Sequence of the iterative redesigned ^{ox}NaqCage GN **7c**; amino acids in red indicate backbone assignments of both the amide proton, where applicable, and H^α can be made. **Q** is the one letter code for Naq. Sites changed during the iterative redesign are indicated with subscript site numbering. Lower case q is the residue dGln GN is generational number. GN **7c** was selected on the basis of CD spectral analysis as a NMR structure target. See text for details.

| Peptide | GN | NMR Backbone Assignments (in Red) |
|-----------------------|-----------|--|
| ^{ox} NaqCage | 7c | T₃P₂E₁Q₁A₂E₃E₄Q₆Q₇KD₉q₁₀GPA₁₃SGRPY₁₈PS |

4.5 Conclusion

There is sufficient evidence to suggest that the ^{15}N NaqCage GN **7c** mutant has the approximate arrangement of secondary structures found in the original TrpCage fold. However, it is difficult to address how compactly folded mutant **7c** actually is. The CD data suggests that in the presence of TFE, mutant **7c** becomes largely helical; at least significantly better than previous generations of the ^{15}N NaqCage. While the NMR confirms this observation showing significant, but overlapping, amide proton-amide proton NOE contacts signifying a helical conformation, there is little in the way of long sequence distance NOE contacts suggesting a compact fold. However, this may simply be due to the fact that the long distance contacts that are likely to occur are between side chains that have not been either found (like Ser₁₄) or ambiguously assigned (like the majority of the Glx residues in the helix). In fact, the 28 long range ($i \rightarrow i+n$, $n \geq 5$) contacts observed in the NMR of the optimized¹² ^{15}N TrpCage almost all occur between residues at equivalent positions that have either ambiguous side chain resonances (like Naq's aromatics) or are not assigned in the NMR of ^{15}N NaqCage. At the same time, large scale chemical shift deviations that Andersen's group observed in the TrpCages, which was attributed to the aromatic ring of Trp₆ shielding or deshielding nearby protons, were similarly observed to a lesser degree as far away as Pro₁₉ in the NMR of **7c**. These effects, indicating close proximity to the Naq_{ox} cofactor, suggest that the ^{15}N NaqCage **7c** may be adopting a compact fold, but without full assignment it is difficult to know whether the effects are truly due to a unique structure being adopted.

It remains to be seen whether Naq can be incorporated into a folded protein of 20 residues, and whether through differential design the midpoint potential of the cofactor can be driven high or lower at will. The *SCADS* output was an essential component to laying the

groundwork for the CD data driven redesigns. This should be clear since the only *SCADS* directed mutant site that was different in the final mutant **7c** was a functionally conservative D4E mutation that *SCADS* also picked up. The remaining mutations that progressed the design were outside of the functional sphere of *SCADS*; namely, adding the sequence TPE to the backbone, and use of a dGln residue in the 3₁₀-helix. These were perhaps the only real inputs the hand re-design process added to the *SCADS* design, but they were possibly as essential to obtaining a peptide that significantly folded in the presence of TFE. It is noteworthy that although the fraction folded was not determined, the CD spectrum of mutant **7c** in the absence and presence of TFE appears qualitatively more helical than the Trp₆ to 2-naphthalanine mutant under the same conditions that Barua and Andersen generated in studying the importance of the Trp₆ hydrogen bond.¹³ In effect, the design process to the ^{ox}NaqCage **7c** created a mutant protein that represents the first apparently successful replacement of the centrally buried tryptophan in a TrpCage-like protein. It is an open question whether **7c** or a subsequent mutant are properly structured.

On a more philosophical note, the project, as conceived, limits the design to only two redox states of Naq when a full six are hypothetically available at neutral pH. The electrostatic effects on the redox chemistry of neutral oxidized to neutral reduced are likely to be somewhat minimized compared to the same effects on negatively charged deprotonation states. In fact, the biggest effect of varying the electrostatic environment around Naq is likely to be seen in the relative stability of the single electron semiquinone state with little to no effect on the midpoint potential. This is perhaps why natural quinone redox sites fall into three families related to the availability of *protons* with little known changes to the midpoint of the neutral hydroquinone to neutral quinone redox couple. This

is one of the reasons why this project was initiated to begin with, natural systems may not have found a method to stably (in the evolutionary sense) take advantage of electrostatic effects on the midpoint potential of neutral cofactors, but protein designers are not limited to the survivability of a bug harboring their protein, at least not in most cases. Of course it is also clear that selective hydrogen bonding can effect changes in redox midpoints, perhaps evidenced by the observed change in midpoint of Naq going from random coil to helix, hence the focus of the initial designs and subsequent mutations (see for instance between GN 6 and GN 7) on finding good specific hydrogen bonding interactions in a structured protein environment. These effects can be substantial – typical hydrogen bond energies of a few kcal mol⁻¹ can produce a substantial 65 mV change in the midpoint of Naq, especially if the hydrogen bond is allowed to one redox state but not the other. Finally, just burying Naq into a well packed hydrophobic environment would represent not only a first for Naq, but also possibly restrict both water and proton access enough to change the kinetics of Naq's reduction and possibly its semiquinone stability through pKa effects. It is perhaps clear, then, why it might be too early to give up on getting a structure of a ^{ox}NaqCage and restarting the *SCADS* design process with backbone coordinates of a protein already incorporating Naq.

4.6 References

- (1) Tommos, C.; Skalicky, J. J.; Pilloud, D. L.; Wand, A. J.; Dutton, P. L. *Biochemistry-Us* **1999**, *38*, 9495.
- (2) Berry, B. W.; Elvekrog, M. M.; Tommos, C. *J Am Chem Soc* **2007**, *129*, 5308.
- (3) Zhu, Z. Y.; Gunner, M. R. *Biochemistry-Us* **2005**, *44*, 82.
- (4) Gunner, M. R.; Madeo, J.; Zhu, Z. *J Bioenerg Biomembr* **2008**, *40*, 509.
- (5) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1999**, *121*, 4952.
- (6) Koder, R. L.; Dutton, P. L. *Dalton Trans* **2006**, 3045.

- (7) Nanda, V.; Rosenblatt, M. M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P. L.; Saven, J. G.; Degrado, W. F. *J Am Chem Soc* **2005**, *127*, 5804.
- (8) Thyme, S. B.; Jarjour, J.; Takeuchi, R.; Havranek, J. J.; Ashworth, J.; Scharenberg, A. M.; Stoddard, B. L.; Baker, D. *Nature* **2009**, *461*, 1300.
- (9) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. *J Phys Chem B* **2006**, *110*, 3759.
- (10) Biswas, P.; Zou, J.; Saven, J. G. *J Chem Phys* **2005**, *123*, 154908.
- (11) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425.
- (12) Barua, B.; Lin, J. C.; Williams, V. D.; Kummler, P.; Neidigh, J. W.; Andersen, N. H. *Protein Eng Des Sel* **2008**, *21*, 171.
- (13) Barua, B.; Andersen, N. H. *Lett. Pept. Sci.* **2001**, *8*, 221.
- (14) Neuweiler, H.; Doose, S.; Sauer, M. *Proc Natl Acad Sci U S A* **2005**, *102*, 16650.
- (15) Williams, D. V.; Barua, B.; Andersen, N. H. *Org Biomol Chem* **2008**, *6*, 4287.
- (16) Dunbrack, R. L., Jr. *Curr Opin Struct Biol* **2002**, *12*, 431.
- (17) Starzyk, A.; Barber-Armstrong, W.; Sridharan, M.; Decatur, S. M. *Biochemistry-U S* **2005**, *44*, 369.
- (18) Harper, E. T.; Rose, G. D. *Biochemistry-U S* **1993**, *32*, 7605.

Chapter 5: Design and Properties of a Single-Chain Four-Helix

Bundle Maquette

5.1 Introduction

With perhaps the exception of a few of the soluble quinoproteins and the vitamin K dependent carboxylases, most proteins that use a quinone cofactor bind more than one redox cofactor. In some sense this confirms general concepts about the ability for natural quinone cofactors to do substrate level redox chemistry, which in the case of the respiratory and photosynthetic membrane complex *para*-quinone cofactors is limited to the occasional *reduction* of the quinone by a hydride source like tetrahydrobiopterin¹ or borohydride. Indeed, given their biological role as membrane soluble proton and electron transport cofactors, it is not surprising that the active sites of the membrane complexes have multiple different redox cofactors capable of processing electrons to and from the substrate quinones. This presents a glaring omission in the work described thus far with Naq; if the biological role of *para*-quinone cofactors is to predominantly act as an electron source or repository, then focusing on the subtle effects of helix-coil transitions on the midpoint potential of Naq (Chapter 3) or

on the possibility of differentiating between the reduced and oxidized state in a pair of designed scaffolds (Chapter 4) completely misses the broad need for better descriptions of active site chemistry and the natural engineering principles that guide quinone oxidoreductases. Indeed, while most of the work that has been presented focuses on setting Naq on a firm chemical, physical and design basis, the long term goal in the development of Naq has been the creation of maquettes which emulate natural function. This topic will be addressed specifically in the Future Directions section of Chapter 6. However, to jump from simple peptide systems, the longest of which was 31 residues (P2NaqExt, Chapter 3), to a protein that contains a single Naq residue and can bind multiple other natural cofactors requires a significant enhancement from the designs of almost all previously reported cofactor binding maquettes.

A maquette, in the biochemical sense, is a protein or peptide that functions to emulate natural biological cofactor binding proteins. The maquette is further differentiated in a strict sense from *de novo* designed cofactor binding proteins through the observation that all true maquettes have had each of their amino acids selected for a particular reason.² This differentiation is critical to the entire maquette design process, in that because amino acids are placed with clear intentions in each iterative round of redesign, interpretation of the overall success of the design process is much easier than when designing a protein wholesale. Each step and mutation can be judged according to its designed purpose. This pseudo-site specific (not entirely specific due to the symmetry of most of the designs) knowledge of amino acid function is perhaps best thought of as a mutational scanning of a natural protein in reverse; each step of the iterative process of redesign introduces more complexity leading to more native-like sequences to the point, in theory, where given the end result maquette

designed in such a manner it would be almost impossible for the roles of the individual amino acids to be determined by inspection. However, because of the process, the individuals who designed the maquette would know each reason why every amino acid was placed, and presumably the linear effects of each mutation. This process mirrors evolution and is always dependent upon the starting point. It must be made clear, however, that the loosest definition of maquette, which broadly incorporates both *de novo* designed cofactor binding proteins and those natural proteins whose function was extended through novel cofactor addition is perfectly legitimate and perhaps closer to the way the term is used in the Dutton laboratory.

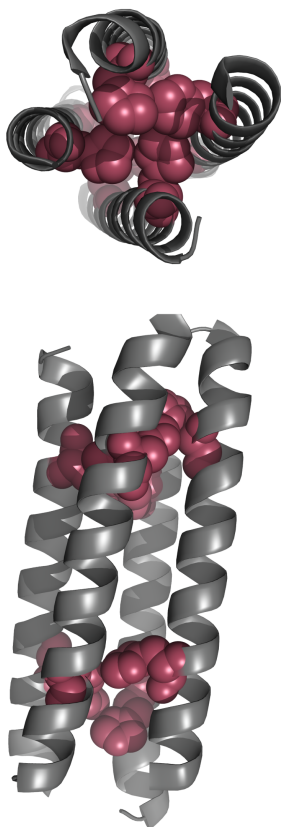


Figure 5.1. Crystal Structure of apo H10H24-L6I-L13F-L31M from end (top) and side (bottom), with heme-ligating histidines highlighted in red.

The earliest maquette was designed to emulate the function of the b_H and b_L bis-histidine heme binding sites in the respiratory bc_1 complex.³ Based upon earlier work by Regan and DeGrado,⁴ the peptide H10H24, named for the sites of heme-ligating histidines, Table 5.1, adopted into a four helix bundle tertiary structure characterized as a noncovalent homodimer of cystine-crosslinked monomer helices. Sequence design was quite minimal, based upon binary patterning of hydrophobic and hydrophilic residues around a heptad repeat with the heme-ligating histidines located along the hydrophobic helix interface. The original designs were highly successful in creating a multi-heme binding protein that was capable of differentiating the redox midpoints of the bound hemes. However, unlike natural proteins, these maquettes proved not

to have a singular structure in either apo or holo form.⁵⁻⁷ Iterative redesign produced a mutant protein, H10H24-L6I-L13F, Table 5.1, which increased the steric bulk of the interhelix volume and allowed both NMR⁶ and crystal structures⁷ to be determined of the apoprotein, Figure 5.1. The apoprotein structure revealed that the non-covalent monomers naturally aligned in *anti*-topology, Figure 5.2A. Subsequent redesign of the sequence focused upon producing a protein that was uniquely structured in the holo form. Several mutations were made which eliminated one of the heme binding sites on each helix, reduced the steric bulk near the remaining heme binding site, introduced non-ionic polar residues at sites which become exposed to solvent upon heme binding, and lastly, deleted a critical arginine which was responsible for heme differentiation in the original maquette design but was disrupting the helical register, HP1, Table 5.1.⁸ HP1 proved to have a singular structure upon heme binding and, while losing the midpoint potential differentiation between the two anti-cooperatively binding heme sites, HP1 maintained the binding affinity of heme.

Table 5.1. Sequences of maquettes discussed in text showing evolution of design. Proteins are separated by covalent architecture. – indicates a deletion of a residue. **Q** represents the position of Naq in a semisynthetic maquette (see text and Chapter 6). **C** is highlighted to show the site of covalent linkage between the two chemically ligated parts of BT3Naq. **H**istidine heme ligation sites are indicated in bold. Naming scheme for peptides follows laboratory convention; literature names may be different or non-existent.

| Maquette | Helix Number | <i>N</i> -loop or N-cap | Helix Sequence | <i>C</i> -loop or Interhelix Loop |
|----------------------------------|--------------|-------------------------|--|-----------------------------------|
| <i>Crosslinked Homotetramers</i> | | | | |
| H10H24 | 1 | Ac-CGGG | ELWKL HE ELLKKFEELLKL HE ERLKKL-NH ₂ | |
| H10H24-L61I-L13F | 1 | Ac-CGGG | EIWKL HE EFLKKFEELLKL HE ERLKKL-NH ₂ | |
| H10H24-L61I-L13F-L31M | 1 | CGGG | EIWKL HE EFLKKFEELLKL HE ERLKKM-NH ₂ | |
| HP1 | 1 | CGGG | EIWK QHE EALKKFEEAL QF EE-LKKL-NH ₂ | |
| BT2 | 1 | TPE | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGC-NH ₂ |
| <i>Crosslinked Homodimers</i> | | | | |
| HP7 | 1 | G | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGCGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| HP8 | 1 | GSP | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGCGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| <i>Single Chain</i> | | | | |
| BT3 | 1 | GMTPE | QIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGSGG |
| | 3 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 4 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| cBT3 | 1 | GMTPE | QIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 2 | | EICK CHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGSGG |
| | 3 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 4 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| BT5 | 1 | GMTPE | QIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 3 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 4 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| BT6 | 1 | G | QIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 3 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 4 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | |
| BT3Naq | 1 | MTPE | QIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGG |
| | 2 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGSGSGSGG |
| | 3 | | EIWK QHE DALQKFEDAL NQ FEE-LKQL | GGSGCGSGG |
| | 4 | | EIWK QHE DAL Q KFEALQFEED-LKQL-NH ₂ * | |

*Contains additional unintentional mutations besides Naq

HP1 was the first of a series of proteins designed to increase the likelihood of obtaining a solution structure of a heme-bound maquette.² While eliminating the sequence homogeneity by introducing diversified surface residues was critical for NMR analysis, the most significant advance was the conversion of the *anti*-oriented homodimer of cystine-crosslinked helical monomers structure of HP1, Figure 5.2A, into the *syn*-oriented cystine-crosslinked homodimer of helix-loop-helix monomers structure of HP7, Figure 5.2B. As described by Gibney et al. with the H10H24 protein,⁹ this conversion only required the elimination of the

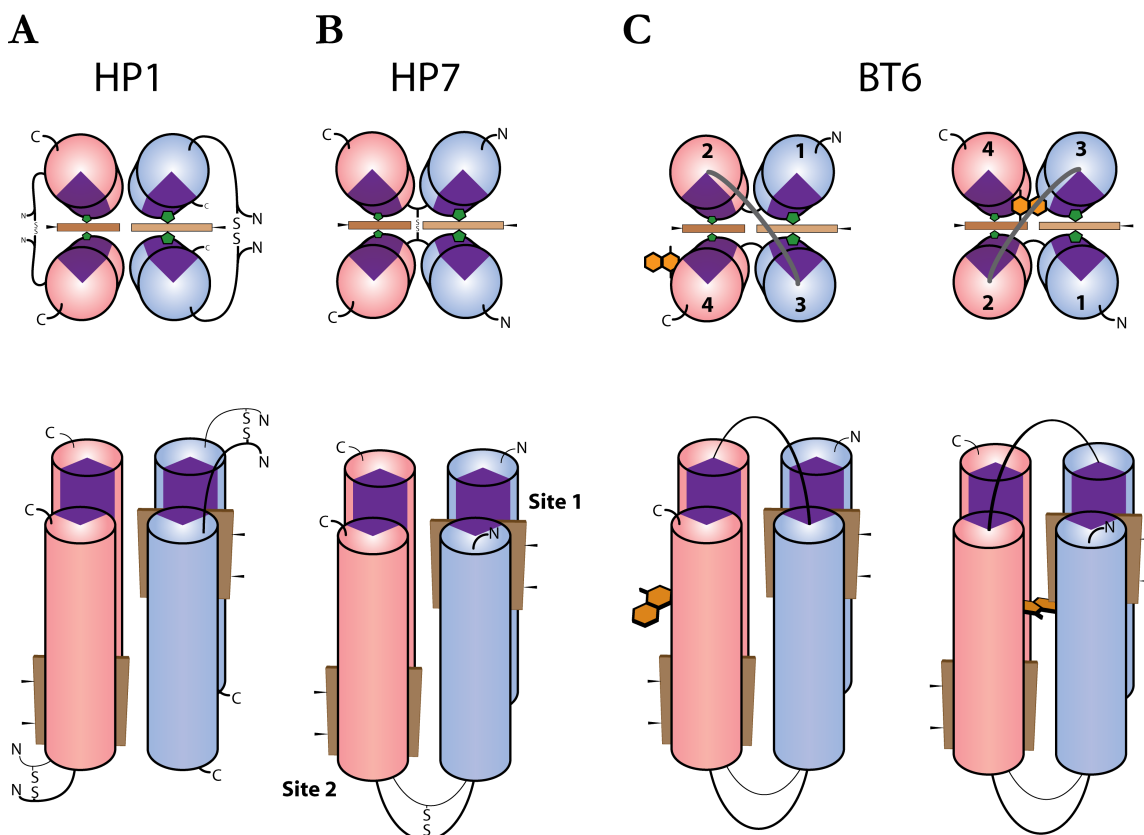


Figure 5.2. Protein topologies of multichain heme-binding maquettes HP1 (A) and HP7 (B), and single chain heme-binding maquette BT6 with hypothetical Naq residue to demonstrate the effect of adding a desymmetrizing cross-bundle loop (C), see text for details. Helices in vary in color by whether they have their N-termini (blue) or C-termini (pink) pointed towards the viewer (top) or up (bottom). Loop rearrangement was the most significant change in topology between HP1 and HP7 as orientation of helices and heme binding site sequence did not vary significantly. Similarly between HP7 and BT6, the most major topology changes involve the construction of a cross-bundle loop between helices 2 and 3 and elimination of the disulfide bridge found between the loops of HP7.

CGGG sequence from the N-terminus of the helical monomers of HP1 and the addition of two loops (sequence GGSGCGSGG) between the *anti*-oriented helices, Table 5.1. The loops, which incorporated central cysteine residues, allowed the creation of a covalent ‘candelabra’ geometry via disulfide formation, though this proved unnecessary for maintenance HP7’s properties.² Once again, HP7 demonstrated nearly identical properties to HP1 while allowing full NMR backbone assignments to be completed.¹⁰ Unfortunately, due to the redundancy of the protein sequence and, hence, overlap of sidechain chemical shifts, NOE-based distance measurements needed to define the structure were impossible to obtain (Ron Koder, personal communication). Holo-HP7 proved to have very similar properties to holo-HP1 with the added feature of being able to form a transient, but observable, oxyferrous state, intriguingly having a slightly higher ferrous affinity for O₂ than CO.² The determined midpoint of HP7 was approximately -300 mV (Ronald Koder, personal communication).

5.2 Design of BT3

The work described in this section is the result of collaboration between Tammer Farid and the author.

HP7 advanced the groundwork set by the uniquely structured holo-HP1. It not only represented an advance in the topological construction of the primary sequence, but, in addition, represented the first four-helix bundle maquette in the Dutton lab to be solely produced by expression given the length of the individual monomers (though more recent maquettes of similar length have been synthesized). However, for all of its benefits, HP7 cannot be used directly for incorporating a single site mutation or covalent modification due

to the homodimeric nature of its structure, and this certainly precludes, without resorting to heterodimeric disulfide bridged helix-loop-helix monomers, incorporation of a single Naq residue. Rather, if possible, it would be much preferable to have a single chain variant of HP7. In work undertaken simultaneously as that presented here, Bender et al. describe the creation of a single chain dual iron(III) porphyrin binding *de novo* designed four-helix bundle protein with a significantly different cofactor binding architecture, though nearly identical topology, than HP7.¹¹

5.2.1 N-Cap design

In the NMR of HP7, it was clear that the N-termini of the monomers were not structured. The same sequence on the N-terminal end of the second helix in the helix-loop-helix monomer was however structured (Ron Koder, personal communication). To further advance solution structure determination by clarifying the NOE constraints of the N-terminal residues, Ron Koder added a non-ideal Ser-Pro sequence to the beginning of the HP7 sequence producing HP8 (Ron Koder, personal communication), Table 5.1. This had the undesired effect of eliminating one of the heme binding sites (data not shown, collected by T. Farid). To test the more ideal N-capping motif TPEQ,¹² BT2 was synthesized as a single helix of HP7 with this N-capping motif in place, Table 5.1, and a C-terminal loop ending at a Cys for covalent cross linking. Presumably forming dimers of cystine-crosslinked dimers BT2 demonstrated full two-heme binding, suggesting the ideal capping sequence TPEQ would not interfere with heme binding in a single chain HP7 variant. Since it was observed that the glycine-serine loops in HP7 were sufficient to cap and stabilize the N-

terminal end of the second helix in the monomer, only one N-terminal TPEQ sequence was needed in the single chain variant.

5.2.2 Cross Bundle Loop Design

The nine residue loops of HP7 are much longer than would be expected to link two adjacent helices.¹³⁻¹⁵ However, in the development of HP7 it became clear that shorter loops increased the amount of higher order oligomeric states of heme bound maquettes as observed by size-exclusion chromatography (Ron Koder, personal communication), which confirmed an earlier report that suggested shorter loops would be disruptive to the four-helix bundle fold,¹⁵ although in that case the variation observed was in loops 3-5 residues long. The need for loops of nine residues belies the fact that the N-terminal and C-terminal ends of the Dutton Laboratory maquettes are still very much identical to the original design and have not been optimized for linking via loops in this manner, especially considering the need to not interfere with cofactor binding. Indeed, the glycine-serine loops in HP7 were designed to be as minimally perturbing as possible so as not to develop a set structure that might interfere with assembly or cofactor binding. Since there was no need to keep the cystine-crosslinks in the single chain HP7 variant, the central cysteine in the found in the HP7 loops was replaced by a homologous serine residue that had been previously found not to globally change the properties of the heme bound HP7C38S.² Additionally, it was envisioned that the cross-bundle loop (from helices 2 to 3), Figure 5.2C, would need to be longer than the loops found in HP7 so the cross-bundle loop length was increased to eleven residues by the addition of two glycines around the central serine in the HP7C38S loop, Table 5.1. This design process contrasts with that used by Bender et al. in the *de novo* redesign

of a dual iron(III) porphyrin binding single chain four-helix bundle.¹¹ Using a database of natural interhelical loop structures, the DeGrado group were able to find a loop that best matched (r.m.s.d.) the positioning of at least five C α backbone atoms at the loop-helix interface of a computationally optimized backbone. This reveals an advantage that *de novo* design methods has over the maquette iterative design procedure in that computational approaches do not depend upon the determination of a protein structure when a backbone target is produced during the design process, thus making it easier to systematically design loops and other structures to stabilize a goal structure.

5.2.3 Addition of an N-Terminal Methionine

To allow direct expression of the new single chain maquette in *E. coli* for use in the intein work detailed in Chapter 6, an N-terminal methionine was appended, giving the final sequence of the single chain HP7 variant, BT3, Table 5.1.

5.2.3 Breaking the Symmetry of HP7

Addition of the cross-bundle loop between helix 2 and helix 3, Figure 5.2C, generates the possibility of two topologically distinct, but nearly structurally identical BT3 conformations. The same conformational rearrangement in HP7 is possible, but it produces a protein that has an indistinguishable topology. This level of symmetry is broken by the addition of the cross-bundle loop. It is anticipated that at least with early BTX designs, this new asymmetry will have limited effects on heme cofactor chemistry since no difference in properties was observed between the non-identical sites in HP7 with histidine knockout

mutants. However, it should be clear from Figure 5.2C that positioning an amino acid like Naq that introduces further asymmetry into BT3 can cause some differentiation between the symmetric pair in a manner that may complicate analysis. To avoid significantly different behavior from the two topological arrangements of helices in such pseudo-symmetric bundles, mutations are effectively limited to pseudo-symmetric positions that are not in different environments with heme bound. Essentially, until one topology can be forced to be favored this requires placement of single site mutations at positions that should remain solvent exposed regardless of topology. The caveat to this design consideration is that mutations that are designed to break the remaining symmetry and effect differentiation of the heme sites will more than likely be found at positions that vary significantly in their environments between the two topologies.

5.3 Confirmation of General Properties of Heme Binding Maquette BT3

The work presented in this section and subsequent sections is based upon data collected by Ross Anderson and Tammer Farid. The author was actively involved in the design process of the maquettes described below, which was mirrored in the work presented in Chapter 6; however, for the sake of brevity, this work will be discussed in the context of the BT3 maquette because the advances of BT3 are not exclusively in the domain of Naq incorporation. While incorporation of Naq into a single site in a four-helix bundle maquette was the original motivation for developing the BT3 protein, it was also evident that the natural advantage of being able to introduce mutations truly site-specifically in a single chain protein would greatly enhance the structural, cofactor and environmental control possible in the maquettes. It is believed that the long term benefits of the development of the BT family

of single chain four helix bundles will be in the diversity that this structural change allows, including the incorporation of one or more Naqs.

All BT proteins were expressed in *E. coli* BL21(DE3)-RIL as a C-terminal fusion to GST. Purification on a nickel column, followed by cleavage from the fusion via a TEV cut site afforded BT3 with an additional N-terminal glycine residue.

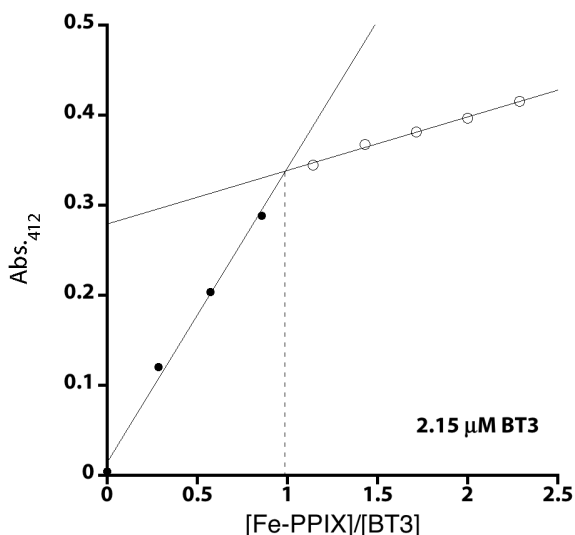


Figure 5.3. Iron(III) protoporphyrin IX (PPIX) endpoint titrations of BT3 at 1.88 μM showing unanticipated one-heme binding in a protein designed with two binding sites. Lines through data and drop line from equivalency point are meant to guide the eye. The titration was followed at 412 nm, the peak position of the bound heme Soret. Filled dots represent titrations in which heme binding was observed; open dots represent those in which heme aliquots contributed to free heme spectra. Equivalents bound is calculated at 0.98 hemes/bundle. Data collected by Tammer Farid.

5.3.1 Heme Binding

BT3 was designed as a single chain variant of HP7, which unambiguously binds two hemes.¹⁰ End point heme titrations at high and low concentrations of BT3 universally showed that BT3 binds only one heme; thus, failing to match the design criteria, Figure 5.3 (Tammer Farid, personal communication). It was apparent that either the cross-bundle loop or the N-cap was interfering with heme binding since no other major structural changes were incorporated in the step from HP7 to BT3.

While this result was surprising, the iterative design process limits the number of factors that needed to be tested to identify the root causes of the property change, something that would be difficult to do in a protein designed from scratch.

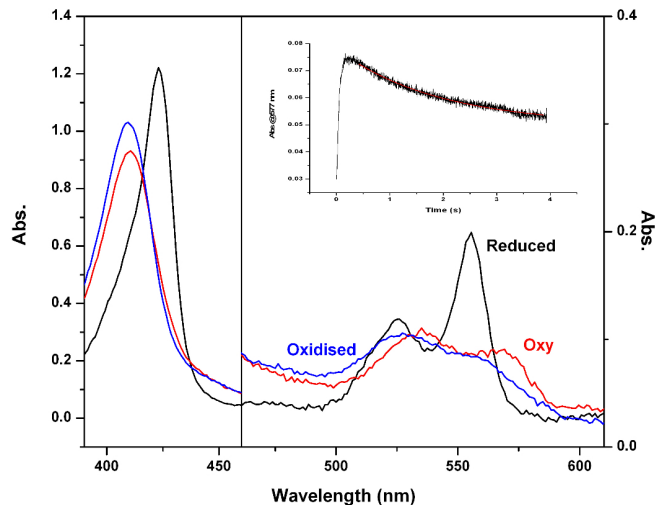


Figure 5.4. Formation of oxy-ferrous state in hemin bound BT3 in a stopped-flow experiment. Spectral panels show Soret (left) and alpha-beta bands (right) of intermediates determined by SVD analysis of the full spectral kinetic traces. Oxidized (blue), reduced (black) and oxy-ferrous (red) states are distinctly visible. Data collected by Ross Anderson.

5.3.2 Molecular Oxygen Binding

Perhaps not surprisingly given that one site of BT3 binds heme and both sites of HP7 independently (and jointly) bind molecular oxygen, ferrous heme holo BT3 binds molecular oxygen with similar, or perhaps slightly better, kinetics as HP7, Figure 5.4.

This suggests that the molecular rate constants involved in this unique property are appreciably maintained even though the protein does not have two functional heme binding sites.

5.3.3 Redesign

The observation that either the loop or the N-cap, or both, were interfering with heme binding at one of the sites was confirmed in an early chemical ligation maquette based upon BT3 that incorporated Naq(OMe/OMe) in the fourth helix, BT3Naq(OMe/OMe), Table 5.1. Incomplete histidine knockout mutagenesis studies appeared to indicate that the N-terminal heme binding site (Site 1) was not binding heme. This weakly suggested that the N-cap, which had not interfered with heme binding in BT2 was the cause of the loss of two-heme binding, however, the histidine knockout studies could not completely exclude the loop from potential interference. Mutagenesis of the N-cap to the simpler and less bulky

Asp, BT4, Table 5.1, produced no further changes in heme binding behavior. Further mutagenesis of the Asp residue to a Gly was underway when two laboratories working on related proteins revealed that a cross-bundle loop length of nine residues identical to the loops found in HP7 was sufficient to maintain heme-binding behavior (Dror Noy and Ron Koder, personal communications). This was quickly confirmed, Figure 5.5, with a new mutant BT6, Table 5.1, which also completely eliminated the N-cap (though it proved unnecessary, mutant BT5, Table 5.1, which maintained the TPEQ N-cap and had shorter loops bound two hemes, Tammer Farid, personal communication).

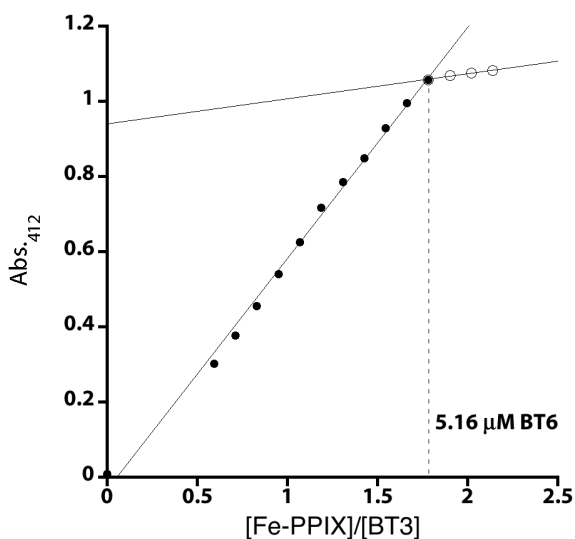


Figure 5.5. ‘Two-heme’ binding restored in the BT6 single chain heme-binding maquette. Iron(III) protoporphyrin IX (PPIX) endpoint titrations of BT6 at 5.16 μM showing a closer to ideal 1.8 equivalents heme bound per bundle. Lines through data and drop line from equivalency point are meant to guide the eye. The titration was followed at 412 nm, the peak position of the bound heme Soret. Filled dots represent titrations in which heme binding was observed; open dots represent those in which heme aliquots contributed to free heme spectra. See text for details and discussion of deviation from optimum 2 equivalents. Data collected by Tammer Farid.

Surprisingly, the exceeding long cross-bundle loop somehow interfered with two-heme binding. It is possible that a loop of such length allows the protein to adopt a lower energy heme binding topology that excludes a second heme-binding event. While the loop is certainly long enough (extended length approximately 37 Å) to span most of the entire helix (~43 Å), it is unclear if this is the cause of the observed properties. As will be discussed below, there is evidence that by forcing BT3 into proper orientation two heme-binding can be restored. More sophisticated loop design will need to be addressed at some point, but

the current generation of maquettes will likely continue to function without intensive loop redesign. It should be noted that although the maquettes have shown unambiguous two-heme binding by NMR,¹⁰ the stoichiometry obtained during a titration is invariably lower. Some amount of this may be due to overestimation of the concentration of the protein or under estimation of the heme stock (determined with pyridine hemochrome¹⁶), but these corrections are not perfectly in line with the data for proteins that only bind one heme. In those cases, the amount of heme bound is nearly exactly one. What this might be suggesting is that in at least all of the recent maquettes the first heme binding mostly occurs at one site with a minority binding in a topology or higher order protein aggregate that prevents a second heme binding. This minority binding site would have to be of a low enough fraction that it is not apparent in NMR studies. It could be that the failure of BT3 to bind two hemes was just an extreme example of this sort of behavior. Unfortunately, without a means to detect such minority binding events, it may not be possible to fully describe this slight deviation from integer binding in the second heme site. It is likely that NMR on heme binding in BT3 would be revelatory.

5.4 cBT3

The work described in this section is almost entirely that of Ross Anderson. While the author was engaged in the protein design process and some of the experimental designs, it is included to further emphasize the long-term benefits of the single chain BT3 family of proteins.

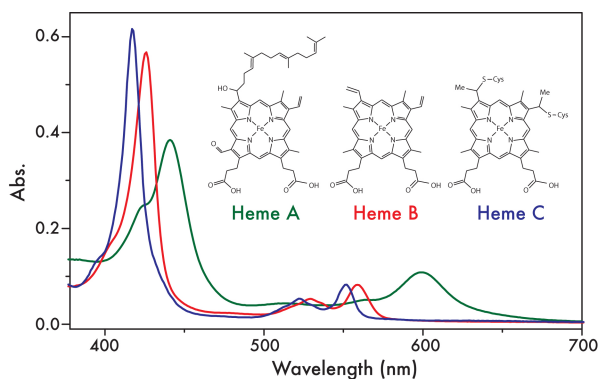


Figure 5.6. Spectra of ferrous Heme A (green), Heme B (red) and Heme C (blue) bound to protein maquettes. With the advent of the BTX family of single chain heme-binding maquettes a *c*-type heme can be incorporated into the maquette family with single site specificity. Data collected and figure generated by Ross Anderson.

5.4.1 In vivo *c*-Type Maturation

The heme and iron(III) porphyrin binding maquettes developed thus far have been mostly focused on the non-covalent attachment of the cofactor ligand in the holo proteins. A notable exception is the work by Ishida et al. in which a helix-loop-helix version of H10H24 was converted from a *b*-heme binding maquette into a covalent *c*-heme binding maquette using the conserved *c*-type binding motif CXXCH.¹⁷ Reaction with iron(III) protoporphyrin IX *in vitro* afforded an undisclosed amount of a singly thioether linked *c*-type monomer creating a dimeric double *c*-type protein.¹⁷ Under similar conditions, cBT3, Table 5.1, with a CXXCH motif at the second helix's heme binding site (Site 2), *c*-type incorporation was minimal (Ross Anderson, personal communication). However, expressing cBT3 with an periplasmic export tag in BL21(DE3) constitutively expressing a *c*-type maturation operon on the pEC86 vector generates a majority of the desired dithioether modified *c*-type, Figure 5.6. Such *in vivo* incorporation of a cofactor at a single site in BT3 represents a critical advance in the maquette family design.

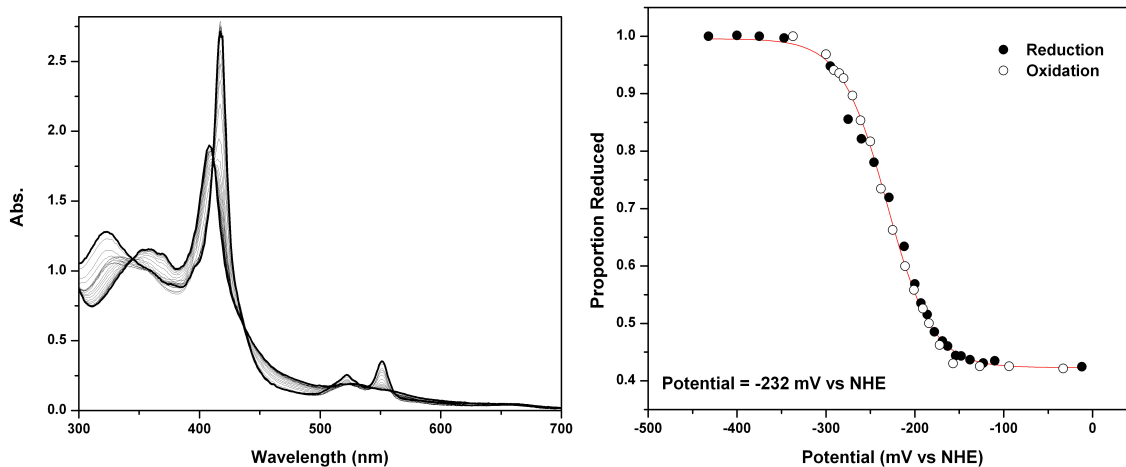


Figure 5.7. Potentiometry of cBT3. Spectra acquired during potentiometry of cBT3 (left) showing classing *c*-type heme responses and reversible one-electron Nernst redox behavior (right) with an E_m of -232 mV, slightly more positive than *b*-hemes in similar constructs. Data collected and figures generated by Ross Anderson.

5.4.2 Potentiometry and Molecular Oxygen Binding

Potentiometric titration of cBT3 was typical, Figure 5.7, with an expected slightly more positive midpoint potential compared to the *b*-type maquette, E_m -232 mV at pH 9.0. Like ferrous BT3 and HP7, fully matured ferrous cBT3 also binds molecular oxygen, Figure 5.8. Intriguingly, when titrated with iron(III) protoporphyrin IX, the *c*-type incorporated cBT3 bound an additional one equivalent of heme, suggesting that the binding of a *c*-type heme covalently at the more C-terminal heme site (Site 2) returns typical two-heme binding properties (Ross Anderson, personal communication). Since the *c*-type is on the second helix which ideally bis-histidine ligates an iron(III) porphyrin as a partner with the forth helix, it is possible that *c*-type heme binding serves to eliminate the flexibility imparted upon the bundle by the long cross-bundle loop by preventing the forth helix from adopting a topology that allows it to improperly histidine ligate an iron(III) porphyrin co-ligated with the first helix.

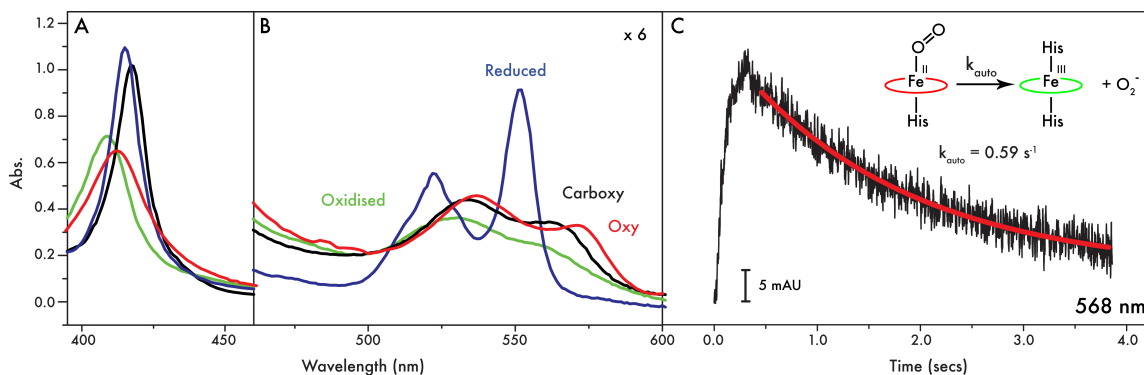


Figure 5.8. Formation of oxy-ferrous cBT3 in a stopped-flow experiment. Soret peak (A) and alpha-beta band (B) absorbance spectra of four forms of cBT3. Reduced (blue), oxidized (green) and oxy (red) were determined by SVD analysis on full spectral kinetic data from the experiment and match literature spectra for the species. Carboxy (black) was determined in a separate experiment. Kinetic trace (C) shows auto-oxidation of oxy-ferrous heme. Data collected and figure generated by Ross Anderson.

5.5 Conclusion

The development of the two-heme binding single chain maquettes BT5, BT6 and cBT3 from the homodimeric HP7 represent a new range of opportunities that have thus far been absent from the heme maquettes. cBT3 is the first direct example of this development showing that *in vivo* incorporation of a redox cofactor in a site specific manner can be fully realized. This also completes one aspect of the long story of natural heme cofactor binding in the maquettes as all three of the most common hemes, heme *a*, heme *b* and heme *c*, have now been shown to bind in a single protein, Figure 5.6. It is now almost trivial to generate a maquette that binds any two of these natural heme cofactors, a feat that was not possible before the development of the single chain maquette family. On a more practical note, the preservation of the critical properties observed in the previous generation maquette HP7 means that differentiation and selection of a subset of properties can be held off for at least another generation. The advancement to a single chain maquette also makes it more practical to introduce mutations that would ease solution structure determination by eliminating the

hurdle of sequence homogeneity. Finally, the BT series of proteins represent the first time that advances in the generation of semisynthetic proteins can be applied to the maquettes, a distinction that allows the incorporation of Naq into designed protein scaffolds capable of emulating natural redox electron transport proteins.

5.6 References

- (1) Schallreuter, K. U.; Rokos, H.; Chavan, B.; Gillbro, J. M.; Cemeli, E.; Zothner, C.; Anderson, D.; Wood, J. M. *Free Radical Biology and Medicine* **2008**, *44*, 538.
- (2) Koder, R. L.; Anderson, J. L.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.
- (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) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976.
- (5) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1997**, *119*, 2323.
- (6) Skalicky, J. J.; Gibney, B. R.; Rabanal, F.; Urbauer, R. J. B.; Dutton, P. L.; Wand, A. J. *J. Am. Chem. Soc.* **1999**, *121*, 4941.
- (7) Huang, S. S.; Gibney, B. R.; Stayrook, S. E.; Dutton, P. L.; Lewis, M. *J Mol Biol* **2003**, *326*, 1219.
- (8) Huang, S. S.; Koder, R. L.; Lewis, M.; Wand, A. J.; Dutton, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5536.
- (9) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry-U.S.* **1998**, *37*, 4635.
- (10) 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.
- (11) Bender, G. M.; Lehmann, A.; Zou, H.; Cheng, H.; Fry, H. C.; Engel, D.; Therien, M. J.; Blasie, J. K.; Roder, H.; Saven, J. G.; DeGrado, W. F. *J Am Chem Soc* **2007**, *129*, 10732.
- (12) Harper, E. T.; Rose, G. D. *Biochemistry-U.S.* **1993**, *32*, 7605.
- (13) Engel, D. E.; DeGrado, W. F. *Proteins* **2005**, *61*, 325.
- (14) Brunet, A. P.; Huang, E. S.; Huffine, M. E.; Loeb, J. E.; Weltman, R. J.; Hecht, M. H. *Nature* **1993**, *364*, 355.
- (15) Predki, P. F.; Regan, L. *Biochemistry-U.S.* **1995**, *34*, 9834.
- (16) Berry, E. A.; Trumpower, B. L. *Analytical Biochemistry* **1987**, *161*, 1.
- (17) Ishida, M.; Dohmae, N.; Shiro, Y.; Oku, T.; Iizuka, T.; Isogai, Y. *Biochemistry-U.S.* **2004**, *43*, 9823.

Chapter 6: Incorporation and Deprotection of Naq in a Four-Helix Bundle Maquette

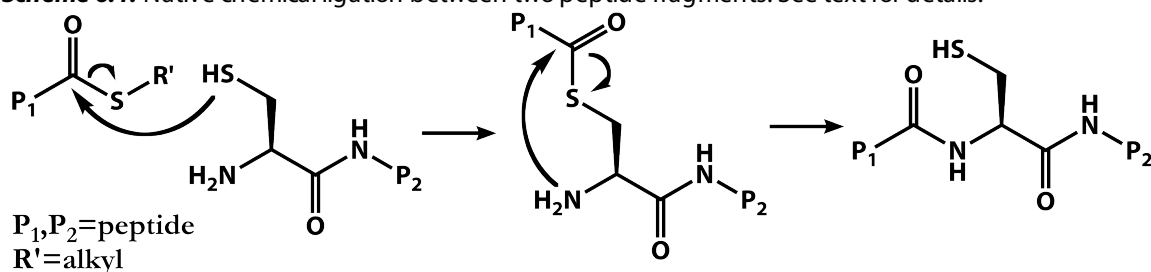
6.1 Introduction

The single chain BTX heme-binding four-helix bundle maquette family (Chapter 5) was conceived as a way to produce four-helix bundle maquettes that permitted single site-specific mutations free from the pseudo-symmetry of all previous generations of maquettes. As was seen in Chapter 5, this allowed the single-site incorporation of a α -type heme in the progenitor single chain maquette BT3. Yet the advantages of the BTX maquettes are not limited to single-site mutagenesis. Indeed, by taking advantage of expressed protein ligation techniques,¹ a single chain BTX mutant can be created with synthetic fragment containing any number of functional moieties not available *in vivo*.

Expressed protein ligation (EPL)¹ built upon the already established native chemical ligation (NCL), in which a synthetic fragment having a C-terminal thioester and a peptide with an N-terminal cysteine are initially coupled via thioester exchange, Scheme 6.1.² A subsequent rearrangement of the crosslink affords a native peptide bond between the two

synthetic fragments. There is a similar system in nature, wherein a protein domain efficiently excises itself post-translationally, leaving intact the functions of the remainder of the protein via cross coupling the N-terminal and C-terminal fragments at the site of excision. In recognizing that these so-called intein domains catalyze their reactions through a thioester intermediate, using a commercially available intein based protein purification vector Muir et al. were able to capture the thioester of an expressed fragment of Csk. Through native chemical ligation they then attached a synthetic fragment containing a phosphorylated tyrosine residue thought to be important for the self-inhibition of the related kinase Src, but absent from wild type Csk, thus generating a full length phosphorylated-Src analog that lacked self-inhibition.¹ Like native chemical ligation, expressed protein ligation results in the incorporation of a cysteine residue at the site of cross-linking. While a full review of the use of native chemical ligation and expressed protein ligation is beyond the scope of this chapter, these techniques have been used to create a sizable number of previously impossible or very difficult proteins from the total synthesis of all D-amino acid proteins,³⁻⁵ to construction a three segment piecewise MBP expression system that allowed selective ¹⁵N labeling of each segment.⁶ The biology of inteins themselves is fascinating – the intein domain is an expressed part of a parasitic genetic element and in most cases serves as a homing endonuclease for propagation of the gene that encodes it. Because the intein splices out of

Scheme 6.1. Native chemical ligation between two peptide fragments. See text for details.



the gene product that it has infected, it minimally perturbs the normal biological function of its host.⁷ Regardless, it is clear that by using expressed protein ligation of the single chain BTX maquette family with a synthetic fragment, it should be possible to incorporate Naq into a semi-synthetic four-helix bundle maquette.

BT3

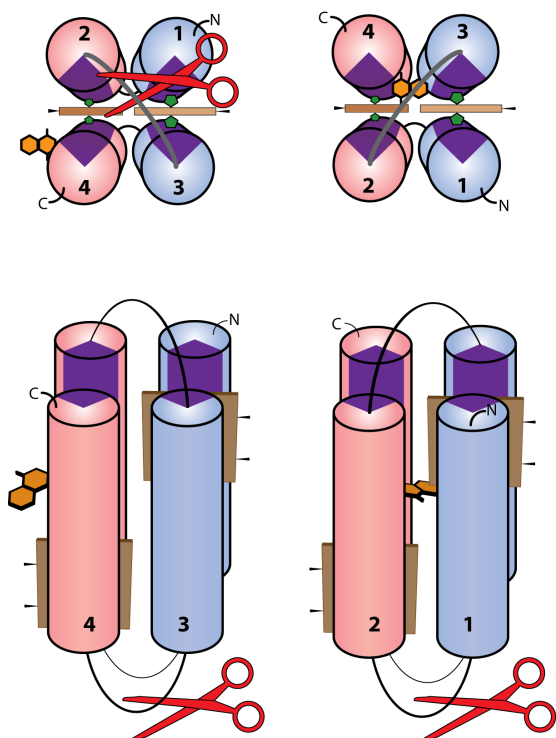


Figure 6.1. Diagram of possible intein joint sites in BT3.

6.2 Building a Maquette Competent

Expressed Protein Ligation Vector

In some ways the single chain BTX maquettes are the ideal starting point for use in expressed protein ligation (EPL) system since they have both long loops between the helices and those loops have virtually no functional requirements in terms of amino acid content. This makes the fact that the expressed protein ligation and native chemical ligation methodologies leave a cysteine at the site of the coupling of minimal concern. There are then three

convenient sites to covalently link synthetic and expressed fragments at each of the three loops in the single chain BT3 maquette, Table 6.1 and Figure 6.1. To minimize the number of fragments needing to be coupled, although there are methods that allow several fragments to be coupled,³⁻⁶ only two of these joint sites are synthetically practical since coupling at the loop between helix 2 and 3 would require chemical synthesis of a 70+ residue fragment.

Both of the two remaining sites of attachment are similar but not identical in the details of the fragments needed to perform EPL. Chemical ligation between a synthetic helix 1 and expressed helices 3-4 requires a synthetic C-terminal thioester of helix 1 and an expressed N-terminal cysteine fragment of helices 3-4. Expression of an N-terminal cysteine fragment is made possible by the fact that the cleavage of the thioester intermediate generated by an intein produces a fragment with an N-terminal cysteine in the absence of cross coupling. On the other hand, chemical ligation of expressed helices 1-3 and synthetic helix 4 requires an expressed thioester fragment and a synthetic helix with an N-terminal cysteine. The second method was arbitrarily chosen, though it is likely in retrospect that the first method is slightly less work since a number of purification steps would be eliminated and excess N-terminal cysteine containing fragments would be easier to produce. The site of chemical ligation in the loop between helices 3 and 4 was the central serine, which in the dimeric HP7 was a cysteine suggesting that the requisite mutation of this residue to cysteine for EPL would likely be minimally perturbing. The cysteine could be blocked with bromoacetamide, or otherwise modified to, for instance, covalently incorporate another cofactor.

Table 6.1. Sequences of maquettes discussed in text. **Q** represents the position of Naq in a semisynthetic maquette. **C** is highlighted to show the cysteine site required for and preserved during chemical ligation.

| Maquette | Helix Number | <i>Nloop</i> or <i>Ncap</i> | Helix Sequence | <i>Cloop</i> or <i>Interhelix Loop</i> |
|--------------------------------------|--------------|-----------------------------|--|--|
| <i>N-terminal Cysteine Fragments</i> | | | | |
| BT0.1 | 1 | C GSGG | EIWKQHEDALQKFEDALNQFEELKQL-NH ₂ | |
| BT0.2 | 1 | C GSGG | EIWKQHEDAL Q KFEALQFEEDLKQL-NH ₂ * | |
| BT0.3 | 1 | C GSGG | EIWKQHEDAL Q KFEDALNQFEELKQL-NH ₂ | |
| BT0.4 | 1 | C GSGG | EIWKQHEDALQKFE Q ALNQFEELKQL-NH ₂ | |
| <i>Single Chain</i> | | | | |
| BT3 | 1 | GMTPE | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 4 | | EIWKQHEDALQKFEDALNQFEELKQL | |
| BT5 | 1 | GMTPE | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 4 | | EIWKQHEDALQKFEDALNQFEELKQL | |
| BT6 | 1 | G | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 4 | | EIWKQHEDALQKFEDALNQFEELKQL | |
| BT3.1 | 1 | MTPE | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSG C GSGG |
| | 4 | | EIWKQHEDALQKFEDALNQFEELKQL-NH ₂ | |
| BT3Naq | 1 | MTPE | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSG C GSGG |
| | 4 | | EIWKQHEDAL Q KFEALQFEEDLKQL-NH ₂ * | |
| BT5.3 | 1 | MTPE | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSG C GSGG |
| | 4 | | EIWKQHEDAL Q KFEDALNQFEELKQL-NH ₂ | |
| BT6.4 | 1 | MGG | QIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 2 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSGSGSGG |
| | 3 | | EIWKQHEDALQKFEDALNQFEELKQL | GGSG C GSGG |
| | 4 | | EIWKQHEDALQKFE Q ALNQFEELKQL-NH ₂ | |

*Contains additional unintentional mutations besides Naq

6.2.1 Constructing a BT3-Intein Expression System

Cloning of the first three helices of BT3 into the commercially available intein purification vector pKYB1, Table 6.2, from NEB was difficult, but ultimately accomplished using conventional cloning methodology at the NdeI, N-terminal, and SapI, C-terminal intein, restriction sites. pKYB1 harbors kanamycin resistance, which was chosen due to cross-contamination of naturally ampicillin resistant photosynthetic organisms in the Dutton laboratory. The expression cassette in pKYB1 follows a T7 promoter and contains a C-terminal VMA intein from *Saccharomyces cerevisiae* followed by a chitin-binding domain (CBD) for affinity purification on chitin columns. The C-terminal intein is normally used for on column cleavage of protein thus allowing single step purification, yet it is easily adapted to chemical ligation by changing the cleavage thiol mixture and eluting the captured thioester. However, because CBD binding to chitin is essentially irreversible (knocking the recombinant protein off involves applying strong base to the column), it is all but impossible to ascertain the efficiency of the cleavage reaction. Nevertheless, expressing the first three helices of BT3 (BT3³) in the commercial vector in BL21(DE3)-RIL cells and attempting purification and cleavage on a chitin column afforded no protein.

Table 6.2. Manipulation of the expressed BT3³ intein fusion protein to obtain soluble expression.

| Expression Vector | N-Terminal Fragment | Intein | Affinity Tag 1 | Affinity Tag 2 | Expressing As |
|---|----------------------------|---------------|-----------------------|-----------------------|----------------------|
| pKYB1 | BT3 ³ | Sce VMA | CBD | | Inclusion Body |
| pKYB1-His ₆ | BT3 ³ | Sce VMA | CBD | His ₆ | Inclusion Body |
| pKYB1-MBP-ΔCBD-His ₆ | BT3 ³ | Sce VMA | MBP | His ₆ | Inclusion Body |
| pKYB1-GyrA-ΔVMA-MBP-ΔCBD-His ₆ | BT3 ³ | Mxe GyrA | MBP | His ₆ | Soluble |

Following the four primer mutagenesis methodology of Chiu et al.,^{8,9} a His-tag was added to the construct, Table 6.2, to allow purification of the intact fusion protein. This was quite revelatory since it became clear that the BT3³-intein construct was expressing as an insoluble inclusion body. Previously expressed maquettes were often fused to proteins known for increasing construct solubility like thioredoxin. It seemed likely that the VMA intein and CBD were insufficient to effect enhanced solubility. MBP, a well established molecular chaperone that lacks any possibly interfering cysteines,¹⁰ was cloned into the CBD His-tag construct by replacing the CBD gene with an MBP gene from another NEB vector (pMal-c2x) using similar methodology as was used in generating the His-tag construct, Table 6.2.¹¹

Surprisingly expression of the MBP containing construct at both 16 °C and 37 °C resulted in further inclusion body formation. Purification of the inclusion body under denaturing conditions worked satisfactorily, but upon refolding a slow aggregation of the protein was observed. Solubility with the MBP fusion partner actually appeared worse than that observed for the intein construct with CBD. It was suspected that since most inteins are endonucleases the VMA intein was likely to have large positively charged DNA binding surface. Ordinarily this might not be a cause for concern, but the BT3 maquette was designed as a highly negatively charged protein with a calculated pI of ~4.7; similarly, MBP's isoelectric point is ~5.2. On the other hand the VMA intein, has a pI of approximately 8, suggesting that under the most common pH conditions, this protein construct would have two highly negatively charged domains (MBO, BT3³) perfectly capable of binding the positively charged VMA intein domain. Aggregation would naturally occur because individual proteins would be able to crosslink through ionic interactions. A search of NEB's

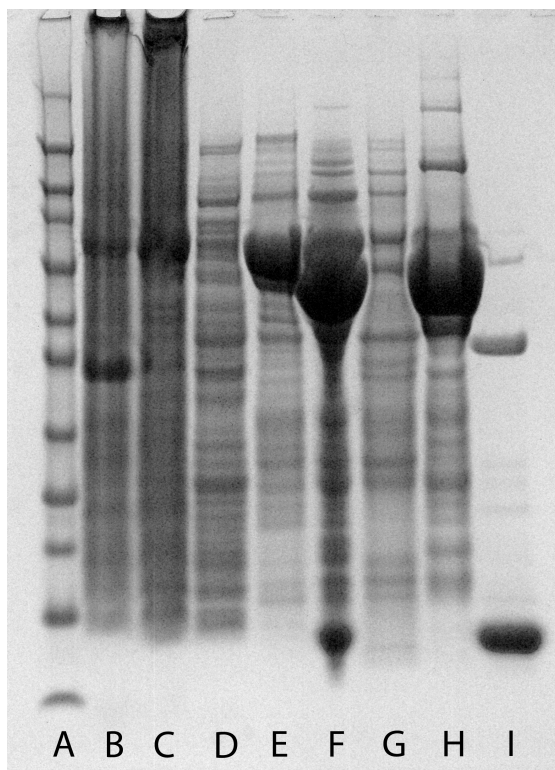


Figure 6.2. SDS-Page protein gel of BT⁶³-intein construct from expression to formation of BT⁶³-mensa thioester. (A) Novex Sharp Standard Ladder; (B) Cell pellet from expression; (C) Supernatant from cell lysis (D) Co-NTA Talon Resin column flowthrough; (E) Co-NTA Talon Resin Column peak; (F) Crude Na-MENSA thioester capture reaction; (G) First 'peak' from Superdex 75 gel filtration column; (H) Second 'peak' from gel filtration column; (I) Third 'peak' from gel filtration column corresponding to BT⁶³-mensa thioester. See text for details.

intein database revealed that the *Mycobacterium xenopi* Gyr A intein has a pI ~5.4. Cloning this intein from NEB's pMXB10 vector over the VMA intein in the MBP construct gave the final expression vector, Table 6.2. Expression of the construct BT³³-MxeGyrA-MBP-His₆ proceeded in high yield, affording ~50 mg/L of the fusion protein from terrific broth, see Figure 6.2 for expression of a subsequent construct under similar conditions.

6.2.2 Chemical Ligation of BTX³ with Synthetic Helices

Purification of BT³³-MxeGyrA-MBP-His₆ was straightforward using a Co-NTA Talon resin affinity column followed by dialysis into cleavage buffer and concentration. Room temperature capture of the BT³³ thioester was also fairly standard, using a large excess of the highly water soluble and odorless thiol sodium mercaptoethanesulfonic acid (mensa) at pH 9.¹² This reaction was found to proceed more completely when the reaction was run under argon (or in a glove box) after the solution had been degassed, presumably because

the thioester capture reaction is competitive with disulfide formation when oxygen is present. Cleavage yields varied, but it is estimated that >90% of the peptide gets cleaved under ideal conditions. The mensa thioester is stable at pH 7 for up to a month.¹² However, hydrolysis rates rapidly increase with pH, so it was imperative to quickly purify the BT3³-mensa thioester product, which was effected by a Superdex 75 size exclusion column in pH 7.5 buffer, which cleanly gave the presumptive thioester, Figure 6.2. After serial dilution-concentration of the appropriate column fractions against water, the BT3³-mensa thioester product was lyophilized. Following the procedure described by Johnson and Kent,¹² low mM BT3³-mensa was reacted with 2-3 equivalents of N-terminal cysteine containing BT0.1, Table 6.1, under denaturing conditions in the presence of 200 mM mercaptophenylacetic acid and the reducing agent TCEP. After completion, the reaction was purified with HPLC since other chromatographic techniques available cannot separate the full-length protein and the three-helix fragment. Purification afforded ~70% by HPLC peak integration of the semisynthetic four-helix bundle BT3.1, Table 6.1, confirmed by MALDI-TOF-MS.

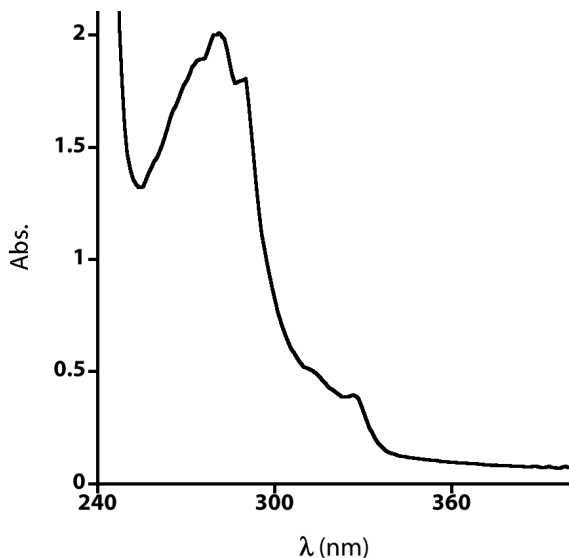


Figure 6.3. UV/vis spectrum of BT5.3 incorporating Naq(OMe/OMe) in 250 mM Sodium Borate, 100 mM KCl, pH 9.0, showing a characteristic tryptophan spectrum over that of Naq(OMe/OMe).

Repeating this procedure with BT0.2 afforded the first four helix bundle with Naq(OMe/OMe), BT3Naq. As described in Chapter 5, this protein, which accidentally incorporated several unintentional mutations in the synthetic helix, only bound one heme like full length BT3. Following the same design path as described in Chapter 5, a new construct was

created which had a shorted loop between helices 2 and 3, BT5³, Table 6.1, which was exactly analogous to BT5. At the same time BT0.3, which was designed to fix the mutations incorporated into BT0.2 and place Naq approximately equidistant from the two heme sites, was synthesized. Chemical ligation as described of BT5³ with BT0.3 afforded BT5.3, Table 6.1 and Figure 6.3.

6.2.3 Deprotecting Naq(OMe/OMe) in a Four Helix Bundle

Unlike the previous Naq(OMe/OMe) incorporating peptides, treatment of BT5.3 with stoichiometric DDQ did not effect activation of Naq. However, there was some evidence of protein modification, possibly due to oxidation of the cysteine or methionine residues, which have a much lower potential than Naq(OMe/OMe) thereby making these residues more likely to become modified by DDQ *before* Naq is activated. As mentioned in Chapter 3 in the discussion of failed deprotection attempts, the mixture TFA/trimethylsilyl-trifluoromethane sulfonic acid/thioanisole may have deprotected Naq(OMe/OMe), but this was quickly followed by both backbone hydrolysis and Naq quinol backbone cyclization. Since BT5.3 was not soluble in methylene chloride, the common solvent used in methyl ether deprotection with BBr₃, the equivalent reagent B(OTFA)₃ was used with BT5.3 in TFA. Backbone hydrolysis appeared to occur before sidechain deprotection. Without a suitable oxidative or Lewis-acid deprotection method for Naq(OMe/OMe) in a four-helix bundle, it proved necessary to rethink the protecting groups on Naq.

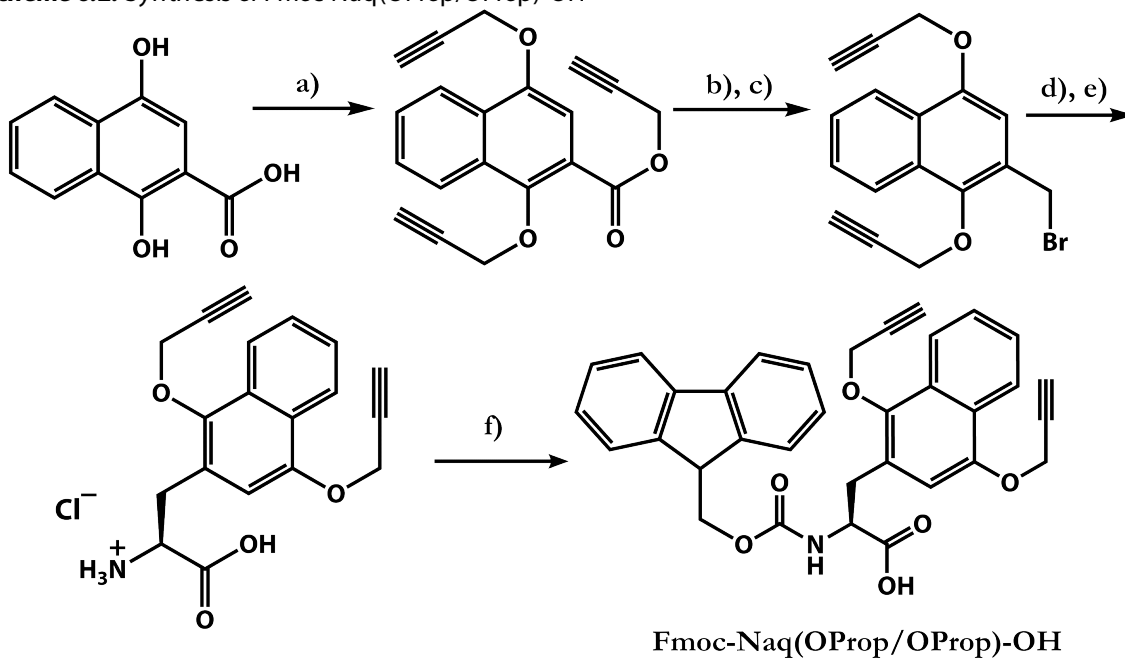
6.3 Reductive Deprotection of Naq in a Four-Helix Bundle

Having searched and tried more labile protecting groups for Naq in the past, the search for a new protecting scheme was effectively limited. Because the major issues confronting using Naq(OMe/OMe) in proteins incorporating natural amino acids of all functional groups were solely due to the fact that oxidative cleavage was the only effective means to activate Naq from Naq(OMe/OMe), the protecting group search was limited to reductively labile alkyl groups. This allowed the development of an orthogonally protected Naq that could, in theory, be used in concert with Naq(OMe/OMe) and be activated under very different conditions. The broad class of aryl alkyl protecting groups, including benzyl, was discounted since these groups tend to be fairly acid labile which would be incompatible with multiple steps of the synthesis route of Naq and peptides incorporating Naq. Similarly, other protecting groups were eliminated from consideration due to concerns of reactivity during the course of peptide synthesis, like the 2-ethylhalides. In the end, though it is not the only option, the underused propargyl protecting group was chosen because of its relative inertness and ease of incorporation into the established synthetic scheme to Fmoc-Naq(OMe/OMe)-OH.

6.3.1 Synthesizing Fmoc-Naq(OProp/OProp)-OH

Since the synthesis of Fmoc-Naq(OMe/OMe)-OH is described in detail in Chapter 2, the synthesis of Fmoc-Naq(OProp/OProp)-OH which follows the same exact synthetic route will only be touched upon, Scheme 6.2. The only major change from the synthesis of Fmoc-Naq(OMe/OMe)-OH is the combination of the first two steps into a single alkylation step in which three propargyl groups are attached to the 2-carboxy and 1- and 4- hydroxyls. The remainder of the synthesis follows without complication due to the propargyl protecting groups. Purification of the initial O'Donnell amino acid synthesis product, a silica column which was of limited use in the synthesis of Fmoc-Naq(OMe/OMe)-OH, was entirely

Scheme 6.2. Synthesis of Fmoc-Naq(OProp/OProp)-OH



- a) Propargyl Bromide, K₂CO₃, KI, DMF, 90 °C, 60% b) LiAlH₄, THF, 100% c) PBr₃/CCl₄ 89%
d) CsOH, Benzophenone Glycinate, 3rd Generation Cinchonidium PTC, Tol, -20 °C
e) EDT, TFA, 68% after two steps f) Fmoc-Osu, NaHCO₃, DMF, 60 °C, 90%

skipped with the rationale that the final two steps, including the purification of Fmoc-Naq(OProp/OProp)-OH would be sufficient to removed any contaminants.

6.3.2 Incorporation of Naq(OProp/OProp) into Single Helix BT0.4

The previously synthesized single helix C-terminal fragments for chemical ligation BT0.1-0.3, either had no Naq (BT0.1) or Naq was placed equidistant from the two heme

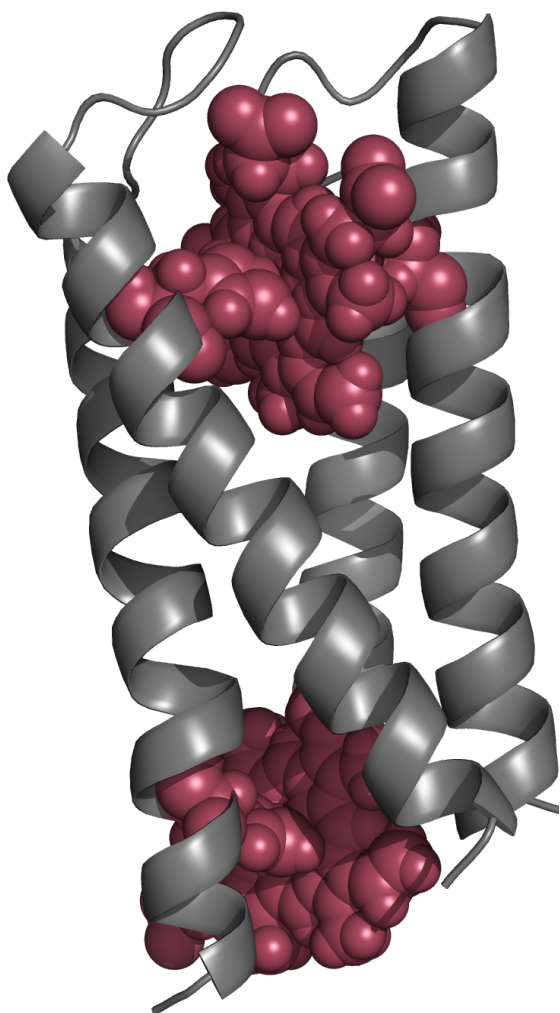


Figure 6.4. Energetically minimized structure of HP7 based upon the crystal structure of H10H24-L6I-L13F-L31M. Hemes and ligating histidines are in red. Structure graciously provided by the Klein group.

binding sites using rough approximations of the topology of the BT3 protein. Upon further examination of a energetically minimized structure of HP7 based upon the crystal structure of H10H24-L6I-L13F-L31M generated by the Klein group, Figure 6.4, the best edge-of-heme to edge-of-Naq equidistant site on the exterior of the fourth helix was a full turn away ($i+4$) from the previously designed site. This error was likely due to a miscalculation of the width of the helices, but gives further motivation to obtaining an actual structure of one of the holo-maquettes.

Fmoc-Naq(OProp/OProp)-OH easily incorporated into synthetic helix BT0.4, Table 6.1, as designed.

6.3.3 Deprotection of Naq(OProp/OProp) in BT0.4

BT0.4, Table 6.1, has nearly all of the side chain functionality of a four-helix bundle and thus was a good framework to attempt deprotection methodology. Propargyl groups have been deprotected by using a variety of methods ranging from BBr_3 ¹³ to $\text{Pd(II)}/\text{TEA}$.¹⁴ The palladium-based methods were particularly attractive since Pd^0 is often used to deprotect alloc groups from peptides. Initial attempts to use TEA as an electron source for palladium(0) tetrakis(triphenyl phosphine) ($\text{Pd(PPh}_3)_4$) reductive deprotection of Naq(OProp/OProp) at 80 °C in DMF consumed the starting material, but the product was not identified. In this and all other reaction with $\text{Pd(PPh}_3)_4$, the palladium catalyst was in large excess compared to the protein to prevent the cyclization of the propargyl group to the quinone ring system at the site of its attached oxygen, a reaction that is known to occur with Pt(II) catalysts.¹⁵ Changing the reductant to BH_3NH_3 appeared to work (MALDI-TOF-MS mass was 8 D off from expected product mass in an uncalibrated run), but produced copious quantities of what appeared to be colloidal palladium.

A room temperature deprotection reaction of Naq(OProp/OProp) in BT0.4 using phenylsilane as the reductant was halted prematurely to allow isolation of intermediates. Three major products were observed during HPLC, each of which had a slightly different 'reduced' Naq spectrum, which matched the observation that Naq(OMe/OMe) has a significantly blue shifted absorbance spectrum in the 300 nm range compared to Naq_{red} . MALDI-TOF-MS indicated that none of peaks separated by HPLC was the starting material, and it appeared that there were three nearly equal mass losses from starting material to smallest product. This was unanticipated, and it was initially believed that given some of the inconsistencies of the MALDI-TOF-MS instrument that was used to collect the data, that

one of the ‘losses’ in mass was simply systematic error in the mass measurement. In order to gauge the exact mass of the species, ESI-MS was performed, which confirmed the MALDI-TOF-MS results but because of its higher mass accuracy presented a fuller picture. The product with the largest mass had a mass that corresponded to the loss of exactly one sulfur atom. The second largest mass observed corresponded to the loss of a single propargyl from the peak with the largest mass, exactly. And finally, the smallest mass corresponded to loss of two propargyls and one sulfur atom from the starting material, exactly. In addition, collected with the smallest mass peak, was a small amount of peptide that was too heavy by a single oxygen atom, suggesting that some amount of the double deprotected Naq_{red} had become modified by oxygen (or oxidized, modified by water and oxidized again).

The observed desulfurization reaction likely produces alanine at the cysteine site. While it is somewhat surprising that the desulfurization reaction occurs and is seemingly significantly faster than the depropargylation, there are other textbook methods of desulfurization of cysteines under reducing conditions, for instance using Raney nickel. Regardless, this fortuitous observation allows the simultaneous deprotection of Naq(OProp/OProp) and elimination of the free cysteine sulfur. It is not clear whether the desulfurization reaction occurs under all conditions (notably whether ammonia borane effects the same transformation – it would seem not given the previously discussed results) or whether the reaction will similarly cleave thioethers like that found in methionine. Nevertheless, it appeared that a method for deprotecting Naq(OProp/OProp) in a nearly fully functionalized helical segment was possible.

6.3.4 Incorporation and Deprotection of Naq(OProp/OProp) in a Four Helix Bundle

At the time of the creation of BT5.4, it was not clear whether the N-cap or the long cross-bundle loop between helix 2 and 3 in BT3 was responsible for changing the heme binding properties of the single chain maquettes. Since there was no need to leave anything to chance, a new maquette intein construct was created analogous to BT6 called BT6³, Table 6.1. For practical reasons it does not exactly match the sequence of BT6 since the expression system created for the maquette inteins requires an N-terminal methionine, so instead of simply having a single N-terminal glycine residue that is present in BT6 due to the TEV cleavage site, the BT6³ fragment begins with the sequence MGG. Chemical ligation of BT6³-mensa thioester as described with BT0.4, produces BT6.4 with Naq(OProp/OProp) in the fourth helix, Table 6.1.

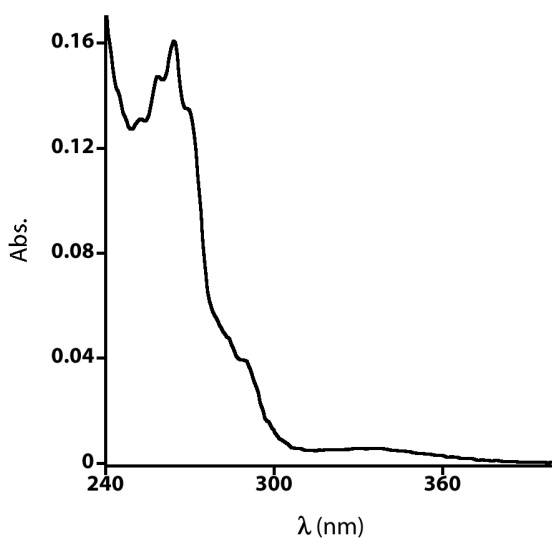


Figure 6.5. Spectrum of BT6.4 after reductive deprotection of Naq(OProp/OProp). Scattering background was subtracted by fitting long wavelength spectral data to the Rayleigh scattering equation and subtracting the fit from the raw data. The band at 335 nm matches that expected for Naq_{ox}. The band at ~260 nm is due to contaminating triphenylphosphine. See text for details. Spectrum was taken in 20 mM K₂PO₄, 100 mM KCl, pH 8.

Following the deprotection protocol described above, a single HPLC purification's worth of BT6.4 was dissolved in a 300 μ L of DMF with 100 μ L phenylsilane. Excess Pd(PPh₃)₄ was added and the reaction vessel was briefly evacuated before being placed under positive pressure argon to eliminate some amount of the atmospheric oxygen. After 2 hours, the reaction was purified. While the HPLC was not ideal, a single peak appeared to contain both BT6.4 with activated Naq and

significant amounts of a species with a phenyl absorption spectrum, likely the triphenylphosphine ligand of the catalyst. Spectroscopy of the acquired product, after correction for a significant scattering background (again, likely triphenylphosphine), produced a spectrum that had some of the hallmarks of an oxidized quinone in a protein with tryptophans, Figure 6.5. While it is not possible to observe the major absorbance at 252 nm for oxidized Naq because it sits under the phenyl absorbance, the band at ~ 335 nm is in approximately the right ratio to the tryptophan band at 290 nm for a protein having four tryptophans and one Naq_{ox}. Obviously this is simply preliminary data, but it is the best evidence to date that the incorporation and deprotection of Naq in a four-helix bundle has been accomplished. In addition, though a full heme titration was not performed, BT6.4 binds at least one equivalent of heme, Figure 6.6A, which can subsequently be reduced by borohydride in the presence of CO, Figure 6.6B. There is clear evidence (data not shown) that Naq in BT6.4 under these conditions goes reduced based upon difference spectra. It is not apparent whether heme reduction with borohydride goes through Naq, or whether this occurs naturally. There are reports of a ferric heme going reduced by borohydride in *Cyt-c* oxidase,¹⁶ but *Cyt-c* oxidase has other redox centers capable of reacting with borohydride and then passing the electrons to the heme, notably the copper centers. In any event, there is tangible evidence that Naq has been functionally incorporated into a four-helix bundle capable of binding hemes, allowing for the first time studies of electron transfer between heme(s) and Naq thus extending the range of maquette functional emulation.

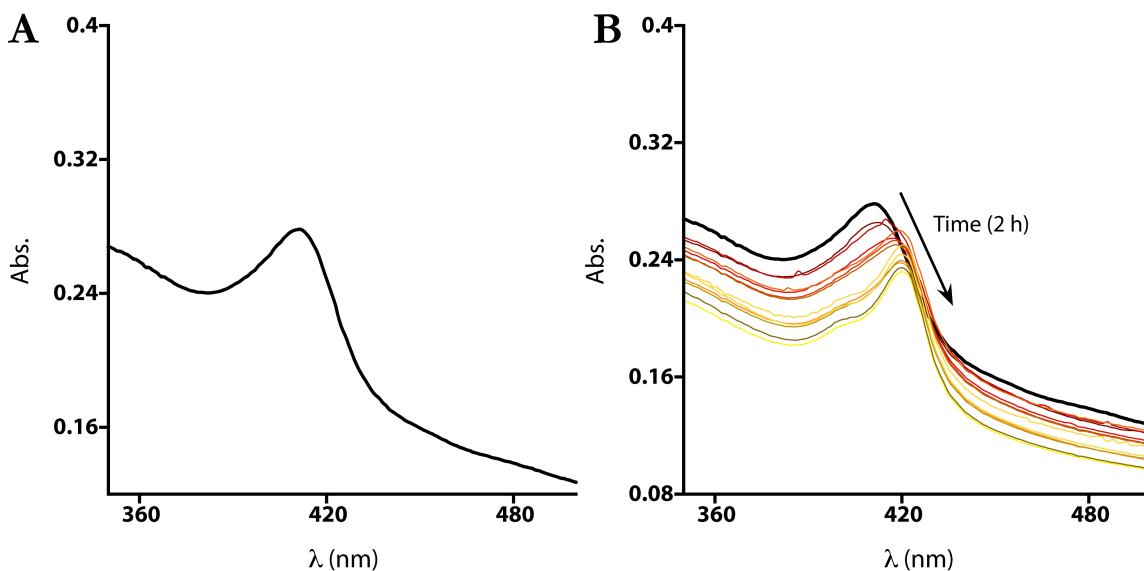


Figure 6.6. Formation of carboxy-ferrous complex of heme bound BT6.4 upon treatment with borohydride. Soret region spectrum of BT6.4 bound to 1 equivalent of ferric heme under CO (A); heme Soret band is near that observed for BT6. BT6.4, under CO, after addition of sodium borohydride showing a slow formation of the carboxy-ferrous state with time (B). Spectra are uncorrected for large scattering background due to insoluble triphenylphosphine. Spectrum was taken in 20 mM K_2PO_4 , 100 mM KCl, pH 8.

6.4 Future Directions

Incorporation of activated Naq into a heme-binding four-helix bundle represents a major advance in the design of non-natural multifactor proteins. However, the critical experiments remain to be done. Importantly, in work done by Sarah Chobot, Naq has also been incorporated and deprotected in an amphiphilic homotetrameric maquette, and while heme incorporation was disrupted in the first design, it is clear that with a better site choice the design possibilities of four helix bundles with Naq are substantial. Experiments that allow the observation and exploration of electron transfer reactions, photosynthetic and respiratory intermediates, and the natural engineering principles that guide semiquinone stability and proton uptake can be cleanly designed with Naq. Practically, the first experiments for observing electron transfer in a heme-binding maquette between a heme and Naq are quite simple, Figure 6.7. Starting with a Naq containing protein like BT6.4 and

binding a single heme, it is possible to capture a high effective potential carboxy-ferrous state by reducing Naq and the heme in the presence of CO, Figure 6.7A to 6.7B. Re-poising the electrochemical potential of the sample by oxidizing just Naq sets the stage for observing transient formation of a semiquinone, Figure 6.7C. Photolysis of the carboxy-ferrous heme bond, regenerates a low potential heme which is sufficiently reducing to generate, transiently, a Naq semiquinone, Figure 6.7C to 6.7E. With values of cofactor spacing obtained from the minimized structure in Figure 6.4, 13 Å, the known solution potentials of the

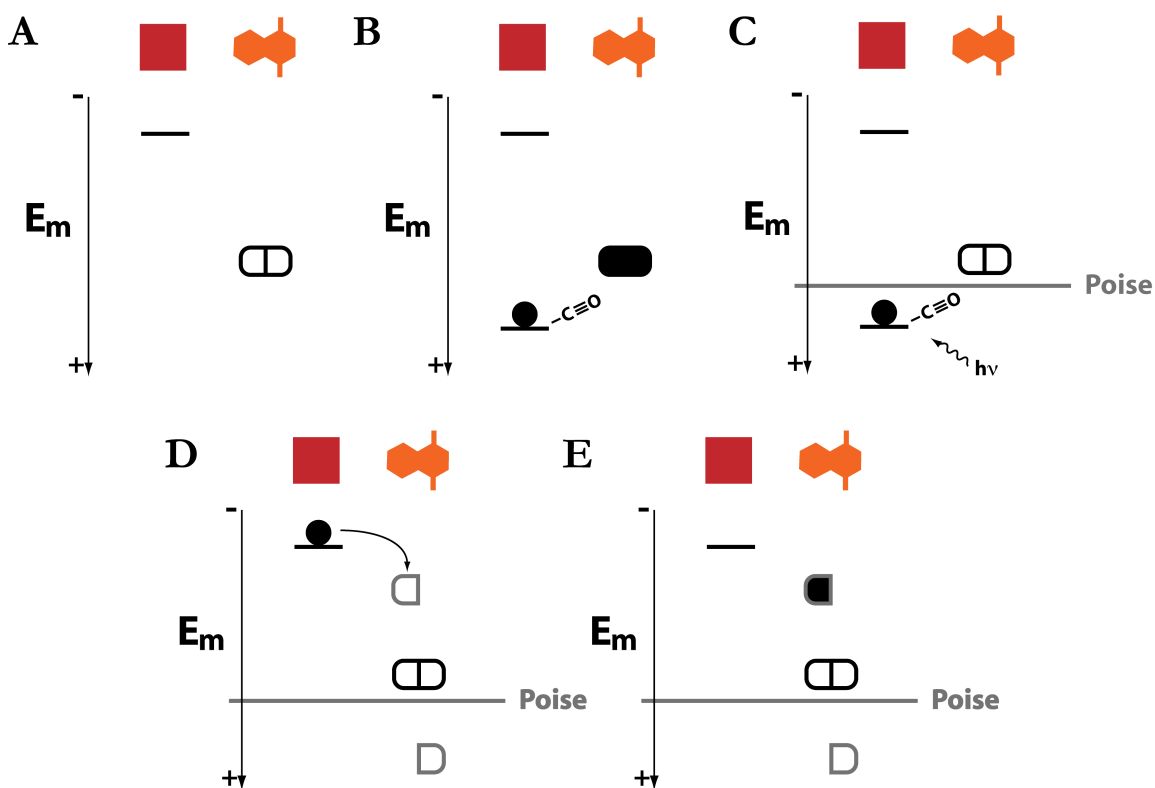


Figure 6.7. Schematic of CO-flash photolysis of a single heme bound BT6.4-like maquette incorporating an activated Naq (A). Cofactor identity is across top of each image; E_m is drawn to arbitrary scale on left, note orientation is with respect to free energy and not midpoint potential. Unoccupied lines or boxes represent oxidized heme or Naq, respectively. Occupied lines or boxes represent reduced heme or Naq respectively. Grey boxes represent midpoint potentials of Naq's two semiquinone couples; black boxes represent two-electron midpoint. Reducing Naq and heme in presence of carbon monoxide generates the carboxy-ferrous state and fully reduced Naq (B). Poising the system such that Naq is oxidized while the carboxy-ferrous heme remains reduced sets the stage for CO photolysis (C). Once no longer ligated to CO, ferrous heme returns to a potential low enough to reduce Naq's semiquinone (D). Transient Naq semiquinone state (E) that will eventually relax back to the state in (C) as long as CO is present. See text for details.

naphthoquinone semiquinone, \sim -200 mV,¹⁷ and heme *b* in HP7, \sim -300 mV, and the kinetics of CO photolysis reactions from studies with HP7,¹⁸ the lifetime of the semiquinone in this construct can be simulated (Chris Moser, personal communication). Significant populations of semiquinone are theoretically observable for up to 1 s under the most favorable conditions and up to \sim 10 ms under the least favorable conditions simulated (Chris Moser, personal communication). While these calculations ignore the collisional exchange of electrons between two Naq_{sq} in solution, observation of that particular exchange route, through non-transient oxidation of the heme, also indicates that semiquinone formation was likely and, regardless, that electron transfer from the heme to Naq was observed. This experiment is immediately at hand because the work presented in this chapter and in Chapter 5 firmly sets the foundation for this simple demonstration of electron transfer between a quinone and heme in a maquette. Further work will start tackling the larger unexplored aspects of quinone electron transfer reactions in the various quinone active sites of photosynthetic and respiratory complexes.

6.4 References

- (1) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc Natl Acad Sci U S A* **1998**, *95*, 6705.
- (2) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776.
- (3) Pentelute, B. L.; Gates, Z. P.; Tereshko, V.; Dashnau, J. L.; Vanderkooi, J. M.; Kossiakoff, A. A.; Kent, S. B. *J Am Chem Soc* **2008**, *130*, 9695.
- (4) Pentelute, B. L.; Gates, Z. P.; Dashnau, J. L.; Vanderkooi, J. M.; Kent, S. B. *J Am Chem Soc* **2008**, *130*, 9702.
- (5) Mandal, K.; Pentelute, B. L.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. *H. J. Am. Chem. Soc.* **2009**, *131*, 1362.
- (6) Otomo, T.; Ito, N.; Kyogoku, Y.; Yamazaki, T. *Biochemistry-U S* **1999**, *38*, 16040.
- (7) Gogarten, J. P.; Hilario, E. *Bmc Evolutionary Biology* **2006**, *6*.
- (8) Chiu, J.; Tillett, D.; Dawes, I. W.; March, P. E. *J Microbiol Methods* **2008**, *73*, 195.
- (9) Chiu, J.; March, P. E.; Lee, R.; Tillett, D. *Nucleic Acids Res* **2004**, *32*, e174.

- (10) Fox, J. D.; Waugh, D. S. *Methods Mol Biol* **2003**, *205*, 99.
- (11) Tillett, D.; Neilan, B. A. *Nucleic Acids Res* **1999**, *27*, e26.
- (12) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640.
- (13) Punna, S.; Meunier, S.; Finn, M. G. *Organic Letters* **2004**, *6*, 2777.
- (14) Pal, M.; Parasuraman, K.; Yeleswarapu, K. R. *Organic Letters* **2003**, *5*, 349.
- (15) Nevado, C.; Echavarren, A. M. *Chemistry* **2005**, *11*, 3155.
- (16) Martin, A. P.; Doyle, G. E.; Stotz, E. *Journal of Biological Chemistry* **1965**, *240*, 1402.
- (17) *Physical Chemistry of Semiquinones*; Swallow, A. J., Ed.; Academic Press: New York, 1982.
- (18) Koder, R. L.; Anderson, J. L.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. *Nature* **2009**, *458*, 305.

Chapter 7: Conclusion

This work was driven by the fact that quinones in the active sites of the natural oxidoreductases of respiration and photosynthesis show a range of electrochemical properties that have yet to be fully explained by more than a half-century of intense study. Mutational analysis and studies in which the natural quinone substrates are exchanged for other quinones have done little to clarify the physical means of biological control of quinone reactivity. Some of the obfuscation of control mechanisms comes from the complexity of the natural proteins themselves. The lack of consensus quinone binding sites in the mitochondrial and chloroplast membrane quinone oxidoreductases makes it nearly impossible to know whether amino acids in quinone active sites are mechanistically relevant or simply involved in quinone binding. Recently determined structures of membrane quinone oxidoreductases provide some insights, but as of yet quinones have not been observed in a representative of each type of quinone binding site. Additionally, because of

the indistinct spectroscopy of multifactor quinone oxidoreductases, it is difficult to monitor single site mutagenesis effects on substrate quinone properties. The work presented in the previous chapters laid the foundation for an entirely new approach to understanding and using the natural engineering principles of quinone dependent redox enzymes by the development of a naphthoquinone amino acid capable of being incorporated into simplified protein maquettes designed to emulate critical natural electron transfer reactions.

The naphthoquinone amino acid, Naq, was designed to maintain quinone function while being synthesized into peptides. Indeed, the full range of natural quinone electrochemistry was observed in the short model peptide heptaNaq (Chapter 2).¹ In aqueous solution, heptaNaq demonstrated typical two-electron, two-proton redox chemistry over a fairly wide pH range, while maintaining an almost identical midpoint to model compound 2-methyl-1,4-naphthoquinone. This contrasts to several literature reports in which the incorporation of a quinone at a cysteine residue resulted in large and inexplicable changes in redox reactivity of the quinones. In addition to maintaining natural reactivity in an aqueous environment, the Naq cofactor in heptaNaq also demonstrated kinetically accessible semiquinone states in aprotic electrochemical studies. With a rapidly changing potential in cyclic voltammetry, proton uptake is bypassed and the anionic semiquinone is formed to the exclusion of all other redox states. At slower scan rates, and with stoichiometric protons in the solution, Naq demonstrates two single-electron electron transfer reactions that are coupled to an intervening protonation event. These three distinct redox behaviors of heptaNaq under different conditions correspond nicely to the three broad classes of active sites found in natural quinone oxidoreductases. heptaNaq's aqueous solution behavior is reminiscent of the activity of the Q_o site in *bc₁* in which there is little

evidence of that a semiquinone naturally exists (Chapter 1). In an aprotic medium, with limiting protons, heptaNaq further demonstrates both the activity of the Q_A site of the photosynthetic reaction center where the anionic semiquinone is only formed transiently, and the Q_i site of bv_1 , where proton uptake is limited enough that a stable semiquinone is formed but is eventually converted to the hydroquinone. These observations both support the general concept that proton uptake is possibly the most important limiting factor in the differentiation of the quinone properties observed at natural binding sites, and demonstrate that Naq is capable of recapitulating the major mechanistic features of quinones in natural oxidoreductases.

While the heptaNaq project described in Chapter 2 demonstrated the full electrochemical mechanistic versatility of Naq in emulating natural quinone cofactor chemistry, it is important to recognize that Naq is not a natural cofactor. Its very nature as an amino acid means that Naq interacts with the backbone of peptides in ways that no natural cofactor could. It was important to confirm that Naq did not interfere with the common alpha-helix secondary structural element and to gauge a maximum change in interaction energy with the backbone upon conformation change from coil to helix. The second property is particularly important since Naq can theoretically be placed in any structural element, and it would be difficult to separate the local environmental effects from the more global backbone topology in determining the redox properties of Naq. Adapting a peptide family that had been used to generate an enthalpic measure of alpha-helix propensity, it was observed that Naq experiences a -16 mV (~ 0.7 kcal mol⁻¹) change in redox potential upon going from an unfolded random coil state to a folded helical state (Chapter 3). This is not a large change in midpoint potential; it is approximately the same energetic

cost as a slightly destabilizing mutation, so it is not believed that a single Naq alone could break a stable alpha-helix upon getting reduced. In addition, while it was observed that Naq acts as a C-capping amino acid for short helices, the unambiguous observation of Naq within a stretch of alpha-helix in an extended alanine-rich helix suggests that when well placed in a stable helix Naq should not interfere with secondary structure formation in either redox state. Otherwise it might be necessary to design specific interactions with Naq to reduce its tendency to break short helices.

Using *SCADS*, a computationally guided design approach, Naq was incorporated into the helix of a miniprotein, NaqCage (Chapter 4). The NaqCage is based upon the 20-residue TrpCage, which has a compact hydrophobic core surrounding a buried Trp residue. By replacing this Trp, the NaqCage was designed to stably incorporate Naq in a similar structural context, allowing investigations on the effects a well-packed protein environment on its redox chemistry. While the initial computational designed target was not appreciably structured due to destabilization of the helix by Naq, the NaqCage was iteratively redesigned by hand to a point where a solution structure could be determined. Analysis of the NMR data was complicated by several factors, but nearly all of the backbone protons have been assigned and further analysis may yield a structure. This would be the first report of stably incorporating another amino acid at the site of the Trp, but more importantly this work generated a water-soluble, synthesizable framework for testing the effects of hydrogen bonding and electrostatic interactions on Naq's redox chemistry in an aqueous environment – a crucial set of information that is woefully lacking in the literature.

Finally, to take full advantage of Naq's potential in the well established heme-binding maquettes, a cross-bundle loop was added to the sequence of the homodimeric HP7. By

adding the loop and removing the symmetry, a single chain four-helix bundle maquette that allowed single site incorporation of mutations (Chapter 5) and the construction of semisynthetic bundles that incorporated Naq (Chapter 6) was produced. The bundles incorporating Naq bind at least one heme, which allows the creation of maquettes designed to explore electron transfer reactions between Naq and heme.

What has been presented here is nothing less than the first stone of a new foundation for maquette work. In making good use of Naq, profound opportunities exist for creating proteins that offer more than just a full-throated mimicry of natural systems. Indeed, maquettes incorporating Naq are not limited to the imagination of nature, and Naq offers methods of exploring diverse protein environments and cofactor arrangements that evolution may never have stumbled upon. While it is clear that the most affirmative knowledge gained from continued use of Naq is in understanding the intricate mechanisms of quinone active sites in natural proteins, Naq's long term scientific applicability to more practical considerations like energy generation and conversion may be its lasting legacy.

7.1 References

- (1) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun (Camb)* **2009**, 168.

Appendix A: Experimental Methods

A.1 Introduction

The majority of the methods described below were published as supplemental data to Lichtenstein et al.¹ Where appropriate additional methodology has been added and described. Raw spectra found in the paper's supplementary data have been removed for brevity. As elsewhere, Naq is represented by the single letter code **Q**, where appropriate.

A.2 General Methods

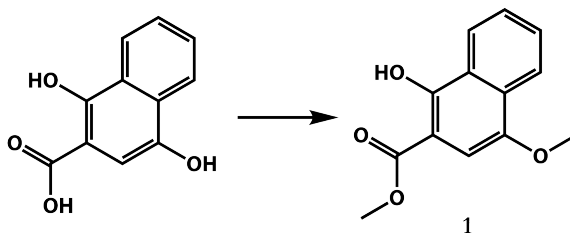
All reactions were carried out under a positive pressure of argon or nitrogen unless otherwise indicated. Commercial reagents were from Sigma Aldrich, unless otherwise indicated, and used as received. HPLC grade diethylether, ethyl acetate, hexanes, methylene chloride, and acetonitrile were purchased from Fisher Scientific. HPLC grade carbon tetrachloride was purchased from Sigma Aldrich. Anhydrous dimethylformamide (DMF) and anhydrous tetrahydrofuran (THF) were purchased from Aldrich and stored under argon.

Sulfuric acid, glacial acetic acid and hydrochloric acid were purchased from Fisher Scientific. All other solvents were reagent grade and used as received. Flash chromatography was performed using Fisher 230-400 silica gel 60 unless otherwise indicated. Analytical TLC was carried out on Analtech silica gel HLF pre-coated glass plates with detection by UV illumination unless otherwise indicated. ^1H and ^{13}C NMR spectroscopy of small molecules was performed on a Varian Inova-500 (at 500 MHz and 125 MHz, respectively) using CDCl_3 as solvent (unless otherwise noted) and referenced with the residual solvent signal (CDCl_3 , 7.24 and 77.0 ppm, respectively) or referenced externally. The following abbreviations are used to describe peak patterns where appropriate: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad). Small molecule mass spectrometry data were obtained by the Penn State University Proteomics and Mass Spectrometry Core Facility. Peptide synthesis was carried out either a Pioneer Peptide Synthesis System or CEM Liberty Microwave Synthesiser using standard Fmoc chemistry (protected natural amino acids from Novabiochem) with HATU/HOAt (Genscript Corporation) as the coupling reagents (4-5.0 mol equivalents of activator/Fmoc-amino acid) on PAL-PEG-PS resin (Applied Biosciences, most peptides) or Fmoc-Ser(tBu)-Wang resin (Anaspec, peptides containing C-terminal serine carboxylates). After cleavage and deprotection of the natural amino acids (Reagent R: 90:5:3:2 TFA/Thioanisole/Ethane Dithiol/Anisole) under argon, the resin was filtered and washed with 10 mL TFA (3x). The TFA was removed *in vacuo*, and the crude peptide precipitated with cold methyl tert-butyl ether. The identity of all peptides was confirmed with MALDI-TOF-MS using either α -cyano-4-hydroxycinnamic acid or sinapic acid (Sigma) as the peptide matrix. MALDI-TOF-MS determinations were performed on a PerSeptive Biosystem Voyager-DE RP. Kjeldahl analysis was performed by Galbraith

Laboratories, Inc. CHN elemental analysis was performed by Midwest Microlab, LLC. UV/*vis* spectra were acquired on an Agilent 8453 UV-Visible Spectrophotometer or a Varian Cary-50 Spectrophotometer. CD Spectra were acquired on either an AVIV 405 or 410 CD Spectrophotometer. Automated titrations were performed with a Hamilton Company automated diluter.

A.3 Naq Synthetic Methods

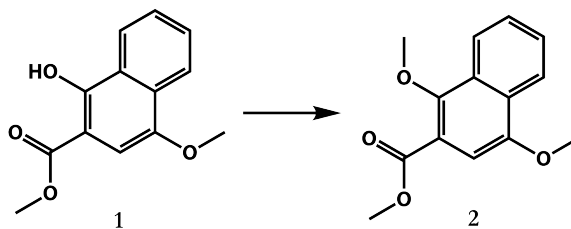
A.3.1 Methyl 1-hydroxy-4-methoxy-2-naphthoate (1)



(For alternative method see Boger and Jacobson²) 40.83 g (200 mmol) 1,4-dihydroxynaphthoic acid was suspended in methanol (200 mL) at room temperature. With stirring, sulfuric acid (30 mL, 2 eq) was added dropwise to the solution such that the reaction mixture was nearly refluxing when the addition was complete. The reaction was refluxed for 24 hr at which point an additional 2 equivalents of sulfuric acid and 50 mL of methanol were added. After an additional 24 hr of refluxing the reaction appeared complete by TLC (eluent toluene). The reaction was diluted into 700 mL of CH₂Cl₂ and was added to 1 L of saturated aqueous sodium bicarbonate into a 2 L beaker with stirring. Solid sodium bicarbonate (Fisher Scientific) was added until the aqueous layer no longer evolved carbon dioxide at which point the layers were separated. The organic layer was dried over anhydrous magnesium sulfate (Fisher Scientific) and the solvent was removed *in vacuo*. Silica column

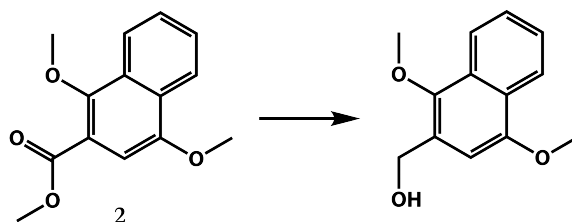
chromatography (eluent toluene) afforded 43.40 g (93.4%) of the methyl 1-hydroxy-4-methoxy-2-naphthoate as a pale green solid. NMR ^1H (500 MHz) δ 3.97 (s, 3H), 4.00 (s, 3H), 7.02 (s, 1H), 7.57 (t, 1H, $J=7.5$ Hz), 7.63 (t, 1H, $J=7.5$ Hz), 8.19 (d, 1H, $J=8.2$ Hz), 8.39 (d, 1H, $J=8.2$ Hz), 11.62 (s, 1H); ^{13}C (125 MHz) δ 52.49, 55.95, 100.71, 104.52, 122.14, 124.04, 125.78, 126.66, 129.26, 130.14, 147.96, 155.88, 171.57; Expected MS: M-H^+ 231.1, Found 231.1.

A.3.2 Methyl 1,4-dimethoxy-2-naphthoate (2)



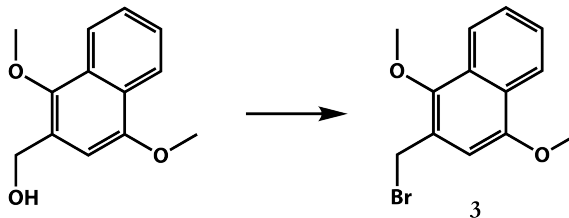
(For alternative method see Newman and Choudhary³) 35.5 g (153 mmol) of methyl hydroxynaphthoate **1** was dissolved in 155 mL of anhydrous DMF. To the solution was added 35.5 g (257 mmol, 1.67 eq) of K_2CO_3 (Fisher Scientific) with stirring. Iodomethane (58 mL, 929 mmol, 6 eq) was added and the reaction was refluxed for 48 hours at which point TLC analysis (eluent toluene) showed that the reaction was complete. The reaction was diluted into 1L ethyl acetate and extracted three times with 1L water, and once with 1L brine. The organics were collected and the solvent removed *in vacuo*. After chromatography (eluent toluene) the trimethyl naphthoate **2** was obtained in 90% yield (33.9 g) as a beige solid. NMR ^1H (500 MHz) δ 3.99 (s, 3H), 4.00 (s, 3H), 4.01 (s, 3H), 7.16 (s, 1H), 7.58 (m, 2H), 8.23 (m, 2H); ^{13}C (125 MHz) δ 52.48, 55.93, 63.51, 103.74, 118.88, 122.50, 123.67, 127.25, 127.94, 128.99, 129.40, 151.57, 152.25, 167.01; Expected MS: MH^+ 247.1, Found 247.1.

A.3.3 1,4-Dimethoxy-2-hydroxymethylnaphthalene



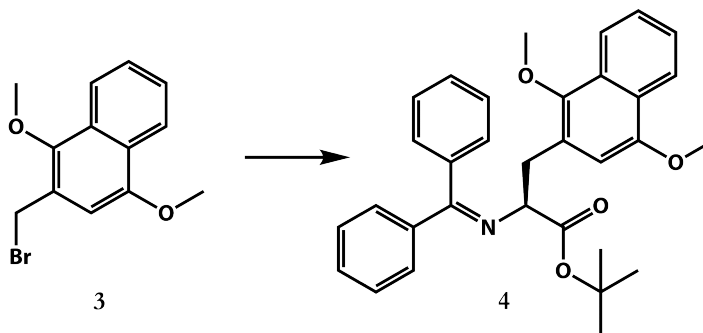
(For alternative method see Bulbule *et al.*⁴) Methyl naphthoate **2** (33.9 g, 138 mmol) in 200 mL anhydrous THF was added dropwise to a stirred suspension of LiAlH₄ (15.7 g, 414 mmol, 3 eq) in 100 mL anhydrous THF at 0°C. The reaction was allowed to warm to room temperature and was stirred for 2 hours. The reaction was quenched by the dropwise addition of 700 mL of saturated sodium potassium tartrate and stirred overnight. The resulting mixture was diluted with 2L ethyl acetate and was extracted twice with 1L water and once with brine. The ethyl acetate layer was collected and the solvent removed *in vacuo* affording the product hydroxymethyl naphthalene (29.7 g, 99%) as a beige solid. NMR ¹H (500 MHz) d 3.90 (s, 3H), 3.96 (s, 3H), 4.87 (s, 2H), 6.79 (s, 1H), 7.47 (t, 1H, J=7.6), 7.53 (t, 1H, J= 7.3 Hz), 8.02 (d, 1H, J=8.3 Hz), 8.22 (d, 1H, J=8.3 Hz); ¹³C (125 MHz) d 55.93, 61.20, 62.87, 104.05, 122.02, 122.68, 125.75, 126.54, 126.94, 128.70, 128.78, 147.20, 152.43; Expected MS: MH+-H₂O 201.1, Found 201.1.

A.3.4 1,4-Dimethoxy-2-bromomethylnaphthalene (**3**)



(For alternative method see Bulbule *et al.*⁴) Hydroxymethyl naphthalene (34.6 g, 159 mmol) was dissolved in 200 mL CCl_4 and cooled to 0°C . PBr_3 (22.4 mL, 238 mmol, 1.5 eq) in 70 mL CCl_4 was added dropwise to this solution. The reaction was stirred for 30 minutes at 0°C and for 1.5 hours at RT. The reaction was quenched by the dropwise addition of saturated aqueous sodium bicarbonate until evolution of CO_2 ceased. The mixture was diluted with 2L CH_2Cl_2 and extracted three times with 1L saturated aqueous sodium bicarbonate, and then with brine. Collection of the organic layer and removal of the solvent *in vacuo* yielded the product aryl bromide **3** (44.6 g, 95.8%) as a tan solid. This product was unstable and was used immediately in the next reaction.

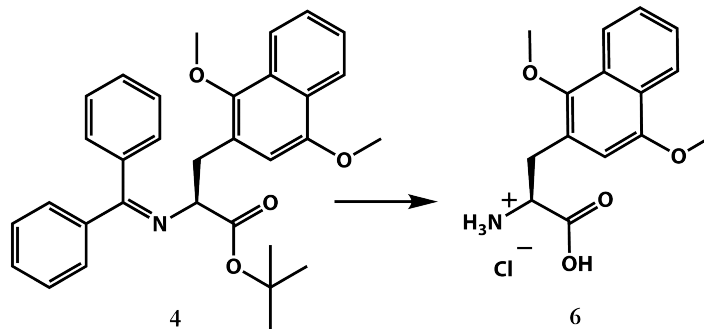
A.3.5 N-(Diphenylmethylene)-L-1,4-dimethoxy-2-naphthalanine tertbutyl ester (**4**)



N-(Diphenylmethylene)-glycine tert-butyl ester (36.0 g, 122 mmol), O-allyl-N-(9-anthracenylmethyl)cinchonidinium bromide (7.38 g, 12.2 mmol, 0.1 eq), $\text{CsOH}\cdot\text{H}_2\text{O}$ (204.54

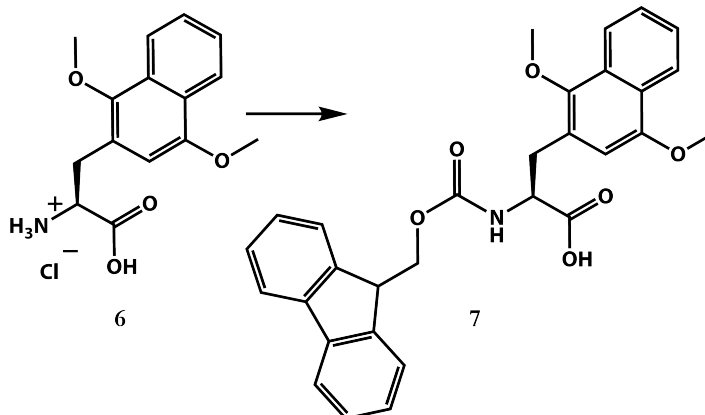
g, 1220 mmol, 10 eq) were suspended in toluene at $-20\text{ }^{\circ}\text{C}$ (ethylene glycol/dry ice bath) with overhead stirring. To this suspension was added dropwise aryl bromide **3** (36.0 g, 125 mmol, 1.02 eq) in 200 mL toluene. The reaction was stirred until TLC analysis indicated completion (eluent 9:1 hexanes/ethyl acetate with 1% triethylamine) after 6 hours. The reaction was poured into 1.5 L H_2O and partitioned between the aqueous layer and 1 L diethylether; a precipitate of the catalyst formed between the two layers and was collected with the aqueous layer. The organic layer was washed with 1 L H_2O and 1 L saturated aqueous sodium chloride and the solvent was removed *in vacuo*. The collected aqueous layers were extracted three times with 500 mL methylene chloride to dissolve the cinchonidinium catalyst which was further purified by recrystallization from methylene chloride/ethylether. The collected organic solids were dry loaded onto silica in methylene chloride with 1% triethylamine. After chromatography (silica, eluent 19:1 hexanes/ethyl acetate with 1% triethylamine followed by 9:1 hexanes/ethyl acetate with 1% triethylamine) 51.85 g (86%) product **4** was obtained as a slightly yellow solid. NMR ^1H (500 MHz) δ 1.47 (s, 9H), 3.31 (dd, 1H, $J_1=10\text{ Hz}$, $J_2=13.5\text{ Hz}$), 3.49 (dd, 1H, $J_1=3.5\text{ Hz}$, $J_2=13\text{ Hz}$), 3.69 (s, 3H), 3.70 (s, 3H), 4.31 (dd, 1H, $J_1=3.5\text{ Hz}$, $J_2=9.5\text{ Hz}$), 6.46 (s, 1H), 6.51 (s, 1H), 7.07 (t, 2H, $J=7.5\text{ Hz}$), 7.23 (t, 1H, $J=7.5\text{ Hz}$), 7.27 (t, 2H, $J=7.5\text{ Hz}$), 7.34 (t, 1H, $J=7.5\text{ Hz}$), 7.42 (t, 1H, $J=7.5\text{ Hz}$), 7.49 (t, 1H, $J=7.5\text{ Hz}$), 7.59 (d, 2H, $J=7\text{ Hz}$), 7.93 (d, 1H, $J=8\text{ Hz}$), 8.15 (d, 1H, $J=8.5\text{ Hz}$); ^{13}C (125 MHz) δ 28.38, 34.06, 55.61, 62.35, 67.25, 81.42, 107.24, 122.14, 122.50, 125.12, 125.80, 126.54, 126.68, 127.92, 128.19, 128.38, 128.98, 130.38, 136.40, 139.77, 148.12, 151.28, 170.37, 121.29; Expected MS: MH^+ 496.2, Found 496.2.

A.3.6 L-1,4-dimethoxy-2-naphthalanine hydrochloride (6)



To 14.4 g (29 mmol) of the protected amino acid **4** was added with stirring 10 mL (119 mmol, 4 eq) ethane dithiol (Sigma-Aldrich) and 60 mL trifluoroacetic acid in quick succession. The reaction was stirred for two hours at which point it was diluted to 2L with water and extracted with 2L 9:1 Hexanes/Ethyl acetate to avoid extraction of the free amino acid into the organics. The organic layer was extracted with 1L 1.5 M aqueous HCl until it no longer appeared to have free amino acid by TLC (eluent 9:1 Hexanes/Ethyl acetate; free amino acid remains at the origin and fluoresces a characteristic blue under UV illumination). The aqueous layers were combined and solvent was removed *in vacuo*. The resulting solid was redissolved in 1.5 M HCl and the solvent was once again removed *in vacuo* affording 8.9 g (98.2%) of the product **6** as a white solid. NMR in D₂O ¹H (500 MHz) d 3.27 (dd, 1H, J₁=14.5 Hz, J₂=8 Hz), 3.42 (dd, 1H, J₁=14 Hz, J₂=5.5 Hz), 3.82 (s, 3H), 3.92(s, 3H), 4.31 (t, 1H, J=6.5 Hz), 6.73 (s, 1H), 7.50t, 1H, J=7.5 Hz), 7.58 (t, 1H, J=7.5 Hz), 7.98 (d, 1H, J=8.5 Hz), 8.11 (d, 1H, J=8 Hz); ¹³C (125 MHz) d 31.19, 54.30, 56.11, 62.24, 106.43, 121.97, 122.18, 122.92, 126.04, 126.49, 127.63, 128.08, 147.72, 152.05, 121.98; Expected MS: MH⁺ 276.1, Found 276.1.

A.3.7 Fmoc-Naq(OMe/OMe)-OH (7)



11.3 g (35.7 mmol) of the amino acid hydrochloride **6** was dissolved in 145 mL DMF. To this stirred solution was added 3 g (35.7 mmol, precisely 1.0 eq) of sodium bicarbonate. The reaction was heated to 60°C and N-(9-Fluorenylmethoxycarbonyloxy)succinimide (13.2 g, 39 mmol, 1.1 eq) was added. After stirring at 60°C for 3h, the reaction was poured into 2L water and extracted twice with ethyl acetate 1L. The organics were combined and extracted twice with 2L water and once with 1L brine. The organics were collected and the solvent removed *in vacuo*. The crude product was dry loaded onto silica for chromatography. After purification (silica column chromatography, eluent 2:1 hexanes/ethyl acetate with 0.1% acetic acid, followed by 5% methanol in methylene chloride), the product Fmoc-Naq(OMe/OMe)-OH was obtained as an off-white solid in an 85% yield (15.1 g). NMR ¹H (500 MHz) d 3.33 (d, 2H, J=6.5 Hz), 3.91 (s, 3H), 3.92 (s, 3H), 4.15 (t, 1H, J=7.5 Hz), 4.24 (m, 2H), 4.61 (q, 1H, J=6.5 Hz), 6.13 (d, 1H, J=7 Hz), 6.62 (s, 1H), 7.11 (t, 1H, J=7.5 Hz), 7.18 (t, 1H, J=7.5 Hz), 7.33 (t, 2H, J=7.5 Hz), 7.46 (m, 3H), 7.56 (t, 1H, J=7.5 Hz), 7.69 (m, 2H), 8.04 (t, 1H, J=8 Hz), 8.23 (d, 1H, J=8 Hz); ¹³C (125 MHz) d 32.00, 32.71, 37.06, 47.28, 55.63, 55.99, 62.31, 105.71, 120.17, 122.14, 122.87,

124.35, 125.46, 125.77, 127.15, 127.30, 127.91, 128.60, 141.49, 143.96, 144.13, 148.02, 152.65, 156.77, 163.58, 175.98; Expected MS: MH+ 498.2, Found 498.2.

A.4 Naq Deprotection Methods

A.4.1 Deprotection of heptaNaq(OMe/OMe) with Cerium Ammonium Nitrate

HeptaNaq(OMe/OMe) was dissolved in enough 0.1 M aqueous TFA to make a solution of about 1 mM as determined by absorbance. 2.1 equivalents of cerium(IV) ammonium nitrate (CAN) was added in an equal volume of 0.1 M aqueous TFA and the solution briefly vortexed to mix. The mixture was allowed to sit on ice for 1 minute, at which point it was diluted to 5 mL with 0.1M TFA. The reaction was incubated on ice for an additional 20 minutes before the deprotected peptide was purified on HPLC. HeptaNaq (PDPQPDQ) was obtained after removal of the solvent *in vacuo* on a lyophilizer taking care not to let the sample thaw by adding additional ddH₂O to dilute the MeCN. Immediately upon drying, solid de heptaNaq was dissolved in ddH₂O, aliquoted into stocks, and flash frozen

A.4.2 Deprotection of P2Naq(OMe/OMe) with Dichlorodicyanoquinone (DDQ)

P2Naq(OMe/OMe) was dissolved in 200 μ L TFA and placed on ice. The concentration of the peptide in this solution was determined quickly by absorbance (ϵ_{301} 5065) and exactly 1 equivalent of DDQ in MeCN (stock ~ 9 mg mL⁻¹) was added with shaking. The solution becomes dark upon DDQ addition and slowly lightens over the course of at least 30 minutes on ice. Dilution of the reaction with ddH₂O is followed immediately by

HPLC purification. After purification samples are treated identically to those of deprotected heptaNaq, above.

A.4.3 Deprotection of Naq(OProp/OProp) in BT6.4

BT6.4 with sidechain protected Naq was dissolved in 200 μ L DMF and transferred to a small round bottom containing a small stir bar. 100 μ L of phenylsilane was added. The solution was frozen with LN₂ and excess (~10 mg) solid Pd(PPh₃)₄ was added on top of the frozen solution. The mixture was evacuated and thawed under vacuum to roughly remove the oxygen from the sample. After backfilling the reaction vessel with argon, the reaction was stirred for 2 hours before being diluted with a mixture of acetonitrile and water. The catalyst immediately precipitated and was removed by centrifugation. The supernatant was purified by HPLC. Purified samples were treated identically to those of purified heptaNaq and P2Naq, above.

A.5 Uncommon Characterization Methods

A.5.1 Determination of Enantiomeric Excess of the Synthesis of Naq

Enantiomeric excess was determined using a Daicel Crownpak CR(+) analytical HPLC column. A solution of **6** in 15% (v/v) aqueous methanol with 0.1% TFA was applied to the column and eluted using the same solvent. The eluent was monitored at 240 nm. Integration of the peaks resulting from the two enantiomers was used to determine the ee of 93%.

A.5.2 Determination of Molar Absorptivity of heptaNaq

Dry heptaNaq(OMe/OMe) was dissolved in pH 7.4 buffer (100 mM sodium phosphate) at various concentrations; absorbance spectra of solutions known dilution of these samples were taken. The concentrations of nitrogen in the original solutions and the buffer were determined by Kjeldahl Analysis at Galbraith Laboratories, Inc. The molar absorptivity (ϵ) at 280 nm is $\sim 3070 \text{ M}^{-1} \text{ cm}^{-1}$ and at λ_{max} (301 nm) is $\sim 5065 \text{ M}^{-1} \text{ cm}^{-1}$. A similar procedure was used to determine the molar absorptivity of deprotected P2Naq.

A.5.3 Spectroelectrochemistry

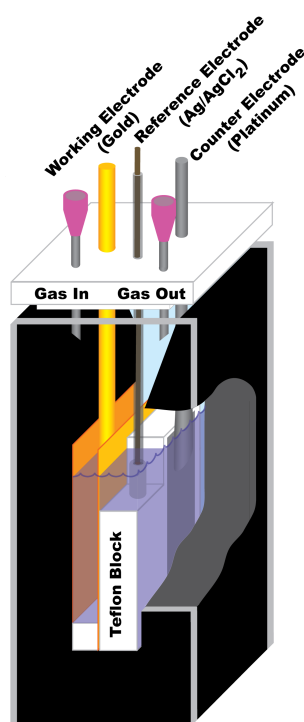


Figure A.1. Schematic of spectroelectrochemical cell built by José Cerda. The light path passes through the gold slides at left.

A schematic representation of the electrochemical cell that was used for the spectroelectrochemistry experiments is depicted in Figure A.1. The cell consisted of two 2 x 1 cm gold slides, a Ag-AgCl electrode (Mircoelectrodes Inc.) and a Pt slide, used as working, reference and counter electrodes, respectively. The gold slides were in electrical contact via a gold wire placed in between them and located on the top cell while a piece of Teflon (1.0 mm thick) was used on the bottom. The Pt slide was separated from the working gold by a Teflon block with a thickness of 2.0 mm. A Pt wire with a flattened tip was squeezed in between the Pt slide and the Teflon block and used as an electrical lead. The tip of the Ag-AgCl microelectrode was held on top of the Teflon block as close as possible to the surface.

The typical sample volume that was used in the cell was 200 -300 mL. Most of the sample

was located in between the gold slides while the remaining solution made contact with the other two electrodes. The electrochemical cell was positioned in the spectrometer so that the UV/vis light beam was centered in between the two gold slides containing the majority of the sample.

Spectroelectrochemical potentials were set using a BAS 100B workstation potentiostat (Bioanalytical Systems, Inc., W. Lafayette, IN). To prevent protein aggregation on the gold surface of the working electrode, the slides were coated with 1-mercapto-6-hexanol. Gold slides are polished with 0.05 mm alumina and rinsed with deionized water, followed by soaking in an ethanol solution containing 2 mM of 1-mercapto-6-hexanol for a period of 24 hours. During the measurements, the solution was kept under oxygen-scrubbed argon atmosphere. The samples all contained approximately 50 μ M peptide diluted from the stocks using buffered solutions at a variety of pHs and 5 mM each of the mediators indigo-trisulphonate (ITS; $E_{m, \text{pH } 7.0}$ -90 mV), pyocyanine (PYO; $E_{m, \text{pH } 7.0}$ -34), and phenazine ethylsulphonate (PES; $E_{m, \text{pH } 7.0}$ 55 mV). Spectra were collected on a Perkin Elmer Lambda 35 UV/Vis spectrometer in the 230 to 600 nm range at 20 °C. For each potential step, various spectra were acquired until no change was further observed, i.e. until equilibrium was reached; experiments were performed in both the reductive and oxidative directions with each sample with no apparent irreversibility over the pH range examined. The time interval between potential steps ranged anywhere from 5 to 20 minutes. The normalized difference in absorbance between 243 nm and 252 nm as a function of redox potential was fit to eq. 1 with Kalaidegraph 4.0 (Synergy Software, Reading PA), derived from the Nernst equation⁵

$$\text{Frac. Reduced} = O + \frac{\Delta}{(1 + 10^{((E_b - E_m)/(58/2))})}$$



Figure A.2. New spectroelectrochemical cell. See text for details.

where O is an offset (approximately 0); Δ is the total change in fraction reduced (approximately 1); E_h is the applied potential; E_m is the measured midpoint potential; 58 mV is the thermal factor at 20 °C; 2 is the number of electrons (allowing the number of electrons to float did not significantly improve the fits). Solutions at a variety of pHs all contained 100 mM KCl and 20 mM respectively of citric acid (pH 5.22), MES (pH 5.97), PIPES (pH 6.79), HEPES (pH 7.52), and CHES (pH 8.33).

A.5.4 New Spectroelectrochemical Cell

While the cell in Figure A.2 was not used to collect data presented in this work, it nevertheless represents some advance over the previously described cell. It is significantly easier to construct and maintain under an inert atmosphere for low potential observations. The cell consists of a custom quartz cuvette designed to minimize the volume of the analyte solution while allowing use of the full beam width of Varian Cary 50 Spectrophotometers. The electrodes are constructed into a Teflon backed silicon septum which inserts into a standard GL14 cap. The working electrode is a large homemade platinum mesh (0.7 cm*1.5 cm), the counter electrode is a coiled platinum wire separated from the analyte solution by a vycor glass frit and the reference is a Ag-AgCl electrode (Mircoelectrodes, Inc.). The solution can be stirred from the side with a conventional or rare-earth stir bar without disruption of the measuring beam that passes along the bottom of the sample volume.

A.5.5 Cyclic Voltammetry of heptaNaq in DMF

Dried fully deprotected heptaNaq was dissolved in anhydrous dimethylformamide to approximately 300 mM. Tetrabutylammonium hexafluorophosphate was added to 100 mM. Cyclic voltammetry experiments were performed in an anaerobic glovebox using a BAS 100B workstation potentiostat. A Pt electrode (1.6 mm diam) was used as the working electrode, a Ag wire was reference, and a Pt wire as the auxiliary electrode. Internal referencing was to ferrocene.⁶ At each scan rate, at least five scans were performed to ensure stability of the system.

A.6 Chemical Ligation Methods

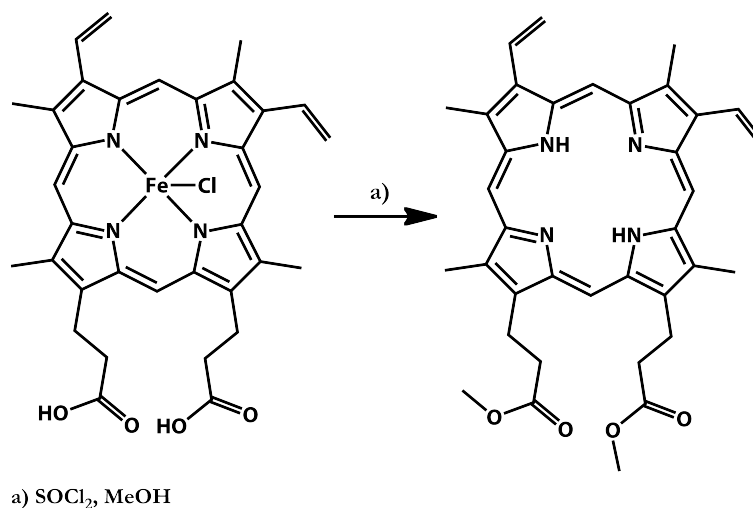
Chemical ligation followed the protocol described by Johnson et al.⁷ Briefly, after expression of the BTX-intein fusion and standard purification over a Co-NTA Talon Resion Affinity column, the BTX-intein fusion was dialyzed into 4 L pH 9 buffer (20 mM CHES, 300 mM KCl) twice and concentrated to a few mL. The BTX-mensa thioester was captured by adding solid sodium mercaptoethanesulfonate to 50 mM to the concentrated sample of the BTX-intein fusion. This reaction was briefly evacuated to eliminate some of the air above the sample and reacted, without stirring, for ~20 hrs under argon at room temperature. After the reaction was complete, the BTX-mensa thioester was purified on a Superdex 75 size exclusion column equilibrated to pH 7.5 (10 mM K₂PO₄). To eliminate some of the salt from the samples, fractions containing the BTX-mensa thioester were concentrated and diluted against water. Lyophilization gave the dry BTX-mensa powder with a small amount of salt contaminant. The thioester and peptide containing an N-terminal cysteine residue (like BT0.4) were individually dissolved in a minimal volume of denaturing

buffer (6 M GdnHCl, 200 mM K₂PO₄, pH 7.5) and their concentrations determined by absorbance using the proper molar absorptivities of the chromophores at 280 nm (Naq 3065, Trp 5690). The thioester and cysteine peptide were then combined in a 1:2 molar ratio and solid TCEP, to 50 mM, and mercaptophenylacetic acid, to 200 mM, were added. The solution was then brought to the reaction pH (7.5) with microliter volumes of 3 M NaOH. The solution was then flash frozen and freeze-pump-thawed three times to remove residual oxygen from the reaction vessel. Back filling with argon, the reaction was stirred for ~24 hours at room temperature and purified on HPLC. It was sometimes necessary to add DTT to the solution prior to HPLC to ensure that there were no disulfide crosslinked species in solution.

A.7 Iron(III) 2,4-Dicyanodeuteroporphyrin IX Synthetic Methods

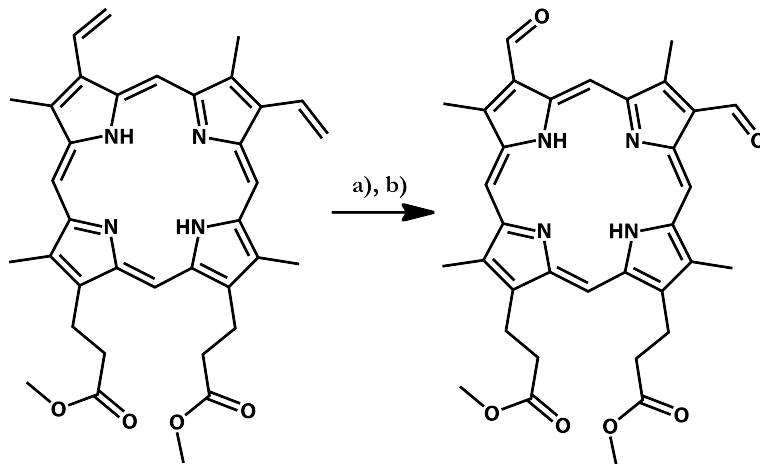
The synthesis of iron(III) 2,4-dicyanodeuteroporphyrin IX described in this section is based upon the general approach of Caughey *et al.*⁸ and La Mar *et al.*⁹ While the dicyanodeuteroporphyrin was not used in the work described in this dissertation, its synthesis proved necessary for a variety of other projects involving heme-binding maquettes, which may be reported elsewhere. The changes to the original syntheses take advantage of recent developments in the selective modification of the porphyrin ring system resulting in a synthesis that allowed a much higher throughput of material.

A.7.1 Protoporphyrin IX dimethyl ester



Hemin (2.5 g) was dissolved in 10 mL pyridine and diluted to 150 mL with methanol. 6.0 g ferrous sulfate was added. Using a dropping funnel, to this stirred reaction, 25 mL of thionyl chloride was added quickly enough to cause the reaction to reflux without external heating. The reaction vessel was left open to the atmosphere until the thionyl chloride addition was complete to prevent over pressurization, which can cause the vessel to shatter. After the addition of the thionyl chloride the reaction was stirred under argon for 30 minutes allowing it to cool to room temperature at which point argon was bubbled through the solution for an additional 30 minutes to remove excess HCl and other volatile acids. It is imperative that the workup be completed as quickly as possible as the free base is sensitive to aqueous acid. The reaction was quenched with 25 mL H₂O and diluted with 300 mL chloroform. The chloroform solution was washed with 300 mL H₂O. The water wash was extracted three times with 100 mL chloroform and the organics combined and further washed with 300 mL H₂O six times followed by neutralization with 300 mL 1 N NH₄OH twice. After drying with magnesium sulfate, the organics were removed *in vacuo*. The product protoporphyrin IX dimethyl ester was confirmed by UV/*vis* spectroscopy and ¹H NMR.

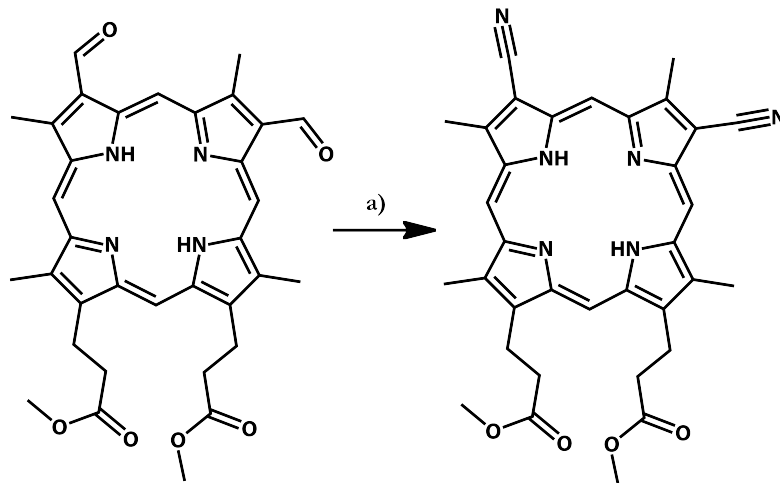
A.7.2 2,4-Diformylprotoporphyrin IX dimethyl ester¹⁰



a) K_2OsO_4 , NMO, H_2O , Dioxane b) Periodic acid, H_2O , THF

The synthesis of 2,4-diformylprotoporphyrin IX dimethyl ester reported by Kahl *et al.*¹⁰ was used with modification. In the first reaction the volatile and toxic osmium tetroxide was replaced by the slightly less toxic and non-volatile potassium tetroxide; otherwise both reactions and workups were identical to that reported.

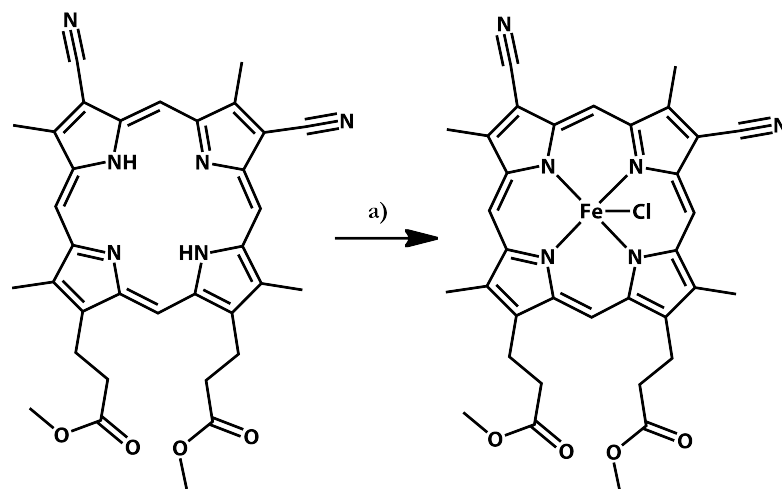
A.7.3 2,4-Dicyanoporphyrin IX dimethyl ester



a) Hydroxylamine hydrochloride, POCl₃, Pyridine, 80 °C

2,4-Diformylprotoporphyrin IX dimethyl ester (388 mg) was dissolved in 100 mL pyridine to which was added 10 equivalents of hydroxylamine hydrochloride (450 mg). After heating at 80 °C for 30 minutes an excess of POCl₃ (7 mL) was added dropwise. The reaction was allowed to stir at this temperature for an additional 3 hours at which point the reaction appeared complete by alumina TLC (eluent chloroform). The reaction was cooled to room temperature and quenched by dropwise addition of saturated aqueous NaHCO₃. The product was extracted into chloroform which was dried with magnesium sulfate and removed *in vacuo*. Purification was effected using an alumina column (eluent CHCl₃/ClCH₂CH₂Cl 8:2) followed by precipitation from chloroform/methanol twice. The product (253 mg, 65%) was confirmed by MALDI-TOF MS and IR spectroscopy.

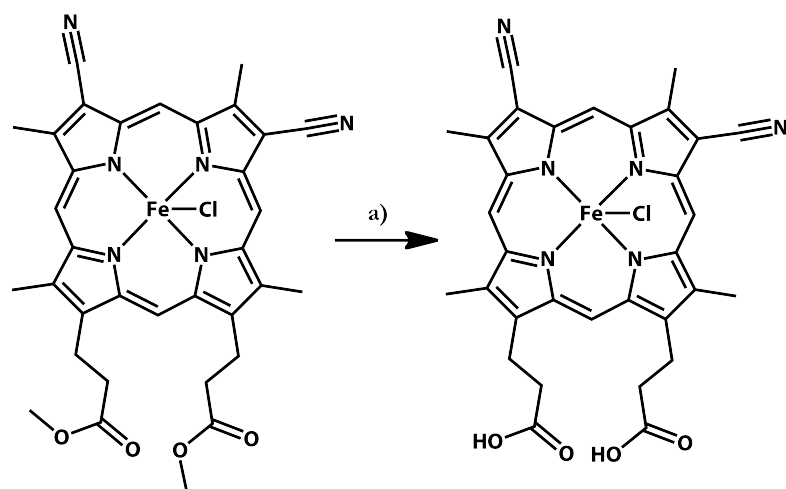
A.7.3 Iron(III) 2,4-Dicyanoporphyrin IX dimethyl ester



a) FeCl_2 , CH_2Cl_2 , MeCN , $50\text{ }^\circ\text{C}$

Iron(III) 2,4-Dicyanoporphyrin IX dimethyl ester was synthesized following the iron insertion procedure of Smith *et al.*¹¹ Better degassing of the reactants was effected by several freeze-pump-thaw cycles. Incomplete reactions as determined by TLC were recommitted to the reaction to increase yield with no obvious ill effects.

A.7.4 Iron(III) 2,4-Dicyanoporphyrin IX



a) Lithium iodide, THF, 90 °C

Removal of the methyl esters from iron(III) 2,4-dicyanoporphyrin IX dimethyl ester was effected by refluxing this compound in THF in the presence of saturated lithium iodide for 48 hours; reaction progress was checked by alumina TLC (eluent CHCl_3 with 0.1% MeOH). The reaction was quenched by the addition of 200 mM aqueous HCl and the product was extracted into a methylene chloride. This extraction was eased by addition of small portions of methanol to the extraction until the product became soluble in the organic layer. The organics were washed once more with 200 mM aqueous HCl followed by brine. It was important that the organics not be dried using magnesium or sodium sulfate as this was associated with the formation of intractable aggregates. The organics were removed *in vacuo*. Attempts at crystallization failed and the product was deemed pure enough by UV/*vis* spectroscopy for use in the maquettes.

A.8 References

- (1) Lichtenstein, B. R.; Cerda, J. F.; Koder, R. L.; Dutton, P. L. *Chem Commun (Camb)* **2009**, 168.
- (2) Boger, D. L.; Jacobson, I. C. *J. Org. Chem.* **1990**, *55*, 1919.
- (3) Newman, M. S.; Choudhary, A. R. *Organic Preparations and Procedures International* **1989**, *21*, 359.
- (4) Bulbule, V. J.; Koranne, P. S.; Munot, Y. S.; Borate, H. B.; Deshpande, V. H. *Synthetic Communications* **2003**, *33*, 587.
- (5) Clark, W. M. *Oxidation/Reduction Potentials of Organic Systems*; The Williams and Wilkins Co.: Baltimore, 1960.
- (6) Prince, R. C.; Dutton, P. L.; Bruce, J. M. *Febs Letters* **1983**, *160*, 273.
- (7) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640.
- (8) Caughey, W. S.; Alben, J. O.; Fujimoto, W. Y.; York, J. L. *Journal of Organic Chemistry* **1966**, *31*, 2631.
- (9) Lamar, G. N.; Viscio, D. B.; Smith, K. M.; Caughey, W. S.; Smith, M. L. *J. Am. Chem. Soc.* **1978**, *100*, 8085.
- (10) Kahl, S. B.; Schaeck, J. J.; Koo, M. S. *Journal of Organic Chemistry* **1997**, *62*, 1875.
- (11) Smith, K. M.; Fujinari, E. M.; Langry, K. C.; Parish, D. W.; Tabba, H. D. *J. Am. Chem. Soc.* **1983**, *105*, 6638.