12-2014

The Structure and Function of Photosystem I and Photosystem I – Hydrogenase Protein Fusions: An Experimental and Computational Study

Bradley Jordan Harris
University of Tennessee - Knoxville, bharri17@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Bradley Jordan Harris entitled "The Structure and Function of Photosystem I and Photosystem I – Hydrogenase Protein Fusions: An Experimental and Computational Study." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Chemical Engineering.

Paul D. Frymier, Major Professor

We have read this dissertation and recommend its acceptance:

Eric T. Boder, Barry D. Bruce, Hugh O'Neill, Paul M. Dalhaimer

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

The Structure and Function of Photosystem I and Photosystem I – Hydrogenase Protein Fusions: An Experimental and Computational Study

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Bradley Jordan Harris
December 2014

DEDICATION

This thesis is dedicated to my wife, Kristen Harris, my parents, Tony and Tammy Harris, my younger brother, Westley Harris, as well as my entire family for their unwavering love, support and encouragement.

ACKNOWLEDGEMENTS

I would like to express my gratitude to many people for their guidance and support during my academic career at the University of Tennessee. First and foremost, I am deeply indebted to my advisor and mentor, Dr. Paul Frymier, for his enthusiastic guidance, tireless patience and support, approachability and constructive criticism throughout this work. Special thanks go to Dr. Eric Boder for his guidance and insight throughout my work, and for serving as a member of my committee. I would also like to thank Dr. Barry Bruce for his help and insight and for serving as a member of my committee. I would like to express my appreciation to Dr. Hugh O'Neill for his contributions and for serving as a member of my committee. My gratitude goes to Dr. Paul Dalhaimer for serving as a member of my committee. Special thanks also go to Dr. Xiaolin Cheng for mentoring me in a new research field and displaying patience and understanding as I continued to learn and develop.

I commend the contributions of Rosemary Le, Megan Farell, Vincent Price, Mariam Raaeszadeh-Sarmarzdeh, Richard Simmerman, and the entire Boder Lab and Bruce Lab members, towards the success of this work. Also a special thanks to my brother Westley Harris for sharing his expertise.

To all my friends in Knoxville, Frank, Ryan, Will, Morgan, Emily M., Emily C., and Jackson thank you for creating an enjoyable environment and being a friend to me through these stressful times.

Lastly, I am deeply indebted to my wife, Kristen Harris, for her love and encouragement, and for her patience in tolerating a husband that was often either working late or deeply engrossed in his work at home.

ABSTRACT

Photosystem I (PSI) is a membrane protein involved in the photosynthetic cycle of plants, algae, and cyanobacteria that is of specific interest due to its ability to harness solar energy to generate reducing power. This work seeks to form an *in vitro* hybrid protein fusion between the membrane integral PSI protein and the membrane-bound hydrogenase (MBH) enzyme, in an effort to improve electron transport between these two proteins.

Small-angle neutron scattering (SANS) was used to characterize the detergent-solubilized solution structure of trimeric PSI from the cyanobacterium *Thermosynechococcus elongatus*, which showed that the detergent interacts primarily with the hydrophobic periphery of PSI. The SANS results were used as a guide to constructing a model of trimeric PSI embedded in a detergent belt. Subsequent all-atom molecular dynamics (MD) simulations of the PSI-detergent complex suggested that the detergent environment could negatively impact the long-term stability of PSI, but is not likely to affect PSI activity or hinder its ligation to the MBH.

Having verified that the solution structure of the PSI-detergent complex will not affect formation of PSI-MBH fusions, the membrane-bound [NiFe]-hydrogenase of *Ralstonia eutropha* was genetically engineered to express a Gly_3 [Gly-Gly-Gly] tag on the N-terminus of the small subunit to allow for site-specific ligation to the psaE subunit of PSI. H_2 [hydrogen] uptake activity results show a complete loss of activity in the mutant *R. eutropha* strain, possibly due to mutations introduced during previous genetic engineering work. In parallel, MD simulations of the PSI-MBH fusion protein indicate this ligation strategy is not optimal for electron transport between these proteins. This MD approach can be used to evaluate other PSI-MBH fusion strategies, possibly targeting other stromal subunits of PSI. Finally, MD simulations of previously studied PSI-[FeFe]-hydrogenase fusions were conducted, revealing significant distortion of the protein structure that could limit their long-term stability.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I
INTRODUCTION

Proteins are large, complex molecules and their interactions within even the simplest of microorganisms are complex and intriguing processes. Gaining insight into proteins and their interactions is crucial to our understanding of the biological world, and could provide technological advancement in a wealth of areas. Here, we are interested in studying hybrid protein fusions of photosystem I (PSI) and hydrogenase enzyme, a unique complex capable of light-driven H_2 evolution *in vitro*, in order to improve electron transport between these proteins. This chapter will include introductory information on PSI and hydrogenase, as well as a discussion of previous research that has motivated this study. The scope of work and specific goals of this study will then be discussed.

1.1 Background

Proteins mediate many biological functions and can possess a diverse set of features, including overall pH-dependent charge, hydrophobic/hydrophilic interactions, highly reactive moieties, metal complexes, etc., all of which affect their interactions with other proteins, inorganic molecules, and surfaces. Membrane proteins are of particular interest, as the membrane surrounding a biological cell is responsible for mediating all cell-environment interactions, and therefore contains a myriad of proteins involved in such processes as cellular recognition, signal transduction, and the transportation of ions and molecules across the membrane. In eukaryotic cells, the internal compartments are also surrounded by their own membranes, each containing their own unique set of membrane proteins. Furthermore, many bacteria have an outer membrane surrounding the inner cell membrane, which provides protection from the environment and is distinctly different in composition from the inner membrane. As a result, membrane proteins comprise approximately 30% of open reading frames, highlighting their biological importance [1]. And yet they represent only a small, although rapidly increasing, fraction of known protein structures, reflecting

the difficulties of both membrane protein expression and crystallization [2]. This problem is partially due to the need for solubilization of the transmembrane surface via a detergent or lipid assembly. These detergents disrupt the native cell membrane and solubilize the membrane proteins in mixed micelles that prevent aggregation by mimicking the native membrane environment [3, 4]. Unfortunately, little is known of the aqueous structure and protein-detergent interactions of solubilized membrane proteins, and how these detergents may interfere with the structural and functional properties of the protein [5, 6].

Photosynthesis is an interesting process that occurs in the thylakoid membranes of plants and organisms, wherein light energy is converted to chemical energy via a macromolecular complex of several metalloproteins (Figure 1) [7], and the efficient transfer of electrons between these dissimilar proteins is accomplished through very specific interactions. Oxygenic phototrophs such as green plants, algae and cyanobacteria transfer electrons generated via water splitting through photosystem II (PSII), to the cytochrome b_6f (Cyt b_6f) complex, a redox mediator, then to photosystem I (PSI), and eventually to ferredoxin (Fd) where it is used by Fd-$NADP^+$ reductase (FNR) to reduce $NADP^+$ to NADPH [8]. The combined activity of PSII, PSI, and the electron

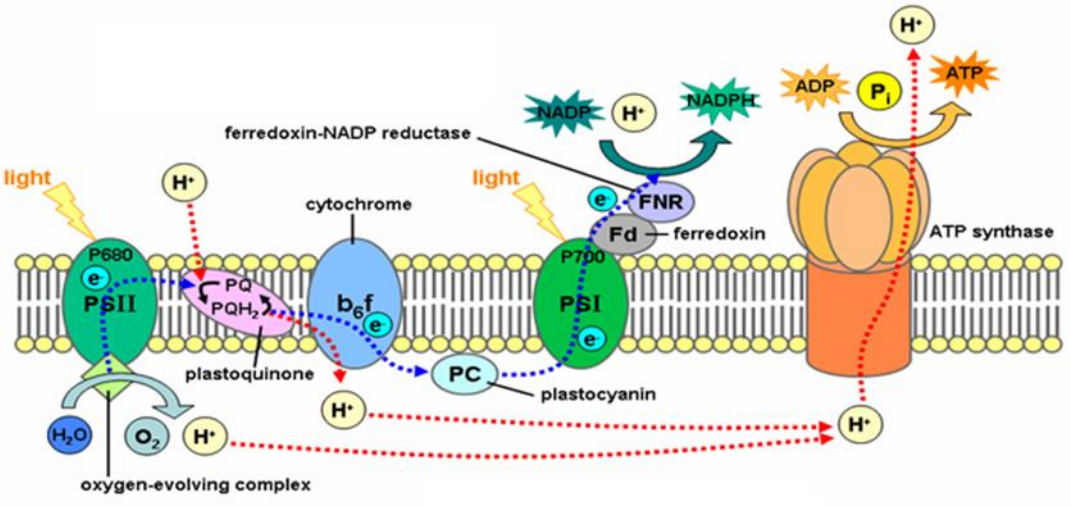

Figure 1. Photosynthetic electron transport chain (Figure adapted from Taiz and Zeger 1998)

transfer chain (ETC) results in the reduction of $NADP^+$ and the transfer of protons across the thylakoid membrane. The translocated protons drive ATP synthesis and the resultant ATP and NADPH drive CO_2 fixation through the Calvin cycle.

Photosystem I (PSI) is one of the three membrane-bound metalloproteins involved in cyanobacterial, algal, and plant photosynthesis. It provides both large antennae for harvesting solar energy and a reaction center for accomplishing stable charge separation. In the cyanobacterium *Thermosynechococcus elongatus* (*T. elongatus*), PSI is located in the thylakoid membrane and exists primarily as a homotrimer of the 12-subunit monomer, which contains several membrane-spanning domains. The crystal structure of trimeric PSI from the thermophilic cyanobacterium *T. elongatus* has been resolved at 2.5 Å (Figure 2) [9] as a large complex (~1 MDa) containing the reaction center pigments, light harvesting chlorophyll *a* (chl *a*), carotenoids, quinones, and iron-sulfur clusters that mediate electron transport to support carbon fixation and other redox reactions [10]. The PsaA and PsaB proteins (shown in green and blue, respectively) form the photoreactive core of the protein, where excitation of the primary electron donor, P_{700}, leads to charge separation stabilized by spatial displacement through a series of redox centers. The electron from excited P_{700} is transferred to the primary acceptor A_0, a Chl *a* monomer, then to the phylloquinone molecule A_1, and onto the [4Fe-4S] cluster F_X. From there, the electron travels through F_A and F_B, the [4Fe-4S] clusters bound to PsaC (shown in red), the terminal acceptor located on the stromal side that donates electrons to Fd. PsaD and PsaE (shown in orange and cyan, respectively) are believed to facilitate docking of and interaction with Fd. On the oxidizing side, PsaF (shown in magenta) forms a docking site for cytochrome c_6. PsaI, PsaJ, PsaK, PsaL, PsaM and PsaX (shown in yellow, olive, wheat, teal, lime and pink, respectively) coordinate the Mg^{2+} ions of antenna Chl *a* molecules directly or through water molecules. In addition, PSI exists naturally as a homotrimer, and PsaL is believed to facilitate this trimerization process [11].

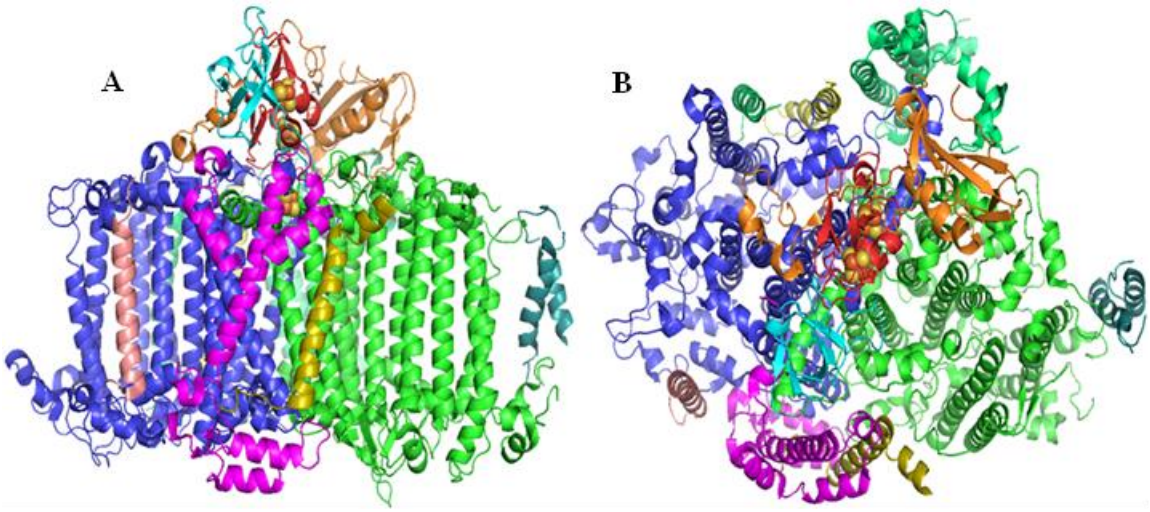

Figure 2. The crystal structure of monomeric PSI from *T. elongatus* (PDB ID: 1JB0): (A) side view along the membrane plane with the stromal side up; (B) top view from the stromal face

PSI accomplishes the unique biological function of converting solar energy into reducing power, and has been the subject of many studies investigating *in vitro* applications for alternative energy solutions, such as hydrogen or direct electricity production as part of devices or materials [12-21]. Greenbaum et al. demonstrated that PSI vectorial arrays prepared on derivatized gold surfaces could function as a biomolecular photovoltaic device, and furthermore that the nature of the derivatized surface (i.e. negative, positive, or hydrophilic) influenced PSI orientation [17]. Researchers continue to investigate various surfaces and PSI deposition methods in an effort to understand the factors that influence PSI surface coverage and orientation [19-22]. Greenbaum and co-workers also demonstrated that colloidal platinum could be precipitated directly onto photosynthetic thylakoid membranes in aqueous solution, and that the resulting suspension was capable of sustained photoevolution of hydrogen and oxygen [12, 23]. This concept has since been extended to isolated PSI, in which the hexachloroplatinate salt is precipitated directly onto its reducing end, thereby substituting negatively charged $[PtCl_6]^{2-}$ for negatively charged Fd and enabling light-driven hydrogen evolution [13, 14, 24].

Another possibility for hydrogen evolution via photosynthetic microorganisms involves the coupling of PSI with a hydrogenase enzyme. Hydrogenases are metalloproteins that are found in Archaea, obligate and facultative anaerobic bacteria, and some eukaryotes that catalyze the reversible oxidation of molecular hydrogen. Hydrogenases can be largely classified into three groups based on the metals present in the active site: (1) the nickel-iron [NiFe]-hydrogenases, (2) the iron only [FeFe]-hydrogenases, and (3) the iron-sulfur-free [Fe]-hydrogenases. These three classes of hydrogenases catalyze the same reaction using very different metal active sites, and are found to be phylogenetically distinct and utilize a completely different set of regulatory and assembly processes [25]. The crystal structures of both the [NiFe]-hydrogenase and the [FeFe]-hydrogenase have been determined, and reveal that these two enzymes are structurally different [26].

The [NiFe]-hydrogenases are heterodimeric periplasmic membrane enzymes. The large subunit (Figure 3A, shown in green) houses the nickel-iron active site and electrons are delivered to this center via three iron-sulfur [FeS] clusters located in the small subunit (Figure 3A, shown in blue) [27]. The [NiFe] active site binds four coordinating cysteine residues, two carbon monoxide (CO) ligands, and one cyanide (CN) ligand. [FeFe]-hydrogenases possess only [FeS] clusters

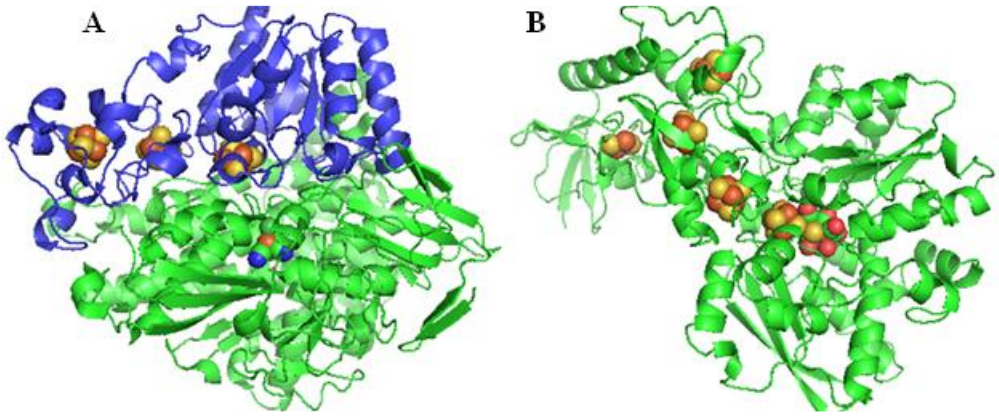

Figure 3. (A) Crystal structure of the [NiFe]-hydrogenase from *Ralstonia eutropha* (PDB ID: 3RGW); (B) Crystal structure of the [FeFe]-hydrogenase from *Clostridium pasteurianum* (PDB ID: 1FEH)

and an iron-iron active site (Figure 3B). The [FeFe] active site binds one coordinating cysteine residue, three CO, two CN, and two sulfur ligands. [FeFe]-hydrogenases are irreversibly inactivated by the presence of oxygen while [NiFe]-hydrogenases are reversibly inactivated [28]. [Fe]-hydrogenases (not shown), as opposed to the [NiFe]- and [FeFe]-hydrogenases, lack iron-sulfur clusters and are only found in a small group of methanogenic archaea [29]. The active site contains a labile light sensitive cofactor [30] with a single iron, most likely Fe(II), which binds two CO ligands [31-33].

The majority of hydrogenases, which operate under anaerobic conditions *in vivo*, are inactivated by trace amounts of O_2. The membrane-bound hydrogenase (MBH) of *Ralstonia eutropha* (*R. eutropha*), however, belongs to the class of so-called “oxygen-tolerant” [NiFe]-hydrogenases which can exhibit high H_2 oxidation activity for sustained periods under aerobic growth. The crystal structure for the MBH of *R. eutropha* has been solved (Figure 3A) [34], and confirms experimental evidence suggesting the conserved presence of two supernumerary cysteine residues around the [FeS]-cluster proximal to the active site. These results support a model for O_2 tolerance of [NiFe]-hydrogenases that stresses that the enzyme, upon attack by O_2, can immediately furnish all four electrons needed to avoid reactive oxygen species that damage or block the active site [35].

Several photosynthetic organisms possess native hydrogenases of the NiFe class, and can evolve hydrogen gas under anoxic conditions. Unfortunately, *in vivo* cyanobacterial hydrogen production falls well below the theoretical maximum due to two factors: (1) the general oxygen sensitivity of hydrogenases and the specific tendency of [NiFe]-hydrogenases to favor the hydrogen uptake reaction, and (2) the electron transfer from PSI to hydrogenase must compete with electron transfer to the Calvin cycle [36]. As a result, numerous studies have investigated *in vitro* systems in which hydrogenase enzymes are coupled to PSI to enable light-driven H_2 production [15, 16, 37, 38].

As PSI-based technologies have great potential in the field of energy conversion, it is important to develop robust, reproducible techniques to control

the orientation of PSI molecules on conductive substrates or in complexes with other proteins. Better knowledge of protein-detergent interactions is crucial to retaining activity and addressing solution-phase aggregation for biophotovoltaic applications, in order to avoid denaturation and/or random orientation of PSI [39]. These interactions could also prove important in the engineering of genetic modifications, such as binding motifs necessary for highly specific binding of PSI to surfaces or in complexes with other proteins, in order to ensure they are not hindered by the presence of detergent. Furthermore, studying the effects of solution dynamics and protein-protein interactions in PSI-hydrogenase fusions could yield new insights into electron transport in redox proteins. Overall, better understanding of protein-detergent and protein-protein interactions could prolong the lifetime of PSI-based energy conversion devices by helping to reduce unbound or inactive protein.

1.2 Scope of this work

In this work, the objective was to form an *in vitro* hybrid protein fusion between the membrane integral photosystem I (PSI) protein and the membrane-bound hydrogenase (MBH) enzyme in an effort to improve electron transport between these two proteins. Because PSI is a membrane-spanning protein and requires solubilizing detergents to isolate it from its native environment, we began by studying the solution structure of this protein to understand the size and shape of the overall PSI-detergent complex, and the orientation of the detergent around PSI. This will help us to understand if the presence of detergent will hinder the ligation of the MBH to the stromal surface of PSI. Therefore, the structure of trimeric PSI from *T. elongatus* stabilized in *n*-dodecyl-β-*D*-maltoside (DDM) detergent solution was investigated using small-angle neutron scattering (SANS) [40]. Scattering curves were collected at 18% D_2O, the contrast match point for the detergent, and 100% D_2O, allowing characterization of the structures of both the protein and protein/detergent complex. We found that the maximum dimension of the PSI/DDM complex was consistent with a monolayer belt of

detergent around the hydrophobic transmembrane domain of PSI, and a dummy atom reconstruction of the complex based on the SANS data shows that the detergent envelope has an irregular shape around the periphery of PSI rather than a uniform, toroidal belt. These results indicate that the presence of detergent will not hinder ligation of the MBH to the stromal surface of PSI, as detergent molecules interact primarily with the transmembrane periphery of the PSI protein.

We then constructed a computational model of the PSI/DDM complex for comparison to the experimental SANS data, and to gain atomic-level insight into the structural changes occurring in PSI in the detergent environment. This was necessary because SANS is a low-resolution technique and cannot tell us if detergent-solubilized conditions induce changes in the atomic structure of PSI that could impact its stability and activity, or obstruct its ligation to the MBH. We performed 200 ns molecular dynamics (MD) simulations of the PSI/DDM complex for in-depth analysis of the structure and dynamics of the PSI trimer embedded in a detergent environment as well as the molecular level protein-detergent interactions [41]. A theoretical scattering curve based on this model provided a reasonable fit to the experimental SANS data. Evaluation of root-mean-squared deviation (RMSD) relative to the known crystal structure shows that the protein is stable within this complex, and root-mean-squared fluctuation (RMSF) analysis indicates regions of high local mobility correspond to solvent-exposed regions such as turns in the transmembrane α-helices and flexible loops on the stromal and lumenal faces. Comparing the protein-detergent complex to a pure detergent micelle, the detergent is less densely packed and the detergent tail regions are more ordered based on analysis of solvent-exposed surface area, dihedral angle order parameters, and diffusion and fluctuation of the detergent. We found that the observed conformational dynamics and protein-detergent interactions also have functional implications, resulting in interesting structural changes in the domains associated with maintaining the trimeric structure of the protein; however, the docking of soluble electron mediators such as cytochrome c_6 and

ferredoxin was not impacted. These findings indicate that the detergent environment could have an adverse effect on the long-term stability of PSI, but that it is unlikely that the detergent will impact PSI activity or induce structural changes that would hinder its ligation to the MBH.

Having determined that the solution structure of the PSI/DDM complex is not likely to obstruct formation of PSI-hydrogenase fusions, we were interested in generating fusions of PSI and the MBH enzyme using a simpler, more direct approach than currently established methods. Using genetically engineered PSI from *Synechocystis* sp. PCC 6803 and the MBH from *R. eutropha*, we planned to generate site-specific fusions of these proteins via sortase-mediated ligation (SML) as a general, straightforward approach to studying electron transport in PSI-hydrogenase protein fusions. In this work, MBH from *R. eutropha* was genetically engineered to express a Gly_3 tag on the N-terminus of the small subunit. This mutation was re-introduced to *R. eutropha* and verified by DNA sequencing. Isolated membrane fractions and solubilized membrane extracts were purified from both the wild-type and mutant *R. eutropha* strains, which exhibited significant differences in terms of MBH localization and H_2 uptake activity. Therefore, lacking experimental data on PSI-MBH fusions, we performed MD simulations on previously developed nanoconstructs of PSI attached to an [FeFe]-hydrogenase (FeFe H_2ase) from *Clostridium pasteurianum* (*C. pasteurianum*), which are known to have high electron transfer rates, to see if we could provide an atomic-level explanation for differences in electron transfer observed in these fusion complexes experimentally. This approach can then be applied to the study of fusions of PSI and MBH in order to identify an optimal linkage strategy for electron transport between these proteins and guide experimental construction of PSI-MBH fusion proteins. We have found that longer linker lengths result in greater structural drift and require longer simulation times to reach a stable complex, likely due to the increased conformational space that complexes with longer linkers are able to explore. We also noted unfavorable interactions occurring in the PSI-FeFe H_2ase fusions that may result

in significant strain in these complexes. We then conducted similar MD simulations on a model of the PSI-MBH protein fusion that we were attempting to generate experimentally, and found that conditions were not optimal for electron transport between PSI and the MBH.

CHAPTER II
LITERATURE REVIEW

The focus of this research is the generation of PSI-hydrogenase fusion complexes for improved electron transport between these proteins. In order to ensure that detergent-solubilized solution conditions will not hinder the formation of PSI-hydrogenase fusions, we need an experimental method for studying membrane protein structure in solution. Here, we will review several methods for studying membrane protein structure experimentally, and highlight why SANS is ideal for our study of the solution structure of PSI. Because SANS is a low-resolution technique, we also need to conduct atomic-level analysis of PSI to ensure that any structural changes in this protein resulting from the detergent environment will not hinder its ligation to hydrogenase. Therefore, we will also discuss how MD simulation can be used to study the dynamics of membrane proteins and multi-protein complexes. Lastly, we also provide an overview of the current state of research on PSI-hydrogenase protein fusions.

2.1 Scattering techniques for studying proteins

Experimental scattering techniques take advantage of the diffraction properties of incident beams of X-rays, electrons, neutrons, etc. to provide valuable structural information for a wide range of substrates, including biomacromolecules. For the purposes of this literature review, we will focus on the following methods as they pertain to the study of membrane proteins and multi-protein complexes: (1) X-ray crystallography, (2) nuclear magnetic resonance (NMR) spectroscopy, and (3) small-angle scattering (SAS), including small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS).

2.1.1 X-ray crystallography

X-ray crystallography is a technique that utilizes the measured angles and intensities of a diffracted beam of incident X-rays upon a crystallized substance to produce a three-dimensional picture of the electron density within the crystal, thus providing information on the atomic and molecular structure of the

substance. A broad range of molecular dimensions (1-1000 Å) can be studied, although high resolution becomes more difficult to achieve as the size of the molecule increases. The proton pump bacteriorhodopsin was the first membrane protein whose low resolution structure was solved by Richard Henderson and Nigel Unwin in 1975 [42], and subsequently the 3 Å resolution structure of the photosynthetic reaction center of the purple bacterium *Rhodopseudomonas viridis* (*R. viridis*) was solved by Deisenhofer et al. in 1985 [43]. In the 29 years since, more than 450 unique structures of membrane proteins have been solved, resulting in over 850 publications [2]. The 1.9 Å resolution structure of bacteriorhodopsin from *Halobacterium salinarum* (*H. salinarum*), solved using X-ray crystallography by Pebay-Peyroula and colleagues in 1999 [44] (Figure 4), showed that the transmembrane α-helix is the fundamental structural motif of plasma membrane proteins. Outer membrane proteins, in contrast, generally have a β-barrel motif [45, 46]. Furthermore, X-ray crystal structures have revealed that proteins involved in proton and electron transport, such as bacteriorhodopsin, form very tight bundles that are nearly impermeable to water, whereas ion/molecule transporters have large, water-filled cavities that span nearly the entire membrane [47-49].

For one such ion transporter, the potassium channel KscA from *Streptomyces*

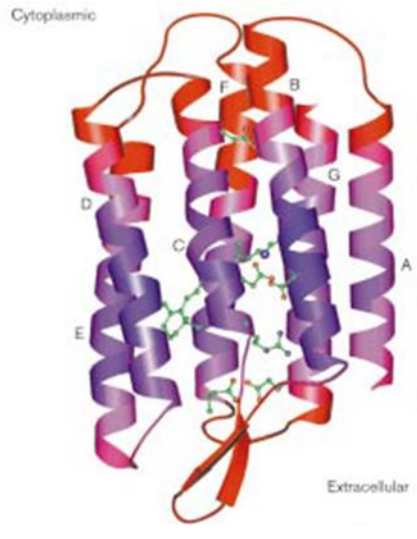

Figure 4. Crystal structure of bacteriorhodopsin (Figure adapted from Edman et al. 1999)

lividans, solving the X-ray crystal structure showed that four identical monomeric subunits create an inverted cone with a 12 Å long selectivity filter at its outer end. The remainder of the pore was found to be wider and lined with hydrophobic amino acids, with a large water-filled cavity and helix dipoles positioned so as to overcome electrostatic destabilization of an ion in the pore at the center of the bilayer. Overall, the architecture of the pore was found to support selective K^+ conduction [47].

X-ray crystallography has also been used to solve ligand-bound protein structures. The crystal structure of the bovine mitochondrial ADP/ATP carrier in complex with the inhibitor carboxyatractyloside revealed six α-helices in a compact transmembrane domain with a deep depression at the surface towards the space between the inner and outer mitochondrial membranes. This structure revealed a possible conformational change as transport substrates bind to the bottom of the cavity, resulting in a transient transition from a ‘pit’ to a ‘channel’ conformation [48].

In a similar case, the crystal structure of the molecular transporter *Escherichia coli* (*E. coli*) lactose permease LacY was solved by Abramson et al. in 2003 [49]. They found this protein to be comprised of symmetrically positioned N- and C-terminal domains, each with six transmembrane helices, with a large internal hydrophilic cavity. By solving the structure with a bound lactose homolog, they were able to locate the sugar-binding site in the cavity and identify residues involved in substrate recognition and proton translocation (Figure 5). Polar surfaces are colored blue (positively charged) and red (negatively charged); the black spheres represent the lactose homolog.

Coupled transporters are another important class of membrane proteins, so-called because they couple the energetically costly movement of one solute to the energetically favorable movement of another. In the case of sodium-coupled transporters, the energy gained from the movement of sodium ions down an electrochemical gradient is used to concentrate substrates, such as aspartate [50] and leucine [51]. X-ray crystallography has revealed that these proteins display pseudo-two-fold symmetry and also exhibit deep water penetration. Their two-fold symmetry allows them to operate as 'rocker switches', alternating access to the two sides of the membrane; these features have been found to be shared by a broad range of transporters [52].

Another class of coupled transporters is referred to as G-protein-coupled receptors (GPCRs). These couplers exhibit the tightly packed 'waterproof' bundle typical of proton/electron transporters, consisting of seven transmembrane α-helices, a remarkably conserved motif also found in bacteriorhodopsin. GPCRs receive an optical or chemical signal on the extracellular surface which initiates signaling cascades in the cytoplasm, and form the largest single class of eukaryotic membrane proteins [53]. The first GPCR crystal structure to be solved

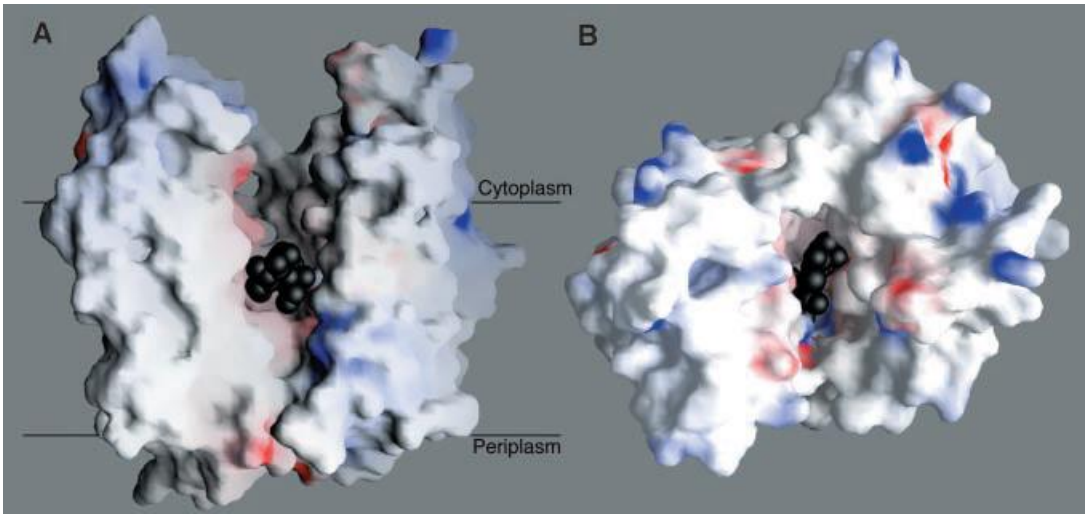

Figure 5. The internal hydrophobic cavity of the *E. coli* lactose permease LacY. (A) View parallel to the membrane; (B) View along the membrane normal from the cytoplasmic side. (Figure adapted from Abramson et al. 2003)

was that of bovine rhodopsin bound to its agonist, the ground-state chromophore 11-*cis*-retinal, by Palczewski et al. [54]; rhodopsins are a member of the largest subfamily of GPCRs, and activated by light to turn on the signaling pathway that leads to vision. This structure identifies key residues whose interactions with the 11-*cis*-retinal chromophore facilitate color discrimination, and also gives information on the mechanism of GPCR activation.

The case of the human β_2 adrenoceptor (β_2AR) GPCR reflects the difficulty of membrane protein crystallization. This and other GPCRs for diffusible hormones and neurotransmitters are conformationally complex, exhibiting basal activity that appears to cause structural instability. To overcome this difficulty, β_2AR had to be crystallized when bound to an inverse agonist and in complex with a monoclonal antibody fragment, allowing new insights into this little studied subfamily of GPCRs [55].

Adenosine triphosphate (ATP) synthase is an interesting example of a proton transporter. This protein produces ATP from adenosine diphosphate (ADP) and inorganic phosphate via energy from respiration or photosynthesis. The enzyme consists of an extramembranous F_1 catalytic domain linked by means of a central stalk to a transmembrane domain F_0. The globular F_1 domain is known to consist of five different subunits with stoichiometry $\alpha_3\beta_3\Upsilon_1\delta_1\varepsilon_1$. The X-ray crystal structure of a subcomplex of this protein extracted from yeast mitochondria shows that the F_0 domain is composed of ten transmembrane two-helix c subunits (Figure 6), and the extensive contact between this ring and the stalk suggests they may rotate as an ensemble during catalysis [56]. The proton motive force generated as protons pass down the electrochemical gradient through F_0 drives the F_1 rotor to produce ATP.

The crystal structure of *T. elongatus* PSII, the first membrane-bound protein complex in the light-dependent reactions of oxygenic photosynthesis (Figure 1), was first solved at 3.8 Å resolution in 2000 by Zouni et al. [57], and further refined to 2.9 Å by Saenger and co-workers in 2009 [58]. This protein captures photons of light to energize electrons provided through water-splitting, transferring these electrons across the membrane to reduce the soluble electron mediator plastoquinone to plastoquinol; protons generated in this process are also shuttled across the membrane by ATP synthase to generate reducing equivalents in the form of ATP. These structures provided insight into the arrangement of the PSII subunits and cofactors, but could not refine the detailed structure of the water-splitting catalytic center. In 2011, Kamiya and colleagues reported the 1.9 Å resolution crystal structure of PSII from the related thermophilic cyanobacteria *Thermosynechococcus vulcanus* (*T. vulcanus*) [59]. They were able to locate all of the metal atoms of the Mn_4CaO_5 cluster, the water-splitting catalytic center, together with all of their ligands. They found five oxygen atoms that served as

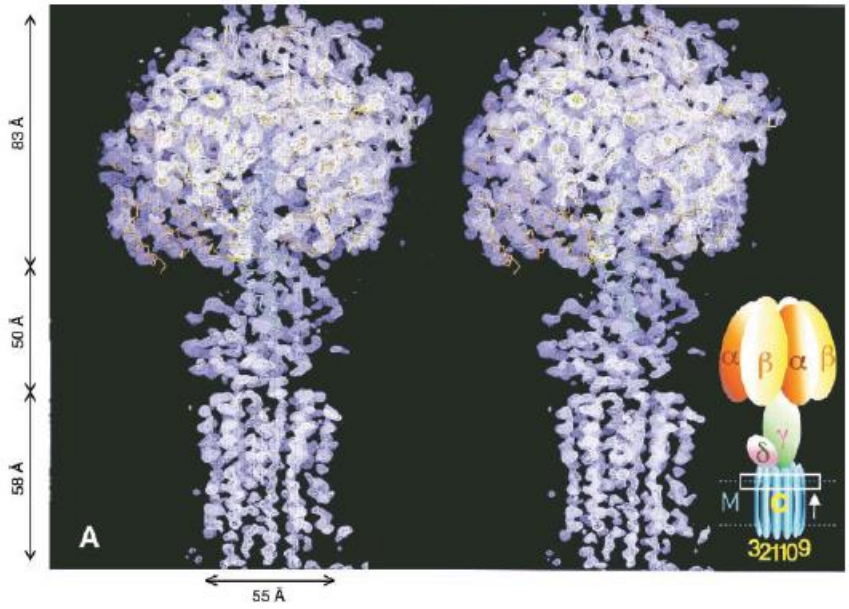

Figure 6. Side view of electron density map of the yeast F_1c_{10} complex. Inset is a model of this enzyme with F_1 subunits α in orange, β in yellow, ϒ in green, ε in pink, and the c subunits of F_0 in blue. (Figure adapted from Stock et al. 1999)

oxo bridges linking the five metal atoms, and found four water molecules bound to the cluster that may serve as substrates for dioxygen formation. They further identified 1,300 water molecules associated with each PSII monomer that form hydrogen-bonding networks that could serve as channels for the reactants and products of the water-splitting process. This information provides key insight on the function of this photosynthetic protein.

The crystal structure of *T. elongatus* PSI, one of the focuses of this study, was solved at 2.5 Å resolution by Jordan et al. in 2001 (Figure 2) [60]. This protein is the second photosystem in the photosynthetic light reactions, using light energy to mediate the transfer of electrons from plastocyanin to ferredoxin (Figure 1). These electrons will eventually be used by ferredoxin-$NADP_+$ reductase to generate reducing equivalents in the form of NADPH. The overall generation of NADPH and ATP from photosynthesis is used to drive carbon fixation in the Calvin cycle. This crystal structure provides atomic level detail of trimeric PSI and the 96 chlorophylls and 22 carotenoids that coordinate with this complex, serving as light-harvesting antennae, as well as the two phylloquinone and three iron-sulfur cluster cofactors that serve as the electron transport chain through this membrane-bound protein. Interestingly, they also were able to identify four lipids associated with PsaA/PsaB. The locations of two of these lipids close to the core to PSI, and the binding of antenna chlorophyll to a third, indicate that they are functionally important in the PSI complex and not mere preparation artifacts [60].

In rare instances, X-ray crystal structures of multi-protein complexes have also been obtained. In 2011, Rasmussen et al. solved the X-ray crystal structure of a ternary complex composed of nucleotide-free Gs heterotrimer, the stimulatory G protein for adenylyl cyclase, and agonist-occupied monomeric β_2AR [61]. In order to crystallize this complex, they had to mix purified GDP-Gs with a molar excess of purified β_2AR bound to a high-affinity agonist in DDM solution, and subsequently add apyrase to hydrolyze GDP released from Gs during complex formation with β_2AR. This crystal structure represented the first atomic resolution view of transmembrane signaling by a GPCR.

Whorton and MacKinnon recently reported the 3.5 Å resolution crystal structure of the mammalian G-protein-gated inward rectifier K^+ (GIRK) channel in complex with βϒ G-protein subunits, the central signaling complex that links GPCR stimulation to K^+ channel activity (Figure 7) [62]. As was the case for Rasmussen et al., their efforts to purify a stable GIRK-Gβϒ complex in detergent solution were unsuccessful. Instead, crystallization of the complex in DDM required combining individually purified GIRK and Gβϒ proteins at twofold to threefold molar excess of Gβϒ and in the presence of tenfold molar excess of the lipid PIP_2. They found the resultant structure to be consistent with previous explanations of GIRK activation by G proteins, and gained new insights into understanding how PIP_2 and intracellular Na^+ ions participate in multi-ligand regulation of GIRK channels.

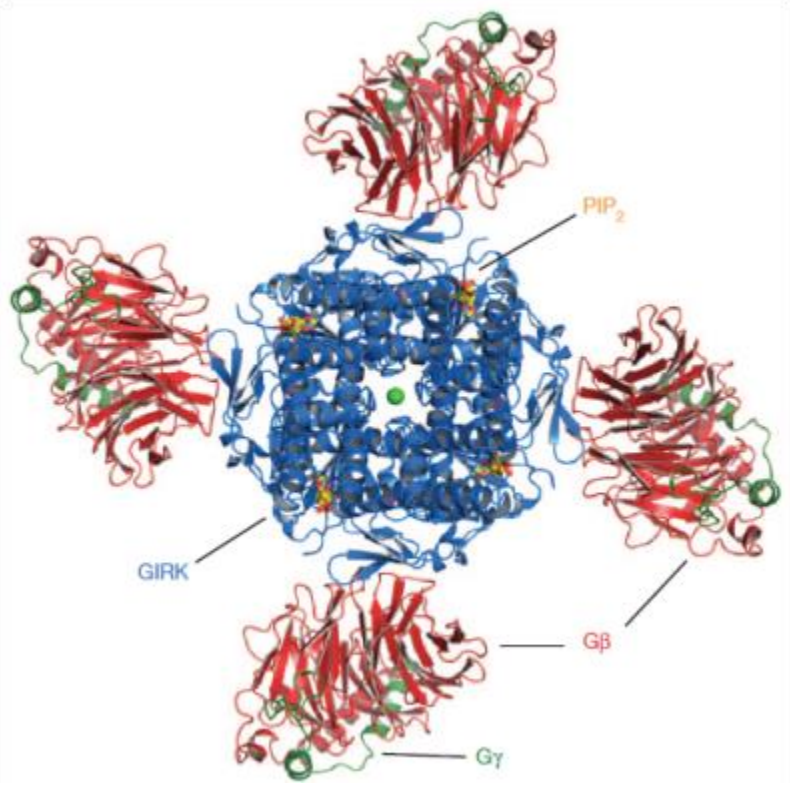

Figure 7. Overall structure of the GIRK-Gβϒ complex. Top-down view of GIRK (blue), Gβ (red) and Gϒ (green). (Figure adapted from Wharton and MacKinnon 2013)

In summary, X-ray crystallography is capable of providing not only atomic level detail of the overall structure of proteins, but also key insights into transport pathways, ligand binding, unique conformational changes, and the chemical makeup of catalytic binding sites. However, there are several limitations to this method as well. The case of β_2AR illustrates the complicated measures sometimes needed to overcome the difficulties of membrane protein crystallization. X-ray crystallography also provides little to no information on protein-lipid and protein-detergent interactions, as most lipid/detergent molecules are either lost in solution or are not sufficiently ordered to be resolved [63]. Bacteriorhodopsin and PSI serve as rare cases in which information regarding associated lipids has been identified. Similarly, crystallization of multi-protein complexes presents a difficult challenge, with technological advances only recently enabling determination of X-ray structures of multiple associated proteins such as G-protein complexes. Perhaps most importantly, X-ray structures provide only a static snapshot of the protein structure at cryogenic temperatures required for crystallization, failing to capture the dynamics of conformational changes and also leaving open the question of whether or not they are representative of the protein structure under typical conditions.

2.1.2 Nuclear magnetic resonance (NMR) spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a research technique that exploits the magnetic properties of certain atomic nuclei to provide detailed information on the structure, dynamics, reaction state, and chemical environment of molecules. For example, a simple NMR spectra can provide information on the degree of folding of a protein, its thermal and temporal stability, and its aggregation propensity. Unlike X-ray crystallography, NMR is a technique that can be applied both in solid-state and in solution, and approaches have been developed to probe proteins that have proven difficult for X-ray diffraction methods, such as inherently flexible or membrane-spanning proteins. Solution NMR enables the study of membrane protein-detergent complexes, whereas solid-state NMR is used to study the structure and dynamics of molecules without

isotropic mobility, such as membrane proteins in lipid bilayers. Furthermore, the analysis of protein interactions, both with small molecule ligands or peptides and with other proteins, is a particular strength of NMR [64].

One of the earliest NMR studies of membrane proteins was conducted by Brown et al., wherein they used solid-state NMR to characterize interactions between the photoreceptor protein rhodopsin and phospholipids in the retinal rod outer segment disc membrane [65]. More recent studies have focused on the use of solid-state NMR to investigate retinal dynamics during light activation of rhodopsin [66-70], and determine how retinal isomerization choreographs intricate changes of the transmembrane helices and loops [71-74]. Results from solution NMR studies of GPCRs have been relatively modest due to difficulties stemming from the size and complexity of these proteins. However, a recent study was conducted on bacterial phototaxis sensory rhodopsin, a GPCR structural analog, in dihexanoylphosphatidylcholine (DHPC) micelles [75], and preliminary studies of CC chemokine receptor 5 and the thromboxane A2 receptor have been reported [76-78].

NMR techniques have also advanced our understanding of ion channel proteins, such as the potassium channel protein KcsA. Baldus and co-workers examined the conformational dynamics of KcsA under different pH, K+ concentration, and inhibitor binding conditions using a combination of electrochemical experiments and solid-state NMR [79-82]. They identified chemical shift changes in the selectivity filter residues that corresponded to transitions of the channel from the high pH closed state to the low pH inactivated state, and found that the conformational changes of the selectivity filter and intracellular activation gate are coupled. Furthermore, they were able to show that two known channel blockers inhibit KcsA in different ways.

Another interesting channel protein, the M2 proton channel of the influenza A virus, has also been extensively investigated using NMR [83]. Schnell and Chou used solution NMR to study tetrameric M2 in DHPC micelles, carefully examining changing in protein dynamics as a function of pH and identifying a potential

mechanism for acid-induced activation of the channel (Figure 8) [84]. Similarly, Hong and co-workers used solid-state NMR to investigate the conformation and dynamics of the M2 inhibitors amantadine and spiro-piperidine [85-87], and even attained atomic-level insight into its proton conduction mechanism by measuring the protonation state, rotameric structure, side-chain dynamics, and pH-dependent hydrogen bonding of the crucial proton-selective His37 residue [88].

NMR studies have also been conducted on multi-protein complexes. In 1997, Ubbink et al. used NMR to characterize a complex of plastocyanin and cytochrome c [89], and studied of a variety of similar transient redox protein complexes have since been reported [90-93]. More recently, Shimada and co-workers used a combination of isothermal titration calorimetry and solid-state NMR to study the multi-protein GPCR complex of GIRK and Gβϒ, finding that the Gβϒ binding site spans two neighboring subunits of the GIRK tetramer. This work yielded new insights into conformational rearrangements between GIRK monomers and a potential mechanism for the binding of these two proteins and the gating of GIRK [94].

In summary, NMR is a powerful experimental technique for studying biomacromolecules that enables analysis of the structure and dynamics of membrane proteins, both in lipid bilayers (solid-state NMR) and in detergent micelles (solution NMR), as well as multi-protein complexes. However, although technological advances continue to be made, the size and complexity of structures that can be studied using NMR is limited. Solution NMR becomes increasingly difficult for complexes over 30 kDa [95], and while solid-state NMR can be used to study biomacromolecules larger than 100 kDa, the abundance of anisotropic features present can potentially limit the features observable in the spectra [96].

2.1.3 Small-angle scattering (SAS)

Small-angle scattering (SAS) techniques, using X-rays or neutrons, provide another approach for investigating the structural properties of proteins. In contrast to X-ray crystallography and NMR, which provide atomic resolution

structural details of substrates, SAS techniques involve the analysis of small-angle (0.1-10°) deflections of collimated radiation to yield lower resolution (10-10,000 Å) information on the size and shape of materials in solution. Although it provides limited data on internal structure, SAS is ideal for characterizing large-scale shape changes of macromolecules and is a powerful tool for studying the structures of biomacromolecules and their complexes in solution [97].

SAS measurements can be conducted via either small-angle X-ray scattering (SAXS) or small-angle neutron scattering (SANS). The key difference is that X-rays are scattered by electrons, therefore X-ray scattering amplitudes increase monotonically with the number of electrons. In contrast, neutrons are scattered by atomic nuclei with no systematic dependence on atomic number, resulting in unique properties that will be further explained later in this section.

An early SAXS study of bovine rhodopsin in complex with the detergent dodecyldimethylamine oxide gave a first look at how solubilizing detergents interact with membrane proteins, finding that rhodopsin existed in a compact, monomeric form with the detergent forming a monolayer belt around the membrane-spanning portions of the structure [98].

A comprehensive study of three membrane proteins in detergent conducted by Heller et al. serves as an example of the interpretation of scattering data and how theoretical curves can be generated for comparison based on known crystal structures. In it, they studied the detergent-solubilized conformations of bacteriorhodopsin (bR) from *H. salinarum* in octyl glucoside (OG), the Ste2p GPCR from the yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) in DDM, and the *E. coli* porin OmpF in octyl polyoxyethylene (OPOE) [99]. OG and OPOE were the detergents used for previously conducted X-ray crystallography studies of bR and OmpF, respectively; the high-resolution structure of Ste2p had not yet been determined. SAXS data sets are typically represented by plotting the intensity distribution (I) versus the scattering vector length q, where $q = 4\pi(\sin\theta)/\lambda$, 2θ is the scattering angle between the incident and scattering beam, and λ is the wavelength. They found this curve to be consistent with the expected compact

conformation in the case of OmpF, and that the data was well fit by a theoretical scattering curve generated for a model of the known crystal structure embedded in disk-like detergent micelle. In the case of the seven transmembrane α-helical (7TM) proteins bR and Ste2p, however, the SAXS data sets were not consistent with a natively folded protein surrounded by a belt of detergent. These results demonstrated that detergent-solubilized membrane proteins may not be properly folded with their hydrophobic segments encapsulated in a belt of detergent as is typically envisioned, but rather that the detergent may instead induce a non-native conformation.

SAS data can also be used to reconstruct a low resolution shape of the scattering particles in solution using *ab initio* modeling tools [100], thus providing an average structural envelope for a given system. In 2004, Hong et al. used SAXS to study the integral membrane protein light-harvesting complex LH2 from the photosynthetic bacteria *Rhodobacter spheroides* (*R. spheroides*) in an *n,n* dimethyldodecylamine *n*-oxide (LDAO) detergent solution [101]. Using *ab initio* shape determination methods, they found that the shape of detergent-solubilized LH2 clearly deviated from the previously reported crystal structure, deforming from a ring structure in the crystal to an ellipsoid in detergent solution (Figure 8).

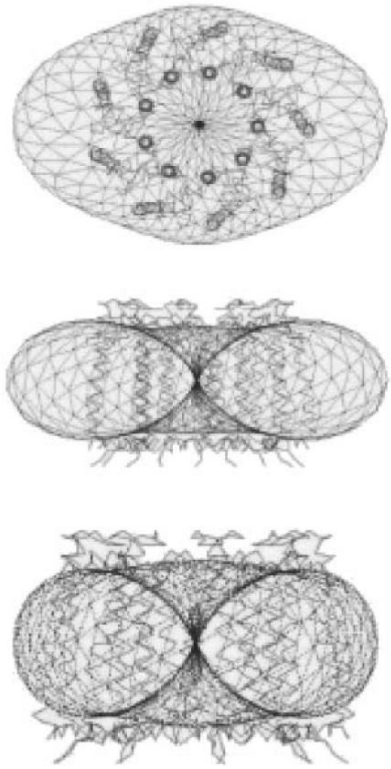

Figure 8. Restored shape envelope of LH2 complex with crystal structure (Figure adapted from Hong et al. 2004)

Greenbaum et al. used SAXS to investigate the structure of isolated spinach PSI complexes stabilized in detergent solution [102]. They investigated two different types of purified PSI preparations: in the first, thylakoid membranes were solubilized using Triton X-100 detergent and purified by density gradient centrifugation; in the second, PSI was isolated as before but was further purified by anion exchange chromatography, resulting in isolated PSI-detergent complexes. Data analysis and model fitting indicated that the first preparation resulted in large scattering objects or microphases with ~68 Å thickness, which is in reasonable agreement with a typical membrane thickness (~60 Å) and the transmembrane thickness of PSI (80-90 Å), and virtually infinite lateral extension. These results suggested a sheetlike structure similar to a bilayer membrane.

SAXS can also be used to study multi-protein complexes. In 2008, Xu et al. demonstrated the use of a combination of SAXS and solution NMR to study the

formation of the short-lived complex of adrenodoxin (Adx) and cytochrome c (Cc) [103], using a cross-linked counterpart to study the conformational freedom of the native complex. Based on molecular envelope reconstructions of the SAXS data, they found that Cc sampled approximately half the surface area of Adx in the native complex. Interprotein paramagnetic effects were observed in NMR studies of the cross-linked system, while extensive averaging occurred in the case of the native complex, consistent with a sampling of multiple orientations.

Small-angle neutron scattering (SANS) is another SAS technique useful for studying the solution structures of biomacromolecules. The advantage of using neutrons over X-rays, as discussed previously, is that the amplitude of X-ray scattering is proportional to the atomic number, whereas neutron scattering is not. This means that neutron scattering lengths of heavier elements can be of similar magnitude to that of lighter elements, and there can be large isotope effects; the most prominent of these is that of hydrogen vs. deuterium [97]. This unique property allows for solvent or contrast matching, wherein the H_2O/D_2O ratio of the solvent is tuned to match the scattering density of one or more components of a complex, thus masking its scattering contribution and allowing study of those components with different scattering densities. In terms of solubilized membrane proteins, one could imagine matching out the scattering contribution of the detergent in order to study the solution structure of the protein without disturbing the overall protein-detergent complex.

One of the earliest SANS studies of membrane proteins was conducted on bovine rhodopsin by Chabre and colleagues, in which they used contrast matching and perdeuteration of the detergent in order to eliminate the detergent-solvent contrast and study the structure of rhodopsin within a detergent micelle [104]. Santonicola et al. used contrast variation SANS to enable measurement of the composition of bacteriorhodopsin-detergent complexes and determine the thickness of the detergent shell bound to the protein [105]. Using a similar approach in their study of the structure of spinach light-harvesting complex II stabilized in *n*-octyl-β-*D*-glucoside (BOG) detergent solution, Cardoso et al. used

SANS with contrast variation to investigate the properties of the protein-detergent complex at 15% D_2O (BOG match point), 42% D_2O (protein match point) and 100% D_2O. They found the restored topological shape of LHCII based on their data to be consistent with the previously determined crystal structure, and further demonstrated the use of a model LHCII/BOG complex to interpret the 100% D_2O data (Figure 9) [106].

Kendall and co-workers showed that SANS can be used to investigate conformational changes induced by ligand binding in their study of the *E. coli* membrane protein SecA in lipid vesicles [107]. A proposed role of SecA is that of membrane protein insertion through ATP hydrolysis, and they found that in the absence of nucleotide SecA exists in dimeric form, while in contrast the presence of either ADP or a non-hydrolyzable ATP analog induces conversion to a monomeric form.

SANS techniques have also been used to characterize multi-protein systems. Comoletti et al. demonstrated the usefulness of selective perdeuteration in their

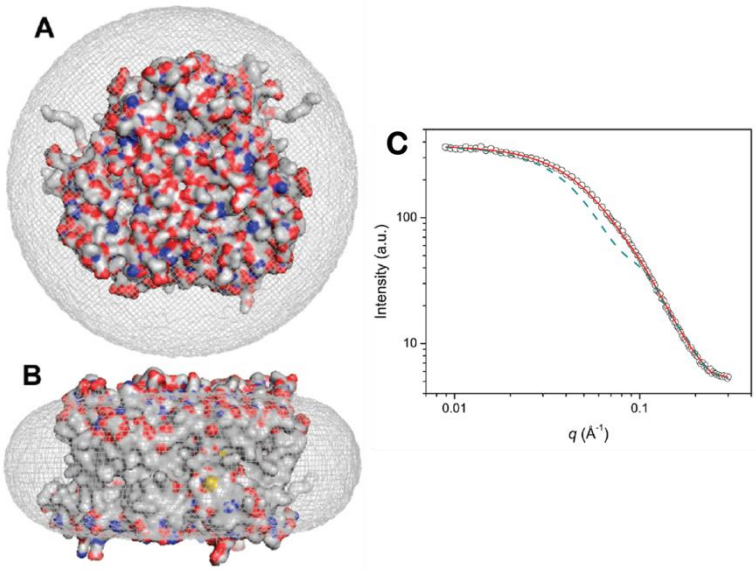

Figure 9. (A) Superior and (B) lateral views of LHCII protein surface embedded in a dislike detergent micelle belt. (C) SANS data from the LHCII/BOG mixture in 100% D_2O (open circles) with the theoretical scattering curve for the belt model (dashed green line). (Figure adapted from Cardoso et al. 2009)

analysis of the postsynaptic cell-adhesion protein neuroligin associated with its presynaptic partner neurexin. By collecting SANS data at 42% D_2O on the neuroligin/neurexin complex with protonated neuroligin and deuterated neurexin, they were able to show that two neurexin monomers bind at symmetric locations on opposite sides of the long axis of the neuroligin dimer. This allowed development of structural models of different neuroligin domains and their partnering molecules, providing a framework for understanding altered recognition by these proteins in neurodevelopmental disorders [108].

To summarize, SAS techniques offer advantages for the study of proteins compared to X-ray crystallography in that they enable determination of the time-averaged structure of the protein complex in solution rather than a static structure under crystallization conditions, and can resolve the scattering contribution of the detergent to provide structural information on the protein-detergent complex as a whole. Furthermore, SAS techniques can be used to study large protein-detergent micelles and multi-protein complexes in solution, whereas size is a limitation for NMR studies. Experimental SAS data can provide insight on structural details of the protein and can be compared to theoretical curves generated based on model fits or known structures, and *ab initio* shape restoration techniques are available for determining the structural envelope of the protein complex. SANS is a particularly useful tool for studying membrane proteins, as the unique properties of neutrons enable the use of contrast matching and/or selective perdeuteration to mask the scattering contributions of individual system components such as the detergent [106] or particular protein subunits [108] without disturbing the overall protein-detergent complex. The drawback is that these techniques are low resolution (10-10,000 Å), preventing atomic level analysis of protein-detergent interactions and their effects on the structure of the protein of interest. Also, SAS studies of solubilized membrane proteins typically require a structural model of the protein-detergent complex in order to interpret the experimental data.

2.2 Molecular dynamics simulations for in-depth analysis of protein-environment, protein-protein interactions

Molecular dynamics (MD) is a computer simulation of the trajectories of atoms and molecules determined by numerically solving Newton's equations of motion for a system of interacting particles, where the potential energy of the system and the forces between particles are determined by molecular mechanics force fields. MD simulation has proven to be a useful computational technique for the study of proteins and other biomacromolecules [109-111], including detergents, lipids, and membrane proteins [112]. Theoretical scattering curves can also be generated based on MD models for comparison to experimental SAS data [113]. For the purposes of this review, we will focus on recent developments in the use of MD simulation for the study of (1) membrane proteins in lipid bilayers, (2) detergent-solubilized membrane proteins, and (3) multi-protein complexes.

2.2.1 MD simulations of membrane proteins in lipid bilayers

MD simulation has been used for over a decade to study the structural properties and conformational dynamics of pure lipid bilayers under a variety of conditions including constant pressure [114-116], constant volume [117, 118], and constant surface tension [119-121]. Recent technological advances in computation have enabled all-atom MD studies of several classes of membrane proteins that span lipid bilayers, such as porins [122-126], ion channels [127-133], G-protein-coupled receptors (GPCRs) [134-136], and even PSII [137-139].

One of the earliest of these studies was that of the *E. coli* OmpF porin in a palmitoyl-oleoyl-phosphatidylethanolamine (POPE) bilayer conducted by Tieleman and Berendsen in 1998 [122]. Their system consisted of an OmpF trimer embedded in a bilayer of 318 POPE lipids solvated in water (Figure 10). After equilibration and a short nanosecond MD production run, they found the OmpF trimer to be stable with C_α RMSD of 0.23 nm relative to the crystal structure, and also noted that regions of high local mobility corresponded to loops and turns of this β-barrel protein based on RMSF analysis. They further

examined properties of the OmpF channel protein, including pore size, conductance, and the diffusion of water inside the pore.

MD simulations of ion channel proteins in lipid bilayers, a large class of proteins linked to a wide range of biological processes, further demonstrate the unique insights that computational techniques can provide concerning protein dynamics and function. Berneche and Roux performed a 4 ns MD simulation of the potassium (K^+) channel KcsA from the bacterium *Streptomyces lividans* in a dipalmitoylphosphatidylcholine (DPPC) bilayer, during which they observed a concerted dynamic transition of three K^+ ions in the pore. They determined this transition to be coupled to fluctuations of known selectivity filter residues within the channel, suggesting that a mechanical response of the channel structure could play an important role in K^+ transport. They further observed that a single

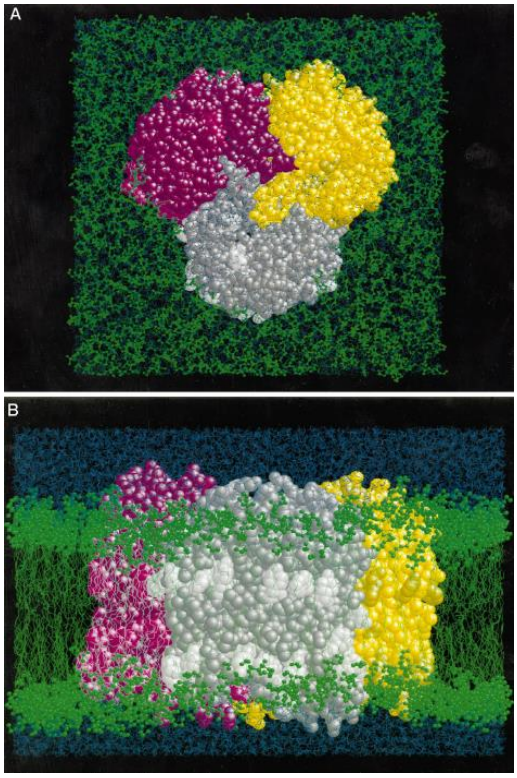

Figure 10. Visual depiction of OmpF in a POPE bilayer after MD simulation: (A) View perpendicular to the membrane; (B) Transmembrane view. Monomer one is colored grey, two yellow, and three magenta, with lipids in green and water in blue. (Figure adapted from Tieleman and Berendsen 1998)

water molecule was stabilized between each pair of ions during this transition, which agreed with experimental measurements [129]. Similar studies have found that the exit of the K^+ ion from the channel is enabled by a transient increase in the diameter of the intracellular mouth of the pore, suggesting that such "breathing" motions may form the molecular basis of experimentally observed channel gating [130]. A recent MD simulation of KcsA in a lipid bilayer by Jensen et al. has demonstrated the transition of the K^+ channel from activated to deactivated states on longer μs timescales [131].

Recent MD studies of GPCRs provide new insights on how protein-lipid interactions can affect the structure and function of membrane proteins. In a 15 ns MD simulation conducted on the GPCR rhodopsin with covalently bound retinal in a 2-oleoyl-1-palmitoyl-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer, Huber et al. found that the lipid environment induced significant fluctuations in the conformational state of a cytoplasmic loop C2 that is preceded by the highly conserved ERY motif essential to receptor activation. They found that changes in the conformation of this C2 loop could alter the protonation state and hydrogen-bonding environment of the ERY motif, thus suggesting a possible mechanism for lipid effects on rhodopsin function [134]. More recently, Kobilka and colleagues used a combination of nuclear magnetic resonance (NMR) spectroscopy and MD to show that in the case of β_2AR, unlike rhodopsin, the link between the agonist and the GPCR is flexible, allowing for conformational heterogeneity that may be important for β_2AR's ability to engage multiple signaling and regulatory proteins [136].

MD simulations have also been used to study the photosynthetic reaction center PSII in lipid bilayers. In 2006, Vasil'ev and Bruce conducted a 1 ns MD simulation of PSII in a pre-equilibrated lipid bilayer to study how protein dynamics affects energy transfer and electron transport efficiency in photosynthesis [137]. Combining MD with quantum mechanical (QM) calculations, they were able to accurately predict the experimental absorbance of the PSII core complex, whereas QM calculations based on the x-ray structure did not. In subsequent

work, Bruce and co-workers used MD to identify channels for water entrance to the water-splitting catalytic center of PSII, also known as the oxygen-evolving complex (OEC). They were able to identify a system of branching pathways of water diffusion in PSII leading to the OEC that connected to distinct entrance points on the lumenal surface [138]. A recent 10 ns MD study of PSII embedded in the thylakoid membrane conducted by Ogata et al. has further suggested that the entrance of water and subsequent exit of protons and oxygen from the OEC occur through distinctly different pathways [139].

2.2.2 MD simulations of membrane proteins in detergent micelles

Purification of membrane proteins from their native environment typically requires the use of detergents that disrupt the native cell membrane and solubilize the membrane proteins in mixed micelles that prevent aggregation by mimicking the native membrane environment [3, 4]. MD studies have been conducted on a variety of detergent micelle types including anionic [140, 141], cationic [142], non-ionic [143, 144], and zwitterionic [145-147]. MD simulations have also been used to study the solution structure and dynamics of several membrane protein-detergent complexes.

Braun et al. conducted 2.5 ns MD studies of wild-type and mutant glycophorin A (GpA), an α-helical membrane protein, embedded in a pre-formed sodium dodecyl sulfate (SDS) micelle. They also performed a 32 ns MD simulation showing the formation of a complete micelle around wild-type GpA from an initially random placement of SDS molecules, the first such study of spontaneous lipid organization in the presence of protein, and found its properties to be indistinguishable from that of the pre-formed protein-detergent micelle [148]. Similarly, Sansom and co-workers conducted 50 ns MD simulations of pre-formed and spontaneously formed protein-detergent micelles for both GpA and the β-barrel membrane protein OmpA in dodecylphosphocholine (DPC) detergent (Figure 11) [149]. They found the end structures for self-assembled OmpA and GpA to be very similar to their preformed counterparts, with the time constant for micelle formation for OmpA being about double that of GpA. These works

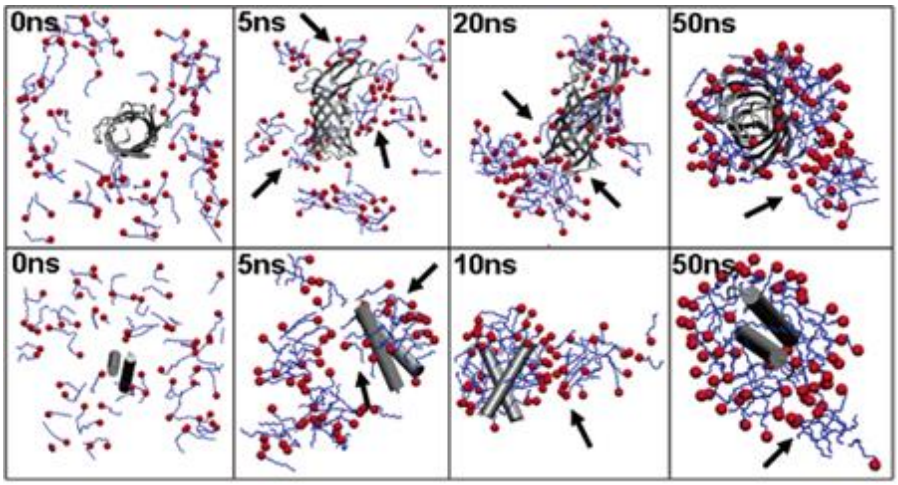

Figure 11. MD snapshots of spontaneous protein-micelle formation for OmpA (top) and GpA (bottom) (Figure adapted from Bond et al. 2004)

validated the established protocols for generation of protein-micelle models, demonstrating their ability to capture the solution structure of protein-detergent complexes and analyze protein-detergent interactions.

MD has also been used to study differences in membrane proteins in a lipid bilayer versus a detergent micelle. In 2004, Bond and Sansom performed two 10 ns MD simulations of OmpA: one in a DPC micelle and one in a dimyristoyl-phosphatidylcholine (DMPC) bilayer. They found the dynamic fluctuations of the protein structure to be ~1.5 times greater in the micelle environment than in the lipid bilayer. They further identified subtle differences in the nature of OmpA-detergent and OmpA-lipid interactions that may explain experimentally observed channel formation by OmpA [150]. In 2012, Rodriguez-Ropero and Fioroni performed 50 ns MD simulations of the membrane channel protein FhuA from *E. coli* in both a lipid bilayer and a detergent micelle. They found significant differences in protein structural and dynamical behavior between the two environments, in particular noting enhancements of the ellipticity of the open state and its fluctuations in the detergent environment that could have implications for further engineering and optimization of FhuA in nanodelivery or nanoreactor technology [151].

MD can also be used in conjunction with experimental techniques for conformational studies of transmembrane proteins in solution, as demonstrated by Koutsioubas et al. [152]. In order to better interpret experimental SAXS data of aquaporin-0 tetramers in DDM detergent solution, they constructed a model based on a 10 ns all-atom MD simulation of the aquaporin-0 tetramer embedded in a corona of DDM molecules. This MD model closely resembled the expected elliptical toroid shape of the protein-detergent complex, and was able to accurately fit the SAXS data and reproduce experimentally-determined parameters such as the radius of gyration and detergent aggregation number.

2.2.3 MD simulations of multi-protein complexes

Fully unrestrained computational MD simulations of the formation and dynamics of multi-protein complexes is not yet entirely feasible in many cases, due to the computational cost of simulating such large systems and the relevant timescales that need to be attained. However, there have been a few computational approaches developed to study multi-protein systems that artificially bypass these restrictions.

In 2006, Bui and McCammon used targeted MD (TMD) simulation to study protein complex formation of the neurotoxin fasciculin-2 (FAS2) and its high-affinity binding protein, acetylcholinesterase (AChE) [153]. In TMD, a subset of atoms in the simulation is guided towards a final target structure by means of steering forces. In this case, TMD was used to describe the binding of AChE to FAS2 using the previously reported crystal structure of the FAS2/AChE complex as a target [154, 155]. Bui and McCammon found that formation of this complex followed a pathway that begins with an encounter complex of one of the apo forms of FAS2 and AChE, followed by rapid conformational rearrangements into an intermediate complex that subsequently converts to the final complex as observed in crystal structures [153]. They also found that FAS2 transitions were significantly faster in the presence of AChE, with the energy barrier between the apo and liganded conformational states being reduced by half.

In 2008, Gsponer et al. demonstrated the use of restrained ensemble-averaged MD simulations to study the free and bound states of calmodulin (CaM). Using structural information and relaxation data obtained via NMR as constraints, they studied the structure and dynamics of CaM in the calcium-bound state (Ca^{2+}-CaM) and in the state bound to myosin light chain kinase (CaM-MLCK) [156]. They found that calcium binding to CaM directs structural fluctuations towards complex-like substrates, enabling initial ligation of MLCK to the C-terminal domain of CaM, which subsequently facilitates binding to the N-terminal domain. These results corroborated previous experimental studies that indicated cooperative binding of MLCK [157], and suggest that a coupled equilibrium shift mechanism controls the efficient binding of CaM to a wide range of ligands (Figure 12) [156].

In summary, MD simulations offer atomic level information on the structure and

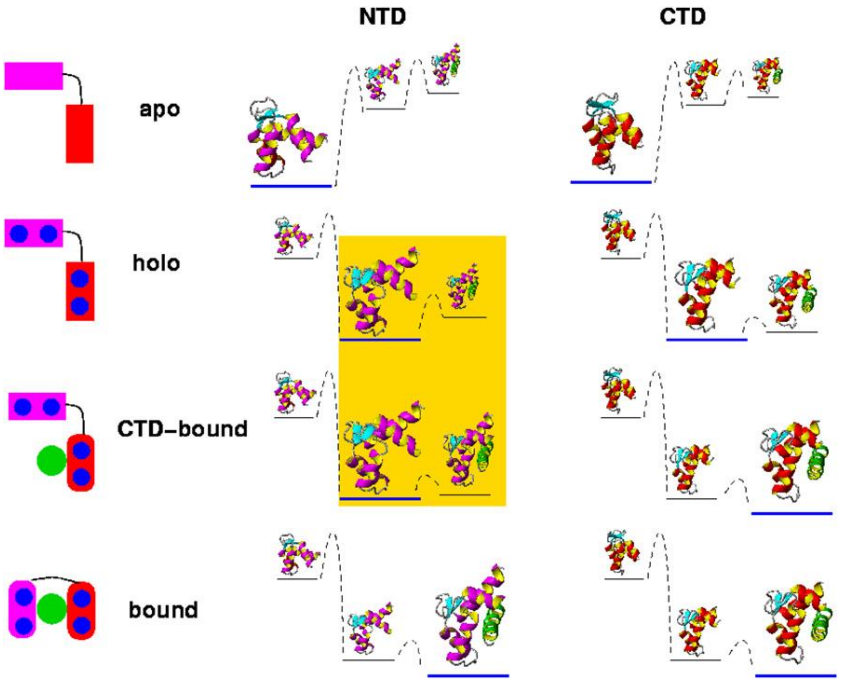

Figure 12. Free energy landscape of CaM illustrating cooperative binding mechanism. Binding of calcium ions shifts equilibrium from the apo to the holo state (Ca^{2+}-CaM). Binding of MLCK then takes place first at the C-terminal domain, which increases the likelihood of binding at the N-terminal domain (highlighted in yellow). (Figure adapted from Gsponer et al. 2008)

conformational dynamics of membrane proteins, whether in a lipid bilayer environment or under detergent-solubilized conditions. MD offers the ability to study the stability and local mobility of these proteins in complex with a variety of lipid or detergent moieties, and in many cases has yielded unique insights into membrane protein function. As membrane proteins mediate a myriad of biologically important processes, and solubilization of these proteins from their native lipid environment via a detergent or lipid assembly is a necessary step in their purification, computational techniques such as MD simulation that can identify the impact of protein-detergent interactions on protein structure and function will be essential to furthering our understanding of the biological world. Furthermore, recent developments have also enabled MD studies of the structure and dynamics of multi-protein complexes. However, these methods typically require artificial constraints, and there are limits concerning the size of the complex to be studied and the timescales attainable.

2.3 PSI-hydrogenase fusions as a novel system for studying multi-protein complexes

As has been stated previously, gaining insight into protein-environment and protein-protein interactions is crucial to our understanding of the biological world, particularly in the case of integral membrane proteins such as PSI that mediate a variety of processes within the cell. Studying the effects of solution dynamics of redox protein complexes like PSI-hydrogenase fusions could yield new insights into electron transport in redox proteins, and the structure/function relationship of multi-protein complexes in general.

Cell-free photoproduction of H_2 has been demonstrated in a variety of PSI-based systems. For the purposes of this review, we will focus on (1) catalyst-based methods, and (2) PSI-hydrogenase fusion systems.

2.3.1 Catalyst-based methods for photocatalytic H_2 production

It has been demonstrated that chemically platinized thylakoid membranes can photocatalytically split water into H_2 and O_2 [12, 23]. This concept can also be

applied to isolated PSI, in which hexachloroplatinate salt ($[PtCl_6]^{2-}$) is photo-precipitated directly onto its reducing end (Figure 13) [13, 14, 24], with a maximum reported rate of 5.5 µmol H_2/h/mg Chl [13]. It has also been shown that metals such as osmium and ruthenium can be photo-precipitated onto photosynthetic membranes, thus enabling light-driven hydrogen production [158].

Photocatalytic hydrogen production has also been demonstrated using PSI/molecular wire/metal nanoparticle bioconjugates [159-161]. Having previously established that a C13G/C33S mutant of PsaC from the cyanobacterium *Synechococcus* sp. PCC 7002 that lacks a native cysteine ligand at a solvent-exposed position on the F_B cluster can be chemically rescued by a thiolated organic molecule [162], Golbeck and co-workers used this method to covalently link gold (Au) and platinum (Pt) nanoparticles to PSI via a 1,6-hexanedithiol molecular wire. They reported H_2 production rates of 3.4 µmol H_2/h/mg Chl for the PSI/1,6-hexanedithiol/Au bioconjugates, and 9.6 µmol H_2/h/mg Chl for the PSI/1,6-hexanedithiol/Pt bioconjugates. Furthermore, they found that addition of 5 µM cytochrome c_6 (cyt c_6) increased the rate of H_2 production for the Pt nanoparticles to 49.3 µmol H_2/h/mg Chl, indicating that H_2 evolution was rate-limited by donor-side reduction of $P700^+$ [159]. In subsequent

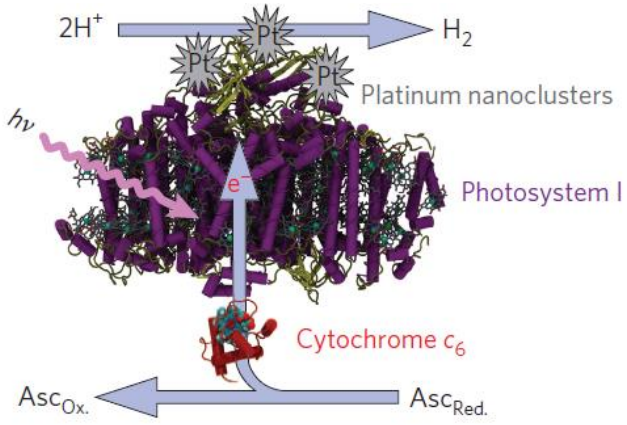

Figure 13. Schematic of electron flow in the PSI-Pt system (Figure adapted from Iwuchukwu et al. 2010)

work, they sought to maximize H_2 production by cross-linking plastocyanin (PC) with mutant PSI, both from spinach, and testing a variety of dithiol linker lengths, reporting a onetime H_2 evolution rate of 312 µmol H_2/h/mg Chl using a 1,4-benzenedithiol molecular wire [160]. Golbeck and co-workers have also incorporated a napthoquinone-molecular wire-Pt nanoparticle into the A_{1A} and A_{1B} sites of PSI from *Synechocystis* sp. PCC 6803 (*Synechocystis*), tapping into the electron transport pathway ahead of the iron-sulfur clusters, and reported a H_2 production rate of 67.3 µmol H_2/h/mg Chl [161].

2.3.2 PSI-hydrogenase fusions for photocatalytic H_2 production

PSI-hydrogenase complexes have also been studied as a means for photoproduction of H_2 [16, 37, 38, 163-165]. Ihara et al. genetically fused the stromal subunit PsaE of PSI from *T. elongatus* to the MBH of *R. eutropha*, finding that this PsaE-hydrogenase fusion protein spontaneously self-assembled with PsaE-free PSI in aqueous solution and that the resulting complex was capable of light-driven H_2 production at a rate of 0.58 µmol H_2/h/mg Chl [37]. However, H_2 evolution activity in this system was completely suppressed by the addition of Fd and FNR. In order to establish a H_2 photoproduction system that was not interrupted by Fd and FNR, PsaE from *Synechocystis* was cross-linked with cytochrome c_3 from *Desulfovibrio vulgaris* and subsequently self-assembled with PsaE-free PSI from *Synechocystis*. Consequently, in the presence of hydrogenase, Fd, and FNR, this system exhibited a H_2 evolution rate of 0.30 µmol H_2/h/mg Chl, suggesting the possibility of its use for H_2 production *in vivo* [163]. Krassen et al. subsequently extended this technique to PSI attached to a gold electrode. By genetically engineering a decahistidine tag on the PsaF subunit of PSI, they were able to orient PsaE-free PSI from *Synechocystis* lumenal side down on the gold surface via histidine/Ni-NTA interactions, which subsequently self-assembled with the PsaE-MBH fusion protein. Electrons were supplied to PSI from the gold electrode via a soluble electron carrier, and subsequently transferred to the hydrogenase upon light excitation of P_{700}. This PSI-MBH complex exhibited a H_2 photoproduction rate of 4500 µmol H_2/h/mg Chl

[164]. Drawbacks to this self-assembly method include the generation of a PsaE-free mutant of PSI in *T. elongatus*, which poses significant difficulty in terms of implementing the appropriate genetic modification without affecting the synthesis, maturation, and stability of the overall PSI complex. Furthermore, the PsaE gene occurs at multiple sites on the genome, making it quite difficult to knock out. This approach also requires the synthesis of a cyanobacterial protein in *R. eutropha*, which presents considerable challenges, such as the ability of the existing cell machinery to accurately synthesize the foreign protein, codon optimization, etc. These difficulties likely contribute to the poor H_2 evolution rates reported in the absence of a constant electron source such as a gold electrode.

Golbeck and co-workers developed a PSI-hydrogenase complex (Figure 14) wherein the terminal [4Fe-4S] cluster F_B of PSI from *Synechococcus* sp. PCC 7002 was directly tethered to the distal [4Fe-4S] cluster of the [FeFe]-

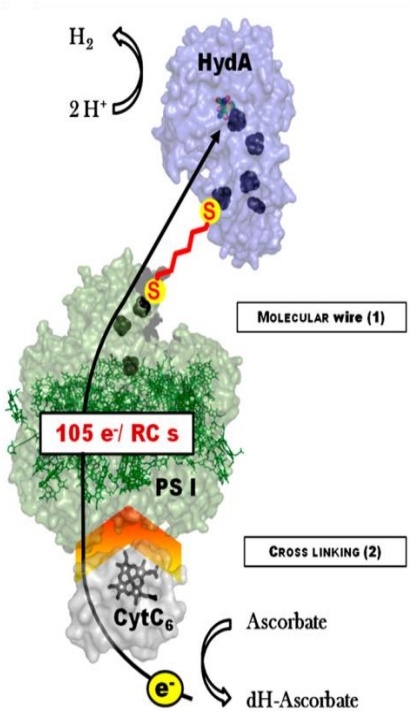

Figure 14. H_2 producing bionanoconstruct consisting of cyt c_6 cross-linked PSI connected via a dithiol molecular wire to an FeFe hydrogenase (Figure adapted from Lubner et al. 2010)

hydrogenase from *Clostridium acetobutylicum* (Hyd) via a thiol “molecular wire” as demonstrated previously with Pt nanoparticles [165]. PSI was reconstituted from P_{700}/F_X cores and recombinant PsaC and PsaD overproduced in *E. coli*, cross-linked to cyt c_6 also overproduced in *E. coli*, and tethered to the [FeFe]-hydrogenase using a 1,6-hexanedithiol ligand. The optimized system evolved H_2 at an average rate of 2,200 µmol H_2/h/mg Chl, exceeding the electron throughput rate of natural photosynthesis [16]. Difficulties associated with this approach stem from the fact that the [4Fe-4S] clusters from both PsaC and Hyd are genetically modified in the same fashion by replacing a coordinating cysteine residue with a glycine, in order to accomplish the ligand exchange. As a result, this method is non-specific and therefore results in the undesirable formulation of PSI-PSI and H_2ase-H_2ase homodimers. In addition, the [FeFe]-hydrogenase from *C. acetobutylicum* is extremely oxygen sensitive, hence this system is only functional in an anoxic environment.

In summary, cell-free biological hydrogen production is an active area of research. Significant research advances have recently been made in improving the H_2 evolution rates of PSI-hydrogenase complexes for photocatalytic H_2 production. However, a lack of understanding of the effects of solubilizing detergents on the stability and activity of integral membrane proteins like PSI, as well as the effects of solution dynamics on electron transport in these fusion complexes, presents a barrier to further advancement of this technology. This review has presented a variety of experimental and computational approaches for studying membrane proteins and mutli-protein complexes, highlighting the successes of small-angle scattering techniques in providing insight into the average structural envelope of protein-detergent and multi-protein complexes in solution, as well as the complementary use of MD simulation to provide atomic level detail on protein-environment and protein-protein interactions in these systems. The work of Koutsioubas et al. and Gsponer et al. in particular demonstrate the effectiveness of a combination of experimental and

computational techniques to comprehensively study protein-detergent complexes [152] and multi-protein systems [156].

In this work, we endeavor to address the lack of knowledge concerning detergent-solubilized membrane proteins and multi-protein complexes. First, we characterize the structure and dynamics of PSI under detergent-solubilized conditions using a combination of SANS and MD in an effort to provide new insights into improving the stability and activity of this integral membrane protein *in vitro*. Second, we present an experimental method for the site-specific ligation of PSI and MBH as an approach for the optimization of H_2 photoproduction via this fusion protein. We also demonstrate the use of a combination of SANS and MD to characterize the solution structure of this multi-protein complex, and attempt to gain atomic level insight into the protein-environment and protein-protein interactions that govern electron transport in this and other previously reported PSI-hydrogenase systems.

CHAPTER III
ANALYSIS OF THE SOLUTION STRUCTURE OF PHOTOSYSTEM I IN DETERGENT USING SMALL-ANGLE NEUTRON SCATTERING AND MOLECULAR DYNAMICS SIMULATION

A version of this chapter was originally published by Le, Rosemary K., Harris, Bradley J., et al.:

Le, R. K.; Harris, B. J.; Iwuchukwu, I. J.; Bruce, B. D.; Cheng, X.; Qian, S.; Heller, W. T.; O'Neill, H.; Frymier, P. D., Analysis of the Solution Structure of Thermosynechococcus elongatus Photosystem I in n-Dodecyl-beta-D-Maltoside using Small-Angle Neutron Scattering and Molecular Dynamics Simulation. *Archives of Biochemistry and Biophysics* 2014. 550-551: p. 50-57 (DOI: http://dx.doi.org/10.1016/j.abb.2014.04.005).

Bradley Harris was responsible for analysis of the SANS data collected for PSI in detergent solution, including pair distance distribution function (P(r)) analysis and low-resolution shape reconstruction. Additionally, he conducted the MD simulations of PSI embedded in a belt of detergent and generated theoretical scattering curves based on this model for comparison to the experimental SANS data. This chapter has been revised from the original manuscript to focus primarily on the SANS data analysis and MD modeling. Co-researchers Rosemary Le and Ify Iwuchukwu were responsible for preparation of the PSI samples for SANS analysis; Rosemary Le was also the primary writer of the original manuscript and did the Guinier fitting of the SANS data. Dr. Shuo Qian supervised SANS data collection and performed the subsequent data reductions. Dr. Barry Bruce, Dr. Xiaolin Cheng, and Dr. William Heller were collaborators on this work and offered their expertise in the areas of biochemistry and molecular biology, computational modeling, and small-angle scattering, respectively. Hugh O'Neill and Paul Frymier were the co-principal investigators (PIs) of this work.

3.1 Introduction

As discussed earlier, there has been considerable interest in investigating *in vitro* applications of the photosynthetic protein PSI for alternative energy solutions towards hydrogen [12-16, 37] or direct electricity production [17-21]. PSI is a membrane-bound protein complex involved in the photosynthetic cycle of plants, algae, and cyanobacteria that uses solar energy to mediate electron transfer from plastocyanin to ferredoxin, thereby furnishing reducing power for use in carbon fixation [10]. In the cyanobacterium *T. elongatus*, PSI exists as a trimeric complex containing the quinones and iron-sulfur clusters necessary for stable charge separation as well as the chlorophylls and carotenoids that make up its light-harvesting antennae [9].

Isolation of PSI from the thylakoid membrane requires the use of non-ionic detergents which disrupt the membrane and solubilize PSI in mixed micelles that prevent aggregation by mimicking the native environment. Although detergents permit the solubilization of membrane proteins in aqueous solution, they can often interfere with analysis of the structural and functional properties of these proteins [5, 6], and little is known of the aqueous structure of solubilized membrane proteins. It is important to understand the effect of protein-detergent interactions on PSI stability and activity in order to advance photosynthesis-based alternative energy solutions, as this knowledge could prove important for retaining PSI activity and addressing issues related to solution-phase aggregation of PSI at a surface.

Small-angle neutron scattering (SANS) is a useful technique for studying the solution structures of membrane proteins complexes [104, 106-108]. Unlike X-rays, neutron scattering lengths of heavier elements can be of similar magnitude to that of lighter elements, and there can be large isotope effects such as that of hydrogen vs. deuterium [97]. This unique property allows for solvent or contrast matching, wherein the H_2O/D_2O ratio of the solvent is tuned to match the scattering density of one or more components of a complex, thus masking its

scattering contribution and allowing study of the remaining components without disturbing the overall complex. The naturally occurring difference in scattering length densities of proteins and detergent molecules make SANS with contrast matching an excellent tool for studying protein-detergent complexes [106].

Molecular dynamics (MD) simulation is a useful computational technique for the study of biomacromolecules [109-111], and has been applied to a variety of detergent micelles [140-147] and protein-detergent complexes [148-152]. Additionally, simulations of protein complexes can be used to generate theoretical scattering curves for comparison to experimental SANS data [113], and low-resolution structural envelopes can be predicted based on SANS data to compare against simulated structures [100]. This allows for a greater understanding of protein-detergent interactions and the structure of the overall complex.

In this chapter, we have used a combination of contrast variation SANS and MD simulation to study the structure of PSI from *T. elongatus* solubilized in DDM detergent solution. SANS data from PSI/DDM complexes with and without detergent contrast matching provides detailed structural information on the protein-detergent complex. The SANS results were used as a guide to constructing an MD model of trimeric PSI in DDM detergent. A theoretical scattering curve based on a 50 ns MD simulation of trimeric PSI in DDM provides a reasonable fit to the experimental SANS data. Structural insights gained here not only improve our understanding of membrane protein-detergent interactions, but also guide strategies for incorporating PSI onto surfaces and in complex with other proteins for energy conversion applications.

3.2 Materials and methods

3.2.1 Isolation and purification of trimeric PSI from T. elongatus

Trimeric PSI used in this study was purified from the cyanobacterial strain *T. elongatus BP-1* [166] using a previously established protocol [13, 14]. Transmission electron microscopy analysis of complexes purified in this manner

confirmed the uniformity and purity of the trimeric PSI nanoparticles [13]. Purified PSI trimers were stored in 20 mM MES buffer with 0.03% DDM (GLYCON Biochemicals, D97002-C). The chl *a* concentration of PSI used for the SANS experiments was 0.59 mg/ml and 0.64 mg/ml for the 18% D_2O and 100% D_2O samples, respectively. The concentration of chlorophyll *a* (chl *a*) present in the samples was determined by Eq. (1):

$$\frac{\mu g\ chl\ a}{mL} = Abs_{665} \times 13.9 \quad (1)$$

as done previously [13]. Trimeric PSI concentration was then approximated by using the following stoichiometric conversion factors: 893 g chl *a* per mol chl *a*, 96 mol chl *a* per mol P_{700}, and 3 mol P_{700} per mol trimer.

3.2.2 Small-angle neutron scattering

SANS experiments were performed on the Bio-SANS beamline [167] located at the High Flux Isotope Reactor (HFIR) facility of the Oak Ridge National laboratory. Samples were measured at 20°C in 1 mm path length quartz cells (Hellma U.S.A., Plainview, NY). Scattered neutrons were collected with a 1 m by 1 m two-dimensional position-sensitive detector with 192 by 192 pixel resolution (ORDELA, Inc., Oak Ridge, TN). The 2D data were corrected for detector pixel sensitivity, as well as the dark current from ambient background radiation and the detector's electronic noise. The 2D reduced data were azimuthally averaged to yield the 1D scattering intensity I(*q*) vs. *q*, where *q* is the scattering vector described by Eq. (2):

$$q = \frac{4\pi \sin(\theta)}{\lambda} \quad (2)$$

and 2θ is the scattering angle from the incident beam and λ is the neutron wavelength (6 Å; Δλ/λ ≈ 0.15%). SANS data were taken at sample-to-detector distances of 1.1, 6.8, and 15.3 m for each sample to cover the expected *q* range. The 1D profiles from these different detector distances were merged to produce a complete scattering intensity plot. Trimeric PSI samples in 0.12% (w/v) DDM were measured at both 18% D_2O, the contrast match point for the detergent

[168], and 100% D_2O. A 0.12% DDM solution was also measured at 100% D_2O. All buffers used were also measured for background correction. The radius of gyration (R_g) was determined from the Guinier region of the SANS data using the Guinier approximation [169], Eq. (3):

$$I(q) = I(0)e^{\frac{q^2 R_g^2}{3}} \quad (3)$$

as described previously [106], where $I(0)$ is the forward scattering intensity, a shape-independent function of the total scattering power of the sample. A linear fit of $\ln[I(q)]$ vs. q^2, known as a Guinier plot, provides $I(0)$ and R_g from the y-intercept and slope, respectively [170]. R_g was also calculated from P(r), as mathematically it is the second moment of P(r) [171]. The data were processed using the small-angle scattering (SAS) data analysis package PRIMUS [172].

3.2.3 Shape restoration and MD simulation

Pair distance distribution function (P(r)) analysis of the SANS data was carried out using the indirect Fourier transform method for a monodisperse arbitrary particle implemented in GNOM [171]. The low-resolution shape of the scattering particles was reconstructed from the SANS data in the range of $q \leq 0.10$ Å^{-1} using the *ab initio* modeling tool DAMMIF [100]. The results of 50 DAMMIF runs using P3 symmetry for the trimeric complex were used to determine the average structure using the program DAMAVER [173], and were aligned to the crystal structure of PSI (PDB ID: 1JB0) via the program SUPCOMB [174]. The resulting overlaid structures were rendered in a low-resolution surface representation using VMD [175].

PSI/DDM models were constructed using the crystal structure of trimeric PSI from *T. elongatus* (Figures 15A and 15B; PDB ID: 1JB0) as a starting point and equilibrated using MD simulation for comparison to the experimental SANS data. For the first model (designated the DDM ring model; Figure 15C), a toroidal ring of approximately 800 DDM molecules was built around the peripheral transmembrane region of the PSI trimer complex. An additional 200 DDM molecules were added into the voids within the PSI trimer in a bilayer-like

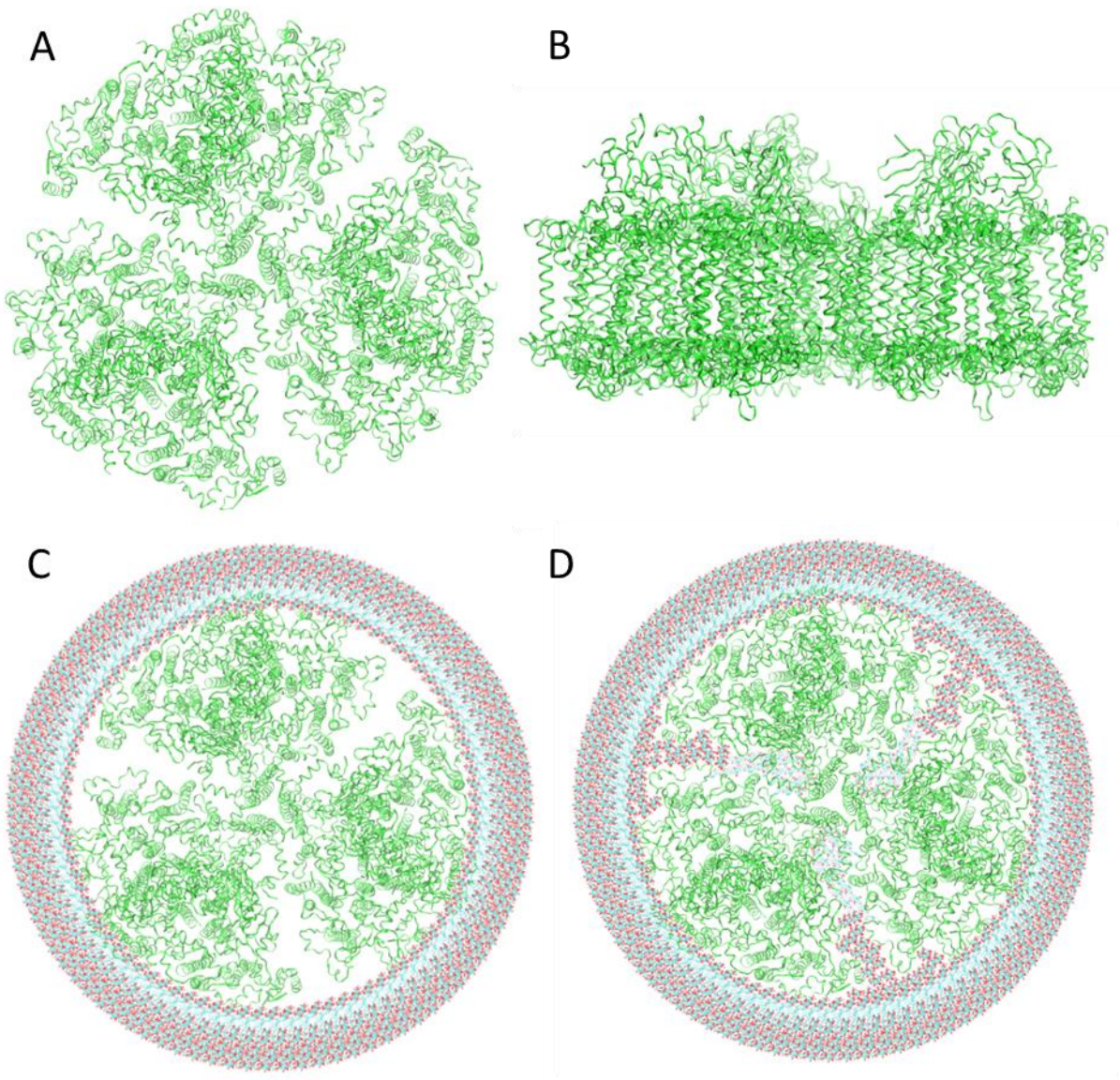

Figure 15. Crystal structure of trimeric PSI (PDB ID: 1JB0) from (A) top view and (B) side view stromal side up. Top view of initial configurations for (C) PSI/DDM ring model and (D) PSI/DDM void-filled ring model. All images are shown stromal side up, with protein in green ribbons. For the MD models, detergent molecules are shown in lines format with carbon atoms in cyan and oxygen atoms in red, and water molecules and counterions omitted for clarity.

configuration to create the second model (referred to as the void-filled DDM ring model; Figure 15D). MD simulations were performed using NAMD [109] with an extended version of the CHARMM27 force field for biomacromolecules [176-180]. Additional CHARMM parameters for the PSI cofactors chlorophyll *a* (chl *a*),

beta-carotene (BCR), and the iron-sulfur clusters (SF4) were obtained from parameterization of cofactors associated with PSII (chl *a* and BCR) [181] and hydrogenase enzymes (SF4) [182]. CHARMM parameters for DDM were taken from previous MD simulations of pure DDM micelles [144]. MD simulations of both models were equilibrated for 5 ns in a water box using an isothermal-isobaric ensemble (310 K, 1 atm). Theoretical scattering curves for both models were generated using ORNL_SAS [113], and were found to be nearly identical. Subsequently, a simulation of the void-filled DDM ring model was run for 50 ns and the MD trajectory was saved every 20,000 fs, resulting in 2,500 frames for final analysis. All simulations were performed on the Kraken supercomputer at the National Institute for Computational Sciences, co-located at the University of Tennessee – Knoxville and Oak Ridge National Laboratory campuses.

3.3 Results and discussion

3.3.1 SANS analysis of pure DDM micelles

SANS data of pure DDM micelles was collected for comparison to the PSI/DDM complex. The resulting scattering profile is shown in Figure 16A. The radius of gyration (R_g) of the DDM micelles in 100% D_2O was determined to be 2.04 ± 0.30 nm by the Guinier approximation (Table 1; Figure 17A). This result is consistent with the value of 2.01 nm determined by Abel et al. [180]. R_g is a summary measure of the distribution of distances of atoms from the center of the molecule, and is representative of the size of a scattering particle in solution. Therefore, considering that the fully extended length of a DDM molecule, as determined by Auer et al. [183], is ~2.3-2.8 nm, these R_g results suggest a micellar structure. R_g can be used to calculate the radius R of a spherical micelle using Eq. (4) [184]:

$$R_g^2 = \frac{3R^2}{5} \quad (4)$$

Based on an R_g of 2.04 nm (Table 1), a spherical DDM micelle would have a radius of 2.63 nm and an approximate surface area of 87.2 nm^2. Using a DDM monomer headgroup area of 0.63 nm^2 (assuming an ellipsoidal projected area

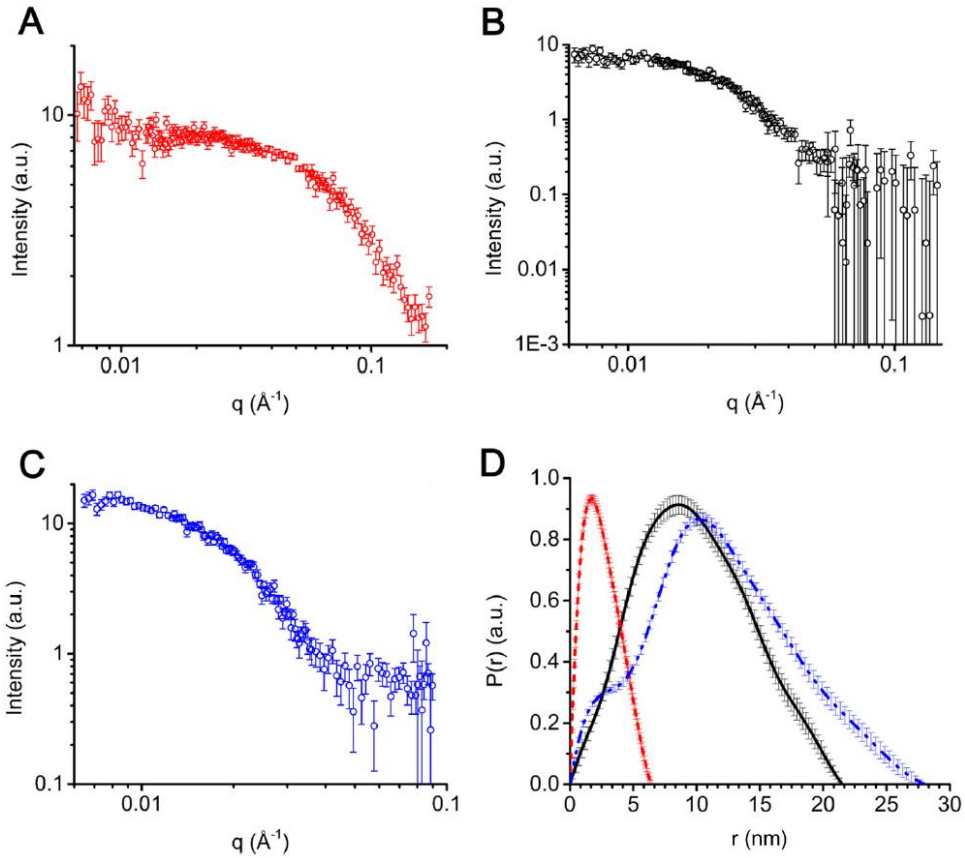

Figure 17. Scattering profiles for: (A) 0.12% DDM in 100% D_2O; (B) PSI in 0.12% DDM in 18% D_2O; (C) PSI in 0.12% DDM and 100% D_2O. (D) P(r) curves for 0.12% DDM in 100% D_2O (red curve), PSI in 0.12% DDM and 18% D_2O (black curve), and PSI in 0.12% DDM and 100% D_2O (blue curve). (Figure from Le et al. 2014)

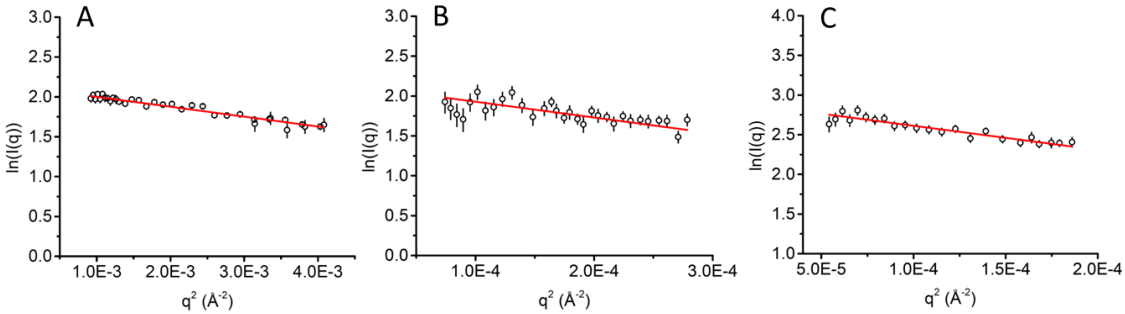

Figure 16. Guinier analysis (ln[I(*q*)] vs. q^2) for (A) 0.12% (w/v) DDM at 100% D_2O; (B) PSI in 0.12% DDM at 18% D_2O; (C) PSI in 0.12% DDM at 100% D_2O. In all cases, SANS data is shown in open circles with the red line representing a linear fit in the low-q regime (qR_g < 1.3). (Figure from Le et al. 2014)

Table 1. Structural parameters of pure DDM and PSI/DDM samples from SANS analysis

Sample	[D_2O] %	[PSI trimer] (mM)	[DDM] (mM)	Guinier R_g (nm)	P(r) R_g (nm)	D_{max} (nm)
DDM	100	0	2.35	2.04±0.30	2.03±0.02	6.4±0.5
PSI/DDM	18	2.27×10^{-3}	2.35	7.79±2.86	7.59±0.09	21.5±1.0
PSI/DDM	100	2.48×10^{-3}	2.35	9.49±2.32	9.31±0.11	28.0±1.0

with a = 1.05 nm and b = 0.76 nm as measured in PyMOL [185]) and assuming densely hexagonal-packed headgroups ($\eta = \frac{\pi}{2\sqrt{3}}$) yields an aggregation number of ~126 DDM molecules per micelle, which is consistent with previously reported values [186-189]. The micelle concentration can then be calculated from Eq. (5):

$$[\text{total detergent}] = [\text{monomer}] + [\text{micelles}] \times AN_{DDM} \quad (5)$$

where [total detergent] = 2.35 mM DDM (0.12% (w/v)), [monomer] = 0.17 mM DDM (critical micelle concentration (CMC) for DDM), and AN_{DDM} = 126 monomers per micelle. This yields a micelle concentration of 1.73×10^{-2} mM, and the distribution of DDM in this system is therefore 92.8% in micelles and 7.2% as non-associated monomer, effectively a monodisperse system. For this reason we used the monodisperse arbitrary particle model for calculating the P(r) distribution in GNOM [171]. The P(r) function represents the probable distribution of distances between the scattering centers within the scattering particle and goes to zero at the maximum particle dimension, D_{max}. In our case, the P(r) curve can provide information on the size and shape of DDM micelles (or PSI/DDM complexes) in solution. An R_g value of 2.03 ± 0.02 nm was obtained from the P(r) fitting (Table 1; Figure 16D, red curve). The agreement between the Guinier and P(r)-derived R_g values further indicates that the DDM micelles are well-behaved, monodisperse particles in solution. However, the assumption of a spherical micelle is not ideal, as the value of the maximum particle size (D_{max}) from the GNOM fit is 6.4 ± 0.5 nm, which is greater than twice the length of a DDM

molecule, and the P(r) curve is not symmetric, indicating that the micelles are not spherical. This is not unexpected, as MD studies of pure DDM micelles indicate an ellipsoidal micelle structure [144].

3.3.2 SANS analysis of PSI at DDM match point

SANS data of PSI/DDM at 18% D_2O, the contrast match point for DDM, was collected to study the structure of trimeric PSI under detergent-solubilizing conditions. The scattering profile and P(r) curve for PSI in 0.12% DDM and 18% D_2O are shown in Figures 16B and 16D (black curve), respectively. R_g values for PSI/DDM in 18% D_2O (Table 1) are 7.79 ± 2.86 nm from Guinier analysis (Figure 17B) and 7.59 ± 0.09 nm from P(r) fitting, which agree well with each other. The P(r) curve is asymmetric, which has been seen previously in membrane proteins [106-108]. It peaks at 8.6 nm and has a D_{max} of 21.5 ± 1.0 nm, suggesting a disk-shaped structure that has been observed in previous SAS studies of PSI solubilized in DDM [102], as well as other membrane protein - detergent complexes [101, 105, 106]. These dimensions are in good agreement with the crystal structure of trimeric PSI (Figures 15A,B) [9].

3.3.3 SANS and MD analysis of PSI/DDM complex

The structure of the PSI/DDM complex was probed by collecting neutron scattering data of PSI in 0.12% DDM at 100% D_2O (Figure 16C), the point at which there is maximum contrast between scattering particles and the solvent, with both protein and detergent contributing to scattering. An R_g value for the PSI/DDM complexes of 9.49 ± 2.32 nm was determined by Guinier analysis (Table 1; Figure 17C). To calculate the DDM aggregation number for the PSI/DDM system, we approximate the complex as a squat cylinder comprising two concentric cylinders, the inner being PSI and the outer DDM (Figure 18); we further assume that DDM exists with the hydrophilic headgroups oriented on the periphery of the cylinder. The radius of this cylinder can be determined based on Eq. (6) [184]:

$$R_g^2 = \frac{R^2}{2} + \frac{L^2}{12} \quad (5)$$

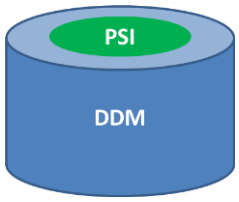

Figure 18. Cylindrical approximation of PSI/DDM complex

where *R* and *L* represent the cylinder radius and height, respectively. For the PSI/DDM complex, *L* can be approximated as the height of the membrane-spanning domain of PSI (6.7 nm, as measured using the crystal structure in PyMOL). Therefore, for an R_g of 9.49 nm (Table 1), the PSI/DDM cylinder would have a radius of 13.1 nm and a peripheral surface area (excluding the top and bottom of the cylinder) of 553 nm^2. Again using the projected surface area of 0.63 nm^2 for a DDM headgroup and assuming dense hexagonal packing, we calculate an aggregation number of ~796 DDM monomers per PSI trimer. The concentration of free DDM micelles can then be calculated using Eq. (6):

$$[\text{total detergent}] = [\text{monomer}] + [\text{PSI}] \times AN_{DDM\text{-}PSI} + [\text{micelles}] \times AN_{DDM} \quad (6)$$

where [total detergent] = 2.35 mM DDM, [monomer] = [CMC] = 0.17 mM DDM, [PSI] = 2.48 x 10^{-3} mM PSI trimer, $AN_{DDM\text{-}PSI}$ = 796 DDM monomers per PSI trimer, and AN_{DDM} = 126 DDM monomers per micelle. This yields a free micelle concentration of 1.63 x 10^{-2} mM. The resulting distribution of DDM monomer is thus 84% associated with PSI trimer, 8.8% in micelles, and 7.2% as non-associated monomer. We therefore assume that the scattering due to DDM micelles is effectively negligible compared to that of the PSI/DDM complexes and use the monodisperse arbitrary particle model for P(r) analysis. P(r) fitting yielded an R_g of 9.31 ± 0.11 nm (Table 1); the agreement of the Guinier and P(r)-derived values as before indicates a monodisperse system. The P(r) curve (Figure 16D, blue curve) has an asymmetric parabolic shape, peaking at 10.1 nm and tapering off to a D_{max} of 28.0 ± 1.0 nm, which is again suggestive of a disk-like structure [101, 102, 105, 106]. The P(r) curve also has a shoulder feature at vector lengths below 5 nm; this feature has not been previously seen in SAS studies of

membrane proteins and cannot yet be explained. Comparing to PSI/DDM in 18% D2O, D_{max} increases from 21.5 ± 1.0 nm to 28.0 ± 1.0 nm. This difference of 6.5 nm is slightly larger than twice the length of an extended DDM molecule (~5 nm), which supports the existence of detergent around the peripheral transmembrane α-helices of the trimer also reported in previous studies of detergent-solubilized membrane proteins [190-194].

To visualize the structural envelope of the PSI/DDM complex, low-resolution shape restoration was performed based on the SANS data from PSI in 0.12% DDM at both 18% and 100% D_2O for $q \leq 0.1$ $Å^{-1}$ using DAMMIF [100]. The superposition of the resulting average structure with the PSI crystal structure can be seen in Figures 19A-C. Due to the low-resolution surface rendering in VMD, the dummy-atom reconstruction of PSI/DDM at 18% D_2O appears slightly larger than the crystal structure (blue ribbons); however, the calculated D_{max} of 21.5 nm

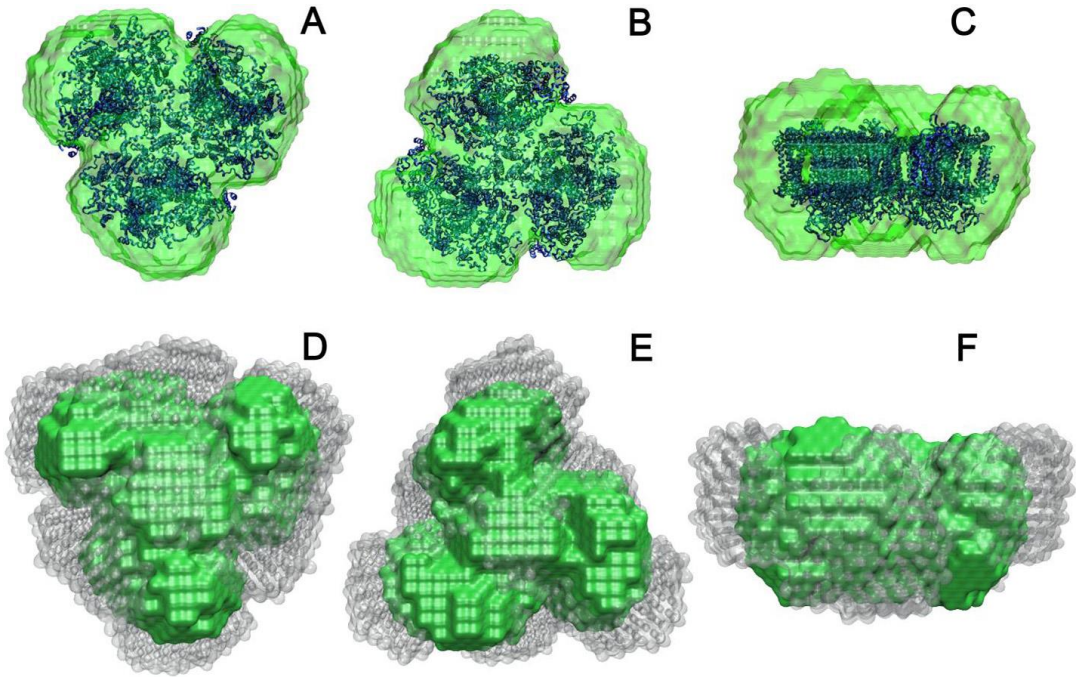

Figure 19. Trimeric PSI crystal structure (blue ribbons) superimposed with dummy-atom reconstruction of PSI/DDM at 18% D_2O (green) shown from (A) stromal side up, (B) lumenal side up, and (C) side view lumenal side up. Reconstruction of PSI/DDM at 18% D_2O (green) superimposed with reconstruction at 100% D_2O (gray) shown from (D) stromal side up, (E) lumenal side up, and (F) side view lumenal side up. (Figure from Le et al. 2014)

(Table 1) is consistent with the expected diameter of PSI from the crystal structure. The reconstructed shape from the SANS data of PSI/DDM at 100% D_2O (gray), overlaid with the reconstruction at 18% D_2O (green) in Figures 19D-F, suggests that DDM interacts with PSI in a non-uniform envelope around the exposed periphery of the trimer. In particular, the detergent appears localized to the peripheral transmembrane α-helices of the individual PSI monomer lobes and the interstitial spaces between these monomers (Figures 19D,E), with very little DDM on the stromal and lumenal faces as can be seen clearly in Figure 19F. This is expected, as analysis in VMD shows no significant hydrophobic patches present on these faces (Figures 20A,B), while the exposed surface of the transmembrane region is a uniform distribution of hydrophobic α-helices (Figure 20C).

Atomic-resolution PSI/DDM models were constructed and equilibrated using MD simulation for comparison to the experimental SANS data. Starting structures for the DDM ring model and void-filled DDM ring model are shown in Figures 21A and 21D, respectively. In both models, the detergent molecules surrounding the periphery of PSI quickly evolve from a uniform belt to conform to the individual monomer lobes and fill the interstitial voids after 5 ns of MD simulation (Figures 21B,C,E,F). Structures of the PSI/DDM complex after 5 ns of simulation for both the ring and void-filled ring models were used to generate theoretical scattering

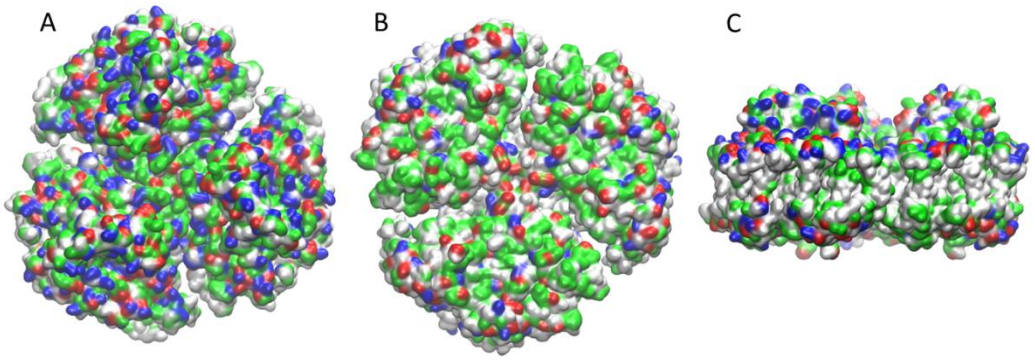

Figure 20. Crystal structure of trimeric PSI rendered according to residue type (non-polar = white, polar = green, acidic = red, basic = blue) shown from (A) stromal side up, (B) lumenal side up, and (C) side view stromal side up.

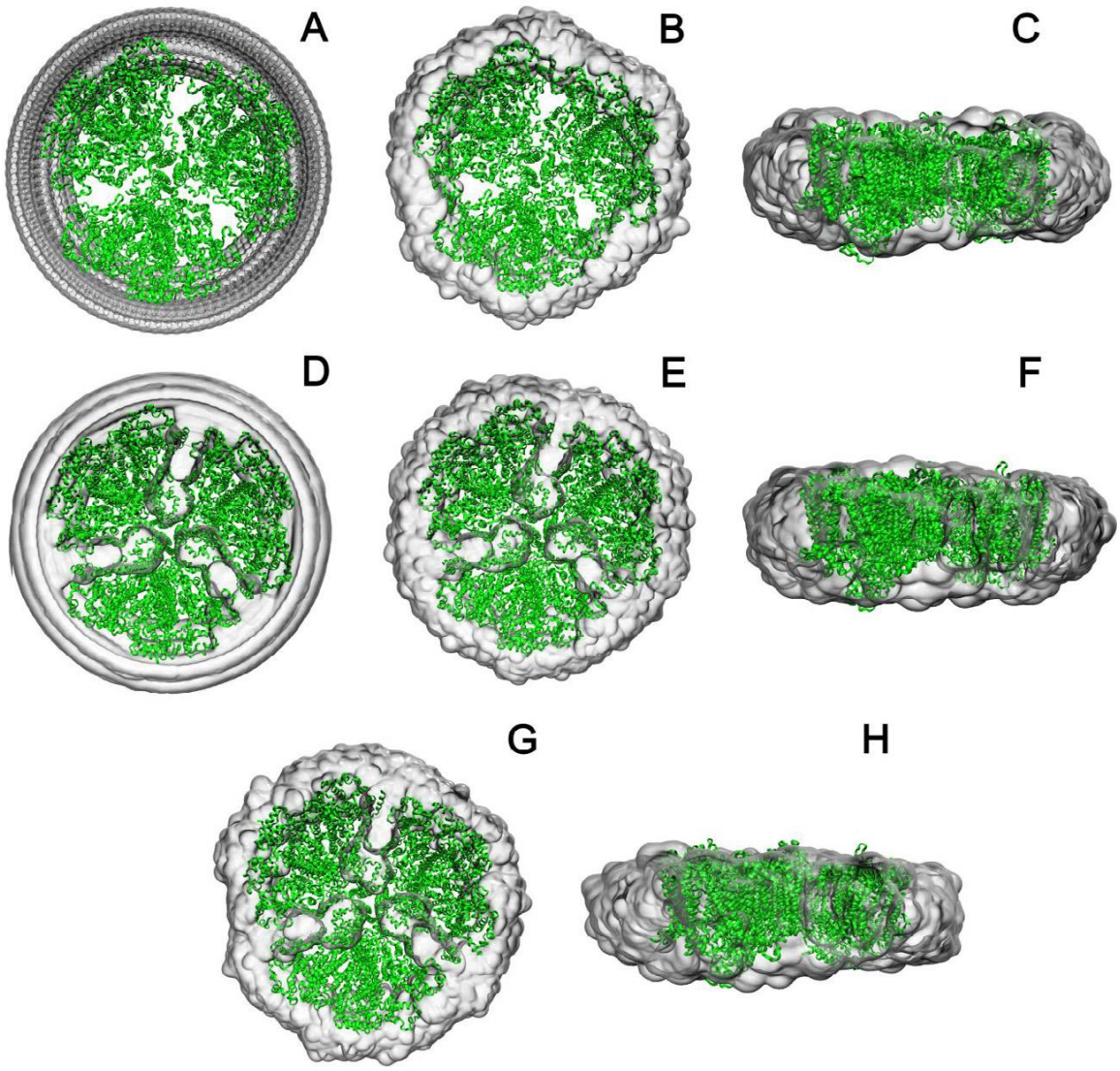

Figure 21. The DDM ring model shown (A) lumenal side up prior to MD simulation, (B) lumenal side up after 5 ns of MD simulation, and (C) side view lumenal side up after 5 ns of MD simulation. The DDM void-filled ring model shown (D) lumenal side up prior to MD simulation, (E) lumenal side up after 5 ns of MD simulation, (F) side view lumenal side up after 5 ns of MD simulation, (G) lumenal side up after 50 ns of MD simulation, and (H) side view lumenal side up after 50 ns of MD simulation. In all cases, the protein is shown in green ribbons and the detergent in gray in low-resolution surface representation. (Figure from Le et al. 2014)

curves using ORNL_SAS for comparison to the SANS data; the curves for these two models were nearly identical when superimposed (Table 2). Therefore, a subsequent 50 ns MD simulation was conducted for only the void-filled ring model to allow further equilibration of the PSI/DDM complex. The C_α root-mean-square deviation (RMSD) of PSI was calculated relative to the simulation starting structure throughout the 50 ns MD trajectory to analyze the stability of this complex; this value represents an average over all three monomers of the PSI trimer. The RMSD quickly rises to ~2.2 Å after 10 ns, drifting to 2.4 Å after 30 ns and remaining there for the duration of the simulation (Figure 22). Additionally, there was little discernible change in the theoretical scattering curves generated after 5 and 50 ns of simulation time (Table 2). Therefore, the 50 ns MD simulated structure of the DDM void-filled ring model was deemed sufficient for comparison

Table 2. Theoretical scattering data for PSI/DDM MD models

Ring		Void-Filled Ring 5 ns		Void-Filled Ring 50 ns	
q ($Å^{-1}$)	I(q) (a.u.)	q ($Å^{-1}$)	I(q) (a.u.)	q ($Å^{-1}$)	I(q) (a.u.)
0.0050	14.53	0.0050	14.63	0.0050	14.65
0.0101	12.17	0.0101	12.23	0.0101	12.24
0.0150	9.17	0.0152	9.05	0.0150	9.18
0.0202	6.09	0.0202	6.08	0.0202	6.07
0.0251	3.82	0.0251	3.80	0.0251	3.80
0.0304	2.35	0.0304	2.33	0.0304	2.33
0.0409	1.31	0.0409	1.29	0.0409	1.29
0.0510	0.97	0.0510	0.95	0.0510	0.95
0.0615	0.73	0.0615	0.72	0.0615	0.72
0.0720	0.65	0.0720	0.65	0.0720	0.65
0.0803	0.62	0.0803	0.63	0.0803	0.63
0.0915	0.61	0.0915	0.61	0.0915	0.62
0.1009	0.61	0.1009	0.62	0.1009	0.61
0.1113	0.59	0.1113	0.60	0.1113	0.60
0.1227	0.59	0.1227	0.60	0.1227	0.60
0.1398	0.59	0.1310	0.60	0.1310	0.60
0.1444	0.58	0.1444	0.60	0.1444	0.59
0.1492	0.58	0.1492	0.59	0.1492	0.59

to the SANS data.

The experimental scattering data for PSI/DDM at 100% D_2O is plotted in Figure 23, overlaid with theoretical scattering curves based on the PSI crystal structure (green curve) and the equilibrated PSI/DDM void-filled ring model (red curve). The MD model results in a reasonable fit to the SANS data, with deviations occurring in the high *q* regime. The theoretical curve based on the crystal structure yields a $\chi^2 = 8.85$ compared to the experimental data, compared to a $\chi^2 = 3.83$ for the equilibrated PSI/DDM MD model curve. Taken together with the DAMMIF dummy-atom reconstructions (Figures 19D-F), these results suggest that the idealized geometry of trimeric PSI embedded in a uniform detergent belt surrounding the peripheral transmembrane domain is an oversimplification of this system. It should be noted that the initial uniform detergent ring continued to conform to the PSI trimer throughout the simulation time, as can be seen in comparing the structure of the PSI/DDM complex after 5 ns of MD simulation (Figures 21E,F) to that after 50 ns (Figures 21G,H). Future work will be focused on achieving longer simulation times to determine whether or not this redistribution of DDM progresses towards the trilobal structure seen in the DAMMIF reconstructions.

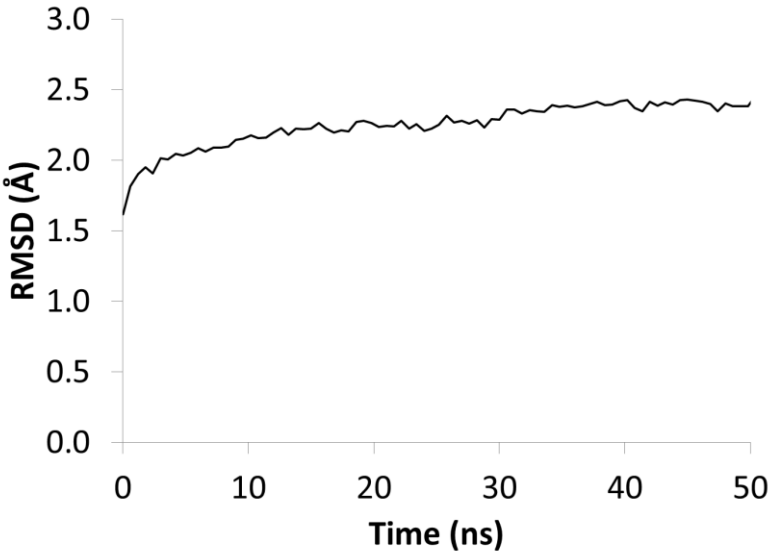

Figure 22. C_α RMSD values versus time for the PSI/DDM void-filled ring model simulation.

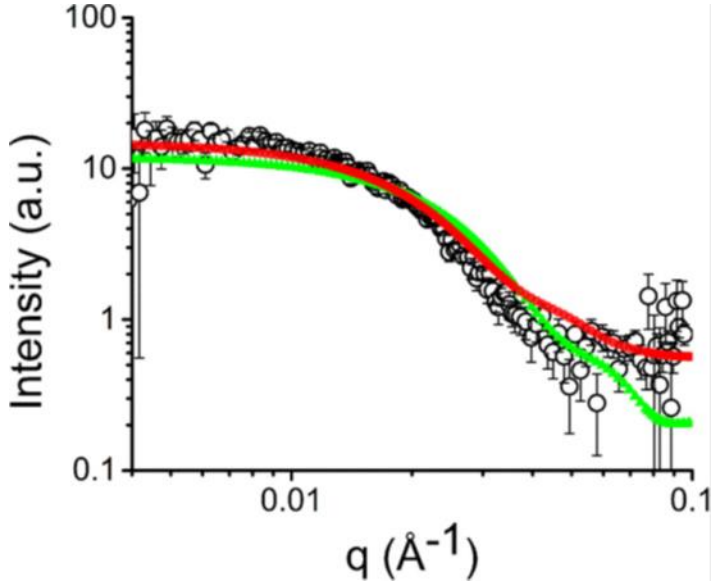

Figure 23. Comparison of scattering data for PSI in 0.12% DDM (w/v) at 100% D_2O (open circles) with theoretical scattering curves based on the DDM void-filled ring model (red continuous line) and the crystal structure of trimeric PSI (green continuous line). (Figure from Le et al. 2014)

3.4 Summary

The solution structure of trimeric PSI under detergent-solubilized conditions was investigated using contrast variation SANS and produced results that have not been seen previously. Analysis of the SANS data and subsequent shape restoration both with and without the scattering contributions of DDM suggests that the detergent exists as a belt surrounding the transmembrane periphery of the PSI complex. MD simulations of trimeric PSI embedded in a uniform belt of DDM were developed based on the SANS shape restoration results. Theoretical scattering profiles generated based on these models demonstrate that the inclusion of a DDM belt produces a reasonable fit to the SANS data, but fails to reproduce key structural features in the high *q* regime. These results suggest that the conventional notion of membrane proteins as being suspended in solution within a uniform detergent structure is an idealization and an oversimplification. This new information changes our understanding of protein-detergent complexes, and in the particular case of PSI from the thermophilic cyanobacterium *T. elongatus* can be utilized in designing detergent solutions that minimize instability and aggregation, thus leading to enhanced performance and longer useful lifetimes for biorenewable energy applications. The MD simulations conducted here will need to be extended to allow for in-depth analysis of the stability and mobility of this PSI/DDM complex and the effects of protein-detergent interactions on the structure and function of PSI.

CHAPTER IV
IN-DEPTH ANALYSIS OF PHOTOSYSTEM I – DETERGENT COMPLEX STABILITY AND INTERACTIONS USING ALL-ATOM MOLECULAR DYNAMICS SIMULATION

A version of this chapter was originally published by Harris, Bradley J. et al.:

Harris, B.J., Cheng, X., Frymier, P.D. All-Atom Molecular Dynamics Simulation of a Photosystem I/Detergent Complex. *Journal of Physical Chemistry B* 2014. 118 (40): p. 11633-11645 (DOI: http://dx.doi.org/10.1021/jp507157e).

Bradley Harris was the primary writer of the original manuscript, and conducted the MD simulations and subsequent data analyses described therein. Dr. Xiaolin Cheng was a collaborator on this work and offered his expertise in computational modeling. Dr. Paul D. Frymier was the principle investigator of this work.

4.1 Introduction

The biological importance of membrane proteins is reflected in the fact that they comprise approximately 30% of open reading frames [1], yet they represent only a small fraction of known protein structures [2]. As discussed earlier, this knowledge barrier is caused by difficulties in membrane protein expression and crystallization, most often arising from the need for solubilization of these proteins from the transmembrane surface via a detergent or lipid assembly [3, 4]. Little is known of the aqueous structure of protein-detergent complexes and how the presence of detergents can impact the stability and activity of membrane proteins [5, 6]. PSI represents an interesting test case for the study of membrane protein complexes, as it is capable of generating reducing power via solar energy and as such has been the subject of much research regarding potential alternative energy applications, as reviewed before. Enhanced understanding of the effects of protein-detergent interactions on the functional properties of PSI could improve the performance and prolong the useable lifetimes of PSI-based energy conversion devices.

In the previous chapter, we discussed our study of the solution structure of PSI from *T. elongatus* using a combination of SANS and MD simulation [195]. We were able to determine that the dimensions of the PSI/DDM complex were consistent with trimeric PSI embedded in a disk-like detergent micelle, and shape restoration based on the SANS data indicated that the detergent conformed to the transmembrane periphery in a non-uniform, trilobal orientation rather than a uniform toroidal belt. MD simulations based on a starting structure of trimeric PSI surrounded by a monolayer DDM belt were used to generate theoretical scattering curves that produced a reasonable fit to the experimental SANS data. However, they failed to produce key structural features in the high *q* regime, suggesting further work might be necessary to verify that these models had reached an equilibrated conformation and better analyze protein-detergent interactions and their implications in this system.

In this chapter, we have conducted extensive simulations of the PSI/DDM complex for in-depth analysis of the structure and dynamics of the PSI trimer under detergent-solubilized conditions. With this work, we are able to examine the following to a level of detail not previously possible: stability and flexibility of PSI in a detergent environment; detailed protein-detergent interactions and their mutual effects on each other's behavior; and possible implications of the observed conformational dynamics for the function of this photosynthetic protein complex.

4.2 Materials and methods

4.2.1 Setup of pure DDM micelles

Pure DDM micelles were initially constructed using Packmol [196] in a spherical orientation with the hydrophobic tails pointed inward and the hydrophilic head groups pointed outward. A base case of 132 DDM molecules was chosen based on the reported aggregation number [144], with micelles of 70 and 200 DDM molecules also simulated as extreme cases. All systems were subjected to 20 ps of MD equilibration *in vacuo*, and subsequently solvated with TIP3P water

molecules. Solvated systems were then subjected to an additional 20 ps of MD equilibration. These equilibration steps were carried out to remove clashes between the atoms, heat the system gradually to the target temperature and reach a kinetic energy equi-partition. This is meant to ensure that the systems do not explore unrealistically high-energy conformational spaces prior to MD production runs. Equilibrated, solvated systems were then simulated for 100 ns, and the resulting trajectories used for subsequent data analyses.

4.2.2 Setup of PSI/DDM micelle systems

The starting structure for all PSI/DDM models was based on the 2.5 Å resolution crystal structure of trimeric PSI from *T. elongatus* (PDB ID: 1JB0) [9]. There are 91 out of a total of ~2,300 residues which were not resolved in the PSI crystal structure and were not included in the simulations. These unresolved regions span less than 15 successive residues except in the cases of psaF and psaK, which are missing 23 and 20 amino acids, respectively. These terminally located, intrinsically disordered regions are expected to play a role in PSI-mediator interactions, but are unlikely to significantly affect the global dynamics of the PSI/DDM complex during simulation. Phylloquinone molecules located in the reaction center core (two per PSI monomer) were not included in the simulations.

A monolayer belt of DDM detergent was constructed around the periphery of PSI consisting of semi-circular planes of DDM densely packed around the hydrophobic exterior transmembrane surface of the protein (Figure 15C). Additional DDM molecules were inserted into the interstitial voids between individual PSI monomers in a bilayer-like orientation, resulting in the so-called void-filled ring model which is the focus of this paper (Figure 15D). An alternative random model was generated by randomly placing DDM molecules around the PSI trimer using Packmol [196], and subsequently removing all those lipids which overlapped with the protein complex. For the associated lipids model, the structure and coordinates of included lipids were based on 1,2-distearoyl-monogalactosyldiglyceride (LMG) and 1,2-dipalmitoyl-phosphatidylglycerol (LHG)

lipids resolved in the crystal structure, and atom types and parameters were taken from the updated CHARMM force field for lipids [179]. All systems were subjected to 50 ps of MD equilibration *in vacuo*, and subsequently solvated with TIP3P water molecules. Sodium and chloride counterions were then added by random replacement of water molecules in order to neutralize each system. The solvated, neutralized systems were subjected to an additional 50 ps of MD equilibration, and then MD production runs were carried out.

4.2.3 Simulation details

All simulations were carried out using NAMD [109], with an extended version of the CHARMM27 force field [176-180]. Parameters for chlorophyll *a* and beta-carotene were taken from the work of Zhang et al. on the parameterization of PSII cofactors [181], and those of the iron-sulfur clusters based on the work of Smith and co-workers [182]. Parameters for DDM were taken from MD studies of pure DDM micelles conducted by Abel et al. [144]. All simulations were performed at constant temperature (300 K), pressure (1 atm), and number of particles. Electrostatics were calculated using particle mesh Ewald [197], with a 1 nm cut-off for the real space calculation; a 1 nm cut-off was also used for van der Waals interactions. System temperature was maintained by controlling the kinetic energy of the system using Langevin dynamics, with a damping coefficient of 10 ps^{-1}, and system pressure was controlled using the Langevin piston method [198], with an oscillation time constant of 200 fs and a damping time constant of 100 fs. The time step for integration was 2 fs for all simulations, and coordinates and velocities were saved every 20 ps. Data analyses used either VMD [175] or locally written code; all images were rendered in VMD.

4.2.4 Predicted docking of soluble electron mediators

The ZDOCK online server [199] was used to generate predicted docked structures for the soluble electron mediators cytochrome c_6 (PDB ID: 1C6S) and ferredoxin (PDB ID: 2CJN) to PSI. Model predictions were generated for every monomer of trimeric PSI using snapshots of the MD trajectory taken for every 20

ns of simulation, with predictions based on the PSI crystal structure also generated for comparison. In the case of cytochrome c_6, only the psaAB subunits of PSI were considered for predicted docking, with W655A and W631B specified as contact residues [200]. For ferredoxin, only the psaCDE subunits were considered, and I11C, T14C, Q15C, K34C, K104D, and R39E were specified as contact residues [201]. C_α RMSD values of the top ten predicted bound mediator structures (if available) were calculated for each time step relative to the crystal structure predictions for each of the three monomers of the PSI trimer, with the minimum C_α RMSD for each time step being reported.

4.3 Results and discussion

In an initial model of randomly placed DDM molecules around the PSI trimer, the system had not begun to converge after >20 ns of MD simulation (Figure 24). Therefore, in order to obtain an equilibrated PSI/DDM complex on an attainable time scale, we chose to use a pre-formed protein-detergent assembly as our

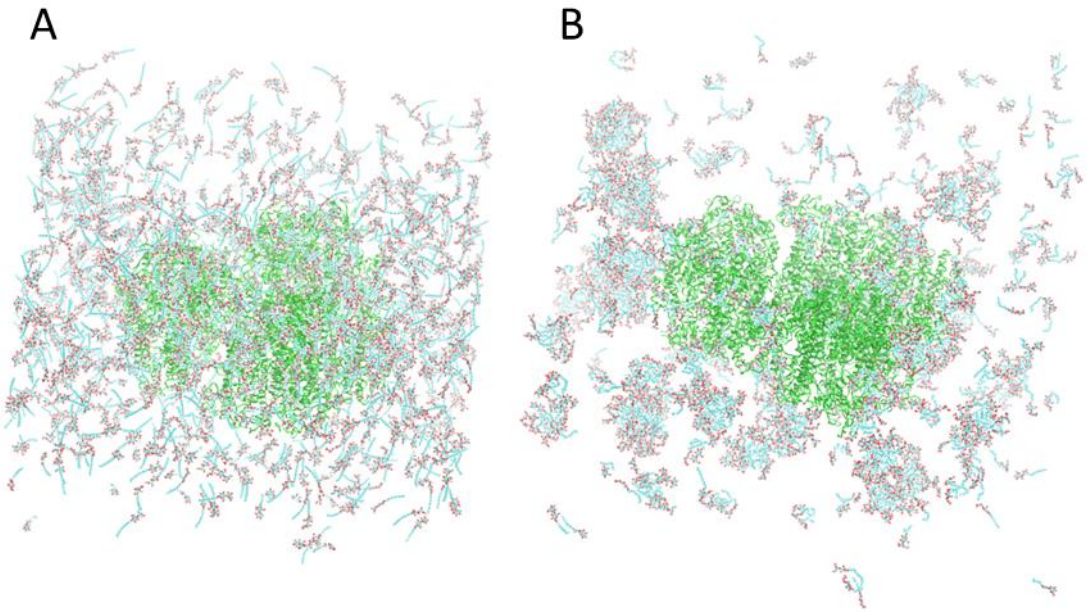

Figure 24. Model of trimeric PSI surrounded by randomly placed DDM molecules (A) prior to MD simulation, and (B) after 20 ns of MD simulation. Protein is shown in green ribbons, with detergent shown in lines format with carbon atoms in cyan and oxygen atoms in red, and water molecules and counterions omitted for clarity.

starting structure. Based on results from dynamic light scattering, the PSI/DDM complex is known to be disk-like, 30 nm in diameter by 9 nm height [202]. Therefore, using the characteristic hydrophobic belts of membrane proteins as a guide, we constructed a uniform, toroidal belt of DDM detergent around the peripheral transmembrane domain of the PSI trimer. In order to fill the interstitial hydrophobic voids between individual PSI monomers, we added additional DDM molecules into the interstitial spaces of the PSI trimer, in a bilayer orientation with the hydrophobic tails pointed inward. Based on our previous SANS study, the aggregation number has been estimated to be ~796 DDM molecules per PSI trimer, assuming uniform packing of detergent around the peripheral transmembrane region [195]. Because this aggregation number does not account for the interstitial voids between PSI monomers, we chose to use 1,000 DDM molecules in our simulations, with 800 detergent monomers in a toroidal belt around the transmembrane domain and 200 in the interstitial voids. The initial configuration is shown in Figures 25C and 25D. This system was energy-minimized, solvated, and equilibrated (please refer to Materials and methods), and subsequently subjected to a 200 ns MD production run. A theoretical scattering curve generated based on this PSI/DDM complex after 50 ns of MD simulation yielded a χ^2 = 3.83 when compared to experimental SANS data [195].

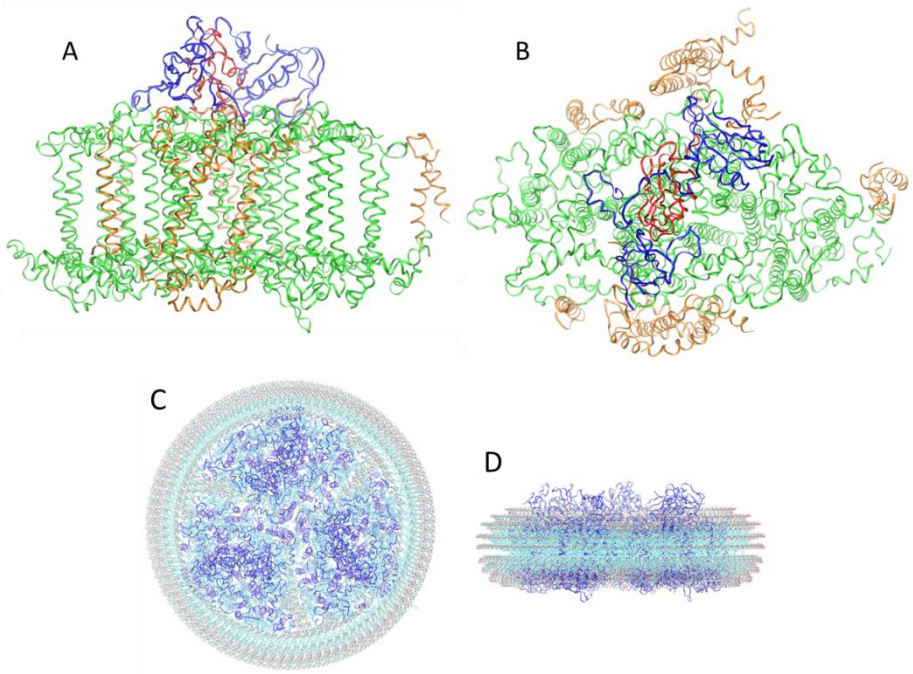

Figure 25. Crystal structure of monomeric PSI from *T. elongatus* shown from (A) side view with the stromal side up, and (B) lumenal side up. The protein is shown in ribbons with the reaction center subunits (psaAB) in green, the terminal electron acceptor psaC in red, the peripheral stromal subunits (psaDE) in blue, and the peripheral transmembrane helices (psaFIJKLMX) in orange. The starting structure for the PSI/DDM void-filled ring model shown from (A) lumenal side up and (B) side view along the transmembrane domain with the stromal side up. Protein is shown in blue ribbons, detergent molecules in lines format, and water molecules and counterions are omitted for clarity.

4.3.1 Protein dynamics

Protein stability

In order to determine the stability of PSI in the DDM detergent environment, we calculated the root-mean-squared deviation (RMSD) of the PSI trimer relative to the initial crystal structure throughout the course of the simulation (Figure 26). The C_α RMSD calculated for all residues climbed to ~2.5 Å after 50 ns, drifting to ~2.6 Å after ~100 ns and stabilizing there for the remainder of the simulation. This extended drift phenomenon has also been observed in a previous MD study of the membrane protein OmpA in a DPC micelle [150], and a comparable backbone RMSD value of ~2.0 Å has previously been reported in a 50 ns MD simulation of the membrane channel protein FhuA in *N*-octyl-hydroxyethyl sulfoxide (OES) detergent [151]. However, these values fall within the same

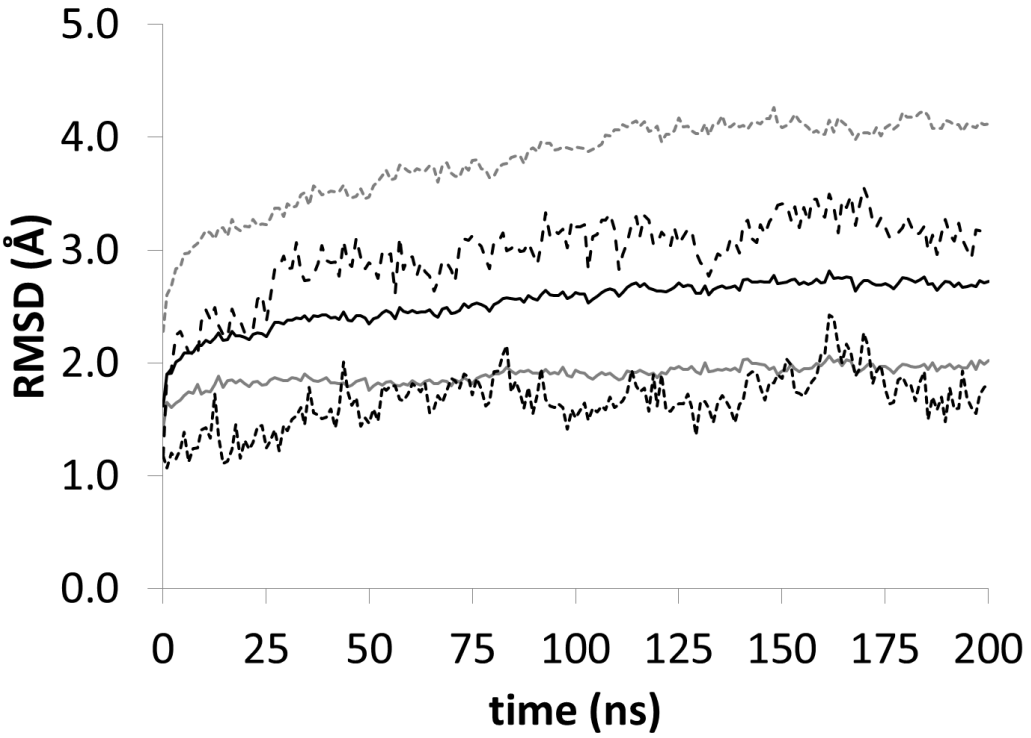

Figure 26. C_α RMSD versus time for the PSI/DDM void-filled ring model simulation. Lines show the C_α RMSD values for: all residues (black continuous line), the reaction center subunits psaAB (gray continuous line), the terminal electron acceptor psaC (black dotted line), the ferredoxin docking subunits psaDE (black dashed line), and the peripheral transmembrane helices psaFIJKLMX (gray dotted line). All curves were generated from the simulation starting structure. (Figure from Harris et al. 2014)

range as those reported from MD studies of membrane proteins in lipid bilayers. For example, C_α RMSD values of 2.0 and 2.3 Å were obtained from 15-20 ns MD simulations of OmpA in a dimyristoyl phosphatidylcholine (DMPC) bilayer and KcsA in a POPC bilayer, respectively [203]. Similar results were also seen in a 10 ns simulation of PSII embedded in the thylakoid membrane [139].

For detailed analysis of the protein structural drift, we decomposed the RMSD values into those of the various individual structural components of PSI, which are highlighted in Figures 25A and 25B. These domains are as follows: subunits psaA and psaB, which house the light-harvesting reaction center and associated pigments; psaC, found on the stromal face and housing the terminal electron acceptors, the iron-sulfur clusters F_A and F_B; subunits psaD and psaE, which are

also located on the stromal face and are postulated to facilitate the docking of the soluble electron mediator ferredoxin [204]; and the peripheral transmembrane α-helices psaI, psaJ, psaK, psaL, psaM, and psaX, as well as psaF, an exterior helix which traverses the membrane and is believed to be involved in the docking of ferredoxin on the stromal face as well as of electron carriers such as cytochrome c_6 and plastocyanin on the lumenal face [10]. The reaction center subunits psaA and psaB were found to display the lowest structural drift with a final C_α RMSD of ~1.8 Å (Figure 26, gray solid line). This is to be expected, as these subunits represent the core of the protein complex and thus have little exposure to the solvent, an explanation which is consistent with root-mean-squared fluctuation (RMSF) analysis discussed below. The terminal electron acceptor psaC also exhibits C_α RMSD values in the range of 1.5-2.5 Å (Figure 26, black dotted line), albeit with larger fluctuations. This can be explained by the fact that this subunit is located on the stromal surface and therefore has little interaction with the detergent but is highly exposed to the solvent. In contrast, the peripheral transmembrane helices exhibit the greatest drift with a maximum C_α RMSD of ~4.0 Å (Figure 26, gray dotted line), a plateau which is reached after 125 ns of MD simulation and is stable for the remainder of the 200 ns simulation time, indicating this region has reached equilibrium. This substantial structural drift is likely due to interactions with the surrounding detergent molecules in the protein-detergent complex, an environment that is quite different from the low temperatures and tight helical packing necessary for obtaining the high-resolution crystal structure which these RMSD values were generated relative to. As is the case for psaK, these protein-detergent interactions can cause the transmembrane helices to bend and contort during the simulation, becoming more compact and kinked as the detergent molecules surrounding them evolve from the initial uniform belt (Figures 27A and 27B). Chandler et al. observed a similar disparity in their study of purple bacterial chromatophores, wherein they conducted a 20 ns MD simulation of a model chromatophore complex of light harvesting complex LH1 and the reaction center (RC) embedded in a lipid bilayer

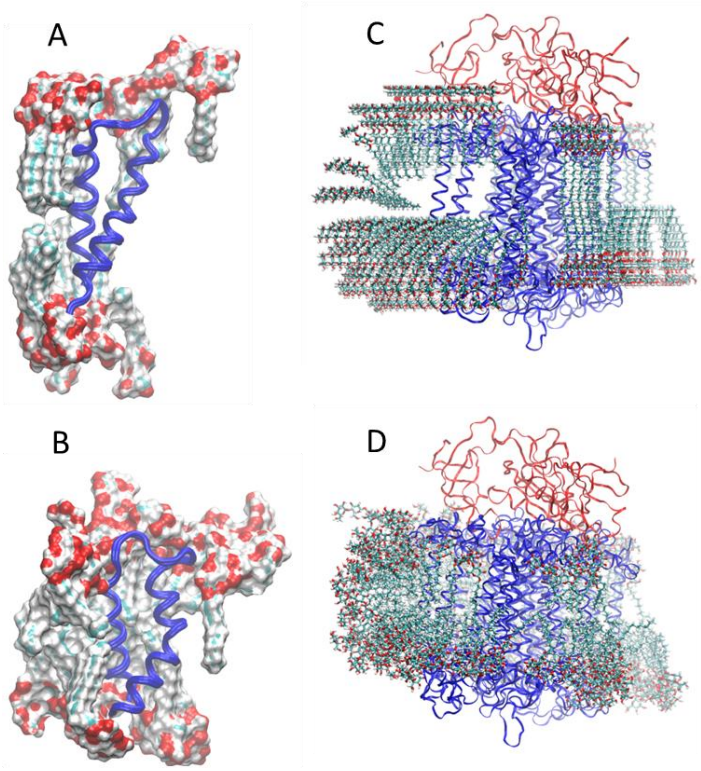

Figure 27. MD snapshots of the peripheral transmembrane helix psaK (A) prior to simulation and (B) after 200 ns of MD simulation. PsaK is shown in blue ribbons with the surrounding detergent molecules in surface format. Snapshots of the core subunits psaA and psaB (C) prior to simulation and (D) after 200 ns of MD simulation. The protein is shown in blue ribbons with the surrounding detergent molecules in licorice format. (Figure from Harris et al. 2014)

and reported RMSD values of ~4 Å for the core RC protein and ~8 Å for the peripheral LH1 domain [205]. C_α RMSD values in the range of 5-9 Å have also been obtained in MD studies of the membrane proteins Mistic [206] and aquaporin[152] in detergent micelles. The outer stromal subunits psaD and psaE show unique behavior, with the C_α RMSD rising quickly to ~2.5 Å, climbing to ~3.0 Å after 50 ns and fluctuating between 2.5-3.5 Å for the remainder of the simulation (Figure 26, black dashed line). This can be explained by the thinning of the detergent belt over the course of the simulation as it conforms to the hydrophobic periphery of the PSI trimer, a rearrangement that results in detergent molecules moving farther away from these stromal subunits (Figures 27C and 27D). In terms of detergent behavior, it should be noted that individual

detergent molecules displayed significant translational mobility in both the protein-detergent complex and pure detergent micelle simulations, but that no loss of detergent was observed in either case.

Protein flexibility

To characterize local protein mobility, we calculated the time-averaged root-mean-square fluctuation (RMSF) during the simulation for each residue of PSI (Figure 28). RMSF is a measure of the deviation of backbone atoms (N, C_α, and C atoms) from their average positions during the MD simulations, and provides insight into the thermal fluctuations and atomic mobility of proteins and protein

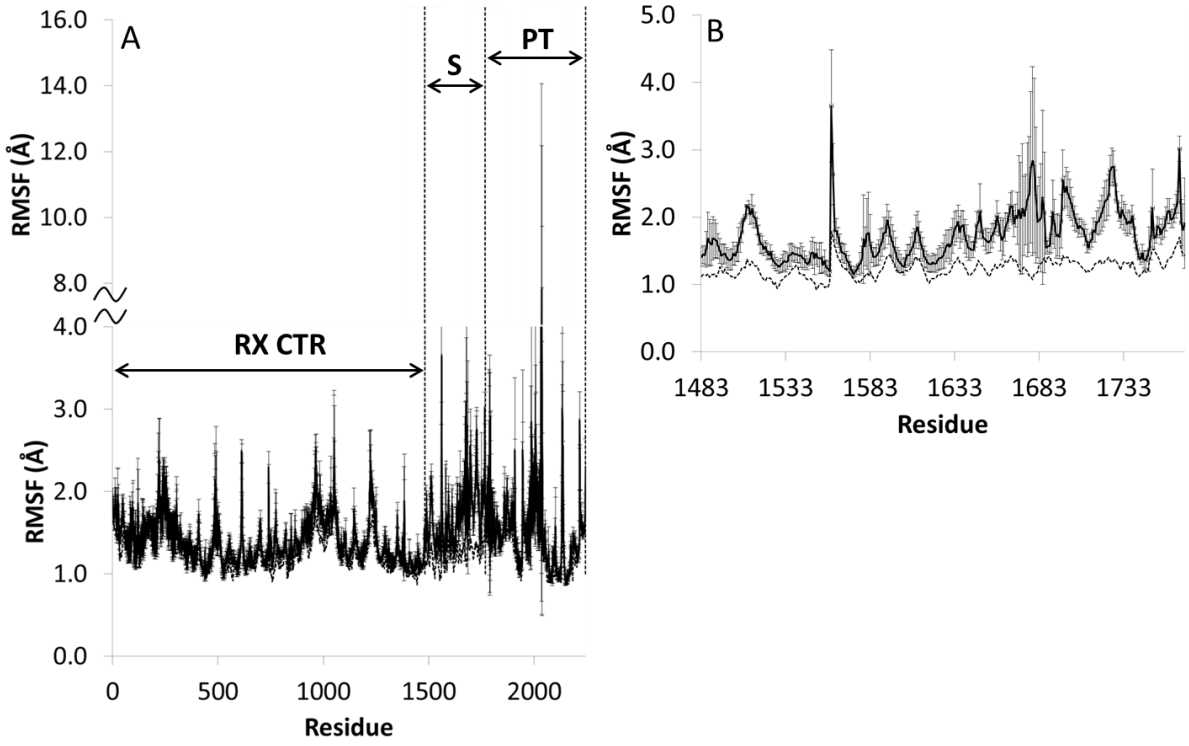

Figure 28. (A) Broken-line plot of C_α RMSF values versus residue number for PSI (solid black line), averaged over the last 100 ns of the PSI/DDM void-filled ring MD simulation. RMSF values based on the temperature values from the X-ray structure (dotted black line) are shown for comparison. RMSF values were averaged over all three monomers of the PSI trimer, with error bars representing the standard deviation. Residues corresponding to the reaction center (RX CTR), stromal (S), and peripheral transmembrane (PT) domains are labeled. (B) Zoomed-in view of time-averaged C_α RMSF values versus residue for the stromal domain. (Figure from Harris et al. 2014)

complexes. For comparison, we also converted the crystallographic B-factors of the 2.5 Å resolution PSI crystal structure to equivalent RMSF values (Figure 28, black dotted line). Qualitatively, the trends of the experimental and simulated curves are remarkably similar, with a clear correlation between local atomic mobility and protein structure. Regions of low local mobility are primarily located in the core of the reaction center subunits which are not solvent-exposed. The RMSF values calculated from the simulation are virtually identical in these regions when compared to the crystal structure-derived values. In contrast, the RMSF values obtained from the simulation for turns in the transmembrane helices of both the reaction center and peripheral transmembrane domains are significantly higher than those obtained from experiment. This is also the case for solvent-exposed loops found on the stromal face (Figure 28B). Overall, the RMSF results show that regions confined to the interior of the protein complex or shielded by detergent exhibit low local mobility, with higher fluctuations seen in solvent-exposed loops and turns. The noticeably high maximum RMSF value of ~7.4 Å corresponds to the N-terminus of psaL, the occurrence of which is due to the extension of this flexible region from the confines of the interstitial void out into the solvent over the course of MD simulation (as shown in Figure 29). This may be due to the absence of associated lipids from the simulation, and will be discussed in detail below. Comparable RMSF values for N- and C-terminal regions of transmembrane helices were observed in an MD study of aquaporin-0 in a lipid bilayer [207].

4.3.2 Micelle dynamics

Micelle shape

We were also interested in comparing the behavior of the detergent in the PSI/DDM complex to that of a pure DDM micelle, in order to elucidate the effects of the protein on detergent behavior and vice versa. Therefore, a 100 ns MD simulation was conducted for a pure DDM micelle using 132 DDM monomers, the previously reported aggregation number for DDM [144]. Simulation details for all cases are shown in Table 3. Visually inspecting changes in the shape of the

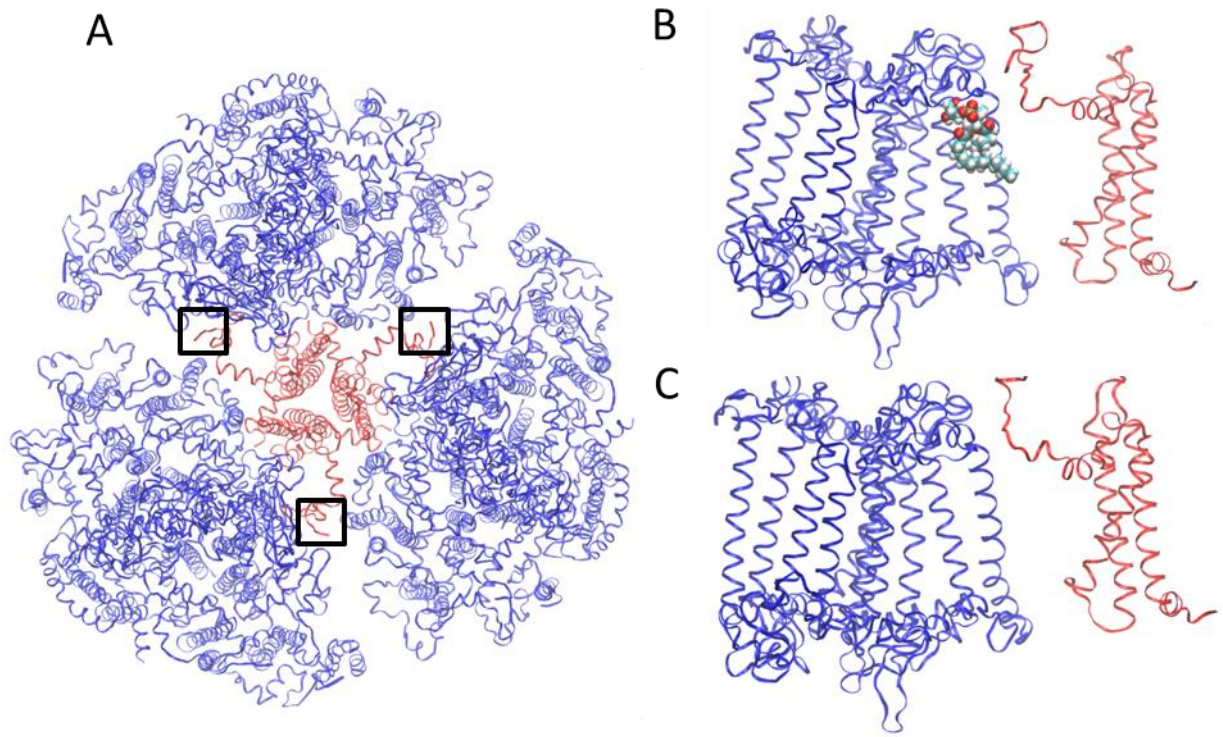

Figure 29. (A) Top view of trimeric PSI shown in blue ribbons, with the psaL subunits shown in red and their N-termini highlighted in boxes. Side view of the psaA (blue) and psaL (red) subunits shown (B) after 40 ns of MD simulation with the associated lipids identified in the crystal structure present (shown as VDW spheres) and (C) after 200 ns of MD simulation without the associated lipids included. (Figure from Harris et al. 2014)

each system throughout the course of MD simulation, we note that in the case of the PSI/DDM complex, the thickness of the detergent belt shrinks during dynamics as it conforms to the non-uniform hydrophobic periphery of the PSI trimer (Figures 30A and 30B). Overall, the protein-detergent complex becomes more ellipsoidal in shape over the course of the simulation. Initially constructed in a spherical configuration, the pure DDM micelle also transitions to an ellipsoid over the course of MD simulation (Figures 30C and 30D).

Interfacial properties

Another interesting aspect of these systems is the packing of detergent molecules and the resulting extent of water penetration, as it can influence the

Table 3. Simulation details for PSI/DDM models

System	Components	No. of water molecules	No. of detergents	No. of lipids	No. of ions	No. of atoms	Box size (nm)	Simulation time (ns)
PSI/DDM void-filled ring model	PSI, DDM	248,861	1,002	-	44 Na^+	977,756	28.0 x 28.0 x 13.5	200
PSI/DDM w/ lipids model	PSI, DDM, LMG, LHG	245,775	1,002	3 LHG, 1 LMG	283 Na^+, 230 CL^-	970,497	28.0 x 28.0 x 13.5	40
PSI/DDM random model	PSI, DDM	912,442	997	-	44 Na^+	2,986,094	35.0 x 35.0 x 25.0	20
DDM micelle	DDM	26,116	132	-	-	89,040	10.0 x 10.0 x 10.0	100

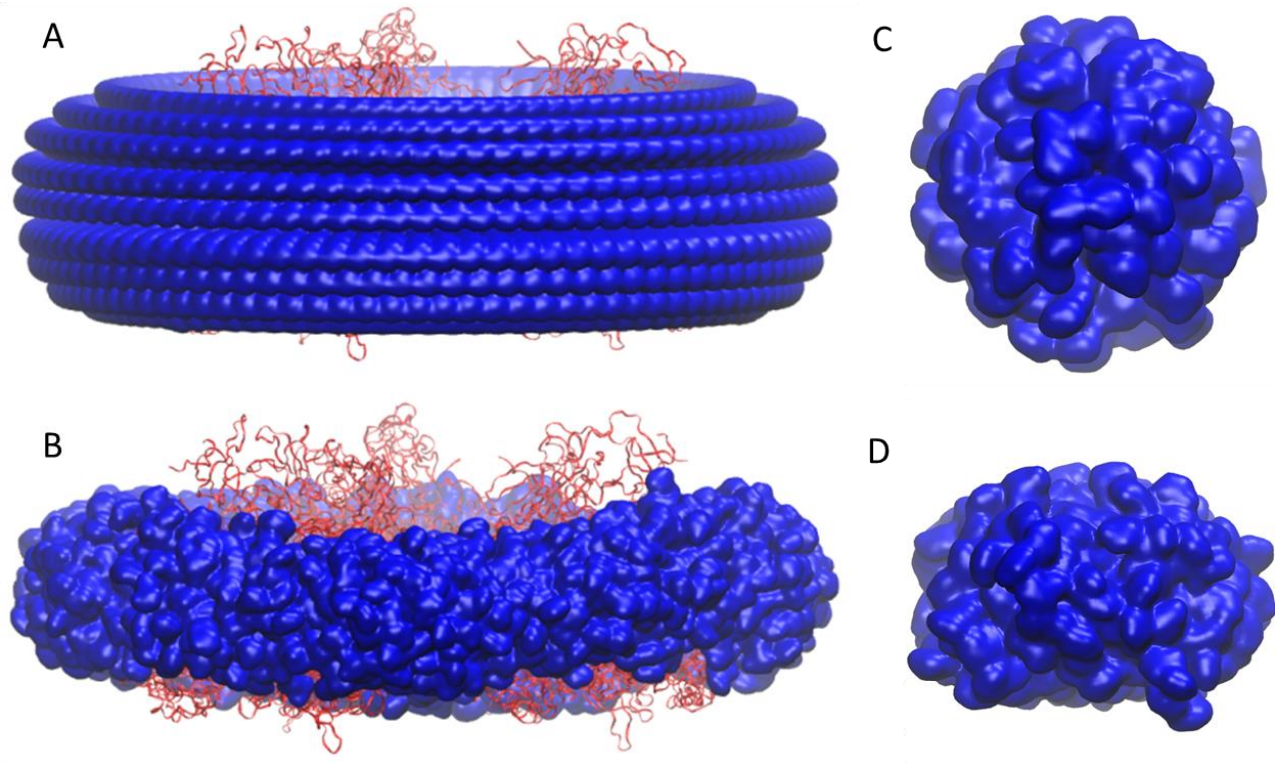

Figure 30. Side view of the PSI/DDM complex (A) prior to MD simulation and (B) after 200 ns of MD simulation. Side view of the pure DDM micelle (C) prior to MD simulation and (D) after 100 ns of MD simulation. The protein is shown in red ribbons, and the detergent in blue in low resolution surface representation. Water and counterions are omitted for clarity. Note that these images are not to scale. (Figure from Harris et al. 2014)

internal dynamics of the micellar aggregate as well as the protein complex embedded in the micelle. Therefore, we generated radial density profiles relative to the center of mass for both the pure DDM micelle and the PSI/DDM complex (Figure 31). In both cases, the detergent atom distributions and solvent-detergent interface are quite broad, as has been observed in previous studies. In the case of the PSI/DDM complex, there is a shift in the distributions of the detergent head and tail atoms compared to the pure DDM micelle, presumably due to protein-

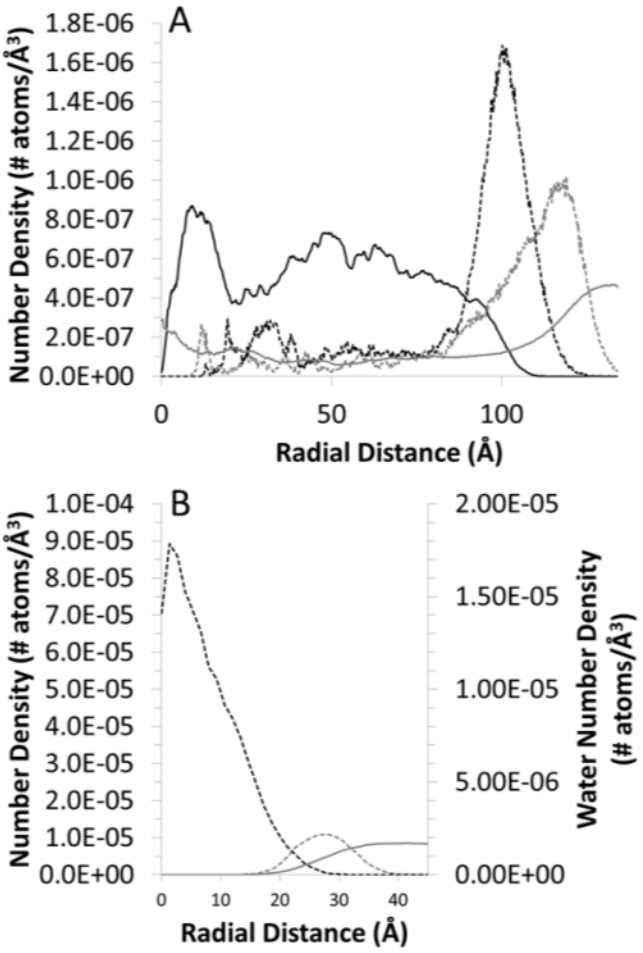

Figure 31. Radial atomic density profiles for (A) the PSI/DDM complex, and (B) the pure DDM micelle. In each case the atomic densities of the system components (protein = solid black line; detergent tails = dotted black line; detergent heads = dotted gray line; water = solid gray line) are plotted as a function of the distance from the center of mass of the system. Note that the water atomic density curve is plotted on a separate scale in (B). (Figure from Harris et al. 2014)

detergent interactions. In the pure detergent micelle system, the density profile of the detergent tail atoms is asymmetric, peaking sharply at 1.0 Å before gradually decreasing toward the center of the micelle, and displaying a width of ~30 Å between minima (Figure 31B, dotted black line). The non-zero density at the center of the micelle illustrates the tightly packed nature of the detergent tails in this system. In the protein-detergent complex, the tail atomic density profile is more symmetric, with a peak at 100 Å that reduces to zero by 80 Å due to the presence of the protein at the center of the complex; the width of the curve also broadens to ~40 Å (Figure 31A, dotted black line). A comparison of the detergent head atom densities for both systems reveals the opposite trend, with the DDM micelle system exhibiting a symmetric head atomic density profile with a peak at 29 Å (Figure 31B, dotted gray line), while the PSI/DDM complex displays an asymmetric curve with a peak at 119 Å (Figure 31A, dotted gray line). The width of the head atom distribution broadens from ~35 Å for the detergent micelle to ~50 Å for the protein-detergent system. In both the pure DDM and PSI/DDM systems, the detergent tails are more localized with high peak atomic densities, whereas the head group peak densities are less pronounced; this trend was also observed in MD simulations of aquaporin-0 in a lipid bilayer [207].

We also plotted the atomic density curves before and after MD simulation for both the pure DDM micelle and PSI/DDM complex to observe how the component profiles are changing (Figure 32). In the case of the detergent micelle, the tail atomic density profile sharpens over the course of MD simulation as the detergent tails re-arrange from the initial configuration to become more densely packed. Meanwhile, the head atomic density curve shifts outward, possibly due to the transition from a spherical to an ellipsoidal micelle. For the PSI/DDM complex, a similar sharpening of the tail atom density profile is observed, while changes in the head atomic density curve are minimal. Table 4 contains the atom densities at various time points for both systems, showing that the radial distances corresponding to the maximum head and tail atomic densities are stable over the course of MD simulation.

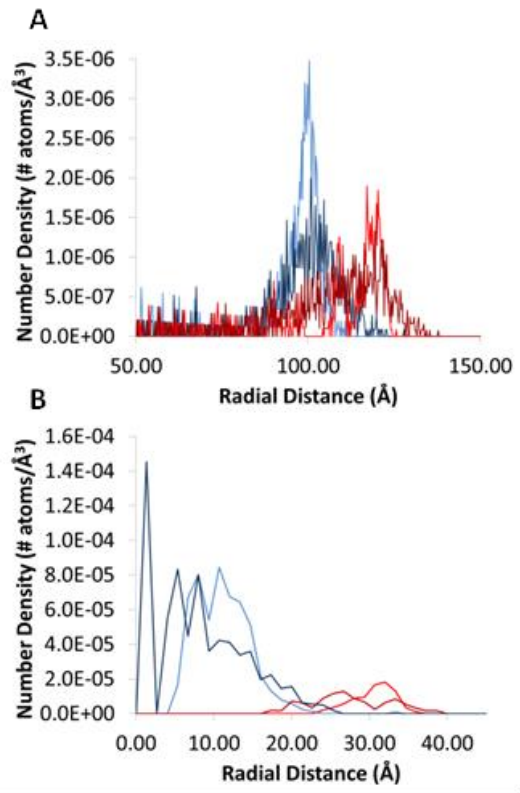

Figure 32. Radial atomic density distributions before and after MD simulation for (A) the PSI/DDM complex, and (B) the DDM micelle. In each case, the atomic densities of the various components (hydrophilic heads before simulation = dark blue line; hydrophilic heads after simulation = light blue line; hydrophobic tails before simulation = dark red line; hydrophobic tails after simulation = light red line) are plotted as a function of the distance from the center of mass of the system. (Figure from Harris et al. 2014)

The water atomic density curve is similar in both the DDM micelle and PSI/DDM systems (Figure 31, solid gray line), approaching the bulk density value before the detergent head atom density has reached zero, indicating solvation of the hydrophilic detergent headgroups. This means that water significantly penetrates the detergent in both systems. Solvation of the DDM headgroups was also observed in an MD study of pure DDM micelles in water [144]. In the case of the PSI/DDM complex, the non-zero water atomic density throughout the protein is due to penetration of water into cavities on the stromal and lumenal faces. We also analyzed the distance between the head and tail atomic density peaks for

Table 4. Radial distances for maximum atomic densities of detergent head and tail atoms

PSI/DDM complex			DDM micelle		
Time (ns)	Detergent Head (Å)	Detergent Tail (Å)	Time (ns)	Detergent Head (Å)	Detergent Tail (Å)
0	117.0	100.3	0	32.0	10.7
50	120.8	100.3	25	28.0	1.3
100	117.8	99.8	50	28.0	8.0
150	117.3	101.5	75	26.7	2.7
200	119.3	100.8	100	26.7	1.3

both systems. This value decreases from 28 Å for the pure DDM micelle to 19 Å in the case of the PSI/DDM complex, which suggests that the head-to-tail length of the detergent molecules is shrinking in order to match the detergent belt thickness to that of the exposed hydrophobic transmembrane region of the PSI trimer. This thinning of the detergent belt is clearly visible in both Figures 27C and 27D and Figures 30A and 30B, where DDM molecules depart from the initial uniform belt configuration as they shrink and conform to the transmembrane periphery of PSI over the course of MD simulation.

Another interesting aspect of these systems is the solvent-accessible surface area (SASA) of the detergent atoms (Table 5). Calculating the SASA requires drawing a mesh of points extended by extending the known radius of each specified atom by the radius of a water molecule, and checking these points against the surface of neighboring atoms to determine if they are buried or accessible. The number of accessible points is then multiplied by the surface area that each point represents to determine the SASA. This value can be used as a measure of water accessibility and thus detergent packing. We have also plotted the SASA versus time for both systems (Figure 33), which shows that the SASA quickly reaches an equilibrium value in each case and further confirms the stability of these systems during MD simulation. The higher SASA value of 285.7 ± 9.6 $Å^2$ for the PSI/DDM complex, versus 247.6 ± 12.3 $Å^2$ for a pure DDM micelle, indicates a slightly more diffuse packing of detergent molecules for the protein-detergent system. Analyzing the SASA of the individual detergent head

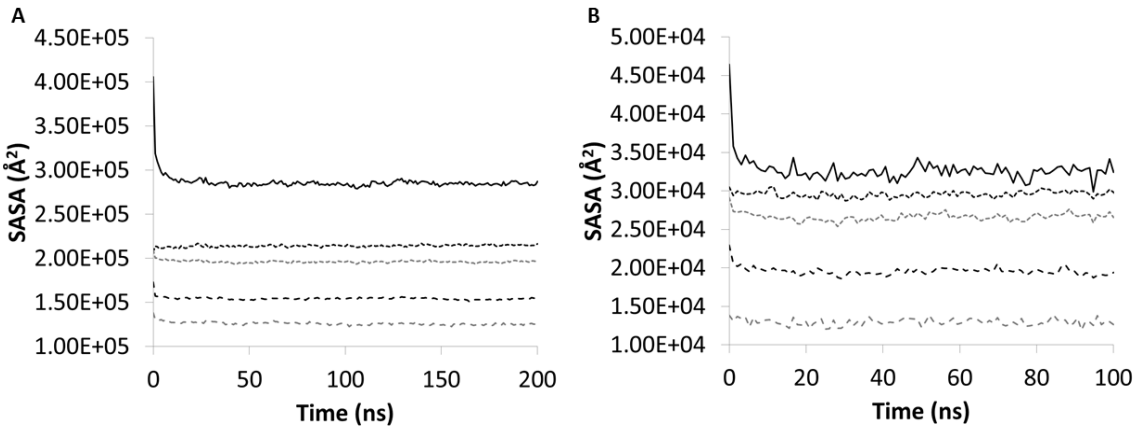

Figure 33. Detergent SASA versus time for (A) the PSI/DDM complex, and (B) the DDM micelle. SASA values are shown for: all detergent atoms (black solid line), the first head group (black dotted line), the second head group (gray dotted line), the upper tail group (black dashed line), and the lower tail group (gray dashed line). (Figure from Harris et al. 2014)

and tail groups (Table 5) reveals that the SASA for the head atoms is actually lower for the PSI/DDM complex than the pure DDM micelle, while the tail atoms exhibit the reverse trend (see Table 6 for single DDM molecule schematic). This

Table 5. Solvent-accessible surface area per detergent

System	Total (Å^2)	1st Head[a] (Å^2)	2nd Head[b] (Å^2)	Upper Tail[c] (Å^2)	Lower Tail[d] (Å^2)
PSI/DDM complex	285.7 ± 9.6	213.7 ± 1.3	195.9 ± 1.5	154.1 ± 1.7	125.9 ± 1.7
DDM micelle	247.6 ± 12.3	224.0 ± 3.1	201.7 ± 3.7	148.2 ± 4.0	98.1 ± 3.3

For each system, SAS was calculated using the VMD plugin with a probe radius of 1.4 Å

[a] The first six carbon and five oxygen atoms (first glucose ring)

[b] The second six carbon and oxygen atoms (second glucose ring)

[c] The first six carbon atoms of the hydrocarbon tail

[d] The last six carbon atoms of the hydrocarbon tail

Table 6. Comparison of the properties of the detergent tail

	Dihedral Angle Order Parameter (S^2)		MSD of Tail Hydrogen Atoms ($Å^2$)	
Carbon #*	**PSI/DDM**	**DDM**	**PSI/DDM**	**DDM**
1	0.057 ± 0.012	0.026 ± 0.036	58.0 ± 22.5	84.7 ± 33.4
2	0.049 ± 0.014	0.020 ± 0.033	56.6 ± 21.9	86.0 ± 33.8
3	0.050 ± 0.014	0.019 ± 0.035	56.0 ± 21.6	85.5 ± 33.6
4	0.052 ± 0.012	0.025 ± 0.033	55.6 ± 21.4	86.1 ± 33.8
5	0.047 ± 0.013	0.021 ± 0.033	55.5 ± 21.2	86.6 ± 33.8
6	0.045 ± 0.013	0.021 ± 0.031	55.8 ± 21.2	88.0 ± 34.3
7	0.042 ± 0.013	0.018 ± 0.037	56.4 ± 21.3	89.4 ± 34.7
8	0.037 ± 0.012	0.015 ± 0.032	57.6 ± 21.7	91.7 ± 35.7
9	0.027 ± 0.012	0.016 ± 0.030	59.2 ± 22.4	94.3 ± 36.8
10	0.020 ± 0.011	0.006 ± 0.032	61.5 ± 23.7	97.7 ± 38.7

*Numbered from the first (closest to the head group) to last CH_2 group of the detergent tail
These values are time-averaged for the last 100 ns and the last 50 ns of MD simulation for the PSI/DDM and DDM micelle systems, respectively.

indicates that the head groups are more tightly packed in the protein-detergent complex compared to the detergent micelle, while the tail groups are more loosely arranged. This suggests that the detergent belt surrounding PSI is becoming more like a bilayer structure than a micellar aggregate over the course of MD simulation.

Internal structure and dynamics

To determine the effect of protein-detergent interactions on the internal dynamics of the detergent, we calculated the dihedral order parameter (S^2) for the detergent tail groups of both the PSI/DDM and DDM micelle systems (Table 6). S^2 is a measure of the equilibrium distribution of the orientation of hydrocarbon chains and gives insight into the chain conformations and

fluctuations in the detergent micelle interior [208]. Values for S^2 can vary from 0.0 to 1.0, spanning a range from random, uninhibited fluctuations to rigidly fixed conformations. Time-averaged S^2 values of tail CH_2 groups (a single DDM molecule is shown with Table 6 for reference) are consistently higher for the PSI/DDM complex compared to pure DDM, suggesting that the presence of protein results in more ordered detergent tail structures. This is in agreement with the favorable interactions we have noted throughout this chapter between the detergent and the hydrophobic transmembrane region of PSI, such as the visible thinning of the detergent belt as it conforms to the protein as well as the shrinking of the average head-to-tail length of a DDM molecule for the PSI/DDM complex compared to a pure DDM micelle. We also note that the variance (standard deviation) of the time-averaged S^2 values is roughly equal to the mean in the case of the DDM micelle, possibly due to the greater fluctuations in this system compared to the PSI/DDM complex, whose S^2 standard deviation equates to ~20% of the mean. We also studied the diffusion of the detergent tail by analyzing the time-averaged mean-square displacements (MSDs) of the hydrogen atoms of each tail CH_2 group (Table 6). The MSD values for the PSI/DDM complex are lower than those obtained for the pure DDM micelle, again indicating that the detergent tails are more ordered (less diffusive) in the protein-detergent system. The results of the order parameter calculations and detergent tail diffusion analysis, combined with the SASA results, indicate that the detergent is overall more loosely arranged in the PSI/DDM complex compared to the DDM micelle, but that the detergent tails are more ordered. This is contrary to what is usually seen in lipid bilayer models, wherein the liquid disordered phase is more loosely packed than the liquid ordered phase [209, 210].

4.3.3 Implications for function

As discussed previously, PSI is an integral membrane protein involved in the photosynthetic cycle of plants and microorganisms, utilizing solar energy to accomplish electron transfer across the thylakoid membrane and generate reducing power for carbon fixation. Having established that our MD model of

trimeric PSI embedded in a DDM detergent ring is a stable complex, we also wanted to discern any potential effects of solution dynamics and protein-detergent interactions on the superstructure and function of this protein.

Changes in Trimeric Structure of PSI

It is known that PSI adopts a trimeric structure in cyanobacteria, but exists as a monomer in algae and higher plants. In the case of plant PSI, the peripheral interfaces of the PSI monomer participate in interactions with peripheral antenna, including the light-harvesting complex II (LHCII) [211]. As these organisms are typically found in land surface and shallow water environments, LHCII recruitment is necessary for regulation and protection from high light intensity and photodamage common to these environs. In contrast, cyanobacterial PSI does not possess an external antenna system and the peripheral interfaces of individual PSI monomers function primarily to maintain the trimeric superstructure of this protein. Cyanobacterial PSI may exist as a trimer in order to provide a larger antenna system for optimal capture of dim light in order to combat the low light conditions common to the natural habitat of these species [212]. The individual psaL subunits and their associated chlorophylls are hypothesized to be central to this process by serving as sites for excitation energy transfer between adjacent monomers in the PSI trimer. Experimental studies have shown that the growth of *T. elongatus* psaL deletion mutants under low light conditions was decreased by a factor of ten relative to wild type, and these mutants were also unable to form trimers [213].

A particularly intriguing result of this simulation study is the high fluctuations of the N-termini of the psaL subunits of the PSI trimer, resulting from the extension of these flexible regions out of the trimer voids into the solvent (Figure 29C). This extension is seen in two of the three psaL subunits of trimeric PSI and is a relatively fast event, occurring after ~20 ns of MD simulation. The X-ray crystal structure of cyanobacterial PSI contains several associated lipids, with one particular phospholipid identified as being bound to psaA and in close proximity

to psaL and the monomer-monomer interface [9]. We therefore conducted MD simulations of the PSI/DDM complex with the identified lipids included (see Materials and methods) in order to ascertain if this might have an effect on the observed extension of psaL, and found that the N-termini of all three psaL subunits remained in the trimer voids without extending into the solvent for the entirety of a 40 ns MD simulation (Figure 29B). The absence of these lipids from our MD model of the PSI/DDM complex could therefore contribute to the extension of the psaL N-termini during dynamics, and could also result in further destabilization of the trimeric superstructure of PSI on longer timescales, as the psaL subunits are vital to the trimerization process and these integral lipids are believed to be functionally important to PSI and not mere preparation artifacts. For comparison, similarly associated phospholipids have been found to be functionally important in the case of the yeast cytochrome bc(1) complex [214], as well as that of the photosynthetic reaction center of the purple bacterium *Rhodobacter sphaeroides* [215].

Docking of Soluble Electron Mediators

The particular role of PSI in photosynthesis is to use solar energy to accomplish the transfer of electrons provided by soluble electron carriers on the lumenal side of the thylakoid membrane to electron mediators on the stromal side. These transferred electrons will be used to reduce $NADP^+$ to NADPH, thus providing reducing power for the cell [10]. Researchers have thoroughly investigated these docking processes, including the docking of cytochrome c_6 (cyt c_6) and plastocyanin (PC) on the lumenal face of PSI, as well as of ferredoxin (Fd) on the stromal face. Sommer et al. determined the major interaction site for cyt c_6/PC docking to be a hydrophobic lumenal indentation formed by psaA and psaB, and also identified two tryptophan residues crucial to this docking process [200]. In terms of Fd docking, mutagenesis studies conducted by several research groups have identified residues of the stromal ridge subunits (psaCDE) that are involved in Fd docking on the stromal face through electrostatic

interactions [201]. Additionally, all-atom MD studies have previously been used to study the docking of cytochrome c_6 and plastocyanin to PSI monomer *in vitro* [216].

Using the docking prediction server ZDOCK [199] and our MD model of detergent-solubilized PSI, we have attempted to investigate the docking of the electron mediators cyt c_6 and Fd to PSI *in vitro* (see Materials and methods). Comparing docking predictions based on snapshots of the PSI/DDM MD simulation trajectory to those obtained using the PSI crystal structure, we were able to determine the effects of solution dynamics and protein-detergent interactions on cyt c_6 and Fd docking to PSI. In general, we found that the presence of detergent had little effect on the docking of soluble electron carriers, as the predicted docking based on the MD snapshots did not differ significantly

Table 7. C_α RMSD values for mediator docking to PSI/DDM MD trajectory

Time (ns)	Cyt c_6 Docking RMSD (Å)	Fd Docking RMSD (Å)
20	1.50	5.25
40	3.58	1.98
60	1.21	2.60
80	2.20	3.07
100	1.55	3.18
120	3.81	3.07
140	3.72	4.28
160	2.11*	3.38
180	1.63	2.83
200	1.77	2.14

*Predicted docking of cyt c6 was successful for only two of three PSI monomers for the 160 ns MD snapshot

the predicted docked structures based on the PSI crystal structure are shown in Table 7. Cyt c_6 docking to MD simulation snapshots differed by an average of from that based on the crystal structure. C_α RMSD values of predicted docked mediator structures based on the PSI/DDM MD simulation trajectory relative to 2.31 ± 1.01 Å relative to that based on the crystal structure, while Fd docking differed by an average of 3.18 ± 0.97 Å. We propose that these slight differences in binding can be attributed to fluctuations in the key binding residues referenced previously, which is indicative of an induced-fit mechanism for mediator docking.

For the case of ferredoxin, the difference in mediator docking can probably be attributed to the ~3.0 Å structural drift of the stromal ridge psaDE subunits relative to the crystal structure over the course of MD simulation (Figure 26, black dashed line), which results in conformational changes that alter Fd docking. Also of note is the fact that several key residues associated with Fd binding display high fluctuations during the simulation, specifically residues T14 and Q15 of psaC and residue K104 of psaD (Table 8). Experimental studies involving mutations of these residues in PSI are known to affect the charge of the binding pocket for Fd [201]. Also, the triple mutation I12V/T15K/Q16R of psaC in PSI from the green algae *Chlamydomonas reinhardtii*, which corresponds to I11V/T14K/Q15R in *T. elongatus* PSI, has been shown to reduce Fd binding by two orders of magnitude [217]. As such, fluctuations in these residues could also contribute to the altered binding of Fd to PSI under detergent-solubilized conditions.

The root cause of differences in the docking of cytochrome c_6 to PSI is not so clear, as this mediator binds to a hydrophobic indentation on the lumenal face formed by the psaAB subunits, which represent the reaction center of PSI and display the lowest structural drift during MD simulation of the PSI/DDM complex relative to the PSI crystal structure (Figure 26, gray solid line). Additionally, the

Table 8. RMSF values for PSI-mediator binding site residues

	Residue*	PSI/DDM MD Model RMSF (Å)	Crystal Structure RMSF (Å)
Stromal	T14C	1.96 ± 0.02	1.26
	Q15C	2.00 ± 0.03	1.25
	K104D	2.74 ± 1.13	1.13
Lumenal	R627B	1.61 ± 0.69	1.07
	D628B	1.64 ± 0.50	1.13
	Y629B	1.89 ± 0.56	1.17
	L630B	1.60 ± 0.55	1.13

*Format for residue identifier is one-letter amino acid symbol, residue number, subunit name

key residues believed to be involved in cyt c_6 binding, W655 from psaA and W631 from psaB, do not fluctuate noticeably. However, there are several residues of the lumenal helix *I* of psaB, specifically R627 through L630, that exhibit unusually high RMSF values (Table 8). Experimental studies involving mutations of helix *I* have shown that this loop is essential for binding and subsequent electron transfer with the soluble electron donors cytochrome c_6 and plastocyanin [218]. As such, it is possible that fluctuations in this luminal helix could contribute to the altered binding of cyt c_6 to PSI under detergent-solubilized conditions that we have observed.

4.4 Summary

We have conducted MD simulations of DDM detergent-solubilized trimeric PSI, one of the largest membrane protein complexes known to have so far been studied. RMSD analysis relative to the crystal structure has shown it to be a stable complex, with C_α RMSD values in a range that suggests the behavior of PSI in a DDM micellar environment is similar to that observed in MD studies of membrane proteins in lipid bilayers. Also, we identified the largest contributors to the structural drift of this protein to be the peripheral transmembrane subunits,

which are in close contact with detergent molecules and in general become more compact and kinked during the MD simulations. Investigating protein local mobility based on RMSF, we observe that regions of low atomic mobility are confined to the core residues of the reaction center which are not solvent-exposed, while regions of high fluctuation correspond to flexible loops of the stromal domains and turns in the transmembrane helices, all of which are exposed to solvent.

We further examined differences in detergent behavior in the protein-detergent complex relative to a pure detergent micelle. We observe a thinning of the detergent belt thickness over the course of MD simulation for the PSI/DDM complex as the detergent conforms to the non-uniform hydrophobic periphery of the protein. The complex as a whole becomes more ellipsoidal in shape over the course of MD simulation, a transition that was also observed in MD studies of the pure detergent micelle. Radial atomic density profiles for various system components indicates that the presence of the protein results in a broadening of the detergent head and tail atom density curves relative to the micelle system, but results in little change to the extent of water penetration. We also note that the distance between the peaks of the head and tail atom densities decreases from 28 Å for the pure DDM micelle to 19 Å for the PSI/DDM complex, indicating that favorable protein-detergent interactions involving the hydrophobic transmembrane region of PSI result in the shrinking of the detergent molecules. Solvent-exposed surface area (SASA) analysis of the detergent atoms suggests that the DDM detergent is more loosely arranged in the presence of protein compared to a pure detergent micelle, but with differing trends in the detergent head and tail groups. Based on evaluation of the dihedral angle order parameters (S^2) of the detergent tail CH_2 groups and MSD values for the associated hydrogen atoms, we have further concluded that the detergent tails are more ordered in the PSI/DDM complex compared to the DDM micelle. Taken in its entirety, we believe the behavior of the transmembrane domain of PSI and the surrounding detergent molecules in our simulations suggests a degree of

plasticity in the structure of the *in vivo* complex. The membrane-shielded region of the protein can adapt to the bilayer thickness and vice versa via kinking and tilting motions of the peripheral transmembrane helices and the rearrangement of surrounding lipid molecules, thus enabling PSI to adapt to different bilayer environments, such as different lipid types or phases, etc.

We have also attempted to ascertain any functional consequences for this photosynthetic protein complex as a result of solution dynamics and protein-detergent interactions. We believe the extension of the N-termini of the psaL subunits from the trimer voids out into the solvent during MD simulation is due to the absence of integral lipids identified in the PSI crystal structure from our model, and believe this behavior could have an impact on the trimeric structure of PSI on longer timescales. We have determined that docking of the soluble electron mediators cytochrome c_6 and ferredoxin is not hindered by the presence of detergent, differing only slightly for predictions based on our MD trajectory of detergent-solubilized PSI relative to predictions based on the known crystal structure of PSI. We propose that the slight changes in binding are due to fluctuations in key binding site residues during the MD simulation, which may imply an induced-fit mechanism for mediator docking. These results yield new insights into the structural integrity and activity of PSI *in vitro*.

In conclusion, the use of all-atom molecular dynamics simulation and subsequent analyses conducted herein demonstrate a novel approach to understanding the solution structure and dynamics of detergent-solubilized membrane proteins at the atomic level. Furthermore, the results yielded by this study in particular offer a solid starting point for understanding the *in vitro* structure and dynamics of the photoactive pigment-protein complex photosystem I, an integral membrane protein that functions as a key component in the photosynthetic cycle of plants and microorganisms. Of particular significance is the indication that integral lipids may be crucial to maintaining the stability of trimeric PSI in detergent solution. Also, the presence of DDM detergent does not appear to interfere with PSI activity.

Having established our ability to computationally model PSI, we can easily translate this MD simulation approach to studying complexes of PSI and other proteins, such as hydrogenase enzyme. Such work can yield new insights into solution dynamics and protein-protein interactions in multi-protein systems, and in the particular case of PSI-hydrogenase fusions, potentially allow us to predict optimal conditions for electron transport between these two proteins.

CHAPTER V
ENGINEERING OF PHOTOSYSTEM I – HYDROGENASE PROTEIN FUSIONS USING SORTASE-MEDIATED LIGATION

5.1 Introduction

As stated previously, gaining knowledge of the effects of protein-environment and protein-protein interactions on the stability and activity of proteins and multi-protein complexes will advance our scientific understanding of the biological world and could have a significant impact for applications in a wealth of areas such as alternative energy, drug delivery, etc. This is especially true in the case of integral membrane proteins such as PSI, which mediate a variety of functions within the cell.

PSI accomplishes the unique biological function of converting solar energy into reducing power, and as such is actively researched as an alternative energy solution, as has been reviewed before. In particular, cell-free photoproduction of H_2 has been demonstrated in a variety of PSI-based systems, including chemically platinized thylakoid membranes [12, 23] and isolated PSI trimers [13, 14, 24], as well as complexes of PSI and hydrogenase enzyme [16, 37, 38, 163].

In previous chapters, we presented our study of the solution structure of DDM detergent-solubilized cyanobacterial PSI using a combination of SANS and MD simulation [195], and our in-depth analysis of this system using all-atom MD simulations of trimeric PSI embedded in a DDM detergent ring. Our SANS study revealed that the dimensions of the PSI/DDM complex were consistent with trimeric PSI embedded in a disk-like detergent micelle. Conducting MD simulations on a PSI/DDM complex constructed based on these SANS results, we have shown this to be a stable system, and have gained new insights into the effects of protein-detergent interactions on detergent organization and packing that suggests a degree of plasticity in the *in vivo* complex of PSI embedded in the thylakoid membrane. Furthermore, this simulation study has revealed new information on the structural integrity and activity of PSI *in vitro* that could prove

important in PSI-based energy conversion devices. Applying a similar combination of experimental and computational approaches to the study of PSI in complex with other proteins could provide new insights on electron transport in redox proteins and the structure and function of multi-protein complexes in general, and the study of PSI-hydrogenase fusions in particular could provide new insights for alternative energy solutions via the photoproduction of H_2.

In this chapter, we present our work towards experimentally generating a site-specific fusion of PSI from *T. elongatus* and the membrane-bound [NiFe]-hydrogenase (MBH) from *R. eutropha*. We accomplish this ligation using sortase, a transpeptidase enzyme found in most Gram-positive bacteria whose *in vivo* function is the covalent anchoring of a variety of surface proteins to the cell wall envelope [219]. The particular isoform used here, *Staphylcoccus aureus* sortase A (SrtA), has been demonstrated as a viable option for a wide variety of protein engineering and bioconjugation applications [220, 221]. In parallel, we have also conducted MD simulations of this PSI-MBH fusion protein, as well as of PSI-molecular wire-[FeFe]-hydrogenase complexes previously studied experimentally by Golbeck and co-workers [16, 38], to allow for atomic-level analysis of the effects of solution dynamics and protein-protein interactions on the structure and function of multi-protein complexes, and potentially predict optimal linkage strategies for electron transport between PSI and hydrogenase.

5.2 Materials and methods

5.2.1 Strains and plasmids

The strains and plasmids used in this study are listed in Table 9. *Ralstonia eutropha* H16 (ATCC 17699, DSM 428) is the wild-type strain harboring the endogenous megaplasmid pHG1, a large operon containing all the structural, accessory, and regulatory genes necessary for the synthesis of active MBH [222]. *R. eutropha* HF387 is a derivative of wild-type *R. eutropha* H16 that lacks the NAD^+-dependent soluble hydrogenase (SH) because of an in-frame deletion

Table 9. Relevant bacterial strain/plasmid information

Strains and plasmids	Relevant characteristic(s)	Source
R. eutropha		
H16	Wild-type	[222]
HF387	Derivative of H16, ΔhoxH	[223]
HF387H	Derivative of HF387H, hoxK-His_6	[224]
NGLY3	Derivative of HF387H, Gly_3-hoxK-His_6	This work
Synechocystis		
PCC 6803	Wild-type	[225]
CLPETG	Derivative of 6803, psaC-LPETG	[226]
DLPETG	Derivative of 6803, psaD-LPETG	[226]
ELPETG	Derivative of 6803, psaE-LPETG	[226]
E. coli		
JM109	*endA1 gyrA96 hsdR17 Δ(lac-proAB) recA1 relA1 supE44 thi-1 F'[lacI^{q}lacZΔM15 proAB$^+$ traD36]*	Promega [227, 228]
S17-1	*Tra+ recA pro thi hsdR, chr.:RP4-2*	
Plasmids		
pLO3	Tet^R sacB RP4 OriT colE1 ori	[229]
pGEM-T Easy	Amp^R, lacZ, f1 ori, T7 gene promoter, Sp6 gene promoter	Promega
pHoxKG	6078 bp fragment containing hoxKG genes in pGEM-T Easy vector	This work
pHoxKGmod	Derivative of pHoxKG containing Gly_3 tag on N-terminus of hoxK	This work
pLO3-HoxKGmod	Derivative of pHoxKG in pLO3	This work

in the structural gene hoxH [223]. This strain was a kind gift from Dr. Oliver Lenz (Institute for Biology, Humboldt University of Berlin). *R. eutropha* HF387H is a derivative of *R. eutropha* HF387 containing a six histidine (His_6) purification tag on the C-terminus of the MBH small subunit hoxK, and was previously constructed in our lab by Iwuchukwu et al. [224].

Synechocystis sp. PCC 6803 was also used for this study. *Synechocystis* CLPETG, DLPETG and ELPETG are derivatives of *Syn.* 6803 containing an LPETG tag on the C-terminus of the psaC, psaD, and psaE subunits, respectively. These strains were constructed in our lab by Rosemary Le [226].

Escherichia coli JM109 was used as the recipient in standard cloning procedures, and *E. coli* S17-1 was used as the donor in conjugative gene transfer [227, 228].

5.2.2 Isolation of R. eutropha megaplasmid DNA

The megaplasmid DNA of *R. eutropha* HF387H was isolated as described previously [230]. Briefly, 1.5 mL of *R. eutropha* HF387H was grown in FGN minimal media for ~30 h. Cells were harvested via centrifugation at 8,000 x g for 1 min, and the cell pellet was gently resuspended in 500 µL of lysis buffer (0.3 M sucrose, 25 mM Tris-HCl pH 8.0, 25 mM EDTA pH 8.0). 5 mg of lysozyme and 25 µg of RNase A were added and the mixture incubated at 37°C for 20 min without shaking. 250 µL of 2% SDS was added, mixed by inverting and then incubated for 5 min at room temperature. 250 µL of phenol/chloroform/isoamyl alcohol (25:24:1) was then added and the mixture emulsified by tube inversion. The aqueous and phenol phases were subsequently separated via centrifugation at 12,000 x g for 5 min. The aqueous phase (~700 µL) was transferred to a clean DNase-free tube, and 0.1 vol. of 3 M sodium acetate pH 4.8 followed by 1 vol. of isopropanol were then added. The mixture was incubated at room temperature for 5 min, and the precipitated DNA recovered via centrifugation at 12,000 x g for 10 min. The supernatant was removed, leaving ~50 µL in the tube, and 500 µL ice-cold (-20°C) 70% ethanol was added. The tube was kept at -20°C for 30 min and the precipitated DNA recovered via centrifugation at 12,000 x g for 30 min at -2°C. The supernatant was discarded and any residual ethanol evaporated by speed vacuum, and the pellet was then dissolved in 50 µL of nuclease-free distilled water.

5.2.3 Genetic construction of N-terminally Gly_3-tagged hoxK

A Gly_3 tag fusion to the N-terminus of the MBH small subunit hoxK was constructed as follows. Using primers 1 and 2 (Table 10) and the megaplasmid DNA of HF387H as template, a 3 kbp fragment containing the hoxKG genes with ~1 kbp of homology on either side was amplified by PCR (*TaKaRa Ex Taq* DNA polymerase, Clontech) and subsequently T/A cloned into the cloning vector pGEM-T easy (Promega) to form pHoxKG. The resulting pHoxKG plasmid was transformed into chemically competent *E. coli* JM109 cells (Promega), and successful inserts selected using blue-white screening on ampicillin-containing media, as pGEM-T Easy contains the lacZ promoter as well as the ampicillin resistance gene ampR. These cells were inoculated in LB/ampicillin media, and the plasmid DNA isolated using the PureYield™ plasmid miniprep system (Promega) and the nucleotide sequence confirmed by sequencing.

The Gly_3 tag was subsequently introduced at the N-terminus of the matured protein, in order to account for cleavage of the membrane translocation signaling peptide [25, 231], via overlapping PCR. First, using primers 1 and 3 and pHoxKG as template, a 1 kbp fragment containing the hoxK gene modified to include the Gly_3 tag at its N-terminus was amplified by PCR (designated Gly_{UP} in Figure 34). Second, using primers 2 and 4 and again using pHoxKG as template, a 2 kbp fragment also containing N-terminally Gly_3 tag modified hoxK was amplified by PCR (designated Gly_{DOWN} in Figure 34). Finally, using primers 1 and 2 and the Gly_{UP} and Gly_{DOWN} PCR fragments as template, the complete hoxKG gene

Table 10. Oligonucleotide primers used for targeted mutagenesis of hoxK gene

Primer #	Designation	DNA Sequence 5'-3'
1	HoxKG_Forward	TTA**GAGCTC**AGCGCTTCGATCTGCAAGAACTAC
2	HoxKG_Reverse	TTACTCGAGGGCCTGTTTATACAGTGTGCCGATGG
3	TriGly_Upper	**GGAGGTGGA**ATGGAAACCAAGCCGCGTACACCA
4	TriGly_Lower	GGCTTGGTTTCCAT**TCCACCTCC**CGCGTGCGCGATCTGC

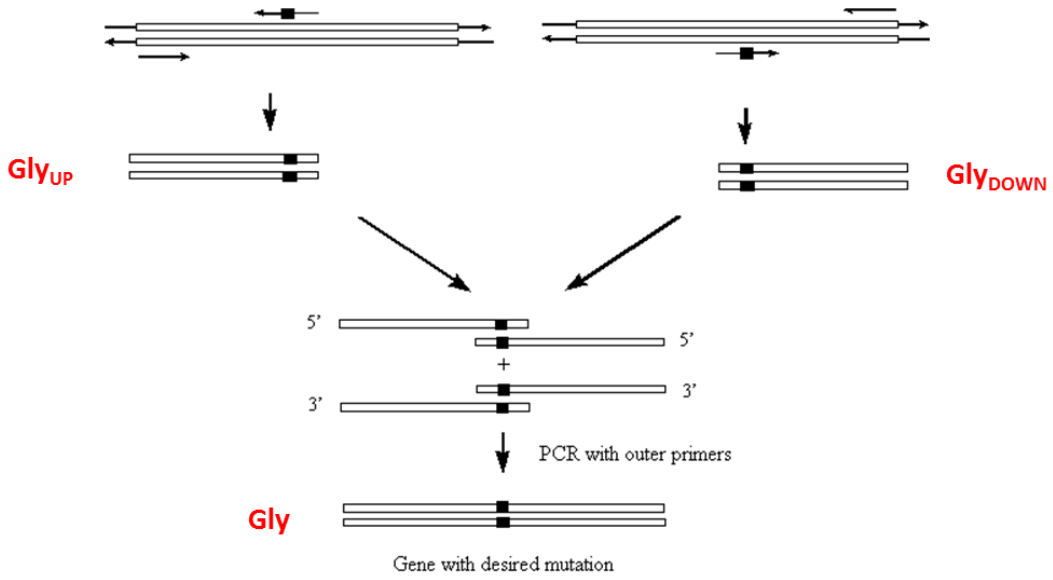

Figure 34. Schematic of overlapping PCR for introduction of Gly_3 tag to N-terminus of hoxK from *Ralstonia eutropha* HF387H.

fragment with the desired Gly_3 tag was amplified via PCR and T/A cloned into pGEM-T Easy to form pHoxKGmod. As before, this plasmid was transformed into *E. coli* JM109 and the plasmid DNA harvested via PureYield™ and sequence verified.

5.2.4 Conjugative plasmid transfer and gene replacement

In order to introduce the mutated enzyme back into *R. eutropha* via homologous recombination, SacI restriction enzyme digested fragments from pHoxKGmod were ligated into pLO3 to form pLO3-HoxKGmod. Conjugal gene transfer was then accomplished as follows (a schematic of this process is shown in Figure 35). Donor (*E. coli* S17-1) and recipient (*R. eutropha* HF387H) strains were grown to late exponential phase. 1 mL of the donor 3 mL of the recipient strains were concentrated via centrifugation (3,500 x g for 5 min) and washed three times with LSLB-MOPS media and then resuspended in 100 µL of LSLB-MOPS media. Mobilizable plasmids were transferred from *E. coli* S17-1 to *R. eutropha* HF387H via conjugation by spot mating 200 µL of the donor/recipient mixture on D-medium. After 16 h of incubation at 30°C in the dark, cells were re-suspended with 700 µL of FGN media and washed once by centrifugation at

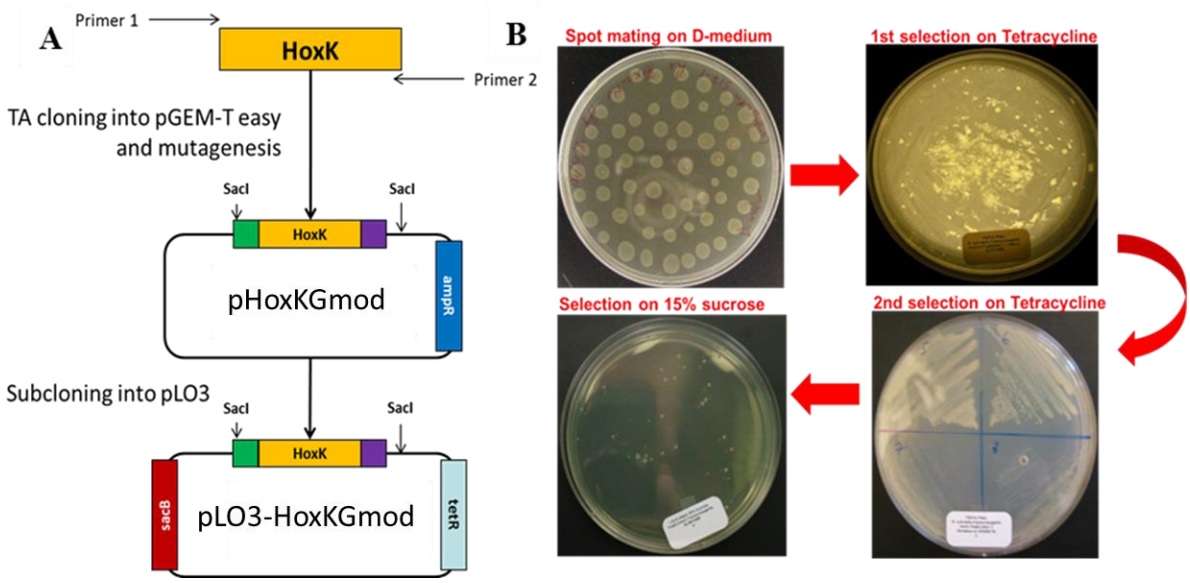

Figure 35. A) Strategy for introducing the Gly_3 tag to *R. eutropha* via conjugation (the green portion represents the Gly_3 tag and the purple portion the His_6 purification tag), and (B) the accompanying screening process (Figure from Iwuchukwu et al. 2011).

3,500 x g for 5 min. Several dilutions of the cell suspension were spread onto FGN plates containing 15 µg/mL tetracycline, with transconjugants appearing after 3-5 days incubation at 30°C. Tetracycline-resistant transconjugants were picked and purified twice by streaking onto FGN/Tet plates, then inoculated in 2 mL LSLB-MOPS and incubated overnight at 30°C. 100 µL of cells were then spread onto LB plates with 15% sucrose, and sucrose-resistant survivors picked for screening by isolation of the megaplasmid DNA and verification of the presence of the Gly_3 tag by DNA sequencing.

5.2.5 Purification of MBH from R. eutropha

R. eutropha strains were grown micro-aerobically as described previously [232], and the MBH was purified according to the method of Wisitruangsakul et al. [233]. Briefly, cells were grown in a modified FGN minimal medium containing 0.05% (w/v) fructose and 0.4% (w/v) glycerol. Four baffled 2 L Erlenmeyer flasks were filled with 1.6 L culture and shaken at 120 rpm and 30°C for approximately one week. Cells were harvested via centrifugation at 5,000 x g and 4°C for 15 min, and washed once with phosphate buffer (71 g L^{-1} $Na_2HPO_4 \cdot 7H_2O$, 15 g L^{-1}

KH_2PO_4). Resulting cell pellets were flash frozen in liquid N_2 and stored at -80°C. Cells were re-suspended in lysis buffer (50 mM KH_2PO_4 pH 8.0, 300 mM NaCl, 20 mM imidazole, 25 µL DNase I (Thermo Scientific), and 1 pellet of EDTA-free protease inhibitor cocktail (PIC, Roche)) at a ratio of 12 g cells (wet weight) to 60 mL lysis buffer. Cells were disrupted via three passages through a chilled French press (Aminco SLC) at a cell pressure of ~30,000 psi. Cell debris was removed via low-speed centrifugation at 4,000 x g and 4°C for 20 min. The pale brown pellet was discarded, and the cloudy reddish brown supernatant was subsequently ultracentrifuged at 36,000 rpm (ice-cold Beckman Type 60 Ti rotor) and 4°C for 45 min. The clear red supernatant was discarded, and the reddish brown membranes were resuspended in solubilization buffer (50 mM KH_2PO_4 pH 8.0, 300 mM NaCl, 20 mM imidazole, 2% Triton X-114, and 1 PIC pellet) at a ratio of 1.0 g membranes (wet weight) to 10 mL solubilization buffer. This mixture was then stirred at 4°C for 2 hours to solubilize the membrane proteins out of the native membrane, and subsequently ultracentrifuged at 36,000 rpm (ice-cold Beckman Type 60 Ti rotor) and 4°C for 45 min. The loose brown pellet was discarded, and the supernatant loaded onto a HisPur Ni-NTA column (3 mL bed volume (BV), Thermo Scientific) equilibrated with 20 mL of solubilization buffer. The column was then washed with 50 mL of solubilization buffer without detergent. Purified MBH was then eluted dropwise using 500 µL of elution buffer (50 mM KH_2PO_4 pH 8.0, 300 mM NaCl, 250 mM imidazole) until the A_{280} absorbance (NanoDrop spectrophotometer) reached zero. All purification steps were performed at 4°C. Fractions were pooled and concentrated using a centrifugal filter (Amicon Ultra-15 (30,000 MWCO), Millipore), buffer exchanged and stored at -20°C in storage buffer (50 mM KH_2PO_4 pH 5.5, 150 mM NaCl, 20% glycerol).

5.2.6 SDS-PAGE and immunoblot analysis

Protein concentrations were determined using the BCA method with bovine serum albumin as standard (Pierce). The purity of samples was estimated via visual inspection by SDS-PAGE electrophoresis of polyacrylamide gels and

subsequent staining with Coomassie Brilliant Blue G-250 (Fisher). For the detection of protein samples containing hoxK/G, the respective samples were resolved on pre-cast 10% Bis-Tris polyacrylamide gels (NuPAGE Novex, Life Technologies) and subsequently transferred to polyvinylidene fluoride (PVDF) membranes (Immobilon, Millipore). For immunological detection of MBH-related proteins, antisera were applied in the following dilutions: anti-hoxK serum (1:20,000) and anti-hoxG serum (1:1000). HoxK antiserum was a kind gift from Dr. Oliver Lenz (Institute for Biology, Humboldt University of Berlin), and hoxG antiserum was a kind gift from Dr. Carrie Eckert (Energy Sciences Division, National Renewable Energy Laboratory (NREL)). Secondary detection was accomplished via chemiluminescence of a horseradish peroxidase conjugate using the Clean-Blot IP Detection Kit (Thermo Scientific).

5.2.7 Hydrogenase activity assays

Hydrogen evolution activity using reduced methyl viologen as electron donor

Hydrogenase activity was measured in two ways: hydrogen (H_2) evolution and hydrogen uptake. The H_2 evolution activity assay was carried out in a 2 mL GC vial containing 1.5 mL of 50 mM KH_2PO_4 (pH 7.0 for cell lysates/membrane extracts, and pH 5.5 for solubilized membrane proteins/purified MBH), 3 mM methyl viologen, and the hydrogenase sample (50-200 µL), as described previously [37]. The headspace was flushed with N_2 for 10 min, and the reaction was catalyzed by the anaerobic addition of 20 mM sodium dithionite (prepared in a glove box under inert atmosphere). H_2 evolution was measured every 50-60 min over the course of ~3 h by withdrawing 100 µL from the headspace and analyzing it with a gas chromatograph (GC) calibrated with H_2, and the activity was measured as the slope of the time-dependent reaction. A 5890 Series III (HP) GC equipped with a thermal conductivity detector and Supelco Carboxen 1000 column was used for these measurements. One unit of activity was defined as 1 µmol H_2 evolved/min/mg protein.

Hydrogen uptake activity using methylene blue as electron acceptor

Hydrogenase activity was also measured by monitoring H_2-dependent methylene blue reduction spectrophotometrically using a modified version of the protocol of Schink and Schlegel [234]. This method was used qualitatively for the monitoring of MBH purifications, and quantitatively for characterization of MBH enzyme. In this assay, 50 mM KH_2PO_4 buffer at pH 7.0 was used as the reaction buffer for membrane fractions, whereas for soluble extracts, solubilized membrane proteins, and purified MBH samples 50 mM KH_2PO_4 at pH 5.5 was used. A gas-tight cuvette containing 2.9 mL of reaction buffer with 200 µM methylene blue was flushed with H_2 for 30 min at room temperature. The protein sample (50-200 µL) was then injected to start the reaction, and the reaction was followed by monitoring the absorbance at 570 nm at room temperature for 10-20 min using a spectrophotometer (Thermo Scientific BioMate 3S). Using the molar extinction coefficient of methylene blue (ε_{570} = 13.1 cm^2 μmol^{-1}) and the protein concentration of the sample, the H_2-dependent methylene blue reduction activity was then calculated based on the slope of the time-dependent reaction. One unit of activity was defined as 1 µmol H_2 oxidized/min/mg protein.

5.2.8 Sortase-mediated ligation of PSI and MBH

Fusions of PSI and MBH were generated via site-specific ligation of the C-terminus of the stromal psaC, D, or E subunits of PSI to the N-terminus of the small subunit hoxK of MBH using the sortase enzyme SrtA [219]. Reactions were carried out in 1X TBS pH 8.0, 60 mM $CaCl_2$ and 0.04% DDM, using a 10:1 excess of MBH:PSI and a 1:1 ratio of sortase to total reactants (e.g. 10 µM PSI-LPETG, 100 µM GGG-MBH, and 110 µM sortase). Reactions were incubated at 38°C for 4 h, and then the reaction was stopped by the addition of EDTA to a final concentration of 10 mM to inactivate the sortase enzyme.

5.2.9 Hydrogen evolution activity using reduced DCIP as electron donor

The activity of the PSI-MBH protein fusions was assayed via light-induced H_2 evolution. Reactions were carried out in a 2 mL GC vial containing 1.5 mL of reaction buffer (50 mM KH_2PO_4 pH 5.5) with 0.04% (w/v) DDM, 100 mM sodium

ascorbate (NaAsc), 1.7 mM dichlorophenolindophenol (DCIP), 30 µM methyl viologen (MV), and the sortase reaction mix (100-250 µL). The reaction buffer and the NaAsc, DCIP, and MV stocks were flushed with N_2 for 10 min. NaAsc, DCIP, and MV were then anaerobically injected into the reaction vial, and the protein sample was then injected. Samples were illuminated under 240 µEinsteins/m^2/sec of white light for ~3 hours at room temperature. 100 µL was then withdrawn from the headspace and analyzed on a GC to quantify the amount of H_2 produced.

5.2.10 MD models of PSI-hydrogenase protein fusion complexes

Monomer-monomer PSI-hydrogenase fusions were built based on the known crystal structures of PSI from *T. elongatus* [9] (PDB ID: 1JB0), MBH from *R. eutropha* [34] (PDB ID: 3RGW), and the [FeFe]-hydrogenase (FeFe H_2ase) from *C. pasteurianum* [235] (PDB ID: 1FEH). As was the case in the previous chapter, the 91 out of ~2,300 residues missing from the PSI crystal structure, including the flexible N-termini of psaF and psaK, were not included in these simulations. Additionally, there are 70 out of 942 residues not resolved in the MBH crystal structure that were also not included in the simulations, including the final 65 residues of the C-terminus of the small subunit hoxK. These terminally located, intrinsically disordered regions are expected to play a role in PSI-MBH interactions, but are unlikely to significantly affect the global dynamics of the PSI-MBH fusion complex during simulation. Phylloquinone molecules located in the reaction center core of PSI were also not included in the simulations.

As PSI is a membrane-integral protein, a monolayer belt of DDM detergent was built around the transmembrane periphery of the PSI monomer in a similar manner to that described in the previous chapter (Figure 36A). For the PSI-MBH fusion model, a predicted structure for the LPETGGG linker located on the N-terminus of hoxK from MBH was generated using MODELLER [236], and PSI and MBH were ligated via a user script in VMD [175] (Figure 36B). For the PSI-FeFe H_2ase fusion models, PSI and FeFe H_2ase were ligated via either an octanedithiol (Figure 36C) or decanedithiol molecular wire with a user script in

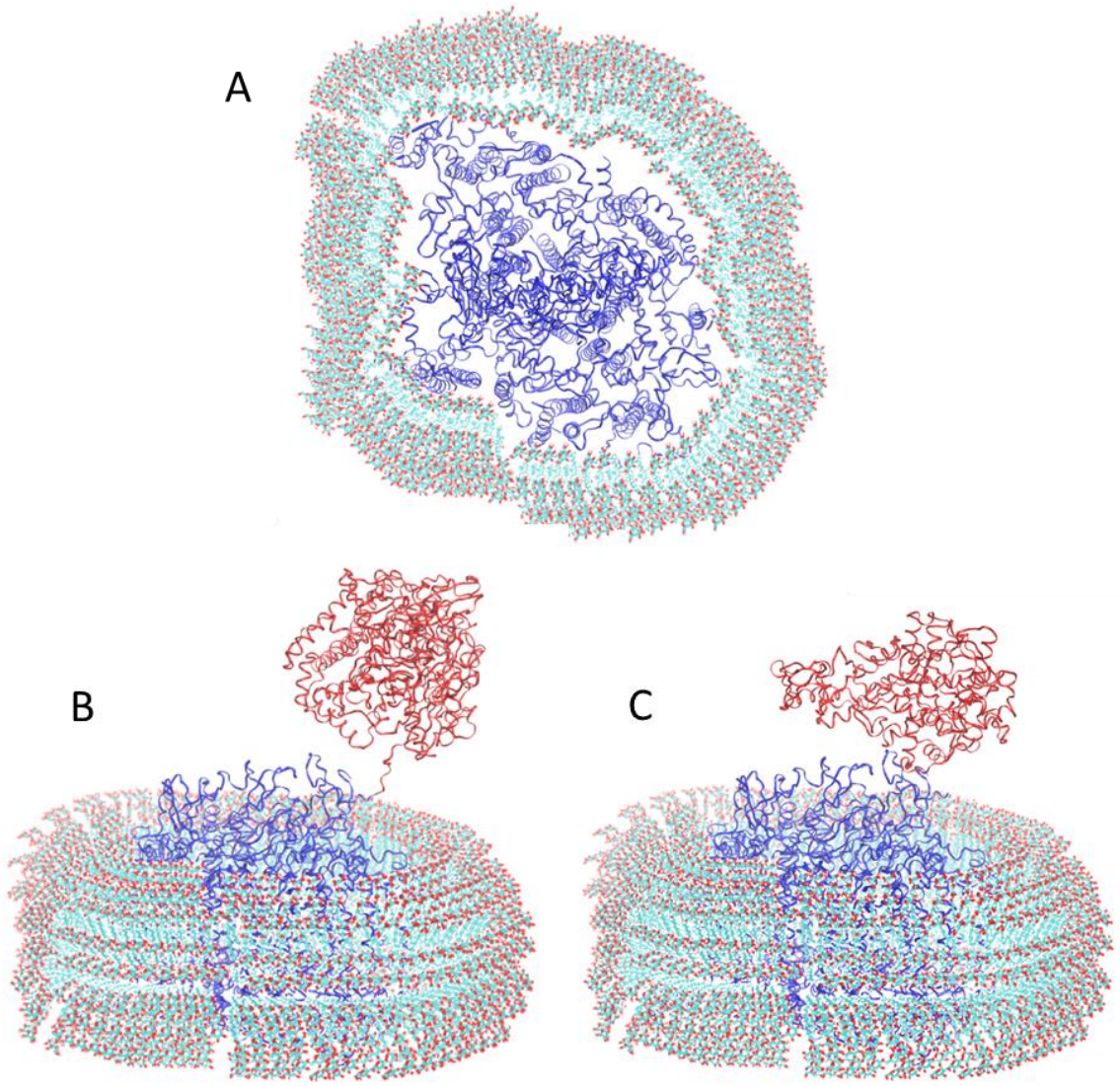

Figure 36. (A) Top view from the stromal side of monomeric PSI surrounded by a belt of DDM detergent molecules. (B) Side view of the PSI-MBH fusion complex, with PSI in blue ribbons and MBH in red ribbons. (C) Side view of the PSI-FeFe H_2ase fusion complex, with PSI in blue ribbons and FeFe H_2ase in red ribbons. In all cases, detergent molecules are shown in lines format, and water and counterions are omitted for clarity.

VMD. This was done according to the experimental method of Lubner et al., in which cysteines coordinating the F_B cluster of PSI (residue C13 from psaC) and the distal [4Fe4S] cluster of the FeFe H_2ase (residue C97) were mutated to glycines, and subsequently chemically rescued via a dithiol molecular wire to form a fusion complex [16]. All systems were subjected to 50 ps of MD

equilibration *in vacuo*, and subsequently solvated with TIP3P water molecules. Sodium and chloride counterions were then added by random replacement of water molecules in order to neutralize each system. The solvated, neutralized systems were subjected to an additional 50 ps of MD equilibration, and then MD production runs were carried out.

5.2.11 MD simulation details

MD simulation details were as described in the previous chapter, with one important addition. Parameters for the active site and iron-sulfur clusters of the FeFe H_2ase were based on the work of Chang and Kim [237].

5.3 Results and discussion

5.3.1 Targeted mutagenesis of hoxK

The complete nucleotide sequence of the *R. eutropha* megaplasmid pHG1 has been determined, and contains all the necessary structural, accessory, and regulatory genes for the synthesis of active MBH [222]. This enzyme is accessible to genetic engineering, and its biosynthesis and biochemical properties have been extensively studied [35, 229, 232, 238, 239].

The megaplasmid DNA (mpDNA) of *R. eutropha* HF387H was isolated, and an ~3 kbp fragment containing the hoxKG genes was amplified from the mpDNA via PCR and subsequently T/A cloned into pGEM-T Easy (Promega) to form pHoxKG (Figure 37A). pGEM-T Easy contains both an ampicillin-resistance cassette (ampR) and the lacZ promoter, allowing for easy selection of successful inserts via blue-white screening on ampicillin-containing media. The forward primer used, HoxKG_Forward, contains the GAGCTC sequence (shown in red in Table 10), thus allowing for future transfer of the DNA fragment via a SacI restriction enzyme digest. The Gly_3 tag was then introduced to the N-terminus of hoxK as follows. Primers HoxKG_Forward and TriGly_Lower were used to amplify the Gly_{UP} fragment from pHoxKG via PCR. TriGly_Lower is a reverse primer targeting the N-terminus of hoxK but also containing the additional nine codons for the Gly_3 tag (shown in green in Table 10). Separately, primers

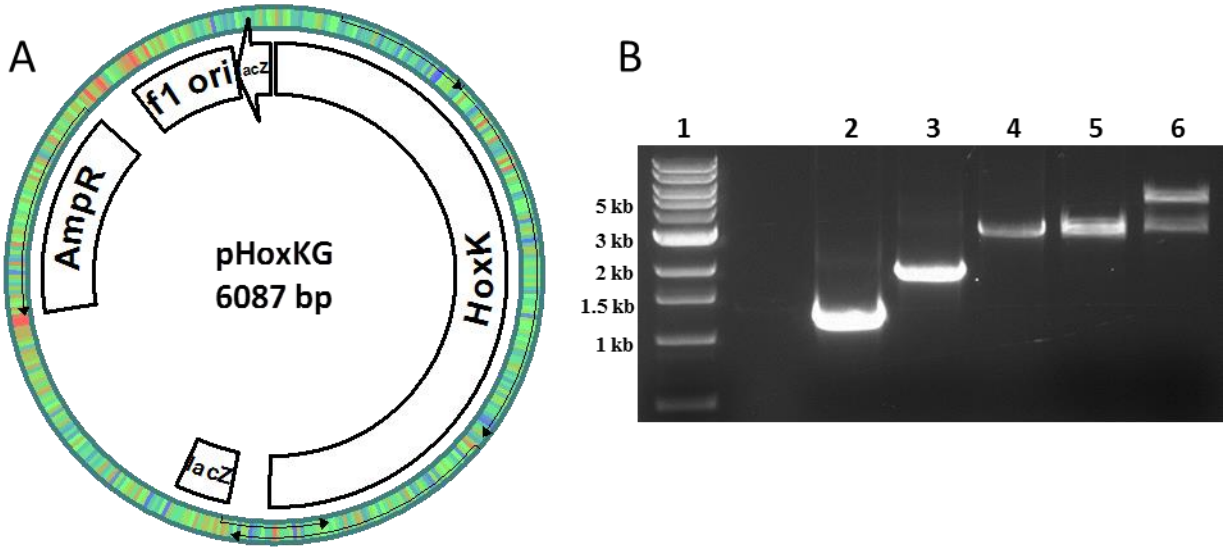

Figure 37. (A) pHoxKG plasmid resulting from T/A cloning of hoxKG PCR fragment into pGEM-T Easy. (B) Ethidium bromide-stained agarose gel electrophoresis results. Lanes are as follows: (1) MW marker, (2) GlyUP PCR fragment, (3) GlyDOWN PCR fragment, (4) Gly PCR fragment, (5) SacI-digested pHoxKGmod, and (6) SacI-digested pLO3-HoxKGmod.

TriGly_Upper and HoxKG_Reverse were used to amplify the Gly_{DOWN} fragment from pHoxKG via PCR. TriGly_Upper contains a complementary sequence to that of TriGly_Lower, including the nine codons for the Gly_3 tag (shown in green in Table 10). Then, using primers HoxKG_Forward and HoxKG_Reverse and both the Gly_{UP} and Gly_{DOWN} PCR fragments as template, the Gly PCR fragment was amplified by PCR to produce the modified hoxK gene. The Gly PCR fragment was then T/A cloned into pGEM-T Easy to produce pHoxKGmod. A schematic of this overlapping PCR approach is shown in Figure 34. Ethidium bromide-stained agarose gel electrophoresis results are shown for various stages of this process in Figure 37B. The expected sizes of 1 kbp for Gly_{UP} (Figure 37B, lane 2) and 2 kbp for Gly_{DOWN} (Figure 37B, lane 3) were obtained after PCR amplification of pHoxKG using the aforementioned primers. Subsequent PCR amplification of these fragments using primers HoxKG_Forward and HoxKG_Reverse yields the 3 kbp Gly fragment (Figure 37B, lane 4). A SacI double digest of pHoxKGmod, the plasmid resulting from T/A cloning of the Gly PCR fragment into pGEM-T Easy, results in a double band

at ~3 kbp corresponding to the sizes of the original vector and the Gly PCR fragment. For a complete explanation of mpDNA isolation, PCR reactions and cloning, see Materials and methods.

5.3.2 Conjugal gene transfer in R. eutropha

Bacterial conjugation is the transfer of genetic material between bacterial cells, either by cell-cell contact or bridge-like connections. As with transformation and transduction, conjugation is a mechanism of horizontal gene transfer. During conjugation, the donor strain provides a conjugative (mobilizable) genetic element, usually a plasmid or transposon. In this case, the mobilizable (suicide) vector pLO3 was used to facilitate the gene exchange procedure. pLO3 is a ColE1 replicon outfitted with the RP4 transfer origin (oriT) and a conditionally lethal *Bacillus subtilis* sacB gene in addition to a tetracycline resistance gene [229]. This vector is competent for mobilization in *R. eutropha* but incapable of autonomous replication in this bacterium [239]. Suicide vectors are used for allelic exchange of a non-selectable marker. The Tet^R phenotype of the plasmid provides a direct selection for integration of the plasmid into the chromosome. Expression of the *sacB* gene is toxic for gram-negative bacteria when grown in the presence of 5% sucrose, providing a direct selection for loss of the plasmid [240]. *E. coli* S17-1 was used as the mobilizable (donor) strain, as it can carry the transfer genes of the broad host range IncP-type plasmid with chromosomally integrated RP4 derivative to promote mobilization, and can utilize any gram-negative bacterium as a recipient for conjugative DNA transfer [227, 228].

In order to introduce the Gly_3 tag mutation back into R. eutropha HF387H via conjugation, the Sac I-digested Gly fragment from pHoxKGmod was sub-cloned into pLO3 to form pLO3-HoxKGmod (Figure 38A). This plasmid was then transformed into *E. coli* S17-1 for conjugal transfer, and the modified hoxK sequence was introduced into R. *eutropha* via allelic exchange. Conjugation was initiated by spot-mating donor and recipient strains. Heterogenote (single crossover) recombinants were selected by screening for tetracycline resistance. Tet^R segregants were picked and purified by streaking. In a subsequent step, homogenote (double crossover) recombinants which have lost the vector sequence, including the *sacB* gene, were selected as sucrose-resistant survivors. A schematic of this process is shown in Figure 35B. The megaplasmid DNA of the resulting mutant strain (designated NGLY3 in Table 9) was isolated,

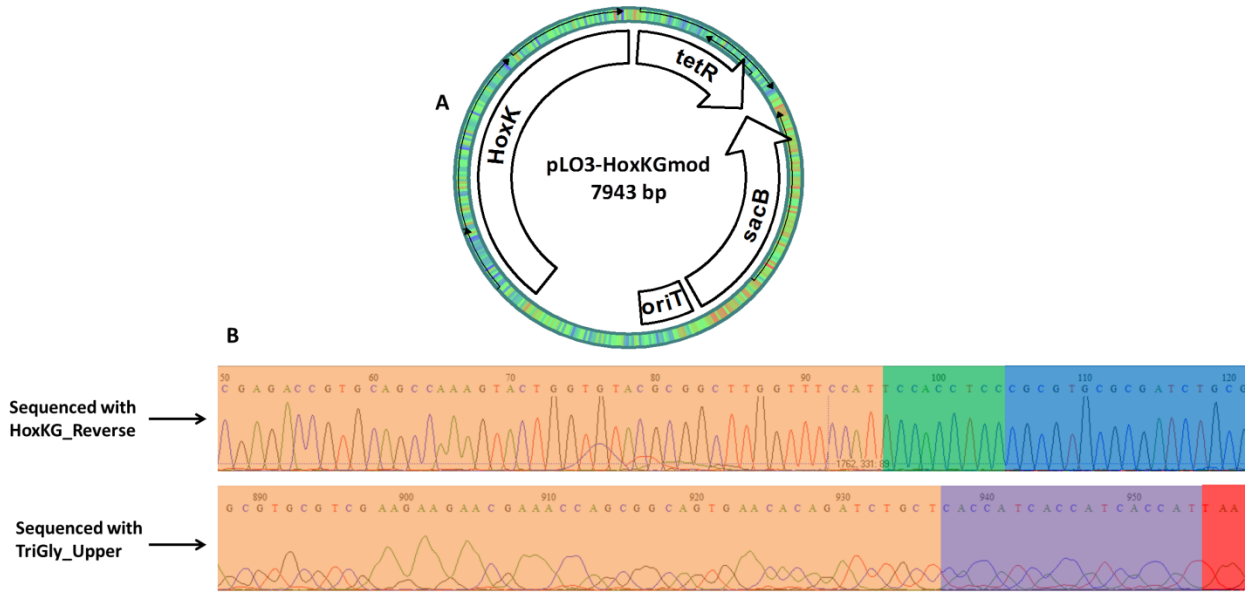

Figure 38. (A) pLO3-HoxKGmod plasmid resulting from ligation of SacI-digested Gly_3 fragment from pHoxKGmod into the suicide vector pLO3. (B) Electeropherograms showing sequence verification of the N-terminally Gly_3 tagged hoxK gene *R. eutropha* mutant. The upper electropherogram was obtained via DNA sequencing using the HoxKG_Reverse primer (see Table 10), with the orange region corresponding to the hoxK gene, the green region representing the N-terminal Gly_3 tag, and the blue region corresponding to the flanking sequence. The lower electropherogram was obtained via sequencing with the TriGly_Upper primer, with the orange region again representing the hoxK gene, the purple region the His_6 purification tag, and the red region corresponding to the hoxK stop codon.

and the presence of the Gly_3 tag on the N-terminus of hoxK, as well as the continued presence of the His_6 purification tag on its C-terminus, was verified by DNA sequencing (Figure 38B). For the complete bacterial conjugation protocol, please refer to Materials and methods.

5.3.3 Purification of the wild-type and mutant MBH

R. eutropha strains were grown micro-aerobically using the method of Lenz et al. [232]. This method involves growing cell cultures in flasks filled four-fifths full with a modified FGN minimal medium containing 0.05% (w/v) fructose and 0.4% (w/v) glycerol, which has been shown to increase the level of hydrogenase expression. Solubilized membranes containing wild-type MBH from *R. eutropha* H16 were isolated as described previously (Figures 39A,B) [233]. SDS-PAGE gel electrophoresis and subsequent Western blot analysis using an antibody against

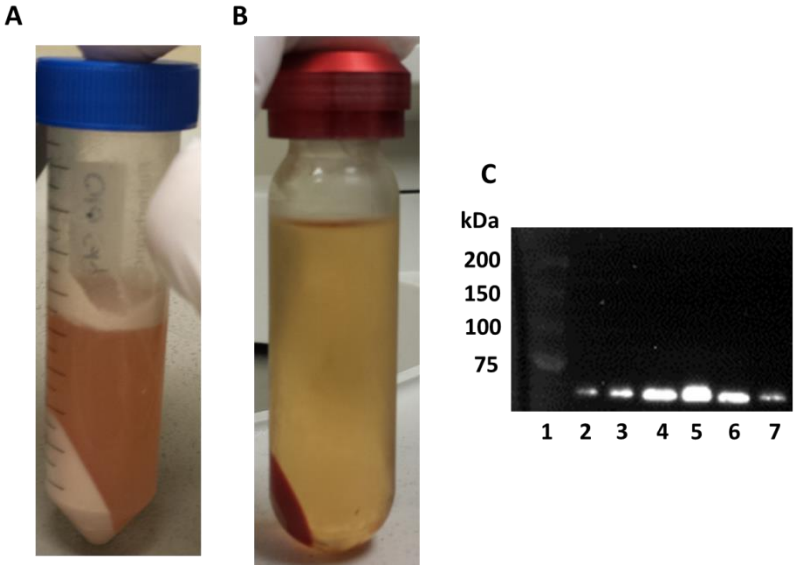

Figure 39. Subcellular localization of MBH in WT *R. eutropha* H16. (A) Soluble and pelleted fractions after low-speed centrifugation of lysed cells at 4,000 x g for 20 min at 4°C, (B) Soluble and membrane fractions after ultracentrifugation of low-speed centrifugation supernatant at ~130,000 x g for 45 min at 4°C. (B) Western blot analysis of WT MBH purification steps using anti-hoxG antibody. Lanes are as follows: (1) MW marker, (2) Low-speed centrifugation (4,000 x g for 20 min at 4°C) supernatant, (3) Low-speed centrifugation pellet, (4) Ultracentrifugation (130,000 x g for 45 min at 4°C) supernatant, (5) Ultracentrifugation pellet (a.k.a. membrane fraction), (6) Ultracentrifugation supernatant after detergent solubilization (a.k.a. solubilized membrane extract), and (7) Ultracentrifugation pellet after detergent solubilization.

hoxG (67 kDa) revealed the presence of MBH throughout the purification process, including in the membrane fraction (Figure 39B, lane 5) as well as in the solubilized membrane extract (Figure 39B, lane 6).

The MBH of *R. eutropha* consists of a large subunit, hoxG, that houses the active site, and a small subunit hoxK, which contains three iron-sulfur clusters. The MBH is tethered to the periplasmic side of the membrane via a C-terminal "anchor" region of hoxK that is connected to the membrane-integral *b*-type cytochrome hoxZ (Figure 40) [238]. The hoxZ gene product is not required for MBH translocation, but is essential for H_2-dependent electron transfer *in vivo* [241]. It has been shown that recombinant strains lacking the C-terminal anchor region of hoxK do not sustain H_2-dependent autotrophic growth due to the fact that the mutant MBH is incapable of establishing a proper connection to hoxZ [15]. Due to the fact that the C-terminal anchor region is still present in the mutated hoxK gene of the *R. eutropha* NGLY strain (Table 9), we expected to find both the hoxG and hoxK subunits in the membrane fraction and detergent-solubilized membrane extract of this strain, as was the case for WT *R. eutropha* H16. However, Coomassie-stained SDS-PAGE gel electrophoresis results for several steps throughout the purification process (Figure 41A) indicated that the

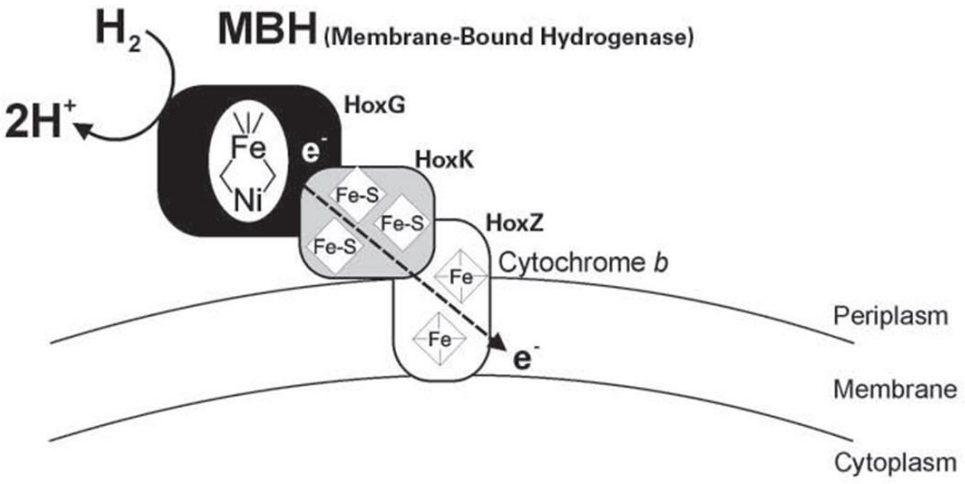

Figure 40. Schematic of membrane-bound hydrogenase of *R. eutropha* H16 depicting organization of individual subunits (Figure adapted from Burgdorf et al.).

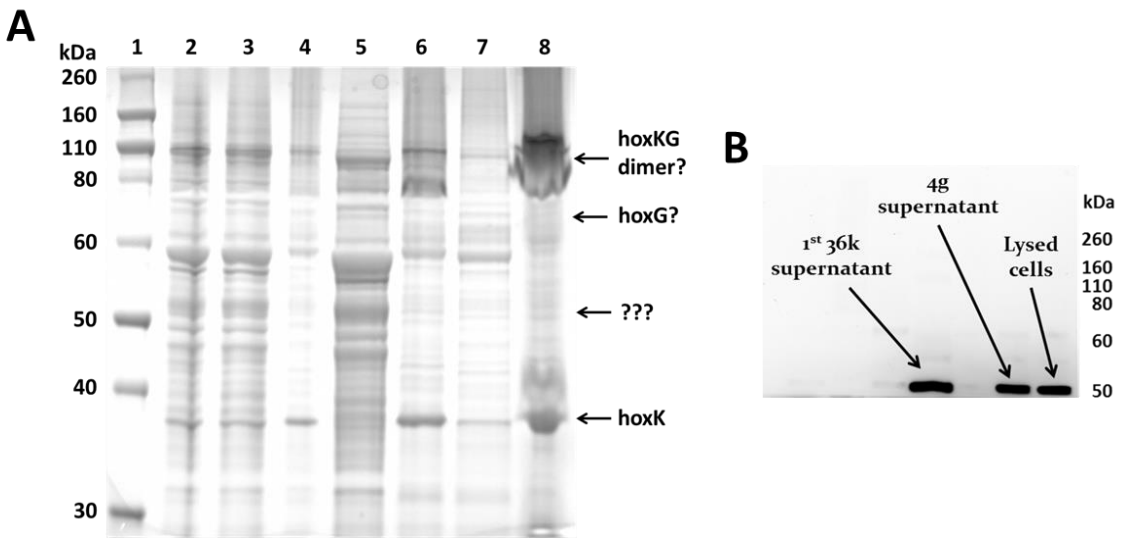

Figure 41. (A) Coomassie-stained SDS-PAGE gel electrophoresis results for the purification of MBH from the mutant *R. eutropha* strain NGLY. Lanes are as follows: (1) MW marker, (2) Lysed cells, (3) Low-speed centrifugation supernatant, (4) Low-speed centrifugation pellet, (5) Ultracentrifugation supernatant, (6) Ultracentrifugation pellet, (7) Ultracentrifugation supernatant after detergent solubilization, and (8) Ultracentrifugation pellet after detergent solubilization. (B) Western blot analysis of purification of MBH from the mutant *R. eutropha* strain HF387H using an anti-hoxG antibody.

MBH large subunit hoxG (67 kDa) was not present in the membrane fraction of the mutant strain as expected. There were also bands throughout the purification process at ~52 and ~110 kDa that could not be explained, the latter of which might be the non-denatured hoxKG heterodimer. The explanation may lie in the maturation and membrane translocation of the MBH enzyme.

The MBH is transported to the periplasmic side of the membrane via a specific protein translocation pathway known as the membrane targeting and translocation (Mtt) [231] or twin-arginine translocation (Tat) pathway [242]. The cytoplasmic, premature hoxK subunit (designated prehoxK) contains a long signaling peptide (43 amino acids) on its N-terminus with a conserved (S/T)RRxFxK motif that serves as signal recognition to target the fully-folded heterodimer to the membrane and the periplasm [25]. It has been demonstrated by Bernhard et al. through experimental mutagenesis studies that the immature MBH small subunit prehoxK co-purifies with the accessory proteins hoxO and

hoxQ, a protein complex that was also isolated from a strain in which the large subunit hoxG had been deleted [243]. These results suggest that hoxO and hoxQ serve to stabilize the hoxK precursor prior to oligomerization with hoxG, possibly serving as chaperones which control MBH assembly and prevent premature MBH from being tagged to the membrane and prematurely delivered to the Tat apparatus [238]. After transport and insertion into the membrane, the signaling peptide is removed by specific processing peptidases. Experimental mutagenesis studies of the processing peptidase cleavage sites of several thylakoid membrane Tat substrates conducted by Frielingsdorf and Klösgen have shown that both the C-terminal segment of the signaling peptide and the N-terminal part of the mature protein play an important role in the maturation process [244]. Efficient cleavage of the signal peptide requires the presence of charged or polar residues in at least one of these regions, while increased hydrophobicity in either segment impairs the process. However, Frielingsdorf and Klösgen also found membrane complexes of high molecular mass, presumably representing Tat complexes, which contained fully translocated passenger proteins that had not yet been terminally processed by peptidases. These results suggest that membrane transport and terminal processing of Tat substrates are independent processes. Therefore, although it is entirely plausible that the introduction of the non-polar Gly_3 tag to the N-terminus of hoxK may interfere with efficient cleavage of the signaling peptide, this does not mean that the mutation will prevent translocation of the fully-folded, active heterodimer to the membrane and periplasm. Furthermore, the presence of an intense band in the size range of hoxK (35 kDa), rather than that of prehoxK (40 kDa), in the cell debris and membrane fractions of partially purified MBH from the mutant *R. eutropha* NGLY strain (Figure 41) suggests that the presence of the Gly_3 tag does not interfere with cleavage of the signaling peptide at all.

Due to the fact that similar issues were encountered when isolating membranes from the mutant strain expressing a C-terminal His_6 purification tag on hoxK (HF387H in Table 9) from which the NGLY strain was derived, the

absence of the MBH may be due to problems associated with this original mutant strain. Interference with membrane translocation due to mutations associated with the C-terminal anchor region of hoxK has been seen in previous work on self-assembling PSI-MBH protein fusions conducted by Schwarze et al. [15]. Generating MBH-psaE hybrid proteins by replacing the C-terminal membrane anchor of the *R. eutropha* MBH small subunit hoxK with the psaE subunit of PSI from *Synechocystis*, they found that the mature hoxK-psaE protein was not detectable in the membrane or the periplasm. Instead, they found that prehoxK-psaE was present in the cytoplasm and in considerable amounts in the membrane, and therefore concluded that modification of the C-terminus of hoxK strongly affected the Mtt/Tat-dependent transport of the hybrid protein.

The mutant *R. eutropha* HF387H strain was originally constructed in our lab by Ify Iwuchukwu in 2011 [224]. The absence of hoxG from the membrane fraction could be the result of unnoticed mutations introduced during the genetic engineering of hoxK to express the His_6 purification tag, or during subsequent homologous recombination to transfer the mutation to the host *R. eutropha* HF387 strain. Additionally, the ~52 kDa band observed in Figure 41 was also identified in Western blots against hoxG for purification preps of both the HF387H (Figure 41B) and NGLY strains, indicating a point mutation may be present in this gene, as it should run much higher (~67 kDa). Therefore, it is recommended that future work troubleshooting this problem begin with sequencing the hoxG gene.

5.3.4 H_2 uptake activity of the WT and mutant MBH

The H_2 uptake activities of the wild-type and mutant *R. eutropha* strains were determined using methylene blue as the electron acceptor (Figure 42A). For qualitative determination of the presence of MBH activity, GC vials containing 1 mL of reaction buffer (see Materials and methods) were flushed with H_2 and monitored for color change after injection of the protein sample for various steps throughout the purification process. For the case of WT *R. eutropha* H16 (Figure 42B), as well as for the soluble hydrogenase-free mutant *R. eutropha* HF387, a rapid color change (<5 min) from blue to colorless was observed for the low-

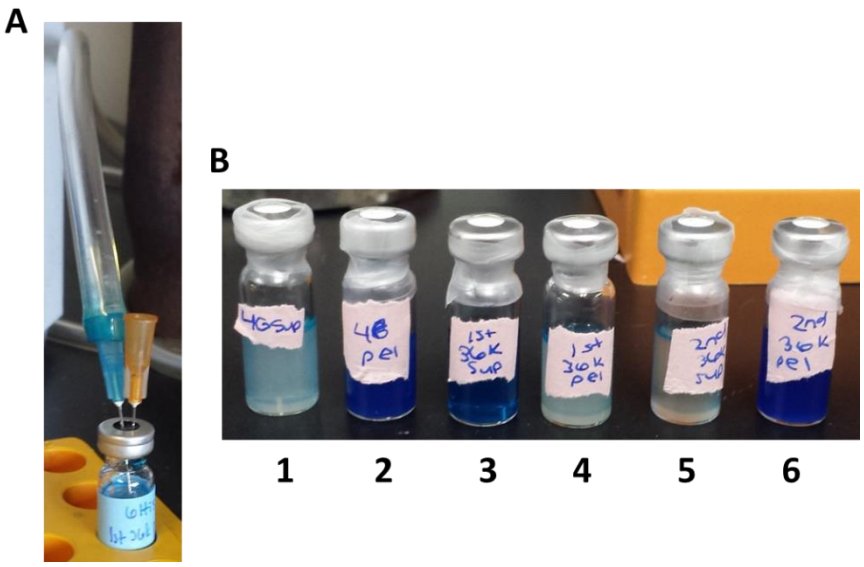

Figure 42. (A) Crimped GC vial containing reaction buffer with methylene blue for H_2 uptake activity assay being flushed with H_2. (B) Qualitative activity results for partial purification of MBH from WT *R. eutropha* H16. Samples are as follows: (1) Low-speed centrifugation supernatant, (2) Low-speed centrifugation pellet, (3) Ultracentrifugation supernatant, (4) Ultracentrifugation pellet, (5) Ultracentrifugation supernatant after detergent solubilization, and (6) Ultracentrifugation pellet after detergent solubilization.

speed centrifugation supernatant after cell lysis, as well as for the membrane fraction and solubilized membrane extract. This is expected, as these fractions should contain the fully-folded, active MBH heterodimer. However, there was no observable activity in any fraction tested for both the HF387H and NGLY mutant *R. eutropha* strains. This serves as further evidence that introduction of the His_6 purification tag to the C-terminus of hoxK, or some overlooked undesirable mutation introduced to hoxKG during construction of this mutant, has interfered with the proper assembly and/or membrane translocation of the active MBH heterodimer. For specific activity, the change in absorbance at 570 nm (A_{570}) over time was measured via a spectrophotometer (Thermo Scientific BioMate 3S). Using the molar extinction coefficient of methylene blue (ε_{570} = 13.1 cm^2 μmol^{-1}) and the protein concentration of the sample (determined using the BCA method), the H_2 uptake activity was calculated based on the slope of the time-dependent reaction. A sample kinetic absorbance curve for partially purified MBH

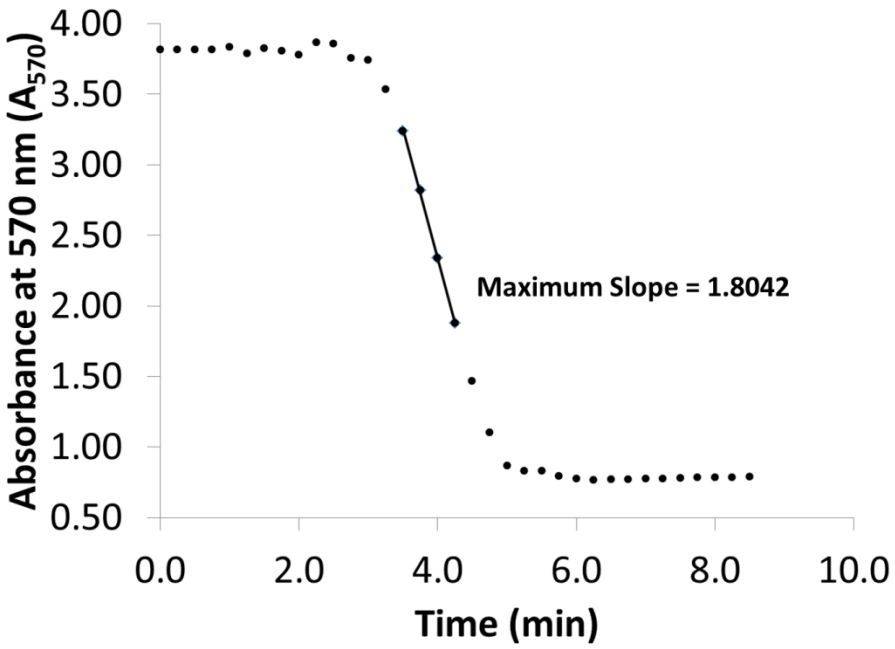

Figure 43. Absorbance versus time for H_2-dependent reduction of methylene blue via partially purified MBH from WT *R. eutropha* H16. Experimental data is shown in black dots, with the black trendline representing the four points chosen for determination of the maximum slope.

from WT *R. eutropha* H16 is shown in Figure 43. Using the maximal slope obtained from this curve and the total protein concentration as determined using the BCA assay (Pierce), we obtain an activity of 0.24 µmol H_2/min/mg protein. This is lower than the expected value of 1-3 µmol H_2/min/mg protein predicted by Dr. Oliver Lenz (personal communication), an expert in hydrogenase research. However, we obtain a value of 2.37 µmol H_2/min/mg protein for the soluble hydrogenase-free *R. eutropha* HF387 mutant strain, which is in agreement with the value of 1.7 µmol H_2/min/mg protein obtained by Lenz et al. [239], indicating that we are able to successfully isolate membranes containing active MBH.

5.3.5 MD modeling of PSI-FeFe H_2ase and PSI-MBH protein fusions

We elected to conduct MD simulations on three distinct monomer-monomer PSI-hydrogenase fusion complexes for in-depth analysis of complex stability and protein-protein interactions: (1) PSI from *T. elongatus* connected to the FeFe H_2ase from *C. pasteurianum* via a 1,8-octanedithiol (OD) molecular wire

(designated the PSI-FeFe H_2ase OD model), (2) PSI from *T. elongatus* connected to the FeFe H_2ase from *C. pasteurianum* via a 1,10-decanedithiol (DD) molecular wire (designated the PSI-FeFe H_2ase DD model), and (3) PSI from *T. elongatus* ligated to the MBH from *R. eutropha* via a sortase linkage (LPETGGG) between the C-terminus of psaE from PSI and the N-terminus of hoxK from the MBH (designated the PSI-MBH model). Constructs corresponding to the PSI-FeFe H_2ase OD and DD models have been assayed experimentally for light-induced H_2 evolution activity by Golbeck and co-workers [16], and the PSI-MBH model serves as a direct comparison to our experimental fusion complex. Based on our previous experience with MD simulations of trimeric PSI in detergent solution, we chose to use a pre-formed structure of PSI monomer embedded in a toroidal belt of DDM detergent. A monolayer belt of 666 DDM detergent molecules was built around the periphery of monomeric PSI consisting of semicircular planes of DDM densely packed around the hydrophobic periphery of the protein. A higher detergent-to-protein ratio relative to the trimeric PSI/DDM model was necessary to adequately cover exposed hydrophobic regions that had previously been shielded by neighboring PSI monomers. Structures for the dithiol linkers were freely available via the Protein Data Bank, and an initial structure for the flexible LPETGGG linker of the PSI-MBH fusion model was predicted using MODELLER [236]. Starting structures for the PSI-FeFe H_2ase and PSI-MBH models are shown in Figure 36, and Table 11 contains simulation details for each system. Each system was energy-minimized, solvated, and equilibrated (see Materials and methods), and subsequently subjected to an MD production run.

Table 11. Simulation details for PSI-hydrogenase fusion models

System	Components	No. of water molecules	No. of ions	No. of atoms	Box size (nm)	Simulation time (ns)
PSI-MBH model	PSI, MBH, DDM	214,282	96 Na^+, 201 Cl^-	760,805	20.5 x 20.5 x 20.0	250
PSI-FeFe H_2ase OD model	PSI, H_2ase, OD, DDM	198,076	96 Na^+, 186 Cl^-	707,532	20.5 x 20.5 x 18.5	115
PSI-FeFe H_2ase DD model	PSI, H_2ase, DD, DDM	198,074	96 Na^+, 186 Cl^-	707,532	20.5 x 20.5 x 18.5	170

Stability and local mobility of the PSI-FeFe H_2ase fusion complexes

The PSI-FeFe H_2ase OD model was simulated for ~115 ns. Analysis of the root-mean-squared deviation (RMSD) of the C_α atoms of both the PSI monomer and the FeFe H_2ase indicates that these proteins have reached a stable complex (Figure 44). Similar to what was seen previously in MD simulations of trimeric PSI in DDM [41], the C_α RMSD of all residues of PSI reaches a maximum value of ~2.5 Å (Figure 44A, black solid line). Backbone RMSDs of ~2.0 Å were previously reported in an MD study of the channel protein FhuA in OES detergent [151], but values in this range are typically associated with MD studies of membrane proteins in lipid bilayers [139, 203]. The FeFe H_2ase exhibits a final C_α RMSD of ~4.0 Å (Figure 44B). This value is much higher than the ~0.8 Å RMSD previously reported by Chang and Kim for MD simulations of this FeFe H_2ase [237]. The high structural drift of the FeFe H_2ase relative to the crystal structure, exceeding that which is expected due to aqueous solution conditions, indicates that this protein is undergoing substantial structural changes in this complex. This is not unexpected, as the octanedithiol (OD) linker closely tethers this protein to PSI, the implications of which will be discussed in detail below.

We again decomposed the C_α RMSDs of PSI into four domains: the reaction center subunits psaAB, the terminal electron acceptor psaC, the ferredoxin-docking stromal subunits psaDE, and the peripheral transmembrane α-helices

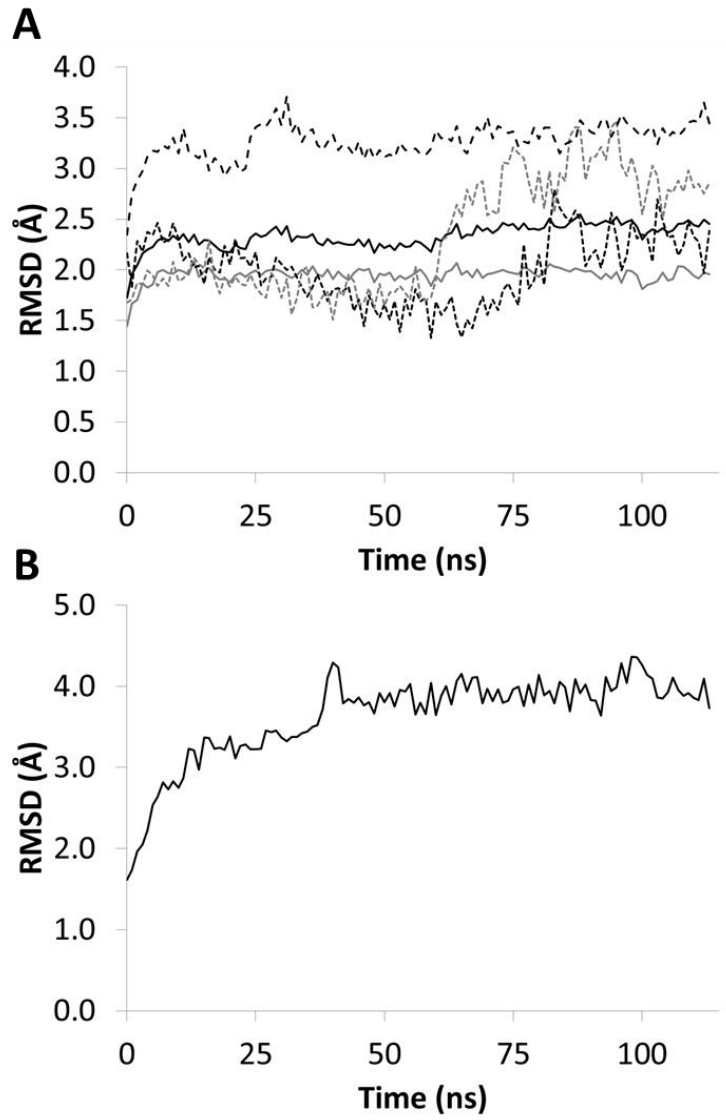

Figure 44. C_α RMSD versus time for the PSI-FeFe H_2ase OD model MD simulation, calculated for (A) PSI monomer relative to the crystal structure (PDB ID: 1JB0). The lines shown the C_α RMSDs for: all residues (black solid line), reaction center subunits psaAB (gray solid line), terminal electron acceptor psaC (black dotted line), ferredoxin docking subunits psaDE (gray dotted line), and the peripheral transmembrane helices psaFIJKLMX (black dashed line). Also C_α RMSDs for (B) all residues of the FeFe H_2ase relative to the crystal structure (PDB ID: 3RGW).

psaFIJKLMX. As seen before, the reaction center subunits psaA and psaB (Figure 44A, gray solid line) display the lowest structural drift with a final C_α RMSD of ~2.0 Å. The slight rise seen here compared to the 1.8 Å C_α RMSD value seen for these subunits in our MD simulations of trimeric PSI in DDM could

be due to the exposure of a larger portion of these residues to the solubilizing detergent, which were previously shielded by adjacent monomers of the PSI trimer. As seen previously, the peripheral transmembrane helices exhibit the highest structural drift of ~4.0 Å (Figure 44A, black dashed line). The stromal subunits show remarkably different behavior from previous work, with psaC (Figure 44A, black dotted line) reaching a maximum RMSD of ~2.5 Å and the psaDE subunits (Figure 44A, gray dotted line) rising as high as 3.5 Å, both of which are much higher than what was seen for trimeric PSI in DDM [41]. Again, this is likely due to the fact that the stromal surface of PSI is in close contact with the FeFe H_2ase throughout the simulation.

We calculated the time-averaged C_α RMSF values for each residue of PSI and the FeFe H_2ase for the final ~60 ns of the OD model MD simulation to analyze the local atomic mobility of the protein fusion (Figure 45). For PSI (Figure 45A), the trend is much the same as that seen in our previous MD study of trimeric PSI in detergent solution, with regions of low atomic mobility primarily confined to the core residues of the reaction center and those of high mobility corresponding to turns in the peripheral transmembrane helices and flexible loops of the stromal subunits [41]. However, higher C_α RMSFs than seen previously were obtained for detergent-exposed regions such as the peripheral α-helices, possibly due to the higher detergent-to-protein ratio. Most noticeably, the N-terminus of psaL does not extend out into the solvent and exhibit the high RMSFs seen previously. This may also be due to the higher detergent-to-protein ratio used in this study compared to the trimeric PSI in DDM simulations. The maximum C_α RMSF of ~4.3 Å corresponds to the C-terminus of psaD, possibly due to interactions with the hydrogenase. For the FeFe H_2ase, the observed C_α RMSFs for the MD simulation (Figure 45B, black solid line) are in the range of ~2-5 Å, which is much higher than the values of ~1 Å determined from the crystal structure (Figure 45B, black dotted line). This provides further evidence, along with the structural drift observed for the FeFe H_2ase and the stromal residues of PSI, that the tethering of this protein to PSI has resulted in significant structural changes.

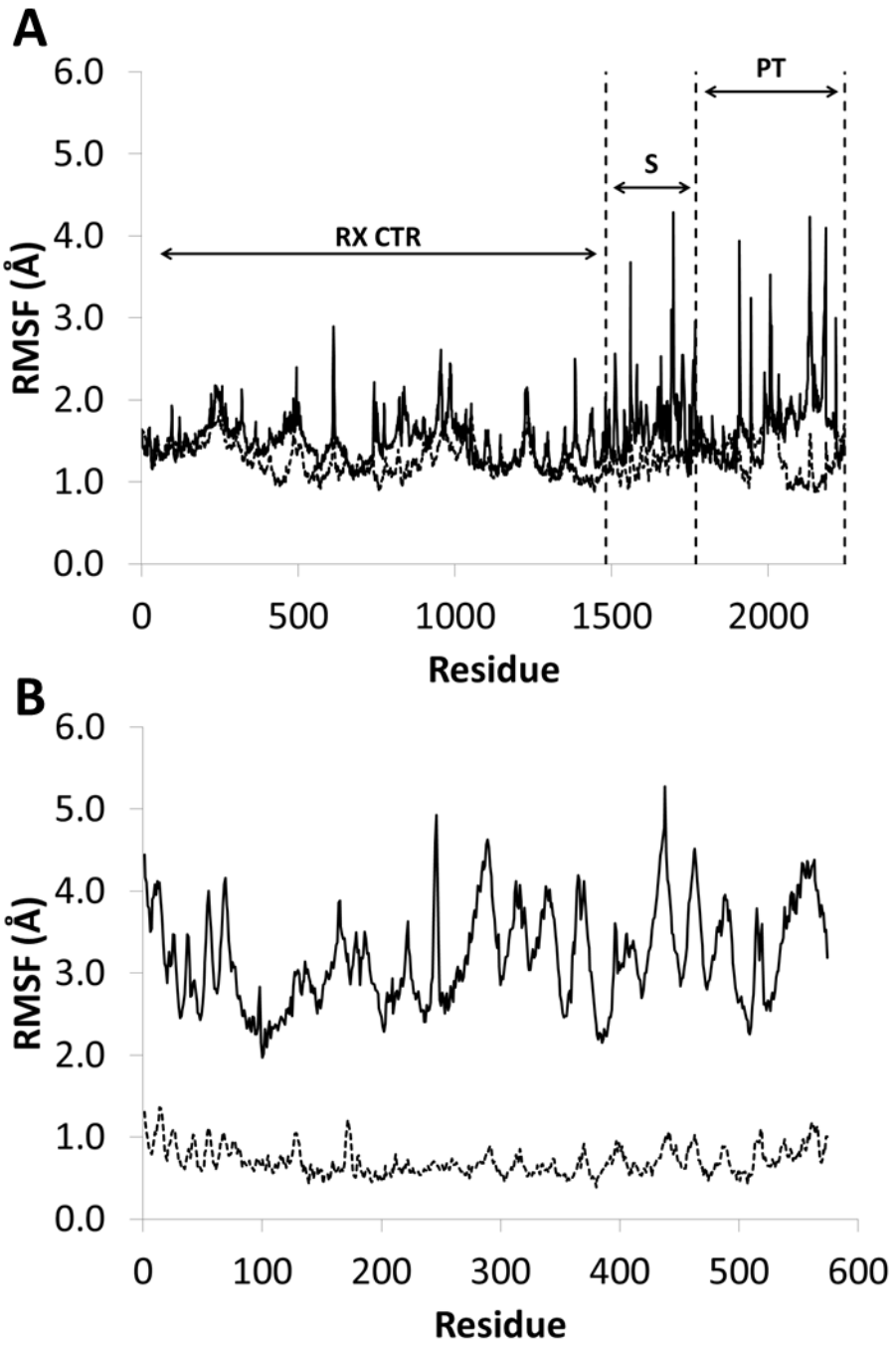

Figure 45. Time-averaged C_α RMSF versus residue number, calculated over the final ~60 ns of the FeFe H_2ase OD model MD simulation, shown for (A) the PSI monomer (solid black black line), with residues of the reaction center (RX CTR), stromal (S), and peripheral transmembrane (PT) domains labeled accordingly, and (B) the FeFe H_2ase (solid black line). In both cases, RMSF values calculated from the temperature factors of the X-ray crystal structures are included for comparison (black dotted line).

The PSI-FeFe H_2ase DD model was simulated for ~170 ns, at which point the complex had stabilized based on C_α RMSD analysis. Maximum C_α RMSD values of 2.6 Å and 6.5 Å were obtained for PSI and the FeFe H_2ase, respectively (Table 12). Comparing these results to that obtained for the PSI-FeFe H_2ase, we see that linker length has a significant effect on the structural drift of the hydrogenase relative to its crystal structure, resulting in an increase from 4.4 Å for the OD model to 6.5 Å for the DD model. This is likely due to the fact that the increase in linker length from octanedithiol to decanedithiol allows the FeFe H_2ase to explore a larger conformational space during the simulation, resulting in greater structural drift. This also means that longer simulation times will be required in order for the system to reach equilibrium, as the DD model had to be simulated for 170 ns in order for the C_α RMSD to stabilize, while the OD model only had to be simulated for 115 ns (Table 11). C_α RMSF results for the DD model were similar in overall trend to what was seen for the OD model, with higher values in the range of 2-7 Å for the FeFe H_2ase (Table 12), reflecting the greater structural changes occurring in the hydrogenase in this system.

Protein-protein interactions of PSI-FeFe H_2ase fusion complexes

Snapshots of the PSI-FeFe H_2ase OD model in its initial configuration and after ~30 ns of MD simulation are shown in Figures 46A and 46C, respectively. The DDM detergent belt has quickly conformed to the hydrophobic periphery

Table 12. Maximum C_α RMSDs for PSI-hydrogenase fusion MD simulations

System	Maximum PSI C_α RMSD (Å)	Maximum H_2ase C_α RMSD (Å)	Maximum H_2ase C_α RMSF (Å)
PSI-FeFe H_2ase OD model	2.6	4.4	5.3
PSI-FeFe H_2ase DD model	2.6	6.5	6.9
PSI-MBH model	2.8	4.6	8.1

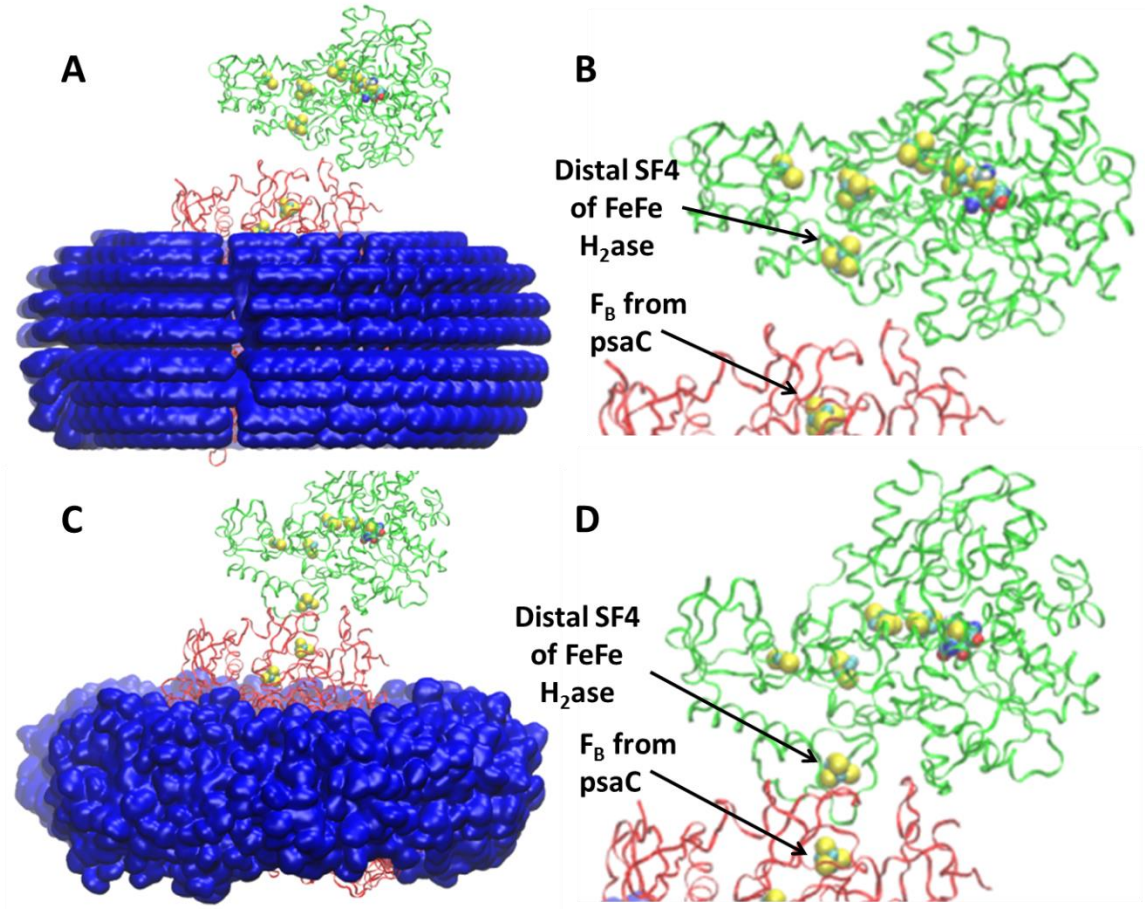

Figure 46. The initial configuration of the PSI-FeFe H_2ase OD model shown from (A) side view, and (B) zoomed-in view of the stromal surface of PSI and the FeFe H_2ase. The PSI-FeFe H_2ase OD model after ~30 ns of MD simulation shown from (C) side view, and (D) zoomed-in view of the stromal surface of PSI and the FeFe H_2ase. In all cases, PSI and the FeFe H_2ase are shown in red and green ribbons, respectively, the detergent is shown in low-resolution surface representation in blue, and the iron-sulfur clusters of psaC from PSI as well as the iron-sulfur clusters and active site of the FeFe H_2ase are shown in VDW format. In (B) and (D), the F_B cluster of psaC from PSI and the distal [4Fe4S] cluster of the FeFe H_2ase are labeled accordingly.

of the PSI monomer, as was the case in MD simulations of trimeric PSI in DDM; this behavior is also seen in the PSI-FeFe H_2ase DD and PSI-MBH simulations. Interestingly, the distance between the F_B cluster of psaC from PSI and the distal [4Fe4S] cluster of the FeFe H_2ase shrinks substantially after just ~30 ns of simulation (Figures 46B and 46D). This is due to the octanedithiol molecular wire connecting these two species (not shown), and enables the high light-driven H_2

evolution rates seen experimentally for this fusion complex [16]. However, the shrinking of this distance is accompanied by significant structural drift of the associated proteins. In particular, the majority of the FeFe H_2ase pulls away from those residues surrounding its distal iron-sulfur cluster over the course of the simulation. Furthermore, the main body of the hydrogenase moves away from the stromal surface of PSI, possibly due to unfavorable electrostatic interactions. This combined behavior is particularly obvious in the case of the PSI-FeFe H_2ase DD model MD simulation, as shown in Figure 47, and implies that there may be significant strain in these complexes. However, there are favorable electrostatic interactions occurring as well, most notably between the negatively charged D79 residue of psaD from PSI and the positively charged K117 and R119 residues of the hydrogenase, as well as between the negatively charged D31 residue of psaD from PSI and the positively charged K389 residue of the FeFe H_2ase.

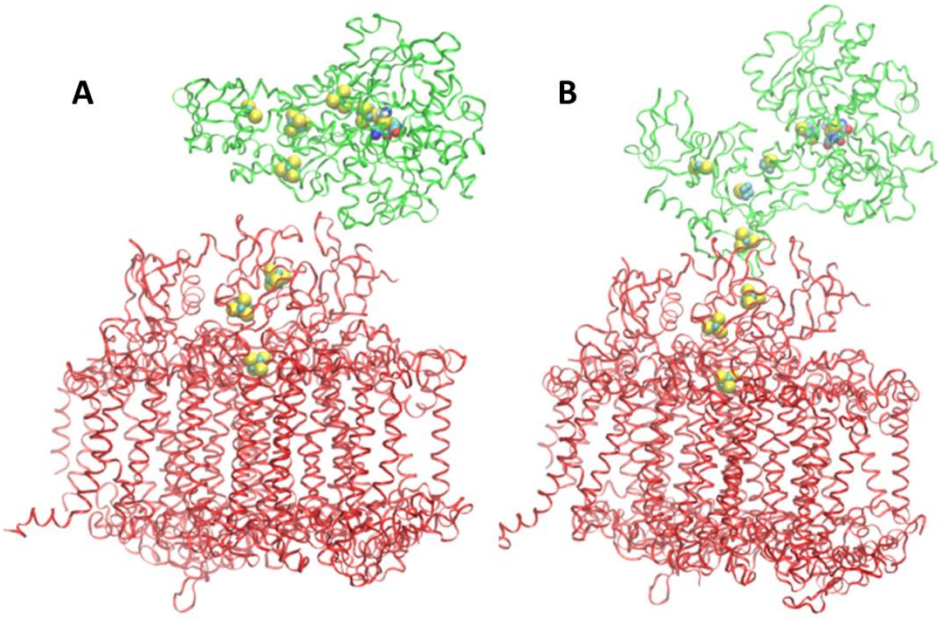

Figure 47. The PSI-FeFe H_2ase DD model (A) before MD simulation and (B) after 170 ns of MD simulation. PSI and the FeFe H_2ase are shown in red and green ribbons, respectively, and the iron-sulfur clusters of PSI as well as the iron-sulfur clusters and active site of the FeFe H_2ase are shown in VDW format.

MD simulations of a PSI-FeFe H_2ase fusion connected via a 1,6-hexanedithiol (HD) ligand were unsuccessful due to molecular overlaps as the initial configuration changes rapidly to accommodate the short alkane linker. Future simulations of this fusion, if successful, could shed light on why this shorter linker leads to lower electron transfer rates. Implications for electron transport based on the PSI-FeFe H_2ase OD and DD model results will be discussed in detail below, after analysis of the PSI-MBH MD simulation.

Stability and local mobility of the PSI-MBH fusion complex

The PSI-MBH model was simulated for 250 ns; the C_α RMSDs versus time of PSI and MBH for the PSI-MBH model are shown in Figures 48A and 48B, respectively. Similar to what was seen in MD simulations of trimeric PSI in DDM detergent solution [41], the C_α RMSD calculated for all residues of PSI (Figure 48A, black solid line) rose to ~2.4 Å after ~25 ns of MD simulation before drifting to ~2.6 Å after ~75 ns and stabilizing there for the remainder of the simulation. The MBH displayed much different behavior, with the C_α RMSD of all residues (Figure 48B, black solid line) rising quickly to ~3.0 Å after only ~10 ns, drifting all the way to ~4.5 Å after ~200 ns and remaining there for the duration of the MD simulation. These values are significantly higher than the C_α RMSD values of ~1.5 Å seen in MD simulations of [NiFe]-hydrogenases from *Desulfovibrio gigas* [245] and *Desulfovibrio fructosivorans* [182] in aqueous solution. As mentioned for the FeFe H_2ase, the high structural drift of the MBH relative to its crystal structure suggests that this protein is undergoing significant structural changes in the PSI-MBH complex. This is not unexpected, as the sortase linkage brings the MBH in close proximity to the stromal surface of PSI, resulting in interactions between charged residues of both of these proteins that will be discussed in further detail below.

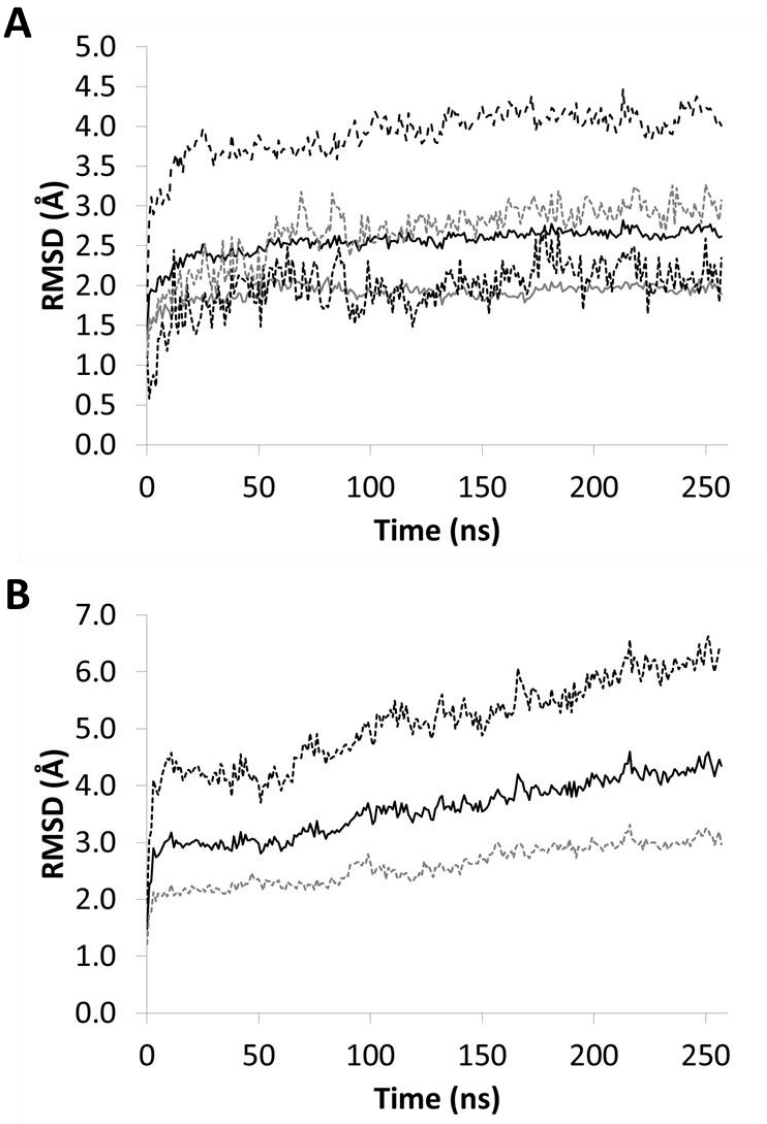

Figure 48. C_α RMSD versus time for the PSI-MBH MD simulation, calculated for (A) PSI monomer relative to the crystal structure (PDB ID: 1JB0). The lines shown the C_α RMSDs for: all residues (black solid line), reaction center subunits psaAB (gray solid line), terminal electron acceptor psaC (black dotted line), ferredoxin docking subunits psaDE (gray dotted line), and the peripheral transmembrane helices psaFIJKLMX (black dashed line). Also C_α RMSDs for (B) MBH relative to the crystal structure (PDB ID: 3RGW). Lines are for: all residues (solid black line), hoxG (dotted gray line), and hoxK (dotted black line).

We also decomposed the structural drift of both PSI and the MBH into those of various individual structural components. As with the PSI-FeFe H_2ase simulations, the main differences concern the stromal subunits. The psaC subunit exhibits C_α RMSDs in the range of 1.5-2.5 Å (Figure 48A, black dotted

line); this value fluctuates considerably over the course of the simulation. This is likely due to protein-protein interactions with the MBH subunits, which are in close proximity to the stromal surface of the PSI monomer. The C_α RMSD of the stromal psaD and psaE subunits (Figure 48A, gray dotted line) also fluctuates noticeably in the range of 2.0-3.0 Å during MD simulation. These subunits are likewise in close proximity to the MBH, and also interact with the detergent belt as was discussed in the previous chapter. Looking at the behavior of the individual MBH subunits, we see that hoxG exhibits lower structural drift with a final C_α RMSD of ~3.0 Å (Figure 48B, gray dotted line), while hoxK has a much higher final C_α RMSD value of ~6.2 Å (Figure 48B, black dotted line). This may be due to the fact that hoxK is tethered to the stromal surface of PSI via the sortase linkage, resulting in substantially different behavior relative to the unrestrained conditions under which the crystal structure of the MBH was determined, the base case from which these RMSD values were derived.

We calculated the time-averaged C_α RMSF values for each residue of PSI and the MBH for the final 125 ns of MD simulation to analyze the local atomic mobility of the fusion complex. The trend for PSI (Figure 49A) was much the same as seen for MD simulations of trimeric PSI in DDM and the PSI-FeFe H_2ase complex. However, higher C_α RMSFs than seen previously were obtained for all residues of PSI, indicating that the PSI protein is more mobile in this complex compared to PSI in DDM alone. The highest observed C_α RMSF of ~5.1 Å corresponds to the C-terminus of psaL. This flexible region, which had previously been shielded by neighboring PSI monomers in the trimeric PSI/DDM MD simulations, now closely interacts with adjacent detergent molecules and the surrounding solvent. For the case of the MBH, the C_α RMSFs of ~2-8 Å observed during MD simulation (Figure 49B, black solid line; Table 12) are much higher than those of ~1 Å obtained from the crystal structure (Figure 49B, black dotted line). This further serves to demonstrate that the tethering of the MBH to PSI via the sortase linkage, and the consequential protein-protein interactions between the MBH and the stromal subunits of PSI, results in substantial

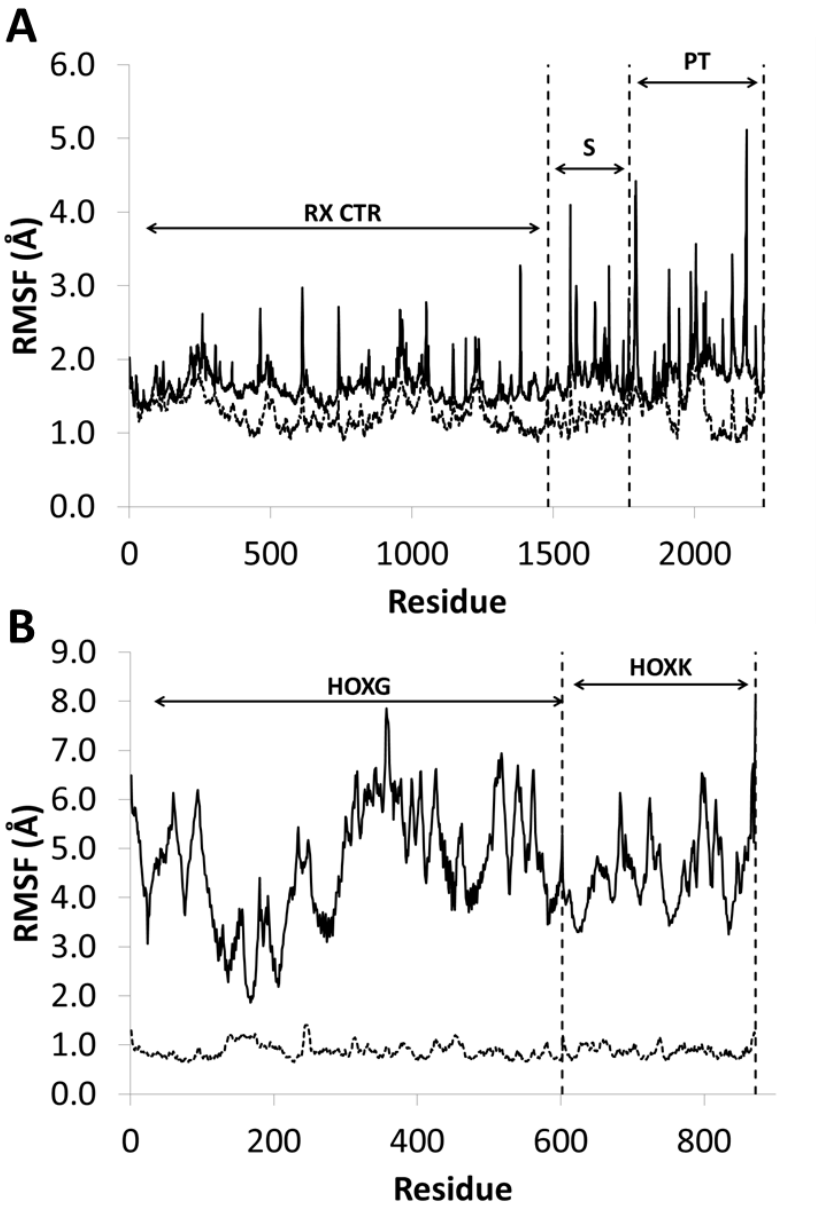

Figure 49. Time-averaged C_α RMSF versus residue number, calculated over the final 125 ns of the PSI-MBH MD simulation, shown for (A) the PSI monomer (solid black line), with residues of the reaction center (RX CTR), stromal (S), and peripheral transmembrane (PT) domains labeled accordingly, and (B) the MBH (solid black line), with residues of the large subunit hoxG and small subunit hoxK labeled correspondingly. In both cases, RMSF values calculated from the temperature factors of the X-ray crystal structures are included for comparison (black dotted line).

structural deviations and fluctuations relative to the crystal structure. Interestingly, however, the highest C_α RMSF values observed for the MBH

correspond not to residues in close proximity to PSI, but rather to flexible loops and termini exposed to the solvent. This suggests that the behavior of this protein during MD simulation can partly be explained by its exposure to aqueous solution conditions. It is recommended that future work include simulations of hydrogenase alone to confirm this.

Protein-protein interactions of PSI-MBH fusion complex

Inspecting the MD trajectory, there is a visually observable contact surface occurring between PSI and the MBH, as shown in Figure 50. This is due to electrostatic interactions occurring between several polar and charged residues of the psaC and psaD subunits of PSI and the hoxG subunit of the MBH over the

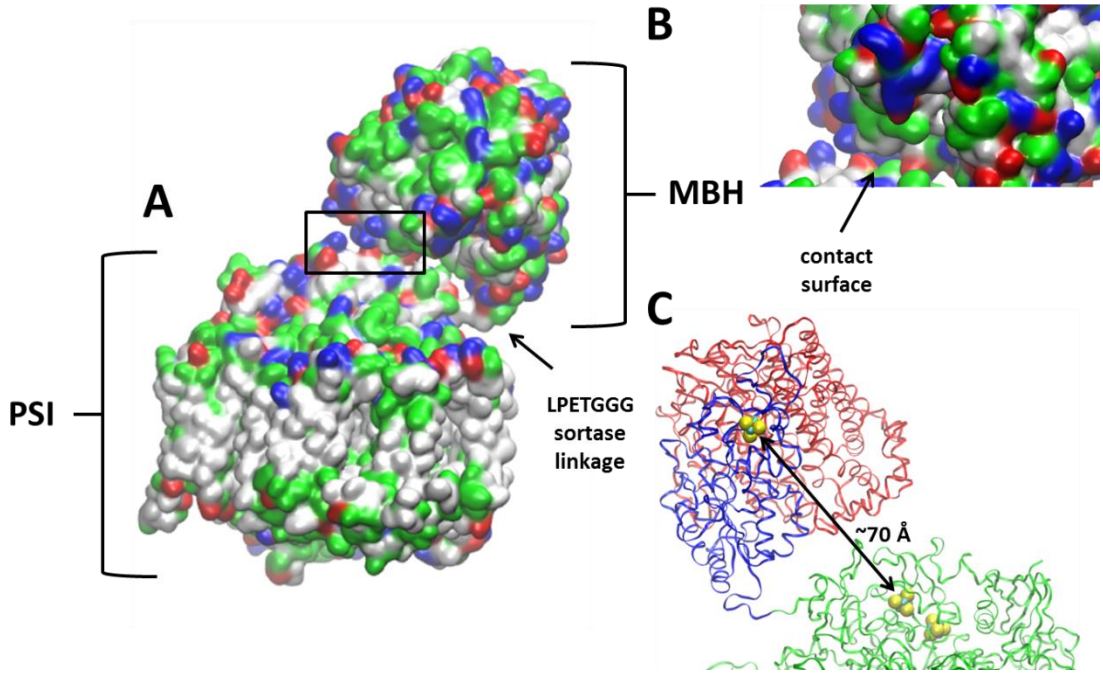

Figure 50. (A) Low-resolution surface representation of the PSI-MBH fusion after MD simulation, colored according to residue type (non-polar = white, polar = green, acidic = red, basic = blue). PSI and the MBH are labeled accordingly, with an arrow indicating the sortase linkage of the fusion complex. The highlighted box corresponds to (B) a zoomed-in view of the fusion complex with an arrow indicating the contact surface of several interacting residues of the stromal subunits of PSI and the MBH. (C) Zoomed-in view of hoxG (red ribbons) and hoxK (blue ribbons) of MBH, and stromal surface of PSI (green ribbons) after 250 ns MD simulation of PSI-MBH protein fusion. Arrow denotes distance between distal iron-sulfur cluster of MBH and F_B cluster of PSI (shown in VDW format).

course of the simulation. In particular, the negatively charged D31 residue of psaC and positively charged K107 residue of psaD from PSI continually interact with the positively charged K138 and negatively charged D140 residues of hoxG from the MBH, respectively, during MD. Unfortunately, there do not appear to be any protein-protein interactions between PSI and the small subunit hoxK of the MBH which would bring the F_B cluster of psaC, the terminal electron acceptor in the electron transport chain of PSI, in close proximity to the distal iron-sulfur cluster of hoxK from the MBH (Figure 50C). These results suggest that the formation of a PSI-MBH fusion complex using a hoxK N-terminal to psaE C-terminal attachment scheme will not provide conditions optimal for electron transport between these two proteins and thus for light-driven H_2 evolution via this fusion complex (see below for details).

Implications for electron transport between PSI and hydrogenase

As a gauge for the efficiency of electron transport, we calculated the average distance between the F_B cluster of PSI and the distal [4Fe4S] cluster of the hydrogenase (FeFe H_2ase or MBH) over the course of the simulation for each system (Table 13). The substantial difference between the average distance for the PSI-FeFe H_2ase OD and PSI-MBH models, 14.5 Å vs. 69.2 Å, serves to demonstrate the effectiveness of the method of Lubner et al. [16] for generating

Table 13. Average cluster-to-cluster distances for PSI-hydrogenase fusion complexes

System	Average cluster-to-cluster distance (Å)
PSI-FeFe H_2ase OD model	14.5 ± 0.5
PSI-FeFe H_2ase DD model	16.8 ± 0.5
PSI-MBH model	69.2 ± 1.9

optimal fusions of PSI and hydrogenase for electron transport. However, as noted above, interactions between PSI and the FeFe H_2ase in these complexes results in a pulling away of the main body of the hydrogenase from the residues surrounding the distal iron-sulfur cluster, which could result in significant strain in these complexes. No such unfavorable interactions were noted for the PSI-MBH MD simulation. This could be because the interacting surfaces of PSI and the MBH are more compatible than that of PSI and the FeFe H_2ase, or because the sortase-mediated linkage puts less strain on the system than the dithiol molecular wire.

To further analyze the dynamics of these multi-protein complexes, we calculated the autocorrelation function (ACF) of the cluster-to-cluster distance for each system. Correlation functions provide a measure of the disorder introduced into a system over time, and the time decay of the ACF of a dynamic variable of interest can provide information on relevant dynamical processes. This analysis has aided in the discovery of collective motions and long-range conformational changes in a variety of protein complexes [246-250]. The rapid initial decay of the ACF for the OD model (Figure 51A, black solid line) is indicative of a fast dynamical process that can be reasonably fit by an exponential decay function (Figure 51A, black dashed line). However, the continued oscillations of the ACF suggest longer timescale dynamics that are not captured by the simulation times studied here. The ACF for the DD model (Figure 51B, black solid line) exhibits an initial decay that is somewhat slower than that seen for the OD model, again with continued oscillations. This strengthens the argument that longer linker lengths allow these proteins to explore a larger conformational space, resulting in longer timescale dynamics. In the case of the PSI-MBH fusion (Figure 51C), the ACF shows substantial oscillations during the simulation, appearing anti-correlated. This suggests that the sortase linkage results in longer timescale dynamics that were not captured in the simulation time evaluated here. These dynamics could potentially lead to favorable conformations with the distal [4Fe4S] cluster of hoxK quite close to the F_B cluster of PSI, which were not seen here.

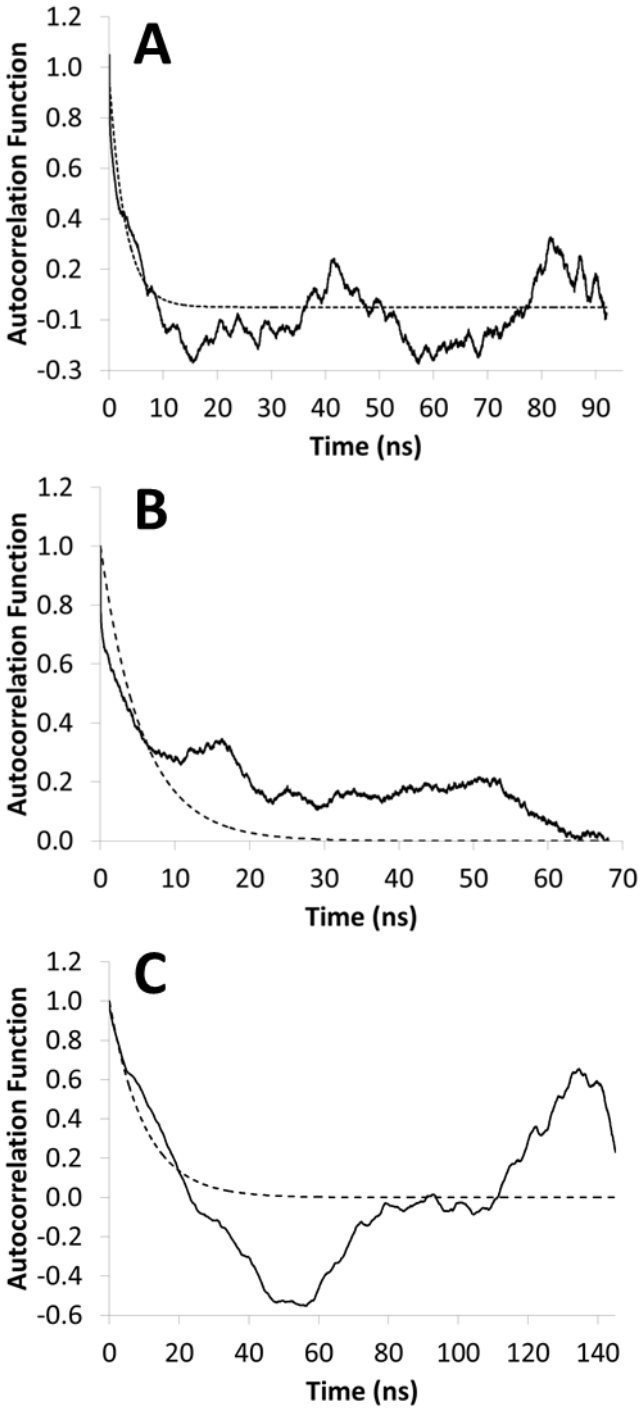

Figure 51. Autocorrelation function (ACF) results versus time for (A) the PSI-FeFe H_2ase OD model and (B) the PSI FeFe H_2ase DD model, and (C) the PSI-MBH model. In all cases, ACF data = black solid line, exponential decay fit = black dashed line.

5.4 Summary

We have genetically engineered the small subunit hoxK of the MBH from *R. eutropha* HF387H to express an N-terminal Gly_3 tag, with the goal of enabling sortase-mediated ligation of this enzyme with mutants of PSI from *Synechocystis* containing C-terminal LPETG tags on either the psaC, psaD, or psaE stromal subunits. This should enable the fusion of hydrogenase via hoxK to complementarily engineered subunits of PSI via sortase-mediated ligation. These protein fusions would be simpler to make than methods based on cysteine replacement and dithiol ligand rescue chemistry, and would also enable the study of a greater variety of fusion complexes.

The mutant hoxK gene has been re-introduced to *R. eutropha* HF387H via homologous recombination and the presence of the N-terminal Gly_3 tag verified by DNA sequencing, resulting in the new mutant *R. eutropha* NGLY strain. We have demonstrated our ability to isolate solubilized membrane extracts containing the fully active MBH heterodimer from wild-type *R. eutropha* H16, confirmed both through Western blot analysis and H_2 uptake activity using methylene blue as the electron acceptor. However, the large subunit hoxG was found to be conspicuously absent from the membrane fraction and solubilized membrane extracts isolated from cell cultures of both *R. eutropha* HF387H and *R. eutropha* NGLY, and instead was located in the soluble fraction at ~52 kDa, instead of its known molecular weight of ~67 kDa. There was also no observable H_2 uptake activity in any fraction tested for either of these mutant strains. We believe this loss of activity most likely stems either from the introduction of an undesirable mutation to the hoxG gene during the design of the original *R. eutropha* HF387H construct containing a His_6 purification tag on the C-terminus of hoxK, or is the result of interference with the Mtt/Tat membrane translocation pathway.

We have also performed MD simulations of PSI-FeFe H_2ase constructs, as well as of the PSI-MBH fusion complex, in order to analyze the stability and protein-protein interactions of these protein fusions and investigate potential

atomistic explanations for observed experimental activity. We have found that longer linker lengths result in greater structural drift and require longer simulation times to reach a stable complex, likely due to the increased conformational space that complexes with longer linkers are able to explore. We also noted highly distorted tertiary structure in the PSI-FeFe H_2ase fusions that could result in significant strain on these complexes. Lastly, autocorrelation analysis suggests that there are dynamics occurring in these multi-protein complexes that are on timescales longer than the simulation times evaluated here, which indicates that coarse-graining may be necessary to capture dynamics relevant to electron transport between these proteins.

CHAPTER VI
CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

This work represents an attempt to form hybrid protein fusions between cyanobacterial photosystem I (PSI) and the membrane-bound hydrogenase (MBH) from *Ralstonia eutropha* via a new method using sortase-mediated ligation. This approach, if successful, has the advantage of being simpler and more versatile than current methods that involve cysteine replacement and dithiol ligand rescue chemistry between the iron-sulfur clusters of the two proteins. Although ultimately unsuccessful in forming these fusion complexes, a number of significant secondary objectives were met.

We conducted a comprehensive study of the solution structure and function of detergent-solubilized trimeric PSI from the cyanobacterium *Thermosynechococcus elongatus*, a membrane integral protein capable of light-driven electron transport with enormous potential for use in biorenewable energy conversion devices. The structure of PSI trimer solubilized in *n*-dodecyl-*β*-maltoside (DDM) detergent was analyzed using contrast variation SANS and yielded results not previously seen. Pairwise distance distribution function (P(r)) analysis of the SANS data and subsequent shape restoration both with and without the scattering contributions of DDM suggests that the detergent exists in a non-uniform trilobal orientation that conforms to the hydrophobic periphery of the PSI trimer.

The SANS results were used as a guide to developing a computational model of trimeric PSI embedded in a disk-like detergent belt. Extensive all-atom MD simulations of trimeric PSI from *T. elongatus* in DDM detergent were conducted for an in-depth analysis of the structure and dynamics of the PSI trimer embedded in the detergent environment, one of the largest membrane protein complexes known to have so far been studied. RMSD analysis of C_α atoms relative to the known crystal structure showed PSI to be stable in the DDM environment, with the largest contributors to the structural drift of this protein

being the peripheral transmembrane helices. We observed a thinning of the detergent belt over the course of the MD simulation as it conforms to the non-uniform hydrophobic periphery of the protein and attribute this to favorable interactions with the hydrophobic transmembrane domain of PSI result in the shrinking of the DDM molecules. Solvent-exposed surface area (SASA) calculations showed that the detergent molecules were less densely packed in the protein-detergent complex relative to a pure detergent micelle, but interestingly dihedral angle order parameter (S^2) and mean-square-displacement (MSD) analyses indicated that the detergent tails were more ordered. These results are contrary to what is seen in most lipid bilayer models, and suggest a degree of plasticity in the *in vivo* complex of PSI embedded in the thylakoid membrane. Lastly, we found interesting implications for the stability of the trimeric superstructure of cyanobacterial PSI *in vitro* due to high fluctuations of trimer-maintaining subunits in the absence of associated lipids. However, the docking of soluble electron mediators, and thus the activity of PSI, was not hindered by the presence of the detergent. These results represent a significant step in atomic-level analysis of the solution structure and conformational dynamics of detergent-solubilized membrane proteins. In particular, this work offers a solid starting point for analyzing the *in vitro* structure and function of PSI, and provides a computational approach that is easily translatable to studying complexes of PSI and other proteins, such as hydrogenase enzyme.

Lastly, this work represents a foundation for the study of PSI-hydrogenase protein fusions, an intriguing multi-protein complex that is capable of light-induced H_2 evolution. We have genetically engineered the MBH of *R. eutropha* to express a Gly_3 tag on the N-terminus of the small subunit hoxK, and have re-introduced this mutation into *R. eutropha* and verified its presence via DNA sequencing. The presence of the N-terminal Gly_3 tag makes possible the site-specific fusing of the MBH with PSI from *Synechocystis* sp. PCC 6803, which is being engineered to express an LPETG tag on the exposed C-termini of several stromal subunits by researchers in our lab, via sortase-mediated ligation (SML).

We have demonstrated a reliable method for the partial purification of the membrane fraction and solubilized membrane extract from wild-type *R. eutropha* H16, verified both through Western blot analysis against the hoxG and hoxK subunits and through assaying for H_2 uptake activity using methylene blue as the electron donor. However, active MBH could not be isolated from the mutant strain expressing an N-terminal Gly_3 tag for sortase-mediated ligation on the MBH small subunit hoxK (designated NGLY in Table 9). Due to the fact that the same result was obtained when working with the mutant strain expressing a C-terminal His_6 purification tag on hoxK (HF387H in Table 9) from which the NGLY strain was derived, we attribute the absence of active MBH to problems associated with this original mutant strain. This may be the result of overlooked undesired mutations to the hoxG gene introduced during construction of the *R. eutropha* HF387H mutant strain. It could also be due to interference of this C-terminal tag with the Mtt/Tat pathway, which serves to target the fully-folded heterodimer to the periplasmic side of the membrane [251]. Interference with membrane translocation due to mutations associated with the C-terminal anchor region of hoxK has been seen in previous work on self-assembling PSI-MBH protein fusions conducted by Schwarze et al. [15]. In parallel, the computational approach demonstrated for trimeric PSI in DDM was applied towards conducting all-atom MD simulations of PSI-FeFe H_2ase and PSI-MBH protein fusions. MD simulations of the PSI-FeFe H_2ase fusions were conducted to attempt to explain the relationship between the length of the dithiol linker and the observed electron transfer rates. These simulations revealed significant distortion of the tertiary structure of the hydrogenase that could limit the stability of these complexes. MD simulations of a PSI-FeFe H_2ase connected via a HD linker were unsuccessful due to molecular overlap caused by rapid structural changes associated with this very short linker. Future simulations of this fusion complex will be interesting, if successful, as they may show why this shorter linker leads to lower electron transfer rates. MD simulations of the PSI-MBH fusion indicated that this ligation strategy did not result in conditions optimal for electron transport. Overall, longer

simulation times may be necessary to adequately capture the dynamics occurring in these systems.

6.2 Future directions

Future research regarding the study of trimeric PSI under detergent-solubilized conditions should be focused on the influence of associated lipids on the structure and function of this photoactive pigment-protein complex. Few, if any, lipids are typically resolved in crystal structures of membrane proteins, with those that are presumably being tightly bound. As is the case for PSI, these lipid molecules (sometimes referred to as non-annular or cofactor lipids) are often located at protein-protein interfaces and are essential for activity [63, 252]. This has been the case for similarly associated phospholipids found to be functionally important in the case of the yeast cytochrome bc(1) complex [214], as well as that of the photosynthetic reaction center of the purple bacterium *Rhodobacter sphaeroides* [215]. Contrast variation SANS of trimeric PSI isolated from *T. elongatus* cell cultures grown in cofactor lipid-containing media could provide unique insights on conformational changes induced by the presence of these lipids and any associated changes in the long-term stability of the PSI trimer. Extending the PSI/DDM with associated lipids MD simulation until a stable complex is reached could also provide new information, such as the binding affinity for these phospholipids to PSI. Also, the contribution to scattering resulting from the presence of DDM micelles in the PSI/DDM SANS data needs to be re-evaluated. The presence of an unexplainable shoulder in the P(r) curve of trimeric PSI in 0.12% (w/v) DDM at vector lengths below 5 nm, as well as the inability of PSI/DDM computational models to reproduce the experimental data in the high q regime, suggests that the scattering contribution of DDM micelles may not be negligible as has been assumed. Collecting SANS data of trimeric PSI solubilized in a lower concentration of DDM may reduce the concentration of free DDM micelles and resolve this issue, giving a clearer picture of the structural envelope of the PSI/DDM complex.

In terms of future work focused on the PSI-hydrogenase protein fusions, the groundwork has been laid for the experimental generation and characterization of sortase-ligated fusions of PSI from *Synechocystis* to the N-terminally Gly_3 tagged hoxK subunit of the MBH from *R. eutropha*. First and foremost, the original mutant strain HF387H needs to be re-visited to ensure that the His_6 purification tag is inserted onto the C-terminus of hoxK in such a way that does not interfere with translocation of the mature heterodimer to the membrane, and that no additional undesired mutations, particularly in the hoxG subunit, are incurred. Also, MBH proteins display very low protein expression levels, making it difficult to characterize the protein due to limited materials. Design of an overexpression vector for the MBH protein in its native host or possibly in *E. coli* would be extremely useful, although previous work has demonstrated the difficulty of this endeavor [239]. Also, FTIR and EPR spectra of wild-type and mutant MBH proteins could provide useful information on the redox states and transitions and intermolecular interactions within the catalytic center [35, 232]. The protocols referenced in this work for sortase-mediated ligation of PSI and MBH, and for assaying the resulting protein fusion for light-induced H_2 evolution, should then provide a solid starting point for experimental characterization of PSI-MBH complexes. SANS analysis of successful fusion complexes can be interpreted using the MD models presented herein. Furthermore, the development of coarse-grained models of all possible permutations for generating PSI-MBH fusions (Figure 52) could provide a simple method for determining the optimum linkage of these two proteins. These coarse-grained models can be developed using a previously demonstrated Brownian dynamics approach [253-256], and may even result in an optimal PSI-MBH fusion with electron throughput rates exceeding that of previously constructed PSI-FeFe H_2ase fusion complexes [16].

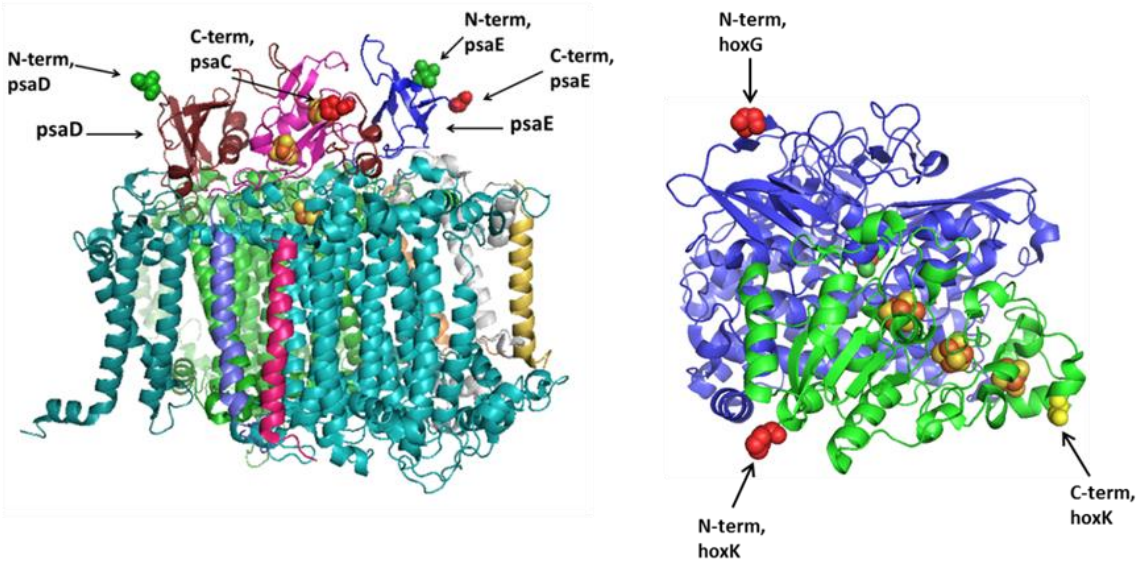

Figure 52. Exposed termini for sortase-mediated ligation of PSI (left) and MBH (right).

LIST OF REFERENCES

1. Wallin, E. and G. von Heijne, Genome-Wide Analysis of Integral Membrane Proteins from Eubacterial, Archaean, and Eukaryotic Organisms. *Protein Science* **1998**. 7(4): p. 1029-1038.
2. White, S.H., Biophysical Dissection of Membrane Proteins. *Nature* **2009**. 459(7245): p. 344-346.
3. White, S.H. and W.C. Wimley, Membrane Protein Folding and Stability: Physical Principles. *Annual Review of Biophysics and Biomolecular Structure* **1999**. 28: p. 319-365.
4. Garavito, R.M. and S. Ferguson-Miller, Detergents as Tools in Membrane Biochemistry. *Journal of Biological Chemistry* **2001**. 276(35): p. 32403-32406.
5. Tamm, L.K. and B.Y. Liang, Nmr of Membrane Proteins in Solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2006**. 48(4): p. 201-210.
6. Lipfert, J. and S. Doniach, *Small-Angle X-Ray Scattering from Rna, Proteins, and Protein Complexes*, in *Annual Review of Biophysics and Biomolecular Structure*. 2007. p. 307-327.
7. Taiz, L. and E. Zeiger, *Plant Physiology*. Plant Physiology. 1998, Sunderland,: Sinauer Associates Incorporated. xxvi + 792.
8. Ghirardi, M.L., et al., *Hydrogenases and Hydrogen Photoproduction in Oxygenic Photosynthetic Organisms*, in *Annu Rev Plant Biol*. 2007, Annual Reviews: Palo Alto. p. 71-91.
9. Jordan, P., et al., Three-Dimensional Structure of Cyanobacterial Photosystem I at 2.5 Angstrom Resolution. *Nature* **2001**. 411(6840): p. 909-917.
10. Golbeck, J.H., Structure and Function of Photosystem-I. *Annual Review of Plant Physiology and Plant Molecular Biology* **1992**. 43: p. 293-324.
11. Chitnis, P.R., Photosystem I. *Plant Physiol* **1996**. 111(3): p. 661-669.
12. Greenbaum, E., Platinized Chloroplasts - a Novel Photocatalytic Material. *Science* **1985**. 230(4732): p. 1373-1375.
13. Iwuchukwu, I.J., et al., Self-Organized Photosynthetic Nanoparticle for Cell-Free Hydrogen Production. *Nature Nanotechnology* **2010**. 5(1): p. 73-79.
14. Iwuchukwu, I.J., et al., Optimization of Photosynthetic Hydrogen Yield from Platinized Photosystem I Complexes Using Response Surface Methodology. *International Journal of Hydrogen Energy* **2011**. 36(18): p. 11684-11692.
15. Schwarze, A., et al., Requirements for Construction of a Functional Hybrid Complex of Photosystem I and [Nife]-Hydrogenase. *Appl Environ Microbiol* **2010**. 76(8): p. 2641-2651.
16. Lubner, C.E., et al., Solar Hydrogen-Producing Bionanodevice Outperforms Natural Photosynthesis. *Proc Natl Acad Sci U S A* **2011**. 108(52): p. 20988-20991.

17. Lee, I., J.W. Lee, and E. Greenbaum, Biomolecular Electronics: Vectorial Arrays of Photosynthetic Reaction Centers. *Physical Review Letters* **1997**. 79(17): p. 3294-3297.
18. Das, R., et al., Integration of Photosynthetic Protein Molecular Complexes in Solid-State Electronic Devices. *Nano Letters* **2004**. 4(6): p. 1079-1083.
19. Ciesielski, P.N., et al., Enhanced Photocurrent Production by Photosystem I Multilayer Assemblies. *Advanced Functional Materials* **2010**. 20(23): p. 4048-4054.
20. LeBlanc, G., et al., Enhanced Photocurrents of Photosystem I Films on P-Doped Silicon. *Advanced Materials* **2012**. 24(44): p. 5959-+.
21. Mershin, A., et al., Self-Assembled Photosystem-I Biophotovoltaics on Nanostructured Tio2 and Zno. *Scientific Reports* **2012**. 2.
22. Lee, I., et al., Measurement of Electrostatic Potentials above Oriented Single Photosynthetic Reaction Centers. *Journal of Physical Chemistry B* **2000**. 104(11): p. 2439-2443.
23. Greenbaum, E., Energetic Efficiency of Hydrogen Photoevolution by Algal Water Splitting. *Biophysical Journal* **1988**. 54(2): p. 365-368.
24. Millsaps, J.F., et al., Nanoscale Photosynthesis: Photocatalytic Production of Hydrogen by Platinized Photosystem I Reaction Centers. *Photochem Photobiol* **2001**. 73(6): p. 630-635.
25. Vignais, P.M. and A. Colbeau, Molecular Biology of Microbial Hydrogenases. *Current Issues in Molecular Biology* **2004**. 6: p. 159-188.
26. Fontecilla-Camps, J.C., et al., Structure/Function Relationships of Nife - and Fefe -Hydrogenases. *Chemical Reviews* **2007**. 107(10): p. 4273-4303.
27. Volbeda, A., et al., Crystal-Structure of the Nickel-Iron Hydrogenase from Desulfovibrio-Gigas. *Nature* **1995**. 373(6515): p. 580-587.
28. Adams, M.W.W., The Structure and Mechanism of Iron-Hydrogenases. *Biochimica Et Biophysica Acta* **1990**. 1020(2): p. 115-145.
29. Zirngibl, C., et al., H-2-Forming Methylenetetrahydromethanopterin Dehydrogenase, a Novel Type of Hydrogenase without Iron-Sulfur Clusters in Methanogenic Archaea. *European Journal of Biochemistry* **1992**. 208(2): p. 511-520.
30. Shima, S., et al., The Cofactor of the Iron-Sulfur Cluster Free Hydrogenase Hmd: Structure of the Light-Inactivation Product. *Angewandte Chemie-International Edition* **2004**. 43(19): p. 2547-2551.
31. Shima, S., et al., Mossbauer Studies of the Iron-Sulfur Cluster-Free Hydrogenase: The Electronic State of the Mononuclear Fe Active Site. *J Am Chem Soc* **2005**. 127(29): p. 10430-10435.
32. Lyon, E.J., et al., Carbon Monoxide as an Intrinsic Ligand to Iron in the Active Site of the Iron-Sulfur-Cluster-Free Hydrogenase H-2-Forming Methylenetetrahydromethanopterin Dehydrogenase as Revealed by Infrared Spectroscopy. *J Am Chem Soc* **2004**. 126(43): p. 14239-14248.

33. Korbas, M., et al., The Iron-Sulfur Cluster-Free Hydrogenase (Hmd) Is a Metalloenzyme with a Novel Iron Binding Motif. *Journal of Biological Chemistry* **2006**. 281(41): p. 30804-30813.
34. Fritsch, J., et al., The Crystal Structure of an Oxygen-Tolerant Hydrogenase Uncovers a Novel Iron-Sulphur Centre. *Nature* **2011**. 479(7372): p. 249-252.
35. Lukey, M.J., et al., Oxygen-Tolerant [Nife]-Hydrogenases: The Individual and Collective Importance of Supernumerary Cysteines at the Proximal Fe-S Cluster. *J Am Chem Soc* **2011**. 133(42): p. 16881-16892.
36. Ducat, D.C., J.C. Way, and P.A. Silver, Engineering Cyanobacteria to Generate High-Value Products. *Trends Biotechnol* **2011**. 29(2): p. 95-103.
37. Ihara, M., et al., Light-Driven Hydrogen Production by a Hybrid Complex of a [Nife]-Hydrogenase and the Cyanobacterial Photosystem I. *Photochem Photobiol* **2006**. 82(3): p. 676-682.
38. Lubner, C.E., et al., Wiring Photosystem I for Direct Solar Hydrogen Production. *Biochemistry* **2010**. 49(3): p. 404-414.
39. Kiley, P., et al., Self-Assembling Peptide Detergents Stabilize Isolated Photosystem I on a Dry Surface for an Extended Time. *Plos Biology* **2005**. 3(7): p. 1180-1186.
40. Le, R., et al., Analysis of the Solution Structure of Thermosynechococcus Elongatus Photosystem I in N-Dodecyl-Beta-D-Maltoside Using Small-Angle Neutron Scattering and Molecular Dynamics Simulation. *Arch Biochem Biophys* **2014**.
41. Harris, B.J., X. Cheng, and P. Frymier, All-Atom Molecular Dynamics Simulation of a Photosystem I/Detergent Complex. *The Journal of Physical Chemistry B* **2014**. 118(40): p. 11633-11645.
42. Henderson, R. and P.N.T. Unwin, 3-Dimensional Model of Purple Membrane Obtained by Electron-Microscopy. *Nature* **1975**. 257(5521): p. 28-32.
43. Deisenhofer, J., et al., Structure of the Protein Subunits in the Photosynthetic Reaction Center of Rhodopseudomonas-Viridis at 3a Resolution. *Nature* **1985**. 318(6047): p. 618-624.
44. Edman, K., et al., High-Resolution X-Ray Structure of an Early Intermediate in the Bacteriorhodopsin Photocycle. *Nature* **1999**. 401(6755): p. 822-826.
45. Buchanan, S.K., Beta-Barrel Proteins from Bacterial Outer Membranes: Structure, Function and Refolding. *Current Opinion in Structural Biology* **1999**. 9(4): p. 455-461.
46. Schulz, G.E., Beta-Barrel Membrane Proteins. *Current Opinion in Structural Biology* **2000**. 10(4): p. 443-447.
47. Doyle, D.A., et al., The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **1998**. 280(5360): p. 69-77.

48. Pebay-Peyroula, E., et al., Structure of Mitochondrial Adp/Atp Carrier in Complex with Carboxyatractyloside. *Nature* **2003**. 426(6962): p. 39-44.
49. Abramson, J., et al., Structure and Mechanism of the Lactose Permease of Escherichia Coli. *Science* **2003**. 301(5633): p. 610-615.
50. Yernool, D., et al., Structure of a Glutamate Transporter Homologue from Pyrococcus Horikoshii. *Nature* **2004**. 431(7010): p. 811-818.
51. Yamashita, A., et al., Crystal Structure of a Bacterial Homologue of Na+/Cl--Dependent Neurotransmitter Transporters. *Nature* **2005**. 437(7056): p. 215-223.
52. Krishnamurthy, H., C.L. Piscitelli, and E. Gouaux, Unlocking the Molecular Secrets of Sodium-Coupled Transporters. *Nature* **2009**. 459(7245): p. 347-355.
53. Lagerstrom, M.C. and H.B. Schioth, Structural Diversity of G Protein-Coupled Receptors and Significance for Drug Discovery. *Nature Reviews Drug Discovery* **2008**. 7(4): p. 339-357.
54. Palczewski, K., et al., Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **2000**. 289(5480): p. 739-745.
55. Rasmussen, S.G.F., et al., Crystal Structure of the Human Beta(2) Adrenergic G-Protein-Coupled Receptor. *Nature* **2007**. 450(7168): p. 383-U384.
56. Stock, D., A.G.W. Leslie, and J.E. Walker, Molecular Architecture of the Rotary Motor in Atp Synthase. *Science* **1999**. 286(5445): p. 1700-1705.
57. Zouni, A., et al., Crystal Structure of Photosystem Ii from Synechococcus Elongatus at 3.8 Angstrom Resolution. *Nature* **2001**. 409(6821): p. 739-743.
58. Guskov, A., et al., Cyanobacterial Photosystem Ii at 2.9-Angstrom Resolution and the Role of Quinones, Lipids, Channels and Chloride. *Nature Structural & Molecular Biology* **2009**. 16(3): p. 334-342.
59. Umena, Y., et al., Crystal Structure of Oxygen-Evolving Photosystem Ii at a Resolution of 1.9 Angstrom. *Nature* **2011**. 473(7345): p. 55-U65.
60. Fromme, P., P. Jordan, and N. Krauss, Structure of Photosystem I. *Biochimica Et Biophysica Acta-Bioenergetics* **2001**. 1507(1-3): p. 5-31.
61. Rasmussen, S.G.F., et al., Crystal Structure of the Beta(2) Adrenergic Receptor-Gs Protein Complex. *Nature* **2011**. 477(7366): p. 549-U311.
62. Whorton, M.R. and R. MacKinnon, X-Ray Structure of the Mammalian Girk2-Beta Gamma G-Protein Complex. *Nature* **2013**. 498(7453): p. 190-+.
63. Lee, A.G., How Lipids and Proteins Interact in a Membrane: A Molecular Approach. *Molecular Biosystems* **2005**. 1(3): p. 203-212.
64. Bieri, M., et al., Macromolecular Nmr Spectroscopy for the Non-Spectroscopist: Beyond Macromolecular Solution Structure Determination. *Febs Journal* **2011**. 278(5): p. 704-715.

65. Brown, M.F., et al., H-1-Nmr Studies of Protein-Lipid Interactions in Retinal Rod Outer Segment Disk Membranes. *Febs Letters* **1976**. 70(1): p. 56-60.
66. Salgado, G.F.J., et al., Deuterium Nmr Structure of Retinal in the Ground State of Rhodopsin. *Biochemistry* **2004**. 43(40): p. 12819-12828.
67. Salgado, G.F.J., et al., Solid-State H-2 Nmr Structure of Retinal in Metarhodopsin I. *J Am Chem Soc* **2006**. 128(34): p. 11067-11071.
68. Struts, A.V., et al., Structural Analysis and Dynamics of Retinal Chromophore in Dark and Metal States of Rhodopsin from H-2 Nmr of Aligned Membranes. *Journal of Molecular Biology* **2007**. 372(1): p. 50-66.
69. Struts, A.V., et al., Retinal Dynamics Underlie Its Switch from Inverse Agonist to Agonist During Rhodopsin Activation. *Nature Structural & Molecular Biology* **2011**. 18(3): p. 392-394.
70. Struts, A.V., G.F.J. Salgado, and M.F. Brown, Solid-State H-2 Nmr Relaxation Illuminates Functional Dynamics of Retinal Cofactor in Membrane Activation of Rhodopsin. *Proc Natl Acad Sci U S A* **2011**. 108(20): p. 8263-8268.
71. Patel, A.B., et al., Coupling of Retinal Isomerization to the Activation of Rhodopsin. *Proc Natl Acad Sci U S A* **2004**. 101(27): p. 10048-10053.
72. Patel, A.B., et al., Changes in Interhelical Hydrogen Bonding Upon Rhodopsin Activation. *Journal of Molecular Biology* **2005**. 347(4): p. 803-812.
73. Crocker, E., et al., Location of Trp265 in Metarhodopsin Ii: Implications for the Activation Mechanism of the Visual Receptor Rhodopsin. *Journal of Molecular Biology* **2006**. 357(1): p. 163-172.
74. Ahuja, S., et al., Helix Movement Is Coupled to Displacement of the Second Extracellular Loop in Rhodopsin Activation. *Nature Structural & Molecular Biology* **2009**. 16(2): p. 168-175.
75. Gautier, A., J.P. Kirkpatrick, and D. Nietlispach, Solution-State Nmr Spectroscopy of a Seven-Helix Transmembrane Protein Receptor: Backbone Assignment, Secondary Structure, and Dynamics. *Angewandte Chemie-International Edition* **2008**. 47(38): p. 7297-7300.
76. Nisius, L., et al., Large-Scale Expression and Purification of the Major Hiv-1 Coreceptor Ccr5 and Characterization of Its Interaction with Rantes. *Protein Expression and Purification* **2008**. 61(2): p. 155-162.
77. Ruan, K.H., V. Cervantes, and J.X. Wu, A Simple, Quick, and High-Yield Preparation of the Human Thromboxane a(2) Receptor in Full Size for Structural Studies. *Biochemistry* **2008**. 47(26): p. 6819-6826.
78. Ruan, K.H., V. Cervantes, and J.X. Wu, Ligand-Specific Conformation Determines Agonist Activation and Antagonist Blockade in Purified Human Thromboxane A2 Receptor. *Biochemistry* **2009**. 48(14): p. 3157-3165.
79. Lange, A., et al., Toxin-Induced Conformational Changes in a Potassium Channel Revealed by Solid-State Nmr. *Nature* **2006**. 440(7086): p. 959-962.

80. Schneider, R., et al., Solid-State Nmr Spectroscopy Applied to a Chimeric Potassium Channel in Lipid Bilayers. *J Am Chem Soc* **2008**. 130(23): p. 7427-7435.
81. Ader, C., et al., A Structural Link between Inactivation and Block of a K+ Channel. *Nature Structural & Molecular Biology* **2008**. 15(6): p. 605-612.
82. Ader, C., et al., Coupling of Activation and Inactivation Gate in a K+-Channel: Potassium and Ligand Sensitivity. *Embo Journal* **2009**. 28(18): p. 2825-2834.
83. Cady, S.D., et al., Structure and Function of the Influenza a M2 Proton Channel. *Biochemistry* **2009**. 48(31): p. 7356-7364.
84. Schnell, J.R. and J.J. Chou, Structure and Mechanism of the M2 Proton Channel of Influenza a Virus. *Nature* **2008**. 451(7178): p. 591-U512.
85. Cady, S.D. and M. Hong, Amantadine-Induced Conformational and Dynamical Changes of the Influenza M2 Transmembrane Proton Channel. *Proc Natl Acad Sci U S A* **2008**. 105(5): p. 1483-1488.
86. Cady, S.D., T.V. Mishanina, and M. Hong, Structure of Amantadine-Bound M2 Transmembrane Peptide of Influenza a in Lipid Bilayers from Magic-Angle-Spinning Solid-State Nmr: The Role of Ser31 in Amantadine Binding. *Journal of Molecular Biology* **2009**. 385(4): p. 1127-1141.
87. Wang, J., et al., Discovery of Spiro-Piperidine Inhibitors and Their Modulation of the Dynamics of the M2 Proton Channel from Influenza a Virus. *J Am Chem Soc* **2009**. 131(23): p. 8066-8076.
88. Hu, F.H., W.B. Luo, and M. Hong, Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State Nmr. *Science* **2010**. 330(6003): p. 505-508.
89. Ubbink, M. and D.S. Bendall, Complex of Plastocyanin and Cytochrome C Characterized by Nmr Chemical Shift Analysis. *Biochemistry* **1997**. 36(21): p. 6326-6335.
90. Worrall, J.A.R., et al., Interaction of Yeast Iso-1-Cytochrome C with Cytochrome C Peroxidase Investigated by N-15,H-1 Heteronuclear Nmr Spectroscopy. *Biochemistry* **2001**. 40(24): p. 7069-7076.
91. Worrall, J.A.R., et al., Myoglobin and Cytochrome B(5): A Nuclear Magnetic Resonance Study of a Highly Dynamic Protein Complex. *Biochemistry* **2002**. 41(39): p. 11721-11730.
92. Crowley, P.B., et al., The Ternary Complex of Cytochrome F and Cytochrome C: Identification of a Second Binding Site and Competition for Plastocyanin Binding. *Chembiochem* **2002**. 3(6): p. 526-533.
93. Worrall, J.A.R., et al., Transient Protein Interactions Studied by Nmr Spectroscopy: The Case of Cytochrome C and Adrenodoxin. *Biochemistry* **2003**. 42(23): p. 7068-7076.
94. Yokogawa, M., et al., Nmr Analyses of the G Beta Gamma Binding and Conformational Rearrangements of the Cytoplasmic Pore of G Protein-Activated Inwardly Rectifying Potassium Channel 1 (Girk1). *Journal of Biological Chemistry* **2011**. 286(3): p. 2215-2223.

95. Sanders, C.R. and F. Sonnichsen, Solution Nmr of Membrane Proteins: Practice and Challenges. *Magnetic Resonance in Chemistry* **2006**. 44: p. S24-S40.
96. Hong, M., Y. Zhang, and F.H. Hu, Membrane Protein Structure and Dynamics from Nmr Spectroscopy. *Annual Review of Physical Chemistry, Vol 63* **2012**. 63: p. 1-24.
97. Wall, M.E., S.C. Gallagher, and J. Trewhella, Large-Scale Shape Changes in Proteins and Macromolecular Complexes. *Annual Review of Physical Chemistry* **2000**. 51: p. 355-380.
98. Sardet, C., A. Tardieu, and V. Luzzati, Shape and Size of Bovine Rhodopsin - Small-Angle X-Ray-Scattering Study of a Rhodopsin-Detergent Complex. *Journal of Molecular Biology* **1976**. 105(3): p. 383-407.
99. Mo, Y., et al., Detergent-Associated Solution Conformations of Helical and Beta-Barrel Membrane Proteins. *Journal of Physical Chemistry B* **2008**. 112(42): p. 13349-13354.
100. Franke, D. and D.I. Svergun, Dammif, a Program for Rapid Ab-Initio Shape Determination in Small-Angle Scattering. *Journal of Applied Crystallography* **2009**. 42: p. 342-346.
101. Hong, X.G., Y.X. Weng, and M. Li, Determination of the Topological Shape of Integral Membrane Protein Light-Harvesting Complex Lh2 from Photosynthetic Bacteria in the Detergent Solution by Small-Angle X-Ray Scattering. *Biophysical Journal* **2004**. 86(2): p. 1082-1088.
102. O'Neill, H., et al., Small-Angle X-Ray Scattering Study of Photosystem I - Detergent Complexes: Implications for Membrane Protein Crystallization. *Journal of Physical Chemistry B* **2007**. 111(16): p. 4211-4219.
103. Xu, X.F., et al., Dynamics in a Pure Encounter Complex of Two Proteins Studied by Solution Scattering and Paramagnetic Nmr Spectroscopy. *J Am Chem Soc* **2008**. 130(20): p. 6395-6403.
104. Osborne, H.B., et al., Structural Study of Rhodopsin in Detergent Micelles by Small-Angle Neutron-Scattering. *Journal of Molecular Biology* **1978**. 123(2): p. 177-206.
105. Santonicola, M.G., A.M. Lenhoff, and E.W. Kaler, Binding of Alkyl Polyglucoside Surfactants to Bacteriorhodopsin and Its Relation to Protein Stability. *Biophysical Journal* **2008**. 94(9): p. 3647-3658.
106. Cardoso, M.B., et al., Insight into the Structure of Light-Harvesting Complex Ii and Its Stabilization in Detergent Solution. *Journal of Physical Chemistry B* **2009**. 113(51): p. 16377-16383.
107. Bu, Z.M., L.O. Wang, and D.A. Kendall, Nucleotide Binding Induces Changes in the Oligomeric State and Conformation of Sec a in a Lipid Environment: A Small-Angle Neutron-Scattering Study. *Journal of Molecular Biology* **2003**. 332(1): p. 23-30.

108. Comoletti, D., et al., Synaptic Arrangement of the Neuroligin/Beta-Neurexin Complex Revealed by X-Ray and Neutron Scattering. *Structure* **2007**. 15(6): p. 693-705.
109. Nelson, M.T., et al., Namd: A Parallel, Object Oriented Molecular Dynamics Program. *International Journal of Supercomputer Applications and High Performance Computing* **1996**. 10(4): p. 251-268.
110. Lindahl, E., B. Hess, and D. van der Spoel, Gromacs 3.0: A Package for Molecular Simulation and Trajectory Analysis. *Journal of Molecular Modeling* **2001**. 7(8): p. 306-317.
111. Case, D.A., et al., The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry* **2005**. 26(16): p. 1668-1688.
112. Hansson, T., C. Oostenbrink, and W.F. van Gunsteren, Molecular Dynamics Simulations. *Current Opinion in Structural Biology* **2002**. 12(2): p. 190-196.
113. Tjioe, E. and W.T. Heller, Ornl_Sas: Software for Calculation of Small-Angle Scattering Intensities of Proteins and Protein Complexes. *Journal of Applied Crystallography* **2007**. 40: p. 782-785.
114. Egberts, E., S.J. Marrink, and H.J.C. Berendsen, Molecular-Dynamics Simulation of a Phospholipid Membrane. *European Biophysics Journal with Biophysics Letters* **1994**. 22(6): p. 423-436.
115. Tu, K., D.J. Tobias, and M.L. Klein, Constant Pressure and Temperature Molecular Dynamics Simulation of a Fully Hydrated Liquid Crystal Phase Dipalmitoylphosphatidylcholine Bilayer. *Biophysical Journal* **1995**. 69(6): p. 2558-2562.
116. Berger, O., O. Edholm, and F. Jahnig, Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. *Biophysical Journal* **1997**. 72(5): p. 2002-2013.
117. Raghavan, K., M.R. Reddy, and M.L. Berkowitz, A Molecular-Dynamics Study of the Structure and Dynamics of Water between Dilauroylphosphatidylethanolamine Bilayers. *Langmuir* **1992**. 8(1): p. 233-240.
118. Venable, R.M., et al., Molecular-Dynamics Simulations of a Lipid Bilayer and of Hexadecane - an Investigation of Membrane Fluidity. *Science* **1993**. 262(5131): p. 223-226.
119. Chiu, S.W., et al., Incorporation of Surface-Tension into Molecular-Dynamics Simulation of an Interface - a Fluid-Phase Lipid Bilayer-Membrane. *Biophysical Journal* **1995**. 69(4): p. 1230-1245.
120. Feller, S.E., Y.H. Zhang, and R.W. Pastor, Computer-Simulation of Liquid/Liquid Interfaces .2. Surface-Tension Area Dependence of a Bilayer and Monolayer. *Journal of Chemical Physics* **1995**. 103(23): p. 10267-10276.

121. Feller, S.E. and R.W. Pastor, Constant Surface Tension Simulations of Lipid Bilayers: The Sensitivity of Surface Areas and Compressibilities. *Journal of Chemical Physics* **1999**. 111(3): p. 1281-1287.
122. Tieleman, D.P. and H.J.C. Berendsen, A Molecular Dynamics Study of the Pores Formed by Escherichia Coli Ompf Porin in a Fully Hydrated Palmitoyloleoylphosphatidylcholine Bilayer. *Biophysical Journal* **1998**. 74(6): p. 2786-2801.
123. Im, W. and B. Roux, Ions and Counterions in a Biological Channel: A Molecular Dynamics Simulation of Ompf Porin from Escherichia Coli in an Explicit Membrane with 1 M Kcl Aqueous Salt Solution. *Journal of Molecular Biology* **2002**. 319(5): p. 1177-1197.
124. Bond, P.J., J.D. Faraldo-Gomez, and M.S.P. Sansom, Ompa: A Pore or Not a Pore? Simulation and Modeling Studies. *Biophysical Journal* **2002**. 83(2): p. 763-775.
125. de Groot, B.L. and H. Grubmuller, Water Permeation across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and Glpf. *Science* **2001**. 294(5550): p. 2353-2357.
126. Zhu, F.Q., E. Tajkhorshid, and K. Schulten, Molecular Dynamics Study of Aquaporin-1 Water Channel in a Lipid Bilayer. *Febs Letters* **2001**. 504(3): p. 212-218.
127. Woolf, T.B. and B. Roux, Structure, Energetics, and Dynamics of Lipid-Protein Interactions: A Molecular Dynamics Study of the Gramicidin a Channel in a Dmpc Bilayer. *Proteins-Structure Function and Genetics* **1996**. 24(1): p. 92-114.
128. Chiu, S.W., S. Subramaniam, and E. Jakobsson, Simulation Study of a Gramicidin/Lipid Bilayer System in Excess Water and Lipid. I. Structure of the Molecular Complex. *Biophysical Journal* **1999**. 76(4): p. 1929-1938.
129. Berneche, S. and B. Roux, Molecular Dynamics of the Kcsa K+ Channel in a Bilayer Membrane. *Biophysical Journal* **2000**. 78(6): p. 2900-2917.
130. Shrivastava, I.H. and M.S.P. Sansom, Simulations of Ion Permeation through a Potassium Channel: Molecular Dynamics of Kcsa in a Phospholipid Bilayer. *Biophysical Journal* **2000**. 78(2): p. 557-570.
131. Jensen, M.O., et al., Mechanism of Voltage Gating in Potassium Channels. *Science* **2012**. 336(6078): p. 229-233.
132. Elmore, D.E. and D.A. Dougherty, Molecular Dynamics Simulations of Wild-Type and Mutant Forms of the Mycobacterium Tuberculosis Mscl Channel. *Biophysical Journal* **2001**. 81(3): p. 1345-1359.
133. Gullingsrud, J., D. Kosztin, and K. Schulten, Structural Determinants of Mscl Gating Studied by Molecular Dynamics Simulations. *Biophysical Journal* **2001**. 80(5): p. 2074-2081.
134. Huber, T., et al., Membrane Model for the G-Protein-Coupled Receptor Rhodopsin: Hydrophobic Interface and Dynamical Structure. *Biophysical Journal* **2004**. 86(4): p. 2078-2100.

135. Periole, X., et al., G Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers. *J Am Chem Soc* **2007**. 129(33): p. 10126-10132.
136. Nygaard, R., et al., The Dynamic Process of Beta(2)-Adrenergic Receptor Activation. *Cell* **2013**. 152(3): p. 532-542.
137. Vasil'ev, S. and D. Bruce, A Protein Dynamics Study of Photosystem Ii: The Effects of Protein Conformation on Reaction Center Function. *Biophysical Journal* **2006**. 90(9): p. 3062-3073.
138. Vassiliev, S., et al., Tracking the Flow of Water through Photosystem Ii Using Molecular Dynamics and Streamline Tracing. *Biochemistry* **2010**. 49(9): p. 1873-1881.
139. Ogata, K., et al., All-Atom Molecular Dynamics Simulation of Photosystem Ii Embedded in Thylakoid Membrane. *J Am Chem Soc* **2013**. 135(42): p. 15670-15673.
140. Tarek, M., S. Bandyopadhyay, and M.L. Klein, Molecular Dynamics Studies of Aqueous Surfactants Systems. *Journal of Molecular Liquids* **1998**. 78(1-2): p. 1-6.
141. Bruce, C.D., et al., Molecular Dynamics Simulation of Sodium Dodecyl Sulfate Micelle in Water: Micellar Structural Characteristics and Counterion Distribution. *Journal of Physical Chemistry B* **2002**. 106(15): p. 3788-3793.
142. Bocker, J., J. Brickmann, and P. Bopp, Molecular-Dynamics Simulation Study of an N-Decyltrimethylammonium Chloride Micelle in Water. *Journal of Physical Chemistry* **1994**. 98(2): p. 712-717.
143. Bogusz, S., R.M. Venable, and R.W. Pastor, Molecular Dynamics Simulations of Octyl Glucoside Micelles: Dynamic Properties. *Journal of Physical Chemistry B* **2001**. 105(35): p. 8312-8321.
144. Abel, S., et al., Molecular Simulations of Dodecyl-Beta-Maltoside Micelles in Water: Influence of the Headgroup Conformation and Force Field Parameters. *Journal of Physical Chemistry B* **2011**. 115(3): p. 487-499.
145. Wendoloski, J.J., et al., Molecular-Dynamics Simulation of a Phospholipid Micelle. *Science* **1989**. 243(4891): p. 636-638.
146. Wymore, T., X.F. Gao, and T.C. Wong, Molecular Dynamics Simulation of the Structure and Dynamics of a Dodecylphosphocholine Micelle in Aqueous Solution. *Journal of Molecular Structure* **1999**. 485: p. 195-210.
147. Abel, S., F.-Y. Dupradeau, and M. Marchi, Molecular Dynamics Simulations of a Characteristic Dpc Micelle in Water. *Journal of Chemical Theory and Computation* **2012**. 8(11): p. 4610-4623.
148. Braun, R., D.M. Engelman, and K. Schulten, Molecular Dynamics Simulations of Micelle Formation around Dimeric Glycophorin a Transmembrane Helices. *Biophysical Journal* **2004**. 87(2): p. 754-763.
149. Bond, P.J., et al., Md Simulations of Spontaneous Membrane Protein/Detergent Micelle Formation. *J Am Chem Soc* **2004**. 126(49): p. 15948-15949.

150. Bond, P.J. and M.S.P. Sansom, Membrane Protein Dynamics Versus Environment: Simulations of Ompa in a Micelle and in a Bilayer. *Journal of Molecular Biology* **2003**. 329(5): p. 1035-1053.
151. Rodriguez-Ropero, F. and M. Fioroni, Structural and Dynamical Analysis of an Engineered Fhua Channel Protein Embedded into a Lipid Bilayer or a Detergent Belt. *Journal of Structural Biology* **2012**. 177(2): p. 291-301.
152. Koutsioubas, A., et al., Ab Initio and All-Atom Modeling of Detergent Organization around Aquaporin-0 Based on Saxs Data. *Journal of Physical Chemistry B* **2013**. 117(43): p. 13588-13594.
153. Bui, J.M. and J.A. McCammon, Protein Complex Formation by Acetylcholinesterase and the Neurotoxin Fasciculin-2 Appears to Involve an Induced-Fit Mechanism. *Proc Natl Acad Sci U S A* **2006**. 103(42): p. 15451-15456.
154. Bourne, Y., P. Taylor, and P. Marchot, Acetylcholinesterase Inhibition by Fasciculin - Crystal-Structure of the Complex. *Cell* **1995**. 83(3): p. 503-512.
155. Harel, M., et al., Crystal Structure of an Acetylcholinesterase-Fasciculin Complex: Interaction of a Three-Fingered Toxin from Snake Venom with Its Target. *Structure* **1995**. 3(12): p. 1355-1366.
156. Gsponer, J., et al., A Coupled Equilibrium Shift Mechanism in Calmodulin-Mediated Signal Transduction. *Structure* **2008**. 16(5): p. 736-746.
157. Malmendal, A., et al., Structural Dynamics in the C-Terminal Domain of Calmodulin at Low Calcium Levels. *Journal of Molecular Biology* **1999**. 293(4): p. 883-899.
158. Lee, J.W., R.T. Collins, and E. Greenbaum, Molecular Ionic Probes: A New Class of Hill Reagents and Their Potential for Nanofabrication and Biometallocatalysis. *Journal of Physical Chemistry B* **1998**. 102(11): p. 2095-2100.
159. Grimme, R.A., et al., Photosystem I/Molecular Wire/Metal Nanoparticle Bioconjugates for the Photocatalytic Production of H-2. *J Am Chem Soc* **2008**. 130(20): p. 6308-+.
160. Grimme, R.A., C.E. Lubner, and J.H. Golbeck, Maximizing H-2 Production in Photosystem I/Dithiol Molecular Wire/Platinum Nanoparticle Bioconjugates. *Dalton Transactions* **2009**(45): p. 10106-10113.
161. Gorka, M., et al., Light-Mediated Hydrogen Generation in Photosystem I: Attachment of a Naphthoquinone-Molecular Wire-Pt Nanoparticle to the a(1a) and a(1b) Sites. *Biochemistry* **2014**. 53(14): p. 2295-2306.
162. Antonkine, M.L., et al., Chemical Rescue of a Site-Modified Ligand to a 4fe-4s Cluster in Psac, a Bacterial-Like Dicluster Ferredoxin Bound to Photosystem I. *Biochimica Et Biophysica Acta-Bioenergetics* **2007**. 1767(6): p. 712-724.
163. Ihara, M., et al., Photoinduced Hydrogen Production by Direct Electron Transfer from Photosystem I Cross-Linked with Cytochrome C(3) to Nife - Hydrogenase. *Photochem Photobiol* **2006**. 82(6): p. 1677-1685.

164. Krassen, H., et al., Photosynthetic Hydrogen Production by a Hybrid Complex of Photosystem I and Nife -Hydrogenase. *Acs Nano* **2009**. 3(12): p. 4055-4061.
165. Lubner, C.E., et al., Wiring an Fefe -Hydrogenase with Photosystem I for Light-Induced Hydrogen Production. *Biochemistry* **2010**. 49(48): p. 10264-10266.
166. Nakamura, Y., et al., Complete Genome Structure of the Thermophilic Cyanobacterium Thermosynechococcus Elongatus Bp-1. *DNA Research* **2002**. 9(4): p. 123-130.
167. Lynn, G.W., et al., Bio-Sans - a Dedicated Facility for Neutron Structural Biology at Oak Ridge National Laboratory. *Physica B-Condensed Matter* **2006**. 385-86: p. 880-882.
168. Timmins, P.A., et al., A Physical Characterization of Some Detergents of Potential Use for Membrane-Protein Crystallization. *Febs Letters* **1988**. 238(2): p. 361-368.
169. Guinier, A., X-Ray Diffraction at Small Angles. *Annales de Physique* **1939**. 12: p. 161-237.
170. Jacques, D.A. and J. Trewhella, Small-Angle Scattering for Structural Biology-Expanding the Frontier While Avoiding the Pitfalls. *Protein Science* **2010**. 19(4): p. 642-657.
171. Semenyuk, A.V. and D.I. Svergun, Gnom - a Program Package for Small-Angle Scattering Data-Processing. *Journal of Applied Crystallography* **1991**. 24: p. 537-540.
172. Konarev, P.V., et al., Primus: A Windows Pc-Based System for Small-Angle Scattering Data Analysis. *Journal of Applied Crystallography* **2003**. 36: p. 1277-1282.
173. Volkov, V.V. and D.I. Svergun, Uniqueness of Ab Initio Shape Determination in Small-Angle Scattering. *Journal of Applied Crystallography* **2003**. 36: p. 860-864.
174. Kozin, M.B. and D.I. Svergun, Automated Matching of High- and Low-Resolution Structural Models. *Journal of Applied Crystallography* **2001**. 34: p. 33-41.
175. Humphrey, W., A. Dalke, and K. Schulten, Vmd: Visual Molecular Dynamics. *Journal of Molecular Graphics & Modelling* **1996**. 14(1): p. 33-38.
176. MacKerell, A.D., et al., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B* **1998**. 102(18): p. 3586-3616.
177. Guvench, O., et al., Additive Empirical Force Field for Hexopyranose Monosaccharides. *Journal of Computational Chemistry* **2008**. 29(15): p. 2543-2564.
178. Guvench, O., et al., Charmm Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *Journal of Chemical Theory and Computation* **2009**. 5(9): p. 2353-2370.

179. Klauda, J.B., et al., Update of the Charmm All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *Journal of Physical Chemistry B* **2010**. 114(23): p. 7830-7843.
180. Raman, E.P., O. Guvench, and A.D. MacKerell, Charmm Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses. *Journal of Physical Chemistry B* **2010**. 114(40): p. 12981-12994.
181. Zhang, L., et al., Force Field Development for Cofactors in the Photosystem Ii. *Journal of Computational Chemistry* **2012**. 33(25): p. 1969-1980.
182. Smith, D.M.A., et al., Force-Field Development and Molecular Dynamics of Nife Hydrogenase. *Journal of Chemical Theory and Computation* **2012**. 8(6): p. 2103-2114.
183. Auer, M., G.A. Scarborough, and W. Kuhlbrandt, Three-Dimensional Map of the Plasma Membrane H+-Atpase in the Open Conformation. *Nature* **1998**. 392(6678): p. 840-843.
184. Hammouda, B., *Probing Nanoscale Structures - the Sans Toolbox*. 2010: National Institute of Standards and Technology Center for Neutron Research.
185. Schrodinger, L., *The Pymol Molecular Graphics System*. 2010.
186. Dupuy, C., X. Auvray, and C. Petipas, Anomeric Effects on the Structure of Micelles of Alkyl Maltosides in Water. *Langmuir* **1997**. 13(15): p. 3965-3967.
187. Strop, P. and A.T. Brunger, Refractive Index-Based Determination of Detergent Concentration and Its Application to the Study of Membrane Proteins. *Protein Science* **2005**. 14(8): p. 2207-2211.
188. Lipfert, J., et al., Size and Shape of Detergent Micelles Determined by Small-Angle X-Ray Scattering. *Journal of Physical Chemistry B* **2007**. 111(43): p. 12427-12438.
189. Barret, L.A., et al., Influence of Hydrophobic Micelle Structure on Crystallization of the Photosynthetic Rc-Lh1-Pufx Complex from Rhodobacter Blasticus. *Journal of Physical Chemistry B* **2013**. 117(29): p. 8770-8781.
190. Roth, M., et al., Detergent Structure in Crystals of a Bacterial Photosynthetic Reaction Center. *Nature* **1989**. 340(6235): p. 659-662.
191. Moller, J.V. and M. Lemaire, Detergent Binding as a Measure of Hydrophobic Surface-Area of Integral Membrane-Proteins. *Journal of Biological Chemistry* **1993**. 268(25): p. 18659-18672.
192. Pebaypeyroula, E., et al., Detergent Structure in Tetragonal Crystals of Ompf Porin. *Structure* **1995**. 3(10): p. 1051-1059.
193. le Maire, M., P. Champeil, and J.V. Moller, Interaction of Membrane Proteins and Lipids with Solubilizing Detergents. *Biochimica Et Biophysica Acta-Biomembranes* **2000**. 1508(1-2): p. 86-111.

194. Muh, F. and A. Zouni, Micelle Formation in the Presence of Photosystem I. *Biochimica Et Biophysica Acta-Biomembranes* **2008**. 1778(10): p. 2298-2307.
195. Le, R.K., et al., Analysis of the Solution Structure of Thermosynechococcus Elongatus Photosystem I in N-Dodecyl-Beta-D-Maltoside Using Small-Angle Neutron Scattering and Molecular Dynamics Simulation. *Arch Biochem Biophys* **2014**. 550-551: p. 50-57.
196. Martinez, L., et al., Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *Journal of Computational Chemistry* **2009**. 30(13): p. 2157-2164.
197. Darden, T., D. York, and L. Pedersen, Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **1993**. 98(12): p. 10089-10092.
198. Feller, S.E., et al., Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *Journal of Chemical Physics* **1995**. 103(11): p. 4613-4621.
199. Pierce, B.G., et al., Zdock Server: Interactive Docking Prediction of Protein-Protein Complexes and Symmetric Multimers. *Bioinformatics* **2014**(Copyright (C) 2014 U.S. National Library of Medicine.).
200. Sommer, F., et al., The Hydrophobic Recognition Site Formed by Residues Psaa-Trp(651) and Psab-Trp(627) of Photosystem I in Chlamydomonas Reinhardtii Confers Distinct Selectivity for Binding of Plastocyanin and Cytochrome C(6). *Journal of Biological Chemistry* **2004**. 279(19): p. 20009-20017.
201. Setif, P., et al., The Ferredoxin Docking Site of Photosystem I. *Biochimica Et Biophysica Acta-Bioenergetics* **2002**. 1555(1-3): p. 204-209.
202. Mukherjee, D., M. May, and B. Khomami, Detergent-Protein Interactions in Aqueous Buffer Suspensions of Photosystem I (Ps I). *Journal of Colloid and Interface Science* **2011**. 358(2): p. 477-484.
203. Deol, S.S., et al., Lipid-Protein Interactions of Integral Membrane Proteins: A Comparative Simulation Study. *Biophysical Journal* **2004**. 87(6): p. 3737-3749.
204. Zilber, A.L. and R. Malkin, Organization and Topology of Photosystem-I Subunits. *Plant Physiol* **1992**. 99(3): p. 901-911.
205. Chandler, D.E., et al., Intrinsic Curvature Properties of Photosynthetic Proteins in Chromatophores. *Biophysical Journal* **2008**. 95(6): p. 2822-2836.
206. Psachoulia, E., P.J. Bond, and M.S.P. Sansom, Md Simulations of Mistic: Conformational Stability in Detergent Micelles and Water. *Biochemistry* **2006**. 45(30): p. 9053-9058.
207. Aponte-Santamaria, C., et al., Molecular Driving Forces Defining Lipid Positions around Aquaporin-0. *Proc Natl Acad Sci U S A* **2012**. 109(25): p. 9887-9892.

208. vanderSpoel, D. and H.J.C. Berendsen, Molecular Dynamics Simulations of Leu-Enkephalin in Water and Dmso. *Biophysical Journal* **1997**. 72(5): p. 2032-2041.
209. Edidin, M., The State of Lipid Rafts: From Model Membranes to Cells. *Annual Review of Biophysics and Biomolecular Structure* **2003**. 32: p. 257-283.
210. Simons, K. and W.L.C. Vaz, Model Systems, Lipid Rafts, and Cell Membranes. *Annual Review of Biophysics and Biomolecular Structure* **2004**. 33: p. 269-295.
211. Kouril, R., et al., Structural Characterization of a Complex of Photosystem I and Light-Harvesting Complex Ii of Arabidopsis Thaliana. *Biochemistry* **2005**. 44(33): p. 10935-10940.
212. Amunts, A. and N. Nelson, Plant Photosystem I Design in the Light of Evolution. *Structure* **2009**. 17(5): p. 637-650.
213. Fromme, P., et al., Structure and Function of Photosystem I: Interaction with Its Soluble Electron Carriers and External Antenna Systems. *Febs Letters* **2003**. 555(1): p. 40-44.
214. Lange, C., et al., Specific Roles of Protein-Phospholipid Interactions in the Yeast Cytochrome Bc(1) Complex Structure. *Embo Journal* **2001**. 20(23): p. 6591-6600.
215. Fyfe, P.K., et al., Disruption of a Specific Molecular Interaction with a Bound Lipid Affects the Thermal Stability of the Purple Bacterial Reaction Centre. *Biochimica Et Biophysica Acta-Bioenergetics* **2004**. 1608(1): p. 11-22.
216. Pendley, S.S., et al., *Challenges and Development of a Multi-Scale Computational Model for Photosystem I Decoupled Energy Conversion*, in *Applications of Molecular Modeling to Challenges in Clean Energy*, G. Fitzgerald and N. Govind, Editors. 2013. p. 177-202.
217. Fischer, N., P. Setif, and J.D. Rochaix, Site-Directed Mutagenesis of the Psac Subunit of Photosystem I - F-B Is the Cluster Interacting with Soluble Ferredoxin. *Journal of Biological Chemistry* **1999**. 274(33): p. 23333-23340.
218. Sommer, F., F. Drepper, and M. Hippler, The Luminal Helix L of Psab Is Essential for Recognition of Plastocyanin or Cytochrome C(6) and Fast Electron Transfer to Photosystem I in Chlamydomonas Reinhardtii. *Journal of Biological Chemistry* **2002**. 277(8): p. 6573-6581.
219. Tsukiji, S. and T. Nagamune, Sortase-Mediated Ligation: A Gift from Gram-Positive Bacteria to Protein Engineering. *Chembiochem* **2009**. 10(5): p. 787-798.
220. Mao, H.Y., et al., Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *J Am Chem Soc* **2004**. 126(9): p. 2670-2671.
221. Parthasarathy, R., S. Subramanian, and E.T. Boder, Sortase a as a Novel Molecular "Stapler" for Sequence-Specific Protein Conjugation. *Bioconjugate Chemistry* **2007**. 18(2): p. 469-476.

222. Schwartz, E., et al., Complete Nucleotide Sequence of Phg1: A Ralstonia Eutropha H16 Megaplasmid Encoding Key Enzymes of H2-Based Lithoautotrophy and Anaerobiosis*. Journal of Molecular Biology* **2003**. 332(2): p. 369-383.
223. Massanz, C., V.M. Fernandez, and B. Friedrich, C-Terminal Extension of the H-2-Activating Subunit, Hoxh, Directs Maturation of the Nad-Reducing Hydrogenase in Alcaligenes Eutrophus*. European Journal of Biochemistry* **1997**. 245(2): p. 441-448.
224. Iwuchukwu, I.J., B. Bruce, and P. Frymier, *Protein Engineering for the Enhanced Photo-Production of Hydrogen by Cyanobacterial Photosystem I: A Ph.D. Thesis*, in *Chemical and Biomolecular Engineering*. 2011, University of Tennessee - Knoxville. p. 1-208.
225. Kaneko, T., et al., Sequence Analysis of the Genome of the Unicellular Cyanobacterium Synechocystis Sp. Strain Pcc6803. Ii. Sequence Determination of the Entire Genome and Assignment of Potential Protein-Coding Regions*. DNA research : an international journal for rapid publication of reports on genes and genomes* **1996**. 3(3): p. 109-136.
226. Le, R.K., Raeeszadeh-Sarmazdeh, M., Boder, E.T., and Frymier, P.D., Sortase-Mediated Ligation of Psae-Modified Photosystem I from Synechocystis Sp. Pcc 6803 to a Conductive Surface for Enhanced Photocurrent Production on a Gold Electrode*. Langmuir* **2014**: p. Manuscript submitted.
227. Simon, R., U. Priefer, and A. Puhler, A Broad Host Range Mobilization System for Invivo Genetic-Engineering - Transposon Mutagenesis in Gram-Negative Bacteria*. Bio-Technology* **1983**. 1(9): p. 784-791.
228. Yanischperron, C., J. Vieira, and J. Messing, Improved M13 Phage Cloning Vectors and Host Strains - Nucleotide-Sequences of the M13mp18 and Puc19 Vectors*. Gene* **1985**. 33(1): p. 103-119.
229. Lenz, O. and B. Friedrich, A Novel Multicomponent Regulatory System Mediates H-2 Sensing in Alcaligenes Eutrophus*. Proc Natl Acad Sci U S A* **1998**. 95(21): p. 12474-12479.
230. Hogrefe, C. and B. Friedrich, Isolation and Characterization of Megaplasmid DNA from Lithoautotrophic Bacteria*. Plasmid* **1984**. 12(3): p. 161-169.
231. Weiner, J.H., et al., A Novel and Ubiquitous System for Membrane Targeting and Secretion of Cofactor-Containing Proteins*. Cell* **1998**. 93(1): p. 93-101.
232. Goris, T., et al., A Unique Iron-Sulfur Cluster Is Crucial for Oxygen Tolerance of a [Nife]-Hydrogenase*. Nat Chem Biol* **2011**. 7(5): p. 310-318.
233. Wisitruangsakul, N., et al., Monitoring Catalysis of the Membrane-Bound Hydrogenase from Ralstonia Eutropha H16 by Surface-Enhanced Ir Absorption Spectroscopy*. Angewandte Chemie-International Edition* **2009**. 48(3): p. 611-613.

234. Schink, B. and H.G. Schlegel, Membrane-Bound Hydrogenase of Alcaligenes-Eutrophus .1. Solubilization, Purification, and Biochemical-Properties. *Biochimica Et Biophysica Acta* **1979**. 567(2): p. 315-324.
235. Peters, J.W., et al., X-Ray Crystal Structure of the Fe-Only Hydrogenase (Cpl) from Clostridium Pasteurianum to 1.8 Angstrom Resolution. *Science* **1998**. 282(5395): p. 1853-1858.
236. Fiser, A. and A. Sali, *Modeller: Generation and Refinement of Homology-Based Protein Structure Models*, in *Macromolecular Crystallography, Pt D*, C.W. Carter and R.M. Sweet, Editors. 2003, Elsevier Academic Press Inc: San Diego. p. 461-+.
237. Chang, C.H. and K. Kim, Density Functional Theory Calculation of Bonding and Charge Parameters for Molecular Dynamics Studies on Fefe Hydrogenases. *Journal of Chemical Theory and Computation* **2009**. 5(4): p. 1137-1145.
238. Burgdorf, T., et al., [Nife]-Hydrogenases of Ralstonia Eutropha H16: Modular Enzymes for Oxygen-Tolerant Biological Hydrogen Oxidation. *J Mol Microbiol Biotechnol* **2005**. 10(2-4): p. 181-196.
239. Lenz, O., et al., Requirements for Heterologous Production of a Complex Metalloenzyme: The Membrane-Bound [Nife] Hydrogenase. *J Bacteriol* **2005**. 187(18): p. 6590-6595.
240. Gay, P., et al., Positive Selection Procedure for Entrapment of Insertion-Sequence Elements in Gram-Negative Bacteria. *Journal of Bacteriology* **1985**. 164(2): p. 918-921.
241. Bernhard, M., et al., Functional and Structural Role of the Cytochrome B Subunit of the Membrane-Bound Hydrogenase Complex of Alcaligenes Eutrophus H16. *European Journal of Biochemistry* **1997**. 248(1): p. 179-186.
242. Sargent, F., et al., Overlapping Functions of Components of a Bacterial Sec-Independent Protein Export Pathway. *Embo Journal* **1998**. 17(13): p. 3640-3650.
243. Bernhard, M., et al., The Alcaligenes Eutrophus Membrane-Bound Hydrogenase Gene Locus Encodes Functions Involved in Maturation and Electron Transport Coupling. *Journal of Bacteriology* **1996**. 178(15): p. 4522-4529.
244. Frielingsdorf, S. and R.B. Klosgen, Prerequisites for Terminal Processing of Thylakoidal Tat Substrates. *Journal of Biological Chemistry* **2007**. 282(33): p. 24455-24462.
245. Teixeira, V.H., A.M. Baptista, and C.M. Soares, Pathways of H-2 toward the Active Site of Nife -Hydrogenase. *Biophysical Journal* **2006**. 91(6): p. 2035-2045.
246. Ichiye, T. and M. Karplus, Collective Motions in Proteins - a Covariance Analysis of Atomic Fluctuations in Molecular-Dynamics and Normal Mode Simulations. *Proteins-Structure Function and Genetics* **1991**. 11(3): p. 205-217.

247. Amadei, A., A.B.M. Linssen, and H.J.C. Berendsen, Essential Dynamics of Proteins. *Proteins-Structure Function and Genetics* **1993**. 17(4): p. 412-425.
248. Garcia, A.E. and J.G. Harman, Simulations of Crp:(Camp)(2) in Noncrystalline Environments Show a Subunit Transition from the Open to the Closed Conformation. *Protein Science* **1996**. 5(1): p. 62-71.
249. Doruker, P., A.R. Atilgan, and I. Bahar, Dynamics of Proteins Predicted by Molecular Dynamics Simulations and Analytical Approaches: Application to Alpha-Amylase Inhibitor. *Proteins-Structure Function and Genetics* **2000**. 40(3): p. 512-524.
250. Henzler-Wildman, K.A., et al., Intrinsic Motions Along an Enzymatic Reaction Trajectory. *Nature* **2007**. 450(7171): p. 838-U813.
251. Schubert, T., et al., Chaperones Specific for the Membrane-Bound Nife - Hydrogenase Interact with the Tat Signal Peptide of the Small Subunit Precursor in Ralstonia Eutropha H16. *Molecular Microbiology* **2007**. 66(2): p. 453-467.
252. Lee, A.G., Lipid-Protein Interactions. *Biochemical Society Transactions* **2011**. 39: p. 761-766.
253. Gabdoulline, R.R. and R.C. Wade, Simulation of the Diffusional Association of Barnase and Barstar. *Biophysical Journal* **1997**. 72(5): p. 1917-1929.
254. Wade, R.C., et al., Electrostatic Steering and Ionic Tethering in Enzyme-Ligand Binding: Insights from Simulations. *Proceedings of the National Academy of Sciences of the United States of America* **1998**. 95(11): p. 5942-5949.
255. Elcock, A.H., et al., Computer Simulation of Protein-Protein Association Kinetics: Acetylcholinesterase-Fasciculin. *Journal of Molecular Biology* **1999**. 291(1): p. 149-162.
256. Gabdoulline, R.R. and R.C. Wade, Biomolecular Diffusional Association. *Current Opinion in Structural Biology* **2002**. 12(2): p. 204-213.

VITA

Bradley Jordan Harris was born in Morristown, TN on May 28th, 1986 to Tony and Tammy Harris. He graduated Summa Cum Laude from the University of Tennessee, Knoxville with a B.S. in Chemical Engineering as a member of the University Honors program in May 2008. From July 2008 to July 2009, he worked as a Chemical Engineer in the Chemicals Development Division at Eastman Chemical Company in Kingsport, TN. He returned to the University of Tennessee, Knoxville in January 2010 to pursue his Ph.D. in Chemical and Biomolecular Engineering. On October 26th, 2013, he married his wife, Kristen Harris.