ACYL CARRIER PROTEIN DYNAMICS AND PARTNER ENZYME INTERACTIONS

by

Peter W. Murphy

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
May 2014

Dedicated in memory of my cousin:

Jeffrey Donald Bates

January 7th, 2000 – December 8th, 2001

The inspiration behind my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Acyl carrier protein (ACP) is essential for the synthesis of fatty acids, phospholipids, lipid A and other primary and secondary metabolites. ACP and ACP-like proteins are found throughout nature; however, differences between prokaryotic and eukaryotic ACP makes bacterial ACP-dependent enzymes attractive antibacterial targets. In the 50 years since its initial discovery, the structures of many bacterial ACPs (including that of *Vibrio harveyi*, the focus of this work) have been determined: all sharing a dynamic four α-helix bundle structure with a hydrophobic binding pocket enclosing attached fatty acyl chains. It has been hypothesized that ACP must undergo a conformational change to expose the sequestered acyl chain to a partner enzyme; the mechanism of and residues important to this conformational change remain unknown.

Towards this end, I have explored the effect of constraining ACP's termini on its conformational stability and function. Employing split-intein technology, a cyclized version of ACP was constructed. Using various biophysical techniques, I demonstrated that cyclic ACP was stabilized in the folded conformation *in vitro* relative to its linear counterpart. Furthermore, *in vivo* complementation assays proved that, counter to the prevailing hypothesis mentioned above, cyclic ACP can functionally replace the linear wild-type protein and support growth of an *E. coli* ACP-null strain.

Additionally, I have expanded the use of fluorescence methods for studying ACP conformation, dynamics, and interaction with partner enzymes. Previous work in our lab established the use of Trp as a fluorescent probe of ACP conformation. To extend the utility of fluorescence to our lab's extensive mutant ACP collection, the lone intrinsic tyrosine (Tyr 71) was tested for efficacy as a probe of ACP's conformation. Although Tyr 71 was sensitive to its environment and to conformational change in ACP mutants, several experimental issues likely preclude its use as a fluorescent probe. To study ACP-partner enzyme interactions, two enzymes that lack endogenous Trp, *E. coli* UDP-*N*-acetylglucosamine acyltransferase (LpxA) and *V. fischeri* holo-ACP synthase (AcpS) were chosen for Trp-substitution at various positions. Characterization of these LpxAs suggests that Trp-substitution is a good probe for measuring ACP-enzyme interactions and transfer of the acyl chain to a partner enzyme's active site.

LIST OF ABBREVIATIONS USED

A. aeolicus	*Aquifex aeolicus*
A. baumannii	*Acinetobacter baumannii*
AAS	2-Acyl-glycerolphosphatidylethanolamine/acyl-ACP synthetase
ACP	Acyl carrier protein
AcpS	Holo-ACP synthase
AcpH	ACP phosphodiesterase
AEDANS	5-(Acetylaminoethyl)aminonaphthalene-1-sulfonic acid
AinS	Autoinducer synthase
Ala	Alanine
Asp	Aspartate
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
AHL	*N*-acyl-homoserine lactone
AUC	Area under the curve
B. anthracis	*Bacillus anthracis*
B. brevis	*Bacillus brevis*
B. subtilis	*Bacillus subtilis*
B. thailandensis	*Burkholderia thailandensis*
BioH	Pimeloyl-ACP methyl ester esterase
BioI	Biotin synthesis cytochrome P450 enzyme
BioA	7,8-Diaminopelargonic acid synthase
BioB	Biotin synthase
BioC	Malonyl-ACP *O*-methyltransferase
BioD	Dethiobiotin synthetase
BioF	8-Amino-7-oxononanoate synthase
BMP	Bitmap file
BSA	Bovine serum albumin
B. subtilis	*Bacillus subtilis*
B.s.AcpS	*Bacillus subtilis* holo-ACP synthase
C. ammoiniagenes	*Chronobacterium ammoiniagenes*
C. crescentus	*Caulobacter crescentus*
C. jejuni	*Campylobacter jejuni*
C. trachomatis	*Chlamydia trachomatis*
CD	Circular dichroism
CDP	Cytidine diphosphate
CDP-DAG	CDP-diacylglycerol
CdsA	CDP-DAG synthase
CHARMM	Chemistry at HARvard Macromolecular Mechanics molecular simulation program
CIF	Chain initiation factor
CMP	Cytidine monophosphate
CoA	Coenzyme A
CLS	Cardiolipin synthases
CSD	Charge state distribution

CSV	Comma separated value
CTP	Cytidine triphosphate
cycF50A	Cyclic ACP with phenylalanine 50 mutated to alanine
cycL46W	Cyclic ACP with leucine 46 mutated to tryptophan
Cys	Cysteine
Desktop1-PC	Windows® 7 Home Premium 64-bit, Intel® Core™ i7 CPU (2.67 GHz); 12 Gb RAM; ATI™ Radeon™ HD 5600)
Desktop2-PC	Windows® XP Professional 32-bit, Intel® Pentium® 4 CPU (3.20 GHz); 1 Gb RAM; NVIDIA® GeForce® FX 5200)
DCD	NAMD Trajectory information file (single precision binary FORTRAN files)
DCP	D-Alanyl carrier protein
DltA	D-Alanine- D-alanyl carrier ligase
DtlB	D-Alanyl-lipoteichoic acid biosynthesis acyltransferase protein
DltC	D-Alanyl carrier protein
DltD	Poly- D-Alanine transfer protein
DTT	1,4-Dithiothreitol
E. coli	*Escherichia coli*
*E.c.*AcpH	*Escherichia coli* ACP phosphodiesterase
*E.c.*AcpS	*Escherichia coli* holo-ACP synthase
EDTA	Eythylene-diamine tetra-acetic acid (disodium salt used)
EMS	Enhanced MS (Applied Biosystems single MS scanning mode)
EntF	Adenylation domain of *E. coli* non-ribosomal peptide synthetase
ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
FabA	β-hydroxydecanoyl-ACP dehydratase/*trans*-2, *cis*-3-decenoyl-ACP isomerase
FabB	β-ketoacyl-ACP synthase I
FabD	Malonyl-CoA:ACP transacylase
FabE	Biotin carboxyl carrier protein of acetyl-CoA carboxylase
FabF	β-ketoacyl-ACP synthase II
FabG	β-ketoacyl-ACP reductase
FabH	β-ketoacyl-ACP synthase III
FabI	Enoyl-ACP reductase I
FabK	Enoyl-ACP reductase II
FabL	Enoyl-ACP reductase III
FabZ	β-hydroxyacyl-ACP dehydratase
FAS	Fatty acid synthase
Fester-HPC	Centos 6.4 (GNU/Linux) High Performance Cluster consisting of: 51 x Dual Quad core Intel Xeon E5404 @ 2 GHz (8 Logical CPUs) with 8 GB RAM. 52 x Infiniband DDR connected to all compute nodes + headnode (16Gbit/s interconnect) 1 x Dual Quad core Intel Xeon E5405 @ 2Ghz (8 Logical CPUs) Headnode/NFS server
FMN	Flavin mononucleotide (oxidized form)

$FMNH_2$	Flavin mononucleotide (reduced form)
FPLC	Fast Protein Liquid Chromatography
Fx_A	Activated Factor X (serine protease)
G3P	Glycerol-3-phosphate
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
GST	Glutathione S-transferase
H. pylori	*Helicobacter pylori*
H. sapiens	*Homo sapiens*
His	Histidine
HEPES	*N*-(2-hydroxyethyl)piperazine-*N*'-(2-ethanesulfonic acid)
HlyC	Hemolysin-activating lysine-acyltransferase
ID	Inner-diameter
IDE	Intelligent development environment
Ile	Isoleucine
Inspiron1-PC	Dell™ Inspiron™ 6400 computer (Windows® Vista Home Premium 32-bit, Intel® Core™ 2 Duo CPU (1.83 GHz); 2 Gb RAM; ATI Mobility Radeon X1400)
Inspiron2-PC	Dell™ Inspiron™ 1525 computer (Windows® Vista Home Premium 32-bit, Intel® Core™ 2 Duo CPU (1.83 GHz); 3 Gb RAM; Mobile Intel® 965 Express Chipset)
IPTG	Isopropyl β-D-thiogalactopyranoside
IscS	Cysteine desulfurase
I-PCR	Inverse PCR
kDa	Kilodalton
KdtA	3-Deoxy-D-manno-octulosonic-acid transferase
KS	Ketosynthase
L. interrogans	*Leptospira interrogans*
L. lactis	*Lactococcus lactis*
Lac. casei	*Lactobacillus casei*
Lac. rhamnosus	*Lactobacillus rhamnosus*
LB	Miller's Luria-Bertani broth
LB-SA	LB medium containing spectinomycin (50 μg/mL) and ampicillin (100 μg/mL)
LC	Liquid chromatography
LC-MS	LC coupled to mass spectrometry
LC-MS/MS	LC coupled to tandem mass spectrometry
LD	Langevin dynamics
LipA	Lipoic acid synthase
LipB	Lipoyl transferase
linF50A	Linear ACP with phenylalanine 50 mutated to alanine
linL46W	Linear ACP with leucine 46 mutated to tryptophan
LPS	Lipopolysaccharide
LpxA	UDP-*N*-acetylglucosamine acyltransferase
LpxB	Lipid A disaccharide synthase

LpxC	UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase
LpxD	UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine *N*-acyltransferase
LpxH	UDP-2,3-diacylglucosamine hydrolase
LpxK	Tetraacyldisaccharide-1-P 4-kinase
LpxM	Lipid A biosynthesis $(Kdo)_2$-(lauroyl)-$lipid_{IV}$ A acyltransferase
LpxL	Lipid A biosynthesis lauroyl acyltransferase.
LSC	Liquid scintillation counting
LS50B	Perkin Elmer LS-50B
LuxCE	Fatty acid reductase complex
LuxD	Myristoyl-ACP thioesterase
LuxI	Acyl-homoserine-lactone synthase
LuxM	Acyl-homoserine-lactone synthase
Lys	Lysine
M. pneumoniae	*Mycoplasma pneumoniae*
M.smegmatis	*Mycobacterium smegmatis*
M. tuberculosis	*Mycobacterium tuberculosis*
MD	Molecular dynamics
MdoH	Glucans biosynthesis glucosyltransferase H
Me. loti	*Mesorhizobium loti*
Met	Methionine
MES	2-(N-morpholino)ethanesulphonic acid
Morticia-PC	Ubuntu® Linux® 10, Quad-Core AMD Athlon CPU (2.33GHz); 4 GB RAM
MS	Mass spectrometry
MS-DOS	Microsoft Disk Operating System
MSMS	Michael Sanner's Molecular Surface
MukB	Chromosome partition protein
MW	Molecular weight
NAMD	Not (just) Another Molecular Dynamics program
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NMR	Nuclear magnetic resonance
NodF	*Rhizobium leguminosarum* nodulation protein
NRPS	Non-ribosomal peptide synthetase
NUP	Natively unfolded protein
P. aeruginosa	*Pseudomonas aeruginosa*
*P.a.*LpxA	*Pseudomonas aeruginosa* UDP-*N*-acetylglucosamine acyltransferase
PA	Phosphatidic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank file (information on format available

	(RCSB. *Protein Data Bank: PDB File Format Information*.))
P. falciparum	*Plasmodium falciparum*
*P.f.*ACP	*Plasmodium falciparum* acyl carrier protein
P.f.$AcpS_{FL}$	*Plasmodium falciparum* holo-ACP synthase (full length)
*P.f.*AcpS-C	C-terminal fragment of the full length *Plasmodium falciparum* holo-ACP synthase
*P.f.*AcpS-N	N-terminal fragment of the full length *Plasmodium falciparum* holo-ACP synthase
PGP	Phosphatidylglycerolphosphate phosphatase
PgsA	Phosphatidylglycerol synthase
PKS	Polyketide synthase
PlsB	*sn*-Glycerol-3-phosphate acyltransferase
PlsC	1-Acyl-*sn*-glycerol-3-phosphate acyltransferase
PlsX	Phosphate acyltransferase
PlsY	Glycerol-3-phosphate acyltransferase
POV	POV-Ray file
PP	4'- Phosphopantetheine
PPTase	Phosphopantetheinyl transferase
preL46W^{mut}	L46W with a mutated (inactivated) split-intein
PS	Phosphatidylserine
PsD	Phosphatidylserine decarboxylase
PSF	Protein structure file (in X-PLOR format)
PssA	Phosphatidylserine synthase
P. yoelii	*Plasmodium yoelii*
QM4CW	Photon Technology International (PTI) QuantaMaster-4CW
Qtrap	Applied Biosystems/Sciex hybrid triple quadrupole linear ion trap
rACP	Recombinant *Vibrio harveyi* ACP
R. glutinis	*Rhodotorula glutinis*
R. leguminosarum	*Rhizobium leguminosarum*
Ra. norvegicus	*Rattus norvegicus*
RCSB	Research Collaboratory for Structural Bioinformatics
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
S. aureus	*Staphylococcus aureus*
S. avermitilis	*Streptomyces avermitilis*
SA	Divalent cation-binding site A mutant ACP (D30N/D35N/D38N)
SAM	*S*-Adenosyl-L-methionine
SA/SB	Divalent cation-binding site A and B mutant ACP (D30N/D35N/D38N/E47Q/D51N/E53Q/D56N)
SB	Divalent cation-binding site B mutant ACP (E47Q/D51N/E53Q/D56N)
S. cerevisiae	*Saccharomyces cerevisiae*
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	Serine

SpoT	Bifunctional (p)ppGpp synthase/hydrolase
S. pneumoniae	*Streptococcus pneumoniae*
Sp. oleracea	*Spinacia oleracea*
SrfA	Surfactin synthase subunit A
St. coelicolor	*Streptomyces coelicolor*
St. rimosus	*Streptomyces rimosus*
St. roseofulvus	*Streptomyces roseofulvus*
T_M	Melting temperature
TCA	Trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
TycC	Tyrocidine synthetase
Tyr	Tyrosine
UCSF	University of California: San Francisco
UCSF Chimera	UCSF Chimera: an Extensible Molecular Modeling System
UDP	Uridine diphosphate
UDP-Glc*N*Ac	Uridine diphosphate *N*-acetylglucosamine
UMP	Uridine monophosphate
Urea-PAGE	Polyacrylamide gel electrophoresis in the presence of urea
Val	Valine
VDW	Van der Waals
VibF	Non-ribosomal protein synthetase necessary for the biosynthesis of vibriobactin
VMD	Visual Molecular Dynamics
V. cholerae	*Vibrio cholerae*
V. fischeri	*Vibrio fischeri*
*V.f.*AcpS	*Vibrio fischeri* holo-ACP synthase
V. harveyi	*Vibrio harveyi*
Val	Valine
XIC	Extracted ion chromatogram
X-PLOR	Program for running molecular simulations; Similar to CHARMM
XPS-PC	Dell™ XPS™ L702X computer (Windows® 7 Home Premium 64-Bit, Intel® Core™ i7-2720QM CPU (2.20 GHz); 8 Gb RAM; NVIDIA® GeForce® GT 555M)
YbgC	Acyl-CoA thioester hydrolase

ACKNOWLEDGEMENTS

Many people have been instrumental in the training and guidance that have helped me get to where I am today. First and most importantly, my supervisor, Dr. David Byers, whose training and mentorship began during my honours project so many years ago. During this time I was given access and trained on a brand new (and quite expensive!) mass spectrometer; an experience I will forever value. Over the years he has supported me in my research, my collaborations, and my extracurricular involvement within the department and student societies. He has been essential in shaping me into the critically thinking scientist I have become.

My supervisory committee (Drs. Bearne, McMaster, and Ewart) has provided invaluable help, guidance, understanding and support over the years, for which I am grateful. Additionally I would like to thank Dr. Archibald and Dr. Ewart for their unwavering support during difficult times.

For their invaluable training, support and technical assistance I would like to thank Anne Murphy (former Lab Technician, Byers Lab) and Elden Rowland (former Lab Manager for the Faculty of Medicine Proteomics Core Facility, Dalhousie University). Anne has been an instrumental part of my training throughout my graduate degree. Elden trained me during my honours project and helped during my graduate degree.

Various faculty members have provided invaluable training, access to equipment and technical assistance over the years: Dr. Blouin for getting me interested in molecular dynamics and modeling, teaching me the basics of Python programming language and providing invaluable help troubleshooting, and graciously giving me access to high

performance computers to perform molecular dynamic simulations, Dr. Bearne for providing me access to his equipment (spectropolarimeter and isothermal titration calorimeter) and for making me feel welcome in his lab, and Dr. McMaster for his invaluable help troubleshooting PCR early in my degree. Lastly, the late Dr. Warren DeLano (creator of PyMol), for his enthusiastic help in troubleshooting problems creating a cyclic ACP *in silico*.

I would also like to thank my fellow graduate students for their help and collaboration. Specifically, Gerrit Volkmann (former Ph.D. student, Liu Lab, Dalhousie University) for the enjoyable and informative collaboration on cyclic ACP and his help in producing mutant AcpSs, David Chan (former Ph.D. student, Vogel Lab, University of Calgary) for his help with creation of cyclic ACP sausage diagrams, NMR work with cyclic ACP and general molecular dynamics discussions, Marissa Le Blanc (former Ph.D. student, McMaster Lab, Dalhousie University) for creating the *Plasmodium falciparum* Trp-substituted ACPs, and Javier Alfaro (former Ph.D. student Roger Lab, Dalhousie University) for his help with Python programming.

I would like to thank past members of the Byers lab for their help, and support. Additionally, the past and present members of the Atlantic Research Center have been an invaluable source of support, technical assistance, and stimulating conversation; it has been a great lab to train in due to the amicable environment, and people's willingness to provide assistance. Specifically I would like to thank Robert Zwicker: he was an invaluable member of the ARC, providing technical assistance, equipment repair, and stimulating conversation.

Last, but not least, I would like to thank my parents: Joanne and Bill. They have always told me I can do anything I put my mind to; this thesis is a testament to that. Their unwavering support over the years means more to me than they know.

CHAPTER 1 INTRODUCTION

1.1 OVERVIEW

Fatty acids, the key hydrophobic component of phospholipids, are synthesized by fatty acid synthase (FAS). To function, FAS complexes require a carrier of activated fatty acyl groups in a thioester linkage: an acyl carrier protein (ACP). ACP and ACP-like proteins exist in many forms (discussed further in Section 1.2), but are highly conserved across species from each main phylogenetic group: archaea, bacteria and eukaryotes (discussed in Section 1.3.1). Thus, all belong to the "ACP Family". These carrier proteins are indispensable cofactors in various metabolic pathways, including the biosynthesis of fatty acids, phospholipids, endotoxin, glycolipids, signalling molecules, polyketides and nonribosomal peptides (Byers & Gong, 2007).

The founding member of the "ACP Family", *E. coli* ACP, was first characterized as an intermediate in fatty acid biosynthesis in a historic series of papers by Vagelos and coworkers (see tribute paper by Kresge *et al*. (2005)). This ACP, like ACP from all bacterial species, is a small (77 amino acids; ~9 kDa), acidic (pI = 4.1) (Vanaman et al., 1968) protein that interacts with a multitude of enzyme partners in bacteria (Butland et al., 2005). ACP has been shown to be a highly conformationally dynamic protein (discussed in Section 1.3.2). In addition to the prototypical *E. coli* ACP, ACPs have been found in a plethora of other bacterial species, including *V. harveyi* (discussed in Section 1.3.3).

Due to its central role in many pathways essential to bacterial growth and pathogenesis and the fundamental differences between bacterial and human ACP (discussed in Sections 1.2.1.1 and 1.2.1.2), ACP-dependent enzymes are attractive

antimicrobial drug targets. Two of these targets (LpxA and AcpS) are discussed more extensively below (see Sections 1.4.2 and 1.4.3, respectively). Additionally, its central role in FAS makes characterization of ACP and ACP-dependent enzymes crucial for the directed production of food oils and biofuels. In plant chloroplasts, FAS (of bacterial origin) functions in much the same manner as its bacterial counterpart (discussed in Section 1.2.1.2), producing C16 and C18 saturated fatty acids that are unsaturated by Δ9-acyl-ACP desaturase (Schmid & Ohlrogge, 2002). Indeed, it has been recognized that the fatty acid composition, nutritional quality and quantity of seed oils can be manipulated by overexpression or alteration of enzymes that participate in FAS, including ACP (Jha et al., 2007). Biodiesel, a typical hydrocarbon biofuel, contains fatty acid methyl esters and fatty acid monoalkyl esters (Lee et al., 2013). Through similar manipulation of FAS genes, *E. coli* strains can be engineered to produce the long-chain and very-long-chain fatty acids required to produce biofuels such as biodiesel (Lee et al., 2013). Non-ribosomal peptide synthases (NRPSs) and polykeytide synthases (PKSs) are involved in the production of natural product pharmaceuticals and agro-chemicals (Beld et al., 2014). Thus, much research has gone into production of these products in heterologous systems (such as *E. coli*) that are easy to culture and maintain in a laboratory or industrial settings by engineering these systems to produce non-natural, directed, products (Khosla & Keasling, 2003).

1.2 Classification of ACP and Carrier Proteins

A broad family of acyl carrier proteins exists and consists of ACP and ACP-like proteins; known as the "ACP Family". ACPs are classified into two types: Type I, in which ACP is a domain of a larger multifunctional fatty acid synthesis (FAS) polypeptide

(discussed further in Section 1.2.1.1), and Type II, in which ACP functions as a discrete protein in FAS synthesis (discussed further in Section 1.2.1.2 and sub-sections). ACP-like proteins include: D-alanyl carrier proteins (DCP), polyketide synthase (PKS) ACPs, and non-ribosomal peptide synthesis (NRPS) peptidyl carrier proteins (PCP). These ACP-like proteins are discussed further in Section 1.2.1.3. Irrespective of the type, all carrier proteins must be activated with a phosphopantetheine moiety by a phosphopantetheinyl transferase (PPTase) (Byers & Gong, 2007). Functional diversification of ACP and ACP-like proteins can be exceptionally important to some bacteria, as demonstrated in *S. avermitilis*, which contains at least 70 FAS, PKS and NRPS systems, corresponding to more than 85 carrier proteins (Lai et al., 2006).

1.2.1 Acyl Carrier Proteins

ACP plays a central role in FAS in both eukaryotes and bacteria. However, the fundamental design of FAS in these two major cell types differs. Eukaryotic FAS is designated Type I and prokaryotic FAS, Type II. Most bacteria contain only Type II FAS, with the exception of species such as *M. tuberculosis*, which contain Type I FAS as well (Brennan & Nikaido, 1995). Conversely, eukaryotic cells have both Type I and II FAS, with Type I occurring in the cytosol and Type II confined to either mitochondria or plastids (Chan & Vogel, 2010).

1.2.1.1 ACP in Eukaryotes (Type I)

Type I FAS consists of one or more large polypeptides with multiple catalytic centers involved in the cyclic condensation of 2-carbon acyl units from malonyl-CoA (Smith et al., 2003). The ACP domain of Type I FAS is on a flexible linker that, in

addition to its own flexibility, gives it access to all active sites in the large polypeptide structure, effectively acting as a swinging arm delivering the covalently attached thioester acyl intermediates to the appropriate active site. Despite the similarities in structure of Type I FAS across species, differences in the quaternary structure of the polypeptide and destination of the end products are present, for example in mammals *versus* yeast.

Mammalian FAS exists as a single homodimeric protein (Smith et al., 2003) with each subunit containing a fully functional unit: ACP and the functional enzyme domains required to carry out the cyclic growth of fatty acids by 2 carbons per cycle. The thioesterase domain of mammalian Type I FAS catalyzes the release of fatty acids (typically palmitate (16:0)), which can then be activated by acyl-CoA synthetase and further elongated by microsomal systems (Cook & McMaster, 2002). Originally, mammalian FAS was proposed to exist in a head-to-tail extended antiparallel configuration (Wakil, 1989), but this has been recently revised with the determination of the crystal structure of porcine type I FAS. This FAS structure suggests two non-identical lateral semicircular reaction chambers (Maier et al., 2006).

In contrast, *S. cerevisiae* type I FAS consists of six copies of two non-identical subunits ($\alpha_6\beta_6$). The α subunit of each copy contains the ACP domain, each of which have access to one of six reaction chambers. These reaction chambers are organized in a hollow globe separated by an equatorial wheel (Leibundgut et al., 2007, Lomakin et al., 2007). As with mammalian Type I FAS, the end product of yeast Type I FAS is a 16C fatty acid; however, unlike the mammalian system, the palmitate is transferred back to CoA from the ACP domain by a malonyl-palmitoyl transacylase.

As mentioned briefly above (Section 1.2.1) some bacterial species, such as *M. tuberculosis*, contain Type I FAS in addition to the ubiquitous Type II system (Brennan & Nikaido, 1995). In *M. tuberculosis*, Type I FAS produces C12–C16 fatty acid products that are further elongated by the Type II FAS system to make very long chain fatty acids (C50–C60), eventually leading to the production of mycolic acids, the predominant and characteristic lipid components of the cell envelope (Schweizer & Hofmann, 2004).

1.2.1.2 ACP in Bacteria (Type II)

Unlike Type I ACP in eukaryotic systems, bacterial Type II ACPs function as a discrete monomeric protein (White et al., 2005, Chan & Vogel, 2010). Type II ACPs are small (~9 kDa), acidic and highly conserved across species as they are essential for bacterial growth and pathogenesis (Byers & Gong, 2007). It has been estimated that ACP makes up ~0.25% of total soluble protein (Rock & Cronan, 1979) or is present at concentrations of about 10 μM (De Lay & Cronan, 2007); however, most exists in the holo as opposed to the acylated form (Rock & Jackowski, 1982). Based on proteomic analysis, ACP has been discovered to have >30 interacting partners in *E. coli* (Butland et al., 2005), which is not surprising given its central role in many bacterial processes (Figure 1). However, its main function is in the synthesis of fatty acids and phospholipids.

1.2.1.2.1 Function in FAS II and Phospholipid Synthesis

As with its ACP, Type II FAS is comprised of discrete proteins, each catalyzing one step of the synthesis process (Rock & Cronan, 1996). As a result, synthetic intermediates remain covalently attached to ACP and must be sequestered in the core of

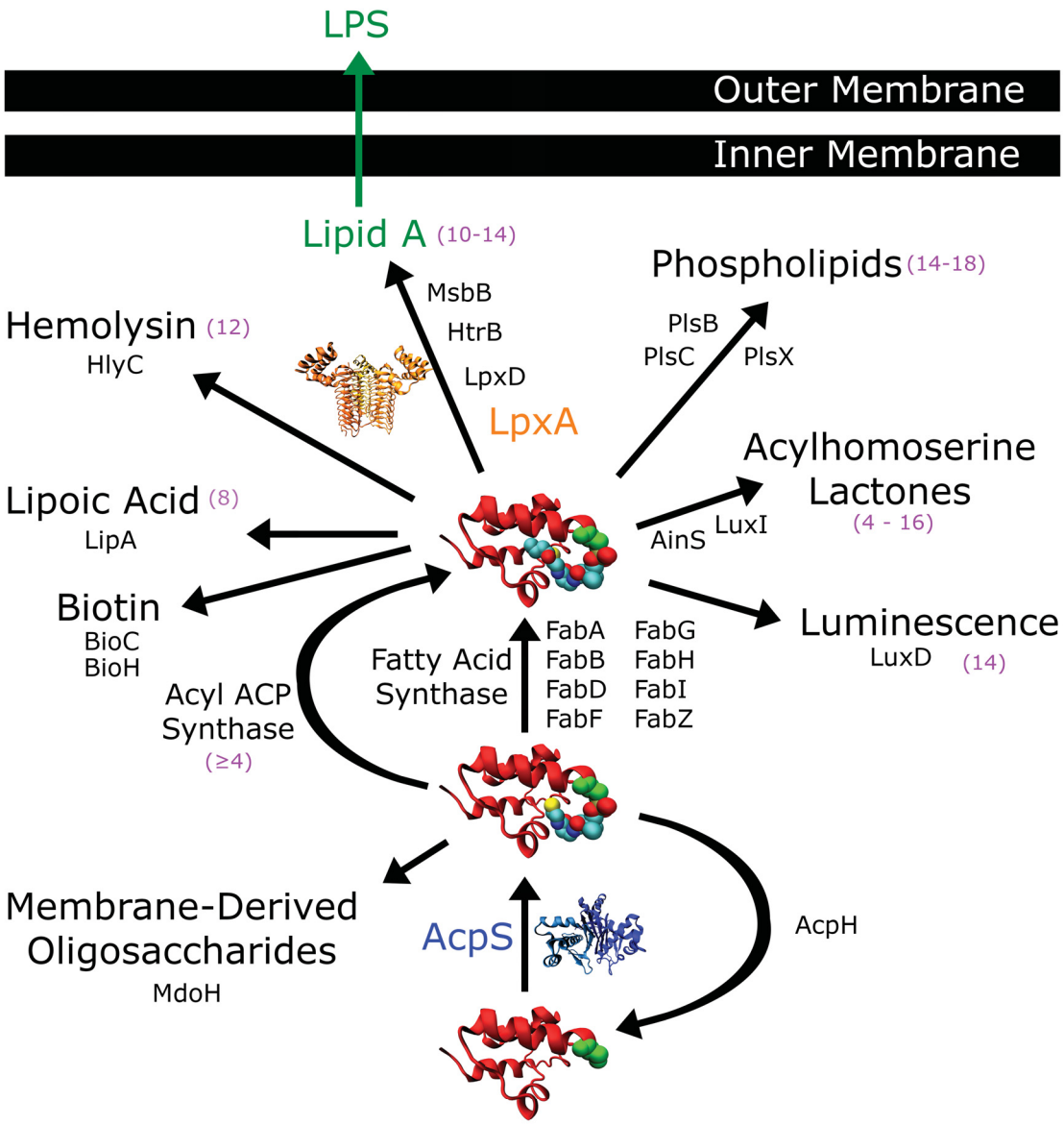

Figure 1 ACP-dependent pathways in bacteria. ACP-dependent enzymes are shown with enzymes of interest highlighted: AcpS (blue), LpxA (orange). ACP (red) is depicted in its various forms: apo (unmodified Ser (green)), holo (phosphopantetheine moiety (PP) attached to Ser) and acyl (acyl chain attached to PP; butryl-ACP is depicted). General acyl chain specificity of pathway enzymes is indicated (purple). ACP, LpxA, and AcpS are displayed using "NewCartoon" representation. The "active" serine is displayed using "VDW" representation. ACP is displayed using "NewCartoon" graphical representation. Three structures were used to create this figure: *E. coli* LpxA bound to UDP-Glc*N*Ac (PDB ID = 2JF3 (Ulaganathan et al., 2007)), *B. subtilis* AcpS (PDB ID = 1F7T (Parris et al., 2000)), and *E. coli* butyryl-ACP (PDB ID = 1L0I (Roujeinikova et al., 2002)). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

the protein between enzyme reactions. This is in stark contrast to ACP in Type I FAS (discussed in detail above), in which ACP acts as a flexible arm, allowing the attached intermediates to be shuttled from one active site to the next in a large polypeptide chain. The *E. coli* Type II FAS system has been extensively characterized and recent work has determined the three-dimensional structure of each component of this prototypical system (reviewed by White *et al*. (2005)); thus, this system will be described.

For ACP to function as an acyl carrier, it first must be modified via a reaction catalyzed by holo-ACP synthase (AcpS; discussed in detail in Section 1.4.3). This reaction adds a phosphopantetheine moiety (PP), which has a nucleophilic thiol group at the distal end. The reverse of this reaction is catalyzed by a divalent cation-dependent ACP phosphodiesterase (AcpH) producing apo-ACP and free PP (Vagelos & Larrabes, 1967); however, it is not essential for growth (Thomas & Cronan, 2005). AcpH is responsible for regeneration of the PP pool *in vivo*, and has been shown to be active with not only holo-ACP, but also ACPs carrying up to 16 carbon acyl chains (Vagelos & Larrabes, 1967, Thomas & Cronan, 2005). *E. coli* AcpH (*E.c.*AcpH) has been shown to be active with FAS ACPs from other species, but is not active with peptide fragments of ACP (Vagelos & Larrabes, 1967), suggesting a full length, folded ACP is required for AcpH activity. Additionally, *E.c.*AcpH is inactive with *R. leguminosarum* nodulation protein NodF (Geiger et al., 1991) and the mitochondrial ACP from *Bos taurus* (Thomas & Cronan, 2005) despite structural homology with bacterial ACPs, suggesting these Type II ACP-like proteins lack important residues for binding or activity.

The first committed step of Type II FAS is a biotin-dependent carboxylation of acetyl-CoA, catalyzed by acetyl-CoA carboxylase (ACC) (Cronan & Waldrop, 2002) resulting in malonyl-CoA (Figure 2). ACC is a heterotetrameric enzyme consisting of AccA (carboxytransferase), AccB (biotin carboxy carrier protein), AccC (biotin carboxylase) and AccD (carboxytransferase) (Rock & Jackowski, 2002). Malonate is then transferred from malonyl-CoA to ACP by malonyl-CoA:ACP transacylase (FabD). This malonyl-ACP then starts a cycle of fatty acid elongation through the β-ketoacyl-ACP synthase III (FabH) catalyzed condensation of acetyl-CoA with malonyl-ACP. FAS is the only known destination for malonyl-CoA in *E. coli* (Cronan & Waldrop, 2002). Following condensation, a reduction reaction is carried out by β-ketoacyl-ACP reductase (Fab G). *Trans*-2-enol-ACP is then formed by either β-hydroxydecanoyl-ACP dehydratase/*trans*-2, *cis*-3-decenoyl-ACP isomerase (FabA) or β-hydroxyacyl-ACP dehydratase (FabZ). The additional *trans*-2, *cis*-3-decenoyl-ACP isomerase function of FabA allows for isomerization of 10 carbon unsaturated fatty acids. The final stage of the first cycle of FAS is catalyzed by enoyl-ACP reductase I, II and III (FabI, FabK, and FabL, respectively). All subsequent rounds of FAS elongation are initiated by either β-ketoacyl-ACP synthase I (FabB) or II (FabF), in which the growing acyl chain is transferred to malonyl-ACP, and the cycle continues with FabG (as above). Substrate specificities of Fab B and Fab F ultimately determine the structure and distribution of fatty acid products (Vagelos & Larrabes, 1967). An illustration of the steps described above can be found in Figure 2. Type II FAS systems also exist in eukaryotic mitochondria and plant plastids, and are thought to be of bacterial origin (Jordan & Cronan, 1997, Miller et al., 2000, Schmid & Ohlrogge, 2002). For more on these systems,

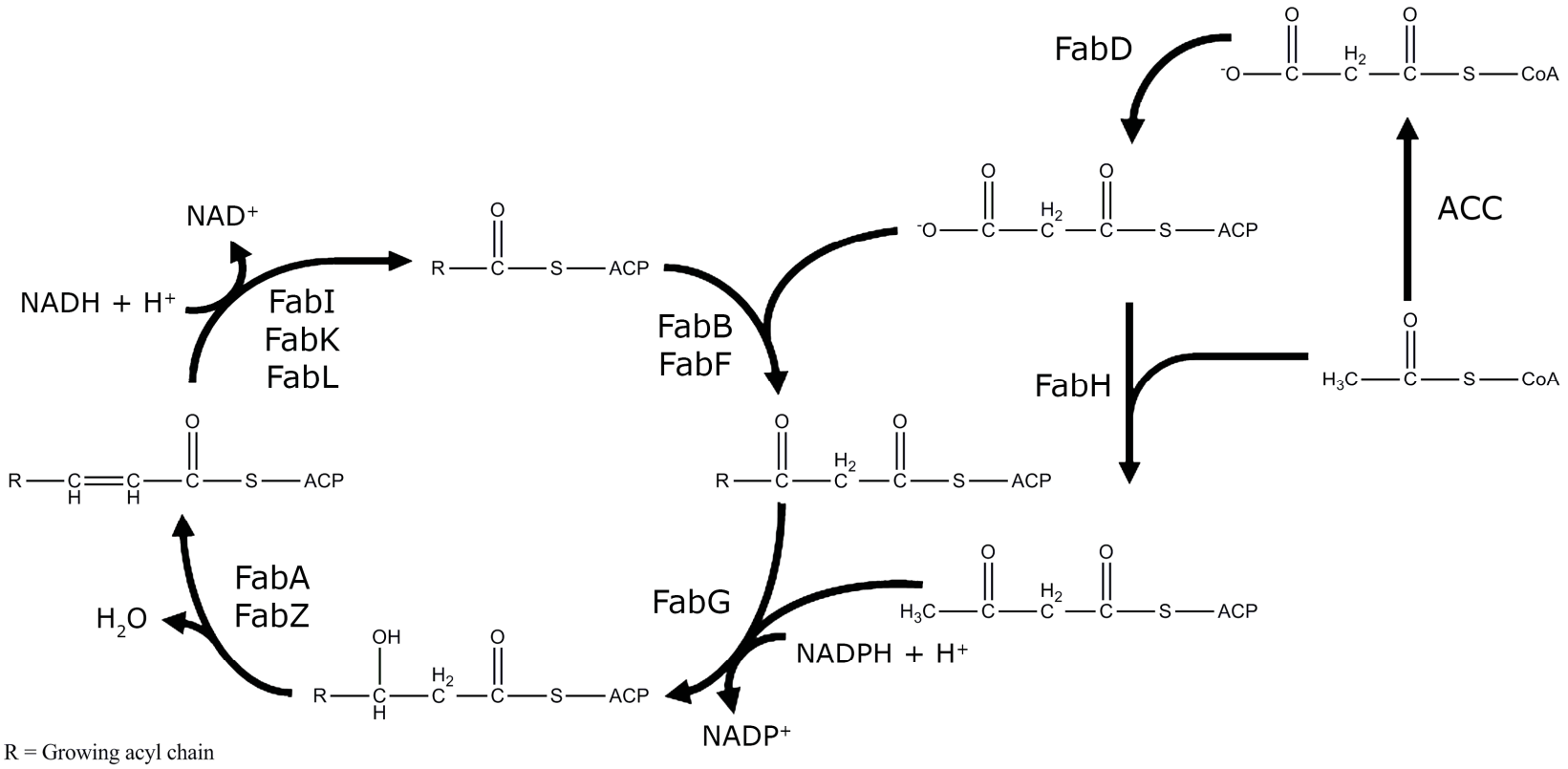

Figure 2 Fatty acid biosynthesis in bacteria. Enzyme abbreviations are as follows: ACC, acetyl-CoA carboxylase; FabA, β-hydroxydecanoyl-ACP dehydratase/*trans*-2, *cis*-3-decenoyl-ACP isomerase; FabB, β-ketoacyl-ACP synthase I; FabD, malonyl-CoA:ACP transacylase; FabF, β-ketoacyl-ACP synthase II; FabG, β-ketoacyl-ACP reductase; FabH, β-ketoacyl-ACP synthase III; FabI/FabK/FabL, enoyl-ACP reductase I/II/III, respectively; FabZ, β-hydroxyacyl-ACP dehydratase. Figure created based on Rock and Jackowski (2002).

the reader is referred to the detailed review by Chan and Vogel (2010).

While the acyl-ACPs have many potential uses (discussed below), the majority of acyl-ACP end products from FAS are used in the formation of membrane phospholipids. In *E. coli*, phospholipid biosynthesis consists of the transfer for two acyl chains from acyl-ACPs to a glycerol-3-phosphate (G3P) backbone.

Acylation of the sn-1 position of G3P can be catalyzed by two proteins. First, *sn*-glycerol-3-phosphate acyltransferase (PlsB (Green et al., 1981)), an inner membrane protein (Bayan & Thérisod, 1989), will transfer an acyl chain to the sn-1 position of G3P using acyl-ACP as a donor. PlsB is essential for bacterial growth in *E. coli*; however, homologues are not widely distributed in bacteria (Heath & Rock, 1999). The second protein that can perform this acyltransferase activity, PlsY, is more widely distributed, but uses acyl-phosphate as the donor (Lu et al., 2006, Lu et al., 2007). Acyl-phosphate is generated by the enzyme PlsX (Lu et al., 2006), which shares structural homology with acylglycerol-3-phosphate acyltransferase (PlsC) (Paoletti et al., 2007). However, unlike PlsB, PlsY and PlsX are not essential in *E. coli* (Yoshimura et al., 2007).

The product of the first transacylation reactions, lysophosphatidic acid, is successively acylated at the sn-2 position by 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (PlsC (Coleman, 1990)), resulting in phosphatidic acid (PA). In *E. coli*, PlsC primarily uses acyl-ACP as the acyl donor but is also able to use acyl-CoAs. In contrast, Gram-positive bacteria such as *B. subtilis* (Paoletti et al., 2007) or *S. pneumoniae* (Lu et al., 2006) use only acyl-ACPs. Recently, a thioesterase (acyl-CoA thioester hydrolase; YbgC) has been discovered and proposed to be involved in phospholipid biosynthesis due to its interactions with PlsB and a phosphatidylserine

synthase (PssA, discussed below), although its exact role has yet to be elucidated. Following synthesis through either the PlsB/PlsC or PlsX/PlsY/PlsC pathways (Dowhan, 2013), PA is converted to CDP-diacylglycerol (CDP-DAG) by CDP-DAG synthase (CdsA). This CDP-DAG, in turn, is converted to phosphatidylserine (PS) by phosphatidylserine synthase (PssA). PS can then be used to form phosphatidylethanolamine through a reaction catalyzed by PsD (Gully & Bouveret, 2006, Dowhan, 2013). Alternatively, CDP-DAG can react with *sn*-glycerol-3-phosphate to form phosphatidylglycerolphosphate which eventually becomes cardiolipin (Dowhan, 2013). An illustration of the pathways described above can be found in Figure 3. For a more detailed description of phospholipid synthesis, the reader is directed to reviews of phosphatidic acid biosynthesis in bacteria (Yao & Rock, 2013) and the use of *E. coli* to study phospholipid synthesis (Dowhan, 2013).

1.2.1.2.2 Other Functions of Bacterial ACP

Beyond its role in fatty acid and phospholipid synthesis, ACP is involved in a number of other essential and non-essential pathways, including the synthesis of lipid A, biotin and lipoic acid, bioluminescence, quorum sensing, protein acylation, and phospholipid reacylation (Table 1).

In addition to the aforementioned roles of ACP in *de novo* phospholipid synthesis, it functions in phosphatidylethanolamine (PE) remodeling, where it regenerates PE through acylation of 2-acyl-glycerolphosphatidylethanolamine formed by transacylation or phospholipase A1 reactions (Rock, 1984). In *E. coli*, the reaction is catalyzed by the membrane associated 2-acyl-glycerolphosphatidylethanolamine/acyl-ACP synthetase

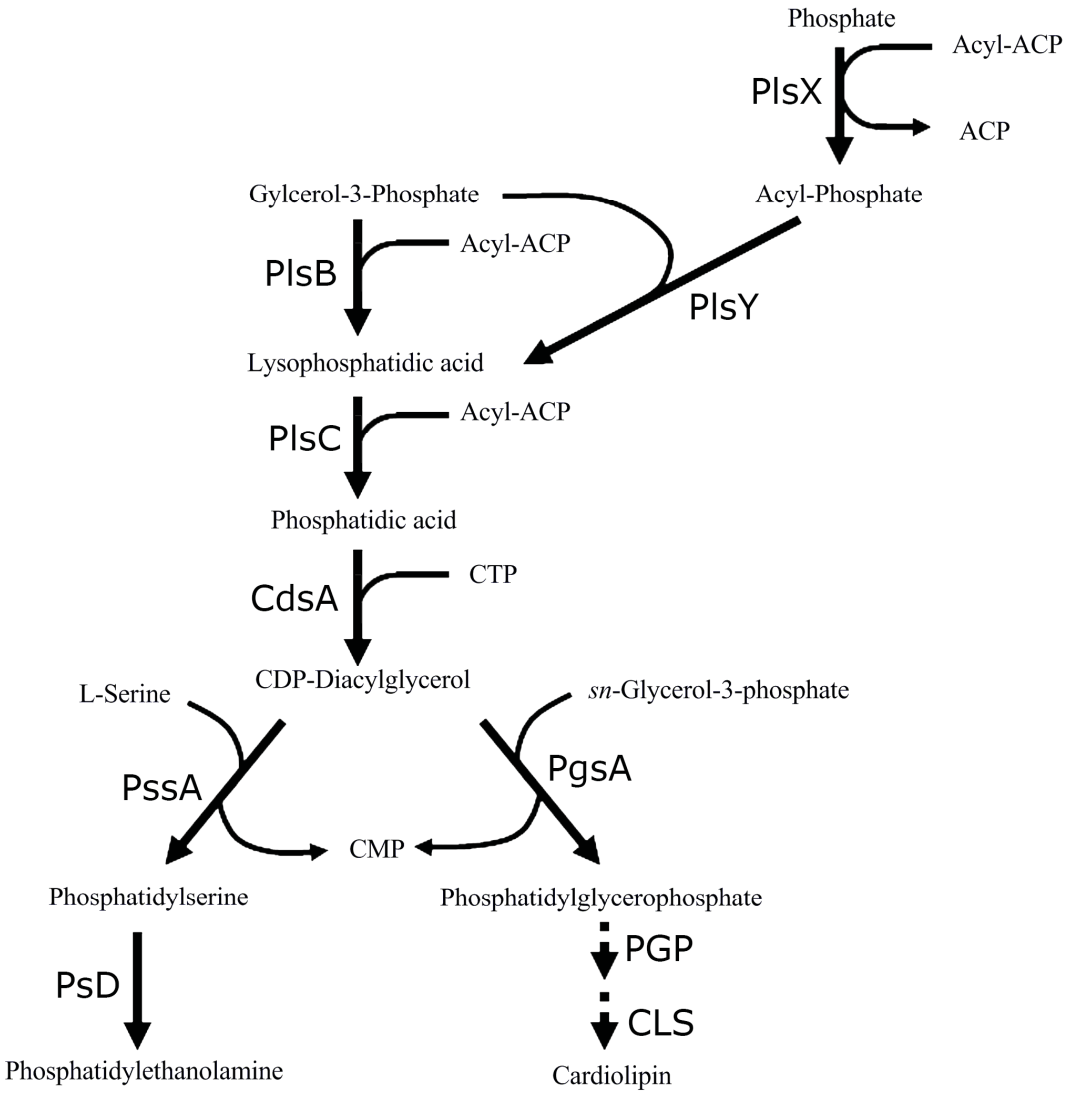

Figure 3 Phospholipid synthesis in bacteria. Abbreviations are as follows: PlsB, *sn*-glycerol-3-phosphate acyltransferase; PlsC, 1-acyl-*sn*-glycerol-3-phosphate acyltransferase; PlsX, phosphate acyltransferase; PlsY, glycerol-3-phosphate acyltransferase; CTP, cytidine triphosphate; CDP, cytidine diphosphate; CMP, cytidine monophosphate; CdsA, phosphatidate cytidylyltransferase; PgsA, phosphatidylglycerol synthase; PssA, phosphatidylserine synthase; PsD, phosphatidylserine decarboxylase; PGP, phosphatidylglycerolphosphate phosphatase; CLS, cardiolipin synthases. This figure was created based on figures from previous work (Gully & Bouveret, 2006, Dowhan, 2013, Yao & Rock, 2013).

Table 1 ACP partners in Gram-negative bacteria. This is not a complete list: additional interactions have been identified by proteomic approaches (Butland et al., 2005). All enzymes listed are *E. coli* except where noted.

Protein(s)	Acyl-ACP[1]	Function	Reference
AcpS	N/A	Synthesis of holo-ACP from apo-ACP through addition of CoA-derived phosphopantetheine	(Lambalot et al., 1996)
AcpH	N/A	Removal of phosphopantetheine moiety from holo-ACP, producing apo-ACP	(Thomas et al., 2007)
FabA, FabB, FabD, FabE, FabF, FabG, FabH, FabI, FabK, FabL, FabZ	2 – 18	*De novo* fatty acid synthesis	(White et al., 2005)
PlsB, PlsC, PlsX	16 – 18	Phospholipid acylation	(Rock & Jackowski, 1982, Lu et al., 2006)
HlyC	14 – 18	Protein acylation (hemolysin)	(Issartel et al., 1991)
LipB	8	Lipoic acid biosynthesis	(Jordan & Cronan, 1997)
LpxA, LpxD, LpxL, LpxM	10 – 14	Lipid A biosynthesis	(Raetz et al., 2007)
LuxI, AinS[2]	4 – 16	Acylhomoserine lactone synthesis; involved in quorum sensing	(Fuqua & Greenberg, 2002)
LuxD	14	Acyl-ACP esterase; involved in bioluminscence	(Byers & Meighen, 1985)
Aas	12 – 16	Acyl-ACP ligation/phospholipid reacylation	(Hsu et al., 1991)
AasS[3]	6 – 14	Acyl-ACP ligation	(Fice et al., 1993, Jiang et al., 2006)
MdoH	N/A	Membrane-derived oligosaccharide synthesis	(Therisod & Kennedy, 1987)
SpoT	Unknown	(p)ppGpp synthesis and hydrolysis	(Battesti & Bouveret, 2006)
IscS	Unknown	Cysteine desulfurase	(Gully et al., 2003)
MukB	Unknown	Chromosome segregation	(Gully et al., 2003)
PssA, YbgC	Unknown	Phospholipid synthesis complex	(Gully & Bouveret, 2006)

Notes: Abbreviations used: N/A, not applicable; (p)ppGpp, guanosine 5′-(tri)diphosphate-3′-diphosphate.
[1] = Prefered acyl chain length [2] = *Vibrio fischeri* [3] = *Vibrio harveyi*

(AAS; encoded by the *aas* gene) and is responsible for the acyl-CoA-independent incorporation of exogenous fatty acids into phospholipids (Hsu et al., 1989). Acyl-ACP intermediates of this reaction remain enzyme associated and thus are not mixed with the *de novo* fatty acid/phospholipid synthesis acyl-ACP pools (Cooper et al., 1989). Soluble forms of AAS are found in other bacterial species, including *R. glutinis* (Gangar et al., 2001) and *V. harveyi* (Byers, 1988, Byers, 1989, Jiang et al., 2006). In contrast to the *E. coli* enzyme, *V. harveyi* AAS (encoded by the *aasS* gene) is capable of activating exogenous fatty acids for both β-oxidation and phospholipid synthesis (Byers, 1989, Shen & Byers, 1994, Jiang et al., 2006) and is able to act on C4 – C16 fatty acids (Fice et al., 1993). Furthermore, AAS in bioluminescent bacteria has been implicated in the ATP-dependent activation of myristate that is then converted to the myristyl aldehyde substrate of luciferase (Meighen, 1988). ACP's role in bioluminescent bacteria is discussed in more detail below.

Lipid A, more commonly known as endotoxin, is a unique glucosamine-based phospholipid comprising the outer membrane of Gram-negative bacteria (Raetz, 1993). In addition to being essential for growth (Galloway & Raetz, 1990, Onishi et al., 1996) and membrane integrity (Vaara, 1993), lipid A potently activates animal innate immunity (Ulevitch & Tobias, 1995) leading to septic shock (Rietschel et al., 1993, David, 2001). The *E. coli* synthesis pathway has been extensively characterized by Christian Raetz and his colleagues and therefore will be the focus of the following paragraphs; however, most Gram-negative bacteria synthesize Kdo_2-lipid A molecules resembling those made by *E. coli* (Raetz et al., 2007) (Figure 4).

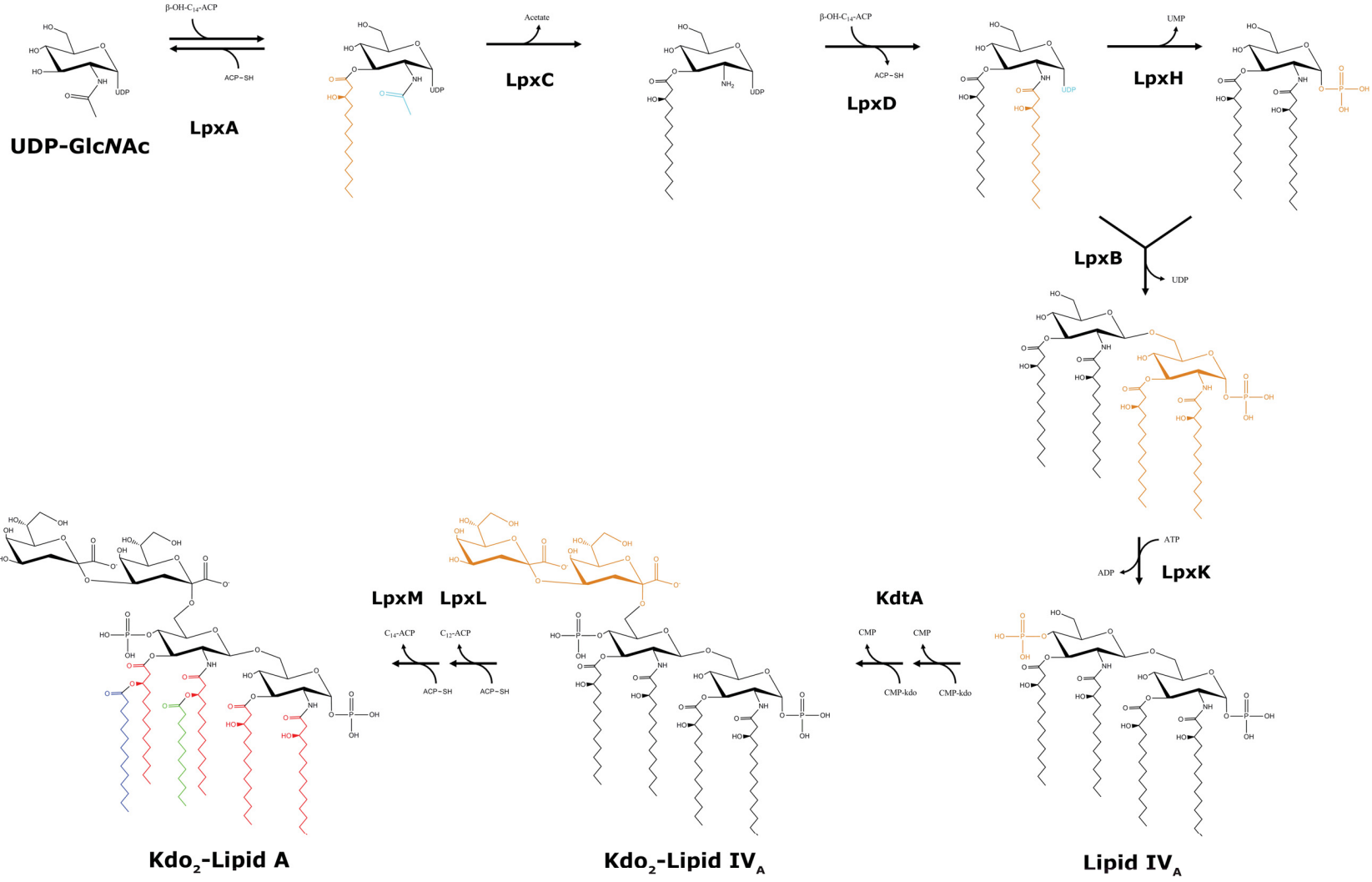
UDP-GlcNAc
β-OH-C14-ACP
ACP-SH
LpxA
Acetate
LpxC
β-OH-C14-ACP
ACP-SH
LpxD
UMP
LpxH
LpxB
UDP
ATP
ADP
LpxK
KdtA
CMP
CMP
CMP-kdo
CMP-kdo
LpxM
LpxL
C14-ACP
C12-ACP
ACP-SH
ACP-SH
Kdo2-Lipid A
Kdo2-Lipid IVA
Lipid IVA

Figure 4 Lipid A biosynthesis in *E. coli*. Coloring of final product (Kdo_2-Lipid A) is indicative of acyl-chain identities: β-OH-myristoyl (red), myristoyl (blue), and lauroyl (green). Coloring in all other steps indicates: moiety added during the previous reaction (orange) and moiety to be removed in following reaction (cyan). Enzyme abbreviations are as follows: LpxA, UDP-*N*-aceytlglucosamine acyltransferase; LpxB, lipid A disaccharide synthase; LpxC, UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase; LpxD, UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine *N*-acyltransferase; LpxH, UDP-2,3-diacylglucosamine hydrolase; LpxK, Tetraacyldisaccharide-1-P 4-kinase; KdtA, 3-deoxy-D-manno-octulosonic-acid transferase; LpxM, lipid A biosynthesis $(Kdo)_2$-(lauroyl)-$lipid_{IV}$ A acyltransferase; LpxL, lipid A biosynthesis lauroyl acyltransferase. Other abbreviation definitions: ACP, acyl carrier protein; ATP, adenosine triphosphate; ADP, adenosine diphosphate; CMP, cytidine monophosphate; UDP, uridine diphosphate; UDP-Glc*N*Ac, uridine diphosphate *N*-acetylglucosamine; UMP, uridine monophosphate. This figure was created based on Sweet *et al*. (2001) with updated enzyme abbreviations.

The first step in lipid A biosynthesis, catalyzed by LpxA, is the acylation of UDP-*N*-acetylglucosamine (UDP-Glc*N*Ac). This reaction, as all other acylation steps (exceptions noted below), uses acyl-ACP as the acyl donor. For LpxA, β-OH-myristoyl-ACP is required. Due to its importance in my thesis, a more detailed discussion of this reaction and enzyme can be found in Section 1.4.2. The newly acylated product is then deacetylated by LpxC, the first committed step of the pathway (Young et al., 1995, Jackman et al., 1999, Jackman et al., 2000), followed by a second acylation catalyzed by LpxD (Kelly et al., 1993). Similar to LpxA, LpxD is specific for β-OH-myristoyl-ACP. The product of the LpxD reaction has two destinations: one is the LpxH catalyzed loss of UMP (Babinski et al., 2002a, Babinski et al., 2002b) and the second is the LpxB catalyzed combination of the diacyl-UDP product of LpxD and the phospho-diacyl-glucosamine product of LpxH (Ray et al., 1984, Radika & Raetz, 1988). This phospho-tetraacyldisaccharide is phosphorylated to produce lipid IV_A, the precursor to lipid A.

Two Kdo-disaccharide moieties must be added to lipid IV_A before the final two acylations can be performed. Kdo moieties are added in succession by KdtA (Brozek et al., 1989, Clementz & Raetz, 1991, Belunis & Raetz, 1992). The Kdo disaccharide of the inner core of *E. coli* lipopolysaccharide is required for the completion of lipid A acylation (Sweet et al., 2001); Kdo is required for recognition by the final enzymes of the pathway (Brozek & Raetz, 1990, Vorachek-Warren et al., 2002). The final two acylations, catalyzed by LpxL and LpxM, require lauroyl- and myristoyl-ACP, respectively (Brozek & Raetz, 1990); however, both enzymes can use acyl-CoAs as an acyl donor (Six et al., 2008). A detailed illustration of the pathway described above can be found in Figure 4. For more detail on enzymes involved in lipid A biosynthesis not covered in this thesis,

the reader is directed to reviews by Christian Raetz and colleagues (Raetz, 1993, Raetz et al., 2007).

ACP plays a central role in lipoic acid synthesis in bacteria and higher organisms. As mentioned above (Section 1.2.1), Type II FAS also occurs in the mitochondria and/or plastids of mammals (Cronan et al., 2005), fungi (Brody & Mikolajczyk, 1988), and higher plants (Gueguen et al., 2000). However, unlike Type II FAS in bacteria (see Section 1.2.1.2.1) a major product of these Type II FAS systems is octanoyl-ACP (Brody et al., 1997). The octanoyl moiety is converted to lipoyl by lipoic acid synthase (LipA) (Morikawa et al., 2001) and ACP donates lipoic acid to the pyruvate dehydrogenase complex in both mitochondria and *E. coli* (Jordan & Cronan, 1997). Demonstrating its importance to cellular respiration, disruption of mitochondrial ACP results in a sharp decrease in lipoic acid content and results in a respiratory-deficient phenotype (Harington et al., 1993). For a more detailed overview of lipoic acid biosynthesis in *E. coli*, the reader is directed to Jordan and Cronan (1997).

ACP has recently been discovered to be involved in the synthesis of the pimeloyl moiety of biotin in both *E. coli* (Lin et al., 2010, Lin & Cronan, 2012) and *B. subtilis* (Cryle & De Voss, 2004); biotin is an essential enzyme cofactor required by all three domains of life (Knowles, 1989, Attwood & Wallace, 2002). In *E. coli*, pimeloyl-ACP is synthesized from malonyl-CoA using the fatty acid biosynthetic pathway (Lin et al., 2010, Lin & Cronan, 2012). In this process, the terminal carboxyl group of a malonyl-CoA is converted to its methyl ester in a *S*-adenosyl-L-methionine (SAM)-dependent process catalyzed by malonyl-ACP *O*-methyltransferase (BioC); this methylation allows for its use in 3 rounds of FAS. Following FAS, pimeloyl-ACP methyl ester esterase

(BioH) hydrolyzes the methyl ester generating pimeloyl-ACP (Lin et al., 2010, Lin & Cronan, 2012). In *B. subtilis*, pimeloyl-ACP can be generated by an alternate pathway involving the biotin synthesis cytochrome P450 enzyme (BioI). BioI binds a long chain fatty acyl ACP and successively catalyzes the hydroxylation of C7 and C8, producing pimeloyl-ACP (Cryle & De Voss, 2004). These two pathways of biotin synthesis are the first to be discovered and characterized; however, genomic analysis suggests that additional complementary pathways likely exist in other organisms (Rodionov et al., 2002). Regardless of its source, pimeloyl-ACP is then used as the acyl donor in the 8-amino-7-oxononanoate synthase (BioF)-dependent acylation of L-Ala (Webster et al., 2000). The remaining steps of the pathway, catalyzed by 7,8-diaminopelargonic acid synthase (BioA) (Eliot et al., 2002), dethiobiotin synthetase (BioD) (Gibson, 1997), and biotin synthase (BioB) (Ugulava et al., 2003), do not involve ACP and thus will not be covered here; however, the final four steps are conserved across species (Fugate & Jarrett, 2012). For a more detailed overview of biotin synthesis the reader is directed to Fugate and Jarrett (2012).

Quorum sensing is the process by which bacteria detect the populations of their own and other species and synchronize population behavior using small molecules termed autoinducers. In Gram-negative bacteria this process is primarily controlled by *N*-acyl-homoserine lactones (AHL). AHLs are synthesized by two known families of AHL synthases: the LuxI-type and the LuxM-type. Both the LuxI and LuxM/AinS families of autoinducer synthase function similarly, but share no structural homology (Cao & Meighen, 1989). Additionally, both families of AHL synthases use acyl-ACPs to produce AHLs, though it has been noted that AinS also efficiently uses acyl-CoA as a substrate

(Hanzelka et al., 1999). As the acyl chain length and modification of each AHL is very specific for each bacterial species and gene regulatory response network, each AHL synthase has a very defined acyl-ACP specificity. For example, *V. fischeri* produces C6-AHLs (Nealson et al., 1970) while *P. aeruginosa* produces both C12-AHLs (Pesci et al., 1997) and C4-HSL (Pearson et al., 1995, Winson et al., 1995). For more information on bacterial quorum sensing, the reader is directed to a review by Li and Nair (2012).

Quorum sensing was first discovered as a cell density-dependent bioluminescence in marine bacteria such as *V. fischeri* (Nealson et al., 1970) and *V. harveyi* (Byers & Meighen, 1985). In these organisms, ACP plays a limited but important role as an acyl-donor, specifically myristoyl-ACP. The fatty acid is removed by myristoyl-ACP thioesterase (LuxD) to produce a free fatty acid; additionally, LuxD has been shown to be active with myristoyl-CoA (Byers & Meighen, 1985). Fatty acids produced by LuxD are transferred to a fatty acid reductase complex (LuxCE) to convert myristate to its aldehyde form in a NADPH-dependent process (Rodriguez et al., 1983, Byers & Meighen, 1985, Wall et al., 1986). This myristaldehyde product is the substrate for luciferase, which uses oxygen and $FMNH_2$ to produce light (Meighen, 1993). For a more detailed description of bioluminescence, the reader is directed to reviews by Edward Meighen (Meighen, 1988, Meighen, 1993).

ACP also has known roles in protein acylation (Stanley et al., 1994, Stanley et al., 1999) and membrane-derived oligosaccharide synthesis (Therisod et al., 1986, Therisod & Kennedy, 1987). For more information on ACP's involvement in these processes the reader is directed to the noted references. Overall, it is clear that the length of the attached fatty acid chain is important to most, but not all, ACP partner enzymes.

1.2.1.3 Other Carrier Proteins

As mentioned earlier, other types of ACP-like proteins exist in the ACP family of proteins: these include peptidyl carrier protein (PCP), polyketide synthase (PKS) ACPs, and D-alanyl carrier protein (DCP). As for ACP, in order for these synthases to function in primary and secondary metabolism, the central conserved Ser residue must be modified by a PPTase, providing a free sulphydryl to which intermediates and products are attached (Beld et al., 2014). PCP is a carrier protein domain of large multifunctional modular megasynthases (similar to that of FAS Type I ACP) found mainly in prokaryotes and lower eukaryotes and is involved in the assembly line-like biosynthesis (Marahiel et al., 1997) of peptide-based natural products, including the antibiotic vancomycin, the iron sequestering enterobactin (Cane & Walsh, 1999) and glycopeptidolipid components of the cell wall (Lautru & Challis, 2004). To function, a NRPS needs at the minimum: a condensation domain, an adenylation domain, and a PCP domain (Weber & Marahiel, 2001). In this system, the adenylation domain recognizes and activates the amino acid substrate through the formation of an aminoacyl adenylate. PCP then forms a covalent thioester bond through its PP moiety. Lastly, the condensation domain cataylzes peptide bond formation between two adjacent NRPS units (Weber & Marahiel, 2001). For more detail on NRPS and the role of carrier proteins, the reader is directed to the review by Lai and colleagues (Lai et al., 2006).

The PKS family of proteins, found in marine filamentous bacteria, fungi and higher plants, is related to NRPS systems and is responsible for producing therapeutically important natural products, including erythromycin and lovastatin (Lai et al., 2006). Unlike NRPS which are typically Type I, PKS systems exists in both the megasynthase

polypeptide (Type I) and dissociated (Type II) classifications (Cane & Walsh, 1999); similar to FAS (discussed above). Type I PKSs are organized similarily to NRPS, with the minimally required ketosynthase (KS), acyltransferase and ACP domains (Cane & Walsh, 1999). First, the acyltransferase domain attaches an acyl chain, primarily malonyl or methylmalonyl, to the ACP domain using the appropriate acyl-CoA. The KS domain then catalyzes the formation of a C-C bond between two adjacent PKS modular units through an essential Cys residue (Cane & Walsh, 1999). As with FAS, Type II PKS functions with discrete proteins rather than a single multifunctional megasynthase. The minimal protein requirements for Type II PKSs to function are: KS, chain initiation factor (CIF) and ACP (Shen & Kwon, 2002). The essential active site Cys involved in KS activity in PKS Type I systems is replaced by Gln in CIF. It has been demonstrated that CIF catalyzes the decarboxylation of malonate, providing the acetyl starter units required for aromatic polyketide assembly. Type II PKS systems lack the acyltransferase domain of Type I systems; however, it is thought that the host FAS malonyl-CoA:ACP transacylase provides this activity (Carreras & Khosla, 1998). In the minimal Type II PKS system described above, the acetyl starter unit is transferred from ACP to KS, and subsequently decarboxylation and condensation transfers the growing polyketide chain to a malonyl-ACP extender, yielding a linear polyketo-acyl-ACP intermediate (Staunton & Weissman, 2001). A Type III PKS system exists in higher plants and some bacteria (Funa et al., 1999); however, it lacks ACP (Hopwood, 1997) and thus will not be discussed here.

Unlike NRPS and PKS ACPs, D-alanyl carrier protein (DCP) only exists as a discrete protein (Type II). DCP function is required for the formation of D-alanyl-

lipoteichoic acid that is a component of the thick peptidoglycan layer of the cell wall of most Gram-positive bacteria (Heaton & Neuhaus, 1994). A decrease in D-alanylation in the cell wall abolishes biofilm production (Gross et al., 2001), impairing bacterial virulence (Collins et al., 2002) and increasing susceptibility to cationic antibiotics (Peschel et al., 2000). The *B. subtilis* system, which seems to be widespread among Gram-positive bacteria (Neuhaus & Baddiley, 2003), consists of DltA (D-alanine- D-alanyl carrier ligase), DtlB (D-alanyl-lipoteichoic acid biosynthesis acyltransferase protein), DltC (DCP) and DltD (poly- D-alanine transfer protein). DltA, which resembles the adenylation domain of NRPS, selects for D-Ala. The adenylated D-Ala is then transferred to the PP moiety of DCP and subsequently to lipoteichoic acids with the aid of DltB and DltD (Debabov et al., 2000, Kiriukhin & Neuhaus, 2001).

1.3 ACP: A Conserved and Dynamic Protein

In this section, the structure of "ACP Family" proteins will be discussed with a focus on the founding member, ACP.

1.3.1 Structural Features of Carrier Proteins

With the sequencing of over 250 bacterial genomes now complete, a noteworthy degree of conservation exists in primary sequences of Type II ACPs (Figure 5). This high degree of conservation is likely due to evolutionary constraints caused by the requirement of these ACPs to interact with multiple enzyme partners (Byers & Gong, 2007). It is most noticeable in the Helix II region of ACP that contains the phosphopantetheine attachment site, where in addition to the conserved Ser residue, to which the PP is attached (Ser 36 in *E. coli* ACP (Boom & Cronan, 1989)), the two surrounding residues are also highly

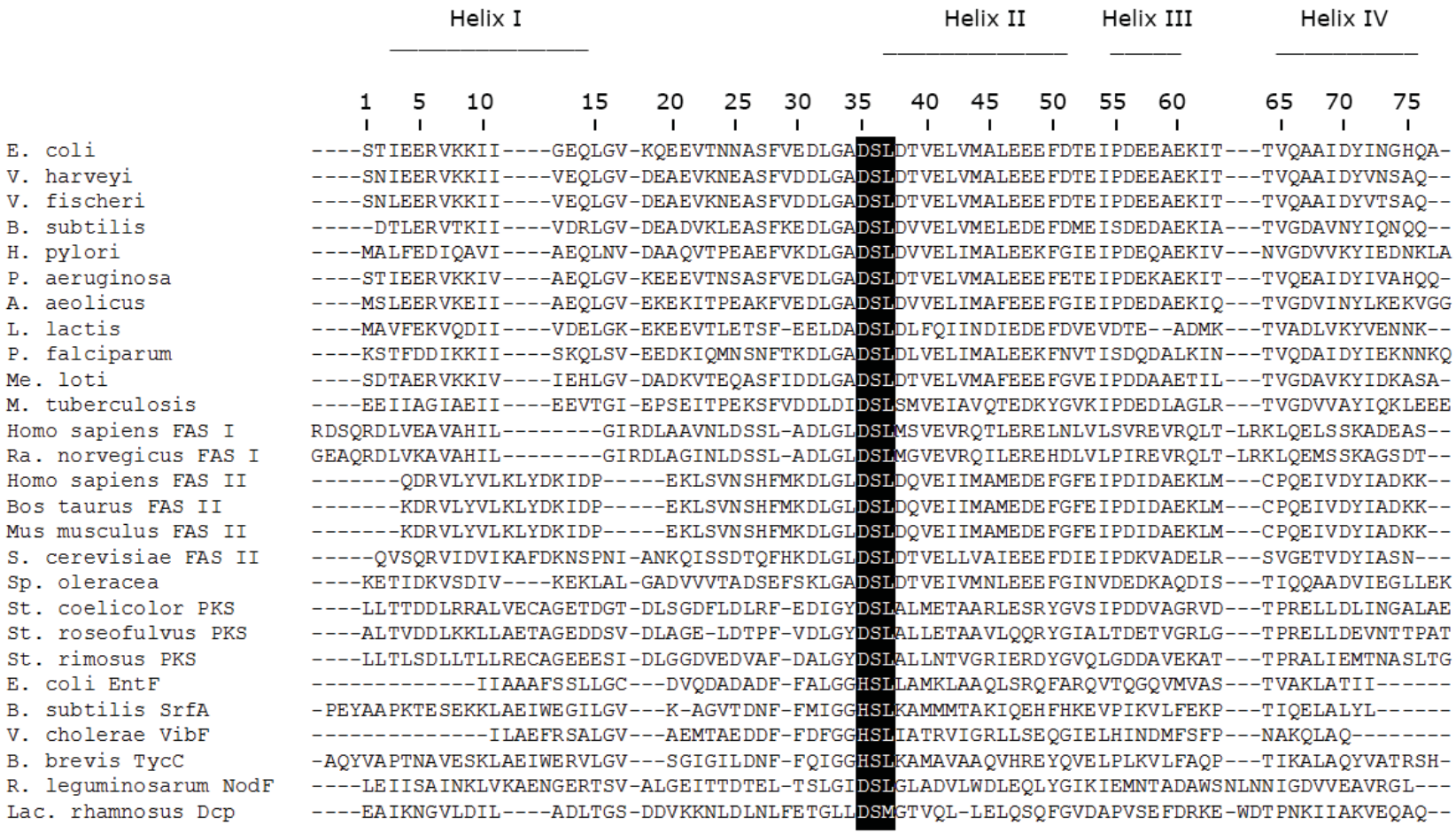

Figure 5 Sequence alignment of selected ACP and ACP-like proteins. Clustal Omega (European-Bioinformatics-Institute. *Clustal Omega*.) was used to produce the above figure. The highly conserved sequence (DSL) recognized by members of the phosphopanthienyl transferase family is highlighted. Numbering of residues and helices are based on *E. coli* ACP.

conserved. These three residues make up the recognition motif DSL (Figure 5; black background) for members of the phosphopantetheinyl transferase family of enzymes (Mofid et al., 2002), which includes AcpS (discussed in detail in Section 1.4.3). The protypical *E. coli* ACP shares >80% sequence identity with ACPs from other Gram-negative bacteria, including *V. harveyi* ACP, which is the focus of this work (Byers & Gong, 2007). Even the more distantly related Type I ACP domain and ACP-like proteins share a significant amount of sequence identity with *E. coli* ACP compared to randomly aligned sequences (<10% sequence identity (Rost, 1997)): 21% – 27% for Type I ACPs, Type II PKS ACPs, and DCP; ~15% for *E. coli* nodulation protein (NodF) and PCPs (Byers & Gong, 2007). Despite a considerably lower sequence identity of ACP homologues, the DSL motif found in Type I and II ACPs is conserved (Figure 5). Furthermore, the DSL motif is consistently positioned at the N-terminal end of Helix II in "ACP Family" proteins (Crump et al., 1996).

ACPs are primarily comprised of α-helices, which account for more than 50% of their protein structure (Schulz, 1975, Kim & Prestegard, 1990). As mentioned previously, Helix II is highly conserved across ACP and ACP-like proteins. This acidic helix has been found to be important for ACP interaction with partner enzymes (Zhang et al., 2003a), including: AcpS (Parris et al., 2000), LpxA (Jain et al., 2004, Williams et al., 2006) and FAS (Zhang et al., 2003a). Interaction of LpxA and AcpS with ACP will be discussed in further detail below (see Sections 1.4.2 and 1.4.3, respectively).

Despite low sequence identity between Type II ACPs and other "ACP Family" proteins, striking structural similarity is present (Figure 6; 0.375 Å RMSD across structures). In all family members, a four-helix bundle fold is adopted, with Helix I being

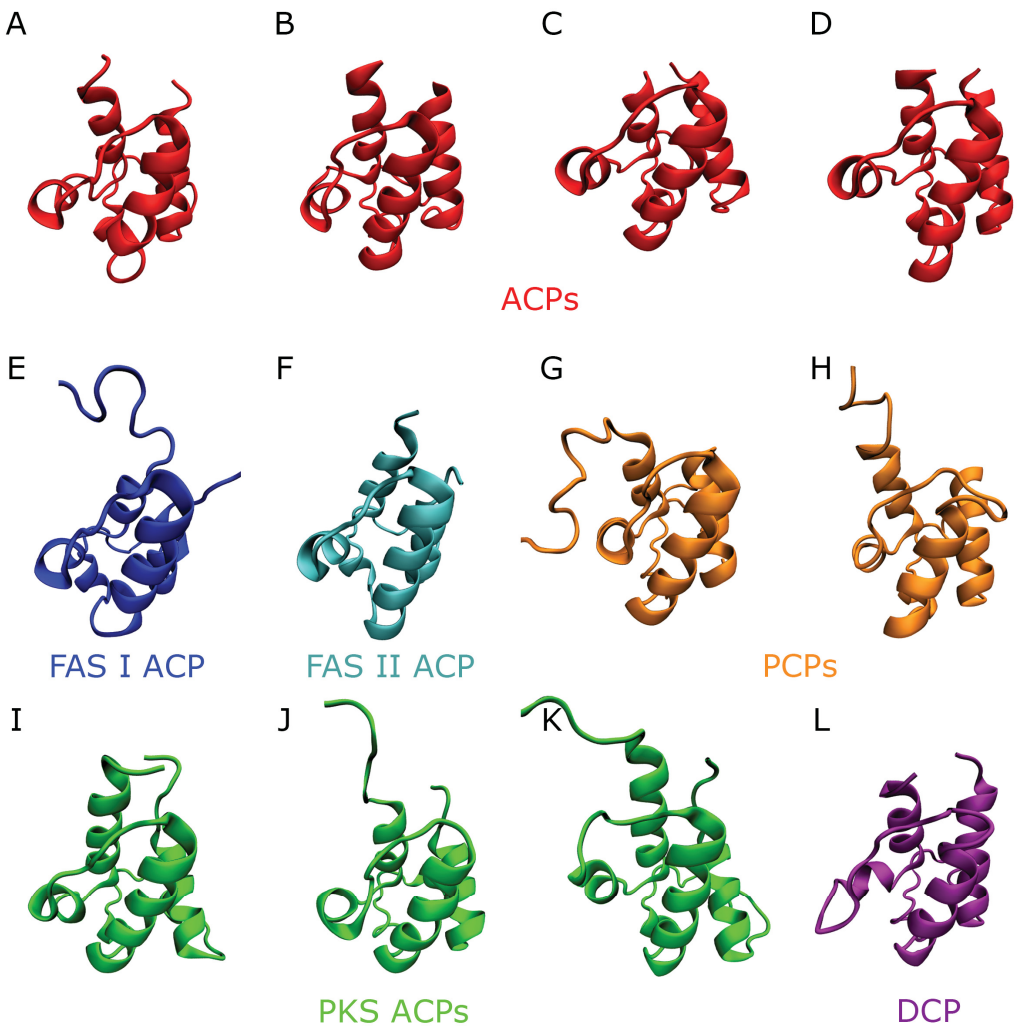

Figure 6 Structures of various ACP and ACP-like proteins. *E. coli* (A: PDB ID = 2K92 (Wu et al., 2009)), *B. subtilis* (B: PDB ID = 1HY8 (Xu et al., 2001)), *S .aureus* (C: PDB ID = 4DXE) and *P. falciparum* (D: PDB ID = 3GZM (Gallagher & Prigge, 2010)) ACPs (red) share a conserved three-dimensional structure with FAS type I ACP (blue; *Ra. norvegicus*, E: PDB ID = 2PNG (Ploskoń et al., 2008)), FAS type II ACP (cyan; *H. sapiens*, F: PDB ID = 2CG5 (Bunkoczi et al., 2007)), peptidyl-carrier proteins (orange; *E. coli* EntF, G: PDB ID = 3TEJ (Liu et al., 2011); *B. subtilis* SrfA, H: PDB ID = 2VSQ (Tanovic et al., 2008)), polyketide carrier proteins (green; *St .coelicolor* PKS, I: PDB ID = 2K0Y (Evans et al., 2008); *St. roseofulvus* PKS, J: PDB ID = 1OR5 (Li et al., 2003); *St .rimosus* PKS, K: PDB ID = 1NQ4 (Findlow et al., 2003)) and D-alanyl carrier proteins (purple; *Lac. casei* DCP, L: PDB ID = 1DV5 (Volkman et al., 2001)). All models are displayed using "NewCartoon" representation. Only one structure out the ensemble was chosen for NMR structures. Prior to image creation, structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD 1.91 and POV-Ray 3.6.

antiparallel to Helices II and IV (Byers & Gong, 2007). Even though a highly conserved global fold exists, superimposition of these structures (Figure 7) reveals slight variations; mainly, the packing and length of the helices in addition to considerable variation in the loop regions connecting the helices (White et al., 2005). Differences exist within Type II ACPs as well; for example, NMR spectroscopy experiments have been used to determined that *E. coli* exists in two or more states in fast exchange (Kim & Prestegard, 1989), while spinach (Kim et al., 1990) and *P. falciparum* (Sharma et al., 2006) ACP exists in two isoforms that undergo slow exchange. Furthermore, NMR spectroscopy experiments have been used to study the interaction of PP and ACP, which varies slightly across species. In *E. coli*, residues Asp 35 – Met 44 (helix II) as well as Ile 54 (loop connecting helix II and III), Glu 60 (helix III), Ile 62 (loop connecting helix III and IV) and Ala 68 (helix IV) of holo-ACP show significant chemical shift perturbations compared to the apo form (Kim et al., 2006). Structurally, the side chains of the aforementioned residues (with the exception of Glu 60) point toward the hydrophobic cleft on the interior of the protein. Chemical shift perturbations between the apo and holo forms of ACP indicate that PP transiently interacts with these residues (Kim et al., 2006). ACP from spinach interacts with PP similarly (Wong et al., 2002), while *P. falciparum* ACP (*P.f.*ACP) demonstrates interactions between PP and the polypeptide backbone (Sharma et al., 2006), a first for ACPs.

It has long been known that ACP's conformation is sensitive to binding of divalent cations; this is thought to be due to charge neutralization of the highly acidic ACP (Schulz, 1975, Therisod et al., 1986), and binding of divalent cations to ACP induces or stabilizes a folded conformation. Early NMR and biophysical characterization

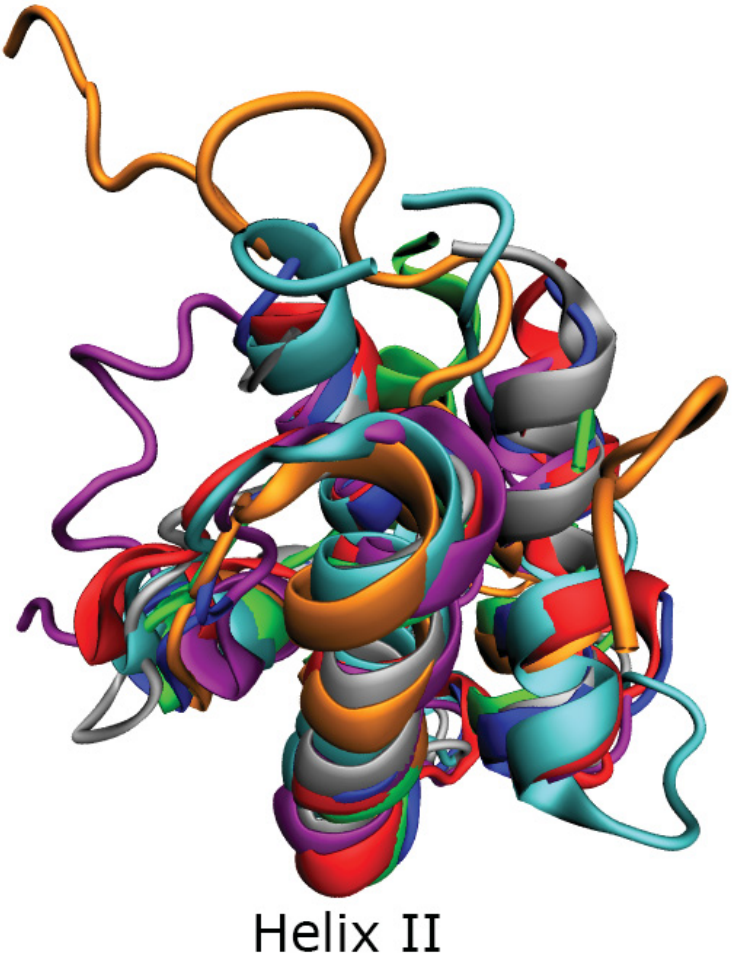

Figure 7 Selected ACP and ACP-like protein structures overlaid. "NewCartoon" representation of *E. coli* ACP (blue: PDB ID = 2K92 (Wu et al., 2009)), *P. falciparum* ACP (red: PDB ID = 3GZM (Gallagher & Prigge, 2010)), *Ra. norvegicus* FAS type I ACP (orange: PDB ID = 2PNG (Ploskoń et al., 2008)), *H. sapiens* FAS type II ACP (green: PDB ID = 2CG5 (Bunkoczi et al., 2007)), *E. coli* peptidyl-carrier protein EntF (purple, G: PDB ID = 3TEJ (Liu et al., 2011)), *St. coelicolor* PKS ACP (cyan: PDB ID = 2K0Y (Evans et al., 2008)) and *Lac. casei* DCP (silver: PDB ID = 1DV5 (Volkman et al., 2001)). Prior to image creation, structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. RMSD across all structures is 1.2 Å as calculated by Chimera during alignment. Images were created using VMD 1.91 and POV-Ray 3.6.

of ACPs (Schulz, 1975, Frederick et al., 1988, Tener & Mayo, 1990) suggested two cation-binding sites in *E. coli* ACP, and recent NMR work has identified and confirmed the existence and location of these sites in *E. coli* ACP (Kim et al., 2006). These sites have been found in ACPs from other bacteria, such as *V. harveyi* (Gong et al., 2007, Chan et al., 2010), and are presumed to exist in all bacterial ACPs. The two sites, termed Sites A and B, are positioned at the N- and C-terminus of Helix II, respectively, and have been shown to bind Mg^{2+} and Ca^{2+} (Kim et al., 2006) (Figure 8), although it is thought that other divalent cations such as Mn^{2+} could bind as well (Tener & Mayo, 1990). In *E. coli* ACP, residues Glu 30, Asp 35 and Asp 38 make up Site A at the N-terminus of Helix II, while Site B is comprised of Glu 47, Asp 51, Glu 53 and Asp 56 (Kim et al., 2006). Like divalent cations, fatty acylation of the PP group has been shown to induce folding in ACPs (Jones et al., 1987a, Jones et al., 1987b) through interaction with specific residues lining its hydrophobic pocket (Gally et al., 1978, Jones et al., 1987a). This early NMR work suggested that Ile 54 and Ala 59 of *E. coli* ACP are directly involved in binding an attached fatty acyl chain (Jones et al., 1987a), although this work was restricted to examination of the C5 position of the fatty acid due to limitations on the availability of fluorinated acyl chains. The crystal structure determination of *E. coli* butyryl-ACP confirmed the involvement of Ile 54 and also implicated Val 7, Leu 46, Met 62, Tyr 71, and Ile 72 (Roujeinikova et al., 2002), supporting the earlier model that ACP sequesters its attached fatty acyl chain in the hydrophobic core of the protein (Mayo & Prestegard, 1985). More recent structural studies on *E. coli* (Roujeinikova et al., 2007) and spinach (Zornetzer et al., 2006) long chain acyl-ACPs have confirmed the above model and suggested an intriguing possibility for ACP partner enzyme selectivity: a flexible

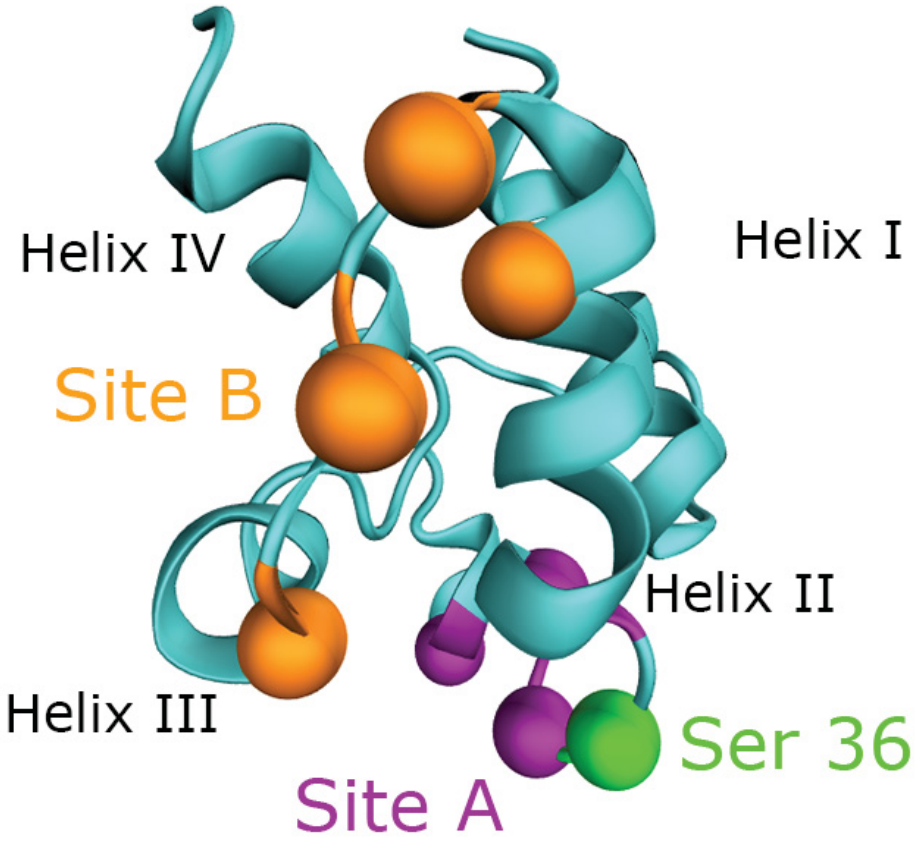

Figure 8 Divalent cation binding sites in *E. coli* ACP. "NewCartoon" representation of *E. coli* ACP (cyan: PDB ID = 2K92 (Wu et al., 2009)) with divalent cation binding sites A (purple) and B (orange), and Ser 36 (green) highlighted using "VdW" representation. Images were created using VMD 1.91 and POV-Ray 3.6.

hydrophobic FA binding pocket, where increasing the acyl chain length causes a shift in Helix I, allowing the acyl chain to interact with Helices II and IV. As the acyl chain length increases, the proximal end of this pocket (containing conserved Ser PP attachment site) and not the distal end of ACP expands laterally (Zornetzer et al., 2006). This is consistent with early evidence demonstrating increasing hydrophobic character of acyl-ACPs longer than C8 (Rock & Garwin, 1979). These studies suggest a potential mechanism for ACP partner enzyme selection of acyl chain length based on the structural differences in acyl-ACPs (Roujeinikova et al., 2007). More recent NMR evidence with various long chain acyl-ACPs support this mechanism (Zornetzer et al., 2010). Taken together, all evidence suggests that ACP is a dynamic protein and that those dynamic properties are essential for function in bacteria (discussed further in Section 1.3.2).

1.3.2 The Importance of Conformational Dynamics to ACP Structure and Function

Conformational dynamics and flexibility are important to ACP function for at least two main reasons. First, ACP has to interact with a multitude of partner enzymes as discussed in Section 1.2.1.2 and illustrated in Figure 1. Most of these enzymes possess different folds and quaternary structure. For example, ACP interacts with LpxA and AcpS; both enzymes are trimeric, but assume very different folds. LpxA exists in a mainly left-handed parallel β-helix structure with small α-helical sections near the C-terminus (Raetz & Roderick, 1995), while AcpS adopts a phosphopanthienyl transferase fold characterized by β-sheets, which help in monomer association, surrounded by α-helices (Parris et al., 2000). Second, as noted previously, many ACP-dependent enzymes require access to the attached acyl chain as part of their catalytic function, so the acyl

chain must be alternately exposed at the active site and sequestered in the hydrophobic pocket in these cases.

Most ACP partner enzymes have basic and/or hydrophobic patches near their active site. These patches have been implicated in an electrostatic interaction with the acidic Helix II of ACP (Zhang et al., 2001). Additionally, the acidic character of the ACP has been implicated in its inherent flexibility due to charge repulsion, with the most dynamic regions often containing the highest concentration of acidic residues (Byers & Gong, 2007). NMR backbone dynamics experiments and hydrogen-deuterium exchange have suggested that Loop I, Helix II, and Helix III of *E. coli* ACP exhibit the greatest flexibility (Andrec et al., 1995, Park et al., 2004, Kim et al., 2006). ACP from *M. tuberculosis* has a large C-terminal extension, which exists as a random coil and is thought to bind the very long chain intermediates in mycolic acid biosynthesis and/or to regulate specific protein–protein interactions (Wong et al., 2002). Despite the inherent flexibility of most ACPs, recent work from our lab and others provides evidence for more stable Type II ACPs, specifically *P. falciparum* ACP (Modak et al., 2007).

Early work indicated that denaturation of ACP was a cooperative two state process without observable stable intermediates (Horvath et al., 1994, Reed et al., 2003, Park et al., 2004). However, more recent studies implicate ACPs in a larger family of proteins known as intrinsically unstructured or natively unfolded proteins (NUPs) (Wright & Dyson, 1999, Tompa, 2002). Although the actual number of NUPs discovered to date is relatively low, neural network predictors (developed to recognize sequences that correspond to structurally disordered regions) have suggested that ~30% – 50% of eukaryotic proteins have at least one long (>50 residues) disordered region (Dunker et al.,

2000, Dunker et al., 2002). These proteins are often involved in transcription, translation, signal transduction and the cell cycle, where multiple partners are often involved (Tompa, 2002). NUPs are characterized by a number of features, including: (i) a high ratio of charged to hydrophobic residues (Dunker et al., 2000, Uversky, 2002) as measured by algorithms such as FoldIndex (Prilusky et al., 2005), (ii) resistance to precipitation by heating or organic solvents, (iii) anomalous behaviour on SDS gels (apparent MW is often 1.2–1.8 times higher than actual MW) and (iv) high intramolecular flexibility (Tompa, 2002).

ACP exhibits all of the above characteristics (Byers & Gong, 2007) and was recently shown to denature on the millisecond timescale using electrospray ionization mass spectrometry (Murphy et al., 2007). Most importantly, the hallmark of NUPs is the coupling of binding and folding. This has been recently demonstrated for *V. harveyi* ACP (Gong et al., 2008) as well as the distantly related PCP (Lai et al., 2006). However, while ACP does seem to fit into the NUP category *in vitro*, it should be noted that even the relatively unstable *V. harveyi* ACP (discussed in further detail below) would likely be folded *in vivo* due to intracellular Mg^{2+} levels; which have been estimated to be 1 – 2 mM (Moncany & Kellenberger, 1981, Alatossava et al., 1985). A functional advantage of NUPs is that the decrease in conformational entropy upon partner binding, which effectively uncouples binding strength from specificity (Tompa, 2002), contributing to the low affinity reversible interactions that are often seen for proteins that have multiple partners.

1.3.3 Insights from *V. harveyi* ACP Model System

Due to its additional role in bioluminescence, our laboratory has characterized the ACP from the luminescent marine Gram negative bacterium *V. harveyi*. *V. harveyi* and *E. coli* ACPs are 86% identical (Shen & Byers, 1996), with only one, of the eleven differences, a non-conservative change in a helical region (Flaman et al., 2001): Val occurs in position 12 of *V. harveyi* ACP compared to Gly in *E. coli*. Despite these differences in primary sequence (Figure 5), these ACPs share a number of biophysical properties, including: similar hydrodynamic radii and an alkaline pH-induced expansion of this radius that can be largely prevented by attachment of fatty acids (de la Roche et al., 1997). Nevertheless, structural differences do exist; first, *V. harveyi* ACP is more acidic than its *E. coli* counterpart (Shen & Byers, 1996, de la Roche et al., 1997), containing two more acidic residues and being one residue shorter. However, the most striking is the difference in conformation at physiological pH: *E. coli* ACP is folded while *V. harveyi* does not adopt its native conformation *in vitro*; although it should be noted that *E. coli* ACP is only marginally stable at this pH (Schulz, 1975). As discussed below, this difference has been partially attributed to the presence of a basic amino acid (His) at position 75 of *E. coli* ACP, replacing Ala in *V. harveyi* ACP (Keating et al., 2002).

Aside from the discovery that *V. harveyi* ACP is unfolded at physiological pH in the absence of divalent cations, Flaman and coworkers (2001) were the first to create *V. harveyi* ACP mutants to investigate the importance of specific residues in ACP structure and function (Table 2). This initial work targeted residues (Phe 50, Ile 54, Ala 59, and Tyr 71) previously implicated in interaction with the first 6 – 8 carbons of an attached

Table 2 *V. harveyi* ACPs created by the Byers Lab to date.

ACP[1]	Description of Protein/Reason Position Chosen	Source Reference
rACP	Recombinant form of *V. harveyi* ACP; contains 4 extra N-terminal residues (GIPL) which remain after purification tag removal.	(Flaman et al., 2001)
V12G	Position 12 of *E. coli* ACP contains Gly. This mutant was made to determine if the position was responsible for differences between *E. coli* and *V. harveyi* ACP.	
F50A	These positions are implicated in binding the first 6 – 8 carbons of an attached acyl chain	
I54A		
I54V		
I54L		
A59G		
Y71A		
D18K	These positions contain Lys and His, respectively, in *E. coli* ACP. Mutants were made to determine if these are responsible for differences between *E. coli* and *V. harveyi* ACP.	(Keating et al., 2002)
A75H		
D18K/A75H		
E41D	These positions were chosen as they had been previously implicated in binding of FAS enzymes.	(Gong & Byers, 2003)
E41K		
A45G		
A45C		
A45W		
V43I	This mutation causes a stabilization of *E. coli* ACP structure, and was made to see if the same effect would be seen in the less stable *V. harveyi* ACP.	
SA	Divalent cation Site A mutant (D30N, D35N, D38N).	(Gong et al., 2007)
SB	Divalent cation Site B mutant (E47W, D51N, E53Q, D56N).	
SA/SB	Double divalent cation site mutant.	
D30N	Single point mutants of Site A.	
D35N		
D38N		
E47Q	Single point mutants of Site B.	
E53Q		
D56N		
E25W	Trp-substitution in Loop I (E25), Helix II (L46, F50) and Helix IV (V72) to examine the microenvironment of these positions using fluorescence	(Gong et al., 2008)
L46W		
F50W		
V72W		
S36A	Mutation of PP-attachment site for ACP.	Unpublished
P55A	Non-conservative replacement in key Helix III region	Unpublished

Note: [1] = Numbering based on wild-type *V. harveyi* ACP. All mutants were created using rACP as a template.

fatty acyl chain (Rock, 1983, Mayo & Prestegard, 1985, Jones et al., 1987a), as well as Val 12, a non-conservative sequence difference between *E. coli* and *V. harveyi* ACP. Circular dichroism (CD) analysis indicated that, like native *V. harveyi* and recombinant *V. harveyi* ACP (rACP), all mutants were unfolded in the absence of and folded upon the addition of Mg^{2+} (Flaman et al., 2001), with the exception of F50A and I54A which were incapable of folding in the presence of divalent cations. This was confirmed using conformationally sensitive native PAGE. These ACP mutants were also tested as substrates for *V. harveyi* AAS and *V. harveyi* myristoyl-ACP thioesterase. Most fatty acid binding pocket mutants (Table 2) exhibited significantly decreased acylation compared to rACP, caused by changes in K_M and V_{max} (Flaman et al., 2001). Of note, the F50A and I54A mutants, which were incapable of adopting a folded conformation, exhibited little to no AAS activity. In contrast, V12G was found to have no effect on acylation by AAS and none of the mutations affected the production of myristate by myristoyl-ACP thioesterase (Flaman et al., 2001).

Keating *et al*. (2002) continued analysis of the 11 amino acids that differ between *V. harveyi* and *E. coli* ACP, in particular focusing on two residues (Glu 18 and Ala 75) in which the charge is reversed (Table 2). CD analysis in the presence and absence of Mg^{2+} indicated that unlike D18K, which acted similarly to rACP, A75H and the double mutant (D18K/A75H) exhibited a folded conformation even in the absence of divalent cations (Keating et al., 2002). AAS activity was not impaired by any of these mutations, indicating that the folded structure attained by A75H was the native conformation. It was concluded that His at position 75 in *E. coli* is likely important in maintaining the conformational stability at physiological pH (Keating et al., 2002), likely due to charge

neutralization because of the unusually high pKa (7.5) of this residue; this was subsequently confirmed by NMR analysis of the *V. harveyi* A75H mutant (Chan et al., 2010). Indeed, it has also been demonstrated that addition of a hexahistidine tag to the C-terminus of *V. fischeri* ACP induces a folded conformation in the absence of divalent cations (D.M. Byers, unpublished observations).

The role of Helix II residues in FAS activity were explored by Gong and Byers (2003). Specifically, both the negative side chain of Glu 41 and small size of Ala 45 have been implicated in the interaction of *E. coli* ACP Helix II with the FAS condensing enzyme FabH (Zhang et al., 2001). None of the replacements at Glu 41 (E41D, E41K) or at Ala 45 (A45G, A45C, A45W) (Table 2) had any apparent effect on ACP conformation or impaired activity of *V. harveyi* AAS. However, E41K (but not E41D) was non-functional with partially purified *V. harveyi* FAS, indicating that an acidic residue at position 41 is essential for FAS, but not AAS, activity (Gong & Byers, 2003).

More recent mutational analysis has focused on the previously identified divalent cation binding sites in Helix II (Schulz, 1975, Therisod et al., 1986, Tener & Mayo, 1990). Gong and coworkers (2007) created individual (SA, SB) and double (SA/SB) site mutants as well as 6 single point mutations within the two sites (Table 2). In all cases, an acidic residue was mutated to the corresponding amide, thus neutralizing the acidic charge. CD was used to demonstrate that both individual site mutants (SA, SB) exhibited a partially stabilized α-helical structure compared to rACP in the absence of Mg^{2+}; upon addition of Mg^{2+}, both mutants folded similarly to rACP (Gong et al., 2007). Furthermore, the double site mutant (SA/SB) adopted a folded, α-helical conformation independent of divalent cations. Interestingly, all of the individual point mutations

(D30N, D35N, D38N, E47Q, E53Q, and D56N) exhibited CD properties similar to rACP in the absence of Mg^{2+} and similar behavior upon the addition of Mg^{2+}. From these results it was concluded that Site A and Site B bind Mg^{2+} independently and have an additive effect on ACP conformation (Gong et al., 2007), supporting the early hypothesis (Schulz, 1975, Therisod et al., 1986) that binding of divalent cations to ACPs induces a folded conformation through charge neutralization (see Section 1.3.1). Interestingly, mutation of Site A (SA) blocked AcpS activity, while mutations in SB preferentially impaired ACP function with AAS and LpxA.

The absence of tryptophan in ACP was exploited recently by Gong and colleagues (2008) by substituting Trp into various positions of ACP (Table 2) to act as intrinsic fluorescent probes of the microenvironment at each position. All Trp-substituted ACPs, including the previously characterized A45W (Gong & Byers, 2003), were effective substrates for FAS and AAS, and CD analysis in the presence and absence of Mg^{2+} showed that they behave as rACP conformationally (Gong et al., 2008). In the absence of Mg^{2+}, all Trp-substituted ACPs exhibited a peak fluorescence emission wavelength of ~350 nm irrespective of position in the folded protein, indicative of a polar microenvironment and thus an unfolded conformation under these conditions. Addition of Mg^{2+} caused a small red and blue shift for E25W and A45W, respectively, indicating that they are fully exposed on the surface of folded ACP (Gong et al., 2008). L46W exhibited a full blue shift, with a peak wavelength of ~310 nm, indicative of a hydrophobic microenvironment for this Helix II residue, which is buried in the fatty acid binding pocket in folded ACP. F50W and V72W exhibited a partial blue shift, with a peak wavelength of ~335 nm in the presence of Mg^{2+}. The effect of fatty acid attachment

on the Trp microenvironment was also examined. Acylation of L46W, F50W and V72W with increasing chain length fatty acyl groups caused a progressive blue shift, while the same treatment of E25W and A45W resulted in small red and blue shifts, respectively (Gong et al., 2008). Comparison of Trp-substituted holo- and acyl-ACPs in the presence and absence of Mg^{2+} suggested that stabilization by divalent cations and acylation do not result in identical ACP conformers. Overall, it was concluded that L46W and A45W are good probes of ACP conformation and Helix II interaction, respectively (Gong et al., 2008).

1.4 Interaction with Partner Enzymes: LpxA and AcpS

1.4.1 General Considerations in ACP-Partner Interactions

Despite the availability of sequence and structural information for a number of ACP enzyme partners, no clear ACP binding motif appears to exist (Byers & Gong, 2007). This is likely due to the wide variation in enzyme classes with which ACP must interact. However, some general observations have been made. As noted above, Helix II of ACP plays an important role in interaction with almost all partner enzymes; it has thus been termed the "recognition helix" of ACP (Parris et al., 2000, Zhang et al., 2003a, Chan et al., 2008). Examination of structural information for ACP partner enzymes reveals a basic-hydrophobic patch located near the active site that has been implicated in electrostatic interaction with the acidic helix II of ACP (Zhang et al. 2001). Another general feature is that ACP exhibits relatively weak interactions with most binding partners; K_M values are typically in the low micromolar range (Byers & Gong, 2007).

Though no specific binding motif exists across ACP-dependent enzymes, many enzymes require specific (and in some cases different) regions and/or residues of ACP for

function. A few known examples of such enzymes are, FabH, FabG, AcpS, LpxA and AAS. FabH interaction with ACP, initially characterized *in silico* and confirmed *in vitro*, has been shown to be dependent on Glu 41 and the N-terminal region of ACP (Zhang et al., 2001). Glu 41 is thought to be required for a crucial interaction with Arg 249 of FabH that, together with interaction of Helix II, orients ACP correctly in the active site, allowing for catalysis (Zhang et al., 2001). Conversely, both FabG and AcpS have been shown to be dependent on residues near Helix III, in addition to Helix II, which are important for binding (Parris et al., 2000, Zhang et al., 2003b). As noted previously, AcpS has been shown to be sensitive to mutations in the N-terminal region of Helix II (divalent cation binding Site A), whereas mutations to the C-terminal region of Helix II (divalent cation binding Site B) greatly affect both LpxA and AAS activity (Gong et al., 2007). LpxA-ACP and AcpS-ACP interactions are discussed in further detail below (see Sections 1.4.2 and 1.4.3, respectively). Many of these enzyme-specific requirements are undoubtedly due to differences in the nature of the (acyl-)ACP substrate involved and the catalytic reaction performed. For example, acyltransferases (e.g. LpxA) and FAS component enzymes would require a conformational change in ACP to allow access to the sequestered fatty acid, while others such as the phosphopantetheinyl transferase AcpS would likely not require such a change.

1.4.2 Bacterial LpxA

As discussed briefly above (see Section 1.2.1.2.2), ACP is involved in the synthesis of lipid A (Figure 4), the main component of the outer membrane of Gram-negative bacteria (Raetz, 1993, Raetz et al., 2007). The first step in the synthesis of lipid A is catalyzed by UDP-*N*-acetylglucosamine acyltransferase (LpxA), a cytoplasmic

member of the acyltransferase family of proteins. *E. coli* LpxA, the product of the *lpxA* gene, is a 262 amino acid protein (~28 kDa) in its monomeric form and catalyzes the transfer of β-hydroxymyristate from ACP to UDP-Glc*N*Ac (Anderson & Raetz, 1987, Coleman & Raetz, 1988). For catalysis to occur, LpxA must form a trimeric complex, as the active site of the enzyme exists in the cleft between two adjacent subunits (Wyckoff & Raetz, 1999). Each monomer is composed of two domains: a N-terminal domain (residues 1 to 186) comprised mainly of β-sheet structures, and a C-terminal domain (residues 187 – 262) comprised mainly of α-helices (Raetz & Roderick, 1995). Examination of LpxA sequences from various Gram-negative bacteria (Figure 9) shows conserved residues, most of which cluster around the active site cleft. Similar to other ACP-dependent enzymes, this cleft is comprised of multiple His and Lys residues (Wyckoff & Raetz, 1999). Due to the three-fold symmetry of LpxA, each multimeric structure contains three active sites (Figure 10).

In the biologically relevant direction, LpxA catalyzes the S-O transfer of β-hydroxymyristoyl group from the PP of ACP to the 3-OH position of UDP-Glc*N*Ac (Anderson & Raetz, 1987, Coleman & Raetz, 1988). However, the equilibrium constant of this reaction favours the substrates rather than the products, potentially due to the hydrophobic stabilization of ACP by the attached acyl chain (Anderson et al., 1993). As a result, the LpxC-catalyzed second step in the lipid A pathway (Figure 4) drives the LpxA reaction through removal of the product, UDP-3-*O*-(R-3-hydroxymyristoyl)-Glc*N*Ac (Anderson & Raetz, 1987), and is thus the committed step of lipid A biosynthesis (Young et al., 1995). Early non-directed mutagenic studies of the endogenous *lpxA* gene resulted in the isolation of a temperature-sensitive strain that is unable to grow at 42ºC unless

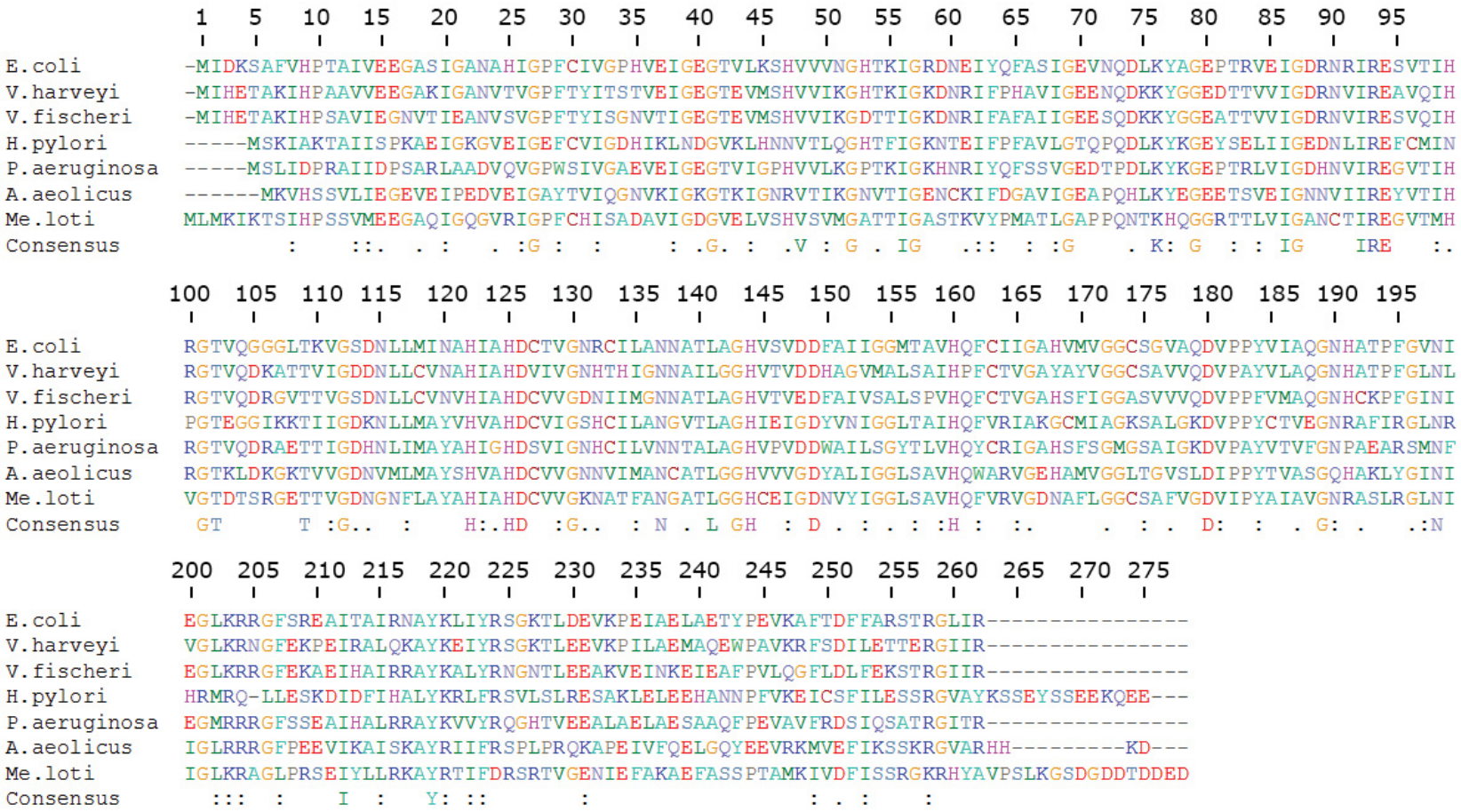

Figure 9 Sequence alignment of LpxA from selected organisms. Clustal Omega (European-Bioinformatics-Institute. *Clustal Omega*.) was used to produce the above figure. Numbering of residues is based on *E. coli* LpxA. Coloring is based on amino acid properties. Consensus sequence shown: amino acid given for fully conserved residue, a colon (:) indicates conservation between groups of strongly similar properties and a period (.) indicates conservation between groups of weakly similar properties as determined by Clustal Omega.

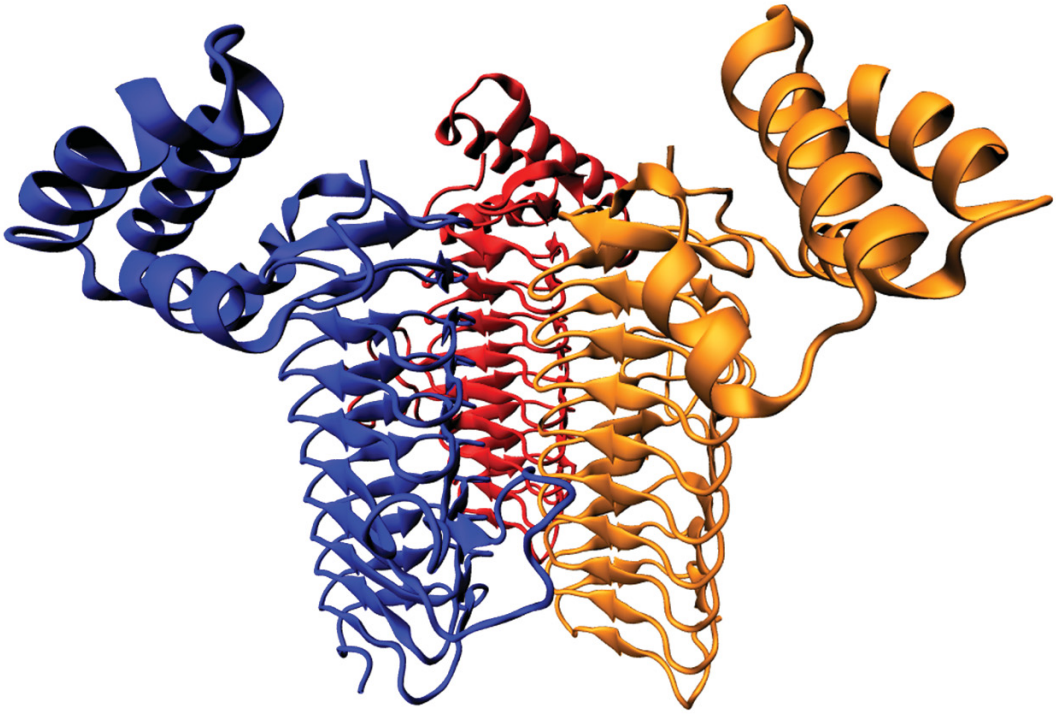

Figure 10 *E. coli* LpxA quarternary structure. “NewCartoon” representation of *E.c.*LpxA colored by subunit: A (blue), B (red) and C (orange). The active site is formed at the interface of adjacent subunits. This figure was created using PDB ID = 2JF3 (Ulaganathan et al., 2007). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

supplemented with wild-type LpxA expressed from a plasmid (Galloway & Raetz, 1990).

LpxA is believed to catalyze the aforementioned S-O transfer (Figure 11) through a general base mechanism (Wyckoff & Raetz, 1999). Mutational and crystallographic studies have implicated a number of residues important for catalysis and substrate binding. His 125 has been shown to be essential for catalytic activity, as mutation of this (but not other nearby His residues) to alanine resulted in no measureable activity (Wyckoff & Raetz, 1999). Additionally, the conserved Asp 126 side chain positions and stabilizes the His 125 side chain by hydrogen bonding (Williams & Raetz, 2007). These two residues form a highly conserved catalytic diad.

As mentioned above, the majority of LpxA conformation is comprised of hexapeptide β-sheets that coil in an extremely regular left-handed triangular helix of 10 coils. This unusual fold was first discovered in LpxA and was termed a left-handed parallel β-helix (Raetz & Roderick, 1995). Six of the 10 coils follow the regular hexapeptide pattern, [LIV]-[GAED]-X_2-[STAV]-X (Vaara, 1992, Jenkins & Pickersgill, 2001), with one turn of the helix comprised of 3 hexapeptide repeats. Two of the remaining four coils (residues 69 to 83 and 99 to 108) are interrupted at corners and disobey the hexapeptide motif by the presence of the polar residues (Asp74 and Gln 104) at the aliphatic residue position. The first and the last coils are abbreviated (Raetz & Roderick, 1995). The shortened final coil connects to the C-terminal domain, comprised of four α-helices (residues 199 to 253). Until recently, LpxA was the only lipid A pathway enzyme with structures determined for the *E. coli* (Raetz & Roderick, 1995) and *H. pylori* (Lee & Suh, 2003) proteins (Table 3). Only very recently have the structures of other enzymes in this pathway been determined, including: LpxC, LpxD, LpxI (a recently

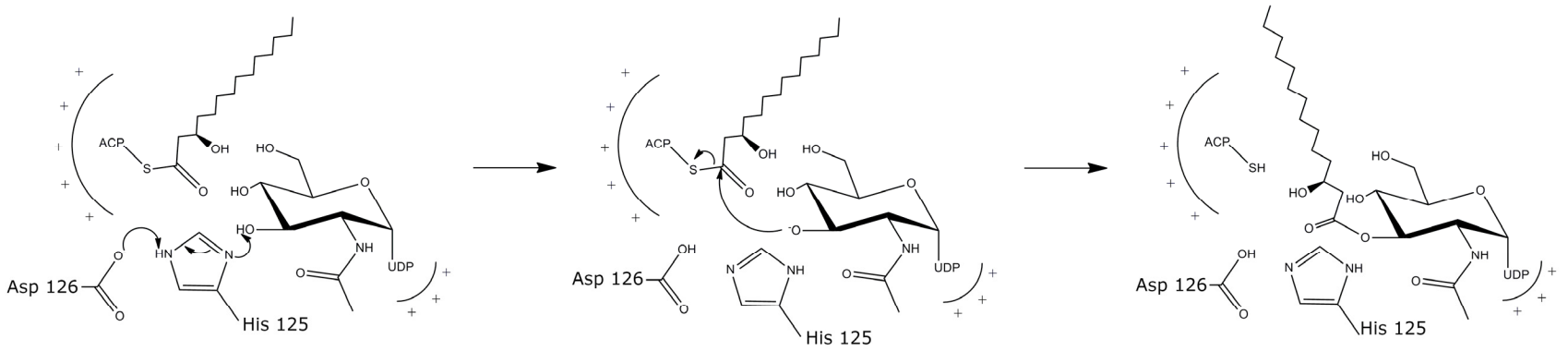

Figure 11 Mechanism of LpxA reaction. Myristoyl-ACP and UDP-Glc*N*Ac bind positively charged residues in the LpxA active site, positioning the 3-OH group on UDP-Glc*N*Ac and thioester bond of β-OH-myristoyl-ACP near His 125 (left), which is positioned and oriented through hydrogen-bonding with Asp 126. This positioning allows His 125 to act as a general base, removing a proton from the 3-OH and allowing for nucleophilic attack of the carbonyl group on the fatty acyl chain (middle); this attack results in the acylation of UDP-Glc*N*Ac at the 3 position of the glucose ring and ACP in its holo form (right). This figure was generated using ChemDraw 12.0 (CambridgeSoft. *ChemDraw*.) and based on Wyckoff and Raetz (1999) and data from Williams and Raetz (2007).

Table 3 Bacterial LpxA structures determined to date.

Species	Bound to	Structure Resolution (Å)	PDB ID	Identity to *E. coli*[1]	Reference
E. coli	N/A[2]	2.60 1.80	1LXA 2JF2	N/A	(Raetz & Roderick, 1995) (Ulaganathan et al., 2007)
	β-hydroxymyristoyl-UDP-Glc*N*Ac[3]	1.74	2QIA		(Williams & Raetz, 2007)
	β-hydroxydecanoyl-UDP-Glc*N*Ac	1.85	2QIV		
	UDP-Glc*N*Ac[4]	3.00	2FJ3		(Ulaganathan et al., 2007)
	P920[5]	1.80	2AQ9		(Williams et al., 2006)
H. pylori	N/A[2]	2.10	1J2Z	42%	(Lee & Suh, 2003)
L. interrogans	N/A[2]	2.10, 2.12	3HSQ, 3I3A	42%	(Robins et al., 2009)
	β-hydroxylauroyl-UDP-Glc*N*Ac[3]	2.10	3I3X		
C. jejuni	N/A[2]	2.30	3R0S	37%	2011-03-23[6]
A. baumannii	N/A[2]	1.80, 1.41	4E6T, 4E6U	53%	(Badger et al., 2012)
B. thailandensis	N/A[2]	1.80	4EQY	53%	(Baugh et al., 2013)
	CMP[3]	2.42	2XCU		

[1] = As determined by Protein BLAST alignment (Altschul et al., 1997, Altschul et al., 2005).
[2] = Structure without substrate bound.
[3] = Reaction product.
[4] = Enzyme substrate
[5] = Antibacterial peptide.
[6] = Structure in PDB Data Bank with out an associated publication; Date of publication to PDB Data Bank given.

discovered replacement for LpxH in *C. crescentus* (Metzger & Raetz, 2010)), LpxK and KdtA.

Crystallographic and mutational analysis has also identified LpxA residues important for substrate binding and specificity. Asp 74, Lys 76, His 122, His 125, His 144, His 160, Gln 161, Asn 198, Glu 200, Arg 204, and Arg 205 are important for proper binding and alignment of the acyl-ACP and UDP-Glc*N*Ac substrates (Wyckoff & Raetz, 1999, Ulaganathan et al., 2007, Williams & Raetz, 2007). Homologous residues are found in *H. pylori* LpxA (Lee & Suh, 2003). In *E. coli* LpxA (*E.c.*LpxA), Gly 173 has been implicated, along with His 191, in acting as a "hydrocarbon ruler", specifically selecting for C14 carbon chains (Wyckoff et al., 1998a). Evidence for the "hydrocarbon ruler" in *E.c.*LpxA comes mainly from mutational analysis. A G173M mutation, mimicking *P. aeruginosa* LpxA (*P.a.*LpxA), changes acyl chain specificity from *R*-3-hydroxymyristoyl to *R*-3-hydroxydecanoyl, the preferred natural substrate of *P.a.*LpxA (Wyckoff et al., 1998a). Furthermore, when the reciprocal mutation, M173G, is made in *P.a.*LpxA, its substrate specificity is changed to *R*-3-hydroxymyristoyl-ACP.

Mutagenesis has been a key approach to determining the various residues of ACP (see Section 1.3.3) involved in LpxA function. It was shown that mutations in divalent cation binding Site B affected LpxA activity more than those in Site A (Gong et al., 2007). This is somewhat surprising since Site B is quite remote from the active site, while Site A is in proximity to catalytic residues His 125 and Asp 126, i.e. LpxA's catalytic diad. Further mutagenic work from our lab (Gong et al., 2008) involving the substitution of Trp into various positions of ACP to measure the microenvironment in both unfolded and folded *V. harveyi* ACP has reinforced the importance of Helix II in ACP-LpxA

interaction. When A45W ACP was incubated with LpxA, the Trp residue was protected from fluorescence quenching (Gong et al., 2008).

Due to its essential role in the synthesis of lipid A, LpxA is an attractive antibiotic drug target. Furthermore, by inhibiting lipid A (endotoxin) synthesis, the host innate immune system isn't likely to be activated and thus less likely to cause septic shock (Wyckoff et al., 1998b), a complication of treatment with current antibiotics that can lead to a build-up of lipid A (Warren et al., 1992, Parrillo, 1993). Despite this, no LpxA-specific inhibitors have been reported to date.

1.4.3 Bacterial AcpS

As discussed earlier (see Section 1.2 and sub-sections within), all carrier proteins, including ACPs, DCPs and PCPs, must be modified by a phosphopantetheine moiety (PP) to actively carry out their essential roles. This addition is catalyzed by one of the many phosphopantetheinyl transferase (PPTase) family enzymes. PPTases have been shown to be essential for cell viability in each of the three phylogenetic classifications: bacteria, archaea and eukaryotes (Lambalot et al., 1996, Beld et al., 2014).

Despite the discovery and initial characterization of *E. coli* AcpS (Reuter et al., 1999b) almost 50 years ago, it wasn't until the mid-90's that this family of proteins was recognized and received large-scale attention (Lambalot & Walsh, 1995, Lambalot et al., 1996, Beld et al., 2014). Enzymes in the PPTase family can be divided into three subgroups based on sequence identity, three-dimensional structure, and their target (Beld et al., 2014). Type I PPTases generally participate in Type II FAS, PKS and NRPS systems with multiple discrete proteins, while Type II PPTases function in Type I systems that are assembled from one or more large multifunctional polypeptides. The

final group, Type III, exist as C-terminal transferases residing in the megasynthases as one of several catalytic domains in Type I FAS. This family of PPTases modifies the ACP domain prior to the formation of the megasynthase (Beld et al., 2014).

The archetypical Type I PPTase is holo-ACP synthase (AcpS), first discovered in *E. coli* by Alberts and Vagelos (1999b). As AcpS is a major focus of this work, it will be discussed in more detail below. *B. subtilis* Sfp, the prototypical Type II PPTase, is responsible for the addition of PP to the PCP of surfactin synthase (Wakil, 1989, Beld et al., 2014). It exists as a pseudo-homodimer, essentially resembling two AcpS monomers with one active site. *B. subtilis* Sfp exhibits a more relaxed substrate specificity than AcpS, having the ability to modify a broad range of PP acceptors (Wakil, 1989, Beld et al., 2014). Type III PPTases are best represented by the C-terminal PPTase domain in Type I FAS megasynthases, such as those in mammals (Smith et al., 2003) and yeast (Leibundgut et al., 2007, Lomakin et al., 2007) (discussed in Section 1.2.1.1). For a more detailed description of PPTases, the reader is directed towards the recent comprehensive review by Beld et al. (2014).

Although the stringency of carrier protein specificity varies across groups, all PPTase family members share a conserved target sequence, DSL, as can be seen by alignment of ACP and ACP-like carrier proteins (Figure 5; black selection). In more distantly related carrier proteins that are substrates for promiscuous PPTases such as *B. subtilis* Sfp (Wakil, 1989, Beld et al., 2014), the first and third positions of the target sequence can vary, with His being a common replacement for Asp and Met or Ile being a substitute for Lys (Figure 5). Promiscuity in these PPTases leads to their ability to both phosphopantetheinylate a broad range of targets, as well as use a broad range of thiol-

substituted CoA substrates (Issartel et al., 1991). These characteristics have allowed for many biotechnological applications, including PPTase-catalyzed site-specific protein labeling both *in vitro* and *in vivo* (Lambalot & Walsh, 1997, George et al., 2004, Altschul et al., 2005, Jacquier et al., 2006, Meyer et al., 2006). One particularly intriguing application is the use of ACP as a tag for plasma membrane proteins, which can be modified using a PPTase and a thiol-substituted CoA derivative. The most notable example is the tagging of the external domain of G protein coupled receptors in intact cells using fluorescent derivatives of CoA, which can then be monitored using microscopy or FRET analysis. This approach has been successfully employed to study various membrane proteins in *E. coli*, *S. cerevisiae*, HEK293 and TRVb cells and neurons (Vivero-Pol et al., 2005, Yin et al., 2005, Meyer et al., 2006, Prummer et al., 2006, Kropf et al., 2008).

E. coli AcpS, the archetypical and founding member of Type I PPTases, is a relatively small (~120 amino acids; ~14.5 kDa) basic (pI = 9.3) protein (Lambalot & Walsh, 1995, Lambalot et al., 1996) and is highly specific towards ACP. Early reports based on gel filtration analysis indicated that *E. coli* AcpS is a homodimer (Lambalot & Walsh, 1995, Lambalot et al., 1996, Flugel et al., 2000). To date, a x-ray crystal or NMR structure of *E.c.*AcpS has yet to be reported; however, its sequence is very similar to all bacterial AcpS species of known crystal structure (Chirgadze et al., 2000, Parris et al., 2000, Dall'aglio et al., 2011, Halavaty et al., 2012). Table 4 contains a complete list of AcpS structures determined to date along with their PDB ID. All of these form trimeric structures with the active site at the subunit interface (three active sites per multimer).

Table 4 Bacterial AcpS structures determined to date.

Species	Bound to	Structure Resolution (Å)	PDB ID	Identity to *V. fischeri*[1]	Reference
B. subtilis	N/A[2]	1.8	1F7T	44%	(Parris et al., 2000)
	CoA	1.8	1F7L		
	holo-ACP	2.3	1F80		
S. pneumoniae	N/A[2]	2.40	1FTE	38%	(Chirgadze et al., 2000)
	N/A[2]	2.05	1FTF		
	3′5′-ADP	1.90	1FTH		
St. coelicolor	N/A[2]	2.0	2JCA	35%	(Dall'aglio et al., 2011)
	CoA	1.6	2JBZ		
	acetyl-CoA	1.5	2WDO		
St. coelicolor H110A Mutant	CoA	1.3	2WDS	35%	
St. coelicolor D111A Mutant	CoA	1.4	2WDY	35%	
M. tuberculosis	N/A[2]	1.95	3HQJ	30%	(Dym et al., 2009)
B. anthracis (str. Ames)	CoA	2.31	3HYK	44%	(Halavaty et al., 2012)
S. aureus	N/A[2]	1.82	4JM7	40%	
V. cholera (O1 biovar eltor)	N/A[2]	1.85	3QMN	75%	
M. smegmatis	N/A[2]	1.70	3GWM	29%	2010-04-07[3]
M. tuberculosis	SO_4 ion	2.51	3NE1	30%	(Gokulan et al., 2011)
	N/A[2]	1.90	3NE3		
C. ammoiniagenes	N/A[2]	2.5	3NE9	27%	
	CoA	1.89	3NFD		
S. aureus (subsp. aureus COL)	ACP	2.52	4DXE	40%	2012-03-14[3]
M. tuberculosis (Rv2523cE77A)	N/A[2]	1.8	4HC6	30%	2012-1-28[3]

[1] = As determined by Protein BLAST alignment (Altschul et al., 1997, Altschul et al., 2005).
[2] = Structure without substrate bound.
[3] = Structure in PDB Data Bank with out an associated publication; Date of publication to PDB Data Bank given.

Thus, AcpS from *E. coli* and other bacteria, such as *V. fischeri,* likely function as a trimer as well.

Of the AcpS structures determined to date (Table 4; sample structure in Figure 12A), those described by Parris *et al*. (2000) were the first to elucidate the structure of a Type I AcpS in complex with ACP (Figure 12B). This structure (PDB ID = 1F80) not only reinforced the theory that Helix II of ACP is the "recognition helix" (discussed in Section 1.4.1), but also identified residues important for AcpS-ACP binding and allowed proposal of a reaction mechanism. The majority of residues involved in ACP-AcpS binding fall within Helix I of AcpS (Parris et al., 2000) and all are conserved across species (Figure 13). Two key electrostatic interactions occur between Arg 14 of AcpS and Asp 35/38 of ACP, functioning to orient Ser 36 of ACP near the β-phosphate of CoA. Additionally, Ser 36 is in close contact with Glu 58, which functions in coordination of magnesium ions. On the distal portion of ACP's Helix II, Asp 48 interacts with Gln 22 and Arg 24 of AcpS, locking ACP into its correct orientation in the AcpS active site. These electrostatic interactions have been confirmed through further crystallographic (Xu et al., 2001) and mutagenic (Gong et al., 2007) studies. The importance of these interactions is further supported by the finding that Type II ACPs are better substrates for AcpS than the less acidic carrier proteins, such as PCP (Finking et al., 2004).

As mentioned above, a reaction mechanism for AcpS was proposed based on the *B. subtilis* ACP-AcpS crystal structure (Parris et al., 2000). Additionally, a more general PPTase family reaction mechanism has been proposed based on the *B. subtilis* Sfp crystal structure (Reuter et al., 1999a). The two proposed mechanisms (Figure 14) are

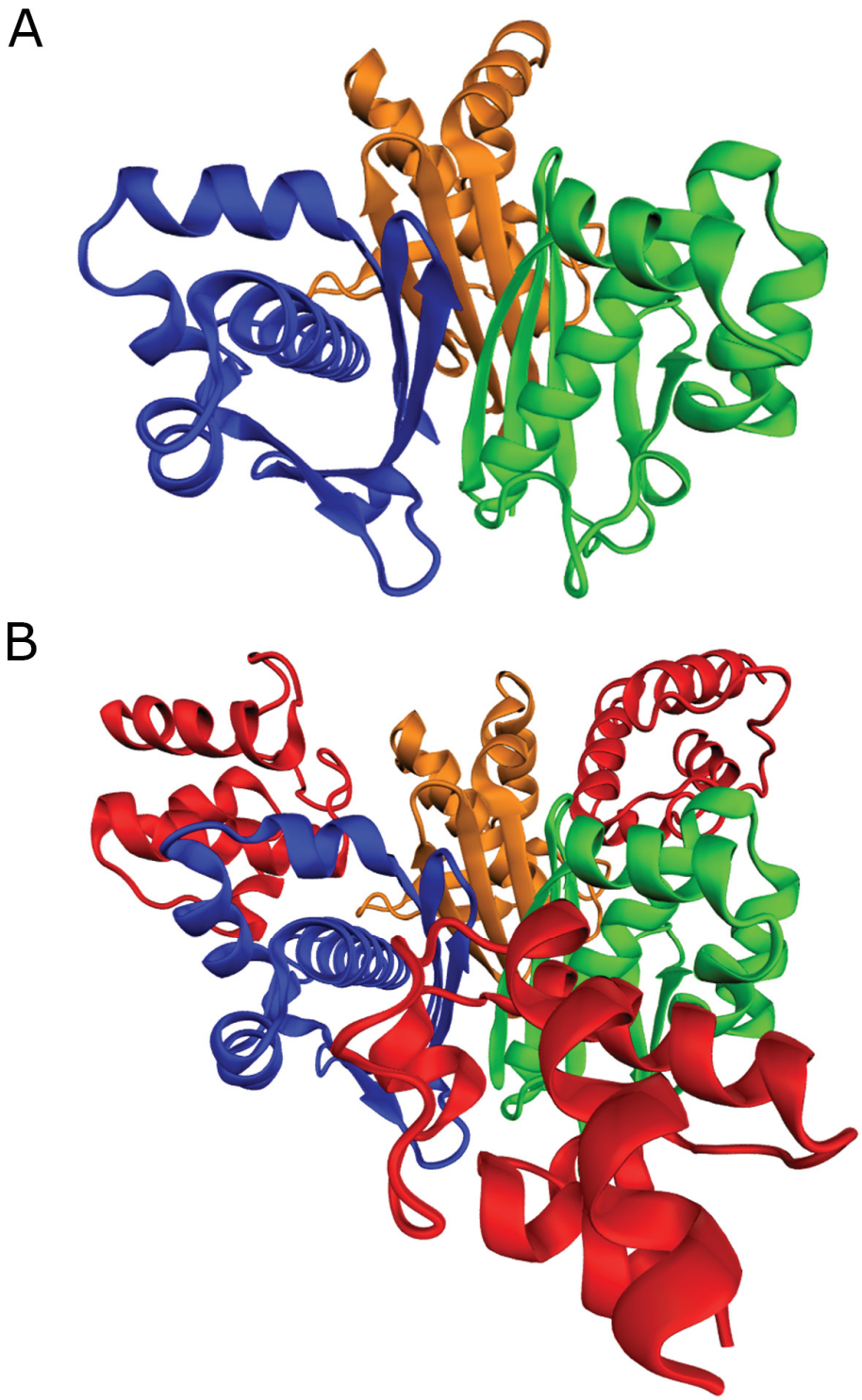

Figure 12 *B. subtilis* AcpS quarternary structure. “NewCartoon” representation of *B.s.*AcpS (A) and *B.s.*AcpS-ACP complex (B). AcpS is colored by subunit: A (blue), B (red) and C (orange) while ACPs are red. This figure was created using PDB ID = 1F7T/1F80 (Parris et al., 2000). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

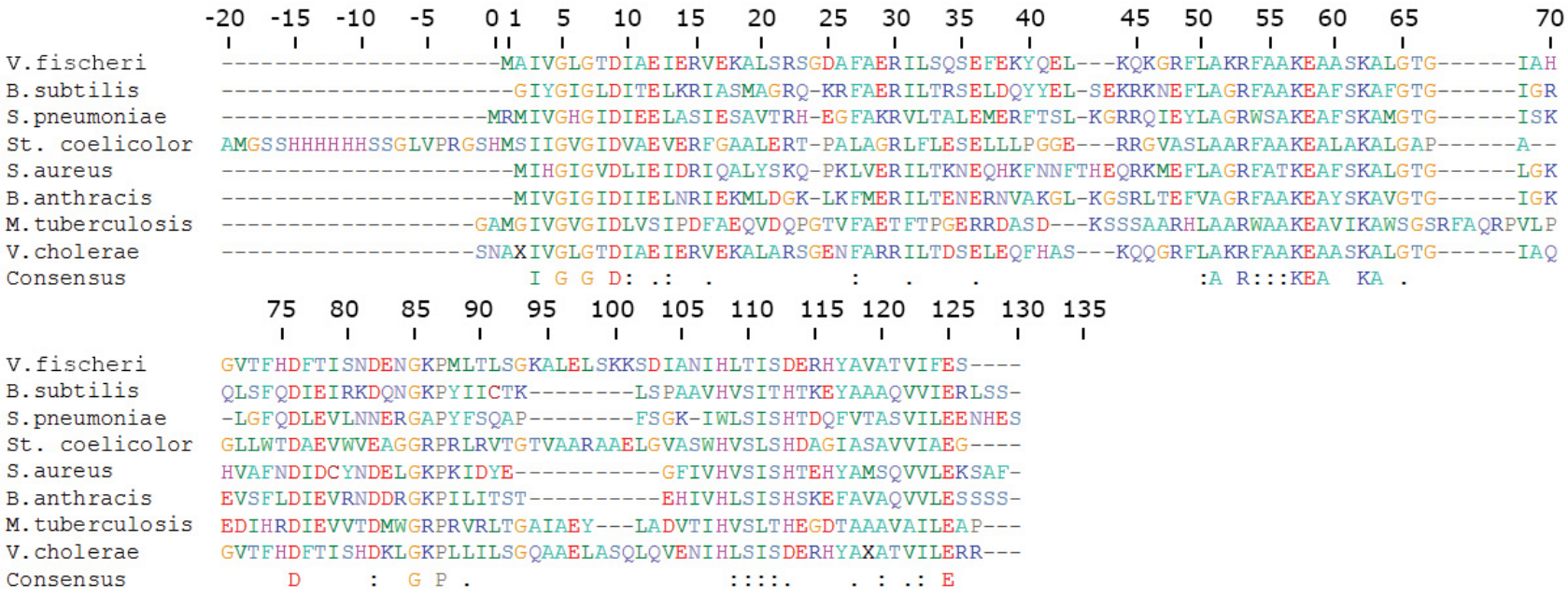

Figure 13 Sequence alignment of AcpS from selected organisms. Clustal Omega was used to produce the above figure. Numbering of residues is based on *V. fischeri* AcpS. Coloring is based on amino acid properties. Consensus sequence shown: amino acid given for fully conserved residue, a colon (:) indicates conservation between groups of strongly similar properties and a period (.) indicates conservation between groups of weakly similar properties as determined by Clustal Omega.

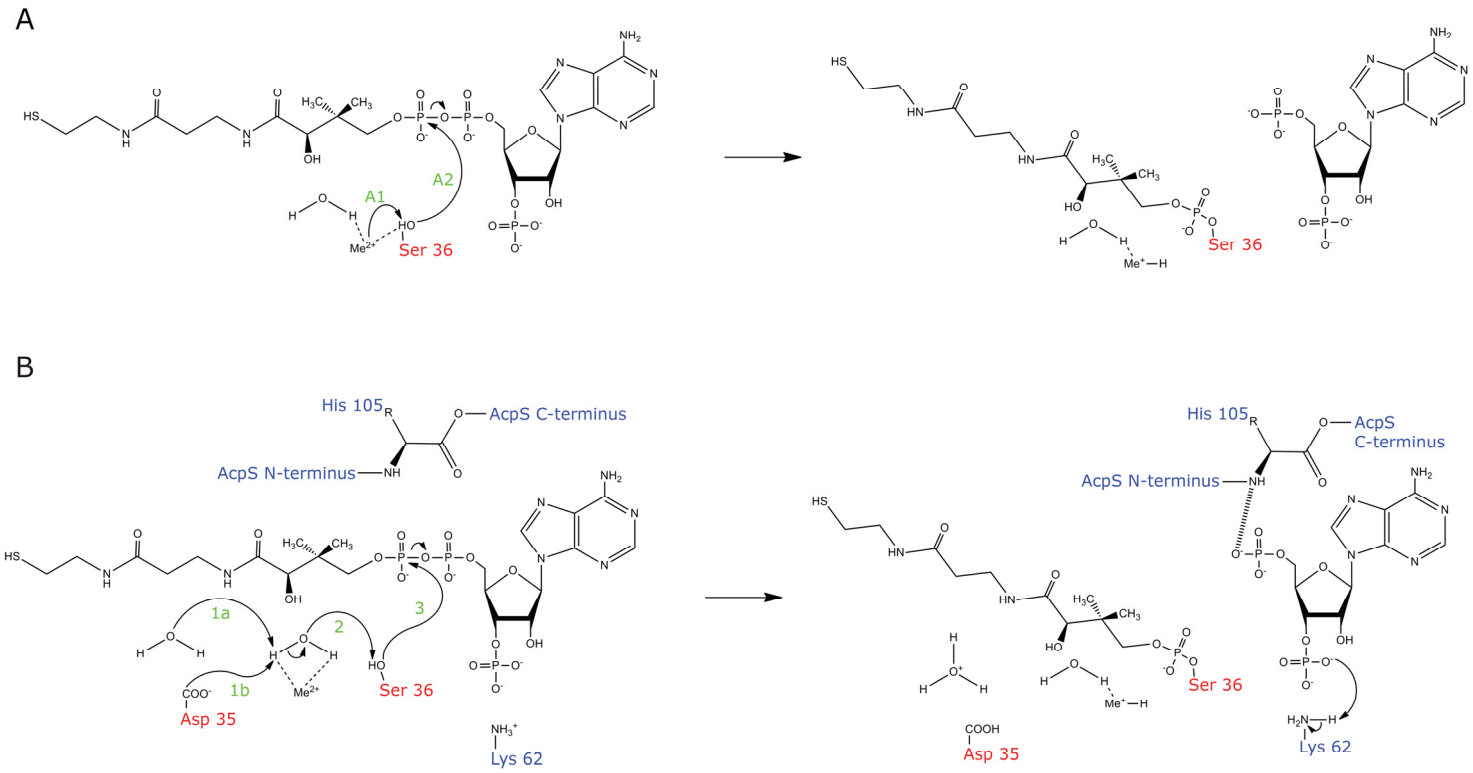

Figure 14 Proposed mechanisms for PPTase reactions. Two mechanisms have been proposed for PPTases based on crystal structures. In the first mechanism (panel A; based on the structure of *B. subtilis* Sfp (Reuter et al., 1999a)), coordination of the metal ion (typically magnesium) by Ser 36 causes deprotonation (arrow A1), allowing for a nucleophilic attack on the β-phosphate of CoA (arrow A2). The second mechanism (panel B; based on the structure of the *B. subtilis* AcpS-ACP complex (Parris et al., 2000)) proposes that a water takes place of Ser 36 in the coordination of the metal ion. This water is deprotonated by either bulk solvent (arrow 1a) or Asp 35 (arrow 1b). This activated water then deprotonates Ser 36 (arrow 2), which then attacks the β-phosphate of CoA (arrow 3). The negatively charged 3′,5′-ADP is stabilized and protonated by the backbone amide of His 105 and the side chain amide of Lys 62, respectively. Amino acids colored by protein: AcpS (blue) and ACP (red). Arrow labels are green. This figure was generated using ChemDraw 12.0.

remarkably similar despite the structural differences between Type I (AcpS-like) and Type II (Sfp-like) PPTases (discussed above). For AcpS, a metal-coordinated water is deprotonated by either bulk solvent or Asp 35 of ACP, which in turns deprotonates Ser 36 of ACP, allowing for nucleophilic attack on the β-phosphate of CoA, resulting in holo-ACP and 3′,5′-ADP (Figure 14B). The other mechanism varies only by the method of deprotonation of Ser 36, which is deprotonated by coordination with the metal ion (Figure 14A). Further studies will be required to establish whether these mechanistic variations typify the Type I and II classes of PPTase.

Initial characterization of *E.c.*AcpS kinetics was completed over 40 years ago by Elovson and Vagelos (1968); this determined the K_M for ACP and CoA to be 400 nM and 150 μM, respectively. The former value is relatively low compared to the typical K_M range for other ACP-dependent enzymes, which tend to be in the low micromolar range (Byers & Gong, 2007). One caveat to the determination of the K_M of ACP for *E.c.*AcpS, and likely for other bacterial AcpSs, is the existence of ACP substrate inhibition. This inhibition exists even at low ACP concentrations, precluding an accurate determination of the K_M value (Elovson & Vagelos, 1968). Further characterization of *E. coli* AcpS revealed severe substrate inhibition of the enzyme by apo-ACP at concentrations >5 μM (Flugel et al., 2000); similar results were determined for *M. pneumoniae* and *S. pneumoniae* enzymes (McAllister et al., 2006). Based on the aforementioned crystal structure of the *B. subtilis* ACP-AcpS, Parris et al. (2000) suggested a possible mechanism for this inhibition. They proposed that CoA binds AcpS first, initiating the formation of a tighter protein assembly and changing the electrostatic potential of the ACP binding site, allowing ACP to bind and the reaction to occur; holo-ACP dissociates,

followed by 3′,5′-ADP. If ACP were to bind first (e.g. at high apo-ACP concentrations), CoA would be unable to bind the complex (Parris et al., 2000). Furthermore, at higher apo-ACP concentrations, it is plausible that upon the release of holo-ACP but before the release of 3′,5′-ADP, apo-ACP could bind, trapping the 3′,5′-ADP in the active site and producing an unproductive complex (Parris et al., 2000).

Due to the essential and ubiquitous role of Type I AcpS enzymes in bacteria and their fundamental difference from *H. sapiens* Type II PPTases, bacterial AcpS is an excellent target for the development of novel antibacterial drugs. As such, our lab started an ongoing collaboration with DeNovaMed Inc. to develop antimicrobial AcpS inhibitors using structure-aided drug design.

Recently, our interest has extended to the malaria parasite *Plasmodium falciparum*, which is somewhat unusual among eukaryotic protist genomes in encoding a single AcpS-like PPTase (Cai et al 2005). *P. falciparum* AcpS (*P.f.*$AcpS_{FL}$), unlike bacterial AcpS, is comprised of two domains: the N-terminal (*P.f.*AcpS-N) and C-terminal domain (*P.f.*AcpS-C). *P.f.*AcpS-N possesses unknown function, but has been shown to share sequence identity with the metal-dependent phosphohydrolase enzyme family (Cai et al., 2005), including *E. coli* ACP phosphodiesterase (AcpH) (Thomas et al., 2007). *P.f.*AcpS-C shares sequence identity (30 – 50%) with bacterial AcpS enzymes. In addition to these two domains, *P.f.*$AcpS_{FL}$ contains an N-terminal targeting sequence that is known to target the protein, as well as *P. falciparum* ACP, to the apicoplast of the protist (Cai et al., 2005). The apicoplast organelle is surrounded by a double membrane (two sets of both inner and outer membranes) and is thought to be derived by two endosymbiotic events (Kalanon & McFadden, 2010).

Our laboratory has recently reported, for the first time to our knowledge, the cloning and expression of *P.f.*AcpS-C, hereafter referred to as *P.f.*AcpS. Additionally, we have demonstrated that this C-terminal domain exhibits holo-ACP synthase activity and that this activity can be blocked by a subset of proprietary DeNovaMed AcpS inhibitors (Tami et al., 2011). In attempts to characterize the native full length protein, *P.f.*$AcpS_{FL}$ was also cloned and expressed, but was not soluble in the expression system used (G. Tami and D. Byers, personal communication).

1.5 Thesis Rationale

Clearly, ACP is an interesting and important protein with respect to its unique structural features, multiple roles and enzyme partners, and applications to both medicine and biotechnology. Within this broader context, my work has two central objectives: first, exploring the effect of constraining the physical structure of ACP on its conformational stability and function; this is described in Chapter 3. The second goal is to expand the use of fluorescence methods in the experimental analysis of ACP conformation and partner interactions (described in Chapters 4 and 5).

The first goal was achieved using split intein-mediated fusion of the N- and C-termini of ACP to form a cyclic derivative that could be studied *in vitro* and *in vivo*. Inteins, first discovered in the early 1990s independently by two labs (Hirata et al., 1990, Kane et al., 1990), are special internal domains within proteins that can spontaneously fold and remove themselves through catalytic self-splicing, resulting in the joining of flanking regions (exteins) through peptide bond formation (Volkmann & Iwai, 2010, Volkmann & Mootz, 2013). Thus, they are analogous to RNA splicing of introns. Three types of inteins have been found in nature to date: bi-functional inteins, mini-inteins and

split-inteins (Volkmann, 2009). Split-inteins, employed in this work to create a cyclic ACP, are mini-inteins where the intein domain is divided into two separate protein fragments (I_N and I_C). These can interact with high affinity to form a catalytically active intein, resulting in self-splicing and the joining of any adjacent fusion partners to which they are attached (Volkmann & Iwai, 2010).

Chapter 4 exploits the possibility of using fluorescence of the lone intrinsic Tyr (Tyr 71) found in *V. harveyi* ACP to monitor conformational transitions in previously-made mutants affecting ACP structure and function; the SA/SB series of divalent cation-binding mutants are used to test this approach. Chapter 5 describes site-specific mutagenic introduction of Trp probes into two enzymes that lack endogenous Trp, *E.c.*LpxA and *V.f.*AcpS, in order to study their interaction with ACP and co-substrates.

CHAPTER 2 MATERIALS AND METHODS

All concentrations given are final concentrations unless otherwise specified. All bacterial strains created or used in this study are shown in Table 5. All primers were purchased from Integrated DNA Technologies (IDT) and were PAGE purified. Complete primer sequences and names are given in Table 6. A description of all plasmids created and used in this work can be found in Table 7. The numbering of all residues in this thesis corresponds to that of the wild-type protein, unless otherwise specified.

2.1 Plasmid Construction and Site-Directed Mutagenesis

2.1.1 Cyclic and Linear ACPs

Polymerase chain reaction (PCR) was carried out using Phusion DNA Polymerase (Finnzymes) and commercially synthesized primers. The cyclic (cycL46W) and linear (linL46W) L46W ACPs were constructed and cloned in the pJET1.2/blunt cloning vector (Fermentas). DNA encoding cyclic and linear F50A (cycF50A, linF50A, respectively) were created using inverse PCR and directly circularized using blunt-end ligation. Additionally, a control with mutant inteins (unable to react catalytically) was created (preL46W^{mut}). All construction and cloning of cyclic and linear ACPs was completed by Gerrit Volkmann (Volkmann, 2009, Volkmann et al., 2010).

2.1.2 Trp-Substituted *Escherichia coli* LpxA Mutants

All Trp substitutions were performed using the QuikChange® II XL Site-Directed

Table 5 Names, genotypes and other relevant information for strains used in this study.

Strain	Genotype/Relevant Features	Description	Source
BL21(DE3)pLysS	*ompT* *dcm*$^+$ *pLysS* λ(DE3)	Derivative of *E. coli* B strain. These cells (pLysS) provide tighter control of protein expression for expression of toxic proteins; addition of the pLysS plasmid adds chloramphenicol resistance.	(Studier & Moffatt, 1986)
CY1861	Δ*acpP::cat* pACYC/*acp*P	Derived from strain DY329. Strain's endogenous *acp*P gene replaced by chloramphenicol resistance gene; arabinose inducible plasmid carries synthetic *acp*P gene and imparts chloramphenicol resistance. Strain useful for testing various ACP constructs *in vivo*.	(Volkmann et al., 2010)
BL21-CodonPlus (DE3)-RIL	*ompT* *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*argU ileY leuW* Cam^r]	Derivative of *E. coli* B strain. Contains extra copies of the *argU*, *ileY,* and *leuW* tRNA genes. These genes encode tRNAs that recognize the arginine codons AGA and AGG, the isoleucine codon AUA, and the leucine codon CUA, respectively. The addition of these tRNAs allows expression of proteins from organisms that have AT-rich genomes. Cells are resistant to chloramphenicol.	[a]
XL10-Gold® ultracompetent cells	Tet^r Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173* *endA1* *recA1 gyrA96* *relA1* *lac* Hte [F′ *proAB* *lacI*qZΔ*M15* Tn*10* (Tet^r) Amy Cam^r]	Derivative of Stratagene's XL2-Blue MRF′ cell line. Contains the Hte phenotype, which increases transformation efficiency of ligated DNA. Cells are endonuclease (*endA1*) and recombination deficient (*recA*) which greatly improves the quality of plasmid preps and insert stability, respectively.	[a]

[a] Stratagene; 11011 North Torrey Pines Road, La Jolla, CA, USA, 92037.

Table 6 Primer names and sequences used in this study. Restriction enzyme (underlined) and mutation (bold) sites indicated.

Primer	**Sequence (5′ → 3′)**
IC-*for*	GGGAGCTCATGGAAGCAGTATTAAATTACAATCAC
IC-*rev*	GACTAATGTATCTCCAGAAAAACAGGAAGAGCATATGCTAGCGCTGTTATGGACAAACAC
ICMUT-r*ev*	GACTAATGTATCTCCAGAAAAAGCGGAAGAGCATATGCTAGCGCTTGCATGGACAAACAC
IN-*for*	GTGTTTGTCCATAACAGCGCTAGCATATGCTCTTCCTGTTTTTCTGGAGATACATTAGTC
INMUT-*for*	GTGTTTGTCCATGCAAGCGCTAGCATATGCTCTTCCGCTTTTTCTGGAGATACATTAGTC
INHIS-*rev*	GGCTGCAGTTAATGGTGATGGTGATGGTGATGACCAGAATCTTCC
ACP-*for*	GGGCTAGCAACATCGAAGAACGCGTAAAGAAAATC
ACP-*rev*	GGGGCTCTTCTACAACCCTGAGCGCTGTTTACG
GSTC-*for*	GGGTGGCCATCATACGTTATATAGCTGACAAGC
GSTC-r*ev*	CGTTCTTCGATGTTGCTAGCAGAACGACCTTCGATCAGATCC
LINL46W-*for*	GGATCTGATCGAAGGTCGTTCTGCTAGCAACATCGAAGAACG
LINL46W-*for2*	CCCATATGTCTGCTAGCAACATCGAAGAACGCG
LINL46W-*rev*	GGGAGCGCTTCATTAACCCTGAGCAGAGTTTACGTAGTCG
LINL46W-*rev2*	CCCTGCAGTCATTAACCCTGAGCAGAGTTTACG
LINF50A-*for*	GAGGCTGACACTGAGATTCCTGATGAAG
LINF50A-*rev*	CTCTTCCAGAGCCATTACTAGCTCTACAGTG
LpxA-Q104W- *for*	GTCACCATTCATCGTGGCACAGTC**TGG**GGCGGTGGATTGACG
LpxA-Q104W- *rev*	TCCTTCGTCAATCCACCGCC**CCA**GACTGTGCCACGATGAATG
LpxA-F162W- *for*	GCAGTCCATCAG**TGG**TGCATCATTGGTGCG
LpxA-F162W- *rev*	CGCACCAATGATGCA**CCA**CTGATGGACTGC
LpxA-G173W- *for*	GCGCACGTGATGGTTGGC**TGG**TGCTCCGGTGTGGCGC
LpxA- G173W- *rev*	CTGCGCCACACCGGAGCA**CCA**GCCAACCATCACGTGCGC
LpxA-S208W- *for*	CTGAAGCGCCGCGGATTC**TGG**CGTGAGGCGAT
LpxA-S208W- *rev*	GGTAATCGCCTCACG**CCA**GAATCCGCGGCGCTT
AcpS-R22	GGAACGTGAAAGGGCTTTTTCAACCC
AcpS-R22E	GGA**TTC**TGAAAGGGCTTTTTCAACCC
AcpS-F27	GGAGATGCGTTTGCAGAGCGCATTTTAAGTCAG
AcpS-F27W	GGAGATGCG**TGG**GCAGAGCGCATTTTAAGTCAG

Table 7 Plasmids (and relevant information) used in this study.

Plasmid	Relevant Features/Description	Source
$pTI_C(NS)I_NH$	Permuted *Synechocystis ssp*. PCC6803 *Gyr*B split-intein sequence cloned into *Sac*I/*Pst*I sites of pT; contains *Nhe*I and *Sap*I sites between I_C and I_N sequences, respectively, for insertion of target protein sequence for *in vivo* cyclization	*a*
pTCYC-L46W	*V. harveyi* ACP L46W mutant cloned into $pTI_C(NS)I_NH$ via *Nhe*I and *Sap*I.	*a*
pTPRE-L46W-mut	Cyclization-deficient control construct of pTCYCL46W. Mutation in I_C (C1A) and in I_N (N435A) inactivate the trans-splicing reaction	*a*
pTCYC-F50A	As pTCYC-L46W, but with F50A mutation in place of L46W	*a*
pGEX-linL46W	*V. harveyi* ACP mutant derivatives pGEX-5X-3 derived plasmids: tac promoter N-terminal (GST) tag Fx_A cleavage site to remove tag Ampicillin resistance	*b*
pGEX-F50A		*c*
pET23b-*P.f.*ACP-A104W	*P. falciparum* ACP mutant derivatives pET23 (b) + derived plasmids: T7 promoter C-terminal hexahistidine (His_6) tag Ampicillin resistance	*d*
pET23b-*P.f.*ACP-L105W		*d*
pGEX-rACP	*V. harveyi* ACP (WT and mutant derivatives) pGEX-5X-3 derived plasmids: tac promoter N-terminal (GST) tag Fx_A cleavage site to remove tag Ampicillin resistance	*e*
pGEX-SA		*e*
pGEX-SB		*e*
pGEX-SASB		*e*

Plasmid	Relevant Features/Description	Source
pGEX-D30N	As for pGEX-rACP	[e]
pGEX-D56N		[e]
pET23b-LpxA	*E. coli* LpxA (WT and mutant derivatives) pET23 (b) + derived plasmids: T7 promoter C-terminal hexahistidine (His_6) tag Ampicillin resistance	[e]
pET23b-LpxA-Q104W		[f]
pET23b-LpxA-F162W		[f]
pET23b-LpxA-G173W		[f]
pET23b-LpxA-S208W		[f]
pET23b-AcpS	*V. fischeri* AcpS (WT and mutant derivatives) pET23 (b) + derived plasmids: T7 promoter C-terminal hexahistidine (His_6) tag Ampicillin resistance	[b]
pET23b-AcpS-F27W		[f]
pET23b-AcpS-R22E		[f]
pET23b-AcpS-R22E/F27W		[f]

[a] Volkmann, G., Murphy, P.W., Rowland, E.E., Cronan, J.E., Jr., Liu, X.Q., Blouin, C., and Byers, D.M., (2010) *J. Biol. Chem.* 285:8605-8614.
[b] Gong, H., Murphy, P. W., Langille, G. M., Minielly, S. J., Murphy, A., McMaster, C. R., and Byers, D. M. (2008) *Biochim Biophys Acta* 1784, 1835-1843.
[c] Flaman, A.S., Chen, J.M., Van Iderstine, S.C., and Byers, D.M. (2001) *J. Biol. Chem.* 276, 35934-35939.
[d] Marissa LeBlanc, former graduate student; C.R. McMaster Lab, Dalhousie University.
[e] Gong, H., Murphy, A., McMaster, C.R., Byers, D.M. (2007) *J. Biol. Chem.* 282, 4494-4503.
[f] This work

Mutagenesis kit (Stratagene) as directed by the manufacturer (Stratagene. *QuikChange® II XL Site-Directed Mutagenesis kit manual*.). See Table 6/Table 8 for primer sequences/combinations used for each mutation and Figure 15 for the PCR conditions.

2.1.3 Trp-Substituted *Vibrio fischeri* AcpS Mutants

Trp substituted *V. fischeri* AcpS (F27W) was created using inverse PCR (I-PCR) using a wild-type *V. fischeri* AcpS (*V.f.*AcpS) plasmid kindly provided by Annette Henneberry (DeNovaMed Inc.). With the I-PCR method used (Volkmann, 2009), phosphorylated primers are necessary. The phosphorylation reaction took place in a 20 µL total volume, containing Buffer A (1X; Fermentas), ATP (1 mM; Sigma), Primer (20 µM; IDT; AcpS-R22 or AcpS-F27W), and T4 Polynucleotide Kinase (2.5% (v/v); Thermo Scientific). Reactions (37°C, 30 min) yielded 20 µM phosphorylated primer that was used directly (without purification) in the I-PCR reaction.

The I-PCR reaction occurred in a total volume of 20 µL, containing HF Buffer (1X; Finnzymes, distributed by NEB), dNTPs (0.2 mM; Thermo Scientific), phosphorylated Primer1 (2.5% (v/v)), phosphorylated Primer2 (2.5% (v/v)), plasmid DNA (from QIAprep® Spin Miniprep Kit (Qiagen); diluted 1:50 in water prior to use; 5.0% (v/v) in reaction mixture), and Phusion® High-Fidelity DNA Polymerase (1.0% (v/v); Finnzymes, distributed by NEB). In addition to the F27W mutation, a charge change mutant was created (R22E) as well as its Trp-containing double mutant (R22E/F27W). See Table 6/Table 9 for primer sequences/combinations used for each mutation and Figure 16 for the I-PCR temperature profile. As with the F27W mutation primers, primers for charge-change mutants were phosphorylated prior to the I-PCR reaction.

Table 8 LpxA Trp-substitutions and the primer combinations used for PCR-based mutagenesis.

LpxA	Template	Primers
Q104W	Wild-Type	LpxA-Q104W- *for* LpxA-Q104W- *rev*
F162W	Wild-Type	LpxA-F162W- *for* LpxA-F162W- *rev*
G173W	Wild-Type	LpxA-G173W- *for* LpxA- G173W- *rev*
S208W	Wild-Type	LpxA-S208W- *for* LpxA-S208W- *rev*

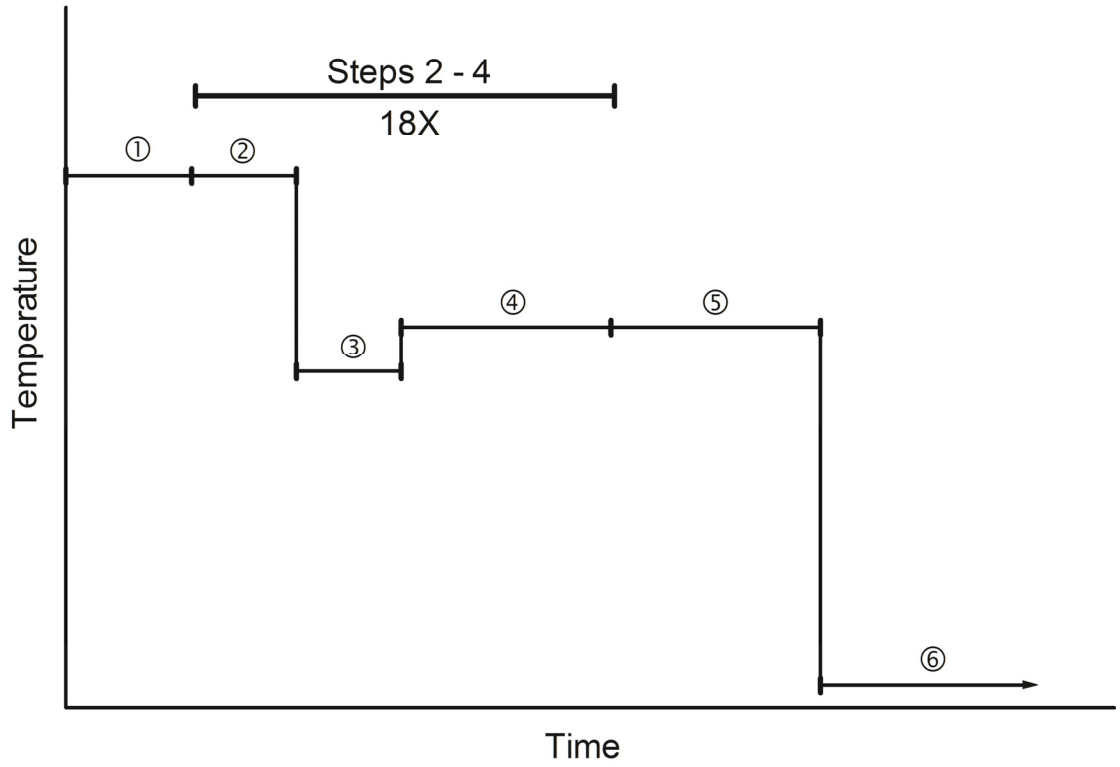

Figure 15 Temperature profile for LpxA PCR mutation reactions using the QuikChange® II XL Site-Directed Mutagenesis kit. Step time and temperatures: ①95°C, 60 s, ②95°C, 50 s, ③60°C, 50 s (temperature dependent on TM of Primers), ④68°C, 360 s (1 min/kbp), ⑤68°C, 7 min, ⑥4°C, ∞.

Table 9 AcpS mutations and the primer combinations used for I-PCR.

AcpS	Template	Primer1	Primer2
F27W	Wild-Type	AcpS-R22	AcpS-F27W
R22E	Wild-Type	AcpS-R22E	AcpS-F27
R22E/F27W	Wild-Type	AcpS-R22E	AcpS-F27W

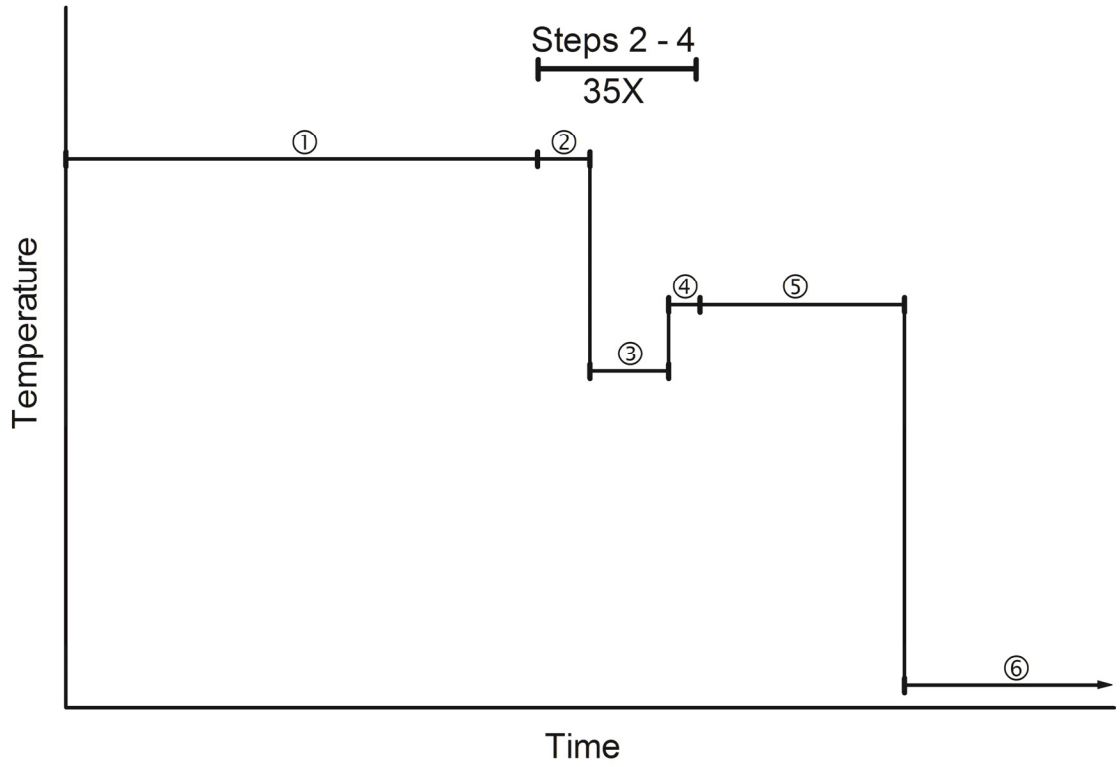

Figure 16 Temperature profile for AcpS mutations using I-PCR. Step time and temperatures: ①98°C, 90 s, ②98°C, 10 s, ③60°C, 15 s (temperature dependent on T_M of Primers), ④72°C, 6 s (1 s/kbp), ⑤72°C, 6 min, ⑥4°C, ∞.

2.2 *In Vivo* Characterization of Cyclic ACP

2.2.1 Confirmation of Protein Cyclization *In Vivo*

E. coli cells (BL21 (DES) pLysS) harbouring either pTCYC-L46W or pTPRE-L46W-mut were grown (37°C) in LB medium containing ampicillin (50 μg/mL) and chloramphenicol (34 μg/mL) to mid-log phase (OD_{595} = 0.5 – 0.6) and induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 0.8 mM). Induced cells were then grown for an additional 18 h at room temperature. Following the induction period, cells were harvested by centrifugation (7,700*g*, 10 min), lysed in reducing SDS-PAGE sample buffer (62.5 mM Tris-HCl (pH 6.8), 6.25% glycerol, 0.625% SDS, 0.025% bromophenol blue, 250 mM β-mercaptoethanol) and boiled for 5 min. To determine the extent of cyclization, the prepared total cell lysates were subjected to SDS-PAGE (12.5% NEXT gel (Mandel Scientific)), 4% Laemmli stacking gel; 160 V, 1 h) followed by either western transfer to a polyvinylidene difluoride membrane or staining with Coomassie Brillant Blue G-250 (Thermo Scientific).

2.2.2 Construction of the *acp*P Deletion Strain CY1861

Strain CY1861, carrying plasmid pCY765 (containing a synthetic *acp*P gene controlled by an arabinose inducible/glucose repressible promoter), was created by Dr. John E. Cronan Jr. (University of Illinois). Plasmid pCY765 was created from plasmids pNRD25 (De Lay & Cronan, 2006), pBAD30 (Guzman et al., 1995), and p34S-Sm2 (Dennis & Zylstra, 1998). Briefly, the *acp*P gene was excised from pNRD25 using restriction enzymes (BamHI and FspI) and ligated into pBAD30. To add the spectinomycin resistance gene to the construct, the resultant plasmid was digested with FspI (inactivating the pBAD30 β-lactamase gene) and ligated to a fragment encoding

spectinomycin resistance (obtained from p34S-Sm2) resulting in pCY765. To create the CY1861 strain, host strain (CY1878) was created from strain DY329 (Yu et al., 2000) by two steps of transduction with P1vir lysates. The first step selected for growth in the absence of biotin/nicotinic acid and no growth at 42ºC. One of the transductants was transformed with pCY765. In the second step chloramphenicol resistance was added and screened for resistance to chloramphenicol/ spectinomycin and arabinose-dependent growth. This resulted in strain CY1861 (Schmid & Ohlrogge, 2002, Volkmann, 2009, Volkmann et al., 2010). Strain CY1861 carries a synthetic *acp*P gene under control of an arabinose-inducible/glucose-repressible promoter.

2.2.3 *In Vivo* Growth Complementation Assays

Single colonies of CY1861 with the plasmid construct to be tested were resuspended in 50 µL of LB medium containing spectinomycin (50 µg/mL) and ampicillin (100 µg/mL) (LB-SA). Five microliters of this resuspension were used to inoculate 245 µL of LB-SA additionally supplemented with either 0.2% arabinose or 0.2% glucose/0.8 mM IPTG in a 96-well plate. The optical density at 595 nm (OD_{595}) was recorded (t = 0), and the plate incubated (37°C and 400 rpm) with measurements of OD_{595} taken every 20 – 30 min. All absorbance values were corrected for initial absorbance (t = 0) for each sample and scaled by 4 to account for culture size. IPTG had no effect on growth in 0.2% arabinose, indicating that none of the constructs are toxic. *In vivo* complementation assays were performed by Gerrit Volkmann (Volkmann, 2009, Volkmann et al., 2010).

2.3 Protein Expression and Purification

GST-tagged proteins were purified using glutathione-Sepharose 4B (GE LifeSciences). Factor X_a was used to cleave the GST tag (Hematological Technologies Inc.; 16 µg/mg of fusion protein). All anion exchange chromatography was performed with a Source15Q column (GE Healthcare, 6 mL packed bed volume) using a Waters 650 Advanced Protein Purification System in Buffer A (10 mM MES (pH 6.0), 2 mM DTT). Note that Buffer A used here differs from Buffer A used for molecular biology (see Section 2.1.3).

His-tagged protein cultures were lysed with B-PER Reagent (Thermo Scientific) and purified using TALON® Metal Affinity Resin (Clontech) unless otherwise specified. Gel filtration was performed on a Waters 650 Advanced Protein Purification System using a Superose 6 10/300 GL size exclusion column (GE Healthcare).

After purification, the concentration of all proteins was determined using the µBCA protein assay (Thermo Scientific) following manufacturer's instructions. All proteins were expressed in *E. coli* BL21 (DE3) pLysS cells unless otherwise specified. Information on specific strains used in this work can be found in Table 5.

2.3.1 Linear L46W (linL46W) and Linear F50A (linF50A)

Proteins for *in vitro* analysis were produced by growing cells containing either pGEX-linL46W (Gong et al., 2008) or pGEX-F50A (Flaman et al., 2001) plasmids in LB media (100 µg/mL ampicillin and 34 µg/mL chloramphenicol) to mid-log phase (OD_{660} = 0.6 – 0.8) at 37°C before inducing plasmid expression with IPTG (1 mM) and allowing the cells to grow at 30°C to an OD_{660} of 1.5–2.0. Cells harvested by centrifugation were resuspended in 1X PBS (4 mL per 500 mL original culture volume), and lysed by

mechanical force by extruding through progressively smaller bore needles prior to sonication (6 – 8 pulses, 30 s each). Mechanical lysis was utilized in addition to sonication due to the thickness of re-suspended pellets. The lysate was incubated (30 min) with Triton X-100 (SurfactAmps; Thermo Scientific) prior to centrifugation (20,000*g*, 20 min, 4°C). The supernatant was then incubated with glutathione-Sepharose 4B (30 min, room temperature). After washing to remove unbound protein, GST-tagged protein (linL46W or linF50A) was eluted with reduced glutathione (Sigma; 10 mM in 50 mM Tris-HCl (pH 8.0); 10 mL used per 500 mL original culture volume). To the eluate, NaCl (100 mM) and $CaCl_2$ (1 mM) were added prior to incubation with Factor X_a (48 hr, 4°C). Following incubation, the solution was heated (5 min, 95°C) and subjected to centrifugation (13,000*g*, 10 min, room temperature) to pellet precipitated proteins (GST and Factor X_a), while ACP remained soluble. The supernatant, containing ACP, was decanted followed by dialysis into Buffer A and filtered (MILLEX®GV 0.22 μM; Millipore) prior to anion exchange chromatography using a Source15Q anion exchange column. Bound protein was eluted using a linear gradient (1 mL/min flow rate) from 0 – 1 M NaCl in Buffer A. SDS-PAGE (15%, 160 V, 60 min) and native PAGE (20%, 160 V, 90 min) (Flaman et al., 2001) were used to identify fractions containing apo and/or holo-ACPs following gel staining with GelCode Blue (Thermo Scientific).

2.3.2 Cyclic L46W (cycL46W) and Cyclic F50A (cycF50A)

To obtain cyclized ACP, cells harbouring either pTCYC-L46W or pTCYC-F50A were grown, induced, harvested, lysed, and resuspended in Buffer A as for linL46W/linF50A (see 2.3.1) with a few changes. Following centrifugation (20,000*g*, 20 min, 4°C) the soluble protein fraction was decanted and heated (10 min, 95°C), cooled

(10 min, ice), and subjected to centrifugation (4°C, 14,000*g*, 10 min). The supernatant, containing heat-stabile ACP, was filtered (0.22 μM) and injected on a Superose 6 10/300 GL size exclusion column (GE Healthcare) equilibrated with Buffer A. Fractions containing cyclic ACP (cycL46W or cycF50A) were pooled (based on gel analysis) and further purified using anion exchange chromatography as with linL46W/linF50A (see 2.3.1). Cyclic ACP prepared in this manner was mainly in its apo form (no phosphopantetheine moiety) and was compared with the corresponding apo form of linear ACP.

2.3.3 *Plasmodium falciparum* Trp-Substituted ACPs

The *P. falciparum* ACP (*P.f.*ACP) gene (without the leader sequence that encodes the putative N-terminal signal and translocation sequence (amino acids 1 to 60)) was cloned into pET28b+ expression vector and was gifted to our lab from Dr. Remo Perozzo (School of Pharmaceutical Services, Geneva 4, Switzerland). The *P.f.*ACP gene was sub-cloned into pET23b+ for mutation. Sub-cloning of wild-type *P.f.*ACP was completed by Helen Chuang, a former honours student in our lab.

Mutagenesis of wild-type *P.f.*ACP was performed by Marissa LeBlanc using the QuikChange® II XL Site-Directed Mutagenesis kit (Stratagene) as directed by the manufacturer (Stratagene. *QuikChange® II XL Site-Directed Mutagenesis kit manual*.). Two Trp substitutions were made independently, A104W and L105W. Numbering used refers to the natively expressed ACP (with leader sequence). A104W and L105W mutations correspond to A45W and L46W mutations in *V. harveyi* rACP as described in Section 3.3.3.2.

E. coli BL21-CodonPlus (DE3)-RIL (Stratagene) harbouring either pET23b-*P.f.*ACP-A104W (A104W mutant) or pET23b-*P.f.*ACP-L105W (L105W mutant) were grown at 37°C in LB medium containing kanamycin (30 μg/mL) and chloramphenicol (34 μg/mL) to mid-log phase (OD_{660} = ~0.6). Plasmid protein expression was induced with IPTG (0.4 mM) and cells were returned to 37ºC to grow to an OD_{660} of 1.5–2.0. Cells were then collected by centrifugation (4°C, 3,000*g*, 30 min), and pellets resuspended (2 mL per 50 mL original culture volume) in HEPES (50 mM, pH 8.0) containing NaCl (500 mM), imidazole (20 mM), glycerol (10% v/v), DTT (2 mM) and Complete Protease Inhibitor tablets (1X, Boehringer-Mannheim). Cells were lysed by sonication (3 pulses, 10 s each). A104W or L104W were then purified from lysates using His Spin Trap mini-columns (GE Health care) or Talon Resin (Clontech). *P.f.*ACP mutants were eluted with 0.6 – 1 ml of elution buffer (50 mM HEPES pH 8.0, 500 mM NaCl, 500 mM imidazole, 20% glycerol and 2 mM DTT). The purified proteins were aliquoted and stored at -70°C.

2.3.4 Wild-type and Mutant ACPs

Wild-type recombinant (pGEX-rACP (Flaman et al., 2001)) and mutant ACPs (pGEX-SA, pGEX-SB, pGEX-SASB, pGEX-D30N, pGEX-D56N (Gong et al., 2007)) were purified as in Section 2.3.1 with the exception that cells were lysed in 1X PBS (8 mL per 2 L original culture) and without mechanical lysis. As with linL46W and linF50A, this method results in more of the apo-form of the protein being produced. To obtain holo-ACPs, coenzyme A (1 mM), $MgCl_2$ (10 mM), DTT (5 mM), and purified AcpS (500 – 800 μg) were added prior to addition of the supernatant to glutathione-Sepharose. This mixture was then incubated with glutathione-Sepharose (2 h instead of

30 min). Elution from this column was done with reduced glutathione (20 mL per 2 L original culture volume) prior to anion exchange chromatography (as with linL46W/linF50A; see Section 2.3.1).

2.3.5 Wild-type and Trp-Substituted LpxA Mutants

Proteins were produced by growing cells containing pET23b-LpxA (wild-type), pET23b-LpxA-Q104W (Q104W mutant), pET23b-LpxA-F162W (F162W mutant), pET23b-LpxA-G173W (G173W mutant) or pET23b-LpxA-S208W (S208W mutant) plasmids in LB media (100 μg/mL ampicillin and 34 μg/mL chloramphenicol) to mid-log phase (OD_{660} = 0.6 – 0.8) at 37°C before inducing plasmid expression with IPTG (0.5 mM) and allowing the cells to grow at 30°C to an OD_{660} of 1.5–2.0. Cells were collected by centrifugation (4°C, 7,000*g*, 10 min), resuspended in B-PER Reagent (Thermo Scientific; 10 mL/250 mL culture), and lysed by mechanical force using progressively smaller bore needles followed by gentle mixing (10 min). Centrifugation (4°C, 27,000*g*, 10 min) was used to pellet all insoluble matter. Talon Resin (Clontech) was used for purification of the His-tagged LpxAs from the resulting supernatant (1 mL packed bed volume/250 mL culture) as directed by the manufacturer. SDS-PAGE (15%, 160 V, 60 min) and GelCode Blue staining was used to assess purification procedure. Purified protein was dialyzed into sodium phosphate buffer (10 mM, pH 7.0) prior to use. This is especially important for LpxAs used in Trp fluorescence experiments as imidazole (used to elute protein bound to Talon Resin during purification) fluoresces strongly when excited at 296 nm.

2.3.6 Wild-type and Trp-Substituted *V. fischeri* AcpS Mutants

Proteins were produced by using the same process described for wild-type and Trp-substituted LpxAs (see section 2.3.5) with the exception that cells contained pET23b-AcpS (wild-type), pET23b-AcpS-F27W (F27W mutant), pET23b-AcpS-R22E (R22E mutant) or pET23b-AcpS-R22E/F27W (R22E/F27W double mutant) plasmids.

2.4 PROTEIN CHARACTERIZATION *IN VITRO*

2.4.1 Fluorescence Spectroscopy

All spectra are corrected for changes in concentration of the analyte and intensity of the fluorophore are adjusted for dilution due to titrations.

2.4.1.1 Steady-State Tryptophan and Tyrosine Fluorescence

To avoid the inner filter effect (and other artifacts), all fluorescence measurements were carried out such that the absorbance of the sample at the excitation wavelength did not surpass 0.1. For the vast majority of experiments, the absorbance was much lower than this upper limit (e.g. $A_{296\ nm}$ for cycL46W = 0.006).

Unless noted, all fluorescence measurements were carried out on a Photon Technology International (PTI) QuantaMaster-4CW (QM4CW) controlled using Felix32 (PTI); some early experiments were performed on a Perkin Elmer LS-50B (LS50B) controlled using FL WinLab (Perkin Elmer). For experiments conducted with the QM4CW, a 3 mm semimicro cuvette (Hellma; 130 µL volume) at 25°C was utilized and slit widths were set at 5 nm (both excitation and emission monochromators) for all Trp emission scan experiments. Experiments conducted with the LS50B used a 10 mm path length cuvette (Perkin Elmer*;* 120 µL) and slit widths were set to 10 nm for excitation

and emission. Cyclic/linear ACPs containing Trp (cycL46W and linL46W), wild-type/mutant LpxAs (WT, Q104W, F162W, G173W, S208W), or wild-type/mutant *P.f.*ACPs (WT, A104W, L105W) were diluted to 1 μM in 10 mM sodium phosphate (pH 7.0), 0.1 mM EDTA and excited at 296 nm with emission spectra recorded from 300 – 450 nm in photon-counting mode (QM) with an integration time of 1 s or 310 – 450 nm at a scan rate of 250 or 500 nm/min (LS50B). All fluorescence intensities (regardless of instrument) were corrected for the Raman band from the solvent and any signal arising from the blank (10 mM sodium phosphate (pH 7.0), 0.1 mM EDTA).

Tyr fluorescence experiments were performed on the QM4CW as above with slit widths of 1 nm and 2 nm for the excitation and emission monochromators, respectively. The excitation wavelength was 280 nm. Wild-type (rACP) and mutant ACPs (SA, SB, SA/SB) were diluted as for Trp fluorescence. As above, all spectra were corrected for the Raman band from the solvent and any signal arising from the blank.

2.4.1.2 Binding Analysis

2.4.1.2.1 Titration of protein with analyte

Trp-substituted (Q104W, F162W, G173W) LpxAs were titrated with increasing concentrations of various holo (h) ACPs (h-rACP, h-D30N and h-D56N) both in the absence and presence of UDP-Glc*N*Ac (200 μM); concentrations are noted in the appropriate Figure legends (Chapter 5). Prior to analysis using a python program that I developed (Appendix 2), data sets were plotted (change in AUC *versus* [analyte]), and outliers removed using Chauvenet's Exclusion criteria (Chauvenet, 1960 (Reprint of 1891)) for non-linear data. AUC was determined from the x-axis. All spectra were

recorded on the LS50B and corrected for the Raman band from the solvent, any signal arising from the blank titration, and changes in fluorescence due to dilution of the fluorophore upon titration.

Wild-type (rACP) and mutant ACPs (SA, SB, SA/SB) were titrated with increasing concentrations of Mg^{2+}, Ni^{2+}, and Zn^{2+}, as indicated in the appropriate Figure legends (Chapter 4). Prior to analysis using a python program I developed (Appendix 4) data sets were plotted (change in peak fluorescence intensity *versus* [analyte]), and outliers removed using Chauvenet's Exclusion criteria (Chauvenet, 1960 (Reprint of 1891)) for non-linear data. All spectra were corrected for the Raman band from the solvent, any signal arising from the blank titration, changes in fluorescence due to dilution of the fluorophore, and photobleaching. Since it was not determined that photobleaching was occurring immediately, control experiments were performed separately, averaged and used to correct the normalized decrease in peak fluorescence intensity. Photobleaching control experiments consisted of performing dummy titrations with buffer instead of metals.

2.4.1.2.2 Python Program

To facilitate and expedite calculation of binding constants (K_Ds), programs were written in python (Python 2.7.2) programming language using Wing IDE Personal 4.0.3-1 (Wingware). The contents of the programs can be found in Appendices 2 (calculation of iterative K_D from change in AUC data) and 4 (calculation of iterative K_D from change in peak fluorescence intensity data), while sample input/output files for each program can be found in Appendices 3 and 5. Both python programs use Equation 1 for non-linear regression.

$$Y = \frac{B_{max} \times [S]}{K_D + [S]}$$

Equation 1 Equation used for curve fitting of fluorescence titration experiments. Y represents the change in area under the curve (for LpxA titrated with ACPs) or the change in fluorescence intensity (for ACP titrated with metals), [S] is the substrate concentration, B_{max} is the maximum value for Y (i.e. saturated binding), and K_D is the dissociation constant. Both B_{max} and K_D are fitted or a given dataset using Prism® 5 and python programs created in this work (Appendices 2 and 4).

2.4.1.3 Steady-State Thermal Denaturation

Measurements were carried out on a Photon Technology International (PTI) QuantaMaster-4CW in a 3 mm semimicro cuvette (Hellma) with Peltier temperature control. Slit widths were set at 1 nm and 2 nm for the excitation and emission monochromators, respectively. ACPs diluted to 1 μM in 10 mM sodium phosphate (pH 7.0), 0.1 mM EDTA were excited at 296 nm and fluorescence emission at 310 nm was recorded over a temperature range (15 – 80°C; 1°C/min).

2.4.1.4 Stopped-Flow Tryptophan and Tyrosine Fluorescence

All stopped-flow experiments were carried out using an SFA-20M stopped-flow apparatus (TgK Scientific. *Rapid Mixing Devices*.) in conjunction with the QM4CW with slit widths of 1 nm and 2 nm for the excitation and emission monochromators, respectively. For Trp fluorescence, L46W was diluted as previously (see Section 2.4.1.1) and excited at 296 nm. Fluorescence emission at 320 nm was recorded over time. During the experiment, 20 mM $MgSO_4$ was injected using a nitrogen-driven (~2 psi) piston. Since the stopped-flow setup used employs two syringes of equal volume to inject the fluorophore and the analyte, injecting 20 mM Mg^{2+} gives a final concentration of 10 mM Mg^{2+}. Similarly, to achieve 1 μM L46W, 2 μM was loaded in the injection syringe.

Tyr fluorescence stopped-flow experiments were performed as above with a few exceptions: excitation occurred at 280 nm, emission at 307 nm was recorded over time, and multiple Mg^{2+} concentrations (as noted in the appropriate Figure legend (Chapter 4)) were used. As for Trp stopped-flow, to achieve these concentrations in the experiment, twice the concentration needs to be loaded into the injector syringe.

2.4.2 Circular Dichroism

Spectra were recorded on a J-810 spectropolarimeter (Jasco) using a 0.1 cm water-jacketed cell (at 25°C). LpxAs (wild-type, Q104W, F162W, G173W, S208W), and ACPs (linL46W, cycL46W) were diluted to 1 μM in 10 mM sodium phosphate (pH 7.0), 0.1 mM EDTA. Spectra were recorded from 190 – 260 nm in continuous mode with a speed of 20 nm/min. For ACPs, spectra were recorded before and immediately after the addition of $MgSO_4$ (10 mM).

2.4.3 Mass Spectrometry of linL46W and cycL46W

All mass spectrometry (MS) was performed at the Faculty of Medicine Proteomics Facility (Dalhousie University). Peptide mapping experiments were performed by facility Manager Elden Rowland, while apo/holo ratio experiments were completed by myself with some assistance from Elden Rowland. MS was performed on a hybrid triple quadrupole linear ion trap (Qtrap) mass spectrometer (Applied Biosystems) with a nanoflow electrospray ionization (ESI) source and spectra were recorded with Analyst® software (Applied Biosystems). All samples were sprayed through a distal coated fused silica needle (75 μm inner-diameter (ID) with 15 μm ID tip (New Objectives)). For liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) experiments, a nanoflow Ultimate system (LC Packings) was interfaced to the nanoflow ESI source of the Qtrap.

2.4.3.1 Confirmation of Cyclization through Peptide Mapping

LC-MS/MS was employed to confirm cyclization of ACP constructs. CycL46W (resolved by native PAGE) was digested (7.5 h, 37°C, standard trypsin:protein ratio) with

sequencing-grade trypsin (Promega) after which the solvent was removed by a SpeedVac SPD111V vacuum concentrator (Thermos Savant; 3 – 4 h minimum, ~35°C). Following concentration, tryptic peptides were resuspended (30 µL of 5% acetonitrile, 0.5% formic acid) and injected onto an Onyx monolithic C18 capillary column (0.1 × 150 mm, Phenomenex). Peptides were eluted with a 3 – 30% gradient of solvent B (0.1% (v/v) formic acid in acetonitrile/water (98:2)) over 35 min at a flow rate of 1 µL/min. All spectra were acquired using Information Dependent Acquisition mode (Volkmann et al., 2010).

2.4.3.2 Determination of Folding State using MS and Charge State Distribution (CSD) Analysis

Purified linL46W or cycL46W was diluted to 1 µM in ammonium acetate (10 mM) and injected into the Qtrap using a Nanospray II ion source (Applied Biosystems) at a flow rate of 100 nL/min – 1 µL/min in positive ion mode (ion source voltage = 2400 V; declustering potential = 5 V). Spectra were recorded in Enhanced MS (EMS) mode. In EMS mode ions are scanned in Q1 to the linear ion trap where they are collected and then scanned out of Q3 to produce single MS type spectra.

2.4.3.2 Determination of Apo:holo Ratio of ACP

Purified protein (linL46W or cycL46W; 1 µM) was desalted using POROS R2 media following the manufacturer's guidelines (Applied Biosystems. *POROS® 50 R1 and R2 Perfusion Chromatography® Bulk Media for Reversed-Phase Chromatography: Operating Instructions*.) and resuspended (10% acetonitrile, 0.5% formic acid). Protein was subjected to LC-MS (Konermann & Douglas, 1997) using an Onyx monolithic C18

capillary column (0.1 × 50 mm, Phenomenex) with a gradient of 10 – 70% solvent B (0.1% (v/v) formic acid in acetonitrile/water (98:2)) over 20 min (3 μL/min) to separate the apo and holo forms. Spectra were acquired in EMS mode (400 –1700 *m*/*z*) both before and after reaction with AcpS (see Section 2.4.4.3 for conditions).

2.4.4 Enzyme Reactions

2.4.4.1 UDP-*N*-acetylglucosamine Acyltransferase (LpxA) Assay

Assays were performed as previously described by Sweet *et al.* (2001) at room temperature. Each reaction mixture (10 μL total volume) contained Na^+-HEPES (40 mM, pH 8.0), bovine serum albumin (BSA; 10 mg/mL), β-hydroxymyristoyl-ACP (2 – 10 μM), purified His_6-tagged LpxA (1.0 μg), and [^{3}H]UDP-Glc*N*Ac (12.5 μM; ~17,000 dpm/pmol). Samples (2 μL) were removed at 2, 5 and 10 min (linear portion of assay) and spotted on a UNIPLATE™ silica gel G (20 × 20 cm; AnalTech) TLC plate, which was developed in chloroform/methanol/acetic acid/water (25:15:2:4). In this system, the acylated product migrates while the unreacted UDP-Glc*N*Ac remains at the origin. The product was located using a Bioscan System 200 Imaging Scanner with Win-Scan software. The region of interest on the TLC plate was subsequently scraped and radioactivity measured by liquid scintillation counting (LSC) using a Beckman LS6500 Multi-Purpose Scintillation Counter. All values determined by LSC were corrected for blank values.

2.4.4.2 Holo-ACP Synthase (AcpS) Assay

A modified version of the assay described by Lambalot and Walsh (1997) was performed at 25°C or 37°C. Reactions (10 μL) contained sodium phosphate (50 mM, pH

7.0), $MgCl_2$ (10 mM), dithiothreitol (DTT; 5 mM), apo-ACPs (50 μM), and purified *E. coli* or *V. fischeri* AcpS (0.25 – 1.00 μg). [^{3}H]Acetyl-CoA (57 μM, 444 dpm/pmol) was added to start the reaction and samples (3 μL) were removed at 1, 2 and 5 min. The reaction was stopped by addition of cold trichloroacetic acid (TCA; 800 μL of 10% w/v) and BSA (20 μL of 25 mg/mL). After incubation on ice for 15 min, the precipitate was obtained by centrifugation (12,000*g*, 5 min), washed twice with cold TCA (800 μL of 10% w/v), and solubilized in 50 μL of formic acid (Gong et al., 2007). Product formation was measured by LSC and corrected for blank values obtained in the absence of added apo-ACP.

2.4.4.3 Conversion to Holo-ACP using AcpS

To convert ACPs from the apo to holo form, conditions for the AcpS assay described above were used with the following modifications: non-radioactive CoA (0.3 – 0.5 mM for rACP; 0.5 – 1.5 mM for cycL46W), and more enzyme (1 – 10 μg *V.f.*AcpS or *E.c.*AcpS/100 μg rACP; 1 – 10 μg *V.f.*AcpS or *E.c.*AcpS/10 μg cycL46W). Furthermore, the reaction was allowed to proceed for 2 – 4 h, with the reaction stopped at various times using 4X Native gel loading buffer instead of TCA. The conversion was examined post-reaction via native PAGE.

2.4.5 Gel Filtration Analysis of LpxA

Standards (ferritin (440 kDa; Pharmacia), aldolase (158 kDa; Pharmacia), BSA (66 kDa; Pierce), ovalbumin (45 kDa; Pharmacia), and RNase A (13.7 kDa; Pharmacia)) and samples (wild-type, Q104W, F162W, G173W, S208W LpxAs) were individually subjected to gel filtration analysis. All proteins (2 mg quantities in 0.5 mL or less) were

first dialyzed into Buffer A (10 mM MES (pH 6.0), 2 mM DTT) and filtered (MILLEX®GV 0.22 μM; Millipore) prior to gel filtration (1 mL/min) using a Superose 6 10/300 GL size exclusion column (GE Healthcare) on a Waters 650 Advanced Protein Purification System. The elution volume of each protein was recorded and standards plotted to allow for estimation of sample protein molecular weight.

2.4.6 Gel Electrophoresis

All gel electrophoresis was carried out using the Mini-PROTEAN® 3 Electrophoresis System (Bio-Rad).

2.4.6.1 SDS-PAGE

Separating (15%) and stacking (4%) gels were prepared using 30% acrylamide/bis solution (37.5:1, 2.6% C; Bio-Rad). Both separating and stacking gels contained Tris-HCl (375 mM), however the separating gel is at pH 8.8 while the stacking gel is pH 6.8 (Laemmli, 1970). Gels were polymerized with 10% ammonium persulphate (25 uL per gel) and TEMED (2.5 uL per gel). All samples were mixed with 4× SDS-PAGE sample buffer (250 mM Tris-HCl (pH 6.8), 25% glycerol, 2.5% SDS, 0.1% bromophenol blue, ~1M β-mercaptoethanol) prior to separation. Electrophoresis was typically carried out at 160 V (~60 min) at room temperature, and gels were stained with GelCode Blue staining reagent (Pierce). Gels were dried using GelWrap (BioDesign).

2.4.6.2 Native PAGE

Separating (20%) and stacking (4%) gels were prepared using 40% acrylamide/bis solution (29:1, 3.3% C; Bio-Rad). Separating gels contained Tris-HCl (390 mM, pH 9.2) while stacking gels contained Tris-HCl (195 mM, pH 6.8). Gels were polymerized as

described for SDS-PAGE (see Section 2.4.6.1). All samples were mixed with 4× native PAGE sample buffer (250 mM Tris-HCl (pH 6.8), 25% glycerol, 0.1% bromophenol blue) prior to separation. Electrophoresis was typically carried out at 160 V (~90 min) at room temperature, and gels were stained and dried as in 2.4.6.1.

2.4.6.3 Urea PAGE

Separating and stacking gels were made as for native PAGE (Section 2.4.6.2) with the addition of urea (6 M; Fischer Scientific, certified ACS quality) to both gels. Samples were diluted using 4× native PAGE sample buffer with the addition of urea (6 M). Electrophoresis occurred as for native PAGE (160 V, ~60 min); however for better and faster separation the electrophoresis apparatus can be heated to 37°C using a low-temperature oven. Gels were stained and dried as in 2.4.6.1.

2.5 Molecular Dynamics (MD)

2.5.1 *In Silico* Creation of Protein Data Bank (PDB) Files

2.5.1.1 LinL46W and cycL46W ACPs

The NMR solution structure of *E. coli* apo-ACP (PDB ID: 2K92 (Wu et al., 2009); 86% sequence identity to *V. harveyi* ACP) was used as a starting point for creation of both files *in silico*. To minimize the changes to existing atoms/residues, one of the 20 structures included in 2K92 (with termini pointed towards each other) was chosen. The mutation equivalent to *V. harveyi* L46W was made in 2K92 using VMD's "Mutate Residue" functionality (version 1.8.7; (Humphrey et al., 1996); see Appendix 6 for details on using "Mutate Residue" and Appendix 7 for how to generate the necessary input files) and the resultant PDB was loaded into PyMol 1.2r1 (De Lano, 2002) to complete the

construction. For cycL46W, Gly, Ser, and Ala were added sequentially to the C-terminus of L46W, adjusting the bond dihedrals to bring the termini as close as possible prior to joining of the N- and C-termini. Minimization steps were performed in PyMol to optimize bond lengths and dihedrals of the newly added residues/bonds before saving the molecule for simulations using NAMD. If this minimization was not done, a number of problems arose with the saved PDB file: slight steric clashes (atoms x, y, z coordinates are close; Figure 17), extreme steric clashes (atoms x, y, z coordinates are the same, resulting in deletion of one of the atoms and creation of bonds, often resulting in triangulated shapes; Figure 18), and improper dihedrals, multiple slight and extreme steric clashes (resulting in missing atoms, extra bonds, and/or strained bonds (due to length of dihedral); Figure 19). For linL46W, Gly was added to the C-terminus, while Ser and Ala were added to the N-terminus to match the linear protein produced upon expression. As residues were added with acceptable dihedrals and bond lengths, no minimization in PyMol was required prior to saving the PDB file for simulations.

2.5.1.2 LpxA and AcpS Mutants

In silico mutant LpxAs and AcpSs were made using VMD's built-in "Mutate Residue" function (as discussed in Appendix 6). To use the "Mutate Residue" functionality, a protein structure file (PSF) must be generated (see Appendix 7) to be used with a PDB file as the input files for a MD simulation. For *E.c.*LpxA, PDB ID = 2JF3 (Ulaganathan et al., 2007) was used as a starting structure to create Q104W, F162W, G173W, and S208W Trp-substitutions. This structure contains one of the substrates for the LpxA reaction (UDP-Glc*N*Ac), which was removed prior to mutation.

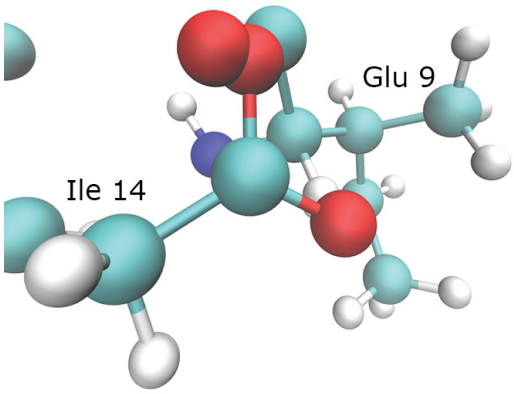

Figure 17 Diagram showing simple steric clashes in structures produced via PyMol when minimization is not performed prior to saving the final PDB for MD simulation. For simplicity only Glu 9 and Ile 14 are shown. Model is displayed using "CPK" graphical representation and is colored by atom. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

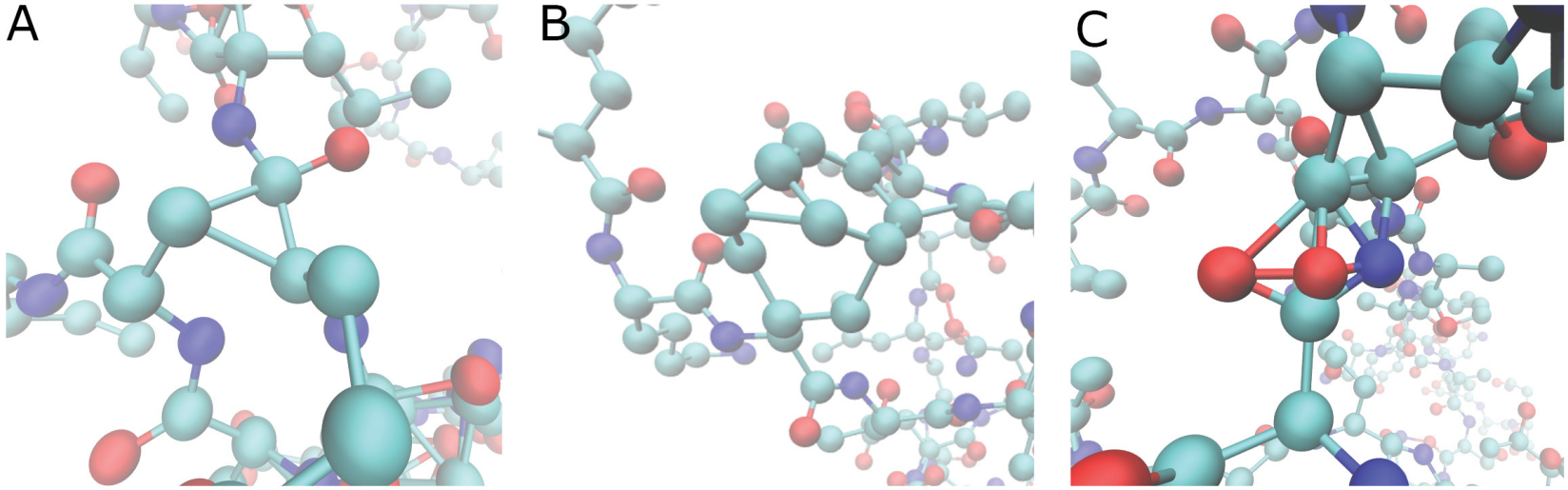

Figure 18 Diagram showing extreme steric clashes in structures produced via PyMol when minimization is not performed prior to saving the final PDB for MD simulation. Panel A depicts a portion of one attempt at creating cycL46W. Panels B, C depict two different portions of a second attempt at creating cycL46W. Models are displayed using “CPK” graphical representation and are colored by atom. Hydrogen atoms are not shown for simplicity. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

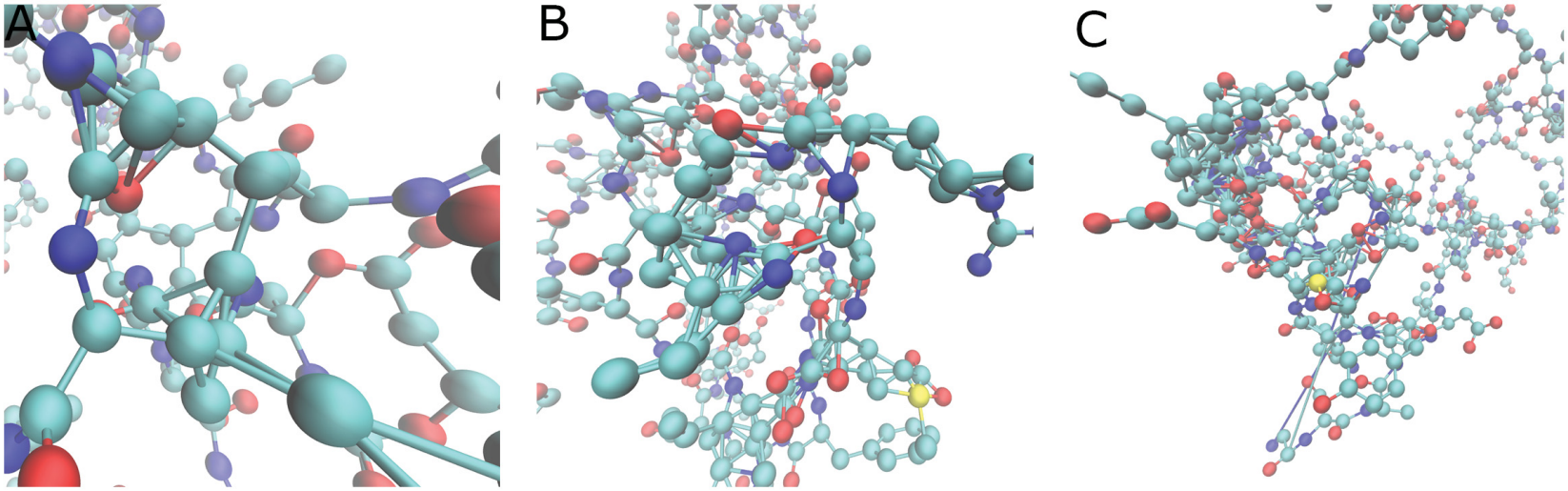

Figure 19 Diagram showing improper bonds (dihedrals and lengths) in addition to extreme steric clashes in structures produced via PyMol when minimization is not performed prior to saving the final PDB for MD simulation. Panels A, B depict two different portions of one attempt at creating cycL46W. Panel C depicts a portion of a second attempt at creating cycL46W. Models are displayed using “CPK” graphical representation and are colored by atom. Hydrogen atoms are not shown for simplicity. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

As a structure of *V. fischeri* AcpS (*V.f.*AcpS) does not exist, *B. subtilis* AcpS (PDB ID = 1F7T (Parris et al., 2000); 44% sequence identity to *V.f.*AcpS) was used as a starting structure to create files for mutants R21E, F25W, and R21E/F25W. For the double mutant (R21E/F25W), mutations were made consecutively. The aforementioned *B. subtilis* AcpS (*B.s.*AcpS) mutations are equivalent to the R22E, F27W, and R22E/F27W mutations made to *V.f.*AcpS *in vitro* (see Section 2.1.3).

Prior to MD simulation, wild-type and mutant LpxAs and AcpSs were converted to their biologically active form (trimer) and solvated. To trimerize each wild-type and mutant monomer, the BIOMT transformations were copied from the relevant section of the wild-type PDB (2JF3 (Ulaganathan et al., 2007) for LpxA and 1F7T (Parris et al., 2000) for AcpS. The BIOMT transformation information is contained in the "REMARK 350" section of the starting PDBs (Worldwide PDB. *PDB File Format Information: Remarks 300 - 999*.). Once the BIOMT records were copied into each wild-type and mutant monomer, VMD and a readily available script (mono2poly.tcl) can be used to produce the final multimers (see Appendix 8 for script contents and usage). Following trimerization, the wild-type and mutant LpxAs and AcpSs were solvated (see Appendix 9) prior to simulation.

2.5.1.3 Homology Modeling of *Plasmodium falciparum* AcpS (*P.f.*AcpS)

P. falciparum AcpS (*P.f.*$AcpS_{FL}$) is made up of two primary sections: the N-terminus (of unknown function) shares sequence identity with the metal-dependent phosphohydrolase enzyme family (Cai et al., 2005), which includes *E. coli* AcpH ACP phosphodiesterase (AcpH) (Thomas et al., 2007), and the C-terminus, which shares

sequence identity with bacterial AcpS. For the purposes of this work, *P. falciparum* AcpS (*P.f.*AcpS) refers to the C-terminal domain unless specified otherwise. As there is no published PDB for *P.f.*AcpS, homology modeling was used to create one. All homology modeling was done using Swiss-PdbViewer: DeepView Version 4.0 or 4.1 (Swiss Institute for Bioinformatics. *Swiss-PdbViewer: DeepView*.) which is directly linked to SWISS-MODEL (Swiss Institute for Bioinformatics. *SWISS-MODEL*., Peitsch, 1995, Arnold et al., 2006, Kiefer et al., 2009), an automated homology modeling server. Homology modeling was done following the DeepView Tutorial on Homology Modeling (Swiss Institute for Bioinformatics. *DeepView Tutorial: Homology Modelling*.).

Three different starting structures were used to homology model *P.f.*AcpS: *Plasmodium yoelii* AcpS (PDB ID: 2BDD (Vedadi et al., 2007); 66% sequence identity), *P. yoelii* AcpS (PDB ID: 2QG8; 66% sequence identity), and *Bacillus subtilis* AcpS (*B.s.*AcpS) (PDB ID: 1F7T (Parris et al., 2000); 30% sequence identity). In all cases, a monomer of each AcpS was saved independent of the starting PDB to facilitate homology modeling.

Unlike bacterial LpxA and AcpS (see Section 2.5.1.2), homology modeled *P. falciparum* AcpS was simulated in its monomer form (see Section 2.5.2), due to a lack of local computing power available at the time. A longer minimization and equilibration period, as well as simulation of the biologically active form (trimer), would be preferable.

Following MD simulation, the final *P.f.*AcpS monomer structure was saved (see Appendix 11 for details on saving the final structure) and trimerized. Unlike bacterial LpxA and AcpS (2.5.1.2), which used BIOMT information within the PDB file (see 2.5.1.2 for more detail) to trimerize the monomer, homology modeled *P.f.*AcpS does not

contain this information; thus a different approach was necessary to trimerize the monomer. The Protein Data Bank (Research Collaboratory for Structural Bioinformatics. *RCSB Protein Data Bank*.) contains both the monomer of *P. yoelii* AcpS (PDB ID: 2BDD), as published by the authors, as well as the biologically active form (generated automatically by The Protein Interfaces, Surfaces and Assemblies (PISA) server (European Bioinformatics Institute. *Protein Interfaces, Surfaces and Assemblies*.). Using this biologically active form PDB file, and 3 copies of the simulation end structure, each monomer can be individually aligned to a subunit of the *P. yoelii* AcpS trimer using UCSF Chimera's (Pettersen et al., 2004, Yang et al., 2012) built in feature, MatchMaker (see Appendix 12). Once aligned, the structures were saved relative to the *P. yoelii* AcpS trimer structure, thus preserving the position of each monomer relative to other monomers across files. Since each of the monomers is a copy of the original monomer, they all have the same chain ID. To combine multiple PDB files, each subunit must have a unique chain ID. This was achieved by manually editing each PDB file such that the chain was changed from X to A, B, and C for subunit 1, 2 and 3, respectively. Loading these newly created subunit files into VMD, multiple PDBs can be combined using the built in Merge Structures plugin. To use this plugin, both PDB and PSF files are required for the structures to be combined, thus PSF files were generated for each subunit PDB file using AutoPSF (as in Appendix 7). Using the Merge Structures plugin (see Appendix 13), subunit A (Figure 20A) and B (Figure 20B) were combined and saved as subunitA-B PDB and PSF files (Figure 20D). SubunitA-B was then combined with subunit C (Figure 20C) to give the final *P.f.*AcpS-Trimer PDB and PSF Files (Figure 20E).

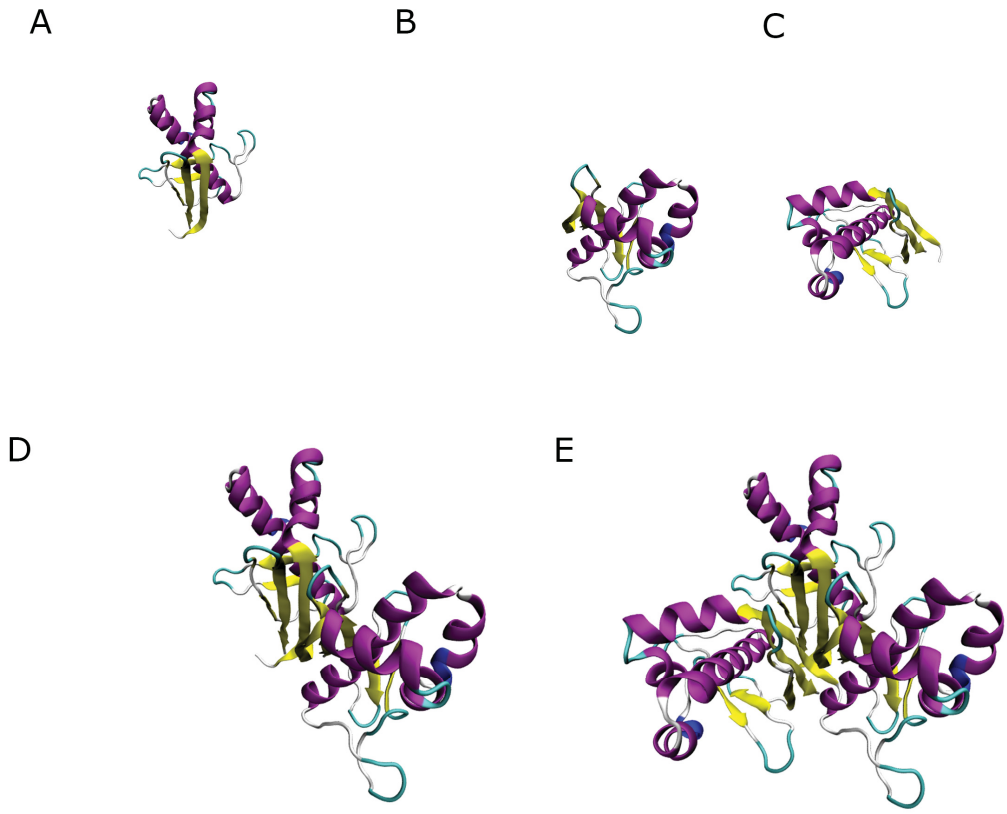

Figure 20 Simulation end structure for homology modeled *P.f.*AcpS depicted during various stages of the conversion to its assumed biologically active form. Structures of the subunits A, B, and C (A, B, C, respectively) positioned by alignment in Chimera, prior to combination, are depicted as well as the two stages of the combination process using VMD's Merge Structures plugin: Subunits A and B (D) and subunits A, B and C (E). Model is displayed using "NewCartoon" graphical representation and is colored by secondary structure. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

As noted above, the N-terminal fragment of the full length *P. falciparum* AcpS (*P.f.*AcpS-N) shares sequence identity with a family of enzymes (Cai et al., 2005) that includes *E. coli* AcpH (*E.c.*AcpH) (Thomas et al., 2007); there is 32% sequence identity between *P.f.*AcpS-N and *E.c.*AcpH. As there isn't a published structure of *E.c.*AcpH available, a structure needed to be created before *P.f.*AcpS-N could be homology modeled. Previous work showed that *E.c.*AcpH (a member of the HD Phosphatase/Phosphodiesterase Family) can be convincingly (29% sequence identity between *E.c.*AcpH and SpoT) modeled on the structure of the hydrolase domain of the bifunctional bacterial protein, SpoT (Thomas et al., 2007). Both were used as starting structures to homology model *P.f.*AcpS-N (without its leader sequence). First, the sequence of *E.c.*AcpH was homology modeled using Swiss-PdbViewer: DeepView Version 4.0 (Swiss Institute for Bioinformatics. *Swiss-PdbViewer: DeepView*.) and SWISS-MODEL as above. Next, the *E.c.*AcpH structure produced by SWISS-MODEL was used as the starting structure for homology modeling of *P.f.*AcpS-N, again using SWISS-MODEL. For the second approach, *P.f.*AcpS-N was homology modeled directly using SpoT as the template structure; there is 31% sequence identity between *P.f.*AcpS-N and SpoT. As a control, BSA (a protein that shares little sequence identity (~20% sequence identity to *P.f.*AcpS-C) and no structural identity with SpoT) was homology modeled using SpoT as a template.

2.5.1.4 Homology Modeling of *Vibrio fischeri* (*V. fischeri*) AcpS

As with *P.f.*AcpS, there isn't a published structure of *V. fischeri* AcpS (*V.f.*AcpS), thus homology modeling was used to create a structure. *Bacillus subtilis* AcpS (*B.s.*AcpS) (PDB ID: 1F7T (Parris et al., 2000); 44% sequence identity to *V.f.*AcpS) was

used as the template. Homology modeling was completed as for *P.f.*AcpS (Section 2.5.1.3) using Swiss-PdbViewer: DeepView Version 4.0 and SWISS-MODEL. Due to lack of local processing power at the time of homology modeling, the structure produced by SWISS-MODEL was not simulated. The monomer of *V.f.*AcpS was trimerized as for *P.f.*AcpS using *B.s.*AcpS (PDB ID: 1F7T) as the template to align each individual monomer. Since the structure was not simulated and was not to be used for further analysis (other than comparison to template structure), the three *V.f.*AcpS monomers were not merged into one PDB file, as was done for *P.f.*AcpS. The three monomers (saved relative to the template) were loaded into VMD and treated as one when producing images (see Section 5.3.2.3).

2.5.1.5 rACP, SA, SB, and SA/SB Mutants

As there is currently no published structure for native *Vibrio harveyi* ACP to use as a starting point, the structure of a mutant of rACP (A75H; generated by a collaborating lab (Chan et al., 2010)) was used. All mutations were made using VMD's built-in "Mutate Residue" function (as discussed in Appendix 6) after creation of corresponding PSF files as discussed in Section 2.5.1.2. For rACP, His 75 was converted back to alanine. This PDB was saved using a selection (set command) and the writepdb command as described in Appendix 14 and the PSF file was created (Appendix 7). To create the divalent cation-binding site A mutant of ACP (SA), the mutations (D30N, D35N, D38N) were made consecutively to rACP, saving a PDB, creating a PSF and re-loading the structure after each mutation until all three had been successfully made. As with SA, SB was created using rACP as the starting structure and making mutations (E47Q, D51N, E53Q, D56N) consecutively. Lastly, to create the divalent cation-binding site double

mutant (SA/SB), the SA structure above was used as the starting structure and the site B mutations (E47Q, D51N, E53Q, D56N) made consecutively as for SA and SB. Once the mutants had been created *in silico*, a PSF of each was created. This final PDB/PSF combination for each ACP (rACP, SA, SB, SA/SB) was used as the input files for creating a solvation sphere surrounding each protein. The script and how it was used to solvate the proteins can be found in Appendix 9. Following solvation, the ACPs were individually minimized and equilibrated in a molecular dynamics (MD) simulation (see Section 2.5.2 for details).

2.5.1.6 Addition of Mg^{2+}, Ni^{2+}, and Zn^{2+} to rACP and mutant ACPs

Following minimization and equilibration simulations of rACP, SA, SB, and SA/SB (Section 2.5.2), an "equilibrated averaged" structure was created and saved as described in Appendix 14. These structures were then aligned using Chimera 1.7 (build 38197) as described in Appendix 1 prior to metal addition. As there is no way to add a specific atom to a defined position in VMD, Chimera, or PyMol, the structure of a protein containing the ion of interest was found in the Protein Data Bank (Research Collaboratory for Structural Bioinformatics. *RCSB Protein Data Bank*.). Magnesium (Mg^{2+}) ions were saved from the CD11a domain (domain 1 of the CD11a/CD18 integrin) (PDB ID: 1ZOO (Qu & Leahy, 1996)). This was done by manually editing 1ZOO.pdb and deleting all content except the Mg^{2+} ions. The "ATOM" information lines for Mg^{2+} were then copied and pasted into the equilibrated and aligned ACP structure created above. Lastly, the newly created rACP PDB (with Mg^{2+} ions present) was loaded in VMD and each Mg^{2+} ion re-positioned (using Mouse/Move/Atom in the VMD Main window) to the two divalent cation binding sites (site A and B), creating rACP+Mg.pdb.

Finally this PDB was saved (as in Appendix 14), a PSF file generated (as in Appendix 7), water solvation sphere added (as in Appendix 9) and the protein minimized and equilibrated (a sample configuration file is shown in Appendix 16). Once the rACP-Mg PDB file had been created, it was duplicated, opened and Mg was changed to nickel (Ni), creating rACP+Ni.pdb. Opening the PDB file in VMD and using "Graphical Representations", it was confirmed that the Mg^{2+} was changed to Ni^{2+}. This procedure was repeated for zinc (Zn^{2+}) (replacing Mg with Zn) to create rACP+Zn.pdb. As with rACP+Mg.pdb, PSF files of rACP+Ni.pdb and rACP+Zn.pdb were created, the protein solvated, and then simulated (minimization and equilibration steps). For additions of metal ions to SA, SB and SA/SB, the information was manually copied from the respective rACP PDB (rACP+Mg, rACP+Ni, or rACP+Zn) and pasted into the PDB for SA, SB, and SA/SB. This was done to ensure the starting position of the atoms was the same in each of the ACP/metal combinations. By ensuring the same starting coordinates, the effect of the site mutations on the position of the ions could be determined. As a control for the metal additions, the ACPs without metal ions added (the starting structures for metal addition) were also simulated in a solvation sphere.

2.5.2 Molecular Simulations

All molecular dynamics (MD) simulations were performed with NAMD version 2.7b1 or 2.9 (Phillips et al., 2005) using CHARMM22 (MacKerell et al., 1998) force fields. Each simulation was carried out in a sphere of explicit water molecules (see Appendix 9 for details on adding a solvation sphere) at 310 K using Langevin temperature control (see Appendices 16 and 17 for details on Langevin temperature control and Langevin dynamics (LD), respectively) and was set up with two periods: (1)

steepest descent minimization and (2) equilibration (see Appendix 16 for a sample configuration file). To ensure constructs had adopted an equilibrated structure, the average root mean square deviation (RMSD) for backbone atoms was calculated over the course of the simulation (see Appendix 18 and 19 for script and usage to calculate RMSD for the raw (non-smoothed) and smoothed trajectory files, respectively). Table 10 provides details on the number of steps of minimization and equilibration, as well as other information, for all MD simulations performed in this thesis.

2.5.3 Simulation Analysis and Image Creation

After completion, the RMSD for the simulation was measured to determine if equilibrium had been reached. This check was performed on the non-smoothed DCD trajectory file (directly from NAMD); however, the majority of analyses (including subsequent RMSD calculations) were performed on the smoothed DCD trajectory file and/or the equilibrated structure as indicated. For detailed information on the script to create a smoothed trajectory and its use see Appendix 20.

2.5.3.1 Distance Measurements

Distances between atoms were measured in VMD (1.8.7 or 1.9.1) using the label feature (See Appendix 21 for a detailed procedure for calculating distances between atoms).

2.5.3.2 Volume and Solvent Accessible Surface Area Calculation

Protein volumes of equilibrated averaged structures from ACP ± metal simulations were measured using UCSF's Chimera version 1.7 (build 38197). See

Table 10 Simulation specific information for MD simulations completed in this study.

Protein(s)	Minimization		Equilibration		Version of NAMD	Computer(s) Used	Average Time to complete
	Steps	Time	Steps	Time			
LpxAs (wild-type, Q104W, F162W, G173W, S208W)	10,000	20 ps	500,000	1 ns	2.6b	Morticia-PC	4 – 6 d
AcpS/ACP Combinations (see Table 11 for combinations)	10,000	20 ps	125,000	0.25 ns	2.6	Morticia-PC	~ 1 wk
linL46W and cycL46W	500,000	1 ns	5,000,000	10 ns	2.71b	Fester-HPC	~2.5 – 3.5 wk
Homology Modeled *P. falciparum* AcpS	100,000	0.2 ns	125,000	0.25 ns	2.71b	Inspiron1-PC	1 – 2 d[1]
ACP Mutations (rACP, SA, SB, SA/SB)	125,000	0.25 ns	5,000,000	10 ns	2.9	XPS-PC	~12.5 h[1]
ACP ±Metals (see Table 12 for combinations)	125,000	0.25 ns	5,000,000	10 ns	2.9	Desktop1-PC, Inspiron2-PC, XPS-PC	Desktop1-PC: ~15 h[2] Inspiron2-PC: 40 - 60 h[1] XPS-PC: ~10 h[1]

[1] = 1 simulation running [2] = 2 simulations running concurrently

Table 11 AcpS-ACP Combinations used in MD simulations.

AcpS	ACP
Wild-type	Wild-type
R22E	Wild-type
Wild-type	E41K
R22E	E41K
F27W	Wild-type
R22E/F27W	Wild-type
F27W	E41K
R22E/F27W	E41K

Table 12 ACP-Metal Combinations used in MD simulations.

ACP	Metal
rACP	None
SA	None
SB	None
SA/SB	None
rACP	Mg^{2+}
SA	Mg^{2+}
SB	Mg^{2+}
SA/SB	Mg^{2+}
rACP	Ni^{2+}
SA	Ni^{2+}
SB	Ni^{2+}
SA/SB	Ni^{2+}
rACP	Zn^{2+}
SA	Zn^{2+}
SB	Zn^{2+}
SA/SB	Zn^{2+}

Appendix 22 for details. Additionally, the solvent accessible surface area (SASA) was determined for all residues (see Appendix 23 for details).

2.5.3.3 Scripts

Due to the length of the MD simulations, scripts were used to automate analysis (RMSD, trajectory smoothing, etc). These scripts are executed using the "Tcl Console" of VMD and can be found in the Appendices (18, 19: RMSD Calculation scripts; 20: trajectory smooth script). RMSD scripts calculated the RMSD from the starting structure for each frame. An alternate approach would be to calculate the RMSD from the final structure.

2.5.3.4 Image and Movie Creation

Using VMD (1.8.7 or 1.9.1), images containing measured distances were rendered using the "Snapshot VMD OpenGL Window" engine producing a bitmap (BMP) file. All other images were saved as POV-Ray files using the POV3 engine (see Appendix 24 for a more detailed description of saving a BMP or POV file) and subsequently rendered externally using POV-Ray (3.6.2) ray-tracing software (32- or 64-bit build, depending on the computer doing the rendering; see Table 13). Movie frames for simulations were created using a script to automate the saving of each frame of the simulation to POV-Ray files (see Appendix 26, Appendix 27 and Appendix 28). Simulation sequences were then rendered using POV-Ray. Due to the limitations of POV-Ray's queuing system (maximum 512 files), a MS-DOS batch file was written to automate rendering (Appendix 29, Appendix 30). Once all frames for a simulation were rendered, movies were created

Table 13 Simulation specific rendering information for MD simulations completed in this study.

Protein(s)	Computer(s) Used for Rendering POV-Ray Files	POV-Ray Version/Build	Approximate Render Time
LpxAs (wild-type, Q104W, F162W, G173W, S208W)	Inspiron1-PC	3.6 32-bit	~72 h[1]
AcpS-ACP Combinations (see Table 11 for combinations)	Inspiron1-PC	3.6 32-bit	~45 – 55 h[1]
linL46W/cycL46W	Inspiron1-PC	3.6 32-bit	36 – 48 h[1]
Homology Modeled *P. falciparum* AcpS	Inspiron1-PC	3.6 32-bit	8 – 10 h[1]
ACP Mutations (rACP, SA, SB, SA/SB)	XPS-PC	3.6 64-bit	~12 h[2]
ACP ±Metals (see Table 12 for combinations)	Desktop2-PC Inspiron2-PC XPS-PC	3.6 Desktop2-PC: 32-bit Inspiron2-PC: 32-bit XPS-PC: 64-bit	Desktop2-PC: 34 – 38 h[2] Inspiron2-PC: 32 – 35 h[2] XPS-PC: ~4 h[3]

[1] = rendered in sets of 512 images over multiple days; render time given is cumulative total

[2] = rendered in 1 set (20501 frames)

[3] = rendered in 4 sections (~5125 frames each; rendering simultaneously) to more efficiently use computer power and time; if rendered in 1 set (20501 frames) average time ~15.5 h; rendering 4 sections simultaneously, 2 simulations can be rendered/night as opposed to 1 if rendered in 1 set.

using VideoMach (Gromada. *VideoMach*.) or FFmpeg (FFmpeg. *FFmpeg: a complete, cross-platform solution to record, convert and stream audio and video*). The video showing potential helix movement in cycL46W exposing the fatty acyl chain was created using Chimera's "Morph Conformations" and built-in movie maker.

"Sausage representation" figures for linL46W and cycL46W were created using MOLMOL 2.6.0 (Koradi et al., 1996) (see Appendix 32 for how these representations were created).

CHAPTER 3 INTEIN-MEDIATED CYCLIZATION OF BACTERIAL ACP

3.1 Introduction and Rationale

The mechanism by which ACP unfolds and transfers its attached and embedded fatty acid to or from a partner enzyme is largely unknown. However, the flexibility and dynamic properties of ACP are thought to be essential for its function (Wu et al., 2009). As the N- and C-termini of ACP are both mobile (Chan et al., 2008) and in relatively close proximity, intein-mediated cyclization was used to constrain this flexibility of *V. harveyi* ACP and test the hypothesis that ACP does not need to completely unfold to be functional. An ACP mutant, L46W, was chosen as the template for this cyclization, as placement of a fluorescent Trp probe at this position is known to confer sensitivity to the conformational state of the protein without affecting either its substrate properties with partner enzymes or its overall secondary structure (Gong et al., 2008). This chapter was the result of a collaboration: Gerrit Volkmann performed construct creation (Section 2.1.1), initial protein expression and gel characterization (Section 2.2.1 and Section 3.2.1), and *in vivo* complementation assays (Section 2.2.3 and discussion (Section 3.5)). I performed all remaining protein purification, biophysical characterization (Section 3.3) and molecular dynamics (Section 3.4).

3.2 Results: Construction and Confirmation of Cyclic ACP

To create a cyclic ACP, the N-terminal splicing domain of the *Ssp* GyrB intein (residues 1 – 111 followed by a His_7 tag; I_NH) (Perler, 2002) was fused to the C-terminus of ACP, while the C-terminal splicing domain (residues 393 – 435; I_C) was fused to the

N-terminus of ACP to create plasmid pTCYC-L46W (Table 7). After expression and *trans*-splicing, an ACP (cycL46W) with its C- and N-termini joined by a three amino acid linker (glycine-serine-alanine; glycine attached to ACP's "C-terminus" and alanine attached to "N-terminus") would be expected (Figure 21). As a control for the effect of cyclization, a L46W construct was created with the additional linker residues attached to the appropriate terminus (N-terminus: SA; C-terminus: G), resulting in plasmid pGEX-linL46W (Table 7). Expression of this protein would be expected to produce a linear folded ACP (Gong et al., 2008) with 3 additional residues (linL46W; Figure 21). A splicing deficient control was created as for cycL46W, with the exception that a mutated *Ssp* GyrB split-intein (C1A/N435A) was used, which is unable to catalyze the *trans*-splicing reaction, resulting in plasmid pTPRE-L46W-mut (Table 7). Upon expression, a folded ACP (linL46W^{premut}) with large C- and N-terminal extensions would be expected (Figure 21). Lastly, to test the effect of cyclization on a folding deficient ACP mutant (F50A (Flaman et al., 2001)), a construct was created as for cycL46W, except using F50A as the ACP template. This resulted in plasmid pTCYC-F50A (Table 7). Following expression and *trans*-splicing, an ACP (cycF50A) with its C- and N-termini joined by the three amino acid linker would be expected (Figure 21; suspected structure shown). As a control for the effect of cyclization on the structure and function of cycF50A, its linear counterpart (without the added linker residues), from our lab's extensive mutant collection, was used. Expression of this plasmid, pGEX-F50A (Flaman et al., 2001), presumably results in an unfolded ACP (linF50A) *in vivo* (Figure 21).

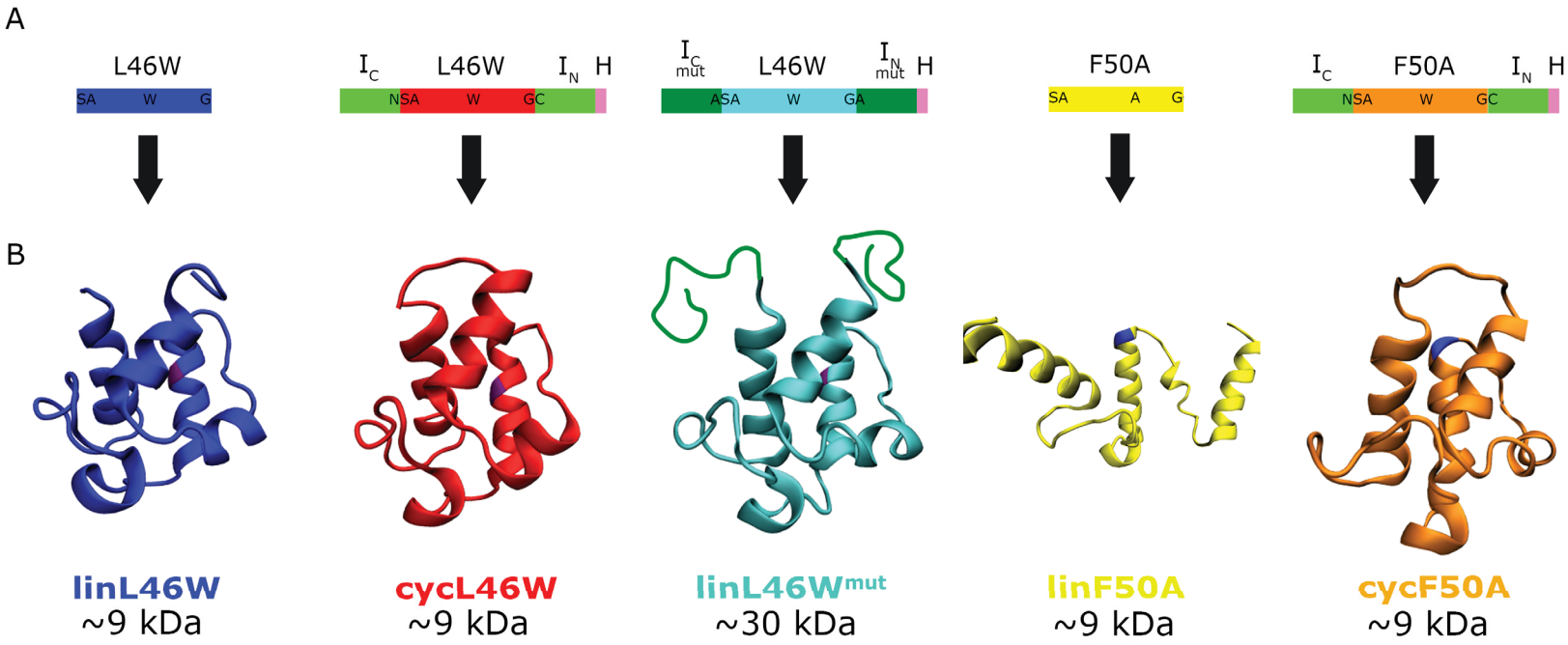

Figure 21 Schematic representations of cyclization constructs and their controls (A) as well as the resultant proteins (B) expressed *in vivo*. (B) "NewCartoon" representations of linL46W (blue), cycL46W (red), preL46W^{mut} (cyan), linF50A (yellow) and cycF50A (orange). Residues of interest highlighted: L46W (purple), F50A (blue). His_6-tag (H) is shown in pink. Structures for linF50A and cycF50A are schematics based on previous data from our lab (Flaman et al., 2001) and this work, respectively. Approximate molecular mass of the final expressed protein is given below structures. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD 1.91 and POV-Ray 3.6.

3.2.1 Protein Expression and SDS-PAGE

Initial protein expression and SDS-PAGE analysis was performed by Gerrit Volkmann (Volkmann, 2009, Volkmann et al., 2010). Briefly, expression of preL46W^{mut} showed an induced protein band of ~31 kDa that reacted with anti-His-tag antibodies (Figure 22A); thus, this protein corresponds to the *trans*-splicing deficient preL46W^{mut}. In contrast, expression of cycL46W resulted in two induced protein bands of ~14 kDa and <10 kDa. The larger of the two fragments reacted with anti-His-tag antibodies and is likely the IN_H split-intein (13.3 kDa) (Figure 22A). The smaller fragment was identified using trypsin digestion and MS/MS to be the I_C split-intein (5.1 kDa). Detecting both split-intein fragments and none of the precursor indicated that cyclization was complete. Although an induced band for cycL46W was not evident in cell extracts using SDS-PAGE, subsequent analysis using native PAGE revealed a candidate protein that could correspond to cyclic ACP (Figure 22B) (Volkmann, 2009, Volkmann et al., 2010).

3.2.2 Mass Spectrometry – Peptide Mapping

To verify the identity of the suspected cycL46W band on native PAGE (Figure 22B; asterisk), the band was excised, cleaved with trypsin, and subjected to LC-MS/MS. Of the peptides detected during this process, one in particular (a $[M+3H]^{3+}$ peptide ion with an m/z value of 846.84) matched the sequence ITTVQAAIDYVNSAQGSASNIEER (Figure 23). Residues 1 – 15 and 19 – 24 of the peptide correspond to the C-terminus and N-terminus of *V. harveyi* ACP, respectively, while residues (16 – 18) correspond to the three amino acid linker left after intein *trans*-splicing (GSA) with the peptide bond between Gly and Ser being the point of ligation. These results confirm that intein-

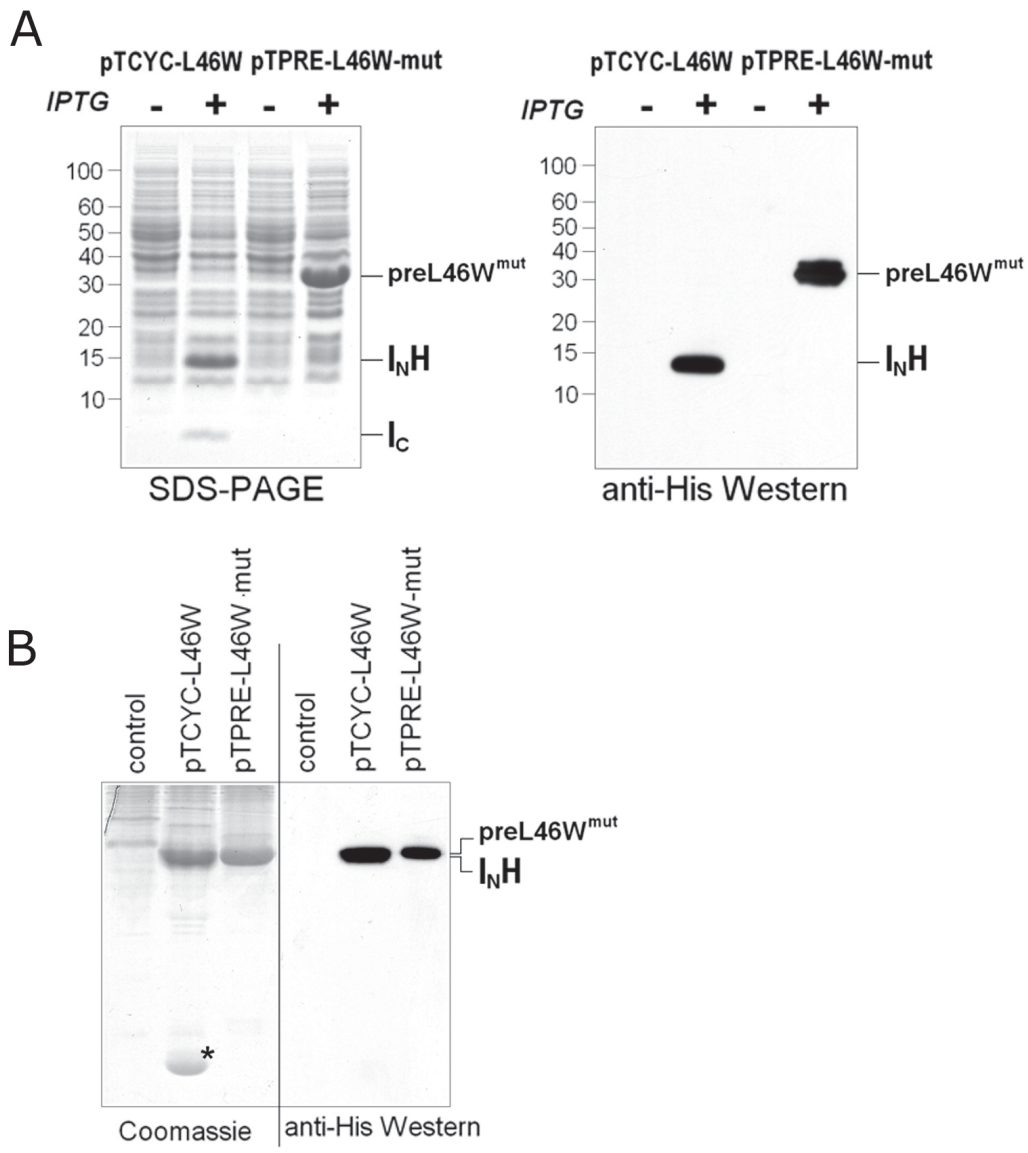

Figure 22 Initial gel confirmation of cyclization *in vivo*. A) *E. coli* BL21(DE3)pLysS cells harboring plasmids pTCYC-L46W or pTPRE-L46W-mut were induced with IPTG (+) or grown in the absence of IPTG (–). Total lysates of cells harbouring either pTCYC-L46W or pTPRE-L46W-mut were analyzed by SDS-PAGE and Coomassie Blue staining (left) and Western blot analysis using anti-His antibodies (right). Relevant protein species are indicated, and molecular masses are given in kDa. B) Native PAGE analysis of soluble protein extracts derived from induced cells harboring the indicated plasmids. Gels were either stained with Coomassie Blue or subjected to Western blotting with anti-His antibodies. The control lanes show lysate from cells not overexpressing ACP-related proteins. Asterisk in B indicates a candidate protein that could correspond to cyclic ACP. This band was subjected to trypsin digestion and LC-MS/MS analysis (Figure 23). This figure is modified from Volkmann *et al*. (2010).

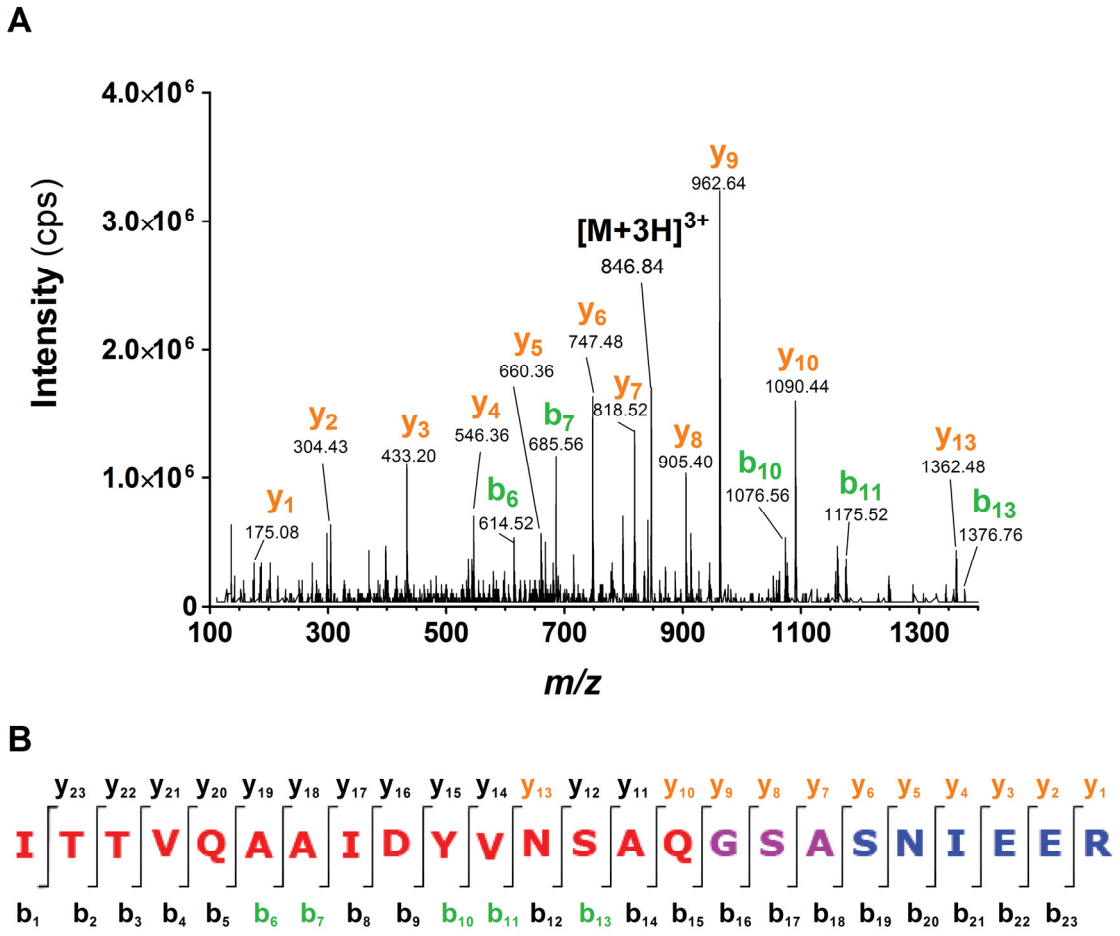

Figure 23 Confirmation of cyclization by tandem mass spectrometry (MS/MS) analysis. The native PAGE band suspected to be cycL46W (Figure 22) was subjected to trypsin digestion and analyzed by MS/MS. The MS/MS spectrum (A) shows the $[M+3H]^{3}+$ ion, as well as selected y (orange) and b (green) ions derived through its fragmentation. The peptide sequence with selected y and b ions indicated (B) represents the C-terminus (red) and N-terminus (blue) of ACP and the linker connecting the two after the split-intein reaction (purple). This figure is modified from Volkmann *et al.* (2010).

mediated *trans*-splicing occurs *in vivo* resulting in cycL46W.

3.3 Results: Biophysical Characterization of Cyclic ACP

To examine the effects of cyclization on ACP's structure *in vitro*, a variety of complementary biophysical methods were used.

3.3.1 Electrophoretic Mobility

Wild-type rACP exhibits anomalous SDS-PAGE migration (de la Roche et al., 1997), giving an apparent molecular mass of ~20 kDa as opposed to its ~9 kDa actual size. Attachment of acyl chains to rACP causes an increase in electrophoretic mobility, presumably due to a decreased hydrodynamic radius caused by fatty acyl induced folding. To examine what effect constraining the C- and N-termini may have on mobility, linL46W and cycL46W were subjected to SDS-PAGE. As expected, linL46W exhibited anomalous behaviour, with an apparent mass of ~20 kDa (Figure 24A). Not surprisingly, cycL46W migrated further than linL46W, but the extent of increased mobility is interesting. Based on previous data (de la Roche et al., 1997), acyl chains up to 10 carbons long increase electrophoretic mobility, with decanoyl-ACP exhibiting an apparent molecular weight of ~14 kDa. Constraining the termini of ACP caused an even larger increase in mobility to an apparent molecular weight of ~9 kDa, the expected mass of ACP (Figure 24A).

To examine this phenomenon further and to determine more accurately the effect of cyclization on hydrodynamic size in the absence of SDS, both proteins were subjected to conformationally sensitive native PAGE (Rock et al., 1981). As with SDS-PAGE, linL46W migrated as expected, with apo-linL46W (Figure 24B; upper band) being

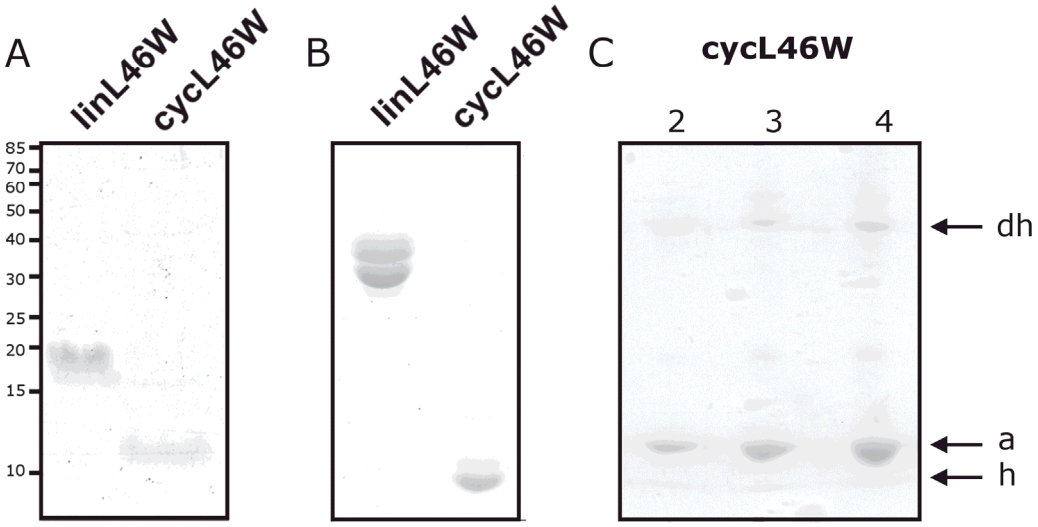

Figure 24 Migration of purified linL46W and cycL46W in SDS-PAGE (A), native PAGE (B) and Urea PAGE (C). Molecular masses in A are given in kDa. (C) Suspected forms of cycL46W are indicated: apo (a), holo (h) and di-holo (dh). Numbers above each lane in C indicate total protein in micrograms.

separated from the faster moving holo form (Figure 24B; lower band). Conversely, cycL46W migrates much further than linL46W and the two forms (apo and holo) could not be easily distinguished (Figure 24B) under native PAGE conditions.

Many methods were examined for their ability to better separate apo- and holo-forms of cycL46W, including using a larger gel system (as opposed to the PROTEAN mini-gel system normally used), increasing acrylamide percentage to 30%, increasing temperature and adding varying concentrations of urea in the gel. I found that adding 6 M urea to the normal native PAGE conditions allowed for the separation of apo and holo forms of cycL46W (Figure 24C). This separation can be enhanced slightly by conducting urea PAGE at 37ºC.

3.3.2 Mass Spectrometry

3.3.2.1 Charge State Distribution Analysis

Electrospray ionization mass spectrometry (ESI-MS) was used to analyze the charge state distribution (CSD) produced upon injection of purified linL46W or cycL46W. It has been demonstrated previously that the CSD of a protein is dependent on the folding state of the protein as it enters the gas phase (Konermann & Douglas, 1997). Unfolded proteins typically exhibit a broad CSD with highly charged states, while folded proteins have a narrow CSD range with relatively low charge states. Purified apo-linL46W produced a broad CSD with charge states ranging from 6^+ to 9^+ (Figure 25A), indicative of an unfolded protein. In contrast, cycL46W demonstrated a narrow distribution with only charge states of 6^+ and 7^+ present (Figure 25B), suggesting a more

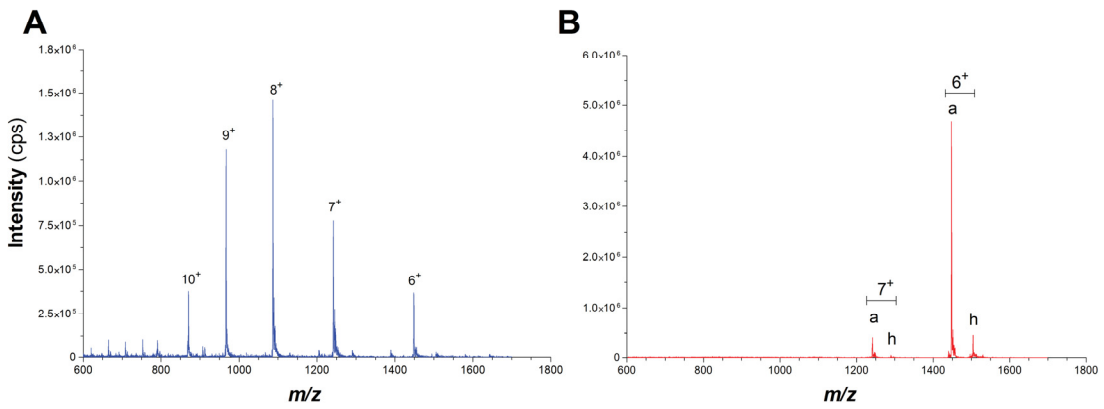

Figure 25 MS spectra for linear and cyclic derivatives of L46W. EMS spectra were obtained in positive ion mode nanospray for intact linL46W (A; apo) and cycL46W (B; mixture of apo and holo forms). The charge of relevant peaks is indicated and the apo (a) and holo (h) forms noted for cycL46W. Masses of the linL46W and the apo/holo forms of cycL46W are 8,685 Da and 8,667/9,007 Da, respectively.

stable folded conformation under these conditions. Both linL46W and cycL46W have the expected mass: apo-linL46W (8,685 Da), apo (8,667 Da) and holo (9,007 Da) cycL46W.

3.3.2.2 CycL46W as a Substrate for AcpS

LC-MS was used to determine whether cycL46W is a substrate for AcpS *in vitro*, based on the expected mass increase upon phosphopantetheine attachment. LinL46W and cycL46W were incubated in the presence or absence of AcpS and extracted ion chromatograms (XICs) were created from the resulting LC-MS data (Figure 26) using Analyst® software. Without addition of AcpS, linL46W had an apo:holo ratio of 1:1.6 (Figure 26A) while cycL46W had a ratio of 2.4:1 (Figure 26B), indicating that cycL46W is less modified than its linear counterpart by endogenous AcpS *in vivo*. After AcpS treatment, these ratios changed to 1:4.5 (Figure 26C) and 1:5.6 (Figure 26D) for linL46W and cycL46W, respectively, indicating that cycL46W is a substrate for the enzyme. Interestingly, cyclization reversed the retention time of the holo and apo forms of ACP on the reversed-phase column relative to linL46W, in which the holo form has an increased retention time compared to the apo form (Figure 26).

3.3.3 Steady-State Trp Fluorescence

3.3.3.1 Effect of Mg^{2+} on linL46W *versus* cycL46W

As mentioned above, the L46W mutation was chosen as a template for cyclization because folding of this ACP causes a pronounced fluorescence blue shift upon addition of Mg^{2+} (Gong et al., 2008). Similarly, in the absence of Mg^{2+} linL46W had a peak fluorescence of ~355 nm which shifted to ~310 nm upon addition of 10 mM Mg^{2+} (Figure 27A). In contrast, cycL46W had a peak fluorescence of ~310 nm in both the

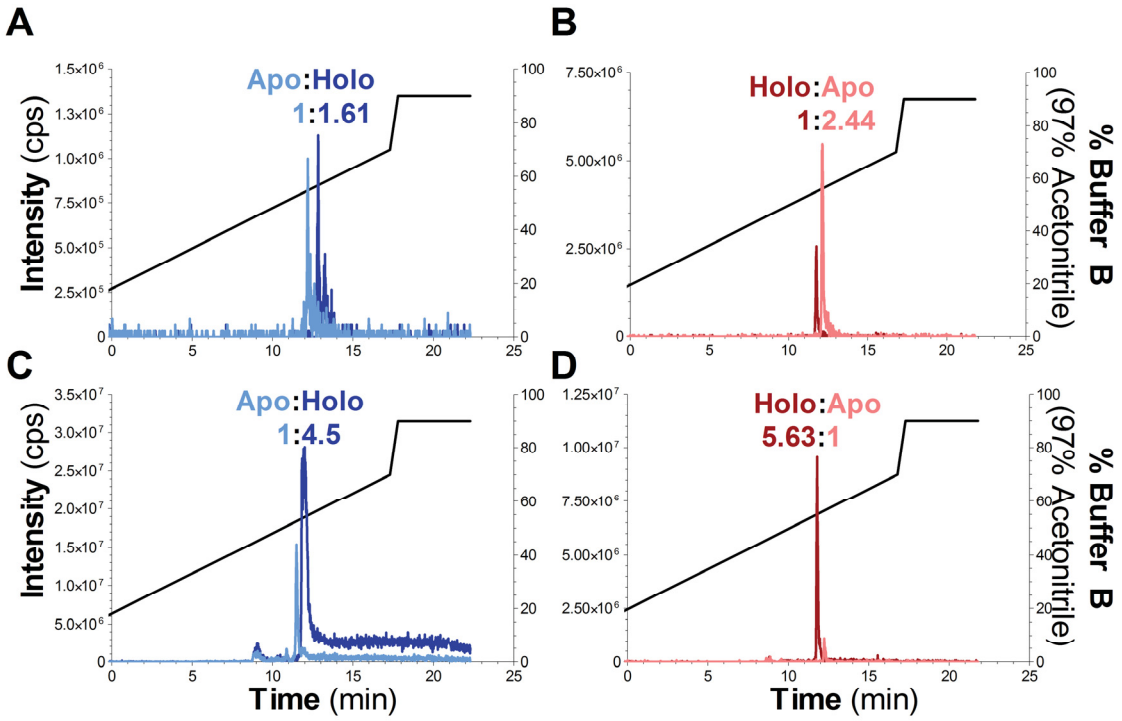

Figure 26 Extracted Ion Chromatogram (XIC) depicting apo:holo ratios for the linear and cyclic derivatives of L46W. Mass spectra were recorded and XICs created for linL46W (A, C) and cycL46W (B, D) in the absence (A, B) or presence (C, D) of prior treatment with AcpS. The apo:holo ratio indicated was determined by integrating each peak. For AcpS reaction conditions see Section 2.4.4.3. Masses of the apo/holo forms of linL46W and cycL46W are 8,685/9,025 Da and 8,667/9,007 Da, respectively. These data are representative of multiple experiments performed on different days.

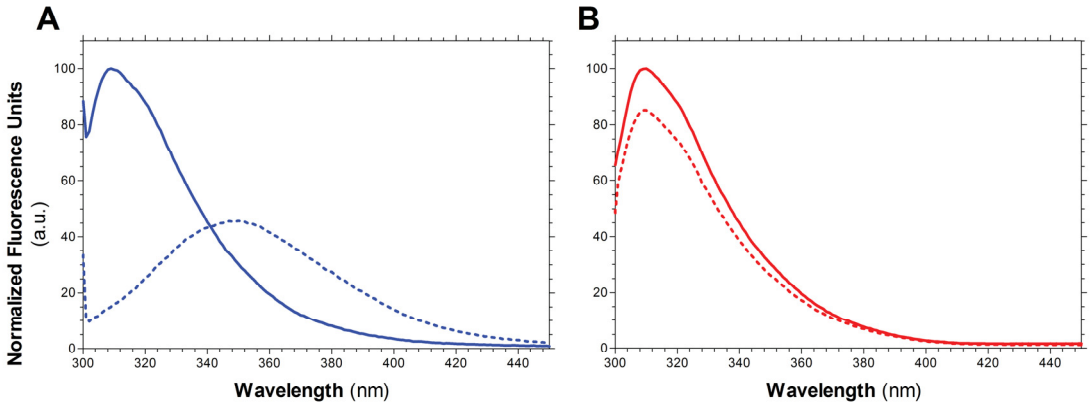

Figure 27 Intrinsic tryptophan fluorescence for linear and cyclic derivatives of L46W. Trp fluorescence spectra (λ_{ex} = 296 nm) for linL46W (A) and cycL46W (B) were recorded in the presence (solid lines) and absence (dotted lines) of 10 mM Mg^{2+}.

absence and presence of Mg^{2+} (10 mM) (Figure 27B), indicating that it is folded independently of Mg^{2+}.

3.3.3.2 Effect of Mg^{2+} on Trp-substituted *Plasmodium falciparum* ACP Mutants

For further validation of Trp probes as an indicator of ACP conformation, Trp replacement mutations were created in the ACP from the malaria parasite *P. falciparum*. This ACP has been shown to exhibit increased stability relative to most bacterial ACPs (Modak et al., 2007). Two Trp-substituted *P.f.*ACPs were created: A104W and L105W. These positions were chosen based on an alignment of wild-type *P.f.*ACP and *V. harveyi* rACP (Figure 28) and correspond to A45 and L46 in *V. harveyi* rACP. Previous work from our lab (Gong et al., 2008) has shown that the A45W probe on Helix II is exposed to the aqueous environment regardless of the presence or absence of Mg^{2+} (as A45 is exposed in folded ACP structures). On the other hand, it would be expected that the L105W substitution would have an emission spectra indicative of a solvent exposed Trp in the absence of Mg^{2+} (as bacterial ACP is unfolded) which shifts to a solvent-shielded Trp in the presence of Mg^{2+} (as L46 is buried in folded bacterial ACP structures).

The Trp in A104W (Figure 29A, blue lines) is more solvent exposed than in L105W (Figure 29B, red lines). This is not wholly unexpected based on *P.f.*ACP structures determined to date (Figure 30). In the absence of Mg^{2+}, A104W exhibits a spectrum indicative of solvent exposed Trp (Figure 29A, blue dotted line; peak = 338 nm) that becomes slightly more exposed upon addition of Mg^{2+} (Figure 29A, blue solid line; peak = 339 nm). L105W, on the other hand, exhibits an emission spectrum indicating a solvent-shielded Trp in the absence of Mg^{2+} (Figure 29B, red dotted line;

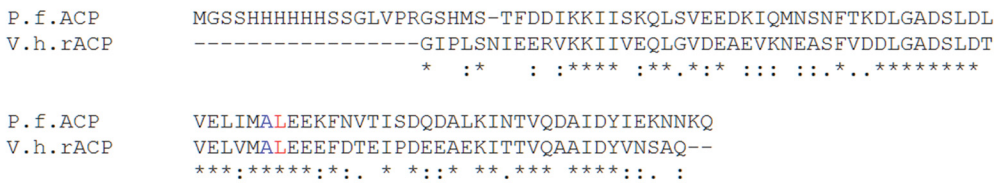

Figure 28 Sequence alignment of *P.f.*ACP and *V. harveyi* rACP. Clustal Omega (European-Bioinformatics-Institute. *Clustal Omega*.) was used to produce the above figure, after which coloring was added using Adobe Photoshop CS6. Ala 45 (blue) and Leu 46 (red) selected; numbering based upon *V. harveyi* rACP.

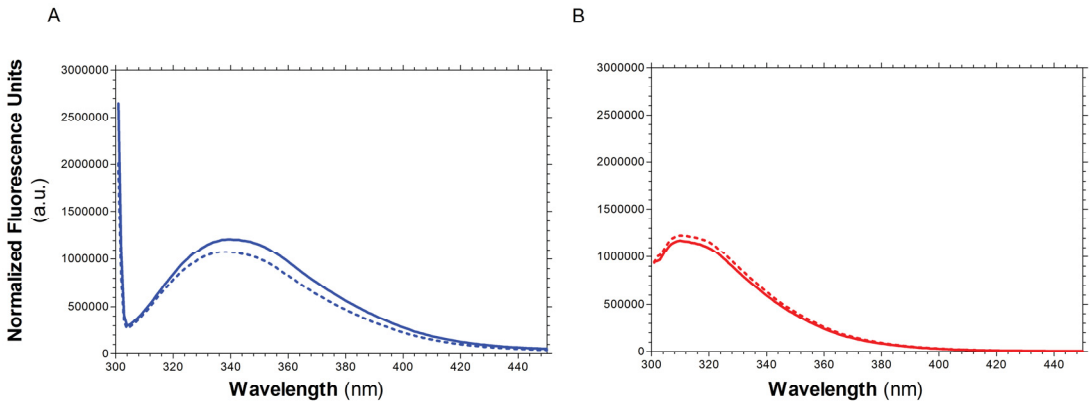

Figure 29 Intrinsic tryptophan fluorescence for Trp-substituted derivatives of *P.f.*ACP. Trp fluorescence (λ_{ex} = 296 nm) spectra for A104W (A) and L105W (B) were recorded in the presence (solid lines) and absence (dotted lines) of 10 mM Mg^{2+}.

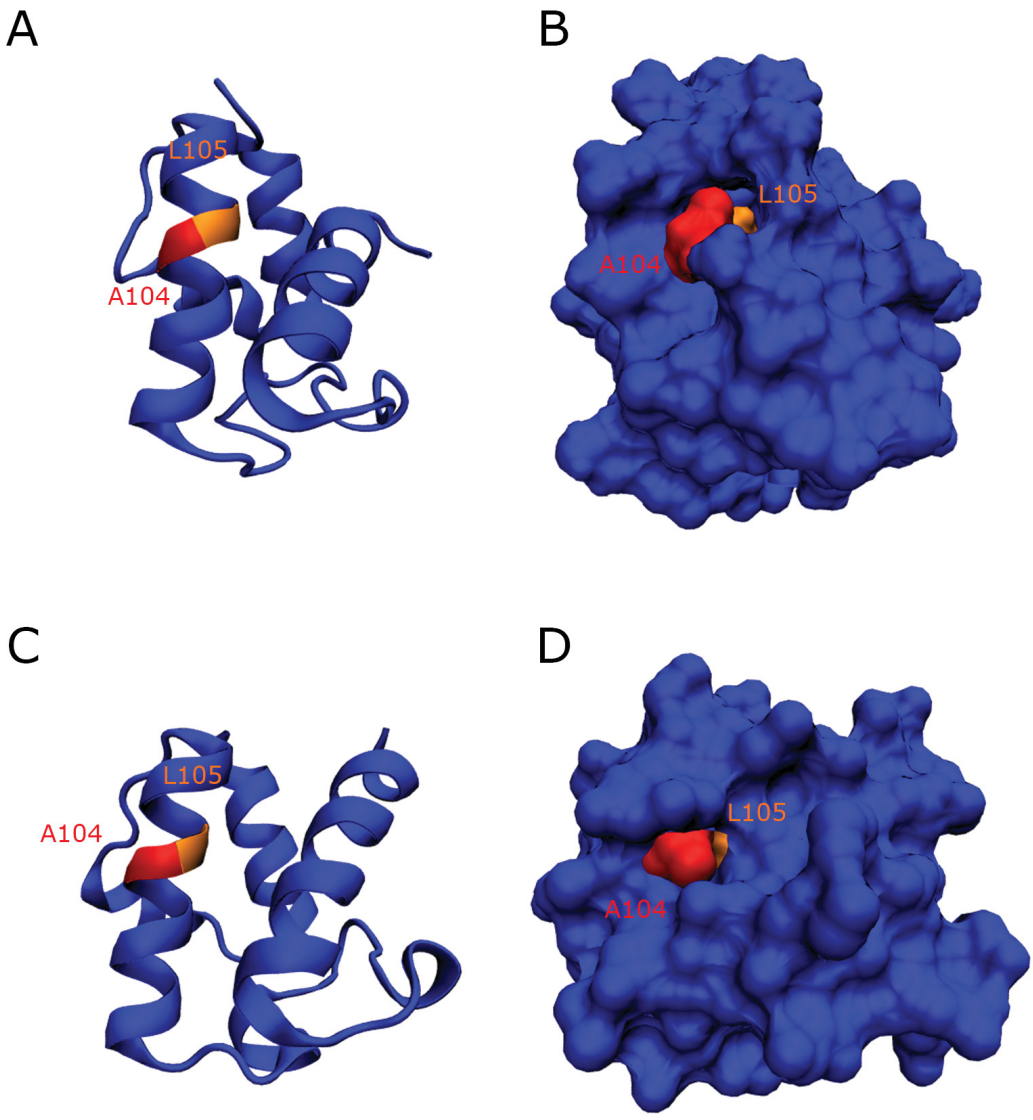

Figure 30 Diagram of selected *P.f.*ACP structures. "NewCartoon" (A, C) and "Surface" (B, D) representations of wild-type *P.f.*ACP (A, B = PDB ID 2FQ0 (Sharma et al., 2006); C, D = PDB ID 3GZM (Gallagher & Prigge, 2010)). Ala 104 (red) and Leu 105 (orange) are highlighted. PDB ID 2FQ0 is an ensemble of structures; a representative structure is depicted here. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

peak = 310 nm), which doesn't change upon the addition of Mg^{2+} (Figure 29B, red solid line; peak = 310 nm), indicating that *P.f.*ACP is folded even in the absence of magnesium.

3.3.3.3 Thermal stability of cycL46W *versus* linL46W

Steady-state fluorescence was also used to determine the effect of constraining the termini of ACP on its thermal stability. The conformational stability of linL46W was measured only in the presence of Mg^{2+}, as it is already unfolded at room temperature in the absence of divalent cations. LinL46W in the presence of Mg^{2+} was found to be the least resistant to heat-induced denaturation, with an apparent melting temperature midpoint (T_M) of approximately 65°C (Figure 31, solid blue line). Interestingly, cycL46W in the absence of Mg^{2+} was even more resistant to denaturation than linL46W in the presence of Mg^{2+}, indicating that cyclization stabilizes ACP's folded state more than Mg^{2+} binding (Figure 31). Addition of Mg^{2+} to cycL46W further stabilized the protein, but the apparent T_m could not be estimated due to restrictions on the upper temperature limit achievable (Figure 31, red lines; dotted and solid for absence and presence of Mg^{2+}, respectively).

3.3.4 Circular Dichroism (CD)

It has been shown previously that addition of Mg^{2+} to wild-type *V. harveyi* ACP increases its α–helical content, as indicated by an increase in the absolute magnitude of the ellipiticity at 220 nm (θ_{220}) (Gong et al., 2008). Not surprisingly, linL46W exhibited a similar transition after addition of Mg^{2+} (Figure 32A). However, cycL46W exhibited a CD spectrum indicative of a folded ACP, even in the absence of Mg^{2+}. Addition of Mg^{2+}

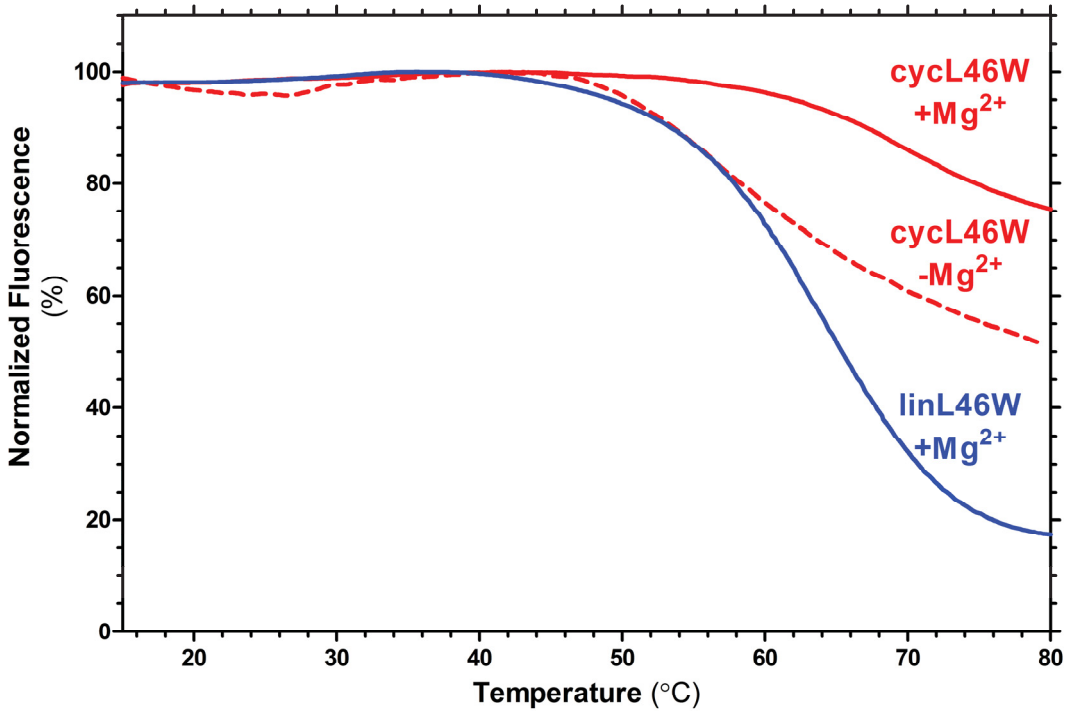

Figure 31 Thermal stability of linear and cyclic derivatives of L46W. LinL46W (blue) and cycL46W (red) were excited at 296 nm and the fluorescence intensity at 310 nm was measured as a function of temperature (15 – 80°C; 1°C/min) in the absence (dashed line) or presence (solid line) of 10 mM Mg^{2+}. This emission wavelength was chosen as it corresponds to the peak wavelength of a folded ACP (linL46W + Mg or cycL46W ±Mg).

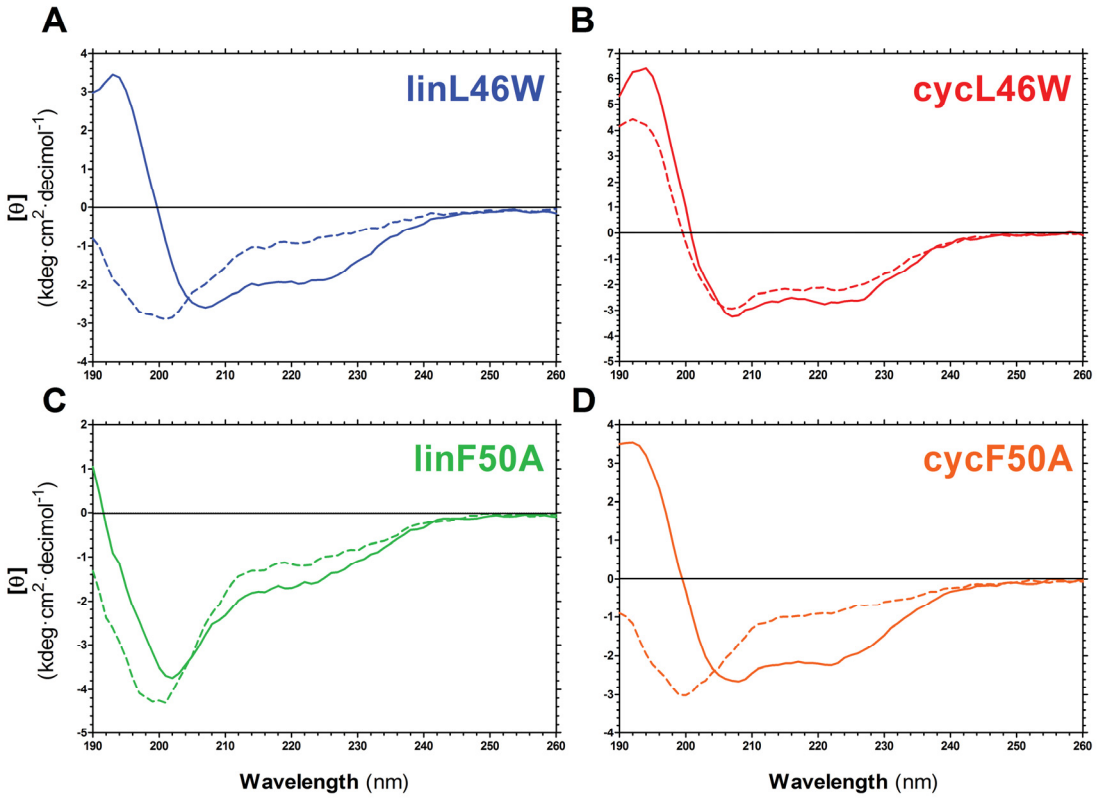

Figure 32 Circular dichroism (CD) analysis of linear and cyclic derivatives of L46W and F50A. CD spectra of linL46W (A), cycL46W (B), linF50A (C), and cycF50A (D) were measured in the presence (solid lines) and absence (dotted lines) of 10 mM Mg^{2+}.

caused a slight increase in θ_{220} (Figure 32B), which is consistent with Trp fluorescence and thermal denaturation results shown above.

Our previous work has revealed that mutation of Phe 50 in the helix II-III region of *V. harveyi* rACP renders the protein incapable of folding even in the presence of Mg^{2+}, and decreases its activity as a substrate of acyl-ACP synthetase by 95% (Flaman et al., 2001). CycF50A was constructed to examine the possible influence of cyclization on this structural defect. Interestingly, cyclization of the F50A mutant fully restored its ability to fold in the presence of Mg^{2+} (Figure 32C and D).

3.4 Results: Molecular Dynamics

Molecular dynamics (MD) simulations of linL46W and cycL46W were performed over an 11 ns period (1 ns minimization; 10 ns equilibration; as described in Section 2.5.2). As mentioned previously, *V. harveyi* ACP, the proteins being modeled, share 86% sequence identity with *E. coli* ACP, the PDB structure used to create the aforementioned ACP structures. A plot of RMSD *versus* time indicates that the two simulations approach equilibrium (Figure 33). No substantial changes to the overall four-helix structure of ACP were observed (Figure 34A); however, the distance between the N- and C-termini of linL46W increased from 11.3 Å at the start of the simulation to 15.0 Å at the end of the simulation (Figure 35, blue line). Conversely, the distance between the same atoms in cycL46W remains fairly constant (11.7 Å to 11.3 Å) during the course of the simulation (Figure 35, red line). Interestingly, cycL46W's "termini" are farther apart than linL46W at the start of the simulation (Figure 35, inset). This is likely due to the restricted contraction movement during minimization. As shown previously (Chan et al., 2008), the regions of greatest flexibility in linL46W include the two termini, the start/end of the

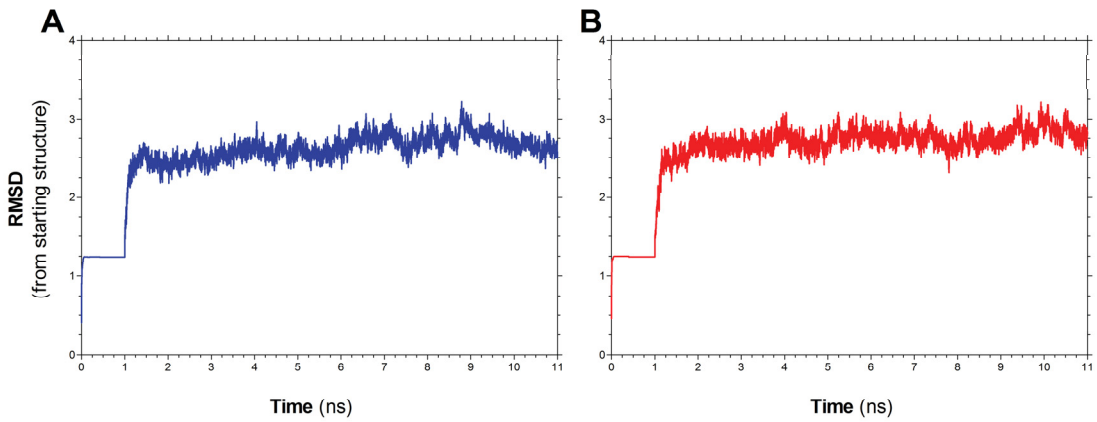

Figure 33 RMSD for backbone atoms of linear and cyclic ACP simulations. RMSD was calculated from the starting structure for linL46W (A) and cycL46W (B). Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

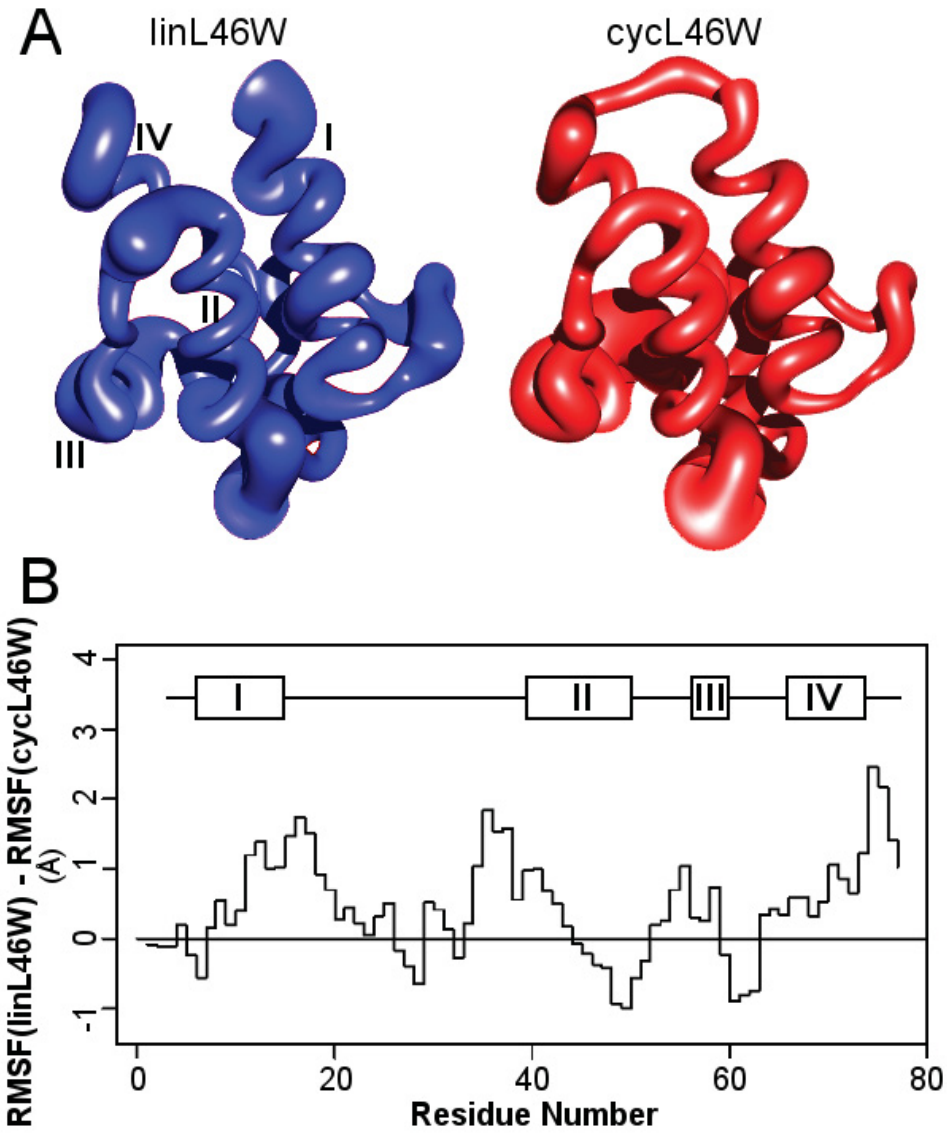

Figure 34 Molecular dynamics simulations of linear and cyclic L46W ACP. Sausage representations (A) of linL46W (*left*) and cycL46W (*right*), where the thickness of the polypeptide chain represents the RMSF of C_α atoms during the final 9 ns of simulation. A scaling factor of 1⁄2 was used, and helices I-IV are indicated. Difference plot (B) of C_α atom RMSF values of linL46W *versus* cycL46W as a function of residue position in the chain. RMSF values measure the movement of a particular atom over the course of the simulation and is calculated in Angstroms. The locations of helical segments of the starting *E. coli* structure (Wu et al., 2009) is shown *above* the plot. Prior to analysis, linL46W and cycL46W structures were aligned using UCSF Chimera (see Appendix 1). Images in A were created using MOLMOL 2.6.0 as directed in Appendix 32. Graph in B was created using R Programming Language (R-Project. *R Programming Language*.) with the help of Dr. Christian Blouin.

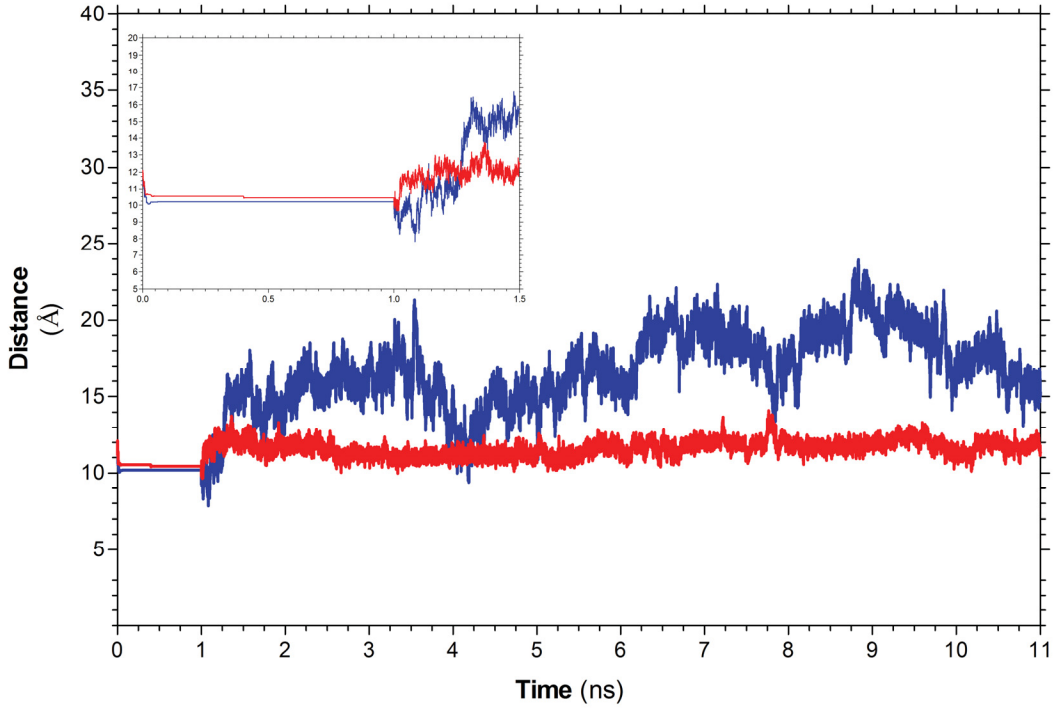

Figure 35 Distance between the terminal residues 1 and 77 over the course of the MD simulation for linear and cyclic derivatives of L46W. Distances measured for linL46W (blue) and cycL46W (red) using VMD as directed in Appendix 21. Inset shows the first 1.5 ns of simulation (1 ns of minimization/0.5 ns of equilibration).

long loop connecting helices I and II, and the region between helix II and the end of helix III (Figure 34B). Cyclization caused a stabilization of these areas and increased the mobility of the end of helix II and the loop connecting helices III and IV (Figure 34B).

3.5 Discussion

Due to the large number (>30) of its interacting partner enzymes (Butland et al., 2005), it is thought that ACP must be conformationally dynamic to carry out its essential roles in fatty acid and complex lipid synthesis. Indeed, ACP exhibits many of the properties of natively unfolded proteins (Byers & Gong, 2007), where structural flexibility and the ability to interact with multiple partners at low affinity is selected over high-affinity binding with a specific enzyme partner. Over a dozen years ago it was proposed that a reversible conformational change must occur to allow ACP to alternately sequester a fatty acyl chain within its hydrophobic binding pocket and then transfer it to the active site of an enzyme (Zhang et al., 2001).

The present study has explored the limits of ACP's conformational flexibility and its functional relationship (see below) through intein-mediated linking of its N- and C-termini to form a cyclic protein. The free termini are normally highly mobile in solution (Chan et al., 2008), yet are close together in most known structures of ACP, and no evidence is presented here that cyclization causes major rearrangements to the four-helix bundle conformation of folded ACP, or that function is severely compromised by the constraint of the joining the N- and C-termini.

Cyclization *in vivo* was very rapid and efficient as evident by the absence of the precursor protein (preL46W) on SDS-PAGE, and was highly specific towards intra-molecular cyclization as head-to-tail polymerized ACPs (i.e. the result of inter molecular

intein reactions) were non-existent in total cell lysates separated by SDS-PAGE. Moreover, cyclization of the folding deficient ACP mutant F50A (Flaman et al., 2001) suggests that this process is driven by the high affinity of the intein fragments for one another rather than any propensity of ACP to adopt a folded conformation

My work clearly demonstrates that cyclization of *V. harveyi* ACP causes a stabilization of its folded helical conformation, as suggested by MD, specifically RMSF measurements, and confirmed by various complementary biophysical methods. Intrinsic Trp fluorescence and far-UV circular dichroism both demonstrated that cyclized ACP adopts a folded conformation under conditions where linear ACP does not (e.g. in the absence of Mg^{2+}), and cyclized ACP was more resistant than linear ACP to thermal denaturation. This stabilization is not wholly unexpected, as constraining the normally free and mobile ACP termini should in fact decrease the conformational entropy of the unfolded state relative to the folded state. In the case of cycF50A, this loss of entropy is not, by itself, sufficient to overcome the folding deficiency in the absence of divalent cations, but does permit folding to occur upon addition of Mg^{2+}.

The effect of cyclization on the CSD of ACP determined by ESI-MS is particularly interesting. In previous work, I demonstrated that *V. harveyi* ACP, like linL46W, exhibits a broad CSD irrespective of its solution state conformation (manipulated by fatty acylation or charge neutralization) (Murphy et al., 2007). We speculated that unlike other proteins (e.g. cytochrome C), which are trapped in their solution phase conformation during gas phase ionization (Konermann & Douglas, 1997, Murphy et al., 2007), ACP's dynamic nature might allow it to unfold during the ionization process. The narrow CSD recorded for cycL46W in the present work suggests

that, in contrast to fatty acylation or charge neutralization, constraining the ACP's termini prevents unfolding on the microsecond to millisecond timescale of the ESI process. This is consistent with the additional thermal stability that cyclization confers to ACP, which is greater than that of magnesium on linear ACP. NMR and molecular dynamics could provide more information about the basis of additional stability conferred by cyclization. MD simulations performed here reveal specific regions (e.g. the helix I-loop-helix II region and the C-terminus) where motion is curtailed relative to the linear form; further MD experiments with Mg^{2+}-ACP or acyl-ACP, or with *P.f.*ACP, might provide additional insight. It should be noted that cyclization also stabilizes *E. coli* DnaB against unfolding under acidic conditions (Watt et al., 2007).

Somewhat surprisingly, linking the N- and C-termini of ACP has only a limited impact on its function *in vivo*. As part of the study described in this thesis, the effect of ACP cyclization on its ability to restore growth by complementation of an ACP-deficient *E. coli* strain (CY1681) was carried out in parallel by Gerrit Volkmann (Volkmann et al, 2010). The reader is referred elsewhere for a more complete description of that work (Volkmann et al., 2010) but the key experiment is shown in Figure 36. Briefly, while cyclization had a slightly negative effect on the ability of L46W to restore growth (compare panels B and C), cyclization actually enhanced *in vivo* complementation of the folding deficient F50A mutant (compare panels E and F). The latter effect correlates with the ability of Mg^{2+} to restore the helical content of cyclic, but not linear, F50A. Based on CD analysis of folding, it would be expected then that under physiological conditions of millimolar Mg^{2+}(Moncany & Kellenberger, 1981, Alatossava et al., 1985), linear F50A

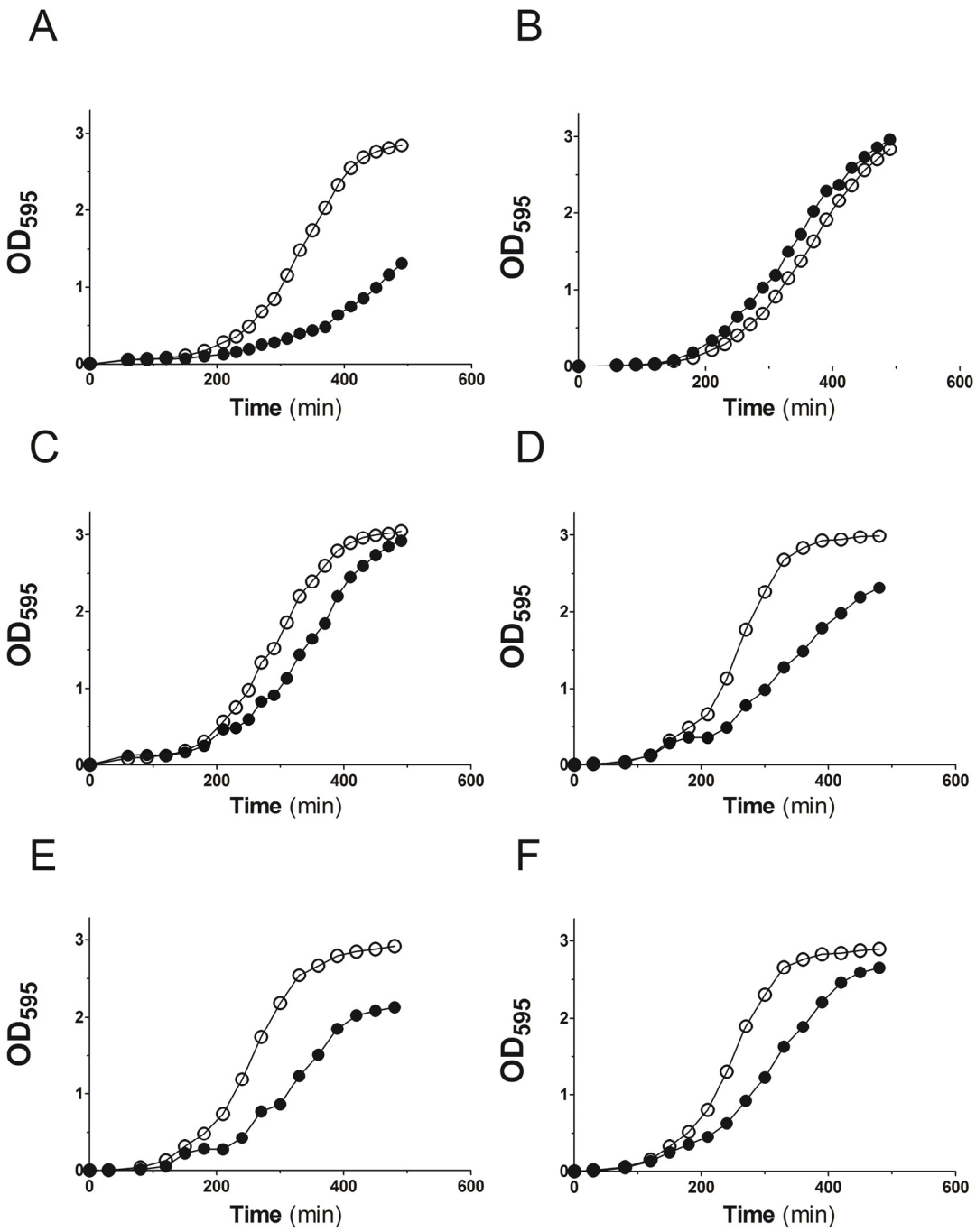

Figure 36 *In vivo* complementation assay. *E. coli* CY1861 was transformed with pMAL-derived plasmids encoding MBP (A), linL46W (B), cycL46W (C), preL46W^{mut} (D), linF50A (E) or cycF50A (F) under control of the IPTG-inducible p_{tac} promoter. Cells were grown in medium containing either arabinose (open circles) to allow expression of a synthetic *E. coli acp*P gene as a positive control, or IPTG plus glucose (closed circles), to induce the construct of interest and repress *E. coli acp*P. The optical density at 595 nm was recorded over a 6 h period. Growth curves representative of three individual experiments are shown.

would be denatured while its cyclic counterpart would be able to fold properly, consistent with the greater ability of cycF50A *versus* linF50A to restore *in vivo* complementation.

Taken together, the structural and functional data indicate that ACP must be able to adopt a folded conformation to be functional. This requirement may be linked to the importance of the acidic Helix II in the initial electrostatic interaction and recognition of ACP by partner enzymes, many of which exhibit complementary basic patches on their surface (Zhang et al., 2001). A second major conclusion is that complete unfolding of ACP (i.e. involving separation of its N- and C-termini) is not required for its function *in vivo*, at least for enzyme partners that are essential for growth in *E. coli*. Indeed, ACP function appears to be relatively tolerant of the status of its terminal residues, as reflected by the fact that bacterial ACPs with C-terminal TAP/SPA (Butland et al., 2005) and Flag (Battesti & Bouveret, 2009) fusion tags are functional *in vivo*. Similarly, mammalian ACP functions properly despite being an internal domain within the large fatty acid synthase (FAS) polypeptide chain (Leibundgut et al., 2008).

An increasing body of evidence has implicated the Helix II-III region in the dynamic conformational behaviour required for acyl chain binding and release (Wu et al., 2009). Very recently, Nguyen *et al*. (2014) have used both NMR spectroscopy and X-ray crystallography to provide molecular detail about the specific amino acid residues in the Helix II-III region involved in this process. Interestingly, despite the overall conformational stabilization conferred by cyclization of ACP, the MD simulations performed here suggest *increased* flexibility in cycL46W (*versus* linL46W) in this region as well as near the end of Helix II, which is implicated in enzyme recognition (Chan et

al., 2008). Perhaps this enhanced mobility partially compensates for any functional deficiencies caused by constraint of the N- and C-termini in cyclized ACP. In any case, ACP does not require full unfolding to pass its acyl chain to or from a partner enzyme and the required movement of Helix III and the surrounding loops can occur without separation of its termini, as schematically proposed in Figure 37.

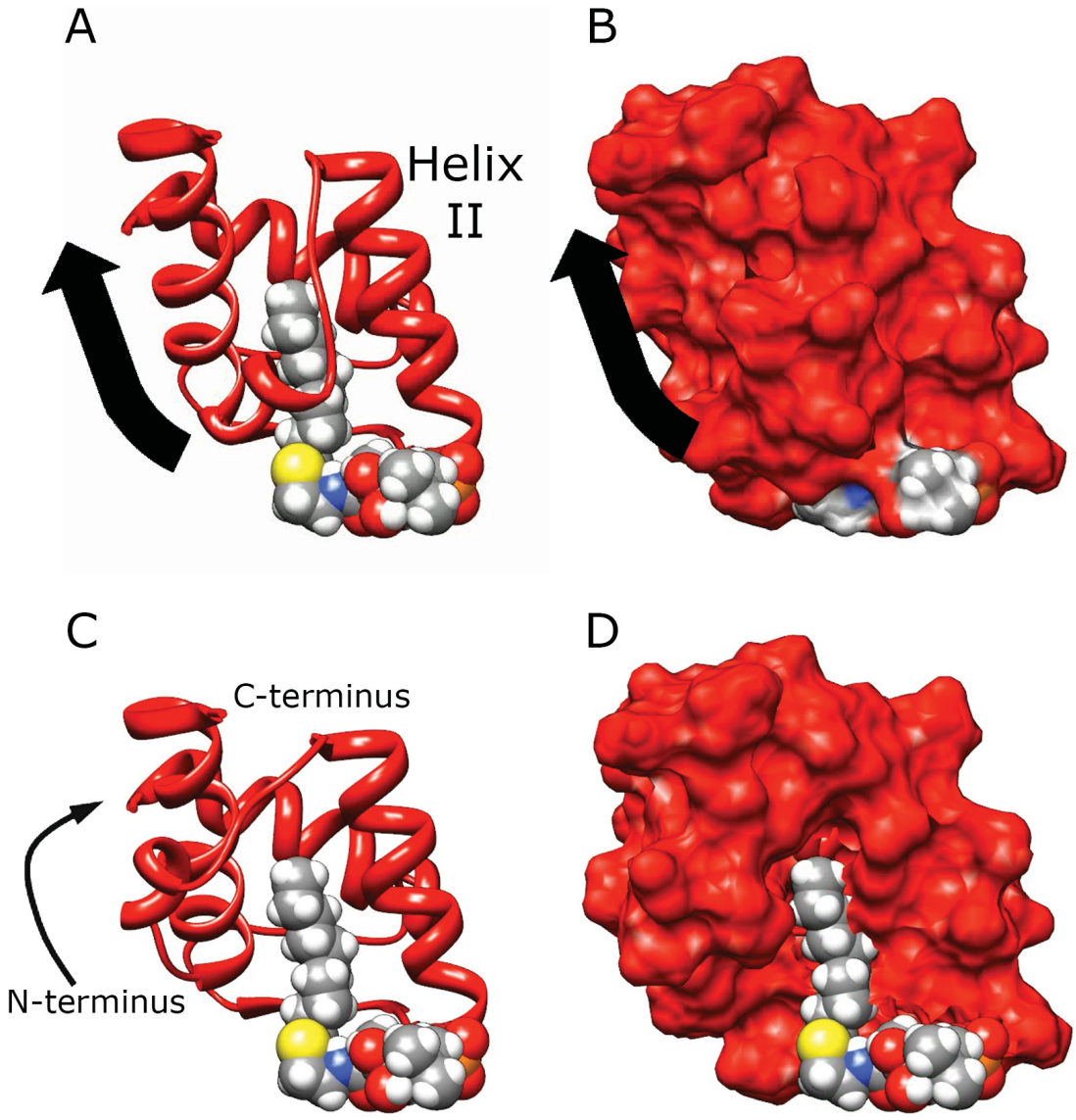

Figure 37 Potential movement in cyclic ACP to allow transfer of attached embedded acyl chain. "Ribbon" (A, C) and "Surface" (B, D) of ACP before (A, B) and after (C, D) proposed movement of Helix III and the connecting loops. This movement (indicated by black arrow) allows the termini to remain together (as is the case for cyclic ACP) but still expose the fatty acid for transfer to a partner enzyme. "Sphere" representation was used to display the decanoyl-phosphopantetheine moiety. Pictures were created from a movie created in Chimera 1.7. The spinach decanoyl-ACP PDB (PDB ID = 2FVF (Zornetzer et al., 2006)) was used to create this figure.

CHAPTER 4 TYROSINE-71 AS A PROBE OF ACP STABILITY AND CONFORMATION

4.1 Introduction and Rationale

Over the past dozen years, the Byers laboratory has accumulated an extensive collection of ACP mutants that have contributed to our understanding of the dynamic conformation and partner interactions of this central protein in fatty acid metabolism (Flaman et al., 2001, Keating et al., 2002, Gong & Byers, 2003, Gong et al., 2007, Gong et al., 2008) (Table 2). A subset of these mutants have Trp introduced as a conformationally-sensitive fluorescence probe (Gong et al., 2008); however, most mutants were constructed prior to this discovery and therefore lack Trp. Fortunately, wild-type *V. harveyi* ACP contains only one intrinsic tyrosine residue near its C-terminus (Tyr 71). If this tyrosine is sensitive to ACP's conformational state, then fluorescence could in principle be used to examine the conformational state and/or interactions of these mutants. The efficacy of Tyr 71 as a fluorescent probe of ACP folding was examined here using the previously characterized recombinant wild-type (rACP) and SA, SB, and SA/SB mutants in which one or both of the site A/B divalent cation binding sites have been modified (Gong et al., 2007).

4.2 Results: Fluorescence Analysis

4.2.1 Steady-State Tyr Fluorescence of rACP and Site A/B Mutants

Our previous work has shown that folding of L46W ACP due to Mg^{2+} binding at sites A and/or B can be monitored by Trp fluorescence (Gong et al., 2008); more recent experiments have revealed that other divalent cations such as Ni^{2+} and Zn^{2+} cause similar

effects at even lower concentrations (A. Murphy and D. Byers, unpublished observations; see Section 4.4). To determine whether Tyr 71 is sensitive to these conformational changes, baseline fluorescence emission spectra were initially measured for rACP, SA, SB, and SA/SB; these emission spectra exhibited similar excitation wavelengths (~300 nm) expected for tyrosine, but with different intensities for each ACP (Figure 38). Notably, the emission intensity was about two-fold greater for each of the SA and SB mutants compared to rACP, and almost double that again for the SA/SB mutant. These differing intensities likely reflect the differing environments of Tyr 71 in each ACP species, possibly related to the solvent accessibility of Tyr 71 and/or the quenching by neighbouring functional groups. Accessibility of the tyrosine was further explored *in silico* (see Section 4.3.1).

4.2.2 Titration of rACP, SA, SB, SA/SB with Divalent Cations

ACPs (rACP, SA, SB, SA/SB) were titrated with Mg^{2+}, Ni^{2+} and Zn^{2+} (Figure 39 – Figure 42). In each case, a saturable decrease in tyrosine fluorescence intensity was observed (Figure 43 – Figure 46). However, it became apparent that at least two effects confounded quantitative interpretation of these observations. First, it was discovered that progressive photobleaching of Tyr 71 was occurring during the experiment; to control for this, emission spectra were obtained for each ACP titrated with buffer (no metal). These spectra were used to correct the normalized decrease in peak fluorescence intensity graphs (below). Second, the metals appeared to have an effect on the fluorescence of tyrosine alone (data not shown). Thus, free L-Tyr was titrated with each cation. Data for these controls were recorded on multiple days and averaged prior to correcting ACP titration data.

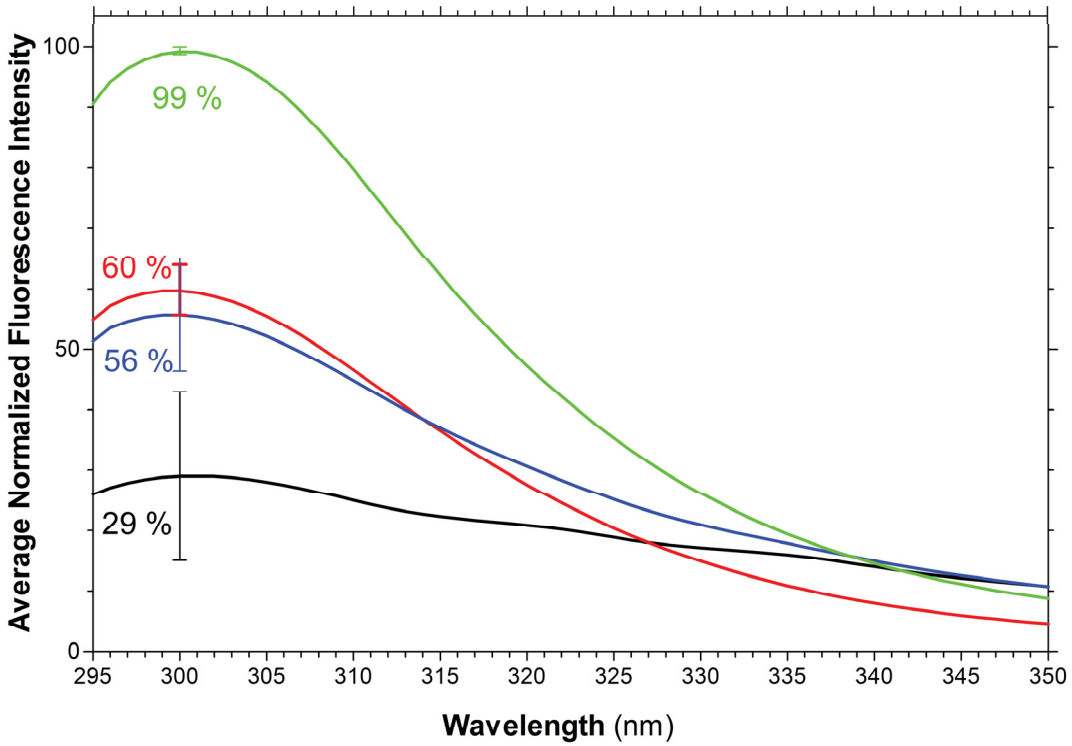

Figure 38 Intrinsic tyrosine fluorescence of rACP, SA, SB, and SA/SB. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded in the absence of divalent cations for rACP (black), SA (red), SB (blue), and SA/SB (green). To produce the figure, spectra were recorded on multiple days and normalized to that of SA/SB for that day; the spectra were then averaged across days to produce the given spectra. The average peak fluorescence of each ACP (as a % relative to SA/SB) is indicated, as well as the SD of 5 independent experiments.

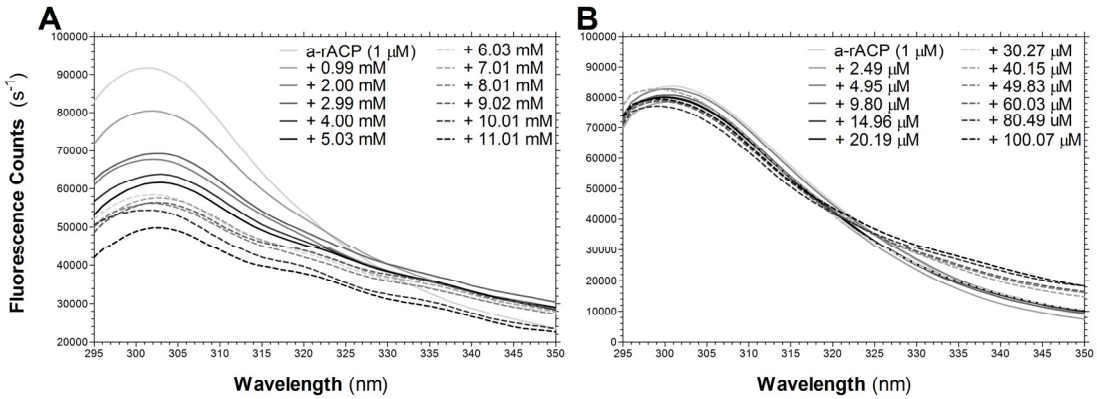

Figure 39 Changes in Tyr emission spectra of rACP titrated with divalent cations. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded in the absence and presence of increasing concentrations of Mg^{2+} (A) and Ni^{2+} (B). Concentrations of Mg^{2+} and Ni^{2+} are indicated on each graph. Spectra were recorded on multiple days, samples shown. Spectra for titration with Zn^{2+} lacked reproducibility across days (not shown).

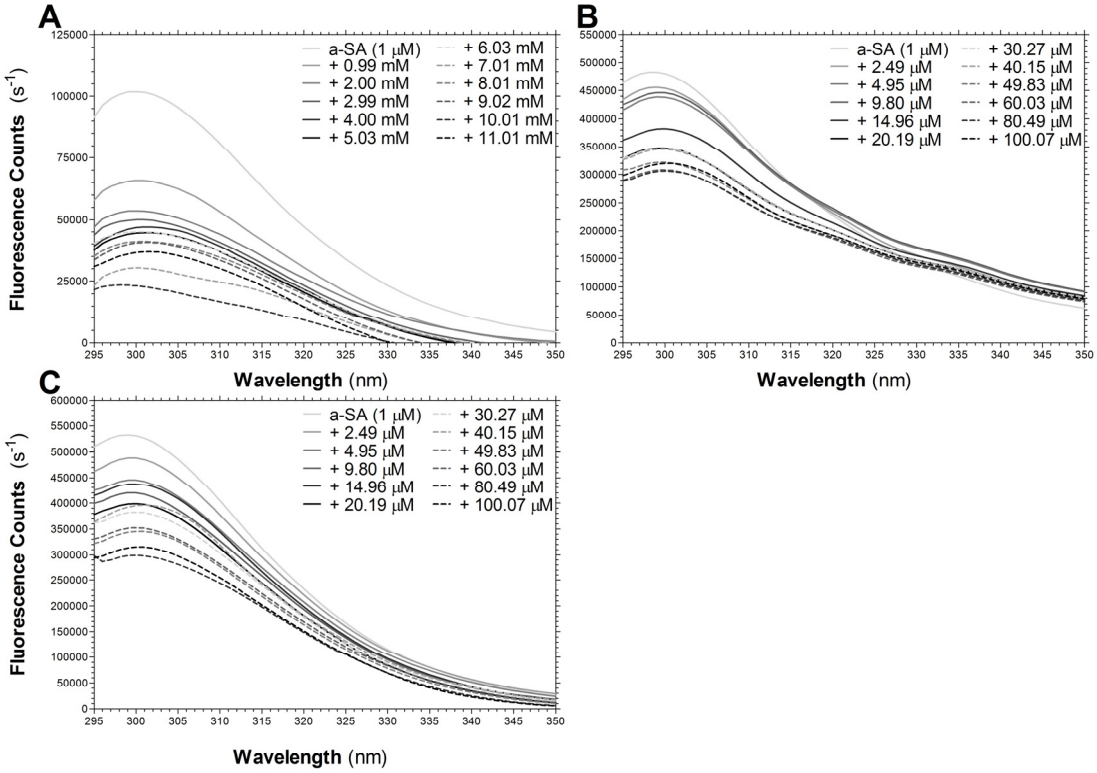

Figure 40 Changes in Tyr emission spectra of SA titrated with divalent cations. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded in the absence and presence of increasing concentrations of Mg^{2+} (A), Ni^{2+} (B), and Zn^{2+} (C). Concentrations of Mg^{2+}, Ni^{2+} and Zn^{2+} are indicated on each graph. Spectra were recorded on multiple days, samples shown.

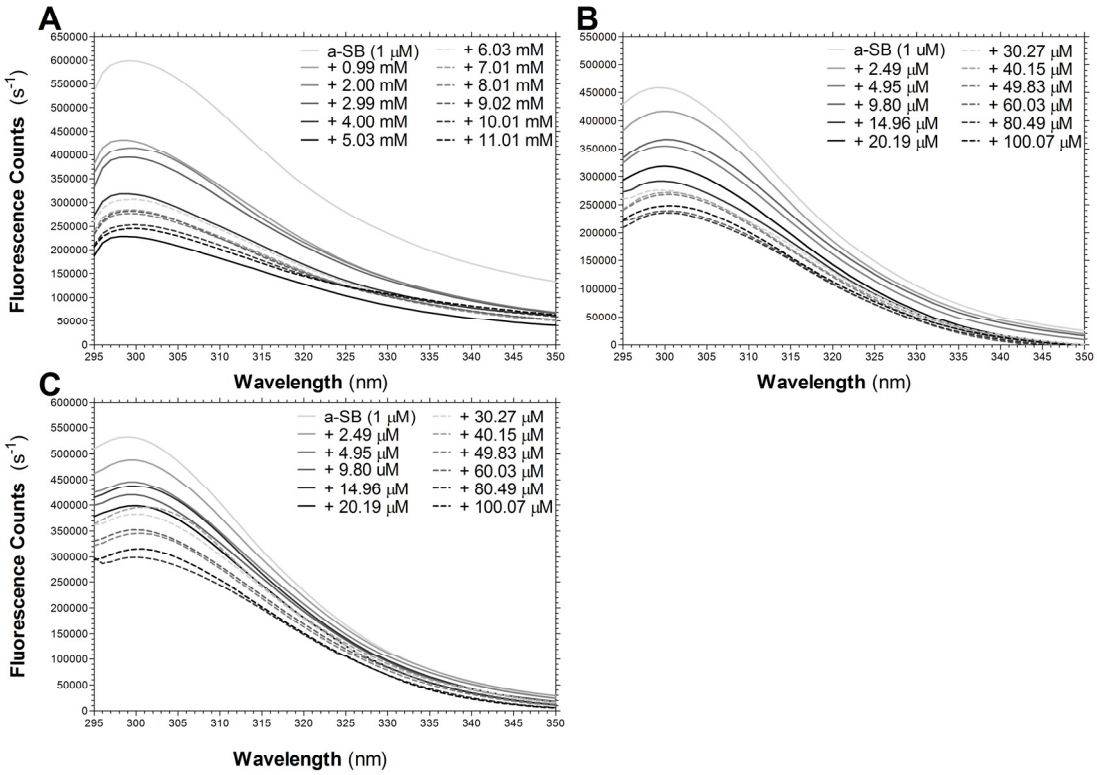

Figure 41 Changes in Tyr emission spectra of SB titrated with divalent cations. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded in the absence and presence of increasing concentrations of Mg^{2+} (A), Ni^{2+} (B), and Zn^{2+} (C). Concentrations of Mg^{2+}, Ni^{2+} and Zn^{2+} are indicated on each graph. Spectra were recorded on multiple days, samples shown.

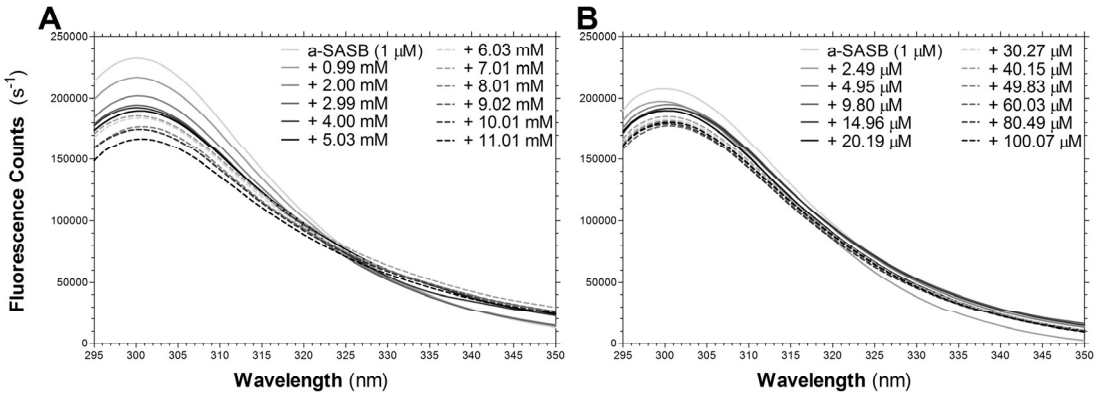

Figure 42 Changes in Tyr emission spectra of SA/SB titrated with divalent cations. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded in the absence and presence of increasing concentrations of Mg^{2+} (A) and Ni^{2+} (B). Concentrations of Mg^{2+} and Ni^{2+} are indicated on each graph. Spectra were recorded on multiple days, samples shown. Spectra for titration with Zn^{2+} lacked reproducibility across days (not shown).

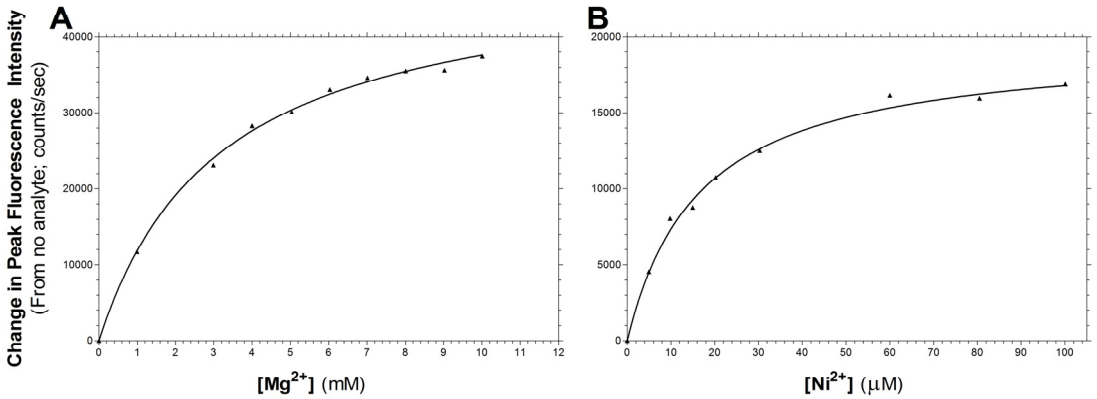

Figure 43 Change in tyrosine peak fluorescence intensity of rACP titrated with various divalent cations. Change in fluorescence intensity for rACP plotted *versus* the concentration of Mg^{2+} (A) and Ni^{2+} (B) for data from Figure 39. GraphPad™ Prism® 5 was used for determination of K_D value pre-python analysis.

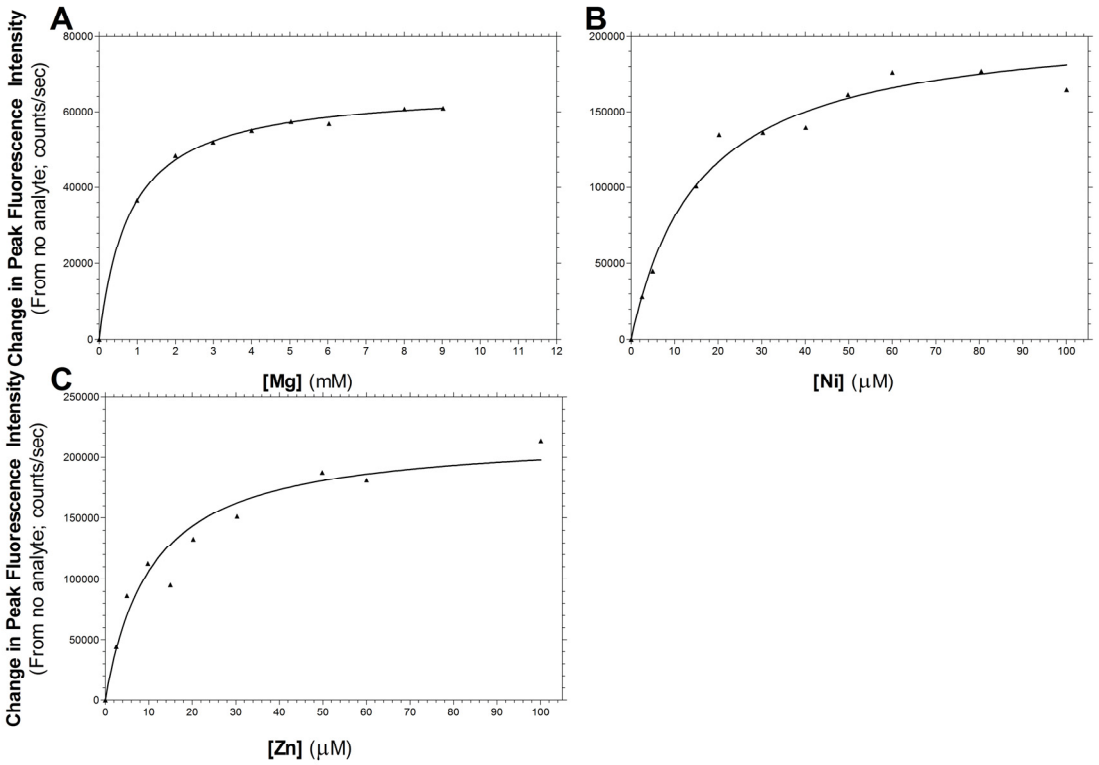

Figure 44 Change in tyrosine peak fluorescence intensity of SA titrated with various divalent cations. Change in fluorescence intensity for SA plotted *versus* the concentration of Mg^{2+} (A), Ni^{2+} (B), and Zn^{2+} (C) for data from Figure 40. GraphPad™ Prism® 5 was used for determination of K_D value pre-python analysis.

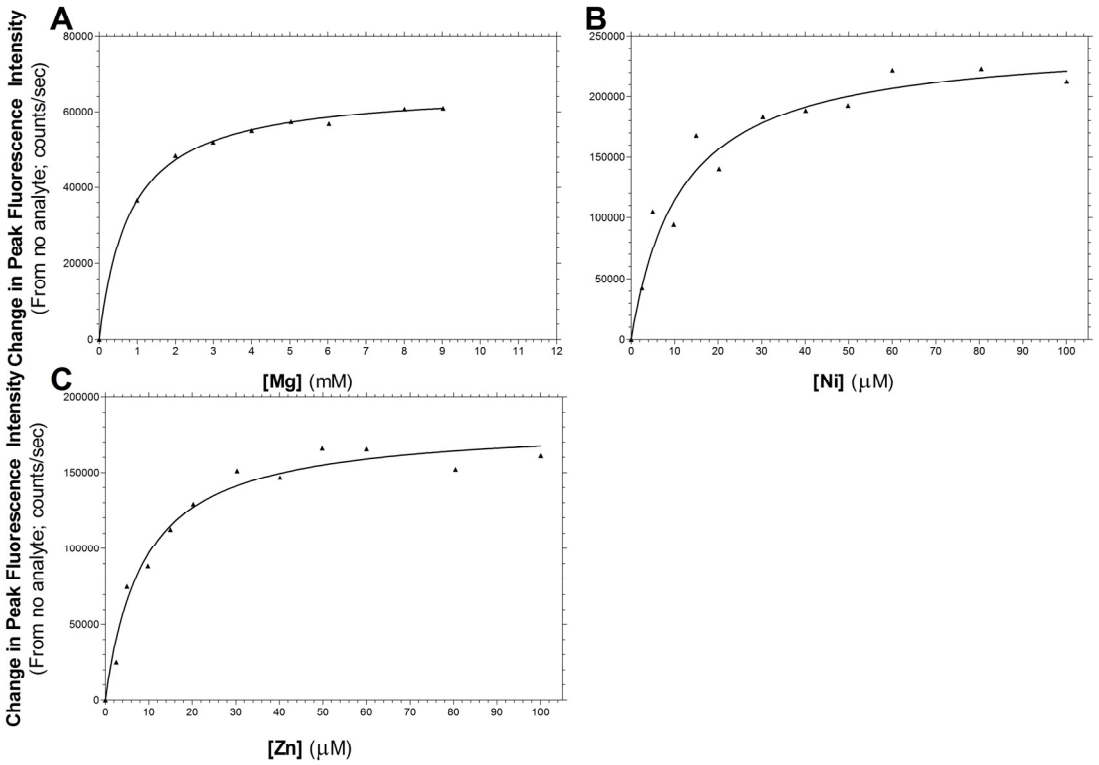

Figure 45 Change in tyrosine peak fluorescence intensity of SB titrated with various divalent cations. Change in fluorescence intensity for SB plotted *versus* the concentration of Mg^{2+} (A), Ni^{2+} (B), and Zn^{2+} (C) for data from Figure 41. GraphPad™ Prism® 5 was used for determination of K_D value pre-python analysis.

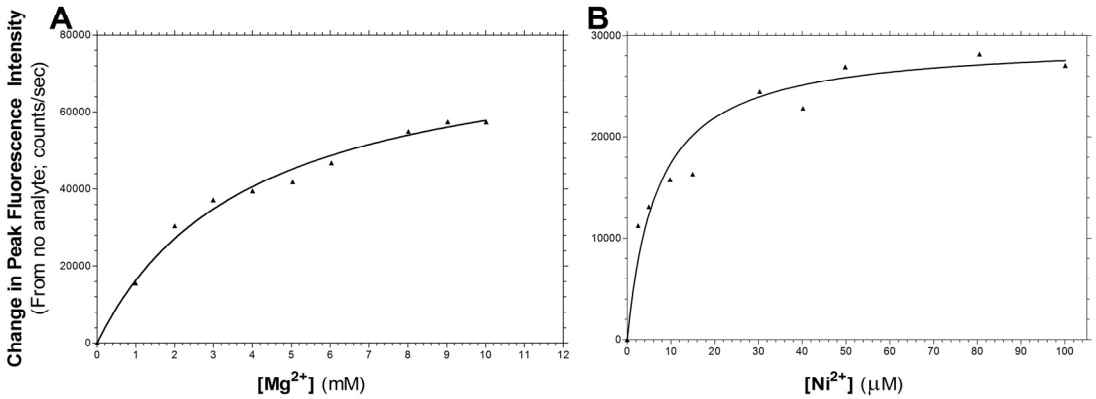

Figure 46 Change in tyrosine peak fluorescence intensity of SA/SB titrated with various divalent cations. Change in fluorescence intensity for SA/SB plotted *versus* the concentration of Mg^{2+} (A) Ni^{2+} (B), and Zn^{2+} (C) for data from Figure 42. GraphPad™ Prism® 5 was used for determination of K_D value pre-python analysis.

The combined and corrected results for ACP titration with metals, relative to the initial differing fluorescence intensity of each ACP, is shown in Figure 47. In general, these data reveal a small but significant decrease in Tyr 71 emission, which occurs over metal concentrations in the micromolar (for Ni^{2+} and Zn^{2+}) or low millimolar (for Mg^{2+}) range. Despite expected differences in metal affinity caused by neutralization of acidic residues in sites A and/or B, the percentage degree of metal-induced decrease in fluorescence intensity was not substantially different among the various ACPs. Nevertheless, these data were used to calculate apparent K_D values (Table 14), by plotting change in corrected fluorescence intensity *versus* metal concentration using GraphPad™ Prism®'s non-linear regression and Equation 1 ("Pre-Python Analysis"), and by using a Python program (Appendix 4) which uses non-linear regression (Equation 1) and iterative calculation ("Post-Python Analysis").

4.2.3 Stopped-Flow Trp and Tyr Fluorescence Analysis

The acquisition during this project of a PTI spectrofluorometer with stopped-flow capabilities enabled the potential experimental analysis of the conformational dynamics of ACP and its mutants. Initially, this was explored using the conformationally sensitive Trp probe of L46W (Figure 48A). Rapid mixing of L46W and Mg^{2+} resulted in a rapid (within 2 – 3 s) decrease in fluorescence emission at 350 nm; recording of spectra before and after mixing (Figure 48B) verified that this was due to the blue shift that accompanies Mg^{2+}-induced folding of ACP (Gong et al., 2008).

To elucidate the potential effect of mutations in the site A and site B regions of rACP, stopped-flow experiments using Tyr 71 as an intrinsic probe were carried out at various Mg^{2+} concentrations (representative data are shown in Figure 49). However,

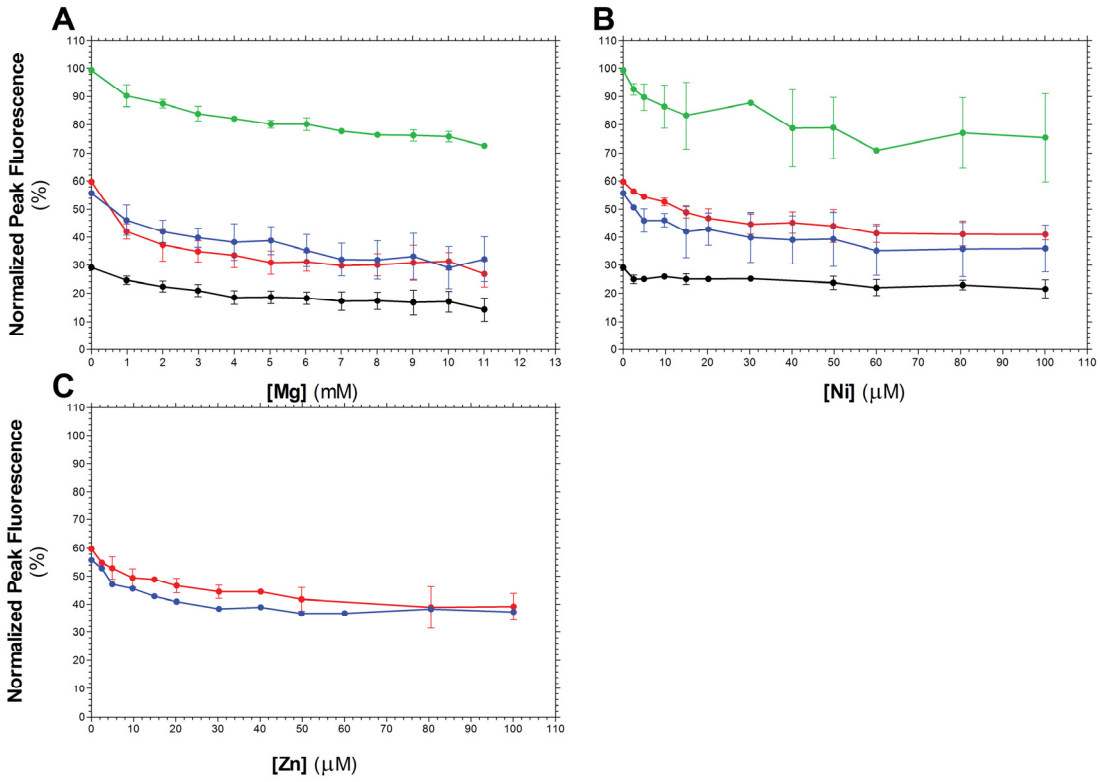

Figure 47 Corrected changes in fluorescence intensity of ACPs titrated with various divalent cations. Tyr fluorescence spectra (λ_{ex} = 280 nm) were recorded for rACP (black), SA (red), SB (blue), and SA/SB (green) in the absence and presence of increasing concentrations of Mg^{2+} (A), Ni^{2+} (B), and Zn^{2+} (C). To produce the figure, spectra were recorded on multiple days and normalized to SA/SB for that day. The normalized peak value for each analyte concentration was scaled based on Figure 38. Error bars represent standard deviation (n=3 to 7). All decreases in peak fluorescence were corrected for Tyr photobleaching (using buffer titrations) as described in Chapter 2. Spectra for rACP and SA/SB titration with Zn^{2+} lacked reproducibility across days (C, data not shown).

Table 14 Calculated K_D values for ACPs titrated with various divalent cations.

ACP	Metal	K_D Pre-Python Analysis[1]	K_D Post-Python Analysis[2]
rACP	Mg^{2+}	1.9 mM	2.9 mM
	Ni^{2+}	3.5 μM	20.9 μM
	Zn^{2+}	ND[3]	ND[3]
SA	Mg^{2+}	0.78 mM	0.9 mM
	Ni^{2+}	16.8 μM	16.2 μM
	Zn^{2+}	7.1 μM	14.0 μM
SB	Mg^{2+}	1.2 mM	2.7 mM
	Ni^{2+}	5.5 μM	9.1 μM
	Zn^{2+}	12.7 μM	9.0 μM
SA/SB	Mg^{2+}	1.2 mM	3.3 mM
	Ni^{2+}	4.7 μM	10.5 μM
	Zn^{2+}	ND[3]	ND[3]

[1] Values were determined by GraphPad™ Prism® 5.
[2] Values were determined using a python program I developed (Appendix 4).
[3] = Not determined; data not reproducible.

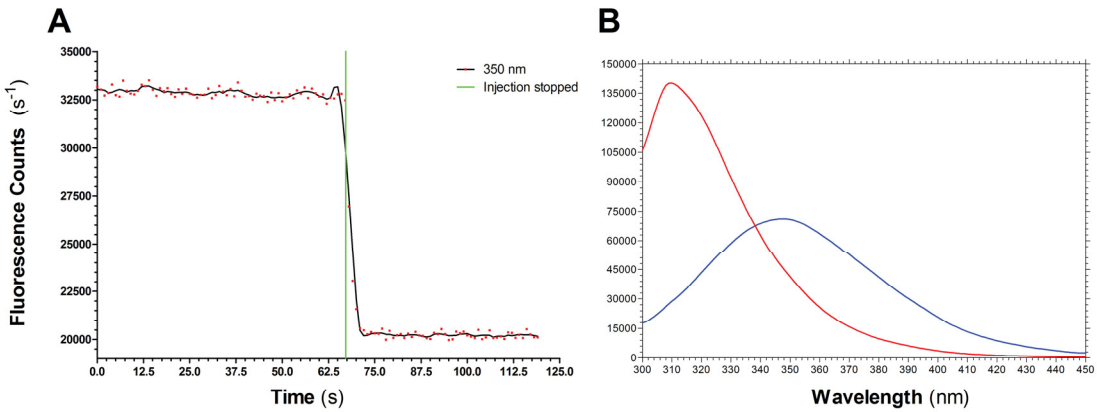

Figure 48 Steady-state Trp stopped-flow and emission spectra of *V. harveyi* L46W folding with Mg^{2+}. (A) Trp fluorescence (λ_{ex} = 296 nm) stopped-flow analysis of L46W, with emission monitored at 350 nm. (B) Emission spectra (λ_{ex} = 296 nm) were recorded before (blue line) and after (red line) rapid mixing of rACP (5 μM) with Mg^{2+} (10 mM). Data shown are representative of multiple experiments recorded on different days. Emission spectra (B) have been corrected for signal arising from the Raman band.

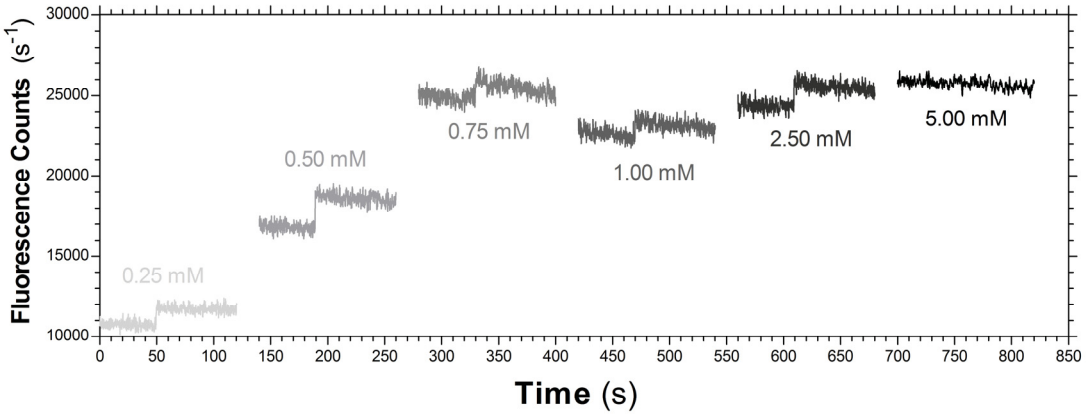

Figure 49 Tyr stopped-flow analysis of rACP. Fresh ACP samples (5 μM) were rapidly mixed with successively increasing concentrations of Mg^{2+} and fluorescence was monitored over the timescale shown. Stopped-flow analysis is shown for each Mg concentration (indicated from left to right, 0.25, 0.50, 0.75, 1.00, 2.50 and 5.00 mM) and exactly as recorded by the instrument. Data representative of multiple experiments are shown.

although altered fluorescence emission upon mixing was observed over the same timescale seen with the Trp probe, several issues confounded the analysis and quantitative interpretation of Tyr stopped-flow analysis. These issues included: (i) the relatively low signal-to-noise due to the lower quantum yield of Tyr, (ii) an unexplained *increase* (rather than the expected decrease) in fluorescence intensity upon mixing with Mg^{2+}, (iii) the lack of a saturable and concentration-dependent effect of Mg^{2+} on the magnitude of this increase, and (iv) the lack of consistency in initial fluorescence. Overall, my results indicate that it should be possible to monitor ACP conformational changes within this timescale using stopped-flow fluorescence with Trp as a probe, but further work will be required to determine whether similar experiments with Tyr 71 can be optimized to yield useful data.

4.3 Results: Molecular Dynamics

4.3.1 *In Silico* Creation of rACP, SA, SB, and SA/SB ACP PDB files

Although no three-dimensional structure of native *V. harveyi* ACP or rACP has yet been determined, the Vogel laboratory recently reported the NMR structure of the *V. harveyi* A75H mutant (Chan et al., 2010), in which the folded conformation is stabilized due to introduction of a basic residue near the C-terminus (Keating et al., 2002). To complement the fluorescence experiments and explore the potential effects of site A and B mutations, the appropriate mutations were made to *V. harveyi* A75H *in silico* to produce PDB files for rACP, SA, SB, and SA/SB (Figure 50A, B, C, D, respectively). Simulations of these newly created PDBs were performed over a 10.25 ns period (0.25 ns minimization; 10 ns equilibration, as described in Section 2.5.2). Plotting RMSD *versus* time indicates that the simulations proceeded to equilibrium (Figure 51). No substantial

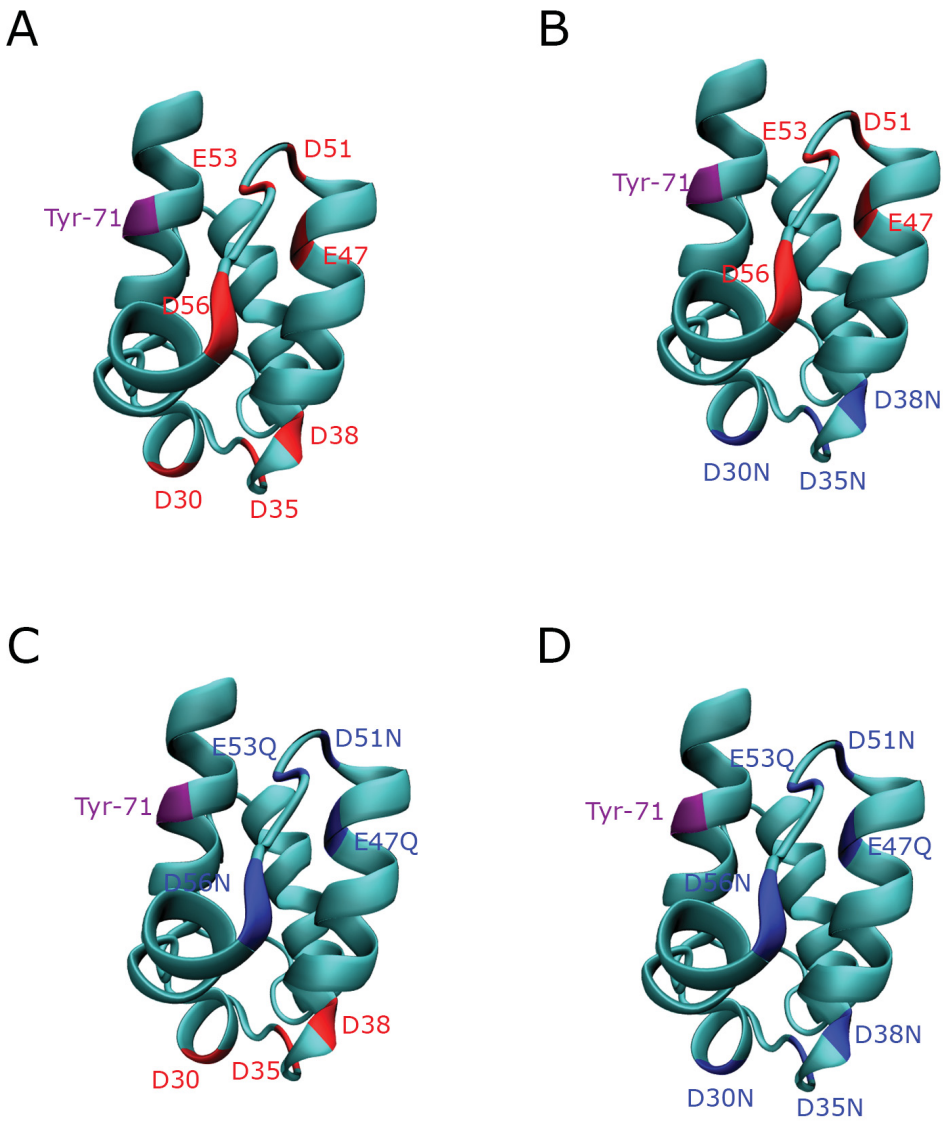

Figure 50 *Vibrio harveyi* ACPs created *in silico* pre-MD simulation. "NewCartoon" representations of rACP (A), SA (B), SB (C), and SA/SB (D). Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Images were created using VMD 1.9.1 and POV-Ray 3.6.

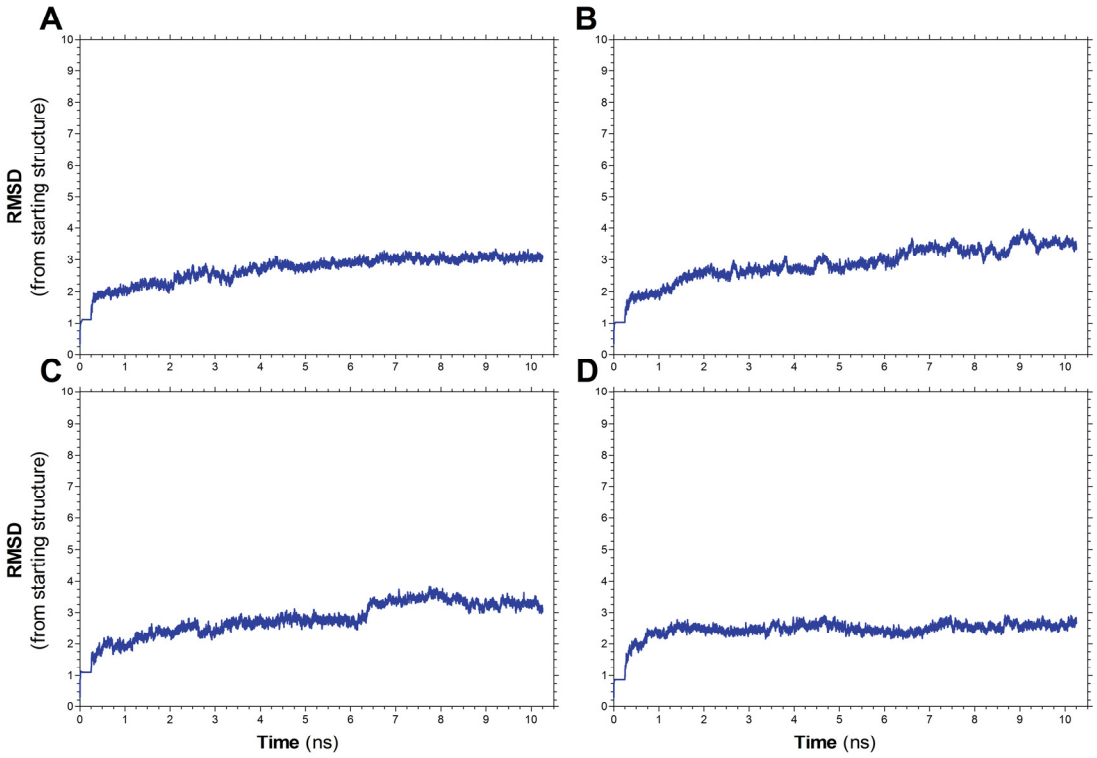

Figure 51 RMSD for backbone atoms of ACP mutant creation simulations (Raw Data). RMSD was calculated from the starting structures for rACP (A), SA (B), SB (C), and SA/SB (D) as shown in Figure 50. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

changes to the overall 4-helix structure of ACP were observed for any of the ACPs, based on the post-simulation structures (Figure 52A, B, C, D).

To examine the possible effect of side chain accessibility on tyrosine fluorescence differences among the ACP constructs (see Section 4.2.1), the solvent accessible surface area (SASA) of all residues was calculated for the equilibrated structures (Figure 53). In most cases, specific residues exhibited accessibility ranging from almost completely buried to fully exposed regardless of construct; for example, Leu 46 (position 50 of rACP due to the N-terminal GIPL tag) is almost completely buried as expected in the folded conformation (Gong et al., 2008). However, the accessibility of some residues did vary among the different simulated ACP structures; in particular the SA mutant has the most solvent exposed Tyr 71 (position 75 of rACP), followed by SA/SB, SB and rACP.

4.3.2 *In Silico* Addition of Mg^{2+}, Ni^{2+} and Zn^{2+} to rACP, SA, SB, and SA/SB ACP PDB Files

Following equilibration of the *in silico* created ACPs, selected divalent cations (Mg^{2+}, Ni^{2+}, Zn^{2+}) were added to the equilibrated, aligned structures created from the previous simulation (Figure 54 – Figure 57); starting control structures are also shown (panel A in each Figure). In each case, two metal ions were placed near sites A and B. All ACP metal combinations, as well as their controls (16 simulations in total), were simulated over a 10.25 ns period (0.25 ns minimization 10 ns equilibration, as described in Section 2.5.2). As previously, RMSD *versus* time plots indicate that all protein structures had reached equilibrium (Figure 58 – Figure 61).

As above, equilibrated, averaged structures saved following each simulation

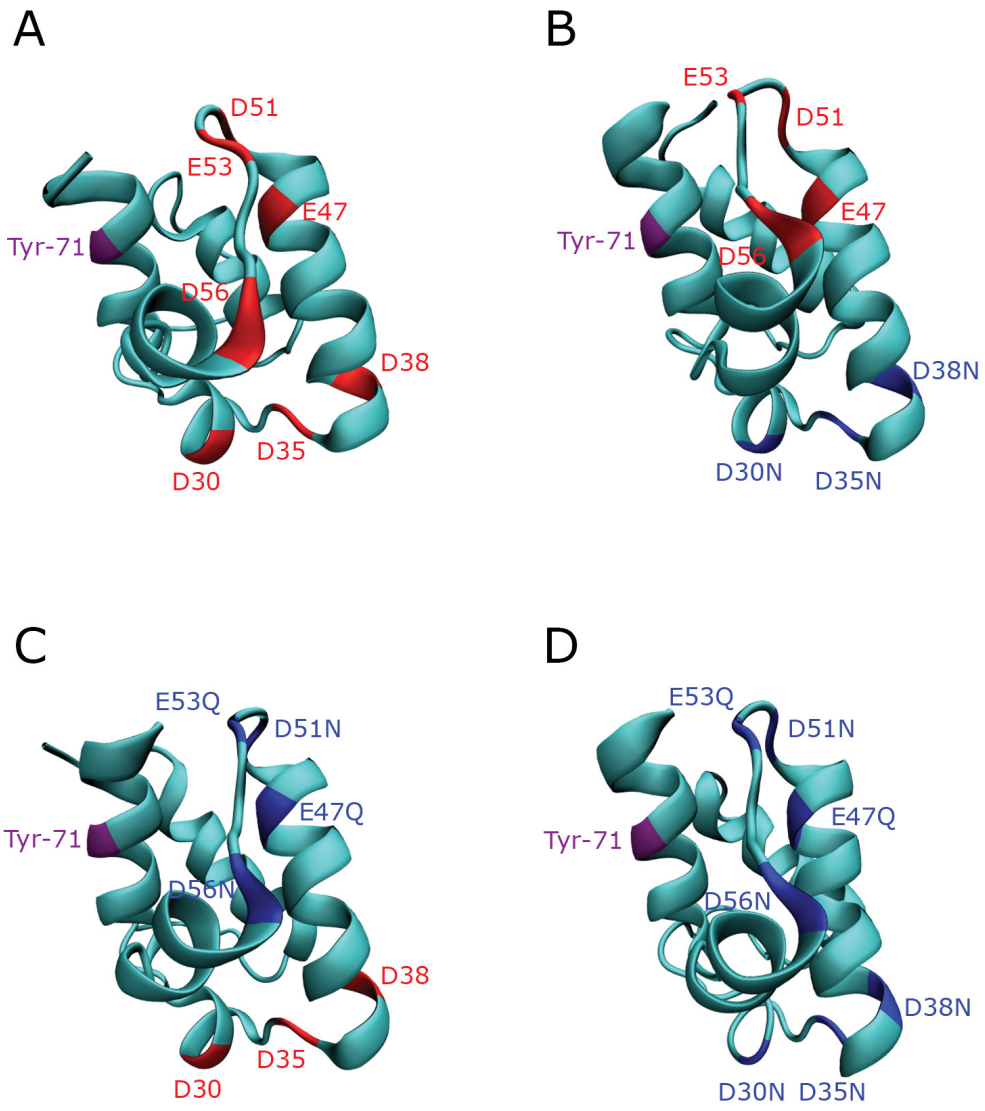

Figure 52 *Vibrio harveyi* ACPs post-MD simulation. "NewCartoon" representations of equilibrated, averaged structure (see Appendix 14 for details) of rACP (A), SA (B), SB (C), and SA/SB (D). Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD and POV-Ray 3.6.

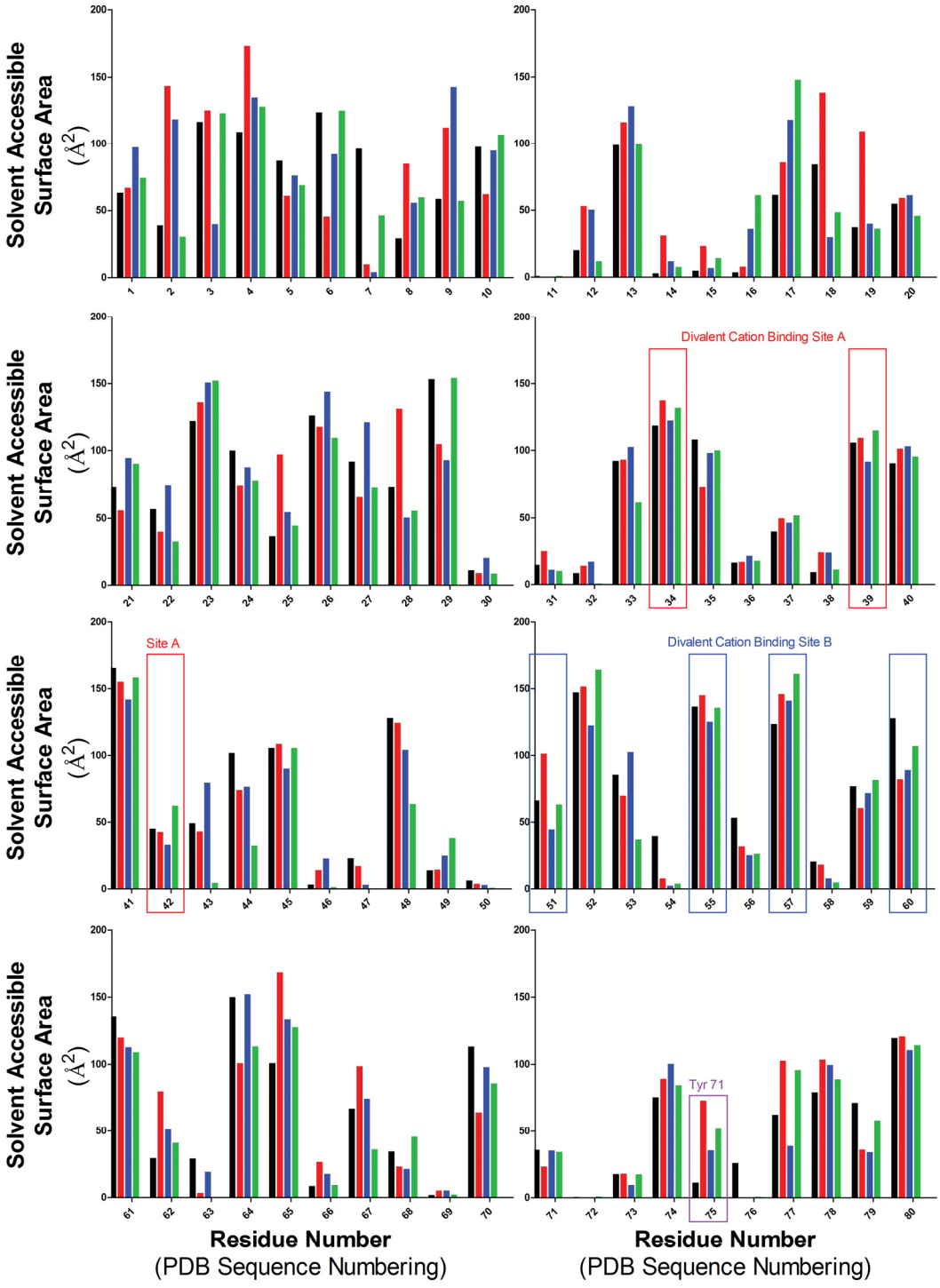

Figure 53 Solvent accessible surface area for each residue following MD simulations of ACPs. SASA was determined for the equilibrated/averaged structure of rACP (black), SA (red), SB (blue) and SA/SB (green). Residue SASA was determined using Chimera as directed in Appendix 23 based on Figure 52. Residue numbering is based on the primary structure of rACP (+4 relative to *V. harveyi* ACP numbering used throughout the thesis). The locations of Tyr 71 and of acidic Site A and Site B residues are indicated by boxes.

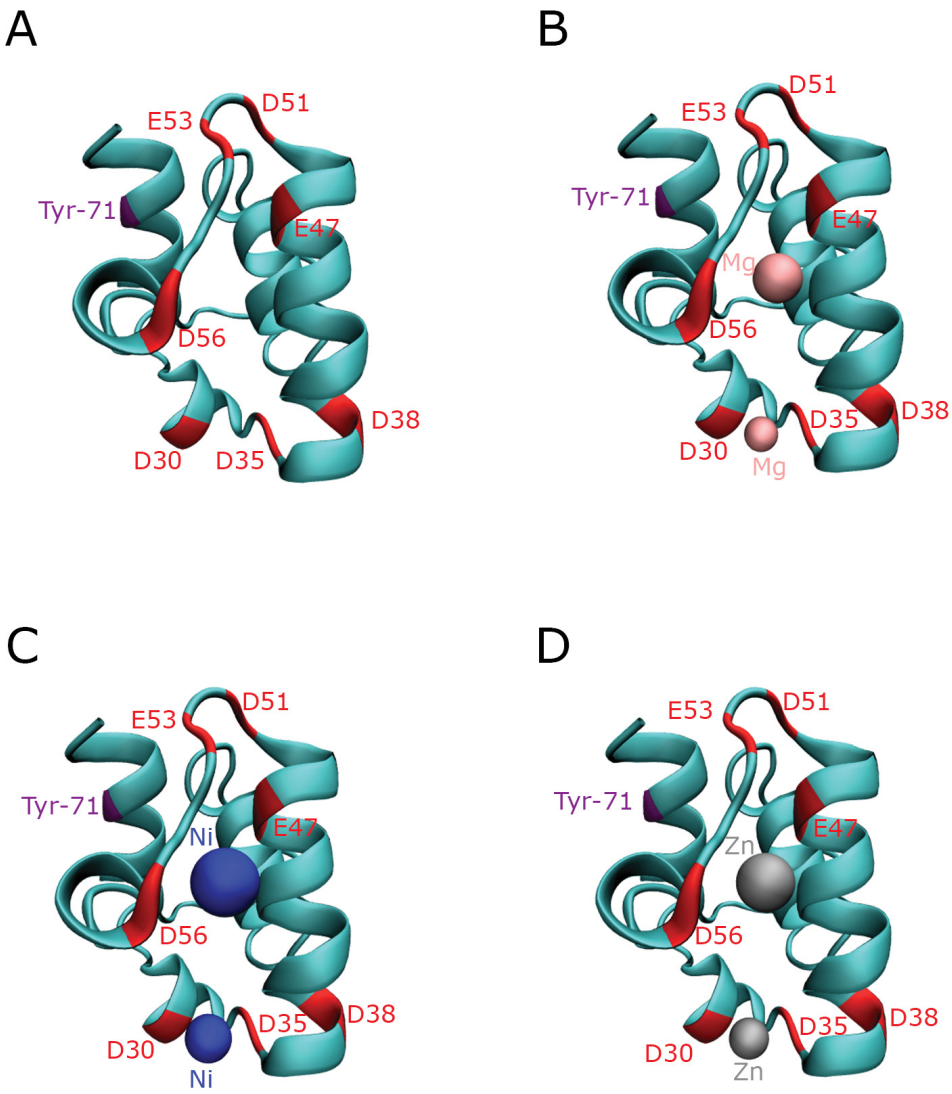

Figure 54 Pre-simulation rACP structures in the presence or absence of Mg, Ni, or Zn created *in silico*. "NewCartoon" representations of rACP (A), rACP + Mg (B), rACP + Ni (C), and rACP + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red). Images were created using VMD and POV-Ray 3.6.

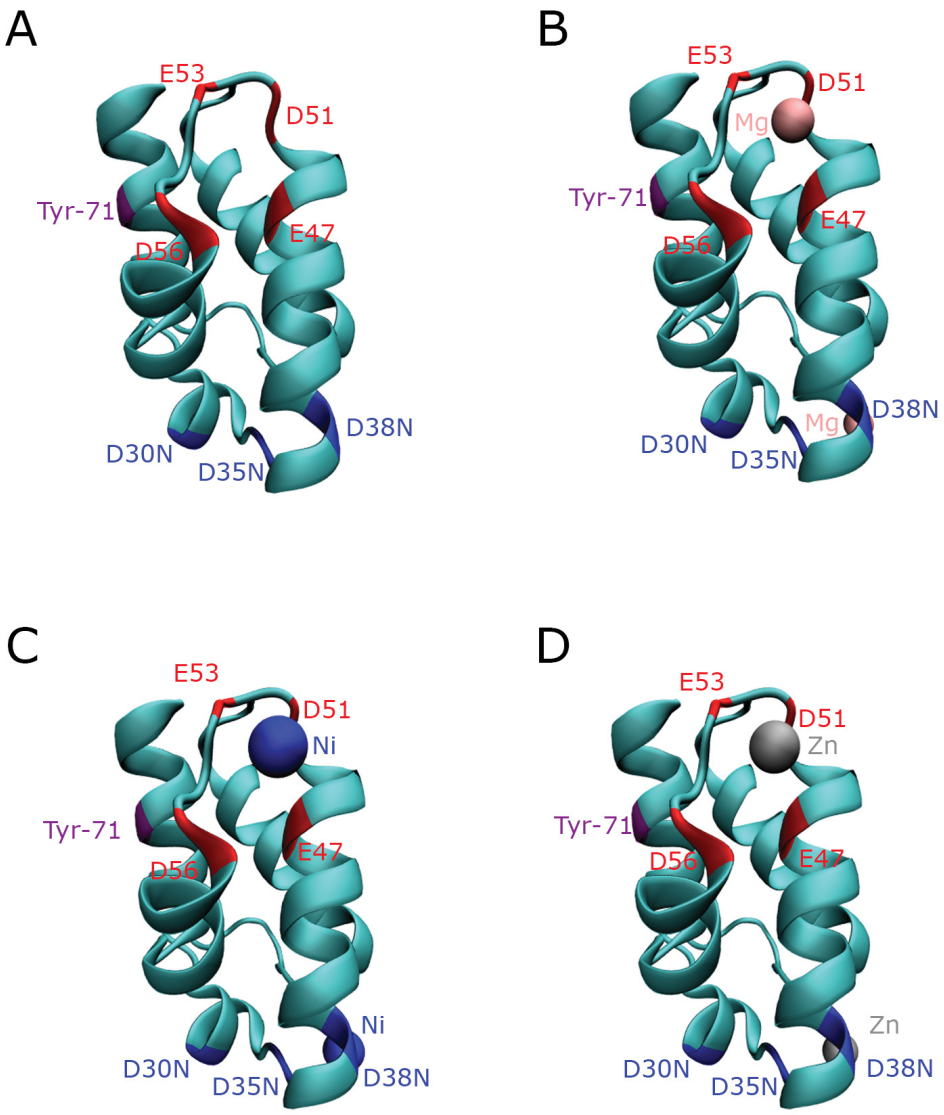

Figure 55 Pre-simulation SA structures in the presence or absence of Mg, Ni, or Zn created *in silico*. "NewCartoon" representations of SA (A), SA + Mg (B), SA + Ni (C), and SA + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Images were created using VMD and POV-Ray 3.6.

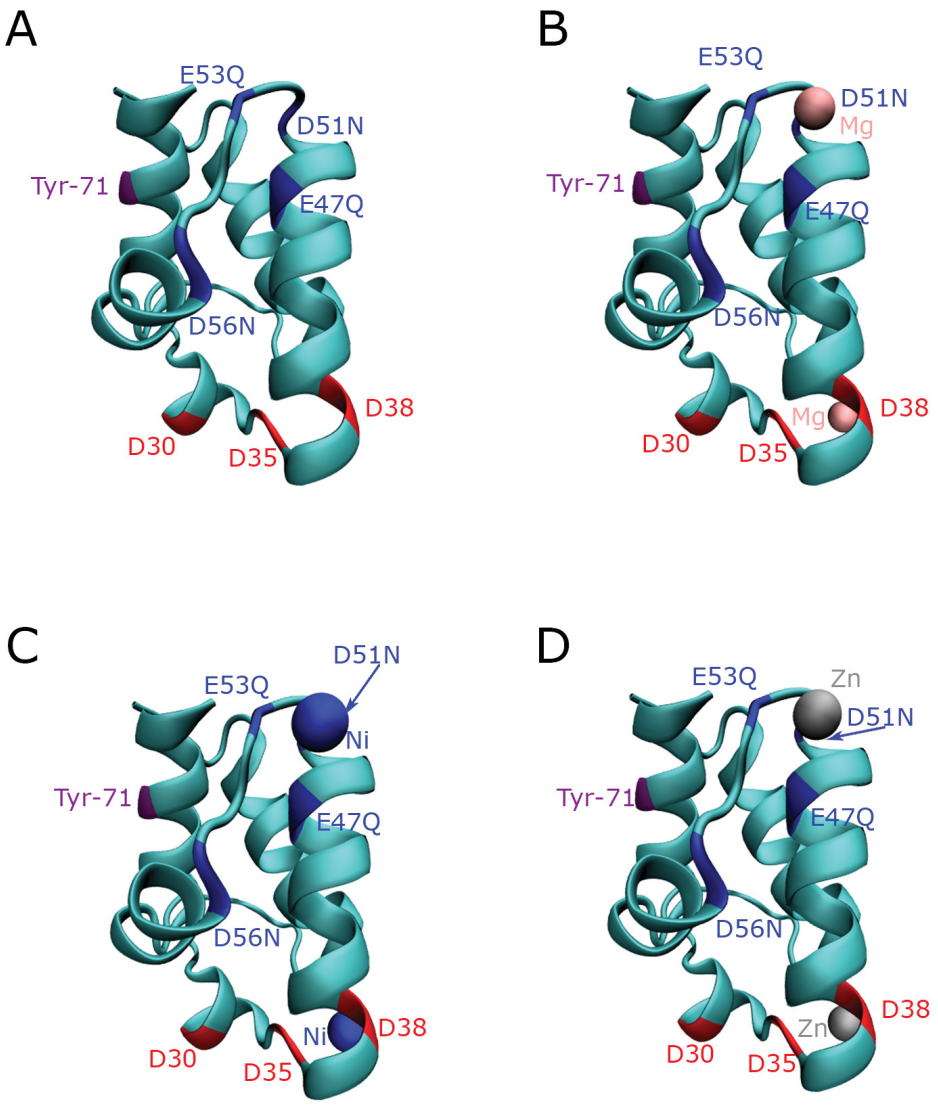

Figure 56 Pre-simulation SB structures in the presence or absence of Mg, Ni, or Zn created *in silico*. "NewCartoon" representations of SB (A), SB + Mg (B), SB + Ni (C), and SB + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Images were created using VMD and POV-Ray 3.6.

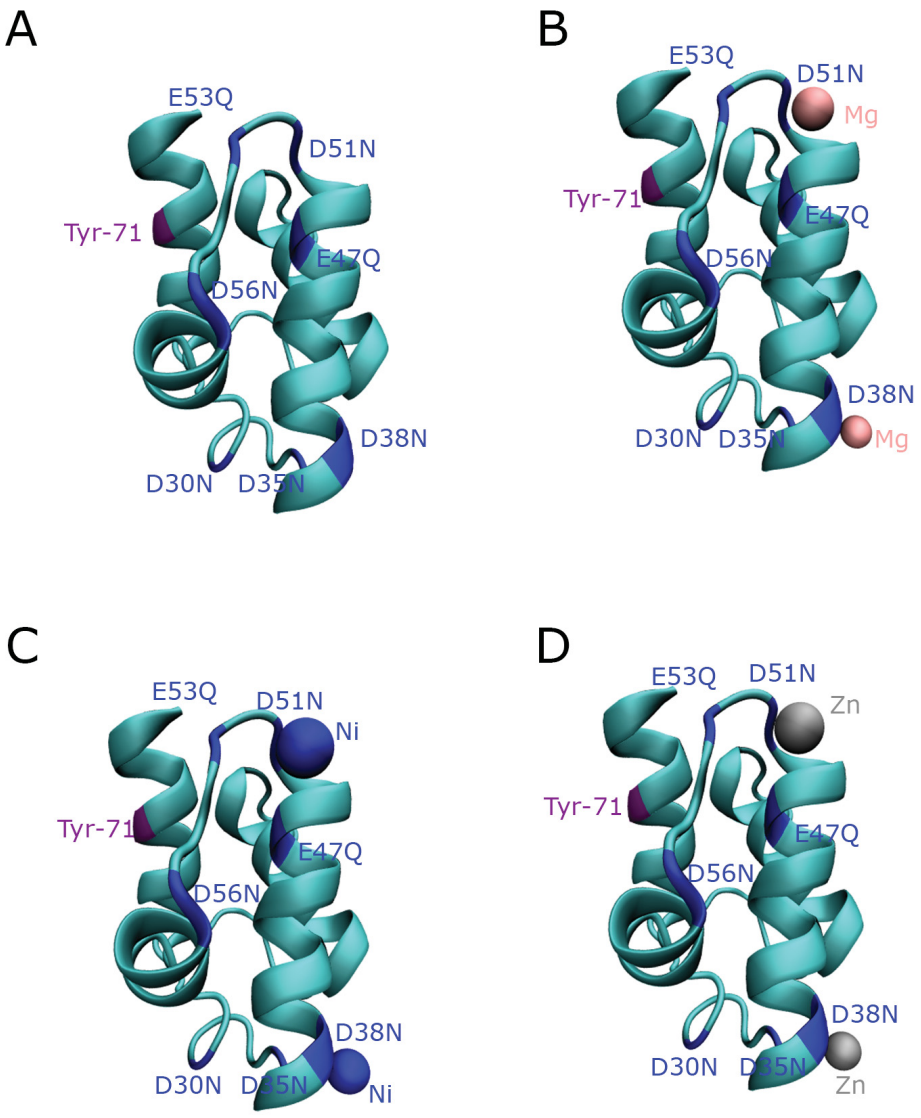

Figure 57 Pre-simulation SA/SB structures in the presence or absence of Mg, Ni, or Zn created *in silico*. "NewCartoon" representations of SA/SB (A), SA/SB + Mg (B), SA/SB + Ni (C), and SA/SB + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: neutral (blue). Images were created using VMD and POV-Ray 3.6.

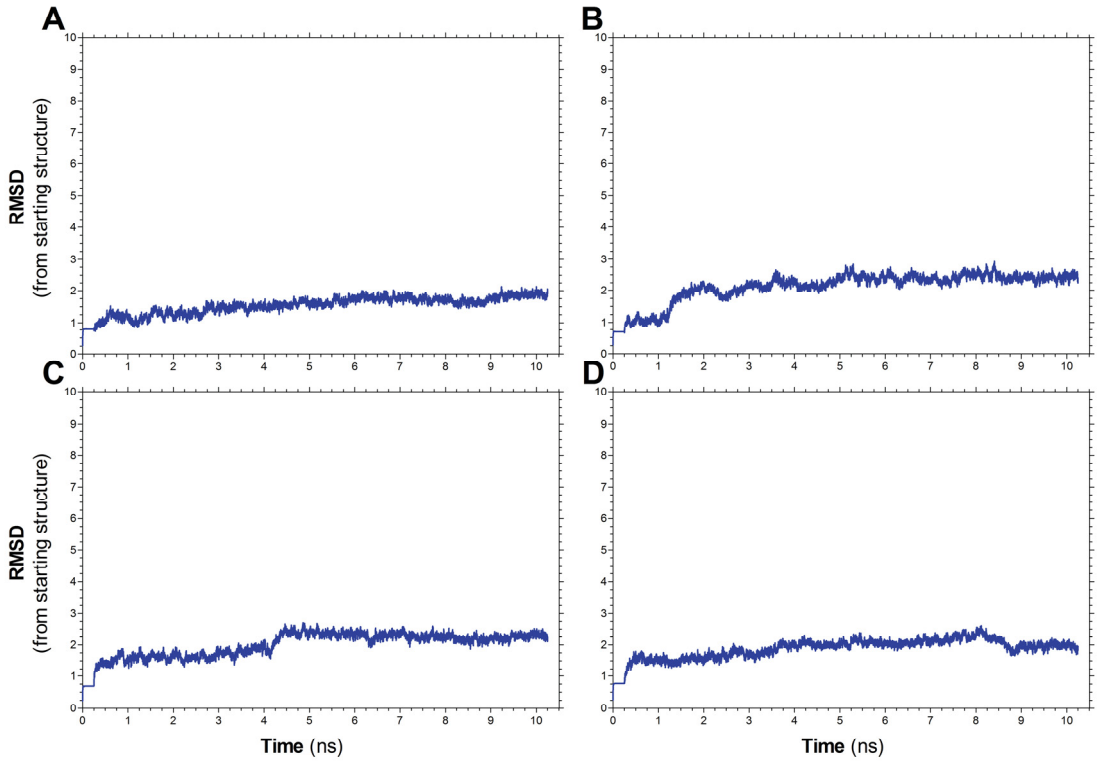

Figure 58 RMSD for backbone atoms of rACP simulations (± metal) (Raw Data). RMSD was calculated from the starting structures for rACP (A), rACP + Mg (B), rACP + Ni (C), and rACP + Zn (D) as shown in Figure 54. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

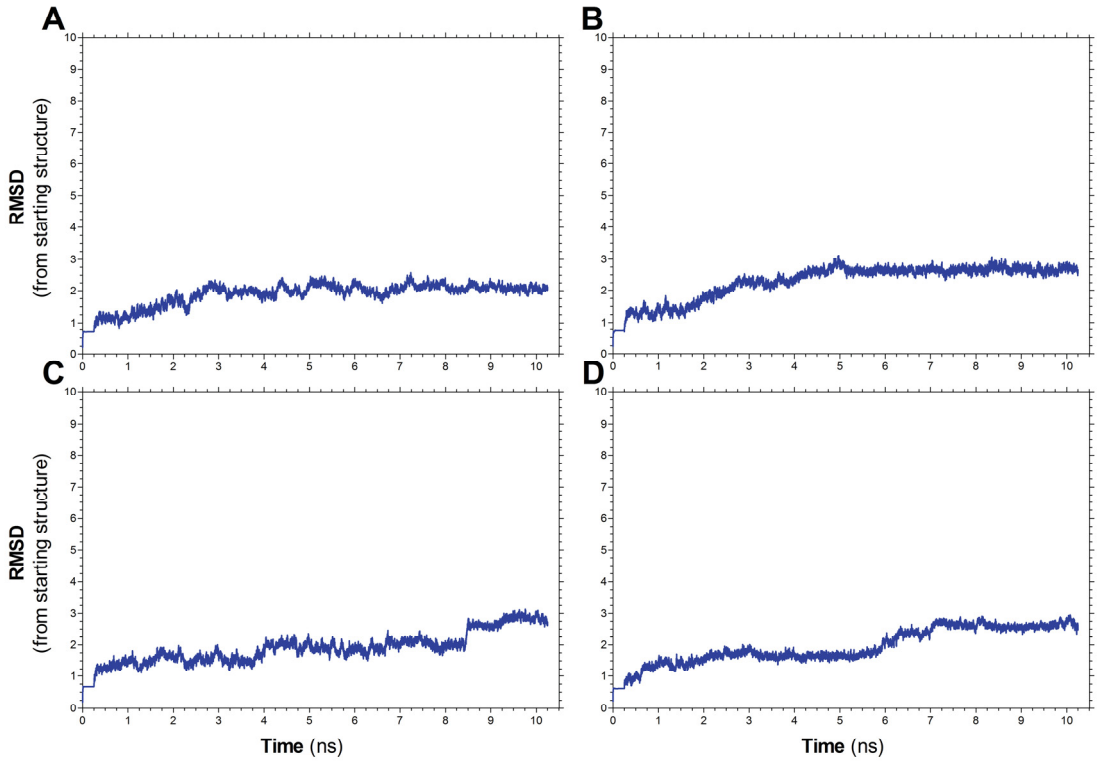

Figure 59 RMSD for backbone atoms of SA simulations (± metal) (Raw Data). RMSD was calculated from the starting structures for SA (A), SA + Mg (B), SA + Ni (C), and SA + Zn (D) as shown in Figure 55. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

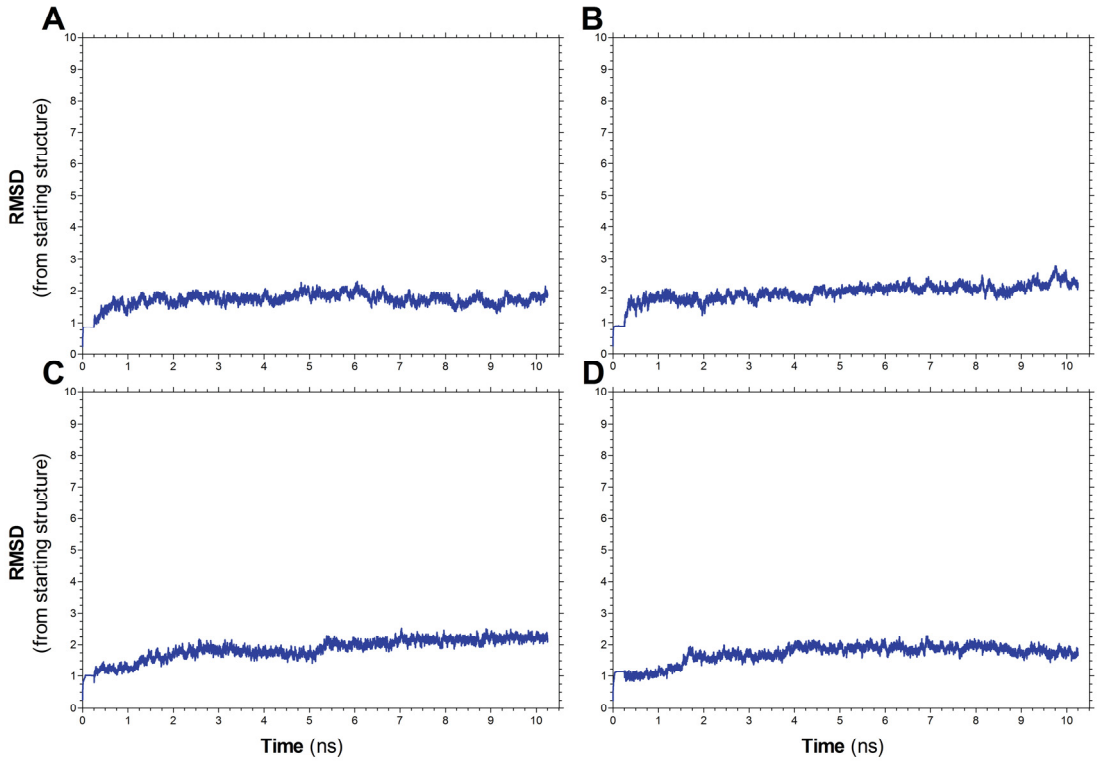

Figure 60 RMSD for backbone atoms of SB simulations (± metal) (Raw Data). RMSD was calculated from starting the structures for SB (A), SB + Mg (B), SB + Ni (C), and SB + Zn (D) as shown in Figure 56. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

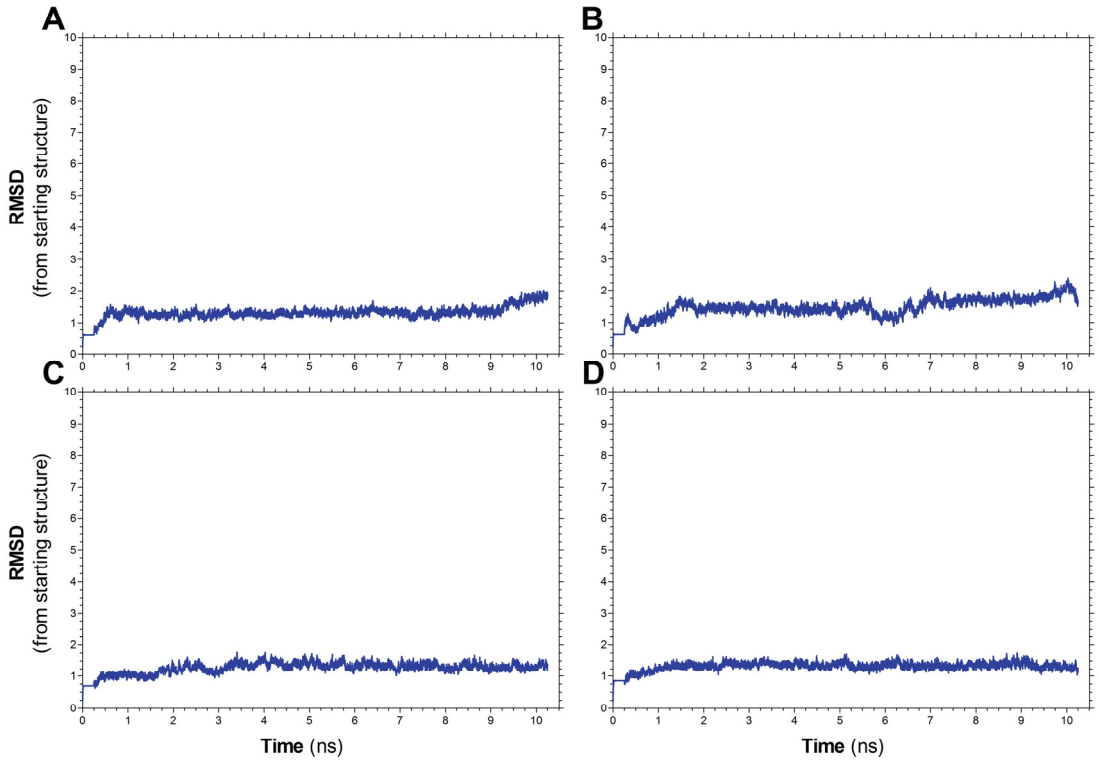

Figure 61 RMSD for backbone atoms of SA/SB simulations (± metal) (Raw Data). RMSD was calculated from the starting structures for SA/SB (A), SA/SB + Mg (B), SA/SB + Ni (C), and SA/SB + Zn (D) as shown in Figure 57. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

indicate the movement (or lack thereof) of metal ions in each case (Figure 62 – Figure 65). To further illustrate the effect of mutation of Site A and Site B on the binding of the various metals *in silico*, distances between the divalent cation placed initially at each site (A or B) and the Cα atom of each acidic residue in that site was measured over the 10 ns simulation (Figure 66 – Figure 69). As expected for the wild-type protein (rACP), all three divalent cations (Mg^{2+}, Ni^{2+} and Zn^{2+}) remained positioned consistently at both sites A and B throughout the course of the simulation (Figure 62 and Figure 66). Mutant SA, in which three acidic residues have been replaced by their neutral amide counterparts in site A, retained each divalent cation at site B, but cations at site A drifted at least twice the distance from their original positions, indicating only transient binding at these sites (Figure 63 and Figure 67). Interestingly, the SB mutant exhibited retention of all divalent cations at both sites despite the considerable charge neutralization at site B (Figure 64 and Figure 68). Lastly, the SA/SB mutant exhibited loose binding of the divalent cations to site A (as with the SA mutant), while only Mg^{2+} significantly drifted from its initial placement at site B (Figure 65 and Figure 69). By comparison, selected water molecules near both site A and B appear to float freely around the surface of the protein (Figure 70).

To determine the effects of metals on the overall size (or degree of compactness) of ACP the protein volume was calculated using a MSMS surface as described in Appendix 35 (Figure 71). Protein volumes prior to metal addition (but following simulation post-mutation to achieve equilibrated structures) were: 9570 $Å^3$, 9428 $Å^3$, 9559 $Å^3$, and 9680 $Å^3$ for rACP, SA, SB and SA/SB, respectively (Figure 71A). Upon addition of Mg^{2+} and further equilibration, volumes of all ACPs except rACP decreased slightly (Figure 71B), while a more pronounced decrease (up to 5% of initial volume)

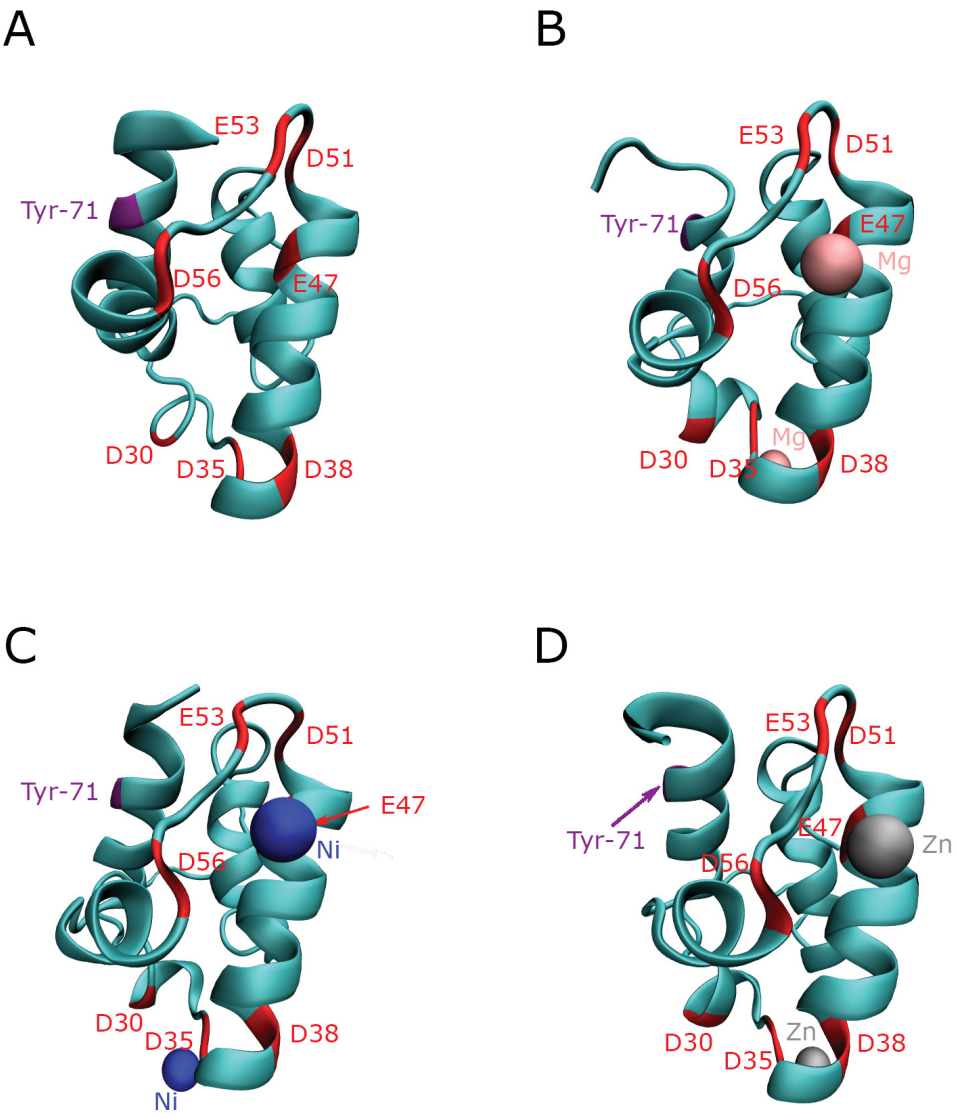

Figure 62 rACP in the presence or absence of divalent cations post-MD simulation. "NewCartoon" representations of equilibrated, averaged structure (see Appendix 14 for details) of rACP (A), rACP + Mg (B), rACP + Ni (C), and rACP + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Equilibrated structures displayed here are the average of the last 1000 frames of the simulation (~500 ps); see Appendix 14. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD and POV-Ray 3.6.

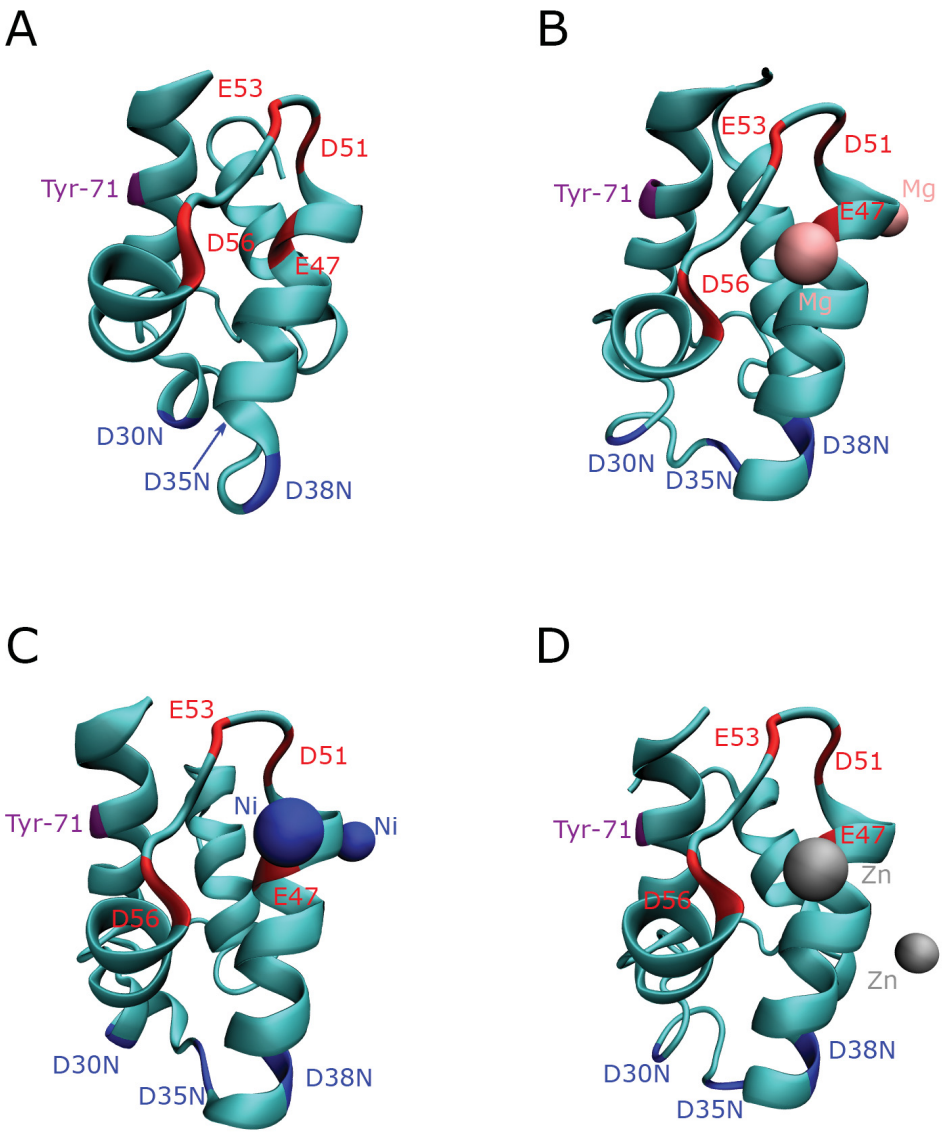

Figure 63 SA in the presence or absence of divalent cations post-MD simulation. "NewCartoon" representations of equilibrated, averaged structure (see Appendix 14 for details) of SA (A), SA + Mg (B), SA + Ni (C), and SA + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Equilibrated structures displayed here are the average of the last 1000 frames of the simulation (~500 ps); see Appendix 14. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD and POV-Ray 3.6.

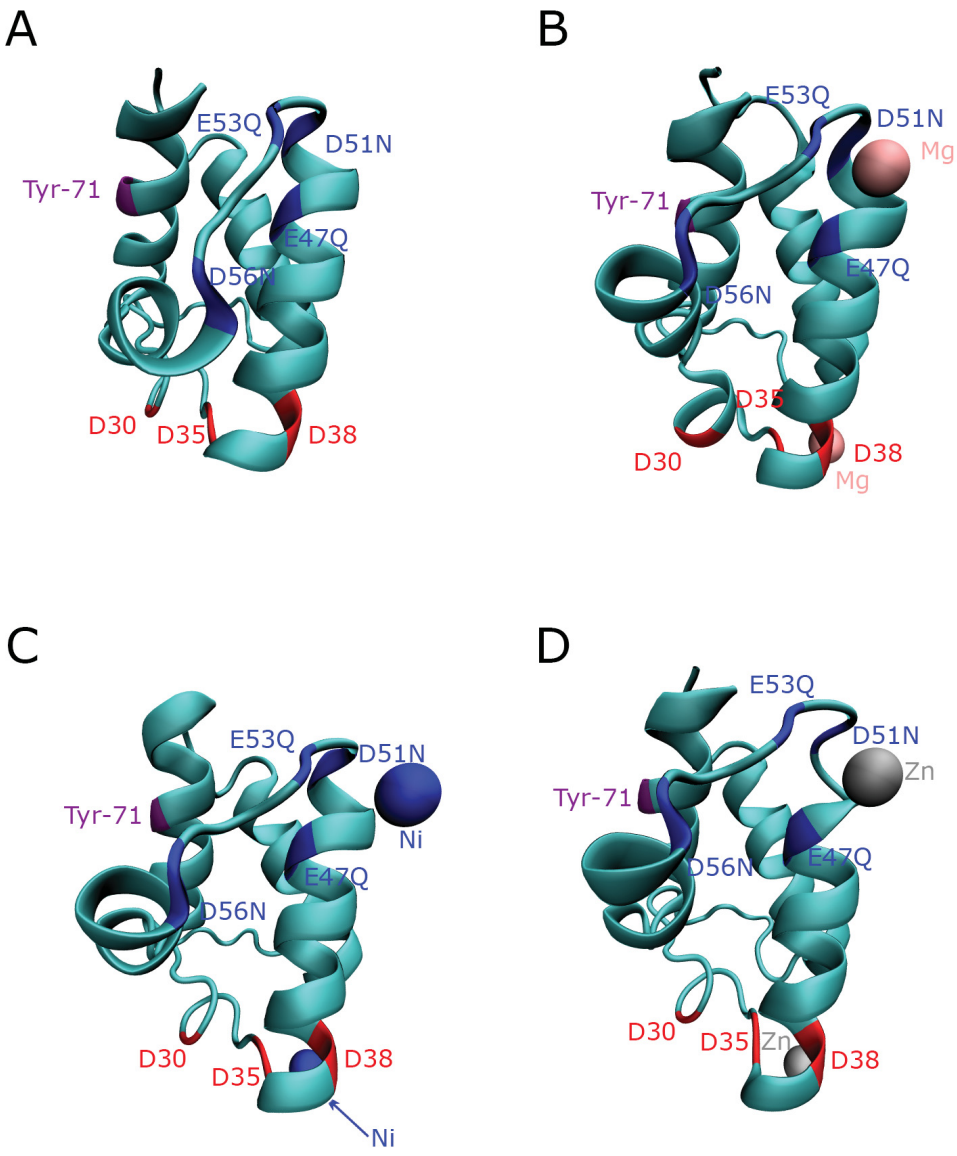

Figure 64 SB in the presence or absence of divalent cations post-MD simulation. "NewCartoon" representations of equilibrated, averaged structure (see Appendix 14 for details) of SB (A), SB + Mg (B), SB + Ni (C), and SB + Zn (D). Metal ions are depicted using "VDW" graphical representation. Positions of interest are noted as well as the charge of selected amino acids: acidic (red) and neutral (blue). Equilibrated structures displayed here are the average of the last 1000 frames of the simulation (~500 ps); see Appendix 14. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD and POV-Ray 3.6.

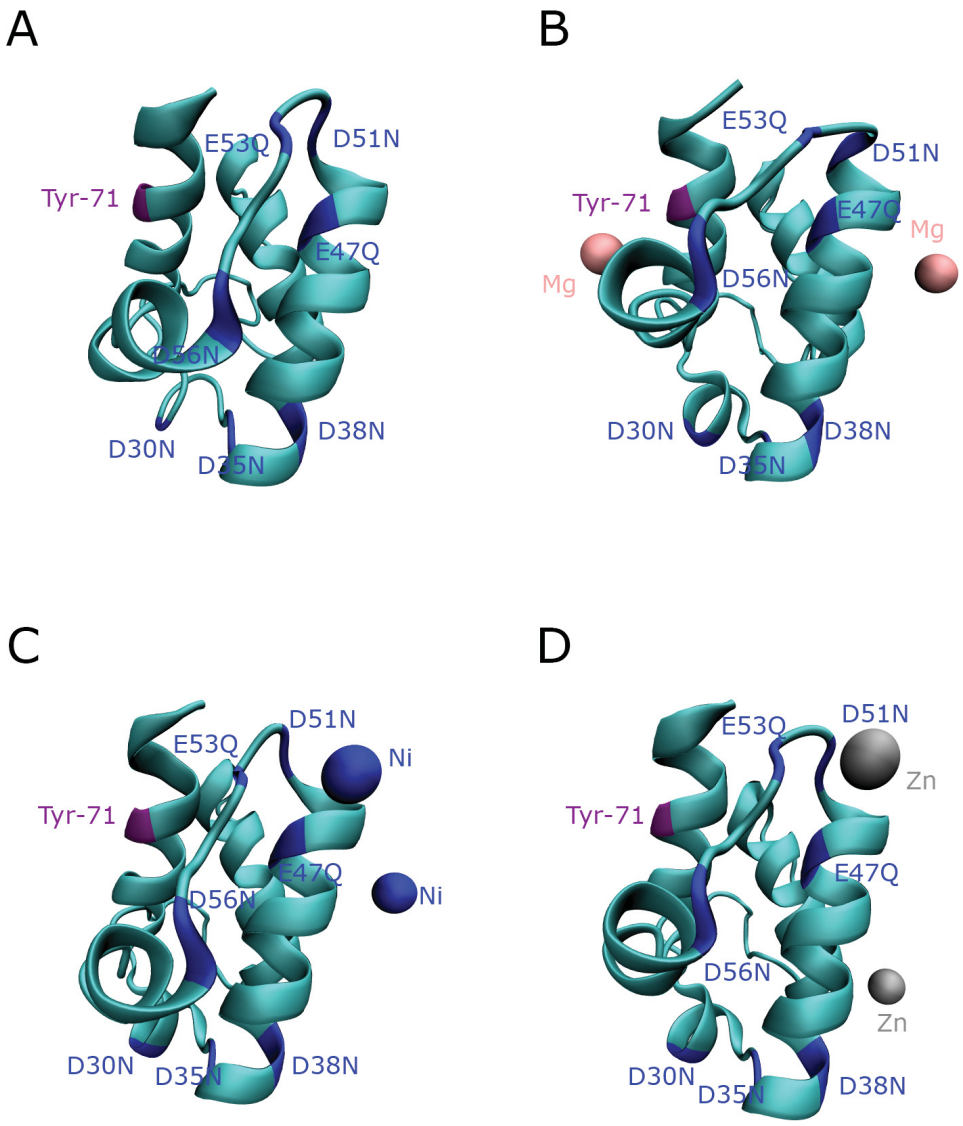

Figure 65 SA/SB in the presence or absence of divalent cations post-MD simulation. "NewCartoon" representations of equilibrated, averaged structure (see Appendix 14 for details) of SA/SB (A), SA/SB + Mg (B), SA/SB + Ni (C), and SA/SB + Zn (D). Metal ions are depicted using "VDW" graphical representation. Equilibrated structures displayed here are the average of the last 1000 frames of the simulation (~500 ps); see Appendix 14. Positions of interest are noted as well as the charge of selected amino acids: neutral (blue). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD and POV-Ray 3.6.

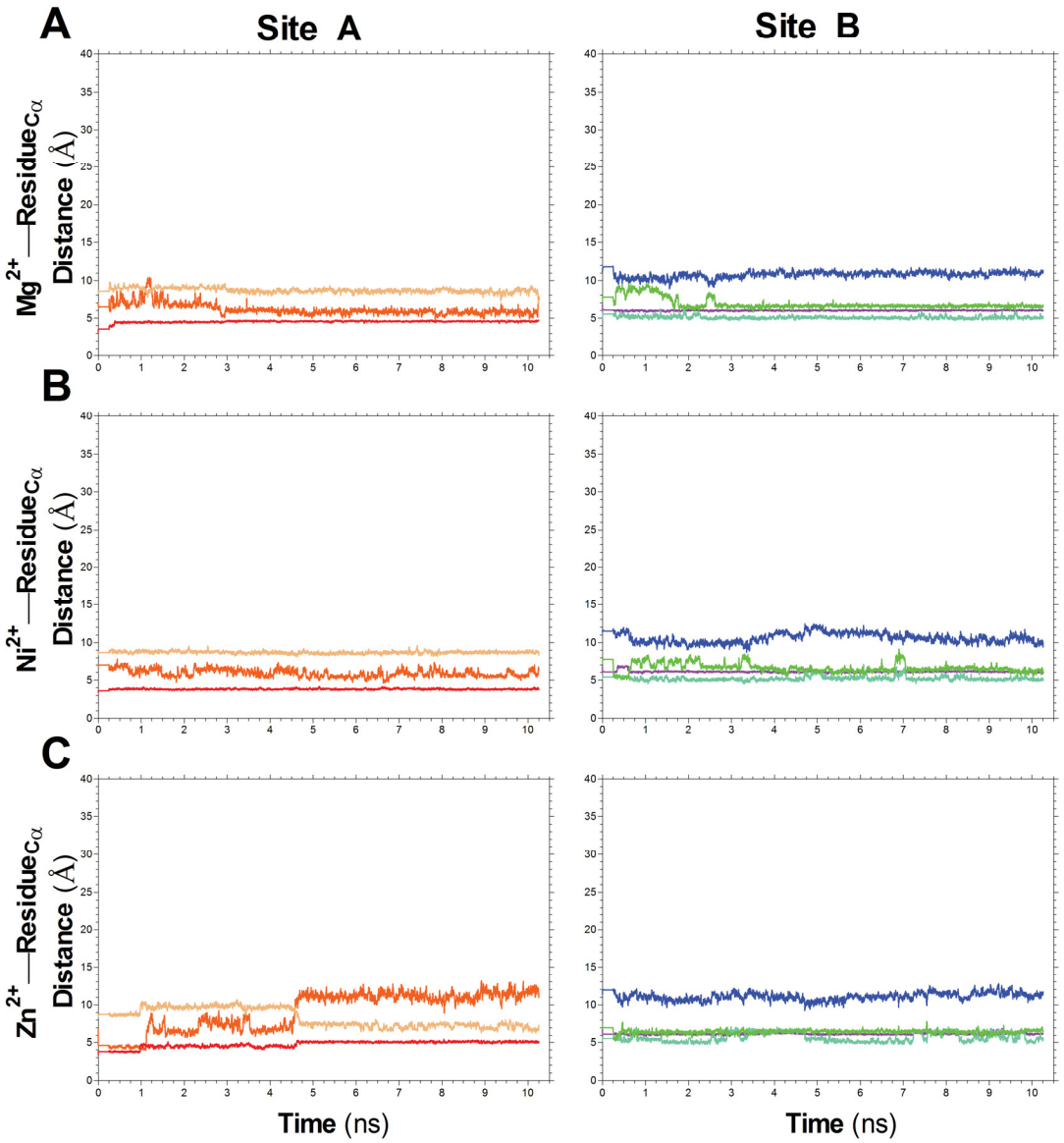

Figure 66 Distance between metal and site A and B Cα residues in rACP metal simulations. Distances were calculated for rACP + Mg (A), rACP + Ni and rACP + Zn (C) simulations for residues: D30 (orange), D35 (red), D38 (tan), E47 (purple), D51 (blue), E53 (cyan) and D56 (green). Values were calculated as described in Appendix 21 using the smoothed trajectory file (sliding window of 9; Appendix 20) to remove movements due to temperature fluctuations.

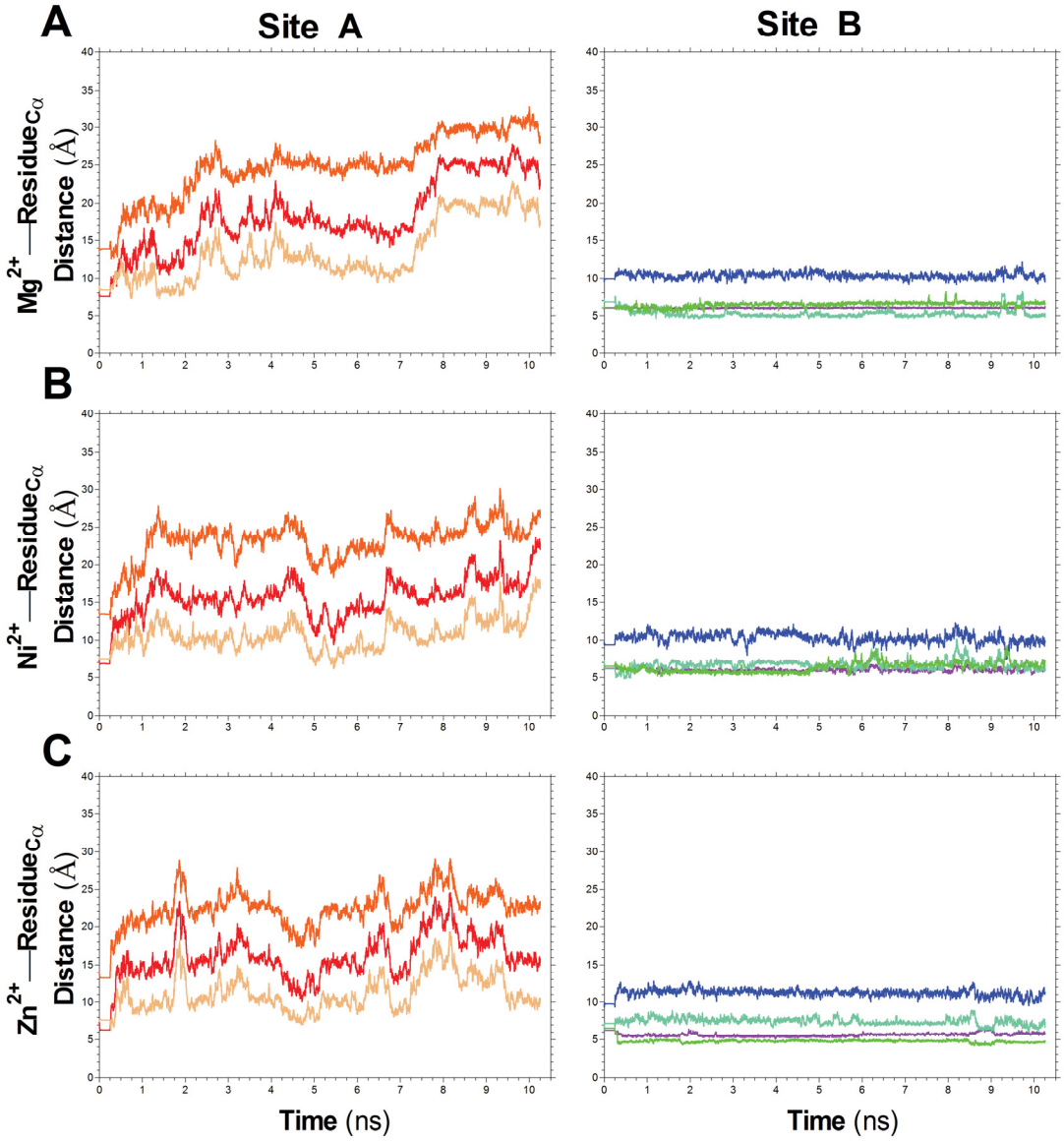

Figure 67 Distance between metal and site A and B Cα residues in SA metal simulations. Distances were calculated for SA + Mg (A), SA + Ni and SA + Zn (C) simulations for residues: D30N (orange), D35N (red), D38N (tan), E47 (purple), D51 (blue), E53 (cyan) and D56 (green). Values were calculated as described in Appendix 21 using the smoothed trajectory file (sliding window of 9; Appendix 20) to remove movements due to temperature fluctuations.

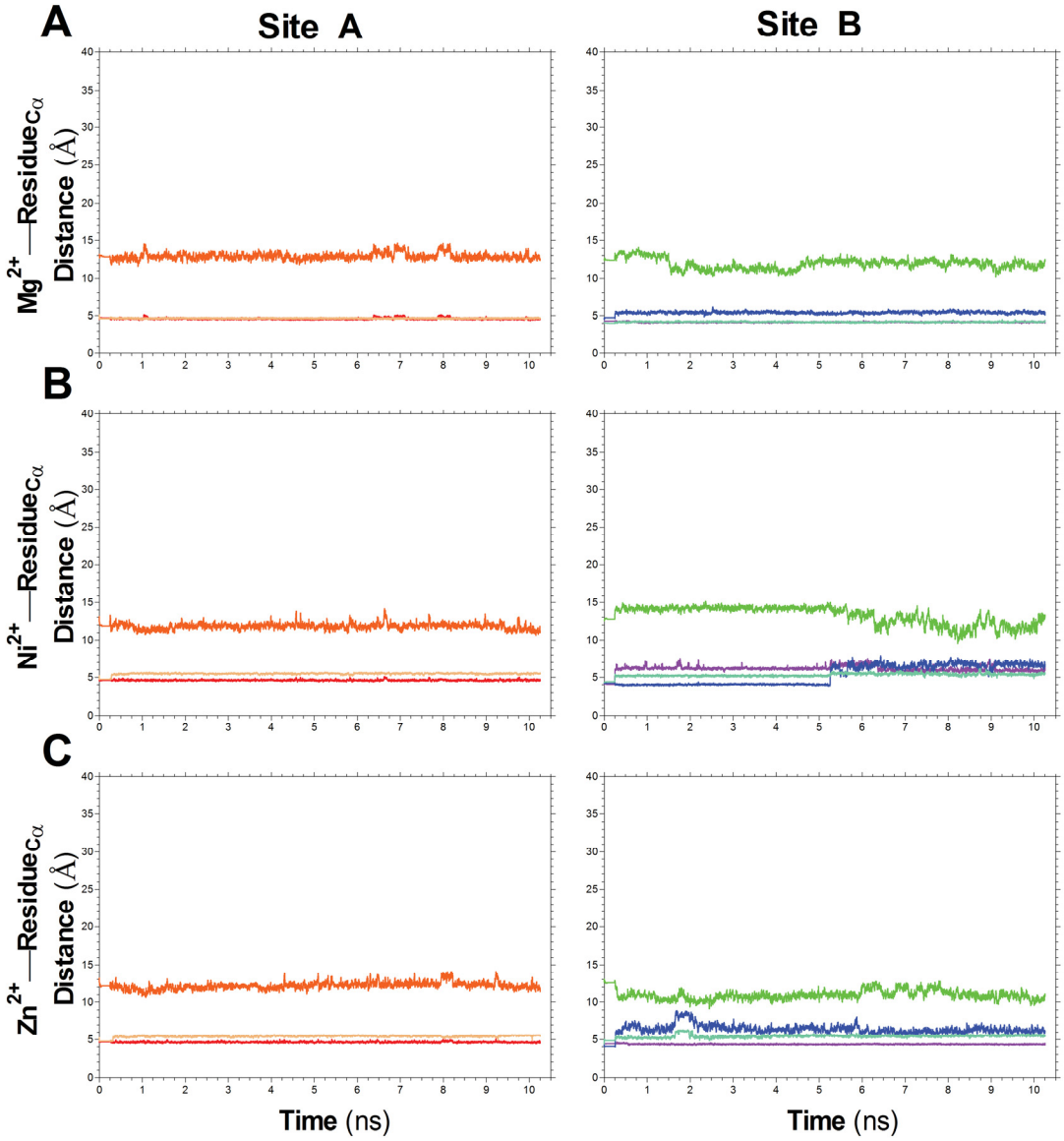

Figure 68 Distance between metal and site A and B Cα residues in SB metal simulations. Distances were calculated for SB + Mg (A), SB + Ni and SB + Zn (C) simulations for residues: D30 (orange), D35 (red), D38 (tan), E47Q (purple), D51N (blue), E53Q (cyan) and D56N (green). Values were calculated as described in Appendix 21 using the smoothed trajectory file (sliding window of 9; Appendix 20) to remove movements due to temperature fluctuations.

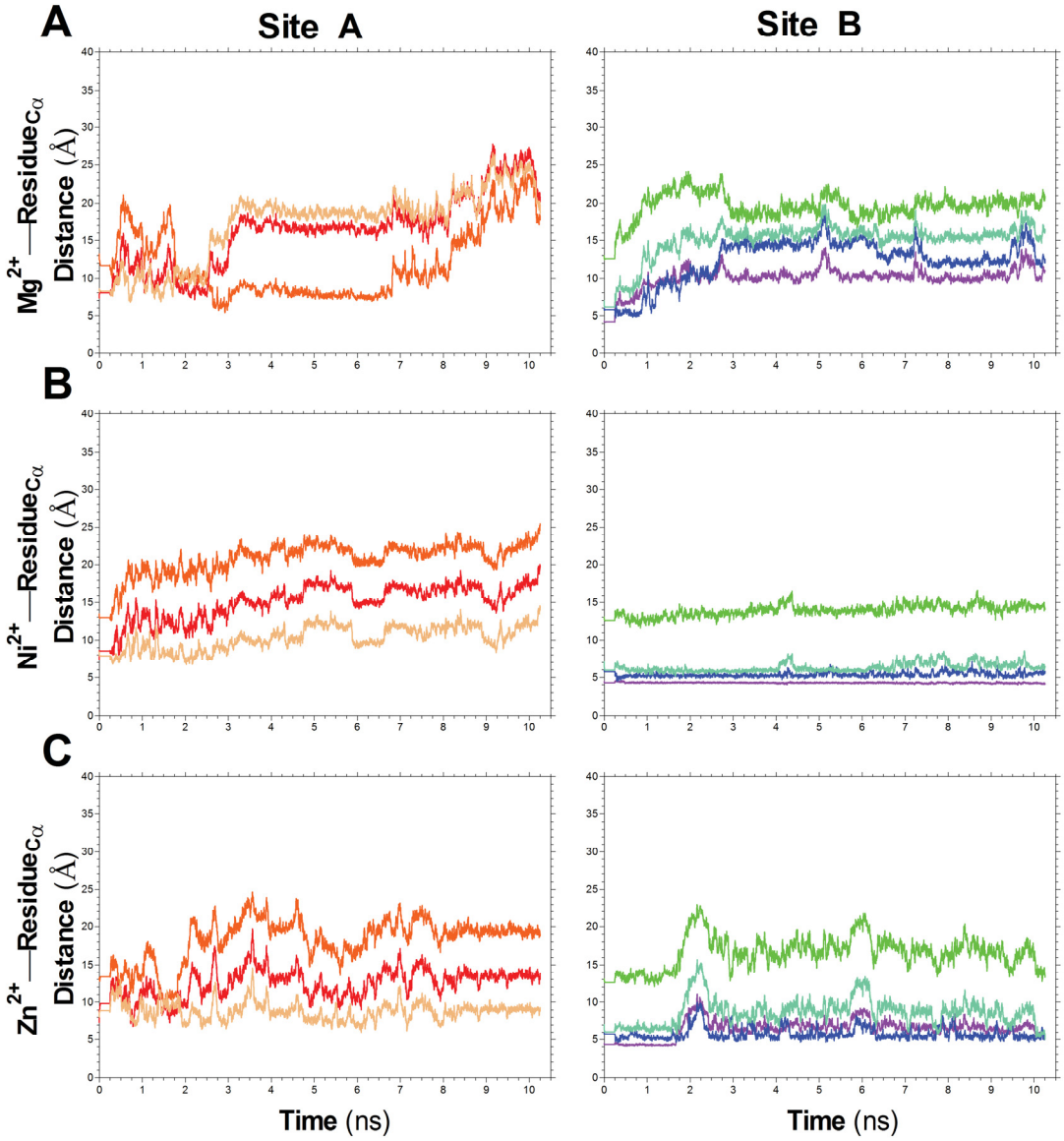

Figure 69 Distance between metal and site A and B Cα residues in SA/SB metal simulations. Distances were calculated for SA/SB + Mg (A), SA/SB + Ni and SA/SB + Zn (C) simulations for residues: D30N (orange), D35N (red), D38N (tan), E47Q (purple), D51N (blue), E53Q (cyan) and D56N (green). Values were calculated as described in Appendix 21 using the smoothed trajectory file (sliding window of 9; Appendix 20) to remove movements due to temperature fluctuations.

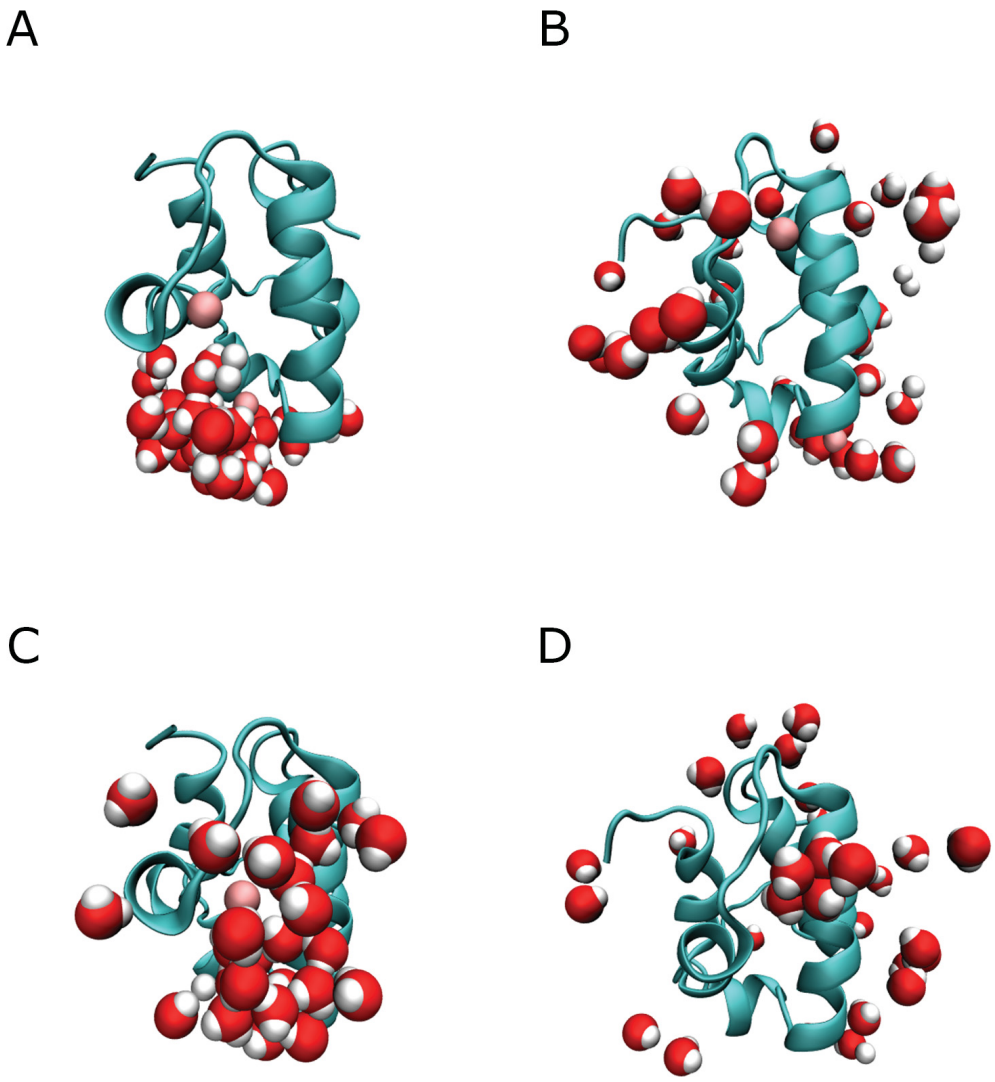

Figure 70 Movement of water molecules near site A or B over the course of a MD simulation. "VDW" representation of selected water molecules at the start (A, C) and end (B, D) of the rACP + Mg simulation. Displayed water molecules are within 10 Å of the Mg^{2+} ion bound to either site A (A, B) or site B (C, D) at the start of the simulation. rACP (cyan) is displayed using "NewCartoon" representation and Mg (pink) displayed using "VDW" representation. One simulation is depicted here, but is representative of all ACP metal simulations. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

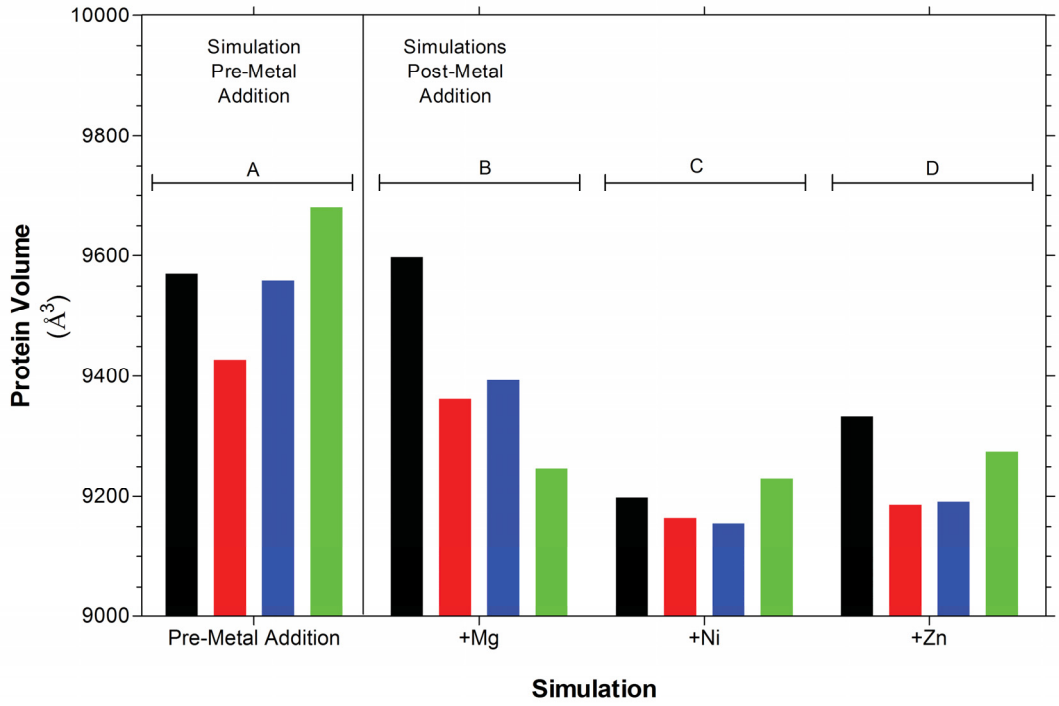

Figure 71 Volume data for ACP MD simulations. Volumes were determined for equilibrated/averaged structures before metal addition (A) and following further 10 ns simulations in the presence (B – D) of metals indicated. Each panel shows rACP (black), SA (red), SB (blue) and SA/SB (green). Volumes were determined using Chimera as directed in Appendix 22.

was observed after addition of Ni^{2+} or Zn^{2+} (Figure 71C, D, respectively). However, a similar control simulation without cations (not shown) also indicated further compaction of the SA, SB and SA/SB mutants, suggesting that more simulation time may be required to determine the contribution of cations to this particular parameter.

4.4 Discussion

Both tyrosine and tryptophan have been employed for decades as intrinsic fluorescent probes of protein conformation and interactions (Lakowicz, 2006). In the context of *V. harveyi* ACP (which lacks endogenous Trp), we have previously placed Trp at an interior position of Helix II (L46W) to demonstrate the stabilization of its folded conformation by divalent cations, acyl group attachment and interaction with partner enzymes LpxA and AcpS (Gong et al., 2008) or by cyclization (Volkmann et al., 2010). Alternatively, Trp placed on the exterior of Helix II (A45W) has been used to reveal the interaction of this helix with LpxA (Gong et al., 2008). Although these replacements do not appear to affect ACP's function *in vitro* (Gong et al., 2008) or *in vivo* (Volkmann et al., 2010), they cannot easily be retrofitted to study the over two dozen ACP mutants made previously in our lab, many of which affect ACP structure and function (Flaman et al., 2001, Keating et al., 2002, Gong & Byers, 2003, Gong et al., 2007).

In this part of my thesis, I have explored the potential of using the single native Tyr 71 found in all of these constructs as a probe for ACP conformation and dynamics. Tyrosine has been used to study dynamic protein interactions in solution (VanScyoc et al., 2002, Lakowicz, 2006, Amaro et al., 2011). In general, Tyr exhibits a lower fluorescent quantum yield than Trp, and correlation of fluorescence with its environment (e.g. variation of peak Trp emission wavelength with environmental polarity) is not as

straightforward to interpret (Lakowicz, 2006). However, with appropriate assumptions, changes in Tyr fluorescence intensity can theoretically be used to provide binding information in simple two-state models of protein interaction or conformational change. Although the results of this work overall suggest that Tyr 71 is not a good probe for ACP for several conceptual and technical reasons, a number of interesting observations have been made.

To examine the feasibility of Tyr 71 as a fluorescent probe of ACP structure, the established Mg^{2+}-induced folding of *V. harveyi* recombinant ACP (rACP) into a helical conformation (Flaman et al, 2001; Gong et al, 2007) was selected as an experimental system. This conformational transition is due to low affinity (mM) binding of Mg^{2+} to acidic residues clustered at two sites (A and B) at either end of the acidic Helix II (Frederick et al., 1988). Neutralization of either site A (D30, D35, D38) or site B (E47, D51, E53, D56) by mutagenic replacement with the corresponding amide residues to form the SA or SB mutant, respectively, has been shown to cause partial folding of rACP, while the combined SA/SB mutant (with seven amide replacements) appears to exhibit a folded conformation that is relatively insensitive to Mg^{2+} (Gong et al., 2007).

More recently, based on the observation that Zn^{2+} inhibits LpxA activity in the 10-100 μM range (D. Byers, personal communication), stabilization of rACP with larger divalent metals such as Zn^{2+} and Ni^{2+} has been investigated (A. Murphy and D. Byers, unpublished results). By monitoring both the blue shift of L46W Trp fluorescence emission (Figure 72) and decrease in CD signal at 220 nM (Figure 73), these experiments reveal that both Zn^{2+} and Ni^{2+} cause conformational transitions similar to that caused by Mg^{2+}, but at concentrations at least 10-fold lower. By curve fitting of these data, I have

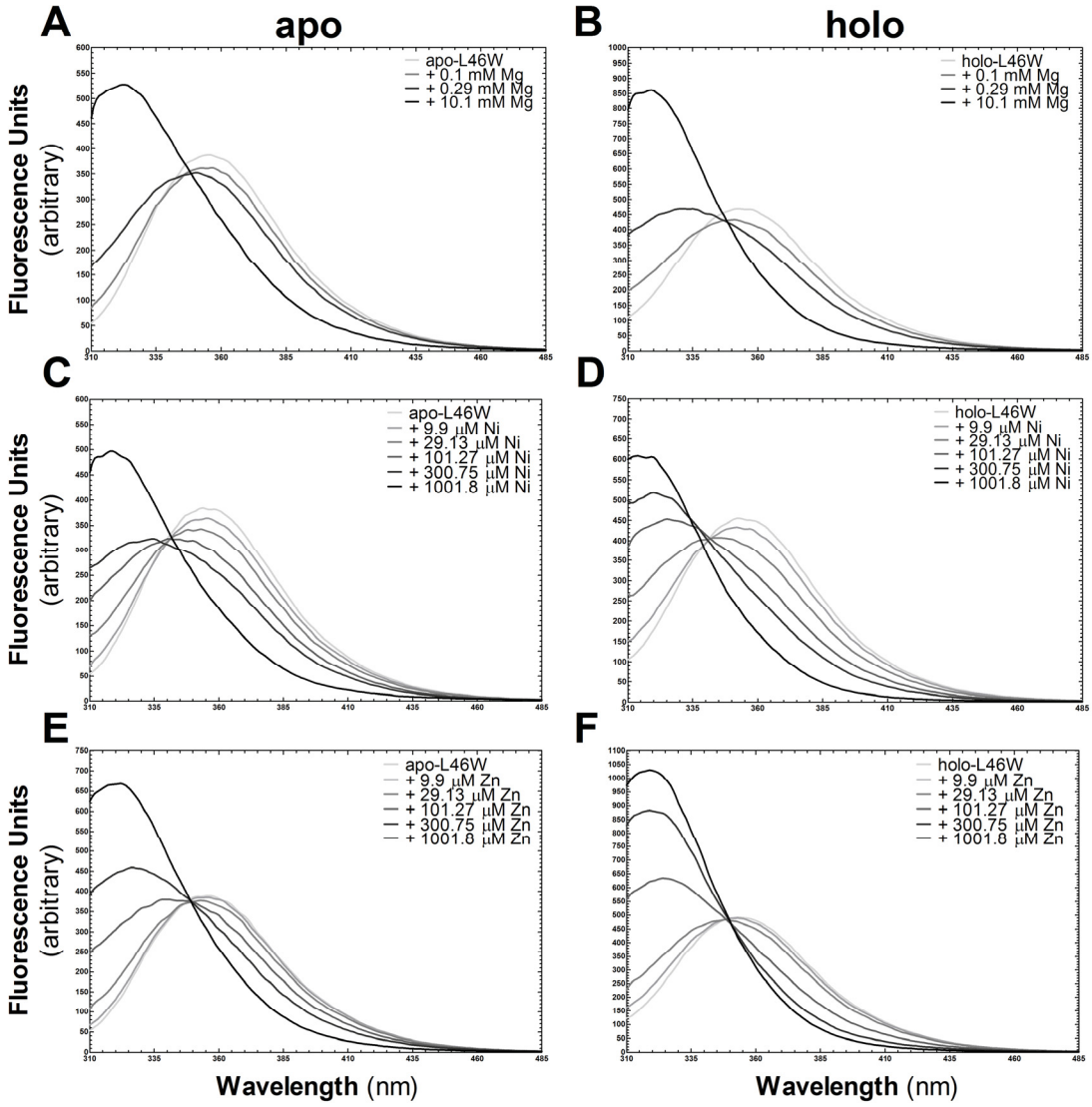

Figure 72 Intrinsic tryptophan fluorescence for L46W titrated with various divalent cations. Trp fluorescence (λ_{ex} = 296 nm) emission spectra for apo-L46W (A, C, E) and holo-L46W (B, D, F) in the presence of increasing concentrations of Mg (A, B), Ni (C, D) and Zn (E, F). L46W (5 μM) was diluted in 10 mM HEPES buffer (pH 7.4). Concentrations of Mg, Ni and Zn are given for each graph. Unpublished data were provided by D. Byers and A. Murphy. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

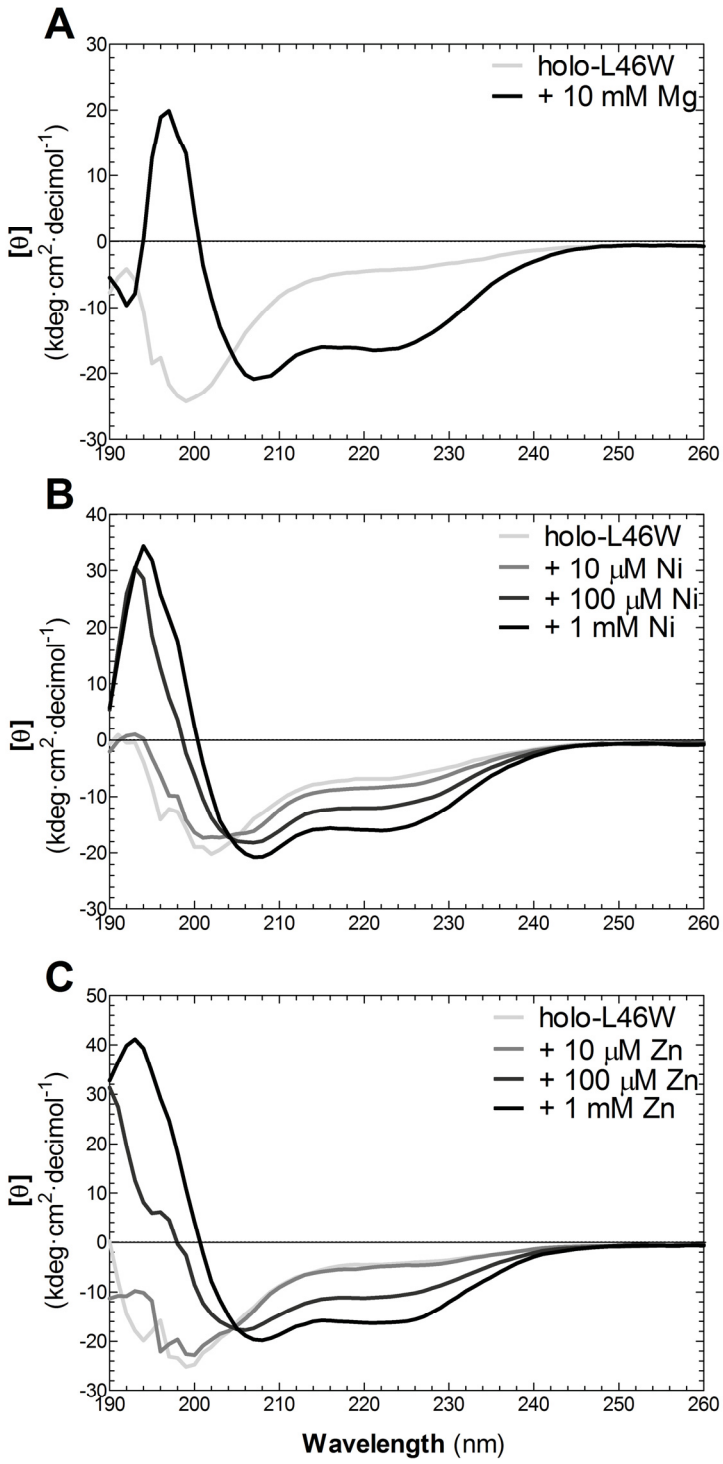

Figure 73 CD analysis of L46W titrated with various divalent cations. CD spectra of holo-L46W in the presence of increasing concentrations of Mg (A), Ni (B) and Zn (C). L46W (5 μM) was diluted in 10 mM HEPES buffer (pH 7.4) containing 1 μM EDTA. Concentrations of Mg, Ni and Zn are given for each graph. Unpublished data were provided by D. Byers and A. Murphy.

estimated the conformational transition midpoint for L46W to be 50 – 100 μM for both Zn^{2+} and Ni^{2+}, compared to the published value for Mg^{2+} of ~ 1 mM (Gong et al, 2008).

After correction for a number of experimental factors, including photobleaching of Tyr 71 and the dilution upon successive addition of metal, only a modest decrease in the fluorescence emission of Tyr 71 was noted (Figure 47). Nevertheless, these changes were observed over concentration ranges similar to those previously observed with Mg^{2+} (Gong et al., 2007, Gong et al., 2008) and other metals (Figure 72 and Figure 73), indicating that changes in Tyr 71 fluorescence do in fact reflect alterations in the environment of this residue due to metal binding and/or conformational transition. However, because of the cumulative error associated with the multiple corrections involved in these titrations, any K_D values determined from Tyr fluorescence data (Table 14) should be regarded with suspicion.

One of the more interesting observations from the fluorescence titration experiments is that the effects of metal addition on site A and/or B mutants appear to be similar to those with the rACP control. Based on previous circular dichroism data showing that SA and SB mutants are both partially folded but can be further folded in the presence of Mg^{2+}, we assumed that amide replacement at each site destroyed its metal binding affinity yet partially stabilized folded ACP through neutralization of the acidic Helix II (Gong et al, 2007). Furthermore, it was concluded that the remaining (unmutated) site was still independently capable of cation binding leading to further conformational stabilization. However, Mg^{2+} binding to each site was not directly measured in that study, and the assumption that amide replacement completely ablates association with divalent cations should perhaps be re-evaluated.

The above conclusions are further supported by molecular dynamics simulations. *In silico* mutation of the *V. harveyi* A75H template structure to introduce amides into sites A and/or B did not have any substantial impact on overall secondary or tertiary structure. However, these experiments do indicate that sites A and B, after mutation, may differ in their ability to retain associated divalent cations, and they also provide a qualitative estimate of the affinities of different metals. As expected, metals added to wild-type rACP appeared to remain bound to both sites A and B during the simulation, while all three cations eventually drifted away from site A in both the SA and SA/SB mutants. In the aforementioned simulations, Mg^{2+} was bound to site A for the shortest time (~2 ns), while Ni^{2+} and Zn^{2+} remained bound for more of the simulation (~5 and ~7 ns, respectively). In contrast, however, neutralization of site B in the SB mutant had little negative effect on cation binding at either site, although some movement of Mg^{2+} away from that site was noted in the double SA/SB mutant. This suggests that the sole remaining acidic site A residue (D31) is incapable of keeping divalent cations at that site, whereas the two remaining acidic site B residues (E48, E49) may be sufficient to retain some binding. Some differences in relative cation affinity ($Zn^{2+} > Ni^{2+} > Mg^{2+}$) and indications of cooperativity between sites were suggested by the time-dependent motions in these simulations, but it may be premature to draw more detailed conclusions on the basis of only a single simulation for each condition. Furthermore, it has been shown very recently that standard CHARMM27 forcefields encompass primarily electrostatic interactions and do not accurately predict solvation energies for ions; especially Zn (~60 kcal/mol difference in CHARMM27 forcefield calculated energy *versus* the values estimated based on experimental data) (Riahi et al., 2013). Therefore, any further MD

study involving divalent cations would benefit from the additional use of either the molecular mechanical Drude polarizable forcefield or the combined quantum mechanical/molecular mechanical method involving both the Drude polarizable forcefields and quantum mechanical regions encompasing the ion and the 6 nearest coordinated oxygens. It would be interesting to see if these trends are supported by multiple replicate simulations of increased duration, but those computational resources were not available during this study.

Replicate simulations (using the above mentioned changes to the forcefields) of increased duration as well as increasing the water sphere/spherical boundary would also strengthen interpretation of protein volume changes as a result of site A/B modification and/or divalent cation binding. Figure 71 suggests that mutagenic replacement of acidic residues does decrease the protein volume relative to rACP *in silico*, as might be expected on the basis of charge neutralization of electrostatic repulsion. However, there is some indication that volume changes (especially for SA/SB) are not complete within the simulation period. Moreover, the lack of effect of Mg^{2+} on rACP volume is at odds with the known effect of this cation on ACP conformation. MD simulations here were constrained by the spherical boundary conditions imposed, which restrict atom movements to a sphere with a set radius and center. This would restrict any hydrodynamic expansion of the very acidic rACP that might be expected to occur in the absence of metal ions.

Another interesting observation in this study was the apparent differences in Tyr 71 fluorescence intensity of rACP and the SA, SB and SA/SB mutants. The increased fluorescence intensity of Tyr 71 in these mutants appears to roughly correlate with loss of

acidic nature: the number of acidic residues in rACP (22) decreases to 19 (SA), 18 (SB), and 15 (SA/SB), while the predicted isolelectric points of these proteins are 3.79, 3.92, 3.92 and 4.09, respectively (ExPASy. *Compute pI/MW Tool*.). Fluorescence intensity would thus also correlate with the extent of helix formation (or fraction of molecules in a helical conformation) in the various ACP constructs (Gong et al , 2007). However, the differences in fluorescence intensity persist in the presence of divalent cations (Figure 47), where all ACPs should be in a more compact folded conformation. One explanation might be differential quenching by amide *versus* carboxyl groups in the vicinity of Tyr 71, as amides are known quenchers of tyrosine fluorescence (Wiczk et al, 2001), but this would only make sense if carboxyl groups caused even greater quenching because rACP exhibited the lowest fluorescence intensity. Carboxyl groups quench Trp fluorescnece (Lakowicz, 2006); however, the quenching ability of an amide *versus* a carboxyl group is not known. Moreover, Tyr 71 is quite distant from many of the acidic residues in sites A and B, even in the folded conformation. Molecular dynamics provides limited insight into the potential environment of Tyr 71 in the folded state: although the SASA of this residue varied significantly among the constructs (Figure 53), no simple correlation with fluorescence intensity was apparent. If solvent accessibility was the cause of differing initial fluorescence it would be expected that SA/SB would have the most accessible tyrosine, followed by SA, SB and rACP (Figure 38). Thus it seems likely that more factors are at play than simply tyrosine accessibility. The origin of differing fluorescence intensities could be further explored using other conformational triggers of ACP folding (e.g. fatty acylation, decreased pH) or by examining the effects of different quenchers (e.g. acrylamide, iodide) on Tyr 71 fluorescence of these constructs.

Finally, my results indicate that stopped-flow analysis of Trp fluorescence may be a promising tool to investigate the importance of specific residues in the conformational transitions involved in the interaction of ACP with its partner enzymes. The blue shift accompanying Mg^{2+}-induced folding of L46W was complete within 2-3 s, much greater than the mixing dead time of the instrument (8 ms). Very recent NMR and X-ray analysis has confirmed that ACP undergoes a significant conformational transition during the transfer of bound acyl groups to the active site of FabA (Nguyen et al., 2014), and fluorescent probes such as L46W and A45W should be well positioned to monitor these motions for other enzymes (such as *E.c.*LpxA) that lack endogenous Trp (see Chapter 5). ACP mutations that inhibit or slow down these transitions can potentially be identified using this approach, although these might have to be retrofitted to Trp probe constructs unless issues preventing the use of Tyr 71 as a stopped-flow probe can be overcome.

CHAPTER 5 FLUORESCENT PROBES FOR MEASURMENT OF ACP INTERACTION WITH LpxA AND AcpS

5.1 Introduction and Rationale

As described in Chapter 3 and in previous work (Gong et al., 2008), introduction of a single well placed Trp residue can provide an informative probe of the conformation and interactions of ACP. Fluorescence can also provide information (albeit less easily interpreted) about the environment of the lone Tyr residue in ACP and site-directed mutants thereof (Chapter 4), but Tyr fluorescence is not suitable for studying ACP interactions with its partner enzymes due to the relative abundance of this amino acid in most proteins. To extend the utility of Trp as a fluorescent probe of ACP-dependent enzymes, this chapter explores the introduction of Trp into two enzymes that naturally lack this residue: *E. coli* LpxA (*E.c.*LpxA) and *V. fischeri* AcpS (*V.f.*AcpS). In addition, molecular modeling of the apicoplast AcpS enzyme from the malaria parasite *Plasmodium falciparum*, a promising anti-malarial drug target, was also performed.

5.2 Results: LpxA

As noted in the Introduction, UDP-N-acetylglucosamine acetyltransferase (LpxA) catalyzes the initial step in the synthesis of Gram negative bacterial lipid A, and is thus an attractive antibacterial drug target due to the unique, essential, and toxic nature of this product. Wild-type *E.c.*LpxA contains no endogenous tryptophan residues, thus Trp can be substituted into various positions in LpxA as binding probes. Four positions were chosen for the Trp to function as a probe of binding of the different reaction components (Figure 74). Q104 (Figure 74; red) and F162 (Figure 74; blue) were chosen for their

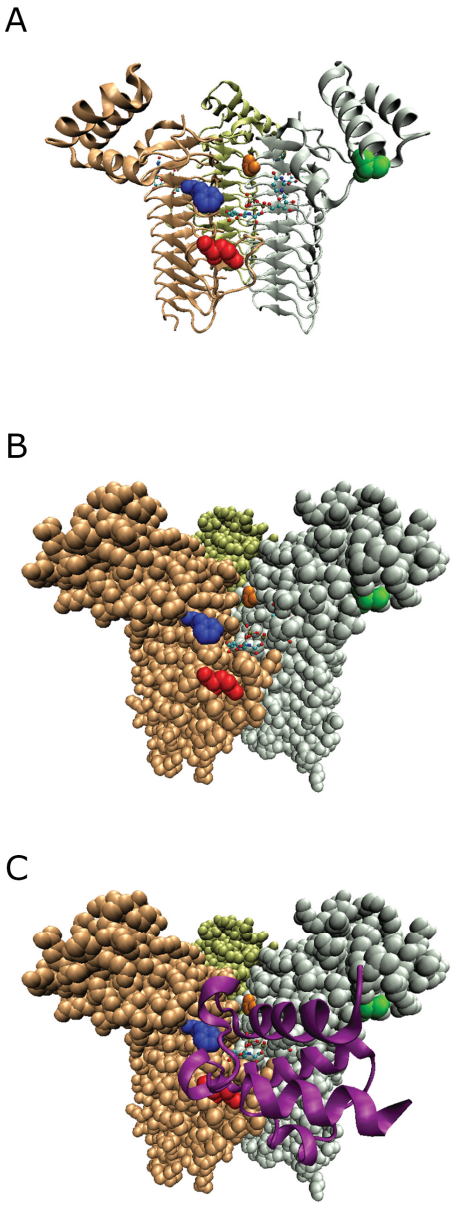

Figure 74 Various representations of the LpxA trimer bound to UDP-Glc*N*Ac and ACP showing the positions chosen for Trp-substitution. A, B) LpxA bound with UDP-Glc*N*Ac. C) LpxA bound with UDP-Glc*N*Ac and ACP. The positions selected for Trp-substitution are highlighted in each panel: Q104 (red), F162 (blue), G173 (orange), and S208W (green). LpxA is displayed using "NewCartoon" (A) or "VDW" (B, C) representation and is colored by subunit (peach, silver, pale yellow). UDP-Glc*N*Ac is displayed using "CPK" representation and is colored by atom in all panels. ACP is displayed using "NewCartoon" graphical representation. Two structures were used to create this figure, LpxA bound to UDP-Glc*N*Ac (PDB ID = 2JF3 (Ulaganathan et al., 2007)) and ACP docked to LpxA (created and provided by Dr. Christopher Barden (DeNovaMed Inc.)). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

position relative to both ACP and UDP-Glc*N*Ac, while S208 (Figure 74; green) is only in proximity to ACP. Lastly, G173W (Figure 74; orange) was chosen for its position at the end of the fatty acid binding site of LpxA.

5.2.1 Enzyme Activity

To examine the functional effect of Trp-substitutions into various positions of LpxA, its enzyme activity was measured (Figure 75). All Trp-substitutions caused a decrease in activity, with the exception of F162W (Figure 75; blue bar), which exhibited about 50% higher activity than wild-type LpxA. The activity of Q104W was about 70% of wild-type, while G173W (20% residual activity) and S208W (10%) were more greatly affected (Figure 75; red, orange, and green bars, respectively). As LpxA is not negatively affected by Trp-substitution at position 162, F162W LpxA was used for most fluorescence binding analyses (below).

5.2.2 Steady-State Trp Fluorescence

5.2.2.1 Emission Spectra of Trp-Substituted LpxAs

The fluorescence emission spectrum of Trp is sensitive to its environment. Steady-state excitation of Trp-substituted LpxAs resulted in spectra indicative of a solvent exposed Trp in most cases (peak wavelength ~350 nm), with S208W being slightly more solvent-shielded (peak wavelength ~345 nm) (Figure 76). This is somewhat surprising due to the apparent exposed location of S208W (Figure 74; green). Based on the structure of *E.c.*LpxA, it was expected that if Trp in any of the chosen positions would be solvent shielded, it would be G173W (Figure 74; orange).

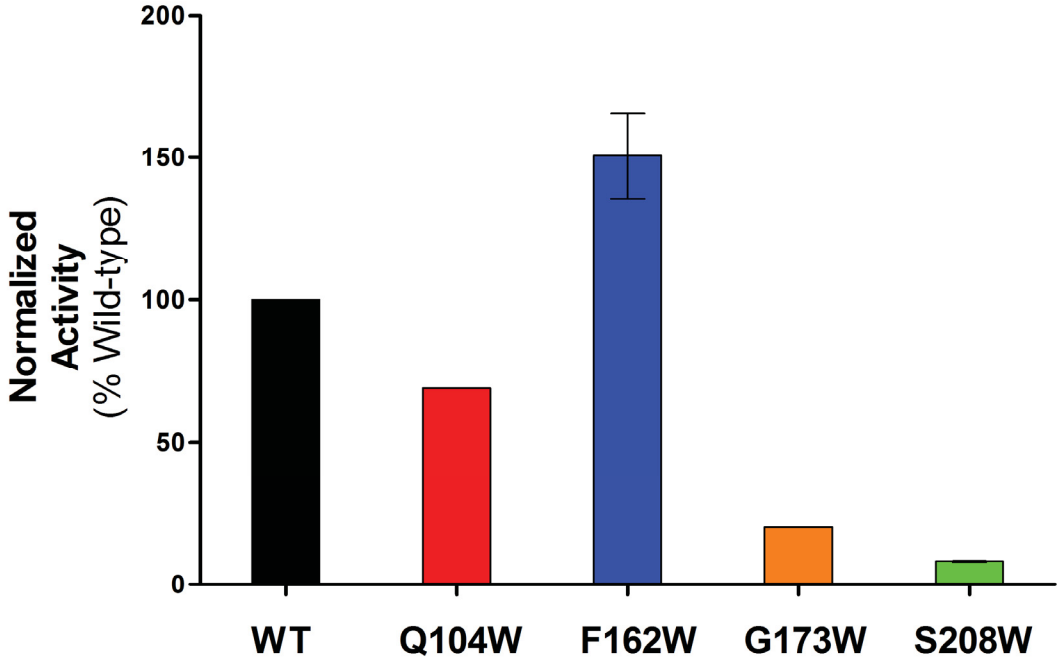

Figure 75 Normalized activity of wild-type and Trp-substituted LpxAs. The activity of wild-type (black), Q104W (red), F162W (blue), G173W (orange), and S208W (green) was determined as per Section 2.4.4.1. Activity values for each mutant LpxA are expressed as a percentage of wild-type LpxA measured on the same day. Error bars shown represent standard deviation (n=3) while values without error bars are the mean of duplicate measurements.

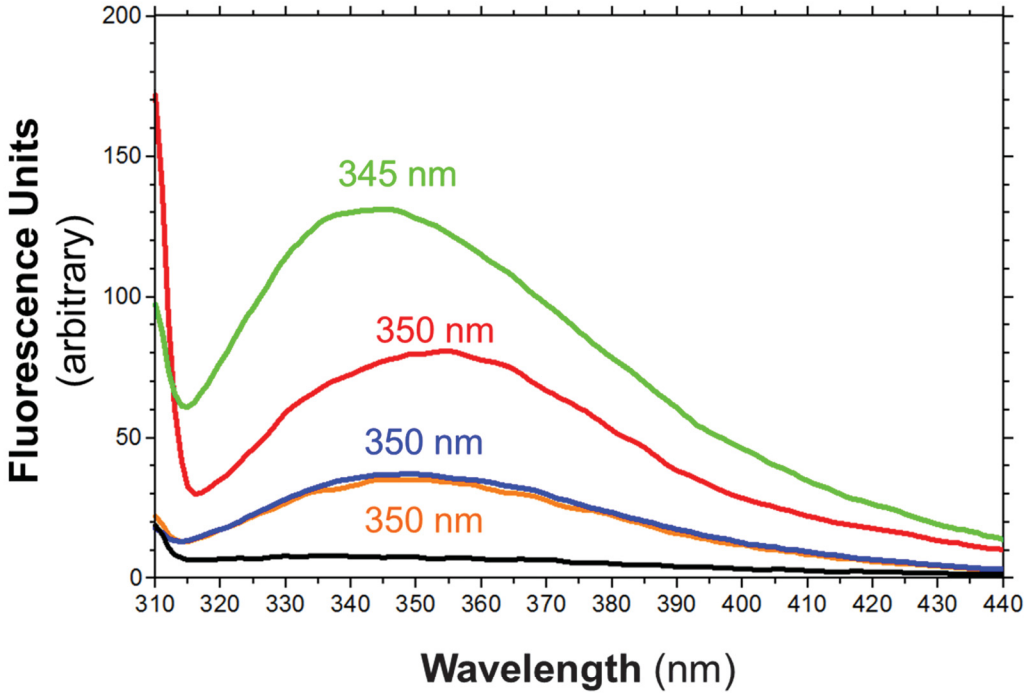

Figure 76 Intrinsic tryptophan fluorescence of wild-type and mutant LpxAs. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for wild-type (black), Q104W (red), F162W (blue), G173W (orange) and S208W (green). Peak wavelength is given for each Trp-containing LpxA. Spectra shown are representative of multiple spectra recorded on different days. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

5.2.2.2 Titration of LpxA with Wild-type and Mutant ACPs

F162W was titrated with wild-type holo-ACP (h-rACP), as well as with ACP mutants representative of divalent cation sites A (h-D30N) and B (h-D56N) (Figure 77, Figure 78 and Figure 79, respectively). Titrations were done in both the absence and presence (Figure 77 – Figure 79A, B, respectively) of the co-substrate UDP-Glc*N*Ac. Q104W and G173W also show activity, although much less than F162W (Figure 75), and thus were only used for some experiments (Figure 80 and Figure 81). S208W was not used for titration experiments as it was essentially inactive (Figure 75).

All Trp-substituted LpxAs show a saturable decrease in fluorescence upon ACP binding with no apparent change in peak emission wavelength (Figure 77 – Figure 81). Apparent K_D values for binding of ACP to LpxA in the absence or presence of co-substrate UDP-Glc*N*Ac were calculated (Table 15, Figure 77C, D – Figure 81C, D), by plotting change in area under the curve *versus* ACP concentration using GraphPad™ Prism®'s non-linear regression and Equation 1 ("Pre-Python Analysis"), and by using a Python program (Appendix 2) which uses non-linear regression (Equation 1) and iterative calculation "Post-Python Analysis"). Although these results were obtained with the less precise LS50B fluorometer, they indicate that F162W interacts with rACP, D30N and D56N with similar affinity in the absence of UDP-GlcNAc, while a lower affinity for the mutant ACPs was noted in the presence of this co-substrate. The apparent affinity of rACP for Q104W was similar to that with F162W, while that for G173W was somewhat lower (Table 15).

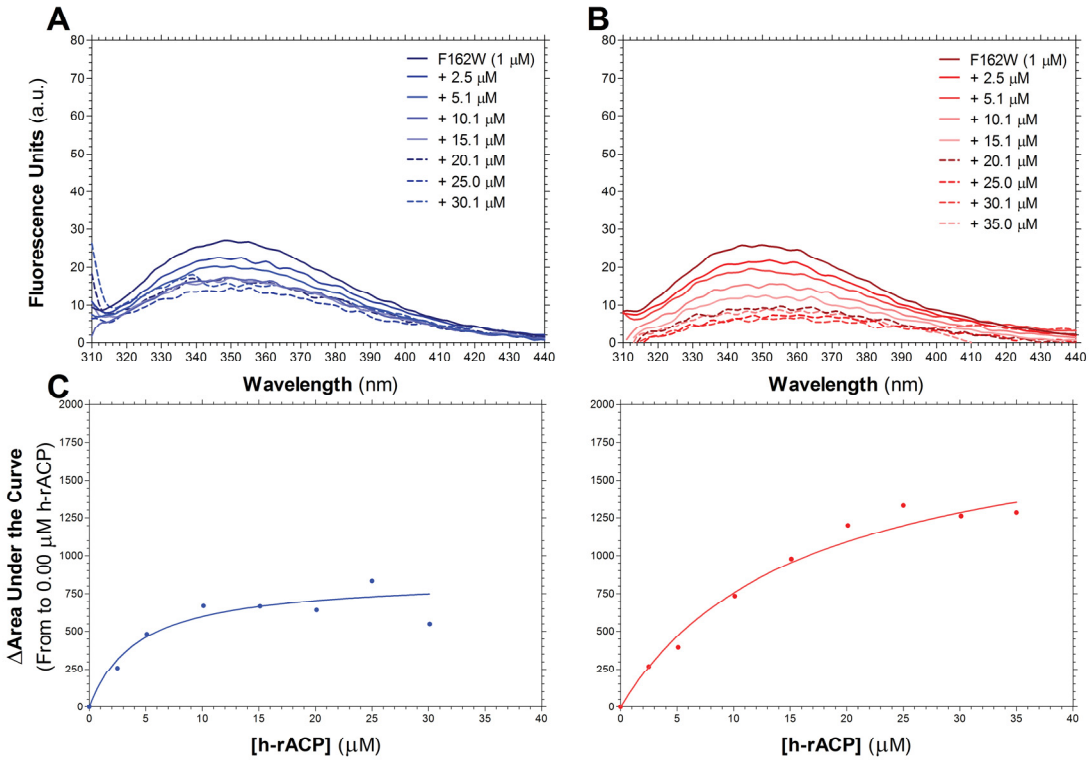

Figure 77 Changes in F162W tryptophan fluorescence upon titration with increasing concentrations of h-rACP. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for F162W in the absence (A) and presence (B) of UDP-Glc*N*Ac. Change in area under the curve K_D determination in the absence (C) and presence (D) of UDP-Glc*N*Ac, based on the spectra shown in A and B. Spectra shown are representative of multiple spectra recorded on different days. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

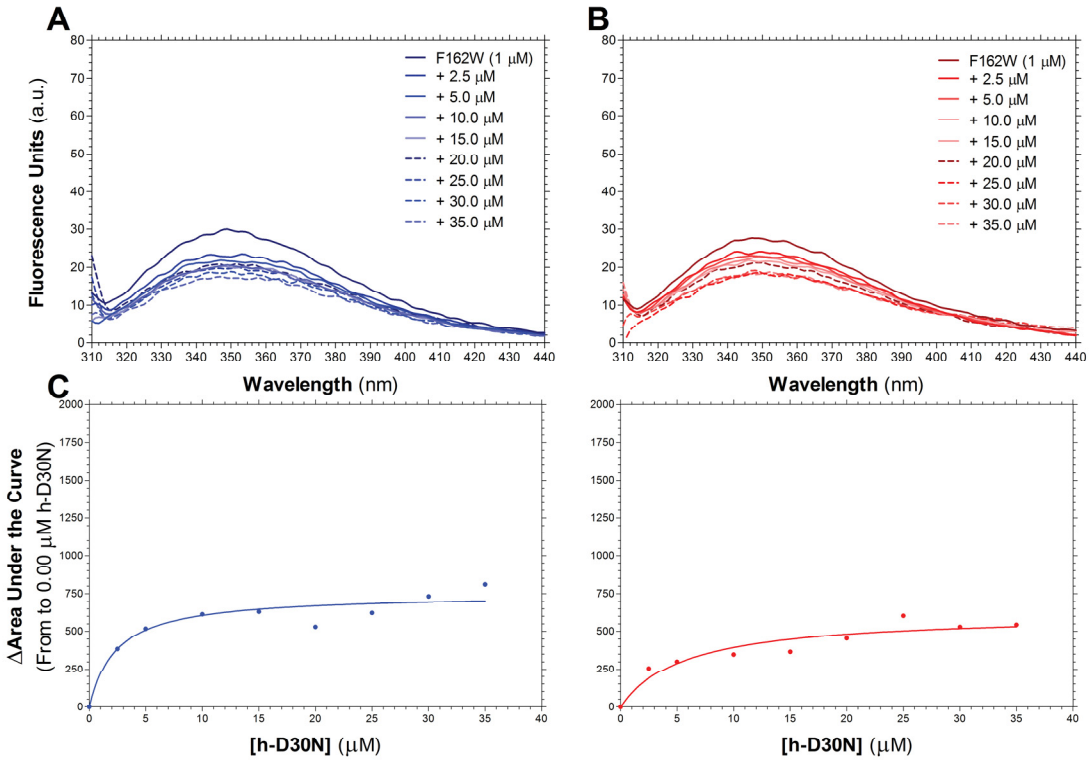

Figure 78 Changes in F162W tryptophan fluorescence upon titration with increasing concentrations of h-D30N. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for F162W in the absence (A) and presence (B) of UDP-Glc*N*Ac. Change in area under the curve K_D determination in the absence (C) and presence (D) of UDP-Glc*N*Ac, based on the spectra shown in A and B. Spectra shown are representative of multiple spectra recorded on different days. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

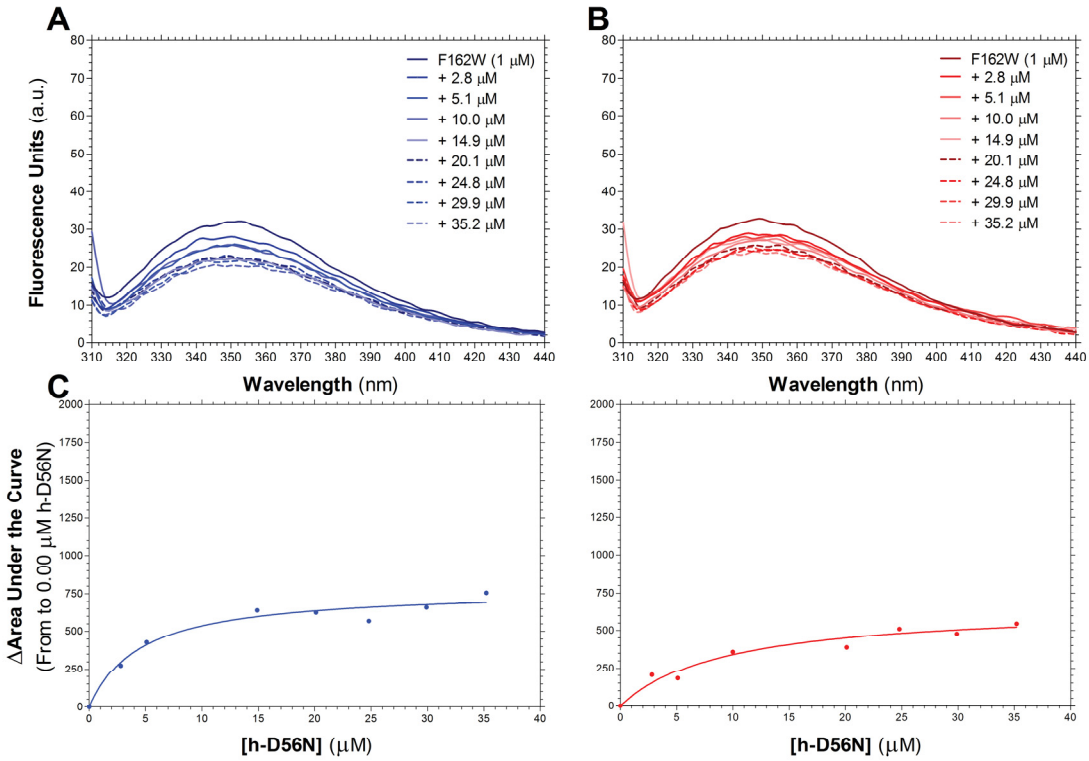

Figure 79 Changes in F162W tryptophan fluorescence upon titration with increasing concentrations of h-D56N. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for F162W in the absence (A) and presence (B) of UDP-Glc*N*Ac. Change in area under the curve K_D determination in the absence (C) and presence (D) of UDP-Glc*N*Ac, based on the spectra shown in A and B. Spectra shown are representative of multiple spectra recorded on different days. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

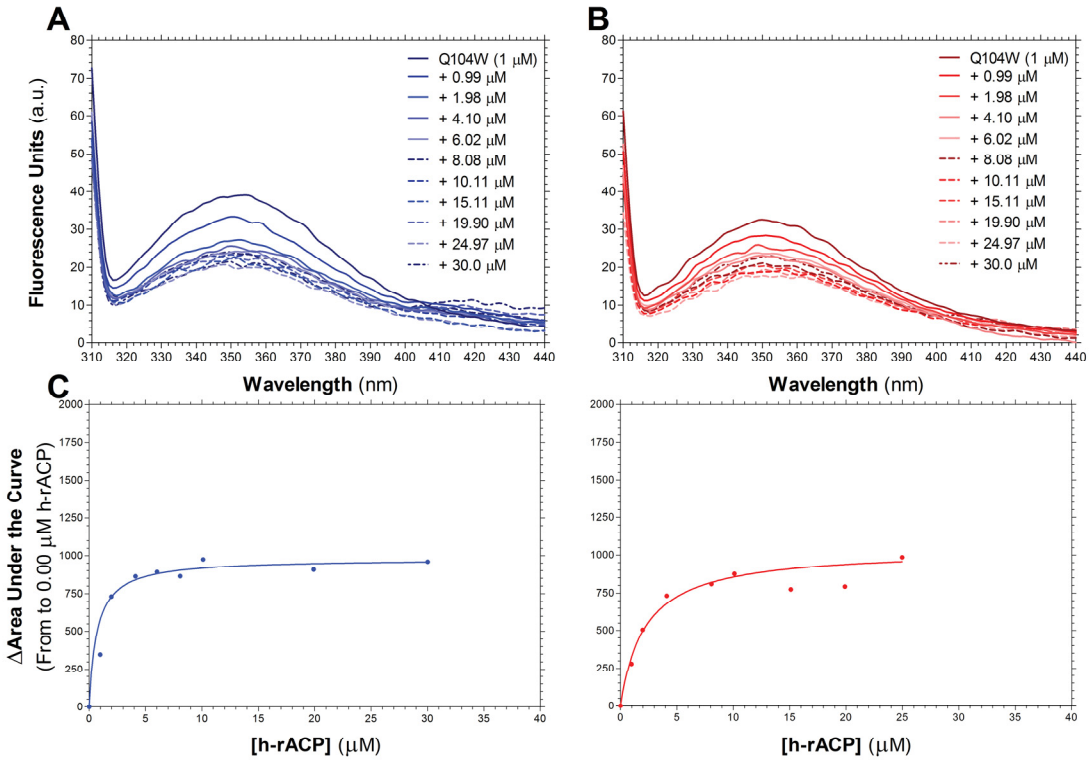

Figure 80 Changes in Q104W tryptophan fluorescence upon titration with increasing concentrations of h-rACP. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for Q104W in the absence (A) and presence (B) of UDP-Glc*N*Ac. Change in area under the curve K_D determination in the absence (C) and presence (D) of UDP-Glc*N*Ac, based on the spectra shown in A and B. Spectra shown are representative of multiple spectra recorded on different days. All spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

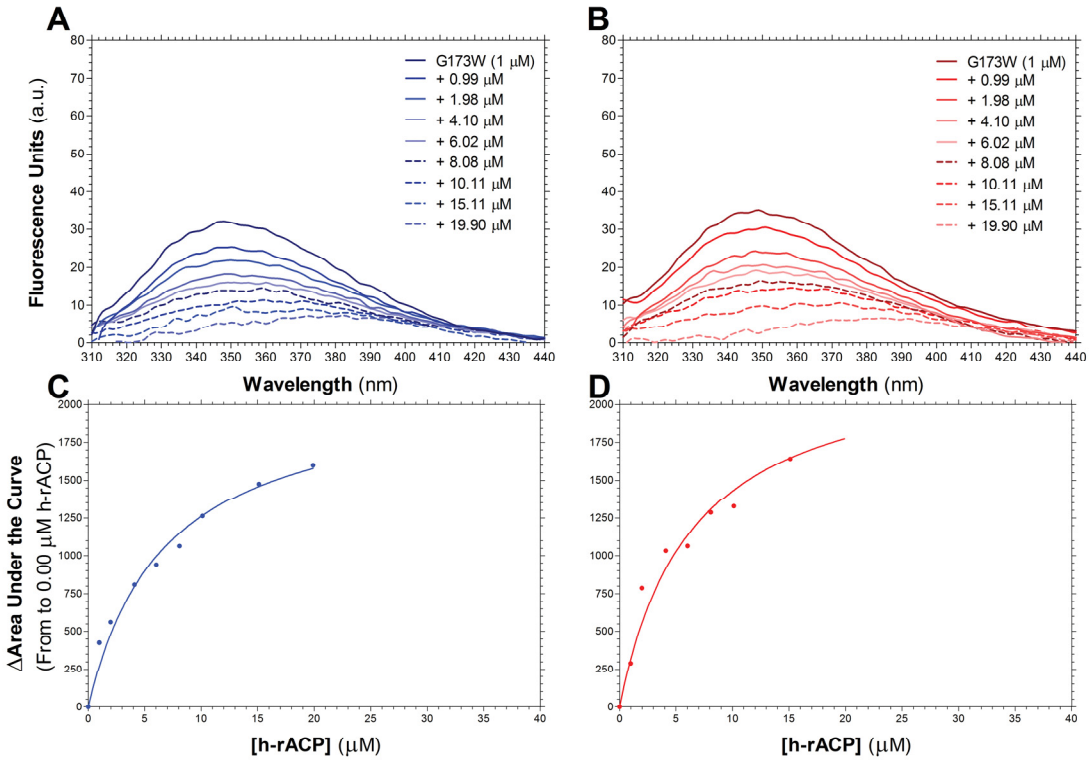

Figure 81 Changes in G173W tryptophan fluorescence upon titration with increasing concentrations of h-rACP. Trp fluorescence spectra (λ_{ex} = 296 nm) were recorded for G173W in the absence (A) and presence (B) of UDP-Glc*N*Ac. Change in area under the curve K_D determination in the absence (C) and presence (D) of UDP-Glc*N*Ac, based on the spectra shown in A and B. Spectra were recorded on the LS50B and have been corrected for signal arising from the Raman band.

Table 15 Calculated K_D values for LpxAs titrated with wild-type and mutant ACPs in the absence and presence (200 μM) of UDP-Glc*N*Ac.

LpxA	ACP	K_Ds in the Absence of UDP-Glc*N*Ac		K_Ds in the Presence of UDP-Glc*N*Ac	
		Pre-Python Analysis[1] (μM)	Post-Python Analysis[2] (μM)	Pre-Python Analysis[1] (μM)	Post-Python Analysis[2] (μM)
F162W	h-rACP	3.2 ± 1.6	2.6 ± 0.9	4.4 ± 2.4	4.4 ± 2.5
	h-D30N	3.9	3.0	13.4	11.6
	h-D56N	3.7	3.7	9.7	9.1
Q104W	h-rACP	3.0	5.5	2.8	3.7
G173W	h-rACP	6.8	6.0	6.4	6.6

[1] Values were determined by GraphPad™ Prism® 5 and error given is standard deviation ($n \geq 3$). Duplicate titrations were performed for F162W titrated with h-D56N and h-D30N, thus no error is given. Similarly, only single or duplicate values were obtained from titrations of Q104W and G173W, thus no error is given.

[2] Values were determined using a python program I developed (Appendix 2) and error given is standard deviation ($n \geq 3$). Only single or duplicate values were obtained from titrations of Q104W and G173W, thus no error is given.

5.2.3 Results: Circular Dichroism of Trp-Substituted LpxAs

The secondary structure of LpxA consists of a left-handed parallel β-helix (Raetz & Roderick, 1995). Due to this unusual secondary structure, normal CD deconvolution methods that calculate percentages of α-helix and β-sheet were not used. However, to ensure that Trp-substitution into various positions of LpxA doesn't greatly affect higher order (e.g. secondary) structure, CD spectra were recorded for wild-type and mutant LpxAs (Figure 82). With the possible exception of S208W, the shape of the CD curve for all Trp-substituted LpxAs was similar to wild-type LpxA (Figure 82) indicating that Trp-substitution into these positions does not dramatically alter protein secondary structure.

5.2.4 Results: Gel Filtration of Trp-Substituted LpxAs

Gel filtration analysis was also performed with wild-type and mutant LpxAs to ensure that Trp-substitution did not affect trimer formation. All Trp-substituted LpxAs eluted from the size exclusion column in the same volume as wild-type LpxA (Figure 83). Based on the standards, elution at this point is consistent with the expected size of an LpxA trimer (84 kDa). The similar elution position of all proteins suggests that Trp-substitution does not affect the quaternary structure of LpxA.

5.2.4 *In Silico* Creation and Simulation of Trp-Substituted LpxA PDB Files

To provide an independent evaluation of the effect of Trp replacement on LpxA structure and to facilitate interpretation, the appropriate mutations were made to *E.c.*LpxA (Figure 84A) *in silico* to produce Q104W (Figure 84B), F162W (Figure 84C),

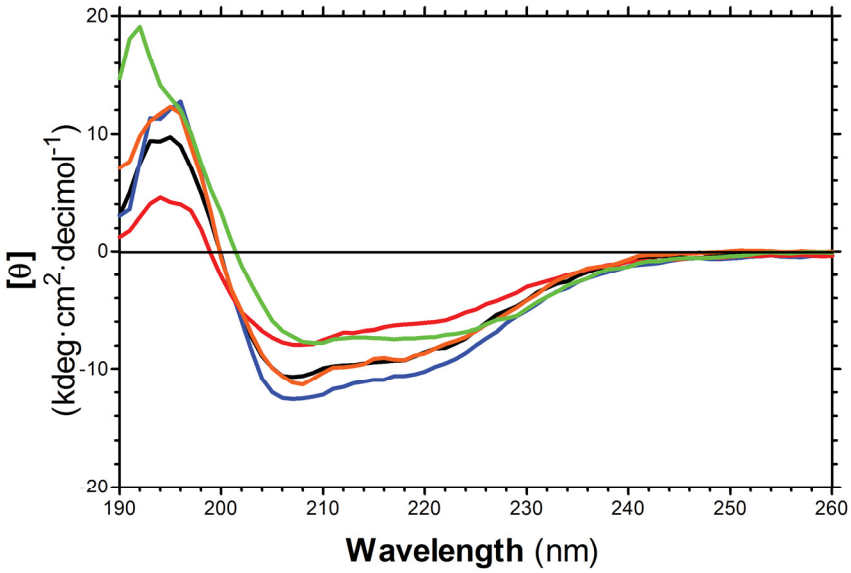

Figure 82 CD analysis of wild-type and Trp-substituted LpxAs. CD spectra of wild-type (black), Q104W (red), F162W (blue), G173W (orange), and S208W (green) were measured at ~1 μM.

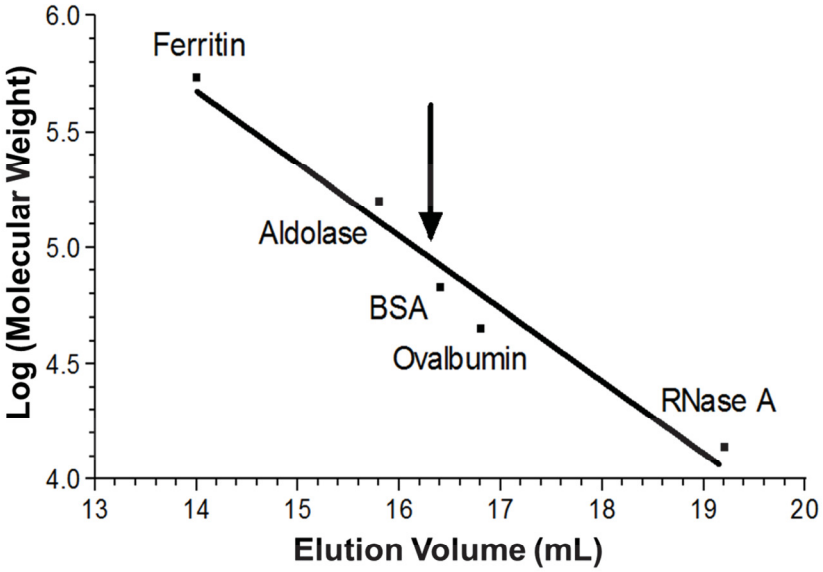

Figure 83 Gel filtration analysis of wild-type and Trp-substituted LpxAs. Elution volume of standards (ferritin, aldolase, BSA, ovalbumin and RNase A) are indicated. The arrow indicates elution point for wild-type, Q104W, F162W, G173W and S208W LpxAs at 16.2 mL. Data shown is representative of multiple experiments.

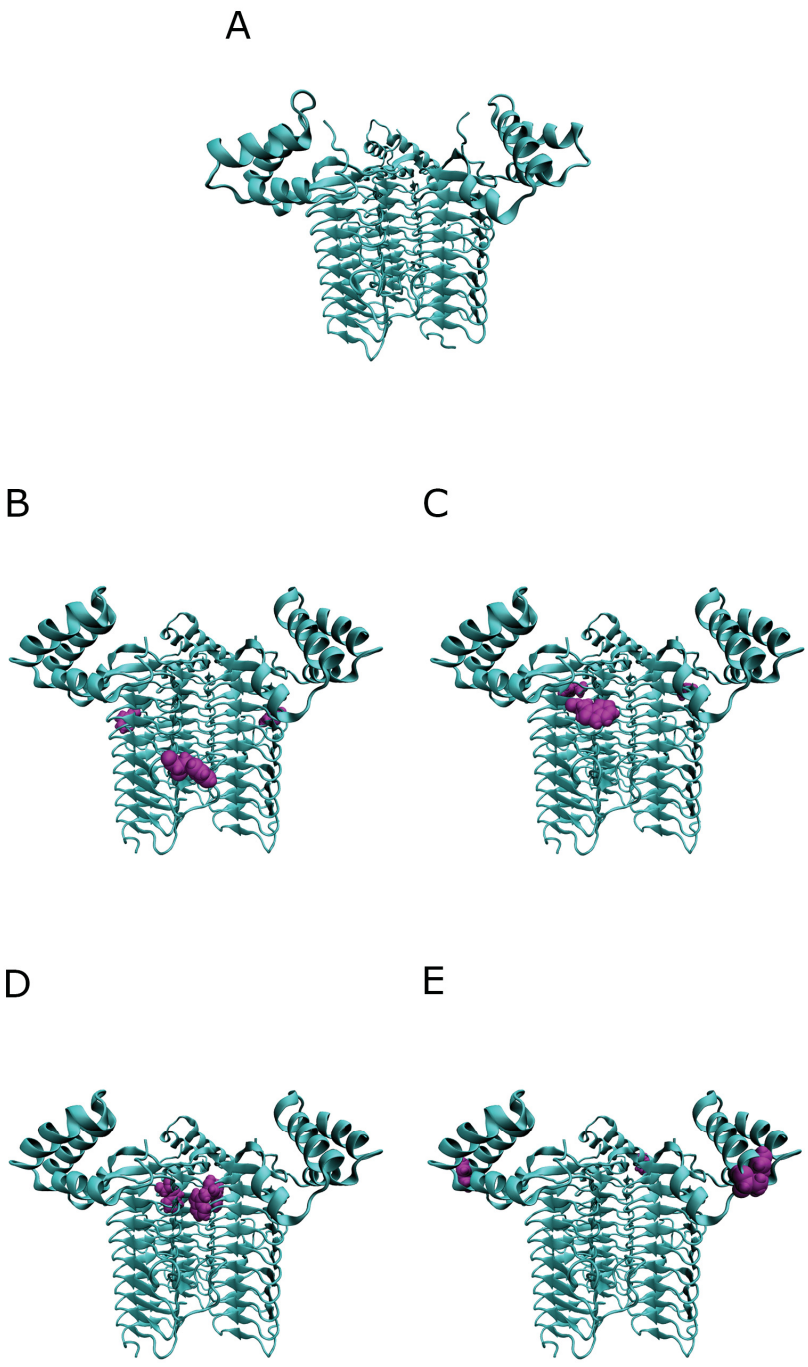

Figure 84 *E.c.*LpxA Trp-substituted LpxA models created *in silico* pre-MD simulation. "NewCartoon" representations of wild-type (A), Q104W (B), F162W (C), G173W (D), and S208W (E). Starting structure for *in silico* mutation is wild-type *E. coli* LpxA (A; PDB ID: 2JF3 (Ulaganathan et al., 2007)). Trp is highlighted in purple using "VDW" graphical representation (B, C, D, E). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

G172W (Figure 84D) and S208W (Figure 84E) LpxAs. Simulations of these newly created PDB files (as well as wild-type LpxA) were performed over a 1.02 ns period (20 ps minimization; 1 ns equilibration; as described in Section 2.5.2). As previously, plotting RMSD over the course of the simulation indicates that an equilibrium structure was being approached in all cases (Figure 85). No substantial local rearrangements are caused by Trp-substitution (Figure 86); as indicated by RMSD for each Trp-substituted LpxA compared to wild-type LpxA: Q104W (0.146 Å), F162W (0.147 Å), G173W (0.158 Å), and S208W (0.150 Å).

5.3 Results: AcpS

Similar to *E.c.*LpxA, wild-type *V.f.*AcpS does not contain any endogenous Trp residues. Thus, a Trp residue can in principle be inserted in the binding site as a probe of ACP binding. Phe 27 (Figure 87; purple) was chosen as the site of replacement as it is not directly involved in binding ACP but is near the active site at the trimer subunit interface. Additionally, the change of Phe to Trp would be a relatively conservative substitution and is therefore less likely to perturb ACP binding.

In addition to F27W, a charge-change mutant of AcpS (R22E; Figure 87B; cyan) was designed with the idea of creating specific AcpS-ACP reaction pairs based on known ionic interactions in enzyme-substrate binding. Specifically, Arg 21 of *B. subtilis* AcpS is known to interact with Glu 41 of ACP Helix II (Parris et al., 2000), so we would anticipate that an acidic residue at this position of AcpS would preferentially interact with the E41K ACP mutant made previously in our laboratory (Gong & Byers, 2003). Lastly, the double mutant (R22E/F27W; Figure 87B; cyan/purple, respectively)) was also created to serve as a probe for binding of E41K ACP.

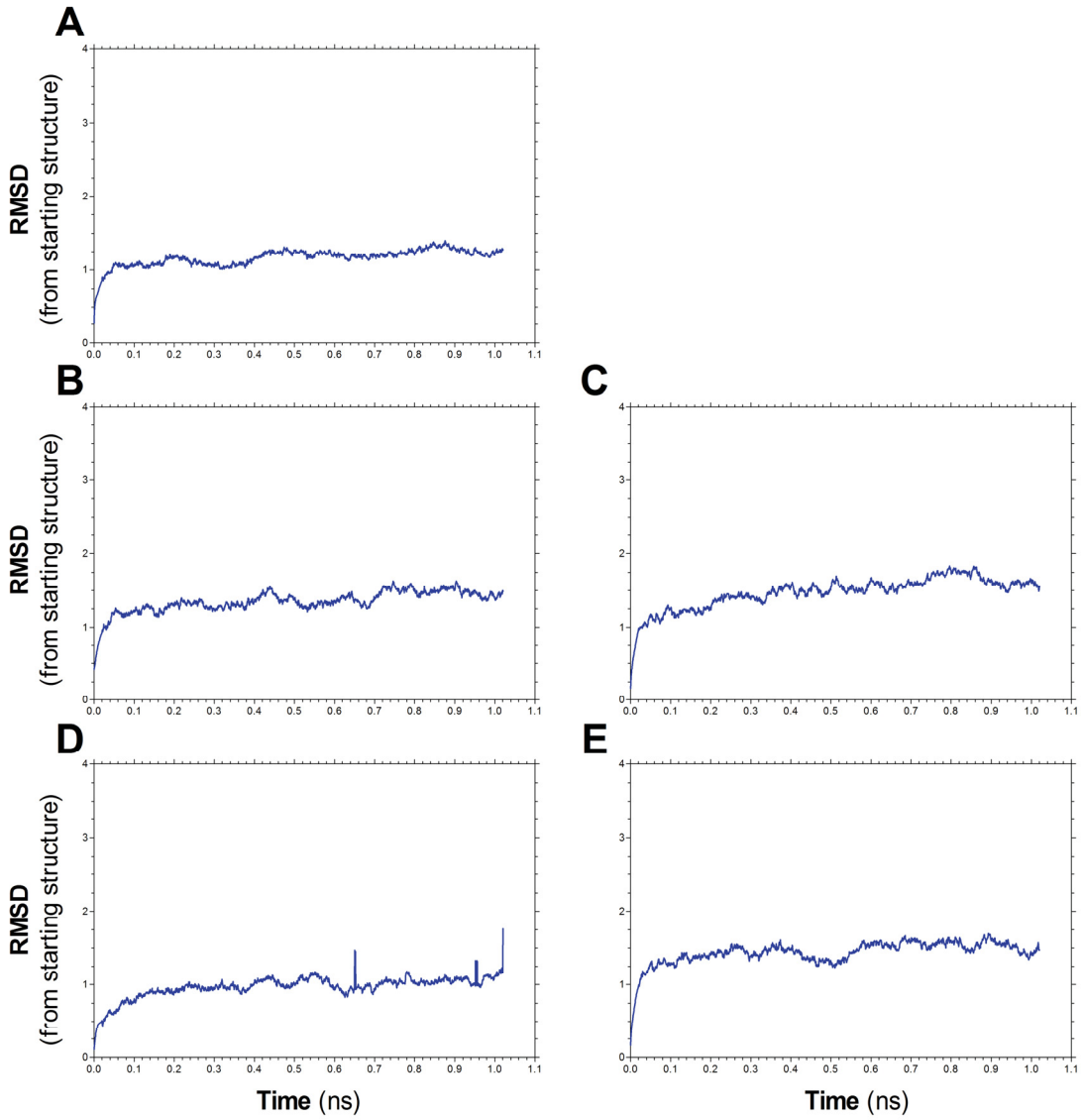

Figure 85 RMSD for backbone atoms of Trp-substituted LpxAs. RMSD was calculated from the starting structure for wild-type (A), Q104W (B), F162W (C), G173W (D) and S208W (E) LpxAs. Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

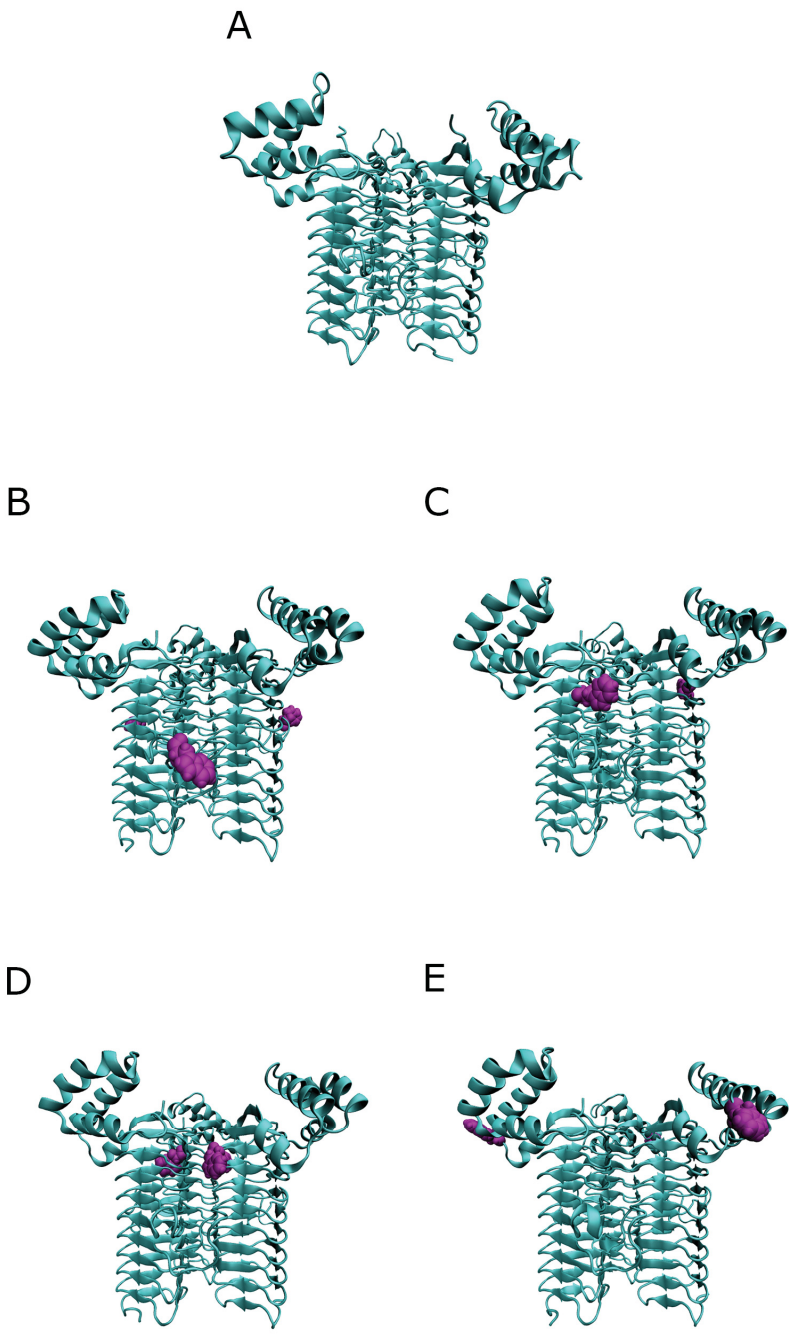

Figure 86 *E.c.*LpxA Trp-substituted LpxA models created *in silico* post-MD simulation. "NewCartoon" representations of wild-type (A), Q104W (B), F162W (C), G173W (D), and S208W (E) LpxAs. Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Trp is highlighted in purple using "VDW" graphical representation (B, C, D, E). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

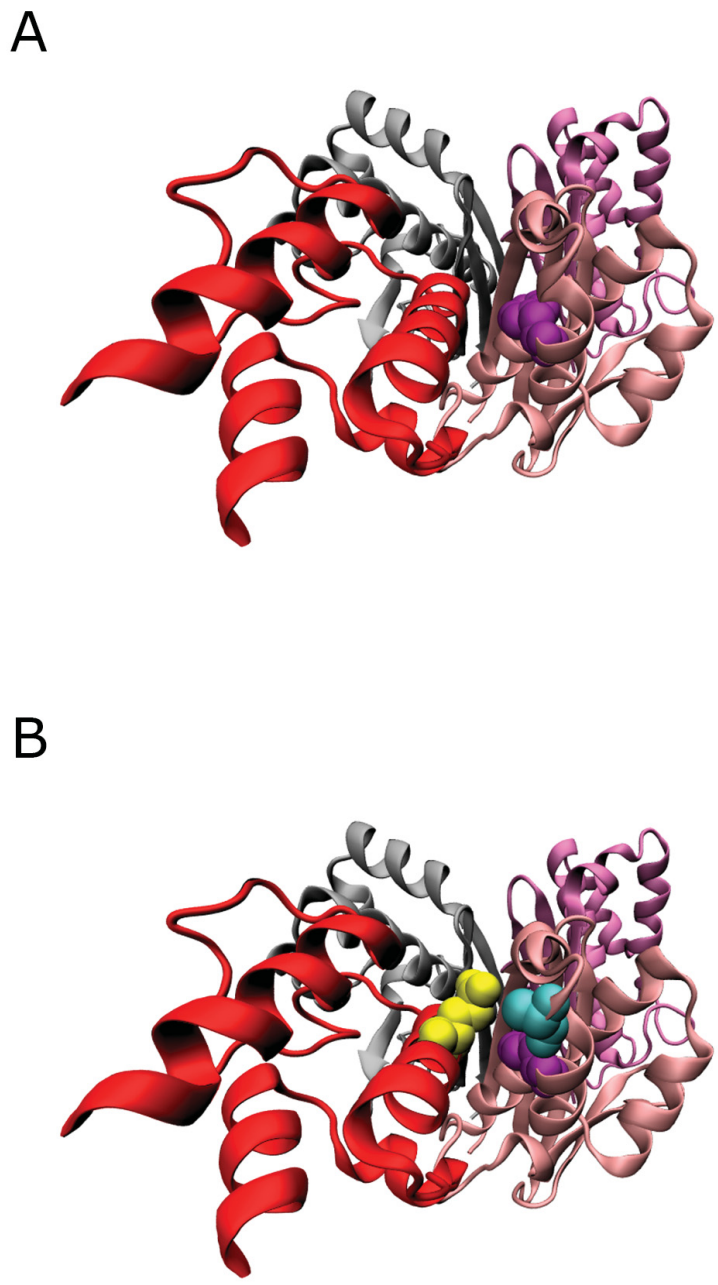

Figure 87 AcpS bound to ACP showing the positions chosen for mutation. A) NewCartoon representation of AcpS bound with ACP with position of Trp-substitution shown (purple; VDW representation). B) NewCartoon representation of AcpS bound with ACP with position of Trp-substitution (F27; purple) and charge-change mutations (R22 of AcpS (cyan) and E41 of ACP (yellow)). AcpS is colored by subunit (pink, magenta, silver) in all panels. For simplicity, only 1 ACP molecule (red) is shown. One PDB structure was used to create this figure, *B.s.*AcpS bound to ACP (PDB ID = 1F80 (Parris et al., 2000)). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

5.3.1 Enzyme Activity

As with LpxAs, the activity of Trp-substituted and charge-change mutants were tested for activity to observe the effect of each mutation holo-ACP synthesis. All mutant AcpSs demonstrated some activity; however, wild-type AcpS wasn't active and thus the activity of the mutants relative to wild-type cannot be calculated; thus, these data are not shown.

5.3.2 Molecular Dynamics

5.3.2.1 *In Silico* Creation and Simulation of Trp-Substituted and Charge-Change AcpS PDB Files

As for LpxA, AcpS mutations made *in vitro* were examined *in silico*. Both Trp-substitution and charge-change mutant PDB files were created along with their respective charge-change mutant of ACP. The known crystal structure of *B. subtilis* AcpS (*B.s.*AcpS) with ACP bound (Parris et al., 2000) was used to make a variety of AcpS-ACP mutational combinations (see Table 11 for list of combinations and Figure 88, Figure 89, and Figure 90 for pre-simulation images). *B.s.*AcpS shares 44% sequence identity with *V.f.*AcpS. Simulations of these AcpS-ACP combinations were performed over a 0.45 ns period (0.20 ns minimization; 0.25 ns equilibration; as described in Section 2.5.2). Upon simulation completion, the RMSD was measured over the course of the simulation (Figure 91, Figure 92, and Figure 93). These plots indicate that more simulation time is necessary for an equilibrium structure to be reached. This was not unexpected, but due to lack of local computing power at the time, short simulations were necessary. As above, structures were saved post-simulation (Figure 94, Figure 95, and

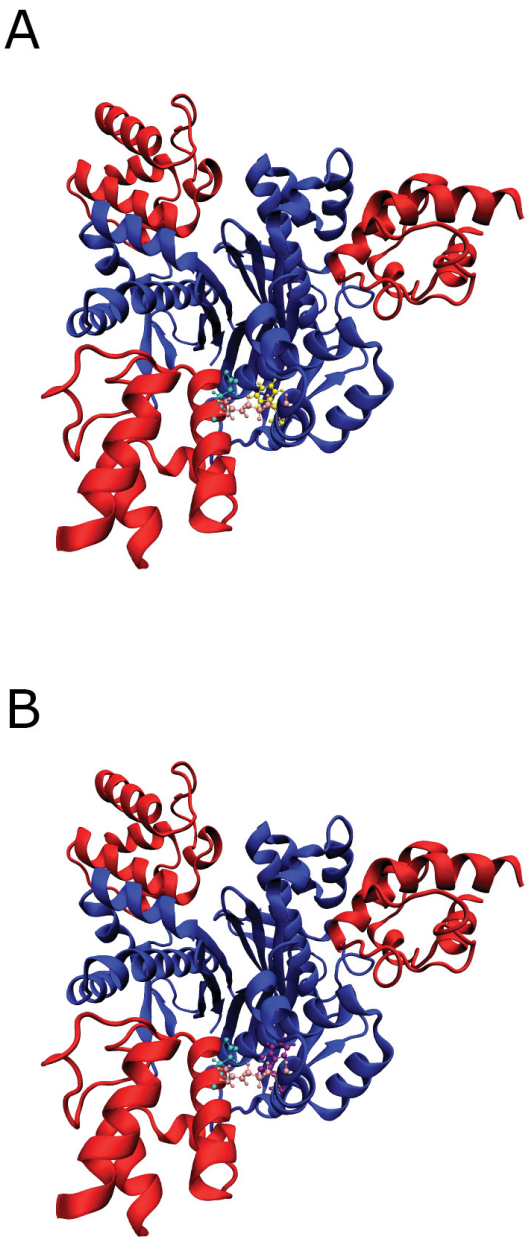

Figure 88 *B.s.*AcpS Trp-substituted model created *in silico* pre-MD simulation. "NewCartoon" representations of AcpS (blue)/ACP (red): A) WT AcpS/wild-type ACP, B) F25W AcpS/wild-type ACP. Starting structure for *in silico* mutation is wild-type *B. subtilis* AcpS bound by ACP (A; PDB ID: 1F80). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using "CPK" graphical representation: Trp 25 (purple), Phe 25 (yellow), Glu 41 (cyan), Arg 21 (pink). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

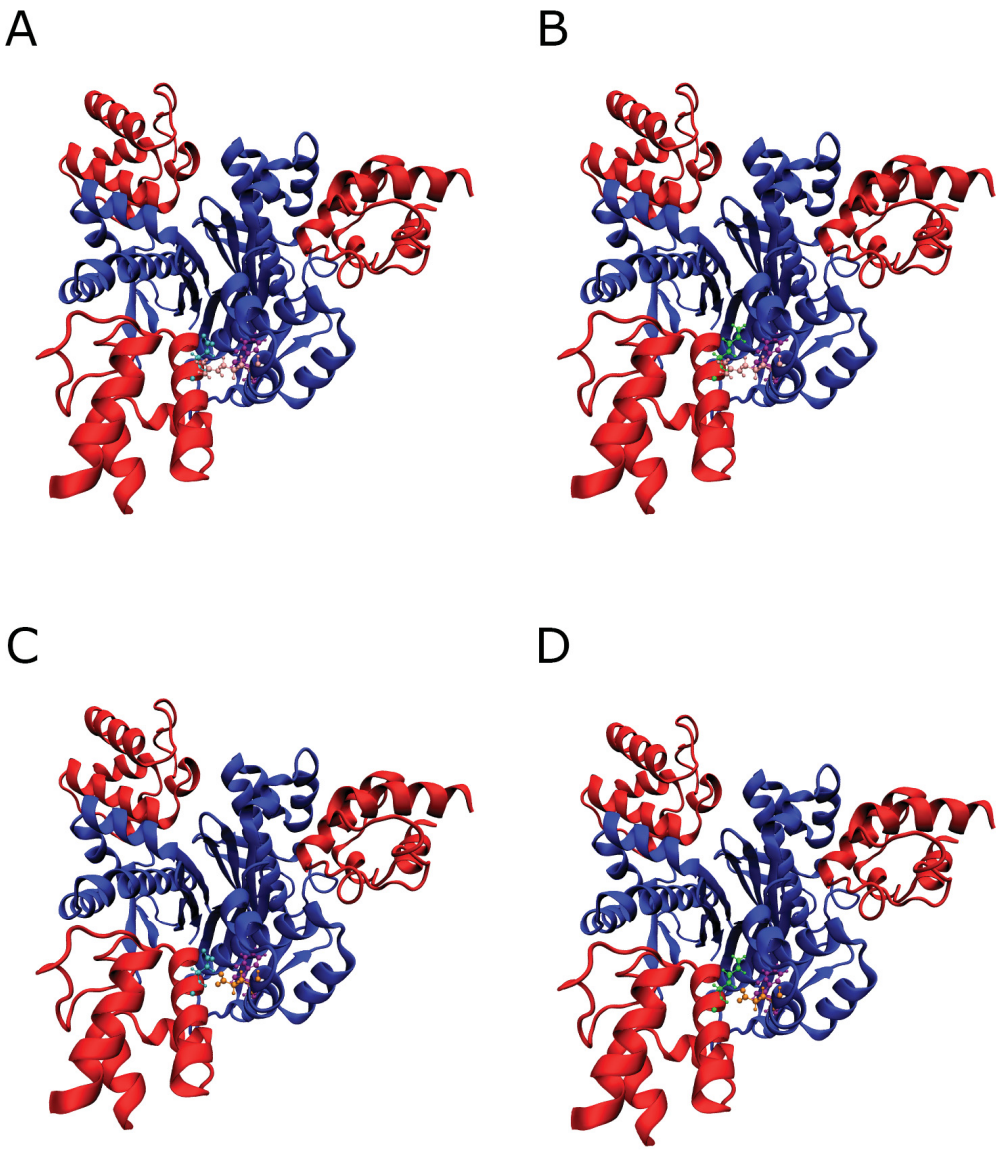

Figure 89 *B.s.*AcpS Trp-substituted charge-change model combinations created *in silico* pre-MD simulation. “NewCartoon” representations of AcpS (blue)/ACP (red) combinations: A) F25W AcpS/wild-type ACP, B) F25W AcpS/E41K ACP, C) R21E-F25W AcpS/wild-type ACP, and D) R21E- F25W AcpS/E41K ACP. Starting structure for *in silico* mutation is wild-type AcpS bound by ACP (PDB ID: 1F80). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using “CPK” graphical representation: Trp 25 (purple), Glu 41 (cyan), Lys 41 (green), Arg 21 (pink), Glu 21 (orange). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

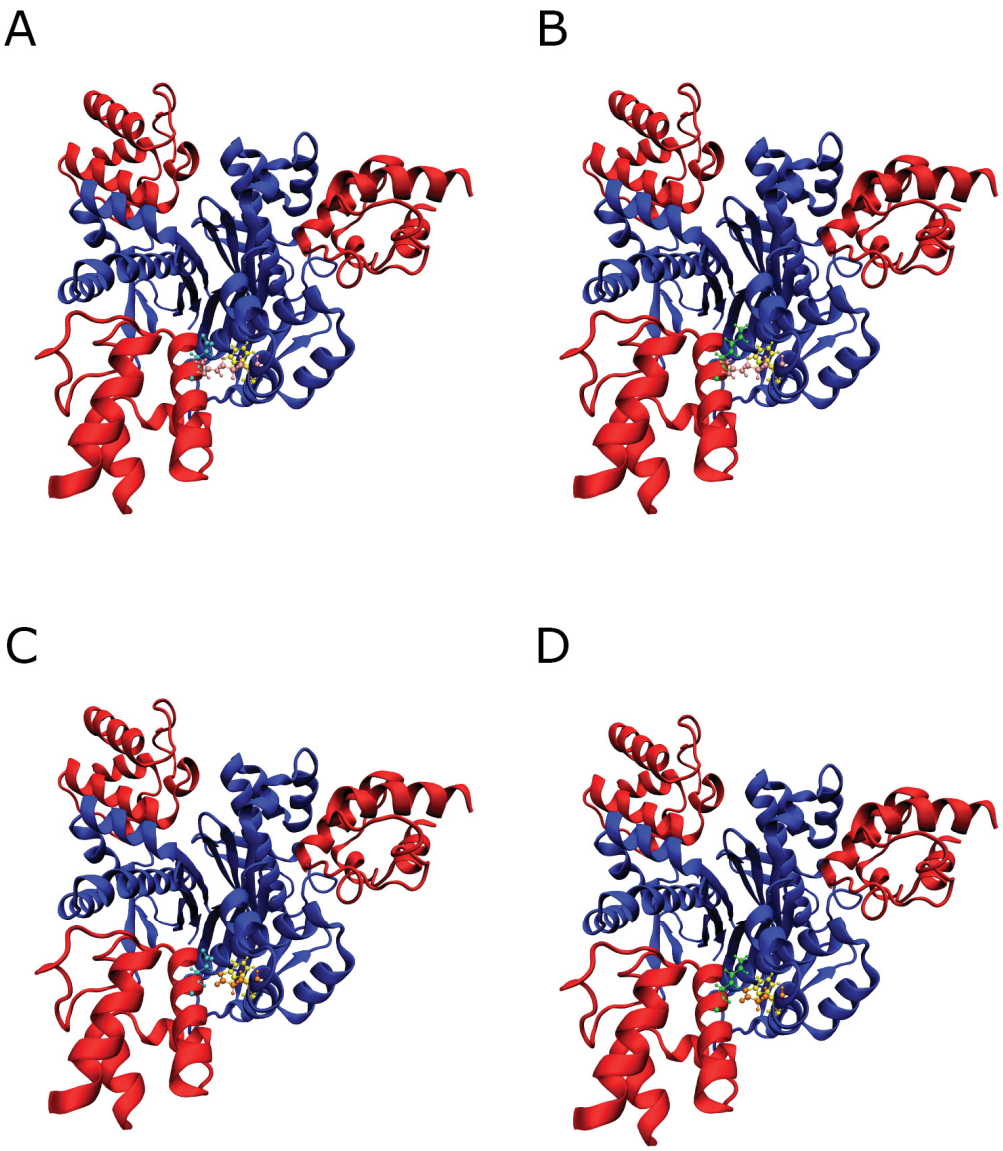

Figure 90 *B.s.*AcpS charge-change model combinations created *in silico* pre-MD simulation. "NewCartoon" representations of AcpS (blue)/ACP (red) combinations: A) wild-type AcpS/wild-type ACP, B) wild-type AcpS/E41K ACP, C) R21E AcpS/wild-type ACP, and D) R21E AcpS/E41K ACP. Starting structure for *in silico* mutation is wild-type AcpS bound by ACP (PDB ID: 1F80; A). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using "CPK" graphical representation: Phe 25 (yellow), Glu 41 (cyan), Lys 41 (green), Arg 21 (pink), Glu 21 (orange). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

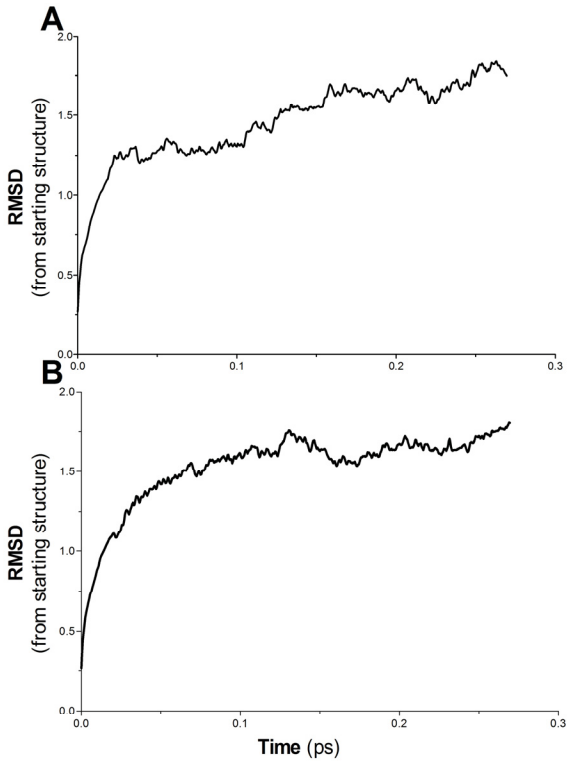

Figure 91 RMSD for backbone atoms of AcpS/ACP simulations pre- and post-Trp-substitution. RMSD was calculated from the starting structures for wild-type AcpS/ACP (A) and F25W AcpS/wild-type ACP (B). Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

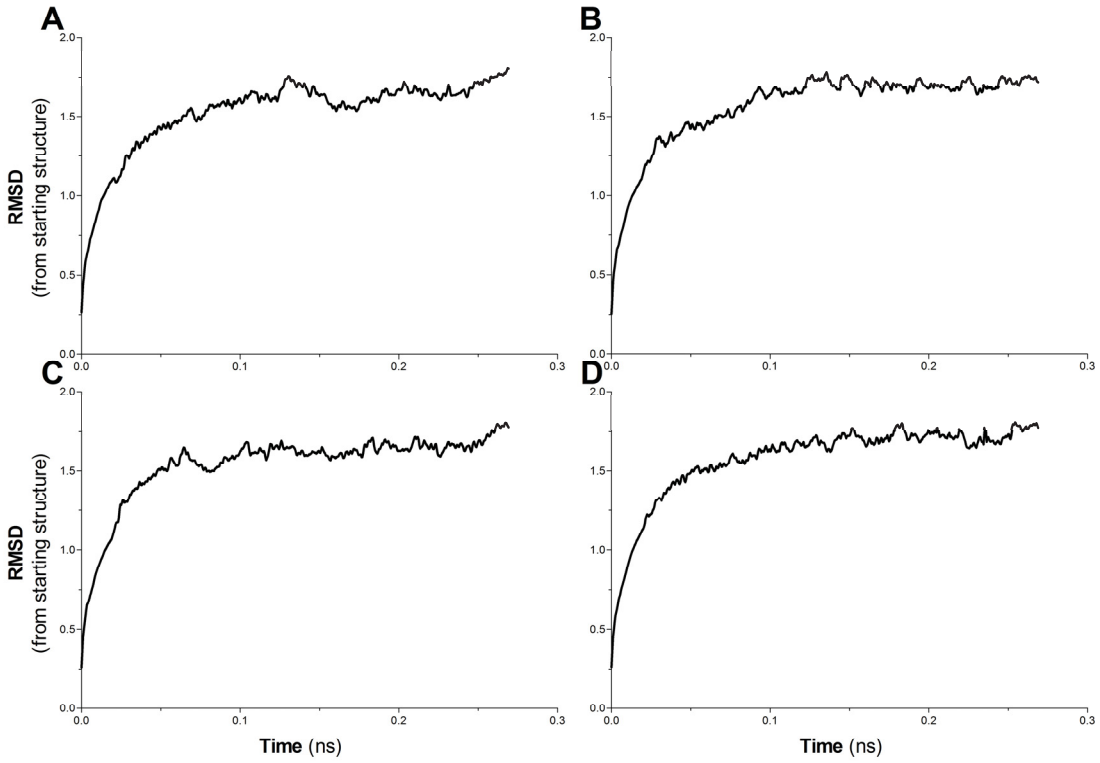

Figure 92 RMSD for backbone atoms of AcpS/ACP simulations pre- and post-Trp-substitution and charge-change mutation. RMSD was calculated from the starting structures for F25W AcpS/ACP (A), R21E-F25W AcpS/wild-type ACP (B), F25W AcpS/E41K ACP (C) and R21E-F25W AcpS/E41K ACP (D). Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

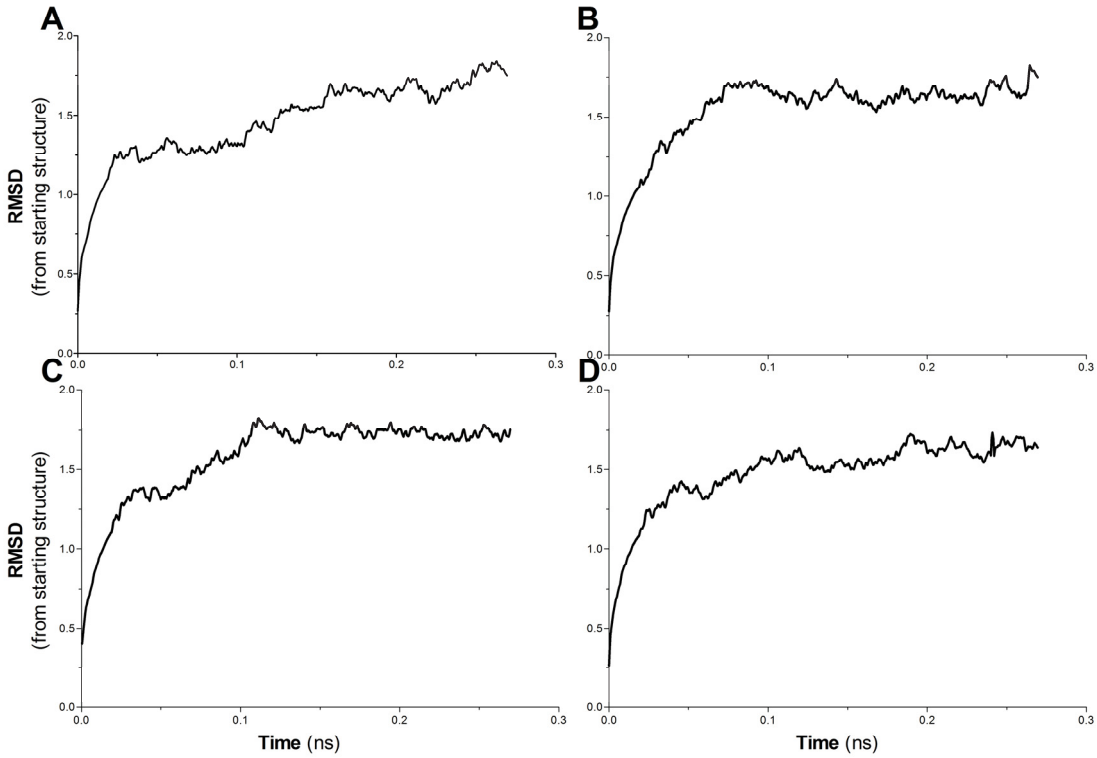

Figure 93 RMSD for backbone atoms of AcpS/ACP simulations pre- and post-charge-change mutation. RMSD was calculated from the starting structures for wild-type AcpS/ACP (A), R21E AcpS/wild-type ACP (B), wild-type AcpS/E41K ACP (C) and R21E AcpS/E41K ACP (D). Values were calculated as described in Appendix 18 using the non-smoothed trajectory file.

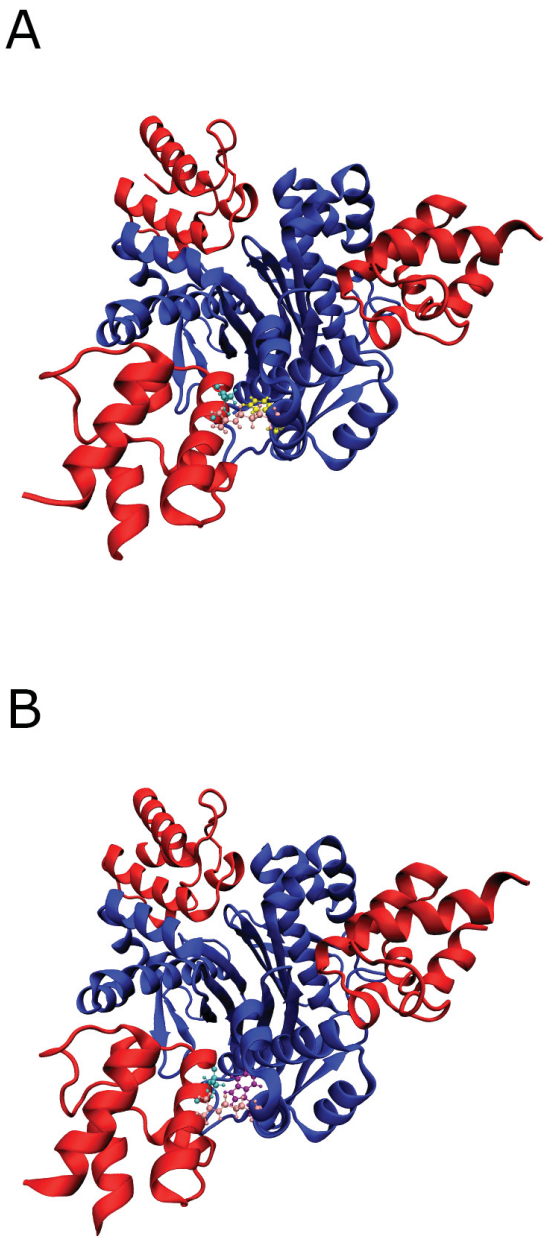

Figure 94 *B.s.*AcpS Trp-substituted model post-MD simulation. “NewCartoon” representations of AcpS (blue)/ACP (red): A) WT AcpS/wild-type ACP, B) F25W AcpS/wild-type ACP. Starting structure for *in silico* mutation is wild-type AcpS bound by ACP (A; PDB ID: 1F80). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using “CPK” graphical representation: Trp 25 (purple), Phe 25 (yellow), Glu 41 (cyan), Arg 21 (pink). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

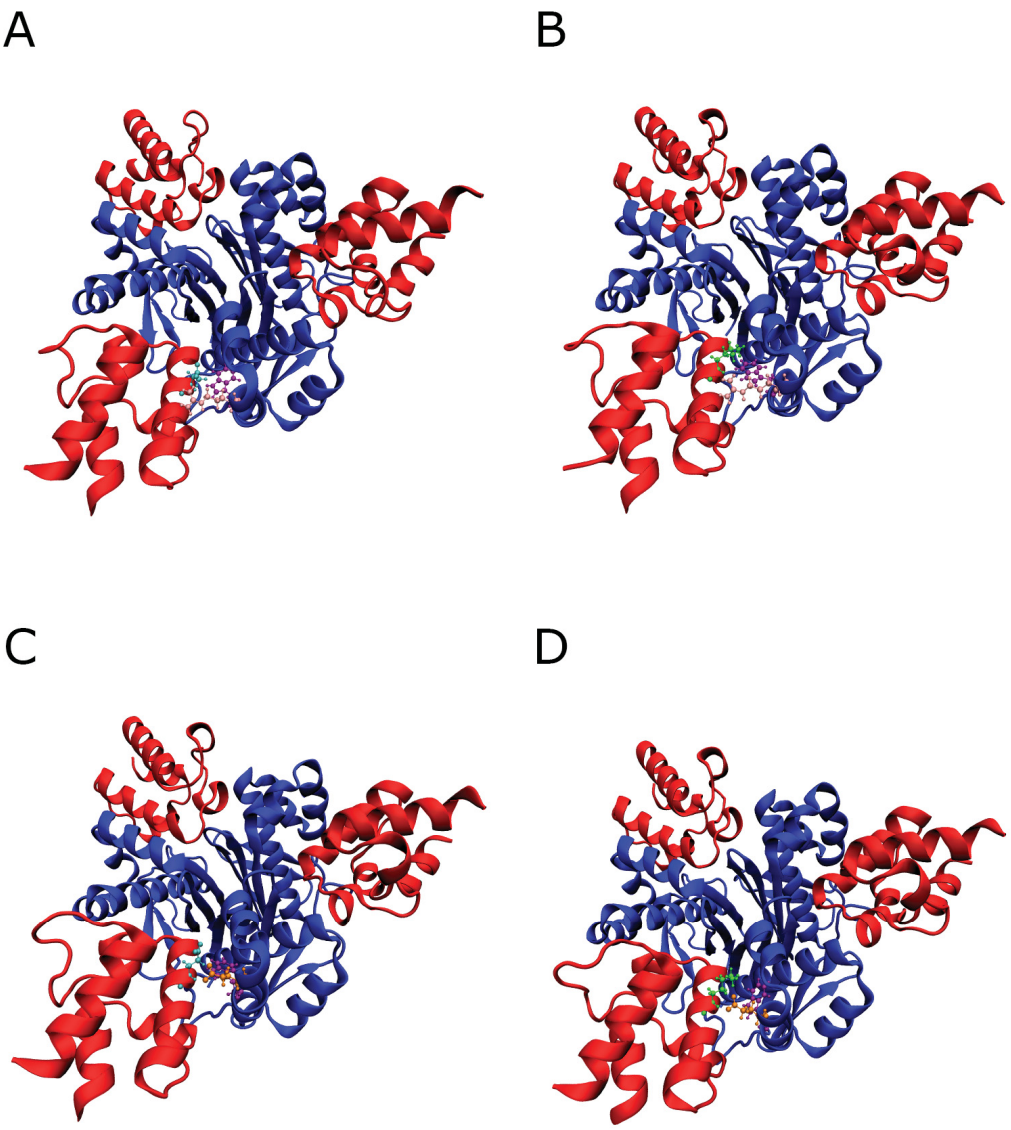

Figure 95 *B.s.*AcpS Trp-substituted charge-change model combinations post-MD simulation. "NewCartoon" representations of AcpS (blue)/ACP (red) combinations: A) F25W AcpS/wild-type ACP, B) F25W AcpS/E41K ACP, C) R21E- F25W AcpS/wild-type ACP, and D) R21E- F25W AcpS/E41K ACP. Starting structure for *in silico* mutation is wild-type AcpS bound by ACP (PDB ID: 1F80). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using "CPK" graphical representation: Trp 25 (purple), Glu 41 (cyan), Lys 41 (green), Arg 21 (pink), Glu 21 (orange). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

Figure 96).

Despite the fact that an equilibrium structure was not achieved during these short simulations, looking at the distance between key residues involved in ACP-AcpS binding can give indications of the effect of Trp-substitution as well as charge-change mutations. Since an aromatic residue not directly involved in AcpS-ACP binding was chosen for Trp-substitution, it isn't expected that this replacement would cause much in the way of local changes. This appears to be the case as distances between key residues involved in binding are not altered drastically (Figure 97; 0.146 A RMSD across 2 simulation end structures). Similarly, distances between key residues in both charge-change mutants (Figure 98; 0.135 A RMSD across 2 simulation end structures) and Trp-containing charge-change mutants (Figure 99; 0.129 A RMSD across 2 simulation end structures). were fairly constant across all simulations, indicating that, at lteast over the length of these simulations, Trp and charge-change substitutions do not have a large affect on local protein structure.

5.3.2.2 Homology Modeling of *P.f.*AcpS

As noted in the Introduction, the genome of the malaria parasite *Plasmodium falciparum* encodes a single bacterial type II AcpS that is localized to the apicoplast (Gardner et al., 2002); this unusual enzyme consists of two regions: a larger N-terminal domain (AcpS-N) of unknown function and a smaller C-terminal domain (AcpS-C) that is 30 – 50% identical to bacterial AcpS enzymes. Our laboratory has recently cloned and expressed the *P.f.*AcpS-C domain and demonstrated its phosphopantheinyltransferase activity (Tami et al., 2011). To complement ongoing structure-function analysis of this enzyme, its structure was predicted by homology modeling and molecular dynamics.

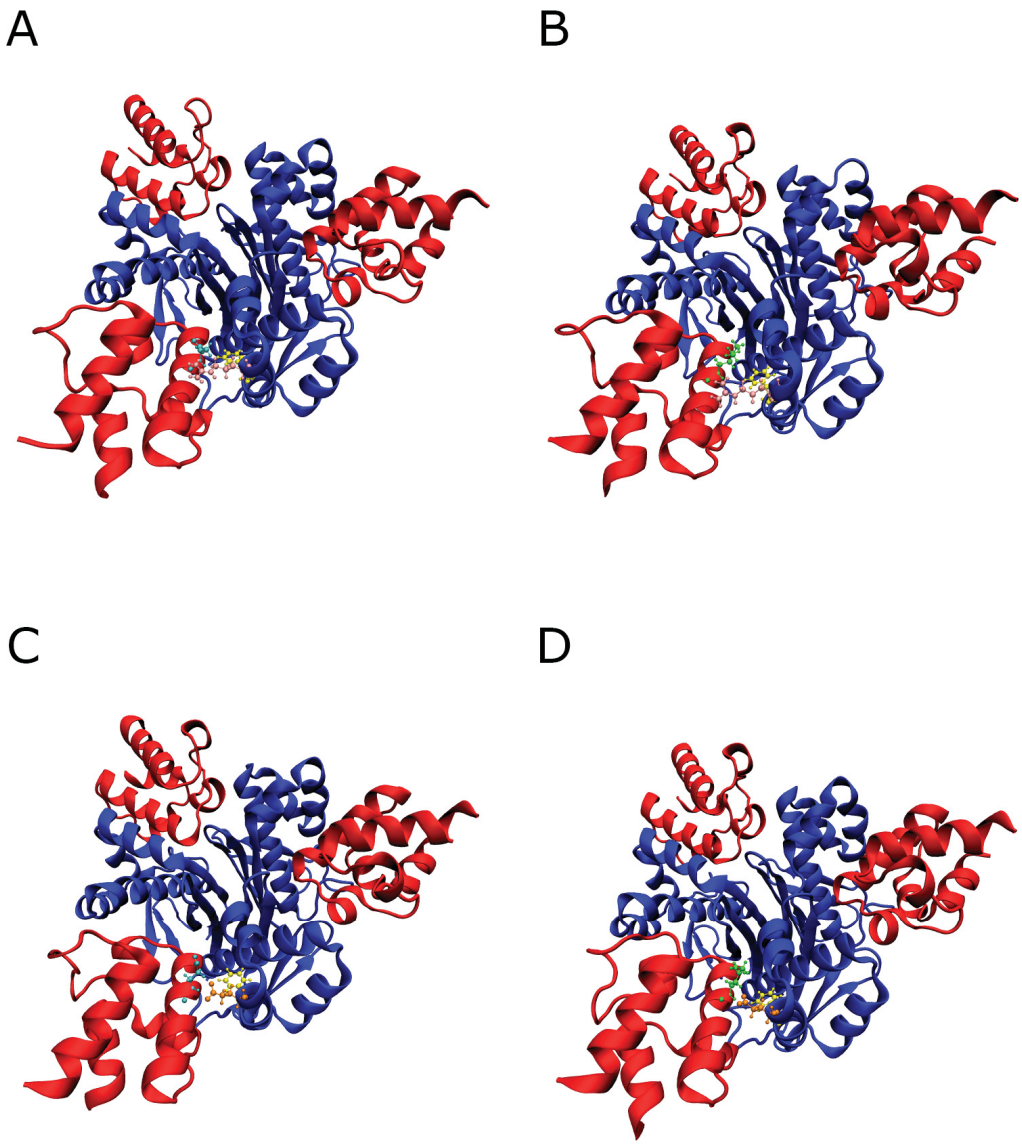

Figure 96 *B.s.*AcpS charge-change model combinations post-MD simulation. "NewCartoon" representations of AcpS (blue)/ACP (red) combinations: A) wild-type AcpS/wild-type ACP, B) wild-type AcpS/E41K ACP, C) R21E AcpS/wild-type ACP, and D) R21E AcpS/E41K ACP. Starting structure for *in silico* mutation is wild-type AcpS bound by ACP (PDB ID: 1F80; A). Equilibrated structures displayed here are the average of the last 40 frames of the simulation (~20 ps). Selected residues are highlighed using "CPK" graphical representation: Phe 25 (yellow), Glu 41 (cyan), Lys 41 (green), Arg 21 (pink), Glu 21 (orange). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

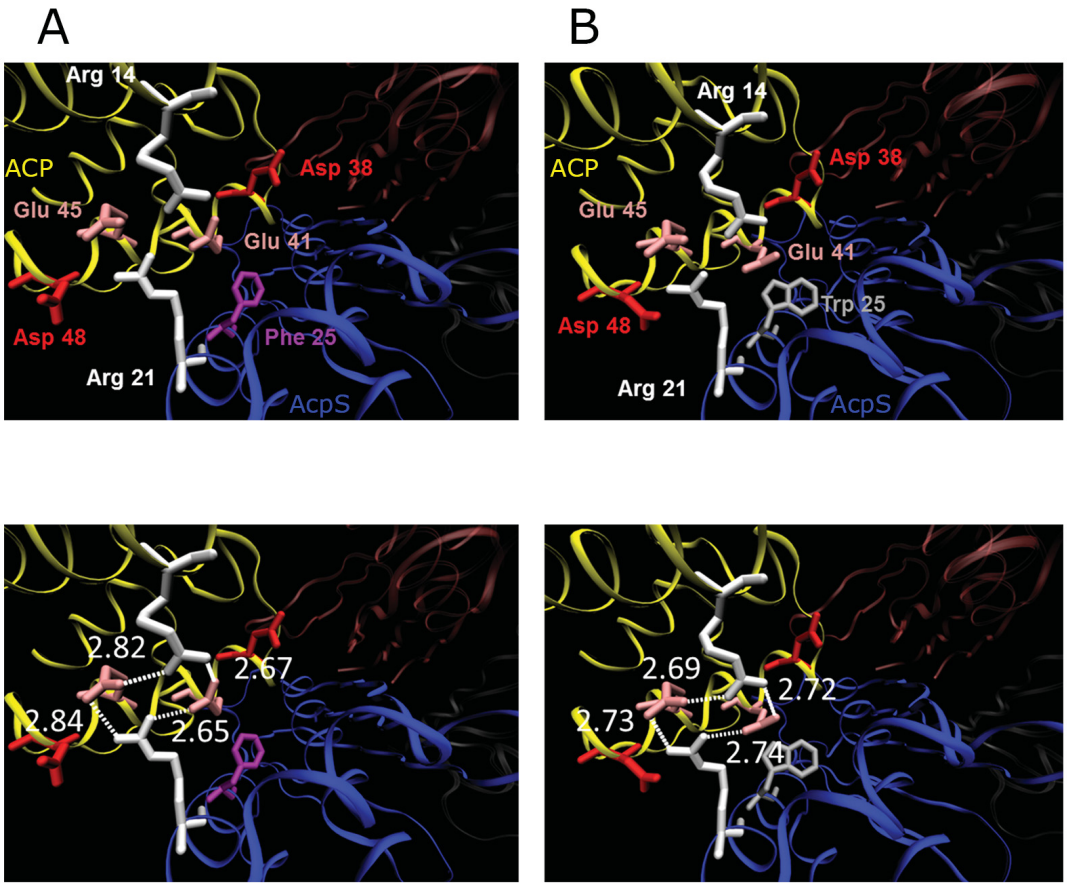

Figure 97 Distances between key residues for wild-type AcpS/wild-type ACP *versus* F25W AcpS/wild-type ACP. "NewRibbon" representations of wild-type AcpS/wild-type ACP (A) and F25W AcpS/wild-type ACP (B). Residues of interest are labeled (top) and distances displayed (bottom). Key residues are displayed in "Licorice" graphical representation: Glu (pink), Asp (red), Arg (white), Phe (purple) and Trp (grey). Distances were measured in VMD as described in Appendix 21. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD and rendered using the "Snapshot" rendering engine in VMD (see Appendix 24).

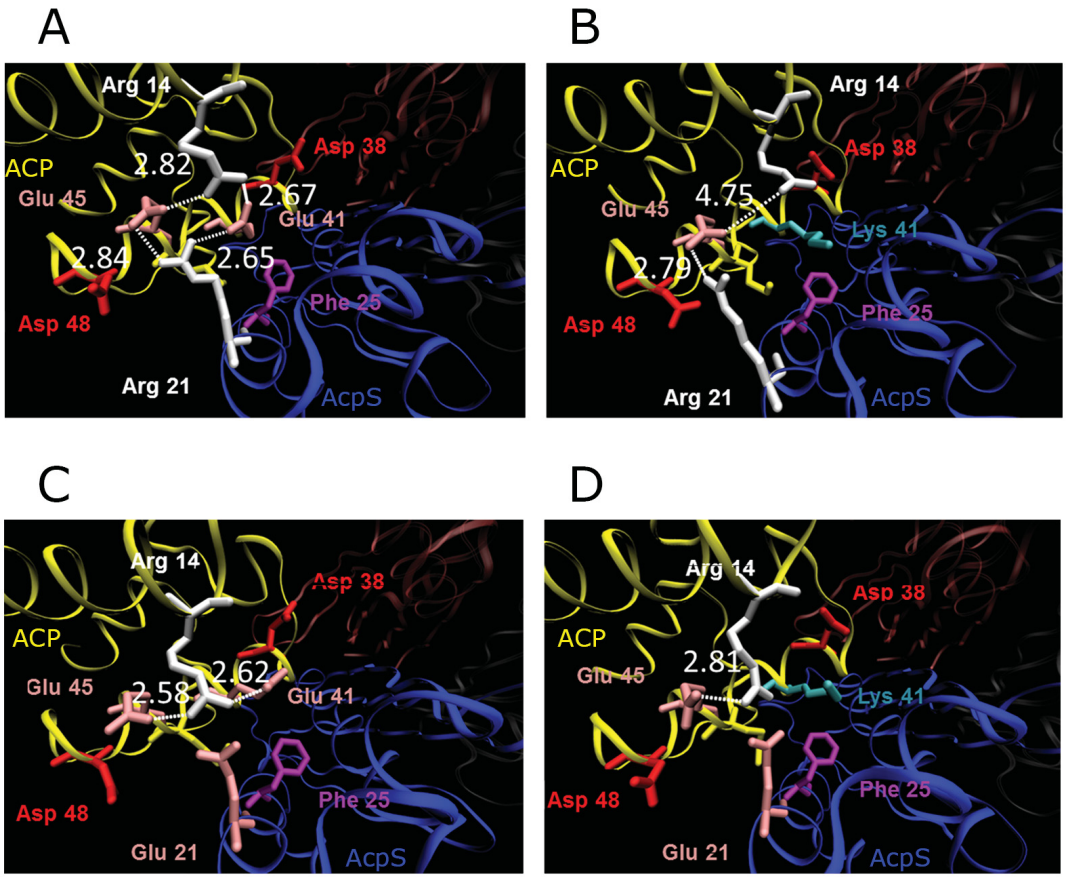

Figure 98 Distances between key residues for charge-change mutants. "NewRibbon" representations of AcpS (blue)/ACP (yellow) combinations: wild-type AcpS/wild-type ACP (A), wild-type AcpS/E41K ACP (B), R21E AcpS/wild-type ACP (C), and R21E AcpS/E41K ACP (D). Residues of interest are labeled (top) and distances displayed (bottom). Key residues are displayed in "Licorice" graphical representation: Glu (pink), Asp (red), Arg (white), Phe (purple) and Trp (grey). Distances were measured in VMD as described in Appendix 21. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD and rendered using the "Snapshot" rendering engine in VMD (see Appendix 24).

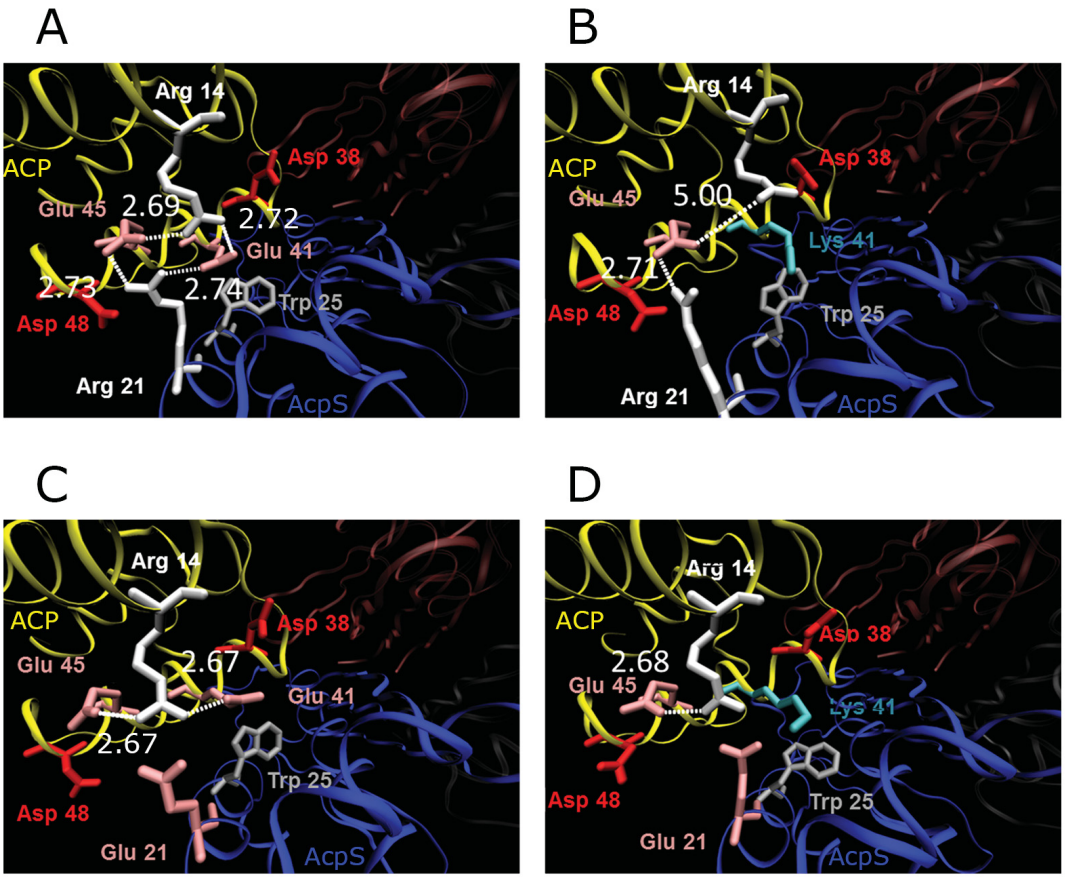

Figure 99 Distances between key residues for Trp-containing charge-change mutants. "NewRibbon" representations of AcpS (blue)/ACP (yellow) combinations: F25W AcpS/wild-type ACP (A), F25W AcpS/E41K ACP (B), R21E- F25W AcpS/wild-type ACP (C), and R21E- F25W AcpS/E41K ACP (D). Residues of interest are labeled (top) and distances displayed (bottom). Key residues are displayed in "Licorice" graphical representation: Glu (pink), Asp (red), Arg (white), Phe (purple) and Trp (grey). Distances were measured in VMD as described in Appendix 21. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD and rendered using the "Snapshot" rendering engine in VMD (see Appendix 24).

Three crystal structures were used as templates to create homology modeled structures of *P.f.*AcpS-C: AcpS-C from *Plasmodium yoelii* without (PDB ID: 2BDD; 66% sequence identity to *P.f.*AcpS-C) and with (2QG8; 66% sequence identity to *P.f.*AcpS-C) bound ADP, and the *B. subtilis* AcpS described previously (1F7T; 30% sequence identity to *P.f.*AcpS-C). Upon completion of modeling of *P.f.*AcpS using SWISS-MODEL, these structures were aligned using Chimera (Appendix 1) so they could be compared (Figure 100). All three modeled *P.f.*AcpSs are nearly identical (RMSD = 0.009) in structure (Figure 101) so only the structure based on *P. yeolii* AcpS (2BDD) was chosen for MD equilibration. Prior to MD simulation, the protein monomer was solvated as described previously (Figure 102A, B). Following the short simulation, the end structure was saved (Figure 102C, D) and the monomer was trimerized (Figure 103). As with previous simulations, the RMSD was calculated over the course of the simulation to determine if the structure had reached equilibrium (Figure 104). Based on the graph, equilibrium had not been reached and a longer simulation would be necessary to use the modeled structure for other simulations or in-depth analysis. Furthermore, it would be best to minimize and equilibrate the biologically active trimer, rather than the monomer of *P.f.*AcpS. Evidence of this can be seen in the final trimerized structure (Figure 103). Examining the inset of the figure it is clear that the termini of each monomer overlap slightly. Moreover, a partial loss of β–sheet secondary structure occurs for a handful of residues (Figure 102D; red arrow). These effects would likely not have occurred if the trimer had been simulated instead of the monomer.

The *P.f.*AcpS-N domain was modeled based on two starting structures. The first was produced using *E.c.*AcpH as a starting structure (Figure 105D; 32% sequence

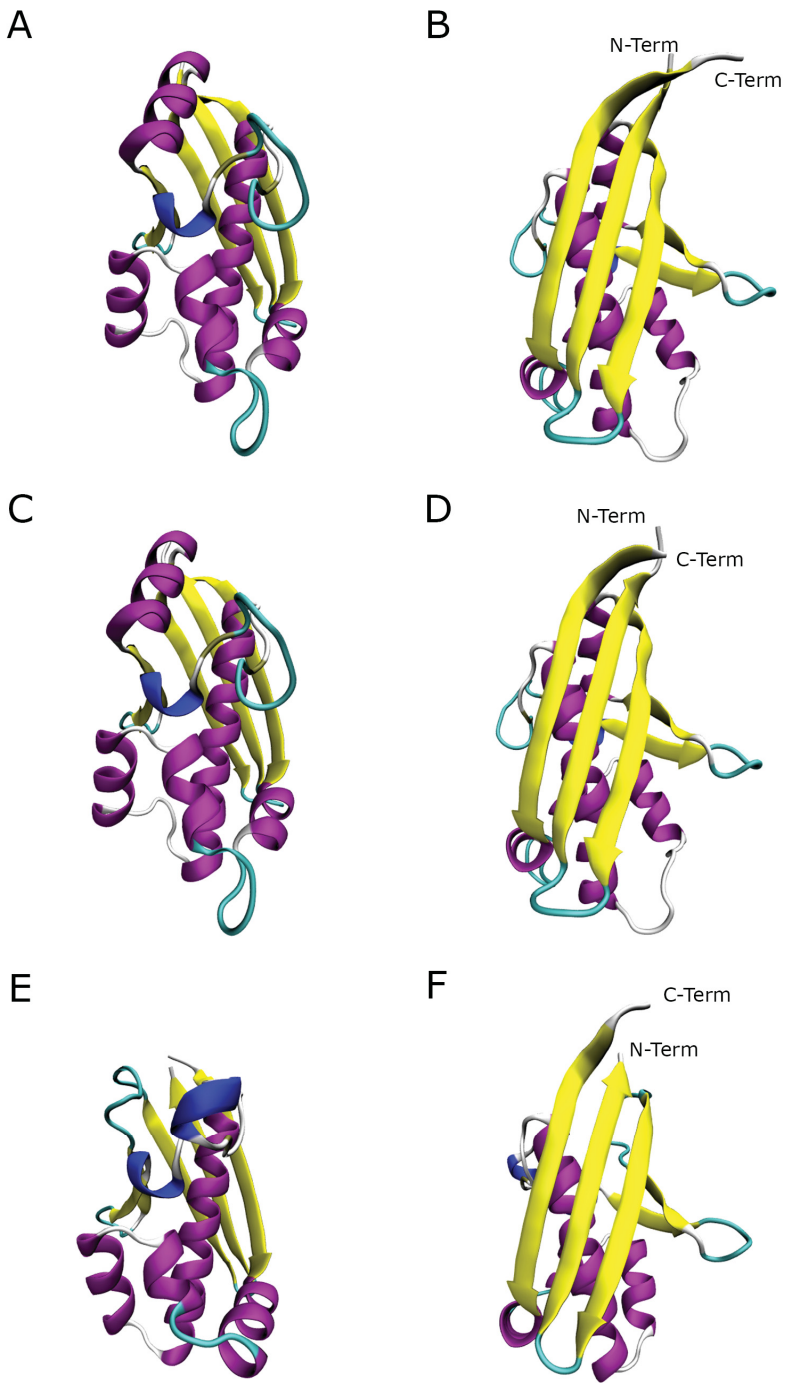

Figure 100 *P.f.*AcpS homology models created *in silico* using SWISS-MODEL. Front (A, C, E) and back (B, D, F) views of homology models of *P.f.*AcpS using different starting structures: PDB ID 2BDD (A, B), PDB ID 2QG8 (C, D), and PDB ID 1F7T (E, F). Models are displayed using "NewCartoon" graphical representation and are colored by secondary structure. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

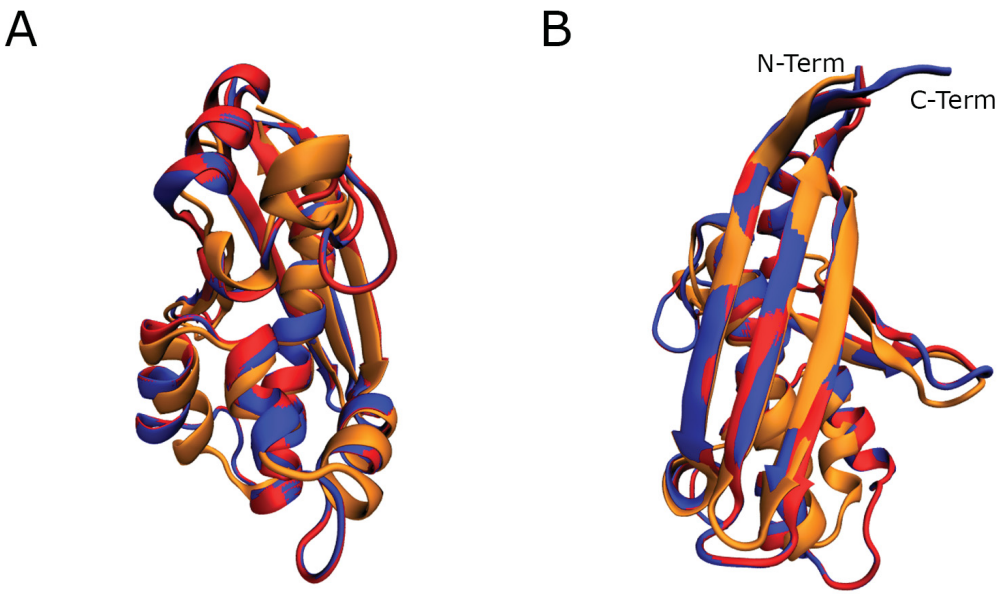

Figure 101 *P.f.*AcpS homology models aligned and overlaid. Front (A) and back (B) views of homology models of *P.f.*AcpS using 2BDD (Blue), 2QG8 (Red), and 1F7T (Orange) as starting structures. Models are displayed using "NewCartoon" graphical representation and are colored by secondary structure. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

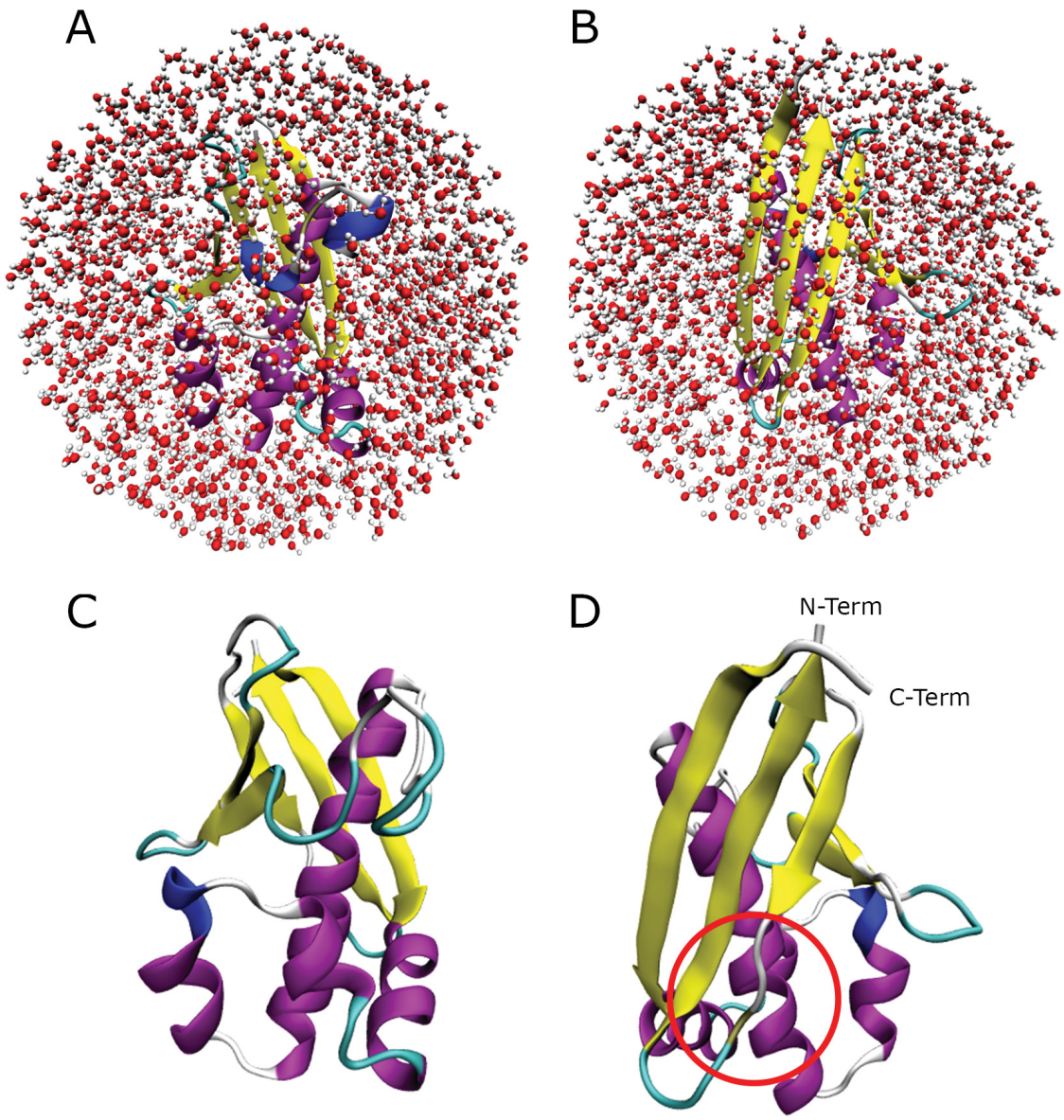

Figure 102 *P.f.*AcpS homology model (based on PDB ID 2BDD) Pre- and Post-Simulation. Front (A, C) and back (B, D) views of *P.f.*AcpS are shown for solvated structure, pre-simulation (A, B) and the structure post-simulation (C, D). Red circle indicates region where β–sheet structure is lost during simulation. Models are displayed using "NewCartoon" graphical representation and are colored by secondary structure. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

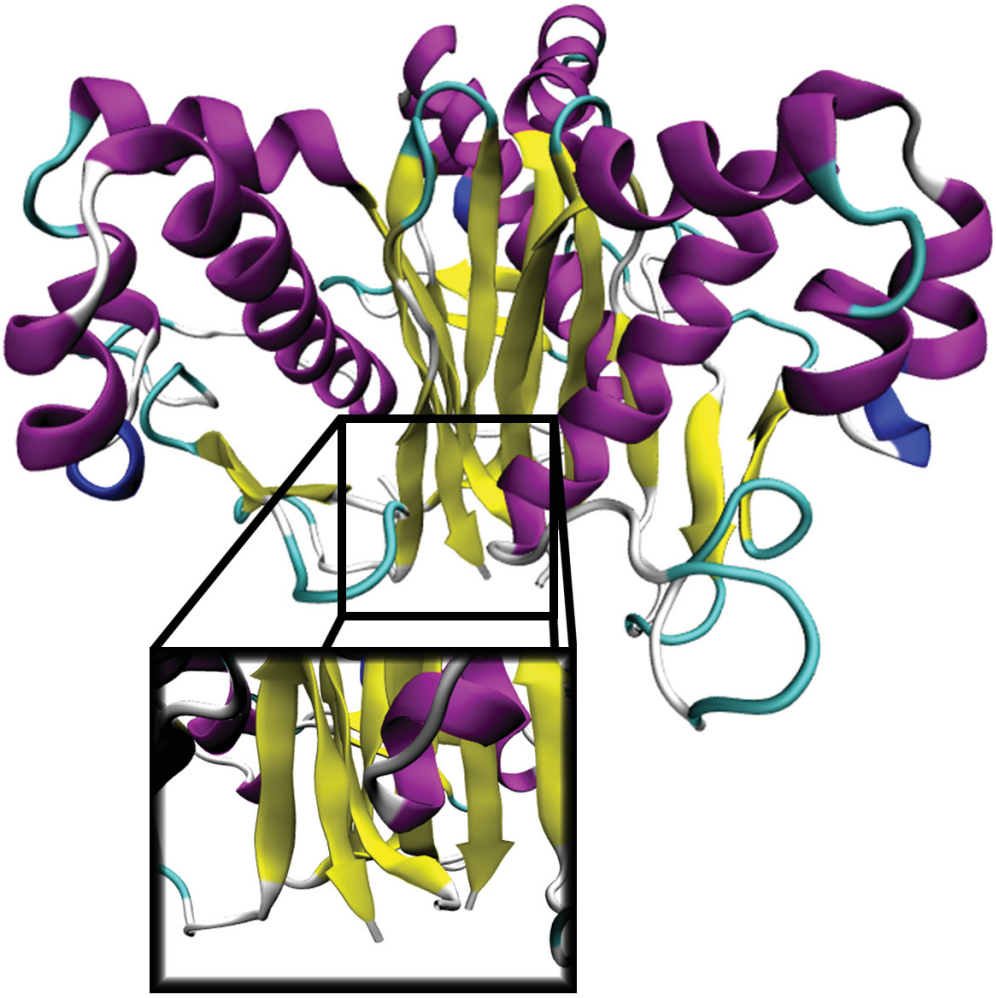

Figure 103 Simulation end structure for homology modeled *P.f.*AcpS depicted in its assumed biologically active trimeric form. Inset shows closer view of subunit interface centered on the termini. End structures displayed here are the atomic positions in the final frame of the simulation (see Appendix 11). Model is displayed using "NewCartoon" graphical representation and is colored by secondary structure. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

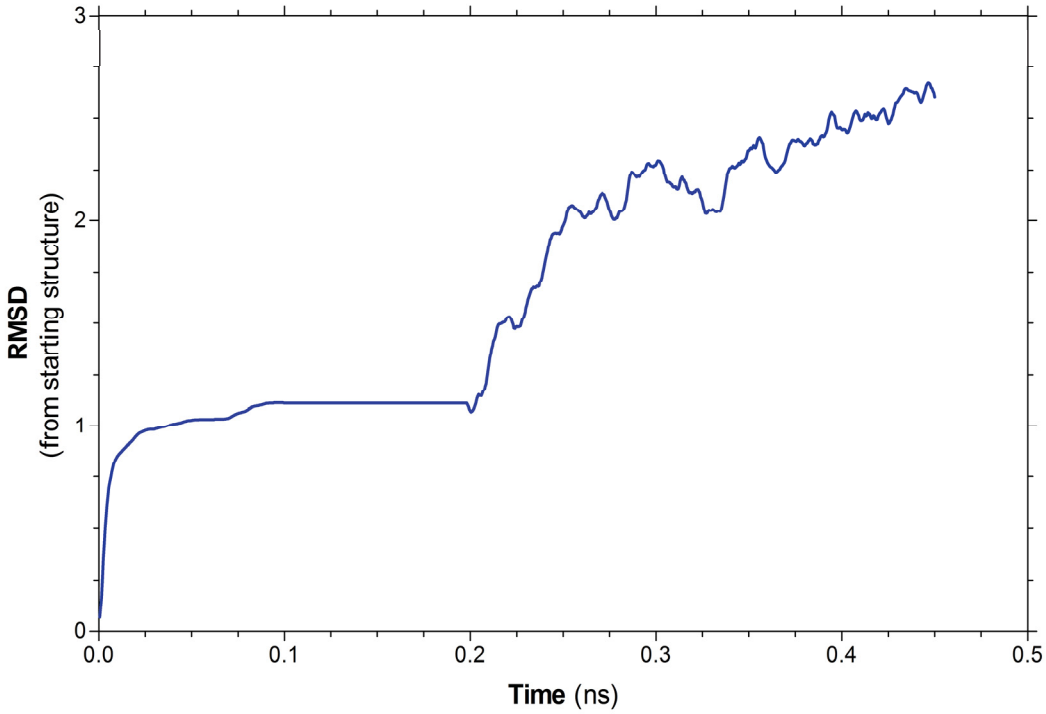

Figure 104 RMSD for backbone atoms of homology modeled *P.f.*AcpS. RMSD was calculated from the starting structure. Values were calculated as described in Appendix 19 using the smoothed trajectory file (sliding window of 9; Appendix 20).

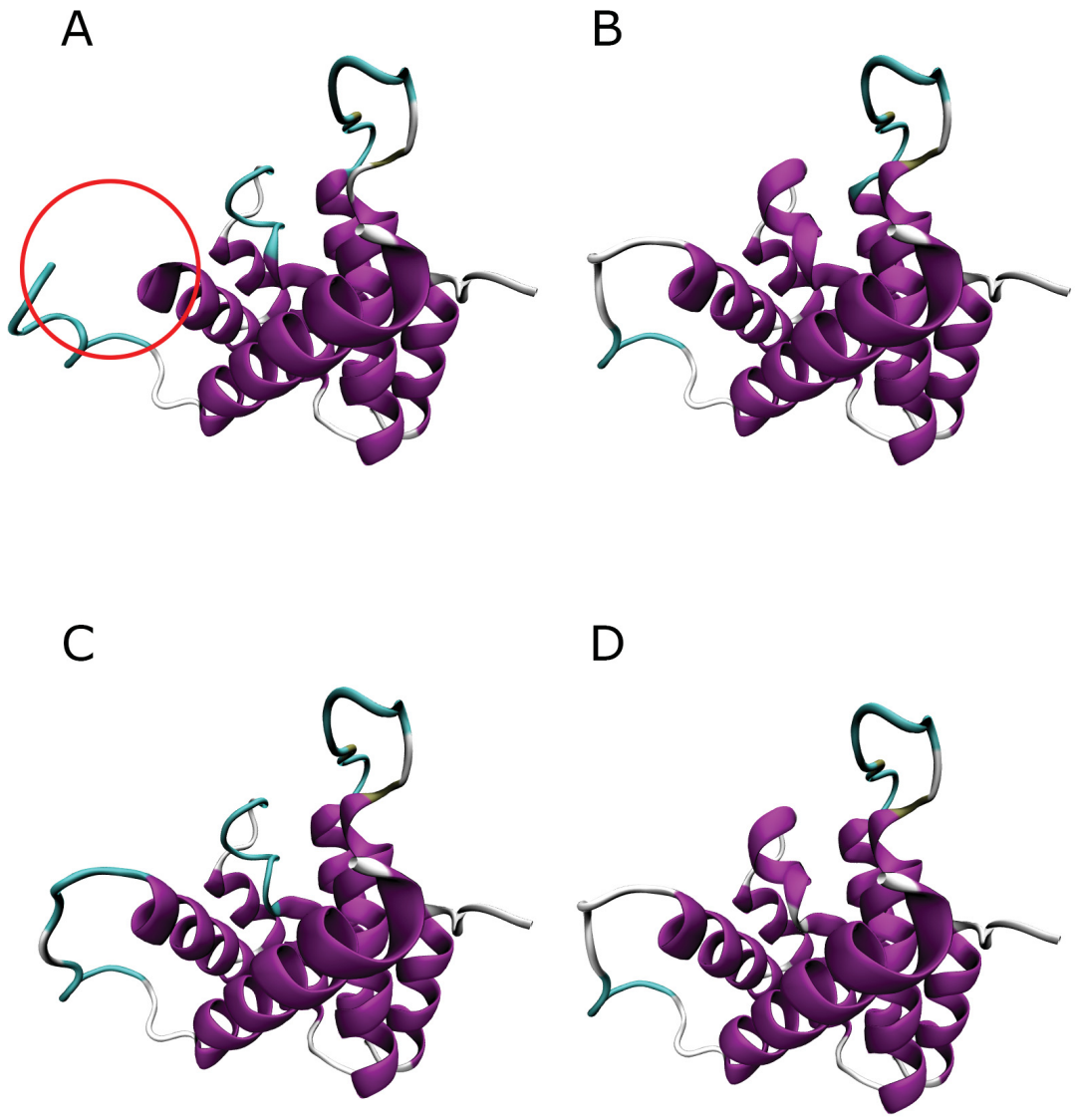

Figure 105 Template and homology modeled structures for *P.f.*AcpS-N. SpoT's hydrolase domain (A) was used as a template to homology model *E.c.*AcpH (B). Both SpoT and *E.c.*AcpH were then used as templates to homology model *P.f.*AcpS-N (C, D, respectively). Models are displayed using "NewCartoon" graphical representation and are colored by secondary structure. The circle in A shows the area of missing amino acids in the starting SpoT structure. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

identity between *P.f.*AcpS-N and *E.c.*AcpH); since there isn't a published structure of *E.c.*AcpH, it in turn is modeled on SpoT (Figure 105B; 29% sequence identity between *E.c.*AcpH and SpoT) based on previous work (Thomas et al., 2007). The second structure was homology modeled directly using SpoT (Figure 105A; 31% sequence identity between *P.f.*AcpS-N and SpoT) as a starting structure (Figure 105C). In the SpoT starting structure, a short sequence of amino acids are missing (Figure 105A; black circle); of note is that this did not affect the resultant homology models. Threading through this region was not an issue. Comparing both *P.f.*AcpS-N structures to SpoT and the homology modeled *E.c.*AcpH it is obvious that all structures are essentially identical (Figure 106A; 0.121 Å RMSD across structures; lower RMSD vslues obtained for individual comparision (see Table 16)). One caveat about homology modeling is that any sequence can be threaded through any structure and produce what appears to be a potential structure (0.100 Å RMSD) as demonstrated by homology modeling BSA (~20% sequence identity to *P.f.*AcpS-C) using SpoT as a template (Figure 106B). If homology modeling is done as here (with DeepView and SWISS-MODEL), this would involve using the "Find Appropriate ExPDB template" search function in Swiss-PdbViewer: DeepView (more information available in the DeepView Tutorial on Homology Modeling (Swiss Institute for Bioinformatics. *DeepView Tutorial: Homology Modelling*.)).

5.3.2.3 Homology Modeling of *V.f.*AcpS

The crystal structure of the *V.f.*AcpS enzyme used as the experimental system for Trp replacement has not been determined, thus SWISS-MODEL was employed to produce a homology modeled structure of *V.f.*AcpS (Figure 107C, D). This model was

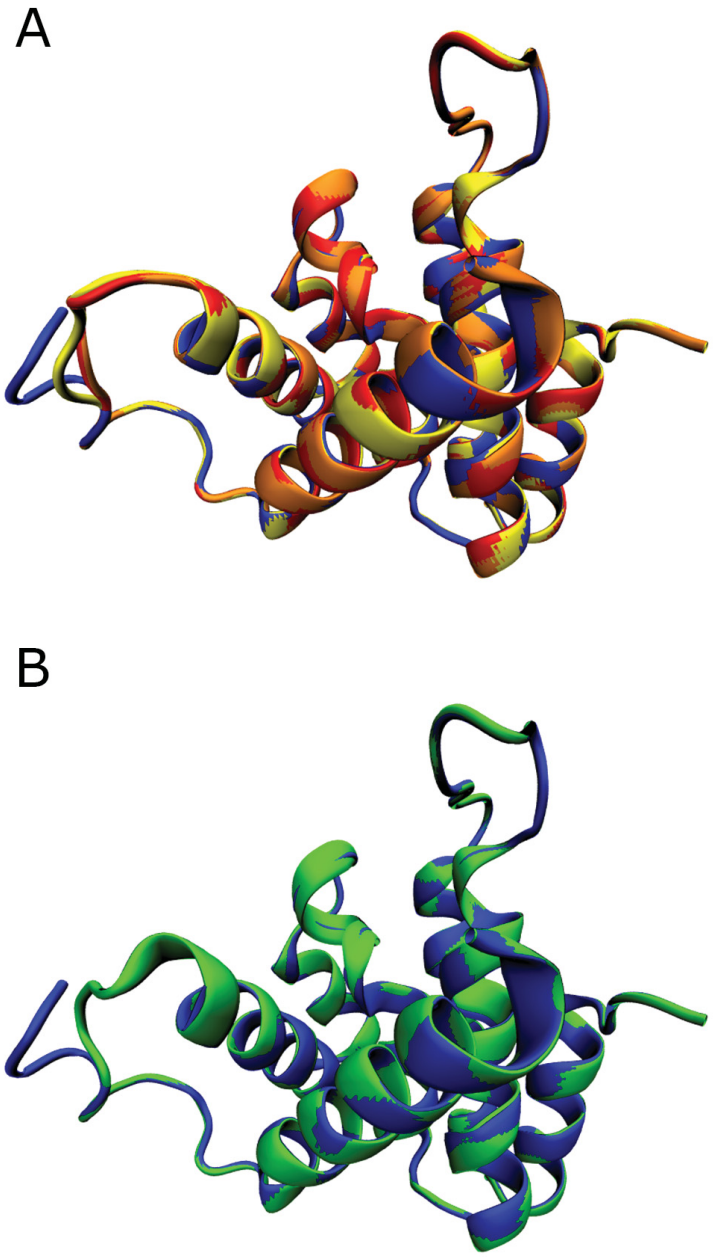

Figure 106 Overlaid template and homology modeled structures for homology modeling of *P.f.*AcpS-N. A, SpoT's hydrolase domain (blue), *E.c.*AcpH (red), *P.f.*AcpS-N (SpoT as template; yellow), and *P.f.*AcpS-N (*E.c.*AcpH as template; orange) are overlaid for easy structural comparison. B, SpoT's hydrolase domain (blue), and BSA (green) are overlaid. Models are displayed using "NewCartoon" graphical representation and are colored by model. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

Table 16 RMSD values for homology models comared to the template structure.

Homology Modeled Protein	Template	RMSD from Template
*E.c.*AcpH	SpoT	0.080 Å
*P.f.*AcpS-N	*E.c.*AcpH	0.067 Å
*P.f.*AcpS-N	SpoT	0.084 Å

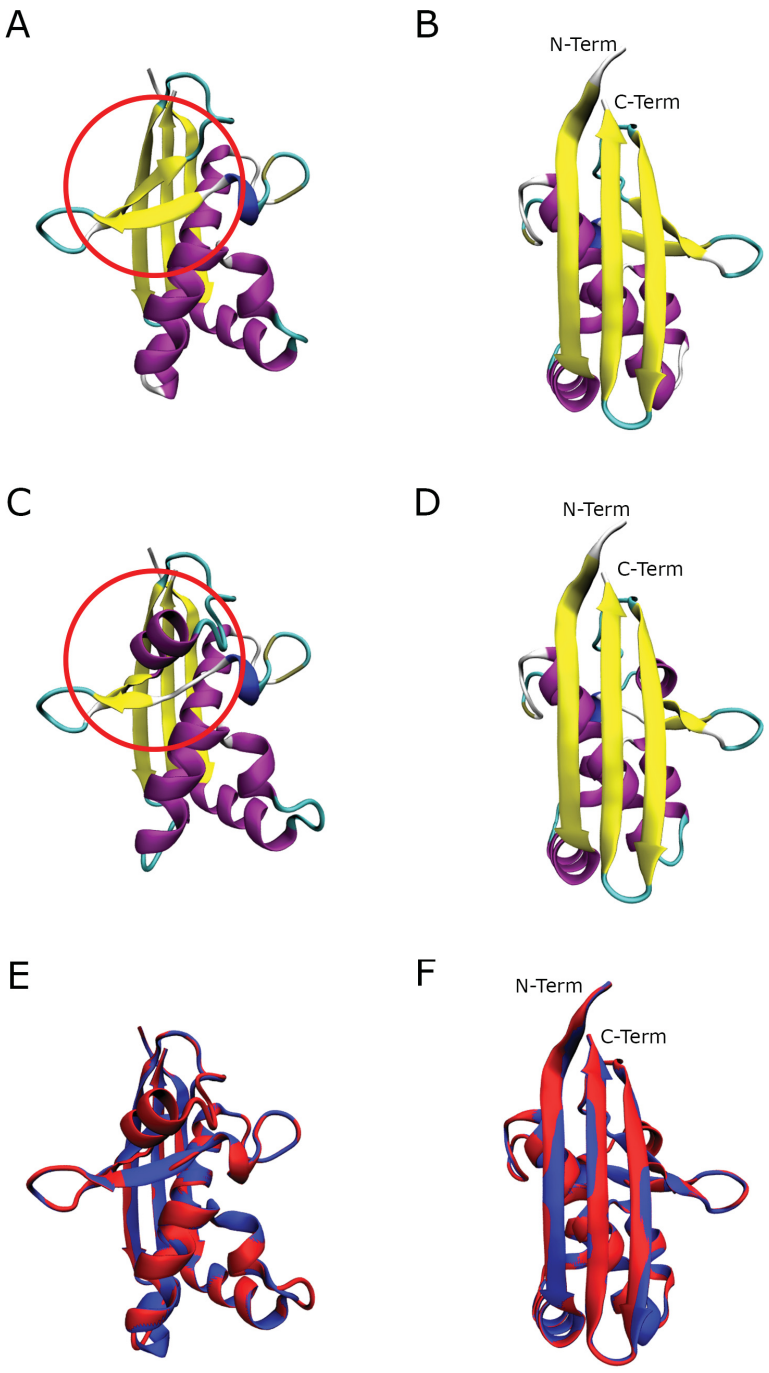

Figure 107 Template and homology modeled structures for *V.f.*AcpS. *B.s.*AcpS (A, B) was used as a template to homology model *V.f.*AcpS (C, D). Models are displayed using "NewCartoon" graphical representation and are colored by secondary structure. The circle in A/C highlights the area of differing secondary structure between the template and homology modeled *V.f.*AcpS. Front (E) and back (F) view of overlaid *B.s.*AcpS (blue) and *V.f.*AcpS (red) structures demonstrates the differences in secondary structure highlighted in A and C. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

not identical to that of its template (Figure 107A, C; red circle): a β-sheet in the template is replaced by an α-helix in the homology modeled structure. Assuming this is the proper secondary structure for this region of *V.f.*AcpS, this change would not be expected to affect trimerization as it is not near the subunit interface (Figure 108; orange selection). Additionally, due to its placement below the ACP binding site, it likely wouldn't have an effect on ACP binding (Figure 108). Despite this difference, the structures are highly similar, with a RMSD of 0.076 Å.

5.4 Discussion

LpxA catalyzes the first step of lipid A synthesis (Raetz, 1986, Anderson & Raetz, 1987, Coleman & Raetz, 1988). Much is known about which residues are involved in binding of ACP and UDP-Glc*N*Ac (Wyckoff & Raetz, 1999, Ulaganathan et al., 2007, Williams & Raetz, 2007) and even residues important for acyl chain specificity (Wyckoff et al., 1998a). However, very little is known about which ACP residues might be important for triggering ACP unfolding and transfer of the acyl chain to UDP-Glc*N*Ac in the LpxA active site. To examine this, a probe is needed to independently measure binding of substrates to LpxA. Tryptophan has been previously validated in our lab as an excellent fluorescent probe of ACP structure and since *E.c.*LpxA lacks endogenous Trp, it can be used as a probe to monitor binding. Four positions were chosen for Trp substitution: Q104, F162, G173, and S208 (Figure 74). Most Trp-substituted LpxA exhibited similar circular dichroism (CD) spectra to wild-type LpxA and all co-eluted with wild-type enzyme when injected on a gel filtration column. Taken together, this suggests that Trp-substitution into these positions does not affect global protein structure and therefore multimer formation. Furthermore, molecular dynamics (MD) simulations suggest that

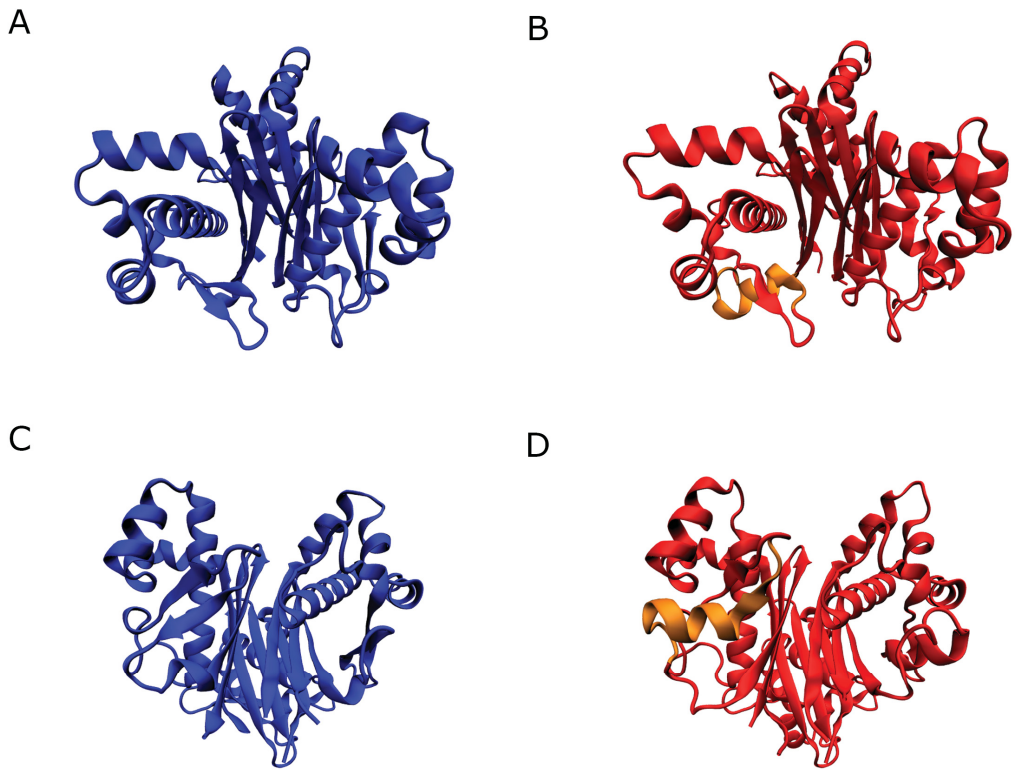

Figure 108 Trimer formation of *B.s.*AcpS and *V.f.*AcpS. *B.s.*AcpS (A, C) trimer (blue) compared to *V.f.*AcpS (B, D) trimer (red). Both top (A, B) and bottom (C, D) views are displayed. The region of secondary structure change between *B.s.*AcpS and *V.f.*AcpS is highlighted on *V.f.*AcpS (orange). *B.s.*AcpS and *V.f.*AcpS share 44% identity. Models are displayed using "NewCartoon" graphical representation. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

there are no major local arrangements to protein secondary or tertiary structure on the nanosecond timescale. Since trimer formation was not affected by Trp-substitution at these positions, any changes in LpxA activity would likely result from the substitution affecting binding and/or catalysis. Previous work from the Raetz and Hunter labs (Wyckoff & Raetz, 1999, Ulaganathan et al., 2007, Williams & Raetz, 2007) identified two residues important for catalysis (His 125 and Arg 126) and several others involved in binding of ACP and UDP-Glc*N*Ac (Asp 74, Lys 76, His 122, His 125, His 144, His 160, Gln 161, Asn 198, Glu 200, Arg 204, Arg 205).

Looking at the position of Phe 162 relative to bound ACP, the acyl chain and key residues on LpxA could give clues to why F162W exhibits higher than wild-type activity (Figure 109). Due to the location of F162, I postulate that the increased activity could be due to possible increased hydrophobic interactions between F162W and the acyl chain as it is passed from ACP to LpxA's FA binding site (Figure 109; silver MSMS surface). This could be tested by comparing the affinity of F162W with holo- *versus* acyl-ACP using fluorescence. Work from Chapter 3 (cyclic ACP) suggests that the short helix (III) between Helix II and IV is likely mobile enough to expose the fatty acyl chain such that it can pass from ACP to partner enzymes (in this case LpxA). This region (Figure 110; orange region) can rotate slightly (Figure 110; orange arrow), which would expose the fatty acid for extraction from ACP. Taken together, this suggests that for the acyl chain to be passed from ACP to LpxA, this small helix moves, the acyl chain is exposed and hydrophobic residues in the area are used to draw the fatty acyl chain from the inner core of ACP. While the acyl chain is being removed, it could be then drawn toward the hydrophobic binding pocket of LpxA. This would position the S-C bond of acyl-ACP in

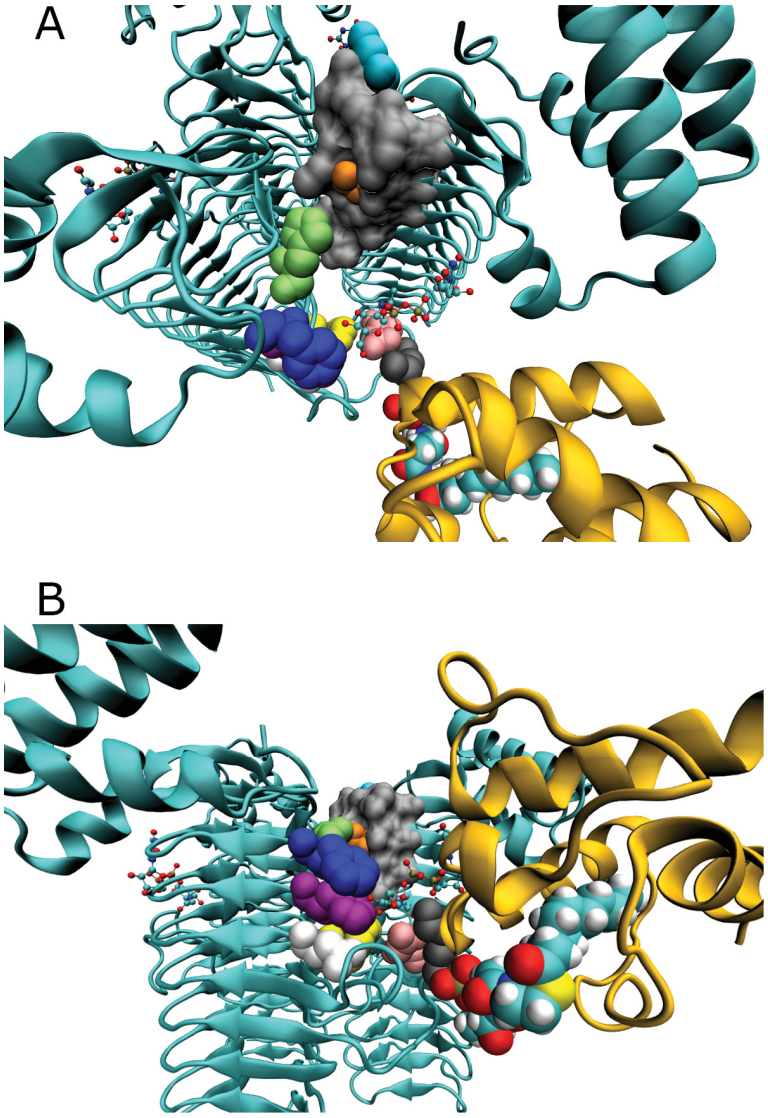

Figure 109 Position of F162 relative to residues involved in binding and catalysis for *E.c.*LpxA. Top (A) and side (B) views of LpxA bound to ACP and UDP-Glc*N*Ac showing position of F162 (blue) relative to important LpxA residues: His 125 (yellow), Asp 126 (white), His 144 (purple), Lys 76 (grey), Asp 74 (pink), His 160 (lime green), His 191 (light turquoise), Gly 173 (orange) and the fatty acid binding cleft (silver). LpxA (cyan) and ACP (gold) are depicted in the "NewCartoon" representation, UDP-Glc*N*Ac (colored by atom) is depicted using "CPK" and decanoyl-phosphopantetheine (colored by atom) is depicted using "VDW" graphical representation. All selected residues are displayed using the "VDW" representation while the fatty acid binding cleft is represented using a MSMS surface. Three structures were used to create this figure: LpxA bound to UDP-Glc*N*Ac (PDB ID = 2JF3 (Ulaganathan et al., 2007)), Spinach decanoyl-ACP (PDB ID = 2FVF (Zornetzer et al., 2006)) and ACP docked to LpxA (created and provided by Dr. Christopher Barden (DeNovaMed Inc.)). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1 and Appendix 12). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

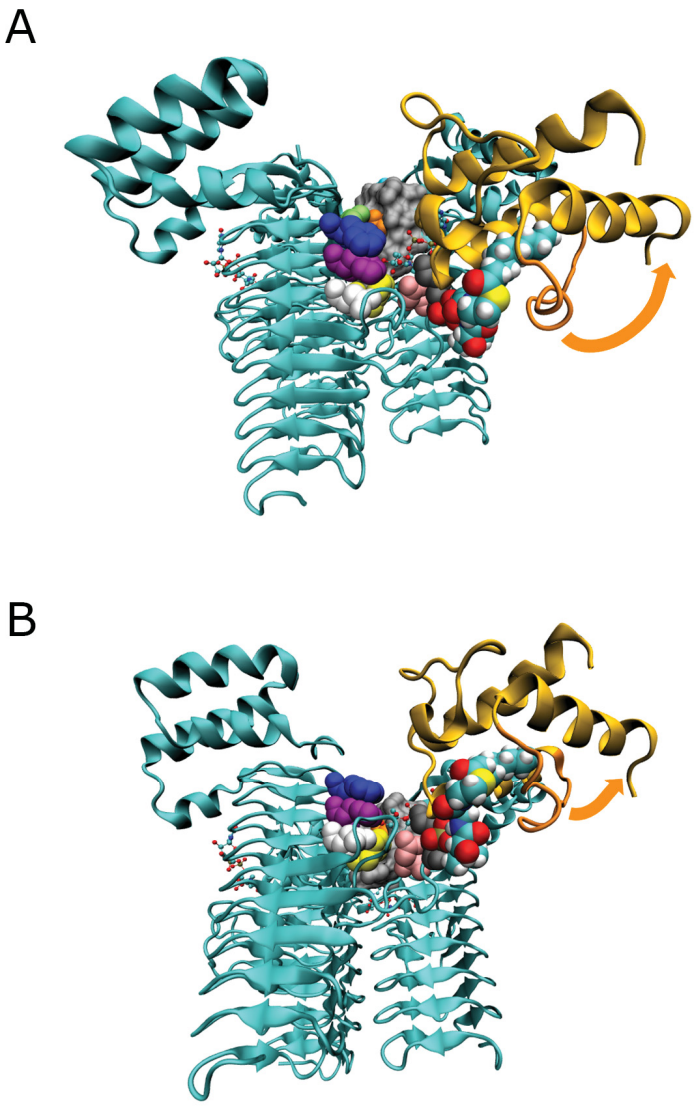

Figure 110 Position of F162 relative to fatty acid and the potential Helix III "flap" on ACP. Top (A) and side (B) views of LpxA bound to ACP and UDP-Glc*N*Ac showing position of F162 (blue) relative to important LpxA residues: His 125 (yellow), Asp 126 (white), His 144 (purple), Lys 76 (grey), Asp 74 (pink), His 160 (lime green), His 191 (light turquoise) and the fatty acid binding cleft (silver). LpxA (cyan) and ACP (gold) are depicted in the "NewCartoon" representation, UDP-Glc*N*Ac (colored by atom) is depicted using "CPK" and decanoyl-phosphopantetheine (colored by atom) is depicted using "VDW" graphical representation. All selected residues are displayed using the "VDW" representation while the fatty acid binding cleft is represented using a MSMS surface. The potential flap (orange) on ACP whose movement (orange arrow) would allow for the transfer of the fatty acid from ACP to the binding cleft (silver) on LpxA. The three structures used to create this figure are described in the legend to Figure 109. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1 and Appendix 12). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

proximity to both His 125 and the 3-O of GlcNAc, allowing for the general base catalyzed transfer of the acyl chain from ACP to UDP-Glc*N*Ac (Wyckoff & Raetz, 1999, Ulaganathan et al., 2007, Williams & Raetz, 2007).

Q104W has somewhat decreased activity relative to wild-type *E.c.*LpxA. As it has been established that this isn't due to global changes in active site formation through trimerization (above), Trp-substitution at this location must perturb binding of the substrates or sterically interfere with catalysis. Structurally, Q104W is positioned at the binding interface of LpxA and ACP, adjacent to Ser 36 of ACP (the phosphopantetheine-fatty acid attachment point (Byers & Gong, 2007)) (Figure 111; black). Inserting a large, rigid ring structure here in place of a long, mobile and polar side chain (glutamine) could easily hinder movement of the acyl chain from ACP to the hydrophobic fatty acid binding cleft on LpxA. Furthermore, this substitution could affect ACP binding. Due to its position, Q104W is unlikely to affect binding of UDP-Glc*N*Ac (Figure 111). Both of these effects could cause the decreased activity observed in this work.

The greatly diminished activity of G173W is not unexpected. Previous work (Raetz 1998, Williams 2007) has implicated Gly 173 as part of a "hydrocarbon ruler" that determines the acyl chain specificity of LpxA (His 191 denotes the end of the ruler in *E.c.*LpxA (Williams & Raetz, 2007)). When this residue is mutated to methionine, to mimic *P. aeruginosa* LpxA (*P.a.*LpxA), the acyl chain specificity of *E.c.*LpxA is switched from *R*-3-hydroxymyristoyl to *R*-3-hydroxydecanoyl, the preferred natural substrate of *P.a.*LpxA (Wyckoff et al., 1998a). Similarly, when the reciprocal mutation (M173G) is made to *P.a.*LpxA, *R*-3-hydroxymyristoyl is the preferred substrate. Thus, substituting a Trp (a residue that is even larger than Met) at this position would be

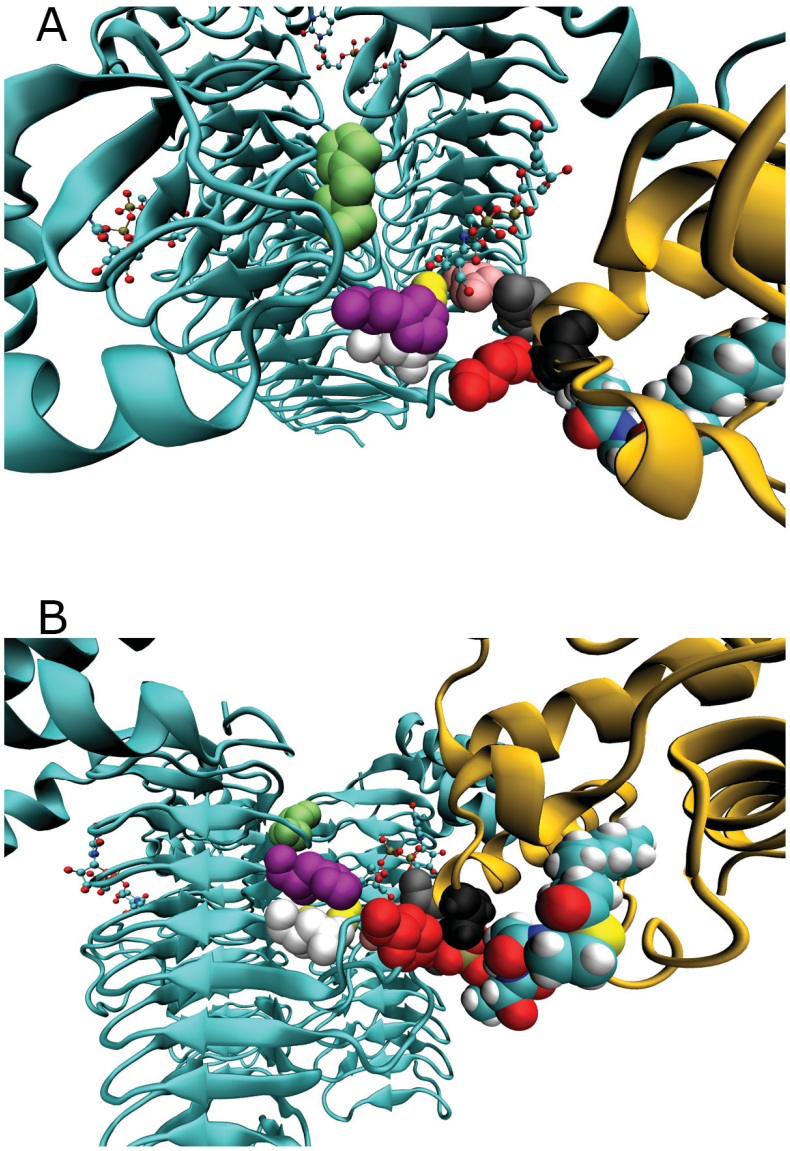

Figure 111 Position of Q104 relative to residues involved in binding and catalysis for *E.c.*LpxA. Top (A) and side (B) views of LpxA bound to ACP and UDP-Glc*N*Ac showing position of Q104 (red) relative to important LpxA residues: His 125 (yellow), Asp 126 (white), His 144 (purple), Lys 76 (grey), Asp 74 (pink), and His 160 (lime green). Additionally, Ser 36 of ACP is highlighted (black). LpxA (cyan) and ACP (gold) are depicted in the "NewCartoon" representation, UDP-Glc*N*Ac (colored by atom) is depicted using "CPK" and decanoyl-phosphopantetheine (colored by atom) is depicted using "VDW" graphical representation. All selected residues are displayed using the "VDW" representation while the fatty acid binding cleft is represented using a MSMS surface. The three structures used to create this figure are described in the legend to Figure 109. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1 and Appendix 12). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

expected to decrease the chain length that *E.c.*LpxA would accept. Observing the structure of LpxA bound to ACP (Figure 112) this postulated effect on acyl chain length specificity is apparent. The exact acyl chain length that G173W prefers was not investigated here, as 3-hydroxy fatty acids of shorter chain lengths are not commercially available. However, given that His 191 of *E.c.*LpxA is positioned near C14 of *R*-3-hydroxymyristoyl while G173 is positioned near C12 (Williams & Raetz, 2007) and considering the structure of LpxA (Figure 112), I posit that G173W would prefer an *R*-3-hydroxyoctanoyl or perhaps 3-hydroxyhexanoyl acyl chain.

At first glance, the almost complete lack of S208W activity might seem somewhat surprising. Structurally, Ser 208 isn't located near either the binding or catalysis sites (~34 Å between Ser 208 C_α and His 125 C_α; ~17 Å between Ser 208 C_α and the closest UDP-Glc*N*Ac atom; ~10 Å between Ser 208 C_α and Arg 205 C_α; these distances were calculated based on Figure 84A). Rather, it is located adjacent to the distal side of ACP (away from Ser 36, and the active site). However, previous work from our laboratory (Gong et al., 2007) has revealed that mutations in divalent cation binding site B of ACP preferentially inhibit LpxA activity. In the docked structure of ACP-LpxA, site B is far from the enzyme active site but is relatively close to Ser 208. As discussed above, the site B/Helix III region of ACP is implicated in structural changes required for binding and release of the acyl chain from ACP, and Ser 208 could be a critical LpxA residue in this process. Due to its positioning (Figure 113), it is possible that substitution of Ser (a relatively small amino acid) for Trp (large and rigid) could affect the ability of the distal side of ACP to bind in the correct location, potentially changing the orientation of ACP in LpxA's binding site. This re-orientation could move the acyl chain outside the catalytic

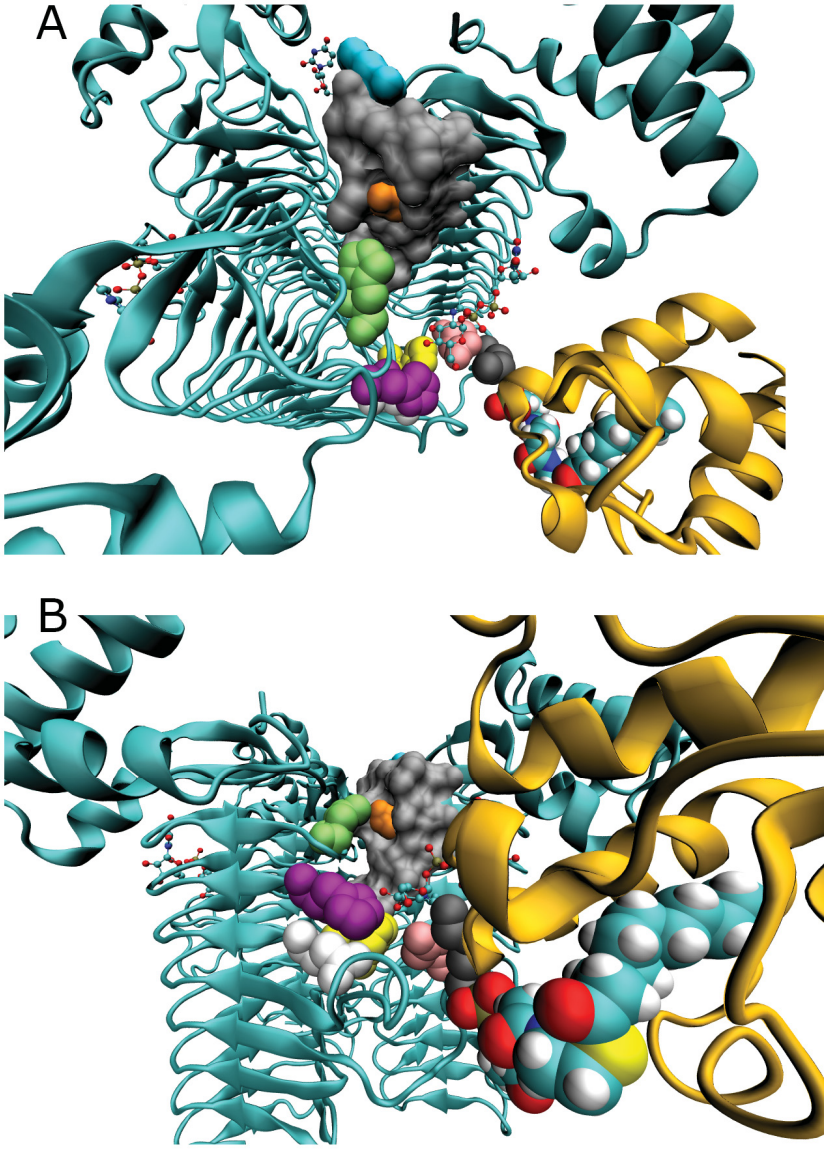

Figure 112 Position of G173 relative to residues involved in binding and catalysis for *E.c.*LpxA. Top (A) and side (B) views of LpxA bound to ACP and UDP-Glc*N*Ac showing position of G173 (orange) relative to important LpxA residues: His 125 (yellow), Asp 126 (white), His 144 (purple), Lys 76 (grey), Asp 74 (pink), His 160 (lime green), His 191 (light turquoise), and the fatty acid binding cleft (silver). LpxA (cyan) and ACP (gold) are depicted in the "NewCartoon" representation, UDP-Glc*N*Ac (colored by atom) is depicted using "CPK" and decanoyl-phosphopantetheine (colored by atom) is depicted using "VDW" graphical representation. All selected residues are displayed using the "VDW" representation while the fatty acid binding cleft is represented using a MSMS surface. The three structures used to create this figure are described in the legend to Figure 109. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1 and Appendix 12). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

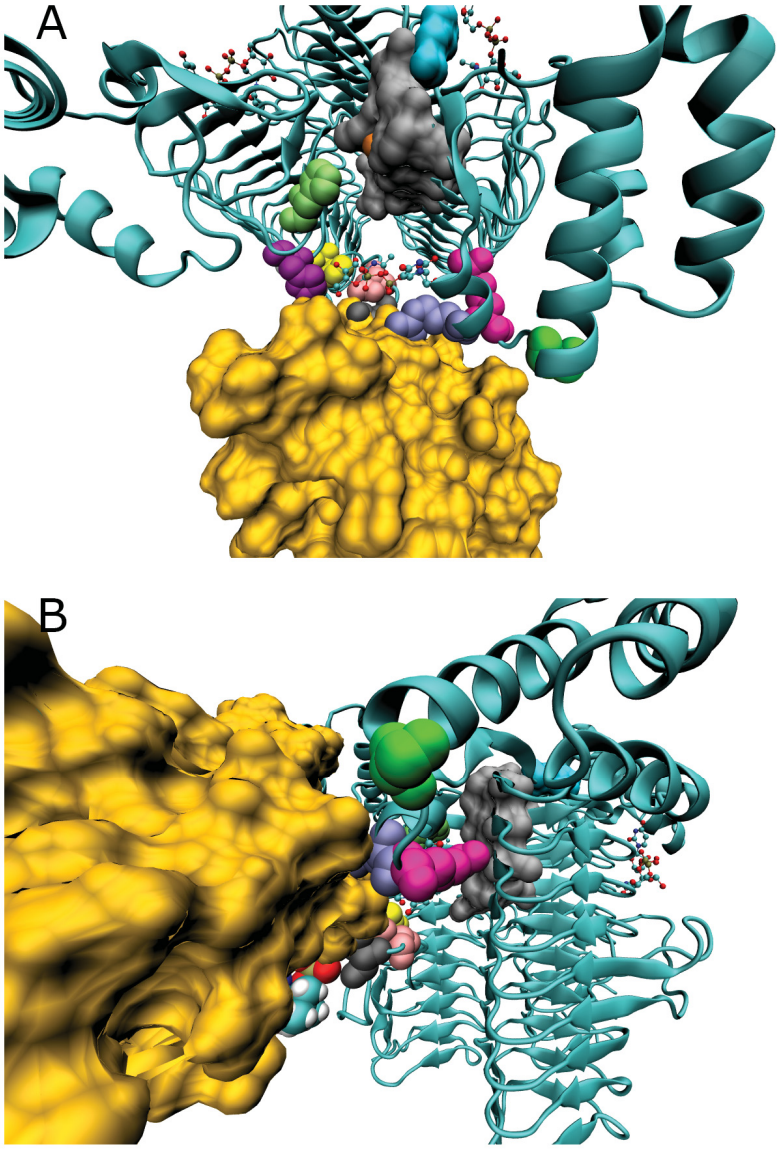

Figure 113 Position of S208 relative to residues involved in binding and catalysis for *E.c.*LpxA. Top (A) and side (B) views of LpxA bound to ACP and UDP-Glc*N*Ac showing the position of S208 (green) relative to important LpxA residues: His 125 (yellow), Asp 126 (white), His 144 (purple), Lys 76 (grey), Asp 74 (pink), His 160 (lime green), His 191 (light turquoise), G173 (orange), Arg 204 (royal blue), Arg 205 (magenta), and the fatty acid binding cleft (silver). LpxA (cyan) and ACP (gold) are depicted in the "NewCartoon" and "Surface" representation, respectively. UDP-Glc*N*Ac is colored by atomand is depicted using "CPK" and decanoyl-phosphopantetheine (colored by atom) is depicted using "VDW" graphical representation. All selected residues are displayed using the "VDW" representation while the fatty acid binding cleft is represented using a MSMS surface. The three structures used to create this figure are described in the legend to Figure 109. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1 and Appendix 12). Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

reach of His 125 and the 3-O of GlcNAc, accounting for the extremely diminished activity demonstrated here.

The dissociation constants determined here (as described in Section 2.4.1.2) are all in the low micromolar range (Table 15), consistent with the relatively high K_M values of most of the partner enzymes (Flaman et al., 2001, Byers & Gong, 2007, Gong et al., 2007) for ACP. Due to the relatively large errors observed in repeated ACP titrations, significant conclusions about ACP residues which are either important for or trigger the conformational change in ACP (allowing the fatty acid to be transferred to LpxA) can not be drawn at this time. However, the K_D values obtained for ACP binding to F162W are similar to the K_M (2 μM) previously estimated for β-OH-myristoyl-ACP, the native substrate for *E.c.*LpxA) (Anderson et al., 1993), and to the K_D of LpxA binding AEDANS-labeled ACP (Gong et al., 2007). Thus I have demonstrated a "proof of concept" for this approach to study interactions between LpxA and ACP mutants under a variety of conditions. Further analyses using the more precise QM4CW spectrofluorometer would likely allow more quantitative interpretations.

AcpS catalyzes the conversion of the inactive apo-ACP to the active holo-ACP through the addition of a phosphopantetheine moiety to Ser 36 of ACP (Polacco & Cronan, 1981, Lambalot & Walsh, 1995, Lambalot & Walsh, 1997). As with LpxA, information on residues that are important for ACP-AcpS interaction is abundant (Parris et al., 2000, Xu et al., 2001); however, unlike LpxA, no major alteration of ACP conformation need be invoked for ACP binding and catalytic activity, as attachment of phosphopantetheine to Ser 36 should not require alterations in the fatty acid binding

pocket of ACP. Thus, mutations in ACP site A (where the action is occurring) have a greater effect on AcpS activity than site B mutations (Gong et al., 2007). Addition of a fluorescent Trp probe to *V. fischeri* ACP (which like *E.c.*LpxA lacks endogenous Trp) provides an experimental system to dissect the role of specific residues in AcpS-ACP interactions. Although experimental characterization and exploitation of this system was not completed and must await further work, molecular modeling was used to provide validation of both the predicted *V. fischeri* AcpS structure and the site(s) of amino acid replacement.

To date, a crystal or NMR structure of *E.c.*AcpS has yet to be determined; however, the structure of AcpS from many other bacterial species has been solved (Chirgadze et al., 2000, Parris et al., 2000, Dall'aglio et al., 2011, Halavaty et al., 2012) (Table 4). When comparing these highly conserved structures (Figure 114) it becomes clear that the likely quaternary structure of all bacterial type II AcpS enzymes is a trimer. Examining the crystal structure of *Bacillus subtilis* AcpS (*B.s.*AcpS) bound to CoA (Figure 115A) and *B.s.*AcpS bound to ACP (Figure 115B) it is also clear that, like LpxA, the active site for AcpS is formed at the subunit interface. Since the active site is highly conserved across species (>80% sequence identity across 264 sequences (Ulaganathan et al., 2007)), it is likely that all bacterial type II AcpS enzymes' active sites are formed at the subunit interface.

Based on this information, a fluorescent probe (Trp) was substituted into the binding region of AcpS. It is known that Helix II of ACP is important for recognition of ACP by partner enzymes (Parris et al., 2000, Zhang et al., 2003a) and that Helix I of AcpS is important for AcpS binding to ACP (Parris et al., 2000, Xu et al., 2001). The

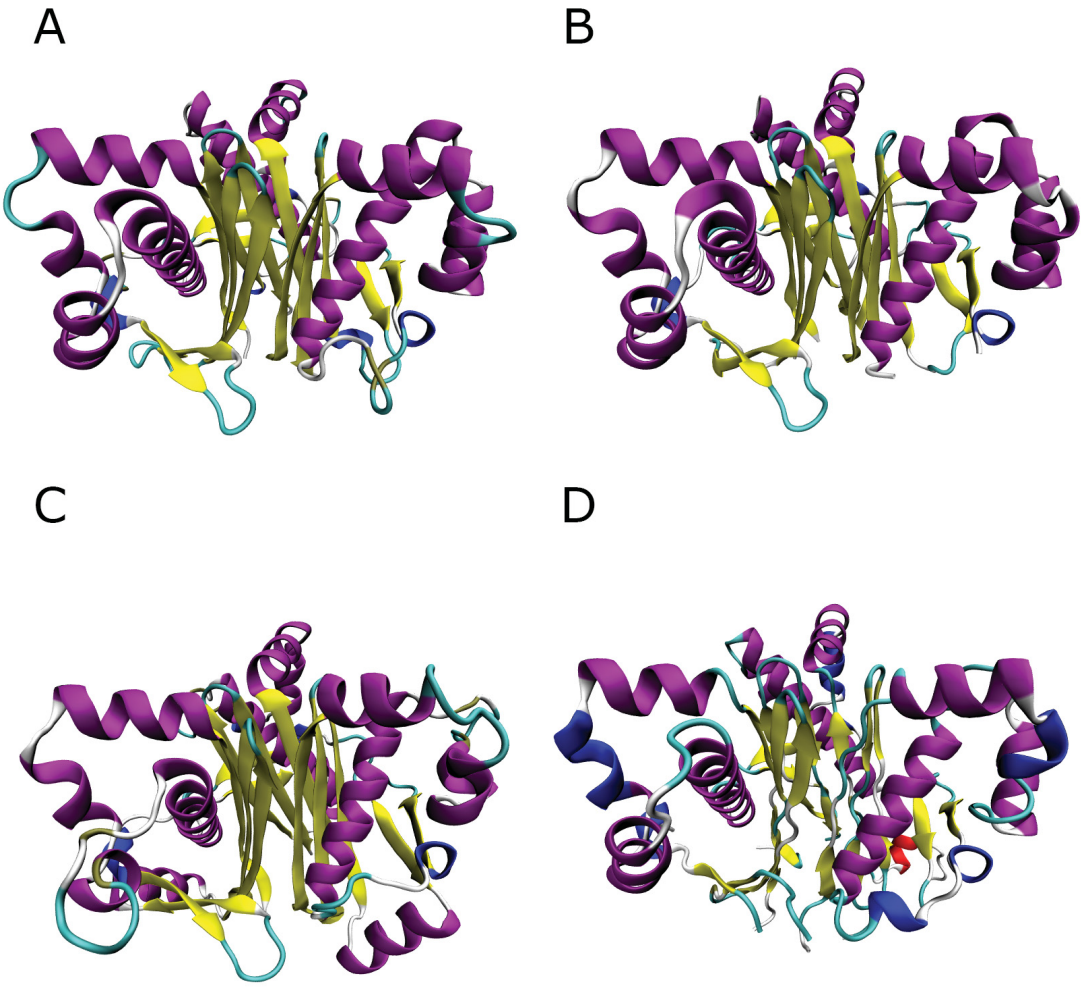

Figure 114 Sample AcpS structures. “NewCartoon” representation of *Bacillus subtilis* (A; PDB ID = 1F7T), *Streptococcus pneumoniae* (B; PDB ID = 1FTE), *Streptomyces coelicolor* (C; PDB ID = 2JCA) and *Staphylococcus aureus* (D; PDB ID = 4JM7) AcpS. All models are colored by secondary structure. Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

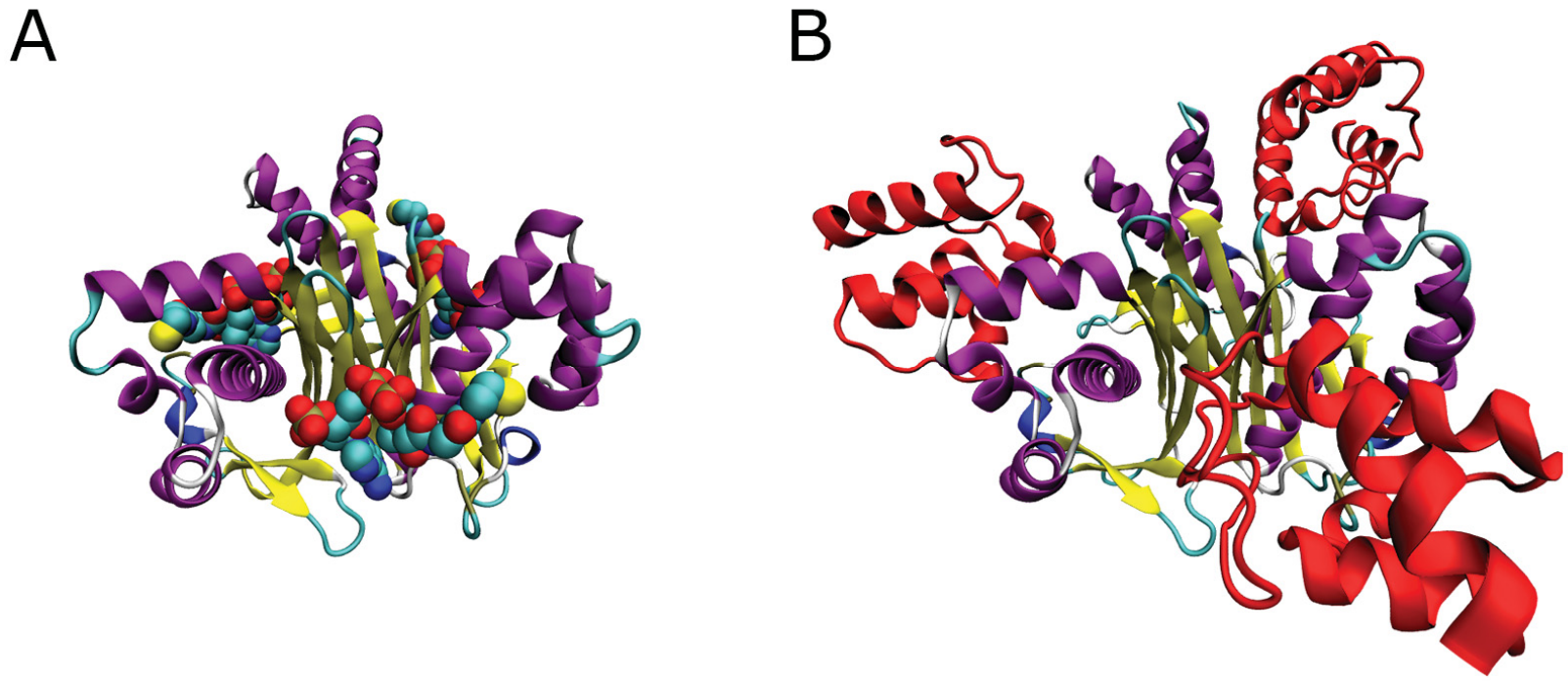

Figure 115 Crystal structure of *B.s.* AcpS bound to coenzyme A (A) and ACP (B). AcpS and ACP (red) are displayed using "NewCartoon" representation; AcpS subunits are colored by secondary structure. CoA is displayed using "VDW" representation and is colored by atom. Two PDBs were used to create this figure: *Bacillus subtilis* AcpS bound to CoA (PDB ID = 1F7L) and ACP (PDB ID = 1F80). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Pictures were created in VMD 1.9.1 and rendered using POV-Ray 3.6.

residue chosen for Trp-substitution, Phe 27, is located in the ACP binding site (Figure 116; purple), but is not directly involved in binding ACP because ACP-enzyme interactions are primarily electrostatic in nature (Zhang et al., 2001).

Previous work has elucidated residues involved in AcpS-ACP binding (Parris et al., 2000, Xu et al., 2001). Specifically, these include interactions between the acidic "recognition" Helix II of ACP and Arg 14, Arg 21, Gln 22 and Arg 24 in Helix I of AcpS (Figure 116; yellow, black, orange and red, respectively). Arg 14 forms a salt bridge with Asp 35 and hydrogen bonds to Asp 38 (Figure 116; blue and silver, respectively), both on ACP. These residues position this end of Helix II of ACP near the bottom of the AcpS binding site which orients Ser 36 (Figure 116; green) for reaction with CoA (Parris et al., 2000, Xu et al., 2001). Arg 21 forms a salt bridge with Glu 41 of ACP (Figure 116; black and cyan, respectively). Asp 48 of ACP is bound by Arg 24 and Gln 22 of AcpS (Figure 116; red and orange, respectively).

As noted in Chapter 1, AcpS can be used to attach fluorescent CoA derivatives to proteins tagged with a consensus ACP phosphopantethine acceptor sequence (Jacquier et al., 2006). In an effort to extend the utility of this system to non-wild-type AcpS-ACP interacting pairs, the charge-change AcpS mutant R21E was made. As mentioned above Arg 21 normally forms a salt-bridge with Glu 41 of ACP. The charge-change AcpS (R21E) was made to be selectively active with the E41K ACP mutant previously made in our lab (Gong & Byers, 2003). To study binding interactions of this charge-change AcpS-ACP pair, a double mutant (R21E/F25W, containing the charge-change and the aforementioned Trp-substitution) was also created.

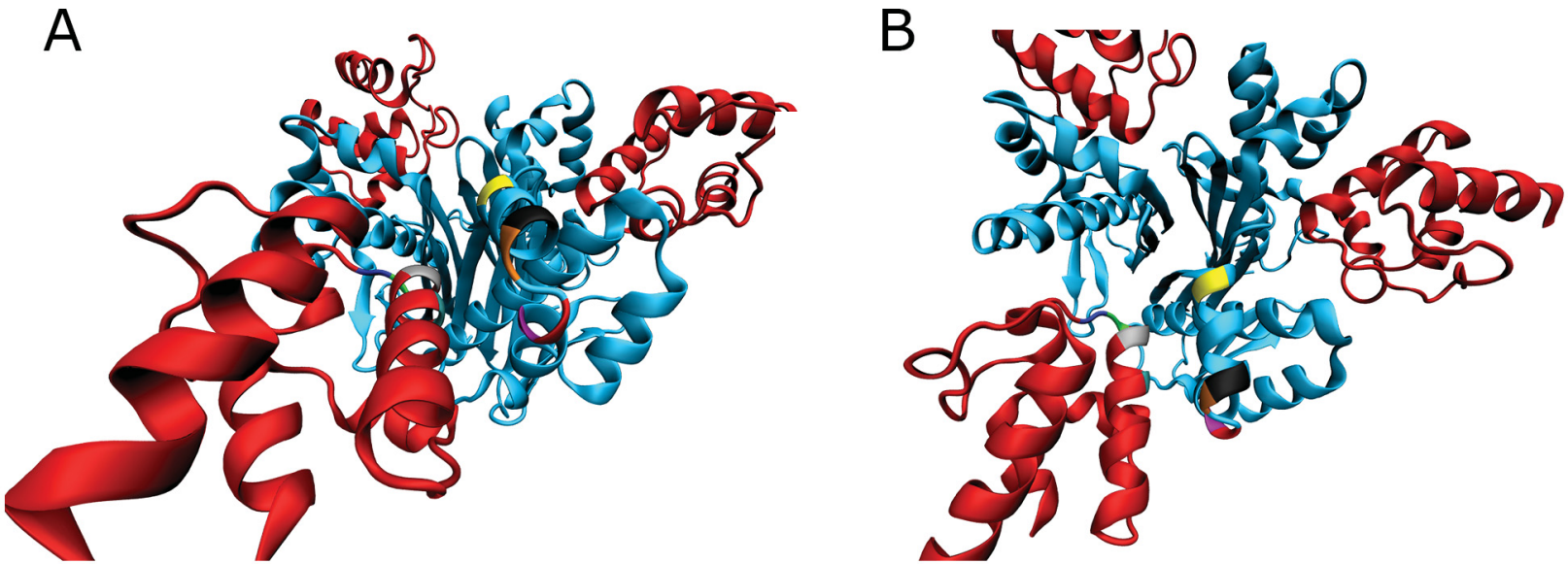

Figure 116 *B.s.*AcpS bound to ACP: location of Trp-substitution relative to important residues. Side (A) and top (B) views of AcpS (light blue) bound to ACP (dark red). Residues involved in binding are highlighted relative to the position selected for Trp-substitution (purple); AcpS: Arg 14 (yellow), Arg 21 (black), Gln 22 (orange) and Arg 25 (red); ACP: Asp 35 (blue), Asp 38 (silver), and Glu 41 (cyan). Ser 36 (green) of ACP is also highlighted as it is the attachment point for the phosphopantetheine moiety to be attached by AcpS. Structure used to create images: *B.s.*AcpS bound to ACP (PDB ID: 1F80). Prior to image creation structures were aligned using Chimera 1.7 (see Appendix 1) to allow for easy comparison. Images were created using VMD 1.91, and POV-Ray 3.6.

Based on MD simulations in this work, Trp-substitution into the *B. subtilis* AcpS-ACP complex of known structure (Parris et al., 2000) did not cause any major changes in local secondary or tertiary structure in either the single (F25W) or double mutant (R21E/F25W); however, longer simulations would be necessary to confirm this. As mentioned in Chapter 2, these mutations are equivalent to the R22E, F27W, and R22E/F27W mutations made to *V.f.*AcpS *in vitro* (see Section 2.1.3). These results are not unexpected as substituting a Trp for Phe is likely to cause the least structural perturbation (as Phe and Trp both have ring-based side chain). Furthermore, distance measurements of these MD simulations suggest that ACP's position in the AcpS active site is virtually unchanged.

Lacking endogenous Trp residues, *V. fischeri* AcpS has been previously used in our laboratory to demonstrate that ACP can adopt its folded conformation in the presence of an enzyme partner (Gong et al 2008). The kinetic properties of this enzyme are essentially indistinguishable from *E. coli* AcpS (A. Murphy and D. Byers, unpublished results), although the predicted pI of *V.f.*AcpS (6.4) is significantly lower than those of either the *E. coli* (9.3) or *B. subtilis* (9.6) enzymes. As a structure of *V.f.*AcpS has not yet been published, homology modeling was used to test if *V.f.*AcpS could potentially adopt the tertiary and quaternary structure suggested by *B.s.*AcpS structures (Parris et al., 2000, Xu et al., 2001). SWISS-MODEL (Swiss Institute for Bioinformatics. *SWISS-MODEL*., Peitsch, 1995, Arnold et al., 2006, Kiefer et al., 2009), utilized for homology modeling (as described in Section 2.5.1.4), produced a potential structure for *V.f.*AcpS. This structure was successfully trimerized *in silico* making it a viable structure for *V.f.*AcpS to adopt *in vivo*. Of note is a small area of secondary structure change (from a β-sheet in the

template to an α-helix in the homology modeled structure). Since this region is located at the periphery of the protein and not in the active site nor near the trimer formation site, activity shouldn't be greatly affected.

As part of an ongoing collaboration with DeNovaMed Inc. to develop antimicrobial AcpS inhibitors, we have recently cloned and expressed the C-terminal domain of *Plasmodium falciparum* AcpS, demonstrated its enzyme activity, and shown that it can be blocked by a subset of proprietary bacterial AcpS inhibitors (Tami et al., 2011). The full length *P. falciparum* AcpS (*P.f.*$AcpS_{FL}$) consists of two domains, not including the transit signal peptides, which locate the enzyme to the apicoplast of the protist (Cai et al., 2005). The N-terminal domain (*P.f.*AcpS-N) is of unknown function but is homologous to the metal-dependent phosphohydrolase enzyme family (Cai et al., 2005), which includes *E. coli* AcpH ACP phosphodiesterase (AcpH) (Thomas et al., 2007). The C-terminal domain (*P.f.*AcpS-C), shares sequence identity (30 – 50%) with bacterial AcpS enzymes. *P.f.*AcpS-C is 65% identical to the C-terminal fragment of *Plasmodium yoelii* AcpS, for which the crystal structure has been determined (PDB ID = 2BDD (Vedadi et al., 2007)) and PDB ID: 2QG8). These two structures, as well as *B.s.*AcpS (PDB ID = 1F7T (Parris et al., 2000)) were used as templates for homology modeling. The three viable structures produced were essentially identical and thus only one was subjected to MD simulation to approach an equilibrated structure. The simulation end structure was trimerized *in silico*, as for *V.f.*AcpS, providing a viable structure for further structure-function analyses of this potentially valuable antimalarial drug target.

The N-terminal fragment of *P.f.*AcpS is hypothesized to have ACP phosphodiesterase activity based on its sequence homology with bacterial AcpH (Cai et al., 2005, Thomas et al., 2007). This raises the intriguing possibility of a bifunctional enzyme unique to the *Plasmodium* apicoplast that is capable of both addition and removal of the PPT (or acyl-PPT) moiety from apo-ACP. To date there isn't a published structure of *E.c.*AcpH; however, previous work demonstrated that *E.c.*AcpH can be convincingly modeled on the structure of the hydrolase domain of the bifunctional bacterial stringent response protein, SpoT (Thomas et al., 2007). Thus, *E.c.*AcpH was homology modeled using SpoT as a template. As expected a viable structure was produced for *E.c.*AcpH. This structure and SpoT were both used as templates for modeling of *P.f.*AcpS-N (without its leader sequence). Both templates produced viable models of *P.f.*AcpS-N. Of note is that the final homology modeled *E.c.*AcpH and *P.f.*AcpS-N structures produced were not affected by missing amino acids in the template, SpoT. However, one major caveat of homology modeling is that essentially any sequence can be threaded through any structure, as demonstrated for BSA homology modeled using SpoT as a template (Figure 106B). Thus, before attempting homology modeling, a search for viable templates must be performed. In other words, a structure with sequence homology to the protein to be modeled must be used as a template.

CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS

In the current work, two main themes have been explored: first, the effect of constraining ACP's physical structure on its conformational stability and function (Chapter 3), and second, expanding the use of fluorescence methods in the study of ACP dynamics and partner enzyme interactions (Chapters 4 and 5, respectively).

Perhaps the most significant contribution of my thesis is a better understanding of the requirements and limitations of ACP conformation for its proper function. All experimental results from Chapter 3 indicate that cyclized rACP is stabilized in its folded conformation relative to linear ACP, yet this has only a moderate impact on ACP growth *in vivo*. Conversely, cyclization of the folding-deficient F50A mutant restores both its ability to fold in the presence of Mg^{2+} *in vitro*, and sustain growth *in vivo*. Thus, we can conclude that, at least for its essential functions, ACP: (i) does not require complete unfolding (i.e. separation of its N- and C-termini), but (ii) does require the ability to adopt a folded conformation.

Whether the modest negative effect of ACP cyclization on growth is due to altered interactions with one or more rate-limiting enzymes is not currently known. This could be explored further through detailed comparison of kinetic parameters for various ACP-dependent enzymes (e.g. holo-ACP synthase, FAS components, and acyltransferases) with cyclic *versus* linear ACP. MD simulations on the apo forms of linear and cyclic ACP indicate that while flexibility is lost in some portions of cyclic ACP, it is gained in others, and further MD simulations involving holo and acyl forms of both ACPs would likely expand our understanding of the dynamics of the various cellular ACP substrates. Alternatively, metabolomic approaches comparing linear and cyclic

ACPs in our complementation system could be used to monitor key intermediates in metabolism and deduce the pathways and enzymes most affected.

The ability of cyclic ACP to restore growth in an ACP-deficient strain could also be related to whether or not it can sequester an attached fatty acid in its hydrophobic core. Although this reversible transition has now been structurally characterized during the catalytic cycle of some ACP-dependent enzymes (Masoudi et al., 2014, Nguyen et al., 2014), we do not know whether cyclization of ACP affects its ability to enclose hydrophobic acyl chains. A constantly exposed acyl chain, akin to that of the internal ACP domain of rat type I fatty acid synthase (Ploskoń et al., 2008), would undoubtedly alter the thermodynamic and kinetic parameters of many ACP-dependent enzymes while not necessarily blocking function completely.

This question could be investigated experimentally through the use of either NMR or fluorescence approaches. ^{19}F-NMR of fluorinated fatty acids attached to linear *versus* cyclic ACP could be used to directly monitor the free *versus* bound environment of the acyl group. Alternatively, attachment of distal brominated fatty acids to cyclic *versus* linear L46W could be used to examine the proximity of the acyl chain to position 46 in the hydrophobic core using fluorescence. Bromine (Br) has been shown to quench fluorescence by a mechanism known as heavy atom quenching (Barber, 1989, Bolen & Holloway, 1990). Based on available acyl-ACP structures, either Br-6:0 or Br-8:0 would be expected to align with and quench the fluorescence of Trp 46 if the acyl chain was sequestered, but not if it was inaccessible to the binding pocket. Preliminary data from our group indicate that *V. harveyi* acyl-ACP synthetase is tolerant of modified FAs and could be used to prepare these acyl-ACPs. Acyl chain sequestration could also be tested

in silico using MD of the holo *versus* acyl forms of cyclic and linear ACP. Longer simulations, with larger spherical boundary conditions and with varying starting positions of both PP and acyl moieties, would provide insight into the location of and dynamics involved in acyl group interaction with cyclic ACP.

Although the results obtained in Chapter 4 indicate that Tyr 71 fluorescence is somewhat sensitive to its environment and to changes in ACP conformation caused by divalent cation binding, the data do not validate this as a generally useful intrinsic probe for ACP conformation and dynamics. The reasons for this are discussed in more detail in Chapter 4, but amount to the relative lack of Tyr fluorescence sensitivity and to the number of control variables involved. The latter was dominated by photobleaching of Tyr fluorescence, which could be controlled in two ways. First, using the programmable shutter on the QM4CW exposure time of Tyr to light could be drastically decreased, potentially eliminating photobleaching, although this would also reduce fluorescence emission such that it may be necessary to increase the concentration of ACPs used for experiments. The second method would be to scan just the peak (rather than the full range) of Tyr emission, thus reducing the amount of time Tyr is exposed to light. This method can be used since there isn't a shift in peak fluorescence emission of Tyr as there is for L46W. The possible advantage of this method over using the programmable shutter is that no loss of signal intensity should be involved since the shutter would be open for the normal integration time. Both methods would likely be required to confidently determine effects of mutations on ACP's dynamics.

Stopped-flow fluorescence analysis of L46W in response to Mg^{2+} addition suggests this as a promising approach to measuring the kinetics of conformational

transitions of ACP in response to various perturbations. This could be expanded to monitoring the interaction kinetics with partner enzymes lacking endogenous Trp residues (or with just one or two, using a newly acquired fluorescence lifetime accessory). However, it is unfortunate that (at least based on the current results) this approach cannot be extended to Tyr 71 as a probe for studying the conformational kinetics of the many existing ACP mutants. Further work will hopefully resolve some of the technical issues encountered here and establish the practicality of Tyr stopped-flow analysis in ACP analysis.

Lastly, experiments from Chapter 5 have established and validated the concept of introducing Trp as a fluorescent probe into two ACP-dependent enzymes that lack endogenous Trp (*E. coli* LpxA and *V. fischeri* AcpS), although more accurate and sensitive instrumentation (such as the QM4CW) will be necessary to fully take advantage of this system. My work has suggested that the increased activity of F162W *versus* wild-type LpxA could be due to possible increased hydrophobic interactions between F162W and the acyl chain, causing easier extraction of the attached fatty acid from ACP's hydrophobic core. This could be tested by comparing the affinity of F162W with holo- *versus* acyl-ACP using the fluorescence method developed in this work.

Further information about how ACP passes an attached acyl chain to a partner enzyme could be explored using the newly acquired fluorescence lifetime excitation source for the QM4CW. This attachment allows for the measurement of fluorescence lifetimes of fluorphores; normal fluorescence lifetimes of Trp are typically ~3.5 and ~0.5 ns (Lakowicz, 2006). As noted above, titrating the appropriate Br-acyl modified L46W with G173W LpxA, it could be possible to examine the transfer of the fatty acid from the

hydrophobic core of ACP to the hydrophobic fatty acid binding cleft of LpxA. While the acyl chain is sequestered, the Trp in L46W would likely be quenched and have a modified fluorescence lifetime. Upon titration of increasing amounts of G173W LpxA, one would expect the acyl chain to be passed to the hydrophobic binding cleft on LpxA, thereby both restoring the fluorescence lifetime and intensity of the ACP Trp 43 to normal and causing a change in the fluorescence lifetime and quenching of Trp 173 in LpxA.

Of course, further fluorescence characterization of the Trp-substituted LpxA and AcpS mutants created in this study using the superior QM4CW instrument would provide more detailed information about the roles of specific enzyme and ACP residues involved in interaction and activity. This could in theory be accompanied by parallel *in vivo* studies using the *E. coli* complementation system developed in our collaboration with John Cronan and described in Chapter 3 (Volkmann et al 2010). Like ACP, both LpxA and AcpS are essential enzymes in *E. coli*, and preparation of vectors for the deletion of the wild-type enzyme and complementation with mutant enzymes should be relatively straightforward given the materials and expertise acquired previously. As demonstrated in my thesis, the combination of biophysical and functional information has the potential provide great insight into the biology of ACP and its partner enzymes.

REFERENCES

Alatossava, T., Jütte, H., Kuhn, A., and Kellenberger, E. (1985). Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized *Escherichia coli* by ionophore A23187. *Journal of Bacteriology* **162**, 413-419

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schaffer, A. A., and Yu, Y. K. (2005). Protein database searches using compositionally adjusted substitution matrices. *The FEBS Journal* **272**, 5101-5109

Amaro, M., Birch, D. J., and Rolinski, O. J. (2011). Beta-amyloid oligomerisation monitored by intrinsic tyrosine fluorescence. *Physical Chemistry Chemical Physics* **13**, 6434-6441

Anderson, M. S., and Raetz, C. R. (1987). Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyltransferase that converts UDP-*N*-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-*N*-acetylglucosamine. *The Journal of Biological Chemistry* **262**, 5159-5169

Anderson, M. S., Bull, H. G., Galloway, S. M., Kelly, T. M., Mohan, S., Radika, K., and Raetz, C. R. (1993). UDP-*N*-acetylglucosamine acyltransferase of *Escherichia coli*. The first step of endotoxin biosynthesis is thermodynamically unfavorable. *The Journal of Biological Chemistry* **268**, 19858-19865

Andrec, M., Hill, R. B., and Prestegard, J. H. (1995). Amide exchange rates in *Escherichia coli* acyl carrier protein: correlation with protein structure and dynamics. *Protein Science* **4**, 983-993

Applied Biosystems. *POROS® 50 R1 and R2 Perfusion Chromatography® Bulk Media for Reversed-Phase Chromatography: Operating Instructions*. Web Page: http://www3.appliedbiosystems.com/cms/groups/psm_support/documents/generaldocuments/cms_041645.pdf.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201

Attwood, P. V., and Wallace, J. C. (2002). Chemical and catalytic mechanisms of carboxyl transfer reactions in biotin-dependent enzymes. *Accounts of Chemical Research* **35**, 113-120

Babinski, K. J., Kanjilal, S. J., and Raetz, C. R. (2002a). Accumulation of the lipid A precursor UDP-2,3-diacylglucosamine in an *Escherichia coli* mutant lacking the lpxH gene. *The Journal of Biological Chemistry* **277**, 25947-25956

Babinski, K. J., Ribeiro, A. A., and Raetz, C. R. (2002b). The *Escherichia coli* gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *The Journal of Biological Chemistry* **277**, 25937-25946

Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H., and Nienaber, V. (2012). Structure determination of LpxA from the lipopolysaccharide-synthesis pathway of *Acinetobacter baumannii*. *Acta Crystallographica Section F* **68**, 1477-1481

Barber, J. (1989). Membrane proteins. Detergent ringing true as a model for membranes. *Nature* **340**, 601

Battesti, A., and Bouveret, E. (2006). Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. *Molecular Microbiology* **62**, 1048-1063

Battesti, A., and Bouveret, E. (2009). Bacteria possessing two RelA/SpoT-like proteins have evolved a specific stringent response involving the acyl carrier protein-SpoT interaction. *Journal of Bacteriology* **191**, 616-624

Baugh, L., Gallagher, L. A., Patrapuvich, R., Clifton, M. C., Gardberg, A. S., Edwards, T. E., Armour, B., Begley, D. W., Dieterich, S. H., Dranow, D. M., Abendroth, J., Fairman, J. W., Fox, D., 3rd, Staker, B. L., Phan, I., Gillespie, A., Choi, R., Nakazawa-Hewitt, S., Nguyen, M. T., Napuli, A., Barrett, L., Buchko, G. W., Stacy, R., Myler, P. J., Stewart, L. J., Manoil, C., and Van Voorhis, W. C. (2013). Combining functional and structural genomics to sample the essential Burkholderia structome. *PloS One* **8**, e53851

Bayan, N., and Thérisod, H. (1989). Evidence for interactions of acyl carrier protein with glycerol-3-phosphate acyltransferase, an inner membrane protein of *Escherichia coli*. *FEBS Letters* **255**, 330-334

Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P., and Burkart, M. D. (2014). The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Natural Product Reports* **31**, 61-108

Belunis, C. J., and Raetz, C. R. (1992). Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from *Escherichia coli*. *Journal of Biological Chemistry* **267**, 9988-9997

Bolen, E. J., and Holloway, P. W. (1990). Quenching of tryptophan fluorescence by brominated phospholipid. *Biochemistry* **29**, 9638-9643

Boom, T. V., and Cronan, J. E. (1989). Genetics and regulation of bacterial lipid metabolism. *Annual Review of Microbiology* **43**, 317-343

Brennan, P. J., and Nikaido, H. (1995). The envelope of mycobacteria. *Annual Review of Biochemistry* **64**, 29-63

Brody, S., and Mikolajczyk, S. (1988). Neurospora mitochondria contain an acyl-carrier protein. *European Journal of Biochemistry* **173**, 353-359

Brody, S., Oh, C., Hoja, U., and Schweizer, E. (1997). Mitochondrial acyl carrier protein is involved in lipoic acid synthesis in *Saccharomyces cerevisiae*. *FEBS Letters* **408**, 217-220

Brozek, K. A., Hosaka, K., Robertson, A. D., and Raetz, C. R. (1989). Biosynthesis of lipopolysaccharide in *Escherichia coli*. Cytoplasmic enzymes that attach 3-deoxy-D-manno-octulosonic acid to lipid A. *The Journal of Biological Chemistry* **264**, 6956-6966

Brozek, K. A., and Raetz, C. R. (1990). Biosynthesis of lipid A in *Escherichia coli*. Acyl carrier protein-dependent incorporation of laurate and myristate. *Journal of Biological Chemistry* **265**, 15410-15417

Bunkoczi, G., Pasta, S., Joshi, A., Wu, X., Kavanagh, K. L., Smith, S., and Oppermann, U. (2007). Mechanism and substrate recognition of human holo ACP synthase. *Chemistry & Biology* **14**, 1243-1253

Butland, G., Peregrin-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* **433**, 531-537

Byers, D. M., and Meighen, E. A. (1985). Acyl-acyl carrier protein as a source of fatty acids for bacterial bioluminescence. *Proceedings of the National Academy of Sciences USA* **82**, 6085-6089

Byers, D. M. (1988). Luminescence-specific synthesis of myristic acid in the bioluminescent bacterium *Vibrio harveyi*. *Biochemistry and Cell Biology* **66**, 741-749

Byers, D. M. (1989). Elongation of exogenous fatty acids by the bioluminescent bacterium *Vibrio harveyi*. *Journal of Bacteriology* **171**, 59-64

Byers, D. M., and Gong, H. (2007). Acyl carrier protein: structure-function relationships in a conserved multifunctional protein family. *Biochemistry and cell biology* **85**, 649-662

Cai, X., Herschap, D., and Zhu, G. (2005). Functional characterization of an evolutionarily distinct phosphopantetheinyl transferase in the apicomplexan *Cryptosporidium parvum*. *Eukaryotic Cell* **4**, 1211-1220

CambridgeSoft. *ChemDraw*. Web Page: https://www.cambridgesoft.com/Ensemble_for_Chemistry/ChemDraw/

Cane, D. E., and Walsh, C. T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chemistry & Biology* **6**, R319-R325

Cao, J. G., and Meighen, E. A. (1989). Purification and structural identification of an autoinducer for the luminescence system of *Vibrio harveyi*. *Journal of Biological Chemistry* **264**, 21670-21676

Carreras, C. W., and Khosla, C. (1998). Purification and *in vitro* reconstitution of the essential protein components of an aromatic polyketide synthase. *Biochemistry* **37**, 2084-2088

Chan, D. I., Stockner, T., Tieleman, D. P., and Vogel, H. J. (2008). Molecular dynamics simulations of the Apo-, Holo-, and acyl-forms of *Escherichia coli* acyl carrier protein. *The Journal of Biological Chemistry* **283**, 33620-33629

Chan, D. I., Chu, B. C. H., Lau, C. K. Y., Hunter, H. N., Byers, D. M., and Vogel, H. J. (2010). NMR solution structure and biophysical characterization of *Vibrio harveyi* acyl carrier protein A75H: Effects of divalent metal ions. *Journal of Biological Chemistry* **285**, 30558-30566

Chan, D. I., and Vogel, H. J. (2010). Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochemical Journal* **430**, 1-19

Chauvenet, W. (1960 (Reprint of 1891)). *A Manual of Spherical and Practical Astronomy V. II*, 5th ed., Dover, N.Y. 474-566.

Chirgadze, N. Y., Briggs, S. L., McAllister, K. A., Fischl, A. S., and Zhao, G. (2000). Crystal structure of *Streptococcus pneumoniae* acyl carrier protein synthase: an essential enzyme in bacterial fatty acid biosynthesis. *The EMBO Journal* **19**, 5281-5287

Clementz, T., and Raetz, C. R. (1991). A gene coding for 3-deoxy-D-manno-octulosonic-acid transferase in *Escherichia coli*. Identification, mapping, cloning, and sequencing. *The Journal of Biological Chemistry* **266**, 9687-9696

Coleman, J., and Raetz, C. R. (1988). First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of the *lpxA* gene. *Journal of Bacteriology* **170**, 1268-1274

Coleman, J. (1990). Characterization of *Escherichia coli* cells deficient in 1-acyl-sn-glycerol-3- phosphate acyltransferase activity. *The Journal of Biological Chemistry* **265**, 17215-17221

Collins, L. V., Kristian, S. A., Weidenmaier, C., Faigle, M., van Kessel, K. P. M., van Strijp, J. A. G., Götz, F., Neumeister, B., and Peschel, A. (2002). *Staphylococcus aureus* strains lacking D-alanine modifications of teichoic acids are highly susceptible to human neutrophil killing and are virulence attenuated in mice. *Journal of Infectious Diseases* **186**, 214-219

Cook, H. W., and McMaster, C. R. (2002). *Biochemistry of lipids, lipoproteins and membranes (Chapter title: Fatty acid desaturation and chain elongation in eukaryotes.)*, Elsevier, Amsterdam. 181 – 204.

Cooper, C. L., Hsu, L., Jackowski, S., and Rock, C. O. (1989). 2-Acylglycerolphosphoethanolamine acyltransferase/acyl-acyl carrier protein synthetase is a membrane-associated acyl carrier protein binding protein. *Journal of Biological Chemistry* **264**, 7384-7389

Cronan, J. E., Fearnley, I. M., and Walker, J. E. (2005). Mammalian mitochondria contain a soluble acyl carrier protein. *FEBS Letters* **579**, 4892-4896

Cronan, J. E., Jr., and Waldrop, G. L. (2002). Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* **41**, 407-435

Crump, M. P., Crosby, J., Dempsey, C. E., Murray, M., Hopwood, D. A., and Simpson, T. J. (1996). Conserved secondary structure in the actinorhodin polyketide synthase acyl carrier protein from *Streptomyces coelicolor* A3(2) and the fatty acid synthase acyl carrier protein from *Escherichia coli*. *FEBS Letters* **391**, 302-306

Cryle, M. J., and De Voss, J. J. (2004). Carbon-carbon bond cleavage by cytochrome P450BioI (CYP107H1). *Chemical Communications*, 86-87

Dall'aglio, P., Arthur, C. J., Williams, C., Vasilakis, K., Maple, H. J., Crosby, J., Crump, M. P., and Hadfield, A. T. (2011). Analysis of *Streptomyces coelicolor* phosphopantetheinyl transferase, AcpS, reveals the basis for relaxed substrate specificity. *Biochemistry* **50**, 5704-5717

David, S. A. (2001). Towards a rational development of anti-endotoxin agents: novel approaches to sequestration of bacterial endotoxins with small molecules. *Journal of Molecular Recognition* **14**, 370-387

de la Roche, M. A., Shen, Z., and Byers, D. M. (1997). Hydrodynamic properties of *Vibrio harveyi* acyl carrier protein and its fatty-acylated derivatives. *Archives of Biochemistry and Biophysics* **344**, 159-164

De Lano, W. (2002). The PyMol Molecular Graphics System. San Carlos, CA

De Lay, N. R., and Cronan, J. E. (2006). Gene-specific random mutagenesis of *Escherichia coli in vivo*: isolation of temperature-sensitive mutations in the acyl carrier protein of fatty acid synthesis. *Journal of Bacteriology* **188**, 287-296

De Lay, N. R., and Cronan, J. E. (2007). *In vivo* functional analyses of the type II acyl carrier proteins of fatty acid biosynthesis. *Journal of Biological Chemistry* **282**, 20319-20328

Debabov, D. V., Kiriukhin, M. Y., and Neuhaus, F. C. (2000). Biosynthesis of lipoteichoic acid in *Lactobacillus rhamnosus*: role of DltD in D-alanylation. *Journal of Bacteriology* **182**, 2855-2864

Dennis, J. J., and Zylstra, G. J. (1998). Improved antibiotic-resistance cassettes through restriction site elimination using Pfu DNA polymerase PCR. *Biotechniques* **25**, 772-774, 776

Dowhan, W. (2013). A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1831**, 471-494

Dunker, A. K., Obradovic, Z., Romero, P., Garner, E. C., and Brown, C. J. (2000). Intrinsic protein disorder in complete genomes. *Genome Informatics. Workshop on Genome Informatics* **11**, 161-171

Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002). Intrinsic disorder and protein function. *Biochemistry* **41**, 6573-6582

Dym, O., Albeck, S., Peleg, Y., Schwarz, A., Shakked, Z., Burstein, Y., and Zimhony, O. (2009). Structure-function analysis of the acyl carrier protein synthase (AcpS) from *Mycobacterium tuberculosis*. *Journal of Molecular Biology* **393**, 937-950

Eliot, A. C., Sandmark, J., Schneider, G., and Kirsch, J. F. (2002). The dual-specific active site of 7,8-diaminopelargonic acid synthase and the effect of the R391A mutation. *Biochemistry* **41**, 12582-12589

Elovson, J., and Vagelos, P. R. (1968). Acyl carrier protein: X. Acyl carrier protein synthetase. *Journal of Biological Chemistry* **243**, 3603-3611

European Bioinformatics Institute. *Clustal Omega*. Web Page: http://www.ebi.ac.uk/Tools/msa/clustalo/

European Bioinformatics Institute. *Protein Interfaces, Surfaces and Assemblies*. Web Page: http://pdb.org/pdb/home/home.do

Evans, S. E., Williams, C., Arthur, C. J., Burston, S. G., Simpson, T. J., Crosby, J., and Crump, M. P. (2008). An ACP structural switch: conformational differences between the apo and holo forms of the actinorhodin polyketide synthase acyl carrier protein. *ChemBioChem* **9**, 2424-2432

ExPASy. *Compute pI/MW Tool*. Web Page: http://web.expasy.org/compute_pi/

FFmpeg. *FFmpeg: a complete, cross-platform solution to record, convert and stream audio and video.* Web Page: http://www.ffmpeg.org/index.html

Fice, D., Shen, Z., and Byers, D. M. (1993). Purification and characterization of fatty acyl-acyl carrier protein synthetase from *Vibrio harveyi*. *Journal of Bacteriology* **175**, 1865-1870

Findlow, S. C., Winsor, C., Simpson, T. J., Crosby, J., and Crump, M. P. (2003). Solution structure and dynamics of oxytetracycline polyketide synthase acyl carrier protein from *Streptomyces rimosus*. *Biochemistry* **42**, 8423-8433

Finking, R., Mofid, M. R., and Marahiel, M. A. (2004). Mutational analysis of peptidyl carrier protein and acyl carrier protein synthase unveils residues involved in protein-protein recognition. *Biochemistry* **43**, 8946-8956

Flaman, A. S., Chen, J. M., Van Iderstine, S. C., and Byers, D. M. (2001). Site-directed Mutagenesis of Acyl Carrier Protein (ACP) Reveals Amino Acid Residues Involved in ACP Structure and Acyl-ACP Synthetase Activity. *Journal of Biological Chemistry* **276**, 35934-35939

Flugel, R. S., Hwangbo, Y., Lambalot, R. H., Cronan, J. E., Jr., and Walsh, C. T. (2000). Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in *Escherichia coli*. *The Journal of Biological Chemistry* **275**, 959-968

Frederick, A. F., Kay, L. E., and Prestegard, J. H. (1988). Location of divalent ion sites in acyl carrier protein using relaxation perturbed 2D NMR. *FEBS Letters* **238**, 43-48

Fugate, C. J., and Jarrett, J. T. (2012). Biotin synthase: insights into radical-mediated carbon-sulfur bond formation. *Biochimica et Biophysica Acta* **1824**, 1213-1222

Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. (1999). A new pathway for polyketide synthesis in microorganisms. *Nature* **400**, 897-899

Fuqua, C., and Greenberg, E. P. (2002). Listening in on bacteria: acyl-homoserine lactone signalling. *Nature Reviews. Molecular Cell Biology* **3**, 685-695

Gallagher, J. R., and Prigge, S. T. (2010). *Plasmodium falciparum* acyl carrier protein crystal structures in disulfide-linked and reduced states and their prevalence during blood stage growth. *Proteins* **78**, 575-588

Galloway, S. M., and Raetz, C. R. (1990). A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis. *The Journal of Biological Chemistry* **265**, 6394-6402

Gally, H. U., Spencer, A. K., Armitage, I. M., Prestegard, J. H., and Cronan, J. E., Jr. (1978). Acyl carrier protein from *Escherichia coli*: characterization by proton and fluorine-19 nuclear magnetic resonance and evidence for restricted mobility of the fatty acid chain in tetradecanoyl-acyl-carrier protein. *Biochemistry* **17**, 5377-5382

Gangar, A., Karande, A. A., and Rajasekharan, R. (2001). Purification and characterization of acyl-acyl carrier protein synthetase from oleaginous yeast and its role in triacylglycerol biosynthesis. *Biochem. J.* **360**, 471-479

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M.-S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M. A., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M., and Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498-511

Geiger, O., Spaink, H. P., and Kennedy, E. P. (1991). Isolation of the *Rhizobium leguminosarum* NodF nodulation protein: NodF carries a 4'-phosphopantetheine prosthetic group. *Journal of Bacteriology* **173**, 2872-2878

George, N., Pick, H., Vogel, H., Johnsson, N., and Johnsson, K. (2004). Specific labeling of cell surface proteins with chemically diverse compounds. *Journal of the American Chemical Society* **126**, 8896-8897

Gibson, K. J. (1997). Isolation and chemistry of the mixed anhydride intermediate in the reaction catalyzed by dethiobiotin synthetase. *Biochemistry* **36**, 8474-8478

Gokulan, K., Aggarwal, A., Shipman, L., Besra, G. S., and Sacchettini, J. C. (2011). *Mycobacterium tuberculosis* acyl carrier protein synthase adopts two different pH-dependent structural conformations. *Acta crystallographica. Section D, Biological Crystallography* **67**, 657-669

Gong, H., and Byers, D. M. (2003). Glutamate-41 of *Vibrio harveyi* acyl carrier protein is essential for fatty acid synthase but not acyl-ACP synthetase activity. *Biochemical and Biophysical Research Communications* **302**, 35-40

Gong, H., Murphy, A., McMaster, C. R., and Byers, D. M. (2007). Neutralization of acidic residues in helix II stabilizes the folded conformation of acyl carrier protein and variably alters its function with different enzymes. *The Journal of Biological Chemistry* **282**, 4494-4503

Gong, H., Murphy, P. W., Langille, G. M., Minielly, S. J., Murphy, A., McMaster, C. R., and Byers, D. M. (2008). Tryptophan fluorescence reveals induced folding of *Vibrio harveyi* acyl carrier protein upon interaction with partner enzymes. *Biochimica et Biophysica Acta* **1784**, 1835-1843

Green, P. R., Merrill, A. H., Jr., and Bell, R. M. (1981). Membrane phospholipid synthesis in *Escherichia coli*. Purification, reconstitution, and characterization of sn-glycerol-3-phosphate acyltransferase. *The Journal of Biological Chemistry* **256**, 11151-11159

Gromada. *VideoMach*. Web Page: http://gromada.com/videomach/

Gross, M., Cramton, S. E., Götz, F., and Peschel, A. (2001). Key role of teichoic acid Net charge in *Staphylococcus aureus* colonization of artificial surfaces. *Infection and Immunity* **69**, 3423-3426

Gueguen, V., Macherel, D., Jaquinod, M., Douce, R., and Bourguignon, J. (2000). Fatty acid and lipoic acid biosynthesis in higher plant mitochondria. *Journal of Biological Chemistry* **275**, 5016-5025

Gully, D., Moinier, D., Loiseau, L., and Bouveret, E. (2003). New partners of acyl carrier protein detected in *Escherichia coli* by tandem affinity purification. *FEBS Letters* **548**, 90-96

Gully, D., and Bouveret, E. (2006). A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in *Escherichia coli*. *Proteomics* **6**, 282-293

Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology* **177**, 4121-4130

Halavaty, A. S., Kim, Y., Minasov, G., Shuvalova, L., Dubrovska, I., Winsor, J., Zhou, M., Onopriyenko, O., Skarina, T., Papazisi, L., Kwon, K., Peterson, S. N., Joachimiak, A., Savchenko, A., and Anderson, W. F. (2012). Structural characterization and comparison of three acyl-carrier-protein synthases from pathogenic bacteria. *Acta crystallographica. Section D, Biological Crystallography* **68**, 1359-1370

Hanzelka, B. L., Parsek, M. R., Val, D. L., Dunlap, P. V., Cronan, J. E., and Greenberg, E. P. (1999). Acylhomoserine lactone synthase activity of the *Vibrio fischeri* AinS protein. *Journal of Bacteriology* **181**, 5766-5770

Harington, A., Herbert, C. J., Tung, B., Getz, G. S., and Slonimski, P. P. (1993). Identification of a new nuclear gene (CEM1) encoding a protein homologous to a β-keto-acyl synthase which is essential for mitochondrial respiration in *Saccharomyces cerevisiae*. *Molecular Microbiology* **9**, 545-555

Heath, R. J., and Rock, C. O. (1999). A missense mutation accounts for the defect in the glycerol-3-phosphate acyltransferase expressed in the plsB26 mutant. *Journal of Bacteriology* **181**, 1944-1946

Heaton, M. P., and Neuhaus, F. C. (1994). Role of the D-alanyl carrier protein in the biosynthesis of D-alanyl-lipoteichoic acid. *Journal of Bacteriology* **176**, 681-690

Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences USA* **89**, 10915-10919

Hirata, R., Ohsumk, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990). Molecular structure of a gene, VMA1, encoding the catalytic subunit of H(+)-translocating adenosine triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **265**, 6726-6733

Hopwood, D. A. (1997). Genetic contributions to understanding polyketide synthases. *Chemical Reviews* **97**, 2465-2498

Horvath, L. A., Sturtevant, J. M., and Prestegard, J. H. (1994). Kinetics and thermodynamics of thermal denaturation in acyl carrier protein. *Protein Science* **3**, 103-108

Hsu, L., Jackowski, S., and Rock, C. O. (1989). Uptake and acylation of 2-acyl-lysophospholipids by *Escherichia coli*. *Journal of Bacteriology* **171**, 1203-1205

Hsu, L., Jackowski, S., and Rock, C. O. (1991). Isolation and characterization of *Escherichia coli* K-12 mutants lacking both 2-acyl-glycerophosphoethanolamine acyltransferase and acyl-acyl carrier protein synthetase activity. *Journal of Biological Chemistry* **266**, 13783-13788

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33-38, 27-38

DeNovaMed Inc. *DeNovaMed: Developing novel antibiotics to fight infections of today and tomorrow.* Web Page: http://www.denovamed.com/index.shtml

Issartel, J. P., Koronakis, V., and Hughes, C. (1991). Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature* **351**, 759-761

Jackman, J. E., Raetz, C. R. H., and Fierke, C. A. (1999). UDP-3-O-(R-3-Hydroxymyristoyl)-*N*-acetylglucosamine deacetylase of *Escherichia coli* is a zinc metalloenzyme. *Biochemistry* **38**, 1902-1911

Jackman, J. E., Raetz, C. R. H., and Fierke, C. A. (2000). Site-directed mutagenesis of the bacterial metalloamidase UDP-(3-O-acyl)-*N*-acetylglucosamine deacetylase (LpxC). Identification of the zinc binding Site. *Biochemistry* **40**, 514-523

Jacquier, V., Prummer, M., Segura, J.-M., Pick, H., and Vogel, H. (2006). Visualizing odorant receptor trafficking in living cells down to the single-molecule level. *Proceedings of the National Academy of Sciences USA* **103**, 14325-14330

Jain, N. U., Wyckoff, T. J., Raetz, C. R., and Prestegard, J. H. (2004). Rapid analysis of large protein-protein complexes using NMR-derived orientational constraints: the 95 kDa complex of LpxA with acyl carrier protein. *Journal of Molecular Biology* **343**, 1379-1389

Jenkins, J., and Pickersgill, R. (2001). The architecture of parallel β-helices and related folds. *Progress in Biophysics and Molecular Biology* **77**, 111-175

Jha, J. K., Sinha, S., Maiti, M. K., Basu, A., Mukhopadhyay, U. K., and Sen, S. K. (2007). Functional expression of an acyl carrier protein (ACP) from *Azospirillum brasilense* alters fatty acid profiles in *Escherichia coli* and *Brassica juncea*. *Plant Physiology and Biochemistry* **45**, 490-500

Jiang, Y., Chan, C. H., and Cronan, J. E. (2006). The soluble acyl-acyl carrier protein synthetase of *Vibrio harveyi* B392 is a member of the medium chain acyl-CoA synthetase family. *Biochemistry* **45**, 10008-10019

Jones, P. J., Cioffi, E. A., and Prestegard, J. H. (1987a). [19F]-1H heteronuclear nuclear Overhauser effect studies of the acyl chain-binding site of acyl carrier protein. *The Journal of Biological Chemistry* **262**, 8963-8965

Jones, P. J., Holak, T. A., and Prestegard, J. H. (1987b). Structural comparison of acyl carrier protein in acylated and sulfhydryl forms by two-dimensional 1H NMR spectroscopy. *Biochemistry* **26**, 3493-3500

Jordan, S. W., and Cronan, J. E. (1997). A new metabolic link: The acyl carrier protein of lipid synthesis donates lipoic acid to the pyruvate dehydrogenase complex in *Escherichia coli* and mitochondria. *Journal of Biological Chemistry* **272**, 17903-17906

Jordan, S. W., and Cronan Jr, J. E. (1997). Biosynthesis of lipoic acid and posttranslational modification with lipoic acid in *Escherichia coli*. in *Methods in Enzymology* (Donald B. McCormick, J. W. S. C. W. ed.), Academic Press. pp 176-183

Kalanon, M., and McFadden, G. I. (2010). Malaria, *Plasmodium falciparum* and its apicoplast. *Biochemical Society Transactions* **38**, 775-782

Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., and Stevens, T. H. (1990). Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science* **250**, 651-657

Keating, M. M., Gong, H., and Byers, D. M. (2002). Identification of a key residue in the conformational stability of acyl carrier protein. *Biochimica et Biophysica Acta* **1601**, 208-214

Kelly, T. M., Stachula, S. A., Raetz, C. R., and Anderson, M. S. (1993). The firA gene of *Escherichia coli* encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine *N*-acyltransferase. The third step of endotoxin biosynthesis. *Journal of Biological Chemistry* **268**, 19866-19874

Khosla, C., and Keasling, J. D. (2003). Metabolic engineering for drug discovery and development. *Nat Rev Drug Discov* **2**, 1019-1025

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research* **37**, D387-392

Kim, Y., and Prestegard, J. H. (1989). A dynamic model for the structure of acyl carrier protein in solution. *Biochemistry* **28**, 8792-8797

Kim, Y., Ohlrogge, J. B., and Prestegard, J. H. (1990). Motional effects on NMR structural data. Comparison of spinach and *Escherichia coli* acyl carrier proteins. *Biochemical Pharmacology* **40**, 7-13

Kim, Y., and Prestegard, J. H. (1990). Refinement of the NMR structures for acyl carrier protein with scalar coupling data. *Proteins* **8**, 377-385

Kim, Y., Kovrigin, E. L., and Eletr, Z. (2006). NMR studies of *Escherichia coli* acyl carrier protein: dynamic and structural differences of the apo- and holo-forms. *Biochemical and Biophysical Research Communications* **341**, 776-783

Kiriukhin, M. Y., and Neuhaus, F. C. (2001). D-Alanylation of lipoteichoic acid: role of the D-alanyl carrier protein in acylation. *Journal of Bacteriology* **183**, 2051-2058

Knowles, J. R. (1989). The mechanism of biotin-dependent enzymes. *Annu Rev Biochem* **58**, 195-221

Konermann, L., and Douglas, D. J. (1997). Acid-induced unfolding of cytochrome c at different methanol concentrations: electrospray ionization mass spectrometry specifically monitors changes in the tertiary structure. *Biochemistry* **36**, 12296-12302

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of Molecular Graphics* **14**, 51-55, 29-32

Kresge, N., Simoni, R. D., and Hill, R. L. (2005). The role of the acyl carrier protein in fatty acid synthesis: the work of P. Roy Vagelos. *Journal of Biological Chemistry* **280**, e32

Kropf, M., Rey, G., Glauser, L., Kulangara, K., Johnsson, K., and Hirling, H. (2008). Subunit-specific surface mobility of differentially labeled AMPA receptor subunits. *European Journal of Cell Biology* **87**, 763-778

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685

Lai, J. R., Koglin, A., and Walsh, C. T. (2006). Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry* **45**, 14869-14879

Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer, New York, NY, USA.

Lambalot, R. H., and Walsh, C. T. (1995). Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *Journal of Biological Chemistry* **270**, 24658-24661

Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. *Chemistry & Biology* **3**, 923-936

Lambalot, R. H., and Walsh, C. T. (1997). Holo-[acyl-carrier-protein] synthase of *Escherichia coli*. *Methods in Enzymology* **279**, 254-262

Lautru, S., and Challis, G. L. (2004). Substrate recognition by nonribosomal peptide synthetase multi-enzymes. *Microbiology* **150**, 1629-1636

Lee, B. I., and Suh, S. W. (2003). Crystal structure of UDP-*N*-acetylglucosamine acyltransferase from *Helicobacter pylori*. *Proteins* **53**, 772-774

Lee, S., Jung, Y., Lee, S., and Lee, J. (2013). Correlations between FAS elongation cycle genes expression and fatty acid production for improvement of long-chain fatty acids in *Escherichia coli*. *Applied Biochemistry and Biotechnology* **169**, 1606-1619

Leibundgut, M., Jenni, S., Frick, C., and Ban, N. (2007). Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase. *Science* **316**, 288-290

Leibundgut, M., Maier, T., Jenni, S., and Ban, N. (2008). The multienzyme architecture of eukaryotic fatty acid synthases. *Current Opinion in Structural Biology* **18**, 714-725

Li, Q., Khosla, C., Puglisi, J. D., and Liu, C. W. (2003). Solution structure and backbone dynamics of the holo form of the frenolicin acyl carrier protein. *Biochemistry* **42**, 4648-4657

Li, Z., and Nair, S. K. (2012). Quorum sensing: How bacteria can coordinate activity and synchronize their response to external signals? *Protein Science* **21**, 1403-1417

Lin, S., Hanson, R. E., and Cronan, J. E. (2010). Biotin synthesis begins by hijacking the fatty acid synthetic pathway. *Nat Chem Biol* **6**, 682-688

Lin, S., and Cronan, J. E. (2012). The BioC O-methyltransferase catalyzes methyl esterification of malonyl-acyl carrier protein, an essential step in biotin synthesis. *Journal of Biological Chemistry* **287**, 37010-37020

Liu, Y., Zheng, T., and Bruner, Steven D. (2011). Structural basis for phosphopantetheinyl carrier domain interactions in the terminal module of nonribosomal peptide synthetases. *Chemistry & Biology* **18**, 1482-1488

Lomakin, I. B., Xiong, Y., and Steitz, T. A. (2007). The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* **129**, 319-332

Lu, Y.-J., Zhang, Y.-M., Grimes, K. D., Qi, J., Lee, R. E., and Rock, C. O. (2006). Acyl-phosphates initiate membrane phospholipid synthesis in Gram-positive pathogens. *Molecular Cell* **23**, 765-772

Lu, Y.-J., Zhang, F., Grimes, K. D., Lee, R. E., and Rock, C. O. (2007). Topology and Active Site of PlsY: The bacterial acylphosphate:glycerol-3-phosphate acyltransferase. *Journal of Biological Chemistry* **282**, 11339-11346

MacKerell, A. D., Bashford, D., Bellott, Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B* **102**, 3586-3616

Maier, T., Jenni, S., and Ban, N. (2006). Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* **311**, 1258-1262

Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chemical Reviews* **97**, 2651-2674

Masoudi, A., Raetz, C. R., Zhou, P., and Pemble, C. W. t. (2014). Chasing acyl carrier protein through a catalytic cycle of lipid A production. *Nature* **505**, 422-426

Mayo, K. H., and Prestegard, J. H. (1985). Acyl carrier protein from *Escherichia coli*. Structural characterization of short-chain acylated acyl carrier proteins by NMR. *Biochemistry* **24**, 7834-7838

McAllister, K. A., Peery, R. B., and Zhao, G. (2006). Acyl carrier protein synthases from gram-negative, gram-positive, and atypical bacterial species: Biochemical and structural properties and physiological implications. *Journal of Bacteriology* **188**, 4737-4748

Meighen, E. A. (1988). Enzymes and genes from the lux operons of bioluminescent bacteria. *Annual Review of Microbiology* **42**, 151-176

Meighen, E. A. (1993). Bacterial bioluminescence: organization, regulation, and application of the lux genes. *The FASEB Journal* **7**, 1016-1022

Metzger, L. E., and Raetz, C. R. H. (2010). An alternative route for UDP-diacylglucosamine hydrolysis in bacterial lipid A biosynthesis. *Biochemistry* **49**, 6715-6726

Meyer, B. H., Segura, J.-M., Martinez, K. L., Hovius, R., George, N., Johnsson, K., and Vogel, H. (2006). FRET imaging reveals that functional neurokinin-1 receptors are monomeric and reside in membrane microdomains of live cells. *Proceedings of the National Academy of Sciences USA* **103**, 2138-2143

Miller, J. R., Busby, R. W., Jordan, S. W., Cheek, J., Henshaw, T. F., Ashley, G. W., Broderick, J. B., Cronan, J. E., and Marletta, M. A. (2000). *Escherichia coli* lipA is a lipoyl synthase: *in vitro* biosynthesis of lipoylated pyruvate dehydrogenase complex from octanoyl-acyl carrier protein. *Biochemistry* **39**, 15166-15178

Modak, R., Sinha, S., and Surolia, N. (2007). Isothermal unfolding studies on the apo and holo forms of *Plasmodium falciparum* acyl carrier protein. *FEBS Journal* **274**, 3313-3326

Mofid, M. R., Finking, R., and Marahiel, M. A. (2002). Recognition of hybrid PCP/ACP acyl carrier proteins in nonribosomal peptide synthetase modules by the 4'-phophopantetheinyl transferases AcpS and Sfp. *Journal of Biological Chemistry*

Moncany, M. L., and Kellenberger, E. (1981). High magnesium content of *Escherichia coli* B. *Experientia* **37**, 846-847

Morikawa, T., Yasuno, R., and Wada, H. (2001). Do mammalian cells synthesize lipoic acid?: Identification of a mouse cDNA encoding a lipoic acid synthase located in mitochondria. *FEBS Letters* **498**, 16-21

Murphy, P. W., Rowland, E. E., and Byers, D. M. (2007). Electrospray ionization mass spectra of acyl carrier protein are insensitive to its solution phase conformation. *Journal of the American Society for Mass Spectrometry* **18**, 1525-1532

NAMD Online Documentation. *Tutorial*. Web Page: http://www.ks.uiuc.edu/Training/Tutorials/namd-index.html#namd

Nealson, K. H., Platt, T., and Hastings, J. W. (1970). Cellular Control of the Synthesis and Activity of the Bacterial Luminescent System. *Journal of Bacteriology* **104**, 313-322

Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**, 443-453

Neuhaus, F. C., and Baddiley, J. (2003). A continuum of anionic charge: Structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews* **67**, 686-723

Nguyen, C., Haushalter, R. W., Lee, D. J., Markwick, P. R., Bruegger, J., Caldara-Festin, G., Finzel, K., Jackson, D. R., Ishikawa, F., O'Dowd, B., McCammon, J. A., Opella, S. J., Tsai, S. C., and Burkart, M. D. (2014). Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **505**, 427-431

Onishi, H. R., Pelak, B. A., Gerckens, L. S., Silver, L. L., Kahan, F. M., Chen, M. H., Patchett, A. A., Galloway, S. M., Hyland, S. A., Anderson, M. S., and Raetz, C. R. (1996). Antibacterial agents that inhibit lipid A biosynthesis. *Science* **274**, 980-982

Paoletti, L., Lu, Y.-J., Schujman, G. E., de Mendoza, D., and Rock, C. O. (2007). Coupling of fatty acid and phospholipid synthesis in *Bacillus subtilis*. *Journal of Bacteriology* **189**, 5816-5824

Park, S. J., Kim, J. S., Son, W. S., and Lee, B. J. (2004). pH-induced conformational transition of *H. pylori* acyl carrier protein: insight into the unfolding of local structure. *Journal of Biochemistry* **135**, 337-346

Parrillo, J. E. (1993). Pathogenetic mechanisms of septic shock. *New England Journal of Medicine* **328**, 1471-1477

Parris, K. D., Lin, L., Tam, A., Mathew, R., Hixon, J., Stahl, M., Fritz, C. C., Seehra, J., and Somers, W. S. (2000). Crystal structures of substrate binding to *Bacillus subtilis* holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites. *Structure* **8**, 883-895

Pearson, J. P., Passador, L., Iglewski, B. H., and Greenberg, E. P. (1995). A second *N*-acylhomoserine lactone signal produced by *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* **92**, 1490-1494

Peitsch, M. C. (1995). Protein modeling by E-mail. *Nat Biotech* **13**, 658-660

Perler, F. B. (2002). InBase: the intein database. *Nucleic acids research* **30**, 383-384

Peschel, A., Vuong, C., Otto, M., and Götz, F. (2000). The D-alanine residues of *Staphylococcus aureus* teichoic acids alter the susceptibility to vancomycin and the activity of autolytic enzymes. *Antimicrobial Agents and Chemotherapy* **44**, 2845-2847

Pesci, E. C., Pearson, J. P., Seed, P. C., and Iglewski, B. H. (1997). Regulation of las and rhl quorum sensing in *Pseudomonas aeruginosa*. *Journal of Bacteriology* **179**, 3127-3132

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera: a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-1612

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26**, 1781-1802

Ploskoń, E., Arthur, C. J., Evans, S. E., Williams, C., Crosby, J., Simpson, T. J., and Crump, M. P. (2008). A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *Journal of Biological Chemistry* **283**, 518-528

Polacco, M. L., and Cronan, J. E. (1981). A mutant of *Escherichia coli* conditionally defective in the synthesis of holo-[acyl carrier protein]. *Journal of Biological Chemistry* **256**, 5750-5754

POV-Ray Online Documentation. *Automatic Bounding Control*. Web Page: http://www.povray.org/documentation/view/3.6.1/223/

POV-Ray Online Documentation. *Adding New Resolutions*. Web Page: http://www.povray.org/documentation/view/3.6.1/223/

Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I., and Sussman, J. L. (2005). FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435-3438

Prummer, M., Meyer, B. H., Franzini, R., Segura, J.-M., George, N., Johnsson, K., and Vogel, H. (2006). Post-translational covalent labeling reveals heterogeneous mobility of individual G protein-coupled receptors in living cells. *ChemBioChem* **7**, 908-911

Python Software Foundation. *Python 2.7*. Web Page: http://www.python.org/download/releases/2.7/

Qu, A., and Leahy, D. J. (1996). The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. *Structure* **4**, 931-942

R-Project. *R Programming Language*. Web Page: http://www.r-project.org/

Radika, K., and Raetz, C. R. (1988). Purification and properties of lipid A disaccharide synthase of *Escherichia coli*. *Journal of Biological Chemistry* **263**, 14859-14867

Raetz, C. R. (1993). Bacterial endotoxins: extraordinary lipids that activate eucaryotic signal transduction. *Journal of Bacteriology* **175**, 5745-5753

Raetz, C. R., and Roderick, S. L. (1995). A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science* **270**, 997-1000

Raetz, C. R. H. (1986). Molecular genetics of membrane phospholipid synthesis. *Annual Review of Genetics* **20**, 253-291

Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007). Lipid A modification systems in Gram-negative bacteria. *Annual Review of Biochemistry* **76**, 295-329

Ray, B. L., Painter, G., and Raetz, C. R. (1984). The biosynthesis of gram-negative endotoxin. Formation of lipid A disaccharides from monosaccharide precursors in extracts of *Escherichia coli*. *Journal of Biological Chemistry* **259**, 4852-4859

RCSB. *Protein Data Bank: PDB File Format Information*. Web Page: http://www.rcsb.org/pdb/static.do?p=file_formats/pdb/index.html

RCSB. *Protein Data Bank*. Web Page: http://pdb.org/pdb/home/home.do

Reed, M. A., Schweizer, M., Szafranska, A. E., Arthur, C., Nicholson, T. P., Cox, R. J., Crosby, J., Crump, M. P., and Simpson, T. J. (2003). The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs. *Organic & Biomolecular Chemistry* **1**, 463-471

Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999a). Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *The EMBO Journal* **18**, 6823-6831

Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999b). Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4′-phosphopantetheinyl transferase superfamily. *The EMBO Journal* **18**, 6823-6831

Riahi, S., Roux, B., and Rowley, C. N. (2013). QM/MM molecular dynamics simulations of the hydration of Mg(II) and Zn(II) ions. *Canadian Journal of Chemistry* **91**, 552-558

Rietschel, E. T., Kirikae, T., Schade, F. U., Ulmer, A. J., Holst, O., Brade, H., Schmidt, G., Mamat, U., Grimmecke, H. D., Kusumoto, S., and et al. (1993). The chemical structure of bacterial endotoxin in relation to bioactivity. *Immunobiology* **187**, 169-190

Robins, L. I., Williams, A. H., and Raetz, C. R. H. (2009). Structural basis for the sugar nucleotide and acyl-chain selectivity of *Leptospira interrogans* lpxA. *Biochemistry* **48**, 6191-6201

Rock, C. O., and Cronan, J. E., Jr. (1979). Re-evaluation of the solution structure of acyl carrier protein. *The Journal of Biological Chemistry* **254**, 9778-9785

Rock, C. O., and Garwin, J. L. (1979). Preparative enzymatic synthesis and hydrophobic chromatography of acyl-acyl carrier protein. *The Journal of Biological Chemistry* **254**, 7123-7128

Rock, C. O., Cronan, J. E., and Armitage, I. M. (1981). Molecular properties of acyl carrier protein derivatives. *Journal of Biological Chemistry* **256**, 2669-2674

Rock, C. O., and Jackowski, S. (1982). Regulation of phospholipid synthesis in *Escherichia coli*. Composition of the acyl-acyl carrier protein pool *in vivo*. *Journal of Biological Chemistry* **257**, 10759-10765

Rock, C. O. (1983). Environment of the aromatic chromophores of acyl carrier protein. *Archives of Biochemistry and Biophysics* **225**, 122-129

Rock, C. O. (1984). Turnover of fatty acids in the 1-position of phosphatidylethanolamine in *Escherichia coli*. *Journal of Biological Chemistry* **259**, 6188-6194

Rock, C. O., and Cronan, J. E. (1996). *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochimica et Biophysica Acta* **1302**, 1-16

Rock, C. O., and Jackowski, S. (2002). Forty years of bacterial fatty acid synthesis. *Biochemical and Biophysical Research Communications* **292**, 1155-1166

Rodionov, D. A., Mironov, A. A., and Gelfand, M. S. (2002). Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. *Genome Research* **12**, 1507-1516

Rodriguez, A., Wall, L., Riendeau, D., and Meighen, E. (1983). Fatty acid acylation of proteins in bioluminescent bacteria. *Biochemistry* **22**, 5604-5611

Rost, B. (1997). Protein structures sustain evolutionary drift. *Folding & Design* **2**, S19-24

Roujeinikova, A., Baldock, C., Simon, W. J., Gilroy, J., Baker, P. J., Stuitje, A. R., Rice, D. W., Slabas, A. R., and Rafferty, J. B. (2002). X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. *Structure* **10**, 825-835

Roujeinikova, A., Simon, W. J., Gilroy, J., Rice, D. W., Rafferty, J. B., and Slabas, A. R. (2007). Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *Journal of Molecular Biology* **365**, 135-145

Schmid, K. M., and Ohlrogge, J. B. (2002). *Biochemistry of lipids, lipoproteins and membranes (Chapter title: Lipid metabolism in plants.)*, Elsevier, Amsterdam. 93 – 126.

Schulz, H. (1975). On the structure-function relationship of acyl carrier protein of *Escherichia coli*. *The Journal of Biological Chemistry* **250**, 2299-2304

Schweizer, E., and Hofmann, J. (2004). Microbial type I fatty acid synthases (FAS): Major players in a network of cellular FAS systems. *Microbiology and Molecular Biology Reviews* **68**, 501-517

Sharma, A. K., Sharma, S. K., Surolia, A., Surolia, N., and Sarma, S. P. (2006). Solution structures of conformationally equilibrium forms of holo-acyl carrier protein (PfACP) from *Plasmodium falciparum* provides insight into the mechanism of activation of ACPs. *Biochemistry* **45**, 6904-6916

Shen, B., and Kwon, H. J. (2002). Macrotetrolide biosynthesis: a novel type II polyketide synthase. *Chemical record (New York, N.Y.)* **2**, 389-396

Shen, Z., and Byers, D. M. (1994). Exogenous myristic acid can be partially degraded prior to activation to form acyl-acyl carrier protein intermediates and lipid A in *Vibrio harveyi*. *Journal of Bacteriology* **176**, 77-83

Shen, Z., and Byers, D. M. (1996). Isolation of *Vibrio harveyi* acyl carrier protein and the fabG, acpP, and fabF genes involved in fatty acid biosynthesis. *Journal of Bacteriology* **178**, 571-573

Six, D. A., Carty, S. M., Guan, Z., and Raetz, C. R. H. (2008). Purification and mutagenesis of LpxL, the lauroyltransferase of *Escherichia coli* lipid A biosynthesis. *Biochemistry* **47**, 8623-8637

Smith, S., Witkowski, A., and Joshi, A. K. (2003). Structural and functional organization of the animal fatty acid synthase. *Progress in Lipid Research* **42**, 289-317

Spanish National Research Council. *MOLMOL Tutorial - Example 3*. Web Page: http://www.wwpdb.org/documentation/format33/remarks2.html

Stanley, P., Packman, L. C., Koronakis, V., and Hughes, C. (1994). Fatty acylation of two internal lysine residues required for the toxic activity *Escherichia coli* hemolysin. *Science* **266**, 1992-1996

Stanley, P., Hyland, C., Koronakis, V., and Hughes, C. (1999). An ordered reaction mechanism for bacterial toxin acylation by the specialized acyltransferase HlyC: formation of a ternary complex with acylACP and protoxin substrates. *Molecular Microbiology* **34**, 887-901

Staunton, J., and Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **18**, 380-416

Stratagene. *QuikChange® II XL Site-Directed Mutagenesis kit manual*. Web Page: http://www.chem.agilent.com/library/usermanuals/Public/200521.pdf

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130

Sweet, C. R., Lin, S., Cotter, R. J., and Raetz, C. R. (2001). A *Chlamydia trachomatis* UDP-*N*-acetylglucosamine acyltransferase selective for myristoyl-acyl carrier protein. Expression in *Escherichia coli* and formation of hybrid lipid A species. *The Journal of Biological Chemistry* **276**, 19565-19574

Swiss Institute for Bioinformatics. *DeepView Tutorial: Homology Modelling*. Web Page: http://spdbv.vital-it.ch/modeling_tut.html

Swiss Institute for Bioinformatics. *Swiss-PdbViewer: DeepView*. Web Page: http://spdbv.vital-it.ch/

Swiss Institute for Bioinformatics. *SWISS-MODEL*. Web Page: http://swissmodel.expasy.org/

Tami, G., Murphy, P. W., and Byers, D. M. (2011). Characterization of *Plasmodium falciparum* holo-ACP synthase: a potential antimalarial drug target. Experimental Biology 2011 (ASBMB; Washington, D.C., USA).

Tanovic, A., Samel, S. A., Essen, L.-O., and Marahiel, M. A. (2008). Crystal structure of the termination module of a nonribosomal reptide synthetase. *Science* **321**, 659-663

Tener, D. M., and Mayo, K. H. (1990). Divalent cation binding to reduced and octanoyl acyl-carrier protein. *European Journal of Biochemistry / FEBS* **189**, 559-565

TgK Scientific. *Rapid Mixing Devices*. Web Page: http://www.hi-techsci.com/products/rapidmixing/

Therisod, H., Weissborn, A. C., and Kennedy, E. P. (1986). An essential function for acyl carrier protein in the biosynthesis of membrane-derived oligosaccharides of *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **83**, 7236-7240

Therisod, H., and Kennedy, E. P. (1987). The function of acyl carrier protein in the synthesis of membrane-derived oligosaccharides does not require its phosphopantetheine prosthetic group. *Proceedings of the National Academy of Sciences USA* **84**, 8235-8238

Thomas, J., and Cronan, J. E. (2005). The enigmatic acyl carrier protein phosphodiesterase of *Escherichia coli*: genetic and enzymological characterization. *The Journal of Biological Chemistry* **280**, 34675-34683

Thomas, J., Rigden, D. J., and Cronan, J. E. (2007). Acyl carrier protein phosphodiesterase (AcpH) of *Escherichia coli* is a non-canonical member of the HD phosphatase/phosphodiesterase family. *Biochemistry* **46**, 129-136

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences* **27**, 527-533

Ugulava, N. B., Frederick, K. K., and Jarrett, J. T. (2003). Control of adenosylmethionine-dependent radical generation in biotin synthase: a kinetic and thermodynamic analysis of substrate binding to active and inactive forms of BioB. *Biochemistry* **42**, 2708-2719

Ulaganathan, V., Buetow, L., and Hunter, W. N. (2007). Nucleotide substrate recognition by UDP-*N*-acetylglucosamine acyltransferase (LpxA) in the first step of lipid A biosynthesis. *Journal of Molecular Biology* **369**, 305-312

Ulevitch, R. J., and Tobias, P. S. (1995). Receptor-dependent mechanisms of cell stimulation by bacterial endotoxin. *Annual Review of Immunology* **13**, 437-457

Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Science* **11**, 739-756

Vaara, M. (1992). Eight bacterial proteins, including UDP-*N*-acetylglucosamine acyltransferase (LpxA) and three other transferases of *Escherichia coli*, consist of a six-residue periodicity theme. *FEMS Microbiology Letters* **76**, 249-254

Vaara, M. (1993). Antibiotic-supersusceptible mutants of *Escherichia coli* and *Salmonella typhimurium*. *Antimicrobial Agents and Chemotherapy* **37**, 2255-2260

Vagelos, P. R., and Larrabes, A. R. (1967). Acyl carrier protein. IX. Acyl carrier protein hydrolase. *The Journal of Biological Chemistry* **242**, 1776-1781

Vanaman, T. C., Wakil, S. J., and Hill, R. L. (1968). The complete amino acid sequence of the acyl carrier protein of *Escherichia coli*. *Journal of Biological Chemistry* **243**, 6420-6431

VanScyoc, W. S., Sorensen, B. R., Rusinova, E., Laws, W. R., Ross, J. B., and Shea, M. A. (2002). Calcium binding to calmodulin mutants monitored by domain-specific intrinsic phenylalanine and tyrosine fluorescence. *Biophysical Journal* **83**, 2767-2780

Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., Wasney, G. A., Gao, M., Hills, T., Brokx, S., Qiu, W., Sharma, S., Diassiti, A., Alam, Z., Melone, M., Mulichak, A., Wernimont, A., Bray, J., Loppnau, P., Plotnikova, O., Newberry, K., Sundararajan, E., Houston, S., Walker, J., Tempel, W., Bochkarev, A., Kozieradzki, I., Edwards, A., Arrowsmith, C., Roos, D., Kain, K., and Hui, R. (2007). Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. *Molecular and Biochemical Parasitology* **151**, 100-110

Vivero-Pol, L., George, N., Krumm, H., Johnsson, K., and Johnsson, N. (2005). Multicolor imaging of cell surface proteins. *Journal of the American Chemical Society* **127**, 12770-12771

VMD Online Documentation. *Tutorials*. Web Page: http://www.ks.uiuc.edu/Research/vmd/current/docs.html#tutorials

VMD Online Documentation. *Molecular Representations in VMD*. Web Page: http://www.ks.uiuc.edu/Research/vmd/allversions/repimages/#representations

VMD Downloads. *VMD Script Library*. Web Page: http://www.ks.uiuc.edu/Research/vmd/script_library/

Volkman, B. F., Zhang, Q., Debabov, D. V., Rivera, E., Kresheck, G. C., and Neuhaus, F. C. (2001). Biosynthesis of D-alanyl-lipoteichoic acid: the tertiary structure of apo-D-alanyl carrier protein. *Biochemistry* **40**, 7964-7972

Volkmann, G. (2009). *Protein Clevage, Labeling and Cyclization Using Non-Canonical Split-Inteins.* Ph.D., Dalhousie University

Volkmann, G., and Iwai, H. (2010). Protein trans-splicing and its use in structural biology: opportunities and limitations. *Molecular BioSystems* **6**, 2110-2121

Volkmann, G., Murphy, P. W., Rowland, E. E., Cronan, J. E., Jr., Liu, X. Q., Blouin, C., and Byers, D. M. (2010). Intein-mediated cyclization of bacterial acyl carrier protein stabilizes its folded conformation but does not abolish function. *The Journal of Biological Chemistry* **285**, 8605-8614

Volkmann, G., and Mootz, H. D. (2013). Recent progress in intein research: from mechanism to directed evolution and applications. *Cellular and Molecular Life sciences* **70**, 1185-1206

Vorachek-Warren, M. K., Ramirez, S., Cotter, R. J., and Raetz, C. R. H. (2002). A triple mutant of *Escherichia coli* lacking secondary acyl chains on lipid A. *Journal of Biological Chemistry* **277**, 14194-14205

Wakil, S. J. (1989). Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**, 4523-4530

Wall, L., Rodriguez, A., and Meighen, E. (1986). Intersubunit transfer of fatty acyl groups during fatty acid reduction. *Journal of Biological Chemistry* **261**, 15981-15988

Warren, H. S., Danner, R. L., and Munford, R. S. (1992). Anti-endotoxin monoclonal antibodies. *New England Journal of Medicine* **326**, 1153-1157

Watt, S. J., Sheil, M. M., Beck, J. L., Prosselkov, P., Otting, G., and Dixon, N. E. (2007). Effect of protein stabilization on charge state distribution in positive- and negative-ion electrospray ionization mass spectra. *Journal of the American Society for Mass Spectrometry* **18**, 1605-1611

Weber, T., and Marahiel, M. A. (2001). Exploring the domain structure of modular nonribosomal peptide synthetases. *Structure* **9**, R3-R9

Webster, S. P., Alexeev, D., Campopiano, D. J., Watt, R. M., Alexeeva, M., Sawyer, L., and Baxter, R. L. (2000). Mechanism of 8-amino-7-oxononanoate synthase: spectroscopic, kinetic, and crystallographic studies. *Biochemistry* **39**, 516-528

White, S. W., Zheng, J., Zhang, Y.-M., and Rock, C. O. (2005). The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry* **74**, 791-831

Williams, A. H., Immormino, R. M., Gewirth, D. T., and Raetz, C. R. (2006). Structure of UDP-*N*-acetylglucosamine acyltransferase with a bound antibacterial pentadecapeptide. *Proceedings of the National Academy of Sciences USA* **103**, 10877-10882

Williams, A. H., and Raetz, C. R. (2007). Structural basis for the acyl chain selectivity and mechanism of UDP-N-acetylglucosamine acyltransferase. *Proceedings of the National Academy of Sciences USA* **104**, 13543-13550

Wingware. *Wing IDE: The Intelligent Developemtn Envrionment for Python Programmers*. Web Page: http://www.wingware.com/

Winson, M. K., Camara, M., Latifi, A., Foglino, M., Chhabra, S. R., Daykin, M., Bally, M., Chapon, V., Salmond, G. P., and Bycroft, B. W. (1995). Multiple *N*-acyl-L-homoserine lactone signal molecules regulate production of virulence determinants and secondary metabolites in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences USA* **92**, 9427-9431

Wong, H. C., Liu, G., Zhang, Y. M., Rock, C. O., and Zheng, J. (2002). The solution structure of acyl carrier protein from *Mycobacterium tuberculosis*. *The Journal of Biological Chemistry* **277**, 15874-15880

Worldwide PDB. *PDB File Format Information: Remarks 300 - 999*. Web Page: http://www.wwpdb.org/documentation/format33/remarks2.html

Wright, P. E., and Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **293**, 321-331

Wu, B. N., Zhang, Y. M., Rock, C. O., and Zheng, J. J. (2009). Structural modification of acyl carrier protein by butyryl group. *Protein Science* **18**, 240-246

Wyckoff, T. J., Lin, S., Cotter, R. J., Dotson, G. D., and Raetz, C. R. (1998a). Hydrocarbon rulers in UDP-*N*-acetylglucosamine acyltransferases. *The Journal of Biological Chemistry* **273**, 32369-32372

Wyckoff, T. J., Raetz, C. R., and Jackman, J. E. (1998b). Antibacterial and anti-inflammatory agents that target endotoxin. *Trends in Microbiology* **6**, 154-159

Wyckoff, T. J., and Raetz, C. R. (1999). The active site of *Escherichia coli* UDP-*N*-acetylglucosamine acyltransferase. Chemical modification and site-directed mutagenesis. *The Journal of Biological Chemistry* **274**, 27047-27055

Xu, G. Y., Tam, A., Lin, L., Hixon, J., Fritz, C. C., and Powers, R. (2001). Solution structure of *B. subtilis* acyl carrier protein. *Structure* **9**, 277-287

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F., Goddard, T. D., Meng, E. C., Sali, A., and Ferrin, T. E. (2012). UCSF Chimera, Modeller, and IMP: an integrated modeling system. *Journal of Structural Biology* **179**, 269-278

Yao, J., and Rock, C. O. (2013). Phosphatidic acid synthesis in bacteria. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1831**, 495-502

Yin, J., Lin, A. J., Buckett, P. D., Wessling-Resnick, M., Golan, D. E., and Walsh, C. T. (2005). Single-cell FRET imaging of transferrin receptor trafficking dynamics by Sfp-catalyzed, site-specific protein labeling. *Chemistry & Biology* **12**, 999-1006

Yoshimura, M., Oshima, T., and Ogasawara, N. (2007). Involvement of the YneS/YgiH and PlsX proteins in phospholipid biosynthesis in both *Bacillus subtilis* and *Escherichia coli*. *BMC Microbiology* **7**, 69

Young, K., Silver, L. L., Bramhill, D., Cameron, P., Eveland, S. S., Raetz, C. R. H., Hyland, S. A., and Anderson, M. S. (1995). The envA permeability/cell division gene of *Escherichia coli* encodes the second enzyme of lipid A biosynthesis: UDP-3-O-(R-3-hydroxymyristoly)-*N*-Acetylglucosamine deacetylase. *Journal of Biological Chemistry* **270**, 30384-30391

Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **97**, 5978-5983

Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001). Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. *The Journal of Biological Chemistry* **276**, 8231-8238

Zhang, Y. M., Marrakchi, H., White, S. W., and Rock, C. O. (2003a). The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. *Journal of Lipid Research* **44**, 1-10

Zhang, Y. M., Wu, B., Zheng, J., and Rock, C. O. (2003b). Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. *The Journal of Biological Chemistry* **278**, 52935-52943

Zornetzer, G. A., Fox, B. G., and Markley, J. L. (2006). Solution structures of spinach acyl carrier protein with decanoate and stearate. *Biochemistry* **45**, 5217-5227

Zornetzer, G. A., Tanem, J., Fox, B. G., and Markley, J. L. (2010). The length of the bound fatty acid influences the dynamics of the acyl carrier protein and the stability of the thioester bond. *Biochemistry* **49**, 470-477

APPENDIX 1 Aligning Monomeric Proteins Using Chimera

UCSF Chimera version 1.7 (build 38197) was used for all structure alignments. Alignments were done for WT and mutant ACPs (post-mutation and simulation), WT and mutant ACPs (post-metal addition and simulation), WT and mutant LpxAs, and WT and mutant AcpSs. The process described below is for WT and mutant ACPs, but the same procedure would also apply for LpxAs and AcpSs (with the WT structures serving as the reference structure to align the mutated structures).

In the UCSF Chimera window, the structures to be aligned were loaded (File/Open and browse to the location of equilibrated structures), and the following protocol can be used. Select the equilibrated structures (PDB) and click Open. The structures will now open in the main Chimera window. To align the different proteins based on structure, Chimera's built-in "MatchMaker" (MM) tool (Tools/Structure Comparison/MatchMaker) was used. Before MM can align the proteins, a reference structure must be set. For all ACP alignments, rACP was used as the reference structure (selected under "Reference structure:" portion of the MM window). Once the reference structure is set, ensure all other structures are selected in the "Structure(s) to match:" portion of the window. Other options can be set (including "Chain pairing", "Alignment algorithm" and the Matrix used for the alignment) depending on the structures loaded, etc. For all ACP alignments, "Chain pairing" was set to "Best-aligning pair of chains between reference and match structure", the alignment was calculated using the Needleman-Wunsch algorithm (Needleman & Wunsch, 1970) and the BLOSUM-62 matrix (Henikoff & Henikoff, 1992). All other settings were set to the default options. Click "OK" and the structures are aligned in the main Chimera window. To save these structures for further simulation or analysis, click File/Save PDB. Each PDB file should be saved individually (specified in the "Save models:" portion of the window. To ensure the saved PDB are aligned correctly (as on screen), save each model (consecutively) relative to your reference structure (set by clocking the checkbox beside "Save relative to model:" and selecting the reference model used in the alignment; in this case rACP.

Note: The PDB files saved at the end of the above process are saved in another directory and/or with a different name so the original structures aren't overwritten.

APPENDIX 2 Python Program for Calculating Iterative K_D Value Based on Area Under the Curve Data

This Python program calculates a K_D value based on AUC input files can be found below (see Appendix 3 for sample input/output files). Colors are indicative of syntax (comments, strings, constants, normal text, etc) in Wing IDE Personal 4.0.3-1 (Wingware). Program written using Python 2.7 (Python Software Foundation. *Python 2.7*.).

Python program:

```
#!C:\python2.7\python.exe
'''
  Process fluorescence area under the curve (AUC) data from LpxA titrated with ACP
experiments, determine iterative Kd, and write the output to a file

'''
# Import Necessary Python Libraries
import sys
import numpy as np
from scipy.optimize import leastsq

# Define Functions
def computedB(params, S):

    # Compute B from Parameters and S Values
    Bmx = params[0]
    Kd  = params[1]
    return (Bmx*S) / (Kd+S)

def residuals(params, experimentalData, S):
    a = experimentalData - computedB(params,S)
    print a
    return a

# Variables
converged = False

# Read Experiment Filesnames
fin = open("input_names.csv", "r")

# Read in Experiment Dates from Above File
exptdate = fin.readline().strip().split(',')[1:]
exptfile = fin.readline().strip().split(',')[1:]
# Read Experiment Date and Description from Blank Corrected Data File
```

```
# NOTE: ".strip("\"").strip()" command from Javier Alfaro (fellow graduate student,
biochemistry department)
fout = open("results.csv", "w")
for Expt_File_i in exptfile:

    # Remove "\r" Characters
    datain = open(Expt_File_i, "r").read()
    fout = open(Expt_File_i, "w")
    fout.write(datain.replace('\r',''))
    fout.close()

    converged = False
    fin = open(Expt_File_i, "r")
    Expt_date = fin.readline().strip().strip(',')
    Expt_desc = fin.readline().strip().strip(',')

    # Skip Two Lines in the Input File
    fin.readline()
    fin.readline()

    #Create Empty Lists for Area Under the Curve Data and ACP Concentrations
    areas = []
    acpconc = []

    # Read in ACP Concentrations and Areas
    for line in fin:
        line = line.split(',')
        acpconc.append(pet(line[0]))
        areas.append(float(line[1]))

# Fitting Data to "B" Using Numpy Library
B = np.array(areas,dtype='float')
S = np.array(acpconc,dtype='float')

# Set Initial Parameters
Kd = Kdold = 0.1
Bmx = Bmxold = 10000.0
xparams = [Bmx, Kd]
tolerance = 0.000000001
while not converged:

    # Optimize Kd and Bmax (Using LeastSquare From Scipy)
    best = leastsq(residuals, xparams, args=(B,S), full_output=1)
    Bmx = best[0][0]
    Kd  = best[0][1]
    # Recompute Analyte (Free)
```

```
    Bi = (Bmx*S)/(Kd+S)
    Ba = Bi/Bmx
    S = S - (Ba*0.001)

    # Has Convergence Been Reached (IE: If both Bmx/Bmxold and Kd/Kdold are within
      0.000000001 of each other)?
    if (Bmx - Bmxold)/Bmxold <= tolerance and (Kd - Kdold)/Kdold <= tolerance:
        converged = True

        # If Convergence Has Been Reached, Save Results to File (Experiment
          Date/Description and Kd) (help from Javier Alfaro)
        fout.write("Experiment Date:, %s,Experiment Description:, %s, KD:,
                %s\n"%(Expt_date, Expt_desc, Kd))
    else:

        # If Convergence Has Not Been Reached, Store the Old Results and Continue
Calculation
        Bmxold = Bmx
        Kdold = Kd

# Close File After Above Convergence Calculation Is Complete for All Days in
"input_names.csv"
fout.close()
```

Script Execution:

Open a command prompt and navigate to the location of the data and script. Execute the script as below:

"python AUC Input Kd Calculation Script.py"

Ensure "input_names.csv" and the CSV files referenced within are in the same directory as the python program file. Sample input and output files can be found below (see Appendix 3).

Note: For the "python" command to work, the directory containing the python executable file must be added to the "PATH" System Environmental Variable (see Appendix 29 for more information on how to do this in windows).

APPENDIX 3 Sample Input and Output Files for Python Program in Appendix 2

Sample Input Files:

Experiment Dates (input_names.csv):

Date,July 29 2007
Filename,2007-07-29_F162W_h-D30N+UDP.csv

Data File (2007-07-29_F162W_h-D30N+UDP.csv):

July 29 2007,
Q104W Titrated with h-D30N +UDP,
[D30N] (uM),NaP Buffer (7.0 pH)
,AUC
0,0
2.5,230
5,329
10,117
15,300
20,402
25,600
30,451
35,573

Sample Output File (results.csv):

Experiment Date:, July 29 2007,Experiment Description:, Q104W Titrated with h-D30N +UDP, KD:, 17.4485475242

APPENDIX 4 Python Program for Calculating Iterative K_D Value Based on Peak Fluorescence Intensity Data

This Python program calculates a K_D value based on the intensity of the fluorescence peak for titration curve data. Input files can be found below (see Appendix 5 for sample input/output files). Colors are indicative of syntax (comments, strings, constants, normal text, etc) in Wing IDE Personal 4.0.3-1. Program written using Python 2.7.

Python program:

```
#!C:\python2.7\python.exe
'''
  Process peak fluorescence intensity data from ACP titrated with metals experiments,
determine iterative Kd, and write the output to a file
'''
# Import Necessary Python Libraries
import sys
import numpy as np
from scipy.optimize import leastsq

# Define Functions
def computedB(params, S):

    # Compute B From Parameters and S Values
    Bmx = params[0]
    Kd  = params[1]
    return (Bmx*S) / (Kd+S)

def residuals(params, experimentalData, S):
    a = experimentalData - computedB(params,S)
    print a
    return a

# Variables
converged = False

# Read Experiment Filesnames
fin = open("input_names.csv", "r")

# Read in Experiment Dates From Above File
exptdate = fin.readline().strip().split(',')[1:]
exptfile = fin.readline().strip().split(',')[1:]
```

```
# Read Experiment Date and Description from Blank Corrected Data File
# NOTE: ".strip("\"").strip()" command from Javier Alfaro (fellow graduate student,
biochemistry department)
fout = open("results.csv", "w")
for Expt_File_i in exptfile:

    # Remove "\r" Characters
    datain = open(Expt_File_i, "r").read()
    fout = open(Expt_File_i, "w")
    fout.write(datain.replace('\r',''))
    fout.close()

    converged = False
    fin = open(Expt_File_i, "r")
    Expt_date = fin.readline().strip().strip(',')
    Expt_desc = fin.readline().strip().strip(',')

    # Skip Two Lines in the Input File
    fin.readline()
    fin.readline()

    #Create Empty Lists for Peak Fluorescence Intensity and Metal Concentrations
    intensity = []
    meconc = []

    # Read in Metal Concentrations and Intensities
    for line in fin:
        line = line.split(',')
        meconc.append(pet(line[0]))
        intensity.append(float(line[1]))

# Fitting Data to "B" Using Numpy Library
B = np.array(intensity,dtype='float')
S = np.array(meconc,dtype='float')

# Set Initial Parameters
Kd = Kdold = 0.1
Bmx = Bmxold = 10000.0
xparams = [Bmx, Kd]
tolerance = 0.000000001
while not converged:

    # Optimize Kd and Bmax (Using LeastSquare From Scipy)
    best = leastsq(residuals, xparams, args=(B,S), full_output=1)
    Bmx = best[0][0]
    Kd  = best[0][1]
```

```
    # Recompute Analyte (Free)
    Bi = (Bmx*S)/(Kd+S)
    Ba = Bi/Bmx
    S = S - (Ba*0.001)

    # Has Convergence Been Reached (IE: If both Bmx/Bmxold and Kd/Kdold are within
      0.000000001 of each other)?
    if (Bmx - Bmxold)/Bmxold <= tolerance and (Kd - Kdold)/Kdold <= tolerance:
        converged = True

        # If Convergence Has Been Reached, Save Results to File (Experiment
          Date/Description and Kd) (help from Javier Alfaro)
        fout.write("Experiment Date:, %s,Experiment Description:, %s, KD:,
                %s\n"%(Expt_date, Expt_desc, Kd))
    else:

        # If Convergence Has Not Been Reached, Store the Old Results and Continue
Calculation
        Bmxold = Bmx
        Kdold = Kd

# Close File After Above Convergence Calculation is Complete For All Days in
"input_names.csv"
fout.close()
```

Script Execution:

Open a command prompt and navigate to the location of the data and script. Execute the script as below:

"python AUC Input Kd Calculation Script.py"

Ensure "input_names.csv" and the CSV files referenced within are in the same directory as the python program file. Sample input and output files can be found below (see Appendix 3).

Note: For the "python" command to work, the directory containing the python executable file must be added to the "PATH" System Environmental Variable (see Appendix 29 for more information on how to do this in windows).

APPENDIX 5 Sample Input and Output Files for Python Program in Appendix 4

Sample Input Files:

Experiment Dates (input_names.csv):

ACP-Me,SA-Ni
Filename,SA-Ni.csv

Data File (2007-07-29_F162W_h-D30N+UDP.csv):

Overall Normalized Data,
SA Titrated with Ni,
[Ni] (uM),NaP Buffer (7.0 pH)
,F Peak
0,0
2.49,3.39744
4.95,5.32015
9.8,7.12462
14.96,10.86399
20.19,12.96794
30.27,15.09595
40.15,14.53413
49.83,15.75705
60.03,18.35775
80.49,18.65989
100.07,18.7569

Sample Output File (results.csv):

Experiment Date:, Overall,Experiment Description:, SA Titrated with Ni, KD:,
16.2421110918

APPENDIX 6 Mutating a Residue *In Silico* Using VMD

All mutations were completed using the "Mutate Residue" functionality of VMD. This functionality is found under "Extensions/Modeling/Mutate Residue" in the main VMD window. All mutations were made using VMD (1.8.7 or 1.9.1).

To mutate a residue, the following protocol can be used:

In the "Mutator" window, under "Input" the PSF and PDB to be mutated are specified (e.g. V.H.-rACP). "MUTATED" is changed to the output name (e.g. V.H.-rACP-D30N). Next, the target residue is specified (under "Target Residue"). When specifying the residue, the VMD residue id (resid) must be used (usually 1 less than the numbering in the PDB file as the first residue is 0, not 1 as in the PDB file). In the case of mutations of ACP to make rACP, SA, SB, and SA/SB, the resid is actually 4 more than the mutation to be made (e.g. D30's resid is 34) as the numbering is based on wild-type ACP and not rACP (which has 4 extra residues at the beginning of the protein left after Fx_A cleavage of the GST tag). For all *in silico* mutations, Free Energy Perturbation (FEP) files were not generated. After all settings are correct, the "Run Mutator" button was used to start the *in silico* mutation process. Upon completion, all structures were removed and reloaded (using the PDB/PSF generated by "Mutator") and the mutation verified using "Graphical Representations". For more information on using "Graphical Representations" see Section 1.3 of the VMD tutorial ((VMD Online Documentation. *Tutorials*.). As long as the correct resid is specified, the mutation should be successful. As suggested by VMD's "Mutator", all structures were minimized/equilibrated in a water sphere (see Appendix 16 for a sample configuration file and Appendix 9 for details on solvating a protein).

APPENDIX 7 Using VMD to Create Protein Structure Files (PSF)

All PSF files were created using VMD (1.8.7 or 1.9.1) using the "Automatic PSF Builder" (PSFGEN; found under Extensions/Modeling/Automatic PSF Builder in the VMD Main window).

To create a PSF file, the following protocol can be used:

After loading the proper PDB file, open the "Automatic PSF Builder". In the AutoPSF window, select the model you want to generate a PSF file for (the PDB that was loaded) and set an output name (in Step 1 portion of window). The "Topology files" section of Step 1 is where any extra topology files would be loaded (e.g. if the loaded PDB has non-standard amino acids, etc). For most proteins, the default topology file is sufficient. After the input/output files have been specified, click "Load input files". In the Step 2 section, specify what will be included in the final PDB/PSF files. For all structure files processed with PSFGEN, "Everything" was selected. Once the selection is set, click "Guess and split chains using current selection". When this process is complete the segments identified in Step 2 will appear in Step 3. If there are errors in the splitting process, they can be edited. For most proteins this should be done correctly by PSFGEN. When correct, click "Create chains. In Step 4 specify any patches necessary and click "Apply patches and finish PDB/PSF. For all PDBs in this work, patches were not necessary. After processing, PSFGEN will write PDB and PSF files to the working directory for the VMD session.

APPENDIX 8 Creation of Biologically Active Form of a Protein using Mono2poly and VMD

LpxA and AcpS trimers were made using VMD (1.8.7 or 1.9.1) and the mono2poly script which is readily available in the VMD script library (VMD Downloads. *VMD Script Library*.). Text after # explains what function the command performs.

Mono2poly.tcl contents:

```
proc parsematrix {orig_file}  {
  set infile [open  $orig_file r]
  set mtnum 0
  while {[gets $infile line]>=0} {
    set title [lindex $line 0]
    set biomt [lindex $line 2]
    if { $title == "REMARK" && [string match BIOMT? $biomt] } {
      set linenum [string index $biomt  end]
      if {$linenum == 1} {
        incr mtnum 1
        set matri($mtnum) {}
      }
      set lineelement [lrange $line 4 7]
      lappend matrix($mtnum) $lineelement
    }
  }
  close $infile

  set matrixlist {}
  foreach name [lsort -integer [array names matri]] {
    lappend matrix($name) {0.0 0.0 0.0 1.0}
    lappend matrixlist $matrix($name)
  }

  if {[llength matrixlist]==0} {
    error "There is no BIOMT REMARK information in this pdb file"
  }

  puts "this is a $mtnum polymer"
  return $matrixlist
}


proc mono2poly_Usage {} {
  puts "you should input:"
  puts "mono2ploy -chain chainlist -o outfilename sel matrix"
```

```
  error ""
}


proc mono2poly {args} {
  set cmdlinelength [llength $args]
  if {$cmdlinelength!=2 && $cmdlinelength<4} {
    mono2poly_Usage
  }
  set sel [lindex $args [expr $cmdlinelength-2]]
  set matrix [lindex $args end]
  set polynum [llength $matrix]
  puts $polynum
  set output poly
  set chainlist {}
  if {$cmdlinelength>=4} {
    set args [lrange $args 0 [expr $cmdlinelength-3]]
    set i 0
    while {$i<[llength $args]} {
      set opt [lindex $args $i]
      switch -exact -- $opt {
        -o {
          incr i
          set output [lindex $args $i]
          incr i
          continue
        }
        -chain {

          incr i
          set chainlist [lindex $args $i]
          if { [llength $chainlist] != $polynum && [llength $chainlist]!=0} {
            error "chain list should have same element number with matrix or empty list"
          }
          incr i
          continue
        }
        default {
          error "Unknown option $opt"
        }
      }
    }
  }
  $sel writepdb ${output}.org.pdb
  set out [open ${output}.pdb w]
```

```
  set i 0
  foreach mat $matrix {
    mol new ${output}.org.pdb waitfor all
    set mono [atomselect top all]
    $mono move $mat
    if { [llength $chainlist] == $polynum } {
      $mono set chain [lindex $chainlist $i]
    } elseif { [llength $chainlist] == 0 } {
      set chainstr
"ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789"
      set chainchar [string index $chainstr $i]
      puts "Warning: making up sequential chain IDs, no chain list provided"
      puts "Using chain code '$chainchar'"
      $mono set chain $chainchar
    } else {
      puts "Warning: not setting monomer chains IDs, mismatched chain list size"
    }

    $mono writepdb ${output}.${i}.pdb
    set channel [open ${output}.${i}.pdb r]
    while {[gets $channel line]>=0} {
      if {[lindex $line 0]=="ATOM"} {
        puts $out $line
      }
    }
    close $channel
    file delete ${output}.${i}.pdb
    mol delete top
    incr i
  }
  close $out
  file delete ${output}.org.pdb
  mol load pdb ${output}.pdb
  set all [atomselect top all]
  $all writepdb ${output}.pdb
  mol delete top
  return "${output}.pdb file have finished"
}
```

To create the biologically active form of a protein using BIOMT records, the following protocol can be used (text preceded by "#" are annotations explaining individual lines):

In the VMD Main window, load the PDB file for the monomer (containing BIOMT transformations) into VMD. In this work, wild-type and mutant LpxAs and AcpSs were processed. The example below is for wild-type LpxA, but the process is the same for the other PDBs.

In the “Tcl Console” of VMD:

```
source mono2poly.tcl                    #Loads script for conversion from monomer to multimer.

set sel [atomselect top all]            #Creates a selection, sel, including all atoms.

set matrix [parsematrix LpxA-WT.pdb]    #Creates a selection, matrix, that contains the BIOMT transformation information. LpxA-WT.pdb is replaced by the filename for the structure file of the LpxA mutants, wild-type AcpS, or AcpS mutants.

mono2poly -o LpxA-WT-Timer $sel $matrix
                                        #Executes the script’s mono2poly function using the selections “sel” and “matrix”, saving the output as “LpxA-WT-Trimer.pdb”.
```

APPENDIX 9 Sample Tcl Script for Solvating a Protein in a Water Sphere

Script used as provided by the VMD tutorial ((VMD Online Documentation. *Tutorials*.); All Tutorial Files link). Sample given is for rACP (no metal). Portions of the script changed for different proteins highlighted (other options given after "#"). Script executed in VMD (1.8.7 or 1.9.1). Script executed as in Appendix 10.

water_rACP.tcl script contents:	#rACP is replaced with SA, SB, SASB, rACP_E-A, rACP_E-A_Mg, rACP_E-A_Ni, rACP_E-A_Zn, SA_E-A, SA_E-A_Mg, SA_E-A_Ni, SA_E-A_Zn, SB_E-A, SB_E-A_Mg, SB_E-A_Ni, SB_E-A_Zn, SASB_E-A, SASB_E-A_Mg, SASB_E-A_Ni, SASB_E-A_Zn, linL46W, cycL46W, LpxA-WT, LpxA-Q104W, LpxA-F162W, LpxA-G173W, LpxA-S208W, AcpS(WT)-ACP(WT), AcpS(R22E)-ACP(WT), AcpS(WT)-ACP(E41K), AcpS(R22E)-ACP(E41K), AcpS(F27W)-ACP(WT), AcpS(R22E/F27W)-ACP(WT), AcpS(F27W)-ACP(E41K), AcpS(R22E/F27W)-ACP(E41K), or P.f.AcpS, depending on which protein is being solvated.
### Script to immerse rACP in a sphere of water just large enough to cover it	#rACP is replaced (as above) depending on which protein is being solvated.
set molname V.H.-rACP	#Sets the variable "molname" to V.H.-rACP. This variable is used below to load the PSF and PDB files into VMD.
mol new ${molname}.psf	#Creates a new molecule in VMD using the PSF file molname.PSF (V.H.-rACP.psf in this case).
mol addfile ${molname}.pdb	#Adds information to the new molecule using molname.pdb (V.H.-rACP.pdb in this case).

Determine the center of mass of the molecule and store the coordinates
set cen [measure center [atomselect top all] weight mass]

```
set x1 [lindex $cen 0]
set y1 [lindex $cen 1]
set z1 [lindex $cen 2]
set max 0

### Determine the distance of the farthest atom from the center of mass
foreach atom [[atomselect top all] get index] {
  set pos [lindex [[atomselect top "index $atom"] get {x y z}] 0]
  set x2 [lindex $pos 0]
  set y2 [lindex $pos 1]
  set z2 [lindex $pos 2]
  set dist [expr pow(($x2-$x1)*($x2-$x1) + ($y2-$y1)*($y2-$y1) + ($z2-$z1)*($z2-
$z1),0.5)]
  if {$dist > $max} {set max $dist}
  }

mol delete top

### Solvate the molecule in a water box with enough padding (15 A).
### One could alternatively align the molecule such that the vector
### from the center of mass to the farthest atom is aligned with an axis,
### and then use no padding
package require solvate
solvate ${molname}.psf ${molname}.pdb -t 15 -o del_water

resetpsf
package require psfgen
mol new del_water.psf
mol addfile del_water.pdb
readpsf del_water.psf
coordpdb del_water.pdb

### Determine which water molecules need to be deleted and use a for loop
### to delete them
set wat [atomselect top "same residue as {water and ((x-$x1)*(x-$x1) + (y-$y1)*(y-$y1)
+ (z-$z1)*(z-$z1))<($max*$max)}"]
set del [atomselect top "water and not same residue as {water and ((x-$x1)*(x-$x1) + (y-
$y1)*(y-$y1) + (z-$z1)*(z-$z1))<($max*$max)}"]
set seg [$del get segid]
set res [$del get resid]
set name [$del get name]
for {set i 0} {$i < [llength $seg]} {incr i} {
  delatom [lindex $seg $i] [lindex $res $i] [lindex $name $i]
  }
writepsf ${molname}_ws.psf
writepdb ${molname}_ws.pdb
```

```
mol delete top

mol new ${molname}_ws.psf
mol addfile ${molname}_ws.pdb
puts "CENTER OF MASS OF SPHERE IS: [measure center [atomselect top all] weight
mass]"
puts "RADIUS OF SPHERE IS: $max"
mol delete top
```

The script outputs two files (molname_ws.psf/molname_ws.pdb) as well as the center of mass of the sphere as well as the sphere radius (both are needed in the NAMD configuration file (see Appendix 16)).

Output files water_rACP.tcl (in working directory):

V.H.-rACP_ws.psf
V.H.-rACP_ws.pdb

Text output for water_rACP.tcl (in "Tcl Console"):

CENTER OF MASS OF SPHERE IS: 0.2240777611732483 1.8840218782424927
0.4377172887325287
RADIUS OF SPHERE IS: 18.935011812984914

Note: for all of the ACP simulations a radius of 21 (sphericalBCr1 setting in Appendix 16) was set for the spherical boundary conditions as the largest water sphere was ~20.5.

APPENDIX 10 Tcl Command for Loading/Executing Script Files in VMD

```
source script_filename.tcl          #Replace “script_filename.tcl” with the
                                    name of the script you want to load and/or
                                    execute.
```

APPENDIX 11 Creation of End Structure from MD simulations in VMD

Simulation end structures created using VMD (1.8.7 or 1.9.1). Text after # explains what function the command performs.

To create these structures, the following protocol can be used:

In the VMD Main window, load the appropriate *P.f.*AcpS PDB/PSF files as well as DCD file generated using trajectory_smooth.tcl (Appendix 20); delete frame 0 as it is included in the averaged DCD file.

In the "Tcl Console" of VMD:

```
Source trajectory_smooth.tcl                #Loads script for smoothing a trajectory.


set AcpS [atomselect top "not water" frame 900]
                                            #Creates a selection, AcpS, including all
                                            non-water atoms for frame 900 (last frame
                                            of the simulation).

$AcpS writepdb AcpS_End.pdb
                                            #Write a PDB file containing atoms in
                                            selection "Simulation" to
                                            "AcpS_End.pdb".
```

APPENDIX 12 Aligning Subunits to Proteins Using Chimera

UCSF Chimera version 1.7 (build 38197) (Pettersen et al., 2004, Yang et al., 2012) was used for all structure alignments. *P.f.*AcpS monomers were aligned with subunits of the template's biologically active form.

As in Appendix 1, load relevant structures and open the structural alignment tool "MatchMaker" (MM). In the case of *P.f.*AcpS, the biologically active form of PDB ID 2BDD as well as 3 copies of the *P.f.*AcpS end structure (produced in Appendix 11). To get the biologically active form of *P.f.*AcpS, 3 alignments must be done (one for each monomer). Before selection of the "Reference structure" and "Structure(s) to match", a setting must be changed to allow for the selection of just one chain (subunit) of the template to be selected. In this case, one chain (monomer) of *P.f.*AcpS is being aligned with a specific chain in the reference structure, thus "Specific chain in reference structure with best-aligning chain in match structure" should be selected under the "Chain Pairing" section of the MM window. If all three monomers had been saved to a single file prior to alignment, the last option ("Specific chain in reference structure with specific chain in match structure") would be used instead. The "Reference structure" and "Structure to match" can now be selected as in Appendix 1, selecting one subunit of the template structure (2BDD) as the reference and one monomer of *P.f.*AcpS as the structure to match. The rest of the alignment process proceeds as in Appendix 1. To finish creation of the biologically active form of *P.f.*AcpS, the above process was repeated for the second and third monomer of *P.f.*AcpS using the second and third subunit of 2BDD, respectively, as the reference structure. As in Appendix 1, the aligned monomers were saved relative to the template (2BDD) to preserve their relative positions.

Note: The PDB files saved at the end of the above process are saved in another directory and/or with a different name so the original structures aren't overwritten.

APPENDIX 13 Combining Multiple Structures using VMD's Merge Structures Plugin

Merge Structures plugin used in version 1.91 of VMD. To use this plugin, PSF files are needed in addition to PDB files. All PSF files created for use in the plugin were made using VMD's AutoPSF (see Appendix 7 for more details).

In the VMD open the Merge Structures plugin (Extensions\Modeling\Merge Structures).

Select the first of two structures to merge (e.g. *P.f.*AcpS Subunit A) using the "Browse" button (load both the PDB and PSF file). If the structures were loaded in VMD prior to opening the Merge Structures plugin, the "Select loaded molecule" button can be used instead of selecting PDB/PSF files. Load the second structure to be merged (e.g. *P.f.*AcpS Subunit B) following the same procedure. Lastly, give a name for the merged file (e.g. *P.f.*AcpS Subunit AB).

For proteins with more than two subunits to be merged (e.g. *P.f.*AcpS), they can be subsequently added to the merged file (e.g. *P.f.*AcpS Subunit AB) by repeating the above process.

APPENDIX 14 Creation of Equilibrated, Averaged Structures from MD simulations in VMD

All equilibrated structures were created using VMD (1.8.7 or 1.9.1). Text after "#" explains what function the command performs.

To create an equilibrated, averaged structure the following protocol can be used:

In the VMD Main window, load the appropriate PDB/PSF files as well as DCD file generated using trajectory_smooth.tcl (Appendix 20); delete frame 0 as it is included in the averaged DCD file.

In the "Tcl Console" of VMD:

```
Source trajectory_smooth.tcl                    #Loads script for smoothing a trajectory.


set sel [atomselect top "all"]                  #Creates a selection, sel, including all
                                                atoms.
sliding_avg_pos $sel 1001 Simulation_Smoothed_1001.dcd first last.
                                                #Smoothes the MD simulation (selection
                                                of atoms) with a sliding window of 1001
                                                and writes the new atom trajectory
                                                information to a new DCD file
                                                Simulation_Smoothed_1001.dcd;
                                                Simulation is replaced with what
                                                simulation was being analyzed at the time
                                                (e.g. SASB). For a more detail account on
                                                using trajectory_smooth.tcl see Appendix
                                                20.
```

In the VMD Main Window, delete the protein and reload the PDB/PSF and the newly created smoothed DCD file Simulation_Smoothed_1001.dcd (as above delete frame 0 as it is included in the averaged DCD file.

In the "Tcl Console" of VMD:

```
set Simulation [atomselect top "not water" frame #]
                                                #Creates a selection, Simulation,
                                                including all non-water atoms for frame #;
                                                Simulation is replaced by the protein
                                                being saved (e.g. SASB+Mg); # is
                                                replaced by a frame number near the end
                                                of the simulation where there are no
                                                improper dihedrals, atom positions, etc
                                                (e.g. 20500).
```

$Simulation writepdb Simulation_equilibrated.pdb

#Write a PDB file containing atoms in selection “Simulation” to “Simulation_equilibrated.pdb”; as above Simulation is replaced by the protein being saved (e.g. SASB+Mg).

For ACP ±Metal simulations:
Instead of having to manually load a simulation, set the selection and write a PDB for each simulation, scripts were written to automate and speed up the process. This was done after the smoothed trajectory (sliding window of 1001) was created for each simulation. See Appendix 35 for a sample script that loads a simulation and saves the equilibrated structure to a PDB file and Appendix 15 for the script that consecutively processes each of the 16 simulations.

APPENDIX 15 Script for Loading Multiple ACP ±Metal Simulations Consecutively and Saving Equilibrated Structures

Sample given is for rACP. Portions of the script changed for different proteins highlighted (other options given after "#"). Script written for and executed in VMD (1.8.7 or 1.9.1). Based on the manipulations required, I provide the instructions below:

First, the "Tcl Console" was opened and the working directory was changed to a directory containing the movie generation scripts (within a general ACP simulation directory, containing all ACP ±Metal simulation folders). The script was then executed as indicated in Appendix 10.

Equilibrated-Structure-Generation.tcl script contents:

```
play rACP.tcl                          #Execute rACP.tcl (see Appendix 35 for
                                       detailed explanation of script contents).

wait 5                                 #Wait 5 s.

puts "rACP PDB Written"                #Outputs the text "rACP PDB Written" in
                                       the "Tcl Console". Having this text output
                                       gives an indication of the progress of the
                                       script.
play rACP_Mg.tcl
wait 5
puts "rACP+Mg PDB Written"
play rACP_Ni.tcl
wait 5
puts "rACP+Ni PDB Written"
play rACP_Zn.tcl
wait 5
puts "rACP+Zn PDB Written"
play SA.tcl
wait 5
puts "SA PDB Written"
play SA_Mg.tcl
wait 5
puts "SA+Mg PDB Written"
play SA_Ni.tcl
wait 5
puts "SA PDB+Ni Written"
play SA_Zn.tcl
wait 5
puts "SA+Zn PDB Written"
play SB.tcl
wait 5
```

```
puts "SB PDB Written"
play SB_Mg.tcl
wait 5
puts "SB+Mg PDB Written"
play SB_Ni.tcl
wait 5
puts "SB+Ni PDB Written"
play SB_Zn.tcl
wait 5
puts "SB+Zn PDB Written"
play SASB.tcl
wait 5
puts "SASB PDB Written"
play SASB_Mg.tcl
wait 5
puts "SASB+Mg PDB Written"
play SASB_Ni.tcl
wait 5
puts "SASB+Ni PDB Written"
play SASB_Zn.tcl
wait 5
puts "SASB+Zn PDB Written"
```

APPENDIX 16 Sample Configuration File for MD Simulations

Below is a sample configuration file used for NAMD MD simulations. Text within the configuration file preceded by "#" are descriptions/titles for each section of the file. Highlighted text preceded by "#" are annotations explaining individual lines of the file and are not included in the file used by NAMD to perform the MD simulation. More in-depth information about sections of a configuration file can be found in the NAMD Tutorial (NAMD Online Documentation. *Tutorial*.)).

```
#############################################################
## JOB DESCRIPTION                                         ##
#############################################################
#
# Minimization and Equilibration of
# ACP (rACP, no metal) in a Water Sphere
#
#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################
```

```
structure          ./rACP_E-A_ws.psf
```
#Indicates the location of structure file (current directory in this case). The structure file tells NAMD how the atoms are connected.

```
coordinates        ./rACP_E-A_ws.pdb
```
#Indicates the location of coordinates file (current directory). The coordinates file contains the X, Y, and Z positions of all atoms in the simulation.

```
set temperature    310
```
#Sets the variable "temperature" to 310, the variable can be called using $temperature (see below).

```
set outputname     rACP_E-A_ws_eq
```
#Sets the output filename.

```
firsttimestep      0
```
#Sets the first timestep of the simulation. This value will always be 0 unless a simulation is being restarted from a frame other than 0.

```
#############################################################
## SIMULATION PARAMETERS                                   ##
#############################################################

# Input
paraTypeCharmm          on                  #Indicates to NAMD that parameter files
                                            given below are in CHARMM format.

parameters          ./par_all27_prot_lipid.inp  #Indicates to NAMD the location of
                                            parameter files needed to perform the
                                            simulation. “par_all27_prot_lipid.inp”
                                            will be sufficient for most protein and/or
                                            lipid simulations. Simulations that contain
                                            nucleic acids, non-standard amino acids,
                                            ligands, etc will need other files. If a
                                            parameter file does not exist for any
                                            simulation component the simulation will
                                            fail. In that case, a parameter file would
                                            need to be found or created for the non-
                                            parameterized portion of the simulation.

temperature         $temperature            #Indicates the temperature at which the
                                            simulation will occur. In this case the file
                                            is using a variable (“temperature”; defined
                                            above as 310 (given in Kelvins (K))).

# Force-Field Parameters
exclude             scaled1-4               #“exclude” specifies which atomic
                                            interactions are to be excluded from
                                            consideration. Setting this to “scaled1-4”
                                            indicates that interactions between atoms
                                            1 and 2, or 1 and 3 (see Figure 117 for
                                            numbering) are neglected and interactions
                                            between atoms 1 and 4 and weakened.

1-4scaling          1.0                     #“1-4scaling” specifies if electrostatic
                                            interactions between atoms 1 and 4 are
                                            taken into account; possible settings are 0
                                            and 1, representing off and on,
                                            respectively.

cutoff              12.                     #Sets the distance (in angstroms) at which
                                            electrostatic and Van der Waals (VdW)
                                            forces are cut-off. If an large cut-off is
                                            used, interactions (although weak)
```

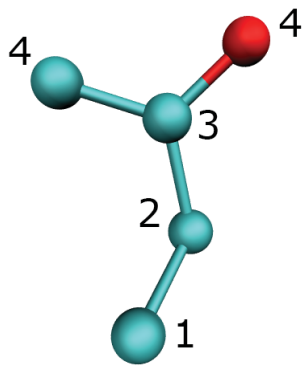

Figure 117 Generic number labels for atoms when determining the effect each subsequent atom has on the primary atom. Labeling is determined by the number of bonds each atom is away from the primary atom (1). For example, both the oxygen and carbon atoms at the opposite end of the molecule to carbon 1 are 3 bonds away from the primary atom and are therefore numbered 4. Figure created in PyMol 1.2r1 based on the NAMD Tutorial (NAMD Online Documentation. *Tutorial*.)). Picture created in VMD 1.9.1 and rendered using POV-Ray 3.6.

	between atoms over long distances will occur, increasing the total number of interactions taken into account for each atom. This causes an increase in processing power and time needed for the simulation (Figure 118A and B).
switching on	#Indicates whether switching functions are on or off. Switching functions allow a smooth transition to zero for electrostatic and VdW forces at the cut-off distance specified above.
switchdist 10.	#Sets the distance at which the switching functions (if on) start tapering the above forces to zero (tapered such that zero is reached at cut-off distance specified above). For a visual representation of effect of switching/switching distance see Figure 118A (dotted line).
pairlistdist 13.5	#Specifies the distance (in angstroms) from an atom that NAMD will search for electrostatic or VdW forces. As with switching, this will decrease computation time. For a visual representation of effect of pairlistdist/cutoff see Figure 118B.
# Integrator Parameters	
timestep 2.0 ;	#The “timestep” parameter tells NAMD how much time occurs per step and is set in femptoseconds (fs); in this case 2.0 fs. This is needed as MD simulation programs such as NAMD use Newton’s laws in a discrete approximation to determine the trajectories of the atoms in a simulation. The “timestep” tells NAMD how to discretize the particle dynamics.
rigidBonds all ;	#“rigidBonds” specifies which bonds involvinghydrogen are kept rigid for the simulation (i.e. no vibrations). For a 2.0 fs timestep, the setting must be all. Setting rigidBonds to all tells NAMD that all bonds involving hydrogen are kept rigid (irrespective of what the adjoining atom is).

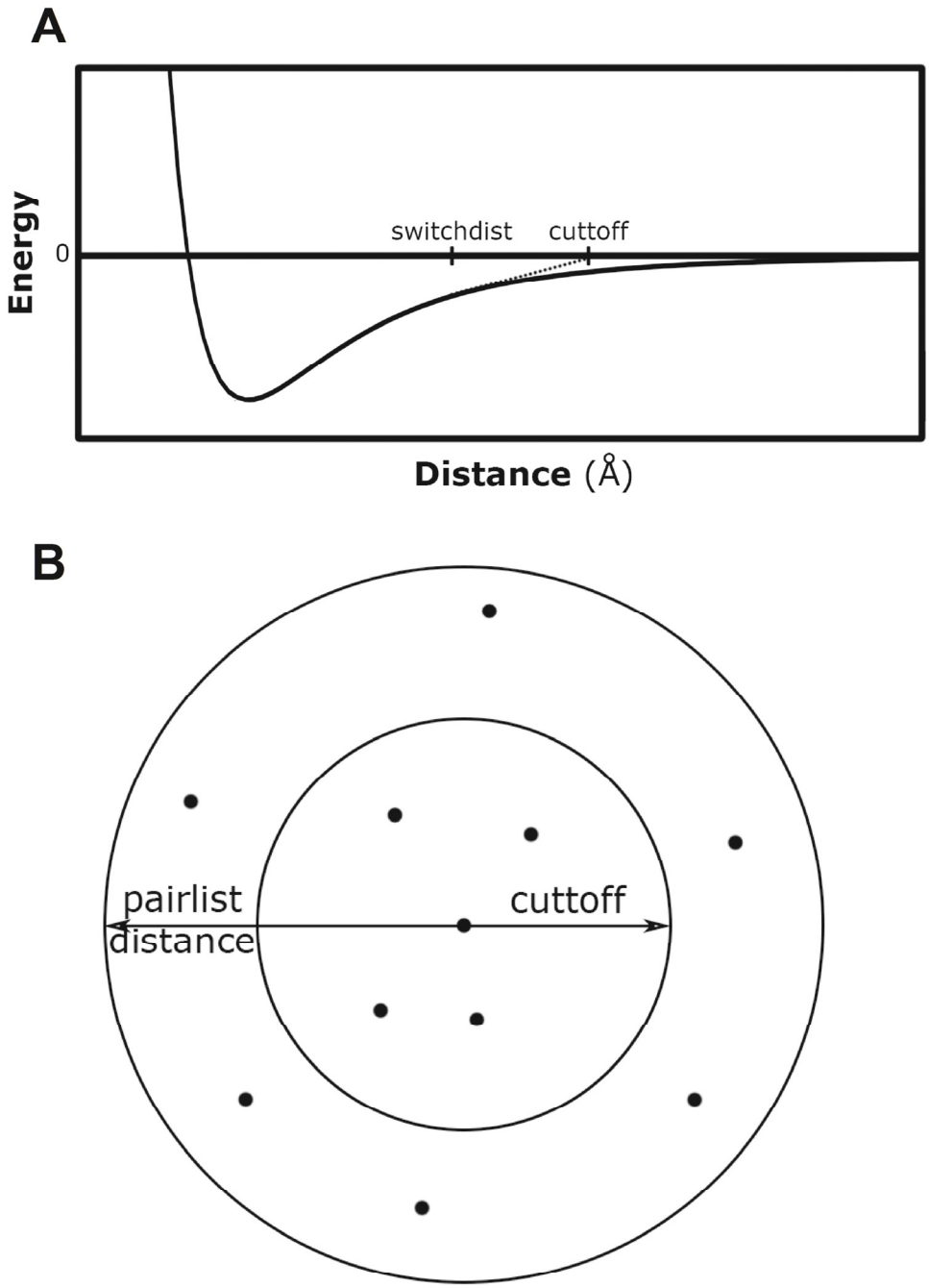

Figure 118 Visual representations of the relationship between "switchdist" and "cutoff" (A) and "pairlistdist" and "cutoff" (B). The plot in A is a theoretical Lennard-Jones potential diagram for energy *versus* distance between atoms. In addition to portraying the relationship between switchdist and cutoff, A also depicts the function of switching visually (dotted line). For the diagram in B, the dots represent atoms surrounding the central atom. This figure was re-created based on the NAMD Tutorial (NAMD Online Documentation. Tutorial.).

```
nonbondedFreq       1
```
#Tells NAMD how often (in “timesteps”) non-bonded interactions (VdW, etc) should be calculated. This can be increased to save computational time.

```
fullElectFrequency  2
```
#Tells NAMD how often (in “timesteps”) full electrostatic interactions are calculated. As with nonbondedFreq, an increase in this number will decrease computational time.

```
stepspercycle       10
```
#Specifies the number of “timesteps” in a cycle. Once a cycle atoms are reassigned pair list identities (identified using pairlistdist and cutoff, as demonstrated in Figure 118B and discussed above).

```
# Constant Temperature Control
langevin            on    ;
```
#Indicates whether or not Langevin dynamics (LD) is used to control the simulation temperature.

```
langevinDamping     5     ;
```
#Sets the Langevin coupling coefficient (γ in Equation 2; see Appendix 17 for a short explanation of LD). This specifies the friction that is applied to the system (decreasing atom speed, etc). It is specified in ps^{-1}; in this case 5 ps^{-1}.

```
langevinTemp        $temperature
```
#Specifies which temperature LD should keep atoms at. In this case the variable “temperature” is used (set above to 310 K).

```
langevinHydrogen    off    ;
```
#Specifies if the LD should be applied to hydrogen atoms. Default is off.

```
# Output
outputName          $outputname
```
#Sets the output filename. In this case the file is using a variable (“outputname”; defined above as rACP_E-A_ws_eq).

```
restartfreq         250     ;
```
#Sets the restart frequency (how many steps before NAMD writes files that can

be used to restart a simulation that has stopped for any reason). In this case every ½ ps.

```
dcdfreq             250                 #Sets the frequency at which NAMD writes coordinate information to the trajectory (DCD) file. In this case every ½ ps.

outputEnergies      100                 #Sets the frequency at which NAMD writes the system energy to the log file for the simulation. In this case every 100 frames.

outputPressure      100                 #Sets the frequency at which NAMD writes the system pressure to the log file for the simulation. In this case every 100 frames.

#############################################################
## EXTRA PARAMETERS                                        ##
#############################################################

# Spherical boundary conditions
sphericalBC         on                  #Indicates that a spherical boundary should be used (instead of a box). A spherical boundary is used in this case as a water sphere was used to solvate proteins prior to simulation. A water sphere was used to solvate instead of a box as it saves simulation time (and therefore processing power; this allows more step (time) to be simulated). If a box is used to solvate, a large portion of the start of the simulation will be rearrangement of the water molecules to form a spherical shape.

sphericalBCcenter   0.2240777611732483 1.8840218782424927 0.4377172887325287
                                        #Sets the XYZ coordinates for the center of the sphere. These coordinates are given after a protein is solvated (see Appendix 9).

sphericalBCr1       21                  #Sets the size of the spherical boundary (i.e where the first bounding potential begins to act). For solvated systems, make this slightly larger than the water sphere radius determined during protein solvation (Appendix 9).
```

```
sphericalBCk1       10                  #Sets the value of the force constant for the first bounding potential. The value is specified in kcal/molÅ^2; in this case the default of 10 kcal/molÅ^2 is used.

sphericalBCexp1     2                   #Sets the value of the exponent used to calculate the potential. Must be set to a positive, even integer. In this case the default of 2 is used.

#####################################################################
## EXECUTION SCRIPT                                                ##
#####################################################################

# Minimization
minimize            125000              #Indicates 1,250,000 steps of minimization at the start of the simulation. With a "timestep" of 2 fs (set above), 125,000 steps is 0.25 ns.

reinitvels          $temperature        #During minimization, atomic velocities are set to zero. "reinitvels" resets the atomic velocities such that the system starts at a specified temperature. In this case the variable "temperature" is used (set above to 310 K).


run 5000000 ;# 10 ns                   #Indicates 5,000,000 steps of simulation (equilibration after energy minimization). With a "timestep" of 2 fs (set above), 5,000,000 steps is 10 ns.
```

APPENDIX 17 Basic Explanation of Langevin Dynamics

To control the temperature (or pressure) of a simulation Langevin dynamics (LD) can be used. LD is a method of controlling the kinetic energy of the system, which in-turn controls temperature. This method uses the Langevin equation for a single particle (Equation 2).

$$m_i \frac{d^2 x_i(t)}{dt^2} = F_i\{x_i(t)\} - \gamma_i \frac{dx_i(t)}{dt} m_i + R_i(t)$$

(1) (2)

Equation 2 Langevin dynamics equation for a single particle used by NAMD to control temperature in a simulation.

In the above equation there are two terms (1, 2 in Equation 2) in addition to the ordinary force a particle experiences. "1" represents a frictional dampening that is applied to the particle (frictional coefficient $\gamma_i m_i$; γ is set by the "langevinDamping" command in the configuration file for a simulation (Appendix 16). "2" represents random forces which act on the particle (e.g. as a result of solvent interaction). These terms are controlled to maintain particle kinetic energy, keeping the system temperature constant.

APPENDIX 18 Tcl Script for Determining RMSD for Protein Backbone in MD simulations (Non-Smoothed)

Modified version of rmsd.tcl script provided in VMD tutorial ((VMD Online Documentation. *Tutorials*.); All Tutorial Files link). Modifications to script highlighted with original content given after the "#" symbol.

rmsd-protein_backbone.tcl script contents:

```
set outfile [open rmsd_Protein_backbone.dat w];                    #rmsd.dat
set nf [molinfo top get numframes]
set frame0 [atomselect top "protein and backbone and noh" frame 0]
set sel [atomselect top "protein and backbone and noh"]
# rmsd calculation loop
for {set i 1 } {$i < $nf } { incr i } {
   $sel frame $i
   $sel move [measure fit $sel $frame0]
   puts $outfile "[measure rmsd $sel $frame0]"
}
close $outfile
```

The following protocol can be used to create a calculate RMSD of backbone atoms over the course of a simulation:

The appropriate PDB/PSF and DCD files were loaded into VMD through the main window. Then, in the "Tcl Console" (in the VMD Main window, click Extensions then "Tcl Console") execute rmsd-protein_backbone.tcl as directed in Appendix 10.

APPENDIX 19 Tcl Script for Determining RMSD for Protein Backbone in MD simulations (Smoothed, Sliding Window of 9 Frames)

Modified version of rmsd.tcl script provided in VMD tutorial ((VMD Online Documentation. *Tutorials*.); All Tutorial Files link). Modifications to script highlighted with original content given after the "#" symbol.

rmsd-protein_backbone_smoothed.tcl script contents:

```
set outfile [open rmsd_Protein_backbone_smoothed.dat w];                    #rmsd.dat
set nf [molinfo top get numframes]
set frame0 [atomselect top "protein and backbone and noh" frame 0]
set sel [atomselect top "protein and backbone and noh"]
# rmsd calculation loop
for {set i 1 } {$i < $nf } { incr i } {
   $sel frame $i
   $sel move [measure fit $sel $frame0]
   puts $outfile "[measure rmsd $sel $frame0]"
}
close $outfile
```

The following protocol can be used to create a calculate RMSD of backbone atoms over the course of a simulation:

The appropriate PDB/PSF and DCD files were loaded into VMD through the main window (**Note:** the smoothed DCD file is used when running this script). Then, in the "Tcl Console" (in the VMD Main window, click Extensions then "Tcl Console") execute rmsd-protein_backbone_smoothed.tcl as directed in Appendix 10.

APPENDIX 20 Tcl Script for Smoothing Movement of Atoms Due to Temperature Fluctuations (sliding window of 9 frames)

Script (trajectory_smooth.tcl) downloaded from the VMD script library (VMD Downloads. *VMD Script Library*.).

trajectory_smooth.tcl script contents:

```
# source ~/vmd/tclproc/statistics.tcl

# SMOOTHING A TRAJECTORY
#--------------------------
# sliding_avg_pos $sel $width $file [beg $firstframe] [end $lastframe]
# [restore 1/0] [crop 0/1]

# Computes the average position in $width frames of the atoms in the
# selection $sel and moves this sliding window through time. Let's say
# you choose a window of width 7, then the current frame, the three
# preceding ones and the three following frames are used to compute the
# average position. This results in a smoothed trajectory which will be
# saved as dcd with the name $file.

# Depending on the window size you will be able to see net movements of
# residues and get rid of some of the thermal noise.

# You can limit the averaging to specified frames using the
# 'beg $firstframe' and 'end $lastframe' options. Instead of $lastframe
# you can simply type 'last' if you mean the last frame of the loaded
# dcd trajectory.
# Note that your original trajectory is transformed into the averaged
# one, you can automatically reload the original using the 'restore 1'
# option.
# Through the option 'crop 1' you can automatically crop the trajectory
# to the size specified in the beg/end statements

# The averaging procedure effects only the selection you specified, so
# might get strange atom distances at the fringes. Therefore you should
# include everything you want to look at in the selection. But careful,
# it could become quite slow then. (for rhodopsin with about 5600 atoms
# it takes about 1-2 seconds per frame on my linux box)

# examples:
# set sel [atomselect top "protein"]
# sliding_avg_pos $sel 9 slide_prot10.dcd beg 0 end 100 crop 1
# sliding_avg_pos $sel 9 slide_prot10.dcd beg 0 end last restore 1
```

```
proc sliding_avg_pos {sel width file args} {
   set first 0
   set last last
   set mol [$sel molid]
   set restore 0
   set crop 0

   foreach {i j} $args {
         if {$i=="beg"} then { set first $j }
         if {$i=="end"}  then { set last $j }
         if {$i=="restore"} then { set restore $j }
         if {$i=="crop"} then { set crop $j }
   }
   if { [expr fmod($width,2)] == 0.0 } then {
         puts "ERROR: Window size must be an odd number to get a symmetric window."
         puts "Try again."
         return 0
   }

   puts "Transforms this dcd file into avg trajectory and writes new file."

   set numframes [molinfo $mol get numframes]
   if {$last=="last"} then {set last [expr $numframes-1]}

   # initialize vector $oldsum
   set oldsum ""
   for {set i 0} {$i < [$sel num]} {incr i} {
         lappend oldsum {0 0 0}
   }
   set zerolist $oldsum

   # initialize $sum and $coordbuf
   set fcount 0
   set coordbuf ""
   set hwidth [expr ($width-1)/2]
   for {set frame [expr $first-$hwidth]} {$frame < [expr $first+$hwidth]} {incr frame} {
         if {$frame>=0} then {
            $sel frame $frame
            set coords [$sel get {x y z}]
            set sum ""
            foreach atom $coords osum $oldsum {
                  lappend sum [vecadd $osum $atom]
            }
            set oldsum $sum
            incr fcount
```

```
        # initialize the coordinate buffer:
        # I must store the frames of the fir\st half of the window in a buffer
        # because the center frame will be overwritten with the new avg positions
        # of the window.
        if {$frame<$first} then {
            lappend coordbuf $coords
        }
    }
}

# get the avgpos for every slice:
for {set slice $first} {$slice<$last} {incr slice} {
    puts $slice

    # get coords which will be subtracted from $sum
    if {$slice>=[expr $first+$hwidth]} then {
        $sel frame [expr $slice-$hwidth]
        set subcoords [lindex $coordbuf 0]
        incr fcount -1

        # remove the coords from the coordinate buffer
        lvarpop coordbuf
    } else { set subcoords $zerolist }

    # get the coords which will be added to sum
    if {$slice<=[expr $last-$hwidth]} then {
        $sel frame [expr $slice+$hwidth]
        set addcoords [$sel get {x y z}]
        incr fcount
    } else { set addcoords $zerolist }

    # update the coordinate buffer with the coordinates added
    $sel frame $slice
    lappend coordbuf [$sel get {x y z}]

    # update $sum
  set addsum ""
    set sum ""
    foreach addatom $addcoords subatom $subcoords osum $oldsum {
        if {$slice<[expr $first+$hwidth]} then {
            # ramp it up
            lappend sum [vecadd $osum $addatom]
        } elseif {$slice>[expr $last-$hwidth]} then {
            # ramp it down
            lappend sum [vecsub $osum $subatom]
```

```
            } else {
                  # add and substract frame data
                  lappend sum [vecsub [vecadd $osum $addatom] $subatom]
            }
         }
         set oldsum $sum

         # scale the vector
         if {$slice<[expr $first+$hwidth]} then {
            set scale [expr 1.0/($hwidth+$slice+1)]
         } elseif {$slice>[expr $last-$hwidth]} then {
            set scale [expr 1.0/($hwidth+$last-$slice-1)]
         } else { set scale [expr 1.0/($width)] }
         set scale [expr 1.0/$fcount]
         set avgpos ""
         foreach atom $sum {
            lappend avgpos [vecscale $atom $scale]
         }

      # set the new positions:
         $sel frame $slice
         $sel set {x y z} $avgpos
   }

   if {$crop==1} then {
         # delete the frames after $last:
         puts "deleting frames [expr $last+1] to [expr $numframes+1]"
         animate delete beg [expr $last+1] end $numframes
         #crop the unused frames at the beginning
         if { $first>0 } then {
            puts "deleting frames 0 to $first"
            animate delete beg 0 end $first
         }
   }

   # write a new dcd file with the avg positions
   animate write dcd $file beg 0 waitfor all

   puts "Your original file was transformed to avg coordinates."
   puts "It contains [molinfo $mol get numframes] frames.\n"

   if {$restore} {
         # load the original file:
         puts "Reloading the original trajectory with:"
         puts "> mol load psf [molinfo $mol get filename] dcd [molinfo $mol get
filename2]"
```

```
        mol load psf [molinfo $mol get filename] dcd [molinfo $mol get filename2]
        puts "It has molid [molinfo top get id]"
    } else {
        puts "You can reload the original trajectory with:"
        puts "> mol load psf [molinfo $mol get filename] dcd [molinfo $mol get
filename2]"
    }
}


# COMPUTE THE AVERAGE POSITIONS
# -----------------------------
# This procedure works similar as 'sliding_avg_pos' but it does not
# slide the window, it generates only one frame with the average
# positions of the selection between 'beg $firstframe' and 'end $lastframe'
# and returns them. Optionally they can be saved as a pdb file with the name
# $file.
# After the 'writesel' keyword you can specify which atoms you want to save
# in the pdb file:
# 'writesel none'    - no pdb file is written
# 'writesel selonly' - only the selected atoms are written
# 'writesel all'     - all atom are written

# examples:
# set sel [atomselect top "index 44 to 67"]
# set avgpos [avg_position $sel avgpos.pdb beg 0 end last writesel selonly]
# mol load pdb avgpos.pdb

# You can compute the mean square deviation using:
# set dev [dev_pos $sel $avgpos]
# mean $dev

proc avg_position { sel file args } {
    set first 0
    set last last
    set mol [$sel molid]
    set writesel none
    set restore 0
    foreach {i j} $args {
        if {$i=="beg"} then { set first $j }
        if {$i=="end"}  then { set last $j }
        if {$i=="writesel"}   then { set writesel $j }
        if {$i=="restore"} then { set restore $j }
    }
    #draw delete all
    if {$last=="last"} then {set last [expr [molinfo $mol get numframes]-1]}
```

```
   set numatoms [$sel num]
   set oldsum ""
   for {set i 0} {$i < $numatoms} {incr i} {
         lappend oldsum {0 0 0}
   }
   for {set frame $first} {$frame <= $last} {incr frame} {
         $sel frame $frame
         set coords [$sel get {x y z}]
      set sum ""
         foreach atom $coords osum $oldsum {
            lappend sum [vecadd $osum $atom]
         }
         set oldsum $sum
   }
   set scale [expr 1.0/($last-$first+1)]
   #puts $scale
   set avgpos ""
   foreach atom $sum {
         lappend avgpos [vecscale $atom $scale]
   }
   #draw color yellow
   #draw arrow {0 0 0} [lindex $avgpos 5]

   # write avg positions to pdf-file
   if {$writesel=="all"} {
         $sel frame 0
         set storepos [$sel get {x y z}]
         $sel lmoveto $avgpos
         set all [atomselect $mol "all"]
         $all writepdb $file
         $sel lmoveto $storepos
   } elseif {$writesel=="selonly"} {
         $sel frame 0
         set storepos [$sel get {x y z}]
         $sel lmoveto $avgpos
         $sel writepdb $file
         $sel lmoveto $storepos
   }
   return $avgpos
}


# Takes avg positions as input and computes the mean square deviation

proc dev_pos { sel avgpos args } {
   set first 0
```

```
    set last last
    set mol [$sel molid]
    foreach {i j} $args {
        if {$i=="beg"} then { set first $j }
        if {$i=="end"}  then { set last $j }
    }

    if {$last=="last"} then {set last [molinfo $mol get numframes]}
    if {[llength $avgpos] != [$sel num]} {
        puts "ERROR: Selection and avgpos don't have same size!"
        puts "[llength $avgpos]"
        puts "[$sel num]"
        return 0
    }
    for {set frame $first} {$frame <= $last} {incr frame} {
        $sel frame $frame
        set coords [$sel get {x y z}]
        set oldmeandev2 ""
      set meandev2 ""
        for {set i 0} {$i < [$sel num]} {incr i} {
          lappend oldmeandev2 0
        }
        set i 0
        foreach atom $coords mean $avgpos {
          set dev2 [expr pow([veclength [vecsub $atom $mean]], 2)]
          lappend meandev2 [expr [lindex $oldmeandev2 $i] + $dev2]
          incr i
        }
        set oldmeandev2 $meandev2
    }
    return $meandev2
}

# Just computes the mean value of a list of values
proc mean { x } {
    set len [llength $x]
    set sum 0
    foreach e $x {
        set sum [expr $sum+$e]
    }
    return [expr $sum/$len]
}

# Computes the average dihedral angle of a bond
# (takes the indices of the four dihedral atoms as input)
proc avg_angle { a1 a2 a3 a4 {first 0} {last last}} {
```

```
  if {$last=="last"} then {set last [molinfo top get numframes]}
#  set dihed [atomselect top "index $a1 $a2 $a3 $a4"]
  set sel1 [atomselect top "index $a1"]
  set sel2 [atomselect top "index $a2"]
  set sel3 [atomselect top "index $a3"]
  set sel4 [atomselect top "index $a4"]
  set sum 0
  for {set frame $first} {$frame < $last} {incr frame} {
    $sel1 frame $frame
    $sel2 frame $frame
    $sel3 frame $frame
    $sel4 frame $frame
    set coord1 [lindex [$sel1 get {x y z}] 0]
    set coord2 [lindex [$sel2 get {x y z}] 0]
    set coord3 [lindex [$sel3 get {x y z}] 0]
    set coord4 [lindex [$sel4 get {x y z}] 0]
    set v1 [vecsub $coord1 $coord2]
    set v2 [vecsub $coord3 $coord2]
    set v3 [vecsub $coord4 $coord3]
    set cross1 [vecnorm [veccross $v2 $v1]]
    set cross2 [vecnorm [veccross $v2 $v3]]
    set dot [vecdot $cross1 $cross2]
    set angle [expr acos($dot)]
    set sum [expr $sum + $angle]
  }
    set avgangle [rad2deg [expr $sum/($last+1)]]
  puts "$frame $avgangle"
}

# Computes the dihedral angle of a bond
# (takes the indices of the four dihedral atoms as input)
proc dihed_angle { a1 a2 a3 a4 } {
  set dihed [atomselect top "index $a1 $a2 $a3 $a4"]
  set coord1 [lindex [[atomselect top "index $a1"] get {x y z}] 0]
  set coord2 [lindex [[atomselect top "index $a2"] get {x y z}] 0]
  set coord3 [lindex [[atomselect top "index $a3"] get {x y z}] 0]
  set coord4 [lindex [[atomselect top "index $a4"] get {x y z}] 0]
  set v1 [vecsub $coord1 $coord2]
  set v2 [vecsub $coord3 $coord2]
  set v3 [vecsub $coord4 $coord3]
  # draw delete all
  # draw color red
  # draw arrow $coord2 [vecadd $coord2 $v1]
  # draw arrow $coord2 [vecadd $coord2 $v2]
  # draw color green
  # draw arrow $coord3 [vecadd $coord3 [vecinvert $v2]]
```

```
  # draw arrow $coord3 [vecadd $coord3 $v3]
  set cross1 [vecnorm [veccross $v2 $v1]]
  set cross2 [vecnorm [veccross $v2 $v3]]
  # draw color yellow
  # draw arrow $coord3 [vecadd $coord3 $cross2]
  # draw arrow $coord2 [vecadd $coord2 $cross1]
  set dot [vecdot $cross1 $cross2]
  set angle [rad2deg [expr acos($dot)]]
  return $angle
}

# Computes the angle between two vectors x and y
proc vecangle {x y} {
    if {[llength $x] != [llength $y]} {
        error "vecangle needs vectors of the same size: $x : $y"
    }
    set ret 0
    foreach t1 $x t2 $y {
        set ret [expr $ret + $t1 * $t2]
    }
    return [rad2deg [expr (acos($ret/([veclength $x] * [veclength $y])))]]
}

proc deg2rad { deg } {
  return [expr ($deg/180*3.14159265)]
}

proc rad2deg { rad } {
  return [expr ($rad/3.14159265)*180]
}
```

The following protocol can be used to create a smoothed trajectory (DCD) file (text preceded by "#" are annotations explaining individual lines):

The appropriate PDB/PSF and DCD files were loaded into VMD (Humphrey et al., 1996) through the main window. Then, in the "Tcl Console" source (load information from) trajectory_smooth.tcl as directed in Appendix 10. Now that the script has been loaded, a selection is needed for the script to perform its functions on. This is done by typing (again, in the "Tcl Console"):

```
set sel [atomselect top "all"]                    #Creates a selection, sel, including all
                                                  atoms.
```

Once the selection is created, the script can be executed with a simple command:

```
sliding_avg_pos $sel 9 Smoothed_9.dcd first last
                                    #“sliding_avg_pos” is the command being
                                    executed in the previously sourced script
                                    #“$sel” tells the script the selection to use.
                                    #“9” is the number of frames to use in the
                                    sliding window (number of frames to use
                                    to create the average position for each
                                    frame (9 means 4 frames on either side of
                                    the current frame (as well as the current
                                    frame) are used.
                                    #“Smoothed_9.dcd” tells the script to
                                    save the averaged trajectory coordinates
                                    to Smoothed_9.dcd (Smoothed_9
                                    replaced by e.g. SASB_Smoothed_9.dcd)
                                    #“first” and “last” tell the script which
                                    frames to analyze (in this case, all frames
                                    are being analyzed).
```

Note: For most of the MD simulations completed in this thesis, a sliding window of 9 was used for smoothing.

APPENDIX 21 Measuring Distances Between Atoms in VMD

All distance measurements were made in VMD (1.8.7 or 1.9.1).

To measure distances, the following protocol can be used:

In the VMD Main window, load the appropriate PDB/PSF files (as well as DCD file if distance(s) over the course of the simulations are desired). Create graphical representation(s) of the structure loaded such that the atoms for the distance to be determined are visible and easy to select (i.e. atoms/residues not necessary for the measurement are hidden). I have found CPK representation style with the correct selection works best for the atoms you want to determine distances between. For more information on using "Graphical Representations" see Section 1.3 of the VMD tutorial (VMD Online Documentation. *Tutorials*.).

In the "VMD OpenGL Display" window, re-orient the view (using scale (keyboard shortcut s), rotate (r) and transform (t) (see Section 1.2 of the VMD tutorial ((VMD Online Documentation. *Tutorials*.) for information on manipulating the viewpoint)) such that the first atom in the atom set to be measured is in front and easily selectable (no other atoms blocking it).

Back in the VMD Main window, click "Mouse/Label/Bonds" or press 2 (keyboard shortcut for "Mouse/Label/Bonds").

Back in the "VMD OpenGL Display" window, notice the cursor has changed to a cross. Use the cross the select the first atom in the atom set to be measured. A text label should appear telling you the residue and atom type (if you are dealing with a protein) of the atom you selected. This can be used to confirm the correct atom is selected. Clicking/dragging the mouse (when there are no atoms under the cross) allows the viewpoint to be rotated. Rotate the viewpoint such that the second atom in the atom set is easily selectable (as above). Use the cross to select the second atom. After selection of the second atom, a label appears (as above) in addition to a dotted line connecting the two selected atoms labeled with the distance of the line (in angstroms). This process can be repeated for any number of atom pairs. If a DCD (trajectory) was loaded in addition to the PDB/PSF combination at the start, this distance will change as the simulation is played. To save the data representing the distance *versus* simulation frame number, click "Graphics/Labels".

In the Labels window all of the atoms that have been selected are shown, selecting Bonds from the drop-down menu (Atoms default choice) the list changes to atom pair (bonds) measured above. Click the atom pair of interest; then click the "Graph" tab. To see a graph of the data (distance *versus* simulation frame number) click the "Graph" button. To save the data, click the "Save" button. Data exported using the "Save" button is in comma separated value (CSV) format. Knowing this is useful if you want to load the file in Microsoft® Excel® to graph, or to extract data for processing.

When rendering images containing bond distances, use the “Snapshot VMD OpenGL Window” engine instead of the “POV-Ray 3.6” rendering engine, as the “bonds” and their distances may not display correctly because the thickness of the line below minimum threshold for POV-Ray to render it (see Appendix 24 for instructions for rendering BMPs and saving POV-Ray files).

APPENDIX 22 Calculation of Protein Volume and Surface Area in UCSF Chimera

UCSF Chimera version 1.7 (build 38197) was used for all calculations. Based on the manipulations required, I provide the following instructions:

Instructions for ACP equilibrated structures (±metal):
In the UCSF Chimera window Click File/Open and browse to the location of equilibrated structure. Select the equilibrated structure (PDB) and click Open. The structure will now open in the main Chimera window. To measure a volume a MSMS surface must be added to the protein. An MSMS surface is calculated using Michael Sanner's Molecular Surface (MSMS) program. This program is written in the C programming language and computes molecular surfaces; it is included as a function of chimera. To do this select the entire protein (this can be done through the Select menu or with the mouse while holding CTRL), then click Actions/Surface/Show. Now that the surface is visible, click Tools/Volume Data/Measure Volume and Area. A new window will pop-up that displays the volume and surface area of the loaded PDB. This information was recorded in Microsoft® Excel® and graphed with GraphPad™ Prism® 5.04.

Instructions for ACP trajectories (±metal):
In the UCSF Chimera window Click Tools, then MD/Ensemble Analysis, then MD Movie. In the new window, select NAMD (PSF/DCD) as the trajectory format and then browse to and locate the PSF/DCD files for the simulation to be analyzed. Click OK and the molecule and trajectory information will be loaded. Once this has finished a new window (MD Movie) will open allowing you to control the simulation. As a script was not available that saved volume/surface area information for each frame to a file was not available (and I was unsuccessful in writing one), the volume and surface area were determined every 500 frames. As above, add a MSMS surface to the protein and open the Measure Volume and Surface Area tool. Ensure "Update Automatically" is checked, and determine the volume/surface area for every 500 frames. This information was recorded in Microsoft® Excel® and graphed with GraphPad™ Prism® 5.04.

Note: at the time of writing, Chimera only included protein in the calculation of the MSMS surface, volume and surface area, but a future version will have an option to include ions.

APPENDIX 23 Determining Solvent Accessible Surface Area (SASA) of Residues in Chimera

All SASA measurements were made in UCSF Chimera version 1.7 (build 38197).

To make SASA measurements, the following protocol can be used:

In the UCSF Chimera window, load the appropriate PDB file(s). Add a MSMS surface to the PDBs (Actions/Surface/Show). Once the surface is shown, the SASA for each residue can be determined. Open the Render by Attribute function (Tools\Structure Analysis\ Render by Attribute. In the "Render/Select by Attribute" window set "Attributes of" to "residues" and select a PDB file (model) to be analyzed. Under the "Render" tab, set the "Attribute" to "areasSAS"; a histogram of the data should appear. The SASA data for this PDB can be saved (File/Save Attributes). Repeat this process for any additional PDB files remembering to save each new data set as a new filename. Exported data is in comma separated value (CSV) format.

In the case of this work, the data was then loaded into Microsoft® Excel® and plotted using GraphPad™ Prism® 5.04.

APPENDIX 24 Rendering Images or POV-Ray files Using VMD

VMD (1.8.7 or 1.9.1) was used to render all images (BMP) and save all POV-Ray files.

The structure to be rendered as a BMP (using Snapshot (VMD Open GL window) renderer) or saved as a POV file is loaded and "Graphical Representations" are set up such that the protein (and any other molecules) are visible in the correct representation (NewCartoon, VDW, etc) For more information on using "Graphical Representations" see Section 1.3 of the VMD tutorial ((VMD Online Documentation. *Tutorials*.). The view is rotated to a viewpoint that displays the structure correctly (to show bonds, a certain residue, etc). The display is resized as necessary. For all images and POV-Ray files saved from VMD, the XYZ axis indicator is turned off (Display/Axes/Off in the VMD Main window). Once the viewpoint, representations, etc are correct a BMP or POV file can be saved. The "File Render Controls" window (opened by clicking File/Render in the VMD Main window) is used to save image (BMP) or raytracing (POV) files.

For BMP files:
In "File Render Controls" window the set the rendering engine under "Render the current scene using:" to "Snapshop (VMD Open GL window)" and give the BMP a file name (including ".BMP" file extension). If just a filename is specified, the BMP file will be saved to the current working directory. To save the file elsewhere, click "Browse" and select the location for the file to be saved as well as a filename (including ".BMP" file extension). Once the filename (and location if necessary) are set, click "Start Rendering" and the BMP will be saved in the working directory or the specified location.

For POV-Ray files:
Saving POV-Ray files uses a similar process to rendering a BMP file with a few exceptions. Before opening the "File Render Controls" window, the materials used in the "graphical Representations" must be changed (e.g. the vast majority of POV-Ray files saved in this work use the "Opaque" material). Open the "Materials" window (Graphics/Materials in the VMD Main window) and select the material to be modified ("Opaque" in most cases). Both "Specular" and "Shininess" sliders should be decreased before saving the POV file as raytracing software (like POV-Ray) handle light differently than the basic "Snapshot" rendering engine in VMD. If these changes are not made the BMPs resulting from rendering the POV-Ray files would have harsh reflections. A setting of 0.07 works well for both "Specular" and "Shininess", but may change dependent on what final look you are trying to achieve (some experimentation may be necessary). Once the changes have been made to all materials used in the "Graphical Representations" set up earlier, the "File Render Controls" window can be opened (as above). Set the "Render the current scene using:" to "POV-Ray 3.6". As above, set the filename (and location if necessary) where the POV file will be saved. Once this has been set the POV file can be saved by clicking "Start Rendering". This file must now be rendered using POV-Ray 3.6 (see Appendix 25). When saving POV-Ray files care should be taken when re-sizing the "VMD OpenGL Display" window as the exact size of the "VMD OpenGL Display" window must be known for POV-Ray to correctly render the file to a BMP. If the wrong dimensions are used, the resultant image will not look correct

due to either stretching or compression of the dimensions. To ensure the correct size it is best to set the “VMD OpenGL Display” window size directly using the “Tcl Console” using the “display resize” command (as in Appendix 27).

APPENDIX 25 Rendering POV-Ray files Using POV-Ray 3.6

All POV-Ray files (rendered using the POV-Ray Graphical User Interface (GUI) or command line (as in Appendix 29 and Appendix 30)) were rendered using POV-Ray 3.6.

To render a single POV-Ray file:
The file to be rendered is opened in POV-Ray. Before rendering the file, the resolution (height and width in pixels) must be set. This is done using the dropdown menu selection box in the top left-hand corner of the window (below "New" in the toolbar). If the required resolution is not available in the list, it must be added to the "resolution INI file". Instructions on how this is achieved are available in the POV-Ray Online Documentation (POV-Ray Online Documentation. *Adding New Resolutions*.). Once the correct resolution is selected, the POV file can be rendered by clicking "Run" on the toolbar.

To render multiple POV-Ray files:
Instead of opening each file in POV-Ray and rendering (as above), multiple files can be rendered in sequence using POV-Ray's built-in "File Queue". Up to 512 files can be set to render using the "File Queue" ("Render/File Queue" in the POV-Ray window, or "Queue" in the toolbar"). In the "File Queue" window click "Add File" and browse to/select the file(s) to be added to the queue. If the files are in multiple locations, this process will need to be repeated. Once the "File Queue" has all POV-Ray files to be rendered, click "OK". If "Auto Render" is on ("Render/Auto Render"; can also be set using checkbox in "File Queue" window), the files will automatically start rendering after clicking "OK". "Auto Render" must be used for the "File Queue" to work properly.

APPENDIX 26 Tcl Script for Saving All Frames of a Simulation to POV-Ray Files

Modified version of trajectory_movie.tcl from the VMD script library (VMD Downloads. *VMD Script Library*.). Modifications to script are highlighted with original content given after #.

trajectory_movie-POV3.tcl script contents:

```
proc take_picture {args} {
 global take_picture

 # when called with no parameter, render the image
 if {$args == {}} {
  set f [format $take_picture(format) $take_picture(frame)]
  # take 1 out of every modulo images
  if { [expr $take_picture(frame) % $take_picture(modulo)] == 0 } {
   render $take_picture(method) $f
   # call any unix command, if specified
   if { $take_picture(exec) != {} } {
    set f [format $take_picture(exec) $f $f $f $f $f $f $f $f $f $f]
    eval "exec $f"
   }
  }
  # increase the count by one
  incr take_picture(frame)
  return
 }
 lassign $args arg1 arg2
 # reset the options to their initial stat
 # (remember to delete the files yourself
 if {$arg1 == "reset"} {
  set take_picture(frame)  0
  set take_picture(format) "./animate.%05d.pov"    #animate.%04d.rgb; %05d
                                                   specifies 5 placeholders in the
                                                   filename (this would be changed
                                                   based on the number of frames in
                                                   the simulation; e.g. ACP ±Metal
                                                   simulations have 20501 frames,
                                                   therefore 5 placeholders are
                                                   necessary).

  set take_picture(method) POV3                    #snapshot
  set take_picture(modulo) 1
  set take_picture(exec)   {}
```

```
    return
  }
  # set one of the parameters
  if [info exists take_picture($arg1)] {
    if { [llength $args] == 1} {
      return "$arg1 is $take_picture($arg1)"
    }
    set take_picture($arg1) $arg2
    return
  }
  # otherwise, there was an error
  error {take_picture: [ | reset | frame | format  | \
  method  | modulo ]}
}
# to complete the initialization, this must be the first function
# called.  Do so automatically.
take_picture reset
```

```
proc make_trajectory_movie_files {} {
        set num [molinfo top get numframes]
        # loop through the frames
        for {set i 0} {$i < $num} {incr i} {
                # go to the given frame
                animate goto $i
           # force display update
           display update
                # take the picture
                take_picture
      }
}
```

To create a POV-Ray file for each frame of a simulation, the following protocol can be used:

To use script, appropriate PDB/PSF and DCD files were loaded into VMD (Humphrey et al., 1996) through the main window. Then, in the “Tcl Console” execute the trajectory_movie-POV3.tcl script as directed in Appendix 10. For ACP ±Metal simulations, this process would take ~1.5 h per simulation (20501 frames), however, since little processing power and memory is needed for this process, 4 simulations (using 4 separate VMD windows) could be exported at once in the same time frame. Additionally, to allow for this process to occur overnight without intervention, scripts (see Appendix 27 for sample script for loading a simulation and saving the POV-Ray

files and Appendix 28 for sample script for saving POV-Ray files for simulations consecutively) were written to automate this process, allowing one simulation to be loaded, displayed correctly, and rotated to the correct angle, saved to POV-Ray files, followed by 3 more. Total export time was ~6 h for 16 simulations (4 simulations at once, each running 3 more consecutively; totaling 328016 frames) on XPS-PC.

APPENDIX 27 Sample Script for Loading an ACP ±Metal Simulation and Saving POV-Ray Files

Sample given is for rACP (no metal). Portions of the script changed for different proteins highlighted (other options given after "#"). Script was written for and executed in VMD (1.8.7 or 1.9.1).

First, the "Tcl Console" was opened and the working directory was changed to a directory containing the movie generation scripts (within a general ACP simulation directory, containing all ACP ±Metal simulation folders). The script was then executed as indicated in Appendix 10.

rACP.tcl script contents:	# rACP is replaced rACP_Mg, rACP_Ni, rACP_Zn, SA, SA_Mg, SA_Ni, SA_Zn, SB, SB_Mg, SB_Ni, SB_Zn, SASB, SASB_Mg, SASB_Ni, or SASB_Zn, depending on which simulation is being loaded.
cd ..	#Leave directory containing movie generation scripts.
cd 14\ -\ rACP\ eq\ no\ Metal\ (Reg\ WS)/	#Enter rACP (no metal) simulation directory. This will change dependent on simulation being loaded.
mol load pdb rACP_E-A_ws.pdb	#Creates a new molecule in VMD using the PDB file rACP_E-A_ws.pdb.
mol addfile rACP_E-A_ws.psf	#Adds information to the new molecule using rACP_E-A_ws.psf.
wait 1	#Wait 1 s for loading to finish.
set id [molinfo top]	#Sets the variable "id" to the molecule loaded above.
wait 1	#Wait 1 s.
mol delrep 0 $id	#Removes the representation of the previously loaded molecule. With nothing to display loading of large DCD files is almost immediate; if the representation had been left on, each frame needs to be loaded (and displayed) before VMD can move on to the next frame.

wait 1	#Wait 1 s to ensure representation has been deleted.
cd 02\ -\ Smoothed\ Trajectory\ File/	#Enter directory containing the smoothed trajectory file (see Appendix 20 for detailed information on the script and the process used to smooth a trajectory).
mol addfile rACP_Smoothed_9.dcd waitfor all	#Adds trajectory information for the simulation to the loaded molecule. The "waitfor all" command after the DCD filename tells VMD to wait until all of the frames have loaded before proceeding with the script. rACP is replaced (as above) depending on simulation being loaded.
wait 1	#Wait 1 s.
animate delete beg 0 end 0	#Delete frame 0 (information from PDB and PSF files). This is done as this information is present in the smoothed DCD trajectory file.
cd ..	#Exit smoothed DCD file directory to root rACP (no metal) simulation directory.
cd ..	#Exit rACP (no metal) simulation directory to main ACP simulations folder.
mol rep NewCartoon	#Sets the representation style to NewCartoon. See (VMD Online Documentation. *Molecular Representations in VMD*.) for examples of representation styles.
mol selection protein	#Sets the representation selection to protein.
mol addrep $id	#Adds a new representation (with the above settings) for the selected atoms in the simulation.
	#Wait 1 s to ensure representation has been added.

```
wait 1
mol rep VDW                                  #Sets the representation style to VDW.
                                             See (VMD Online Documentation.
                                             Molecular Representations in VMD.) for
                                             examples of representation styles.

mol selection "not protein and not water"    #Sets the representation selection to atoms
                                             that are neither protein nor water.

mol material Opaque                          #Sets the representation material to
                                             Opaque.

mol addrep $id                               #Adds a new representation (with the
                                             above settings) for the selected atoms in
                                             the simulation.
cd 42\ -\ Create\ Movies\ of\ Metal\ Simulations\ (Script)/
                                             #Enter directory containing movie
                                             generation scripts.

rotate x by -55                              #Rotate the view by -55° on the X-axis.

rotate y by -30                              #Rotate the view by -30° on the Y-axis.
                                             #Positioning the view as done here in the
                                             script allows for a similar starting view
                                             for all ACP ±Metal simulations.

material change Specular Opaque 0.07         #Change Specular settings for Opaque
                                             materials to 0.07.

material change Shininess Opaque 0.07        #Change Shininess settings for Opaque
                                             materials to 0.07.

material change Specular Transparent 0.07    #Change Specular settings for
                                             Transparent materials to 0.07.

material change Shininess Transparent 0.07   #Change Shininess settings for
                                             Transparent materials to 0.07.
                                             #These changes in Specular and Shininess
                                             are necessary when saving POV-Ray files
.                                            as Raytracing software (like POV-Ray)
                                             handle light differently than the basic
                                             “Snapshot” rendering engine in VMD. If
                                             these changes were not made the BMPs
                                             resulting from rendering the POV-Ray
                                             files would have harsh reflections.
```

#A decrease in “Specular” and “Shininess” must be set prior to saving a POV file as raytracing software (such as POV-Ray) handles “light” in a different way than does the VMD Display window or the VMD “Snapshot” rendering engine. By reducing these values greatly (to ~0.07) greatly reduces reflection/shininess of final images produced by POV-Ray.

color Display Background white	#Changes the display windows background to white.
axes location off	#Turns the XYZ axis indicator off.
display resize 1280 720	#Re-size the display window to 1280x720 pixels (WxH).
scale by 1.4	# Zoom in 1.4x.
source trajectory_movie-POV3.tcl	#Load POV file generation script (Appendix 26).
cd I:\	#Change working directory to an external hard drive. Two empty lines are required after “cd I:\” to ensure the change in drives has been made.
cd Simulation\Movie\ Making\ Temp\ Directory/rACP/	#Enter Directory where rACP (no metal) POV-Ray files should be saved, rACP will change (as above) depending on which simulation is being saved to POV-Ray files.
wait 1	#Wait 1 s.
make_trajectory_movie_files	#This command executes the process in trajectory_movie-POV3.tcl (Appendix 26).
cd C:\	#Change working directory back to the C drive. As above two empty lines are needed to ensure drive change.
cd Documents/DAL/Atlantic\ Research\ Centre/ACP\ Mutant\ Simulations/42\ -\ Create\ Movies\ of\ Metal\ Simulations\ (Script)/	#Enter directory containing movie generation scripts.

```
mol delete $id                    #Deletes the current molecule. This
                                  prepares VMD for the next simulation to
                                  be loaded and processed.
```

APPENDIX 28 Sample Script for Loading Multiple ACP ±Metal Simulations Consecutively and Saving POV-Ray Files

Sample given is for rACPs. Portions of the script changed for different proteins highlighted (other options given after "#"). Script was written for and executed in VMD (1.8.7 or 1.9.1).

First, the "Tcl Console" was opened and the working directory was changed to a directory containing the movie generation scripts (within a general ACP simulation directory, containing all ACP ±Metal simulation folders). The script was then executed as indicated in Appendix 10.

Movie-Generation-rACPs.tcl script contents:

```
play rACP.tcl                          # Execute rACP.tcl (see Appendix 27 for detailed explanation of script contents)
wait 5                                 #Wait 5 s.
puts "rACP Movie Frames Exported (POV)"   #Outputs the text "rACP Movie Frames Exported (POV)" in the "Tcl Console". Having this text output gives an indication of the progress of the script.
play rACP_Mg.tcl
wait 5
puts "rACP+Mg Movie Frames Exported (POV)"
play rACP_Ni.tcl
wait 5
puts "rACP+Ni Movie Frames Exported (POV)"
play rACP_Zn.tcl
wait 5
puts "rACP+Zn Movie Frames Exported (POV)"
```

APPENDIX 29 MS-DOS Batch File for Rendering POV-Ray Files Using the 64-bit Build of POV-Ray 3.6.2

Batch File:

```
@ECHO Off
for /f %%a IN ('dir /b *.pov') do START /wait pvengine64 acp-sims.ini /EXIT %%a
```

Explanation of batch command:

@ECHO Off #"ECHO off" makes each command typed not visible in MS-DOS; in the case of this batch file, it keeps the execution of pvengine64 (Pov-Ray; 64-bit) from printing in the command window saving time and processing power
#the @ symbol before "ECHO off" means that the command applies for all lines of the batch file instead of just the one immediately following "ECHO off"

for /f %%a IN ('dir /b *.pov') do START /wait pvengine64 acp-sims.ini /EXIT %%a
#This is a basic for "item" in "a list" do programming command (the /f switch on for indicates that the list should be read line by line with the do command being executed after each line is read). The item is "%%a", the list is "dir /b *.pov" (dir is the directory list command; the /b switch tells MS-DOS to use bare format (no headers or summary just filenames; *.pov indicates that only files with the .pov extension should be contained in the list) and the command to execute (do) is "START /wait pvengine64 acp-sims.ini /exit" (START tells MS-DOS to execute pvengine64; acp-sims.ini tells Pov-Ray that the settings used to render each pov-ray file and the /EXIT switch tells pov-ray to close the window after each render (necessary to start the rendering the next frame).

Note: For this batch file to work, the directory containing the POV-Ray executable file (pvengine64) must be added to the "PATH" System Environmental Variable. If this was not done, the batch file and all of the POV-Ray files would have to be inside the POV-Ray install folder, and multiple renders could not proceed concurrently. To add a folder to the PATH variable on Windows Operating System, right click on computer and click properties. On the "Advanced" tab, click "Environmental Variables". Under system variables scroll down to PATH an click edit. Go to the end of the variable value and type ";"and paste the location of the POV-Ray executable (the folder containing the exe file; do not include the exe filename in the variable). Click "OK" until all windows are closed.

Note: Contents of ACP-Sims.ini can be found in Appendix 31.

Note: This batch file was only used on XPS-PC.

APPENDIX 30 MS-DOS Batch File for Rendering POV-Ray Files Using the 32-bit Build of Pov-Ray 3.6.2

Batch File:

```
@ECHO Off
for /f %%a IN ('dir /b *.pov') do START /wait pvengine acp-sims.ini /EXIT %%a
```

Explanation of batch command:

Explanations (and notes) from Appendix 29 apply here with the exception that (POV-Ray 32-bit build; pvengine) in executed instead of pvengine64 (POV-Ray 64-bit build).

Note: This batch file was only used on Inspiron2-PC and Desktop2-PC.

APPENDIX 31 Settings file (ACP-Sims.ini) Used to Render ACP ±Metal Simulation Frames with POV-Ray 3.6.2

Highlighted text preceded by "#" are annotations explaining individual lines.

```
;                    PERSISTENCE OF VISION RAY TRACER
;
;                         POV-Ray VERSION 3.6
;
;                             ACP-Sims.INI FILE
;
; Width of image in pixels.  Accepts integer values.
;
Width = 1280
;
; Height of image in pixels.  Accepts integer values.
;
Height = 720
;
;
;
;
;
;
```

#The width and height of the image will be determined by the size of the "VMD OpenGL Display" window. For the ACP ±metal simulations, 1280x720 was set (in addition to viewing angle, etc; see Appendix 27) in the script used to load each simulation to ensure all simulations were rendered at the same size.

```
Bounding_Threshold = 3
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
;
```

Sets the minimum number of objects in a scene before automatic bounding is used. Using automatic bounding decreases rendering time for scenes with many objects. Any protein scene exported from either VMD or Chimera will contain a large number of objects and therefore bounding is a useful option. For in-depth information on how/why POV-Ray uses bounding for large scenes see the POV-Ray Online Documentation (POV-Ray Online Documentation. *Automatic Bounding Control.*).

```
; Turn display Off
Display=Off                                 #Indicates to POV-Ray that
;                                           the render window should not
;                                           be displayed. This decreases
;                                           the time to render each frame.
; Turn verbose mode on
Verbose=On                                  #Setting verbose mode to on
;                                           tells POV-Ray to report on
;                                           the render in progress in the
;                                           main POV-Ray window. The
;                                           information reported by
;                                           POV-Ray when verbose
;                                           mode is on includes the line
;                                           number being rendered (of
;                                           total lines), the elapsed time
;                                           for the current frame, and
;                                           other basic information about
;                                           the render. The default is
;                                           “Off”, but having this feature
;                                           on is useful for tracking
;                                           render progress.
```

Note: Lines commencing with “;” are ignored by the software and are merely for explanations of the following command or for formatting purposes

APPENDIX 32 Creating Sausage Representations in MOLMOL

Sausage representations for linL46W and cycL46W protein simulations were created in MOLMOL 2.6.0. The script to generate these representations was generously provided by David Chan (Vogel Lab). The script is a modification of the script provided in the MOLMOL Tutorial (Spanish National Research Council. *MOLMOL Tutorial - Example 3*). Text within the MOLMOL script preceded by "#" are descriptions/titles for each section of the file.

The MOLMOL script below uses values from the temperature factor (B-factor) column of the PDB file for each Cα atom to determine the thickness of the sausage in the final diagram. Columns 61 – 66 in for any ATOM record of any PDB file represents the B-factor (Worldwide PDB. *PDB File Format Information: ATOM records*.).To use RMSF, these values must be entered into these columns of the PDB file being analyzed. This can be done manually, which can be time consuming, or with a simple python script (see Appendix 33).

Contents of generate_sausage_b_column_v2.0.mac:

```
# This macro should create a sausage diagram based on the values # in the B-factor
column of the pdb loaded. It's basically cut   # out from sausage.mac (David I. Chan).

# Select CA atoms of mean structure
SelectAtom 'name = "CA"'

# Set atom radius (which will be used as spline radius)
# to the B-factor, which was calculated as average
# displacement to the mean structure (RMSF)
RadiusAtom bfactor

# Add the spline
AddRibbon

# Use atom radius as spline radius
StyleRibbon as_is as_is as_is atom

# Restore old selection, with newly created spline added
SelectMol 'prev_sel | name = "mean"'
SelectRes 'prev_sel | mol.name = "mean"'
SelectAtom 'prev_sel | mol.name = "mean"'
```

To use generate_sausage_b_column_v2.0.mac, the following instructions can be used:

In MOLMOL, open the PDB to be processed. Then execute the script by typing "XMacStand sausage.mac" into the MOLMOL command line. Alternatively, the script can be executed under the File Menu (File->Macro->Execute Standard). This will

generate a sausage diagram based on RMSF values. An image can then be created normally (see the MOLMOL Tutorial).

APPENDIX 33 Python Script to Enter RMSF Data into Temperature Factor Section of a PDB File Prior to Sausage Diagram Creation

The Python script that enters the RMSF value for each Cα residue into the temperature factor column of a PDB file in preparation for sausage diagram creation (Appendix 32). Input files can be found below (see Appendix 34 for sample input/output files). Colors are indicative of syntax (comments, strings, constants, normal text, etc) in Wing IDE Personal 4.0.3-1 and above. Script was written using Python 2.7 with the help of Dr. Christian Blouin.

Python program:

```
'''
  Read in RMSF data per residue and enter that data into temperature factor (B-factor) columns of a PDB file.
'''

# Import necessary python libraries
import sys

# Tells the script to determine the input filenames from the command line used to execute the script
pdbfile = sys.argv[1]
datfile = sys.argv[2]

# Create an Empty List for the data and populate the list with the data
data = []
for i in open(datfile):
    data.append(i[:4])

# Set starting points for index (related to residue number)
currentAA = ''
index = -1

# Open input and output files
fin = open(pdbfile)
fout = open(pdbfile+'.out.pdb', 'w')

# For each atom line, determine residue number. If current amino acid = residue number then replace B-factor with RMSF value for current amino acid; iterate over all ATOM lines
for line in fin:
    if line.startswith('ATOM'):
        aa = line[24:26]
```

```
    else:
        fout.write(line)
        continue
    if aa != currentAA:
        index += 1
        currentAA = aa
    # New Line
    newline = line[:62] + data[index] + line[66:]
    fout.write(newline)

# Close input and output files
fout.close()
fin.close()
```

Script Execution:

Open a command prompt and navigate to the location of the data and script. Execute the script as below:

"python mapdata.py myfile.pdb myfile.txt"

Where myfile.pdb is replaced with the PDB to be processed and myfile.txt is replaced with the CSV file containing the RMSF pre residue data. Sample input and output files can be found below (see Appendix 34).

Note: For the "python" command to work, the directory containing the python executable file must be added to the "PATH" System Environmental Variable (see Appendix 29 for more information on how to do this in windows).

APPENDIX 34 Sample Input and Output Files for Python Script in Appendix 33

PDB ATOM Records for residue "1" of cycL46W in input PDB file (highlighted text is temperature factor (B-factor)):

```
ATOM      1  N   SER P   1      -3.863 -14.054  -6.836  1.00  0.00      ACP  N
ATOM      2  HT1 SER P   1       -3.858 -14.730  -7.573  1.00  0.00       ACP  H
ATOM      3  HT2 SER P   1       -2.999 -13.551  -6.829  1.00  0.00       ACP  H
ATOM      4  HT3 SER P   1       -4.735 -14.195  -7.350  1.00  0.00       ACP  H
ATOM      5  CA  SER P   1       -4.024 -14.754  -5.541  1.00  0.00       ACP  C
ATOM      6  HA  SER P   1       -4.999 -15.251  -5.553  1.00  0.00       ACP  H
ATOM      7  CB  SER P   1       -2.971 -15.859  -5.365  1.00  0.00       ACP  C
ATOM      8  HB1 SER P   1       -3.126 -16.354  -4.525  1.00  0.00       ACP  H
ATOM      9  HB2 SER P   1       -1.982 -15.393  -5.367  1.00  0.00       ACP  H
ATOM     10  OG  SER P   1       -3.036 -16.844  -6.393  1.00  0.00       ACP  O
ATOM     11  HG1 SER P   1       -2.869 -16.481  -7.297  1.00  0.00       ACP  H
ATOM     12  C   SER P   1      -4.029 -13.702  -4.420  1.00  0.00      ACP  C
ATOM     13  O   SER P   1      -4.068 -12.488  -4.755  1.00  0.00      ACP  O
```

PDB ATOM Records for residue "1" of cycL46W in output PDB file (highlighted text is temperature factor (B-factor)):

```
ATOM      1  N   SER P   1      -3.863 -14.054  -6.836  1.00  5.22      ACP  N
ATOM      2  HT1 SER P   1       -3.858 -14.730  -7.573  1.00  5.22       ACP  H
ATOM      3  HT2 SER P   1       -2.999 -13.551  -6.829  1.00  5.22       ACP  H
ATOM      4  HT3 SER P   1       -4.735 -14.195  -7.350  1.00  5.22       ACP  H
ATOM      5  CA  SER P   1       -4.024 -14.754  -5.541  1.00  5.22       ACP  C
ATOM      6  HA  SER P   1       -4.999 -15.251  -5.553  1.00  5.22       ACP  H
ATOM      7  CB  SER P   1       -2.971 -15.859  -5.365  1.00  5.22       ACP  C
ATOM      8  HB1 SER P   1       -3.126 -16.354  -4.525  1.00  5.22       ACP  H
ATOM      9  HB2 SER P   1       -1.982 -15.393  -5.367  1.00  5.22       ACP  H
ATOM     10  OG  SER P   1       -3.036 -16.844  -6.393  1.00  5.22       ACP  O
ATOM     11  HG1 SER P   1       -2.869 -16.481  -7.297  1.00  5.22       ACP  H
ATOM     12  C   SER P   1      -4.029 -13.702  -4.420  1.00  5.22      ACP  C
ATOM     13  O   SER P   1      -4.068 -12.488  -4.755  1.00  5.22      ACP  O
```

RMSF Input Data File (highlighted value used in above sample input/output PDB files):

```
5.2243332862854
4.497013568878174
3.7670469284057617
3.3950915336608887
3.5286357402801514
2.9388372898101807
1.9680670499801636
2.4783458709716797
2.688375949859619
```

1.9739489555358887
1.742240309715271
2.72348690032959
3.3609049320220947
2.962346315383911
2.3925459384918213
3.1952691078186035
3.5288727283477783
4.440845966339111
4.049095630645752
5.065846920013428
4.887911319732666
3.686123847961426
4.3002777099609375
4.195681571960449
4.54545259475708
4.670831680297852
3.7512781620025635
3.228506565093994
3.030280113220215
4.062873840332031
4.637818813323975
4.4419145584106445
4.17212438583374
3.4782371520996094
4.053452491760254
4.7344889640808105
3.985020399093628
3.5541467666625977
3.579050302505493
4.311070919036865
3.5892937183380127
2.6053218841552734
3.483659267425537
4.179526329040527
3.3272228240966797
3.0686609745025635
4.369908809661865
4.723226070404053
4.025538921356201
4.096055030822754
5.096864223480225
4.647741317749023
4.1278605461120605
3.392721176147461
3.8727564811706543

4.326551914215088
4.096563816070557
3.0348639488220215
2.9929184913635254
4.026630878448486
3.9492268562316895
3.104670286178589
2.3711819648742676
2.6182942390441895
2.520613670349121
3.410625457763672
3.0942752361297607
2.3500139713287354
3.0367629528045654
3.7028403282165527
3.523616313934326
3.6498630046844482
4.367389678955078
4.856426239013672
5.139342308044434
5.330323219299316
5.35849142074585
5.31622314453125
5.230905055999756
4.747288227081299

APPENDIX 35 Sample Script for Loading an ACP ±Metal Simulation and Saving an Equilibrated Structure

Sample given is for rACP (no metal). Portions of the script changed for different proteins highlighted (other options given after "#"). Script written for and executed in VMD (1.8.7 or 1.9.1). Based on the manipulations required, I provide the following instructions:

First, the "Tcl Console" was opened and the working directory was changed to a directory containing the equilibrated/averaged structure generation scripts (within a general ACP simulation directory, containing all ACP ±Metal simulation folders). The script was then executed as indicated in Appendix 10.

rACP.tcl script contents:	#rACP is replaced rACP_Mg, rACP_Ni, rACP_Zn, SA, SA_Mg, SA_Ni, SA_Zn, SB, SB_Mg, SB_Ni, SB_Zn, SASB, SASB_Mg, SASB_Ni, or SASB_Zn, depending on which simulation is being loaded.
cd ..	#Leave directory containing equilibrated/averaged structure generation scripts.
cd 14\ -\ rACP\ eq\ no\ Metal\ (Reg\ WS)/	#Enter rACP (no metal) simulation directory. This will change dependent on simulation being loaded.
mol load pdb rACP_E-A_ws.pdb	#Creates a new molecule in VMD using the PDB file rACP_E-A_ws.pdb.
mol addfile rACP_E-A_ws.psf	#Adds information to the new molecule using rACP_E-A_ws.psf.
wait 1	#Wait 1 s for loading to finish.
set id [molinfo top]	#Sets the variable "id" to the molecule loaded above.
wait 1	#Wait 1 second.
mol delrep 0 $id	#Removes the representation of the previously loaded molecule. With nothing to display loading of large DCD files is almost immediate; if the representation had been left on, each frame needs to be

```
                                        loaded (and displayed) before VMD can
                                        move on to the next frame.

wait 1                                  #Wait 1 s to ensure representation has
                                        been deleted.

cd 04\ -\ Equilibrated\ Structure\ Creation/   #Enter directory containing the over-
                                        smoothed (sliding window of 1001)
                                        trajectory file (see Appendix 20 for
                                        detailed information on the script and the
                                        process used to smooth a trajectory).

mol addfile rACP_Smoothed_1001.dcd waitfor all
                                        #Adds trajectory information for the
                                        simulation to the loaded molecule. The
                                        "waitfor all" command after the DCD
                                        filename tells VMD to wait until all of the
                                        frames have loaded before proceeding
                                        with the script. rACP is replaced (as
                                        above) depending on simulation being
                                        loaded.

mol addrep $id                          #Adds a new representation (with the
                                        default settings) for the atoms in the
                                        simulation. The default settings are:
                                        selection, all; color, name; style, lines.

set rACP [atomselect top "not water" frame #"]
                                        #Creates a selection, rACP, including all
                                        non-water atoms for frame #; "rACP" is
                                        replaced by the protein being saved (as
                                        above); "#" is replaced by a frame number
                                        near the end of the simulation where there
                                        are no improper dihedrals, atom positions,
                                        etc (e.g. 20501).

wait 1                                  #Wait 1 s.
$rACP writepdb rACP_No-Metal_equil.pdb
                                        #Write a PDB file containing atoms in
                                        selection "rACP" to "rACP_No-
                                        Metal_equil.pdb". As above, rACP is
                                        replaced by the protein being saved.

wait 1                                  #Wait 1 s to ensure PDB has saved.
```

```
mol delete $id                                  #Deletes the current molecule. This
                                                prepares VMD for the next simulation to
                                                be loaded and processed.

cd..                                            #Leave directory containing the over-
                                                smoothed (sliding window of 1001)
                                                trajectory file.

cd ..                                           #Leave rACP (no metal) simulation
                                                directory.
cd 32\ -\ Average\ Structure\ Generation\ Scripts/
                                                #Enter directory containing
                                                equilibrated/averaged structure generation
                                                scripts.
```