University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2016

Engineering Principles Of Photosystems And Their Practical Application As Long-Lived Charge Separation In Maquettes

Zhenyu Zhao
University of Pennsylvania, zzyhorde@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Biochemistry Commons, and the Biophysics Commons

Recommended Citation

Zhao, Zhenyu, "Engineering Principles Of Photosystems And Their Practical Application As Long-Lived Charge Separation In Maquettes" (2016). *Publicly Accessible Penn Dissertations*. 2667.
https://repository.upenn.edu/edissertations/2667

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2667
For more information, please contact repository@pobox.upenn.edu.

Engineering Principles Of Photosystems And Their Practical Application As Long-Lived Charge Separation In Maquettes

Abstract

Light-activated electron transfer reactions between cofactors embedded in proteins serve as the central mechanism underlying numerous biological processes essential to the survival and prosperity of most organisms on this planet. These processes range from navigation, to DNA repair, to metabolism, and to solar energy conversion. The proper functioning of these processes relies on the creation of a charge-separated states lasting for a necessary length of time, from tens of nanoseconds to hundreds of milliseconds, by the arrays of cofactors in photosystems. In spite of decades of experiments and theoretical frameworks providing detailed and extensive description of the behavior of the photosystems, coherent and systematic understanding is lacking regarding the underlying structural and chemical engineering principles that govern the performance of charge-separation in photosystems, evaluated by the fraction of the input energy made available by the photosystem for its intended function. This thesis aims to establish a set of engineering principles of natural and man-made photosystems based on the fundamental theories of electron transfer and the biophysical and biochemical constraints imposed by the protein environment, and then to apply these engineering principles to design and construct man-made photosystems that can excel in charge-separation while incurring minimal cost in their construction. Using the fundamental theories of electron transfer, this thesis develops an efficient computational algorithm that returns a set of guidelines for engineering optimal light-driven charge-separation in cofactor-based photosystems. This thesis then examines the validity of these guidelines in natural photosystems, discovering significant editing and updating of these guidelines imposed by the biological environment in which photosystems are engineered by nature. This thesis then organizes the two layers of engineering principles into a concise set of rules and demonstrates that they can be applied as guidelines to the practical construction of highly efficient man-made photosystems. To test these engineering guidelines in practice, the first ever donor-pigment-acceptor triad is constructed in a maquette and successfully separates charges stably for >300ms, establishing the world record in a triad. Finally, this work looks ahead to the engineering of the prescribed optimal tetrads in maquettes, identifying what's in place and what challenges yet remain.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biochemistry & Molecular Biophysics

First Advisor
Peter L. Dutton

Keywords
Charge separation, Electron transfer, photosynthesis, protein design, Solar energy conversion

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2667

Subject Categories
Biochemistry | Biophysics

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2667

ENGINEERING PRINCIPLES OF PHOTOSYSTEMS AND THEIR PRACTICAL APPLICATION AS LONG-LIVED CHARGE SEPARATION IN MAQUETTES

Zhenyu Zhao

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

2016

Supervisor of Dissertation

____

P.Leslie Dutton, PhD, FRS

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

Graduate Group Chairperson

______

Kim A. Sharp, PhD
Associate Professor of Biochemistry and Biophysics

Dissertation Committee:

E. James Petersson, PhD., Associate Professor of Chemistry
Bohdana Discher, PhD., Research Associate Professor of Biochemistry and Biophysics
Brian Chow, PhD., Assistant Professor of Bioengineering
Walter Englander, PhD., Jacob Gershon-Cohen Professor of Biochemistry, Biophysics, & Medical Science
Kim Sharp Ph.D., Associate Professor of Biochemistry and Biophysics
Robert Stanley, PhD., Professor of Chemistry at Temple University

ENGINEERING PRINCIPLES OF PHOTOSYSTEMS AND THEIR PRACTICAL APPLICATION

AS LONG-LIVED CHARGE SEPARATION IN MAQUETTES

COPYRIGHT

2016

Zhenyu Zhao

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

https://creativecommons.org/licenses/by-nc-sa/3.0/us/

Dedication page

To love, and all of those who love generously

ACKNOWLEDGMENT

It is very difficult to imagine completing this thesis without the amazing amount of support, help, and care I have had the honor and fortune to receive over the past five years as a student in the Dutton lab, in the BMB program, and at the University of Pennsyvalnia.

First of all, I would like thank my thesis advisor, my mentor, my supervisor, my artist in resident, and one of the most supportive and caring figures I have ever had the honor of having in my life, P. Leslie Dutton. He is much more than the professor to whom I am a mere student working on a project. He is the source of the most brilliant inspiration in both science and my personal life. I cannot accurately count how many moments of Eureka there had been simply because of a seemingly drawn-out conversation between us. Despite the differences in age and appearances, Les and I are rather alike in many ways. I will only mention the most important one here. We both love arts immensely, he a practitioner and a producer and I more of a consumer, but I think this common appreciation of a creation process that’s the organic complement to science really allowed us to approach science together from a unique and revealing perspective. It’s also something that I will take and keep with me for very very long time.

While Les inspires, Christopher C. Moser helps and teaches, in the most detailed and nuanced way that it is impossible to not learn something useful in every interaction with him, and I never was good at staying away from any opportunity to benefit from his great intellect and uncanny sense of the right approach to a scientific problem. Chris has been instrumental to my entire thesis project in every way possible. He saw the engineering of photosystems as an optimization problem of a function describing the charge-separated state; he knew that SVD-based model fitting can sift through the overlapping mess of spectra in the triad experiments. Chris’s ability to help is not just mere assistance, but rather, he urges you on, gives you the necessary tools, literally and figuratively, to tackle not just the immediate problem, but all the ones behind it.

There would be no Dutton lab without the extreme amount of care and love that Bohdana Discher showers us with. She organizes weekly lab meeting and ensures we are well-fed during these. She makes sure we are working as hard as we can to stay on-schedule to graduation. She cares about us the same way she cares about people she loves. this care and love are really helpful for us students, when any time a failed experiment could mean your months of hard work had been for nothing, and this care and love is rare and hard to come by, but Bohdana is there with it, full of it. I am thankful to have her in the Dutton lab, and on my committee, and in the future.

Since we are on the topic of care and love that's has been so crucial in getting me through grad school, there is no other person in this world who had given me more care and love throughout the course of my career at Penn both as an undergrad and a graduate student than Glennis Logsdon. She embodies love, both in the romantic and altruistic senses, and love is indeed palpable when you are around her. The genuine, true happiness that radiates out of her is probably a good indicator of it, but stay a while longer, you can sense it from her every word and every action.

Life also would have been a lot different, a lot worse, had there not been the amazing friendships in and outside of lab. I met Bryan Fry on the bus trip to Les' house during the recruitment weekend. It was snowing and the bus was slow in traffic. That gave us time to talk, and we did, and we did even more afterwards, and we are still talking today, about a lot of different things. Chris Bialas and Josh Mancini, they joined the Dutton lab at the same time as I did, and they really have made this journey a unique experience. Science would have been a lot less fun without them, and maybe a tad bit more productive? But that's okay, that bit of productivity is made up by Nathan Ennist, who is one of the hardest working and most intelligent graduate students I know. I have him to thank for the maquette that made the triad possible, and he deserves every bit of it. High resolution crystal structure doesn't happen everyday, and it had never happened for a monomer maquette, until him. I wish them the best and all the luck in the world. Godspeed friends!

My thesis committee, James Petersson, Brian Chow, Bohdana Discher, Walter Englander, have been there with me since the beginning of this thesis project, and have seen all the blunders I have made and all the misdirections I have followed, and for each of those have offered everything they can to get me back on track, and it’s an understatement to say that this work would have been complete without them.

Last but absolutely not least, my family, my parents and grandparents, the people dearest to my heart, who have been standing by me since the very beginning and have never wanted anything from me than seeing me smile, seeing me happy. Thanks and appreciations alone are not enough for their love for me, but at least it’s good to know that I will always be there for them as well.

ABSTRACT

ENGINEERING PRINCIPLES OF PHOTOSYSTEMS AND THEIR PRACTICAL APPLICATIONS IN MAQUETTES

Zhenyu Zhao

P.Leslie Dutton

Light-activated electron transfer reactions between cofactors embedded in proteins serve as the central mechanism underlying numerous biological processes essential to the survival and prosperity of most organisms on this planet. These processes range from navigation, to DNA repair, to metabolism, and to solar energy conversion. The proper functioning of these processes relies on the creation of a charge-separated states lasting for a necessary length of time, from tens of nanoseconds to hundreds of milliseconds, by the arrays of cofactors in photosystems. In spite of decades of experiments and theoretical frameworks providing detailed and extensive description of the behavior of the photosystems, coherent and systematic understanding is lacking regarding the underlying structural and chemical engineering principles that govern the performance of charge-separation in photosystems, evaluated by the fraction of the input energy made available by the photosystem for its intended function. This thesis aims to establish a set of engineering principles of natural and man-made photosystems based on the fundamental theories of electron transfer and the biophysical and biochemical constraints imposed by the protein environment, and then to apply these engineering principles to design and construct man-made photosystems that can excel in charge-separation while incurring minimal cost in their construction. Using the fundamental theories of electron transfer, this thesis develops an efficient computational algorithm that returns a set of guidelines for engineering optimal light-driven charge-separation in cofactor-based photosystems. This thesis then examines the validity of these guidelines in natural photosystems, discovering significant editing and updating of these guidelines imposed by the biological environment in which photosystems are engineered by nature. This thesis then organizes the two layers of engineering principles into a concise set of rules and demonstrates that they can be applied as guidelines to the practical construction of highly efficient man-made photosystems. To test these engineering guidelines in practice, the first ever donor-pigment-acceptor triad is constructed in a maquette and successfully separates charges stably for >300ms, establishing the world record in a triad. Finally, this work looks ahead to the engineering of the prescribed optimal tetrads in maquettes, identifying what's in place and what challenges yet remain.

TABLE OF CONTENTS

ACKNOWLEDGMENT IV

ABSTRACT VII

LIST OF TABLES AND SCHEMES XV

LIST OF ILLUSTRATIONS XVI

PREFACE XXII

PART I: FORMULATION OF COMPREHENSIVE ENGINEERING GUIDELINES FOR ELECTRON-TRANSFER BASED PHOTOSYSTEMS IN BIOLOGICAL CONTEXT VIA COMPUTATIONAL MODELING AND OPTIMIZATION 1

CHAPTER I: THEORETICAL BASIS OF MOSER-DUTTON RULER BASED ELECTRON TRANSFER ENGINEERING GUIDELINES 2

1.1 Introduction: 2

1.2 Electron transfer formulated as a non-adiabatic tunneling phenomenon in solution. 3

1.3 Modeling the nuclear term of the Fermi's expression for electron transfer 6

1.4 Empirical simplification of the ET rate equation without loss of generalizability 12

1.5 Successful applications of the Moser-Dutton rate expression 16

1.6 Conclusions 16

1.7 References 17

CHAPTER II: ENGINEERING BLUEPRINTS FOR OPTIMAL DESIGNS OF PHOTOACTIVATED CHARGE SEPARATION IN ET BASED SYSTEMS: RULES AND PRINCIPLES 20

2.1 Introduction: 20

2.2 Abstracting photosystems to cofactors and their roles: the *n-ad* paradigm. 23

2.3 Performance parameters as metrics for optimality of design 25

2.4 Analytical paradigm and the mathematical infrastructure of the formulation of the engineering blueprints 27

2.5 Comprehensive engineering blueprints of optimal charge-separating dyads 30

2.6 Engineering blueprints for optimal triads. 40

2.7 Engineering blueprints of optimal photochemical tetrads 44

2.8 Equivalence of donors and acceptors in triads and beyond 49

2.9 Increasing tolerance for design parameters as n-ads become more complex 50

2.10: effect of the lifetime of the excited state on the engineering of photosystems 55

2.11: Conclusion: Central engineering principles for optimal charge-separating photosystems based on the fundamental theories of electron transfers. 56

2.12: Computational methods 58

2.13: References 60

CHAPTER III: THE UNIVERSAL CATALYTIC QUARTET AND ITS SUBVERSION AND REWRITING OF THE ENGINEERING BLUEPRINTS OF CHARGE-SEPARATING PHOTOSYSTEMS 62

3.1: Introduction: 62

3.2 Results and Discussion: 71

3.2.1 Rhodobacter Sphaeroides reaction center as the reference model system par excellence for illustrating the catalytic quartet model 71

3.2.2 Resolving the apparent paradox within the initial electron transfer rates in R.Sphaeroides RC 75

3.2.3: A combination of unrelaxed and fully relaxed protein dielectrics as the only appropriate model for experimentally observed ET kinetics. 77

3.2.4 Uniformly high reorganization energy is a central engineering feature universal to natural photosystems 81

3.2.5 Natural photochemical quartets rely on incomplete dielectric relaxation, ultra-fast excited state lifetime, and closely placed initial cofactors to minimize the lifetimes of the intermediate CS states. ... 87

3.2.6: tetrads have significant engineering advantage over triads for effective long-lived charge-separation at high reorganization energies. 91

3.2.8: Protection against the potential damage by long-lived intermediate charge-separated states is a central engineering constraint in optimal designs of photosystems. 98

3.2.9 Revised engineering blueprints for optimal photosystems prescribe designs highly similar to known natural photosystems 100

3.2.10: four cofactors are the minimal requirement for high performance charge-separating photosystems when the engineering dogma of damage protection is included 103

3.2.11: Natural catalytic quartet anticipates membrane potential effect with suboptimal cofactor energetics 105

3.3 Conclusions: 110

3.4 Computational methods... 111

3.5 References 112

PART II: EXPERIMENTAL APPLICATION AND TESTING OF OPTIMAL ENGINEERING BLUEPRINTS IN MAQUETTE PLATFORM...................................... 115

Preface:.. 115

CHAPTER IV: MAQUETTE AS THE IDEAL VERSATILE PLATFORM FOR ENGINEERING OF PHOTOCHEMICAL CHARGE-SEPARATING DEVICE............... 117

4.1: The necessity of de-novo designed protein as engineering platform for optimal photosystems ... 117

4.2: Overview of major approaches in de-novo protein design. 119

4.3 The Maquette approach: origin and development 124

4.4 Maquettes as the ideal testing platform of design blueprints and engineering guidelines for photochemical multi-ads for charge-separation... 130

4.5 Structured, stable, versatile, and adaptable: The newest generation of maquette with crystal structure 131

4.6 Conclusion:.. 132

4.6 References: 134

CHAPTER V: TOWARDS A PROOF-OF-PRINCIPLE PHOTOCHEMICAL TRIAD CHARGE-SEPARATING DEVICE IN MAQUETTE... 136

5.1 Introduction: 136

5.2 Methods:... 138

5.2.1: In-vivo incorporation of C-type heme in maquettes expressed in E.Coli 138

5.2.2: in-vivo incorporation of propargyl tyrosine in maquettes. 139

5.2.3: Click-chemistry between the azido-containing label and the PPY-containing maquette............ 139

5.3 Results: 140

5.3.1 Engineering of an extended maquette to accommodate three redox/ET active cofactors 140

5.3.2 Engineering of a single-his binding site for anchoring of Zn-porphyrin and other Zn-based pigment cofactors 141

5.3.3 Engineering of CXXCH motif for in-vivo attachment of C-type heme for cofactor site-specificity 143

5.3.4 Non-covalently ligated and electrochemically poised DADPIX as the electron donor in the triad 145

5.3.5 Incoporation of unnature amino acids as site-specific covalent anchoring point for additional cofactors 147

5.3.6 Click chemistry based covalent incorporation of cofactors into maquettes 149

5.4.1: Challenges of bringing everything together: consistent failures in simultaneous application of multiple methods presented in this chapter. 149

5.4.2: Sky's the limit: maquette approach's first-principled, functionally neutral nature makes it susceptible to any type of functionalization. 151

5.5 Conclusion 153

5.6 References 155

CHAPTER VI: ENGINEERING OF WATER-SOLUBLE CHARGE-SEPARATING DONOR-PIGMENT-ACCEPTOR(DPA) TRIADS IN A HIGHLY STRUCTURED 4-HELICAL SCAFFOLD PROTEIN TO ACHIEVE WORLD RECORD FOR LONGEST LIFETIME OF CHARGE-SEPARATION AND TO TEST THE THEORETICAL DESIGN BLUEPRINTS OF GENERAL CHARGE-SEPARATING DEVICES 156

6.1 Introduction: 156

6.2: Methods 159

6.2.1 covalent modification of maquette via cysteine-maleimide reaction 160

6.2.2: Fluorescence emission spectroscopy: 160

6.2.3: Redox and electrochemical measurements: 160

6.2.4: Transient absorption spectroscopy of the monad, dyad, and triads: 160

6.2.5: SVD-based global analysis and kinetics model fitting of the transient absoprtions data 161

6.3 Results: 162

6.3.1: High resolution X-ray crystal structure of a novel generation of highly stable and cofactor-specific maquette as the appropriate template for DPA triad design. 162

6.3.2: Engineering landscape of the DPA triad using the structured maquette as design platform. 164

6.3.3 Incorporation of ferrocene as the donor cofactor into the donor slot of the maquette scaffold at near and far donor positions 168

.... 180

6.3.7 Electron transfer and charge separation kinetics of the P-A dyad in the scaffold maquette 181

6.3.8: Electron transfer and charge-separation kinetics of the canonical ferrocene-ZnP-DADPIX triad. 187

6.4: Discussion 197

6.4.1 Agreements and discrepancies between the model and empirical kinetics and their implications. 197

6.4.2 The positives and negatives of the triplet excited state in practical charge-separation engineering. 198

6.4.2: Routes to improve the current DPA triad design 200

6.4.2.1 Constitutive donor cofactors with the right potential. 200

6.4.3.2 Acceptor with appropriate energetics to allow for singlet-driven charge separation 201

6.5: Conclusion 202

6.6 References 203

CHAPTER VII: APPLICATION OF THE UPDATED OPTIMAL ENGINEERING GUIDELINES TO THE CONSTRUCTION OF TETRADS IN A NEW GENERATION OF MAQUETTES 204

7.1 Introduction: 204

7.2: Prescribing an optimal tetrad meeting all practical constraints in the new maquette 204

7.3: Engineering the maquette towards the optimal tetrad. 206

7.4: Looking even further ahead: the greatest unresolved challenges 208

7.5: References: 210

APPENDIX 211

A.1: Maquette sequences engineered during the process towards constructing the DPA triad 211

A.2: Mathematical details. **212**

A.3: Common experimental methods.. **213**

A.4: redox spectra of cofactors used in this thesis... **216**

LIST OF TABLES AND SCHEMES

SCHEME 2.1 : ANALYTICAL STEPS TO OBTAIN THE COMPLETE KINETIC INFORMATION OF A ET-BASED PHOTOSYSTEM AS A TIME-DEPENDENT FUNCTION OF ALL ENGINEERING PARAMETERS. 29

TABLE 3.1: EXPERIMENTAL AND MODELED RATES OF INITIAL ELECTRON TRANSFER OF SPHAEROIDES RC. 76

SCHEME 4.1: LONG-LIVED CHARGE-SEPARATION IN SYNTHETIC SUPRAMOLECULAR TRIADS AND TETRADS. 117

TABLE 5.1: SOME REPRESENTATIVE SEQUENCES OF GEN1 MAQUETTES 139

TABLE 6.1: DISSOCIATION CONSTANTS AND EMS, ALONG WITH OTHER PROPERTIES, OF THE TETRAPYRROLES 159

TABLE 6.2: TESTING THE PERFORMANCE OF SVD-BASED MODEL FITTING UNDER VARIOUS NOISE LEVELS 177

SCHEME A1: SYNTHESIS OF AZIDO-FERROCENE FROM METHYL HYDROXYL FERROCENE. 213

LIST OF ILLUSTRATIONS

Figure 1.1: Electronic and nuclear components of electron transfer. 3

Figure 1.2: Non-adiabatic and adiabatic nuclear potential surfaces of electron transfer reaction participants. 5

Figure 1.3: The reorganization energy of an electron transfer reaction and the inverted region. . 6

Figure 1.4: Hopfield's quantized model of electron transfer successfully explains the low-temperature behavior of reaction centers. 8

Figure 1.5: Comparison of classical Marcus and semi-classical Hopfield rate-driving force parabolas.. 10

Figure 1.6: Various models' abilities to accurately fit experimental ET rates. 11

Figure 1.7: A single characteristic frequency is capable of accurately describing the electron transfer rates in biological context 14

Figure 1.8: Common log-linear dependence of electron transfer rates on distances between the cofactors. 15

Figure 2.1: Multitudes of natural photosystems that rely on electron transfer driven charge separation to achieve essential functions in biology. 21

Figure 2.2: Cofactor-centric, role-based abstract representation of all photosystems and representation of their mechanism 24

Figure 2.3: Performance metrics to evaluate and compare different charge-separating photosystems.. 26

Figure 2.4: Engineering blueprints of photochemical dyads, demonstrating various aspects of the optimal engineering parameters at selected lifetime of charge-separation ranging from ns to megaseconds. 32

Figure 2.5: Indirect charge-recombination via uphill thermal repopulation of preceding, higher-lying energy states and the calculation of its rate. 35

Figure 2.6: Trikes for optimal dyads at representative lifetimes. The energy levels and rates between the states are shown. Forward ET rates are shown in green arrows while recombinations are shown in red 37

Figure 2.7: Optimal engineering blueprints for PAA triad at selected lifetimes from nano- to kiloseconds for photon energies from 1.0eV to 3.0eV. 39

Figure 2.8: Electron transfer details of optimal triads at various selected lifetimes. 43

Figure 2.9: Optimal engineering blueprints of charge-separating tetrads at selected lifetimes under different incident photon energies at 0.7eV reorganization energy, demonstrating the following characteristics of the optimal tetrads 45

Figure 2.10: Engineering details of optimal charge-separating tetrads at uniform 0.7eV reorganization energy and 1.8eV incident photon energy. 47

Figure 2.11: Isomorphism of various configurations of triads and tetrads. 48

Figure 2.12: Tolerances for design parameters in dyads and triads. 52

Figure 2.13: Engineering tolerances for the design parameters of the second and final acceptors in a PAAA tetrad 53

Figure 2.14: Effects of the excited state lifetimes on the optimal engineering efficiencies of photochemical devices 54

Figure 3.1: Natural photosystems cover a broad range of the solar radiation spectrum. 63

Figure 3.2: The universal PAAA organization of the functional cores of natural photosystems of various species. 65

Figure 3.3: Universally conserved catalytic quartet in photosystems across organisms. 66

Figure 3.4: Structural details of the electron transfer active cofactors of Sphaeroides reaction center and the distances separating them 71

Figure 3.5 Predicted electron transfer kinetics of Sphaeroides catalytic quartet using uniform reorganization energies 74

Figure 3.6: Rates of initial electron transfers in R.Sphaeroides at various uniform and mixed reorganization energies compared to experimentally measured rates 79

Figure 3.7: Marcus ET kinetics of initial electron transfer in natural photosystems described by both unrelaxed and relaxed dielectric protein environments at different timescales. 80

Figure 3.8: PSII electron transfer kinetics demonstrate unrelaxed low reorganization energy behavior. 81

Figure 3.9: Sphaeroides reaction center specifically engineered for high performance at uniformly high reorganization energies. 86

Figure 3.10: Photosystem I/II adopted generic water-soluble protein environment characterized by high uniform reorganization energy. Left: Structural and energetic details of Sphaeroides reaction center according to the quartet model. Right: Quantum yield and engineering efficiencies at increasing reorganization energies of three types of tetrads. Green: fully optimal tetrads under the constraint of curved geometry and the unrelaxed lambda of 0.3eV. Red: semi-optimized tetrad where the distances are held constant but driving forces are allowed to vary 87

Figure 3.11: ET distance-rate model with unrelaxed low reorganization energy considered. 88

Figure 3.12: Effects of reorganization energy on the engineering of tetrads from the perspective of varying lifetimes of charge-separation 90

Figure 3.13: Effects of increasing reorganization energy on dyads, triads, and tetrads 93

Figure 3.14: Effects on the engineering efficiencies of abstract multi-ads viewed from the reorganization energy perspective, with fixed selected lifetime of 100ms 94

Figure 3.15: The curving of the cofactors of natural photosystem to favor direct recombination from the final charge-separated state to the pigment. A: the rates of direct RC from QA to P at various effective curvature of the reaction center, representing a simple one-angle bending of the P-QA distance. Red line represents the rate of the QA to P if the RC was arranged linearly like in B. Green line represents the experimentally observed rate. B: linear and curved models of the catalytic quartet. 95

Figure 3.16: The curving of natural photosystems to accelerate direct recombination from Pheophytin to the ground state. The rates of direct RC from H_B to P at various effective curvature of the reaction center, representing a simple one-angle bending of the P-H distance. Red line represents the rate of the H_B to P if the RC was arranged linearly........ 96

Figure 3.17: Optimal tetrads under the revised engineering guidelines display high similarity to natural photosystems. 101

Figure 3.18: tetrads are the minimal engineering requirement for high performance when protection against damage is included as an essential engineering constraint. 104

Figure 3.19: Quantum yield’s dependence on the membrane potential experienced by the Sphaeroides reaction center. Reproduced from (41) 105

Figure 3.20 The shut-down effect imposed upon the reaction center of Sphaeroides by increasing membrane potentials. 106

Figure 3.21 Effects of the membrane potential on the performance of Spaheroides reaction center 108

Figure 3.22: Effects of membrane potentials on the details of electron transfer kinetics in the Sphaeroides reaction center. 109

Figure 3.23: Shut-down effect of higher membrane potential on the Sphaeroides reaction center. 110

Figure 4.1: The iterative processes of directed evolution. 121

Figure 4.2: Iterative process of computational protein design using Rosetta3........ 122

Figure 4.3: The highly non-uniform distribution of probability masses for the optimal sequences in the protein sequence space. 123

Figure 4.4: A 2-dimensional map illustrating the relative positions of major protein design approaches........ 124

Figure 4.5: Cofactors successfully incorporated into maquettes categorized by their chemical properties........ 127

Figure 4.6: The iterative process of the maquette approach. 131

Figure 4.7: X-ray crystal structure of the water-soluble maquette MZH3 at a resolution of 2.02 Å. 133

Figure 5.1: Electron-transfer based charge-separating dyads in previous generations of short maquettes produced by the Dutton group........ 137

Figure 5.2: example sequences of Gen1 single-chain maquettes with binding sites for hemes. 141

Figure 5.3: Further functionalization of Gen2 maquette to include Zn-porphyrin binding site and C-heme incorporation capability 142

Figure 5.4: Gen2 maquette's ability to bind both Iron- and Zn-porphyrins with high affinities .... **Error! Bookmark not defined.**

Figure 5.5: Spectral electrochemical measurement of the redox potential of the ground state of the ZnP 144

Figure 5.6: In-vivo incorporation of C-heme in the long BT6 maquette. 144

Figure 5.7: Redox midpoint potential of C-type heme in Gen2 maquette determined by two independent methods. 145

Figure 5.8: Full assembly of the first DPA triad in a maquette using DADPIX-ZnP-C-heme as the cofactors, with ascorbate poised DADPIX as the donor. 147

Figure 5.9: chemical characterization of propargyl tyrosine using Left: 148

Figure 5.10: Incorporation of propargyl tyrosine into maquettes. 152

Figure 5.11: Incorporation of azido-ferrocene in maquettes via click-chemistry. 153

Figure 6.1: Available tetrapyrrole cofactors successfully incorporated into maquettes in past studies. 157

Figure 6.2: Crystal structures of photochemical maquette 163

Figure 6.3: Photophysical properpties of the ZnP pigment cofactor for the triad design. 165

Figure 6.4: Design guidelines of the DPA triad in the maquette 167

Figure 6.5: Rotamer preferences and their impacts on the effective donor-pigment distances in the G164C(near) and Y168C(far) variants of the DPA triad designs 171

Figure 6.6: covalent incorporation of ferrocene into the maquette via maleimide-cysteine linkage 172

Figure 6.7: Maquette's high affinity for the acceptor heme cofactor and the pigment zinc-porphyrin cofactor after the incorporation of ferrocene. 173

Figure 6.8: Redox properties of the acceptor and donor cofactors of the DPA triad. 174

Figure 6.9: Assembly of the DPA, DADPIX-ZnP-Ferrocene triad in the scaffold maquette. 175

Figure 6.10. Synthetic spectra generated to validate the singular value decomposition (SVD) model-fitting algorithm. 176

Figure 6.11: Structural and kinetics model of the P-monad in the maquette scaffold. 178

Figure 6.12: Transient absorption spectra and SVD-based global analysis of the P-monad kinetics. 180

Figure 6.13: Assembly of a P-A dyad in the maquette scaffold using DADPIX and ZnP without modifying the protein with ferrocene. 181

Figure 6.14: Continuous illumination of the P-A dyad in the presence of sacrificial donor 182

Figure 6.15: Transient absorption spectra of the P-A dyad in the maquette. 183

Figure 6.16: SVD-based global analysis of the P-A dyad in the maquette. 184

Figure 6.17: Electron transfer kinetics and charge-separating efficiency of the P-A dyad in the maquette. 185

Figure 6.18: The near- and far- versions of the DADPIX-ZnP-Fc triad. 186

Figure 6.19: Formation of long-lived charge-separated state directly observable from the raw data of the DADPIX-ZnP-Fc triads. 187

Figure 6.20: Direct estimate of yields of the near- and far-versions of the DADPIX-ZnP-Fc triads. 189

Figure 6.21: Time evolution profiles of the third principle components of the near and far variants of the DADPIX triads. 190

Figure 6.22: SVD-based global kinetics analysis of the far(Top row) and near(bottom row) variants of the DPA triad in maquette 191

Figure 6.23: Experimental and theoretical ET kinetics and charge-separation performances of the near variant of the DPA triad. 194

Figure 6.24: Electron transfer kinetics of the far-DPA triad and its yield and lifetime, determined by both SVD-based model fitting and theoretical model using predictive analysis. 196

Figure 7.1: Optimal tetrad in the new generation of maquettes 205

Figure 7.2: Multi-helical bundles to provide curved geometry for maquettes 207

Figure A.1: Solutions to the system of ordinary differential equations for the kinetics model of the DPA triad in Chapter III. 213

Figure A.2: Redox spectra of ferrocene. 216

PREFACE

One of the most exciting and interesting aspects of my thesis, I find, is the theory-practice duality that permeates its entirety and defines its unique nature, which is largely reflected in the organization and presentation of this thesis at a supra-chapter level into two parts. The first part, which constitutes the relatively larger half of this thesis, focuses on the theoretical exploration and discovery of the fundamental engineering constraints and principles that govern the behavior and optimality of photosystems, based in both the fundamental physical theories of electron transfer as well as the need for survival, that is, protection of the photosystem against damages. The first part then formulates a series of design blueprints based on its combined understanding of these constraints and then delivers them to the second part of thesis, where the stage is set for these design blueprints to be carried out in small, water-soluble, man-made proteins. Part II of this work thus details the efforts in establishing the experimental infrastructure, understanding and overcoming the practical constraints involved in the specific engineering environment, and in achieving the first ever and longest ever man-made charge-separation by an electron-transfer based triad system in biological environment.

Although I have separated the thesis into two parts, with Part I involving mostly the theoretical side of the work and Part II the experimental, never is there a chapter or a section in this thesis where the theory or the experiment stands completely in isolation. Each chapter in the more theoretical Part I is either preoccupied with the validation of its theoretical findings in practice or seeks to modify the theories with limits and constraints imposed by the boundaries of practice. Similarly, every experimental result in Part II is evaluated against the theoretical model upon which the analysis relies.

Part I: Formulation of comprehensive engineering guidelines for electron-transfer based photosystems in biological context via computational modeling and optimization

Chapter I: Theoretical basis of Moser-Dutton ruler based electron transfer engineering guidelines

1.1 Introduction:

Electron transfer reactions within, across, and from proteins form the fundamental building blocks of almost all essential biological functions across species, families, and even kingdoms. Involving direct movement of electrons from one molecule to another, it is the basis of respiration that powers the metabolism of every living cell. It empowers the most efficient and clean source of energy of this planet, photosynthesis in both plants and bacteria. It also drives the protection and repair mechanism that safeguards against DNA-damage. As the significance of biological electron transfer cannot be understated, scientific efforts striving towards a complete and straightforward understanding of this process have been consistently surging forward over the last few decades. From the ground-breaking efforts of Chance(*1*) and Devault(*2*), Marcus(*3*-*5*), Hopfield(*6*) that set the theoretical foundation to the seminal recent works by Moser/Dutton(*7*) and Gunner(*8*, *9*), Gray, and Winkler(*10*), we have obtained an in-depth understanding of electron transfer in both inorganic and organic(biological) systems that transformed electron transfer from a scientific novelty to a well-described and readily modeled process. This in-depth understanding has allowed us to reduce the complexity and high computational cost of the complete quantum mechanical theory of electron transfer to an elegant and easy-to-implement semi-classical empirical expression that's arithmetic in nature. This ostensibly simplistic expression, now widely known as the Moser-Dutton ruler, comes out of analysis of large amount of experimental results on ET rates in biological and semi-biological systems(*11*-*13*). This empirical simplification in the presentation of the underlying principles governing electron transfer has opened doors to previously impractical and unviable computational simulations of electron transfer kinetics of natural and man-made photosystems in the fields of photosynthesis, DNA-repair, signal-sensing and transduction, charge-separation, and more. Since systematic application of computational simulation of photosystem kinetics serves as the theoretical foundation of both halves of my thesis, namely the engineering blueprints of natural and man-made photosystems, I have devoted the opening chapter of my thesis to describe and trace its development from the very beginning of systematic formulation of theories to explain the phenomenon of electron transfer until the current forefront of research on electron transfer, where the elegance of the Moser-Dutton ruler has become an essential component of the essential toolset of photosynthetic research.

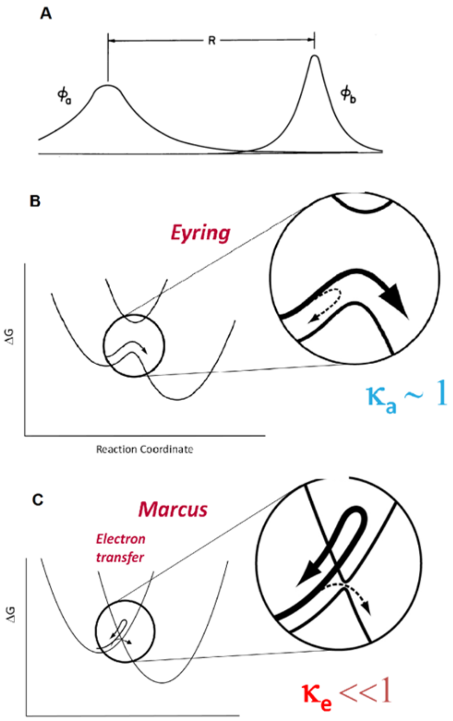


Figure 1.1: Electronic and nuclear components of electron transfer. A: the electronic wavefunction overlap between that of the acceptor and of the donor. The degree of overlap depends on the distance and is seen as the value of the matrix element H_{AB}. Adapted from (Hopfield reference) B: Non-adiabatic crossing of nuclear potential functions of the acceptor and donor, known as the Eyrin limit on electron transfer rates. C: Adiabatic, normal electron transfer with low transition probability given by the FC term of Fermi's Golden Rule.

1.2 Electron transfer formulated as a non-adiabatic tunneling phenomenon in solution.

The phenomenon of electron transfer between an acceptor and a donor molecule is based in electron tunneling across a potential barrier, a fundamental quantum mechanical phenomenon and a popular pedagogical example that had been described in detail applied to electron transfers as early as the 1920s(*14*, *15*).Because of the particle-wave duality of the electron and its long wavelength, the wave function describing the probability density of the electron is able to penetrate a potential barrier, representing the insulating surrounding that separates the electron at its donor from its acceptor. This penetration gives rise to a non-zero probability that the electron will be found at the donor, shown in Figure 1.1A. Unlike traditional chemical reactions, which are largely adiabatic processes described by transition-state theories involving reactants overcoming an activation energy barrier at the crossing point of the potential energy surfaces to form the products(*16*), electron transfer reactions are largely non-adiabatic, in the sense that there does not exist a strong enough coupling between the nuclei of the donor and acceptor to split the potential curves as shown in Figure 1.1C to cause the transition probability from donor to acceptor to be 1, and necessitates an entirely different theoretical framework.(*3*, *17*) To quantatively describe the probability of the electron tunneling from the donor to the acceptor, Fermi's Golden Rule, an expression of fundamental importance to transition from state A to state B in light of quantum mechanics, is invoked. In its original form, Fermi's golden rule is stated as following:

$$W_k = \frac{2\pi}{\hbar}\left|\langle i|H_p|k\rangle\right|^2 \rho_k \qquad [1]$$

, where W_k refers to the transition probability between state *i* and state *k*, h-bar is the adjusted Boltzmann constant, H_p is the perturbation Halmitonian applied to state *k*, while ρ_k represents the density of the target states per unit of energy(*18*).

Fermi's golden rule was derived using time-dependent perturbation theory in quantum mechanics to compute the probability of a transition when a system residing in an initial state is perturbed by a certain potential.(citation for time-dependent perturbation theory, Feyman lectures here). It is stated in the most general form in equation [1], as the expression can be applied to all types of transitions and not just electron transfers. (*19*)

When applied to tunneling based electron transfers, according to the derivation summarized by Devault(*2*), Fermi's golden rule takes on a slightly different form, :

$$r_{ab} = \frac{2\pi}{h}|H_{AB}(r)|^2 FC \qquad [2]$$

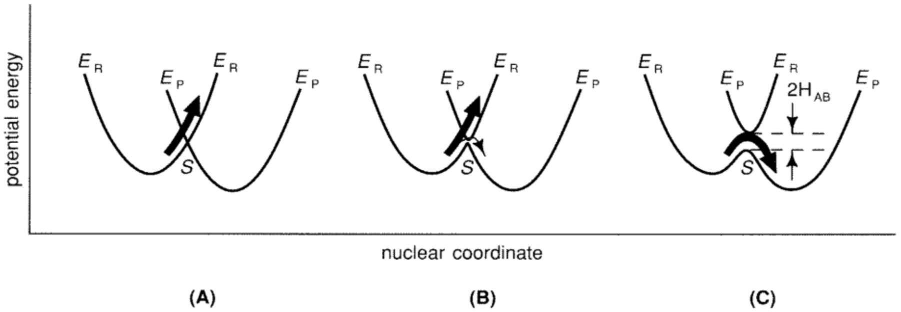


Figure 1.2: Non-adiabatic and adiabatic nuclear potential surfaces of electron transfer reaction participants. A: Reactant and product intersection with 0 transition probability, where no coupling between the electronic wavefunctions of the reactant and product exists. **B:** Non-adiabatic transition represented by small transition probability resulting from weak coupling between the recant and product. **C:** adiabatic reaction characterized by large transition probability from large overlap.

Written as the product of a "matrix element" term(H_{ab}) measuring the interaction between the initial and final electronic state of the transition and the overlap between the vibrational states that are part of the potential energy surfaces(FC) of the acceptors and donors involved in the transition(in our case the electron transfer), Fermi's Golden Rule provides the most direct and comprehensive fundamental mathematical description of the rate of electron transfer between two molecules. It separates contributions to the electron transfer rate into an electronic term and a vibronic(nuclear) term. The electronic term, the matrix element H_{AB} of tunneling, gives the rate of the electron transfer when the nuclei of the donor and acceptors are both in the appropriate locations to effect the electron transfer without violating the Frank-Condon principle(*20*) and conservation of energy. It is measured by the overlap of the electronic wave-functions of the acceptor and donor within the tunneling medium that attenuates exponentially as the distance between them increases, and can be written as the following:

$$|H_{AB}(r)|^2 = V_0^2 \exp(\beta R) \qquad [3]$$

, where the term V^2_0 represents the maximal overlap between the electronic wavefunctions of the acceptor and donor, β is the pre-exponential factor coefficient of the decay of the electronic coupling. In addition to visualizing the matrix element as the overlap of the electronic wave-functions, however, it is helpful to consider H_{AB} as the spacing between the two potential surfaces at their intersection as shown in Figure 1.2.

That is, the greater the overlap between the electronic wave-functions, the larger the coupling between the nuclear surfaces between the reactants of the electron transfer reaction and, consequently, the less non-adiabatic it becomes. Indeed, when overlap integral is large enough, the coupling becomes so great that a break in the potential surfaces takes place and the transition probability at the intersection becomes unity, giving us the behavior of a transition-state driven chemical reaction rather than a tunneling-based ET reaction. Because of the simple correspondence between the overlap of the wave-functions to the matrix element, this model of the electronic term has remained as the consensus model since the very early days of understanding electron transfers in terms of tunneling events(*13*, *21*, *22*).

1.3 Modeling the nuclear term of the Fermi's expression for electron transfer

The nuclear term, on the other hand, represents the probability of bringing the acceptor and donor nuclei to the appropriate location on the reaction coordinate. As already discussed, in the electron transfer/tunneling formulation of Fermi's golden rule, the energy density is represented by the Frank-Condon overlap integral between the nuclear potential energy surfaces of the acceptor and

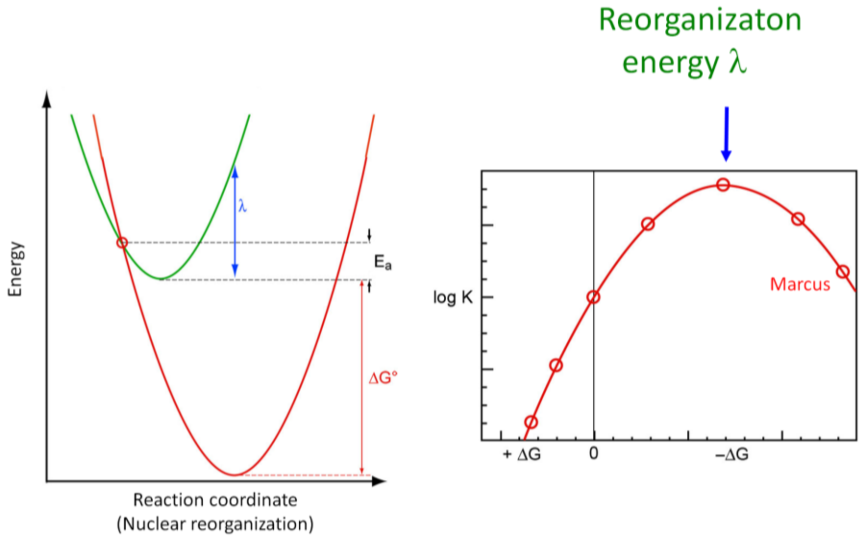


Figure 1.3: The reorganization energy of an electron transfer reaction and the inverted region. Left: the parabolic potential energy surfaces of the reactant and product in the electron transfer reaction, with the driving force ΔG^o, activation energy E_A, and reorganization energy λ, of the reaction shown with appropriate arrows. Right: parabolic dependence of the rate of electron transfer on the driving force of the reaction. The inverted region effect is shown as the decreasing rates past the reorganization energy maximum.

donor, essentially describing the overlap of the vibrational states that make up the potential energy surfaces. The earliest description of the FC term was formulated by Marcus using a combination of the formalisms of both transition state theory and the Landau-Zener formulation of Fermi's "golden rule" formulation.(*3*, *23*) Treating the nuclear potential energy surfaces of the reactant and product as purely classical harmonic oscillators, as seen in Figure 1.3, Marcus determined the effective activation energy barrier analog for the non-adiabatic electron transfer reaction to be:

$$E^{\pm} = \frac{(\lambda - \Delta E)^2}{4\lambda} \qquad [4]$$

This activation energy is then used in the derivation of the Frank-Condon nuclear term. Using the Landau-Zener model of the Fermi's golden rule(*24*), Marcus derived the Frank-Condon contribution to the ET rate in his model from the difference in the slope of the slopes of the reactant and product nuclear potential surfaces as well as the velocity of crossing the intersection of the surfaces.

$$P = \exp\left\{-\frac{2\pi}{\hbar v}\frac{\left|H_{12}\right|^2}{\left|\Delta F_{12}\right|}\right\} \qquad [5]$$

, where v is the velocity of crossing the intersection of the surfaces and is proportional to the activation energy stated in equation [4], and ΔF_{12} is proportional to the reorganization energy. By replacing v and ΔF_{12} in equation [5] with the corresponding terms, Marcus obtained a nuclear term that's shown in equation [6]. Here the FC nuclear term is thus a Gaussian distribution with its characteristic mean and variance dependent on two important quantities: The free energy difference between the acceptor and donor states, and the reorganization energy between the acceptor and donor state:

$$FC_M \propto e^{-\frac{(\lambda - \Delta E)^2}{4\lambda k_B T}} \qquad [6]$$

The mean of the Gaussian, when viewed from the free energy difference's perspective, is the reorganization energy, while the variance is the product between that and k_BT, the ambient temperature thermal energy of the system. Marcus' formulation was purely classical, regarding the nuclear energy surfaces of the reactants as classical harmonic oscillator. One of the most significant and surprising predictions of the Marcus formulation of the electron transfer rate model is the presence of an inverted-region. That is, because of the parabolic dependence of the electron transfer rate on the driving force of the reaction, when the driving force of the electron transfer reaction is equal to the reorganization energy of the system, the rate reaches its maximum, and further increases in driving forces leads to decreases rather than expected increases in the rate of the electron transfer, as shown in Figure 1.3B. The driving forces greater than the reorganization energy are called the inverted-region, which exerts a slowing effect on the rate of electron transfer. The existence of the inverted region, when demonstrated experimentally, had been one of the most convincing evidence in support of Marcus' version of the electron transfer theory. The inverted region indicates that electron transfer reaction can be slowed down by adopting a small reorganization energy and a appropriately large driving force, a device called inverted-region stabilization that would receive significant attention in the upcoming chapters.

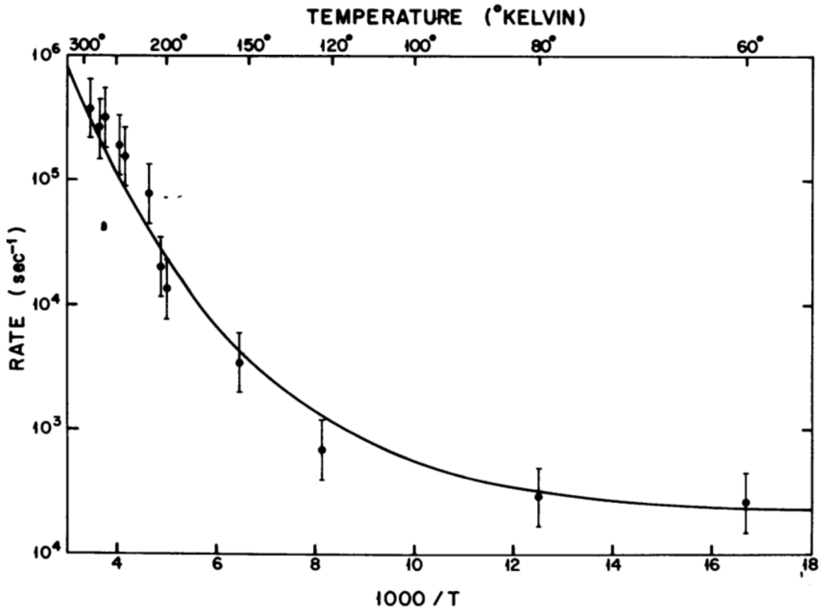


Figure 1.4: Hopfield's quantized model of electron transfer successfully explains the low-temperature behavior of reaction centers. The points plotted are from the experimentally determined ET rates in *Chromatium.* The curve is drawn using the Hopfield semi-classical model of the ET rate. Reproduced from (6)

However, the classical Marcus description of the ET rates soon was discovered to be insufficient in describing ET behaviors at low temperature. Specifically, while the classical Marcus theory predicts continual decreases in ET rates as temperature decreases, the experimentally observed rates begin to level off and stop decreasing as temperature falls below a threshold value. (*1*, *25*) This is seen in Figure 1.4. Consequently, further studies by Hopfield and others have added quantum mechanical corrections that included characteristic vibrational frequency of the nucleus of the acceptor and donor molecules to account for the disappearance of observed temperature dependence of the electron transfer at low temperature.(*26*) Unlike Marcus with his transition state theory approach, Hopfield arrived at his expression by treating electron transfer as an analogue to energy transfer between an acceptor and donor, essentially invoking Fermi's golden rule in the quantum mechanical way, calculating the FC factor term as an integral of two energy functions representing the effects of taking away and adding an electron to and from the participants of the ET reaction respectively. Similar to the electronic absorption and emission spectra used to compute the FC factor in energy transfer and spectroscopy, Hopfield derived the Frank-Condon factor for electron transfer as the integral between the nuclear "spectra" of the reactant and product. I will not go through the details of the derivation, as those who are mathematically inspired can easily follow the elegant flow of Hopfield's thoughts in his article. However, it is worth emphasizing that the FC term in Hopfield's formulation is calculated from a novel and inventive cross integral of probability density functions of various energies of the reactant and product. It is not surprising that the effective result of Hopfield's derivation is a modified version of the Marcus model, as both approaches are based largely in Fermi's golden rule(Marcus used the Landau-Zener formulation to arrive at his expression), in spite of the significant differences in the motivations and points-of-departure in the derivation of these two models. The semi-classical ET rate equation modifies the effective variance of the Gaussian to be dependent on the quantized characteristic frequency of the acceptor and donor molecule as a harmonic oscillator, modulating the width of the Gaussian curve with a cotangent function and thereby resolving the low-temperature disagreements between experimental data and the classically based Marcus theory. At high temperature limits, the temperature dependence behavior predicted by the classical Marcus theory resumes. Hopfield's semi-classical model, as shown in Figure 1.5, modifies the slowing down effect of the inverted region by widening the parabola that describes the rate dependence on the driving force, and indeed represents an improvement in fitting of temperature-dependent experimental data. The inverted-region slowing effect becomes more gradual in the Hopfield model, requiring a much larger driving force past the reorganization energy to achieve the same level of slowing, as seen in Figure 1.5.

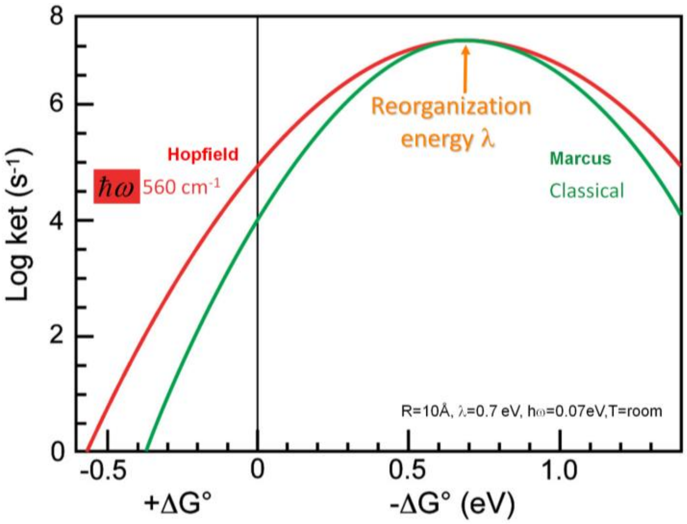


Figure 1.5: Comparison of classical Marcus and semi-classical Hopfield rate-driving force parabolas. Red trace represents the Hopfield and green represents the Marcus model. The remaining parameters used to plot these two traces are in the bottom right corner.

Hopfield's formulation is described as semi-classical because it only incorporates a single quantum mechanical character of the system into the model, the quantized vibrational energy levels of the potential energy surfaces of the reactant and product. Jortner(*27*), Levich, and Dogonadze(*28*) have proposed a fully quantum mechanical model, in which an additional overlap integral between the vibrational wavefunctions of the reactant and product of the electron transfer reaction is included, representing the true Frank-Condon term in the quantum mechanical sense:

$$C(nn') = \int \chi_n \chi_{n'} \; dx \qquad [7]$$

The updated, fully QM version of the FC term is then:

$$FC_{QM} = \sum_{n=0}^{\infty} C^2(n, n+p) e^{-\beta h\omega n}(1 - e^{-\beta h\omega}) \qquad [8]$$

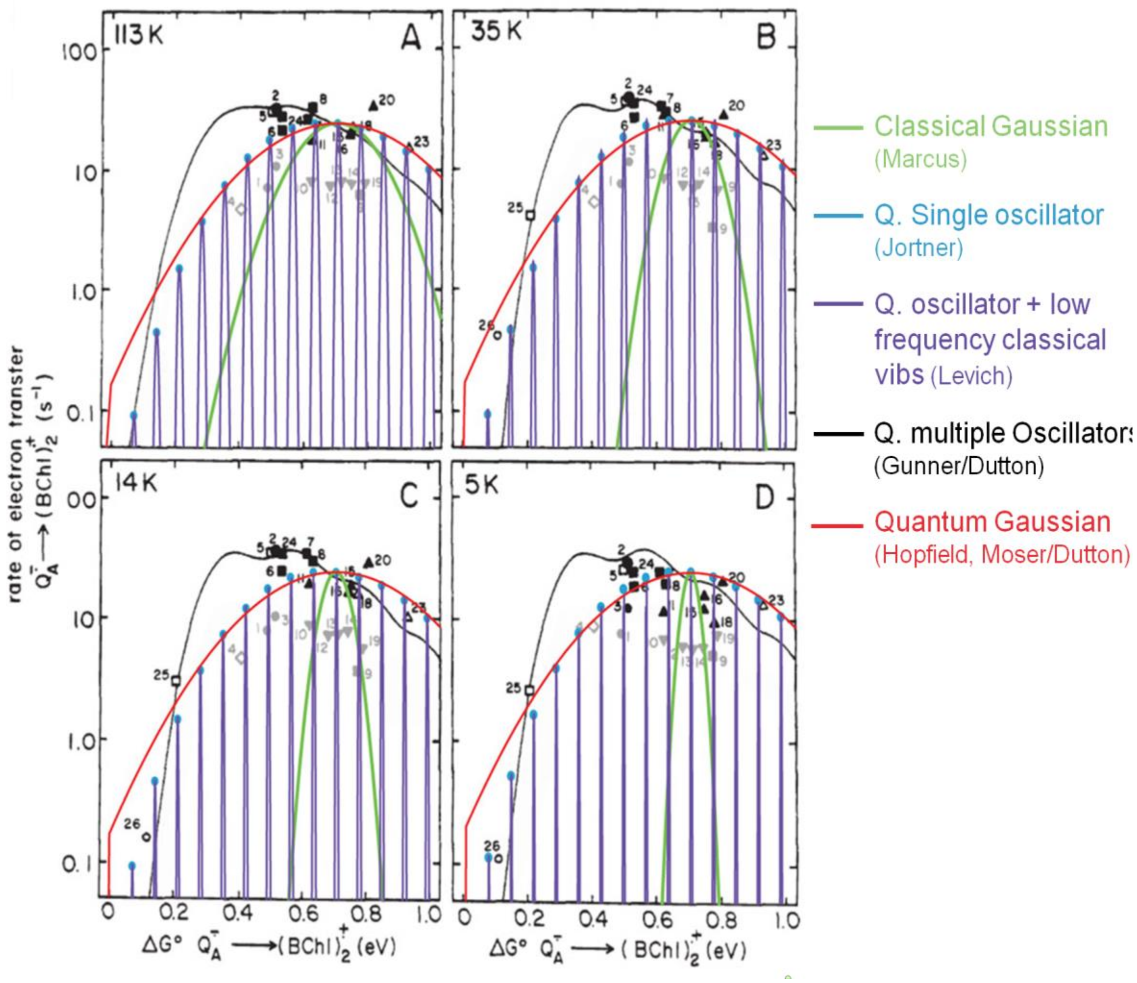


Figure 1.6: Various models' abilities to accurately fit experimental ET rates. Electron transfer rates between the final acceptor of the Sphaeroides reaction center and the ground state of the pigment at various temperatures, along with model fits provided by the models described in this chapter. The relevant models include the classical Marcus Gaussian in green, the quantum mechanical modified Marcus model in red(Hopfield), and the full quantum mechanical model in purprle. The multiple oscillator model is described in the Gunner work but not included in the discussion in this chapter. Figure reproduced and adapted from (*8*)

This model introduces coupling to temperature dependent vibrational levels, modifying the FC term with relative contributions from various vibrational states occupied at certain temperatures. When fully expanded and written out, equation [7] includes a numerically cumbersome and computationally expensive series of summation terms that can be prohibitive in repeated application of the rate equation for modeling. This highly precise model introduces periodic rise and fall into the driving force Gaussian curve but, given the uncertainties involved in the experimental measurements of ET rates, poses great danger of over-fitting when applied as the general model for biological and chemical electron transfers. Moreover, the large number of parameters involved in the expression as well as the discreet summations makes any kind of repeated application of such model difficult and impractical. The complexity of the QM model can be seen in Figure 1.6, where all of the above-described models are used, in addition to the Moser-Dutton empirical approximation of the Hopfield semi-classical expression we will introduce in the next section.

1.4 Empirical simplification of the ET rate equation without loss of generalizability

While the Hopfield semi-classical and the fully QM-based implementation of Fermi's Golden Rule in the context of biological electron transfer have been proven to be an accurate and reliable theoretical model(*29, 30*), its mathematical complexity, in terms of both the large number of parameters involved, use of discreet summation terms with varying lengths depending on the temperature, and the presence of a Bessel function within the expression, makes its application in systematic and repeated computational applications costly and impractical. The goal of this thesis is to formulate comprehensive and practical engineering guidelines for ET-based photosystem. Consequently, we cannot directly apply either the Jortner model or the Hopfield model. Nevertheless, the Hopfield formulation is the perfect point of departure from which a computationally efficient model can be constructed, using experimental rates as the guide for complexity reduction.

In the early 1990s Gray and Winkler conducted an extensive series of experiments in which electron transfer rates between various Ruthenium metal complexes bound to modified Azurin and Cyt C with different driving forces and distances were measured with high accuracy.(*31-33*) Around the same time, Gunner from the Dutton group(*8, 9*) conducted an extensive and encompassing study on the electron transfer between the pheophytin of a bacterial reaction center protein and its natural electron acceptor Qa, along with a large array of substituted quinones of various chemical and therefore redox properties. Additionally, the availability of the X-ray crystal structures of bacterial reaction centers(*34, 35*) led to a wide variety of studies in ET rates of natural photosystems using mutants to adjust distances as well as driving forces between the cofactors. This abundance of experimentally reported electron transfer rates and their accompanying distances and driving forces within protein environment provided the necessary foundation for an empirically based and computationally friendly formulation of the rate expression in equation (2) that allows for ET rate prediction of moderate to high accuracy while drastically reducing the number of parameters required and complexity of computation involved.

The effort towards a more mathematically tractable and pragmatic rate expression began with the simplification of the semi-classical(Hopfield) model of the nuclear term using large number of empirical measurements of electron transfer reaction rates between reactants with known adjustments in driving forces at fixed room-temperature.(*7*) Fitting of experimental measurements of ET rates from natural reaction centers modified with a large variety of quinones and semi-quinones as well as semi-synthetic Ru(II/III) containing ET proteins as shown in Figure 1.7A and B, in addition to fully synthetic ET systems, to the semi-classical model demonstrates a persisting and general independence of electron transfer rates of the characteristic frequency at different temperatures. That is, for ET reactions in natural, synthetic, or non-protein environments between various types of cofactors at room temperature, a single characteristic frequency of $h\omega$=70mV satisfactorily produces models that accurately describe their electron transfer rates. The prevalence of a single characteristic frequency in all of the experimentally determined ET rates strongly suggests that the Gaussian of the nuclear term in the Hopfield ET rate expression can be drastically simplified to depend only on the driving force and the reorganization energy, with the contribution of all other parameters reduced to a computable constant, thereby greatly reducing the computational complexity of modeling the rate of electron transfer in biological context. The extent and elegance of simplification becomes most satisfying when the ET rate is presented in its log form. The exponent disappears and results in a simple, arithmetic nuclear term that varies with respect to the driving force quadratically.

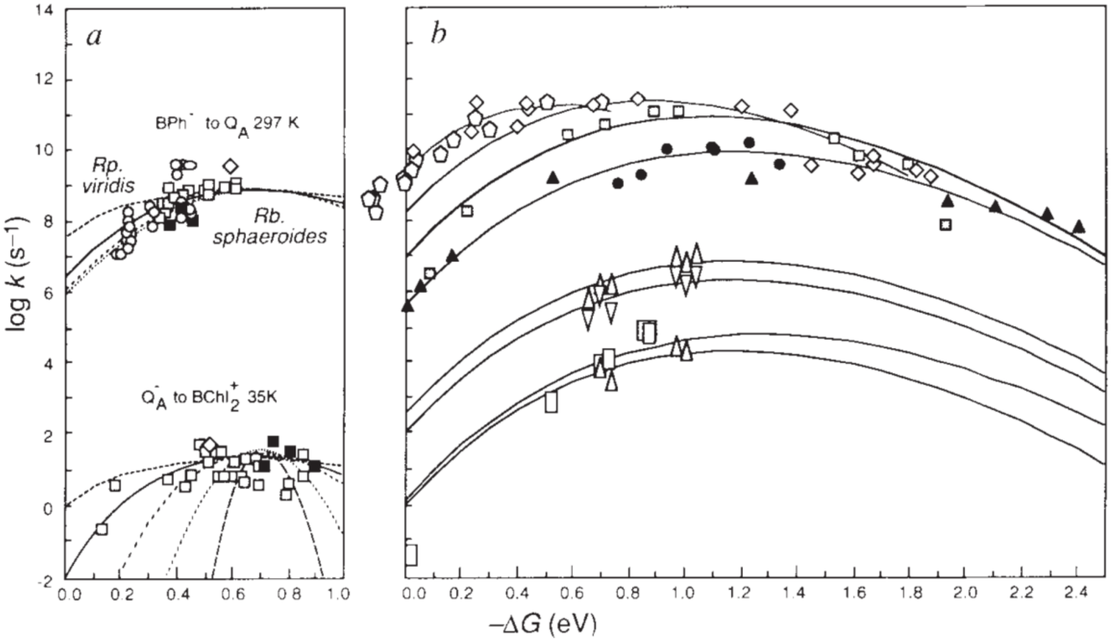


Figure 1.7: A single characteristic frequency is capable of accurately describing the electron transfer rates in biological context. A: electron transfer rates between BPh and Q_A in Sphaeroides with varying driving forces generated by native and non-native quinones, with models fitted with the quantimzed semi-classical Hopfield equation with various characteristic frequency, with the solid trace being hω=70mV. B: Electron transfer rates between semi-synthetic ruthenium-cytochrome, ruthenium-myoglobin, bridged porphyrin-quinone. All models are fitted using hω=70mV with the semi-classical Hopfield model. Figure reproduced from (*7*)

While the nuclear term simplification was achieved via fitting experimentally determined ET rates between reactants of constant distances and known driving forces, the electronic term proves to be more straightforward. Since the Frank-Condon term of the Hopfield rate expression is a Gaussian, its contribution to the ET rate is maximal when the driving force equals the mean of the Gaussian, specifically the reorganization energy of ET reaction. With maximal contribution from the FC term, the remaining parameters that fully describe the ET rate are the distance between the cofactors and the exponential coefficient that describes the "height" of the intervening potential energy barrier. By examining and tabulating the natural logarithm of ET rates against various ET reactions with known distances and known driving forces, as shown in Figure 1.8, Moser et al demonstrated a remarkably consistent linear relationship with a constant slope between the distances and the ET rates across 12 orders of magnitude in various protein-based systems. Moreover, examination of ET reactions with known distances in non-protein environment also demonstrated linear relationship, with a slope different from that of protein ETs. Consequently, Moser(*7*) et al concluded that, without significant loss of accuracy, a universal exponential factor for protein electron transfer exists, and when compared to exponential factors for other tunneling media, protein interior resemble glassy organic solvent. However, a more accurate statement would be that the average of protein interior when acting as the tunneling barrier between cofactors in the context of electron transfer is fairly constant and similar across various protein species. This conclusion allows for an elegant simplification of the electronic term of the Hopfield semi-classical rate expression similar to that of the nuclear term. Together, the respective empirical approximations of the nuclear and electronic terms allow us to write equation [2] in the following form:

$$log k_{et}^{exer} = 13 - \frac{\beta}{2.3}(R - 3.6) - 3.1\frac{(\Delta G+\lambda)^2}{\lambda} \qquad [9]$$

, where 13 is the Eyring limit of rates, obtained from $k_B T/h$ with T=298K, and β is the characteristic coefficient of the intervening medium between the acceptor and donor of the electron transfer(*36*, *37*), 3.6 represents the minimal Van der Waal contact distance cutoff, and 3.1 is the log-transformed numerical coefficient obtained from using the standard 70mV characteristic frequency as shown above.

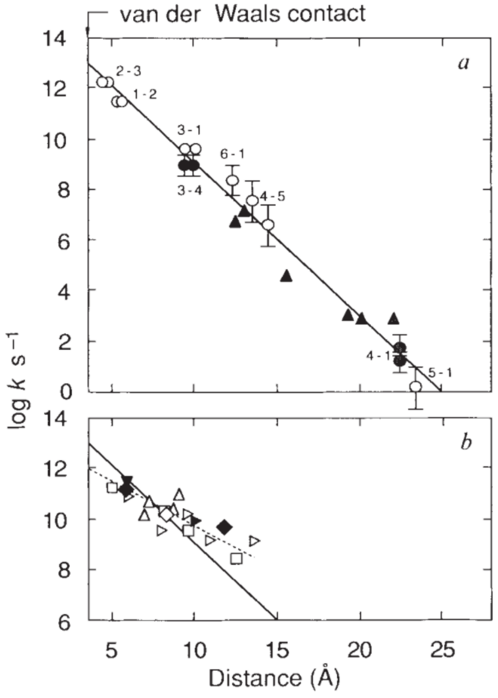


Figure 1.8: Common log-linear dependence of electron transfer rates on distances between the cofactors. Distances and rates are obtained from reported experimental measurements in natural and semi-synthetic ET systems. Figure reproduced from (*7*)

This expression significantly reduces the complexity of electron transfer rate prediction while maintains a moderately high level of accuracy. It has not only allowed for predictions of single rates between pairs of electron-transfer active cofactors, but also, and much more importantly, enabled fast and robust model building for complex electron transfer systems consisting of multiple cofactors.

1.5 Successful applications of the Moser-Dutton rate expression

This expression, dubbed by many as the Moser-Dutton ruler, has seen extensive application since its formulation in 1992 in various experimental and theoretical studies of natural and man-made photosystems as well as electron-transfer driven natural enzymes that are not light-activated. In addition to predicting individual electron transfer rates for experimentalists, it was used by theorists who are interested in modeling ensembles of electron-transfer active cofactors, either for the kinetics of an entire reaction center or for abstract exploration of design principles for electron transfer devices.

On the experimental side, the rate expression is used to model and estimate expected ET rates in photolyaze(*38*, *39*) and cryptochrome, cytochromes(*40*), PS II and I(*12*, *41-43*), and Viridis reaction center(*44*).

On the theoretical side, applications of the Moser-Dutton rate expression was used to formulate some engineering principles that determine whether a particular arrangement of a triad was optimal or not, without having demonstrated either the intent or results to systematically provide a comprehensive engineering blueprints for all photosystems(*45*). The expression was also used to examine photosystems' engineering tolerances for driving forces and distances(*13*), also to explore and discover consistent trends in electron-transfer based natural enzymes. (*11*)

1.6 Conclusions

In the opening chapter of this thesis, I summarized the various theoretical models, with increasing levels of complexities, which describe the rates of electron transfer between any pairs of cofactors. These models are either computationally too costly or insufficient in accuracy to be implemented with high efficiency in computational modeling of ensemble of electron transfer cofactors. Consequently, this chapter also described the seminal efforts by Moser et al. in their formulation of the empirical rate expression that reduces the complexity of modeling the rates of electron transfer to 3 variables without significant loss of generality. With numerous applications that demonstrated its accuracy and elegance, the rate expression known as Moser-Dutton ruler offers the ideal computational infrastructure for systematic analysis that seeks to discover the fundamental principles governing the efficiency of ET-based charge-separation and formulate engineering blueprints for optimal charge-separating photosystems.

1.7 References

1. B. Chance, D. Devault, V. Legallais, L. Mela, *Kinetics of Electron Transfer Reactions in Biological Systems* (1967).
2. D. Devault, Quantum Mechanical Tunnelling in Biological Systems. *Quarterly reviews of biophysics*. **13**, 387 (1980).
3. R. A. Marcus, Chemical and Electrochemical Electron-Transfer Theory. *Annual Review of Physical Chemistry* (1964).
4. H. Sumi, R. A. Marcus, Dynamical Effects in Electron Transfer Reactions. *The Journal of chemical physics*. **84**, 4894–4914 (1986).
5. R. A. Marcus, Electron Transfer Reactions in Chemistry. Theory and Experiment. *Reviews of Modern Physics* (1993).
6. J. J. Hopfield, Electron Transfer Between Biological Molecules by Thermally Activated Tunneling. *PNAS*. **71**, 3640–3644 (1974).
7. C. C. Moser, J. M. Keske, K. Warncke, R. S. Farid, P. L. Dutton, Nature of Biological Electron Transfer. *Nature*. **355**, 796–802 (1992).
8. M. R. Gunner, P. L. Dutton, Temperature and -DeltaG Degree Dependence of the Electron Transfer From BPh to Q_A In Reaction Center Protein From Rhodobacter sphaeroides with Different Quinones as Q_A. *J. Am. Chem. Soc.* **111**, 3400 (1989).
9. M. R. Gunner, Y. Liang, D. K. Nagus, *Variation in Rates of Electron Transfer in Photosynthetic Reaction Centers with the Primary Ubiquinone Substituted with Other Quinones* (1982).
10. H. B. Gray, J. R. Winkler, Electron Transfer in Proteins. *Annual review of biochemistry* (1996).
11. C. C. Moser, J. L. R. Anderson, P. L. Dutton, Guidelines for Tunneling in Enzymes. *Biochimica et Biophysica Acta (BBA)*. **1797**, 1573–1586 (2010).
12. C. C. Moser, P. L. Dutton, in *Photosystem I* (Springer Netherlands, Dordrecht, 2006), vol. 24 of *Advances in Photosynthesis and Respiration*, pp. 583–594.
13. C. C. Page, C. C. Moser, X. Chen, P. L. Dutton, Natural Engineering Principles of Electron Tunnelling in Biological Oxidation-Reduction. *Nature*. **402**, 47–52 (1999).
14. G. Gamow, Zur Quantentheorie Des Atomkernes. *Z. Physik*. **51**, 204–212 (1928).
15. R. W. Gurney, R. H. Fowler, The Quantum Mechanics of Electrochemistry. II. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* (1932), doi:10.1887/0750308397/b1107c6.
16. E. Wigner, The Transition State Method. *Transactions of the Faraday Society*. **34**, 29 (1938).
17. M. Baer, Adiabatic and Diabatic Representations for Atom-Molecule Collisions: Treatment of the Collinear Arrangement. *Chemical Physics Letters*. **35**, 112–118 (1975).
18. H. J. Reittu, Fermi"s Golden Rule and Bardeen"s Tunneling Theory. *American Journal of Physics*. **63**, 940–944 (1995).
19. J. L. McHale, *Molecular spectroscopy* (1999).
20. A. S. Coolidge, H. M. James, R. D. Present, A Study of the Franck-Condon Principle. *The Journal of chemical physics*. **4**, 193 (1936).

21. H. B. Gray, Long-Range Electron-Transfer in Blue Copper Proteins. *Chemical Society Reviews* (1986).
22. S. Fukuzumi, Artificial Photosynthesis for Production of Hydrogen Peroxide and Its Fuel Cells. *Biochim. Biophys. Acta*. **1857**, 604–611 (2015).
23. R. A. Marcus, N. Sutin, Electron Transfers in Chemistry and Biology. *Biochimica et Biophysica Acta (BBA)* (1985).
24. C. Zener, Non-Adiabatic Crossing of Energy Levels. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **137**, 696 (1932).
25. B. Chance, M. Nishimura, On the Mechanism of Chlorophyll-Cytochrome Interaction: the Temperature Insensitivity of Light-Induced Cytochrome Oxidation in Chromatium. *PNAS*. **46**, 19 (1960).
26. J. J. Hopfield, Electron Transfer Between Biological Molecules by Thermally Activated Tunneling. *PNAS* (1974).
27. J. Jortner, Temperature dependent activation energy for electron transfer between biological molecules. *The Journal of chemical physics*. **64**, 4860–4867 (1976).
28. V. G. Levich, R. R. Dogonadze, *Theory of Non-Radiation Electron Transitions From Ion to Ion in Solutions* (1959).
29. J. Jortner, M. Bixon, T. Langenbacher, Charge Transfer and Transport in DNA. *PNAS*. **95**, 12759 (1998).
30. C. Wan *et al.*, Femtosecond dynamics of DNA-mediated electron transfer. *PNAS*. **96**, 6014–6019 (1999).
31. T. J. Meade, H. B. Gray, J. R. Winkler, Driving-Force Effects on the Rate of Long-Range Electron Transfer in Ruthenium-Modified Cytochrome C. *Journal of the American Chemical Society*. **111**, 4353–4356 (1989).
32. J. R. Winkler, H. B. Gray, Electron Transfer in Ruthenium-Modified Proteins. *Chem. Rev.* **92**, 369 (1992).
33. D. G. Nocera, J. R. Winkler, K. M. Yocom, E. Bordignon, H. B. Gray, Kinetics of Intermolecular and Intramolecular Electron Transfer From Ruthenium (II) Complexes to Ferricytochrome C. *Journal of the American Chemical Society*. **106**, 5145–5150 (2002).
34. J. P. Allen, G. Feher, T. O. Yeates, H. Komiya, D. C. Rees, Structure of the Reaction Center From Rhodobacter Sphaeroides R-26: the Cofactors. *PNAS*. **84**, 5730–5734 (1987).
35. C. H. Chang *et al.*, Structure of Rhodopseudomonas sphaeroides R-26 reaction center. *FEBS Letters*. **205**, 82–86 (1986).
36. J. J. Rooney, Eyring Transition-State Theory and Kinetics in Catalysis. *Journal of Molecular Catalysis A: Chemical*. **96**, L1–L3 (1995).
37. H. Eyring, The Resultant Electric Moment of Complex Molecules. *Physical Review*. **39**, 746–748 (1932).
38. M. Byrdin, A. P. M. Eker, M. H. Vos, K. Brettel, Dissection of the Triple Tryptophan Electron Transfer Chain in Escherichia Coli DNA Photolyase: Trp382 Is the Primary Donor in Photoactivation. *PNAS*. **100**, 8676–8681 (2003).
39. Z. Liu *et al.*, Electron Tunneling Pathways and Role of Adenine in Repair of Cyclobutane Pyrimidine Dimer by DNA Photolyase. *Journal of the American Chemical Society*. **134**, 8104–8114 (2012).
40. A. Jasaitis, M. P. Johansson, M. Wikström, Nanosecond Electron Tunneling Between the Hemes in Cytochrome Bo3. *PNAS*. **104**, 20811 (2007).
41. A. G. Yakovlev, V. A. Shuvalov, Reversible Charge Separation in Reaction Centers of Photosynthesis: a Classical Model. *Doklady. Biochemistry and Biophysics*. **450**, 143 (2013).
42. C. C. Moser, C. C. Page, P. L. Dutton, Tunneling in PSII. *Photoch Photobio Sci*. **4**, 933–939 (2005).
43. S. Santabarbara, P. Heathcote, M. Evans, Modelling of the Electron Transfer Reactions in Photosystem I by Electron Tunnelling Theory: the Phylloquinones Bound to the PsaA and the PsaB Reaction Centre Subunits of PS I Are Almost Isoenergetic to the Iron–Sulfur Cluster FX. *Biochim. Biophys. Acta* (2005).

44. C. C. Page, C. C. Moser, X. Chen, P. L. Dutton, Natural Engineering Principles of Electron Tunnelling in Biological Oxidation–Reduction. *Nature*. **402**, 47–52 (1999).
45. A. Punnoose *et al.*, Fundamental Limits on Wavelength, Efficiency and Yield of the Charge Separation Triad. *PLoS ONE*. **7**, e36065 (2012).

Chapter II: Engineering blueprints for optimal designs of photoactivated charge separation in ET based systems: rules and principles.

2.1 Introduction:

Light-activated charge separation is the central process in a myriad of essential biological functions ranging from photosynthesis to DNA repair and signaling. (*1*-*5*)They share a common structural and engineering motif where the light activation of a pigment cofactor creates an excited electron/hole that, via electron transfers between multiple cofactors, quickly evolve into a charge-separated state, with a reducing negative charge residing on a distant acceptor cofactor and a oxidizing donor on a donor cofactor, as seen in Figure 2.1. Understanding of this fundamental common theme and descriptions of photosynthetic energy conversion structures, metabolism and genomes of terrestrial and aquatic plants and microorganisms have opened the door to their reengineering by biologists and emulation by synthetic chemists. Synthetic biologists and biochemists intend to maintain the unique power of plant photosystems to oxidize water as a vast source of reducing electrons but, rather than directing electrons along their natural course to reduce CO_2 to carbohydrates, they are steering them more directly toward generation of H_2 and reduced carbon or nitrogen compounds as fuels. In parallel, chemists aim to synthesize molecular mimics of early photochemical steps of natural photosynthesis with the same goal of generating chemical fuels in response to mankind's rapidly growing energy needs.

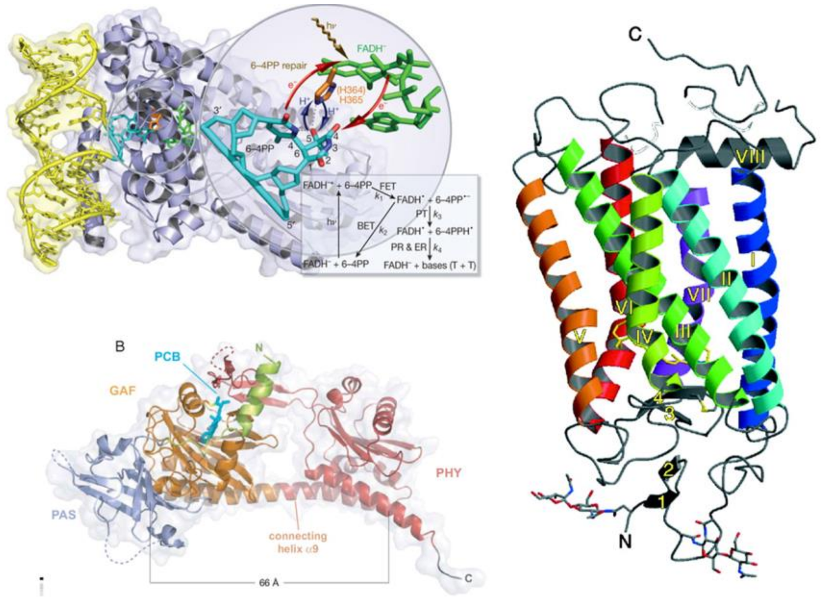


Figure 2.1: Multitudes of natural photosystems that rely on electron transfer driven charge separation to achieve essential functions in biology. Top Left: crystal structure of photolyase in complex with its substrate, DNA, as well as the energy diagram describing the electron transfer details. **Bottom Left:** crystal structure of phytochrome, reproduced from (*5*). Right: crystal structure of rhodopsin. (*4*)

The experimental and theoretical advances outlined in the previous chapter have allowed us to approach electron transfer events as kinetics problems where the rates of ET can be computed in an elegant manner using a small set of parameters and require minimal computational resources. Specifically, the Moser-Dutton ruler(*6*), formulated as an simplified representation of the semi-classical Hopfield-modification of the Marcus theory(*7*, *8*), based on a large collection of various ET reactions in biological and synthetic settings, shown here as equations [1]A and [1]B, allows efficient and fast calculation of electron transfer rates between any pair of cofactors given the distance, free energy difference, and reorganization energy between the two. The derivation and specific terms of this set of equations had been covered in great detail in Chapter I. Although the fundamental assumptions made during the formulation of the Moser-Dutton ruler limit its applications ostensibly to “long-range” electron transfers in protein, in practice the ruler has been demonstrated to be effective for ET reactions whose distances are near Van der Waal distance. Given any electron transfer driven system, we can fully describe its kinetics as long as the distances between the participating cofactors, their respective redox potentials, and the nature of the medium for the electron transfers are known.

$$\log k_{et}^{exer} = 13 - \beta(R-3.6)/2.3 - 3.1(\Delta G+\lambda)^2/\lambda \qquad [1A]$$

$$\log k_{et}^{ender} = \log k_{et}^{exer} - \Delta G/0.06 \qquad [1B]$$

Past studies of our group have demonstrated successful and effective application of the Moser Dutton ruler in reproducing the electron transfer kinetics in well-studied light activated charge separation process of bacterial photosystem I and photosystem II using only the reported photophysical properties of the chromophores and the reported redox midpoint potential values of the cofactors and the distances between them.(*9*, *10*) Similarly, application of the Moser-Dutton ruler by other research groups have shown its capability to model electron transfer rates in bacterial and plant photosystems, as well as other electron-transfer based enzymes. (*11-13*)Our past results applying our empirical rate equations to various photosystems successfully demonstrated that, given all the necessary parameters of any ET-based photosystems, models constructed from Moser-Dutton ruler can accurately describe the system in terms of its non-equilibrium kinetics. This description of the ET-kinetics in turn allows us to compute an essential quantity that we from here on consider as the characteristic metric of any light-activated, electron-transfer based charge-separating device: the yield of the charge-separated state at a specific selected lifetime.

With the growing demand and desire to engineer and design light-activated photosystems in biological and chemical settings without relying on modification of the structurally complex natural reaction centers, we noticed an apparent lack of systematic and comprehensive analysis or available analytic tools to provide a reliable and robust set of guidelines to identify what optimally efficient molecular photochemical systems, natural or manmade, would comprise in terms of photo-pigment and redox cofactor numbers, properties and size and geometry to catalyze oxidation of water for the production of reduced chemicals and fuels. (*14-16*) Certain attempts were made previously by other research groups in analyzing and exploring the effects of specific engineering parameters on the efficiency of the photosystem(*17*, *18*), but, without offering an extensive review here, none has taken the comprehensive and exhaustive, first-principle based approach we envisioned to be necessary for formulating a complete set of engineering principles for ET-based photosystems. Here we develop an efficient, simple, yet highly accurate analytical paradigm using the Moser-Dutton ruler to create engineering blueprints that identify and prescribe optimal design parameters for ET-based photochemical devices, relying solely on the fundamental theories of electron transfer introduced in the previous chapter. Due to the simplicity of the underlying theory and the non-probabilistic nature of our analytical algorithm, our computational system is implementable in all common and local programming environments without the requirement of computing clusters or cloud services. At the same time, because of the lack of engineering principles and constraints other than the Moser-Dutton ruler incorporated into the algorithm, the engineering principles formulated in this chapter represent the purest and most general set of rules that are universal to all photosystems that rely on ET-driven charge separation as the core mechanism. However, it also means that applications of these principles to designs and engineering in specific environments, such as the biological context represented by protein environment, would require additional modifications and amendments to improve their effectiveness, as will be demonstrated in detail in Chapter III.

2.2 Abstracting photosystems to cofactors and their roles: the *n-ad* paradigm.

In order to devise a systematic analytic approach for all photosystems, natural or man-made, using the Moser-Dutton ruler's computational elegance and simplicity, we consider only the electron transfer active cofactors of the photosystems, abstracting away the surrounding environment, e.g. protein residues in the case of natural photosystems, as an essential parameter for the rate equation, namely the reorganization energy. This abstraction results in what we term as the *n-ad paradigm,* where n is the number of cofactors in the photosystem under consideration, and we always have n>0, as shown in Figure 2.2A. Consequently, the minimal photosystem is a photochemical monad, consisting of a single photoactivatable pigment cofactor, P, which upon photo-excitation forms an excited state P* and subsequently returns to the ground state with a characteristic excited state decay rate. Many pigments also have excited states that undergo singlet to triplet conversion to produce excited state with significantly slower(spin-forbidden) decay rate. In our n-ad paradigm, however, such conversion is equivalent to replacing the pigment with one that has a lower photon energy and slower decay rate, without loss of generality. Our analysis later will reveal and discuss the advantages and disadvantages of including a triplet excited state P in the engineering of a photosystem.

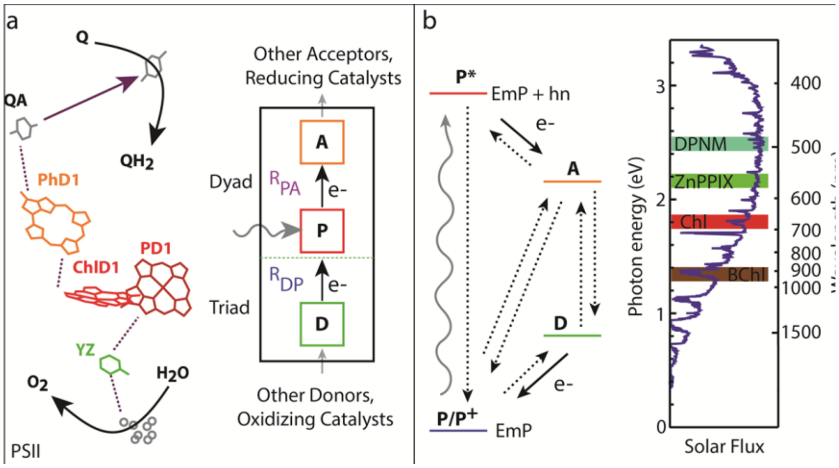

a
Q
QA
QH2
PhD1
ChlD1
PD1
YZ
O2
H2O
PSII
Other Acceptors,
Reducing Catalysts
A
Dyad
RPA
e-
P
RDP
e-
Triad
D
Other Donors,
Oxidizing Catalysts
b
P*
EmP + hn
e-
A
D
e-
P/P+
EmP
Photon energy (eV)
3
2
1
0
400
500
600
700
800
900
1000
1500
DPNM
ZnPPIX
Chl
BChl
Solar Flux

Figure 2.2: Cofactor-centric, role-based abstract representation of all photosystems and representation of their mechanism. A: natural photosystems, sush as PSII, represented as cofactor multi-ads, with the essential functions of the cofactors reduced to a dyad/triad representation on the right. **B:** The details of the electron transfer kinetics of the n-ad paradigm, as well as the coverage of the solar energy spectrum by various pigments of the photochemical n-ads.

Additional cofactors in the photosystems can play the roles of acceptors, A, or donors, D. Acceptors receive an electron from the excited state of the P or from a previous acceptor and form either intermediate or final charge separated states. Donors donate an electron to the oxidized ground state of the P after the excited state of P has given its electron off to the acceptor or the preceding donor. Unless otherwise noted, all cofactors of the photochemical n-ads are arranged in a linear fashion, with the acceptors lined up to one side of the pigment, and donors on the other side. Consequently, the linear arrangement ensures that the full diameter of the cofactors will contribute to the distances between cofactors that are not adjacent to each other. The additional distances provided by the size of the intervening cofactor is one of the fundamental reasons why multi-ads offer significant performance gains over dyads. Rather than having the same distance for charge-separation and charge-recombination and therefore fully relying on the inverted region of the parabolic Marcus curve to stabilize the charge-separated state, multi-ads exploit the size of the cofactors to gain additional effective distance for the charge-recombination electron transfer. This will be discussed in greater detail in some of the upcoming sections in this chapter.

In spite of the possibility to include both donor and acceptor cofactors, we will demonstrate in an upcoming section that without any loss of generality in the case of triads, and with almost no loss of generality in the case of tetrads and beyond, to consider only n-ads that consist of pigments and acceptors, as donor-containing triads(DPA), when only the fundamental principles of electron transfer are concerned, are mathematically equivalent to PA-only triads, and therefore would provide the same optimal performance as pigment-acceptor(PAA) triads, while donor-containing tetrads are almost congruent to pigment-acceptor-only tetrads, and will confer marginal gain(DPAA) or no gain(DDPA) over pigment-acceptor tetrads(PAAA). However, in practical design settings, as will become evident in Chapter V and VI when we attempt to construct a proof-of-principle triads experimentally, including a donor rather than an additional acceptor offers the practical advantage of a much larger selection of cofactors from a already limited pool of candidates.

2.3 Performance parameters as metrics for optimality of design

It is impossible to establish any sense of optimality of design without having a standard and universal metric to compare the designs and determine which is more or less optimal. Since the goal of most known natural and man-made photochemical device is to harvest and convert solar energy into high-energy charges separated state for subsequent usage, usually in the form of diffusion limited catalysis, we identify and propose two parameters that together form the metric that determines the performance of any charge-separating photosystems and allow for direct comparisons between them: engineering efficiency at any selected lifetime of charge separation. Its definition is shown in Figure 2.3, as a product of the quantum yield(QY) of the charge separated state at selected lifetime and the ratio of energy preserved(REP) in the charge separated state. Engineering efficiency at any selected lifetime offers the most appropriate metric for the performance of photosystems in most circumstances because it accurately measures the percentage of the incident photon energy that is captured and held available by the charge separation for at least as long as the specified lifetime. The selected lifetime of choice is usually on the order of 1-100ms, since this range represents the timescale of diffusion-limited chemical reaction typical of the type of catalysis coupled to photoactivated charge separation, such as hydrogen evolution(*19*, *20*). This quantity therefore directly demonstrates the capability of the given photosystem in solar energy conversion at a timescale that allows it to be functionally relevant.

However, it is important to notice that as the product of ratio of energy preserved and quantum yield at selected lifetime, optimal engineering efficiency sometimes does not correspond to either maximal quantum yield or maximal ratio of energy preserved, but rather a compromise between the two or a preference for one of the two. In certain circumstances, this means also that the appropriate metric for a photosystem's performance is no longer the engineering efficiency, but rather QY or REP. It is straightforward to see that neither QY nor REP can be optimized without trivializing the problem unless additional constraint on the other parameter is enforced. Specifically, without

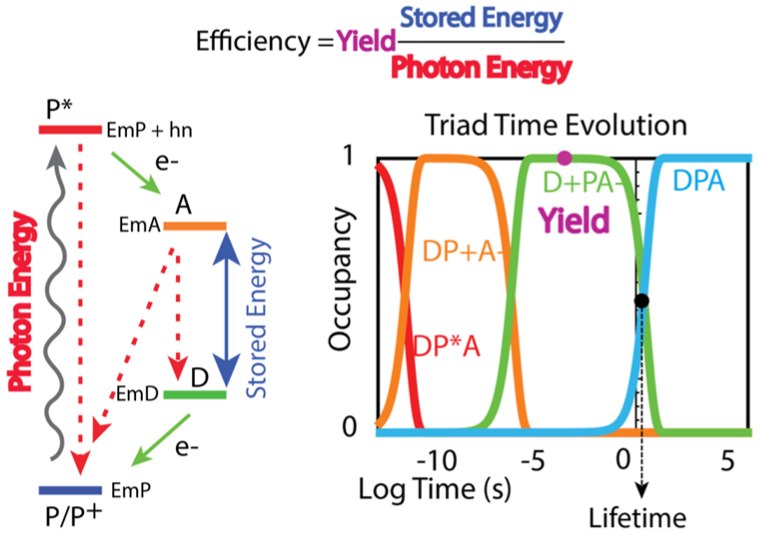


Figure 2.3: Performance metrics to evaluate and compare different charge-separating photosystems. Left: Definition of "stored energy" and "photon energy" in the context of a DPA triad. Red dashed arrows represent premature recombinations. Green solid traces represent charge-separating electron transfers. **Right**: Definition of the "yield" of the charge-separated in the context of a DPA triad. It is crucial to note that these definitions are universal across all forms of photochemical charge-separating devices, and dependent on the configuration of the device, only the specific representation will differ, while the concept remains the same.

setting a lower limit on the REP, optimizing QY alone will lead to the trivial solution of minimal REP, since QY has a somewhat inverse correlation with the energy preserved. Similarly, optimizing for REP without constraining QY will also lead to an arbitrarily high value of REP but minimal QY. This concept will become important in Chapter III, when we examine the engineering of conserved core cofactor units of natural photosystems and their adherence to and deviation from the optimal engineering blueprints we propose here in this chapter. It becomes apparent that natural photosystems preferentially “chose” to optimize for quantum yield while allowing the REP to be suboptimal when biologically relevant constraints are also included in the engineering landscape.

2.4 Analytical paradigm and the mathematical infrastructure of the formulation of the engineering blueprints.

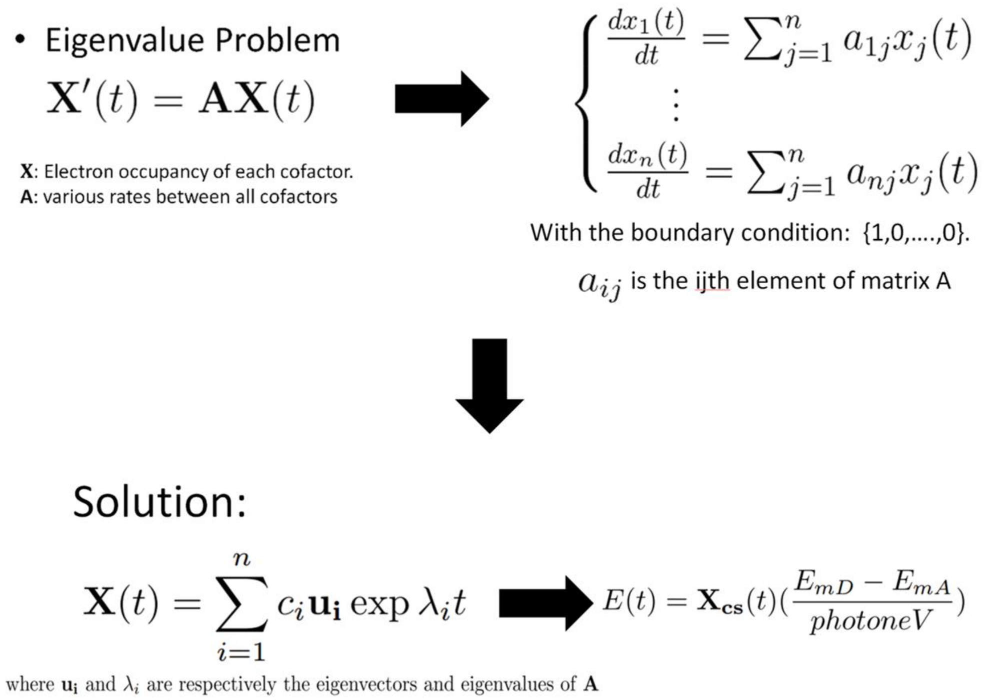
Our analytical paradigm consists of three progressive steps/stages towards the final goal of obtaining the optimal engineering parameters that give us photosystems with the highest performance metrics outlined above, and the three stages are: descriptive, predictive, and prescriptive.

The technical details of the implementation of the analytical paradigm described in this section using Mathematica 10.0 are included at the end of this chapter in section 2.12.

The descriptive stage involves direct application of the Moser-Dutton ruler detailed in Chapter 1 to obtain estimates of electron transfer rates between the cofactors of any photosystems with arbitrary parameters, thereby providing a descriptive illustration of the kinetics of electron and shedding some lights on the performance of the photosystem.

In the predictive stage, we seek to obtain either closed form or numerical but mathematically tractable and computationally robust representation of the metrics and then use such solution to compute the metrics given any arbitrary configuration of the parameters of the photosystems. Chris C. Moser has contributed most significantly to the development of the essential algorithms and many of the analytical procedures for the predictive and the following prescriptive stages of the analysis.

To arrive at such expressions for the efficiency of any given photochemical device, we describe the electron transfer kinetics between all cofactors of the photosystem as the semi-analytic solution of a system of differential equations generated from computing the rates between all ET-active “microstates” of the photosystems. Each of such microstate represents a non-equilibrium snapshot of the photosystem that captures the location of the excited electron created initially by photo-excitation of the pigment cofactor. For example, in an abstract 3-cofactor system, termed PAA triad, 4 microstates exist: $\{[P^*A_1A_2], [P^+A^-_1A_2], [P^+A_1A^-_2], [P^+A_1A_2]\}$. We can then write a rate equation for each of the 4 microstates as a function of rates between the microstates and the population of the microstates. As summarized in Scheme 2.1, the solution to this system of ordinary differential equations contains the full kinetic information of all electron transfer events following the photo-excitation of the pigment cofactor, listed as time-dependent functions of the population of each of the microstates. One of these functions describes the population of the final charge-separated state for a specific photosystem, depending on the number of cofactors involved. Evaluation at selected time allows us to discover the quantum yield(abbreviated from now on as QY) at a selected lifetime. When multiplied by the fraction of energy preserved in the charge-separated state, we obtain a mathematical expression for our desired metric, engineering efficiency at a given lifetime. This expression is a numerical function with the parameters that set up the system of ODEs for the ET kinetics of the photosystem as independent variables, including but not limited to: the distances and driving forces between all pairs of cofactors, reorganization energies for each ET reactions, photophysical properties of the pigment such as photo energy and excited state decay rates, and so on.



• Eigenvalue Problem
$\mathbf{X}'(t) = \mathbf{A}\mathbf{X}(t)$
X: Electron occupancy of each cofactor.
A: various rates between all cofactors
$\frac{dx_1(t)}{dt} = \sum_{j=1}^{n} a_{1j}x_j(t)$
$\vdots$
$\frac{dx_n(t)}{dt} = \sum_{j=1}^{n} a_{nj}x_j(t)$
With the boundary condition: {1,0,….,0}.
a_{ij} is the ijth element of matrix A
Solution:
$\mathbf{X}(t) = \sum_{i=1}^{n} c_i \mathbf{u_i} \exp \lambda_i t$
$E(t) = \mathbf{X_{cs}}(t)(\frac{E_{mD} - E_{mA}}{photoneV})$
where $\mathbf{u_i}$ and λ_i are respectively the eigenvectors and eigenvalues of $\mathbf{A}$

Scheme 2.1 : Analytical steps to obtain the complete kinetic information of a ET-based photosystem as a time-dependent function of all engineering parameters. Top left: Formulation of the kinetics problem as a typical eigenvalue problem, where the operation is the differentiation and the eigenvalue matrix consists of the rates. Top right: specific differential equations along with the boundary/starting condition. Bottom left: solution of the eigenvalue problem written as a sum of products of eigenvectors and exponentials of eigenvalues. Bottom right: mathematical expression for the engineering efficiency at selected lifetime.

Finally, in the prescriptive stage, we subject the numerical function of the performance metrics to various optimization processes with different constraints and free parameter spaces depending on the problem of interest. The result of the optimization is a list of design parameters that together will generate the photosystem with the optimal performance given the constraints. Although any of the parameters that determine the value of the engineering efficiency of the charge-separated state can be fixed and considered as constraints, we particularly choose a few parameters as the constraints and regard the rest as design parameters. Specifically, we believe that certain parameters are much more difficult to vary and adjust in comparison to others, and these include the incident photon energy, lifetime of the excited state of the pigment, the reorganization energy of the system and so on. These are “hard” parameters that cannot be easily modulated without completely throwing away a design, while the “soft” parameters, including the distances driving forces between cofactors, can be modified significantly faster and more easily in practice. Consequently, in this and the next chapter, we usually consider the “hard” parameters as constraints and the “soft” parameters as accessible engineering inputs.

Also in the prescriptive stage of our analysis, we can examine the engineering tolerance towards a certain subset of the parameters that determine the optimal value of the EE. Specifically, given all the parameters that are returned by our prescriptive algorithm to be the optimal engineering parameters, we can fix most of them and allow certain others to vary, while computing the resulting engineering efficiency. This gives us a contour, in x-dimension, with x being the number of parameters allowed to vary, that demonstrate the shape of the engineering landscape of this subset of parameters in the vicinity of optimality. It allows us to understand how much tolerance to changes in parameters the engineering of photosystems provides, thereby giving us an additional way to measure the practicality of the design, since in most real engineering circumstances, it is near impossible to design following the exact specification of the blueprints, as will become evident in Chapter V and VI.

Throughout all three stages of our analytical paradigm, it makes the important assumption that the photosystem under consideration only involves one excited electron, and all cofactors within the system are Nernst n=1 redox centers, that is, their conversion from reduced to oxidized state only involves the transfer of a single electron. This is a simplifying but reasonable assumption, since almost all known natural photosystems employ only n=1 redox centers as ET cofactors as their core components.

2.5 Comprehensive engineering blueprints of optimal charge-separating dyads

To explore the design landscape and to create a set of comprehensive engineering blueprints for man-made or natural photosystems, we consider the prescriptive stage of our analytical paradigm, examining the values of the design parameters as well as the tolerance range of these parameters that produce the optimally performing photosystems according to our metrics. However, our expression of the engineering efficiency of the photochemical device depends on more variables than those that can be straightforwardly and concisely illustrated. In order to provide a clear picture without sacrificing the comprehensiveness of our study, we will explore and present the optimal engineering of photochemical multi-ads iteratively to allow us to focus on a few design parameters at a time, while holding the remaining parameters locally/temporarily constant, allowing them to vary across separate "experiments". Many of the parameters that are held constant are those mentioned above as "hard" parameters, or constraints.

Specifically, we apply our prescriptive analysis with fixed and/or constrained parameters that represent certain situational design and engineering limitations. We begin by demonstrating the optimal engineering for a "generic" n multi-ad, namely, the n-ads engineered in a protein-interior whose reorganization energy is 0.7eV for all the electron transfer events within the photosystem, using a typical chromophore whose excited state is singlet and decays via fluorescence with a lifetime of 6ns. Using these locally constrained "generic" parameters, we sequentially examine the optimal engineering efficiency for dyads, triads, and tetrads at various selected lifetimes with various incident photon energies (representative of different available pigment cofactors) at a fixed uniform reorganization energy. For each level of complexity of the multi-ads we explore the distances and driving forces between the cofactors that are necessary to produce the optimal engineering efficiency under the given conditions. We then relax these constraints and examine the effects of varying these previous held parameters on the engineering efficiency. By observing and analyzing the patterns and trends of the correlation of the engineering parameters with the engineering efficiency, we formulate corresponding design blueprints that are universal and fundamental to any electron transfer based photosystems, applicable to practical engineering regardless of the setting and the specific environment in which the photosystem is designed. However, the lack of any other constraints than the fundamental theories of electron transfer that govern these design blueprints also means that the design blueprints formulated at this stage are also subjected to modifications when additional constraints are included for engineering in specific settings, as will become apparent in Chapter III.

It is particularly attractive to go through the analysis of the multi-ads from dyads to tetrads and beyond, since dyad represents the fundamental and most basic form of charge-separating device. It allows us to fully understand the immediate interactions of physical principles that promote long-lived, highly efficient charge separation. With the behaviors of dyads fully understood, we can then isolate and focus on the additional interaction of physical principles that are provided by the additional cofactors and the corresponding electron transfers. However, it is important to point out here that in order to fully illustrate the engineering specifics of the dyad, the choice of "generic" reorganization of

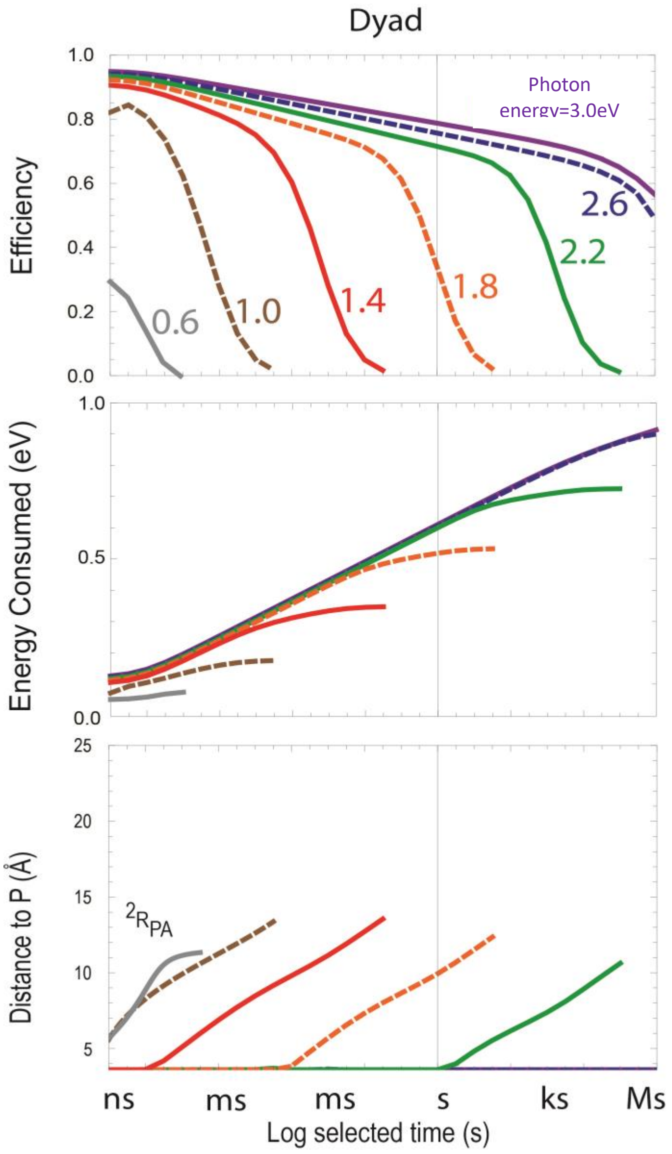


Figure 2.4: Engineering blueprints of photochemical dyads, demonstrating various aspects of the optimal engineering parameters at selected lifetime of charge-separation ranging from ns to megaseconds. Top: Engineering efficiency of P-A dyads of various incident photon energy(singlet excited state), with a 0.3eV reorganization energy. **Upper mid:** energy gap between the singlet excited state of the pigment and the acceptor for dyads of various incident photon energy, plotted at 0.3eV reorganization energy. **Bottom:** P-A distance of the optimal dyad at selected lifetime for various incident photon energies.

0.7eV represents too harsh of a condition and prevents us from presenting a complete engineering landscape of the dyad. That is, in an environment whose reorganization energy is 0.7eV, dyad is incapable of sustaining stable charge-separation for lifetimes greater than microseconds. Instead, for the sake of clear demonstration of the engineering principles governing the behavior of dyads, we choose 0.3eV as the default reorganization energy, and will return to the default value of 0.7eV for the analysis of the triads and tetrads.

Panel A(top) of Figure 2.4 demonstrates the optimal engineering efficiency of a PA dyad, at selected lifetimes ranging from 10ns to 100s for a series of incident photon energies. The B and C panels of Figure 2.4 illustrate the corresponding driving forces(represented as energy consumed) and distances between the cofactors required to achieve the optimal efficiency for the various incident photons shown in panel A. For all the incident photons, panel A shows that the optimal efficiency achievable by a photochemical dyad decreases monotonically with respect to increasing lifetime of charge-separation until the engineering efficiency drops to 0. Much more interestingly, note that this monotonic relation between the efficiency and the lifetime consists of two separate phases: an approximately linear “slow phase” and a non-linear, quasi-exponential “fast phase”. The approximately linear “slow phase” of the decrease in engineering efficiency corresponds to the necessary energy “consumed” in order to generate and stabilize the charge-separated state for a given lifetime, since according to equation 1B, the uphill electron transfer to repopulate the excited state(in the case of the dyad) or the preceding intermediate charge-separated state(in the case of the triad, tetrad, and beyond) is slowed down by 1 decade per 60mV of additional driving force, representing the additional energy required per Boltzmann distribution to populate the preceding state in the electron transfer.(21) This association led us to coin the term “Boltzmann phase” for the approximately linear portion of the engineering efficiency decay. This relation is demonstrated by the 60mV/decade slope of the linear phase of the energy consumed curves of the optimal dyads in Figure 2.4B. This represents the theoretical minimal cost in the engineering of charge-separating photosystems that is expected to last a certain lifetime.

This constant cost of charge-separation, therefore the Boltzmann phase of the engineering efficiency, however, is only applicable when the charge-separated state is stabilized via the inverted region of the Marcus curve. That is, when the driving force of the direct charge-recombination from the CS state to the ground state is so far in the inverted region(see section 1.5 of Chapter I) that it is faster for the electron in the CS state to recombine through uphill thermal repopulation of the excited state or the preceding intermediate charge separated state followed by its own decay route; in the case of the dyad it is therefore the singlet excited state decay. The exact rate of this uphill repopulation-decay combination can be computed in the way shown in Figure 2.5, as a product of the equilibrium constant of the forward and reverse ET and the rate of the CR of the repopulated state. Figure 2.5 also demonstrates the validity of this method of computing the uphill-repopulation driven recombination with experimental data obtained from observations of such up-and-over events in modified bacterial reaction center. (*21, 22*)The effective decay via the combination of the thermal repopulation of the preceding state and its own decay mechanism appears here, when only the principles of electron transfer are considered, to be the more efficient and therefore superior method of charge-recombination, but as will be discussed in further details in chapter III, it poses one of the most severe challenges faced by the engineering of photosystems when biological context are considered.

As the rate of the direct recombination becomes faster as a result of the optimal engineering adjusting to increasing design constraints, such as increasing lifetime of charge-separation(shown in Figure 2.4A following the progress of each individual trace), decreasing incident photon energy(shown in Figure 2.4A at a fixed lifetime across multiple traces), and so on, the engineering efficiency enters the “fast phase”, where recombination routes faster than uphill thermal recombination become dominant. In the case of dyads, the “fast phase” corresponds to faster direct recombination as the driving forces from the CS state to the ground state is no longer large enough to exploit the stabilizing effect of the Marcus inverted region. We thus term the fast decaying phase the Marcus phase. Consequently, this allows to conclude that the increasing design

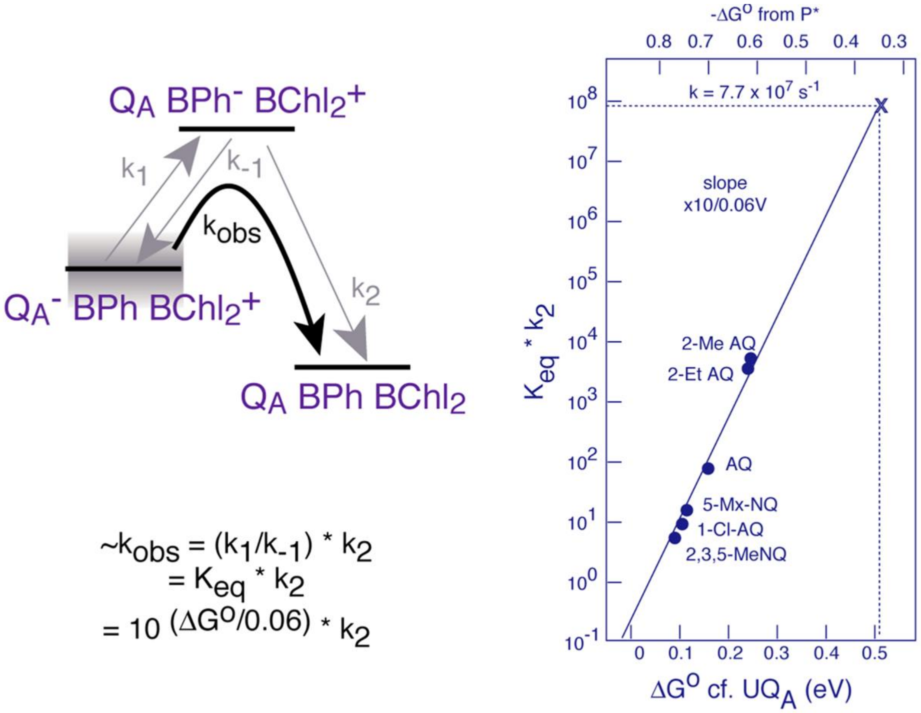


Figure 2.5: Indirect charge-recombination via uphill thermal repopulation of preceding, higher-lying energy states and the calculation of its rate. Top left: energy-diagram cartoon illustrating the steps involved in recombination via uphill thermal repopulation. Right: experimentally measured rates of recombination in bacterial reaction center with substituted quinones follow the model of recombination by thermal repopulation. (*21*)Bottom left: computing the exact rates of recombination by thermal repopulation.

constraints listed above all correspond to the common effect of pushing the dyad out of the Boltzmann phase by lessening the stabilizing effect of the inverted region as the difference between the A^- to P^+ driving force and the reorganization energy becomes smaller. Because of its ability to effectively magnify the slowing effect of the Marcus inverted region, a small reorganization energy appears to be, and has indeed been accepted as the consensus in the field of photosynthesis, the optimal and necessary engineering condition for successful charge-separating device. However, as will become evident later in this chapter as well as the next chapter, deliberate adaptation of low dielectric environments(small reorganization energy) is not an engineering guideline followed by natural photosystems. In fact, as will be revealed by the analysis of Chapter III, the engineering of natural photosystems appear to have proceeded directly against the guideline proposed here by completely avoiding the Marcus inverted region stabilization and recombination by thermal uphill repopulation of preceding intermediate charge-separated states or the excited state.

Once the dyads' engineering enters the "fast phase", additional engineering devices must be engaged to offer other means of stabilization than the Marcus inverted region. As seen in Figure 2.4C, P-A distance becomes the next stabilizing factor exploited to provide additional stabilization when the direct recombination rate becomes comparable to that of the uphill thermal repopulation of the excited state. However, while slowing down the direct recombination rate, increased P-A distance also decreases the rate of the forward electron transfer to create the charge-separated state equally. This statement is of utmost important and deserves detailed explanation and

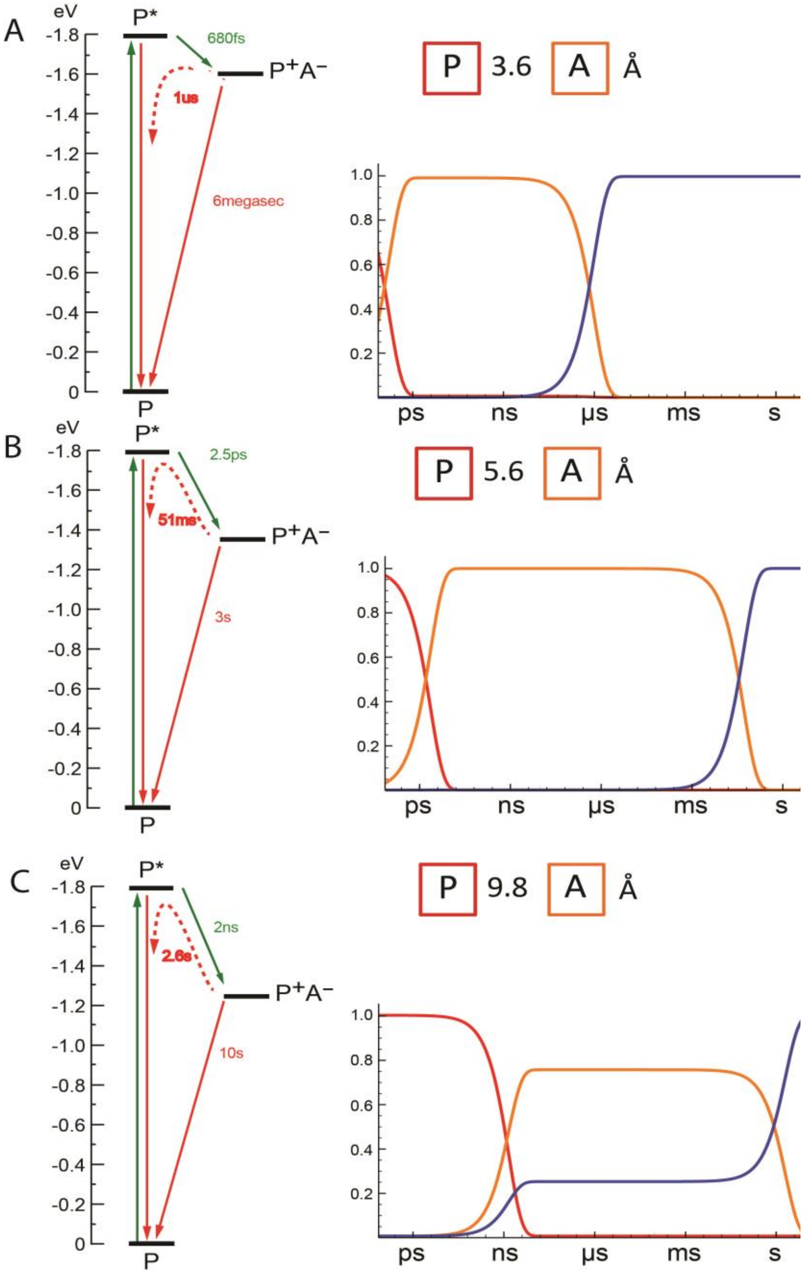


Figure 2.6: Trikes for optimal dyads at representative lifetimes. The energy levels and rates between the states are shown. Forward ET rates are shown in green arrows while recombinations are shown in red. A: details of optimal PA dyad designs with 1.8eV incident photon energy and 0.3eV reorganization energy, optimized for 10ns. B: same as A but for 1us. C: same as B but for 100ms.

illustration.

In fact, to better understand the detailed interactions of the all the factors influencing the yield and efficiency of various photosystems from dyads to other multi-ads, we present in a new type of plots we term “Trikes”, as exemplified by Figure 2.6. Unique to each configuration of an ET-based photosystem, “Trike” consists of the energy diagram of the cofactors of the photosystem displaying the rates of each significant ET reaction of the photosystem; the kinetic profile of the electron transfer reactions of the photosystem, and the cofactor arrangement of the photosystem. Together the three components of a “Trike” illustrates the details and specifics of the complete electron transfer kinetics of a photosystem under a specific set of constraints/environmental variables, revealing main loss mechanisms and reasons of failures, as well as providing a detailed and straightforward understanding of the complete kinetic information of the specific design of the photosystem.

Specifically for dyads, as seen in Figure 2.6A-C, forward electron transfer must compete with the inherent decay rate of the excited state, for which we have used the standard singlet fluorescence lifetime of tetrapyrrole pigments of 6ns. Specifically, the difference between Figure 2.6B and C demonstrates the challenges dyads cannot overcome with the engineering devices available to guarantee high yield at the desired lifetime. Between 1ms and 1s, the dyad must “pay” additional 180mV in energy to ensure that the recombination via uphill repopulation is on the timescale of seconds, but the additional cost in energy results in a smaller driving force to the ground state, and in order to slow down the direct recombination, increasing the P-A distance becomes necessary. When the optimal dyads continue to increase the P-A distance to slow down the direct recombination rate to allow the charge-separated state to achieve longer and longer lifetimes, the forward rate eventually becomes too slow and the charge-separated state is prevented from forming. This results in the engineering efficiency of the dyads begin to leave the Boltzmann phase and enter the Marcus phase, quickly falling to 0 at lifetimes that are often inadequate for diffusion-limited chemical reactions.

We have fully explored and illustrated the possible engineering devices available to photochemical dyads to generate stable and long lived charge-separation in response to stronger and harsher constraints. It is evident that as the most basic form of photosystem for charge-separation, dyads are highly limited in their performance and display high sensitivity to increasing engineering constraints. The poor performance of dyads is rooted in the lack of effective additional engineering devices to rely upon when its primary method of stabilization of charge-separation, namely the Marcus inverted region, becomes inadequate due to increasing constraints. This limitation explains well the observation that no effective, dyad charge-separating photosystems has been observed in nature, and engineering efforts at constructing charge-separating dyads(*23*, *24*) have been largely unsuccessful or unimpressive at best, with the longest charge-separated reported to be 200us.

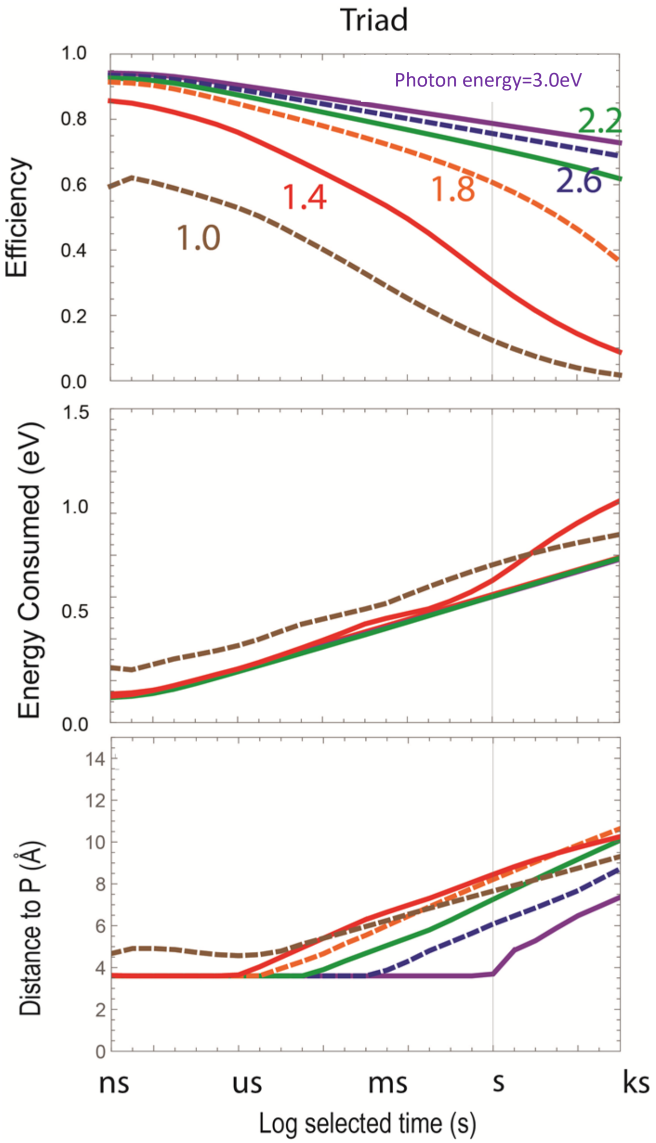


Figure 2.7: Optimal engineering blueprints for PAA triad at selected lifetimes from nano- to kiloseconds for photon energies from 1.0eV to 3.0eV. Top: Engineering efficiencies of the optimal triads at various lifetimes with various incident photon energies. **Mid**: The amount of energy used in terms of driving forces to create the final CS state with optimal efficiency at selected lifetimes for various incident photon energies. Bottom: The distance between the pigment and the acceptor at various incident photon energies.

2.6 Engineering blueprints for optimal triads.

The engineering of charge-separating multi-ads demonstrate the fundamental phenomenon that allows us to draw one of the central design guidelines of electron-transfer based photosystems. The addition of another cofactor in the form of either another acceptor cofactor linearly placed further ahead of the primary acceptor, or a donor cofactor placed on the other side of the acceptor beyond the donor, immensely improves the engineering efficiency achieved by the charge-separating photosystem at lifetimes that are relevant to chemistry. Here we focus our analysis of the PAA triad due to the structural and organizational similarity to the PA dyad as well as the simplicity with which its efficiency and energy-preserved metrics could be discussed, although as will be demonstrated in section 2.6, the behavior of the DPA triad is in fact identical to that of the PAA triad.

As shown in Figure 2.7, the optimal engineering of triads significantly outperforms that of dyads, sustaining charge-separation of much higher efficiency at much longer lifetimes and much lower incident photon energy. However, similar to the dyads, the engineering efficiency curves still consist of the Boltzmann and Marcus phases, although the Boltzmann phase of triads tend to persist significantly longer than that of dyads and the distinction between the two phases are more blurred, with a smoother transition between the two phases. Indeed, as seen in Figure 2.7C, the cost of charge-separation for triads largely replicate and extend the trends observed in Figure 2.4C, that is, triads continue to give up 60mV per each additional decade for the lifetime of the charge-separation, until the Boltzmann phase ends and the Marcus phase begins. To demonstrate effect of the increasing engineering constraints on the distances of triads in the most direct and straightforward manner, we chose to plot the mean edge-to-edge distance between P-A1 and A1-A2 in Figure 2.7C. Here a trend similar to the one seen in Figure 2.4 is also observed. Specifically, for longer lived charge-separation, distances must be used to extend the Boltzmann-based stabilization of the charge-separated state when the inverted-region slowing down of the direct recombination is no longer enough. The significant extension of the Boltzmann phase, as seen in both Figure 2.7B and C, accounts for the impressive performance gain in triads over dyads, but the question remains, how?

The presence of the second acceptor allows the charge-separating device to utilize an additional engineering device that fundamentally alters the engineering landscape of charge-separating devices that characterize the necessity of multi-ads in successful designs of long-lasting charge-separation. The extra cofactor(s) allow the photosystem to extend the “linear”, Boltzmann phase of the efficiency/lifetime curve by stabilizing the charge-separated with distances when the Marcus inverted region is no longer applicable. But how is this different from the dyads? Indeed, dyads also attempted to exploit distance, but the increased distance also slowed down the electron transfer reaction to create the CS state. Because of the additional cofactor and its linear placement away from the pigment, triads can enjoy the benefit of the distance-based stabilization of the CS state without slowing down the forward ET reaction to form the charge-separated state. In fact, as seen in the Trikes of Figure 2.6, triads can be seen as dyads that have separate distances for the forward ET reaction and the charge-recombination reaction, with the distance for the forward ET reaction significantly shorter than that of the recombination reaction. Indeed, because electron transfer distances are calculated from the side-to-side distances between cofactors, the size of the primary acceptor effectively become the additional distance between the final CS state and the pigment cofactor that provides the extra stabilization(slowing down the charge-recombination) without slowing down the charge-separation. This extra distance allows the charge-separated state to continue to recombine via the uphill thermal repopulation of the preceding CS state/the P* even though the driving force of the A2 to P^+ reaction can no longer keep the direct recombination in the Marcus inverted region. Thanks to the stabilization by the extra distance offered by the size of the cofactor, the forward ET reactions, both from P* to A1 and from A1 to A2 can be engineered so that the electron can arrive at the final charge-separated state while incurring minimum amount of loss in both the quantum yield and the energy preserved given the constraint of required lifetime of charge-separation, of course.

As the constraints continue to become harsher, triads gradually begin to lose the ability to sustain the charge-separation using entirely the Boltzmann phase. Unlike the dyads, the Marcus phase of triads manifests itself from multiple sources of failures and therefore appear more gradual and smooth than what's seen in the dyads, where a precipitous drop in the efficiency results from the slowing down of the forward electron transfer from P to A to an extent where the forward ET rate is slower than the lifetime of the excited state. In triads, premature charge recombination to the ground state can happen at two rather than one branching points for triads, namely the P* to A1 and the A1 to A2, where forward electron transfers must compete against the rate of decay of the singlet excited and the $A1^-$ to P^+ recombination. The presence of two rather than one branching point appears to present additional challenges not faced by the single branching point in dyads, providing more opportunities for the photosystem to fail. In reality, however, the two branching points allow the triad to distribute the necessary extra distance for Boltzmann-stabilized longer lifetimes between the P-A1 and A1-A2 pairs. This ensures that, for the lifetimes at which dyads fail completely and beyond, triads do not see its engineering efficiency suffer a sudden decay to 0 but rather a much more gradual and smooth trend, as can be seen from a comparison of the curves representing the same photon energy(say, 1.4eV) in Figure 2.7 and Figure 2.4.

Figure 2.8A-C illustrate triads' significantly improved abilities to deal with pressures from increasing constraints in comparison to dyads, specifically emphasizing the ability to sustain increasingly long-lived charge-separation without suffering a drastic failure as seen in Figure 2.4. It is important to note that only at lifetimes as long as 100s, the optimal triad enters the Marcus phase, in which the direct recombination to the ground state proceeds faster than the uphill-thermal repopulation driven recombination, as seen in Figure 2.8C. Yet even at this

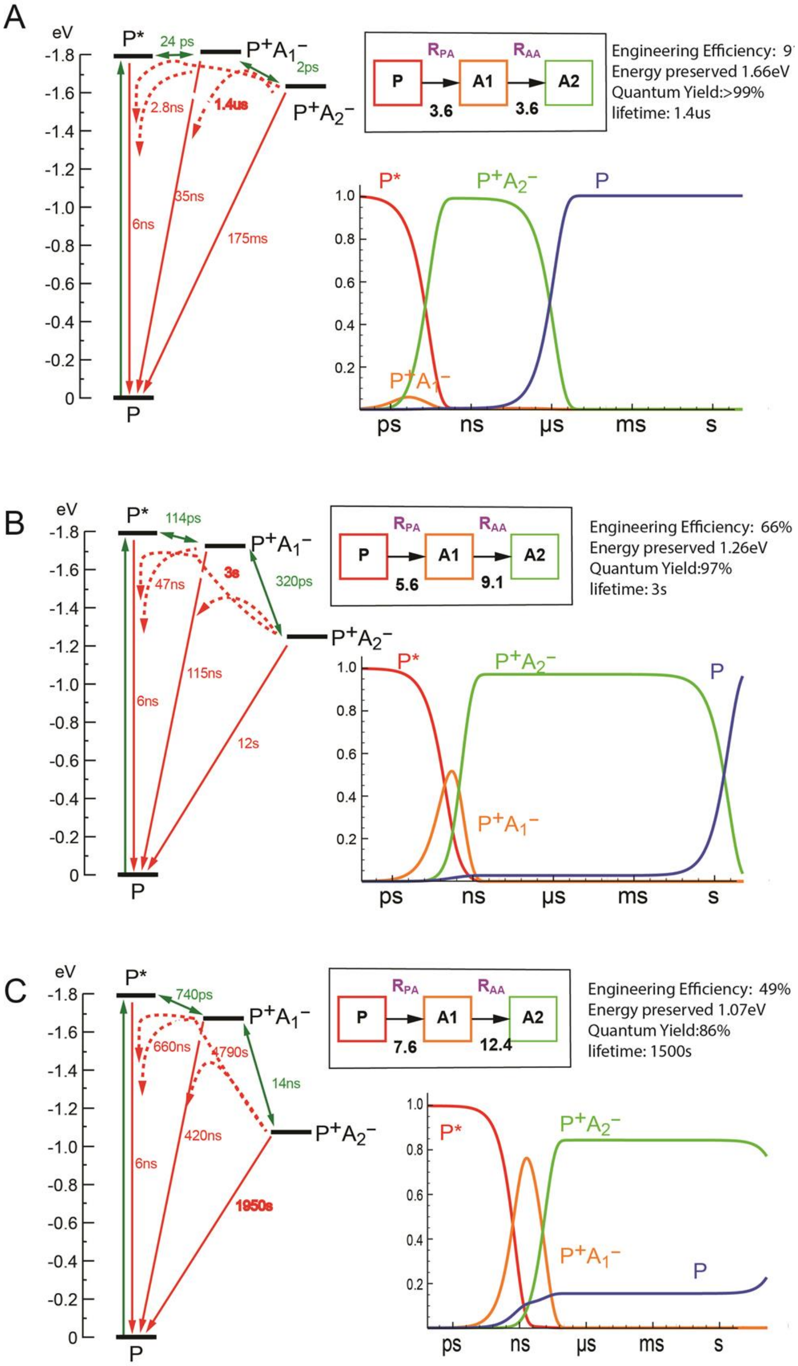


Figure 2.8: Electron transfer details of optimal triads at various selected lifetimes. The green arrows and numbers indicate the rate of the forward electron transfers, while the red are recombination reactions, with the solid traces representing direct and dashed traces represent uphill-repopulation followed by recombination. The cartoon and the numbers indicate the distances between the cofactors, while the kinetic profile shows the populations of the various states during the charge-separation. A: optimal triad for a 10ns charge separation. B: 100ms CS. C: 100s CS.

long-lived lifetime, the major source of loss in efficiency stems from the premature recombination occurring at the two branching points when the forward ET rates are slowed down by the increased distances to stabilize the final CS state, rather than lowering of the energy level of the final charge-separated state or the complete failure in ability to send the electron to the final charge-separated state. Indeed, to generate and stabilize a 100s charge-separation using the minimal Boltzmann cost, we would expect around 720mV of the incident photon energy to be expended, Figure 2.8C shows that the energy preserved in the charge-separated state is 1.06eV. That is, 740mV of the incident energy is used to generate and stabilize the final charge separated state, matching the expected minimal cost of the charge-separation since in order to generate a 100s charge-separation with optimal efficiency the actual lifetime of the CS state must be at least 1 order of magnitude greater, and we begin counting the 60mV/decade cost at the lifetime of the excited state. With the forward rate form P* to A1 at 740ps and the singlet excited state decaying at 6ns, 12% loss is incurred at this branching point, and another 4% is incurred at the second branching point. In stark contrast to the 33% loss at the only branching point in the PA dyad at a lifetime 3 decades earlier seen in Figure 2.6C, triads' ability to split the necessary increase in distance to the two sub-steps along the formation of the final CS state, along with the "free" additional distance provided by the diameter of the cofactor, represents the central engineering mechanism that distinguishes it from the dyads.

2.7 Engineering blueprints of optimal photochemical tetrads

After the significant improvement provided by the third cofactor we have seen in the previous section, it is therefore natural to expect that the addition of the 4th cofactor would provide a similarly significant improvement in the engineering. However, Figure 2.9 demonstrates that, unlike the dramatic improvements seen in the transition from a dyad to a triad, the incorporation of a 4th cofactor does not grant the same level of improvements in the engineering efficiency we had seen by going from 2 to 3. Instead, for most incident photon energies at most selected lifetimes, the engineering efficiency traces of the tetrads appear to be highly similar to those of the triads, offering marginal increases at

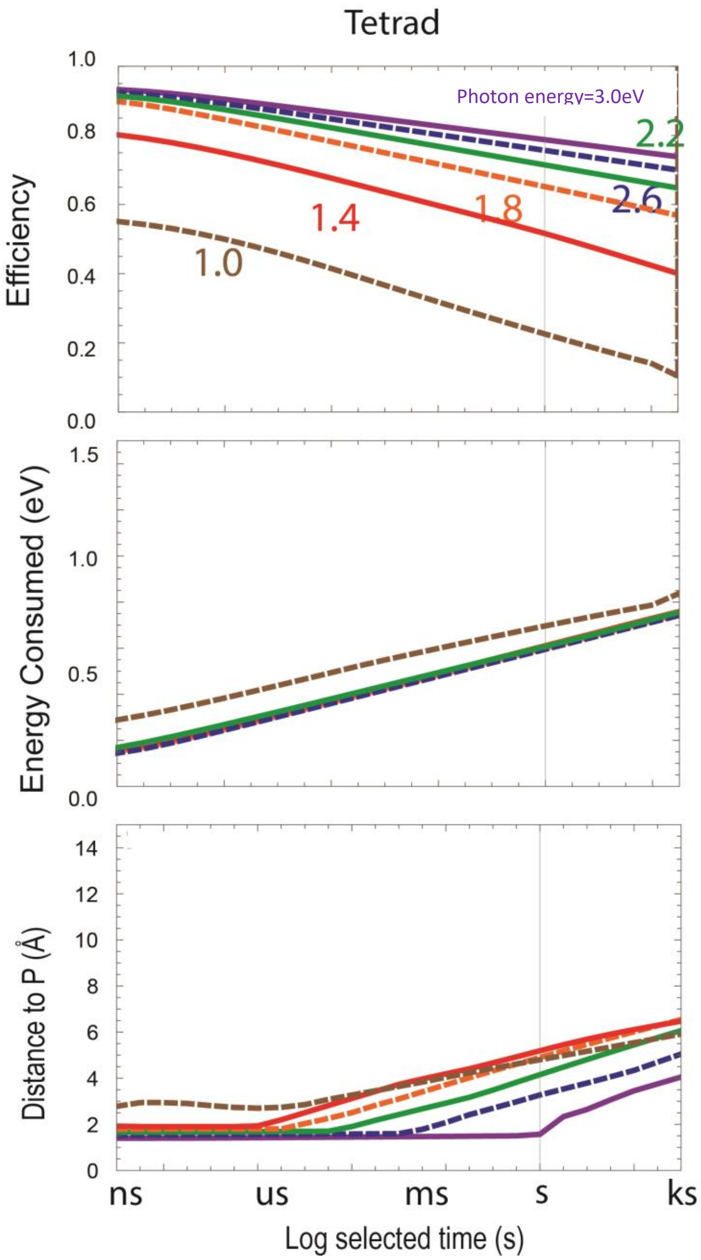


Figure 2.9: Optimal engineering blueprints of charge-separating tetrads at selected lifetimes under different incident photon energies at 0.7eV reorganization energy, demonstrating the following characteristics of the optimal tetrads: Top: Engineering efficiency. **Mid**: Energy used to create the final charge-separated state. **Bottom**: Average inter-cofactor distances.

lifetimes longer than necessary. Only at lifetimes well past the 100ms mark of practicality do we observe significant differences between the tetrads and triads, or at photon energies lower than 1.8eV.

One of the most noticeable differences between the traces of triads and tetrads is that tetrads now display almost no noticeable phase change between the “Boltzmann” and the “Marcus” within the window of selected lifetimes that we are examining. Instead, the engineering efficiency curves of tetrads with various incident photons appear to look either completely linear, with increasing negative slopes as the incident photon energy becomes lower. This behavior is unsurprising given our understanding of the mechanistic details that account for the differences between the dyads and the triads. Specifically, the addition of yet another cofactor allows for additional distance-based stabilization of the charge-separated state, allowing the charge-separation to take further advantage of the Boltzmann stabilization. The 4th cofactor provides not only additional “free” distance in

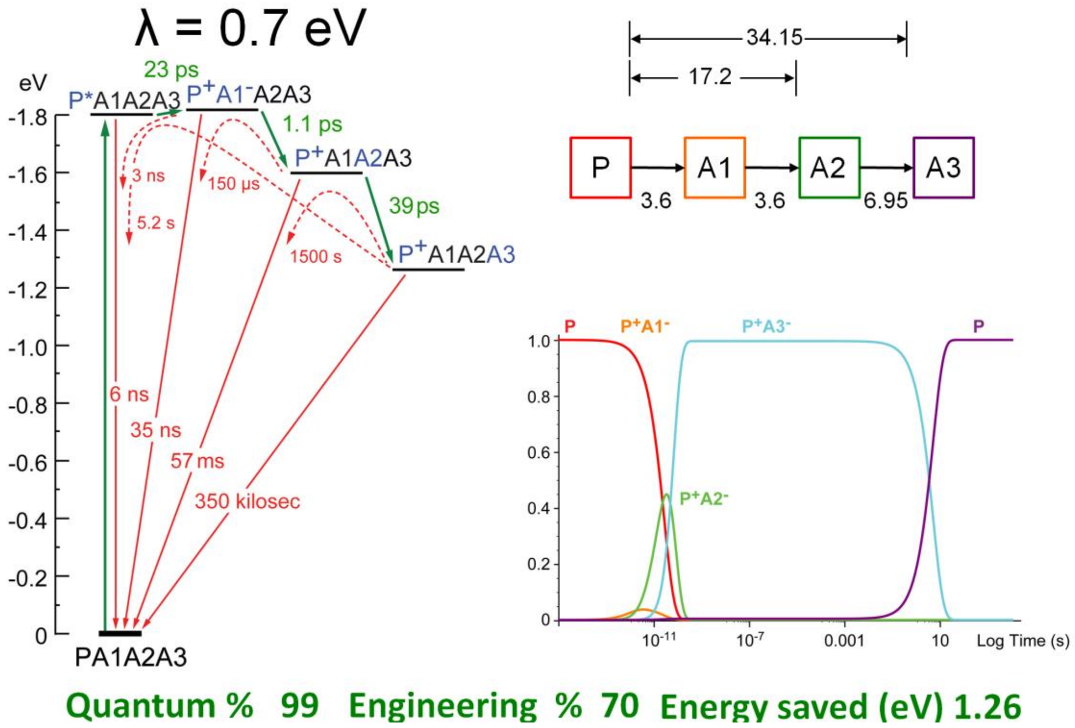


Figure 2.10: Engineering details of optimal charge-separating tetrads at uniform 0.7eV reorganization energy and 1.8eV incident photon energy. The energy diagram on the left shows the energy levels of the various CS states and the rates between them. The cartoon on the right lists the distances between the cofactors, while the kinetics profile at the bottom demonstrates the population of the various CS states at different lifetimes.

the form of cofactor diameter, but also another branching point to mediate the loss caused by premature charge-recombination. This is simply a continuation of the trend we have observed in the previous sections and therefore will not be significantly elaborated except for an exemplary representation of an optimal triad producing uniform quantum yield and theoretically maximal energy preserved, seen in Figure 2.10. The extra branching also allows the average inter-cofactor distances for the tetrads to be much smaller than the triads, as can been see in a comparison between the corresponding distance panels of Figure 2.9 and Figure 2.7. Moreover, the lifetime at which the optimal tetrads must adopt non-van der Waal contact distance is also delayed in tetrads, thanks to the additional cofactor and branching allowing more flexibilities to distribute the necessary distance stabilization when the inverted-region slowing becomes inadequate.

This extra distance-based stabilization is almost inconsequential, when only the fundamental principles of electron transfer are considered, for engineering charge-separation that last no longer

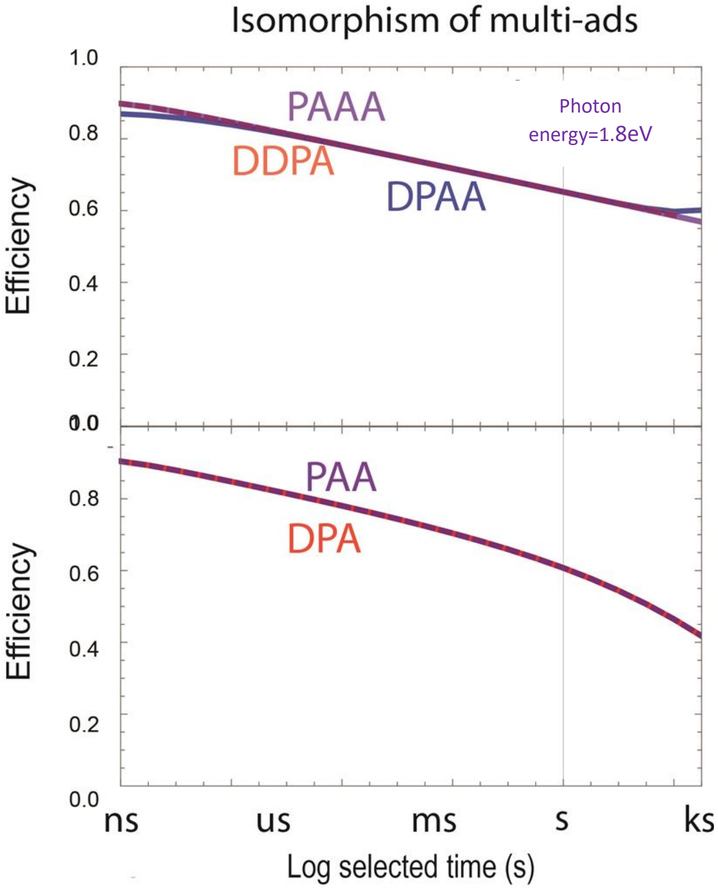


Figure 2.11: Isomorphism of various configurations of triads and tetrads. Top: traces of engineering efficiency of the 3 different possible tetrad configurations at various selected lifetimes, with the colored labels corresponding to the various forms of tetrads. Fixed photon energy of 1.8eV. **Bottom:** engineering efficiency of the 2 possible triad configurations at various selected lifetime with the same optimization algorithm under the same set of constraints. The deviation at the ks of the DPAA from the other two tretrads results from the numerical instability of the model.

than 100ms, since, as already seen in Figure 2.8, the distance-assisted Boltzmann stabilization by triads with 1.8eV incident photon only begin to fail well past 100ms.

2.8 Equivalence of donors and acceptors in triads and beyond

In our exposition of the optimal designs of triads and tetrads in the previous section, we have only examined the engineering of PAA triad and PAAA tetrad, while the DPA configuration for triad and the DPAA or DDPA configurations for tetrad exist. Here we demonstrate that the choice to focus on acceptor-only n-ads significantly simplifies the challenge of presentation as well as avoids numerical computational instabilities, without any loss of generality. That is, the optimality of engineering, measured by the metric engineering efficiency, is the same given the same set of engineering constraints for the various possible configurations of triads and tetrads, although the engineering parameters to produce such optimality might be different. In fact, we make the stronger claim that the optimization problem for all different configurations of mult-ads, as long as they share the same number of cofactors, can be considered isomorphic within the scope of this work and our engineering blueprints for optimal charge-separating device.

We first present the isomorphism for the triads, shown in the bottom half of Figure 2.11. For triads, both PAA and DPA have four microstates. More importantly, the rate coefficient matrices for the system of ODEs for PAA and DPA have the same non-zero terms. That is, electron transfer reactions between all microstates are equivalent for the two configurations of triads. Consequently, during the optimization process to discover the maximal engineering efficiency at various lifetimes, the optimal solutions of the system of ODEs, namely, the electron transfer kinetics, of the two configurations, will also be equivalent, although the parameters that generate the optimal performance will differ, since the donor cofactor sends its electron to the ground state of P. Notice that the optimal engineering efficiency curve for the PAA triad overlaps completely with the curve for the DPA triad.

For tetrads, DDPA arrangement and PAAA arrangements have the same symmetric identity as that between DPA and PAA. They both have the same number of microstates eligible for electron transfer, specifically 5, and the rate matrices are equivalent upon rearrangements in the exact same way as those of the triads. Indeed, as shown in the top half of Figure 2.11, the optimal engineering efficiency of an arbitrary DDPA and PAAA tetrad are identical throughout all timescales. However, with DPAA, there are 6 rather than 5 microstates, making DPAA mathematically not identical to the other two. However, when the optimization algorithm is applied to the DPAA, we have observed that other than small numerical fluctuations at extreme conditions(as seen at the far end of the DPAA trace in Figure 2.11 bottom), DPAA's optimal behavior is identical to that of PAAA and DDPA. Although further discussion to fully understand the differences between DPAA and PAAA/DDPA could prove to be interesting, given the scope of the current work and the myriad of other topics of interest, I have decided to limit the discussion in this thesis to the practical indistinguishibility of the different variants of tetrads and triads that could be considered in the design process.

2.9 Increasing tolerance for design parameters as n-ads become more complex

While the analysis in the previous sections have sought to identify the optimal engineering efficiencies attainable by photochemical multi-ads and the necessary engineering parameters required to achieve such optimality, we have not considered the engineering tolerance at the optimalities. That is, within what range of values of the engineering parameters prescribed by our blueprint for optimality does the engineering efficiency of charge-separation remain sufficiently close to optimal. The extent to which the engineering parameters are allowed to deviate from the optimal values without noticeable decrease in the efficiency is defined as the engineering tolerance. The tolerance is essentially a reflection of the overall smoothness of the engineering landscape; greater tolerance suggests that the engineering landscape, in terms of the parameters that influence the engineering efficiency of the photochemical device, is relatively flatter, while smaller tolerance corresponds to a more jagged landscape. Practically, engineering tolerance carries great significance since when the engineering blueprints are realized in practical designs, many additional constraints that are not considered in the formulation of the engineering blueprints in this chapter could make it impossible to actually use the prescribed optimal engineering parameters. A greater tolerance therefore corresponds to a higher likelihood that the realized design would replicate the engineering efficiency that's promised by the blueprints. In fact, strong tolerances in certain engineering parameters can be considered to be much more important in naturally occurring photosystems than the absolute engineering efficiency attainable, as will be revealed and examined in this section and in Chapter III of this thesis.

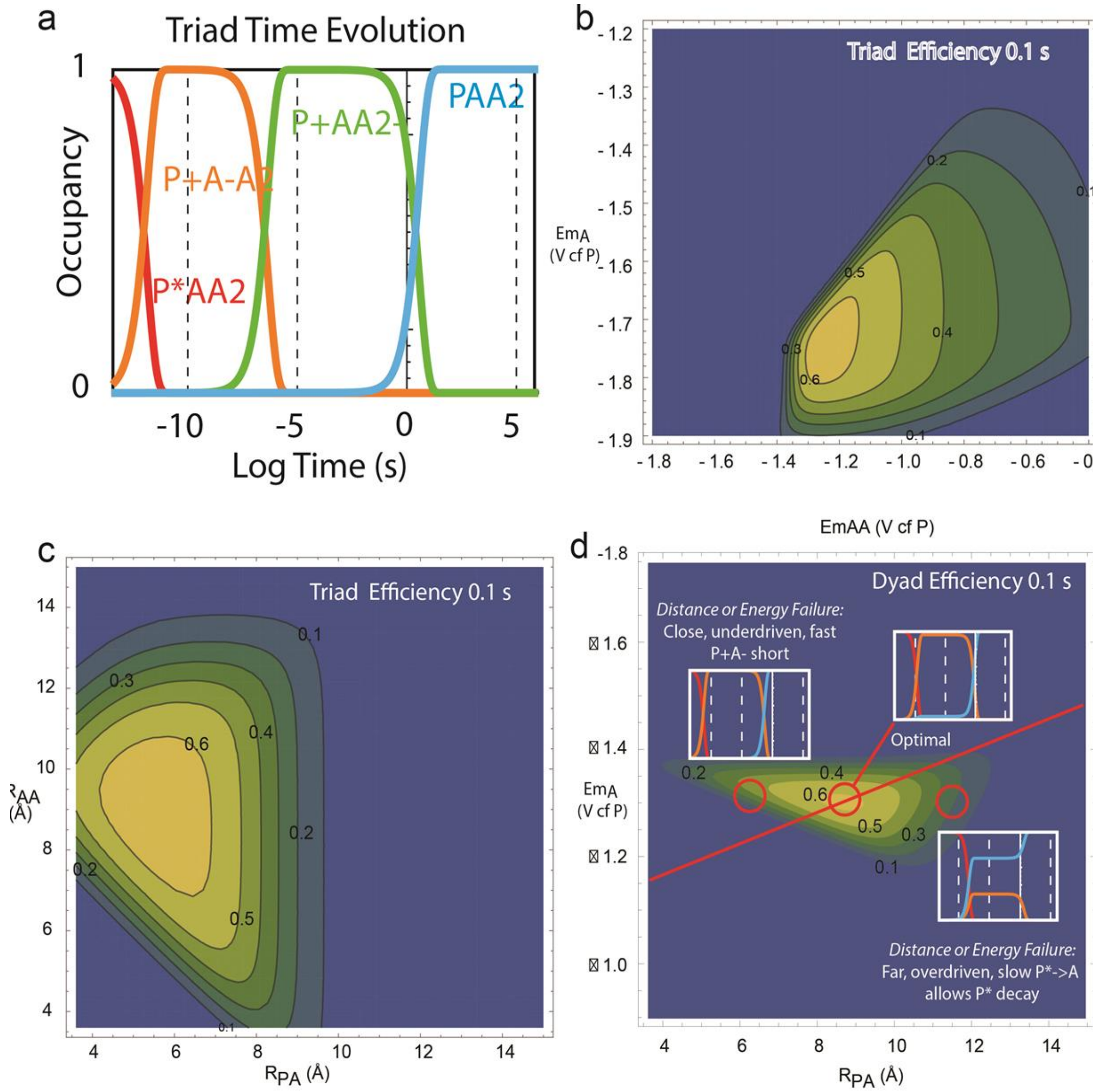


Figure 2.12: Tolerances for design parameters in dyads and triads. A: Examplary kinetic traces of the dyads and triads shown in this figure. B: engineering tolerance of the redox potentials of the two acceptors in an optimally designed triad at 100ms, with contours of equal engineering efficiency. C: engineering tolerances of the distances between the cofactors in an optimally designed triad at 100ms, with contours of equal engineering efficiency. D: engineering tolerance of the acceptor redox potential and distance to the pigment for an optimal dyad at 100ms, with the red segment bisecting the contours into two regions of dominant loss routes, further illustrated by the inserted kinetic details. The left insert represents suboptimal engineering of the dyad when the distance is too short and therefore cannot provide enough distance-based stabilization when Boltzmann stabilization is no longer adequate at 100ms. The right insert shows when too much distance stabilization slows down the charge-separating

Following the order of the analysis of the previous sections, we begin by examining the engineering tolerances of the optimal dyads. Figure 2.12d illustrates the contours of the distances and driving forces of dyads when a charge-separated state lasting 100ms is required. For the purpose of illustration we have chosen the special constraint of 0.3eV reorganization energy, since otherwise no effective charge-separation can be observed at 100ms for dyads. It is not difficult to conclude that even with such a generous constraint of 0.3eV reorganization energy, only very small deviations from the exact values of the optimal parameters are allowed before the efficiency begins to deteriorate. In fact, when the practicality of engineering is considered, in terms of both the P-A distance and the Em of the acceptor, only the optimality is allowed. That is, in a practical setting, the design of the optimally performing dyad offers no tolerance and therefore poses yet another challenge in attempts to construct such devices. Figure 2.12d also demonstrates the consequences of suboptimal engineering. When the P-A distance is to large, the yield of the CS state decreases as the forward ET slows down and gets outcompeted by the excited state relaxation rate. When the P-A distance is too small, faster charge-recombinations, largely in the form of uphill thermal repopulation of the P*, causes lower

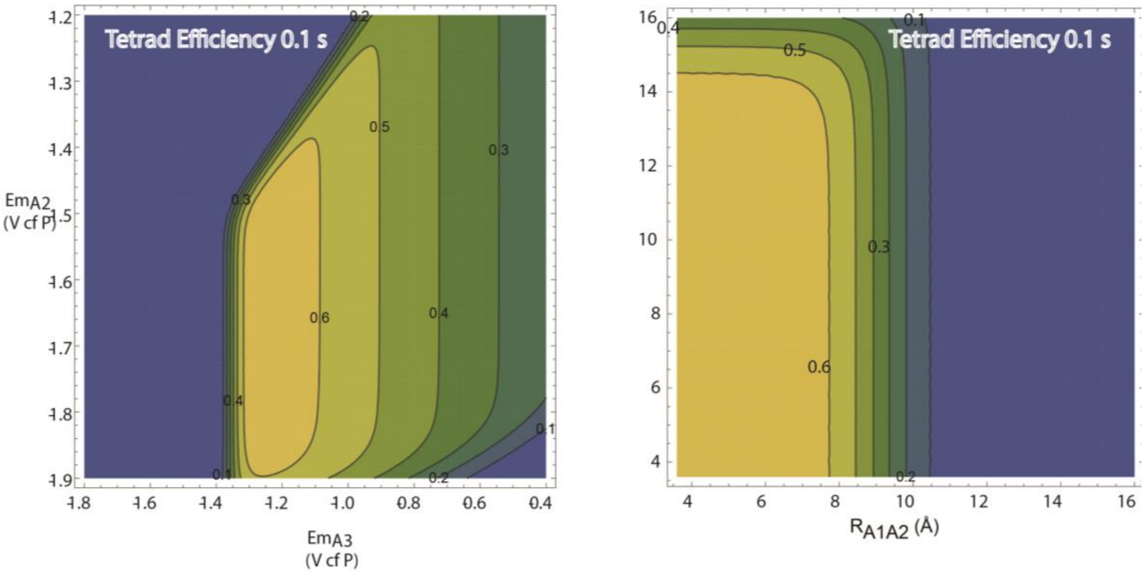


Figure 2.13: Engineering tolerances for the design parameters of the second and final acceptors in a PAAA tetrad, with an incident photon energy of 1.8eV under normal reorganization energy of 0.7eV, examined when all other parameters are fixed at optimal values and two are allowed to vary. Left: engineering tolerance for the Ems of the second and final acceptors. Right: engineering tolerances of the distances between the first-second and second-final acceptors.

yield at the selected lifetime.

It is also interesting to note that the dyads offer greater tolerances in the distance than in the Em of the acceptor. This reflects the importance of the total reliance on the Boltzmann stabilization offered by the slowing effect of the inverted-region. Any significant deviation in the Ems of the acceptor would cause the driving force to the ground state of the P to become either too great or too small, annulling the stabilization effect.

As we increase the number of cofactors from 2 to 3, not only do we see significant improvements in the absolute performance, but also some increases in the engineering tolerances, as seen in panel B and C of Figure 2.12, more so in the tolerances for the distances than the Ems of the cofactors. In fact, the "optimality plateau" for the Ems of the acceptors of the triad is only 150mV wide at maximal difference. In design practices, this range is not large enough to offer different options of cofactors since the differences in the Ems of different families of cofactors could easily be greater than 200mV. The distances offer somewhat larger flexibility, but the only practically relevant tolerance is for the A1-A2 distance with a range of 3.5Å. This range is also considered rather small, especially when the photosystem is to be engineered in a de-novo man-made protein environment as will be seen in Chapter V and VI.(*25*)

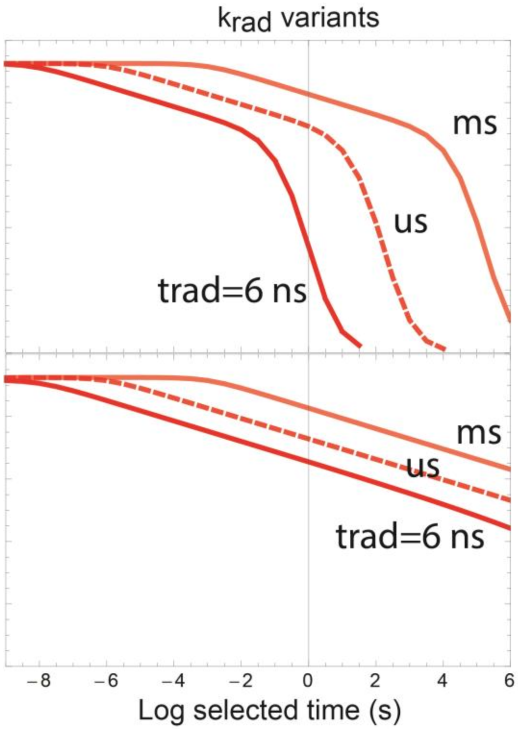


Figure 2.14: Effects of the excited state lifetimes on the optimal engineering efficiencies of photochemical devices. Top: optimal P-A dyads' engineering efficiency at lifetimes from ns to megaseconds under excited lifetimes of 6ns, 1us, and 1ms. Bottom: optimal engineering efficiencies of triads at the 3 lifetimes of excited states.

While the move from 3 to 4 does not offer the same kind of dramatic improvements in absolute optimal engineering efficiency as from 2 to 3 in the chemically relevant range of lifetimes, it makes up by providing significant gains in the engineering tolerances of the parameters. Consider the comparison between Figure 2.13 and the two triad panels of 2.12. In both the distance and energetic perspectives, the region of optimal efficiency is significantly greater for the tetrads. In fact, the level of tolerances offered by tetrads for the Ems of the cofactors is finally large enough to be meaningful in design practices. The range of 250mV for acceptor A3 and ~500mV for acceptor A2 is large enough that either different choices of different cofactors or modifications of the local environments of the platform housing the photosystem can be allowed. Similarly, the large range of values of distances for both the secondary and ternary acceptors are also large enough that allow for various attempts at designs in practice. The range of >10A of tolerance for the A2A3 distance and the >5A range for the A1A2 distance make the design of optimal tetrads a significantly less challenge than that of the tetrads.

This significant contrast between the engineering tolerances of the tetrads and triads will figure heavily in the results of the next chapter, where we explore the engineering principles when biological contexts are considered. The significantly greater engineering tolerances of tetrads allow them to provide a much more robust and resilient engineering landscape when additional engineering constraints not considered in this chapter are incorporated in Chapter III in order to establish a more specific and applicable version of the engineering blueprints for optimal photosystems in biological contexts.

2.10: Effect of the lifetime of the excited state on the engineering of photosystems

Now that we have a good understanding of the trends of engineering parameters in the shifting landscape of optimal dyads, triads, and tetrads, it is time to examine the effects of the constraint, “hard” parameters that have remained fixed so far. In the various EE/lifetime traces in Figure 2.4, the effect of increasing the constraint represented by the incident photon energy has already been demonstrated. Given a fixed, uniform reorganization energy, decreasing the incident photon energy leads to smaller stabilization of the charge-separated state via the Marcus inverted region and earlier onset of the Marcus phase, resulting in a much shorter Boltzmann phase in the efficiency/lifetime curves and therefore a harsher engineering landscape for optimal performance.

One of the most interesting and non-trivial constraint that we have not considered is the lifetime of the excited state of the pigment, although it can exert significant impact on the optimal engineering of the photosystems. As shown in Figure 2.14, increasing the lifetime of the excited state greatly increases the length of the Boltzmann stabilized phase of the efficiency/lifetime curve, suggesting that lifetime of the excited state should be extended as long as possible to minimize cost in charge-separation and therefore efficiency in the engineering. We will soon reveal in Chapter III that this statement in the engineering guideline established in this chapter, along with many others, are the result of optimization based on an incomplete set of constraints and therefore not practically viable unless specific measures are taken during the practice of constructing these photosystems to remove or alleviate those constraints. Nevertheless, it is important to remember that longer lifetime of the excited state indeed represent a significant improvement in the engineering landscape of any photosystems, and the choice to adopt a charge-separated state with a ms lifetime rather than ns is one of the central engineering devices that made our near-second charge-separation possible in the face of overwhelmingly suboptimal practical constraints, as will be described in Chapter V and VI.

2.11: Conclusion: Central engineering principles for optimal charge-separating photosystems based on the fundamental theories of electron transfers.

Throughout this chapter, we have established an abstract, cofactor-focused analytical paradigm that reliably and accurately describes the details of electron transfer kinetics of charge-separating photosystems, natural or man-made, from a small set of input variables that can be considered as the essential engineering parameters of photosystems. Using this paradigm, we have developed an elegant and straightforward algorithm that allows for fast and accurate determination of the performance of any photosystems, measured as engineering efficiency at a specific lifetime of charge-separation. We have exploited the computationally inexpensive nature of our algorithm to conduct prescriptive analysis on the engineering of ET driven charge-separating devices, delineating the engineering principles for optimal design of n-ads that are based in fundamental theories of electron transfer alone.

As we carefully studied the prescribed optimal engineering parameters of dyads, triads, and tetrads given by our n-ad paradigm model, we discovered a series of conserved mechanisms that explain dyads' consistent failure to sustain long-lived charge-separation while triads and tetrads success at such tasks. We consider these mechanisms the fundamental engineering principles of ET-based charge-separating photosystems, dictating the blueprints for optimal designs. We were able to condense the trends and patterns manifested in the optimal engineering into the following set of principles that, as long as the fundamental theories of electron transfer are the only constraints considered, allow the construction of optimal engineering of ET-based charge-separating photosystems:

(1) 60mV/decade Boltzmann stabilization represents the minimal energy cost for charge-separation that lasts for any desired length of time.
(2) Uphill-repopulation based recombination is the optimal and preferred route to dissipate the charge-separated state.
(3) Stabilization of the charge-separated by maximally slowing down the rate of direct charge-recombination via low reorganization energy and long distance leads to optimal performance at any selected lifetime.
(4) Incorporation of additional cofactors to make n-ads with n>2 introduces essential distance-based stabilization of the CS state from the diameter of the cofactor as well as multiple branching points to offer multiple loss routes to avoid drastic decay in performances.

These fundamental engineering principles prescribe optimal designs of ET-based photosystem that are characterized by a few engineering motifs, which would result in the highest quantum yield of the CS state while preserving the greatest fraction of the incident photon energy at the longest lifetime. These motifs would characterize an optimal charge-separating photosystem as a triad with the cofactors, one pigment and either two acceptors or one donor and one acceptor, arranged in a linear fashion, assembled in an environment that would be as low in reorganization energy as possible, and with redox potentials of the cofactors no greater than 60mV per each decade past the lifetime of the excited state less in energy than the excited state. The choice of triad over tetrads comes from the notion that when reorganization energy is specifically engineered to be low, thee marginal improvements of tetrads over triads do not justify the significant additional experimental effort to engineer a tetrad.

However, until this point, all of the engineering prescriptions demonstrated have been derived from optimization studies using the Moser-Dutton ruler alone, a model for the rates of electron transfer reactions. The intended application of such engineering blueprints is in protein-based biological environments. Consequently, it is necessary to evaluate whether and to what extent the engineering blueprints established in this chapter are followed by natural photosystems.

2.12: Computational methods

All analysis performed and discussed in this chapter are implemented and carried out using Wolfram Mathematica, with version 9.0 for the results obtained prior to 2016 and version 10.0 for the those obtained in 2016.

For each variant of the model n-ads ranging from dyads(P-A or D-P) to tetrads(PAAA/DPAA/DDPA/DDDP), we have defined and implemented in Mathematica a function that returns the engineering efficiency of the n-ad, which from now on will be referred to as ***E***. ***E*** takes as input variables all the engineering parameters of the n-ad, including: selected lifetime of the charge-separation, ***t***; pair-wise distances between the cofactors within the n-ad, ***R***; driving forces between the cofactors, ***ΔG***; photon energy of the pigment to generate the excited state, ***hv*;** reorganization energy of the overall electron transfer environment, ***λ***; the lifetime of the excited state, ***τ***; the diameter of the cofactors in the n-ad, ***PD***; and redox potential of the ground state of the pigment, ***E_mP***. Based on the full electron transfer kinetics energy diagram of the n-ad, we then generate the *m*-by-*m* electron transfer rate matrix **A**, where *m* is the number of available states that are electron transfer competent, as explained in section 2.4. Each element of **A**, namely, A_{ij}, refers to the rate of all the possible electron transfers, with the off-diagonal elements representing the ET rates from the jth microstate to the ith microstate and the diagonal elements the electron transfer rates leaving microstate i heading to all possible destination microstates. We then defined each of the rates included in matrix **A** using the Moser-Dutton ruler introduced in section 2.2, except for the rate between the microstate representing the excited state of the pigment and the ground state of the pigment, which is computed using the inverse of the lifetime of the excited state, as part of the engineering parameters. The implementation of ***E*** then uses Mathematica's *Eigensystem* function to find the eigenvalues and eigenvectors of the matrix **A,** which then are used to construct the general solution of the time-dependent population profile of the specific n-ad as shown in Scheme 2.1 as a sum of exponentials. The initial condition that describes the state of the n-ad upon a photoexcitation event is then used to compute the coefficients in front each of the exponential terms in the general solution calling the "*Solve"* function of Mathematica, thereby producing the special solution, which contains the time-dependent electron occupancy for each of the cofactor. The implementation of **E** then returns the value of the population of the charge-separated state(one of the microstates whose population is described by the specific solution) multiplied by the ratio of the energy preserved in the charge-separated state and the input variable ***hv*** described above. The energy preserved in the charge-separated state is computed as the difference between the final acceptor and the donor or the ground state E_m of the pigment when no donor is present.

The implementation of **E** described above allows us to perform both descriptive and predictive analysis on an n-ad, as it returns the metric of the performance of this n-ad given a set of design parameters. To implement the prescriptive analysis, we took advantage of the "NMaximize" function of Mathematica. The NMaximize function is called with **E** and a subset of the engineering parameters used to compute **E** as input variables, along with a set of constraints that restrict the possible values of the parameters to be tested during the optimization process. As described in sections 2.4 and 2.5, we treated distances and driving forces between cofactors as the soft design parameters and thus used them as the varying parameters for each calling of the NMaximize function, while keeping ***hv, λ, τ, PD,*** $\boldsymbol{E_mD}$ all fixed. To generate the plots shown in figures 2.4, 2.7, 2.9, and 2.11, we have run the NMaximize function on **E** for a set of selected lifetimes t ranging from 1ns to 1000s at the interval of every half log decade and a set of ***hv*** ranging from 3.0eV to 1.0eV at 0.4eV interval and plotted the optimal engineering efficiency as a function of the increasing selected lifetime for each ***hv.*** To examine the trend of optimal engineering efficiency with respect to other design parameters, "NMaximize" is simply called with a different set of variables at specific intervals and plotted with respect to the desired variable. For all of the analysis performed in this and the following chapter, NMaximize function is used with the option of "DifferentialEvolution" as the selected method of optimization with "RandomSeed" flag set to 1.

To produce the contour plots shown in Figure 2.12, **E** is evaluated at a set of points representing a range of two of the many design parameters with other parameters are held constant either at optimal values or values specified by the user.

To generate the trike plots shown in Figure 2.6,2.8, and 2.10. a modified version of **E** was implemented, which we will term **P**. Rather than returning the numerical value of the engineering efficiency, **P** returns the plot of the time-evolution of the populations of all the microstates involved in the electron transfer as well as the rates of each individual electron transfer reaction involved in the n-ad, with the uphill-based recombination rate computed as shown in Figure 2.5.

2.13: References

1. *ChemInform*, in press, doi:10.1002/chin.199201368.
2. A. R. Cashmore, J. A. Jarillo, Y.-J. Wu, D. Liu, Cryptochromes: Blue Light Receptors for Plants and Animals. *Science (Washington, D.C.); (United States)*. **284**, 760–765 (1999).
3. C. C. Moser, C. C. Page, R. Farid, P. L. Dutton, Biological Electron Transfer. *Journal of Bioenergetics and Biomembranes*. **27**, 263–274 (1995).
4. R. R. Birge, Photophysics of Light Transduction in Rhodopsin and Bacteriorhodopsin. *Annual review of biophysics and bioengineering*. **10**, 315 (1981).
5. A. T. Ulijasz, R. D. Vierstra, Phytochrome Structure and Photochemistry: Recent Advances Toward a Complete Molecular Picture. *Current opinion in plant biology*. **14**, 498–506 (2011).
6. C. C. Moser, J. M. Keske, K. Warncke, R. S. Farid, P. L. Dutton, Nature of Biological Electron Transfer. *Nature*. **355**, 796–802 (1992).
7. R. A. Marcus, Chemical and Electrochemical Electron-Transfer Theory. *Annual Review of Physical Chemistry* (1964).
8. J. J. Hopfield, Electron Transfer Between Biological Molecules by Thermally Activated Tunneling. *PNAS*. **71**, 3640–3644 (1974).
9. C. C. Moser, C. C. Page, P. L. Dutton, Tunneling in PSII. *Photoch Photobio Sci*. **4**, 933–939 (2005).
10. C. C. Moser, P. L. Dutton, in *Photosystem I* (Springer Netherlands, Dordrecht, 2006), vol. 24 of *Advances in Photosynthesis and Respiration*, pp. 583–594.
11. T. J. Meade, H. B. Gray, J. R. Winkler, Driving-Force Effects on the Rate of Long-Range Electron Transfer in Ruthenium-Modified Cytochrome C. *Journal of the American Chemical Society*. **111**, 4353–4356 (1989).
12. S. Santabarbara, P. Heathcote, M. Evans, Modelling of the Electron Transfer Reactions in Photosystem I by Electron Tunnelling Theory: the Phylloquinones Bound to the PsaA and the PsaB Reaction Centre Subunits of PS I Are Almost Isoenergetic to the Iron–Sulfur Cluster FX. *Biochim. Biophys. Acta* (2005).
13. D. Zhong, Electron Transfer Mechanisms of DNA Repair by Photolyase. *Annu.Rev.Phys.Chem* **66**, 691–715 (2015).
14. D. Noy, C. C. Moser, P. L. Dutton, Design and Engineering of Photosynthetic Light-Harvesting and Electron Transfer using Length, Time, and Energy Scales. *Biochimica et Biophysica Acta (BBA)*. **1757**, 90–105 (2006).
15. T. Wydrzynski, W. Hillier, B. Conlan, Engineering Model Proteins for Photosystem II Function. *Photosynth Res*. **94**, 225–233 (2007).
16. B. P. Fingerhut, W. Zinth, R. de Vivie-Riedle, The Detailed Balance Limit of Photochemical Energy Conversion. *Phys. Chem. Chem. Phys.* **12**, 422–432 (2009).
17. A. Punnoose *et al.*, Fundamental Limits on Wavelength, Efficiency and Yield of the Charge Separation Triad. *PLoS ONE*. **7**, e36065 (2012).
18. B. P. Fingerhut, W. Zinth, R. de Vivie-Riedle, Design Criteria for Optimal Photosynthetic Energy Conversion. *Chemical Physics Letters*. **466**, 209 (2008).
19. A. R. Finkelmann, M. T. Stiebritz, M. Reiher, Kinetic Modeling of Hydrogen Conversion at [Fe] Hydrogenase Active-Site Models. *J Phys Chem B*. **117**, 4806–4817 (2013).
20. A. Bar-Even *et al.*, The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry*. **50**, 4402 (2011).
21. M. R. Gunner, D. E. Robertson, P. L. Dutton, Kinetic Studies on the Reaction Center Protein From Rhodopseudomonas Sphaeroides: the Temperature and Free Energy Dependence of Electron Transfer Between Various Quinones in the QA Site and the Oxidized Bacteriochlorophyll Dimer. *The Journal of Physical Chemistry*. **90**, 3783–3795 (1986).

22. M. R. Gunner, P. L. Dutton, Temperature and -DeltaG Degree Dependence of the Electron Transfer From BPh to Q_A In Reaction Center Protein From Rhodobacter sphaeroides with Different Quinones as Q_A. *J. Am. Chem. Soc.* **111**, 3400 (1989).
23. S. Fukuzumi, Artificial Photosynthesis for Production of Hydrogen Peroxide and Its Fuel Cells. *Biochim. Biophys. Acta*. **1857**, 604–611 (2015).
24. S. Fukuzumi, K. Ohkubo, T. Suenobu, Long-Lived Charge Separation and Applications in Artificial Photosynthesis. *Accounts Chem Res*. **47**, 1455 (2014).
25. T. A. Farid *et al.*, Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).

Chapter III: The universal catalytic quartet and its subversion and rewriting of the engineering blueprints of charge-separating photosystems

3.1: Introduction:

While in the previous chapter we developed a comprehensive set of engineering blueprints and guidelines for optimal photo-activated charge-separating device based purely on the fundamental rules of physics governing electron transfer, our goal remains to carry out designs of highly efficient charge-separating photosystems in biologically relevant environments such as the interior of small, hydrophilic proteins. Consequently, here in this chapter we attempt to examine whether and how much the engineering principles discovered in the previous chapter continue the development of the engineering guidelines for optimal photosystems by discovering and understanding more specific and practical constraints imposed upon naturally occurring electron transfer processes by biophysical, biochemical, and biological environments and conditions. This allows us to correct, refine, and enrich the engineering blueprints we obtained from Chapter II and to carry out designs that are practical and applicable to biological contexts. In order to discover these additional engineering constraints enforced by the biological setting and the corresponding modifications to the existing design principles to achieve optimality, the ideal object to study is nothing other than the charge-separating multi-ads that are engineered by nature through the process of evolution, the various biological electron-transfer based photosystems responsible for some of the most fundamental processes of living organisms.

Many biological functions essential to the survival of organisms across different kingdoms are accomplished by electron transfer based photoactivated charge separation processes, ranging from DNA repairs(*1*, *2*) to photosynthesis, exploiting a significant portion of the solar spectrum with high efficiency, as seen in Figure 3.1.(*3*) These processes are driven by a series of small molecule cofactors encapsulated in complex protein frameworks called photosystems. Despite the seemingly enormous differences in the function and apparent structures of the protein frameworks representing them, these biological photosystems share a highly conserved structural engineering motif as their core functional unit. A catalytic quartet of four redox active and electron transfer

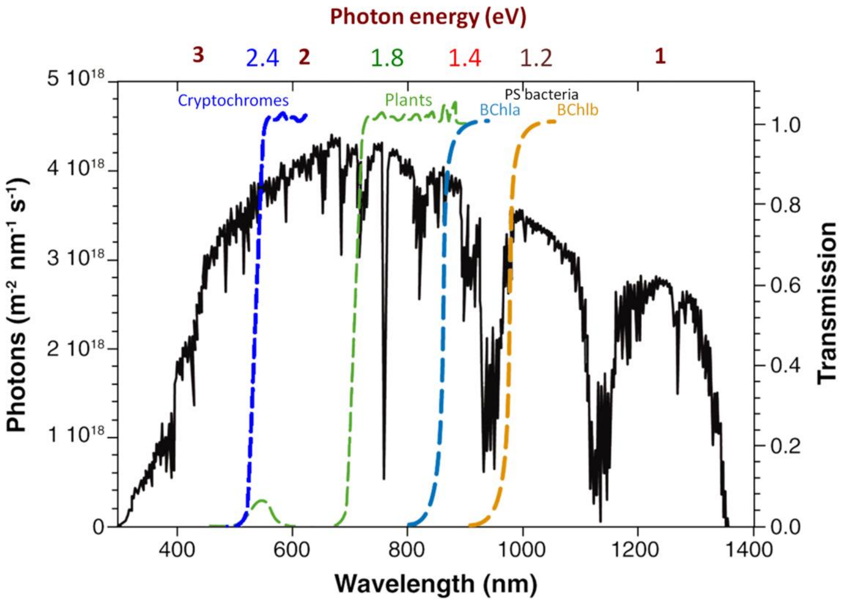


Figure 3.1: Natural photosystems cover a broad range of the solar radiation spectrum. The black trace represents the black-body radiation of the modern sun at visible and near- to mid-IR range, shown as both raw photon flux on the left vertical axis and transimission on the right vertical axis. The four colored traces represent the transmission spectra of various prominent natural photosystems. Namely they indicate wavelengths at which the pigments of natural photosystems absorb photons and produce excited singlet states to drive charge-separation. The numbers on the top indicate the energy of the absorbed photon as well as the energy of the excited state of the corresponding photosystems. Reproduced and modified from Blankenship et al(3).

competent cofactors tightly bound to the center of the photosystems form the basis of the functional core unit. This functional core is highly conserved in various organisms, as summarized in Figure 3.2, and all share the same P-A-A-A organization of cofactors(*4*).(*5-8*)That is, the catalytic quartet consists of a pigment cofactor followed by three acceptor cofactors that relay the excited electron produced by photo-excitation of the pigment cofactor to form the final charge separated state. In the bacterial photosystems, this quartet consists of the light-activated bacterial chlorophyll dimer as the pigment, and the bacterial chlorophyll monomer, bacterial pheophytin, and Qa as the chain of three acceptors. Similarly, in the plant photosystem I and II, P680/P700 dimer serve as the pigment, a bacterial chlorophyll serve as the primary acceptor, a bacterial pheophytin play the role of the secondary acceptor A2, while a quinone(QA) that does not diffuse away and remains integral to the photosystem. Additionally, signaling and DNA repair enzymes in the crytochrome/photolyase family contains a photosensitive Flavin as the P

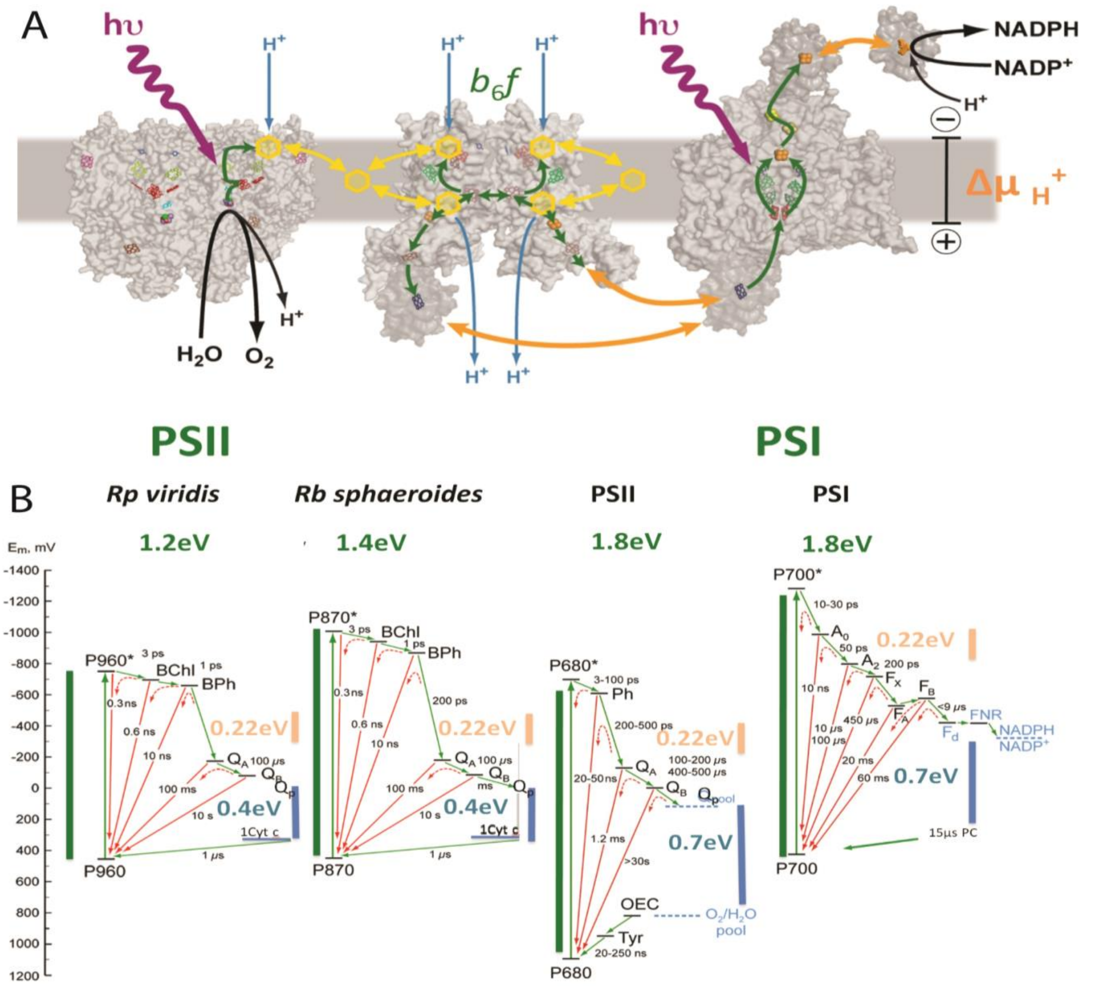

A
hυ
H⁺
H⁺
H⁺
b6f
hυ
NADPH
NADP⁺
H⁺
⊖
Δμ H⁺
⊕
H⁺
H₂O
O₂
H⁺
H⁺
PSII
PSI
B
Rp viridis
Rb sphaeroides
PSII
PSI
1.2eV
1.4eV
1.8eV
1.8eV
Em, mV
-1400
-1200
-1000
-800
-600
-400
-200
0
200
400
600
800
1000
1200
P960*
3 ps
BChl
1 ps
BPh
0.3ns
0.6 ns
10 ns
0.22eV
QA 100 μs
QB
Qp
100 ms
10 s
0.4eV
1Cyt c
1 μs
P960
P870*
3 ps
BChl
1 ps
BPh
0.3ns
0.6 ns
10 ns
200 ps
0.22eV
QA 100 μs
QB
ms
Qp
100 ms
10 s
0.4eV
1Cyt c
1 μs
P870
P680*
3-100 ps
Ph
200-500 ps
0.22eV
QA
100-200 μs
400-500 μs
QB
Qpool
20-50ns
1.2 ms
>30s
0.7eV
OEC
O2/H2O pool
Tyr
20-250 ns
P680
P700*
10-30 ps
A0
50 ps
A2
200 ps
0.22eV
Fx
FB
10 ns
FA
<9 μs
FNR
Fd
NADPH
NADP⁺
10 μs
100 μs
450 μs
20 ms
60 ms
0.7eV
15μs PC
P700

Figure 3.2: The universal PAAA organization of the functional cores of natural photosystems of various species. (A) The cofactors of natural photosystems are embedded within complex matrices of natural reaction centers, significant portion of which is membrane-bound. **B**: The essential cofactors involved in the charge-separation processes of various natural reaction centers, with their redox midpoint potentials shown relative to a common scale. Electron transfer reactions of the natural photosystems are shown as arrows of different colors.

and a chain of three tryptophans as the acceptors.

This consensus is quite surprising, in light of the optimal engineering blueprints we have formulated in chapter II, where we demonstrated that the marginal benefit, measured in differences in the engineering efficiency gained from adopting a 4-cofactor linear tetrad design rather than a triad, is significantly less than that of the transition from dyads to triads. Moreover, other than extreme cases where the engineering constraints(incident photon energy or reorganization energy) become very prohibitive, triads offer performances that are adequate and not significantly worse than those of tetrads, especially for charge-separation that last hundreds of milliseconds for the purpose of driving subsequent catalytic reactions. The design blueprints based purely on the physics of electron transfer, therefore, creates the expectation that triads are sufficiently competent and would be adopted widely in nature. However, as already become evident, this is not the case. This startling observation thus prompts us to examine the catalytic quartets of natural photosystems with greater care in order to discover trends and underlying principles that would allow us to expand, correct, and improve our design guidelines.

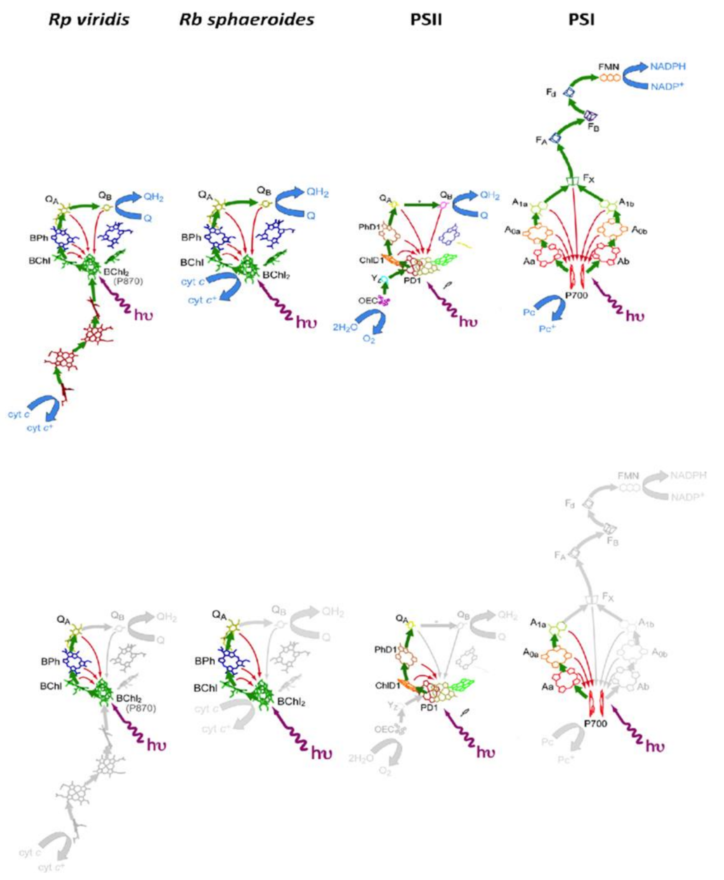


Figure 3.3: Universally conserved catalytic quartet in photosystems across organisms. Top: organization of electron transfer active cofactors in four representative photosystems of various organisms. Bottom: The highly conserved quartet consisting of cofactors structurally and functionally integral to the photosystems.

A closer examination of photosystems that contain the core quartet reveals the following list of characteristics common to all:

1. The quartet consists of a single photoactivatable pigment cofactor and a chain of three acceptors or donors.
2. The quartet functions solely through tunneling-based electron transfer reactions.
3. Three out of the four cofactors of the quartet belong to one family of chemical molecules while the remaining member belongs to another family with vastly different chemical and redox properties.
4. The presence of three acceptors/donors in a chain with close spacing between the first two acceptors promotes ultra-fast formation of the final charge separated state
5. The quartet favors a curved configuration rather than linear arrangement of the chain of cofactors.
6. The electron transfer between the penultimate and the final acceptor involves a significant driving force.
7. Processes involved in the quartet not coupled to any bond-formation or breaking.

Together, these traits suggest that, in addition to the apparent similarity in terms of cofactor organization, the catalytic quartets of natural photosystems share highly conserved fundamental characteristics that would imply the presence of a novel set of engineered principles not considered in Chapter II, that govern the design of the natural photosystems. In addition, the lack of naturally occurring triads and the abundance of the tetrads in natural photosystems led us to conclude that the engineering blueprints we developed throughout the previous chapter are in fact incomplete, inadequate, or simply incorrect, due to the fact they are derived and formulated solely from the fundamental physics of electron transfer while all naturally occurring photosystems function as components of living organisms. We therefore hypothesize that the biological context of natural photosystems exert significant influence on the engineering of photosystems, imposing additional constraints, thereby establishing novel engineering guidelines that revise or replace those formulated previously.

In order to reveal these additional constraints and uncover the novel engineering guidelines, we aim to study the natural photosystems using the same methodologies and approach we used for the abstract photohchemical multi-ads in the previous chapter. That is, we construct accurate but computationally efficient models of the catalytic quartets of the natural photosystems in order to gain systematic understanding of the transfer kinetics of the photosystems and their performances, measured in engineering efficiencies, enabling us to analyze trends and patterns and finally identify and establish the engineering principles in addition to the physics of electron transfer reactions. In the past, efforts by the Dutton group have successfully reproduced various important components of the kinetics of complete natural reaction centers with all of their cofactors of primary, secondary, and even tertiary functions by applying the Moser-Dutton ruler using the distances and driving forces known from experiments and fitting the remaining less straightforward parameters using the descriptive and predictive stages of our analytical paradigm introduced in Chapter II(*9-11*). However, in the above cited studies, our models have only accurately matched the rates that had been experimentally determined at the time of their publication. Many rates, even the existence of certain ET reactions and their corresponding intermediate species, remained elusive and unresolved. These happened to be electron transfer reactions during the initial events of the catalytic quartets of the biological photosystems. Their transient nature, lasting for as long as 3ps, coupled with the fact that the rates of their formation being slower than the rates of their disappearance, made them difficult to detect. Because of the lack of available experimental data at the time of the previous studies, the initial rates of electron transfer have therefore not received the attention they deserve, as will be seen in sections below. Recent experimental advances(*12-15*) have successfully elucidated most of these early events of electron transfer in biological photosystems and presented us with additional impetus to provide a complete model of the core catalytic quartets of natural photosystems, in addition to uncovering the universal engineering principles that led to this common design motif.

Here in this chapter we begin with constructing complete and accurate kinetic models of the core catalytic quartets of all of the listed biological RCs, using the most up-to-date distances and driving forces, referencing them to the consensus electron transfer rates determined by experiments(*12-15*) and applying our Moser-Dutton ruler based methodology from the previous chapter along with various necessary modifications and improvements as explained below. Our model reveals that the catalytic quartets of natural photosystems employed a surprising set of novel engineering approaches that contradict and violate the engineering guidelines for optimal multi-ads. These engineering approaches include:

1: the ultra-fast, ~200-ps lifetime of the singlet excited state of the pigment cofactor,

2: the uniform high reorganization energy of 0.7eV that characterizes the protein environment of the natural photosystems,

3: the close spacing between the first three cofactors in the quartet that limits inter-cofactor distances to below 5.5A,

4: the highly pronounced curved arrangement of the cofactors that significantly decreases the distances between the acceptors and the pigment while still maintaining the tetrad design, and,

5: the sub-optimal choices for the redox potentials of the cofactors when no external potential is present.

These engineering approaches produced tetrads that violate and contradict the engineering guidelines from Chapter II, characterized by the ultra-fast formation of a highly stable final CS state with near-unity yield whose engineering efficiency further increases under physiological membrane potential and who undergoes direct charge-recombination to the ground state of the pigment without revisiting the various highly unstable high energy, harmful intermediate CS states that are made highly transient and short-lived as well as difficult to reform via thermal repopulation. Careful studies of these novel engineering approaches and devices allow us to identify a single, essential, and previously not considered engineering constraint based in the biological context of natural photosystems: protection against damages to the structural and functional elements of the photosystem by high energy, unstable side species of the electron transfer active cofactors. We discover that this single engineering principle is responsible for all of the novel engineering devices employed by natural photosystems that are not prescribed by the engineering blueprints formulated in Chapter II.

Finally, we demonstrate that we can apply the same type of prescriptive optimization algorithms used in Chapter II to formulate a new set of engineering guidelines with our revised engineering constraints based on the combination of biological engineering constraints and fundamental theories of electron transfer. Such application generates engineering prescriptions for a new set of optimal charge-separating devices that are characteristically highly similar to the catalytic quartets observed in nature. That is, the new optimal photosystems returned by our revised algorithm share many of the common traits listed above, thereby proving that our reformulated design guidelines have largely captured the underlying engineering principles that resulted in the universal catalytic quartet among natural systems. However, in spite of such organizational similarities, there are significant potentials for improvements in the engineering efficiency of natural photosystems that can be unlocked by designs that parameters provided by the prescriptive studies, even when the protein environments used for the design become even less hydrophobic than those of the natural photosystems. Furthermore, we show that if we only partially revise our engineering blueprints from the previous chapter to accommodate only some of the biological constraints while leaving others to be accounted for by practical experimental details, we have the capability to improve the performance of photosystems to an unprecedented level of high efficiency. In the end, I provide a look ahead to the second half of this thesis, showing that maquettes, a minimalist, man-made protein scaffold, would perfectly satisfy the need for a platform on which the engineering blueprints and guidelines we have formulated in both Chapter II and the current chapter can be tested and realized with the highest level of control and efficiency. I show that the application of maquette would significantly facilitate the optimal engineering of a catalytic quartet in water-soluble protein environment.

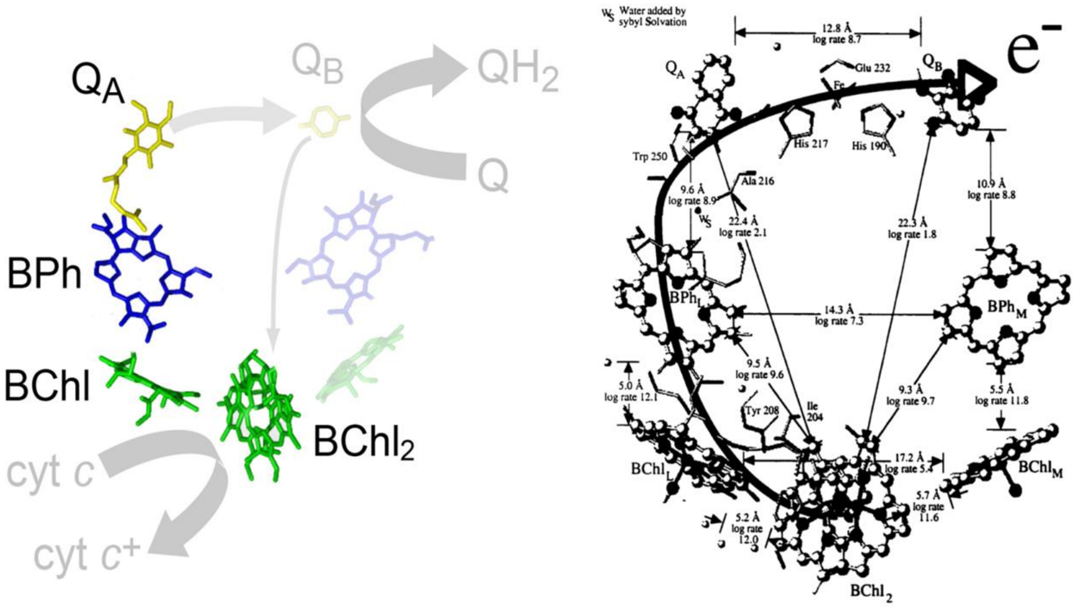

Q_A
QB
QH2
Q
BPh
BChl
BChl2
cyt c
cyt c+
WS Water added by sybyl Solvation
12.8 Å log rate 8.7
e⁻
QA
Glu 232
QB
Fe
His 217
His 190
Trp 250
Ala 216
9.6 Å log rate 8.9
WS
22.4 Å log rate 2.1
22.3 Å log rate 1.8
10.9 Å log rate 8.8
BPhL
14.3 Å log rate 7.3
BPhM
9.5 Å log rate 9.6
5.0 Å log rate 12.1
Tyr 208
Ile 204
9.3 Å log rate 9.7
5.5 Å log rate 11.8
17.2 Å log rate 5.4
BChlL
BChlM
5.7 Å log rate 11.6
5.2 Å log rate 12.0
BChl2

Figure 3.4: Structural details of the electron transfer active cofactors of Sphaeroides reaction center and the distances separating them. Left: cartoon representation of the path of the electron from P($BChl_2$) to A3(Q_A), outlining the catalytic quartet. **Right:** detailed distances between pairs of electron transfer active cofactors and the protein environment surrounding these cofactors. Reproduced with permission from Chris Page's thesis.

3.2 Results and Discussion:

3.2.1 Rhodobacter Sphaeroides reaction center as the reference model system par excellence for illustrating the catalytic quartet model

It is very difficult to construct and test a highly accurate model for the kinetics of electron transfer within the catalytic quartet without first having a well-studied and fully understood natural quartet as reference. Among all the naturally occurring photosystems, those of the Archae family, specifically of the species Viridis and Sphaeroides have been best studied and understood. High resolution crystal structures of the bacterial reaction centers have become available since late 1980s(*6*, *7*, *16*, *17*). Soon after the structure of the reaction center of Sphaeroides became available, researches have independently recognized the special importance of a subset of the cofactors within the Sphaeroides reaction center that are integral to both its structure and function: the Bacterial chlorophyll dimer(abbreviated from now on as P), the bacterial chlorophyll monomer(Ba), the bacterial pheophytin(H), and the non-diffusive quinone(Q_A). Numerous mutants have been generated to study almost all aspects of these reaction centers, mostly focusing on the electron transfer kinetics and how it is impacted by changes in both the protein surrounding and the properties of the cofactors involved.

Moreover, the electron transfer kinetics of the Sphaeroides reaction center has been described almost completely over the years. Well recorded rates for each step exist in multiple sources of literature. The once highly controversial details of the initial steps(*18-22*)of the electron transfer mechanisms have also recently seen significant advances that help bring about a rough consensus, thereby making Sphaeroides RC one of the rare natural photosystems whose electron transfer kinetics of the core cofactors are known to very minute details at each individual step. (*12-15*) Additionally, extensive studies that modify the components of the bacterial reaction center have also been performed, including mutations that result in measurable changes in the midpoint potentials of the ground state of the pigment, the Pheophytin, and the Qa cofactors, (*13*, *23*, *24*)further validating the existing experimental rates that together produce the complete kinetics of electron transfer and charge separation.

Given the ample experimental results as references, the conserved catalytic quartet, the cofactors integral and central to the reaction center of R.Sphaeroides, serves as the most appropriate model to understand the engineering principles governing the recurring natural catalytic quartet. Therefore, we begin our analysis of the engineering principles of the catalytic quartet by modeling the electron transfer kinetics of the conserved catalytic quartet of the Sphaeroides bacterial photosystem as a PAAA tetrad, with the same programing details as described in Chapter II. Unlike the past studies of the Dutton group on the kinetic details of natural photosystems(*9*, *10*)which focused on exploring parameters that would allow the models to reproduce the experimentally observed electron transfer rates, in the present study the main goal is to uncover the underlying interaction of already understood engineering principles of photosystems and novel principles not included in the existing guidelines. Consequently, our model treats the catalytic quartet of Sphaeroides reaction center as an instance of the PAAA tetrad with preset parameters obtained from the crystal structure and electrochemical studies(*7*, *25*)on the energetic details of the reaction center. We model the bacterial chlorophyll dimer as the pigment, monomer B_A as A_1, H as A_2, and Q_A as A_3. This representation of the catalytic quartet of the Sphaeroides reaction center allows us to apply all three stages of our analytical paradigm, namely, descriptive, predictive, and prescriptive as discussed in Chapter II to the Sphaeroides RC. The application of the first two stages of the analysis to the Sphaeroides tetrad, when compared to the results from Chapter II, allow us to discover and examine the various novel engineering principles that are not provided by the fundamental physics of electron transfer. With the prescriptive stage, we gain the advantage of looking at the performance of and deviation from optimality of the natural photosystems under various levels of constraints imposed upon by newly discovered engineering principles, beginning with the minimal set of constraints that represent the fundamental physics of electron transfers, equivalent to what we used to obtain the engineering principles in Chapter II. For each new constraint

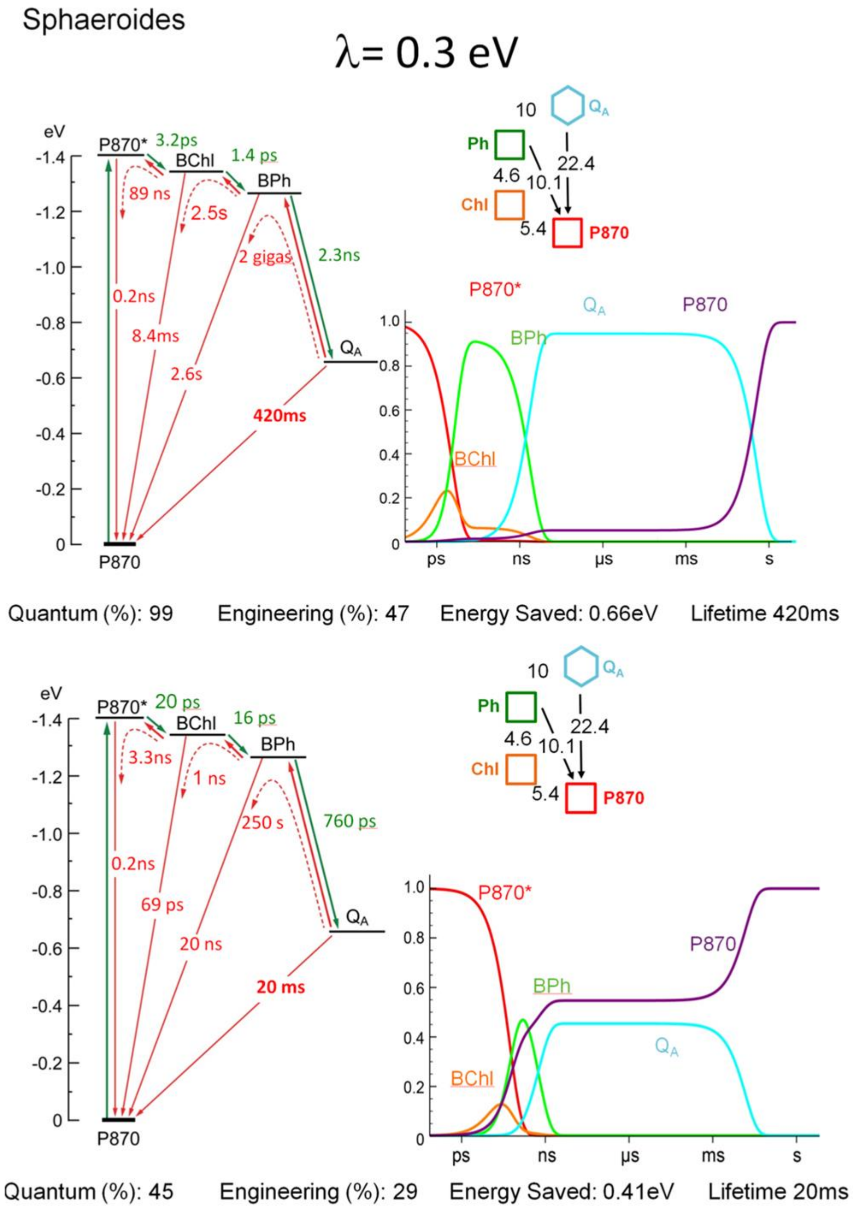


Figure 3.5 Predicted electron transfer kinetics of Sphaeroides catalytic quartet using uniform reorganization energies. Top: Energy levels, rates of ETs, distances, and kinetic profile generated using a uniform lambda of 0.3eV. **Bottom:** same as top, generated using a uniform lambda of 0.7eV.

discovered using the previous two stages of the analytical paradigm, we will update our algorithm to reflect the effect of the constraint on the performance of the model quartet. That is, we incorporate the new constraints into our model by fundamentally changing the way we compute the engineering efficiency from the input parameters. The details of how each additional constraint affects the implementation of the algorithm will be explained individually in the following sections. This allows us to consider the catalytic quartet of the Sphaeroides reaction center not as a special entity but rather as one of the many possible configurations of a general charge-separating tetrad. Therefore, we can utilize the structural and functional information of the reaction center as an initial condition/starting point for optimization studies under existing and novel constraints. By examining the extent of optimality and deviation from optimality of the Sphaeroides RC under various levels of constraints, especially from the trends in the engineering parameters, we can evaluate to what extent does our revised collection of engineering principles approach the real engineering blueprints nature followed to produce the catalytic quartet we have now; more importantly, we can examine the effects of artificially removing individual or combination of these constraints, exposing the catalytic quartet to potential risk and vulnerabilities and seeing what kind of performance gain results from such modification.

3.2.2 Resolving the apparent paradox within the initial electron transfer rates in R.Sphaeroides RC

In the process of establishing an accurate model for the Sphaeroides quartet, we discovered the first major biology-driven revision to our established engineering principles from Chapter II. To establish our model, we obtained the necessary parameters of the bacterial reaction centers from the canonical crystal structures and the seminal studies. (*7*, *25*) The distances and the redox potentials/driving forces of the catalytic quartet of Sphaeroides reaction center are shown in Figure 3.4. Given these parameters, we generated the expected electron transfer kinetics of the quartet at various reorganization energies uniform across the entire reaction center, presented in the forms of trikes as those shown in Chapter II. The results are summarized in Figure 3.5. A careful comparison to the experimentally determined rates revealed that, unlike the analysis we performed in Chapter II, no single uniform reorganization energy can fully and completely correctly account for the experimental rates, and neither can a special, low reorganization energy of 0.3eV for the initial step. The discrepancy mainly involves the initial steps of the charge separation, as shown in greater detail in Figure 3.6 and summarized in Table 3.1, with step-by-step comparison to the experimentally determined electron transfer rates. Here it shows the expected electron transfer rates of the initial charge separation between the pigment P780* and the primary acceptor, Chla, and the charge recombination from the Chla- to the ground state of P780+ at both low and high reorganization energies, 0.3eV and 0.7eV respectively. We noticed that while the experimentally determined value of the forward electron transfer rate from the singlet excited state of the pigment to the $\mathrm{Chl_A}$ can be described by a Marcus curve with a reorganization energy of 0.3eV, the corresponding charge recombination can instead only be described by the Marcus curve with a reorganization energy of 0.7eV. This apparent contradiction indicates that neither the simple uniform reorganization energy model, nor the model where the reorganization energy for the initial step is held extremely low at 0.3eV, long deemed to be the correct model for the ET kinetics of natural RCs(*26*-*28*)is successful in delivering a satisfying explanation of the experimental results. Indeed, a low initial reorganization energy model cannot account for the rapid direct recombination from the Chla to the ground state of P780+. Similarly, a low reorganization energy cannot account for the 15ns recombination rate from the Pheophytin to the P780 ground state. At the same time, as shown in Figure 3.6, a uniform lambda of 0.7eV fails to accurately reproduce the observed rates of Sphaeroides reaction center. Consequently, a different and previously encountered model is needed to accurately describe all the rates in the reaction center.

Table 3.1: Experimental and modeled rates of initial electron transfer of Sphaeroides RC.

ET step	**Experimental rates**	**Theoretical rates @ 0.3eV**	**Theoretical rates @ 0.7eV**	**Theoretical 0.3eV/0.7eV relaxable lambda mixture**

P* -> B_A	3.5ps	3.7ps	69ps	3.7ps
B_A^- -> P^+(direct recombination)	200-700ps	110ms	100ps	100ps
B_A^- -> P^+(uphill thermal recombination)	NA	8ns	8ns	66ns
B_A^- -> H_A	0.9ps	1.2ps	25ps	1.2ps
H_A^- -> P^+(direct recombination)	20ns	1.3ms	12ns	12ns
H_A^- -> P^+(via uphill thermal repopulation of excited B^-)	Possibly 4ns (from the second of the multi-exponential fits made by Gibasiewicz)	80ns(assuming a further 8ns up-and-over recombination from B_A^-)	5ns	130ns
H_A^- -> Q_A	200ps	4ns	400ps	4ns
Q_A^- -> P^+ (direct recombination)	100ms	1s	50ms	50ms
Q_A^- -> P^+ (via uphill-thermal repopulation of H_A^-		120s	30s	13s

3.2.3: A combination of unrelaxed and fully relaxed protein dielectrics as the only appropriate model for experimentally observed ET kinetics.

As seen in the previous section, both models that use uniform reorganization energy and a special low reorganization energy for the initial step fail to explain the all the experimentally measured electron transfer kinetics in these photosystems, especially the ultra-fast initial steps. Here we propose a time-dependent model for the dielectric protein environment, manifested in our model as the reorganization energy, where an uniform high reorganization energy of 0.7eV fully describes the completely relaxed dielectric protein environment at timescales slower/longer than 10ps, while at timescales faster/shorter than 10ps, a low reorganization energy of 0.3eV describes yet unrelaxed dielectric protein environment. The central assumption of this model uses a well-proven notion about protein dielectrics when ultra-fast changes in charge distribution occurs within the protein environment, typical of fast electron transfers as seen in the initial steps of bacterial reaction centers and type 2 reaction centers such as PSII.

As the dominant tunneling medium for electron transfer reaction in protein, as introduced in Chapter I of this thesis, the protein environment surrounding the acceptors and donors of the electron transfer reaction exert significant influences upon the ET rates via its dielectric properties. In most circumstances, the dielectrics of protein environment can be treated statically and regarded almost as isotropically uniform, as in the algorithm we used throughout Chapter II. However, in electron transfer reactions of natural reaction centers such as those described above, electrostatic charges on the amino acid residues that make up the protein dielectric environment experience redistribution as a result of protein's dynamic response to the transfer of charges that occur, however, only on the order of >10ps. When electron transfer reactions taking place on timescales faster than 10ps, however, the dielectric relaxation of the surrounding protein environment has not yet taken place as the protein dynamics has an upper bound on its timescale as shown by Woodbury(*29, 30*). Consequently, an electron transfer reaction between the same pair of acceptor and donors, when allowed to proceed at timescales faster than 10ps, will experience an effective reorganization energy, on the order of 0.3eV, that is significantly lower than the uniform reorganization energy that represents the overall protein environment at post-relaxation, equilibrium timescales that is appropriate for the type of protein under consideration. This transient small effective reorganization has been recognized by the field of photosynthetic research since the 1980s but has not received its deserved attention as one of the central engineering devices adopted by natural photosystems.(*31-34*) This unrelaxed low reorganization at <10ps timescales, summarized in Figure 3.7, can now fully account for the apparent paradox summarized in Figure 3.5 and the left 3 panels of 3.6. Indeed, using this model and the native distances and driving force parameters, we were able to accurately predict the kinetics of the Rb.Sphaeroides RC, as shown in the far right panel of Figure 3.6, where all rates agree to very small margin of errors with the experimentally determined rates.

Unsurprisingly, the reaction center of Sphaeroides is not the only natural photosystems that demonstrate the paradoxical initial electron transfer rates that require the unlreaxed protein dielectrics for a satisfying model. The same combination of fast forward ET and fast direct recombination is also observed in photosystem II and I. Figure 3.8 demonstrates the same comparison between recently reported rates of early electron transfer events and the effectiveness of the mixture model in accounting for such rates. However, it is important to note that for PS II and PS I, the lack of a strong consensus in the rates of electron transfers and even the actual mechanism of electron transfer, make these two photosystems less than ideal candidates for modeling and also removes them from the focus of our study in this chapter. Nevertheless, because of the accuracy with

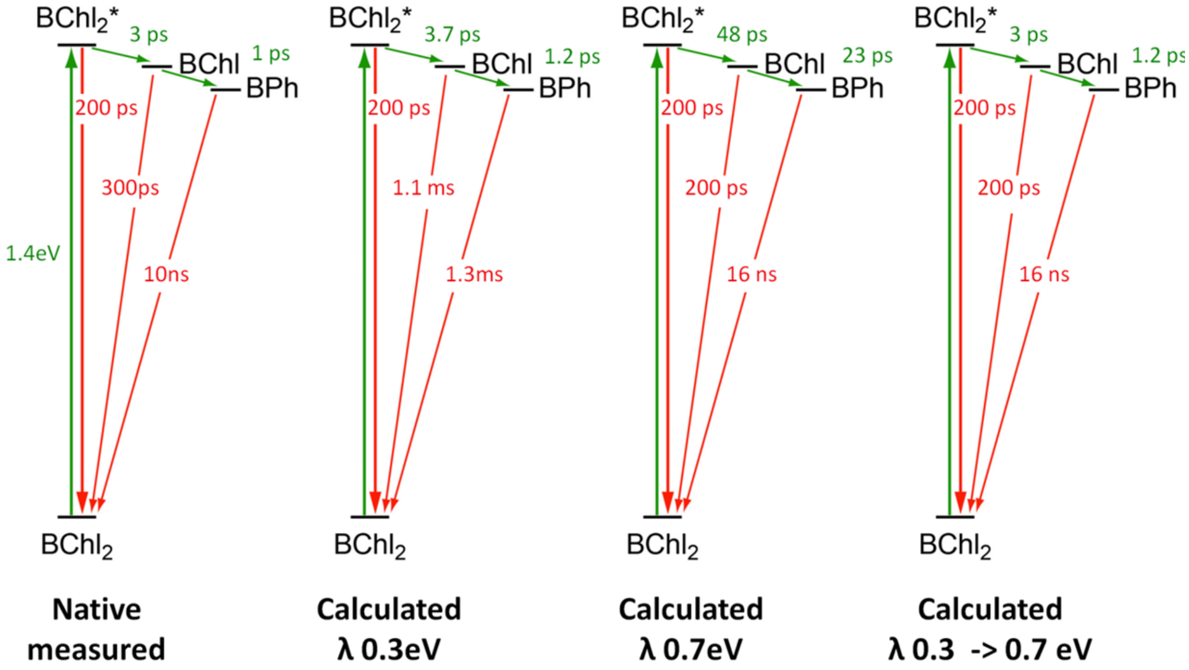

BChl2*
3 ps
BChl
1 ps
BPh
200 ps
300ps
1.4eV
10ns
BChl2
Native measured
BChl2*
3.7 ps
BChl
1.2 ps
BPh
200 ps
1.1 ms
1.3ms
BChl2
Calculated λ 0.3eV
BChl2*
48 ps
BChl
23 ps
BPh
200 ps
200 ps
16 ns
BChl2
Calculated λ 0.7eV
BChl2*
3 ps
BChl
1.2 ps
BPh
200 ps
200 ps
16 ns
BChl2
Calculated λ 0.3 -> 0.7 eV

Figure 3.6: Rates of initial electron transfers in R.Sphaeroides at various uniform and mixed reorganization energies compared to experimentally measured rates. Left most: experimentally determined ET rates of R.Sphaeroides. **Mid left:** model-determined ET rates using an uniform reorganization energy of 0.3eV. **Mid Right:** model-determined ET rates using a uniform reorganization energy of 0.7eV. **Far right:** model-determined ET rates using the mixed 0.3eV/0.7eV unrelaxed/relaxed dielectric environment model.

which our model has predicted the few rates that were successfully measured by experiments, I believe that our model-predicted ET rates for the remaining steps of PS I and PS II can serve as references and benchmarks for future experiments aimed to determine these yet unknown rates. Our discussion of them within this thesis will only go as far as predicting the rates of electron transfers using the unrelaxed/relaxed model.

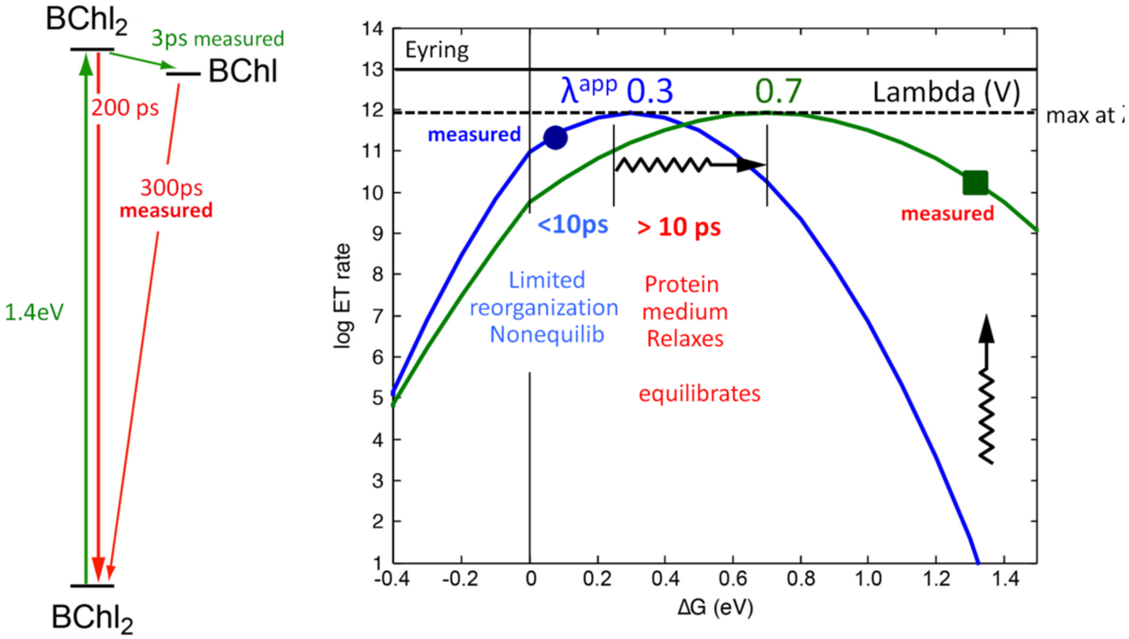


Figure 3.7: Marcus ET kinetics of initial electron transfer in natural photosystems described by both unrelaxed and relaxed dielectric protein environments at different timescales. The blue trace describes the log ET of the forward and reverse electron transfer between the P870 and the initial acceptor Ba in R.Sphaeroides if the reorganization is as low as 0.3eV. The green trace represents the log ET rate of the same electron transfers if the reorganization energy is 0.7eV. The filled circle and square on the two curves represent the experimentally observed rates.

The simultaneous occurrence of ultra-fast(<5ps) forward ET and ultra-fast(<1ns) initial direct recombination rates can only result from an electron transfer environment that is characterized by a uniformly high reorganization energy at equilibrium(timescales >10ps) and a low reorganization energy at ultra-fast timescales. Moreover, the presence of this combination of unrelaxed low dielectrics and high equilibrium reorganization energy in all of the major natural photosystems suggest that rather than an evolutionary exception and coincidence, such an environment of overall high dielectrics typical of interior or polar protein, is in fact a widely adopted, universal novel engineering device not predicted by the engineering principles we have formulated in Chapter II. In the next section, I will explore and understand the effect of this special engineering device on the performance of the natural photosystems and discover the biological design constraints natural engineering overcame with this device.

3.2.4 Uniformly high reorganization energy is a central engineering feature universal to natural photosystems.

The results of the previous section led us to conclude that the conserved catalytic quartet of natural photosystems operate in an environment that is characterized by uniformly high reorganization energy of at least 0.7eV, contrary to both long-held view of the field and the engineering principles we developed in Chapter II. In the traditionally held model where an isolated and special low reorganization energy is assigned to the initial step of electron transfer in natural photosystems,(*26*, *27*)it is believed that such low reorganization is necessary to slow down the recombination from the various acceptors to

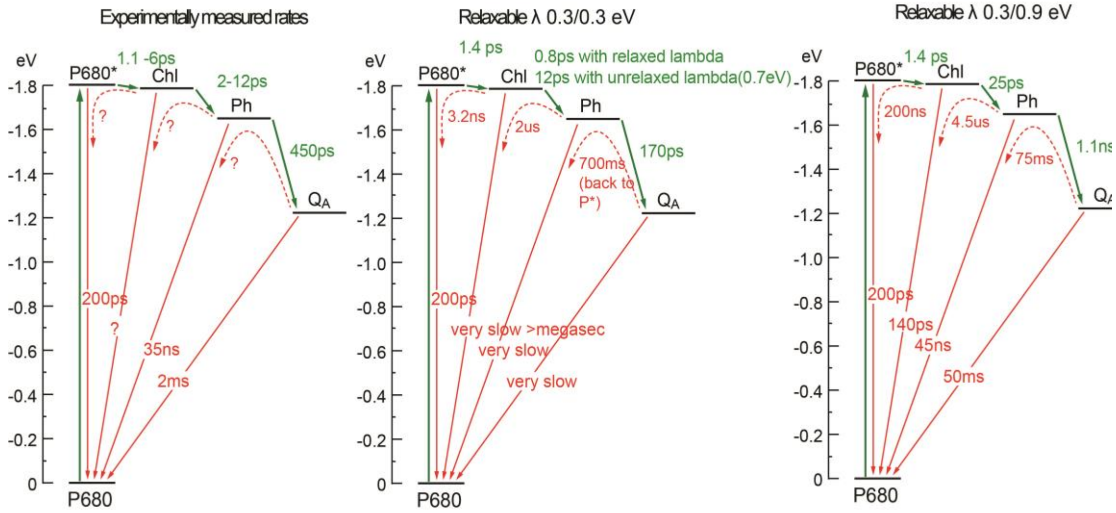

Experimentally measured rates
Relaxable λ 0.3/0.3 eV
Relaxable λ 0.3/0.9 eV
eV
-1.8
-1.6
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
P680*
Chl
Ph
QA
P680
1.1 -6ps
2-12ps
450ps
?
200ps
35ns
2ms
1.4 ps
0.8ps with relaxed lambda
12ps with unrelaxed lambda(0.7eV)
3.2ns
2us
700ms (back to P*)
170ps
very slow >megasec
very slow
25ps
200ns
4.5us
75ms
1.1ns
140ps
45ns
50ms

Figure 3.8: PSII electron transfer kinetics demonstrate unrelaxed low reorganization energy behavior. Left: experimentally determined rates of electron transfers in PS II. Question marks indicate rates that have not been determined or reported. The ranges of the known rates reflect the differing rates reported by independent studies(citations). Mid: ET kinetics predicted by the model specified in the text with a uniform reorganization of 0.3eV without the unrelaxed low reorganization at timescale <10ps. Rates that are slower than $10^6 s^{-1}$ are considered very slow and not specifically written out. Right: ET kinetics predicted by the model using a 0.3eV/0.9eV unrelaxed/relaxed mixture constraint added to the kinetics algorithm as specified in the text.

the ground state of the pigment and thereby stabilize all of the charge separated states, including both the intermediate and the final CS states. Similarly, the engineering blueprints we formulated in Chapter II indicates that the optimal method to stabilize the charge-separated state for n-ads is the Boltzmann stabilization, namely slowing down the direct recombination via a combination of inverted-region effect and increasing the distance of the recombination reaction, forcing the CS state to recombine via the uphill-repopulation of either the intermediate CS state or the excited state. Consequently, when only the physics of electron transfer are concerned, reorganization energy as low as possible would represent the optimal engineering option, as lower reorganization energies tend to significantly decrease the rate of the direct charge-recombination and force the photosystem to choose the decay route that incurs the least energetic cost. On the contrary, the conclusion we arrived at by the end of the previous section suggests that natural engineering of the biological photosystems did not favor or specifically engineer towards low reorganization energy in the form of low protein dielectrics, and the presence of transient low reorganization energy of 0.3eV provided by the unrelaxed protein dielectrics during the ultrafast forward electron transfer reactions serves a purpose that’s completely opposite of the said expectation, whose effect we will leave until section 3.2.6 to further elaborate upon. Instead of exploiting the Marcus inverted region to slow down the direct charge recombinations and stabilizing the intermediate CS states prior to the formation of the final charge-separated state, natural engineering of the photosystems employed novel engineering devices not prescribed by our engineering principles from Chapter II that strongly favor and promote fast but non-premature direct recombination, competing favorably against uphill-thermal repopulation of the preceding CS or the excited states followed up recombination. As seen in Figure 3.5 and 3.6, reorganization energy as high as 0.7eV uniform across the entire protein accelerates the charge-recombination between the initial and secondary charge-separated state to the ground state of the pigment to be on the order of 100ps and 10ns respectively, rather than the millisecond lifetime that would have been with a special low initial reorganization energy. This result indicates that the engineering of the natural photosystems did not follow the optimal engineering principles from Chapter II but rather took the completely opposite direction. It sought to destabilize rather than stabilize the intermediate charge-separated state, minimizing their lifetimes rather than maximizing them as normally expected.

In fact, here we show that the natural photosystems have engineered their placements and choices of cofactors so that they can deliver optimal charge-separation under uniformly high reorganization energy. Rather than modifying thousands of residues to obtain a more favorable dielectric environment, natural reaction center instead chose an optimal configuration of four cofactors in terms of their distances and redox potentials that could separate charge with near uniform quantum yield at high reorganization energy. It is important to note that here rather than using engineering efficiency as the metric we choose to use the quantum yield of the photosystem, since we are examining the performance of a single instance of a tetrad with fixed distances and energies under various conditions. The blue traces of the right panel in Figure 3.9 illustrates the quantum yield and engineering efficiency of Sphaeroides RC as a function of increasing equilibrium lambda with an unrelaxed lambda staying at 0.3eV. In addition to the native Sphaeroides reaction center shown in blue, we also show the semi-optimized case where the Ems of the Sphaeroides were allowed to vary with the distances fixed, or where both the Ems and distances are allowed to vary. These two traces will be discussed in greater detail later in this chapter. Here we observe that all three Sphaeroides RCs demonstrate different extents of sensitivity towards increasing reorganization energy, relative to the degree of freedoms they were allowed to have in the optimization processes. In spite of the differences in extents of sensitivity, we notice that all three cases show a drastic decrease in efficiency when the overall reorganization energy moves beyond 0.7eV, the value associated with average interior of polar proteins. Similarly, the performance, measured in quantum yield, of PS II is also examined across increasing reorganization

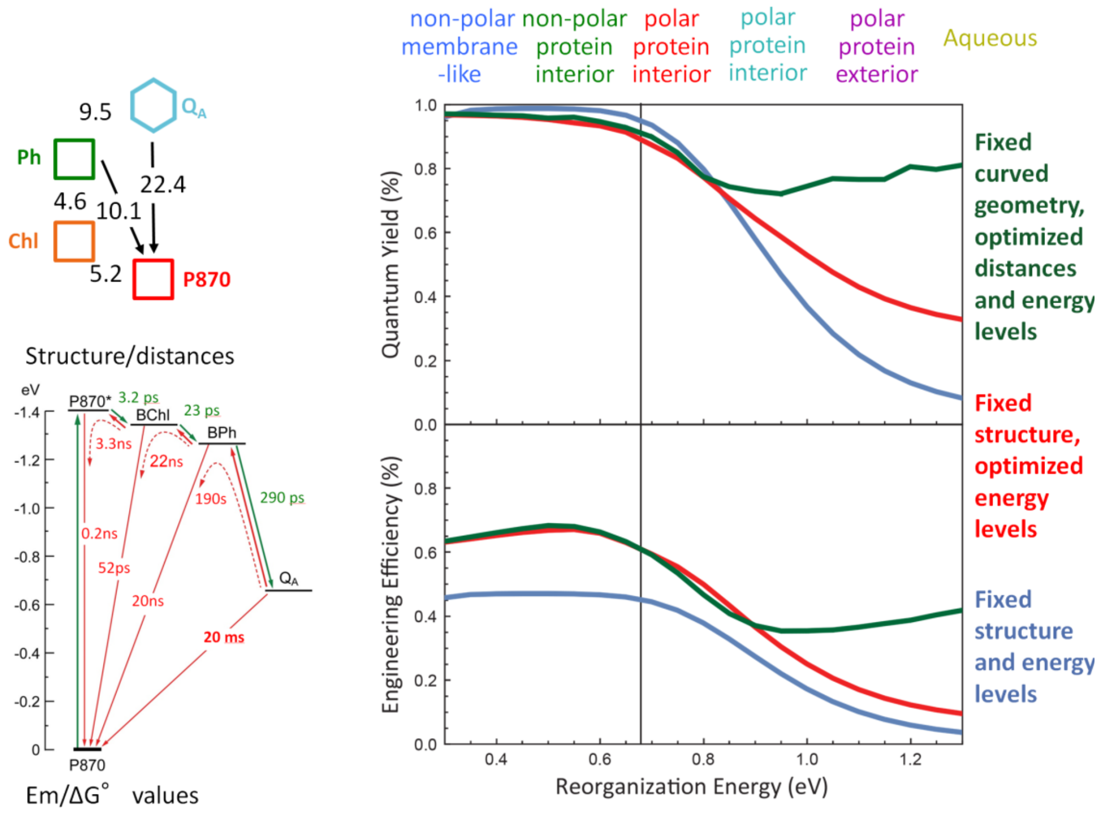


Figure 3.9: Sphaeroides reaction center specifically engineered for high performance at uniformly high reorganization energies. Left: Structural and energetic details of Sphaeroides reaction center according to the quartet model. Right: Quantum yield and engineering efficiencies at increasing reorganization energies of three types of tetrads. Green: fully optimal tetrads under the constraint of curved geometry and the unrelaxed lambda of 0.3eV. Red: semi-optimized tetrad where the distances are held constant but driving forces are allowed to vary.

energy and shown in Figure 3.10. It demonstrates a trend that's highly similar to that seen in Figure 3.9, with a somewhat higher critical point at which the efficiency experiences a sudden drop. Together these two results strongly reinforce our claim that the cofactors of the functional core of natural photosystems are engineered to function effectively in an environment of high reorganization energy, rather than the opposite, which has been a long held belief in the field of photosynthetic research.

3.2.5 Natural photochemical quartets rely on incomplete dielectric relaxation, ultra-fast excited state lifetime, and closely placed initial cofactors to minimize the lifetimes of the intermediate CS states.

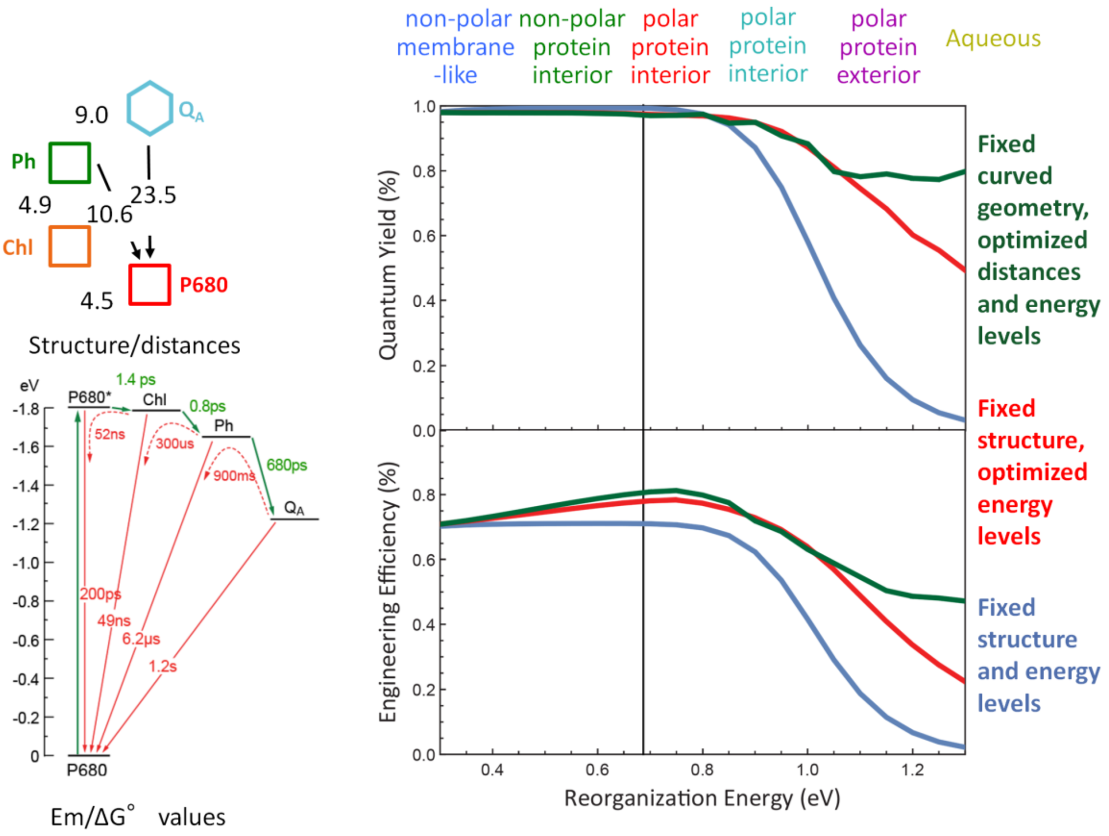


Figure 3.10: Photosystem I/II adopted generic water-soluble protein environment characterized by high uniform reorganization energy. Left: Structural and energetic details of Sphaeroides reaction center according to the quartet model. Right: Quantum yield and engineering efficiencies at increasing reorganization energies of three types of tetrads. Green: fully optimal tetrads under the constraint of curved geometry and the unrelaxed lambda of 0.3eV. Red: semi-optimized tetrad where the distances are held constant but driving forces are allowed to vary

As we have seen from section 3.2.2, although natural photosystems are engineered in an environment whose reorganization energy at equilibrium(>10ps) is uniformly high at 0.7eV, at timescales <10ps, however, a much lower reorganization energy is available and is indeed used by the engineering of the natural photosystems to ensure ultra-fast formation of the final charge-separated state without noticeable loss in quantum yield.

According to the engineering blueprints for optimal tetrads formulated in Chapter II, even under reorganization energies as high as 0.7eV/0.9eV and above, tetrads are capable of sustaining highly efficient, long-lived charge-separation.

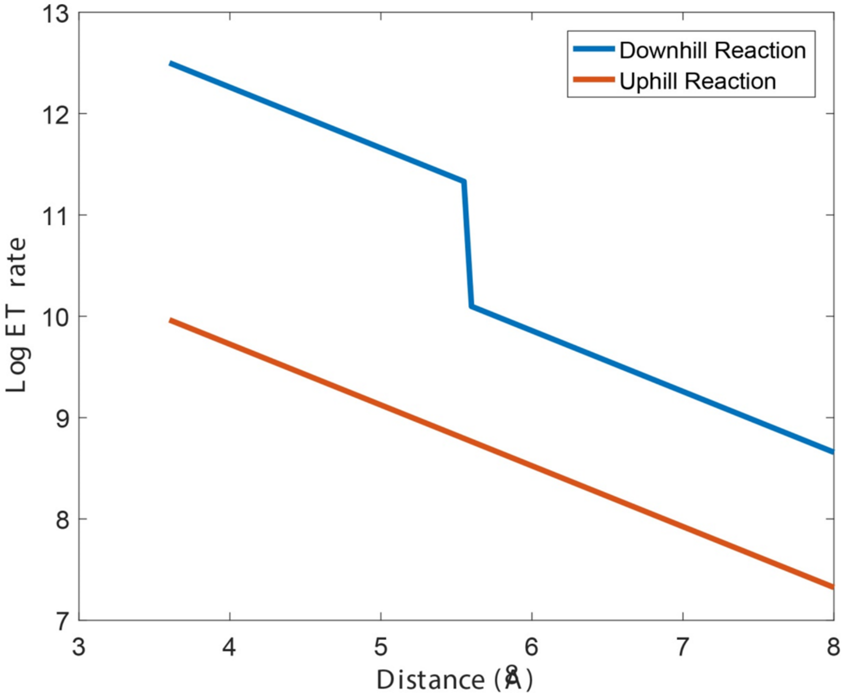


Figure 3.11: ET distance-rate model with unrelaxed low reorganization energy considered. The distance dependence of the rate of an electron transfer reaction with driving forces of 80mV(blue trace) and -80mV(red trace) at a general reorganization energy of 0.7eV at timescales >10ps and an unrelaxed reorganization energy of 0.3eV at timescales < or = 10ps.

However, in order to achieve that, the engineering guidelines specify that the initial electron transfer from the excited state of the pigment to the primary acceptor must be on the order of hundreds of picoseconds and no faster than that. However, the experimental evidence reveals that natural photosystems have adopted pigments whose singlet excited states have lifetimes on the order of hundreds of picoseconds,(*13*)rather than the 6ns singlet excited state lifetime we have used throughout our formulation of optimal engineering blueprints in the previous chapter. The ultra-fast lifetime of the excited state means that the rate of the initial forward electron transfer must be significantly faster than hundreds of picosecond to avoid premature charge-recombination, if high quantum yield of the charge-separated state is to be achieved. The significantly faster lifetime makes it necessary for the natural catalytic quartet to take advantage of the transient low reorganization energy to promote and facilitate the fast forward transfers of the excited electron to the next acceptor along the acceptor chain to efficiently and quickly generate the final charge separated state, which is then stabilized for ~100ms in order to carry out its desired function of sustaining catalysis.

To effectively take advantage of this low reorganization energy at fast timescales, natural photosystems exploited yet another engineering device that is not prescribed by the blueprints from Chapter II: minimizing the distances between the cofactors involved in the initial steps of the charge-separation to near Van der Waal contact to maximize the effective ET rate. Figure 3.10 demonstrates that a bi-phasic relationship between inter-cofactor distances and the resulting ET rates exists when the unrelaxed low reorganization is included in the model and when the driving force is favorable. A critical distance separates the fast from the slow phase, at which a sharp decrease in electron transfer rate occurs, representing the shift from the faster rate given by the unrelaxed protein dynamic to the slower rate under fully relaxed, equilibrium dynamic. It is important to note that the critical distance at which the two phases separate, shown in Figure 3.10 to be around 5.5A, will depend on the driving force of the electron transfer reaction. For the natural photosystems, whose driving forces for the initial steps are, on average, about 80mV downhill, the critical distance of 5.5A is extremely relevant and crucial. In fact, all the distances observed up-to-date in the first two steps of natural photosystems are strictly bounded by 5.5A. Consequently, it becomes apparent that the engineering of natural reaction centers preferred smaller inter-cofactor distances to encourage ultra-fast forward electron transfer, so that the intermediate charge-separated state, specifically in the case of the Sphaeroides RC the B_a^- and the H^-, are as short lived as possible.

More importantly, also shown in Figure 3.11 is the behavior of the rate of the reverse(uphill) electron transfer between the same pair of cofactors when the unrelaxed low reorganization energy is considered. Unlike the downhill reaction rate curve, the one for the reverse reaction does not involve a phase transition and therefore does not have a fast phase corresponding to that of the downhill reaction. Moreover, we notice that for the distances chosen by natural engineering, namely <5.5A, the difference between the downhill and uphill rates are significantly larger than that when the distances of the electron transfer is greater than 5.5A. In fact, the critical distance mentioned above not only creates a fast and a slow phase for the forward electron transfers, it also produces two phases for the uphill thermal repopulation based recombination, recalling Figure 2.4 from the previous chapter detailing how the recombination via uphill-repopulation is computed. Indeed, the fast phase not only increases the forward electron transfer rate, but also practically eliminates any potential charge-recombination that goes through the uphill-thermal-repopulation route.

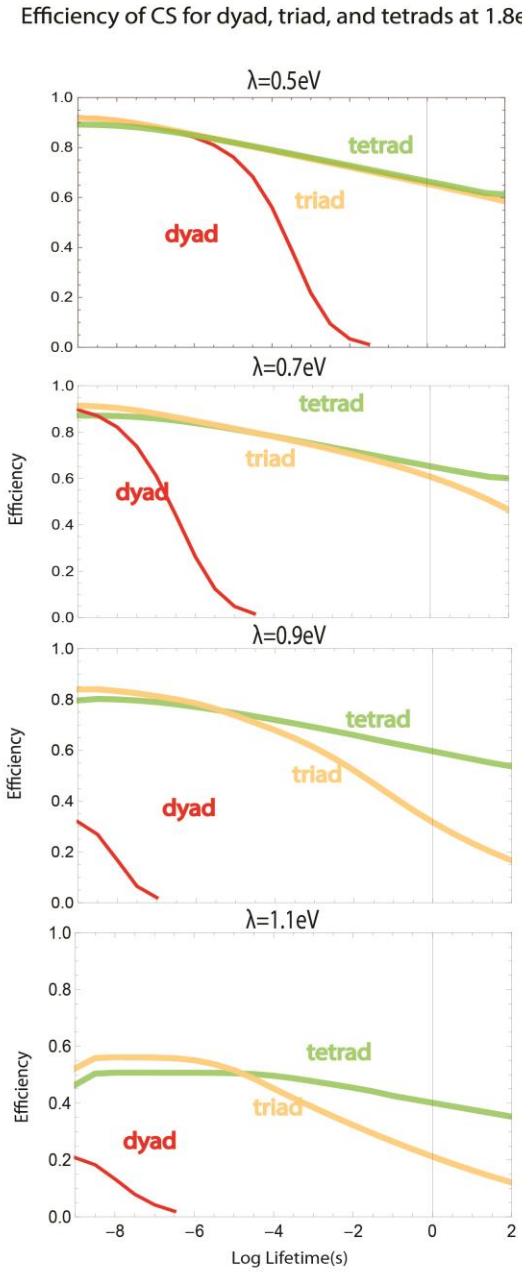


Figure 3.12: Effects of reorganization energy on the engineering of tetrads from the perspective of varying lifetimes of charge-separation. From top to bottom: each panel displays the optimal engineering efficiency of dyads, triads, and tetrads at CS lifetimes ranging from ns to 100s, with an uniform reorganization energy at the level shown at the top of the panel.

Together we can conclude that the close spacing between the initial cofactors, as one of the central engineering devices employed by natural photosystems, serves to ensure that the single dominant event in the early steps of natural photosystems is the ultra-fast and essentially irreversible formation of the charge-separated state while minimizing the possibility of any intermediate states lingering.

The transient low reorganization energy for the forward electron transfer, coupled with the high uniform reorganization energy given by the protein environment, adds another engineering mechanism to destabilize any intermediate charge-separated state. In addition to the acceleration of the direct recombination to the ground state of the pigment, natural engineering also organized the cofactors in such configurations as to both increase the electron-transfer rates from the excited state along the entire chain of acceptors to the final charge-separated state, and to eliminate any practical possibility of repopulation of the intermediate CS states. Together these engineering devices constitute the greatest subversion of the engineering blueprints formulated in Chapter II as well as long-held understanding of engineering principles of natural photosystems. Natural photosystems are engineered to avoid any possible existence of the intermediate charge-separated states or the excited state of the pigment beyond the timescale of 200ps, at which the final charge-separated is formed in an almost irreversible manner and kept alive for 100ms.

3.2.6: tetrads have significant engineering advantage over triads for effective long-lived charge-separation at high reorganization energies.

The presence of uniform high reorganization throughout natural photosystems prompts us to examine the effect of various reorganization energies on the performance of photosystems, which, in our formulation of the fundamental design blueprints of photosystems in Chapter II, did not receive much attention, since it largely represents the effects of the environment of the platform upon which the photosystem is engineered and therefore better suited for the discussions. However, after having identified the high reorganization energy as one of the fundamental engineering devices exploited by natural(biological) photosystems to produce highly efficient charge-separated state under the constraints that were not considered in the previous chapter, this inquisition is now necessary. Figure 3.12, Figure 3.13, and Figure 3.14 together demonstrate the effect of increasing reorganization energy on the optimal engineering efficiencies of linear dyads, triads, and tetrads at the range of lifetimes of interest as well as at specifically 100ms in spectacular fashion. It is important to note that these are the optimal engineering efficiencies generated using the algorithm shown in Chapter II, where none of the engineering approaches of natural photosystems is modeled. Nevertheless, even in an engineering landscape where far less constraints are considered, the tetrads demonstrate significant advantage over both dyads and triads in maintaining high performance at reorganization energy equal to or greater than 0.7eV. We have demonstrated in section 2.8 that photosystems with higher number of cofactors confer significantly greater engineering tolerance and therefore can perform much better in more constrained engineering landscape. These results therefore suggest that the absence of triads and the universal presence of tetrads in the core functional units of natural photosystems are the manifestation of another novel engineering device employed by natural photosystems, this time as a secondary response to the adoption of the high reorganization energy environment. While it can be argued that at the experimentally determined uniform reorganization energy of 0.7eV of natural photosystems, the difference in the optimal engineering efficiency between triads and tetrads do not differ significantly, the difference becomes significantly greater when the novel engineering approaches discovered in this chapter are included in the model, as will be seen in section 3.2.10.

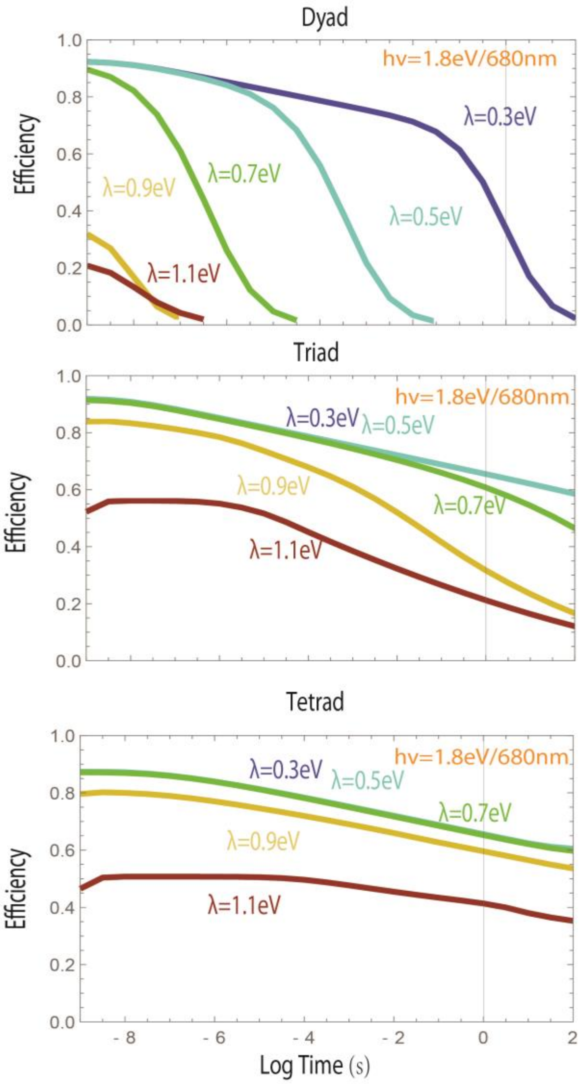


Figure 3.13: Effects of increasing reorganization energy on dyads, triads, and tetrads. From top to bottom: optimal engineering efficiency of dyads, triads, and tetrads plotted with increasing uniform reorganization energy at lifetimes from ns to 100s.

3.2.7 Natural photosystems adopt curved geometry to promote direct charge recombination at long lifetimes

While the previously characterized engineering devices all aimed to minimize the lifetimes of the intermediate charge-separated state, one of the most important and impactful engineering approaches by natural photosystems focuses mainly on the final charge-separated state, redirecting its route of its recombination from the Boltzmann-stabilized uphill-repopulation to direct electron transfer reaction to the ground state of the pigment. According to the engineering principles established in Chapter II, multi-ads allow stable charge-separations that last hundreds of milliseconds by asymmetrically slowing down direct charge recombination via a chain of acceptors placed linearly. This allows the diameter of the intermediate cofactors to maximally increase the effective distance between the final charge-separated state and the pigment, thereby maximizing the lifetime of the CS state. While the engineering of natural photosystems does utilize the chain of acceptors, surprisingly it did not arrange the cofactors linearly but rather in a curved geometry. As already discussed above, the high reorganization energy protein environment of the natural photosystems promotes fast direct charge recombinations from the intermediate charge-separated states during the electron transfer events leading up to the final charge separation. The high direct charge recombination rates ensure that these high energy intermediate CS states will not be long-lived when the path to the final charge-separated is somehow blocked. This protects the photosystem from undesirable side-reactions by these high energy intermediate states that could cause severe damage to the photosystem. However, the final charge separated state, although significantly lower in energy due to the significant drop in the energy levels described above as part of the essential engineering, must also be engineered so that direct recombination to the ground state of the pigment is preferred to the uphill repopulation of one of the intermediate CS states. Because the lifetime of the final CS state is on millisecond timescale at minimum and many times significantly longer, high reorganization energy alone is inadequate at promoting direct recombination, if the photosystem is engineered in a linear geometry as we have discovered in Chapter II. Linear arrangement of the cofactors in a photochemical triad/tetrad inevitably leads to uphill thermal repopulation of the proceeding charge separated state. In order to ensure the final CS state recombines directly to the ground state, natural photosystems adopted the fourth major

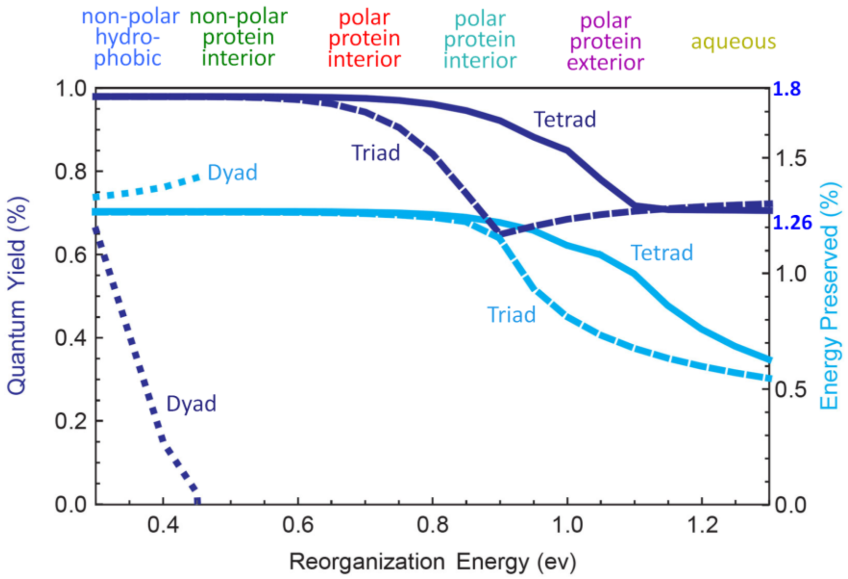


Figure 3.14: Effects on the engineering efficiencies of abstract multi-ads viewed from the reorganization energy perspective, with fixed selected lifetime of 100ms. Dark blue traces represent the quantum yield of the respective dyad(dotted), triad(dashed), and tetrad(solid) photosystems. Light blue represent the energy preserved in the final charge separated state.

engineering device, bending the arrangement of the cofactors into a curved geometry in order to decrease the distance between the final acceptor and the pigment.

This represents another major deviation from our existing engineering blueprints that are based solely in the physics of electron transfer theories: maximizing the distance between the pigment and the final acceptor to allow for maximal distance-based CS stabilization and take full advantage of the Boltzmann phase of the engineering efficiency decay. Instead, the charge-separated state is now engineered to have only adequate distance away from the pigment to sustain charge-separation long enough before it directly recombines with the ground state of the pigment. Figure 3.15 showcases the dramatic effect of promoting direct recombination via the curving effect. The curved engineering of the cofactors by an effective 50 degree brought the final acceptor from 35Å away in the linear arrangement to 22.4 Å away in the curved arrangement, resulting in 8 orders of magnitude of acceleration of ET rates and allowing the final charge-separated state to recombine directly rather than through uphill repopulation routes. This result convincingly determines the main effect of adopting a curved configuration of cofactors.

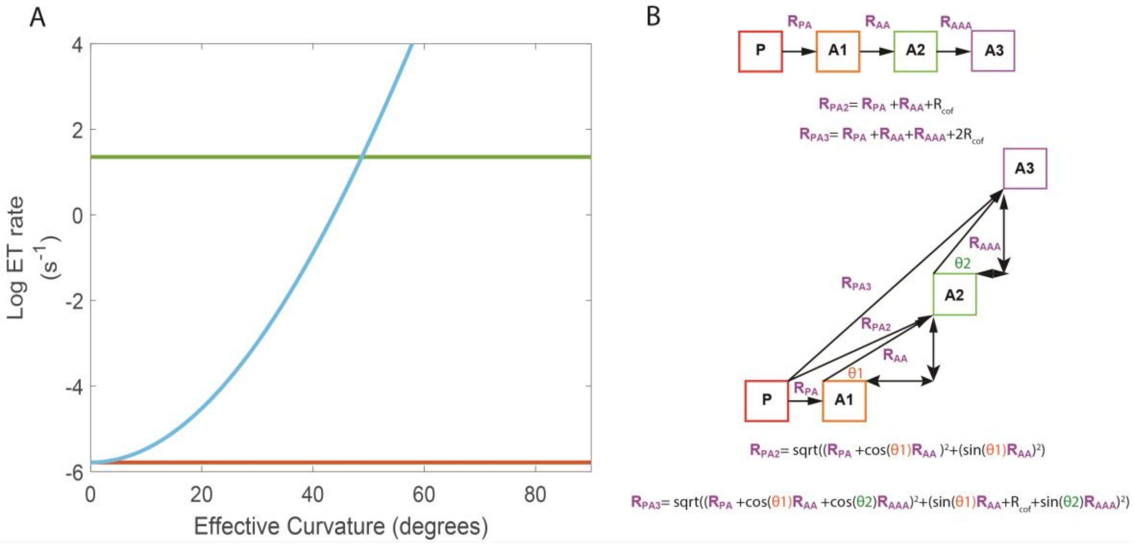


Figure 3.15: The curving of the cofactors of natural photosystem to favor direct recombination from the final charge-separated state to the pigment. A: the rates of direct RC from QA to P at various effective curvature of the reaction center, representing a simple one-angle bending of the P-QA distance. Red line represents the rate of the QA to P if the RC was arranged linearly like in B. Green line represents the experimentally observed rate. B: linear and curved models of the catalytic quartet.

Moreover, the curved design of the natural photosystems not only increases the rate of direct recombination from the final acceptor to the pigment, it also exerts the same effect on the direct recombination from the secondary acceptor, Pheophytin, to the ground state. This acceleration effect is shown in Figure 3.16. Instead of the near millisecond direct recombination rate when the cofactors are arranged linearly even with a uniform high reorganization energy of 0.7eV, the curved arrangement accelerates the direct recombination rate by 5 orders of magnitude to a few nanoseconds, ensuring that if somehow Pheophytin becomes reduced and cannot transfer its electron to the final acceptor Q_A it will not repopulate either the excited state of the pigment or the primary acceptor B_A, or have the time to undergo intersystem conversion to become the triplet state.

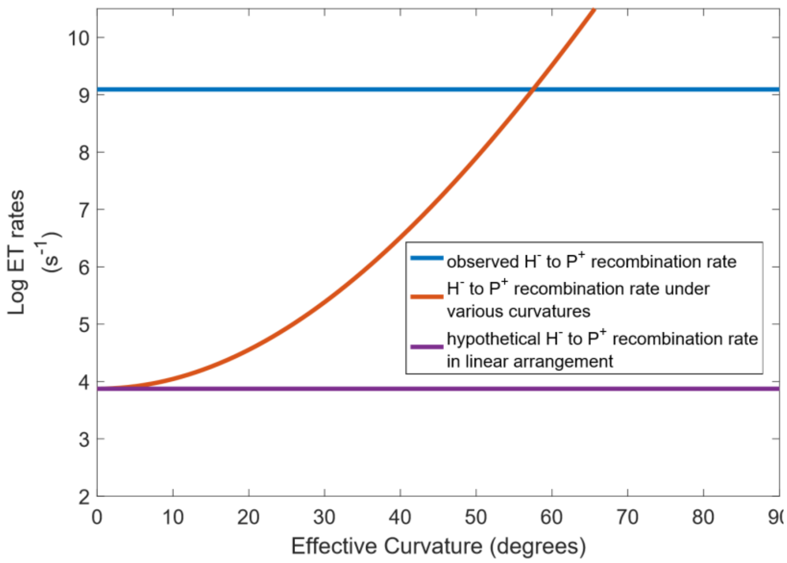


Figure 3.16: The curving of natural photosystems to accelerate direct recombination from Pheophytin to the ground state. The rates of direct RC from H_B to P at various effective curvature of the reaction center, representing a simple one-angle bending of the P-H distance. Red line represents the rate of the H_B to P if the RC was arranged linearly.

In addition to recognizing the adoption of curved cofactor arrangement as an important engineering device of natural engineering, it is of greater significance to be able to incorporate this into our n-ad computational model of photosystems in order obtain a updated version of optimal engineering guidelines that reflect the previously unconsidered constraint of protection against damage. While the three-dimensional arrangements of the cofactors in natural photosystems cannot be exactly described by a two-dimensional representation, previous sections in this chapter have successfully shown that the using the individual distances for each pair of cofactors from the crystal structures we have generated highly accurate kinetics model. This suggests that if we can discover a 2-dimensional projection of the 3d cofactor arrangements that allow us to obtain pair-wise distances that are comparable to those from the crystal structure, then this 2-dimensional project is a model that adequately describes the real arrangement of cofactors. After carefully examining various possible options, the 2D representation of the curving of the natural catalytic quartet, shown in the bottom half of Figure 3.13B, stood out as the model that gave the best distances. This model describes the curved geometry by placing the acceptors A2 and A3 at two angles, θ1 and θ2. Moreover, the diameters of the cofactors are included in the horizontal projection of the A1-A2 and A2-A3 distances. That is, the curving of the cofactors in this model allows the cofactors to overlap to certain extent from the horizontal perspective. This is necessary to model the amount of effective distance contraction described above in a 2-dimensional representation of the curving. The effective distances between the non-adjacent cofactors are computed according to the equations shown at the bottom of Figure 3.15B.

Up to this point, we have discovered and described a series of major engineering devices used by natural photosystems, all of which appear to violate and contradict the prescriptions of the engineering blueprints established in the previous chapter. All of these engineering devices contribute to the common goal of minimizing the lifetime of the intermediate CS states as well as prevention of uphill repopulation of these intermediate states when direct recombination routes are too slow. This suggests that the presence of these intermediate states at timescales longer than hundreds of picoseconds represents highly unfavorable engineering conditions for the natural photosystems. Next we proceed to address this question central to this chapter: what is nature of this novel constraint based in biological context of the natural photosystems that led to the adaptation and adoption of these engineering approaches.

3.2.8: Protection against the potential damage by long-lived intermediate charge-separated states is a central engineering constraint in optimal designs of photosystems.

The simultaneous adoption and adaptation of a series of engineering devices by the natural photosystems unanimously serve to minimize both the lifetime and the probability of re-formation of any of the states prior to the final charge-separated state in the electron transfer kinetics scheme, limiting the timescale of the presence of these intermediate states to a few to <200ps in the best case, and a few ns in the worst case. Here we demonstrate that the presence of the intermediate states at timescales longer than nanoseconds becomes sources of significant damage to both the functionality and integrity of the photosystem, thereby representing a central and essential engineering constraint based in the biological context of natural photosystem design. This constraint is the main underlying basis for natural photosystems' deviation from and contradiction against the blueprints formulated from electron transfer theory. Consequently, the incorporation of this constraint into the algorithm shown in Chapter II would allow us to formulate a new set of engineering blueprints for photosystems that confer damage-free optimal performance in protein environments.

While the final acceptors of the natural photosystems tend to be moderate in its energy levels, the reduced radical states of the preceding cofactors that constitute the intermediate charge-separated states are significantly higher in energy and much less stable. With the ground state Em ranging from 470mV to >1000mV, the excited state of the pigments can be as reductive as -1000mV, and even on the higher end(PS II), the pigment's singlet excited state has an E_m of -780mV. Other than the excited states, the intermediate charge-separated states of the natural photosystems also tend to be highly reductive, since the optimal engineering principle aims to minimize energy loss prior to the final charge-separated state. Their redox midpoint potentials range from -930mV to -650mV. Consequently, in addition to the ability to reduce their intended electron transfer partners, specifically the next acceptor in the cofactor chain, these intermediate cofactors can also reduce any other redox active entities in the vicinity when the proper conditions are met, thanks to the same fundamental theories of electron transfer that allow the reduction of their intended electron acceptors. Because of their extremely low reduction potential, the intermediate states can easily reduce redox-active side-chains and even parts of the main chain of the protein surrounding protein framework. Moreover, because the cofactors are all n=1 redox agents, their reduction would produce highly reactive and dangerous radicals that can be destructive to both the functional and structural integrity of the photosystem. Unlike their intended ET partners, however, the possible unintended targets are usually significantly farther away and the ET rates for these highly dangerous and undesirable reactions would be much slower than those with their intended partners. Therefore, the unintended ET reactions would only take place when the intermediate charge-separated states somehow remain active or reappear at much longer timescales. If the design blueprints formulated in Chapter II are strictly followed, where the charge-separated states, both the intermediate and the final, are Boltzmann stabilized via a combination of long distance and low reorganization energy, the existence of the intermediate states at timescales as long as tens or hundreds of milliseconds is inevitable, due to the preference for charge-recombination via thermal uphill-repopulation of the preceding states and the slow direct recombination. In fact, the optimal engineering blueprints established in Chapter II lead to designs of photosystems that are optimal at self-destruction. It is therefore no surprise that these design blueprints are not followed by the engineering of photosystems in nature.

In addition to direct damages to the protein environment, 3 out of the 4 cofactors of the core quartet are tetrapyroles and therefore can easily form triplet states when they remain in the high energy radical states.(*35*, *36*) Triplet excited states of tetrapyroles can easily react with oxygen molecules, which are present in the environment of functional natural photosystems. The reaction of triplet excited states of tetrapyroles with oxygen generates reactive singlet oxygen(ROS) species, exposing the photosystem to structural and function damage. (*37*, *38*).The rate of intersystem conversion, the process of forming the triplet excited state from the singlet excited state, is usually on the order of 10-100ns. Consequently, the engineering of natural photosystems must reduce the lifetimes of any intermediate CS states to faster than 10ns, in case the intended electron transfer to the next acceptor in the cofactor chain fails to take place.

3.2.9 Revised engineering blueprints for optimal photosystems prescribe designs highly similar to known natural photosystems

With the understanding of the fundamental basis of natural photosystems' novel engineering approaches we have identified and described throughout the chapter, we are now in place to modify our optimization algorithm from Chapter II to incorporate all of the novel engineering approaches in order to formulate new optimal blueprints that reflect both the physics of electron transfer and the biological context of engineering photosystems in protein environments.

First of all, we have modified the method that computes the rate of electron transfer between any two cofactors to include the unrelaxed low dielectrics at fast enough timescales. Specifically, if the rate of the ET reaction between any pairs of cofactors occurs faster than 10ps then we use 0.3eV as the reorganization energy between this pair, otherwise the reorganization energy for this specific electron transfer would resume to the uniform reorganization energy reflecting the equilibrium environment of the photosystem.

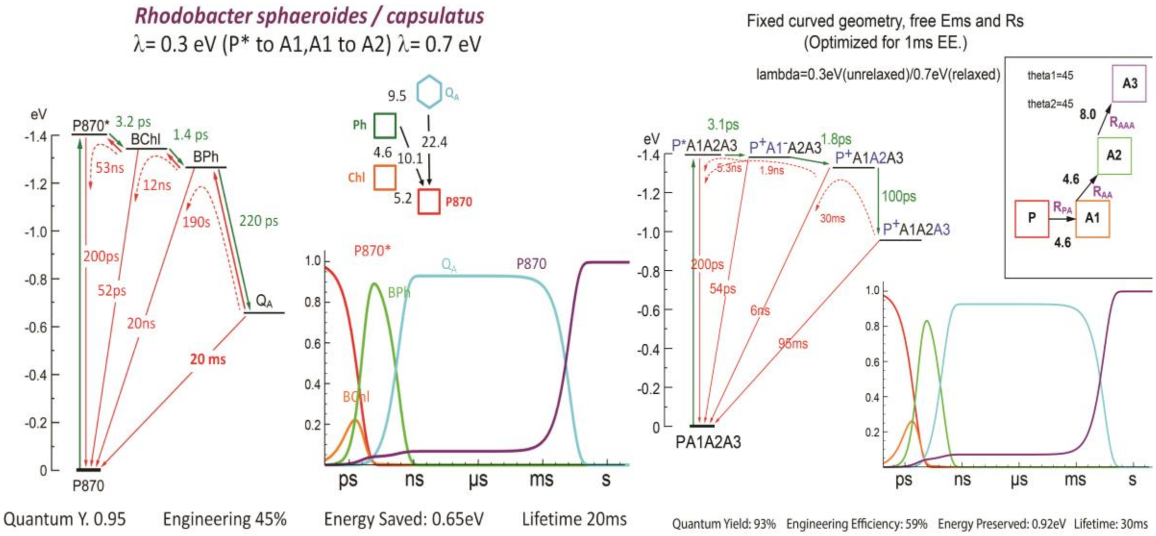


Figure 3.17: Optimal tetrads under the revised engineering guidelines display high similarity to natural photosystems. Left: electron transfer details of the reaction center of Sphaeroides, with the rates of each step of the ET kinetics shown next to the arrows as in previous trikes, while the distances between the cofactors are shown on the right top, and the populations of the various states shown at bottom right. **Right:** electron transfer details of the optimized tetrads using the revised algorithm that incorporates the engineering devices described in this chapter.

In addition, as shown in Figure 3.15B, rather than arranging the cofactors in a completely linear fashion as we have done in Chapter II, cofactors are now placed following a curved geometry, characterized by two additional engineering parameters not present in the previous versions of the algorithm: θ1 and θ2, representing the angles at which the cofactors are curving. With these two additional parameters, we can then compute the effective distances between all pairs of the cofactors in the curved arrangement using simple trigonometry.

These new optimal design blueprints also allow us to examine the extent of the optimality of the engineering of natural photosystems in the presence of the additional engineering devices we have discovered in this chapter, by comparing the performance of natural photosystems with that of the optimized tetrads from the revised algorithm. Figure 3.17 demonstrates the specific engineering details of optimal tetrads with incident photon energy of 1.4eV from the revised optimization algorithm in comparison to the natural photosystems, specifically the Sphaeroides RC. We notice that the optimized tetrads under the revised algorithm display the same engineering motifs as the natural photosystems, although minute details do differ, such as the specific redox potentials and the distances. The rates of the electron transfers, especially the first two forward steps as well as all of the direct recombination steps display practically no difference from one another. Moreover, the final forward ET step in both the optimal model and the natural photosystems, namely between A2 and A3 and H and Q_A, demonstrate a significant increase in free energy difference in comparison to the previous steps. This large drop-off, a well-known phenomenon in natural photosystems, is shown here to be the optimal engineering measure, intended to ensure that the recombination from the final acceptor state occurs via direct ET with the ground state of the pigment. The most impressive similarity between the model tetrad and the natural photosystem is in fact the kinetic profiles, also shown in Figure 3.17. The time evolutions of the populations look largely identical, with very minor differences that cannot be easily perceived. This demonstrates that the overall engineering of photosystems can be considered optimal.

However, while both the optimal model tetrads and the natural photosystems display similar engineering motif and kinetics, a revisit to Figure 3.10 and 3.11 demonstrates the performance of the fully optimized tetrad(when both the distances and midpoint potentials are allowed to vary) as well as the semi-optimized tetrad(fixing the distances to be identical to the distances of Sphaeroides reaction center) display higher engineering efficiency but very similar, practically identical quantum yield in the range of reorganization energy that’s relevant to the engineering of photosystems in protein environment(0.6eV to 0.9eV). The similarity in quantum yield between the natural and the optimal tetrads reaffirm our above-stated result, that the kinetics of natural photosystem is designed to the level of optimality. The discrepancy in engineering efficiency, however, indicates that the absolute energy preserved in the final charge-separated according to the newly revised design blueprints is higher than that of the natural photosystem. This can also be seen from the two energy diagrams in Figure 3.18, where the energy level of Q_A lies noticeably lower than that of the final acceptor in the optimal model tetrad. We will analyze and account for this discrepancy with our newly revised optimal engineering guidelines in section 3.2.11, which in fact reflects an additional but secondary engineering constraint that are faced specifically by the natural photosystems that are engineered in trans-membrane environment.

3.2.10: four cofactors are the minimal requirement for high performance charge-separating photosystems when the engineering dogma of damage protection is included

Although we have already demonstrated in 3.2.6 that tetrads hold significant engineering advantage over dyads and triads when the photosystem is engineered in environments characterized by uniform high reorganization energy, the advantage at the low end of the range of high reorganization energy tetrad holds over triads is not significant enough to fully justify the complete absence of triads in natural photosystems. However, this conclusion was reached rather prematurely in section 3.2.6 using the model established in the previous chapter, where the novel engineering devices were not considered. It is necessary to examine the dependence of multi-ads on reorganization energy under the revised optimal model, although we hypothesize that since additional and stronger constraints have been added to the model, tetrads' advantage could only become greater.

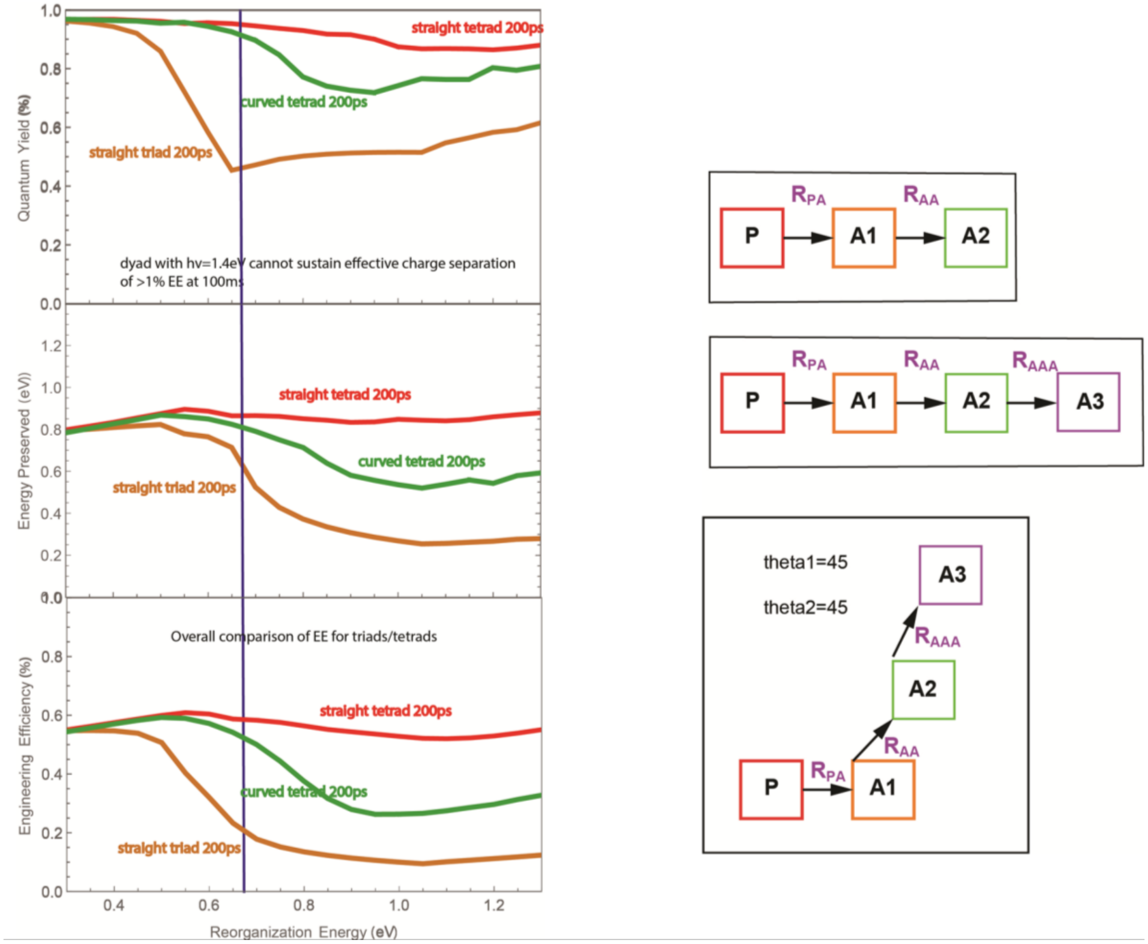


Figure 3.18: tetrads are the minimal engineering requirement for high performance when protection against damage is included as an essential engineering constraint. Left: optimal quantum yield, energy preserved, and engineering efficiency of optimized linear(red) and curved(green) tetrads and linear triad(brown) at increasing reorganization energies from 0.3eV to 1.3eV. **Right**: cartoons representing the structural organization of the triads and tetrads examined in this figure.

It is not surprising to discover that, under the revised engineering guidelines of this chapter, triads can no longer supply the same or similar level of high efficiency at lifetimes required for chemical catalysis at reorganization energies that are at least 0.7eV, and become completely unviable as an engineering option, as seen in Figure 3.18. Figure 3.18 illustrates that when the additional engineering guidelines provided by biochemical and biophysical constraints are engaged and the time-dependent, relaxation enabled dielectric environment is applied, tetrads' advantage over dyads and triads becomes significantly more pronounced than what was shown above in section 3.2.6. It is important to note that the efficiency of optimal triads shown in this figure is obtained

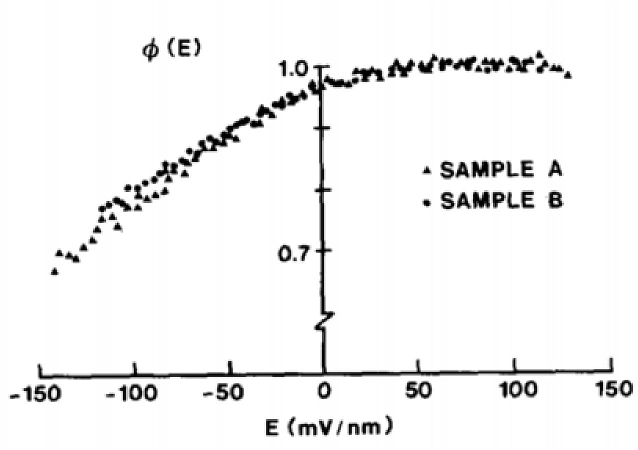


Figure 3.19: Quantum yield's dependence on the membrane potential experienced by the Sphaeroides reaction center. Reproduced from (41)

using a model does not include the curved geometry, which in fact over-estimates the optimal performance of triads, and even in this comparison, a significant advantage in is seen for tetrads at the relevant reorganization energies.

This result confirms and further solidifies a previously stated conclusion that 4 cofactors is the minimal engineering requirement for a photosystem characterized by uniform or near-uniform quantum yield and optimal engineering efficiency in an environment representative of typical interior of polar proteins.

3.2.11: Natural catalytic quartet anticipates membrane potential effect with suboptimal cofactor energetics

In the comparison of the optimal tetrads according to the revised engineering principles to existing natural photosystems, it appears that the natural catalytic quartet of a 1.4eV photon pigment is only partially optimal, as seen in both Figure 3.9 and 3.10. Only the behavior of the quantum yield of the natural photosystem appears to agree with the prescribed optimal tetrads, while the engineering efficiency remains much higher than that of the natural photosystem. This means that both the semi-optimized(using the same inter-cofactor distances as the native with Ems allowed to vary) and the fully-optimized(both distances and Ems allowed to vary, with cofactors arranged in the curved geometry described above) versions of the said quartet display higher energy preserved than the native Sphaeroides quartet at relevant reorganization energies. That is, the Em of the final acceptor in the Sphaeroides is lower in energy than what the engineering blueprints prescribes as optimal. It appears that natural engineering incurred additional energetic costs that are not demanded by the revised engineering guidelines.

Here we describe an additional engineering constraint that's faced specifically by natural photosystems such as the reaction center of Sphaeroides that span the membrane of thykloids and serve to transport charges from one side of the membrane to the other. This constraint is therefore the effect of a trans-membrane electric field generated by the intended function of the reaction center over time, resulting in significant changes in the redox midpoint potentials of the cofactors at equilibrium. Various studies have demonstrated that the cofactors of Sphaeroides reaction center experience significant modulation from a membrane

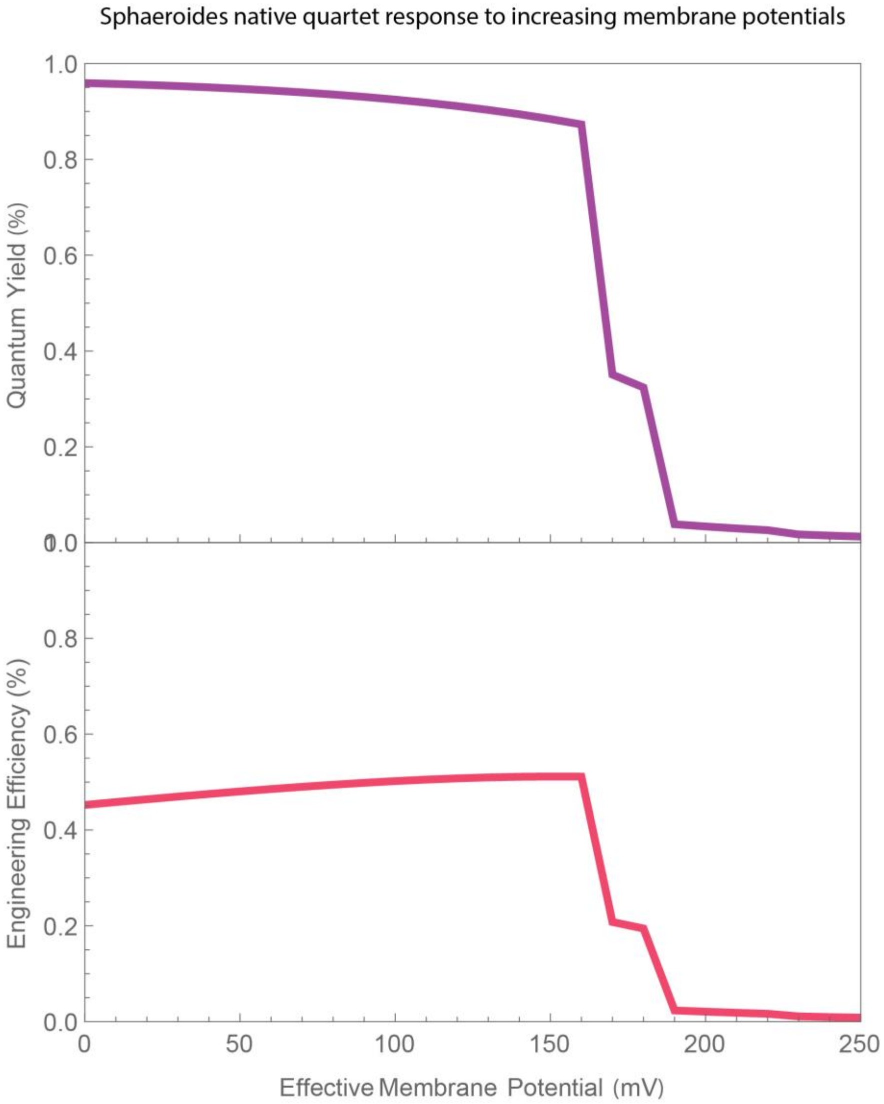


Figure 3.20 The shut-down effect imposed upon the reaction center of Sphaeroides by increasing membrane potentials. Top: effect of increasing membrane potential on the quantum yield of the Spaheroides reaction center. **Bottom:** effect of increasing membrane potential on the engineering efficiency of the Sphaeroides reaction center.

potential as strong as 200mV/nm, established at equilibrium when continual formation of the charge-separated state results in an artificial electric field across the membrane where the reaction center resides. (*39*, *40*)Moreover, the modulating effect of the membrane potential is applied unevenly to the various cofactors in the natural reaction center of Sphaeroides, since the cofactors are arranged along the direction of the field. Therefore, the more distal cofactors will experience greater changes in potentials than the more proximal ones. Specifically, the E_m of Q_A experiences the full effect of the membrane potential, while that of the chlorophyll monomer experiences the least. Figure 3.20 demonstrates the effect of increasing membrane potentials on the quantum yield of the native Sphaeroides reaction center.(*41-43*) The classic study reveals that there is a parabolic relationship between the quantum yield of the reaction center and the increasing membrane potentials, thereby specifying an optimal level of membrane potential beyond which the photosystem begins to fail. It has been known that the strength of the generated membrane potential causes the effective midpoint potentials of the cofactors in the reaction center to change, thereby impacting both the energy preserved and quantum yield of the charge separation.

The effect of quantum yield can be easily incorporated into our quartet model of the natural photosystems. Figure 3.20 demonstrates the impact of increasing effective membrane potential on both the quantum yield and the engineering efficiency of a native Sphaeroides reaction center operating at its standard ET environment represented by a reorganization energy of 0.7eV, with an unrelaxed low dielectric environment represented by a 0.3eV reorganization energy. Our predictive analysis shows that there is indeed an optimal membrane potential prior to which the engineering efficiency of the reaction center continues increasing and past which a critical failure essentially shuts down the system. Figure 3.21 demonstrates the performance of the Sphaeroides reaction center after applying the effects of the membrane potential on the potentials of the cofactors. We notice that the optimal tetrads do display improved engineering efficiency over the membrane-potential free case, as the quantum yield remains unaffected while the energy preserved in the final CS state has increased. However, it is important to note that the direct recombination rates from the intermediate acceptors to the ground state of the pigment are slowed down by the effects of the membrane potential, which increases the driving force for the direct recombination electron transfer and thus induces a greater inverted-region effect.

We claim that the shut-down effect at higher membrane potentials is another engineering device employed by the natural photosystems for the protection against damage. At higher membrane potentials, the secondary acceptor Pheophytin would no longer have a direct recombination rate that's faster than the <10ns timescale that we have identified to be essential for the protection effect. This is seen in Figure 3.22, where the direct-recombination rate from H to P^+ has been slowed to 210ns due to the membrane potential effect, making its presence highly dangerous. However, the natural catalytic quartet is engineered to anticipate this. Due to the unequal effect the membrane potential exerts on the various cofactors, at high membrane

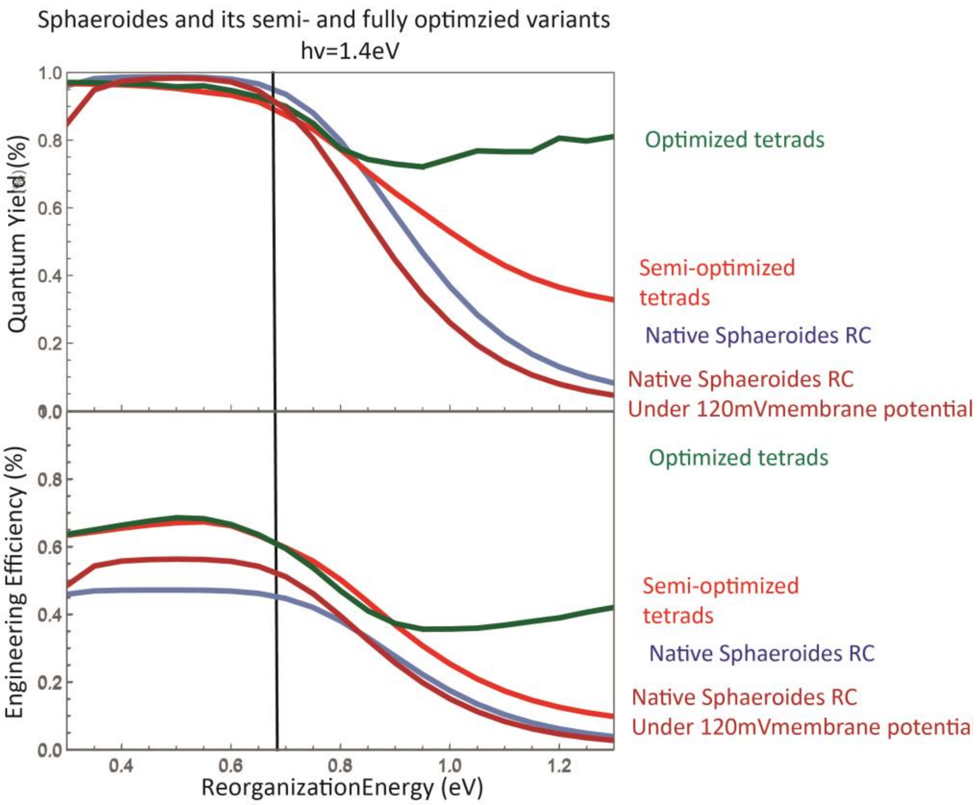
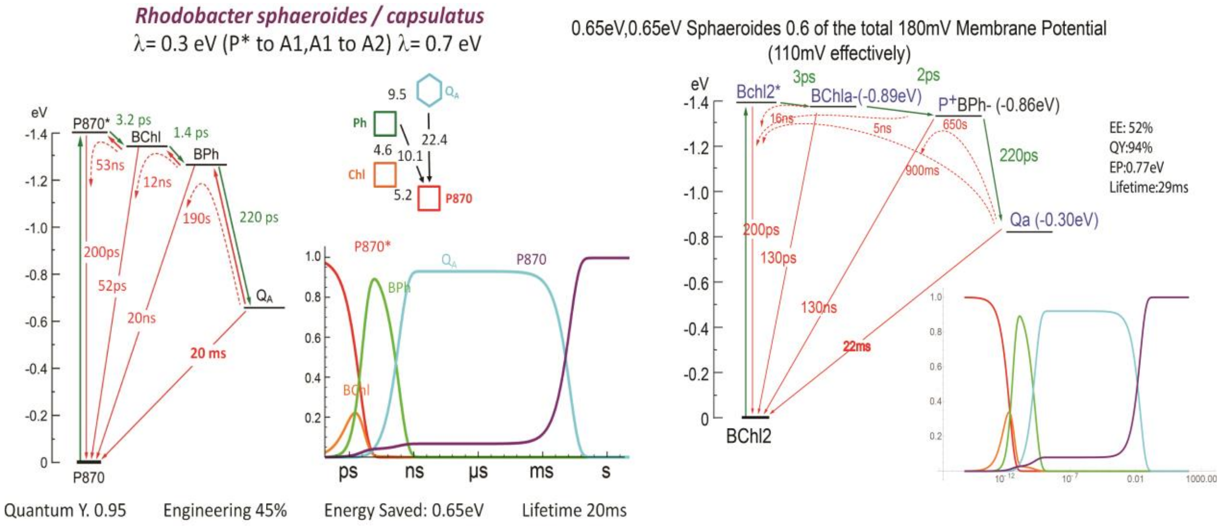


Figure 3.21 Effects of the membrane potential on the performance of Spaheroides reaction center. A modified version of figure 3.11, with the addition of the performance of Sphaeroides RC when the effects of the membrane potential is added.

potentials, the once favorable forward electron transfer from B_A to H is now uphill and becomes highly unfavorable, resulting in a greater rate from H back to B_A, thereby causing the apparent shut-down effect. At an effective membrane potential of 170mV, we see that the shut-down effect is partially engaged, with the quantum yield of the Sphaeroides reaction center reduced to 32%. With further increases in the membrane potential, the unfavorability of the B_A to H reaction will keep aggravating and eventually leads to the complete failure we observe in Figure 3.20.

However, it is important to note that we chose not to present a discussion of the effects of the equilibrium membrane potential on the effect of our optimal design blueprints, as the effects of membrane potential is not an essential constraint like the protection against damages by the intermediate charge-separated states most of this chapter has focused on. The membrane potential is a by-product of designing a trans-membrane photosystem whose function helps generate the membrane potential. The actual application of the revised engineering blueprints for an optimal charge-separating tetrad in practice, as will be shown in chapter VII, will be carried out in a small, man-made, and water-soluble protein, absolving the engineering of any necessity to compromise its compromise while



Rhodobacter sphaeroides / capsulatus
λ= 0.3 eV (P* to A1,A1 to A2) λ= 0.7 eV
eV
P870*
3.2 ps
BChl
1.4 ps
BPh
53ns
12ns
190s
220 ps
200ps
52ps
20ns
20 ms
QA
P870
9.5
22.4
10.1
4.6
5.2
Ph
Chl
P870
P870*
BPh
BChl
QA
P870
ps
ns
μs
ms
s
Quantum Y. 0.95
Engineering 45%
Energy Saved: 0.65eV
Lifetime 20ms
0.65eV,0.65eV Sphaeroides 0.6 of the total 180mV Membrane Potential
(110mV effectively)
eV
3ps
2ps
Bchl2*
BChla-(-0.89eV)
P+BPh- (-0.86eV)
16ns
5ns
650s
900ms
220ps
EE: 52%
QY:94%
EP:0.77eV
Lifetime:29ms
Qa (-0.30eV)
200ps
130ps
130ns
22ms
BChl2

Figure 3.22: Effects of membrane potentials on the details of electron transfer kinetics in the Sphaeroides reaction center. Left: details of the electron transfer kinetics of the Sphaeroides reaction center when no membrane potential is involved. **Right**: kinetics of the electron transfer of the reaction center when 110mV of effective membrane potential is applied. Rates are shown along with the arrows indicating the direction and partners of the electron transfers.

incurring extraneous structural complexity. Indeed, the results and discussion of this section aimed to further advance the importance of protection against any possible sources of damage as the central engineering constraint of ET-based photosystems in addition to the fundamental theories of electron transfers.

3.3 Conclusions:

Throughout this chapter, we began with a set of engineering principles that we have shown to be the guidelines for optimal engineering of charge-separating photosystems, formulated on a complete understanding of the interactions of the physics of electron transfer reactions between cofactors. The examination of naturally occurring photosystems revealed the surprising fact that nature did not follow the engineering blueprints we have shown to be optimal and prompted us to discover additional engineering principles that are based in the fundamental theories of electron transfer.

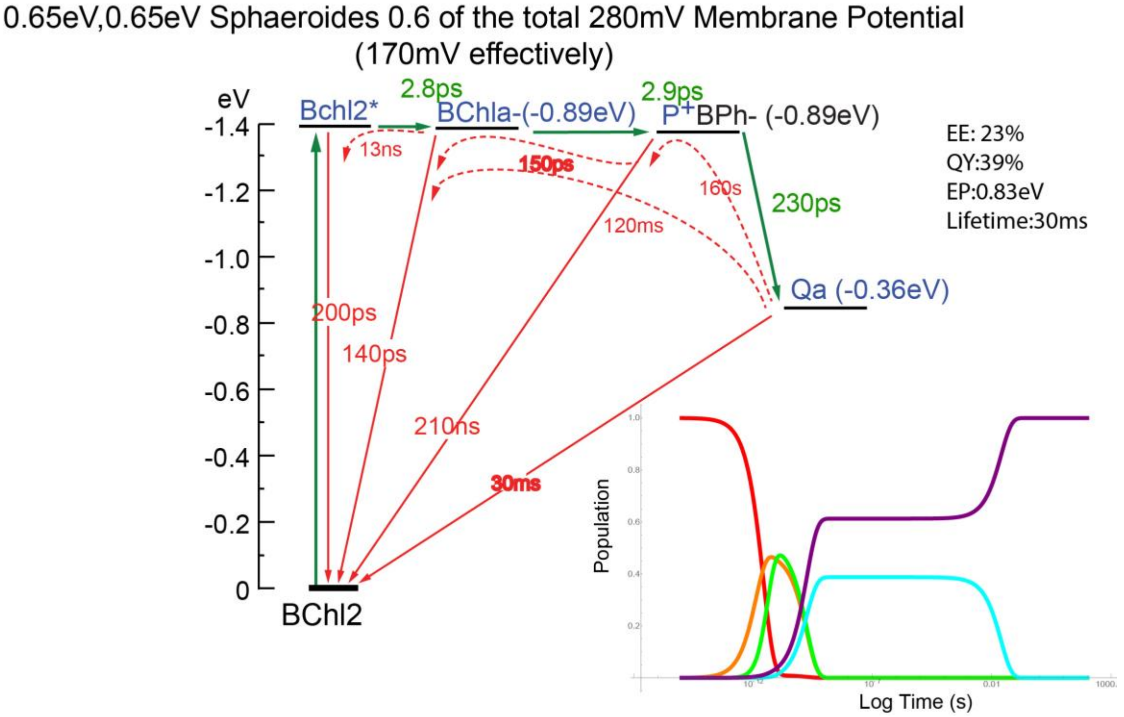

0.65eV,0.65eV Sphaeroides 0.6 of the total 280mV Membrane Potential
(170mV effectively)
eV
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
Bchl2*
2.8ps
BChla-(-0.89eV)
2.9ps
P+BPh- (-0.89eV)
13ns
150ps
160s
120ms
230ps
Qa (-0.36eV)
200ps
140ps
210ns
30ms
BChl2
EE: 23%
QY:39%
EP:0.83eV
Lifetime:30ms
Population
1.0
0.8
0.6
0.4
0.2
10^-12
10^-7
0.01
1000.
Log Time (s)

Figure 3.23: Shut-down effect of higher membrane potential on the Sphaeroides reaction center. Details of the electron transfer kinetics of the reaction center of Sphaeroides under high membrane potentials, at which partial shut-down effect is observed. Red trace represents the excited state of the pigment, while light blue represents the final excited, and purple is the ground state of the pigment.

Instead, natural engineering of photosystems adopted various devices and approaches that prioritize both the efficiency of the charge-separation and the safety and integrity of the photosystem. These novel engineering devices serve to protect the photosystem from the potential damages induced by the high-energy intermediate radical states preceding the final charge-separated state. Ironically, the engineering principles that are considered optimal when only the theories of electron transfer are involved turn out to be highly unfavorable and lead to completely unviable designs. However, upon understanding that protection against damage is an essential engineering constraint along with the limits set by the physics of electron transfer, a new set of optimal engineering blueprints were successfully formulated by modifying the theoretical model to incorporate the series of engineering approaches and devices employed by nature to address damage protection. This new set of engineering guidelines prescribe designs that are highly similar to known natural photosystems in spite of small discrepancies and variations, indicating that the revised model is adequate and sufficient for practical application.

This chapter finishes the theoretical side of my thesis, after having extensively explored the engineering landscape of photosystems from the perspectives of fundamental electron transfer theories as well as practical constraints imposed upon by the biological environment of natural photosystems. The results and conclusions of this half of my thesis becomes the guidelines that the efforts described in the second half of the thesis will follow, to the best of their ability thanks to the many additional constraints that are highly specific to particular experimental approaches we have access to in the engineering of photosystems. These practical constraints do not represent universal limitations but rather lack of resources or accesses and therefore are not considered or included in the guidelines of this chapter.

3.4 Computational methods

The analysis performed in this chapter and their corresponding results utilized computational methods that are extensions and further developments of those described in Computational Methods section of Chapter 2. We will continue using the notational system of the methods section of Chapter 2.

To model the electron transfer kinetics of the quartet of Sphaeroides photosystem, we implemented another variant of the function **P** described in the methods section of Chapter 2 and will call it **Sp.** Rather than specifying only the pair-wise distance between adjacent cofactors and use the linear geometry to infer the pairwise distances between all the cofactors, **Sp** requires the user to input pairwise distances between all pairs of cofactors, in order to reflect the non-linear arrangement of cofactors in the natural reaction centers. Additionally, **Sp** also includes a conditional statement that checks whether each electron transfer is taking place at timescales faster than 10ps when computing the rate of electron transfers between each pair of the cofactors using the Moser-Dutton ruler and uses the value of 0.3eV as the reorganization energy if the condition is evaluated to be True.

To reflect both the dynamic reorganization energy and the curved geometry in generating the optimal engineering blueprints, we implemented a variant of **E,** and will call it **Ec. Ec** includes two extra input parameters in addition to those required by **E,** $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$**,** which are used to compute the effective distances between the pigment and the secondary and tertiary acceptors as shown in Figure 3.15. The calculation of the effective distances involves calling trigonometric functions that are built-in components of Mathematica 10.0. Depending on the analysis, parameters $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_2}$ are either kept constant at 45° and 60° or allowed to vary as part of the varying design parameters that are the objects of optimization using the NMaximize function. **Ec** is optimized in a similar way as **E** as described in Chapter 2, using the “DifferentialEvolution” flag as the selected method of optimization and “1” as the selected “RandomSeed” flag.

We also implemented a version of **E** to examine the level of optimality in the energetics of the natural Sphaeroides reaction center and will denote it as **ESp. ESp** fixes all the pairwise distances between the cofactors to be the values used in **Sp. ESp** thus does not take distances as input parameters. **ESp** is optimized under a series of constraints using the “NMaximize” built-in function of Mathematica similar to **E**, with the “DifferentialEvolution” method flag and the “RandomSeed” flag of 1, but only with respect to the driving forces between the cofactors rather than both distances and driving forces.

To reflect the effect of increasing membrane potentials on the performance of Sphaeroides reaction center, we implemented a function **MemPot** to compute the E_m of all the cofactors under a user-specified level of membrane potential. The implementation of **MemPot** involves a distance-discounted modification of the redox potentials of the cofactors, returning a set of new redox potentials which are then further used in **Sp** to calculate the electron transfer kinetics in Sphaeroides RC under various membrane potentials shown in Figures 3.20-3.22.

3.5 References

1. R. E. Blankenship, *Molecular mechanisms of photosynthesis* (2013).
2. D. Zhong, Electron Transfer Mechanisms of DNA Repair by Photolyase. *http://proxy.library.upenn.edu:2146/10.1146/annurev-physchem-040513-103631*. **66**, 691–715 (2015).

3. R. E. Blankenship *et al.*, Comparing Photosynthetic and Photovoltaic Efficiencies and Recognizing the Potential for Improvement. *Science*. **332**, 805 (2011).
4. N. Kamiya, J. R. Shen, Crystal Structure of Photosystem II (2003), doi:10.2210/pdb1izl/pdb.
5. A. Ben-Shem, F. Frolow, N. Nelson, Crystal Structure of Plant Photosystem I (2004), doi:10.2210/pdb1qzv/pdb.
6. J. Deisenhofer, H. Michel, in *The Photosynthetic Bacterial Reaction Center* (Springer US, Boston, MA, 1988), pp. 1–3.
7. J. P. Allen, G. Feher, T. O. Yeates, H. Komiya, D. C. Rees, Structure of the Reaction Center From Rhodobacter Sphaeroides R-26: the Cofactors. *PNAS*. **84**, 5730–5734 (1987).
8. K. Hitomi, A. S. Arvai, J. A. Tainer, E. D. Getzoff, (6-4) Photolyase Crystal Structure (2009), doi:10.2210/pdb3fy4/pdb.
9. C. C. Moser, C. C. Page, P. L. Dutton, Tunneling in PSII. *Photoch Photobio Sci*. **4**, 933–939 (2005).
10. C. C. Moser, P. L. Dutton, in *Photosystem I* (Springer Netherlands, Dordrecht, 2006), vol. 24 of *Advances in Photosynthesis and Respiration*, pp. 583–594.
11. C. C. Moser, C. C. Page, P. L. Dutton, *Photosynthesis: Bacterial Reaction Center*.
12. K. Dubas, M. Baranowski, A. Podhorodecki, M. R. Jones, K. Gibasiewicz, Unified Model of Nanosecond Charge Recombination in Closed Reaction Centers from Rhodobacter sphaeroides: Role of Protein Polarization Dynamics. *J Phys Chem B*. **120**, 4890–4896 (2016).
13. M. Saggu *et al.*, Putative Hydrogen Bond to Tyrosine M208 in Photosynthetic Reaction Centers from Rhodobacter capsulatus Significantly Slows Primary Charge Separation. *J Phys Chem B*. **118**, 6721 (2014).
14. J. Zhu, I. van Stokkum, L. Paparelli, M. R. Jones, Early Bacteriopheophytin Reduction in Charge Separation in Reaction Centers of Rhodobacter Sphaeroides. *Biophysical Journal* (2013).
15. Y. Kakitani, A. Hou, Y. Miyasako, Y. Koyama, Rates of the Initial Two Steps of Electron Transfer in Reaction Centers From Rhodobacter Sphaeroides as Determined by Singular-Value Decomposition Followed by Global Fitting. *Chemical Physics Letters*. **492**, 142 (2010).
16. Constructing a Man-Made C-Type Cytochrome Maquette in Vivo: Electron Transfer, Oxygen Transport and Conversion to a Photoactive Light Harvesting Maquette. **5**, 507–514 (2014).
17. J. Deisenhofer, H. Michel, *Three-Dimensional Structure of the Reaction Center of Rhodopseudomonas viridis* (1993).
18. J. M. Peloquin *et al.*, Time-Dependent Thermodynamics During Early Electron Transfer in Reaction Centers From Rhodobacter sphaeroides. *Biochemistry*. **33**, 8089–8100 (2002).
19. W. Zinth, W. Kaiser, *Time-Resolved Spectroscopy of the Primary Electron Transfer in Reaction Centers of Rhodobacter sphaeroides and Rhodopseudomonas viridis* (1993).
20. H. XU, Theoretical Studies on the Influence of Molecular Interactions on the Mechanism of Electron Transfer in the Photosynthetic Reaction Center of Rps. Viridis. *Science in China Series B*. **45**, 570 (2002).
21. C. Kirmaier, L. Laporte, C. C. Schenck, The Nature and Dynamics of the Charge-Separated Intermediate in Reaction Centers in which Bacteriochlorophyll Replaces the Photoactive Bacteriopheophytin. 2. The Rates and Yields of Charge Separation and Recombination. *The Journal of Physical Chemistry*. **99**, 8910 (1995).
22. N. W. Woodbury, M. Becker, D. Middendorf, W. W. Parson, Picosecond Kinetics of the Initial Photochemical Electron-Transfer Reaction in Bacterial Photosynthetic Reaction Centers. *Biochemistry*. **24**, 7516–7521 (2002).
23. Jessica I Chuang, Steven G Boxer, A. Dewey Holten, Christine Kirmaier, High Yield of M-Side Electron Transfer in Mutants of Rhodobacter capsulatus Reaction Centers Lacking the L-Side Bacteriopheophytin. *Biochemistry*. **45**, 3845–3851 (2006).
24. C.-K. Tang, J. C. Williams, A. K. W. Taguchi, J. P. Allen, N. W. Woodbury, P + H A - Charge Recombination Reaction Rate Constant in Rhodobacter sphaeroides Reaction Centers Is Independent of the P/P + Midpoint Potential †. *Biochemistry*. **38**, 8794 (1999).
25. D. A. Moss, M. Leonhard, M. Bauscher, W. Mäntele, Electrochemical Redox Titration of Cofactors in the Reaction Center From Rhodobacter sphaeroides. *FEBS Letters*. **283**, 33–36 (1991).
26. M. Iwaki, S. Itoh, Structure of the Phylloquinone-Binding (Q.Vphi.) Site in Green Plant

Photosystem I Reaction Centers: the Affinity of Quinones and Quinonoid Compounds for the Q.Vphi. Site. *Biochemistry*. **30**, 5347–5352 (1991).

27. W. W. Parson, Z. T. Chu, A. Warshel, Reorganization Energy of the Initial Electron-Transfer Step in Photosynthetic Bacterial Reaction Centers. *Biophysical Journal* (1998).
28. I. McConnell, G. H. Li, G. W. Brudvig, Energy Conversion in Natural and Artificial Photosynthesis. *Chem Biol*. **17**, 434–447 (2010).
29. J. P. Allen, J. Williams, N. W. Woodbury, Unusual Temperature Dependence of Photosynthetic Electron Transfer Due to Protein Dynamics. *J Phys Chem B*. **113**, 818 (2008).
30. H. Wang *et al.*, Protein Dynamics Control the Kinetics of Initial Electron Transfer in Photosynthesis. *Science*. **316**, 747 (2007).
31. Z. Guo, N. W. Woodbury, J. Pan, S. Lin, Protein Dielectric Environment Modulates the Electron-Transfer Pathway in Photosynthetic Reaction Centers. *Biophysical Journal*. **103**, 1979 (2012).
32. L. I. Krishtalik, The Medium Reorganization Energy for the Charge Transfer Reactions in Proteins. *Biochimica et Biophysica Acta (BBA)*. **1807**, 1444–1456 (2011).
33. L. I. Krishtalik, Fast Electron Transfers in Photosynthetic Reaction Centre: Effect of the Time-Evolution of Dielectric Response. *Biochimica et Biophysica Acta (BBA)*. **1228**, 58–66 (1995).
34. C. C. Moser, C. C. Page, R. Farid, P. L. Dutton, Biological Electron Transfer. *Journal of Bioenergetics and Biomembranes*. **27**, 263–274 (1995).
35. S. Santabarbara *et al.*, Chlorophyll Triplet States Associated with Photosystem I and Photosystem II in Thylakoids of the Green Alga Chlamydomonas Reinhardtii. *Biochimica et Biophysica Acta (BBA)*. **1767**, 88 (2007).
36. S. Santabarbara, E. Bordignon, R. C. Jennings, D. Carbonera, Chlorophyll Triplet States Associated with Photosystem II of Thylakoids †. *Biochemistry*. **41**, 8184 (2002).
37. I. Vass, Molecular Mechanisms of Photodamage in the Photosystem II Complex. *Biochimica et Biophysica Acta (BBA)*. **1817**, 209 (2012).
38. H. J. M. Hou, in *Photosynthesis. Energy from the Sun* (Springer Netherlands, Dordrecht, 2008), pp. 1473–1478.
39. J. P. Armitage, M. C. W. Evans, Chemotactically Induced Increase in the Membrane Potential of Spheroplasts of Rhodopseudomonas Sphaeroides. *FEBS Letters*. **112**, 5–9 (1980).
40. D. E. Robertson, P. L. Dutton, The Nature and Magnitude of the Charge-Separation Reactions of Ubiquinol Cytochrome C2 Oxidoreductase. *Biochimica et Biophysica Acta (BBA)*. **935**, 273–291 (1988).
41. Z. D. Popovic, G. J. Kovacs, P. S. Vincett, G. Alegria, P. L. Dutton, Electric Field Dependence of Recombination Kinetics in Reaction Centers of Photosynthetic Bacteria. *Chemical Physics*. **110**, 227–237 (1986).
42. C. C. Moser *et al.*, Initial Charge Separation Kinetics of Bacterial Photosynthetic Reaction Centers in Oriented Langmuir-Blodgett Films in an Applied Electric Field. *Chemical Physics*. **197**, 343–354 (1995).
43. Z. D. Popovic, G. J. Kovacs, P. S. Vincett, P. L. Dutton, Electric Field Modulation of Charge Transfer Processes in Reaction Centers of Photosynthetic Bacteria. *Chemical Physics Letters*. **116**, 405–410 (1985).

Part II: Experimental application and testing of optimal engineering blueprints in maquette platform

Preface:

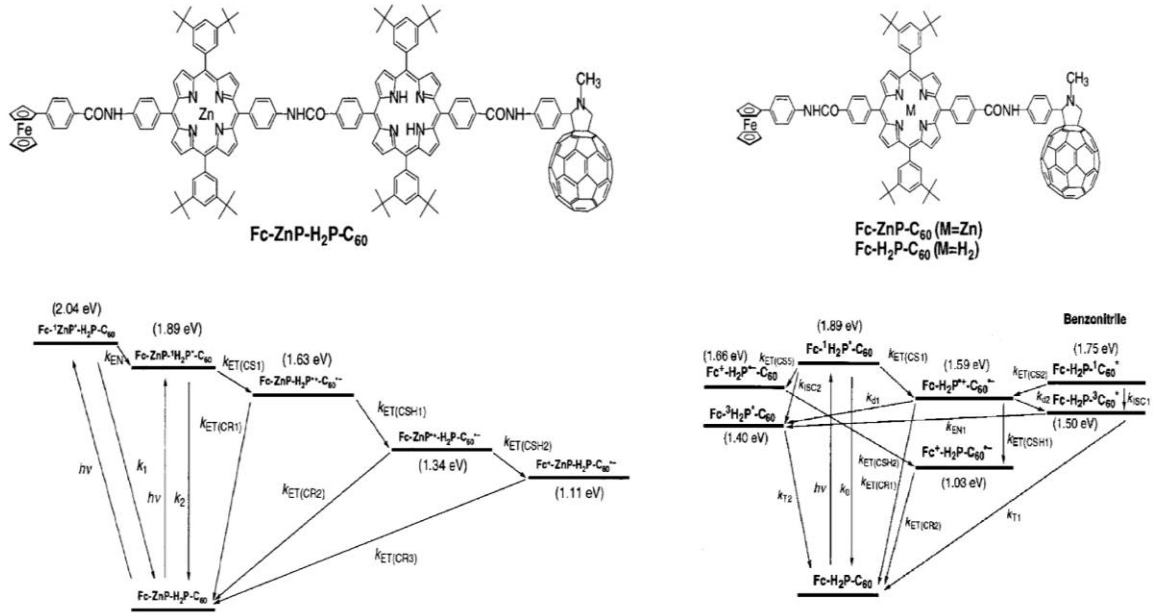
The previous three chapters of this work have focused mainly on exploration of the fundamental physical, biological, biochemical, and biophysical principles, constraints, and approaches that govern the engineering of an optimal charge-separating device within a protein-based, biologically relevant design platform. Although these chapters gave significant consideration to the practical constraints and influences on the engineering of charge-separation devices thanks to nature of the protein environment that natural reaction centers selected, all the simulation conducted have not specifically included any component of the protein environment, and all the biological constraints have been incorporated into the algorithm as heuristics-based modifications of the solutions to the system of ordinary differential equations. The end goal of such engineering blueprints, however, is the practical realization of the optimal design, and consequently, in order to truly test the validity and effectiveness of the engineering principles we have discovered and organized throughout the first half of this thesis, experimental testing and verification are the necessary next steps.

The second part of my thesis therefore focuses on the practical application and examination of the engineering principles and guidelines arrived at in the first three chapters. I begin the empirical testing of the engineering guidelines by a brief exploration of the essential tools for the testing of these principles in Chapter IV, giving an contextualizing summary, by no means comprehensive or complete, that allows the readers to understand and appreciate the choice of maquette as the appropriate platform upon which practical testing of the optimal engineering blueprints. In Chapter V, I present a series of novel experimental techniques successfully applied to maquettes that greatly expanded the horizons of biophysical and biochemical functionalities achievable by maquettes, leading to the successful engineering and construction of a charge-separating triad in a maquette that supports long-lived charge-separation with lifetimes up to 428ms, reported in Chapter VI, a proper capstone for my thesis, demonstrating that the engineering guidelines formulated in the first part of the thesis indeed carries great potential in practical design of charge-separating photosystems. The combined results of part I and II successfully demonstrate that, given the proper experimental infrastructures, the optimal designs prescribed by the engineering guidelines will lead to photosystems built in small water-soluble proteins with unparalleled efficiencies and tolerances to their functional environments.

Chapter IV: Maquette as the ideal versatile platform for engineering of photochemical charge-separating device

4.1: The necessity of de-novo designed protein as engineering platform for optimal photosystems.

The engineering guidelines for long-lived charge separation in multiads detailed in Chapter II and III have demonstrated that a quartet comprising of a photo-activatable pigment along with three additional acceptor molecules arranged in a curved chain with close spacing between the first two acceptors, comprises the most concise form of engineering guidelines for optimal charge-separating device, that's robust enough to deliver near-unity quantum yield at hundreds of milliseconds. Chapter III also demonstrated that natural photosystems largely follow the prescribed optimal engineering, except for the energy expended beyond the amount required by the optimal



Fe
CONH
Zn
NHCO
NH
HN
CH₃
Fc-ZnP-H₂P-C₆₀
M
Fc-ZnP-C₆₀ (M=Zn)
Fc-H₂P-C₆₀ (M=H₂)
(2.04 eV)
Fc-¹ZnP*-H₂P-C₆₀
(1.89 eV)
Fc-ZnP-¹H₂P*-C₆₀
k_EN
k_ET(CS1)
(1.63 eV)
Fc-ZnP-H₂P•+-C₆₀•−
k_ET(CR1)
k_ET(CSH1)
Fc-ZnP•+-H₂P-C₆₀•−
(1.34 eV)
k_ET(CSH2)
Fc•+-ZnP-H₂P-C₆₀•−
(1.11 eV)
hν
k₁
hν
k₂
k_ET(CR2)
k_ET(CR3)
Fc-ZnP-H₂P-C₆₀
Benzonitrile
(1.89 eV)
Fc-¹H₂P*-C₆₀
(1.66 eV)
k_ET(CS5)
Fc⁺-H₂P•−-C₆₀
k_ISC2
k_ET(CS1)
(1.59 eV)
Fc-H₂P•+-C₆₀•−
(1.75 eV)
k_ET(CS2)
Fc-H₂P-¹C₆₀*
k_d1
k_d2
Fc-H₂P-³C₆₀*
k_ISC1
Fc-³H₂P*-C₆₀
(1.40 eV)
k_EN1
(1.50 eV)
k_ET(CSH1)
Fc⁺-H₂P-C₆₀•−
k_ET(CSH2)
(1.03 eV)
k_T2
hν
k₀
k_ET(CR1)
k_ET(CR2)
k_T1
Fc-H₂P-C₆₀

Scheme 4.1: Long-lived charge-separation in synthetic supramolecular triads and tetrads. Top Left: Charge-separating tetrad consisting of a Zn-porphyrin, free-base porphyrin, ferrocene, and fullerene. Bottom left: electron transfer kinetics of the tetrad, with the lifetime of the CS determined to be 380ms. Top Right: Charge-separating triad, same as the tetrad minus the free-base porphyrin. Bottom Right: electron transfer kinetics of the triad, with the lifetime of the CS determined to be 16us. Reproduced from (*7*) engineering guidelines to compensate for the membrane potential effect(see Chapter III) and to further ensure that no uphill repopulation of the intermediate CS states takes place. However, the final product of a single natural photosystems' charge-separation cannot be directly used by mankind, and the immense structural complexity, coupled with their native trans-membrane environment, makes any effective and efficient reengineering of the natural photosystems too difficult.

Because of the abovementioned difficulty of reengineering existing natural photosystems into small and water-soluble devices with the same ability to separate charge highly efficiently at long lifetimes, designs and construction of electron-transfer based charge separating device consisting of multi-ads of redox active cofactors have been largely limited to synthetic supramolecules. Within the last two decades, extensive efforts have been exerted on constructing such synthetic multi-ads.(*1-6*) While these systems have the capabilities to achieve very long charge separation, (lasting for >100ms), they require extensive effort in terms of complex preparation as the synthetic process requires incremental incorporation of cofactors one after another. This naturally leads to a multiplication of necessary steps in the synthesis scheme and consequently a prohibitive low yield and high costs, severely hampering the practicality of such synthetic multi-ads in their application for solar energy conversion. For example, the reported scheme for the photochemical tetrad that separates charges for 380ms by Fukuzumi et al. (*7*)consists of 8 separate steps, with an accumulative yield of 0.61%, assuming perfect transition between the steps. In fact, the scale of the scheme begins on the order of 10g and ends with a total of 17mg of product.

Moreover, the synthetic multi-ads that have been reported to achieve >100ms charge-separation all include more than three cofactors. The longest lifetime of charge-separation reported in a synthetic triad is only 16us(*8*),as shown in Scheme 4.1. This suggests that the designs of the chemical supramolecular photosystems have not followed the optimal engineering guidelines developed in the first part of this thesis, and, more importantly, development of a more appropriate engineering platform can address this shortcoming.

In addition to the difficulty of construction, such synthetic chemical systems have very limited means to modulate distances between the cofactors, the fundamental parameter that influences the yield and lifetime of charge-separation. Literature suggests that the most reliable method of distance modulation for such synthetic multi-ads is via incorporation of additional cofactors, as simply increasing the covalent chemical linker between the cofactors results in loss of rigidity of the device and therefore would not contribute to increase in inter-cofactor distances. Yet incorporation of additional cofactors to extend distances inevitably leads to further complication of the synthetic complexity and therefore practicality of the device.

Finally, it remains to be seen how readily such synthetic chemical devices could be linked with the corresponding catalytic components that would allow the stable charge separation to achieve its desired function, since creating such linkage requires further steps in the final synthetic schemes, leading to further reduction in the already low yield of the synthesis.

Consequently, engineering of charge-separation multi-ads in simple, small, but structurally and functionally stable protein scaffolds that also offer significant degree of fine-grained control over its structural details stands out as the appropriate practical approach. A few well-tested experimental methods in de-novo engineering of proteins serve as potential candidates, which I will describe in the following section.

4.2: Overview of major approaches in de-novo protein design.

Design and engineering of de-novo protein is not a novel idea, and in fact, the first appearance of this concept goes back to 1902, when Emil Fischer delivered his Nobel prize speech, claiming that the future of “physiological chemistry will not only make extensive use of the natural enzymes as agents, but when it will also prepare synthetic ferments for its purposes” .(*9*) Since then, numerous researches from both experimental and computational perspectives have undertaken the quest for the Holy Grail of intelligent, effective, and complete method of de-novo protein design. Among the efforts three main approaches can be identified, based on their attitudes towards existing natural proteins: repeated mutagenesis followed by screening and selection with natural protein sequences and motifs as starting points, computational optimization of permutation and concatenation of short natural protein sequences, and first-principle heuristics based sequence design free of mimicry of natural proteins.

In the first approach, also known as directed evolution(*10*, *11*), the DNA of an existing protein is used as the starting point for a set of error-prone mutagenesis that generates a library of specific size, usually on the order of 10^{6-10} different sequences. The quality and success of directed evolution largely depends on the ability of the library-generation to sample as uniformly and as widely as possible the relevant protein sequence space. High throughput screening using various functional assays based on the desired enzymatic activity to be engineered and sequencing is then employed as the method of artificial selection. In addition to functional assays, in-vivo compartmentalization of the target protein followed by subsequent fluorescence-based sorting of the cell expressing the protein has also become available for functions that are not easily assayed directly. The process is then repeated through certain number of iterations until either the functional assay metrics converges or the improvements in metrics per iteration per cost becomes so small that it can no longer justify further cycles. The pipeline of directed evolution is summarized in Figure 4.1.

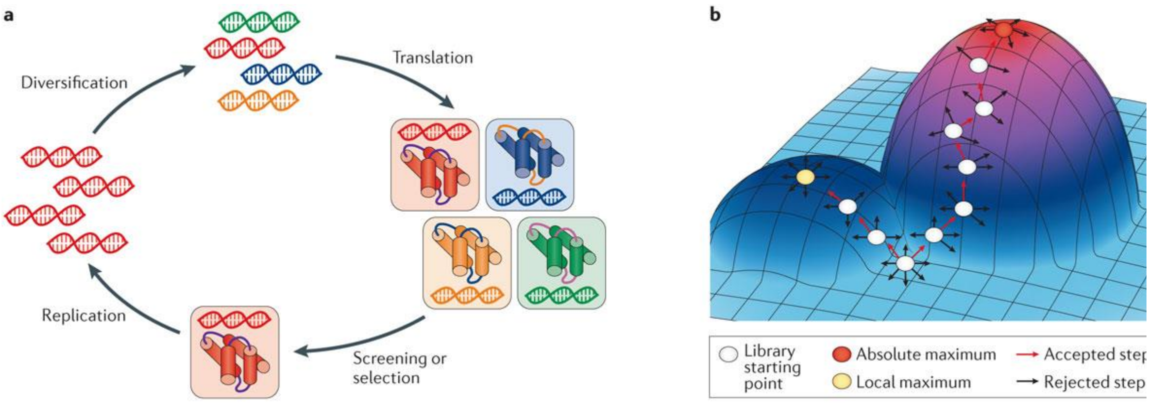


Figure 4.1: The iterative processes of directed evolution. **a** | The process of directed evolution in the laboratory mimics that of biological evolution. A diverse library of genes is translated into a corresponding library of gene products and screened or selected for functional variants in a manner that maintains the correspondence between genotype (genes) and phenotype (gene products and their functions). These functional genes are replicated and serve as starting points for subsequent rounds of diversification and screening or selection. **b** | Although the mutational space is multidimensional, it is conceptually helpful to visualize directed evolution as a series of steps within a three-dimensional fitness landscape. Library generation samples the proximal surface of the landscape, and screening or selection identifies the genetic means to 'climb' towards fitness peaks. Directed evolution can arrive at absolute maximum activity levels but can also become trapped at local fitness maxima in which library diversification is insufficient to cross 'fitness valleys' and access neighboring fitness peaks. Figure and legends reproduced from (10)

In the second approach, best exemplified and represented by the “Rosetta” method(*12-14*), de-novo design of protein with catalytic functions is accomplished via a cascade of computational optimizations. The first optimization involves a “grafting” process where a “theoenzyme” of desired catalytic function, abstracted in the format of a series of coordinates and constraints, is matched with the optimal protein scaffold sequence from known proteins to provide the best fitting active site for the function. After a candidate scaffold has been selected, the next step involves repeated application of two optimization processes in series. First applied is the minimization of a limited energy function involving the backbone and the catalytic residues of the active site. It returns the best position for the substrate to interact with the active site. It is followed by the minimization of the same energy function while trying out all possible amino-acids for residues in the active site that are not directly catalytic. The new active site is now used as input for another iteration of the optimization that searches for the optimal locations for the residues involved in the actual catalysis. Then a repeated optimization for the best sequence takes place. After 3-4 iterations a final optimization without heavy constraints in the earlier steps is executed to ensure that the design has converged to a structure that indeed produces a stable active site. Due to the stochastic nature of the sequence and position sampling of the Rosetta algorithm, each design process, given the same initial condition, produces a different output. Consequently, it is usual to consider many design outputs and select the optimal design after examining and comparing their projected activities using molecular dynamics. The entire process of protein design by Rosetta is summarized in Figure 4.2.

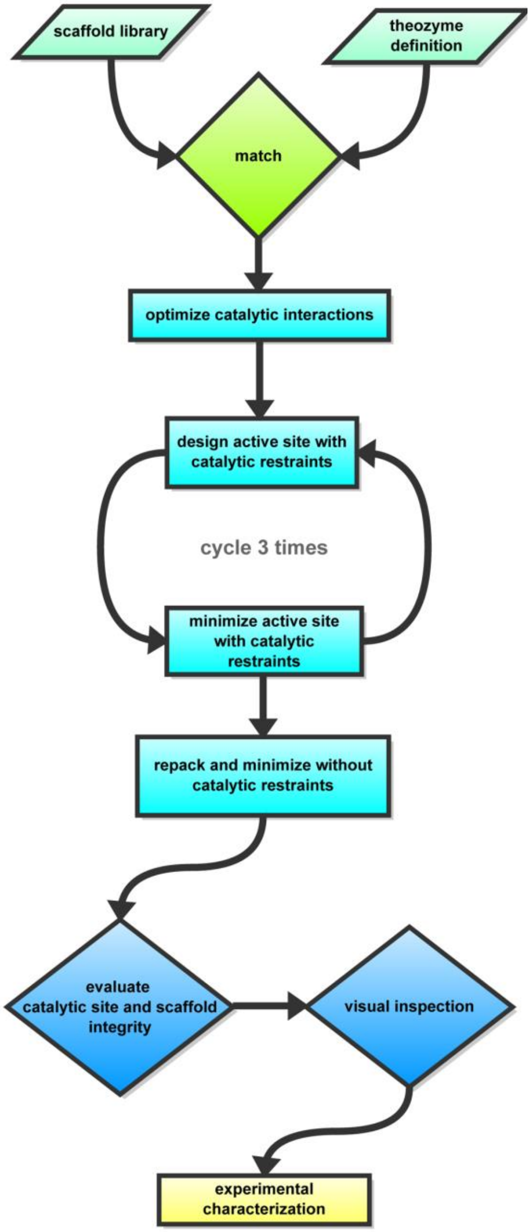

scaffold library
theozyme definition
match
optimize catalytic interactions
design active site with catalytic restraints
cycle 3 times
minimize active site with catalytic restraints
repack and minimize without catalytic restraints
evaluate catalytic site and scaffold integrity
visual inspection
experimental characterization

Figure 4.2: Iterative process of computational protein design using Rosetta3. Figure reproduced from (12).

While both methods and even combinations of the two have at times generated impressive results that are regarded as milestones in protein design, such as the genesis of a novel catalyst for a chemical reaction previously never catalyzed by any natural agent, none of the designs have produced enzymes that deliver catalytic performance on par with or even within orders of magnitude of the natural counterparts. The lack of successes have been largely attributed to the enormous difference between the size of the search space of protein sequences (on the order of 10^{300} for a average-sized protein) and the ability to sample this space by directed evolution.(*15*) Moreover, the Rosetta design protocol depends on a reliable and clear understanding of the structure and property of the ideal active site, which relies on prior knowledge that is independent of the design process. Similarly, the Rosetta method’s initial matching process, due to practical constraints on our current computational capabilities, cannot sample the entire sequence of the candidate protein scaffold, but instead must require the user to judiciously select a subset of the sequence. Together these two steps indicate that the Rosetta algorithm in fact uses a mixture of heuristics and stochastic sampling of the sequence space. This in turn implies that it is not necessary and sufficient, even given enough computational resources, Rosetta would be able to discover the optimality in the sequence space.

More importantly, as recently stated by Woolfson (*16*), the real sequence space of naturally occurring proteins is also highly non-uniform, with the probability masses of optimally functional sequences concentrated in a very small subspace of the entire possible "protein universe", as shown in Figure 4.3. While the lack of optimal sequences in the space where natural proteins did not favor is still heavily debated, it is straightforward to show that sequence space known to produce functional structure is adequate in supporting and sustaining an immense variety of functions, and this known sequence space can be summarized by a simple set of biophysical and biochemical principles while abstracting away the fine details of the individual amino-acids.

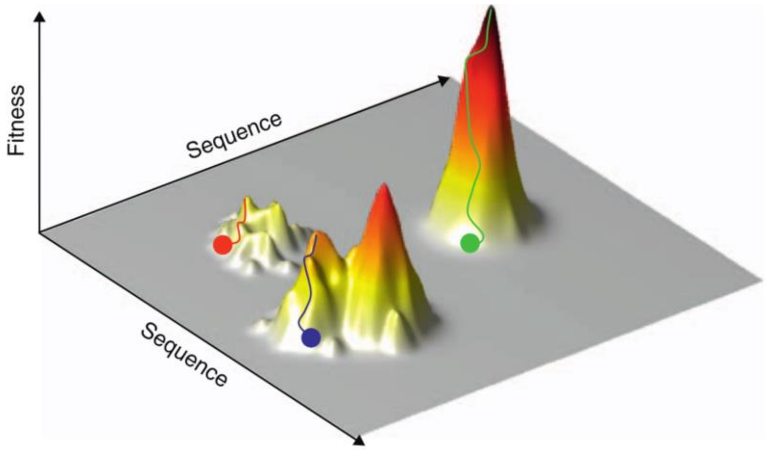


Figure 4.3: The highly non-uniform distribution of probability masses for the optimal sequences in the protein sequence space. The protein sequence space is shown in a simplified, abstract 3D space. Fitness measures the optimality of the specific sequence. Figure reproduced from (*16*)

While the previously mentioned approaches could not free itself fully from heuristic elements, we here introduce a third approach that completely isolates itself from any stochastic element in the design process, and focuses entirely on rule-based, heuristics-central engineering revolved around fundamental biophysical and biochemical principles that determine protein secondary and tertiary structures as well as protein-cofactor interactions and cofactor-based protein functions. The comparison and contrast between the methods of protein design discussed so far are summarized in Figure 4.4.

4.3 The Maquette approach: origin and development

The so-called "maquette" approach takes on the challenge of protein design from a heuristic, first-principle perspective. Rather than trying to discover the optimality of design within the entire protein sequence space or a significant subspace of it, the maquette approach limits itself to highly specific and special regions of the sequence search space where the probability of discovering sequence optimality for the desired

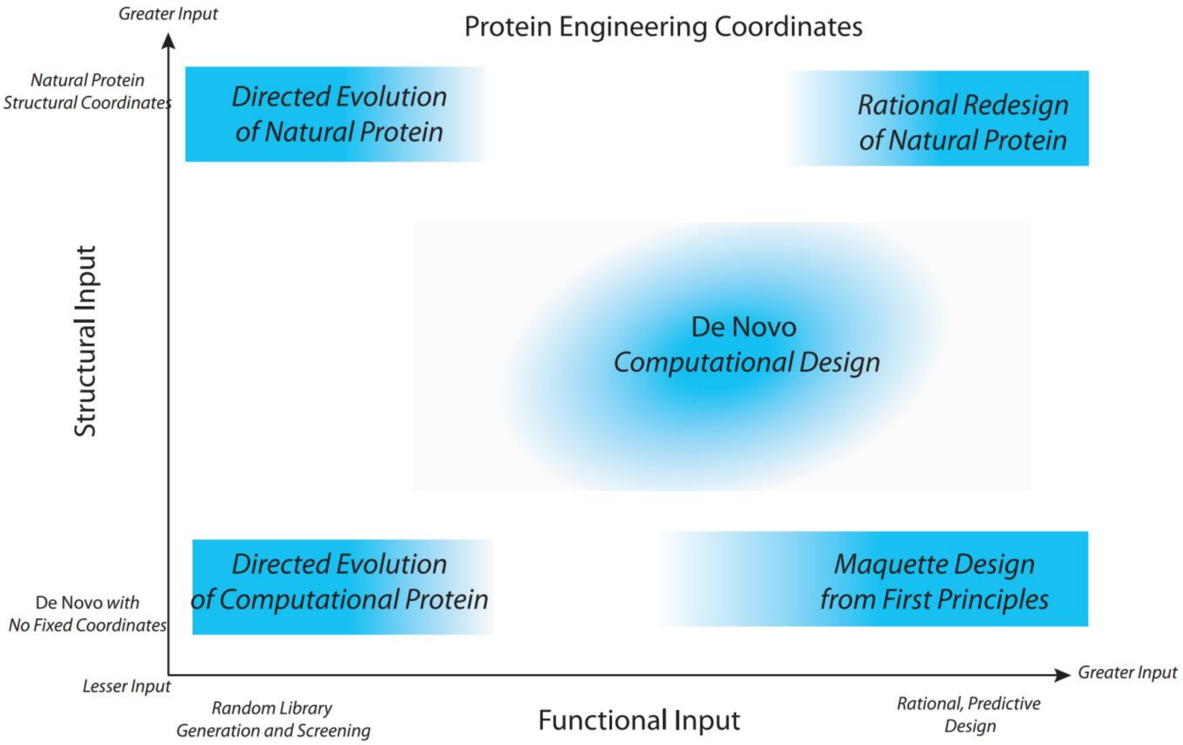

Protein Engineering Coordinates
Greater Input
Natural Protein Structural Coordinates
Directed Evolution of Natural Protein
Rational Redesign of Natural Protein
Structural Input
De Novo Computational Design
Directed Evolution of Computational Protein
Maquette Design from First Principles
De Novo with No Fixed Coordinates
Lesser Input
Greater Input
Random Library Generation and Screening
Functional Input
Rational, Predictive Design

Figure 4.4: A 2-dimensional map illustrating the relative positions of major protein design approaches with respect to the type and degree of constraints imposed upon the design, or the strength of the prior applied to the distribution of the sequence space. The horizontal axis represents the level of heuristics based constraints based on fundamental biophysical and biochemical principles, while the vertical axis indicates the level of constraints on the sequence space based on heuristics of naturally occurring sequences. Maquette approach relies on strong biophysical/biochemical heuristic constraints but no natural protein based constraint, while directed evolution is the opposite. Figure adapted from the PhD thesis of Tammer Farid.

function is expected to be high according to a set of heuristics. Moreover, rather than comprehensively searching through the now limited sequence space using random or computational methods, the maquette approach further restricts and contracts the sequence space by increasing the granularity of the sequence from amino-acid level to biochemical-property level, significantly reducing the complexity of the optimization problem. From a Bayesian perspective, the maquette approach implements a set of strong priors on the problem of protein design and generates a posterior distribution of the sequence space that's highly inhomogeneous and non-isotropic. Together these characteristics make maquette approach the least complex and at the same time most reliable method of the three. At the same time, because of the strong prior on the sequence space, the maquette approach is limited in its scope of structural and functional diversity in comparison to the other methods, making it high attractive for only a small subset of protein functions and less attractive for the rest.

The maquette approach began in 1990s as an attempt to construct a functionally versatile but structurally stable testing ground for oxidoreductase functions.(*17, 18*) From the beginning on, the maquette approach has chosen 4-alpha-helical bundle as the fundamental structural requirement, because the engineering requirements for 4-helical bundles were well-understood and attractively simple, at the same time 4-helical bundles offering a stable yet non-rigid platform upon which various functions could be implemented without significant departures from the fundamental. This decision immediately reduces the search space for the optimal sequences for the maquette approach, limiting it to a small sub-space of the much larger search space considered by the methods described above. However, this reduction in the search space for sequences does not automatically translate to a reduction in the potential range of functions that can be designed in maquettes. Alpha-helical bundles can provide protein environments ranging from highly hydrophobic to hydrophilic, and active site designs for numerous functionalities are possible since alpha-helix represents one of the fundamental and most common secondary structures for protein.

The choice of helical-bundle structural motif induces the first and also the strongest set of rule-based priors on the sequence space: binary patterning and heptad repeat, as seen in Figure 4.6. When hydrophilic and hydrophobic residues are placed alternatingly in a 7-residue repeat in an aqueous solution,(*19*, *20*) the electrostatic and hydrophobic interactions between the solvent and the protein itself force the polypeptide to adopt a helical secondary structure, with the hydrophilic residues on one side of the helix and the hydrophobic ones on the other, buttressed by the hydrogen-bonds formed between the backbones of residues on the same side of the helix. When multiple such polypeptides are present, the hydrophilic and hydrophobic interactions further compel them to form tetramers that create a hydrophobic core isolated from the hydrophilic exterior. However, the tetrameric version of the helical bundle does not allow the engineering of the topology of the maquette, as the direction of monomer in the tetramer is determined based on thermodynamic partitioning and random chance. The next set of rule-based priors, therefore, is the placement of loop residues between 4 near-repeats of the would-be monomeric polypeptide, resulting in a 4-helical bundle consists of a single polypeptide chain with a single topology. This is necessary to ensure that the intended anchoring positions of cofactors to achieve the desired function follows the specification of the design. Until now, these two sets of rule-based sequence constraints have provided the necessary basis for constructing a stable and functionally neutral scaffold.

The next sets of rule-based sequence constraints are intended to supply the functional specifications of the design. These include but are not limited to selections of: primary and secondary residues that comprise the binding sites of cofactors for the desired function; level of hydrophobicity of the hydrophobic interior of the 4-helical bundle; overall charge-patterning of the exterior of the bundle; length of the loops between the component helices of the bundle; level of hydrophobicity of the hydrophilic exterior of the bundle; and so on.

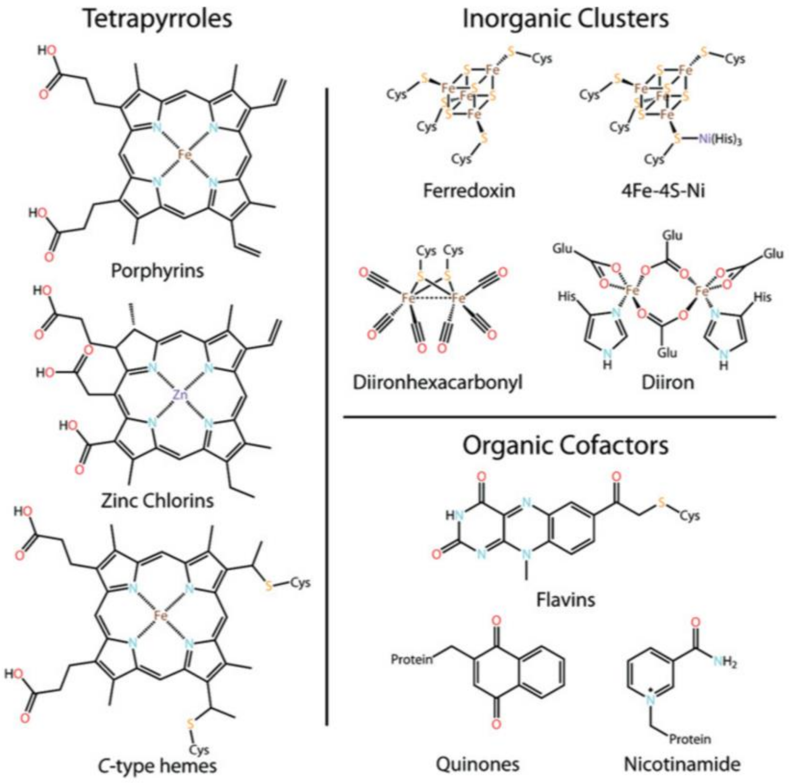

Tetrapyrroles
Inorganic Clusters
Porphyrins
Zinc Chlorins
C-type hemes
Ferredoxin
4Fe-4S-Ni
Diironhexacarbonyl
Diiron
Organic Cofactors
Flavins
Quinones
Nicotinamide

Figure 4.5: Cofactors successfully incorporated into maquettes categorized by their chemical properties. The tetrapyroles have been the main focus of cofactor bindings in the past, but the other cofactors have also demonstrated their fitness for maquette applications. Reproduced from (*21*)

The primary residues of the cofactors binding sites simply determine the location as well as the type of cofactors required by the desired function. (*22-24*)Since the cofactors of interest in maquette designs are mostly redox active molecules, the primary binding sites are usually on positions that are facing the hydrophobic interior of the bundle. Up to now, binding sites for metal tetrapyrroles such as heme, zn-porphyrin, chlorin, and so on, have allowed for high affinity ligation for these cofactors in various maquettes.(*25*) The secondary residues of cofactor anchoring, however, require significantly more planning and thoughts. They are the residues that provide either stabilizing or destabilizing effects on the cofactors of interest. For example, the binding site of a heme cofactor requires the residue at the i+4 position to be of hydrophilic nature in order to increase binding affinity(*26*). In addition to tetrapyroles, a large variety of other cofactors have also demonstrated successful incorporation and functions in maquettes. Figure 4.5 demonstrates the multitude of cofactors that have been shown to have high affinity binding to the maquettes(*21*).

The hydrophobicity of the protein interior determines the rigidity and flexibility of the overall maquette structure, as well as its thermal stability. This characteristic of the maquette must be modulated depending on its intended function. Many functions require or prefer the local environment of the functional cofactors to be highly dynamic and flexible, thereby allowing facile exchange of ligands or access to substrates. For example, the binding of heme to maquettes that have histidine binding sites on adjacent helices require the interior to be dynamic enough that concerted outward rotation can take place so that the proper His rotamer can form(*27*, *28*). This therefore requires a mixing of polar and non-polar residues in the interior-facing part of the maquette to anticipate the helical rotation prior to heme binding. In general, when the overall maquette structure is beyond a certain threshold of rigidity, it has difficulty in rearranging itself structurally to accommodate for the cofactor, which, although mostly hydrophobic, could have minor structural peculiarities that would require the maquette to dynamically adapt. Finally, in the context of electron transfer, highly hydrophobic protein environments correspond to low dielectric environment and tend to decrease the reorganization energy for the electron transfer.

The charge-patterning of the maquette surface represents another significant source of sequence space constraints that depend on the intended function of the maquette. The distribution as well as number of charged residues on the exterior facing parts of the maquette directly determines its ability to interact with itself or, when present, another partner protein.(*29*) This becomes important when the maquette is designed to serve either as a donor or an acceptor in functions requiring inter-protein electron transfers. Highly charged surfaces allow the protein to interact strongly with its intended partner and consequently increase the rate of electron transfer. Moreover, specifically increasing the charges of a region of the maquette can preferentially strengthen the inter-protein interaction between the maquette and its intended partner at the intended site, giving further options to engineering. As an example, Bryan Fry of the Dutton group has designed maquettes whose varied charge patterning and distribution on their exteriors allow them to differentially interact with natural proteins such as cytochrome c(Manuscript in preparation). When equipped with the appropriate cofactors, the interactions with other protein partners allow the maquette to transfer electrons from and to the appropriate partner proteins. This is highly desirable in the context of engineering charge-separating photosystems in maquettes, as the partnership with other proteins removes the necessity of engineering the down-stream catalytic function in the same maquette that also houses the charge-separating photosystem. However, discussion of inter-protein electron transfer is outside of the scope of this thesis, but has been an ongoing effort and the center of studies by my colleague Bryan Fry in the Dutton lab, whose thesis will focus on the interaction between electron-transfer competent maquettes and natural proteins.

The length of the loops connecting the helices has been shown to strongly modulate both the thermal stability of the protein as well as its affinity to ligands whose ligation sites are engineered at the ends of the maquette.(*30*) Longer loops allow more dynamic movements of the individual helices and therefore decrease the rigidity of the maquette. The increased dynamic movements in turn lead to a less stable hydrophobic core of the helical bundle of the maquette, resulting in changes in cofactor affinities. Moreover, the loop region, completely exposed to the aqueous environment surrounding the protein, serves as ideal sites for catalytic modules such as cobalt-clusters, although as already mentioned above, the need to engineer catalytic site in the same maquette already tasked with charge-separation is small, and therefore the structurally regulatory function of the length of the loops remains the main constraint on the sequence space of the loop residues.

Finally, the exterior of the maquette could be engineered to comprise mostly of hydrophobic residues that are usually found in the exterior of membrane proteins.(*31*-*33*)) This option allows for engineering of membrane-friendly maquettes that are either completely embedded in membranes or partly embedded, exposing hydrophilic ends and thereby acting as trans-membrane maquettes. This engineering option greatly expands the horizon of potential maquette function, including artificial ion channels, receptors, and any biological functions associated with membrane or trans-membrane proteins.

4.4 Maquettes as the ideal testing platform of design blueprints and engineering guidelines for photochemical multi-ads for charge-separation

Because of the context of its historical development, the maquette approach has been applied mostly to protein functions and enzymatic activities that revolve around anchoring and exploiting redox-active cofactors: oxidation and reduction of substrate molecules(*34*), light-harvesting(*35*), and since its first application, more than three decades of continual research efforts have been devoted to understanding and improvements of the heuristics that provide the proper priors for the sequence space, as extensively discussed above. Together with the stability of the helical bundle structural motif and the functionally neutral nature of maquettes as a scaffold, they make maquette approach the ideal choice for designing and engineering protein-based platform to test the engineering blueprints for charge-separation devices demonstrated in the first part of this thesis.(*36*) Indeed, our engineering guidelines for man-made photochemical multi-ads, aimed to unlock the full potential of the catalytic quartet, consider the placement of a set of optimally selected set of redox-active cofactors and pigments within a non-specific and highly stable water-soluble protein environment to be essential to the successful design. The functionally neutral maquette further excels as the platform of choice due to the quantized but exact control over the location of the cofactors, allowing the engineering to modulate both the distances between and the redox potentials of the cofactors. Additionally, the simple, minimalist nature of the maquettes, in terms of its sequence, structure, and size, removes possible interferences from otherwise complex and delicate protein environment that could impact the testing of engineering blueprints. Finally, the practical ease and speed with which maquettes could be expressed and reengineered means that experimental examination of the design blueprints can proceed with high efficiency. This allows us to exploit an iterative procedure involving design-test-redesign that can ensure that we converge to the design that's optimal and most appropriate with respect to both the theoretical specifications and practical constraints. This iterative aspect is summarized in Figure 4.6, with certain steps, such as III.2 and III.3, variable to a great extent depending on the intended function of the maquette designed.

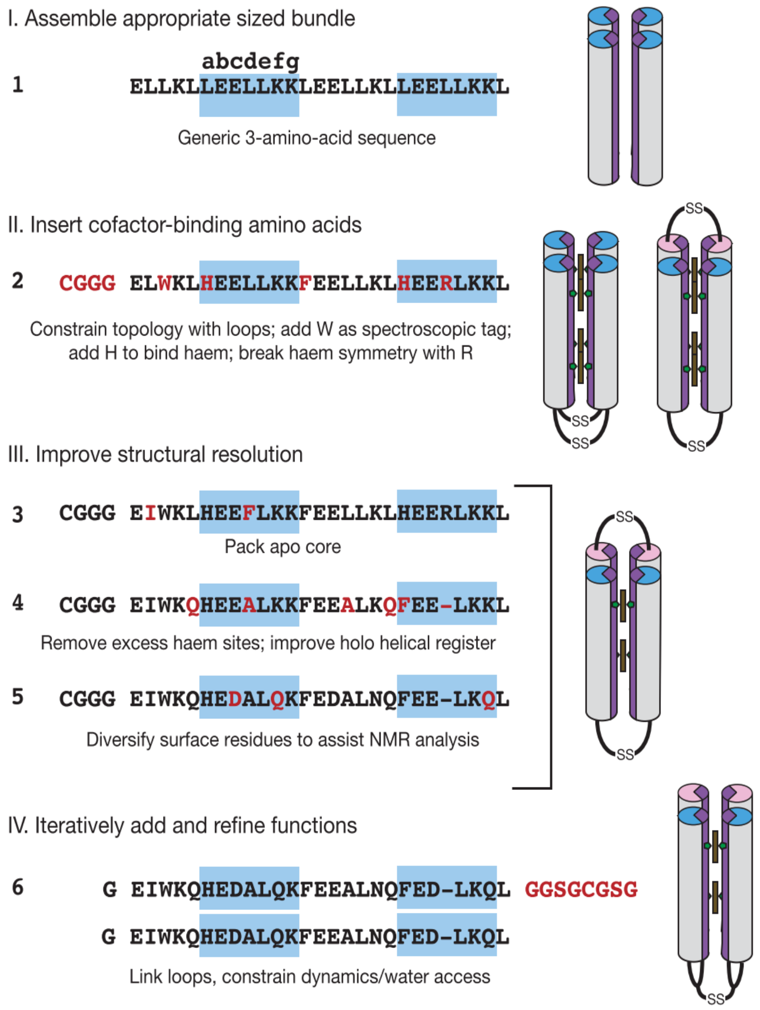

I. Assemble appropriate sized bundle
abcdefg
1
ELLKLLEELLKKLEELLKLLEELLKKL
Generic 3-amino-acid sequence
II. Insert cofactor-binding amino acids
2
CGGG ELWKLHEELLKKFEELLKLHEERLKKL
Constrain topology with loops; add W as spectroscopic tag;
add H to bind haem; break haem symmetry with R
SS
SS
SS
SS
III. Improve structural resolution
3
CGGG EIWKLHEEFLKKFEELLKLHEERLKKL
Pack apo core
4
CGGG EIWKQHEEALKKFEEALKQFEE-LKKL
Remove excess haem sites; improve holo helical register
5
CGGG EIWKQHEDALQKFEDALNQFEE-LKQL
Diversify surface residues to assist NMR analysis
SS
SS
IV. Iteratively add and refine functions
6
G EIWKQHEDALQKFEEALNQFED-LKQL GGSGCGSG
G EIWKQHEDALQKFEEALNQFED-LKQL
Link loops, constrain dynamics/water access
SS

Figure 4.6: The iterative process of the maquette approach. Major steps are listed using Roman numerals while minor steps are in listed in Arabic numbers. Figure modified and reproduced from(28)

4.5 Structured, stable, versatile, and adaptable: The newest generation of maquette with crystal structure

A common and persistent criticism of the maquette approach had been the lack of high resolution structural information to confirm the realization of the rule-based design. This lack of detailed structural information has largely been the caused by past designs' choices of design priors that prioritize the functional neutrality and versatility of the design product. That is, these designs did not choose a single function and employ a large collection of engineering devices to ensure that the product of the design delivers only that single intended function. Consequently, the resulting maquettes serve as ideal scaffolds not only for its intended function, but also many other unintended yet similar functions. Recently, a new generation of maquettes, designed with a highly specific function in mind, has become the first family of single-chain maquettes to produce high-quality crystal that produces high-resolution X-ray diffraction and subsequently 3D structures. Discussion of the specifics of this family of maquettes will soon be seen in the thesis of my colleague Nathan Ennist and will therefore be kept brief here. As seen in Figure 4.7, the design of this maquette aimed to create a high-affinity, high specificity iron-tetrapyrole with a diagonally ligated bis-his ligation site residing on one end of the maquette, a single-his, high-affinity zinc-tetrapyrole binding site in the center, and two auxiliary functional sites consisting of a tyrosine residue and metal binding center on the other end of the maquette. Rather than using simple heptad repeats to ensure the helical structure and following the streamline shown in Figure 4.6, this generation of maquettes incorporated additional designs that are residue specific, considering the primary and secondary shells of ligation for both of the tetrapyroles, as well as the steric and charge-based stabilization of cofactor binding. With the available crystal structure, this maquette has become the new workhorse of the Dutton group's experimental side, as the structural information allows for accurate and targeted changes in the structure for either modification of existing or introduction of new functions. As will be seen in Chapter V and VI, this maquette, along with the various mutants designed to examine the principles of electron transfer, will serve as the platform upon which the proof-of-principle set of triads are engineered and tested.

4.6 Conclusion:

This chapter offered an overview of the major methods of de-novo design and engineering of proteins, all of which represent different approaches to the highly daunting problem of optimizations of a fitness/functionality measure within the search-space of all available protein sequences. Amongst the available approaches, the maquette approach

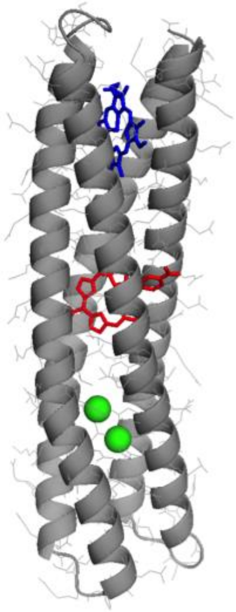


Figure 4.7: X-ray crystal structure of the water-soluble maquette MZH3 at a resolution of 2.02 Å. The protein is ligating a heme (blue), a synthetic zinc porphyrin (red), and 2 Mn atoms (green). Figure adapted from crystal structure solved by Nate Ennist.

was selected as the most appropriate because it offers the greatest level of distance and redox E_m control while incurring the least amount of experimental cost in terms of both computational requirement and the amount of time involved in developing the needed platform. The task of engineering a charge-separating photosystem as specified by the engineering guidelines from Part I of this thesis presents a set of requirements that are very different from other de-novo protein design that are better addressed by either directed evolution or Rosetta-based computational design or a combination of both, namely, the lack of a theo-enzyme, a prototype active site and a well-defined single reaction. The choice of the maquette approach as the ideal platform is further encouraged by the availability of a structurally elucidated family of maquettes, which offers unprecedented simplicity and accuracy with which target functional modification of the maqutte can be carried out. This provides the necessary experimental infrastructure to not only construct a single proof-of-principle photochemical triad to demonstrate that the predicted performance jump from dyads to triads can be realized, but also, more importantly, to examine the effects of varying some of the engineering parameters on the performance of the charge-separation in practice.

4.6 References:

1. Darius Kuciauskas *et al.*, *Photoinduced Electron Transfer in Carotenoporphyrin−Fullerene Triads: Temperature and Solvent Effects* (American Chemical Society, 2000), vol. 104.
2. T. A. Moore *et al.*, Multistep Photoinitiated Charge Separation In A Molecularpentad. *[1990] Proceedings of the Twelfth Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, doi:10.1109/iembs.1990.691991.
3. S. Caffarri, T. Tibiletti, R. C. Jennings, S. Santabarbara, A Comparison Between Plant Photosystem I and Photosystem II Architecture and Functioning. *Curr Protein Pept Sc*. **15**, 296–331 (2014).
4. K. Ohkubo *et al.*, Fullerene Acting as an Electron Donor in a Donor–Acceptor Dyad to Attain the Long-Lived Charge-Separated State by Complexation with Scandium Ion. *Chem. Commun.*, 589–591 (2007).
5. K. Ohkubo *et al.*, Production of an Ultra-Long-Lived Charge-Separated State in a Zinc Chlorin–C60 Dyad by One-Step Photoinduced Electron Transfer. *Angewandte Chemie*. **116**, 871–874 (2004).
6. S. Fukuzumi, K. Ohkubo, T. Suenobu, Long-Lived Charge Separation and Applications in Artificial Photosynthesis. *Accounts Chem Res*. **47**, 1455 (2014).
7. Charge Separation in a Novel Artificial Photosynthetic Reaction Center Lives 380 ms. *Journal of the American Chemical Society*. **123**, 6617–6628 (2001).
8. H. Imahori *et al.*, Modulating Charge Separation and Charge Recombination Dynamics in Porphyrin Fullerene Linked Dyads and Triads: Marcus-Normal Versus Inverted Region. *Journal of the American Chemical Society*. **123**, 2607–2617 (2001).
9. H. Grünewald, Nobel Lectures Chemistry 1901–1921 und 1922–1941. Herausgeg. von der Nobel Foundation. Elsevier Publishing Company, Amsterdam-London-New York 1966. Band 1901–1921: XII, 409 S., mehrere Abb., geb. Dfl. 80-; Band 11922–1941: 536 S., mehrere Abb., geb. Dfl. 80. *Angewandte Chemie*. **80**, 52 (1968).
10. N. J. Turner, Directed Evolution Drives the Next Generation of Biocatalysts. *Nat. Chem. Biol.* (2009).
11. P. A. Romero, F. H. Arnold, Exploring Protein Fitness Landscapes by Directed Evolution. *Nature Reviews Molecular Cell Biology*. **10**, 866 (2009).
12. D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* (2008).
13. A. Leaver-Fay, S. Bjelic, D. Baker, S. D. Khare, F. Richter, De Novo Enzyme Design Using Rosetta3. *PLoS ONE*. **6**, e19230 (2011).
14. P. Gainza, H. M. Nisonoff, B. R. Donald, Algorithms for Protein Design. *Curr. Opin. Struct. Biol.* **39**, 16–26 (2016).
15. D. Baker, An Exciting but Challenging Road Ahead for Computational Enzyme Design. *Protein science* (2010).
16. D. N. Woolfson *et al.*, De Novo Protein Design: How Do We Expand into the Universe of Possible Protein Structures? *Curr. Opin. Struct. Biol.* **33**, 16–26 (2015).
17. C. Choma *et al.*, Design of a Heme-binding Four Helix Bundle. *Journal of Inorganic Biochemistry*. **51**, 58 (1993).
18. T. P. Quinn, D. C. Richardson, J. S. Richardson, De Novo Design, Expression and Characterization of Felix: A Four-helix bundle proteins of native-like sequence (1991), doi:10.2210/pdb3flx/pdb.
19. S. F. Betz, W. F. DeGrado, Controlling Topology and Native-like Behavior of de Novo-designed Peptides: Design and Characterization of Antiparallel Four-stranded Coiled Coils. *Biochemistry*. **35**, 6955–6962 (1996).
20. R. Fairman, G. McLendon, M. H. Hecht, A Protein Designed by Binary Patterning of Polar and Nonpolar Amino Acids Displays Native-like Properties. *J. Am. Chem. Soc.* **119**, 5302 (1997).
21. B. R. Lichtenstein *et al.*, Engineering Oxidoreductases: Maquette Proteins Designed from Scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).
22. B. R. Gibney, P. L. Dutton, De novo Design and Synthesis of Heme Proteins. *Advances in Inorganic Chemistry* (2000).

23. J. M. Shifman, B. R. Gibney, R. E. Sharp, P. L. Dutton, Heme Redox Potential Control in de Novo Designed Four-α-Helix Bundle Proteins †. *Biochemistry*. **39**, 14813 (2000).
24. B. R. Gibney, S. E. Mulholland, F. Rabanal, P. L. Dutton, Ferredoxin and Ferredoxin-Heme Maquettes. *PNAS*. **93**, 15041–15046 (1996).
25. M. R. Razeghifard, T. Wydrzynski, Binding of Zn−Chlorin to a Synthetic Four-Helix Bundle Peptide through Histidine Ligation. *Biochemistry*. **42**, 1024 (2003).
26. C. Negron, C. Fufezan, R. L. Koder, Geometric Constraints for Porphyrin Binding in Helical Protein Binding Sites. *Proteins: Structure, Function, and Bioinformatics*. **74**, 400–416 (2009).
27. J. Zhuang *et al.*, Design of a Five-Coordinate Heme Protein Maquette: A Spectroscopic Model of Deoxymyoglobin. *Inorg Chem*. **43**, 8218 (2004).
28. R. L. Koder, J. Anderson, L. A. Solomon, K. S. Reddy, Design and Engineering of an O2 Transport Protein. *Nature* (2009).
29. B. A. Fry, L. A. Solomon, P. L. Dutton, C. C. Moser, Design and Engineering of a Man-made Diffusive Electron-Transport Protein. *Biochim. Biophys. Acta*. **1857**, 513–521 (2015).
30. A. D. Nagi, L. Regan, An Inverse Correlation Between Loop Length and Stability in a Four-helix-bundle Protein. *Folding and Design*. **2**, 67 (1997).
31. B. A. Fry, G. R. Wiedman, C. C. Moser, P. L. Dutton, B. M. Discher, Design of Transmembrane Electron Transport Chain within Amphiphilic Protein Maquettes. *Biophysical Journal*. **100**, 131a (2011).
32. Bohdana M Discher *et al.*, *Design of Amphiphilic Protein Maquettes: Controlling Assembly, Membrane Insertion, and Cofactor Interactions* (American Chemical Society, 2005), vol. 44.
33. G. N. Goparaju, B. A. Fry, P. L. Dutton, B. M. Discher, Design and Characterization of a Single Chain Amphiphilic Maquette for Membrane Insertion and Electron Transfer. *Biophysical Journal*. **104**, 661a (2013).
34. B. R. Lichtenstein *et al.*, Engineering Oxidoreductases: Maquette Proteins Designed from Scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).
35. D. Noy, C. C. Moser, P. L. Dutton, Design and Engineering of Photosynthetic Light-Harvesting and Electron Transfer using Length, Time, and Energy Scales. *Biochimica et Biophysica Acta (BBA)*. **1757**, 90–105 (2006).
36. T. A. Farid *et al.*, Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).

Chapter V: Towards a proof-of-principle photochemical triad charge-separating device in maquette

5.1 Introduction:

As demonstrated in the previous chapter, the maquette approach represents the most appropriate method for engineering a versatile and stable platform upon which practical application and examination of the optimal engineering guidelines formulated in Part I of this thesis can be carried out, thanks to its ease with which the locations and environments of the cofactors can be easily and efficiently customized and the conditions of the experiments involved in the construction of the multi-ads can be easily controlled. However, in spite of the extensive functionalities already achieved in the past using the maquette approach both within and without the Dutton group,(*1-5*) construction of photochemical multi-ads, especially when n≥3, within maquettes as charge-separation devices presents a few significant engineering challenges that have never been encountered in the past applications of the maquette approach. As one of the core functionalities that are intended for maquettes, ET-based charge-separation has, however, only seen very limited application in maquettes and without demonstrating much significant successes. Up to now, only three cases of ET-based charge-separation in maquettes have been reported, two of which are shown in Figure 5.1, and all involve only P-A dyads with suboptimal performance, in terms of engineering efficiency and lifetime of the charge-separated state. In fact, the engineering of maquettes up to now had not considered the incorporation of more than two cofactors with different redox and photophysical properties. As seen in the P-A dyad shown in Figure 5.1B, the pigment and the acceptor are located at the two ends of the 4-helical bundle of the maquette, suggesting that significant engineering modifications to existing maquettes as well as changes in the fundamental design philosophy of maquettes are necessary to achieve this breakthrough.

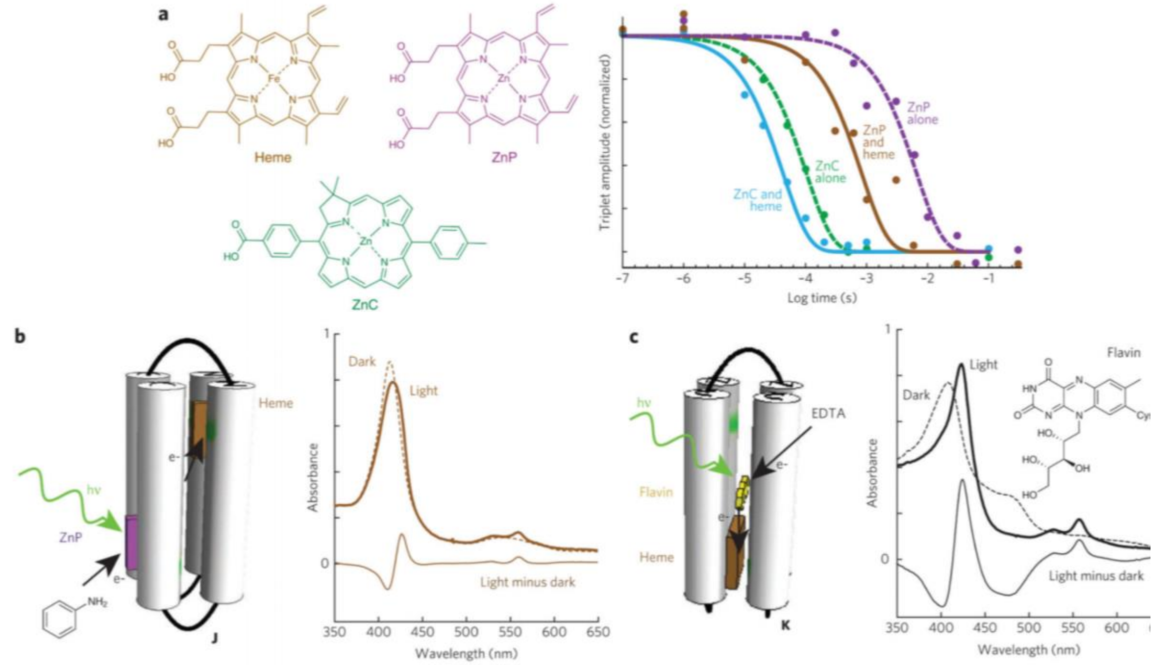


Figure 5.1: Electron-transfer based charge-separating dyads in previous generations of short maquettes produced by the Dutton group. A: decay of the bleach signal of the soret signal repreenting the ground state of the pigments in ZnP/ZnC-Heme dyads. B: reduction of heme bound in 8 μm maquette following photoexcitation of ZnP. dashed and bold traces are absorption spectra of the sample before and after 20 min of white light illumination, respectively. experiments were performed in 20 mm cHeS, 150 mm Kcl, pH 9.0, and 0.1 mm aniline. C: reduction of bound heme following photoexcitation of 8-bromoriboflavin covalently attached to 8 μm maquette. dashed and bold traces show absorption spectra before and after 60 min of white light illumination, respectively. experiments were performed in 20 mm cHeS, 150 mm Kcl, pH 9.0, and 1 mm edTA. experiments in b and c were done in duplicate. Figure and legends adapted from (3)

The prospect of engineering photochemical multi-ads in maquettes with n>2 presents the following list of previously unseen experimental challenges in the maquette design environment:

1) stably and readily anchoring at least 3 unique cofactors with differing physical and chemical properties within the framework of a single four-helical bundles.

2) guaranteeing the site specificity of all cofactors specified by the design blueprints of the multi-ad via varying the cofactor ligation/incorporation method.

3) providing the necessary inter-cofactor distances as prescribed by the design blueprints.

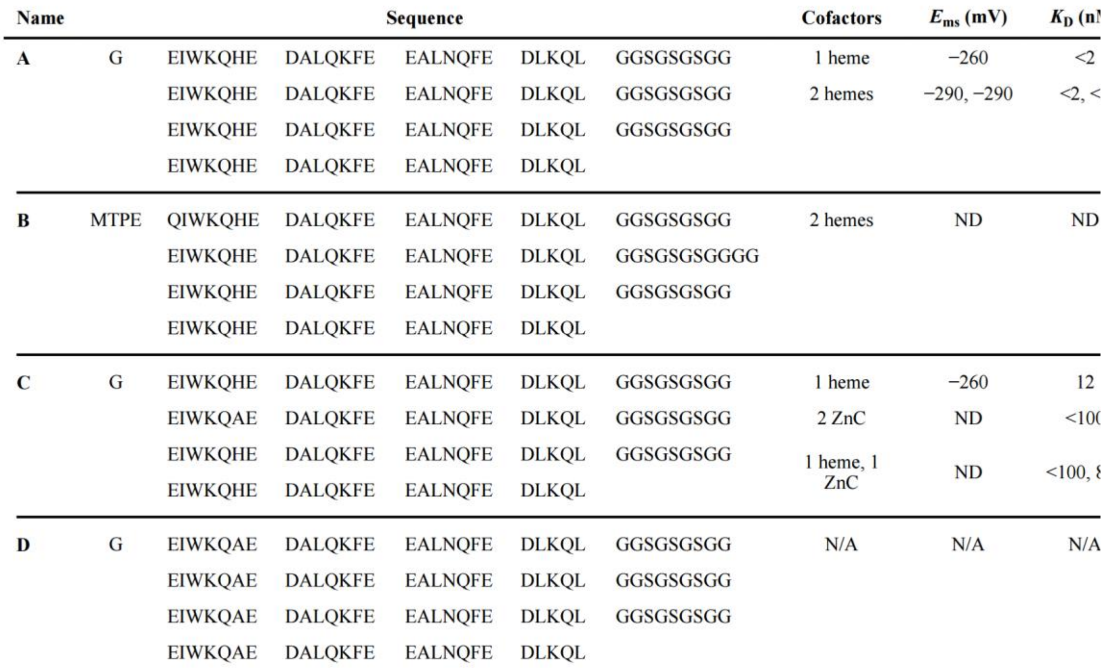
Moreover, in addition to the challenges faced by the engineering of maquttes themselves, even greater constraints are imposed upon by the highly limited selection of cofactors to follow the optimal midpoint potentials prescribed by the optimal engineering blueprints. The difficult to design cofactors with customizable redox potentials covering the range prescribed by the optimal engineering blueprint will be further examined in the next two chapters and indeed remains as the greatest challenge in accurately following the optimal engineering guidelines.

In order to address the novelty and difficulty presented by the above-listed engineering challenges, a multitude of biophysical and biochemical methods completely new to the context of maquette design have been tested in combinations in the many attempts at constructing a stable and functional charge-separating triad, yielding results of varying effectiveness. While most of the attempted methods have not survived to the final, successful design that produced record-breaking charge-separation lifetime and yield that will be the centerpiece of Chapter VI, they nevertheless deserve their places in this thesis since, in spite of not being part of the final solution, they have served to greatly expand the capabilities and potentials of the maquette approach. They represent incremental knowledge in maquette designs that will inevitably prove useful in future designs. Here in the remainder of this chapter I aim to describe each of the novel engineering methodologies applied to the maquette approach, the extent to which they contributed to the successful engineering of the charge-separating multi-ads in maquettes, and their corresponding significance to the future of maquette design.

5.2 Methods:

Standard biophysical and biochemical methods are summarized at the end of the thesis in Appendix A1, but methods special and specific to this chapter are presented here to emphasize their relevance and novelty in the context of maquette approach.

5.2.1: In-vivo incorporation of C-type heme in maquettes expressed in E.Coli



Name		Sequence					Cofactors	E_{ms} (mV)	K_D (n
A	G	EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	1 heme	−260	<2
		EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	2 hemes	−290, −290	<2, <
		EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG			
		EIWKQHE	DALQKFE	EALNQFE	DLKQL				
B	MTPE	QIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	2 hemes	ND	ND
		EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGGGG			
		EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG			
		EIWKQHE	DALQKFE	EALNQFE	DLKQL				
C	G	EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	1 heme	−260	12
		EIWKQAE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	2 ZnC	ND	<10
		EIWKQHE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	1 heme, 1 ZnC	ND	<100,
		EIWKQHE	DALQKFE	EALNQFE	DLKQL				
D	G	EIWKQAE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG	N/A	N/A	N/A
		EIWKQAE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG			
		EIWKQAE	DALQKFE	EALNQFE	DLKQL	GGSGSGSGG			
		EIWKQAE	DALQKFE	EALNQFE	DLKQL				

Table 5.1: Some representative sequences of Gen1 maquettes. The cofactors ligated and their respective redox midpoint potentials and dissociation constants are shown. Table reproduced and adapted from (2)

In addition to the standard protein expression protocols, an additional plasmid encoding the CCM enzymes are co-transformed into BL-21 strains and selected with both antibiotics resistances by the plasmid containing the maquette DNA and the plasmid containing the C-type maturation enzymes. Cell cultures(2L) are inoculated by 5ml of overnight pre-culture and induced when the cell density, represented by OD of the culture, reaches 0.6 after incubation at 37C^o. Induction is achieved with IPTG, but to encourage heme production, 100uM of $FeCl_2$ and levulinic acid are added to the culture. The culture is kept at 37C^o for 30mins and then transferred to 20C^o for overnight incubation.

5.2.2: in-vivo incorporation of propargyl tyrosine in maquettes.

Standard mutagenesis was used to introduce a TAG Amber codon into the plasmid encoding the Gen2 maquette. The plasmid was then transformed into a pre-made strain of E. coli BL21 (DE3) competent cell already containing the pDule2-AzfRS plasmid containing an M.jannaschii mutant tRNA synthetase and tRNA CUA pair. Single colonies were selected and grown in 4 mL of LB using Amp (100 µg/mL) and streptomycin (Strep, 100 µg/mL) for overnight hours at 37 °C with shaking at 250 rpm. The 4 mL primary culture was inoculated into 1 L of minimal media containing 6 g of Na 2 HPO 4 , 3 g of KH 2 PO 4 , 0.5 g of NaCl, and 1 g of NH 4 Cl: 1 mL of 2 M MgSO 4,1 mL of 15 mg/mL FeCl 2 (in 1.0 M HCl), 1 mL of 15mg/mL ZnCl 2 (in acidified water), 2 mL of 10% Bacto yeast extract (w/v), 25 mL of 40% glucose (w/v) supplemented with Amp and Strep and grown at 37 °C with shaking at 250 rpm until OD 600 1.0. Protein expression was induced with 1.0 mM p-propargyloxyphenylalanine (pPf) and 1.0 mM IPTG. Cells were incubated at 25 °C with 250 rpm shaking for Overnight. Cell lysis and purification were performed using the standard methods described in the Appendix of this thesis.

5.2.3: Click-chemistry between the azido-containing label and the PPY-containing maquette.

Catalysts for the click reaction was first prepared by mixing 100uM of Cu(1.25µL of 80mM $CuSO_4$) and 500µM THPTA ligand(10µL of 50mM) and spinning down to remove any precipitations. 1mM ascorbate was added to the catalyst mixture to reduce the mixture for 5 mins. 3-5 equivalents of the maquette is then added to the reduced catalyst mixture dropwise. The mixture is then allowed to react for 1-3 hours at room temperature. The mixture is then examined by MALDI to ensure the completion of the reaction, and purification is carried out using standard HPLC method shown in the Appendix.

5.3 Results:

5.3.1 Engineering of an extended maquette to accommodate three redox/ET active cofactors

Prior to the attempt to engineer a charge-separating multi-ad in maquette, incorporation of more than two unique cofactors into maquettes had not been considered, and consequently, there had not been any existing maquette design that even provides the appropriate dimension to accommodate three tetrapyrroles in terms of spatial requirements. As seen in Table 5.1 as well as figure 5.2, the past generations of maquettes designed to ligate two cofactors all have individual helices that are 27 amino acids residues long.(*3*) Since the cofactor binding site must not be situated at the beginning or the end of the sequence and the cofactor themselves, usually tetrapyrroles, have their own dimensions, the actual maximal edge-to-edge distance obtainable in such maquettes is 14 Å. The first novel design effort therefore was to extend existing maquette vertically so that three binding sites for three different cofactors could be engineered. This extension is shown in figure 5.3. According to our design blueprint, an optimal triad with its cofactors arranged linearly must have at least 30 Å of distance between the binding site on one end of the maquette to the other binding site on the other end of the maquette, providing enough space for 3 cofactors as large as tetrapyroles to be present simultaneously inside the maquette with adequate(~5Å) distances between each pair of the cofactors. While the need to have longer helical bundles is apparent, it is non-trivial to determine how to extend the helices, specifically, what amino acids to add to the sequence so that while the helices are elongated, minimal impact on the biophysical properties of the maquette is expected. The rather obvious choice is to extend the maquette helices by adding yet another heptad repeat that consists of the middle region of the existing generation of maquettes. This heptad repeat inserted is thus AEEALNQ, with the F replaced by A as the addition of four phenylalanine simultaneously to an already stable maquette would disproportionally increase the thermal stability and consequently the rigidity of the maquette and cause ability to bind cofactors with high affinity to suffer.

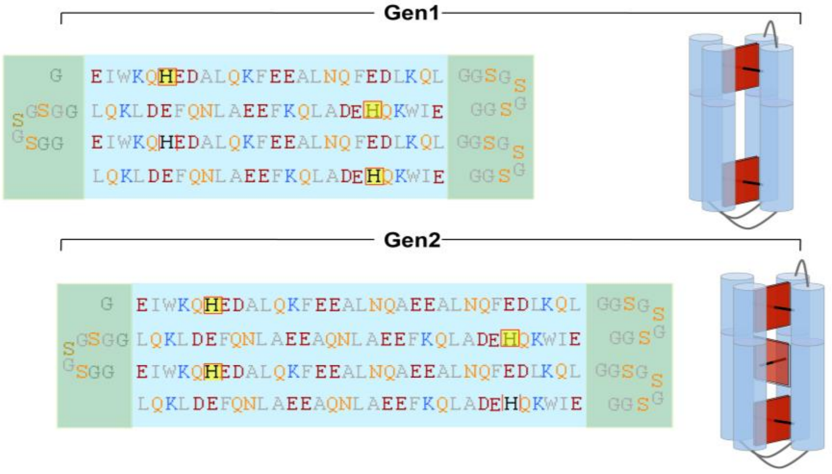

Gen1
G EIWKQHEDALQKFEEALNQFEDLKQL GGSGS
SGSGG LQKLDEFQNLAEEFKQLADEHQKWIE GGSG
GSGG EIWKQHEDALQKFEEALNQFEDLKQL GGSGS
LQKLDEFQNLAEEFKQLADEHQKWIE GGSG
Gen2
G EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGS
SGSGG LQKLDEFQNLAEEAQNLAEEFKQLADEHQKWIE GGSG
GSGG EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGS
LQKLDEFQNLAEEAQNLAEEFKQLADEHQKWIE GGSG

Figure 5.2: example sequences of Gen1 single-chain maquettes with binding sites for hemes. . Regions shaded in green represent the loop region of the maquette, while shading in light blue represents the helical region. Top: Gen1 maquette sequences, with 27 residues per helix. The placement of the cofactors are shown in the cartoon on the right. Bottom: Gen2 maquette with the helical region extended by a heptad repeat.

The extended maquette's sequence in shown in Figure 5.3, demonstrates heme affinity similar to that of its predecessor, as shown in Figure 5.4 With heme affinity on the order of 10nM, the extended maquette satisfies the initial requirement of serving as the design platform of the multi-ad photosystem envisioned by the engineering blueprints from Chapter II and III.

5.3.2 Engineering of a single-his binding site for anchoring of Zn-porphyrin and other Zn-based pigment cofactors

The extended maquette, as shown in Figure 5.2, has two bis-his ligation sites for B-type(non-covalent) heme binding, inherited from its predecessor generation of maquettes. In order to allow the maquette to serve as the design platform for photochemical multi-ads, at least one additional cofactor binding site with site specificity needs to be created. Unlike bis-his binding sites, which is capable of high affinity ligation of iron porphyrins, single-his site provides site-specific ligation of Zn-based tetrapyroles such as chlorins and porphyrins with similarly high affinities(*6*). Single-his site is utilized in the engineering of the photochemical dyad in the Gen1 maquettes shown in Figure 5.1. Consequently, the logical implementation of the necessary Zn-porphyrin binding site for the Gen2 maquette is also a single-his site, however, in the middle of the helical bundle rather at the end since we aim to engineer a donor-pigment-acceptor triad, which has been shown in Chapter III to be equivalent to the model PAA triad used throughout both Chapter II and III in terms of optimality under the same set of constraints. That is, the DPA triad and the PAA triad would be able to separate charges with the same efficiency at the same lifetimes as long as the engineering parameters prescribed by the optimal blueprints are followed, although the engineering parameters themselves could be significantly different. Figure 5.3 demonstrates the modification of the Gen2 maquette to include a single-his binding site in the middle of the helical bundle, while Figure 5.4 demonstrates this maquette's ability to ligate both heme and Zn-porphyrin cofactors with high affinities. Special thanks are given here to Tatianna Esipova from the Vinogradov group for her generous assistance in the preparation of the specific ZnP used throughout this thesis.

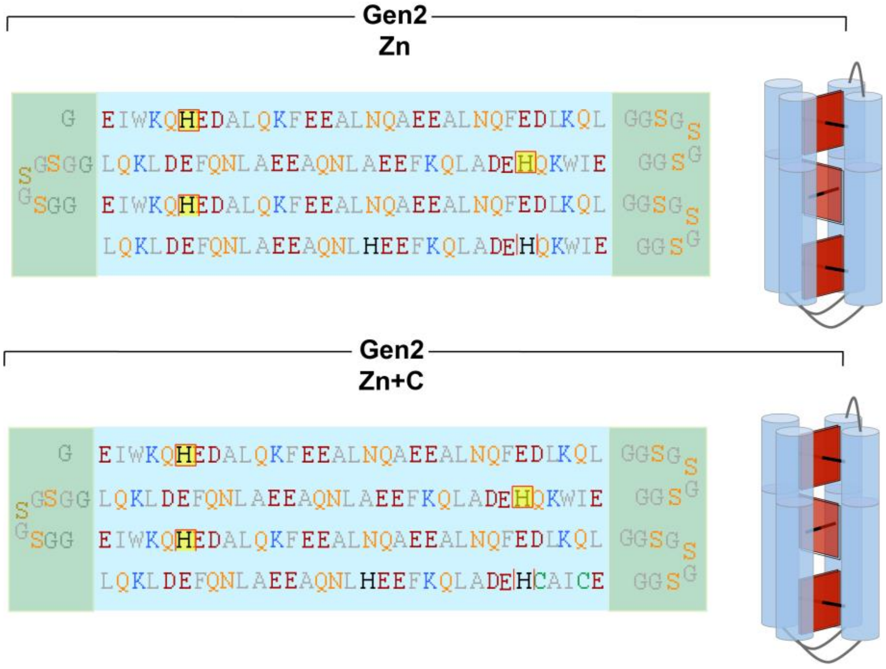

Gen2
Zn
G EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSG S
S GSGG LQKLDEFQNLAEEAQNLAEEFKQLADEHQKWIE GGS G
G SGG EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSG S
LQKLDEFQNLAEEAQNLHEEFKQLADEHQKWIE GGS G
Gen2
Zn+C
G EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSG S
S GSGG LQKLDEFQNLAEEAQNLAEEFKQLADEHQKWIE GGS G
G SGG EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSG S
LQKLDEFQNLAEEAQNLHEEFKQLADEHCAICE GGS G

Figure 5.3: Further functionalization of Gen2 maquette to include Zn-porphyrin binding site and C-heme incorporation capability. Same as in Figure 5.2, but this time the Gen2 maquette now includes a single-his binding site in the middle of helix 4(top) and the CXXCH motif in helix 4(bottom).

We have attempted to determine the redox potential of the ZnP/ZnP$^+$ transition in the Gen2 maquette using standard spectral electrochemical method summarized in the Appendix. However, due to the high potential of the species and the lack of appropriate redox mediators to facilitate the electron transfer, a rather significant hysteresis was observed, as shown in Figure 5.5. This resulted in a rather rough estimate of the ZnP potential in the Gen2 maquette. In Chapter VI I will demonstrate a significantly improved set of measurements of the redox potentials of the same ZnP in the new generation of maquettes.

5.3.3 Engineering of CXXCH motif for in-vivo attachment of C-type heme for cofactor site-specificity

With the extended maquette now capable of accommodating 3 tetrapyrroles, although two of which are the same non-covalently ligated hemes, the next step of design is to impart the binding sites of the cofactors with site specificity so that cofactors intended to serve the roles of D, P, and A can be incorporated at the correct locations. At the time of the design of the Gen2 maquettes, iron tetrapyrroles of different midpoint potentials were the main candidates of interest for all the electron transfer roles. Past efforts in the Dutton group have demonstrated that the presence of the CXXCH motif(*7*) on a maquette allows in-vivo covalent attachment of C-type heme to maquettes. This is accomplished by co-expressing the maquette at appropriate condition with a cascade of chaperone proteins that catalyze the formation of covalent linkages between the cysteine residues and the vinyl groups of the B-heme in reduced environments of the extracellular matrix of the *E.coli*. It was especially advantageous in solving the problem of site-specificity for three different tetrapyrroles that have very similar physical properties to have the protein expressed already with one of the three components covalently attached, since iron-tetrapyroles would ligate to any available bis-his sites that are not occupied, unless heavy engineering of the local environments of the binding site cause the binding site for the intended acceptor to strongly favor the acceptor. This type of engineering is time-consuming, requiring multiple cycles of iterative design and testing, rather against the underlying philosophy of the maquette approach. Consequently, in-vivo c-heme maturation became the logical next step towards the construction of the complete DPA triad with three functionally and biophysically different cofactors. While C-type maturation had indeed been demonstrated to be viable in de-novo proteins, the maquettes utilized for that demonstration was much smaller in size and did not include multiple binding sites for hemes. Consequently, in order to maximize the chance of success with the Gen2 maquette, we implemented additional experimental techniques to increase the indigenous production of hemes in bacterial.(*8-10*)

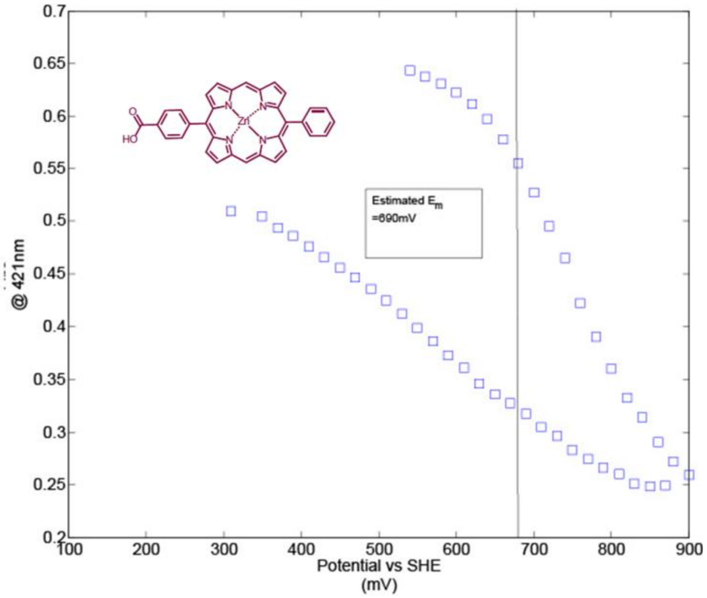


Figure 5.5: Spectral electrochemical measurement of the redox potential of the ground state of the ZnP. vertical axis represents the absorbance of the soret peak under increasing/decreasing potentials shown on the horizontal axis. 4uM of protein with sub-stoichiometric amount of ZnP bound, in standard PBS buffer used throughout this thesis. No mediator was used.

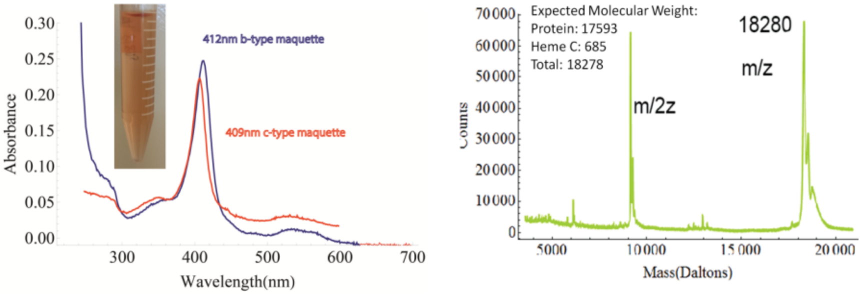


Figure 5.6: In-vivo incorporation of C-heme in the long BT6 maquette. Left: UV-Vis spectrum of the purified, C-heme containing maquette(red) plotted along the same maquette with non-covalently attached B-heme(blue).

Figure 5.6 demonstrates the engineering, expression, and yield of the in-vivo c-heme maturation on the extended Gen2 maquette. C-type maquette demonstrates a unique spectral signature, with the soret peak slightly red-shifted by 3nm in comparison to the non-covalently ligated B-heme, allowing for easy identification of successful c-heme incorporation. However, over the course of the purification of the C-type maquettes, I realized that, at least in the C-type maquette discussed in this chapter, the covalent linkage of the heme to the protein via the two cysteine residues is highly sensitive to reduction in the presence of oxygen. The spectral signature of C-heme disappears upon reduction of the C-heme by strong reducing agents such as β-mercaptolethanol and sodium dithionite. This leads to unexpected difficulty in preparation of C-type containing maquettes to concentrations greater than 1mM, resulting in the difficulty described in section 5.3.7 when a few of the techniques described in this chapter are needed simultaneously to generate the desired triad with proper cofactor site-specificities.

5.3.4 Non-covalently ligated and electrochemically poised DADPIX as the electron donor in the triad

After securing the site-specificity of one of the two bis-his binding sites via in-vivo maturation of C-type heme, the remaining bis-his site automatically gained site-specificity since iron-tetrapyroles do not ligate readily to single-his sites. The next step towards the complete construction of the DPA triad in maquettes is to decide whether the C-heme should act as acceptor or donor, and which iron-tetrapyrole to incorporate for the

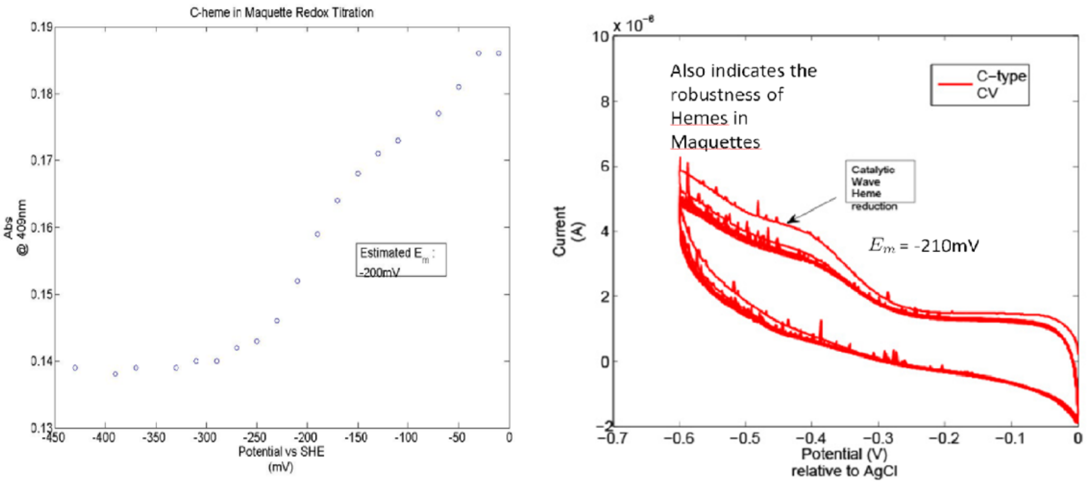
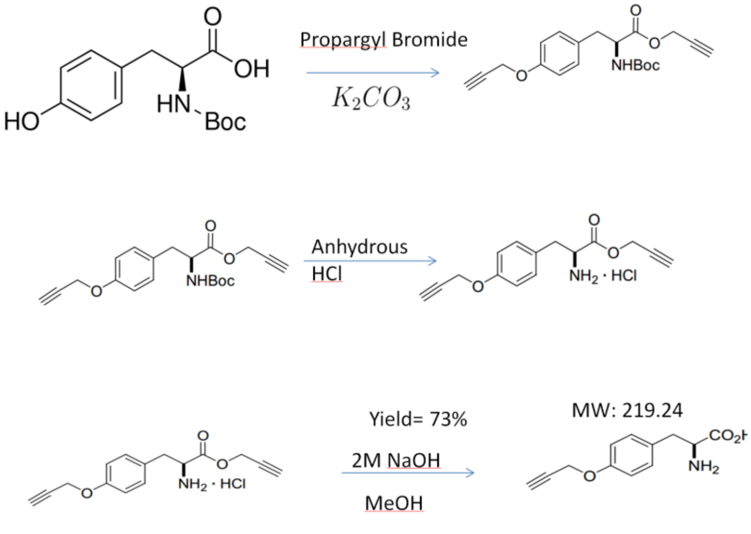


Figure 5.7: Redox midpoint potential of C-type heme in Gen2 maquette determined by two independent methods. Left: spectral electrical chemical determination of the C-heme potential in 4uM of maquette in 200mM NaCl, 20mM sodium phosphate, 2mM Tris-HCl at pH 7.4.Vertical axis represents the absorbance of the soret peak of the C-heme at 409nm. Horizontal axis shows the potential of the environment in which the protein is placed. **Right:** cyclic voltammetry of 20mM of the Gen2 maquette containing c-heme in standard PBS buffer using the method described in the Appendix.

remaining bis-his site.

To determine the proper role of the C-heme, it is necessary first to determine its redox midpoint potential within the maquette. Using standard electrospectral chemical techniques summarized in the Appendix, the redox potential of C-heme in the Gen2 maquette was determined to be -200mV, as shown in Figure 5.7. Since at this point, the choices of the final cofactor for the triad is limited to iron-tetrapyroles, the range of possible potential is rather small, and the most appropriate role for the C-heme would be the acceptor A. In order to produce a charge-separated state with reasonable amount of energy preserved and a proper driving force for the donor to give off its electron to the P^+ radical, the candidate for the donor cofactor must either be already reduced in the normal redox environment of a aqueous solution with decent amount of salt or can be poised by an external factor, either a chemical or electrodes, to remain in the reduced state, without affecting the redox states of the other cofactors. With the redox environment of aqueous solution having a potential around +250-350mV, the first option of having tetrapyroles already reduced is not viable, and therefore DADPIX became the ideal choice as the donor cofactor, since it can be selectively poised to remain in the reduced state with a weak reductant such as ascorbate. With a redox potential of ~50mV, ascorbate can reduce DADPIX only without affecting the C-heme or the pigment of Zn-porphyrin. This redox poising by ascorbate is shown in Figure 5.8, where the first DPA triad in a maquette is successfully assembled.



Scheme 5.1: Synthesis scheme of propargyl tyrosine. Adapted from the supplementary information of (11)

5.3.5 Incoporation of unnature amino acids as site-specific covalent anchoring point for additional cofactors

The combination of the inability to keep all of the DADPIX in its reduced state and the lack of reliable techniques to isolate the ascorbate from participating in the electron transfer reaction resulted in transient absorption experiment results that are difficult to interpret. I will not include the data here in this thesis since it does not offer beneficial information to future designs and engineering of functional maquettes. Nevertheless, this demonstrated that the choice of 3 tetrapyroles for the DPA triad is not viable, as the range of redox potential is too small and the biophysical properties of the cofactors are too similar.

Unnatural amino acid has become an attractive method of introducing chemical modifications to proteins in recent years.(*11*, *12*) I would like to express my sincere appreciation for the expertise of Professor Petersson, whose research group focuses on the various applications of unnatural amino acids to expand the horizons of protein functions. With his advice as well as generous assistances from John Warner and Rebecca Wissner, students mentored by Professor Petersson, I was able to apply

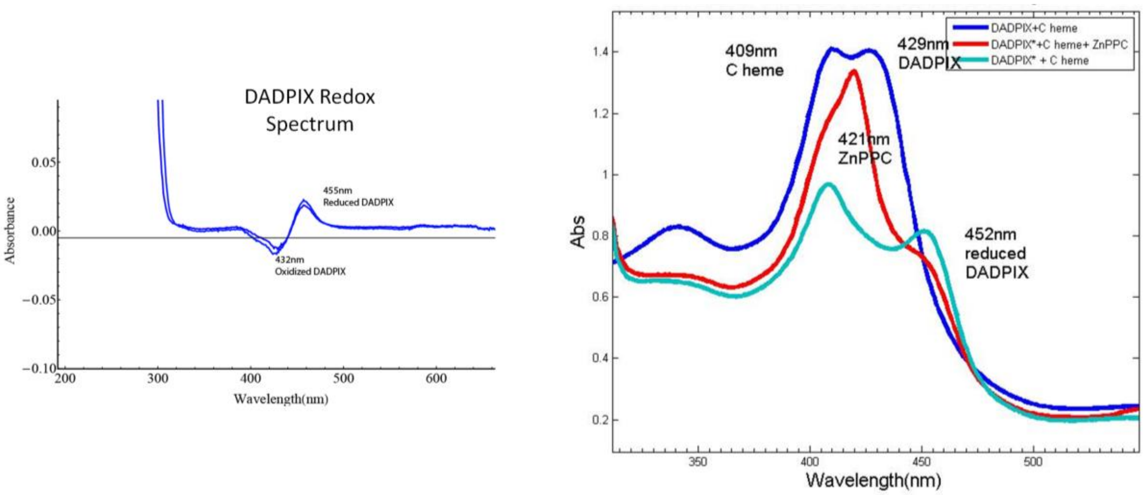


Figure 5.8: Full assembly of the first DPA triad in a maquette using DADPIX-ZnP-C-heme as the cofactors, with ascorbate poised DADPIX as the donor. Left: Redox spectrum of DADPIX bound to 5uM of maquette by in presence and absence of 10mM ascorbate. Right: Assembly of the DPA triad at various stages during the preparation, in 5uM of maquette in standard phosphate buffer(200mM NaCl, 20mM sodium phosphate, 2mM Tris-HCl @pH7.4).

unnatural amino to maquettes to expand the selection of potential cofactors for the construction of a functional DPA triad. Specifically, the Gen2 mauqette was expressed in a strain of E.coli that co-expresses a modified tRNA synthetase and tRNA that recognizes the amber(TAG) codon not as a stop codon but rather a coding codon. The modified tRNA synthetase charges the tRNA that interacts with the amber codon with the unnatural amino acid of choice instead. This leads to the incorporation of the unnatural amino acid into the maquette. However, the efficiency of amber-codon suppression is far from 100%, and more frequently than not, TAG is interpreted as the stop-codon by the bacteria, leading to translation of a fragment of the intended maquette and therefore decrease in the yield of the maquette with the unnatural amino acids. The specific strain provided generously by the Petersson lab and used with the Gen2 maquette contains the tRNA and its synthetase for propargyl tyrosine, shown in Figure 5.9. It is a alkyne derivative of normal tyrosine, having the ability to undergo click-chemistry with molecules containing azide moiety(*13*)The chemical preparation of the unnatural amino acid from normal tyrosine is summarized in Scheme 5.1 and Figure 5.9, and the subsequent successful incorporation of propargyl tyrosine into the Gen2 maquette is shown in Figure 5.10.

Although this experiment successfully demonstrated the applicability of unnatural amino

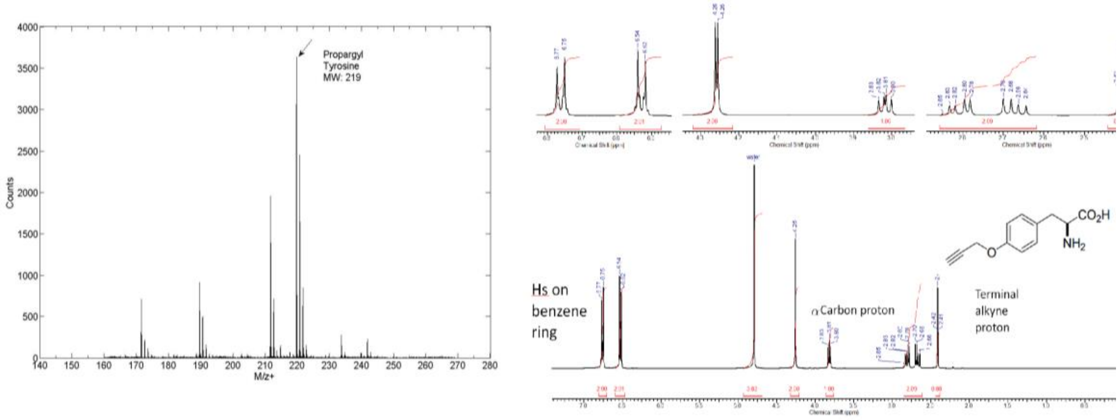


Figure 5.9: chemical characterization of propargyl tyrosine using Left: MALDI-TOF with 1mM sinapinic acid as the matrix, and Right: 1D proton NMR, with the structure of the molecule shown as insert, and the various chemical shift peaks labeled to reflect the hydrogens they correspond to.

acid in de-novo engineered proteins that have no structural or sequence resemblance to existing natural proteins, the yield of the protein is far less than ideal as will be discussed in section 5.4.1.

5.3.6 Click chemistry based covalent incorporation of cofactors into maquettes

The successful expression of maquettes containing unnatural amino acid opens up yet another route of novel cofactor incorporation into the maquettes. As already mentioned above, the unnatural amino acid successfully incorporated in the maquettes, propargyl tyrosine, is click-chemistry ready. The presence of the terminal alkyne allows reactions with molecules containing the azido-group to take place in mild conditions that are ideal for small proteins such as maquettes. The details of the click-chemistry reaction are summarized in the methods section, while Figure 5.11 demonstrates the successful modification of the propargyl-tyrosine containing maquette with a azido derivative of ferrocenes. The choice of ferrocene as the donor is explained in further detail in Chapter VI, as this chapter is mainly devoted to showcasing the various novel techniques that have been individually successfully applied to the maquette approach in order to expand the possible methodologies of customizing maquettes for the intended function. However, because incomplete reaction between the protein and the modifier is possible, another round of purification by HPLC is required. This leads to a further decrease in the yield of the final desired product.

5.4 Discussion

5.4.1: Challenges of bringing everything together: consistent failures in simultaneous application of multiple methods presented in this chapter.

Up to this point, all the engineering methodologies described in this chapter have proved to be successful to various extents in their ability to expand and improve site-specific incorporation of cofactors with distinct intended roles individually. However, none of the methods presented here alone brings site-specific incorporation of the D-P-A method to completion. Instead, a combination of methods is necessary. However, when multiple methods described in this chapter were applied to maquettes simultaneously in an attempt, as is necessary to ensure the necessary site-specificies for the construction of the complete triad, we were surprised by the lack of synergies between almost all of these methods introduced. That is, simultaneous applications of more than one of the above presented methods result in significant reduction in the final yield of the desired maquette with cofactors incorporated. As already shown above, the combination of in-vivo maturated C-heme and the chemically poised DADPIX resulted in suboptimal experimental condition that prevented proper detection of the formation of the charge-separated state. On the other hand, in order to incorporate both the propargyl tyrosine and the C-heme into the maquette, as both methods occur in-vivo during the expression of the protein, the maquette must undergo a series of mechanisms required for both maturation schemes before it is completely expressed. It is highly probable that a strong incompatibility exists between these two in-vivo incorporations, as the yield of maquette that have successfully incorporated both C-heme and propargyl tyrosine is negligible. A closer examination of the mechanisms of these two incorporations reveals that the C-heme maturation takes place after an exportation of the nascent maquette to the periplasmic space(*10, 14*)with a significantly more reducing environment than the cytoplasm, due to the presence of a small export tag sequence near the N-terminus of the expressed protein, while the incorporation of the propargyl tyrosine takes place during translation, when the tRNA is successfully charged with the unnatural amino acid and recognizes TAG. However, for the cases when the translation machinery did not succeed at amber codon suppression and the translation is terminated, the product still contains the export tag sequence and promptly exported to the periplasmic space. This causes the periplasmic space to be filled with a large amount of peptides that do not contain the proper heme binding site, causing the CCM chaperones to have increased difficulty in locating their substrates.

However, it is indeed worth pointing out that rather than applying two or more of the methods demonstrated in this chapter simultaneously, single application of the click-chemistry upon successful incorporation of the propargyl tyrosine would impart the necessary site-specificity for a DPA triad if the acceptor role is played by either a DADPIX or a B-heme. This experiment was indeed planned but was transformed and carried out in the more promising form of engineering a DPA triad using the same strategy in the new generation of maquettes with a high resolution crystal structure, as described in detail in the upcoming Chapter VI.

5.4.2: Sky's the limit: maquette approach's first-principled, functionally neutral nature makes it susceptible to any type of functionalization.

While none of the novel modifications of maquettes described in this chapter was used in the engineering of the final DPA triad showcased in Chapter VI, thanks to the valiant effort by Nathan Ennist in obtaining the crystal structure of a maquette, these modifications were all successful as far as their immediate purposes were concerned. Indeed, throughout this chapter, it has been repeatedly demonstrated that customization of maquettes by biophysical and biochemical techniques that could be problematic to natural proteins, tend to produce successful, positive results. This reflects one of the central characteristics of maquettes and underscores a major advantage of the maquette approach as a method of protein engineering: the functionally neutral nature of maquettes as medium or a platform. As demonstrated in Chapter IV, the single set of first-principled

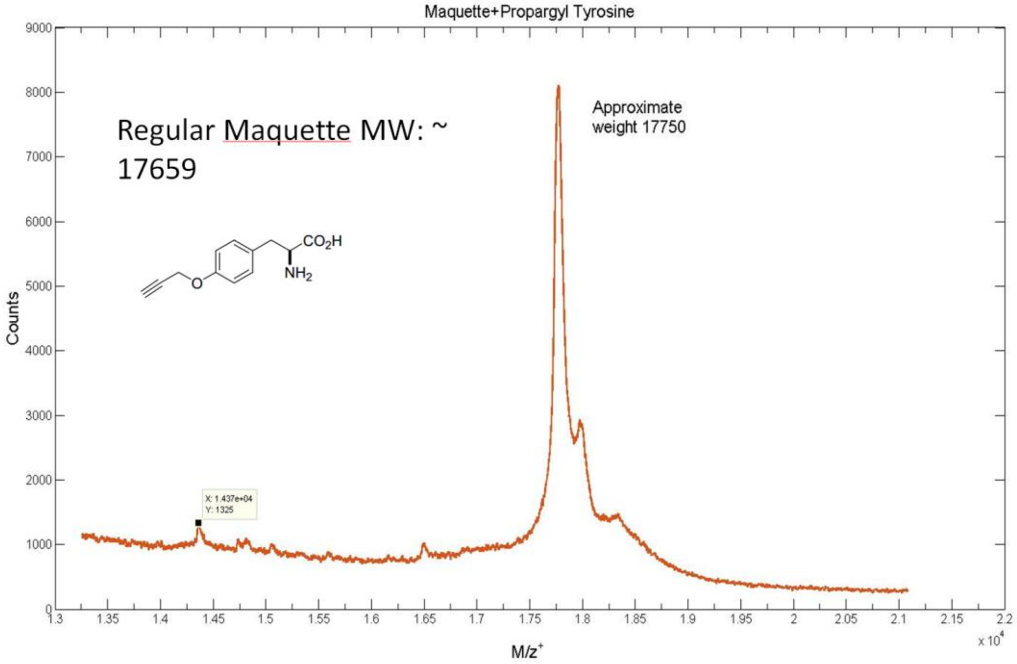


Figure 5.10: Incorporation of propargyl tyrosine into maquettes. Mass spectrum of the purified maquette using the procedure described in the methods section. The mass of the mauqette without the propargyl tyrosine is shown for reference.

based constraints invariant to all maquette designs is the combination of the binary patterning and the heptad repeat of the residues. This set of fundamental rules leaves the function side of maquettes to be essentially tabula rasa, further constrained to a various but still not a great extent by the additional engineering of the specific sequences. The maquettes showcased in this chapter, Gen1 and Gen2 specifically, are excellent examples of minimally engineered sequences that are largely functionally unconstrained. Other than the binding sites for the cofactors, the maquette sequences demonstrated so far are only engineered to ensure that the four-helical bundle motif is realized, without overwhelming rigidity in the hydrophobic core to preclude cofactor binding. This functionally neutral chacteristic, to the best of my understanding, allows the maquettes to be highly compatible with engineering techniques, underscoring one of maquette approach's unique strengths among the various protein design methods.

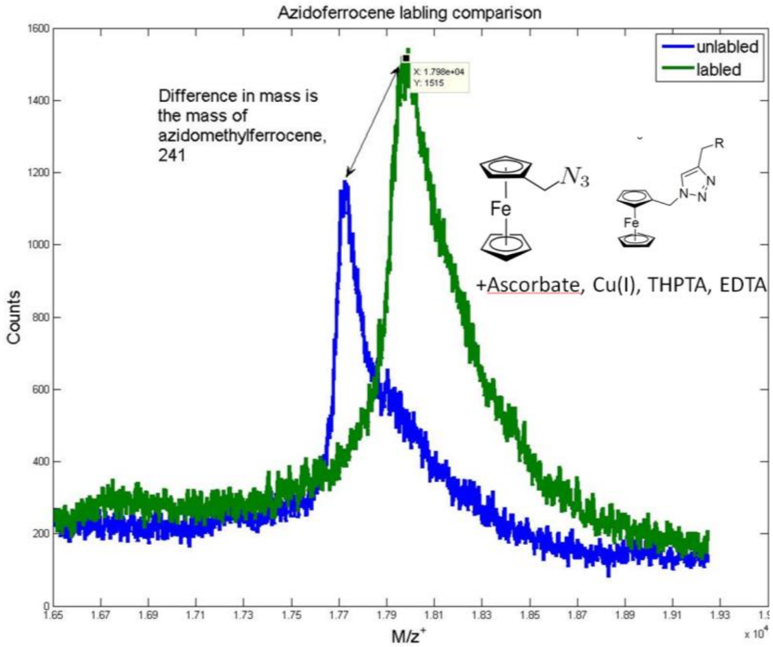


Figure 5.11: Incorporation of azido-ferrocene in maquettes via click-chemistry. Mass spectra of the product of the click-chemistry reaction(green) in comparison against the reactant(blue), with the difference being the mass of the azido-ferrocene. The insert illustrates the mechanism of the reaction.

5.5 Conclusion

Chpater V of my thesis is a special place where a large collection of experimental results that would otherwise receive no attention, in terms of being showcased in presentations in front of crowds or included as figures or even supplementary figures in manuscripts, get their due recognition. Although they did not contribute to the record-breaking charge-separating triad that takes on the main role of the next chapter, these results have significantly expanded the horizon of functional engineering of maquettes, demonstrating the great potential of maquette approach has in the application of functional derivatization of proteins by biochemical and biophysical methods. These results also exposed the maquette engineering to possibilities of cofactors that have never been associated with protein environments in the past, such as the ferrocene and fullerenes, synthetic molecules with redox midpoint potentials that are ideal as electron transfer partners, either acceptors or donors, for pigments such as Zn-porphyrins and chlorins. Indeed, the exploratory experiments with azido-ferrocenes shown in this chapter led to the adoption of ferrocene as the donor, albeit via a different linkage mechanism, by the triads in Chapter VI.

5.6 References

1. B. R. Gibney, P. L. Dutton, De novo Design and Synthesis of Heme Proteins. *Advances in Inorganic Chemistry* (2000).
2. B. R. Lichtenstein *et al.*, Designing Light-Activated Charge-Separating Proteins with a Naphthoquinone Amino Acid. *Angewandte Chemie*. **127**, 13830–13833 (2015).
3. T. A. Farid *et al.*, Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).
4. B. R. Lichtenstein *et al.*, Engineering Oxidoreductases: Maquette Proteins Designed from Scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).
5. G. N. Goparaju, B. A. Fry, P. L. Dutton, B. M. Discher, Design and Characterization of a Single Chain Amphiphilic Maquette for Membrane Insertion and Electron Transfer. *Biophysical Journal*. **104**, 661a (2013).
6. M. R. Razeghifard, T. Wydrzynski, Binding of Zn−Chlorin to a Synthetic Four-Helix Bundle Peptide through Histidine Ligation. *Biochemistry*. **42**, 1024 (2003).
7. Constructing a Man-Made C-Type Cytochrome Maquette in Vivo: Electron Transfer, Oxygen Transport and Conversion to a Photoactive Light Harvesting Maquette. **5**, 507–514 (2014).
8. H. Schulz, Prototype of a Heme Chaperone Essential for Cytochrome c Maturation. *Science*. **281**, 1197 (1998).
9. E. M. Harvat, C. Redfield, J. M. Stevens, S. J. Ferguson, Probing the Heme-Binding Site of the Cytochrome c Maturation Protein CcmE†. *Biochemistry*. **48**, 1820 (2009).
10. J. W. A. Allen, The Escherichia coli Cytochrome c Maturation (Ccm) System Does Not Detectably Attach Heme to Single Cysteine Variants of an Apocytochrome c. *Journal of Biological Chemistry*. **277**, 33559 (2002).
11. K. Lang, J. W. Chin, Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **114**, 4764 (2014).
12. A. Deiters *et al.*, Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. *J. Am. Chem. Soc.* **125**, 11782 (2003).
13. J. E. Hein, V. V. Fokin, Copper-catalyzed Azide–alkyne Cycloaddition (CuAAC) and Beyond: New Reactivity of Copper(i) Acetylides. *Chemical Society Reviews*. **39**, 1302–1315 (2010).
14. H. Schulz, R. A. Fabianek, E. C. Pellicioli, H. Hennecke, L. Thony-Meyer, Heme Transfer to the Heme Chaperone CcmE During Cytochrome C Maturation Requires the CcmC protein, Which May Function Independently of the ABC-transporter CcmAB. *PNAS*. **96**, 6462 (1999).

Chapter VI: Engineering of water-soluble charge-separating Donor-Pigment-Acceptor(DPA) triads in a highly structured 4-helical scaffold protein to achieve world record for longest lifetime of charge-separation and to test the theoretical design blueprints of general charge-separating devices

6.1 Introduction:

In the previous chapter, I described an array of novel engineering modifications of the original generation of single-chain amphiphilic maquettes(*1*, *2*) aimed to bring about an engineering breakthrough that allows single maquettes to accommodate more than two different cofactors of differing redox and photophysical properties with high affinities, without requiring inter-protein interactions or the presence of small redox molecules in the surrounding environment. This represents an evolutionary step in the development of the maquette approach in protein engineering that enables the construction of a donor-pigment-acceptor triad as the first, proof-of-principle step towards the eventual engineering of the optimal photochemical multi-ads specified by the engineering blueprints. In spite of the individual successes to various extents achieved separately, these engineering modifications have not seen successful simultaneous implementation due to various practical limitations of the Gen2 maquettes used, as described at end of the previous chapter. Consequently, the challenge to ensure site specificity while maintaining reasonable yield of the final product, namely the intended charge-separating triad design specified by the engineering blueprints, had not been satisfied. Moreover, the lack of high resolution structural information on that generation of maquette from either an NMR or X-ray 3D has also limited the confidence with which the necessary reengineering of the existing maquettes were carried out. Indeed, no DPA triad that demonstrates formation of small, long-lasting charge-separation with lifetimes greater than 100ms based on this generation of maquettes had been successfully engineered. Fortunately, thanks to the efforts of Nathan Ennist of the Dutton group(manuscript in preparation), a new generation of maquettes that offer both detailed structural information and high site-specificities for cofactors became available as the ideal platform and testing ground for charge-separating multi-ads, obviating any further need to engineer the Gen2 maquette reported in Chapter V.

In addition to the evolution from dyads to triads and beyond in the form of maquette design, another necessary fundamental change in the paradigm of practical CS engineering necessary as learned from the efforts described in Chapter V is that tetrapyrroles alone are insufficient to produce even proof-of-principle charge-separating triads, as they are only able to provide a limited number of midpoint potentials within a range that's too small to obtain desirable driving forces for the lifetime and yield of charge separation.(*3*) When the pigment cofactor availability is limited to Zn-porphyrin/Zn-chlorin due to practical constraints and the usage of triplet excited state is tolerated by artificially subjecting the photosystem to anaerobic environments, hemes'

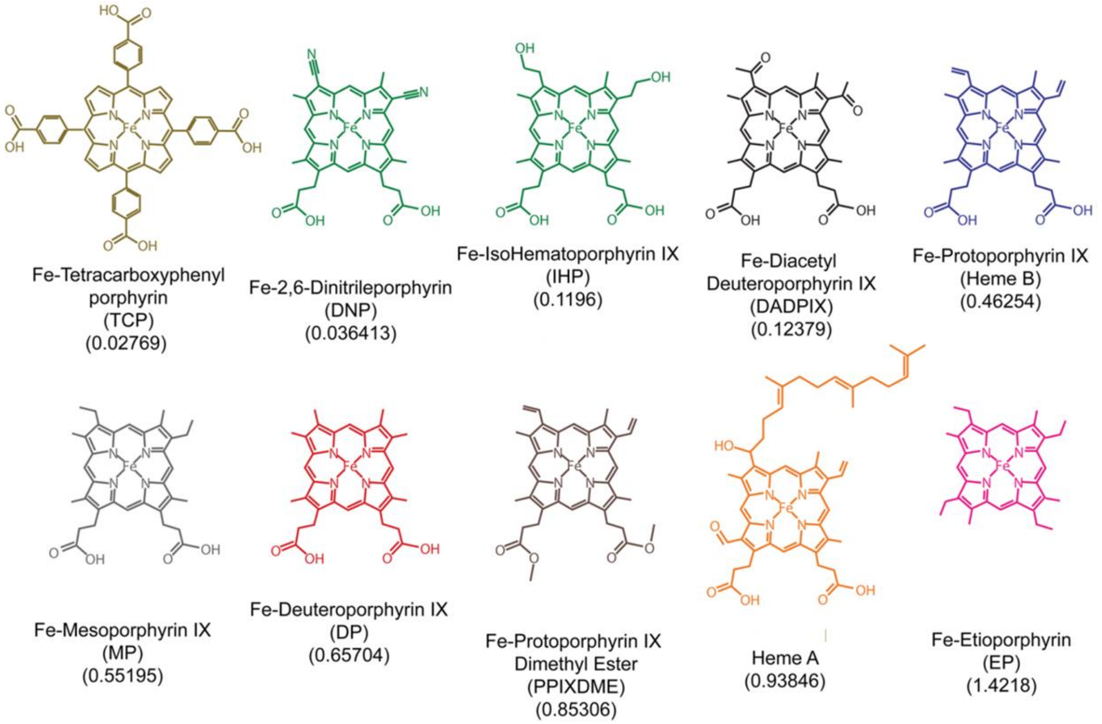
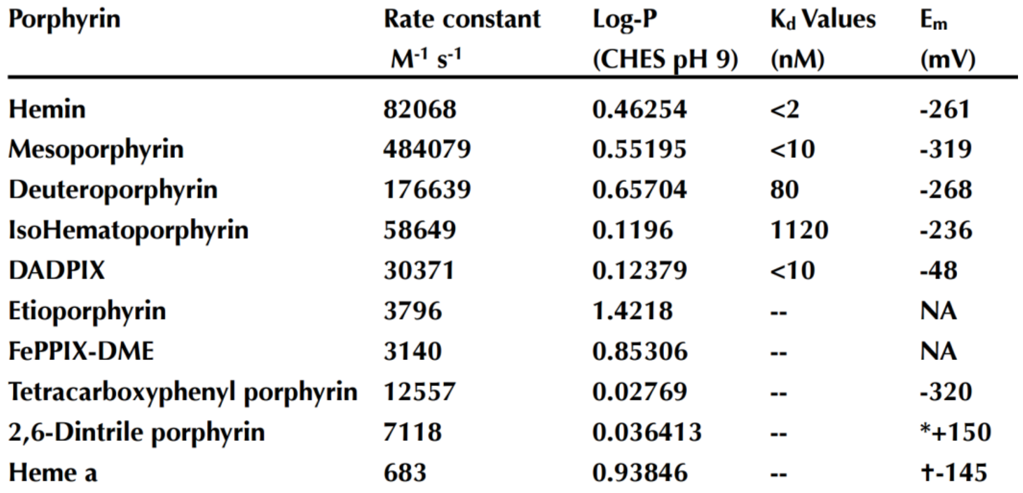

Fe-Tetracarboxyphenyl porphyrin
(TCP)
(0.02769)
Fe-2,6-Dinitrileporphyrin
(DNP)
(0.036413)
Fe-IsoHematoporphyrin IX
(IHP)
(0.1196)
Fe-Diacetyl Deuteroporphyrin IX
(DADPIX)
(0.12379)
Fe-Protoporphyrin IX
(Heme B)
(0.46254)
Fe-Mesoporphyrin IX
(MP)
(0.55195)
Fe-Deuteroporphyrin IX
(DP)
(0.65704)
Fe-Protoporphyrin IX Dimethyl Ester
(PPIXDME)
(0.85306)
Heme A
(0.93846)
Fe-Etioporphyrin
(EP)
(1.4218)

Figure 6.1: Available tetrapyrrole cofactors successfully incorporated into maquettes in past studies. The corresponding midpoint potentials and binding affinities to maquettes are listed in Table 6.1. Figure reproduced from (3) with author's permission. The number shown below each cofactor refers the cofactor's partition coefficient between octanol and hexane.

range of midpoint potential makes them the adequate candidates for the acceptor role, especially when more appropriate acceptors with much lower midpoint potentials are not available due to a simple lack of available candidates. Consequently, the restriction of cofactors selection to tetrapyrrole forces the choices for donors to be difficult and highly suboptimal. Specifically, the tetrapyrroles that have been successfully incorporated into maquettes in previous studies fall within the range of -220mV(B-heme) and 50mV, as shown in Figure 6.1. In comparison to the expected 600mV or higher potential(shown in chapter V and examined in detail again in this chapter) of the ground state of the pigment molecule, none of the tetrapyrroles can serve as adequate donor, as we have seen in Chapter II that the optimal Em of the donor should be slightly above that of the pigment ground state to produce efficient long-lived charge separation. In order to approach the optimality prescribed by the engineering blueprints in Chapter II and III, it is therefore necessary to expand the types of cofactors beyond tetrapyrroles. Incorporation of many frequently used non-tetrapyrrole cofactors capable of participating in electron transfer reactions, as seen in the previous chapter and as to be demonstrated in this chapter, usually involves covalent modifications of side chains or backbones of the maquette, with a linker that influences the effective inter-cofactor distances Without detailed structural information of a given maquette, however, it is significantly more difficult, more risky, and impractical to design modification sites without existing structural information of the protein platform.

In this chapter we report the successful design and engineering of a series of proof-of-principle donor-pigment-acceptor(DPA) charge-separating devices, built on top of a platform of a new generation of highly ordered and structurally stable maquettes engineered by Nathan Ennist of the Dutton group(manuscript in preparation), whose structural details of up to 1.9 Å resolution had been determined by X-ray crystallography. We demonstrate the successful incorporation and utilization of ferrocene, a highly effective cofactor used previously exclusively in synthetic engineering of charge separation that had never been applied to protein-based electron transfer.(*4*, *5*) We investigate the specific engineering landscape of our proof-of-principle triads given the structural and practical constraints using the engineering guidelines we established in Chapter II and III. We remark the successful and significant generation of the longest living charge-separated state, over 331ms, ever observed and engineered in a triad device, to the best of our knowledge, also the longest lasting charge-separation ever observed in any protein-based designs. Finally, we report the empirical effects of varying distances and driving forces on the lifetime, quantum yield, and engineering efficiency of the charge-separation achieved by our DPA triad and compare the observation to the expected changes described by our theoretical model.



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

Table 6.1: Dissociation constants and Ems, along with other properties, of the tetrapyrroles that have been successfully incorporated into maquettes. The names of the tetrapyrroles correspond to those shown in figure 6.2. Reproduced from (3) with author's permission.

6.2: Methods:

Standard molecular biology, biophysical, and biochemical methods are summarized in the Appendix rather than reported here in order to make the presentation of this already large chapter less overbearing. However, methods unique to this chapter and essential to the assembly of the triad are described here. Moreover, the details of the SVD-based global analysis of the transient absorption data are also reported here, since it constitutes one of the main advances this study actually resulted in, an effective and straightforward standard method used by other members of the laboratory.

6.2.1 covalent modification of maquette via cysteine-maleimide reaction:

N-ferrocenyl maleimide was incorporated into the maquette by mixing the molecule in DMF with 50uM of the maquette in 6M Guanadinium HCl and 5mM TCEP at 10:1 ratio and rotating overnight with protection from light. The unreacted protein was separated from the reacted via HPLC using a gradient of water/AcCN mixture going from 35%AcCN:65% water to 65% AcCN and 35% water. The unlabeled protein eluted at 51%AcCN:49%water while the labeled protein eluted later at 55%AcCN and 45% water. The purity of the labeled protein fraction is examined and verified using MALDI-TOF, assisted by the matrix of saturated solution of synaptic acid at 45% laser intensity.

6.2.2: Fluorescence emission spectroscopy:

Fluorescence emission spectra of the ZnP bound to the maquette were measured using fluoremeter, The sample was excited at 540nm and emission spectra were collected from 550nm up to 1100nm, the detection limit of the instrument. The sample was examined both in aerobic(atmospheric) and anerobic conditions. The anaerobic condition was achieved by aeration by dried argon on the top of the sample with stirring for 2 hours, coupled with addition of the oxygen scavenging system(*6*) using a mixture of 1μM glucose, 10nM glucose oxidase, and 1nM catalase, as described in the reference.

6.2.3: Redox and electrochemical measurements:

The redox potentials of the cofactors in the maquette were measured using electrochemical spectroscopy and square-wave voltammetry. 30-50uM of cofactor bound maquette in 50mM phosphate, 200mM NaCl was mixed with mediators of appropriate potentials for the specific cofactor of interest. For DADPIX and B-heme, a mixture of Anthra-quinone-2-sulfonate (100µM), Benzly Viologen (50µM), Methyl Viologen (50µM),Sulfanilamide (50µM), Indigo trisulfonate (50µM), Phenazine (50µM) was used, while for the Zn porphyrin, a mixture of high potential mediators including K2[Ir], 2,6-p-quinone, was used. The samples were subjected to increasingly oxidizing potentials in the range of +_200mV of the estimated potentials of the cofactors at 20mV increments. UV-vis spectra at each potential was recorded after the sample has stabilized after each potential change. The samples were then subjected to increasingly reducing potentials, again using 20mV decrements.

6.2.4: Transient absorption spectroscopy of the monad, dyad, and triads:

The photodynamics of the monad, dyad, and triad are all probed via nano-second transient absorption spectroscopy. All samples were kept under anaerobic conditions via a combination of aeration with dried argon for 2 hours and the addition of a oxygen scrubbing system consisting of 5nM catalase,10nM glucose deoxygenase, and 1uM of glucose. A Nd:YAG laser excites the sample at 532nm, then a pulsed Xe lamp is used to probe the post-excitation sample at various delay times of interest. The instrument design limits our time resolution to as fast as 50ns and as slow as seconds. The probe Xe lamp light, dispersed by a monochrometer, is captured by a PIMAX CCD camera. The captured image is compared with the same image captured when no laser excitation occurs, thereby generating a difference spectrum.

6.2.5: SVD-based global analysis and kinetics model fitting of the transient absoprtions data:

Singular Value Decomposition(SVD) is applied to the data matrix obtained from preprocessing of the raw transient absorption data. The implementation of SVD in this specific thesis involves the usage of Matlab 2014b. The analysis calls the built-in “svd” function of Matlab on the raw data. The raw data is represented by matrix A, with dimensions l by t, where l is the number of wavelengths and t is the number of time-resolved snapshots taken during the experiment. Calling the “svd” function using A as the input variable returns a left matrix U with dimensions l by l, a central matrix S with dimensions l by t, and a right matrix V with dimensions t by t.

The right matrix V resulting from the SVD is used to fit a kinetics model described by solutions of ODEs depending on the configuration of the photosystem examined. The model fitting follows a modified version of the method described by Hofrichter(*7*) where a linear transformation matrix C is applied to the expected population matrix P computed from the model. The L_2 norm of the difference between the right t by t matrix and the product CP is minimized by the standard minimization function “fminfunc” available in MATLAB 2014b, in which the rates of various electron transfers in the kinetics model along with the elements of the C matrix are allowed to vary. The minimization function is instructed to utilize the quasi-newton minimization procedure. Various initial conditions are tried in order to obtain the best fitting of the data. The results of the minimization is examined by comparing the residual of the minimization to other values obtained from different sets of initial conditions to ensure that the minimum discovered by the algorithm is, as far as the information reveals, the global minimum or a satisfying local minimum. The values of the C matrix as well as the rates of the electron transfer between the states are then used to compute the reconstructed elemental spectra that make up the observed raw data as well as the electron transfer kinetics that best explains the observed data.

6.3 Results:

6.3.1: High resolution X-ray crystal structure of a novel generation of highly stable and cofactor-specific maquette as the appropriate template for DPA triad design.

Thanks to the efforts of Nathan Ennist of the Dutton Group(manuscript in preparation), a new interpretation of the maquette approach was adopted where the focus of design shifts from a general, omnibus like scaffold capable of supporting and sustaining multiple dissimilar functions to a more specific and particular platform that specializes in a small set of highly similar functions, namely highly stable and independent cofactor binding. The change in the implementation of the maquette approach led to the engineering of the first single-chain 4-helix bundle maquette that has been successfully crystalized and yielded X-ray structure at 1.9A resolution, in both holo- and apo- forms. This design created a 4-helical bundle that contains three independent, modular cofactor "slots" specifically designed for the three roles of a charge-separating triad, as shown in Figure 6.2. The maquette exploits a bis-his ligation to bind iron-porphyrin at one end of the maquette as the A slot, a single-his ligation to bind Zn-porphyrin/Zn-chlorin in the middle of the maquette as the P slot, and finally a less structured D slot consisting of a metal binding site for future catalysis and a generic ligation site for incorporation of high potential, non-tetrapyrrole cofactors.

However, the high specificity is achieved by trading off the versatility and adaptability of the designs of the past generations of maquettes, as both the bis-his and single-his sites are highly engineered, with residues constituting the primary and secondary shells surrounding the cofactors chosen particularly for optimal binding of specific tetrapyrroles, namely, hemes and Zn-porphyrins respectively. As shown in Figure 6.2, the highly specific binding sites for the two tetrapyrroles ligate both the iron- and zinc-porphyrin in place so stably that the electron density for both factors, more so for the heme and less for the zn-porphyrin, can be clearly seen. As intended by design, the maquette shows similar 4-helix bundle width, hydrophobic core packing, rotameric states of heme-ligating histidines, and histidine hydrogen bonds with nearby threonines. However, it was unexpected that heme would be held in such a fixed orientation (low crystallographic temperature factor) as no facet of the design was consciously included to prevent free rotation of the porphyrin ring about the His-Fe ligands while keeping propionate groups solvent exposed. Furthermore, the orientation

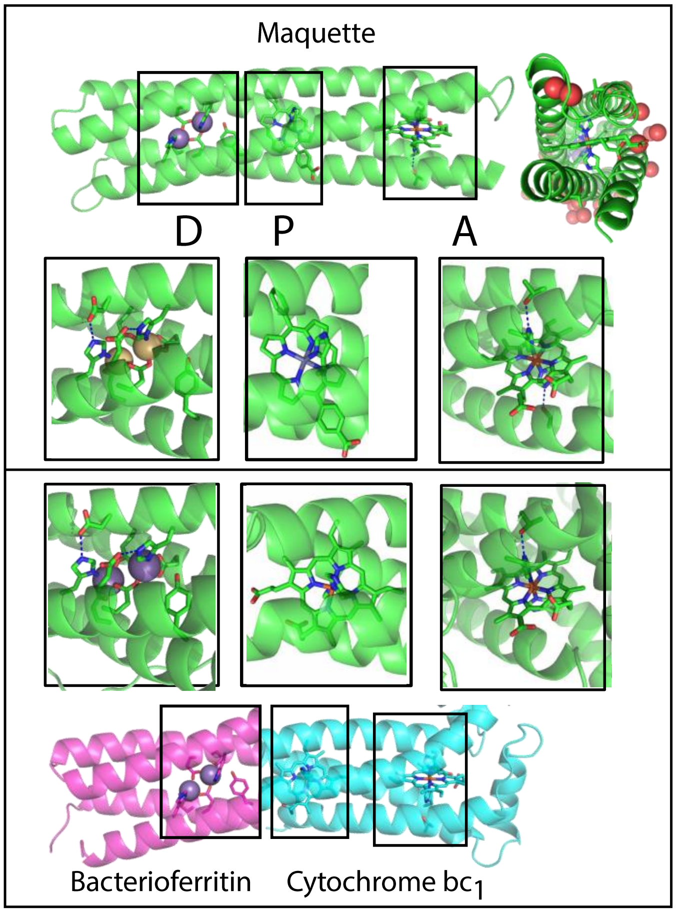


Figure 6.2: Crystal structures of photochemical maquette, top, compared to Mn bound bacterioferritin (PDB ID: 1BFR, magenta) and cytochrome bc_1(PDB ID: 2A06, cyan), bottom. Row 2: enlarged resolved structures around the A, P and D slots with a Tyr and Mn_2 pair, ZnP, and heme. Row 3: close up of the bacterioferritin Mn_2 Tyr center and the two b heme site of bc_1. Row 4: juxtaposed 4-helix units of cytochrome bc_1 and bacterioferritin. Reproduced from Zhao, Ennist, et al(manuscript in preparation).

of the maquette heme relative to the superhelical bundle axis is conspicuously similar to the that observed in cytochrome bc_1, suggesting that the particular conformation of heme in the natural protein is not precisely controlled for any functional purpose, but rather is an accident of the structural constraints demanded by this protein fold. Indeed, in spite of the de-novo nature of the amino-acid sequences comprising the two binding sites, there exist significant similarities between the secondary and tertiary structures of the maquette and naturally occurring cytochrome bc1 and b6f.

The availability of abundant structural information of the new generation of maquette, along with its highly engineered and specific functional slots for acceptor, pigment, and donor meets the requirements not satisfied by the previous generation of maquettes and makes it the appropriate starting point to carry out our design of the DPA triad. However, the highly designed nature of the maquette imposes additional constraints on the design of the triad, limiting the ease with which the acceptor and the pigment binding sites could be modified or moved. We therefore examine the engineering landscape of the DPA triad in this maquette under the constraints of fixed binding sites of the P and A.

6.3.2: Engineering landscape of the DPA triad using the structured maquette as design platform.

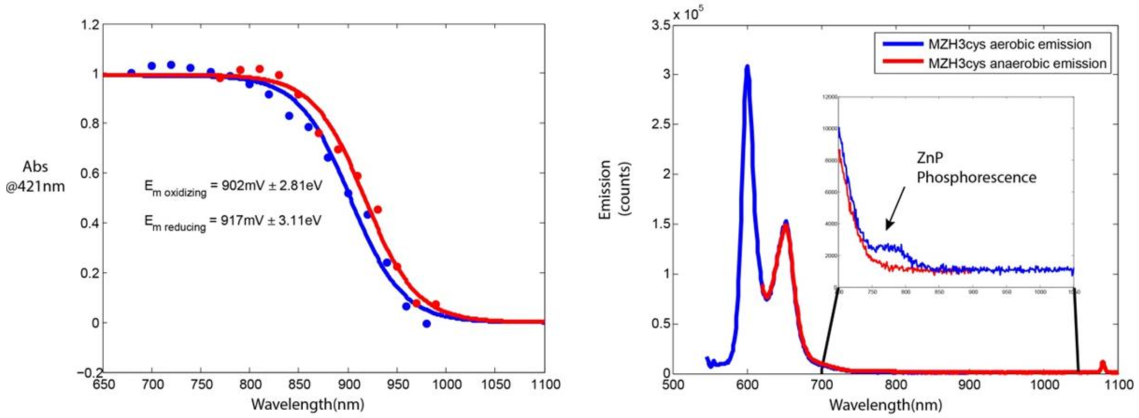


Figure 6.3: Photophysical properpties of the ZnP pigment cofactor for the triad design. Left: electrochemical redox titration of the ZnP in both oxidizing(blue dots) and reducing directions(red dots), fitted to n=1 Nernst equations(red and blue curves) after normalization. **Right**: fluorescence and phosphorescence of the Zn-porphyrin pigment used in the triad construction. The cofactor in maquette is excited at 532nm and its subsequent fluorescence and phosphorescence was collected in both anaerobic and aerobic conditions.

Using the maquette that produced the crystal structure as a starting point, we examine the engineering landscape given the constraints we face, namely fixed distance between the P and A binding sites. This leaves us with the freedom of choosing the exact location of the binding site for the donor(D) within the donor slot as well as its midpoint potential. Additionally, the midpoint potential of the acceptor(A) also serves as one of the engineering parameters that can be explored experimentally in a short amount of time without incurring additional complex and low-yield synthesis to generate other varieties of tetrapyrroles with certain redox and photophysical properties.

We explore the engineering landscape using the type of contour plots as seen in Chapter II, generated from the numerical functional representation of the yield of the charge-separated state in terms of the fundamental parameters of photosystem engineering, viewed from the perspective of keeping everything but two variables fixed. Unlike the previous plots, however, where we explored the theoretical range of design parameters tolerated by a photochemical multi-ads without any practical constraints other than the essential parameters we specified, here we are significantly more limited in our choices of potentials and distances. Specifically, while in the abstract engineering guidelines we focused on the term engineering efficiency as the product of the quantum yield of the charge-separation and the ratio of energy preserved in the charge separation, treating it as a real-valued continuous variable, here the ratio of energy preserved in charge separation is highly discreet thanks to the limited options of acceptor and donor cofactors we have available. Consequently, it is more appropriate for us to examine the engineering contour with respect to the quantum yield of the charge-separated state rather than the engineering efficiency of the charge-separation, as it provides a more direct and fairer metrics for the performance of the device.

As shown in Figure 6.3, the Zn-porphyrin tetrapyrrole shown in the crystal structure of the maquette has a measured redox potential (Em) of 0.9 V and efficiently forms a triplet state by intersystem crossing when excited at 532 nm and persists for 20 ms in the absence of D or A. An excited triplet redox potential of 690mV can be estimated from the observed 780 nm phosphorescence. Using this cofactor as the pigment(P) of the triad, the engineering landscape of the edge-to-edge distances between the P and A or D and the influence of the A and D redox potentials are illustrated in Figure 6.4. The contours follow the anticipated yield of the D+PA- state at 100 ms, a time chosen to facilitate future development of an active catalytic metal center. In panel A, a notable performance "cliff" is observed at the D-P distances longer than 10 Å, as the electron transfer rate of the donor reducing the P^+ radical becomes too slow and can no longer compete with the direct charge recombination from the A^- to the P^+. A similar effect is seen in the redox midpoint potentials of acceptor, represented in practice by the two different iron-porphyrins available, DADPIX and B-heme. The choice of B-heme would result in a significantly lower performance. We therefore favored the choice of DADPIX in our construction of the triad, in addition to a more practical reason relating to the spectral congestion as will be explained below.

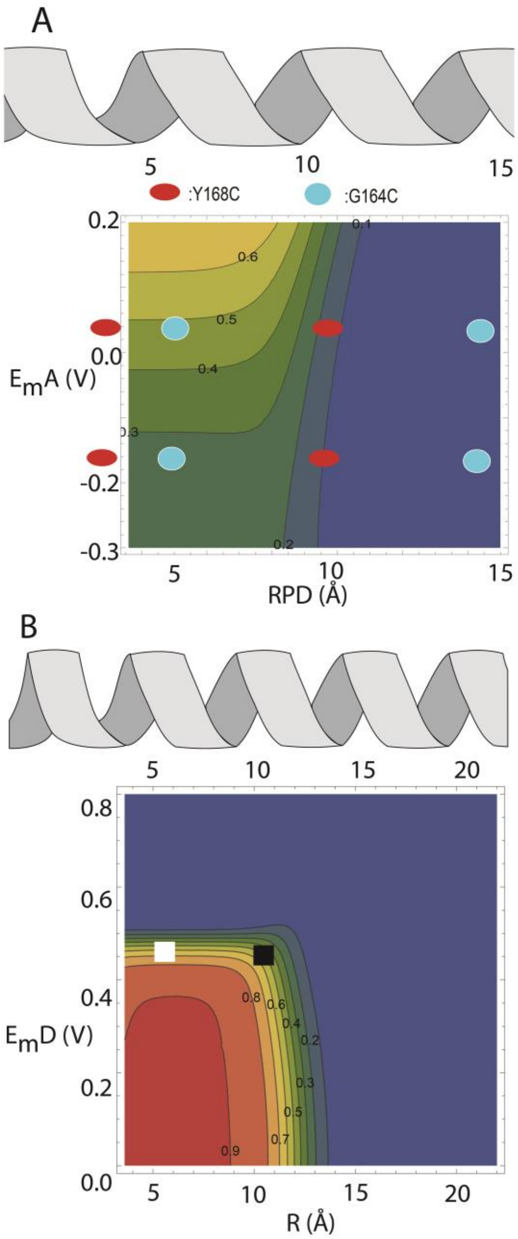


Figure 6.4: Design guidelines of the DPA triad in the maquette platform in the same style of contour plots as in Chapter II, but here of the quantum yield of the charge separated state D+PA- at 100 ms. A: Effect of varying the P to D distance (horizontal axis) and the acceptor redox potential (vertical axis), with the donor potential and P-A distance fixed at 0.5 V and 12 Å. Labels on the top represent changes in spacing arising from one more alpha helical turn (top) between cofactor anchoring amino acids. The red dot represents the variant of the DPA where, thanks to the steric effect described in Figure 6.5, the donor is farther away from the pigment than the other variant represented by the light blue dots. B: corresponding yield contours varying the P to D distance with the acceptor potential and P to A distance fixed at -0.15 mV and 12 Å respectively. Symbols represent four test designs with varying P to D distances and acceptor potentials.

Due to quantization of distance engineering in our maquette as described above, only select regions in the contour plots are accessible by our design, as shown by the helical guide at the top of the figure. We have engineered variants of the DPA triad taking advantage of the quantized nature of the binding sites, and have examined their performances relative to the prediction made here by our model. Because of the flexibility involved with the maleimide linkage to the cysteine residue, there are two orientations of the donor and therefore two possible distances per variant. This will be discussed in detail in the next section.

More importantly, the bottom panel of Figure 6.4 indicates that there is a even more drastic performance drop caused by increases in the redox potential of the donor cofactor chosen. The redox potential of the P/P^+ determines that the donor cofactor must not have redox potentials higher than 550mV, in order to maintain a necessary driving force for the $D\text{->}P^+$ ET to take place.

6.3.3 Incorporation of ferrocene as the donor cofactor into the donor slot of the maquette scaffold at near and far donor positions

The previous section and Figure 6.4B clearly demonstrated the necessity for a donor cofactor that can supply redox potentials in the range above 450mV, far above the upper limit of known tetrapyrroles. As shown in chapter V and other previous studies in the Dutton group, non-tetrapyrrole cofactors have been added to maquettes using click-chemistry or direct modifications via unnatural amino acids.(*8*) However, cofactors that offer midpoint potentials above 300mV had not been attempted, largely due to the general lack of biologically friendly cofactors within this potential range. One of the most common cofactors used as donors for Zn-porphyrin pgiments in the engineering and design of chemical synthetic photochemical systems is ferrocene. Ferrocene has served as the donor cofactor in synthetic triads, tetrads, and pentads that successfully produced charge-separation lasting as long as 510ms(pentad) However, applications of ferrocene as donors in biological photochemical designs have not been reported, to the best of our knowledge. Chapter V has already demonstrated that protein environment, specifically the environment provided by de-novo designed, small, water-soluble proteins, is well suited for ferrocene. Nevertheless, the click-chemistry based method shown in Chapter V requires the incorporation of an unnatural amino acid before the functional moiety could be added, creating a bottleneck in a process where yield is a significant factor. Consequently, a different method for the incorporation of ferrocene in the donor position is required. Due to the availability of a high-resolution crystal structure, however, methods that were once considered unviable for maquettes have now become applicable, such as maleimide-based modification of cysteine residues.

Because of the absence of cysteine residue within the maquette as well as the availability of N-ferrocynyl-maleimide and the high resolution structure of the maquette, maleimide modification was deemed the ideal method of incorporation for the ferrocene cofoactor. It is important to note that maleimide linkage in this specific case is ~3A in length in addition to the dimensions of the side-chain of the cysteine residue and the ferrocene. In the past, when no structural information was available, this additional distance, along with the subsequent uncertainty in the position of the cofactor incorporated via the maleimide linkage, made this method undesirable. However, the crystal structure of the maquette minimizes this uncertainty. Based on modeling using the crystal structure of the holo-protein in pymol, we have identified two optimal anchoring sites, namely, the Y168C and G164C mutations(the G164C mutation includes also the Y168L mutation to remove the tyrosine as a potential donor even though Figure 6.4 has already demonstrated that tyrosine at its native redox potential of 1050mV(*9*, *10*)cannot properly serve as the donor for this DPA triad design, for the maleimide linked ferrocene that provides a near- and a far- configuration for the ferrocene donor, where the effective edge-to-edge distances to the pigment are 6 Å and 12 Å respectively. The near-position, however, requires the placement of the cysteine residue further from the histidine residue ligating the pigment than the far-position, since, as seen in Figure 6.5A-B, the Y168C mutation is too close to the pigment cofactor that due to the steric clashes the maleimide-linked ferrocene must adopt the downward-pointing position. On the other hand, the G164C ligating site, as shown in Figure 6.5C-D, allows the maleimide-linked ferrocene to adopt both the upward and downward pointing positions, but since the downward-pointing position would result in a D-P distance(>15Å) too large to sustain effective electron transfer, the G164C ligating site, although further away from the Y168C, ends up providing the “near-“ ligation for the ferrocene. The difference in the D-P distances provided by the two ligation sites are reflected clearly in the electron transfer kinetics of these two different triads, as demonstrated below in section 6.3.7. However, it must be noted here that the arguments presented above are based on observations made from pymol modeling using a simulated structure of the N-ferrocenyl-maleimide connected to a cysteine sidechain. To obtain additional support, molecular dynamics simulation to examine the preference of the cysteine rotamer will be helpful. Additionally, when the yield of the maleimide modification and subsequent purification could be improved, crystallization of the maleimide-ferrocene modified maquette would be able to offer the most convincing evidence for or against the arguments made here.

Upon deciding where to place the ligating cysteines, we proceeded to express the maquettes with cysteines at the two ligation sites and incorporate the maleimide-ferrocene via standard maleimide-cysteine reaction as described in the methods section. Figure 6.6 demonstrates the successful incorporation of the ferrocene into the maquette, as seen from the comparison of the MALDI-TOF signals of the unlabeled and labeled maquette, as well as successful separation of the reacted fraction of the protein from the unreacted using HPLC. However, it is important to note that the maleimide reaction and the subsequent HPLC separation puts a practical limit on the concentration and volume of the finished ferrocene-maquette adduct to around 20uM/ml at 3-5ml. Moreover, as seen in Figure 6.9, the extinction coefficient of the ferrocene cofactor is around 100-200 times less than that of the acceptor and pigment cofactors. Consequently, it is impractical to determine the redox midpoint potential of the ferrocene in the maquette using the same spectral electrochemical methods used to generate the redox titration data shown above. Instead, methods optimal for low concentration, small sample-size are preferred, such as square-wave step-sized voltammetry were used to determine the in-maquette redox potential of the maleimide ligated ferrocene. The bottom two panels of Figure 6.8 demonstrates the redox midpoint potential of the N-ferrocenyl-maleimide prior to its ligation to the maquette via the above-described cysteine-maleimide reaction using both cyclic voltammetry and spectral electrochemical redox titration. These measurements were conducted with highly concentrated sample of the ferrocene(500mM) so that either a strong enough reducing/oxidizing wave can be observed on the CV or a large enough peak can be observed by the UV-vis spectroscopy.

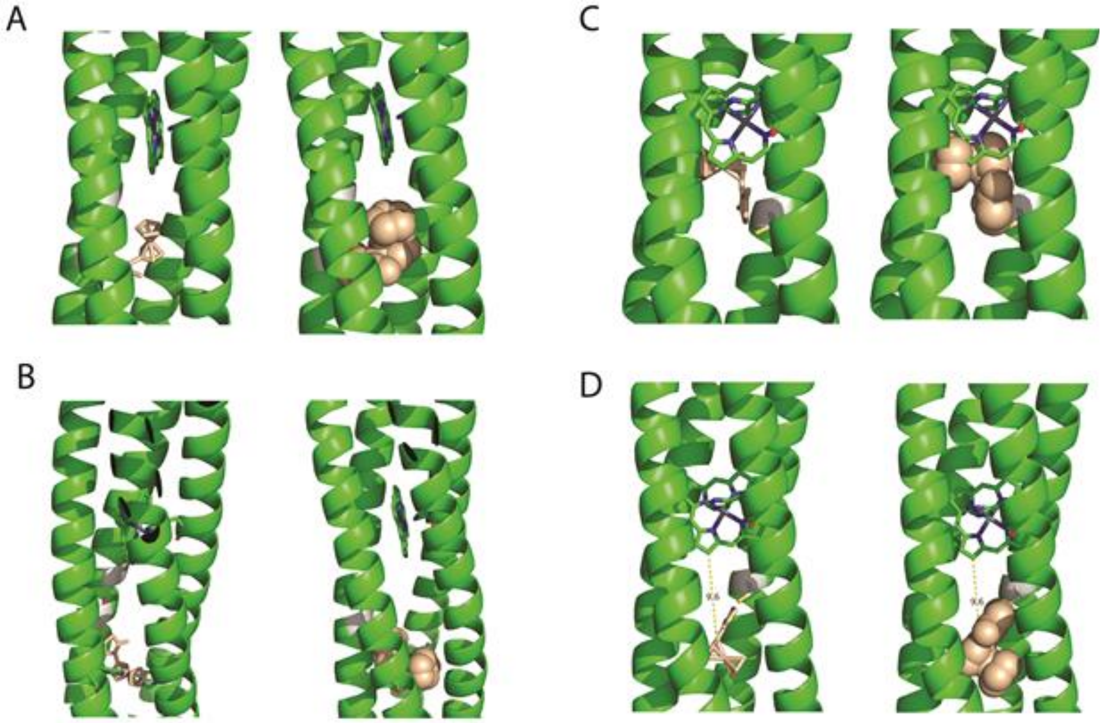


Figure 6.5: Rotamer preferences and their impacts on the effective donor-pigment distances in the G164C(near) and Y168C(far) variants of the DPA triad designs. A: pyMOL model of the pigment and the cysteine ligated maleimide-ferrocene in the G164C mutant when the cysteine rotamer points to the up position, seen in both stick and space-filling representations. This results in an effective D-P distance of 5A. **B:** same as A, except the cysteine rotamer points in the down position, with an effective D-P distance of 15A. **C:** model of the pigment and cysteine ligated maleimide-ferrocene in the Y168C mutant when the cysteine rotamer was allowed to adopt the "up" position, resulting in severe steric clashes with both the pigment cofactor and the protein backbone, as seen in both the stick and space-filling representation of the structure. D: same as C, except with the cysteine rotamer pointing down, causing no steric clashes and resulting in a D-P distance of 9.6A.

6.3.4 Assembly of the Heme-ZnP dyad and Heme-ZnP-Ferrocene charge-separating triads.

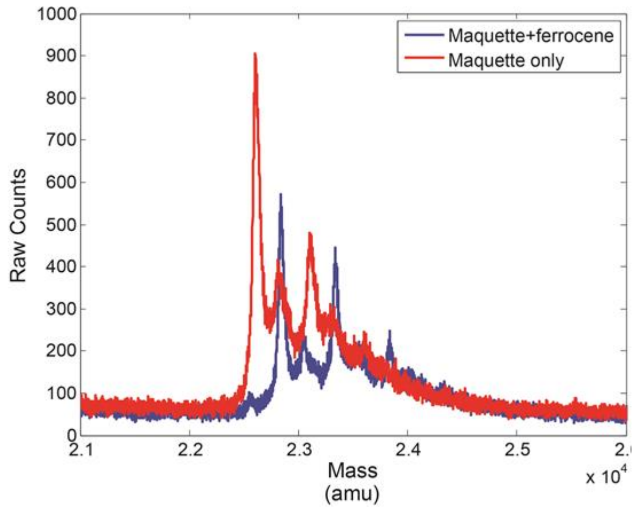


Figure 6.6: covalent incorporation of ferrocene into the maquette via maleimide-cysteine linkage. MALDI-TOF of labeled(blue trace) and unlabeled(red trace) maquettes. Notice the ~250amu difference between the main peaks of the labeled and unlabeled maquette, corresponding to the mass of the ferrocenyl-maleimide incorporated via the labeling reaction.

With the ferrocene covalently anchored to the two intended donor positions, the next step is to examine the binding of the acceptor cofactor and the pigment cofactor, thereby demonstrating that the incorporation of the ferrocene cofactor does not significantly alter the structure and environment of the maquette scaffold and allows for the successful assembly of the DPA triad in the maquette. To ensure site specificity, the assembly proceeds sequentially with the addition of the bis-his ligated iron-porphyrin and the single-his ligated zinc-porphyrin after the ferrocene-cysteine maleimide reaction had been carried out and the reacted species separated from the unreacted. Figure 6.7A and B demonstrates the high affinity of the maquette to both the iron- and zinc-porphyrin with the presence of the ferrocene. Figure 6.8A-B shows the redox midpoint potentials of the two acceptor cofactors of choice, DADPIX and B-heme, obtained via standard spectral electrochemical titration detailed in the methods section, while 6.8C shows the redox potential of the free N-ferrocenyl-maleimide prior to labeling reaction with the maquette, measured using standard cyclic voltammetry. It is interesting to notice that in comparison to previous generations of maquettes shown in Chapter IV and V, this new generation of maquette provides environments for Fe-porphyrin that causes their redox potential to shift towards the positive direction by about 100mV, a significant change based simply on the environment of the cofactors alone. Figure 6.9 shows the UV-Vis absorption spectra of the individual redox spectra of the cofactors of the acceptor, pigment, and donor position, as well as the overall spectral profile of the completed dyad and triads.

While the final goal is the construction and examination of a DPA triad, and successful charge-separating dyads had already been demonstrated in maquettes, in order to ensure the reliability and correctness of our results and analysis of the DPA triad charge separation, it is necessary to understand the complete photophysical and electron transfer details of the P monad, and P-A dyad leading up to the complete triad. The necessity to understand their behaviors is further increased by the fact that the maquette platform upon which the intended triad is constructed belongs to a new generation of single-chain

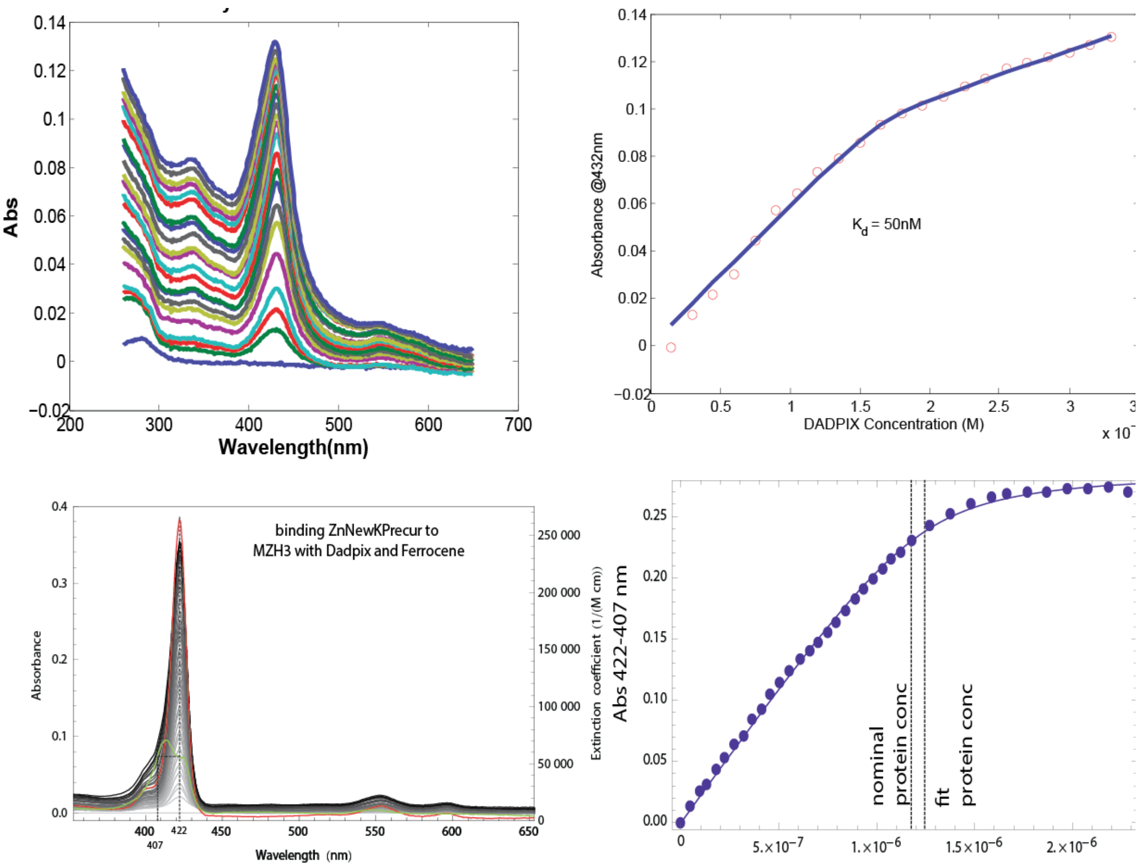


Figure 6.7: Maquette's high affinity for the acceptor heme cofactor and the pigment zinc-porphyrin cofactor after the incorporation of ferrocene. Top left: UV-Vis spectra of a titration of the acceptor cofactor DADPIX into the ferrocene-modified maquette at 20mM Sodium Phosphate and 200mM NaCl., pH 7.4. **Top right:** fitting of the DADPIX binding titration to a standard dissociation constant curve, with the fitted dissociation constant shown in the middle. **Bottom left**: UV-Vis spectra of a titration of the pigment cofactor ZnP into the ferrocene and DADPIX containing maquette at 20mM Sodium Phosphate and 200mM NaCl, pH 7.4. **Bottome Right:** fitting of the ZnP binding titration to standard 1-site binding equation using 1.2μM of the protein in standard PBS buffer(200mM NaCl,20mM $NaPO_4$, 2mM Tris-HCl) at pH7.4

maquettes that had not been used to demonstrate any charge-separating functions.

To assemble the P monad, we simply added the same Zn-porphyrin that had been used in Chapter V, which was also featured in the X-ray crystal structure of the maquette, to the Y168C and the G164C maquettes at 1 molar equivalent. The presence of the unreacted cysteine residue, due to its presence in the highly stable hydrophobic core of the maquette, does not lead to significant dimerization of the maquette, as evidenced by the

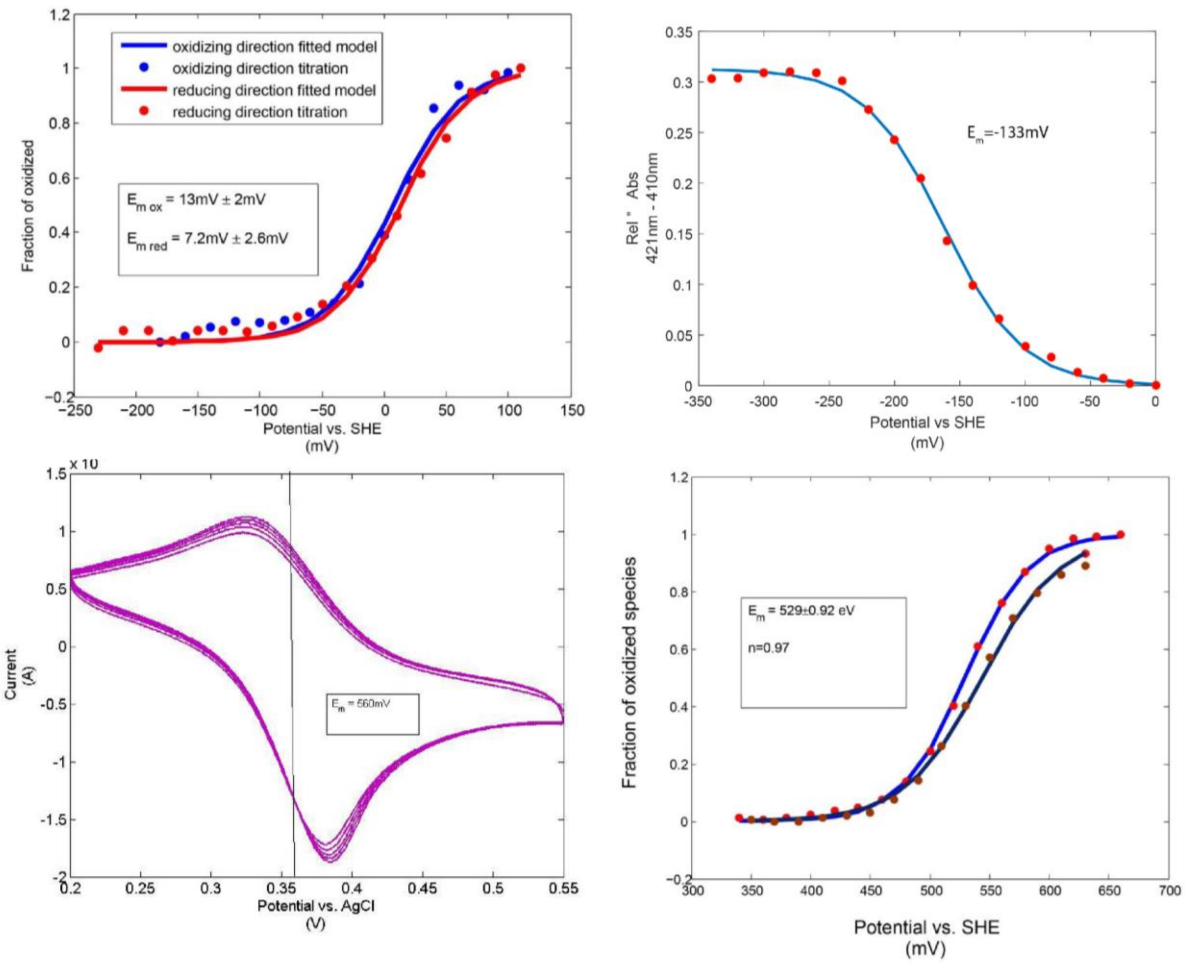


Figure 6.8: Redox properties of the acceptor and donor cofactors of the DPA triad. Top Left: spectral electrochemical redox titration of DADPIX in the maquette. Blue and red dots correspond to the oxidizing and reducing iteration of the titration, while the blue and red traces represent fits to n=1 Nernst equation with the midpoint potential obtained from the fit shown in the middle. **Top Right:** redox teitratin of B-heme in the maquette, similar to top left. Bottom left: cyclic voltammetry of ferrocene-maleimide performed in DMF with 100mM ferrocene-maleimide at room temperature. Bottom right: electrochemical redox titration of 100mM ferrocene-maleimide in DMF, measuring the 450nm absorbance of the ferrocene, whose redox spectra are included in the Appendix.

monomeric mass observed in the MALDI experiments as well as the earlier elution of the unreacted maquette peak in the HPLC purification of the ferrocene labeling reaction. Because of the highly engineered nature of the single-his and bis-his binding sites for their respective ligating partners, the partition ratio of the 1 molar equivalent of the pigment cofactor between the intended P-slot and the empty A-slot will be high. Moreover, due to the lack of electron transfer partners and the overall similarity of the environments of the P and the A slots, P cofactors should display identical photophysical behaviors regardless of the slot it resides in. This will become obvious when the photophysical behavior of the monad is analyzed using a SVD-based global analysis algorithm demonstrated in sections below.

To assemble the P-A dyad, we first ligate the acceptor cofactors, either DADPIX or B-heme, to the bis-his site of the maquette without subjecting it to the ferrocene labeling reaction first, same as in the assembly of the P monad. Then upon a de-salting column crude purification, we titrate in sub-stoichiometric amount of the ZnP pigment. This is essentially identical to the assembly of the DPA triad, except we use the unreacted mutant of the maquette, where the single free cysteine residue remains buried in the

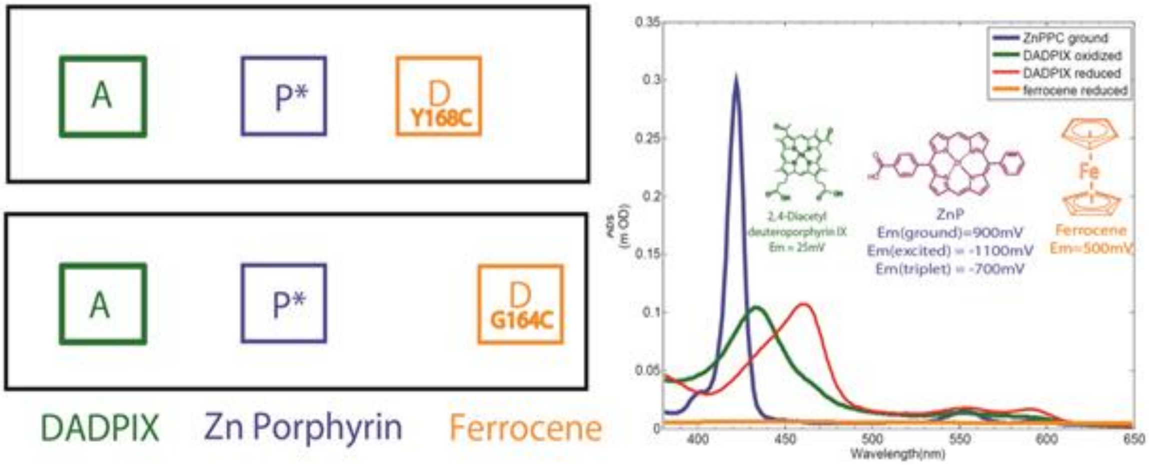


Figure 6.9: Assembly of the DPA, DADPIX-ZnP-Ferrocene triad in the scaffold maquette. Left: Cartoon representation of the triad with the ferrocene anchoring site shown in both the near and far positions. Right: UV-Vis spectra of the redox spectra of the acceptor cofactor(DADPIX) and the spectrum of the pigment(ZnP), and the redox spectra of the donor(ferrocene) molecule.

hydrophobic core.

6.3.5 Illustrating the effectiveness of the SVD-based global analysis with simulated data

Because of the highly overlapping spectral signals of the cofactors that make up the DPA triads as seen in Figure 6.9, we intend to construct due to the limited available pool of off-the-shelf candidates, single-wavelength analysis of the transient absorption spectra is expected to be ineffective and has a high probability of returning inaccurate or erroneous results. Consequently, an analysis approach that can uncouple the overlapped signals and

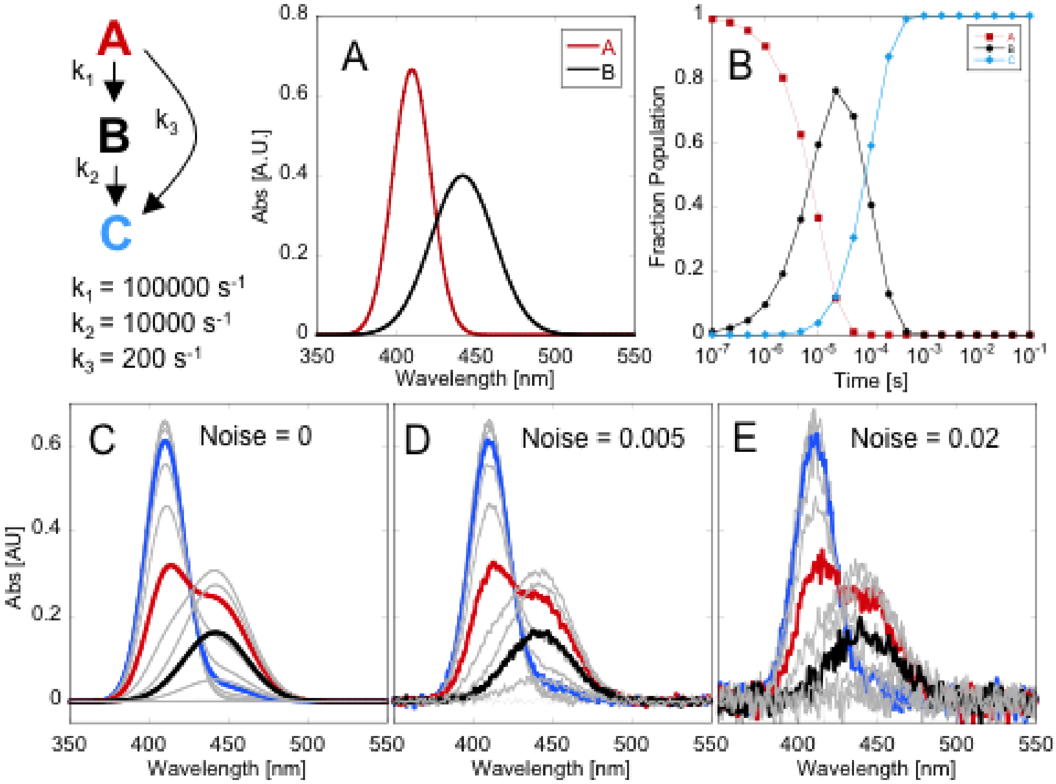


Figure 6.10. Synthetic spectra generated to validate the singular value decomposition (SVD) model-fitting algorithm. Synthetic "observed" spectra were generated based on the 3-state kinetic model shown in the upper-left. **A)** Absorption spectra for the imaginary species "A" and "B" in the model. Species "C" is spectrally silent. **B)** The populations of the three states at various timepoints, given the rates in the upper-left. **C)** The "observed" spectra, computed by multiplying the pure spectra in **A** and the populations in **B.** This represents a perfect observation with no noise. **D)** and **E)** show the same spectra with the addition of Gaussian noise. **Table 6.2** shows the computed rates for various noise amplitudes. Figure generated by Bryan Fry, with simulated spectral and kinetics data provided by me.

capture their independently varying nature is highly preferred. As described in the methods section, SVD-based global analysis is commonly used for time-dependent spectral data consisting of multiple indepdently varying signals with high autocorrelation but low correlation with each other throughout the entire time domain.

Noise Level	**k_1 =100000 s^{-1}**			**k_2 = 10000 s^{-1}**		
	%Err ($\overline{k_1}$)	%Stdev (k_1)	Max %Err (k_1)	%Err ($\overline{k_2}$)	%Stdev (k_2)	Max %Err (k_2)
0.001	0.006	0.10	0.22	0.036	0.12	0.28
0.002	0.009	0.22	0.62	0.041	0.24	0.67
0.005	0.095	0.70	1.7	0.15	0.56	1.5
0.01	0.32	0.97	2.2	0.31	1.1	2.4
0.02	0.014	2.1	4.5	0.10	2.7	8.1
0.05	3.8	5.0	16.5	4.1	6.2	15.0
0.1	22.3	23.3	80.5	71.3	129	400

Table 6.2: testing the performance of SVD-based model fitting under various noise levels

To illustrate the correctness and the effectiveness of the said method of analysis described in detail in the methods section, we generated simulated time-resolved spectral data following a two state kinetics model shown in Figure 6.10A that’s similar to the one that describes the ET kinetics of the P-A dyad. The spectral component of the data consists of two independently varying Gaussian signals with varying degrees of Gaussian noises added. The kinetics profile of the 3 species is shown in Figure 6.10C, which should be returned by the global analysis algorithm upon successful SVD and model fitting analysis. The ability of the global analysis to correctly uncover the underlying true component spectral with increasing levels of noise is shown in table 6.1. It is important to note that this method remains robust and noise-resistant, providing acceptable accuracy in terms of the kinetics parameter discovered from fitting until the noise level begins to overwhelm the signal even by naked eye, as seen in Figure 6.10 Together these results demonstrate the reliability and effectiveness of the global analysis using SVD we have developed to overcome the overlapping spectral signals of the cofactors constituting our proof-of-principle triads.

6.3.6 Examining the photophysics of the P monad in the scaffold maquette.

Upon demonstrating the successful assembly of the DPA photochemical triad with the intended site-specificities for all three cofactors as well as the assembly of monads and dyads that serve as necessary controls, the next step is to examine the photophysical and photochemical details of the charge-separation sequentially from the pigment-only monad to the P-A dyad, and finally to the complete DPA triad. For each configuration, we report the details of the electron transfer reactions kinetics examined using nanosecond transient absorption spectroscopy and subsequently analyzed using a SVD-based global analysis method, described in detail in the methods section.

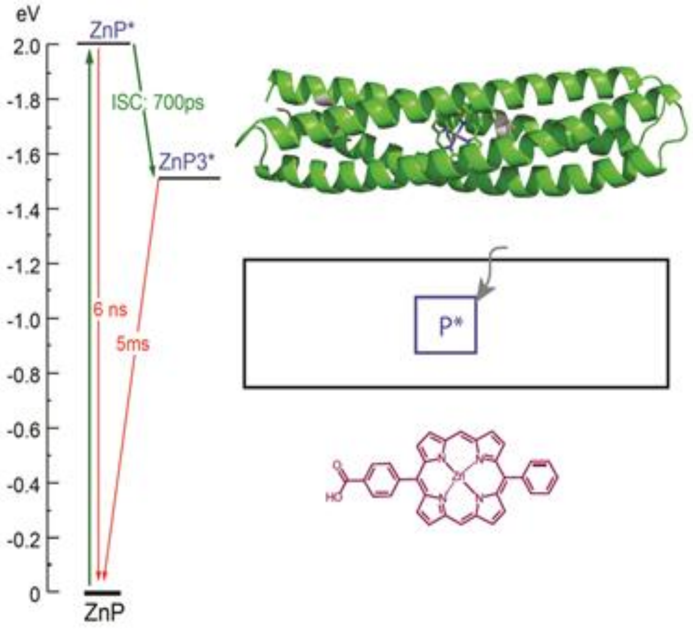


Figure 6.11: Structural and kinetics model of the P-monad in the maquette scaffold. Left: energy diagram and kinetics model of the P-monad. The scale on the left indicates the energy of the lowest lying singlet excited of the ZnP used, which converts to a triplet excited state via intersystem crossing(ISC), with some loss in energy. **Right:** Representation of the electron transfer active components of the monad relative to the maquette.

Figure 6.11 describes the structural and expected kinetics model of the P-only monad, while Figure 6.12 demonstrates the raw transient absorption spectra of the monad upon laser excitation and the SVD-based global analysis and the resulting electron transfer kinetics it displays when examined under transient absorption spectroscopy. The raw data suggests that the recovery of the ground state soret bleach occurs on the same timescale as the disappearance of the broad signal from 440nm to 470nm, indicative of the triplet excited state. The global analysis reveals that the kinetics consists of a single significant component, whose time evolution can be fitted to a single exponential model with a lifetime of ~20ms. Together the results suggest that when the Zn-porphyrin is the only cofactor ligated to the scaffold maquette, the product is indeed a photochemical monad, whose electron “transfer” kinetics follows the model shown in Figure 6.12. While the monad might seem extraordinarily uninteresting as is, it serves as the necessary and ideal control for the dyad and triad assembly soon to be explored. It ensures that the triad design is stable and the site specificity for the pigment cofactor is good enough that a uniform and singular kinetics is observed in spite of the fact that when the acceptor site is empty pigment cofactors can be ligated there as well as in the intended pigment site.

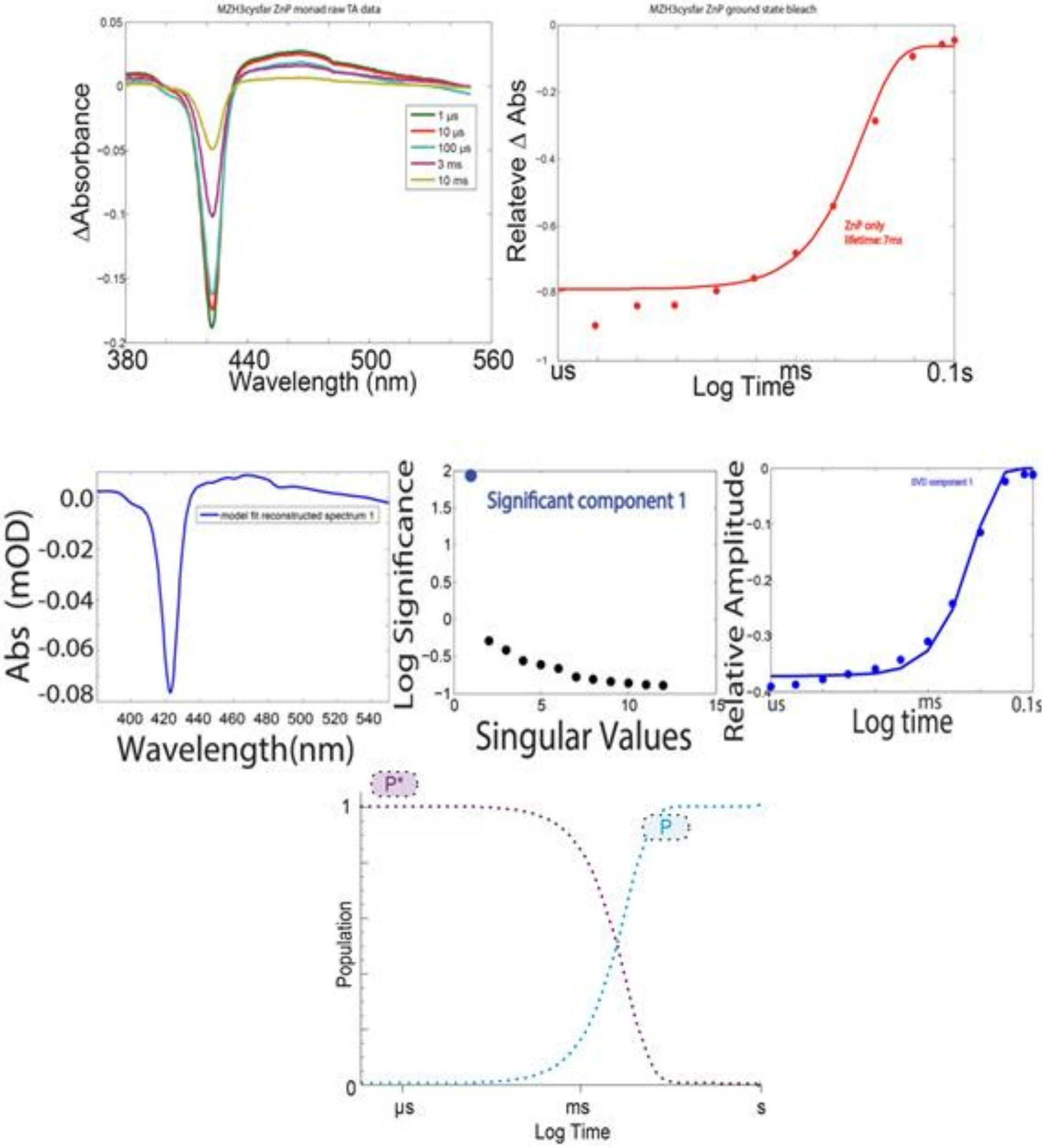


Figure 6.12: Transient absorption spectra and SVD-based global analysis of the P-monad kinetics. Top left: Transient absorption difference spectra of the P-monad, with traces corresponding to various timepoints after laser excitation. Experiments performed in 20mM Sodium Phosphate, 200mM NaCl, pH7.4, with 5uM maquette, 1uM glucose, 20nM glucose oxidase/catalase suite for oxygen scrubbing function. Two significant spectral signatures can be observed: the bleach of the ground state absorption at 424nm, and the appearance of the triplet state excited state between 440nm to 480nm. **Top Right:** Time evolution of the ground state bleach recovery throughout the transient absorption experiments. Red trace represents a single exponential fit of the experimental data. Mid Left: SVD generated spectral component 1 of the P-monad. Mid mid: Distribution of the singular values of the SVD analysis. Mid Right: Quality of fit for the model shown in Figure 6.11. Bot: Complete kinetics model of the P-monad using rates obtained from the global analysis.

6.3.7 Electron transfer and charge separation kinetics of the P-A dyad in the scaffold maquette

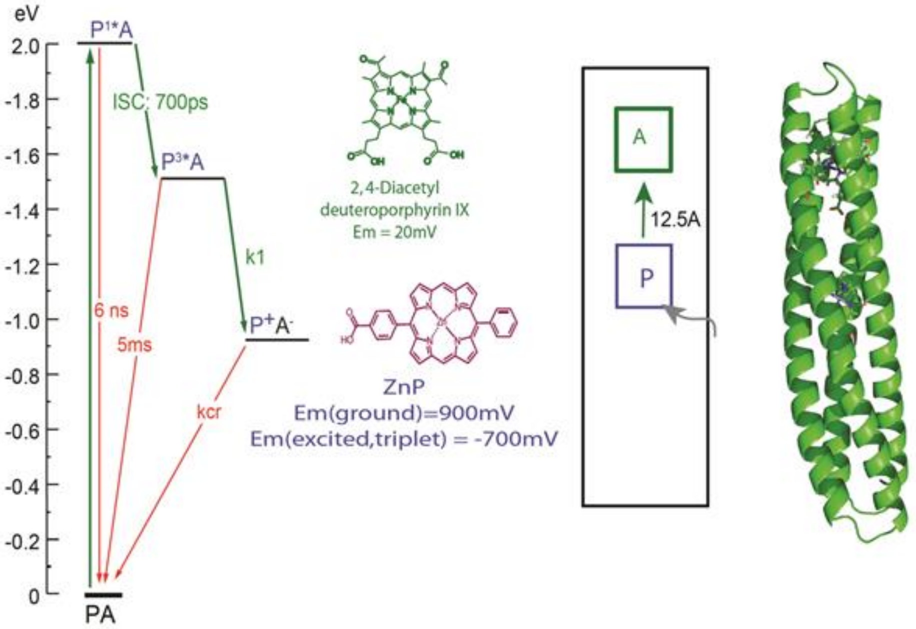


Figure 6.13: Assembly of a P-A dyad in the maquette scaffold using DADPIX and ZnP without modifying the protein with ferrocene. Left: expected kinetics model of the ZnP-DADPIX dyad. In addition to the intersystem crossing from the singlet excited state of the P to the triplet excited state, a forward charge-separating electron transfer to form the PA CS state is now possible, with a rate of k1. A charge recombination to reform the ground state with the rate of kcr is also possible. Mid: structures and redox midpoint potentials of the cofactors used for the dyad. Right: structural details of the PA dyad in the maquette scaffold.

Figure 6.13 illustrastes the structural and kinetics model we expect of the PA, DADPIX-ZnP, dyad assembled in the scaffold maquette. Unlike the P monad, the P-A dyad in the scaffold maquette is now capable of sustaining charge-separation resulting from electron transfer between two cofactors. In order to ensure that electron transfer is indeed taking place, a continuous illumination experiment was first conducted to demonstrate that the acceptor is driven to accumulate in the reduced state while the dyad is under continuous illumination of strong wide-spectrum light in the presence of a sacrificial donor. This ensures that, even when the forward electron transfer from the excited state of P to A is slow and thereby making the charge-separation kinetically unfavorable, we can still obtain evidence of the successful formation of the charge-separated state. The results of the continuous illumination on a DADPIX-ZnP dyad is shown in Figure 6.14. We indeed observed that over the course of 40mins of continuous exposure to white light, the originally fully oxidized DADPIX acceptor gradually becomes fully reduced, while the ZnP component of the spectrum did not experience significant changes. The successful observation of the accumulation of the redox-difference spectral signal of the acceptor indicates that the forward electron transfer from P to A does indeed take place. The next step is to examine whether a stable charge-separation can be sustained by the dyad or not.

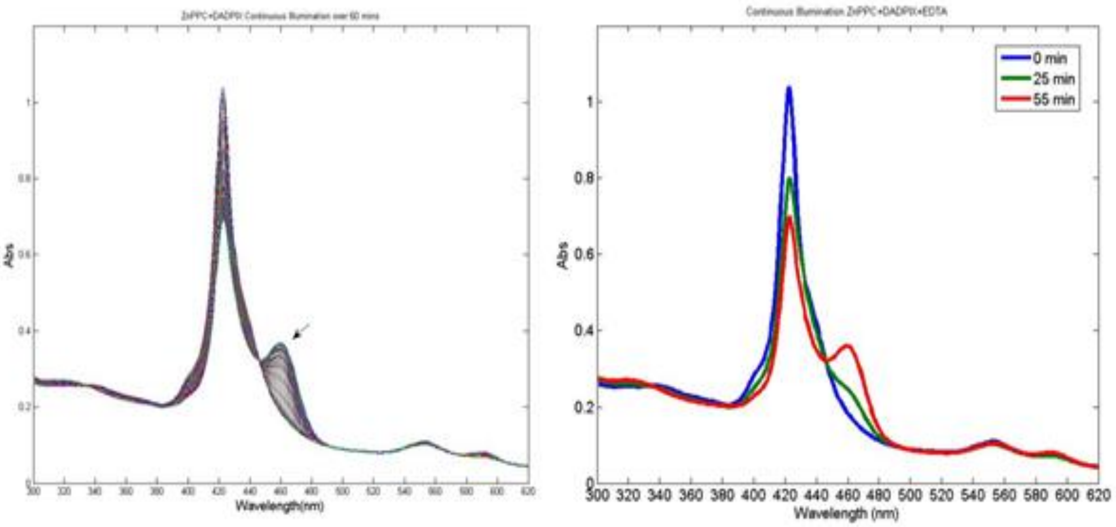


Figure 6.14: Continuous illumination of the P-A dyad in the presence of sacrificial donor. Left: UV-Vis absorption spectra of the PA dyad(DADPIX-ZnP) under continuous illumination of strong white light. The arrow indicates the appearance and accumulation of a spectral signature at 461nm, indicative of reduced DADPIX. Spectra were collected at 1 min interval in 20mM Sodium Phosphate, 200mM NaCl, in the presence of the oxygen scrubbing suite described in the methods section. **Right:** Examplary spectra of key time points during the continuous illumination experiment, demonstrating the decrease in the soret peak of the oxidized DADPIX and the appearance of the reduce DADPIX.

In order to fully understand the electron transfer kinetics of the dyad, however, it is necessary to obtain time-resolved information using transient absorption, with the same setup as in the previous section. As shown in Figure 6.15, the P-A dyad, unlike the P monad, has a significantly faster ground-state bleach recovery rate, implying that electron transfer is indeed taking place. A further examination using global analysis, shown in figure 6.16, informs us that the dyad is able to form a charge separated state at around 5us after the laser excitation, and the charge-separated state recombines to form the ground state in 100us. The global analysis identifies two independently varying signals, one corresponding to the bleach recovery of the ground, while the other represents the charge-separated state between the pigment and the acceptor. Consequently, we were able to fit the SVD-separated independent signals to a 3-state model shown in Figure 6.13 and obtain the rates described above. The quality of the model fit in the analysis demonstrates that the 3-state model accounts for the experimental kinetics observed in the transient absorption, as the model-generated component difference spectra correspond well with the expected difference spectra of the two species involved in the P-A electron transfer. Difference spectrum 1 should correspond to the superposition of the triplet excited state spectrum and the negative of the ground state spectrum of the pigment, without any contribution from the acceptor spectrum. Difference spectrum 2 should consist of the superposition of the spectrum of P^+, the redox difference spectrum of A, and the negative of the ground state spectrum of the P. We notice that the reconstructed spectra 1 and 2 from the SVD-based global analysis, shown in Figure 6.15 as well, largely agree with the spectral signatures of the species we would expect if the kinetics of the dyad-electron transfer follows the model shown in Figure 6.16. Especially encouraging is the observation of the positive spectral signal near wavelength 405nm in difference spectrum 2, a signature of the cation of zinc-porphyrins as seen by various experimental observation of the cation outside of efforts by our lab.

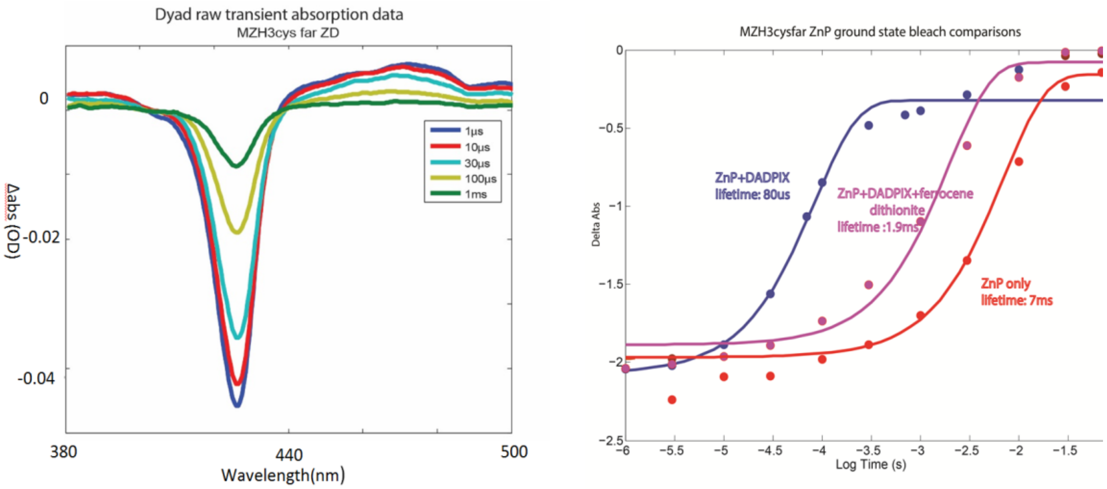


Figure 6.15: Transient absorption spectra of the P-A dyad in the maquette. Left: raw data from the transient absorption experiment, with 6 snapshots at various delays after the laser pulse shown in different colors. Right: Estimates of the ET kinetics based on the recovery of the ground state bleach at

Importantly, the P-A dyad experiment also serves as a necessary and essential control for the following DPA triad experiments, since it demonstrates that the formation of the long-lived charge-separated state in the triad experiments is unequivocally due to the presence of the ferrocene donor that is not present in the P-A dyad assembly, which did not produce any evidence of reduced acceptor cofactor that lasts for hundreds of miliseconds.

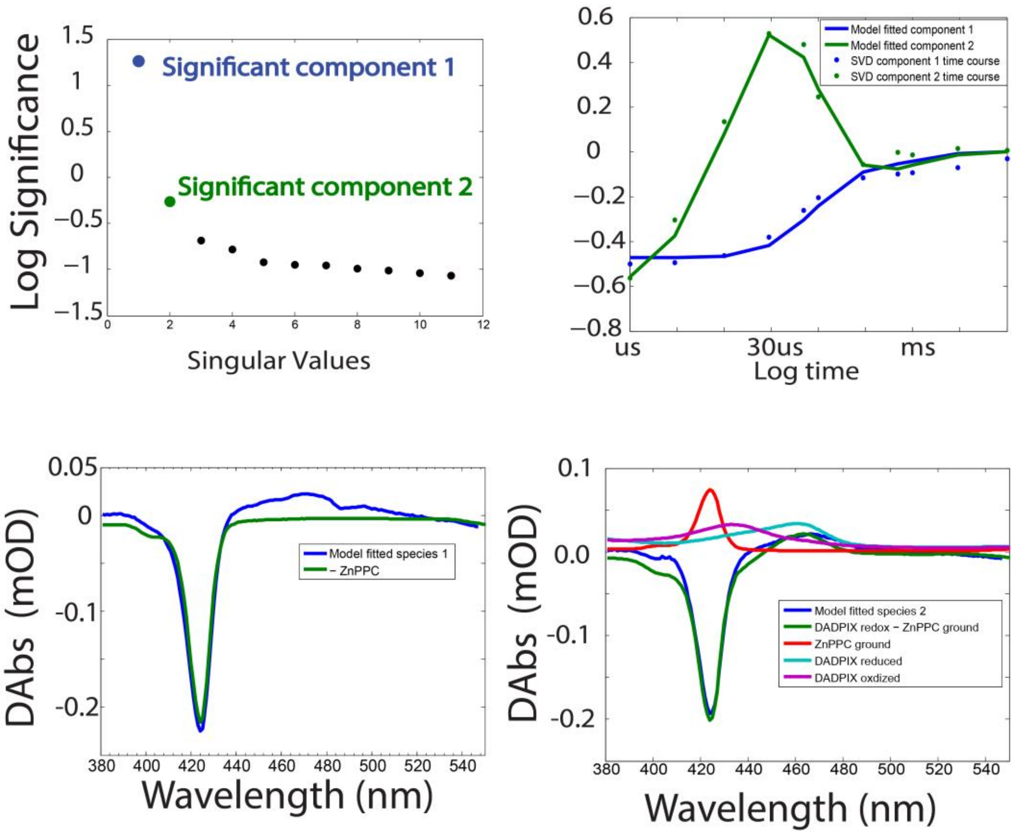


Figure 6.16: SVD-based global analysis of the P-A dyad in the maquette. Top Left: Significance of the components identified by the SVD on the raw data. Two significant components are expected from the dyad kinetics model. Top Right: quality of fitting a two state model to the first two principle components, dots represent the SVD-generated time evolution of the two principle components, while lines are the transformed time evolution of the two species using the fitted constants.

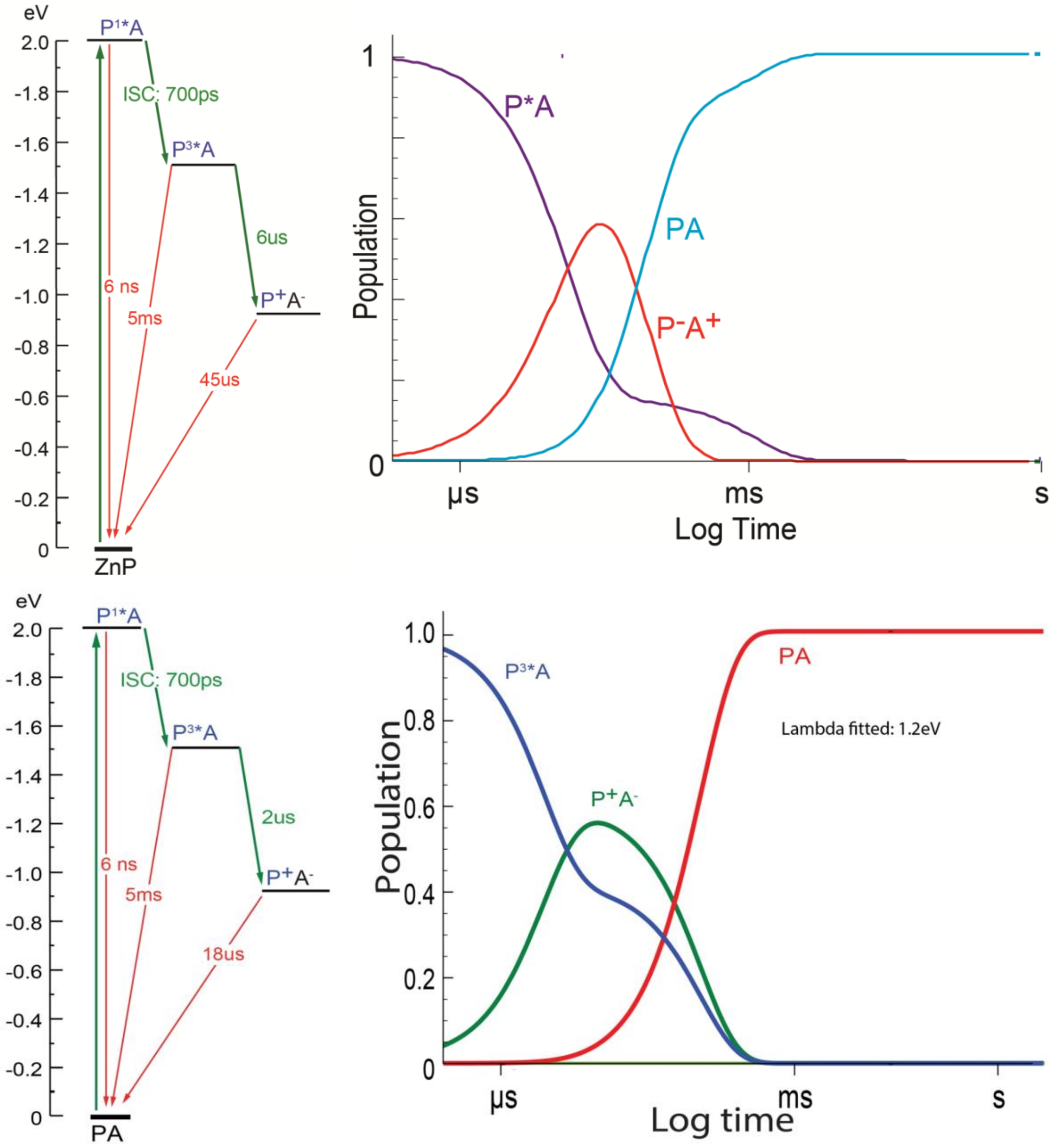


Figure 6.17: Electron transfer kinetics and charge-separating efficiency of the P-A dyad in the maquette. Top: details of the kinetics of electron transfers obtained from the model fitting of the SVD-based global analysis. The ISC, 6ns rate of the singlet excited state decay are from prior knowledge and not determined from experimental results. Bottom: kinetics of electron transfer generated from the predictive analysis stage of a P-A model dyad from Chapter II, using 12Å as the ZnP-DADPIX distance estimated from the crystal structure, and measured Ems as described in the chapter. Λ=1.2eV is the best fitted reorganization energy.

Interestingly, the quality of the model fitting of the global analysis is significantly improved when the 3-state model shown in figure 6.13 is supplemented with a mixture

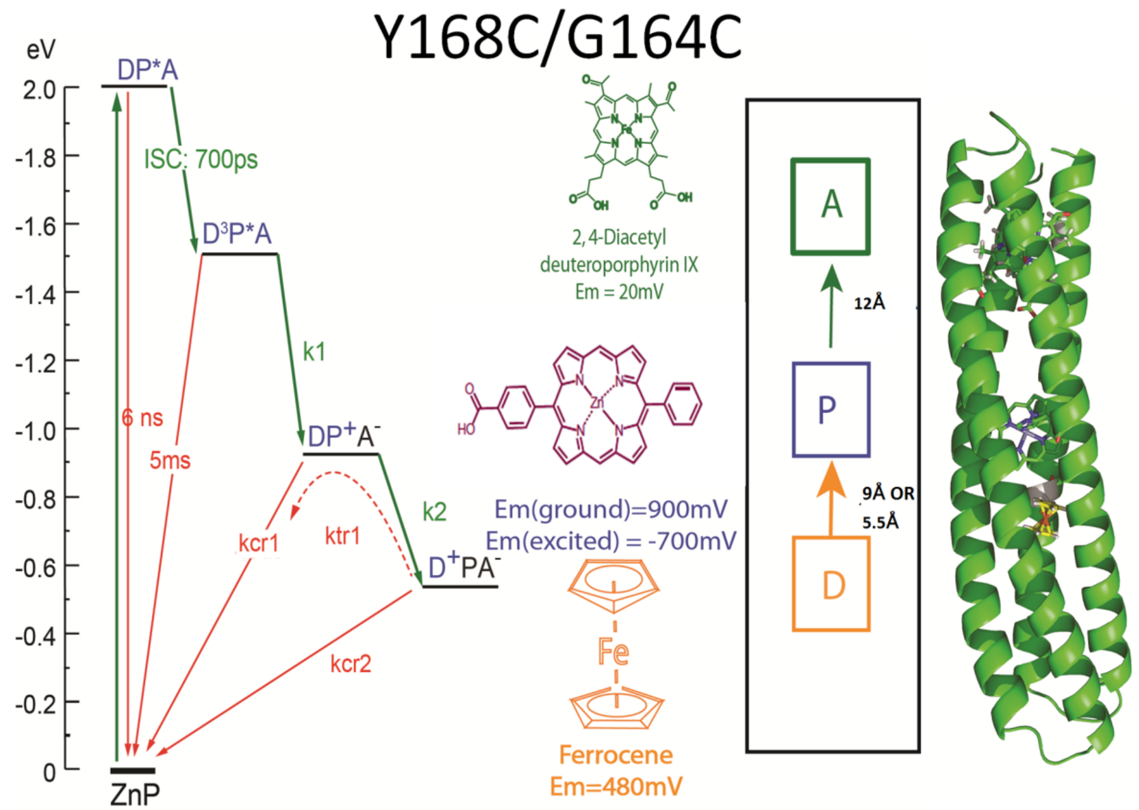

Y168C/G164C
eV
2.0
-1.8
-1.6
-1.4
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0
DP*A
ISC: 700ps
D³P*A
k1
DP⁺A⁻
k2
D⁺PA⁻
6 ns
5ms
kcr1
ktr1
kcr2
ZnP
2,4-Diacetyl deuteroporphyrin IX
Em = 20mV
Em(ground)=900mV
Em(excited) = -700mV
Fe
Ferrocene
Em=480mV
A
12Å
P
9Å OR 5.5Å
D

Figure 6.18: The near- and far- versions of the DADPIX-ZnP-Fc triad. Left: The energy diagram of the various ET states and corresponding rates of electron transfer between them. The rates are written in abstract form, although the energy levels are drawn accurately to scale. Middle: the cofactors used to assemble the triad and their distances. Right: the model structure of the DPA triad.

with the P-monad model shown in Figure 6.10, with an additional parameter in the fitted model representing the fractional contribution of the 3-state and the 2-state model. The minimization process of the model fitting returned a fraction variable valued at 0.9, implying that 10% of the P-A dyad assembled in the maquette scaffold contains pigment or acceptor cofactors that did not participate in the electron transfer reaction, causing a small population of the pigment cofactor to reproduce the P-monad behavior shown in the previous section. This can be seen in Figure 6.16 as the small population of the triplet excited state, shown as residual ground state bleach that persisted until 5ms before it fully recovers.

Figure 6.17 also illustrates the kinetics of a P-A dyad modeled using the predictive analytical algorithm from Chapter II, with an inter-cofactor distance of 12Å, estimated from replacing the B-heme in the crystal structure with a DADPIX in pymol, driving force of -600mV, and a fitted reorganization energy of 1.2eV. We notice that the amount of discrepancies between the experimentally determined and theoretically predicted rates is within half an order of magnitude, demonstrating decent agreements between the theories and experimental results.

6.3.8: Electron transfer and charge-separation kinetics of the canonical ferrocene-ZnP-DADPIX triad

With the kinetics of the P monad and the P-A dyad fully elucidated, we are now in a position to examine the electron transfer kinetics of the four D-P-A triads we have engineered. We begin with the ET kinetics of the ferrocene-ZnP-DADPIX triad constructed with the Y168C mutant, where the ferrocene adopts the downward pointing orientation that results in a greater, ~9A distance between the pigment and the donor. Figure 6.19 demonstrates the raw transient absorption data, while Figure 6.20 illustrates the SVD-based model fitting analysis as described in the previous sections. Unlike the P monad and the PA dyad, the DPA triad demonstrates direct evidence of long-lived

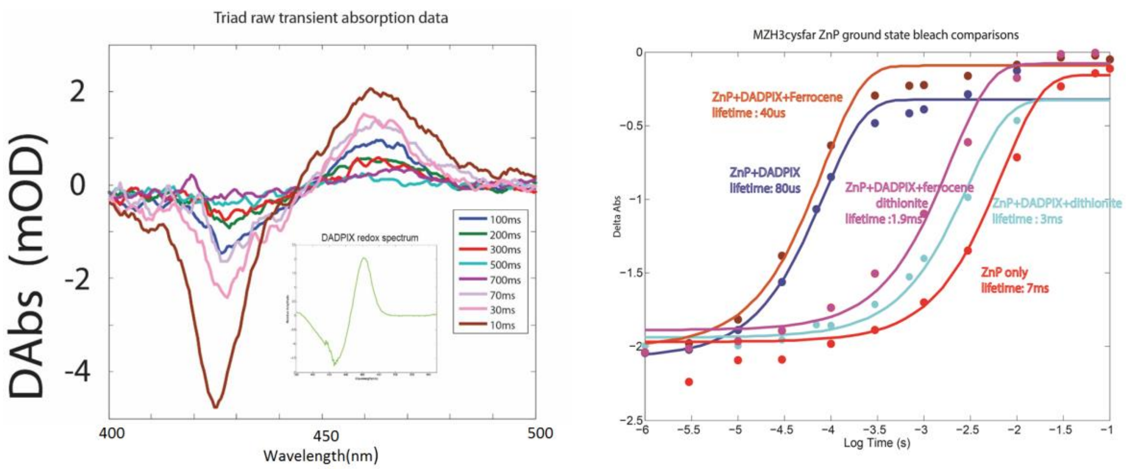


Figure 6.19: Formation of long-lived charge-separated state directly observable from the raw data of the DADPIX-ZnP-Fc triads. Left: raw transient absorption spectra of the DZF(near) triad at delay times after 10ms, all the way up to 700ms. The insert is the redox difference spectrum of the DADPIX for comparison. Right: Time evolution of the ZnP ground state bleach signal(ΔAbs @425nm) for monad, dyads under reduced and oxidized environments, and triads under reduced and oxidized environments.

charge-separation in the raw transient absorption data before the application of the SVD-based global analysis to uncover the spectrum of the charge-separated state and its rates of formation and disappearance, as the lifetime of the charge-separated state is many orders of magnitude and almost 2 orders of magnitudes greater than that of the ground-state soret bleach. This not only provides direct and strong evidence for the successful generation of charge-separated state that lives for hundreds of miliseconds, but also gives us a very straightforward and independent method to estimate the yield of the charge-separation without the global analysis followed by kinetics model fitting. Figure 6.20 illustrates the details of yield estimation for the DADPIX-ZnP-ferrocene triad in both the near- and far-mutants. It's interesting to note that, as discussed in section 6.3.3 and predicted in Figure 6.4, the Y168C mutant, whose ferrocene is further away from the pigment, indeed generated charge-separation with lower yield than the G164C mutant with a closer positioned ferrocene donor. In fact, the difference in the yields of the two variants of the triad as shown in Figure 6.20 corresponds very well with the predicted differences in the yields shown in Figure 6.4. The yield estimate shown in Figure 6.20 takes advantage of the fact that at early timepoints, the transient absorption data is dominated by the signal of the ground state bleach of the pigment, while at much later timepoints, the signal consists solely of the charge-separated state. Since the charge-separated state spectrally consists mostly of the redox difference spectrum of the acceptor, in this case DADPIX, whose extinction coefficient is known to us from standard spectra, this allows us to estimate the concentration of the charge-separated state as well as the concentration of the pigment that is excited by the laser. The ratio of the two gives us an estimate of the yield. Using this method, as shown in Figure 6.20, the yields of the two variants of the DADPIX-ZnP-Fc triad are calculated to be 62% for the near-variant and 31% for the far-variant, in accordance with the predicted yields from Figure 6.4.

Although there is enough direct evidence from the raw data to suggest that we have successfully engineered a triad that separates charges for record long lifetimes, it is necessary to carry out the more rigorous and detailed global analysis to ensure both the authenticity of the charge-separated state and the exact length of its lifetime. SVD-based global analysis requires fitting the experimental data to a kinetics model, and in the case of the complete DPA triad, the full kinetics model is shown in Figure 6.18A. The full model consists of five separate rate constants and four separate kinetics states. Consequently, the full model requires a rather messy set of linear combinations of exponentials, generated from the analytical solutions of the system of ODEs that describe the movements of electrons in the system. Additionally, the full model would also require using the top 3 rather than 2 principal components of the kinetics data, as the third principle component represents the intermediate charge-separated state P^+A^-. However, the triads engineered and analyzed in this chapter have their cofactors arranged linearly, which, as demonstrated in both Chapter II and III, causes the charge-recombination from

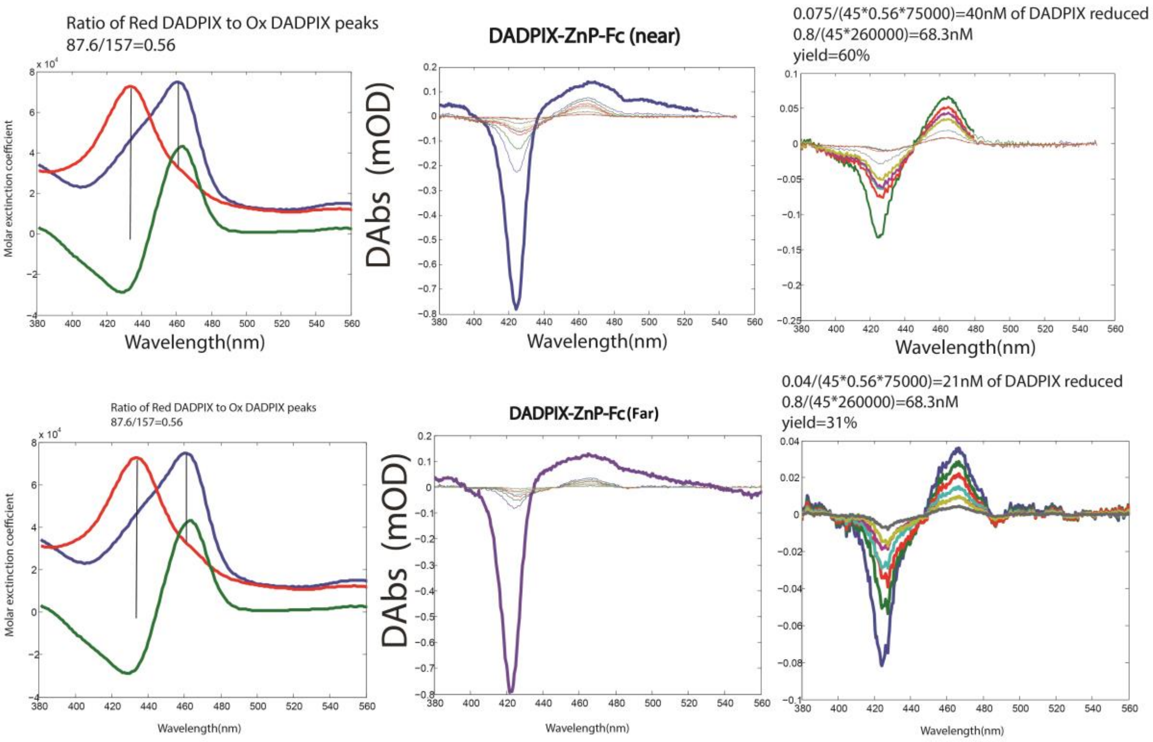


Figure 6.20: Direct estimate of yields of the near- and far-versions of the DADPIX-ZnP-Fc triads. Top Row: yield estimation of the near-variant. Bottom row: yield estimation of the far-variant. Left to right: reduced and oxidized spectra of the DADPIX with molar extinction coefficients as units; raw TA spectra of the triad at timescales where the Zn ground state bleach dominates(blue, thick) and where the long-lived charge-separating triad dominates(thin); zoomed-in view of the CS state at timescales >10ms.

the final CS state to take the uphill-repopulation route rather than the direct recombination with the donor radical D^+. This results in the reformation of the P^+A^- intermediate charge-separated that’s on the same time scale as the disappearance of the final charge-separated state. In fact, by simply examining the time evolutions(vectors of the V^T matrix of the singular-value decomposition of the transient absorption data) of the third principle components of the near and far triads, as shown in Figure 6.21, we can

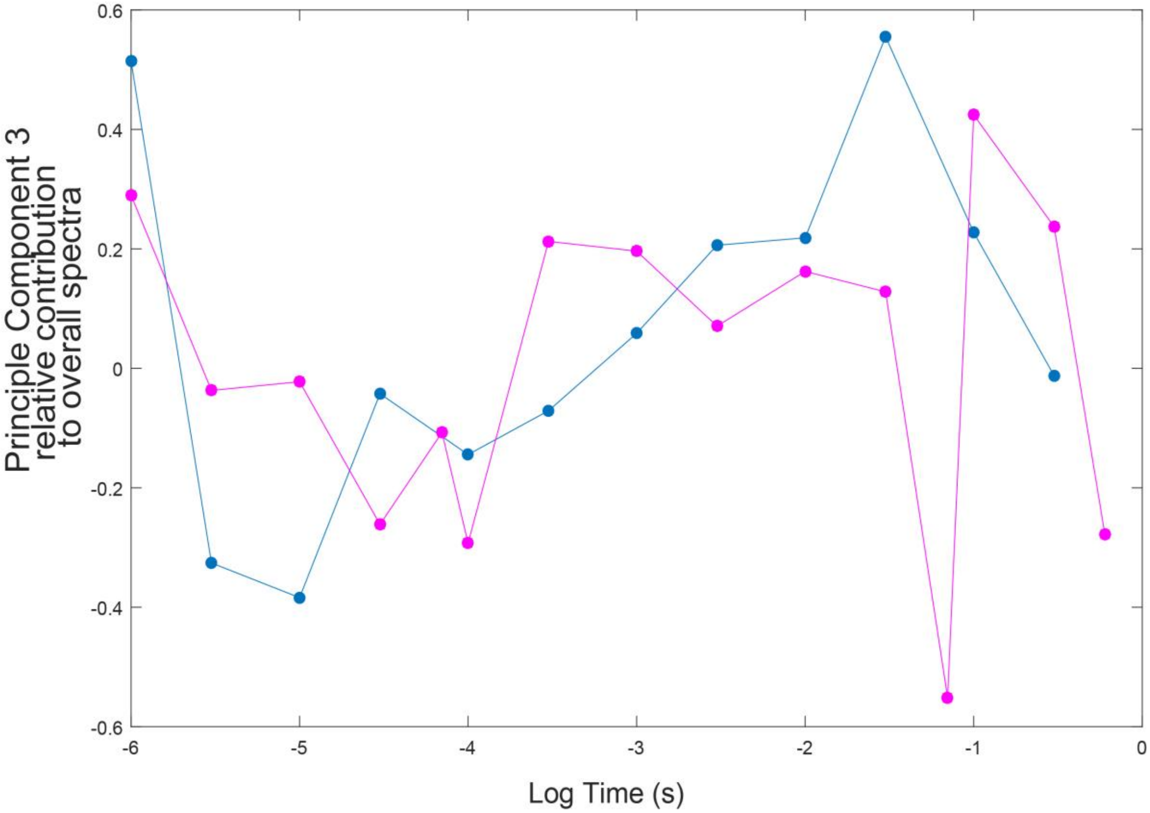


Figure 6.21: Time evolution profiles of the third principle components of the near and far variants of the DADPIX triads. Pink dots and line are the time profile of the near variant while the light blue represents the time profile of the third principle component of the far variant.

observe that the 3rd principle components for both of these triads, representing the intermediate P^+A^- state, contribute to the overall observed transient difference spectra at two separate timescales, one corresponding to its formation and disappearance prior to the formation of the charge separated state , while the other represents the reformation of the intermediate CS state on a timescale slightly ahead of the lifetime of the final charge-separated state. This result conclusively demonstrates that the intermediate charge-separated is reformed and the final charge-separated state goes through the uphill-repopulation route of recombination as predicted by the engineering principles from Chapter II.

In order to provide a proper model to describe this behavior, a much more complex set of ODEs akin to those used in Chapter II and III that model all possible electron transfers in all directions, including forward and reverse, within the triad system becomes necessary. This complete set of ODEs will most likely not have simple, closed-form solutions as those for the ODEs that were sufficient to describe the dyads in the previous section. This would make the model fitting of the experimental result a computationally intractable problem and therefore is not considered further in this chapter, especially after having proven that the third principle component does kinetically represent the reformation of the P^+A^- state. Instead, for the sake of efficiency without significant loss of the accuracy

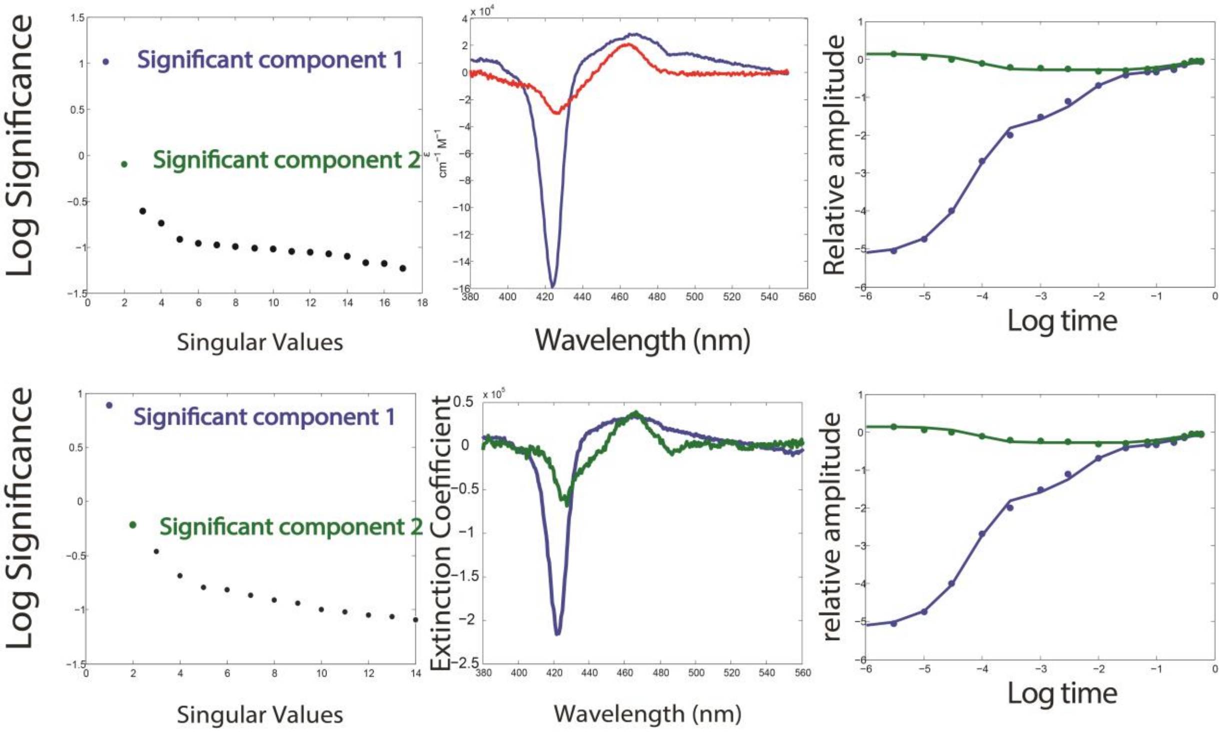


Figure 6.22: SVD-based global kinetics analysis of the far(Top row) and near(bottom row) variants of the DPA triad in maquette. Left: singular values of the principle components listed in order. Mid: reconstructed spectra of the first and second components of the triads, with molar extinction coefficients as the unit. Right: quality of fit for the two triads.

of the model, I am proposing the following two alternative models, which, although do not attempt to describe the reformation of the intermediate CS state acknowledged and proven above, can still offer an accurate model of the lifetime and yield of the charge-separated state, two metrics that have been regarded as essential since they were defined in Chapter II.

Alternative to the full model, a closer look at the driving forces and distances of this specific DPA triad reveals that we can also describe its kinetics with a simpler model given that certain assumptions are met. Moreover, as shown in the previous section, the rate of charge recombination from the P^+A^- to the ground state is 10 times slower than the forward charge-separation rate. This means that, as long as the ET between the donor and the cation of the pigment(D to P^+) is significantly faster than that of the P^{3*} to A electron transfer, we can apply the 3-state model simplification in the global analysis without loss of generality. The approximation to the full kinetics model via the simpler 3-state model combines the electron transfer reactions from P^{3*} to A and from D to P^+ as a single reaction that forms the final charge-separated state D^+PA^- directly from the triplet excited state, treating the system as identical to the dyad shown in Figure 6.12, with the D^+PA^- playing the role of the P^+A^- state.

The simplified 3-state model should be able to account for the ~60% yield of the near-variant of the DPA triad, since the analysis of the dyad ET kinetics from the previous section has demonstrated that the yield of the charge-separated state of the dyad is ~60%, as seen in Figure 6.16. However, the apparent yields of 23% of the far variant DPA triad would suggest that the 3-state model would not be applicable in this case, since the low yield can only result from a D to P^+ ET rate that's comparable or slower than that of the recombination from A^- to P^+, as it is highly unlikely that 77% of the pigment cofactor remained inactive during the ETs of the triad.

To establish a model that could account for the low yield of the far-triad without having to use the third principle component in the global analysis, I implemented a modified version of the full model, where only the top two most significant components are considered in the model fitting. This is possible because in the solution of the system of ODE, too complex to include in this chapter and shown in Appendix C instead, the expression for the population of the final charge-separated includes all of the rates in the kinetics scheme. That is, the information contained in the time-evolution of the excited state as well as the final charge-separated would allow us to implicitly figure out the rates that are ostensibly not directly related to the states. This modified implementation of the full kinetics model for the triad has been highly effective in the analysis of the triad kinetics data.

Figure 6.22 demonstrates the application of the SVD-based global analysis using the modified full kinetics model on the transient absorptions data of the near- and far-variants of the triad. It is noticeable here that the third principle components, as seen by the first non-colored dot in the left two panels of Figure 6.21 are much closer to the first two principle components than the third principle components of the monad and dyad described in previous sections. This increased significance of the third principle component indicates that the triad indeed displays a charge-separating behavior, namely a set of electron transfer reactions, that is fundamentally different from that of the monad and dyad, proving that the incorporation of the donor has successfully effected an evolutionary development in the engineering of photosystems in maquettes, opening the door of multi-ad engineering to maquette-based photosystems.

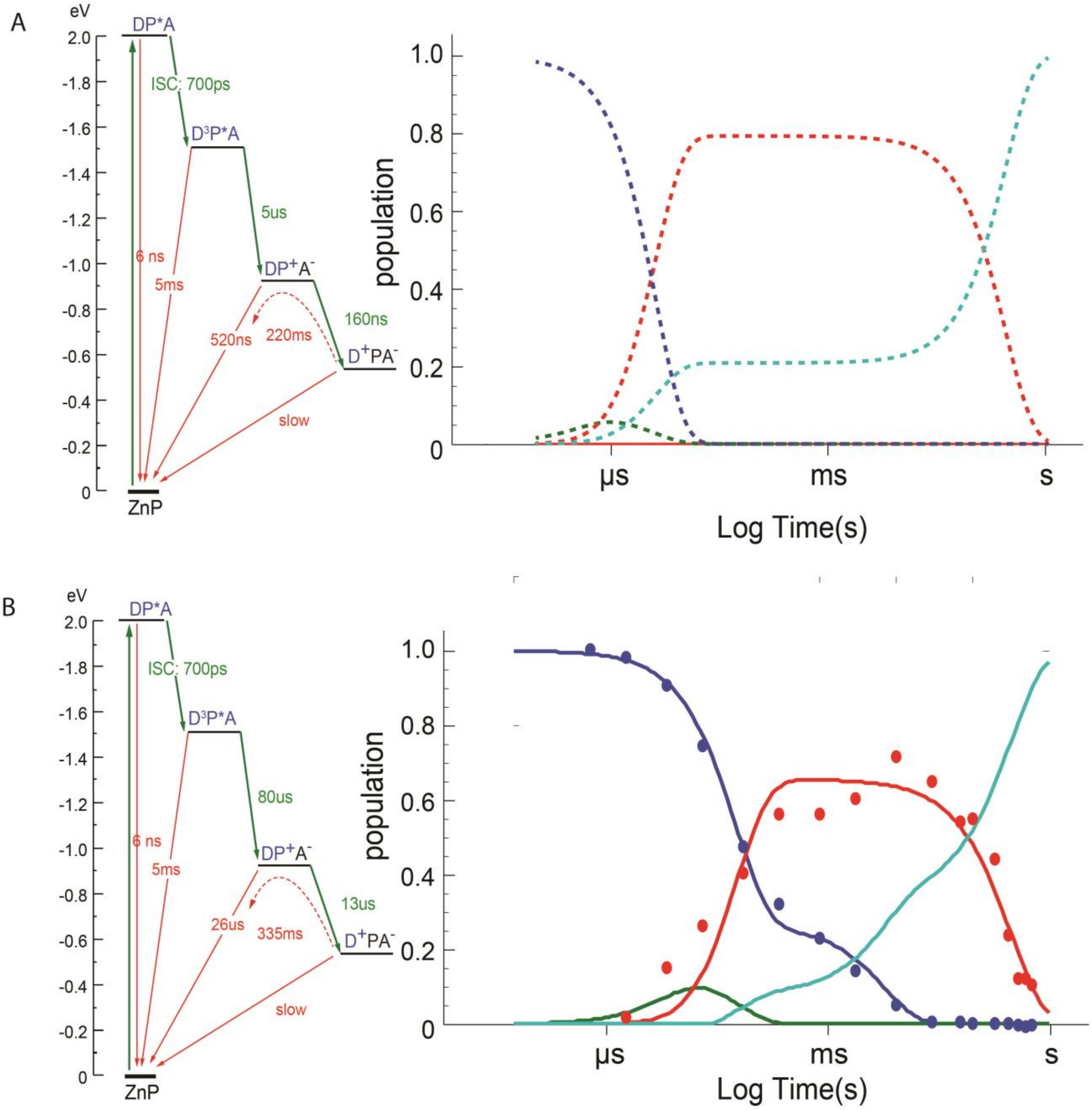


Figure 6.23: Experimental and theoretical ET kinetics and charge-separation performances of the near variant of the DPA triad. A: theoretically expected electron transfer kinetics of the near DPA triad. Rates are computed using a 12A P-A distance, 7A D-P distance, redox potentials as shown in the energy diagram on the left. Reorganization energy of 1.5eV is used to obtain the kinetics. B: experimentally determined ET kinetics of the near-variant of the DPA triad. Dots in the kinetics profile represent transformed vectors of the V^T matrix, representing the population of the specific state in the model.

Figure 6.23 and 6.24 illustrate the experimentally determined rates of the near and far variants of the DADPIX-based triads in the styles of trikes, familiarized in Chapters II and III in comparison to the theoretical counterparts. The rates are sufficient information to determine both the yield and the lifetimes of the charge-separated states. In contrast to the dyads, however, the triads involve a more significant level of deviation from the theoretical model. While the reason behind this increased discrepancy between the model and the experimental values will be examined in the discussion session, it is important to note here that rather than focusing on the differences in the numerical values of the rates, it is more meaningful to notice the similarity in the relationship between the three rates at the branching point of the P^+A^- state between the model and the experimental results for the near and the far triads. In the near triad, the rate of the recombination from the acceptor to the pigment is slower than the ET rate from the reduced donor to the radical P^+ in both the model and the experimental rates, and the ratios of these two rates in the two sets of kinetics are rather similar. This similarity explains for the agreements in yields of the theoretical model and the experimental data. The same similarity is also observed with the ratios of these two rates in the model and experimental kinetics of the far triad. Together they suggest that the fitted kinetics models obtained from the SVD analysis has managed to capture the characteristic rates of the underlying true electron transfer events occurring in the triad, although the specific routes taken by the electrons might differ.

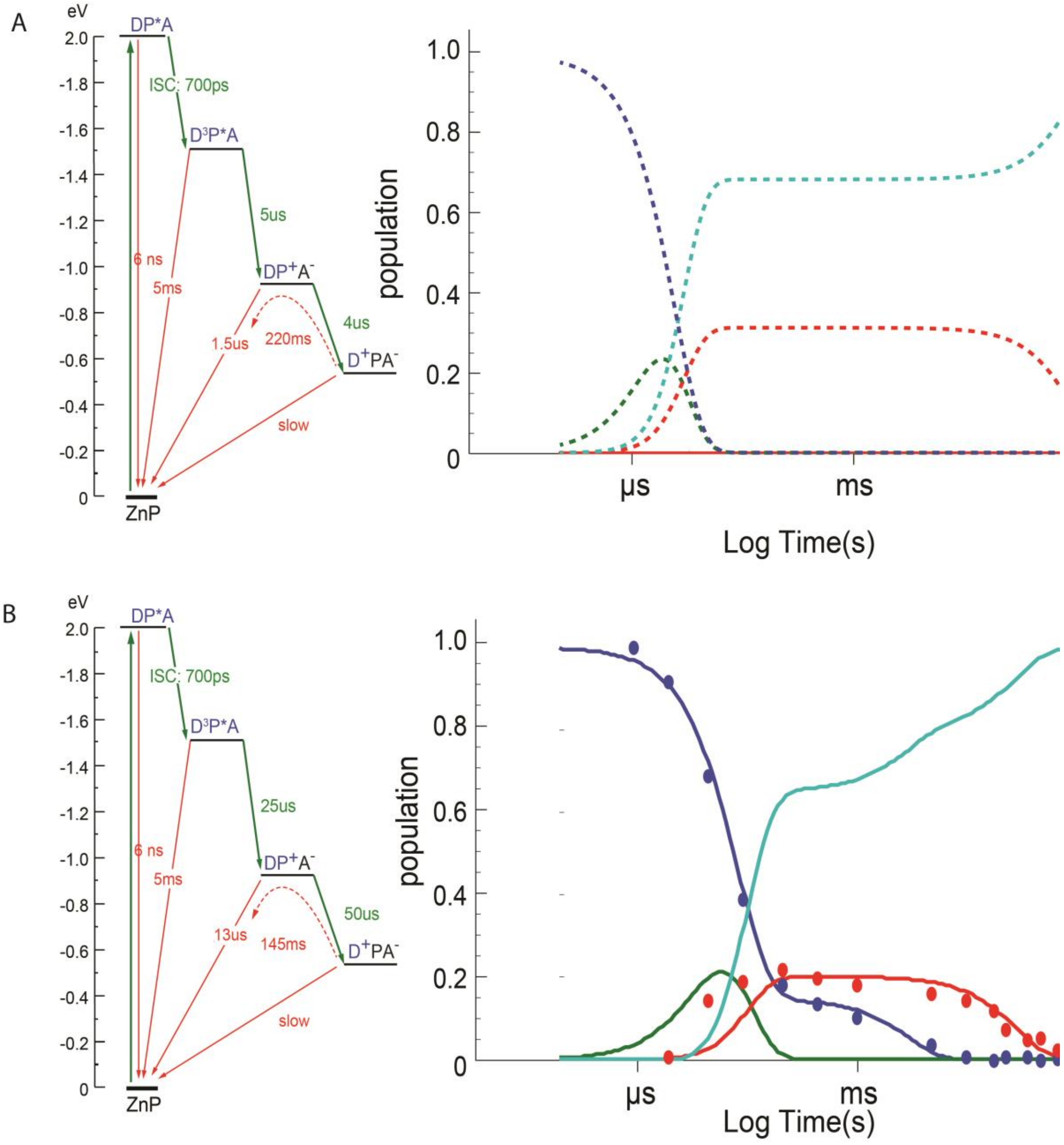


Figure 6.24: Electron transfer kinetics of the far-DPA triad and its yield and lifetime, determined by both SVD-based model fitting and theoretical model using predictive analysis. A: theoretically expected electron transfer kinetics of the far DPA triad. Rates are computed using a 12A P-A distance, 10A D-P distance, redox potentials as shown in the energy diagram on the left. Reorganization energy of 1.5eV is used to obtain the kinetics. B: experimentally determined ET kinetics of the near-variant of the DPA triad. Dots in the kinetics profile represent transformed vectors of the V^T matrix, representing the population of the specific state in the model.

6.4: Discussion

6.4.1 Agreements and discrepancies between the model and empirical kinetics and their implications.

In the results sections we have compared the experimentally determined ET kinetics of the P-A dyad as well as both versions of the DPA triads with their theoretical model counterparts. It was interesting to discover that the experimental and theoretical kinetics of the dyad have a significantly smaller amount of disagreements than that of the two versions of the triads. While the discrepancies of the dyad's rates are less than an order of magnitude and can therefore be considered as experimental/statistical error, the deviation seen in the two triads are much greater and likely reflect differences between the Moser-dutton ruler based theoretical model and the SVD-based analytical algorithm's model.

Here we offer a few possible explanations.

First of all, the presence of a third cofactor in the triad greatly increases the complexity of the system, and this increase in complexity, as already discussed in a previous section above, cannot be properly reflected in the kinetics model used in the SVD-based analysis, for the sake of practicality. In the case of linearly arranged triads, moreover, it is demonstrated with emphasis in Chapter II and III that one of the additional kinetic complexity not possible to include in the experimentally fitted model plays an important role. For example, the lack of representations of the reverse electron transfer reactions prevents the model from being able to explain the uphill-directed recombination typical of linear multi-ads. When the kinetics model being fitted to is simpler than the model used to compute the theoretically expected kinetics, it is inevitable that discrepancies in the experimentally determined rates and the model predicted rates exist.

Moreover, It is important to note that in order to generate such a predictive kinetics profile, not only are the parameters such as inter-cofactor distances and driving forces necessary, but also and more importantly, parameters that are less concrete and well-determined, specifically the reorganization energies between the components of the triad in the protein scaffold. To generate the expected/"theoretical" kinetics shown, we used the pigment-acceptor distance from the crystal structure, the pigment-donor distance from the model shown in figure 6.10, and the midpoint potentials of the various cofactors measured in the scaffold maquette using standard electrochemical procedures summarized in the methods section. For the more nebulous reorganization energy, we generated model kinetics using a few different uniform reorganization energy, as well as combinations of individual reorganization energy for each step of the reactions, and ended up choosing a uniform reorganization energy of 1.3eV as the number that simultaneously produces the best fit to the experimental kinetics and best describe the high-dielectric environment of the small protein interior of the maquette.

6.4.2 The positives and negatives of the triplet excited state in practical charge-separation engineering.

Whereas our design blueprints detailed in Chapter II and III have demonstrated the advantages of using the singlet excited state of the pigment cofactor for the overall robustness and stability of the photosystem and henceforth prescribed all optimal engineering with only the singlet excited state, the practical setting of the proof-of-principle triad engineering in the maquette scaffold described above, however, limits us to the choice of the triplet excited state as the distance between the pigment and acceptor slots are too far for effective formation of the initial P^+-A^- charge-separated state before the inherent intersystem crossing of the singlet excited state to the triplet excited state takes place. While the immediate reaction might be to decrease the P-A distance and engineer a singlet-driven charge-separation, a closer inspection at the specific parameters of our proof-of-principle triad reveals that the usage of triplet excited state as the main driver of charge-separation is indeed the better choice over singlet given the available redox midpoint potentials of the acceptor provided by the B-heme and DADPIX are significantly higher than prescribed by the optimal engineering blueprints. In fact, to make the driving force between the excited state and the acceptor as close to the optimal prescribed by the engineering guidelines as possible, the lower-lying triplet excited state is indeed the better choice.

In addition, the choice of triplet excited state as the starting point of charge-separation offers significant freedom and tolerance for the engineering of the remaining components of the charge-separating device, as the long lifetime of the triplet excited state, as shown above to be around 10-20ms, allows the forward electron transfer to be as slow as hundreds of microseconds, thereby relaxing the constraints on distance and on redox potentials of the acceptors.

As already discussed in Chapter II and III, the long lifetime and the ease with which the triplet excited state reacts with oxygen molecules to form reactive singlet oxygen species mark the principal drawback of using the triplet excited state for long-lived charge-separation. Indeed, natural photosystems engineered away from the triplet excited states of their pigment cofactors by decreasing the distance and driving forces between the pigment and the initial acceptors in order to promote the forward ET to generate the CS state in less than 1ns. This prevents the intersystem crossing, normally on the order of hundreds of picoseconds, from taking place. Moreover, natural photosystems committed a significant amount of the possible energy preservable in the final charge-separated state to ensure that no facile uphill thermal repopulation of the singlet or the triplet excited state could occur. All of these engineering devices were adopted by natural photosystems because they do not and can not carry out their designed functions without the presence of oxygen, nor do they possess an effective mechanism to remove oxygen local to the site of the pigment. Consequently, the presence of the triplet excited state of the pigments employed by natural photosystems represents a significant risk that must be avoided. However, in the setting of practical engineering of charge-separating devices, triplet excited state is less of a threat, at least when the immediate goal is to produce long-lasting charge-separated state and does not involve any oxygen. As already seen above, application of standard and well-reported oxygen removal techniques, ranging from purging the experimental environment with inert gases to active removal with effective oxygen scrubbing enzyme combinations that have been shown to be electron transfer inert, can remove the threat of the singlet oxygen and enables triplet excited state to be a viable engineering option. (*6*, *11*, *12*)

In the end, however, the long-lived charge-separation produced by our man-made photosystem must be applicable to practical catalysis through further engineering. The coupling of the charge-separation unit to the catalytic unit will impose additional constraints on the design of the system, and many constraints could very possibly require the presence of oxygen, thereby making the triplet excited state a undesirable option.

6.4.2: Routes to improve the current DPA triad design

6.4.2.1 Constitutive donor cofactors with the right potential.

While ferrocene successfully demonstrated its capability as an effective donor cofactor even in protein-based biological setting of charge-separation design, the steps necessary to incorporate and purify the ferrocene cofactor lead to significant decreases in the final yield of the completed triad, and the size of the maleimide linker resulted from the labeling reaction creates uncertainty in the final distance between the pigment and the donor, thereby complicating the design process. To overcome these challenges, we desire a donor cofactor that is constitutive to the maquette scaffold presents the optimal alternative, that is, an amino acid residue that is redox active and has the ability to reduce the ground state of pigment cofactors whose potentials are usually higher than 500mV. An obvious choice is indeed the aromatic residues such as tyrosine and tryptophan. However, the redox potentials of tyrosine and tryptophan in small, hydrophilic proteins are significantly higher than the potential of the ground state of the pigment cofactor we currently possess(*10*, *13*). Therefore, the utilization of tyrosine/tryptophan as the constitutive donor cofactor, as is the case in natural PSII, requires the redox potential of the pigment cofactor, when incorporated into the maquette scaffold, be increased significantly, by at least 200mV. This can be achieved by modifying the P slot's local environment in terms of its charge distribution and hydrophobicity(*14*), or customizing the chemical substituents on the pigment's tetrapyrrole ring(*15*) or the environment of the protein(*3*, *16*). Additionally, artificially increasing the pH experienced by the maquette, either by modifying the charge distribution of the donor slot or by subjecting the entire maquette to a solution whose pH is high, changes the effective midpoint potential of the tyrosine/tryptophan residue, with higher pH resulting in lower potentials.(*13*)

Alternatively, since unnatural amino acids have already been successfully incorporated into maquettes, constitutive incorporation of cofactors in the form of unnatural amino acid residues is yet another possibility to consider. Existing amino acid incorporation schemes require specific pairs of unnatural amino acids and their corresponding t-RNA synthase, product of directed evolution.(*17*)

6.4.3.2 Acceptor with appropriate energetics to allow for singlet-driven charge separation

While the triplet excited state offered a large range of acceptor and donor potentials and distance to be viable for successful production of long-lived charge-separation, the extreme danger to damage it imposes upon the photosystem, mitigated only by operating the photosystem in a highly impractical fully anaerobic environment. In order to engineer practically functional and applicable photosystems for efficient and long-lasting charge-separation at large scale, however, it is necessary to follow the design blueprints specified by Chapter III. Namely, a singlet driven charge-separating device with close energy gaps between the singlet excited state of the pigment and the subsequent acceptors is indeed the ultimate goal of design. To prevent the formation of the triplet state via intersystem crossing as have seen in the experimental results presented above, faster forward electron transfers than the intersystem crossing rate of the pigment must be designed. As we have seen in Chapter II and III, such fast forward electron transfers, when designing against a generic protein environment, require the driving force of the initial electron transfer to be sufficiently small, with the midpoint potential of the primary acceptor almost identical to that of the excited singlet state of the pigment. The secondary acceptor's, or in the case of a donor-containing photosystem, the donor's redox potential must also be either very close to that of the primary acceptor or of the ground state of the pigment.

In order to realize the prescribed singlet driven designs, it is of utmost importance to have the ability to customize the cofactors so that their redox midpoint potentials can be adjusted at will. Many studies have reported modulations of redox potentials of cofactors based on altering either the chemical structure of the cofactors or their environments. However, cofactors whose midpoint potentials as low as -700mV and are stable within the interior of small water-soluble proteins have not been reported. The need to customize the potentials of cofactors at will be further discussed in Chapter VII.

6.5: Conclusion

With its high affinity and specificity for the iron-porphyrin as the acceptor, the Zn-porphyrin as the pigment, and its high resolution crystal structure that allows for the site-specific engineering and incorporation of a ferrocene as the donor via a cys-maleimide reaction, the new generation of maquette has provided the perfect platform upon which the first proof-of-principle donor-pigment-acceptor(DPA) triad is successfully constructed, producing a stable charge-separated state between the donor and the acceptor that lasts over 300ms. This represents the longest lifetime achieved by any charge-separating triads designed either in biological and synthetic chemical contexts in the world, to the best of our knowledge. Moreover, high resolution structure also allowed for practical examination of the engineering principles formulated in the first half of this thesis. With the successful experimental observation of the reformation of the intermediate CS state, the engineering principle that linearly arranged multi-ads induce uphill-repopulation based recombination over direct recombination is demonstrated. Additionally, the structural information of the maquette also allowed us to vary the donor-pigment distance and observe a corresponding change in the yield of the CS state as expected by the theoretical model. Together the results of this chapter represent a major evolutionary step in the de-novo engineering of multiads following the prescribed engineering guidelines from Chapter II. They proved that given the necessary binding sites, distances, and proper cofactors, multiads can be engineered and will produce performances as predicted by the engineering blueprints. The successes of the theoretical engineering principles in practice naturally lead to the final chapter of this thesis, the full application of the optimal engineering prescriptions based on both the fundamental theories of electron and biological constraints in man-made water-soluble proteins.

6.6 References

1. T. A. Farid *et al.*, Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).
2. B. R. Lichtenstein *et al.*, Engineering Oxidoreductases: Maquette Proteins Designed from Scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).
3. L. A. Solomon, G. Kodali, C. C. Moser, Engineering the Assembly of Heme Cofactors in Man-made Proteins. *J Am Chem Soc.* **136**, 3192–3199 (2014).
4. J. C. D. Brand, W. Snedden, Electron-transfer Spectra of Ferrocene. *Transactions of the Faraday Society*. **53**, 894–900 (1957).
5. J. R. Pladziewicz, Electron Transfer Reactions of Ferrocene (1971), doi:10.2172/4772898.
6. C. E. Aitken, R. A. Marshall, J. D. Puglisi, An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments. *Biophysical Journal*. **94**, 1826 (2008).
7. E. R. Henry, J. Hofrichter, [8] Singular value decomposition: Application to analysis of experimental data. *Methods in enzymology* (1992).
8. B. R. Lichtenstein *et al.*, Designing Light-Activated Charge-Separating Proteins with a Naphthoquinone Amino Acid. *Angewandte Chemie*. **127**, 13830–13833 (2015).
9. J. J. Warren, J. R. Winkler, H. B. Gray, Redox Properties of Tyrosine and Related Molecules. *FEBS Letters*. **586**, 596–602 (2012).
10. Cecilia Tommos, Jack J Skalicky, Denis L Pilloud, A. A Joshua Wand, P. L. Dutton, De Novo Proteins as Models of Radical Enzymes †. *Biochemistry*. **38**, 9495–9507 (1999).
11. V. Linek, P. Beneš, J. Sinkule, O. Holeček, V. Malý, Oxidation of D-Clucose in the Presence of Glucose Oxidase and Catalase. *Biotechnology and Bioengineering*. **22**, 2515–2527 (1980).
12. 4414334 Oxygen Scavenging with Enzymes: Donald O Hitzman Assigned to Phillips Petroleum Company. *Biotechnology Advances*. **1**, 330 (1983).
13. S. D. Glover *et al.*, Photochemical Tyrosine Oxidation in the Structurally Well-Defined α 3 Y Protein: Proton-Coupled Electron Transfer and a Long-Lived Tyrosine Radical. *J. Am. Chem. Soc.* **136**, 14039 (2014).
14. H. Ishikita, W. Saenger, J. Biesiadka, B. Loll, E. W. Knapp, How Photosynthetic Reaction Centers Control Oxidation Power in Chlorophyll Pairs P680, P700, and P870. *PNAS*. **103**, 9855 (2006).
15. P. J. Dandliker *et al.*, Dendritic Porphyrins: Modulating Redox Potentials of Electroactive Chromophores with Pendant Multifunctionality. *Angewandte Chemie International Edition in English*. **33**, 1739 (1994).
16. C. Olea Jr, J. Kuriyan, M. A. Marletta, Modulating Heme Redox Potential Through Protein-Induced Porphyrin Distortion (2010), doi:10.2210/pdb3nvr/pdb.
17. A. Deiters *et al.*, Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. *J. Am. Chem. Soc.* **125**, 11782 (2003).

Chapter VII: Application of the updated optimal engineering guidelines to the construction of tetrads in a new generation of maquettes

7.1 Introduction:

Having formulated a set of comprehensive engineering blueprints for the optimal charge-separating photosystems with regards to both the physics of electron transfer and the biochemistry and biophysics of protection against damage and having demonstrated the applicability of constructing multi-ads in man-made proteins, I now finish this thesis by laying out the specifics of the necessary engineering devices to construct the optimal tetrad that provides the same level of performance of charge-separation as natural photosystems but is freed from the overwhelming structural complexity characteristic of natural photossytems. This tetrad represents the culmination of the theoretical analysis of part I and the experimental results of part II of my thesis, as it combines the advances achieved in both parts and produces something novel and at the same time reasonably realistic: a photosystem that satisfies all of the engineering constraints that made natural reaction centers the way they are but is housed within a man-made 4-helical bundle that’s already available and have been shown to be capable of supporting engineering of multi-ad photosystems.

7.2: Prescribing an optimal tetrad meeting all practical constraints in the new maquette

We have observed that when the revised optimal engineering blueprints are duly followed, adopting every engineering device discovered in Chapter III for the purpose of damage protection, the resulting catalytic quartet does not differ much from known natural photosystems, as already shown in Figure 3.18. However, as shown throughout Chapter VI, even the most advanced generation of maquettes, which have been the platform of the successful >300ms charge-separation, will have difficulty in binding cofactors following the curved geometry prescribed by Figure 3.18. Therefore, it is important to examine the effects on optimal engineering of photosystems when we relax some of the constraints and remove them from the model. The results of Chapter III have also shown that, as seen in Figure 3.19, the incorporation of the closely spaced cofactor chain engineering device without enforcing the curved design allows the engineering of

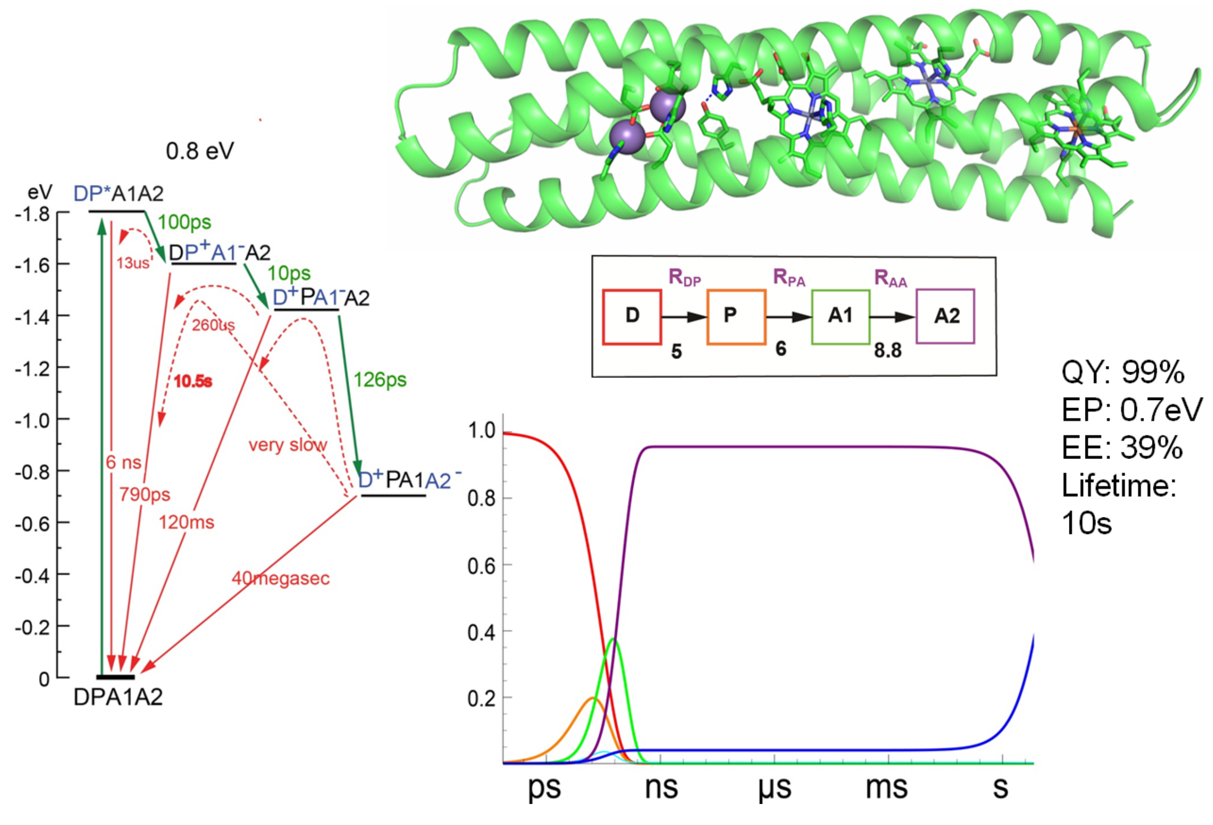


Figure 7.1: Optimal tetrad in the new generation of maquettes using practically available cofactors and following the revised engineering blueprints from Chapter II as much as possible. Top: model structure of the tetrad in the maquette, with ZnP as the pigment, ZnPPIX as the primary acceptor, DADPIX as the final acceptor, and tyrosine/ferrocene as the donor. Left: Energy diagram of the DPAA tetrad, shown here with ZnPPIX as the additional acceptor with a guessed driving force ΔG_{PA} of -150mV, while all else remain the same as the DPA triad in the previous chapter.

high performance charge-separating tetrads under reorganization energy as high as 1.5eV, essentially removing the environment surrounding the cofactors as an engineering consideration. This is a significant deviation from the engineering blueprints formulated in chapter III, where protection of the photosystem against damages holds equal importance as maximizing the performance of the photosystem. Nevertheless, the second half of my thesis has also demonstrated that the maquette platform upon which we carry out our design of photosystems allows for complete control over the environments of the electron transfer. Consequently, we have the luxury of relaxing or even intentionally ignoring some of the engineering requirements, with the knowledge that the effects of such requirements can be achieved via direct control of the ET environment. For example, the requirement to restrict the lifetime of the intermediate CS states to <10ns can be relaxed when the photosystem is engineered in an oxygen free environment, a measure that's easily achievable when the platform of the photosystem design is a small, water-soluble protein rather than a cumbersome transmembrane complex.(*1*, *2*) Rather than diminishing the practical merit of the photosystem engineered, this example perfectly demonstrates the power and usefulness of the complete understanding of the engineering principles and constraints for photosystem we established throughout this chapter. That is, when we are in full control of the engineering environment, the engineering approaches that were shown to be essential can be selectively relaxed in order to satisfy more pressing practical design constraints, in this case the linear geometry of the core of the four helical bundles.

7.3: Engineering the maquette towards the optimal tetrad.

Figure 7.1 demonstrates the optimal engineering blueprints for the tetrad to be constructed in the new generation of maquettes with the high resolution crystal structure described in Chapter VI. This design is the product of the theoretical and experimental advances made throughout this thesis, and represents a concluding remark offered by my thesis. It demonstrates that the optimal engineering demanded by the revised prescriptive algorithm in Chapter III can be satisfied without the need for another generation of maquettes but rather reengineering of the current generation of maquettes that have been shown to be highly compatible with charge-separating multi-ads.

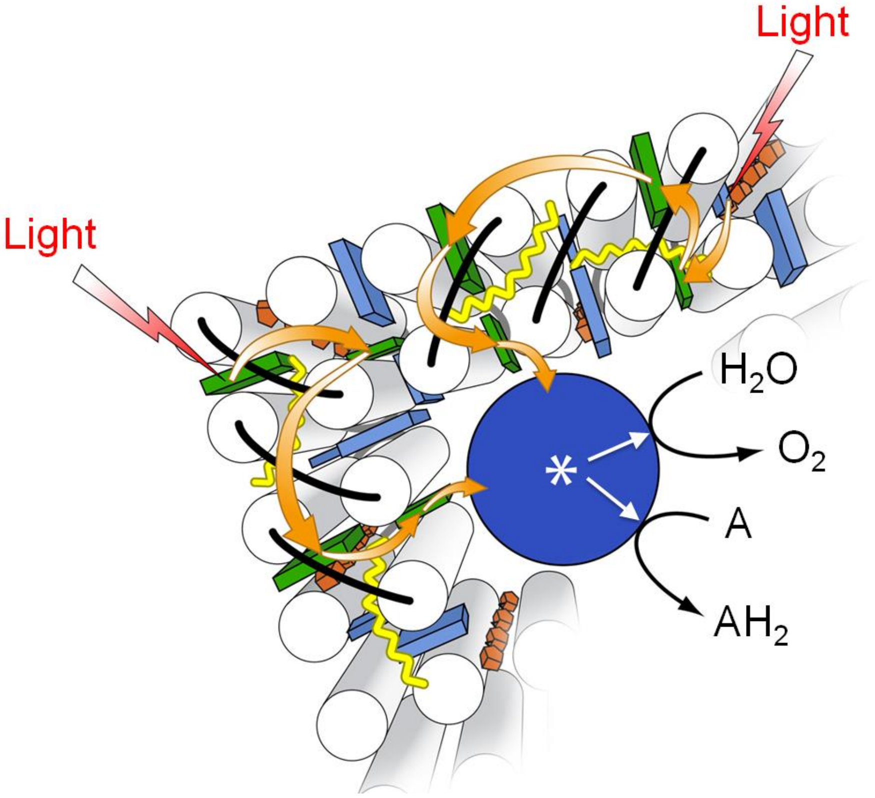


Figure 7.2: Multi-helical bundles to provide curved geometry for maquettes. The need to have curved geometries in maquette-based photosystems can be achieved by bundles made with more than 4 helices.

To practically realize this design, it demands a series of reengineering of both the maquette scaffold and the cofactors participating in the electron transfer and charge-separation. The additional engineering needed on the maquette is realistic and easily achievable. Indeed, the current maquette scaffold design already boasts three functionally independent slots, optimized for their respective roles. The tetrad design would necessitate the inclusion of an additional acceptor slot and the repurposing of one of the existing slots from the donor role to the role of a pigment. The model shown in figure 7.1 is created based on the crystal structure of the maquette as shown in figure 6.3 without any extension of the helices. This therefore suggests that the current maquette scaffold's dimension allows for incorporation of four cofactors, with 3 tetrapyroles forming the P-A-A chain while a ferrocene plays the role of the donor. This also suggests that the engineering would focus mainly on adjusting the packing of the hydrophobic core of the maquette to prevent steric clashes when the space originally intended for 2 tetrapyroles is appropriated for 3.

The necessary engineering required for the cofactor, as has always been the case throughout this thesis, represents a much greater challenge. This design essentially adds another acceptor cofactor in between the P-A in the DPA triad featured in Chapter VI. Given the limited available off-the-shelf, we have chosen Zn-protoporphyrin IX, a standard light-activatable cofactor that has a radical state whose potential is expected to be slightly below that of the ZnP used in the previous chapter.(*3*) This allows ZnPPIX to act as an intermediate acceptor before giving the electron off to the final acceptor, still played by DADPIX. However, additional experiments are necessary to determine the redox potential of its negative radical state to ensure that it can properly serve as the intended intermediate acceptor. Photophysical experiments are also required to ensure that laser excitation will preferentially target the ZnP rather than ZnPPIX. The choice of ZnPPIX represents a suboptimal compromise and exposes the lack of ability to customize cofactors with desired redox potentials as the greatest practical constraint on the realization of the optimal engineering blueprints, and consequently, the central focus of future research efforts if construction of highly efficient photosystem in small and water-soluble man-made protein is to become a streamlined engineering commonality.

7.4: Looking even further ahead: the greatest unresolved challenges

This chapter has illustrated that with this new generation of maquettes, not only is the breakthrough from dyads to triads possible, as already shown in Chapter VI, but also, given the necessary engineering infrastructures, the next leap forward to constructing a tetrad that’s not far off from the optimal engineering blueprints for tetrads formulated in Chapter III based on a combination of electron transfer theories and practical protection against damage. However, throughout the experimental half of my thesis, a recurring and unrelenting challenge has always been the highly limited range of potentials available from the cofactors that can be utilized. This challenge has not been properly addressed even with the application of chemical modifications of maquettes via either click-chemistry or maleimide-cysteine reaction, since the availability of cofactors whose midpoint potentials meet the engineering blueprints did not improve. In order to unlock the full potential of the optimal engineering blueprints of photosystems in small and water-soluble protein platforms, the ability to customize the redox potentials of cofactors either chemically(directly modifying the chemical structure of cofactor) or biochemically(modifying the environment of the cofactors). This involves formulation of another set of engineering blueprints, in this case for the redox midpoint potentials of cofactors. Because of the ease with which small molecules could be represented by a rather large and comprehensive collection of features representative of their unique chemical properties, as frequently used by the pharmaceutical industries in the process of drug-discovery, the redox potentials of cofactors can be engineered via statistical predictive modeling, exploiting recent advances in machine learning algorithms. This can either be the primary method of design or a pre-experimental sieve to narrow down possible experimental targets. (*4*)(*5*)

The second challenge, as seen in section 7.2 of this chapter, is the ability to allow maquettes to adopt non-linear geometries. This can be addressed by using maquettes that have multiple hydrophobic core regions rather than only one in the middle of the 4-helical bundle. Indeed, bundles as large as 8 helices have been constructed by the Dutton group, as shown in Figure 7.2. This larger bundle would allow the placement of cofactors no longer in a line but in a curved geometry as prescribed by the optimal design shown in Figure 3.19.

The final challenge, lying much further ahead but does represent the ideal destination of our efforts, is to allow the engineered optimal tetrad photosystem to be produced and assembled by bacteria in-vivo. While the theoretical and practical discussion and exploration is outside of the scope of this work, it is necessary, especially at the end of this thesis, to place the advances made in this work into perspective.

It warms the heart to see that what little I have done with the past five years has actually gotten us a little bit closer to something nice.

7.5: References:

1. V. Linek, P. Beneš, J. Sinkule, O. Holeček, V. Malý, Oxidation of D -glucose in the presence of glucose oxidase and catalase. *Biotechnology and Bioengineering*. **22**, 2515 (1980).
2. B. MISTRY, D. B. MIN, Reduction of Dissolved Oxygen in Model Salad Dressing by Glucose Oxidase-Catalase Dependent on pH and Temperature. *Journal of Food Science*. **57**, 196 (1992).
3. I. Hamachi *et al.*, Electron Transfer from Zn-Protoporphyrin IX to Ruthenium Ammine Attached at His63 of Reconstituted Cytochrome b562. *Chemistry Letters*, 551 (1999).
4. M. Kuhn, K. Johnson, *A Summary of Solubility Models* (2013).
5. N. T. Hansen, I. Kouskoumvekaki, F. S. Jørgensen, S. Brunak, S. Ó. Jónsdóttir, Prediction of pH-Dependent Aqueous Solubility of Druglike Molecules. *Journal of Chemical Information and Modeling*. **46**, 2601 (2006).

APPENDIX

A.1: Maquette sequences engineered during the process towards constructing the DPA triad

This section of the appendix provides the specific sequences of all the maquettes engineered and constructed throughout the experimental part of this thesis.

Gen1 Maquette:

GEIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQLGGSGSGSGG
EIWKQHEDALQKFEDALNQFEDLKQL

This maquette is featured in Chapter IV and V as the Gen1 maquette. The basic single-chain maquette that provided the fundamental platform upon which a series of engineering devices are applied to.

Gen2 Maquette: Zn+C

FSASALAKHHHHHHGSSGEFGDGENLYFQG
EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
ECIACHEDALQKFEEALNQHEEALNQFEDLKQL

This maquette is the second generation of the single-chain maquette, after it has been extended by one heptad-repeat i the middle and given a single his binding site for zinc-porphyrin and a CXXCH motif for the in-vivo maturation of c-heme. A periplasmic export tag was also added to the front of the sequence.

Gen2 Maquette: Zn+C+PPY

FSASALAKHHHHHHGSSGEFGDGENLYFQG
EIWKQ(Amber)EDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
EIWKQHEDALQKFEEALNQAEEALNQFEDLKQL GGSGSGSGG
ECIACHEDALQKFEEALNQHEEALNQFEDLKQL

This maquette is the second generation of the single-chain maquette, modified to contain an amber stop condon for the incorporation of the propargyl tyrosine.

SM Maquette: structural maquettte:

MGKGGHHHHHHGGDGENLYFQG
SPELRQEHQQLAQEFQQLLQEIQQLGRELLKGELQGIKQLREAS
EKARNPEKKSVLQKILEDEEKHIELLETLQQTGQEAQQLLQELQQTGQELWQL
GGSGG
PELRQKHQQLAQKIQQLLQKHQQLGAKILEDEEKHIELLETIL
GGSGG
DELRELLKGELQGIKQYRELQQLGQKAQQLVQKLQQTGQKLWQLG

This is the new generation of maquette engineered by Nathan Ennist that has provided the high resolution crystal structure shown in Chapter and became the basis upon which the DPA triad has been successfully engineered.

SM Maquette Y168C (far) variant.

MGKGGHHHHHHGGDGENLYFQG
SPELRQEHQQLAQEFQQLLQEIQQLGRELLKGELQGIKQLREAS
EKARNPEKKSVLQKILEDEEKHIELLETLQQTGQEAQQLLQELQQTGQELWQL
GGSGG
PELRQKHQQLAQKIQQLLQKHQQLGAKILEDEEKHIELLETIL
GGSGG
DELRELLKGELQGIKQCRELQQLGQKAQQLVQKLQQTGQKLWQLG

This is the mutant of the SM maquette that allows for the assembly of the far-variant of the DPA triad in Chapter VI.

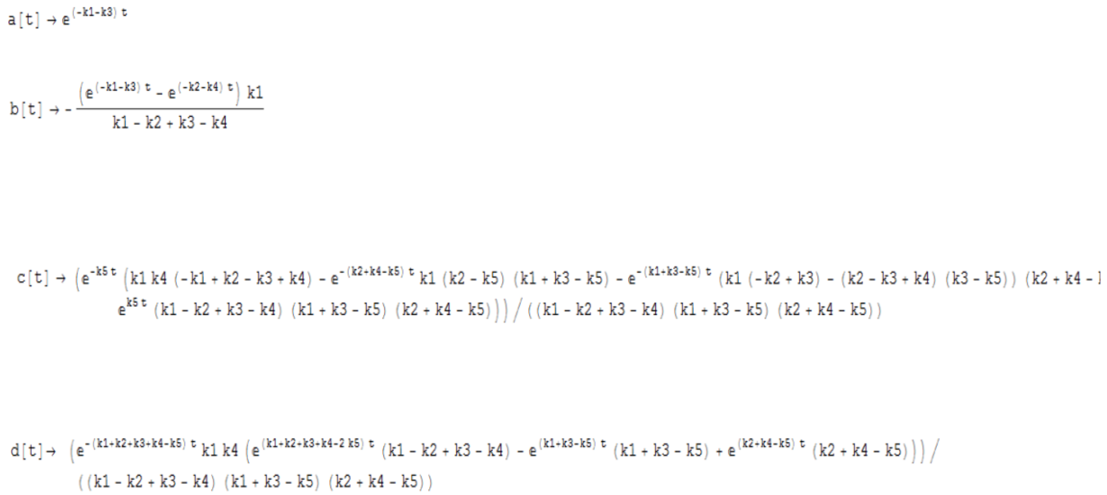
SM Maquette G164C (near) variant.

MGKGGHHHHHHGGDGENLYFQG
SPELRQEHQQLAQEFQQLLQEIQQLGRELLKGELQGIKQLREAS
EKARNPEKKSVLQKILEDEEKHIELLETLQQTGQEAQQLLQELQQTGQELWQL
GGSGG
PELRQKHQQLAQKIQQLLQKHQQLGAKILEDEEKHIELLETIL
GGSGG
DELRELLKGELQCIKQLRELQQLGQKAQQLVQKLQQTGQKLWQLG

This is the mutant of the SM maquette that allows for the assembly of the near-variant of the DPA triad in Chapter VI.

A.2: Mathematical details.

The solution to the system of ordinary differential equations for the kinetics model used in the SVD-based model fitting, representing a simplified kinetics model for the DPA triads, is included here as Figure A.2



$$a[t] \to e^{(-k1-k3)\,t}$$

$$b[t] \to -\frac{\left(e^{(-k1-k3)\,t} - e^{(-k2-k4)\,t}\right) k1}{k1 - k2 + k3 - k4}$$

$$c[t] \to \Big(e^{-k5\,t} \big(k1\,k4\,(-k1 + k2 - k3 + k4) - e^{-(k2+k4-k5)\,t}\,k1\,(k2 - k5)\,(k1 + k3 - k5) - e^{-(k1+k3-k5)\,t}\,(k1\,(-k2 + k3) - (k2 - k3 + k4)\,(k3 - k5))\,(k2 + k4 -$$
$$e^{k5\,t}\,(k1 - k2 + k3 - k4)\,(k1 + k3 - k5)\,(k2 + k4 - k5)\big)\Big) \Big/ \big((k1 - k2 + k3 - k4)\,(k1 + k3 - k5)\,(k2 + k4 - k5)\big)$$

$$d[t] \to \Big(e^{-(k1+k2+k3+k4-k5)\,t}\,k1\,k4\,\big(e^{(k1+k2+k3+k4-2\,k5)\,t}\,(k1 - k2 + k3 - k4) - e^{(k1+k3-k5)\,t}\,(k1 + k3 - k5) + e^{(k2+k4-k5)\,t}\,(k2 + k4 - k5)\big)\Big) \Big/ \big((k1 - k2 + k3 - k4)\,(k1 + k3 - k5)\,(k2 + k4 - k5)\big)$$

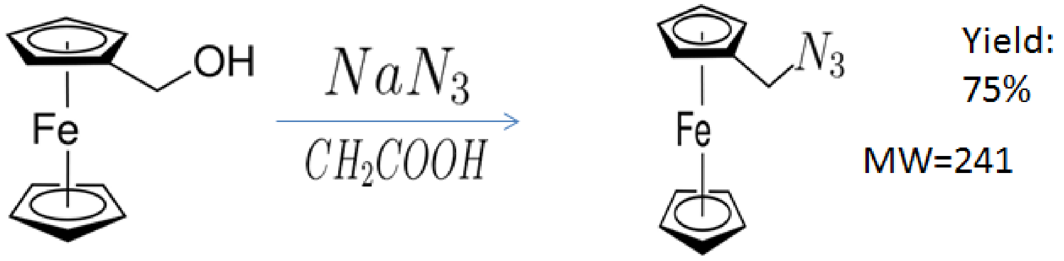
Figure A.1: Solutions to the system of ordinary differential equations for the kinetics model of the DPA triad in Chapter III. The solution is obtained using Mathematica 10.0.

A.3: Common experimental methods

Here are the experimental methods that represent general biophysical and biochemical techniques utilized throughout the experimental parts of this thesis, rather than the specific techniques featured in the methods sections of Chapter V and VI.

Cofactors:

DADPIX and B-heme used in this thesis were purchased from Frontier Scientific. The zinc porphyrin used as



Scheme A1: synthesis of azido-ferrocene from methyl hydroxyl ferrocene.

pigment for all the experiments were synthesized and kindly provided to me by Dr. Tatianna Esipova from the Vinogradov group. Both hemes and Zn-Porphyrin are dissolved in DMSO to make 10mM stock solutions, which are then frozen at -20C^{o}. Concentrations are determined via UV-vis spectroscopy and comparison to standard molar extinction coefficients.

Ferrocene came in two forms: N-ferrocenyl-maleimide and methyl-hydroxyl ferrocene. Ferrocene-carboxylic acid is used to synthesize Azidoferrocene using scheme A1.

All non-porphyrin chemicals used in this thesis were purchased purchased from Sigma.

Proteins/Maquettes:

Genes encoding maquettes with an N-terminal His-tag and TEV cleavage sequence was purchased from DNA2.0 in a pJexpress414 vector with codons optimized for expression in *E. coli*.

The protein was expressed as a thioredoxin fusion with a His-tag in E. coli BL21 (DE3) RIL cells (Stratagene) for 5 hours at 37˚C, after induction with 0.5 mM IPTG. The cells were harvested by centrifugation, resuspended in KH2PO4 buffer with 1% OTG, and lysed by sonication with a micro-tip attachment. Lysate was centrifuged at 25,000xg for 25 minutes, with supernatant applied to a NiNTA superflow resin (Qiagen) on an Akta FPLC. The fusion protein was cleaved by recombinant tobacco etch virus N1a protease overnight, and final purification was via Waters reverse-phase HPLC. Molecular weight was assayed by either MALDI or ESI mass spectrometry.

Cloning and Mutagenesis:

Primers for cloning the Gen2 maquette out of its original plasmid and for general mutagenesis were purchased from Invitrogen. Gen2 maquette DNA were amplified using primers containing appropriate restrictive digest sites for the new plasmid to which it was intended to be cloned. The PCR product were examined using Gel-electrophoresis and the proper band was excised and purified using standard purification techniques, using Invitrogen kits. The fragment is then digested with the proper enzymes and mixed with predigested new plasmids. Ligation reaction was then run on this mixture to obtain the new plasmid.

For mutagenesis, wildtype plasmids were PCR amplified with primers containing the desired mutations. The mixture is then treated with restriction enzyme DPN1 to digest away plasmids that had been directly purified from bacteria and would contain methylations. The remaining DNA was then transformed into BL21(DE3) or DH5α and followed by sequencing analysis to examine whether the correct mutagenesis was obtained.

UV-vis spectroscopy:

UV-vis spectra were recorded on a Perkin–Elmer Lambda 2 spectrophotometer using quartz cells of 1.0-cm path lengths. Peptide concentrations were between 1 and 50 μM as determined spectrophotometrically using ε_{280} = 5600 $M^{-1}\cdot cm^{-1}$ for Trp. Maquettes containing cofactors are usually diluted to 4uM and examined at room temperature in aqueous buffers, usually 200mM NaCl, 20mM $NaPO_4$, 4mM Tris-HCl at pH7.4

K_d determination for cofactors:

Various Maquette's K_d are usually determined by binding titrations, with protein concentration well below the expected K_d if possible, and if not then with the concentration as low as within the limit of detection. Normally concentration of 0.5-1μM of maquettes are used. Cofactors are dissolved in the appropriate buffer or solvent, usually DMSO when not soluble in aqueous buffers, but regular PBS buffer when they are. Cofactors are added at 0.1 molar equivalent of the protein per data point. Spectrum was recorded immediately after the addition of the cofactors and after 10mins. When no changes are noticed the next equivalent is then added.

The spectra are then collected and analyzed either via SVD-based model fitting or direct fitting of the absorbance value at specific wavelength to the following binding equation:

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

CD Spectropolarimetry:

CD spectra were recorded on an Aviv Associates (Lakewood, NJ) model 62DS spectropolarimeter using rectangular quartz cells of 0.2- and 1.0-cm path lengths. Thermal control was maintained by a thermoelectric module with a Neslab Instruments (Portsmouth, NH) CFT-33 refrigerated recirculating water bath as a heat sink. Peptide concentrations for UV-CD experiments were between 5 and 10 μM as determined spectrophotometrically using ε_{280} = 5600 $M^{-1}\cdot cm^{-1}$ for Trp.

HPLC purification of labeling reactions:

Completed reaction mixtures in appropriate solvent were injected into Waters reverse-phase HPLC. Gradients starting at 25%/75% water/AcCN (v/v) and ending at 70%/20%, and any gradient within the range, can be used depending on the nature of the modification.

Spectral electrochemical redox titrations:

Experiments were carried out with a platinum working electrode, gold counter wire, and a Ag/AgCl reference electrode. A CH Instruments (Austin Texas USA) Electrochemical workstation was used with the CH Instruments interface program. In a cuvette with a 1mm path-length, 20-100μM maquette with desired cofactors bound was dissolved in standard PBS buffer described in this thesis. For the titration of hemes, standard ensemble of mediators used were: Anthra-quinone-2-sulfonate (100µM), Benzly Viologen (50µM), Methyl Viologen (50µM),Sulfanilamide (50µM), Indigo trisulfonate (50µM), Phenazine (50µM), Pyocyanin (50µM), and Hydroxy-Napthquinone (50µM). The computer software controls the adjusts and maintains the potential of the experimental environment based on the user's intention. The net current in the local environment of the working electrode is monitored, and when the current is stabilized, UV-vis spectrum of the reaction mixture was taken. This was repeated at 20mV intervals from 250mV below to 250mV above the expected midpoint potential of the cofactor. This was done multiple times (2 oxidative, and 2 reductive) to see if there was any hysteresis. The raw data was then analyzed either via SVD(outside the scope of this thesis) or simple fit to n=1 Nernst equations.

A.4: redox spectra of cofactors used in this thesis.

The redox spectra not already shown in the thesis is that of the ferrocene at the scale where its spectral features can be seen. As discussed in the main body of the text, ferrocene's absorbance is very weak in the visible range, with highest exctinction coefficient of ~100$M^{-1}cm^{-1}$. This redox difference spectrum is shown here, obtained during an electrical spectrochemical redox experiment involving ferrocene in DMF.

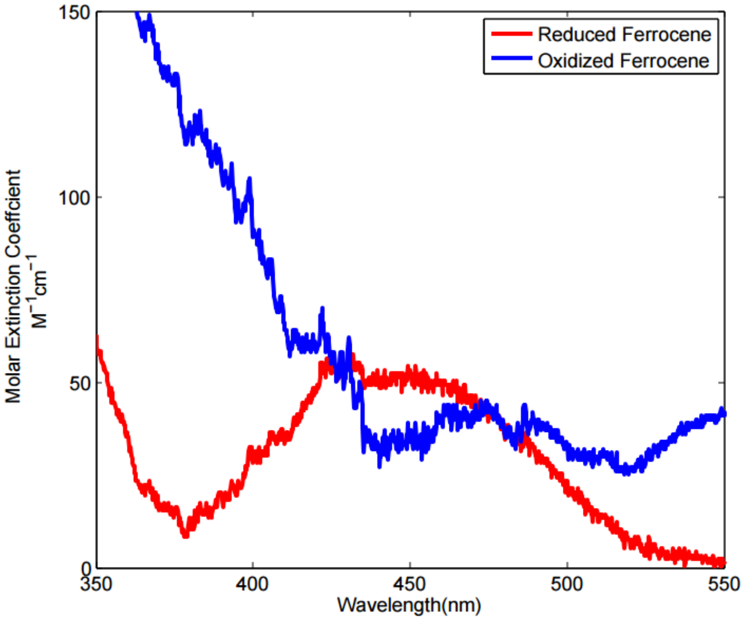


Figure A.2: Redox spectra of ferrocene.