

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

At the interface of experiment and computation: explorations of heme protein redox partner interactions, water behavior on organic surfaces and other systems

Permalink

<https://escholarship.org/uc/item/9fj2b88j>

Author

Hollingsworth, Scott Alan

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

At the interface of experiment and computation: explorations of heme protein redox
partner interactions, water behavior on organic surfaces and other systems

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Scott Alan Hollingsworth

Dissertation Committee
Chancellor's Professor Thomas L. Poulos, Chair
Professor Celia W. Goulding
Associate Professor David Mobley

2016

Chapter 2 © 2015 The Protein Society
Chapter 3 © 2016 Proceedings of the National Academy of Sciences of U.S.A.
Chapter 4 © 2015 American Chemical Society
Chapters 5 and 11 © 2016 The Protein Society
All other materials © 2016 Scott Alan Hollingsworth

DEDICATION

To my parents, Roger and Claudia Hollingsworth,
who have supported me at every turn

to my grandparents, Betty and Darold Calavan,
(who remains to this day the wisest man I have known)

to Mindy Tauberg,
who lead me back from dark shadows and continues to be the light of my life

and finally to all my family and friends,
for their continuing love, support and kindness.

Just when they think they got the answers, I change the questions
~ Roderick George Toombs

– In Memory of Martina Roeselová –

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	xi
ACKNOWLEDGMENTS	xii
CURRICULUM VITAE	xiii
ABSTRACT OF THE DISSERTATION	xviii
CHAPTER 1: An Introduction to Heme Proteins and Their Redox Partners	1
CHAPTER 2: Molecular Dynamics of the P450cam-Pdx Complex Reveals Complex Stability and Novel Interface Contacts	32
CHAPTER 3: On the affect of redox partner binding and conformational state of P450cam through the application of molecular dynamics	51
CHAPTER 4: “Bind and Crawl” Association Mechanism of Leishmania Major Peroxidase and Cytochrome c Revealed by Brownian and Molecular Dynamics Simulations	78
CHAPTER 5: Elucidating Nitric Oxide Synthase Domain Interactions B Molecular Dynamics	114
CHAPTER 6: An Introduction to Self-Assembled Monolayers	132
CHAPTER 7: On the Wetting Effects and Water Penetration of Defective Self-Assembled Monolayers	144
CHAPTER 8: Wetting Effects and Droplet Formation on Rough Hydrophobic Self-Assembled Monolayers	171
CHAPTER 9: On the Effect of Chain Length and Oxidation State on Water and Monolayer Behavior in Organic Self-Assembled Monolayers	193
CHAPTER 10: A Brief Introduction to Other Works	215
CHAPTER 11: Beyond Basins: ϕ, ψ Preferences of a Residue Depend Heavily on the ϕ, ψ Values of its Neighbors	218
CHAPTER 12: On the Docking Dynamics of the GRP1 PH-Domain Bound to a	230

Target Lipid Bilayer

CHAPTER 13: Crystal Structure of <i>Bacillus subtilis</i> Arginase at 2.077 Å	243
CHAPTER 14: Summary and Conclusions	255
BIBLIOGRAPHY	263
APPENDIX A: Detailed Comparison of the Interfacial Interactions of P450cam-Pdx	299
APPENDIX B: Time Evolution of the Major Interdomain Contacts of Nitric Oxide Synthase	300
APPENDIX C: Hydrogen Bonds Per Molecule Distribution Along the Z-Axis for Defect Free SAM Surfaces	302
APPENDIX D: Hydrogen-Bond Orientation Distributions Along the Z-Axis for Defect Free SAM Surfaces	305
APPENDIX E: Density Profiles for the Water in All Defective Systematically Varied SAM Systems	308
APPENDIX F: Conformational Dependence of Central Residue on Neighboring Conformations By Basin	313

LIST OF FIGURES

		Page
Figure 1.1	Generalized protein redox partner association	3
Figure 1.2	The cytochrome P450 catalytic mechanism	6
Figure 1.3	Rebound mechanism for the hydroxylation of substrate by compound	7
Figure 1.4	The “Pull” Effect: proposed hydrogen bonding arrangement in the oxy-state of P450cam	8
Figure 1.5	Valance electron arrangement between low and high spin state of the heme iron	9
Figure 1.6	Structure of a model P450 enzyme, P450cam	12
Figure 1.7	Comparison of the open and closed state of P450cam	13
Figure 1.8	The P450cam-Pdx crystal structure	15
Figure 1.9	Mechanism of heme peroxidase	16
Figure 1.10	Cytochrome c binding to CCP	20
Figure 1.11	Comparison of the CCP-Cytc and LmP LmCytc complexes	21
Figure 1.12	The production of nitric oxide by nitric oxide synthase	23
Figure 1.13	Proposed mechanisms for the oxidation of N ω -hydroxy-L-arginine to L-citrulline	26
Figure 1.14	Architecture of nitric oxide synthase	27
Figure 1.15	Conformational changes in nitric oxide synthase from input to output states	29
Figure 1.16	Structural overlay of the three mammalian NOS isoforms	30
Figure 2.1	Stability of the P450cam-Pdx throughout simulation	36
Figure 2.2	Rotation of Pdx in relation to P450cam	37
Figure 2.3	New contacts formed in the P450cam-Pdx complex	41

Figure 2.4	Additional contacts in the P450cam-Pdx complex	42
Figure 2.5	Distance evolution of the five newly formed MD contacts	43
Figure 2.6	Distance evolution of the Asp251 ion pairs in P450cam	44
Figure 2.7	The Arg186-Asp251 ion pair is weakened but present in the open Pdx-bound complex	45
Figure 2.8	Major contacts observed in the P450cam-Pdx complex	47
Figure 3.1	Root mean squared displacement analysis of P450cam and Pdx	56
Figure 3.2	Distance between the P450cam heme iron and Pdx iron-sulfur cluster cofactors	57
Figure 3.3	Root mean squared fluctuation analysis of P450cam and Pdx	58
Figure 3.4	Analysis of the intermolecular ion pairs of P450cam Asp251	61
Figure 3.5	Overlays of the equilibrated P450cam-Pdx complexes following 100 nanosecond molecular dynamics simulations	63
Figure 3.6	Calculated buried surface area in the P450cam-Pdx complexes	64
Figure 3.7	Conformational dynamics of P450cam in differing conformational and complex states	65
Figure 3.8	Comparisons between the interfaces of the three simulated P450cam-Pdx complexes	67
Figure 3.9	RMSD analysis of the accelerated molecular dynamics simulations of Pdx binding to the closed and open forms of P450cam	69
Figure 3.10	P450cam-Pdx complex evolution following 490 nanoseconds of aMD simulation.	70
Figure 3.11	Comparison of camphor stability on the protonation state of P450cam Asp297	75
Figure 4.1	Calculated rate constants for LmP-LmCytc	87
Figure 4.2	Calculated rate constants for CCP-Cytc	88
Figure 4.3	Experimental kinetics of LmP mutants	93

Figure 4.4	Structure of LmP in the LmP-LmCytc complex from BD simulations at 50 mM ionic strength	97
Figure 4.5	Structure of mutant LmP-LmCytc complexes from BD simulation	99
Figure 4.6	Overview of the LmP-LmCytc system	100
Figure 4.7	Further structures of mutant LmP-LmCytc complexes	100
Figure 4.8	Crystal structure of the D211N LmP mutant	101
Figure 4.9	Differences between individual and co-crystal structures of LmP and LmCytc.	103
Figure 4.10	Representative Brownian dynamics trajectory illustrating the “bind and crawl” of LmCytc at helix A of LmP	105
Figure 4.11	Molecular dynamics of the LmP Helix A centered complex	107
Figure 4.12	Structural parameters from the LmP-LmCytc molecular dynamics simulation	108
Figure 4.13	Comparison of the helix A sequence of LmP and CCP	110
Figure 4.14	Structure of the CCP-Cytc complex from BD simulations	111
Figure 4.15	Visualization of the formed CCP-Cytc complexes in BD simulations at lower contour levels	112
Figure 5.1	Nitric oxide synthase architecture	116
Figure 5.2	MD simulations of the oxy-FMN-CaM reveal complex stability	118
Figure 5.3	Distance measured between heme and FMN cofactors	119
Figure 5.4	Heat maps highlighting domain-domain interactions in full-length iNOS	121
Figure 5.5	Fit of the equilibrated bidomain model into published EM density map of the output state	122
Figure 5.6	Interfacial contacts detected between calmodulin, oxygenase domain and the FMN subdomain of iNOS	123
Figure 5.7	Overlay of FMN-Calmodulin crystal structure and proposed model Structure	126

Figure 5.8	Overview of no-CaM simulation of the iNOS oxy-FMN complex	127
Figure 5.9	Complex stability of the no-CaM complex simulation	128
Figure 6.1	Generalized siloxane self-assembled monolayer structure	134
Figure 6.2	Formation of self-assembled monolayers	135
Figure 6.3	Water at a hydrophobic SAM surface	137
Figure 6.4	Possible defect causes in a SAM surface	139
Figure 6.5	Structure of a mixed hydrophobic SAM surface	140
Figure 6.6	Structure of a hydrophilic SAM surface	142
Figure 7.1	Structure of the self-assembled monolayer system	150
Figure 7.2	Schematic of designed defective SAM systems	151
Figure 7.3	Example MD generated IR spectra from the SAM7 system	154
Figure 7.4	Deconstruction of MD-generated bulk water IR spectrum for peak assignments	154
Figure 7.5	Deconstruction of MD-generated penetrating water spectrum for peak assignment confirmation	155
Figure 7.6	ATR-IR spectra from 7-octenyltrichlorosilane SAM	157
Figure 7.7	Density profiles of each SAM system at varying water coverages	159
Figure 7.8	Hydrogen-bonds per molecule along the z-axis	161
Figure 7.9	Order parameters calculated for simulated monolayers near defect sites at varying hydration levels	164
Figure 7.10	Order parameters calculated for full simulated monolayers at varying hydration levels	166
Figure 7.11	Visual representation of SAM ordering in the presence of water penetration	167
Figure 7.12	MD produced relative population shift in the IR spectra of water in the SAM	168

Figure 8.1	Schematic of all simulated mixed SAM systems	176
Figure 8.2	Density profiles of different hydration levels for mixed SAM systems	178
Figure 8.3	SAM ordering of the terminal carbons of the C22 chains	179
Figure 8.4	Droplet formation on the 50A-C18 SAM	181
Figure 8.5	Droplet formation with buried water in the patterned 50D-C18 SAM	182
Figure 8.6	RMS roughness of SAM surfaces	182
Figure 8.7	Water uptake on smooth and rough SAM systems	184
Figure 8.8	Water penetration as a function of percentage of C18 in mixed SAMs	184
Figure 8.9	Distribution of hydrogen-bonds per molecule as a function of the z-axis	186
Figure 8.10	Hydrogen-bond orientation distributions as a function of the z-axis	187
Figure 8.11	Trends for hydrogen-bond orientation preference for each of the five hydration levels	188
Figure 8.12	Full hydrogen-bond orientation distributions as a function of the z-axis	190
Figure 9.1	Systematic variations of self-assembled monolayer surfaces	198
Figure 9.2	Density profiles of the siloxane carbons atoms for defect-free SAM surfaces	200
Figure 9.3	Characterization of the overall SAM order	201
Figure 9.4	Dependence of chain length and solvent exposed terminal group on SAM tilt angle	203
Figure 9.5	Water density profiles for defect free systems	205
Figure 9.6	Snapshots of equilibrated water on each of the four different terminal groups of a C22 SAM	206
Figure 9.7	Hydrogen bonds per molecule distribution along the z-axis for defect free SAM surfaces	207
Figure 9.8	Snapshots of water penetration for each of the four terminal groups for the C8 system	212

Figure 11.1	Variations in Ramachandran plot distributions	220
Figure 11.2	Dependence of central residue conformational on neighboring Residues	224
Figure 12.1	GRP1 PH domains bound to a target lipid bilayer	233
Figure 12.2	The EPR-derived model is stable in long molecular dynamics Simulations	235
Figure 12.3	Quantitative comparison between spin label depth parameters and membrane penetration depth in the simulation	237
Figure 12.4	Major residues of the PH domain that make contact with the lipid bilayer	238
Figure 12.5	Major residues of the PH domain that make contact with the bilayer by lipid species	239
Figure 13.1	The arginase catalytic mechanism	245
Figure 13.2	Crystals of <i>Bacillus subtilis</i> arginase	248
Figure 13.3	Crystal structure of <i>Bacillus subtilis</i> arginase	251
Figure 13.4	Primary sequence alignment of human arginase I and <i>B. subtilis</i> Arginase	252
Figure 13.5	Structural alignment of known arginase enzymes reveals strong conservation	253

LIST OF TABLES

		Page
Table 2.1	Top time averaged contact pairs over the final 25 nanoseconds	38
Table 2.2	Selected comparisons of the interactions at the interface of P450cam and Pdx	39
Table 3.1	Calculated binding free energies of the three studied P450cam-Pdx complexes	68
Table 3.2	Additional parameters for P450-cam-Pdx complex accelerated molecular dynamics simulations	76
Table 4.1	Crystallographic data and refinement statistics of LmP D211N	95
Table 4.2	Second order association rates of LmCytc and LmP from Brownian dynamics simulations initiated from the co-crystal structures	97
Table 4.3	Kinetic parameters obtained from steady state assays of LmCytc oxidation by LmP	98
Table 5.1	Major domain-domain interactions in the equilibrated oxy-FMN-CaM complex. Contacts are defined as a distance $<3.5\text{\AA}$	120
Table 9.1	Average tilt angle of varying defect-free SAM surfaces	202
Table 9.2	Average tilt angle for all studied SAM systems	204
Table 9.3	Water penetration dependence on chain length, terminal group and defect size	209
Table 11.1	Number of observations for each of the 100 possible conformational contexts of the central residues	223
Table 12.1	2D MSD of Lipid Species in the presence and absence of PH Domain binding	240
Table 13.1	Arginase data collection and refinement statistics	249

ACKNOWLEDGMENTS

First and foremost, I must thank my advisors, Drs. Thomas Poulos and Douglas Tobias for providing me with the freedom to chase down new approaches, ideas and projects from day one and acting as immeasurable sources of both inspiration and guidance. Nothing that I accomplished in my time at UC Irvine would have been possible without both of these two incredible scientists opening up their labs to me. I am also immeasurably in debt to Dr. Abraham Stern for his mentoring and guidance as I started to establish myself in both lab environments and pushed me to always aim higher.

From the Tobias Lab, I must acknowledge Dr. Matthias Heyden for his insights and guidance during his tenure at UC Irvine. I wholeheartedly thank Krista Parry, James Fields and Andrew Geragotelis for their kindness and friendship that made my time in the Tobias Group that much brighter. I also thank Dr. Theresa McEntire for her guidance and feedback throughout my SAM research as well as Dr. Martina Roeselová from the Academy of Sciences of the Czech Republic for her contagious enthusiasm and joy who sadly passed away well before her time.

In the Poulos Lab, I must thank Drs. Huying Li, Irina Sevrioukova, Yergalem Meharena and Dipanwita Batabyal for sharing their years of experience with me throughout my graduate career and pushing me to new heights that I would have never reached without their aid. The countless conversations I've shared with them mean more to me than I can hope to express. I also thank Joumana Jamal for her tireless work to ensure that the Poulos Lab continued to run despite the lab's best efforts. I must also acknowledge my undergraduate Brian Nguyen who not only aided in the research presented in chapter three, but also drove me to be a better scientist and mentor. I would also like to thank all of the enlightened and kind graduate students that I have been fortunate enough to call both peer and friend over the past five years; Dr. Vickie Jasion, Dr. Yarrow Madrona, Dr. Jeffrey Holden, Georges Chreifi, and Alec Follmer. Of particular note however, I must single out Matthew Lewis, who I have been fortunate to call not only a lab mate but a friend in both the Poulos Lab at UC Irvine as well as in the Dr. P. Andrew Karplus Lab at Oregon State University (who, along with Dr. Dale Tronrud, I will forever be indebted to for giving me my start in research).

In addition, I gladly acknowledge Drs. Celia Goulding and David Mobley for serving on both my advancement and dissertation committee for their valuable insights, guidance and support, as well as Drs. Melanie Cocco and Donald Senear for their helpful suggestions and feedback as part of my advancement committee. I must also thank Dr. Mike Mulligan and Dr. David Fruman for their guidance and support throughout my graduate career.

I gladly acknowledge a pair of institutional training grants that supported both my research and myself; a predoctoral fellowship from the Cancer Biology Training Grant (NIH/NCI T32CA009054) as well as a predoctoral fellowship from the Chemical and Structural Biology Training Grant (NIH/NIGMS T32GM108561) as well as the Francisco J. Ayala Graduate Fellowship. I also greatly acknowledge the American Chemical Society as well as John Wiley and Sons for their permission to reproduce chapters 2, 4 and 5 in this document.

CURRICULUM VITAE

SCOTT ALAN HOLLINGSWORTH

Email: scott.alan.hollingsworth@gmail.com

EDUCATION

Oregon State University

Bachelor of Science in Biochemistry and Biophysics

GPA: 3.86/4.0 (*Summa Cum Laude*)

Corvallis, Oregon

June 2010

University of California, Irvine

Doctor of Philosophy in Biological Sciences

Thesis Advisor: Chancellor's Professor Thomas L. Poulos

GPA: 3.94/4.0

Irvine, California

May 2016

RESEARCH EXPERIENCE

September 2011 - Present

Graduate Student Researcher: University of California, Irvine, California

Research Advisor: Dr. Thomas L. Poulos

Structural, experimental and computational analysis of heme proteins and their associated protein redox partners including cytochrome P450, peroxidase and nitric oxide synthase systems

December 2011 – June 2012

Graduate Rotation Student: University of California, Irvine, California

Research Advisor: Dr. Douglas J. Tobias

Probed membrane protein dynamics and the behavior of water on organic self-assembled monolayer surfaces using molecular dynamics simulations

January 2007 – June 2010

Undergraduate Researcher: Oregon State University, Corvallis, Oregon

Research Advisor: Dr. P. Andrew Karplus

Computational studies of commonly occurring protein motifs through the application of machine learning and database analysis to large sets of high-resolution protein structures

PUBLICATIONS

1. **Hollingsworth SA**, Berkholtz DS & Karplus PA. (2009). "On the occurrence of linear groups in proteins." *Protein Sci.* **18**:1321-25

2. **Hollingsworth SA** & Karplus PA (2010). "A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins." *Bimolecular Concepts.* **1**:271-83. (*Review*)

3. **Hollingsworth SA**, Lewis MC, Berkholtz DS, Wong WK, Karplus PA. (2012). “ $(\phi, \psi)_2$ -motifs: a purely conformation-based, fine-grained enumeration of protein parts at the two-residue level.” *J. Mol. Bio.* **416**:78-93.
 4. Madrona Y, **Hollingsworth SA**, Khan B, Poulos TL. (2013). “P450cin active site water: implications for substrate binding and solvent accessibility.” *Biochemistry.* **52**:5039-50.
 5. Nishino N, **Hollingsworth SA**, Stern AC, Roeselová M, Tobias DJ & Finlayson-Pitts RJ. (2014). “Interactions of Gaseous HNO₃ and Water with Individual and Mixed Alkyl Self-Assembled Monolayers at Room Temperature.” *Phys. Chem. Chem. Phys.* **16**: 2358-67.
 6. Madrona Y, **Hollingsworth SA**, Stripathi S, Fields JB, Rwigema C, Tobias DJ, Poulos TL. (2014). Crystal structure of cindoxin, the P450cin redox partner. *Biochemistry.* **53** (9): 1435-46.
 7. Holden JK, Kang S, **Hollingsworth SA**, Li H, Chen S, Huang H, Xue F, Tang W, Silverman RB & Poulos TL. (2015). “Structure-based design of bacterial nitric oxide synthase inhibitors.” *J. Med. Chem.* **58**: 994-1004.
 8. **Hollingsworth SA** & Poulos TL. (2015). “Molecular dynamics of the P450cam-Pdx complex reveals complex stability and novel interface contacts.” *Protein Sci.* **24**: 49-57.
 9. Chreifi G, **Hollingsworth SA**, Li H, Tripathi S, Arce AP, Magaña-Garcia HI & Poulos TL. (2015). “Enzymatic mechanism of *Leishmania major* peroxidase and the critical role of specific ionic interactions.” *Biochemistry.* **54**: 3328-3336.
 10. Fields JB and **Hollingsworth SA**, Chreifi G, Heyden M, Arce AP, Magaña-Garcia HI, Poulos TL & Tobias DJ. (2015) “‘Bind and Crawl’ Association mechanism of *Leishmania major* peroxidase and cytochrome c revealed by Brownian and molecular dynamics simulations.” *Biochemistry.* **54**:7272-82. (*Co-First Author*)
 11. **Hollingsworth SA** and Holden JK, Li H, Poulos TL. (2016) “Elucidating nitric oxide synthase domain interactions by molecular dynamics.” *Protein Sci.* **25**:374-382 (*Co-First Author*)
 12. **Hollingsworth SA**, Lewis MC & Karplus PA. “Beyond basins: ϕ, ψ preferences of a residue depend heavily on the ϕ, ψ values of its neighbors.” *Protein Sci.* (Accepted) (*Co-First Author*)
 13. **Hollingsworth SA**, Batabyal D, Nguyen BD & Poulos TL. “Conformational selectivity in cytochrome P450 redox partner interactions.” *Proc. Natl. Acad. USA.* (Accepted) (*Co-First Author*)
- In Preparation
14. **Hollingsworth SA**, Nguyen BD, Chreifi G, Arce AP & Poulos TL. “On the disassociation mechanism of the *Leishmania major* peroxidase and cytochrome c complex.” (Submitted) (*Co-First Author*)

15. **Hollingsworth SA**, Freites JA, White SH, Falke JJ & Tobias DJ. “The dynamics of the docking of the PH-domain to a target PIP3 modified lipid in a complex bilayer.” (In Preparation)

16. **Hollingsworth SA**, McIntire T, Stern AC, Roeselová M, Hemminger JC & Tobias DJ. “Wetting effects on the surface and water penetration of perfect and defective self-assembled monolayers.” (In Preparation)

17. **Hollingsworth SA**, Stern AC, Roeselová M & Tobias DJ. “Wetting effects and droplet formation on rough hydrophobic self-assembled monolayers.” (In Preparation)

18. **Hollingsworth SA** & Tobias DJ. “On the effect of chain length and oxidation state on water and monolayer behavior in organic self-assembled monolayers.” (In Preparation)

19. Geragotelis A, Wood ML, Goeddeke H, Riahi S, **Hollingsworth SA**, Freites JA, Tombola F and Tobias DJ. “Multi-microsecond molecular dynamics simulations of the Hv1 Proton Channel.” (In Preparation)

FELLOWSHIPS

Predocctoral Chemical and Structural Biology Training Grant Fellowship (*UC Irvine, 2014-Present*)

Predocctoral Cancer Biology Training Grant Fellowship (*UC Irvine, 2013*)

Francisco J. Ayala Graduate Fellowship (*UC Irvine, 2011*)

Howard Hughes Medical Institute Summer Undergraduate Research Fellowship (*OSU, 2008 & 2009*)

INVITED TALKS

• **Hollingsworth, SA**. (March 2015). Molecular dynamics of the P450cam-Pdx complex reveals complex stability and possible Pdx-dependent behaviors of P450cam.” Presented at the 22nd West Coast Protein Crystallography Workshop in Monterey, CA, USA.

• **Hollingsworth, SA**. (January 2014). Studies of water in defects on SAMs. Presented at the 2014 AirUCI Workshop in Laguna Beach, CA, USA.

• **Hollingsworth, SA** (April 2013). Insights into the association of LmP & LmCytC through the application of Brownian dynamics. Presented at the 2013 UCI Cancer Biology Training Grant Retreat in Palm Springs, CA.

• **Hollingsworth, SA** (August 2009). Identification of common protein motifs through the application of machine learning. Presented at the 2009 HHMI Undergraduate Summer Symposium at Oregon State University (*Winner of Best Talk Award*)

- **Hollingsworth, SA** (August 2008). Beyond Pauling & Corey: Conformation and peptide geometry of linear groups in proteins. Presented at the 2008 HHMI Undergraduate Summer Symposium at Oregon State University

POSTER PRESENTATIONS

- **Hollingsworth SA**, Fields JB, Chreifi G, Heyden M, Arce AP, Magaña-Garcia HI, Poulos TL & Tobias DJ. (February 2016). Association mechanism of *Leishmania major* peroxidase and cytochrome c revealed through Brownian and molecular dynamics. Poster presented at the 60th Biophysical Society Meeting in Los Angeles, CA, USA.
- Gregagotelis A, Wood ML, Goeddeke H, Riahi S, **Hollingsworth SA**, Freitas JA, Tombola F, & Tobias DJ. (February 2016). Multi-Microsecond molecular dynamics simulations of the HIV1 proton channel. Poster presented at the 60th Biophysical Society Meeting in Los Angeles, CA, USA.
- **Hollingsworth SA**. (January 2016). Insights into the association of LmP & LmCytc through the application of Brownian dynamics. Poster presented at the 2nd Annual Ewha-UCI Symposium in Irvine, CA, USA.
- **Hollingsworth SA**, Lewis MC & Poulos TL. (March 2015). Beyond basins: ϕ, ψ preferences of a residue depend heavily on the ϕ, ψ values of its neighbors. Poster presented at the 22nd West Coast Protein Crystallography Workshop in Monterey, CA, USA.
- **Hollingsworth SA**. (May 2013). Insights into the association of LmP & LmCytc through the application of Brownian dynamics. Poster presented at the 2013 UCI Cancer Research Institute Symposium on Basic Cancer Research in Irvine, CA, USA.
- **Hollingsworth SA**. (June 2013). Insights into the association of LmP & LmCytc through the application of Brownian dynamics. Poster presented at the 18th International Conference on Cytochrome P450 in Seattle, WA, USA.

AUTHORED GRANTS

- “Visualizing the dynamic movements in heme proteins.” PI: Poulos TL. Agency: XSEDE. Type: Stampede Computing Resources. MCB130001. Effective Dates: 01/01/2016 – 12/31/2016
- “Modeling diffusion of membrane-bound signaling proteins on the microsecond timescale.” PI: Tobias DJ. Agency: NRC/PSC. Type: Anton Computing Time. PSCA15007P. Effective Dates: 11/01/2015 – 07/31/2016.
- “Visualizing the dynamic movements of heme proteins: heme redox partner complexes and drug design.” PI: Poulos TL. Agency: XSEDE. Type: Stampede Computing Resources. MCB130001. Effective Dates: 01/01/2015 – 12/31/2015

- “Visualizing the dynamic movements in heme proteins.” PI: Poulos TL. Agency: XSEDE. Type: Stampede Computing Resources. MCB130001. Effective Dates: 01/01/2014 – 12/31/2014
- “Visualizing dynamic movements in heme proteins.” PI: Poulos TL. Agency: NRC/PSC. Type: Anton Computing Time. PSCA13012P. Effective Dates: 11/01/2013 – 07/31/2014.
- “Molecular dynamics simulations of heme proteins.” PI: Poulos TL. Agency: XSEDE. Type: Stampede Computing Resources. MCB130001. Effective Dates: 01/01/2013 – 12/31/2013.

HONORS AND AWARDS

- Cover Image of *Protein Science* (January 2015)
- Joseph H. Stephens Memorial Prize (June 2014)
- NSF Graduate Fellowship Honorable Mention (2012)
- Featured article in *Protein Science* (June 2009)
- Best Talk in Oregon State HHMI Undergraduate Summer Symposium (August 2009)
- Carroll DeKock Scholarship (2008-2009)
- Drucilla Shepard Smith Scholastic Award (2006-2007)
- Diversity Achievement Scholarship (2006-2010)

COMPUTATIONAL & LABORATORY SKILLS

Platforms: Windows, Mac OS, Linux/Unix (Gentoo, Fedora, Ubuntu)

General Applications: Excel, PowerPoint, Word, R Studio, Grace, GIMP, gnuplot

Simulation and Structural Software: PyMol, Coot, CCP4, PHENIX, Chimera, VMD, NAMD, Amber, AutoDock, Vippar, CHARMM GUI

Programming: HTML, Bash, Python, C, Tcl, R

Laboratory Skills: X-ray Crystallography, Protein Expression and Purification, Enzyme Kinetics, Molecular Dynamics, Brownian Dynamics, MDFF, Protein Modeling, Docking, Bioinformatics, Data Analysis

TEACHING

Teaching Assistant – Bio M114L, Biochemistry laboratory *January – June 2014*

- Independently ran a 20 student undergraduate biochemistry laboratory with minimal faculty supervision. Created and gave short 10-20 minute lectures before each lab.

Teaching Assistant – Bio 98, Biochemistry lecture *January – March 2014*

- Prepared teaching aids, lead discussion groups of 30-60 students over the course of a 10 week term and graded exams.

Teaching Assistant – Bio 99, Molecular Biology lecture *April – June 2014*

- Prepared teaching aids, lead individual discussion groups of 60 students for a total of over 400 students.

PROJECT MANAGEMENT

Coordinating Chemical and Structural Biology Training Grant Journal Club

Organize and coordinate a weekly journal club for twenty-six different laboratories across four different departments, (*UC Irvine, February 2016 – Present*).

Laboratory Meeting Coordinator

Prepares and organizes weekly group meetings for laboratory members, (*UC Irvine, October 2014 – Present*).

Undergraduate Research Advisor

Mentors undergraduates in a laboratory research setting, all supervised undergraduates have produced results leading to co-author manuscripts and have gone on to pursue a Ph.D., (*Oregon State, January 2010-June 2010; and UC Irvine, January 2015 – Present*).

ACADEMIC AFFILIATIONS

2013 – Present: AirUCI at University of California, Irvine

2016 – Present: Biophysical Society

REFEENCES

Dr. Thomas L. Poulos – Ph.D. Advisor

Chancellor's Professor. Departments of Molecular Biology and Biochemistry, Chemistry and Pharmaceutical Sciences. University of California, Irvine. (949)-824-7020. poulos@uci.edu.

Dr. David Mobley – Thesis Committee Member

Assistant Professor. Departments of Chemistry and Pharmaceutical Sciences. University of California, Irvine. (949)-385-2436. dmobley@uci.edu.

Dr. Celia Goulding – Thesis Committee Member

Professor. Departments of Molecular Biology and Biochemistry and Pharmaceutical Sciences. (949)-824-0337. celia.goulding@uci.edu.

Dr. P. Andrew Karplus – Undergraduate Research Advisor

Professor. Department of Biochemistry and Biophysics. Oregon State University. (541)-737-3196. karplusp@science.oregonstate.edu.

ABSTRACT OF THE DISSERTATION

At the interface of experiment and computation: explorations of heme protein redox partner interactions, water behavior on organic surfaces and other systems

By

Scott A. Hollingsworth

Doctor of Philosophy in Biological Sciences

University of California, Irvine 2016

Professor Thomas L. Poulos, Chair

In part one of this thesis, the interactions and resulting complexes between heme proteins and their associated electron-transferring redox partner proteins, termed here as heme protein redox partners, are carefully studied. Such complexes navigate a delicate balance between achieving fast turnover and tight, specific binding. To probe how Nature has balanced these seemingly opposing forces, a series of studies were conducted on three separate heme protein redox partner systems that have achieved a working balance using different approaches; Cytochrome P450cam and its redox partner putidaredoxin, *Leishmania major* peroxidase and its partner *Leishmania major* cytochrome c and finally the multi-domain nitric oxide synthase. Each system was initially probed using computational methods including molecular dynamics, Brownian dynamics, and modeling before being experimentally validated against previous findings, novel studies or both.

In part two, the focus is shifted from protein interactions to water behavior on organic surfaces. Studying the behavior of water on such organic surfaces can have wide-ranging impacts on our understanding of not only fundamental water dynamics, but also atmospheric chemistry

and chemistry on urban surfaces. Self-assembled monolayers (or SAMs) have long served as a model system to probe how differences in organic surface chemistry may affect water behavior. To this point, the behavior of water on a variety of SAMs were investigated using molecular dynamics in a series of studies including one with experimental validation to deepen our understanding of the dependence of water behavior on the structural makeup of both a pristine idealized surface and more realistic defective SAM systems.

Finally, three additional separate studies are presented in the third and last part of this thesis. These distinct and independent studies 1) cover the conformational dependence of a central protein residue on the conformation of its neighbors, 2) computationally study the binding of a membrane associated protein to a target lipid in a lipid bilayer and 3) experimentally determine the crystal structure of *Bacillus subtilis* arginase. When combined with the two series of studies previously presented in this thesis, the great potential in applying a combined experimental and computational approach to address unanswered questions is highlighted and explored.

Chapter 1

An introduction to Heme Proteins and Their Redox Partners

One of the most abundant metalloporphyrins in Nature is the iron-centric heme or iron protoporphyrin IX (1). Heme containing proteins, or heme proteins for short, have been the focus of considerable study for well over a century dating back to the first descriptions of a heme protein, horseradish peroxidase, in what is now considered to be a historical study by Bach and Chodat (2). The heme cofactor can carry out a wide variety of reactions both reductive and oxidative (1). While the work herein focuses nearly exclusively on heme enzymes that catalyze oxidative reactions, it is important to note that heme enzymes can also catalyze reductive chemistry as well. Oxidative heme enzymes fall into two primary categories that differ in function and mechanism; oxygenases and peroxidases (1). Oxygenases enzymes, such as cytochrome P450 and nitric oxide synthase, bind molecular oxygen to the heme iron to oxidize their respective substrates normally through means of an oxygenation reaction. Peroxidases however bind H_2O_2 to the heme iron and oxidize substrates though normally through the utilization of oxygenation chemistry. Both of these reactions require an input of electrons to power the oxidative power of the respective enzyme.

It should not have escaped the reader that in these examples and in those that follow, there must be an electron donor for the required electron transfer process to occur to power these reactions (3, 4). The source of these electrons can be a small molecule, a separate but connected protein domain or even a separate protein. In these cases where a separate protein acts as the source of the electrons, such pairs of proteins or domains can be described as protein redox partners. Herein, we focus our attention to those redox partners that transfer electrons ultimately to a heme group which henceforth will be referred to as heme protein redox partners.

The kinetics of protein redox partners represent a delicate balancing act that Nature must carefully navigate in order to weigh specificity against quick turn over. On one hand, such protein redox partners must have a high degree of specificity between proteins through electrostatic interactions, hydrophobic regions, or a combination of the two. On the other hand however, the complex itself must be dynamic, that is associating and disassociating quickly in order to achieve rapid turnover. If the interactions that promote complex formation are too strong, then the complex would not be dynamic as the complex itself would be long lived. If though the interactions are too weak as to encourage rapid turnover, then the complex may never form in the specific orientation to promote turnover. As such, any pair of protein redox partners must achieve a delicate balance between strong, specific interactions and rapid turnover.

Electron transfer between protein redox partners can be further broken down into three generalized steps (Figure 1.1). Firstly, the two proteins must associate to form a complex. Importantly, there has been a great deal of work that has shown that the initial encounter complex that is formed upon complex formation may not be the ET active complex structure (5-12). In these cases where the encounter complex and active complex are different, the formation step is broke down into two phases, the formation of the encounter complex and the local sampling that leads to the ET active complex. The second step is the ET step itself between the protein redox partners and due to the nature of electron transfer, is seldom the rate-limiting step. Following the ET event, the complex then disassociates in the third step before the process can begin again. Understanding how each of these steps occur (either association or dissociation as well as studying the complexes that result from the interactions of protein redox pairs) is vital in deepening our understanding of the structure-function relationships in protein redox partners.

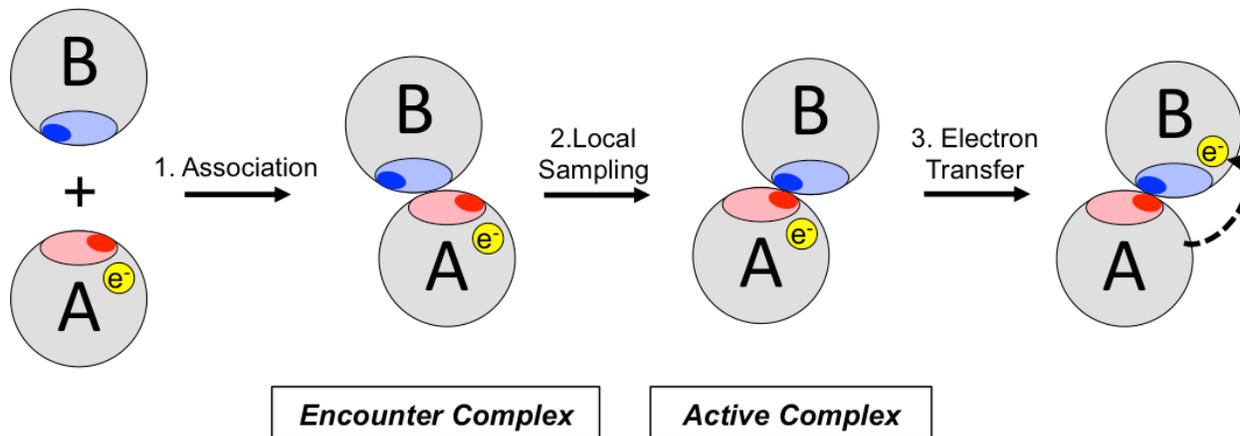


Figure 1.1. Generalized protein redox partner association. Protein redox pair association can be divided into three major stages. During the initial association (1), the complimentary surfaces of each protein (shown as pale red and blue) make contact resulting in the encounter complex that is not active as the reactive patches are not yet in position (dark red and blue). The complex then undergoes local sampling (2) which reduces the conformational sampling to a two-dimensional problem from the initial three dimensional search. Upon formation of the active complex (3), the interprotein electron transfer occurs and the complex disassociates.

What follows are three introductions that focus on three separate heme-containing enzymes that at their heart, rely on heme protein redox partners to help power their reactions and will be the focus of a majority of the research included here. In Part I, Cytochrome P450 enzymes are introduced and discussed, focusing on the incredible power and remaining mysteries of the redox partner binding of the P450cam-Pdx protein redox partner complex. In Part II, peroxidases are highlighted, focusing on a pair of heme-containing proteins that make up a protein redox partners that allows us to further probe how these redox partners associate and form the active electron transfer complex. Finally, in Part III the multi-domain nitric oxide synthase is presented, emphasizing the incredible flexibility of a single polypeptide to undergo tremendous conformational changes in order to carry out two separate electron transfer reactions and complexes that have eluded and puzzled many researchers for decades.

Part I: Cytochromes P450 Enzymes

Cytochrome P450s (or CYPs) are a super-family of heme containing enzymes that activate otherwise unreactive C-H bonds in many different compounds including xenobiotics, steroids, and in some cases compounds that are oxidized as energy sources (13, 14). Named for the distinctive peak at 450nm when bound to CO (15, 16), P450 enzymes have been the focus of a great deal of research for many years. P450s are found across all kingdoms of life as well as viruses with over 18,000 different CYP genes identified to date including at least 57 CYP genes reported in humans as of 2015 (1, 17). In order to distinguish between the many different enzymes, each enzyme is designated as CYPabc where CYP stands for cytochrome P450 and “abc” is a unique three-character code. However, as several notable P450 or CYP enzymes are known more commonly by other names than their unique CYP designator (normally “P450” followed by a short additional designator based on the source of the enzyme), both descriptors will be used interchangeably herein. While there is a great deal of diversity regarding function and substrate specificity among CYP enzymes, each enzyme carries out the same generalized hydroxylation of its substrate to produce product (Reaction 1):



As P450s are able to activate an otherwise inactive C-H bond and oxidize a wide variety of substrates, these enzymes have been lovingly dubbed “Nature’s blow torch” to acknowledge both the awesome power and multiple uses that this enzyme can provide. Through this singular generalized reaction, P450 enzymes are able to carry out the detoxification of xenobiotics (14, 18, 19), aid in the synthesis of antibiotics (20, 21), or provide bacteria with unexpected sources of carbon (22). P450 enzymes play vital roles in a multitude of metabolic pathways, particularly

in the metabolism of drugs which has been the focus of a great deal of research (18, 23, 24). To illustrate the importance of understanding P450 enzyme function, one can look no further than CYP3A4 from *homo sapiens*. This single P450 enzyme, one of 57 in humans, has been shown to be responsible for metabolism of over 50% of the most commonly prescribed medications on Earth (1). The targets for CYP3A4 range wildly in both substrate size and shape (18). If one were able to understand how to inhibit this enzyme, perhaps derived from a deep understanding of the enzymatic function or structure, it could lead to lower doses of these medications retaining the activity of their currently prescribed dosage levels due to the inhibition of the enzyme primarily involved in their metabolism.

Cytochrome P450 Mechanism

Cytochrome P450 have earned their moniker as “Nature’s blow torch” through their ability to hydroxylate otherwise unreactive alkanes in a region- and stereo-selective fashion (25). Through the activation of the otherwise inert C-H bond, P450s are able to modify a wide variety of compounds the likes of which organic and inorganic chemists have yet to reliably reproduce. Through many decades of research, the P450 mechanism or P450 cycle has been well characterized and is displayed in Figure 1.2. The substrate first enters the active site pocket and displaces a water molecule that acted as the six axial ligand for the ferric state iron in the active site heme (A in Figure 1.2). Once the substrate has bound (B), an electron is donated by a redox partner protein which shifts the iron from the ferric(III) to the ferrous(II) state (C) which increases the binding affinity of the heme to molecular oxygen which then quickly binds (D). Upon molecular oxygen binding, internal electron rearrangements return the heme to the

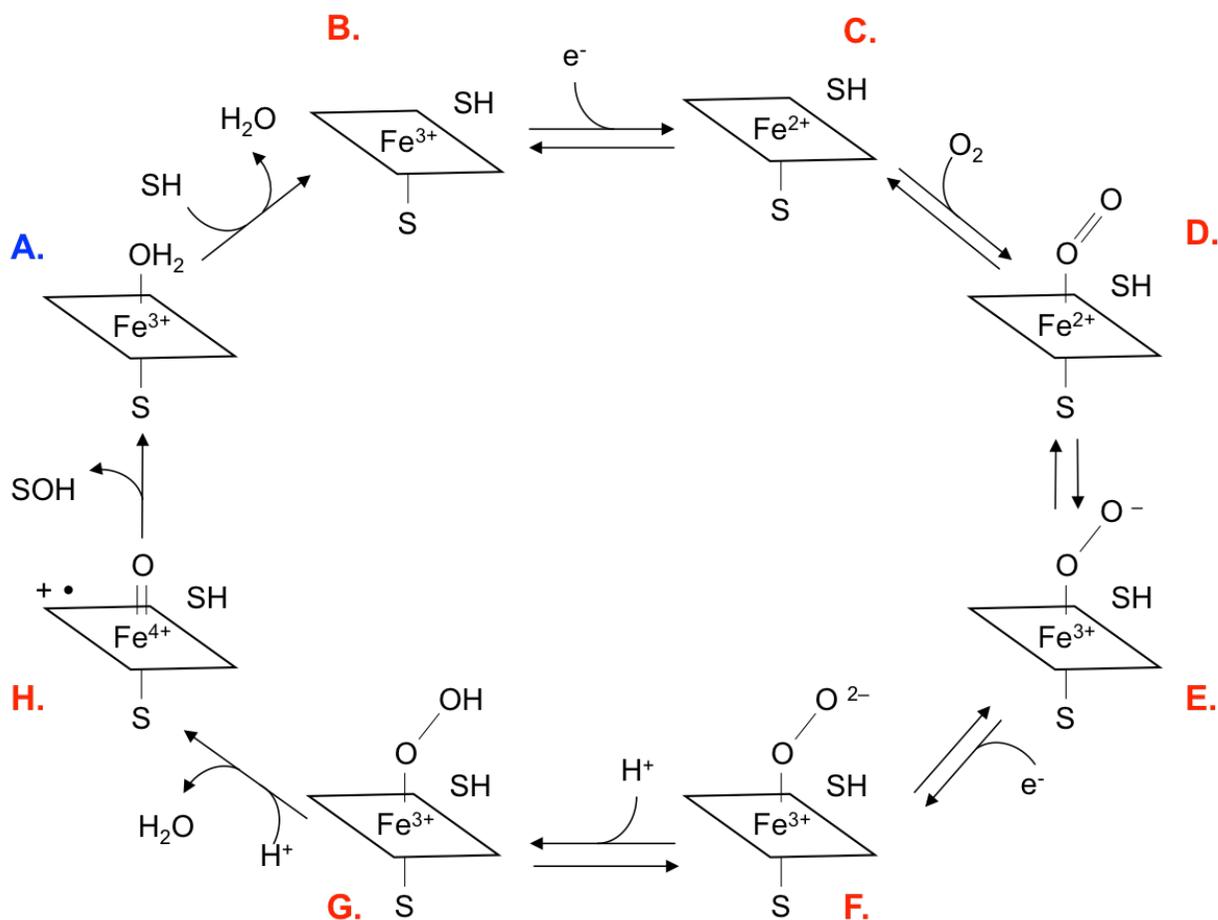


Figure 1.2. The cytochrome P450 catalytic mechanism. The P450 catalytic mechanism (or P450 cycle) is shown. Intermediates are labeled A to H as described in the text. Intermediate labels are colored blue or red to represent low or high spin states respectively.

ferric(III) state resulting in the ferric oxy-complex or ferric-superoxide or $\text{Fe}^{3+}\text{-O}_2^-$ (E). At this point, a second electron is donated from a protein redox partner that begins the process of the O-O bond cleavage through the formation of a ferric peroxo complex with a -2 charge on the distal oxygen (F). The addition of a proton results in a ferric hydroperoxo complex which has been experimentally observed through EPR and ENDOR methods (G) (26). At this point in the reaction cycle, the addition of a second proton results in the loss of a water molecule the Fe(IV), porphyrin cation radical known as compound 1 (H). This unstable intermediate has been shown to be responsible species for the hydroxylation of the substrate by Rittle and Green who were the

first to trap and characterize the P450 compound I (27). The actual hydroxylation has been proposed occur through the so-called rebound mechanism, displayed in Figure 1.3 (28). During hydroxylation, the unstable compound I abstracts a proton from the substrate leaving a substrate radical and a hydroxyl group bound to the heme iron. The remaining substrate then rebounds to attack the oxygen to form the hydroxylated product while returning the heme iron to the ferric state. The resulting product leaves the active site pocket while a solvent water molecule returns to acts as the six axial iron ligand returning the iron to its original state (A), ready to accept the next substrate molecule to begin the cycle anew.

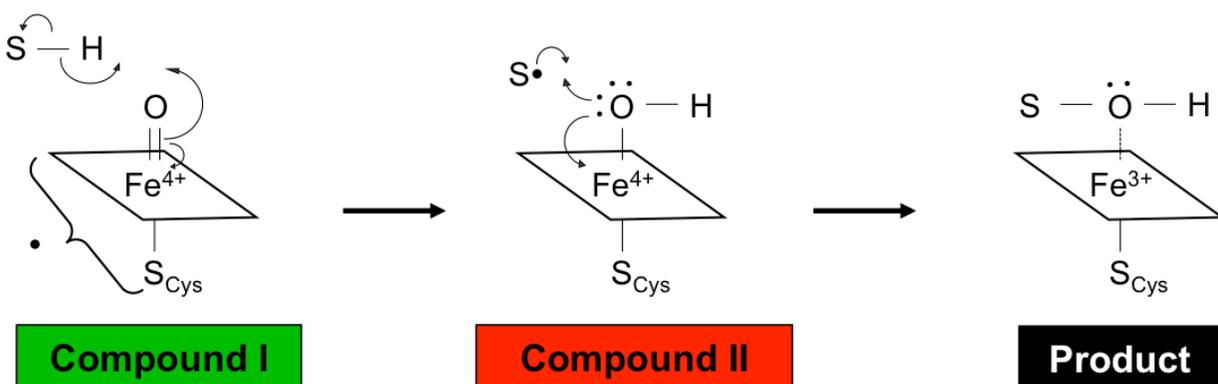


Figure 1.3. Rebound mechanism for the hydroxylation of substrate by compound 1.

Following compound I formation, it is proposed that hydroxylation of the substrate occurs via a rebound mechanism whereupon a hydrogen is abstracted from the substrate leaving a carbon radical that “rebounds: to break the iron-oxygen bond resulting in the product that serves as a heme iron coordinating ligand.

One additional aspect of the P450 mechanism denotes further mention with regards to the role that it plays in aiding P450 function, that of the so-called Push-Pull Effect (29-31). The Push-Pull effect in P450 enzymes can be broken into two parts. The first of these parts, the “Push”, has to do with the stabilizing of the reactive compound I intermediate. The heme ligating Cys residue is a strong electron donating ligand that aids in the promotion of the cleavage of the O-O bond. Upon formation of compound I, the mixing of two separate resonance forms, the

porphyrin-based positive radical shown in Figure 1.2 (the active species for hydroxylation) and a Cys-sulfur based radical further lengthens the lifespan of compound I. The second portion, the so-called Pull effect, described the hydrogen bonding pattern that arises in the oxy-state active site of P450 (Figure 1.4). An active site water bridges the bound molecular oxygen to the sidechains of Asp251 and Thr252 (1). These hydrogen bonds have been proposed to further promote cleavage of the O-O bond by pulling on the distal oxygen atom away through this hydrogen bonding arrangement.

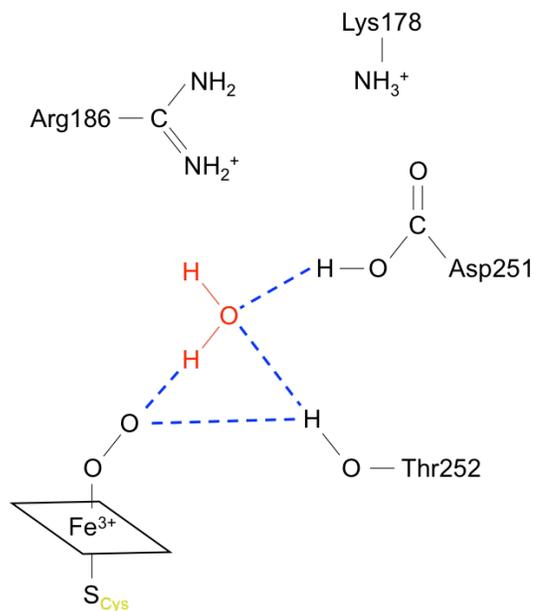


Figure 1.4. The “Pull” Effect: proposed hydrogen bonding arrangement in the oxy-state of P450cam. The proposed hydrogen-bonding network near the heme active site that helps to promote breakage of the O-O bond in the “Pull” effect.

The paramagnetic nature of the heme iron plays a role in changing unique spectroscopic characteristics due to the arrangement of the iron’s valence electrons (32). While five of the ligands to the heme iron, the four nitrogen atoms of the heme porphyrin and one from a coordinating residue in the protein (often a Histidine nitrogen or Cysteine sulfur), the binding of a strong or medium field ligand (substrate) can displace the water molecule that acts as the sixth and final axial ligand. This displacement results in a large energy split between the d_{xy} , d_{xz} , d_{yz} and d_x^2 , $d_x^2 - y^2$ orbitals due to the overlap of the ligand field with the d_x^2 , $d_x^2 - y^2$ orbitals. In ferric heme, four of the five valence electrons are paired. This final unpaired electron results in a spin state described as $S = +1/2$. This spin state has been termed “low spin.” However, upon binding of a strong or medium ligand field substrate and resulting displacement of the axial water, the drop in energy of the d_z^2 , $d_x^2 - y^2$ orbitals results in all five electrons each occupying a single

orbital and a spin state of $S = +5/2$ or the “high-spin” state of the heme iron (Figure 1.5). Each spin state produces a unique Soret peak, ~412nm for low-spin and 392nm for high-spin, which can be tracked using spectroscopy to probe substrate binding.

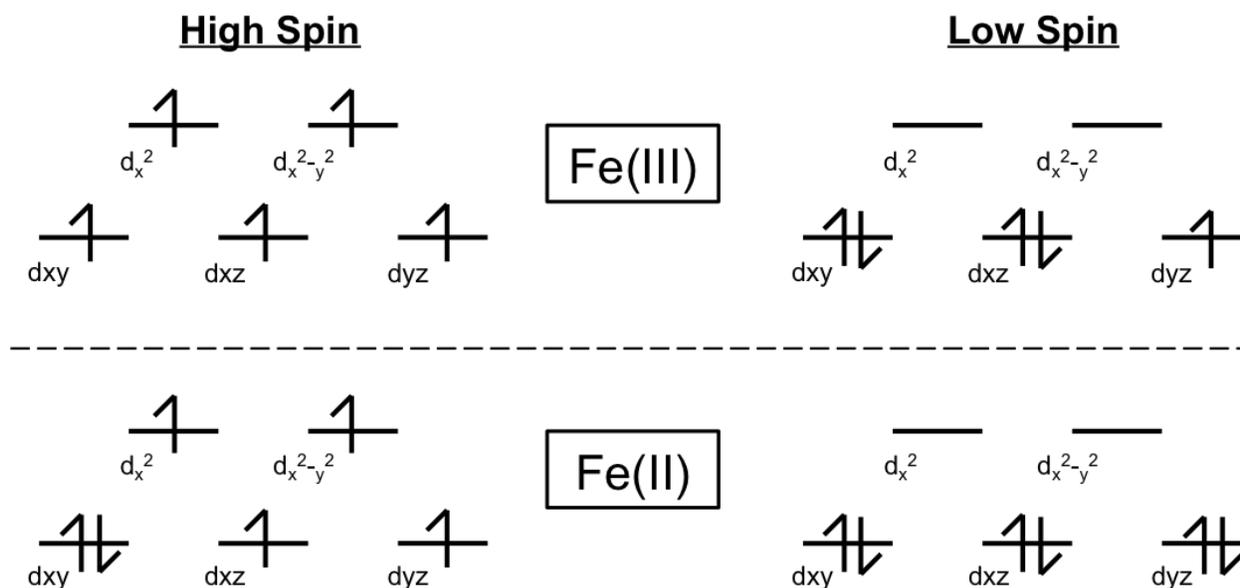


Figure 1.5. Valence electron arrangement between low and high spin state of the heme iron. The binding of a strong field ligand to a iron in either the Fe(II) or Fe(III) state resulting in a lowering of the energy of the d_{x^2} and $d_{x^2-y^2}$ orbitals, resulting in internal electron rearrangements and a higher quantum spin state.

Class I vs. Class II CYP Enzymes

The power source to “Nature’s blow torch,” Cytochrome P450, comes from the donation of electrons provided by NAD(P)H which are relayed through several protein redox partners to the heme of P450 as shown previously. The localization of these protein redox partners have traditionally served to separate P450 enzymes into two classes. The two electrons described in Reaction 1, derived from NAD(P)H, travel through a series of flavin or iron-sulfur cluster containing proteins that may or not be on the same polypeptide chain as the heme active site. In Class I systems, which are found primarily in bacteria and mammalian mitochondria, electrons from NAD(P)H are transferred to a FAD containing reductase domain protein, then to a separate

Fe₂S₂ cluster protein before this second protein transfers the electron to the heme of the P450 enzyme where molecular oxygen is activated, powering the blow torch for activity. Importantly, each of the three proteins are separate polypeptides. In bacterial forms, each of the three enzymes are soluble while in mammalian mitochondria only the iron-sulfur shuttle protein is soluble as the reductase is tethered to the inner mitochondrial matrix and the P450 enzyme associated with a membrane. Class II systems, most often associated with P450s associated with the inner membrane of the endoplasmic reticulum in eukaryotes, consist of a fused reductase domain constructed of the FAD containing domain and an FMN containing domain in place of the iron-sulfur containing protein from Class I, while the P450 is still a separate polypeptide. Commonly, both the fused reductase domain and the P450 enzyme are both anchored to the membrane.

While the distinction between Class I and Class II has been rooted in historical reasons, it is important to note that there are prominent examples of P450 systems that do not fit into either category, of particular note to this author is P450cin from the soil bacterial *Citrobacter braakii* (13, 22, 33, 34). As a bacterium, it would be expected that P450cin would belong to Class I with FAD and iron-sulfur containing proteins acting at the first and second proteins during electron transfer from NAD(P)H to the heme of P450cin respectively. However, P450cin has a FMN containing protein in place the iron-sulfur protein expected in Class I, which while a defining characteristic of Class II, exists as a separate polypeptide from the FAD containing reductase unlike the expected fused nature of expected Class II enzymes. P450cin has been the focus of ongoing research in recent years to study not only the conformational flexibility of P450cin, but also how the final electron transfer complex of P450cin and the FMN containing cindoxin or Cdx associates and forms. While not included in this document, I played a supporting role in the study of P450cin and its interactions with Cdx, which played a large role in informing how the

work that is included herein was designed and carried out. More information regarding these initial studies can be found in Madrona et al. 2013 and Madrona et al. 2014 (33, 34).

Cytochrome P450 Structure

P450 enzymes share a conserved structural domain architecture with a size ranging from 40 to 55 kDa (Figure 1.6). These enzymes are primarily α -helical forming a roughly triangular when viewed from a plane parallel to the heme cofactor. The longest of these helices, the I helix runs from one side of the enzyme to the other, through the center of the enzyme and over the heme active site while interacting with substrates and dioxygen upon binding. Of particular note, P450 enzymes share a distorted I helix kink which results in the disruption of the regular α -helical hydrogen bonding pattern of residue i to residue $i+4$ to two occurrences of π -helical like where residue i hydrogen bonds to residue $i+5$ (35-37). While the I helix kink can vary amongst different P450 enzymes, one region that is highly conserved is found at the heme binding Cys residue. Located near the C-terminus of the L helix, the conformation of this residue is tightly controlled through local hydrogen bonding with neighboring residues. In P450cam, the ligating Cys357 participates in hydrogen bonding with Gly359 and Phe351. This combination of a backbone-to-backbone hydrogen bond between the carbonyl group of Phe351 and the backbone NH of Cys357 and a hydrogen bond between the backbone NH group of Gly359 and the sulfur atom of Cys357 plays vital roles in modulating the redox potential of the heme iron (38-40).

As should be expected, P450 enzymes are very dynamic and flexible in order to both bind and protect the substrate during hydroxylation before release of product and in the case of less substrate-selective enzymes to accommodate substrates of differing sizes and shapes. Amongst the varying conformational states that these enzymes take on, two extreme conformations exist

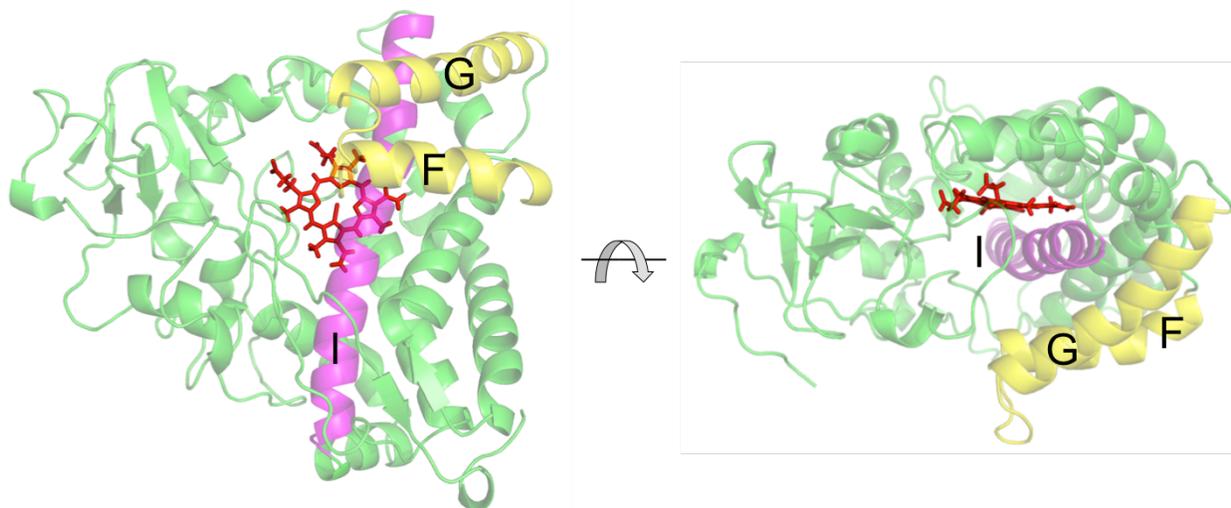


Figure 1.6. Structure of a model P450 enzyme, P450cam. The overall structure of P450cam in the open state is shown in two different views. The critical I helix is shown in purple while the F and G helices that undergo conformational changes upon substrate binding are shown in yellow. The heme is shown in red sticks.

that have been experimentally observed, the open and closed states (Figure 1.7). In order to probe the important differences in the open and closed state, we will continue our focus on P450cam as a model of P450 enzyme structure and function. In the open state, the F and G helices (as well as the connecting F/G loop) on the proximal side of the heme move away from the core of the enzyme revealing the substrate access channel to solvent and allowing substrate to bind to the enzyme. In this conformation, the catalytically vital residue Asp251, which acts in the proton relay network by delivering protons from the solvent to the active site to power the hydroxylation reaction, is free in the active site. Through previous work, the open conformation of P450cam had been experimentally solved only in the absence of substrate or in the presence of large substrates that prevent the enzyme from closing (41, 42). Upon substrate binding, the F and G helices move towards the core of the enzyme, closing the substrate access channel resulting in the closed conformation of the enzyme. In this conformation, Asp251 forms two intramolecular ion pairs with Lys178 and Arg186 that prevent the Asp251 sidechain from participating in the

proton relay network. Experimentally observed for both substrate bound forms of P450cam as well as many catalytic important transient states (43, 44), it was believed until recently that P450cam existed in the open state only in the absence of substrate.

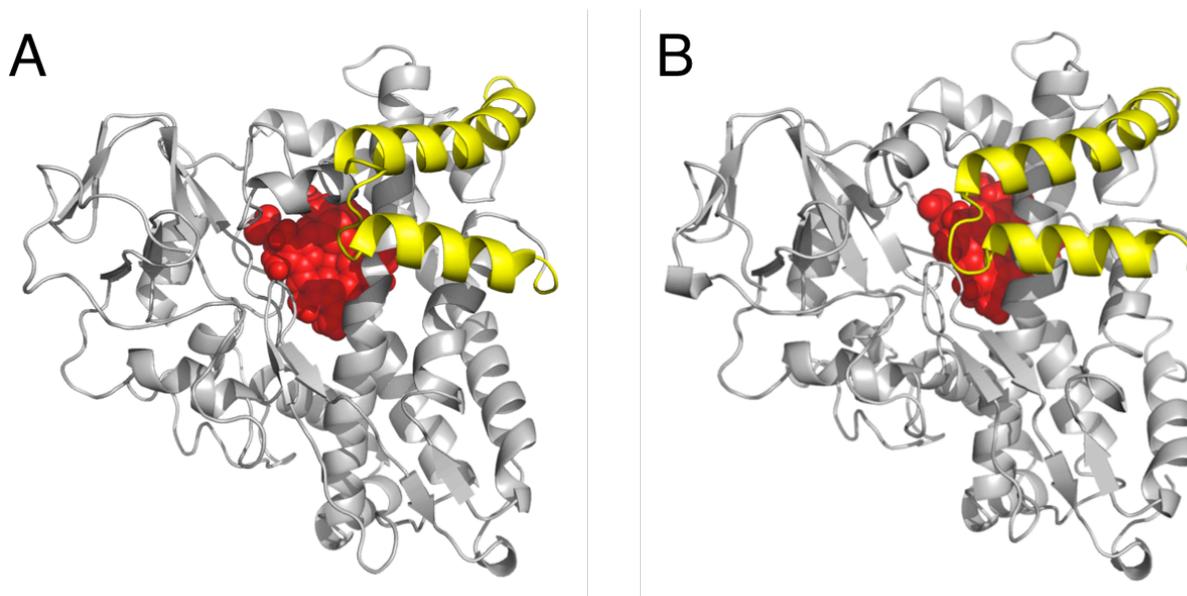


Figure 1.7. Comparison of the open and closed state of P450cam. Movements in the F and G helices upon substrate binding cause P450 to shift conformational state from the substrate-free open state which exposes the heme active site to solvent through the substrate access channel (A) to the substrate bound closed state in which the substrate access channel is closed to solvent (B).

Despite the fascination on these two extreme conformations, it should be noted that other conformations of the enzymes exist that play equally vital roles in activity, chiefly amongst them the oxy-complex state of P450cam (35). In this state, which can generously be described as being in the closed conformation with regards to the F and G helices, important structural changes occur upon binding molecular oxygen. In the ferric, closed state of P450cam, the side chain of Thr252 at the site of the I helix kink forms an irregular hydrogen bond with the backbone of Gly248 causing the I helix kink described previously. Upon formation of the oxy-complex, this hydrogen bond is weakened to the point of breaking which widens the groove in the I helix that allows a water molecule to enter the active site and establish a new hydrogen bond network. This

helical conformational change also changes the orientation of Asp251, which returns to normal α -helical hydrogen bonding patterns. This new water molecule has also been implicated as the source of protons for the hydroxylation of the reaction (See Figure 1.4).

P450 Protein Redox Partner Complexes

To date, there exists a limited degree of structural information regarding how P450 enzymes binds to appropriate redox partner proteins for electron transfer. Only three such complexes have been solved to date; P450BM3 (CYP102A1) bound to the FMN domain of its respective reductase domain (45), P450scc (CYP11A1) bound to the redox partner adrenodoxin or Adx (46), and finally the P450cam-Pdx complex (47, 48). Of the three complexes, only the P450BM3 and P450cam complexes were completely solved as Arx was mostly disordered in the P450scc-Arx complex structure. As this introduction has used P450cam as a model thus far, we shall continue to focus on P450cam and on the resulting redox partner complex structure.

In the absence of substrate, the redox potentials of P450 enzymes are high when compared to other heme containing enzymes. However upon substrate binding and the subsequent closing of the enzyme around the substrate, P450 undergoes a redox potential shift from approximately -340mV to -170mV. This alteration of redox potential play an important role in controlling electron transfer between P450cam and its redox partner Pdx as the electron transfer reaction is only thermodynamically favorable in the closed, substrate bound state as the potential of Pdx is approximately -239mV (49). This unique complex property acts to control the flow of electrons to P450cam in the closed, substrate bound form only. This finding has been used as further evidence that when Pdx binds to P450cam to transfer electrons, particularly the second electron transfer, Pdx would bind to the closed state of P450cam.

The P450cam-Pdx complex crystal structure, which many believed would help clear up a multitude of questions regarding the Pdx-effector role and redox partner interactions, instead lead to several new questions when it was solved in 2013 (Figure 1.8) (48). Surprisingly, P450cam was found in the open state when in complex with Pdx, the opposite of what was expected. While this structure was covalently cross-linked which could have trapped an inactive complex, a second, independently solved crystal structure was solved without the aid of cross-linking. Further evidence that this complex was in fact active was that product was found in the active site of the crystal structure suggesting that hydroxylation had occurred during data collection, likely due to the exposure to the x-ray beam. While forcing a reevaluation of expectations of a closed state complex, the result agrees well with all previous experimental data

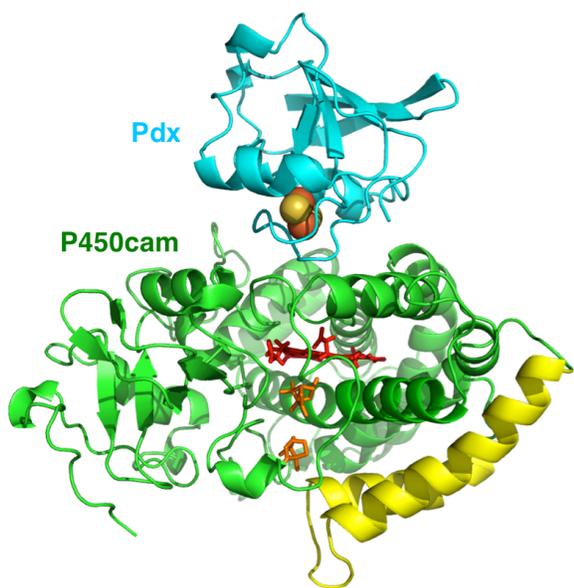


Figure 1.8. The P450cam-Pdx crystal structure. The P450cam-Pdx complex crystal structure as solved by Tripathi and coworkers in 2013. Pdx is displayed in cyan while P450cam is green with the F and G helices are highlighted in yellow to emphasize the open conformation of the enzyme. The heme and substrates and shown as red and orange sticks respectively.

(50-52). It has been known for many years that Pdx plays an effector role when binding to P450cam for the second electron transfer, altering the conformation of P450cam in the active site, which in context of the complex crystal structure suggested that the effector role was to shift P450cam from the closed state to the open state (53, 54). However, even with the solved crystal structure, several important questions still remained. Firstly, there are question whether the complex structure, trapped in a crystalline form, was in fact stable in solution. Second, is the open conformation the state that Pdx initially binds?

While the crystal structure suggests that the complex exists in the open state, the redox potential of P450cam has previously shown that the Pdx can only transfer electrons to P450cam in the closed conformation. This fact raises the possibility that the complex itself may be dynamic, with Pdx binding to the closed state, transfer an electron, then shifting P450cam to the open conformation. These questions are addressed in chapters two and three, using molecular dynamics, experimental methods, and revelations of previous structures to probe the dynamic interactions of the P450cam-Pdx complex which should provide valuable insights into the structure-function relationship into the widely studied P450cam model system. These findings could have wide ranging implications to our understanding of P450 enzyme function that may be applied to other systems beyond the P450cam model.

Part II: Peroxidases

Heme peroxidases, found in many different organisms, utilize peroxides to carry out a variety of physiological processes including H_2O_2 detoxification and biosynthetic pathways. Due in no small part to the ease of the preparation of large quantities of the earliest studied peroxidase Horseradish peroxidase (HRP), these enzymes have played a large roll in the early days of not only not only heme protein studies, but enzymology as well (2). These initial studies of HRP lead to the initial discovery of the catalytic cycle for peroxidases (Figure 1.9) (55, 56).

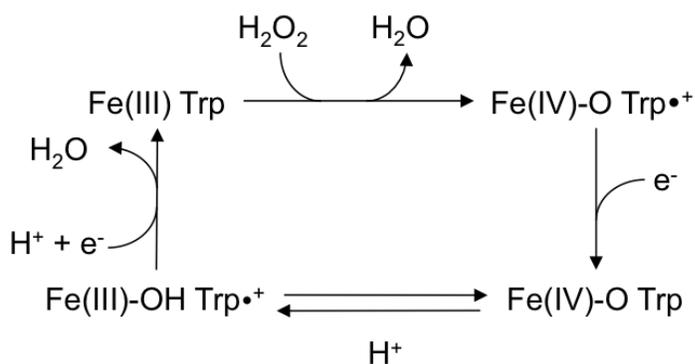


Figure 1.9. Mechanism of heme peroxidase. Following the binding of peroxide and production of compound I, the addition of an electron reduces the resulting $\text{Trp}^{\bullet+}$ cationic radical. An internal electron transfer and proton addition reduces Fe(IV)=O to Fe(III)-OH (compound II). Finally, a second donated electron returns the system to the Fe(III) Trp resting state.

In peroxidases, H_2O_2 reacts with the Fe(III) or ferric state iron to first form a transient intermediate described as an Fe(III)-hydroperoxy complex before being heterolytically cleaved, resulting in the loss of a water and formation of the high oxidation state and highly reactive compound I composed of a oxo-Fe(IV) or oxo-feryl center and a substrate based cation radical. Depending on the system in question, the radical is either based on the porphyrin ring of the heme or on a nearby amino acid side chain (57-59). Next, a single electron transfer reaction that reduced the radical results in the formation of the canonical compound II. Finally, a second electron transfer results in the return of the Fe(III) state from the Fe(IV) and the production of a water molecule that remains bound until the reaction mechanism begins anew (1).

While both cytochrome P450 and peroxidase enzymes share the same intermediates in compound I and II, the reactivities of these enzymes are very different. Like many things in science, the difference in activity cannot be reduced down to a single change between the enzymes but rather a combination of several important differences. Firstly, the distal pocket of P450 enzymes are non-polar which require an intricate water and hydrogen-bonding network to deliver protons (1). Peroxidases however have polar distal pockets and utilize His-mediated acid-base catalysis (60, 61). Second, the proximal heme ligand also differs between the enzyme families with a strong electron donating Cys in P450 enzymes and His in peroxidases that acts to further alter the redox potential of the heme. This is further altered through second sphere effects where the P450 Cys ligand is surrounded by nearby hydrogen bonding donors that increases the redox potential. Conversely, the peroxidase His ligand is surrounded by nearby hydrogen bonding acceptors that decreases the redox potential. Thirdly, the push effect described earlier makes the porphyrin a poor electron acceptor which make it more likely to abstract a hydrogen

atom from a substrate than accept another electron. This is in striking contrast to peroxidases where no such push effect exists as the His ligand does not donate electron density towards the porphyrin, making it a better electron acceptor. Finally, when comparing the structures of compound II, the pKa of the P450 compound II oxygen is approximately 12 while the Fe-O bond is roughly 1.82 Å making it very poor proton donor. In peroxidases, due to the lack of the protonated state of the iron-bound oxygen atom, the iron-oxygen bond distance is shorter at 1.7 Å and a pKa of approximately 3.5 Å.

Peroxidase enzymes can be designated into one of three classes based on sequence, host organism and structure (62). Class I peroxidases are intercellular peroxidases that lack disulfide bonds in the folded tertiary structure. Notable examples of Class I peroxidases include two of the most characterized peroxidase enzymes; mitochondrial yeast cytochrome c peroxidase or CCP and *Leishmania major* peroxidase or LmP. Class II peroxidases consist of fungal peroxidases that exist extracellularly. Finally, Class III peroxidases are extracellular plant peroxidases. The most notable Class III enzyme is the previously described HRP that paved the way in initial studies of these enzymes. Importantly, while Class I lack disulfide bonds, Class II and III contain a number of disulfide bonds as well as additional metal ion binding sites for additional enzyme stability in addition to glycosylation sites.

One of the most well understood heme peroxidases is yeast cytochrome c peroxidase or CCP. Originally discovered in 1939 in baker's and brewer's yeast (63), it was not until the 1960s that the activity of CCP has been well studied (64-66). From these studies came the first studies that were able to deduce characteristics of the elusive compound I intermediate including the resulting oxidizing agent, presence of a π -cation radical and presence of the Fe(IV) redox state iron atom (67-69). With the solving of the the CCP crystal structure acting as the final piece in

the puzzle, the model of compound I formation was deduced through the elucidation of the Poulos-Kraut model (70). The structure of CCP also aided in the identification of the location of the sidechain based radical. Unlike other peroxidases, CCP does not have a porphyrin based radical but rather a sidechain radical, that of Trp191, in compound I as was observed previously through previous studies (57, 58, 71, 72).

In addition to serving as a model to study peroxidase function and compound I formation, CCP has been used as a model system for interprotein electron transfer. CCP and yeast cytochrome c were and still are, like HRP before them, easy to prepare in large quantities which made early studies of this redox partner pair possible. While early studies underlined the importance of electrostatic surface complementarity in CCP-Cytc complex formation (73), it was the solving of the CCP-Cytc crystal structure that began a detailed examination of the complex structure and formation (Figure 1.10). The initial complex structure reported in 1987 provided the first hints at the dynamic nature of the complex when in the crystallographic dimer, only CCP was visible (74). Given the fact that Cytc is itself stable, it was first suggested that this lack of electron density was in fact due to Cytc being not disordered internally but with respect to orientation to CCP, supporting the idea of the idea of dynamic complex association and disassociation events. The full complex structure was solved five years later in 1992 which revealed that the complex surprisingly had no strong ionic interactions at the interface given previous experiments that suggested that the complex was electrostatically driven (75). An engineered S-S crosslinked complex structure further validated the orientation and binding site of the complex revealing that a small number of water molecules served to bridge the polar groups at the interface (76).

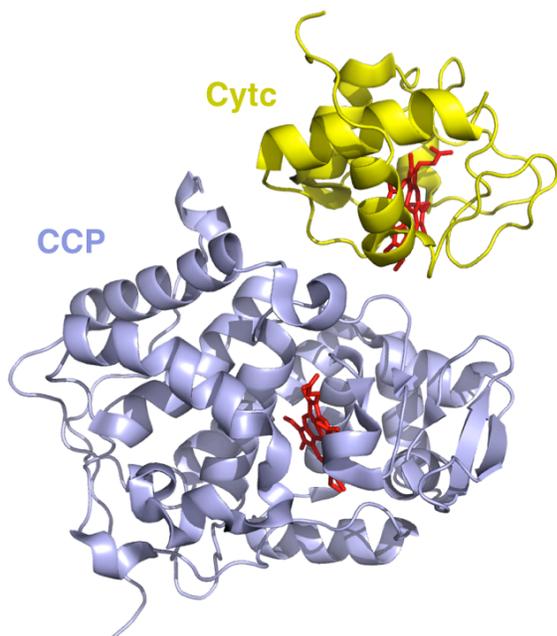


Figure 1.10. Cytochrome c binding to CCP. A) The generally accepted active complex conformation of the yeast cytochrome c peroxidase (CCP, light purple) and cytochrome c (Cyt c, yellow) complex as solved by Pelletier and Kraut.

While the complex crystal structure is now generally accepted as the active conformation, ongoing work has studied not only the dynamic nature of the complex itself but also the presence of alternate binding sites for Cyt c on CCP. Extensive experimental and theoretical work has shown that there are indeed other, lower affinity binding sites on the opposite side of CCP as the active binding site (Figure 1.10) (77-81). Given the importance of Trp191 to act as the radical site during electron transfer, it is very unlikely that this second binding site plays an important role in activity.

Intriguingly however, studies of crosslinked CCP-Cyt c complexes revealed that even when the complex is trapped, Trp191 is still reduced though at a much lower rate. These results may suggest that while other Cyt c binding sites may exist, the reactions are much slower and regardless of the “entry” point of the electron, it will find its way to Trp191 (82). While there is continuing work to study the CCP-Cyt c complex, it is also pertinent to look at similar experimental systems that through comparison may be able to shed light onto an older system.

Leishmania major is the unicellular parasitic causative agent of the tropical disease leishmaniasis (83). Upon infection, the parasite has to resist a barrage of reactive oxygen species produced by the host in an attempt to neutralize the foreign threat (84). While *L. major* has a variety of methods to battle the host response (83-90), one of the more intriguing methods of

defense is through the production of a heme peroxidase, *Leishmania major* peroxidase (or LmP) and a cytochrome c (termed LmCytc). Recent work has shown that the expression of LmP helps protect the parasite from reactive oxygen species (83), essentially turning the host's defenses against itself by using what should be a weapon to weaken or kill the invader instead as a source of further power while also removing toxic and reactive oxidative species out of the environment.

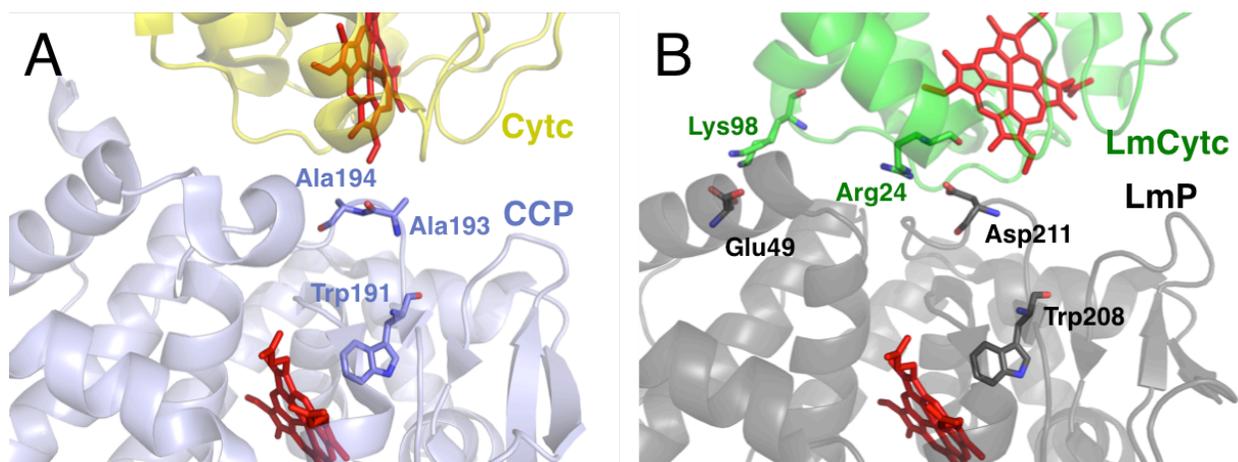


Figure 1.11. Comparison of the CCP-Cytc and LmP LmCytc complexes. A) In the model yeast CCP-Cytc interface (displayed in an identical color scheme as Figure 1.10), the transferred electron travels from the heme of Cytc to CCP through the contacting CCP residues Ala193 and Ala194 before reaching Trp191. B). In the LmP-LmCytc system, structural changes at the interface results in the transferred electron traveling from the heme to LmP through the interfacial ion pair of LmCytc R24 to LmP D211 before traveling to Trp208.

The LmP-LmCytc and CCP-Cytc systems have many similarities though do have several unique characteristics. To begin, both systems carry out identical redox chemistries and both share the same Trp based radical intermediate, suggesting they also carry out the same mechanism (91-93). However, while LmP obeys Michaelis-Menten kinetics (91, 93), yeast CCP does not. Indeed, there are many differences between the CCP and LmP kinetics that would suggest that while these proteins are close in sequence, chemistries and mechanism, their activity is somehow different. Adding further intrigue to this conundrum, the recent solved LmP-LmCytc

co-crystal structure reveals a strikingly similar complex to the yeast CCP-Cytc complex though the interface is dominated by entirely different contacts (Figure 1.11). One of the major differences between the complexes is the presence of an ionic-pair of LmP residue D211 and LmCytc residue R24 where no such interfacial ionic pair exists in CCP, resulting in very different interfaces.

Comparison between the seemingly closely related LmP and CCP systems should yield a new level of understanding of not only the novel LmP system but provide new insights into the paradigm CCP system. With this goal in mind, in chapter 4 I investigate the association and complex formation of the LmP-LmCytc complex through a combination of molecular dynamics, Brownian dynamics, mutagenesis and kinetic analysis. Through this combinatorial approach, I identify the “bind and crawl” mechanisms of association of the LmP-LmCytc complex and reveal the absence of such a mechanism in yeast CCP (94). The results not only help to shed light on the function of LmP-LmCytc, but by revealing the absence of such action in the model CCP system, can serve as a new avenue to probe how this mechanism may have been lost through evolution, thereby potentially providing further insight into the structure function relationship in not only these two peroxidases but perhaps the family as a whole.

Part III: Nitric Oxide Synthase

Nitric oxide synthase (NOS) is a heme-thiolate protein that catalyzes the oxidation of L-arginine to L-citrulline and nitric oxide (NO) via a two-step mechanism (Figure 1.12) (1, 95). NO itself is an integral signaling molecule in the nervous, immune and cardiovascular systems (96, 97). Produced by mammalian endothelial NOS (or eNOS) in a thin layer of cells on the inside of blood vessels known as the endothelium or by neuronal NOS (nNOS) in neurons in the

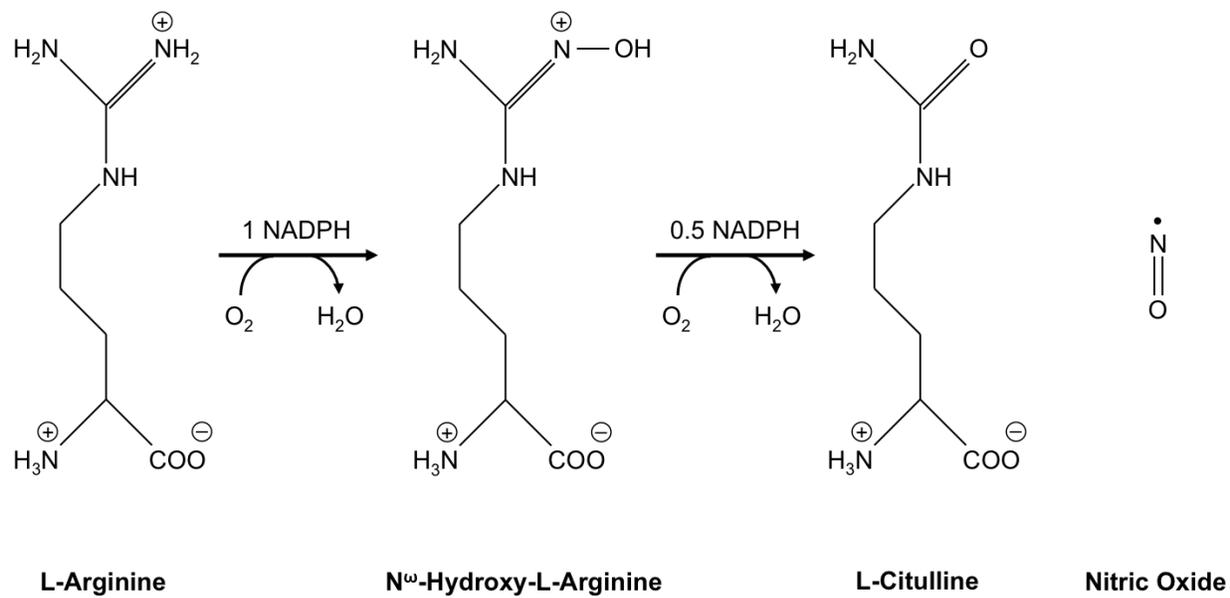


Figure 1.12. The production of nitric oxide by nitric oxide synthase. The generalized mechanism of the production of nitric oxide by the oxidation of L-arginine to L-citrulline by nitric oxide synthase via the intermediate N^ω-Hydroxy-L-Arginine.

brain, NO acts as a second messenger in the activation of guanylyl cyclase (or GC) (98-100). GC itself produces the messenger molecule cyclic guanosine monophosphate (or cGMP) that relaxes blood vessel walls resulting in vasodilation and plays critical regulatory roles in the nervous system(101). Both eNOS and nNOS have been implicated in a variety of diseases where eNOS has been a target for treatments for hypertension, blood pressure and stroke prevention (102, 103) while nNOS has not only been implicated in a variety of neurological disorders including Parkinson's and Alzheimer's diseases (104-108), but also as a novel and viable drug target for treating melanoma (109). In addition, a third mammalian NOS isoform exists in inducible NOS (iNOS) that produces NO not as a messenger molecule but rather as a cytotoxic agent as part of a greater immune response (110). Each of the three isoforms are regulated by Ca²⁺/calmodulin levels in the environment (111) where high concentrations of Ca²⁺/calmodulin activate NOS activity in eNOS and nNOS though this effect is not as prevalent in iNOS. The exact role that

calmodulin plays in modulating the function of NOS will be described in more detail after first describing the mechanism of nitric oxide production by NOS.

It is also important to note that while the focus of this introduction is on mammalian isoforms of nitric oxide synthase, other organisms have NOS enzymes that have proven to be point of interest for research, particularly the bacterial form of NOS. The bacterial form of NOS, or bNOS, has been identified to serve as an additional barrier against not only antibiotics in Gram-positive bacteria, but also against oxidative species produced by the host to attack the bacteria by its own NOS enzyme (iNOS) (112-115). Recent work has identified this bacterial enzyme as a potent target against a variety of pathogens including *Bacillus anthracis* and *Staphylococcus aureus* and a growing number of studies have probed how to design molecules to selectively target this form of NOS (116-120), including work that I participated in that is not included in this document but was nonetheless completed during my time at UCI (118).

Mechanism of NO Production by NOS

As outlined in Figure 1.12, NOS catalyzes the conversion of L-arginine to L-citrulline and NO through a two step reaction. The first step of the reaction, the conversion of L-arginine to N^o-hydroxy-L-arginine, is thought to occur in a nearly identical fashion as the hydroxylation reaction carried out by cytochrome P450 enzymes as described earlier with only small differences(1, 104). The most important difference is that the essential cofactor tetrahydrobiopterin (BH₄) acts as the source of electrons to reduce the oxy-complex to the peroxy intermediate (121-123). During this reduction process, a pterin based cation radical is formed which is thought to be returned to its previous state by the NOS flavin reductase which will be

discussed in following sections (124). The second portion of the NOS catalysis however is still not well understood despite a great deal of work to fill this knowledge gap (125-127).

Two different mechanisms have arisen that attempt to describe how the second step of the NOS mechanism occurs at a molecular level which differ on the importance and roll of the essential cofactor BH_4 in the reaction(1). Both of these proposed mechanisms are displayed in Figure 1.13. In the first of these proposed mechanisms, the BH_4 -Radical mechanism (A in Figure 1.13), an electron is donated from the BH_4 cofactor to the oxy-complex (1, 127). Upon donation of the electron, the molecular oxygen abstracts two protons, one from the solvent and one proton from the hydroxylated nitrogen of N^ω -hydroxy-L-arginine which results in the loss of resonance in the remaining guanidinium group. Loss of the hydroxyl proton then results in the formation of a bond between the central sidechain carbon atom and the proposed compound I intermediate, which quickly forms a double bond with the resulting loss of nitric oxide and final L-citrulline product. In the second proposed mechanism, the No- BH_4 -Radical mechanism, the BH_4 cofactor plays no such role in the catalysis (1). Instead, the electron is donated by the substrate though abstracting of the NH proton resulting in a hydroperoxy intermediate. Following the abstraction, the proximal oxygen atom attacks the substrate central carbon atom carbon accompanied by the loss of two protons and leaving negative radical oxygen where the hydroxyl group originated. This intermediate undergoes internal electron rearrangements resulting in the loss of nitric oxide, L-citrulline and through the gain of two protons, the loss of a water. While both mechanisms are balanced and each have their arguments for why they match experimental data, no firm conclusion has been reached as to which of these proposed mechanisms, or any yet proposed alternative, is indeed correct and remains a point of study for the field at large (122, 126, 128).

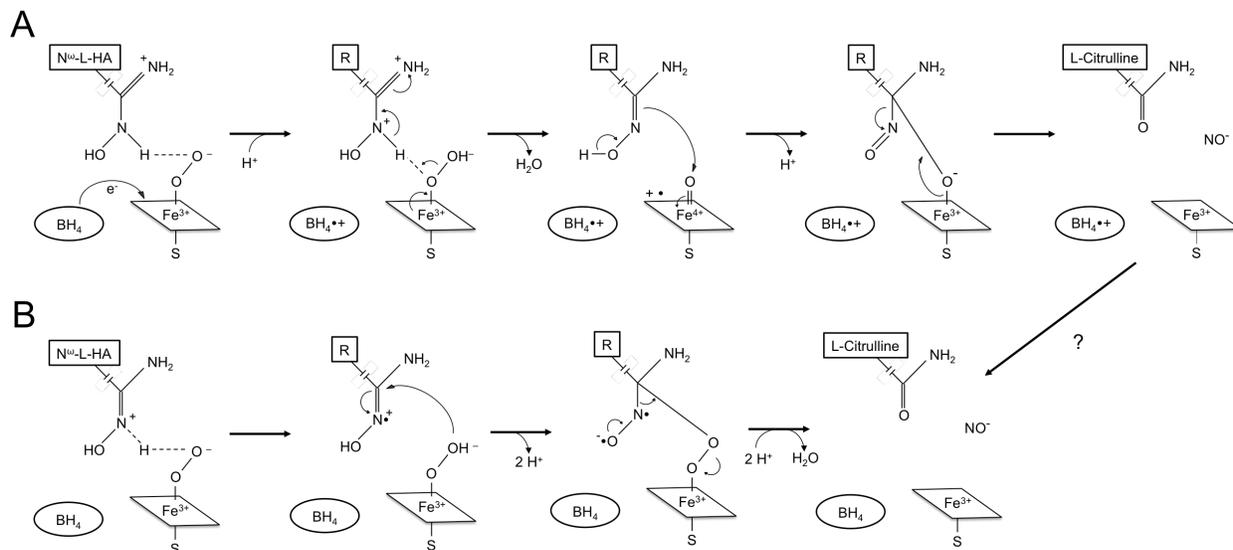


Figure 1.13. Proposed mechanisms for the oxidation of N^ω-hydroxy-L-arginine to L-citrulline. Unlike the conversion of L-Arginine to N^ω-Hydroxy-L-Arginine (which proceeds identical to a P450 reaction), the conversion of N^ω-Hydroxy-L-Arginine to L-citrulline and nitric oxide synthase is not well understood. Based on existing experimental study, two mechanisms have been proposed which are displayed here; A) The BH₄ Radical and B) The No BH₄ Radical mechanism.

NOS Structure

While there are many remaining questions concerning the full architecture of the holo-NOS enzyme in solution, extensive experimental work has been able to shed a great deal of light on its structure. To begin, NOS itself exists as a dimer in solution, where each monomer is constructed of a single polypeptide chain that included multiple domains (Figure 1.14) (1, 95). A single monomer contains three separate domains (in order from N-to-C terminus); the flavin adenine dinucleotide (FAD)-containing FAD domain, the flavin mononucleotide (FMN)-containing FMN domain, and the heme containing heme or oxygenase domain. When the FMN and FAD domains form a tight complex, combined it is referred to as the reductase domain. Dimerization of the monomers occur at the heme domain resulting in the holo-NOS enzyme. Despite extensive crystallographic work to solve the structures of each domain for a single

mammalian isoform, no such set exists though a structure of each domain has been solved across all isoforms.

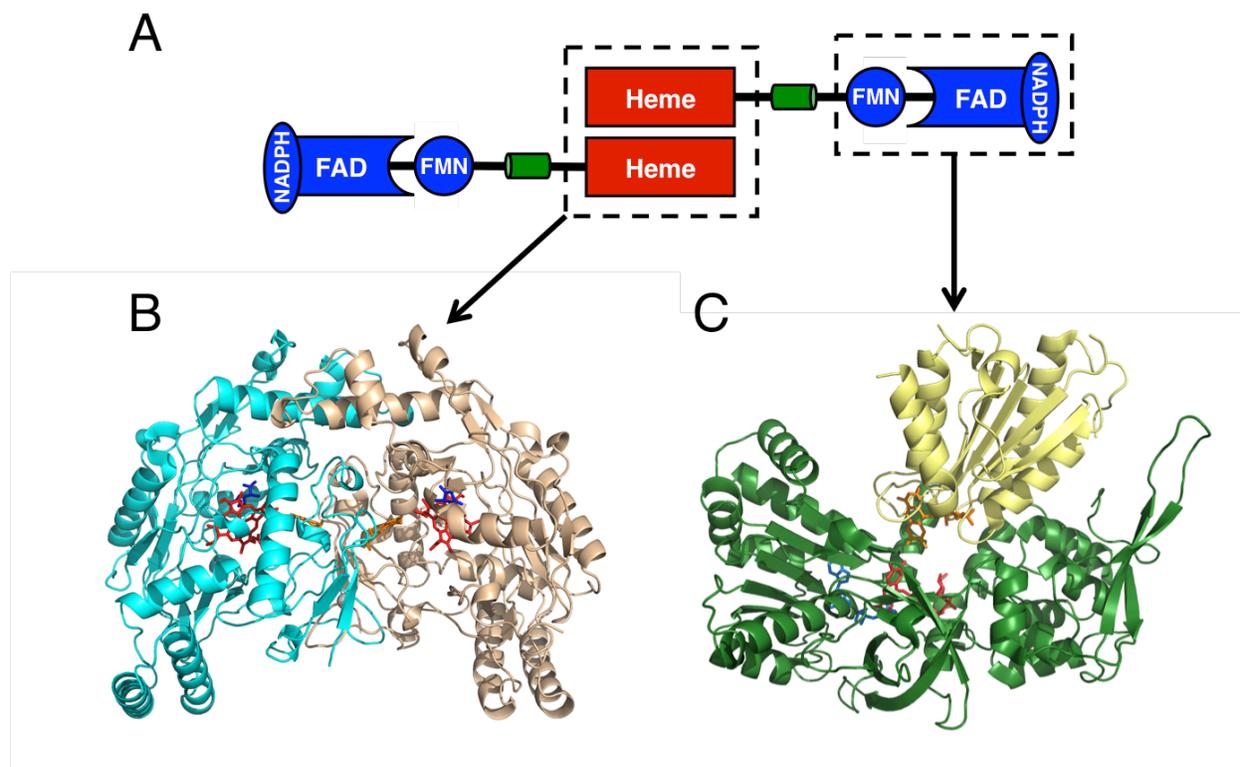


Figure 1.14. Architecture of nitric oxide synthase. A) Cartoon representation of the holo-NOS dimer domain architecture in the input or extended conformational state. B) Crystal structure of the iNOS heme or oxygenase domain dimer (PDBID: 1NSI) where each monomer is colored cyan and light brown. The heme, L-arginine substrate and BH₄ cofactor are shown as red, blue and orange sticks respectively. C) Crystal structure of the reductase domain (FMN and FAD domains in complex) from nNOS (PDBID: 1TLL) with the FMN and FAD domains in yellow and green respectively. The FMN, FAD and NADP cofactors are shown as orange, red and blue sticks respectively.

While the architecture of NOS does recall many of the same features that have previously P450 enzymes, there are several important differences. To begin, while P450 redox partners which include the FMN and FAD cofactors exist as separate polypeptides from the heme domain, all three domains exist on the same polypeptide chain in the NOS enzyme though both enzymes follow the same pattern of electrons flowing from NAD(P)H to FAD to FMN and finally to the heme. Next, P450 enzyme active sites are dynamic, particularly in the opening and

closing motions of the substrate access channel that can prevent solvent from entering the active site upon substrate binding. NOs enzymes on the other hand undergo no such opening and closing motions as the active site remains far more exposed to solvent regardless of substrate binding. Finally, while both enzymes share a Cys-ligated heme at the heart of the enzyme active site, P450 enzymes rely on larger contributions from nonpolar interactions in substrate binding where NOS substrate participate in polar interactions with not only the protein but the heme propionate.

NOS Electron Transfer and Regulation by Calmodulin

To power the two step conversion of L-arginine to L-citrulline and NO, NOS accepts two electrons from NAD(P)H which then undergo two further electron transfer steps within the enzyme itself. To begin, electrons are donated from NAD(P)H to the FAD cofactor in the FAD domain. In this initial stage of accepting electrons, the FAD domain is bound to the FMN domain to make up the full reductase domain which itself is not thought to make contact with the oxygenase domain dimer. This conformation is referred to as the input state or extended state (Figure 1.15) (129). Following the first electron transfer, electrons will then move from the FAD cofactor and domain to the FMN cofactor and domain. Before the third and final electron transfer reaction can take place, NOS itself undergoes a significant conformational change wherein the reductase domain disassociates and the FMN domain moves to form a complex with the heme domain of the opposite monomer (Figure 1.15) (130). This new conformation is termed the output state. In the output state, electrons are transferred from the FMN to the heme group by way of a Trp residue that is positioned between the cofactors.

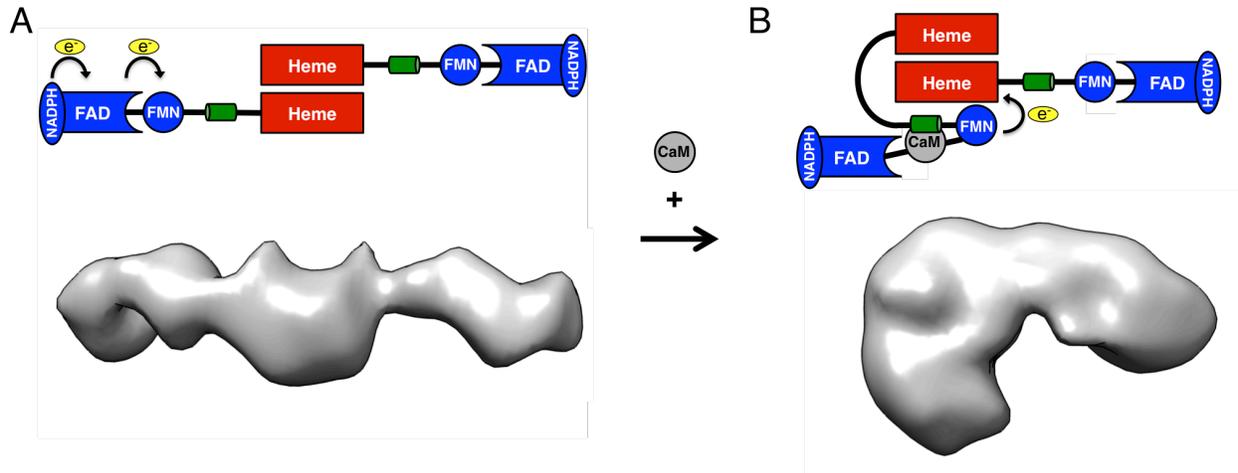


Figure 1.15. Conformational changes in nitric oxide synthase from input to output states. A) In the NOS input state, NADPH binds to the FAD domain which itself is bound to the FMN domain making up the full reductase domain. Electrons are transferred from NADPH to the FAD and finally to the FMN cofactor. The input state is shown as a cartoon (top) as well as visualized through electron microscopy (bottom, EMMAPID:XXXX). B) Upon binding of calmodulin (CaM), the reductase domain of a single monomer breaks so the FMN domain can bind to the heme domain of the opposite monomer to transfer an electron from FMN to the heme, termed as the output state. Both the cartoon (top) and low resolution electron microscopy architecture of the NOS output state are shown.

The constitutive eNOS and nNOS isoforms are regulated by the presence of Ca^{2+} -calmodulin (referred to as CaM henceforth). CaM binds to a helical region located between the FMN and heme domains of NOS and facilitates electron transfer between these two domains in the output state (131). It has been predicted that the binding of CaM promoted the output state by restricting conformational freedom of each monomer and stabilizing the output state (132, 133). Evidence for the conformational change that occurs upon CaM binding has been observed by FMN fluorescence and mutagenesis (134-136). Being able to tightly regulate this final electron transfer reaction is important because if the electron transfer is uncoupled, dangerous peroxy species can form and cause serious damage to the cell (107, 137, 138). In contrast to eNOS and nNOS however, iNOS binds CaM tightly and retains the ability to transition between the input and output states at low levels of Ca^{2+} (139). Despite the unique functions, each of the three

mammalian isoforms of NOS demonstrate a striking degree of conserved secondary and tertiary structure (Figure 1.16).

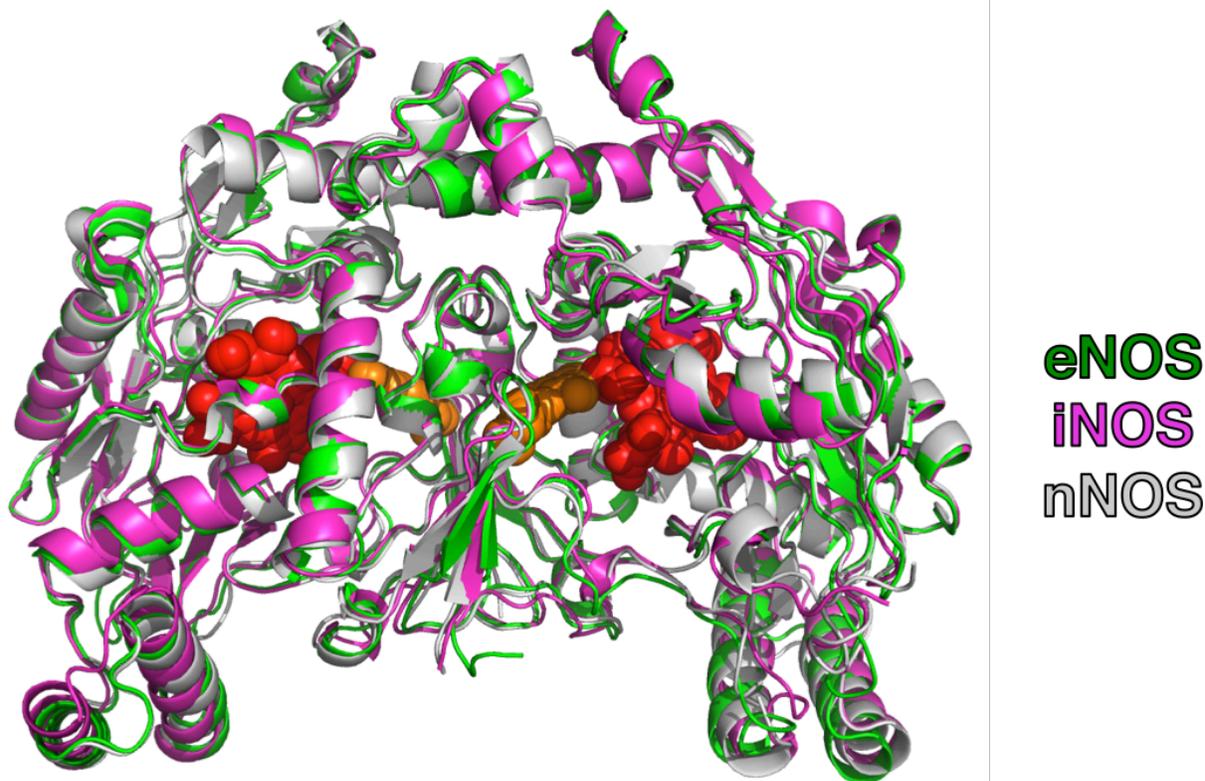


Figure 1.16. Structural overlay of the three mammalian NOS isoforms. Shown is a structural overlay of representative structure of endothelial NOS, inducible NOS, and neuronal NOS to display the high degree of structural similarity between the three isoforms.

To address how these conformational changes may occur it is important to solve the structure of holo-NOS which to date has remained unsolved. Despite a growing collection of crystal structures of individual NOS domains (140-146), no single atomic structure of holo-NOS exists. However, a series of recent experimental studies carried have for the first time provided a glimpse of the structure of holo-NOs at low resolution. The first of these studies, which utilized hydrogen-deuterium exchange of FMN and heme domains as well as CaM, was able to begin to map out the regions of each of these domains that contacted each of the other domains (147).

This study was followed by three independent studies that studied the low resolution structure of

holo-NOS in both the input and output state through electron-microscopy methods (129, 148, 149).

Despite the new advances and findings attempting to unravel the mystery of the holo-NOS structure, there has been to date no study to produce an atomic-level of detail model of NOS in the important output state following the disassociation of the reductase domain. In chapter five, such a study is documented that uses the recent hydrogen-deuterium exchange data to produce such a model that is then tested using molecular dynamics and further validated by comparison of many of the experimental studies that have been carried out on holo-NOS over the past several decades in an attempt to bring more of the holo-NOS structure into focus.

Reproduced with permission from Hollingsworth S.A. and Poulos T.L. (2015). Molecular dynamics of the P450cam-Pdx complex reveals complex stability and novel interface contacts. *Protein Science*. **24**: 49-57. Copyright 2015 The Protein Society.

Chapter 2

Molecular Dynamics of the P450cam-Pdx Complex Reveals Complex Stability and Novel Interface Contacts

Summary

Cytochrome P450cam catalyzes the stereo and regio-specific hydroxylation of camphor to 5-*exo*-hydroxylcamphor. The two electrons for the oxidation of camphor are provided by putidaredoxin (Pdx), a Fe₂S₂ containing protein. Two recent crystal structures of the P450cam-Pdx complex, one solved with the aid of covalent cross-linking and one without, have provided a structural picture of the redox partner interaction. To study the stability of the complex structure as well as the minor differences between the recent crystal structures, a 100 nanosecond molecular dynamics simulation of the cross-linked structure, mutated *in silico* to wild type and the linker molecule removed, was carried out. The complex was stable over the course of the simulation though conformational changes including the movement of the C helix of P450cam further towards Pdx allowed for the formation of a number of new contacts at the complex interface that remained stable throughout the simulation. While several minor crystal contacts were lost in the simulation, all major contacts that had been experimentally studied previously were maintained. The equilibrated MD structure contained a mixture of contacts resembling both the cross-linked and non-covalent structures as well as the newly identified interactions. Finally, the reformation of the P450cam Asp251-Arg186 ion pair in the MD simulation mirrors the ion pair observed in the more promiscuous CYP101D1 and suggests that the Asp251-Arg186 ion pair may be important.

Introduction

Cytochromes P450 (P450 or CYPs) are heme-containing protein monooxygenases that catalyze the oxidation of a wide range of organic compounds (1, 150). Prevalent in all kingdoms of life, there have been over 18,000 CYP genes identified with at least 57 CYP genes reported in humans (1, 46, 151). While diverse in function and substrate specificity, P450 enzymes activate molecular dioxygen for the following generalized hydroxylation reaction:



Due in no small part to their different specificities for substrates and reactions, P450s have been a large focus in the fields of pharmaceutical research, biosensor and bioconversion among many others (152-155).

P450cam (CYP101A1) from *Pseudomonas putida*, one of the most extensively characterized P450 enzymes in terms of structure and function, catalyzes the stereo- and regiospecific hydroxylation of camphor to 5-hydroxycamphor (156, 157). Like almost all P450 enzymes, P450cam accepts electrons from a redox partner to initiate the hydroxylation reaction. In the P450cam system, electrons are shuttled from NADH by the flavin adenine dinucleotide (FAD)-containing putidaredoxin reductase (Pdr) to the Fe₂S₂-containing putidaredoxin (Pdx), which then delivers two electrons, one at a time, to P450cam (54, 156). Previous work has shown that Pdx plays an effector role upon binding to P450cam for the donation of the second electron that induces structural changes in P450cam (53, 54). In addition to the known structures of all three individual components of this relay pathway, the co-crystal structure of Pdr-Pdx has also been solved, leaving only one complex left unsolved in the P450cam electron transfer pathway (43, 158-161). Despite a great deal of effort, however, the P450cam-Pdx complex structure had remained elusive until very recently.

In two recent and separate studies, the three-dimensional structure of the P450cam-Pdx complex was solved through x-ray crystallography and NMR spectroscopy. Using covalent cross-linking chemistry, Tripathi and coworkers were able to trap the complex through flexible linkers and select mutagenesis resulting in the first structure of P450cam-Pdx (henceforth referred to as the cross-linked structure) (48). Immediately following this finding, Hiruma and coworkers solved the complex structure without the use of covalent chemistry employing only a series of mutations on both P450cam and Pdx (henceforth referred to as the non-covalent structure) in addition to the NMR structure (47). While solved independently and utilizing different approaches, the complexes are the same. The crystal structures showed significant differences to the working computational models that had been developed previously, notably that Pdx rotated by ninety degrees from what had been predicted (50). Both structures also revealed Pdx binding to the open form of P450cam, which is consistent with extensive previous studies (50-52, 162, 163). However, there are several differences between the cross-linked and non-covalent structures at the important P450cam-Pdx interface. As the complex is believed to be very dynamic, it is possible that these small differences could be hints at the dynamic nature of the interactions between P450cam and Pdx. As molecular dynamics has proven to be a useful tool in studying both P450 individually as well as in related redox pairs, this method could provide a great deal of insight into the newly solved structures (160, 164, 165).

As each of these structures represents a snapshot of the possible complex trapped in a crystal, it is imperative to study this complex in a dynamic fashion to not only reveal what contacts remain in an equilibrated structure but if any new contacts emerge through simulation that may have been disallowed in the crystal. What we present here is such an analysis. We carried out a 100ns molecular dynamics simulation of the P450cam-Pdx complex to study the

complex's stability and interface. The resulting equilibrated structure revealed not only a hybrid complex of both crystal structures, but several new contacts not seen in either structure appear with small conformational arrangements that may represent previously unobserved important structural contacts. In addition, the reformation of the Arg186-Asp251 ion pair in the MD simulation, broken in all known structures of the complex, has led to a reevaluation of the electron density of the previously solved crystal structures with regards to the proton relay network.

Results

The cross-link free reduced P450cam-Pdx complex solved by Tripathi et al (48) was used as the starting point of the 100ns simulation. Once the initial minimization restraints were removed from the system, both P450cam and Pdx remained both individually stable with few large scale conformational changes while the complex itself showed a remarkable degree of stability over the course of the simulation (Figure 2.1). The changes in root-mean squared displacement or RMSD (Figure 2.1A) upon the beginning of the simulation were caused by minor conformational changes in the loop regions on both P450cam and Pdx with one notable exception. Helix C of P450cam was found to move towards Pdx by approximately 5Å early in the simulation and remains stable in its new conformation throughout the simulation. Root mean squared fluctuation (RMSF) analysis of the C α atoms of each protein reinforce the relative stability of each protein as shown in Figure 2.1B & C, where the most dynamic portions of either P450cam or Pdx occur on loops far removed from either active site and the complex interface.

While the RMSD suggests that minor changes occurred in the complex structure, a structural overlay of both the cross-linked and non-covalent crystal structures with an

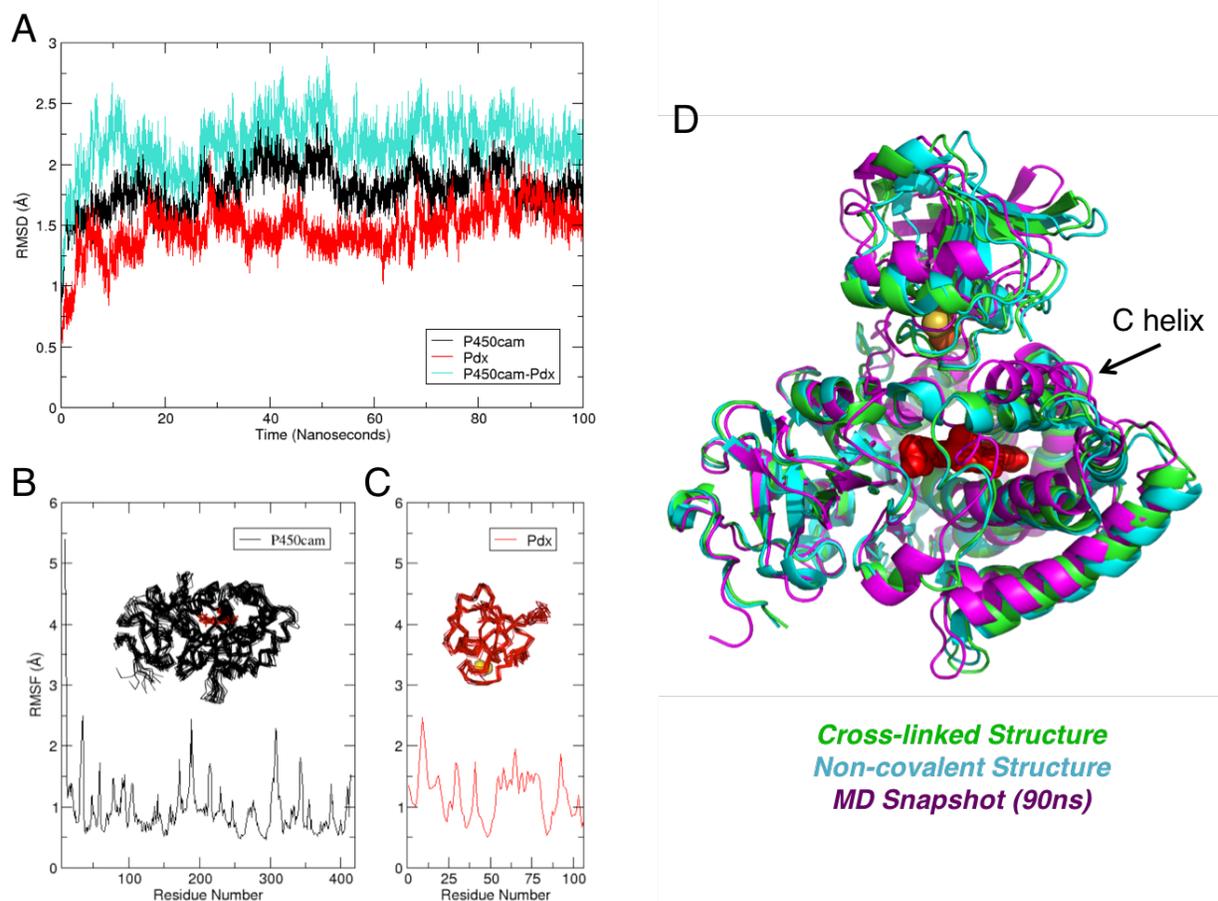


Figure 2.1. Stability of the P450cam-Pdx throughout simulation. A) Shown are the RMSD evolutions for the backbone C_{α} atoms of P450cam (black), Pdx (Red) and the full P450cam-Pdx complex (Teal) throughout the course of the 100 nanosecond simulation. B) Calculated RMSF values for each C_{α} atom and an overlay of snapshots from every 10ns throughout the simulation for both P450cam and C) Pdx. Both P450cam and Pdx show a high degree of stability individually as well as only minor changes as a complex. D) The three P450 structures, the cross-linked reduced of Tripathi et al. (green), the non-covalent complex of Hiruma et al (Cyan) and a snapshot taken from near the end of the 100ns simulation presented in this report (magenta) are overlaid by aligning all three along the backbone of P450cam. While the crystal structures overlap greatly, the MD structure shows significant differences in both helix C and a rotation of Pdx relative to P450cam in addition to small conformational changes in loops at the P450cam-Pdx interface.

equilibrated structure from the 100 nanosecond trajectory reveals several new changes in the structure (Figure 2.1D). Several of the loops of both P450cam and Pdx opposite from the helix C movement noted above have undergone changes to generate new contacts between the redox

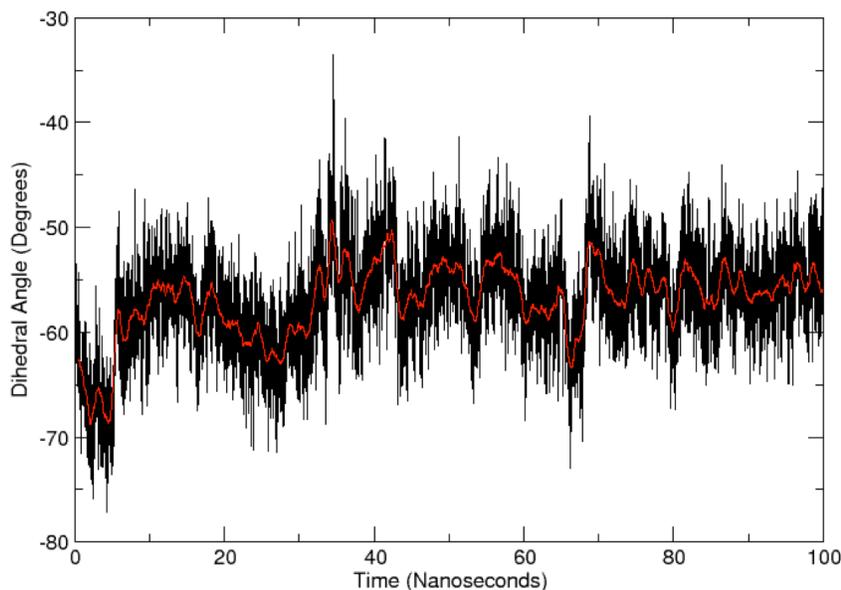


Figure 2.2. Rotation of Pdx in relation to P450cam. Pdx undergoes a small rotation with a fulcrum at the Fe_2S_2 cluster, monitored by calculated the dihedral angle formed by the NA and NC atoms of the P450cam heme and the C_α atoms of V6 and K2 of the central β -strand of a three stranded β -sheet of Pdx which was fairly stable through the simulation, causing it to be chosen to study the movement of Pdx relative to the heme.

partners. In addition, Pdx has undergone a small rotation centered at a fulcrum of the iron-sulfur cluster. To monitor this rotation evolution, the dihedral formed by the C_α atoms of Pdx_{Val6} and Pdx_{Lys2} making up a stable β -strand away from the interface and two of the iron-coordinating nitrogen atoms of the P450cam heme was calculated over the simulation which can be seen in Figure 2.2. As Pdx was shown to be very stable through both RMSD and RMSF analysis, the beta strand was chosen to study the Pdx rotation. The resulting Pdx orientation began near a value of -66° before equilibrating near 40ns to a value near -56° , resulting in a ten degree rotation of Pdx in relation to the heme group of P450cam. The mean distance from the iron-sulfur cluster to the heme iron over the final 25 nanoseconds was found to be $16.05\text{\AA} \pm 0.27\text{\AA}$ (one sigma), showing good agreement with both the cross linked (16.1\AA) and non-covalent

(16.3Å) crystal structures demonstrating that this rotation does not affect the heme-to-cluster distance.

As the interface between P450cam and Pdx is of great importance, time averaged contacts defined by a distance cutoff of 3.5Å between non-hydrogen atoms averaged over time of the simulation were calculated over the final 25 nanoseconds of the simulation as well as over the full simulation to be directly compared to those observed in both the cross linked and non-covalent structures. This analysis over the final equilibrated 25 nanoseconds is presented in Table 2.1 by summing all contacts made between any given pair of residues. Of the top time averaged contacts, only six and four of the contacts had been observed in the non-covalent and cross-linked structures respectively with an additional nine novel contacts. Further analysis of the contacts observed in either the cross-linked or non-covalent structures revealed that a number of contacts were broken early in the simulation and replaced by a host of new contact pairs, which are summarized by Table 2.2 and in more detail in Appendix A.

Table 2.1. Top Time Averaged Contact Pairs Over the Final 25 Nanoseconds

P450cam Residue	Pdx Residue	Time Averaged Contact Pairs	Previously Documented*
Glu76	Ser42	3.59	
Arg109	Cys45	2.9396	CL
Val345	Gln25	2.39	
Arg112	Asp38	2.3604	CL, NC
Gln360	Asp38	2.11	
Arg109	Trp106	1.7171	CL, NC
Asn116	Trp106	1.6392	CL, NC
His352	Gly40	1.4621	
Arg109	Gln105	1.4154	CL
Lys344	Gly20	1.3913	
Ala113	Trp106	1.3179	CL, NC
Gly353	Gly41	1.2992	
Gln343	Ala18	1.2017	
Gly353	Gly40	1.0825	
Gln343	Val17	1.0096	

* – CL, cross-linked structure (4JX1); NC, non-covalent structure (3W9C)

Table 2.2. Selected comparisons of the interactions at the interface of P450cam and Pdx

P450cam Residue	P450cam Atom	Pdx Residue	Pdx Atom	Cross-Linked Distance (Å)	Non-Covalent Distance (Å)	MD Distance & σ of Final 25ns (Å)
Glu76	OE1	Ser42	OG	5.8	7.0	3.43 (0.81)
Glu76	O	Arg66	NH1	2.8	2.9	6.60 (1.85)
Arg109	NH2	Cys45	O	3.0	6.3	2.89 (0.15)
Arg109	NH2	Thr47	CG2	3.3	9.5	3.52 (0.24)
Arg109	NH2	Gln105	O	3.0	9.8	2.88 (0.13)
Arg109	NH1	Trp106	OXT	4.8	9.6	3.70 (0.98)
Arg112	NH2	Asp38	OD2	3.0	3.4	2.88 (0.22)
Arg112	NH2	Asp38	CG	3.3	3.8	3.86 (0.27)
Ala113	CB	Trp106	CE3	3.5	3.6	3.55 (0.18)
Asn116	OD1	Trp106	NE1	3.1	5.0	3.28 (0.43)
Asp125	OD2	Tyr33	CZ	3.5	5.5	5.40 (0.71)
Asp125	CB	Tyr33	OH	3.5	3.2	4.33 (0.72)
Gln343	OE1	Ala18	N	15.3	13.4	3.14 (0.49)
Val345	O	Gln25	NE2	8.1	11.2	3.17 (0.42)
His352	CE1	Gly40	O	6.6	7.0	3.34 (0.49)
Gly353	O	Cys39	O	3.6	4.3	4.56 (0.31)
Gly353	CA	Ser42	OG	3.5	4.0	8.28 (0.55)
Gly353	O	Gly40	N	5.6	6.2	3.23 (0.32)
Gly353	O	Gly41	N	6.7	7.0	3.24 (0.49)
Leu356	CB	Cys39	O	3.4	3.5	7.52 (0.20)
Gln360	NE2	Asp38	OD1	6.7	8.2	3.50 (0.68)

Focusing on the contacts that had been previously described by the two crystal structures, the final MD structure resembles a cross between both the cross-linked and non-covalent structures. The salt-bridge of P450cam_{Arg112} and Pdx_{Asp38} is especially important in the second electron transfer step, reduction of the oxy-complex to generate Compound I that then hydroxylates the substrate (54, 166, 167). Other major electrostatic interactions observed in both structures, however, are weakened in the MD structure when compared to either crystal structure.

The P450cam_{Glu76}-Pdx_{Arg66} interaction (48, 54). observed in both structures is broken quickly in the MD structure and never regained, equilibrating at a distance nearly double that of either crystal distance. The third electrostatic between P450cam_{Asp125}-Pdx_{Tyr33}^(47, 54) is weakened from what was seen in the cross-linked structure and closely resembles the non-covalent structure at the end of the simulation. The final three contact pairs previously described, which all include P450cam_{Trp106}, serve to further reinforce the importance of that residue in partner binding and activity (54, 168). The interfacial hydrogen bond that was observed in the cross-linked structure, P450cam_{Asn116}-Pdx_{Trp106}, was conserved throughout the full simulation. Finally, the major van der Waals interactions from the cross-linked structure such as P450cam_{Ala113}-Pdx_{Trp106} and P450cam_{Arg109}-Pdx_{Trp106} more closely resemble the cross-linked structure after equilibration and 100ns of simulation.

In addition to the small number of contacts that had been previously described in the crystal structure, there were a number of new contacts that were formed in the simulation (Figure 2.3). The strongest of these new interactions was a newly formed hydrogen-bond between P450cam_{Glu76} and Pdx_{Ser42} highlighted in Figure 2.3A. This contact was formed owing to a conformational change on a loop of Pdx that occurred within the first five nanoseconds of the simulation. The interaction is very dynamic as it switches between a single and bifurcated hydrogen-bond throughout the simulation. A similar rearrangement on a Pdx-facing surface exposed loop of P450cam and a corresponding Pdx loop resulted in the formation of two additional hydrogen bonds between P450cam_{Val345}-Pdx_{Gln25} and P450cam_{Gln343}-Pdx_{Ala18} shown in Figure 2.3B. These interactions are significant as in both crystal structures the corresponding distances between these contact pairs were 9.65Å and 14.35Å, respectively. This rearrangement represents the largest change at the P450cam-Pdx interface in the MD structure. Finally, a

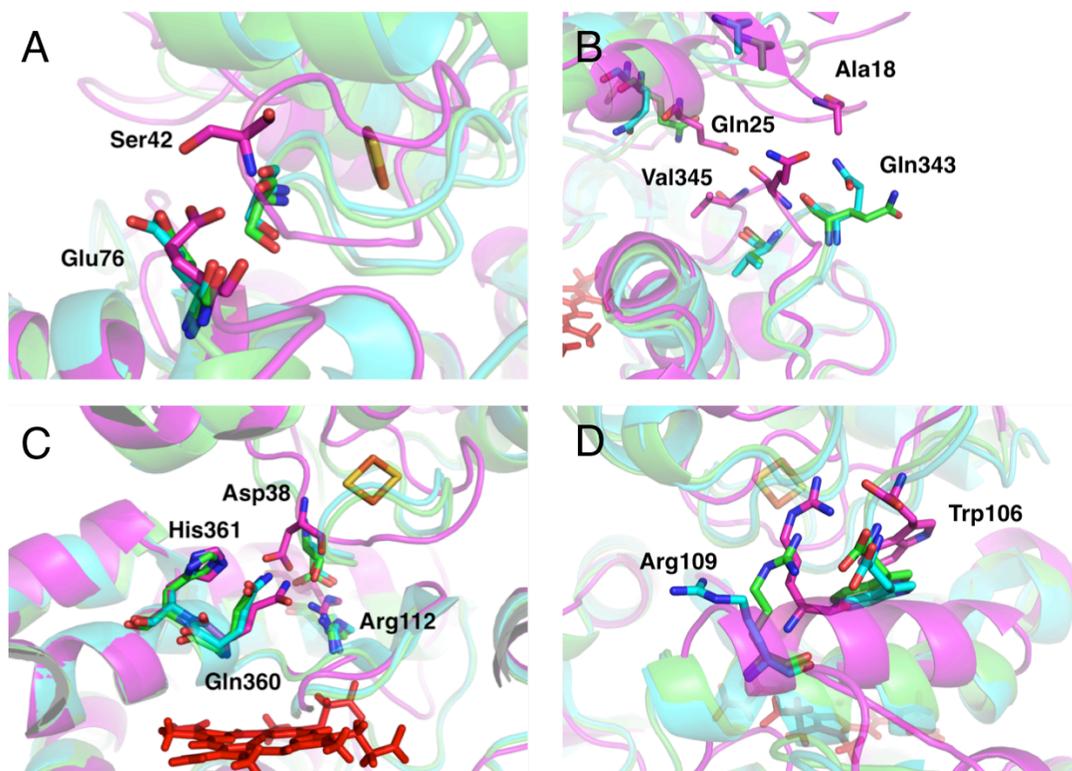


Figure 2.3. New contacts formed in the P450cam-Pdx complex. Shown are several of the new contacts that were formed in the P450cam-Pdx complex that were not observed in either of the published crystal structures. In all figures, the cross-linked structure is shown in green, non-covalent structure in cyan and the final MD structure in magenta. A) Conformational changes in a Pdx interface loop allow for the formation of an interaction between P450cam_{Glu76} and Pdx_{Ser42} that exists as either a single or bifurcated hydrogen bond. B) Loop rearrangements on both P450cam and Pdx allow for the formation of a pair of new hydrogen bonds; P450cam_{Gln343}-Pdx_{Ala18} and P450cam_{Val345}-Pdx_{Gln25}. C) Two P450 neighboring residues interact with the previously identified important intermolecular salt bridge of P450cam_{Arg112} and Pdx_{Asp38}. The new interactions are caused when Pdx_{Asp38} rotates away from its salt bridge partner. The original salt bridge remains a strong interaction despite the addition of two additional interacting residues. D) Helix C of P450cam moves closer to Pdx than what was observed in either crystal structure to allow for a stronger interaction between P450cam_{Arg109} and Pdx_{Trp106}.

change in the iron-sulfur cluster coordinating loop of Pdx allowed for P450cam_{Gly353} to a hydrogen bond to either Pdx_{Gly40} or Pdx_{Gly41} though a bifurcated hydrogen-bond with both residues was never observed (Figure 2.4). However, unlike the hydrogen bonds described previously, this interaction could also be characterized as a van der Waal interaction with a transient hydrogen bond that is closely connected to the van der Waal interaction of P450cam_{His352}-Pdx_{Gly40} which is also highlighted in Figure 2.4.

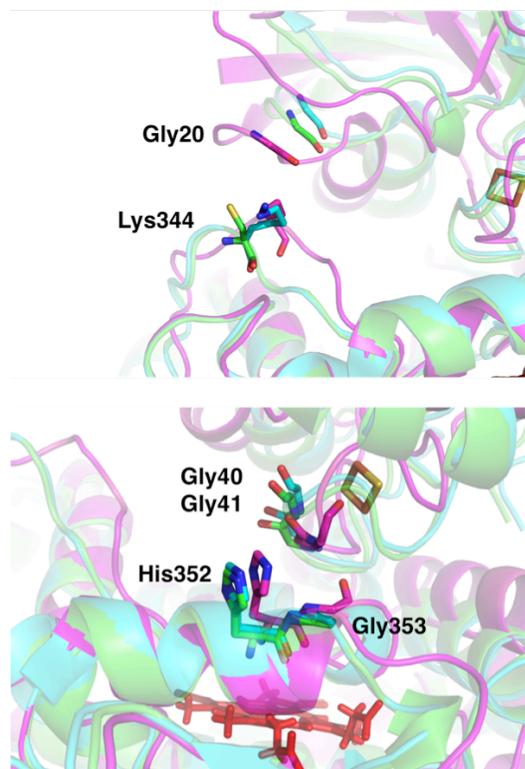


Figure 2.4. Additional contacts in the P450cam-Pdx complex. Two additional minor interfacial contacts identified in the MD simulation in addition to those identified in Figure 3 are shown. A) Conformational changes in loops in both P450cam and Pdx allow for the formation of a van der Waals interaction between P450cam_{Lys344} and Pdx_{Gly20}. B) A small conformational change in the iron-sulfur coordinating loop of Pdx allows for a network of interactions involving the Pdx dipeptide of Gly40 and Gly41. P450cam_{His352} interacts with the Pdx Gly-Gly dipeptide through van der Waals interactions while a weak hydrogen bond between either but never both of the Gly-Gly dipeptide transiently forms with P450cam_{Gly353}.

Two of the newly identified contacts surprisingly interact with the P450cam_{Arg112}-Pdx_{Asp38} salt bridge that has been previously identified to be important. While the salt bridge remains one of the strongest contacts in the MD simulation, P450cam_{Gln360} and P450cam_{His361} interact with Pdx_{Asp38} as well (Figure 2.3C). The first interaction of the two is created due to the breaking of the van der Waals interaction between P450cam_{Gln360} and Pdx_{Gly40} that was observed in the non-covalent structure in favor for the interaction of the polar glutamine with the negatively charged aspartic acid of Pdx. After the formation of this new interaction that supplements the previously described salt-bridge, P450cam_{His361} begins to transiently hydrogen bond with the side chain of Pdx_{Asp38} (Figure 2.5). However, despite both of these new interactions with Asp38, the original salt bridge remains one of the strongest contact pairs in the complex.

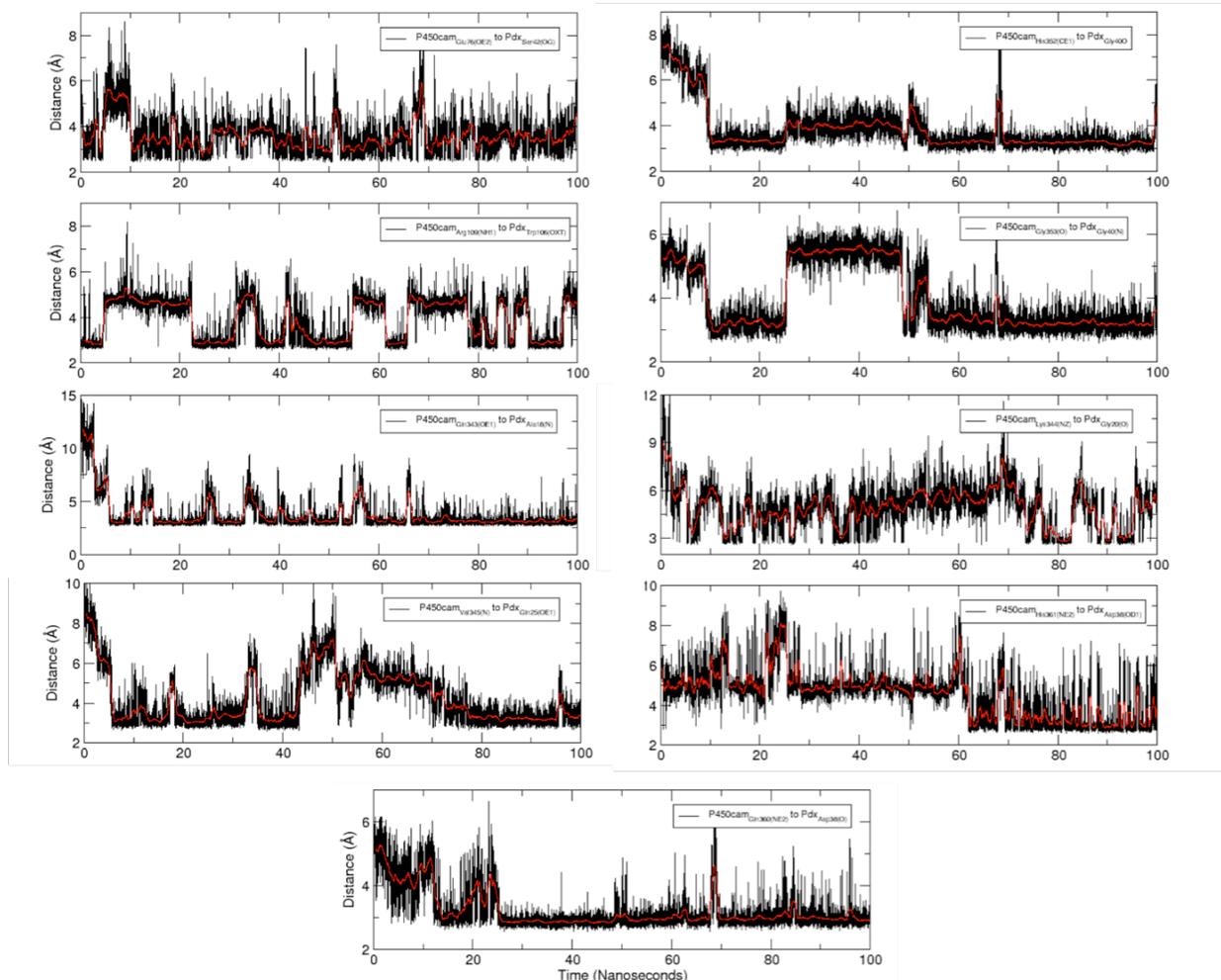


Figure 2.5. Distance evolution of the five newly formed MD contacts. Shown are the time evolutions of eight of the newly identified contacts formed between P450cam and Pdx throughout the simulation. The distance was calculated every 10 picoseconds for the duration of the simulation and plotted in black while a running average is shown in red. From top to bottom, the interactions are; (left column) P450cam_{Glu76(OE2)} to Pdx_{Ser42(OG)}, P450cam_{Arg109(NH1)} to Pdx_{Trp106(OXT)}, P450cam_{Gln343(OE1)} to Pdx_{Ala18(N)}, P450cam_{Val345(N)} to Pdx_{Gln25(OE1)}; (right column) P450cam_{His352(CE1)} to Pdx_{Gly40(O)}, P450cam_{Gly353(O)} to Pdx_{Gly40(N)}, P450cam_{Lys344(NZ)} to Pdx_{Gly20(O)}, P450cam_{His361(NE2)} to Pdx_{Asp38(OD1)}; (bottom) P450cam_{Gln360(NE2)} to Pdx_{Asp38(O)}.

Additionally, one contact that had been previously observed in both crystal structures, P450cam_{Arg109}-Pdx_{Trp106}, was significantly strengthened due to the movement of P450cam helix C towards Pdx as was described previously (Figure 2.3D). While a movement of helix C towards Pdx was observed in both crystal structures, the movement of helix C towards Pdx continues further towards Pdx in the MD simulation. Through the movement of helix C, P450cam_{Arg109} is

able to reorient towards $\text{Pd}_{\text{XTrp106}}$ and allow for the formation of a transient hydrogen bond interaction in addition to the van der Waals interactions that had been previously described.

While each protein remained stable throughout the simulation within P450cam itself, the environment surrounding the essential Asp251, which participates in the proton relay network (169), undergoes a possibly important change. Upon the beginning of the simulation, the Asp251-Arg186 ion pair that breaks in the crystal structure was reformed and remained in contact for the majority of the simulation though it was again temporarily broken twice before reforming (Figure 2.6). However, the second ion pair of Asp251-Lys178 that was broken in

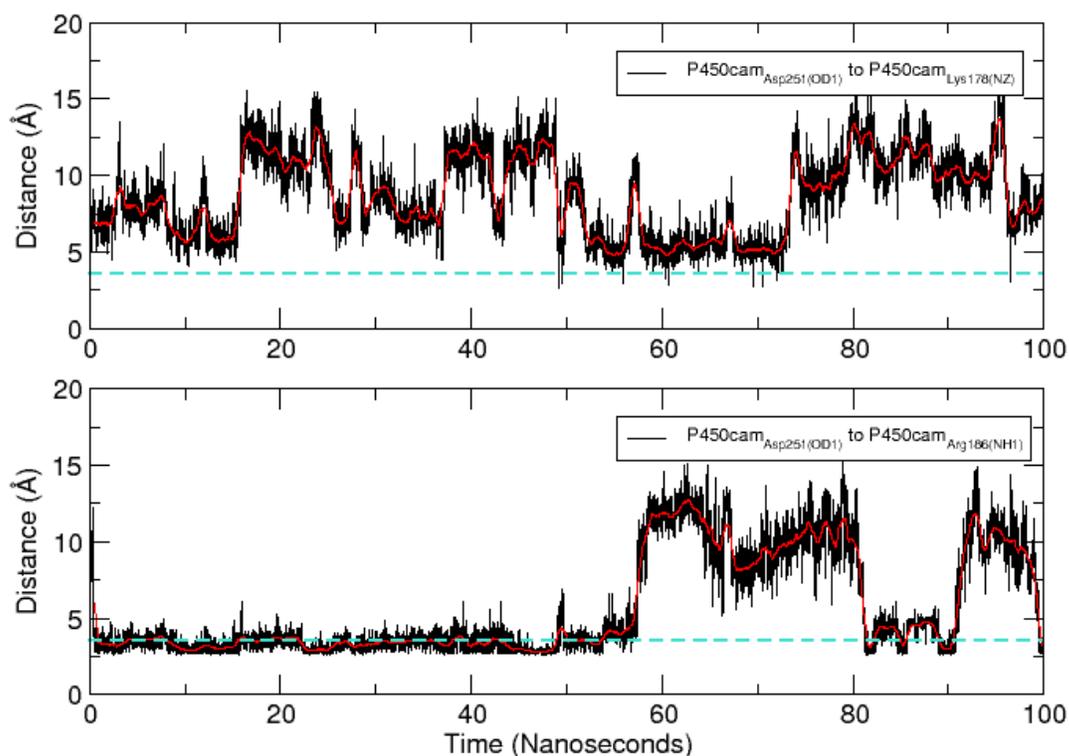


Figure 2.6. Distance evolution of the Asp251 ion pairs in P450cam. Shown are the time evolutions of the two ion pairs that prevent Asp251 from participating in the proton relay network; Asp251-Lys178 (top; from Asp251_{OD1} to Lys178_{NZ}) & Asp251-Arg186 (bottom; from Asp251_{OD1} to Arg186_{NH1}). The distance was calculated every 10 picoseconds for the duration of the simulation and plotted in black while a running average is shown in red. The dashed turquoise line at 3.5Å represents the definition of a contact used in our analysis.

each crystal structure remained broken throughout the simulation. These observations prompted us to re-examine the electron-density map obtained from the refined covalent complex. There are 4 molecules in the asymmetric unit and in two of these the side chain of Arg186 is not visible while in a third asymmetric unit weak density indicates that Arg186 extends toward the surface away from Asp251. However, in the fourth asymmetric unit the 2Fo-Fc electron density map showed that Arg186 possibly adopts two orientations, one toward the surface and one near Asp251. We therefore carried out a few rounds of refinement with Arg186 placed in multiple orientations and the resulting map clearly shows that Arg186 interacts with Asp251 very similar to the MD structure (Figure 2.7). It thus appears that even in the open conformation, the Asp251-Arg186 interaction is present but weakened relatively to the closed conformation.

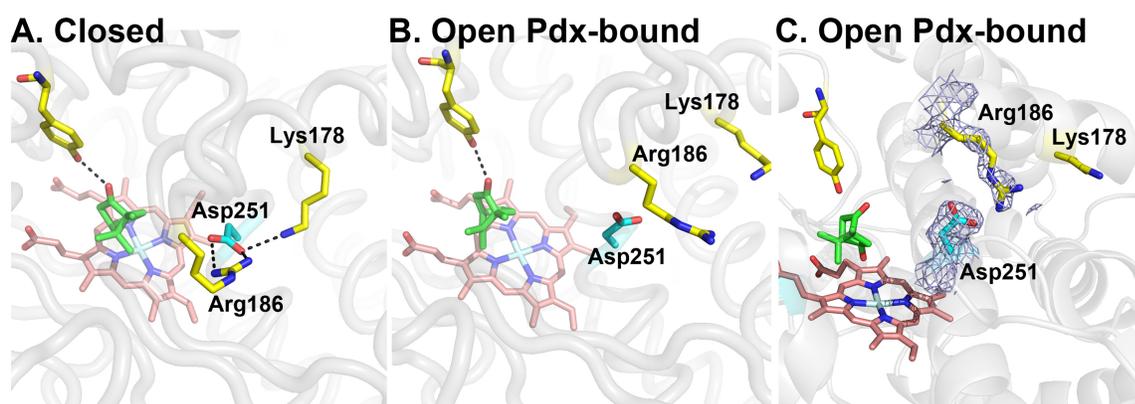


Figure 2.7. The Arg186-Asp251 ion pair is weakened but present in the open Pdx-bound complex. Panels A and B shows the environment of Asp251 in the closed and open structures as initially observed by Tripathi et al.(48) A reexamination of the electron density map, however, reveals that in one of the four P450cam molecules in the asymmetric unit Arg186 interacts with Asp251 (Panel C). The map shown is a 2Fo-Fc map contoured at 0.8σ while the density breaks up at 1.0σ . This indicates a relatively high mobility of Arg186 further suggesting that the Asp251-Arg186 interaction is not as strong as in the closed conformation.

Discussion

The recent structures of the P450cam-Pdx complex represent one last missing pieces in the P450 electron transfer puzzle, although several questions remain unanswered. The crystal

structures represent a single snapshot in time of the complex and the two crystal structures do show some differences in relation to the complex interface. In addition, dynamics clearly are important since the Pdx induced shift of P450cam from the closed to open conformation is closely tied to arming the proton coupled electron transfer machinery required for O₂ activation. Since structural changes are so intimately tied to the activation process, molecular dynamics simulations can provide important insights into functional important structural changes.

The complex is very stable throughout over the 100ns simulation. The most important change is that the C helix of P450cam, which has been previously identified as important to partner recognition, undergoes a conformational change that moves it closer to Pdx than what was observed in the complex crystal structures (47, 48, 168). This emphasizes the importance of the C helix as this movement brings P450cam into closer proximity to Pdx and strengthens the intermolecular interactions that had previously been observed as the largest singular conformational change observed in the simulation. This movement of the C helix tightens the interactions with Pdx_{Trp106}, a residue known to be essential for activity (54, 168). Additionally, Pdx undergoes a small rotation in relation to P450cam that allows for the formation of several new interactions between the redox partners and an enlargement of the interface. Despite several minor contacts that had been observed in the crystal structure being broken by these changes, the formation of a host of new contacts keeps the overall complex very stable. When the top contacts from the MD simulation are used to describe the P450cam-Pdx interface, a striking observation emerges. Contacts that were observed in both crystal structures represent only a small number of contacts localized between only the intermolecular salt bridge and P450cam helix C while the new contacts make up the remainder of the interface (Figure 2.8). In the MD simulations, these large surfaces of Pdx and P450cam relax to each other increasing the interface surface

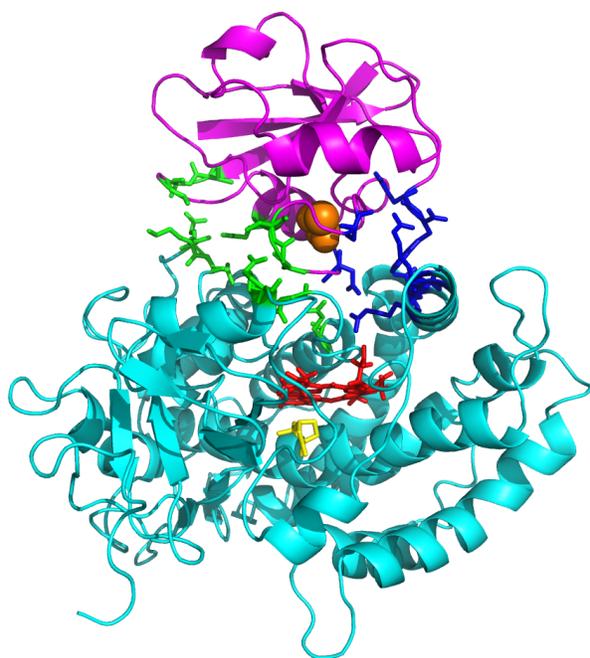


Figure 2.8. Major contacts observed in the P450cam-Pdx complex. Snapshot from the equilibrated P450cam (cyan) – Pdx (magenta) complex. The iron-sulfur cluster, heme and camphor are shown in orange, red and yellow for reference. While several of the major contacts observed in both crystal structures are conserved throughout the MD simulation (shown as blue sticks), the newly formed contacts in the MD simulation all occupy a different part of the P450cam-Pdx interface where minimal contacts were observed structurally. The residue contact pairs are taken from Table 1. Recent experimental work has shown that contacts in this new region however may still play a role in partner recognition.

significantly through small conformational changes and a small rotation of Pdx in relation to P450cam. These movements significantly increase the intermolecular contacts in the complex.

An important question to address is why these contacts were not observed in either crystal structure. The complex in the cross linked structure will have its conformational freedom restricted by the covalent linkage and, in fact, several of the new contacts would be directly inhibited by the presence of these molecules. The initial cross-linked structure presented by Tripathi et al. would restrict conformational freedom near the P450cam_{Glu76}-Pdx_{Ser42} and P450cam_{Gly353}-Pdx_{Gly40/41} hydrogen bonds while the second cross linked structure would directly inhibit the formation of the hydrogen bond between P450cam_{Val345} and Pdx_{Gln25} due to the presence of the cross link site at the neighboring P450cam_{Lys344} residue. Although the non-covalent crystal structure does not have these restrictions, in both crystal forms, Pdx is tightly packed with not just its P450cam redox partner but other P450 molecules in the crystal lattice. While what is observed in the crystal structure very likely represents the lowest energy complex,

the energy well in such electron transfer complexes is shallow in order to maintain rapid off-rates required for high turnover. MD thus provides a clearer picture on states that are nearly isoenergetic with the lowest free energy state.

An important hypothesis to have emerged from the P450cam-Pdx structure is the role of the essential Asp251 (Figure 2.7). The Asp251Asn mutant dramatically lowers activity (169) and what little activity is left exhibits a large increase in the kinetic solvent isotope effect (170). This suggests that Asp251 plays an important role in the delivery of protons to the iron-linked O₂ molecule that is essential for cleaving the O-O bond. However, in the closed P450cam structure Asp251 is ion paired with Arg186 and Lys178 which is totally inconsistent with its proposed role as a proton shuttling residue. When Pdx binds and the active site opens up, the Asp251 ion pairs are broken thus freeing Asp251 to serve its proposed role in proton transfer. The MD simulation, however, shows that the Asp251-Arg186 ion pair reforms for much of the simulation unlike the Asp251-Lys178 ion pair that is never reformed which is supported by our re-examination of the electron density maps. This is particularly relevant in light of recent studies on a close homolog of P450cam, CYP101D1 (171). CYP101D1 catalyzes exactly the same reaction as P450cam, at the same rate, and uses a similar Pdx-like ferredoxin called Arx. In CYP101D1 its Asp251 homolog is also essential for activity but this Asp is ion paired with an Arg while the residue analogous to Lys178 in P450cam is Gly180 in CYP101D1. CYP101D1 is quite promiscuous with respect to redox partner since Pdx can support CYP101D1 camphor hydroxylation while the CYP101D1 redox partners exhibit no activity with P450cam (171). This clearly means that CYP101D1 does not require the assistance of specific redox partner binding to activate the proton coupled electron transfer process very likely because its catalytically critical Asp259 has only one ion pair with an Arg residue and thus can more readily serve in a proton relay network

without a large conformational adjustment. Our MD structure of the P450cam-Pdx structure with the Asp-Arg ion pair reforming thus closely resembles the ion pairing in CYP101D1. This further suggests that the key to shutting down the active role of Asp251 is the ion pair with Lys178 in the closed structure and might point toward an important role for Arg186. These observations can be readily tested by suitably designed mutagenesis experiments.

Materials and Methods

The ferric P450cam-Pdx complex as solved by Tripathi et al. was used as the starting point of the simulation (48). The covalent cross-linker molecules were removed and the cross-links sites mutated back to their wild type counterparts (Pdx_{Asp19} and P450cam_{Lys344}). The system was then placed in an octahedral unit cell with an angle of 109.47 degrees and solvated with 11,811 water molecules and neutralized through the addition of sodium ions. The full system consisted of a total 43,495 atoms. The Amber 12.0 suite (<http://ambermd.org/>) was used for all calculations (172). The ff10 forcefield provided with the Amber 12.0 package was used for the protein, while the heme-Cys ligand parameters were taken from Shahrokh et al (173). Parameters for the camphor were derived with antechamber and the gaff forcefield (174) using the BCC charging scheme (175, 176). Partial charges and optimal geometry for the Pdx iron-sulfur cluster were obtained using density functional theory and the 6-31G* basis set as implemented in Jaguar (Shrodinger) while average bond distances and angles were taken from the refined Pdx structure. The P450cam-Pdx crystal structure including crystallographic identified waters was solvated with box of TIP3 waters using 10Å cushion. Counterions were added to maintain a net neutral charge.

The structure was prepared for production MD runs by first energy minimization for 1000 cycles with all heavy atoms except water molecules fixed in position followed by another 1000 cycles where all atoms were allowed to move. Production runs were carried out with a 1 fs time step and coordinates saved every 10 ps. Temperature and pressure were held constant through weak coupling with a 1 ps pressure relaxation time and Langevin dynamics using a collision frequency of 1 ps^{-1} . Periodic boundary conditions were used with a Particle Mesh Ewald implementation of the Ewald sum for the description of long-range electrostatic interactions (177). A spherical cutoff of 10.0 \AA was used for nonbonded interactions. Bonds involving hydrogen atoms were constrained using SETTLE (178). Analysis was carried out by using the Amber software suite, VMD and locally developed analysis tools (179, 180).

Chapter 3

On the affect of redox partner binding and conformational state of P450cam through the application of molecular dynamics

Summary

Cytochrome P450cam carries out the hydroxylation of its substrate camphor following the transfer of two electrons from Putidaredoxin or Pdx. Recently, two opposing viewpoints have arisen, both supported by experimental data, of the conformational state of P450cam upon Pdx binding for the second and final electron transfer reaction. While recent crystal structures and EPR data has favored the open state hypothesis, NMR work supports the closed state. In order to probe the stability of the two proposed complexes, we carry out extensive molecular dynamics simulations of P450cam in the open, closed and intermediate oxy states both individually and bound to Pdx. We find that the open state complex structure is more stable and results in a more favorable binding free energy than either the closed or oxy Pdx bound complexes though all three are found to be favorable. The findings herein are then compared to previous experimental data to support the recently proposed two step Pdx binding hypothesis upon which Pdx binds to the closed state and triggers a conformational change to the energetically more favorable open state to allow for catalytic chemistry to occur at the active site which fits all previous experimental and computational findings.

Introduction

Cytochrome P450s (P450s or CYPs) are a super family of heme containing protein monooxygenases that catalyze the oxidation of a wide variety of organic compounds that play

vital roles in the metabolism of sex hormones, vitamin D, xenobiotics and other small molecules across all kingdoms of life (1, 46, 150, 152, 181). Due to wide variety of functions of differing substrate specificity in the 20,000 different known CYP gene identified to date, including 57 in humans, there has been large degree of study into the structure and function of CYP enzymes (1, 151, 152). A significant portion of this research has been studying the interactions between P450 enzymes and the protein redox partners that donate the necessary electrons to power the P450 oxidation reaction (34, 45, 46, 50, 51, 54, 156, 166, 167, 182, 183). To date, only three such P450 redox complexes have been solved experimentally (45-48). While these structures have provided the first glimpses of how these complexes form, the dynamic nature of these complexes still remain in question.

P40cam (or CYP101A1), from the bacteria *Pseudomonas putida*, is the most structurally characterized member of the CYP super family and has served as a paradigm for P450 structure-function studies for many years (157). P450cam catalyzes regio- and stereospecific hydroxylation of camphor to 5-hydroxycamphor through the addition of two protons, molecular oxygen and two electrons (156, 184). The electrons required for the hydroxylation of camphor are shuttled to the active site heme iron from NADH to the FAD cofactor of putidaredoxin reductase (Pdr) which are then donated to the iron-sulfur containing putidaredoxin (Pdx) which then donates two electrons, one at a time, to P450cam. In addition to donating electrons, Pdx has been established to play an effector role upon binding to P450cam and donating the second electron inducing conformational changes in the P450cam active site (53, 54, 163, 185). While both Pdr and Pdx are relatively rigid, P450cam itself is very dynamic (41, 186).

P450cam has been observed to have two extreme conformational states that have both been experimentally observed, the open and closed states. In the open state, the F and G helices

on the proximal side of the heme move to open the substrate access channel that exposes the active site to solvent and allows camphor to enter the active site and bind to P450cam (1). In this state, P450cam_{Asp251}, which participates in the proton relay network by delivering the required protons to the active site, is free in the active site from any intramolecular ion pairs. Crystal structures of P450cam in the open state have previously only been observed in the absence of substrate or with the aid of tethered substrates that force the substrate access channel to remain open (41, 42). In contrast, when P450cam is in the closed state the F and G helices move towards the core of the enzyme and close the channel, trapping substrate in the active site and preventing solvent from entering the active site. In the closed state, P450cam_{Lys178} and P450cam_{Arg186} ion pair to the essential P450cam_{Asp251} preventing it from participating in the proton relay network. This closed conformation has been observed in crystal structures of P450cam in the presence of substrate as well as several trapped catalytic intermediate states of the enzyme (43, 44). Recent EPR work has also shown that the open to closed transition occurs upon substrate binding (187). Based on the differing environments for these crystal structures, it was believed that P450cam existed in the open state whereupon substrate binding shifted the enzyme into the closed state throughout the catalytic hydroxylation before opening to release the substrate and begin the cycle anew. However, upon the solving of the P450cam-Pdx crystal structure, this approach has had to be reevaluated in light of what the newly solved structure revealed.

In 2013, a cross-linked crystal structure of the P450cam-Pdx complex was first solved wherein P450cam was unexpectedly observed to be in the open state (48). This study was shortly followed by a second, non-covalently linked and independently obtained complex structure that also observed P450cam in the open state when bound to Pdx (47). While this result was surprising, upon reflection it is consistent with many previous experimental studies (50-52) as

well as new EPR studies that showed that P450cam existed in the open state when bound to oxidized Pdx (162), which suggested that Pdx may bind initially to the closed state, transfer the second electron, then play its effector role in shifting P450cam to the open state as observed in the crystal structure. Recently however, an NMR study has shown that P450cam exists in a closed state in solution when in complex with Pdx at the long time scales that NMR can probe (188). This study troubling argues that the crystal structures, including their own, are likely crystallographic artifacts and that much of the previous data collected on the complex is not applicable due to the different conditions that each study carried with them which disregards all data that does not fit their result. However, in this proposed closed state complex, P450cam_{Asp251} cannot act in the proton relay network that should prevent P450cam from functioning. While the hypothesis produced by Myers and coworkers of a two step process of P450cam-Pdx binding can still fit all of the data, there are important questions regarding the dynamics of P450cam in both differing conformational states and in the presence and absence of Pdx.

What we present here is a systematic computational study of the dynamics of P450cam in different conformational states through the application of a series of 100 nanosecond molecular dynamics trajectories. In addition to the open and closed states, we also probe the dynamics of P450cam in the catalytically important intermediate oxy state where molecular oxygen has bound to the heme that has been shown to exist between the open and closed conformational states. Additionally, to probe any Pdx-dependent dynamic changes and compare complex stabilities, simulations were carried out on the open-complex crystal structure, an NMR closed-complex model that has Pdx in complex with P450cam in the closed state, and a homology modeled oxy-P450cam-Pdx complex. We find that while the P450cam-Pdx complex is energetically favorable and a subset of previously identified interfacial contacts are preserved in

all three studied complexes, the open complex is found to be the most stable over both 100 nanosecond classical molecular dynamics simulations as well as 490 nanosecond accelerated molecular dynamic simulations as well as more energetically favorable through MMPBSA calculations than any other complex. In light of both these new results and previous experimental results, we agree with the proposed two-step Pdx-binding hypothesis of Myers et al. (162) in which Pdx initially binds to the closed state and triggers a conformational change to the open state of P450cam and discuss how these new findings fit that hypothesis.

Results and Discussion

Dynamic differences between P450cam conformational state and Pdx binding.

In order to study the dynamic nature and stability of P450cam in both different conformational states as well as the presence and absence of Pdx, seven different systems were constructed and simulated over 100 ns molecular dynamics trajectories to systematically study possible dynamic tendencies. Three separate P450 conformations, the open or P450cam(O), closed or P450cam(C) and oxy P450cam(Oxy) states, were used to probe dynamic difference based on conformation in P450cam as well as a simulation consisting of Pdx alone. The three complex structures used as starting points for study were the crystal structure solved by Tripathi and coworkers (48) and previously studied using MD (183) which is termed the P450cam(O)-Pdx system, the NMR-derived structure from Hiruma and coworkers (47) which solved the orientation of Pdx binding to a closed state model P450cam which is termed the P450cam(C)-Pdx system, and a homology modeled oxy state P450cam bound to Pdx termed the P450cam(Oxy)-Pdx system (see Methods for more details on system construction).

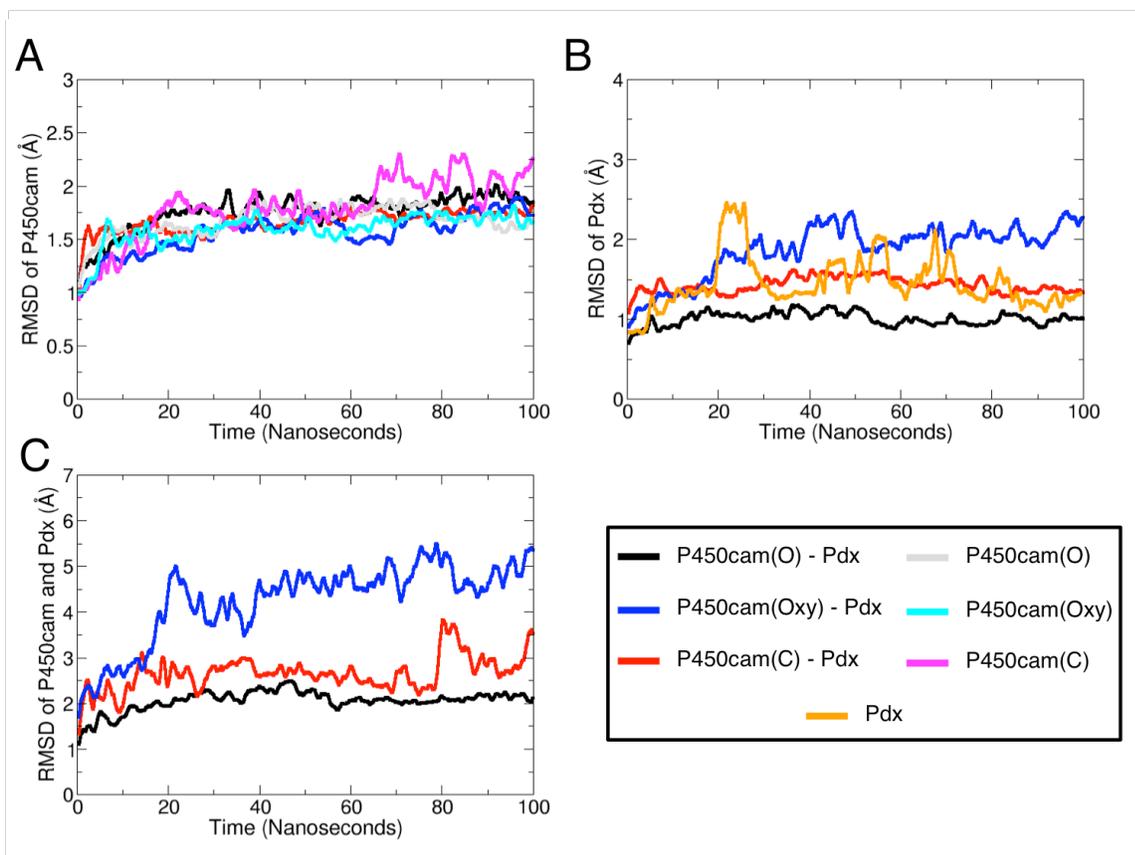


Figure 3.1. Root mean squared displacement analysis of P450cam and Pdx. A running average over a 100 picosecond window of the root mean squared displacement (RMSD) of the C_{α} atoms of A) P450cam only, B) Pdx only and C) P450cam-Pdx complex for the full 100 nanosecond trajectories for the seven P450cam-Pdx systems described in the figure legend over the respective 100 nanosecond trajectories.

Root mean squared displacement (RMSD) analysis of each of the seven systems reveals that the stability of P450cam in each system is very comparable (Figure 3.1A). Interestingly, the addition of Pdx does not seem to have a large affect on the stability of P450cam when compared to the corresponding systems where Pdx is absent. The corresponding Pdx analysis reveals that the equilibrated Pdx structure is similar for all systems with the exception of the P450cam(Oxy)-Pdx system which shows an increase in RMSD (Figure 3.1B). Similar to the Pdx RMSD, analysis of the RMSD of each of the three full complex structures reveals that the P450cam(Oxy)-Pdx system underwent a much larger conformational change than either the

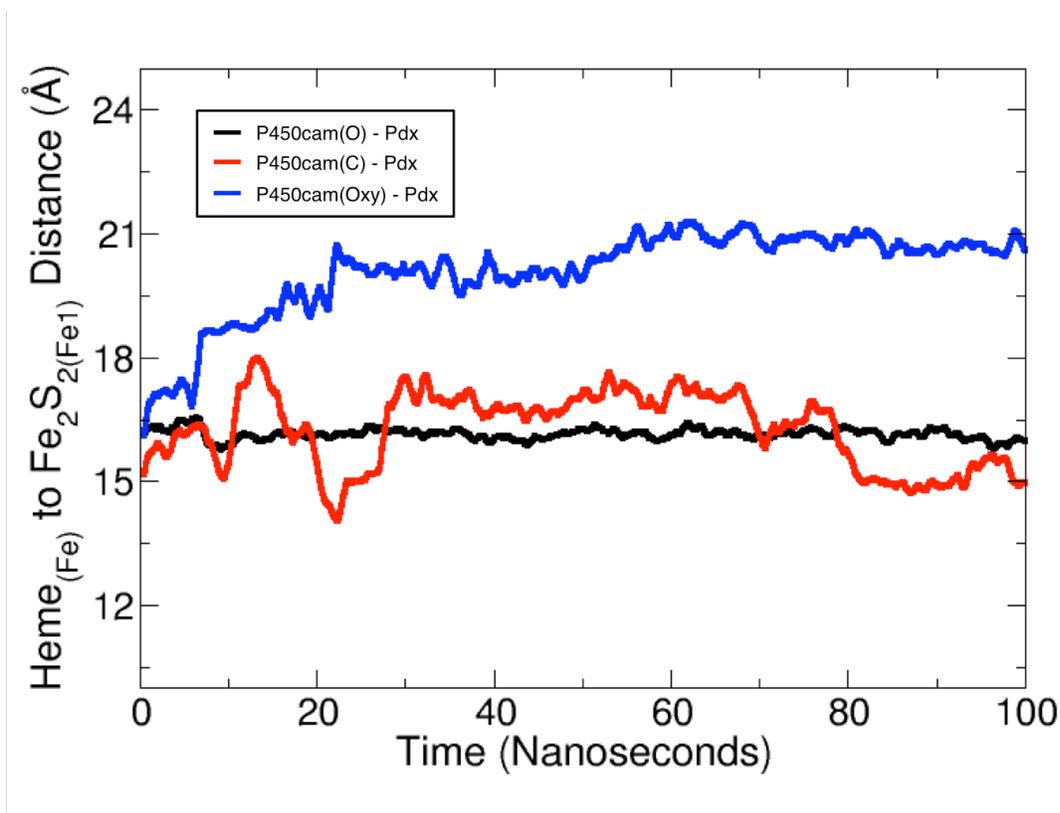


Figure 3.2. Distance between the P450cam heme iron and Pdx iron-sulfur cluster cofactors. The distances between the P450cam heme iron and the Fe1 atom of the Pdx Fe_2S_2 cluster over the course of the 100 nanosecond trajectories for the three P450cam-Pdx simulations as labeled in the upper left corner of the plot. The distances are plotted as a running average over a window of 100 picoseconds.

crystal structure of NMR model of the complex (Figure 3.1C). Given the low RMSD values for the individual components of these systems which would suggest that each protein alone is stable and has not undergone a conformation change, this large increase in the P450cam(Oxy)-Pdx system suggests that the orientation of the complex has changed which is further confirmed by an analysis of the distance between the heme and iron sulfur cofactors (Figure 3.2). While the P450(O)-Pdx system equilibrated to a inter-cofactor distance of $16.14 \pm 0.34\text{\AA}$ over the final 50 ns and the P450(C)-Pdx system equilibrated to $16.14 \pm 1.03\text{\AA}$, the P450cam(Oxy)-Pdx system saw a significant increase in the cofactor distance to $20.78 \pm 0.42\text{\AA}$ over the final 50 ns.

In order to probe for changes at the secondary structure or residue level, root mean squared fluctuation (RMSF) analysis was carried out for each of the six systems (Figure 3.3). While each of the three systems in the absence of Pdx have comparable RMSF values, the P450cam(Oxy)-Pdx system shows an increase in fluctuations not only in the C-terminus end of the sequence, but also in the F helix suggesting that the dynamics of P450cam(Oxy)-Pdx are significantly different than P450cam(Oxy) alone. Focusing on the C-helix, which has been strongly implicated in Pdx binding as well as the opening and closing motion of P450cam along with the B', F and G helices, P450cam in the closed state and in the presence and absence of Pdx shows a higher degree of flexibility when compared to the other states while the presence of Pdx

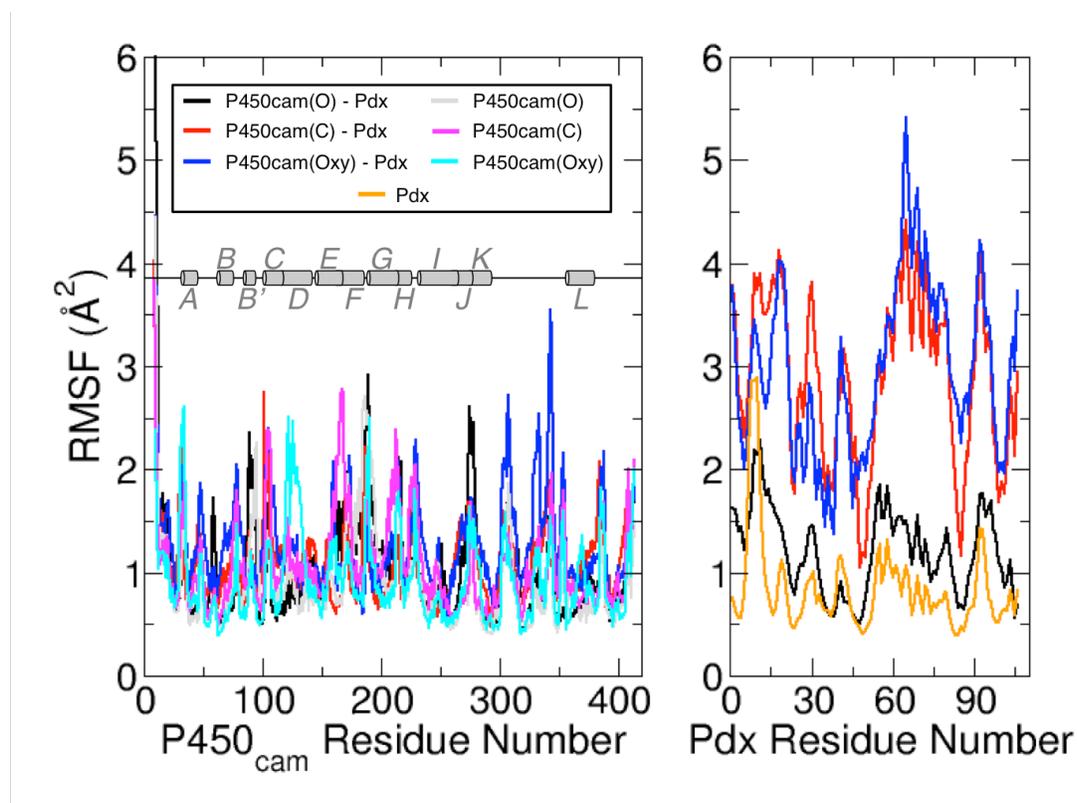


Figure 3.3. Root mean squared fluctuation analysis of P450cam and Pdx. Root mean squared fluctuation (RMSF) analysis for each of the seven systems described in the figure legends for P450cam (left) and Pdx (right).

in the oxy state decreases fluctuations. Unlike the C helix, the F helix shows very similar flexibility over the 100 ns timescale for all systems while N-terminus of the G helix showed high fluctuations for all systems. Additionally, the B' helix showed more flexibility in only the open state both in the presence and absence of Pdx. Interestingly, the RMSF of residue P450cam_{L358}, which in the L358P mutant can mimic the Pdx bound state and has been connected to Pdx-binding (185), is twice as high in the P450cam(Oxy)-Pdx system than the P450cam(Oxy) system. Finally, while Pdx bound to the open state of P450cam is similar to Pdx alone, there are significant increases in the fluctuations seen throughout Pdx when in the presence of P450cam in either the closed or oxy states. These would suggest that Pdx may be more stable when bound to the open state of P450cam, which will be investigated further in the following sections.

Combined, the results show that P450cam in the open state is more flexible in the dynamic regions implicated with the open and closing motions of the enzyme while these regions are lower in the closed state with the exception of the initial residues on the N-terminus of the G helix that remains relatively dynamic in every state while Pdx binding will indeed affect the flexibility of the C helix of P450cam.

Intramolecular P450cam_{Asp251} ion pairs

One of the major questions concerning the activity of P450cam in solution concerns how P450cam_{Asp251}, which is prevented from donating protons by dual ion pairs to P450cam_{Lys178} and P450cam_{Arg186} in the closed state, can participate in the proton relay network for the hydroxylation of camphor. One of the major findings in our previous study concerning the dynamics of P450cam(O)-Pdx complex was that P450cam_{Arg186} both dynamically forms and breaks its ion pair with P450cam_{Asp251} in the open state observed in the crystal structure(183).

For this reason, we tracked the formation of these ion pairs in each of the six systems that included P450cam (Figure 3.4). The first of these ion pairs, P450cam_{Lys178}-P450cam_{Asp251}, has previously been hypothesized to not play a major role in activity through mutagenesis and comparison to similar systems while previous study of the P450cam(O)-Pdx complex showed that the ion pair remained broken in the Pdx-bound open structure(183). In the present study, each of the simulations where P450cam revealed that the ion pair was never formed over the course of the respective 100ns trajectories while the ion pair is broken in every system where P450cam began in the closed, ion-pair formed state (Figure 3.4A and B). Given both the recent experimental evidence and the weakness of this ion-pair in P450cam regardless of conformational state of presence of Pdx binding, we can begin to conclude that this ion pair may not play a major role in either activity nor functions to hold the structure closed as has been the previous hypothesized.

The second of the two ion pairs, P450cam_{Asp251}-P450cam_{Arg186}, has proven to be more elusive than the first with regards to a deep understanding of structure function relationships. While this ion pair was ever only observed to form in the closed state until recently, our previous MD study of the P450cam(O)-Pdx system unexpectedly revealed that it dynamically formed and broke over a 100 nanosecond trajectory. These results lead to a reevaluation of the original crystal structure to reveal unseen density for the ion pair to be in both the open and closed state. Extending this analysis to the present systems, it is observed that this dynamic behavior is not only observed in the P450cam(O)-Pdx systems, but in the P450cam(O) system as well suggesting that this behavior is not Pdx-dependent, but rather may be a characteristic of P450cam in the open state. In the closed state in both the presence and absence of Pdx, this ion pair remained formed over the course of the full 100ns simulation while it also never broke in the

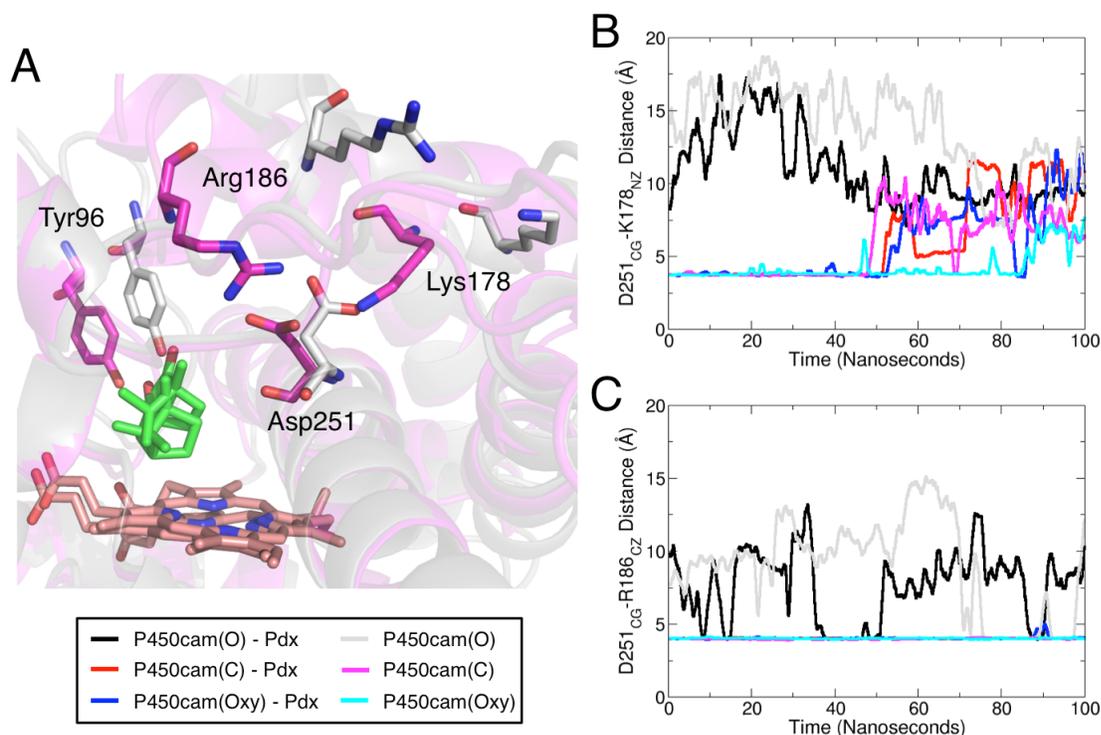


Figure 3.4. Analysis of the intermolecular ion pairs of P450cam Asp251. A) Structural overlay of the P450cam active site using the starting points for the P450cam(O) and P450cam(C) systems. In the closed state of P450cam (pink), P450cam_{Asp251} forms two intermolecular ion pairs with P450cam_{Lys178} and P450cam_{Arg186} while both of these ion pairs are broken in the open state (grey) due to movements in the F and G helices. B) The Asp251_{CG}-Lys178_{NZ} distance is used to probe the state of the Asp251-Lys178 ion pair in each of the six systems that contain P450cam. C) Similarly, the Asp251_{CG}-Arg186_{CZ} distance is used to probe the state of the P450cam_{Asp251}-P450cam_{Arg186} ion pair. For B and C, the displayed values are a running average with a window of 100 picoseconds over the full displayed 100 nanosecond trajectory.

P450cam(Oxy) system. Interestingly, there was a short breakage of approximately 5 nanoseconds of the ion pair in the P450cam(Oxy)-Pdx system which may suggest that Pdx may in fact begin to weaken the P450cam_{Asp251}-P450cam_{Arg186} ion pair slightly, but this would require further study to reach a strong conclusion. However, in light of these findings, the computational work herein suggests that this ion pair may never been completely broken in the open form. Given the new experimental findings that Arg186 mutants show reduced activity but still turn over and can be rescued in the presence of increased concentrations of Pdx, it might be suggested

that this ion pair may not play as large of a role as was previously suggested as the ion pair does not seem to break completely in the open form as previously suggested.

Comparison of the P450cam-Pdx complex structures.

An important question to probe with these simulations is the validity and stability of the different P450cam-Pdx complexes, in particular in the open and closed states that have been determined by crystallography and NMR techniques respectively (47, 48, 188). There has been debate in the recent months regarding the conformational state of P450cam in complex. While recent crystallography, EPR and spectroscopic studies have all supported the hypothesis that the complex exists in the open form (48, 50-52, 162), a recent NMR study showed that P450cam exists in the closed state when Pdx is bound (188). In light of the seeming contradiction, the question remains how Pdx binds to P450cam including whether the process occurs via two step process as originally put forward by Myers and coworkers (162) where Pdx binds initially to P450cam in the closed state whereupon following the donation of an electron, P450cam shifts to the open state. It is important to note that this two-step hypothesis can support all of the previously documented studies.

Structural overlays of the three equilibrated complexes studied herein reveal that the P450(Oxy)-Pdx system has undergone significant changes in the orientation of the complex when compared the to open and closed forms that are remarkable similar (Figure 3.5). During the 100 ns trajectories, the P450cam(O)-Pdx system underwent the same conformational change as observed previously(183) as Pdx rotates to form a larger interface with P450cam with the final equilibrated complex having a buried interface surface area of $1,607.0 \pm 73.6 \text{ \AA}^2$ (calculated as the difference between the sum of the surface areas of P450cam and Pdx and the surface area of

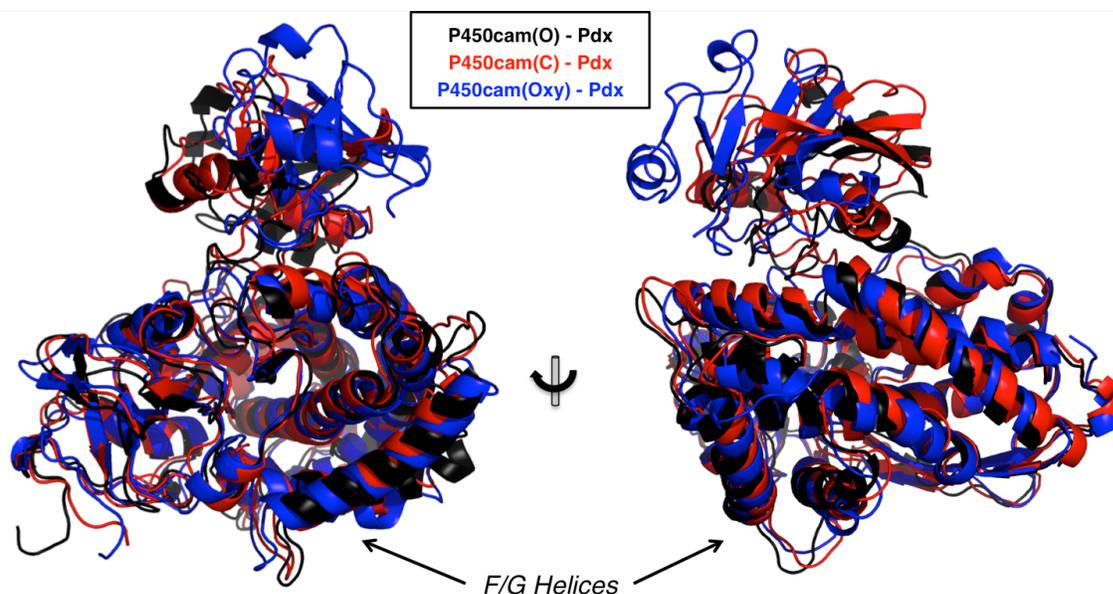


Figure 3.5. Overlays of the equilibrated P450cam-Pdx complexes following 100 nanosecond molecular dynamics simulations. The final equilibrated P450cam-Pdx complex structures from the open-crystal structure (black), closed-NMR model (red), and homology modeled oxy complex (blue) are overlaid and rotated to highlight the similarities between each of the three resulting equilibrated complex. The F/G helices that are involved in the open and closing motions are also highlighted for reference.

the entire complex over the final 50 ns) as compared to $946.6 \pm 72.0 \text{ \AA}^2$ for the closed-state complex and $848.1 \pm 63.2 \text{ \AA}^2$ for the oxy-state complex (Figure 3.6).

To probe any possible dynamical conformational dependence of P450cam on the presence or absence of Pdx, Principal Component Analysis (PCA) was carried out on P450cam over the course of each of the six trajectories that included P450cam. Previous work by Markwick and coworkers helped map out principal component space for P450cam using the first two principal components (186), including noting that the closed state of P450cam existed broadly around $(-28.0, -5.0)$ while the open state exists broadly around $(18.0, 0.0)$, with an energy minimum centered at $(18, -18)$. The projection of the present six systems show good agreement with the previous mapping of principal component space (Figure 3.7). Importantly, no system was ever observed to shift from the open state to the closed state suggesting that the

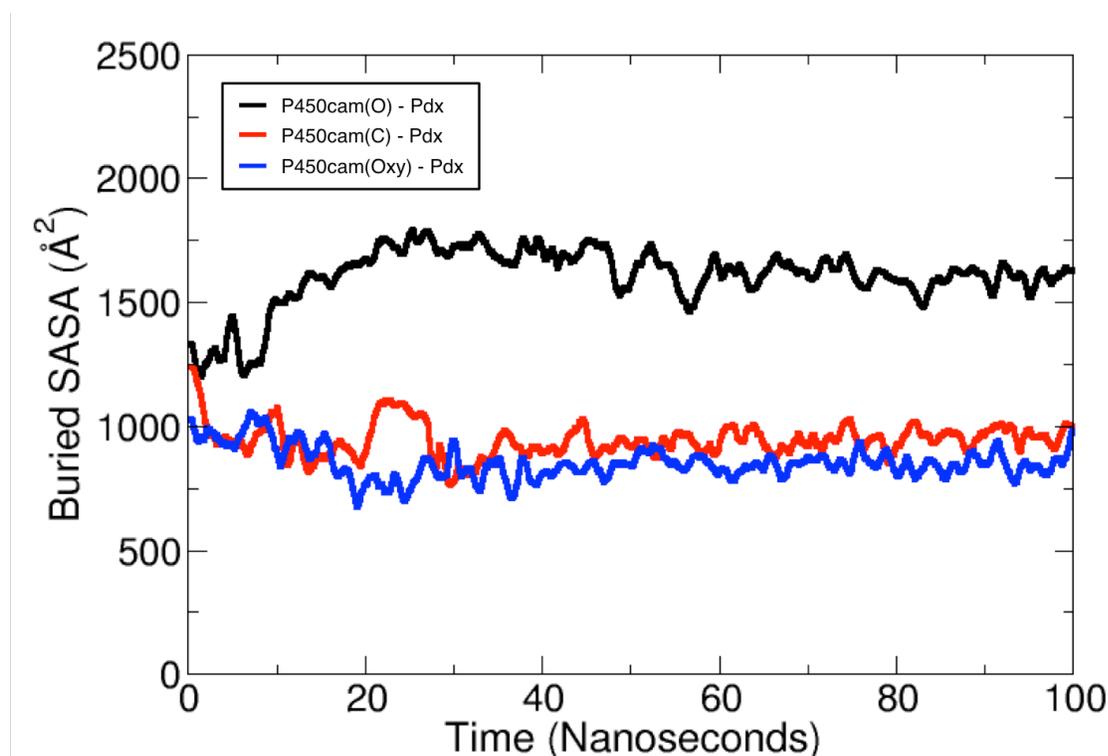


Figure 3.6. Calculated buried surface area in the P450cam-Pdx complexes. Buried surface area was calculated by calculating the sum of the surface areas of P450cam and Pdx separately then subtracting the surface area of the P450cam-Pdx complex for the P450cam(O)-Pdx (black line), P450cam(C)-Pdx (red), and P450cam(Oxy)-Pdx (blue) systems.

P450cam-Pdx complex was stable in regards to the open and closed conformations for each complex over each 100 ns trajectory.

Comparing the systems that differ only in the presence or absence of Pdx, it is possible to probe any possible dependence in the conformational sampling of P450cam when binding to Pdx. In the oxy-state, there is little change in the sampled regions of principal component space (In Figure 3.7, with Pdx in blue and without Pdx in cyan) or in the closed-state (with Pdx in red and without Pdx in magenta). Interestingly, there does seem to be a difference in the sampled conformations as in the absence of Pdx the open form of P450cam samples a large region centered around (20, 15) while in the presence

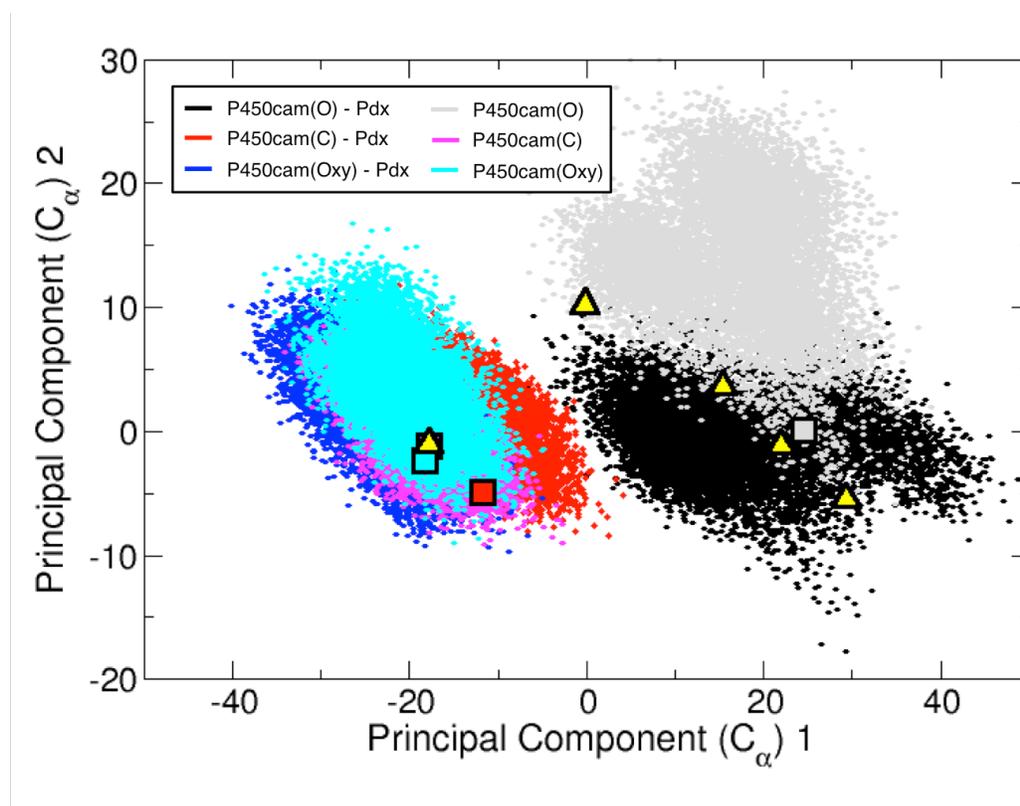


Figure 3.7. Conformational dynamics of P450cam in differing conformational and complex states. Principal component analysis (PCA) was used to project the conformational dynamics of P450cam onto principal component space every 10 picoseconds over the full 100 nanosecond trajectory for each of the six systems that included P450cam. The boxes overlaid onto the plot designate the simulation starting points following minimization; P450cam(O)-Pdx & P450cam(O) – (24.47, 0.06), P450cam(C)-Pdx & P450cam(C) – (-11.66, -4.95), P450cam(Oxy)-Pdx & P450cam(Oxy) – (-18.90, -2.39). In an effort to provide points of reference in principal component space, five crystal structures are also labeled as yellow triangles; 5CP4 (P450cam in a closed conformation) – (-17.84, -1.07), 3P6X (P450cam in the most open conformation observed) – (29.30, -5.51), and three additional intermediates solved (1RE9, 1RF9 & 3P6T) that used tethered substrates to trap intermediates between open and closed at (-0.16, 10.23), (15.38, 3.61) and (21.96, -1.14) respectively similar to the mapping done by Markwick and coworkers (186).

of Pdx the distribution is centered near (20, 0). Of particular interest is the locations of the intermediate-open conformations that have experimentally observed previously by Goodin and coworkers (yellow triangles in Figure 3.7). P450cam(O)-Pdx better samples the P450cam conformations of these intermediate-open structures than P450cam(O) alone, suggesting that the presence of Pdx may aid in P450cam sampling these intermediate states between the closed and open states as observed by crystallography.

Focusing on the interfaces for these three equilibrated complexes, we carried out an analysis of the contacts that made up the P450cam-Pdx interface to probe any difference that may depend on the conformational state of P450cam (Figure 3.8). Similar to previous published results, the equilibrated P450cam(O)-Pdx complex results in a larger surface area for the interface (Figure 3.6) that results in a large number of residues making stable contacts between P450cam and Pdx. This complex results in the largest number of atom-to-atom contacts with an average of 33.64 atom-to-atom contact pairs per frame over the final 50 ns of the trajectory (see Methods for more details on the calculation of atom-to-atom contact pairs). The P450cam(C)-Pdx complex on the other hand has fewer contacts that were found to be stable over the trajectory and which results in fewer residues at the interface that were found to play a role in specific interfacial interactions as seen in Figure 3.7, resulting in an average of 13.26 contact pairs over the final 50ns. Finally, the P450cam(Oxy)-Pdx complex is similar to that of the closed state complex, averaging 14.80 contact pairs over the final 50 ns. This finding is particularly interesting given that the oxy-state complex was previously shown to have undergone significant changes when compared to either of the experimentally observed structures which suggests that while the interface in P450cam(Oxy)-Pdx is smaller than that of P450cam(C)-Pdx, the residue interactions that stabilize the oxy-state complex are still as stable as the closed-state complex.

Given the three different binding model of P450cam-Pdx, it is possible to probe the differences in several of the residues that have been noted previously to play important roles in complex formation and electron transfer between each complex structure. Firstly, the P450cam_{Arg112}-Pdx_{Asp38} intermolecular ion pair that plays an important role in particularly the second electron transfer that results in the production of compound I from the oxy-state of P450cam (54), remains a stable interaction in each complex over the full trajectories.

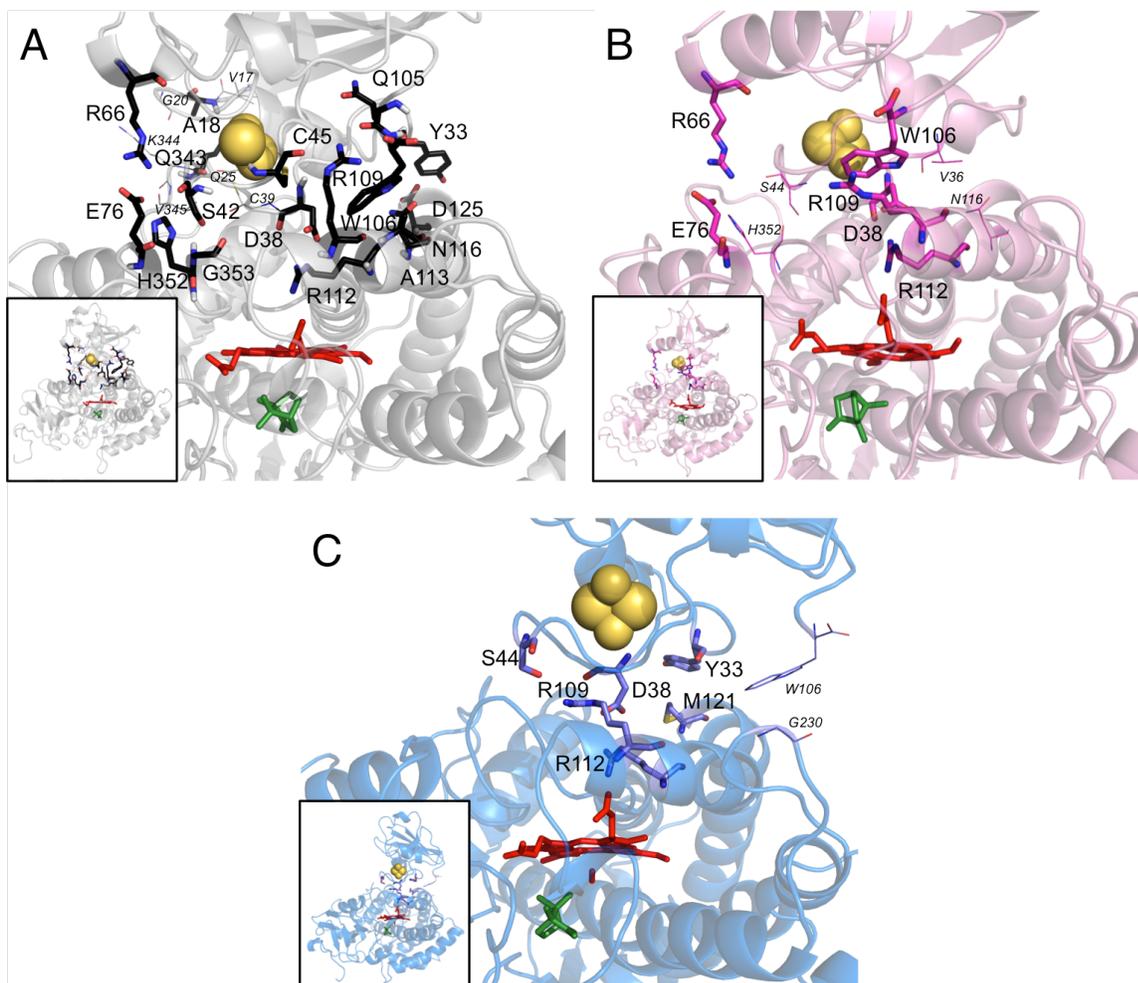


Figure 3.8. Comparisons between the interfaces of the three simulated P450cam-Pdx complexes. The interface of P450cam-Pdx complexes are shown for the A) P450cam(O)-Pdx, B) P450cam(C)-Pdx, and C) P450(Oxy)-Pdx systems. For each system, the iron sulfur cluster is shown as yellow spheres, the heme (and O₂ for the oxy complex) in red sticks and camphor in green sticks for reference. Interfacial residues that were found to average at least one atom-to-atom contact of 3.5Å or less per frame over the full 100 nanoseconds (or termed strong contacts) are shown as thick sticks and labeled with larger residues numbers while those residues that were found to average between 0.5-0.999 atom-to-atom pairs per frame (those residues that play smaller roles at the interface) are shown as thin sticks and are labeled using smaller numbers in italic. See Methods for more details on the calculation of contact pairs.

Interestingly, Pdx_{Trp106}, which has been noted to play an important role in redox partner binding and activity (48, 54, 168) seems to contact different residues in each of the three conformations. In P450cam(O)-Pdx, Pdx_{Trp106} forms stable interactions with P450cam_{Arg109}, P450cam_{Ala113}, and P450cam_{Asn116}. In the closed state complex however, only P450cam_{Arg109} forms a stable

interaction with Pdx_{Trp106}. In contrast to the weaker closed-state, Pdx_{Trp106} forms only transient contacts with a number of residues on P450cam in the oxy-state complex. The strongest of these interactions, that with P450cam_{Gly230}, averages less than one atom-to-atom contact per frame, lower than all of the previously described residues that averaged at least one such contact per frame.

Table 3.1. Calculated binding free energies of the three studied P450cam-Pdx complexes.

Contributions to binding free energy*	P450cam(O)-Pdx: Crystal Structure		P450cam(C)-Pdx: NMR		P450cam(Oxy)-Pdx: Homology Model	
	Mean (kcal/mol)	σ (kcal/mol)	Mean (kcal/mol)	σ (kcal/mol)	Mean (kcal/mol)	σ (kcal/mol)
$\Delta G_{np.MGB}$	-9.1	0.74	-4.4	2.2	-5.1	0.7
$\Delta G_{elec.MGB}$	-982.4	43.9	-1022.5	125.4	-1445.7	59.8
ΔG_{sol}	-991.5	43.7	-1027.0	127.2	-1450.8	60.0
ΔG_{elec}	30.2	10.5	32.7	15.2	37.2	10.8
ΔG_{total}	-60.2	12.6	-12.4	15.0	-9.7	8.8

$\Delta G_{np.MGB}$: non-polar solvation free energy, $\Delta G_{elec.MGB}$: electrostatic solvation free energy, ΔG_{sol} : non-polar and polar contributions to solvation, ΔG_{elec} : electrostatic solvation free energy and electrostatic energy, ΔG_{total} : estimated binding free energy

* - All values in kcal/mol

In an effort to further characterize the differences in Pdx binding to different conformations of P450cam, binding free energies were calculated for the P450cam-Pdx complex using MM-PBSA as included in the Amber suite of programs. The binding free energy was calculated for every 10 frames over the full 100 ns trajectory for each of the three P450cam-Pdx complexes and the results are displayed in Table 3.1. Matching the lowest RMSD, larger buried surface area and increased interfacial contacts, the binding of P450cam in the open state and Pdx was found to result in the lowest binding free energy of -60.2 ± 12.6 kcal/mol. Pdx binding to P450cam in either the closed or oxy-state both resulted in lower binding free energies of $-12.4 \pm$

15.0 and -9.7 ± 8.8 kcal/mol respectively. These results further suggest that while Pdx may form stable complexes with P450cam in the open, closed or oxy-states, the open-state is the more stable complex and should be more energetically favorable than either the closed or oxy conformations.

Exploring longer timescale dynamics with accelerated molecular dynamics

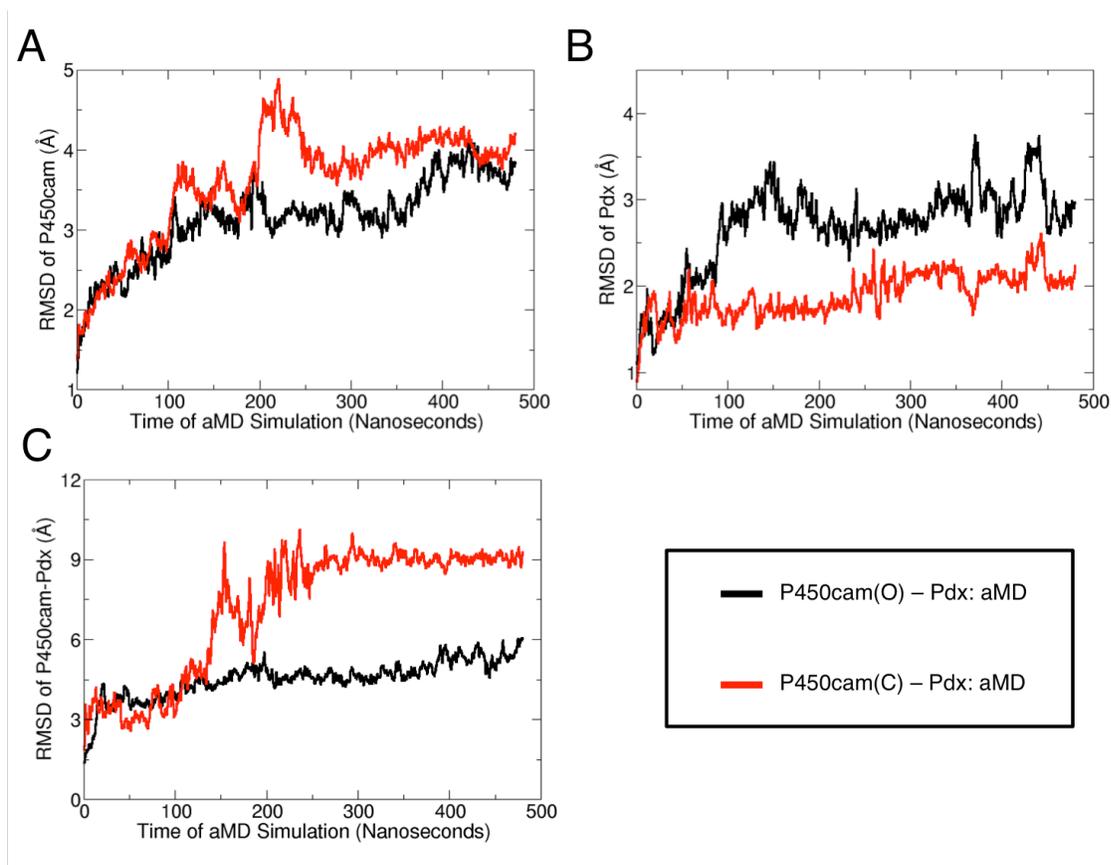


Figure 3.9. RMSD analysis of the accelerated molecular dynamics simulations of Pdx binding to the closed and open forms of P450cam. Similar to Figure 3.1, RMSD analysis was carried out on A) P450cam only, B) Pdx only, and C) the full P450cam-Pdx complex for the 490 ns aMD simulations for P450cam(O)-Pdx (black) and P450cam(C)-Pdx (red) complexes.

Finally, accelerated molecular dynamics (aMD) simulations were employed to probe the dynamics of the P450cam-Pdx complex at longer timescales that would otherwise be unavailable in standard molecular dynamics simulations. Previous work has not only demonstrated the utility

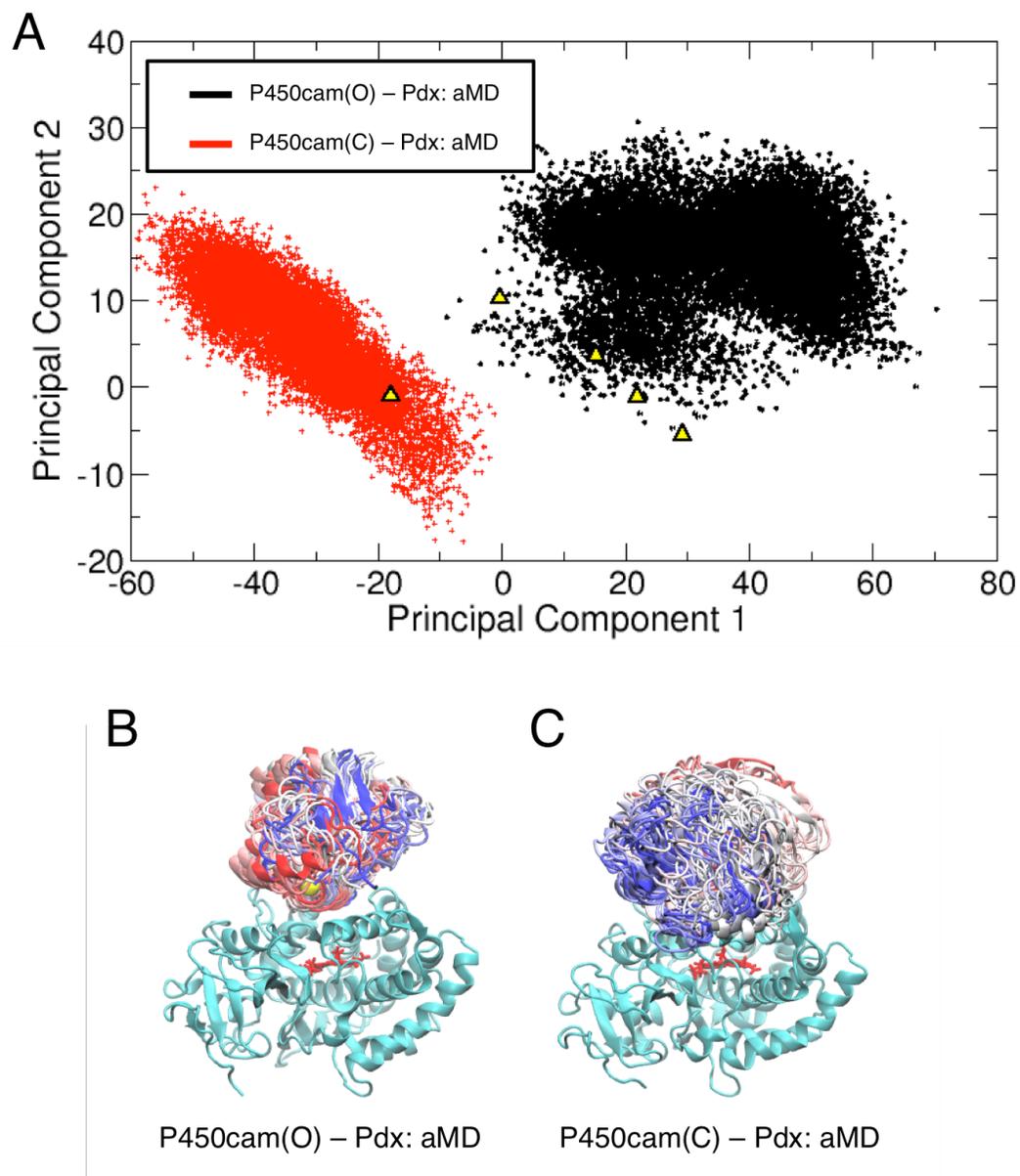


Figure 3.10. P450cam-Pdx complex evolution following 490 nanoseconds of aMD simulation. A) Principal component analysis as in Figure 3.7 of P450cam(O)-Pdx (black) and P450cam(O)-Pdx (red) over the full 490ns aMD simulation revealing P450cam in neither complex underwent a conformational change as mapped out in Figure 3.7. B&C) Snapshots of Pdx every 10 nanoseconds over the aMD simulation in relation to P450cam for the P450cam(O)-Pdx complex (B) and P450cam(C)-Pdx (C). The coloring of Pdx reflects the time in the simulation where blue is at zero nanoseconds and red the end of the simulation.

of accelerated simulations (189-192), but this method has previously been applied to substrate-free P450cam which allows for a easy starting point in analyzing P450cam in the presence of its redox partner (186).

490 nanosecond aMD trajectories were collected for both the P450cam(O)-Pdx and P450cam(C)-Pdx complexes in order to determine complex stability at long time scales. Previous studies employing aMD simulations have compared the sampling of a 500ns aMD trajectory, similar to those included here, to sampling of millisecond timescale events (193). RMSD analyses of the resulting trajectories reveal that while P450cam and Pdx individually are fairly stable, P450cam(C)-Pdx undergoes a significant conformation change in complex orientation (Figure 3.9). While P450cam in neither complex underwent a conformational change as verified by principal component analysis (Figure 3.10A), visual inspection of the resulting complex structures reveals the dynamics of Pdx binding to P450cam is very different (Figure 3.10B & C). While the orientation of Pdx in relation to P450cam does not significantly change in the open-complex, the majority of interactions at the interface between Pdx and P450cam in the closed state are lost save for the P450cam_{Arg112}-Pdx_{Asp38} intermolecular ion pair that allows Pdx to rotate around the intermolecular ion pair, resulting in a more dynamic complex (Figures 3.9 and 3.10).

Conclusions

Recently, there has been some lively discussion in the literature regarding the conformation of P450cam in the P450cam-Pdx complex (47, 48, 162, 183, 188). While crystal structures of the complex have revealed that P450cam is in the open conformation when bound to Pdx that has been validated by EPR studies, recent NMR work has begun to present an alternative viewpoint wherein P450cam is instead in the closed conformational state. Adding additional confusion to this problem, it is known that prior to P450cam binding exists in a closed, substrate-bound state, and the redox potential of P450cam is only capable of electron transfer from Pdx in this closed state. However, in the closed state, Asp251 is locked in an intramolecular ion pair and would be unable to act in the proton relay network, which would prevent

hydroxylation and release of the product from occurring. It is with these unanswered questions that we set out to try and understand the dynamics of the P450cam-Pdx complex in more detail by carrying out molecular dynamic studies of seven different P450cam and Pdx systems.

Interestingly, analysis of the resulting complex structure dynamics reveal that the P450cam(O)-Pdx complex is more stable in terms of RMSD, buried surface area, number of interfacial contacts, stability of Pdx-orientation upon binding as well as free energy of binding when compared to either the theoretical homology model of the P450cam(Oxy)-Pdx complex, or the NMR-based P450(C)-Pdx complex, though all three complexes were found to be thermodynamically favorable in terms of Pdx-binding. These results would suggest that Pdx would bind tighter to P450cam in the open state and be less dynamic than Pdx binding to the closed state of P450cam. However, there has been no experimental evidence that P450cam exists in anything other than a closed conformational state when substrate is bound. This contradiction would seem to suggest that the results here do not match up perfectly with what is known about the system. However, one previously proposed mechanism of the dynamics of this complex proposed by Myers et al. (162) not only fits the results found here but fits almost all data, both computational and experimental, on the P450cam-Pdx complex. Myers et al found that reduced Pdx favored binding to the closed state of P450cam while oxidized Pdx would bind to P450cam in the open state instead. Given their results, they proposed that Pdx would first bind to the closed state of P450cam upon which the electron would be transferred between the iron-sulfur cluster of Pdx to the heme of P450cam, followed by a redox-state-dependent conformational change of P450cam from closed to open. In this hypothesis, the effector role of Pdx that has long been sought after would be triggering the conformational change of P450cam to the open state, which would neatly explain how Asp251 would be freed to act in the proton-

relay network to allow for oxidation of the substrate while also allowing for Pdx to bind to the closed-state of P450cam initially.

The results presented herein can provide strong support to the two-step Pdx-binding hypothesis originally proposed by Myers et al. (162). Key to the two-step hypothesis is that the final conformational state of P450cam-Pdx would be Pdx bound to the open state of P450cam. While no conformational change from closed to open was ever captured herein, the data clearly shows that the open centric complex is more energetically favorable than the closed-state complex which would suggest that the complex will tend to favor the open state following Pdx binding. Importantly, Pdx binding to P450cam in the closed state is favorable and results in a relatively stable complex, though less so than the experimentally observed open-state complex. Following this reasoning, the results herein provide further evidence that the effector role of Pdx would be to favor binding to P450cam in the open-state, which frees up Asp251 to play its vital role in catalysis. The only missing piece of this puzzle is at what timescale the Pdx-dependent conformational change occurs as no such shift was captured in our study but would be required to fit all previously published results concerning Pdx binding to P450cam.

In order to piece together the final missing pieces of this puzzle, further studies that probe the Pdx-binding to P450cam in both the open and closed, substrate-bound states. Complicating the search for this answer is the relative instability of P450cam in the substrate-free and open state. As P450cam is not stable in the absence of camphor, studies that probe this final detail of Pdx-binding will require not only freshly purified P450cam, but would need to occur at a short of a timescale as possible in order to properly probe P450cam in the open state prior to degradation on the enzyme in addition to be accurate enough to pick up even small difference in binding

energies. However, such methods may prove to be the only way that this issue will ever be solved.

Methods

The respective crystal structures of P450cam in the open, closed and oxy states were used as the starting points for the P450cam(O), P450cam(C), and P450cam(Oxy) simulations. Each of these systems were then placed in an octahedral unit cell of water with an angle of 109.47 degrees and solvated with 11,684, 12,219, and 12,952 water molecules and neutralized through the addition of sodium ions. The resulting systems consisted of 41,526, 43,152, and 45,343 atoms. For the complex simulations, the P450cam(O)-Pdx system used the complex structure solved by Tripathi et al.(48) with cross-linkers removed and the system mutated to wild-type (48) identical to the starting point used in our previous MD study (183). The P450cam(C)-Pdx structures used the published NMR structure of Pdx in complex with P450cam in the ferric, closed state (PDBiD: 2M56)(47). Finally, the P450cam(Oxy)-Pdx complex was generated using homology modeling by overlaying the P450cam(Oxy) structure onto the known crystal complex structures to model the position of Pdx onto the new P450cam conformation. As the protonation state of P450cam_{Asp297} remains unclear, two separate simulations were carried out using the closed state of P450cam alone where one simulation included a protonated state and a second included the unprotonated side chain. As the protonated state of P450cam_{Asp297} resulted in a large increase of the time the nearby substrate camphor spent in the expected orientation with the hydroxylation target atom C5 in close proximity to the heme iron as opposed to camphor rotating freely around the hydrogen bond to P450cam_{Tyr96} (Figure 3.11), the protonated state was used for all following simulations herein. These complex systems were placed in orthogonal unit cells with a 10Å cushion around the complex, resulting in the addition of 18,981, 18,070, and 18,250

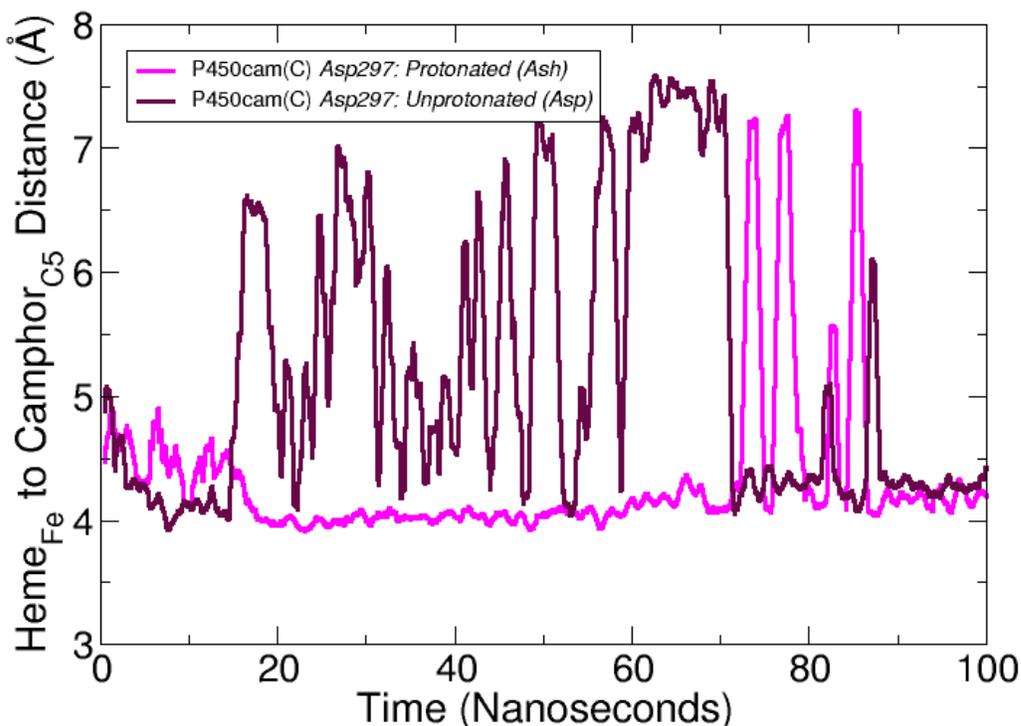


Figure 3.11. Comparison of camphor stability on the protonation state of P450cam Asp297.

In order to test whether P450cam residue Asp297 should be protonated in all simulations included herein, two simulations of P450cam in the closed state were carried out in protonated and unprotonated states and the distance between the hydroxylated target C5 atom of the substrate camphor and the heme iron was tracked. A running average using a window of 100 picoseconds was used to produce the figure above. As the protonated state of Asp297 resulted in increased camphor stability much closer to that expected when bound in the P450cam active site, all systems in the present study used the protonated Ash residue for residue 297.

water molecules for the P450(O)-Pdx, P450(C)-Pdx and P450(Oxy)-Pdx. Each system was then neutralized through the addition of sodium ions. The resulting open, closed, and oxy P450cam-Pdx complex systems consisted of a total of 65,005, 63,895, and 62,818 atoms respectively.

Simulations for all classical molecular dynamics simulations were carried out by the Amber 12 simulation package (180). The parameters employed herein are identical to those used in a previous study (183), but are discussed briefly here. The ff10 force fields were used to

model protein while the the Cys-heme ligand parameters were taken from Shahokh et al (173). The parameters for the Fe₂S₂ iron-sulfur cluster were taken from (194), and parameters for the substrate camphor were derived from antechamber with the gaff forcefield (174) employing the BCC charging scheme (176). The TIP3P water model was employed to model the water in each simulation and counter-ions were added to achieve a net neutral charge for each system. The parameters for the accelerated molecular dynamics simulations were carried out and calculated identically as has been described previously in the literature (186). Table 3.2 includes the additional calculated parameters for the accelerated molecular dynamics simulations listed for both the P450cam(O)-Pdx and P450cam(C)-Pdx simulations.

Each structure was prepared for MD simulation in the NPT ensemble by first carrying out 1000 cycles of energy minimization with all heavy atoms except water molecules held fixed in position, followed by an additional 1000 cycles where the fixed

Table 3.2. Additional parameters for P450-cam-Pdx complex accelerated molecular dynamics simulations.

Parameter Name	P450cam(O)-Pdx Value	P450cam(C)-Pdx Value
α^{dih}	360 kcal/mol	359 kcal/mol
α^{tot}	10,400 kcal/mol	10,223 kcal/mol
E_b^{dih}	7,340 kcal/mol	7,320 kcal/mol
E_b^{tot}	-189,614 kcal/mol	-185,748 kcal/mol

restraints were removed. Production runs for each system were completed with a one-femtosecond time step with coordinates saved every 10 picoseconds for analysis. The temperature and pressure were held constant in the simulation through weak coupling with a 1ps pressure relaxation time and Langevin dynamics using a collision frequency of one-picosecond⁻¹. Periodic boundary conditions were employed while electrostatic interactions were calculated

with a Particle Mesh Ewald implementation of the Ewald sum (177). A spherical cutoff of 10.0Å was used for the calculation of nonbonded interactions while bonds involving hydrogen atoms were constrained using SETTLE (178).

Analysis of the resulting trajectories was carried out using the Amber software suite (180), VMD (179), the Bio3D R package (195, 196) and locally developed analysis tools. Buried surface area over the course of each complex trajectory was calculated by subtracting the solvent accessible surface area (SASA) of the full complex as calculated by VMD using a probe radius of 1.4Å from the combined surface areas of P450cam and Pdx separately. Analysis of the interface contacts as seen in Figure 3.7 was carried out locally developed analysis tools that have previously been used to document major atom-to-atom contacts (197) but is briefly described here. For every frame of a given trajectory, atom-to-atom contact pairs are recorded for every atom-to-atom pair that is within 3.5Å. The resulting contacts are then summed up over the full trajectory and further summarized by residue to produce a list of the number of residue-to-residue contacts over the full trajectory. To aid in designating the staying power of the contact over the full trajectory, each sum is then divided by the number of frames where cutoffs of ≥ 1.0 atom-to-atom contact per frame are designated as major contacts while $1.0 > X \geq 0.5$ contacts are designated as a minor contact as displayed in Figure 3.7. Finally, MM-PBSA energy calculations for the complexes in Table 3.1 used the single trajectory method for calculating the energies for Pdx, P450cam and the complex itself from the complex trajectory (198, 199) as described both in the Amber manual and described in Ref (199).

Reproduced with permission from Fields J. B., Hollingsworth S. A., Chreifi G., Heyden M., Arce A.P., Magaña-García H.I., Poulos T.L. & Tobias D.J. (2015). “Bind and Crawl” Association mechanism of *Leishmania major* peroxidase and cytochrome c revealed by Brownian and molecular dynamics simulations. *Biochemistry*. **54**: 7272-7282. Copyright 2015 American Chemical Society.

Chapter 4

“Bind and Crawl” Association Mechanism of *Leishmania Major* Peroxidase and Cytochrome c Revealed by Brownian and Molecular Dynamics Simulations

*Published as co-first author with James B. Fields. Author Contributions: SAH, TLP, and DJT conceived the study. JBF, SAH, and MH set up, carried out and analyzed the Brownian dynamics simulations. SAH, GC, APA, and HIM carried out the experimental kinetics assays and GC solved the crystal structure of the D211N mutant of LmP. SAH carried out and analyzed the molecular dynamics simulation and the helix A sequence alignment. SAH, JBF, GC, TLP, and DJT prepared the manuscript.

Summary

Leishmania major, the parasitic causative agent of leishmaniasis, produces a heme peroxidase (LmP), which catalyzes the peroxidation of mitochondrial cytochrome c (LmCytc) for protection from reactive oxygen species produced by the host. The association of LmP and LmCytc, which is known from kinetics measurements to be very fast ($\sim 10^8 \text{ M}^{-1}\text{s}^{-1}$), does not involve major conformational changes, and has been suggested to be dominated by electrostatic interactions. We used Brownian dynamics simulations to investigate the mechanism of formation of the LmP-LmCytc complex. Our simulations confirm the importance of electrostatic interactions involving the negatively charged D211 residue at the LmP active site, and reveal a previously unrecognized role in complex formation for negatively charged residues in helix A of LmP. The crystal structure of the D211N mutant of LmP reported herein is essentially identical to that of wild type LmP, reinforcing the notion that it is the loss of charge at the active site, and not a change in structure, that reduces the association rate of the D211N variant of LmP. The Brownian dynamics simulations further show that complex formation occurs via a “bind and crawl” mechanism, in which LmCytc first docks to a location on helix A that is far from the

active site, forming an initial encounter complex, and then moves along helix A to the active site. An atomistic molecular dynamics simulation confirms the helix A binding site, and steady state activity assays and stopped flow kinetics measurements confirm the role of helix A charges in the association mechanism.

Introduction

Visceral and cutaneous leishmaniasis, two manifestations of a tropical disease transmitted by sand flies, is caused by certain parasitic protozoans from the genus *Leishmania*. *Leishmania major* can survive the onslaught of host macrophages and adverse oxidative environment by scavenging toxic and reactive oxidative species such as hydrogen peroxide (200). To carry out this scavenging process, *L. major* produces a mitochondrial heme peroxidase (*L. major* peroxidase, LmP) (201) and a mitochondrial cytochrome c (*L. major* cytochrome c, LmCyt_c), which act in unison as an inter-protein electron transfer complex (91).

Electron transfer (ET) complexes require a precise combination of both specificity and efficiency. On the one hand, such complexes must include a degree of specificity where the complex binding surfaces complement each other through hydrophobic and/or electrostatic interactions. On the other hand, the complexes must be dynamic, associating and disassociating quickly, for rapid ET and turnover to occur. If the redox partners form too tight a complex, rapid turnover cannot be achieved, so Nature must strike a balance between the requirements of specific binding and turnover. ET between two proteins can be broken up into three steps: (i) formation of the complex; (ii) electron transfer reaction; (iii) dissociation and separation of the complex. Here we focus on the first step for the LmP-LmCyt_c system and, particularly, on the question of whether or not it actually comprises two steps, namely, the formation of an initial

encounter complex, followed by local sampling of the binding interface that leads ultimately to the formation of a relatively tight ET complex.

Saccharomyces cerevisiae cytochrome c peroxidase (yeast CCP) has long served as a paradigm for studying protein-protein interactions and the mechanism of interprotein ET (202). The peroxidase that is the focus of the present study, LmP, also uses a cytochrome c as its substrate (91). The overall ET mechanisms for yeast CCP and LmP are similar (203-205). The enzyme is first oxidized by peroxide to give compound I, which consists of a ferryl iron and Trp radical, $\text{Fe}^{\text{IV}}=\text{O};\text{Trp}^{\cdot}$. The Trp radical is reduced by cytochrome c to give compound II, $\text{Fe}^{\text{IV}}=\text{O};\text{Trp}$. This is followed by an intramolecular equilibrium between the iron and Trp to give $\text{Fe}^{\text{III}}-\text{OH};\text{Trp}^{\cdot}$. Finally, a second molecule of cytochrome c reduces the compound II Trp $^{\cdot}$ radical to give the resting enzyme.

Although yeast CCP and LmP carry out the same redox chemistry (91), there is strong evidence that the mechanisms of their associations with their cytochrome c redox partners are different. First, LmP obeys Michaelis-Menten kinetics (91, 93), while yeast CCP does not (206). Second, while K_{M} steadily increases with increasing ionic strength, k_{cat} is effectively independent of ionic strength in LmP. This is quite different from yeast CCP, where k_{cat} steadily increases with ionic strength until a maximum near 150 mM salt (206). This has been attributed to a change in rate limiting step at low ionic strength from product dissociation to intramolecular ET at high ionic strength. For LmP the rate limiting step remains intramolecular ET at all ionic strengths (203). The larger increase in K_{M} with ionic strength in LmP is due to the greater dependence on intermolecular ion pairing, while yeast CCP depends more on nonpolar interactions. The latter observations suggest that electrostatic interactions drive the formation of the LmP ET complex, whereas nonpolar interactions drive the formation of the yeast CCP

complex. This hypothesis is supported by inspection of the crystal structures of the respective ET complexes. In the LmP complex, LmP D211 and LmCytc R24 form a salt bridge at the center of the binding interface and LmP E49 forms an ion pair with LmCytc K98 at the periphery of the interface (93). In contrast, there are close contacts between nonpolar side chains (e.g., yeast CCP A193 and Cytc V28) but no ion pairs in the yeast CCP-Cytc interface (75). Not only does the yeast CCP-Cytc complex lack the inter-protein ion pairs found in the LmP-LmCytc complex, but calorimetric studies show that the DH of association is -2.1 kcal/mol, while $-TDS$ is -6.0 kcal/mol (207). Although initial complex formation may well be driven by complimentary electrostatic surfaces, the predominance of the entropic term in the association free energy is another indication that the formation of the yeast CCP-Cytc complex is driven primarily by nonpolar interactions.

Here we have used Brownian dynamics (BD) simulations to investigate the mechanism of LmP-LmCytc complex formation. The initial association reaction is often very difficult to dissect experimentally since many of these reactions are near diffusion controlled. BD simulations, in which protein molecules are treated as rigid bodies with atomically detailed interactions in an implicit solvent, have been employed extensively to model the formation of many ET protein complexes, thus providing molecular level details on the initial association reaction that are often experimentally inaccessible (11, 12, 34, 208-224). In the case of the LmP-LmCytc system, comparison of the individual and co-crystal structures reveals no major conformational changes and only minimal reorientations of side chains at the interface upon complex formation. Thus, rigid-body BD simulations should be capable of accurately describing the association of LmP and LmCytc. We validated our BD simulations of the formation of the LmP-LmCytc complex by comparing computed second order association rates with measured association rate constants

determined from single turnover experiments. Finally, we carried out BD simulations of the formation of the yeast CCP-Cytc complex to gain insights into the differences in the mechanisms of association of LmP and yeast CCP with their respective cytochromes c.

Materials And Methods

Brownian Dynamics Simulations.

The BD method that we employ models proteins as atomically-detailed rigid bodies, and generates translational and rotational displacements using the Ermak-McCammon algorithm (225). The SDA program (226), versions 6 and 7 (227), was used to carry out the BD simulations. SDA version 6 was used to compute the probability distributions and version 7 was used to calculate the second order association rate constants (both methods are detailed below). The SDA energy function, from which the forces that drive the BD simulations are derived, contains four contributions (226): interactions of the charges on one protein with the electrostatic potential due to another protein; electrostatic desolvation due to the charges on one protein entering the low dielectric cavity of another protein; nonpolar desolvation due to the burial of solvent-accessible surface atoms on one protein in the cavity of another protein; and short-ranged repulsion modeled using an exclusion volume that prevents overlap of the proteins.

The electrostatic interactions were treated in the BD simulations by using a well-established approach based on the effective charge approximation (228). The electrostatic potentials were obtained by numerical solution of the finite-difference Poisson-Boltzmann (PB) equation, taking into account the inhomogeneous dielectric medium and the surrounding electrolyte solution, using the APBS 1.4 program package (229). We utilized a manual multigrid calculation with a single Debye-Hückel boundary condition for each solute on a $200 \times 200 \times$

200 grid of 1 Å spacing. The CHARMM22 force field (230) was used to assign atomic charge and radii parameters for all atoms of the protein and heme groups. The dielectric boundary was defined by the Van der Waals surface of the protein. Harmonic dielectric boundary smoothing and charge antialiasing (231) were employed as implemented in APBS 1.4. The solute and protein dielectric constants were set to 78 and 1, respectively, at 298.15 K for all electrostatic potential grid calculations. Ionic strengths of 50 mM and 150 mM NaCl were used for LmP-LmCytc to match experimental conditions, and 50 mM, 100 mM and 300 mM NaCl were used for the yeast CCP-Cytc to compare probability distributions with LmP-LmCytc and to compare to rates experimentally measured and computed from BD simulations. Ionic charges of +1 and -1 and radii of 1.1 and 1.7 Å, respectively, were used to represent Na⁺ and Cl⁻, and the solvent radius was set to 1.4 Å. Charges at effective charge sites on basic and acidic side chains were fit so that, in a shell around the protein, the electrostatic potential that was precomputed with an inhomogeneous dielectric is reproduced with a homogeneous dielectric (228). The exclusion probe radius and skin thickness were set to 4 and 3 Å, respectively. For consistency, we used the same ionic strengths and solvent dielectric constant as the computed electrostatic potential. For each BD timestep, forces and torques acting on each protein are computed by placing its effective charges over the other protein's electrostatic potential grid (226).

The electrostatic desolvation energy grids (200 × 200 × 200 with a 1 Å spacing) were computed using equation 2 in reference (219) with the scaling factor α set to 1.67. Ionic strengths were varied to be consistent with the values used to compute the electrostatic potential grids with APBS. A single ionic radius of 1.5 Å was used. The nonpolar desolvation energy grids (200 × 200 × 200 with a 1 Å spacing) were computed using equation 3 in reference (219) with the distance parameters a and b set to 3.10 Å and 4.35 Å, respectively, and the normalization

constants c and b set to 0.5 and -0.0065 kcal/(mole·Å²), respectively. The exclusion grids were calculated as described in reference (211) using a probe radius of 1.77 Å.

In our BD simulations of the formation of ET complexes, the larger of the two proteins (LmP or yeast CCP) is held fixed and the other (LmCytc or yeast Cytc), is allowed to move. The larger protein was placed at the center of a sphere of radius 300 Å while the smaller protein was allowed to diffuse from a random starting position on the surface of the sphere. A single trajectory can either result in a reaction event (defined below) or an escape event, in which the center-of-mass to center-of-mass displacement between the two proteins reaches 400 Å. The integration timestep was linearly decreased from 20 ps to 1 ps over a protein surface-to-surface distance interval of 90 Å to 60 Å to allow for the simulation to capture the finer details of the protein dynamics at close proximity, while maintaining efficient sampling at large separation distances. The diffusion coefficients that appear in the Ermak-McCammon algorithm were calculated using the Hydropro 10 program (232). The translational and rotational diffusion coefficients used in our BD simulations were: 8.634×10^{-3} Å²/ps and 1.021×10^{-5} radians²/ps, respectively, for LmP, 1.161×10^{-2} Å²/ps and 2.439×10^{-5} radians²/ps, respectively, for LmCytc, 8.264×10^{-3} Å²/ps and 8.734×10^{-6} radians²/ps, respectively, for yeast CCP, and 1.177×10^{-2} Å²/ps and 2.482×10^{-5} radians²/ps, respectively, for yeast Cytc.

Two separate sets of structures were used to initiate BD simulations of the LmP-LmCytc complex: (i) the individual structures of each protein solved separately (PDBID 3RIV for LmP (92) and PDBID 4DY9 for LmCytc (91)), and (ii) the structure of each protein in the co-crystal structure of the complex (PDBID 4GED (93)). There are no major conformational differences between the two sets of LmP and LmCytc structures, but rather a series of small rearrangements of side chains at the binding interface, with the largest change being the movement of Arg24 of

LmCytc into what is believed to be the active conformation for ET. The same BD simulation protocols were applied to both sets of initial structures to investigate the roles that subtle rearrangements of the binding interface play in complex formation. In addition to wild type LmP we considered the following mutants in which negatively charged side chains were changed to neutral side chains that preserve the side chain orientation: D211N, D47N, E49Q, D50N, E54Q, D47N-D50N-E54Q. We also considered the D47A-D50A-E54A mutant because it provided an additional opportunity to compare computed and experimentally determined association rates. The *in silico* mutations were carried out using the psfgen plugin of VMD starting with the wildtype LmP structure and manually truncating uncommon atoms of the sidechain to preserve the sidechain orientation in the mutated residue. This was done for the D to N and E to Q mutations by truncating the terminal oxygen (OD2), while the corresponding D to A and E to A mutations required truncation to the beta-carbon (CB). A total of 200,000 BD trajectories were generated at 298.15 K for each of the two sets of starting structures for each LmP variant. To investigate the consequences for complex formation of differences in the sequences of cytochrome c peroxidases from *L. major* and yeast, we also generated 200,000 BD trajectories for the yeast CCP-Cytc system, starting from the protein structures in the ET complex (PDBID 2PCC (75)).

Calculation of Second Order Association Rates.

BD association rates are calculated using a method developed by Northrup et al. (233) according to which the diffusion-controlled reaction rate is given by:

$$k = k_D(b)p,$$

where p is the probability that the pair of proteins, starting at separation $R = b$, will associate and react (i.e., not diffuse apart, ultimately to infinite separation), $k_D(b)$ is the rate constant for pairs with $R > b$ to reach $R = b$, regardless of whether or not they ultimately react. In practice, p is not calculated directly from BD trajectories. Rather, trajectories are terminated when R exceeds a cutoff distance q , with $q > b$, and the following expression for the rate constant, which accounts for the truncation of trajectories that lead to $R > q$, is used:

$$k = \frac{k_D(b)\beta}{1 - (1 - \beta)\Omega}$$

Here b is the fraction of trajectories in which the complex forms before the proteins diffuse to $R > q$, and $W = k_D(b)/k_D(q)$. This equation is based on the assumption that every trajectory in which an active site collision occurs, as defined by a specified criterion (typically, a set of interprotein distances; see below), produces a reaction. For fast reactions such as ET, this assumption is justified. In practice, b and q are chosen to be sufficiently large that the potential of mean force (PMF) for the interaction between the two proteins is centrosymmetric, so that $k_D(r)$ ($r = q$ or b) may be calculated using the following analytical expression:

$$k_D(r) = \left[\int_r^\infty dR \frac{e^{U(R)/k_B T}}{4\pi R^2 D(R)} \right]^{-1}$$

where $U(R)$ is the PMF and $D(R)$ is the relative diffusion constant of the two proteins. In the present work, we used $q = 400 \text{ \AA}$ and $b = 300 \text{ \AA}$. Error estimates on the computed second order association rates were obtained by using the bootstrapping analysis implemented in SDA 7 (179, 227).

The association of LmP and LmCytc was tracked by monitoring distances between two pairs of side chains (LmP E49–LmCytc K98 and LmP D24–LmCytc R211) that form inter-protein salt bridges in the crystal structure of the complex. The specific distances monitored

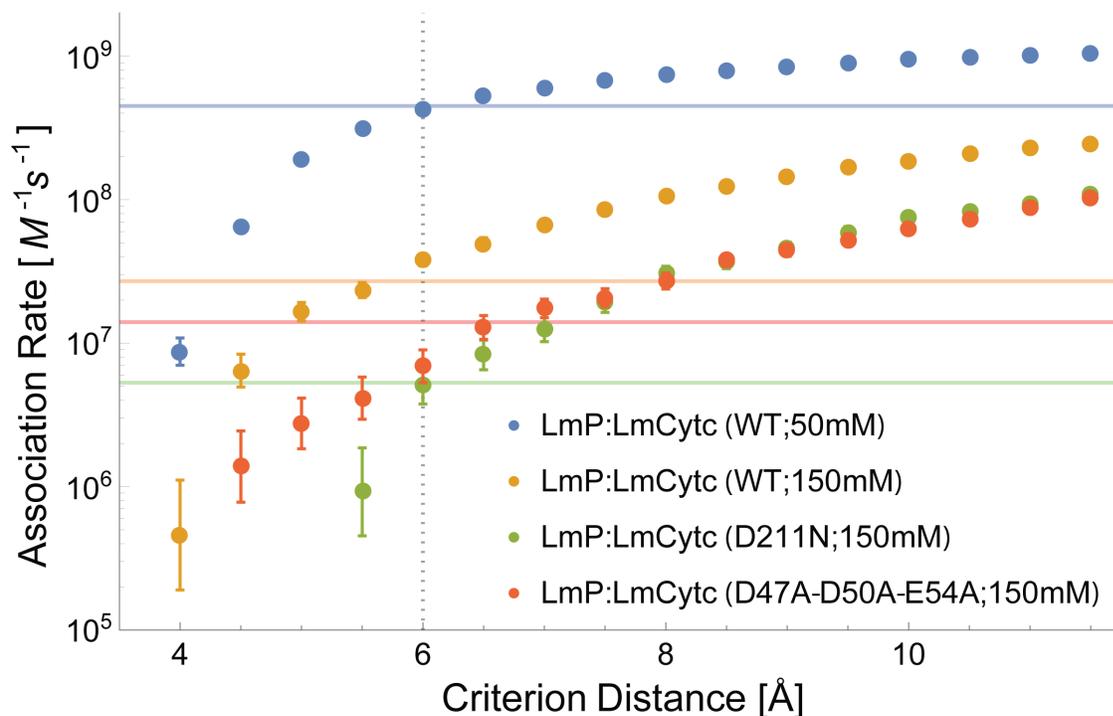


Figure 4.1. Calculated rate constants for LmP-LmCytc. Calculated second order association rate constants for LmP:LmCytc from 200,000 BD trajectories, plotting the occurrence of two crystal contacts as function of criterion distance. The solid colored lines indicate the experimentally determined association rate constants (k_1) for LmP:LmCytc for the different conditions/mutations indicated in the legend. The dotted vertical line at 6 Å shows the criterion distance that produces the best overall agreement with the experimentally determined rates.

were those between OE1 and OE2 of E49 and NZ of K98, and OD1 and OD2 of D211 and NE and NH1 of R24, respectively. An association “reaction” was considered to occur when any of the distances in both pairs of sidechains were less than a criterion distance, d_{\max} . The particular value of d_{\max} used, 6 Å, produced second order association rates that are in good agreement with experimental data on wild type LmP and its D211N and D47A-D50A-E54A variants (Figure 4.1).

Because the yeast CCP-Cytc complex does not contain inter-protein salt bridges, we used the `nda-pairs` function in SDA 7 to produce a list of 10 independent, nonspecific contacts (donor-acceptor atom pairs between the two proteins in the complex structure with a maximum distance

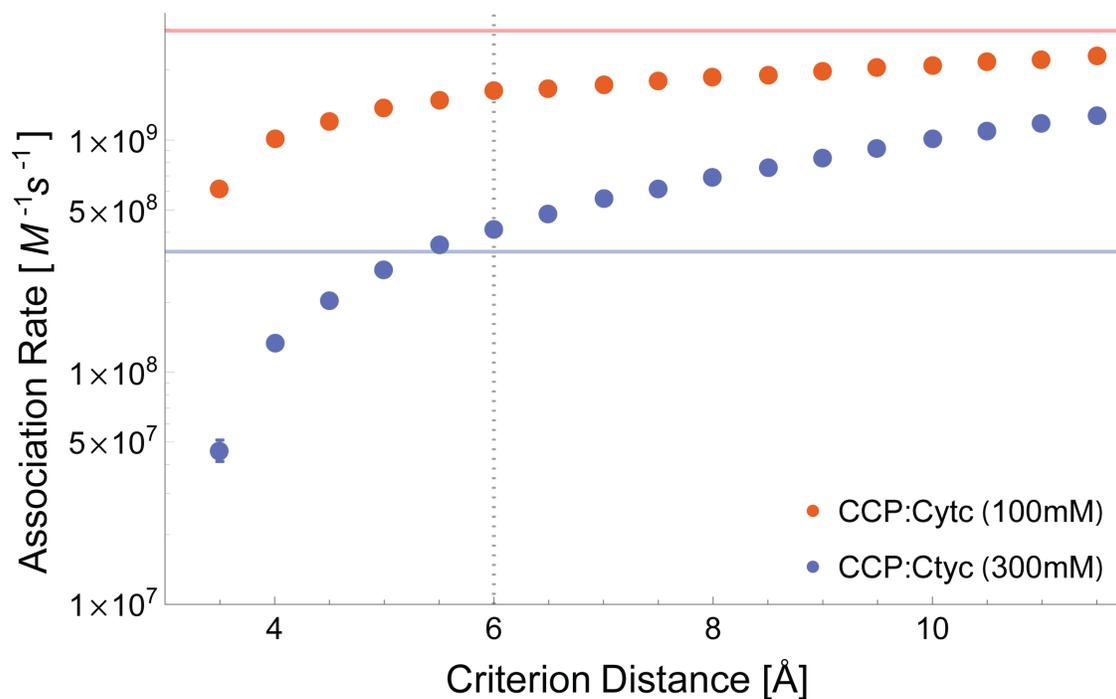


Figure 4.2. Calculated rate constants for CCP-Cytc. Calculated second order association rate constants for yeast CCP:Cytc from 200,000 BD trajectories, using the occurrence of two non-specific contacts as function of criterion distance. The solid red and blue lines indicate the experimentally measured association rate constants (k_1) for yeast CCP:Cytc at 100 mM and 300 mM ionic strength, respectively. The dotted vertical line indicates the criterion distance used in the calculation of the second order association rate from our BD trajectories.

of 4.5 Å between each pair) to define the association of yeast CCP and Cytc. An association reaction was considered to occur when two atom pairs achieve a separation distance less than d_{\max} simultaneously. The value of d_{\max} that we used to define association of yeast CCP and Cytc, 6 Å, provided good agreement with experimental association rates measured at 100 mM and 300 mM ionic strengths (Figure 4.2).

Probability Distributions.

In order to identify the formation of particular complexes during the course of thousands of BD trajectories, probability histograms of the position of the center-of-mass of the mobile protein are constructed and displayed as three-dimensional isosurfaces (contours), superimposed on a display of the fixed protein. We used a voxel width of 1 Å on a cubic grid of 200 × 200 × 200 voxels and a sampling sphere radius of 60 Å to capture configurations generated with a 1 ps timestep. Probabilities within each voxel were calculated by dividing the number of trajectory points within each voxel by the total number of trajectory points within the sampling sphere. The resulting probability distributions include orientational information implicitly because sharp features correspond to consistently oriented complexes. All of the probability isosurfaces we report are contoured at the same value of 2.2×10^{-3} so that isosurfaces from different systems may be readily compared to one another.

Molecular Dynamics Simulation.

In order to validate an initial encounter complex that was prevalent in the BD trajectories initiated from the individual structures, we carried out an atomistic molecular dynamics (MD) simulation. We selected a configuration from a BD simulation in which LmCytc was docked to helix A of LmP and we moved LmCytc about 15 Å away from LmP prior to initiating the MD simulation. The individual crystal structures of LmP (PDBID 3RIV (92)) and LmCytc (PDBID 4DY9 (91)) were used as the initial configurations of the two proteins in the MD simulation. Hydrogen atoms were added to both the protein and the crystallographic water molecules using psfgen with VMD 1.9.1 (179). Patches were used to connect the heme with the coordinating His residues in both structures, and an extra bond was added to describe the Met-Fe coordination in LmCytc. The Ca^{2+} and K^{+} metal ions present in each of the crystal structures were modeled by

adding extra bonds between the coordinating atoms and the metal ion to aid in preserving the coordination site. The solvent box was built using a 15 Å cushion in all three directions around the proteins. The resulting system contained 219,540 atoms.

The MD simulation was performed using the program NAMD, version 2.9 (234). The CHARMM22 force field (230) was used for the proteins, and the TIP3P model (235) for water. The system was subjected to 1,000 steps of conjugate gradient energy minimization prior to MD simulation. The simulation was carried out for 100 ns at a constant temperature of 300 K and a constant pressure of 1 atm using Langevin dynamics for temperature control and a Nosé-Hoover-Langevin piston for pressure control (236, 237). A multiple timestep algorithm (238) was used to integrate the equations of motion with timesteps of 4 fs for the electrostatic interactions and 2 fs for the bonded and short-ranged nonbonded interactions. The smooth particle mesh Ewald algorithm (239) was used to treat the electrostatic interactions, and the real-space part of the Ewald sum and the Lennard-Jones interactions were smoothly switched off between 10 Å and 12 Å. The SHAKE algorithm (240) was used to constrain the lengths of bonds to hydrogen atoms.

Protein Cloning, Expression, and Purification.

LmP: The wild type LmP construct expressed without the N-terminal hydrophobic tail as $\Delta 34$ LmP was supplied by Dr. Subrata Adak (201). The gene was first amplified by PCR, purified using the Nucleospin® gel and PCR cleanup extraction kit, and subsequently ligated into the pET23a (Novagen) vector from the NdeI site at the 5'-end to the XhoI site at the 3'-end. The pET23a vector includes a 6X-His tag at the C-terminus. Site-directed mutagenesis was carried out by PCR: the Stratagene® Site-Directed Mutagenesis kit was used for D211N mutant preparation, while the TaKaRa PrimeSTAR Max (Clontech) was used to generate each mutation

in the D47A-D50A-E54A helix A triple mutant. Each plasmid was fully sequenced to ensure the fidelity of the PCR reaction.

The LmP plasmid was transformed into *E. coli* BL21(DE3) cells and plated onto LB (Lysogeny Broth) agar with ampicillin (100 µg/mL). A single colony was used to inoculate each 5 mL of LB starter culture (100 µg/mL ampicillin). The culture was incubated for 8 hours at 37 °C and 220 rpm agitation. Each liter of TB (Terrific Broth) medium (ampicillin 100µg/mL) was inoculated with 2mL of LB starter culture. The cells were grown at 37°C with 220 rpm agitation in a New Brunswick Scientific C25KC incubator until an OD600 of 1.8 was obtained. Protein expression was induced with 0.4 mM δ-aminolevulinic acid, 0.5 mM Isopropyl β-D-thiogalactoside, and a new dose of ampicillin was added. Post-induction cells were then incubated for 24 hours, after which they were harvested by centrifugation and stored at –80 °C.

Cells were thawed overnight at 4 °C and resuspended by stirring for 3 hours at 4 °C in Buffer A (50mM sodium phosphate, pH 7.5, 10% glycerol, 1 mM PMSF, 1mM DTT). Additionally, 1 µg/mL pepstatin, and 1 µg/mL leupeptin were added to Buffer A before lysis. Cells were lysed by passing through a microfluidizer at 18 kpsi (Microfluidics International Co). The soluble fraction was isolated by centrifugation at 17,000 rpm and 4 °C for 1 hour. The crude extract was then loaded onto a Ni²⁺-nitrilotriacetate column previously equilibrated with 10 column volumes of Buffer A. The column was washed with 10 column volumes of 10 mM L-histidine in Buffer A before elution with 150 mM L-Histidine in Buffer A. Major colored fractions were pooled and concentrated in a 10,000 molecular weight cut-off Amicon concentrator at 4 °C and then loaded onto a Superdex 75 16/60 column previously equilibrated with 2 column volumes of Buffer B (50 mM potassium phosphate, pH 7.0 and 5% glycerol). Fractions were pooled according to the following criteria: an $R_z(A_{408}/A_{280}) > 1.2$ and

spectroscopic observation of the formation of Compound I upon addition of stoichiometric amounts of hydrogen peroxide. Sample homogeneity was determined by SDS-PAGE. Pooled fractions were concentrated to 30 mg/mL.

LmCytc: The sequence of cytochrome c from *L. major* (strain Friedlin) was obtained from GenBank (241). LmCytc was synthesized by GenScript, with optimized codons for *E. coli* expression, into the construct pBPCYC1, replacing yeast cytochrome c. The LmCytc plasmid was expressed in *E. coli* BL21(DE3) competent cells and purified as previously described (91).

Steady-state Activity Assays.

Spectrophotometric steady-state activity assays monitoring the oxidation of reduced LmCytc were performed at 25°C on a Cary 3E UV-Visible spectrophotometer (Varian/Agilent). Reduction of LmCytc was performed by the addition of granules of sodium dithionite followed by an incubation period of 30 minutes on ice. Excess remaining sodium dithionite was then removed by passing the reduced sample through an Econopac 10DG desalting column (Bio-Rad) previously equilibrated with 50 mM potassium phosphate, pH 6.5. The eluate was then concentrated in a 10,000 molecular weight cut-off centricon at 4°C. The concentration of reduced LmCytc was determined using the previously determined molar extinction coefficient $\Delta\epsilon_{558}$ of 29 $\text{mM}^{-1}\text{cm}^{-1}$ (ref. (91)). The assay spanned two hours, during which negligible amounts of LmCytc spontaneously oxidized. The LmP concentration was determined using the Soret molar extinction coefficient ($\Delta\epsilon_{408} = 113.6 \text{ mM}^{-1}\text{cm}^{-1}$)(40). The hydrogen peroxide concentration was also measured using the molar extinction coefficient ($\Delta\epsilon_{240} = 0.0436 \text{ mM}^{-1}\text{cm}^{-1}$).

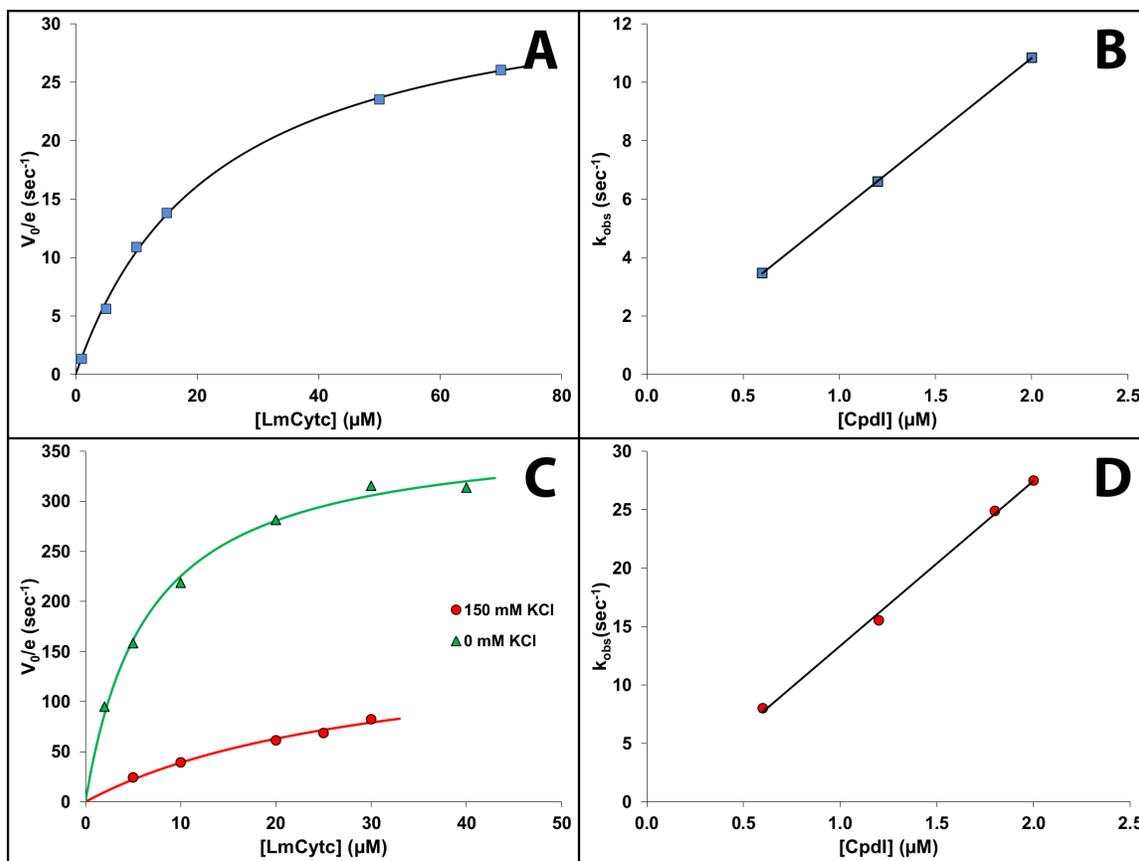


Figure 4.3 Experimental kinetics of LmP mutants. (A) V_0/e vs LmCytc concentration for LmP D211N in 25 mM potassium phosphate buffer, pH 6.5. (B) Plot of first-order k_{obs} as a function of LmP Compound I concentration obtained from single turnover stopped-flow experiment for LmP D211N. (C) Superposed V_0/e vs LmCytc concentration for triple helix A mutant LmP D47A-D50A-E54A in 25 mM potassium phosphate buffer, pH 6.5. The measurements depicted by the green triangles had no KCl added, while the red circles were measured in a reaction buffer containing an additional 150 mM KCl. The data in panels (A) and (C) were fit to the Michaelis-Menten equation (see Methods section). (D) Plot of first-order k_{obs} as a function of Compound I concentration obtained from single turnover stopped-flow experiment for LmP D47A-D50A-E54A. The reaction buffers used for all stopped-flow experiments was supplemented with 150 mM KCl.

In each experiment, the reaction buffer used was 25 mM potassium phosphate, pH 6.5 and the ionic strength was increased up to 150 mM by adding KCl. For each experiment, the reaction was initiated by adding 0.18 mM of hydrogen peroxide and the absorbance change was monitored at 558 nm for 1 minute. The turnover number was calculated using the previously determined $\Delta\epsilon_{558} = 19.4 \text{ mM}^{-1} \text{ cm}^{-1}$ (ref. (91)). The data were analyzed by fitting to the Michaelis-Menten equation:

$$v = \frac{V_{\max} [\text{LmCyt}c]}{K_M + [\text{LmCyt}c]}$$

Stopped-flow kinetics.

Stopped-flow kinetic measurements were performed as described previously (203). Compound I reduction was second-order, and the bimolecular rate-constant k_1 was calculated by fitting the kinetic traces using the software provided with the stopped-flow instrument (Applied Photophysics) according to the following standard single exponential equation:

$$A_{420} = Ce^{-k_{\text{obs}}t} + b,$$

where C is the amplitude term, k_{obs} is the observed rate constant for the decay of oxidizing LmCyt c , and b is an offset value. The association rate constant k_1 was determined as the slope of a fit of the (linear) concentration dependence of k_{obs} .

Crystal Preparation.

The LmP D211N protein sample was prepared for crystallization as previously described (203). Crystals were grown at room temperature in 10% PEG 5,000 MME (w/v), 0.1 M

MES:NaOH pH 6.5 and 7.5 mM praseodymium^{III} acetate hydrate in a hanging drop vapor diffusion setup. Freshly grown crystals were harvested after 24 hrs and passed stepwise through a cryo-protectant solution containing 30% glycerol (v/v) for 4 hrs at 4 °C. Crystals were then flash cooled with liquid nitrogen.

X-ray Diffraction Data Collection, Processing, and Structure Refinement.

Cryogenic (100 K) X-ray diffraction data were collected remotely at the Advanced Light Source (ALS) facility, using the data collection control software Blu-Ice (242), and a crystal

Table 4.1. Crystallographic data and refinement statistics of LmP D211N

PDB entry	5AMM
Radiation source	ALS BL 12.3.1
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell dimensions a, b, c (Å)	46.25, 78.31, 160.98
Data resolution (Å) (highest-resolution shell)	50.0 - 2.09 (2.17 - 2.09)
X-ray wavelength (Å)	1.01
Total no. of observations (highest shell)	69311 (6417)
No. of unique reflections (highest-resolution shell)	34997 (3268)
Completeness (%) (highest-resolution shell)	99.15 (94.86)
Rmerge (highest-resolution shell)	0.02603 (0.5488)
I/sigma (highest-resolution shell)	18.69 (1.35)
CC1/2 (highest-resolution shell)	0.999 (0.477)
Redudancy (highest-resolution shell)	2.0 (2.0)
B factor, Wilson plot (Å ²)	44.34
Number of protein atoms	4254
Number of heteroatoms	90
Number of waters	95
Disordered residues	(A) 301-303 (B) 301-303
R _{work} /R _{free}	0.1842/0.2445
Root-mean-squared deviation for bond lengths (Å)	0.016
Root-mean-squared deviation for bond angles (°)	1.700

mounting robot. An ADSC Q315r detector at beamline 12.3.1 was used for data collection. Raw data frames were indexed, integrated and scaled using XDS (243). The LmP D211N structure was solved by molecular replacement using phaser (244) and the wild type LmP structure (PDB entry 3RIV). The initial difference Fourier map was then calculated and the model was refined using REFMAC (245). Water molecules were also modeled using REFMAC and checked in COOT (246). The TLS (Translation-Liberation-Screw-rotation model) (247) protocol was implemented in the final stage of refinement with each chain as one TLS group and the refined structure was validated in COOT. Coordinates and structure factors were deposited in the RCSB Protein Data Bank. Crystallographic data collection and structure refinement statistics are summarized in Table 4.1.

Results And Discussion

Brownian Dynamics Simulations Initiated from the Co-crystal Structures Reveal the Importance of Negatively Charged LmP Side Chains in the Formation of the LmP-LmCytC Complex.

A probability isosurface for the center-of-mass of LmCytC around wild type LmP obtained from BD trajectories initiated from the co-crystal structures is shown in Figure 4.4A. The most prominent feature in the isosurface coincides with the position of the center-of-mass of LmCytC in the co-crystal structure (Figure 4.4B). Thus, the BD simulations of the co-crystal structures recapitulate the binding of LmCytC to the LmP active site, as established in the crystal structure of the complex. In the probability isosurface shown in Figure 4.4A, there is a second, relatively small feature corresponding to an interaction between LmCytC and helix A of LmP, which will be discussed in more detail below.

The second order association rate computed from the BD trajectories of the wild type co-crystal structures at 50 mM (NaCl) ionic strength, $4.3 \pm 0.1 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$ (Table 4.2), is in good

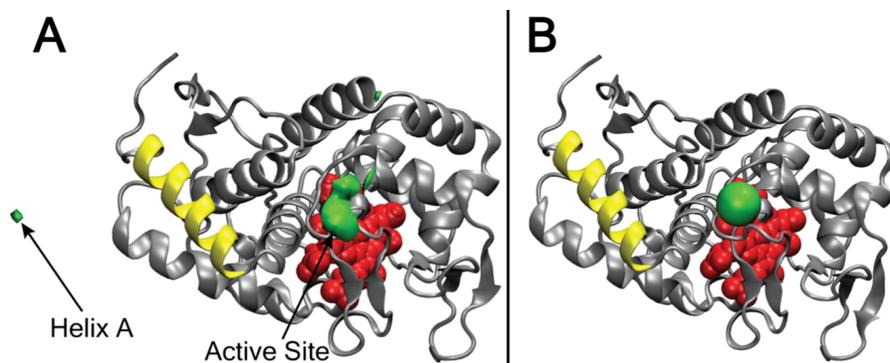


Figure 4.4. Structure of LmP in the LmP-LmCytc complex from BD simulations at 50 mM ionic strength. The backbone is drawn as a gray ribbon, except for helix A, which is drawn as a yellow ribbon. The heme is depicted with red van der Waals spheres. (A) The green surface is the probability isosurface for the LmCytc center-of-mass calculated from BD simulations initiated from the co-crystal structures contoured at a value of 2.2×10^{-3} . At this contour level, there are two features in the isosurface, the larger of which corresponds to LmCytc binding to the active site of LmP, and a smaller feature corresponding to LmCytc interacting with helix A of LmP. (B) The position of the center-of-mass of LmCytc in the co-crystal structure (93) is depicted as a green sphere. The molecular graphics images and probability isosurfaces in this figure were rendered using VMD (179).

Table 4.2 Second order association rates of LmCytc and LmP from Brownian dynamics simulations initiated from the co-crystal structures.

LmP variant ^a	ionic strength, mM	BD rate, $M^{-1}s^{-1}$
wild type	50	$4.3 \pm 0.1 \times 10^8$
	150	$3.8 \pm 0.4 \times 10^7$
D211N (AS)	50	$3.3 \pm 0.4 \times 10^7$
	150	$5.1 \pm 1.5 \times 10^6$
D47N (HA)	50	$2.1 \pm 0.1 \times 10^8$
	150	$4.6 \pm 1.4 \times 10^6$
E49Q (HA)	50	$1.3 \pm 0.1 \times 10^8$
	150	$2.8 \pm 1.0 \times 10^6$
D50N (HA)	50	$1.6 \pm 0.1 \times 10^8$
	150	$2.8 \pm 1.0 \times 10^6$
E54Q (HA)	50	$1.8 \pm 0.1 \times 10^8$
	150	$4.1 \pm 1.3 \times 10^6$
D47A-D50A-E54A (HA)	150	$6.9 \pm 1.8 \times 10^6$

^aAS denotes that the mutation is active site of LmP. HA denotes that the mutation is in helix A of LmP.

agreement with the measured association rate constant k_1 , $4.5 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$, determined by single turnover assay with 50 mM KCl in the reaction buffer (Table 4.3) (203). While k_{cat} for LmP is independent of ionic strength, K_M increases substantially, thereby resulting in a decrease in k_{cat}/K_M , with increasing ionic strength (203). Consistent with this observation, increasing the ionic strength from 50 mM to 150 mM in the BD simulations reduces the second order association rate by an order of magnitude (Table 4.2), in good agreement with the experimental data (Table 4.3).

Table 4.3. Kinetic parameters obtained from steady state assays of LmCytc oxidation by LmP^a

Enzyme	KCl added, mM	k_{cat} , s^{-1}	K_M , μM	k_1 , $\text{M}^{-1} \text{s}^{-1}$
wild type	0	409 ± 9^b	4.1 ± 0.3^b	
	50	531 ± 35^b	13 ± 2.1^b	4.5×10^8
	150	531 ± 74^b	83 ± 17^b	$2.7 \times 10^7^b$
D211N	0	34.5 ± 0.8	23 ± 1.3	
	150			5.3×10^6
D47A/D50A/E49A	0	372 ± 10	7 ± 0.6	
	150	162 ± 10	31 ± 8.6	1.4×10^7

^aFits to determine the kinetic parameters are shown in Figure S1.

^bValues previously reported (203).

To assess the role of electrostatic interactions, several charge neutralizing mutations in LmP were considered. When D211, which is located at the active site binding interface, is mutated to N *in silico*, binding is disrupted, as is evident from the lack of a feature at the active site in the probability isosurface shown in Figure 4.5A, middle column. The D211N mutation results in a substantial reduction of both the BD association rate (Table 4.2) and the experimental k_1 (Table 4.3) at 150 mM ionic strength. The good agreement between the BD predicted and

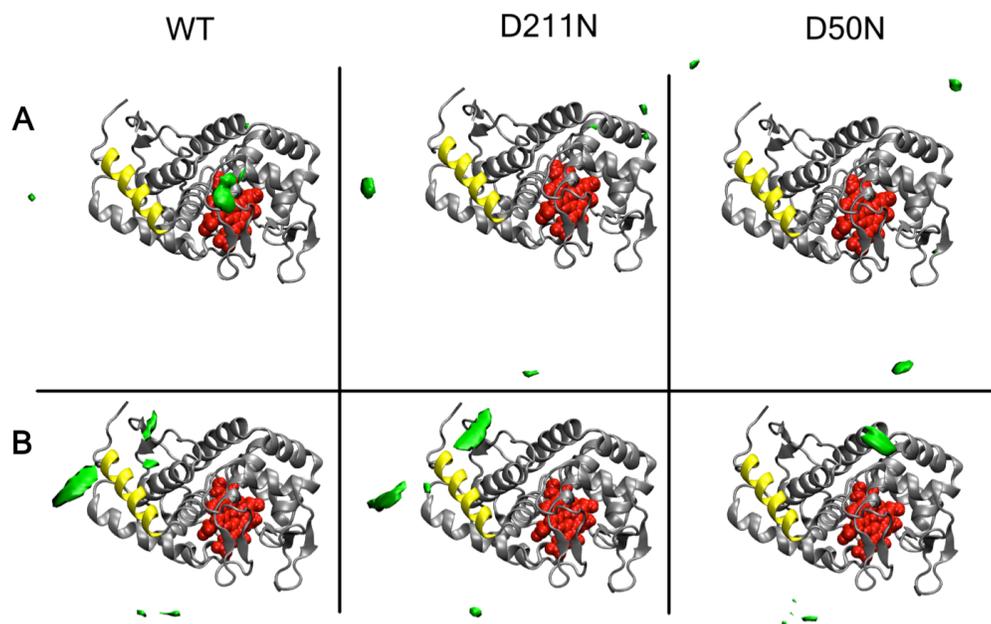


Figure 4.5. Structure of mutant LmP-LmCytc complexes from BD simulation. Probability isosurfaces contoured at a value of 2.2×10^{-3} (green) for the LmCytc center-of-mass computed from BD simulations at 50 mM ionic strength superposed on the structure of LmP (drawn and colored as in Figure 4.4). (A) BD simulations initiated from the co-crystal structures; (B) BD simulations initiated from the individual structures.

experimentally measured consequences of the D211N mutation provides further validation of the BD simulations.

The most significant interaction between LmCytc and D211N LmP occurs in the vicinity of helix A of LmP (Figure 4.5A, middle column). Helix A contains several negatively charged residues that protrude away from the active site interface (Figure 4.6). Elimination of negatively charged side chains on helix A in the D50N, D47N, E49Q, E54Q, and D47N-D50N-E54Q variants also reduces binding of LmCytc to the active site of LmP (Figure 4.5A, right column; Figure 4.7A). With the exception of E49Q, all of the charge neutralizing D to N and E to Q mutations also reduce the interaction of LmCytc with helix A of LmP. The computed association rates were significantly lower than wild type for all of the helix A D to N and E to Q mutants (Table 4.2).

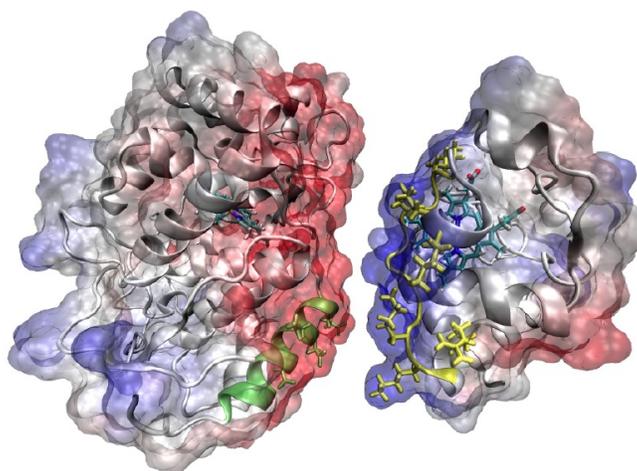


Figure 4.6. Overview of the LmP-LmCytC system. The LmP-LmCytC complex as solved by Jasion et al. is shown where LmP (left) and LmCytC (right) have been separated to highlight the highly complimentary electrostatic surfaces. The backbone of each protein is shown in cartoon representation and electrostatic potentials are displayed on transparent solvent-accessible surfaces where negative potential is red and positive is blue. The backbone of helix A of LmP is colored green. Negatively charged side chains on helix A of LmP are shown in green licorice and positively charged side chains of LmCytC at the binding interface are shown in yellow licorice.

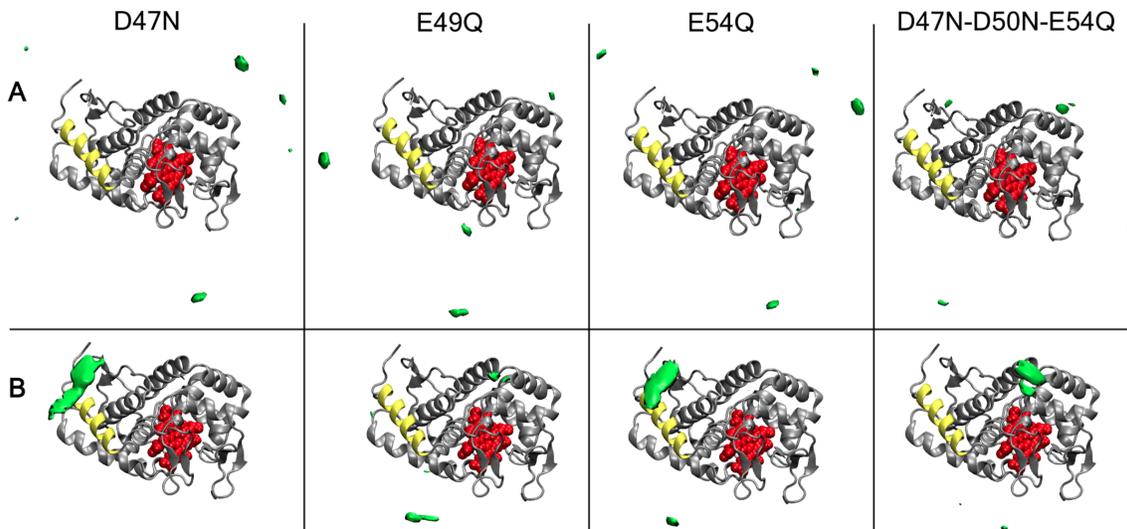


Figure 4.7. Further structures of mutant LmP-LmCytC complexes. Probability isosurfaces contoured at a value of 2.2×10^{-3} (green) for the LmCytC center-of-mass computed from BD simulations at 50 mM ionic strength superposed on the structure of LmP, drawn and colored as in Figures 1 and 2 for four additional mutants of helix A in LmP (D47N, E49Q, E54Q and D47N-D50N-E54Q). (A) BD simulations initiated from the co-crystal structures; (B) BD simulations initiated from the individual structures.

The importance of the helix A charges was confirmed by steady state assays of the LmP D47A-D50A-E54A triple mutant: at 150 mM ionic strength the k_1 for the triple mutant is ~19 times lower than that of wild type LmP (Table 4.3). The ~6X reduction in the second order association rate of the triple mutant vs. wild type computed from the BD trajectories (Table 4.2) agrees qualitatively with the reduction observed experimentally.

Structure of the D211N Mutant of LmP.

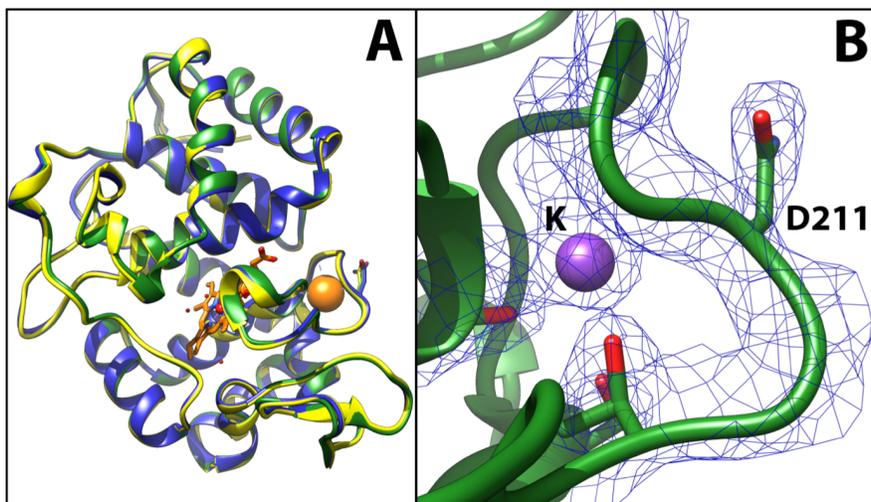


Figure 4.8 Crystal structure of the D211N LmP mutant. (A) Superposition of the α -carbon backbones of the wild type LmP crystal structure (blue), the LmP D211N crystal structure (green), and the LmP D211N structure that was prepared by *in silico* mutation for the BD simulations (yellow). (B) Potassium ion binding site, with the $2F_o - F_c$ density map contoured at 1.0σ . The molecular graphics images in this figure were made using UCSF Chimera (248).

LmP shares a common feature with many other peroxidases in that it has a structurally important cation binding site near the proximal side of the heme (Figure 4.8B). In LmP, this is either K^+ or Na^+ and the carbonyl O atom of D211 provides one of the ligands. Since we were concerned that the D211N mutation might cause structural perturbations around the cation site, we determined the crystal structure of the D211N mutant. As shown in Figure 4.8B, the cation

site remains fully occupied and there are no significant structural perturbations. Therefore, any change in kinetics is due solely to the mutation and not to structural changes. The D211N structure used in the BD simulations, which was prepared by *in silico* mutation before the D211N crystal structure was solved, is essentially identical to the crystal structure (Figure 4.8A).

Brownian Dynamics Simulations Initiated from the Individual Structures Suggests a Key Role for Helix A of LmP in the Formation of the LmP-LmCytC Complex.

In the co-crystal structures, the side chain conformations in the two proteins are optimized at the binding interface for forming a tight ET complex. In the individual structures, the conformations of the side chains at the binding interface are slightly different (Figure 4.9), and are presumably more representative of the conformations of the proteins when they first encounter one another in solution, before the formation of the tight ET complex. Thus, BD simulations initiated from the individual structures permit us to sample configurations of the initial encounter complex, and determine which interactions bring the two proteins together initially.

The probability isosurface obtained from BD trajectories of the individual wild type structures is shown in Figure 4.5B, (left column). In contrast to the case of the wild type co-crystal structures (Figure 4.5A, left column), there are no features in the isosurface near the LmP active site for the individual wild type structures. This suggests that minor adjustments of the conformations of side chains at the binding interface are required for the formation of the tight ET complex. The tiny feature near the N-terminal end of helix A of LmP in the isosurface from the simulations of the co-crystal structures (Figure 4.5A, left column) is significantly enhanced in the isosurface from the simulations of the individual structures (Figure 4.5B, left column). The

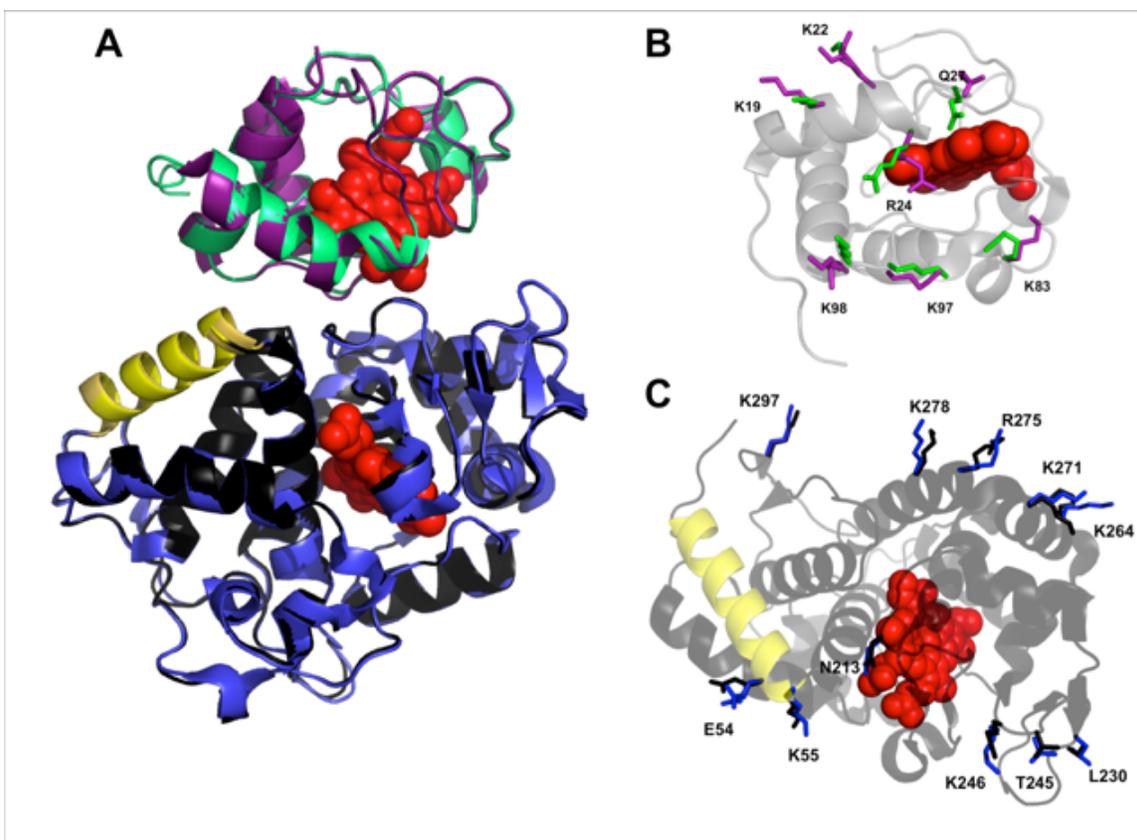


Figure 4.9. Differences between individual and co-crystal structures of LmP and LmCytC. (A) Overlays of the individual and co-crystal structures of LmCytC (top) and LmP (bottom). For LmP, the individual structure is shown in black and helix A is colored yellow while the co-crystal structure is colored in blue and helix A in orange. For LmCytC, the individual structure is shown in green while the co-crystal is displayed in purple. In both cases, the heme is shown as red spheres. Views of the (B) LmCytC and (C) LmP interfaces are shown highlighting residues that show different sidechain conformations between the individual and co-crystal structures. In (B) the sidechains of the individual structure are shown in green and the co-crystal structure in purple. In (C) the sidechains of the individual structure are shown in black and the co-crystal structure in blue.

broadness of this feature in the case of the individual structures is indicative of local sampling of LmCytC around helix A of LmP. The existence of the feature near helix A suggests that interactions at the binding interface that are required for the formation of the final complex are not necessary to bring the proteins together initially. This hypothesis is supported by the persistence of the helix A feature in the isosurfaces obtained from both sets of simulations of the

D211N mutant (middle column of Figure 4.5A and 4.5B), which do not display a feature in the isosurface near the active site in either set of simulations (co-crystal or individual structures).

The key role for negative charges on helix A suggested by the BD simulations of the co-crystal structures of mutants in which helix A charges were neutralized is reinforced by simulations of the corresponding mutants of the individual structures. The helix A feature is absent in the isosurfaces from the simulations of the D50N, E49Q, and D47N-D50N-E54Q variants initiated from the individual structures (Figure 4.4A and Figure 4.7B, second and fourth columns). The helix A feature is present, but shifted significantly away from its position in the isosurfaces of both the wild type and active site D211N mutant in the D47N and E54Q variants (Figure 4.7B, first and third columns). Thus, consistent with the simulations initiated with the co-crystal structures, the simulations initiated from the individual structures predict that interactions of LmCytC with helix A of LmP are altered by elimination of negative charges on helix A.

Two-step “Bind and Crawl” Mechanism for Formation of the LmP-LmCytC Complex.

The BD simulations initiated from the co-crystal structures exposed two binding sites for LmCytC on wild type LmP, one at the active site and the other near helix A of LmP (Figure 4.4A). BD simulations initiated from the individual structures, in which the side chain conformations are not optimized for active site binding, show persistent interactions between LmCytC and helix A of wild type LmP (Figure 4.5B, left column).

The observation that LmCytC binds both to the active site and helix A of LmP raises the question: does helix A compete with the active site for LmCytC, or is binding to helix A a prelude to binding to the active site? Visual inspection of the BD trajectories reveals that

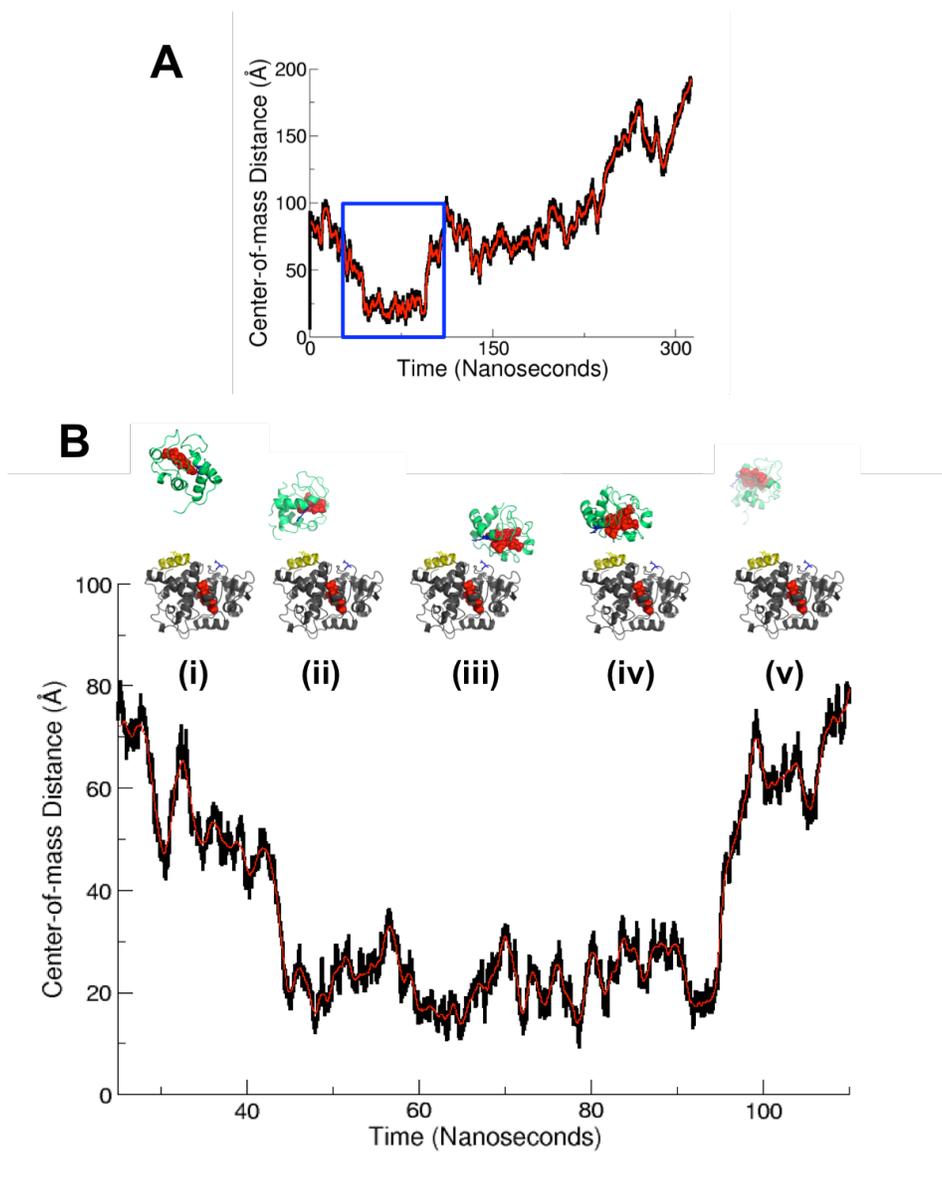


Figure 4.10. Representative Brownian dynamics trajectory illustrating the “bind and crawl” of LmCytc at helix A of LmP. (A) Distance between the electron transfer residues of LmP D211(C γ) and LmCytc R24(C ζ). The blue box designates the region highlighted in panel B. (B) Zoom of the helix A bind and crawl association of LmCytc to LmP. Snapshots from the trajectory are shown above the plot to visualize the association and disassociation of LmCytc to LmP, where LmP is shown in black with helix A in yellow, LmCytc in green, respective heme cofactors in red, and the catalytically ET active residues LmP D211 and LmCytc R24 in blue sticks. Association can be broken down into five phases: (i) approach of LmCytc to LmP; (ii) binding of LmCytc to helix A of LmP; (iii) movement of LmCytc from helix A to the co-crystal active site of LmP; (iv) movement of LmCytc back from the active site of LmP to helix A; and (v) disassociation of the LmP-LmCytc complex.

LmCytc first binds to helix A, and then travels down helix A towards the co-crystal interface, stopping at the end of the helix before jumping from the end of the helix to the interface where it begins local sampling of the active site region of LmP. A representative BD trajectory, initiated from the individual structures, in which LmCytc comes within 3 Å (center-of-mass distance, see Methods) of the co-crystal structure, is depicted in Figure 4.10. By tracking the LmCytc center-of-mass in the BD trajectories that were initiated from the co-crystal structures, we found that LmCytc binding to the active site was preceded by binding to helix A in ~70% of the reactive trajectories.

The BD trajectories suggest that helix A plays two roles in the formation of the complex, one as the initial point of contact for LmCytc on LmP, and the second as a runway to guide LmCytc to the ET interface. Based on the BD trajectories, we propose that the formation of the LmP-LmCytc ET complex proceeds, as conjectured previously (93), via a two-step “bind and crawl” mechanism in which LmCytc first binds to helix A, and subsequently migrates to the active site, with concomitant changes in the conformations of side chains at the binding interface enabling the formation of a tight complex.

Molecular Dynamics Simulation of Encounter Complex Formation.

To examine the importance of helix A in more detail, we performed an atomistic MD simulation. A configuration from a BD simulation of the wild type proteins initiated from the individual structures, in which LmCytc is close to helix A of LmP, was selected as the starting point. The two proteins were then moved further away from each other along the separation vector connecting their geometric centers to the point where their respective surfaces were 15 Å apart. Upon initiation of the simulation, LmCytc returned rapidly (within 2 ns) to helix A near its

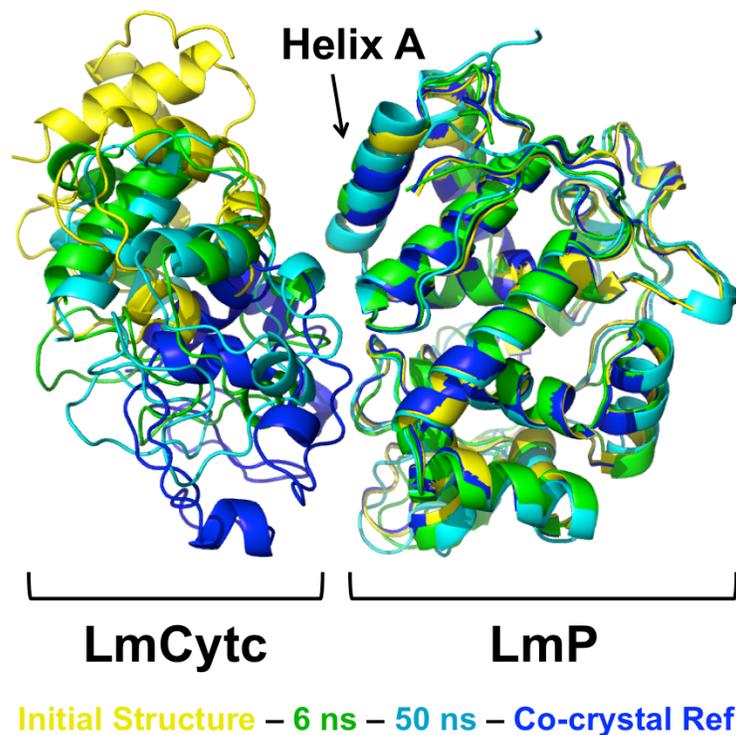


Figure 4.11. Molecular dynamics of the LmP Helix A centered complex. Snapshots after 6 ns (green) and 50 ns (cyan) from the MD simulation of LmP and LmCytc initiated from a configuration that was prepared by selecting a configuration from a BD simulation in which LmCytc was docked onto helix A of LmP and then displacing LmCytc by 15 Å. For comparison, the positions of the two proteins in the complex (co-crystal structure (93)) are shown in dark blue while the initial starting configuration of the system derived from a Brownian dynamics simulation is shown in yellow. This molecular graphics images in this figure were drawn using PyMol (www.pymol.org).

location in the configuration selected from the BD simulation. Once bound to helix A, LmCytc began to travel down the helix towards its location in the co-crystal structure. The movement down the helix, depicted in successive snapshots in Figure 4.11, was completed in 20 ns, reaching the closest approach to the co-crystal structure at approximately 50 ns. At that point, LmCytc reached the end of the helix and remained there without jumping to the active site interface for the remainder of the 100 ns MD simulation. The motion of LmCytc relative to LmP was also monitored by measuring the distance between the C_ζ atom of R24 of LmCytc and the C_γ

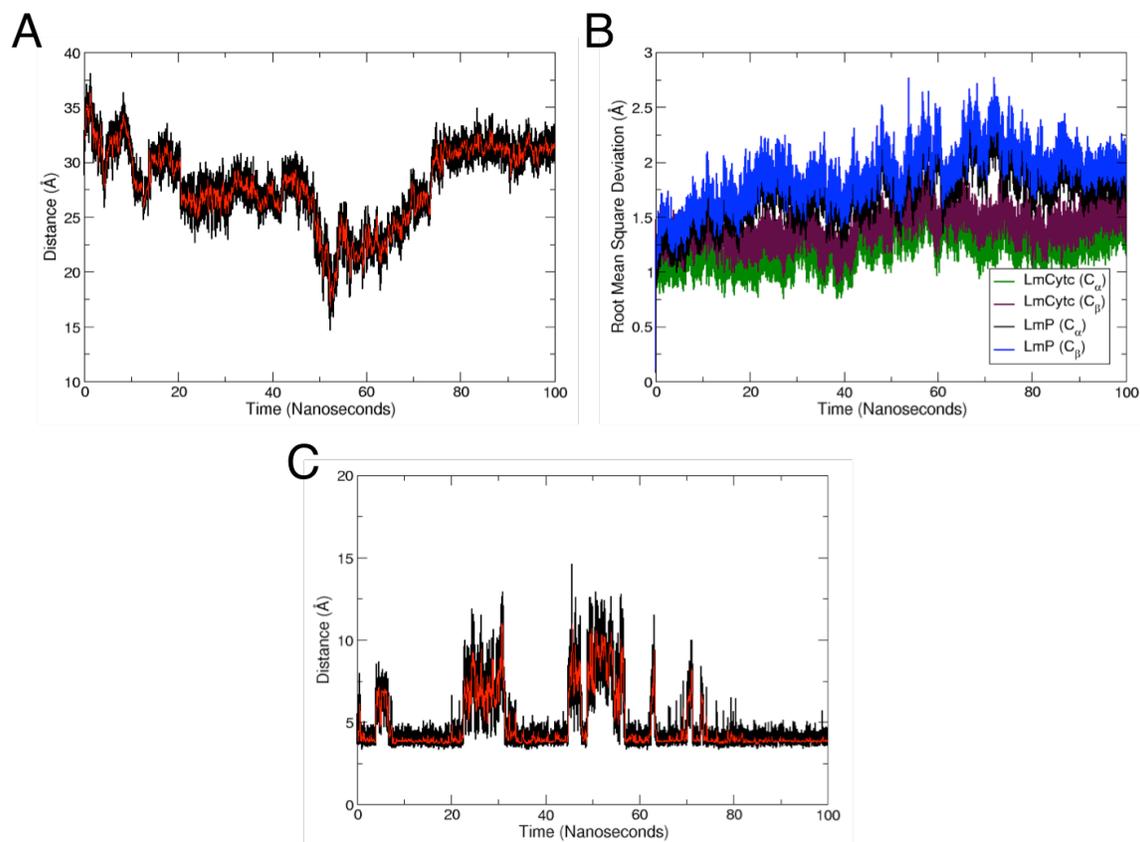


Figure 4.12. Structural parameters from the LmP-LmCytc molecular dynamics simulation. (A) Evolution of the distance between LmP D211(C_{γ}) and LmCytc R24(C_{ζ}) is shown in black with a running average with a window of 100 picoseconds shown in red. Upon the binding to helix A of LmP, LmCytc begins to move towards the LmP active site, reaching its closest approach at approximately 52 ns, before LmCytc moving back and forming a stable complex with helix A for the remainder of the simulation. (B) Root mean-squared deviation (RMSD) of all C_{α} and C_{β} atoms for LmP and LmCytc as labeled. Each structure achieves a stable state within in the first 30 ns of the simulation. (C) Evolution of the distance between LmCytc R24(C_{ζ}) and LmCytc E101(C_{δ}) to track the intermolecular R24-E101 ion pair. A running average with a window of 100 frames is displayed in red. The closest approach of LmP D211 to LmCytc R24 in (A) occurred during one of the short periods when the intermolecular ion pair was broken.

atom of D211 of LmP, which is approximately 6 Å in the co-crystal structure. This distance reaches its maximum value of ~38 Å when LmCytc first binds to helix A of LmP, and is gradually reduced to its minimum value of ~15 Å as LmCytc is guided toward the active site interface along helix A (Figure 4.12A).

The root mean-squared deviations (RMSDs) of the backbone C_α and side chain C_β atoms remained in the range 1.75 Å to 2 Å for LmP and 1 Å to 1.75 Å for LmCytc during the latter half of the MD simulation (Figure 4.12B). The low values of the RMSDs are consistent with the observation that the conformations of both proteins are very similar in the crystal structures of the individual proteins and the structure of the complex (93). The main difference is that, in the complex, R24 of LmCytc is in a different rotameric state that enables it to form a salt bridge with D211 in LmP, whereas in the individual structure of LmCytc R24 forms an intramolecular salt bridge with LmCytc E101. We posit that the change in the rotameric state of R24 is a crucial step along the way from the initial encounter complex, where LmCytc is docked onto helix A of LmP, to the active ET complex. The intramolecular salt bridge remained intact during a large portion of the MD simulation (Figure 4.12C), and this likely explains why the LmCytc R24–LmP D211 distance remained larger than 15 Å throughout the simulation.

Effect of Helix A Mutations on Kinetics.

To assess the role of helix A on activity, both steady state and single turnover kinetics for the helix A triple mutant, D47A-D50A-E54A, were measured. The k_{cat} for the triple mutant is about the same as WT LmP and, as with WT LmP, k_{cat} is independent of ionic strength and is not limited by formation of the LmP-LmCytc complex or dissociation of product. Instead, the rate limiting process remains intramolecular ET from the LmP active site Trp to $\text{Fe}^{\text{IV}}=\text{O}$. We next used stopped flow spectroscopy to measure the rate of LmCytc oxidation by LmP Compound I. As in our previous study (203), an excess of Compound I over LmCytc was used to ensure that only reduction of the Trp radical was measured. The rate of this process is about 50% slower in

the mutant (Table 4.3). These results are consistent with the BD and MD simulations and support the view that the helix A plays an important role in forming the initial encounter complex.

Brownian Dynamics Simulations Show That Yeast Cytc Does Not Associate with Helix A of Yeast CCP.

The combination of BD and MD simulations suggests that the charges on helix A play a vital role in the association of LmP and LmCytc. Three of the four charges found on helix A of LmP are not present in helix A of yeast CCP (Figure 4.13). This suggests that the heme peroxidases of the more complex organisms may have evolved to have a different mechanism of association, a notion that is supported by previous studies showing that the formation of the yeast CCP-Cytc complex is driven by nonpolar rather than electrostatic interactions (93, 207).

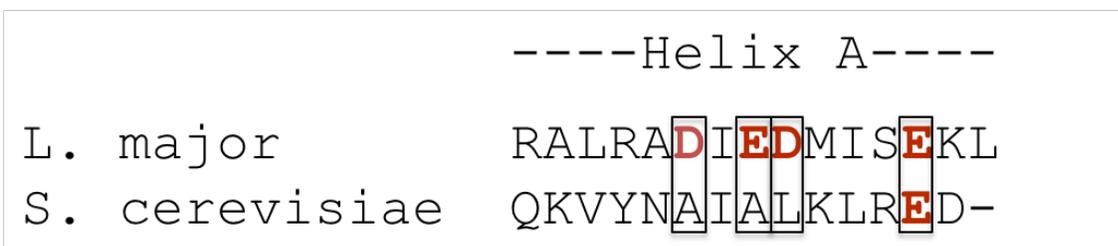


Figure 4.13. Comparison of the helix A sequence of LmP and CCP. Sequence alignment of helix A in the cytochrome c peroxidase LmP of *L. major* and that of the well studied and hydrophobically driven yeast CCP system from *S. cerevisiae*. Note that yeast CCP does not have three of the four charged residues on helix A that are found on helix A of LmP.

The lack of charges on helix A makes the yeast CCP system a useful control for further validation of the role of helix A charges in the formation of LmP-LmCytc encounter complexes. The probability isosurface obtained from BD trajectories of yeast Cytc around yeast CCP, initiated using the structures of the proteins in their ET complex, is depicted in Figure 4.14A.

For comparison, the location of the center-of-mass of Cytc in the structure of the complex is indicated in Figure 4.14B. There are several small features distributed broadly over the surface of CCP but, in contrast to the case of LmP (left panels of Figure 4.5A and 4.5B), there are no features near the N-terminal end of helix A of yeast CCP. Thus, it is apparent that the helix A charges are required for the formation of the initial encounter complex in which LmCytc is docked onto helix A of LmP.

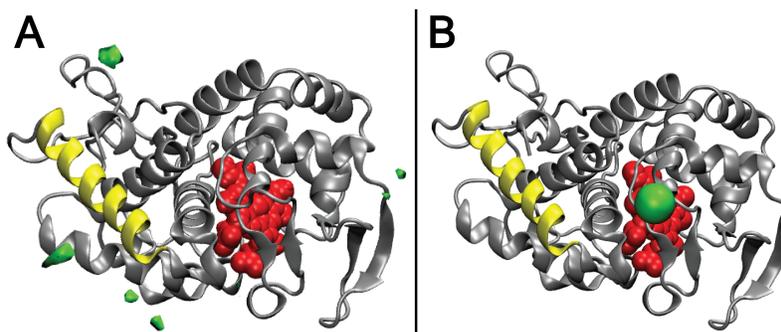


Figure 4.14. Structure of the CCP-Cytc complex from BD simulations. Structure of yeast CCP in the CCP-Cytc complex from BD simulations at 50 mM ionic strength, drawn and colored as in Figure 4.4. (A) The green surface is the probability isosurface for the yeast Cytc center-of-mass calculated from BD simulations initiated from the co-crystal structures contoured at a value of 2.2×10^{-3} . At this contour level, there are several features distributed over the surface of yeast CCP, but none indicating a strong interaction of Cytc with helix A of CCP. (B) The position of the center-of-mass of yeast Cytc in the structure of the complex (75) is depicted as a green sphere.

The second order association rates computed from our BD trajectories of the yeast CCP-Cytc system at 100 mM and 300 mM ionic strength, $1.65 \pm 0.02 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ and $0.414 \pm 0.012 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, respectively, are in good agreement with the corresponding experimental values, $2.95 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ and $0.22 \pm 0.01 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, respectively (211, 249). Our rates are lower than the values $6.71 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ and $0.856 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, respectively, which were computed previously by Gabdoulline and Wade using a similar BD approach, and also agree well with the experimental values (211). Moreover, at first glance, the probability isosurface shown in Figure 4.14A is dramatically different than an isosurface of the Boltzmann factor calculated based on

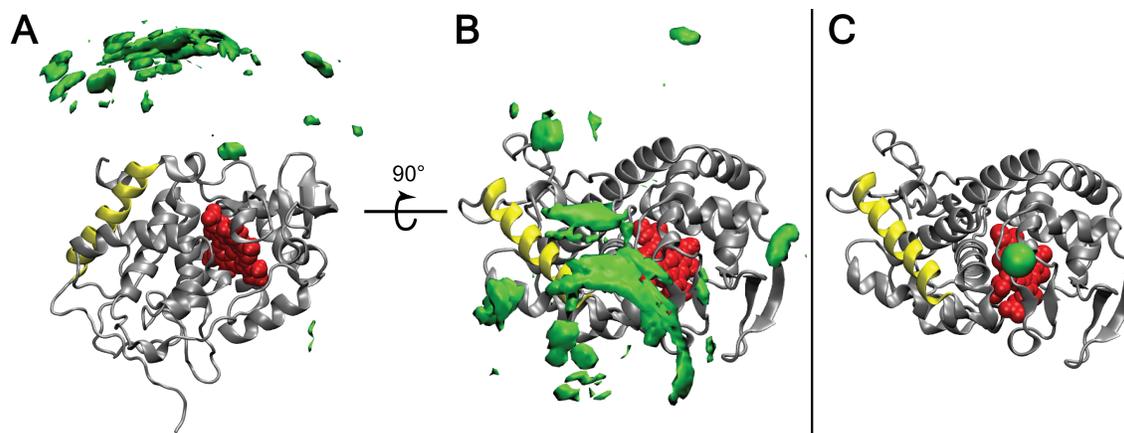


Figure 4.15. Visualization of the formed CCP-Cytc complexes in BD simulations at lower contour levels. Shown in (A) and (B) are two views of the structure of yeast CCP in the CCP-Cytc complex from BD simulations at 50 mM ionic strength, drawn and colored as in Figure 4.14. The orientation of yeast CCP in the same as in Figure 1 in (B), and is similar to Figure 5d of ref. (211) in (A). The green surface is the probability isosurface for the yeast Cytc center-of-mass calculated from BD simulations initiated from the co-crystal structures contoured at a value of 2.2×10^{-4} . Note that this contour level is ten times lower than that in Figure 5. (C) Hence, the features in the isosurface shown here are much broader and encompass much more volume than the isosurface depicted in Figure 5. The position of the center-of-mass of yeast Cytc in the structure of the complex is depicted as a green sphere. The orientation of yeast CCP in this panel is the same as in panel (B).

electrostatic interaction energies (orange surface in Figure 5d of ref. (211)). According to the latter, Cytc has a broad distribution of configurations on the side of yeast CCP that contains the active site, which is clearly not evident in the probability isosurface shown in our Figure 4.14A. However, when we lower the value at which we contour the probability isosurface by a factor of 10 (i.e., to 2.2×10^{-4}), we also observe a broad distribution of Cytc around the active site surface of yeast CCP (Figure 4.15). The differences between our rates and those computed by Gabdoulline and Wade presumably reflect differences in the details of the BD simulation protocols used in the two studies.

Conclusions

This study is an example of how computational approaches can provide novel insights into the mechanism of an important protein-protein interaction that were not at all obvious from previous experimental work, including crystal structures. Indeed, the computational effort guided the experimental work reported herein and, together, the simulations and experiments revealed new aspects of the dynamics of the initial encounter complex formed by LmP and LmCytc. The calculated rates for both wild type and mutant LmP agree very well with experimental rates, thereby validating the BD approach. The most important new insight into the LmP-LmCytc system, predicted by the simulations and confirmed by the experiments, is the unexpected role of helix A of LmP. The dynamic picture that emerges from these studies is that LmCytc rapidly forms an initial encounter complex with helix A of LmP and then “crawls” toward the ET active site. The encounter complex is transient, with no specific intermolecular ion pairing, but rather an interaction between two large and electrostatically complementary surfaces that ensures very rapid complex formation. However, the final ET active complex does require specific interactions such as the LmP:D211-LmCytc:R24 ion pair. The probability of this ion pair forming upon initial complex formation is low, which implies that the system is designed to rapidly form a nonspecific interaction with the A helix that reduces the search for the ET active complex from a relatively slow three-dimensional search to a more rapid two-dimensional search. This ensures that complex formation and dissociation are not rate limiting, which is exactly what we observe experimentally. The interplay between computational and experimental approaches used here should prove useful in applications to many other systems where complementary electrostatic surfaces play a major role in complex formation.

Reproduced with permission from Hollingsworth S. A., Holden J.K., Li H., Poulos T.L. (2016) Elucidating nitric oxide synthase domain interactions by molecular dynamics. *Protein Science*. **25**: 374-382. Copyright 2016 The Protein Society.

Chapter 5

Elucidating Nitric Oxide Synthase Domain Interactions by Molecular Dynamics

*Published as co-first author with Jeffrey K. Holden. Author Contributions: SAH and JKH set up the initial model as well as set up the MD simulation and validated the results against previous experimental data. SAH ran and analyzed the resulting simulation. SAH, JKH, and TLP wrote and edited the manuscript.

Summary

Nitric oxide synthase (NOS) is a multidomain enzyme that catalyzes the production of nitric oxide (NO) by oxidizing L-Arg to NO and L-citrulline. NO production requires multiple interdomain electron transfer steps between the FMN and heme domain. Specifically, NADPH derived electrons are transferred to the heme containing oxygenase domain *via* the FAD and FMN containing reductase domains. While crystal structures are available for both the reductase and oxygenase domains of NOS, to date there is no atomic level structural information on domain interactions required for the final FMN-to-heme electron transfer step. Here, we evaluate a model of this final electron transfer step for the heme-FMN-CaM NOS complex based on recent biophysical studies using a 105 nanosecond molecular dynamics trajectory. The resulting equilibrated complex structure is very stable and provides a detailed prediction of interdomain contacts required for stabilizing the NOS output state. The resulting equilibrated complex model agrees well with previous experimental work and provides a detailed working model of the final NOS electron transfer step required for NO biosynthesis.

Introduction

Nitric oxide (NO) is an important signaling molecule required for fundamental physiological processes in the neuronal, immune, and cardiovascular systems (1, 250, 251). Disruptions to NO signaling have been linked to a wide variety of cardiovascular (251) and neurodegenerative disorders (252-255). NO biosynthesis requires the heme containing enzyme nitric oxide synthase (NOS), substrate L-Arg, cofactor 6R-5,6,7,8-tetrahydrobiopterin (H₄B), molecular oxygen, and electrons derived from NADPH (1, 256). At the NOS heme active site, L-Arg is oxidized to L-citrulline and NO. Electrons from NADPH are transferred to the oxygenase domain active site *via* the FAD and FMN containing reductase domain of the opposing monomer in a calmodulin (CaM) dependent mechanism (Figure 5.1) (1). CaM facilitates interdomain electron transfer by binding to an α helical linker between the heme and FMN subdomain (131). It is generally thought that the required NOS conformational changes induced by CaM binding re-orient the FMN subdomain from an electron accepting conformation (input state) to an electron donating conformation (output state) as illustrated in Figure 5.1 (132, 133). For both endothelial NOS (eNOS) and neuronal NOS (nNOS), the CaM induced conformational change between input and output states is dependent on the concentration of Ca²⁺. In sharp contrast, inducible NOS (iNOS) binds CaM very tightly and retains the ability to transition between the input and output states at basal Ca²⁺ concentrations (139).

The calmodulin dependent interdomain electron transfer between the NOS reductase and oxygenase domains represents a key step in NO production. CaM is thought to function by first destabilizing the FMN-FAD interaction (257). Destabilization of the FMN-FAD subdomains interaction allows for the reductase domain to split and the FMN subdomain to approach the oxygenase domain in the output state. Electron transfer from the FMN subdomain to the

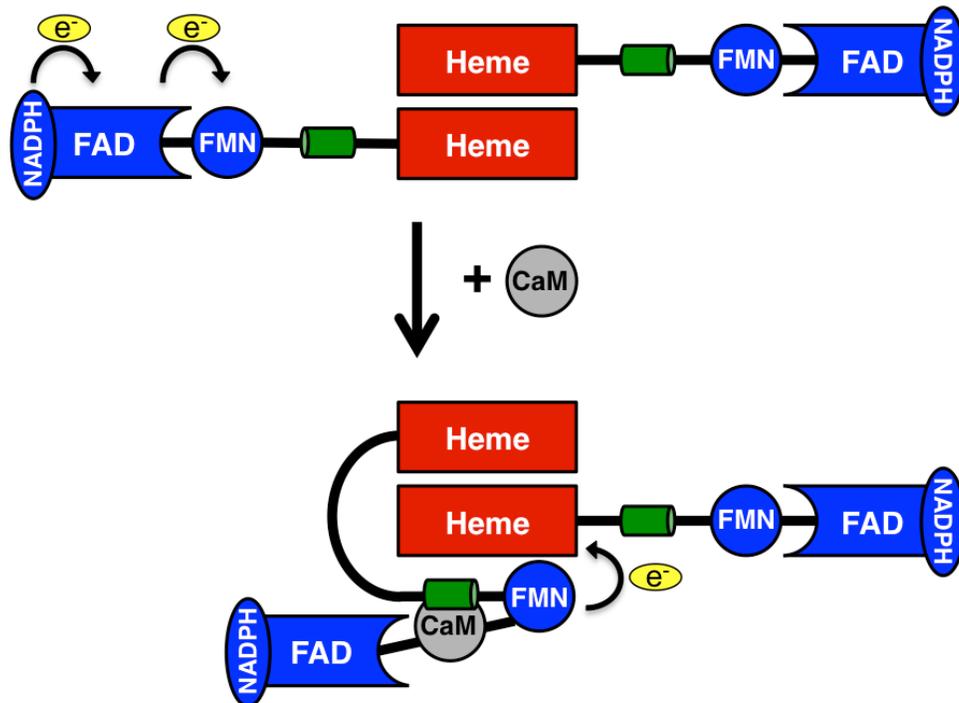


Figure 5.1. Nitric oxide synthase architecture. Each monomer of nitric oxide synthase consists of the oxygenase or heme domain (red) and the reductase domain (blue) which is composed of FMN and FAD containing subdomains. The oxygenase domain forms the heme active site and is the site of dimerization for holo-NOS. In the input state (top), NADPH binds and reduces FAD, which in turn reduces FMN. Calmodulin (CaM, grey) binding on the α -helical linker between the FMN and oxygenase domain (shown in green) is believed to induce a conformational switch towards the output state (bottom) to transfer electrons from FMN to the heme group.

oxygenase domain is gated by the required FMN conformational change from the input to the output state (130). Regulation of this step is critical because if NOS electron transfer is uncoupled, cell damaging peroxy species can form (107, 137, 138). Details on the mechanism of these CaM induced changes resulting in the output state remain unknown.

While crystal structures of individual NOS domains have been solved (140-146), a full length NOS crystal structure has remained elusive. Recent spectroscopic (258, 259), hydrogen-deuterium (HD) exchange (147), and single particle electron microscopy studies (129, 148, 149) have for the first time begun to provide the structural and biophysical details of the interdomain electron transfer step in NOS and the required conformational changes of the NOS FMN domain.

One key piece of evidence was the quantification of the FMN and heme group distance required for interdomain electron transfer at 18.8 Å using pulsed EPR (258).

Molecular dynamics simulations that are consistent with the available data have proven to be a powerful approach for developing detailed atomic-level complexes in the absence of experimental structures and probing important conformational changes (260, 261). In particular, such methods have been used to study similar reversibly binding electron transfer proteins (262, 263). The stability of such complexes, based on convergence of the model RMSD, over a reasonably long trajectory can provide confidence that the model is reasonably close to a functionally relevant structure. These simulations also are useful to cross-validate previous experimental data relative to the model being simulated. Therefore, we developed an output state model of the human iNOS oxy-FMN-CaM complex that is based on domain interactions resolved using HD exchange (147). The output state model remains stable over a 105 nanosecond molecular dynamics simulation and agrees well with previous experimental data and provides new insights into the stability of this iNOS conformational state. Our model provides a structural framework for further experimental testing.

Results and Discussion

Molecular dynamics and architecture of iNOS oxy-FMN-CaM complex.

Surface residues on the human iNOS oxygenase domain (defined in our model as residues 83 to 502) and CaM that were previously identified as complimentary binding sites by HD exchange (147) were used to place CaM within our initial model. Similarly, the FMN subdomain (residues 536 to 699) was placed within the initial model based on previous HD exchange data (147). Linkers connecting the heme and FMN domains based on the iNOS

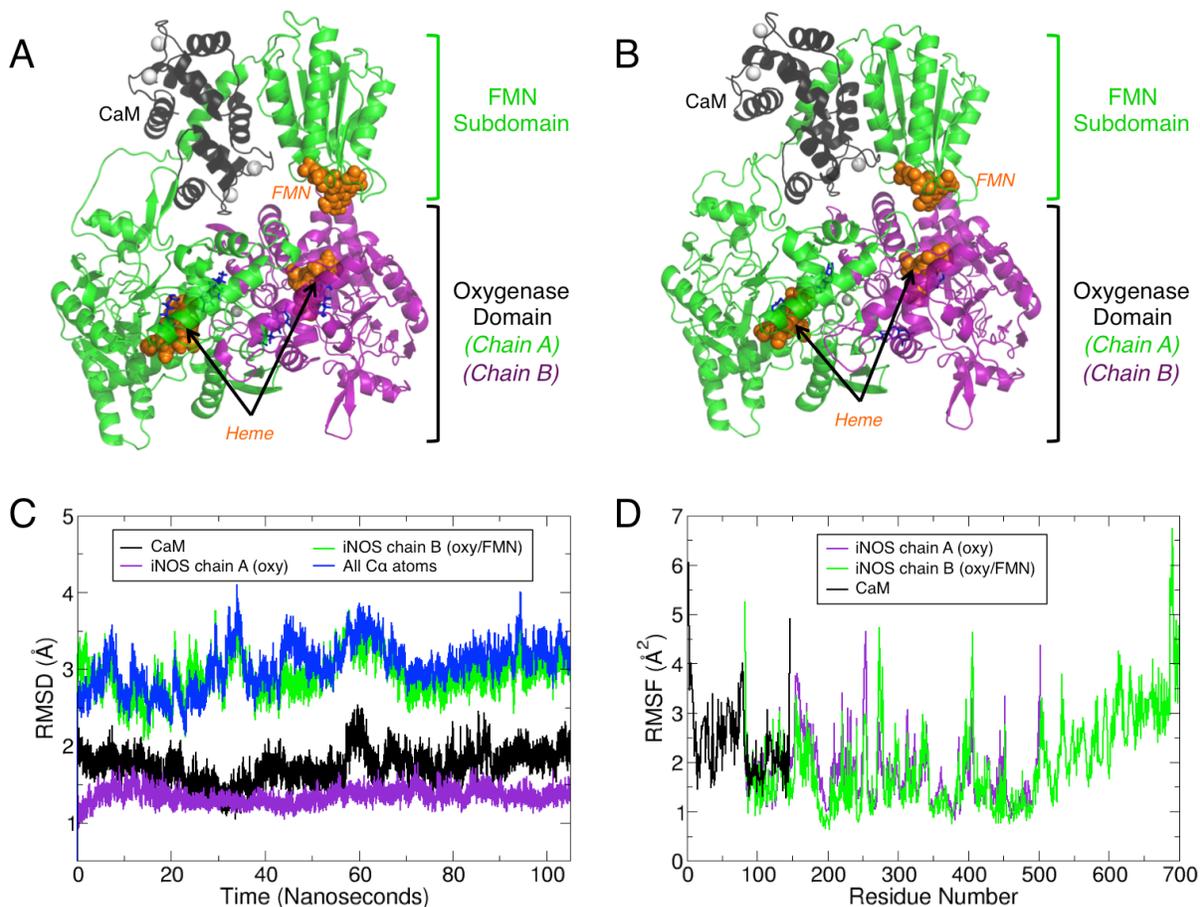


Figure 5.2. MD simulations of the oxy-FMN-CaM reveal complex stability. A) The initial system for the oxy-FMN-CaM system, as described in the results and methods, is displayed and B) a snapshot taken at 80 nanoseconds which represents an equilibrated complex structure. For both A) and B), chain A of the oxygenase domain is colored purple, chain B which includes an oxygenase domain, CaM binding helix and FMN subdomain is colored green and CaM is shown in black. The heme and FMN cofactors are displayed as orange sticks, the substrate L-Arg and H₄B in blue sticks, and Ca²⁺ and Zn²⁺ as white spheres. The equilibrated CaM and FMN subdomain model is closer to chain A of the heme domain than in the initial starting model. C) Root mean squared distance (RMSD) analysis over the full 105 nanosecond trajectory shows that the structure remains relatively unchanged over the course of the trajectory following initial minimization and initiation of the simulation for each major chain of the oxy-FMN-CaM. D) Root mean square fluctuations (RMSF) measured for the sidechains for CaM (black), chain A heme domain (purple) and chain B heme/FMN domain (green).

sequence were added to the initial HD-derived model and were modeled using ϕ, ψ peptide constraints to complete the oxy-FMN model (residues 503 to 535). The constructed model (Figure 5.2A) then was used as the starting point for the 105 nanosecond simulation.

The system underwent a quick conformational shift following the initial minimization and remained relatively unchanged for the remainder of the simulation as noted by RMSD analysis (Figure 5.2C). For the initial 40 nanoseconds of the simulation, the FMN domain did not consistently interact with the heme domain. Visualization of the trajectory revealed that the FMN domain was constantly sampling different conformations at the FMN-heme domain interface as if it were undergoing a “bind and crawl” mechanism of association. Analysis of the FMN and heme cofactor positions indicates the distance between the heme iron and the N5 atom of FMN to be greater than the previously determined 18.8 Å distance during this initial phase of the trajectory (258) (Figure 5.3). However, after approximately 40 nanoseconds the FMN domain moved into a position where it formed close contacts with the heme domain that remained stable for the remainder of the simulation. This equilibrated output model coincides with the formation of several non-covalent contacts that likely function to facilitate formation of the output state (Table 5.1 and Figure 5.4). Moreover, the FMN to heme distance, as measured between the heme iron and the N5 atom of FMN, equilibrated to 18.0 ± 0.6 Å which agrees well with the previously EPR-derived distance of 18.8 Å (258).

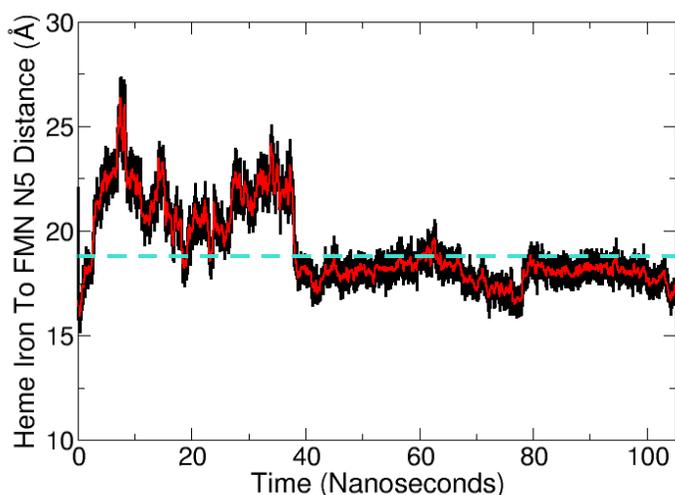


Figure 5.3. Distance measured between heme and FMN cofactors. The distance between the oxygenase heme iron and the N5 atom of the FMN group was calculated for the duration of the simulation. The experimentally determined heme to FMN distance of ~ 18.8 Å (258), is represented as a dashed turquoise line. A running average with a window of 100 picoseconds is traced in red on the full data set in black.

Table 5.1. Major domain-domain interactions in the equilibrated oxy-FMN-CaM complex. Contacts are defined as a distance $\leq 3.5\text{\AA}$.

Residue A	Residue B	Average contacts per picosecond	Selected Atom-Atom Pair	MD Distance and St. Dev. Over Final 50 ns [†]	Previous Experimental Work*
<i>Chain A heme domain to Chain B FMN domain</i>					
HemeA-Arg452	FMN-Glu546	4.331	A-NH2 / B-OE1	$3.11 \pm 0.51\text{\AA}$	(264, 265)
HemeA-Asn449	FMN-Glu546	1.891	A-ND2 / B-OE2	$4.13 \pm 1.10\text{\AA}$	(264, 265)
HemeA-Lys155	FMN-Glu603	1.626	A-NZ / B-OE1	$4.79 \pm 2.14\text{\AA}$	(266)
HemeA-Glu156	FMN-Lys607	1.406	A-OE1 / B-NZ	$5.44 \pm 2.74\text{\AA}$	Reported here
HemeA-Ser207	FMN-Asn595	1.219	A-O / B-ND2	$3.79 \pm 0.94\text{\AA}$	Reported here
HemeA-Asn208	FMN-Gly594	1.063	A-ND2 / B-O	$3.58 \pm 1.15\text{\AA}$	(267)
<i>Chain A of heme domain to Calmodulin</i>					
HemeA-Arg86	CaM-Asp122	5.224	A-NH1 / C-OD2	$2.78 \pm 0.27\text{\AA}$	(147)
HemeA-Arg83	CaM-Glu127	4.357	A-NH2 / C-OE1	$3.71 \pm 1.88\text{\AA}$	(268)
HemeA-Gln97	CaM-Asp133	1.291	A-NE2 / C-O	$3.27 \pm 0.63\text{\AA}$	Reported here
<i>Chain B FMN domain to Calmodulin</i>					
FMN-Arg536	CaM-Glu47	5.076	B-NH2 / C-OE1	$2.99 \pm 0.34\text{\AA}$	(146)
FMN-Arg576	CaM-Glu83	4.568	B-NH1 / C-OE2	$3.23 \pm 0.66\text{\AA}$	Reported here
FMN-Glu551	CaM-Lys94	2.805	B-OE1 / C-NZ	$3.55 \pm 1.22\text{\AA}$	Reported here
FMN-Lys568	CaM-Glu87	2.443	B-NZ / C-OD2	$3.32 \pm 0.92\text{\AA}$	(146)
FMN-Lys574	CaM-Asp93	1.854	B-NZ / C-OD2	$2.67 \pm 0.09\text{\AA}$	Reported here

* - Residue previously identified to be important for either enzyme activity or domain interactions as referenced in corresponding study in either iNOS or an analogous residue in either eNOS or nNOS.

[†] - For a more in depth view of the nature and dynamics of each contact, refer to Figure 5.4.

Based on previous NOS kinetic analyses, the two electron reduced hydroquinone, FMNH₂, transfers electrons to the heme group of the opposing subunit *via* a conserved tryptophan (147, 267, 269). In our equilibrated MD model, the center of mass of this HemeA_{W372} residue (HemeA_{W366} in murine iNOS) sidechain is within $11.67 \pm 0.7\text{\AA}$ of the N5 atom of the

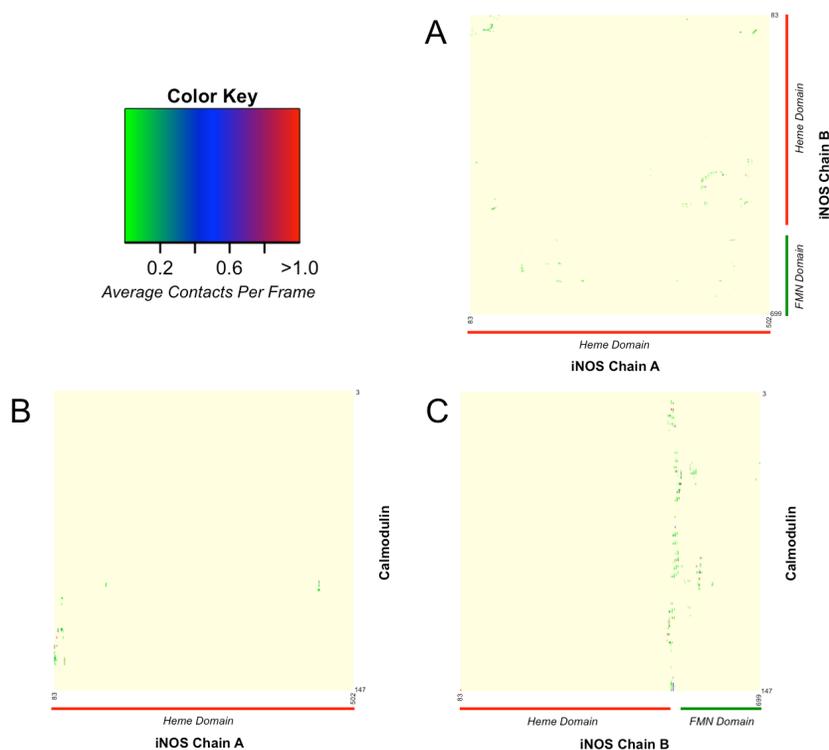


Figure 5.4. Heat maps highlighting domain-domain interactions in full-length iNOS.

Contact analysis that tracks the average of number of atom-to-atom contacts per frame over the final 50 nanoseconds of the bidomain trajectory presented for a given combination of two polypeptides; A) iNOS chain A (the heme domain only) and chain B (heme and FMN domains), B) chain A and calmodulin and C) chain B and calmodulin. The number of contacts per frame are colored as shown in the spectrum in the top left, where residue pairs that have more than one contacts per frame are shown in dark red, which are highlighted in Figure 5.6.

FMN group and $9.46 \pm 0.2 \text{ \AA}$ of the heme iron. The distance between HemeA_{W372} and the redox partners is important, as recent work has implicated this conserved Trp residue to mediate electron transfer between the two redox cofactors and the observed distances between the cofactors and heme are similar to those predicted previously (147).

In order to further test our model we also docked our equilibrated model into a previously reported single particle NOS EM map (EMD 2748) at 74 Å resolution (129). Our model had a cross correlation value of 0.6824 using the “fit in map” function of Chimera with only 93 of 9,716 atoms falling outside the EM density at a map value of 1.287 (Figure 5.5).

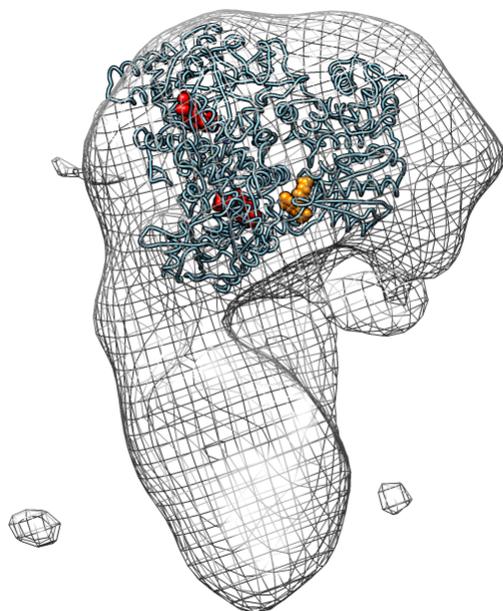


Figure 5.5. Fit of the equilibrated bidomain model into published EM density map of the output state. Shown is the equilibrated bidomain (oxy-FMN-CaM iNOS) model fit into an EM map of holo-iNOS (which includes an additional FMN domain and both FAD domains not included in our bidomain structure) in the output state of approximately 67Å (EMD-2748). The fit was done using the “fit in map” function of Chimera, which resulted in a cross correlation coefficient of 0.6824 at an average map contour value of 1.287, for which only 93 atoms of the 9,716 atoms of the equilibrated model were outside the displayed contour.

FMN and heme domain interface of the oxy-FMN-CaM complex.

In the final equilibrated structure, several contacts were observed to facilitate binding of the FMN domain within close proximity to the heme domain (Figures 5.6, Table 5.1 and Appendix B). The observed domain contacts within our model are largely facilitated by side chain electrostatic interactions. Based on the average number of contacts per frame (where a contact is an atom-to-atom distance of ≤ 3.5 Å), the strongest of these interactions is an ion pair interaction between residues FMN_{E546} and HemeA_{R452}. The FMN_{E546}-HemeA_{R452} ion pair interaction was observed to form upon complex formation and remained stable for the remainder of the simulation. Previous mutagenesis analysis on the FMN_{E546N} in iNOS shows that the FMN-to-heme electron transfer rate decreases by about 2.5-fold owing primarily to a 3-fold increase in the activation entropy (264, 265). In addition to FMN_{E546} contacting HemeA_{R452}, FMN_{E546} contacts HemeA_{N449}. Analysis of the FMN_{E546}-HemeA_{N449} interaction indicates the HemeA_{N449} sidechain nitrogen to switch between contacting the FMN_{E546} carboxyl group and the carbonyl oxygen of FMN_{E456}.

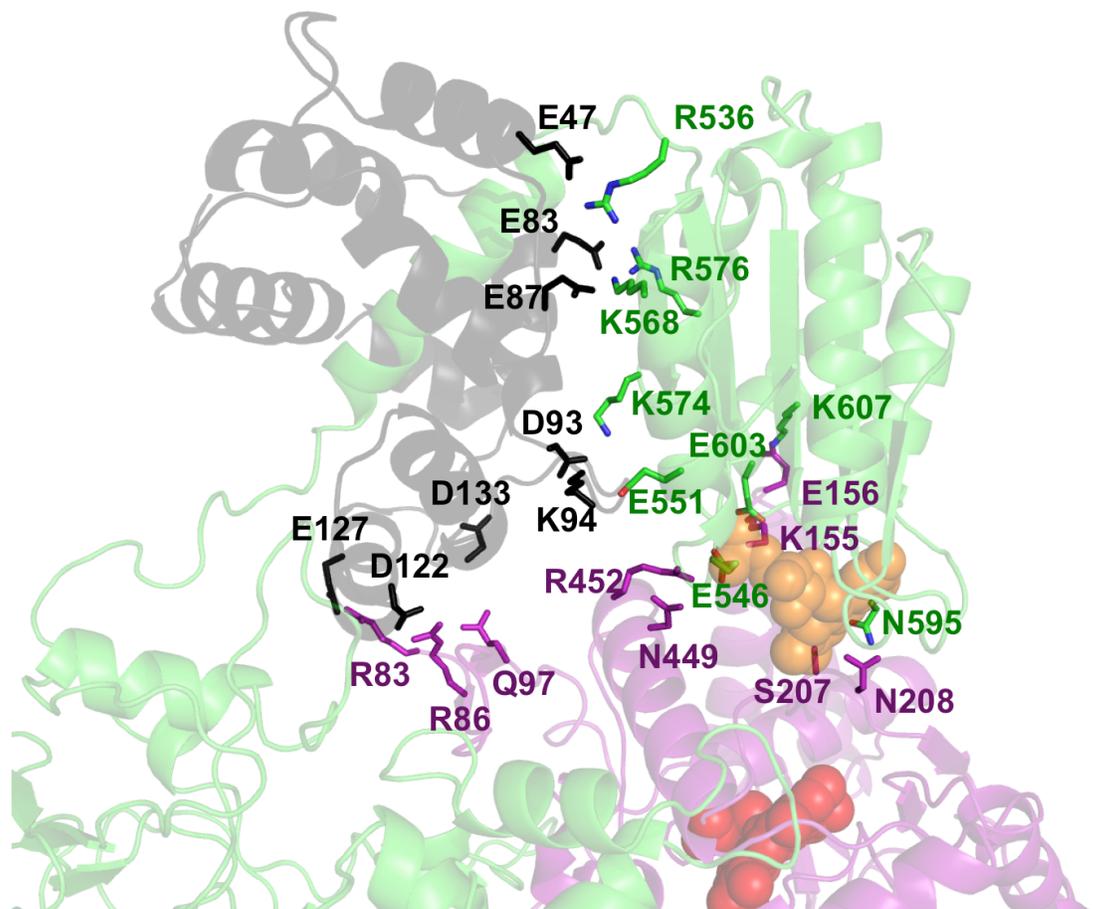


Figure 5.6. Interfacial contacts detected between calmodulin, oxygenase domain, and the FMN subdomain of iNOS. View of the NOS domain interface composed of the heme domain, calmodulin, and FMN subdomain (colored in the same fashion as Figure 5.2) in an equilibrated MD structure of the oxy-FMN-CaM complex, taking a snapshot from roughly 80 nanoseconds into the simulation. Residues listed in Table 1 are shown as sticks and labeled to illustrate the interactions that facilitate formation of the output state. The transient contacts observed throughout the simulation that averaged less than one contact per frame are not shown.

Our model also indicates the heme-FMN interface of human iNOS to be further stabilized by electrostatic interactions between FMN_{K607} -HemeA $_{\text{E156}}$, FMN_{E603} -HemeA $_{\text{K155}}$ and FMN_{N595} -HemeA $_{\text{S207}}$. Human iNOS FMN_{E603} is analogous to rat nNOS FMN_{E819} that previously has been shown to facilitate charge-pairing interactions within the FAD domain, which has been predicted to use the same FMN domain surface region to interact with the oxygenase domain (129, 266). Between our model and previous biochemical studies, FMN_{E603} (iNOS) / FMN_{E819} (rat nNOS) may also play a critical role in regulating electron transfer between the input and output states.

In our equilibrated model we also observed a weak hydrogen-bond interaction between the backbone oxygen of FMN_{G594} and the sidechain of HemeA_{N208}. These residues form a stable interaction between the two domains with an atom-to-atom distance of 3.6 ± 1.1 Å. To the best of our knowledge, these residues have not been experimentally validated by mutagenesis studies. However, both FMN_{G594} and HemeA_{N208} were broadly implicated to contribute towards domain-domain interactions based on HD exchange studies (147). A series of secondary or less frequent and longer-range interactions at the interface of our output state model were also identified and displayed in heatmap form (Figure 5.4). These secondary interactions may supplement the stronger contacts that facilitate binding of the FMN subdomain to the heme domain. When combined with the major contacts described above, the FMN-oxygenase interface averages 16.34 atom-to-atom contacts per frame over the final 50 nanoseconds of the simulation.

Possible role of CaM in the stabilization of the oxy-FMN-CaM complex.

Our model is consistent with the proposed role of CaM in stabilizing the output state by directly interacting with the heme domain (129, 147). In our output state model, CaM binds to the FMN-heme linker (Table S1) to facilitate formation of stable contacts between the oxygenase domain and FMN subdomain and is consistent with HD exchange data (147), as CaM was found to also interact directly with the heme domain (Figures 5.6, Appendix B and Table 5.1). As a result, CaM works together with the FMN domain to form a complex with the heme domain. Analysis of the final 50 nanoseconds of the simulation demonstrates five residues of CaM to form noncovalent interactions with the FMN subdomain and three residues with the heme domain of iNOS. Of the CaM-FMN subdomain interactions observed in our model, ion pairs between charged residues dominate the noncovalent interactions, as detailed in Table 5.1. One

ion pair in particular, CaM_{Asp122}-HemeA_{Arg86}, is of interest as mutations at CaM_{Asp122} have previously been shown to lower NADPH oxidation rates, increase cytochrome-c activity, and decrease NO production (147). These results provide additional support that CaM_{Asp122} plays an important role in facilitating interdomain electron transfer. Moreover, our data implicates HemeA_{Arg86} of human iNOS as being required to stabilize CaM binding in the output state. We also observed several additional transient interactions that play a minor role in facilitating the CaM-FMN subdomain interaction, as demonstrated by the contact heatmaps (Figure 5.4). When combined with the major interactions described in Table 1, the CaM-FMN subdomain interaction averages 21.23 atom-atom contacts per frame over the final 50 equilibrated nanoseconds of the trajectory, more than the contacts observed between the FMN and oxygenase domains. These data suggest that the CaM-FMN subdomain interactions are mediated by several noncovalent interactions that may play a more integral role in stabilizing the output state than previously predicted.

Following equilibration of the output state, NOS domain interactions were further stabilized by three noncovalent interactions between both CaM and chain A of the heme domain (Figure 5.6 and Table 5.1). Complimentary charged and polar residues dominate the contact interactions at this interface, similar to the aforementioned CaM-FMN interface. The entire CaM-oxygenase domain interface interaction averages 13.79 atom-to-atom contacts per frame over the final 50 nanoseconds of the trajectory, making it the weakest of the domain-domain interactions observed in our MD simulation of the output state. One key difference between both our equilibrated model and the iNOS CaM-FMN crystal structure (PDB 3HR4) are the contacts made between the FMN subdomain and CaM that make up this final interface in the oxy-FMN-CaM complex. The resulting equilibrated complex structure conserves two contacts that include

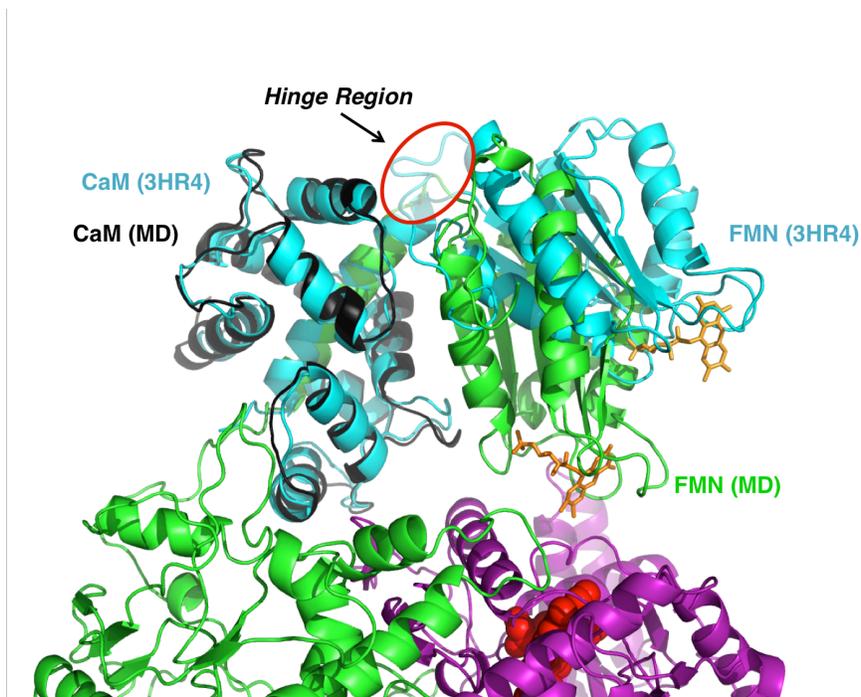


Figure 5.7. Overlay of FMN-Calmodulin crystal structure and proposed model structure.

The co-crystal structure of iNOS FMN domain and calmodulin is overlaid onto the calmodulin of the equilibrated bidomain model. The bidomain is colored as in Figures 2 and 4 (chain A heme domain in purple, chain B heme and FMN domains in green and calmodulin in black), while the co-crystal structure is shown in cyan. The hinge region between the FMN domain and calmodulin is highlighted by both labeling and a red oval showing the flexible linker that accounts for the change in conformation from co-crystal to equilibrated model possible.

residues that were described by the previously described crystal structure (146), FMN_{R536} - CaM_{E47} and FMN_{K568} - CaM_{87} , and the addition of three novel contacts observed in our model are the result of a conformational change within the linker between the CaM-binding helix of chain B and the FMN domain (Figure 5.7). In short, our output state model requires more residue contacts between the FMN subdomain and CaM than previously observed by x-ray crystallography. This is not too surprising since our model includes the heme domain together with information of recent HD and EM data (129, 147) while the crystal structure includes only CaM and the FMN domain.

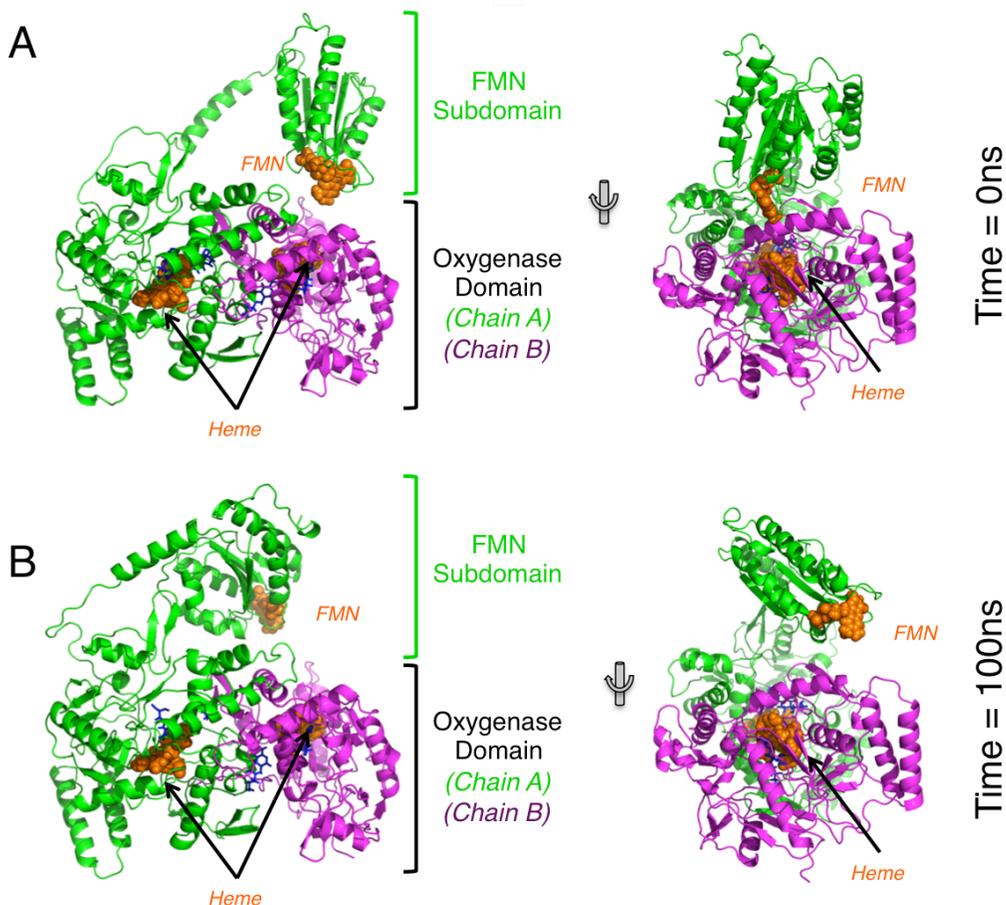


Figure 5.8. Overview of no-CaM simulation of the iNOS oxy-FMN complex. A) The same initial starting model, shown here in two different orientations, was used in the no-CaM simulation as was used in the full oxy-FMN-CaM simulation. B) The resulting equilibrated no-CaM complex following 100ns of simulation. Over the course of 100ns, the FMN domain moved away from the heme domain and no tight interdomain complex was formed. All domains are colored identically as displayed in Figure 2.

In order to further probe the importance of CaM binding for the stability of the oxy-FMN domain complex, a second simulation was carried out with the same starting model where calmodulin was removed (Figure 5.8A). It would be expected that if CaM does play a role in stability of the overall complex, that in the absence of CaM no stable complex would be formed between the FMN and heme domain. Indeed, after 15 ns, the complex undergoes a large conformation change as observed by RMSD analysis (Figure 5.9A) corresponding to the FMN domain moving away from the heme domain dimer. After 20 ns, all significant contacts between

the FMN and heme domains were lost and the complex was broken for the remainder of the 100ns simulation (Figures 5.8B and 5.9A). The FMN-to-Heme cofactor distance was recalculated as was done in Figure 3 and shows that the distance remained much larger than the experimentally observed value of 18.8 Å (Figure 5.9B). Taken together with the results of the initial simulation, this underscores the important role CaM plays in stabilizing the interactions between the FMN and heme domains during the second electron transfer step.

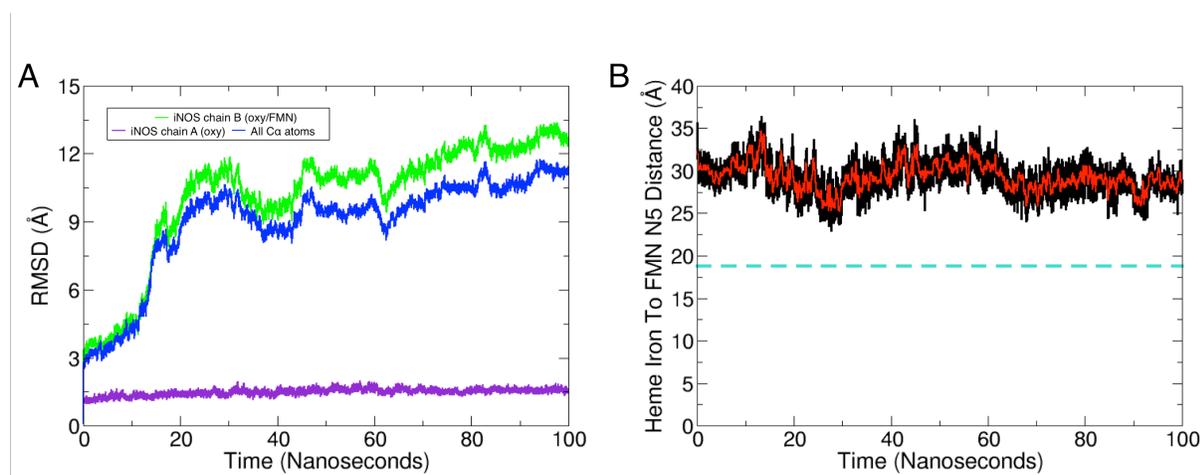


Figure 5.9. Complex stability of the no-CaM complex simulation. A) Root mean squared deviation (RMSD) analysis over the full 100ns simulation of the no-CaM simulation reveals that the complex undergoes significantly larger conformational changes than the oxy-FMN-CaM complex (Figure 2) corresponding to large changes in chain B of the heme-FMN chain of the iNOS complex as displayed in Figure 5.8. B) The distance between the N5 atom of the FMN and heme iron in the no-CaM system is tracked identically as it was done for the oxy-FMN-CaM complex in Figure 5.3, revealing no agreement with the experimentally observed distance (dashed turquoise line).

Conclusion

Through the implementation of modeling and molecular dynamics, we have produced an atomic level model of the iNOS output state. Following equilibration, our output state model is in excellent agreement with previous biochemical, biophysical, and site directed mutagenesis studies. MD has also identified additional surface contacts that may play an important role in the final iNOS interdomain electron transfer step. One of the more important observations in this

study is the role of CaM in activating iNOS catalysis. As first proposed by Smith et al (147), CaM plays an active role by directly interacting with the heme domain. Thus CaM may work synergistically with the FMN domain of iNOS to form an ET competent output state complex. However, the domain arrangement of the input state must be substantially different than our output state model. If we assume that the nNOS reductase domain structure, for which the FAD and FMN are within 3.4 Å (146), is a true representation of the input state then the CaM linker segment and heme domain must be positioned quite differently in the output state. Without such motion there would be severe overlap between CaM and the FAD domain. The simplest dynamic model that emerges from these observations is of an equilibrium mix between the input state where the FMN and FAD are in close contact and the output state where CaM and the FMN domain interact with the heme domain. The function of CaM is thus not so much to destabilize the FMN-FAD interaction but rather to stabilize the ET competent output state long enough for heme reduction. Neither input nor output states must be very stable or long-lived owing to the large domain rearrangements that are necessary for ET first from FAD to FMN and then FMN to heme. As a result, NOS is conformationally heterogeneous as evidenced by the recent cryoEM studies (129, 148, 149, 270). With this new information, it may now be possible to probe the NOS output state through additional biochemical and biophysical studies in order to better understand the dynamic equilibrium between the input and output states of NOS.

Methods

Construction of the initial oxy-FMN-CaM system.

Atomic coordinates for iNOS oxy-FMN-CaM were constructed from PDB 3HR4 and 1NSI. The positioning of CaM bound to the CaM binding peptide (PDB 3HR4) relative to the

iNOS oxygenase domain (PDB 1NSI) was based on previously reported HD exchange data (141, 146, 147). Similarly, FMN subdomain (coordinates from PDB 3HR4) was then positioned relative to the oxygenase domain and CaM sequence. Domains of iNOS were positioned using Pymol. Peptide linkers determined from the human iNOS sequence were then modeled and built using COOT (271).

Molecular dynamics of the oxy-FMN-CaM system.

Hydrogens were added to the system using psfgen of VMD 1.9.1 (272). The system was then hydrated with TIP3P water with a 15 Å cushion for a total of 40,660 additional waters residues. Crystallographic waters from PDB 1NSI were also included within the initial start model. The final system had a unit cell box dimensions of (128Å x 105Å x 117Å) and 141,497 atoms. Initial simulations were carried out on Greenplanet (gplogin1.ps.uci.edu) before productions runs were carried out on the Stampede supercomputer (stampede.tacc.utexas.edu) using NAMD 2.9 (273). The CHARMM force fields used to model the protein, heme, FMN and H₄B cofactors were identical to those used in previous studies (33, 118, 274). A smooth partial Ewald mesh method was used for the calculation of Columbic forces, a Langevin thermostat was used to maintain a constant temperature of 300K, and a Nose-Hoover-Langevin piston employed for constant pressure control (275, 276).

Positional restraints were placed on the central atoms of each heme group in order to reproduce the heme conformation that has been observed in crystal structures as well as an additional extra bond to preserve the hydrogen bonding of the H₄B cofactor to the heme domain as has been done in a previous study (118). Similarly, extra bonds were added to the simulation to preserve the experimentally observed coordination of the heme domain Zn²⁺ and the four Ca²⁺

atoms in CaM. No restraints were placed on the protein backbone of the iNOS model nor were any inter-domain restraints included in the simulation. Equilibration was carried out initially by minimizing the structure for 1000 fsec before the simulation was allowed to propagate for 5 nanoseconds at 1.0 fsec time step. After equilibration, the timestep was increased to 2.0 fsec for the remainder of the simulation. Positional restraints were placed on central atoms of the heme to ensure that the heme coordination was in agreement with previously solved crystal structures of the oxygenase domain. Frames were saved every 2 picoseconds and saved for analysis which was carried out on local systems using locally developed tools and VMD (272) and Chimera (248).

Chapter 6

Introduction to Self-Assembled Monolayers

Amongst some of the most basic phenomena in Nature than has eluded a deep scientific understanding to date, one of the areas that has seen the most progress concerns the behavior of water at interfaces. The structure and dynamics of water at either biological interfaces such as the surface of proteins (277-282), vegetation (283), atmospheric interfaces such as dust or airborne particles (284, 285), or at synthetic surfaces such as buildings or window surfaces (286-290) has been the focus of tremendous work over the past several decades (291). Understanding how water behaves not in bulk solvent but at surfaces is instrumental in a broad range of processes ranging from possible pollutant uptake (292, 293) to protein association. Particularly pertinent to the preceding chapters is how proteins associate to form complexes, such as cytochrome P450 with its protein redox partner to power hydroxylation reactions or peroxidases complexes that remove harmful reactive oxygen species from the environment. It has been well documented that as proteins interact the solvating shells of water surrounding them play vital roles in determining how complex association occurs (294) and that the behavior of these solvation shells is very different from that of bulk water (295-301). The alteration of water behavior due to the presence of these surfaces, which include a wide spectrum from hydrophobic to hydrophilic, can also have drastic effects on the kinetics and even mechanisms of reactions that occur near or at these interfaces (302, 303). For these reasons, it is very important to probe the changes in water behavior at such interfaces. However, given the high variability of such surfaces and the difficulty of preparation and study of such surfaces, a large degree of work has turned to a model system that is not only relatively simple to prepare and study using both experimental and

computational methods but has already been utilized in a variety of fields for a multitude of purposes; self-assembled monolayers.

Self-Assembled Monolayers

Self-assembled monolayers (or SAMs) are made up a series of simple organic molecules that are highly modifiable that undergo self-assembly into single monolayers at the surface of reactive metallic substrates (304-307) and have commonly been employed as model systems to study interfacial chemistry (293, 308-311). Due to the large control that one can have in designing the resulting surface, SAMs are utilized in a variety of biotechnological processes. Perhaps one of the most commonly utilized characteristics of SAMs in biology is as a protein adsorption resistant surface (305, 312, 313). Having such a surface of which the experimentalist can exert a large degree of control or prevent unwanted chemical reactions from occurring in the physical space being studied is obviously a desirable effect that has been utilized many times. SAM technology has also been used in the creation of metal-free quantum dots (314) which are used in a number of fields ranging from bio-labeling agents (315, 316) to solar cells (317), as well as in the immobilization of macromolecules to create reactive surfaces for catalysis and biosensors (318). Applications of SAMs can also be found in material sciences (319, 320), nanotechnology (312, 321, 322) and medical imaging (323, 324) to name a few, underscoring both the wide variety of uses of SAMs and their already widespread use in those fields despite a lack of understanding of the interactions between the SAM surface and the water they are exposed to.

Self-assembly in nature is the process in which molecules will self-organize into a supermolecular series of interlocking or repeating patterns that that result in a wide variety of complex systems (306). A single SAM molecule, as shown in Figure 6.1, includes a tail group

SAMs form spontaneously through the adsorption of an active surfactant (tail group in Figure 6.1) on a solid surface (306). The kinetics, mechanism of formation, and variations in the formation of a SAM surface has been well studied (325-327). A generalized description of this process is shown in Figure 6.2. Surfaces that have been used to generate SAMs include a variety of metals and silicon oxides crystals. Amongst the important findings derived from computational studies of model SAM systems was addressing the level of cross-linking that occurs in the final SAM. While it had initially been proposed that remaining hydroxyl groups could react near the SAM surface to cross link neighboring chain molecules, work by Stevens showed that in order to form dense, fully-saturated SAMs, such cross linking would not be possible due to steric interactions (328). Following work has further probed the interfacial chemistry of the SAM surface that occurs upon formation, demonstrating that even lab to lab variations in the production of each SAM may lead to different SAM structures due to lower

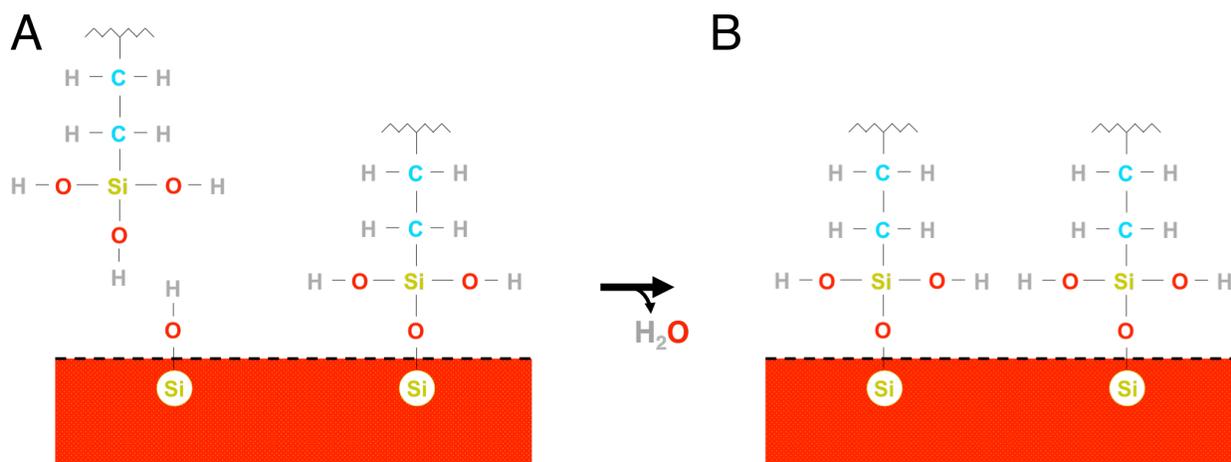


Figure 6.2. Formation of self-assembled monolayers. The generalized formation of self-assembled monolayers on a SiO₂ surface is displayed. For simplicity, only surface silicon atoms of the substrate are shown. A) An incoming organic chain will react with a hydroxyl group bound to a substrate silicon atom. B) Through the loss of a water molecule, the incoming chain covalently bonds to the substrate, expanding the substrate surface covered by the newly formed SAM.

densities of surfactant molecules which make such cross-linking possible (325). The exact degree of cross-linking that is occurring at the substrate-surfactant interface is still studied and remains an unresolved questions concerning SAM structure. However, one question that has been thoroughly studied is the behavior of water at the water-air interface of a variety of SAM surfaces.

There has been a great deal of computational and experimental work done to probe the behavior of water on SAM surfaces to better understand the behavior of water at such interfaces. This work can be divided into three broad categories; water behavior on hydrophobic SAMs, hydrophilic SAMs and rough or mixed SAMs. Given the obvious differences in water dynamics and structure at such vastly different interfaces, each of these three fields of study are highlighted below separately, both as a summary of the field as well as how the work in chapters seven through nine tie into unresolved questions of water behavior at such interfacial regions.

Water Behavior on Hydrophobic SAM Surfaces

The first computational study of a hydrophobic SAM was published by Hautman and Klein in 1989 who probed the structure of a fully reduced organic surface on a metallic surface (329). Along with their follow-up study in 1990 (330), the general structure of a SAM surface was first analyzed showing that while both the core and tail region of the SAM is well ordered, the head groups exposed to vacuum or water have a larger degree of variability (Figure 6.3) that can be altered by temperature which was further detailed by Mar and Klein (331). Further work revealed that the length of the chain itself, or to be more precise whether the number of carbons in the chain was even or odd resulted in different behaviors in a hydrophobic SAM (332). While there has been a great deal of work probing the structure of a dehydrated SAM surface both by

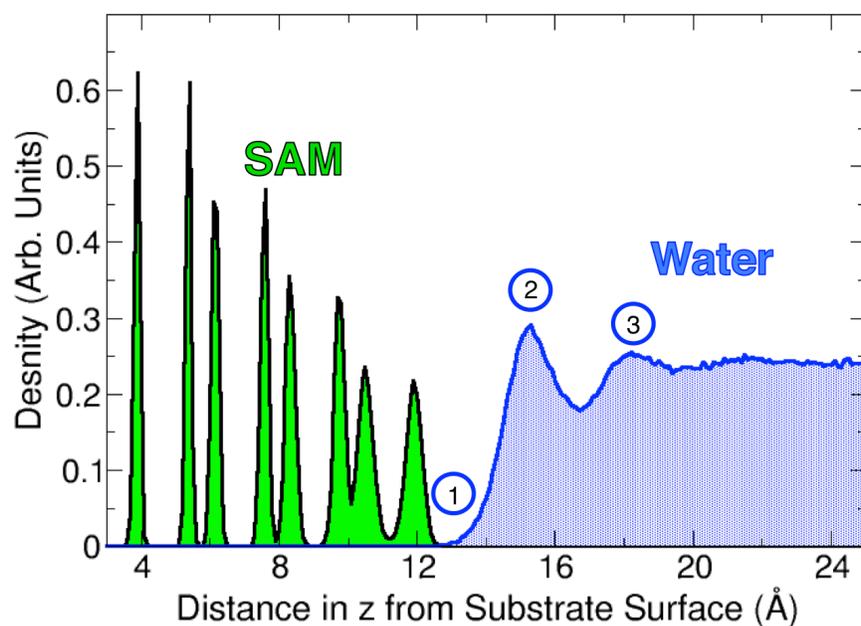


Figure 6.3. Water at a hydrophobic SAM surface. Density profile for a representative eight-carbon alkane siloxane SAM exposed to four monolayers of water. The SAM carbons are shown in green and water in blue. The SAM carbons nearer to the substrate show a higher degree of order than those exposed to solvent. A hydrophobic gap exists between the SAM surface and the droplet (Highlight 1), while there exists two structured layers of water at the droplet surface exposed to the SAM (Highlight 2, and 3).

itself (333) and contacting other SAM surfaces in frictional studies (334, 335), the structure and dynamics of water on hydrophobic SAM surfaces had not been the focus of study.

The dynamics and structure of water at a hydrophobic SAM surface was first probed in 1991 by Hautman and Klein (336). On a hydrophobic surface, exposed water molecules will form a droplet and attempt to minimize the surface area of the droplet that contacts the hydrophobic surface. Over the following decades, a multitude of work has been done to probe the structure of the water at such interfaces as experimental work has shown that even a hydrophobic surface is never truly absent of water (337). Amongst the most important findings of this work has included the presence of structured water at the surface interface that displays unique behaviors and spectroscopic characteristics as well as intermediate layers of water that

bridge the gap between the bulk water and the surface (338-341). Experimental work as confirmed the presence of a so-called “hydrophobic gap” at the surface of hydrophobic SAM surfaces, describing the hydrophobic layer of water exposed the unwelcoming surface (Figure 6.3) (342). Despite this intense study, a great many things concerning the structure and dynamics of water still remains unknown. Pertinent to the work in chapter seven, it is important to note that for all of the computational studies described, an idealized SAM surface with perfect coverage on the substrate was employed. However, such idealized surfaces may not reflect reality when such surfaces are prepared.

While fully formed and idealized single component hydrophobic SAM surfaces have been well studied, the presence of defects in these surfaces, and their affects on water behavior at such interfaces, has remained a mystery. While perfect SAM formation and coverage on a substrate can be possible, the likelihood that any such perfect SAM system is ever formed is incredibly small in even the best of circumstances. Defects in the SAM surface can be caused by a variety of issues including impurities of the substrate itself, imperfect coverage of the substrate by the organic surfactants or even during handling of the SAM surface or repeated use of a single substrate (343) (Figure 6.4). While there has been a single study focusing on the effects of large-scale defects in SAMs, like those caused by improper handling (344), no such studies of small-scale defects have been carried out. This lack of study is particularly puzzling given recent studies that show that water has been experimentally observed to penetrate the hydrophobic surface (345-347). Experimental work has shown that this penetration may even be dependent on chain length (348). In an attempt to probe the affect of small scale defects on water behavior and how it might address previous seemingly contrary experimental and computational findings, I

carried out a systematic study of such a SAM surface with a series of small defects which is presented in chapter seven.

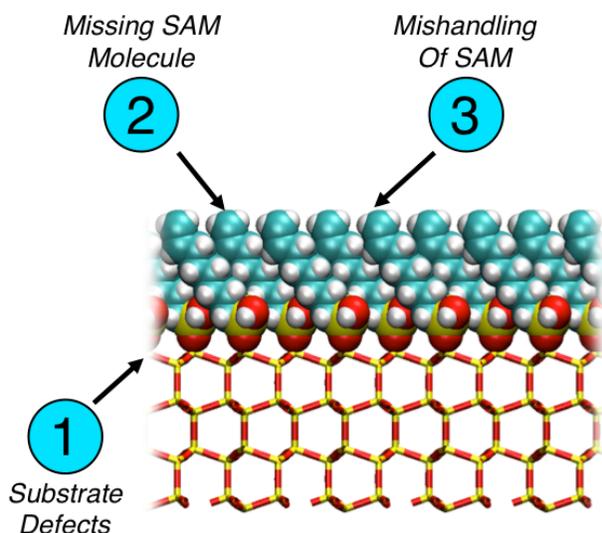


Figure 6.4. Possible defect causes in a SAM surface. Defects can be introduced to a SAM surface in a number of ways. 1) Defects in the substrate such as a rough substrate surface can introduce uneven SAM surfaces and imperfect coverage. Reuse of the same substrate for multiple SAM generations can also generate uneven surfaces. 2) Due to imperfect coverage, siloxane chains can be missing creating a hole in the SAM coverage. 3) Mishandling of the SAM can introduce large holes in the SAM surface which can have large scale effects on the behavior of water near the SAM.

Water on Rough and Mixed SAM Surfaces

Single component SAMs make up only a small part of SAM surfaces that have been studied or utilized to date. Mixed SAM surfaces, SAMs that are made up of multiple, different and unique components to form a more complex resulting molecular surface, can be split up into two groups; SAM surfaces that differ in only chain length but not hydrophobicity resulting in a rough surface, or surfaces that are composed of components that vary in hydrophobicity which result in more complex and difficult to study surfaces. Each of these mixtures can allow a greater deal of control over the resulting surface. Of note, one of the hallmarks of achieving superhydrophobicity (an extreme degree of superhydrophobicity otherwise known as the Lotus effect, named for the superhydrophobic lotus leaf), is the patterning of different length surfaces resulting in a corrugated pattern that can result in a large increase in hydrophobicity (349).

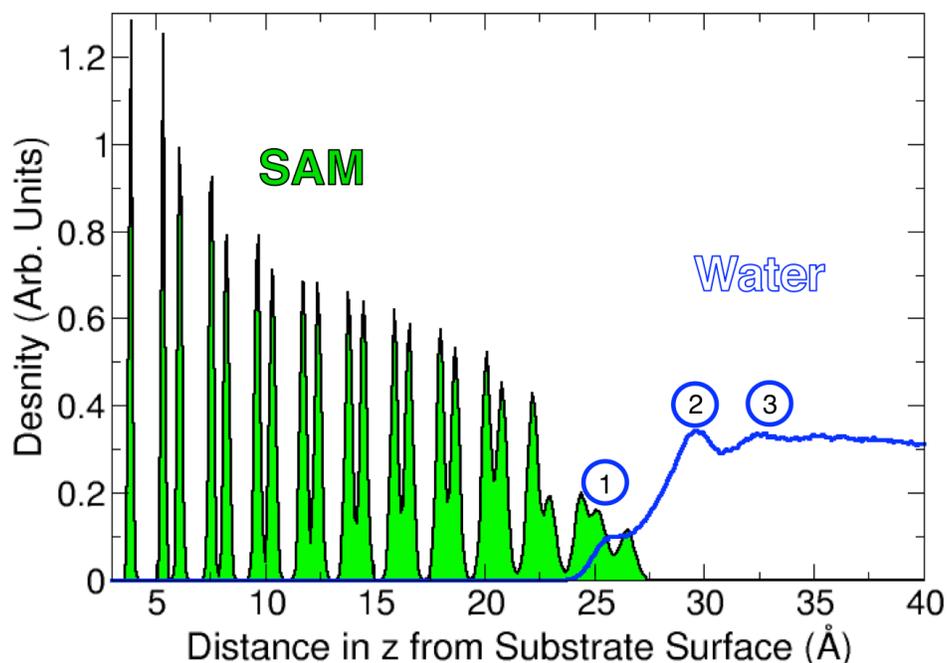


Figure 6.5. Structure of a mixed hydrophobic SAM surface. Density profile for a representative 50:50 mixture of 18-carbon and 22-carbon alkane siloxane molecules resulting in a rough SAM exposed to water. The SAM carbons are shown in green and water in blue. The final four carbons of the longer 22-carbon chains show significantly increased mobility when compared to any other region of the siloxane. Despite the hydrophobic environment in the pockets created by the outstretched terminal four carbons of the longer chains, water penetrates to the lower 18-carbon surface (highlight 1) before the previously observed structured layers of water are once again observed (Highlights 2 and 3).

Some of the earliest studies on such rough, mixed SAMs was done by Rudich and coworkers in 2000 that studied the dynamics of water on the surface of a two component mixed SAM as compared to a corresponding single component system (309). Unlike simple hydrophobic systems, there has been limited work on rough SAMs following this initial computational study. Experimental work on the wetting behavior of rough SAMs revealed a working model that the longer chains of these rough surfaces would not be rigid and show a large degree of flexibility (Figure 6.5) which would not be ideal in designing superhydrophobic surfaces (350). Computational studies have since confirmed that the protruding longer chains reorient to lay against the surface created by the shorter chain lengths (293, 309, 351).

With regards to water dynamics on these rough surfaces, one proposed way that rough hydrophobic surfaces further increase hydrophobicity is through the inhibition of droplet formation (309). While early SAM coarse-grained simulations of water on rough SAMs prevented an expected water from forming when exposed to a hydrophobic surface (309), recent molecular dynamics simulations has begun to shown that such inhibition may not in fact exist (351). To address this seemingly contrasting pair of results, in chapter eight we carry out a systematic study of a series of rough SAM surfaces (matching the roughness of the pioneering work of Rudich and coworkers) exposed to varying amounts of water and probe droplet formation.

Water Behavior on Hydrophilic SAM Surfaces

Hydrophilic SAM surfaces have become a major point of interest in SAM research due in no small part to their value in a experimental setting. Hydrophilic SAMs, particularly SAMs with a oligo-ethylene-glycol terminal group (OEG) have been shown to generate an archetypical protein resistant coating (313). Previous studies have demonstrated that such surfaces are able to resist protein absorption due to the greatly increased entropic penalty onto the SAM surface itself (352). However, it has remained a question at large whether the SAM itself, or the behavior of the water at the surface of the SAM is what controls the prevention of protein absorption (305). And of course, to truly answer this question, a deeper understanding of water structure and dynamics at such hydrophilic SAM surface is required.

The first simulations of hydrophilic SAMs were completed by Hautman and Klein in 1991 which probed the wetting dynamics of hydroxyl terminated SAMs in comparison to the fully reduced alkane terminated SAMs (336). This work was followed by Pertsin and coworkers

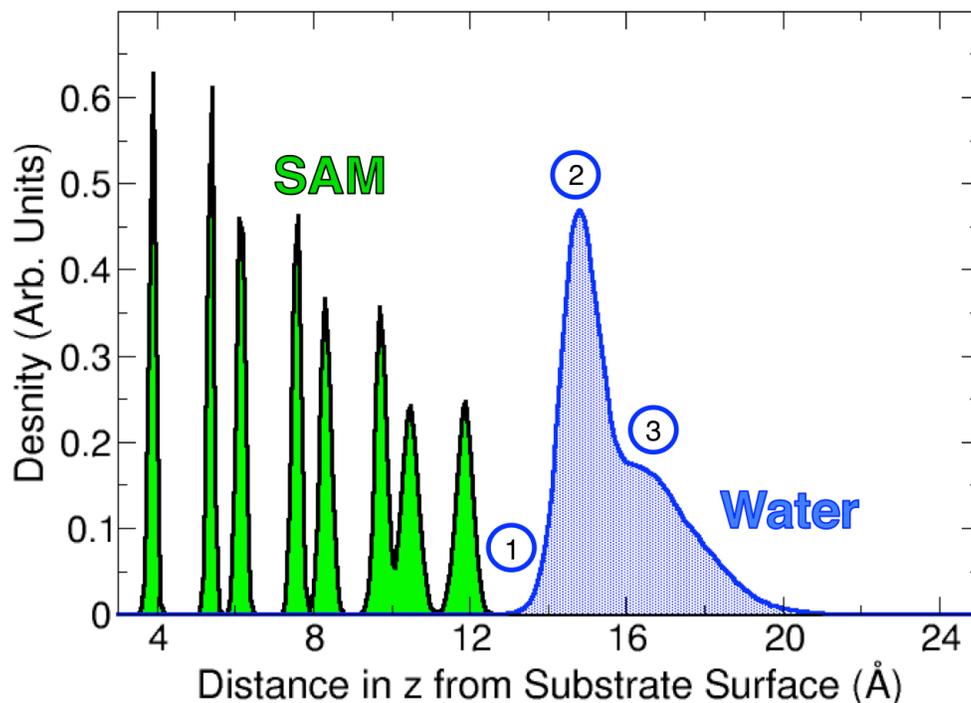


Figure 6.6. Structure of a hydrophilic SAM surface. Density profile for a representative eight-carbon carboxylic-acid terminated siloxane SAM exposed to four monolayers of water. The SAM carbons are shown in green and water in blue. Unlike the hydrophobic surfaces, formation of true monolayers are observed instead of a droplet. The gap between the peaks of the terminal carbon and initial water layer is smaller than observed in Figure 6.3 (Highlight 1), while the two distinct peaks in the water density profile correspond to water hydrogen bonding with the carboxyl (Highlight 2) and hydroxyl group (Highlight 3) of the terminal carboxylic acid.

who were amongst the first to probe the structure of OEG-terminated SAMs and the corresponding wetting behaviors (353, 354). Unsurprisingly, water at the hydrophilic SAM interface was found to form distinct layers of water near the surface (Figure 6.6) which was further probed by many studies that also probed the difference between hydrophilic terminal groups (305, 355-357). These early studies also compared hydrophilic SAMs to hydrophobic SAMs in order to probe for further differences in water dynamics and structure. One similarity that was observed was that the odd vs. even effects observed in hydrophobic surfaces (332) was also found to affect wetting behaviors in hydrophilic surfaces (358, 359). Interestingly, unlike different hydrophobic terminal groups which showed little effect on wetting behaviors,

differences in hydrophilic terminal groups resulted in different water structures at the SAM interface (355).

While these initial comparisons between terminal groups have begun to probe dependencies of wetting behavior on SAM chemical characteristics, many such dependencies have not yet been studied. When comparing the systems that have been described in the preceding section, many different chain lengths and terminal groups have been studied and rarely do two studies use similar systems. However, comparison between different SAMs are commonly carried out despite the lack of a systematic study of the wetting behaviors on such differences. In chapter nine, such a systematic study of the dependencies of terminal group, chain length, and defect size on wetting behaviors is presented to address this knowledge gap in the field.

Chapter 7

On the Wetting Effects and Water Penetration of Defective Self-Assembled Monolayers

Contribution note: The experimental validation in this study was carried out by Theresa M. McIntire of the John C. Hemminger group at UC Irvine. The computational work, analysis, and manuscript was completed primarily by the author.

Summary

Water behavior on organic surfaces can lead to a variety of complex interactions with a wide reaching affect on our understanding of atmospheric chemistry through the deposition of organics onto various surfaces. These organic surfaces will not be perfect surfaces as there will be defects present in the surface caused by missing molecules or defects in the substrate. To better understand how the presence of defects in organic surfaces may affect water behavior, molecular dynamics (MD) simulations and attenuated total reflectance infrared spectroscopy (ATR-IR) have been used in combination to study the behavior of water at the interface of a hydrophobic Self-Assembled Monolayer (SAM) of 7-octenyltrichlorosilane. ATR-IR spectra reveal that as water increases at the surface, the relative population of water molecules that are well hydrogen bonded, that is participating in three to four hydrogen bonds, increases which is also seen in computationally derived MD spectra. Both ATR-IR and MD simulations reveal a global ordering effect on the SAM surface in the presence of water when compared to a dry surface as water penetrates into the defect to make contact with the hydrophobic SiO₂ substrate. Finally, MD simulations of seven different defective SAM systems reveal that water penetration to the hydrophilic substarte begins to occur with a defect characterized by three missing adjacent siloxane molecules while smaller defects are able to prevent water penetration by internal rearrangements. Together, this study shows how drastically the presence of small defects can have on these model surfaces and how they can change water behavior when compared to a perfect defect-free system.

Introduction

Silicates are common components for a variety of windows and building materials in urban areas which are often coated in a variety of different organic and inorganic compounds (289, 290, 360, 361). The organic compounds can be hydrophobic or hydrophilic depending on the chemical makeup of the compound and the oxidation state. It is vital to understand at the molecular level the interactions of such organic surfaces with water as thin water films can be found at the surface and deposition can allow the surface to take up a variety of different molecules and can affect the mechanisms and kinetics of reactions at these interfaces (286, 292, 293, 303). The uptake of these molecule as well as water can also have a drastic effect on optical properties (311).

There has been a large body of work done to study the interfacial chemistries of water with a multitude of surface including glass, quartz, metals, lipid membranes in addition to those done on organic surfaces (286, 338, 362-371). Many of these studies, done on both organic and nonorganic surfaces identified two populations of water molecules that exist at the interfacial region that can be characterized by their hydrogen bonding properties. One population of water molecules at these surfaces exists in an ordered, bulk-like fashion that can be characterized as “well hydrogen bonded,” that is water molecules that are participating in three to four hydrogen bonds. The second population is made up of water molecules that are participating in a smaller number of hydrogen bonds and can be described as “poorly hydrogen bonded.” It is important to note that there have been a variety of nomenclatures that have been developed to describe these populations, but we will use a generalized descriptive term to avoid unnecessary confusion where those water molecules participating in three to four hydrogen bonds are “well hydrogen bonded” and those that are participating in less than three as “poorly hydrogen bonded.” A

variety of unique water structures have been observed and proposed as well in addition to the generalized populations through the use of IR techniques (338). Despite the collective work however, water on organic surfaces is still not well understood.

To better study water on organic surfaces, one possible approach is to employ a model system that can be altered to create defects to identify changes in behavior upon the introduction of such defects (334, 338, 372). To this point, self-assembled monolayers (or SAMs) have been widely used as a simple model for organic surfaces (335, 336, 351, 358, 363, 373, 374). SAMs are made up of a series of generally simple molecules that can form a thin layer on a reactive substrate surface that alter the surface environment by the choice of the terminal SAM moiety opposite from the substrate used (375, 376). As SAM systems can be fairly simple and can be broken down into a small unit cell for study, molecular dynamics has been a vital tool to investigate interactions at the gas-phase exposed surface of SAMs. A variety of simulations have been done to study such systems with different chain lengths and moieties using SAMs as a model system to study water behavior, uptake and roughness (292, 293, 311, 325, 339, 350, 372, 373, 377, 378). It has been known that the process of creating such monolayers or mechanical stress of the final product can lead to defects consisting of missing SAM molecules in the new surface. These defects could allow water to penetrate the monolayer to reach the substrate that could dramatically alter the environment at the surface. Recent work has begun to shed light onto the importance of defects in understanding the behavior of water at organic SAM surfaces (372). Defects of the organic surface due to either missing SAM molecule(s) or imperfections (e.g., step defects) of the substrate can alter the behavior of molecules adsorbed at the surface and can play an important role in the uptake of molecules onto the surface (292, 293). Specifically, how defects in these organic surfaces may affect the behavior of water is poorly understood. While

some work has been done addressing the issue of interactions between water with a defected SAM, it focused predominantly on the impact of water on the tribological properties of the SAMs (372). However, there has been little work to study the effects of small defects consisting of only a handful of SAM molecules on water behavior, in particular for the low-hydration regime when (sub)monolayer amounts of water are present on the SAM surface.

We report here the results of molecular dynamics simulations and infrared spectroscopy (IR) experiments that were used to investigate the behavior of water on the surface of an alkenylsilane SAM on silica substrate. By using molecular models of defective SAMs, the aim is to determine the molecular-level details of water behavior near the defect sites and to obtain more direct correspondence between the experimental and model systems compared to previous work, in which MD simulations were done on perfect, defect-free SAMs (338). Of specific interest in the present study is the ability of water to penetrate the SAM surface to make contact with the hydrophilic SiO₂ crystal substrate. We show that at low levels of water on the silica substrate, computationally generated IR spectra from the molecular dynamics simulations match the results that are seen experimentally, that as more water is introduced to the system, the relative population of well hydrogen bonded water molecules increases. Moreover, the addition of water to a defective SAM surface produces a global ordering effect on the SAM, which is confirmed both experimentally and computationally. Finally, to shed further light on the ability of these waters to penetrate into defective SAMs, the water penetration and structure is studied on a variety of SAM defect systems.

Methods

Preparation of ATR crystals.

A ZnSe or Si attenuated total reflectance (ATR) crystal (Pike Technologies, 80 mm × 10 mm × 4 mm) allowing 10 internal reflections along the length of the crystal was used both as a substrate and as an optical element with which to detect adsorbed water vapor and to monitor changes in the attached SAM film. Prior to use and chemical derivatization, the ATR was boiled in ethanol, boiled in dichloromethane, dried with UHP nitrogen (Praxair, 99.999%) and placed in an argon (Praxair, 99.999%) plasma discharge cleaner (Harrick Scientific Plasma Cleaner/Sterilizer PDC-32G, medium power) for ~30 min to remove organic contaminants. After removal from the plasma cleaner, the clean ATR crystal was rinsed with purified water (Milli-Q Plus, 18.2 MΩ-cm) and dried with UHP nitrogen. The clean ATR crystal was coated with 7-octenyltrichlorosilane [H₂C=CH(CH₂)₆SiCl₃] (Sigma-Aldrich, mixture of isomers, 96%) in hexadecane to produce a self-assembled monolayer, designated hereafter as C8=. A detailed description of the procedure for depositing the SAMs on ATR crystal surfaces has been described in detail previously (310, 338, 343, 363, 379).

Exposure of surfaces to water vapor.

Using a horizontal trough plate ATR flow cell (Pike Technologies) inside the sampling compartment of a FTIR spectrometer (Nicolet 6700 or Avatar 370, ThermoFisher Scientific), the SAM-coated ATR crystal was purged overnight with dry UHP nitrogen.

Water vapor was generated by bubbling dry nitrogen through two water-filled borosilicate fritted glass bubblers in a thermostated water bath at 298 ± 2 K. Desired relative humidity was generated by mixing measured flows of the dry and humidified nitrogen resulting in a total flow ~200-220 mL min⁻¹ (286, 310, 338). A portion of the humidified flow (70–80 mL min⁻¹) was diverted to the horizontal ATR (HATR) flow cell (286, 310, 338). The HATR cell

was purged with the humidified flow for ~15 minutes after which sample scans were collected. The depth of penetration of the evanescent wave from a ZnSe ATR crystal was calculated to be 2.01 μm , at 1000 cm^{-1} region, much greater than the thickness of the organic (~1.3 nm) thin film. The infrared beam interrogates the entire organic monolayer and provides a spectrum similar to a transmission spectrum. All ATR measurements were taken at atmospheric pressure. For each experiment, either 2000 or 1024 scans were collected at 4 cm^{-1} resolution. The spectra are calculated as $-\log(S/S_0)$, where S_0 is the single beam intensity of the dry C8= coated ZnSe ATR crystal and S is that of the hydrated C8= coated ZnSe ATR crystal.

Molecular Dynamics Simulations

The substrate for all simulated SAMs was modeled as an ordered lattice of an alpha-tridymite SiO_2 crystal, which has been described in a previous study(293). The substrate dimensions were $78.66 \times 75.60 \times 14.00$ Å. The SAM systems were constructed from 270 individual C8= (7-octenyltrihydroxylsilane, $\text{H}_2\text{C}=\text{CH}(\text{CH}_2)_6\text{Si}(\text{OH})_3$) molecules. The molecules were covalently bonded to one surface of the SiO_2 crystal slab via the hydroxyl sites (removing one terminal oxygen and two hydrogen atoms). Half of the Si atoms of the opposite surface of the SiO_2 slab were removed in order to maintain charge neutrality of the simulation cell. (Note that the opposite surface of the SiO_2 slab was not chemically capped with H atoms.) The full system is seen in Figure 7.1.

A 1.0 kcal/mol/Å² constraint was placed on the positions of all atoms in the substrate as well as to the terminal O and Si atoms of the alkenylsiloxane chains. A reflective wall potential located at $z = 75$ Å was used to prevent water molecules in the vapor phase from interacting with the unsaturated SiO_2 surface through periodic boundary conditions. A layer of vacuum included

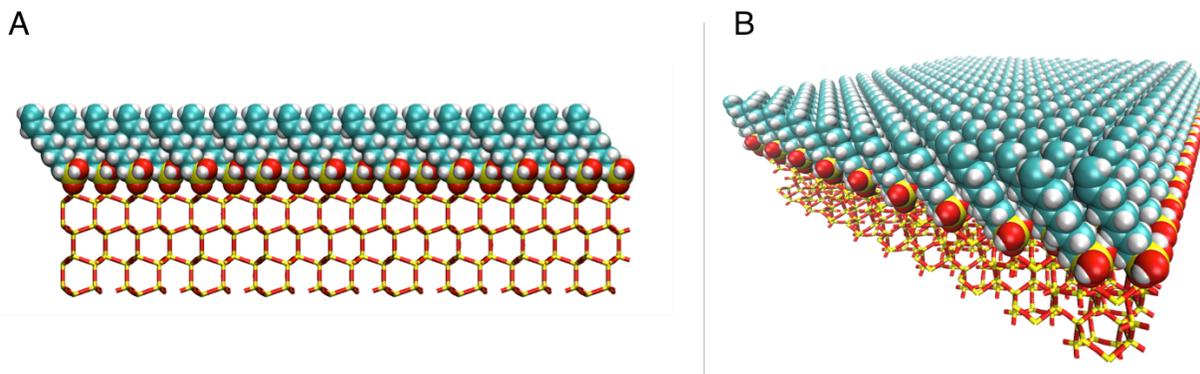


Figure 7.1. Structure of the self-assembled monolayer system. Shown is the experimental system that was used in this investigation. The experimental system consists of a solid SiO₂ crystal slab with 270 7-octenyltrichlorosilane SAM molecules with a terminal double bond positioned over the top layer of Si atoms of the crystal slab. Two views of the systems are shown; a side on view of the system is shown in (A) and a view from above the system in (B).

by extending the simulation cell by an additional 25 Å along the z-axis to ensure that the thickness of the layer of vacuum (vapor phase) separating the periodic replicas of the SAM-coated SiO₂ crystal was large enough for the long-range intermolecular interactions to be negligible. The resulting unit cell for the simulations was $78.66 \times 75.60 \times 100 \text{ \AA}^3$. Because our model explicitly included an inherently anisotropic solid substrate, possessing a permanent dipole, we employed a correction term to both the forces and the electrostatic potential which accounted for the lack of periodicity in the z dimension and the permanent dipole moment of the system. The alkenylsiloxane force field parameters were adapted from the CHARMM27 force field which has been successfully applied to MD simulations of alkylsiloxane SAMs on SiO₂ (293) as well as alkylthiolate SAMs on gold (338). The parameters for the solid SiO₂ crystal were taken from the work of Stevens et al (380). Water molecules were modeled using SPC/E parameters (381).

Defects were created in the monolayers by removing individual alkenylsiloxane molecules and replacing each with a hydroxyl group. By removing varying numbers of molecules at varying positions from the defect-free SAM, six different SAM systems with

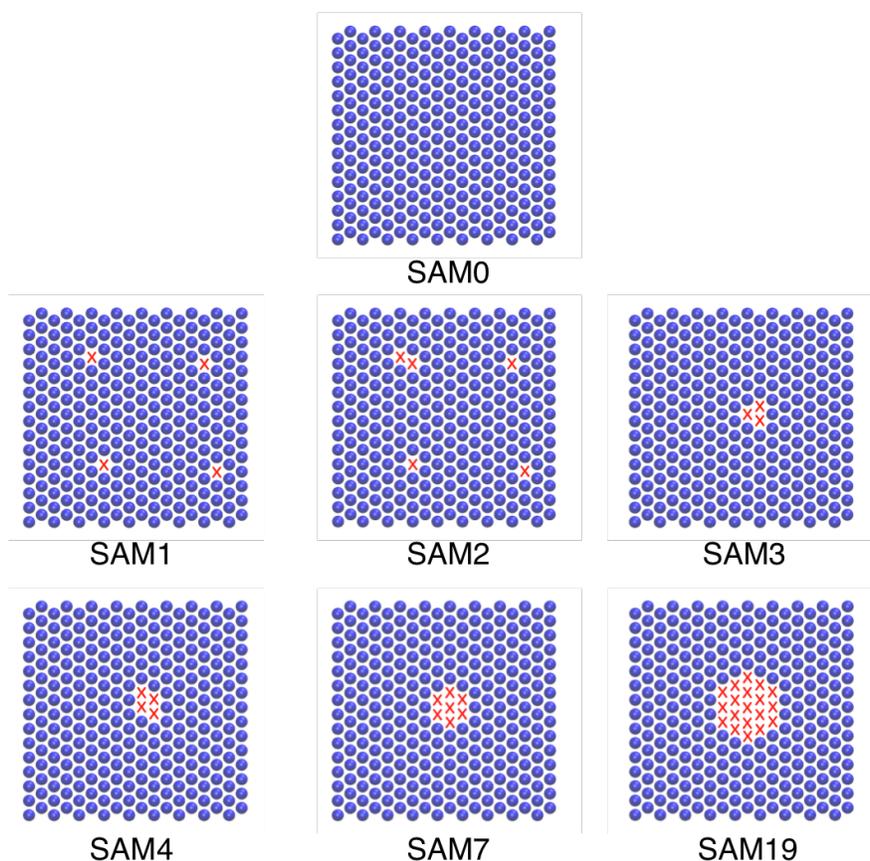


Figure 7.2. Schematic of designed defective SAM systems. Shown are schematic figures of each of the seven systems that were studied with a blue sphere representing a SAM molecule and a red “X” representing a defect (a missing SAM molecule and a terminal -OH group on the crystal slab in its place). The first two defects study pinprick defects and recreate the studies from a previous study (293). The following systems study defects of size, 3, 4, 7 and 19. Each system is named as SAMX, where “X” is the size of the largest defect present in the system.

defects of varying size and distribution throughout the SAM were created, designated as SAM1 through SAM6 (see Figure 7.2). The defect-free SAM is referred to as SAM0. The smallest defects consisted in a single missing alkenylsiloxane chain, followed by the defects of size 2, 3, 4, 7, and 19 missing chains. Each SAM was first equilibrated for 10 nanoseconds in the absence of water. Next, a defect-free perfect SAM was hydrated by bulk liquid water using the solvate tool of VMD (179). The hydrated system was then equilibrated for 10 nanoseconds. From the final configuration of the equilibrated bulk run, three different water coverages were prepared to be used in further simulations by selecting all water molecules that were less than 1, 2, and 4 Å

away from the SAM surface. The resulting amounts of water correspond to 191, 412 and 822 water molecules, approximately an equivalent of one quarter, one half and one monolayer. An additional sub-monolayer coverage consisting of the 26 water molecules closest to the surface of the perfect SAM was also prepared. These four different water amounts (with their initial configurations determined by the above process) were then added to SAM systems depicted schematically in Figure 7.2, resulting in the total of 28 initial conditions for further simulations. The microhydrated SAM systems were propagated for 10 nanoseconds, the final 5 nanoseconds of these trajectories were used for structural analysis, with configurations saved at 1ps intervals. An additional 20ps trajectory, which continued from the end of the above simulation, was collected for each system, writing out atomic positions every 5 fs for the generation of IR spectra (see below).

The simulations were carried out using the NAMD 2.7 software package. All simulations were performed in the isochoric-isothermal (NVT) ensemble at 300 K using the Langevin thermostat with a 1ps coupling coefficient. The classical equations of motion were integrated by Verlet algorithm propagator with a timestep of 1 fs. The short-range interactions were truncated at 20 Å. Partial Ewald mesh was used to calculate the columbic interactions with an external electric field of magnitude $-5.44544 \text{ kcal}/(\text{mol } \text{Å } e)$ as the z-periodicity correction described earlier. All bonds with hydrogen were held constant by the SHAKE and SETTLE algorithms (382).

IR Spectra Calculation

The IR spectrum was computed following the method of Corcelli et al. (383) In this method, the OH stretch band is calculated by using an empirical relationship between the electric

field and the anharmonic stretch frequency calculated from density functional theory. This relationship, in conjunction with the dipole autocorrelation function, allows for the efficient calculation of infrared absorption line shapes from a molecular dynamics trajectory. See Corcelli et al. 2004 for more details (383).

The computed IR spectra can be broken down to identify features based on hydrogen bonding characteristics. Corcelli et al. showed that by isolating water molecules based on their hydrogen bonding patterns, the peaks that were identified in the IR spectra could be attributed to hydrogen bonding patterns (383). Verification of these peak identifications can be found in Figures 7.3-5. To determine the relative populations of the high and low-hydrogen-bonding populations from our own simulations, two Gaussian curves were fit to each computationally generated spectra with the location of the center of each peak held constant to the position observed for each species (well hydrogen bonded and poorly hydrogen bonded) using the following formula:

$$y = a_1(2\pi b_1^2)^{-0.5} e^{-\frac{1(x-c_1)^2}{2b_1^2}} + a_2(2\pi b_2^2)^{-0.5} e^{-\frac{1(x-c_2)^2}{2b_2^2}} \quad (1)$$

where x is the wavenumber for the calculated spectrum, y is the observed absorbance in the computational spectrum, a a part of the normalization coefficient, b the standard deviation for the Gaussian distribution and c the wavenumber at which the distribution is centered. During curve fitting, c (the location in wavenumbers of an identified peak in the computational spectrum) was held constant to match the corresponding well hydrogen bonded and poorly hydrogen bonded peaks while a and b were allowed to fluctuate to achieve the best curve fit. The relative populations were then calculated using the normalization coefficients of the distributions:

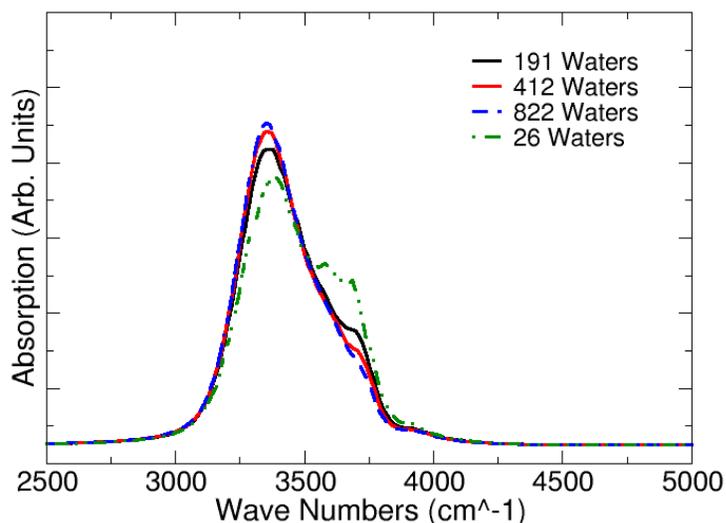


Figure 7.3. Example MD generated IR spectra from the SAM7 system. Shown are the four MD generated IR spectra from the four different hydrated simulations on the SAM7 system. Note that the peak centered at 3600cm^{-1} which has been identified as the poorly-hydrogen bonded peak recedes as more water is added to the system while the peak centered at 3400cm^{-1} identified as the well-hydrogen bonded peak grows in with increased water. Similar trends were seen for all systems, as seen in Figure 8.

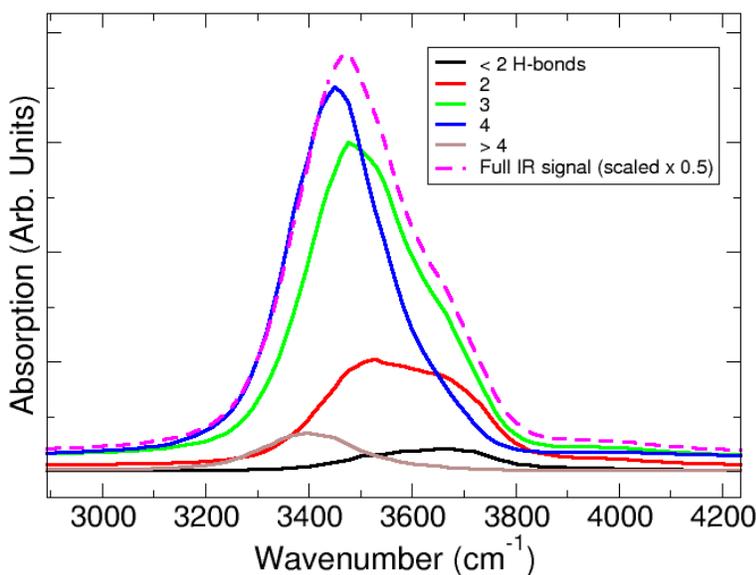


Figure 7.4. Deconstruction of MD-generated bulk water IR spectrum for peak assignments.

A simulation of bulk water was used to generate an example IR spectra, shown in a dashed pink line. The spectra is then deconstructed by separately calculating the IR spectra for water molecules participating in less than two (black), two (red), three (green), four (blue) and greater than four hydrogen bonds (brown). Those waters that are poorly hydrogen bonded correlate with a larger peak centered near 3600cm^{-1} while those spectra

corresponding to well-hydrogen bonded molecules correlate to a peak centered near 3400cm^{-1} . On the basis of the two peaks identified in our simulated systems, this deconstructed spectra and the extreme, low hydrated, SAM-penetrating deconstructed spectra in Figure S4, the 3400cm^{-1} peak was assigned to well-hydrogen bonded water and the 3600cm^{-1} peak to poorly hydrogen bonded water.

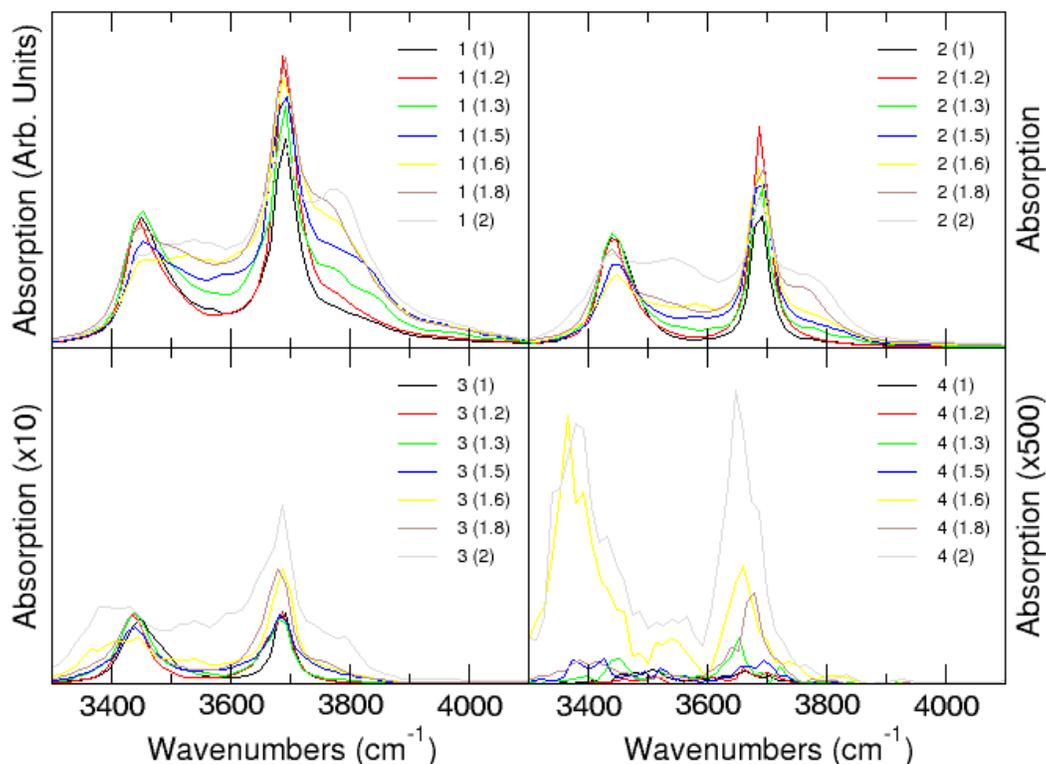


Figure 7.5. Deconstruction of MD-generated penetrating water spectrum for peak assignment confirmation. An eighth SAM system was generated containing 19 defects made up of two adjacent missing organic chains, the smallest defect that water penetration of any kind was observed in our present simulations, distributed evenly throughout the surface. Seven different hydration levels were carried out by placing one molecule of water in each defect, and then randomly placing a second water molecule in three defects to achieve the next hydration level with an average of 1.2 water molecules per defect. This process was repeated until each defect had two waters inside. The repulsive wall described in the methods was lowered to $z = 28\text{\AA}$ to insure that all waters remained in the defect throughout the simulation, which then followed the same procedure as described in the Methods section. MD-generated spectra were then generated and deconstructed as described in Figure S3. From the top right and going clockwise, the first panel shows the spectra of those waters participating in one, two, three and four hydrogen bonds as designated by the first number in each legend. The following value in parenthesis describes the average number of waters per defect. Like Figure S3, the identity of both the well hydrogen bonded peak at 3400cm^{-1} and poorly hydrogen bonded peak at 3600cm^{-1} are reaffirmed as the 3600cm^{-1} peak shrinks as the 3400cm^{-1} peak grows in with increasing number of hydrogen bonds. The slight shift in peak location for each peak when compared to the bulk water spectra is due to the presence of the SiO_2 crystal slab in the IR calculation, but as shown above, has no affect on the identity of each peak with regards to hydrogen bonding.

$$P_R = \frac{a_1(2\pi b_1^2)^{-0.5}}{(a_1(2\pi b_1^2)^{-0.5} + a_2(2\pi b_2^2)^{-0.5})} \quad (2)$$

Results and Discussion

Experimental ATR-FTIR Spectra of dry and hydrated C8= SAM.

The infrared methylene stretching region of the dry C8= SAM on a Si ATR crystal is shown in Figure 7.6A. IR spectra of this SAM have been recorded by multiple studies in the past, and are well understood (310, 338, 343, 363, 379). Spectral bands in the region from 3000-2800 cm^{-1} are attributed to the symmetric and antisymmetric C–H stretches of the CH_2 groups of SAM chain, with the FWHM indicating the degree of order of the alkyl chains (larger FWHM corresponding to less well-ordered SAM).

The extent of water adsorption to the C8= SAM was monitored using ATR-FTIR by adding a controlled amount of HOD vapor in the sealed ATR cell. Figure 7.6B shows a series of IR spectra of the SAM exposed to HOD vapor as a function of relative humidity at 298 K. The adsorption of water (HOD) by the SAM produced a broad O–H peak centered at $\sim 3500 \text{ cm}^{-1}$ that increased as the RH increased up to 45% RH. Concomitantly, the spectra also showed positive methylene stretching peaks that shifted to lower wavenumbers, indicating increased chain order, as the SAM adsorbed water (384). This increased chain ordering was also seen for longer chain-length SAMs (data not shown). Figure 7.6C shows an example of the deconvolution of the OH stretching region of the IR spectra obtained at 40% RH on C8= SAM into 2 separate contributions using Equation 1. This deconvolution showed a more prominent (relative to the dry SAM) broad peak centered at ca. 3498 cm^{-1} and a lower peak centered at 3305 cm^{-1} , attributed to amorphous and structured water, respectively. The deconvolution into these two contributions was performed for all collected spectra. From the deconvoluted spectra, the relative

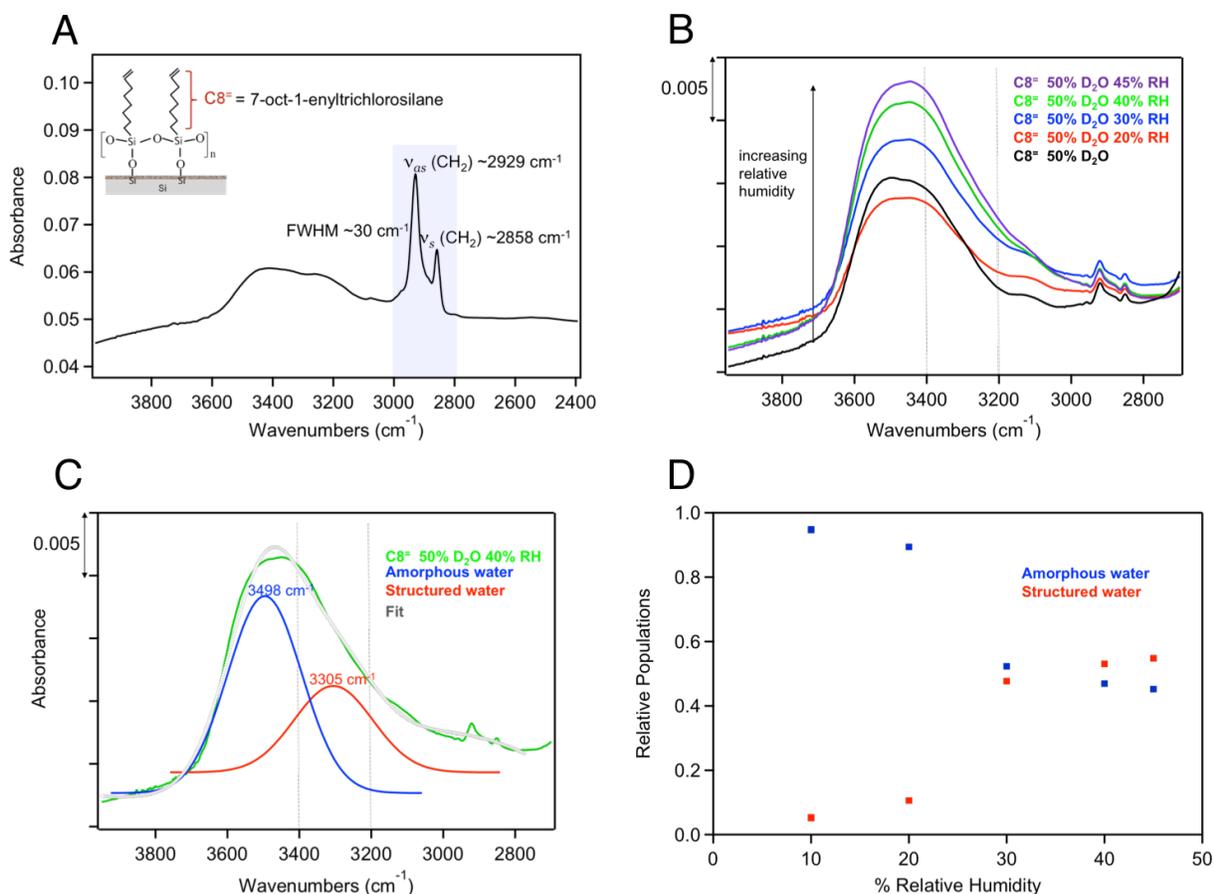


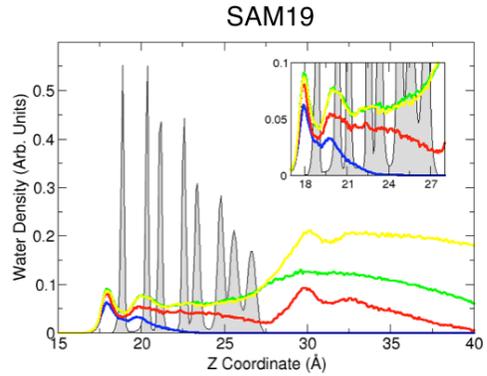
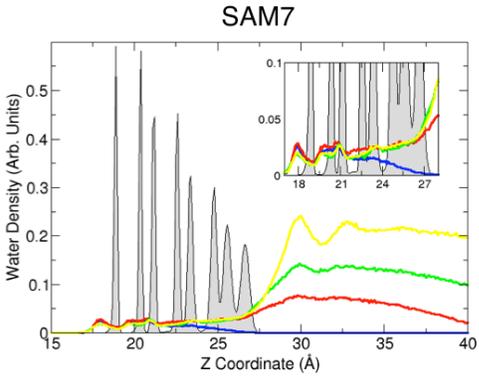
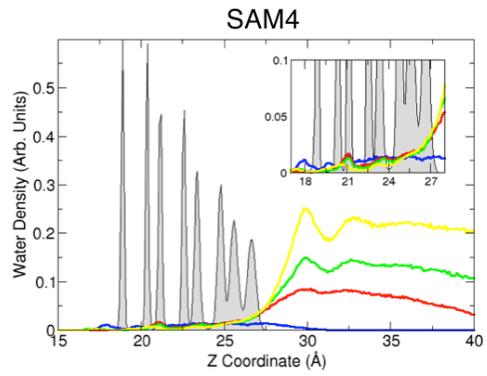
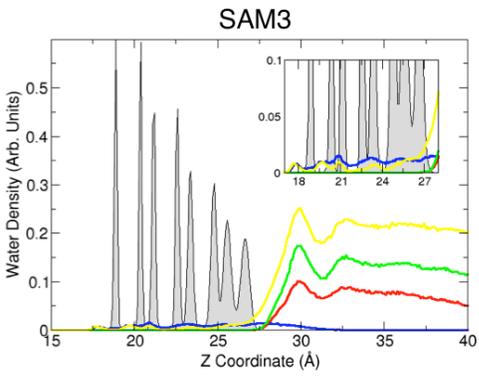
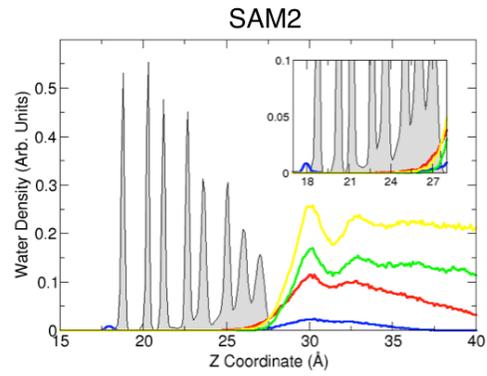
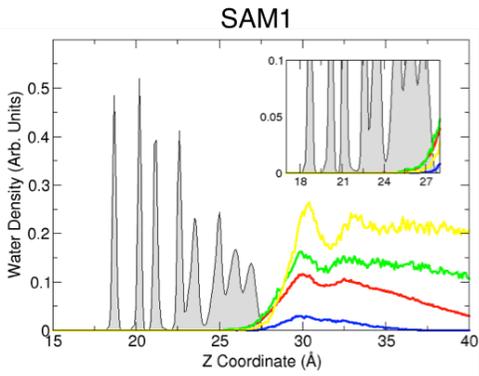
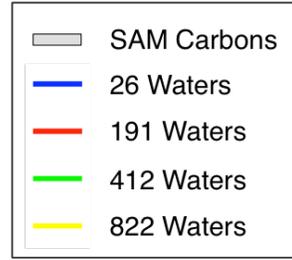
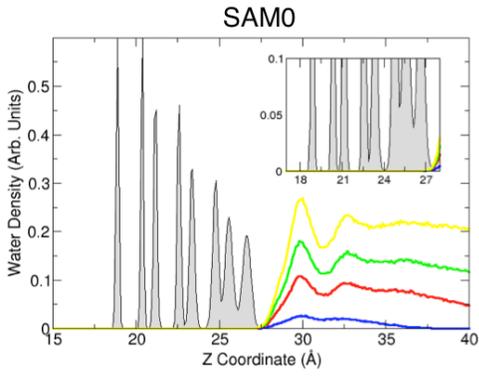
Figure 7.6. ATR-IR spectra from 7-octenyltrichlorosilane SAM. A) Spectra of the dry 7-octenyltrichlorosilane SAM, highlighting peaks associated with methylene peaks that indicate a “less-ordered” SAM. The broad band is the uncrosslinked lateral OHs in the O-Si-O siloxane backbone. B) Spectra of 7-octenyltrichlorosilane SAM with increasing water. Positive methylene stretching peaks near 3000 cm^{-1} are observed to shift to lower wavenumbers which indicates a more ordered SAM. C) The deconvolution of the 40% relative humidity that is used for the assignment of relative populations of water species. D) The relative population of well hydrogen bonded water increases as water increases.

populations of the amorphous and structured water was obtained as a function of relative humidity, as seen in Figure 7.6D. The relative contribution of the higher frequency ($\sim 3500 \text{ cm}^{-1}$) band attributed to the amorphous water is decreasing with increasing RH, whereas the relative contribution of the structured water ($\sim 3300 \text{ cm}^{-1}$) band is increasing. It is possible that the source of the 3300 cm^{-1} band is water penetration through the SAM and its interaction with both the uncross-linked SAM chains and or substrate. Water penetration through SAMs has been

previously observed in contact with the substrate despite the presence of a SAM by sum frequency generation spectroscopy and ATR-IR (369, 385) as well as at defects and edges (345). It is pertinent to note that this trend is different to the trend observed by Moussa et al (338), which saw a decrease in the structured water band with increased relative humidity. However, the two systems being studied here are different in several key ways. While the previous study was done on glass, the present work was done on a SiO₂ crystal. This possibly important change was also reflected in the following molecular dynamics simulations by modeling the crystal substrate where previously no crystal substrate was present.

Molecular dynamics simulations of water on defective SAMs

To study the possible dependence of water behavior on a hydrophobic SAM surface with the presence and size of present defects, seven representative SAM systems made up of double bond terminated 7-octenyltrichlorosilane, H₂C=CH(CH₂)₆SiCl₃ siloxane molecules (or C8=) were simulated in the presence of five different hydration levels; dry, 26 waters, 191 waters, 412 waters and 822 waters (or roughly one monolayer of coverage). See Figures 7.1, 7.2 and Methods for a more detailed description of how each system was prepared for study. In all simulation, the resulting equilibrated waters formed small droplets on the surface of the hydrophobic surface with water from these droplets found to penetrating into the SAM surfaces in the presence of a defect. Density profiles for each of the systems (Figure 7.7) reveal several important characteristics of the simulation at large. First, the presence or absence of water on a global scale does not affect the density profiles for the SAM carbon atoms in any system, as the profiles for each water coverage show no significant difference. This effect is due to the low defect density in these model systems that are lower than what would be expected in an



(Previous Page) Figure 7.7. Density profiles of each SAM system at varying water coverages. Shown are the density profiles of the each of the seven SAM systems studied herein with varying degrees of water present. The density profiles of the SAM carbon atoms shown in gray did not change significantly between the five different water levels (including simulations in the absence of water) so the density profile for the sub-monolayer run is displayed. Inserts are shown that highlight the water density within the monolayer for each system. The water profiles displayed are for 26 water molecules (blue), 191 (red), 412 (green) and 822 (yellow). The number in the name of each system (i.e. SAM19) corresponds to the largest defect present in each system, see Figure 2 for more details.

experimental surface. Second, water penetration is firstly observed in SAM2 (SAM surface with largest defect size of two adjacent siloxane molecules). In the SAM0 and SAM1 systems, no water penetration is observed in any simulation. In the SAM2 system, a single system showed appreciable penetration with a single peak corresponding to a single penetrating water molecule in the smallest wet simulation, which occurred in the largest defect in the system made of two missing adjacent organic chains. The SAM3 system, containing three adjacent missing molecules, is the smallest defect wherein water sizable water density is observed all the way through the SAM, extending from the droplet on the SAM surface to the SiO₂ substrate as seen from the non-zero water density both within the organic monolayer and in the direct vicinity of the substrate. When the defect is grown by an additional missing siloxane, full water penetration is observed at all hydration levels, a trend that continues for the remaining SAM7 and SAM19 systems. Such transient water penetration has also been observed in a previous study (293) though interestingly, transient water penetration was observed for a defect of only one missing siloxane which was not observed in our present simulations. However, the previous system used a saturated SAM chain while in this study we use an unsaturated double-bond terminated chain. This finding may suggest that even such a small change may affect the ability of water to penetrate to the substrate. Finally, the density profiles also show a degree of ordering in the water adsorbed at the surface of the SAMs as observed in the distinct peak in the water densities

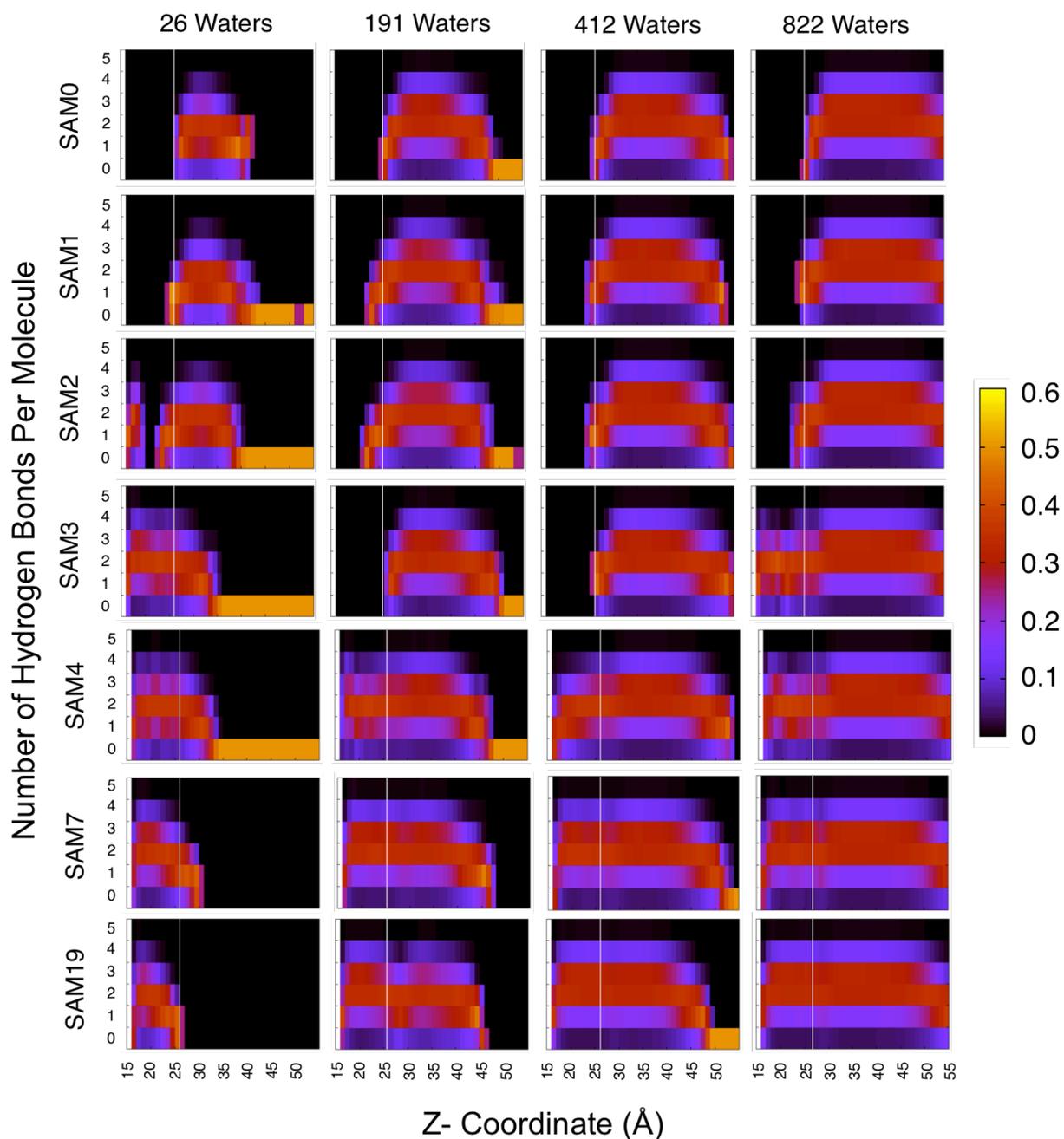


Figure 7.8. Hydrogen-bonds per molecule along the z-axis. Intensity maps of the average number of hydrogen bonds per molecule as a function of z-coordinate of the oxygen atom of the water molecule. All waters that occupied a given 1 Å range in the Z-axis (normal to SAM surface) at any point during the final nanosecond of the simulation were used to generate a histogram for all waters in that slab of the Z-axis by how many hydrogen bonds the water molecules participated in. This data was then combined to generate an intensity map for the full system. In some cases a only a single water molecule inhabits a given slab over the final nanosecond, which due to poor sampling, gives a single orange band for that slab.

centered at $Z \sim 30 \text{ \AA}$ in all simulations suggesting that there is a structured layer of water in contact with the SAM surface.

To further probe the water penetration in each of the defected SAMs, the average number of hydrogen bonds per water molecule along the Z -axis of the simulation cell (normal to the SAM surface) was calculated and plotted as an intensity map (Figure 7.8) These Z -resolved hydrogen-bonding maps corroborate the information obtained from the density profile analysis. It is important to note that the hydrogen-bond maps are built as a series histograms of all waters at any time point in the final nanosecond while the density profiles were averaged over the final five nanoseconds. Because of this difference, the presence/absence of water in the defect within the SAM in Figure 7.8 does not necessarily correlate exactly to water penetration seen in the density profiles. For example, if a single water molecule inhabits a given Z -range for only a brief moment but was the only water molecule to inhabit that zone for the entire simulation, their would be minimal density in the water profile but a large spike in hydrogen-bond map. This phenomenon occurred in several simulations as a single water molecule enters gas phase causing the single band of high intensity in the hydrogen bond maps because that single water which never participated in a hydrogen bond, was the only molecule to inhabit those Z -ranges in the time analyzed. However, for the majority of the system, there were enough waters present to achieve good sampling for the maps.

The analysis in Figure 7.8, together with that of Figure 7.7, reveals that the water adsorbed on the SAM-coated SiO_2 substrate can be divided into the following three categories according to the water – water hydrogen bonding characteristics:

- (1) In the cases that show no or minimal water penetration into the SAM (defects size of one and two), water forms a droplet at the hydrophobic SAM surface with under-coordinated

water molecules (predominantly making 1 – 2 hydrogen bonds) in contact with the SAM surface ($Z \sim 25 \text{ \AA}$) as well as in the outermost parts of the droplet exposed to the vapor phase, whereas the hydrogen bonding environment inside the droplets is more bulk-like, with 3 – 4 hydrogen bonds per water molecule;

- (2) In the systems containing defects of size two, three and four, the hydrogen-bonding patterns of water within the SAM show a water-wire like distribution, with the water molecules inside the defect poorly hydrogen bonded (1 – 3 hydrogen bonds per water molecule). These defects are obviously too small to support bulk-like water.
- (3) In the defects that allow larger amounts of water to penetrate all the way to the substrate and fill the defect (defects size of seven and nineteen), the water inside the defect exhibits more bulk-like character, similar to the water in the interior of the surface droplets, with 3 – 4 hydrogen bonds per water molecule.

Water Penetration into a defect leads to an ordering effect

In order to study the possible ordering of the SAM surface through the presence of water, the 2nd order Legendre polynomial of the cosine of the angle formed by each C-H bond and a vector normal to the SAM surface for every SAM carbon individually as shown below:

$$P_2 = \frac{1}{2}(3 \cos \varphi^2 - 1) \quad (3)$$

The order parameters were calculated for all the micro-hydrated SAM systems that had been used to generate the computational IR-spectra to analyze the ordering effects that had been experimentally observed. As these SAM surfaces have a much lower defect density than those

studied experimentally, analysis is focused on the defect neighboring siloxane residues by selecting those molecules for which the terminal Si atom was within 8\AA of a defect to more accurately represent a realistic defective SAM surface.

Studying the order at each of the eight carbons of the SAM chains shows a trend of higher order in the presence of water when compared to dry systems as seen in Figure 7.9. In the largest defect systems, this trend begins to correlate SAM order and water present in the system,

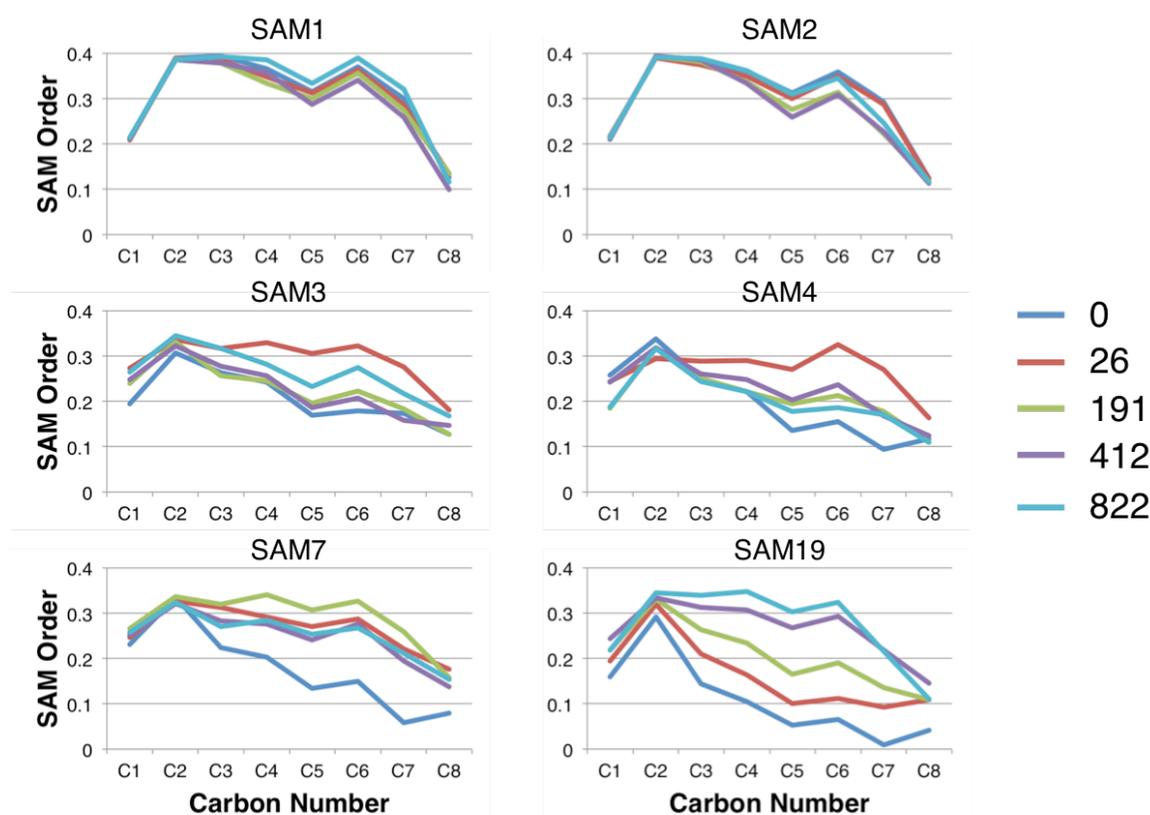


Figure 7.9. Order parameters calculated for simulated monolayers near defect sites at varying hydration levels. Shown are the average 2nd order Legendre polynomial (order parameter) of the cosine of the angle that each C-H bond makes with a vector normal to the SAM surface as a function of carbon number for defect-neighboring SAM molecules as defined by a SAM molecules whose terminal Si atom is within 8\AA of a defect site. The smallest defects (SAM1 and SAM2) show minimal ordering effect in the presence of varying amounts of water. The SAM3 and SAM4 systems show the largest increase in order in the smallest hydration level, while the largest defects show increasing order with increasing water coverage. As the SAM molecules used herein were defined by being close to a defect, the perfect surface SAM0 is not shown. Full system order parameter distributions (including those not near defects) are shown in Figure 7.10.

though this effect is not as apparent in the small defects. The ordering effect is more pronounced in carbons two through seven with the initial and final carbons seemingly unaffected by the increasing hydration levels. This finding is due to the fact that there is much less conformational freedom for carbon-1 for the SAM as it is in close proximity to the crystal substrate while the final carbon that is exposed to the solvent show the same degree of disorder due to the exposure to solvent. In a perfect SAM (devoid of defects), there is no observable ordering in the presence of any amount of water (Figure 7.10) while in the presence of small defects or size one or two, there is only a small degree of ordering present. This absence of ordering effect of water on this perfect SAM is due to the water only being able to influence the arrangement of the terminal –CH₃ groups due to the absence of defects to allow for significant water penetration. Larger defects however reveal a dramatic increase in order with additional water as seen in the SAM19 system. One interesting trend was that in the smaller defect systems, the smallest level of hydration produced the largest ordering effect. It was observed in these systems, water could not enter the defects at higher hydration levels which could be due to the fact that the potential energy barrier for these waters to leave the well hydrogen bonded environment of the droplet to a poor-hydrogen bonding environment in the defect was too large to observe. A visual representation of the degree to which defects in SAMs rearrange in the presence of water is presented in Figure 7.11 which highlights the large changes that occur at an originally dry defect site when it is filled with water.

MD calculated IR spectra

With the ability to derive IR spectra from molecular dynamics simulations of defective SAM surfaces (383), it is now possible to study the behavior of water not only at the SAM

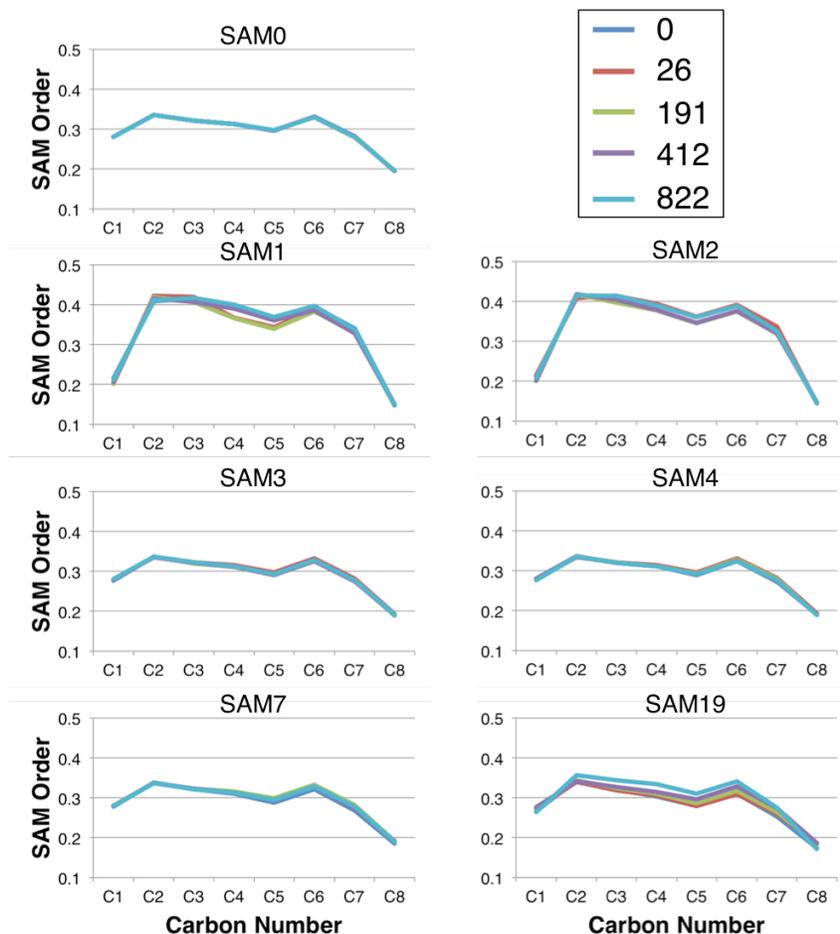


Figure 7.10. Order parameters calculated for full simulated monolayers at varying hydration levels. Shown are the average 2nd order Legendre polynomial (order parameter) of the cosine of the angle that each C-H bond makes with a vector normal to the SAM surface as a function of carbon number for all SAM molecules in each system. At a global scale, there is no significant difference in the SAM ordering present in the SAM0-SAM7 systems. There is a slight difference in order in the largest defect system (SAM19), but it does not correlate well with water coverage.

surface, but also those waters that are able to penetrate through the SAM to reach the hydrophilic substrate. Of particular interest is studying the IR spectra of the water in these two scenarios. As more water is added to the surface of a hydrophobic SAM and a larger droplet forms, it would be expected that the relative population of well hydrogen bonded water molecules would grow faster than the poorly hydrogen bonded molecules. As the size of the droplet formed on a hydrophobic SAM increases, the volume of the droplet “sphere” (which corresponds to the well

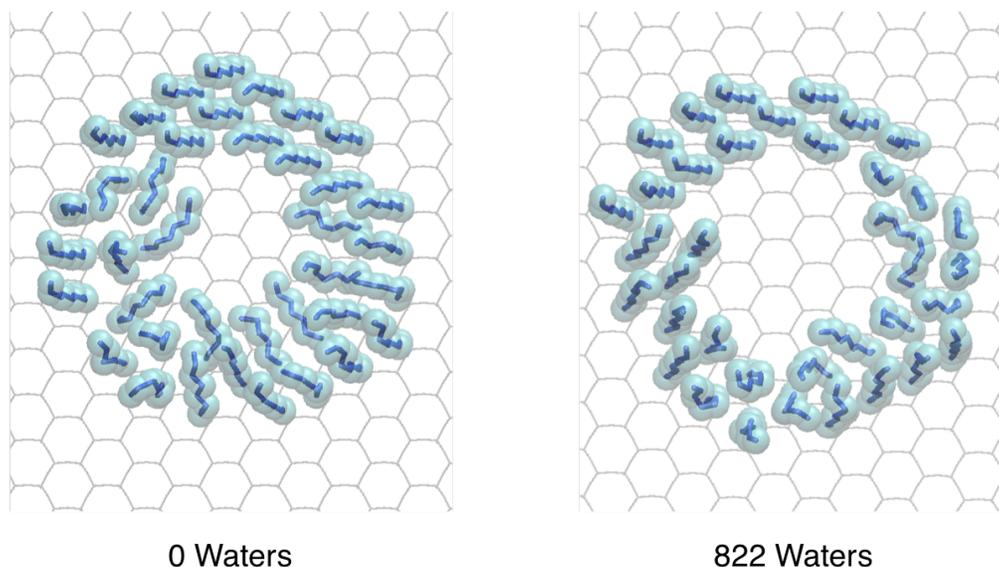


Figure 7.11. Visual representation of SAM ordering in the presence of water penetration. Shown are siloxane molecules near the defect site from the largest defective SAM system studied made up of nineteen missing siloxane molecules (SAM19). The architecture of the SiO_2 substrate is shown for reference. On the left is the equilibrated dry SAM system. Molecules near the defect site attempt to cover the defect site by changing conformation to cover as much of the hydrophilic substrate as possible. On the right is the same defect equilibrated in the presence of 822 water molecules. The water moves into the defect and forces the siloxane molecules to stand straight, perpendicular to the surface to expose as much of the hydrophilic substrate as possible.

hydrogen bonded molecules that resemble bulk water) will increase faster than the surface of the droplet (which corresponds to the poorly hydrogen bonded water molecules). Previous studies that attempted to characterize the behavior of water at the surface of SAMs have not been able to generate good agreement between experimental and computational conclusions (338). However, with the computationally derived IR spectra to compare to the experimental spectra, a more direct comparison can now be made than could be done in the past.

A MD-generated IR spectra of each of the 28 simulations described herein were analyzed through curve fitting to the two identified separate peaks corresponding to both structured and amorphous waters to determine the relative populations of each water species in the simulations (see Methods). Figure 7.3 shows an example set of spectra taken from the SAM7 system for all four hydration states. The assignments of the two peaks identified as well hydrogen-bonded and

poorly hydrogen-bonded were made from a deconstruction of both a bulk water system (Figure 7.4) and a deconstruction of a microhydrated SAM with multiple defects (Figure S4). As seen in Figure 7.12, the relative contribution of the $\sim 3500\text{ cm}^{-1}$ band, attributed to the population of waters that are well hydrogen bonded (i.e. 3-4 hydrogen bonds per water molecule or structured waters), increases with the addition of water into the system whereas the relative contribution of the $\sim 3300\text{ cm}^{-1}$ band corresponding to the population of the less well ordered (amorphous) waters is decreasing. This trend is not dependent on the presence or absence of defects, nor is it

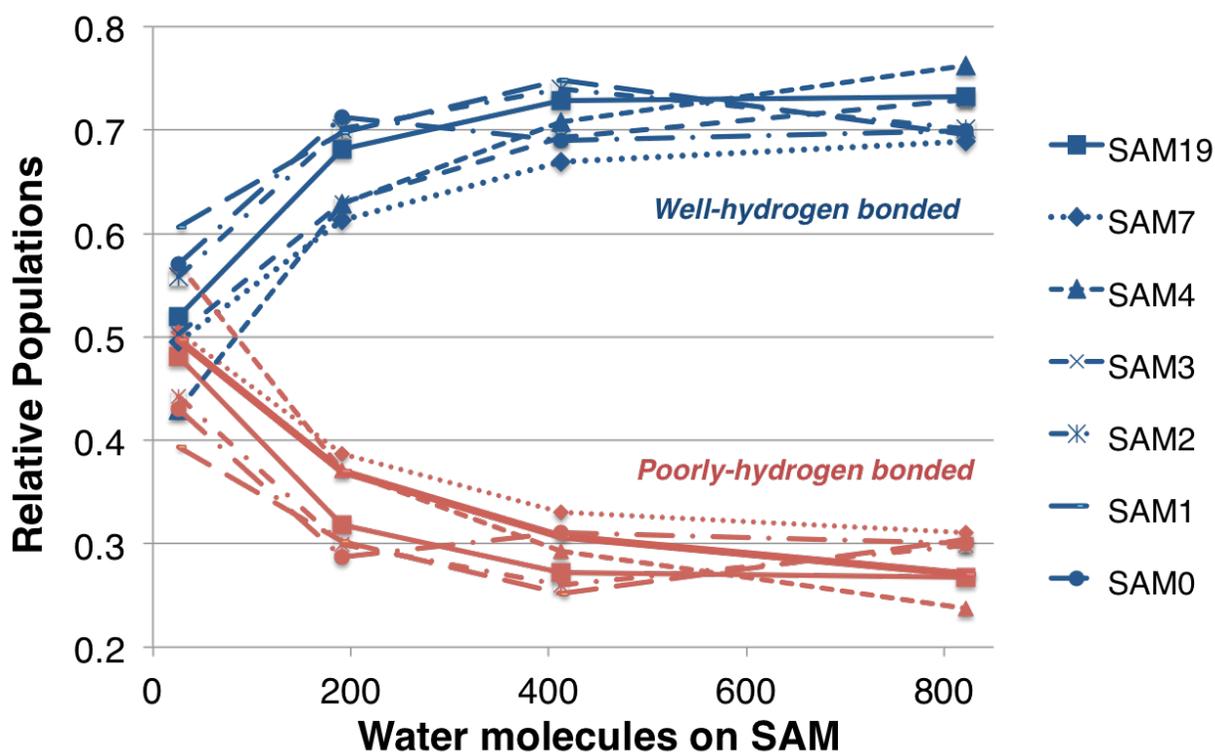


Figure 7.12. MD produced relative population shift in the IR spectra of water in the SAM. IR spectra corresponding to the O–H stretch generated from the MD simulations of all SAM systems in the present studied were analyzed using curve fitting to the two identified peaks that were found to correspond to well-hydrogen bonded (3-4 hydrogen bonds per water molecule) and poorly-hydrogen bonded (less than 3 hydrogen bonds per molecule) to determine the relative populations of each species present in the spectra similar to what was done experimentally in Figure 3. For all systems, as more water is introduced into the system the relative population of water that are well-hydrogen bonded increases while the relative population that is poorly hydrogen bonded decreases, agreeing with new experimental trend observed in Figure 3D for the C8= system on a SiO₂ crystal slab.

dependent on the size or shape of the defect. In all simulations, the water added to the SAMs forms a single droplet, with a select few water molecules entering the defect site if water penetration into the SAM occurs. As more water is added to the simulation, the volume of the water droplet, which is made of the well hydrogen bonded species, increases faster than the surface of the droplet which is made up of the less-hydrogen bonded species. The water molecules inside the defect, which are poorly hydrogen bonded, makes up only a fraction of the signal for the overall IR spectra, as the droplet dominates the IR signal in all systems. These results agree with the experimental results presented above that suggest that structured water contribution which would be more likely to have 3-4 hydrogen bonds per molecule grows faster than the amorphous water that would be poorly hydrogen bonded by comparison.

Taken together, this study sheds light on water adsorption on the surface of defective SAMs and the affects that even low levels of water penetration has on the SAM surface. The simulations herein show that at defects of size corresponding to three missing organic chains or larger, the SAM coating is no longer able to prevent water from contact with the substrate through the defects while in the SAMs with smaller defects the surrounding organic chains are able to reorient to cover the defect, thus preventing the water from penetrating to the substrate through a “healing effect” of SAM molecule rearrangements near the defect site as has been seen in previous studies (292, 293, 372). As the defect grows larger, more and more water can enter between the SAM chains, starting with transient penetration of individual water molecules through the occurrence of water-wire-like structures, until the defects grow to a sufficient size for the penetrating water molecules to begin to exhibit bulk-like characteristics. Both experiment and simulation show that the presence of water has an ordering effect on the SAM, which the MD simulations reveal to be correlated with both the size of the defects and the degree of water

penetration in the SAM itself. Combining these two results could be the basis for a simple test of SAM surface quality, as a SAM that has defects larger than two adjoining molecules would show such an ordering effect in the presence of water while the MD simulations suggests that ideal SAM surfaces or SAM with very small defects (one or two missing chains) would not. In addition, the differences observed for water penetration for the double bonded terminal siloxane molecules observed here and the fully saturated organic chains of Nishino et al suggest that this water behavior may also be dependent on the head group of the SAM and chain length, which should provide ample ground for further research into the complex behavior of water on organic SAM surfaces.

Chapter 8

Wetting Effects and Droplet Formation on Rough Hydrophobic Self-Assembled Monolayers

Summary

Studying the behavior of water on organic surfaces is important to our understanding of both water behavior on building sides and windows in urban areas as well as on organic dust particles in the atmosphere. Self-assembled monolayers have been used as a model system to study the behavior of water in these circumstances focusing on the dependence of wetting properties on SAM characteristics. Of particular interest has been the effect of mixing different organic molecules to form a rough SAM surface that models suggest may affect water uptake when compared to smooth counterparts. While previous coarse grained work on rough hydrophobic SAM surfaces suggest that droplet formation may be inhibited, recent atomistic simulations of similar systems have shown that these effects may not exist. To study this phenomenon closer, we carried out atomistic molecular dynamics simulations of ten SAMs with varying ratios of octadecyl trihydroxylsilane and dodecyl trihydroxylsilane (C18 and C22 chains) including two smooth surfaces as controls at five different hydration levels, matching the circumstances in which droplet inhibition had been observed. Unlike the original studies, we observe droplet formation on both smooth and all rough surfaces. Visualization and comparison suggests that previous studies may not have achieved equilibration and observation of droplet formation before simulations were terminated. In these systems, water has a minimal ability to affect the SAM surface. Further analysis of the orientation of water-water hydrogen bonds near

the SAM surface reveals the existence of four distinct structured layers of water and their on both the ratio of the individual SAM molecules and their placement.

Introduction

The behavior of water on organic surfaces is both complex and important to a variety of fields of biology and chemistry. These organic surfaces can be constructed from a variety of compounds and can be either hydrophobic or hydrophilic depending on the chemical makeup of the surface itself. As has been demonstrated by past studies, even small changes in the surface can lead to drastically different water behavior (292, 372). Such interactions and complex behaviors play a vital role in the atmosphere where water can interact with aerosol organic particles in an addition to variety of other organic surfaces such as vegetation, windows and buildings (289, 290, 361, 386). The presence of water on surfaces such at these has been shown to affect a variety of reactions and properties of the surface (286, 293, 302, 311, 387). Due to the importance of understanding the behavior of water on these organic surfaces, there has been a large body of work done to study the interfacial water chemistry including studies on glass, lipid membranes, metals and quartz (286, 295, 362-367, 369-371, 388).

One particular surface that has been invaluable in our development of a deeper understanding of water on organic surfaces are Self-Assembled Monolayers (or SAMs). SAMs generally consist of simple organic molecules that can spontaneously form a thin layer over a reactive substrate such as gold or SiO₂ through a reactive head group (312, 375, 376). The surface environment formed by the creation of these SAM surfaces can be tuned by choice of the interface exposed terminal group to alter hydrophobicity and further tuned by the length of the organic molecules which is normally a saturated hydrocarbon chain (351, 355). Due to their

simple repeating structure, they have been well suited for study using molecular dynamics and Monte Carlo simulations that have probed water uptake and behavior at a variety of surfaces (335, 336, 358, 363, 373, 374, 389, 390).

SAM surfaces can also be formed through the mixture of different organic molecules. One of the most common mixed SAM surfaces is made up of a mix of two organic chains of similar relativities but differing chain lengths. By using organic molecules that differ in only the length of a saturated hydrocarbon chain, a rough surface can be generated which is believed to be more hydrophobic than a simple smooth one component system. Such systems have been studied in a variety of different studies (309, 374, 391). In particular, Rudich and coworkers used coarse grained molecular dynamics to simulate varying amounts of water on the surface both of a smooth octadecyl trihydroxylsilane (C18) and a rough 50/50 mixture of C18 and dodecyl trihydroxylsilane (C22) molecules (309). They noted that while the water formed a droplet on the smooth hydrophobic surface, no such droplet formation was observed on a rough SAM when exposed to equal hydration levels. In a later study using shorter organic chains and a smaller, two $-CH_2$ length difference, Szőri and coworkers found no such behavior under similar environments using Monte Carlo simulations (351). This new finding, which suggests that further work is needed to elucidate what occurs at the surface of these rough SAMs, has lead us to reevaluate the C18/C22 mixture in more detail to attempt to reconcile the contrasting results.

What we present here is a series of atomistic molecular dynamics simulations of ten SAM systems consisting of varying ratios of C18 and C22 molecules similar to those carried out by Rudich et. al (309) including four different 50/50 C18/C22 systems to reproduce the original work. Five different hydration levels are applied to each dry-equilibrated system before 10 nanosecond simulations were carried out and analyzed. Droplet formation is observed for all

SAM surfaces at all hydration levels with the exception of those surfaces exposed to bulk water where droplet formation is inherently inhibited. Visualization of the droplet formation on rough systems suggest that the previous simulations may not have been long enough to achieve droplet formation. It is observed that water has minimal ability to affect the SAM surfaces, however it is observed that the SAM can have affect the local water structure at the hydrophobic surface. Finally, analysis of the orientation of water-water hydrogen-bonds near the SAM surface reveals four distinct structured layers of water that show a dependence on the component makeup of the SAM system.

Methods

An ordered lattice comprised of alpha-tridymite SiO_2 crystal was used to model the crystal substrate for all systems studied herein, identical in size (78.66 x 75.60 x 14.00 Å) to those used in previous studies (293). The SAM itself was constructed from 270 individual siloxane molecules. The two individual species of siloxane molecules used in this study were octadecyl trihydroxylsilane [$\text{CH}_3(\text{CH}_2)_{17}\text{Si}(\text{OH})_3$] henceforth referred to as C18, and dodecyl trihydroxylsilane [$\text{CH}_3(\text{CH}_2)_{11}\text{Si}(\text{OH})_3$], referred to as C12. Siloxanes were covalently bonded to one the exposed Si atoms on one side of the SiO_2 crystal slab, while half of the surface exposed Si atoms on the opposite side were removed to maintain charge neutrality. A 1.0 kcal/mol/Å² constraint was placed on all substrate atoms as well as the terminal Si-O atoms that were bonded to the substrate. A reflective wall was placed at $z = 90$ Å (where z is normal to the SAM surface) to prevent waters from interacting with the unsaturated periodic image of the substrate, with a layer of vacuum (30 Å) large enough to cause interactions between water and the substrate periodic image to be negligible. An external electric field was applied to the system to correct for

the model's inherent anisotropic nature (see chapter seven for a more detailed description). Force fields for the siloxane molecules were adapted from CHARMM27 force fields as has been done in previous studies (293). The SiO₂ parameters were taken from the work of Stevens and coworkers (380) while the water molecules were modeled using the TIP3P water model (392).

Ten different SAM systems were generated by varying the ratio of C18 and C22 molecules present in each system (Figure 8.1) as follows. To begin, a system consisting of C18 molecules was generated, referred to as 100-C18, where the number denotes the percentage of the system made up of C18 molecules. Next, a given percentage of the C18 molecules were randomly selected to be switched with C22 molecules through a custom C script to generate 90%, 75%, 25% and 10% C18 systems. The process was repeated to generate three different 50% C18 systems, denoted as A, B, and C in Figure 8.1. A perfectly patterned 50% system with perfect channels for possible water uptake was then generated (50D-C18 in Figure 8.1). While this system may be unrealistic, it does represent the best possible chance to observe water uptake on a mixed SAM system. Finally, a C22 only system was also generated (0-C18 in Figure 8.1) to mirror the C18 only system.

Each system was initially equilibrated in the absence of water for at least 10 nanoseconds. Following the initial runs, the equilibrated 0-C18 system was hydrated with bulk water by the solvate tool in the VMD program (179) from the surface of the SAM to reflective wall at $z = 90 \text{ \AA}$ and equilibrated for 10 additional nanoseconds. From this equilibrated hydrated system, four different water coverages were created by selecting only those water molecules that were 1, 2, 4 and 8 \AA from the SAM surface. This process resulted in the creation of water coverages that corresponded to 259, 508, 891 (approximately one monolayer of coverage) and

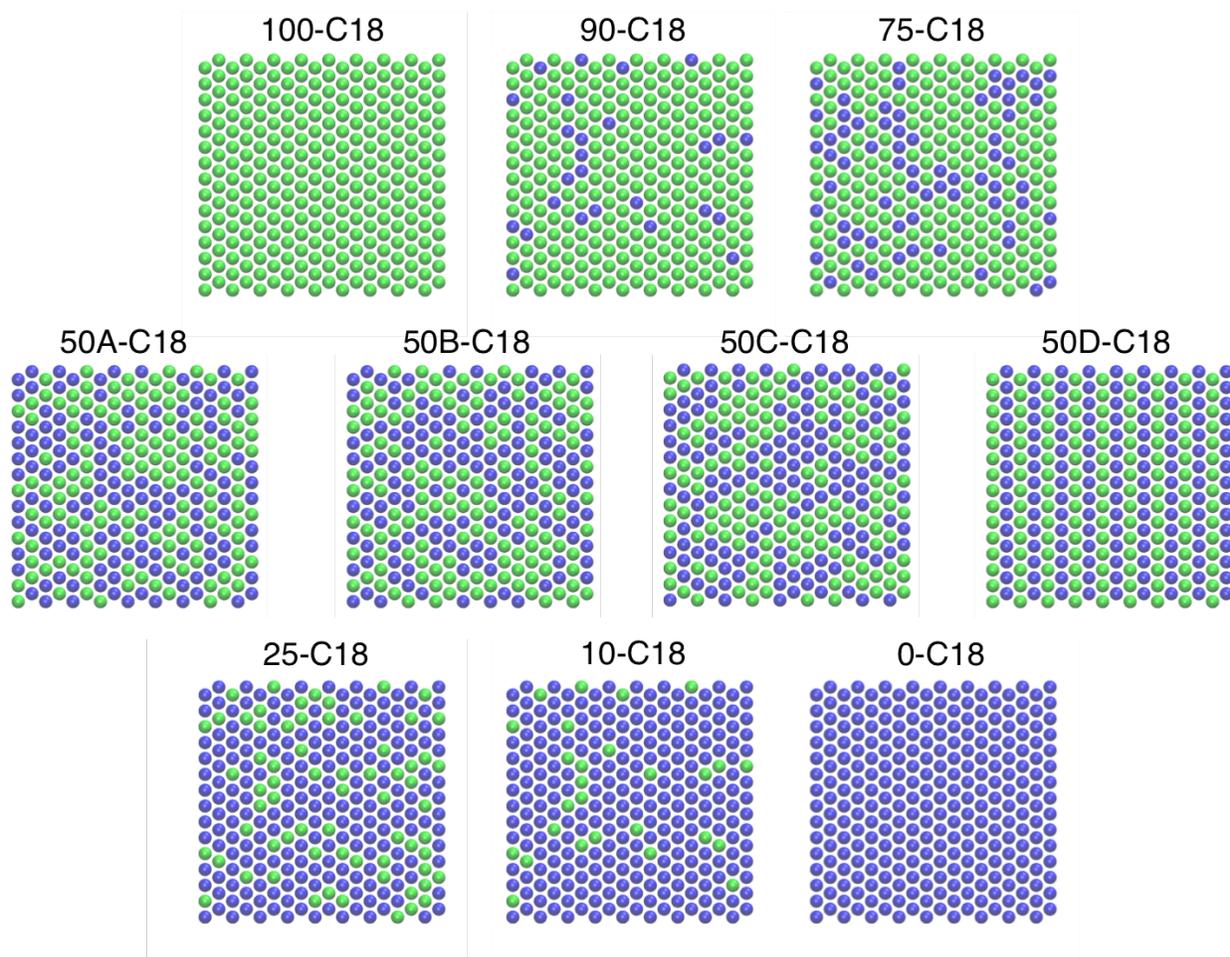


Figure 8.1. Schematic of all simulated mixed SAM systems. The ten SAM surfaces that were built for this study are shown in schematic form, showing the Si atom of each siloxane molecule to display the placement of each organic chain in the SAM. Octadecyl trihydroxylsilane or C18 molecules are shown as blue spheres and dodecyl trihydroxylsilane or C22 molecules as green spheres. Systems were named using the percentage of C18 molecules that comprised each system (i.e. 90-C18 is a system made of 90% C18 molecules and 10% C22 molecules), while a letter was added to distinguish between the four different 50-C18 systems.

1,734 water molecules (or two monolayers). Each of these water coverages were added onto the surfaces of each SAM system, as well as the equilibrated bulk water from the 0-C18 system (8,753 water molecules) which acted as the fifth water coverage studied herein. A sixth water coverage was generated for the 50D-C18 system only where the smallest equilibrated water coverage of 259 molecules were moved entirely into the grooves of the patterned system before

being equilibrated for 10 nanoseconds to study an extreme case of possible water uptake on mixed SAMs.

Simulations were carried out using the NAMD 2.7 software package in the isochoric-isothermal (NVT) ensemble (234). A Langevin thermostat with a 1ps coupling coefficient was used to keep the system at 300K. The short-range interactions were truncated at a distance of 20 Å. The classical equations of motion were integrated by the use of the Verlet algorithm with a timestep of 1 fs. All bonds with hydrogen were held constant through the use of the SHAKE and SETTLE algorithms (382). An external electric field of $-4.53786 \text{ kcal}/(\text{mol } \text{Å } e)$ was used to correct for the lack of z-periodicity in the system (see above and chapter seven for more details). Analysis was carried out using VMD (179) and locally developed analysis tools.

Results and Discussion

Molecular dynamics simulations of mixed SAMs

In order to study the wetting effects on mixed hydrophobic SAM surfaces and possible relationships to the make up of the monolayer with respect to its siloxane molecules, ten different SAM surfaces were constructed from 270 organic chain molecules of varying amounts of octadecyl trihydroxylsilane (or C18) and dodecyl trihydroxylsilane (C22) as seen in Figure 8.1. Each system was equilibrated in the absence of water for 10 nanoseconds before the addition of one of five different hydration levels consisting of 259, 508, 891, 1,734 and bulk water and an additional ten nanoseconds of simulation (see Methods for a more detailed description of each system).

Density profiles of each system at every simulated level of hydration as displayed in Figure 8.2 reveal several key characteristics while the terminal carbon profiles of each system at

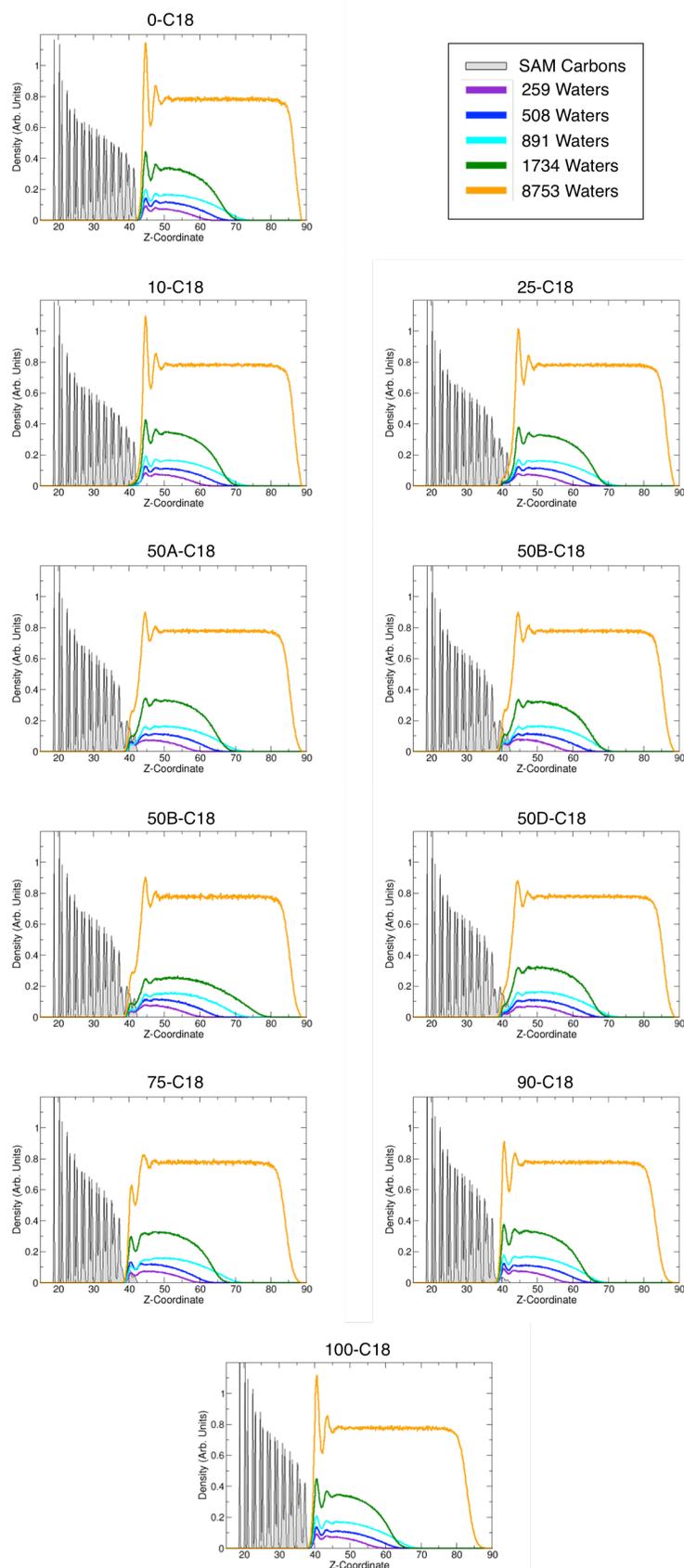


Figure 8.2. Density profiles of different hydration levels for mixed SAM systems. Density profiles are shown for each of the 10 SAM systems simulated in this study as named by Figure 1. As the SAM carbon profiles were identical for each hydration level, the profile at the lowest hydration level is displayed as gray. The water profiles (normalized by the number of water molecules present in each hydration level) that are shown correspond to 259 waters (purple), 508 waters (blue), 891 waters (cyan), 1734 waters (green) and 8753 waters or bulk water (orange).

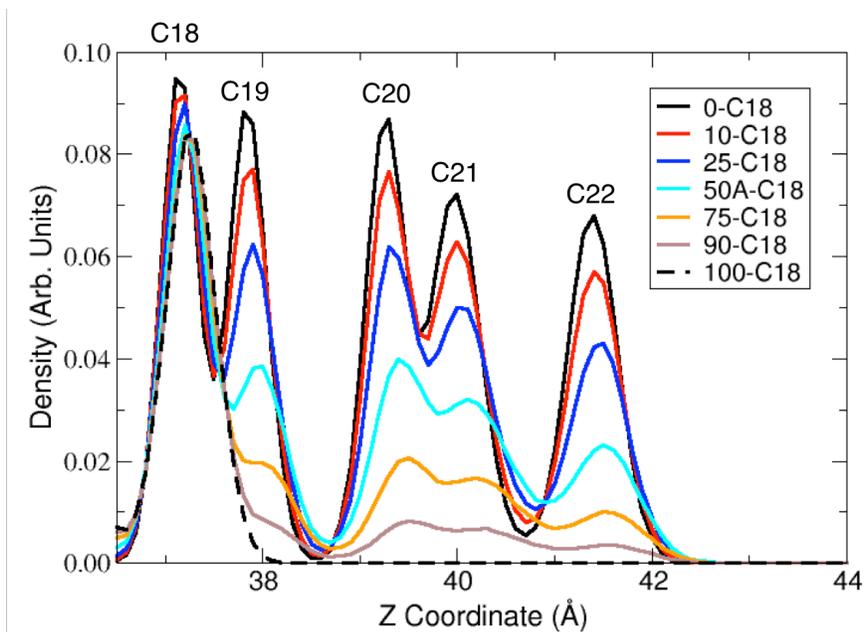


Figure 8.3. SAM ordering of the terminal carbons of the C22 chains. The density profiles of seven of the ten SAM surfaces are overlaid showing the broadening of the density of the terminal four carbons of the C22 chains with the decreasing percentage of C22 chains in the SAM surface at the 259 molecule water coverage. The peaks of the C18 to C22 atoms are labeled. As the profiles for 50A-, 50B-, 50C- and 50D-C18 systems overlapped, only one (50A-C18) is displayed.

the 259 molecule water coverage are displayed in Figure 8.3. As seen in chapter seven, the density profiles of the SAM carbons do not change in the presence or absence of any amount of water. Because the profiles of the carbons for each hydration level was a near perfect match for all other hydration levels for each system, only the 259 water SAM carbon profile is displayed. However, while the profiles were in internal agreement with each system, the density of the nineteenth through twenty-second carbon atoms of the C22 molecules does change significantly depending on the makeup of the system (Figure 8.2 and Figure 8.3). The density of the final carbon molecules broadens with the addition of the shorter C18 siloxanes as the final four carbons are allowed more conformational freedom by the absence of the tightly packed SAM around them. This finding matched what was observed by Rudich and coworkers in previous studies of mixed SAM systems.

Focusing on the water profiles in Figure 8.2, the water in each simulation below the bulk run formed a droplet on the SAM surface. The profile for each system shows the characteristics of water near a hydrophobic surface, with a peak near the surface representing an ordered layer of water resembling that of the expected radial distribution function of an equilibrated liquid before the density extends away from the surface for the bulk of the droplet that formed on each SAM surface. This is in contrast with Rudich et. al which noted that while droplets formed on the smooth systems in a one nanosecond simulation, no such formation was observed in the mixed 50-50 system at low hydration levels. It was possible that the droplet formed much later in our simulations, as our simulations were ten times as long as the previous study. However, visualization of the droplet formation shows that formation occurs within the first nanosecond of each simulation. This formation process and time frame is highlighted in Figure 8.4 which tracks the formation of a droplet formation on the 50A-C18 system in the presence of approximately one monolayer of water, reproducing the coverage from Rudich et al. It is interesting to note that the water structures formed on the mixed SAM at 100ps in our simulation resembles the structures believed to be the equilibrated water in the previous study (seen in the inlays of Figure 3). In addition, it is interesting to note that the density profiles of the droplets on each hydrated SAM surface in Figure 2 bear a close resemblance to the water profile that was observed on not only the smooth surface but also the rough surface which suggests that droplet formation may have occurred by the time that the water profile was generated in the previous work.

Three different random 50/50 C18/C22 mixtures (50A, 50B and 50C) as well as a perfectly patterned 50/50 mixture (50D) were studied in order to further probe this disagreement (Figures 1 & 2). At all hydration levels, water droplet formation is observed in the first nanosecond of simulation and remained stable for the extent of each simulation as observed by

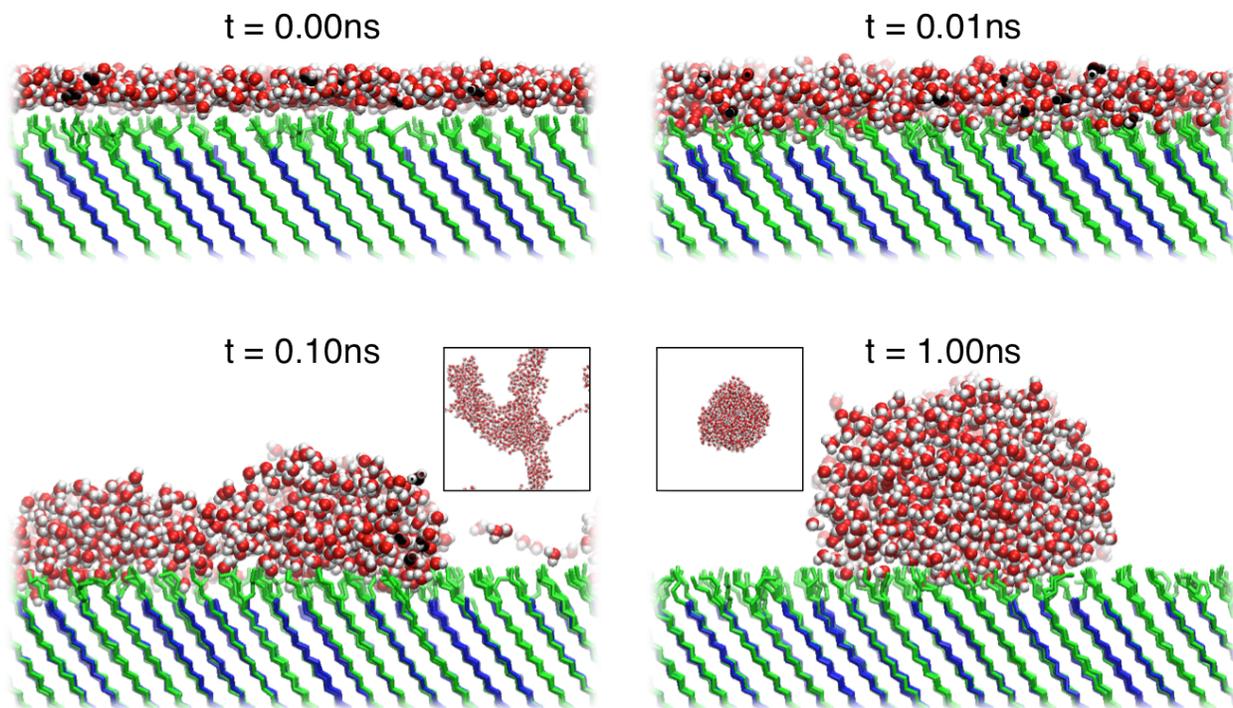


Figure 8.4. Droplet formation on the 50A-C18 SAM. Time-lapse snapshots tracing the formation of a droplet on the 50A-C18 system in the presence of approximately one monolayer of water, reproducing the process shown in Figure 9 of Rudich et al. At $t = 0$ nanoseconds, a monolayer of water that was equilibrated on the surface of a pure C22 SAM is placed on the surface of the hydrophobic surface. After 1 picosecond, the water begins to penetrate the C22 surface level. After 100 picoseconds, the water forms patterns similar to those seen by Rudich et al. as the droplet continues to form on the mixed surface. At $t = 1.0$ nanosecond, the droplet has formed on the surface of the SAM, while waters at the bottom of the droplet are still able to penetrate the C22 surface. The timescale and process of droplet formation is similar in all simulated systems.

the density profiles. In order to see if droplet formation could be prevented under ideal circumstances, an additional simulation of the patterned 50D-C18 system was generated with the smallest hydration level of 259 waters positioned within the C18 channels between C22 siloxane rows. Even in this idealized environment droplet formation was observed in the beginning of the simulation (Figure S2), providing further evidence that water will in fact form a droplet on any mixed system and not remain largely in contact with the hydrophobic surface.

Finally, the RMS surface roughness of each SAM system was calculated for the terminal SAM carbons as a measurement of SAM roughness, shown in Figure 8.6. The RMS roughness

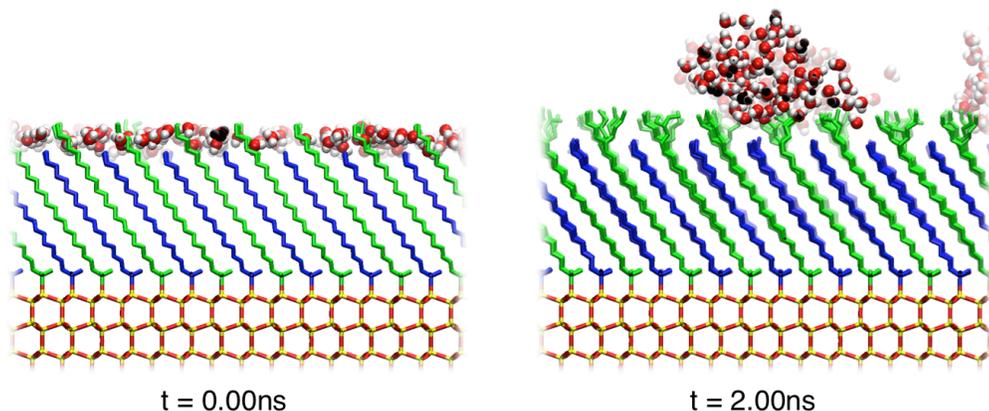


Figure 8.5. Droplet formation with buried water in the patterned 50D-C18 SAM. An additional three nanosecond simulation was carried out of the patterned 50D-C18 system with the lowest hydration level of 259 waters moved into the patterned channels between rows of C22 molecules. After two nanoseconds, the waters had exited the channel and formed a droplet on the surface of the rough system.

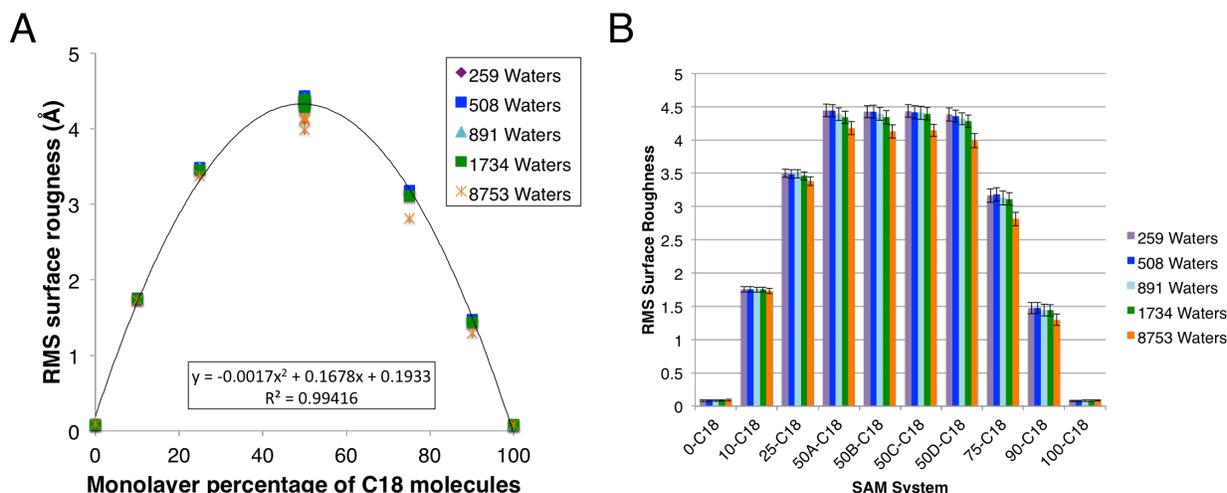


Figure 8.6. RMS roughness of SAM surfaces. Calculations of the average (root mean square) RMS over the final equilibrated nanosecond of the terminal carbon atoms of each SAM surface as a measure of surface roughness. A) A trend line was fit to the RMS measurements and is displayed along with the equation and the corresponding R^2 value. B) Error bars are overlaid on each system representing one standard deviation over the same time period. While each of the four 50-C18 systems are not statistically different from each other and the smooth systems are not either, all other systems show a statistically significant difference in the surface roughness. The only cases where a single hydration level displayed a significant difference from all other hydration levels were the 50C-C18, 50D-C18 and 75-C18 8,753 waters (bulk water).

can easily be fit to a parabola with good statistical agreement ($R^2 = 0.99416$). There is a statistically significant difference in the roughness in a system consisting of 25% C18 and 75% C22 molecules when compared to one consisting of 75% C18 and C22 (Figure 8.6). The RMS roughness of the 75-C18 and 90-C18 are lower than that of the 25-C18 and 10-C18 respectively as the water present in each system are able to penetrate into the SAM to occupy the progressively smaller gaps between the C22 chains and act as the missing four carbons of the siloxane molecule to increase the order of the system and reduce the surface roughness. The 8,753 water run of the 50C-, 50D and 75-C18 shows the extreme case of this healing effect, as these runs are the only cases where a the roughness of a single hydration level was statistically different than all other hydration levels for a given system. Focusing on the four 50-C18 systems, while the patterned 50D-C18 system does show lower RMS roughness values than any of the random surfaces, the differences are not statistically significant.

Water penetration on mixed SAMs

Analysis of the water on the surfaces of each system reveals a sharp contrast in the water penetration between smooth and rough systems. In smooth systems (0-C18 and 100-C18), the water at every hydration level never penetrates into the SAM surface, forming a hydrophobic layer of water at the bottom of a droplet on the smooth hydrophobic surface (Figure 8.7A). This is in good agreement with previous studies of smooth surfaces of saturated and double bonded terminated organic chains (293).

In each of the eight mixed systems, water penetrates into the SAM surface to reach the C18 surface regardless of the SAM makeup (Figure 8.8), with increasing water penetration to the C18 surface with the increasing proportion of the SAM made of C18 chains. The degree of water

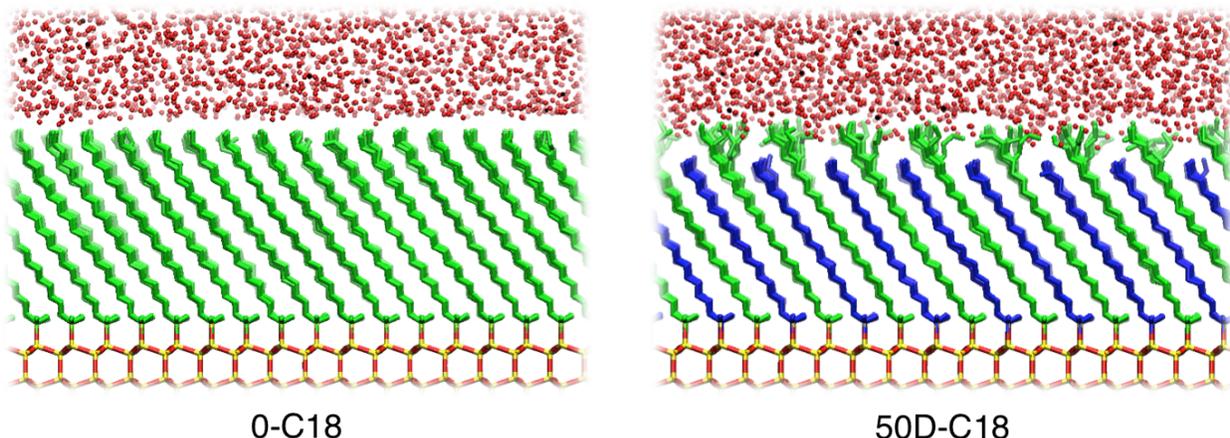


Figure 8.7. Water uptake on smooth and rough SAM systems. Snapshots from the end of the simulations of a representative smooth SAM surface (0-C18) and a rough mixed SAM (50D-C18) display the differences in water uptake in each system. C22 molecules are shown as green sticks, C18 as blue sticks, the SiO₂ slab as sticks colored yellow for Si and red for O, and water molecules as small red spheres. In the smooth surface, water molecules form an even layer of water at the hydrophobic SAM surface. In the mixed SAM, despite the extra flexibility of the final four carbons of the C22 molecules that can reorient to cover the channels formed by the neighboring C18 molecules, water molecules are still able to intercalate between the C22 molecules to reach the C18 surface of the SAM.

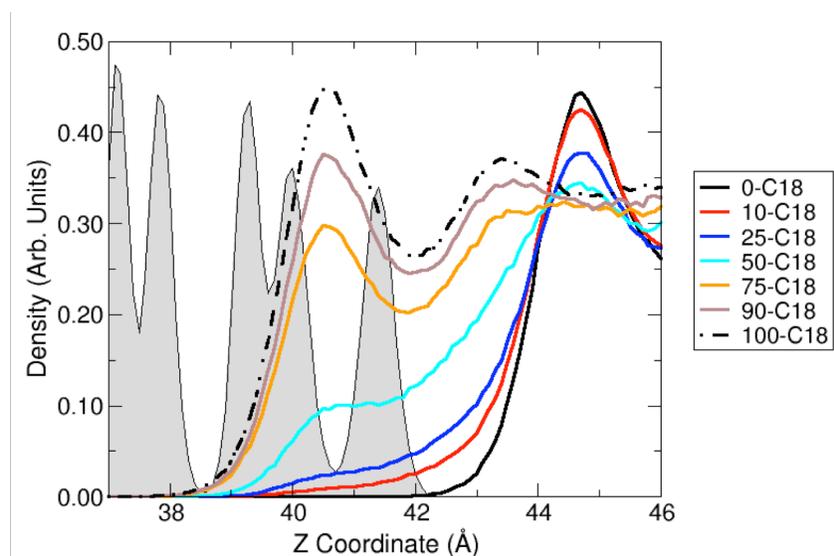


Figure 8.8. Water penetration as a function of percentage of C18 in mixed SAMs. Density profiles of the 1734 water coverage are shown for each of the ten systems studied as described in the legend. The profile for the 0-C18 SAM is shown for reference of the final carbon locations. Water penetration through the C22 surface is observed in all systems that contain C18 molecules.

uptake observed in each of the four 50/50 C18/C22 systems is in good agreement with the pioneering work done by Rudich and coworkers. The water molecules that reach the C18 surface in these mixed systems are still hydrogen-bonded to the water at the C22 surface as the mixed

system is unable to support isolated water uptake into the SAM surface at the hydration levels studied here. Analysis of the number of hydrogen bonds per molecule as a function of the z-axis also confirms that water penetration to the C18 level occurs in every system with C18 molecules present (Figure 8.9). Figure 8.7B visualizes the water penetration on the patterned 50D-C18 system in contrast to the smooth 0-C18 surface of Figure 8.7A. The water molecules are able to reach the C18 level despite the relatively high mobility of the terminal carbons of the C22 molecules to flex and cover the shallow defect in the surface, similar to the healing effect observed previously in the literature (292, 293, 344).

Water structure on mixed SAM systems

In order to characterize the water structure at the surface of the smooth and mixed surfaces, the final equilibrated two nanoseconds of each simulation was analyzed by calculating the angle formed between each water-water hydrogen bond and a vector normal to the SAM surface. This data was broken down by the location along the z-axis of the center of the hydrogen bond into bins of 0.5 Å in the z-axis. The angle formed between the hydrogen bond and the SAM surface normal vector was separated into 10 bin with all angles translated to a value of between 0 and 180 degrees. Each z dataset was normalized by dividing by the total number of hydrogen-bonds in all 10 angle bins for a given z-slab of 0.5 Å. The full set of these intensity maps for all systems at each of the five hydration levels can be found in Figure 8.11 and a detailed example shown in Figure 7A, which was taken from the 25-C18 SAM exposed to 1,734 water molecules. Analysis of the intensity maps from each simulation reveals four distinct layers of water present at the SAM surface as defined by the presence of a preferred Hydrogen-bond orientation, designated as A, B, C and D in Figure 7. Each of the four structured water layers is present at

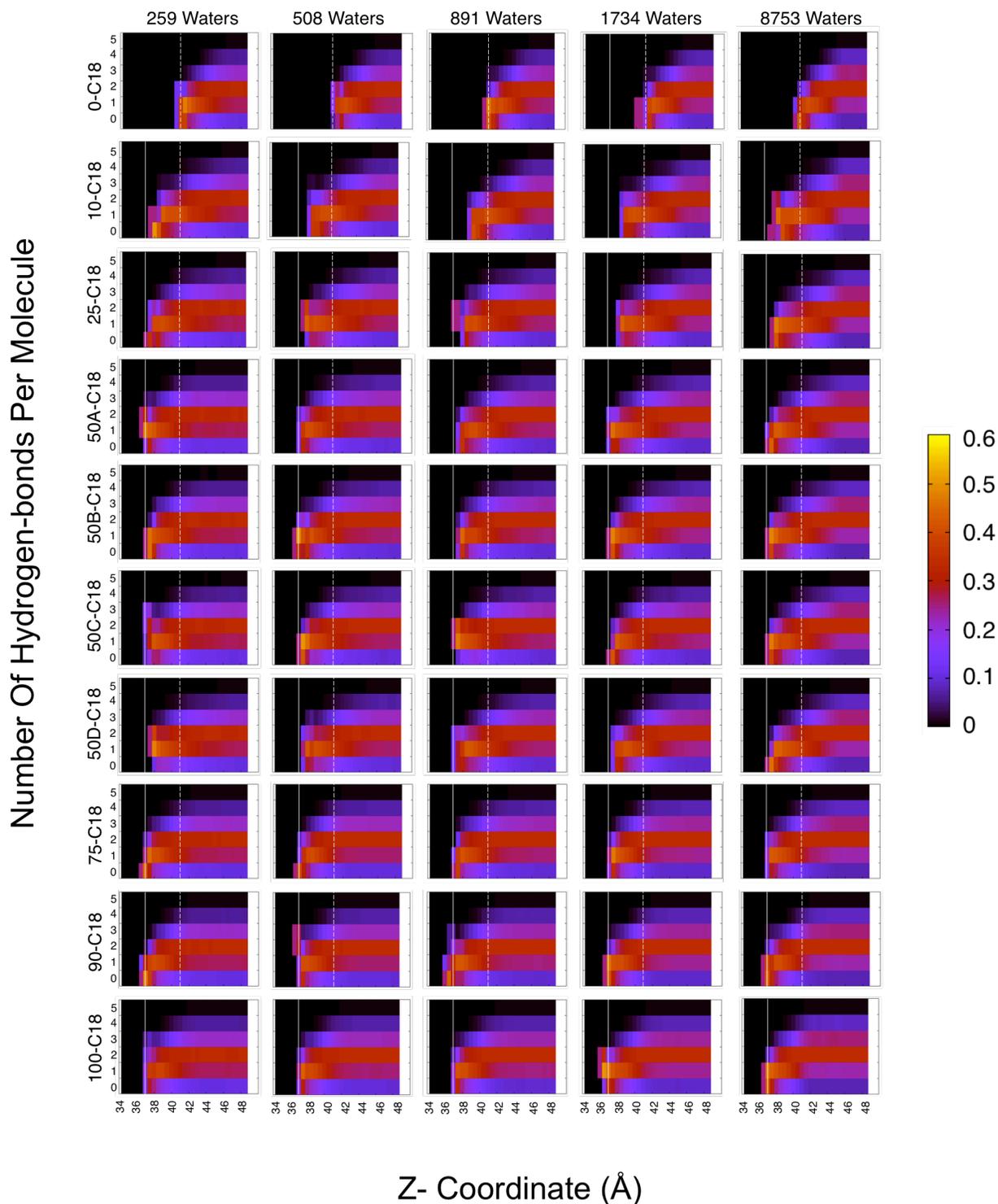


Figure 8.9. Distribution of hydrogen-bonds per molecule as a function of the z-axis. Similar to the analysis in chapter seven, the final two nanoseconds of each simulation was analyzed to generate a histogram of the number of number of hydrogen bonds per water molecule as a function of the Z-axis (normal to the SAM surface). The surface region of the SAMs are displayed. The solid white lines in each plot approximate the C18 surface while the dashed lines do the same for the C22 surface.

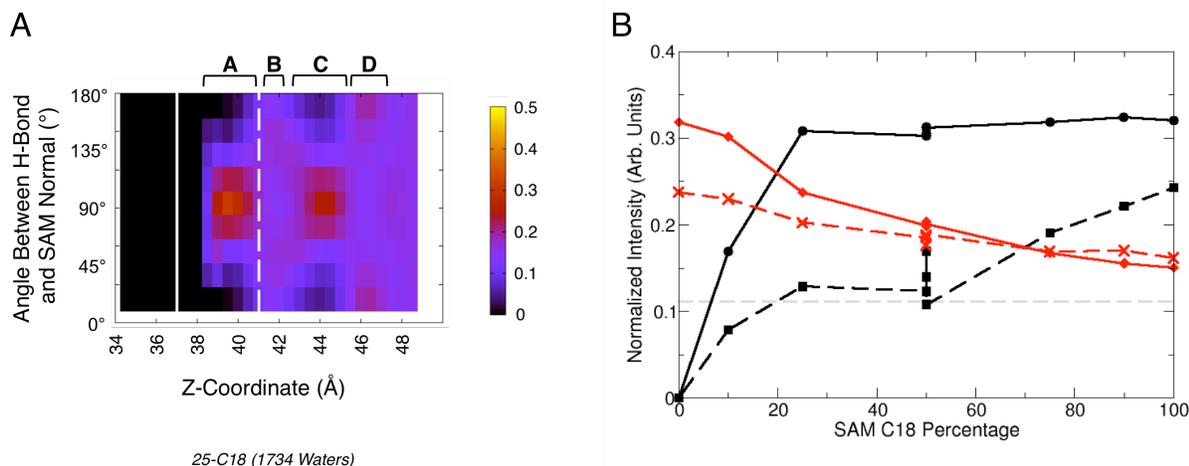


Figure 8.10. Hydrogen-bond orientation distributions as a function of the z-axis. A) An intensity map of the angle formed between all water-water hydrogen bonds and a vector normal to the SAM surface (see Results and Discussion for more details on generation of the data). The solid white line approximates the C18 surface while the dashed white line approximates the C22 surface. Higher intensity corresponds to a preference for the orientation of water-water hydrogen bonds present at that Z-coordinate. Four identified structured layers of water correspond to the letters A, B, C and D at the top of the map. While layer B is not noticeable in this figure, it is noticeable in systems with higher C18 makeup in Figure S6. B) The normalized intensity of the preferred orientations of each structures water layer ($\sim 90^\circ$ for A and C, $\sim 180^\circ$ for B and D) were calculated at the position of each structured layer and plotted as a function of % of C18 molecules in the SAM surface. The grey dashed line corresponds to a perfectly distribution among H-bond orientation.

varying degrees on each mixed SAM surface, which was quantified by plotting the normalized relative intensity of the preferred orientation bin as a function of percentage of C18 chains in the SAM surface. As the location of the C18 and C22 surface is identical in all systems, the same z-bins of 39.5, 42, 44 and 46Å were used to analyze water layers A, B, C and D respectively. The trends observed for this hydrogen-bond quantification between the five hydration levels show good agreement as seen in Figure 8.11, while that of the 1734 water hydration level (corresponding to approximately two monolayers) is displayed in Figure 8.10B.

The first identified water layer (A) is located at the C18 surface and shows a clear preference to form a hydrophobic layer of water that places hydrogen-bonds parallel to the SAM surface. While absent in the 0-C18 system due to the lack of the C18 surface in the pure smooth

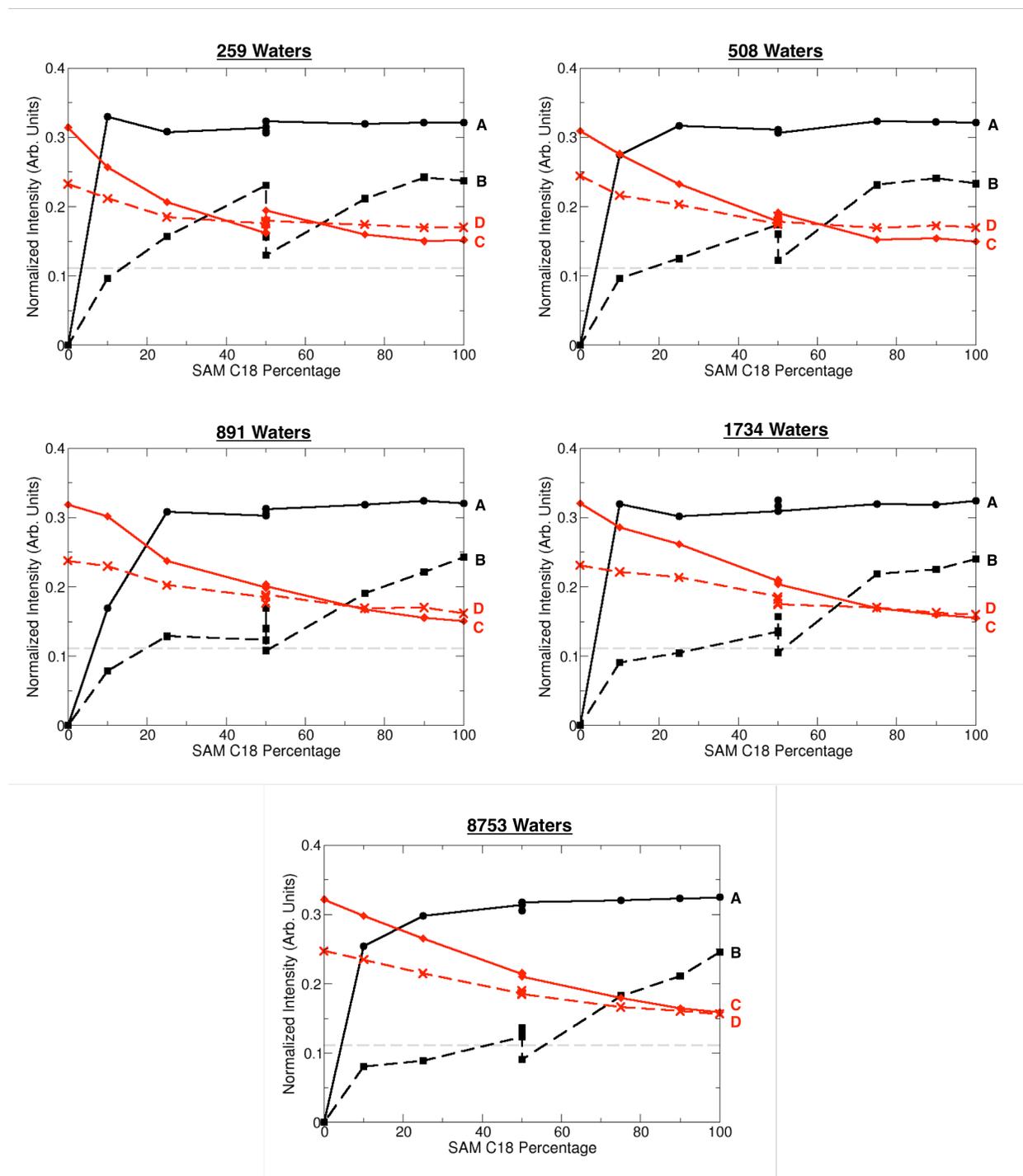


Figure 8.11. Trends for hydrogen-bond orientation preference for each of the five hydration levels. The normalized intensity of the preferred orientations of each structures water layer ($\sim 90^\circ$ for A and C, $\sim 180^\circ$ for B and D) were calculated at the position of each structured layer and plotted as a function of % of C18 molecules in the SAM surface for each of the five hydration levels analyzed in this study, similar to that shown in Figure 7B. The grey dashed line corresponds to a perfectly distribution among H-bond orientation.

C22 system, in all other studied systems the preferred parallel hydrogen-bond orientation produces the strongest signal for a structured layer of water amongst the four identified layers as seen in Figure 8.10B. With regards to the four different 50/50 systems, all equally mixed SAMs shows good agreement with similar preferences for hydrogen-bond orientations in the A layer.

The second structured water layer B is positioned above the hydrophobic layer of A, and in surfaces that are 75% or more C18 molecules show a preference to orient water-water hydrogen-bonds normal to the SAM surface. In systems that are made of 25% or less of C18 molecules, the normal orientation vanishes, replaced by a random orientation with no clear preference (Figures 8.10B and 8.12). Unlike the A layer, the four 50/50 systems can show a great deal of variation in the intensity of the B layer, with the patterned 50D-C18 system showing the weakest signal for this oriented water layer, suggesting a lower degree of structure. This result may suggest that this perfect patterning may be better at disrupting this layer of structured water than a random 50/50 system. In the 100-C18 system, only the A and B structured water layers are observed before bulk water and no preferred orientation dominates the intensity map.

The final two water layers, C and D, are identical to A and B but reside at the C22 surface instead of the C18 surface. Similarly, the preferred hydrogen-bond orientations of parallel and normal to the SAM surface for C and D respectively are highest when the C22 character of the SAM surface is highest and the preference fades to a random orientation in the presence of more C18 character. This random distribution with no preference for hydrogen-bond orientation corresponds to an equilibrated droplet where the water within the droplet will have no forces favoring one orientation with respect the SAM over another and is observed in every system studied herein. In the 0-C18 surface only the C and D structured water layers are observed as the A and B layers are dependent on the presence of a C18 SAM surface.

Angle Formed Between Hydrogen Bond And A Vector Normal To The SAM Surface

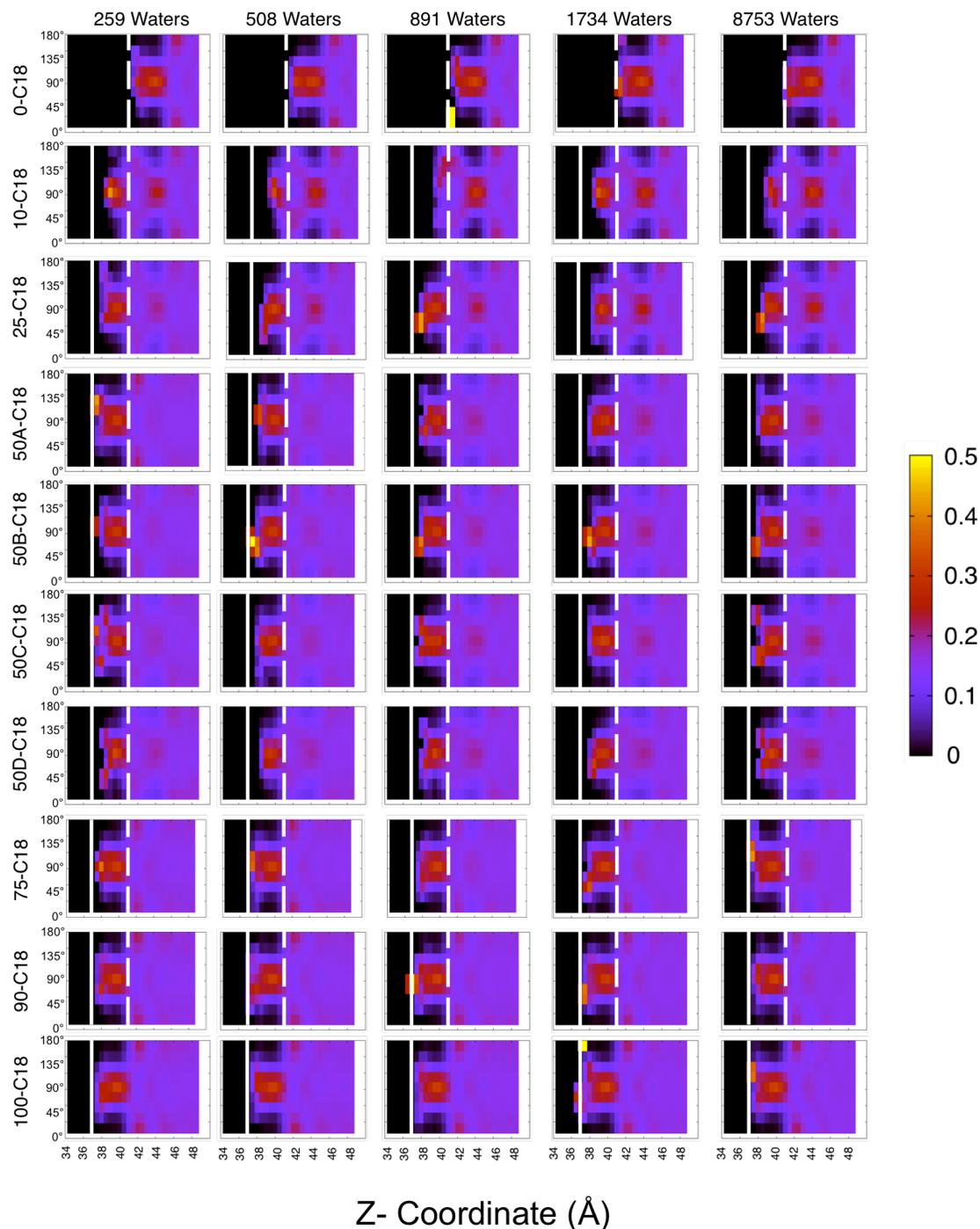


Figure 8.12. Full hydrogen-bond orientation distributions as a function of the z-axis. Intensity maps of the angle formed between all water-water hydrogen bonds and a vector normal to the SAM surface are shown for all simulated systems and hydration levels. The solid white lines in each plot approximates the C18 surface while the dashed white lines approximates the C22 surface. Higher intensity corresponds to a preference for the orientation of water-water hydrogen bonds present at that Z-coordinate. (See Results and Discussion for more details on the creation of the intensity maps.)

Conclusions

Molecular dynamics simulations of ten different systems made up of varying degrees of C18 and C22 molecules demonstrated that water on each of the surfaces studied will form droplets regardless of the hydration level on the surface (Figures 8.3 and 8.4). While previous coarse-grained work by Rudich and coworkers (309) on mixed C18/C22 systems suggested that droplet formation would not occur on 50/50 C18/C22 mixed systems, the similarity of their observed water structures on mixed system and those observed during the equilibration phase of our simulations in which droplet formation always occurs suggests that equilibration may not have been achieved for the previous simulations. Further supporting this idea of a lack of equilibration, droplet formation was observed at four different hydration levels on a large variety of surfaces; three differently random 50/50 systems, a perfectly patterned 50/50 system that maximizes the possibility of water uptake, as well as four additional C18/C22 systems made of 10, 25, 75 and 90% C18 chains. In each of these systems, the hydration level showed minimal effect on the SAM surface (Figures 8.2 and 8.6), suggesting that water has a minimal ability to affect the structure of the mixed hydrophobic SAMs.

In all of mixed systems, water penetration into the SAM to reach the C18 chain surface was observed at all hydration levels (Figures 8.7-9). In addition, four different structured layers of water were identified on the surface of mixed SAMs, corresponding to a pair of layers, a hydrophobic layer of water with water-water hydrogen-bonds oriented parallel to the SAM surface and a second layer that orients hydrogen-bonds normal to the layer beneath it, above both the C18 and C22 surfaces. A mixture of all four structured layers are visible in all mixed systems and the intensity of each layer was found to be dependent on the makeup of the SAM itself (Figure 8.10). As further evidence that the SAM affects the water structure in more ways than

just makeup, the patterned 50/50 mix was found to do a better job of disrupting the second ordered layer of water (which orients hydrogen-bonds normal to the SAM surface below the C22 surface) than the three randomly generated surfaces, suggesting that ordering effects will indeed act to disrupt water structure in addition to the ratio of siloxane molecules making up a SAM.

The water penetration through the four terminal carbons of the C22 chains to reach the C18 surface in every simulation studied also begins to shed light on the dependence of siloxane chain length with regards to water behavior of SAMs. The relative ease at which water penetrates the length of four carbons of the longer C22 chains is at odds with a previous study that focused on water penetration through a double-bonded terminated eight-carbon SAM chain to reach the hydrophilic crystal slab. The behavior is also very different than that of previous studies of hydrophilic SAM surfaces that contain a hydrogen-bonding group at the surface exposed end of the siloxane molecules. These differences in water behavior suggest that SAM chain length and the terminal chemical group identity do indeed play a significant role in water penetration into a SAM surface, a trait that has been hinted at in the past (293, 351), and still provides a new avenue for research.

Chapter 9

On the Effect of Chain Length and Oxidation State on Water and Monolayer Behavior in Organic Self-Assembled Monolayers

Summary

Self-assembling monolayers (SAMs) have been used as a model system to study complex chemical reactions and water behavior for both organic surfaces in urban environments and atmospheric airborne dust particles for many years. The large variety of SAMs that have been studied include systems that hydrophobic or hydrophilic as well as surfaces constructed of short or long hydrocarbon chains. While the results from studies concerning this wide variety of organic surfaces have been compared and contrasted, to date there has been no systematic study for any possible dependence of the structure or order of the SAM as well as the wetting behavior of the water exposed to the SAM on the chain length and solvent exposed terminal group that characterizes the SAM. Without such a study, it is possible that the comparisons that may be made between two different SAM surfaces may miss a key dependence on their differences where such an affect has been assumed to have minimal effect. What we present is such a systematic study of eight different SAM surfaces with four different chain lengths ranging from 4 to 22 carbons in length, four different solvent exposed terminal groups that range from hydrophobic to fully oxidized and five different siloxane layouts to model defects in the system. The results are able to characterize which portions of the SAM surface display dependence on which SAM characteristic. Among the results is the surprising finding that while chain length and defect does largely affect water penetration into the surface, terminal group shows only a minimum affect, against what was previous thought. These results can serve as a guide for future

comparisons of model systems to ensure that the proper dependencies are taken into account to remove any possible bias in analysis.

Introduction

Understanding the complex behavior of water at organic surfaces, such as those found on the sides of buildings and windows in urban environments or those upon airborne dust particles in the atmosphere, has been the focus of a great deal of work. Studying the nuanced behavior at such varied surfaces have provided countless insights into a variety of fields including but not limited to atmospheric chemistry, biology and inorganic chemistry (286, 289, 290, 311, 371, 386). While the effects of water on a variety of different substrates has been thoroughly studied (295, 361-364, 367, 393), even small differences in the chemical structure of these surfaces have been shown to have drastically different effect on water behavior (292, 293, 311, 338, 343, 387). Given that the variation between such surfaces can play such a large role in wetting behavior and other vital characteristics, several model systems have emerged can be easily altered and controlled in a laboratory, among them self-assembled monolayers (375).

Self-assembled monolayers (or SAMs) consist of simple organic molecules that spontaneously form a monolayer in the presence of a reactive substrate such as gold or varying forms of silicon oxide (312, 375, 376). The monomer of the SAM surface consists of three major group; the reactive head group that bonds to the substrate to form the monolayer, the solvent exposed terminal group that upon monolayer formation is oriented to face the solvent, and the chain that encompasses the region between the head group and terminal group. Each of these three components can be varied, with some of the most common molecules having a Si(OH₃) head group (which gives them the general term siloxane molecules) and a saturated hydrocarbon

chain of varying lengths with varying terminal groups. Variation of both the chain length and the choice of terminal group can lead to fine tuning of the surface environment which has been demonstrated by a variety of previous studies (305, 309, 332, 333, 351, 355, 359, 373, 391, 394). In addition, the density of the surface or presence of missing siloxanes termed as defects in the SAM surface have shown to have a large influence in both water behavior and the behavior of small organic molecules.

A large variety of SAM systems have been characterized by either experimental or computational work in the past and a tempting analysis of such systems would be to compare wetting behaviors and structural characteristics between different systems (351). While some comparisons have been made between very similar or even identical SAMs (305, 332), there has also been a number of instances where the conclusions based on one SAM has been compared with that of a second surface with a vastly different chemical composition (351, 374). While such analyses can provide a great deal of insight, what has not been completed to date is a systematic study that reveals which characteristics are a function of the variation between the surfaces themselves, primarily chain length and solvent exposed terminal groups. Such a study could not only potentially act as a guide for future comparisons between systems by describing where comparisons between different systems are relevant, but could also help to put previous analysis in a new light by highlighting which characteristics depend on which variable. While several studies may have compared a single system against another, a large scale consisting of systematically varied surfaces has not been presented in the literature.

What we present here is such a systematically varied study of SAM makeup and their resulting dependencies on both water and SAM structure and behavior. Eighty different SAMs were generated for study from four different chain lengths varying from 4 to 22 carbons in

length, four increasingly oxidized solvent exposed terminal groups (alkane, alkene, hydroxyl, and carboxylic acid), and one of five different defect layouts including a defect-free layout. The resulting analysis suggest that while the structure of the core or chain region of the SAM depends on chain length alone, the order of the core region shows no such dependence. In the terminal region of the SAM, chain length dependence diminishes and the influence of the terminal group increases displaying a dependence on the final three carbon atoms of the siloxane molecule. Even small changes between similar hydrophobic (alkane vs. alkene) or hydrophilic (hydroxyl vs. carboxylic acid) surface has a significant affect on the structure and position of water on the surface as unique structured layers of water can be identified for each surface. Finally, a more complex wetting behavior, that of water penetration through the SAM surface to the hydrophilic substrate, is analyzed to reveal that surprisingly while chain length and defect size play integral roles, the solvent exposed terminal group appear to play a a very minor or even no role in all in preventing or allowing water penetration.

Methods

An ordered lattice of the alpha-tridymite form of SiO_2 was used as a model for the solid crystal substrate for all SAM systems, identical to previous studies (293). Each Sam surface was constructed from 270 individual siloxane molecules covalently bonded to solvent exposed Si atoms. On the opposite side of the crystal slab, half of the surface exposed Si atoms were removed to maintain charge neutrality. The species of each SAM varied from SAM to SAM with a generalized chemical formula of $\text{X}(\text{CH}_2)_Y\text{Si}(\text{OH})_3$ where X is the solvent exposed terminal group and neighboring carbon atom and Y the total number of carbons minus the two carbons in X. The four terminal groups or X in the general formula used were alkane (CH_3CH_2),

alkene (CH_2CH), hydroxyl (HOCH_2CH_2), and carboxylic acid (HO_2CCH_2). The values of Y used in this study were 2, 6, 16 and 20 for total chain length of 4, 8, 18 and 22 carbons respectively (Figure 9.1A and 9.1B). A $1.0 \text{ kcal/mol/\text{\AA}^2}$ constraint was placed on the position of all substrate atoms in addition to the terminal Si and O atoms of the siloxane molecules. A reflective wall was placed at $z = 90 \text{ \AA}$ and an additional 30 \AA cushion of vacuum included in the unit cell (where z is normal to the organic surface) to prevent water molecules from interacting with the periodic image of the substrate and to render any remaining forces negligible. The final unit cell size of each system was $78.66 \times 75.60 \times 120.00 \text{ \AA}$. To correct for the models anisotropic nature, an external electric field equal to $-4.53786 \text{ kcal}/(\text{mol \AA } e)$ was applied to each system (see previous studies for a further explanation of this correction) (395, 396). Force fields for the SiO_2 substrate were taken from the work of Stevens and coworkers (380) while the force fields for the siloxanes were adapted from CHARMM27 force fields (293). Water molecules were modeled using the TIP3P water model (392).

In addition to the sixteen, pure defect-free SAM systems described above, a series of defective surfaces were generated for each of the sixteen pure surfaces. Defects were modeled by removing a varying number of adjacent siloxane molecules and replacing each with a hydroxyl group covalently bonded to the affected substrate Si atom. Defects of size 1, 2, 3, and 4 missing siloxane molecules were generated for every pure system (Figure 9.1C). These variations resulted in a total of 80 unique systems that were studied through molecular dynamics.

SAM systems were equilibrated in the absence of water for at least ten nanoseconds. Following dry equilibration, 422 water molecules (approximately one monolayer of water) were added to the final frame of the equilibration trajectory to generate the hydrated simulation. The water molecules were generated from a previously equilibrated simulation of bulk water in the

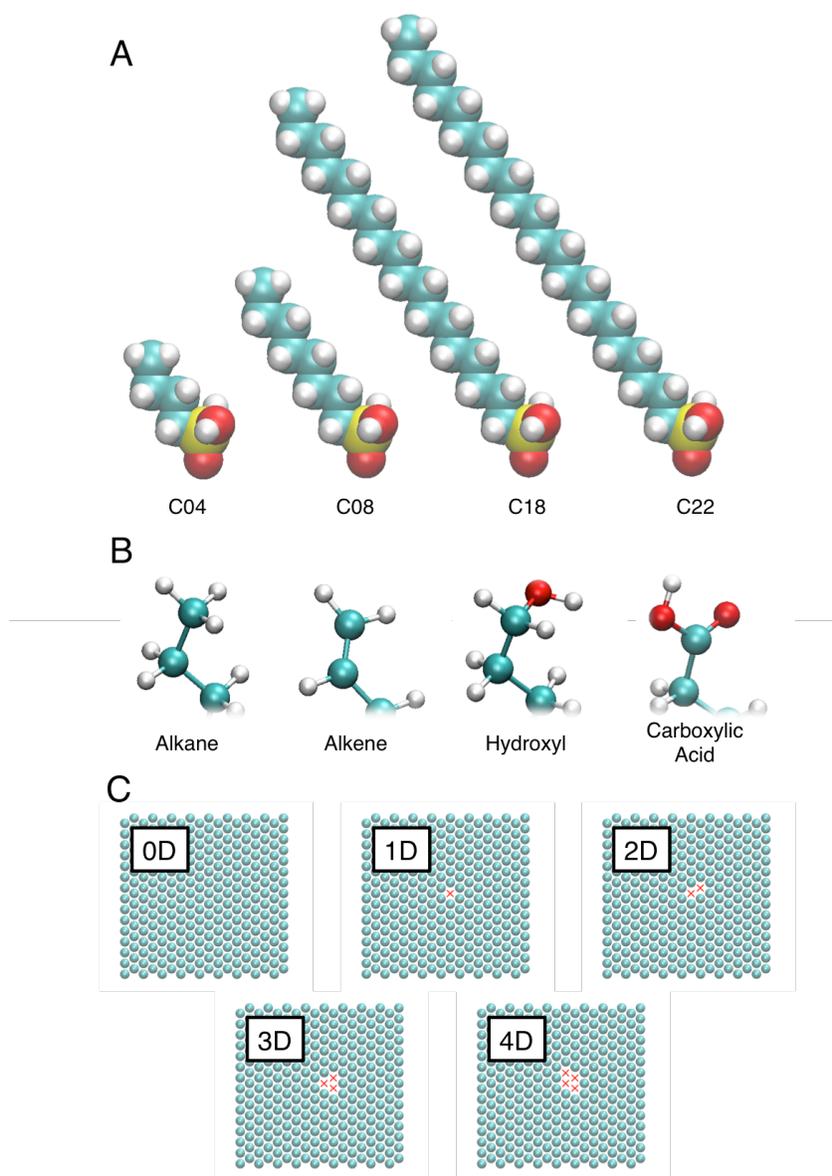


Figure 9.1. Systematic variations of self-assembled monolayer surfaces. Eighty different self-assembled monolayer surfaces were generated in an effort to study any possible dependence of chain length, solvent exposed terminal group and defect size on a variety of characteristics. A) Each of the four chain lengths used in this study, defined by the number of carbons present in the organic chain; C4, C8, C18, and C22. All chains are shown in their fully saturated alkane form. B) The four different solvent exposed terminal group used in this study (in order of increasing oxidation): alkane (-CH₃), alkene (-CH=CH₂), hydroxyl (-CH₂OH), and carboxylic acid (-COOH). C) Schematics of the five different SAM surfaces probed in this study. Spheres represent the first carbon atom in the idealized SAM surface while “X” denotes the presence of a solvent exposed substrate hydroxyl and a missing siloxane molecule. The surface is denoted by a two character code (XD), where X is the number of missing siloxanes and D refers to “defect.”

presence of a hydrophobic C22 SAM surface (chapter eight). The hydrated systems were then equilibrated for an additional ten nanoseconds. The final two nanoseconds of each equilibrated hydrated simulation was then analyzed using VMD and locally developed analysis tools (179).

Simulations were carried out using the NAMD 2.7 package using the isochoric-isothermal (NVT) ensemble (234). Short range interactions were truncated at a distance of 20 Å. Classical equations of forces were integrated by the Verlet algorithm using a timestep of 1 femtosecond. A Langevin thermostat with a 1ps coupling coefficient was employed to maintain the system at a constant temperature of 300K. All bonds with hydrogen in each system were held constant by SHAKE and SETTLE (382).

Results and Discussion

In order to probe any possible dependence of SAM structure, water structure or wetting behavior on varying chain length, oxidation state of solvent exposed terminal group or present defect size, a total of eighty SAM systems were carried out in the presence of approximately one monolayer of water. Each of these systems were consisted of a homogenous surface of one of four possible chain lengths, one of four solvent exposed terminal groups and one of five possible defect layouts (Figure 9.1). To easily approach the results and analysis of this large number of independent systems, SAM structure, water structure and wetting behavior/water penetration will be analyzed separately for any possible dependence on the varying SAM characteristics in Figure 9.1.

I. SAM Structure and Order Dependencies

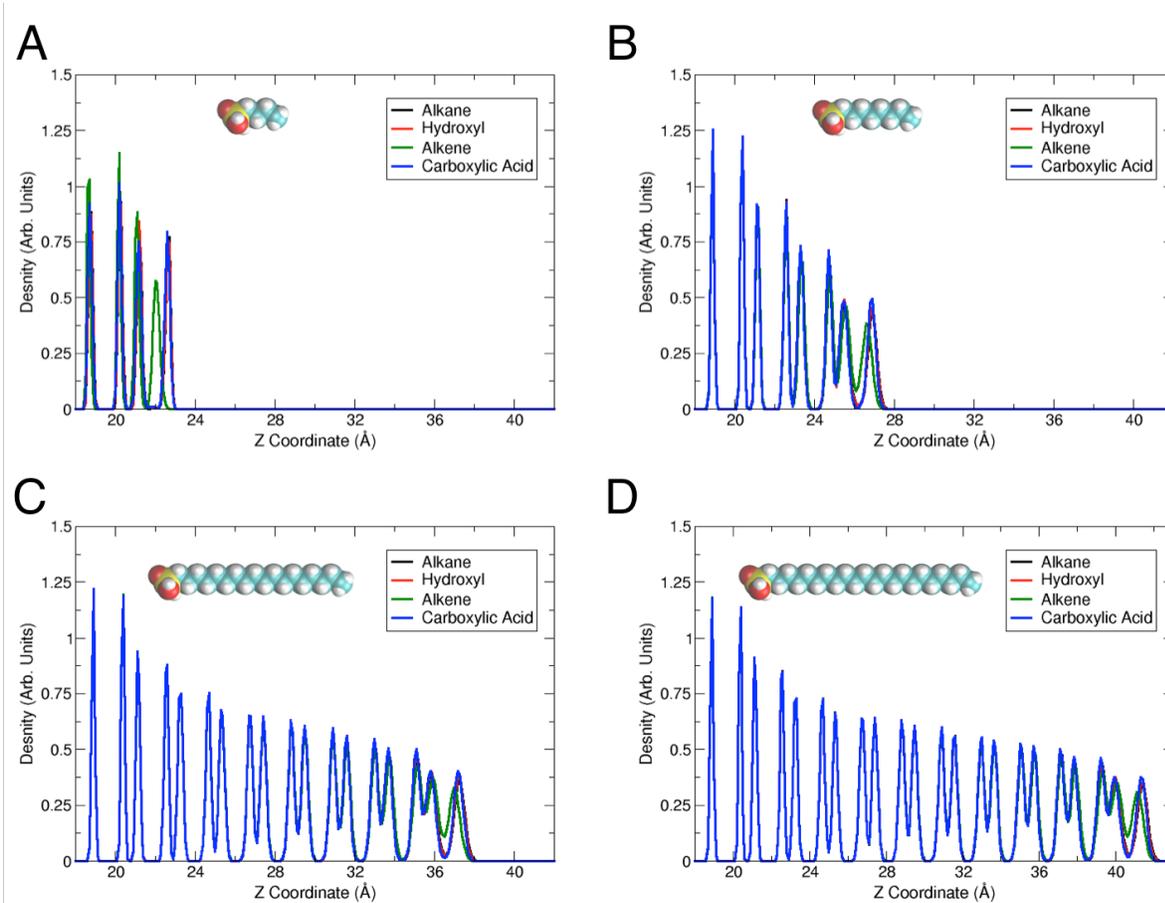


Figure 9.2. Density profiles of the siloxane carbons atoms for defect-free SAM surfaces. Density profiles calculated over the final equilibrated two nanoseconds of the hydrated simulations of each perfect (defect-free) SAM surface with a chain length A) C4, B) C8, C) C18 and D) C22. In each figure, a snapshot of the alkane chain is used to further denote the chain length and terminal groups are color coded as follows: alkane (black), hydroxyl (red), alkene (green), and carboxylic acid (blue).

To probe for any possible differences in the structure of the varying SAM surfaces, density profiles of the SAM carbons for the defect free systems were generated and plotted according to the chain length of the SAM in Figure 9.2. The differences in solvent exposed terminal group have minor differences on the SAM profiles, with the largest change occurring in the positions of the terminal carbons of the alkene-terminated monolayers when compared to other surfaces. This shift, resulting in the central peak of the terminal carbon being shifted by approximately 0.5\AA towards the substrate, is due to the shortened terminal double bond present

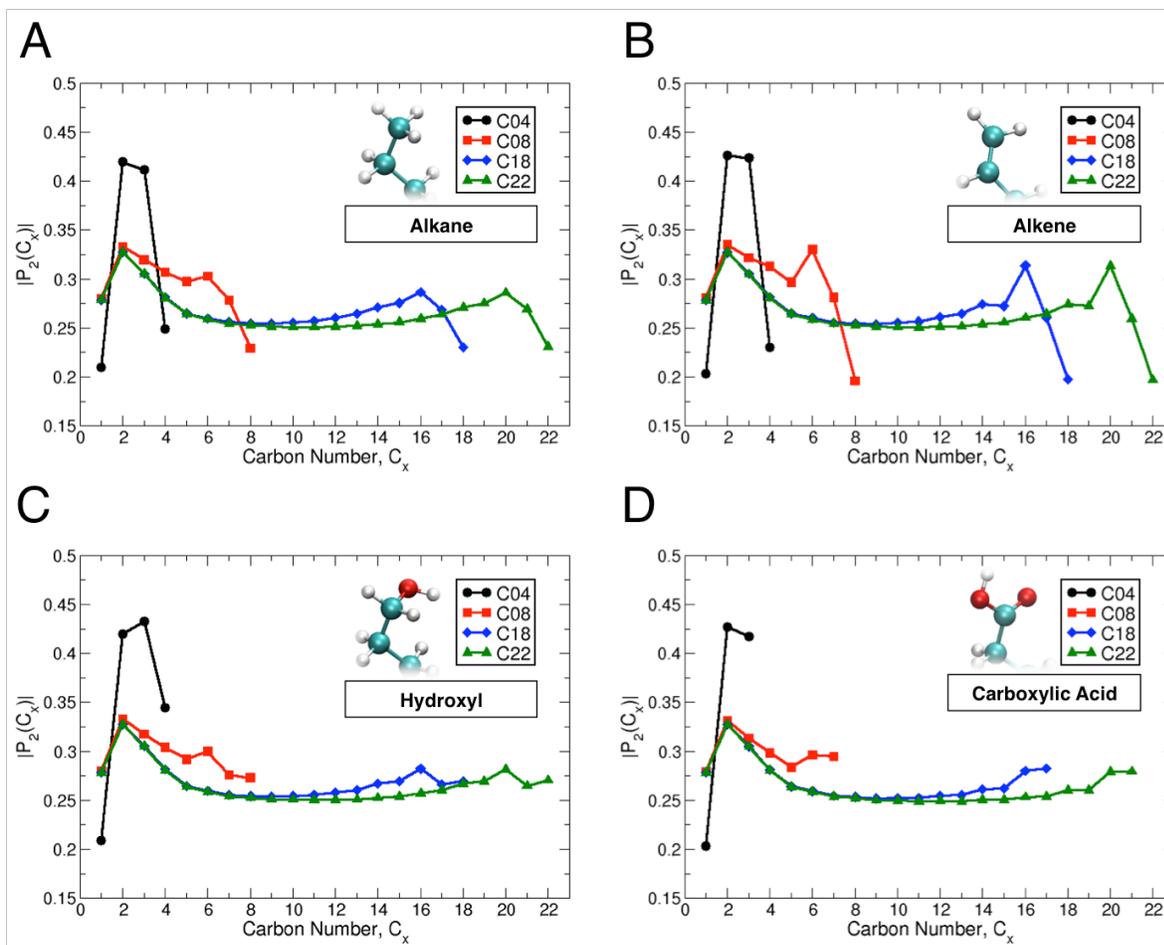


Figure 9.3. Characterization of the overall SAM order. Shown are the absolute value of the average 2nd order Legendre polynomial of the cosine of the angle formed between each C-H bond and a vector normal to the SAM surface where larger values denote a higher degree of order for all perfect (defect-free systems). Each of the sixteen perfect systems are organized by terminal group; A) alkane, B) alkene, C) hydroxyl, D) carboxylic acid. While the shorter chain lengths are suggested to have a higher degree of order, analysis of the average tilt angles of the surfaces reveal that the changes are due to a smaller tilt angle for shorter chain lengths (Table 1). The majority of the difference in order between terminal groups is found in the terminal three carbons of the siloxane.

in the alkene system as compared to the terminal single bonds of all other studied surfaces. The alkene effect is also dependent in chain length, as the shorter chain length shows a greater deviation for the position of the final carbon.

The order of the SAM surface itself was characterized by calculating the absolute value of the second order Legendre polynomial for the cosine of the angle formed by each C-H bond

and a vector normal to the SAM surface itself. The results of this analysis are displayed in Figure 9.3. Focusing on chain length effects first, the data suggests that the shorter chains are more ordered due to higher order parameter values as compared to the longer chains for all terminal groups. However, the order analysis does not take into account the possible variation in the tilt of the SAM surface itself which itself may be dependent on chain length or terminal group.

Average tilt angles, defined as the angle between a vector defined as the position of the first and final carbons and a vector normal to the SAM surface, reveals that there is a clear dependence of tilt angle on chain length (Table 9.1 and Figure 9.4). Analysis of the accompanying defective systems shows that the presence of small defects has no global dependence on tilt angle (Table 9.2). When these analyses are taken together, while the tilt of SAMs will vary between chain lengths, the difference in order will be minimal.

Table 9.1. Average tilt angle of varying defect-free SAM surfaces.

Chain Length	Terminal Group			
	Alkane	Alkene	Hydroxyl	Carboxylic Acid
C04	$9.98 \pm 0.26^\circ$	$23.14 \pm 0.25^\circ$	$10.33 \pm 0.29^\circ$	$9.91 \pm 0.28^\circ$
C08	$26.36 \pm 0.30^\circ$	$26.52 \pm 0.34^\circ$	$26.66 \pm 0.29^\circ$	$26.99 \pm 0.26^\circ$
C18	$31.76 \pm 0.18^\circ$	$31.36 \pm 0.18^\circ$	$31.91 \pm 0.17^\circ$	$32.09 \pm 0.16^\circ$
C22	$32.49 \pm 0.15^\circ$	$32.54 \pm 0.17^\circ$	$32.47 \pm 0.15^\circ$	$32.80 \pm 0.15^\circ$

Focusing on the differences between terminal group in SAM order in Figure 9.3, the core of the SAM (defined here as the region between the first carbon up until but not including the terminal two carbons) shows very little variation. Interestingly, the third to final carbon ($C_{(n-3)}$ where n is the chain length) displays a significantly higher order than both the terminal carbons as well as all but the initial carbons in the siloxanes. The final carbons show a higher degree of disorder, suggesting the final two carbons of each SAM surface will display a higher degree of disorder when exposed to water. The terminal region predictably shows a much greater

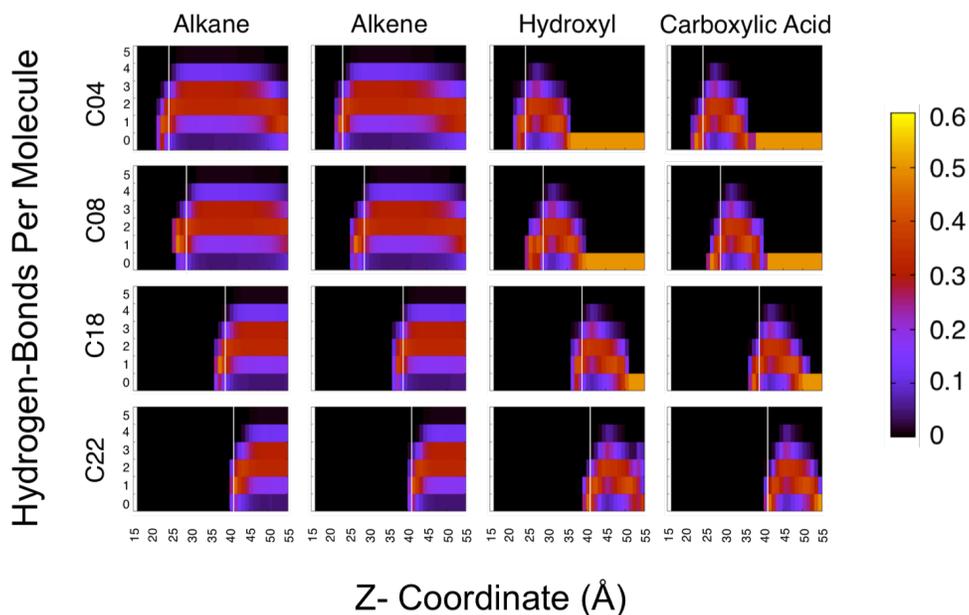


Figure 9.4. Dependence of chain length and solvent exposed terminal group on SAM tilt angle. Displayed are the distributions of the average tilt angle, calculated over the final two equilibrated nanoseconds from each simulation, for all SAM systems. The raw data and standard deviations can be found in Table S1. Each system is designated by a single x that is colored by the appropriate solvent exposed terminal group (alkane, black; alkene, red; hydroxyl, green; carboxylic acid, blue). Logarithmic regressions of each of the four terminal groups are shown in the same color as described above. While terminal group has minimal affect on the tilt, the alkene has a large effect on observed tilt angle for the shortest chain length (C4). All other combinations of terminal group to a given chain length are very similar and all but the previously described alkene produce similar regressions.

dependence on terminal group with a lower degree of order in the alkene when compared to the alkane for the hydrophobic surfaces. For the hydrophilic surfaces, the carboxylic acid surface shows a larger order than the terminal region of the hydroxyl surface. However, as there were no C-H bonds on the terminal carbon of the carboxylic acid siloxane, no data could be collected for that terminal carbon. Snapshots of each terminal group from an equilibrated simulation are included in both Figure 9.1 and 9.3.

Table 9.2. Average tilt angle for all studied SAM systems.

Chain Length	Terminal Group	Size of Defect				
		No Defects	One Siloxane	Two Siloxanes	Three Siloxanes	Four Siloxanes
C04	Alkane	9.98 ± 0.26°	10.07 ± 0.26°	10.19 ± 0.26°	10.13 ± 0.29°	10.32 ± 0.28°
	Alkene	23.14 ± 0.25°	23.23 ± 0.26°	23.32 ± 0.26°	23.33 ± 0.28°	23.41 ± 0.26°
	Hydroxyl	10.33 ± 0.29°	10.41 ± 0.28°	10.65 ± 0.30°	10.82 ± 0.29°	10.72 ± 0.32°
	Carboxylic Acid	9.91 ± 0.28°	9.96 ± 0.26°	10.21 ± 0.28°	10.36 ± 0.45°	10.42 ± 0.27°
C08	Alkane	26.36 ± 0.30°	26.38 ± 0.31°	26.41 ± 0.31°	26.30 ± 0.34°	26.12 ± 0.32°
	Alkene	26.52 ± 0.34°	26.51 ± 0.35°	26.50 ± 0.35°	26.71 ± 0.35°	26.95 ± 0.37°
	Hydroxyl	26.66 ± 0.29°	26.73 ± 0.29°	26.79 ± 0.29°	26.89 ± 0.31°	26.64 ± 0.33°
	Carboxylic Acid	26.99 ± 0.26°	26.89 ± 0.27°	27.09 ± 0.26°	27.05 ± 0.28°	27.06 ± 0.29°
C18	Alkane	31.76 ± 0.18°	31.80 ± 0.18°	31.89 ± 0.18°	32.00 ± 0.19°	32.08 ± 0.19°
	Alkene	31.36 ± 0.18°	31.89 ± 0.19°	31.94 ± 0.20°	32.11 ± 0.19°	32.03 ± 0.21°
	Hydroxyl	31.91 ± 0.17°	31.97 ± 0.17°	32.02 ± 0.18°	32.10 ± 0.18°	32.22 ± 0.19°
	Carboxylic Acid	32.09 ± 0.16°	32.14 ± 0.16°	32.24 ± 0.17°	32.46 ± 0.17°	32.51 ± 0.17°
C22	Alkane	32.49 ± 0.15°	32.60 ± 0.15°	32.66 ± 0.16°	32.76 ± 0.16°	32.87 ± 0.16°
	Alkene	32.54 ± 0.17°	32.63 ± 0.18°	32.61 ± 0.17°	32.85 ± 0.15°	33.02 ± 0.16°
	Hydroxyl	32.47 ± 0.15°	32.47 ± 0.15°	32.60 ± 0.16°	32.85 ± 0.15°	33.03 ± 0.16°
	Carboxylic Acid	32.80 ± 0.15°	32.70 ± 0.15°	32.87 ± 0.14°	32.85 ± 0.15°	33.08 ± 0.15°

II. Water Structure Dependencies

Density profiles of the water in each of the defect free simulations were calculated to compare the wetting behavior on each of the SAM surfaces (Figure 9.5). Focusing on terminal group effects first, droplet formation is observed on the hydrophobic alkane and alkene systems, visualizations of which can be seen in Figure 9.6A and 9.6B. While the droplets are similar in profile, there are several important differences between them. The position of the droplet in relation to the SAM surface is shifted further away for the alkane and equal distance to the difference in the terminal bond in the siloxane, though this movement is increased in the C4 system due to the larger difference in SAM tilt angle. The resulting droplets were not observed to be significantly different over the length of each trajectory. Though the alkene surface would generally produce droplets that would make more contacts with the SAM surface, such contacts were not significantly different from those formed on alkanes. Each of the two hydrophobic surfaces formed a poorly hydrogen-bonded layer of water that is in contact with the SAM

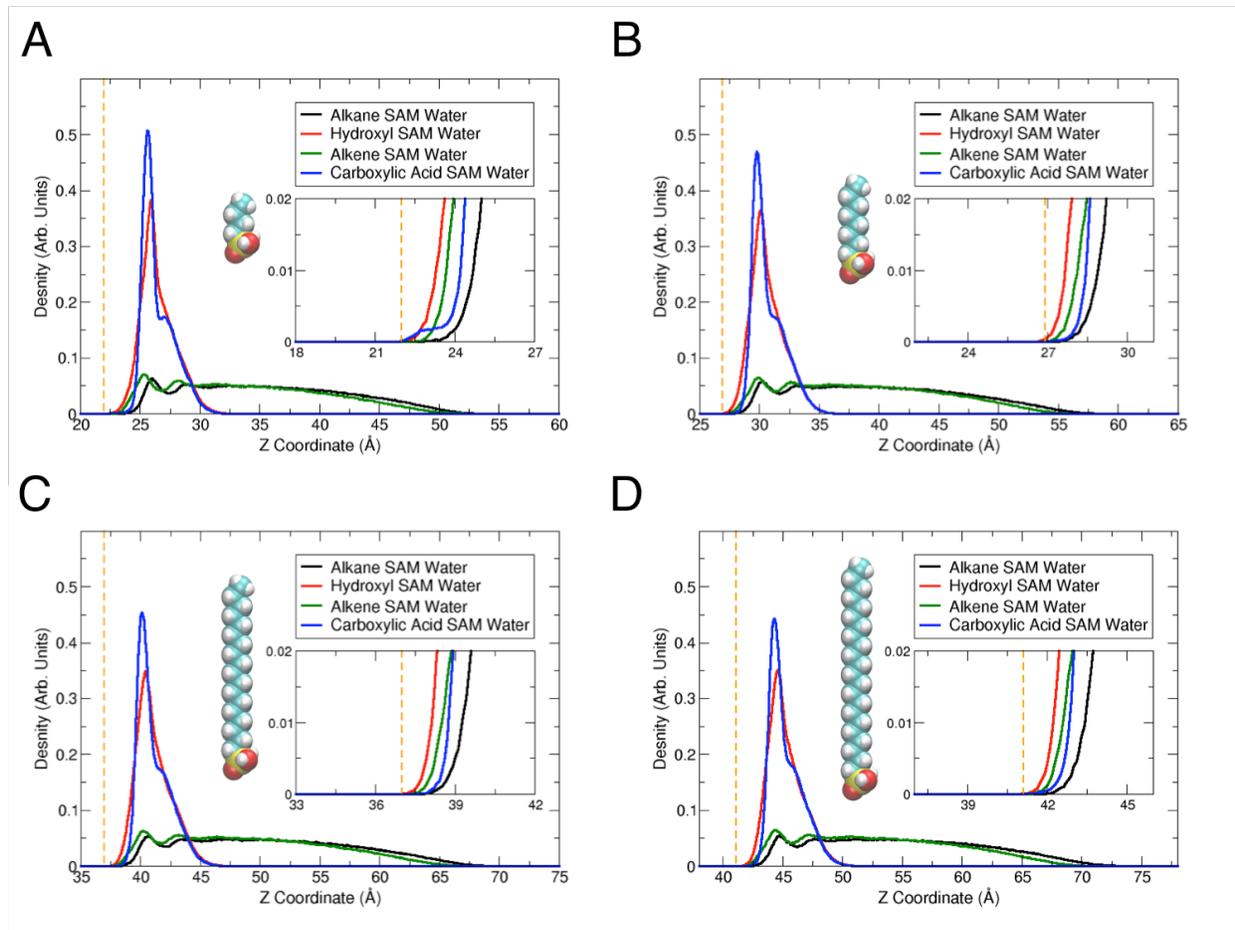


Figure 9.5. Water density profiles for defect free systems. Shown are density profiles for the water in each of the sixteen perfect (defect-free) systems for chain lengths of A) C4, B) C8, C) C18, and D) C22. The different terminal groups on each of the surfaces are distinguished by the color of the profile, shown in the figure legend. The orange dashed line represents the average position of the terminal carbon profile shown in Figure 9.2 as to represent the SAM surface. The insert of each plot is a zoom of the SAM surface. While the profiles of the alkane and alkene surfaces are very similar with differences only in position relative to the SAM surface, those of the two hydrophilic surfaces show significant differences beyond surface proximity.

surface, as shown by the peak in the density profile near the surface and analysis of the number of hydrogen-bonds per molecule (Figure 9.7). Each droplet can be separated into several distinct structural phases as revealed by hydrogen-bond orientation analysis (Figure 9.7). Starting at the SAM surface is a layer of water that shows a clear preference to orient its hydrogen bonds normal to the SAM surface. This initial preferred orientation is followed by a second layer that prefers to arrange hydrogen bonds parallel to the SAM surface. Beyond this second layer the

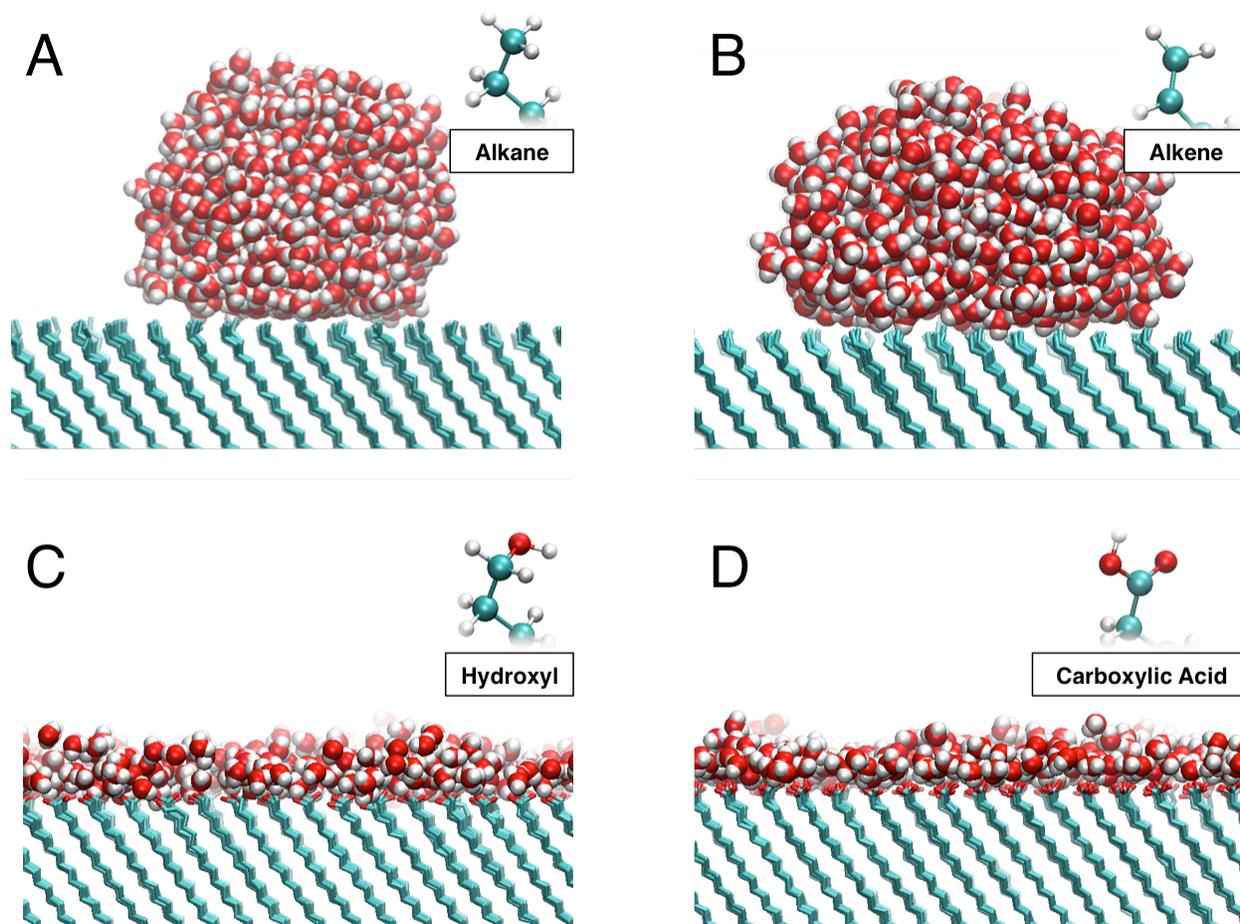


Figure 9.6. Snapshots of equilibrated water on each of the four different terminal groups of a C22 SAM. Snapshots of the C22 perfect SAM systems after equilibration show how the formed droplet interacts with the SAM surface. A) On the alkane surface, a droplet forms to minimize the contact between the hydrophobic surface and water molecules. B) The alkene surface similarly forms a droplet after equilibration. C) The water of the hydroxyl terminal surface equilibrates into a monolayer which hydrogen bonds with the exposed terminal hydroxyl group. D). The water on the carboxylic acid surface forms a monolayer of water that can be further divided into two separate structured layers dependent on the hydrogen-bonding partner of the solvent exposed terminal group.

droplet shows no clear preference for hydrogen-bond orientation suggesting that bulk-like, randomly oriented, well-hydrogen bonded waters dominate.

The monolayers formed by the hydrophilic surfaces terminated by hydroxyl and carboxylic acid terminal groups both form simple monolayers of water on the SAM surface though there are slight differences between the two terminal groups. The density profiles suggest

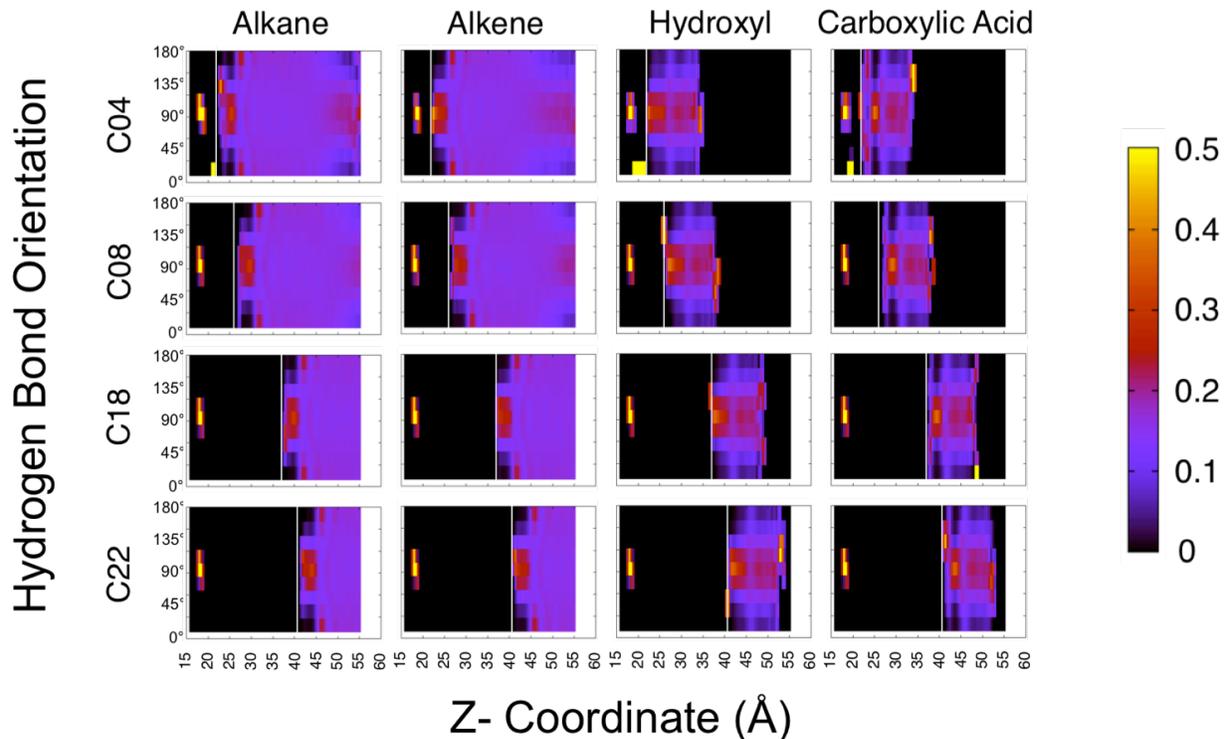


Figure 9.7. Hydrogen bonds per molecule distribution along the z-axis for defect free SAM surfaces. Intensity maps of the average number of hydrogen bonds per water molecule along the z-axis are shown for each of the perfect (defect-free) systems over the final two nanosecond of each hydrated simulation. The z-axis was discretized into 0.5\AA bins while each z slab was further divided into six bins based on the number of hydrogen-bonds per water molecule (0 to 5 hydrogen-bonds). Data was normalized by dividing each bin by the total number of occurrences for the entire given z-slab. The solid white line represents the approximate center of the terminal carbon density as taken from Figure 9.2. The hydrophobic surfaces both form droplets that reach well-hydrogen bonded bulk like water $10\text{-}15\text{\AA}$ away from the surface and a poorly hydrogen bonded layer of water near the surface. The hydrophilic surfaces however have well-hydrogen bonded water located directly at the SAM surface and poor hydrogen bonding for neighboring waters. For the hydrophilic surfaces, there are regions that correspond only to water with no hydrogen-bonding patterns which denotes a water molecule that has escaped the monolayer into the vapor phase.

that while the hydroxyl surface generates a single peak or monolayer, the carboxylic acid surface shows a second peak approximately 2\AA further from the surface. Analysis of the hydrogen-bonding patterns of these surfaces show that while the hydroxyl surface generates a single monolayer that has a strong preference to orient hydrogen-bonds normal to the surface, carboxylic acid surfaces generate two distinct water layers corresponding to the two distinct

possible partners in the carboxylic acid terminal group. The layer closest to the surface orients hydrogen-bonds parallel to the surface (with the carbonyl of the terminal group, centered closer to the surface than the hydroxyl surface layer), while the second structured layer of water hydrogen bonds normal to the surface (with the hydroxyl group, further from the surface than the single hydroxyl water layer). Despite this structural difference, the number of hydrogen bonds per molecule near the surface is very similar for both hydrophilic terminal group.

III. Water Penetration Dependencies

While the previous section focused on water on the surface of an idealized SAM surface, true SAMs seldom reach perfect coverage. How water interacts with these organic surfaces in the presence of defects (such as missing siloxane molecules) are critical to understanding the complex behavior of water on these surfaces in a more real world application. The so called “healing effect” of these organic surfaces, the ability for defect neighboring siloxanes to reorient themselves to “heal” the gap in the surface and prevent water from penetrating to the substrate is one such complexity that has been the focus of recent work. To this point, for each of the sixteen SAM systems, four additional systems were generated with defects of one, two, three and four missing siloxane molecules were generated to probe the dependence of water penetration into a defect on chain length, terminal group, and defect size. A total of 80 systems were analyzed for this analysis (see Methods).

If there is no such dependence on water penetration (or point at which the SAM healing fails) on chain length or terminal, then we would expect to observe water penetration at the same size of defect for all systems. However, if a longer chain leads to a greater healing effect, then we would expect larger defects to observe penetration for such systems. Each of the systems was

exposed to one monolayer of water and the final two nanoseconds of the equilibrated trajectory analyzed for the presence of water penetration into the defect. For the purposes of this analysis, water penetration was characterized by analyzing the hydrogen-bond orientation as in Appendix C, where the presence of hydrogen-bonding beyond the SAM surface denotes water penetration where the SAM surface is defined at the center of the density profile for the terminal carbon. The result of this analysis is summarized in Table 9.3 and in more detail in Appendix D. Water penetration was categorized into three classes; no penetration (no hydrogen-bonding observed by orientation analysis beyond 5Å of the SAM surface), partial penetration (hydrogen-bonding observed by orientation analysis beyond 5Å of the SAM surface but not to the terminal OH groups of the siloxanes or defect hydroxyl), and full penetration (hydrogen-bonding observed by

Table 9.3. Water penetration dependence on chain length, terminal group and defect size.

Chain Length	Terminal Group	Size of Defect				
		No Defects	One Siloxane	Two Siloxanes	Three Siloxanes	Four Siloxanes
C04	Alkane	X	✓	✓	✓	✓
	Alkene	X	✓	✓	✓	✓
	Hydroxyl	X	~	✓	✓	✓
	Carboxylic Acid	X	✓	✓	✓	✓
C08	Alkane	X	X	X	✓	✓
	Alkene	X	X	✓	✓	X
	Hydroxyl	X	X	✓	✓	✓
	Carboxylic Acid	X	X	~	✓	✓
C18	Alkane	X	X	X	X	X
	Alkene	X	X	X	X	~
	Hydroxyl	X	X	X	X	~
	Carboxylic Acid	X	X	X	~	~
C22	Alkane	X	X	X	X	X
	Alkene	X	X	X	X	X
	Hydroxyl	X	X	X	X	X
	Carboxylic Acid	X	X	X	~	~

X – Minimal Penetration

~ – At least 5Å of Penetration

✓ – Penetration to Substrate

orientation analysis throughout the entire SAM surface). Density profiles similar to that of Figure 9.5 for each defective system can be found in Appendix E.

As seen in Table 9.3, it is clear that both chain length and solvent exposed terminal group affect water penetration, though the effects are very different. Firstly, it is clearly apparent that water penetration is strongly dependent on chain length. Focusing on a single terminal group, such as the hydroxyl surface, water penetration is observed starting at a defect size of one for C4 chains, two for C8, 4 for C18 and no observed at all for C22. This dependence is strong for all terminal groups of varying defect size and chain length. The affect that the terminal group plays in this behavior is not nearly as defined as chain length. While a small dependence can be observed by the fact that water penetration is observed only for carboxylic acid surfaces of a chain length of C22, or that for a chain length of C8 all terminal group show at least partial penetration aside from the most hydrophobic alkane, the trend is not well defined nor strong enough to draw a clear conclusion on the strength of the dependence. However, there are a few small anomalies that merit further discussion.

There was one system where water penetration was expected to occur where it was not observed, the alkane C8 system under the presence of a defect size of three missing siloxane molecules. Previous simulations carried out in an identical system observed water penetration into the defect. Visualization of this simulation revealed that after the initial 10 nanoseconds of dry equilibration, the defect was perfectly “healed” by neighboring siloxanes at the end of the equilibration. Once water was added to the simulation, with no way to penetrate into the SAM surface, the droplet formed and moved away from the defect. Over three separate hydrated runs of this perfect healed system, the lifetime of this healed defect SAM ranged from 1 to 4 nanoseconds after water was added, though by the time the defect was exposed, the droplet had

moved or formed away from the defect. In comparison, at no point in the 4-defect system was the defect perfectly healed and only the smallest defect (one missing siloxane) was able to remain healed for the full 20 nanoseconds of analysis (10 nanoseconds dry and 10 nanoseconds hydrated). We believe that this observation does show that while water penetration can be prevented at larger defect sizes than required, the healing of these systems are temporary.

Finally, the actual penetrating water is not significantly different for a given terminal group, as shown in Figure 9.8. Snapshots from C8 surfaces with a defect size of three, where penetration was observed for all surfaces, show that water extends in a water wire like structure from either the droplet or monolayer on the surface to reach either the terminal hydroxyl of the siloxane molecules or the defect hydroxyl of the substrate. The snapshots also demonstrate the variety of water penetration phenomena observed in this study. In Figure 9.8A, the defect in the alkane surface has been completely filled by penetrating water and the defect neighboring siloxanes restored back to a near perfect order by the penetrating water. In the alkene system, the defect neighboring siloxane molecules were able to reorient themselves to heal the defect, but only at the price of exposing their own terminal hydroxyl groups that are now free to hydrogen-bond with water preventing water only from interacting with the substrate but still finding a hydrogen-bonding partner deep within the monolayer. Finally, in both the hydroxyl and carboxylic acid surfaces, only partial healing is achieved, as defect neighboring siloxanes are able to protect only a portion of the defect but not prevent water wires from forming within the defect to span the gap between the monolayer and the substrate. From these results we can conclude that chain length and defect size play much stronger roles in this wetting behavior than terminal group.

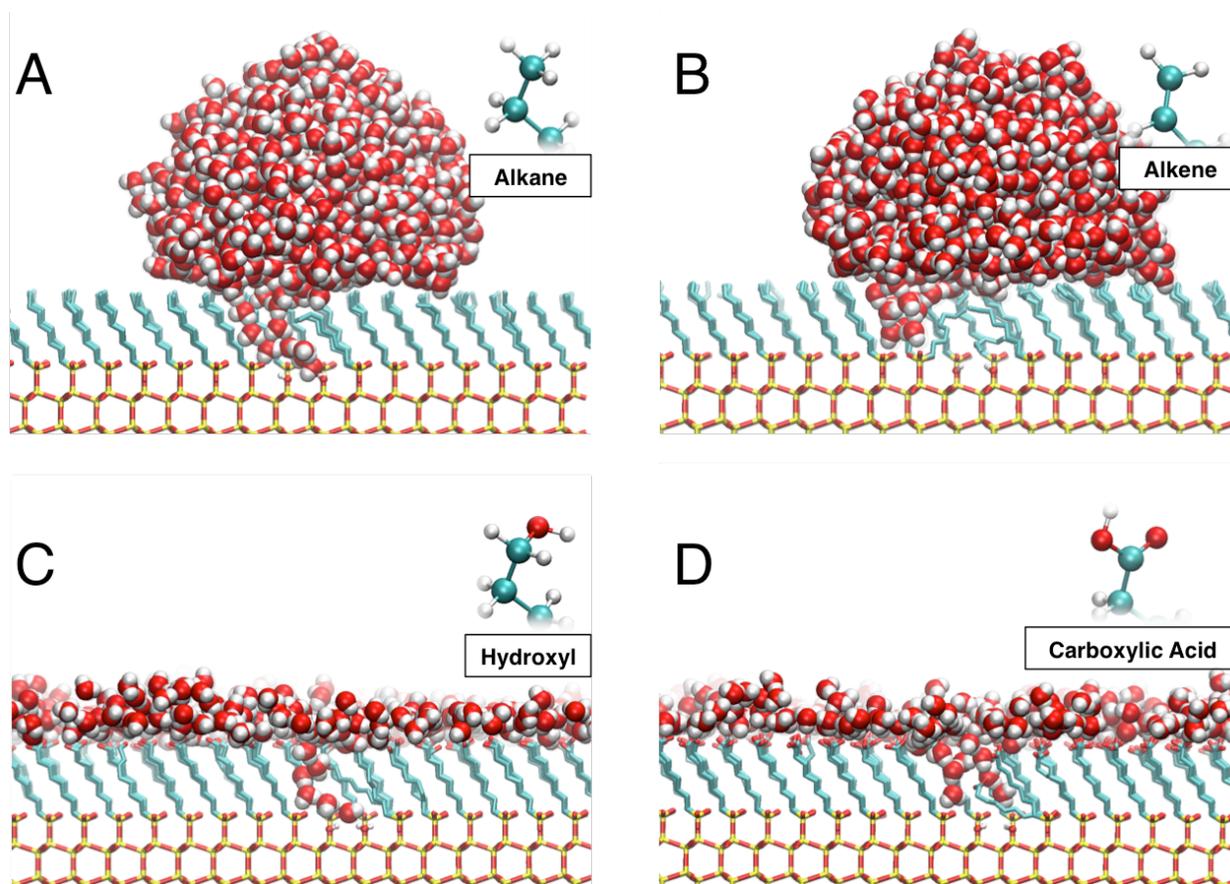


Figure 9.8. Snapshots of water penetration for each of the four terminal groups for the C8 system. Examples of water penetration for each of the four possible solvent exposed terminal group are shown via snapshots of the equilibrated C8 system with a defect size of three. A). In the alkane system, a water wire like structure forms to anchor the droplet to the defect site as the water hydrogen bonds with the defect hydroxyl group. With a defect size of three, two separate water wires can be generated in the defect. B) The alkene system, a different penetration method can be observed as the defect neighboring siloxanes molecules can partially heal the defect site, though at the expense of exposing the terminal hydroxyl groups of the siloxane to the water wires connected to the droplet. C) The hydroxyl surface is able to penetrate into the surface despite partial healing by neighboring siloxanes. The formed water wire forms a hydrogen bonding network that connects the defect to the monolayer above the surface. D) Similar to the hydroxyl surface, water wire structures extend downwards from the monolayer to reach not only the defect hydroxyls, but also the terminal hydroxyls of the siloxanes similar to the alkene example in B.

Conclusions

Taken together, the analysis presented herein can predict what aspects of SAM and water structure as well as behavior will vary between different SAM surfaces, which can in turn provide a set of guidelines that can determine whether results from one SAM surface can be

compared to another. As the results presented herein show, the behavior of the SAM and water can vary greatly depending on both chain length and solvent exposed terminal group, so it is important to know when SAM surface can be directly compared and which aspects of a SAM surface are translatable despite differences in siloxane chemical makeup.

With regards to the SAM structure itself, we have shown that the core order of an idealized SAM surface shows minimal dependence on either chain length or terminal group, though the terminal carbon order can be influenced by the terminal group in a small way. Additionally, global average tilt of the SAM surface shows a large dependence on chain length. The only instance of a terminal group effect on tilt was found in an alkene-terminated C4 SAM, wherein the short chain was unable to adapt as well as longer chains to the presence of such a conformationally restricting terminal group. These results would suggest that characteristics of the core of the SAM could be directly compared for SAM with similar chain lengths regardless of the terminal group present, while the terminal region shows the opposite effect.

The water structure on the surface of a defect free SAM surface shows no dependence on chain length as expected while even small changes in the terminal group can have a large change in the resulting water structure. While there was minimal differences between the hydrophobic surfaces probed in this study with the position of the resulting droplet depending on the tilt of the siloxane and structure of the terminal group itself. The differences between the hydrophilic hydroxyl- and carboxylic acid-terminated surfaces on the water structure were much more pronounced. While both would form a layer of water on the surface, the carboxylic acid surface resulted in the generation of two distinct structured layers of water while the hydroxyl surface only had a single monolayer. Based on these results, it would seem that it would be safe to

compare details of water structure on SAMs only where the terminal group is the same regardless of SAM chain length.

Finally, with regards to the healing effect behavior of SAMs to prevent water penetration, chain length and size of the defect play a large role in water behavior while the solvent exposed terminal group plays a very minimal role. This finding is particularly intriguing as while many of the previous results may be expected, it may run contrary to what would be expected to occur at this surface. With regards to comparisons between SAM surfaces, these results would suggest that studies of water penetration could be compared between SAMs of varying terminal groups given that they are of the same chain length.

These findings help to underscore the importance of studying these organic surfaces as even though this model system has been the focus of study for many years, we are still gaining new insights into the complex water behavior that may be unexpected. These results also display how much of either SAM or water structure/order can be altered by even small alterations in chain lengths and terminal groups. Other characteristics of other complex behaviors such as water penetration may reveal other intriguing dependencies on the chemical makeup of the SAM itself. These results should serve as a guide for future studies when it is time to compare the results of a given SAM system with those carried out on a different system.

Chapter 10

A Brief Introduction to Other Works

In addition to the work included herein concerning heme protein redox partners and self-assembled monolayers, I also studied three other systems during my tenure at the University of California, Irvine that warranted inclusion in my thesis. Unlike the preceding work on heme protein redox partners and self-assembled monolayers that were major points of interest however, these works played smaller though still important rolls in my thesis studies. As each of these projects lie outside the focus of previous sections, they are included in chapters eleven, twelve and thirteen and are briefly introduced here. The introduction of each respective chapter serves as a more in depth introduction of each subfield of research.

Protein Conformation: More Than Meets The Eye

The first of these additional studies focuses on the relationships between the conformation of a central protein residue and the conformation of its neighbors. One of the founding principles of protein structure and conformation research is the Flory Isolated Pair Hypothesis (397) which states that the conformation of a given residue (which can be summarized by its (ϕ , ψ) backbone torsion angles) does not affect the conformation of its neighbors. Despite this hypothesis serving as a basis for protein conformation work for nearly five decades, multiple studies have begun to show that this hypothesis is in fact false (398-400). While these new studies have begun to change the way we view conformational dependencies in proteins, this founding principle has not yet been entirely disregarded. Previous work in identifying the most common two-residue motifs in proteins also began to shed light on the degree to which the conformation of a given residue does indeed affect the conformation of its

neighbors (401). In chapter eleven, this previous work is further developed to show the degree of which the conformation of a given residue alone is dependent on the conformation of its neighbors.

This work is an extension of a series of studies that I carried out as a member of the P. Andrew Karplus Lab at Oregon State University as an undergraduate from 2007 until 2010 that I continued as a graduate student at UCI as an ongoing independent collaboration with Dr. Karplus. For a more detailed introduction into protein conformation, the reader is referred to a review that was authored as part of my undergraduate work, Hollingsworth and Karplus 2010 (402), which introduces the current state of the field.

Pleckstrin Homology Domain and Cascade Pathways

The second study included herein concerns the binding of the pleckstrin homology (PH) domain to a target phosphatidylinositol-3,4,5-trisphosphate (PIP₃) modified lipid in a membrane. The PH domain, a structurally conserved and widespread protein domain, is recruited to the membrane bilayer through the presence of the modified PIP₃ head group and acts a second messenger of a much larger kinase cascade pathway in many different organisms (403, 404). Understanding the role that the PH domain plays in the kinase cascade pathway is important as alteration of the specificity of the PH domain to the PIP₃ headgroup has been shown to lead to the development of cancer (405). To begin to understand how the PH domain functions, we decided to probe the stability and dynamics of the PH domain bound to its target headgroup by carrying out a simulation of the PH domain in complex with its target based off recent EPR data (406) from the Falke group, which is presented in chapter twelve.

This work was carried out in collaboration with the Dr. Stephen H. White group at UCI as well as the Joseph J. Falke group at the University of Colorado at Boulder through the Tobias group at UCI.

Crystal Structure of Bacillus subtilis Arginase

The third and final chapter included here is, unlike all chapters preceding it, a purely experimental study. I sought to solve the novel crystal structure of the enzyme arginase from the Gram-positive bacteria *Bacillus subtilis*. *B. subtilis* has served as a model organism for the Poulos Lab in recent years in the development of bacterial nitric oxide synthase (bNOS) inhibitors in an effort to create new drugs to aid in the treatment of methicillin-resistant *Staphylococcus aureus* (more commonly known as MRSA) and anthrax (116-119). Importantly however, NOS activity is tightly connected with the enzyme arginase, which share a common substrate L-arginine but carry out two different reactions (407-409). Given the interplay between these enzymes in not only humans but also in bacteria where such regulation may be an employed strategy to protect the pathogen from the host, it is important to study the activity and relationship between arginase and bNOS in *B. subtilis*. The first step was to solve the structure of this enzyme, which is presented in chapter thirteen.

This final study was carried out in collaboration with both Dr. Jeffrey Holden and Matthew Lewis of the Poulos Lab at UCI.

Chapter 11

Beyond Basins: ϕ, ψ Preferences of a Residue Depend Heavily on the ϕ, ψ Values of its Neighbors

*Work carried out as co-first author with Matthew C. Lewis. Author Contributions: SAH and MCL generated the datasets, carried out visualization and analysis. SAH, MCL and P. Andrew Karplus wrote and edited the manuscript.

Summary

The Ramachandran plot distributions of non-glycine residues from experimentally determined structures are routinely described as grouping into one of six major basins: β , P_{II} , α , α_L , ξ (Pre-Pro) and γ' . Recent work describing the most common conformations adopted by pairs of residues in folded proteins [i.e. $(\phi, \psi)_2$ -motifs] showed that commonly described major basins are not true single thermodynamic basins, but are composed of distinct subregions that are associated with various conformations of neighboring residues. Here, as documentation of the extent to which the conformational preferences of a central residue are influenced by the conformations of its two neighbors, we present a set of ϕ, ψ -plots that are delimited by the ϕ, ψ -angles of neighboring residues. The level of influence seen here is typically greater than the influence associated with considering the identities of neighboring residues, implying that the use of this heretofore untapped information can improve the accuracy of structure prediction algorithms and low resolution protein structure refinement.

Introduction

Proteins can be described as a series of local conformations or motifs that are strung together to make the 3-dimensional structure (410). Assuming trans peptide bonds, the local protein conformations themselves can be largely described by two torsion angles, ϕ and ψ ,

adopted by each residue. Plotting these two variables against each other leads to the well-known Ramachandran plot, which has served as a basis of understanding protein conformation for fifty years (Figure 11.1) (411-416). Over the decades, various nomenclatures have been used for the populated regions within the plot (e.g. Hollingsworth & Karplus (402)), with recent systems being more closely based on the natural groupings of residues so that each of these apparent “thermodynamic basins” (417) is given a name (Figure 11.1A).

In the 1960s, in support of calculations related to understanding polypeptide conformations, the Flory Isolated-Pair Hypothesis posited the simplifying assumption that the conformation of a given residue does not affect the conformation of its neighbors (397). This hypothesis still guides much thinking about protein conformation, even though studies have shown that it does not, in general, hold for polypeptides (398-400). Notably, Pappu et al (398) used a dataset of Monte Carlo generated Ala-polypeptides to show that simple steric considerations cause the Isolated-Pair Hypothesis to break down quickly with chain length.

While much work has been done on the relationships between amino acid sequence and conformation (400, 418-420), less has been done to explore the details of how the conformation of a residue depends on the *conformation* of its neighboring residues, irrespective of amino acid type (400, 421, 422). Recently, we published a study of local protein conformations that focused on what we called $(\phi, \psi)_2$ -motifs, meaning motifs defined by the conformations of two adjacent residues: $\phi_{i+1}, \psi_{i+1}, \phi_{i+2}, \psi_{i+2}$ (401). In that study, we developed a list of the 101 most common two-residue conformational motifs in proteins (401). Interestingly, a plot of all of the two central ϕ, ψ values associated with each of these motifs shows that most of the populations are not at the centers of the well-defined basins, but are rather broadly distributed within them (Figure 11.1A).

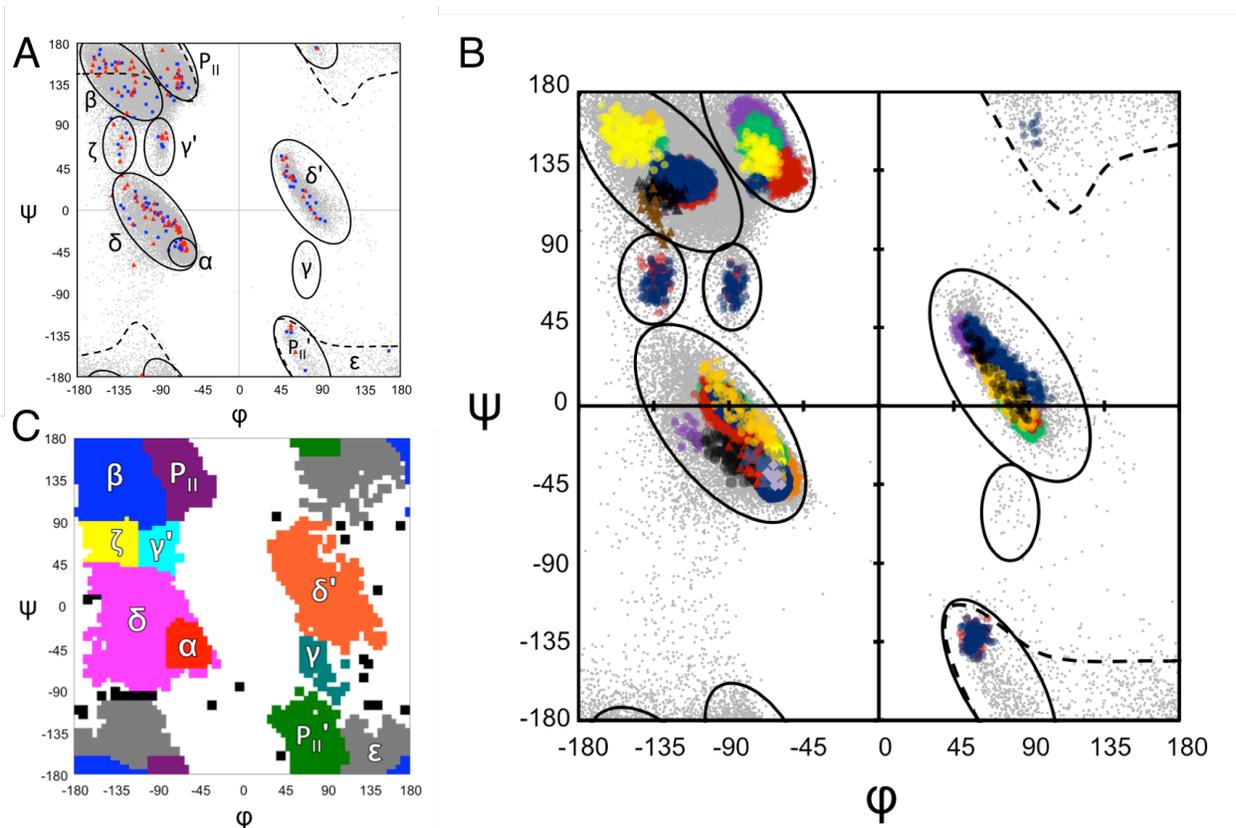


Figure 11.1. Variations in Ramachandran plot distributions. A) Ramachandran region nomenclature based on the natural distributions (e.g. Karplus (423) and Hollingsworth & Karplus (402)) overlaid on the 236,494 residue dataset used in this study. B) Displayed are the natural distributions of some of the most common protein parts with the full accompanying dataset displayed in grey and commonly described basins outlined for reference. The distributions displayed correspond to the many most prominent motifs from those identified by Hollingsworth et al 2010 (401). Each distribution was generated by using the “Everest” method¹⁹ by which natural distributions of each motif were isolated prior to the point at which the motif would be merged with another uniquely identified motif. Several motifs could not be isolated due to merger with another motif before a significant sample size could be achieved, while several are omitted due to overlaps with other motifs. The identities of each of the distributions are given by the nomenclature used in Hollingsworth et al, grouped by basin. **α and δ basin distributions:** $\alpha\alpha$.1 (both residue i and $i+1$), dark purple circles; $\alpha\delta$.1 (residue i), blue circles; $\alpha\delta$.1 (residue $i+1$), green circles; $P\alpha$.1, light green circles; $\beta\alpha$.1, orange circles; $\zeta\alpha$.1, light yellow circle; $\delta\beta$.1, black circles; $\alpha\zeta$.1, red triangles; $\delta'\alpha$.1, blue triangles; $\epsilon\delta$.1, purple diamonds; δP .1, red circles; $\delta\delta'$.1, green triangles; $P\delta$.1, purple circles; $P'\delta$.1, orange triangles. **δ' basin distributions:** $P\delta'$.1, green circles; $\delta\delta'$.1, blue circles; $\delta'P$.1, red circles; $\delta'\delta'$.1 (residue i), purple circles; $\delta'\delta'$.1 (residue $i+1$), orange circles; $\delta'\alpha$.1, black circles. **β basin distributions:** $\beta\beta$.1 (residue i), blue circles; $\beta\beta$.1 (residue $i+1$), red circles; βP .1, green circles; $P\beta$.1, purple triangles; $\beta\alpha$.1, orange circles; $\delta\beta$.1, yellow circles; $\beta\delta'$.1, black triangles; βP .1, brown triangles. **P_{II} basin distributions:** $P\alpha$.1, green circles; PP .1 (both residue i and $i+1$), purple circles; $P\delta'$.1, red circles; δP .1, yellow circles; $P\beta$.1, blue circles; $\delta'P$.1, orange circles. **P_{II}' basin distributions:** $P'\delta$.1, blue circles; $\beta P'$.1, red circles. **γ' basin distributions:** $\gamma'\delta$.1, blue circles; $\delta\gamma'$.1, red circles. **ζ basin distributions:** $\zeta\alpha$.1, blue circles;

$\alpha\zeta.1$, red circles. *ϵ basin distribution*: $\epsilon\delta.1$, blue circles. C) Exact boundaries used for mapping the observations in this data set into the 10 defined conformations as labeled; observations not categorized are in black.

This observation reveals that the thermodynamic basins are not homogeneous populations, but are an aggregate of many discrete subpopulations that depend on the ϕ,ψ -values adopted by the neighbors of the residue in question.

Illustrating this point further, we plot in Figure 11.1B sets of populations from the $(\phi,\psi)_2$ -motif analysis that show the extent to which natural distributions, especially in the broader β , P_{II} , δ , and δ' displaying distinct populations, divide the individual regions into subregions depending on the ϕ,ψ -region of the preceding or following residue. If the Flory isolated-pair hypothesis were true, each of these distributions would overlap perfectly. These results imply that there is another level of detail to the Ramachandran regions than has not been previously explored. Since a residue's populations are influenced by the conformations of both of its neighbors, an effective probe of these relationships would simultaneously take into account both neighbors. So here, we build on the earlier observation by documenting in a simple and direct manner the extent to which the conformations adopted by a central residue are impacted by the conformations of both of its neighbors.

Results & Discussion

The Dataset

The dataset for our $(\phi,\psi)_2$ -motif study, generated using the Protein Geometry Database (424) as described previously (401), was a set of 72,318 four-residue segments with trans peptide bonds and backbone B-factors $\leq 25 \text{ \AA}^2$ from structures determined at 1.2 \AA resolution or better and having less than 25% sequence identity to any other structure in the dataset. For the work

here analyzing three-residue segments, we sought a larger data set because three residue segments represent a 6-dimensional space (ϕ and ψ for each of three residues) that will be more sparsely populated than the 4-dimensional space described in the $(\phi, \psi)_2$ -motif study. We generated a larger number of observations by relaxing the resolution cutoff to 1.5 Å, but otherwise using the same search criteria. While the relaxed resolution cutoff would lead to a larger degree of uncertainty in the torsion angles used herein, the 1.5 Å cutoff will still result in only high resolution structures with a low enough uncertainty for the purposes of this study. This relaxed cutoff resulted in a data set of 236,494 segments.

The influence of conformational context.

To test how the conformational properties of a central residue depend on the conformation of its two neighbors, the new dataset was divided into populations defined by the conformations of the $i-1$ (i.e. preceding) and $i+1$ (i.e. following) neighbors. We then assigned the $i-1$ and $i+1$ residues to conformational regions by discretizing ϕ, ψ -space into $5^\circ \times 5^\circ$ bins and grouping sets of bins to create ten regions that represent the commonly defined populations (Figure 11.1C). These conformations are referred to using the shorthand designators α , δ , β , P_{II} , γ' , ζ , ϵ , P_{II}' , δ' , and γ . An eleventh designator 'X' was used to imply a wild card, meaning the conformation of the residue could be any one of the 10 regions. In all, 99.2% of the residues in the data set (or 234,715 residues), had both residue $i-1$ and $i+1$ in one of the ten categories and contributed to our analyses. With each of the neighbors having 10 possible categories, the central residues included in the analysis could be in one of 100 ($= 10^2$) conformational contexts. The contexts are described by a shorthand such as δ -X- α indicating the central residue can adopt any conformation ('X'), while the residue before and after are limited to a particular conformation, in

this case δ and α , respectively. The populations of residues in each of the 100 contexts are provided in supplementary Table 11.1, and images of the ϕ, ψ -distributions of all 100 populations are in Appendix F.

Table 11.1. Number of observations for each of the 100 possible conformational contexts of the central residues. The first column and top row define the context by giving, respectively, the conformational region of the residue preceding ($i-1$) and following ($i+1$) the central residue. The final row and column, labeled with an “X” reports the sums of the entries in that row or column.

$(i+1) \backslash (i-1)$	α	β	δ	ϵ	γ	P_{II}	ζ	δ'	γ'	P_{II}'	X
α	59596	7951	6160	596	40	7073	513	1457	605	524	84588
β	3601	34525	8480	1118	49	12846	495	4229	480	686	66589
δ	11253	6203	3933	376	21	4552	495	1279	292	330	28776
ϵ	384	1245	420	58	5	606	35	144	23	51	2976
γ	27	62	30	9	0	22	1	2	3	1	158
P_{II}	3897	11877	6659	515	29	7198	347	3101	372	478	34514
ζ	597	688	319	38	4	331	18	91	10	24	2124
δ'	3385	2815	1879	168	8	1635	181	406	100	108	10695
γ'	316	615	398	28	3	343	28	155	28	42	1958
P_{II}'	351	793	459	58	5	430	32	132	33	43	2345
X	83462	66860	28773	2973	164	35074	2153	11018	1948	2290	-

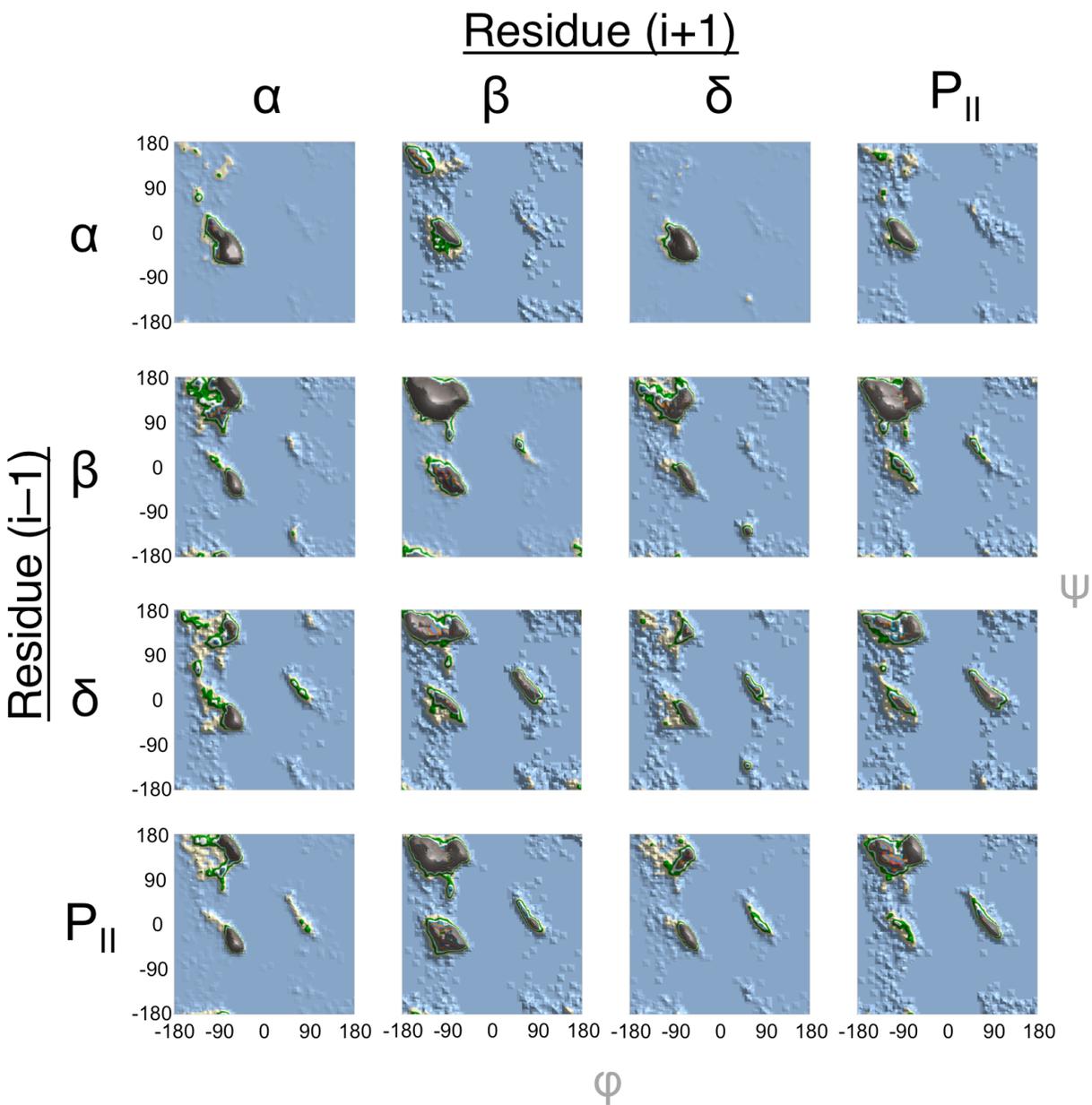


Figure 11.2. Dependence of central residue conformational on neighboring residues.

Conformations adopted by a central residue (residue i) in the context of neighboring residues on either side ($i-1$ and $i+1$) that are in one of the top four most populated regions of conformation as defined by the mask in Figure 1B. This subset of 16 tripeptides accounts for 82.79% of the full dataset. The populations of each tripeptide plot are as follows: $\alpha X \alpha - 59,596$; $\alpha X \beta - 7,951$; $\alpha X \delta - 6,160$; $\alpha X P_{II} - 7,073$; $\beta X \alpha - 3,601$; $\beta X \beta - 34,525$; $\beta X \delta - 8,480$; $\beta X P_{II} - 12,846$; $\delta X \alpha - 11,253$; $\delta X \beta - 6,203$; $\delta X \delta - 3,933$; $\delta X P_{II} - 4,552$; $P_{II} X \alpha - 3,897$; $P_{II} X \beta - 11,877$; $P_{II} X \delta - 6,659$; $P_{II} X P_{II} - 7,198$. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

Because of the dominance of the α , β , δ , and P_{II} regions, over 82% of the data are accounted for by the 16 conformational contexts in which a residue has neighbors in just these four regions. Again, if the Flory isolated pair hypothesis were true, then each of the 16 distributions in Figure 11.2 would be identical. However, a number of substantial differences are observed. For example, focusing on the P_{II} region, some contexts, like β -X- α , roughly match the conventional P_{II} distribution, but others context display different distributions such as P_{II} -X- δ residues show a preference for a larger phi and smaller psi values, δ -X- α residues mostly occupy the central part of the region, and the α -X- α and α -X- δ residues which only minimally occupy the P_{II} region but show an even lower preference for other major regions such as β or δ' . The origins of these preferences can also be studied by comparing the conformation preferences displayed in Figure 11.2 with the 101 common motifs that were identified in our original study (401). In that study, many of the most common two residue motifs that were identified were conformational caps, that is a P_{II} residue that followed at the beginning or end of a linear group such as an alpha helix or beta strand. In the present distributions, such conformational caps can be observed to heavy bias the preferred conformation in the P_{II} -X- α context where the central residue is most commonly a second P_{II} as a continuation of the conformational cap or a α conformation which would signify the beginning of a turn or helix. Similar patterns can be observed after α residue in the α -X- P_{II} or β -X- P_{II} contexts, reinforcing the importance of the conformational cap motifs that were identified previously.

Similar differences can be seen in other broad regions throughout the plots, which can be tied back to previous studies for further context. The β -region is similarly split with some contexts, such as β -X- α , showing a preference to occupy a small region with lower phi and higher psi values while P_{II} -X- β tripeptides occupy the full β -region (with the two distinct peaks

corresponding to the lower and upper beta regions noted previously to exist as linear groups (425)). Similarly, the δ' region shows distinctive subregions dependent on neighboring conformations with the β -X- β and β -X-P_{II} conformations strongly preferring the upper (high psi) portion (mirroring the α region) and others mostly populating the lower portion or tail of the δ' region. This upper subregion within the δ' basin, the left-handed mirror for the small α region as defined in Figure 11.1C, was found most commonly associated with β -bulge motifs that insert a δ' , particularly a residue centered in this upper subregion inside a larger β -strand (401), helping to explain the increased preference for this subregion over the full basin.

Smaller regions in conformational space also show dependence on context, though the differences cannot be as striking as those in broadly populated regions. Both the α and δ regions are highly dependent on local hydrogen bonding patterns which can highly restrict the available conformations as seen in the α -X- α tripeptide which is dominated by helical structures such as the α - and 3_{10} -helix. In the same context the ζ or pre-Pro region is a strong feature while it is not highly populated on a global scale. This shift in distribution is due to the presence of proline-based helix caps in such helical contexts such as the L-shaped $\alpha\alpha$ turn and proline α -C-Cap as identified in previous work (401). In a similar fashion, the relatively minor γ' region is more heavily pronounced in the β -X- β context than either global distributions or any of the other distributions in Figure 11.2. In this example, the shift is due to single residue γ' turn found between anti-parallel β -strands, which previous analysis found to be the most common occurrence for residues found within the γ' region (401).

The potential value of accounting for conformational context in modeling

That the populations filling the major basins are a composite of many subpopulations with distinct behavior is little appreciated. However, now with large numbers of high accuracy protein structures available, it is increasingly possible to define the discrete subpopulations and understand their origins. Especially noteworthy is that the extent to which the conformational context of a residue influences its conformational preferences is at least as great as the extent to which the amino acid types of neighbors influence conformation. Such distributions as those shown in Figure 11.2 can be converted to pseudopotentials that can aid protein modeling in a wide variety of areas. Two examples of such fields are protein structure prediction/modeling and protein structure refinement programs.

Rosetta has emerged as one of the most powerful tools in not only modeling protein interactions, but also in the modeling of unknown loop regions and the prediction of protein structure (426-428). During the model building process, successive fragments of protein structure from a local library are added to the growing polypeptide which are further refined by a Ramachandran potential that pushes residues to fit a global basin (400, 426). One could hypothesize that by having multiple sub-basins based on a neighboring residue context would increase the accuracy of the structure prediction. Recent work by Ting and coworkers carried out such a context-based approach by analyzing Ramachandran distributions of a central residue based on neighboring residue type (i.e. amino acid identity) in loop regions of proteins. By identifying the distinct differences due to neighboring residue type, the authors were able to demonstrate that by incorporating neighbor identify, protein loop conformation prediction was significantly improved over the use of a single global Ramachandran basin (400). When comparing recent analysis of neighboring residue type context focusing on loop regions (400) and that of the present neighboring conformational context for all residues, the observed

differences in Ramachandran distribution variation by taking into account conformational context are at least as striking than differences observed by accounting for residue type context. Further supporting this hypothesis, Ting et al. noted that residues preceding proline showed that largest effect on Ramachandran distributions, stating “Pro[line] on the right exerts the largest effect and this has been well-studied previously” (400). This observed large effect on a pre-Pro residue is in fact a conformational dependent effect, which hints that further conformational dependent effects described herein will continue to be more striking than previously documented residue type context effects. These findings would suggest that by taking conformational context into account, similar to what has been done for residue type effects, the increase of accuracy of structure prediction programs would be larger than that observed by Ting et al.

Similar to Rosetta, the automated protein structure refinement package PHENIX initially used a single global Ramachandran potential that was applied to all residues regardless of identity or neighbor context (429). Recent work to improve Ramachandran accuracy by using four separate basins based on residue identity (430) (general, proline, pre-proline and glycine) was shown to improve the performance of PHENIX refinement (431). By expanding the Ramachandran basins to take neighboring residue conformation into account through the application of context-dependent sub-basins, the accuracy of the torsion angle refinement should increase. Similar increases in accuracy have already been observed with comparable context driven sub-basins for refinement (432). Such information could be incorporated as a late stage refinement procedure once the identity of the conformation for each residue has been roughly determined, at which point the refinement could shift from a generic description for conformation to a neighbor environment delimited population function based on the ten primary basins in the Ramachandran plot. As the number of available high-resolution structures continues

to grow, it should be possible to go beyond broad ϕ, ψ -space to capture the more detail similar to the classification of common motifs (401) in our previous studies once these databases are more populated.

Chapter 12

On the Docking Dynamics of the GRP1 PH-Domain Bound to a Target Lipid Bilayer

Summary

The pleckstrin homology (PH) domain plays a key role in a variety of cell process from signaling to apoptosis by binding to a second messenger lipid PIP₃ (phosphatidylinositol-3,4,5-trisphosphate) in the inner leaflet of the plasma membrane. Recently, using EPR side-directed spin labeling, a model was proposed for the docking geometry of a PH domain bound to a PIP₃-modified lipid embedded in an anionic phospholipid bilayer which was studied using molecular dynamics. We further test the viability of this proposed model using atomistic molecular dynamics simulations focusing on PH-lipid interactions. Our simulations are consistent with the experimental depth parameters deduced from the EPR data over a 500 ns trajectory, and provide further insight when compared to the results of previous MD studies by revealing two separate stable conformations in complex with the bilayer. The specific interactions between the PH domain and the membrane bilayer, which have not been described previously in detail, help support the recently proposed hypothesis of the importance of four different loops that bind to the membrane bilayer.

Introduction

One of the hallmarks of any cell-signaling pathway is the ability of those proteins involved to recognize modified lipids. The ability to recognize and bind to these target lipids is key to the cells ability to interpret a signal and transmit it beyond the cell membrane. One such modified lipid is the second messenger phosphatidylinositol-3,4,5-trisphosphate (PIP₃), which is

generated by phosphoinositide-3-kinase (PI3K) during cell signaling (433-435). Proteins recruited to the plasma membrane by the presence of PIP₃ modified lipids contain, predominantly, the PIP₃-specific pleckstrin homology (PH) domain (403, 404). The PH domain is widespread, occurring in over 500 proteins in humans playing a variety of roles, including membrane remodeling, DNA synthesis, cell growth and apoptosis (436-439). Loss or alteration of the PIP₃ specificity in the PH domain has recently been shown to lead to disease including cancer in humans (405).

Despite the widespread importance of the PH domain and its interaction with its target PIP₃ modified lipid, little is known regarding the lipid-protein configurations and interactions at the docking interface. Recently, there has been a wealth of information that has begun to shed light on this interaction including the crystal structure of various PH-domains in complex with a PIP₃ analogue, NMR studies of the PH domain bound to a small micelle and molecular dynamics studies of the PH domain bound to a PIP₃ on a simple lipid bilayer (440, 441). Recently, this information was expanded to include Brownian dynamics studies of the PH domain binding to the target lipid to a simple lipid bilayer (441). None of these studies, however, have been able to probe the docking interface in the presence of a complex lipid bilayer.

Recently, EPR spin labeling and relaxation methods were used to determine the depth of the PH domain bound to a target bilayer and its angular orientation relative to an ideal phosphate plane, providing further insight into the docking orientation of the PH domain (406). The study by Lietzke et al used the PIP₃ specific PH domain of the General Receptor for Phosphoinositides 1 (GRP1, NCBI Gene ID 9265, CYTH3) (440). GRP1 is an Arf6 guanidine-nucleotide exchange factor that involved in the catalyzed activation of Arf6-GDP to Arf6-GTP at the surface of the plasma membrane. The crystal structure of this domain was solved at 1.54Å-resolution in

complex with the PIP₃ analogue PIP₄, which was used in the construction of the EPR derived binding model. This work was then used as a basis in a atomistic molecular dynamics simulation to probe both how the PH can find the lipid bilayer and tested the viability of the EPR derived PH Domain model (442). This new study revealed the importance of the negatively charged POPS lipids in the lipid bilayer to the PH domain binding to the bilayer. In the atomistic EPR simulations, four loops were identified to be important in lipid binding. While this study was able to identify the loops that make contact with the lipid bilayer and the minimal degree with which the PH domain was found to bind to the membrane bilayer, the individual residues in these loops that were found to make contact with the bilayer were not studied. Understanding residue-by-residue the contact that exists between the PH domain and the membrane bilayer would shed a new light on the complexities of these interactions and provide new possible avenues for research.

Here we analyze the proposed docking model in explicit membrane environment using atomistic molecular dynamics simulations. We find that the PH domain does not undergo major conformational changes and it remains bound to the membrane, at a depth and with an orientation that are within the EPR experimental errors. Furthermore, the experimentally determined EPR depth parameters are consistent with the simulation data. The contacts made throughout the simulations reveal the docking interface is stabilized by specific interactions between amino acid residues and both anionic and zwitterionic lipids.

Methods

The crystal structure of the GRP1 PH domain (PDB ID 1FGY) bound to inositol 1,3,4,5-tetraphosphate (IP₄) was docked onto a 3:1 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) to

1-palmitoyl-2-oleoyl-phosphatidylserine (POPS) bilayer in excess water with the orientation and depth prescribed by the model of Cheng et al. (406). The IP₄ moiety of the crystal structure was used to build a modified PIP₃ DPPI lipid molecule as docking site replacing a POPC molecule. To improve sampling and to prevent asymmetric distortions to the bilayer, a PH domain was docked onto to each membrane bilayer leaflet (Figure 12.1). The final system consisted of 256 lipids, two proteins of 126 residues each, 21,498 waters molecules and 74 counterions for a total of 115,375 atoms. The proteins were positioned in an offset fashion to avoid any interactions with each other through either periodic boundaries or through the bilayer itself.

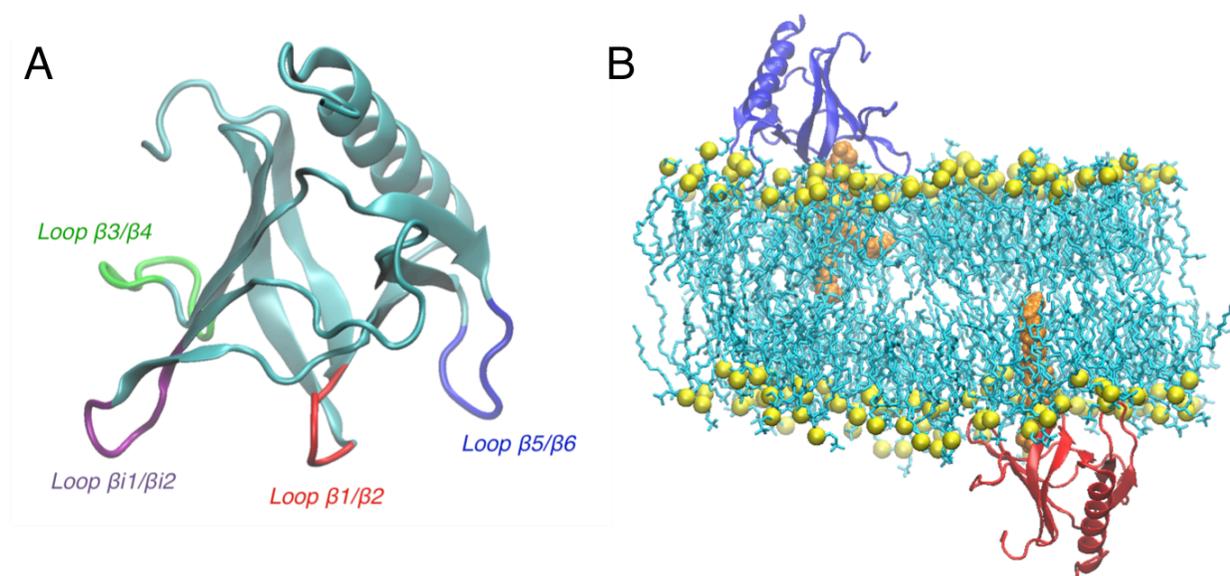


Figure 12.1. GRP1 PH domains bound to a target lipid bilayer. A) Structure of the PH domain. The overall structure and fold of the PH domain is displayed in panel a. The four loops that have been predicted to make contact with the lipid bilayer are highlighted and labeled. B) Snapshot configuration from the end of the simulation trajectory. To prevent asymmetric distortions of the target lipid bilayer, one PH domain (shown in secondary structure representation) was docked onto each bilayer leaflet according to the model of Chen et al. (406). The POPC and POPS lipids are colored cyan with their phosphate P atoms shown as yellow spheres. The PIP₃-modified DPPI lipids are shown in orange in filled-sphere representation.

The simulation system was relaxed using 5000 frames at a 1 femtosecond time step of conjugate-gradient energy minimization followed by a constant volume simulation at 300 K over 30 ps with the protein fully restrained. An additional 470 ns were then collected with the

restraints removed, for a full 500 ns trajectory generated at 1 atm and 300 K. To ensure that during the initial stage of the simulation, the docking would match the proposed model of Chen et al. with the improvement of a fluid complex bilayer in place of an idealized phosphate plane, harmonic restraints were placed on the backbone of each PH domain and the headgroup of the DPPI lipid for the first 30 ns using a timestep of 1 fs for the first 10 ns and 2 fs thereafter. The restraints were released stepwise over the next 4 ns, at which point the system was allowed to run freely. All simulations were performed with the NAMD 2.8 software package (443). The CHARMM22 and CHARMM36 force fields were used for the protein and lipids, respectively, the PIP₃ headgroup was modeled using the CHARMM36 force field for carbohydrates, and the TIP3P model was used for water (444, 445). The smooth particle mesh Ewald method was used to calculate electrostatic interactions. Short-range real-space interactions were cut off at 12 Å, employing a switching function (177). A reversible multiple time-step algorithm was employed to integrate the equations of motion with a time step of 4 fs for electrostatic forces, 2 fs for short-range non-bonded forces, and bonded forces. All bond lengths involving hydrogen atoms were held fixed using the SHAKE and SETTLE algorithms (382). A Langevin dynamics scheme was used for temperature control and a Nosé–Hoover–Langevin piston was used for pressure control(237). Molecular graphics and trajectory analyses were performed using VMD 1.8.9 (179).

Results

Two PH domains were aligned in an offset fashion on a membrane bilayer as shown in Figure 12.1 with each PH Domain studied separately to provide better sampling during the study of PH domain & lipid bilayer contacts. During the 500 nanosecond simulation, both domains

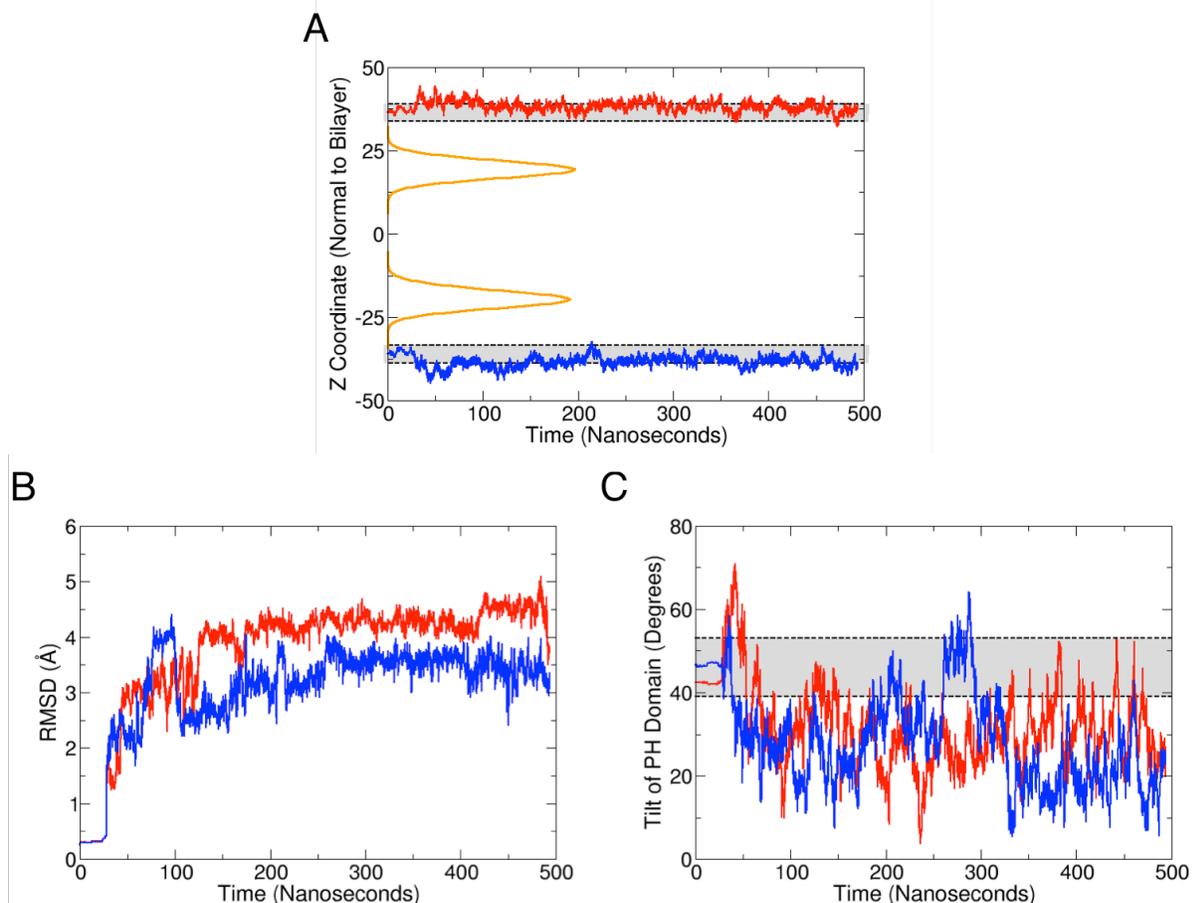


Figure 12.2. The EPR-derived model is stable in long molecular dynamics simulations. In all panels, red lines correspond to the PH domain placed above the lipid bilayer, while blue corresponds to the domain placed below the bilayer, corresponding to Figure 1B. A) Plots of RMSD versus time for each of the PH domains in the simulation. B) The evolution of the PH domains center of mass position along the transmembrane direction is shown using the same color scheme as in A (light and dark color tones correspond to the restrained and unrestrained portions of the trajectory, respectively). The PH domain depth and orientation in the docking model of Chen et al.(406) were reported using the plane described by the mean position of the lipid phosphate moieties as a reference. To compare with the COM trajectories, we take the corresponding widths of the lipid phosphate distributions as a measure of the uncertainty (shown as grey bars). The corresponding lipid phosphate distributions are shown in yellow. C) The evolution of the PH domains orientation is assessed by the tilt angle of the vector form by the α -carbons of C292 and F296, which is the operational definition of the axis of the PH domain core β -sandwich used by Chen et al. Both the depth and orientation of the PH domain in the docking model of Chen et al. are conserved along the simulation trajectory.

remained identical until the restraints were removed after the initial constrained 30 nanoseconds and underwent no major conformational changes throughout the remainder of the simulation as

evident in the RMSD plots in Figure 12.2A. The change in RMSD was found to be due to changes in the loop regions of the domains and not in the secondary structures. As both domains were found to be stable in the presence of the mixed bilayer, the center of mass of each domain with respect to the center of the membrane bilayer was computed throughout the simulation (Figure 2B). The center of mass analysis revealed that both domains remained bound to the bilayer throughout the simulation. In addition, the location of the center of mass remained near the EPR predicted location, represented within one standard deviation in Figure 12.2B as the grey bars. Finally, to study the tilt angle of each domain, the tilt angle of a vector described by α -carbons of C292 and F296 and the membrane normal was computed for every frame of the simulation and plotted in Figure 12.2C. Upon the beginning of the simulation, the different membrane environments caused the tilts to diverge between the domains. The tilt angles of both domains were found to stable at a smaller angle than previously predicted, but both domains stabilized at similar angles.

To further test whether the domains remained similar to the original EPR model orientations, the previously calculated depth parameter was plotted against the distance from the mean phosphate plane for both the initial restrained 30 nanoseconds and the final 100 of the full 500 nanosecond simulation (similar to the plot found in Chen et al. (406)). Figure 12.3 shows the restrained and unrestrained data for both PH domains. Each domain provides a different characteristic trend line though both sets show very little difference between the restrained and unrestrained portion of the simulation, as no unrestrained point is beyond one standard deviation of the corresponding restrained point. The major difference between the PH domains in the simulations lie in the penetration of each domain into the bilayer as the positions of those residues near the bilayer are closer to the center of the bilayer than the other domain.

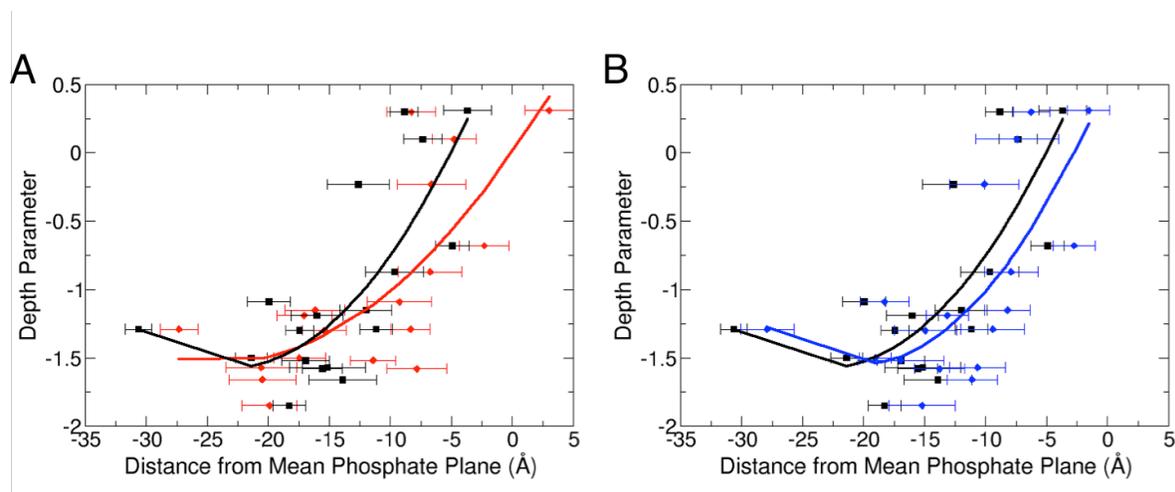


Figure 12.3. Quantitative comparison between spin label depth parameters and membrane penetration depth in the simulation. The experimentally determined depth parameters(406) are plotted against the distance of the C β atoms from the mean of the lipid phosphate distribution. The red dots and line represent the retrained portion of the simulation trajectory and therefore represent the original model with the error bars representing the standard deviation during that time span. The black lines represent the same residues but during the last 100 ns of the simulation with the restraints removed and the model equilibrated. At every point there is strong overlap between the theoretical model and the unrestrained simulations, providing further support for the proposed binding model.

Focusing on lipid penetration and the PH domain-lipid contacts that are formed upon complex formation, time averaged contacts per frame were calculated over the final 100 nanoseconds of the simulation for each domain. The residues that were found to make at least one contact per frame are highlighted in Figure 12.4, which breaks down the contacts by which part of the lipid the domain is making contacts with using the definitions of Lai et al. The majority of residues that were found to make contact with the bilayer were located in one of the four loops that were highlighted in a previous study and proposed to be involved in not only bilayer binding but in the domain's ability to crawl along the bilayer towards negative charge (Figure 12.1A). Between the two domains, there are differences between the number of contacts per residue and which of these four loops are favored. In PH1, the β 3/ β 4 loop made minimal contacts while in PH2 it made contacts with the lipid headgroups, while the β i1/ β i2 loop made

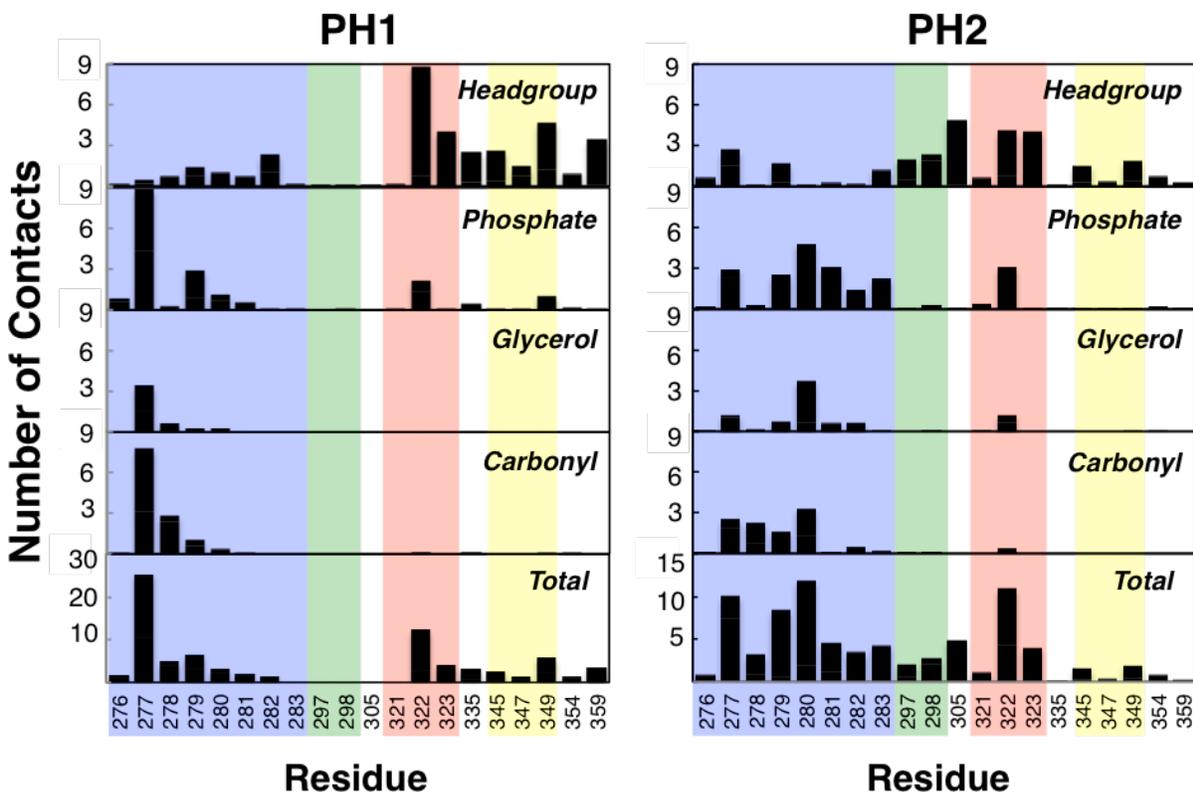


Figure 12.4. Major residues of the PH domain that make contact with the lipid bilayer. Time averaged contacts between the PH domain and the lipid bilayer are displayed, showing only those residues that were found to average at least one contact pair for either PH domain over the last 100 nanoseconds. The residues are further categorized by highlighted color corresponding to the four major loops that had been predicted to make contact with the lipid bilayer: Loop $\beta 1/ \beta 2$ – blue, Loop $\beta 3/ \beta 4$ – green, Loop $\beta 5/ \beta 6$ – red and Loop $\beta 1/ \beta 2$ – yellow.

more contacts with the bilayer in the PH1 domain when compared to the PH2 domain. In both domains the $\beta 1/ \beta 2$ loop makes the most contacts with the lipid bilayer as seen in the bottom panels of Figure 12.4. Also consistent with Figure 12.3, PH1 made more contacts with the lipid bilayer beyond the headgroups of the lipids, particularly with residue R277 which plays not only a major role in PH-lipid contacts, but has been shown to be involved with binding the PIP₃ group itself. Looking deeper into the nature of these contacts, they were further broken up by the species of lipid with which the PH Domain was making contact, the results of which can be seen in Figure 12.5. It is apparent that the PH Domain makes more contacts with the POPS lipid than

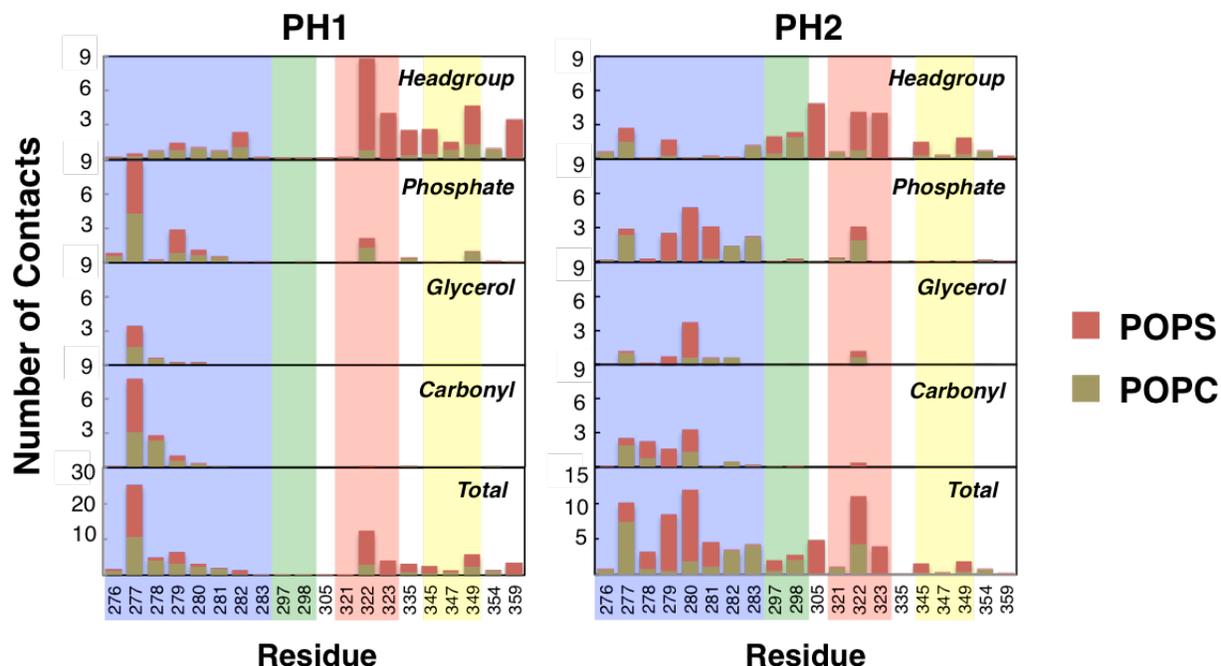


Figure 12.5. Major residues of the PH domain that make contact with the bilayer by lipid species. Time averaged contacts between the PH domain and the lipid bilayer are displayed, showing only those residues that were found to average at least one contact pair for either PH domain over the last 100 nanoseconds. Contacts are divided up into those contacts that were found involving POPS (red) and POPC (tan) lipids. The residues are further categorized by highlighted color corresponding to the four major loops that had been predicted to make contact with the lipid bilayer: Loop $\beta 1/ \beta 2$ – blue, Loop $\beta 3/ \beta 4$ – green, Loop $\beta 5/ \beta 6$ – red and Loop $\beta i1/ \beta i2$ – yellow.

POPC lipids in the bilayer, with several residue making contacts nearly exclusively with POPS lipids.

Finally, to study whether the binding of the PH domain has an affect on the diffusion of lipids in the bilayer itself upon binding, the 2D MSD in a plane parallel to the bilayer was computed for both the PH domain system and a the same bilayer in the absence of the PH domain for 500ns. The results of these calculations can be seen in Table 12.1. In the +PH system, the DPPI lipid diffusion is observed to drop for both DPPI lipid in the bilayer system shown in Figure 1. POPS and POPC lipid diffusion is observed to increase in the presence of the PH domain, though the change is not observed to the same degree in each leaflet of the bilayer.

Table 12.1. 2D MSD of Lipid Species in the presence and absence of PH Domain binding.

		2D MSD (cm ² /s) + PH Domain	2D MSD (cm ² /s) - PH Domain
<i>DPPI</i>	+z	0.2984 ± 0.0878	0.1417 ± 0.17966
	-z	0.0644 ± 0.2970	-0.0096 ± 0.0472
<i>POPS</i>	+z	0.1917 ± 0.1436	0.2003 ± 0.0055
	-z	0.3349 ± 0.0238	0.1446 ± 0.2545
<i>POPC</i>	+z	0.1579 ± 0.1666	0.1545 ± 0.0672
	-z	0.2967 ± 0.0516	0.1198 ± 0.2032

Discussion

Much work has been done to understand the ability of the PH domain to not only bind its PIP₃ target, but also to the membrane bilayer itself, including crystal structures, EPR and molecular dynamics. However, through this work only a general view has emerged on the contacts between the PH domain and the lipid bilayer at a residue-by-residue level. Here we provide further support for the EPR derived model of PH domain binding but also provide an in depth view of the residues that make contact with the bilayer upon binding to the DPPI lipid.

Figures 12.2 and 12.3 reveal that the PH domain is indeed very stable in the presence of a mixed bilayer and that no conformational changes occur upon binding to the lipid bilayer. These results help to further validate both the EPR derived model, as well as the initial computational work done by Lai and coworkers. The new simulations average more contacts than the initial simulations, though the regions of the domain that are making contact with both the PIP₃ group and the lipid bilayer are nearly identical, while both simulations agree well with the general EPR

proposed docking model. The tilt angle of the domains in the new simulation stabilized outside the predicted tilt angle, however the stabilized angle still allows each of the four major loops shown in Figure 12.1A to make contacts with the lipid bilayer. This underscores the flexibility of the PH domain to bind to and make strong contacts with the bilayer while not penetrating into the bilayer itself at a variety of tilt angles near the EPR predicted value. The major difference between the two models in this study lies in the ability of the PH domain to penetrate, particularly V278, into the bilayer. Recent work has suggested that lipid defects created by the neighboring residue 277 and 279 are necessary to create an optimum environment to allow V278 to penetrate the phosphate plane during the lipid binding process. It is important to note in these simulations that even in the absence of V278 penetration that was predicted by other studies, the PH domain is still able to form a stable complex with both DPPI and the lipid bilayer.

It was also revealed that there is a preference of the PH domain to make contacts with the negatively charged POPS lipids in the bilayer over the POPC lipids. This finding agrees well with the findings of Lai et al. that showed that the PH domain would simply bounce off a lipid bilayer unless POPS lipids were present and the inherent dipole moment of the PH domain to seek out and move towards negative charges once bound to the membrane bilayer. An analysis of the contacts made between the PH domain and the lipid bilayer reveals that all major protein-lipid contacts are of an electrostatic nature. Without these negatively charged lipid headgroups present, the PH domain would not bind nearly as strong to the membrane bilayer, which suggests that the PH Domain binding strength may be proportional to the number of negative charges present in the bilayer. Indeed, previous simulation showed that in the absence of negatively charged headgroups, the PH domain is unable to bind to the bilayer. Supporting the evidence that POPS lipid contacts are important to PH domain binding can be found in the increased diffusion

of POPS lipids in the presence of a bound PH domain, which may be due to the PH domain positive charge attracting the negatively charged lipids in the bilayer to form new contacts.

With general agreement between the residue-by-residue results with the loop contacts from Lai et al., we can also compare the nature of the PH domain binding to the lipid bilayer with the initial messenger C2 domain. Unlike the C2 domain, which causes a large deformation in the bilayer upon binding and penetrates deeply from the mean phosphate plane, the PH domain binds to the surface of the bilayer and only minimally penetrates the membrane. These two proteins represent two very different means to bind to the lipid bilayer in the same pathway.

These new results reveal that each of the four loops that had been previously identified are each making extensive contacts with the bilayer with the $\beta 1/\beta 2$ loop playing the most important role in binding to the membrane bilayer, with the other three groups playing smaller roles in comparison. The two PH domains in this simulation also underline the importance of the flexibility of the four membrane contacting loops which had been proposed previously (442). The nature of these contacts are electrostatic, lending further support to the emerging hypothesis that the PH domain is dependent on negatively charged headgroups in the lipid bilayer to bind to the membrane.

Chapter 13

Crystal Structure of *Bacillus subtilis* Arginase at 2.077 Å

Summary

Arginase is an enzyme that is found in all kingdoms of life, catalyzes the conversion of L-arginine to urea and L-ornithine through a carefully coordinated binuclear manganese active site. Competing for its substrate L-arginine with nitric oxide synthase or NOS, these two enzymes have been closely linked in both humans and bacteria is both regulating the homeostasis of L-arginine but also in mediating the reaction of the other by both pathogen and host at the sites of infection. Pathogens in particular can not only express their own arginase in order to decrease NO production by NOS in an effort to survive, but can also hijack the host's own enzymes to work against it. Given the increased focus on developing bacterial NOS inhibitors against the Gram-positive model organism *Bacillus subtilis*, we have chosen to turn our focus to the *B. subtilis* arginase enzyme in the first step of studying how inhibitors of one enzyme may affect the other. To this point, we have solved the crystal structure of *B. subtilis* arginase to 2.077 Å resolution. The resulting structure, which demonstrates a large degree of conservation with previously determined arginase structures, can be used a start point for future studies of such inhibitors off-target effects.

Introduction

Arginase is a binuclear manganese metalloenzyme that catalyzes the reaction of L-arginine to urea and L-ornithine in the urea or Krebs-Henseleit cycle (408, 409, 446, 447). This catalytic conversion is carried out through the attack of a metal coordinated hydroxyl ion to cleave the guanidino group from the substrate L-arginine (Figure 13.1). In humans, two different

isoforms of arginase have been extensively studied; arginase I and II. Arginase I exists as a trimeric cytosolic protein found primarily in the liver that acts in the final steps of the urea cycle (407, 408, 448, 449). Meanwhile, arginase II exists as a trimeric mitochondrial protein that is expressed in numerous tissues including kidney, brain and muscle (408, 450, 451). Unlike arginase I, arginase II functions not in the urea cycle but in the homeostasis of L-arginine which in turn regulates the activity of other enzymes that metabolize L-arginine, most notably nitric oxide synthase or NOS (407, 409, 451-456). One such example of this tight-knit relationship between arginase and NOS lies in the inducible isoform of NOS (iNOS) as extensive work has shown that the activities of these two enzymes at both the site of an infection and in macrophages are closely related and may mediate the reactions of the other (407, 455, 457-461). Importantly, while the K_m for L-arginine binding to NOS is in the μM range compared to the weaker mM range for arginase, at physiological conditions the V_{max} of arginase is 10^3 times higher than NOS which result in each enzyme consuming roughly an equivalent amount of substrate in their competition for substrate (407).

This interplay between arginase and nitric oxide synthase extends into bacterial systems where pathogens have been shown to turn the host's own arginase against it to reduce production of nitric oxide (NO) through increased competition with NOS (407). NO is a central reactant in macrophages as well as an effective antimicrobial agent against pathogenic agents such as *Heliobacter pylori* and *Leishmania major* (407, 462). In the Gram-negative bacteria *Heliobacter pylori* has been shown to quench L-arginine from host iNOS through both the expression of its own arginase (463-465) as well as inducing expression of host arginase II (466). Similarly, the protozoan parasite *Leishmania major*, which would otherwise be exposed to the potentially

damaging NO species through NOS activity, instead hijacks the host arginase pathway to enhance polyamine production through an increased L-ornithine concentration essentially

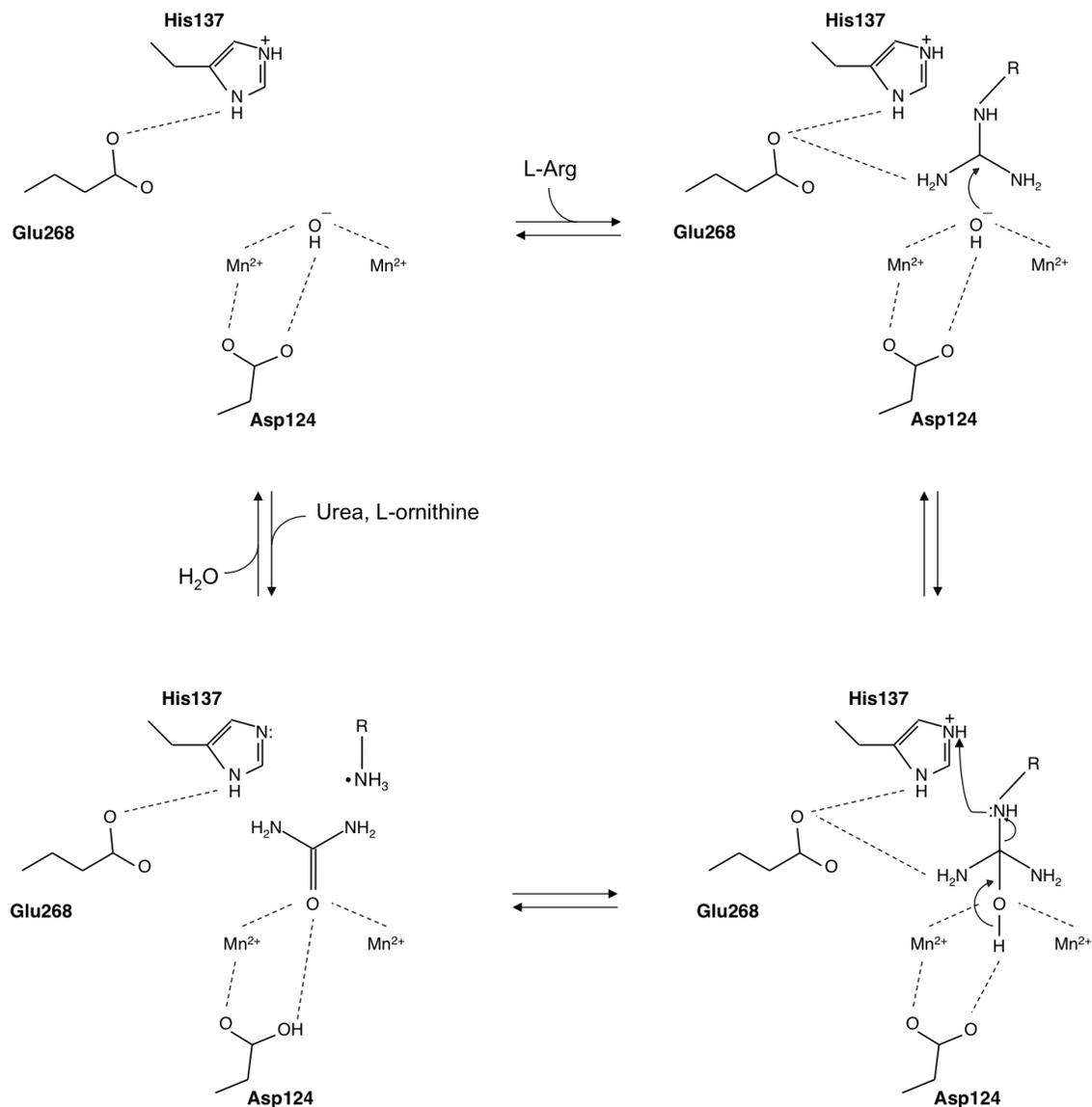


Figure 13.1. The Arginase catalytic mechanism. The enzyme Arginase catalyzes the conversion of L-arginine to urea and L-ornithine through attack of the metal bridging hydroxide ion on the guanidinium group of L-arginine. In the active site, Glu268 aids in coordinating the substrate while His137 acts as a proton shuttle donating a proton to the resulting L-ornithine and abstracting a proton from an entering water molecule that results in the return of a metal-bridging hydroxide ion.

causing the host to further compete with itself in order to produce NO to combat the pathogen (467, 468).

Interestingly, it is possible for bacteria to attempt to bias this delicate balance in their favor by expressing their own bacterial NOS (bNOS) enzyme (469-471). In Gram-positive bacteria, such as *Bacillus subtilis*, NO has been shown to act in both protection from oxidative stress as well as in the detoxification of antimicrobials (112, 119, 472). Organisms such as *Bacillus anthracis*, *Staphylococcus aureus* or *Bacillus subtilis* express their own bNOS in order to up regulate catalase activity in the host which in turn suppresses otherwise harmful Fenton chemistry (472, 473). Given the obvious strategy of inhibiting Gram-positive bacteria growth by removing their ability to produce NO, there has been an increased body of work focusing on developing bNOS specific inhibitors using the model organism *Bacillus subtilis* (116-119).

B. subtilis, in addition to its bacterial NOS enzyme, also has an arginase that has not been well characterized. Given the increased focus on the function of bNOS in the Gram-positive bacteria *B. subtilis*, it is important that we achieve a greater understanding of the interplay between bNOS and the corresponding bacterial arginase. As both enzymes share a common substrate, it follows that there may be some cross-target effects between the designed inhibitors binding to their intended target bNOS and arginase. To this point, we have expressed, purified and solved the structure of *B. subtilis arginase* to 2.077 Å resolution. The resulting structure shows a large degree of conservation in both overall fold and active site geometry with previously determined arginase structures and assembles into a hexameric quaternary structure as seen previously only in the *Bacillus caldovelox* arginase structure. The active site of the bacterial arginase enzyme is significantly different than the bacterial NOS active site which decreases the

possibility that the large, bNOS designed inhibitors bind to Arginase, though future work will need to be carried out to investigate any possible cross-target effects in more detail.

Methods

Expression and Purification

The known sequence for *B. subtilis* arginase was cloned into a pET28a vector and overexpressed in *E. coli* BL21(DE3) cells with a fused 6XHis-Tag. Expression was induced at an OD₆₀₀ of 0.57 with 0.5mM IPTG for 12 hours prior to lysing the cells. The lysed sample purified over a Ni-affinity column in 100 mM Hepes (pH 7.4), 10% glycerol and 10mM MnCl₂, followed by size exclusion chromatography using 20 mM MOPS (pH 7.5). The resulting purified protein, concentrated to 1.67 mM, was stored prior to crystallization at -80 C.

Crystallization

Small rod-shaped crystals of Arginase were grown using the sitting drop vapor-diffusion method at room temperature where the drops contained equal volumes of purified arginase at 28 mg/mL and reservoir solution (90 mM BisTris (pH 7.7), 10mM Bis Tris (pH 9.0), 10mM MnCl₂, 11% PEG8000, and 6% PGA). These initial hits were then used to produce a seed stock for multiple rounds of seeding whereupon the protein concentration was dropped to 14 mg/mL while the reservoir solution was held constant. Crystals were harvested and cryoprotected in well solution with 70% glycerol v/v. Crystals were then flash frozen using liquid nitrogen.

Structure Determination

Diffraction data was collected at SSRL beam-line 14-1. The resulting datasets were pre-processed and scaled by MOSFILM (474). Initial phases were determined by molecular replacement (MR) using chain A of 2CEV (475) by phaser (244) in RefMac (476). The initial MR model was modified by the PHENIX (429) plugin Sculptor to mutate the previously known structure to match the primary sequence of *B. subtilis* arginase. The resulting model was refined through sequential rounds of refinement using PHENIX implementing non-crystallographic symmetry restraints as well as TLS protocols (247) with each chain as a single TLS group. The final structure was visually validated with COOT (246) and statistics generated using PHENIX (Table 13.1).

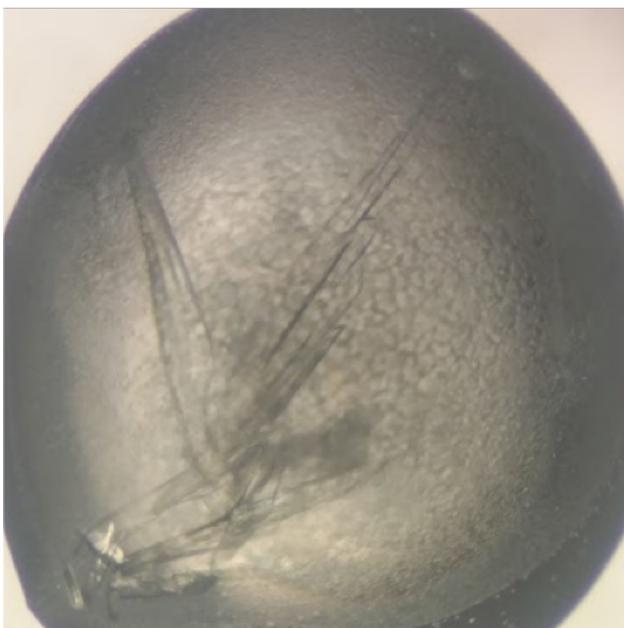


Figure 13.2 Crystals of *Bacillus subtilis* arginase. Shown are the long rod-like crystals of arginase that were obtained following multiple rounds of seeding. The rod on the top left was the crystal of which diffraction data was collected.

Table 13.1 Arginase data collection and refinement statistics

	Arginase
Wavelength	0.979493 Å
Resolution range	46.08 – 20.77 (2.151 – 2.077)
Space group	P 1 21 1
Unit cell	77.1399 128.55 82.02 90 112.69 90
Total reflections	168905 (16255)
Unique reflections	88598 (8819)
Multiplicity	1.9 (1.8)
Completeness (%)	100 (100)
Mean I / sigma(I)	6.26 (1.92)
Wilson B-factor	28.04
R-merge	0.05977 (0.2799)
R-meas	0.08453 (0.3958)
CC_{1/2}	0.993 (0.803)
CC*	0.998 (0.944)
Reflections used in refinement	88581 (8815)
Reflections used for R-free	2009 (209)
R-work	0.2044 (0.2609)
R-free	0.2314 (0.3155)
Number of non-hydrogen atoms	13125
Macromolecules	12578
Ligands	12
Protein residues	1692
RMS (bonds)	0.004
RMS (angles)	0.74
Ramachandran plot	
Favored (%)	95
Allowed (%)	3.6
Outliers (%)	1.8
Rotamer outliers (%)	7.8
Clashscore	7.45
Average B-factor	40.52
Macromolecules	40.61
Ligands	25.89
Solvent	38.71
Number of TLS groups	6

Statistics for the highest-resolution shell are shown in parentheses

Results and Discussion

Overall structure of B. subtilis arginase

B. subtilis arginase itself assembles into a hexamer in the unit cell of the resulting crystal structure (Figure 13.3A). Each monomer is constructed of a central beta-sheet consisting of eight parallel beta strands packed on either side by three and four alpha-helices respectively with two Mn^{2+} ions located near the C-terminus of the central beta strands at the protein surface (Figure 13.3B). The overall symmetry and structure of the bacterial arginase is very similar to the previously solved structure of *B. caldovelox* arginase in both the oligomerization and monomeric structures that also found that at pH 7.0 the Gram-positive bacterial arginase exists as a hexamer (475).

Focusing on the active site (Figure 13.3C and D), the two Mg ions each retain the unique geometries that have been previously observed. The metal ion furthest from solvent exists in a square pyramidal geometry, coordinated by His97, Asp120, Asp124, Asp223, and a coordinated solvent molecule which would represent the known bridging hydroxide ion. The second metal ion, closest to the surface of the protein, exhibits a distorted octahedral geometry, coordinated by Asp120, His 122 Asp223, Asp225 (in a bidental fashion), and the bridging hydroxide ion. Interestingly, following refinement, the bridging hydroxide ion is located further from the metal ions than previous structures have shown, with metal to oxygen distances of 2.6 and 2.7 Å where previous structures have found similar distances of 2.4 Å. This difference in distance however may fall within the range of uncertainty given the resolution of the structure here. Finally, while Asp124 coordinates the first metal ion, the remaining side chain oxygen atom is 2.9 Å from the bridging hydroxide ion that during catalysis it will temporarily coordinate.

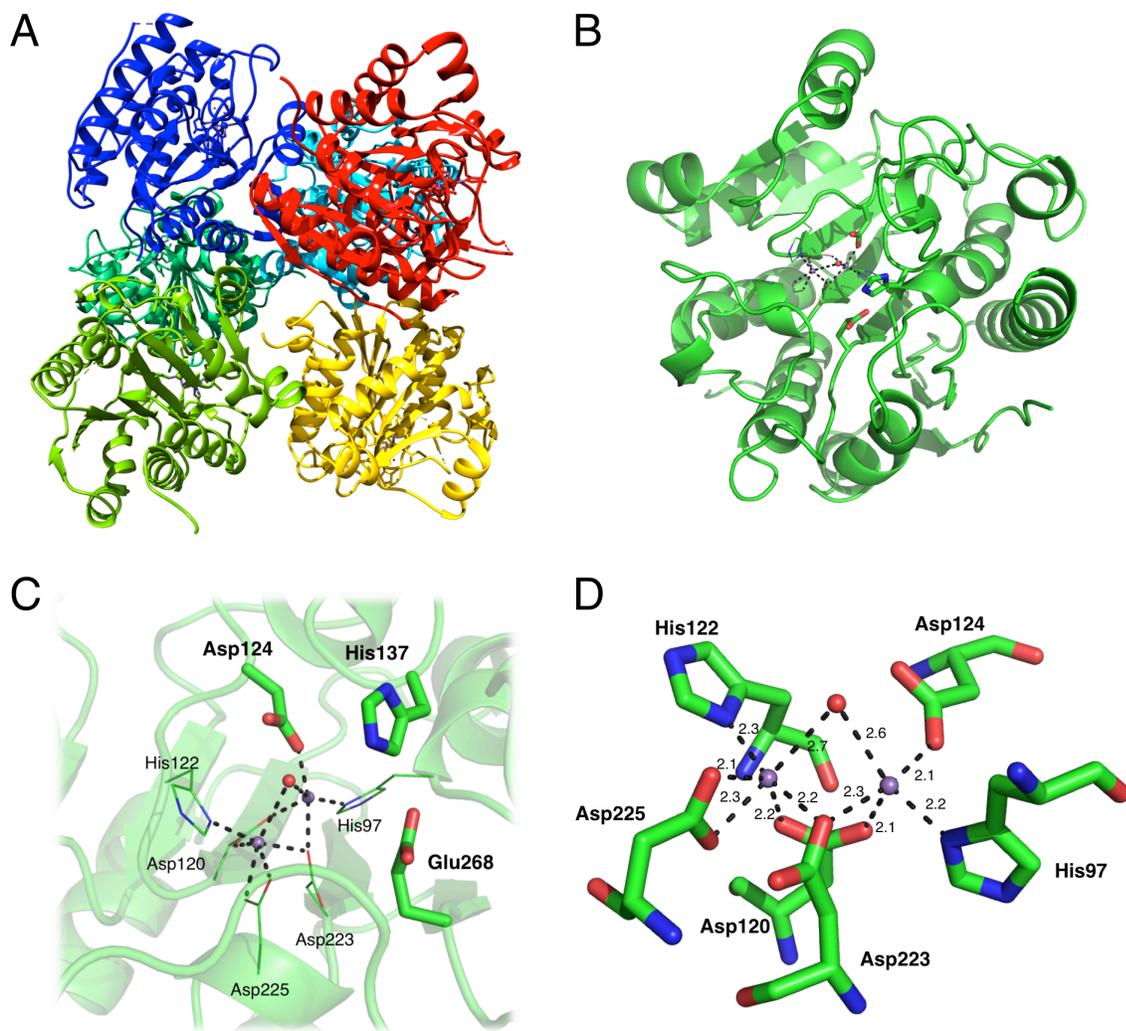


Figure 13.3. Crystal structure of *Bacillus subtilis* arginase. A) Asymmetric unit of the arginase crystal structure with six individual chains present (each colored separately here). B) A single monomer of arginase (chain A) is displayed here highlighting active site residues that are further labeled in C) where the active site residues Asp124, Glu268 and His 137 are shown as sticks while the metal coordinating sites are displayed as lines. D) The metal coordination around the two Mn ions is further highlighted showing the distances between the two metal ions and all coordinating atoms in the arginase structure. Figures created using Chimera (248) and PyMol.

Comparison of B. subtilis arginase to previously known structures

```

hArgI 6 RTIGIIGAPFSKGPGRGGVEEGPTVLRKAGLLEKLKEQECVDKDYGDLPF--ADIPNDSP 63
+TI +IG P GQ R GV+ GP+ +R A L+E+L + V+D GD+P I ND
bsArg 3 KTISVIGMPMDLGQARRGVDMGSPAIRYAHLIERLSDMGYTVEDLGDIPINREKIKNDEE 62

hArgI 64 FQIVKNPRSVGKASEQLAGKVAEVKKNGRISLVLGGDHSIAIGSISGHARVHPDLGVIW 123
KN SV +E+LA KV +V + + LVLGGDHS+AIG+++G A+ + +LGVIW
bsArg 63 L---KNLNSVLAGNEKLAQKVNKVIEEKKFPLVLGGDHSIAIGTLAGTAKHYDNLGVIW 119

hArgI 124 DAHDIINTPLTTTSGNLEGPVVSFLLKELKGIKIPDVPGFSSWVTPCISAKDIVYIGLRDVD 183
DAH D+NT T+ SGN+HG P++ L + ++ G++ P I +++V IG R +D
bsArg 120 DAHGD LNTLETSPSGNLEGMPLAVSLGIGHESLVNLEGYA---PKIKPENVVIIIGARSLD 176

hArgI 184 PGEHYILKTLGIKYFSMTEVDRLGIGKVMETLSYLLGRKKRPIHLSFMDGLDPSFPA 243
GE +K G+K ++M E+DRLG+ KV+EETL YL +HLS D+DGLDP+ P
bsArg 177 EGERKYIKESGMKVYTMHEIDRLGMTKVIEETLDYL--SACDGVHLSIDLDGLDPNDAPG 234

hArgI 244 TGTPVVGGLTYREGLYITEEIIYKTGLLSGLDINEVNP SLGKTPEEVTRTVNTAVAITLAC 303
GTPVVG++YRE E +Y G+++ + +EVNP L + +T TAV + +
bsArg 235 VGTTPVVGGISYRESHLAMEMLYDAGIITSAEFVENVNPI L----DHKNKTGKTAVELVESL 290

hArgI 304 FG 305
G
bsArg 291 LG 292

```

Figure 13.4. Primary sequence alignment of human arginase I and *B. subtilis* arginase. A sequence alignment of the human arginase I (in magenta) and the newly solved *B. subtilis* arginase (green). Active site residues as highlighted in Figure 13.1 are boxed with a solid line while additional residues that are involved in coordination of the dual Mg ions are boxed in dashed lines.

The sequence of *B. subtilis* arginase is 42% identical to that of human arginase I with all metal coordinating residues and active site residue conserved (Figure 13.4). This is unsurprising as to date there has been no known arginase structure that has any change in these residues. Expanding this comparison to include other known arginases, the overall fold and indeed active site residues show an incredible degree of homology (Figure 13.5). Again, such conservation would be expected for such a widely conserved enzyme through all kingdoms of life. Such high conservation in the active site may prove *B. subtilis* arginase a similar drug target as bNOS in that designing a molecule that would bind only the bacterial form of arginase while not affecting the host enzyme may take considerable effort.

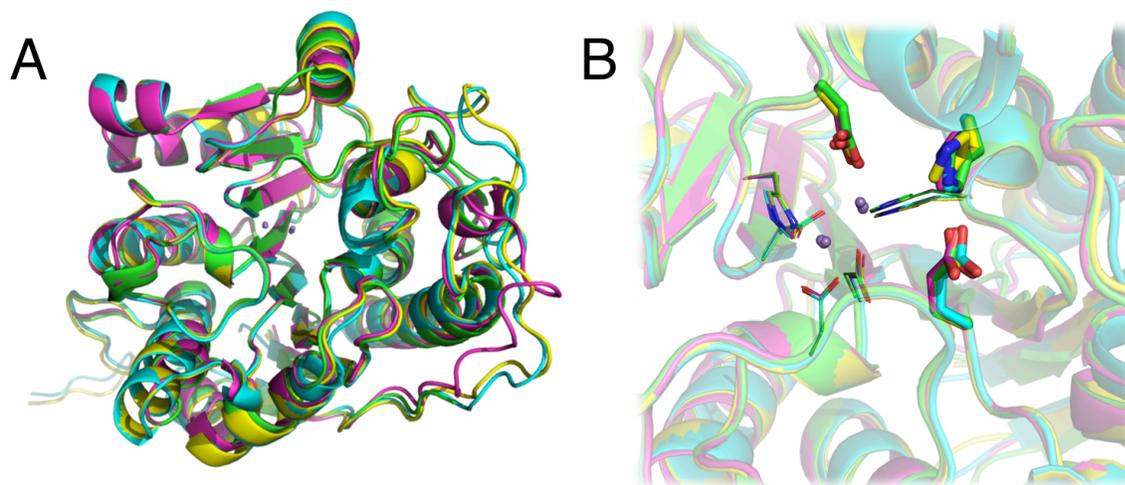


Figure 13.5. Structural alignment of known arginase enzymes reveals strong conservation. The known structures of human arginase I (cyan, PDBID: 2AEB), rat liver arginase (yellow, 1RLA), *Bacillus caldovelox* arginase (magenta, 2CEV) and *Bacillus subtilis* arginase (present study) are aligned to display the strong conservation of both fold (A) and active site coordination (B).

Conclusions and Future Directions

This initial study resulted in the solving of the previously unknown structure of *B. subtilis* arginase. The structure is very similar to previously known arginase structures and has revealed no major changes. The active site of *B. subtilis* arginase is significantly different from that of the known active site of bNOS. The arginase active site lies close to the surface of the enzyme that results in a small targetable pocket. This shallow pocket is significantly dissimilar from the large, deep cavity in bNOS that leads to the heme-centric active site. In addition, the bNOS active site includes a secondary cofactor in BH_4 which has served as an essential drugable pocket in bNOS inhibitor design whereas no such secondary pocket exists in Arginase. Given these differences, it is highly unlikely that these bNOS designed inhibitors, which take advantage of designing molecules to bind to multiple pockets at once, would be true off-target effects and affect arginase

activity. However, to truly study this possible off-target result, further biochemical characterization of *B. subtilis* arginase would be required and is currently underway.

Chapter 14

Summary and Conclusions

The thread that ties the work in chapters one through five is the delicate dance that redox partner proteins must navigate on the edge of a fine knife between specificity, turnover and function. As discussed in chapter one, protein redox pairs must balance high specificity to ensure that electron transfer between the two proteins can readily occur while ensuring that the resulting complexes are weak enough that the complex is not long lived to ensure a high rate of turnover. To probe these effects, three different heme protein redox partner systems were investigated to learn how such systems may be able to navigate this fragile balance in Nature.

In chapters two and three, the dynamic cytochrome P450cam and relatively rigid putidaredoxin (Pdx) redox pair were studied in an attempt to make further sense of recent surprisingly results and to probe what role conformational changes play in redox partner association. Despite the long-held belief that Pdx would bind to a substrate bound-P450cam in a closed conformation, the crystal structure of the complex revealed that Pdx bound to P450cam in an open conformation that had previously only been observed in the absence of substrate. Molecular dynamics simulations on the Pdx-bound, open state of the complex were carried out in chapter two which revealed that the complex was indeed stable and resulted was further stabilized by minor conformational changes that were prevented in the crystal structure by packing. In addition, this work lead to a reevaluation of the published complex crystal structure as though the initial structure showed both the P450cam_{Asp251}-P450cam_{Lys178} and P450cam_{Asp251}-P450cam_{Arg186} ion pairs broken, long used as the definition of the open form of P450cam, the MD simulation revealed that the P450cam_{Asp251}-P450cam_{Arg186} ion pair was able to dynamically

shift between broken and formed, which was further confirmed by the identification of electron density for this pair in the original crystal structure that had initially been overlooked which could have significant mechanistic implications given P450cam_{Asp251}'s role in the proton relay network.

To further this analysis, a systematic study of three different P450cam-Pdx complexes that differ in the conformational state of P450cam as well as simulations of each protein individually to probe the conformational and dynamical effects of P450cam conformation and Pdx binding in chapter three. Numerous analyses of the resulting classical and accelerated molecular dynamics simulations including free energy of binding and principle component analysis revealed that the open-state of P450cam bound to Pdx was significantly more stable than either the closed- or oxy-state complex though all three were found to be energetically favorable. Importantly, the more depth results carried out in this chapter agreed with the previous chapter regarding the addition of new interface contacts in the open state complex upon MD equilibration. Given that Pdx must initially bind to the closed state but the P450cam proton relay network has been believed to only be active in the open state, this data can support the two-step Pdx binding hypothesis put forth by Myers and coworkers in 2013 (162) whereupon Pdx first binds to the closed-state of P450cam followed by the electron transfer which is then followed by a longer-timescale event where Pdx shifts P450cam into the open state which frees up the proton relay network to catalyze the product. Our findings here would suggest that upon Pdx binding, the thermodynamically favorable complex might indeed be the open-state complex and would provide further evidence for this hypothesis. However, additional experimental data will be required before this hypothesis gains more wide-spread acceptance.

While the P450cam-Pdx redox pair introduces conformational changes into the already difficult to study process of redox partner complex formation, the *Leishmania major* peroxidase (LmP) and *Leishmania major* cytochrome c (LmCytc) complex was used to study such systems where no conformational changes in either proteins is involved that was presented in chapter four. Analysis of Brownian dynamics simulations of the LmP and LmCytc complex using both the individual and co-crystal structure revealed the importance of the negatively charged helix A of LmP which serves as the initial point of association for LmCytc to LmP. *In silico* mutagenesis to remove the charges of helix A revealed a significant decrease in the predicted kinetics of the complex, which was validated by accompanying experimental mutagenesis that confirmed the importance of this once overlooked helix. The validity of the Brownian method, as well as the *in silico* mutagenesis was carried out by studying the association and structure of the LmP D211N both experimentally and theoretically. The *in silico* mutant predicted for LmP D211N was revealed to closely match the true structure through comparison to the newly solved crystal structure, while the predicted second order rate constant for the reaction from Brownian dynamics displayed good agreement with the experimental work in a blind study.

Finally, the redox pair complex formed in the output state of nitric oxide synthase (NOS) was probed as a case study for the association and complex formation for redox partners that exist as a single polypeptide. Unlike P450cam-Pdx and LmP-LmCytc, the redox partners for NOS exists as a single polypeptide which means that the association event of two proteins encountering each other at random is severely altered as the reactive domains are always in somewhat close proximity. Despite ongoing work that has begun to bring in the structure and architecture of these domains into focus, an atomic level model of the full complex has remained elusive. The construction and validation of such a model, as presented in chapter five, represents

the first time that an atomic model has been validated. Of interest, the binding of calmodulin, which had long been thought to only alter conformational freedom, actually plays an integral role in stabilizing the redox pair complex between the NOS heme domain dimer and the reactive FMN domain. Supporting these conclusions, the resulting model was validated against a wealth of experimental evidence collected over the past two decades, all of which supports the resulting equilibrated output state model. These results would suggest that a third party, in this case calmodulin, has been recruited through evolution to aid NOS in navigating the delicate balance between the formed and broken Heme-FMN complex which to date not been observed for any other system.

Moving onwards, chapters six through nine move the focus from redox partner proteins to the behavior of water in the presence of self-assembled monolayers as a model to study water at the interface of organic surfaces. There has been a great deal of work that has identified the important role that water plays at the surfaces of proteins, which serves as the basis for the work presented therein. The sensitive nature of water behavior at such organic surfaces has been highlighted through these three different studies.

In chapter seven, the ability of water to penetrate into small-scale defects in SAM surfaces is for the first time detailed. Analysis of the numerous simulations carried out of various defective SAMs reveal that water can penetrate into defects consisting of only a small number of missing siloxane chains. This penetration through the hydrophobic organic surface to the hydrophilic SAM substrate can have drastic effects on the measureable IR spectra of the water exposed to SAM surfaces. Interestingly, beyond the novel findings concerning water behavior on a defective organic surface and penetration through the hydrophobic layer, these findings could serve as a novel test of the purity of a laboratory produced SAM surface. Perfect SAMs, or

SAM's that do not contain defects of three sequential missing siloxane chains should not show the distinctive IR peaks corresponding to poorly hydrogen-bonded water that this work has been able to identify arises from the penetration of water through the substrate, while defective SAMs would reveal such a peak. However, further experimental work would need to be carried out to see if further support for such a quality check test of SAM purity is indeed possible as hypothesized by this work.

Next, the behavior of water on rough SAM surfaces are probed revealing that despite earlier work that suggested such surfaces would inhibit droplet formation, previous work had not reached passed the equilibration phase and droplet formation still occurs in chapter eight. Successive recreation of previously predicted water structures on rough surfaces was achieved but all such structures occurred prior to equilibration which corresponded to droplet formation on the hydrophobic surfaces. Interestingly, it was observed that even in idealized rough surfaces, water was able to penetrate down to the lowest level of the surface, beyond the longest chain length to reach the surface generated by the shorter chains in the SAM. However, one aspect that was not fully investigated was whether a larger difference in SAM chain length might prevent droplet formation as the difference used in this study of four carbon atoms may be too small to achieve superhydrophobicity. Further simulations of additional rough SAM mixtures with different chain length differences could provide further insight into these changes.

Finally, many different variables in the structural makeup of a SAM siloxane chain are varied systematically to present what it is believed to be the first such systematic comparison of chain differences on water behavior in chapter nine. The importance of this study lies in the varied nature of SAM surfaces studies by the field to date. Often, different SAM chains are compared to each other while no comparison has been done between such systems has ever been

carried out until now. This study, which documents the differences both big and small in water behavior and water penetration in a wide variety of SAM chains, could serve as a basis for new comparisons between different SAMs by using this systematic study to ensure that the differences observed between different systems is not system dependent but rather general trends that can be applied for further analysis.

While the work in chapters one through five and six through nine represent two major focuses of my work while at UCI in heme redox partner complexes and water behavior of self-assembled monolayers respectively, the work in chapters eleven and twelve round out this document as additional studies that are unrelated to those two major points of focus. Chapter eleven, which probes the conformational dependence of a central protein residue on the conformation of its neighbor, is the culmination of a long standing collaboration with Dr. Andy Karpus at Oregon State University with whom I worked as an undergraduate studying protein conformation. Building upon our earlier studies, this work continues to evolve our understanding of protein conformational freedom by highlighting the slight dependences that we can now observe due to the explosion of high-resolution crystal structures. Next, in chapter twelve I studied the binding of the PH domain to a target lipid in a lipid membrane building off recent experimental work to improve our understanding of how this protein binds to the membrane. This work serves not only to deepen our understanding of how such proteins bind to targets in a bilayer but also serves as the basis for upcoming work to study lipid protein diffusion on the microsecond timescale that is currently underway.

Finally, in chapter thirteen the unknown structure of *B. subtilis* arginase was solved for the first time in what is the first step in much larger study of the interactions between bNOS and arginase for both activity and inhibition. While work is currently ongoing probing the activity of

the *B. subtilis* arginase enzyme and any possible off-target effects of the previously designed bNOS inhibitors, this particular chapter deserves special note. One of the major goals I had before beginning my graduate career was, on the advice of my undergraduate advisor Dr. P. Andrew Karplus following my extensive work on computational protein conformation work, to get my hands dirty and solve a unknown crystal structure. While I began to move further towards computation as my graduate career carried on onwards, I continuously kept an eye out for a possible lead to scratch that itch. Though solved at the tail end of my time as a graduate student, it should be noted that solving this novel structure is amongst the results that I am most proud of and in doing so was able to achieve what I set out to do as I left Corvallis back in 2010.

In conclusion, the work presented here illustrates the incredible potential that can be harnessed when experimental and computational work intertwines. Indeed, not only do the highlights of this thesis exist solely because of the willingness of both an experimentalist in Dr. Thomas Poulos and a theoretician in Dr. Douglas Tobias to take me in and allow me to experience both worlds, but even some of the purely computational work presented here exists because of the ability to view several of these problems from the point of view of both an experimentalist and theorist. Too often, one viewpoint often disregards or underestimates the insights from the other in such a melding of the minds. Experimentalists can often be guilty of disregarding theoretical work in favor of the bench while ignoring the insights that leads that only computational work can provide. Meanwhile, new theoreticians can get sucked into a narrow path that begins to overlook the valuable scientific point of view of an experimentalist. But when these worlds can come together, they can accomplish much more than either can hope to achieve alone. Despite the rocky road that one may face when presented with the chance to not only live at least partly in both worlds; whether it is the sheer multitude of possible avenues to

take with regard to research direction, or the many long nights lost to learning the intricate details of both a new experimental method or the mathematics behind a new simulation, or the unpredictable swerves when graduate school lovingly decides to try and break a student's spirit, such an opportunity has proven without a doubt both fruitful and fulfilling to this author. And with numerous stories left untold, the work presented herein should hopefully represent the basis for numerous future studies to continue to provide new insights into these fields using this dual view approach.

Bibliography

1. Poulos TL (2014) Heme enzyme structure and function. *Chem. Rev.* 114(7):3919-3962.
2. Bach A & Chodat R (1903) Testings on the importance of peroxides in the chemistry of the alive cell IV. On peroxidase. *Ber Dtsch Chem Ges* 36:600-605.
3. Migliore A, Polizzi NF, Therien MJ, & Beratan DN (2014) Biochemistry and theory of proton-coupled electron transfer. *Chem Rev* 114(7):3381-3465.
4. Winkler JR & Gray HB (2014) Electron flow through metalloproteins. *Chem Rev* 114(7):3369-3380.
5. Spaar A & Helms V (2005) Free energy landscape of protein-protein encounter resulting from Brownian dynamics simulations of Barnase : Barstar. *J Chem Theory Comput* 1(4):723-736.
6. Kim YC, Tang C, Clore GM, & Hummer G (2008) Replica exchange simulations of transient encounter complexes in protein-protein association. *P Natl Acad Sci USA* 105(35):12855-12860.
7. Camacho CJ, Weng ZP, Vajda S, & DeLisi C (1999) Free energy landscapes of encounter complexes in protein-protein association. *Biophysical Journal* 76(3):1166-1178.
8. Gabdoulline RR & Wade RC (1998) Brownian dynamics simulation of protein-protein diffusional encounter. *Methods-a Companion to Methods in Enzymology* 14(3):329-341.
9. Tang C, Iwahara J, & Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. *Nature* 444(7117):383-386.
10. Gabdoulline RR & Wade RC (1999) On the protein-protein diffusional encounter complex. *J Mol Recognit* 12(4):226-234.
11. Northrup SH, *et al.* (1993) Effects of Charged Amino-Acid Mutations on the Bimolecular Kinetics of Reduction of Yeast Iso-1-Ferricytochrome-C by Bovine Ferrocyclochrome-B(5). *Biochemistry* 32(26):6613-6623.
12. Northrup SH, Boles JO, & Reynolds JCL (1988) Brownian Dynamics of Cytochrome-C and Cytochrome-C Peroxidase Association. *Science* 241(4861):67-70.
13. Madrona Y, Tripathi S, Li HY, & Poulos TL (2012) Crystal Structures of Substrate-Free and Nitrosyl Cytochrome P450cin: Implications for O-2 Activation. *Biochemistry* 51(33):6623-6631.
14. Brodie BB, *et al.* (1955) Detoxication of Drugs and Other Foreign Compounds by Liver Microsomes. *Science* 121(3147):603-604.

15. Garfinkel D (1958) Studies on Pig Liver Microsomes .1. Enzymic and Pigment Composition of Different Microsomal Fractions. *Archives of Biochemistry and Biophysics* 77(2):493-509.
16. Klingenberg M (1958) Pigments of Rat Liver Microsomes. *Archives of Biochemistry and Biophysics* 75(2):376-386.
17. Guengerich FP & Munro AW (2013) Unusual Cytochrome P450 Enzymes and Reactions. *Journal of Biological Chemistry* 288(24):17065-17073.
18. Sevrioukova IF & Poulos TL (2015) Current Approaches for Investigating and Predicting Cytochrome P450 3A4-Ligand Interactions. *Advances in experimental medicine and biology* 851:83-105.
19. Denisov IG, Grinkova YV, Baylton JL, Tajkhorshid E, & Sligar SG (2015) Mechanism of drug-drug interactions mediated by human cytochrome P450 CYP3A4 monomer. *Biochemistry* 54(13):2227-2239.
20. Kumar S (2010) Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert opinion on drug metabolism & toxicology* 6(2):115-131.
21. Guengerich FP (2002) Cytochrome P450 enzymes in the generation of commercial products. *Nat Rev Drug Discov* 1(5):359-366.
22. Hawkes DB, Adams GW, Burlingame AL, de Montellano PRO, & De Voss JJ (2002) Cytochrome P450(cin) (CYP176A), isolation, expression, and characterization. *Journal of Biological Chemistry* 277(31):27725-27732.
23. Sevrioukova IF & Poulos TL (2013) Understanding the mechanism of cytochrome P450 3A4: recent advances and remaining problems. *Dalton transactions* 42(9):3116-3126.
24. Sevrioukova IF & Poulos TL (2013) Dissecting cytochrome P450 3A4-ligand interactions using ritonavir analogues. *Biochemistry* 52(26):4474-4481.
25. Green MT, Dawson JH, & Gray HB (2004) Oxoiron(IV) in chloroperoxidase compound II is basic: implications for P450 chemistry. *Science* 304(5677):1653-1656.
26. Davydov R, *et al.* (2001) Hydroxylation of camphor by reduced oxy-cytochrome P450cam: mechanistic implications of EPR and ENDOR studies of catalytic intermediates in native and mutant enzymes. *J Am Chem Soc* 123(7):1403-1415.
27. Rittle J & Green MT (2010) Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics. *Science* 330(6006):933-937.
28. Behan RK & Green MT (2006) On the status of ferryl protonation. *Journal of inorganic biochemistry* 100(4):448-459.

29. Ogliaro F, de Visser SP, & Shaik S (2002) The 'push' effect of the thiolate ligand in cytochrome P450: a theoretical gauging. *Journal of inorganic biochemistry* 91(4):554-567.
30. Groves JT (2014) Enzymatic C-H Bond Activation Using Push to Get Pull. *Nat Chem* 6(2):89-91.
31. Poulos TL (1996) The role of the proximal ligand in heme enzymes. *Journal of Biological Inorganic Chemistry* 1(4):356-359.
32. Bertini I, Gray HB, Stiefel EI, & Valentine JS (2007) *Biological Inorganic Chemistry: Structure and Reactivity* (University Science Books, Sausalito, CA).
33. Madrona Y, Hollingsworth SA, Khan B, & Poulos TL (2013) P450cin active site water: implications for substrate binding and solvent accessibility. *Biochemistry* 52(30):5039-5050.
34. Madrona Y, *et al.* (2014) Crystal structure of cindoxin, the p450cin redox partner. *Biochemistry* 53(9):1435-1446.
35. Poulos TL (2007) Structural biology of p450-oxy complexes. *Drug metabolism reviews* 39(2-3):557-566.
36. Cooley RB, Arp DJ, & Karplus PA (2010) Evolutionary origin of a secondary structure: pi-helices as cryptic but widespread insertional variations of alpha-helices that enhance protein functionality. *J Mol Biol* 404(2):232-246.
37. Ascitutto EK, Young MJ, Madura J, Pochapsky SS, & Pochapsky TC (2012) Solution structural ensembles of substrate-free cytochrome P450(cam). *Biochemistry* 51(16):3383-3393.
38. Ueyama N, Yamada Y, Okamura Ta T, Kimura S, & Nakamura A (1996) Structure and Properties of [Fe(4)S(4){2,6-bis(acylamino)benzenethiolato-S}(4)](2)(-) and [Fe(2)S(2){2,6-bis(acylamino)benzenethiolato-S}(4)](2)(-): Protection of the Fe-S Bond by Double NH.S Hydrogen Bonds. *Inorg Chem* 35(22):6473-6484.
39. Langen R, Jensen GM, Jacob U, Stephens PJ, & Warshel A (1992) Protein control of iron-sulfur cluster redox potentials. *J Biol Chem* 267(36):25625-25627.
40. Adman E, Watenpaugh KD, & Jensen LH (1975) NH---S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin, and Chromatium high potential iron protein. *Proc Natl Acad Sci U S A* 72(12):4854-4858.
41. Hays AM, *et al.* (2004) Conformational states of cytochrome P450cam revealed by trapping of synthetic molecular wires. *J Mol Biol* 344(2):455-469.
42. Lee YT, Wilson RF, Rupniewski I, & Goodin DB (2010) P450cam visits an open conformation in the absence of substrate. *Biochemistry* 49(16):3412-3419.

43. Poulos TL, Finzel BC, & Howard AJ (1987) High-resolution crystal structure of cytochrome P450cam. *J Mol Biol* 195(3):687-700.
44. Schlichting I, *et al.* (2000) The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* 287(5458):1615-1622.
45. Sevrioukova IF, Li H, Zhang H, Peterson JA, & Poulos TL (1999) Structure of a cytochrome P450-redox partner electron-transfer complex. *Proc Natl Acad Sci U S A* 96(5):1863-1868.
46. Strushkevich N, *et al.* (2011) Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc Natl Acad Sci U S A* 108(25):10139-10143.
47. Hiruma Y, *et al.* (2013) The structure of the cytochrome p450cam-putidaredoxin complex determined by paramagnetic NMR spectroscopy and crystallography. *J Mol Biol* 425(22):4353-4365.
48. Tripathi S, Li H, & Poulos TL (2013) Structural basis for effector control and redox partner recognition in cytochrome P450. *Science* 340(6137):1227-1230.
49. Fisher MT & Sligar SG (1985) Control of Heme Protein Redox Potential and Reduction Rate - Linear Free-Energy Relation between Potential and Ferric Spin State Equilibrium. *J Am Chem Soc* 107(17):5018-5019.
50. Pochapsky SS, Pochapsky TC, & Wei JW (2003) A model for effector activity in a highly specific biological electron transfer complex: the cytochrome P450(cam)-putidaredoxin couple. *Biochemistry* 42(19):5649-5656.
51. Unno M, *et al.* (1997) Resonance Raman investigations of cytochrome P450(cam) complexed with putidaredoxin. *J Am Chem Soc* 119(28):6614-6620.
52. Nagano S, *et al.* (2003) Infrared spectroscopic and mutational studies on putidaredoxin-induced conformational changes in ferrous CO-P450cam. *Biochemistry* 42(49):14507-14514.
53. Lipscomb JD, Sligar SG, Namtvedt MJ, & Gunsalus IC (1976) Autooxidation and hydroxylation reactions of oxygenated cytochrome P-450cam. *J Biol Chem* 251(4):1116-1124.
54. Kuznetsov VY, Poulos TL, & Sevrioukova IF (2006) Putidaredoxin-to-cytochrome P450cam electron transfer: differences between the two reductive steps required for catalysis. *Biochemistry* 45(39):11934-11944.
55. Chance B (1949) The Properties of the Enzyme-Substrate Compounds of Horse-Radish and Lacto-Peroxidase. *Science* 109(2826):204-208.
56. Chance B (1949) The primary and secondary compounds of catalase and methyl or ethyl hydrogen peroxide; kinetics and activity. *J Biol Chem* 179(3):1341-1369.

57. Sivaraja M, Goodin DB, Smith M, & Hoffman BM (1989) Identification by ENDOR of Trp191 as the free-radical site in cytochrome c peroxidase compound ES. *Science* 245(4919):738-740.
58. Coulson AF & Yonetani T (1972) Oxidation of cytochrome c peroxidase with hydrogen peroxide: identification of the "endogenous donor". *Biochemical and biophysical research communications* 49(2):391-398.
59. Dolphin D, Forman A, Borg DC, Fajer J, & Felton RH (1971) Compounds I of catalase and horse radish peroxidase: pi-cation radicals. *Proc Natl Acad Sci U S A* 68(3):614-618.
60. Meharena YT, Doukov T, Li HY, Soltis SM, & Poulos TL (2010) Crystallographic and Single-Crystal Spectral Analysis of the Peroxidase Ferryl Intermediate. *Biochemistry* 49(14):2984-2986.
61. Berglund GI, *et al.* (2002) The catalytic pathway of horseradish peroxidase at high resolution. *Nature* 417(6887):463-468.
62. Welinder KG (1992) Superfamily of plant, fungal and bacterial peroxidases. *Curr. Opinion Struct. Biol.* 2:388-393.
63. Altschul AM, Abrams R, & Hogness TR (1939) Soluble cytochrome C oxidase. *Journal of Biological Chemistry* 130(1):427-428.
64. Yonetani T (1965) Studies on Cytochrome C Peroxidase .2. Stoichiometry between Enzyme H₂O₂ and Ferrocycytochrome C and Enzymic Determination of Extinction Coefficients of Cytochrome C. *Journal of Biological Chemistry* 240(11):4509-&.
65. Nicholls P (1964) Observations on Oxidation of Cytochrome C. *Archives of Biochemistry and Biophysics* 106(1-3):25-&.
66. Beetlestone J (1960) The Oxidation of Cytochrome C by Cytochrome C Peroxidase. *Archives of Biochemistry and Biophysics* 89(1):35-40.
67. Lang G, Spartalian K, & Yonetani T (1976) Mossbauer Spectroscopic Study of Compound Es of Cytochrome-C Peroxidase. *Biochimica et biophysica acta* 451(1):250-258.
68. Iizuka T, Kotani M, & Yonetani T (1971) Magnetic Susceptibility Measurements of Cytochrome-C Peroxidase and Its Complexes. *Journal of Biological Chemistry* 246(15):4731-&.
69. Schonbau.Gr & Lo S (1972) Interaction of Peroxidases with Aromatic Peracids and Alkyl Peroxides - Product Analysis. *Journal of Biological Chemistry* 247(10):3353-&.
70. Poulos TL & Kraut J (1980) The Stereochemistry of Peroxidase Catalysis. *Journal of Biological Chemistry* 255(17):8199-8205.

71. Miller MA, *et al.* (1988) Site-Directed Mutagenesis of Yeast Cytochrome-C Peroxidase Shows Histidine-181 Is Not Required for Oxidation of Ferrocycytochrome-C. *Biochemistry* 27(26):9081-9088.
72. Yonetani T & Schleyer H (1966) Studies on Cytochrome C Peroxidase .7. Electron Paramagnetic Resonance Absorptions of Enzyme and Complex Es in Dissolved and Crystalline Forms. *Journal of Biological Chemistry* 241(13):3240-&.
73. Kang CH, Brautigam DL, Osheroff N, & Margoliash E (1978) Definition of Cytochrome-C Binding Domains by Chemical Modification - Reaction of Carboxydinitrophenyl-Cytochromes-C and Trinitrophenyl-Cytochromes-C with Bakers-Yeast Cytochrome-C Peroxidase. *Journal of Biological Chemistry* 253(18):6502-6510.
74. Poulos TL, Sheriff S, & Howard AJ (1987) Cocystals of Yeast Cytochrome-C Peroxidase and Horse Heart Cytochrome-C. *Journal of Biological Chemistry* 262(29):13881-13884.
75. Pelletier H & Kraut J (1992) Crystal-Structure of a Complex between Electron-Transfer Partners, Cytochrome-C Peroxidase and Cytochrome-C. *Science* 258(5089):1748-1755.
76. Guo ML, Bhaskar B, Li HY, Barrows TP, & Poulos TL (2004) Crystal structure and characterization of a cytochrome c peroxidase-cytochrome c site-specific cross-link. *P Natl Acad Sci USA* 101(16):5940-5945.
77. Van de Water K, Sterckx YGJ, & Volkov AN (2015) The low-affinity complex of cytochrome c and its peroxidase. *Nat Commun* 6.
78. Castro G, Boswell CA, & Northrup SH (1998) Dynamics of protein-protein docking: Cytochrome c and cytochrome c peroxidase revisited. *J Biomol Struct Dyn* 16(2):413-424.
79. Stemp EDA & Hoffman BM (1993) Cytochrome-C Peroxidase Binds 2 Molecules of Cytochrome-C - Evidence for a Low-Affinity, Electron-Transfer-Active Site on Cytochrome-C Peroxidase. *Biochemistry* 32(40):10848-10865.
80. Mauk MR, Ferrer JC, & Mauk AG (1994) Proton Linkage in Formation of the Cytochrome-C Cytochrome-C-Peroxidase Complex - Electrostatic Properties of the High-Affinity and Low-Affinity Cytochrome Binding-Sites on the Peroxidase. *Biochemistry* 33(42):12609-12614.
81. Zhou JS & Hoffman BM (1994) Stern-Volmer in Reverse - 2/1 Stoichiometry of the Cytochrome-C Cytochrome-C Peroxidase Electron-Transfer Complex. *Science* 265(5179):1693-1696.
82. Poulos TL (2010) Thirty years of heme peroxidase structural biology. *Archives of Biochemistry and Biophysics* 500(1):3-12.

83. Dolai S, Yadav RK, Pal S, & Adak S (2008) Leishmania major ascorbate peroxidase overexpression protects cells against reactive oxygen species-mediated cardiolipin oxidation. *Free Radical Bio Med* 45(11):1520-1529.
84. Murray HW & Nathan CF (1999) Macrophage microbicidal mechanisms in vivo: Reactive nitrogen versus oxygen intermediates in the killing of intracellular visceral Leishmania donovani. *J Exp Med* 189(4):741-746.
85. Barr SD & Gedamu L (2003) Role of peroxidoxins in Leishmania chagasi survival - Evidence of an enzymatic defense against nitrosative stress. *Journal of Biological Chemistry* 278(12):10816-10823.
86. Barr SD & Gedamu L (2001) Cloning and characterization of three differentially expressed peroxidoxin genes from Leishmania chagasi - Evidence for an enzymatic detoxification of hydroxyl radicals. *Journal of Biological Chemistry* 276(36):34279-34287.
87. Miller MA, *et al.* (2000) Inducible resistance to oxidant stress in the protozoan Leishmania chagasi. *Journal of Biological Chemistry* 275(43):33883-33889.
88. Ghosh S, Goswami S, & Adhya S (2003) Role of superoxide dismutase in survival of Leishmania within the macrophage. *Biochemical Journal* 369:447-452.
89. Chan J, *et al.* (1989) Microbial Glycolipids - Possible Virulence Factors That Scavenge Oxygen Radicals. *P Natl Acad Sci USA* 86(7):2453-2457.
90. Channon JY & Blackwell JM (1985) A Study of the Sensitivity of Leishmania-Donovani Promastigotes and Amastigotes to Hydrogen-Peroxide .2. Possible Mechanisms Involved in Protective H₂O₂ Scavenging. *Parasitology* 91(Oct):207-217.
91. Jasion VS & Poulos TL (2012) Leishmania major peroxidase is a cytochrome c peroxidase. *Biochemistry* 51(12):2453-2460.
92. Jasion VS, Polanco JA, Meharena YT, Li H, & Poulos TL (2011) Crystal structure of Leishmania major peroxidase and characterization of the compound i tryptophan radical. *J Biol Chem* 286(28):24608-24615.
93. Jasion VS, Doukov T, Pineda SH, Li H, & Poulos TL (2012) Crystal structure of the Leishmania major peroxidase-cytochrome c complex. *Proc Natl Acad Sci U S A* 109(45):18390-18394.
94. Fields JB, *et al.* (2015) "Bind and Crawl" Association Mechanism of Leishmania major Peroxidase and Cytochrome c Revealed by Brownian and Molecular Dynamics Simulations. *Biochemistry* 54(49):7272-7282.
95. Li H & Poulos TL (2005) Structure-function studies on nitric oxide synthases. *Journal of inorganic biochemistry* 99(1):293-305.

96. Griffith OW & Stuehr DJ (1995) Nitric Oxides Synthases - Properties and Catalytic Mechanism. *Annu Rev Physiol* 57:707-736.
97. Breddt DS & Snyder SH (1994) Nitric-Oxide - a Physiological Messenger Molecule. *Annual review of biochemistry* 63:175-195.
98. Kerwin JF, Lancaster JR, & Feldman PL (1995) Nitric-Oxide - a New Paradigm for 2nd-Messengers. *Journal of Medicinal Chemistry* 38(22):4343-4362.
99. Dinerman JL, Lowenstein CJ, & Snyder SH (1993) Molecular Mechanisms of Nitric-Oxide Regulation - Potential Relevance to Cardiovascular-Disease. *Circ Res* 73(2):217-222.
100. Moncada S, Palmer RMJ, & Higgs EA (1991) Nitric-Oxide - Physiology, Pathophysiology, and Pharmacology. *Pharmacol Rev* 43(2):109-142.
101. Knowles RG, Palacios M, Palmer RM, & Moncada S (1989) Formation of nitric oxide from L-arginine in the central nervous system: a transduction mechanism for stimulation of the soluble guanylate cyclase. *Proc Natl Acad Sci U S A* 86(13):5159-5162.
102. Endres M, Laufs U, Liao JK, & Moskowitz MA (2004) Targeting eNOS for stroke protection. *Trends Neurosci* 27(5):283-289.
103. Huang PL, *et al.* (1995) Hypertension in mice lacking the gene for endothelial nitric oxide synthase. *Nature* 377(6546):239-242.
104. Poulos TL & Li H (2013) Structural basis for isoform-selective inhibition in nitric oxide synthase. *Accounts of chemical research* 46(2):390-398.
105. Hantraye P, *et al.* (1996) Inhibition of neuronal nitric oxide synthase prevents MPTP-induced parkinsonism in baboons. *Nature medicine* 2(9):1017-1021.
106. Cho DH, *et al.* (2009) S-nitrosylation of Drp1 mediates beta-amyloid-related mitochondrial fission and neuronal injury. *Science* 324(5923):102-105.
107. Calabrese V, *et al.* (2007) Nitric oxide in the central nervous system: neuroprotection versus neurotoxicity. *Nature reviews. Neuroscience* 8(10):766-775.
108. Przedborski S, *et al.* (1996) Role of neuronal nitric oxide in 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced dopaminergic neurotoxicity. *Proc Natl Acad Sci U S A* 93(10):4565-4571.
109. Yang Z, *et al.* (2013) Targeting nitric oxide signaling with nNOS inhibitors as a novel strategy for the therapy and prevention of human melanoma. *Antioxidants & redox signaling* 19(5):433-447.
110. Nathan C (1997) Inducible nitric oxide synthase: what difference does it make? *The Journal of clinical investigation* 100(10):2417-2423.

111. Feng C (2012) Mechanism of Nitric Oxide Synthase Regulation: Electron Transfer and Interdomain Interactions. *Coordination chemistry reviews* 256(3-4):393-411.
112. Gusarov I, Shatalin K, Starodubtseva M, & Nudler E (2009) Endogenous nitric oxide protects bacteria against a wide spectrum of antibiotics. *Science* 325(5946):1380-1384.
113. van Sorge NM, *et al.* (2013) Methicillin-resistant *Staphylococcus aureus* bacterial nitric-oxide synthase affects antibiotic sensitivity and skin abscess development. *J Biol Chem* 288(9):6417-6426.
114. Shatalin K, *et al.* (2008) *Bacillus anthracis*-derived nitric oxide is essential for pathogen virulence and survival in macrophages. *P Natl Acad Sci USA* 105(3):1009-1013.
115. Gusarov I & Nudler E (2005) NO-mediated cytoprotection: Instant adaptation to oxidative stress in bacteria. *P Natl Acad Sci USA* 102(39):13855-13860.
116. Holden JK, *et al.* (2015) Inhibitor Bound Crystal Structures of Bacterial Nitric Oxide Synthase. *Biochemistry* 54(26):4075-4082.
117. Holden JK, *et al.* (2015) Nitric Oxide Synthase as a Target for Methicillin-Resistant *Staphylococcus aureus*. *Chemistry & biology* 22(6):785-792.
118. Holden JK, *et al.* (2015) Structure-based design of bacterial nitric oxide synthase inhibitors. *J Med Chem* 58(2):994-1004.
119. Holden JK, *et al.* (2013) Structural and biological studies on bacterial nitric oxide synthase inhibitors. *Proc Natl Acad Sci U S A* 110(45):18127-18131.
120. Holden JK, Lim N, & Poulos TL (2014) Identification of Redox Partners and Development of a Novel Chimeric Bacterial Nitric Oxide Synthase for Structure Activity Analyses. *J Biol Chem*.
121. Wei CC, *et al.* (2001) Rapid kinetic studies link tetrahydrobiopterin radical formation to heme-dioxy reduction and arginine hydroxylation in inducible nitric-oxide synthase. *J Biol Chem* 276(1):315-319.
122. Wei CC, Wang ZQ, Hemann C, Hille R, & Stuehr DJ (2003) A tetrahydrobiopterin radical forms and then becomes reduced during Nomega-hydroxyarginine oxidation by nitric-oxide synthase. *J Biol Chem* 278(47):46668-46673.
123. Hurshman AR, Krebs C, Edmondson DE, Huynh BH, & Marletta MA (1999) Formation of a pterin radical in the reaction of the heme domain of inducible nitric oxide synthase with oxygen. *Biochemistry* 38(48):15689-15696.
124. Stoll S, *et al.* (2010) Nitric oxide synthase stabilizes the tetrahydrobiopterin cofactor radical by controlling its protonation state. *J Am Chem Soc* 132(33):11812-11823.

125. Marletta MA (1993) Nitric oxide synthase structure and mechanism. *J Biol Chem* 268(17):12231-12234.
126. Huang H, Hah JM, & Silverman RB (2001) Mechanism of nitric oxide synthase. Evidence that direct hydrogen atom abstraction from the O-H bond of NG-hydroxyarginine is not relevant to the mechanism. *J Am Chem Soc* 123(11):2674-2676.
127. Daff S (2010) NO synthase: Structures and mechanisms. *Nitric Oxide-Biol Ch* 23(1):1-11.
128. Crane BR, *et al.* (2000) Structures of the N(omega)-hydroxy-L-arginine complex of inducible nitric oxide synthase oxygenase dimer with active and inactive pterins. *Biochemistry* 39(16):4608-4621.
129. Campbell MG, Smith BC, Potter CS, Carragher B, & Marletta MA (2014) Molecular architecture of mammalian nitric oxide synthases. *Proc Natl Acad Sci U S A* 111(35):E3614-3623.
130. Li W, Fan W, Elmore BO, & Feng C (2011) Effect of solution viscosity on intraprotein electron transfer between the FMN and heme domains in inducible nitric oxide synthase. *Febs Lett* 585(16):2622-2626.
131. Alderton WK, Cooper CE, & Knowles RG (2001) Nitric oxide synthases: structure, function and inhibition. *Biochem J* 357:593-615.
132. Li WB, Fan WH, Elmore BO, & Feng CJ (2011) Effect of solution viscosity on intraprotein electron transfer between the FMN and heme domains in inducible nitric oxide synthase. *Febs Lett* 585(16):2622-2626.
133. Serafim RAM, Primi MC, Trossini GHG, & Ferreira EI (2012) Nitric Oxide: State of the Art in Drug Design. *Curr Med Chem* 19(3):386-405.
134. Li HY, *et al.* (2008) Exploring the Electron Transfer Properties of Neuronal Nitric-oxide Synthase by Reversal of the FMN Redox Potential. *Journal of Biological Chemistry* 283(50):34762-34772.
135. Narayanasami R, *et al.* (1997) The influence of chaotropic reagents on neuronal nitric oxide synthase and its flavoprotein module. Urea and guanidine hydrochloride stimulate NADPH-cytochrome c reductase activity of both proteins. *Nitric Oxide-Biol Ch* 1(1):39-49.
136. Gachhui R, *et al.* (1996) Characterization of the reductase domain of rat neuronal nitric oxide synthase generated in the methylotrophic yeast *Pichia pastoris* - Calmodulin response is complete within the reductase domain itself. *Journal of Biological Chemistry* 271(34):20594-20602.
137. Alderton WK, Cooper CE, & Knowles RG (2001) Nitric oxide synthases: structure, function and inhibition. *Biochem J* 357(Pt 3):593-615.

138. Piazza M, Guillemette JG, & Dieckmann T (2015) Dynamics of nitric oxide synthase-calmodulin interactions at physiological calcium concentrations. *Biochemistry* 54(11):1989-2000.
139. Balligand JL, *et al.* (1994) Cytokine-inducible nitric oxide synthase (iNOS) expression in cardiac myocytes. Characterization and regulation of iNOS expression and detection of iNOS activity in single cardiac myocytes in vitro. *The Journal of biological chemistry* 269(44):27580-27588.
140. Fischmann TO, *et al.* (1999) Structural characterization of nitric oxide synthase isoforms reveals striking active-site conservation. *Nature structural biology* 6(3):233-242.
141. Li H, *et al.* (1999) Crystal structures of zinc-free and -bound heme domain of human inducible nitric-oxide synthase. Implications for dimer stability and comparison with endothelial nitric-oxide synthase. *The Journal of biological chemistry* 274(30):21276-21284.
142. Li H, *et al.* (2002) The novel binding mode of N-alkyl-N'-hydroxyguanidine to neuronal nitric oxide synthase provides mechanistic insights into NO biosynthesis. *Biochemistry* 41(47):13868-13875.
143. Raman CS, *et al.* (1998) Crystal structure of constitutive endothelial nitric oxide synthase: a paradigm for pterin function involving a novel metal center. *Cell* 95(7):939-950.
144. Crane BR, *et al.* (1998) Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. *Science* 279(5359):2121-2126.
145. Zhang J, *et al.* (2001) Crystal structure of the FAD/NADPH-binding domain of rat neuronal nitric-oxide synthase. Comparisons with NADPH-cytochrome P450 oxidoreductase. *The Journal of biological chemistry* 276(40):37506-37513.
146. Xia C, Misra I, Iyanagi T, & Kim JJ (2009) Regulation of interdomain interactions by calmodulin in inducible nitric-oxide synthase. *The Journal of biological chemistry* 284(44):30708-30717.
147. Smith BC, Underbakke ES, Kulp DW, Schief WR, & Marletta MA (2013) Nitric oxide synthase domain interfaces regulate electron transfer and calmodulin activation. *Proc Natl Acad Sci U S A* 110(38):E3577-3586.
148. Yokom AL, *et al.* (2014) Architecture of the nitric-oxide synthase holoenzyme reveals large conformational changes and a calmodulin-driven release of the FMN domain. *The Journal of biological chemistry* 289(24):16855-16865.
149. Volkmann N, *et al.* (2014) Holoenzyme structures of endothelial nitric oxide synthase - an allosteric role for calmodulin in pivoting the FMN domain for electron transfer. *Journal of structural biology* 188(1):46-54.

150. Poulos TL (2005) Structural biology of heme monooxygenases. *Biochemical and biophysical research communications* 338(1):337-345.
151. Nelson DR (2009) The cytochrome p450 homepage. *Human genomics* 4(1):59-65.
152. Ortiz de Montellano PR (2005) *Cytochrome P450 : structure, mechanism, and biochemistry* (Kluwer Academic/Plenum Publishers, New York) 3rd Ed pp xx, 689 p.
153. Bistolos N, Wollenberger U, Jung C, & Scheller FW (2005) Cytochrome P450 biosensors - a review. *Biosens Bioelectron* 20(12):2408-2423.
154. Guengerich FP (1999) Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annual review of pharmacology and toxicology* 39:1-17.
155. Rendic S & Guengerich FP (2012) Contributions of human enzymes in carcinogen metabolism. *Chemical research in toxicology* 25(7):1316-1383.
156. Sevrioukova IF & Poulos TL (2011) Structural biology of redox partner interactions in P450cam monooxygenase: a fresh look at an old system. *Arch Biochem Biophys* 507(1):66-74.
157. Katagiri M, Ganguli BN, & Gunsalus IC (1968) A soluble cytochrome P-450 functional in methylene hydroxylation. *J Biol Chem* 243(12):3543-3546.
158. Sevrioukova IF, Garcia C, Li H, Bhaskar B, & Poulos TL (2003) Crystal structure of putidaredoxin, the [2Fe-2S] component of the P450cam monooxygenase system from *Pseudomonas putida*. *J Mol Biol* 333(2):377-392.
159. Sevrioukova IF, Li H, & Poulos TL (2004) Crystal structure of putidaredoxin reductase from *Pseudomonas putida*, the final structural component of the cytochrome P450cam monooxygenase. *J Mol Biol* 336(4):889-902.
160. Sevrioukova IF, Poulos TL, & Churbanova IY (2010) Crystal structure of the putidaredoxin reductase x putidaredoxin electron transfer complex. *J Biol Chem* 285(18):13616-13620.
161. Churbanova IY, Poulos TL, & Sevrioukova IF (2010) Production and characterization of a functional putidaredoxin reductase-putidaredoxin covalent complex. *Biochemistry* 49(1):58-67.
162. Myers WK, Lee YT, Britt RD, & Goodin DB (2013) The conformation of P450cam in complex with putidaredoxin is dependent on oxidation state. *J. Am. Chem. Soc.* 135(32):11732-11735.
163. Zhang W, Pochapsky SS, Pochapsky TC, & Jain NU (2008) Solution NMR structure of putidaredoxin-cytochrome P450cam complex via a combined residual dipolar coupling-spin labeling approach suggests a role for Trp106 of putidaredoxin in complex formation. *J Mol Biol* 384(2):349-363.

164. Mehareenna YT & Poulos TL (2010) Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450. *Biochemistry* 49(31):6680-6686.
165. Winn PJ, Ludemann SK, Gauges R, Lounnas V, & Wade RC (2002) Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc Natl Acad Sci U S A* 99(8):5361-5366.
166. Koga H, *et al.* (1993) Essential role of the Arg112 residue of cytochrome P450cam for electron transfer from reduced putidaredoxin. *FEBS letters* 331(1-2):109-113.
167. Nakamura K, *et al.* (1994) Significant contribution of arginine-112 and its positive charge of *Pseudomonas putida* cytochrome P-450cam in the electron transport from putidaredoxin. *Biochimica et biophysica acta* 1207(1):40-48.
168. Sligar SG, Debrunner PG, Lipscomb JD, Namtvedt MJ, & Gunsalus IC (1974) A role of the putidaredoxin COOH-terminus in P-450cam (cytochrome m) hydroxylations. *Proc Natl Acad Sci U S A* 71(10):3906-3910.
169. Gerber NC & Sligar SG (1994) A role for Asp-251 in cytochrome P-450cam oxygen activation. *J Biol Chem* 269(6):4260-4266.
170. Vidakovic M, Sligar SG, Li H, & Poulos TL (1998) Understanding the role of the essential Asp251 in cytochrome p450cam using site-directed mutagenesis, crystallography, and kinetic solvent isotope effect. *Biochemistry* 37(26):9211-9219.
171. Batabyal D & Poulos TL (2013) Crystal structures and functional characterization of wild-type CYP101D1 and its active site mutants. *Biochemistry* 52(49):8898-8906.
172. Case DA, *et al.* (2005) The Amber biomolecular simulation programs. *J. Comp. Chem.* 26(16):1668-1688.
173. Shahrokh K, Orendt A, Yost GS, & Cheatham TE, 3rd (2012) Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J Comput Chem* 33(2):119-133.
174. Wang J, Wolf RM, Caldwell JW, Kollman PA, & Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25(9):1157-1174.
175. Jakalian A, Bush BL, Jack DB, & Bayly CI (2000) Fast, efficient generation of high-quality atom charges. AM1-BCC model: I. Method. *J. Comp. Chem.* 21:132-146.
176. Jakalian A, Jack DB, & Bayly CI (2002) Fast, efficient generation of high-quality atom charges. AM1-BCC model: II. Parameterization and validation. *J. Comp. Chem.* 23:1623-1641.

177. Darden T, Perera L, Li L, & Pedersen L (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* 7(3):R55-60.
178. Miyamoto S & Kollman PA (1992) SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water molecules. *J. Comput. Chem.* 13:952-962.
179. Humphrey W, Dalke A, & Schulten K (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* 14(1):33-38, 27-38.
180. Case DA, *et al.* (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668-1688.
181. Gunsalus IC, Pederson TC, & Sligar SG (1975) Oxygenase-catalyzed biological hydroxylations. *Annual review of biochemistry* 44:377-407.
182. Hiruma Y, *et al.* (2014) Hot-spot residues in the cytochrome P450cam-putidaredoxin binding interface. *Chembiochem : a European journal of chemical biology* 15(1):80-86.
183. Hollingsworth SA & Poulos TL (2015) Molecular dynamics of the P450cam-Pdx complex reveals complex stability and novel interface contacts. *Protein science : a publication of the Protein Society* 24(1):49-57.
184. Brewer CB & Peterson JA (1988) Single turnover kinetics of the reaction between oxycytochrome P-450cam and reduced putidaredoxin. *J Biol Chem* 263(2):791-798.
185. OuYang B, Pochapsky SS, Dang M, & Pochapsky TC (2008) A functional proline switch in cytochrome P450cam. *Structure* 16(6):916-923.
186. Markwick PRL, Pierce LCT, Goodin DB, & McCammon JA (2011) Adaptive Accelerated Molecular Dynamics (Ad-AMD) Revealing the Molecular Plasticity of P450cam. *J Phys Chem Lett* 2(3):158-164.
187. Stoll S, *et al.* (2012) Double electron-electron resonance shows cytochrome P450cam undergoes a conformational change in solution upon binding substrate. *Proc. Natl. Acad. Sci. U S A* 109(32):12888-12893.
188. Skinner SP, *et al.* (2015) Delicate conformational balance of the redox enzyme cytochrome P450cam. *Proc Natl Acad Sci U S A* 112(29):9022-9027.
189. Markwick PR, *et al.* (2009) Toward a unified representation of protein structural dynamics in solution. *J Am Chem Soc* 131(46):16968-16975.
190. Grant BJ, Gorfe AA, & McCammon JA (2009) Ras conformational switching: simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLoS computational biology* 5(3):e1000325.

191. Hamelberg D, Mongan J, & McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* 120(24):11919-11929.
192. Hamelberg D, de Oliveira CA, & McCammon JA (2007) Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. *J Chem Phys* 127(15):155102.
193. Pierce LC, Salomon-Ferrer R, Augusto FdOC, McCammon JA, & Walker RC (2012) Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J Chem Theory Comput* 8(9):2997-3002.
194. Chang CH & Kim K (2009) Density Functional Theory Calculation of Bonding and Charge Parameters for Molecular Dynamics Studies on [FeFe] Hydrogenases. *J Chem Theory Comput* 5(4):1137-1145.
195. Grant BJ, Rodrigues AP, ElSawy KM, McCammon JA, & Caves LS (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 22(21):2695-2696.
196. Skjaerven L, Yao XQ, Scarabelli G, & Grant BJ (2014) Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC bioinformatics* 15:399.
197. Hollingsworth SA, Holden JK, Li H, & Poulos TL (2016) Elucidating nitric oxide synthase domain interactions by molecular dynamics. *Protein science : a publication of the Protein Society* 25(2):374-382.
198. Massova I & Kollman PA (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect Drug Discov* 18:113-135.
199. Gohlke H & Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem* 25(2):238-250.
200. Pal S, Dolai S, Yadav RK, & Adak S (2010) Ascorbate peroxidase from Leishmania major controls the virulence of infective stage of promastigotes by regulating oxidative stress. *PLoS One* 5(6):e11271.
201. Adak S & Datta AK (2005) Leishmania major encodes an unusual peroxidase that is a close homologue of plant ascorbate peroxidase: a novel role of the transmembrane domain. *Biochem J* 390(Pt 2):465-474.
202. Volkov AN, Nicholls P, & Worrall JA (2011) The complex of cytochrome c and cytochrome c peroxidase: the end of the road? *Biochimica et biophysica acta* 1807(11):1482-1503.
203. Chreifi G, *et al.* (2015) Enzymatic Mechanism of Leishmania major Peroxidase and the Critical Role of Specific Ionic Interactions. *Biochemistry* 54(21):3328-3336.

204. Miller MA (1996) A complete mechanism for steady-state oxidation of yeast cytochrome c by yeast cytochrome c peroxidase. *Biochemistry* 35(49):15791-15799.
205. Wang K, *et al.* (1996) Design of a ruthenium-cytochrome c derivative to measure electron transfer to the radical cation and oxyferryl heme in cytochrome c peroxidase. *Biochemistry* 35(47):15107-15119.
206. Kang CH, Ferguson-Miller S, & Margoliash E (1977) Steady state kinetics and binding of eukaryotic cytochromes c with yeast cytochrome c peroxidase. *J Biol Chem* 252(3):919-926.
207. Pielak GJ & Wang X (2001) Interactions between yeast iso-1-cytochrome c and its peroxidase. *Biochemistry* 40(2):422-428.
208. Nocek JM, Sishta BP, Cameron JC, Mauk AG, & Hoffmann BM (1997) Cyclic electron transfer within the [Zn-myoglobin, cytochrome b(5)] complex. *J Am Chem Soc* 119(9):2146-2155.
209. Pearson DC & Gross EL (1998) Brownian dynamics study of the interaction between plastocyanin and cytochrome f. *Biophysical Journal* 75(6):2698-2711.
210. De Rienzo F, Gabdoulline RR, Menziani MC, De Benedetti PG, & Wade RC (2001) Electrostatic analysis and Brownian dynamics simulation of the association of plastocyanin and cytochrome F. *Biophysical Journal* 81(6):3090-3104.
211. Gabdoulline RR & Wade RC (2001) Protein-protein association: Investigation of factors influencing association rates by Brownian dynamics simulations. *Journal of Molecular Biology* 306(5):1139-1155.
212. Brittain T, Kidd RD, & Baker EN (2002) Electron transfer between cytochrome b(5) and some oxidised haemoglobins: the role of ionic strength. *Journal of inorganic biochemistry* 88(3-4):328-334.
213. Furukawa Y, Matsuda F, Ishimori K, & Morishima I (2002) Investigation of the electron-transfer mechanism by cross-linking between Zn-substituted myoglobin and cytochrome b(5). *J Am Chem Soc* 124(15):4008-4019.
214. Gross EL & Pearson DC (2003) Brownian dynamics Simulations of the interaction of Chlamydomonas cytochrome f with plastocyanin and cytochrome c(6). *Biophysical Journal* 85(3):2055-2068.
215. Lin JP & Beratan DN (2005) Simulation of electron transfer between cytochrome c(2) and the bacterial photosynthetic reaction center: Brownian dynamics analysis of the native proteins and double mutants. *J Phys Chem B* 109(15):7529-7534.
216. Gross EL (2007) A Brownian dynamics computational study of the interaction of spinach plastocyanin with turnip cytochrome f: the importance of plastocyanin conformational changes. *Photosynth Res* 94(2-3):411-422.

217. Zollner A, Pasquinelli MA, Bernhardt R, & Beratan DN (2007) Protein phosphorylation and intermolecular electron transfer: A joint experimental and computational study of a hormone biosynthesis pathway. *J Am Chem Soc* 129(14):4206-4216.
218. Long H, Chang CH, King PW, Ghirardi ML, & Kim K (2008) Brownian dynamics and molecular dynamics study of the association between hydrogenase and ferredoxin from *Chlamydomonas reinhardtii*. *Biophysical Journal* 95(8):3753-3766.
219. Gabdouliline RR & Wade RC (2009) On the Contributions of Diffusion and Thermal Activation to Electron Transfer between *Phormidium laminosum* Plastocyanin and Cytochrome f: Brownian Dynamics Simulations with Explicit Modeling of Nonpolar Desolvation Interactions and Electron Transfer Events. *J Am Chem Soc* 131(26):9230-9238.
220. Xiong P, *et al.* (2010) Faster Interprotein Electron Transfer in a [Myoglobin, b(5)] Complex with a Redesigned Interface. *Science* 330(6007):1075-1078.
221. Cruz-Gallardo I, Diaz-Moreno I, Diaz-Quintana A, & De la Rosa MA (2012) The cytochrome f-plastocyanin complex as a model to study transient interactions between redox proteins. *FEBS letters* 586(5):646-652.
222. Grove TZ, Ullmann GM, & Kostic NM (2012) Simultaneous true, gated, and coupled electron-transfer reactions and energetics of protein rearrangement. *Journal of inorganic biochemistry* 106(1):143-150.
223. Keinan S, Nocek JM, Hoffman BM, & Beratan DN (2012) Interfacial hydration, dynamics and electron transfer: multi-scale ET modeling of the transient [myoglobin, cytochrome b(5)] complex. *Physical Chemistry Chemical Physics* 14(40):13881-13889.
224. Moreno-Beltran B, *et al.* (2014) Cytochrome c(1) exhibits two binding sites for cytochrome c in plants. *Bba-Bioenergetics* 1837(10):1717-1729.
225. Ermak DL & McCammon JA (1978) Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352-1360.
226. Gabdouliline RR & Wade RC (1998) Brownian dynamics simulation of protein-protein diffusional encounter. *Methods* 14(3):329-341.
227. Martinez M, *et al.* (2015) SDA 7: A modular and parallel implementation of the simulation of diffusional association software. *Journal of Computational Chemistry* 36(21):1631-1645.
228. Gabdouliline RR & Wade RC (1996) Effective charges for macromolecules in solvent. *J Phys Chem-US* 100(9):3868-3878.
229. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *P Natl Acad Sci USA* 98(18):10037-10041.

230. MacKerell AD, *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586-3616.
231. Bruccoleri RE, Novotny J, & Davis ME (1997) Finite difference Poisson-Boltzmann electrostatic calculations: Increased accuracy achieved by harmonic dielectric smoothing and charge antialiasing. *Journal of Computational Chemistry* 18(2):268-276.
232. Ortega A, Amoros D, & de la Torre JG (2011) Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal* 101(4):892-898.
233. Northrup SH, Allison SA, & Mccammon JA (1984) Brownian Dynamics Simulation of Diffusion-Influenced Bimolecular Reactions. *J Chem Phys* 80(4):1517-1526.
234. Phillips JC, *et al.* (2005) Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26(16):1781-1802.
235. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79(2):926-935.
236. Feller SE, Zhang YH, Pastor RW, & Brooks BR (1995) Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *J Chem Phys* 103(11):4613-4621.
237. Martyna GJ, Tobias DJ, & Klein ML (1994) Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101(5):4177-4189.
238. Grubmuller H, Heller H, Windemuth A, & Schulten K (1991) Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-Range Interactions. *Mol Simulat* 6(1-3):121-142.
239. Essmann U, *et al.* (1995) A Smooth Particle Mesh Ewald Method. *J Chem Phys* 103(19):8577-8593.
240. Ryckaert JP, Ciccotti G, & Berendsen HJC (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* 23(3):327-341.
241. Ivens AC, *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309(5733):436-442.
242. McPhillips TM, *et al.* (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J Synchrotron Radiat* 9:401-406.
243. Kabsch W (2010) Xds. *Acta Crystallogr D* 66:125-132.
244. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.

245. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240-255.
246. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126-2132.
247. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D* 57:122-133.
248. Pettersen EF, *et al.* (2004) UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13):1605-1612.
249. Mei HK, *et al.* (1996) Control of formation and dissociation of the high-affinity complex between cytochrome c and cytochrome c peroxidase by ionic strength and the low-affinity binding site. *Biochemistry* 35(49):15800-15806.
250. Moncada S & Higgs EA (1991) Endogenous Nitric-Oxide - Physiology, Pathology and Clinical Relevance. *Eur J Clin Invest* 21(4):361-374.
251. Forstermann U & Sessa WC (2012) Nitric oxide synthases: regulation and function. *Eur Heart J* 33(7):829-+.
252. Dorheim MA, Tracey WR, Pollock JS, & Grammas P (1994) Nitric-Oxide Synthase Activity Is Elevated in Brain Microvessels in Alzheimers-Disease. *Biochem Bioph Res Co* 205(1):659-665.
253. Norris PJ, Waldvogel HJ, Faull RLM, Love DR, & Emson PC (1996) Decreased neuronal nitric oxide synthase messenger RNA and somatostatin messenger RNA in the striatum of Huntington's disease. *Neuroscience* 72(4):1037-1047.
254. Sims NR & Anderson MF (2002) Mitochondrial contributions to tissue damage in stroke. *Neurochem Int* 40(6):511-526.
255. Zhang L, Dawson VL, & Dawson TM (2006) Role of nitric oxide in Parkinson's disease. *Pharmacol Therapeut* 109(1-2):33-41.
256. Stuehr DJ & Griffith OW (1992) Mammalian Nitric-Oxide Synthases. *Advances in enzymology and related areas of molecular biology* 65:287-346.
257. Ilagan RP, *et al.* (2009) Regulation of FMN subdomain interactions and function in neuronal nitric oxide synthase. *Biochemistry* 48(18):3864-3876.
258. Astashkin AV, Elmore BO, Fan W, Guillemette JG, & Feng C (2010) Pulsed EPR determination of the distance between heme iron and FMN centers in a human inducible nitric oxide synthase. *Journal of the American Chemical Society* 132(34):12059-12067.

259. Astashkin AV, *et al.* (2014) Pulsed electron paramagnetic resonance study of domain docking in neuronal nitric oxide synthase: the calmodulin and output state perspective. *The journal of physical chemistry. A* 118(34):6864-6872.
260. Karplus M & McCammon JA (2002) Molecular dynamics simulations of biomolecules. *Nature structural biology* 9(9):646-652.
261. Adcock SA & McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews* 106(5):1589-1615.
262. Salemme FR (1976) An hypothetical structure for an intermolecular electron transfer complex of cytochromes c and b5. *Journal of molecular biology* 102(3):563-568.
263. Wendoloski JJ, Matthew JB, Weber PC, & Salemme FR (1987) Molecular dynamics of a cytochrome c-cytochrome b5 electron transfer complex. *Science* 238(4828):794-797.
264. Sempombe J, *et al.* (2011) Mutation in the flavin mononucleotide domain modulates magnetic circular dichroism spectra of the iNOS ferric cyano complex in a substrate-specific manner. *Inorganic chemistry* 50(15):6859-6861.
265. Feng C, *et al.* (2014) Dissecting regulation mechanism of the FMN to heme interdomain electron transfer in nitric oxide synthases. *Journal of inorganic biochemistry* 130:130-140.
266. Haque MM, Bayachou M, Fadlalla MA, Durra D, & Stuehr DJ (2013) Charge-pairing interactions control the conformational setpoint and motions of the FMN domain in neuronal nitric oxide synthase. *Biochem J* 450(3):607-617.
267. Li H, *et al.* (2008) Exploring the electron transfer properties of neuronal nitric-oxide synthase by reversal of the FMN redox potential. *The Journal of biological chemistry* 283(50):34762-34772.
268. George SE, Su Z, Fan D, Wang S, & Johnson JD (1996) The fourth EF-hand of calmodulin and its helix-loop-helix components: impact on calcium binding and enzyme activation. *Biochemistry* 35(25):8307-8313.
269. Tejero J, Hannibal L, Mustovich A, & Stuehr DJ (2010) Surface charges and regulation of FMN to heme electron transfer in nitric-oxide synthase. *The Journal of biological chemistry* 285(35):27232-27240.
270. Persechini A, Tran QK, Black DJ, & Gogol EP (2013) Calmodulin-induced structural changes in endothelial nitric oxide synthase. *Febs Lett* 587(3):297-301.
271. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.
272. Humphrey W, Dalke A, & Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph Model* 14(1):33-38.

273. Kale L, *et al.* (1999) NAMD2: Greater scalability for parallel molecular dynamics. *J Comput Phys* 151(1):283-312.
274. Mackerell AD, Feig M, & Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25(11):1400-1415.
275. Darden T, York D, & Pedersen L (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* 98(12):10089-10092.
276. Martyna GJ, Tobias DJ, & Klein ML (1994) Constant-Pressure Molecular-Dynamics Algorithms. *J Chem Phys* 101(5):4177-4189.
277. Ebbinghaus S, *et al.* (2007) An extended dynamical hydration shell around proteins. *P Natl Acad Sci USA* 104(52):20749-20752.
278. Heyden M & Tobias DJ (2013) Spatial Dependence of Protein-Water Collective Hydrogen-Bond Dynamics. *Phys Rev Lett* 111(21).
279. Schiro G, *et al.* (2015) Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nat Commun* 6:6490.
280. Hummer G, Garde S, Garcia AE, Paulaitis ME, & Pratt LR (1998) Hydrophobic effects on a molecular scale. *J Phys Chem B* 102(51):10469-10482.
281. Israelachvili J & Wennerstrom H (1996) Role of hydration and water structure in biological and colloidal interactions. *Nature* 379(6562):219-225.
282. Ball P (2008) Water as an active constituent in cell biology. *Chem Rev* 108(1):74-108.
283. Tang XL, Dong JF, & Li XF (2008) A comparison of spreading behaviors of Silwet L-77 on dry and wet lotus leaves. *J Colloid Interf Sci* 325(1):223-227.
284. Seinfeld JH, *et al.* (2004) ACE-ASIA - Regional climatic and atmospheric chemical effects of Asian dust and pollution. *B Am Meteorol Soc* 85(3):367-+.
285. Gao Y & Anderson JR (2001) Characteristics of Chinese aerosols determined by individual-particle analysis. *J Geophys Res-Atmos* 106(D16):18037-18045.
286. Sumner AL, *et al.* (2004) The nature of water on surfaces of laboratory systems and implications for heterogeneous chemistry in the troposphere. *Physical Chemistry Chemical Physics* 6(3):604-613.
287. Ewing GE (2006) Ambient thin film water on insulator surfaces. *Chem Rev* 106(4):1511-1526.

288. Gingrich SE & Diamond ML (2001) Atmospherically derived organic surface films along an urban-rural gradient. *Environ Sci Technol* 35(20):4031-4037.
289. Lam B, *et al.* (2005) Chemical composition of surface films on glass windows and implications for atmospheric chemistry. *Atmos Environ* 39(35):6578-6586.
290. Law NL & Diamond ML (1998) The role of organic films and the effect on hydrophobic organic compounds in urban areas: An hypothesis. *Chemosphere* 36(12):2607-2620.
291. Netz RR (2004) Water and ions at interfaces. *Curr Opin Colloid In* 9(1-2):192-197.
292. Moussa SG, *et al.* (2013) Experimental and theoretical studies of the interaction of gas phase nitric acid and water with a self-assembled monolayer. *Physical Chemistry Chemical Physics* 15(2):448-458.
293. Nishino N, *et al.* (2014) Interactions of gaseous HNO₃ and water with individual and mixed alkyl self-assembled monolayers at room temperature. *Physical chemistry chemical physics : PCCP* 16(6):2358-2367.
294. Raschke TM (2006) Water structure and interactions with protein surfaces. *Curr Opin Struc Biol* 16(2):152-159.
295. Verdaguer A, Sacha GM, Bluhm H, & Salmeron M (2006) Molecular structure of water at interfaces: Wetting at the nanometer scale. *Chem Rev* 106(4):1478-1510.
296. Thiel PA & Madey TE (1987) The Interaction of Water with Solid-Surfaces - Fundamental-Aspects. *Surf Sci Rep* 7(6-8):211-385.
297. Shen YR & Ostroverkhov V (2006) Sum-frequency vibrational spectroscopy on water interfaces: Polar orientation of water molecules at interfaces. *Chem Rev* 106(4):1140-1154.
298. Henderson MA (2002) The interaction of water with solid surfaces: fundamental aspects revisited. *Surf Sci Rep* 46(1-8):1-308.
299. Gragson DE, McCarty BM, & Richmond GL (1997) Ordering of interfacial water molecules at the charged air/water interface observed by vibrational sum frequency generation. *J Am Chem Soc* 119(26):6144-6152.
300. Bagchi B (2005) Water dynamics in the hydration layer around proteins and micelles. *Chem Rev* 105(9):3197-3219.
301. Tian CS, Ji N, Waychunas GA, & Shen YR (2008) Interfacial structures of acidic and basic aqueous solutions. *J Am Chem Soc* 130(39):13033-13039.
302. Finlayson-Pitts BJ & Pitts JN (2000) *Chemistry of the Upper and Lower Atmosphere-Theory, Experiments and Applications* (Academic Press, San Diego).

303. Finlayson-Pitts BJ, Wingen LM, Sumner AL, Syomin D, & Ramazan KA (2003) The heterogeneous hydrolysis of NO₂ in laboratory systems and in outdoor and indoor atmospheres: An integrated mechanism. *Physical Chemistry Chemical Physics* 5(2):223-242.
304. Vericat C, *et al.* (2014) Self-assembled monolayers of thiolates on metals: a review article on sulfur-metal chemistry and surface structures. *Rsc Adv* 4(53):27730-27754.
305. Stevens MJ & Grest GS (2008) Simulations of water at the interface with hydrophilic self-assembled monolayers. *Biointerphases* 3(3):FC13-22.
306. Ulman A (1996) Formation and Structure of Self-Assembled Monolayers. *Chem Rev* 96(4):1533-1554.
307. Ulman A (1991) *An introduction to ultrathin organic films : from Langmuir-Blodgett to self-assembly* (Academic Press, Boston) pp xxiii, 442 p.
308. Voges AB, *et al.* (2007) Insights into heterogeneous atmospheric oxidation chemistry: Development of a tailor-made synthetic model for studying tropospheric surface chemistry. *J Phys Chem C* 111(4):1567-1578.
309. Rudich Y, *et al.* (2000) Wetting of hydrophobic organic surfaces and its implications to organic aerosols in the atmosphere. *J Phys Chem A* 104(22):5238-5245.
310. McIntire TM, *et al.* (2005) Unusual aggregates from the oxidation of alkene self-assembled monolayers: a previously unrecognized mechanism for SAM ozonolysis? *Physical Chemistry Chemical Physics* 7(20):3605-3609.
311. Rudich Y (2003) Laboratory perspectives on the chemical transformations of organic matter in atmospheric particles. *Chem Rev* 103(12):5097-5124.
312. Love JC, Estroff LA, Kriebel JK, Nuzzo RG, & Whitesides GM (2005) Self-assembled monolayers of thiolates on metals as a form of nanotechnology. *Chem Rev* 105(4):1103-1169.
313. Prime KL & Whitesides GM (1991) Self-Assembled Organic Monolayers - Model Systems for Studying Adsorption of Proteins at Surfaces. *Science* 252(5009):1164-1167.
314. Cheng XY, Lowe SB, Reece PJ, & Gooding JJ (2014) Colloidal silicon quantum dots: from preparation to the modification of self-assembled monolayers (SAMs) for bio-applications. *Chem Soc Rev* 43(8):2680-2700.
315. He Y, *et al.* (2009) Photo and pH Stable, Highly-Luminescent Silicon Nanospheres and Their Bioconjugates for Immunofluorescent Cell Imaging. *J Am Chem Soc* 131(12):4434-4438.

316. Li ZF & Ruckenstein E (2004) Water-soluble poly(acrylic acid) grafted luminescent silicon nanoparticles and their use as fluorescent biological staining labels. *Nano Lett* 4(8):1463-1467.
317. Robel I, Subramanian V, Kuno M, & Kamat PV (2006) Quantum dot solar cells. Harvesting light energy with CdSe nanocrystals molecularly linked to mesoscopic TiO₂ films. *J Am Chem Soc* 128(7):2385-2393.
318. Samanta D & Sarkar A (2011) Immobilization of bio-macromolecules on self-assembled monolayers: methods and sensor applications. *Chem Soc Rev* 40(5):2567-2592.
319. Kumar A, Biebuyck HA, & Whitesides GM (1994) Patterning Self-Assembled Monolayers - Applications in Materials Science. *Langmuir* 10(5):1498-1511.
320. Turchanin A, *et al.* (2011) Conversion of Self-Assembled Monolayers into Nanocrystalline Graphene: Structure and Electric Transport. *Acs Nano* 5(5):3896-3904.
321. Losic D, Shapter JG, & Gooding JJ (2002) Integrating polymers with alkanethiol self-assembled monolayers (SAMs): blocking SAM defects with electrochemical polymerisation of tyramine. *Electrochem Commun* 4(12):953-958.
322. Celestin M, Krishnan S, Bhansali S, Stefanakos E, & Goswami DY (2014) A review of self-assembled monolayers as potential terahertz frequency tunnel diodes. *Nano Res* 7(5):589-625.
323. Tu CQ, Ma XC, Pantazis P, Kauzlarich SM, & Louie AY (2010) Paramagnetic, Silicon Quantum Dots for Magnetic Resonance and Two-Photon Imaging of Macrophages. *J Am Chem Soc* 132(6):2016-2023.
324. Erogbogbo F, *et al.* (2012) Bioconjugation of luminescent silicon quantum dots to gadolinium ions for bioimaging applications. *Nanoscale* 4(17):5483-5489.
325. Yamamoto H, Watanabe T, & Ohdomari I (2008) A molecular simulation study of an organosilane self-assembled monolayer/SiO₂ substrate interface. *J Chem Phys* 128(16).
326. Schreiber F (2000) Structure and growth of self-assembling monolayers. *Prog Surf Sci* 65(5-8):151-256.
327. Guo Z, *et al.* (2008) On the chain length dependence of CH₃ vibrational mode relative intensities in sum frequency generation spectra of self assembled alkanethiols. *Surf Sci* 602(23):3551-3559.
328. Stevens MJ (1999) Thoughts on the Structure of alkylsilane monolayers. *Langmuir* 15(8):2773-2778.
329. Hautman J & Klein ML (1989) Simulation of a Monolayer of Alkyl Thiol Chains. *J Chem Phys* 91(8):4994-5001.

330. Hautman J & Klein ML (1990) Molecular-Dynamics Simulation of the Effects of Temperature on a Dense Monolayer of Long-Chain Molecules. *J Chem Phys* 93(10):7483-7492.
331. Mar W & Klein ML (1994) Molecular-Dynamics Study of the Self-Assembled Monolayer Composed of S(CH₂)₁₄CH₃ Molecules Using an All-Atoms Model. *Langmuir* 10(1):188-196.
332. Srivastava P, Chapman WG, & Laibinis PE (2005) Odd-even variations in the wettability of n-alkanethiolate monolayers on gold by water and hexadecane: A molecular dynamics simulation study. *Langmuir* 21(26):12171-12178.
333. Hautman J, Bareman JP, Mar W, & Klein ML (1991) Molecular-Dynamics Investigations of Self-Assembled Monolayers. *J Chem Soc Faraday T* 87(13):2031-&.
334. Mazyar OA, Jennings GK, & McCabe C (2009) Frictional Dynamics of Alkylsilane Monolayers on SiO₂: Effect of 1-n-Butyl-3-methylimidazolium Nitrate as a Lubricant. *Langmuir* 25(9):5103-5110.
335. Chandross M, Grest GS, & Stevens MJ (2002) Friction between alkylsilane monolayers: Molecular simulation of ordered monolayers. *Langmuir* 18(22):8392-8399.
336. Hautman J & Klein ML (1991) Microscopic Wetting Phenomena. *Phys Rev Lett* 67(13):1763-1766.
337. James M, *et al.* (2011) Nanoscale condensation of water on self-assembled monolayers. *Soft Matter* 7(11):5309-5318.
338. Moussa SG, *et al.* (2009) Experimental and Theoretical Characterization of Adsorbed Water on Self-Assembled Monolayers: Understanding the Interaction of Water with Atmospherically Relevant Surfaces. *J Phys Chem A* 113(10):2060-2069.
339. Szori M, Tobias DJ, & Roeselova M (2009) Microscopic Wetting of Mixed Self-assembled Monolayers: A Molecular Dynamics Study. *J Phys Chem B* 113(13):4161-4169.
340. Setny P & Geller M (2006) Water properties inside nanoscopic hydrophobic pocket studied by computer simulations. *J Chem Phys* 125(14).
341. Shibukawa M, Kondo Y, Ogiyama Y, Osuga K, & Saito S (2011) Interfacial water on hydrophobic surfaces recognized by ions and molecules. *Physical Chemistry Chemical Physics* 13(35):15925-15935.
342. Mezger M, *et al.* (2006) High-resolution in situ x-ray study of the hydrophobic gap at the water-octadecyl-trichlorosilane interface. *P Natl Acad Sci USA* 103(49):18401-18404.
343. McIntire TM, *et al.* (2006) Substrate changes associated with the chemistry of self-assembled monolayers on silicon. *Langmuir* 22(13):5617-5624.

344. Lane JM, Chandross M, Lorenz CD, Stevens MJ, & Grest GS (2008) Water penetration of damaged self-assembled monolayers. *Langmuir* 24(11):5734-5739.
345. Diez-Perez I, *et al.* (2004) Interaction of water with self-assembled monolayers of alkylsilanes on mica. *Langmuir* 20(4):1284-1290.
346. Dai JY, *et al.* (2009) Some thoughts on the existence of ion and water channels in highly dense and well-ordered CH₃-terminated alkanethiol self-assembled monolayers on gold. *Biosens Bioelectron* 24(5):1074-1082.
347. Zhang HP, Romero C, & Baldelli S (2005) Preparation of alkanethiol monolayers on mild steel surfaces studied with sum frequency generation and electrochemistry. *J Phys Chem B* 109(32):15520-15530.
348. Waring C, Bagot PAJ, Raisanen MT, Costen ML, & McKendrick KG (2009) Dynamics of the Reaction of O(P-3) Atoms with Alkylthiol Self-assembled Monolayers. *J Phys Chem A* 113(16):4320-4329.
349. Nosonovsky M & Bhushan B (2009) Superhydrophobic surfaces and emerging applications: Non-adhesion, energy, green engineering. *Curr Opin Colloid In* 14(4):270-280.
350. Daub CD, Wang JH, Kudesia S, Bratko D, & Luzar A (2010) The influence of molecular-scale roughness on the surface spreading of an aqueous nanodrop. *Faraday Discuss* 146:67-77.
351. Szori M, Jedlovsky P, & Roeselova M (2010) Water adsorption on hydrophilic and hydrophobic self-assembled monolayers as proxies for atmospheric surfaces. A grand canonical Monte Carlo simulation study. *Physical Chemistry Chemical Physics* 12(18):4604-4616.
352. Szleifer I (1997) Protein adsorption on surfaces with grafted polymers: A theoretical approach. *Biophysical Journal* 72(2):595-612.
353. Pertsin AJ & Grunze M (2000) Computer simulation of water near the surface of oligo(ethylene glycol)-terminated alkanethiol self-assembled monolayers. *Langmuir* 16(23):8829-8841.
354. Pertsin AJ, Hayashi T, & Grunze M (2002) Grand canonical Monte Carlo Simulations of the hydration interaction between oligo(ethylene glycol)-terminated alkanethiol self-assembled monolayers. *J Phys Chem B* 106(47):12274-12281.
355. Tu A, Kwag HR, Barnette AL, & Kim SH (2012) Water Adsorption Isotherms on CH₃-, OH-, and COOH-Terminated Organic Surfaces at Ambient Conditions Measured with PM-RAIRS. *Langmuir* 28(43):15263-15269.

356. Ismail AE, Grest GS, & Stevens MJ (2007) Structure and dynamics of water near the interface with oligo(ethylene oxide) self-assembled monolayers. *Langmuir* 23(16):8508-8514.
357. Winter N, Vieceli J, & Benjamin I (2008) Hydrogen-bond structure and dynamics at the interface between water and carboxylic acid-functionalized self-assembled monolayers. *J Phys Chem B* 112(2):227-231.
358. Grimm RL, Tobias DJ, & Hemminger JC (2010) D2O Water Interaction with Textured Carboxylic Acid-Terminated Monolayer Surfaces Characterized by Temperature-Programmed Desorption and Molecular Dynamics. *J Phys Chem C* 114(3):1570-1579.
359. Duffy DM & Harding JH (2005) Modeling the properties of self-assembled monolayers terminated by carboxylic acids. *Langmuir* 21(9):3850-3857.
360. Simpson AJ, *et al.* (2006) Assessing the organic composition of urban surface films using nuclear magnetic resonance spectroscopy. *Chemosphere* 63(1):142-152.
361. Verdaguer A, *et al.* (2007) Growth and structure of water on SiO₂ films on Si investigated by Kelvin probe microscopy and in situ x-ray spectroscopies. *Langmuir* 23(19):9699-9703.
362. Michel D, Kazansky VB, & Andreev VM (1978) Study of Interaction between Surface Hydroxyls and Adsorbed Water-Molecules on Porous Glasses by Means of Infrared Spectroscopy. *Surf Sci* 72(2):342-356.
363. Dubowski Y, *et al.* (2004) Interaction of gas-phase ozone at 296 K with unsaturated self-assembled monolayers: A new look at an old system. *J Phys Chem A* 108(47):10473-10485.
364. Li I, Bandara J, & Shultz MJ (2004) Time evolution studies of the H₂O/quartz interface using sum frequency generation, atomic force microscopy, and molecular dynamics. *Langmuir* 20(24):10474-10480.
365. Yamamoto S, Beniya A, Mukai K, Yamashita Y, & Yoshinobu J (2005) Water adsorption on Rh(111) at 20 K: From monomer to bulk amorphous ice. *J Phys Chem B* 109(12):5816-5823.
366. Jo SK, Kiss J, Polanco JA, & White JM (1991) Identification of 2nd Layer Adsorbates - Water and Chloroethane on Pt(111). *Surf Sci* 253(1-3):233-244.
367. Binder H (2007) Water near lipid membranes as seen by infrared spectroscopy. *Eur Biophys J Biophys* 36(4-5):265-279.
368. Gragson DE, McCarty BM, & Richmond GL (1996) Surfactant/water interactions at the air/water interface probed by vibrational sum frequency generation. *J Phys Chem-US* 100(34):14272-14275.

369. Ye S, Nihonyanagi S, & Uosaki K (2001) Sum frequency generation (SFG) study of the pH-dependent water structure on a fused quartz surface modified by an octadecyltrichlorosilane (OTS) monolayer. *Physical Chemistry Chemical Physics* 3(16):3463-3469.
370. Asad A, Mmereki BT, & Donaldson DJ (2004) Enhanced uptake of water by oxidatively processed oleic acid. *Atmos Chem Phys* 4:2083-2089.
371. Donaldson DJ & Vaida V (2006) The influence of organic films at the air-aqueous boundary on atmospheric processes. *Chem Rev* 106(4):1445-1461.
372. Lane JMD, Chandross M, Lorenz CD, Stevens MJ, & Grest GS (2008) Water penetration of damaged self-assembled monolayers. *Langmuir* 24(11):5734-5739.
373. Xu Z, Song K, Yuan SL, & Liu CB (2011) Microscopic Wetting of Self-Assembled Mono layers with Different Surfaces: A Combined Molecular Dynamics and Quantum Mechanics Study. *Langmuir* 27(14):8611-8620.
374. Szori M, Roeselova M, & Jedlovsky P (2011) Surface Hydrophilicity-Dependent Water Adsorption on Mixed Self-Assembled Monolayers of C-7-CH₃ and C-7-COOH Residues. A Grand Canonical Monte Carlo Simulation Study. *J Phys Chem C* 115(39):19165-19177.
375. Ulman A (1996) Formation and structure of self-assembled monolayers. *Chem Rev* 96(4):1533-1554.
376. Porter MD, Bright TB, Allara DL, & Chidsey CED (1987) Spontaneously Organized Molecular Assemblies .4. Structural Characterization of Normal-Alkyl Thiol Monolayers on Gold by Optical Ellipsometry, Infrared-Spectroscopy, and Electrochemistry. *J Am Chem Soc* 109(12):3559-3568.
377. Hopkins AJ, McFearin CL, & Richmond GL (2011) SAMs under Water: The Impact of Ions on the Behavior of Water at Soft Hydrophobic Surfaces. *J Phys Chem C* 115(22):11192-11203.
378. Halverson JD, Maldarelli C, Couzis A, & Koplik J (2010) Atomistic simulations of the wetting behavior of nanodroplets of water on homogeneous and phase separated self-assembled monolayers. *Soft Matter* 6(6):1297-1307.
379. McIntire TM, Ryder O, & Finlayson-Pitts BJ (2009) Secondary Ozonide Formation from the Ozone Oxidation of Unsaturated Self-Assembled Monolayers on Zinc Selenide Attenuated Total Reflectance Crystals. *J Phys Chem C* 113(25):11060-11065.
380. Stevens MJ & Grest GS (2008) Simulations of water at the interface with hydrophilic self-assembled monolayers. *Biointerphases* 3(3):Fc13-Fc22.
381. Berendsen HJC, Grigera JR, & Straatsma TP (1987) The Missing Term in Effective Pair Potentials. *J Phys Chem-US* 91(24):6269-6271.

382. Ryckaert J-P, G. Ciccotti, and H. J. C. Berendsen (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.
383. Corcelli SA, Lawrence CP, & Skinner JL (2004) Combined electronic structure/molecular dynamics approach for ultrafast infrared spectroscopy of dilute HOD in liquid H₂O and D₂O. *J Chem Phys* 120(17):8107-8117.
384. Netzer L, Iscovici R, & Sagiv J (1983) Adsorbed Monolayers Versus Langmuir Blodgett Monolayers - Why and How .2. Characterization of Built-up Films Constructed by Stepwise Adsorption of Individual Monolayers. *Thin Solid Films* 100(1):67-76.
385. Angst DL & Simmons GW (1991) Moisture Absorption Characteristics of Organosiloxane Self-Assembled Monolayers. *Langmuir* 7(10):2236-2242.
386. Diamond ML, *et al.* (2000) Evidence for organic film on an impervious urban surface: Characterization and potential teratogenic effects. *Environ Sci Technol* 34(14):2900-2908.
387. Mashburn CD, Frinak EK, & Tolbert MA (2006) Heterogeneous uptake of nitric acid on Na-montmorillonite clay as a function of relative humidity. *J Geophys Res-Atmos* 111(D15).
388. Puibasset J & Pellenq RJM (2005) Water adsorption in disordered mesoporous silica (Vycor) at 300 K and 650 K: A Grand Canonical Monte Carlo simulation study of hysteresis. *J Chem Phys* 122(9).
389. Croteau T, Bertram AK, & Patey GN (2009) Simulation of Water Adsorption on Kaolinite under Atmospheric Conditions. *J Phys Chem A* 113(27):7826-7833.
390. Giovambattista N, Rossky PJ, & Debenedetti PG (2009) Phase Transitions Induced by Nanoconfinement in Liquid Water. *Phys Rev Lett* 102(5).
391. Viecelli J & Benjamin I (2002) Adsorption at the interface between water and self-assembled monolayers: Structure and electronic spectra. *J Phys Chem B* 106(32):7898-7907.
392. Mahoney MW & Jorgensen WL (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112(20):8910-8922.
393. Vanzo D, Bratko D, & Luzar A (2012) Wettability of pristine and alkyl-functionalized graphane. *J Chem Phys* 137(3).
394. Nagumo R, *et al.* (2012) Molecular dynamics simulations for microscopic behavior of water molecules in the vicinity of zwitterionic self-assembled monolayers. *Polym J* 44(11):1149-1153.

395. Yeh IC & Berkowitz ML (1999) Ewald summation for systems with slab geometry. *J Chem Phys* 111(7):3155-3162.
396. Yeh IC & Wallqvist A (2011) On the proper calculation of electrostatic interactions in solid-supported bilayer systems. *J Chem Phys* 134(5).
397. Flory PJ (1969) *Statistical mechanics of chain molecules* (Interscience Publishers, New York,) pp xix, 432 p.
398. Pappu RV, Srinivasan R, & Rose GD (2000) The Flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding. *Proc Natl Acad Sci U S A* 97(23):12565-12570.
399. Betancourt MR & Skolnick J (2004) Local propensities and statistical potentials of backbone dihedral angles in proteins. *J Mol Biol* 342(2):635-649.
400. Ting D, *et al.* (2010) Neighbor-dependent Ramachandran probability distributions of amino acids developed from a hierarchical Dirichlet process model. *PLoS computational biology* 6(4):e1000763.
401. Hollingsworth SA, Lewis MC, Berkholz DS, Wong WK, & Karplus PA (2012) (ϕ,ψ)(2) motifs: a purely conformation-based fine-grained enumeration of protein parts at the two-residue level. *J Mol Biol* 416(1):78-93.
402. Hollingsworth SA & Karplus PA (2010) A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomolecular concepts* 1(3-4):271-283.
403. DiNitto JP, Cronin TC, & Lambright DG (2003) Membrane recognition and targeting by lipid-binding domains. *Sci STKE* 2003(213):re16.
404. Falke JJ (2007) Membrane Recruitment as a Cancer Mechanism: A Case Study of Akt PH Domain. *Cellscience* 4(2):25-30.
405. Carpten JD, *et al.* (2007) A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* 448(7152):439-444.
406. Chen HC, Ziemba BP, Landgraf KE, Corbin JA, & Falke JJ (2012) Membrane docking geometry of GRP1 PH domain bound to a target lipid bilayer: an EPR site-directed spin-labeling and relaxation study. *PLoS One* 7(3):e33640.
407. Das P, Lahiri A, Lahiri A, & Chakravorty D (2010) Modulation of the arginase pathway in the context of microbial pathogenesis: a metabolic enzyme moonlighting as an immune modulator. *PLoS pathogens* 6(6):e1000899.
408. Christianson DW (2005) Arginase: structure, mechanism, and physiological role in male and female sexual arousal. *Accounts of chemical research* 38(3):191-201.

409. Jenkinson CP, Grody WW, & Cederbaum SD (1996) Comparative properties of arginases. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology* 114(1):107-132.
410. Perskie LL & Rose GD (2010) Physical-chemical determinants of coil conformations in globular proteins. *Protein science : a publication of the Protein Society* 19(6):1127-1136.
411. Ramachandran GN & Sasisekharan V (1956) Cylindrical lattice structure of collagen. *Arch Biochem Biophys* 63(1):255-257.
412. Sasisekharan V (1960) Stereochemical criteria for polypeptide and protein structures. *Collagen*:39-78.
413. Ramachandran GN, Ramakrishnan C, & Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95-99.
414. Wilmot CM & Thornton JM (1990) Beta-turns and their distortions: a proposed new nomenclature. *Protein engineering* 3(6):479-493.
415. Porter LL & Rose GD (2011) Redrawing the Ramachandran plot after inclusion of hydrogen-bonding constraints. *Proc Natl Acad Sci U S A* 108(1):109-113.
416. Voet D & Voet JG (2011) *Biochemistry* (John Wiley & Sons, Hoboken, NJ) 4th Ed pp xxv, 1428, 1453 p.
417. Perskie LL, Street TO, & Rose GD (2008) Structures, basins, and energies: a deconstruction of the Protein Coil Library. *Protein science : a publication of the Protein Society* 17(7):1151-1161.
418. Solis AD (2014) Deriving high-resolution protein backbone structure propensities from all crystal data using the information maximization device. *PLoS One* 9(6):e94334.
419. Hagarman A, *et al.* (2011) Amino acids with hydrogen-bonding side chains have an intrinsic tendency to sample various turn conformations in aqueous solution. *Chemistry* 17(24):6789-6797.
420. Carrascoza F, Zaric S, & Silaghi-Dumitrescu R (2014) Computational study of protein secondary structure elements: Ramachandran plots revisited. *Journal of molecular graphics & modelling* 50:125-133.
421. Amir ED, Kalisman N, & Keasar C (2008) Differentiable, multi-dimensional, knowledge-based energy terms for torsion angle probabilities and propensities. *Proteins* 72(1):62-73.
422. Lennox KP, Dahl DB, Vannucci M, & Tsai JW (2009) Density Estimation for Protein Conformation Angles Using a Bivariate von Mises Distribution and Bayesian Nonparametrics. *Journal of the American Statistical Association* 104(486):586-596.

423. Karplus PA (1996) Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Protein science : a publication of the Protein Society* 5(7):1406-1420.
424. Berkholz DS, Krenesky PB, Davidson JR, & Karplus PA (2010) Protein Geometry Database: a flexible engine to explore backbone conformations and their relationships to covalent geometry. *Nucleic acids research* 38(Database issue):D320-325.
425. Hollingsworth SA, Berkholz DS, & Karplus PA (2009) On the occurrence of linear groups in proteins. *Protein science : a publication of the Protein Society* 18(6):1321-1325.
426. Rohl CA, Strauss CE, Misura KM, & Baker D (2004) Protein structure prediction using Rosetta. *Methods in enzymology* 383:66-93.
427. Rohl CA, Strauss CE, Chivian D, & Baker D (2004) Modeling structurally variable regions in homologous proteins with rosetta. *Proteins* 55(3):656-677.
428. Bowers PM, Strauss CE, & Baker D (2000) De novo protein structure determination using sparse NMR data. *Journal of biomolecular NMR* 18(4):311-318.
429. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
430. Lovell SC, *et al.* (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* 50(3):437-450.
431. Headd JJ, *et al.* (2012) Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta crystallographica. Section D, Biological crystallography* 68(Pt 4):381-390.
432. Moriarty NW, Tronrud DE, Adams PD, & Karplus PA (2014) Conformation-dependent backbone geometry restraints set a new standard for protein crystallographic refinement. *The FEBS journal*.
433. Czech MP (2000) PIP2 and PIP3: complex roles at the cell surface. *Cell* 100(6):603-606.
434. Lemmon MA & Ferguson KM (2000) Signal-dependent membrane targeting by pleckstrin homology (PH) domains. *Biochem J* 350 Pt 1:1-18.
435. Wang F, *et al.* (2002) Lipid products of PI(3)Ks maintain persistent cell polarity and directed motility in neutrophils. *Nat Cell Biol* 4(7):513-518.
436. Calleja V, *et al.* (2007) Intramolecular and intermolecular interactions of protein kinase B define its activation in vivo. *PLoS Biol* 5(4):e95.

437. Calleja V, Laguerre M, Parker PJ, & Larijani B (2009) Role of a novel PH-kinase domain interface in PKB/Akt regulation: structural mechanism for allosteric inhibition. *PLoS Biol* 7(1):e17.
438. Okuzumi T, *et al.* (2009) Inhibitor hijacking of Akt activation. *Nat Chem Biol* 5(7):484-493.
439. Wu WI, *et al.* (2010) Crystal structure of human AKT1 with an allosteric inhibitor reveals a new mode of kinase inhibition. *PLoS One* 5(9):e12913.
440. Lietzke SE, *et al.* (2000) Structural basis of 3-phosphoinositide recognition by pleckstrin homology domains. *Mol Cell* 6(2):385-394.
441. Lumb CN, *et al.* (2011) Biophysical and computational studies of membrane penetration by the GRP1 pleckstrin homology domain. *Structure* 19(9):1338-1346.
442. Lai CL, *et al.* (2013) Molecular mechanism of membrane binding of the GRP1 PH domain. *J Mol Biol* 425(17):3073-3090.
443. Kalé L, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten (1999) NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys. J. Comput. Phys.* 151:283-312.
444. Mackerell AD, Feig, M. and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25(1400-1415).
445. Klauda JB, *et al.* (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B* 114(23):7830-7843.
446. Kossel A & Dakin HD (1904) Arginase. *H-S Z Physiol Chem* 41(4):321-331.
447. Kanyo ZF, Scolnick LR, Ash DE, & Christianson DW (1996) Structure of a unique binuclear manganese cluster in arginase. *Nature* 383(6600):554-557.
448. Di Costanzo L, *et al.* (2005) Crystal structure of human arginase I at 1.29-Å resolution and exploration of inhibition in the immune response. *Proc Natl Acad Sci U S A* 102(37):13058-13063.
449. Herzfeld A & Raper SM (1976) The heterogeneity of arginases in rat tissues. *Biochem J* 153(2):469-478.
450. Morris SM, Jr., Bhamidipati D, & Kepka-Lenhart D (1997) Human type II arginase: sequence analysis and tissue-specific expression. *Gene* 193(2):157-161.

451. Cederbaum SD, *et al.* (2004) Arginases I and II: do their functions overlap? *Molecular genetics and metabolism* 81 Suppl 1:S38-44.
452. Wu G & Morris SM, Jr. (1998) Arginine metabolism: nitric oxide and beyond. *Biochem J* 336 (Pt 1):1-17.
453. Cox JD, Kim NN, Traish AM, & Christianson DW (1999) Arginase-boronic acid complex highlights a physiological role in erectile function. *Nat Struct Biol* 6(11):1043-1047.
454. Kim NN, *et al.* (2001) Probing erectile function: S-(2-boronoethyl)-L-cysteine binds to arginase as a transition state analogue and enhances smooth muscle relaxation in human penile corpus cavernosum. *Biochemistry* 40(9):2678-2688.
455. Baggio R, *et al.* (1999) Biochemical and functional profile of a newly developed potent and isozyme-selective arginase inhibitor. *The Journal of pharmacology and experimental therapeutics* 290(3):1409-1416.
456. Xu L, *et al.* (2003) Arginase and autoimmune inflammation in the central nervous system. *Immunology* 110(1):141-148.
457. Shearer JD, Richards JR, Mills CD, & Caldwell MD (1997) Differential regulation of macrophage arginine metabolism: A proposed role in wound healing. *Am J Physiol-Endoc M* 272(2):E181-E190.
458. Albina JE, Mills CD, Henry WL, & Caldwell MD (1990) Temporal Expression of Different Pathways of L-Arginine Metabolism in Healing Wounds. *J Immunol* 144(10):3877-3880.
459. Shapiro BJ, *et al.* (1993) A Novel Role for Arginase All in Macrophage Killing - Possible Arginine Titration for Nitric-Oxide (No) Synthesis. *Am J Hum Genet* 53(3):1629-1629.
460. Moncada S, Palmer RM, & Higgs EA (1989) Biosynthesis of nitric oxide from L-arginine. A pathway for the regulation of cell function and communication. *Biochemical pharmacology* 38(11):1709-1715.
461. Nakamura H, Saheki T, & Nakagawa S (1990) Differential cellular localization of enzymes of L-arginine metabolism in the rat brain. *Brain research* 530(1):108-112.
462. Vogiatzi P, *et al.* (2007) Helicobacter pylori as a class I carcinogen: physiopathology and management strategies. *Journal of cellular biochemistry* 102(2):264-273.
463. Gobert AP, *et al.* (2001) Helicobacter pylori arginase inhibits nitric oxide production by eukaryotic cells: a strategy for bacterial survival. *Proc Natl Acad Sci U S A* 98(24):13844-13849.

464. Chaturvedi R, *et al.* (2007) L-arginine availability regulates inducible nitric oxide synthase-dependent host defense against *Helicobacter pylori*. *Infection and immunity* 75(9):4305-4315.
465. McGee DJ, Radcliff FJ, Mendz GL, Ferrero RL, & Mobley HL (1999) *Helicobacter pylori* rocF is required for arginase activity and acid protection in vitro but is not essential for colonization of mice or for urease activity. *Journal of bacteriology* 181(23):7314-7322.
466. Gobert AP, *et al.* (2002) *Helicobacter pylori* induces macrophage apoptosis by activation of arginase II. *J Immunol* 168(9):4692-4700.
467. Kropf P, *et al.* (2005) Arginase and polyamine synthesis are key factors in the regulation of experimental leishmaniasis in vivo. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 19(8):1000-1002.
468. Iniesta V, *et al.* (2005) Arginase I induction during *Leishmania major* infection mediates the development of disease. *Infection and immunity* 73(9):6085-6090.
469. Adak S, Aulak KS, & Stuehr DJ (2002) Direct evidence for nitric oxide production by a nitric-oxide synthase-like protein from *Bacillus subtilis*. *J Biol Chem* 277(18):16167-16171.
470. Adak S, *et al.* (2002) Cloning, expression, and characterization of a nitric oxide synthase protein from *Deinococcus radiodurans*. *Proc Natl Acad Sci U S A* 99(1):107-112.
471. Pant K, Bilwes AM, Adak S, Stuehr DJ, & Crane BR (2002) Structure of a nitric oxide synthase heme protein from *Bacillus subtilis*. *Biochemistry* 41(37):11071-11079.
472. Gusarov I & Nudler E (2005) NO-mediated cytoprotection: instant adaptation to oxidative stress in bacteria. *Proc Natl Acad Sci U S A* 102(39):13855-13860.
473. Shatalin K, *et al.* (2008) *Bacillus anthracis*-derived nitric oxide is essential for pathogen virulence and survival in macrophages. *Proc Natl Acad Sci U S A* 105(3):1009-1013.
474. Read RJ & Sussman J (2007) *Evolving methods for macromolecular crystallography : the structural path to the understanding of the mechanisms of action of CBRN agents* (Springer Verlag, Dordrecht) pp viii, 191 p.
475. Bewley MC, Jeffrey PD, Patchett ML, Kanyo ZF, & Baker EN (1999) Crystal structures of *Bacillus caldovelox* arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily. *Structure* 7(4):435-448.
476. Vagin AA, *et al.* (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D* 60:2184-2195.

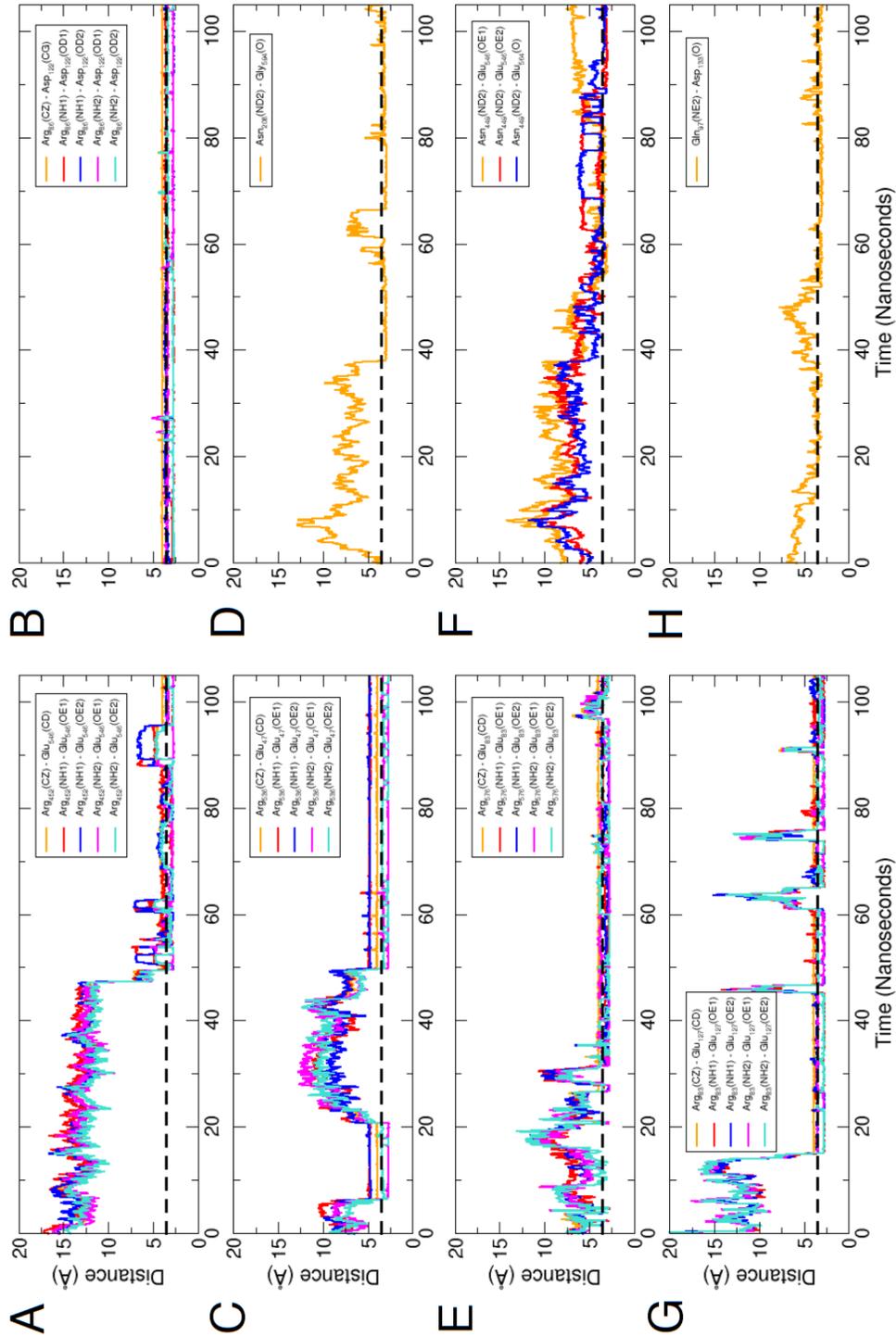
Appendix A

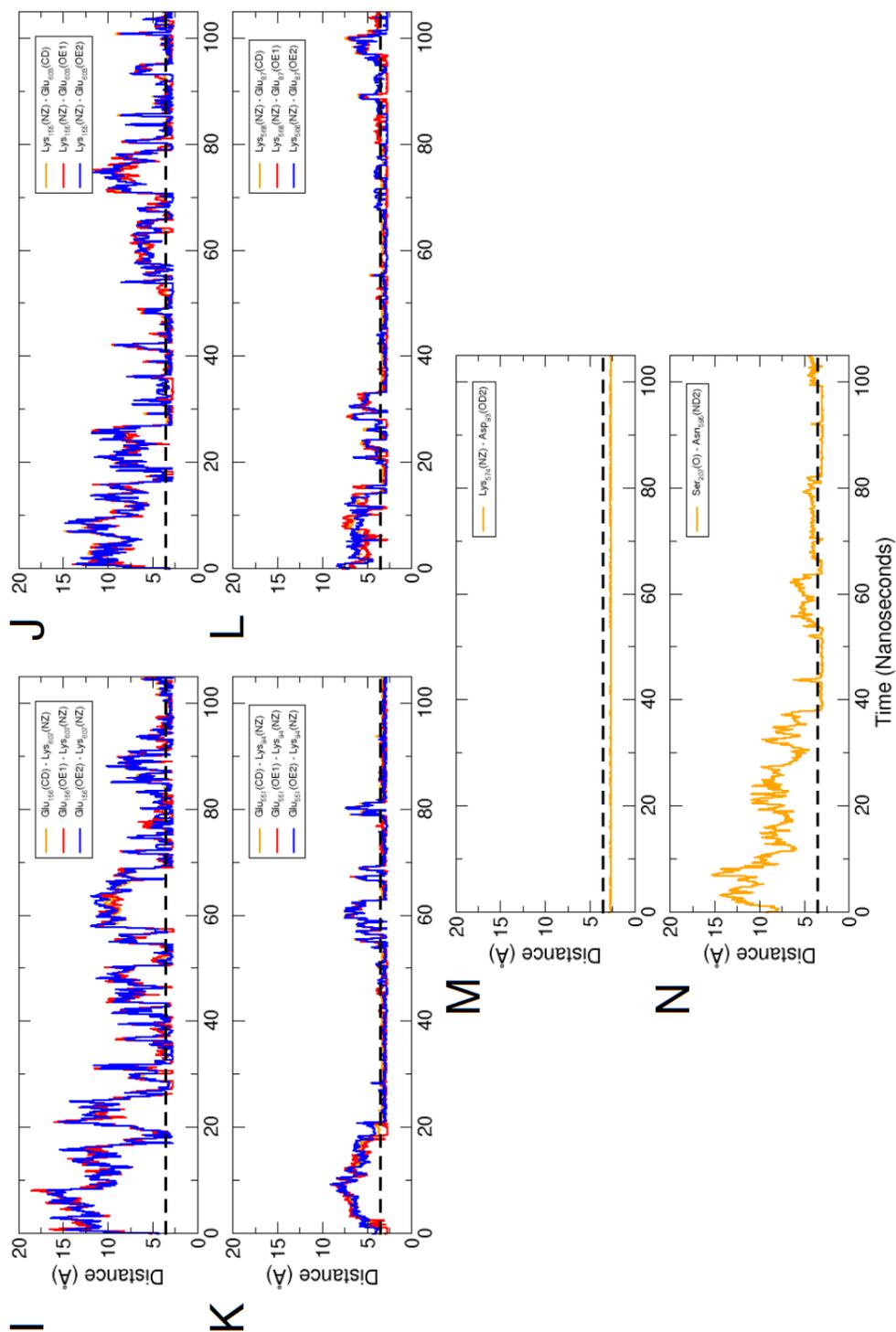
Detailed Comparison of the Interfacial Interactions of P450cam-Pdx

P450cam Residue	P450cam Atom	Pdx Residue	Pdx Atom	4JX1 (Reduced) Distance (Å)	3W9C Distance (Å)	MD Distance & σ of Final 25ns (Å)
Glu76	OE1	Ser42	OG	5.8	7.0	3.43 (0.81)
Glu76	OE2	Ser42	OG	4.9	5.3	3.58 (0.70)
Glu76	OE1	Ser42	N	7.2	8.7	3.47 (0.73)
Glu76	OE2	Ser42	N	6.1	6.7	3.75 (0.77)
Glu76	O	Arg66	NH1	2.8	2.9	6.60 (1.85)
Glu76	O	Arg66	NH2	3.2	3.3	6.46 (1.99)
Glu76	O	Arg66	CZ	3.4	3.5	6.76 (1.62)
Glu76	OE1	Arg66	NE	3.5	4.4	7.08 (1.45)
Glu76	CB	Arg66	CZ	3.6	3.7	7.39 (1.66)
Glu76	CB	Arg66	NH2	3.5	3.4	7.10 (1.90)
Arg109	NE	Cys45	O	3.0	5.4	2.91 (0.15)
Arg109	CZ	Cys45	O	3.5	5.3	3.30 (0.12)
Arg109	NH2	Cys45	O	3.0	6.3	2.89 (0.15)
Arg109	NH2	Thr47	CG2	3.3	9.5	3.52 (0.24)
Arg109	NH2	Gln105	O	3.0	9.8	2.88 (0.13)
Arg109	NH1	Trp106	O	3.1	10.8	4.07 (0.98)
Arg109	NH1	Trp106	OXT	4.8	9.6	3.70 (0.98)
Arg112	NH2	Asp38	OD1	3.4	3.6	4.31 (0.45)
Arg112	NE	Asp38	OD2	2.7	3.6	3.30 (0.51)
Arg112	NH2	Asp38	OD2	3.0	3.4	2.88 (0.22)
Arg112	NH2	Asp38	O	2.8	3.3	4.53 (0.56)
Arg112	NE	Asp38	CG	3.6	3.5	4.46 (0.52)
Arg112	NH2	Asp38	CG	3.3	3.8	3.86 (0.27)
Arg112	CZ	Asp38	OD2	3.3	4.0	3.51 (0.31)
Arg112	CZ	Asp38	OD1	4.0	3.6	5.25 (0.47)
Ala113	CB	Trp106	CD2	3.6	3.7	3.58 (0.18)
Ala113	CB	Trp106	CE3	3.5	3.6	3.55 (0.18)
Asn116	OD1	Trp106	NE1	3.1	5.0	3.28 (0.43)
Asn116	CB	Trp106	CZ2	3.6	4.0	3.82 (0.25)
Asn116	OD1	Val36	CG1	3.4	4.1	4.99 (0.66)
Asn116	CG	Trp106	NE1	3.7	5.0	3.55 (0.36)
Met121	CE	Val28	O	3.6	3.3	4.78 (1.00)
Pro122	CG	Tyr33	CD1	3.6	4.1	4.33 (0.47)
Asp125	OD2	Tyr33	CE1	3.6	6.8	6.29 (0.70)
Asp125	OD2	Tyr33	CZ	3.5	5.5	5.40 (0.71)
Asp125	OD2	Tyr33	OH	2.7	4.4	4.33 (0.78)
Asp125	CB	Tyr33	OH	3.5	3.2	4.33 (0.72)
Asp125	CG	Tyr33	OH	3.5	3.2	4.14 (0.74)
Gln343	OE1	Val17	CA	17.7	15.8	3.52 (0.37)
Gln343	OE1	Ala18	N	15.3	13.4	3.14 (0.49)
Val345	O	Gln25	NE2	8.1	11.2	3.17 (0.42)
Val345	N	Gln25	OE1	10.0	12.4	3.32 (0.48)
His352	CE1	Gly40	O	6.6	7.0	3.34 (0.49)
Gly353	O	Cys39	O	3.6	4.3	4.56 (0.31)
Gly353	CA	Ser42	OG	3.5	4.0	8.28 (0.55)
Gly353	O	Gly40	N	5.6	6.2	3.23 (0.32)
Gly353	O	Gly41	N	6.7	7.0	3.24 (0.49)
Ser354	N	Ser44	CB	3.6	3.8	5.81 (0.57)
Leu356	CB	Cys39	O	3.4	3.5	7.52 (0.20)
Leu356	CD1	Cys39	O	3.3	3.8	5.63 (0.26)
Gln360	NE2	Asp38	O	4.5	4.4	3.02 (0.33)
Gln360	NE2	Asp38	OD1	6.7	8.2	3.50 (0.68)
Heme	Fe	Camphor	H15	N/A	N/A	5.83 (1.38)
Tyr96	OH	Camphor	OH	8.0	N/A	5.98 (1.05)
Heme	Fe	Fe ₂ S ₂	Fe1	16.1	16.3	16.05 (0.27)

Appendix B

Time Evolution of the Major Interdomain Contacts of Nitric Oxide Synthase





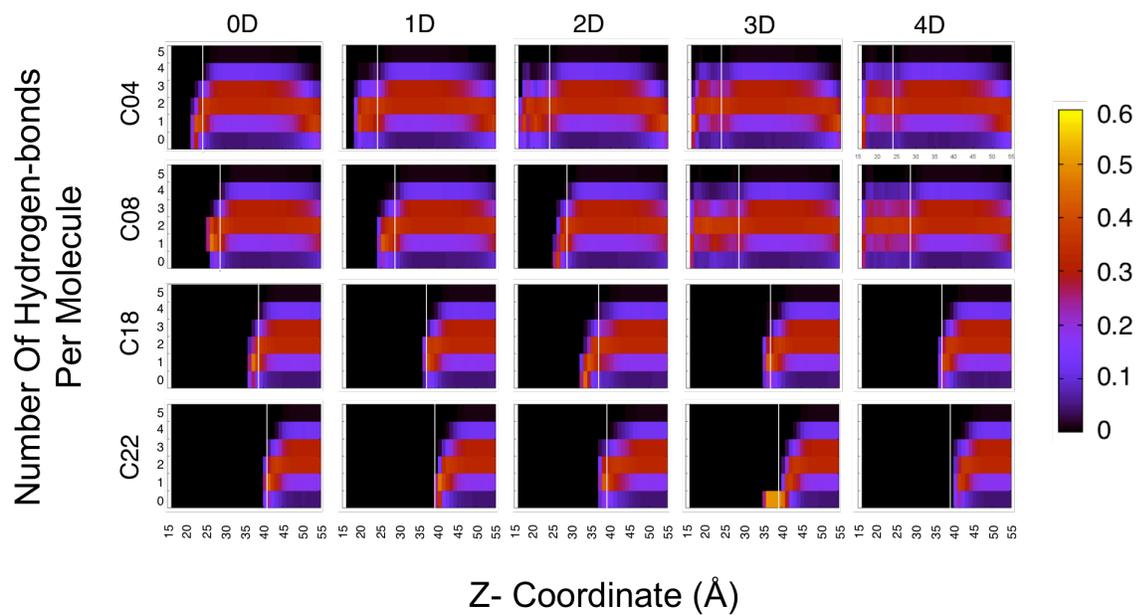
Time evolutions of major interdomain contacts. Shown is a set of time distance evolutions for select atom-to-atom distances for each of the 14 residue pairs that averaged a contact (defined by a distance of 3.5 Å) over the final 50 nanoseconds of iNOS bidomain simulation. Multiple lines are included to capture the dynamic movement of each contact and are labeled within each panel.

Appendix C

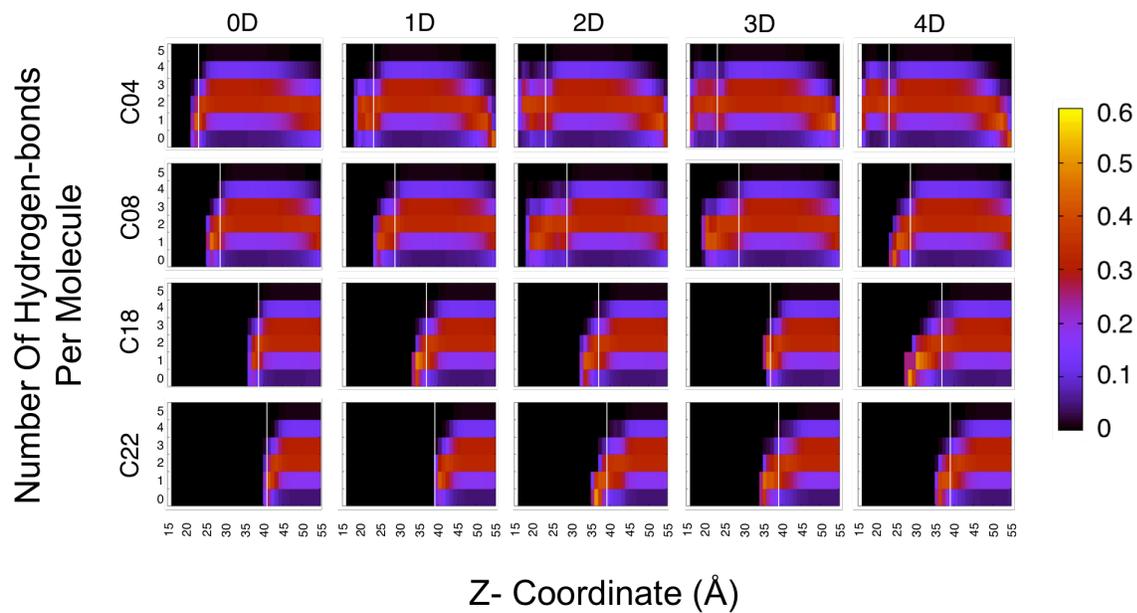
Hydrogen Bonds Per Molecule Distribution Along the Z-Axis for Defect Free SAM Surfaces

Hydrogen bonds per molecule distribution along the z-axis for defect free SAM surfaces. Intensity maps of the average number of hydrogen bonds per water molecule along the z-axis are shown for each of the perfect (defect-free) systems over the final two nanosecond of each hydrated simulation. The z-axis was discretized into 0.5Å bins while each z slab was further divided into six bins based on the number of hydrogen-bonds per water molecule (0 to 5 hydrogen-bonds). Data was normalized by dividing each bin by the total number of occurrences for the entire given z-slab. The solid white line represents the approximate center of the terminal carbon density as taken from Figure 9.2. The hydrophobic surfaces both form droplets that reach well-hydrogen bonded bulk like water 10-15Å away from the surface and a poorly hydrogen bonded layer of water near the surface. The hydrophilic surfaces however have well-hydrogen bonded water located directly at the SAM surface and poor hydrogen bonding for neighboring waters. For the hydrophilic surfaces, there are regions that correspond only to water with no hydrogen-bonding patterns which denotes a water molecule that has escaped the monolayer into the vapor phase.

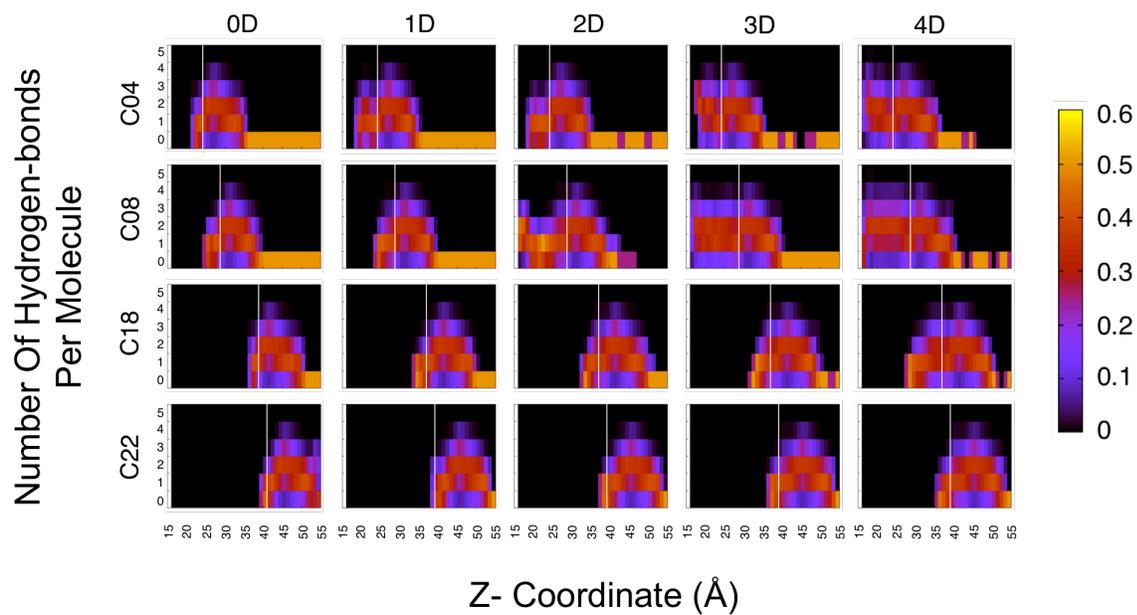
Alkane



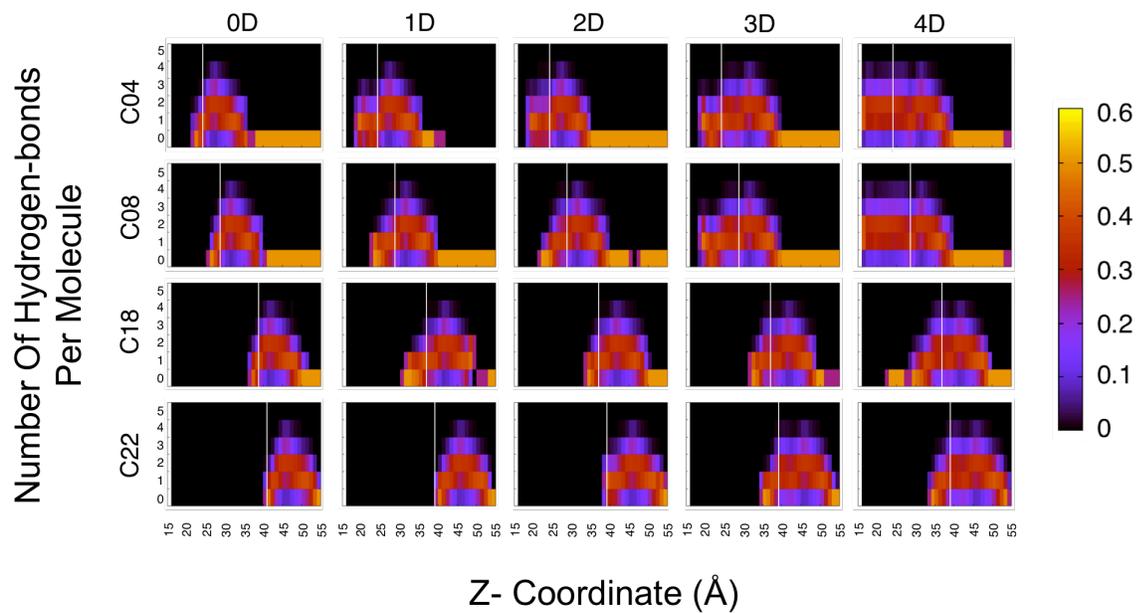
Alkene



Hydroxyl



Carboxylic Acid

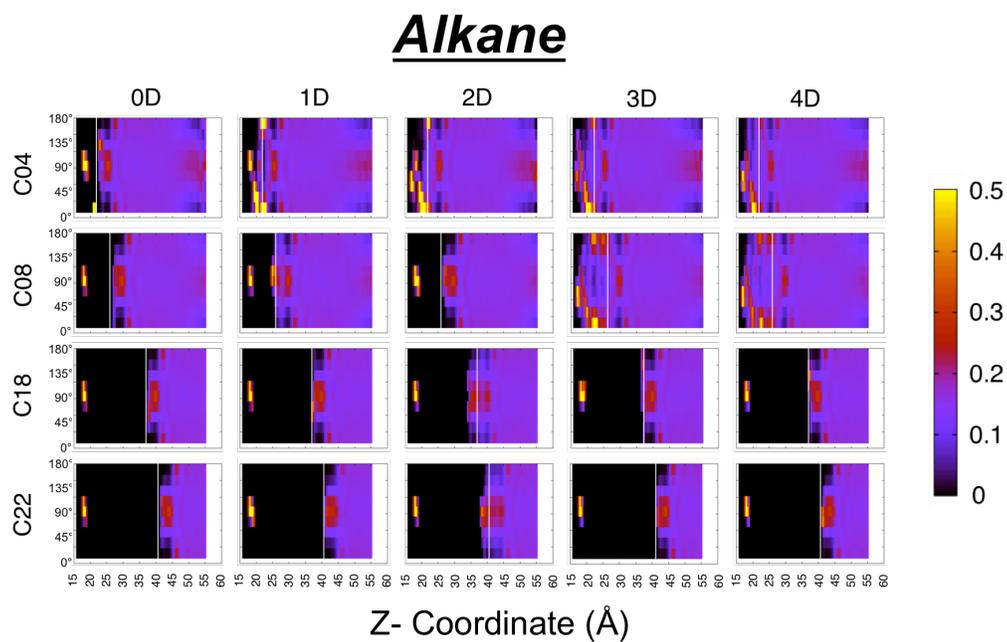


Appendix D

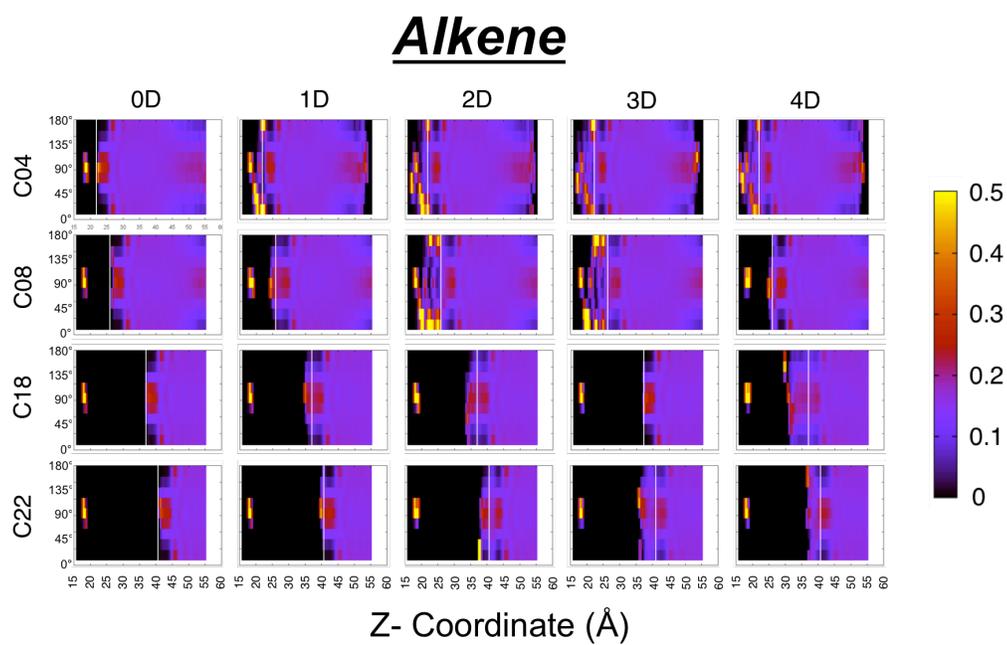
Hydrogen-Bond Orientation Distributions Along the Z-Axis for Defect Free SAM Surfaces

Hydrogen-bond orientation distributions along the z-axis for defect free SAM surfaces. Intensity maps of the angle formed between a given hydrogen bond and a vector normal to the SAM surface as a function of the z-axis. The z-axis was divided up into 0.5Å bins while the angles formed for a given z-bin was divided into 9 bins. The resulting data was then normalized by dividing the number of occurrences for each angle bin by the total number of observations for all angle bins for a given z-bin. The solid white line represents the approximate center of the terminal carbon density as taken from Figure 9.2. This analysis was carried out over the final two nanoseconds of each hydrated simulation. The peak located beneath the SAM surface are hydrogen bonds between the siloxanes themselves, which also allows for further location context. For a given system, structured layers of water can be identified by a strong preference for a given hydrogen-bond orientation value, normally either at 90° or 180°. There appears to be no significant difference dependent on chain length on structured layers of water. See results and discussion for a more discussion of the observed structure water layers on each terminal group.

Angle Between Hydrogen Bond
And Vector Normal To SAM Surface

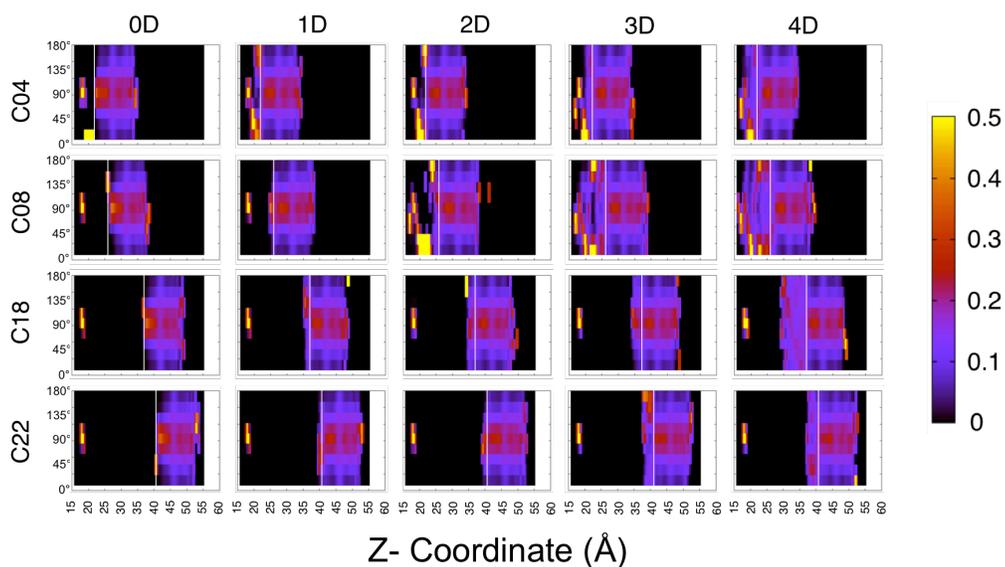


Angle Between Hydrogen Bond
And Vector Normal To SAM Surface



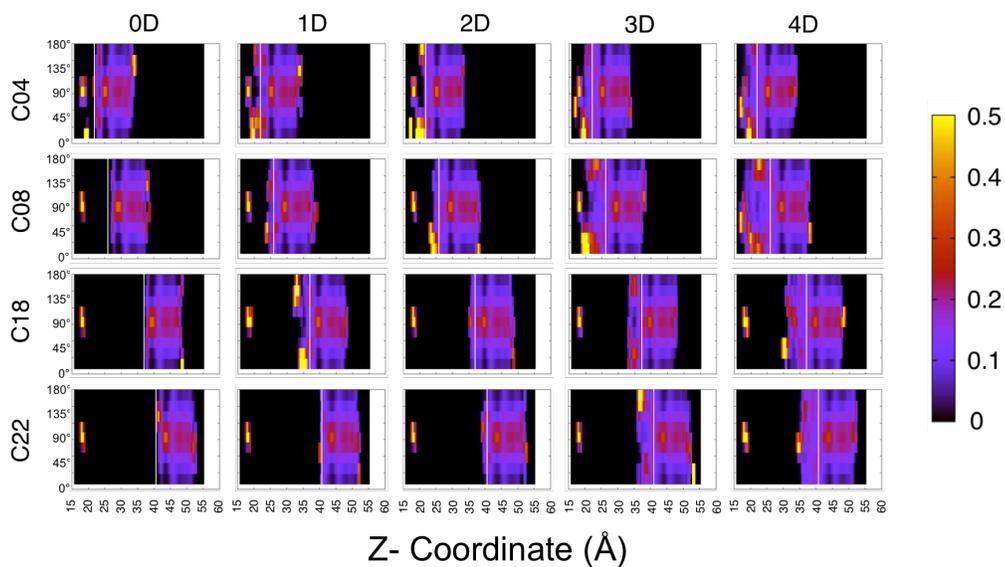
Angle Between Hydrogen Bond
And Vector Normal To SAM Surface

Hydroxyl



Angle Between Hydrogen Bond
And Vector Normal To SAM Surface

Carboxylic Acid

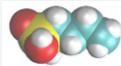


Appendix E

Density Profiles for the Water in All Defective Systematically Varied SAM Systems

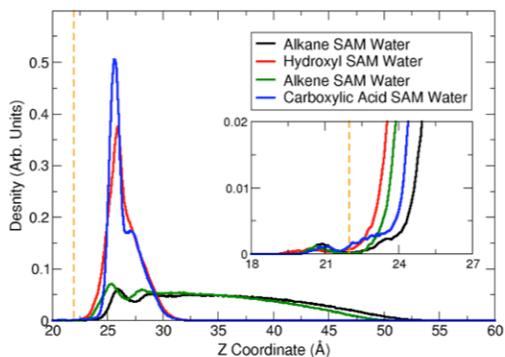
Density Profiles for the water in all defective systematically varied SAM systems.

Shown are density profiles similar to that shown in Figure 4 for all defective systems with a chain length of A) C4, B) C8, C) C18 and D) C22. For each panel, the chain length is designated by the descriptor in the top box, the size of the defect is labeled above each plot while the solvent exposed terminal group can be identified by the color of the profile in each plot.

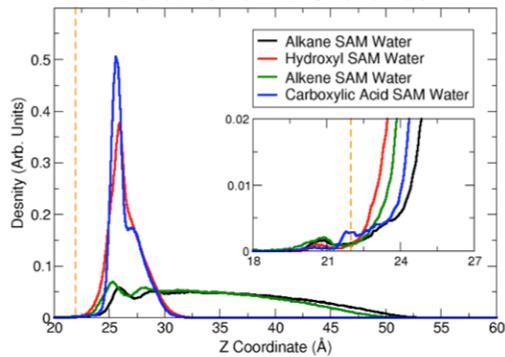


Chain Length: C4

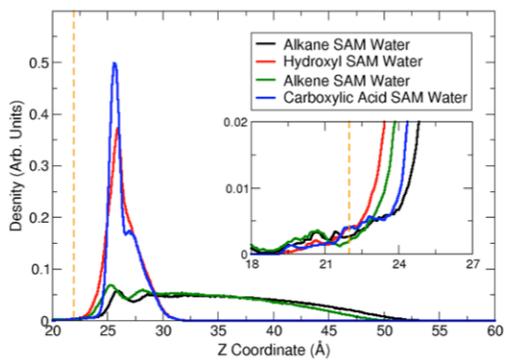
Defect size = 1 Siloxane



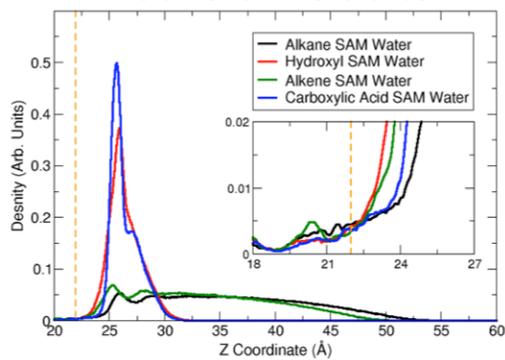
Defect size = 2 Siloxanes

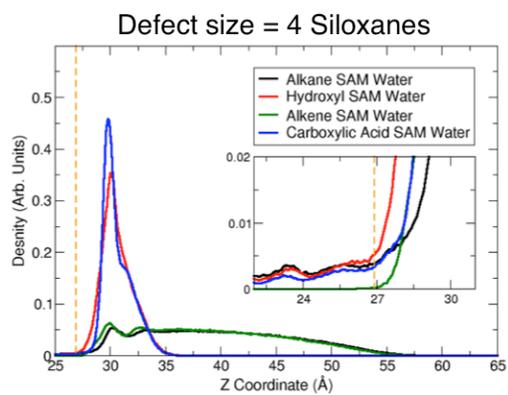
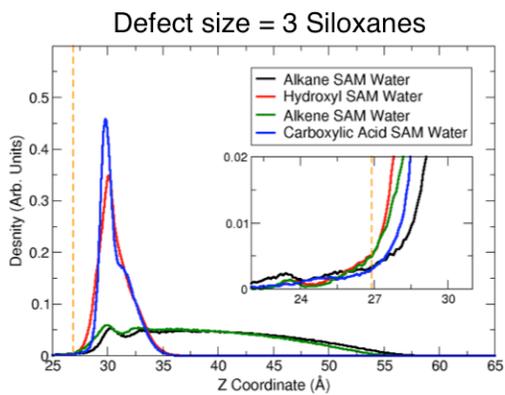
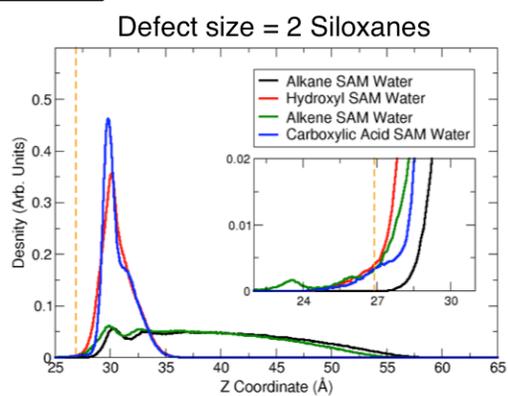
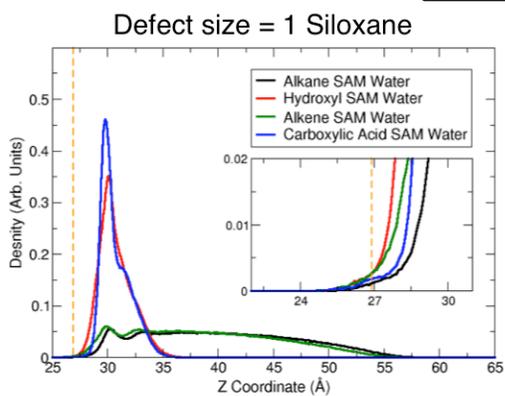
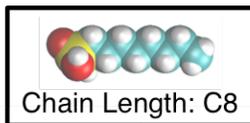


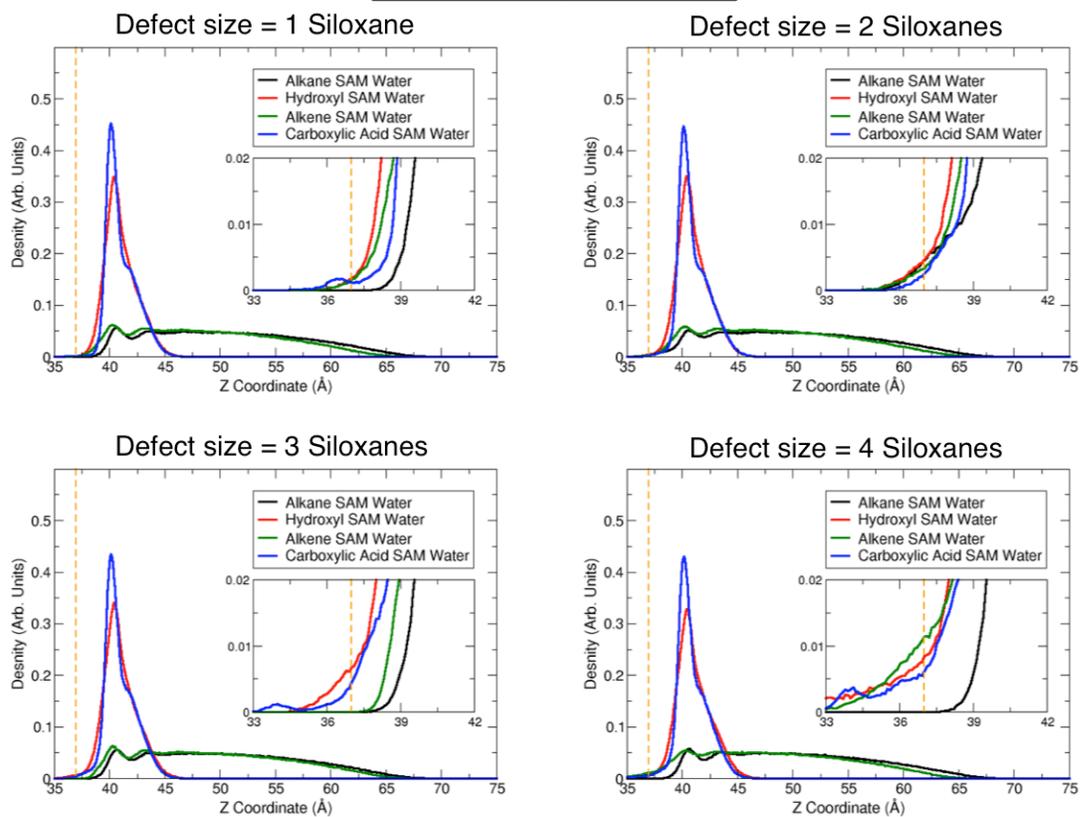
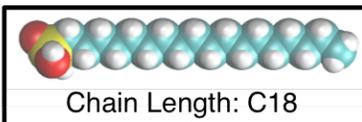
Defect size = 3 Siloxanes

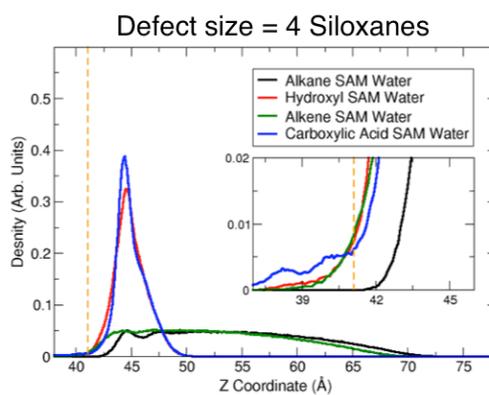
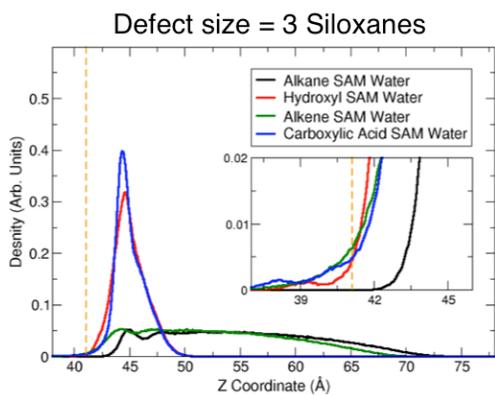
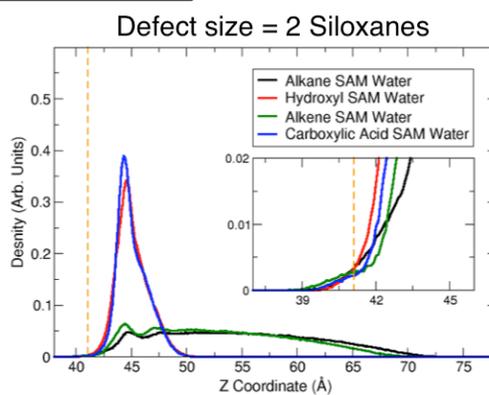
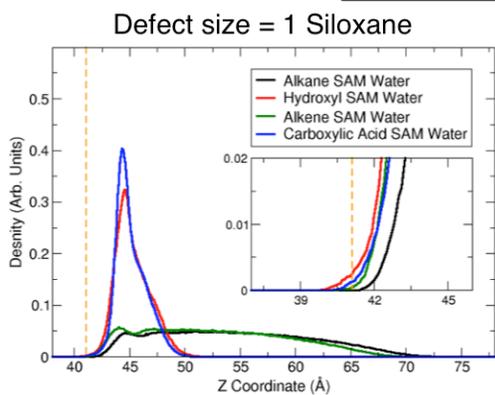
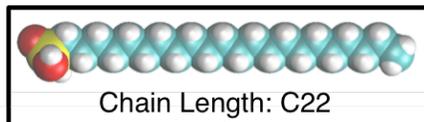


Defect size = 4 Siloxanes









Appendix F

Conformational Dependence of Central Residue on Neighboring Conformations By Basin

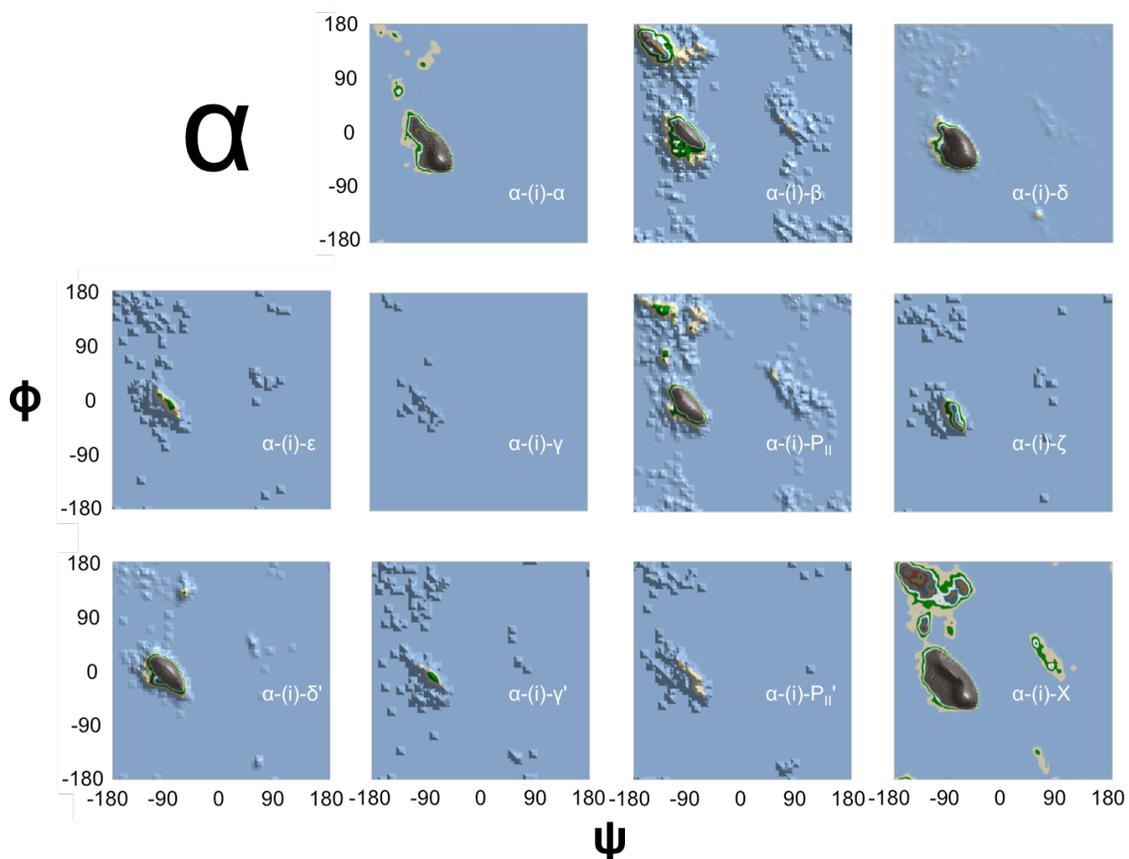


Figure F.1. Neighbor dependence on the central residue of an α -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue $i-1$ is restrained to the α region while residue $i+1$ is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

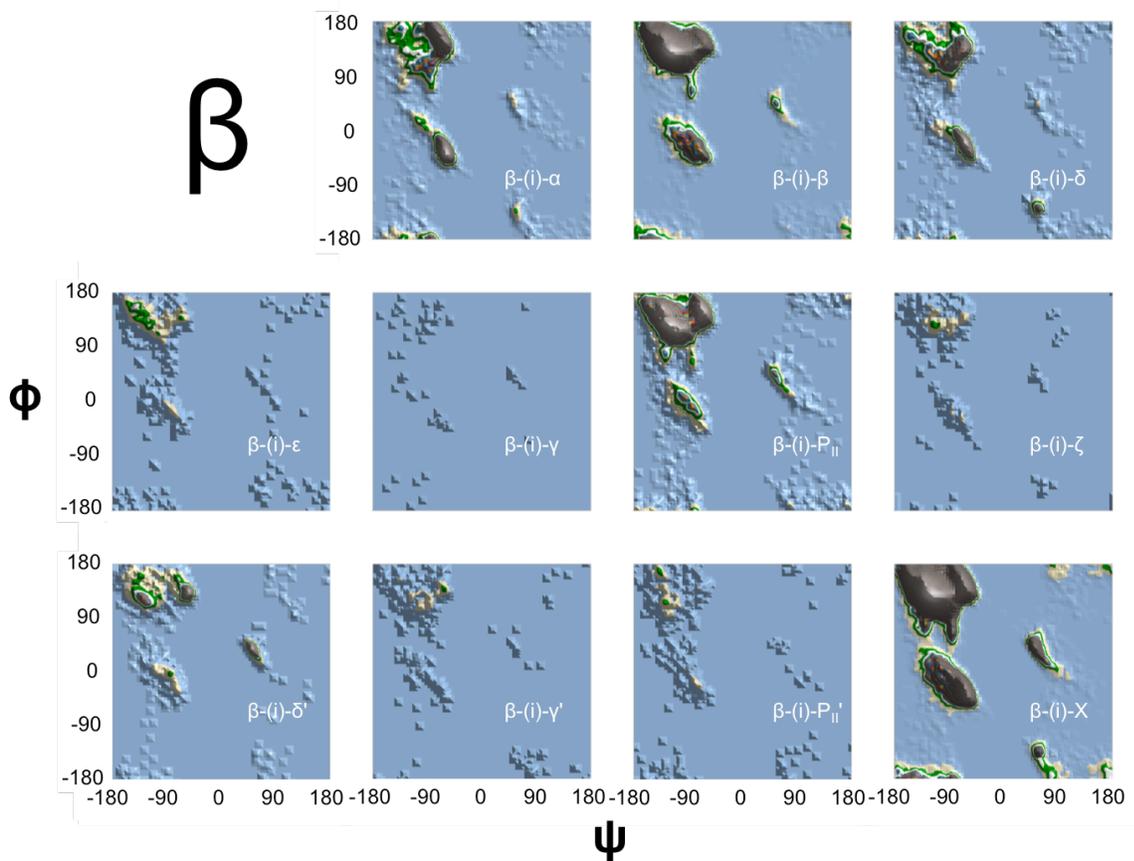


Figure F.2. Neighbor dependence on the central residue of a β -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue $i-1$ is restrained to the β region while residue $i+1$ is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

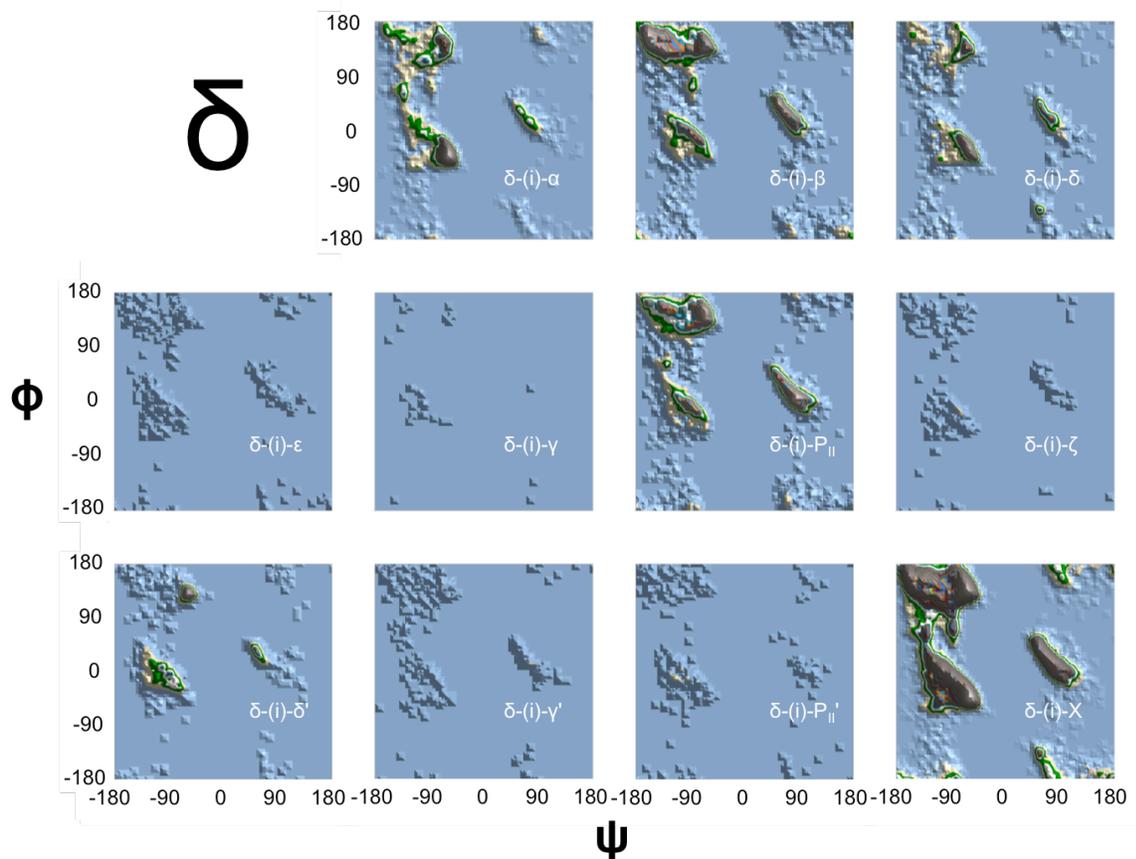


Figure F.3. Neighbor dependence on the central residue of a δ -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the δ region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

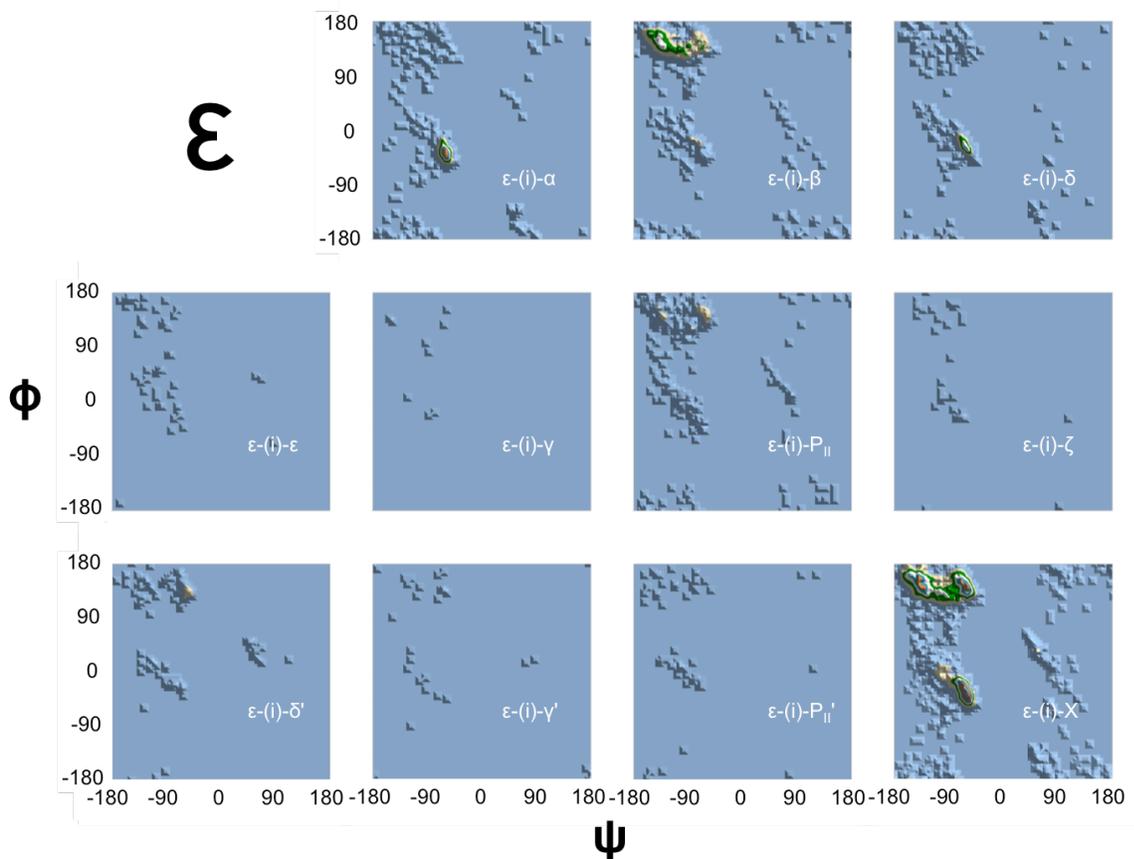


Figure F.4. Neighbor dependence on the central residue of a ϵ -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the ϵ region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

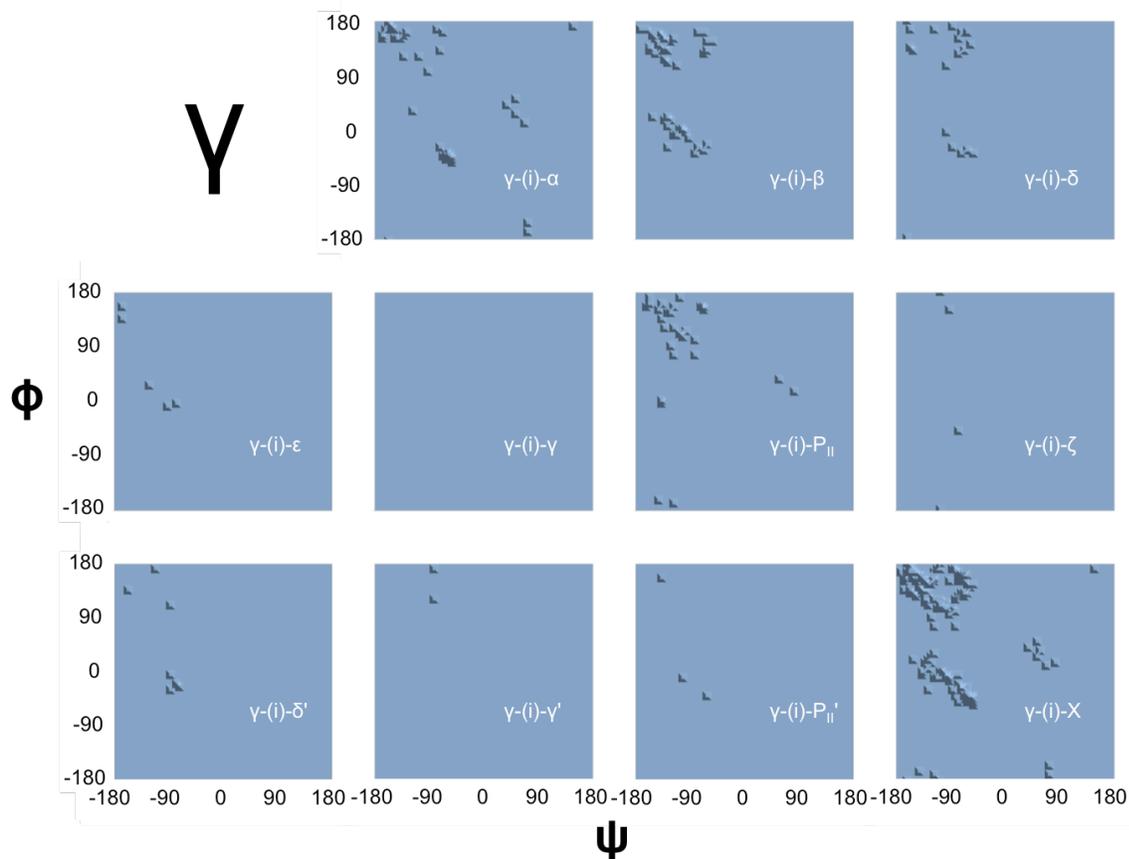


Figure F.5. Neighbor dependence on the central residue of a γ -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the γ region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

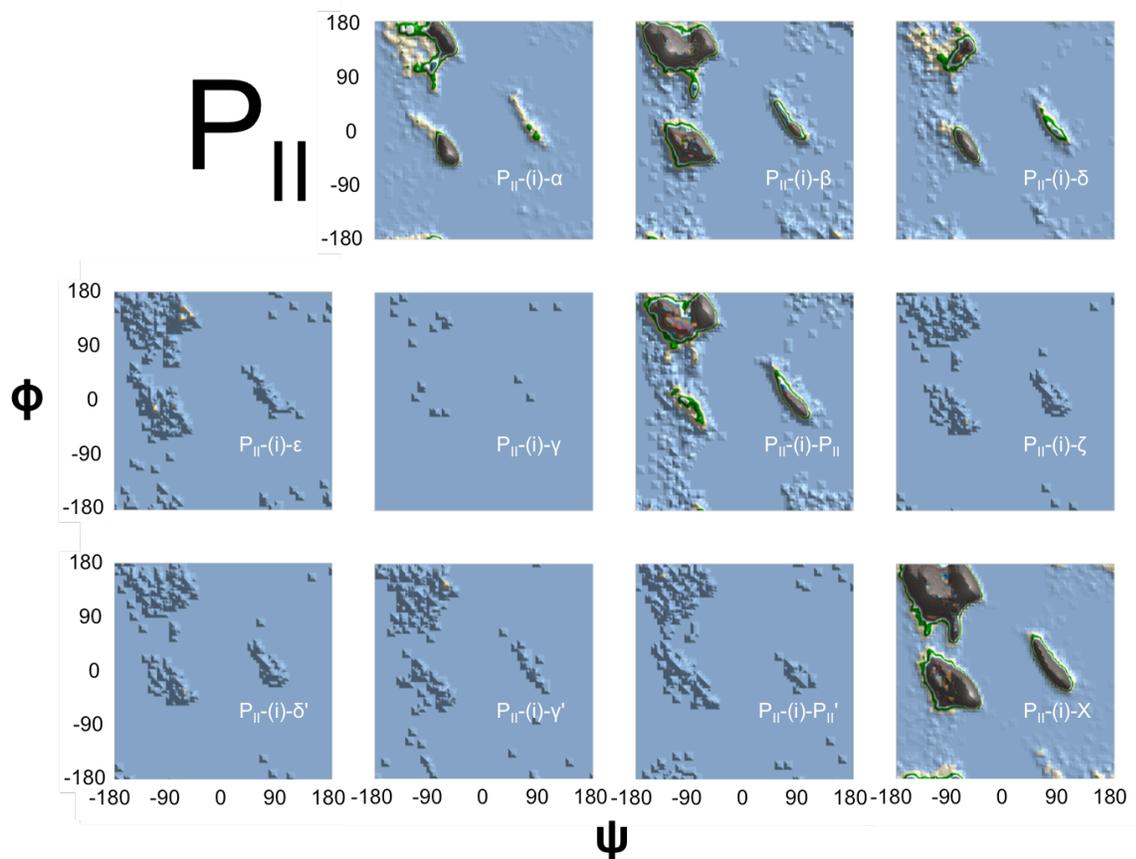


Figure F.6. Neighbor dependence on the central residue of a P_{II} -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue $i-1$ is restrained to the P_{II} region while residue $i+1$ is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

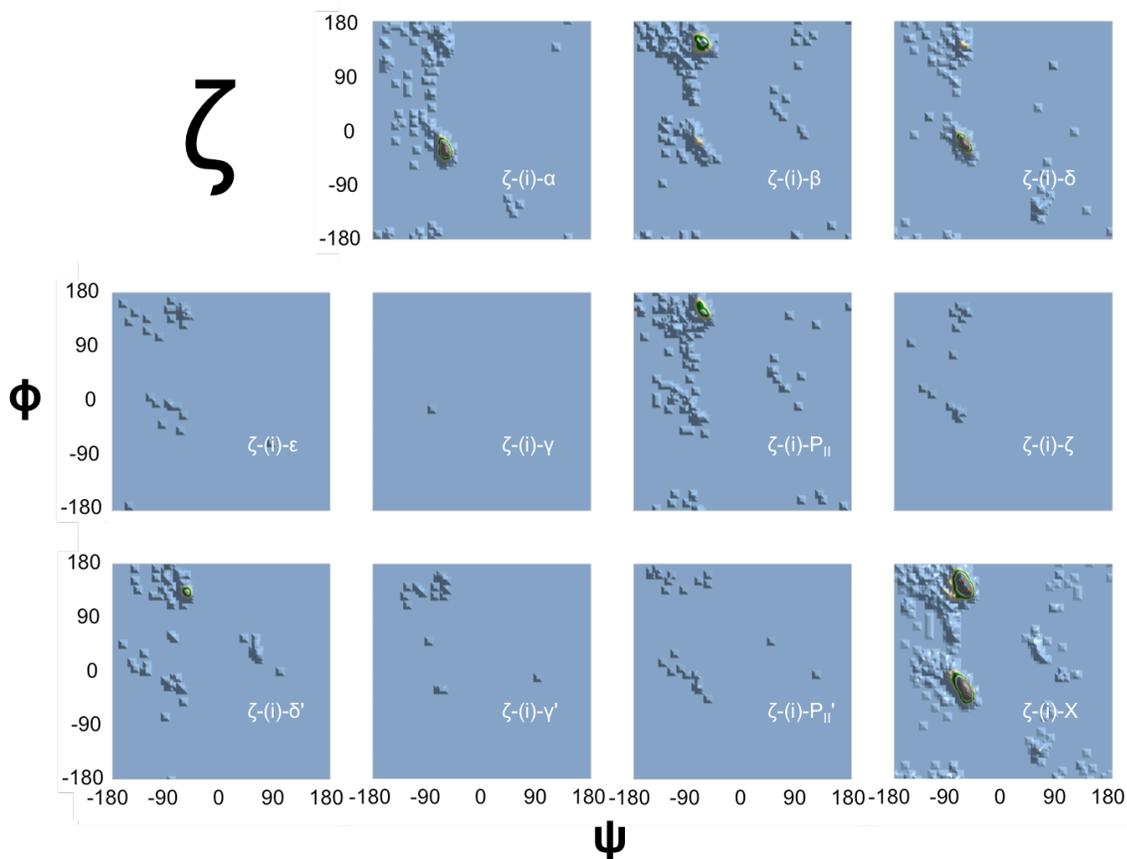


Figure F.7. Neighbor dependence on the central residue of a ζ -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the ζ region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

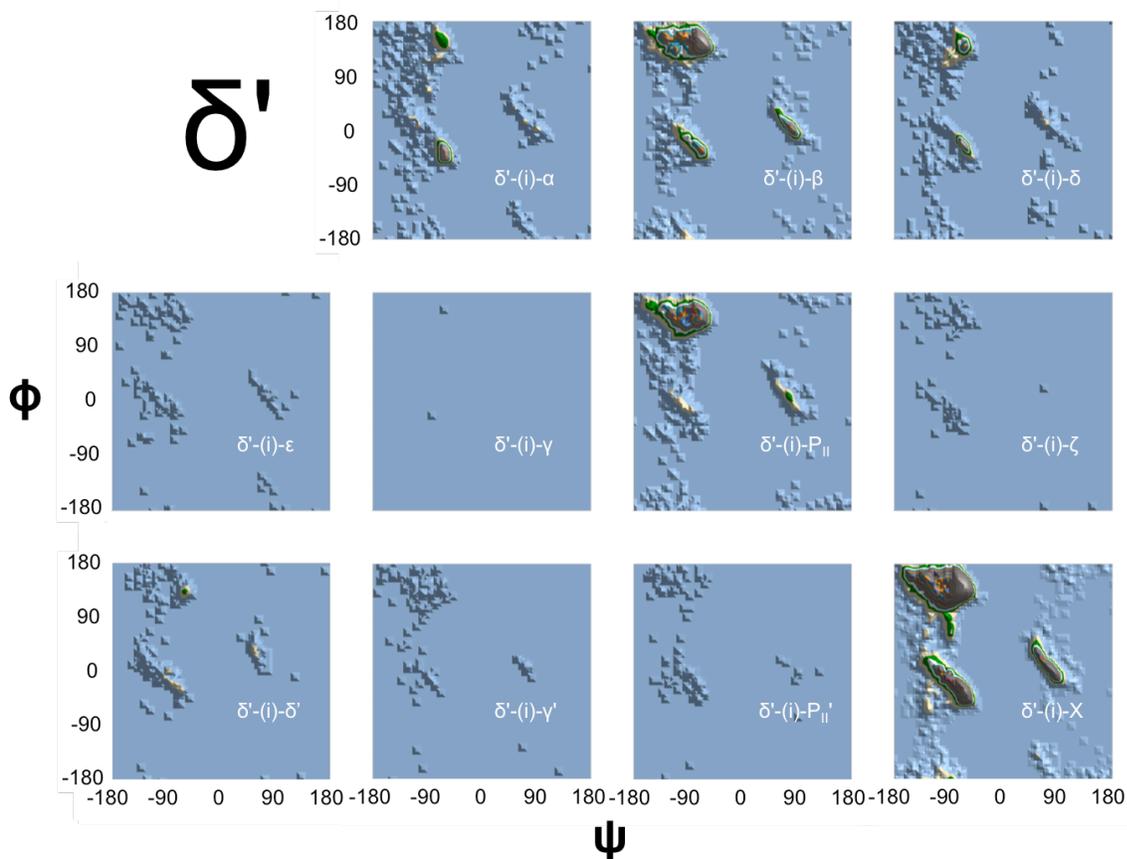


Figure F.8. Neighbor dependence on the central residue of a δ' -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the δ' region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

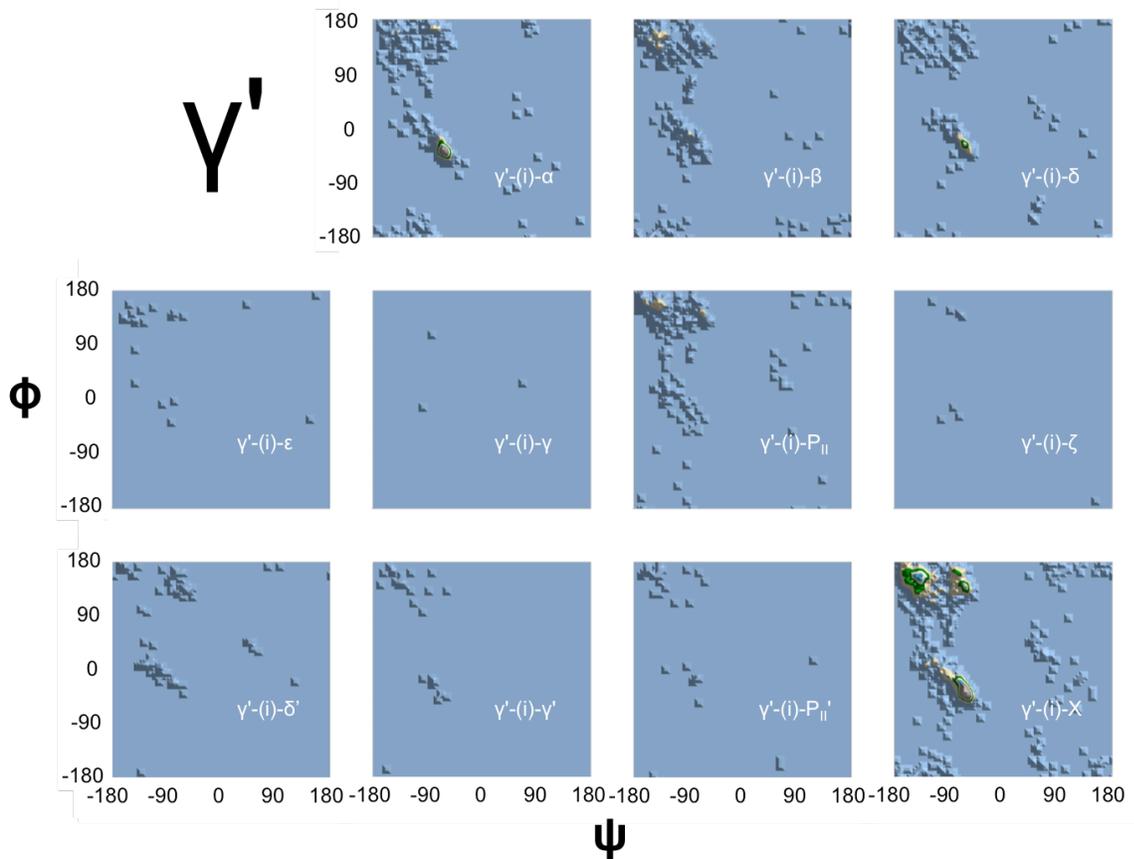


Figure F.9. Neighbor dependence on the central residue of a γ' -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the γ' region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

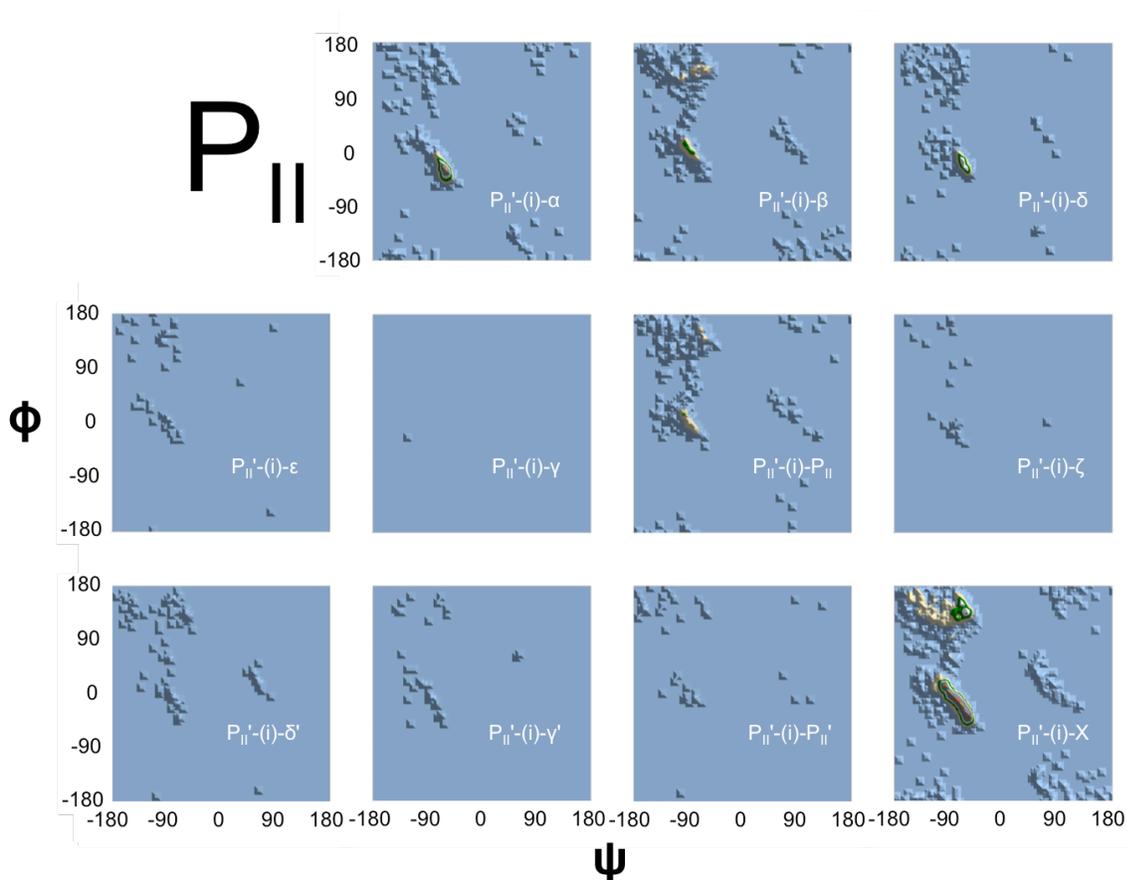


Figure F.10. Neighbor dependence on the central residue of a P_{II}' -X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for a tripeptide where residue i-1 is restrained to the P_{II}' region while residue i+1 is restrained to the designated region as defined by Figure 11.1C. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.

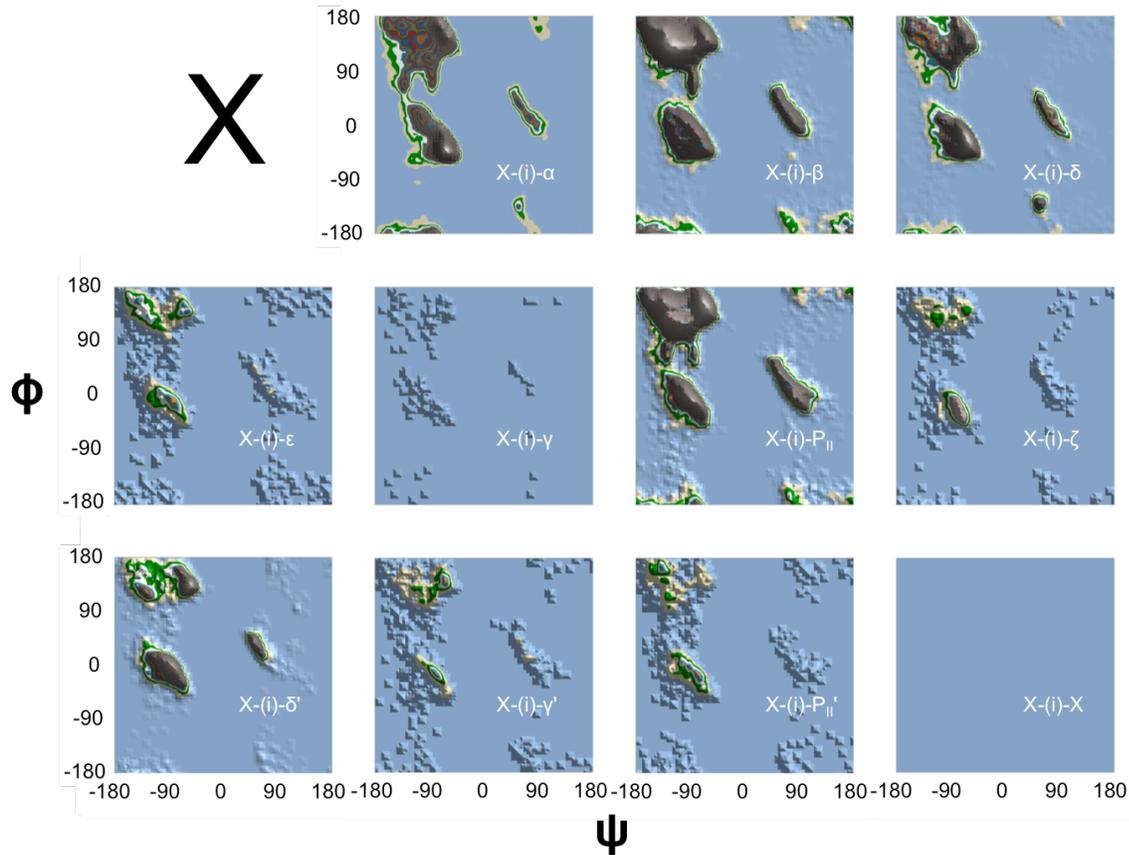


Figure F.11. Neighbor dependence on the central residue of a X-X-X tripeptide. Displayed are contoured Ramachandran plots for the central residue (residue i) for an unrestrained tripeptide. The populations for each distribution are as given in Table 11.1. Colored low contour levels in the observed distributions are: blue (0-10), light brown (11-20), green (21-30), white (31-40). Larger contour values then enter a repeated pattern of teal, orange, blue, red, green and purple.