UNIVERSITY OF CALIFORNIA
SANTA BARBARA

Investigations of the Structure-Function Relationship of Transmembrane Proteins by studies of Proteorhodopsin

A dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

by

Chung-Ta Han

Committee in Charge:

Professor Songi Han, Chair

Professor Michelle A. O'Malley

Professor Siddharth Dey

Professor Mark S. Sherwin

September 2020

The dissertation of Chung-Ta Han is approved:

Professor Michelle A. O’Malley

Professor Siddharth Dey

Professor Mark S. Sherwin

Professor Songi Han, Chair

August 2020

Investigations of the Structure-Function Relationship of Transmembrane Proteins by Studies of Proteorhodopsin

ACKNOWLEDGEMENTS

I feel thankful to all the people that have directly or indirectly helped me to get through the journey of Ph.D. Without those supports, I cannot be the same person as I am now. This is a period that I have learned a lot and become more mature not only professionally but also mentally. There are a lot of people that I need to thank, although it is hard to identify every single person that has helped me and some of those may have lost contact with throughout the past years but still want to thank. First, I would like to thank my advisor, Songi Han. Songi is an incredible advisor that always supports us to explore different research directions. When I faced any challenges in my research, she always can find a way to guide me to success, but at the same time not giving me free answers right away. Therefore, I can learn from the struggle and become a better researcher than I was before. She is a very talented advisor that can always build up a clear roadmap to glory and lead the whole research group to achieve unbelievable goals. Other than the professional aspect, she is also a very fun supervisor to work with. Throughout the past years, she has established a very good vibe in the research group and can let the members in her group to be closely connected. There are myriads fun memories in the group beyond those professional discussions and meetings that we have been through, including but not limited to the pool party at Songi's house, the awesome food road trip to Los Angeles as a whole research group, the ski trip to the Big Bear Resort, and countless potluck and beach barbeque memories. These all make my Ph.D. experience unique compared to others.

I thank all the people that first met when I started my research here at UC Santa Barbara, including those Chemical Engineering classmates that started in the same year. You all give me as a foreigner a friendly environment to start with. Specifically, Dr. Neil Eschmann and Dr. Nikki Schonenbach (now they are Mr. and Mrs. Eschmann) are the two senior colleagues that later become friends of mine. They are two friendly people that not only accommodated me when I first joined the group and shared with me lots of American cultures that I didn't know before I came to the United States. When I haven't bought my car here in Santa Barbara, they offered me countless free rides around the town so that I can participate a lot of social events. Dr. Matt Idso, who worked closely with me on the PR material project, is another graduate student that I want to thank. Matt is enthusiastic about conducting scientific research. I still cannot forget there was one time we went to the Test Pilot in the funk zone of downtown Santa Barbara together for drinks. It was a great time and he just talked to a random citizen in the town about his cool scientific ideas. These senior members helped me to have a great starting point early in my Ph.D.

I want to give a big shout out to all the undergraduate students that have worked with me on the PR project in different periods. Evelyn Chang is the first undergraduate that I worked with. She transferred lots of biochemistry knowledge to me when I first joined the group without a senior graduate student as a mentor. Sirish Narayanan and Tristan Chan are the following two cool people that I worked with. We hung out and had great times together in lots of pint night in Woodstock's Pizza. We also introduced the golden era of "Shot Friday" in the lab. Professionally, they both are great researchers and now are pursuing their Ph.D. careers. Khrysta Pruett and Carl Song are the two next-generation undergrads that are always willing to dedicate time into researching despite their busy schedule during their senior year.

They are also cool people to hang out with and are active in joining lab social events. Elsa Winslow and Annie Chang are the two final generation undergraduates that helped me a lot with wrapping up my research in my 5th year. They also have a strong motivation in conducting research. Despite the unusual campus shutdown time, they are still willing to discuss fundamental knowledge related to magnetic resonance spectroscopy with me and I also have learned a lot from our discussion. Eventually, I must thank all of them for being my “PR protein factory” and “pKa spectral titration robot” so that I can get this dissertation done.

In the professional aspect, I want to thank all my thesis committee: Prof. Songi Han, Prof. Michelle O’Malley, Prof. Siddharth Dey, and Prof. Mark Sherwin. You all have provided lots of great advice in my candidacy exam and annual reviews so that I can have greater confidence in my research plan. The guidance from you all also helped me to sharpen my presentation skills. Mark is also a close collaborator that worked directly with me on the Gd distance measuring project, which will later turn into an awesome work that can film protein structural dynamics. I also want to thank Prof. Brad Chmelka as another close collaborator on the PR material research. He offered me an opportunity to step out of the biophysics and biochemistry world to extend my knowledge to the material science field. With all the meetings that we have been through, I learned from him that a great researcher needs to be well-prepared in all different aspects of the field and should not limit myself. I also appreciate other professors that I have discussed research with through collaboration or met in person in conferences, including but not limited to Prof. Alex Smirnov, Prof. Janet Lovett, Prof. Arnab Mukherjee, Prof. Blake Mertz, and Prof. Klaus-Dieter Kreuer. I want to thank Dr. Jerry (Jingui) Hu, Dr. Asif Equbal, Dr. Sheetal Jain and Miranda (Yuanxin) Li for all the help on NMR experiments and thank Dr. Timothy Keller, Dr. Blake Wilson, Dr. Jessical Clayton, Dr. Marzieh Kavand and now Prof. Yann Fichou for all the discussions and helps on EPR experiments. All the other Han lab members that aren’t listed above are appreciated as well. You all are awesome people to work with. I especially hope that Shiny Maity, Max Berkow, and Brad Price can continue my work and in the end finish more than I can foresee now.

I also want to thank my Taiwanese friends that have shared the time with me in Santa Barbara or have visited me disregard the distance. You all have provided me the connections outside the local communities. Without all you folks, life would be different in the past five years. I list some of these friends in no specific order: Dr. Chih-Cheng Chang, Kangching Fang, Dr. Szu-Ying Chen, Yen-Ti Chen, Shao-Ting Chen, Yung-Chung Hsiao, Dr. Kaiwen Hsiao, Dr. Chu (Marco) Lam, Sheng-Ping Liang, Wei-Tsung Lin, Dr. Po-Yu Peng, Emmy Lin, Cheng-Jung Kuo, and Dr. Shu-Kai Hu. On the other hand, I want to thank Steve Ota, a ballroom dancing teacher at both UCSB and Dojo in Goleta. You offered me lots of chances to sharpen my dancing skills and had a great getaway from my daily research routine. I also want to thank Yi-An Lee, Georgia Pollard, and Miye Ota for all the dancing techniques instructed and all the people that I have danced with on campus and in the local community. Dancing is a great way to get me re-fueled so that I can accomplish more works.

Last but most importantly, I want to thank my parents for always being so supportive. You both are number one mom and dad that are willing to accompany me and helped me to get through those lows during my graduate school. It is great to have you both here visited me at the end of 2019 so that I can show you both how my life at UCSB is like. I want to thank my girlfriend Chia-Ju Lin for all the support during the time while I am busy writing this dissertation and finding the next career to pursue. We will together and get through the next challenges that I will face in the future.

Curriculum Vitae

Chung-Ta Han

EDUCATION

2020	Ph.D. in Chemical Engineering, University of California, Santa Barbara, California
2014	M.S. in Chemical Engineering, National Taiwan University, Taipei, Taiwan
2012	B.S. in Chemical Engineering, National Taiwan University, Taipei, Taiwan

PROFESSIONAL EXPERIENCES

2016 – 2020	Graduate Student Researcher, UCSB, Santa Barbara, California
2012 – 2014	Graduate Student Researcher, National Taiwan University, Taipei, Taiwan
2011 Jul-Aug	Summer Intern, CTCI Corporation, Taipei, Taiwan

PUBLICATIONS AND PATENT APPLICATION

- C.-T. Han, J. Song, C. Pruett, Tristan Chan, S. Han, “Electrostatic Environment of Proteorhodopsin Affects the pKa of Its Buried Primary Proton Acceptor”, Biophys. J. (2020)
- C.Y. Hong, C.-T. Han, L. Chao, “Nonspecific Binding Domains in Lipid Membranes Induced by Phospholipase A2”, Langmuir (2016)
- C.-T. Han, L. Chao, “Using Patterned Grating Structure to Create Lipid Bilayer Platforms Insensitive to Air Bubbles”, Lab on a Chip (2015)
- C.-T. Han, L. Chao, “Creating Air-Stable Supported Lipid Bilayers by Physical Confinement Induced by Phospholipase A2”, ACS Applied Materials & Interfaces (2014)
- C.-T. Han, L. Chao, “Using physical confinement to create lipid bilayer biosensing platforms insensitive to air-water interface”, Taiwan Patent No. 201614075

MANUSCRIPTS IN PREPARATION

- C.-T. Han, S. Hussain, M. Kinnebrew, M.N. Idso, N. Baxter, E. Chang, E. Aye, S. Han, “Role of the lipid membrane in the oligomeric assembly and function of Proteorhodopsin”
- C.-T. Han, Y, Li, C. Pruett, E. Winslow, S. Han, “Internal electrostatic network of proteorhodopsin correlated with its activity modulation revealed by DNP-enhanced MAS NMR”

PRESENTATIONS

Feb. 2020	Poster presentation, Biophysical Society 64th Annual Meeting, San Diego, CA, U.S.A., “Electrostatic environment of Proteorhodopsin affects the pKa of its buried primary proton acceptor”, Han, C.-T., Song, J., Chan, T., Pruett, C., , Han, S.

Oct. 2019	Oral presentation, 12th Annual Amgen-Clorox Graduate Student Symposium, Santa Barbara, CA, U.S.A., “Electrostatic environment of Proteorhodopsin tunes the pKa of its primary proton acceptor”, Han, C.-T., Han, S.
Apr. 2019	Oral presentation, 2019 UC Santa Barbara Glad Slam Competition, Santa Barabra, CA, U.S.A., “A Generalized Approach to Studying Protein’s Function Through its Structural Movement”, Han, C.-T.
Sep. 2018	Poster presentation, 18th International Conference on Retinal Proteins, Ontario, Canada, “Role of the Oligomer Assembly and Environments on Modulation Proteorhodopsin Function”, Han, C.-T., Song, J., Pruett, K., Hussain, S., Idso, M.N., Han, S.
Feb. 2018	Poster presentation, Biophysical Society 62nd Annual Meeting, San Francisco, CA, U.S.A., “Role of Lipid Membrane on the Oligomeric Assembly and Function of Proteorhodopsin”, Han, C.-T., Hussain, S., Idso, M.N., Narayanan, S., Chan, T., Han, S.
Nov. 2013	Oral presentation, Taiwan Chemical Engineering Annual Meeting, Taipei, Taiwan, “Using Phospholipase A2 to Create Air-stable Supported Lipid Bilayers”, Han, C.-T., Chao, L.
Nov. 2013	Oral presentation, 2013 AIChE Annual Meeting, San Francisco, U.S.A., “Using Phospholipase A2 to Create Air-stable Supported Lipid Bilayers”, Han, C.-T., Chao, L.

AWARDS

2013	Excellent Award in graduate student English oral presentation competition, Taiwan Chemical Engineering Annual Meeting, Taipei, Taiwan
2013	4th LCY Chemical Corp. Scholarship for outstanding students, Kaohsiung, Taiwan

TEACHING AND MENTORING

2016-2020	- Mentored total 7 undergraduate students: S. Narayanan (Chemistry and Biochemistry), T. Chan (Chemistry and Biochemistry), K. Pruett (Chemical Engineering), J. Song (Chemical Engineering), N. Atam (Chemistry and Biochemistry), E. Winslow (Chemistry and Biochemistry), and A. Chang (Physics) - Trained and guided undergraduate researchers to accomplish research projects by establishing achievable goals and leading weekly meetings.
2016-2019	Teaching Assistant (TA) - TA for Chemical Engineering Lab (2016, 2019) - TA for Chemical Reaction Engineering (2018) - Reader for Advanced Chemical Reaction Engineering (2017)

ABSTRACT

Investigations of the Structure-Function Relationship of Transmembrane Proteins by Studies of Proteorhodopsin

Chung-Ta Han

Transmembrane proteins are essential biomolecules that reside in cell membranes and play important roles in biological functions including ion transport, signal transductions, and enzymatic reactions. These proteins are surrounded by a highly heterogeneous environment. They have hydrophilic domains that are exposed to water molecules and solutes in solution and hydrophobic membrane-spanning domains that are embedded within a lipophilic environment. This unique combination of environment around the transmembrane protein is not only required to preserve a proper folding of its structure but can also take part in regulating its functional activity. Due to the complexity of the environment, biophysicists are still developing approaches to elucidate the role of different factors in modulating the function of transmembrane proteins.

This dissertation aims to investigate the influence of environmental factors on the structure and function of transmembrane proteins by studies of Proteorhodopsin (PR), a light-activated proton pump that originated from marine bacteria. PR is an excellent model for studying transmembrane protein function as its proton transport function can be directly assessed by optical absorbance spectroscopic techniques. PR also shares the same seven-helical transmembrane (7TM) structure with a large family of human receptors, which makes the structural knowledge learned from PR studies potentially transferable to other physiologically

important transmembrane proteins (e.g. G protein-coupled receptors). On the other hand, PR has excellent robustness and is found to be stable in a variety of environments. This allows it to be studied in environments not limited to native-like bacteria membrane but also other biomimetic platforms with different factors that can be independently controlled. PR also has its significance in protein engineering applications. The proton motive force generated by PR through its vectorial proton transport can be utilized as an approach to harvest solar energy. All these together make the studies of PR appeal to objectives of both fundamental biophysical understanding and bioengineering applications on transmembrane proteins.

To achieve our objectives, the functions of PR under the influences of different modulators are examined in biomimetic environments including synthetic liposomes, lipid nanodiscs, and synthetic host materials. With controls over the surrounding lipid composition and the oligomeric distribution of PR between its monomeric and oligomeric forms, we identify that protein-protein interaction (i.e. the formation of oligomers) is a key factor to determine the proton transport kinetic of PR. Nevertheless, in the lipid membrane environment, the protein-protein interaction shows no impact on the protonation behavior of PR’s embedded ionizable amino acid D97, a switch that controls the population of active PR with proton transport capability. Instead, we find that the electrostatic environment around PR (e.g. the ion type or concentration in the buffer and the net charge of lipid headgroups) can significantly modulate the pKa of this embedded D97 (pKa_{D97}), and in turn affects the population of active PR. Most importantly, the same concept is found to be transferable to PR reconstituted in synthetic host materials with increased thermal and mechanical stability for bioengineering applications.

To further understand the correlation between the structure and function of transmembrane proteins, complementary magnetic resonance spectroscopic tools were used to acquire the structure of PR with different functional outcomes. This dissertation interrogates the structural rearrangement of PR associated with the pKa_{D97} modulation by different electrostatic environments. Local conformation changes at the third intracellular loop of PR are uncovered by using electron paramagnetic resonance (EPR) analysis and Overhauser dynamic nuclear polarization (ODNP) relaxometry through a change of local environment experienced by a site-directly introduced spin label on the loop. Moreover, by conducting dynamic nuclear polarization (DNP)-enhanced solid-state nuclear magnetic resonance (ssNMR) measurements on PR either site-specifically or uniformly labeled with NMR active nuclei, we identify the residues that rearranged while PR is in different electrostatic environments. On the other hand, to have a structural-based understanding of the functional impact of oligomerization, this dissertation also seeks to develop a novel approach to map out the structural rearrangement of PR activation by high field EPR spectroscopy with Gadolinium-based spin labels. The study of PR provides an opportunity of directly correlating the measured structural dynamics with the kinetics of proton transport function measured by its transient absorbance change.

Altogether, PR studies presented here elucidate the roles of oligomerization in modulating its proton transport function and underscore the importance of the biomimetic environment on affecting the function of transmembrane proteins. For bioengineering and biomedical applications relying on transmembrane protein systems, this dissertation offers a guideline for optimizing functions of proteins that are controlled by a similar mechanism as PR. Furthermore, the change of structural properties observed here by magnetic resonance tools from PR with different functional outcomes adds value to biophysical understandings on the relationship between structure and function of transmembrane proteins that share a similar structure as PR.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Membrane Proteins

All beings on the Earth are composed of a myriad of biomolecules. These biomolecules have to work in a collaborative effort to sustain different physiological functions. Among all these essential molecules, transmembrane proteins are the ones that reside in a highly heterogeneous cell plasma membrane environment. As the cell membrane acts as a barrier to separate the interior of the cell from the outside environments, transmembrane proteins can play important roles in functions including ion transport,[1-2] redox reaction,[3] sensoring,[4] and catalysis[5] (Figure 1.1). These proteins, composed of twenty amino acids, have both hydrophobic transmembrane segments that must be embedded in assemblies of amphiphilic lipid molecules and hydrophilic pore and surfaces that are exposed to water and solvents. Most of the transmembrane protein functions can be regulated and have different levels of activity. The activity of functions can be affected by interactions between the transmembrane protein and surrounding environments including lipid molecules[6-11] and structural water coupled on protein surfaces.[12-13] Moreover, a formation of higher-order oligomeric assemblies that introduce interactions between transmembrane proteins can also affect their functional outcome. For example, G-protein coupled receptors (GPCRs) are a big family of transmembrane proteins that are responsible for triggering cascades of different signaling pathways in cells and are common drug targets (~36 % of approved drugs target GPCRs in 2017).[14] The functions of GPCRs can be regulated by formations of homo- or hetero-oligomeric complexes[15-21] or interactions with surrounding lipids.[22] As a result, the effectiveness of developed drug candidates on regulating GPCR signal transductions to cure or mitigate diseases can significantly be affected by these other factors originally present in the biological system. Therefore, understanding the mechanistic basis of transmembrane protein functional regulations by potential modulators is important for both physiological understanding and for drug developments.

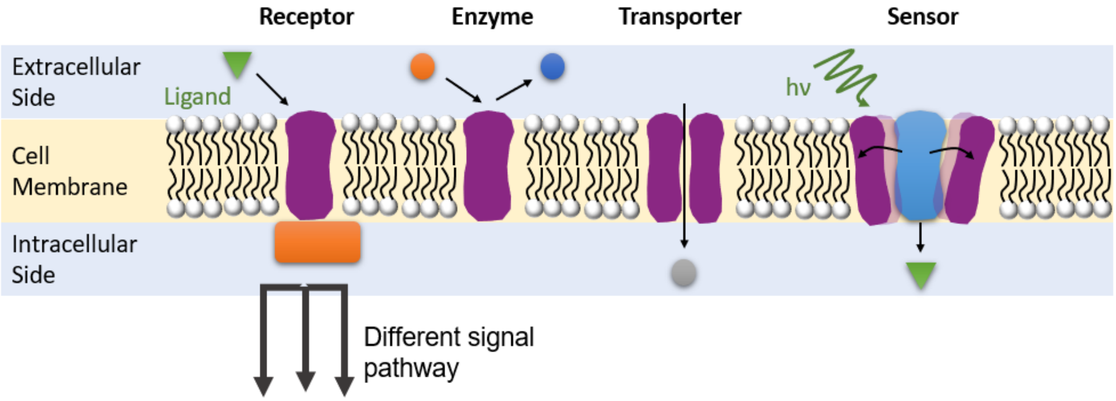


Figure 1.1: Membrane Proteins. Membrane proteins reside in the cell plasma membrane and can function as receptors, enzymes, transporters, and sensors. In some cases, the functions of proteins can be triggered only when they associate with each other to form oligomers.

The knowledge on transmembrane protein functional modulation is also beneficial for bioengineering objectives, as an increasing number of biotechnology applications utilize the functions of transmembrane proteins in biosensing,[23] catalysis,[24] bioanalytics,[25] or energy conversion[26] applications. In these applications, transmembrane proteins are reconstituted into synthetic host materials that can maintain their native structure and function with increased thermal and mechanical stability. While most of these host material designs can preserve the native-like function of transmembrane proteins, a fundamental understanding of how environmental factors modulate transmembrane protein functions can further provide a general guideline on the design of host materials to increase the activity of incorporated transmembrane proteins.

Nevertheless, the study of transmembrane protein functional modulations is experimentally challenging, especially for human receptors. The challenges of transmembrane protein functional studies arise from their structural instability while the proteins are not in the native *in vivo* environment, which excludes the application of traditional *in vitro* characterization tools. On the other hand, the read-out of the transmembrane protein function usually relies on functional assay kits, and these assays are not necessarily available for the protein system of interest or could be experimentally challenging for obtaining unambiguous results. Direct *in vivo* studies on transmembrane proteins is also challenging, as these proteins are often present at a low population in the cell plasma membrane, and the complexity of different signal pathways in cells makes it extremely hard to separate the effect of interest from all different modulators. To overcome these challenges, a model transmembrane protein with excellent structural robustness and accessible functions would be required. Together with technological and methodological innovations on resolving the structure and function relationship of transmembrane proteins, the knowledge gain from the model transmembrane protein system could potentially apply to other systems that either share a similar function or structural motif.

The overarching goal of this dissertation is to elucidate the functional impact of protein oligomerization and biomimetic host environment on modulating the function of transmembrane proteins. The approach here is by varying potential functional modulators one at a time to see how the activity of a model transmembrane protein, proteorhodopsin (PR), changes while it is reconstituted in different biomimetic environments. To get a fundamental understanding of the observed functional tuning, we investigate the structure of our model transmembrane protein through complementary magnetic resonance spectroscopic techniques. We also aim to develop a novel methodology that can capture the structural dynamic of transmembrane proteins during their activation, while using our model protein system to test the developing tools. In the end, we examine whether the knowledge of tuning transmembrane protein functions learned in one biomimetic environment can be transferred to others. Overall, the outcome of this dissertation can assist in establishing a connection between protein function, structure, and surrounding environmental modulators. The following introduction will provide an overview of the potential modulators that could affect the structure or function of transmembrane proteins, justify the use of the light-activated proton pump PR as a model system in this dissertation, and introduce magnetic resonance spectroscopic tools that can characterize the structure of transmembrane proteins.

1.1.1 Protein-Protein Interactions

The oligomerization of transmembrane proteins is proposed to be an important factor to regulate their function. The regulation is sought to be achieved by the formation of specific cross-protomer interactions, where these interactions can at the same time enhance protein stability.[27] For example, it is known that the activation of GPCRs are generally coupled with an outward movement of the sixth transmembrane helix to enable G protein coupling.[28-29] The formation of GPCR oligomers was hypothesized to hinder or promote this conformation change and lead to different functional consequences or different cell signaling pathways.[30-32] However, studying the actual functional impact of oligomer formation is experimentally challenging. Even if the functional outcome of transmembrane proteins can be directly accessed, in most cases, the protein monomer can be functional by itself and delineating the functional difference between oligomers and monomers is challenging. On the other hand, since the interactions between proteins in oligomers are generally weak, oligomers formed in the native environment can easily dissociate after proteins are incorporated into different biomimetic environments. As a result, current knowledge on the functional impact of transmembrane protein oligomers formation is limited.[33]

On the other hand, evidence has shown that the surrounding environment in which transmembrane proteins reside can affect their oligomeric distribution between oligomers and monomers. Both lipid bilayer curvature and hydrophobic mismatch between transmembrane proteins and lipids can drive the formation of rhodopsin oligomers.[34] Functional differences were also found from rhodopsins in the tested lipid membrane environments with different oligomeric distributions. Nevertheless, it is still unclear whether the observed functional modulation is solely caused by protein oligomerization, or the difference in lipid bilayer stress exerted on the protein could have a role in modulating protein function. A study that can distinguish the effect of protein-lipid interactions from protein-protein interactions on modulating protein functions will be beneficial for understanding the role of each factor.

1.1.2 Protein-Lipid Interactions

Increasing evidence suggests a lipid membrane is not only a substrate to host transmembrane proteins but also can actively take part in regulating the activities of protein functions. In the 1970s, S.L. Singer and G.L. Nicolson proposed the fluidic mosaic model to describe the two-dimensional cell plasma membrane composed of proteins, phospholipids, cholesterols, and carbohydrates.[35] In the model, these molecules in the cell membrane constantly moving in a fluidic like fashion and can freely associate with and dissociate from other molecules in the lipid bilayer. Later studies found the distribution of the lipid molecules and transmembrane proteins in the cell plasma membrane is not random and symmetric, but certain lipid species including sphingomyelin and cholesterol can associate with transmembrane proteins to form subdomains called "lipid raft". The formation of lipid raft is proposed to regulate the signal transduction pathway of certain human receptors. Nevertheless, the actual role of membrane protein interactions in modulating the functions of transmembrane proteins is still unclear, as the formation and dissociation of these raft domains can be at the rate of ~100 ns timescale.[36] On the other hand, like the challenge in studying the functional role of protein oligomeric assemblies, lipid-protein interactions are also generally weak and can be disrupted by characterization techniques.[37-38] Overall, the role of complex protein-lipid interactions on modulating protein activity is still under discovery.

There are two main approaches to explain the role of protein-lipid interactions on modulating protein functions. One approach focuses on finding the specific binding domains on the protein for lipids or cholesterol molecules to regulate its structure and function.[39-40] For example, cholesterol in the cell plasma membrane could chemically bind on the M2 protein of influenza A virus and induces conformation changes on the amphipathic helix of the protein to stabilize viral filament formation.[41] The other approach focuses more on the effect of mesoscopic material properties of the lipid membrane on changing the energy landscape of protein conformations.[42] For example, the formation of a reverse hexagonal phase by certain non-bilayer-forming lipids that exist in the retinal rod membrane can affect the conformation transition of the human rhodopsin between its active and inactive form.[6] Different from the fluid mosaic model, this flexible surface model describes how the work from protein-lipid interactions affects the conformational changes of transmembrane proteins during its activation process, and the extent of this work is highly dependent on material properties including membrane curvature elastics and the hydrophobic matching at the lipid-protein interface. In general, these two approaches both suggest that the lipid membrane is not only a solvent to host functional transmembrane proteins but itself could play an active role in initiating or regulating certain cellular pathways. This dissertation focuses on altering the charisteristics of the lipid membrane including lipid curvature, water pore formation probability, and the net charge of lipid headgroups to study the effect of these factors on modulating transmembrane protein functions.

1.1.3 Biomimetic Systems

For *in vitro* biophysical studies, biomimetic systems are required to reproduce the features of the cell plasma membrane to preserve the native structure and function of transmembrane proteins. Nevertheless, the oligomeric distribution of transmembrane proteins could vary in different environments and the lipid molecules could affect protein functions. Therefore, understanding the impact of biomimetic host systems on the function of transmembrane proteins is reasonably important, but often not carefully examined in most biophysical studies.

Figure 1.2 summarizes the usual biomimetic systems used for transmembrane protein structural and functional studies, and these systems are used in this dissertation. The most common platform for studying transmembrane protein is detergent micellar environments. This is done by directly extracting transmembrane proteins from the cell plasma membrane or the cell system used for overexpressing proteins. Nevertheless, this spherical micellar environment cannot reproduce the planar property of the cell plasma membrane, including the fact that the micelles have different dynamics and lateral pressure profile.[13, 43] Therefore, the functional or structural understandings of transmembrane proteins reconstituted in detergent micellar environments sometimes cannot represent their natural behavior in the cell plasma membrane.

Vesicle-based liposome is another common biomimetic platform for transmembrane proteins with most of the cell plasma membrane characteristics. Since the components of synthetic lipids for constituting liposomes can be controlled, it enables systematic studies on the effect of different factors on the structure and function of transmembrane proteins. However, these vesicles are generally large (> 100 nm in diameter). This bigger size of vesicles can introduce undesired light scattering in optical spectroscopy studies and is not ideal for solution-based NMR studies that require faster tumbling of molecules. Another thriving biomimetic platform with lipid bilayer characteristics is the lipid nanodisc platform.[44] These nanodiscs are lipid bilayer patches constituted with transmembrane proteins, while the hydrophobic region of the lipid bilayer is protected by either amphipathic copolymers or membrane scaffold proteins. The smaller size of these nanodiscs (~ 10nm in diameter) makes them compatible with spectroscopy techniques for protein biophysical studies. The formation of nanodiscs can be either from synthetic liposomes with controlled lipid compositions or directly from cells without other intermediate purification steps that could perturb the structure of proteins. This characteristic makes the nanodisc a more versatile platform for different biophysical studies, including the thriving cryo-electron microscopy for resolving protein structures with atomic-resolution.

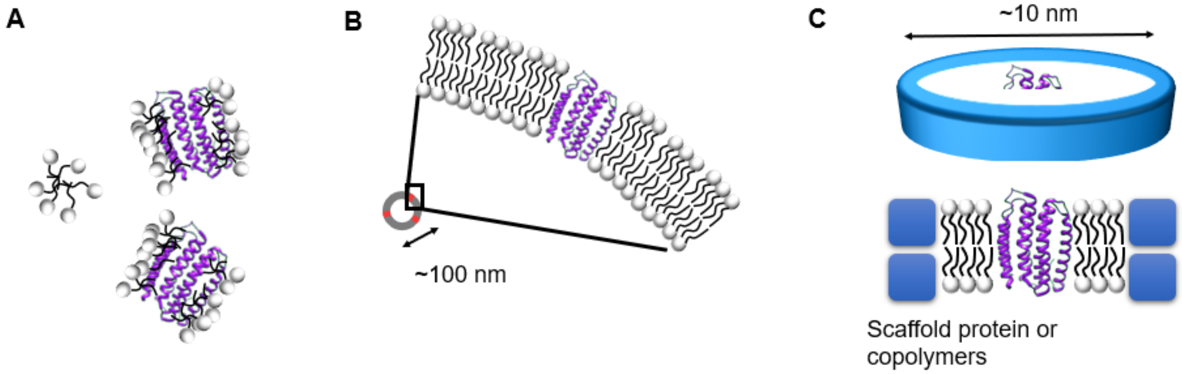


Figure 1.2: Biomimetic Systems. A. Detergent micelles, **B.** vesicle-based liposomes, and **C.** lipid nanodiscs are common biomimetic systems that can preserve the structure and function of transmembrane proteins in *in vitro* studies.

While numerous biomimetic platforms could at the same time preserve the native structure and function of transmembrane proteins, comparing the functional outcome of these transmembrane proteins in different systems can provide additional information on whether the knowledge learned in one system is valid or not. Even the common biomimetic platforms described above cannot reproduce all the detailed features of the cell plasma membrane (e.g. cytoskeleton, glycocalyx elements), as a result, it is important to select an appropriate biomimetic platform for the structural and functional study of transmembrane proteins. This dissertation focuses on studying the structure and function of our model transmembrane protein PR in vesicle-based liposome environments under the influence of different modulators, and at the same time examines whether the knowledge in this biomimetic environment is transferable to lipid nanodisc or mesostructured surfactant-silica film for engineering applications.

1.2 Proteorhodopsin

Proteorhodopsin (PR) is a light-activated proton transferring transmembrane protein that was observed in 2000 from the inner membrane of marine bacterioplankton residing in the Monterey Bay, California.[45] It is responsible for the solar regulation of function and survival of marine bacteria.[46-48] Upon light-activation, PR transfer proton from the intracellular to the extracellular side of the bacterial inner membrane to build up a proton motive force (PMF) across the membrane. The PMF is used to activate other proteins in the vicinity, including driving the ATP synthetase to produce ATP – a chemical energy source in cells and bacteria. PR is estimated to exist in up to 80% of oceanic bacteria, and such abundance makes ecologists believe that PR can play a significant role in energy balance on the Earth in harvesting solar energy.[46] The rhodopsin family that PR belongs to shares a similar seven transmembrane (7TM) structure and has a retinal chromophore connected to a lysine residue for triggering its light-activated function. Nevertheless, despite their structural similarity, the proteins in this family still have a versatile spectrum of functions as ion pumps or ion channels specialized for different ions. This makes PR and other transmembrane proteins in the rhodopsin family an interesting case study for understanding the structure-function relationship of transmembrane proteins with a similar 7TM structure.

With a tunable functionality, PR stands out from most of the other rhodopsins as a perfect model for protein biophysical studies.[49-50] Unlike another light-activated proton pump bacteriorhodopsin (bR) from halophilic archaea, the vectorial of PR transport is found to have pH-dependency, where the percentage of active PR with the same proton transport direction could be significantly altered by pH or other environmental modulators under the seawater ambient environment. On the other hand, PR has variants that yield different optimal absorbing wavelength range, which could be determined by a switch residue buried inside the protein to change the protein structure around the light-absorbing retinal chromophore. In the ocean, there are two main variants of PR with different maximum absorbing range. The green-light-absorbing PR (GPR) is populated in marine bacterioplankton that is closer to the seawater surface and the blue-light-absorbing variant (BPR) is instead mainly found in the deeper region. The proton transport function of PR is also found to be coupled with the conformation of intracellular loops that are far away from the proton pathway of PR. Interestingly, similar functional regulation behavior and conformational coupling can also be found from human G-protein coupled receptors (GPCRs) that play important roles in cell signal transduction pathway. These receptors are physiologically important,[51] but knowledge on their structure and function relationship is still elusive.[33] Direct biophysical studies on GPCRs are generally challenging because of their poor *in vitro* stability. Even though PR is evolutionarily and functionally unrelated to GPCRs, by sharing a similar 7TM structure and having its functional regulation behavior, PR studies still have a great potential to become stepping-stones toward functional studies of GPCRs. This makes the functional studies on PR in this dissertation have a broader impact on not only retinal proteins but also other more complex transmembrane proteins in mammalian or eukaryotic cell membranes.

1.2.1 The structural property of PR

The PR construct used in this dissertation is the “green-light-absorbing” GPR variant derived from the first PR gene sequences (BAC31A8).[45] The structure of GPR has not yet be resolved at the atomic level by x-ray crystallography or cryogenic electron microscopy (cryo-EM), but a structure resolved by solution NMR still provides coarse information on its secondary structure.[52] Taking advantage of the similarity between PR and the other well-studied photoactive proton pump bR, functionally important residues that are in charge of the proton transport process were identified, including a Schiff base K231 and corresponding proton donor E108 and proton acceptor D97 of the Schiff base (Figure 1.3 A). All other identified functional-relevant residues are highlighted on the snake diagram in Figure 1.3 B. A retinal chromophore is covalently linked to the Schiff base. While PR is illuminated by green light, the retinal chromophore undergoes an *all-trans* to *13-cis* isomerization. This event triggers relocation of the proton on the Schiff base K231 to the proton acceptor D97 and initiates a cascade of proton relocation processes.[53-55]

Despite sharing a similar proton transport function, there are some features in the PR proton transport pathway that are substantially different from the one of bR. First, PR does not have a proton release complex, where bR has this complex consisted of two glutamic acids and three structural water molecules to facilitate the release of protons to the extracellular side of the protein in its photochemical reaction cycle. The proton translocation pathway from the Schiff base K231 to the proton acceptor D97 of PR is also not as well-understood as the case of bR.[56] Moreover, PR has a conserved histidine H75 locates on the B helix that was found to interact with the proton acceptor D97. The interaction between H75 and D97 affects the protonation behavior of D97 and in turn determines the population of active PR that can transport proton in the same direction.[49, 57-59] This interaction also has been suggested to slow down the proton transport kinetics of PR, even not as an essential interaction for the proton transport to occur.[49] The third intracellular loop (E-F loop) is also another interesting feature in the structure of PR. Magnetic resonance spectroscopy studies have unveiled a twisting and tilting movement at this α-helical E-F loop when PR is photoactivated.[52, 60-62] Despite being a long distance away from the proton transport pathway, a mutation at site 178 on the E-F loop has a long-range effect on both the absorbance of the retinal chromophore and the protonation behavior of the D97 residue.[61-64] This evidence demonstrates that this unique motif of PR has its functional significance in PR proton transport function, and the secondary structure of E-F loop could be a structural constraint to affect the conformation of the whole PR.[62] L105 is another site that is closer to the retinal binding pocket and can affect the light absorbance range of GPR. By mutating this residue into glutamine (Q), GPR shifted its absorption toward shorter wavelengths to be more like its blue-light absorbing BPR variant.[65-66]

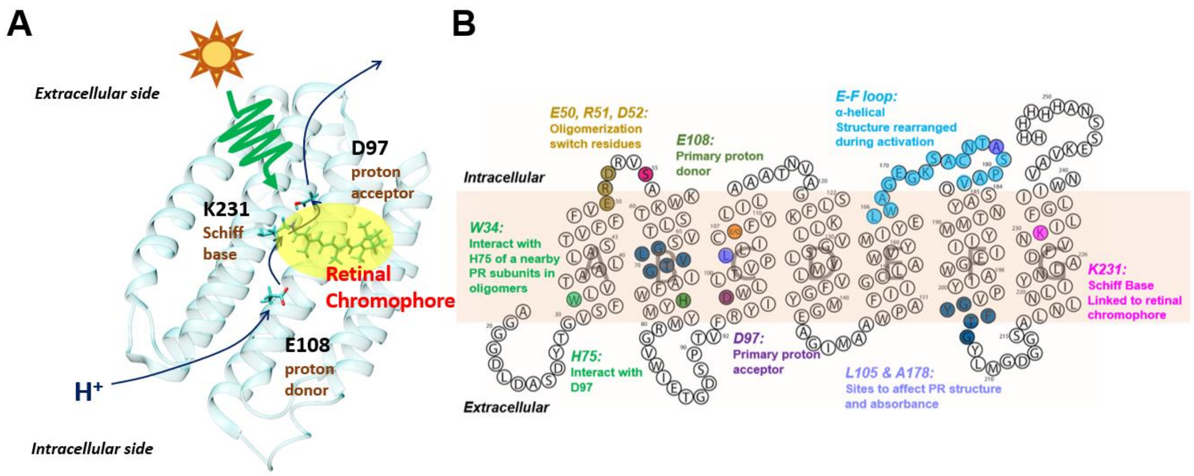


Figure 1.3: Secondary Structure of PR. A. Ribbon Diagram of PR with the light-absorbing retinal chromophore and important residues that form proton transport pathway highlighted. **B.** Snake plot of PR, as structure resolved by solid-state NMR studies. Functionally important residues are highlighted. E108, K231, and D97 are the residues that cooperatively transport proton during the photochemical reaction cycle of PR. H75 is can interact with D97 to affect its pKa. E-F loop is the third intracellular loop that shows conformational change upon the light-activation of PR. E50, R51, and D52 are the interfacial residues that drive the formation of PR pentamers or hexamers. W34 is another cross-protomer interfacial residue that can affect the H75-D97 interaction in the oligomers of PR.

PR can associate with each other to form higher-order protein complexes with functional impacts. PR oligomer complexes is first found in 2D crystalline membrane with a diameter of ~40 Å by electron microscopy.[67] These PR oligomer complexes were identified to exist in both hexameric and pentameric forms by atomic force microscopy.[68] For BPR variants, atomic-level structures of both pentameric and hexameric PR were observed by X-ray crystallography,[69] but such high-resolution structure of GPR pentameric or hexameric oligomers is still elusive. In the case of GPR, the inter-subunit distances of certain residues between the closest neighboring GPR in its hexamers has been measured by electron paramagnetic resonance (EPR), and the distances agreed with the distances estimated from the crystal structure of BPR hexamer,[70-71] supporting the existence of GPR hexamer. Furthermore, with an assist from the dynamic nuclear polarization (DNP) sensitivity enhancement, a solid-state NMR (ssNMR) study identified oligomeric interfacial residues E50, R51 and D52 that carry out specific cross-protomer interactions for holding together GPR oligomeric complexes,[72] with R51 and D52 found to be switches to affect the relative population between GPR hexamers and pentamers.[72] W34 was identified to form another cross-protomer interaction with a H75 of the closest neighboring protomer on PR oligomers, while this interaction could affect a structural movement of H75 in the proton transport reaction and the protonation state of the proton acceptor D97 residue.[73]

Despite the fact that both oligomeric and monomeric forms of PR have the same proton transport function, quantitative functional difference still can be observed between monomeric PR and oligomeric PR in detergent micellar environments.[74] Oligomeric PR was found to have a slower light-driven conformation change by comparing to monomeric PR in DDM detergent environments. In the same detergent micellar environment, the oligomer formation of PR was also found to affect the protonation state of D97 residue to affect the population of active PR. This dissertation is going to further discuss the role of oligomer formation in affecting the functional properties of PR in liposomes, which is a more native-like biomimetic environment. The knowledge of PR oligomers can shed light on understanding the functional role of homo-oligomer formation in other transmembrane protein systems.

1.2.2 Light-Activated Proton Transport Function

The light-activated proton transport function of PR can be characterized in two different aspects. One is the percentage of active PR that is capable of transporting proton, which is determined by the protonation behavior of the proton receptor D97. The other aspect is the proton transport rate of PR in its photochemical reaction cycle. Since the optical absorbance of the retinal chromophore can reflect the structural rearrangement of PR associated with its proton transport function, the two aspects of PR function can be easily quantified by static or transient optical absorption spectroscopy. With accessible functions, PR becomes a perfect model for achieving the goal of this dissertation: understanding how the functions of transmembrane proteins are tuned by different modulators.

The proton acceptor D97 of PR needs to be deprotonated ($PR^{alkaline}$) under its resting state to trigger the light activation of PR. Otherwise, if this proton acceptor D97 is protonated (PR^{acidic}) under the resting state, the proton on the protonated Schiff base K231 then cannot transport its proton to D97 after the isomerization of the retinal chromophore. With a different photoactivated response, the function of PR^{acidic} is still under debate. One school of studies believed that PR^{acidic} transport protons in an opposite direction compared to active $PR^{alkaline}$,[50, 73] while another group of studies suggests that PR^{acidic} does not have net proton transport in its photochemical reaction cycle.[55] These contradictory results make researchers in the retinal protein field suspect that the function of PR could be more than simply proton transport, but could be a sensory protein to trigger other functions. Since the protonation state of D97 is the key factor to determine whether PR is in active or acidic form, the pKa of this D97 residue (pKa_{D97}) becomes an important measure to give us an idea on the amount of active PR under an environment with a given pH.

The pKa_{D97} of PR can be directly determined by spectral titration measurements, since the protonation state of D97 can affect the absorbance of PR through changing the local environment of the nearby retinal chromophore.[64, 75] PR^{basic} with a deprotonated D97 has a maximum absorbance (λ_{max}) at ~515 nm, exhibiting a pink/red color in the solution.[54] PR^{acid} with a protonated D97 has its λ_{max} at ~540 nm, which appears to be more purple (Figure 1.4). Given this pH-dependent color transition, previous studies have established that the pKa_{D97} can be readily measured by optical absorption spectroscopy of PR as a function of bulk solution pH.[53, 76-77] The pH-dependent absorbance change can be tracked in two different ways. One is by tracking the λ_{max} shift between the two extreme cases (between ~515 and ~540 nm).[50] The other widely used method is to track the difference absorbance at 570 nm under different pH,[78-80] which can provide a better resolution on the absorbance change. Unlike the pKa of an analogous proton acceptor D85 that is deprotonated under a wide range of pH (pKa~2.6),[81] The measured apparent pKa_{D97} has been reported to vary from 6.5 to 8.0 depending on the experimental condition.[50, 82-83] These values are close to the pH of the ocean where the marine bacterioplanktons reside, providing a hypothesis that these marine bacteria are utilizing the regulation of pKa_{D97} to tune the overall proton transport capability by PR.

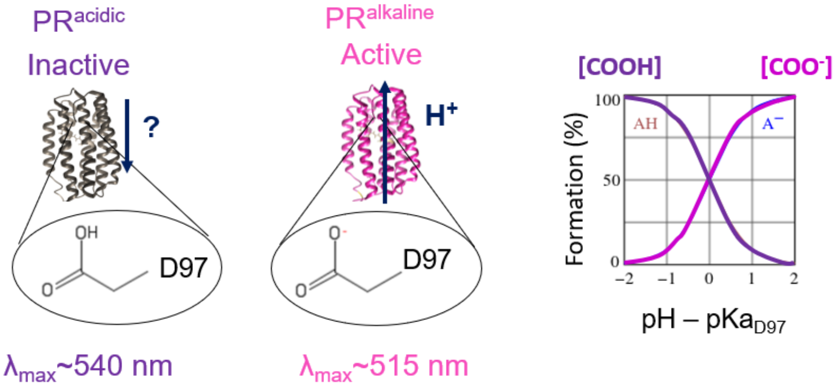


Figure 1.4: Optical Absorption of PR. PR under acid condition (PR^{acidic}) with a protonated D97 residue has a maximum absorbance (λ_{max}) at ~540 nm and PR under basic condition (PR^{alkaline}) with a deprotonated D97 residue has a λ_{max} ~ 515 nm. A spectral titration experiment can determine the pH at which an equal quantity of PR^{acidic} and PR^{alkaline} presence, namely the pKa of the D97 residue (pKa$_{D97}$).

The second aspect of PR function is the kinetics of its proton transport rate. Such kinetics has been widely studied by tracking the transient optical absorbance change after the excitation of PR with a pulsed green-light laser.[53-54, 84] Photoactivated PR undergoes a series of conformational changes that perturb the local environment of the retinal chromophore, which in turn affects PR's overall optical absorbance. Based on a similar light-activated response to bR, several photointermediates with distinct maximal absorption wavelengths have been identified that constitute the photochemical reaction cycle. [53] These include the transient conformational states sequentially labeled as K, M_1/M_2, N, and PR' intermediates, each of which has a signature absorbance spectral peak at 555 nm, 410 nm, 560 nm, and 520 nm, respectively. By tracking the growth and the decay of difference absorbance at wavelengths close to these, the kinetics of each step in the proton pumping reaction cycle can be mapped out.

The steps of the proton transport reaction are shown in Figure 1.5. After PR is photoactivated by light, the fast isomerization (~ps) of the retinal chromophore (all-trans to 13-cis) leads to a formation of the red-shifted K intermediate. The follow-up transfer of a proton from the Schiff base K231 to the proton acceptor D97 leads to a formation of blue-shifted M intermediates. The transitions from and to the M intermediates are hypothesized to accompany with the greatest conformational change between PR helices, with a possible outward movement of the F helix and a slight inward movement of the G helix based on BR studies. The M intermediates consist of M_1 and M_2 with a spectrally silent transition between the two intermediates. The Schiff base K231 is then protonated by a proton contributed from the proton donor E108 to form the N intermediate. This is followed up by uptake of a proton by the proton donor E108 from the intracellular side to form O intermediate. In the end, PR arrives back to its resting state and can be triggered again for a new reaction cycle.[54] It is noted that the mutant E108Q can slows down the whole PR photochemical reaction cycle by more than an order of magnitude and especially prolongs the M intermediates. This is done by slowing down the protonation of the Schiff base K231 during the M-N transition, as the glutamine cannot be a proton donor.[85] This dissertation utilizes the mutant E108Q to populate the M intermediates under the equilibrium of continuum green-light illumination, facilitating a structural comparison between the resting state and the M intermediates.[56, 73]

Figure 1.5: PR Photochemical Reaction Cycle. Schematic of PR photochemical reaction cycle with the conformation and the maximum absorbance wavelength (λ_{max}) of each photointermediate.[85] The M_1 and M_2 intermediates are the part that PR hypothetically experiences the greatest structural rearrangement with a blue-shift of its λ_{max}.

1.2.3 Bioengineering Applications of Microbial Rhodopsins

The photoresponsive proton transport function of PR makes it a great candidate in bioenergy applications that convert solar energy into electrochemical energy. Sunlight illumination can provide 200-250 W/m^2 energy on the Earth's surface.[86] If this energy source can be efficiently utilized, it can become a promising approach to meet up the increasing need for energy due to human population growth and technology development. Although the proton motive force (PMF) generated by PR cannot be directly converted into usable energy, learning from marine bacteria, this PMF could be used by biological systems to boost the biochemical reactions that can synthesize energy storage compounds,[87] as the PMF can subsequently be converted into cell energy source ATP by ATP synthetase. The co-expression of PR in bacteria for energy product synthesis can substitute the oxidative phosphorylation pathway in cells and enable the processes of reduced product generation to perform anaerobically.[88-89] On the other hand, the PMF generated by PR could also be directly used by proteins that can generate energy-storing products. Hydrogenase is one of the examples that can produce hydrogens H_2 from the PMF generated by PR (Figure 1.6 A).[90-92] The coordination between proton pumps as PR and other proteins can also be achieved by genetic expressing proton pumps in hosts for biofuel production or co-assemble both protein species in bio-enabled materials.[93-94]

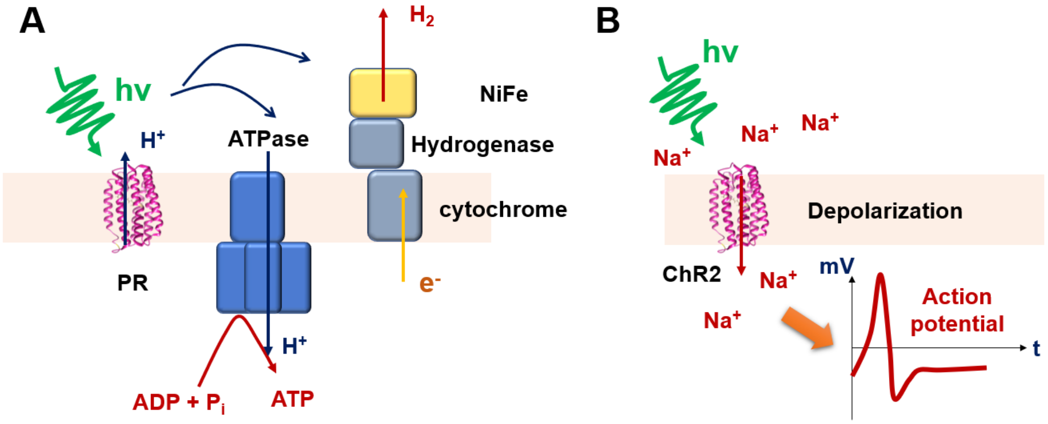


Figure 1.6: Engineering Applications of Rhodopsins. A. PR can be engineered in microbial to generate a proton motive force (PMF) that can be used for powering biofuel synthesis.[88-89] This PMF can be either used by ATPase to synthesize ATP, which can enable biofuel-generating microbial to work under anaerobic conditions, or can be used by hydrogenase to directly generate energy-storing product H_2. Adapted from Kim.[92] **B.** Channelrhodopsin-2 (ChR2) as a light-gated sodium channel enables a trigger of membrane depolarization and action potential by light illumination.

The excellent robustness of PR also allows it to maintain its structure and function not only in biological systems but also in other *in vitro* biomimetic environment, broadening its potential to be engineered in different platforms. PR has shown to be stable in a variety of biomimetic systems including 2D lipid crystals,[68] detergent micelles,[74-75, 95-96] polymersomes,[97] liposomes,[98-99] and lipid nanodiscs.[79, 100-101] Enhanced thermal and mechanical stability were achieved by incorporating PR into mesostructured silica-surfactant composites.[102] In the case of PR, even its structural integrity is preserved in these environments, it is also important to ensure a predominance of its active species with optimal functional performances while reconstituted into biomimetic materials. Therefore, a fundamental understanding of tuning the functionality of these transmembrane proteins in biomimetic materials is important for their bioengineering applications.

Throughout past decades, a variety of light-activated microbial rhodopsins beyond PR has been discovered and engineering applications of these rhodopsins have also been developed. Despite the structural similarity within the family, these microbial rhodopsins have distinct functional differences between species. Functions can be categorized into ion channel and ion pump specified for different cations or anions, or even photosensors that can regulate the photomobility or the photosynthesis of microbial.[103] Some variants have also been engineered as fluorescence sensors of membrane potential.[104-105] Among all the engineering applications developed in microbial rhodopsin studies, the optogenetic application is a particularly intriguing direction that enables a spatial and temporal control of biochemical processes and neuronal activities by local light illuminations. This can be done by genetically targeted expressing ion channels or pumps that can transport sodium, potassium, and chloride in neuron cells. Namely, channelrhodopsin-2 (ChR2) from *Chlamydomonas reihardtii* as a light-gated channel for transporting sodium can be used to trigger the depolarization of membrane potential and anion channelrhodopsin (ACRs) from *Guillardia theta* as a light-gated channel for transporting chloride can trigger the hyperpolarization of membrane potential (Figure 1.6 B).[106-107] Interestingly, these microbial rhodopsins also have a switch buried residue inside the protein to determine its activity, as the case of proton acceptor D97 inside PR. Therefore, the functional and structural understanding of how the pKa_{D97} of PR that controls its activity is affected by different modulators has a great potential to be applied in optimizing the functions of these different microbial rhodopsins in engineering applications. On the other hand, despite the functional difference, these microbial rhodopsins also share a similar photochemical reaction cycle upon light-activation as the case of PR. The development of the tool that can acquire the structural movement of PR in this dissertation could potential ass well be applied to these microbial rhodopsins upon light-activation. The potential differenece in structural dynamics between these proteins and PR could provide mechanistic insight on the reason of their different functionalities.

1.3 Magnetic Resonance Spectroscopy

Resolving the structure of proteins is necessary for uncovering how protein machinery achieve their function. Although the function of the model transmembrane protein PR used in this dissertation can be directly assessed by optical absorption techniques, many transmembrane proteins still do not have functional essays to quantify their function. The studies of PR here provide an opportunity to directly compare the difference in protein static structure or structural dynamics when it has different functional outcomes measured by optical absorption techniques. The structure-function information learned from PR could be helpful to understand how other transmembrane proteins without assessable functions but share a similar structural motif operate.

Among all the developed techniques, magnetic resonance spectroscopic techniques are major approaches to obtain atomic-level structural information of biomolecules other than protein crystallography and cryogenic electron microscopy. These techniques are widely utilized to investigate the microscopic-level structure of biomolecules, given its non-invasive nature to the sample during their measurements. In general, magnetic resonance techniques acquire signals related to the nuclear magnetism and the nuclear spin of chosen atoms on biomolecules under an external magnetic field to get protein structural information. From the interpretation of signals, information including spatial positioning of atoms or the chemical properties of a specific part on the protein construct can be obtained. In addition, the system of interest in magnetic resonance spectroscopy can be as large as the entire human body or organs, like the magnetic resonance imaging (MRI) in hospitals for investigating the condition of soft tissues, or as small as nanometer-scale proteins or other biomolecules for studying microscopic molecule interactions.

Our introduction here focuses on the fundamentals of magnetic resonance and the principles of techniques used in this dissertation to map out the structure of PR under an influence of various functional modulators. The structural information of PR obtained from various magnetic resonance tools is compared to the proton transport functions of PR measured by optical spectroscopy tools to establish an understanding of the structure and function relationship of transmembrane proteins. Electron paramagnetic resonance (EPR) and Overhauser dynamic nuclear polarization (ODNP) are integrated to study the local dynamics of the protein side chain and the local water dynamics on the protein surface, respectively, around the solvent-exposed interhelical loop of PR. Solid-state nuclear magnetic resonance (ssNMR) spectroscopy with dynamic nuclear polarization (DNP) enhancement is utilized here to explore the entire structure of PR, including the interior region. Finally, high field EPR on PR labeled with gadolinium (Gd) spin labels is developing to acquire not only inter-helix distances on PR but also the dynamic of these distances during its proton transport function.

1.3.1 Fundamentals

Magnetic resonance spectroscopy is a powerful approach for investigating properties of both inorganic and organic matter at an atomic level. The microscopic information of matter is obtained through intrinsic physical properties called nuclear spin and nuclear magnetism. These physical properties have no direct effect on other physical or chemical properties of the matter, but can be a probe to retrieve information on other properties without disturbing them by magnetic resonance spectroscopic measurements The spin-related properties (e.g. ground nuclear spin state) of the target nucleus are determined by its spin configuration composed by particles including protons and neutrons. Under the influence of an external magnetic field, a nuclear ground state with spin I would degenerate into $2I + 1$ energy level. Throughout the acquisition of magnetic resonance experiments, the distribution of spins between different energy levels is excited from its equilibrium by electromagnetic irradiations. The spin behavior after excitation then is detected for retrieving desired information of spins.

One of the important constants in magnetic resonance spectroscopy is gyromagnetic ratio γ, it is a proportionality constant between the spin angular momentum S and the magnetic moment μ. This constant decides the frequency for an electron or a nucleus precess under an external magnetic field B_0 by a relationship $\omega_0 = -\gamma B_0$, where ω_0 is called the Larmor frequency. The spins can be excited from a lower energy level to a higher one with an energy difference ΔE only when the frequency of applied electromagnetic irradiation matches the Larmor frequency ω_0 of the spin. Since the gyromagnetic ratio of electron ($\gamma_e = 1.76085963023 \times 10^{11}\ s^{-1} \cdot T^{-1}$) is almost three orders of magnitude higher than the ones of nuclei (One of the highest is ^{1}H, with $\gamma = 4.257747818 \times 10^7\ s^{-1} \cdot T^{-1}$), the electromagnetic irradiation required for EPR experiments is under the microwave range (~GHz) while the irradiation required for NMR experiments is under the radio frequency range (~MHz). The higher frequency of electromagnetic irradiation for EPR spectroscopy makes it instrumentally more challenging for data acquisition or complicated spin manipulation.

Gyromagnetic ratio also determines the sensitivity of a spin and the strength of dipolar interaction between spins in magnetic resonance spectroscopy experiments. A spin with a larger gyromagnetic ratio gains a higher sensitivity because of a greater spin population difference between states that have a greater energy difference. For common nuclei in protein NMR spectroscopy, ^{1}H has the greatest sensitivity, which is followed up by ^{13}C, with a gyromagnetic ratio ($\gamma = 1.07084 \times 10^7\ s^{-1} \cdot T^{-1}$) that is one-fourth of the value of ^{1}H. ^{15}N is also another common nucleus in protein NMR but with a lower sensitivity comparing to the other two nuclei, with a gyromagnetic ration ($\gamma = 4.257747818 \times 10^7\ s^{-1} \cdot T^{-1}$) that is less than one-tenth of the value of ^{1}H. Comparing to NMR nuclei, EPR sensitive paramagnetic species with unpaired electrons has ~660 times larger gyromagnetic ratio. This property makes through-space dipolar interactions between EPR sensitive spins a lot stronger than the ones between NMR sensitive nuclei. The much greater dipolar interactions between electrons allow us to exact inter-spin distances under a longer distance range (2-10 nm, varies between different techniques and sample conditions). The following sections will describe the strength and details of different NMR and EPR techniques used in this dissertation to obtain different aspects of protein structural information.

1.3.2 Continuous-wave Electron Paramagnetic Resonance (cw EPR)

An easier and the most popular way for EPR study is the continuous-wave (cw) method. Since the electron is a spin 1/2 system, it degenerates into two different energy states under an external magnetic field B_0, which is also called Zeeman energy splitting. The energy of each state is determined by

$$E = m_s g \mu_B B_0, \tag{1.1}$$

where m_s is the azimuthal quantum number, which can be either +1/2 or -1/2 in the case of electrons, g is the g-factor that characterize the electron spin-orbital coupling behavior, μ_β is the Bohr magneton constant. The energy splitting ΔE between these two energy states can be described as

$$\Delta E = g \mu_B B_0. \tag{1.2}$$

Here the energy splitting is proportional to the magnetic field strength as illustrated in Figure 1.7 A. If the electrons are irradiated with microwave irradiations at a frequency ν that fulfill the resonance condition, which means the energy of the microwave irradiations matches the energy splitting of electrons under a specific external magnetic field strength, the EPR absorbance peak would appear. In practical cw EPR measurements, the microwave irradiation frequency is fine-tuned to a value that yields an optimized power in the cavity with a target sample. The external magnetic field B_0 from the electromagnet around the EPR cavity is swept in between a range that the resonance condition could happen. To achieve higher sensitivity, the signal acquisition is done with an assist from a lock-in amplifier. This lock-in detection allows us to extract desired signals from a noisy background with phase information. It is done by an application of a small additional oscillating magnetic field ΔB_0 during the hall field sweeping and results in a reflected microwave with an amplitude change of ΔV (Figure 1.7 B). This amplitude change is plotted against the change of hall field B_0 to make the resulting signal a first derivative of the absorption spectrum.

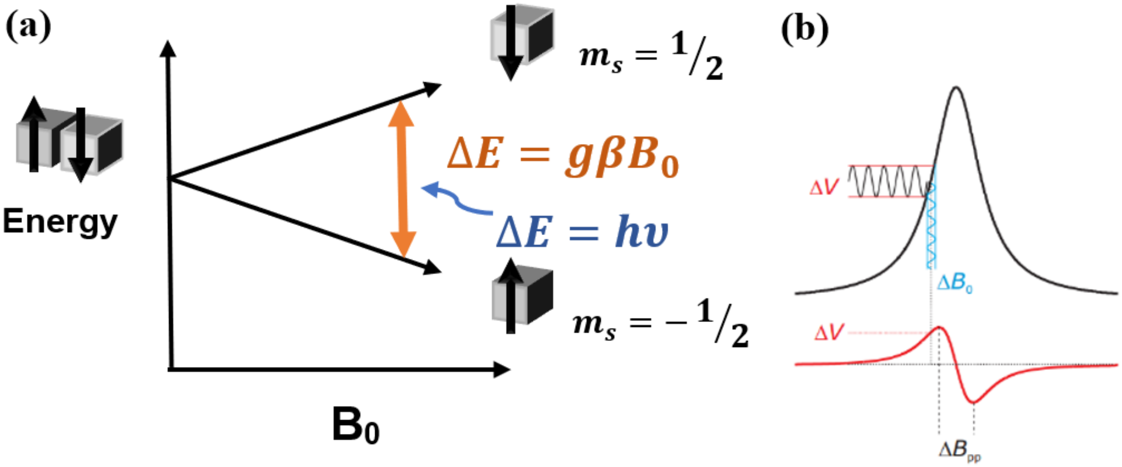


Figure 1.7: Origins of cw EPR spectra. A. The energy splitting of electrons (S=1/2) by Zeeman interactions under different external magnetic field B_0. Under a resonance condition, the microwave irradiation frequency must match the energy gap between two energy levels. **B.** The EPR signal is acquired by a lock-in detection and resulted in the first derivative of the absorption. Adapted from Jeschke et al.

The example in Figure 1.8 A shows the simplest situation for a spin ½ electron with solely Zeeman interaction in the system. Hyperfine interaction is another interaction between the unpaired electron and a nucleus that could affect the energy diagram of the electron. For example, MTSL (S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate), a common nitroxide-based EPR spin label, has the hyperfine interaction between the unpaired electron and a coupled nitrogen atom in the vicinity (Figure 1.8 A). Since the coupled ^{14}N is a spin I=1 nucleus, it degenerates into three energy states under an external magnetic field. The energy state of the coupled unpaired electron splits accordingly into three states through hyperfine couplings, with one state to be the same as the original one and the other two states to be a value higher and lower than the original state. This results in a spectrum with resonances occur under three different magnetic fields.

Since the unpair electron on different spin labels could occupy different orbitals and the whole spin label could not be local highly symmetric, the above-mentioned interactions have an orientational dependency between the spin label and the external magnetic field. Quantum mechanical description, with both Zeeman interaction $\boldsymbol{g}$ and hyperfine interaction $\boldsymbol{A}$ expressed in a tensor form, of the Hamiltonian is

$$\hat{\mathcal{H}} = \mu_B B_0 \cdot \boldsymbol{g} \cdot \hat{S} + \hat{S} \cdot \boldsymbol{A} \cdot \hat{I}, \tag{1.3}$$

where S and I are the electron and nuclear spins, respectively. Both Zeeman tensor **g** and hyperfine tensor **A** are anisotropic, this makes both interactions change with the orientation of spin labels with respect to the external magnetic field. In the case of nitroxide-based spin labels, the hyperfine interaction along the z-direction is greater than the other two directions, which results in a greater splitting between the three resonances (Figure 1.8 B). In actual EPR measurement, the resulting spectrum of the nitroxide-based spin label does not look like one of any individual orientation. Depending on the rotational correlation time t_c of the spin label, it can be an average of all three directions (fast-limit) or a superposition of the three directions (rigid-limit), or anywhere in between the two previous extreme cases for intermediate correlation times (t_c between 100 ps and 1 μs) (Figure 1.8). Therefore, cw EPR spectra can provide information on how fast the spin label or the system of interest that the spin label is conjugated to rotates. Additionally, the local microenvironment around the spin label, specifically the polarity of solvents, can also affect the resulting EPR spectra by changing the hyperfine coupling between the ^{14}N nucleus and the unpaired electron. Overall, cw EPR can provide us fruitful information on the local structure of protein around the site with a spin label attached.

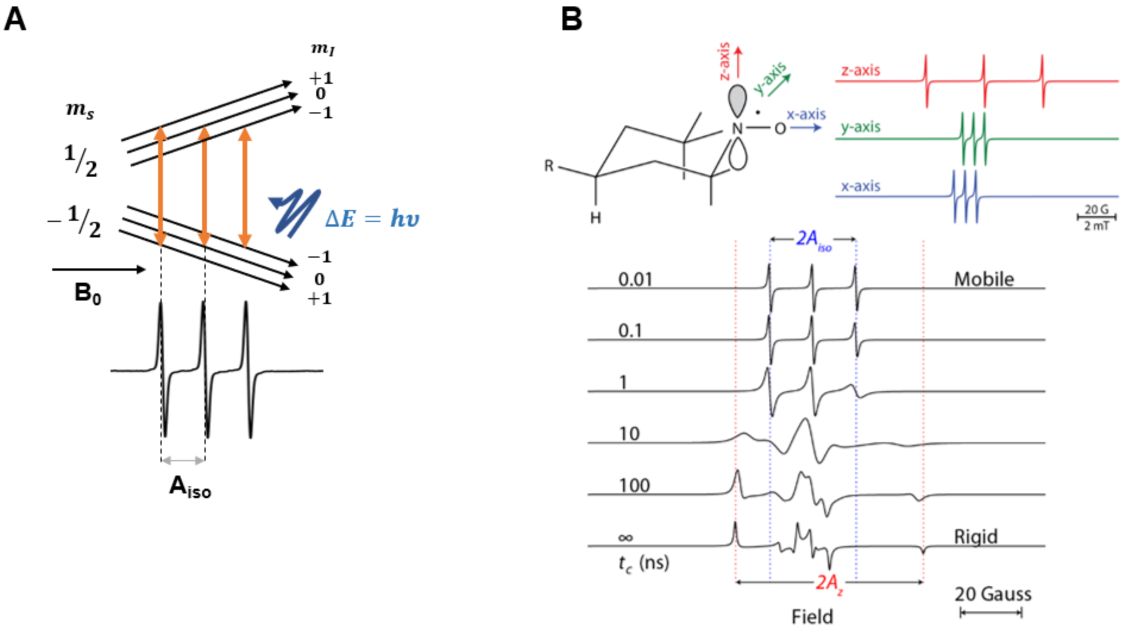


Figure 1.8: cw EPR spectra of nitroxide-based spin labels. A. Hyperfine interaction between an electron spin S and a nitrogen nucleus I split the single absorption peak of spin ½ electron into three resonances. **B.** The molecular structure of a nitroxide radical with the hyperfine splitting of the three orientations in the Cartesian coordinate. The bottom part demonstrates a change of nitroxide radical's cw EPR spectra from the fast limit to the rigid limit as the correlation time t_c increase.

cw EPR lineshape analysis can also be used to extract the distance between a pair of electron spins. This is done by investigating the extent of spectral broadening caused by the dipolar interaction between a pair of spins that are within an upper distance limit. The lineshape broadening is more significant while the paired electron spins come to a shorter distance. The upper distance threshold that the dipolar interaction can affect the cw EPR spectral lineshape is highly dependent on factors including the strength of the external magnetic field and the intrinsic linewidth of the spin label used. For example, experiments on nitroxide-based spin label like MTSL under X-band (~0.35 T) can resolve inter-spin dipolar interaction up to 2 nm. This distance is shorter than most of the distance of interest in the transmembrane protein system, as intermolecular distances between subunits in protein oligomers or intramolecular distances between helices in a protein are usually above 2 nm. As a result, different advanced methods have been developed to extend the inter-spin distance that cw EPR lineshape analysis can resolve.

High field (8.6 T) cw EPR technique using gadolinium-based (Gd-based) spin labels developed at UCSB is an approach that successfully extends the detectable distance of cw EPR lineshape analysis to ~3.5 nm.[108] As Gd^{3+} complexes are spin 7/2 system, it has a longer-range dipolar interaction between spins by comparing to spin ½ nitroxide-based spin labels, enabling a longer accessible distance in cw EPR lineshape analysis. On the other hand, the central transition $|-1/2\rangle \rightarrow |\,1/2\rangle$ of Gd ions become narrower under a higher external magnetic field. Under 8.6 T, this central transition becomes significantly narrower than all the other transitions in the energy diagram, resulting in an absorption spectrum with a single narrow absorption peak over a broad background. This allows the spectra broadening from weaker dipolar interactions not to be hindered by the intrinsic linewidth of the resonance and can provide a better resolution (sub-nm) on the inter-spin distance between 2-3 nm.[108] In membrane protein structural studies, Gd^{3+} is introduced on proteins by Gd-based labels that have ligands to chelate a Gd^{3+} and a linker that is capable to conjugate with thiol functional groups on the protein.[109-111] Most importantly, high field (8.6 T) cw EPR technique using Gd-based spin labels has the potential to measure inter-spin distances under room temperature. This characteristic opens the possibility to measure not only static distances on protein from the trapped state but also the dynamics of conformational changes under physiologically relevant temperatures.

In the developed Gd-Gd cw EPR distance measuring technique, inter-spin distance is obtained from the difference of peak-to-peak linewidth between the system of interest with a pair of Gd^{3+} complexes and a reference control sample with solely one Gd^{3+} complex from the selected pair.[112] In PR structural measurements, we are interested in measuring the distance between sites on different helices and the change of this distance during its proton transport function. To prepare the PR sample for Gd-Gd cw EPR distance measurements, a pair of Gd-based spin labels is introduced to a monomeric with two selected solvent-exposed sites that are mutated into cysteine residues (Figure 1.9). The peak-to-peak linewidth of this doubly-Gd-labeled PR (ΔH_{dl}) is then compared to the one of a reference sample with only one of the sites being labeled (ΔH_{sl}), which is the case without any dipolar broadening effects on its spectrum. By using the calibration between linewidth broadenings and the known inter-spin distances of Gd-ruler molecules, the actual distance between helices then can be estimated from the observed peak-to-peak linewidth broadening of the doubly-Gd-labeled PR ($\Delta\Delta H$).[108] In the case of measuring distance changes between the two Gd-labeled sites, the peak-to-peak linewidth from the cw EPR spectra of the target doubly-Gd-labeled PR (ΔH_{dl}) can be directly compared between different trapped states in the photochemical reaction cycle. Any possible distance change between the two Gd-labeled sites would then be reflected in the change of peak-to-peak linewidth.

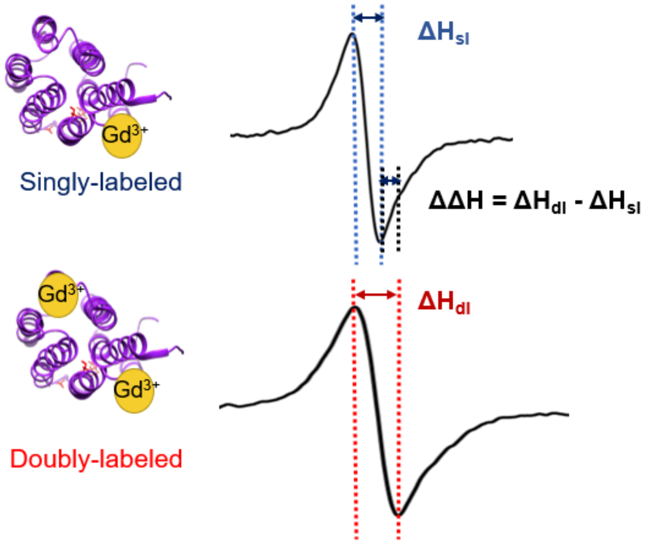


Figure 1.9: 240 GHz cw EPR spectra. The spectra illustrate how the peak-to-peak linewidth broadening is get from the linewidth difference between the doubly-Gd-labeled PR and the singly-Gd-labeled.

1.3.3 Overhauser Dynamic Nuclear Polarization (ODNP)

^{1}H Overhauser dynamic nuclear polarization (ODNP) NMR relaxometry can measure local water dynamics within 2-4 water layers above a spin-labeled biological surface.[113-114] This technique has been widely applied to measure hydration dynamics on surfaces of lipid membranes,[115-116] intrinsically disordered protein,[117] water-soluble proteins,[118] and transmembrane proteins,[95, 119] where the hydration dynamic information could encode the relation between structural water on protein surfaces and protein function.[120] This hydration dynamic information on a spin-labeled surface is extracted from the ^{1}H NMR enhancement profile of water around the spin label, typically nitroxide-based MTSL, under various microwave irradiation power, where the characteristic of the enhancement profile encodes the translational dynamics of water between picoseconds to nanoseconds regime. The ^{1}H NMR signal enhancement is originated from the time-dependent dipolar coupling between the unpaired electron on the spin label and the ^{1}H nuclei on water molecules. Once microwave irradiation is applied to the system, the greater polarization of electrons will transfer their energy to surrounding protons on water through cross-relaxation. Since the transition is dominated by the short-range and distance-dependent dipolar interaction between the electron spin and the proton spin, this enhancement is limited to water molecules that are less than ~10 Å away from the spin label.

To explain the origin of NMR signal enhancements through cross-relaxation, a four-level energy diagram between an electron spin S and a proton spin of water I is needed (Figure 1.10). It is noted that the electron and the proton are both spin ½ and degenerate into two energy levels in an external magnetic field. The electron spin S can exist in one of the two spin states, either the higher energy +1/2 state or the lower energy -1/2 state. On the contrary, the proton spin with a positive gyromagnetic ratio prefers to align itself with the magnetic field as the +1/2 becomes the lower energy state and -1/2 becomes the higher energy state. Under an external magnetic field, the two energy levels of the coupled spins are separated by a quantity proportional to the field strength and the gyromagnetic ratio of spins, so the energy difference between two electron states are greater. Assuming no other transitions exist, the population of spins at all four energy levels would be determined by Boltzmann distribution. But the case here is both scalar and dipolar interactions exist between the two coupled spins, all the possible transitions between the four states that can affect the population of spins is demonstrated in Figure 1.10.

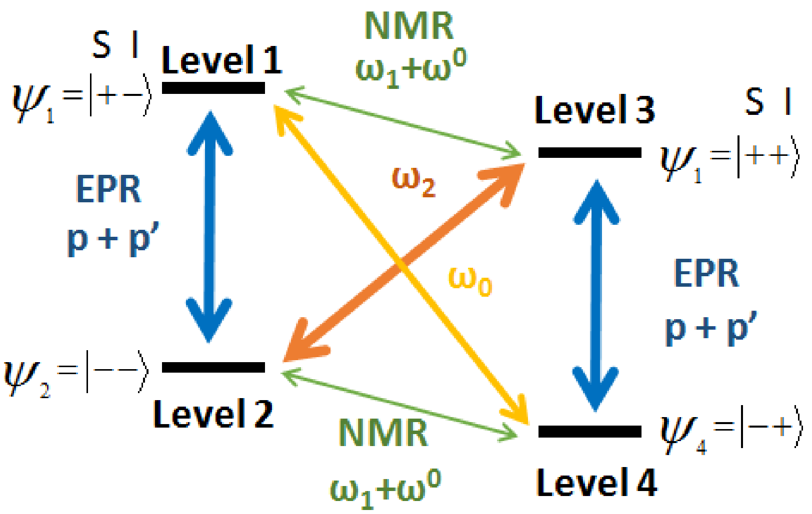


Figure 1.10: Energy Level Diagram of a two-spin system. The four-level diagram is between coupled proton spins (Spin I) and electron spins (Spin S). Arrows show all possible transitions and corresponding transition rates ω_i between two energy levels.

Zero-quantum transition ω_0 involves simultaneous flips of the electron and the proton from the lowest energy level to the highest energy level, which contains both scalar and dipolar terms. Double-quantum transition ω_2 involves flips of two spins in the same direction, which contains only dipolar terms. Single-quantum transitions (ω_1 and p') involves flips with only either electron or proton spin. In addition, transition rates ω_0 and p are related to the longitudinal relaxation of the proton and the electron without any couplings between two spins. In the case with continuous microwave irradiation, electron spins are excited to the two higher energy levels. Since the double quantum transition ω_2 is 6 times faster than the zero-quantum transition ω_0, this will populate the spin population at Level 2 compared to Level 4, and lead to a negative enhancement on the ^{1}H NMR signal.

The relationship between the NMR signal enhancement profile and hydration dynamics can be derived from the equation of motion that describes the polarization of proton nuclear spins interacting with nitroxide electron spins as

$$\frac{d\langle I_z\rangle}{dt} = -(\omega_0 + 2\omega_1 + \omega_2 + \omega^0)(\langle I_z\rangle - I_0) - (\omega_2 - \omega_0)(\langle S_z\rangle - S_0), \qquad (1.4)$$

where $\langle I_z\rangle$ and $\langle S_z\rangle$ are nuclear and electron magnetizations, respectively, I_0 and S_0 are nuclear and electron spin equilibrium magnetizations without microwave irradiation, respectively, ω_0, ω_1, and ω_2 are zero-, single-, and double-quantum transition rates, respectively. The three quantum transition rates represent the transition probabilities of the spin population between different energy states. The first term here can be separated into two parts, where $\omega_0 + 2\omega_1 + \omega_2$ can be defined as self-relaxation rate ρ, and ω^0 here is the nuclear spin relaxation (${T_{1,0}}^{-1}$) unrelated to paramagnetic relaxation. The second term $\omega_2 - \omega_0$ can be defined as a cross-relaxation rate σ. After substituting in these definitions, the steady-state solution of Equation $\frac{d\langle I_z\rangle}{dt} = -(\omega_0 + 2\omega_1 + \omega_2 + \omega^0)(\langle I_z\rangle - I_0) - (\omega_2 - \omega_0)(\langle S_z\rangle - S_0)$, (1.4) under continuous microwave becomes

$$E = 1 - \frac{\sigma}{\rho} \cdot \frac{\rho}{\rho + {T_{1,0}}^{-1}} \cdot \frac{S_0 - \langle S_z\rangle}{S_0} \cdot \frac{|\gamma_S|}{\gamma_I}, \qquad (1.5)$$

where $\frac{|\gamma_S|}{\gamma_I}$ is the ratio of electron gyromagnetic ratio to the one of proton, which is close to 660. The parameters in Equation $E = 1 - \frac{\sigma}{\rho} \cdot \frac{\rho}{\rho + {T_{1,0}}^{-1}} \cdot \frac{S_0 - \langle S_z\rangle}{S_0} \cdot \frac{|\gamma_S|}{\gamma_I}$, (1.5) can be further categorized into three dimensionless ODNP parameters that could affect the NMR enhancement E, including coupling factor ξ, leakage factor f, and saturation factor s:

$$\xi = \frac{\sigma}{\rho} = \frac{C_{SL} k_\sigma}{C_{SL} k_\rho} = \frac{k_\sigma}{k_\rho} \tag{1.6}$$

$$\mathrm{f} = \frac{\rho}{\rho + \mathrm{T}_{1,0}{}^{-1}} = \frac{\mathrm{T}_1{}^{-1} - \mathrm{T}_{1,0}{}^{-1}}{\mathrm{T}_1{}^{-1}} \tag{1.7}$$

$$\mathrm{s} = \frac{S_0 - \langle S_z \rangle}{S_0} \tag{1.8}$$

The coupling factor ξ describes the efficiency of cross-relaxation σ between the electron spin and the proton spin. This factor contains the relative dynamic information between spin labels on biomolecules and ^{1}H spins on water. The cross-relaxation rate σ can be expressed as the product of spin label concentration C_{SL} and cross-relaxivity k_σ, which is exclusively sensitive to the fast water translational movement (~ps). Self-relaxation rate ρ can be expressed as the product of spin label concentration C_{SL} and self-relaxivity k_ρ, which is a contribution from the dipolar self-relaxation of the studied system conjugated with spin labels. Leakage factor f is the ratio of the dipolar relaxation rate to the total relaxation rate, which can be experimentally obtained by measuring the spin-lattice relaxation times T_1 and $T_{1,0}$ with and without the presence of spin labels, respectively. Finally, saturation factor s is a factor related to the microwave power p, which can asymptotically approach to a maximum value s_{max} while the microwave power p increase to an infinite value.

To obtain the cross-relaxivity k_σ of the spin-labeled protein site. We can further rearrange Equation $E = 1 - \frac{\sigma}{\rho} \cdot \frac{\rho}{\rho + \mathrm{T}_{1,0}{}^{-1}} \cdot \frac{S_0 - \langle S_z \rangle}{S_0} \cdot \frac{|\gamma_S|}{\gamma_I}$, (1.5) into the following form by extrapolating it to infinite power

$$k_\sigma = \frac{1}{C_{SL}} \frac{|\gamma_S|}{\gamma_I} \lim_{p \to \infty} \left(\frac{1 - \mathrm{E}(p)}{T_1(p)} s(p) \right), \tag{1.9}$$

where $T_1(p)$ is the spin-lattice relaxation time of the system with spin labels presence as a function of microwave power. Both spin-lattice relaxation time $T_1(p)$ and NMR signal enhancement E(p) can be experimentally measured by inversion recovery and ^{1}H direct detection measurements under various microwave irradiation powers. Spin label concentration C_{SL} can be measured separately by the double integral of the cw EPR spectrum. An example of the cross-relaxivity k_σ enhancement series from the combination of E(p) and $T_1(p)$ under different microwave irradiation powers is shown in Figure 1.11. The enhancement curve is fitted by Equation $k_\sigma = \frac{1}{C_{SL}} \frac{|\gamma_S|}{\gamma_I} \lim_{p\to\infty} \left(\frac{1-\mathrm{E}(p)}{T_1(p)} s(p) \right)$, (1.9) and extrapolate to infinite microwave power to get the cross-relaxivity k_σ. The cross-relaxivity k_σ then is used to compare the fast water diffusion (~ps) under different conditions.

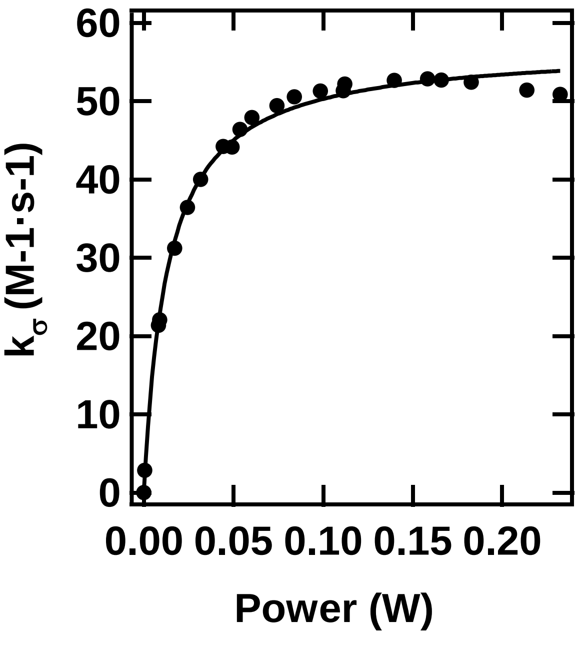


Figure 1.11: ODNP Enhancement Series. ^{1}H Overhauser dynamic nuclear polarization (ODNP) enhancement series of PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes as an example of how the fitting is done to obtain the cross-relaxivity k_σ.

1.3.4 Pulse EPR

Besides cw EPR measurements introduced earlier in this dissertation, pulsed EPR is another powerful approach that can extract information on the local environment of electron spins labeled on biomolecules. With an extendibility of utilizing different pulse patterns, pulsed EPR experiments can filter out the interaction of interest from all other interactions in the system with a higher resolution. The same resolution is hard to be achieved by cw EPR measurements as deconvoluting different interactions from a single cw spectrum could not be trivial. Moreover, spectral components with different spin-magnetic moments or relaxation times can be separated in a pulsed EPR approach by using appropriate pulse sequences. Among all the designed pulsed EPR experiments, double electron-electron resonance (DEER) is a popular technique to measure the distance between electron spins based on the strength of their dipolar coupling.[121] This technique can measure inter-spin distances from 2 nm up to ~8 nm, which covers up a wide distance range toward the distances between peripheral sites on different domains of transmembrane protein assemblies.[122]

The general principle of DEER is to detect how the spin behavior of a group of electron spins is affected by the spin state of the second group of electron spins. If the two groups of electron spins are dipolar coupled (inter-spin distances < 8 nm), a change of the dipolar field induced by the second group of electron spins will then affect the local field experienced by the first group of electron spins. To separate electron spins into two groups, the excitation of electron spins is done under two separate frequency ranges. One frequency is for observer excitation and the other one is for pump excitation. The four-pulse sequence is one of the most widely used pulse sequences in DEER (Figure 1.12).[123] At the observer excitation frequency, a refocused primary echo sequence is used. At the pumping frequency, a π flipping pulse is placed at a variable time t between the first Hahn echo and the second π pulse at the observer frequency. When the two groups of electron spins are dipolar coupled, the position of the flipping pulse at the pumping frequency can affect the intensity of the refocusing echo formed by the observer spins, as part of the observer spins could no longer refocus while experiencing a different local magnetic field. A time-domain DEER trace is collected by tracing the integration of the refocusing echo as we change the location of the π flipping pulse along the variable time t. The distance information between spins is encoded in the time-domain trace by periodic oscillations induced by spin-spin dipolar interactions. This oscillation will have a longer period if the inter-spin distance is longer. As a result, reliable information on longer inter-spin distances can be extracted if we can record a longer time-domain DEER trace. The available range of this variable time t can be extended if the time between the first Hahn echo and the second π pulse (d_2) is longer. However, a longer d_2 also means a smaller signal from the refocusing echo, as spin-spin relaxation is another factor that can reduce the echo signal. This could end up with a poor signal-to-noise ratio in the time-domain DEER trace and more signal averaging will be required to get analyzable data.

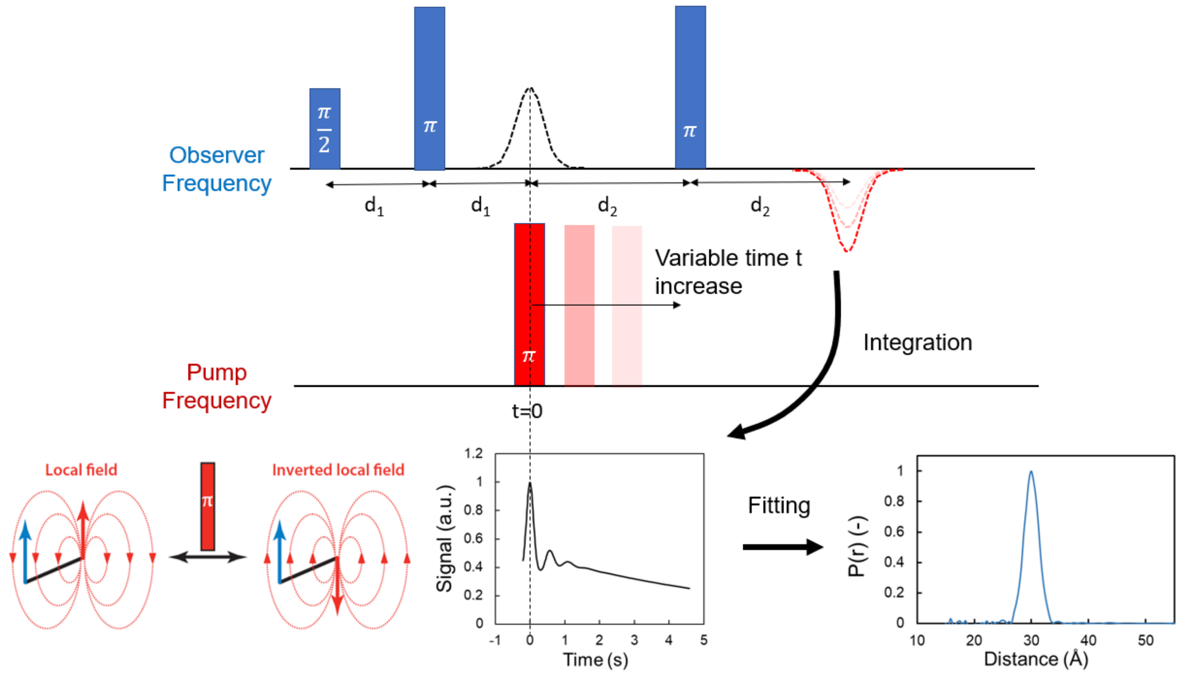


Figure 1.12: Concept of four-pulse DEER. Electron spins are excited under two separate frequencies. The integration of refocusing echo formed by electrons at the observer frequency can be affected by a flip of coupled electrons at the pumping frequency. The modulation effect on the echo integration is collected while the pump pulse position is changed along the variable time t. The time-domain DEER trace is then fitted to obtain the distance distribution of the system of interest. Adapted from Jeschke.[122]

The distance distribution is commonly extracted from the dipolar oscillation in time-domain DEER data through Tikhonov regularization.[124] An additional constraint is added to model-free Tikhonov regularization to ensure that the probability of all distances is greater or equal to zero. The fitting of the time-domain DEER trace can also be done by a model-based approach, typically by assuming the distance distribution to be a Gaussian distribution or a combination of multiple Gaussian distributions.[125] Using either the model-based Gaussian fitting or the model-free fitting usually depends on whether additional information on the expected distance distribution of the measuring system is available or not.

The DEER measurement is done under cryogenic temperatures ranging from 10 to 85 K. Under this temperature range, the relaxation rate of electron spins is slow enough for us to have a decent refocusing echo signal even if the d_2 in the pulse sequence is set to a large value (usually d_2 is less than 4 μs for spin-labeled transmembrane protein sample). Deuterating the sample or the solvent is common in sample preparation to increase the transverse relaxation time T_2 of electron spins. As DEER measurements need to be done under cryogenic temperatures, samples are flash-freezing by liquid nitrogen before being inserted into the microwave cavity. A glassing agent is added to the sample to prevent spin cluster formation during the flash freezing treatment. The operation of DEER measurements under cryogenic temperatures limits it from dynamically tracking protein conformation changes from one state to another *in situ*. However, a comparison of protein structure between different trapped equilibrium is still available by using this technique.

Gd^{3+}-Gd^{3+} DEER experiment is a novel approach by comparing to common DEER experiments with nitroxide-based spin labels as there are several benefits with using Gd^{3+}.[126] First, the spin-lattice relaxation rate T_1 of Gd^{3+} is generally faster than nitroxide at cryogenic temperatures. This allows a higher repetition rate on experiments and more scans on time-domain trace can be collected within a certain time. Next, as a higher spin 7/2 system, Gd^{3+}-Gd^{3+} DEER has the potential to extend the long-distance limit of DEER to above 8 nm. Most importantly, the central transition of Gd^{3+} has the weakest anisotropy, which makes the Gd^{3+}-Gd^{3+} DEER be lack of orientation selectivity. Nevertheless, the pseudo-secular term in the dipolar interactions between Gd^{3+} could introduce artifacts in Gd^{3+}-Gd^{3+} DEER time-domain traces.[127] This pseudo-secular artifact is more significant when the inter-spin distance is shorter. A way to prevent this artifact is to increase the frequency separation between the observer and the pumping frequency. Overall, the Gd^{3+}-Gd^{3+} DEER is a good complementary to the nitroxide-based DEER that has optimal performances under a different distance measuring range.[128-129]

1.3.5 Solid-state Nuclear Magnetic Resonance (ssNMR)

Solid-state nuclear magnetic resonance (ssNMR) is widely used to characterize the structure of transmembrane proteins. Oppose to solution NMR that can have higher resolution only when the size of proteins is smaller to have a faster tumbling, ssNMR is more ideal for transmembrane protein samples that generally have larger assemblies with additional surfactant molecules absorbed for structural stabilization. To reduce anisotropic features from the resulting spectra, ssNMR is often operated with the sample packed in a rotor and spin at a magic angle $\theta_m = 54.74°$ with respect to the direction of the magnetic field. The spinning rate of the rotor for magic-angle spinning (MAS) ranges from 1 to 130 kHz. The maximum spinning rate is highly dependent on the size of rotors and the construct of probes used for the experiment. Higher spinning rates can lead to a higher resolution spectrum with effects of orientation-dependent interactions (e.g. dipolar interaction, chemical shift anisotropy) reduced or removed. To achieve higher sensitivity and resolution of ssNMR spectra acquired from transmembrane protein samples, development on superconducting magnets with a higher magnetic field[130] and MAS probes with a higher spinning rate[131] are still ongoing goals in NMR instrument development. Developing advanced pulse sequences that can achieve better efficiency of spin decoupling or coupling is also another common approach of improving the sensitivity or resolution of protein ssNMR.[132]

Common target nuclei in transmembrane protein ssNMR studies are proton, carbon, and nitrogen. These atoms naturally exist in all amino acids on transmembrane proteins and have spin ½ isotopes. Fluorine, a spin ½ atom with a high gyromagnetic ratio, is also another available option. ^{19}F can be labeled on certain amino acids (e.g. tryptophan) during recombinant protein expression.[133] In the case of carbon and nitrogen atoms, the spin ½ isotope form of both nuclei (i.e. ^{13}C and ^{15}N) are not naturally abundant. Therefore, transmembrane protein samples for ssNMR measurements are prepared by recombinant protein expression in isotopically labeled media to increase the total number of NMR active nuclei.[134] The expressed protein can be either uniformly labeled with the NMR active nuclei or site-specifically labeled on selected amino acids to simplify the resulting spectra with desired resonances. The difference between uniformly labeling or site-specifically labeling can be determined by the composition of culture media used for the recombinant protein expression.

Detections on ^{13}C and ^{15}N that have lower gyromagnetic ratios in ssNMR are often assisted by cross-polarization (CP)[135] and dynamic nuclear polarization (DNP)[136] to increase their sensitivity. The approach is to transfer greater polarizations from proton and electron spins to ^{13}C and ^{15}N nuclei with lower polarization. In protein ssNMR measurements, the combination of DNP and CP can yield a 1-2 order of magnitude enhancement of sensitivity on ^{13}C or ^{15}N measurements, and this can significantly reduce the time required for getting a spectrum with decent signal-to-noise from weeks down to hours.[137] The transmembrane protein samples for ssNMR with DNP enhancement are usually prepared in a DNP juice (d8-glycerol/D_2O/H_2O, 60/30/10, v/v/v) doped with a polarization agent that has unpaired electrons. Common polarization agents are biradicals (e.g. AMUPOL, TOTAPOL) that can utilize polarization mechanisms including three-spin thermal mixing and cross effect.[138-139] A cw microwave irradiation is generated by a gyrotron and guided to the sample with polarization agents in the NMR probe by corrugated waveguides. The NMR probe and the sample are cooled down to cryogenic temperatures (< 100 K). Under this low temperature, electron spins would have an optimal relaxation time that facilitates the transfer of polarization from electrons to other nuclei.

While DNP enhances the polarization of proton spins, the polarization can be further transferred to ^{13}C or ^{15}N through CP (Pulse sequence in Figure 1.13). This polarization transfer is done by applying two contact pulses to two groups of spins with the Hartman Hahn condition being fulfilled.[135] A $\pi/2$ pulse is first applied to rotate the ^{1}H into the transverse plane. A spinlock pulse that is in phase with the precessing ^{1}H spins is followed up. The other contact pulse is adjusted to a correct strength so that the Zeeman splitting of the target spin (^{13}C or ^{15}N) induced by the B_1 field is the same as the one of the ^{1}H spin. Once the polarization is transferred, the free induction decay signal (FID) from the target spin can be collected and analyzed. In one-dimensional (1D) NMR data analysis, the time-domain FID is Fourier transformed to get a frequency-domain NMR spectrum. Nuclei under different local chemical environments can result in resonances with different chemical shifts on the NMR spectrum. It is noted that the relative intensity of peaks in DNP and CP experiments is not necessarily equal to the relative quantity of nuclei with different chemical shifts. This is because the enhancement factor for nuclei with different solvent accessibility or local proton concentration could be different. A quantitative comparison of intensities between resonances will require a direct excitation and detection experiment on the target spin with a lower gyromagnetic ratio.

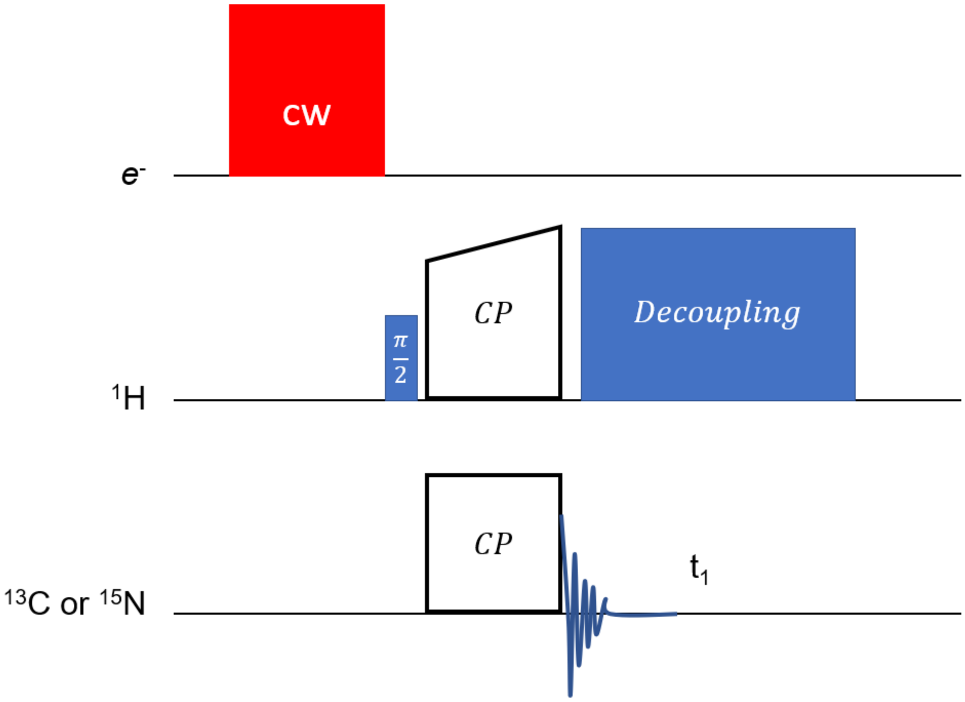


Figure 1.13: Pulse Sequence of CP-MAS with DNP enhancements. The polarization of nuclei for detection (in our case ^{13}C and ^{15}N with lower gyromagnetic ratios) is increased by both DNP and CP to enhance the NMR signal. A continuous-wave microwave is irradiated on the biradicals added in the solution. The polarization of electron spins is first transferred to protons through DNP effects. The hyperpolarization of ^{1}H is then transferred to the nuclei for detection through CP. During the detection state, ^{1}H is decoupled.

Two-dimensional (2D) NMR is an approach to collect scalar[140] or dipolar coupling[141] information between spins, while this information is typically hard to be resolved in 1D NMR experiments. In addition, resonances can overlap on the 1D NMR spectra of transmembrane proteins, as many resonances from the same type of amino acids could have similar chemical shifts. 2D NMR can provide a better resolution by extending the spectrum toward a second dimension. The off-diagonal cross-peaks on 2D NMR encode extra information on nuclei that are close in space and coupled, which together with resonance assigning experiments can help us to identify the resonances from different parts of the protein.

2D NMR correlation experiments are done with four stages: preparation, evolution, mixing, and detection stage. The time of evolution stage t_1 is varied to create the second frequency dimension. During the mixing time, coherences can be transferred between the same spin, which ends up being a diagonal peak, or between different spins, which end up being an off-diagonal cross-peak. An example of a 2D ssNMR experiment is the dipolar assisted rotational resonance (DARR)[142] used in this dissertation. It is a dipolar-based experiment to detect through-space interactions between different ^{13}C spins on a transmembrane protein. Figure 1.14 is the pulse sequence in a DARR experiment. The magnetization of ^{1}H is transferred to ^{13}C through CP, and then the chemical shift of ^{13}C is evolved during the evolution time t_1. The magnetization of ^{13}C is then flipped to z-direction for mixing. The magnetization can be transferred between the coupled ^{13}C if they are close enough to each other in space. A longer mixing time can lead to more coupled ^{13}C and can result in more cross-peaks on the final spectrum. In DARR experiments, ^{1}H is recoupled with ^{13}C during the mixing state to assist the magnetization transfer between ^{13}C spins. This leads to a more efficient transfer of magnetization by comparing to another similar 2D ssNMR technique proton-driven spin diffusion (PDSD). The final detection stage is also done with ^{1}H decoupling and the FID of ^{13}C during the detection time t_2 is encoded with the ^{13}C-^{13}C dipolar coupling information. The obtained array of FID taken under different evolution time t_1 is two-dimensional Fourier transformed to get a 2D contour plot. With shorting mixing time, the off-diagonal correlations on the spectra are generally from nuclei that are one chemical bond away within the same residue. With longer mixing time, some correlations between nuclei that are further away in space, including nuclei that are on different residues but close in place could be resolved.

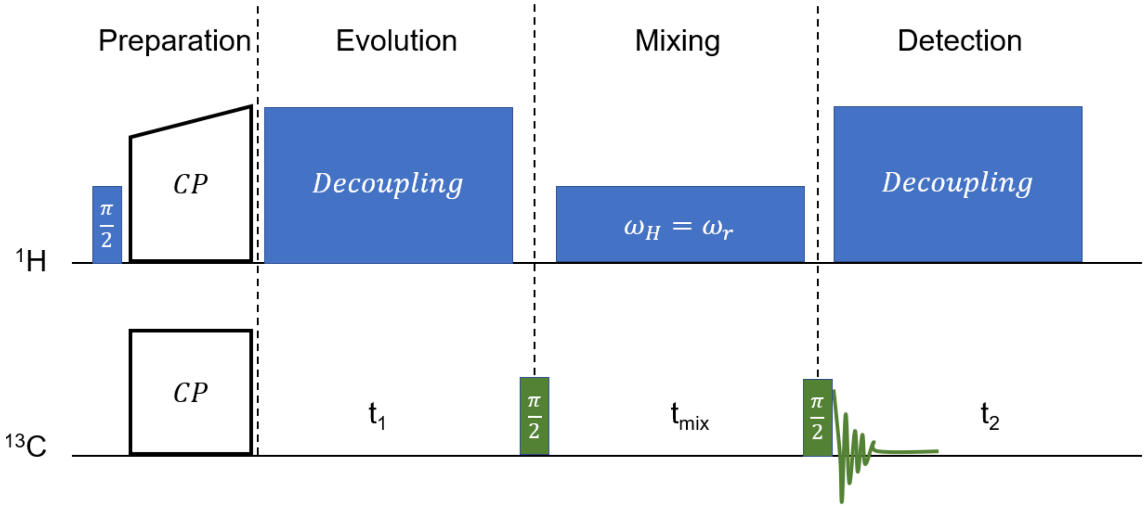


Figure 1.14: Pulse Sequence of ^{13}C-^{13}C DARR. The magnetization of ^{1}H is first transferred to ^{13}C through CP. The chemical shift of ^{13}C is evolved during the evolution time t_1, then a 90° pulse is applied to flip the magnetization of ^{13}C to the z-direction and enter the mixing stage. During the mixing stage, ^{1}H is recoupled with ^{13}C to assist the transfer of polarization from one carbon to others that are close in space through dipolar interactions. Another 90° pulse is applied to flip the magnetization of ^{13}C back to the xy plane for detection.

Chapter 2

Functional Consequences of PR Oligomerization in Lipid Membrane

This chapter discusses the impact of PR oligomeric assembly on its proton transport function in a lipid membrane environment. Two important aspects of PR proton transport function described in Chapter 1 are examined, including the protonation behavior of the primary proton acceptor D97 that controls the population of active PR and the photochemical reaction cycle kinetics that determine the rate of proton transport. PR was found to have both hexamers and pentamers as the dominate oligomeric species.[67-69 70-71] The functional impacts of PR hexameric and pentameric assemblies have been comprehensively studied in the detergent micellar environment.[74] Those studies found a lower pKa of the primary proton acceptor (pKa_{D97}) and a slower photochemical cycle dynamics for pentameric/oligomeric PR compared to its monomeric form while PR was reconstituted in DDM detergent micelles. The lower pKa_{D97} (~6.7) of hexameric/pentameric implied a larger fraction of PR has deprotonated D97, corresponding to the active state of PR, at the higher pH ~8 of the native ocean environment. Another study suggested the type of surfactant used for reconstituting PR could also play a role in modulating pKa_{D97} to affect the activity of PR.[80] In native-like E.coli membrane and proteoliposome environment, the existence of both hexameric and pentameric PR was found by SDS-PAGE with crosslinkers to stabilize the oligomeric structure.[143]

Nevertheless, it is still unclear the functional role of hexameric and pentameric PR while they are in the lipid membrane environment.

This chapter first investigates the oligomeric state of PR within a native-like *E.coli* lipid membrane and a proteoliposome environment by blue native polyacrylamide gel electrophoresis (BN-PAGE), which is a technique that can preserve protein oligomeric assemblies during gel electrophoresis without using chemical crosslinking. In the case of E.coli membrane lysates, Western blotting was used after the gel electrophoresis to immunodetect PR that has a 6x His-tag. To answer the outstanding question of whether oligomeric and monomeric PR still has a functional difference in lipid bilayer environments, we used a mutation, W34D, identified to disrupt a cross-protomer interaction within PR's functional core,[62, 73] as well as a mutation, E50Q, known to weaken the physical association between PR in oligomers .[72] The functions of these PR mutants that represent monomeric PR were compared to WT PR with predominantly oligomeric PR in liposome environment.

2.1 Formation of PR Oligomers in Lipid Membrane Environment

2.1.1 *E.coli* Cell Lysate

The first question we ask is whether PR within the native membrane of the *E.coli* host exists in an oligomeric state. The cell lysates from *E. coli* overexpressing both WT PR and the E50Q mutant of PR, a mutation thought to disrupt oligomers formation,[72] along with a negative control from *E. coli* that is not overexpressing PR were analyzed by gel electrophoresis techniques. First, we used blue native polyacrylamide gel electrophoresis (BN-PAGE), a common technique to identify the oligomeric states of PR (Figure 2.1 A).[72, 144] The total protein stain of the BN-PAGE gel by using Coomassie blue dyes was not sufficient to resolve differences between the three conditions, and therefore a western blot was applied subsequently to the BN-PAGE using a chemiluminescence antibody that binds specifically to the His-tag of PR. This method revealed the existence of a high molecular-weight PR species that corresponds to an oligomeric form in lysates containing WT PR (lane 1, Figure 2.1 B). Lysates from cells overexpressing PR E50Q presented a much smaller, likely monomeric form of PR (lane 2, Figure 2.1 B). As expected, no signal was detected in the negative control sample without PR (lane 3, Figure 2.1 B). These results reinforced our previous observation using SDS-PAGE with crosslinkers that the oligomeric form of PR presence in the native-like *E. coli* membrane environment[74] and the ability of the E50Q mutation to disrupt oligomers.[72] Furthermore, the ~200 kDa apparent molecular weight of the native PR form (lane 1, Figure 2.1 B) suggests the predominance of a hexamer or pentamer species compared to other possible oligomers (e.g. dimer or trimer), while the resolution of western blotting here is not sufficient to provide the information on whether either or both PR oligomeric species presence in the native *E. coli* membrane environment.

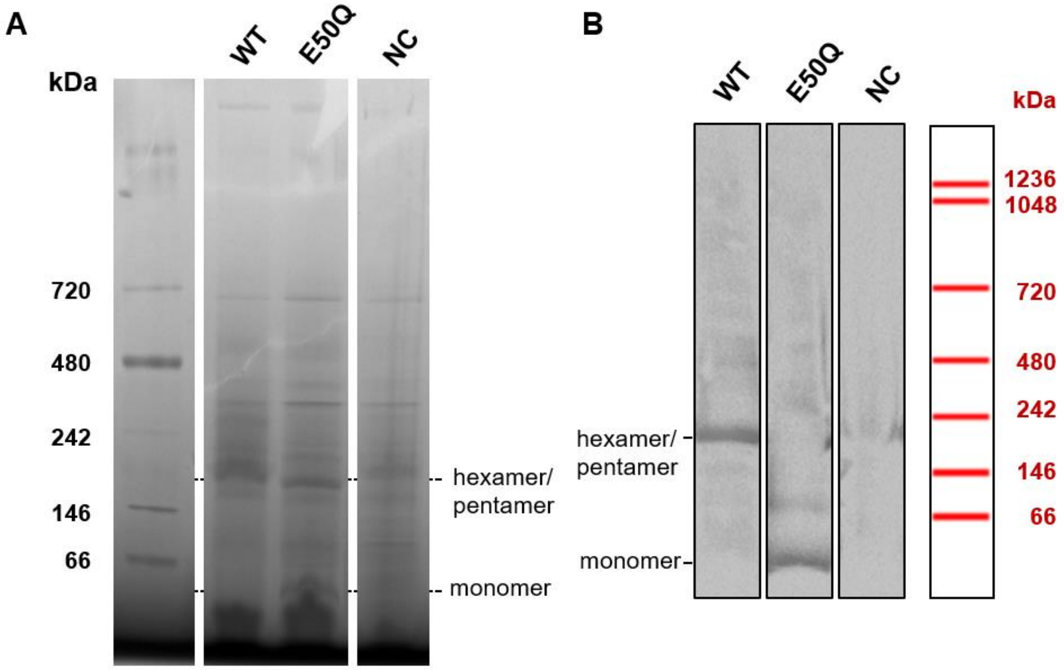


Figure 2.1: BN-PAGE and Western Blot of PR in E.coli Cell Lysate. A. BN-PAGE of cell lysates from *E. Coli* overexpressing WT PR (lane 1), PR E50Q (lane 2), and negative control (NC) from lysates without overexpressing PR (lane 3). **B.** Western Blot analysis of proteins transferred from a BN-PAGE of cell lysates from *E. Coli* overexpressing WT PR (lane 1), PR E50Q (lane 2), and negative control (NC) from lysates without overexpressing PR (lane 3).

2.1.2 Liposomes

We next evaluated the oligomeric state of PR in a reconstituted lipid system. We used POPC/POPG liposomes (80/20, mol/mol) to reconstitute PR, as the anionic PG headgroup lipids would give the membrane surface a negative charge similar to that of the native bacterial membrane, while POPC would impart membrane fluidic properties. Also, past work has shown that the orientation of PR in this exact membrane composition is more uniform than for neutral bilayers (with its N-terminus exposed on the outside of the proteoliposomes).[145] Upon performing BN-PAGE on purified PR in DDM micelles (lane 1, Figure 2.2) and reconstituted into POPC/POPG liposomes (lane 2, Figure 2.2), both samples exhibit a predominant band around the same size as observed with WT PR pentamer or hexamer in *E.*

coli membrane lysates (Figure 2.1 B). There are also several larger species present in DDM micelles, which may be due to nonspecific interactions between multiple PR-DDM micelle complexes, or simply due to impurities that are removed during the reconstitution procedure (lane 1, Figure 2.2). This result agrees with what was observed for PR reconstituted in liposomes with a different composition of DMPC/DMPG (9:1, w/w) that both hexameric and pentameric PR exists.[72]

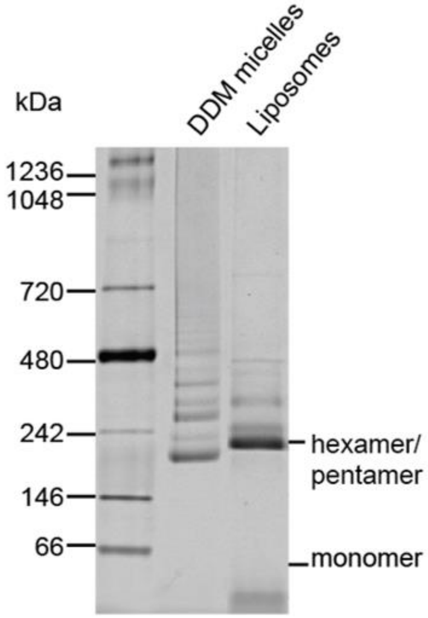


Figure 2.2: BN-PAGE of WT PR in Liposomes. BN-PAGE of purified PR in DDM surfactant micelles (lane 1) and reconstituted into POPC/POPG (80/20, mol/mol) (lane 2).

2.1.3 Oligomeric Assembly of PR in Liposomes

A predominance of oligomers, in both hexameric and pentameric forms, was found while WT PR was reconstituted in liposomes. A cryo-EM study on PR solubilized in Cymal-5 detergent micelles argued the existence of hexameric PR to be an artificial oligomeric form caused by the His-tag added on the PR construct for purification purpose.[96] They found a predominance of sole pentamer from a PR construct without additional His-tag by using BN-PAGE analysis. We also used a tobacco etch virus (TEV) enzyme that can cleave the 6x His-tag on the C-terminus of PR after the affinity column purification. In contrary to what was found in the literature, our BN-PAGE analysis found both hexameric and pentameric form of oligomers from WT PR in either the DDM detergent environment or the POPC/POPG liposome environment, regardless the existence of His-tag or not (Figure 2.3). In the BN-PAGE, PR with its 6x His-tag before the cleavage reaction (lane 1), PR with its 6x His-tag upon mixing with the TEV enzyme (lane 3), and PR with its 6x His-tag after the cleavage reaction in both DDM detergent (lane 4) and POPC/POPG liposome (lane 5) environments were compared. The only notable difference caused by the His-tag removal is a slight shift of bands that could be the molecular weight difference of the His-tag between the two PR constructs (lane 1 vs. lane 4, Figure 2.3). We also analyzed the TEV enzyme on its own (lane 2), and no significant bands appeared as its molecular weight (~25 kDa) is outside the range represented here. A similar result was also found in the literature that used a PR construct with three native cysteines preserved, where both hexameric and pentameric form of oligomers exist before and after the His-tag removal.[72] These all together fortify the statement that the hexameric form of PR oligomers represents a native form and is not an artifact. The difference in dominant oligomeric species between studies could be caused by the difference of biomimetic environment used for reconstituting PR[62] or the difference of protein purification methods.[96] The first argument is supported by a finding from detergent-solubilized PR, where the concentration and type of detergent can significantly affect the oligomeric distribution of PR.[72, 80]

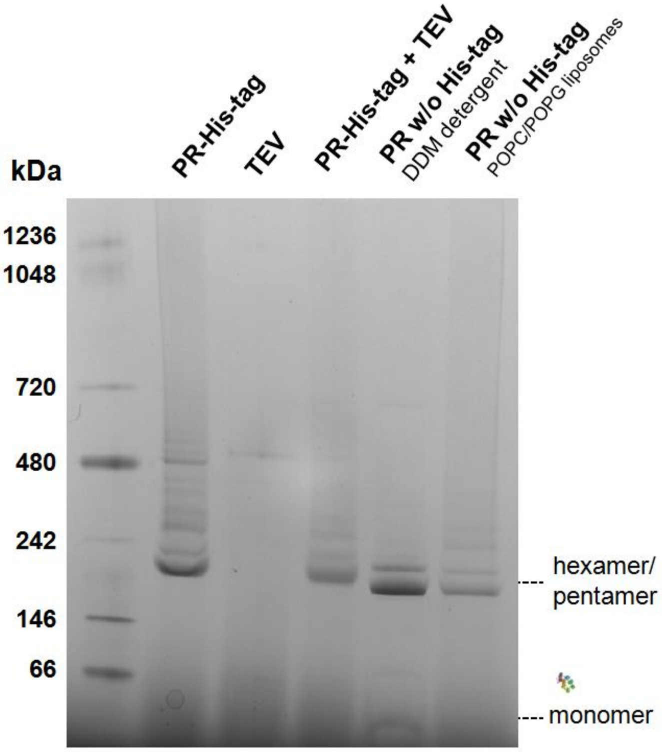


Figure 2.3: Oligomeric Distribution of PR with His-tag removed. BN-PAGE of WT PR before the His-tag removal (lane 1), during the His-tag removal process by adding TEV protease (lane 3), after the His-tag removal process reconstituted in a buffer supplemented with 0.05 wt% DDM detergent (lane 4), and after the His-tag removal but reconstituted in POPC/PG liposomes. No bands can be observed from pure TEV protease (lane 2) in the range displayed because of its smaller molecular weight (~25 kDa).

We also found that the compactness of liposomes could affect the oligomeric distribution of PR between hexamer and pentamer, and this finding could support the hypothesis that the biomimetic environment used for reconstituting PR can also affect the ratio between pentameric and hexameric PR. By replacing 30 mol% of POPG with POVPC that has the same zwitterionic headgroup but a reoriented sn-2 hydrocarbon chain to make the lipid bilayer leakier,[146] a significant decrease of hexameric PR population can be observed from the BN-PAGE (Figure 2.4). This result further supports that the surrounding environment used to reconstitute PR can influence the distribution of PR oligomers between its pentameric and hexameric form. Overall, both hexameric and pentameric form of PR oligomers naturally exists in both native-like *E.coli* membrane and synthetic liposome environment, supported by the gel electrophoresis analysis results. Since it is hard to separate hexameric and pentameric PR due to their similar size, in the latter part of this chapter, we referred both hexameric and pentameric PR as oligomeric PR, and the functions of these oligomers were compared to the monomeric PR.

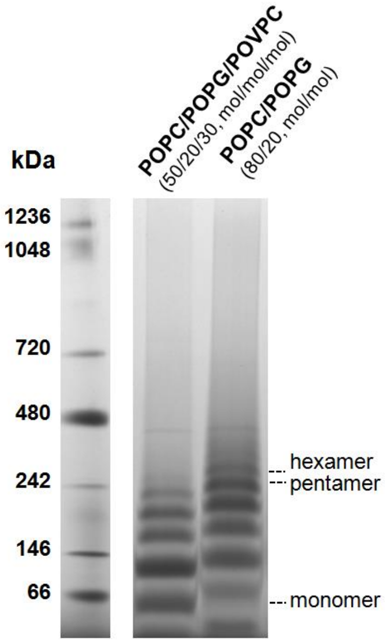


Figure 2.4: BN-PAGE of PR in leakier liposomes. BN-PAGE of PR E50Q reconstituted into POPC/POPG (80/20, mol/mol) liposomes (lane 1) and liposomes with 30 mol% of zwitterionic lipids substituted by POVPC oxidized lipids with a reoriented hydrocarbon chain that makes the lipid bilayer leakier.

2.2 Functional Studies of PR Oligomers

We next seek to understand the functional role of oligomeric PR within the lipid membrane. Our prior study of PR within detergent micelle environments indicated that the oligomeric state has a more prominent effect on PR's functional properties (light-absorption and photoactivation kinetics) than the surfactant type.[74, 80] Here, we applied previously established measures to characterize PR functions, now within a more physiologically-relevant proteoliposome environment.

2.2.1 pH-dependent Color Transition

The first benchmark used to quantify PR function is the apparent pKa_{D97}, which we consider to be indicative of the population of active PR molecules capable of transporting a proton from the cell interior to its exterior. The measured apparent pKa_{D97} of PR has been found to vary from 6.5 to 8.0 depending on the experimental condition[50, 82-83, 147] and the fitting method used to obtain pKa_{D97}.[78] Notably, in detergent environments, the pKa_{D97} of monomeric PR in detergent separated by SEC is 7.4-7.8, significantly higher than oligomeric PR (6.5 – 6.7).[74] The pKa_{D97} of monomeric PR in detergent is remarkably close to the pH of the ocean (8.2[48, 148]), implying that theoretically only half of the PR molecules would be active (with deprotonated D97). In contrast, a majority population of oligomeric PR in detergent micelles is deprotonated at the physiologically-relevant pH, suggesting oligomer formation could be a plausible mechanism used by an organism to modulate PR's proton transport function. However, this has yet to be explored for PR in the more native-like lipid environment. Since PR tends to form oligomers in both DDM detergent micellar and proteoliposome environments and concentrating enough monomeric WT PR for proteoliposome formation process is challenging as shown in the past study, an alternative method is applied here to create a monomeric version of lipid-embedded PR.

Here we introduced the E50Q mutation, replacing a key acidic residue at the oligomeric interface,[72] to populate the monomeric PR in proteoliposomes. Upon BN-PAGE of PR E50Q purified in DDM (lane 1, Figure 2.5 A), we found a mixture of oligomeric and monomeric PR, consistent with the literature.[72] The trend is kept for PR E50Q reconstituted in POPC/POPG lipid vesicles (lane 2, Figure 2.5 A). It is noted that all the lower-order oligomers including the monomer from the PR E50Q in liposomes migrate slightly slower than the one in DDM, as was observed with the WT samples (Figure 2.2); the difference could be attributed to altered size and shape properties of PR in different environments. The E50Q mutation has enriched lower-order oligomers including the monomer compared to WT PR in POPC/POPG liposomes (shown in Figure 2.2), enabling a comparison of possible functional differences between PR monomer and oligomers.

The pKa_{D97} of monomeric-enriched PR E50Q was compared to oligomeric WT PR (Figure 2.5 B). Within DDM detergent micelles, the E50Q mixture exhibits a high pKa_{D97} (7.2), similar to the WT monomer or E50Q monomer separated via SEC, both of which have a pKa of 7.4 (triangle, Figure 2.5 C). Therefore, in detergent monomer-like function seems to dominate the E50Q mixture, as expected for this oligomer-breaking mutation. Interestingly, however, upon reconstitution of PR E50Q into lipids, while BN-PAGE verifies the presence of a mixture of oligomeric species including monomers (lane 2, Figure 2.5), the pKa_{D97} (square, Figure 2.5 B) was similar to the case of oligomeric WT PR in proteoliposomes (square, Figure 2.5 B), where a pKa of ~6.4 was observed in both cases. This result suggested that the oligomeric distribution of PR in the liposome environment did not affect the pKa_{D97} as in the detergent micellar environment.

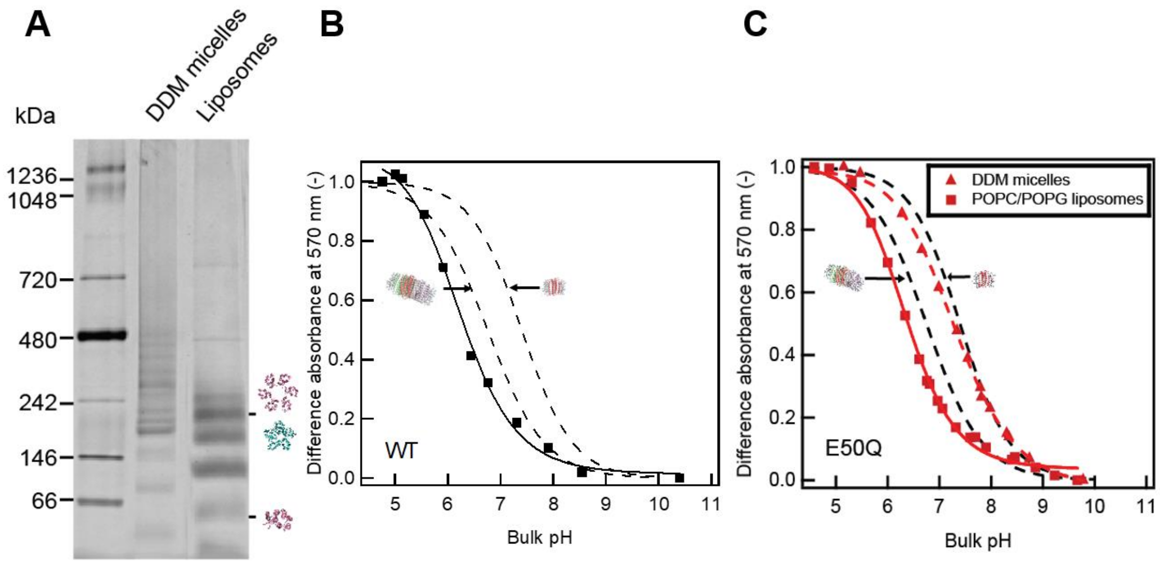


Figure 2.5. PR E50Q in Liposomes. A. BN-PAGE of PR E50Q in DDM detergent micellar environment or POPC/POPG (80/20, mol/mol) liposome environment. pH-dependent absorbance transition at 570 nm of **B.** WT PR and **C.** PR E50Q within DDM detergent micelles (triangle) and reconstituted into POPC/POPG (80/20, mol/mol) liposomes (square). Illustrated titration curves that represent oligomeric and monomeric WT PR within DDM detergent are added to B for comparison (black dashed lines).

To further verify that the pKa_{D97} of PR in liposomes is insensitive to the formation of oligomers, we then applied another kind of mutation at the oligomer interface to explore the functional attributes of PR self-association, as visualized by the blue PR crystal structure.[62] In this structure, site W34 at the oligomer interface was seen to interact via hydrogen bond with residue H75 of an adjacent PR, which itself interacts with the D97 proton acceptor residue and is thought to be responsible for the relatively high pKa_{D97} of PR.[62, 73] We, therefore, introduced the W34D mutation seeking to break this function-relevant intermolecular contact without physically disrupting oligomer formation, as done with the E50Q mutation. BN-PAGE verified the predominant population of hexamers and pentamers for purified PR W34D in DDM detergent micelles (lane 1, Figure 2.6 A). Its pKa_{D97} (7.3) is much higher than WT PR (6.7) and rather similar to the pKa_{D97} found for monomeric PR reconstituted in DDM (triangle, Figure 2.6 B). This result indicates that the intermolecular W34-H75 interaction identified in the blue PR crystal structure[62] can indeed modulate the pKa_{D97} of the green PR oligomer to establish its low experimentally determined value (6.7), consistent with the literature.[73] PR W34D maintained the predominant population of hexamers and pentamers when reconstituted into POPC/POPG liposomes (lane 2, Figure 2.6 A). Remarkably, however, the pKa_{D97} enhancing effect is not reproduced upon the reconstitution of PR W34D in POPC/POPG liposomes. Rather, we observe the same low pKa_{D97} (6.6) as was measured for both WT and PR E50Q (square, Figure 2.6 B) in proteoliposomes. The lipid environment therefore apparently compensates somehow for the absent W34-H75 interaction between adjacent PR.

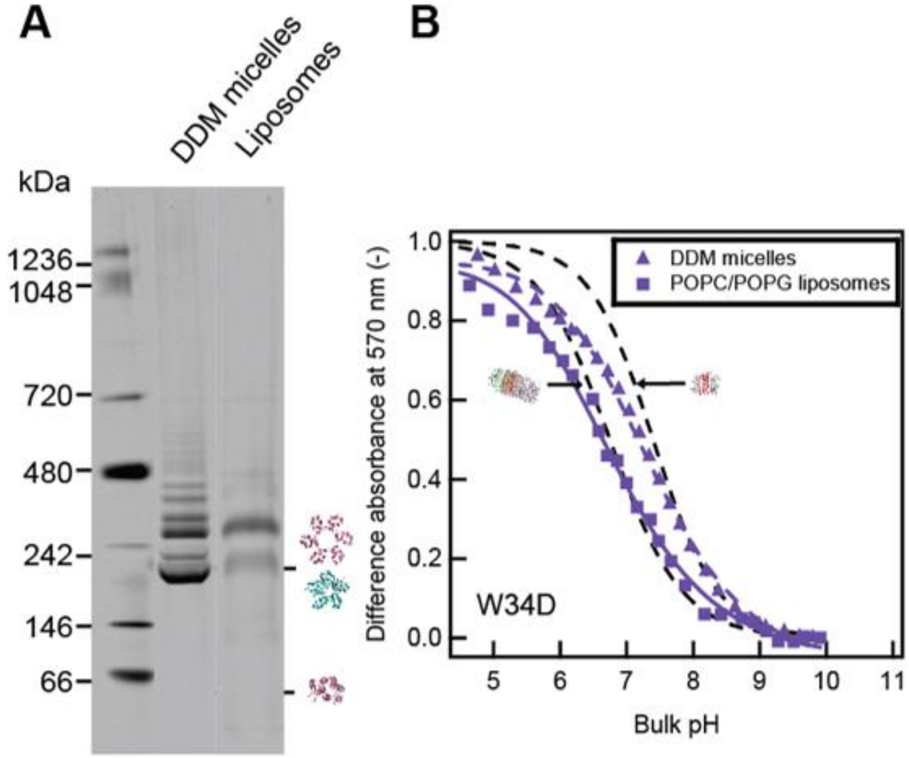


Figure 2.6: PR W34D in Liposomes. A. BN-PAGE and **B.** pH-dependent absorbance transition at 570 nm of PR W34D within DDM detergent micelles and reconstituted into POPC/POPG (80/20, mol/mol) liposomes. Illustrated titration curves that represent oligomeric and monomeric WT PR within DDM detergent are added to B for comparison (black dashed lines).

In proteoliposomes, neither the E50Q mutation that physically disrupts the formation of hexameric and pentameric PR nor the W34D that alters the cross-protomer W34-H75 interactions in hexameric or pentameric PR was found to affect the pKa of PR's proton acceptor residue (D97). Together with the pKa_{D97} measured from hexameric/pentameric and dimeric WT PR that show minimal variations (6.4-6.5), our results suggest that the pKa_{D97} is not modulated by PR's oligomeric state in proteoliposome environments, which is in stark contrast to what was found with PR reconstituted in DDM detergent micelles.[74, 80].

2.2.2 Photochemical Reaction Cycle

We next examined the key PR function by interrogating the kinetics of its proton transport function. Such kinetics has been widely studied by tracking the time-resolved optical absorbance change after the excitation of PR with a pulsed green-light laser.[53-54, 84] Photoactivated PR undergoes a series of conformational changes that perturb the local environment of the retinal chromophore, which in turn affects PR's overall optical absorbance. Several photointermediates with distinct maximal absorption wavelengths have been identified that constitute the photochemical reaction cycle. These include the transient conformational states sequentially labeled as K, M_1/M_2, N, and PR' intermediates, each of which has a signature absorbance spectral peak at 555 nm, 410 nm, 560 nm, and 520 nm, respectively. The detail of these has already been described in Chapter 1 of this dissertation. By tracking the growth and the decay of difference absorbance at wavelengths close to the wavelengths of photointermediates, the kinetics of each step in the proton pumping reaction cycle can be mapped out. In detergent environments, oligomerization has a substantial effect on PR's photocycle kinetics, with monomeric PR displaying a 2-3 times faster rate compared to pentameric/hexameric PR.[74, 80] Here we characterized PR's photochemical reaction cycle for various oligomeric states within POPC/POPG liposomes by comparing WT PR to E50Q mutant, which enriches the monomer and lower-order oligomers (Figure 2.5 A).

A different light-activated transient absorbance response was observed between WT PR and the monomer-enriched PR E50Q within proteoliposomes, revealing that the photochemical reaction cycle kinetics is tuned by oligomerization. The main observable differences are both the difference absorbance at 410 nm (purple solid lines, Figure 2.7 A vs. B) and 590 nm (orange solid lines, Figure 2.7 A vs. B), associated with the accumulation and decay of M and K/N intermediates in the photochemical reaction cycle, respectively. The differential absorbance at 410 nm peaked at longer timescales for oligomeric WT PR (~500 μs) compared to monomer-enriched E50Q PR (~200 μs) (highlighted by dotted lines). This result establishes that for the PR oligomers in proteoliposomes, the M intermediate accumulates in greater quantities relative to other intermediates. Furthermore, the difference absorbance at 590 nm (orange solid lines, Figure 2.7) reached its second local maximum intensity later for oligomeric WT PR (~4 ms) than monomer-enriched PR E50Q (~2 ms), accompanied by a significantly lower intensity for the oligomer, indicating that the self-association of PR causes it to populate the N intermediates more slowly.

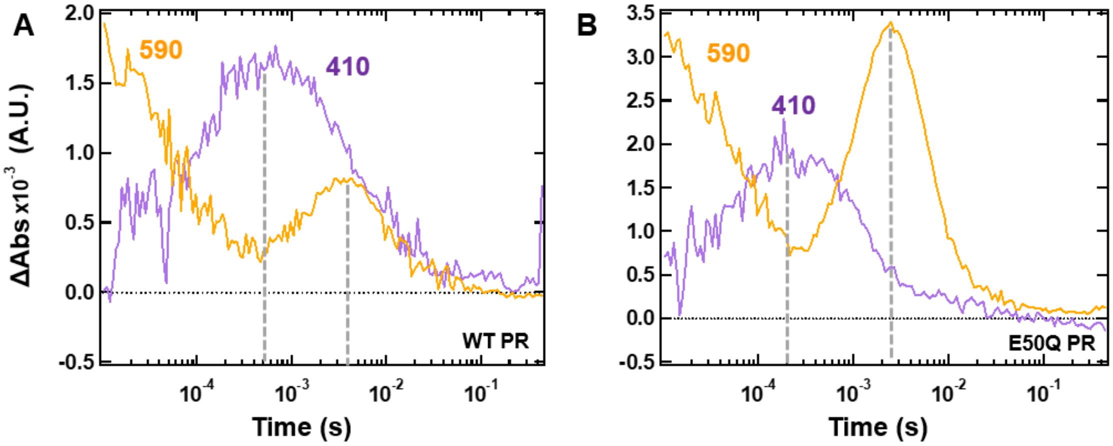


Figure 2.7: Transient absorbances at 410 and 590 nm of PR in liposomes. Transient difference absorbance data of (A) WT PR and (B) PR E50Q reconstituted in POPC/POPG (80/20, mol/mol) liposomes. Measurements were performed at pH 8.5 and ~293 K. The transient absorbance changes at 410 nm (purple) and 590 nm (orange) were collected after PR is photoactivated by a green-light pulse laser. The gray vertical lines indicate the approximate timings for the transient absorbance at 410 nm and 590 nm to reach the local maximum.

Differences can also be observed at other wavelengths (470 nm, 500 nm, 550 nm, and 630 nm) between WT PR and monomer-enriched PR E50Q (Figure 2.8 A vs. B). As the growth and decay of transient absorbance at these wavelengths could be due to combinations of different intermediates or lower resolution (particularly at 630 nm), it is difficult to attribute the differences to any individual photointermediate. Overall, our results qualitatively agree with what has been observed for PR in the DDM detergent micelle environment: hexameric PR populates the M intermediates and the N intermediate at a slower rate compared to monomeric PR.[80]

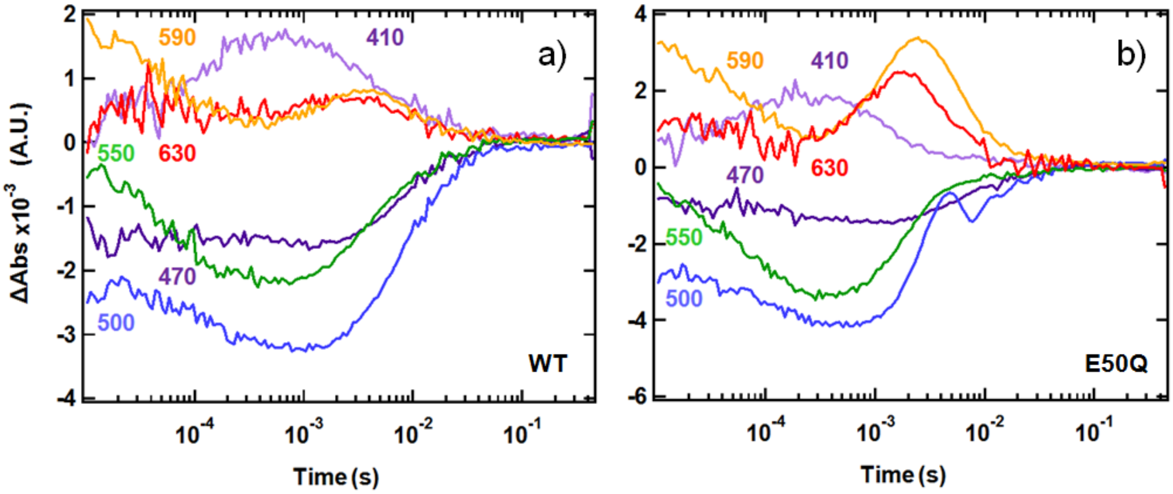


Figure 2.8: Transient Absorbance of PR in Liposomes. Transient difference absorbance data of **A.** WT PR and **B.** PR E50Q in POPC/POPG (80/20, mol/mol) liposomes. Measurements were performed at pH 8.5 and ~293 K. The transient absorbance changes after PR is photoactivated at 470 (deep purple line), 500 (blue line), 550 (green line), and 630 (red line) nm were added here on top of the 410 nm (light purple line) and 590 nm (orange line) in Figure 2.7.

A more quantitative analysis of the transient absorbance data at 410 nm for both WT PR and PR E50Q within liposomes was carried out by fitting the traces with a biexponential model.[149] Here, we presumed the difference absorbance at 410 nm is mainly contributed by the blue-shifted M intermediates, with contributions from both M_1 and M_2 intermediates combined, as the absorbance of other intermediates (e.g. K, N, PR, PR’) is significantly smaller at this wavelength.[54] With this assumption, the growth and decay of the difference spectra at 410 nm represent the accumulation and decay of the M intermediates. A simplification is made by considering the reverse reactions of the K-M and M-N transitions to be negligible. As a result, a biexponential model for a simple 1st-order two-step consecutive reaction model can be used to fit the differential absorbance at 410 nm:

$$\Delta Abs_{410nm} = a\frac{k_1}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t}), \tag{2.1}$$

where a is a parameter that reflects the magnitude of the differential absorbance, k_1 represents the rate constant of the K-M transition step, and k_2 represents the rate constant of the M-N transition step. The difference absorbance data at 410 nm from Figure 2.7 A and B (purple lines) were fitted by using the above reaction model (smooth black curves). The K-M transition rate constant k_1 for oligomeric WT PR reconstituted in liposomes was 8.7 ms^{-1}, ~2 times slower than k_1 for the monomer-enriched PR E50Q ($k_1 = 16.9\, ms^{-1}$). Furthermore, a ~4.5 times slower M-N transition rate constant k_2 was found for oligomeric WT PR reconstituted in liposomes ($k_1 = 0.11\, ms^{-1}$) compared to monomer-enriched PR E50Q ($k_2 = 0.5\, ms^{-1}$). Interestingly, when the same analysis was applied to WT PR and PR E50Q in the DDM detergent environment, a similar trend was observed. The transient absorbance data at 410 nm and 590 nm for both WT PR and PR E50Q in the DDM detergent environment are shown in Figure 2.9 A and B, and the transient absorbance at 410 nm were fitted by the same biexponential model. The values of fitting rate constants for PR reconstituted in lipids vs. detergent were slightly different, with all rate constants to be slightly faster for the same PR in the detergent than in the liposome environment. However, the k_1 and k_2 of oligomeric WT PR in DDM ($k_1 = 10.1\, ms^{-1}, k_2 = 0.16\, ms^{-1}$) were also ~2 times and ~4.5 times slower, respectively, than monomer-enriched PR E50Q in DDM ($k_1 = 21.7\, ms^{-1}, k_2 =$

$0.77\ ms^{-1}$). These fitting results further confirmed that PR oligomerization modulates its photocycle kinetics, specifically the accumulation and decay of the M intermediates, both for PR embedded within a liposome environment and within the detergent environment of our previous study.[74, 80]

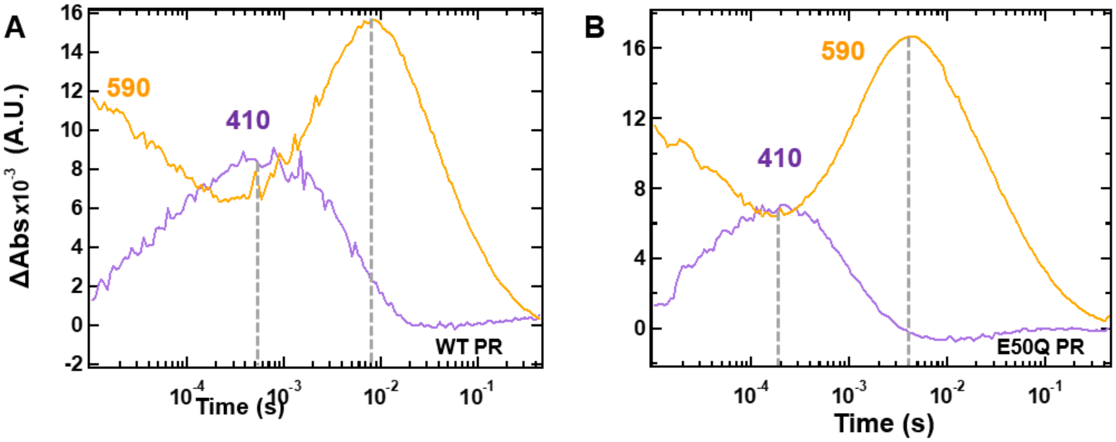


Figure 2.9: Transient absorbances at 410 and 590 nm of PR in DDM detergent micelles. Transient difference absorbance data of **A.** WT PR and **B.** PR E50Q reconstituted in DDM detergent micelles. Measurements were performed at pH 8.5 and ~293 K. The transient absorbance changes at 410 nm (purple) and 590 nm (orange) were collected after PR is photoactivated by a green-light pulse laser. The gray vertical lines indicate the approximate timings for the transient absorbance at 410 nm and 590 nm to reach the local maximum.

On top of a substantially slower accumulation and decay of the M intermediates for oligomeric-enriched WT PR than the monomeric-enriched PR E50Q in the liposome environment, differences in the kinetics of later transitions in the photocycle can also be found between PR reconstituted in the POPC/POPG liposome environment and the DDM detergent environment. A faster transition from the N intermediate back to the ground state PR was found for both WT PR and PR E50Q in the liposome environment. While the difference absorbance at 590 nm for PR in the liposome environment recovered back to 0 between 0.01 and 0.1 s after light-activation, the difference absorbance at 590 nm for PR in the DDM detergent environment was still in the progress of converging back to zero at 0.1 s after its light-activation (Figure 2.7 and Figure 2.9). Notably, oligomeric WT PR integrated with liposomes (Figure 2.7 A) has a significantly smaller population of N intermediate compared to WT PR oligomers in the DDM detergent environment (Figure 2.9 A). The corresponding low intensity at 590 nm, even lower than the relative intensity at 410 nm in the millisecond range, is unique among all the other conditions tested in the current study (oligomeric WT PR in DDM detergent, and monomer-enriched PR E50Q in both DDM detergent and POPC/POPG liposome environments). This difference between oligomeric WT PR in the two environments could result from a relatively fast N-PR' and PR'-PR transitions in the liposome environment, so that the N intermediate cannot accumulate to form a large second local maximum absorbance at 590 nm. These results suggest that the choice of a more native-like liposome environment compared to the DDM detergent environment for reconstituting PR can affect the N-PR' and PR'-PR transitions in the photocycle, while the oligomerization of PR mainly slows down the accumulation and decay of the M intermediates in both environments tested. A detailed structural study on both PR oligomers and monomer in the DDM detergent environment and liposome environment, possibly comparisons of various trapped photointermediate in different environments, will be required in the future to explain the mechanisms of these observed kinetic differences.

2.3 Discussion: Functional Consequences of PR Oligomers

This chapter showed that oligomerization remains an important factor for tuning PR function in near-native lipid bilayer environments. Our gel electrophoresis data have verified the predominance of oligomeric PR, both hexamers and pentamers, in both native-like *E.coli* membrane and synthetic liposome environments. Based on the photocycle kinetic comparison between WT PR and PR E50Q that inhibits oligomer formation in the lipid membrane environment, the oligomeric state of PR can tune the part of the photocycle that involves the M intermediates as has been shown in detergent environments, with a ~2 times slower K-M intermediates transition rate and ~5 times slower M-N intermediates transition rate for oligomers-enriched WT PR. As the transitions between M and other photointermediates were proposed to be the steps where PR experiences the greatest structural movement,[150-152] the observed delay in M intermediate population may be due to a particular protein conformation that exists for oligomers to restrict the K-M and M-N transitions.[62, 73] The H75 and W34 cross-protomer contact has been suggested to be one of the interactions in PR oligomers to affect its structural movement, specifically the tautomerization movement of H75 found between the ground state and the cryo-trapped M intermediates. A more rigorous structural dynamic study will be required to resolve how the slower K-M and M-N transitions observed here for oligomeric WT PR is connected to this H75 tautomerization movement. Whatever the structural basis of the tuning mechanism on the photochemical reaction cycle is, our results suggest it should be different from the ones determining the protonation state of D97.

In contrast to what was found in the DDM detergent environment,[80, 143] the pKa_{D97} that controls the population of active PR did not change with varying oligomeric distributions in the lipid bilayer environments, as long as the electrostatic environment of the lipid bilayer and the surrounding solution was kept constant. The different functional outcomes of oligomerization between the surfactant micellar environment and the lipid bilayer environment reinforce the importance of selecting the appropriate biomimetic platform for the structure-function studies of transmembrane proteins, as showcased here on the exemplary study of PR in this dissertation. Our next goal was to find the dominance factor that determined the pKa_{D97} while PR was reconstituted in the liposome environment.

Chapter 3

Role of Lipid Membrane on the pKa of Embedded D97 (pKa_{D97})

This chapter explores the key factors that modulate the pKa of embedded D97 (pKa_{D97}), a switch that determines the population of active PR with the proton transport capability from the inner to the outer part of the membrane. This D97 is not located on the surface but buried inside the protein [153]. Generally, the pKa of such an embedded ionizable residue inside a transmembrane protein is different from the pKa of the same amino acid in aqueous solution, and the value could vary from protein site to site for the same amino acid and is highly dependent on the local environment around the embedded residue [154-155]. This pKa generally shifts in the direction that favors the neutral form of the amino acid when these residues are embedded in the hydrophobic interior of the protein, as the D96 residue inside bacteriorhodopsin has a high pKa and is mostly protonated [156]. Interestingly, the pKa of embedded residues could be affected by the environment in which the transmembrane protein resides.[78] As the oligomerization of PR reconstituted in liposomes was found not to significantly affect the pKa_{D97}, demonstrated in Chapter 2, there must be certain factors in the lipid environment as the determinants to modulate pKa_{D97}. Therefore, the goal of this chapter is to identify the key external environmental factors that determine the pKa of embedded ionizable amino acids. The results in this chapter have been previously published in *Biophysical Journal*.[99]

Most of the past studies on the pKa of embedded ionizable amino acids were done on a water-soluble protein staphylococcal nuclease (SNase) as a model system.[155, 157-162] These studies mainly focused on how the pKa of an ionizable amino acid (e.g. lysine) could change while it was introduced to a different part of the SNase by [163] by applying continuous constant pH molecular dynamics (CpHMD) or the virtual mixture of multiple states (VMMS) method.[164-167] Despite the pKa of these embedded ionizable amino acids estimated by the simulation could be different from the one measured by NMR techniques,[166, 168-172] these simulation studies still managed to identify local environmental factors including ion penetration in the vicinity [159, 173-174], charge-charge interactions with other ionizable groups [160, 163, 175], charge-polar interactions [176], and hydrogen bonding of the embedded charged residue can affect the pKa of embedded ionizable residues on the SNase.[49] In the case of transmembrane proteins, charge-charge interactions between ionizable residues or water interactions were also suggested to be the determinants of embedded ionizable residue's pKa, shown by a case study on bacteriorhodopsin (bR).[163] Overall, these studies indicated that the microenvironments around the embedded ionizable residue in proteins can sensitively modulate its pKa, but did not identify how the external environment outside the protein can affect these microenvironments to affect the pKa of these embedded residues.

Transmembrane protein is embedded in a lipid bilayer with a densely charged surface formed by lipid headgroups and a low dielectric interior made of hydrophobic lipid hydrocarbon chains.[8, 177] The protonation state of embedded ionizable residues in transmembrane proteins reconstituted in such an environment could potentially be affected by the charge of lipid headgroups, the orientation of water at the lipid bilayer surface [178], the population of penetrated water inside lipid bilayers, or ion condensation around the surface of lipid bilayers including protons [163]. In fact, it has been suggested that differences in the local proton concentration on a charged lipid bilayer surface, which give rise to a spatially varying surface proton concentration compared to the bulk pH, affect the apparent pKa of embedded D85 in bR [179-182]. Motivated by biophysical understanding objectives, this chapter tested the potential factors listed above to identify the key external environmental factors in the lipid membrane environment that determine the pKa_{D97} of PR. We also asked whether the deviation of surface proton concentration from the bulk pH on liposomes with different electrostatics is the only major factor that affects the embedded pKa_{D97}, or in fact the external factors also affect the microenvironment around the embedded charge through structural rearrangement to change its intrinsic pKa, as suggested by the simulation studies.

In this chapter, we systematically altered the electrostatic environment of PR reconstituted in liposomes by varying the salt type and concentration inside the aqueous solution and the net charge of the liposome headgroup composition and examined the effect of these factors on affecting the apparent pKa_{D97}. The intrinsic pKa_{D97} that takes the surface pH difference into account was also examined. The comparison of the apparent and the intrinsic pKa_{D97} revealed whether both the surface pH deviation and the protein intrinsic properties are in charge of determining the pKa_{D97}.

3.1 Apparent pKa_{D97} altered by electrostatics

We examined the effect of the external electrostatic environment on the apparent pKa_{D97} of PR by changing the net charge of liposomes and the concentration of ions in buffer solutions. The net charge of liposomes was adjusted by mixing POPC with a zwitterionic headgroup with either POPG with a negatively-charged headgroup or DOTAP with a positively-charged headgroup (Structures shown in Figure 3.1 A). The apparent pKa_{D97} of PR was measured in the negatively-charged POPC/POPG (80/20, mol/mol) liposomes and positively-charged POPC/DOTAP (80/20, mol/mol) liposomes in the same 10 mM HEPES buffer, while successively increasing the NaCl concentration from 0 to 150 mM. In the case of negatively-charged POPC/POPG (80/20, mol/mol) liposomes, pKa_{D97} decreased gradually from 7.62 ± 0.02 to 6.43 ± 0.01 when the NaCl concentration in the buffer was increased (blue, Figure 3.1 B). In contrary, an opposite incremental trend of pKa shift was observed in the case of positively-charged POPC/DOTAP (80/20, mol/mol) liposomes while NaCl concentration was increased from 0 mM ($pKa_{D97} = 5.63 \pm 0.01$) to 150 mM ($pKa_{D97} = 6.14 \pm 0.02$) in the HEPES buffer (red, Figure 3.1). The sharpest pKa_{D97} difference, which is an up to ~2 pH unit change, between the negatively-charged POPC/POPG and positively-charged POPC/DOTAP liposomes can be observed when there is no additional NaCl in the HEPES buffer. With the presence of more NaCl in the buffer, the difference of measured apparent pKa_{D97} between PR in these two types of liposomes with opposite net charges became smaller.

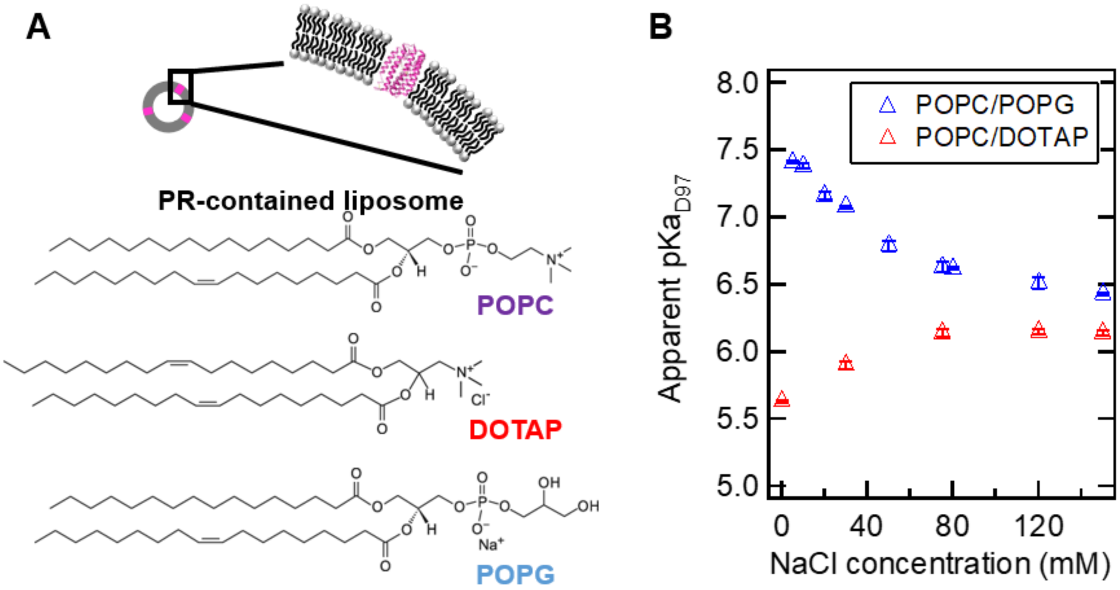


Figure 3.1: Apparent pKa$_{D97}$ of PR in liposomes with different electrostatics. A. Schematic diagram of PR-contained liposomes and chemical structures of the lipids used in this study: zwitterionic POPC, positively-charged DOTAP, and negatively-charged POPG. **B.** Apparent pKa$_{D97}$ of WT PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes (blue triangle) and positively-charged (80/20, mol/mol) liposomes (red triangle) in a 10 mM HEPES buffer. The HEPES buffer contained different concentrations of NaCl between 0 mM and 150 mM.

Besides the apparent pKa$_{D97}$, we also examined whether the titration curves under these electrostatic environments have similar behavior. If the deviation of surface pH from the bulk pH on charged liposome surfaces in the low NaCl concentration range was the main or sole factor to affect the spectral titration curves, a distortion in the titration curves at low NaCl concentrations can be observed. In fact, as we examined all the spectral titration curves of the fitted apparent pKa$_{D97}$ shown in Figure 3.1 B, no qualitative distortion on the titration curves in the range of low ion concentrations can be observed for PR in both negatively-charged POPC/POPG (Figure 3.2 A) and positively-charged POPC/DOTAP (Figure 3.2 B) liposomes.

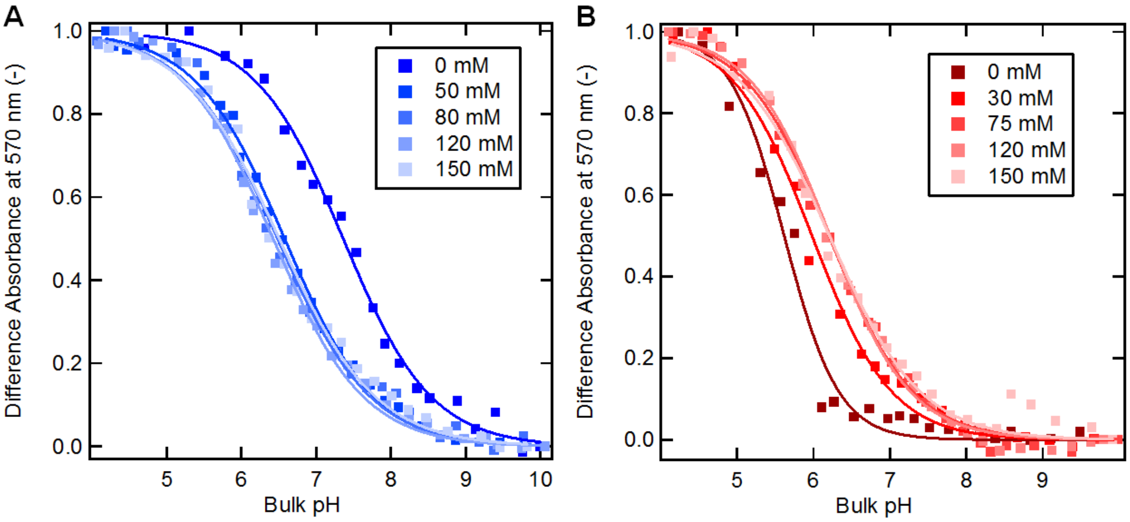


Figure 3.2: Spectral Titration Curves of PR in Liposomes with Different Electrostatics. Difference optical absorbance at 570 nm under various bulk pH from WT PR reconstituted in A. negatively-charged POPC/POPG (80/20, mol/mol) liposomes and B. positively-charged POPC/DOTAP (80/20, mol/mol) liposomes in HEPES buffer contained different concentrations of NaCl.

The non-distortion behavior of titration curves in low NaCl concentration range can also be quantitatively determined by another important fitting parameter, the Hill coefficient, in the Henderson-Hasselbalch fitting. This Hill coefficient n can describe whether the protonation or deprotonation of D97 on a PR can affect the protonation or deprotonation of D97 on another PR in the vicinity. The value of n is less than 1 while it has an anti-cooperative behavior (i.e. the protonation of D97 of a PR will make the protonation of D97 on another PR to be non-favorable), and is greater than 1 while it has a cooperative behavior (i.e. the protonation of D97 of a PR will make the protonation of D97 on another PR to be favorable). Besides the abovementioned characteristic, since this constant is determined by the slope of the sigmoid curve, it can also provide the information on whether the spectral titration curves were distorted by the deviation of surface pH if it was the dominant factor. In line with the qualitative observation in Figure 3.2, our data showed no dependence on the Hill coefficient on the ion concentration in the buffer under the range tested (Figure 3.3). These results suggested that the deviation of surface pH on charged liposome surface is not the main and sole factor to change the apparent pKa_{D97} for PR reconstituted in liposomes with different electrostatic environments.

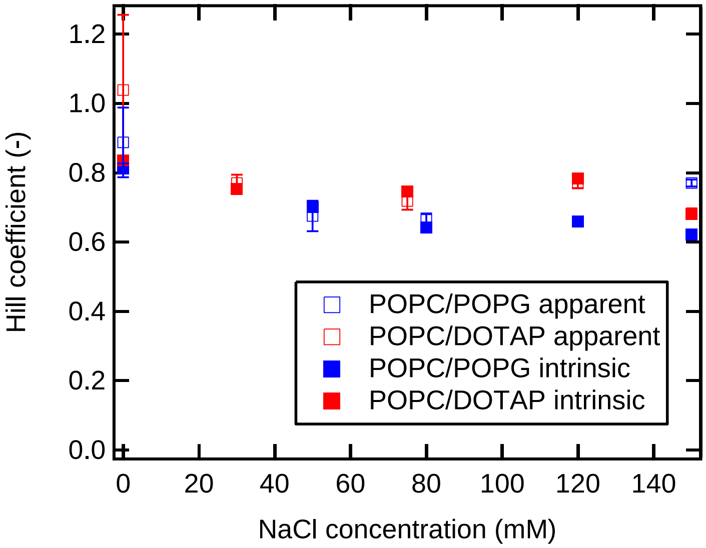


Figure 3.3: Hill Coefficient of PR in Liposomes with Different Electrostatics. Hill coefficients from the Henderson-Hasselbalch fitting for the apparent pKa_{D97} of WT PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes (blue) and positively-charged (80/20, mol/mol) liposomes (red) in a 10 mM HEPES buffer. The HEPES buffer contained different concentrations of NaCl between 0 mM and 150 mM.

We also examined the pKa_{D97} of PR reconstituted in liposomes constituted of pure zwitterionic POPC. The measured pKa_{D97} lied in between the ones from positively-charged POPC/DOTAP (80/20, mol/mol) liposomes and negatively-charged POPC/POPG (80/20, mol/mol) liposomes for both buffers with and without 150 mM NaCl (Figure 3.4). These results all follow a systematic trend, in which the apparent pKa_{D97} shifted toward a lower value when the charge of the liposome surface changed from net negative to net positive.

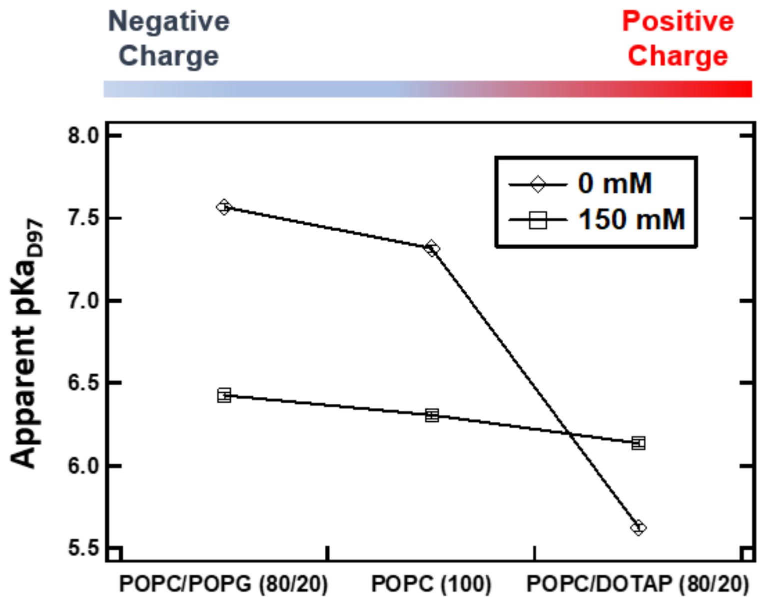


Figure 3.4: Apparent pKa$_{D97}$ of PR in Liposomes with Different Net Charge. Apparent pKa$_{D97}$ of WT PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol), pure zwitterionic POPC, and positively-charged POPC/DOTAP (80/20, mol/mol) liposomes in a HEPES buffer without additional NaCl (diamond) and the same buffer but with 150 mM NaCl (square).

3.2 Intrinsic pKa$_{D97}$ Altered by Electrostatics

To further prove that the observed pKa shift is not only an apparent effect dictated mainly by surface pH, we used the Gouy-Chapman (GC) model to calculate the surface pH on charged liposomes under the conditions tested in Figure 3.1 B (detail derivation included in Appendix C) [183-184]. After calculating the surface pH on charged liposome surfaces and re-plotting the spectral titration data against the surface pH instead of the bulk pH, we found that the intrinsic pKa$_{D97}$ obtained from the Henderson-Hasselbalch fitting still changed systematically under the electrostatic environment tested (solid triangle, Figure 3.5). The greatest deviation in the intrinsic pKa$_{D97}$ from the apparent pKa$_{D97}$ was found in buffers without additional NaCl. Furthermore, the intrinsic pKa$_{D97}$ at 0 mM NaCl is distinctly different for PR in POPC/POPG

(80/20, mol/mol) vs. POPC/DOTAP (80/20, mol/mol) liposome environments, after taking surface proton concentration difference into account (5.11 vs. 7.64). In the presence of more NaCl in the buffer, the intrinsic pKa$_{D97}$ shifts in the opposite direction to the apparent pKa$_{D97}$, while the two values converge. This shows us that the electrostatic effect exerted on PR's pKa$_{D97}$ by its liposome environment is screened out at high NaCl concentrations. If surface pH was the only factor to affect the measured apparent pKa$_{D97}$ in Figure 3.1 B, then comparable or similar intrinsic pKa$_{D97}$ would be expected in Figure 3.5 under all conditions tested. In fact, we found an opposite trend for the intrinsic pKa$_{D97}$ compared to the apparent pKa$_{D97}$, while we also found that the intrinsic pKa$_{D97}$ changes for PR reconstituted in liposomes with different electrostatic environments. Our calculations here reinforced the observed apparent pKa$_{D97}$ change is not solely an apparent effect caused by the deviation of proton concentrations on charged liposome surfaces compared to the bulk solution, but due to actual changes of its acid dissociation constant.

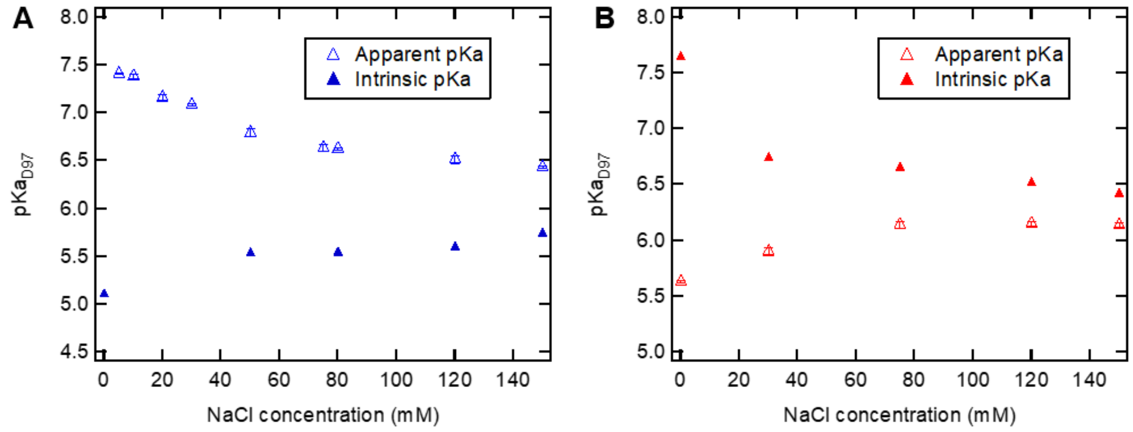


Figure 3.5: Intrinsic pKaD97 of PR in Liposomes with Different Electrostatics. Intrinsic pKa$_{D97}$ (solid triangle) of WT PR reconstituted in A. negatively-charged POPC/POPG (80/20, mol/mol) liposomes and B. positively-charged (80/20, mol/mol) liposomes in a HEPES buffer contained different concentrations of NaCl between 0 mM and 150 mM. The apparent pKaD97 (hollow triangle), adapted from Figure 3.1 B, is reproduced here as a reference.

It is also noted that all the spectral titration curves plotted against surface pH calculated by the GC model showed a similar shape without significant distortion (Figure 3.6), as they were plotted against bulk pH shown in Figure 3.2. This can also be further supported by quantitatively comparing the Hill coefficient from the Henderson-Hasselbalch fitting, where a similar Hill coefficient was found between the spectral titration curves plotted against bulk pH or surface pH under the same condition tested (Figure 3.3 vs. Figure 3.7). These results further support our hypothesis that the deviation of surface pH on charged liposome surface is not the only, nor main driving, factor affecting the apparent pKa_{D97} (Figure 3.1 B) for PR reconstituted in liposomes with different electrostatic environments.

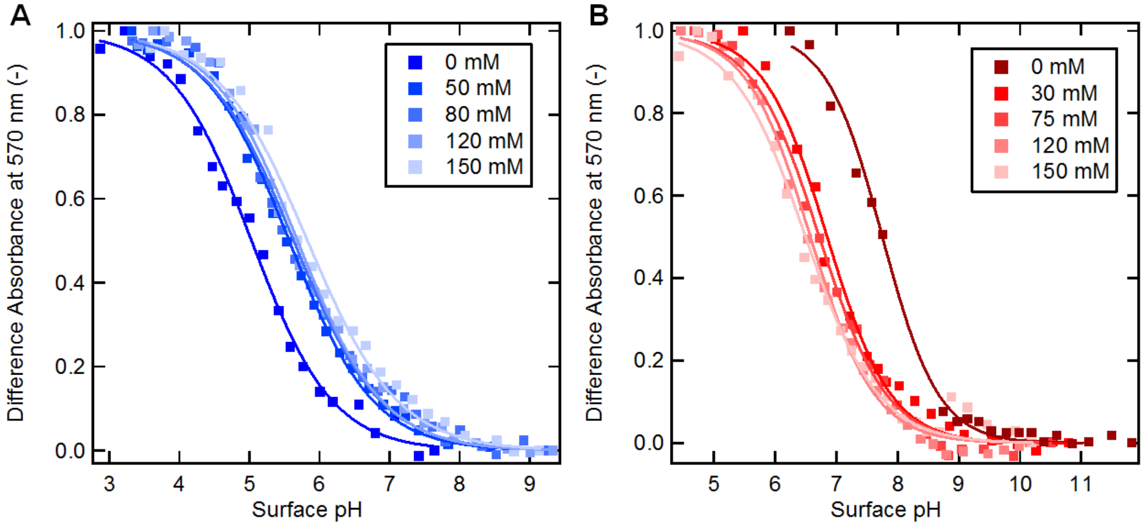


Figure 3.6: Spectral Titration Curves Against Surface pH of PR in Liposomes with Different Electrostatics. Difference optical absorbance at 570 nm under various surface pH from WT PR reconstituted in **A.** negatively-charged POPC/POPG (80/20, mol/mol) liposomes and **B.** positively-charged POPC/DOTAP (80/20, mol/mol) liposomes in HEPES buffer contained different concentrations of NaCl.

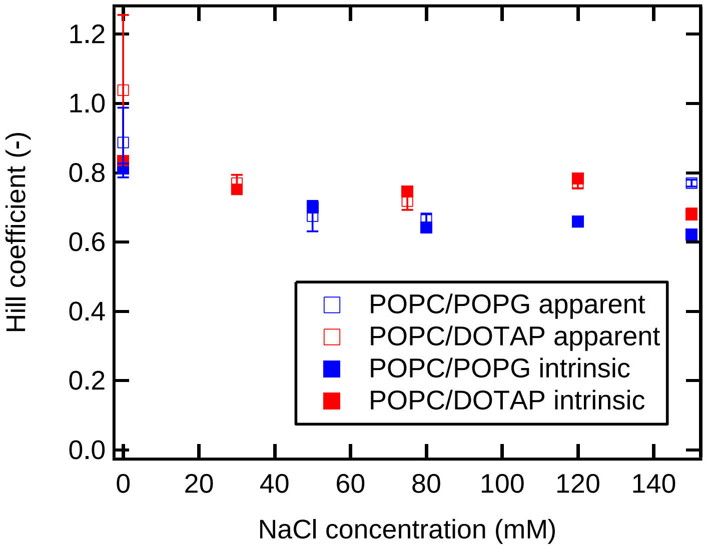


Figure 3.7: Hill Coefficient from Intrinsic pKa Fittings of PR in Liposomes with Different Electrostatics. Hill coefficients from the Henderson-Hasselbalch fitting for the intrinsic pKa$_{D97}$ of WT PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes (blue) and positively-charged (80/20, mol/mol) liposomes (red) in a 10 mM HEPES buffer. The HEPES buffer contained different concentrations of NaCl between 0 mM and 150 mM.

3.3 Effects of Other Lipid Membrane Characteristics

3.3.1 Lipid Hydration

As the addition of cations in the buffer can change both the extent of hydration and the electrostatic environments of PR-containing liposomes [185-186], some might attribute the observed pKa$_{D97}$ shift to be a result of the hydration change in liposomes. To clarify this, we examined several modulators that adjust only the extent of hydration of the PR-containing liposomes, one at a time. First, the PR-contained liposome was dehydrated by adding either propranolol, a pharmaceutical drug blocker to rigidify lipid headgroup region [187], or ethanol, which reduces transient water pore formation across lipid bilayers [188]. The addition of 1 mM propanolol to a HEPES buffer with 150 mM NaCl was found to have a negligible impact on the apparent pKa_{D97}, with a drop from 6.43 ± 0.02 to 6.36 ± 0.07 (Table 3.1). A similarly negligible pKa shift can be observed upon adding 5 wt% ethanol to the HEPES buffer containing 150 mM NaCl (pKa_{D97} = 6.75 ± 0.12). The extent of hydration of PR-containing liposomes was increased by substituting 30 mol% of the POPC with an oxidized lipid POVPC to test its effect on the apparent pKa_{D97}. The oxidized lipid POVPC with a reoriented sn-2 hydrocarbon chain facilitates water penetration into the lipid bilayer to increase its hydration[146, 189]. Again, no statistically significant difference was observed on the measured pKa_{D97} (6.36 ± 0.08 v.s. 6.43 ± 0.02). Overall, the applied modulators that are known to change the extent of hydration of PR-containing liposomes did not induce a significant apparent pKa_{D97} shift compared to the effect of salt concentration of the buffer. These observations suggest that the dehydration of lipid bilayers may not exert dominant effects on the pKa_{D97} inside PR, or that the experimental conditions and measurements are not sensitive to the relevant changes.

Table 3.1: Apparent pKa_{D97} under Different Lipid Hydration. Apparent pKa_{D97} of PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes in a HEPES buffer with 150 mM of NaCl under different hydration environments adjusted by different modulators.

Condition	Effect on liposome's hydration	Apparent pKa_{D97}
Control		6.43 ± 0.02
1 mM propanolol added in the buffer	Decrease[187]	6.36 ± 0.07
5 wt% EtOH added in the buffer	Decrease[188]	6.75 ± 0.12
30 mol% POVPC added in the liposome[a]	Increase[146, 189]	6.36 ± 0.08

[a] 30 mol% of POVPC were applied to substitute the POPG to make POPC/POVPC/POPG (50/30/20, mol/mol/mol) liposomes.

3.3.2 Lipid Membrane Fluidity and Curvature

Other possible factors related to the characteristics of the lipid membrane, including membrane fluidity and membrane curvature of liposomes, were also tested to find their potential effects on modulating pKa_{D97}. We changed these factors while keeping the charge of lipid headgroup the same and the resulting pKa_{D97} were similar to the ones from liposomes with the same headgroup net charge (Table 3.2). Both the pure zwitterionic DMPC liposomes that had a shorter lipid hydrocarbon chain length and the DOPC/DPPC that was a slightly more rigid membrane mixture with binary phases[190-191] showed a similar pKa_{D97} compared to the pure zwitterionic POPC membrane ($pKa_{D97} = 6.31 \pm 0.02$). The effect of lipid membrane curvature was examined by comparing DOPC/DOPA (80/20, mol/mol) liposomes to DOPE/DOPA (80/20, mol/mol) liposomes, where the DOPE with a smaller phosphoethanolamine headgroup compared to the phosphocholine headgroup has a negative spontaneous curvature. Even though lipid membrane curvature can affect both the conformation and the active population of human rhodopsin based on a flexible surface model, in the case of PR, the change of lipid membrane curvature did not significantly affect the apparent pKa_{D97} that determines the population of active PR.

Table 3.2: Apparent pKa$_{D97}$ under Different Membrane Rigidity. pKa$_{D97}$ of PR reconstituted in liposomes without changing the electrostatic environment.

Net Charge	Liposome	Apparent pKa$_{D97}$
Zwitterionic	Pure DMPC	6.21 ± 0.04
	DOPC/DPPC (50/50, mol/mol)	6.65 ± 0.05
Negative	DOPC/DOPA (80/20, mol/mol)	6.74 ± 0.02
	DOPE/DOPA (80/20, mol/mol)	6.86 ± 0.02

3.4 Photochemical Reaction Cycles

We also examined how the photochemical reaction cycle of PR is affected by lipid membrane characteristics that are either a key determinant of modulating pKa$_{D97}$ (e.g. ion concentration in buffer) or aren't the main factor in determining the pKa$_{D97}$ (e.g. lipid hydration). The light-activated transient absorbance of WT PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes in a HEPES buffer with or without an additional 150 mM NaCl were compared (Figure 3.9). While the apparent pKa$_{D97}$ of the condition without NaCl was higher than the one with 150 mM NaCl (7.62 vs. 6.43), both the transient absorbances at 410 nm (Figure 3.9), represent the growth and decay of M intermediates in the photochemical reaction cycle, and the transient absorbance at 590 nm (Figure 3.8), associated with the decay of K intermediate and the growth and decay of the late N intermediate, were nearly identical traces. This suggested that the surrounding electrostatic environment in which PR resides might not change its proton-pumping kinetics, at least under the condition tested. Nevertheless, it is still valuable to examine the transient absorbance of PR reconstituted in the positively-charged POPC/DOTAP (80/20, mol/mol) liposomes that yield a lower pKa_{D97} compared to the two conditions shown here.

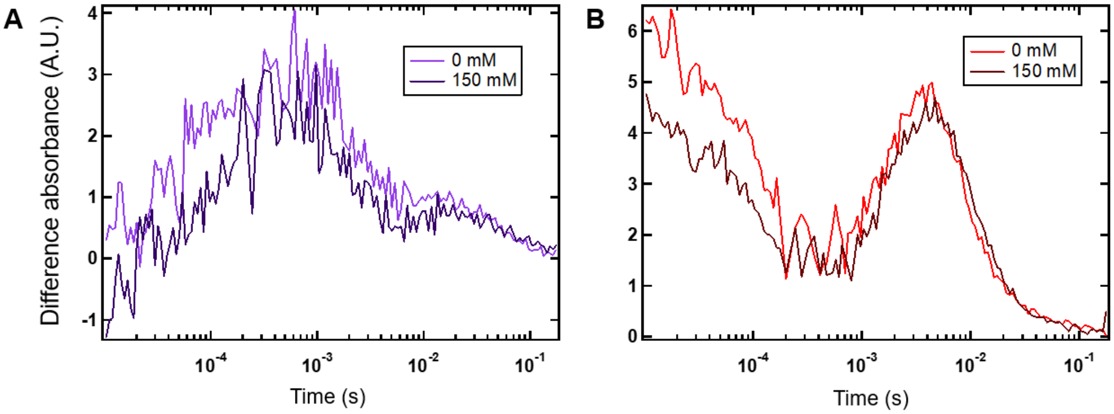


Figure 3.8: Transient Absorbance of PR in Liposomes with Different Electrostatics. Light-activated transient absorbance of WT PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes in a HEPES buffer with and without additional 150 mM NaCl. The transient absorbance was collected under **A.** 410 nm and **B.** 590 nm. The measurements were done at pH 8.5.

We also examined how adding a polycationic polymer polyethyleimine (PEI) in the system with PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes can affect the functions of PR, both the apparent pKa_{D97} and the transient absorbance under light-activation. This PEI polymer can absorb on the liposomes to increase the hydration of lipids by increasing the probability of water pore formation across the lipid membrane. Also, the polycationic polymer itself absorbed on the lipid membrane can change the net charge of liposomes from a net negative charge to a net positive charge. Compared to the control sample of PR reconstituted in negatively-charged POPC/POPG liposomes, the addition of PEI polymers at a concentration of 12 μg/mL reduced the apparent pKa_{D97} from 6.43 to 5.72 (Figure 3.9 A). This is in line with what was observed between the PR in liposomes with a net negative charge and a net positive charge, as the absorption of PEI on the POPC/POPG liposome surface can overturn its net charge. In the case of the photochemical reaction cycle, a slight difference can be found from the transient absorbance at 410 nm with the addition of 12 μg/mL PEI in the buffer (Figure 3.9 B). We applied the same biexponential model introduced in Chapter 2 to fit the growth and the decay of the transient absorbance at 410 nm, and a slight decrement on the k_1 that represents the growth of M intermediates accompanied with a similar k_2 that represents the decay of M intermediates were found with the presence of the PEI in the buffer. The results suggested even the electrostatic or the hydration environment of liposomes that modulated the apparent pKa_{D97} of PR can affect the transient absorbance of PR upon light-activation, the modulation effect was still not as significant as the one done by the oligomerization of PR presented in Chapter 2.

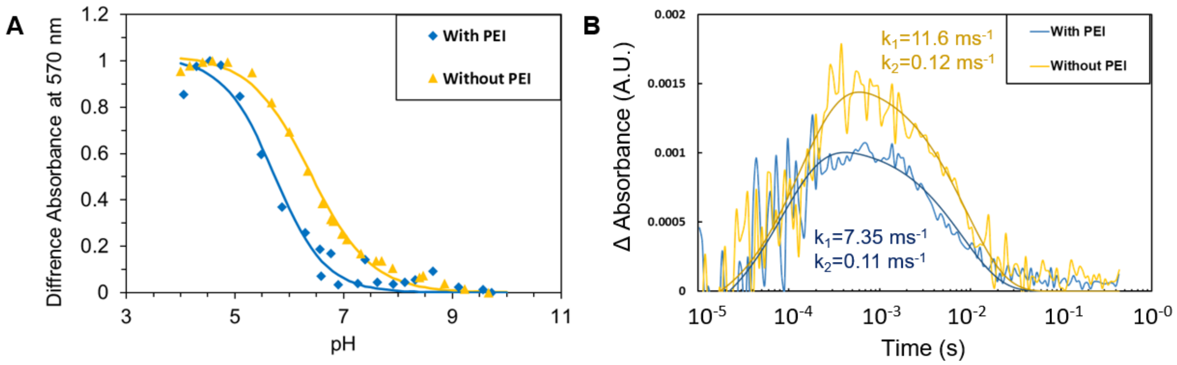


Figure 3.9: PR in Liposomes with Polycationic Polymers. A. Spectral titration curves of WT PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes with and without 12 μg/mL of polycationic polymer PEI in the 10 mM HEPES buffer with 150 mM NaCl. **B.** Transient absorbance at 410 nm measured for the same PR proteoliposome with and without the addition of PEI.

3.5 Discussion: The Role of Lipid Membrane on pKa_{D97}

Significant and systematic changes in both the apparent and the intrinsic pKa_{D97} were found for PR reconstituted in liposomes while we varied the net charge of phospholipid headgroups in the liposomes, as well as the ion type and their concentration in the buffer solution. On the contrary, a change of lipid hydration, lipid membrane curvature, membrane rigidity did not significantly affect the pKa_{D97}. These together suggest the electrostatic environment around PR reconstituted in liposomes should be the major factor determining its pKa_{D97}.

Other retinal protein studies also revealed a transition in the protonation/deprotonation equilibrium of its embedded proton acceptor that depended on the ion concentration in the buffer. These studies inferred that the transition is induced either by specific binding of ions to the membrane protein or by a deviation of the surface pH on the membrane surface. When adding 150 mM of NaCl into the buffer, we found an opposite trend for the pKa_{D97} shift between PR reconstituted in positively-charged liposomes and negatively-charged liposomes (Figure 3.1 B). This observation excludes ion-specific bindings as the key mechanism, as pKa shift in the same direction should occur with specific binding, regardless of the liposome's net surface charge. Furthermore, a much higher cation concentration was required to induce a measurable pKa shift in the here studied PR compared to bR. With bR, a complete transition of pKa was observed when 10 μM of Ca^{2+} was added to the purple membrane sample containing a bR concentration of 1.6 μM, which corresponds to only a six-fold molar excess ions [192-193]. In the current study, the same transition reaches saturation with Ca^{2+} concentration in the 10 mM range, corresponding to three orders of magnitude higher ion concentration compared to the PR concentration of 10 μM (Figure 3.4). On the other hand, although a steeper apparent pKa_{D97} change was found for divalent cations compared to monovalent cations, this could be due to a stronger association between the divalent cations and the negatively-charged POPG lipid headgroups compared to monovalent cations [194]. This stronger association may effectively screen the negative charge on liposome surfaces at lower ion concentration, and so unify surface and bulk proton concentrations. This result suggests the deviation of the surface from bulk pH due to interactions between proton and charged lipid surface to be one of the most plausible factors that affect the apparent pKa of embedded charge residue.

Interestingly, the intrinsic pKa_{D97} modeled with surface pH calculated for charged liposome surfaces using GC models showed an opposite trend from the apparent pKa_{D97}, as the net charge of liposomes and the concentration of NaCl in buffers were altered (Figure 3.5). Furthermore, intrinsic pKa_{D97} was found to be still significantly modulated under different external electrostatic environments. These observations reinforced that the surface pH on charged liposome surfaces cannot be the main factor to give rise to changes in the apparent pKa_{D97}. Even though the GC model might not quantitatively describe the surface potential of liposome systems with mixed lipids and in solution with different electrolyte concentrations, the difference between the experimentally determined surface potential ψ_0 for POPC/POPG (80/20, mol/mol) liposomes in a buffer with 50 mM electrolyte ($\psi_0 = -102\ mV$) reported in the literature [184] and the one predicted by the GC model ($\psi_0 = -72\ mV$) under similar conditions are still comparable, suggesting the GC-derived surface potential and pH to represent physical trends. In the future, a more careful study, probably directly measuring the surface pH on PR-contained liposomes, will be required to obtain a more precise surface potential and intrinsic pKa_{D97}.

Taken all observations together, we hypothesize that long-range interactions or mediated short-range interactions between charged lipid headgroups and the embedded charged residue inside protein must be at play, in addition to the effects of surface potential and pH, for altering the embedded apparent pKa_{D97}. While the value of apparent pKa_{D97} is important for practical applications, as it reports the actual distribution between active PR and inactive under a given environment and pH, the value of intrinsic pKa_{D97} sheds light on the state of protein intrinsic structural properties and has a greater biophysical significance. As a result, the structural properties of PR in these environments with different intrinsic pKa_{D97} were then examined to understand the mechanical basis of the observed modulation.

Chapter 4

Structural Rearrangement Associated with the Modulation of pKa_{D97}

Since the modulation of embedded pKa_{D97} by the external electrostatic environment of PR in liposomes was not solely an apparent effect dictated by the local proton concentration gradient, a structural-based study is required to understand the protein-intrinsic structural properties that are related to the observed change in intrinsic pKa_{D97}. In this chapter, complementary magnetic resonance spectroscopy tools were used to explore the possible structural rearrangement of PR in different electrostatic environments that reported different intrinsic pKa_{D97} in Chapter 3. The continuous wave electron paramagnetic resonance (cw EPR) and the Overhauser dynamic nuclear polarization (ODNP) relaxometry measurements on probing the local structure of E-F loop on PR, together with the results in Chapter 3, have been previously published in *Biophysical Journal*.[99]

The pKa of embedded ionizable residues (e.g. aspartic acid) has been suggested to couple with the structure of the entire protein in simulation studies. Most of these computation works used a water-soluble protein staphylococcal nuclease (SNase) as a model system and found that altering the protonation state of embedded ionizable amino acids inside the protein can induce structure rearrangement.[195-197] These studies also found evidence that showed a structural rearrangement of proteins can in turn affect the pKa of these ionizable amino acids, presumably through a change of microenvironment around the embedded ionizable residue. Another computational study on a catalytic enzyme Histone Deacetylase 8 found that the position of some residues remote from the active site of this enzyme can affect its activity by changing the enzymatic structure constructed by a zinc ion and five ionizable amino acids.[198] These studies all stated the importance of protein structure, including the structural water inside the protein,[199] on determining the protonation behavior of embedded ionizable residues inside the protein. Nevertheless, a mechanical understanding of how the external environment affects the internal structural properties of proteins and the microenvironment around the embedded ionizable residue to affect its pKa is still elusive.

To understand how the pKa of embedded D97 is modulated by the external environment through long-range interaction or mediated short-range interactions between lipid headgroups and amino acids, the structure of PR while reconstituting into liposomes with different net charge and equilibrated in buffer with different ion concentrations was examined by magnetic resonance techniques. These measurements were performed at a sufficiently high pH (8.5) where the majority of the D97 remains deprotonated to examine whether these factors that altered pKa_{D97} induce a structural rearrangement of active PR. Here the magnetic resonance spectroscopic tools applied to study the structure of PR can be categorized into the ones that focus on local structural information or the one that can provide global structural information of PR. The cw EPR lineshape analysis and ODNP relaxometry measurements can obtain local protein steric or water dynamic information, respectively, experienced by a conjugated nitroxide-based spin label on the selected solvent-exposed residue. Through site-specifically isotopic labeling on selected amino acids of PR, nuclear magnetic resonance (NMR) spectroscopy can be used to study the structural information of residues not limited to the solvent-exposed ones but also the embedded ones inside the hydrophobic core of PR, which is generally not accessible by EPR-based spin labels. On the other hand, NMR sensitive isotopes can also be uniformly labeled to entire PR to monitor the global structural information. In this chapter, we examined the structure of PR as it was reconstituted in liposomes under the four representative conditions of the pKa_{D97} modulation study presented in Chapter 3, which are negatively-charged POPC/POPG (80/20, mol/mol) liposomes or positively-charged POPC/DOTAP (80/20, mol/mol) liposomes in 10 mM HEPES buffers with 0 mM NaCl or with 150 mM NaCl. We first focused on two parts of PR structure that have shown to correlate with the pKa_{D97} value: The E-F loop that is remote from the embedded D97 and a conservative histidine H75 that can form a hydrogen bond with D97.[49, 62] Next, we investigated the entire structure of uniformly ^{13}C-labeled PR under the influence of external electrostatic environments through dipolar-based correlation experiments. Lastly, we attempted to mutate ionizable residues that could participate in the electrostatically coupled network between the embedded D97 and the externally charged lipid headgroup to see whether the mutation can affect pKa_{D97}.

4.1 Conformation Change of the E-F loop

cw EPR and ODNP were performed on PR with an MTSL spin label attached to site 174 on the E-F loop of PR (Figure 4.1), which is an interfacial site that has been utilized as a sensitive reporter to the light-activated structural movement of PR at the E-F loop.[95] Upon light-activation, the MTSL spin label conjugated on site 174 experienced a more hindering environment, leads to a change in its cw EPR spectra and a decrement of water dynamics around the spin label measured by ODNP relaxometry. Here, instead of light-activation, we utilized the same spin-labeled interfacial site to probe any possible structural rearrangement at the E-F loop of PR reconstituted in liposomes with different electrostatic environments to see whether we can capture any structural movement at the E-F loop.

Figure 4.1: Schematic Diagram of the E-F loop on PR. Nitroxide-based MTSL spin label is conjugated on site 174. This spin-labeled site 174 was shown to be located at the interface between the region with fast water dynamics (blue shading) and the region with slow water dynamics (red shading) as a sensitive reporter to the structural movement of PR at the E-F loop. The chemical structure of the MTSL spin label is also shown along the side.

4.1.1 cw EPR of the E-F Loop

X-band cw EPR spectra of PR spin-labeled at site 174 were measured and compared between two differently charged liposomes (negatively-charged POPC/POPG vs. positively-charged POPC/DOTAP) at two different NaCl concentration (0 mM vs. 150 mM) in HEPES buffer and pH = 8.5. The locations of the restricted immobile (“i”) component and the faster moving mobile (“m”) component are highlighted at the low-field region of the spectra (Figure 4.2). The relative spectral amplitude of these two components conveys the extent of tertiary contacts that the nitroxide label on site 174 experienced. The more steric contacts experienced by the nitroxide label, the higher relative amplitude of the restricted immobile (“i”) component was reflected in the cw EPR spectra. PR reconstituted in negatively-charged POPC/POPG liposomes in a HEPES buffer with 0 mM NaCl that showed the lowest intrinsic pKa_{D97} (5.11 ± 0.02) is the sample that revealed a dominance of the faster moving mobile ("m") component in the cw EPR spectrum. (Figure 4.2, top left). In contrast, the sample under the condition with the highest intrinsic pKa_{D97} (7.64 ± 0.01, positively-charged POPC/DOTAP, 0 mM NaCl) revealed a prominent immobile component ("i") and a concurrent reduction of the mobile component ("m") (Figure 4.2, top right). PR reconstituted in the same positively-charged liposome, but in the presence of 150 mM NaCl in the HEPES buffer, revealed a recovery of the mobile ("m") component (Figure 4.2, bottom right), although still less mobile compared to the condition that yielded the lowest intrinsic pKa_{D97} for PR. No significant cw EPR lineshape difference was observed when varying the NaCl concentrations (0 mM vs. 150 mM) for the PR sample reconstituted in negatively-charged POPC/POPG liposomes (Figure 4.2, blue spectra). No significant difference was observed between the spin-labeled variant and the WT PR. Taken together, these results show a trend of a cw EPR lineshape change at site 174 that coincides with the observed intrinsic pKa_{D97} shift, with a higher fraction of the immobile component when the PR has a higher intrinsic pKa_{D97}. These observations suggest that the same changes in the electrostatic environment that modulate the pKa_{D97} of the buried D97 residue also cause structural modulation in PR's peripheral E-F loop. We note that all PR-liposome samples studied by cw EPR were equilibrated at pH = 8.5, so that at least 88% or more of the buried D97 is deprotonated (even for the highest apparent $pKa_{D97} = 7.62$). Thus, the changes seen in cw EPR lineshapes are not a consequence of structural changes that follow the protonation or the deprotonation of D97. It is possible that the electrostatic environment couple to D97 is mediated by ordered surface water, and is reflected in the dynamics of spin labels on the E-F loop. Hence next, we examine the surface water diffusion dynamics around site 174.

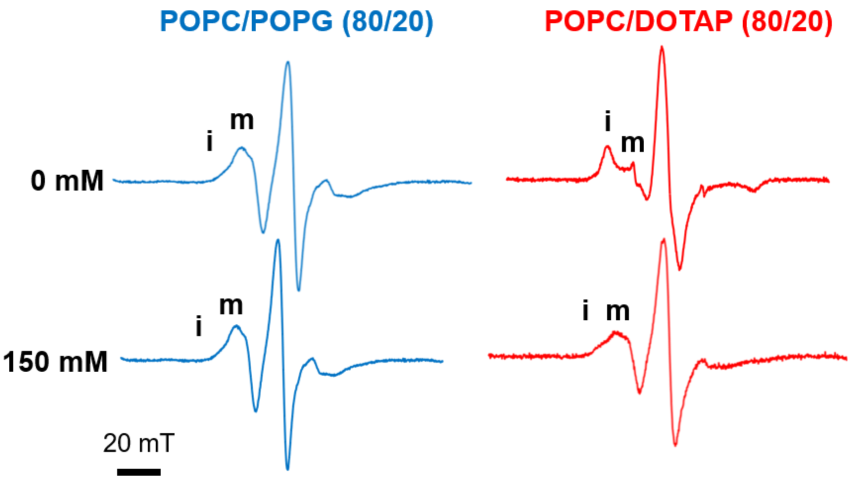


Figure 4.2: X-band cw EPR spectra of site 174 on PR. MTSL-labeled PR at site 174 reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes (blue spectra) or positively-charged POPC/DOTAP (80/20, mol/mol) liposomes (red spectra) in 10 mM HEPES buffers with 0 mM NaCl or with 150 mM NaCl. All the samples were equilibrated at pH 8.5.

4.1.2 ODNP of the E-F Loop

^{1}H ODNP relaxometry sensitive to changes of coupled surface water dynamics within ~10 Å of a nitroxide-labeled protein surface is used here to reveal changes in the hydration dynamics associated with the subtle structural movement around site 174 found by cw EPR lineshape analysis. The hydration dynamic information around site 174 was extracted from the electron-^{1}H spin cross-relaxivity (k_σ) that represents the dynamic of diffusing, loosely bounded (~tens of ps timescale) water. The k_σ at site 174 of PR under the same four conditions used in the cw EPR study were measured and compared (Figure 4.3). Interestingly, we found consistent trends for changes in k_σ with that in intrinsic pKa_{D97}. Overall, the condition (positively-charged POPC/DOTAP, 0 mM NaCl) with the highest intrinsic pKa_{D97} yielded the lowest k_σ among the conditions tested, which agrees with a more immobile cw EPR spectrum (top right, Figure 4.2). For PR reconstituted in positively-charged POPC/DOTAP liposomes, 150 mM NaCl in the HEPES buffer led to a higher k_σ (46.9 ± 1.1 s^{-1}) compared to with 0 mM NaCl (k_σ = 29.5 ± 4.0 s^{-1}), accompanying the decrease in intrinsic pKa_{D97} from 7.64 ± 0.01 to 6.42 ± 0.02 with 150 mM NaCl. For PR reconstituted in negatively-charged POPC/POPG liposomes, the presence of 150 mM NaCl in the HEPES buffer led to a lower k_σ (44.2 ± 2.5 s^{-1}) at site 174 in comparison to the condition without additional NaCl salt (k_σ = 52.6 ± 4.5 s^{-1}). This reduction in k_σ accompanies an increase in intrinsic pKa_{D97} from 5.11 ± 0.02 to 5.74 ± 0.02. Here, the ODNP measurements revealed a change in the hydration environment of site 174 that was not reflected in the cw EPR lineshape between these two conditions.

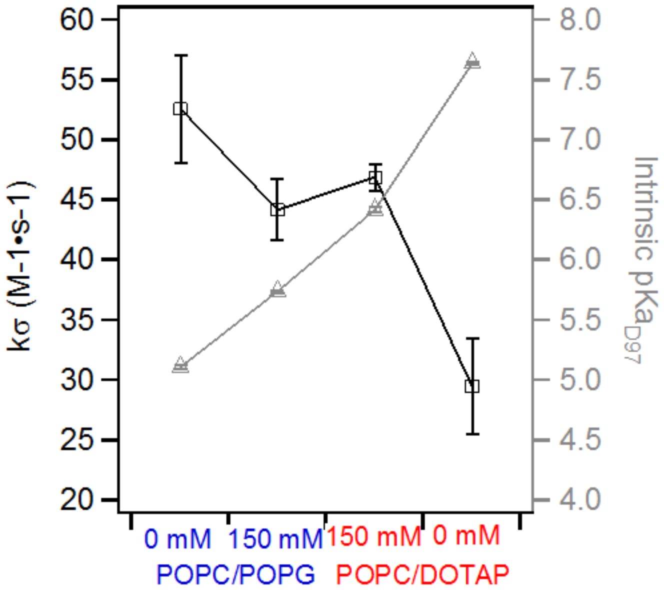
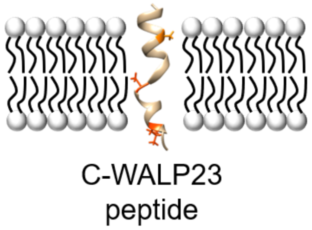


Figure 4.3: ODNP of Site 174 on PR. Cross-relaxivities k_σ that reflect the hydration dynamics (~tens of ps timescale) around MTSL-labeled site 174 measured by ODNP on PR reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes or positively-charged POPC/DOTAP (80/20, mol/mol) liposomes, in HEPES buffers with 0 mM NaCl or with 150 mM NaCl. The trend of intrinsic pKa_{D97} is also plotted along the side as a reference. All the samples were equilibrated at pH 8.5.

Control experiments were done on a model single α-helical WALP23 peptide to confirm the observed hydration dynamic change reflected solely a structural effect at the E-F loop of PR but not a change of hydration at the lipid-water interface with a different salt concentration in the buffer. The WALP peptide is commonly used as a model of transmembrane proteins embedded in a lipid bilayer.[200-202] The name of this peptide is an abbreviation of alternating hydrophobic alanines (A) and leucines (L) that form an α-helical transmembrane structure and four membrane-anchoring tryptophans (W) at the lipid-water interface. The WALP23 peptide used here had a cysteine residue located right above its α-helix (C-WALP23) at the lipid-water interface to mimic the location of site 174 on PR in lipid bilayers (schematic diagram and peptide sequence in Figure 4.4). This cysteine residue is conjugated with a MTSL spin label for measuring the hydration dynamics around the lipid-water interface also using the same ODNP technique.



Name	Peptide Sequence	MW
C-WALP23	Ac-C(sl)GWWLALA LALALALALALAL WWA-NH_2	2643.512

Figure 4.4: Schematic Diagram of C-WALP23. The cysteine residue conjugated with MTSL spin labels located on the top of the α-helix structure of WALP23 peptides mimics the location of the E-F loop on PR reconstituted in lipid bilayers. The peptide sequence and its molecule weight of the C-WALP23 are included in the inlet table.

The trend of k_σ change under different electrostatic environments measured from C-WALP23 was compared to the case of PR in Figure 4.3. For MTSL-labeled C-WALP23 reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes, 150 mM NaCl in the HEPES buffer led to a higher cross-relaxivity k_σ ($28.3 \pm 2.5\ s^{-1}$) compared to with 0 mM NaCl ($k_\sigma = 26.4 \pm 1.6\ s^{-1}$), in contrast to a drop of k_σ from $52.6 \pm 4.5\ s^{-1}$ to $44.2 \pm 2.5\ s^{-1}$ with the presence of 150 mM measured from site 174 on the E-F loop of PR. The opposite trend of hydration changes found here further consolidated that the observed decrement of k_σ on site 174 is caused by the change of local structure on PR, but not solely the change of hydration at the lipid-water interface.

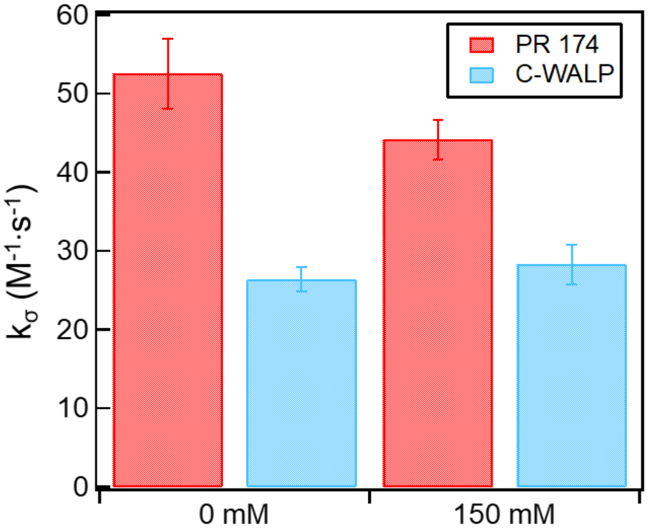


Figure 4.5: ODNP of C-WALP23. Cross-relaxivities k_σ of MTSL-labeled C-WALP and PR with an MTSL spin label on site 174 (same as the data in Figure 4.3) reconstituted in POPC/POPG (80/20, mol/mol) liposomes in a HEPES buffer with or without 150 mM NaCl. All the samples were equilibrated at pH 8.5.

Both cw EPR lineshape analysis and ODNP relaxometry measurements unveiled that the site 174 on the E-F loop move towards a more restricted region while PR is reconstituted in the environment with a higher intrinsic pKa_{D97}. In other words, the observed change in the intrinsic pKa_{D97} may be associated with the structural rearrangement of PR at the E-F loop induced by changing the electrostatic environment. Although this is the first study to unveil that the external electrostatic environment can trigger a structural rearrangement at the E-F loop, several studies have suggested that the structure of the E-F loop could be one of the factors to affect pKa_{D97}. An NMR study has revealed that substituting the alanine to a bulkier amino acid (e.g. arginine) at site 178 could unwind the α-helical E-F loop on the E-F loop and induce a shift of intrinsic pKa_{D97} toward a higher value.[62] Other studies also showed that the mutation at site 178 on the E-F loop can affect the maximum absorbance of the retinal chromophore and the intrinsic pKa_{D97}.[61, 63-64] The NMR study suggested that the structural rearrangement effect is not only a local effect at the E-F loop but can affect the chemical shifts of up to 31 residues from site 178, including the D97 residue.[62] Our observation here could also be another evidence to reinstate the importance of the structure at the E-F loop on affecting the structure and function of PR, specifically the pKa_{D97}.

Interestingly, the E-F loop movement that we found while PR was reconstituted into positively-charged POPC/DOTAP liposomes coincided with the light-activation movement found from PR populated with the M intermediates in its photochemical reaction cycle, where a more restricted environment was also experienced by the nitroxide-based spin label attached on site 174.[95] The M intermediates is the part of the photochemical reaction that the proton on the Schiff base K231 is transferred to the primary proton acceptor D97. This also means that the structure of PR, or more precisely the microenvironment around D97, in the M intermediates is energy-favorable for the D97 residue to be protonated, a shift of the intrinsic pKa_{D97} toward a higher value. Interestingly, a higher intrinsic pKa_{D97} was also found while PR was reconstituted into positively-charged POPC/DOTAP liposomes compared to the ones reconstituted into negatively-charged POPC/POPG liposomes. This might infer that PR reconstituted into positively-charged POPC/DOTAP could have a structure more similar to the structure of M intermediates. However, high-resolution structural studies on the E-F loop and other parts of the PR structure under different electrostatic environments will be required to conclude whether these two movements were the same.

4.2 H75-D97 Interaction

The conserved H75-D97 interaction is a special feature in PR structure that has been proposed to stabilize the protonated D97 through a hydrogen bond interaction, and hence increase the pKa_{D97}.[49] This H75-D97 interaction is not present in another microbial proton pump bacteriorhodopsin where its primary proton acceptor D85 of the Schiff base has a lower pKa (pKa~2.6).[81] In PR studies, mutating H75 into different amino acids that cannot form a hydrogen bond with D97 was found to bring down the pKa_{D97} from 7 to values between 5 and 6.[49, 57-59] Since we found significant changes of pKa_{D97} as PR was reconstituted in liposomes with different electrostatic environments, we next examine whether this H75-D97 interaction is perturbed and has a correlation with the observed pKa_{D97} shift.

A schematic diagram of the protonation behavior of both the D97 residue and the surrounding H75 residue under different pH is plotted in Figure 4.6. Typically, the pKa of this Nδ1 atom ($pKa_{N\delta1}$) is around 6.5 and this value could vary from one protein site to another as other embedded amino acids.[154-155] Under a pH that is above the pKa of Nδ1 ($pKa_{N\delta1}$) and below the pKa_{D97}, a majority of PR would have a deprotonated Nδ1 atom on the H75 that acts as a hydrogen bond acceptor to form a hydrogen bond with the protonated D97. This interaction was sought to be the reason that the primary proton acceptor D97 on PR has a higher pKa, where the protonated D97 can be stabilized through this hydrogen bond interaction.

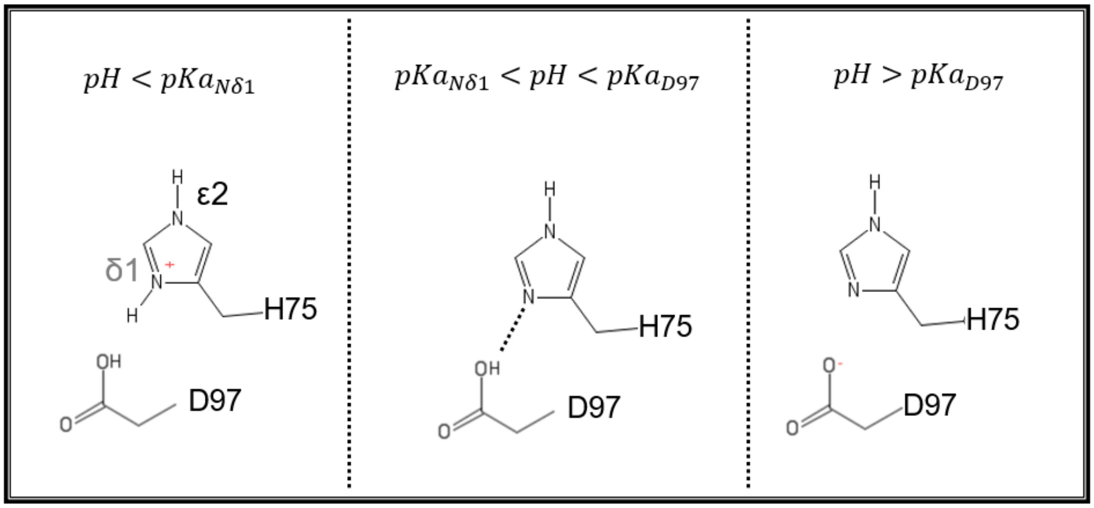


Figure 4.6: Schematic Diagram of H75 and D97 at different pH. The protonation state of both Nδ1 atom on H75 and the carboxyl side chain on D97 changes under different pH. When the pH is lower than the pKa of D97 (pKa_{D97}) and the pKa of Nδ1 atom ($pKa_{N\delta1}$), the majority of PR would have both a protonated Nδ1 atom and a protonated D97. When the pH is between pKa_{D97} and $pKa_{N\delta1}$, the majority of PR would have a deprotonated Nδ1 atom and a protonated D97, and a hydrogen bond could form between these two amino acids to stabilize protonated D97. When the pH is higher than pKa_{D97} and $pKa_{N\delta1}$, the majority of PR would have both a deprotonated Nδ1 atom and a deprotonated D97. This mechanism was proposed by Hempelmann et al.[49]

The possible correlation between the H75-D97 interaction and our observed intrinsic pKa_{D97} change can be tracked by the protonation behavior of the Nδ1 atom on the H75 while PR is reconstituted into liposomes with different electrostatic environments with different pKa_{D97}. If the observed modulation of pKa_{D97} by external electrostatics correlates with the H75-D97 interaction, we can expect that the $pKa_{N\delta1}$ would also shift while PR was reconstituted in different electrostatic environments, and the direction of this $pKa_{N\delta1}$ shift should be in the same direction as what we have observed from the change of pKa_{D97}. Nevertheless, the protonation behavior of the Nδ1 atom cannot be directly tracked by the optical absorbance change of PR as the case of D97. Here, solid-state magic angle spinning (MAS) NMR with dynamic nuclear polarization (DNP) enhancements was applied to track the protonation state of this Nδ1 atom at different pH and conditions. The measurements were done on PR selective-labeled with ^{15}N on the histidine residue (^{15}N-His PR). Since PR only has one histidine over the entire protein, the collected ^{15}N NMR signal is mainly contributed by the target H75 residue. To stand a chance of resolving the resonances of the Nδ1, the state-of-the-art DNP enhancement with polarization transferred from high-γ electrons on a biradical polarizing agent AMUPOL doped in PR samples to the low-γ target nuclei of detection can significantly reduce the time required for signal averaging on the selective-labeled sample from days to hours. In a previous study, the resonances from the protonated Nδ1 (~160 ppm) and the deprotonated Nδ1 (~240 ppm) have been identified and a transition between these two populations was observed under different pH.[49] Since directly mapping out $pKa_{N\delta1}$ would be experimentally challenging as the NMR spectra would need to be collected under several pHs to complete the titration curve, here, instead of altering pH, we aim to identify whether the relative intensity of the two resonances of Nδ1 also change at the same pH but under different electrostatic environment to find any potential coherence between the protonation behavior of Nδ1 and the pKa_{D97}.

Since the PR samples for solid-state MAS DNP were prepared in a DNP juice solvent (d_8-glycerol/D_2O/H_2O, 6/3/1, v/v/v) with 20 mM AMUPOL biradical[203] instead of in water as in previous chapters, we first verified whether changing the net charge of liposomes and the concentration of NaCl in this solvent can still modulate the pKa_{D97} of PR. The apparent pKa_{D97} of PR liposome samples dissolved in DNP juice and water were compared in Figure 4.7, while the apparent pKa_{D97} in water was the same as the ones in Chapter 3. Among the four electrostatic environments tested, the apparent pKa_{D97} of PR liposomes prepared in DNP juice still shared the same trend as the ones prepared in water presented in Chapter 3. The apparent

pKa_{D97} was found to be the highest when PR was reconstituted in negatively-charged POPC/POPG (80/20, mol/mol) liposomes and a DNP juice with 0 mM NaCl (pKa_{D97} = 7.7), and the lowest can be found when PR was reconstituted in positively-charged POPC/DOTAP (80/20, mol/mol) liposomes and a DNP juice with 0 mM NaCl (pKa_{D97} = 6.0). The addition of 150 mM NaCl in the DNP juice also converged the difference in pKa_{D97} to intermediate values, with the apparent pKa_{D97} of PR in negatively-charged POPC/POPG liposomes (pKa_{D97} = 7.2) is still slightly higher than the one from PR in positively-charged POPC/DOTAP liposomes (pKa_{D97} = 6.8). These results suggest that the pKa_{D97} can still be modulated by both the net charge of lipid and the concentration of NaCl in the solution for PR liposomes prepared in DNP juice. The next step was to examine the protonation behavior of Nδ1 atom on selective-labeled ^{15}N-His PR by DNP-enhanced ssNMR.

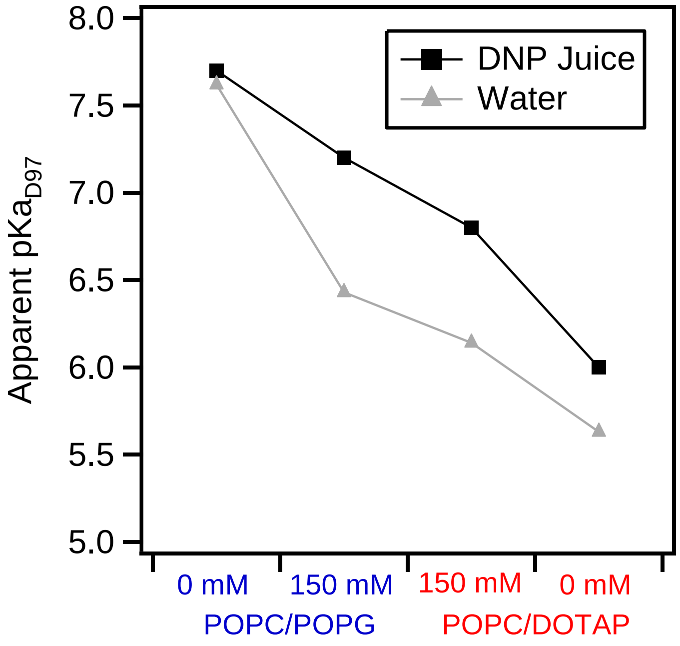


Figure 4.7: Apparent pKa_{D97} of PR in DNP juice. Apparent pKa_{D97} of PRin liposomes with different electrostatics equilibrated in DNP juice (d8-glycerol/D_2O/H_2O, 60/30/10, v/v/v) supplemented with 20 mM AMUPOL is compared with the ones in 10 mM HEPES buffer made with water.

Next, we examined whether we can reproduce the protonation and deprotonation of Nδ1 while PR is prepared in one electrostatic environment but different pH as in the literature. ^{15}N-cross-polarization-MAS (CP-MAS) spectra of ^{15}N-His PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes and a DNP juice (d_8-glycerol/D_2O/H_2O, 6/3/1, v/v/v) supplemented with 20 mM AMUPOL and 150 mM NaCl are shown in Figure 4.8. The sample was equilibrated at 3 different pH to track the change of Nδ1 protonation state. The regions for the resonances of deprotonated ^{15}N (N:) and protonated ^{15}N (N-H) on the imidazole side chain of H75 are highlighted by blue and red shades, respectively. At pH 4.5, both Nδ1 and Nε2 are protonated and the resonances were at ~168.0 ppm. Due to the fast exchange of resonating protons at this pH, there is no clear distinguishment between the resonances from protonated Nδ1 and Nε2. At pH 7, two resonances at 163.6 ppm and 157.8 ppm corresponded to the protonated Nδ1 and the protonated Nε2, respectively, can be distinguished. A lower intensity resonance can also be found in the blue shaded region, corresponded to the deprotonated Nδ1 (244.9 ppm). When the pH further increased to 8.5, a clear decrement of the protonated Nδ1 resonance can be observed in the red shaded region, accompanied by a growth of the deprotonated Nδ1 resonance at 244.9 ppm, showing continue deprotonation of the Nδ1 on the imidazole ring. Overall, the ^{15}N-CP-MAS spectra agreed with what was found from PR in negatively-charged liposomes composed by different lipids, with the resonances from deprotonated and protonated Nδ1 identified.[49] It is noted that the resonance of deprotonated Nδ1 was smaller than the other two resonances from the protonated Nδ1 and Nε2. The difference may originate from the lower cross-polarization efficiency of the deprotonated nitrogen without a proton in the vicinity. It could also be due to a larger chemical shift anisotropy of the deprotonated nitrogen with an electron lone pair to broaden the deprotonated Nδ1 resonance. As it is not as easy and fast as the spectral titration experiments in determining pKa_{D97}, instead of mapping out the $pKa_{N\delta1}$ to compare its value under different electrostatic environments, an alternative method is to fix the pH at an intermediate value (pH = 7) with resonances from both protonated and deprotonated Nδ1 and see whether the change of external electrostatic environment can affect the protonation state of Nδ1.

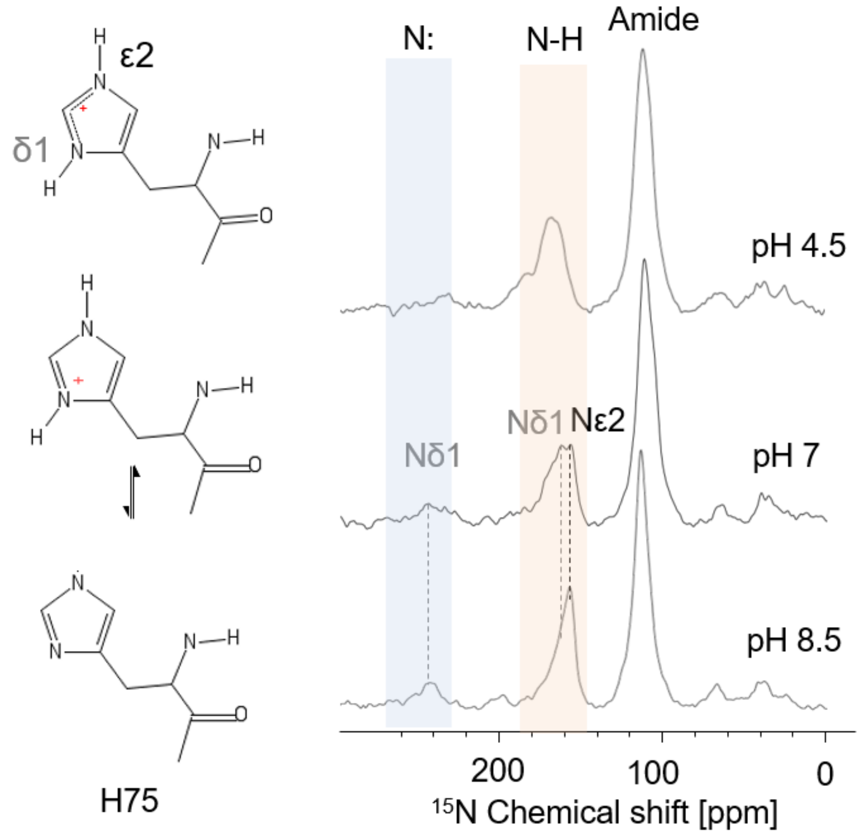


Figure 4.8: ^{15}N-CP-MAS at different pH. ^{15}N CP-MAS spectra of ^{15}N-His-PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes. The red shaded region (~160 ppm) indicated the resonances of protonated Nδ1 and Nε2 on H75. The blue shaded region (~250 ppm) highlighted the resonances of deprotonated Nδ1 and Nε2 on H75. The major protonation state of H75 at the three tested pH are shown at the side.

^{15}N-CP-MAS spectra of ^{15}N-His PR at the same pH 7 but in different electrostatic environments were compared in Figure 4.9. Here we focused on the intensity change of the protonated Nδ1 resonance with a better resolution on any changes compared to the deprotonated Nδ1 resonance. Interestingly, despite the PR samples were equilibrated at the same pH 7, the protonated Nδ1 resonance was smaller in the case of PR reconstituted in positively-charged POPC/DOTAP (80/20, mol/mol) liposomes equilibrated in the DNP juice with 0 mM NaCl compared to other conditions. This result inferred a lower $pKa_{N\delta1}$ for PR reconstituted in positively-charged POPC/DOTAP liposomes, as a lower $pKa_{N\delta1}$ can result in more PR with deprotonated Nδ1 under the same pH 7. As $pKa_{N\delta1}$ shifted in the same direction as the pKa_{D97} did, our result suggested that the H75-D97 interaction could play a role in modulating the pKa_{D97} while PR was reconstituted in different external electrostatic environments. The shift of both $pKa_{N\delta1}$ and pKa_{D97} changed the pH range for the hydrogen bond to form between the deprotonated Nδ1 on H75 and the protonated D97. In the case of PR reconstituted in POPC/DOTAP liposomes with 0 mM NaCl here, the pH range for H75-D97 hydrogen bond formation (the center condition described in Figure 4.6) shift lower compared to other conditions. Nevertheless, even though the ^{15}N-CP-MAS spectra of ^{15}N-His PR unveiled a clear shift of $pKa_{N\delta1}$ that parallels with the observed pKa_{D97} modulation by the external electrostatic environment, it did not reveal potential changes of other residues that could involve the internal electrostatic network between lipid headgroups and D97. To understand which other residues also participated in the structural rearrangement, we next applied a 2D dipolar-based correlation experiment on uniformly isotope-labeled PR.

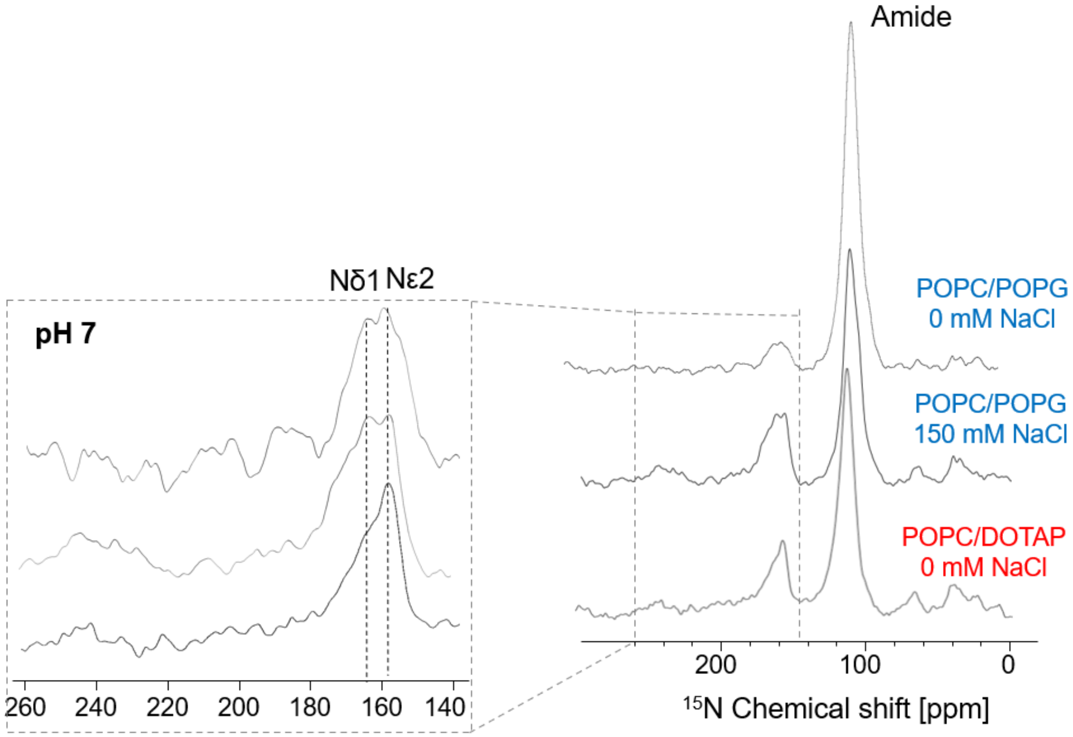


Figure 4.9: ^{15}N CP-MAS spectra of PR in liposomes with different electrostatics. ^{15}N CP-MAS spectra of ^{15}N-His-PR reconstituted in three different electrostatic environments: POPC/POPG (80/20, mol/mol) liposomes with 0 mM NaCl in the solvent, the same liposomes but with 150 mM NaCl in the solvent, and POPC/DOTAP (80/20, mol/mol) liposomes with 0 mM NaCl in the solvent. The resonances of amide backbone, protonated Nδ1, and protonated Nε2 from H75 are highlighted in the spectra. All the samples were equilibrated at pH 7. The NMR measurements were done under 85 K at 400 MHz and 8 kHz of MAS.

4.3 Structural changes beyond H75-D97 and E-F loop

The global structure of PR under different electrostatic environments was obtained through 2D dipolar-based correlation experiments on PR that were uniformly labeled by NMR sensitive nuclei. In the 2D NMR experiment, through-space correlations between nuclei that are dipolar coupled with each other are encoded in the cross-peaks on the off-diagonal region of the 2D NMR spectra. The 2D correlation experiment with an extra dimension is useful for the structural studies of large protein macromolecular to resolve different overlapping resonances from nuclei on the uniformly-labeled protein, and hence can be used to resolve the part of the entire protein structure that experiences structural rearrangements under different conditions. The 2D ssNMR technique used here to resolve PR structure under different electrostatic environments is the dipolar assisted rotational resonance (DARR). PR samples for DARR measurements were uniformly labeled with ^{13}C. The ^{13}C nuclei that are close in place would be dipolar coupled and result in cross-peaks on the spectrum. Under a short mixing time, the cross-peaks are mainly the correlations between the ^{13}C that are one chemical bond away on the same residue. Under a longer mixing time, correlations between ^{13}C on different residues but close in space can appear on the 2D spectrum as well.

^{13}C-^{13}C DARR spectra of uniformly ^{13}C-labeled PR prepared in POPC/POPG (80/20, mol/mol) liposomes with 150 mM NaCl in the solvent (blue) and in POPC/DOTAP (80/20, mol/mol) liposomes with 0 mM NaCl in the solvent (red) were compared in Figure 4.10. The spectra were acquired with a 50 ms mixing time, which is a mixing time that can resolve correlations between carbons that are one chemical bond away or the ones that are on different residues but close in space. Here we focused on the region with correlations mainly from the ones between the Cα and the Cβ (between 10 to 80 ppm). The chemical shift assignments by previous ssNMR studies on PR reconstituted in liposomes were overlaid to interpret our spectra.[60, 62, 204] Distinct intensity differences can be found from the correlations of alanines (Ala) and threonines (Thr), highlighted on the 2D spectra in Figure 4.10. Stronger Cα-Cβ correlations from alanines (A168, A174, A178, A185) were found in the case of PR reconstituted in positively POPC/DOTAP liposomes with 0 mM NaCl in the DNP juice, which is the environment with a lowest apparent pKa_{D97} (6.0). A stronger correlation suggested a decrement of the backbone mobility among these residues. Interestingly, these resonances included the Cα-Cβ correlation of A174 on the E-F loop, together with other alanines from the E-F loop and nearby helices (A168, A178, A185). The stronger correlation found from A174 here coincided with the results of cw EPR lineshape analysis and ODNP relaxometry, where a more steric environment was experienced by the spin label attached on site 174. In the case of threonines resonances (T29, T86, T101, T177), a weaker Cα-Cβ correlation accompanied with a stronger Cβ-Cγ correlation were found from PR in the positively-charged POPC/DOTAP liposomes, also suggesting a change of backbone rigidity among these amino acids upon changing the external electrostatic environment.

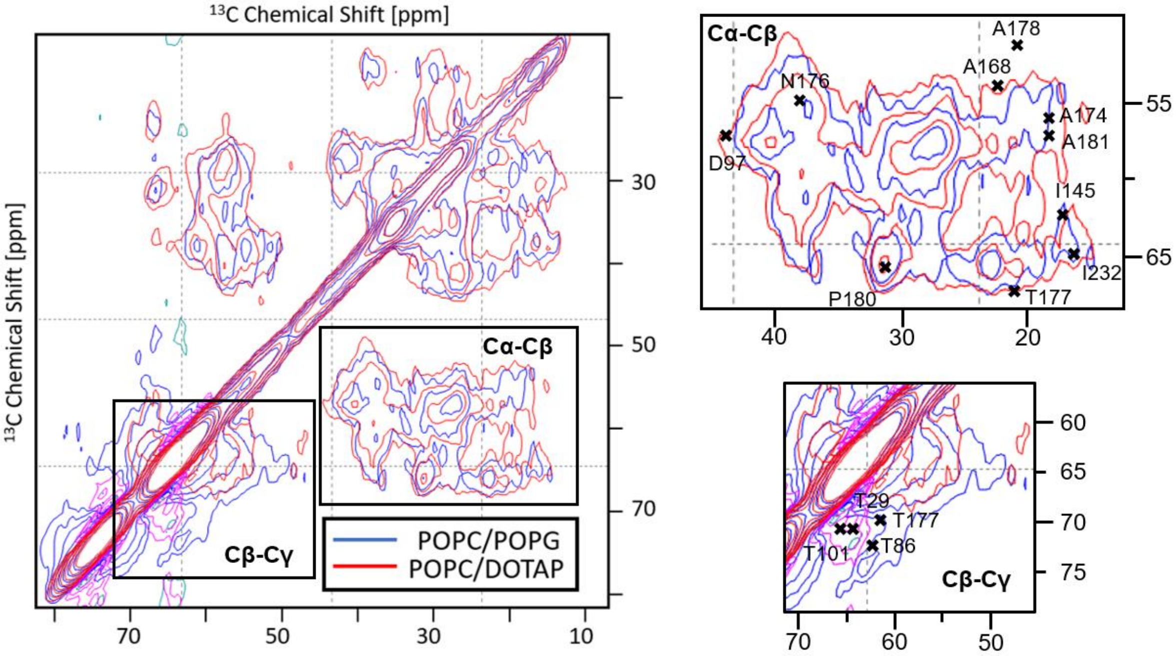


Figure 4.10: ^{13}C-^{13}C DARR of PR in Liposomes with Different Electrostatics. ^{13}C-^{13}C DARR of PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes equilibrated in the DNP juice with 150 mM NaCl (blue) and in POPC/DOTAP (80/20, mol/mol) liposomes equilibrated in the DNP juice with 0 mM NaCl (red). This region mainly encodes Cα-Cβ and Cβ-Cγ correlations. Major differences between the two conditions were highlighted. The chemical shift assignment from literature in a similar environment was overlaid for interpretation.[60, 62]

We also inspected a separate region that encoded the correlations from CO-Cα on the protein backbone of all amino acids and the correlations from the carboxyl functional group of certain amino acids (e.g. aspartic acid, glutamic acid) as shown in Figure 4.11 A. The assigned chemical shifts in literature were also highlighted in the figure for spectrum interpretaton.[60, 62] Among all the correlations, the Cγ-CD correlation of E85 experienced a chemical shift perturbation, which can be observed from a change of its chemical shift between the two electrostatic environments. This observation suggested that E85 could be a residue that experienced a structural rearrangement while PR was reconstituted in different electrostatic environments. Other resonances of glutamic acids or aspartic acids (e.g. D88, E142, D227) were hindered by the resonances from CO-Cα on the backbone of amino acids. As a result, it was hard to resolve whether these residues were affected by external environments or not. We also examined another region with resonances from the side-chain carbons of H75 (Figure 4.11 B). However, possibly due to a lower signal level from the only histidine (i.e. H75) over the entire PR, we cannot find all of the correlations from H75 as highlighted in the literature, specifically, the correlations between the Cγ on D97 and the Cγ on H75.[49] Even other correlations related to H75 were found, these correlations overlapped with resonances of amino acids including tyrosine, tryptophan, and phenylalanine, as these amino acids were not reversely labeled in our PR sample.[60, 62] As a result, we cannot resolve the possible change of relative positions between H75 and D97 under different electrostatic environments. In the future, the same experiments can be done on PR with ^{13}C labeled on only aspartic acids and the H75 to increase the resolution of these resonances.

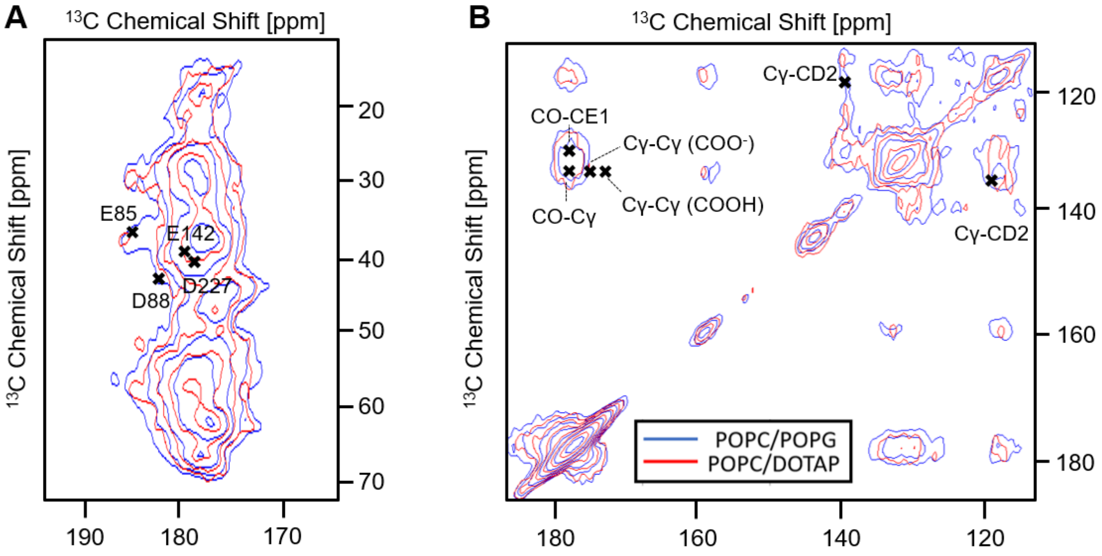
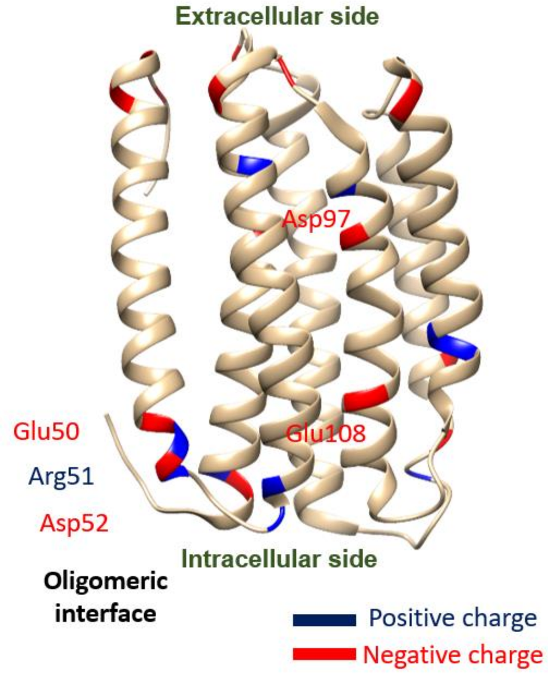


Figure 4.11: ^{13}C-^{13}C DARR of PR in Liposomes with Different Electrostatics. ^{13}C-^{13}C DARR of PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes equilibrated in the DNP juice with 150 mM NaCl (blue) and in POPC/DOTAP (80/20, mol/mol) liposomes equilibrated in the DNP juice with 0 mM NaCl (red), with the regions of correlations from **A.** CO-Cα backbones and amino acids with a carboxylic functional group and **B.** side-chain carbons on histidines, tyrosines, tryptophans, and phenylalanines. The assignments on resonances done by the literature were also overlaid for comaparison.[60, 62]

Overall, the 2D dipolar-based correlation experiments suggested that the entire structure of PR, beyond the E-F loop and the H75-D97 cluster, my be rearranging when PR is reconstituted in different electrostatic environments. This structural rearrangement could have a direct relationship with the observed pKa$_{D97}$ modulation by the external electrostatic environment. Nevertheless, the spectral resolution here of Figure 4.10 and Figure 4.11 acquired by 400 MHz MAS DNP at UCSB were insufficient to identify changes at a single residue level. To identify every individual residue involved in the internal electrostatic network to couple with lipid surface charge and modulate the pKa$_{D97}$, the same 2D DARR correlation experiment will need to be done under a higher magnetic field with higher sensitivity, and also under a higher MAS frequency to achieve higher resolution.

4.4 Effect of Point Mutation on pKa_{D97}

Other than using the spectroscopic tool to locate residues that were affected by the change of external electrostatic environment, an alternative approach to identify the internal electrostatic network was to directly mutate other amino acids on PR that could take part in this network and examine whether the pKa_{D97} was modulated. If the mutation can induce a pKa_{D97} shift, it is highly possible that the mutated site is part of the network that determines the protonation state of D97. A plausible guess on the residues that could involve in this network would be ionizable residues on PR other than the embedded aspartic acid D97, especially those ionizable residues that are closer to the surface of PR and locate between the embedded D97 and lipid headgroups. Therefore, we identified and highlighted all ionizable residues on the surface of PR that could be part of the internal electrostatic network in Figure 4.12. Among these residues, the resonances of D85, E88, and E142 on the extracellular surface of PR have been assigned on the ssNMR spectra in literature.[60, 62] Specifically, our ^{13}C-^{13}C DARR spectra unveiled that E85 experienced a chemical shift perturbation as a sign of structural rearrangement in different external electrostatic environments (Figure 4.10). As a result, we chose these ionizable residues as a starting point and mutated the ionizable aspartic acid (D) and glutamic acid (E) into neutral asparagine (N) and glutamine (Q), respectively, and compared the apparent pKa_{D97} of these mutants to the ones from WT PR to see whether the mutation can induce a shift of apparent pKa_{D97}.



Extracellular	Intracellular
D22	D245
D24	E50
D27	D52
E85	E170
D88	R51
E142	K53
D212	K57
	K59
	K172
	K244

Figure 4.12: Ionizable Residues on PR. Ionizable amino acids including aspartic acid (D), glutamic acid (E), arginine (R), and lysine (K) are highlighted on PR structure, color-coded by the net charge of their ionized form. The amino acids on the surface of PR are categorized into the ones from the extracellular sides and the ones from the intracellular side.

Apparent pKa_{D97} of PR E85Q, D88N, and E142Q mutants were compared to the pKa_{D97} of WT PR in both DDM detergent micellar and POPC/POPG (80/20, mol/mol) liposome environment (Figure 4.13). Overall, the differences in apparent pKa_{D97} induced by mutation were larger while PR was in the DDM detergent micelle environment compared to the ones in POPC/POPG (80/20, mol/mol) environment. Among these mutants, the apparent pKa_{D97} of E142Q (circle, Figure 4.13) was significantly lower than the ones of WT PR (triangle). No significant apparent pKa_{D97} shift was observed from the E85Q and D88N mutants, despite the fact that a chemical shift perturbation was found from E85 on the ^{13}C-^{13}C DARR spectra between PR reconstituted into liposomes with different electrostatic environments (Figure 4.10). Nevertheless, these results supported our hypothesis that a mutation on an ionizable residue which is not in the close vicinity of the embedded D97 residue can still affect its

pKa_{D97}. More importantly, the direction of pKa_{D97} shift toward a lower value observed from the E142Q mutant agreed with what was expected, where a local environment with fewer negative charges drives the deprotonation of D97.

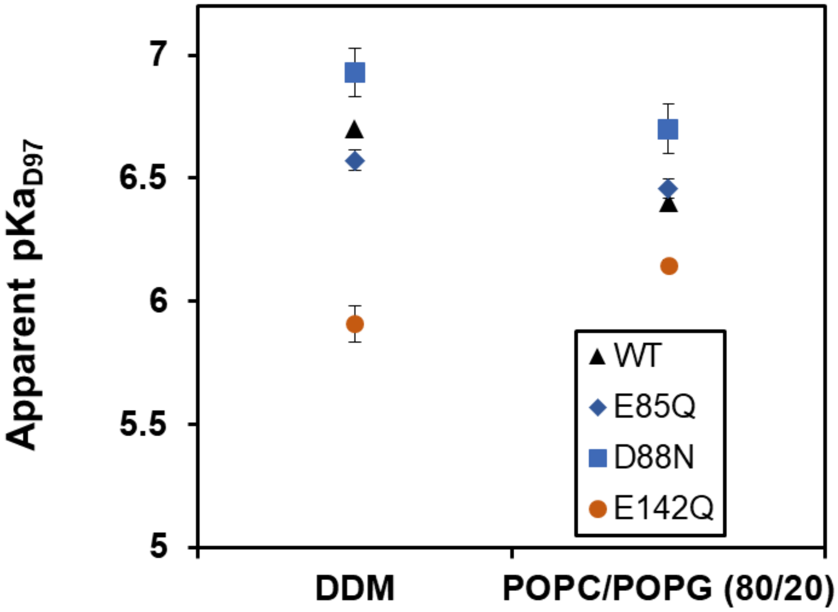


Figure 4.13: Apparent pKa_{D97} of PR with Mutations on Charged Residues. The pKa_{D97} of E85Q (diamond) and D88N mutants (square) with a charged residue substituted by a neutral residue, in both DDM detergent micelle and POPC/POPG (80/20, mol/mol) liposomes environment. The pKa_{D97} are compared to the ones of WT PR (triangle).

In fact, the E142Q mutation at a corresponding glutamic acid on the blue-light-absorbing PR (BPR) has also been shown to induce a decrement of pKa_{D97}, together with a blue shift on the maximum absorbance wavelength.[205] In the case of BPR, this mutation also slows down its photochemical reaction cycle by slowing down the decay of the K intermediate. These all infer that the E142 could have a functional role in regulating both the activity and the proton transport rate of BPR. Our work here concurred that this site could be part of the internal electrostatic network to regulate the pKa_{D97} of the green-light-absorbing variant used in this dissertation and in turn affect the population of active PR. Nevertheless, the unclear part is whether the PR structure at site E142 is affected by the change of external electrostatic environment, as the resonance of E142 was obscured by other resonances from the CO-Cα correlations of other amino acids on the ^{13}C-^{13}C DARR spectra of uniformly-^{13}C-PR (Figure 4.10). In the future, the same experiment on PR with selective ^{13}C-labeling on glutamic acids would be helpful to resolve the E142 resonance to find any potential structural rearrangement at this site with the change of surrounding electrostatic environment.

We also looked into the literature to understand the exact protonation state of these mutated glutamic acids and aspartic acid at pH~8.5. Both solvent-exposed E85 and D88 were shown to populate with their negatively charged deprotonated state, and the E142 at the end of the helix D was shown to populate with its neutral protonated state.[60] This is an interesting finding, as the mutation on both negatively-charged residues to neutral amino acids did not significantly affect the apparent pKa_{D97}. On the contrary, even the E142 was populated with its neutral state, a mutation on this neutral glutamic acid into a neutral glutamine induced a greater pKa_{D97} shift. This suggested that the net charge of each ionizable amino acid on PR may not be the key determinant to affect the pKa_{D97}, but other factors like the capability of salt bridge formation[205] or the orientation of the side chain[163] can be the determinants. In fact, a simulation study on a catalytic enzyme Histone Deacetylase 8 identified several remote amino acids that can be critical points to affect the charge density at its active site. The study also suggested that the charge of critical points was not the key determinant, but the relative positions of all the critical points to the embedded site could be the main factor to affect the local microenvironment around the active site, in our case the embedded D97 residue. As a result, future studies on identifying the internal electrostatic network of PR that determines

pKa_{D97} should extend to residues not limited to the ionizable ones but also other polar or non-polar residues.

4.5 The outlook of PR Structural Measurements

In this chapter, by applying complementary magnetic resonance techniques, we observed structural rearrangements of PR while it was reconstituted into different electrostatic environments with different pKa_{D97}. Our observation further fortified that protein intrinsic structural properties must play a role in modulating pKa_{D97}. The structural rearrangement of PR was not limited to the sites that are in the vicinity of the embedded D97 residue, but also remoted sites that are far away from this D97 residue. While the change of the $pKa_{N\delta1}$ on H75 was found to be in the same direction as the pKa_{D97} shift as PR was reconstituted in different electrostatic environments (Figure 4.9), suggesting a change on the pH range for the H75-D97 hydrogen bond formation, we also found some remote residues, like the ones on the E-F loop and the E142 residue, that could also be parts of the internal electrostatic network coupled with charged lipid headgroups to affect pKa_{D97} (Figure 4.10 and Figure 4.13). Taking all our observations together, we inferred that mediated short-range interactions between charged lipid headgroups and the embedded D97 inside PR form an internal electrostatic network that determines the protonation state of D97 residue, and multiple residues beyond the local H75 could be part of this network.

The 2D dipolar-based DARR spectra in this chapter revealed changes in the PR backbone rigidity at different regions over the entire protein. However, the resolution is insufficient for us to identify each individual residue that was rearranged by the change of external environment and could take part in the internal electrostatic network to determine pKa_{D97}.

Undoubtedly, repeating these experiments under a higher magnetic field (800 MHz) and a higher MAS frequency (above 10 kHz) together with the signal enhancement from DNP stands a great chance to resolve all the residues that participate in this internal network coupled with the embedded D97. Since a similar long-range modulation effect on the protonation behavior of embedded ionizable residues has been also observed in other protein systems, including a catalytic enzyme Histone Deacetylase 8[198] and another proton pump bacteriorhodopsin.[163] Further elucidating the role of the internal electrostatic network on determining the pKa of embedded ionizable residue as the case of D97 here has significance in understanding the regulatory mechanisms of proteins that have functions controlled by a similar mechanism as PR.

Chapter 5

Optimizing PR Function in Other Biomimetic Materials

In Chapter 3, we learn that the electrostatic environment of PR-constituted liposomes can affect both the apparent and intrinsic pKa_{D97}, which in turn determined the population of active PR with vectorial proton transport capability from the inner to the outer side of the lipid bilayer. For energy-related bioengineering applications utilizing the vectorial proton transport capability of PR, a lower apparent pKa_{D97} is preferred, as more active PR would present to maximize its proton transport performance. On the other hand, for biophysical studies, the apparent pKa_{D97} of PR in different biomimetic environments provides a guideline for researchers to decide at which pH the structural or functional characterization experiments should be conducted to get the information from active PR instead of a mixture of both species. In this chapter, we aim to examine whether the knowledge of adjusting the apparent pKa_{D97} of PR by the external electrostatics in the liposome environment can be applied to other biomimetic environments. The two environments examined here are the lipid nanodisc, an environment ideal for protein structural studies using cryo-EM, and the mesostructured silica film, a synthetic material that can host transmembrane proteins for bioengineering applications.

5.1 PR Nanodisc

Lipid nanodisc is a biomimetic platform with transmembrane proteins embedded in a small patch of lipid bilayers (typically ~10 nm in diameter). The hydrophobic part of the lipid bilayer is protected by either scaffolding proteins or polymers. Past studies have successfully reconstituted PR into lipid nanodiscs formed by membrane scaffolding proteins (MSPs). Monomeric PR was found to be the major form in nanodiscs formed by DMPC lipids and MSP1D1 scaffold proteins, with an average diameter of ~7.4 nm found by size exclusion chromatography.[79] In the same study, both the spectral titration and time-resolved transient absorbance experiments confirmed that PR prepared in lipid nanodisc was functionally active, with a similar light-activated response compared to PR in liposomes. ^{13}C-^{13}C dipolar-based correlation experiments by ssNMR on PR reconstituted in pure DMPC nanodiscs also unveiled a similar structure compared to the PR in the liposome environment.[101] While these studies focused on ensuring the formed PR nanodisc was structurally and functionally robust, this chapter took a step forward and studied how lipid composition used for nanodisc formation and the oligomerization of PR can affect the function of PR in nanodisc environment.

Here, instead of using the MSP to drive PR nanodisc formation from liposomes, we used amphipathic styrene-maleic acid (SMA) copolymer to form PR nanodisc from the PR proteoliposomes. This SMA polymer was synthesized by copolymerization of polar maleic acid and apolar styrene. While SMA was added to a solution with liposomes, the net negatively-charged SMA polymers can bind on the surface of lipid bilayers. The polymer then would insert into the hydrophobic part of the lipid bilayer and drive the formation of nanodiscs.[206] In contrary to the traditional MSP method, SMA polymers can not only form nanodisc from transmembrane proteins reconstituted in synthetic liposomes but has a great potential to directly extract transmembrane proteins from cell lysate used for protein overexpression.[207-208] This can minimize the possibility of disturbing the native structure and function of transmembrane proteins during detergent-based purification steps. Some studies named SMA polymer-formed lipid particles as SMA-lipid particles (SMALP).[209] In this dissertation, we will use the name "PR nanodisc" for all SMA polymer-formed lipid particles constituted with PR.

5.1.1 Characteristics of SMA-formed PR nanodisc

The SMA polymer-formed PR nanodisc was formed from positively-charged POPC/DOTAP (80/20, mol/mol) synthetic liposomes constituted with monomer-enriched PR E50Q. The PR liposome sample as the one used for studies in previous chapters was mixed with hydrolyzed SMA polymers to drive the formation of PR nanodiscs. To examine whether PR nanodiscs were successfully formed, the particle size distribution of the sample was measured by dynamic light scattering (DLS) after undissolved liposomes and impurities were removed by ultracentrifugation. The average diameter of formed monomeric PR nanodiscs was 6.7 ± 2.4 nm (Figure 5.1 A), agreed with the 6 – 8 nm diameter size reported in the literature estimated by size exclusion chromatography.[79] On the other hand, no significant particle size from liposomes with a ~200 nm diameter size was observed, which suggested that the centrifugation step after the nanodisc formation had removed remaining liposomes from the sample. To examine the functional property of PR, we next measured the optical absorption of the formed PR nanodisc sample between 400 and 750 nm. The absorption spectrum was compared to the ones from the same PR E50Q but in the DDM detergent micellar environment (0.05 wt % DDM in the buffer) and in POPC/POPG (80/20, mol/mol) liposome environment (Figure 5.1 B). All the samples were equilibrated at the same pH 8.5. The absorption spectrum of the PR nanodisc sample (blue) was more similar to the one of PR in detergent micelles (red), with a sharp characteristic maximum absorbance at ~520 nm. On the contrary, the spectrum of the liposome sample (orange) had a broader absorbing range due to the light scattering effect from larger liposomes. It is noted that the maximum absorbance of PR E50Q nanodisc was at 527 nm, which is higher than the one of PR E50Q in the DDM detergent micellar environment (513 nm). This inferred that there was more acidic PR presence in the nanodisc sample at the same pH 8.5. Overall, the optical absorbance spectrum of PR nanodisc suggested a successful formation of SMA polymer-formed PR nanodisc with complete removal of residual liposomes.

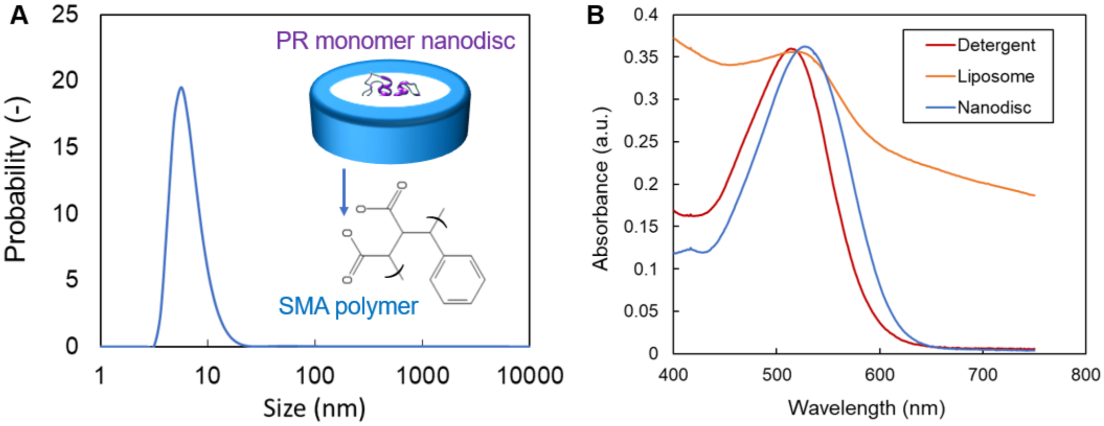


Figure 5.1: SMA Polymer-formed PR Nanodisc. A. Particle size distribution showed the average diameter of formed monomer-enriched PR E50Q nanodisc with POPC/POPG (80/20, mol/mol) lipids is 6.7 ± 2.4 nm. Schematic diagram of an SMA polymer-formed PR nanodisc with the chemical structure of the SMA repeat unit. **B.** The optical absorbance of PR E50Q mutant in DDM detergent micelles (red), in POPC/POPG (80/20, mol/mol) liposomes, and SMA polymer-formed nanodisc (blue) with the same POPC/POPG lipid composition measured at pH 8.5.

5.1.2 Oligomerization and Lipid Net Charge

We compared the apparent pKa_{D97} of PR nanodiscs constituted by either oligomeric WT PR or monomer-enriched PR E50Q, and with lipid compositions of either negatively-charged POPC/POPG (80/20, mol/mol) or positively-charged POPC/DOTAP (80/20, mol/mol). Among all four conditions, oligomeric WT PR in POPC/POPG liposomes was the only condition without nanodisc formation after the addition of SMA polymers and ultracentrifugation. The spectral titration curves of the other three conditions were compared in Figure 5.2. The apparent pKa_{D97} of PR nanodisc under the three conditions were all higher than the ones observed from PR in DDM detergent micelle (the highest was 7.4) and in liposome environment (the highest was 7.6). The monomer-enriched PR E50Q nanodisc composed of POPC/DOTAP has a higher pKa_{D97} (8.78) compared to the oligomeric WT PR nanodisc with the same lipid composition (7.79). No significant apparent pKa_{D97} difference was observed between PR E50Q nanodisc composed of positively-charged POPC/DOTAP and negatively-charged POPC/POPG (pKa_{D97} = 8.73).

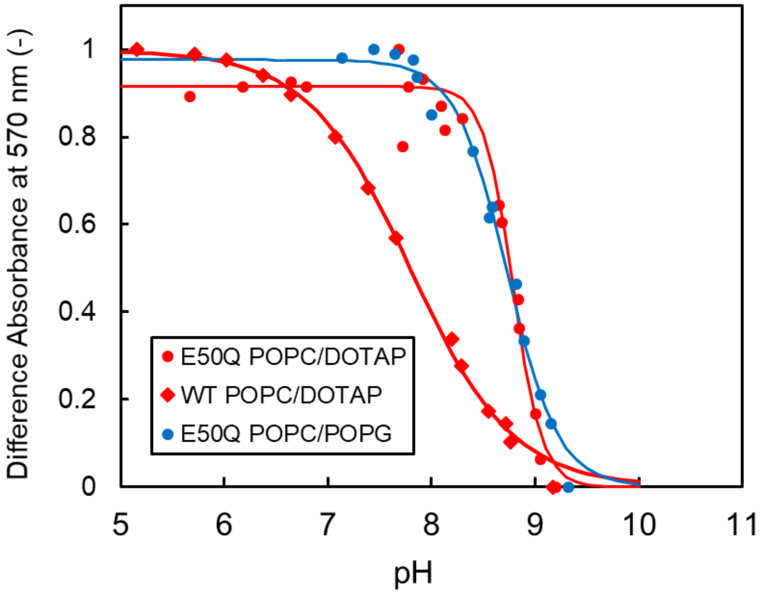


Figure 5.2: Spectral Titration Curves of PR Nanodisc. pH-dependent absorbance transition at 570 nm of PR E50Q nanodisc composed of POPC/DOTAP (80/20, mol/mol) (red circle) or POPC/POPG (80/20, mol/mol) (red diamond) and WT PR nanodisc composed with POPC/DOTAP (80/20, mol/mol).

The higher apparent pKa_{D97} found for all PR nanodisc suggested that PR preferred to be in its inactive state with a protonated D97. This is also supported by a red-shifted maximum absorbance wavelength (λ_{max} = 527 nm) of PR nanodisc at pH 8.5 (Figure 5.1). Learned from the studies in Chapter 3 and 4, the higher apparent pKa_{D97} of PR in nanodisc environment can be due to a higher intrinsic pKa_{D97} of PR reconstituted in nanodisc, or a lower surface pH around the SMA polymer-formed nanodisc. It is noted that the SMA polymer itself is negative-charge at neutral and basic pH.[210] The negatively charged polymer can either attract protons and decrease the surface pH around PR reconstituted in SMA-formed nanodisc or directly change the intrinsic electrostatic network of PR to affect its intrinsic pKa_{D97}. The latter case will need to be further confirmed by structural studies on PR in SMA polymer-formed nanodisc.

On the other hand, opposed to what was found in the liposome environment, no significant apparent pKa_{D97} difference was found between monomeric PR E50Q nanodisc composed of lipids with an opposite net charge. Since the average size of a monomeric PR E50Q nanodisc (6.7 nm) is only slightly larger than the estimated size of PR (~5 nm), this inferred that there could be only one or two layers of lipids between the SMA polymer belt and the protein in a nanodisc. In this situation, the surface potential of the PR nanodisc may not be dominated by the net charge of lipid headgroups but determined by the protein itself or the surrounding negatively-charged SMA polymers. Interestingly, the average size of WT PR nanodisc composed of POPC/DOTAP is 17.2 ± 4.5 nm is significantly larger than the size of pentameric and hexameric PR estimated by cryo-EM[96] (~12 nm) and AFM[68] (~10 nm) studies. As a result, in the case of oligomeric WT PR nanodisc, more than two layers of lipids could presence around the protein, and this could increase the contribution of surface potential from lipid headgroups compared to the monomeric PR E50Q nanodisc. Our interpretation can be supported by a lower apparent pKa_{D97} of oligomeric WT PR nanodisc composed of POPC/DOTAP, as the increasing number of positively-charged DOTAP lipids can increase the surface potential to a less negative value. Experimentally, we also found a less negative zeta potential for WT PR nanodisc composed of POPC/DOTAP (-6.59 mV) compared to the monomer-enriched PR E50Q nanodisc sample (-16.7 mV). A more accurate surface potential measurement on PR nanodisc will need to be done in the future to understand the effect both negatively-charged SMA polymers and the number of lipids in a nanodisc on the surface potential experienced by PR. Moreover, the measurements here were done in a HEPES buffer with 150 mM of NaCl, an ion concentration that could screen out most of the electrostatic effect. Systematic measurements on the apparent pKa_{D97} of PR nanodisc in buffers with different NaCl concentrations will help us to understand the mechanical basis of pKa_{D97} modulation observed in nanodisc environments.

5.1.3 Outlook

We successfully formed both oligomeric WT PR and monomer-enriched PR E50Q nanodiscs for the first time by SMA polymer. It is noted that the addition of SMA polymer to POPC/POPG (80/20) liposomes constituted with WT PR was the only condition tested without PR nanodisc formation. The formation of nanodiscs required the SMA polymer to first absorb on the liposome surface and then wrapped around PR with lipids to form the nanodiscs. The first step can be affected by the net charge of liposomes. Even in a buffer with high ionic strength, it is still more energy favorable for negatively-charged SMA polymers to absorb on a POPC/DOTAP liposome surface compared to a POPC/POPG liposome surface, which could end up with a greater population of SMA polymers penetrated the POPC/DOTAP liposome. Since PR oligomers are larger than its monomeric form, the more penetrated SMA polymer into POPC/DOTAP liposomes may overcome the energy barrier for SMA polymers to wrap around the larger oligomeric WT PR, while the SMA polymers in POPC/POPG liposomes might not be sufficient to wrap around the larger oligomeric WT PR and drive the nanodisc formation. For studies that require oligomeric PR nanodisc, specifically cryo-EM studies on resolving the hexameric or pentameric assemblies of PR in nanodisc, our work provides a guideline on using the positively-charged PR proteoliposomes to facilitate the nanodisc formation for oligomeric species.

A higher apparent pKa_{D97} (> 7.8) found from PR nanodisc reinforced the importance of biomimetic environments on affecting the function of PR. Not the same as PR in liposome environment, here the apparent pKa_{D97} of PR in nanodisc could depend on either the oligomerization of PR or the number of lipids in a nanodisc. Our results also suggested that the negatively-charged SMA polymer itself could be an additional factor to affect the function and structure of PR in nanodisc compared to the synthetic liposome environment. Nevertheless, the effect of SMA nanodisc on the structure and function of PR yet needs to be confirmed by magnetic resonance studies. Comprehensive structural studies on PR nanodiscs as the ones done for PR liposomes in Chapter 4 will be helpful for structural understandings.

5.2 PR in Mesostructured Silica Film

Bioengineering applications generally prefer host materials that can increase the thermal, mechanical, and chemical stability of the incorporated transmembrane proteins to broaden the operating range of the platform with functional proteins. These characteristics typically cannot be achieved by traditional biomimetic systems with lipid bilayers, as the lipid bilayer has poor mechanical stability and can easily be destroyed by air-water interfacial force even on a solid substrate. Materials including silica gels or glasses, block copolymers, and hydrogels can increase the mechanical and thermal stability of incorporated transmembrane proteins but exhibit poor processability. At UCSB, a novel approach has been developed by successfully incorporating transmembrane proteins into mesostructured silica films with great mechanical and thermal stability while retaining the native-like function of transmembrane proteins. The mesostructured silica material is an excellent host medium as transmembrane proteins and surrounding surfactant species can be anisotropically aligned in its mesoscopic structures (e.g. hexagonal or worm-like). The alignment can improve the performance of the vectorial transport function of incorporated ion pumps or channels, like PR used in this dissertation. On the other hand, mesostructured silica films are transparent, which is beneficial for proteins with functions triggered by light-activation. The developed synthesis method of surfactant-directed mesostructured silica film can host a high concentration of proteins (up to 15 wt% for PR). While the surfactant–silica mesostructured film can preserve the native-like proton transport function of PR, it has not been optimized to maximize the population of active PR with deprotonated D97. This can be achieved by lowering the apparent pKa_{D97} of PR in mesostructured silica film to expand the operating pH range with over 90 percent of active PR. This is especially important with the fact that the synthesized film is more stable under neutral or mildly acidic pH while hydrated in the aqueous solution. Having an apparent pKa_{D97} of PR that is less than 7 would significantly increase the amount of active PR while the material is operated under this pH range. Learning from the studies in Chapter 3, it is plausible that an introduction of positively-charged lipids (e.g. DOTAP) into the surfactant composite of this surfactant–silica host material could lower the apparent pKa_{D97} to optimize PR activity. We then examined whether this approach can lower the apparent pKa_{D97} of PR in mesostructured silica film.

5.2.1 Incorporating Positively Charged Lipids

The PR-contained mesostructured silica films were typically composed of hydrolyzed silica precursors, PR solubilized in DDM detergents, and pure 1,2-Diheptanoyl-sn-Glycero-3-Phosphocholine (diC_7PC) surfactant with a zwitterionic headgroup (Structures are shown in Figure 5.3 A). The role of the surfactant composite, both DDM and diC_7PC, was to interact with PR to maintain its native-like structure and function and at the same time direct the assembly of mesostructured silica material. Among these two components, diC_7PC is the surfactant component with a lipid-like structure, with two short hydrocarbon tails and a hydrophilic zwitterionic headgroup, which is an ideal candidate to be substituted by DOTAP lipid also with two hydrocarbon tails but a positively-charged headgroup. Since pure DOTAP has low solubility in aqueous solution and cannot meet the concentration (7.1 wt%) required for synthesizing the PR-surfactant-silica mesostructured film, we co-solubilized DOTAP and the zwitterionic POPC at a 1:4 molar ratio as the composition used for liposome samples and substituted the 7.1 wt% diC$_7$PC with this POPC/DOTAP (80/20, mol/mol) mixture for film synthesis. The formed mesostructured silica film made with POPC/DOTAP loaded with 5 wt% PR had a similar macroscopic appearance compared to the well-characterized film made with diC$_7$PC. The film formed on polydimethylsiloxane (PDMS) slab was hydrated by a 10 mM HEPES buffer at pH 8.5 for optical absorbance measurement (Figure 5.3 B). A similar optical absorbance, with a maximum absorbance at ~540 nm, was observed from the two compositions of PR-contained mesostructured silica films, suggesting successful preservation of PR's native structure in the mesostructured film formed by the positively-charged POPC/DOTAP lipid mixture. It is noted that the baseline difference between the two samples in Figure 5.3 B could be caused by the difference in PDMS slab thickness used to support and stabilize the film submerged in the buffer. The red-shifted maximum absorbance (λ_{max} = 540 nm) of PR in mesostructured silica film at alkaline pH (pH 8.5) was also found in the literature. This could be a characteristic of PR incorporated into the mesostructured silica film. Overall, our results suggested a successful formation of a PR-surfactant-silica mesostructured film with the positively-charged DOTAP lipid. The next goal was to examine whether the apparent pKa$_{D97}$ is lower in this environment as we predicted.

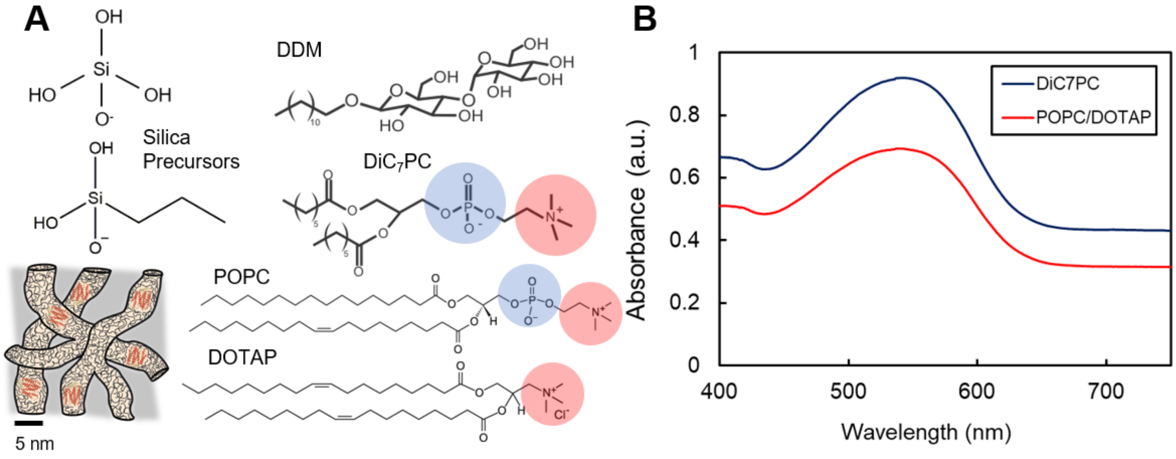


Figure 5.3: PR-surfactant-silica Mesostructured Film. A. Molecular structure of silica precursors, DDM detergent, and DiC$_7$PC used for synthesizing PR-surfactant-silica mesostructured film. The schematic diagram depicts how PR surrounded by surfactant molecules were incorporated into the mesochannels formed by silica. Adapted from Jahnke et al. **B.** Optical absorbance of 5 wt% WT PR loaded in mesostructured silica film composed of 59.4 wt% DDM, 28.5 wt% SiO$_2$, and 7.1 wt% surfactant composites of either pure zwitterionic DiC$_7$PC (blue) or positively-charged POPC/DOTAP (80/20, mol/mol) lipids (red).

5.2.2 Apparent pKa$_{D97}$

Apparent pKa$_{D97}$ of 5 wt% WT PR loaded in mesostructured silica film composed of zwitterionic DiC$_7$PC and positively-charged POPC/DOTAP (80/20, mol/mol) lipids were compared. Since the optical absorbance of PR in mesostructured silica film red-shifted by ~30 nm, pH-dependent absorbance changes of PR in mesostructured silica film were then tracked at 600 nm instead of at 570 nm in previous studies.[74, 145] A clear decrement of apparent pKa$_{D97}$ was observed for PR-contained mesostructured silica films composed of positively-charged POPC/DOTAP lipids (pKa$_{D97}$ = 6.13 ± 0.39) compared to the ones composed of pure zwitterionic diC$_7$PC (pKa$_{D97}$ = 8.16 ± 0.22) in the surfactant composite (Figure 5.4 A). This apparent pKa_{D97} shift of PR induced by changing the net charge of host materials was comparable to the one found in the liposome environment in Chapter 3 (Figure 5.4 B). The results suggested that the incorporation of positively-charged DOTAP lipids into the surfactant composite of the mesostructured silica film can decrease the apparent pKa as we hypothesized. The ~2 pH units drop of apparent pKa_{D97} was significant, as the population of active PR with deprotonated D97 at an environment of pH 8 would increase from 50 to 99 %.

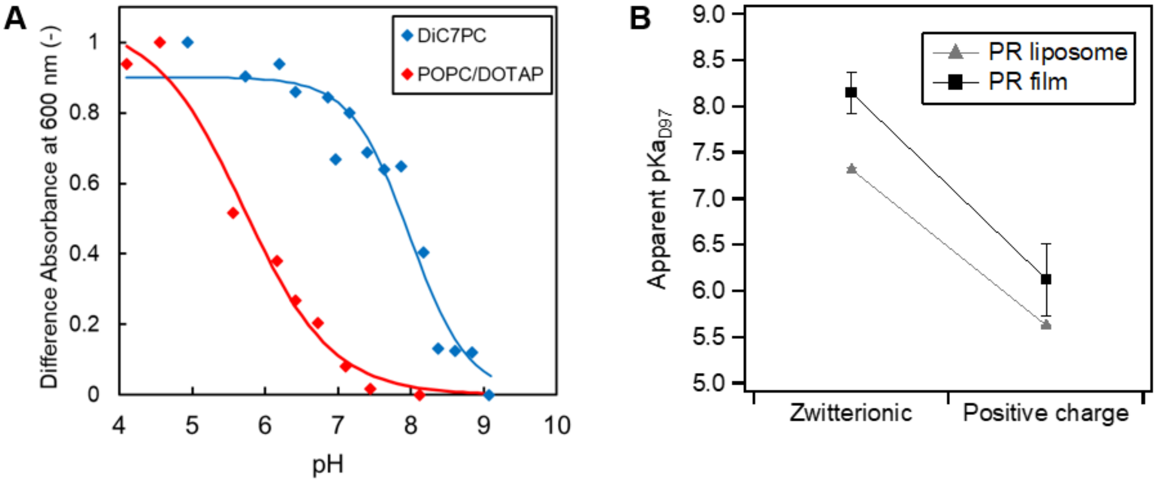


Figure 5.4: Spectral Titration of PR-surfactant-silica Film. A. pH-dependent absorbance transition at 600 nm of 5 wt% WT PR loaded in mesostructured silica film composed of 59.4 wt% DDM, 28.5 wt% SiO_2, and 7.1 wt% surfactant composites of either pure zwitterionic DiC_7PC (blue) or positively-charged POPC/DOTAP (80/20, mol/mol) lipids (red). **B.** Comparisons of apparent pKa_{D97} measured from WT PR reconstituted in liposomes (triangle) or mesostructured silica films (square) with either a pure zwitterionic or a positive net charge

5.2.3 Outlook

A success increment of active PR population with vectorial proton transport capability in mesostructured silica films was achieved by incorporating positively-charge lipids to the surfactant composite of the host material. Our results reinforced the importance of surrounding electrostatic environment on determining the activity of PR, or other transmembrane proteins with functions that are controlled by a similar mechanism, for bioengineering applications. The versatility of the surfactant components used for synthesizing mesostructured silica films facilitates the adjustment on the net charge of host materials in protein engineering applications. In the case of PR-containing mesostructured silica materials, further improvements including aligning PR into the same direction will still need to be addressed in order to maximize the light-activated proton transport flux of PR in materials.

Chapter 6

Structural Movements of PR in its Photochemical Reaction Cycle

The kinetics of PR proton transport function is often characterized by transient optical absorbance technique, as introduced in previous chapters. This technique measures the function-determining structural movement of PR during its proton transport reaction cycle indirectly by the transient absorbance change of the retinal chromophore on PR.[50, 53-54] While it is a powerful approach to examine the proton transport rate of PR under different conditions, however, the transient absorbance cannot unveil the structural movement of different compartments on PR with spatial resolution. This spatiotemporal information on the structural movement could potentially unravel how PR as machinery transports proton from one residue to another to achieve the vectorial proton transport. In addition, this information could also provide structural insight into how microbial rhodopsins have diverse functionality despite sharing a similar light-activated transient absorbance response.

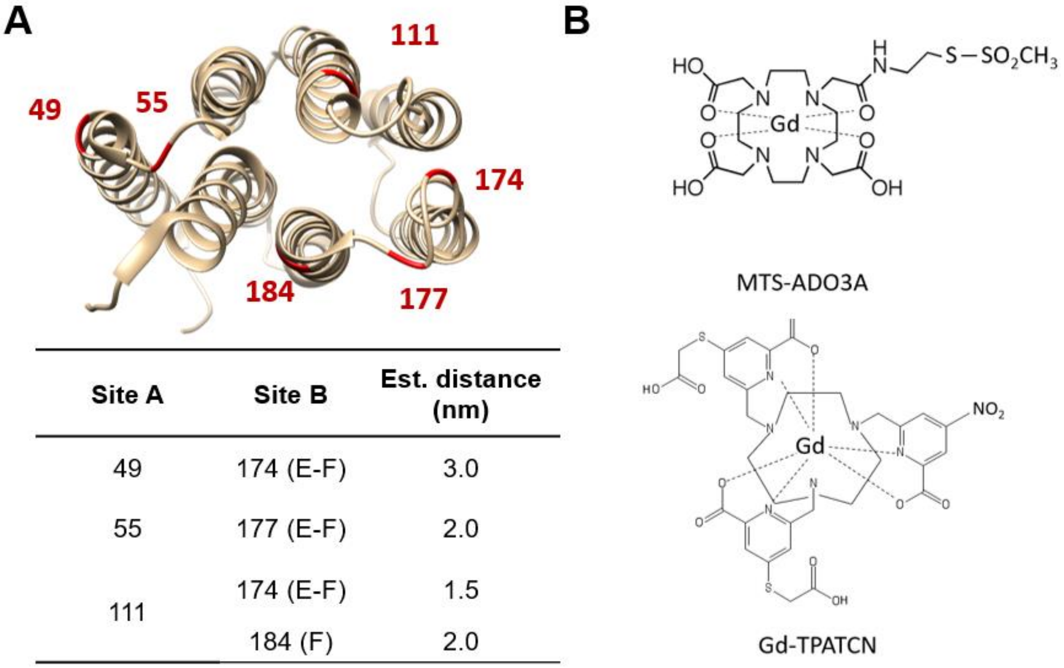
Limited techniques can directly measure the structural dynamics of proteins under a sub-millisecond time range. Even though techniques including NMR spectroscopy and pulse EPR spectroscopy can identify the structural differences between cryo-trapped intermediates in the photochemical reaction cycle of microbial rhodopsin, no kinetic information on the structural movement between trapped states can be retrieved. In the case of PR, an outward structural movement of the E-F loop with temporal information has been resolved by X-band EPR spectroscopy, but only from the variant with an E108Q mutation that slows down the photochemical reaction to tens of seconds.[74] Such an approach with a lower time-resolution is not fast enough to capture the same movement of WT PR with most of the transitions occurs within tens of milliseconds. The current state-of-the-art methodology to measure the function-related structural dynamics of proteins is the time-resolved serial femtosecond crystallography (SFX) using X-ray free-electron lasers (FELs).[211] It has resolved the structural dynamics of bacteriorhodopsin (bR) at logarithmically spaced times between 16 ns and 1.7 ms after the light-activation of bR with an atomic-scale resolution (2.1 Å).[212] However, this method cannot be applied to proteins with disorder regions that are hard to form crystals, such as PR used in this dissertation.

We aim to apply a developing time-resolved Gd-Gd EPR (TiGGER) technique to measure the time-resolved distance change between different segments of PR during its photochemical reaction cycle with a sub-millisecond time resolution. This technique acquires distance information between a pair of residues labeled with Gd chelates through cw EPR lineshape difference. When the pair of Gd chelates are separated by less than ~3.5 nm, the central transition of the cw EPR spectra measured from Gd spins under 240 GHz broadens because of a dipolar interaction between Gd spins, follows a $1/r^3$ dependence. The relationship between Gd-Gd inter-spin distances and the observed lineshape broadening has been established by a Gd-ruler study with known inter-spin distances.[108] The same study also unveiled the potential of conducting this measurement not only under cryogenic temperatures but also under physiologically-relevant room temperatures, which is essential if we want to measure the dynamics of PR structural movements. In fact, the Gd-Gd distance measuring technique has successfully measured the inter-PR distances between the interfacial residue S89 of closest-neighboring PR in its hexamer.[112] Nevertheless, the success of distance measuring on PR by this technique is only limited to sites that have sufficiently high spin labeling efficiency (> 80%) with Gd-based spin labels. Several challenges still need to be overcome before biological TiGGER experiments can be done under room temperature for protein structural dynamics measurements. First, the labeling of Gd chelators on protein samples need to be further improved to reach a higher labeling yield on different sites. On the other hand, a robust method needs to be established to slower the tumbling of Gd-labeled protein under room temperatures to preserve the anisotropic lineshape broadening on cw EPR spectra with distance information.

This part of the dissertation continues the development of Gd-Gd distance measuring technique on biological samples by using PR as a case study. We selected several pairs of residues with distances that could change upon the light-activation of PR due to its structural rearrangement. These pairs of residues were labeled with Gd-based spin labels. To further improve the sample preparation, we introduced a novel approach developed by collaborators to increase the spin labeling efficiency by removing non-fully-labeled PR in the sample. Static cryo-trapped PR structure was compared between samples under its resting state and the one populated with M intermediates by a continuous green-light laser illumination during flash-freezing. The distance change upon light-activation between Gd-labeled residues were first verified by a well-established pulsed EPR technique DEER under a Q-band frequency (34 GHz), and then measured by Gd-Gd distance measuring technique. Finally, the Gd-labeled PR was reconstituted into liposomes with a slower molecular tumbling as an approach of enabling room-temperature distance measurements. All these developments established a standard sample preparation protocol for future TiGGER measurements on the structural dynamics of protein, not limited to the case study of PR here.

6.1 PR Labeled with Gd-based Spin Labels

Variants of PR, each with a pair of cysteine mutations, were designed for measuring inter-residue structural movements associated with the light-activation of PR. The locations of the chosen pairs of sites on PR and the corresponding inter-residue distances are listed in Figure 6.1 A. The principle is to use reference sites (Site A), hypothesized to be static with respect to neighboring helices during the light-activation of PR, to track the movement of residues (Site B) on the E-F loop (174 & 177) and the F-helix (184) of PR, as these residues showed structural differences between cryo-trapped photointermediates of the homologous human rhodopsin.[213] The estimated inter-residue distances here are all < 3.5 nm and fall within the detectable range of both DEER and Gd-Gd distance measuring techniques while using Gd-MTS-ADO3A[214] as the Gd-based spin label to label PR (Figure 6.1 B). To quantitatively extract inter-residue distances between the pair of cysteines labeled with Gd-MTS-ADO3A (doubly-labeled PR), the cw EPR spectrum of doubly-labeled PR can be compared to the spectrum of a reference sample with only one of the cysteines being labeled (singly-labeled PR). In the case of a qualitative comparison of inter-residue distances between the resting state (light-off) and the excited state populated with M intermediates (light-on), the extent of lineshape broadening on the cw EPR spectra of the doubly-labeled sample under the two conditions can be directly compared to find the direction of inter-residue distance changes upon light-activation.



Site A	Site B	Est. distance (nm)
49	174 (E-F)	3.0
55	177 (E-F)	2.0
111	174 (E-F)	1.5
	184 (F)	2.0

Figure 6.1: Sites for Gd labeling on PR. A. List of doubly-Gd-labeled PR variants. Each with two selected sites mutated into cysteines that are later conjugated with Gd-based spin labels. Site A is the reference site and Site B is the site that hypothetically moves upon the light-activation of PR. The estimated inter-residue distances are listed in the table. **B.** Gd-MTS-ADO3A and Gd-TPATCN are the Gd-based spin labels used for labeling PR.

6.2 Enrichment of Doubly-labeled PR

For Gd-Gd distance measuring technique, the ideal doubly-labeled PR sample would be the one with all of the PR labeled with two Gd-MTS-ADO3A labels conjugated on the two designated sites. However, this requires the yield of Gd-MTS-ADO3A labeling reaction to achieve 100%, which is not possible in reality. Therefore, the cw EPR spectrum from a doubly-labeled PR sample could contain contributions from undesired non-fully-labeled PR. The non-fully-labeled PR with one cysteine site being labeled and the other not being labeled can give rise to an unwanted background signal, which can significantly dilute the lineshape broadening from the cw EPR spectra of doubly-labeled PR, which encodes the inter-residue distance information. Strategies including removing PR aggregates with low labeling efficiency by size exclusion chromatography (SEC), conducting the labeling reaction under room temperatures, and applying E50Q mutation to enrich monomeric PR with high labeling efficiency have been collaboratively implemented to effectively improve the labeling efficiency of Gd-MTS-ADO3A on PR.[112] Nevertheless, the labeling efficiency could still vary from site to site and often is less than 80%. Similar to the concept of process design in the chemical engineering field, where separation processes can always follow up a chemical reactor to further increase the purity of desired products. Instead of further optimizing the labeling reaction toward a yield that is close to 100%, selecting out the desired doubly-labeled PR from other impurities is an alternative approach to prevent the presence of unwanted background signals from non-fully-labeled PR in Gd-Gd distance measurements.

An unpublished protocol shared by our collaborator Dr. Janet Lovett has shown to effectively enrich the amount of doubly-labeled water-soluble myoglobin in their EPR samples for DEER measurement. This was done by removing the non-fully-labeled protein after the Gd-labeling reaction. Here, we tested whether the same process can be applied to PR as a transmembrane protein. Figure 6.2 demonstrates the flowchart of the doubly-labeled protein enrichment process. The imperfectly labeling of Gd-based spin label would lead to non-fully-labeled PR with free cysteines in the protein sample. By adding biotin-maleimide linkers to the spin labeled protein sample, the maleimide functional group can conjugate with the free cysteines on PR, and the non-fully-labeled PR then becomes the ones with a biotin ligand. The non-fully-labeled PR with a biotin ligand can be captured and removed by a streptavidin-agarose resin column. The flowthrough collected from the column turns into a sample enriched with doubly-labeled PR.

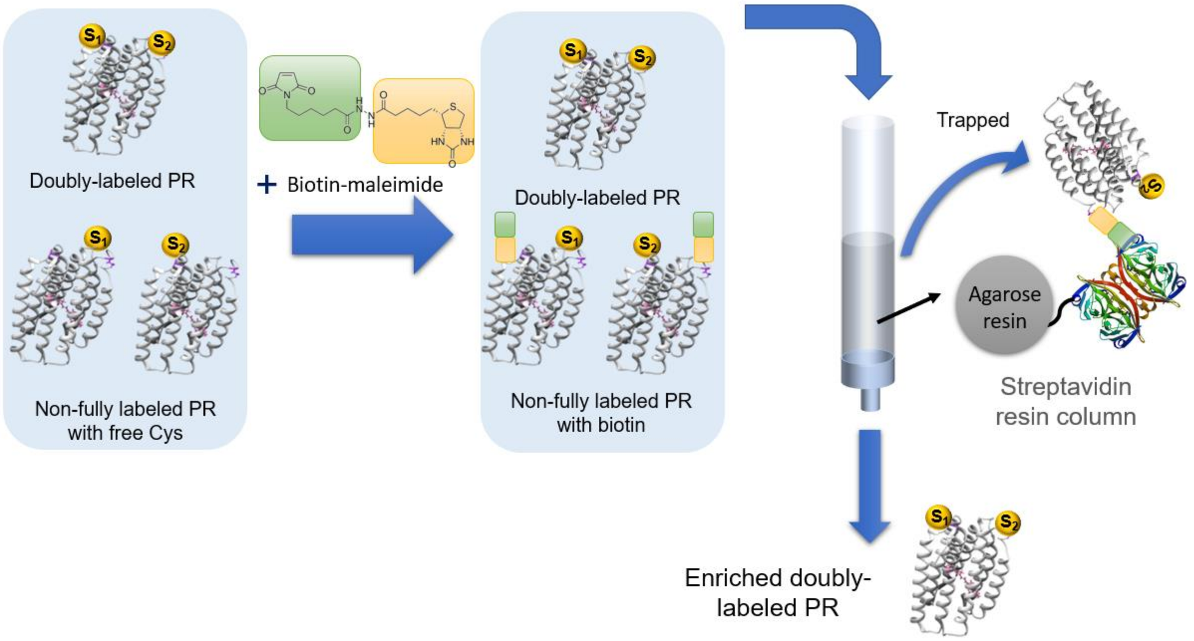


Figure 6.2: Enrichment of Doubly-labeled PR. The Gd-labeled PR samples are treated with biotin-maleimide that can conjugate with the non-fully labeled PR with free cysteines. The biotin-tagged PR can then be trapped by a streptavidin agarose resin column. The flow-through with the enriched double-labeled PR is then concentrated for spectroscopic measurements.

The effectiveness of this process on enriching doubly-labeled PR can be examined by Q-band DEER measurements. DEER is a pulse EPR technique can not only acquire the inter-residue distance of the doubly-labeled PR but also the modulation depth of the time-domain DEER data, which inherits the amount of electron spin pairs with dipolar modulation effect, can be utilized to examine whether the amount of doubly-labeled PR has been enriched through the above process. If the enrichment process successfully eliminates the non-fully-labeled PR in the sample, we can expect a deeper modulation depth from the time-domain DEER trace.

Time-domain DEER traces of both doubly-labeled 174-49 and 184-111 PR before and after the enrichment process were compared in Figure 6.3. For the doubly-labeled 174-49 PR, a 10 to 14% increase of the modulation depth can be found from the estimation by the LongDistances software after background subtraction and a fitting on the raw time-domain traces. In the case of doubly-labeled 184-111 PR, not as significant as the case of 174-49, a distinguishable increment of the modulation depth from 4 to 5% can still be found after the enrichment treatment. These data suggested that the enrichment treatment successfully removed part of the undesired non-fully-labeled PR without a pair of Gd spins.

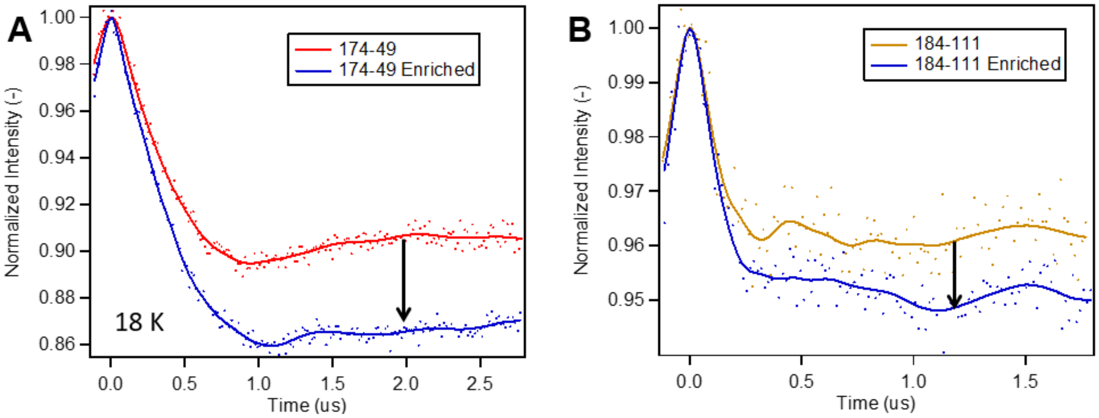


Figure 6.3: Time-domain DEER traces of Enriched PR. A comparison of time-domain DEER traces before and after the enrichment treatment measured from **A.** 174-49 doubly-labeled PR and **B.** 184-111 doubly-labeled PR. The measurements were done under 18 K.

To further examine whether the inter-residue distances of the two doubly-labeled mutants are valid, the time-domain DEER traces of both doubly-labeled 174-49 and 184-111 PR after the enrichment treatment were analyzed using a Gaussian distribution model to get the average inter-residue distances. For doubly-labeled 174-49 PR, the average inter-residue distance of the Gaussian fitting is 4.2 nm. This inter-residue distance measured by DEER is ~1.4 nm longer than the estimated inter-residue distance from PR’s PDB structure (Figure 6.4 A). The discrepancy between the measured distance and the estimated one could be contributed from the length of the linker on the Gd-MTS-ADO3A molecule. In the case of DEER measurement, the inter-residue distances were the distance between Gd ions on the attached spin labels. However, the estimated inter-residue distances in Figure 6.1 A only took the distance between Cα atoms on site 174 and site 49 into account. The distance between the Cα atom on the amino acid and the Gd ion on the attached spin label can add up additional distances in the DEER measurement compared to the estimated inter-residue distances. This is in line with the longer distance we observed here from the DEER data. The same longer measured average distance (3.0 nm) can also be observed from the doubly-labeled 184-111 PR compared to its estimated distance (2.0 nm).

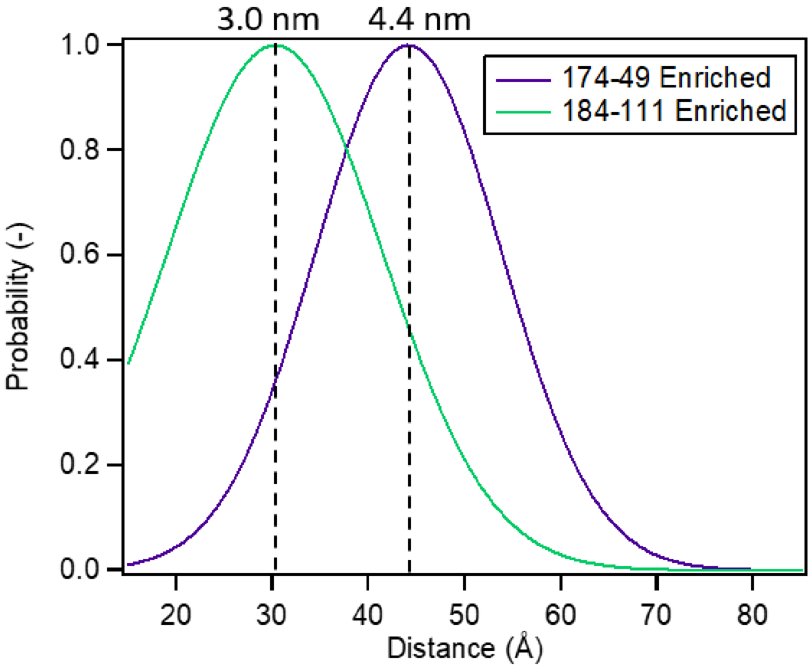


Figure 6.4: Inter-residue DEER Distance Distributions. Distance distributions fitted by the Gaussian model of both doubly-labeled 174-49 (purple) and 184-111 (green) PR after the enrichment treatment.

A justification of the longer inter-residue distances measured by Q-band DEER was done by multiscale modeling of macromolecules (MMM) open-source MATLAB toolbox.[215] MMM can simulate the inter-spin distance distributions between the pair of spin labels that are attached on the protein, with different conformers of both the protein structure and the spin label considered. After the protein structure is imported from the protein data bank to the MMM, a rotamer library approach is used to add precomputed spin label rotamers that fit into the spatial constraints created by surrounded molecules on the protein. Here, Gd-DOTA from the MMM databank that is similar to the Gd-MTS-ADO3A used in DEER measurements but with a different maleimide linker is added to both site 174 and site 49 on the PR solution NMR structure (PDB: 2L6X[52]). All the possible conformations of the protein and the attached Gd-DOTA on the two sites are simulated and presented in Figure 6.5 A. The ribbons are the backbone of the twenty PR structure rotamers and the ball-and-stick part represents all the possible rotamers of attached Gd-DOTA spin labels on the two sites. The distance distribution between the Gd ions on each doubly-labeled PR structure rotamer was simulated by the DEER simulation tool built-in MMM toolbox. Here we selected six representative distance distribution plots from the twenty rotamers and presented in Figure 6.5 B. The average distances of these distributions range from 3.7 to 4.6 nm, which is close to the average distance found from our DEER data (4.2 nm). These distance distribution plots together with the simulated attached spin labels concur that the inter-residue distance measured from DEER could be 0.7 to 1.7 nm longer than the one estimated from the Cα between the two sites. The longer distances here can be contributed from the linker length on the Gd-based label together with the propensity of the labels to point away from the center of the PR.

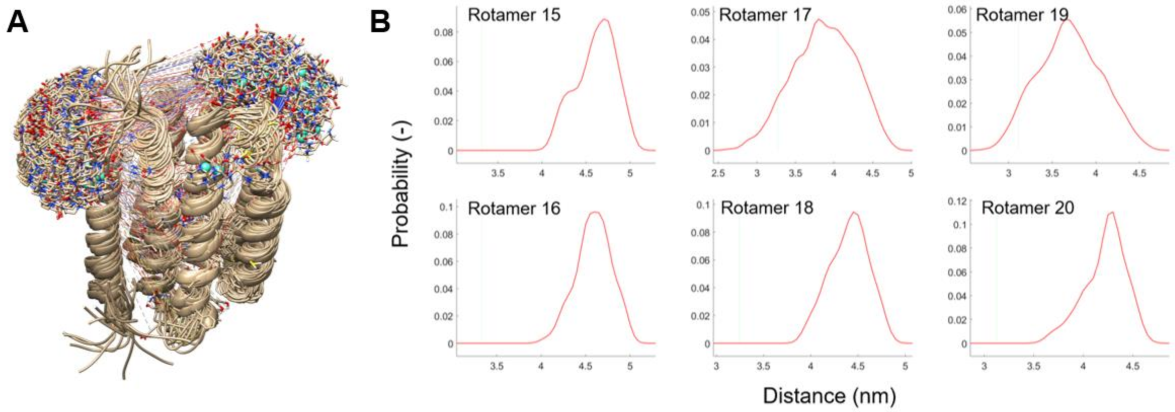


Figure 6.5: Distance Distributions Simulated by MMM. A. All possible rotamers of Gd-DOTA spin labels conjugated on both site 174 and site 49 of PR. The structure of PR was adapted from PDB 2L6X[52] and the figure was created using the UCSF Chimera package.[216] The ribbons are the backbone of PR and the ball-and-stick modeled molecules are the Gd-DOTA. **B.** Simulated distance distributions between the pair of Gd spins attached on site 174 and site 49 by using the MMM toolbox.[215] The six distance distribution plots are from the 15th to 20th rotamers among the 20 rotamers in the solution NMR structure of PR.

The next step was to examine how the enrichment process improved the cw EPR spectra by TiGGER of the two doubly-labeled (174-49 & 184-111) PR with Gd-MTS-ADO3A (Figure 6.6). The spectrum of enriched 174-49 had a greater signal to noise ratio by comparing the one from the same protein but before the enrichment treatment, despite the two samples had a similar protein concentration (~200 μM). Not as significant as the case of 174-49, the cw EPR spectrum of enriched 184-111 still had a slightly better signal to noise ratio. With all the samples prepared with a similar protein concentration, these results suggest that the spin labeling concentration is higher after the enrichment process. This higher spin concentration of the enriched samples can be a result of the non-fully-labeled PR removal after the enrichment process. Overall, the enrichment process improved the signal-to-noise ratio and removed the non-fully-labeled PR to improve the accuracy of TiGGER spectra by removing the undesired background.

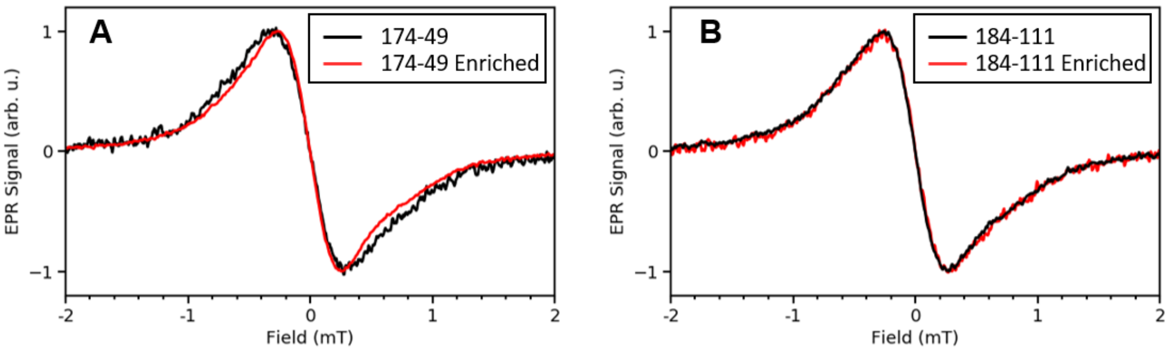


Figure 6.6: cw EPR on Doubly-labeled PR after Enrichment Treatment. The cw EPR spectra acquired from doubly-labeled **A.** 174-49 and **B.** 184-111 PR with Gd-MTS-ADO3A spin labels. The measurements were done under 30 K.

6.3 Light-activation of PR

The light-activation structural movements on PR between the pair of labeled sites were examined by both Q-band DEER measurements and TiGGER measurements. The well-established DEER measurements on Gd-labeled PR can provide more exact information on the change of distance distribution after light-activation. This can provide a guideline on what can be expected to observe on the lineshape broadening in the TiGGER measurements. Here, DEER distance distributions of doubly-labeled 174-49, 174-111, and 184-111 PR were compared between their own cryo-trapped resting state (light-off) and light-activated states populated with M intermediates (light-on).

Light-induced changes in DEER time-domain traces and fitted Gaussian distance distributions can be observed from doubly-labeled 174-49 as shown in Figure 6.7. While the time-domain DEER trace of the light-off state here is the same one as presented in Figure 6.3 A, the light-on activated state of 174-49 had a time-domain DEER trace with a longer dipolar oscillation period (Figure 6.7 A). This was also reflected in a shift of the average distance from 4.2 nm to 4.4 nm in the fitted distance distribution (Figure 6.7 B). The same longer average distance of 4.4 nm can also be reproduced from the same sample to be re-trapped into the light-on activated state for the second time ("on 2" in Figure 5.7), which justified the observed distance change is a true structural movement between site 174 and site 49 but not an artifact introduced from repeating free-thaw cycles during the sample preparation. Although the average distance change between site 174 and 49 upon light-activation was within the errors from the fitting, the reproducibility of the same longer average distance from the light-on condition made the observed change more convincing. As a twisted and tilted movement at the E-F loop of PR has been suggested in its light-on activated state, the DEER distance distribution here confirmed the movement at site 174 that has been unveiled by X-band cw EPR and ODNP relaxometry,[95] and further provide the information on the quantitative distance change between the helices across the PR.

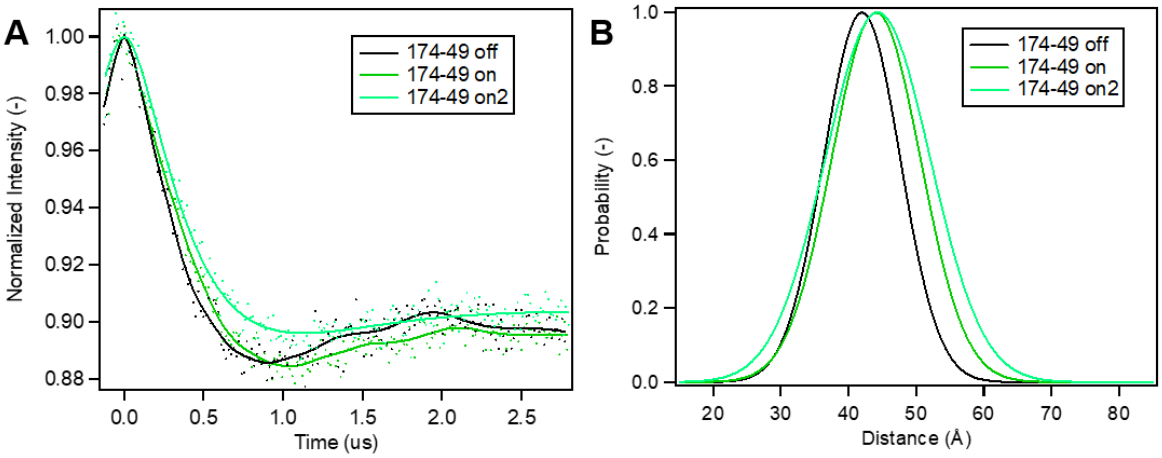


Figure 6.7: DEER of PR 174-49 upon Photoactivation. A. Time-domain DEER traces of the light-off (black) and the light-on (green) cryo-trapped doubly-labeled PR with Gd-MTS-ADO3A at site 174 and site 49. The dots represent the original time-domain data and the lines represent the fitted time-domain traces. **B.** DEER distance distributions of the light-off (black) and the light-on (green) PR fitted by a Gaussian model. These measurements were done under 18 K. The sample was treated with the previously described enrichment process to remove non-fully-labeled PR.

Slight distance changes in DEER distance distributions can also be observed from the other two doubly-labeled PR (184-111 and 174-111) upon photoactivation as shown in Figure 5.8. Since the modulation depth in the time-domain traces of these two doubly-labeled PR sample were small even after the enrichment treatment (< 10%), a more scattered time-domain raw data can be found for both light-on and light-off PR (Figure 6.8 A&C). This led to fittings that cannot fully describe the dipolar oscillation encoded in the raw data. Nevertheless, the light-on 184-111 PR still showed a 0.1 nm shorter average distance compared to the light-off case (3.0 nm vs. 2.9 nm). Similar distance distribution of the light-off PR can be found from the same sample that was once trapped into light-on state and re-trapped into the light-off state for the second time (off 2, Figure 6.8 B), fortified the credibility of the obtained distance distributions even under the low modulation depth here in the DEER raw data. The light-activated movement at site 184 observed here is new to PR, as this site is not on the E-F loop but the F helix of PR. A similar outward movement on the F-helix has been unveiled in the case of human rhodopsin,[213] which is a retinal-bonded GPCR with a similar 7TM structure as the case of PR. On the other hand, the average distance of 174-111 PR between the two labeled sites in its light-off state is 2.2 nm, while a 0.1 nm longer average distance can be found from the light-on state (Figure 6.8 D). All these data together suggest the chosen inter-residue distances on PR should change between the light-off resting state and the light-on activated state populated with M intermediates, and the change of the average distance induced by light-activation is ~0.2 nm. More importantly, the change of distances observed by DEER reassured all these doubly-labeled PR could be great candidates for tracking the structural dynamics of PR function by TiGGER.

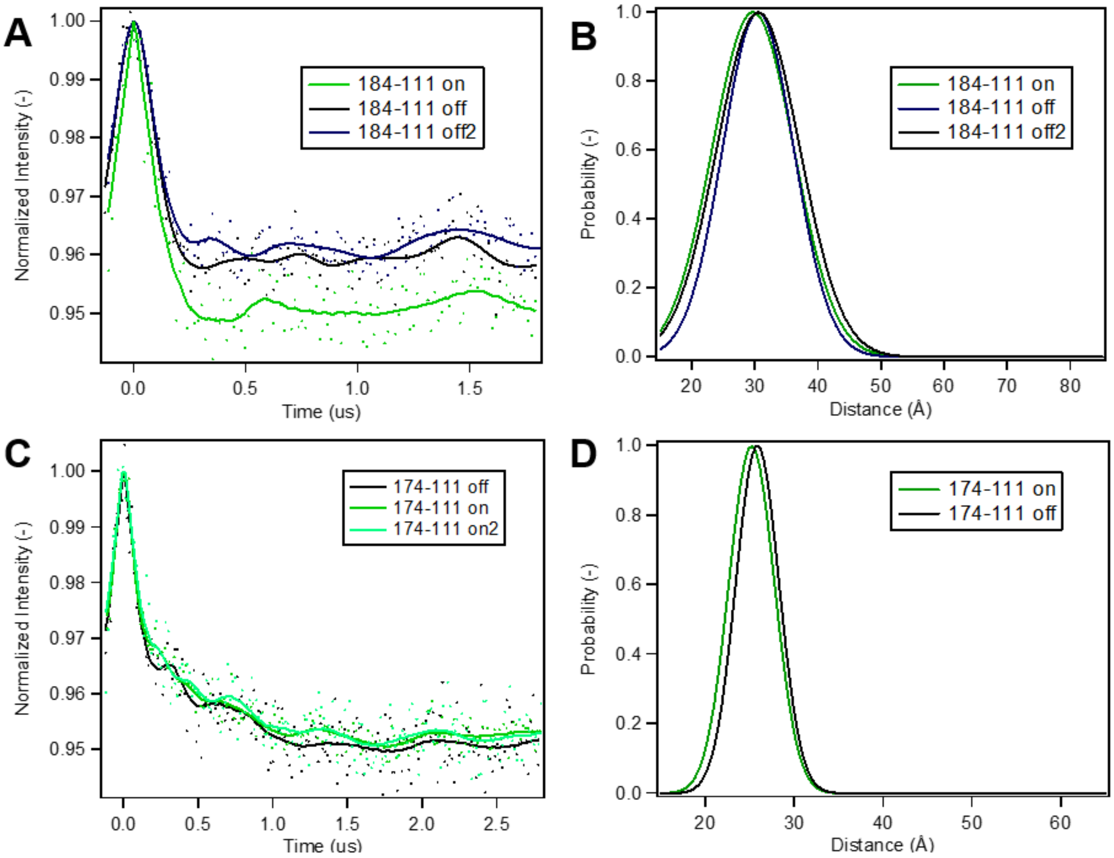


Figure 6.8: DEER of PR 184-111 and 174-111 upon Photoactivation. A. Time-domain DEER traces of the light-off (black), the light-on (green) and the second light-off (blue) cryo-trapped doubly-labeled PR with Gd-MTS-ADO3A at site 184 and site 111. The dots represent the original time-domain data and the lines represent the fitted time-domain traces. **B.** DEER distance distributions of the light-off (black), the light-on (green) and the second light-off (blue) PR 184-111fitted by a Gaussian model. **C.** Time-domain DEER traces of the light-off (black), the light-on (green) and the second light-on (cyan) cryo-trapped doubly-labeled PR with Gd-MTS-ADO3A at site 174 and site 111. The dots represent the original time-domain data and the lines represent the fitted time-domain traces. **D.** DEER distance distributions of the light-off (black), the light-on (green) and the second light-on (cyan) PR 174-111 fitted by a Gaussian model. These measurements were done under 18 K.

We then examined whether the light-activated movement of doubly-labeled PR can also be observed from the TiGGER measurements. The cw EPR spectra of cryo-trapped PR under its resting state (light-off) were compared with the ones under its light-activated state (light-on) populated with M intermediates. No significant lineshape difference can be observed between the light-on (orange) and the light-off condition in the case of doubly-labeled 174-49

PR (Figure 6.9 A). This is reasonable as the average inter-residue distance between site 174 and site 49 estimated by DEER is 4.2 nm, in which most of the inter-spin distances are above the detectable limit (3.5 nm) of using Gd-MTS-ADO3A spin label in Gd-Gd distance measurements. The small average distance change (0.2 nm) upon light-activation under this long inter-spin distance could be hard to be reflected in a change of lineshape broadening in the cw EPR spectra. Interestingly, in the case of doubly-labeled 184-111 PR (Figure 6.9 B) that has an average distance of 3.0 nm under its resting state, a slight broadening can be observed from the cw EPR spectra of the under the light-on condition (orange) compared to the light-off condition (blue). The broadening can mainly be observed from the wing of the central transition (~ 1 mT above the center field), but not significant enough to broaden the peak-to-peak linewidth. Nevertheless, the broadening of the cw spectra infers a shorter inter-residue distance while PR is in light-on condition, which coincides with the shorter distance between site 184 and site 111 measured by DEER upon light-activation (2.9 nm vs. 3.0 nm).

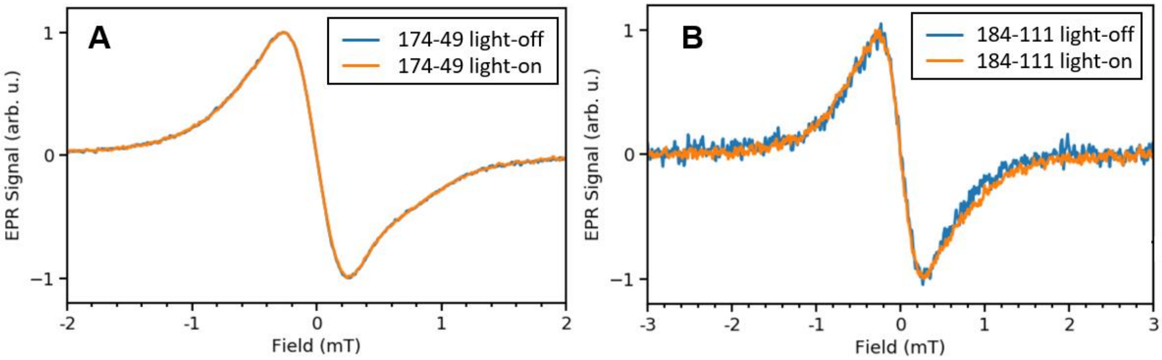


Figure 6.9: cw EPR of PR 174-111 and 184-111 upon Photoactivation. Comparing the cw EPR lineshape of doubly-labeled PR **A.** 174-49 and **B.** 184-111 between the resting state (blue) and the light-activated state populated with M intermediates (orange). The measurements were done under 30 K.

6.4 Gd-Gd Distance Measurement under RT

To achieve the goal of using TiGGER to measure the dynamics of PR light-activated movement, the Gd-Gd cw EPR distance measurements will need to be conducted under room temperature instead of the cryogenic temperature (30 K) used in previous sections. The difficulty arises from the molecule tumbling of spin labeled PR in the solution under room temperature. This tumbling could average out the anisotropic dipolar interaction between Gd spins. In this case, even the pair of Gd-chelates are within the detectable 3.5 nm distance, the cw EPR spectra could have no lineshape broadening because of the averaging effect from molecular tumbling. In the case of Gd-rulers, Gd-Gd cw EPR measurements under 288 K were done by immobilizing the rulers in the dehydrated trehalose matrix.[108] This glassy surrounding slowed down the tumbling of Gd-rulers and prevented the dipolar interaction between spins to be averaged to zero. This method worked perfectly in the case of Gd-rulers as the same trend of dipolar lineshape broadening along with the change of inter-spin distances can be observed under either 30 K or 288 K. While it is possible to apply the same approach by preparing the Gd-labeled PR sample inside the trehalose matrix, however, a slower tumbling of PR can also be achieved by reconstituting PR into a larger liposome vesicle (above 200 nm in diameter). The latter method is a superior one for membrane protein studies as lipid bilayer is a more native-like environment for transmembrane proteins compared to surfactant micelles. As a result, the next goal is to examine whether reconstituting doubly-labeled PR into liposomes can preserve its dipolar broadening in the cw EPR spectra.

The cw EPR spectra of doubly-labeled 174-49, singly-labeled 174, and singly-labeled 49 reconstituted in either DDM detergent micelles or POPC/POPG (80/20, mol/mol) liposomes taken under 288 K are compared in Figure 6.10. In the DDM detergent micellar environment, no significant lineshape broadening was observed from the doubly-labeled 174-49 (green) compared to the other two singly-labeled references (174 & 49) without pairs of Gd-MTS-ADO3A on a PR (Figure 6.10 A). On the contrary, while PR was reconstituted in liposomes with a slower tumbling (Figure 6.10 B), a clear lineshape broadening was observed from the doubly-labeled 174-49 (red) compared to the singly-labeled 174 reference (purple). The cw EPR spectra of the same two mutants but in DDM detergent micelle environments in Figure 5.10 A is also added in Figure 6.10 B for comparison. It is clear that the cw EPR spectrum of doubly-labeled 174-49 PR reconstituted in liposomes have the broadest peak-to-peak broadening. This is in line with what we expected, as the reconstitution of PR into liposomes slows down the tumbling of the protein and preserves the dipolar-induced lineshape broadening in the cw EPR spectra. The extra peak-to-peak broadening for the liposome samples under room temperature could also increase the resolution of the Gd-Gd distance measurement. Nevertheless, no lineshape difference on the cw EPR spectra of doubly-labeled 174-49 PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes can be observed upon light-activation (data not shown). The same test will need to be conducted on a doubly-labeled PR with a shorter inter-residue distance (e.g. 184-111) so that the light-induced distance change could be resolved. Once there is a clear lineshape difference observed under room temperature, the next step would be to conduct TiGGER to track the dynamics of the lineshape change upon the light activation of PR.

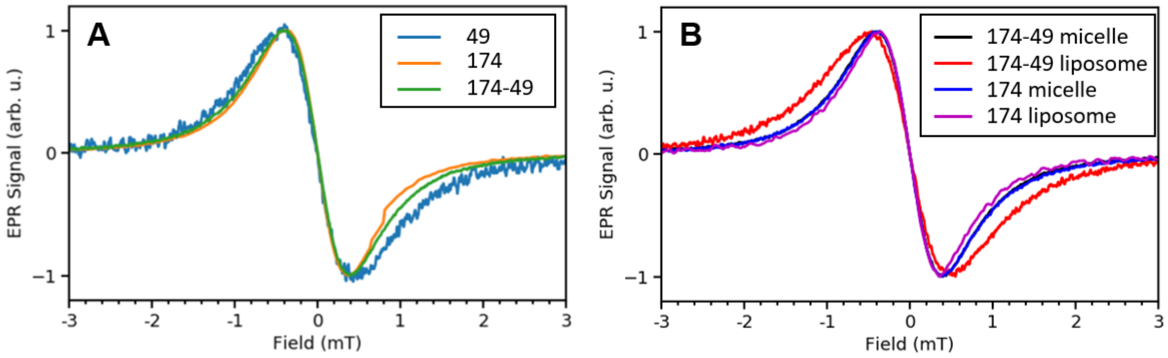


Figure 6.10: cw EPR of PR 174-49 under Room Temperature. Comparing the cw EPR lineshape of doubly-labeled 174-49 and singly-labeled references (174 & 49) while the PR are in **A.** DDM surfactant micelles or are reconstituted into **B.** POPC/POPG (80/20, mol/mol) liposomes. The measurements were done under 288 K. The doubly-labeled 174-49 sample was treated with the previously described enrichment process to remove non-fully-labeled PR.

6.5 Outlook for Structural Dynamic Measurements

In this chapter, we optimized the sample preparation for the developing TiGGER technique that aims to measure the dynamics of protein functionally relevant structural movements. The two main hurdles of the Gd-based cw EPR measurements on biological samples were the low labeling efficiency of Gd-based spin labels on proteins and the limitation on operating measurements under cryogenic temperatures. Here, the introduction of the novel enrichment process after the labeling of Gd-chelates on proteins effectively removed non-fully-labeled PR, proven by a better modulation depth in DEER measurements and a better signal-to-noise ratio in Gd-Gd distance measurements. On the other hand, the reconstitution of PR into liposomes slowed down its tumbling and preserved the dipolar broadening feature in the cw EPR spectra of a doubly-labeled PR sample under room temperature (288 K). This opened the possibility for the Gd-Gd distance measuring technique to measure the functional movements of proteins under physiological temperatures.

Nevertheless, the sensitivity of this technique still needs to be improved to capture the structural rearrangement of PR in its photochemical reaction cycle, where the inter-helical distance changes during the reaction can be as small as 1 or 2 Å, based on what we have observed from the doubly-labeled 174-49 and 184-111 PR using DEER. The sensitivity of the Gd-Gd distance measuring technique and the time-resolved version of it (i.e. TiGGER) is highly dependent on the inter-spin distance that we are measuring. Even the upper limit for the dipolar broadening between Gd ions to be resolvable is ~3.5 nm in the case of the Gd-MTS-ADO3A spin label, the quantity of lineshape broadening difference induced by a distance change is significantly smaller while the inter-spin distance is closer to this limit. Therefore, judiciously selecting the pair of moving and reference residues to have a 2-3 nm distance between the attached pair of spin labels is more ideal to achieve a higher sensitivity in this technique. The study here also unveiled the importance of taking the linker length on Gd-chelates into consideration while estimating the inter-residue distances. The case of doubly-labeled 174-49 has a 4.2 nm inter-spin distance measured by DEER, which is longer than the estimated distance between the two amino acids on the PR structure. As a result, even a ~2 Å average distance change can be observed between site 174 and site 49 upon light-activation of PR by DEER, the 4.2 Å distance is too long for the Gd-Gd distance measuring technique to capture any distance change through the change of cw EPR lineshape broadening. In the future, doubly-labeled PR with a shorter inter-residue distance should be selected for probing the light-activated structural change of PR, including the future TiGGER measurement.

Using a spin label with a narrower intrinsic linewidth (i.e. a smaller zero-field splitting) can further increase the upper distance limit in Gd-Gd distance measuring technique. This can also provide a higher resolution on distance changes within 3.5 nm. Gd-TPATCN,[217] with a zero-field splitting that is ~1.5 times smaller than the one of Gd-MTS-ADO3A, is one of the options that can achieve the goal of increasing the sensitivity of the Gd-Gd distance measurement and the TiGGER measurement on protein samples in the future. On the other hand, the shorter linker length on this spin label can also make it more rigidly conjugated to the target protein, and in turn, becomes a label that can more royally reflect the position of the labeled amino acid on the protein. With all the above developments, the goal of measuring the structural dynamics of PR and understanding the mechanical basis of its proton transport function can be achieved in the future by TiGGER. Moreover, the fruitful sample preparation guideline here can also be applied to protein systems beyond PR, in the case of using the same TiGGER technique to measure triggered functional dynamics of proteins, for example the structural rearrangement of GPCR after a ligand-binding event.

Chapter 7

Conclusion and Outlook

This dissertation studied the effect of several potential modulators on the structure and function of the light-activated proton pump PR in the native-like liposome environment. Two important aspects of the PR proton transport function were examined, the protonation state of embedded D97 that controls the population of active PR and the photochemical reaction cycle kinetics that determines the rate of proton transport. Our studies found these two aspects of PR proton transport function were independently tuned by separate factors in the native-like liposome environment. Protein-protein interactions of PR oligomeric assembly slowed down the growth and the decay of M intermediates in its photochemical reaction cycle (Chapter 2), same as the case for PR in the detergent micellar environment. In contrast, the pKa_{D97} that controls the population of active PR was not affected by the different oligomeric distribution of PR in the liposome environment but can be systematically modulated by the net charge of liposomes and the ion type or concentration in aqueous solution in which PR was reconstituted (Chapter 3). These observations together suggested that PR's photocycle kinetics and pKa_{D97} were adjusted through different underlying mechanisms.

For the protonation behavior of the embedded D97 that controls the population of active PR, this dissertation found that both apparent pKa and the intrinsic pKa of the D97 residue were modulated by the electrostatic environment of PR-constituted liposomes (Chapter 3). Since the intrinsic pKa modeled with the Gouy-Chapman equation had already taken the deviation of local pH from bulk pH on charged liposome surfaces into account, the change of intrinsic pKa inferred that the intrinsic structural properties of PR must play a role in adjusting the local microenvironment around the embedded D97. The long-range interactions between charged lipid headgroups and the embedded D97 were likely mediated by a change of the internal electrostatic network constructed by other amino acids on PR. Amino acids that were rearranged by different external electrostatic environments were identified by using complementary magnetic resonance spectroscopic techniques, and these residues were identified as candidates that take part in the internal electrostatic network coupled with D97 (Chapter 4). We found a change in the protonation behavior of the Nδ1 atom on the H75 in the vicinity of D97, where the change was in line with what was found from the D97 residue. This inferred that the pH range for the H75-D97 hydrogen bond formation can be affected by the external electrostatic environment. On the other hand, we also found structural rearrangements on amino acids that are remote from the embedded D97 residue, including the residues on the E-F intracellular loop. These all helped to construct the internal electrostatic network that could play a role in determining the pKa of the embedded D97 while PR is reconstituted in different external electrostatic environments.

Our findings underscored the importance of the environments, including the valency and the concentration of cations in buffer or the net charge of surrounding biomimetic environment, in modulating the activity of transmembrane proteins. The knowledge was successfully utilized to optimize the proton transport capability of PR in a biomimetic material—surfactant-silica mesostructured film that can increase the thermal and the mechanical stability of incorporated proteins. The population of active PR with vectorial proton transport capability increased while positively-charged lipids were introduced into the silica-surfactant host material (Chapter 5). Our guideline on tuning the function of transmembrane proteins by adjusting the electrostatics of the biomimetic environment can also be applied to other proteins with functions controlled by a similar mechanism as PR.

The developing time-resolved Gd-Gd EPR (TiGGER) technique has a great potential to measure the structural dynamics of proteins. This information can help us to resolve challenging questions like how oligomerization affects the function of transmembrane proteins with structural movements involved. By using PR as a case study, the comparison of its structural dynamics and its photocycle kinetics can provide insights on structure-function relationships. Our preliminary results provided a guideline on selecting proper pairs of sites for spin labeling to measure distances, increasing the resolution of measured cw EPR spectra, and enabling distance measurements on transmembrane protein samples at room temperature (Chapter 6). These all can be applied to future TiGGER structural dynamic studies on proteins not limited to PR. The unique structural dynamic information from TiGGER can resolve how proteins operate as machines to achieve functions triggered by different factors.

While this dissertation mainly focused on the structure-function relationship of PR under the influence of different modulators, our findings have a great potential to be generalized to studies of other transmembrane proteins, especially to other optogenetic microbial rhodopsins that shares a similar light-triggered ion transport function or drug targeting GPCRs that shares a similar 7TM structure. Despite breakthroughs in the past decades have uncovered fruitful knowledge on the functions of transmembrane proteins in different organisms, there is still a myriad of protein functionalities and structural behaviors yet to be discovered. We hope our work can be a cornerstone to resolve part of these long-standing scientific mysteries related to transmembrane proteins.

Appendix A Material and Method: Preparation of PR Samples and Biochemistry

A.1 PR Expression and Purification

The PR construct used in this dissertation is the “green-light-absorbing” GPR variant derived from the first PR gene sequences (BAC31A8). The WT PR construct used for the studies in this dissertation has three natural cysteines (C107, C156, C175) mutated into serines, and a 6x-His tag is added on its C-terminus for the purification purpose. As cysteines could oxidize over time, the removal of natural cysteines can improve the functional stability of PR for biophysical studies and showed no functional impact compared to the natural “wild-type” PR.[85] For all the EPR studies in Chapter 4 and 5, single or double cysteines were introduced to the desired sites, and these cysteines were later conjugated with EPR sensitive spin labels. For ssNMR studies in Chapter 5, we remove the 6x-His tag on PR by a tobacco etch virus (TEV) protease (Sigma Aldrich, MO) that recognized and cleaved the ENLYFQS sequence between the recombinant PR and the 6x-His tag after the affinity column purification. Functional-relevant residues were also mutated intentionally for different experimental purposes. For the PR oligomer functional studies in Chapter 2, the E50Q mutant that increased the population of monomeric PR and the W34D mutant that disrupted the cross-protomer interaction were utilized as monomeric PR references in liposomes. For the study of comparing the structure of light-activated and resting PR in Chapter 5, the E108Q mutant was introduced to slow down the M-N transition of the photochemical reaction cycle and populate the desired M intermediates under continuous green-light illumination.

A.1.1 Cloning and Expression of PR

The PR gene was cloned into a pET26b (+) vector (Novagen, Madison, WI). In the case of non-WT mutants, desired mutants were made by a site-directed mutagenesis method with designed mutation-induced primers and polymerase chain reactions (PCR). A two-stage PCR technique was also incorporated here to prevent primers from forming self-dimer instead of binding to the PR template during the PCR process. The expression of recombinant PR was done by transforming prepared vectors into *E. Coli* strain BL21(DE3) competent cells (Agilent), growing bacteria in liter-scale flasks with LB media, and inducing overexpression by adding isopropyl-β-D-thiogalactopyranoside (IPTG) and *all-trans*-retinal. The overexpressed PR in *E. Coli* cells with a pink-colored appearance was then collected by centrifugation for downstream purification.

In the case of isotope-labeled PR expression for NMR sample preparation, *E. Coli* strain BL21(DE3) with transformed vectors were cultured in a 1L LB media until it reached an optical density between 0.8 and 1.2. The cells were collected by centrifugation at 4,000 rpm for 10 min and washed by an equal volume of M9 minimal media. The cells were spin down again at 4,000 rpm for 20 min and resuspended in a 250 mL final isotope-enriched culture medium for PR overexpression. The lower volume of the medium used here led to a higher concentration of cells and was reported to have a better expression yield. The medium with resuspended cells was first shaken under 200 rpm at 37^{o}C for 1 hr and was added with IPTG and all-*trans* retinal to inoculate the overexpression. The rest of the purification steps were the same as the one used for non-isotope-labeled PR.

The recipe of the final isotope-enriched culture medium can vary between uniformly-isotope-labeled expression and selective-labeled expression. In the case of selective-labeled PR with ^{15}N on the histidine, the medium recipe was adapted from Mehler et al.,[218] but with the histidine substituted by L-$^{15}N_3$-Histidine (Cambridge Isotope Laboratories, Inc., MA). In the case of uniformly-^{13}C-labeled PR, the medium recipe was adapted from Marley et al.,[134] with the D-glucose substituted by D-glucose-$^{13}C_6$.

A.1.2 Purification of PR

Our purification protocol involved cell lysis to digest *E. Coli* cell wall, freeze-fracture, and sonication to break the inner membrane into small fragments so that PR in cell lysate can later be extracted and solubilized by a buffer (50 mM potassium phosphate dibasic buffer, 150 mM potassium chloride, pH 8.2) with 2 % (w/v) of β-D-dodecyl maltoside (DDM) detergents. The DDM-solubilized PR was then further purified by a Ni-NTA affinity resin column that can selectively collect PR with the 6x-His tag. After washing the resin column with the buffer that has a mild concentration of imidazole (50 mM) for at least six times to remove other impurities, the DDM-solubilized PR on the resin was then eluted by a buffer (50 mM potassium phosphate dibasic buffer, 150 mM potassium chloride, 0.05 % (w/v) DDM, pH 8.2) with 300 mM of imidazole. The PR was then buffer exchanged into the same buffer without imidazole and concentrated by Amicon Ultra centrifugal filter units with a 50 kDa MWCO (Millipore, Burlington, MA) for storage or later sample preparation.

A.1.3 Spin Labeling of PR

For labeling of both the nitroxide-based 1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl methanesulfonothioate spin label (MTSL) and the Gd-based spin label, PR was pre-treated with 10 mM DTT at 4°C overnight to ensure the thiol groups on cysteines were reduced. The DDT in the buffer was then removed by a PD-10 desalting column (GE Healthcare, Chicago, IL) containing Sephadex G-25 resin, immediately followed by adding a ten-time molar excess of the spin label into the PR solution to initiate the labeling reaction. The mixture was gently shaken overnight under room temperature for the reaction to complete. The excess spin label was removed via size exclusion chromatography, where PR was run through a HiLoadTM 16/600 SuperdexTM 200 pg column (GE Healthcare, Chicago, IL) connected to a NGC^{TM} Medium-Pressure Liquid Chromatography System (BioRad, Herculus, CA).

In the case of the Gd-based spin label, Gd-chelates were prepared by mixing Gadolinium chloride ($GdCl_3$) and MTS-ADO3A spin tags (provided by Weizmann Institute of Science) in MilliQ water with a 1.1 to 1 molar ratio. The mixture was constantly adjusted to pH 5.8 by adding 1M NaOH and incubated under room temperature for ~2.5 hr to drive the chelation reaction. The prepared Gd-MTS-ADO3A spin labels were either directly used for the labeling reaction described above or stored in the freezer for later use.

To enrich the doubly-Gd-labeled PR by removing non-fully-labeled PR in the sample, the spin labeled PR sample was mixed with a 0.01 mg/mL maleimide-biotin stock solution prepared in the same buffer for solubilizing PR with 0.05 % (w/v) DDM to achieve a 1 to 1 molar ratio between cysteines in the system the maleimide-biotin. The mixture was incubated under room temperature for 30 min for the maleimide-biotin to bind on the non-Gd-labeled cysteines on PR. After the incubation, the mixture was added to the Pierce® Streptavidin agarose resin (Thermo Scientific) pre-equilibrated with the same buffer at an amount that can capture all the biotins in the solution. The mixture with the Streptavidin resins was gently shaken at 4^oC for 1 hour and resins bind with biotinylated non-fully-labeled PR were removed. The enriched PR sample was then concentrated for spectroscopic measurements.

A.2 Preparation of PR Liposomes

Large unilamellar vesicles (LUVs) with the desired composition were prepared by a lipid extrusion method. Lipid stocks dissolved in chloroform were purchased from Avanti Polar Lipids (AL) and mixed to achieve a desired molar ratio between different lipid species. The lipid mixture was then dried under a nitrogen stream and further desiccated under vacuum overnight to ensure the removal of residue chloroform. The dried lipids were reconstituted with a HEPES buffer (10 mM HEPES, 150 mM NaCl, pH 6.7), and lipid vesicles were extruded through the Avanti mini-extruder for 19 times using filters with 200 nm pore size. The prepared lipid vesicles were then mixed with a DDM surfactant solution to achieve a final DDM concentration 2 times the DDM critical micelle concentration (CMC, 0.0088 w/v %). The mixture was gently shaken for 1 hr to form lipid-surfactant complexes, and the desired variant of PR was then added to the mixture with a 1:50 PR-to-lipid molar ratio. The DDM surfactants in the PR-lipid-DDM complex were then removed by using six vials of ~160 mg polystyrene BioBeads SM2 (Bio-Rad) to drive the formation of PR proteoliposomes.

A.3 Preparation of PR Styrene-maleic Acid Particles (SMALPs)

PR styrene-maleic acid particle (SMALP) samples were prepared from PR proteoliposome samples. The PR proteoliposome sample was first concentrated to a 30 mM total lipid concentration. An equal volume of 2.5 wt% SMA polymer solution (Polyscience, Netherlands) with a 2.3:1 styrene-to-maleic-acid ratio was then added to the concentrated PR proteoliposome sample dropwise with constant mixing. The mixture was equilibrated under room temperature for 30 min and then further equilibrated overnight at 4^{o}C. The formed PR SMALPs were centrifuged under 40,000 g for 30 min under room temperature to remove non-solubilized impurities.

A.4 Preparation of PR in Mesostructured Silica Film

Purified WT PR in 50 mM potassium phosphate dibasic buffer with 150 mM potassium chloride and 0.05 wt% DDM was first concentrated to ~300 μM by Amicon Ultra centrifugal filter units with a 50 kDa MWCO (Millipore, Burlington, MA). Hydrolyzed silica precursor mixture for film synthesis was prepared by mixing tetraethoxysilane (TEOS), n-propyltriethoxysilane, and 4 mM HCl aqueous solution (3:1:80, w/w) at room temperature with constant stirring for 2 h. After the silica precursor mixture was hydrolyzed, concentrated PR, additional DDM, and surfactant composite were added to meet the desired composition (5 wt% PR, 59.4 wt% DDM, 7.1 wt% surfactant composite, 28.5 wt% SiO_2). The mixture solution was adjusted to pH~4 by 1 M NaOH and cast directly onto a PDMS slab to allow drying in a close container with controlled humidity (~97% relative humidity) for up to 3 days. During the drying process, the assembly of a mesostructured silica framework was directed by DDM and surfactant composite. To control the net charge of the host material, the surfactant composite used here was either a pure zwitterionic DiC_7PC or a positively charged POPC/POPG (80/20, mol/mol) mixture. For pH-dependent spectral titration measurements on PR incorporated in mesostructured silica film, the formed PR-surfactant-silica film on PDMS was inserted into a polystyrene cuvette and rehydrated by a 10 mM HEPES buffer.

A.5 Native Gel Electrophoresis and Western Blotting

Blue native polyacrylamide gel electrophoresis (BN-PAGE) was performed using an XCell SureLock® Mini-Cell apparatus with NativePAGE™ Bis-Tris Protein Gels having a 3-12% acrylamide gradient (ThermoFisher). PR sample preparation and gel electrophoresis were performed according to the instructions provided by the manufacturer. The destained gel was then imaged with a ChemiDoc MP imaging system (Bio-Rad) using the default protocol for Coomassie blue dyes. Western blotting of BN-PAGE was performed using an XCell II^{TM} Blot Module by following the instructions for Western blotting of NativePAGE™ Gels provided by the manufacturer. The blotted PVDF membrane was rinsed with deionized water and blocked in a TBST buffer containing 5% (w/v) dry milk. The immunodetection was then done by applying HRP conjugated Anti-6X His tag® antibody (Abcam, Cambridge, MA, USA; ab1187) targeted to the six-histidine tag at the N-terminus of PR at 1:5000 in a TBST buffer containing 0.5% (w/v) milk. The membrane was then imaged with a ChemiDoc MP imaging system (Bio-Rad) using the default protocol for chemifluorescence dyes.

A.6 Preparing WALP Peptides in Liposomes

The C-WALP23 peptide was synthesized by RS Synthesis (Louisville, KY). We mainly followed the protocol published by Lueders et al. to prepare MTSL spin labeled WALP23 peptides [219]. The spin labeled WALP peptides dissolved in trifluoroethanol (TFE) was added to lipids with a 1:2200 peptide-to-lipid molar ratio. The lipid vesicles with WALP peptides were prepared by the extrusion method described in Appendix A.2 to form lipid vesicles with 200 nm in diameter. All the WALP-contained liposomes were equilibrated in the same HEPES buffer as used for PR samples with the desired NaCl concentration.

Appendix B Methods: Spectroscopy

B.1 Optical Absorption Spectroscopy

B.1.1 Spectral Titration Experiment

PR pH-dependent optical absorbance change was recorded by a UV-1800 spectrophotometer (Shimadzu, Kyoto, Japan), and the data analysis for getting the pKa_{D97} was done by the processing and fitting algorithm described in previous studies.[80] Prepared PR-containing liposomes were diluted by the HEPES buffer with the desired salt concentrations to achieve a total volume of 750 μL and an initial optical density between 0.3 and 0.5 at 520 nm. For all PR samples, optical absorbance spectra between 400 nm and 750 nm were recorded with a 0.5 nm and under interval at up to 30 different pH values between 4 and 10. Acidic PR, with a protonated D97 residue, has a maximum absorbance at ~535 nm, which shifts to ~518 nm upon deprotonation. An example of the pH-dependent optical absorbance change of PR reconstituted in liposomes is plotted in Figure B.1 A. The difference absorbance was calculated by subtracting each spectrum with the one measured under the most basic condition in that set of spectral titration experiment as shown in Figure B.1 B.

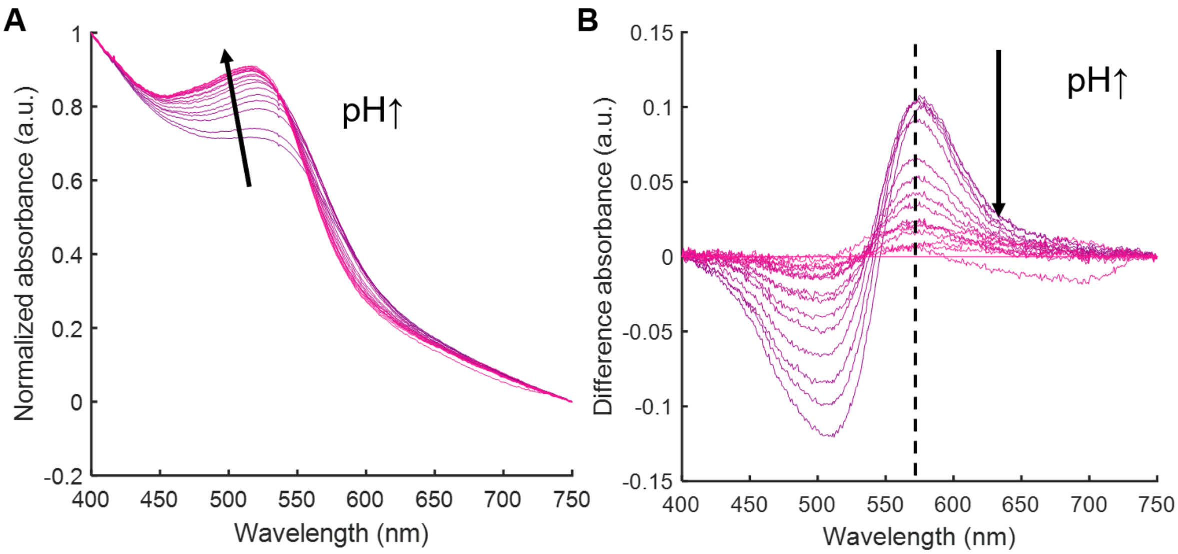


Figure B.1: Spectral Titration for Obtaining pKa$_{D97}$ of PR. A. pH-dependent optical absorbance of PR under a range between 400 and 750 nm after background subtraction. **B.** Difference absorbance of PR between 400 and 750 nm under various pH in the spectral titration experiment. The dotted line indicates the difference absorption at 570 nm.

The trend of difference absorbance at 570 nm $\Delta \text{Abs}_{570\ nm}$, indicated by the dotted line, was normalized and plotted against the measured bulk pH to get the pH-dependent titration curve as shown in Figure B.2. The pKa$_{D97}$ was then obtained from the Henderson–Hasselbalch fitting

$$\Delta \text{Abs}_{570\ nm} = \frac{1}{1+10^{n(pH-pKa_{D97})}}, \tag{B.1}$$

where n is the Hill coefficient that represents the cooperative behavior of the protonation/deprotonation of D97. The protonation behavior is anti-cooperative when this value is between 0 and 1.

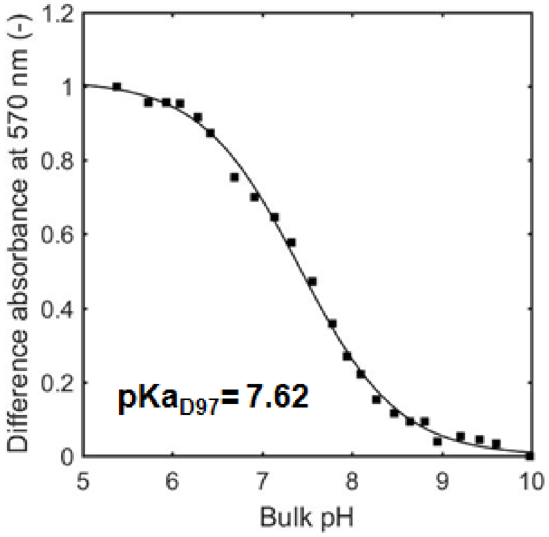


Figure B.2: Henderson-HasselBatch Fitting of a Spectral Titration Curve. The difference absorbance at 570 nm was normalized to the difference between the maximum and the minimum. The pH-dependent absorbance curve was fitted by the Henderson–Hasselbalch equation to get the apparent pKa$_{D97}$. The sample condition shown here is WT PR reconstituted in POPC/POPG (80/20, mol/mol) liposomes and a 10 mM HEPES buffer with 0 mM NaCl. All the pKa$_{D97}$ presented in this dissertation were processed using the same approach as the example here.

B.1.2 Transient Absorbance Measurement

Transient absorbance measurement was conducted using a SpectraPhysics Nd:YAG laser with a monochrometer. PR samples were excited by a 532 nm pulse laser with 10 ns duration, and the PR transient absorbances under various wavelengths were monitored over a time span of 10 μs to 0.5 s by an oscilloscope. Over 5096 data points were collected with a time resolution of 20 ns, 20 μs, and 2 nm, and these data were composed together to form a complete transient absorbance spectrum. The absorbance spectrum under each wavelength was averaged over approximately 200 logarithmically spaced time points. The biexponential fitting analyses on the transient spectra at 410 nm was done by MATLAB (Mathworks, MA) using home-built codes.

B.2 Magnetic Resonance Spectroscopy

B.2.1 X-band cw EPR

The X-band (0.35 T) cw EPR spectra from PR liposome samples were recorded on an EMXplus spectrometer (Bruker, Billerica, MA) with the Bruker TE011 high-sensitivity cylindrical microwave cavity (ER 4119HS-LC High Sensitivity Probehead). The prepared spin labeled PR samples were loaded into 0.6 mm ID × 0.84 mm OD quartz capillaries (VitroCom, NJ) at a volume of 3.5 μL. All the measurements were conducted with a microwave power of 20 mW, modulation amplitude of 1 G, and a total sweep width of 150 G. The spectra were post-processed with background subtraction using the Bruker WinEPR software.

B.2.2 ODNP

ODNP relaxometry measurements on PR reconstituted in liposomes were carried out in the 0.35 T Bruker EMXPlus magnet. A homebuilt NMR probe was added within a TE011 microwave cavity to hold the sample, and the probe was connected to a Bruker Avance 300 NMR spectrometer to detect ^{1}H NMR signal from the sample. PR liposome samples were loaded in capillary tubes with a PR concentration of 200-500 μM and a volume of 3.5 μL. The ^{1}H NMR signal from water molecules around the spin labeled site on PR was collected under different powers of microwave irradiation to get the ^{1}H NMR enhancement series. The spin-lattice relaxation time T_1 under different microwave irradiation powers was collected by inverse recovery experiments for temperature-correction purposes due to sample heating issues. The detail of the automated ODNP experiment setup is described by Franck et al. [114] The data were processed by a home-built Python script.

B.2.3 DEER

Four-pulse DEER on doubly-Gd-labeled samples were done on Q-band (34 GHz) Bruker E580 ELEXSYS pulse EPR spectrometer equipped with a TWT amplifier (300 W). The four-pulse DEER sequence used in this dissertation is described in Chapter 1. Gaussian observe pulses were set to 40 ns and 80 ns for the $\pi/2$ pulse and the π pulse, respectively. The power of the observe pulses was tuned down to ~60 % of the maximum to prevent the ring-down effect. An adiabatic chirp π pulse with a length of 100 ns was used for the electron-electron double resonance (ELDOR) pump pulse. The observer frequency was set to 33.8 GHz, and the pump frequency ranged from 300 MHz to 400 MHz higher than the observer frequency. The magnetic field was set to a value that the range of pump frequency was overlapped with the central transition of the Gd spin to maximize the modulation effect. Time-domain DEER traces were collected at a length between 2 to 3.5 μs to have a decent signal-to-noise ratio on the refocusing echo. The trace was signal averaging for over 300 scans. The time-domain DEER trace was background subtracted and fitted with Gaussian distribution by LongDistances software written by C. Altenbach.

B.2.4 240 GHz cw EPR

Gd-Gd distance measurements on PR were done on the 240 GHz cw EPR spectrometer built in Sherwin lab, UCSB. The spectrometer used a solid-state frequency-multiplied source

(Virginia Diodes, Inc.) to multiply a 15 GHz microwave source to an output of 240 GHz microwave irradiated to the samples. The power of the microwave under cw mode is ~50 mW. The spectrometer was operated in induction mode detection. The 240 GHz microwave irradiation was guided by a quasi-optical bridge, attenuated by a cross wire-grid polarizers, and guided to the sample that sits at the end of a waveguide (Thomas Keating, Inc.). The sample was sat on a home-built sample hold and loaded in a quartz capillary containing 0.5 μL of PR reconstituted in either DDM micelles or liposomes. The waveguide was surrounded by a continuous-flow cryostat (Janis Research Company, LCC), equilibrated at 30 K for cryo-trapped sample measurements and 290 K for room temperature measurements. The cryostat was placed in a 12.5 T superconducting magnet (Oxford Instrument plc) and a smaller sweeping coil was used for the field sweep in a range of ± 0.06 T of the static field to achieve cw measurements.

B.2.5 DNP-MAS ssNMR

All DNP-enhanced MAS ssNMR measurements were done on a 400 MHz (9.4 T) ASCEND DNP-NMR spectrometer and a 263 GHz gyrotron microwave source. A 3.2 mm MAS DNP-NMR triple resonance broadband X/Y/H probe was used for all measurements. All spectra were recorded at 95 K and the rotor was spin at 8 kHz MAS. The prepared PR proteoliposomes for DNP-MAS ssNMR measurements were first collected by centrifugation and reconstituted with a DNP juice (d_8-glycerol/D_2O/H_2O, 60/30/10, v/v/v) supplemented with 20 mM AMUPOL polarizing agent,[220] 10 mM HEPES, and in some case 150 mM NaCl. The resuspended proteoliposomes were well mixed and packed into a 3.2 mm sapphire rotor. The rotor packed with the sample was then evenly distributed by spinning at room temperature using a Bruker rotor testing device, and subsequently loaded into the probe for measurements.

^{15}N-CP-MAS spectra were recorded with an 800 μs contact time. ^{13}C-^{13}C DARR spectra were recorded with a 50 ms mixing time, 16 scans in the direct dimension, and 256 increments in the indirect dimension with a dwell time of 3.33 μs.

Appendix C

Gouy-Chapman Theory on Estimating Intrinsic pKa

Proton concentration on a charged liposome surface can differ from its concentration in the bulk solution. This concentration difference highly depends on the electrostatic potential on the charged liposome surface. Briefly, protons can be absorbed on a negatively-charged surface based on the double-layer theory and results in a higher surface proton concentration in comparison to the bulk concentration. On the contrary, a positively-charged liposome surface can repel protons and results in a lower surface proton concentration in comparison to the bulk concentration. The relationship between the proton concentration difference and the surface potential ψ_0 can be described by the Nernst equation

$$pH_{surface} = pH_{bulk} + \frac{zF\psi_0}{2.303RT}, \tag{C.1}$$

where z is the ion charge, which is +1 for proton, F is the Faraday constant, R is the ideal gas constant and T is the absolute temperature. The surface potential ψ_0 is related to the surface charge density σ. With an assumption that the liposome surface is static and flat, we can use the Gouy-Chapman model to correlate the surface potential ψ_0 with the membrane surface charge density σ

$$\sigma = \sqrt{8000kT\varepsilon_0\varepsilon C_{el}N_A} \cdot \sinh\left(\frac{zF\psi_0}{2RT}\right), \tag{C.2}$$

where k is the Boltzmann constant, ε_0 is the vacuum permeability constant, N_A is the Avogadro constant, ε is the dielectric constant, for water $\varepsilon = 78$, and C_{el} is the electrolyte concentration, which is the concentration of applied salt with monovalent or divalent cations in the buffer. To resolve the surface potential ψ_0, we have to calculate the surface charge density σ. The surface charge density of a charged species i can be expressed as

$$\sigma_i = \frac{ze\alpha}{A_{L,i}}, \tag{C.3}$$

where e is the elementary charge, $A_{L,i}$ is the surface area per ionizable group of lipid (0.597 nm^2 for PC [221], 0.61 nm^2 for DOTAP [222], 0.65 nm^2 for POPG [184]), and α is the degree of dissociation of an ionizable group that determines the percentage of it in its charged form. For example, in the case of POPC, it contains a phosphorous group that can distribute between a negatively charged form and a neutral form while it is associated with protons or other cations (e.g. Na). The degree of dissociation of this phosphorous group can be expressed as

$$\alpha_{PC,PO_4^-} = \frac{1}{1+\left(\frac{[H_3O^+]}{K_H}+\frac{[Na^+]}{K_{Na}}\right)exp\left(-\frac{F\psi_0}{RT}\right)}, \tag{C.4}$$

where $[H_3O^+]$ and $[Na^+]$ are the proton and sodium concentration in the buffer, respectively. K_H and K_{Na} are the dissociation constant for the binding of proton and sodium to the phosphorous group, respectively. All the parameters for calculating the surface charge density of different ionizable headgroups are obtained from the literature.[194] For the phosphorous group on the PC headgroup, its K_{Na} is infinity as it is hard for sodium cations to access. For the phosphorous group on the PG headgroup, we can use the same Equation $\alpha_{PC,PO_4{}^-} = \frac{1}{1+\left(\frac{[H_3O^+]}{K_H}+\frac{[Na^+]}{K_{Na}}\right)exp\left(-\frac{F\psi_0}{RT}\right)}$, (C.4) to get $\alpha_{PG,PO_4{}^-}$. In this case, its K_{Na} is 1/0.65 M, which means the association between sodium ions and its phosphorous group should be considered. For the choline group on the PC headgroup that can be either negative charge or neutral if it is associated with a chloride anion, the degree of dissociation can be expressed as

$$\alpha_{PC,\ NH_3{}^+} = \frac{1}{1+\frac{[Cl^-]}{K_{Cl}}exp\left(-\frac{F\psi_0}{RT}\right)}, \tag{C.5}$$

where K_{Cl} is the dissociation constant of a chloride anion on a choline group (The K_{Cl} for the PC headgroup is 1/0.9 M). The same equation can also be applied to the choline group on the DOTAP headgroup, with an assumption that its dissociation constant is the same as the one for the PC headgroup. The surface charge density σ of all the ionizable groups can then be combined based on the ratio between different lipid species using the following equation

$$\sigma = \sum_i x_i\, \sigma_i \tag{C.6}$$

By solving the nonlinear equations between Equation C.2 and Equation C.6, the surface potential ψ_0 on the charged liposome surface can be obtained. The surface potential ψ_0 under different bulk pH then can be used to calculate the surface pH. In this dissertation, negatively-charged POPC/POPG (80/20, mol/mol) liposomes and positively-charged POPC/DOTAP (80/20, mol/mol) liposomes were used for reconstituting PR. When these charged liposomes are prepared in a buffer with 0 mM NaCl, the surface pH would significantly deviate from the bulk pH as shown in Figure C.1 A. Increasing C_{el} can diminish the difference between the surface pH and the bulk pH. In the case of 150 mM NaCl presence in the buffer, the surface pH is close to the bulk pH as shown in Figure C.1 B.

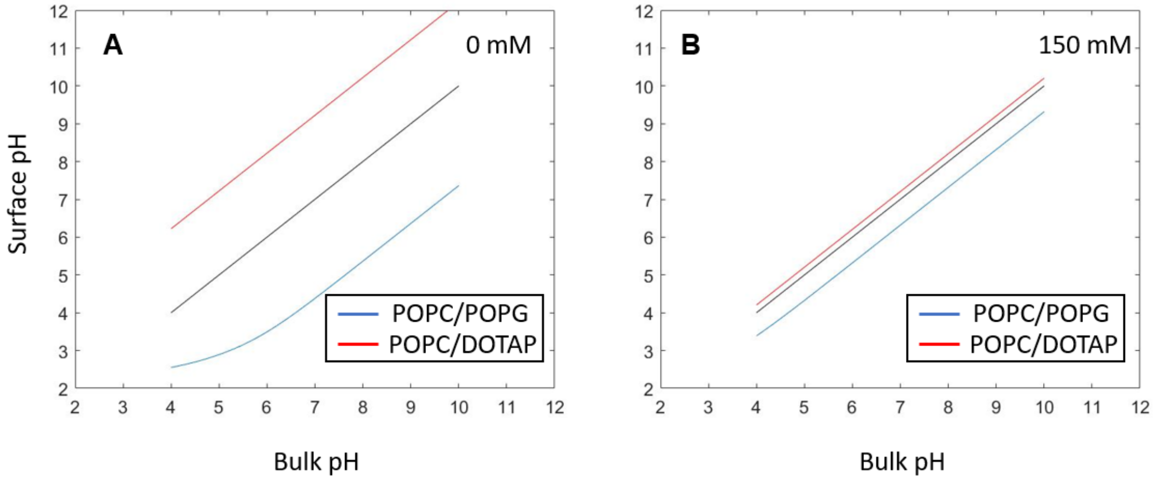


Figure C.1: Surface pH Predicted by Gouy-Chapman Model. Surface pH of negatively-charged POPC/POPG (80/20, mol/mol) liposomes (blue) and positively-charged POPC/DOTAP (80/20, mol/mol) liposomes (red) equilibrated in a buffer with **A.** 0 mM of NaCl and **B.** 150 mM NaCl. The black line is a reference when the surface pH is equal to bulk pH.

The bulk pH measured in the spectral titration experiment was converted into surface pH by using the above equations with the C_{el} set to the NaCl concentration used in each condition. The converted surface pH series for each spectral titration experiment then was fitted by the Henderson-Hasselbalch equation to get the intrinsic pKa_{D97}. This value represents the intrinsic acid dissociation constant of the D97 residue, while the apparent pKa does not take the surface proton concentration difference into account. The intrinsic pKa_{D97} still clearly changed while

PR was reconstituted in liposomes with different net charges and in buffers with different NaCl concentrations (Table C.1)

Table C.1: Gouy-Chapman Model Predicted Surface Potential and Intrinsic pKa. Intrinsic pKa_{D97} and corresponding apparent pKa_{D97} of PR reconstituted in liposomes with different net charges and in buffers with different NaCl concentrations. The surface potential of liposomes is predicted by the Gouy-Chapman model.

POPC/POPG (80/20)				POPC/DOTAP (80/20)			
NaCl (mM)	$pKa_{D97, intrinsic}$	$pKa_{D97, apparent}$	ψ_0 (mV)	NaCl (mM)	$pKa_{D97, intrinsic}$	$pKa_{D97, apparent}$	ψ_0 (mV)
0	5.11	7.62	-142	0	7.64	5.63	116
50	5.53	6.79	-72	30	6.74	5.90	48
80	5.54	6.62	-62	75	6.65	6.14	29
120	5.60	6.51	-53	120	6.52	6.14	22
150	5.74	6.43	-40	150	6.42	6.14	16

Bibliography

1. Luecke, H.; Richter, H.-T.; Lanyi, J. K., Proton transfer pathways in bacteriorhodopsin at 2.3 angstrom resolution. *Science* **1998,** *280* (5371), 1934-1937.
2. Yoshikawa, S.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yamashita, E.; Inoue, N.; Yao, M.; Fei, M. J.; Libeu, C. P.; Mizushima, T., Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. *Science* **1998,** *280* (5370), 1723-1729.
3. Churg, A.; Warshel, A., Control of the redox potential of cytochrome and microscopic dielectric effects in proteins. *Biochemistry* **1986,** *25* (7), 1675-1681.
4. Liman, E. R., Use it or lose it: molecular evolution of sensory signaling in primates. *Pflugers Archiv : European journal of physiology* **2006,** *453* (2), 125-131.
5. Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W., Modeling electrostatic effects in proteins. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2006,** *1764* (11), 1647-1676.
6. Brown, M. F., Modulation of rhodopsin function by properties of the membrane bilayer. *Chemistry and physics of lipids* **1994,** *73* (1-2), 159-180.
7. Hsia, C.-Y.; Richards, M. J.; Daniel, S., A review of traditional and emerging methods to characterize lipid–protein interactions in biological membranes. *Analytical Methods* **2015,** *7* (17), 7076-7094.
8. Lee, A., Lipid–protein interactions in biological membranes: a structural perspective. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2003,** *1612* (1), 1-40.
9. Maxfield, F. R.; Tabas, I., Role of cholesterol and lipid organization in disease. *Nature* **2005,** *438* (7068), 612-621.
10. Shaw, A. S., Lipid rafts: now you see them, now you don't. *Nature immunology* **2006,** *7* (11), 1139-1142.
11. de Planque, M. R.; Killian*, J. A., Protein–lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring (Review). *Molecular membrane biology* **2003,** *20* (4), 271-284.
12. Raschke, T. M., Water structure and interactions with protein surfaces. *Current opinion in structural biology* **2006,** *16* (2), 152-159.
13. Zhou, H.-X.; Cross, T. A., Influences of membrane mimetic environments on membrane protein structures. *Annual review of biophysics* **2013,** *42*, 361-392.
14. Sriram, K.; Insel, P. A., G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? *Molecular pharmacology* **2018,** *93* (4), 251-258.
15. Ferre, S.; Casado, V.; Devi, L. A.; Filizola, M.; Jockers, R.; Lohse, M. J.; Milligan, G.; Pin, J. P.; Guitart, X., G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives. *Pharmacological reviews* **2014,** *66* (2), 413-34.
16. Ferre, S.; Euler, G. v.; Johansson, B.; Fredhol, B. B.; Fuxe, K., Stimulation of high-affinity adenosine A2 receptors decreases the affinity of dopamine D2 receptors in rat striatal membranes. *Proc Natl Acad Sci U S A* **1991,** *88*, 7238-7241.
17. Fuxe, K.; Ferre, S.; Canals, M.; Torvinen, M.; Terasmaa, A.; Marcellino, D.; Goldberg, S. R.; Staines, W.; Jacobson, K. X.; Lluis, C.; Woods, A. S.; Agnati, L. F.; Franco,

R., Adenosine A2a and Dopamine D2 Heteromeric Receptor Complexes and Their Function. *Journal of Molecular Neuroscience* **2005,** *26*, 209-219.

18. Hasbi, A.; O'Dowd, B. F.; George, S. R., Dopamine D1-D2 receptor heteromer signaling pathway in the brain: emerging physiological relevance. *Molecular brain* **2011,** *4*, 26.

19. Perreault, M. L.; O'Dowd, B. F.; George, S. R., Dopamine receptor homooligomers and heterooligomers in schizophrenia. *CNS neuroscience & therapeutics* **2011,** *17* (1), 52-7.

20. So, C. H.; Varghese, G.; Curley, K. J.; Kong, M. M.; Alijaniaram, M.; Ji, X.; Nguyen, T.; O'Dowd B, F.; George, S. R., D1 and D2 dopamine receptors form heterooligomers and cointernalize after selective activation of either receptor. *Molecular pharmacology* **2005,** *68* (3), 568-78.

21. So, C. H.; Verma, V.; Alijaniaram, M.; Cheng, R.; Rashid, A. J.; O'Dowd, B. F.; George, S. R., Calcium signaling by dopamine D5 receptor and D5-D2 receptor hetero-oligomers occurs by a mechanism distinct from that for dopamine D1-D2 receptor hetero-oligomers. *Molecular pharmacology* **2009,** *75* (4), 843-54.

22. Lee, A. G., How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2004,** *1666* (1-2), 62-87.

23. Reimhult, E.; Kumar, K., Membrane biosensor platforms using nano-and microporous supports. *Trends in biotechnology* **2008,** *26* (2), 82-89.

24. Shieh, F.-K.; Wang, S.-C.; Yen, C.-I.; Wu, C.-C.; Dutta, S.; Chou, L.-Y.; Morabito, J. V.; Hu, P.; Hsu, M.-H.; Wu, K. C.-W., Imparting functionality to biocatalysts via embedding enzymes into nanoporous materials by a de novo approach: size-selective sheltering of catalase in metal–organic framework microcrystals. *Journal of the American Chemical Society* **2015,** *137* (13), 4276-4279.

25. Ispas, C.; Sokolov, I.; Andreescu, S., Enzyme-functionalized mesoporous silica for bioanalytical applications. *Analytical and bioanalytical chemistry* **2009,** *393* (2), 543-554.

26. Oda, I.; Hirata, K.; Watanabe, S.; Shibata, Y.; Kajino, T.; Fukushima, Y.; Iwai, S.; Itoh, S., Function of membrane protein in silica nanopores: incorporation of photosynthetic light-harvesting protein LH2 into FSM. *The Journal of Physical Chemistry B* **2006,** *110* (3), 1114-1120.

27. González-Maeso, J., GPCR oligomers in pharmacology and signaling. *Molecular brain* **2011,** *4* (1), 1.

28. Kobilka, B. K., G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2007,** *1768* (4), 794-807.

29. Kimata, N.; Reeves, P. J.; Smith, S. O., Uncovering the triggers for GPCR activation using solid-state NMR spectroscopy. *Journal of Magnetic Resonance* **2015,** *253*, 111-118.

30. Vafabakhsh, R.; Levitz, J.; Isacoff, E. Y., Conformational dynamics of a class C G-protein-coupled receptor. *Nature* **2015,** *524* (7566), 497-501.

31. Cordomí, A.; Navarro, G.; Aymerich, M. S.; Franco, R., Structures for G-protein-coupled receptor tetramers in complex with G proteins. *Trends in biochemical sciences* **2015,** *40* (10), 548-551.

32. Gahbauer, S.; Böckmann, R. A., Membrane-mediated oligomerization of G protein coupled receptors and its implications for GPCR function. *Frontiers in physiology* **2016,** *7*, 494.

33. Schonenbach, N. S.; Hussain, S.; O'Malley, M. A., Structure and function of G protein-coupled receptor oligomers: implications for drug discovery. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **2015,** *7* (3), 408-427.
34. Botelho, A. V.; Huber, T.; Sakmar, T. P.; Brown, M. F., Curvature and hydrophobic forces drive oligomerization and modulate activity of rhodopsin in membranes. *Biophysical journal* **2006,** *91* (12), 4464-4477.
35. Singer, S. J.; Nicolson, G. L., The fluid mosaic model of the structure of cell membranes. *Science* **1972,** *175* (4023), 720-731.
36. East, J.; Melville, D.; Lee, A., Exchange rates and numbers of annular lipids for the calcium and magnesium ion dependent adenosine triphosphatase. *Biochemistry* **1985,** *24* (11), 2615-2623.
37. Kusumi, A.; Suzuki, K., Toward understanding the dynamics of membrane-raft-based molecular interactions. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2005,** *1746* (3), 234-251.
38. Foster, L. J.; Chan, Q. W., Lipid raft proteomics: more than just detergent-resistant membranes. In *Subcellular Proteomics*, Springer: 2007; pp 35-47.
39. Hong, M.; DeGrado, W. F., Structural basis for proton conduction and inhibition by the influenza M2 protein. *Protein Science* **2012,** *21* (11), 1620-1633.
40. Schnell, J. R.; Chou, J. J., Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **2008,** *451* (7178), 591-595.
41. Rossman, J. S.; Jing, X.; Leser, G. P.; Balannik, V.; Pinto, L. H.; Lamb, R. A., Influenza virus m2 ion channel protein is necessary for filamentous virion formation. *Journal of virology* **2010,** *84* (10), 5078-5088.
42. Brown, M. F., Curvature forces in membrane lipid–protein interactions. *Biochemistry* **2012,** *51* (49), 9782-9795.
43. Bordag, N.; Keller, S., α-Helical transmembrane peptides: a “divide and conquer” approach to membrane proteins. *Chemistry and physics of lipids* **2010,** *163* (1), 1-26.
44. Dörr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schäfer, M.; van Walree, C. A.; Killian, J. A., The styrene–maleic acid copolymer: a versatile tool in membrane research. *European Biophysics Journal* **2016,** *45* (1), 3-21.
45. Béja, O.; Aravind, L.; Koonin, E. V.; Suzuki, M. T.; Hadd, A.; Nguyen, L. P.; Jovanovich, S. B.; Gates, C. M.; Feldman, R. A.; Spudich, J. L., Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **2000,** *289* (5486), 1902-1906.
46. Beja, O.; Spudich, E. N.; Spudich, J. L.; Leclerc, M.; DeLong, E. F., Proteorhodopsin phototrophy in the ocean. *Nature* **2001,** *411* (6839), 786-789.
47. DeLong, E. F.; Beja, O., The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS biology* **2010,** *8* (4), e1000359.
48. Olson, D. K.; Yoshizawa, S.; Boeuf, D.; Iwasaki, W.; DeLong, E. F., Proteorhodopsin variability and distribution in the North Pacific Subtropical Gyre. *The ISME journal* **2018,** *12* (4), 1047.
49. Hempelmann, F.; Hölper, S.; Verhoefen, M.-K.; Woerner, A. C.; Köhler, T.; Fiedler, S.-A.; Pfleger, N.; Wachtveitl, J.; Glaubitz, C., His75− Asp97 cluster in green proteorhodopsin. *Journal of the American Chemical Society* **2011,** *133* (12), 4645-4654.
50. Friedrich, T.; Geibel, S.; Kalmbach, R.; Chizhov, I.; Ataka, K.; Heberle, J.; Engelhard, M.; Bamberg, E., Proteorhodopsin is a light-driven proton pump with variable vectoriality. *Journal of molecular biology* **2002,** *321* (5), 821-838.

51. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L., How many drug targets are there? *Nature reviews Drug discovery* **2006,** *5* (12), 993-996.

52. Reckel, S.; Gottstein, D.; Stehle, J.; Lohr, F.; Verhoefen, M. K.; Takeda, M.; Silvers, R.; Kainosho, M.; Glaubitz, C.; Wachtveitl, J.; Bernhard, F.; Schwalbe, H.; Guntert, P.; Dotsch, V., Solution NMR structure of proteorhodopsin. *Angewandte Chemie (International ed. in English)* **2011,** *50* (50), 11942-6.

53. Dioumaev, A. K.; Brown, L. S.; Shih, J.; Spudich, E. N.; Spudich, J. L.; Lanyi, J. K., Proton Transfers in the Photochemical Reaction Cycle of Proteorhodopsin. *Biochemistry* **2002,** *41* (17), 5348-5358.

54. Váró, G.; Brown, L. S.; Lakatos, M.; Lanyi, J. K., Characterization of the Photochemical Reaction Cycle of Proteorhodopsin. *Biophysical Journal* **2003,** *84* (2), 1202-1207.

55. Lakatos, M.; Lanyi, J. K.; Szakács, J.; Váró, G., The photochemical reaction cycle of proteorhodopsin at low pH. *Biophysical Journal* **2003,** *84* (5), 3252-3256.

56. Ni, Q. Z.; Can, T. V.; Daviso, E.; Belenky, M.; Griffin, R. G.; Herzfeld, J., Primary transfer step in the light-driven ion pump bacteriorhodopsin: An irreversible U-turn revealed by dynamic nuclear polarization-enhanced magic angle spinning NMR. *Journal of the American Chemical Society* **2018,** *140* (11), 4085-4091.

57. Rangarajan, R.; Galan, J. F.; Whited, G.; Birge, R. R., Mechanism of spectral tuning in green-absorbing proteorhodopsin. *Biochemistry* **2007,** *46* (44), 12679-12686.

58. Bergo, V. B.; Sineshchekov, O. A.; Kralj, J. M.; Partha, R.; Spudich, E. N.; Rothschild, K. J.; Spudich, J. L., His-75 in proteorhodopsin, a novel component in light-driven proton translocation by primary pumps. *Journal of Biological Chemistry* **2009,** *284* (5), 2836-2843.

59. Pfleger, N.; Wörner, A. C.; Yang, J.; Shastri, S.; Hellmich, U. A.; Aslimovska, L.; Maier, M. S.; Glaubitz, C., Solid-state NMR and functional studies on proteorhodopsin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2009,** *1787* (6), 697-705.

60. Shi, L.; Ahmed, M. A.; Zhang, W.; Whited, G.; Brown, L. S.; Ladizhansky, V., Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump—structural insights. *Journal of molecular biology* **2009,** *386* (4), 1078-1093.

61. Yoshitsugu, M.; Yamada, J.; Kandori, H., Color-Changing Mutation in the E− F Loop of Proteorhodopsin. *Biochemistry* **2009,** *48* (20), 4324-4330.

62. Mehler, M.; Scholz, F.; Ullrich, S. J.; Mao, J.; Braun, M.; Brown, L. J.; Brown, R. C.; Fiedler, S. A.; Becker-Baldus, J.; Wachtveitl, J., The EF loop in green proteorhodopsin affects conformation and photocycle dynamics. *Biophysical journal* **2013,** *105* (2), 385-397.

63. Yoshitsugu, M.; Shibata, M.; Ikeda, D.; Furutani, Y.; Kandori, H., Color change of proteorhodopsin by a single amino acid replacement at a distant cytoplasmic loop. *Angewandte Chemie International Edition* **2008,** *47* (21), 3923-3926.

64. Yamada, K.; Kawanabe, A.; Kandori, H., Importance of alanine at position 178 in proteorhodopsin for absorption of prevalent ambient light in the marine environment. *Biochemistry* **2010,** *49* (11), 2416-2423.

65. Man, D.; Wang, W.; Sabehi, G.; Aravind, L.; Post, A. F.; Massana, R.; Spudich, E. N.; Spudich, J. L.; Béja, O., Diversification and spectral tuning in marine proteorhodopsins. *The EMBO journal* **2003,** *22* (8), 1725-1731.

66. Kralj, J. M.; Spudich, E. N.; Spudich, J. L.; Rothschild, K. J., Raman spectroscopy reveals direct chromophore interactions in the Leu/Gln105 spectral tuning switch of proteorhodopsins. *The Journal of Physical Chemistry B* **2008,** *112* (37), 11770-11776.

67. Shastri, S.; Vonck, J.; Pfleger, N.; Haase, W.; Kuehlbrandt, W.; Glaubitz, C., Proteorhodopsin: characterisation of 2D crystals by electron microscopy and solid state NMR. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2007,** *1768* (12), 3012-3019.

68. Klyszejko, A. L.; Shastri, S.; Mari, S. A.; Grubmüller, H.; Muller, D. J.; Glaubitz, C., Folding and assembly of proteorhodopsin. *Journal of molecular biology* **2008,** *376* (1), 35-41.

69. Ran, T.; Ozorowski, G.; Gao, Y.; Sineshchekov, O. A.; Wang, W.; Spudich, J. L.; Luecke, H., Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta crystallographica. Section D, Biological crystallography* **2013,** *69* (Pt 10), 1965-80.

70. Edwards, D. T.; Ma, Z.; Meade, T. J.; Goldfarb, D.; Han, S.; Sherwin, M. S., Extending the distance range accessed with continuous wave EPR with Gd 3+ spin probes at high magnetic fields. *Physical Chemistry Chemical Physics* **2013,** *15* (27), 11313-11326.

71. Stone, K. M.; Voska, J.; Kinnebrew, M.; Pavlova, A.; Junk, M. J.; Han, S., Structural insight into proteorhodopsin oligomers. *Biophysical journal* **2013,** *104* (2), 472-481.

72. Maciejko, J.; Mehler, M.; Kaur, J.; Lieblein, T.; Morgner, N.; Ouari, O.; Tordo, P.; Becker-Baldus, J.; Glaubitz, C., Visualizing Specific Cross-Protomer Interactions in the Homo-Oligomeric Membrane Protein Proteorhodopsin by Dynamic-Nuclear-Polarization-Enhanced Solid-State NMR. *Journal of the American Chemical Society* **2015,** *137* (28), 9032-9043.

73. Maciejko, J.; Kaur, J.; Becker-Baldus, J.; Glaubitz, C., Photocycle-dependent conformational changes in the proteorhodopsin cross-protomer Asp–His–Trp triad revealed by DNP-enhanced MAS-NMR. *Proceedings of the National Academy of Sciences* **2019,** *116* (17), 8342-8349.

74. Hussain, S.; Kinnebrew, M.; Schonenbach, N. S.; Aye, E.; Han, S., Functional consequences of the oligomeric assembly of proteorhodopsin. *Journal of molecular biology* **2015,** *427* (6), 1278-1290.

75. Sharaabi, Y.; Brumfeld, V.; Sheves, M., Binding of Anions to Proteorhodopsin Affects the Asp97 p K a. *Biochemistry* **2010,** *49* (21), 4457-4465.

76. Wang, W.-W.; Sineshchekov, O. A.; Spudich, E. N.; Spudich, J. L., Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *Journal of Biological Chemistry* **2003,** *278* (36), 33985-33991.

77. Ikeda, D.; Furutani, Y.; Kandori, H., FTIR study of the retinal Schiff base and internal water molecules of proteorhodopsin. *Biochemistry* **2007,** *46* (18), 5365-5373.

78. Sharaabi, Y.; Brumfeld, V.; Sheves, M., Binding of Anions to Proteorhodopsin Affects the Asp97 pKa. *Biochemistry* **2010,** *49* (21), 4457-4465.

79. Ranaghan, M. J.; Schwall, C. T.; Alder, N. N.; Birge, R. R., Green proteorhodopsin reconstituted into nanoscale phospholipid bilayers (nanodiscs) as photoactive monomers. *Journal of the American Chemical Society* **2011,** *133* (45), 18318-18327.

80. Idso, M. N.; Baxter, N. R.; Narayanan, S.; Chang, E.; Fisher, J. M.; Chmelka, B. F.; Han, S., Proteorhodopsin Function is Primarily Mediated by Oligomerization in Different Micellar Surfactant Solutions. *The Journal of Physical Chemistry B* **2019**.

81. Balashov, S. P.; Imasheva, E. S.; Govindjee, R.; Ebrey, T. G., Titration of aspartate-85 in bacteriorhodopsin: what it says about chromophore isomerization and proton release. *Biophysical journal* **1996,** *70* (1), 473-481.

82. Kelemen, B. R.; Du, M.; Jensen, R. B., Proteorhodopsin in living color: diversity of spectral properties within living bacterial cells. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2003,** *1618* (1), 25-32.
83. Partha, R.; Krebs, R.; Caterino, T. L.; Braiman, M. S., Weakened coupling of conserved arginine to the proteorhodopsin chromophore and its counterion implies structural differences from bacteriorhodopsin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2005,** *1708* (1), 6-12.
84. Váró, G.; Lanyi, J., Distortions in the photocycle of bacteriorhodopsin at moderate dehydration. *Biophysical journal* **1991,** *59* (2), 313-322.
85. Dioumaev, A. K.; Wang, J. M.; Bálint, Z.; Váró, G.; Lanyi, J. K., Proton transport by proteorhodopsin requires that the retinal Schiff base counterion Asp-97 be anionic. *Biochemistry* **2003,** *42* (21), 6582-6587.
86. Walter, J. M.; Greenfield, D.; Liphardt, J., Potential of light-harvesting proton pumps for bioenergy applications. *Current opinion in biotechnology* **2010,** *21* (3), 265-270.
87. Atsumi, S.; Cann, A. F.; Connor, M. R.; Shen, C. R.; Smith, K. M.; Brynildsen, M. P.; Chou, K. J.; Hanai, T.; Liao, J. C., Metabolic engineering of Escherichia coli for 1-butanol production. *Metabolic engineering* **2008,** *10* (6), 305-311.
88. Stephanopoulos, G., Challenges in engineering microbes for biofuels production. *Science* **2007,** *315* (5813), 801-804.
89. Lang-Hinrichs, C.; Queck, I.; Büldt, G.; Stahl, U.; Hildebrandt, V., The archaebacterial membrane protein bacterio-opsin is expressed and N-terminally processed in the yeast Saccharomyces cerevisiae. *Molecular and General Genetics MGG* **1994,** *244* (2), 183-188.
90. Weusthuis, R. A.; Lamot, I.; van der Oost, J.; Sanders, J. P., Microbial production of bulk chemicals: development of anaerobic processes. *Trends in biotechnology* **2011,** *29* (4), 153-158.
91. Claassens, N. J.; Volpers, M.; dos Santos, V. A. M.; van der Oost, J.; de Vos, W. M., Potential of proton-pumping rhodopsins: engineering photosystems into microorganisms. *Trends in biotechnology* **2013,** *31* (11), 633-642.
92. Kim, J. Y.; Jo, B.; Jo, Y.; Cha, H., Improved production of biohydrogen in light-powered Escherichia coli by co-expression of proteorhodopsin and heterologous hydrogenase. *Microbial cell factories* **2012,** *11* (1), 2.
93. Janke, C.; Scholz, F.; Becker-Baldus, J.; Glaubitz, C.; Wood, P. G.; Bamberg, E.; Wachtveitl, J.; Bamann, C., Photocycle and vectorial proton transfer in a rhodopsin from the eukaryote Oxyrrhis marina. *Biochemistry* **2013,** *52* (16), 2750-2763.
94. Hildebrandt, V.; Fendler, K.; Heberle, J.; Hoffmann, A.; Bamberg, E.; Büldt, G., Bacteriorhodopsin expressed in Schizosaccharomyces pombe pumps protons through the plasma membrane. *Proceedings of the National Academy of Sciences* **1993,** *90* (8), 3578-3582.
95. Hussain, S.; Franck, J. M.; Han, S., Transmembrane Protein Activation Refined by Site-Specific Hydration Dynamics. *Angewandte Chemie International Edition* **2013,** *52* (7), 1953-1958.
96. Hirschi, S.; Kalbermatter, D.; Ucurum, Z.; Fotiadis, D., Cryo-electron microscopic and X-ray crystallographic analysis of the light-driven proton pump proteorhodopsin reveals a pentameric assembly. *Journal of Structural Biology: X* **2020**, 100024.
97. Gaitzsch, J.; Hirschi, S.; Freimann, S.; Fotiadis, D.; Meier, W., Directed Insertion of Light-Activated Proteorhodopsin into Asymmetric Polymersomes from an ABC Block Copolymer. *Nano letters* **2019,** *19* (4), 2503-2508.

98. Lindholm, L.; Ariöz, C.; Jawurek, M.; Liebau, J.; Mäler, L.; Wieslander, Å.; von Ballmoos, C.; Barth, A., Effect of lipid bilayer properties on the photocycle of green proteorhodopsin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2015,** *1847* (8), 698-708.

99. Han, C.-T.; Song, J.; Chan, T.; Pruett, C.; Han, S., Electrostatic environment of Proteorhodopsin affects the pKa of its buried primary proton acceptor. *Biophysical Journal* **2020**.

100. Roos, C.; Zocher, M.; Müller, D.; Münch, D.; Schneider, T.; Sahl, H.-G.; Scholz, F.; Wachtveitl, J.; Ma, Y.; Proverbio, D., Characterization of co-translationally formed nanodisc complexes with small multidrug transporters, proteorhodopsin and with the E. coli MraY translocase. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2012,** *1818* (12), 3098-3106.

101. Mörs, K.; Roos, C.; Scholz, F.; Wachtveitl, J.; Dötsch, V.; Bernhard, F.; Glaubitz, C., Modified lipid and protein dynamics in nanodiscs. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2013,** *1828* (4), 1222-1229.

102. Jahnke, J. P.; Idso, M. N.; Hussain, S.; Junk, M. J.; Fisher, J. M.; Phan, D. D.; Han, S.; Chmelka, B. F., Functionally active membrane proteins incorporated in mesostructured silica films. *Journal of the American Chemical Society* **2018,** *140* (11), 3892-3906.

103. Govorunova, E. G.; Sineshchekov, O. A.; Li, H.; Spudich, J. L., Microbial rhodopsins: diversity, mechanisms, and optogenetic applications. *Annual review of biochemistry* **2017,** *86*, 845-872.

104. Engqvist, M. K.; McIsaac, R. S.; Dollinger, P.; Flytzanis, N. C.; Abrams, M.; Schor, S.; Arnold, F. H., Directed evolution of Gloeobacter violaceus rhodopsin spectral properties. *Journal of molecular biology* **2015,** *427* (1), 205-220.

105. McIsaac, R. S.; Engqvist, M. K.; Wannier, T.; Rosenthal, A. Z.; Herwig, L.; Flytzanis, N. C.; Imasheva, E. S.; Lanyi, J. K.; Balashov, S. P.; Gradinaru, V., Directed evolution of a far-red fluorescent rhodopsin. *Proceedings of the National Academy of Sciences* **2014,** *111* (36), 13034-13039.

106. Richards, R.; Mondal, S.; Weinstein, H.; Dempski, R. E., Channelrhodopsin-2 Function is Modulated by Residual Hydrophobic Mismatch with the Surrounding Lipid Environment. *Applied Sciences* **2019,** *9* (13), 2674.

107. Govorunova, E. G.; Sineshchekov, O. A.; Janz, R.; Liu, X.; Spudich, J. L., Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science* **2015,** *349* (6248), 647-650.

108. Clayton, J. A.; Qi, M.; Godt, A.; Goldfarb, D.; Han, S.; Sherwin, M. S., Gd 3+–Gd 3+ distances exceeding 3 nm determined by very high frequency continuous wave electron paramagnetic resonance. *Physical Chemistry Chemical Physics* **2017,** *19* (7), 5127-5136.

109. Abdelkader, E. H.; Lee, M. D.; Feintuch, A.; Cohen, M. R.; Swarbrick, J. D.; Otting, G.; Graham, B.; Goldfarb, D., A new Gd3+ spin label for Gd3+–Gd3+ distance measurements in proteins produces narrow distance distributions. *The journal of physical chemistry letters* **2015,** *6* (24), 5016-5021.

110. Matalon, E.; Huber, T.; Hagelueken, G.; Graham, B.; Frydman, V.; Feintuch, A.; Otting, G.; Goldfarb, D., Gadolinium (III) spin labels for high-sensitivity distance measurements in transmembrane helices. *Angewandte Chemie International Edition* **2013,** *52* (45), 11831-11834.

111. Martorana, A.; Bellapadrona, G.; Feintuch, A.; Di Gregorio, E.; Aime, S.; Goldfarb, D., Probing protein conformation in cells by EPR distance measurements using Gd3+ spin labeling. *Journal of the American Chemical Society* **2014,** *136* (38), 13458-13465.

112. Clayton, J. A. High-field CW EPR with Gd (III) spin labels for structure studies of membrane proteins. UC Santa Barbara, 2017.

113. Armstrong, B. D.; Han, S., Overhauser dynamic nuclear polarization to study local water dynamics. *Journal of the American Chemical Society* **2009,** *131* (13), 4641-4647.

114. Franck, J. M.; Pavlova, A.; Scott, J. A.; Han, S., Quantitative cw Overhauser effect dynamic nuclear polarization for the analysis of local water dynamics. *Progress in nuclear magnetic resonance spectroscopy* **2013,** *74*, 33-56.

115. Cheng, C.-Y.; Wang, J.-Y.; Kausik, R.; Lee, K. Y. C.; Han, S., Nature of Interactions between PEO-PPO-PEO Triblock Copolymers and Lipid Membranes: (II) Role of Hydration Dynamics Revealed by Dynamic Nuclear Polarization. *Biomacromolecules* **2012,** *13* (9), 2624-2633.

116. Cheng, C.-Y.; Olijve, L. L.; Kausik, R.; Han, S., Cholesterol enhances surface water diffusion of phospholipid bilayers. *The Journal of chemical physics* **2014,** *141* (22), 12B613_1.

117. Pavlova, A.; Cheng, C.-Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S., Protein structural and surface water rearrangement constitute major events in the earliest aggregation stages of tau. *Proceedings of the National Academy of Sciences* **2016,** *113* (2), E127-E136.

118. Barnes, R.; Sun, S.; Fichou, Y.; Dahlquist, F. W.; Heyden, M.; Han, S., Spatially heterogeneous surface water diffusivity around structured protein surfaces at equilibrium. *Journal of the American Chemical Society* **2017,** *139* (49), 17890-17901.

119. Cheng, C.-Y.; Varkey, J.; Ambroso, M. R.; Langen, R.; Han, S., Hydration dynamics as an intrinsic ruler for refining protein structure at lipid membrane interfaces. *Proceedings of the National Academy of Sciences* **2013,** *110* (42), 16838-16843.

120. Franck, J. M.; Han, S., Chapter Five - Overhauser Dynamic Nuclear Polarization for the Study of Hydration Dynamics, Explained. In *Methods in Enzymology*, Wand, A. J., Ed. Academic Press: 2019; Vol. 615, pp 131-175.

121. Milov, A. D.; Ponomarev, A. B.; Tsvetkov, Y. D., Electron-electron double resonance in electron spin echo: Model biradical systems and the sensitized photolysis of decalin. *Chemical Physics Letters* **1984,** *110* (1), 67-72.

122. Jeschke, G., DEER Distance Measurements on Proteins. *Annual Review of Physical Chemistry* **2012,** *63* (1), 419-446.

123. Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W., Dead-time free measurement of dipole–dipole interactions between electron spins. *Journal of Magnetic Resonance* **2011,** *213* (2), 316-325.

124. Chiang, Y.-W.; Borbat, P. P.; Freed, J. H., The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *Journal of Magnetic Resonance* **2005,** *172* (2), 279-295.

125. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C.; Hilger, D.; Jung, H., DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* **2006,** *30* (3-4), 473-498.

126. Goldfarb, D., Gd 3+ spin labeling for distance measurements by pulse EPR spectroscopy. *Physical Chemistry Chemical Physics* **2014,** *16* (21), 9685-9699.

127. Cohen, M. R.; Frydman, V.; Milko, P.; Iron, M. A.; Abdelkader, E. H.; Lee, M. D.; Swarbrick, J. D.; Raitsimring, A.; Otting, G.; Graham, B., Overcoming artificial broadening in Gd 3+–Gd 3+ distance distributions arising from dipolar pseudo-secular terms in DEER experiments. *Physical Chemistry Chemical Physics* **2016,** *18* (18), 12847-12859.

128. Doll, A.; Qi, M.; Wili, N.; Pribitzer, S.; Godt, A.; Jeschke, G., Gd (III)–Gd (III) distance measurements with chirp pump pulses. *Journal of Magnetic Resonance* **2015,** *259*, 153-162.

129. Bahrenberg, T.; Rosenski, Y.; Carmieli, R.; Zibzener, K.; Qi, M.; Frydman, V.; Godt, A.; Goldfarb, D.; Feintuch, A., Improved sensitivity for W-band Gd (III)-Gd (III) and nitroxide-nitroxide DEER measurements with shaped pulses. *Journal of Magnetic Resonance* **2017,** *283*, 1-13.

130. Nakamura, T.; Tamada, D.; Yanagi, Y.; Itoh, Y.; Nemoto, T.; Utumi, H.; Kose, K., Development of a superconducting bulk magnet for NMR and MRI. *Journal of Magnetic Resonance* **2015,** *259*, 68-75.

131. Lalli, D.; Idso, M. N.; Andreas, L. B.; Hussain, S.; Baxter, N.; Han, S.; Chmelka, B. F.; Pintacuda, G., Proton-based structural analysis of a heptahelical transmembrane protein in lipid bilayers. *Journal of the American Chemical Society* **2017,** *139* (37), 13006-13012.

132. **!!! INVALID CITATION !!! {}**.

133. Crowley, P. B.; Kyne, C.; Monteith, W. B., Simple and inexpensive incorporation of 19 F-tryptophan for protein NMR spectroscopy. *Chemical communications* **2012,** *48* (86), 10681-10683.

134. Marley, J.; Lu, M.; Bracken, C., A method for efficient isotopic labeling of recombinant proteins. *Journal of biomolecular NMR* **2001,** *20* (1), 71-75.

135. Hartmann, S.; Hahn, E., Nuclear double resonance in the rotating frame. *Physical Review* **1962,** *128* (5), 2042.

136. Maly, T.; Debelouchina, G. T.; Bajaj, V. S.; Hu, K.-N.; Joo, C.-G.; Mak–Jurkauskas, M. L.; Sirigiri, J. R.; van der Wel, P. C.; Herzfeld, J.; Temkin, R. J., Dynamic nuclear polarization at high magnetic fields. *The Journal of chemical physics* **2008,** *128* (5), 02B611.

137. Bajaj, V. S.; Mak-Jurkauskas, M. L.; Belenky, M.; Herzfeld, J.; Griffin, R. G., Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization–enhanced solid-state NMR. *Proceedings of the National Academy of Sciences* **2009,** *106* (23), 9244-9249.

138. Hu, K.-N.; Yu, H.-h.; Swager, T. M.; Griffin, R. G., Dynamic nuclear polarization with biradicals. *Journal of the American Chemical Society* **2004,** *126* (35), 10844-10845.

139. Hu, K.-N.; Song, C.; Yu, H.-h.; Swager, T. M.; Griffin, R. G., High-frequency dynamic nuclear polarization using biradicals: A multifrequency EPR lineshape analysis. *The Journal of chemical physics* **2008,** *128* (5), 052302.

140. Bax, A.; Clore, G. M.; Driscoll, P. C.; Gronenborn, A. M.; Ikura, M.; Kay, L. E., Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of 13C-labeled proteins. *Journal of Magnetic Resonance (1969)* **1990,** *87* (3), 620-627.

141. Bloembergen, N., On the interaction of nuclear spins in a crystalline lattice. *Physica* **1949,** *15* (3-4), 386-426.

142. Takegoshi, K.; Nakamura, S.; Terao, T., 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical physics letters* **2001,** *344* (5-6), 631-637.

143. Hussain, S., *Studies of proteorhodopsin to investigate transmembrane protein function and dynamics*. University of California, Santa Barbara: 2014.

144. Wittig, I.; Braun, H.-P.; SchÃ¤gger, H., Blue native PAGE. *Nature protocols* **2006,** *1* (1), 418.
145. Tunuguntla, R.; Bangar, M.; Kim, K.; Stroeve, P.; Ajo-Franklin, C. M.; Noy, A., Lipid bilayer composition can influence the orientation of proteorhodopsin in artificial membranes. *Biophysical journal* **2013,** *105* (6), 1388-1396.
146. Lis, M.; Wizert, A.; Przybylo, M.; Langner, M.; Swiatek, J.; Jungwirth, P.; Cwiklik, L., The effect of lipid oxidation on the water permeability of phospholipids bilayers. *Physical Chemistry Chemical Physics* **2011,** *13* (39), 17555-17563.
147. KÃ¶hler, T.; Weber, I.; Glaubitz, C.; Wachtveitl, J., Proteorhodopsin Photocycle Kinetics Between pH 5 and pH 9. *Photochemistry and photobiology* **2017,** *93* (3), 762-771.
148. Yoshizawa, S.; Kawanabe, A.; Ito, H.; Kandori, H.; Kogure, K., Diversity and functional analysis of proteorhodopsin in marine Flavobacteria. *Environmental microbiology* **2012,** *14* (5), 1240-1248.
149. Tamogami, J.; Sato, K.; Kurokawa, S.; Yamada, T.; Nara, T.; Demura, M.; Miyauchi, S.; Kikukawa, T.; Muneyuki, E.; Kamo, N., Formation of M-Like Intermediates in Proteorhodopsin in Alkali Solutions (pH â‰¥ âˆ¼8.5) Where the Proton Release Occurs First in Contrast to the Sequence at Lower pH. *Biochemistry* **2016,** *55* (7), 1036-1048.
150. Hirai, T.; Subramaniam, S., Protein Conformational Changes in the Bacteriorhodopsin Photocycle: Comparison of Findings from Electron and X-Ray Crystallographic Analyses. *PLOS ONE* **2009,** *4* (6), e5769.
151. Andersson, M.; Malmerberg, E.; Westenhoff, S.; Katona, G.; Cammarata, M.; WÃ¶hri, A. B.; Johansson, L. C.; Ewald, F.; Eklund, M.; Wulff, M., Structural dynamics of light-driven proton pumps. *Structure* **2009,** *17* (9), 1265-1275.
152. Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G., Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *science* **1996,** *274* (5288), 768-770.
153. Reckel, S.; Gottstein, D.; Stehle, J.; Löhr, F.; Verhoefen, M.-K.; Takeda, M.; Silvers, R.; Kainosho, M.; Glaubitz, C.; Wachtveitl, J.; Bernhard, F.; Schwalbe, H.; Güntert, P.; Dötsch, V., Solution NMR Structure of Proteorhodopsin. *Angewandte Chemie International Edition* **2011,** *50* (50), 11942-11946.
154. Grimsley, G. R.; Scholtz, J. M.; Pace, C. N., A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Science* **2009,** *18* (1), 247-251.
155. Isom, D. G.; Castañeda, C. A.; Cannon, B. R., Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **2011,** *108* (13), 5260-5265.
156. Gerwert, K.; Hess, B.; Soppa, J. r.; Oesterhelt, D., Role of aspartate-96 in proton translocation by bacteriorhodopsin. *Proceedings of the National Academy of Sciences* **1989,** *86* (13), 4943-4947.
157. Stites, W. E.; Gittis, A. G.; Lattman, E. E.; Shortle, D., In a staphylococcal nuclease mutant the side-chain of a lysine replacing valine 66 is fully buried in the hydrophobic core. *Journal of Molecular Biology* **1991,** *221* (1), 7-14.
158. García-Moreno, B. E.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E., Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophysical chemistry* **1997,** *64* (1-3), 211-224.

159. Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S.; Stites, W. E., High apparent dielectric constants in the interior of a protein reflect water penetration. *Biophysical Journal* **2000,** *79* (3), 1610-1620.
160. Harms, M. J.; Schlessman, J. L.; Chimenti, M. S.; Sue, G. R.; Damjanović, A.; García-Moreno, B. E., A buried lysine that titrates with a normal pKa: Role of conformational flexibility at the protein-water interface as a determinant of pKa values. *Protein Science* **2008,** *17* (5), 833-845.
161. Takayama, Y.; Castaneda, C. A.; Chimenti, M.; García-Moreno, B. E.; Iwahara, J., Direct evidence for deprotonation of a lysine side chain buried in the hydrophobic core of a protein. *Journal of the American Chemical Society* **2008,** *130* (21), 6714-6715.
162. Goh, G. B.; García-Moreno, B. E.; Brooks III, C. L., The high dielectric constant of staphylococcal nuclease is encoded in its structural architecture. *Journal of the American Chemical Society* **2011,** *133* (50), 20072-20075.
163. Bashford, D.; Gerwert, K., Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. *Journal of Molecular Biology* **1992,** *224* (2), 473-486.
164. Wallace, J. A.; Wang, Y.; Shi, C.; Pastoor, K. J.; Nguyen, B. L.; Xia, K.; Shen, J. K., Toward accurate prediction of pKa values for internal protein residues: the importance of conformational relaxation and desolvation energy. *Proteins: Structure, Function, and Bioinformatics* **2011,** *79* (12), 3364-3373.
165. Wu, X.; Brooks, B. R., A virtual mixture approach to the study of multistate equilibrium: Application to constant ph simulation in explicit water. *PLoS computational biology* **2015,** *11* (10), e1004480.
166. Wu, X.; Lee, J.; Brooks, B. R., Origin of p K a Shifts of Internal Lysine Residues in SNase Studied Via Equal-Molar VMMS Simulations in Explicit Water. *The Journal of Physical Chemistry B* **2016,** *121* (15), 3318-3330.
167. Gunner, M. R.; Baker, N. A., Continuum electrostatics approaches to calculating pKas and Ems in proteins. In *Methods in enzymology*, Elsevier: 2016; Vol. 578, pp 1-20.
168. Markley, J. L., Observation of histidine residues in proteins by nuclear magnetic resonance spectroscopy. *Accounts of Chemical Research* **1975,** *8* (2), 70-80.
169. Forsyth, W. R.; Antosiewicz, J. M.; Robertson, A. D., Empirical relationships between protein structure and carboxyl pKa values in proteins. *Proteins: Structure, Function, and Bioinformatics* **2002,** *48* (2), 388-403.
170. Alexov, E.; Mehler, E. L.; Baker, N.; M. Baptista, A. n.; Huang, Y.; Milletti, F.; Erik Nielsen, J.; Farrell, D.; Carstensen, T.; Olsson, M. H., Progress in the prediction of pKa values in proteins. *Proteins: Structure, Function, and Bioinformatics* **2011,** *79* (12), 3260-3275.
171. Vorobjev, Y. N.; Scheraga, H. A.; Vila, J. A., Coupled molecular dynamics and continuum electrostatic method to compute the ionization pKa's of proteins as a function of pH. Test on a large set of proteins. *Journal of Biomolecular Structure and Dynamics* **2018,** *36* (3), 561-574.
172. Cvitkovic, J. P.; Pauplis, C. D.; Kaminski, G. A., PKA17—A Coarse-Grain Grid-Based Methodology and Web-Based Software for Predicting Protein pKa Shifts. *Journal of computational chemistry* **2019**.
173. Kougentakis, C. M.; Grasso, E. M.; Robinson, A. C.; Caro, J. A.; Schlessman, J. L.; Majumdar, A.; García-Moreno, B. E., Anomalous Properties of Lys Residues Buried in the Hydrophobic Interior of a Protein Revealed with 15N-Detect NMR Spectroscopy. *The Journal of Physical Chemistry Letters* **2018,** *9* (2), 383-387.

174. Wu, X.; Brooks, B. R., Hydronium Ions Accompanying Buried Acidic Residues Lead to High Apparent Dielectric Constants in the Interior of Proteins. *The Journal of Physical Chemistry B* **2018,** *122* (23), 6215-6223.

175. Harms, M. J.; Castañeda, C. A.; Schlessman, J. L.; Sue, G. R.; Isom, D. G.; Cannon, B. R.; García-Moreno, B. E., The pKa Values of Acidic and Basic Residues Buried at the Same Internal Location in a Protein Are Governed by Different Factors. *Journal of Molecular Biology* **2009,** *389* (1), 34-47.

176. Kim, J.; Mao, J.; Gunner, M., Are acidic and basic groups in buried proteins predicted to be ionized? *Journal of Molecular Biology* **2005,** *348* (5), 1283-1298.

177. Peetla, C.; Stine, A.; Labhasetwar, V., Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Molecular pharmaceutics* **2009,** *6* (5), 1264-1276.

178. Chen, Y.; Okur, H. I.; Lütgebaucks, C.; Roke, S., Zwitterionic and charged lipids form remarkably different structures on nanoscale oil droplets in aqueous solution. *Langmuir* **2017,** *34* (3), 1042-1050.

179. Szundi, I.; Stoeckenius, W., Effect of lipid surface charges on the purple-to-blue transition of bacteriorhodopsin. *Proceedings of the National Academy of Sciences* **1987,** *84* (11), 3681-3684.

180. Szundi, I.; Stoeckenius, W., Purple-to-blue transition of bacteriorhodopsin in a neutral lipid environment. *Biophysical Journal* **1988,** *54* (2), 227-232.

181. Szundi, I.; Stoeckenius, W., Surface pH controls purple-to-blue transition of bacteriorhodopsin. A theoretical model of purple membrane surface. *Biophysical Journal* **1989,** *56* (2), 369-383.

182. Voinov, M. A.; Smirnov, A. I., Ionizable nitroxides for studying local electrostatic properties of lipid bilayers and protein systems by EPR. In *Methods in enzymology*, Elsevier: 2015; Vol. 564, pp 191-217.

183. Wang, Y.; Botelho, A. V. r.; Martinez, G. V.; Brown, M. F., Electrostatic properties of membrane lipids coupled to metarhodopsin II formation in visual transduction. *Journal of the American Chemical Society* **2002,** *124* (26), 7690-7701.

184. Voinov, M. A.; Rivera-Rivera, I.; Smirnov, A. I., Surface electrostatics of lipid bilayers by EPR of a pH-sensitive spin-labeled lipid. *Biophysical Journal* **2013,** *104* (1), 106-116.

185. Váró, G.; Brown, L. S.; Needleman, R.; Lanyi, J. K., Binding of calcium ions to bacteriorhodopsin. *Biophysical Journal* **1999,** *76* (6), 3219-3226.

186. Song, J.; Kang, T. H.; Kim, M. W.; Han, S., Ion specific effects: decoupling ion–ion and ion–water interactions. *Physical Chemistry Chemical Physics* **2015,** *17* (13), 8306-8322.

187. Först, G.; Cwiklik, L.; Jurkiewicz, P.; Schubert, R.; Hof, M., Interactions of beta-blockers with model lipid membranes: molecular view of the interaction of acebutolol, oxprenolol, and propranolol with phosphatidylcholine vesicles by time-dependent fluorescence shift and molecular dynamics simulations. *European Journal of Pharmaceutics and Biopharmaceutics* **2014,** *87* (3), 559-569.

188. Jansen, M.; Blume, A., A comparative study of diffusive and osmotic water permeation across bilayers composed of phospholipids with different head groups and fatty acyl chains. *Biophysical Journal* **1995,** *68* (3), 997-1008.

189. Beranova, L.; Cwiklik, L.; Jurkiewicz, P.; Hof, M.; Jungwirth, P., Oxidation changes physical properties of phospholipid bilayers: fluorescence spectroscopy and molecular simulations. *Langmuir* **2010,** *26* (9), 6140-6144.

190. Veatch, S. L.; Keller, S. L., Separation of liquid phases in giant vesicles of ternary mixtures of phospholipids and cholesterol. *Biophysical journal* **2003,** *85* (5), 3074-3083.

191. Veatch, S. L.; Keller, S. L., Seeing spots: complex phase behavior in simple membranes. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **2005,** *1746* (3), 172-185.

192. Chang, C.-H.; Chen, J.-G.; Govindjee, R.; Ebrey, T., Cation binding by bacteriorhodopsin. *Proceedings of the National Academy of Sciences* **1985,** *82* (2), 396-400.

193. Eliash, T.; Weiner, L.; Ottolenghi, M.; Sheves, M., Specific binding sites for cations in bacteriorhodopsin. *Biophysical Journal* **2001,** *81* (2), 1155-1162.

194. Cevc, G., Membrane electrostatics. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes* **1990,** *1031* (3), 311-382.

195. Alexov, E.; Gunner, M., Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophysical Journal* **1997,** *72* (5), 2075-2093.

196. Damjanović, A.; Brooks, B. R.; García-Moreno, B. E., Conformational relaxation and water penetration coupled to ionization of internal groups in proteins. *The Journal of Physical Chemistry A* **2011,** *115* (16), 4042-4053.

197. Liu, J.; Swails, J.; Zhang, J. Z. H.; He, X.; Roitberg, A. E., A Coupled Ionization-Conformational Equilibrium Is Required To Understand the Properties of Ionizable Residues in the Hydrophobic Interior of Staphylococcal Nuclease. *Journal of the American Chemical Society* **2018,** *140* (5), 1639-1648.

198. Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N., Quantified electrostatic preorganization in enzymes using the geometry of the electron charge density. *Chemical science* **2017,** *8* (7), 5010-5018.

199. Gerwert, K.; Freier, E.; Wolf, S., The role of protein-bound water molecules in microbial rhodopsins. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2014,** *1837* (5), 606-613.

200. Nielsen, R. D.; Che, K.; Gelb, M. H.; Robinson, B. H., A ruler for determining the position of proteins in membranes. *Journal of the American Chemical Society* **2005,** *127* (17), 6430-6442.

201. Holt, A.; Killian, J. A., Orientation and dynamics of transmembrane peptides: the power of simple models. *European Biophysics Journal* **2010,** *39* (4), 609-621.

202. Segawa, T. F.; Doppelbauer, M.; Garbuio, L.; Doll, A.; Polyhach, Y. O.; Jeschke, G., Water accessibility in a membrane-inserting peptide comparing Overhauser DNP and pulse EPR methods. *The Journal of chemical physics* **2016,** *144* (19), 194201.

203. Gerfen, G. J.; Becerra, L. R.; Hall, D. A.; Griffin, R. G.; Temkin, R. J.; Singel, D. J., High frequency (140 GHz) dynamic nuclear polarization: polarization transfer to a solute in frozen aqueous solution. *The Journal of chemical physics* **1995,** *102* (24), 9494-9497.

204. Shi, L.; Lake, E. M.; Ahmed, M. A.; Brown, L. S.; Ladizhansky, V., Solid-state NMR study of proteorhodopsin in the lipid environment: secondary structure and dynamics. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2009,** *1788* (12), 2563-2574.

205. Kralj, J. M.; Bergo, V. B.; Amsden, J. J.; Spudich, E. N.; Spudich, J. L.; Rothschild, K. J., Protonation state of Glu142 differs in the green-and blue-absorbing variants of proteorhodopsin. *Biochemistry* **2008,** *47* (11), 3447-3453.

206. Scheidelaar, S.; Koorengevel, M. C.; Pardo, J. D.; Meeldijk, J. D.; Breukink, E.; Killian, J. A., Molecular model for the solubilization of membranes into nanodisks by styrene maleic acid copolymers. *Biophysical journal* **2015,** *108* (2), 279-290.
207. Tighe, B.; Tonge, S., Lipid-containing compositions and uses thereof. *UK patent number WO/1999/009955* **2000**.
208. Tonge, S.; Tighe, B., Responsive hydrophobically associating polymers: a review of structure and properties. *Advanced drug delivery reviews* **2001,** *53* (1), 109-122.
209. Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y.-P.; Dafforn, T.; Overduin, M., Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *Journal of the American Chemical Society* **2009,** *131* (22), 7484-7485.
210. Banerjee, S.; Pal, T. K.; Guha, S. K., Probing molecular interactions of poly (styrene-co-maleic acid) with lipid matrix models to interpret the therapeutic potential of the co-polymer. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2012,** *1818* (3), 537-550.
211. Johansson, L. C.; Stauch, B.; Ishchenko, A.; Cherezov, V., A bright future for serial femtosecond crystallography with XFELs. *Trends in biochemical sciences* **2017,** *42* (9), 749-762.
212. Nango, E.; Royant, A.; Kubo, M.; Nakane, T.; Wickstrand, C.; Kimura, T.; Tanaka, T.; Tono, K.; Song, C.; Tanaka, R., A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* **2016,** *354* (6319), 1552-1557.
213. Altenbach, C.; Kusnetzow, A. K.; Ernst, O. P.; Hofmann, K. P.; Hubbell, W. L., High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proceedings of the National Academy of Sciences* **2008,** *105* (21), 7439-7444.
214. Thonon, D.; Jacques, V.; Desreux, J. F., A gadolinium triacetic monoamide DOTA derivative with a methanethiosulfonate anchor group. Relaxivity properties and conjugation with albumin and thiolated particles. *Contrast media & molecular imaging* **2007,** *2* (1), 24-34.
215. Jeschke, G., MMM: A toolbox for integrative structure modeling. *Protein Science* **2018,** *27* (1), 76-85.
216. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004,** *25* (13), 1605-1612.
217. Shah, A.; Roux, A.; Starck, M.; Mosely, J. A.; Stevens, M.; Norman, D. G.; Hunter, R. I.; El Mkami, H.; Smith, G. M.; Parker, D., A gadolinium spin label with both a narrow central transition and short tether for use in double electron electron resonance distance measurements. *Inorganic chemistry* **2019,** *58* (5), 3015-3025.
218. Mehler, M.; Eckert, C. E.; Leeder, A. J.; Kaur, J.; Fischer, T.; Kubatova, N.; Brown, L. J.; Brown, R. C.; Becker-Baldus, J.; Wachtveitl, J., Chromophore distortions in photointermediates of proteorhodopsin visualized by dynamic nuclear polarization-enhanced solid-state NMR. *Journal of the American Chemical Society* **2017,** *139* (45), 16143-16153.
219. Lueders, P.; Jäger, H.; Hemminga, M. A.; Jeschke, G.; Yulikov, M., Multiple Pathway Relaxation Enhancement in the System Composed of Three Paramagnetic Species: Nitroxide Radical–Ln3+–O2. *The journal of physical chemistry letters* **2012,** *3* (10), 1336-1340.
220. Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency. *Angewandte Chemie* **2013,** *125* (41), 11058-11061.

221. Petrache, H. I.; Tristram-Nagle, S.; Nagle, J. F., Fluid phase structure of EPC and DMPC bilayers. *Chemistry and physics of lipids* **1998,** *95* (1), 83-94.
222. Dreier, L. B.; Nagata, Y.; Lutz, H.; Gonella, G.; Hunger, J.; Backus, E. H.; Bonn, M., Saturation of charge-induced water alignment at model membrane surfaces. *Science advances* **2018,** *4* (3), eaap7415.