2018

Molecular Dynamics and Nonlinear Dynamics Studies of Chemical Systems

Sadegh Faramarzi Ganj Abad

Molecular Dynamics and Nonlinear Dynamics Studies of Chemical Systems

Sadegh Faramarzi Ganj Abad

Dissertation submitted
to the Eberly College of of Arts and Sciences
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

John B. Mertz, Ph.D., Chair
Kenneth Showalter, Ph.D., Chair
Terry Gullion, Ph.D.
Justin Legleiter, Ph.D.
James P. Lewis, Ph.D.
Mark Tinsley, Ph.D.

C. Eugene Bennett Department of Chemistry

Morgantown, West Virginia
2018

Keywords: Molecular Dynamics Simulations, Proteorhodopsin, Surfactants, Photosensitive BZ reaction, Phase Response Curves, Synchronization, Chimera States

ABSTRACT

Molecular Dynamics and Nonlinear Dynamics Studies of Chemical Systems

Sadegh Faramarzi Ganj Abad

Proteorhodopsin (PR) is a microbial proton pump that is ubiquitous in marine environments and may play an important role in the oceanic carbon cycle. Photoisomerization of the retinal chromophore in PR leads to a series of proton transfers between specific acidic amino acid residues and the Schiff base of retinal, culminating in a proton motive force to facilitate ATP synthesis. In this work, we study the effect of the ionization of a key residue called proton donor on the hydration and dynamics of PR. We also study the effect of surfactant environment on the stability of PR.

Synchronization is the adjustment of the rhythms of oscillating systems due to the mutual interaction or an external force, and it is observed in many natural systems, such as heart cells, firing in populations of fireflies and the human brain. Studying the different ways by which unsynchronized chemical oscillators transform to synchronized sets or subsets allows for better understanding of the more complex oscillatory systems in nature. In this work, we study the synchronization of photosensitive Belouzov-Zhabotinsky (BZ) oscillators as a model for complex biological systems such as neurons.

Acknowledgement

I would like to thank my advisors, Dr. John B. Mertz and Dr. Kenneth Showalter for providing a great research atmosphere and helping me during my graduate education. I also want to express my gratitude to all the committee members: Dr. Terry Gullion, Dr. Justin Legleiter, Dr. James P. Lewis and Dr. Mark Tinsley for evaluating my dissertation. I also thank the research group members in both Mertz and Showalter laboratories: Dr. Choongkeun Lee, Dr. Jun Feng, Chitrak Gupta, Yue Ren, Eric Sefah, Violetta Burns Casamayor, Corey Cleavenger, Nicolas Frazee, Austin Clark, Dr. Razan Snari, Dr. Desmond Yengi, Dr. Darrell Collosin, Syed Jazli Syed Jamaluddin, David Mersing, and all undergraduate researchers for their help during my graduate education. My special thanks go to my family for their patience and support throughout my PhD.

Contents

List of Figures

List of Tables

Chapter 1

1. Proteorhodopsin as a Natural Proton Pump

1.1 Introduction

Proteorhodopsin (PR) is a microbial membrane protein, discovered in 2000[1]. It was expressed in *Escherichia coli* and has a chromophore called retinal that makes it an active, light-driven proton pump by generating a proton, H^+, gradient across the cell membrane[1]. The light-harvesting properties of PR and its ability to function in different environments makes it a promising renewable energy resource. The proton pump occurs in a multi-step process known as photocycle, that involves the ionization of several residues in protein interior, as shown in Fig. 1.2. At the end of the photocycle, a proton is transfered from cytoplasm into the periplasm using the energy of a photon of sun light (Fig. 1.1).

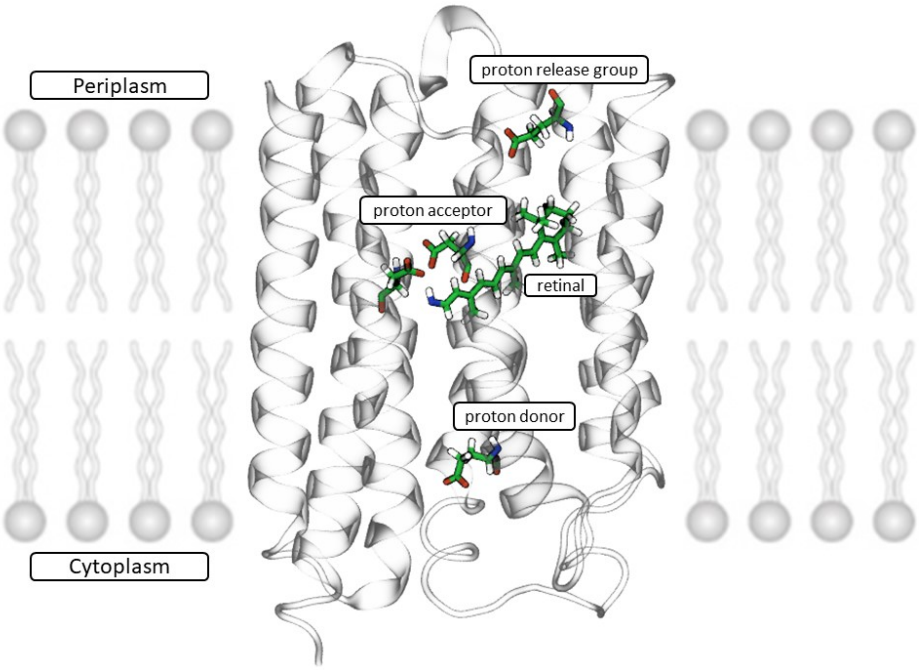


Figure 1.1. Proton pump in proteorhodopsin. The proton pump is initiated by photon absorption, leading to isomerization of retinal from all-trans to 13-cis. This culminates in a proton being pumped from the cytoplasmic side (CY) to the periplasmic side (PP) of the protein, with multiple steps that involve exchange of proton between retinal and proton donor and proton acceptor residues.

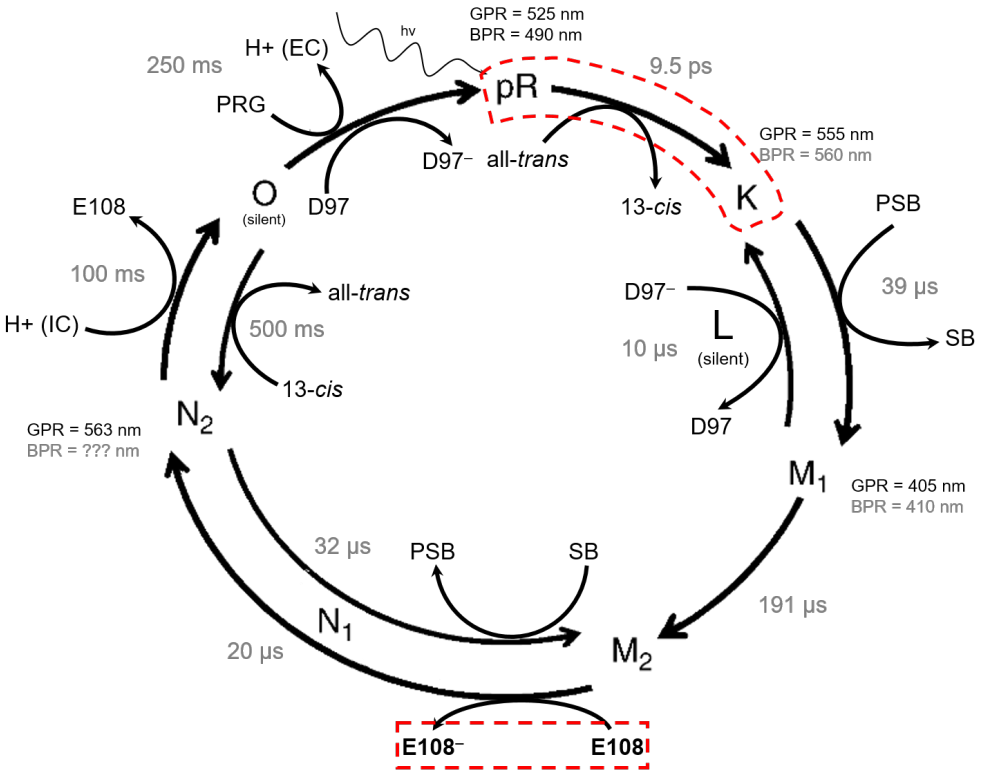


Figure 1.2. Photocycle of proteorhodopsin.The process starts with all-*trans* to 13-*cis* isomerization of the retinal, known as K state. A proton from a Schiff base in retinal is transfered to a neighboring aspartic acid residue, known as the proton acceptor (M state). In the next step (N), a proton from a glutamic acid on the cytoplasmic side, E108, is donated to the retinal Schiff base to make up the loss of a proton in the previous state. The proton donated by the proton donor is replaced by a proton from the bulk water in the cytoplasm (O state). Eventually, a proton from the proton acceptor, D97, is transfered to a putative proton release group, E142, and the protein returns back to the initial state, known as the dark state. An aspartic acid in the vicinity of the retinal Schiff base stabilizes the Schiff base with hydrogen bonding[2]. All the steps are indicated a letter.

The retinal plays an essential role in the proton pump. Its long polyene chain allows for the isomerization of retinal upon absorbing a photon of sun light (Fig. 1.1). Ionization of a Schiff base at the end of the polyene chain allows for the proton exchange between retinal and the proton acceptor and the proton donor residues in the photocycle. Protorhodopsin has two variants, green and blue (GPR and BPR, respectively). A major difference between the two variants is the residue 105, which is leucine is GPR and glutamine in BPR. This residue is responsible for color tuning in PR[3]. GPR and BPR have maximum wavelengths at 525 and 490 nm, respectively[4]. GPR has a faster photocycle rate and higher proton pumping activity than BPR[5]. The X-ray structure exists only for BPR.

The photocycle starts with the light-induced ultrafast isomerization of retinal from all-*trans* to 13-*cis* that leads to K state (Fig. 1.2). The K state then decays to M state that is marked by deprotonation of Schiff

base. A neighboring D97, known as proton acceptor, receives the proton from Schiff base (Fig. 1.1). Schiff base is reprotonated from E108 in cytoplasmic side of the protein, known as proton donor, to reach an intermediate known as N. E108 in turn is reprotonated from the bulk water in cytoplasmic side. Meanwhile, retinal reisomerization occurs and the system reaches to O state. At the end of the photocycle, a proton from the proton acceptor is released to the periplasmic side, putatively via E142 known as proton release group. PR returns to dark state and it is ready to start a new photocycle by absorbing another photon of sun light.

Due to difficulties associated with extraction of membrane proteins, the structural and functional details of retinal proteins, including PR, have been unknown for decades. The structure of bacteriorhodopsin (bR) that was resolved with X-ray crystallography in 1999[6] was an exception. Bacteriorhodopsin is the most widely studied bacterial membrane protein and has similarities with PR, such as having 7 transmembrane helices, retinal and several ionizable residues in the hydrophobic cavity of the protein. In the following chapters, we will compare our observations with the known facts about bR.

1.2 Role of the Hydration in PR Photocycle

Among different conformational changes that facilitate proton pump function, the ones that control the accessible surface area of cytoplasmic (CY) side of PR are of particular importance, since they control the dynamics of water molecules in protein interior and around the key residues. Experimental evidences infer that reducing the humidity[7] or adding osmotically active solutes (dehydration)[8] inhibits the proton exchange between the proton donor residue and the Schiff in bR[9], highlighting the role of water molecules in the proton transfer inside protein. However, the conformational changes that lead to the opening of the CY side of PR have not been well understood. Generally, conformation of different photointermediates of the bacterial seven-helix transmembrane proteins reported in different X-ray studies are not consistent[10]. Although E108 is known to be a proton donor, it is located far from bulk water molecules, i.e. 7 Å. Being in the hydrophobic region is known to be a reason for unusually high pK_a of E108[11]. Still E108 needs water molecules around it to serve as a proton donor[12]. Therefore, water flow inside the protein interior is a critical step in the proton pump function of PR[9]. Furthermore, wide angle X-ray scattering (WAXS) experiments show that a major rearrangement of water molecules inside protein and around Schiff base must happen prior to its deprotonation to decrease its pH[13]. The effect of the ionization of E108 on the hydration of PR in dark and K states of the photocycle are discussed in Chapter 2.

1.3 Surfactant as Membrane Substitute

Amphiphilic molecules possess both hydrophilic and hydrophobic properties, due to a polar headgroup and hydrocarbon chain(s). One of the common groups of amphiphilic molecules are lipids, and they play a vital role in living systems, separating the interior of cells from the aqueous environment via the hydrophobic effect [14]. At the same time, the membrane allows polar molecules and ions necessary for biological function to pass through via the hydrophilic effect. Without such elegant regulations, life on earth would not be possible.

Surfactants (i.e., detergents) are another group of amphiphilic molecules with the ability to self-assemble into micelles and are commonly used in membrane protein extraction and solvation [15]. Depending on their effect on the stability and functionality of the protein, surfactants are categorized as harsh or mild. Some empirical rules are used to predict whether a surfactant is harsh or mild. For example, Fig. 1.2 represents some of these rules. Despite the widespread use of surfactants in the extraction, purification and crystallization of membrane proteins, an atomistic-level perspective of the physical properties that lead to membrane protein solvation remains relatively unknown.

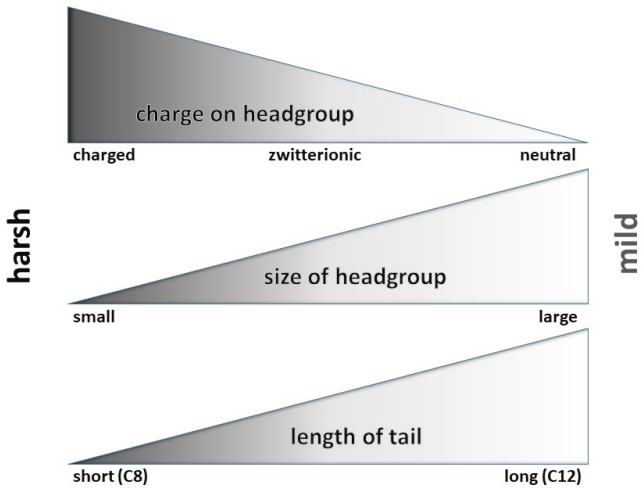


Figure 1.3. General trends observed in mild and harsh surfactants. The figure is modified from Ref. [16]. If a detergent has a charged small headgroup and a short tail, it is categorized as harsh.

In chapter 3, we study physical properties of pure surfactant micelles such as self-assembly starting from random distributions of detergent molecules and compare them with simulations of preformed micelles,

using molecular dynamics (MD) simulations. In chapter 4, we use PR as a canonical model for studying the effect of surfactant environment on the dynamics and function of membrane proteins, as it is the most abundant member of retinal superfamily[16]. Similar to many membrane proteins, PR has seven transmembrane helices. Also, the structure and function of PR is sensitive to the environment. With site-directed spin-labeling, continuous-wave electron spin resonance (ESR) and Overhauser dynamic nuclear polarization (DNP) methods, Stone and co-workers found that PR forms mostly hexamers in *n*-dodecyl-β-D-maltoside (DDM) detergent environment. In high concentration of DDM, monomers are the predominant species, possibly because the detergents shield PRs from each other by covering their hydrophobic regions, hence preventing them from reconstitution into oligomers[17]. Also, PR is present as monomers in a zwitter ionic detergent, *n*-dodecylphosphocholine (DPC) environment[18]. Moreover, the pK_a of the proton acceptor, D97, is different in the two detergent environments. All these facts make PR an appropriate model for studying proteomicelle properties. Solid state NMR studies show that a salt bridge between R51 from a monomer and D52 from the neighboring monomer plays a significant role in oligomerization of PR[19]. The salt bridge stabilizes pentamers, so that the mutation of any of the two residues destroys the pentamers and facilitates formation of hexamers[19]. Salt bridge is a noncovalent interaction between amino acid side chains with opposite full-electron charges, namely between Arg/Lys and Asp/Glu, all in neutral form. Simulations of the proteomicelle systems answers some of questions about the oligomeric states of PR.

1.4 Molecular Dynamics Simulations

The time-dependent pathway of a particle i, can be predicted by knowing its position and velocity and applying Newton's second law of motion:

$$\frac{\mathrm{d}^2\mathbf{x_i}}{\mathrm{d}t^2} = \frac{F_{x_i}}{m_i} = -\frac{\mathrm{d}V_{x_i}}{m_i \mathrm{d}\mathbf{x_i}}, \tag{1.1}$$

where x_i is the position of a particle of mass of m_i, and F_{x_i} and V_{x_i} are the applied force and potential energy, respectively. If we consider each atom as an independent particle, the forces or potential energies applied on it arise from its interaction with other atoms in the system. All of these interactions fall into two general categories: 1) bonded and 2) non-bonded interaction:

$$E_{total} = E_{bonded} + E_{nonbonded}, \tag{1.2}$$

where

$$E_{bonded} = E_{bond} + E_{angle} + E_{dihedral} + E_{improper}, \tag{1.3}$$

and

$$E_{nonbonded} = E_{VanderWaals} + E_{electrostatic}, \tag{1.4}$$

Followings formulas are used for each energy term in Eq. 1.3 and 1.4

$$E_{bond} = \sum k_i^{bond}(r_i - r_0)^2, \tag{1.5}$$

$$E_{angle} = \sum k_i^{angle}(\theta_i - \theta_0)^2, \tag{1.6}$$

$$E_{dihedral} = \sum k_i^{dihe}[1 + cos(n_i\phi_i + \delta_i)], \tag{1.7}$$

$$E_{VanderWaals} = \sum_{i=1}\sum_{j\neq i} 4\varepsilon_{ij}\left[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^{6}\right], \tag{1.8}$$

$$E_{electrostatic} = \sum_{i=1}\sum_{j\neq i} \frac{q_i q_j}{\varepsilon r_{ij}}, \tag{1.9}$$

where k_i^{bond}: the bond energy constant, r_i: the distance between any pair of atoms, r_0: the equilibrium distance between any pair of atoms, k_i^{angle}: the angular energy constant , θ_i: the angle between three atoms bond together, θ_0: the equilibrium angle between three atoms bond together, k_i^{dihe}: dihedral energy constant, ϕ_i: the dihedral angle between 4 atoms, n_i: multiplicity, δ_i: the angle where the potential passes through its minimum value, ε_i: the depth of the potential well, σ_{ij}: the finite distance at which the inter-atomic potential is zero, r_{ij}: the distance between two atoms, q_i and q_j: atomic charges, and ε: Coulomb's constant.

The functions and parameters defining these potential energies are referred to as a force field. Once the force field is defined, computer programs are used to simulate the dynamics of a given molecular system, starting with the initial coordinates of the atoms that are usually taken from the X-ray structures. There is always a compromise between accuracy and computational efficiency used to calculate these potential energies, and different MD packages are categorized based on their respective approaches to modeling these interactions. The parameters for the potential energies are obtained by a combination of quantum mechanical calculations and experiment.

Since biological systems contain a vast number of particles, it is impossible to calculate macroscopic properties such as pressure and temperature by analytically solving the potential energy equations at the microscopic level. Rather, we use statistical mechanics as a bridge from the microscopic level to macroscopic experimental observables, by studying average behavior of the systems in the context of probability theory. MD simulations are able to capture the atomistic details of the micelle assembly and behavior over ns-μs timescales, making it an effective complementary tool to experimental methods.

References

1. Béjà, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science (New York, N.Y.)* **289**, 1902–1906 (2000).

2. Feng, J. & Mertz, B. Proteorhodopsin Activation Is Modulated by Dynamic Changes in Internal Hydration. *Biochemistry* **54**, 7132–7141 (2015). URL http://dx.doi.org/10.1021/acs.biochem.5b00932.

3. Amsden, J. J. *et al.* Different Structural Changes Occur in Blue- and Green-Proteorhodopsins during the Primary Photoreaction. *Biochemistry* **47**, 11490–11498 (2008). URL https://doi.org/10.1021/bi800945t.

4. Man, D. *et al.* Diversification and spectral tuning in marine proteorhodopsins. *The EMBO Journal* **22**, 1725–1731 (2003). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC154475/.

5. Jung, J. Y., Choi, A. R., Lee, Y. K., Lee, H. K. & Jung, K.-H. Spectroscopic and photochemical analysis of proteorhodopsin variants from the surface of the Arctic Ocean. *FEBS Letters* **582**, 1679–1684. URL https://febs.onlinelibrary.wiley.com/doi/abs/10.1016/j.febslet.2008.04.025.

6. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution 11edited by D. C. Rees. *Journal of Molecular Biology* **291**, 899–911 (1999). URL http://www.sciencedirect.com/science/article/pii/S0022283699930279.

7. Váró, G. & Lanyi, J. K. Distortions in the photocycle of bacteriorhodopsin at moderate dehydration. *Biophysical Journal* **59**, 313–322 (1991). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1281148/.

8. Cao, Y. *et al.* Water is required for proton transfer from aspartate-96 to the bacteriorhodopsin Schiff base. *Biochemistry* **30**, 10972–10979 (1991). URL http://dx.doi.org/10.1021/bi00109a023.

9. Brown, L. S., Váró, G., Needleman, R. & Lanyi, J. K. Functional significance of a protein conformation change at the cytoplasmic end of helix F during the bacteriorhodopsin photocycle. *Biophysical Journal* **69**, 2103–2111 (1995). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1236444/.

10. Hirai, T., Subramaniam, S. & Lanyi, J. K. Structural snapshots of conformational changes in a seven-helix membrane protein: lessons from bacteri-

orhodopsin. *Current Opinion in Structural Biology* **19**, 433–439 (2009). URL http://www.sciencedirect.com/science/article/pii/S0959440X09001043.

11. Oka, T., Yagi, N., Tokunaga, F. & Kataoka, M. Time-resolved X-ray diffraction reveals movement of F helix of D96n bacteriorhodopsin during M-MN transition at neutral pH. *Biophysical Journal* **82**, 2610–2616 (2002). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1302049/.

12. Wang, T. *et al.* Deprotonation of D96 in Bacteriorhodopsin Opens the Proton Uptake Pathway. *Structure* **21**, 290–297 (2013). URL http://www.sciencedirect.com/science/article/pii/S0969212613000130.

13. Andersson, M. *et al.* Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17**, 1265–1275 (2009). URL /structure/abstract/S0969-2126(09)00290-1.

14. Bagchi, B. Water in Biological and Chemical Processes: From Structure and Dynamics to Function (2013). URL /core/books/water-in-biological-and-chemical-processes/CDC338A7F87DB86352D54421764E5E53.

15. Koley, D. & J. Bard, A. Triton X-100 concentration effects on membrane permeability of a single HeLa cell by scanning electrochemical microscopy (SECM). *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16783–16787 (2010).

16. Bamann, C., Bamberg, E., Wachtveitl, J. & Glaubitz, C. Proteorhodopsin. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**, 614–625 (2014). URL http://www.sciencedirect.com/science/article/pii/S0005272813001667.

17. Stone, K. *et al.* Structural Insight into Proteorhodopsin Oligomers. *Biophysical Journal* **104**, 472–481 (2013). URL http://www.sciencedirect.com/science/article/pii/S0006349512050783.

18. Hussain, S., Kinnebrew, M., S. Schonenbach, N., Aye, E. & Han, S. Functional Consequences of the Oligomeric Assembly of Proteorhodopsin. *Journal of Molecular Biology* **427**, 1278–1290 (2015).

19. Maciejko, J. *et al.* 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* **137**, 9032–9043 (2015). URL http://dx.doi.org/10.1021/jacs.5b03606.

Chapter 2[1]

2. Allosteric Effects of the Proton Donor on the Microbial Proton Pump, Proteorhodopsin

2.1 Abstract

Proteorhodopsin (PR) is a microbial proton pump that is ubiquitous in marine environments and may play an important role in the oceanic carbon cycle. Photoisomerization of the retinal chromophore in PR leads to a series of proton transfers between specific acidic amino acid residues and the Schiff base of retinal, culminating in a proton motive force to facilitate ATP synthesis. The proton donor in a similar retinal protein, bacteriorhodopsin, acts as a latch to allow influx of bulk water. However, it is unclear if the proton donor in PR, E108, utilizes the same latch mechanism to become internally hydrated. Here we used molecular dynamics (MD) simulations to model the changes in internal hydration of PR during photoactivation with the proton donor in protonated and deprotonated states. We find that there is a stark contrast in the levels of internal hydration of the cytoplasmic half of PR based on the protonation state of E108. Instead of a latch mechanism, deprotonation of E108 acts as a gate, taking advantage of a nearby polar residue (S61) to promote formation of a stable water wire from bulk cytoplasm to the retinal binding pocket over hundreds of ns. No large-scale conformational changes occur in PR over the μs timescale. This subtle yet clear difference in the effect of deprotonation of the proton donor in PR may help explain why the photointermediates that involve the proton donor (i.e., M and N states) have timescales that are orders of magnitude different from the archaeal proton pump, bacteriorhodopsin. In general, our study highlights the importance of understanding how structural fluctuations lead to differences in the way that retinal proteins accomplish the same task.

2.2 Introduction

Proteorhodopsin (PR) is a microbial proton pump that was first identified in marine phytoplankton[1], and its discovery spawned the field of metagenomics. Two variants of PR exist – green and blue[2] – that have

[1]This chapter is submitted to Biophysical Journal, 2018

adapted to optimize the depth-dependent wavelengths of light at which they reside in the ocean. Since their initial discovery, PR-encoding genes have also been found in freshwater bacteria[3], soil-bound bacteria[4], viruses[5], and eukaryotes[6]. Although the exact biological role of PR is undetermined, it may be a vital component in several marine ecosystems, notably the carbon cycle for degradation of complex organic matter[7] and as a secondary source of ATP synthesis in iron-deficient regions of the ocean[8]. In addition, PR has the potential for applications in solar-to-chemical energy conversion[9].

PR is a member of the microbial retinal protein family, proteins with a seven-transmembrane (7TM) alpha-helical topology and a retinal chromophore that is covalently bound to a lysine residue on helix G. Absorption of a photon of light leads to an all-*trans* to 13-*cis* isomerization, initiating the photocycle[10]. This structural rearrangement of the retinal binding pocket leads to a shift in the pK_a of key acidic residues in the interior of PR that allows for proton transfers to take place – the proton acceptor (D97), proton donor (E108), and proton release group (as yet unidentified)[11] (Fig. 2.1). Each proton transfer is characterized by a spectroscopically distinct photointermediate, with the end product being a net transport of one proton across the inner membrane to the periplasmic space of the bacterial cell. This proton gradient can then be used for synthesis of ATP.

The majority of what we know about microbial proton pumps comes from decades of research on bacteriorhodopsin (bR)[12,13,14]. Although many similarities exist between bR and PR, proteorhodopsin possesses several unique characteristics that must be elucidated to fully understand how it generates a proton motive force. The timescale for the photocycle of PR is an order of magnitude faster than that of bR[14]. In addition, the pK_a of the proton acceptor is noticeably higher (7.1-7.6[10,15]) than bR (2.6[16]), leading to an ability for PR to invert the direction of proton pumping at low pH[17]. In bR, the proton release group is comprised of a pair of glutamic acids[18], but in PR there is a lack of corresponding residues. Finally, a single residue in the retinal binding pocket at position 105 – a leucine in green PR and a glutamine in blue PR – acts as a color-tuning switch between the two variants of PR[19].

The cytoplasmic half of both bR and PR undergo large-scale conformational changes that allow for an influx of bulk water in the later stages of their respective photocycles[14]. This influx is necessary to form a stable water wire to bridge the 10-15 Å gap that exists between the proton donor and the protonated Schiff base (PSB) of retinal and allow for proton transfer through a Grotthuss mechanism[20,21,22,23]. In bR, this outward movement of the TM helices are controlled by the proton donor; deprotonation of this residue leads to opening of the protein on the ns timescale[24]. This rapid release of the helical bundle in bR (i.e., "latch mechanism") has a two-fold purpose: 1) formation of the water wire and 2) to prevent backflow of protons that impede the photocycle[25]. Although PR has a glutamic acid (E108) in the same position as the proton

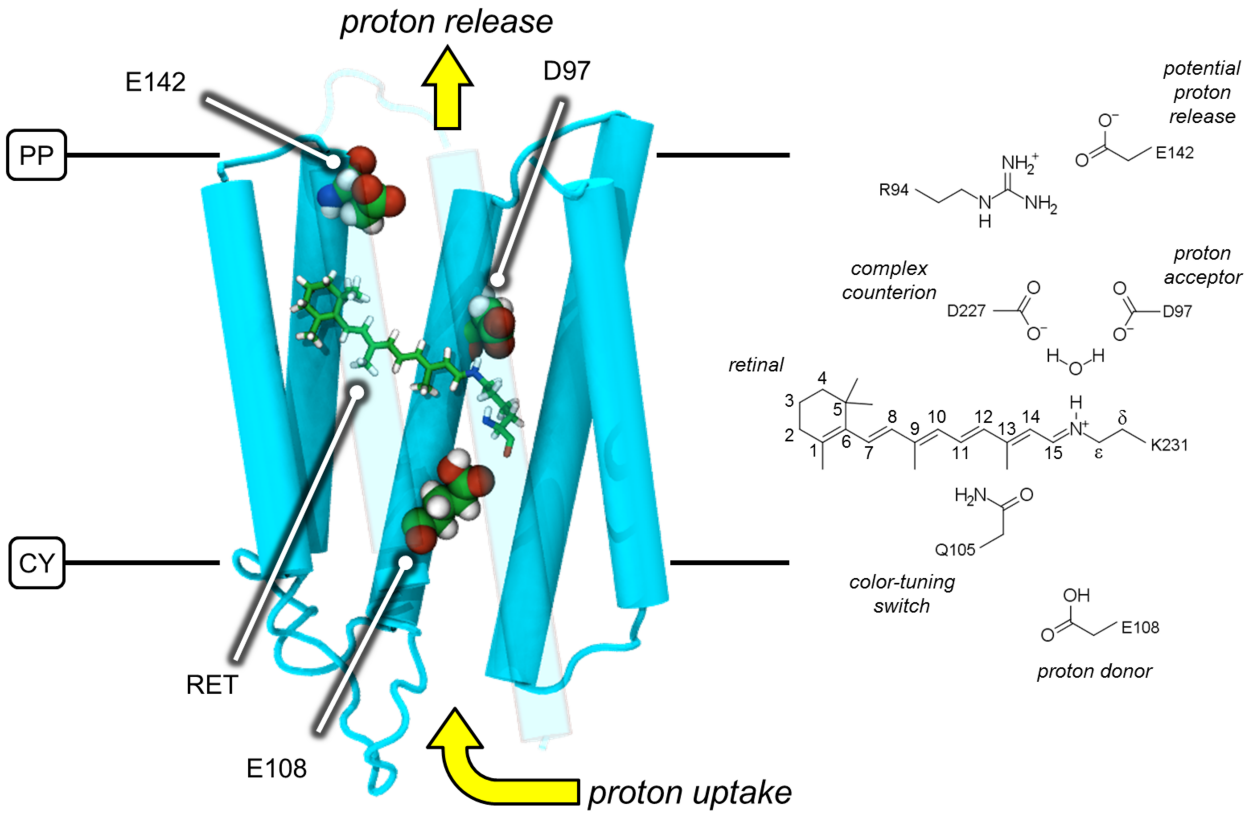


Fig. 2.1. Proteorhodopsin functions as a proton pump. Left: tertiary structure of blue proteorhodopsin (PDB 4JQ6). Right: chemical structure of retinal and important residues in the photocycle of PR. Upon photoisomerization of retinal (RET), PR proceeds through its photocycle, utilizing a series of proton transfers to facilitate proton pumping. The initial proton transfer occurs from the protonated Schiff base (PSB) to the proton acceptor, D97. The SB then becomes reprotonated by the proton donor, E108. E108 is then protonated from bulk cytoplasm, and D97 donates its proton to the putative proton release group, E142, allowing an excess proton to be released in the the periplasmic space. PP: periplasmic side of the inner membrane; CY: cytoplasmic side of the inner membrane.

donor in bR, it lacks a pair of glutamic acids that act as a proton uptake complex[26] as well as a hydrogen-bonding partner in helix B that lies directly across from the proton donor[12]. In addition, an acidic residue as the proton donor is not conserved among other proteorhodopsin-like proton pumps, such as a lysine in *E. sibiricum* rhodopsin[27] and a histidine on helix B in eubacterial proton pumps[28]. Our recent computational study of PR hinted that the protonation state of the proton donor in PR may have a more subtle influence on the internal hydration of protein, as we did not observe the latch mechanism seen in bR[29].

In this work, we set out to characterize the effect of protonation of the proton donor in blue PR, E108, on the early photointermediates of the PR photocycle using molecular dynamics (MD) simulations. The most reliable tertiary structure of PR is from the blue variant in the dark state[30], limiting us to a μs-timescale investigation of the dark and K states. We find that deprotonation of E108 controls the interior hydration of the cytoplasmic half of PR, with full hydration occurring over hundreds of ns. Unlike in bR, there is no

large-scale conformational change associated with deprotonation of the proton donor – rather, E108 acts as a gate that opens and closes to allow or deny access to the internal water channel. In addition, we observe a significant degree of crosstalk between E108 and the color-tuning switch in PR as well as differences in interactions of the retinal binding pocket that may help explain how proton release occurs in PR.

2.3 Methods

Simulation setup. The X-ray crystal structure of BPR from Med12 in the dark state (PDB 4JQ6)[30] was used as the starting structure. Missing loops were modeled based on the loops of the NMR structure of GPR (PDB 2L6X)[31]. We have truncated the N- and C-termini (D9 to S231 are included) and have numbered residues based on green PR.

Two sets of equilibrium MD simulations were carried out based on the protonation state of E108. In the protonated state, the $O_{\delta 2}$ oxygen of the carboxylic group is protonated, and in the deprotonated state, both oxygens of the carboxylic terminal group are deprotonated. The ionization state of all other residues are identical between the two sets of simulations. Based on our earlier simulations of PR[29], we protonated H75 on the N_{δ} atom of the imidazole ring. The putative proton release group, E142, is deprotonated since its pK_a is reported to be < 7.5 in blue PR[32].

PR with protonated E108 was placed within 160 lipid molecules (80 per leaflet) with a molar ratio of 3:1 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE):1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG) as a proxy for a gram-negative inner membrane[33] using the replacement method[34,35] of the CHARMM-GUI server (http://www.charmm-gui.com)[36,37]. After solvation with 7867 TIP3P water molecules and ionization with 44 sodium ions (to neutralize the system) the total system size was 47,288 atoms. The corresponding system with deprotonated E108 has 129 lipids (65 and 64 in upper and lower leaflets, respectively) with the same 3:1 molar ratio of POPE:POPG. There were 7,030 water molecules, 38 sodium ions, and 1 chloride ion for a total of 40,880 atoms.

The systems were minimized and equilibrated based on the CHARMM-GUI protocol with the CHARMM c36 force field for lipids and proteins[38]. The minimization and equilibration protocol involves 6 steps that gradually remove restraints from the protein backbone, lipids, water molecules, and ions over 375 ps[34]. Retinal force field parameters were obtained from Scott Feller[39,40]. After equilibration, the systems and force fields were converted to AMBER-compatible formats in CHAMBER[41], and all production simulations were run with AMBER 12 and AMBER 16 packages with GPU acceleration[42,43]. All equilibrium MD simulations were run in the *NPT* ensemble at 300 K and 1 bar using a Langevin thermostat and a

Berendsen barostat (relaxation time of 8 ps) with a 2 fs timestep. A cutoff of 8 Å was applied for non-bonded interactions. To initiate K state simulations through all-*trans* to 13-*cis* isomerization of retinal, the torsional potential of the C13=C14 bond was temporarily lowered, as has been done previously for PR[29] and rhodopsin[44].

Analysis. VMD[45] was used for analysis of angles, interatomic distances, solvent accessibility, hydrogen bonds, and root mean squared deviation (RMSD). Images of representative snapshots were rendered in either VMD or PyMOL[46]. Water density maps were calculated with MDAnalysis[47]. CAVER 3.0[48,49] was used to detect cavities inside PR. The probe radius was 0.8 Å, with a shell radius of 3 Å and a shell depth of 4 Å. A threshold of 10 was used for clustering tunnels. For hydrogen bond analysis, a distance threshold of 3 Å and an angle threshold of 20° (between the donor heavy atom - hydrogen - acceptor heavy atom) is used. *N*-body information theory (NbIT) was used to identify the correlation of dynamics between distant residues. The methodology uses configurational entropy of molecules in MD trajectories to identify distal residues that communicate with each other or organize dynamics within functional sites[50,51].

2.4 Results and Discussion

Deprotonation of E108 leads to an increase in hydration from bulk cytoplasm

A detailed examination of the proton donor, E108, reveals that titration of the sidechain has a very specific purpose: to prohibit internal hydration and subsequent interaction with the PSB. At the most fundamental level, this function manifests itself in the orientation of the E108 sidechain. In the dark state, the χ_1 dihedral has a dominant conformation at -70°, regardless of protonation state. However, the χ_2 dihedral is highly dependent on protonation state of E108: when protonated, χ_2 predominantly has a $\pm$ 180° orientation, leading to a linear conformation that occludes the interior channel of the cytoplasmic half of PR from bulk cytoplasm (Fig. 2.2 A and Supporting Information Fig. S2.1 and S2.2). In contrast, when E108 is deprotonated the χ_2 dihedral is -65°, corresponding to an outward rotation towards bulk cytoplasm.

Upon isomerization of retinal and the transition to the K state, the sidechain orientation when E108 is deprotonated is largely unaffected. In contrast, when E108 is protonated, both the χ_1 and χ_2 dihedral angles shift to a bimodal distribution, with a loss of sampling in the region of -70°, 60° (χ_1,χ_2) (Fig. 2.2 B). The corresponding rearrangement of the sampling landscape of the E108 sidechain in the protonated state indicates that the sidechain prefers to be in a completely linear orientation. This indicates that there is a degree of coupling existing between E108 and retinal that will be discussed below. In addition, we can further characterize the hydration of the sidechain of E108 by calculating its solvent-accessible surface area

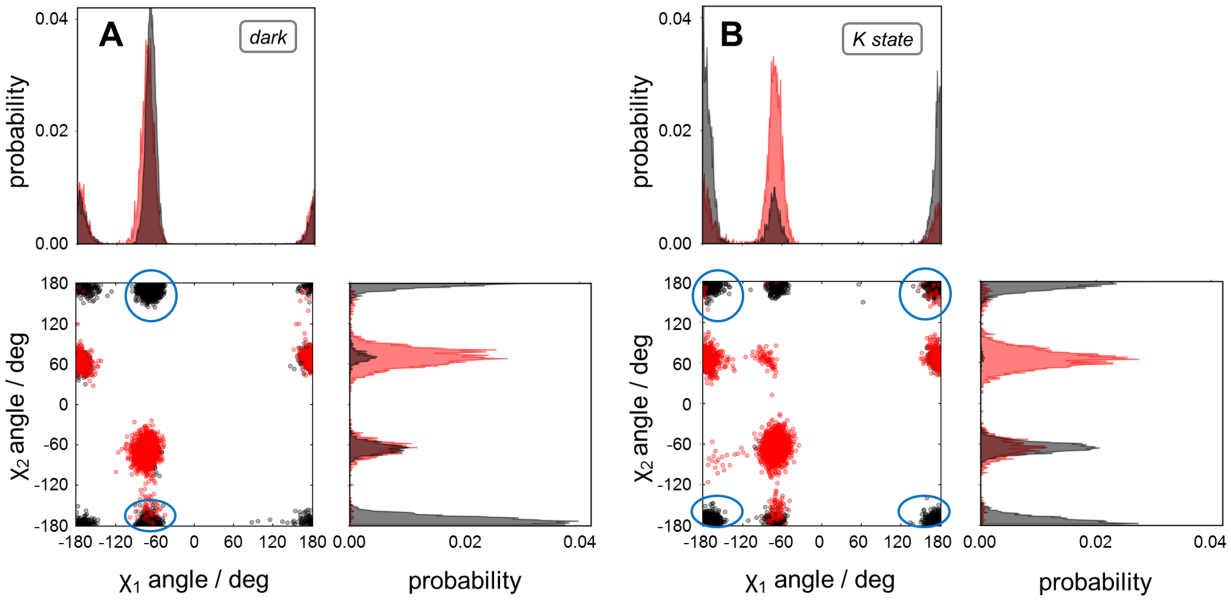


Fig. 2.2. Protonation state of E108 strongly influences its sidechain orientation. (A) Normalized histograms of the χ_1 and χ_2 dihedral angles of E108 in the dark state. When E108 is protonated (*black*), the dominant orientation of the χ_2 dihedral is $\pm 180^\circ$, leading to a fully-extended sidechain and blockage of the interior channel from the cytoplasmic end of PR to the retinal binding pocket. When E108 is deprotonated (*red*), the dominant χ_2 dihedral is -65°, i.e. allowing the sidechain to be oriented away from the retinal binding pocket and opening the cytoplasmic half-channel inside the protein for hydration from bulk solvent. Unlike χ_2, the dominant orientation of the χ_1 dihedral for E108 in both protonated (*black*) and deprotonated (*red*) states is -60°, leading to the conclusion that the χ_2 dihedral angle is the determining factor in sidechain orientation of E108 in the dark state. (B) Normalized histograms of the χ_1 and χ_2 dihedral angles of E108 for PR in the K state. No noticeable changes occur for the sidechain orientation of the deprotonated state of E108 for both χ_1 and χ_2 (*red*). However, the bimodal distribution of the χ_1 and χ_2 sidechain orientation of E108 between closed and open conformations when protonated (*black*) undergoes a shift – χ_2 becomes more evenly distributed whereas χ_1 shifts towards $\pm 180^\circ$, leading to transient access of bulk waters to the interior of the protein. Circles highlight the shift of the χ_1 and χ_2 sidechain orientation from the dark to the K state when E108 is protonated.

(SASA). It appears that a dihedral orientation of $\pm 180^\circ$ at either of the χ dihedrals is sufficient to prevent the interior channel of PR from becoming hydrated: in both the dark and K states, the SASA of the sidechain of E108 when protonated is less than 10 $Å^2$ (i.e., essentially dehydrated) (Fig. S2.3). In contrast, the SASA of E108 when it is deprotonated fluctuates from 20 to 80 $Å^2$ (Fig. S2.3).

When expanding the scope of analysis beyond E108, we see that the protonation state of E108 also has a marked effect on its interactions with a polar residue in helix B, S61 (Table 2.1). In the protonated state, the sidechain of E108 is much less likely to lie within a range where it is possible to form a hydrogen bond with S61. However, upon deprotonation, the distance between the sidechains of E108 and S61 becomes noticeably smaller (1-3 Å), to the point where a stable hydrogen bond is formed on the μs timescale (Figs. 2.3 A and S2.4). Inspection of the cytoplasmic half of PR shows that the hydrogen bond between E108 and S61 is essential to facilitate interior hydration of the cytoplasmic half of the protein (Fig. 2.3 B).

When E108 is protonated, the sidechain of the residue is oriented towards the PSB, occluding the proximal space of the binding pocket and preventing formation of a coordinated system of water molecules. In contrast, when E108 is deprotonated, the hydrogen bond formed with S61 provides a stable checkpoint for intercalated waters to diffuse into and out of the interior channel, connecting the retinal binding pocket with bulk cytoplasmic solvent. This result is consistent with our previous study[29] and is also conserved in other retinal proteins, most notably bR. In bR, the proton donor, D96, forms a hydrogen bond with T46. In crystal structures of bR in the dark and N states, this hydrogen bond rearranges to accommodate an influx of waters from the cytoplasm[12,52]. Specifically, D96 recruits waters from bulk cytoplasm, and in the progression towards the N photointermediate, the interaction of the T46 and D96 sidechains provides a key stabilizing force for a water molecule to diffuse between the two. The unique aspect of the conserved behavior between PR and bR with respect to this hydrogen bond is that S61 lies 1.5 turns further into the interior of PR on helix B: the sidechain of E108 is predisposed to have a χ_2 dihedral that favors an orientation that facilitates formation of the hydrogen bond (Fig. S4).

Table 2.1. E108 shows a noticeably higher propensity to form a hydrogen bond with S61 in the deprotonated state. Average number of hydrogen bonds between the carboxylic side chain of E108 and the hydroxyl group of S61. The number of hydrogen bonds is consistent when E108 is protonated. In contrast, when E108 is deprotonated, it increases the tendency for a hydrogen bond to form. This is consistent with our previous results[29], and shows similarities to the behavior of bR. In bR, there is a conserved threonine (T46) that forms a hydrogen bond with the proton donor, D96. S61 is one turn away on helix B in PR compared to T46 in bR, making formation of a hydrogen bond less likely. There is some conservation among other retinal proteins closely related to PR: in *E. sibiricum* sibiricum rhodopsin this residue is a threonine, and in xanthorhodopsin and the sodium pump KR2 it is a serine. So there may be some functional relevance to other microbial retinal proteins, but it is difficult to determine this with certainty.

system	dark state	K state
protonated	0.18±0.4	0.20±0.5
deprotonated	0.55±0.5	0.40±0.5

A marked contrast from bR in our simulations is that there is no large-scale conformational change in the cytoplasmic side of the protein. MD simulations of the different photointermediate states of bR, with the proton donor either protonated or deprotonated, showed that deprotonation of the proton donor led to a rapid opening of the cytoplasmic side of the protein[24] (i.e., within 50 ns). It was hypothesized that the proton donor serves as a "latch", with deprotonation triggering the release of this latch. A subsequent study on a triple mutant of bR in which the dark state conformation had an open cytoplasmic end showed that this fast response of the latching mechanism may play a vital role in preventing backflow of protons into bulk cytoplasm[25]. More recently, it was revealed that proton uptake in bR utilizes a proton uptake cluster

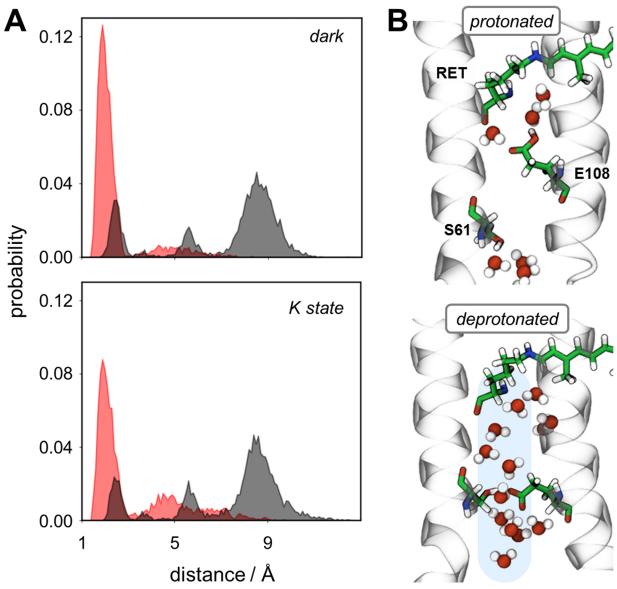


Fig. 2.3. E108 shows a noticeably higher propensity to form a hydrogen bond with S61 in the deprotonated state. (A) Histogram of the distance between the center of mass of the two terminal oxygens of E108 and the hydrogen of S61. When E108 is deprotonated (*red*), the sidechain of the acidic residue is in much closer proximity to the sidechain of S61, facilitating formation of a hydrogen bond. In contrast, when E108 is protonated (*black*), the distance between the sidechains of E108 and S61 is usually too large to form a direct hydrogen bond. This relationship stays consistent for both the dark (top) and K (bottom) states. (B) Representative snapshots of the cytoplasmic interior of PR when E108 is in the protonated (top) and deprotonated (bottom) states. Sticks: residues involved in the interaction of waters from bulk cytoplasm to the retinal binding pocket; spheres: water molecules; ribbons: helices B (left) and C (right) of PR.

to stabilize interactions of the proton donor with bulk cytoplasmic waters, leading to modulation of the later stages in the photocycle[26]. PR does not possess a corresponding aspartic acid residue, so it would make sense that the proton donor behaves differently from bR. Based on our simulations of PR in the dark and K states, it appears that the protonation-dependent orientation of E108 is sufficient to control passage of waters into the cytoplasmic half-channel of PR. We do not observe any major conformational fluctuations of the protein based on root-mean squared fluctuations (RMSF) (Fig. S2.5).

Long-range effects of the protonation of E108 on hydration in PR

Inspection of the entire protein reveals that the protonation state of E108 has long-range effects as well. In most retinal proteins, the cytoplasmic half-channel between the proton donor and the Schiff base undergoes subtle rearrangements immediately after photoactivation. It is only in the later stages of the photocycle (N and O states) that large-scale conformational changes in the cytoplasmic side of PR take place[14]. One example is the distance between E108 and the PSB. Since deprotonation of E108 results in a sidechain orientation away from the retinal binding pocket, the distance from the PSB is much larger than when E108 is protonated (Fig. 2.4 A). Regardless of photointermediate (dark or K state), this distance

fluctuates but remains fairly stable. However, when E108 is deprotonated, there is little change in the corresponding distance, regardless of state. It appears that the sidechain orientation of E108 is the major determining factor in internal hydration of the cytoplasmic half-channel: when E108 is protonated, 1-2 water molecules diffuse into the channel, whereas when E108 is deprotonated, we observe an influx of water (up to ten water molecules in the K state) (Fig. 2.4 B). Interestingly, this increase in hydration for the deprotonated form of E108 is not restricted to the cytoplasmic half of PR. The periplasmic half of PR remains largely dehydrated in the dark state and with E108 protonated; when E108 is deprotonated, 2-3 times as many waters diffuse into the protein (Fig. 2.4 C). After isomerization of retinal, the levels of hydration between the protonation states of E108 becomes indistinguishable. This behavior indicates that E108 may be subtly influencing the retinal binding pocket, as will be discussed in more detail below.

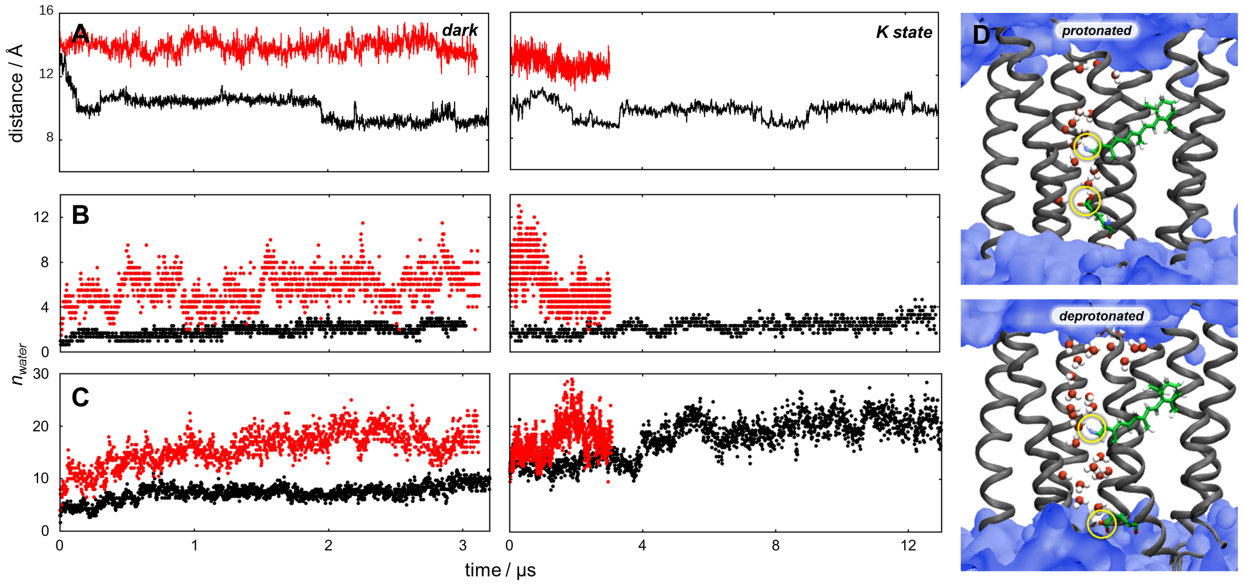


Fig. 2.4. Deprotonation of E108 has proximal and distal effects on hydration of BPR. (A) Distance between the PSB and carboxylate side chain of E108 in the deprotonated (*red*) versus the protonated (*black*) states in the dark (left) and K (right) states. Simulations with E108 in the deprotonated state have a much larger separation in the cytoplasmic half-channel, due to orientation of the E108 sidechain away from the retinal binding pocket. (B) Number of water molecules in the cytoplasmic half-channel of PR, defined as the region inside PR between the nitrogen of the Schiff base and the carboxylic oxygen of E108. The number of waters when E108 is deprotonated (*red*) steadily increases in the dark state (left), whereas the cytoplasmic half-channel stays dehydrated when E108 is protonated (*black*). In the K state (right), hydration fluctuates more but remains around four waters when E108 is deprotonated, while when E108 is protonated, the protein remains dehydrated. (C) Number of water molecules in the periplasmic (PP) half-channel of PR, defined as the region inside PR between the nitrogen of the Schiff base and the carboxylic carbon of E142. In the dark state (left), both titration states of E108 lead to an increase in internal hydration, with a more noticeable increase when E108 is deprotonated (red). In the K state (right), the number of waters fluctuates but remains at about twenty water molecules with E108 in the deprotonated state. When E108 is protonated (black), hydration increases after isomerization of retinal, stabilizing around twenty waters. (D) Representative snapshots of PR with E108 protonated (top) or deprotonated (bottom). Surface: bulk water; ribbons: backbone of PR; sticks: retinal; spheres: individual internal water molecules. Schiff base and E108 sidechains are highlighted for clarity.

If we look at the average water density of the interior of the protein for both protonation states of E108 in the dark and K states, a clearer picture of the mechanism of proton pumping emerges. When E108 is deprotonated, two stable half-channels of water, connected by retinal, form from the cytoplasmic to the periplasmic side of the protein (Fig. S2.6). This continuous water channel exists regardless of the dark or K state. In contrast, when E108 is protonated, only the periplasmic half-channel becomes hydrated, and only in the K state. If we analyze formation of these half-channels, identifying potential half-channels that exist at a given frame in our trajectories, we observe similar behavior. A significantly larger number of pathways exist in the cytoplasmic side of PR when E108 is deprotonated, while few pathways exist when E108 is protonated. In the periplasmic half of PR, a large number of pathways exist in both photointermediates when E108 is deprotonated, and when E108 is protonated, a transition from few to many pathways occurs from the dark to the K state (Fig. 2.5). The gating mechanism we have observed with E108 has been determined as a function of the ionization of E90 in channelrhodopsin with FTIR experiments and MD simulations[53]. To the best of our knowledge, this is the first time that such an effect has been elucidated for E108 in PR.

Deprotonation of the proton donor perturbs the retinal binding pocket in PR

All retinal proteins require a coordinated hydrogen-bonded network within the retinal binding pocket to stabilize the protonated Schiff base in the dark state. In bR, the closest homolog to PR, this network is formed by a pentameric arrangement of three water molecules with the D85 and D212 counterions[12]. The crystal structure of blue PR does not have a high enough resolution to identify any waters within the binding pocket[30], but the arrangement of the PSB and counterions is very similar to the retinal binding pocket of bR in the dark state. We observe that after diffusion of water into the interior of PR, a stable pentameric arrangement of waters forms, consistent with other microbial proton pumps[54,55] (Fig. 2.6 A). In addition, the protonation state of E108 does not perturb the formation of the pentameric water cluster. Upon all-*trans* to 13-*cis* isomerization of retinal to generate the K photointermediate, the hydrogen of the PSB rotates away from the aspartate counterions and is no longer available to form hydrogen bonds with the water cluster. The behavior of the retinal polyene chain after isomerization is consistent with the slight out-of-plane rotation around the C14-C15 bond that was recently observed with solid-state NMR in the dark to K state photoactivation of green PR[56] (Table 2.2). This leads to a subsequent shift in the waters and counterions; one water molecule remains to coordinate with D97 and D227, but the sidechains of the counterions now form direct hydrogen bonds with R94 (Fig. 2.6 B). This rearrangement within the retinal binding pocket is the initial event that leads to an influx of waters from the periplasmic space, as shown from our previous study of PR[29].

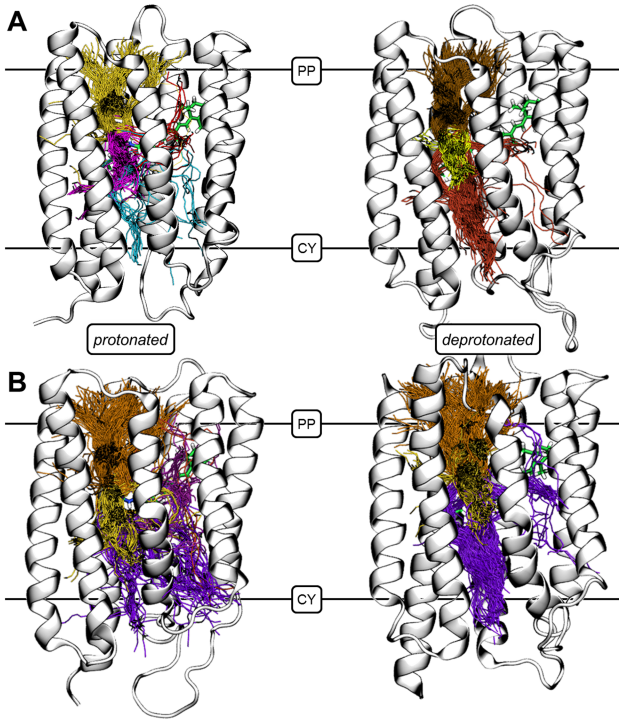


Fig. 2.5. Water pathways change as a function of protonation state of E108 and photoactivation of PR. (A) CAVER 3.0 was used to identify tunnels large enough to accommodate diffusion of water from outside PR to the retinal binding pocket in the dark state. Significantly more water-accessible pathways exist in the cytoplasmic half-channel of PR when E108 is deprotonated. Same-colored tunnels represent pathways from a particular cluster. Left: protonated E108; right: deprotonated E108; sticks: retinal. PP: periplasmic; CY: cytoplasmic. (B) Water-accessible tunnels in PR in the K state, identified with CAVER. Although the greatest number of tunnels are still found when E108 is deprotonated, the number of tunnels in both the periplasmic and cytoplasmic half-channels of PR increases when E108 is protonated. Coloring and representations are the same as in A.

In particular, of the two aspartic acid residues in the binding pocket (D97 and D227), D227 is the only one that directly forms a hydrogen bond with the PSB (Table 2.3). The closer interaction of D227 with the PSB compared to D97 is structurally consistent with the X-ray structure of blue PR[30]. In addition, this stable interaction under dynamic conditions helps validate earlier experimental investigations of the structure-function relationship of PR. One of the first studies to examine the PR photocycle with a D227N point mutant determined that D227 was responsible for selectivity of the retinal photoisomerization (predominantly 13-*cis*[57]). A more recent pump-probe and flash photolysis study showed that D227 is able to modulate the pK_a of the Schiff base, partially contributing to the abnormally high pK_a of the proton acceptor ($\sim$ 7.1-7.6[10,15]) with respect to that of bR (2.6)[58]. We observe this direct interaction only in the dark state; after isomerization of retinal, the PSB rotates away from D97 and D227 as mentioned above (Figs. S2.9 and S2.10). Deprotonation of E108 does not have any significant effect on this arrangement.

The residue at position 105 is the major determinant in the difference of spectral absorption between

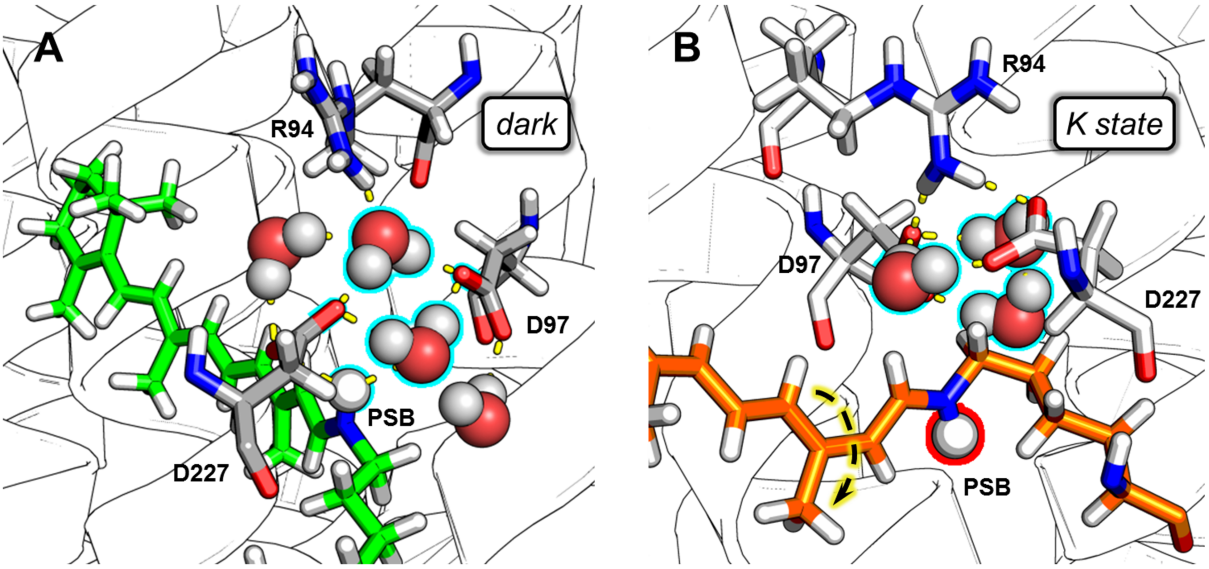


Fig. 2.6. Pentameric water cluster is disrupted upon photoactivation of PR. (A) Snapshot of the retinal binding pocket of PR with protonated E108 in the dark state. In the dark state of PR, a pentameric water cluster is formed between the Schiff base, the proton acceptor D97, the complex counterion D227, and 3 water molecules, consistent with other microbial proton pumps such as bR[12]. The presence of the water molecules stabilizes the PSB by creating a hydrogen-bonded network. Sticks: retinal (*green*) and important amino acid residues (*gray*) in the binding pocket; spheres: waters involved in the hydrogen-bonded network with the PSB and counterions; red: oxygen; white: hydrogen; blue: nitrogen; yellow dashes: hydrogen bonds; PSB: protonated Schiff base. The hydrogen on the PSB is rendered as a sphere. Waters and the hydrogen of the PSB are highlighted in cyan. (B) Snapshot of the retinal binding pocket of PR in the K state. After all-trans to 13-cis isomerization of retinal (dashed arrow), the PSB rotates away from the counterions, leading to a rearrangement of the water cluster. Although a water molecule remains coordinated between D97 and D227, the sidechains shift to create direct hydrogen bonds with R94. This shift propagates movements towards the periplasmic side of the protein, allowing an influx of bulk water into PR. Orange sticks: retinal. All other color schemes are the same as A, with the hydrogen of the PSB highlighted in red.

green and blue variants of PR: in blue PR it is a glutamine, whereas in green PR it is a leucine. The glutamine residue in green PR is able to alter the electrostatic environment of the Schiff base[19]. Earlier quantum mechanical (QM) calculations on a homology model of blue PR suggested that Q105 specifically influences the electronic environment around H15 of the retinal polyene chain[59], and subsequent Fourier transform infrared (FTIR) spectroscopy studies showed that Q105 interacts with the Schiff base in both the dark and K states[32]. In addition, a solid-state NMR study on a L105Q mutant of green PR showed that the presence of the glutamine sidechain disrupts the conjugated π system of the polyene chain, leading to a defect in the electronic environment[60]. Interestingly, we observe that the protonation state of E108 has a long-range effect on the interaction of Q105 with the retinal chromophore. In the dark state, the carbonyl oxygen in the sidechain of Q105 is too far from H14 to form a hydrogen-bonding interaction (i.e., $>$ 5 Å,

Fig. 2.7 A and S2.11). However, when E108 is protonated, the sidechain of Q105 remains $\sim$5 Å from H14 on the microsecond timescale (Fig. S2.12), close enough to influence the electrostatic environment of the retinal polyene chain. Upon photoisomerization of retinal to initiate the K state, the distance between Q105 and H14 decreases, to the point that a stable hydrogen bond could form between the two (Fig. 2.7 B). This distance has a bimodal distribution for both protonation states of E108; however, when E108 is protonated, it is more than twice as likely to form a hydrogen bond (i.e., < 3 Å) than when it is deprotonated.

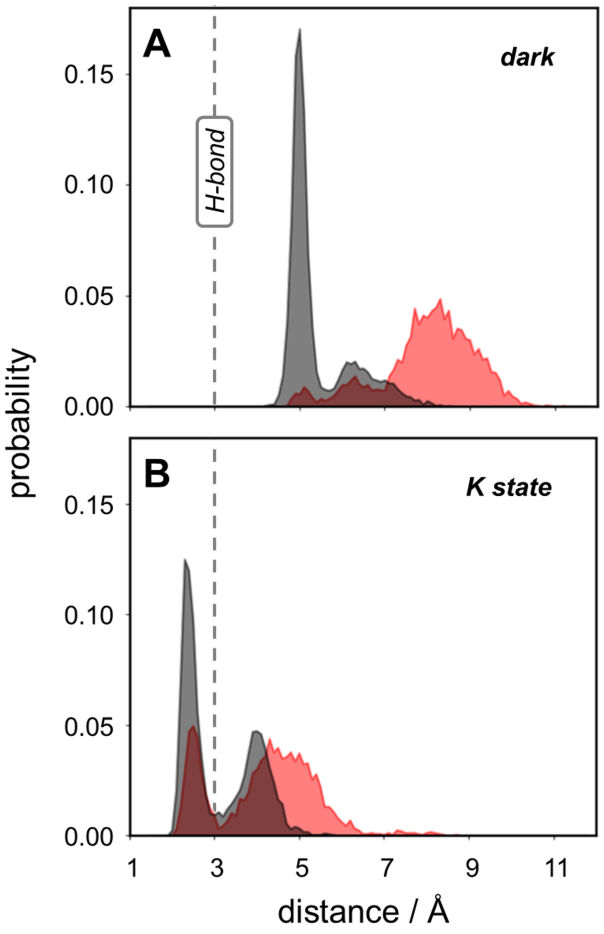


Fig. 2.7. Deprotonation of E108 destabilizes interactions of Q105 with retinal. (A) Normalized distance histograms of the carbonyl oxygen of the side chain of Q105 with H14 of retinal in the dark state with E108 either protonated (*black*) or deprotonated (*red*). A clear dependence exists in the interaction of Q105 with retinal based on the protonation state of E108. Deprotonation of E108 (and subsequent hydration of the cytoplasmic interior) leads to rearrangement of the retinal binding pocket in the vicinity of Q105. Kandori et al. hypothesized that the color-tuning mechanism of PR is based on spatial effects of residue 105 interacting with retinal, and that internal hydration could play a role in stabilizing interactions with retinal[61]. (B) Normalized distance histograms of the carbonyl oxygen of the side chain of Q105 with H14 of retinal in the K state with E108 either protonated (*black*) or deprotonated (*red*). Similar to the dark state, Q105 and retinal tend to be more separated when E108 is deprotonated. The all-trans to 13-cis isomerization of retinal positions Q105 significantly closer to H14.

To further quantify the relationship between E108 and Q105, we carried out an NbIT analysis. There is evidence for correlation (> 0.5) between E108 and Q105 in one of the three dark trajectories when E108 is protonated and in one of the two K state trajectories when E108 is deprotonated (Table S2.1). In addition, in all cases, the correlation between between the dynamics of Q105 and retinal is less when E108 is deprotonated, i.e. when the protein interior is fully hydrated. Finally, a definite correlation exists for the

distance between Q105 and H14 when E108 is protonated versus when E108 is deprotonated, indicating that some degree of cross talk exists with the proton donor and the color-tuning switch in blue PR.

Table 2.2. Out-of-plane rotation observed for C14-C15 dihedral in PR. In the crystal structure of BPR, the C14-C15 dihedral is 180.48°. Our simulations of the dark to K state transition show dynamic behavior within the range of the crystal structure, regardless of the protonation state of E108. We observe an interesting trend in the photoactivation process, where the C14-C15 dihedral undergoes a negative out-of-plane twist. This behavior is consistent with a recent solid-state NMR study on GPR[56], where it was observed that the C14-C15 dihedral shifted from 158 ± 2° in the dark state to 154 ± 2° in the K state.

system	dark state	K state
protonated	186.90° ± 12.46°	176.27° ± 12.33°
deprotonated	182.03° ± 11.84°	177.43° ± 11.80°

Table 2.3. Isomerization of retinal during photoactivation leads to key rearrangements of hydrogen-bonding networks with the binding pocket of PR.

system	intermediate	PSB to D97[a]		PSB to D227		ASP to water[b]
		$O_{\delta 1}$	$O_{\delta 2}$	$O_{\delta 1}$	$O_{\delta 2}$	
protonated	dark state	0	0.1%	11.9	22.3	1.3%
	K state	0	0	0	0	0.3
deprotonated	dark state	0.2	0.1	25.1	16.3	0.7
	K state	0	0	0	0	0.7

[a] Average occupancy (in percent) of hydrogen bonds between N atom of retinal Schiff base and side chain oxygen atoms of neighboring aspartic acids. D227 is the only counterion that is close enough to the Schiff base to directly form hydrogen bonds. However, isomerization of retinal effectively abolishes these direct interactions. D227 is believed to influence isomerization of retinal[57]. This effect might be due to proximity of D227 to the retinal Schiff base. [b] Average occupancy (in percent) of hydrogen bonds between aspartic acid residues in retinal pocket with water molecules.

To further quantify the relationship between E108 and Q105, we carried out an NbIT analysis. There is evidence for correlation (> 0.5) between E108 and Q105 in one of the three dark trajectories when E108 is protonated and in one of the two K state trajectories when E108 is deprotonated (Table S2.1). In addition, in all cases, the correlation between between the dynamics of Q105 and retinal is less when E108 is deprotonated, i.e. when the protein interior is fully hydrated. Finally, a definite correlation exists for the distance between Q105 and H14 when E108 is protonated versus when E108 is deprotonated, indicating that some degree of cross talk exists with the proton donor and the color-tuning switch in blue PR.

2.5 Conclusion

We have studied the effect of the protonation state of the proton donor on the hydration and color tuning of blue PR in the early stages of photoactivation. When E108 is protonated, its sidechain is oriented towards the retinal binding pocket, preventing internal hydration of the cytoplasmic half-channel. Upon deprotonation, the sidechain of E108 rotates outward, facilitating formation of a stable water channel from bulk cytoplasm to the retinal binding pocket, in essence acting as a gating mechanism. In addition, deprotonation of E108 has distal effects on PR. In the dark state, deprotonation of E108 leads to an increase in hydration of the periplasmic half-channel that is normally characteristic of the K state. In addition, deprotonation of the proton donor destabilizes interactions of the color-tuning switch, Q105, with the retinal polyene chain. Although a comprehensive picture of the proton transfers that occur during the photocycle of PR remain to be fully elucidated, Our study highlights the ability of MD simulations to aid in characterizing subtle details that lead to differences in retinal proteins that act as proton pumps.

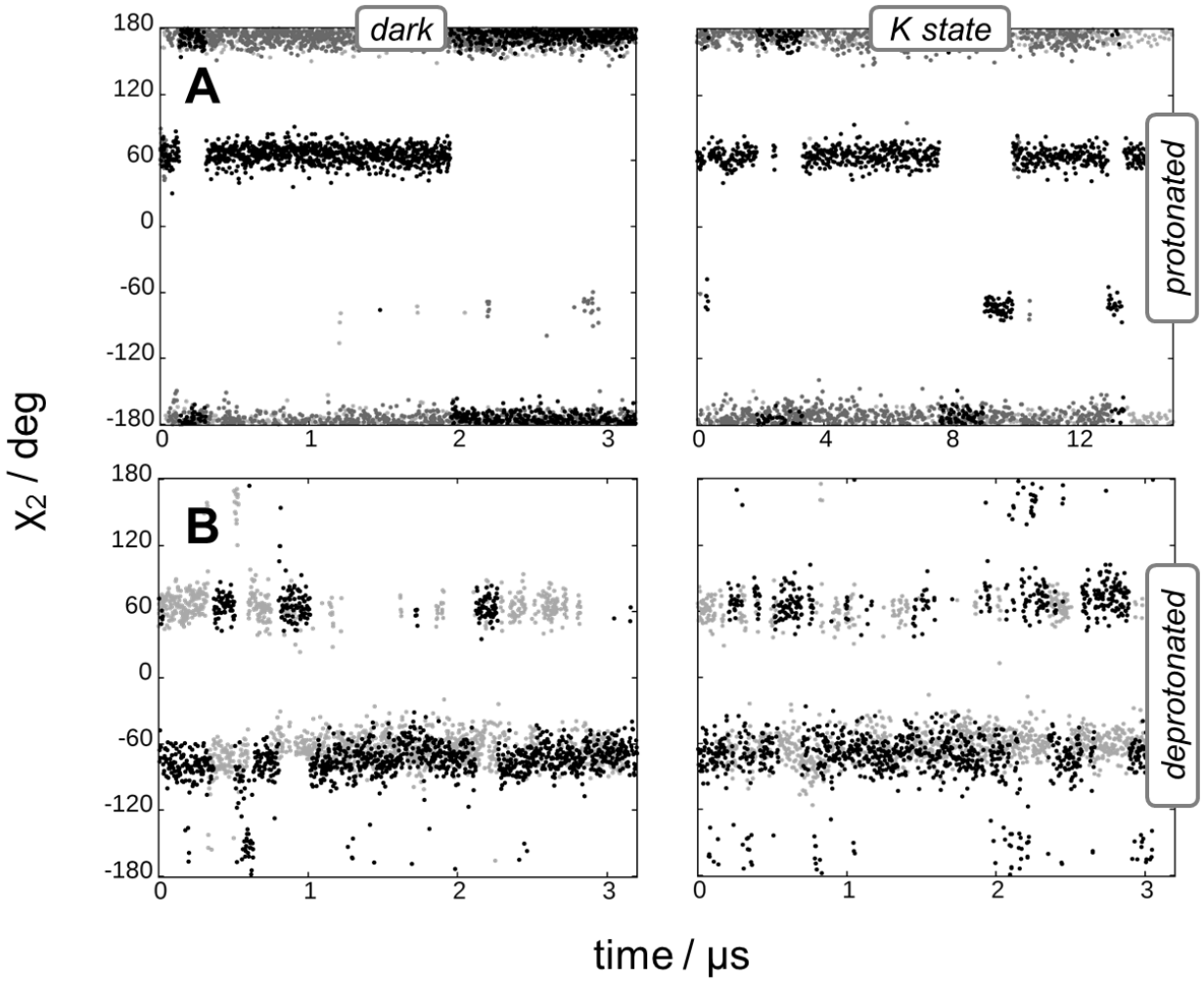


Figure S2.1. Each protonation state of E108 possesses a dominant χ_2 sidechain orientation. (A) The χ_2 dihedral angle of protonated E108 in the dark (left) and K (right) states remains at about $\pm 180^\circ$, orienting the side chain toward retinal. In , with transient orientation towards CY side, corresponding to dihedral angle of $\pm 60^\circ$. Black: run 1 for protonated E108; dark gray: run 2 for protonated E108; light gray: run 3 for protonated E108. (B) However, the χ_2 dihedral angle of deprotonated E108 in both the dark (left) and K (right) states is predominantly $\pm 60^\circ$, indicating a stable kink of the side chain towards bulk water from the cytoplasm. Black: run 1 for deprotonated E108; light gray: run 2 for deprotonated E108.

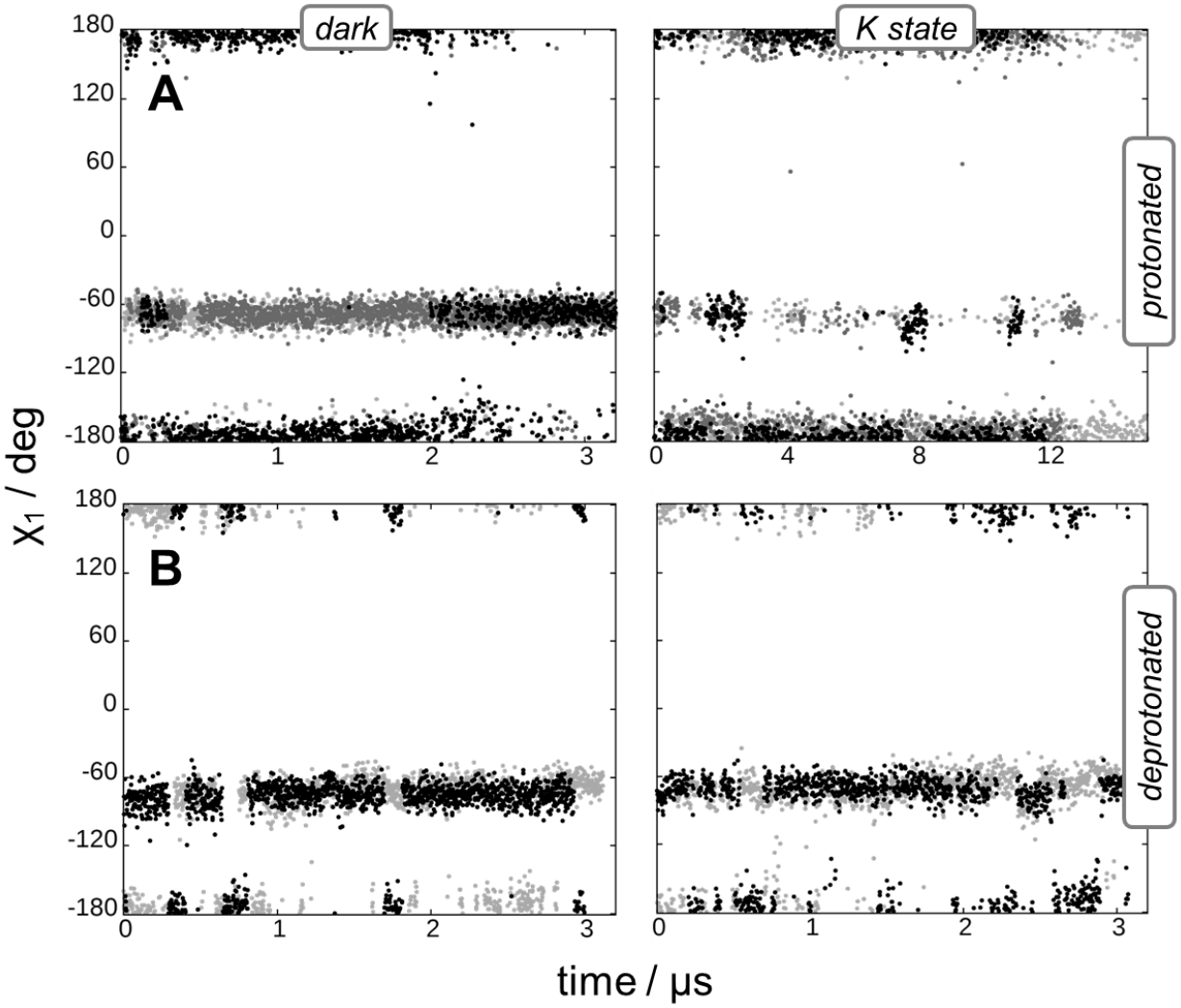


Figure S2.2. Each protonation state of E108 possesses a dominant χ_1 sidechain orientation. (A) χ_1 dihedral angle of protonated E108 in the dark (left) and K state (right) stays around $\pm 180^\circ$ that is associated with the orientation of the side chain toward the retinal, with transient orientation towards CY side, corresponding to dihedral angle of $\pm 65^\circ$. Black: run 1 for protonated E108; dark gray: run 2 for protonated E108; light gray: run 3 for protonated E108. (B) However, χ_1 dihedral angle of deprotonated E108 in both dark (left) and K state (right) is -65° for the majority of the simulations, orienting the sidechain towards bulk water in the cytoplasm. Black: run 1 for deprotonated E108; light gray: run 2 for deprotonated E108.

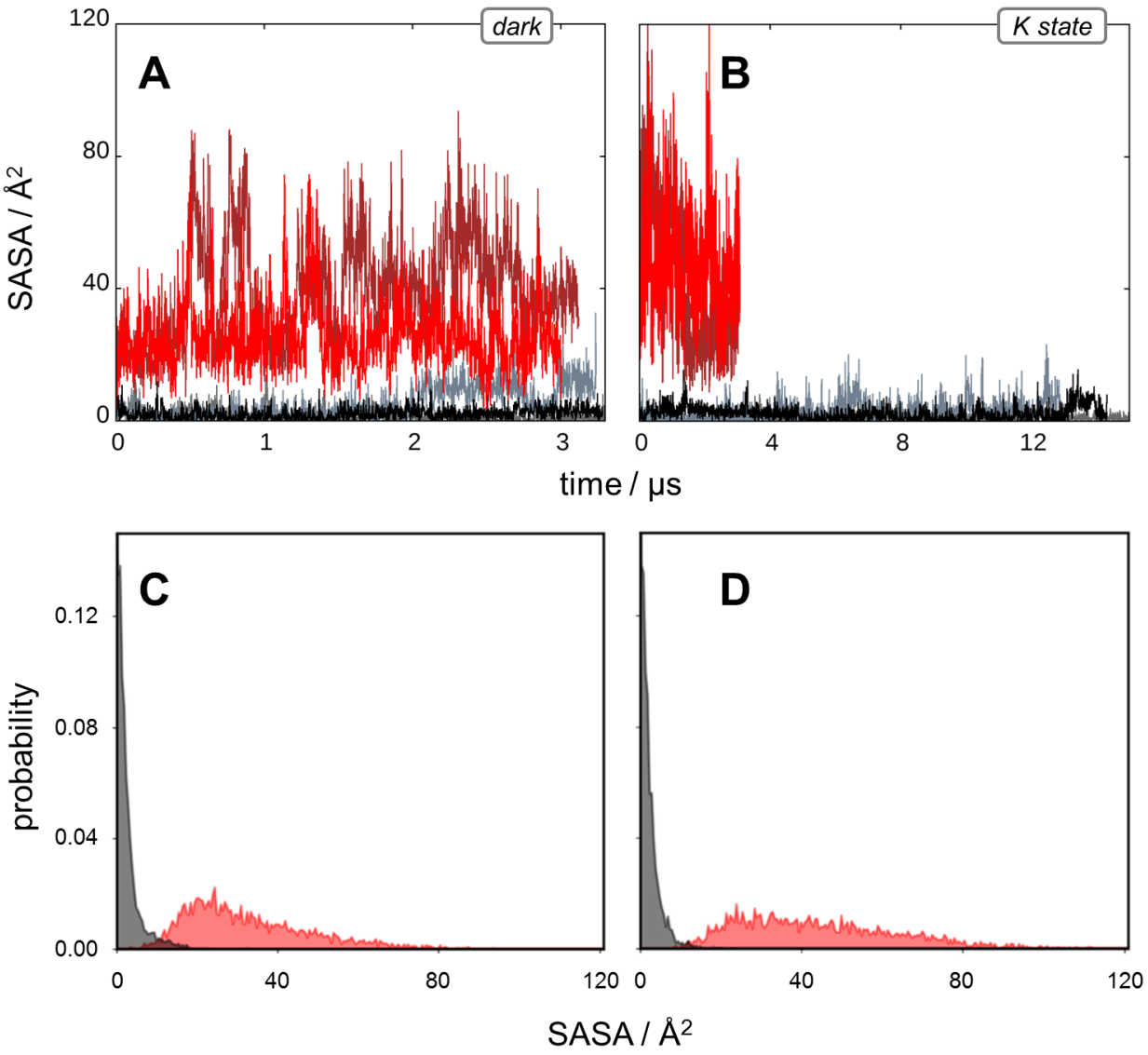


Figure S2.3. E108 is significantly more hydrated in the deprotonated state. (A) Solvent accessible surface area (SASA) of E108 sidechain in the dark state of PR. SASA remains stable on the microsecond timescale, regardless of protonation state of E108. However, when E108 is deprotonated, the sidechain becomes more exposed to solvation by bulk waters from the cytoplasm. Dark red: run 1 for deprotonated E108; light red: run 2 for deprotonated E108; light gray: run 1 for protonated E108; dark gray: run 2 for protonated E108; black: run 3 for protonated E108. (B) SASA of E108 in the K state of PR. After all-*trans* to 13-*cis* isomerization of retinal, the SASA of E108 in the deprotonated state increases, followed by fluctuations after one microsecond. In contrast, the SASA of protonated E108 remains the same. Color scheme is same as A. (C) Normalized histogram of the solvent accessible surface area (SASA) of the E108 sidechain in the dark state of PR. There is a clear dependence of SASA on the protonation state of E108: when E108 is deprotonated, the sidechain becomes more exposed to solvation by bulk waters from the cytoplasm. Red: deprotonated E108; black: protonated E108. (D) Normalized histogram of the SASA of E108 in the K state of PR. After all-*trans* to 13-*cis* isomerization of retinal, the SASA of E108 in the deprotonated state sees an increase characterized by larger fluctuations. In contrast, the SASA of protonated E108 remains the same. Color scheme is identical to part C. Taken altogether, this titration-dependent distribution of sidechain orientations allows E108 to act as a gate, providing access to the protein interior for hydration from bulk solvent in the deprotonated state of E108.

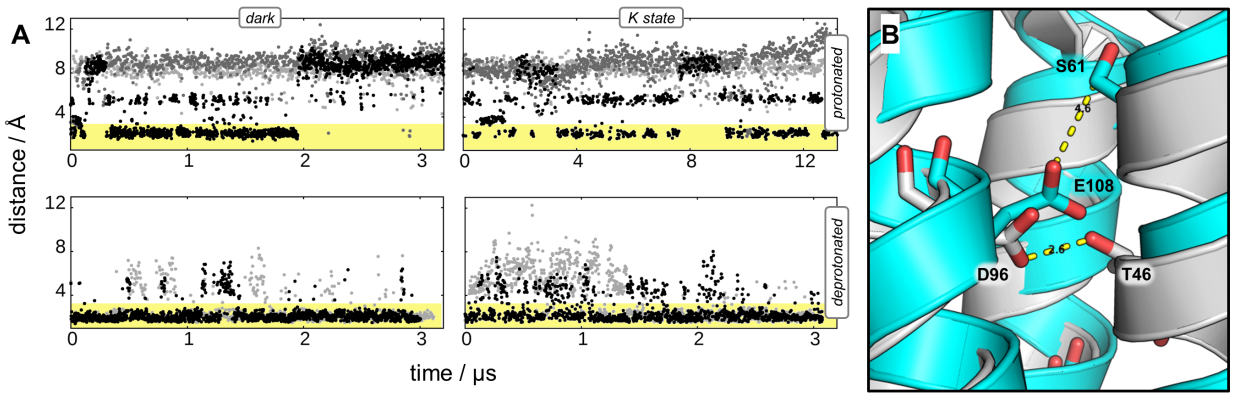


Figure S2.4. Deprotonation of E108 stabilizes hydrogen-bonding interactions with S61. (A) Time-dependent distances between the center of mass (COM) of the carboxyl oxygens on E108 and the hydroxyl hydrogen of S61 when E108 is protonated (top) or deprotonated (bottom) in the dark (left) and K (right) states. There are few differences in the interaction of E108 and S61 in the dark versus K states, but a significant difference in the interaction when E108 is protonated versus deprotonated. When E108 is protonated, the majority of the time it lies too far from S61 to form a hydrogen bond. However, when E108 is deprotonated, this interaction changes, with most interactions lying within the threshold for formation of a hydrogen bond (*yellow highlights*). (B) Superposition of crystal structures from blue PR (cyan, PDB 4JQ6) and bacteriorhodopsin (grey, PDB 1C3W). In bR, the proton donor, D96, and T46 in helix B lie directly across from each other, easily forming a hydrogen bond. In contrast, S61 in blue PR lies 1.5 turns above the location of T46 in bR, outside the range of a typical hydrogen bond (4.6 Å, whereas a hydrogen bond is normally formed >3 Å).

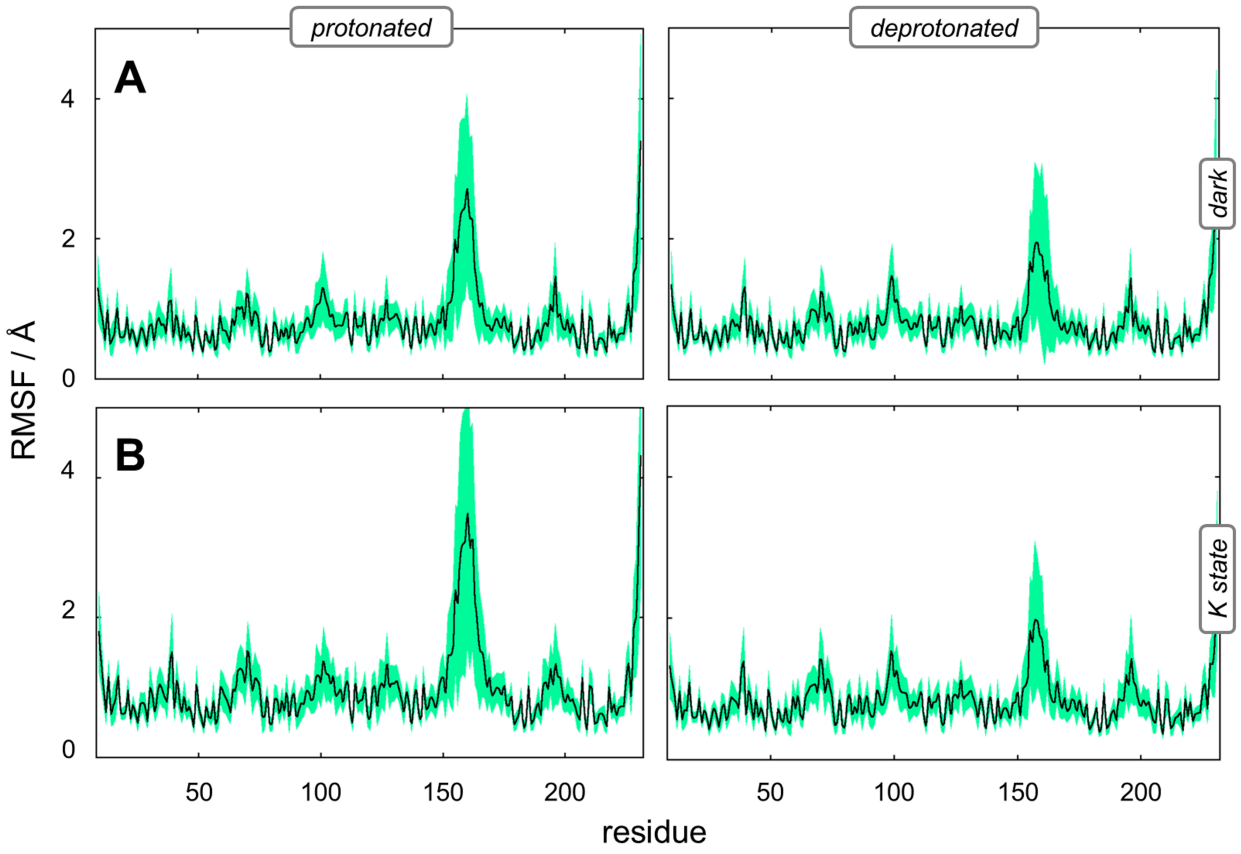


Figure S2.5. Internal hydration does not perturb the tertiary fold of PR. (A) Average root mean square fluctuation (RMSF) of PR residues in dark state when E108 is protonated (left) and deprotonated (right). Peaks in the RMSF represent loops in PR, troughs represent transmembrane helices. Note the largest fluctuation in the RMSF is for the cytoplasmic EF loop that is the location of the A178R mutation that leads to color-tuning in PR and may be directly implicated in function. Shaded region (*green*) represents the standard error. (B) Corresponding plots for the K state, respectively.

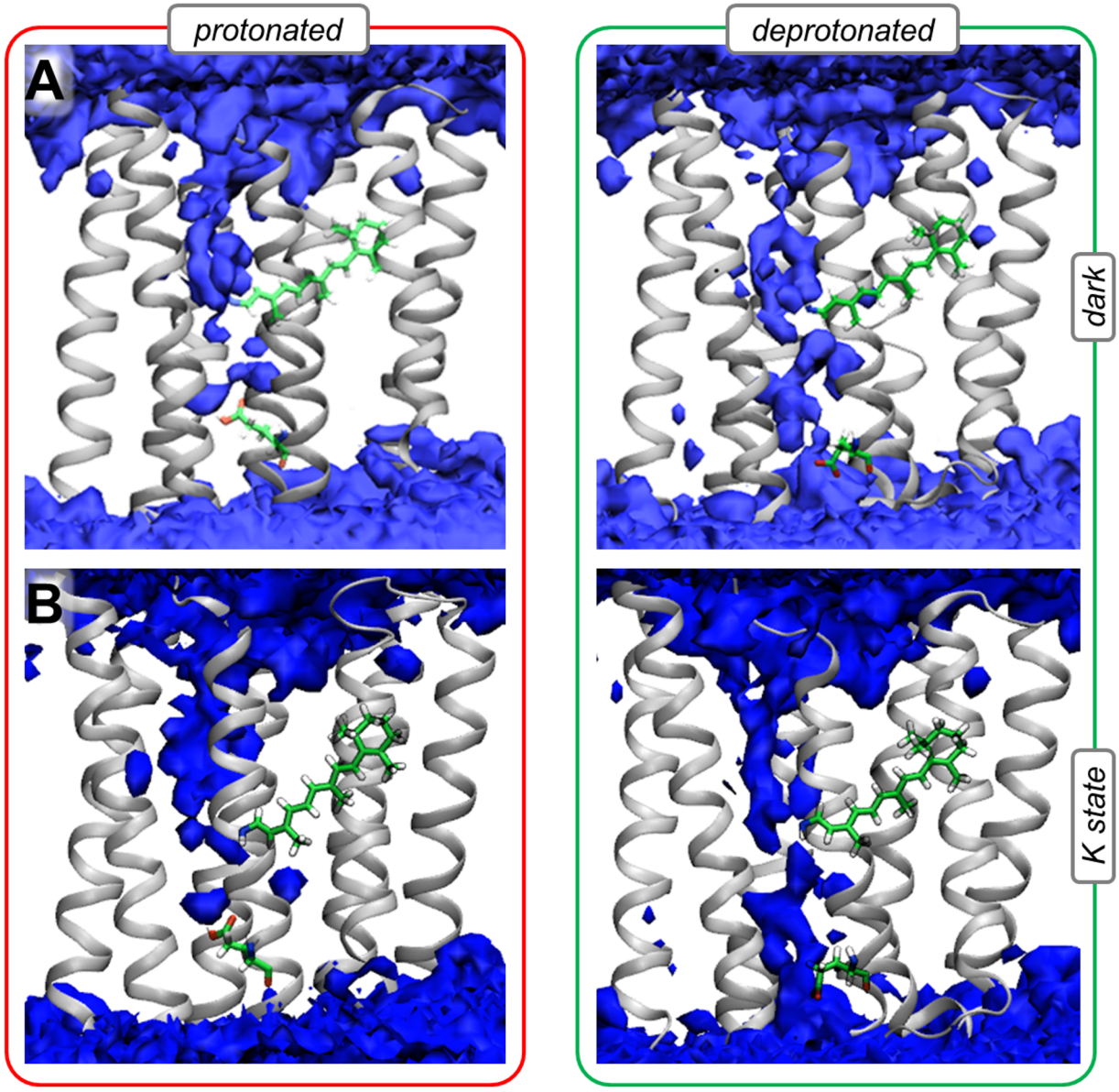


Figure S2.6. Deprotonation of E108 leads to formation of stable interior water channel along the entire length of PR. (A) Average water density inside PR in the dark state with E108 protonated (left) and deprotonated (right) Lipid bilayer and ions are not shown for clarity. Surface: water density; ribbons: TM helices of PR; sticks: retinal and E108. (B) Average water density inside PR in the K state with E108 protonated (left) and deprotonated (right). Rendering and color scheme is same as A.

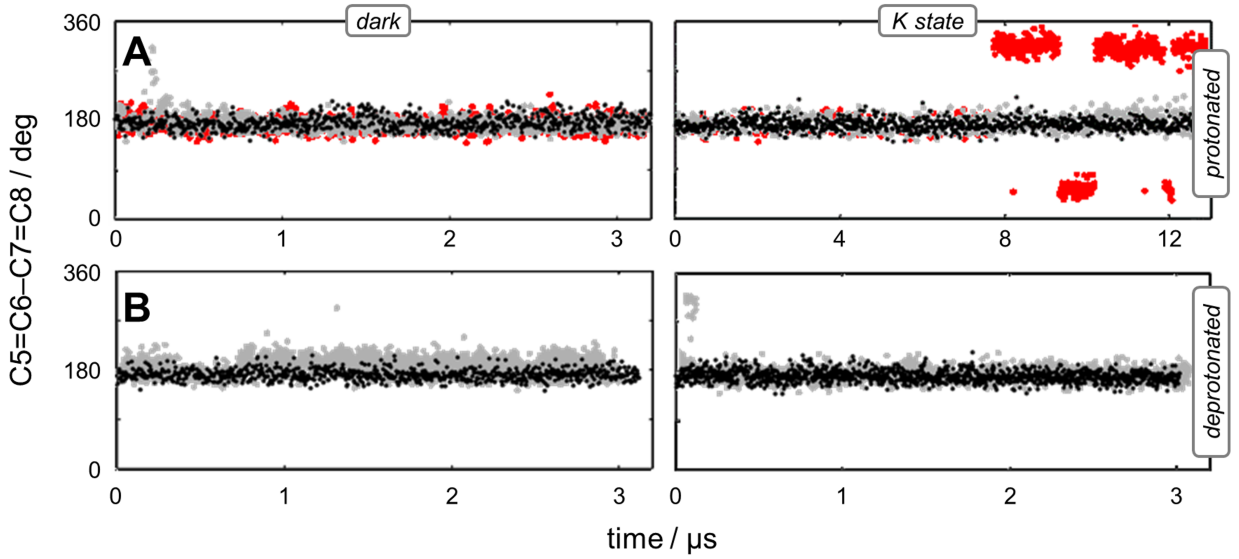


Figure S2.7. Protonation of E108 has little direct effect on dynamics of retinal in PR. (A) C5=C6-C7=C8 dihedral of retinal with E108 protonated. In the dark state (left) the C6-C7 dihedral of the β-ionone ring remains planar (i.e., 180°). After photoisomerization of retinal, the C6-C7 dihedral remains planar in 2 of the 3 simulations (right). However, in the third trajectory, the ring fluctuates between $\pm$60° deg. This fluctuation is consistent with the ability for the orientation of the β-ionone ring to modulate the pKa of the protonated Schiff base, as with rhodopsin[39]. (B) C5=C6-C7=C8 dihedral of retinal with E108 deprotonated. In both the dark (left) and K (right) states, the β-ionone ring remains planar.

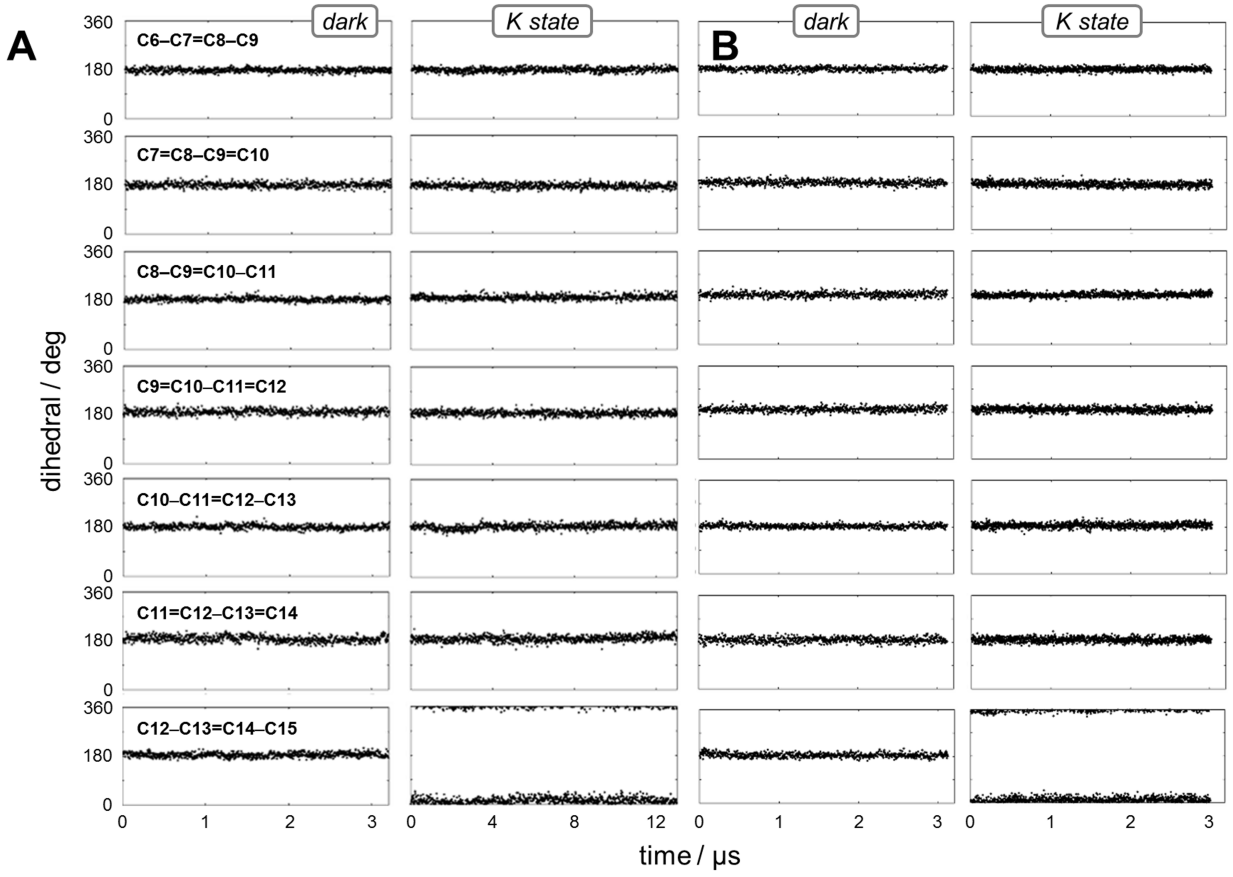


Figure S2.8. Dihedral angles along the polyene chain of retinal remain largely unchanged regardless of protonation state of E108. (A) Time-dependent dihedral angles along the polyene chain of retinal when E108 is protonated in the dark (left) and K (right) states. The dihedral angles remain stable over the μs timescale, with the only shift occurring in the C12-C13=C14-C15 dihedral angle after the all-*trans* to 13-*cis* isomerization of retinal. (B) Time-dependent dihedral angles along the polyene chain of retinal when E108 is deprotonated in the dark (left) and K (right) states. The behavior of retinal is very similar to that when E108 is protonated.

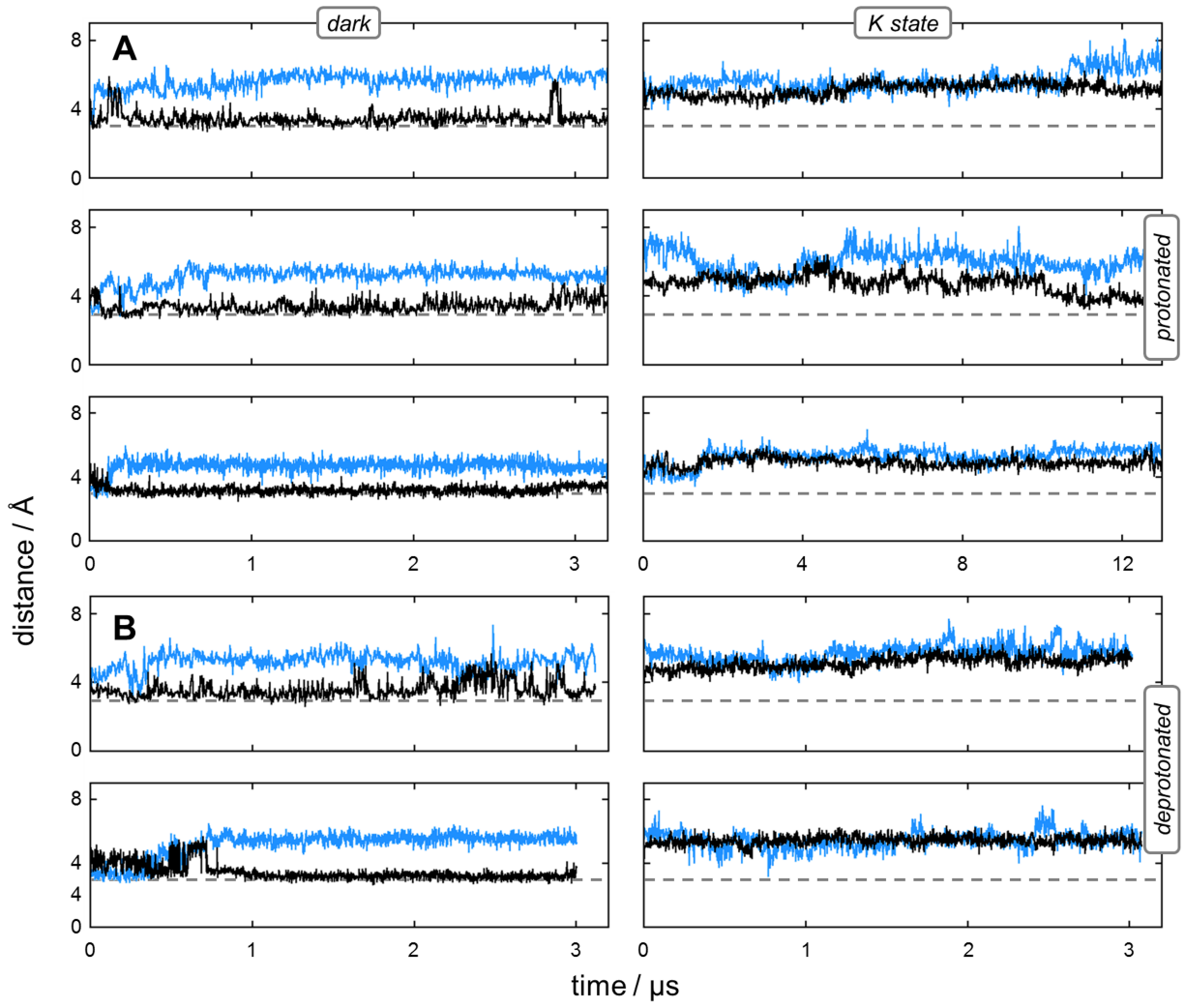


Figure S2.9. D227 stabilizes the retinal Schiff base in the dark state when E108 is protonated. (A) Left: distance between N atom of the retinal Schiff base and the center of mass of $O_{\delta 1}$ and $O_{\delta 2}$ atoms of neighboring D227 (*black*) and D97 (*blue*) residues shows that only D227 directly interacts with the Schiff base in the dark state. Right: after all-*trans* to 13-*cis* isomerization of retinal to the K state, the direct interaction between D227 and the protonated Schiff base is abolished. (B) Distance between the N atom of the retinal Schiff base and the COM of the carboxyl oxygens of D227 and D97 when E108 is deprotonated. The interactions between the retinal Schiff base and D227 and D97 are very similar to the interactions observed when E108 is protonated. Coloring scheme is the same as in part A.

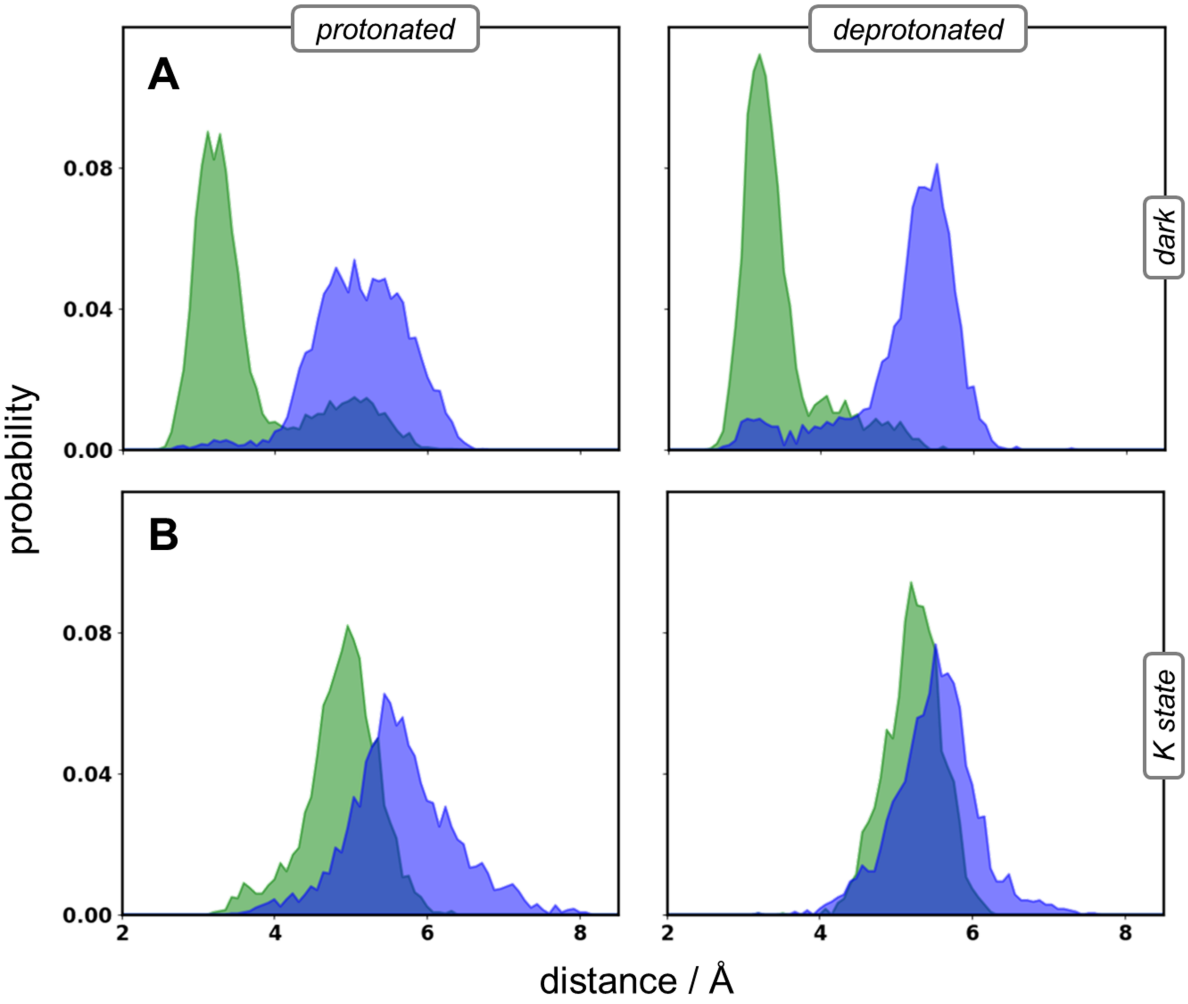


Figure S2.10. D227 interacts with the retinal Schiff base only in the dark state, regardless of the protonation state of E108. (A) Distribution of the distance between the protonated Schiff base (PSB) and the carboxylic oxygens of D227 (*green*) and D97 (*blue*) when E108 is protonated (left) or deprotonated (right) in the dark state of PR. Only D227 directly interacts with the PSB (i.e., <3 Å). (B) Distribution of the distance between the protonated Schiff base (PSB) and the carboxylic oxygens of D227 (green) and D97 (blue) when E108 is protonated (left) or deprotonated (right) in the K state of PR. Isomerization of retinal rotates the Schiff base away from the counterions.

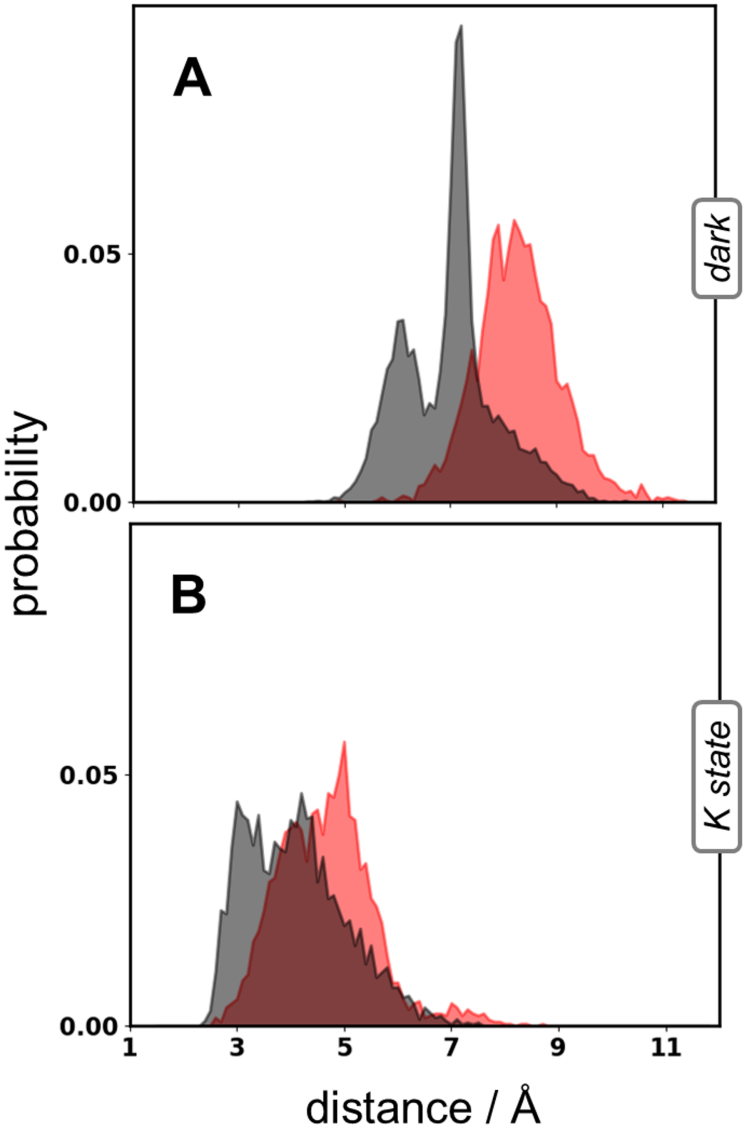


Figure S2.11. Deprotonation of E108 destabilizes interactions of Q105 with retinal. (A) Normalized distance histograms of N atom of Q105 side chain with hydrogen of C14 in retinal in dark state with protonated (*black*) and deprotonated (*red*) E108. (B) Normalized distance histograms of N atom of Q105 side chain with the hydrogen of C14 in the K state with protonated (*black*) and deprotonated (*red*) E108. Similar to the dark state, Q105 and retinal are more separated when E108 is deprotonated. However, due to the all-*trans* to 13-*cis* isomerization of retinal, Q105 is closer to retinal in both protonation states.

Interactions of retinal with the proton donor and color-tuning switch in PR

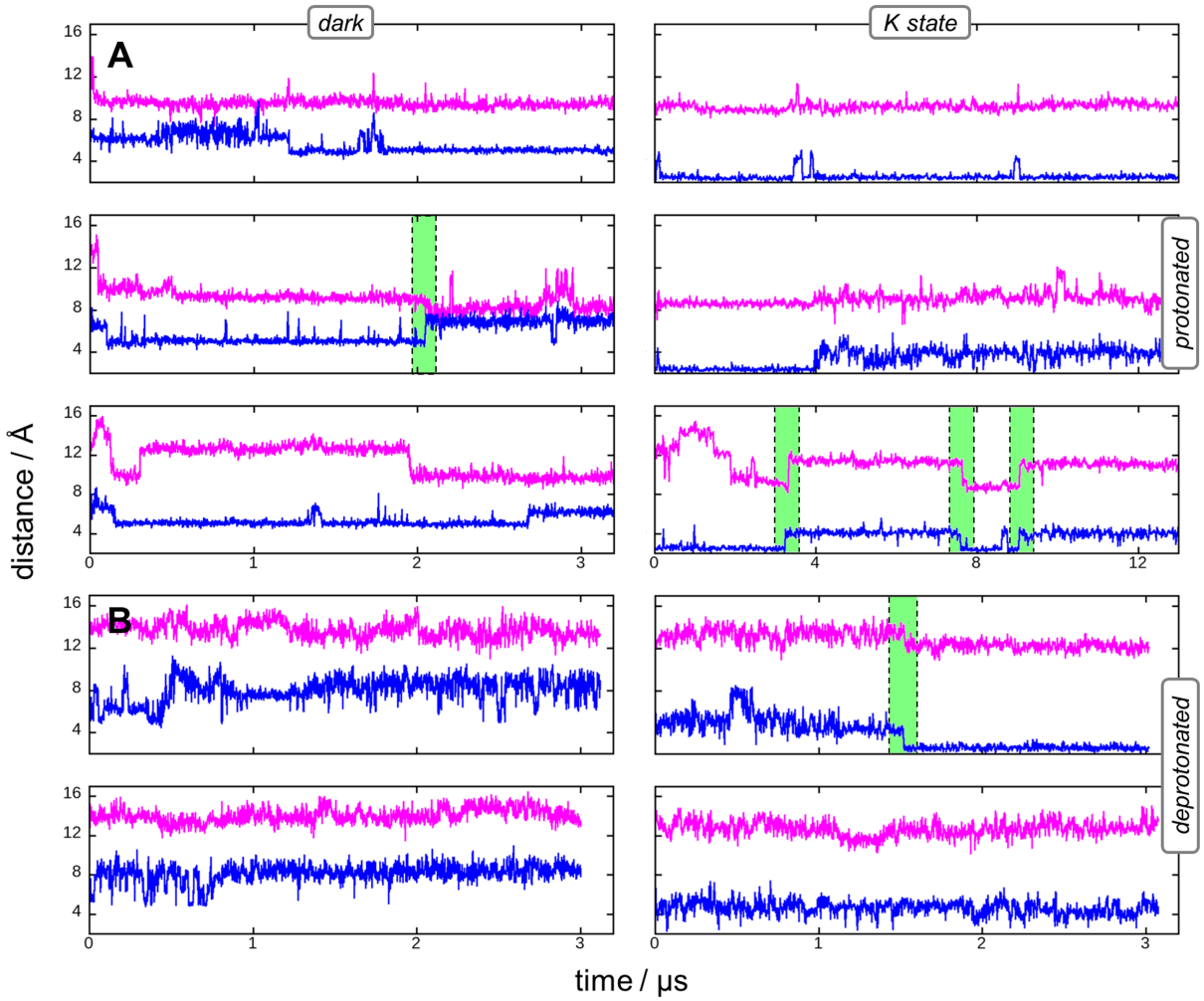


Figure S2.12. Interactions of retinal with the proton donor and color-tuning switch in PR. (A) In two out of three trajectories, the distance between the O atom of the Q105 side chain and the H atom bonded to the C14 atom of retinal (*blue*) is correlated with the distance between the protonated Schiff base and the side chain of E108 (*pink*) when it is protonated. Left: dark state of PR; right: K state of PR. (B) When E108 is deprotonated, there is almost no apparent correlation between the E108-PSB and Q105-PSB distances. Color scheme is the same as A.

Table S2.1. NbIT analysis confirms the correlation between dynamics of retinal and Q105 when E108 is protonated. Generalized correlation coefficient can have any value between 0 and 1 corresponding to the lack of correlation between dynamics of two species (atoms, residues, etc.) and full correlation, respectively[51]. The correlation between coordinates of H atom of C14 in retinal and terminal O and N atoms of Q105 is generally more correlated when E108 is protonated in dark and K state. Interestingly, in a copy of the dark state simulations of PR with protonated E108 (bold), the coordinates of side chain of Q105 is highly correlated to the OE2 atom of the side chain of E108 although they are located far from each other. Q105 is generally more correlated when E108 is protonated in dark and K state. Interestingly, in a copy of the dark state simulations of PR with protonated E108 (bold), the coordinates of side chain of Q105 is highly correlated to the OE2 atom of the side chain of E108 although they are located far from each other.

system	intermediate	Ret_{H141}–$Q105_{NE2}$	Ret_{H141}–Q105 OE1	Ret_{NZ}–$E108_{OE2}$	$Q105_{NE2}$–$E108_{OE2}$	$Q105_{OE1}$–$E108_{OE2}$
protonated	dark state	0.40	0.37	0.19	0.24	0.17
		0.49	0.47	0.42	**0.62**	**0.61**
		0.31	0.40	0.21	0.25	0.32
	K state	0.59	0.68	0.41	0.41	0.41
		0.52	0.52	0.30	0.30	0.24
		0.51	0.62	0.33	0.40	0.37
deprotonated	dark state	0.35	0.35	0.31	0.43	0.42
		0.30	0.21	0.27	0.25	0.22
	K state	0.39	0.49	0.36	0.56	0.56
		0.28	0.30	0.23	0.28	0.32

References

1. Béjà, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science (New York, N.Y.)* **289**, 1902–1906 (2000).

2. Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786–789 (2001).

3. Atamna-Ismaeel, N. *et al.* Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *The ISME Journal* **2**, 656–662 (2008). URL https://www.nature.com/articles/ismej200827.

4. Petrovskaya, L. *et al.* Predicted bacteriorhodopsin from Exiguobacterium sibiricum is a functional proton pump. *FEBS Letters* **584**, 4193–4196 (2010). URL http://onlinelibrary.wiley.com/doi/10.1016/j.febslet.2010.09.005/abstract.

5. Yutin, N. & Koonin, E. V. Proteorhodopsin genes in giant viruses. *Biology Direct* **7**, 34 (2012). URL https://doi.org/10.1186/1745-6150-7-34.

6. Slamovits, C. H., Okamoto, N., Burri, L., James, E. R. & Keeling, P. J. A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nature Communications* **2**, 183 (2011). URL https://www.nature.com/articles/ncomms1188.

7. Gómez-Consarnau, L. *et al.* Proteorhodopsin light-enhanced growth linked to vitamin-B1 acquisition in marine Flavobacteria. *The ISME Journal* **10**, 1102–1112 (2016). URL https://www.nature.com/ismej/journal/v10/n5/full/ismej2015196a.html.

8. Marchetti, A., Catlett, D., Hopkinson, B. M., Ellis, K. & Cassar, N. Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *The ISME Journal* **9**, 2745–2748 (2015). URL https://www.nature.com/articles/ismej201574.

9. Walter, J. M., Greenfield, D., Bustamante, C. & Liphardt, J. Light-powering Escherichia coli with proteorhodopsin. *Proceedings of the National Academy of Sciences* **104**, 2408–2412 (2007). URL http://www.pnas.org/content/104/7/2408.

10. Dioumaev, A. K. *et al.* Proton Transfers in the Photochemical Reaction Cycle of Proteorhodopsin. *Biochemistry* **41**, 5348–5358 (2002). URL https://doi.org/10.1021/bi025563x.

11. Váró, G., Brown, L. S., Lakatos, M. & Lanyi, J. K. Characterization of the Photochemical Reaction Cycle of Proteorhodopsin. *Biophysical Journal* **84**, 1202–1207 (2003). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1302695/.

12. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution 11edited by D. C. Rees. *Journal of Molecular Biology* **291**, 899–911 (1999). URL http://www.sciencedirect.com/science/article/pii/S0022283699930279.

13. Facciotti, M. T. *et al.* Structure of an Early Intermediate in the M-State Phase of the Bacteriorhodopsin Photocycle. *Biophysical Journal* **81**, 3442–3455 (2001). URL http://www.sciencedirect.com/science/article/pii/S0006349501759760.

14. Andersson, M. *et al.* Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17**, 1265–1275 (2009). URL /structure/abstract/S0969-2126(09)00290-1.

15. Hempelmann, F. *et al.* His75-asp97 Cluster in Green Proteorhodopsin. *Journal of the American Chemical Society* **133**, 4645–4654 (2011). URL http://dx.doi.org/10.1021/ja111116a.

16. 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* **70**, 473–481 (1996). URL http://www.sciencedirect.com/science/article/pii/S0006349596795917.

17. Friedrich, T. *et al.* Proteorhodopsin is a Light-driven Proton Pump with Variable Vectoriality. *Journal of Molecular Biology* **321**, 821–838 (2002). URL http://www.sciencedirect.com/science/article/pii/S0022283602006964.

18. Spassov, V. Z., Luecke, H., Gerwert, K. & Bashford, D. pKa calculations suggest storage of an excess proton in a hydrogen-bonded water network in bacteriorhodopsin11edited by G. von Heijne. *Journal of Molecular Biology* **312**, 203–219 (2001). URL http://www.sciencedirect.com/science/article/pii/S0022283601949022.

19. 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* **278**, 33985–33991 (2003). URL http://www.jbc.org/content/278/36/33985.

20. Freier, E., Wolf, S. & Gerwert, K. Proton transfer via a transient linear water-molecule chain in a membrane protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11435–11439 (2011). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136293/.

21. Váró, G. & Lanyi, J. K. Distortions in the photocycle of bacteriorhodopsin at moderate dehydration. *Biophysical Journal* **59**, 313–322 (1991). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1281148/.

22. Cao, Y. *et al.* Water is required for proton transfer from aspartate-96 to the bacteriorhodopsin Schiff base. *Biochemistry* **30**, 10972–10979 (1991). URL http://dx.doi.org/10.1021/bi00109a023.

23. Brown, L. S., Váró, G., Needleman, R. & Lanyi, J. K. Functional significance of a protein conformation change at the cytoplasmic end of helix F during the bacteriorhodopsin photocycle. *Biophysical Journal* **69**, 2103–2111 (1995). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1236444/.

24. Wang, T. *et al.* Deprotonation of D96 in Bacteriorhodopsin Opens the Proton Uptake Pathway. *Structure* **21**, 290–297 (2013). URL http://www.sciencedirect.com/science/article/pii/S0969212613000130.

25. Wang, T. *et al.* Stable Closure of the Cytoplasmic Half-Channel Is Required for Efficient Proton Transport at Physiological Membrane Potentials in the Bacteriorhodopsin Catalytic Cycle. *Biochemistry* **53**, 2380–2390 (2014). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4004217/.

26. Lorenz-Fonfria, V. A., Saita, M., Lazarova, T., Schlesinger, R. & Heberle, J. pH-sensitive vibrational probe reveals a cytoplasmic protonated cluster in bacteriorhodopsin. *Proceedings of the National Academy of Sciences* 201707993 (2017). URL http://www.pnas.org/content/early/2017/11/30/1707993114.

27. Gushchin, I. *et al.* Structural insights into the proton pumping by unusual proteorhodopsin from nonmarine bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12631–12636 (2013). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3732926/.

28. Harris, A. *et al.* A new group of eubacterial light-driven retinal-binding proton pumps with an unusual cytoplasmic proton donor. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 1518–1529 (2015). URL http://www.sciencedirect.com/science/article/pii/S0005272815001590.

29. Feng, J. & Mertz, B. Proteorhodopsin Activation Is Modulated by Dynamic Changes in Internal Hydration. *Biochemistry* **54**, 7132–7141 (2015). URL http://dx.doi.org/10.1021/acs.biochem.5b00932.

30. Ran, T. *et al.* Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta Crystallographica Section D: Biological Crystallography* **69**, 1965–1980 (2013). URL http://scripts.iucr.org/cgi-bin/paper?be5237.

31. Reckel, S. *et al.* Solution NMR Structure of Proteorhodopsin. *Angewandte Chemie International Edition* **50**, 11942–11946 (2011). URL http://onlinelibrary.wiley.com/doi/10.1002/anie.201105648/abstract.

32. Amsden, J. J. *et al.* Different Structural Changes Occur in Blue- and Green-Proteorhodopsins during the Primary Photoreaction. *Biochemistry* **47**, 11490–11498 (2008). URL https://doi.org/10.1021/bi800945t.

33. Dowhan, W. Molecular Basis for Membrane Phospholipid Diversity: Why Are There So Many Lipids? *Annual Review of Biochemistry* **66**, 199–232 (1997). URL https://doi.org/10.1146/annurev.biochem.66.1.199.

34. Jo, S., Kim, T. & Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLOS ONE* **2**, e880 (2007). URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000880.

35. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophysical Journal* **97**, 50–58 (2009). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2711372/.

36. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **29**, 1859–1865 (2008). URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20945.

37. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **12**, 405–413 (2016). URL https://doi.org/10.1021/acs.jctc.5b00935.

38. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **31**, 671–690 (2010). URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.21367/abstract.

39. Zhu, S., Brown, M. F. & Feller, S. E. Retinal Conformation Governs pKa of Protonated Schiff Base in Rhodopsin Activation. *Journal of the American Chemical Society* **135**, 9391–9398 (2013). URL https://doi.org/10.1021/ja4002986.

40. Mertz, B., Lu, M., Brown, M. F. & Feller, S. E. Steric and Electronic Influences on the Torsional Energy Landscape of Retinal. *Biophysical Journal* **101**, L17–L19 (2011). URL http://www.sciencedirect.com/science/article/pii/S0006349511007181.

41. Crowley, M. F., Williamson, M. J. & Walker, R. C. CHAMBER: Comprehensive support for CHARMM force fields within the AMBER software. *International Journal of Quantum Chemistry* **109**, 3767–3772 (2009). URL http://onlinelibrary.wiley.com/doi/10.1002/qua.22372/abstract.

42. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **9**, 3878–3888 (2013).

43. D.A. Case,. R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.

44. Martínez-Mayorga, K., Pitman, M. C., Grossfield, A., Feller, S. E. & Brown, M. F. Retinal Counterion Switch Mechanism in Vision Evaluated by Molecular Simulations. *Journal of the American Chemical Society* **128**, 16502–16503 (2006). URL https://doi.org/10.1021/ja0671971.

45. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38, 27–28 (1996).

46. The PyMOL Molecular Graphics System, version 1.3r1 (2010) Schrödinger, LLC, New York.

47. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry* **32**, 2319–2327 (2011). URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.21787/abstract.

48. Pavelka, A. *et al.* CAVER: Algorithms for Analyzing Dynamics of Tunnels in Macromolecules. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **13**, 505–517 (2016).

49. Chovancova, E. *et al.* CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLOS Computational Biology* **8**, e1002708 (2012). URL http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002708.

50. LeVine, M. V. & Weinstein, H. NbIT - A New Information Theory-Based Analysis of Allosteric Mechanisms Reveals Residues that Underlie Function in the Leucine Transporter LeuT. *PLoS Computational Biology* **10** (2014). URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4006702/.

51. LeVine, M. V., Perez-Aguilar, J. M. & Weinstein, H. N-body Information Theory (NbIT) Analysis of Rigid-Body Dynamics in Intracellular Loop 2 of the 5-HT2a Receptor. *arXiv:1406.4730 [physics, q-bio]* (2014). URL http://arxiv.org/abs/1406.4730. ArXiv: 1406.4730.

52. Schobert, B., Brown, L. S. & Lanyi, J. K. Crystallographic Structures of the M and N Intermediates of Bacteriorhodopsin: Assembly of a Hydrogen-bonded Chain of Water Molecules Between Asp-96 and the Retinal Schiff Base. *Journal of Molecular Biology* **330**, 553–570 (2003). URL http://www.sciencedirect.com/science/article/pii/S002228360300576X.

53. Eisenhauer, K. *et al.* In Channelrhodopsin-2 Glu-90 Is Crucial for Ion Selectivity and Is Deprotonated during the Photocycle. *Journal of Biological Chemistry* **287**, 6904–6911 (2012). URL http://www.jbc.org/content/287/9/6904.

54. Kandori, H. *et al.* Water-Mediated Proton Transfer in Proteins: An FTIR Study of Bacteriorhodopsin. *Journal of the American Chemical Society* **117**, 2118–2119 (1995). URL https://doi.org/10.1021/ja00112a036.

55. Hatanaka, M., Kandori, H. & Maeda, A. Localization and orientation of functional water molecules in bacteriorhodopsin as revealed by polarized Fourier transform infrared spectroscopy. *Biophysical Journal* **73**, 1001–1006 (1997). URL http://www.sciencedirect.com/science/article/pii/S0006349597781335.

56. Mehler, M. *et al.* Chromophore Distortions in Photointermediates of Proteorhodopsin Visualized by Dynamic Nuclear Polarization-Enhanced Solid-State NMR. *Journal of the American Chemical Society* (2017). URL http://dx.doi.org/10.1021/jacs.7b05061.

57. Imasheva, E. S., Balashov, S. P., Wang, J. M., Dioumaev, A. K. & Lanyi, J. K. Selectivity of Retinal Photoisomerization in Proteorhodopsin Is Controlled by Aspartic Acid 227. *Biochemistry* **43**, 1648–1655 (2004). URL https://doi.org/10.1021/bi0355894.

58. Herz, J. *et al.* Critical Role of Asp227 in the Photocycle of Proteorhodopsin. *Biochemistry* **51**, 5589–5600 (2012). URL https://doi.org/10.1021/bi3003764.

59. Hillebrecht, J. R. *et al.* Structure, Function, and Wavelength Selection in Blue-Absorbing Proteorhodopsin. *Biochemistry* **45**, 1579–1590 (2006). URL http://dx.doi.org/10.1021/bi051851s.

60. Mao, J. *et al.* Structural Basis of the Green–Blue Color Switching in Proteorhodopsin as Determined by NMR Spectroscopy. *Journal of the American Chemical Society* **136**, 17578–17590 (2014). URL http://dx.doi.org/10.1021/ja5097946.

61. Ozaki, Y., Kawashima, T., Abe-Yoshizumi, R. & Kandori, H. A Color-Determining Amino Acid Residue of Proteorhodopsin. *Biochemistry* **53**, 6032–6040 (2014). URL https://doi.org/10.1021/bi500842w.

Chapter 3[1]

3. Molecular Dynamics Simulations as a Tool for Accurate Determination of Surfactant Micelle Properties

3.1 Abstract

Molecular dynamics (MD) simulations were used to characterize the equilibrium size, shape, hydration, and self-assembly of dodecylphosphocholine (DPC) and dodecyl-β-D-maltoside (DDM) micelles. We show that DPC molecules self-assemble to form micelles with sizes within the range reported in the experimental literature. The equilibrium shape of DPC and DDM micelles as well as associated micellar radii are in agreement with small angle X-ray scattering (SAXS) experiments and theoretical packing parameters. In addition, we show that hydration of the micelle interior is limited; however, flexibility of the acyl chains leads to dynamic encounters with the solvated outer shell of the micelle, providing an explanation for long-standing differences in models of micelle hydration. Altogether, our results provide fundamental understanding of physical characteristics of micelles that can be utilized to study other types of detergents and proteomicelle complexes.

3.2 Introduction

Surfactants (i.e., detergents) are a class of amphiphilic molecules with the ability to self-assemble into micelles and are commonly used in membrane protein extraction and solvation.[1] In fact, the detergent chosen for purification can alter the functionality of a membrane protein.[2] Recent studies have shown that detergent type and concentration modulate the degree of oligomerization and function of membrane proteins,[2] further extending their usefulness as a tool

[1] This chapter is published in Langmuir journal, 33, 9934-9943 (2017) DOI: 10.1021/acs.langmuir.7b02666

for investigating the structure-function relationship of membrane proteins. In addition, surfactants are used in applications such as cleaning, drug delivery,[3] oil recovery,[4] and as fuel additives. Despite their widespread use, the biophysical properties of detergent micelles at the molecular level are poorly understood. The vast majority of studies on surfactants have focused on either bulk properties or derivation of microscopic properties via indirect measurements (e.g., viscosity, conductivity, NMR).[5] Such methods suffer from a lack of precision in evaluating physical parameters such as radius of gyration, aggregation number, and micelle shape profile.[6]

Detailed knowledge of detergent-only systems is a prerequisite for any study of protein-detergent complexes, providing clues to factors governing detergent aggregation such as detergent-detergent and detergent-solvent interactions. Detergent micelles are commonly described by the following properties: 1) aggregation number, a_N (i.e., the average number of detergent molecules present in a micelle), 2) shape, 3) radius of gyration, and 4) internal hydration. However, in many cases the literature reflects a large distribution of values for these properties. For example, the a_N of micelles formed from *n*-dodecyl-β-D-maltoside (DDM) and *n*-dodecylphosphocholine (DPC), two common surfactants, have been reported to be anywhere between 78 and 149 and between 54 and 80 molecules, respectively.[7]

Theoretical studies on micelles have been conducted with several approaches. The most straightforward ones have been all-atom molecular dynamics (MD) simulations with initial configurations of preformed micelles at a_N's determined from experiment,[8–12] or self-assembly of micelles starting from a random distribution of detergent molecules near the a_N.[13–15] Other studies have used coarse-grained MD simulations[16] or Monte Carlo simulations[17] to investigate micelle behavior near the critical micelle concentration (CMC). However, it remains extremely difficult to relate these computational results to corresponding experimental data. For example, the radius of gyration obtained from MD simulations is often below experimental values despite using the experimentally-derived a_N. In addition, the timescales for micelle self-assembly can often be on the order of tens to hundreds of ns, which have only recently become widely accessible to the modeling community.[8,11,18]

Several models have been developed to describe interactions between aqueous solvent and micelles. However, it is unclear to what extent the interior of a micelle is hydrated and how the packing efficiency of a micelle is related to hydration.[19,20] Gruen developed a model of detergent micelles based on statistical mean field calculations, in which the central region is almost

completely dehydrated, but also acknowledged that atoms of the hydrophobic chain can be proximal to the hydrophilic shell of the micelle.[21] Hargreaves and coworkers came to a similar conclusion based on a combination of wide q-range neutron diffraction experiments and Monte Carlo-based simulations.[5] In addition, MD simulations determined that the interior of the micelle was devoid of water.[5,20] However, small angle scattering x-ray experiments (SAXS) propose that water molecules penetrate into the interior of the micelle.[22] A similar conclusion was reached from ^{13}C NMR experiments.[23] Clearly it is necessary to formulate an approach that allows for direct comparison of theoretical results to experimental ones and that quantitatively determines the extent to which the micelle interior is hydrated.

In this work, we have studied self-assembly of micelles starting from random distributions of detergent molecules and have compared them with simulations of preformed micelles. We have also utilized geometrical packing models in conjunction with MD simulations to characterize the shape of micelles. In addition, theoretical SAXS curves of micelles with different a_N were calculated from MD simulations and compared to experimental curves, allowing for explicit determination of the aggregation number. Finally, we have carried out a detailed study on hydration of the micelle interior, with results that help resolve discrepancies between currently proposed models of hydration. The commonly used detergents DPC and DDM chosen for this study possess hydrocarbon tails of the same length but different headgroups (zwitterionic for DPC versus polar carbohydrate for DDM) (Fig. 3.1). Our results show that the shape of the micelles from MD simulations agrees remarkably well with predictions based on theoretical packing parameters. Furthermore, a_N can be determined by a combination of simulations (i.e., self-assembly and preformed micelles) and calculation of theoretical SAXS curves based on these simulations. Finally, we observed that the micelle core is consistently dehydrated, despite noticeable conformational flexibility of the hydrocarbon tails of detergent molecules.

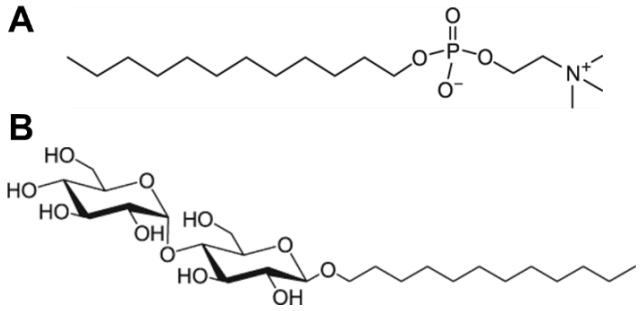


Figure 3.1. Detergents used in this study. (A) *n*-dodecylphosphocholine (DPC); (B) *n*-dodecyl-β-D-maltoside (DDM).

3.3 Computational Methods

Simulation setup. Single detergent molecules were constructed and minimized with Avogadro.[24] Initial configurations for both self-assembly and preformed micelle simulations were prepared with Packmol.[25] The distance cutoff of all atoms from other detergent molecules was 2.0 Å. VMD[26] was used to generate coordinate files and solubilize all systems. For each system, at least 10 Å of water padding was added around the micelle or randomly distributed detergents. Ions were added to reach a final concentration of 0.100 M NaCl. The CHARMM36 force field[14,27] was used for surfactants and ions in each system, and the TIP3P water model[28] was used for solvent. All simulations were carried out in the *NPT* ensemble at 310 K and 1 atm, using the Langevin thermostat and Nosé-Hoover Langevin barostat. Nonbonded interactions were treated with a switching function at 10 Å and cutoff at 12 Å. Dynamics were carried out with a 2 fs time step. All systems were minimized for 20 ps. All simulations were run in NAMD 2.10.[29]

Analysis. VMD[26] was used to visualize trajectories and compute radii of gyration, moments of inertia, solvent accessible surface area (SASA), water contact, hydrogen bonding, and energies. Representative snapshots (determined by *k*-means clustering in LOOS) were uploaded to the WAXSiS web server[30] to obtain theoretical SAXS curves. An in-house version of LOOS[31] was used for cluster analysis of self-assembly simulations to determine the distribution of sizes of detergent aggregates. The maximum inter-residue squared distance for each cluster was 4 $Å^2$. Average radial distribution functions were also calculated with LOOS. VMD and PyMOL[32] were used to render all molecular images. Volume of the dry region inside the micelle was computed with the 3V web server.[33] Criteria for hydrogen bonds were a cutoff distance of 3.9 Å between

donor and acceptor atoms[34] and an angle tolerance of 60° between acceptor-donor hydrogen atoms.[10]

3.4 Results and Discussion

Self-assembly of detergent molecules. To estimate the timescale of micelle formation, a series of simulations was carried out starting from random distributions of 54 and 80 DPC molecules (for a complete list of simulations, see Table S3.1 in Supporting Information). These values for DPC were chosen because they represent the lower and upper limits of aggregation numbers determined from experiment.[35–37] All simulations with 54 DPC molecules formed a single micelle within 506 ns (Fig. 3.2 A). The kinetics of self-assembly can be described as follows:

$$\Delta N(t) = N(t) - N(\infty) = c_1 e^{-t/\tau_1} + c_2 e^{-t/\tau_2} \quad 3.1$$

where N is the number of aggregates at any time and c_1 and c_2 are associated with the number of aggregates in the faster and slower aggregation periods. $N(\infty)$ is assumed to be 1. Similar to previous studies,[13,38,39] we observed fast and slow regimes in micelle formation, indicated by two different relaxation times (τ_1 and τ_2, respectively) (Table 3.1). For the case of the 54-molecule DPC systems, initial aggregation was one order of magnitude faster than previously reported (slow aggregation was essentially the same),[13] most likely due to smaller minimum intermolecular distances (2 Å) between detergents in our starting configurations. As can be seen by monitoring the radius of gyration during a representative trajectory (Fig. 3.2 B), smaller aggregates are nucleation points for self-assembly, after which stable micelle formation occurs over hundreds of ns.

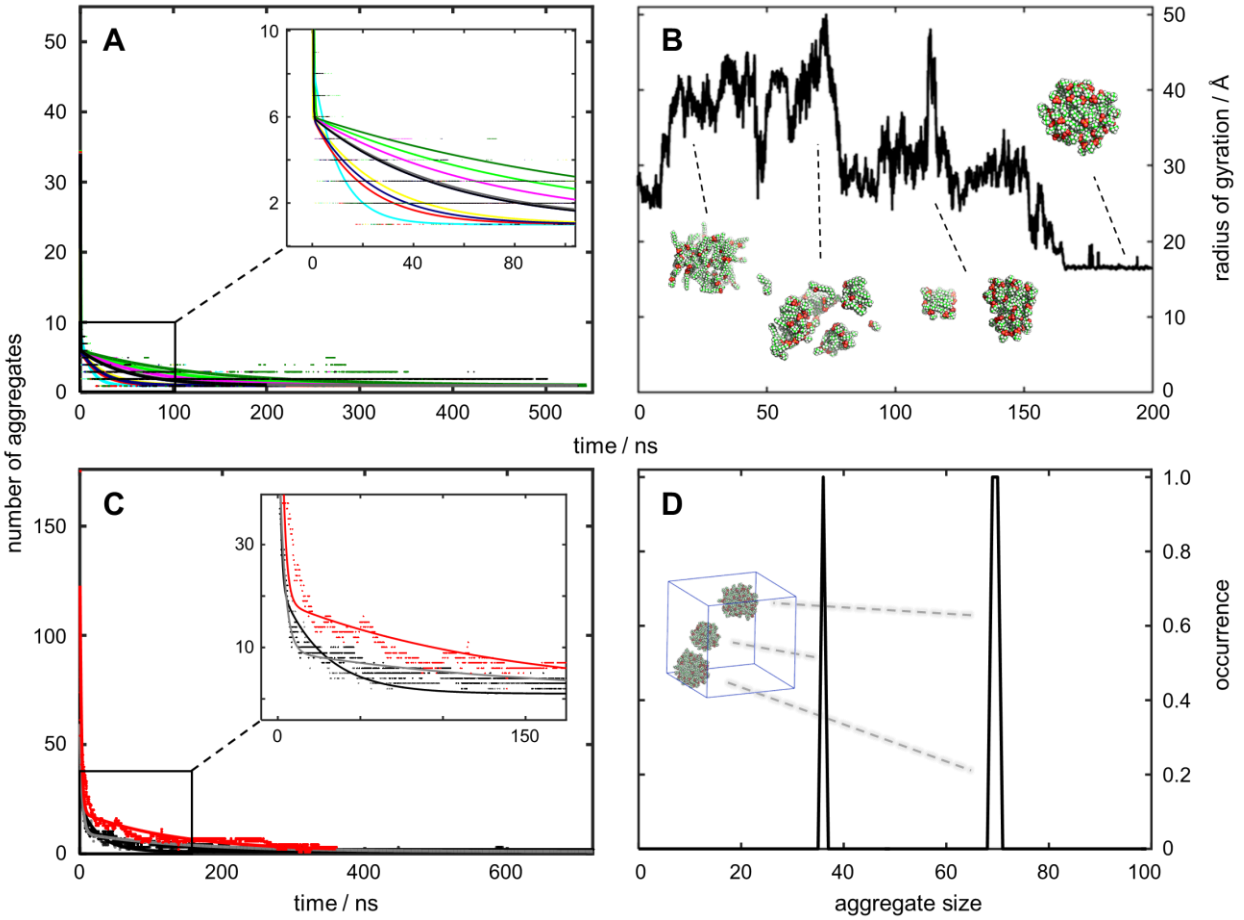


Figure 3.2. **Formation of DPC micelles occurs through a series of aggregations**. (A) Starting from an initial configuration of 54 randomly distributed DPC molecules, a single micelle forms in all our simulations in the range of 150-506 ns. All simulations started with the same initial configuration. Continuous lines show fitting curves. (B) Sharp increases in the radius of gyration (R_g) mark addition of free DPC molecules to the aggregate, followed by compaction into a micelle. Representative snapshots of the micelle are shown. (C) Micelle formation of DPC molecules occurs through a series of aggregation steps, starting from a random distribution of 80, (*black*), 115 (*gray*), and 175 (*red*) DPC molecules. All systems stabilize in ~350 ns. Continuous lines show corresponding fitting curves. (D) Distribution of detergent aggregate sizes for self-assembly simulation of 175 DPC molecules. Besides the small aggregate, two micelles with 69 and 70 molecules are within the range determined from X-ray experiments.[7] Snapshot of corresponding aggregates is shown to the left. Water molecules and ions are omitted for clarity.

Table 3.1 Kinetics fitting parameters for self-assembly simulations of DPC and DDM detergents into micelles.

detergent	n	c_1	c_2	τ_1 / ns	τ_2 / ns
DPC†	54	54	5.25±0.8	0.16±0.1	52.8±36.7
DPC*	70	70	10	0.07	0.07
DPC	80	80	21.75	1.12	26.48
DPC	115	51.69	8.38	3.28	150
DPC‡	175	116.5	18.1	2.14	135.1
DDM	110	109.9	5.77	0.03	0.63
DDM⁑	110	110	24.19	0.47	36.97
DDM	130	130	1.13	0.47	27.70
DDM	200	200	7.08	0.50	11.72

† Average of 10 simulations.
‡ Average of 2 simulations.
* Concentrated (1703 times CMC). DPC systems typically ranged from 100–300 times CMC (1.1 mM).
⁑ Dilute (379 times CMC). DDM systems typically ranged from 700–3000 times CMC (0.2 mM).

In contrast to our 54-molecule simulations, the 80-molecule DPC systems did not converge to a single micelle, even after hundreds of ns. We hypothesize that this may be due to the fact that initial formation of smaller aggregates (e.g., 40 molecules) is more energetically favorable than formation of a larger micelle (Fig. 3.2 C). In a similar manner, after about 350 ns, our 115-molecule DPC system stabilized into two micelles between 50-60 molecules (Fig. 3.2 C). For the 175-molecule system, two micelles (69 and 70 molecules, respectively) formed after 360 ns, along with a smaller aggregate with 36 molecules (Fig. 3.2 D). The general trend for these simulations is that the size of these micelles lies within the experimentally determined range of a_N = 54-80.[7] It appears that the upper limit for self-assembly of DPC micelles is around seventy molecules, based on the simulations of larger systems. The kinetics of aggregation and micelle formation are similar to, but slower than, the 54 DPC systems (i.e., with respect to τ_1). Initial aggregation takes place within the first 20 ns, followed by a slower growing process towards stable micelles (Fig. 3.2 D). Similar self-assembly behavior was observed for DDM systems, although the kinetics of self-assembly are directly influenced by system concentrations. The effect of simulating micelle self-assembly far above the CMC and practical aspects of setting up these types of simulations are discussed in the Supporting Information.

Packing factor as a determinant of micelle shape. By determining the uniformity and shape of micelles, we can qualitatively compare the stability of micelles with different a_N. Simple

steric factors[40,41] are useful tools to carry out this comparison. According to Israelachvili's formulation, the shape of a micelle is defined by a packing parameter, P:

$$P = \frac{v}{a_0 \ell_c} \quad 3.2$$

where a_0 is the optimal headgroup area, v is the volume of the hydrocarbon chain(s), and ℓ_c is the maximum effective length of the tail (i.e., critical chain length). Tanford[42] developed approximate formulas for v and ℓ_c:

$$\ell_c = (0.154 + 0.1265 n_c) \text{ nm} \quad 3.3$$

$$v = (27.4 + 26.9 n_c) \times 10^{-3} \text{ nm}^3 \quad 3.4$$

where n_c is the number of carbon atoms in the hydrophobic tail. Using a spherical model for our micelle, we have

$$a_N = 4\pi R^2/a_0 = 4\pi R^3/3v \quad 3.5$$

with R being the radius of the micelle. The relation in Eq. 3.5 means that the aggregation number (or size of a preformed micelle) is the total area of the micelle divided by the headgroup area of a detergent or the total volume of a micelle divided by the volume of a single acyl chain. Solving for R in E. 3-5 gives

$$R = 3v/a_0 \quad 3.6$$

Eq. 3.6 can be used to calculate a_0 and subsequently the packing parameter, P.

We determined values of P for preformed micelles with 54, 80, 115, and 175 DPC molecules. Micelle sizes of 115 and 175 molecules are 2-3 times larger than reported values of aggregation numbers for DPC (a_N = 54-80)[7] and were chosen to characterize how micelle shape would be affected significantly above the aggregation number. Our simulations of preformed DPC micelles produce spherical, ellipsoid, and elongated ellipsoid shapes, respectively. In addition, we also simulated a preformed micelle system of 300 DPC molecules (~5 times the a_N). After hundreds of ns, the micelle deformed from a spherical to a cylindrical or rodlike shape, corresponding to the hexagonal phase and in agreement with the relevant packing parameter ($P \approx ½$) (Fig. 3.3). Similar results were obtained for DDM simulations over a series of micelle sizes (60, 80, 110, 130, 149, and 200 DDM molecules).

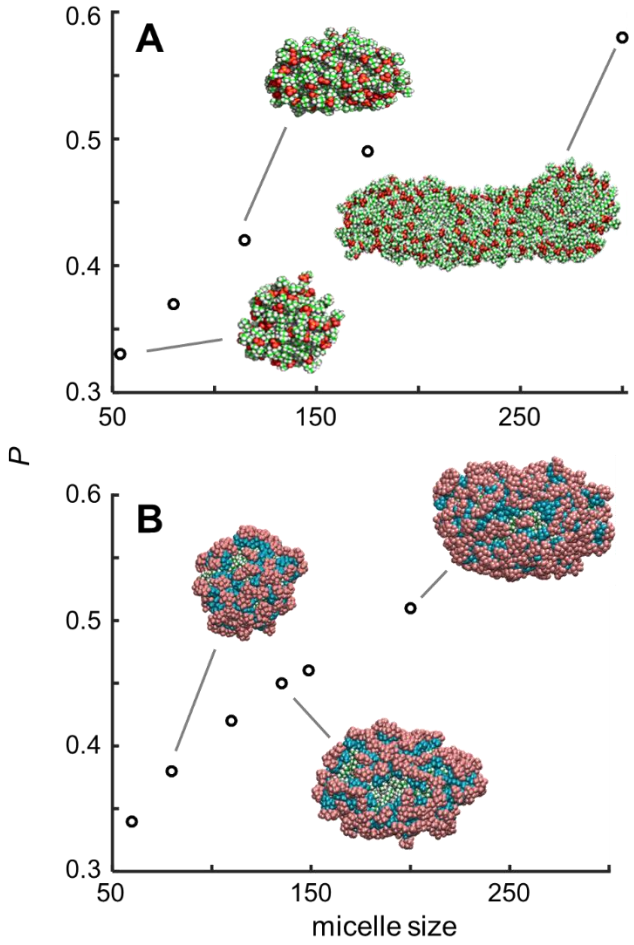


Figure 3.3. Packing parameters and aggregation number are positively correlated. (A) Plot of packing parameter versus micellar size and corresponding shapes. After 200 ns, preformed micelles with a_N = 54, 80, 115, and 175 and P values of 0.33, 0.37, 0.42 and 0.49, take spherical, ellipsoid, and elongated ellipsoid shapes, respectively. A spherically preformed micelle with a_N = 300 deformed to a cylindrical micelle (P = 0.58) after 700 ns. All observations are in agreement with geometrically-based packing models (Israelachvili). (B) A similar trend is observed for DDM micelles.

There is a surprisingly consistent level of agreement between theoretical packing predictions[40] and the micellar shapes reported in the literature,[7,35] and our modeling of these systems provides computational validation of theory and experiment. P and a_N are positively correlated: as the cubed root of a_N increases, so does P. Our systems progress in shape from spherical (a_N = 54) to prolate ellipsoidal (a_N = 80) to a cylindrical micelle (a_N = 300). These results indicate that the shape of a micelle can be predicted by knowing the aggregation number and using it in conjunction with the packing parameter for each respective type of detergent. According to theory, the hydrophobic effect governs micelle formation by attracting the hydrocarbon tails together, while the headgroups tend to remain exposed to water and ions.[40] As a result of these competing forces, there is an optimal headgroup area a_0 for which the interaction energy between

detergent molecules has a minimum.[40,41] The combination of a_0, v, and ℓ_c determines the most favored shape of a micelle.

Moments of inertia/asymmetry. To further characterize micelle structures, principal moments of inertia, I_1, I_2, and I_3 and the asymmetry parameter ($\alpha = (2\ I_1 - I_2 - I_3) / (I_1 + I_2 + I_3)$) were calculated for preformed micelles composed of DPC or DDM. For a perfect sphere, $\alpha = 0$, and any micelle with $\alpha \leq 0.05$ is still considered spherical.[13] Comparing the values of α for the DPC systems with 54 or 80 molecules shows that while the former slightly deviates from a spherical shape, the larger micelle is completely ellipsoid, in agreement with the corresponding packing parameters (Table 3.2). These results are also consistent with a previous study using the CHARMM36 force field.[13] As system size increases, the asymmetry parameter continues to increase. Likewise, a similar trend is observed for the preformed DDM micelles. We did not observe significant differences in moments of inertia and asymmetry parameters for self-assembled versus preformed systems (see Table S3.2 in Supporting Information).

Table 3.2 Average principal moments of inertia for preformed DPC and DDM micelles. Each simulation was at least 200 ns. The first 10 ns of each simulation were excluded from calculations.

surfactant	*n*	I_1*	I_2	I_3	I_1/I_3	I_1/I_2	I_2/I_3	α
DPC	54	4.00 ± 0.16	3.69 ± 0. 13	3.28 ± 0. 13	1.22	1.08	1.12	0.09 ± 0.03
DPC	80	7.92 ± 0.30	7.13 ± 0.35	5.95 ± 0.25	1.33	1.13	1.20	0.13 ± 0.04
DPC	115	1.58 ± 0.69	1.42 ± 0.12	0.97 ± 0.75	1.63	1.12	1.46	0.20 ± 0.06
DPC	175	2.20 ± 0.11	1.85 ± 0.79	1.26 ± 0.47	1.76	1.19	1.47	0.24 ± 0.38
DDM	60	13.3 ± 0.5	12.2 ± 0.5	11.0 ± 0.4	1.21	1.09	1.11	0.09 ± 0.03
DDM	80	13.3 ± 0.5	12.2 ± 0.5	11.0 ± 0.4	1.21	1.08	1.11	0.09 ± 0.03
DDM	110	22.1 ± 0.4	20.1 ± 0.5	17.5 ± 0.4	1.26	1.10	1.14	0.11 ± 0.02
DDM	135	32.1 ± 1.2	28.3 ± 1.0	24.8 ± 0.8	1.30	1.13	1.14	0.13 ± 0.03
DDM	149	39.0 ± 1.8	33.0 ± 1.2	29.4 ± 1.9	1.32	1.18	1.12	0.15 ± 0.04
DDM	200	74.4 ± 7.0	59.1 ± 5.7	51.2 ± 2.3	1.45	1.26	1.15	0.21 ± 0.07

* All moments of inertia are x 10^6 amu-Å^2.

Determining better estimates of aggregation numbers consistent with theory and experiment. As seen above, the average size of DPC micelles obtained from self-assembly simulations lies within the range Lauterwein and coworkers determined by quasi-electric light scattering (56 ± 5).[36] However, depending on the experimental technique used, different aggregation numbers have been determined for DPC, namely 70-80.[7,43] One of the most frequently used techniques in studying micelles is small-angle X-ray scattering (SAXS). SAXS does not

require crystallized samples, making it an excellent technique for providing low-resolution insights into a large range of molecules.[44,45] In addition, since molecules are inherently dynamic, MD simulations are well-suited to connect microscopic phenomena to macroscopic ensemble-averaged observations (i.e., SAXS experiments). The biggest advantage of reproducing SAXS profiles by MD simulations is that we can unambiguously characterize the size, shape, electron density, and exact number of atoms of the structure. Numerous studies have theoretically determined SAXS curves that are in general agreement with experiment.[30,46–48] More recent methodological developments by Hub and coworkers,[30,48] in which they calculate SAXS curves from MD simulations that use explicit solvent models, eliminate common issues encountered through the more conventional use of implicit solvent models. Explicitly modeling the solvation shell around the solute molecule(s) accounts for changes in solvent density when transitioning from the hydration shell to bulk solvent and also for the solvent that is inherently part of experimental SAXS curves.[48] This allows an accurate determination of micelle size and shape. Calculations of SAXS curves for the DPC and DDM systems over a range of micelle sizes clearly show that a particular system size definitively reproduces experimental data (Fig. 3.4).[7] The simulations closest to experiment are the 80-molecule DPC and 149-molecule DDM systems, respectively. Both micelles have an ellipsoid shape, consistent with their classification as prolate (DPC) and oblate (DDM) micelles. However, the correlation with experiment is poorer for DPC than DDM systems. A possible reason is wider distribution of experimental micelle size of DPC compared to DDM; if we consider the upper limit of the experimental aggregation number, a_N, the relative error in a_N is 0.09 for DDM and 0.15 for DPC.[7]

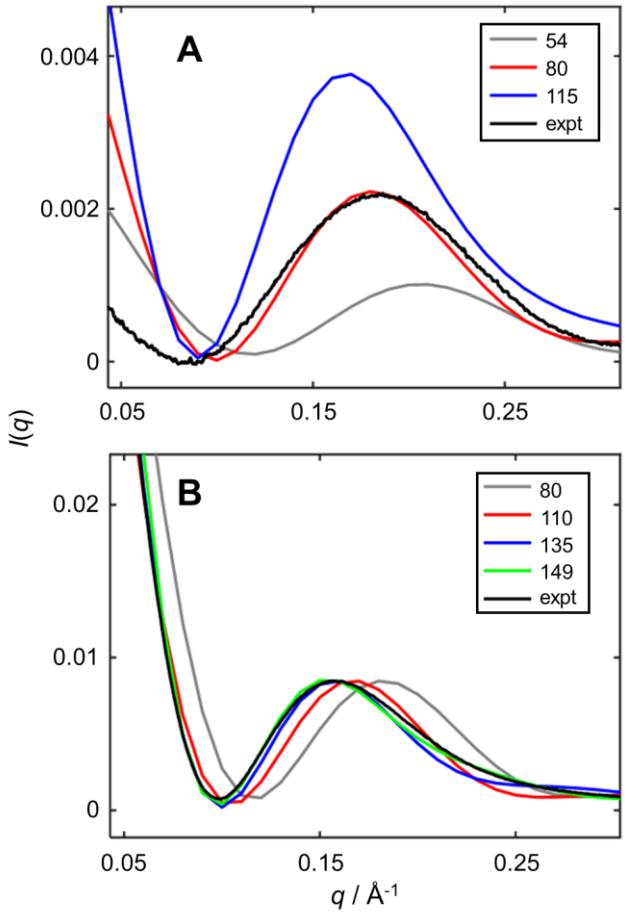


Figure 3.4. SAXS curves derived from MD simulations can be used for accurate determination of surfactant aggregation number. Comparison of SAXS curves generated with the WAXSiS web server[30] from simulations of preformed micelles to experimental curves from Columbus and coworkers.[7] (A) Position of minimum and maximum of curve for the 80-molecule DPC system (*dash-dot line*) is most similar to experimental curve (*solid line*). (B) The curve for the 149-molecule DDM system (dashed line) is most similar to the experimental one (*solid line*). Each simulation is minimum 200 ns long.

Theoretical SAXS curves can be used in conjunction with packing parameters to determine the size and shape of micelles. As seen from snapshots of our simulations (Fig. 3.5), DPC micelles have a prolate shape and DDM micelles are more oblate, in agreement with experiment.[7] The ability to accurately model micellar shape allows for predictions of micelle properties important for solubilization of membrane proteins. For example, the degree of hydrophobic mismatch between membrane proteins and native micellar shapes is correlated to the oligomerization states of the membrane protein as well as the ability to maintain the fold and biological function of the protein.[49,50] In addition, accurate modeling provides an additional tool for structural biologists, as it opens the possibility to screen different surfactants to predict their micellar properties before testing their ability to solubilize membrane proteins in the lab.

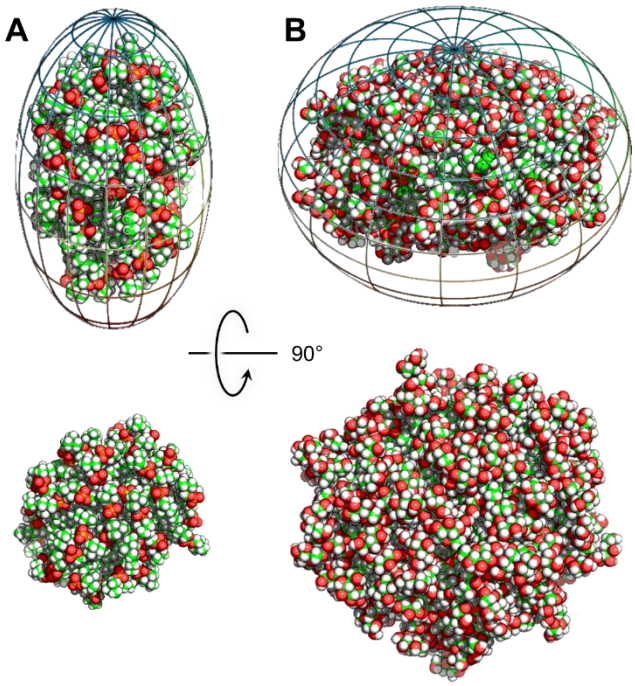


Figure 3.5. DPC and DDM simulations consistently reproduce shape of their respective micelles. Snapshots from simulations of (A) an 80-molecule preformed DPC micelle with a prolate shape and (B**)** a 149-molecule preformed DDM micelle with an oblate shape. *Top*: front-facing view of micelles; *bottom*: top-down view of micelles. Prolate ellipsoids occur when $c > a$, where c is the symmetry axis (in z direction) and a is equatorial axis of ellipsoid. Oblate ellipsoids are occur when $c < a$. Geometries of each shape are shown in front-facing view (*top*).

Radius of gyration. The average radius of gyration (R_g) for a preformed micelle of 54 DPC molecules is 16.44 ± 0.01 Å, similar to the value reported by Abel and co-workers.[13] For a preformed 80-molecule DPC micelle it is 18.71 ± 0.18 Å. However, the R_g for DPC micelles determined from SAXS experiments is 34.5 ± 0.8 Å.[7] The reason for this discrepancy is because the hydration shell associated with the solute (i.e., the micelle) is included in experimental calculations,[30,51] whereas MD simulations give us the ability to calculate an exact R_g based on solute molecules alone. Since the WAXSiS web server includes the effect of the solvent layer through use of a spatial envelope (this also accounts for slight fluctuations in the micelle), R_g values determined from WAXSiS are much closer to experimental values. Thus, the radius of gyration calculated from theoretical SAXS curves of the DPC micelle with 80 molecules is 31.17 Å, a relative error of 9.6% compared to experiment. Likewise, for the 149-molecule DDM system R_g is 30.85 Å, with a relative error of 9.7% (Table 3.3). It is worth noting that the R_g of preformed micelles is very similar to that of the corresponding randomly assembled systems (Table S3.1 in Supporting Information). In fact, all properties of the preformed and randomly assembled systems are very similar, regardless of initial configuration, in agreement with previous simulations.[8,18]

Table 3.3 Radius of gyration of preformed DPC and DDM micelles calculated from the WAXSiS server, which includes the first hydration shell in its determination of *Rg*.

DPC	R_g (Å)	DDM	R_g (Å)
54	25.45	80	29.97
70	30.25	110	29.99
80	31.17	135	30.85
115	33.74	149	32.59
experiment[7]*	34.5 ± 0.8	experiment[7]*	31.8 ± 0.1

* Experimental radii are for DPC and DDM micelles with 68-80 and 135-149 detergent molecules, respectively.

Internal hydration of the micelle. Disagreements within the field still exist as to the extent of hydration within the micellar core. While MD simulations do not show noticeable water penetration into the core,[5,20] experimental studies have found that the micelle interior is highly hydrated (up to 40%), even for carbon atoms far from the headgroup.[19,52] In experiments, for example, a solvent-sensitive property of detergents such as NMR shielding is measured in solutions with concentrations below the CMC (when there are no aggregates present) and above the CMC (i.e., when micelles are present) or in organic solvent and in water. Comparing data obtained from the system in two different conditions reveals the degree of hydration of the hydrocarbon tail in the micelle. These properties can also be measured by neutron scattering, SAXS, or electron paramagnetic resonance (EPR).[19] Menger et al.[23] used ^{13}C NMR to characterize the hydration of a carbonyl group positioned in the center of the hydrocarbon tail of a cationic surfactant (8-ketohexadecyltrimethylammonium bromide), both in organic and aqueous solvent. The difference between NMR peaks in the two media is directly related to the extent of hydration of the carbonyl group. Their results showed that the carbonyl group is in contact with water above the CMC (i.e., when the carbonyl groups are inside a micelle). Courchene calculated the hydrated volume of micelles in aqueous solutions with viscosity and sedimentation measurements.[52] His results showed that 10-12 water molecules are associated with each detergent in a micelle. In a similar study, Vetter hypothesized that water molecules penetrate the micelle to some extent.[53] The general conclusion of these studies is that water penetrates into the micelle; however, does hydration of the hydrocarbon tail necessarily mean that the micellar core is hydrated?

Hydration analysis of an 80-molecule preformed DPC micelle shows that the entire hydrocarbon tail is in close contact with several water molecules, in agreement with experiments (Fig. 3.6).[19] However, a completely dry region in the micelle interior exists, as shown by a radial

distribution function. The reason for this seemingly irreconcilable difference is that the hydrophobic tail is hydrated only when it approaches the micelle surface, an action made possible by the highly flexible nature of the acyl chain (Fig. 3.6). Capponi and co-workers have observed a similar phenomenon in lipid bilayers, in which the terminal methyl group of the acyl chains has enough flexibility to explore the water-membrane interface.[54]

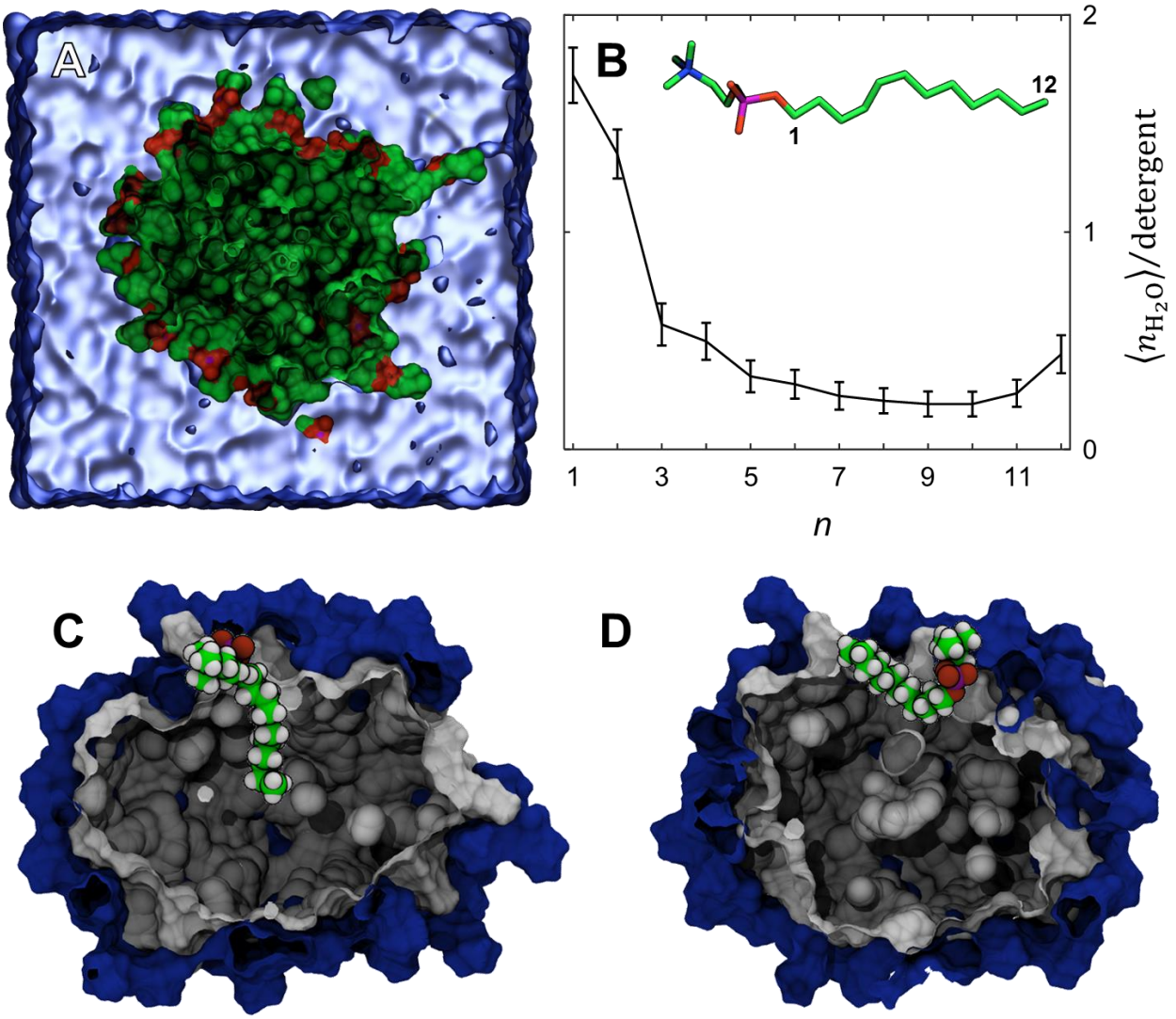


Figure 3.6. DPC micelles undergo minimal solvation of the interior, but acyl chain flexibility allows for proximal interactions with the first hydration shell. (A) Snapshot of a preformed DPC micelle containing 80 molecules with sliced view of the micelle interior. No waters are present within the micelle. *Blue*: bulk water. The volume of the dry region is 23623 Å^3, with an effective radius of 12.28 Å. (B) Average number of water molecules within 4 Å from carbon atoms in the hydrocarbon tail of DPC for a preformed micelle of 80 DPC molecules. *n* corresponds to numbering of the acyl chain (*n* = 1 is bound to the phosphate group and *n* =12 is the chain terminus). The hydrophobic tail is hydrated, in agreement with Raman spectroscopy.[19] The apparent contradiction between A and B can be resolved by the fact that the hydrocarbon tails are hydrated when they curl towards the surface of the micelle. Similar behavior is observed for DDM micelles (Fig. S3.7). (C) Snapshot from a simulation of an 80-molecule preformed DPC micelle. The acyl chain of a representative DPC molecule is shielded from the surrounding water molecules (*left*). After 30 ns, the acyl chain of the same detergent curls towards the surface of the micelle and is hydrated, while the core of the micelle is still dry (*right*). *Green*: carbon atoms of the single DPC molecule; *white*: hydrogen; *pink*: phosphorus; *red*: oxygen; *blue*: hydrophilic shell of micelle; *gray*: hydrophobic core.

To compare our results with experiments that report water penetration into the micelle interior,[19] we also carried out similar simulations on a preformed micelle of sodium dodecanoate. The number of detergent molecules was 64, corresponding to the reported aggregation number.[55] After 160 ns, no water penetration to the micellar core was observed (Fig. S3.6 in Supporting

Information). Our results are consistent with the work of Chun, et. al. on sodium dodecyl sulfate micelles, in which they found that the hydrophobic acyl chains were in close proximity to the micelle surface, but that the hydrophobic effect made both detergent dissociation and water penetration energetically unfavorable.[12]

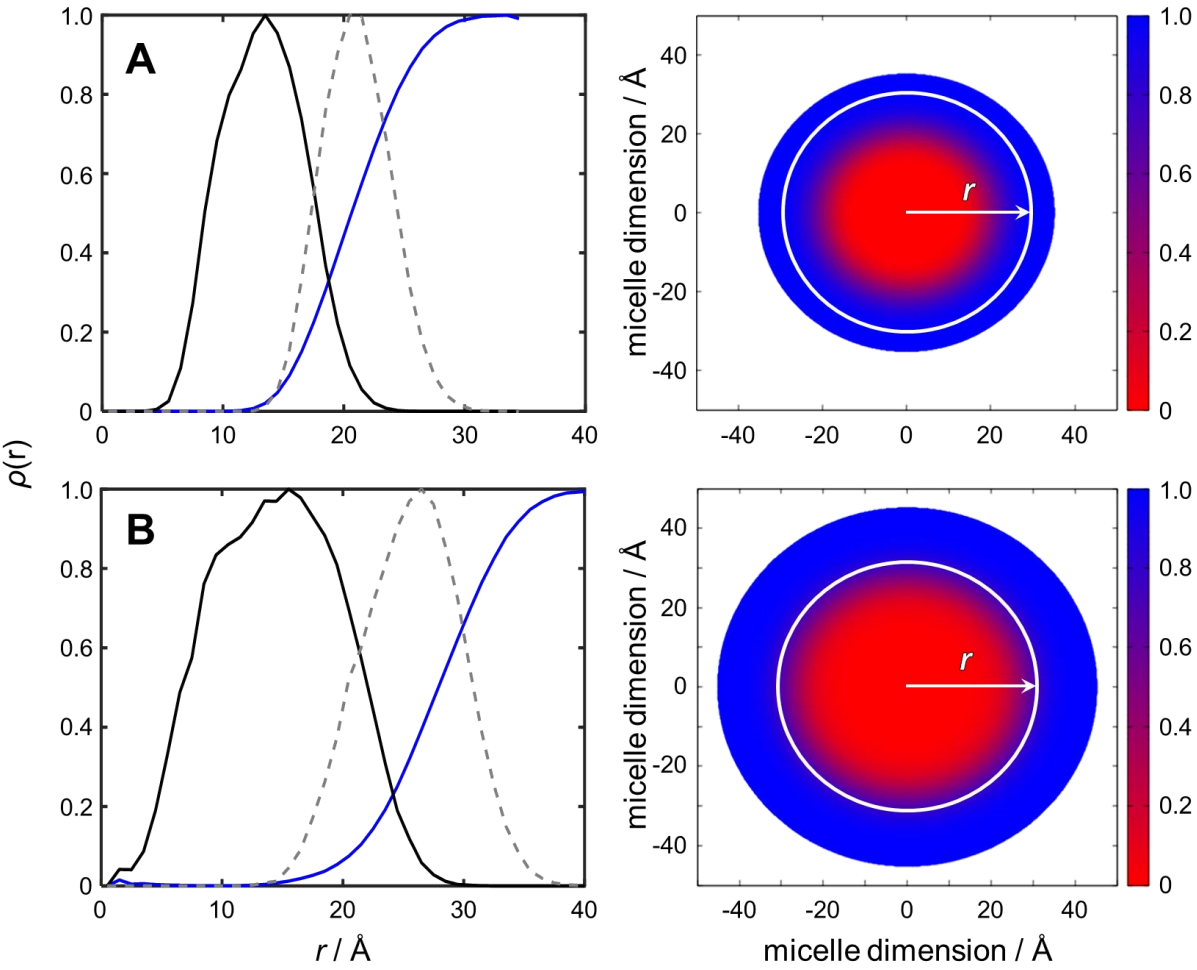


Figure 3.7. DPC micelles have higher levels of solvent penetration than DDM micelles. Normalized average radial distribution function of water (*blue*), headgroup (*dashed gray*) and tail (*black*) of (A) preformed 80-molecule DPC micelle and (B) preformed 135-molecule DDM micelle, averaged over 200 ns. r = 0 is the center of mass of the micelle. The cores of each micelle are dehydrated up to ~12 Å and ~14 Å, respectively. Heat maps on the right show corresponding normalized water densities. *White*: radii calculated from theoretical SAXS data on WAXSiS web server.

Radial distribution functions (RDF) show that the micellar core of both DPC and DDM systems are completely dehydrated (Fig. 3.7). This phenomena has also been observed for other detergent micelle simulations.[12,56] However, the overlapping shoulders of the distributions for water and the acyl tails confirm that the tails are partially hydrated. Even higher overlap occurs between water and the headgroup region, showing that the outer shell of each type of micelle is

highly hydrated. The degree of hydration is different for each type of detergent. The DDM headgroups pack the surface of the micelle more effectively than DPC, preventing penetration of water molecules into the interior of the micelle.

The degree of hydration of micelles can be further illustrated by examining the solvent-accessible surface area (SASA) per molecule for the respective surfactant molecules. For DPC, SASA is largest for the smallest micelle (54 molecules), steadily decreasing as micelle size increases (Table 3.4 and Fig. 3.8). For DDM, a similar trend is observed, albeit with a smaller slope. In general, solvent accessibility decreases with an increase in micelle size up to the limit of the micellar phase regime, indicating that packing becomes more efficient as size increases. Upon transitioning from the micellar to the hexagonal phase, as in the case of our preformed 300-molecule DPC micelle, the change in shape leads to a corresponding increase in SASA.

Table 3.4 Average Solvent Accessible Surface Area (SASA) per detergent molecule for preformed DPC and DDM micelles.

surfactant	*n**	$SASA_{total}$ ($Å^2$)†	$SASA_{headgroup}$ ($Å^2$)
DPC	54	237.48 ± 8.21	188.81 ± 8.73
DPC	80	215.36 ± 6.28	172.91 ± 4.31
DPC	115	192.57 ± 5.97	162.05 ± 7.34
DPC	175	175.21 ± 30.02	114.29 ± 7.25
DPC	300	232.09 ± 7.56	159.48 ± 2.76
DDM	60	266.69 ± 10.03	237.63 ± 8.62
DDM	80	237.82 ± 7.57	215.89 ± 6.30
DDM	110	209.62 ± 4.93	191.44 ± 4.14
DDM	135	195.95 ± 8.30	179.88 ± 5.74
DDM	149	190.27 ± 8.34	174.52 ± 5.74
DDM	200	187.43 ± 9.92	164.24 ± 4.81

* Abbreviations: *n*: number of surfactant molecules; SASA: solvent-accessible surface area; total: entire surfactant molecule; headgroup: portion of detergent without the C_{12} acyl chain.

† SASA results analyzed for 200 ns simulations.

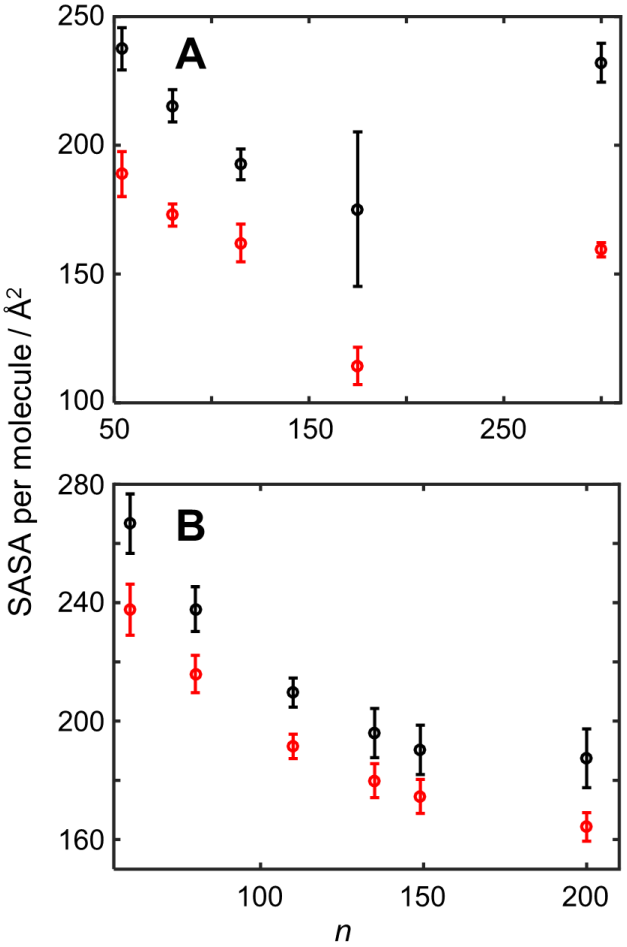


Figure 3.8. **Exposure of surfactant to solvent generally decreases with increasing micelle size in the micellar phase**. (A) Plot of the solvent-accessible surface area (SASA) per molecule for the entire DPC molecule (*black*) and the headgroup (*red*) as a function of micelle size. (B) Corresponding data for DDM micelles.

Having 4 ether oxygens and 7 hydroxyl groups in the headgroup allows DDM to form networks of detergent-detergent and detergent-water hydrogen bonds (Table 3.5). The higher number of hydrogen bonds per detergent in DDM may contribute to the lack of penetration by water molecules into the interior of DDM micelles compared to DPC micelles (Fig. 3.7). Also, as the size of DDM micelles increases the number of detergent-water hydrogen bonds slightly decreases, while detergent-detergent hydrogen bonds slightly increase, implying more efficient packing of larger micelles. Although our cutoff distance for detection of hydrogen bonds is slightly higher than previous work on DDM (3.9 Å compared to 3.5 Å), the number of detergent-water hydrogen bonds for the 130-molecule micelle, 14.2 ± 0.4, is comparable with their 130-molecule micelle, namely, 11.97 ± 0.04.[57] In addition, the number of detergent-detergent hydrogen bonds is significantly less than detergent-water hydrogen bonds. The higher frequency of detergent-solvent

hydrogen bonds compared to detergent-detergent hydrogen bonds has been previously observed for maltosides and glycoside micelles;[10,57] this behavior can be attributed to the fact that the glucopyranosyl rings of neighboring DDM molecules must adopt very specific orientations to form hydrogen bonds in comparison to forming hydrogen bonds with water molecules.

Table 3.5 Average number of hydrogen bonds per detergent molecule for preformed DPC and DDM micelles.

DPC		**DDM**		
n	detergent-water	*n*	detergent-water	detergent-detergent
54	7.6±0.7	60	16.1±0.6	3.7±0.9
70	6.6±0.4	80	16.2±1.0	4.1±0.8
80	8.0±0.2	110	15.6±0.8	4.3±0.7
115	7.6±0.3	130	14.2±0.4	4.7±0.9
175	4.4±0.4	149	13.9±0.6	4.8±0.6
300	6.6±0.2	200	12.3±0.3	4.8±0.6

3.5 Summary

Physical properties of *n*-dodecylphosphocholine (DPC) and *n*-dodecyl-β-D-maltoside (DDM) micelles have been studied computationally and compared with experiment. Our findings show that theoretical packing considerations are useful in determining the shape of micelles, although they do not explicitly account for complex phenomena such as hydrogen bonding. All micellar shapes we have obtained are in agreement with the ones predicted based on their packing parameters and SAXS experiments, and can help us in better understanding micellar geometry for applications such as solvation of membrane proteins. Self-assembly of DPC and DDM micelles was also carried out, leading to micellar sizes in agreement with experiment. In addition, our generation of theoretical SAXS scattering curves agrees with these experiments, and correctly identifies the shapes for DPC (prolate) and DDM (oblate) micelles. Moreover, our results show that the micellar core is dehydrated, although the hydrocarbon tails of detergent molecules are conformationally flexible enough to approach the hydration shell proximal to the headgroup region. This type of physical characterization of detergent micelles used in membrane protein

extraction is a necessary fundamental step in studying physical properties of protein-micelle complexes.

Supporting Information

Kinetics of self-assembly for DDM micelles is more rapid than DPC micelles due to higher system concentrations. DDM simulations with 110 molecules show similar behavior to the DPC systems, forming a single micelle; however, the characteristic times for micelle formation are an order of magnitude faster than for the DPC systems (Fig. S3.1). The most likely reason for faster aggregation times is because our DDM simulations are 1000-1500 times the CMC (0.2 mM), whereas our DPC simulations had much lower concentrations in comparison to their CMC (1.1 mM) (about 180 times). Previous studies have shown that working with surfactants far above the CMC leads to rapid aggregation, as the vast majority of intermolecular interactions are detergent-detergent, not detergent-solvent.[58,59]

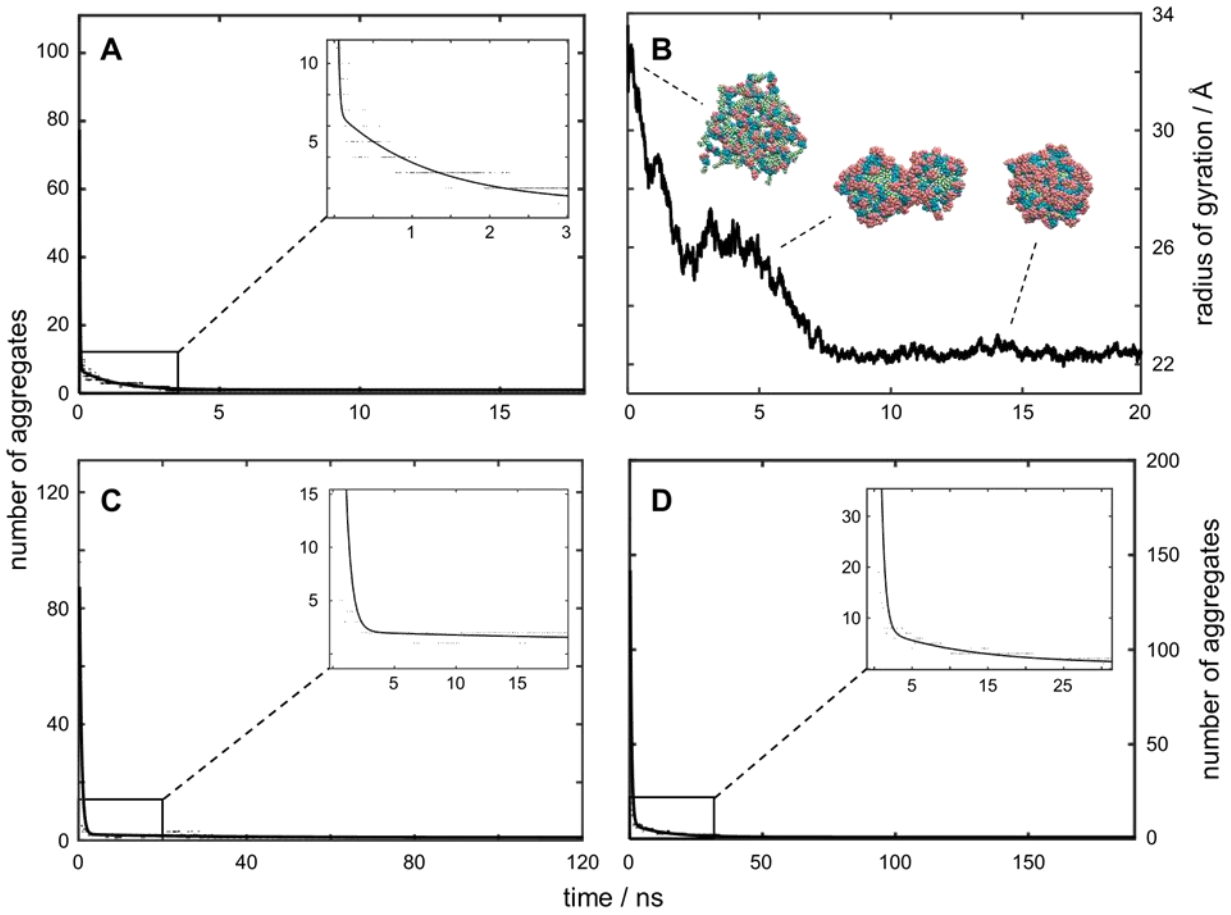


Figure S3.1. Kinetics of DDM micelle assembly are significantly faster than for DPC micelles. (A) Clustering of 110 randomly distributed DDM molecules, showing fast formation of a micelle. Continuous lines show fitting curves. (B) Radius of gyration of same DDM system. Sharp increases in R_g reflect similar behavior to DPC micelle formation, with addition of single DDM molecules to the aggregate. Similar plots for 130- (C) and 200-molecule (D) systems of randomly distributed DDM molecules, respectively.

Similar rates of self-assembly were observed for other DDM systems as well (Fig. S3.1 C and S3.1 D). However, decreasing the concentration of DDM detergent significantly increases the timescale of the fast and slow process (Fig. S3.2 A). On the other hand, increasing concentration of DPC system accelerates

the rate of micelle formation (Fig. S3.2 B). These results indicate the necessity of considering concentration with respect to CMC of self-assembly systems in MD simulations.

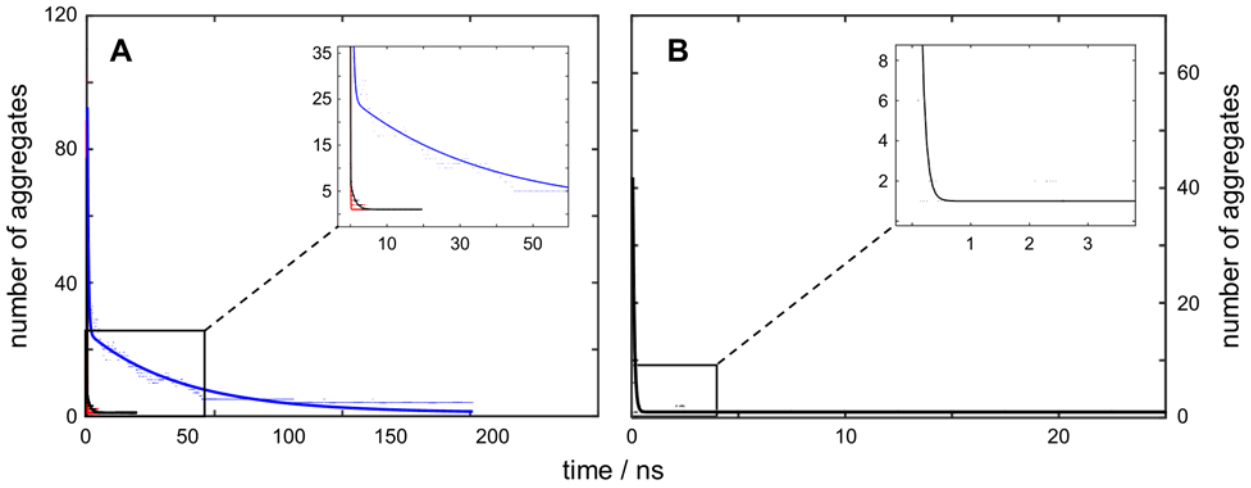


Figure S3.2. Kinetics of micelle assembly depends on system concentration. (A) Dilute system of 110 DDM molecules (*blue*) forms a micelle significantly more slowly than concentrated systems (*black and red*). (B) Kinetics of 70-molecule DPC system at 1700 times the CMC. This markedly higher concentration significantly accelerates micelle formation, especially compared to 80-DPC system at much more dilute concentration (see Table 3.1 in main manuscript).

Micelles are energetically favorable beyond a_N. If we calculate the energies for preformed micelles of DPC or DDM, the energy per detergent molecule generally increases with increasing micelle size. We originally postulated that the lowest energy per molecule (i.e., the most energetically stable micelle) would determine the optimal aggregation number. Our observed energy minima at 50 DPC molecules and 80 DDM molecules exist at the low end of the literature range of aggregation numbers for both DPC (54–80 molecules) and DDM (78–149 molecules) (Fig. S3.3).[7] This trend is in agreement with our kinetics data; specifically, the largest micelles we observed that spontaneously formed for DPC consisted of 70 molecules. This also implies that subtler factors besides energy may be influencing the optimum configuration of a micelle. Two other insights can be gained from our energy analysis. First, the increase in energy per detergent molecule begins to level out as DPC and DDM micelle size goes beyond 200 molecules. This is consistent with a transition from a micellar shape to a cylindrical shape and can be explained by theoretical packing parameters (see main paper for discussion). Second, even at micelle sizes well beyond the aggregation number (e.g., 300 molecules for DPC), the energy per detergent molecule is still negative. This means that it is energetically unfavorable (and thus, highly unlikely) for a detergent molecule to dissociate from the micellar complex.

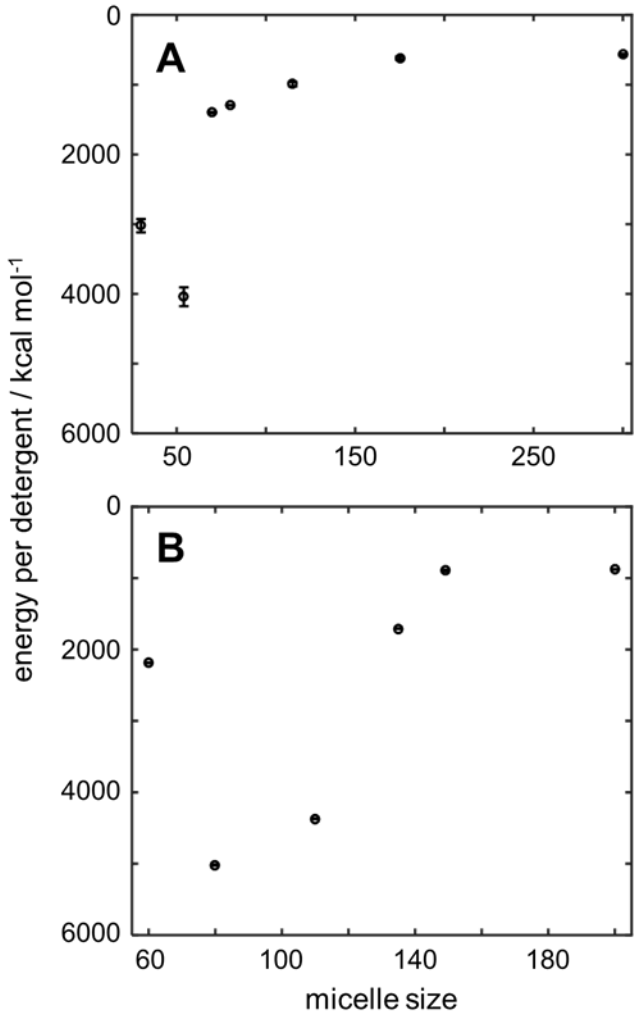


Figure S3.3. In general, energy per detergent molecule increases with an increase in micelle size. Total energies per detergent molecules of (A) DPC and (B) DDM micelles (kcal-mol^{-1}) as a function of micelle size. Energies were calculated from simulations of pre-formed micelles except for the 30-molecule DPC system, which was calculated after micelle assembly.

Snapshots representative of micelle fluctuations produce similar SAXS curves. The WAXSiS server takes the user-provided structure and carries out a short timescale (20 ps) molecular dynamics simulation to generate a theoretical SAXS curve. WAXSiS uses the YASARA MD engine with the AMBER03 force field with TIP3P water model. Force field parameters for modified side chains and ligands are generated by YASARA's AutoSMILES method. To verify that the SAXS curves generated by the WAXSiS server are representative of an entire MD trajectory, we used *k*-means clustering on a sample trajectory to define three clusters with different I_1/I_3 ratios, where I_1 and I_3 are the largest and smallest principal moments of inertia (Fig. S3.4). At lower q values (0.15 ≤), the minimum and maximum are consistent for all three clusters. This region provides more detailed information about micelle size; the similarity between all three clusters confirms that each snapshot is representative of the entire production trajectory.

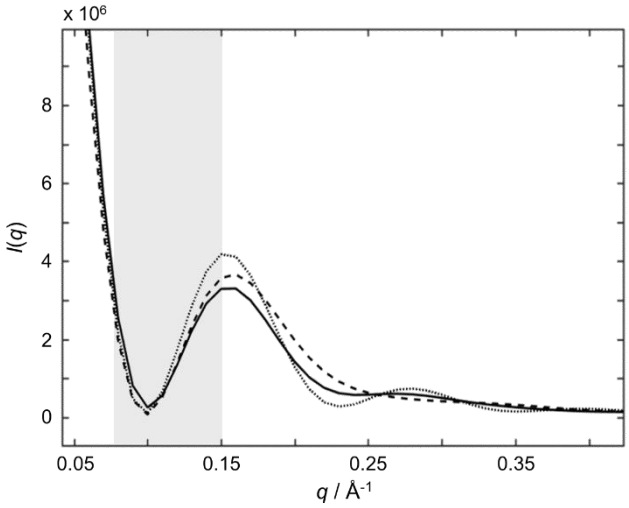


Figure S3.4. SAXS curves of micelle fluctuations show high degree of similarity in size region. SAXS curves of three snapshots of a 135-molecule preformed DDM micelle, with I_1/I_3 ratios of 1.18 (*black*), 1.29 (*dashed*), and 1.40 (*dotted*), respectively. Shaded area represents region of similarities for minimum and maximum of curves. Curves were obtained from WAXSiS web server.[30]

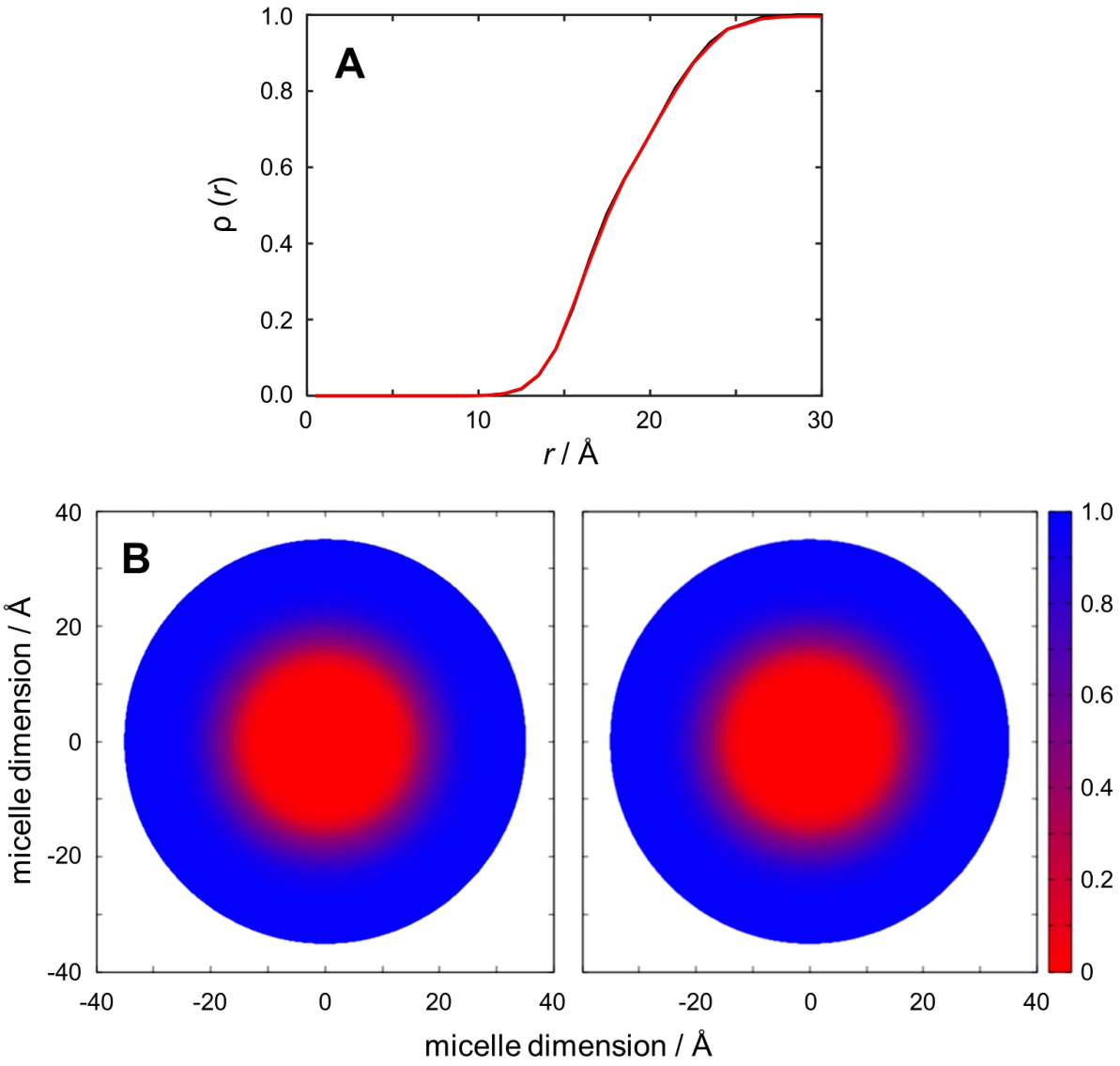


Figure S3.5. Radial distribution function of a micelle is independent of its initial configuration. (A). Normalized average radial distribution function curves of water for a preformed (*black*) and self-assembled (*red*) DPC micelle from 54 detergent molecules. Internal hydration (< 20 Å) is virtually indistinguishable between the two types of micelles. (B) Normalized water density for the preformed (*left*) and self-assembled (*right*) DPC micelles.

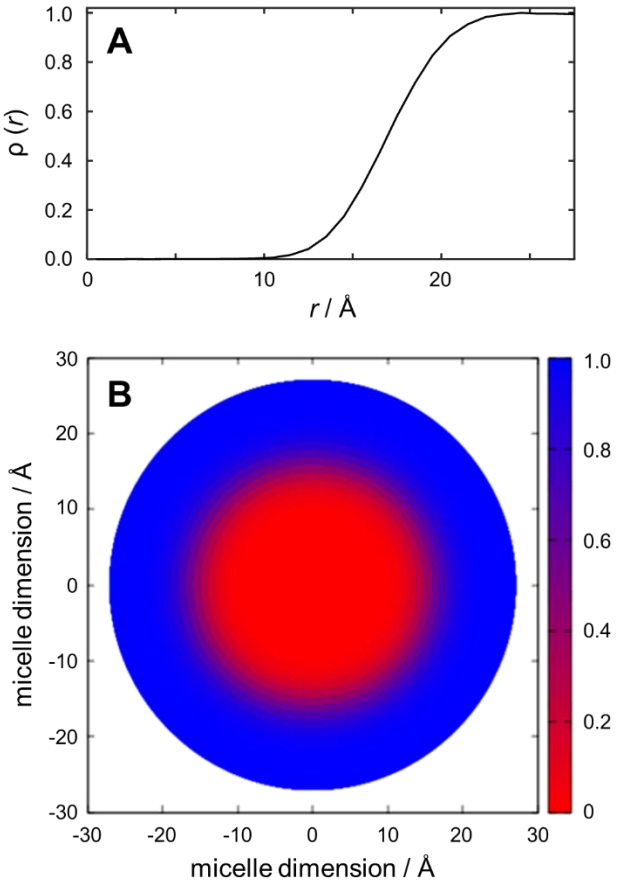


Figure S3.6. Sodium dodecanoate micelles have a dehydrated core. (A) Normalized average radial density function of a 64-molecule preformed micelle of sodium dodecanoate. The micelle core is completely dry as far as 10 Å from the micelle center of mass. Radius of gyration of the micelle is 14.54 ± 0.18 Å. (B) Normalized water density. The CHARMM c36 force field was used for these simulations.[60]

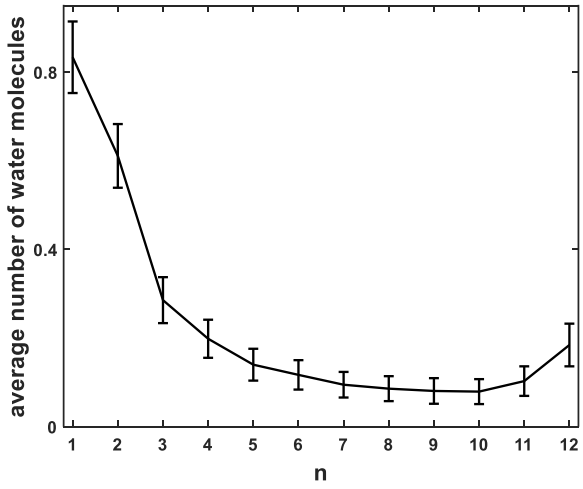


Figure S3.7. **Hydrocarbon tail of DDM micelle is partially hydrated, similar to DPC.** Average number of water molecules within 4 Å of the hydrophobic tail of DDM molecules within a micelle. The trend is similar to DPC micelle (Fig. 3.6 C).

Table S3.1 List of simulations carried out in this study.

Detergent[1]	*n*	# H2O / detergent	starting configuration	concentration (×CMC)[2]	# of runs	time (ns)
DPC	1	4170	n/a	12	1	40
DPC	30	313	random	162	1	40
DPC	54	278	random	182	10	2000
DPC	54	409	preformed	123	6	12000
DPC	70	30	random	1703	1	25
DPC	80	131	preformed	384	1	200
DPC	80	473	random	107	2	640
DPC	115	390	random	130	1	930
DPC	115	98	preformed	513	1	200
DPC	175	383	random	132	1	360
DPC	175	54	preformed	935	3	400
DPC	300	54	preformed	933	1	700
DDM	60	770	preformed	424	1	206
DDM	60	102	random	3196	1	206
DDM	80	716	preformed	456	1	206
DDM	80	529	preformed	618	1	200
DDM	110	478	preformed	683	1	200
DDM	110	215	random	1520	1	206
DDM	110	862	random	379	1	150
DDM	130	459	preformed	711	1	182
DDM	130	187	random	1747	1	120
DDM	135	196	preformed	1665	1	200
DDM	149	114	preformed	2868	1	200
DDM	200	346	preformed	944	1	159
DDM	200	225	random	1452	1	150
DDM	300	790	random	413	2	140
SD	64	88	preformed	25	1	160

[1] Abbreviations: DDM: *n*-dodecyl-β-D-maltoside; DPC: *n*-dodecylphosphocholine; SD: sodium dodecanoate.
[2] Concentrations are calculated by dividing the number of moles of detergents by volume of water molecules, assuming the density of water to be 0.9982 g/mL.

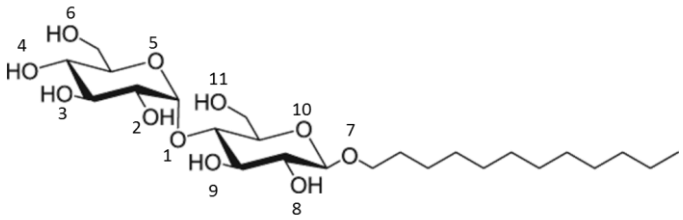
Table S3.2 Average principal moments of inertia for self-assembled DPC and DDM micelles. Moments of inertia are calculated after a stable micelle forms.

surfactant	*n*	I_1*	I_2	I_3	I_1/I_3	I_1/I_2	I_2/I_3	α
DPC	54	4.09 ± 0.37	3.77 ± 0. 35	3.31 ± 0. 13	1.23	1.08	1.14	0.10 ± 0.03
DDM	110	22.5 ± 0.7	20.1 ± 0.6	18.4± 0.5	1.22	1.12	1.09	0.10 ± 0.02

* All moments of inertia are x 10^6 amu-Å^2.

Table S3.3 Average number of hydrogen bonds between headgroups and headgroup-solvent for a DDM micelle with 130 detergent molecules.

	inter-headgroup		headgroup-solvent	
atom	donor	acceptor	donor	acceptor
O1		0.04±0.02		0.14±0.03
O2	0.28±0.04	0.47±0.05	0.83±0.05	1.11±0.07
O3	0.17±0.03	0.09±0.03	0.96±0.06	1.43±0.08
O4	0.13±0.03	0.17±0.03	1.02±0.06	1.08±0.07
O5		0.44±0.05		0.30±0.07
O6	0.31±0.05	0.30±004	0.77±0.06	1.31±0.07
O7		0.25±0.04		0.58±0.06
O8	0.26±0.04	0.23±0.03	0.59±0.05	0.86±0.07
O9	0.52±0.05	0.36±0.05	0.63±0.06	0.88±0.06
O10		0.14±0.03		0.30±0.04
O11	0.27±0.04	0.31±0.04	0.45±005	0.89±0.07
total	1.94	2.80	5.25	8.88
total	4.7		14.2	



References

1. Koley, D. & J. Bard, A. Triton X-100 concentration effects on membrane permeability of a single HeLa cell by scanning electrochemical microscopy (SECM). *Proc. Natl. Acad. Sci. U. S. A.* **107,** 16783–16787 (2010).
2. Hussain, S., Kinnebrew, M., S. Schonenbach, N., Aye, E. & Han, S. Functional Consequences of the Oligomeric Assembly of Proteorhodopsin. *J. Mol. Biol.* **427,** 1278–1290 (2015).
3. Saltzman, W. M. & Torchilin, V. P. Drug delivery systems. *AccessScience McGraw-Hill Educ.* (2014). doi:http://dx.doi.org/10.1036/1097-8542.757275
4. Shah, D. O. *Micelles, Microemulsions, and Monolayers*. (Marcel Dekker, Inc., 1998).
5. Hargreaves, R., Bowron, D. T. & Edler, K. Atomistic Structure of a Micelle in Solution Determined by Wide Q-Range Neutron Diffraction. *J. Am. Chem. Soc.* **133,** 16524–16536 (2011).
6. Lipfert, J., Columbus, L., B. chu, V., A. Lesley, S. & Doniach, S. Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering. *J Phys Chem B* **11,** 12427–12438 (2007).
7. Oliver, R. C. *et al.* Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS ONE* **8,** e62488 (2013).
8. Tieleman, D. P., van der Spoel, D. & Brendensen, H. J. C. Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. *J Phys Chem B* **104,** 6380–6388 (2000).
9. Cheng, X., Jo, S., Lee, H. S., Klauda, J. B. & Im, W. CHARMM-GUI Micelle Builder for Pure/Mixed Micelle and Protein/Micelle Complex Systems. *J. Chem. Inf. Model.* **53,** 2171–2180 (2013).

10. Chong, T. T., Hashim, R. & Bryce, R. A. Molecular dynamics simulation of monoalkyl glycoside micelles in aqueous solution: influence of carbohydrate headgroup stereochemistry. *J. Phys. Chem. B* **110,** 4978–4984 (2006).
11. Bruce, C. D., Berkowitz, M. L., Perera, L. & Forbes, M. D. E. Molecular Dynamics Simulation of Sodium Dodecyl Sulfate Micelle in Water: Micellar Structural Characteristics and Counterion Distribution. *J. Phys. Chem. B* **106,** 3788–3793 (2002).
12. Chun, B. J., Choi, J. I. & Jang, S. S. Molecular dynamics simulation study of sodium dodecyl sulfate micelle: Water penetration and sodium dodecyl sulfate dissociation. *Colloids Surf. Physicochem. Eng. Asp.* **474,** 36–43 (2015).
13. Abel, S., Dupradeau, F.-Y. & Marchi, M. Molecular Dynamics Simulations of a Characteristic DPC Micelle in Water. *J. Chem. Theory Comput.* **8,** 4610–4623 (2012).
14. Lee, S. *et al.* CHARMM36 United Atom Chain Model for Lipids and Surfactants. *J. Phys. Chem. B* **118,** 547–556 (2014).
15. Sanders, S. A., Sammalkorpi, M. & Panagiotopoulos, A. Z. Atomistic Simulations of Micellization of Sodium Hexyl, Heptyl, Octyl, and Nonyl Sulfates. *J. Phys. Chem. B* **116,** 2430–2437 (2012).
16. Marrink, S. J., deVries, A. H. & Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108,** 750–760 (2004).
17. Santos, A. P. & Panagiotopoulos, A. Z. Determination of the critical micelle concentration in simulations of surfactant systems. *J. Chem. Phys.* **144,** 044709 (2016).
18. Marrink, S. J., Tieleman, D. P. & Mark, A. E. Molecular Dynamics Simulation of the Kinetics of Spontaneous Micelle Formation. *J. Phys. Chem. B* **104,** 12165–12173 (2000).

19. Long, J. A., Rankin, B. M. & Ben-Amotz, D. Micelle Structure and Hydrophobic Hydration. *J. Am. Chem. Soc.* **137,** 10809–10815 (2015).
20. Mizuguchi, T., Ishizuka, R. & Matubayasi, N. Effect of diffuseness of micelle boundary on the solute distribution upon solubilization. *Chem. Phys. Lett.* **624,** 19–23 (2015).
21. Gruen, D. W. R. A model for the chains in amphiphilic aggregates. 2. Thermodynamic and experimental comparisons for aggregates of different shape and size. *J. Phys. Chem.* **89,** 153–163 (1985).
22. Svens, B. & Rosenholm, B. An investigation of the size and structure of the micelles in sodium octanoate solutions by small-angle X-ray scattering. *J. Colloid Interface Sci.* **44,** 495–504 (1973).
23. Menger, F. M., Jerkunica, J. M. & Johnston, J. C. The water content of a micelle interior. The fjord vs. reef models. *J. Am. Chem. Soc.* **100,** 4676–4678 (1978).
24. Hanwell, M. D. *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **4,** 1 (2012).
25. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30,** 2157–2164 (2009).
26. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38, 27–28 (1996).
27. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31,** 671–690 (2010).

28. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).
29. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
30. Knight, C. J. & Hub, J. S. WAXSiS: a web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics. *Nucleic Acids Res.* **43,** 225–230 (2015).
31. Romo, T. D. & Grossfield, A. LOOS: an extensible platform for the structural analysis of simulations. *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **2009,** 2332–2335 (2009).
32. The PyMOL Molecular Graphics System, version 1.3r1 (2010) Schrödinger, LLC, New York.
33. Voss, N. R. & Gerstein, M. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38,** W555–W562 (2010).
34. Torshin, I. Y., Weber, I. T. & Harrison, R. W. Geometric criteria of hydrogen bonds in proteins and identification of 'bifurcated' hydrogen bonds. *Protein Eng.* **15,** 359–363 (2002).
35. Göbl, C. *et al.* Influence of phosphocholine alkyl chain length on peptide-micelle interactions and micellar size and shape. *J. Phys. Chem. B* **114,** 4717–4724 (2010).
36. Lauterwein, J., Bösch, C., Brown, L. R. & Wüthrich, K. Physicochemical studies of the protein-lipid interactions in melittin-containing micelles. *Biochim. Biophys. Acta BBA - Biomembr.* **556,** 244–264 (1979).
37. Kallick, D. A., Tessmer, M. R., Watts, C. R. & Li, C. Y. The use of dodecylphosphocholine micelles in solution NMR. *J. Magn. Reson. B* **109,** 60–65 (1995).
38. Nyrkova, I. A. & Semenov, A. N. On the Theory of Micellization Kinetics. *Macromol. Theory Simul.* **14,** 569–585 (2005).

39. Zhang, J. & Liu, S. Kinetics of thermo-induced micelle-to-vesicle transitions in a catanionic surfactant system investigated by stopped-flow temperature jump. *Phys. Chem. Chem. Phys.* **13,** 12545–12553 (2011).
40. Israelachvili, J. N. *Intermolecular and surface forces*. (Academic Press, 2011).
41. Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *J. Chem. Soc., Faraday Trans. 2* **72,** 1525–1568 (1976).
42. Tanford, C. Micelle shape and size. *J. Phys. Chem.* **76,** 3020–3024 (1972).
43. Arora, A. & Tamm, L. K. Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol.* **11,** 540–547 (2001).
44. Andersson, M. *et al.* Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17,** 1265–1275 (2009).
45. Malmerberg, E. *et al.* Time-Resolved WAXS Reveals Accelerated Conformational Changes in Iodoretinal-Substituted Proteorhodopsin. *Biophys. J.* **101,** 1345–1353 (2011).
46. Svergun, D. I., Barberato, C. & Koch, M. H. J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28,** 768–773 (1995).
47. Grishaev, A., Guo, L., Irving, T. & Bax, A. Improved Fitting of Solution X-ray Scattering Data to Macromolecular Structures and Structural Ensembles by Explicit Water Modeling. *J. Am. Chem. Soc.* **132,** 15484–15486 (2010).
48. Chen, P. & Hub, J. S. Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-ray Scattering Data. *Biophys. J.* **107,** 435–447 (2014).

49. Fernández, C., Hilty, C., Wider, G. & Wüthrich, K. Lipid–protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. *Proc. Natl. Acad. Sci.* **99,** 13533–13537 (2002).

50. Columbus, L. *et al.* Mixing and matching detergents for membrane protein NMR structure determination. *J. Am. Chem. Soc.* **131,** 7320–7326 (2009).

51. Perkins, S. J. X-Ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits. *Biophys. Chem.* **93,** 129–139 (2001).

52. Courchene, W. L. Micellar Properties from Hydrodynamic Data. *J. Phys. Chem.* **68,** 1870–1874 (1964).

53. Vetter, R. J. Micelle Structure in Aqueous Solutions of Colloidal Electrolytes. *J. Phys. Colloid Chem.* **51,** 262–277 (1947).

54. Capponi, S., Freites, J. A., Tobias, D. J. & White, S. H. Interleaflet mixing and coupling in liquid-disordered phospholipid bilayers. *Biochim. Biophys. Acta BBA - Biomembr.* **1858,** 354–362 (2016).

55. Vikingstad, E. Partial molal volumes and compressibilities of n-alcohols in micellar solutions of sodium alkylcarboxylates. *J. Colloid Interface Sci.* **72,** 75–80 (1979).

56. Munusamy, E., Luft, C. M., Pemberton, J. E. & Schwartz, S. D. Structural Properties of Nonionic Monorhamnolipid Aggregates in Water Studied by Classical Molecular Dynamics Simulations. *J. Phys. Chem. B* **121,** 5781–5793 (2017).

57. Abel, S., Dupradeau, F.-Y., Raman, E. P., MacKerell, A. D. & Marchi, M. Molecular Simulations of Dodecyl-β-maltoside Micelles in Water: Influence of the Headgroup Conformation and Force Field Parameters. *J Phys Chem B* **115,** 487–499 (2011).

58. Bogusz, S., Venable, R. M. & Pastor, R. W. Molecular Dynamics Simulations of Octyl Glucoside Micelles: Dynamic Properties. *J. Phys. Chem. B* **105,** 8312–8321 (2001).

59. Bonincontro, A., Briganti, G., D'Aprano, A., La Mesa, C. & Sesta, B. Dielectric Behavior of Octyl β-d-Glucopyranoside Micelles in Water and in Water−Glycine Solutions. *Langmuir* **12,** 3206–3210 (1996).

60. Lee, S. *et al.* CHARMM36 United Atom Chain Model for Lipids and Surfactants. *J. Phys. Chem. B* **118,** 547–556 (2013).

Chapter 4

4. Effect of Mild and Harsh Detergents on the Stability of the Model Bacterial Membrane Protein, Proteorhodopsin

4.1 Introduction

Surfactants are widely used in membrane protein extraction. The surfactant used for the purification of a protein may alter its structure and function[1]. As an example, PR is known to be present predominantly as a monomer in DPC and a hexamer in DDM[2], as mentioned in Chapter 1. In addition, the pK_a of the proton acceptor, D97, decreases and the photocycle slows down in the oligomeric state. Also, PR-DPC complexes are not stable for more than a week[2]. DPC is a harsh detergent, and it is expected to destabilize the protein, as discussed in Section 1.3. It is not clear how surfactants affect the photocycle or even how they determine the oligomeric state of PR.

The oligomerization of PR is also observed in 1,2-dioleolyl-*sn*-glycero-3-phosphocholine (DOPC) membrane with high resolution atomic force microscopy (AFM), where the hexamers are dominant[3]. Solid state NMR experiments show that salt bridges, i.e., electrostatic interactions between charged amino acid side chains, must exist between Asp/Glu and Arg/Lys in the oligomeric interfaces[4]. In green variant of PR, a salt bridge between R51 from a monomer and D52 from the neighboring monomer favors the formation of a pentamer; mutation of R51 or D52 destroys the pentamers and results in formation of hexamers. A neighboring glutamic acid, E50, plays an essential role in oligomerization, possibly via a hydrogen bond with R51[4]. In the blue PR structure studied here (PDB 4JQ6[5]), the corresponding residue set for E50-R51-D52 is E32-R33-S34-D35. One would expect that having S33 would make the position of D35 less favorable for a salt bridge with R33 from a neighboring monomer to form a pentamer. Therefore, pentamers are not dominant oligomers in blue PR, and possibly a hexamer rearrangement will allow the formation of an Arg-Asp salt bridge[4].

To investigate the effect of the surfactant on the structure of membrane proteins, we have studied the blue PR in DDM and DPC detergent environments, as examples of mild and harsh surfactants, respectively, with

equilibrium MD simulations. The starting structure of BPR is from the atomic coordinates from the X-ray structure[5]. The proteomicelles are in aqueous solutions. Our results show that PR is unstable in DPC, in agreement with experimental observations. The surfactant - water interactions play an essential role in the stabilization of the solvated protein, that helps in explaining the degree of mildness in surfactants. DPC molecules interact with water more than DDM. As a result, DPC does not effectively shield the hydrophobic part of PR from water molecules, and destabilizes PR. Also, the side chain orientation of R51 is different in two surfactant environments. Although our systems are monomers, the insight we get will help understanding the stability and rearrangement of oligomers. Specially, the orientation of the oligomeric interface residues, E32-R33-S34-D35, and their interaction with surrounding surfactants provide explanations for the preferred oligomeric states in different surfactant environment. Finally, the solvent accessibility and side chain orientation of D97 is different in two surfactant environment, that may help explaining the effect of micelles on its pK_a.

4.2 Computational Methods

Simulation setup. The X-ray crystal structure of BPR from Med12 in the dark state (PDB 4JQ6)[5] was used as the starting structure. Missing loops were modeled based on the NMR structure of GPR (PDB 2L6X)[6]. The N- and C- termini have been truncated so that helix A starts with D9 and helix G ends with S231. We use the residue numbering of the green PR. Based on our earlier simulations of PR[7], we protonated H75 on the N_δ atom of the imidazole ring. Also, E142 is deprotonated since its pK_a is reported to be < 7.5 in blue PR[8].

The PR-DDM system is composed of BPR with 284 surfactant molecules, 161 sodium ions, and 156 chloride ions in a 140 $\times$ 141 $\times$ 136 Å water box. The PR-DPC has 283 surfactant molecules, 105 sodium ions and 102 chloride ions in a 125 $\times$ 123 $\times$ 120 Å water box. Packmol[9] is used for building the initial surfactant corona around BPR. The minimum water thickness between the proteomicelles and the edges of the periodic box was 20 Å. Although the aggregation number of the surfactants are much lower, i.e., 149 for DDM and 79 for DPC[10,11], we added extra surfactant molecules to make sure that all of the hydrophobic area of the protein is covered. We assume that all extra surfactants will move apart from the protein as simulations run. The pure DDM micelle has 149 DDM molecules, with 100 mM NaCl in a 150 $\times$ 125 $\times$ 114 Å water box. The pure DPC micelle has 80 DPC molecules, with 100 mM NaCl in a 73 $\times$ 70 $\times$ 71 Å box. We also used a PR in lipid bilayer as a control system. 160 lipid molecules (80 per leaflet) with a

molar ratio of 3:1 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE):1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG), were built around the protein using the CHARMM-GUI server (http://www.charmm-gui.com). The system was neutralized with adding 44 sodium ions. The total number of atoms is 47,288.

The proteomicelle systems were minimized for 1 ps, followed by 4 heating steps at 50, 100, 200 and 310 K, each 0.5 ns long with 10 kcal/mol constraints on the protein backbone. The minimization and heating were done with NAMD 2.11[12] using CHARMM c36 force field for lipids and proteins[13]. Then, the structures and force field were converted to AMBER-compatible formats using CHAMBER[14], and all production simulations were run with AMBER 12 and AMBER 16 packages with GPU acceleration[15,16]. The initial 140 ns simulations were run with protein backbone being constrained to 10 kcal/mol. The constraints were decreased to 5, 4, 3, 2, 1, 0.8, 0.5, 0.4, 0.3, 0.2, 0.1 and 0.0 kcal/mol, respectively, with each step being 20 ns long. All equilibrium MD simulations were run in the *NPT* ensemble at 310 K and 1 bar using a Langevin thermostat and a Berendsen barostat with a 1 fs timestep. A cutoff of 12 Å and a force switching of 10 Å were applied for non-bonded interactions.

Analysis. VMD[17] was used for analysis of angles, inter-atomic distances, solvent accessibility, hydrogen bonds, root mean squared fluctuation (RMSF), energy and image rendering. Average density maps of water are calculated with MDAnalysis[18]. A cutoff distance of 3 Å and an angle threshold of 20 degrees between the donor heavy atom - hydrogen - acceptor heavy atom is used in the analysis of hydrogen bonds.

4.3 Results and Discussion

Dynamics of proteomicelles is explained by nonbonded interactions. Visual inspection of the trajectories shows a clear difference in the shape and dynamics of the two proteomicelles. In 1 μs, PR-DDM proteomicelle stays in an almost uniform structure as seen from the snapshots in Fig. 4.1 A and the RMSF plot in Fig. S4.1 B. The DPC corona however drastically changes to a non-uniform micellar structure. The plots of the radius of gyration in Fig. S4.1 A also show that the PR-DPC system is generally larger and more dynamic than DDM, although the size of the pure DPC micelle is smaller than DDM. This rapid expansion of the PR-DPC complex is mainly due to the translational velocity of DPC molecules. The average mean square deviation, MSD, of surfactant headgroups in Fig. 4.3 shows this difference. In a μs, the average deviation of a DPC detergent from its initial coordinate is greater than DDM. As seen from the slope of the curves in Fig. 4.3, the velocity of molecules is also greater for DPC. As a result, a greater area of PR

is exposed to water in DPC environment. The cylindrical density of the proteomicelles in Fig. 4.2 shows that DPC makes a thinner layer around PR, compared to DDM or bilayer. The lateral thickness of DPC proteomicelle along xy plane is also the smallest between the three.

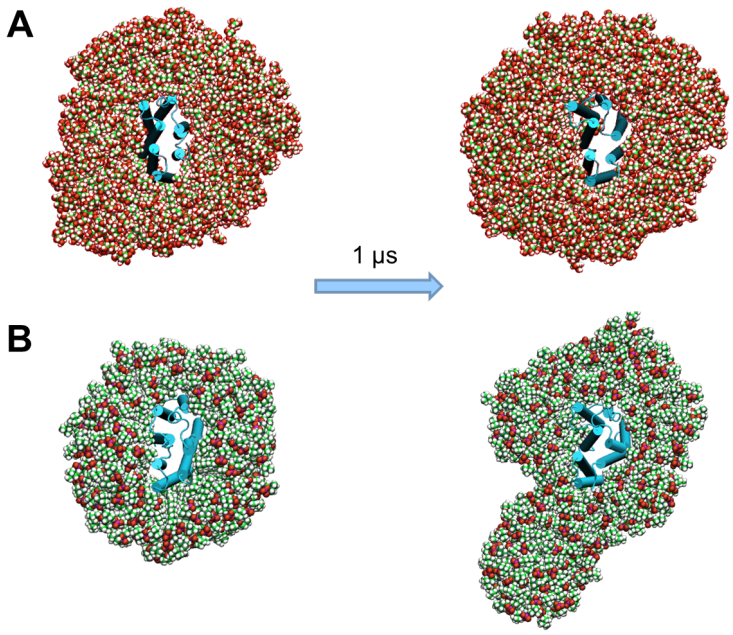


Figure 4.1. PR-DPC proteomicelle is less stable than PR-DDM. (A) There is a negligible difference between the starting and final shape of PR-DDM micelle. (B) In contrast, the DPC corona drastically changes to a non-uniform thickness around PR. This is in agreement with the experiment where PR-DPC micelles are not stable for more than a week.

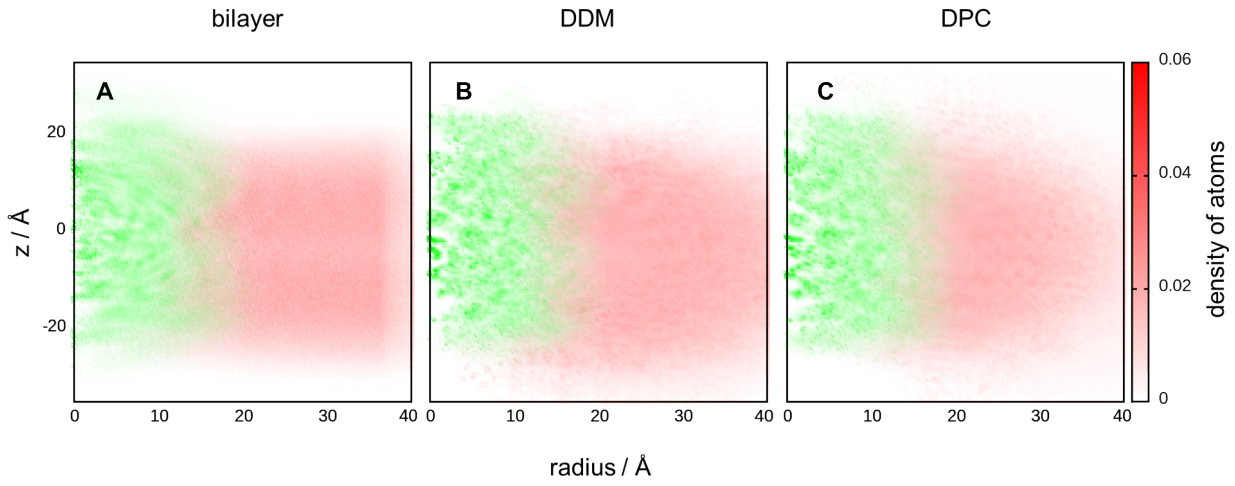


Figure 4.2. DPC covers transmembrane part of PR less than bilayer and DDM. Cylindrical density profiles of lipids and detergents along the membrane normal in bilayer (A), DDM (B) and DPC (C), shown in red. The green area is PR that is aligned to Z axis.

In the solvation of the membrane proteins, the surfactant-protein and surfactant-water interactions are critical. Therefore, one would consider the balance between these two sets of interactions. Fig. 4.4 shows the sum of the Vand der Waals, electrostatic and non-bonded interactions for surfactant - protein and surfactant - water molecules in the two systems. The protein - surfactant interactions in two systems are comparable. However, DPC - water interactions are almost an order of magnitude more negative that DDM - water, despite the smaller size of the DPC headgroup.

The higher DPC - water affinity is also reflected on the solvent accessibility of DPC molecules, specially the acyl chains (Fig. 4.5).

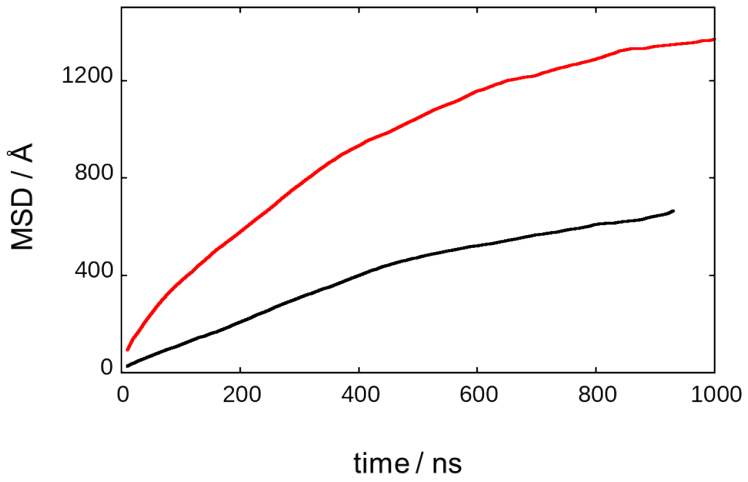


Figure 4.3. DPC molecules are more dynamic than DDM. Average MSD of the etheric oxygen of DDM (*black*) and phosphorous atom of DPC (*red*) headgroups in proteomicelles. DPC molecules move for longer distances than DPC. Also, the slope of MSD is higher for DPC, indicating a higher translational velocity.

Strong interactions between DPC and water compete with the DPC - PR interactions. As a result, DPC molecules leave the hydrophobic part of the protein more exposed to water, and destabilize the protein. The destabilizing effect is seen from the RMSF plots (Fig. S4.1 B). In other words, DPC prefers the water molecules to the protein, and consequently, the helical part of the protein is less protected against polar water molecules. It seems intuitive that the ionic or zwitterionic surfactants destabilize the protein by interacting with the ionic residues on the protein, but our results show that the electrostatic interactions between the surfactant and water molecules are more important.

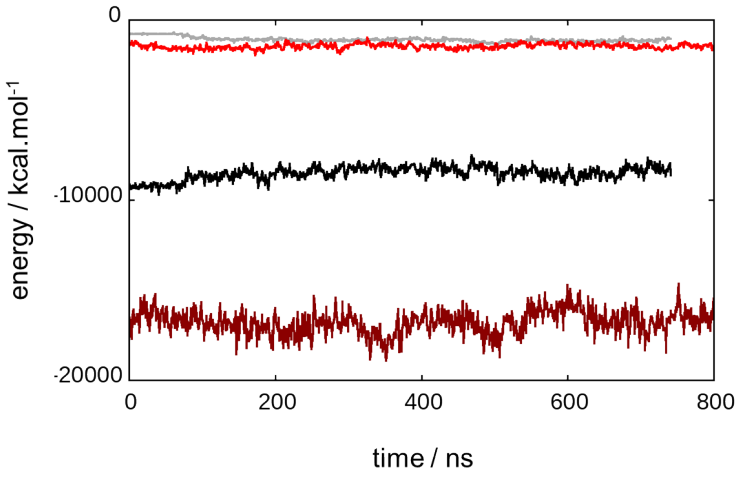


Figure 4.4. Water interaction with DPC is more favorable than DDM. The sum of VdW, electrostatic and non-bonded interaction energies between PR and DDM (*gray*) and DPC (*red*) are close. However, DPC – water interactions (dark red) are almost an order of magnitude more negative than DDM (*black*), despite the smaller size of DPC head-group. Only the surfactant molecules within 20 Å of the protein are considered. The water - DPC affinity is a reason for destabilizing the extracted membrane proteins. The data are collected after the constraints are removed from PR backbone.

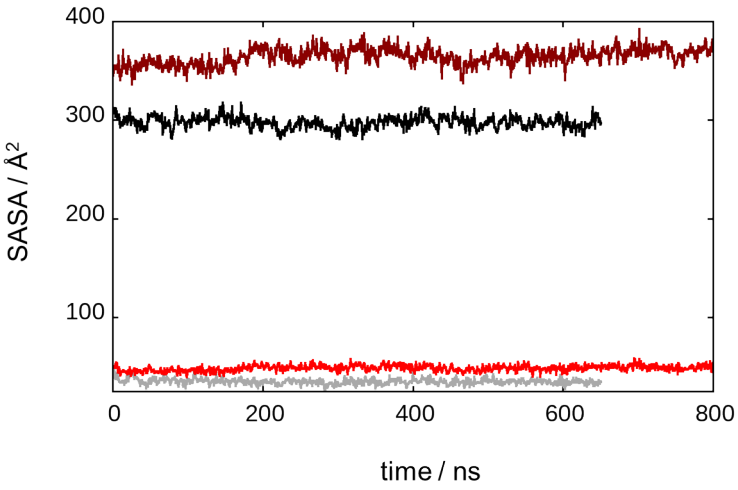


Figure 4.5. DPC is more exposed to water than DDM. In line with the energy analysis, the higher interaction of DPC molecules with water is reflected on the SASA per detergent molecule for DPC headgroup (*red*) and acyl chain (*dark red*) that is higher than DDM (*gray* and *black*, respectively). Since DDM and DPC acyl chains have the same length, this difference is attributed to the headgoup. Only the surfactant molecules within 20 Å of the protein are considered. The data are collected after the constraints are removed from PR backbone.

Potential influence of surfactant environment on the oligomerization. Among the residues that are proposed to be critical in oligomerization, R51 has the longest side chain, and as mentioned earlier, it mediates the formation of the pentamer with formation of a salt bridge with D52 from the neighboring

monomer. Fig. 4.6 shows that in the DDM and bilayer environment, the side chain of R51 Yo predominantly is oriented towards the cytoplasmic bulk solvent. However, in DPC the upward orientation of the side chain is dominant. This orientation might be due to the electrostatic interactions between the positively charged R51 side chain and the negatively charged oxygen atoms of the DPC headgroup. The radial pair distribution function between the R51 side chain and the polar atoms in DPC headgroup has a peak around 2.5 Å that is larger than DDM, indicating stronger interactions between DPC and R51 (Fig. S4.3). Interestingly, the plot of DDM is more similar to bilayer than DPC. We did not find a significant difference between side chain orientation of the other oligomerization residues in the two proteomicelles.

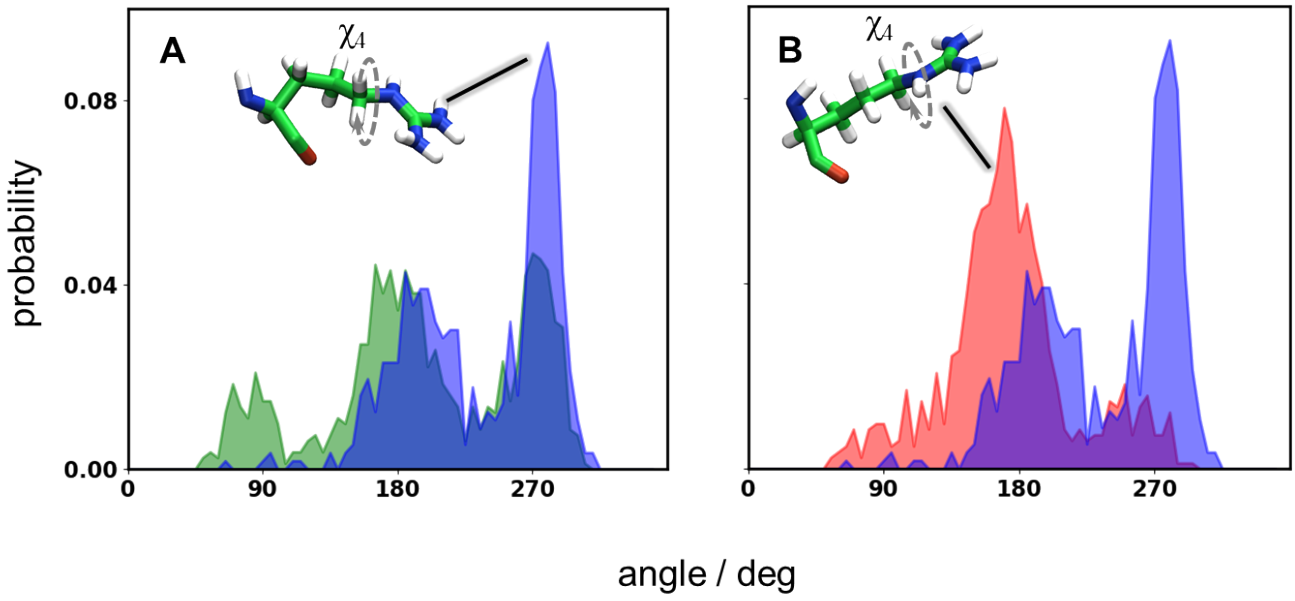


Figure 4.6. The side chain orientation of R51 in PR-DDM is similar to the lipid bilayer and different from PR-DPC. (A) The distribution of the χ_4 dihedral angle of R51 in DDM (*blue*) is similar to bilayer (*green*). The similar side chain orientation might correspond to the similar oligomeric preference. (B) In contrast, the χ_4 dihedral angle in DPC environment has a broad distribution around 180 degrees corresponding to the upward orientation of the side chain.

Influences on the proton acceptor. As mentioned earlier, the oligomerization slows down the conformational changes in protein that occur during photocycle, and it decreases the pK_a of D97[2]. A possible reason for this difference is the effect that the micelle has on the dynamics of protein backbone and the water molecules in protein interior in the oligomer. We saw that PR backbone is more dynamic in DPC, where the monomer is the dominant species. PR in DDM is more stable. In addition, DPC allows more wa-

ter molecules to diffuse inside protein, that potentially mediates the proton exchange between key residues in photocycle[7]. Comparing the SASA of D97 in the two surfactant environment shows a higher solvent accessibility of D97 in DPC (Fing 4.7 A). Also, a clear difference in the side chain orientation of D97 is observed (Fig. 4.7 B). In DPC, the D97 side chain kinks towards the periplasmic water molecules. Interestingly, for the most of time there is a hydrogen bond between the carboxylic oxygen atoms of D97 and H75 in both systems (Fig. S4.5). The accelerating effect of DPC on the photocycle can also be explained by the destability of PR in DPC. As seen in other retinal proteins, detergent - protein interactions can displace the native interactions and mediate the transition to the active forms[19]. Therefore, the speed of the photocycle may be directly related to the RMSF the protein backbone.

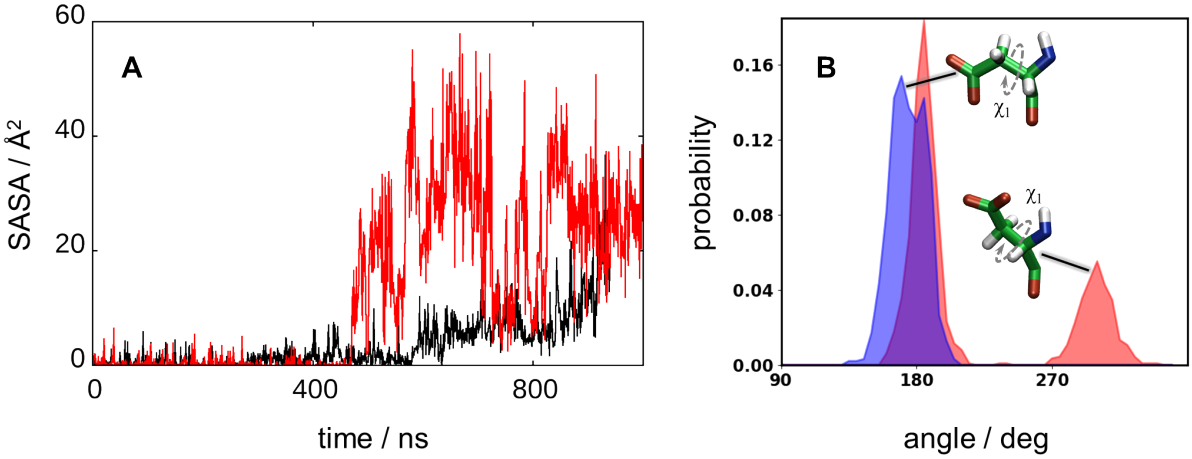


Figure 4.7. D97 is more hydrated in DPC. (A) The SASA of D97 is almost an order of magnitude higher in DPC (*red*) than DDM (*black*). The difference in pKa of D97 in two surfactant environment can be explained by the difference in solvent accessibility. (B) The χ_1 dihedral angle of D97 in DDM environment has a uni-modal distribution (*blue*). In DPC (*red*), another peak is observed at 300°, corresponding to the upward orientation of the side chain, that allows for nonbonded interactions between the terminal carboxylic group and the periplasmic water molecules.

4.4 Conclusion

Dynamics of PR in DDM and DPC micelles has been studied. Our results show that interactions between surfactant and water molecules has a critical role in the stability of the membrane protein. This might be an explanation for the destabilizing effect of ionic and zwitterionic surfactants that is observed in experiment. The micelle also affects the oligomerization state of PR possibly by altering the side chain orientation of R51. DPC micelle also allows for diffusion of more water molecules than DDM. The increased hydration potentially facilitates the proton exchange between ionizable residues inside PR, and speeds up the photo-

cycle, as known from experiment. The work presented here explains the phenomena observed in protein extraction with surfactants in atomistic details, and it helps in selecting the efficient surfactants for purification purposes.

Supplemental Data

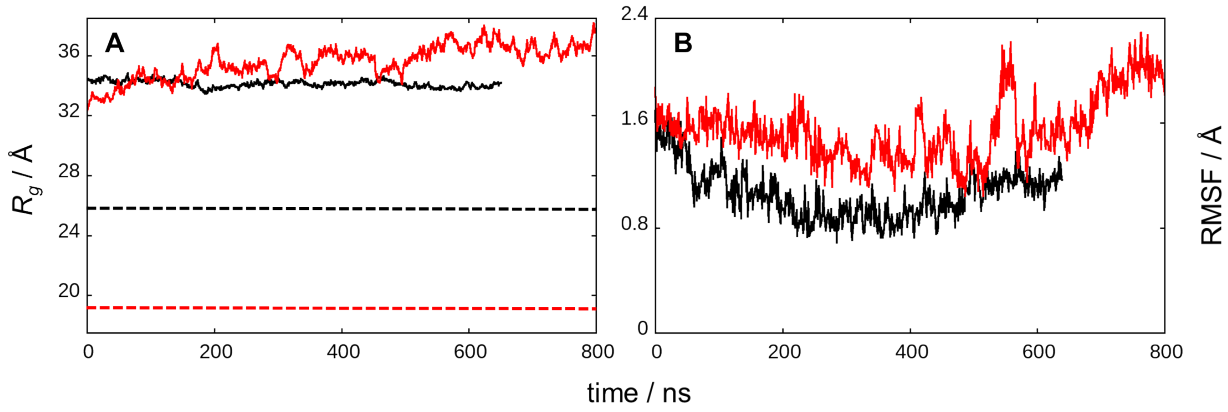


Figure S4.1. Overall size of proteomicelle complex is consistent regardless of detergent, but less stable in DPC. (A) PR-DPC (*red*) system has a dynamic radius of gyration (*Rg*) that is generally larger than PR-DDM (*black*). The dashed lines are *Rg* of corresponding detergent-only micelles from Ref [[20]]. (B) RMSF of the PR backbone in DDM (*black*) and DPC (*red*). The data are collected after the constraints on PR backbone are removed.

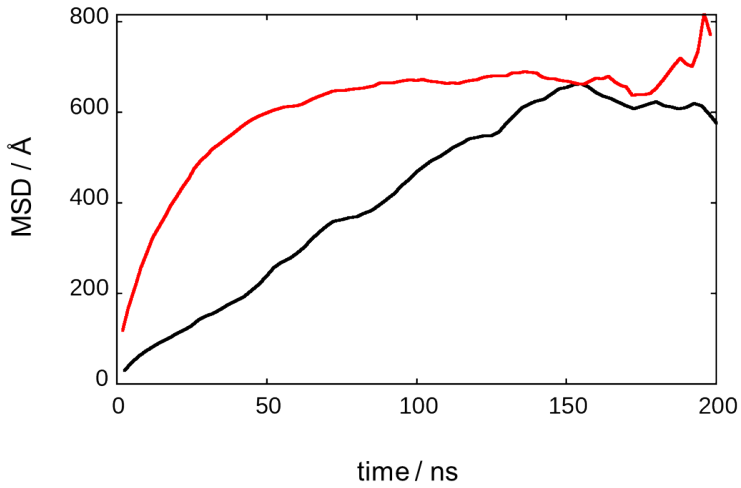


Figure S4.2. Surfactants in a pure micelle are generally more dynamic than proteomicelle. Average MSD of the etheric oxygen of DDM (*black*) and phosphorous atom of DPC (*red*) headgroups in pure micelles with 149 and 80 molecules, respectively. The trend of MSD is less clear than proteomiclles, although DPC molecules are more dynamic.

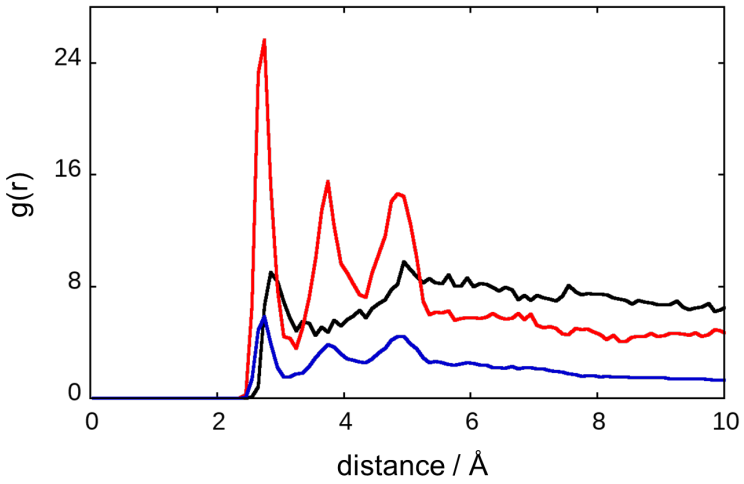


Figure S4.3. DPC headgroup has stronger interactions with R51 side chain than DDM. Radial pair distribution function between the R51 side chain and the polar atoms in DPC headgroup (*red*) has a peak around 2.5 Å, indicating an electrostatic interaction. The plot for DDM (*black*) is smoother and more similar to bilayer (*blue*). Detergents within 30 Å of protein are included in calculations.

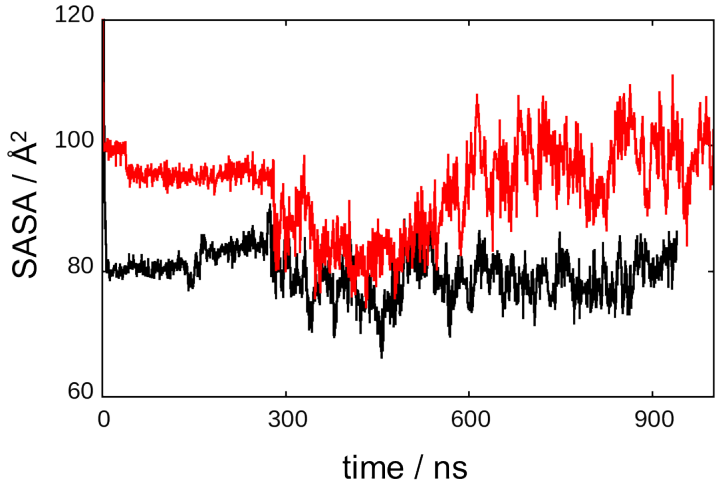


Figure S4.4 D97 is more hydrated in DPC. SASA of PR in DPC (*red*) is higher than DDM (*black*). The higher SASA is a sign of disability of PR in DPC.

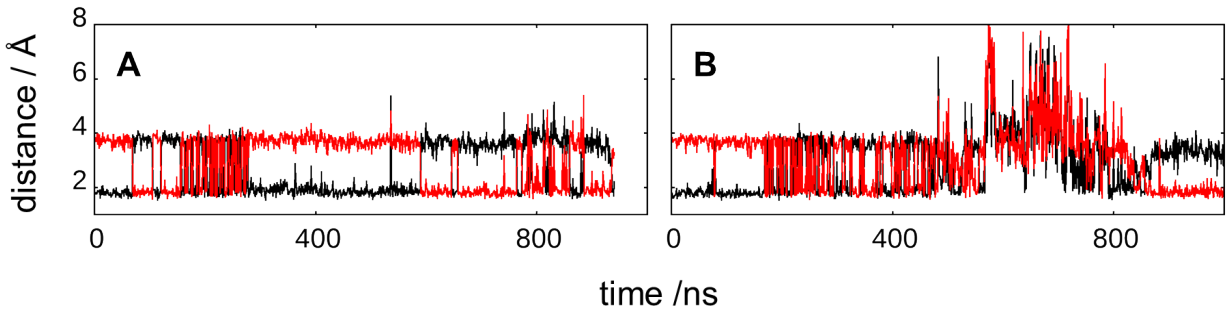


Figure S4.5 Despite the side chain orientation of D97, there is a stable hydrogen bond between D97 and H75 side chains. The distance between OD1 (*black*) and OD2 (*red*) of D97 side chain and HD1 of H75 in DDM (A) and DPC (B) environments. For the most of time, D97 provides an oxygen for the hydrogen bond.

References

1. Oliver, R. C., Pingali, S. V. & Urban, V. S. Designing Mixed Detergent Micelles for Uniform Neutron Contrast. *The Journal of Physical Chemistry Letters* **8**, 5041–5046 (2017). URL http://dx.doi.org/10.1021/acs.jpclett.7b02149.

2. Hussain, S., Kinnebrew, M., S. Schonenbach, N., Aye, E. & Han, S. Functional Consequences of the Oligomeric Assembly of Proteorhodopsin. *Journal of Molecular Biology* **427**, 1278–1290 (2015).

3. Klyszejko, A. L. *et al.* Folding and Assembly of Proteorhodopsin. *Journal of Molecular Biology* **376**, 35–41 (2008). URL http://www.sciencedirect.com/science/article/pii/S002228360701501X.

4. Maciejko, J. *et al.* 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* **137**, 9032–9043 (2015). URL http://dx.doi.org/10.1021/jacs.5b03606.

5. Ran, T. *et al.* Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta Crystallographica Section D: Biological Crystallography* **69**, 1965–1980 (2013). URL http://scripts.iucr.org/cgi-bin/paper?be5237.

6. Reckel, S. *et al.* Solution NMR Structure of Proteorhodopsin. *Angewandte Chemie International Edition* **50**, 11942–11946 (2011). URL http://onlinelibrary.wiley.com/doi/10.1002/anie.201105648/abstract.

7. Feng, J. & Mertz, B. Proteorhodopsin Activation Is Modulated by Dynamic Changes in Internal Hydration. *Biochemistry* **54**, 7132–7141 (2015). URL http://dx.doi.org/10.1021/acs.biochem.5b00932.

8. Amsden, J. J. *et al.* Different Structural Changes Occur in Blue- and Green-Proteorhodopsins during the Primary Photoreaction. *Biochemistry* **47**, 11490–11498 (2008). URL https://doi.org/10.1021/bi800945t.

9. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **30**, 2157–2164 (2009).

10. Oliver, R. C. *et al.* Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS ONE* **8**, e62488 (2013). URL http://dx.doi.org/10.1371/journal.pone.0062488.

11. Lipfert, J., Columbus, L., B. chu, V., A. Lesley, S. & Doniach, S. Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering. *J. Phys. chem B* **11**, 12427–12438 (2007).

12. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26**, 1781–1802 (2005). URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.20289/abstract.

13. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **31**, 671–690 (2010). URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.21367/abstract.

14. Crowley, M. F., Williamson, M. J. & Walker, R. C. CHAMBER: Comprehensive support for CHARMM force fields within the AMBER software. *International Journal of Quantum Chemistry* **109**, 3767–3772 (2009). URL http://onlinelibrary.wiley.com/doi/10.1002/qua.22372/abstract.

15. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **9**, 3878–3888 (2013).

16. D.A. Case,. R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.

17. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38, 27–28 (1996).

18. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry* **32**, 2319–2327 (2011). URL http://onlinelibrary.wiley.com/doi/10.1002/jcc.21787/abstract.

19. Grossfield, A., Feller, S. E. & Pitman, M. C. A role for direct interactions in the modulation of rhodopsin by ω-3 polyunsaturated lipids. *Proceedings of the National Academy of Sciences* **103**, 4888–4893 (2006). URL http://www.pnas.org/content/103/13/4888.

20. Faramarzi, S. *et al.* Molecular Dynamics Simulations as a Tool for Accurate Determination of Surfactant Micelle Properties. *Langmuir* **33**, 9934–9943 (2017). URL http://dx.doi.org/10.1021/acs.langmuir.7b02666.

Chapter 5

5. The Photosensitive BZ Reaction

5.1 Phase Oscillators and Synchronization

For any oscillating system that periodically returns to a reference state, one can define the phase as the state of the system with respect to the reference state. A set of phase oscillators that influence each other can reach a state in which their oscillations are in harmony or synchronized. Synchronization is the adjustment of the rhythms of oscillating systems due to the mutual interaction or an external force[1], and it is observed in many natural systems, such as heart cells[2], firing in populations of fireflies[3] and the human brain[4]. The first reported study of synchronization is the work of Hugens on two coupled pendulum clocks suspended from a wood beam[5]. He observed that two pendulums become frequency synchronized after a period of time, and even if they are perturbed, the synchronization state is recovered in a short time. He found that the synchronization is simply due to the motion of the beam.

Later, Kuramoto developed a simple model for phase oscillators and introduced a threshold for the coupling strength, $K_{critical}$, above which they start to synchronize[6]. In his model, the phase dynamics of an oscillator i, θ_i, in a set of n oscillators is described by the following equation:

$$\dot{\theta}_i = \omega_i - \frac{K}{n}\sum_{j=1}^{n} \sin(\theta_i - \theta_j), \qquad 5.1$$

where θ_j is the phase of oscillator j, ω_i is the natural frequency and K is the coupling strength. A set of coupled phase oscillators starts to synchronize once K reaches a critical threshold, called critical coupling strength, $K_{critical}$, which is a function of the natural frequency distribution[7].

In order to quantify the degree of the synchronization of a set of phase oscillators, Kuramoto introduced the order parameter, r: [6]

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{n} e^{i\theta_j}, \qquad 5.2$$

where θ_j is the phase of oscillator j, ψ is the average phase, $i = \sqrt{-1}$ and r is a number between 0 and 1 depending on the degree of synchronization. If we represent the phase of each oscillator j in a circle in the complex plane, the average phase will be the angle between the r vector and the x axis, while the magnitude of r is the degree of synchronization, as shown in Fig. 5.1.

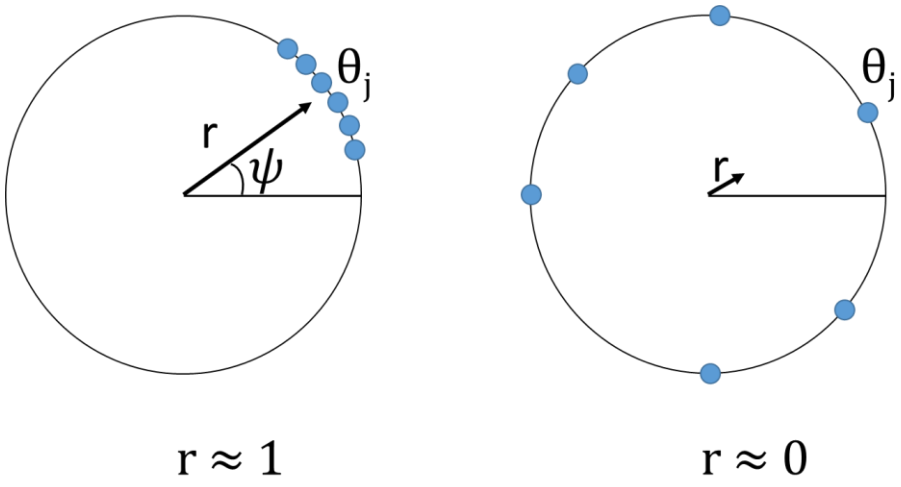


Figure 5.1. Synchronization of phase oscillators is defined by the order parameter. The θj are the phases of individual oscillators and ψ is the average phase. The magnitude of the order parameter, r, shows the degree of the synchronization in a system.

5.2 Chemical Oscillators and the Belouzov-Zhabotinsky Reaction

5.2.1 The Belouzov-Zhabotinsky Reaction Mechanism

There is a group of oscillating chemical reactions in which the concentration of some species changes periodically. Among them, the Belouzov-Zhabotinsky (BZ) reaction has been extensively studied in recent decades[8]. The BZ reaction is a complex multi-step reaction that involves the oxidation of an organic compound, malonic acid, for example, in the presence of a metal complex, usually ferroin or ruthenium. The net reaction is represented as follows:

$$2BrO_3^- + 3CH_2(COOH)_2 + 2H^+ \rightarrow 2BrCH(COOH)_2 + 3CO_2 + 4H_2O. \quad 5.3$$

Field, Körös and Noyes have proposed a three step mechanism for the BZ reaction, known as the FKN mechanism[9]. The mechanism involves the consumption of bromide ion, followed by the autocatalytic production of bromous acid, and the oxidation of the organic species that is associated with the reproduction of bromide ion. Having bromide ion from the last step, the oscillatory reaction is reset and another cycle starts. Net reactions of the three steps are as follows:

Process A

$$BrO_3^- + 5Br^- + 6H^+ \rightarrow 3Br_2 + 3H_2O, \quad 5.4$$

Process B

$$BrO_3^- + HBrO_2 + 2M^{2+} + 3H^+ \rightarrow 2HBrO_2 + 2M^{3+} + H_2O, \quad 5.5$$

Process C

$$2M^{3+} + MA + BrMA \rightarrow 2M^{2+} + fBr^- + (\text{other products}), \quad 5.6$$

where M^{n+} is the metal catalyst, BrMA is brominated malonic acid and f is a stoichiometric coefficient. When the concentration of bromide drops to a critical value in Process A, the autocatalytic production of bromous acid, Process B, starts. However, when bromous acid concentration reaches its maximum, the following redox reaction inhibits further production of bromous acid

$$2HBrO_2 \rightarrow HBrO_3 + HOBr, \quad 5.7$$

therefore, with accumulation of M^{3+}, Process C can take place, allowing for the reproduction of bromide ion that resets the reaction.

Later on, Zhabotinsly, Bochholtz, Kiyatkin and Epstein developed the FKN mechanism to a more detailed reaction scheme, known as ZBKE mechanism[10] (Table 1).

Table 5.1. ZBKE reaction model from Ref. [10].

$HOBr + Br^- + H^+ \rightleftarrows Br_2 + H_2O$	R1
$HBrO_2 + Br^- + H^+ \rightleftarrows 2HOBr$	R2
$HBrO_3 + Br^- + H^+ \rightleftarrows HBrO_2 + HOBr$	R3
$2HBrO_2 \rightleftarrows HOBr + HBrO_3$	R4a
$HBrO_2 + H^+ \rightleftarrows H_2BrO_2^+$	R4b
$HBrO_2 + H_2BrO_2^+ \rightleftarrows HOBr + HBrO_3 + H^+$	R4c
$H^+ + BrO_3^- \rightleftarrows HBrO_3$	R5a
$HBrO_3 + H^+ + HBrO_2 \rightleftarrows HBrO_2^+ + BrO_2^\bullet + H_2O$	R5b
$BrO_2^\bullet + H^+ \rightleftarrows HBrO_2^+$	R5c
$M^{n+} + HBrO_2^+ \rightleftarrows M^{(n+1)} + HBrO_2$	R6
$M^{(n+1)} + CHBr(COOH)_2 \rightleftarrows M^{n+} + CBr(COOH)_2^\bullet + H^+$	R7
$H_2O + CBr(COOH)_2^\bullet \rightarrow H^+ + Br^- + COH(COOH)_2^\bullet$	R8
$H_2O + CHBr(COOH)_2 \rightarrow CHOH(COOH)_2 + H^+ + Br^-$	R9
$2COH(COOH)_2^\bullet \rightarrow CHOH(COOH)_2 + CO(COOH)_2$	R10
$COH(COOH)_2^\bullet + CBr(COOH)_2^\bullet \rightarrow CHBr(COOH)_2 + CO(COOH)_2$	R11
$HOBr + CHBr(COOH)_2 \rightarrow CBr_2(COOH)_2 + H_2O$	RA-1
$Br_2 + CHBr(COOH)_2 \rightarrow CBr_2(COOH)_2 + H^+ + Br^-$	RA-2

Zhabotinsly and co-workers assumed that RA-1 and RA-2 are fast enough to neglect all other reactions of HOBr and Br_2, and also that R5c considerably shifts to the right[10]. With these assumptions they developed the mathematical equations to model the reactions listed in Table X[11].

Table 5.2. Simplified ZBKE reaction.

$X + Y + H \xrightarrow{k_2} 2\mathrm{P}$	5-8
$A + Y + H \xrightarrow{k_3} \mathrm{X} + \mathrm{P}$	5-9
$2X \xrightarrow{k_4} \mathrm{P} + \mathrm{A}$	5-10
$A + X + H \underset{k_{-5}}{\overset{k_5}{\rightleftarrows}} 2\mathrm{U}$	5-11
$U + (C - Z) \underset{k_{-6}}{\overset{k_6}{\rightleftarrows}} \mathrm{X} + \mathrm{Z}$	5-12
$Z + B \underset{k_{-7}}{\overset{k_7}{\rightleftarrows}} (C - Z) + \mathrm{B}^{\bullet} + \mathrm{H}$	5-13
$\mathrm{B}^{\bullet} \xrightarrow{k_8} \mathrm{qY} + \mathrm{Q}$	5-14
$B \xrightarrow{k_9} \mathrm{Y}$	5-15

$$\frac{dX}{dt} = -k_2 h_0 XY + h_3 h_0 AY - 2k_4 X^2 - h_5 h_0 AX + k_{-5} U_{ss}^2 + k_6 U_{ss}(C - Z) - k_{-6} XZ, \quad 5.16$$

$$\frac{dY}{dt} = -k_2 h_0 XY - k_3 h_0 AX + q\left(\frac{k_7 k_8 BZ}{k_{-7} h_0 (C-Z) + k_8}\right) + k_9 B, \quad 5.17$$

$$\frac{dZ}{dt} = k_6 U_{ss}(C - Z) - k_{-6} XZ - \left(\frac{k_7 k_8 BZ}{k_{-7} h_0 (C-Z) + k_8}\right), \quad 5.18$$

where $X = [\mathrm{HBrO_2}]$, $Y = [\mathrm{Br^-}]$, $U = [\mathrm{HBrO_2^+}]$, $Z = [\mathrm{ox}]$, $A = [\mathrm{HBrO_3}] = h_0[\mathrm{NaBrO_3}]_0/(0.2 + h_0)$, $B = [\mathrm{CHBr(COOH)_2}]$, $\mathrm{B}^{\bullet} = [\mathrm{CBr(COOH)_2^{\bullet}}]$, $C = Z + [\mathrm{red}]$, P = [HOBr], and h_0 = Hammett acidity function. The parameter q is a stoichiometric factor and Q is organic products.

U_{ss}, the steady state concentration of $\mathrm{HBrO_2^+}$ is defined as:

$$U_{ss} = \frac{1}{4k_{-5}}\left(k_6(C - Z) + [k_6^2(C - Z)^2 + 16k_5 k_{-5} h_0 AX + 8k_{-5} k_{-6} h_0 XZ]^{1/2}\right). \quad 5.19$$

By introducing the scaling parameters listed in Table 3, we reach the following dimensionless equations:

$$\frac{dx}{dt} = \frac{1}{\varepsilon_1}(-x^2 - x + \varepsilon_2 \gamma u_{ss}^2 + u_{ss}(1 - z) + \delta xz + \mu y - xy), \quad 5.20$$

$$\frac{dy}{dt} = \frac{1}{\varepsilon_4}\left(-xy - \mu y + q\left[\frac{\alpha z}{\varepsilon_3+1-z} + \beta\right]\right), \quad 5.21$$

$$\frac{dz}{dt} = u_{ss}(1-z) - \delta xz - \frac{\alpha z}{\varepsilon_3+1-z}, \quad 5.22$$

where *x, y* and *z* are concentrations of $HBrO_2$, Br^-, and the oxidized form of the catalyst, respectively.

Table 5.3. Scaling parameters for the dimensionless ZBKE reaction model.

variable	scaling	parameter	scaling	parameter	scaling
X	$\frac{k_5 h_0 A}{2k_4} x$	ε_1	$\frac{k_5 h_0 A}{2k_4 C}$	β	$\frac{2k_4 k_9 B}{(k_5 h_0 A)^2}$
Y	$\frac{k_5 A}{k_2} y$	ε_2	$\frac{(k_5 h_0 A)^2}{2k_4 k_6 C}$	μ	$\frac{2k_3 k_4}{k_2 k_5 h_0}$
Z	Cz	ε_3	$\frac{k_8}{k_{-7} h_0 C}$	Υ	$\frac{k_{-5}}{k_6}$
U	$\frac{(k_5 h_0 A)^2}{2k_4 k_6 C} u$	ε_4	$\frac{k_5 A}{k_2 C}$	δ	$\frac{k_{-6} C}{k_5 h_0 A}$
t	$\frac{2k_4 C}{(k_5 h_0 A)^2} z$	α	$\frac{2k_4 k_7 k_8 B}{k_5^2 k_{-7} h_0^3 A^2}$		

If we treat [*Y*] with a steady state approximation, the model can be further reduced to a two-variable model[10]

$$\frac{dX}{dt} = \left(\frac{-k_2 X + k_3 A}{k_2 X + k_3 A}\right)\left(q\left[\frac{k_7 k_8 BZ}{k_8 + k_{-7} h_0 (C-Z)}\right] + k_9 B\right) - 2k_4 X^2 - k_5 h_0 AX + k_{-5} U_{ss}^2 + k_6 U_{ss}(C - Z) - k_{-6} XZ, \quad 5.23$$

$$\frac{dZ}{dt} = k_6 U_{ss}(C - Z) - k_{-6} XZ - \left(\frac{k_7 k_8 BZ}{k_{-7} h_0 (C-Z) + k_8}\right). \quad 5.24$$

5.2.2 Photosensitive BZ Reaction Model

If a photosensitive complex of ruthenium, such as tris(2,2'-bipyridine)ruthenium(II) ($Ru(bpy)_3^{2+}$) is used as the metal catalyst, the concentration of the species in the BZ reaction can be influenced by visible light[12]. If the concentration of bromate and acid are higher than malonic acid and bromide, the illumination of the catalyst promotes the oxidation of $Ru(bpy)_3^{2+}$, and consequently the autocatalytic reaction[13]. In the ZBKE model, the excited metal catalyst, *E*, reacts with *A* ($HBrO_3$) to generate *X*,

$$E + A \rightarrow X + 2Z, \quad 5.25$$

or alternatively it reacts with *B* ($CHBr(COOH)_2$) to generate *Y*,

$$E + B \rightarrow Y + Z. \qquad 5.26$$

Vanag and co-workers introduced a term, φ, to account for the production of Y,[14] which was further modified to account for the production of X as well. Taking into account φ, the final version of the ZBKE model used in our work is as follows[15]

$$\frac{dx}{dt} = \frac{1}{\varepsilon_1}\left(\varphi - x^2 - x + \varepsilon_2 \gamma u_{ss}^2 + u_{ss}(1-z) + \left[\frac{(\mu - x)}{(\mu + x)}\right]\left[\frac{q\alpha z}{(\varepsilon_3 + 1 - z)} + \beta\right]\right), \qquad 5.27$$

$$\frac{dz}{dt} = \varphi + u_{ss}(1-z) - \left(\frac{\alpha z}{(\varepsilon_3 + 1 - z)}\right). \qquad 5.28$$

Fig. 5.2 shows the time evolution of the concentration of x and z, obtained from the simulation of Eq. 5.27 and 5.28 with the parameters in Table 4.

Table 5.4. Parameters used in photosensitive ZBKE mode represented by Eq. 5.27 and 5.28.

parameter	value	parameter	value
ε_1	0.11	β	1.7×10^{-5}
ε_2	1.7×10^{-5}	γ	1.2
ε_3	0.016	μ	2.4×10^{-4}
α	0.1	q	0.8

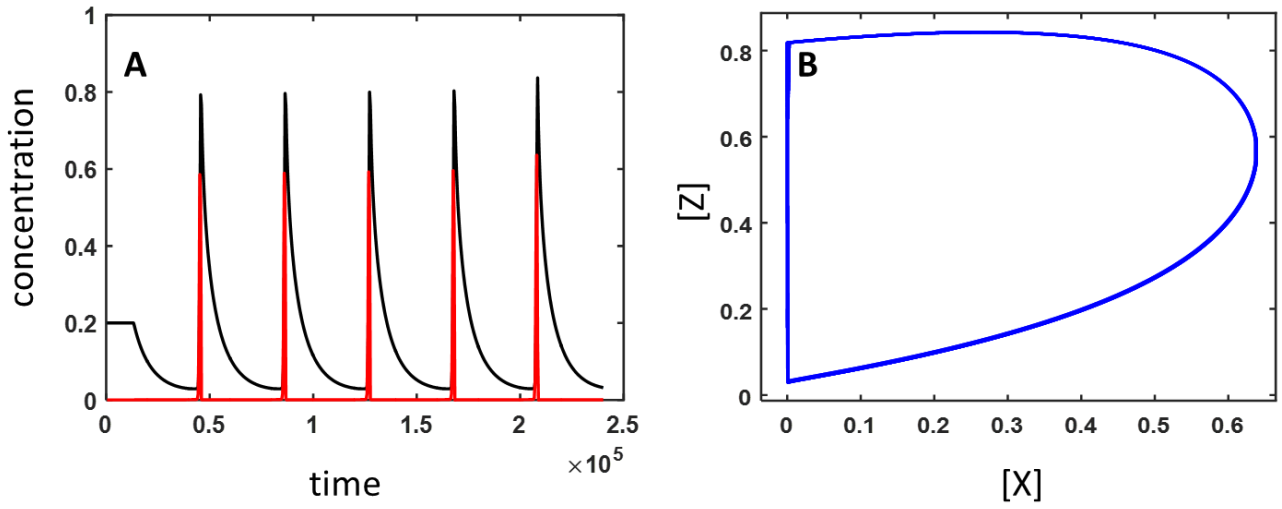


Figure 5.2. (A) Time evolution of the concentration of X (*red*) and Z (*black*) species in ZBKE model. (B) Periodic change in the concentration of Z as a function of the concentration of X.

5.2.3 Experimental Setup

Preparation of Catalyst-Loaded Beads. In the experiments, we used the tris(2,2'-bipyridyl)ruthenium(II)chloride, $[Ru(bpy)_3]Cl_2.H_2O$, catalyst (from Acros Organics, NJ, US) loaded on cation exchange particles with an average diameter of 270 microns[15,16] (DOWEX 50WX2-100, Sigma-Aldrich, Carlsbad, CA, USA). 1.0 ml of a 25.0 mM solution of $Ru(bipy)_3^{2+}$ was mixed with 3.0 g of the ion exchange particles and 7.0 ml of deionized water. The mixture was stirred for 24 hours to allow for homogeneous absorption of the catalyst on the particles. The mixture was then decanted and kept in dark for at least a weak until it was completely dried. Loading the catalyst on the beads allows for having discrete light sensitive micro-oscillators that are easy to monitor with a camera and perturbed by projecting light on them.

Preparation of Catalyst-Free BZ Solution. For each experiment, stock solutions of 0.8 M malonic acid, 1.0 M sodium bromide (NaBr), 3.0 M sulfuric acid (H_2SO_4) and 2.0 M sodium bromate ($NaBrO_3$) were used to prepare a 25.0 ml BZ solution in a volumetric flask. All chemicals were obtained from Fisher Scientific (Fair Lawn, NJ, US). Deionized water was used to prepare and dilute all solutions.

Computer-Aided Experiments. The catalyst changes between oxidized and reduced states with green and orange colors, respectively. The periodic color change of the oscillator allows measuring the period, and consequently the phase of an oscillator at any time. For each experiment, 30-40 micro-oscillators were placed in a petri dish containing 2.5 ml of BZ solution, with a minimum spacing between them of 1 mm. Since $Ru(bipy)_3^{2+}$ is a photosensitive catalyst, the period and phase of each oscillator is altered by illuminating it with 440 nm light. The oscillators are coupled by monitoring the gray level of each oscillator with a CCD camera and applying the appropriate feedback to them with a spatial light modulator (SLM). Fig. 5.3 shows a schematic representation of the experimental setup.

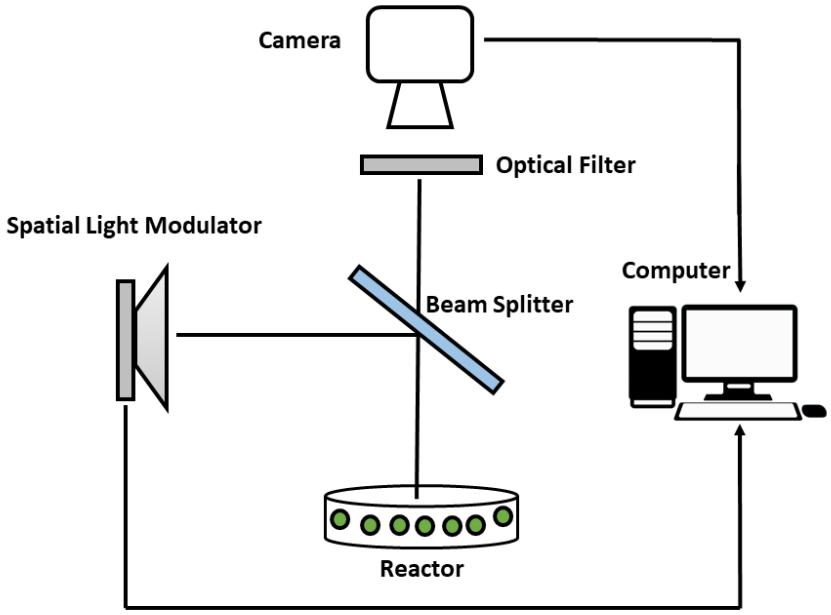


Figure 5.3. Experimental setup of BZ reaction. The gray level of the photosensitive micro-oscillators is monitored with a CCD camera and the appropriate feedback is applied with a SLM.

The light intensity projected on oscillator *j* is determined by the following equation:

$$\phi_j = \sum_{i=1}^{n} k(I_i\ (t-\tau) - I_j(t)), \qquad 5.\text{-}29$$

where *n* is the total number of oscillators, *k* is the coupling constant, *I* is the gray level of each oscillator monitored by the CCD camera, *t* is time and *τ* is the time delay of the coupling. The minimum and maximum light intensities are 0 and 250 (2.31 mWcm^{-2}) on the gray scale, respectively. The background light projected onto all oscillators was 140 or 180 in gray scale, depending on the experiment. The period of an oscillator, *T*, was measured as the time between two successive peaks in light intensity corresponding to the oxidized state of the metal catalyst. The phase of an oscillator at any time *t* is a value between 0 and 2π, corresponding to the start and end of a cycle, respectively, and is calculated from Eq. 5.30:

$$\theta = \frac{2\pi(t-t_{peak})}{T}, \qquad 5.30$$

where t_{peak} is the time of the previous peak.

5.2.4 Phase Response Curves

Since the illumination of the metal catalyst promotes its oxidation, it affects the entire oscillatory reaction. The response of an oscillator to any external perturbation, including the illumination, is measured with the Phase Response Curve (PRC), which is a plot of the change in the phase due to the perturbation with light, $\Delta\theta$, versus the phase at which the perturbation occurs, θ.[4] The PRC allows for understanding the behavior and predicting the dynamics of a set of oscillators that are influenced by each other, due to coupling or by an external perturbation. To generate the PRC of our BZ oscillators, they are perturbed for 2 s by a beam of light with the maximum intensity of 250 gray scale, with intervals between perturbations of at least two periods. The periods of the oscillators are calculated from Eq. 5.30, and are compared to the previous value of the period. Fig. 5.4 shows a representative PRC of a set of 40 oscillators.

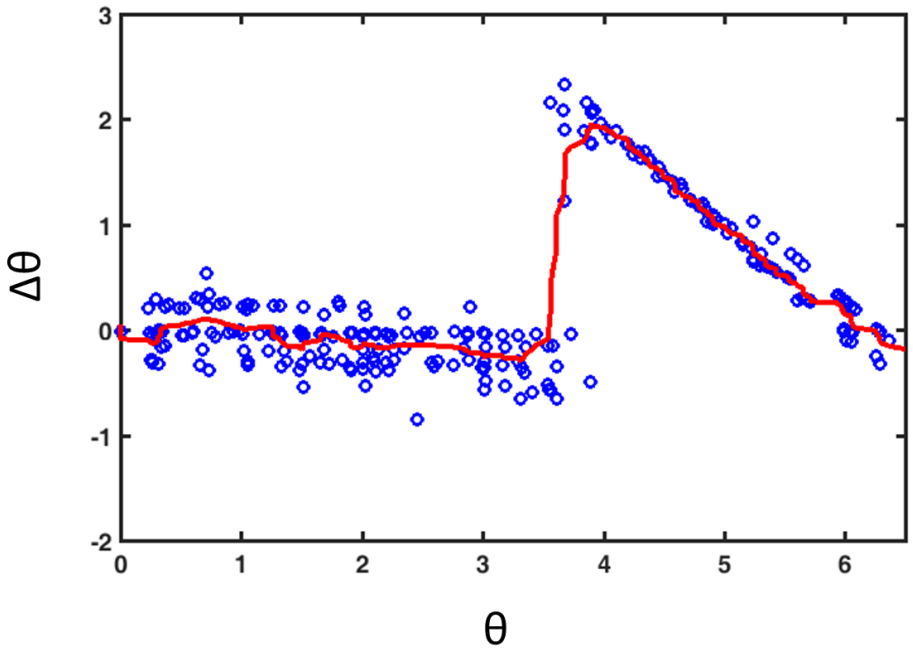


Figure 5.4. A representative PRC experiment with a positive slope. The red curve is the smooth response fit of the experimental data (blue circles).

In the example shown above, perturbations occurring at phases greater than 3 result in a phase advance, i.e., a decrease in the period. In the following chapter we will discuss the other possible types of PRCs.

References

1. Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: A Universal Concept in Nonlinear Sciences*. (Cambridge University Press, 2003).
2. Michaels, D. C., Matyas, E. P. & Jalife, J. Dynamic interactions and mutual synchronization of sinoatrial node pacemaker cells. A mathematical model. *Circ. Res.* **58,** 706–720 (1986).
3. Mirollo, R. & Strogatz, S. Synchronization of Pulse-Coupled Biological Oscillators. *SIAM J. Appl. Math.* **50,** 1645–1662 (1990).
4. Gonze, D., Bernard, S., Waltermann, C., Kramer, A. & Herzel, H. Spontaneous Synchronization of Coupled Circadian Oscillators. *Biophys. J.* **89,** 120–129 (2005).
5. Huygens, C. *Oeuvres complètes de Christiaan Huygens. Publiées par la Société hollandaise des sciences.* (M. Nijhoff, 1888).
6. Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence (Springer-Verlag, New York, 1984).*
7. Dörfler, F. & Bullo, F. Synchronization in complex networks of phase oscillators: A survey. *Automatica* **50,** 1539–1564 (2014).
8. Zhabotinsky, A. M. A history of chemical oscillations and waves. *Chaos, Interdiscip. J. Nonlinear Sci.* **1,** 379–386 (1991).
9. Field, R. J., Körös, E. & Noyes, R. M. Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in the bromate-cerium-malonic acid system. *J. Am. Chem. Soc.* **94,** 8649–8664 (1972).
10. Zhabotinsky, A. M., Buchholtz, F., Kiyatkin, A. B. & Epstein, I. R. Oscillations and waves in metal-ion-catalyzed bromate oscillating reactions in highly oxidized states. *J. Phys. Chem.* **97,** 7578–7584 (1993).
11. Toth, R., Taylor, A. F. & Tinsley, M. R. Collective Behavior of a Population of Chemically Coupled Oscillators. *J. Phys. Chem. B* **110,** 10170–10176 (2006).
12. Kádár, S., Amemiya, T. & Showalter, K. Reaction Mechanism for Light Sensitivity of the $Ru(bpy)_3^{2+}$-Catalyzed Belousov−Zhabotinsky Reaction. *J. Phys. Chem. A* **101,** 8200–8206 (1997).

13. Toth, R. & Taylor, A. F. The tris(2,2′-bipyridyl)ruthenium-catalysed Belousov–Zhabotinsky reaction. *Prog. React. Kinet. Mech.* **31,** 59–115 (2006).
14. Vanag, V. K., Yang, L., Dolnik, M., Zhabotinsky, A. M. & Epstein, I. R. Oscillatory cluster patterns in a homogeneous chemical system with global feedback. *Nature* **406,** 389–391 (2000).
15. Taylor, A. F. *et al.* Phase clustering in globally coupled photochemical oscillators. *Eur. Phys. J. Spec. Top.* **165,** 137–149 (2008).
16. Snari, R. *et al.* Desynchronization of stochastically synchronized chemical oscillators. *Chaos Interdiscip. J. Nonlinear Sci.* **25,** 123116 (2015).

Chapter 6[1]

6. Synchronization in the BZ Oscillators with External Perturbations and the Non-local Coupling

6.1 Synchronization in Neurons

As discussed in Chapter 5, synchronization is a ubiquitous phenomenon in nature. In this chapter, we discuss two ways of synchronizing phase oscillators, by applying a periodic entertainment signal and a noise entertainment signal. The synchronization also occurs via the global coupling, as discussed in Chapter 7. We also investigate the coexistent of the synchronized and unsynchronized subsets of oscillators, known as chimera states. Chimera states arise from a special pattern of the coupling with a time delay in the coupling.

Studying the different ways by which unsynchronized chemical oscillators transform to synchronized sets or subsets allows for better understanding of the more complex oscillatory systems in nature. For example, the abnormal synchronization of neurons in basal ganglia is proposed to be a reason for impaired movements in Parkinson's disease[1]. High frequency deep-brain stimulation (DBS) is suggested to relieve the symptoms of Parkinson's disease by desynchronizing neurons in globus pallidus[2]. DBS needs to be periodic to be effective[2,3], and it is proposed to interfere with synchronized sets of neurons either by entraining them or disruption of their synchrony[2]. In this chapter, we will discuss the desynchronization with a periodic external signal that interrupts the synchronized oscillators. First, we introduce different ways to make a set of BZ oscillators synchronized.

6.2 Synchronization with the Periodic Entrainment

In these experiments, 30 BZ micro-oscillators were used in an experimental setup explained in Section 5.2.3. The oscillators were allowed to oscillate without any external perturbation for at least 126 s. The background light intensity was 180 in gray scale (1.67 $mWcm^{-2}$). Then we applied periodic positive pulses with the light intensity of 240 in gray scale (2.22 $mWcm^{-2}$) and a duration of 9.0 s. To monitor the average natural frequency of the oscillators, we let 5 oscillators remain unperturbed for the entire experiment. The period of these "spectator" oscillators was monitored as the experiment was running, and

[1] Sections 6.1 to 6.3 are published in Chaos, an Interdisciplinary Journal of Nonlinear Science. 25, 123116 (2015), DOI: 10.1063/1.4937724

the period of the entertaining pulse was adjusted accordingly; the frequency of the applied signal was maintained close to the natural frequency of the oscillators for the entire experiment. The oscillators quickly respond to the signal and begin to synchronize with the same frequency as of the signal as seen in Fig. 6.1. After 1200 s, a negative periodic signal was applied while the positive pulse was still being applied, as seen in Fig. 6.1 A and D. The only difference between Fig. 6.1 A and D is the timing of the negative signal. In Fig. 6.1 A the negative signal was applied at the first half of the positive signal, i.e., the first half of the period of the synchronized oscillators, and as seen from Fig. 6.1 B and C, this leads to desynchronization of the oscillators. On the other hand, when the negative pulses occur at the second half of the positive signal, they do not affect the synchrony of the oscillators, as shown in Fig. 6.1 E and F.

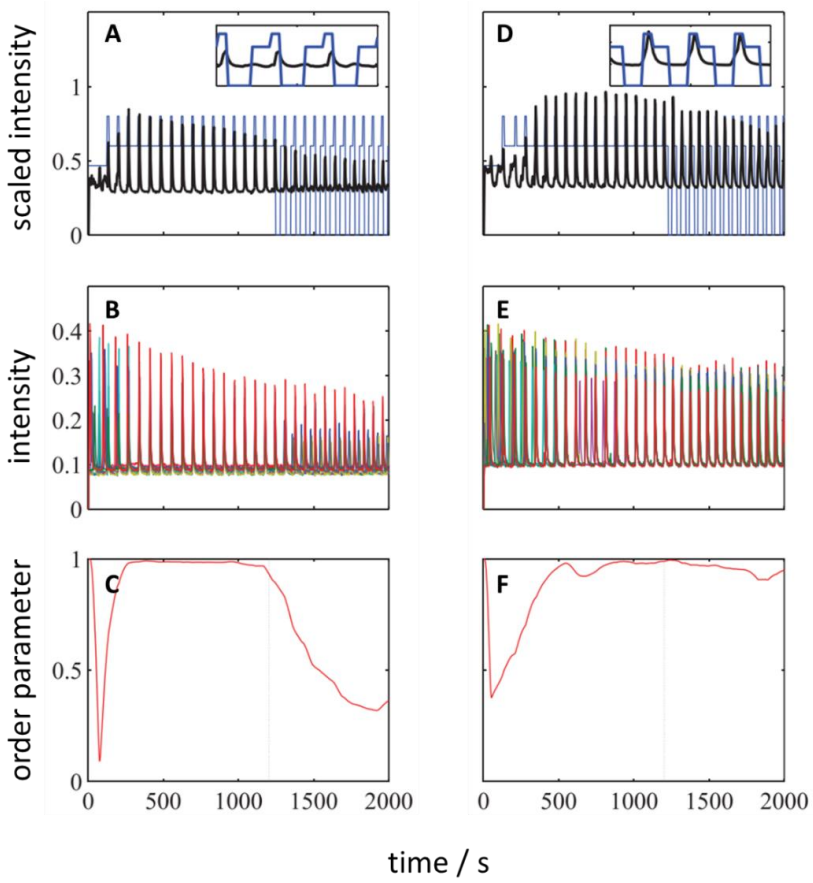


Figure 6.1. The photosensitive oscillators quickly start to synchronize with a periodic signal. (A) A periodic positive pulse (blue) with the duration of 9.0 s is applied to 30 BZ oscillators after they are left unperturbed for 120 s. After 1200 s, a negative periodic signal with a duration of the half of the positive signal was applied. The mean signal of the oscillators is shown in black. The negative signal is adjusted to occur in the first half of the period of the positive signal. (B) The time evolution of 10 representative oscillators. (C) The time evolution of the Kuramoto order parameter. The negative signal has a desynchronizing effect on the oscillators. D-F are the same as A-C, repectively, except that the negative signal is adjusted to occure in the scond half of the positive signal. In this case, the nagative signal does not affect the synchrony of the oscillators. For all experiments, the background light intensity was 180 in gray scale (1.67 $mWcm^{-2}$). The negative and positive signals have intensities of 0 and 240 in gray scale (0.0 and 2.22 $mWcm^{-2}$), respectively. The figure is from Ref. [4].

The two different responses of the oscillators to the same negative signal with different timings is explained with the experimental PRCs of the positive and negative perturbations, shown in Fig. 6.2. In PRC experiments, the perturbation signals were applied for 2 s with intervals of at least 1 period. Each data point in Fig. 6.2 is obtained from $Z(\theta) = \frac{\Delta\theta}{u\Delta t}$, with u being the absolute value of the signal amplitude and Δt being the duration of the perturbation. Since the slope of the PRC for the negative perturbation (Fig. 6.2 A) is positive in the range of θ = [0, 4.4], any perturbation in this region will lead to desynchronization of the oscillators. The slope is negative for θ = [4.4, 6.28], meaning that the perturbations applied in that region will not affect the synchrony of the oscillators. There is no region with positive slope in the PRC with the positive perturbations in Fig. 6.2 B. Therefore, the positive perturbations will either have no effect (if applied in the θ range of [0, 2.7]) or will have a synchronization effect on the oscillators (when applied between θ = 2.7 and the end of the period).

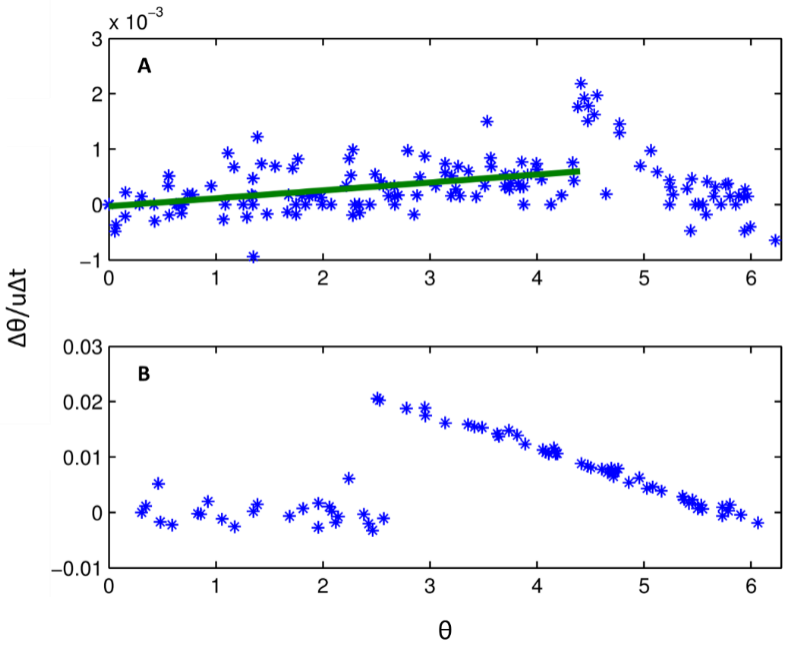


Figure 6.2. The positive and negative perturbations have two different PRCs. (A) The experimental PRC associated with the negative perturbations. The green line is fitted to the data in the range θ = [0, 4.4] with the equation $Z\times10^4 = 1.50\theta - 0.04$. (B) The PRC with positive perturations. The positive and negative perturbations are caused by light intensitties of 240 and 70 in gray scale, (2.22 and 0.64 mWcm^{-2}), respectively. Every pulse was applied for 2 s. The background light was 180 in gray scale (1.67 mWcm^{-2}). The figure is adopted from Ref. [4].

6.3 Synchronization with a Noise Entrainment Signal.

In a different approach, we applied a noise signal to mediate the synchronization in the BZ oscillators. The phenomenon of noise-induced synchronization has been observed in neurons[5,6]. We generated a noise signal, *X(t)*, using Ornstein-Ohlenbeck equation[7,8]

$$\dot{X}(t) = -\phi(X(t) - X_0) + \sqrt{2\phi\sigma^2}\eta(t). \quad 6.1$$

The magnitude of ϕ determines the rate of the mean reversion. The parameter X_0 is the mean value, σ^2 is the variance, and η is Gaussian white noise. The blue curve in Fig. 6.3 D is an example of the noise signal generated with this method. As seen from Fig. 6.3 D-F, the noise signal leads to an intermittent synchronization in the oscillators. Fig. 6.3 A-C is an unperturbed system as the control experiment.

To further verify the results of Fig. 6.3, we carried out 10 experiments in the absence of the external perturbation (set I) and 10 experiments with applying the noise signal (set II). The average order parameters for sets I and II were 0.16 ± 0.03 and 0.53 ± 0.08, respectively, indicating a clear distinction between the two systems.

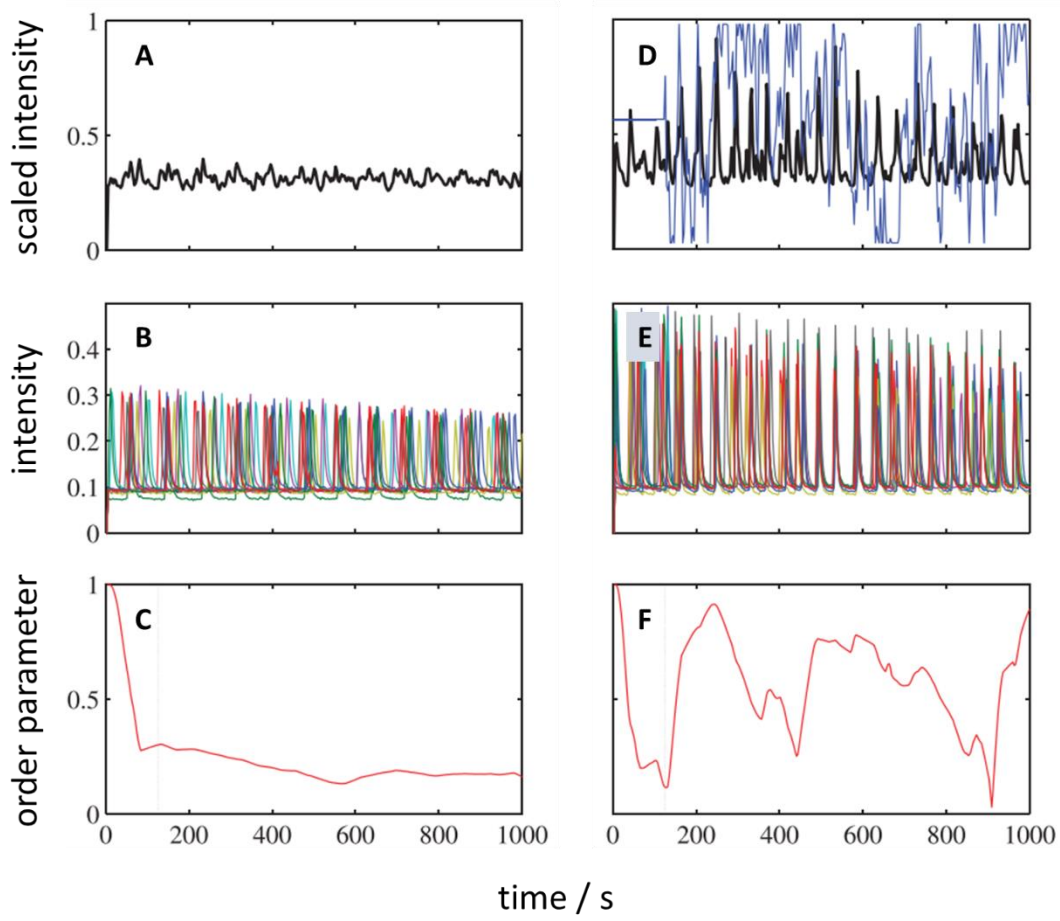


Figure 6.3. A noise signal causes an intermittent synchrony in a set of uncoupled BZ oscillators. (A) A representative experiment with no external perturbation. The balck curve is the mean signal of 30 oscillators. (B) The time evolution of 10 BZ oscillators. (C) The time evolution of the order parameter for 30 oscillators. (D) The experiment with the same conditions as part A, except for applying a noise signal (*blue*). (E) The time series of 10 representative oscillators show an intermittent synhcronization, that is also seen from the plot of the order parameter as a function of time (F). All experiments were run for 2000 s. Only 1000 s of the experiments are shown for clarity. The noise parameters are ϕ = 0.08 and σ = 50. The background light intensity was 140 in gray scale (1.29 mWcm^{-2}). The figures are adopted from Ref. [4].

Similar to Section 6.1, we applied the negative pulses to disrupt the noise-induced synchronization in the BZ oscillators, as shown in Fig. 6.4 A-C. Again, the frequency of the pulses was adjusted according to the frequency of the spectator oscillators. Since the synchronized states are now intermittent, the negative pulses were not timed to occur at any particular time with respect to the phase of the oscillators. Still, the negative pulses have a net desynchronizing effect on the oscillators. The average order parameter for 10 experiments carried out under this condition is 0.27 ± 0.05, compared to 0.53 ± 0.08 for 10 systems perturbed with noise signal, indicating a desynchronizing effect of the negative pulses, although the order parameter is not as low as 0.16 ± 0.03 that was observed for the unperturbed system. We also applied a

positive pulse signal on top of the noise signal as shown in Fig. 6.4 D-F. As expected from the PRC of the positive perturbations (Fig. 6.2 B), the oscillators eventually synchronize, as seen from the plot of the order parameter as a function of time in Fig. 6.4 F.

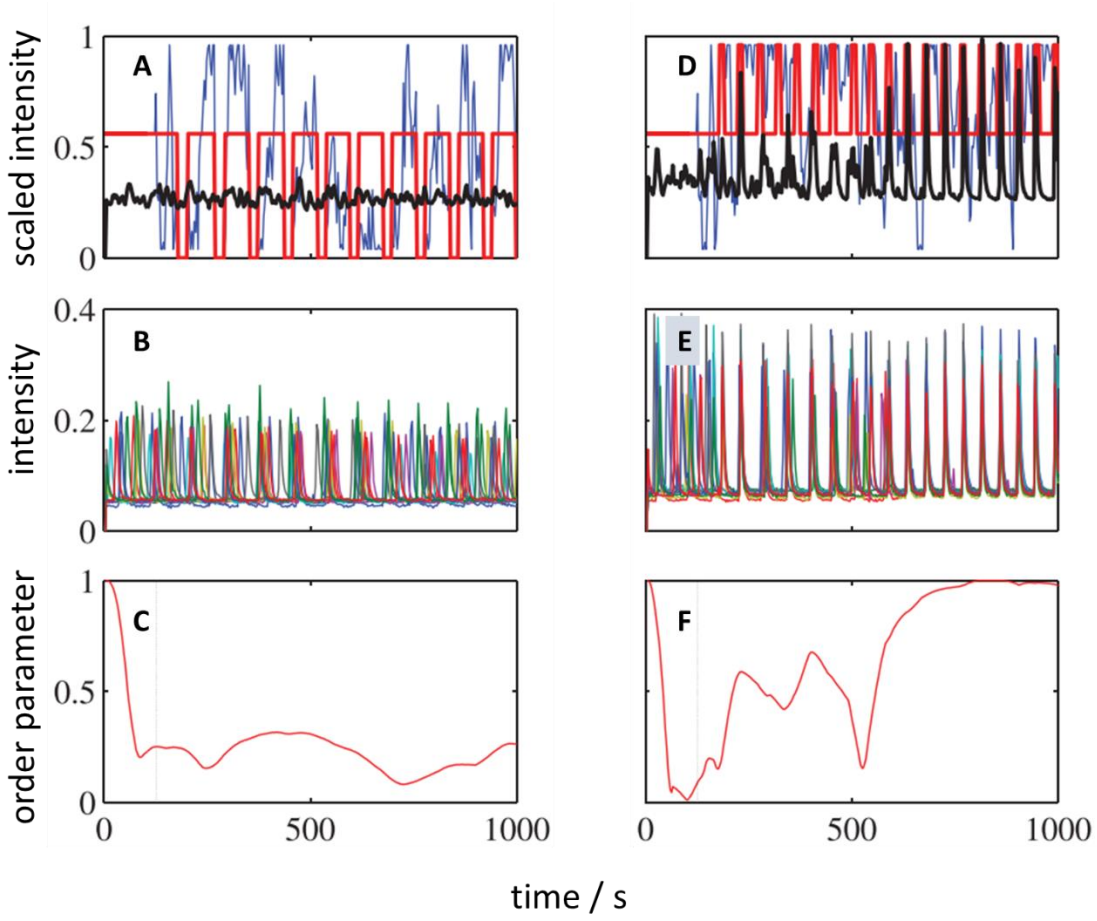


Figure 6.4. Negative periodic pulses abate the noise-induced synchrony in a set of the BZ oscillators. (A) The noise signal (*blue*) and the negative pulses (*red*) are simultaneously applied to a set of uncoupled oscillators. The period of the negative pulses is 9 s shorter than the average period of the spectator oscillators. (B) The time evolution of 10 BZ oscillators. (C) The time evolution of the order parameter of the oscillators indicates a low level of synchrony in the system. (D-F) The same as A-C, respectively, except that the negative pulses are replaced with positive pulses. As expected, the positive signal results in a synchronized state (E and F). The duration of both positive and negative pulses are ¼ of their period. The background light intensity is 140 in gray scale (1.29 $mWcm^{-2}$). The inensity of the negative and positive pulses are 0 and 240 in gray scale (0 and 2.22 $mWcm^{-2}$), respectively. The figure is adopted from Ref. [4].

6.4 Controlled Chimera States

6.4.1 Introduction

In a set of phase oscillators, if each oscillator is weekly non-locally coupled to its neighbors, the system may reach a state in which two subsets of synchronized and unsynchronized oscillators coexist. The state is known as a chimera and was first explained by Kuramoto and co-workers[9]. A schematic representation of the coupling pattern, known as nonlocal coupling[10,11] in a ring configuration is shown in Fig. 6.5 A. In the photosensitive BZ model (Eq. 5.27 and 5.28), the light intensity projected on the oscillator j at each time is defined with Eq. 6.2:

$$\phi_j = \phi_0 + \sum_{\rho=j-n}^{j+n} k(I_\rho\ (t-\tau) - I_j(t)), \qquad 6.2$$

where ϕ_0 is the background light intensity, τ is the delay in the coupling, ρ is the index of the neighboring oscillators, I is the light intensity of each oscillator in gray scale, and n is the coupling radius, i.e., the number of neighboring oscillators coupled to the oscillator j on each side. The coupling constant here is defined as $k = k'\exp(-\kappa|\rho - j|)$, where k' and κ are constants. The latter determines the decay of the effective coupling going from j to the farthest neighbor coupled to it. The time evolution of the local order parameter for a set of 40 oscillators is shown in Fig. 6.5 B. For the oscillator j, the local order parameter is calculated for a subset of 5 oscillators including j and two nearest neighbors on each side, i.e., j-2 to j+2.

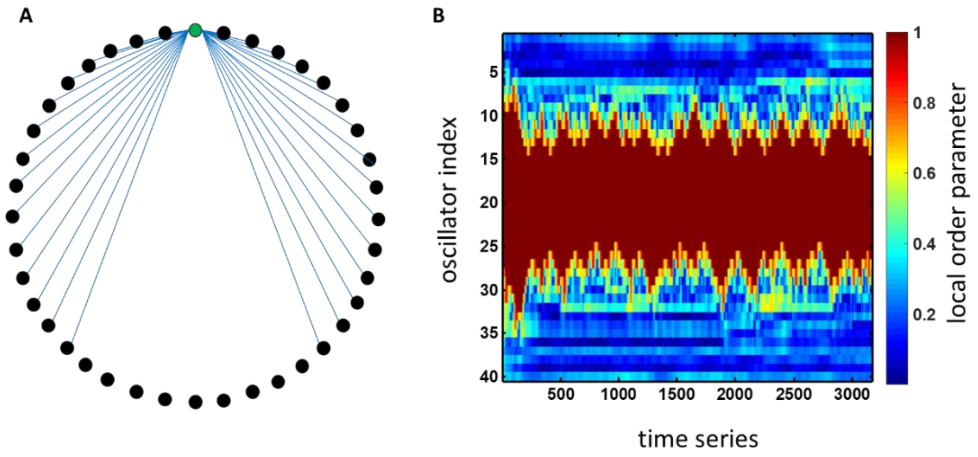


Figure 6.5. A simulation of a chimera state in a set of 40 BZ oscillators. (A) Oscillator j (*green*) is coupled to 15 nearest neighbor ocillators, *i* (*black*) on each side. All 40 oscillators in the ring have the same connection with their nearest neighbors as oscillator j. (B) The time evolution of the local order parameter (Eq. 5.2) for subsets of 5 oscillators (j-2 to j+2). The simulation is based on the ZBKE model, Eq. 5.27 and 5-28. The time delay of the coupling is 35×10^3 simulation timesteps. The aveagre period is $41\pm2.1 \times 10^3$ timesteps. Simulation parameters are k= 0.3. κ = 0.4, n =15.

The chimera state has been observed in several experimental systems, such as coupled-map lattices[12], mechanical oscillators[13,14], electronic oscillators[15,16] and photosensitive BZ oscillators[11,17,18]. Chimera states are intrinsically unstable[19]. Also, the boundary between the synchronized and unsynchronized regions of the chimera meanders over time[20]. Omelchenko and co-workers have shown that chimera states in Van der Pol oscillators are stabilized and feedback-controlled by applying an asymmetric coupling strength to each oscillator in a ring configuration[21]. In this work we show a method of controlling chimera states in the BZ oscillators.

6.4.2 Biased Chimera States

We have used a method to control the boundary between the synchronized and unsynchronized domains by adding a constant to the coupling feedback each oscillator receives from the other oscillators. For oscillator *j*, the intensity of projected light is similar to Eq. 6.2, with an additional term times the oscillator index:

$$\phi_j = \phi_0 + \left\{\sum_{\rho=j-n}^{j+n} k(I_\rho\ (t-\tau) - I_j(t))\right\}(-j^2 c), \qquad 6.3$$

where *c* is a constant and all variables are the same as in Eq. 6.2. As the result, the light intensity projected to oscillator *j* decreases proportional to the square of the oscillator index, moving from $j = 1$ to $j = 40$. Alternatively, one can reverse the bias towards the larger oscillator indices by using the Eq. 6.4:

$$\phi_j = \phi_0 + \left\{\sum_{\rho=j-n}^{j+n} k(I_\rho\ (t-\tau) - I_j(t))\right\}(-(40-j)^2 c), \qquad 6.4$$

Fig. 6.6 shows a comparison between an unbiased chimera state and the biased forms of Eq. 6.3 and 6.4. As expected, the synchronized domain moves towards the biased oscillator indices.

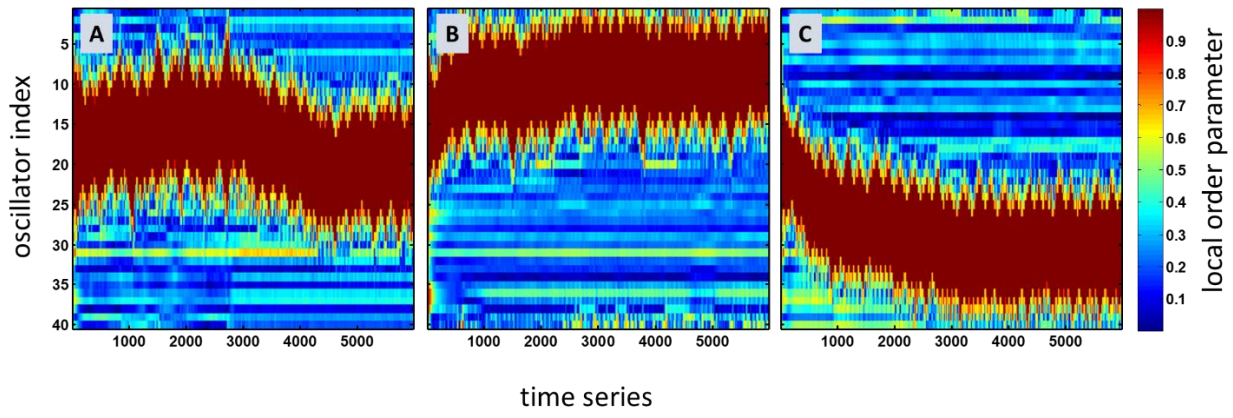


Figure 6.6. A biased feedback moves the synchronized domain towards the desired direction. (A) An unbiased chimera state formed based on the light projected on oscillator *j* in Eq. 6.2. The synchronized region meanders erratically. (B) and (C) The same system with a bised feedback based on Eq. 6.3 and 6-4, respectively. The time delay in the coupling is 35×10^3 simulation timesteps. The aveagre period is $41 \pm 2.1 \times 10^3$ timesteps. Simulation parameters are $k = 0.3$. $\kappa = 0.4$, $n = 15$.

We set to determine the repose of the system to a change in the biased feedback. In a simulation shown in Fig. 6.7, the direction of the bias is reversed from the larger oscillator indices to the smaller ones. The system quickly adjusts itself to the change in bias.

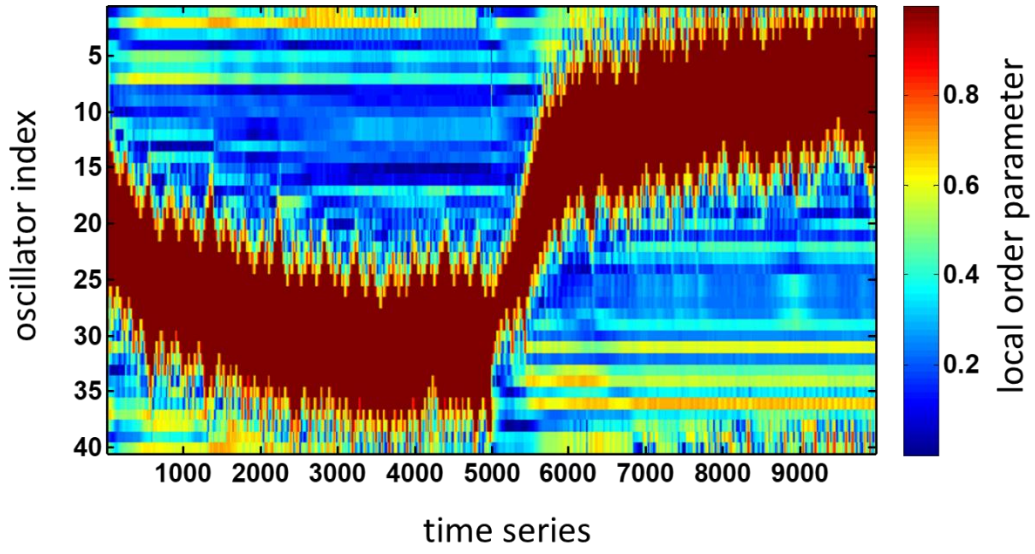


Figure 6.7. The synchronized domain adjusts intsef to a change in the biased feedback. The synchronized domain moves towards larger oscillator indices as a bised feedback is applied according to Eq. 6.3, from the beginning of the simulation to 5000 timesteps. Then, the bias is switched off and another biased feedback towards the smaller oscillator indices is applied according to Eq. 6.4. The system adjusts itself to the new conditons in ~ 1000 timesteps. All simulation parameters are the same as in Fig. 6.6.

6.4.3 The Feedback Control of Chimera States

The respose of the system to a change in the biased feedback can be used to lead the synchronized domain toward a desired position, for example, toward a target scillator. This is done by setting a threshold for the local order parameter of the target oscillator and 4 nearest neighbor oscillators. The feedback from all other oscillators is biased towards the target oscillator. The farther the synchronized domain is from the target, the stronger biase will be applied. Suppose that the oscillator *i* is the target. The distance, *d*, between the oscillator *i* and any oscillator *j* in a ring configuration is the smaller value between |*i*-*j*| or |40-*i*+*j*|, due to the periodic nature of the ring. The light projected on the oscillator *j* is defined by:

$$\phi_j = \phi_0 + \left\{\sum_{\rho=j-n}^{j+n} k(I_\rho\ (t-\tau) - I_j(t))\right\}(-(0.9-r)d^2c), \qquad 6.5$$

where *r* is the local order parameter of the target oscillator and 4 nearest neighbors and 0.9 is the threshold for the order parameter. The feedback to the oscillator *j* drops square-wise as *d* increases. This makes a steep gradient in feedback towards the target. Fig. 6.8 shows a chimera state with a target oscillator changing form oscillator 1 towards oscillator 40. As seen, the synchronized domain follows the target.

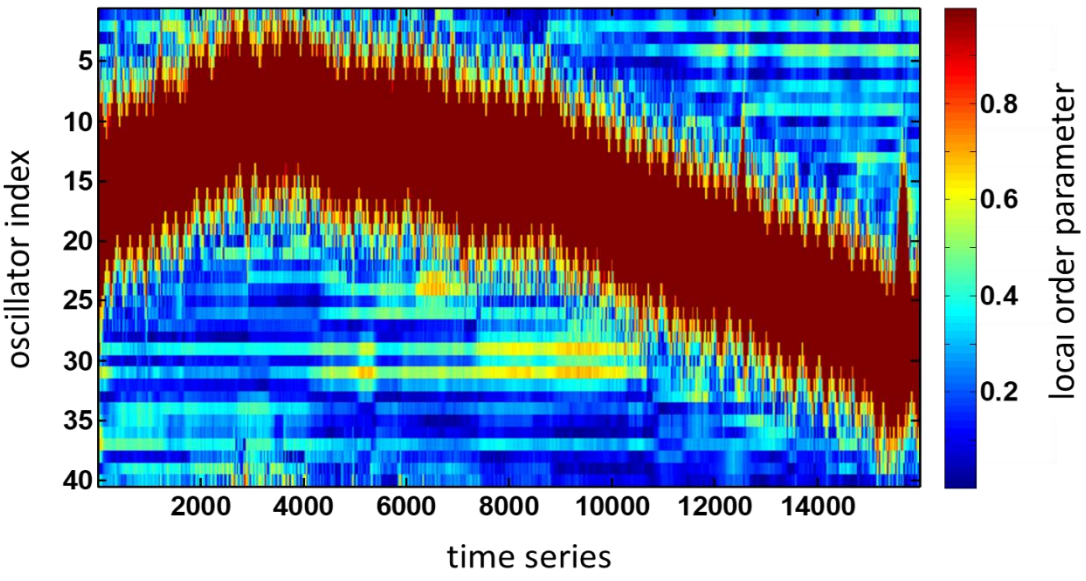


Figure 6.8. The Feedback control of the synchronized domain of a chimera state. The target oscillator changes from 1 to 40 with increments of 400 simulation timesteps. The feedback to oscillator *j* is applied according to Eq. 6.5. All other simulation parameters are the same as in Fig. 6.6.

6.5 Summary

The synchrony in a BZ system occurs by applying a common external perturbation or the internal coupling of the oscillators. The effective synchronization - as well as the desynchronization - depends on the timing of the perturbation with respect to the phase of the oscillators. The PRCs are effective tools in tuning the external signals to transfer an oscillatory system to a desired state.

The synchronized domain of a chimera state moves towards the desired direction by applying a nonlinear biased feedback on top of the usual non-local coupling. The method we developed can also be used to prevent the meandering of the synchronized domain observed in many oscillatory systems.

References

1. Brown, P. Abnormal oscillatory synchronization in the motor system leads to impaired movement. *Curr. Opin. Neurobiol.* **17,** 656–664 (2007).
2. Wilson, C. J., Beverlin, B. & Netoff, T. Chaotic desynchronization as the therapeutic mechanism of deep brain stimulation. *Front. Syst. Neurosci.* **5,** (2011).
3. Dorval, A. D., Kuncel, A. M., Birdno, M. J., Turner, D. A. & Grill, W. M. Deep brain stimulation alleviates Parkinsonian Bradykinesia by regularizing Pallidal activity. *J. Neurophysiol.* **104,** 911–921 (2010).
4. Snari, R. *et al.* Desynchronization of stochastically synchronized chemical oscillators. *Chaos, Interdiscip. J. Nonlinear Sci.* **25,** 123116 (2015).
5. Mainen, Z. F. & Sejnowski, T. J. Reliability of spike timing in neocortical neurons. *Science* **268,** 1503–1506 (1995).
6. Fellous, J.-M., Rudolph, M., Destexhe, A. & Sejnowski, T. J. Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. *Neuroscience* **122,** 811–829 (2003).
7. Goldobin, D. S., Teramae, J., Nakao, H. & Ermentrout, G. B. Dynamics of limit-cycle oscillators subject to general noise. *Phys. Rev. Lett.* **105,** 154101 (2010).
8. Yoshimura, K., Valiusaityte, I. & Davis, P. Synchronization induced by common colored noise in limit cycle and chaotic systems. *Phys. Rev. E* **75,** 026208 (2007).
9. Kuramoto, Y. & Battogtokh, D. Coexistence of coherence and incoherence in nonlocally coupled phase oscillators. *Nonlinear Phenom. in Complex Syst.* **5**, 380-385 (2002).
10. Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence (Springer-Verlag, New York, 1984).*
11. Nkomo, S., Tinsley, M. R. & Showalter, K. Chimera states in populations of nonlocally coupled chemical oscillators. *Phys. Rev. Lett.* **110,** 244102 (2013).
12. Hagerstrom, A. M. *et al.* Experimental observation of chimeras in coupled-map lattices. *Nat. Phys.* **8,** 658–661 (2012).

13. Martens, E. A., Thutupalli, S., Fourrière, A. & Hallatschek, O. Chimera states in mechanical oscillator networks. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 10563–10567 (2013).
14. Wojewoda, J., Czolczynski, K., Maistrenko, Y. & Kapitaniak, T. The smallest chimera state for coupled pendula. *Sci. Rep.* **6,** 34329 (2016).
15. Larger, L., Penkovsky, B. & Maistrenko, Y. Virtual chimera states for delayed-feedback systems. *Phys. Rev. Lett.* **111,** 054103 (2013).
16. Gambuzza, L. V. *et al.* Experimental investigation of chimera states with quiescent and synchronous domains in coupled electronic oscillators. *Phys. Rev. E* **90,** 032905 (2014).
17. Tinsley, M. R., Nkomo, S. & Showalter, K. Chimera and phase-cluster states in populations of coupled chemical oscillators. *Nat. Phys.* **8,** 662–665 (2012).
18. Nkomo, S., Tinsley, M. R. & Showalter, K. Chimera and chimera-like states in populations of nonlocally coupled homogeneous and heterogeneous chemical oscillators. *Chaos, Interdiscip. J. Nonlinear Sci.* **26,** 094826 (2016).
19. Wolfrum, M. & Omel'chenko, O. E. Chimera states are chaotic transients. *Phys. Rev. E* **84,** 015201 (2011).
20. Omel'chenko, O. E., Wolfrum, M. & Maistrenko, Y. L. Chimera states as chaotic spatiotemporal patterns. *Phys. Rev. E* **81,** 065201 (2010).
21. Omelchenko, I., Omel'chenko, O. E., Zakharova, A., Wolfrum, M. & Schoell, E. A tweezer for chimeras in small networks. *Phys. Rev. Lett.* **116,** (2016).

Chapter 7

7. Synchronization of Chemical Oscillators in Response to Strong and Weak Coupling

7.1 PRCs with Weak and Strong Perturbations.

As discussed in Section 5.2.4, the PRC for an excitatory BZ system usually has regions either with approximately zero or positive slope. The positive part of the PRC occurs at the second half of the natural period of the oscillation, as seen in Fig. 5.4, meaning the perturbations that occur in that region have the maximum influence on the oscillators. A perturbation to an oscillator can occur with an external signal as in Fig. 5.4 and Fig. 6.2, or with coupling to other oscillators. If we have a system that is fully synchronized and apply a delay in the coupling, the phase at which a perturbation due to the coupling occurs depends on the time delay. If the time delay is a small fraction of the period of the synchronized oscillators, the effect of the coupling on the phase is smaller than when it is close to the period. Delay in coupling is ubiquitous in natural systems such as neurons[1]. In this chapter we show that there are two sets of critical parameters to keep a group of BZ oscillators synchronized when they are coupled with a time delay that is a small or large fraction of their period. Since we start with a fully synchronized systems, the critical parameters are called upper bounds of synchronization.

To understand the response of a BZ system to weak and strong perturbations, we have simulated the PRC of BZ oscillators using the photosensitive model in Eq. 5.27 and 5.28. The equations are

$$\varepsilon_1 \frac{dx}{dt} = \left(\frac{\Delta\rho}{1000} - x^2 - x + \varepsilon_2 \gamma u_{ss}^2 + u_{ss}(1-z) + \left[\frac{(\mu - x)}{(\mu + x)}\right]\left[\frac{q\alpha z}{(\varepsilon_3 + 1 - z)} + \beta\right]\right), \quad 7.1$$

$$\frac{dz}{dt} = \frac{\Delta\rho}{500} + u_{ss}(1-z) - \left(\frac{\alpha z}{(\varepsilon_3 + 1 - z)}\right), \quad 7.2$$

where $\Delta\rho$ is the difference between the background and applied light intensities, where background light, ρ_0, is 0.54. Fig. 7.1 shows the PRCs with $\Delta\rho = 0.05$ and $\Delta\rho = 1.0$ lasting for 0.5 timesteps, as examples of weak and strong perturbations, respectively. The PRC is defined as $\Delta\theta/\Delta\rho\Delta t$.[2] For the small values of θ, the weak and strong perturbations result in almost identical PRCs. However, for $\theta > 4.3$, PRC changes smoothly when the perturbations are weak, compared to an almost immediate resetting of θ for the strong perturbations. Figures 7.1 C and 7.1 E show these two different regimes.

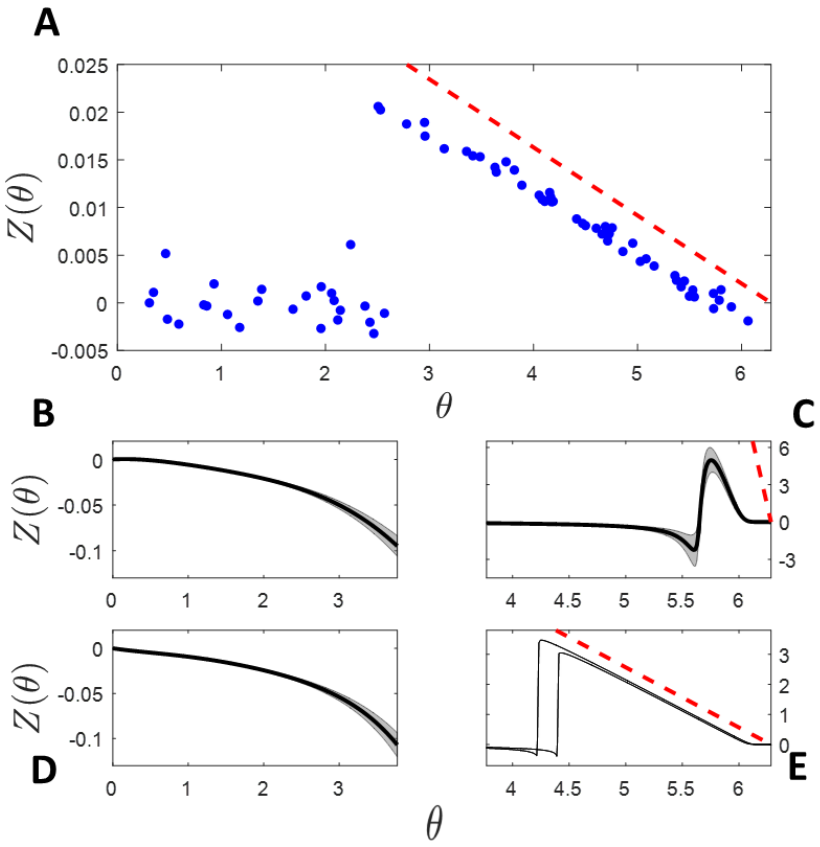


Figure 7.1. Weak and strong perturbations result in two different PRCs. (A) A representative PRC measured from experiment. The oscillators are perturbed with pulses with light intesity of 2.31 mWcm^{-2}. The background light intensity is 1.67 mWcm^{-2}. (B) and (C) are simulations with the perturbations, $\Delta\rho$, with the magnitude of 0.05. (D) and (E) are the same system with $\Delta\rho$ having a magnitude of 1.0. The gray bands show results of simulations for systems with q in the range of [0.68, 0.80]. The two black curves in (E) are for the two limits of the range. Dashed lines show the upper bounds for PRC. The figure is from Ref [2].

The PRC associated with weak perturbations is called the infinitesimal PRC, or iPRC. The PRC resulting from strong perturbations is called the spike threshold curve, STC. One would expect that the difference in the PRC in the two regimes will result in differences in synchronization in oscillators when they are weekly or strongly coupled. We set out to determine the critical parameters for maintaining a set 40 BZ oscillators synchronized under these two regimes. Similar to Eq. 5.29, the light projected on oscillator j is calculated from

$$\rho_j = \rho_0 + \frac{k}{n}\sum_{j=1}^{n}(I_j\ (t-\tau) - I_i(t)), \qquad 7.3$$

where ρ_0 is the background light intensity. All other parameters are the same as in Eq. 5.29. The background light ρ_0 is 0.54 and 1.29 mWcm^{-2} for simulations and experiments, respectively. The value of q in Eq. 7.1 is chosen from a random distribution with a mean of 0.6 and the maximum variance of 0.1. The average

frequency of the oscillators is 34. Fig. 7.2 shows the order parameter, r, as a function of the coupling strength, k, in simulations and experiments in the weak and strong coupling regimes. In experiments, a high initial k is applied to fully synchronize the system. Then desired k and τ values are applied and the average order parameter is calculated. Each data point in Fig. 7.2 B and D is for an experiment. In the weak coupling regime, a gradual shift to higher values of r is observed. In the strong coupling regime, an abrupt transition from lower to higher values of r occurs. There are also some points with relatively high values of k and low values of r. As we will see the order parameter in this regime depends on the maximum period among the oscillators. For high values of the maximum period, the systems stays weekly synchronized in spite of applying a high k.

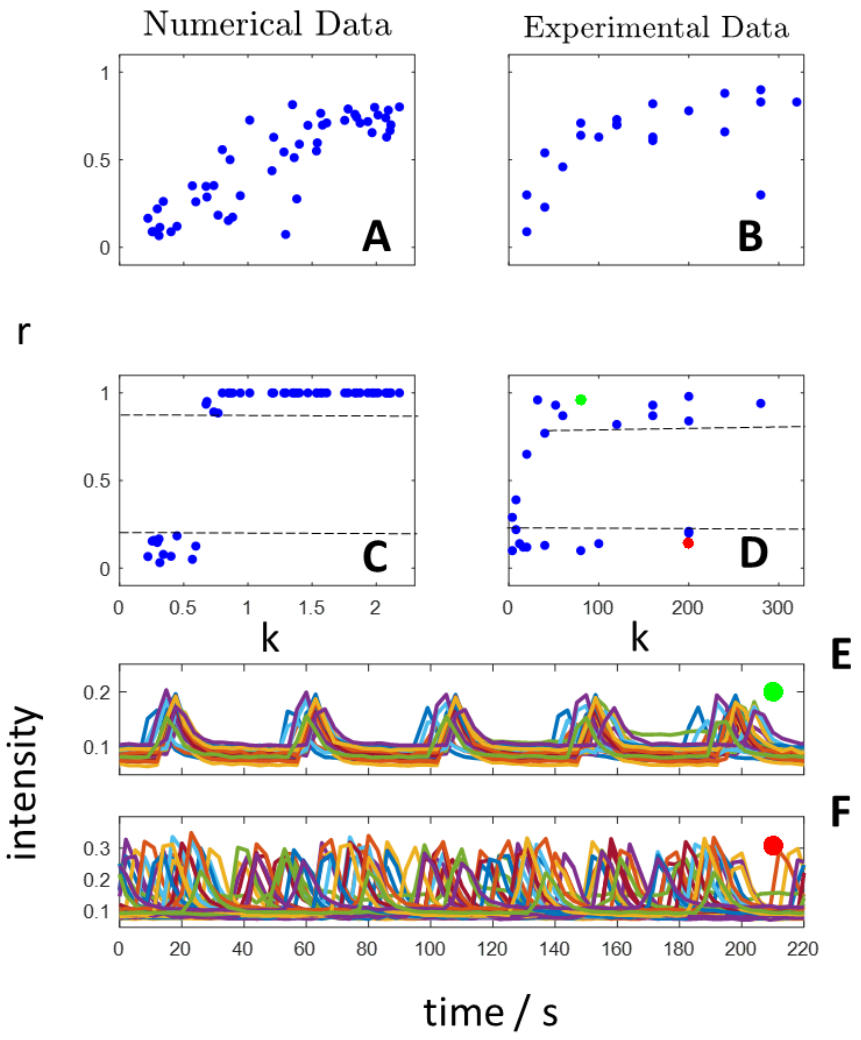


Figure 7.2. Weak and strong perturbations result in different patterns of synchronization. Simulatoin of Eq. 7.1 and 7.2 with time delays of 17 (A) and 82 (B) percent of the average period. (B) and (D) are corresponding experiments with time delays of 9 and 83 percent of the average period, respectively. (E) and (F) show the long time behavior of two representative systems shown in green and red dots in (D), respectively. The figure is from Ref [2].

7.2 Critical Parameters for Synchronization with Weak and Strong Coupling.

Simulations of Eq. 7.1 and 7.2 with time delays as a small and large fraction of the mean period show that there are two different set of critical parameters required to maintain a set of 40 oscillators synchronized. For the weak coupling, the requirements can be explained with Eq. 7.4:

$$k_c \leq \frac{\omega_{max}-\omega_{min}}{\beta}, \qquad 7.4$$

where k_c is the upper bound for the critical coupling, and ω_{max} and ω_{min} are the maximum and minimum frequency of the oscillators, respectively. The parameter β is the phase difference in the coupling, which determines the synchronizing influence from the weak coupling, and it depends on the heterogeneity of the system[2]. Since we are not able to measure the iPRC of the system experimentally, β is not involved in the results presented below. For a mathematical proof of Eq. 7.4 see Ref [3].

In the case of strong coupling, i.e., when τ is close to the average period, we do not observe a direct dependency of k_c on the distribution of natural frequencies (see Table 7.1). Since in the STC regime we observe an immediate phase resetting due to a perturbation (Fig. 7.1 E), we set to obtain *S(θ)* for θ in range of [0, 2π]. Fig. 7.3 shows the results of the simulations and experiments. Each data point in Fig. 7.3 A is a PRC experiment with a different perturbation light intensity, *S(θ)*, as a function of θ at which the maximum phase resetting occurs. For θ > 5, any small perturbation results in an immideate phase reset. When oscillators are coupled, any coupling feedback that is greater than *S(θ)* at that phase will result in an immediate phase resetting, which is expected from the STC. In this regime the contribution from the iPRC can be neglected. If the coupling feedback is less than *S(θ)*, the behavior of the oscillator is explained by the iPRC. The condition for the immediate phase resetting is expressed by Eq. 7.5:

$$\frac{k}{n}\sum_{j=1}^{n} f(\theta_i, \theta_j) > S_i(\theta_i), \qquad 7.5$$

where f is a function of θ_i and θ_j and $S_i(\theta_i)$ is the STC of oscillator *i*. The expression on the left side of Eq. 7.5 is simply the coupling feedback to oscillator *i* from the rest of the oscillators. In order to cause an immediate resetting in oscillator *i*, the feedback from the rest of oscillators has to be greater than $S_i(\theta_i)$. If we have an estimation of the term on the left side of Eq. 7.5, we can express the critical coupling, k_c, as a function of the STC of the oscillator with the largest period. The oscillator with the maximum period needs the strongest perturbation to become synchronized with the group, and once this condition is met, the rest of the oscillators receive the minimum feedback for synchronization as well. Therefore, $\mathcal{S}$ is a function of the maximum period and the time delay. Since the system is initially synchronized, i.e., the oscillators have equal phases, $f(\theta_i, \theta_j) \approx f(\theta_j)$. The maximum value of $f(\theta_i, \theta_j)$, M_j, occurs at τ timesteps after all

oscillators spike, a region in the PRC that perturbations will result in the maximum value of phase reset. Since the PRCs of the oscillators do not differ significantly, we define $\bar{M}$ as the average values of M_j. With all of these assumptions we reach a critical coupling strength in Eq. 7.6:

$$k > \frac{\max[S_i(\tau\omega_i)]}{\bar{M}}, \tag{7.6}$$

or

$$k_c \leq \frac{S_{max}(\tau\omega_{min})}{\bar{M}}, \tag{7.7}$$

where $S_{max}(\theta) = \max(S(\theta))$. As mentioned earlier, the slowest oscillator determines the k_c required to keep the oscillators synchronized. Therefore, ω_{min} appears in Eq. 7.7. The importance of the Eq. 7.7 is the fact that the properties of the group such as $S_{max}(\theta)$ and ω_{min} are sufficient for determination of the k_c, and the analytical equations for all individual oscillators do not need to be solved, which is mathematically tedious. Similar conditions exist for Eq. 7.4 for the weak coupling. Fig. 7.4 shows plots of the order parameter as a function of the critical parameters in Eq. 7.4 and 7.7. The dashed vertical lines are the critical thresholds for synchronization.

In the experiments we fit $S(\theta)$ to θ with the equation $S(\theta) = a\theta + b$, where $a = -58$ and $b = 281$. Replacing $S(\theta)$ in Eq. 7.7 gives $\frac{k\bar{M}-b}{\omega_{min}} > a\tau$. The experimental results in Fig. 7.5 B and C show a similar pattern to the simulation results in Fig. 7.4.

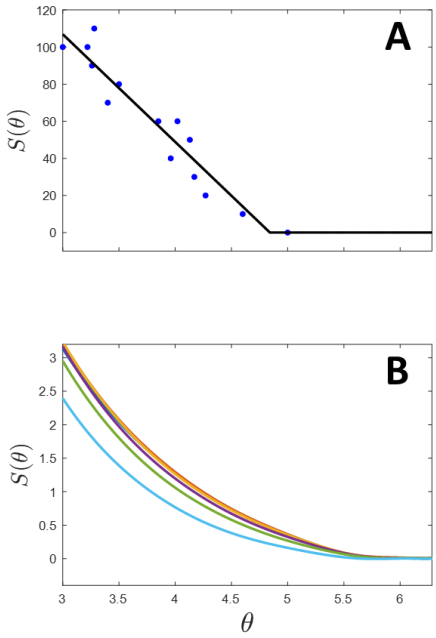


Figure 7.3. Spike threshold is reversely proportional to the phase of the oscillators. (A) Experimental measurment of *S(θ)* as a funcitonal of *θ*. The closer the peruturbations are to the end of the cycle, the weaker the light intensities are required for phase resetting. The black curve is the linear piecewise fit of the experimental data. (B) The same curves obtained from simulations with *q* in the range of [0.5, 0.87]. The figure is from Ref [2].

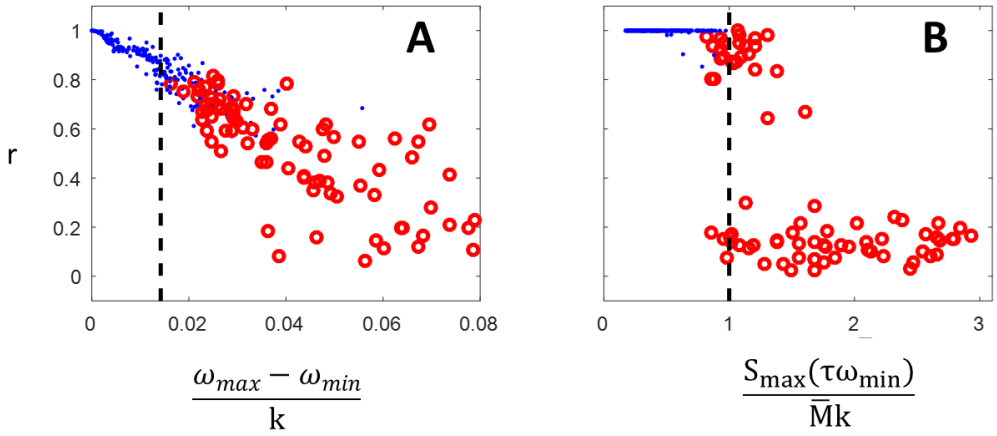


Figure 7.4. Evaluation of critical parameters for synchronization with weak and strong coupling. Numerically determined order parameter as a function of the parameters in Eq. 7.4 (A) and Eq. 7.7 (B). The time delay for (A) and (B) is 6 and 28 timesteps, respectively. The dashed lines are the critical bounds for synchronization determined from simulations. Closed and open circles show the oscillator sets that are phase coherent and phase incoherent, respectively. The figure is from Ref [2].

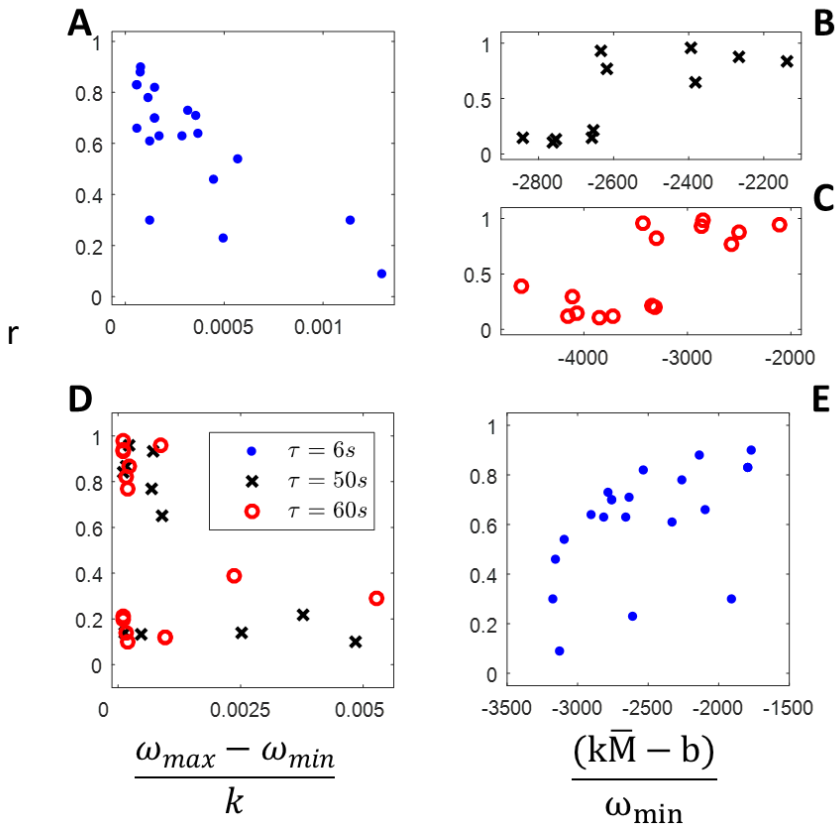


Figure 7.5. Experimentally evaluated critical parameters for synchronization with weak and strong coupling. (A) Order parameter as a function of critical parameters in Eq. 7.4 for 40 oscillators with a time delay of 6 s. (B) and (C) are systems with time delays of 50 and 60 s, respectively. The order parameter is a function of parameters in Eq. 7.7. (D) and (E) are the plots of *r* with the other set of parameters, i.e., Eq. 7.4 for the strong coupling and Eq. 7.7 for the weak cupling, respectively. No clear pattern is observed in (D) and (E). The figure is from Ref [2].

As seen from Eq. 7.4 and 7.7, there are different criteria for synchronization in weak and strong coupling conditions. For the former, the distribution of the periods determines the synchronization, and for the latter the maximum period is the most important parameter. Figures 7.4 and 7.5 are the results of simulations and experiments, respectively. The pattern seen in Fig. 7.4 A for the weak coupling is in qualitative agreement with the experimental results of Fig. 7.5 A. Also, the simulation results in Fig. 7.4 B agrees with Fig. 7.5 B and C, where time delays are on the order of the period of the synchronized oscillators. However, the plot of the order parameter as a function of the critical parameters obtained for the iPRC when the time delay is large (Figure 7.5 D), or as a function of the critical parameters obtained for the STC when the time delay is small (Fig. 7.5 D) does not show any particular trend. These results necessitate a distinction between weak and strong coupling regimes for the synchronization of BZ oscillators.

Table 7.1. Order parameter as a function of crirtical coupling parameters. For a small time delay, the order parameter depends on the distribution of periods. When the time delay is large, the maximum period determines the synchronization of the system.

time delay (s)	k	T_{min} (s)	$T_{average}$ (s)	T_{max} (s)	r
6	8	60.8	65.8	73.7	0.83
	7	56.8	63.1	66.7	0.83
	7	53.7	60.3	65.8	0.90
	7	51.1	63.0	71.0	0.30
	6	60.1	66.0	72.6	0.88
	6	61.5	65.7	71.2	0.66
	5	56.2	63.5	70.7	0.78
	4	55.7	77.2	67.5	0.61
	4	57.7	75.0	77.0	0.63
	4	57.5	65.9	73.4	0.82
	3	51.7	68.5	75.1	0.73
	3	61.5	66.8	74.4	0.70
	3	61.5	66.8	74.4	0.70
	2.5	55.0	66.5	73.4	0.63
	2	51.1	60.8	66.5	0.71
	2	54.6	65.8	73.3	0.64
	1.5	58.1	68.2	77.2	0.46
	1	58.0	65.9	73.4	0.54
	1	51.8	57.0	61.9	0.23
	0.5	55.5	64.9	72.0	0.09
	0.5	57.8	66.2	73.1	0.30
50	5	54.5	65.0	82.1	0.14
	2	48.1	53.6	56.2	0.96
	1.5	45.4	53.1	56.0	0.87
	1.3	47.2	53.2	60.8	0.93
	1	51.6	58.0	63.1	0.13
	1	47.2	52.0	59.9	0.77
	0.5	46.5	50.3	53.9	0.65
	0.4	50.0	53.3	60.3	0.11
	0.3	49.0	52.8	59.8	0.14
	0.2	49.8	54.3	59.6	0.22
	0.1	51.2	57.9	61.9	0.10
60	7	62.7	65.0	66.1	0.94
	5	66.3	72.7	78.2	0.98
	5	64.4	74.3	89.9	0.20
	5	72.4	78.7	91.8	0.21
	4	65.6	68.9	71.5	0.93
	4	48.1	59.4	64.9	0.87
	3	50.6	57.0	63.0	0.77
	3	57.6	70.3	84.7	0.82
	2.5	67.1	86.0	99.9	0.14
	2	75.6	83.0	92.5	0.10
	0.8	68.1	73.0	79.4	0.96
	0.5	70.3	79.0	94.9	0.12
	0.4	69.8	75.8	84.4	0.12
	0.2	81.6	88.6	103.8	0.39
	0.1	70.5	79.3	92.2	0.29

7.3 Summary

Different sets of critical parameters are determined for BZ oscillators coupled together with weak and strong coupling. In the two regimes, the phase equations for all oscillators do not need to be solved, since the properties of the group such as the distribution of periods and the maximum period are sufficient to determine the critical coupling strength, which significantly reduces the complexity of the equations. Our methodology is useful especially for biological systems in which the exact phase equations or PRCs are difficult to determine[4].

References

1. Izhikevich, E. M., Gally, J. A. & Edelman, G. M. Spike-timing Dynamics of Neuronal Groups. *Cereb. Cortex* **14,** 933–944 (2004).
2. Wilson, D., Faramarzi, S., Moehlis, J., Tinsley, M. & Showalter, K. Synchronization of Heterogeneous Oscillator Populations in Response to Weak and Strong Coupling. *To be Published* (2018).
3. Wilson, D. & Moehlis, J. Analytical bounds on the critical coupling strength in a population of heterogeneous biological oscillators. in *2016 American Control Conference (ACC)* 5772–5778 (2016). doi:10.1109/ACC.2016.7526574
4. Wilson, D., Holt, A. B., Netoff, T. I. & Moehlis, J. Optimal entrainment of heterogeneous noisy neurons. *Front. Neurosci.* **9,** (2015).