Crystallographic Studies on Archaeal Light-driven Proton Pumps

Siu-Kit CHAN

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Philosophy Doctor
in
Physics

Supramolecular Biophysics Laboritary (J-Lab)
Graduate School of Science
Nagoya University
February 3, 2016

Abstract of thesis entitled:
Crystallographic Studies on Archaeal Light-driven Proton Pumps
Submitted by Siu-Kit CHAN
for the degree of Philosophy Doctor
at Nagoya University in February 3, 2016

To acquire energy for metabolism, organisms have developed different strategies to obtain energy from the environment. One of the most important energy sources is sunlight. Plants and algae convert light energy to chemical energy by photosynthesis, in which carbon dioxide and water are converted to glucose and oxygen by using chlorophyll. On the other hand, archaea and some other microorganism convert light energy to a proton gradient by using retinal-containing membrane proteins which function as light-driven proton pump.

Since the discovery of the first light-driven proton pump – bacteriorhodopsin (bR) – in early 1970s, efforts have been made to understand the structures and mechanism of light-driven proton pumps for more than four decades. To date, several proton-pumping rhodopsins from various types of organisms have been crystallized and their structures have been determined. Comparison of these different structures would provide an insight into essential structural motifs that are relevant to the pumping activity. In this thesis we report the crystal structure and spectroscopic properties of cruxrhodopsin-3, which is a newly found member of archaeal light-driven proton pump; we also discuss about the spectroscopic properties of archaerhodopsin-3.

Cruxrhodopsin-3 (cR3) is a new member of archaeal light-driven proton pump that is found in *Haloarcula vallismortis*. Recently we have successfully crystallized cR3 by the membrane fusion method into crystals with space group *P321*, which diffracts X-ray to 2.1 angstrom resolution. Similar to the crystal structures of other archaeal rhodopsins (bacteriorhodopisin, archaerhodopsin-2, deltarhodopsin-3), cR3 consists of seven transmembrane alpha helices, with a beta sheet located at the BC loop region on the extracellular side. Our structural data showed that the proton-releasing group, which consists of two glutamates facing each other, is maintained by a low barrier hydrogen bond. This structural feature is commonly seen among all the crystal struc-

tures of archaeal proton-pumping rhodopsins which were crystallized by the membrane fusion method.

A detailed structural analysis shows that cR3 consists of the following specific structural properties: i) cR3 has a remarkably long DE loop which interacts with a neighboring subunit, strengthening the trimeric structure; ii) three positively charged residues at the cytoplasmic end of helix F are pointing outward, influencing the arrangement of cR3 molecules in the cell membrane; iii) the cytoplasmic part of helix E is bent in such a manner that an open pathway for proton uptake from the extracellular medium is created; iv) a short amphiphilic helix formed near the C terminal end covers the cytoplasmic surface of the protein; v) the retinal binding pocket of cR3 is more rigid than that of bR. It is suggested that the rigidity of the retinal binding pocket restricts the motional freedom of retinal, affecting formation/decay kinetics of the early reaction intermediates in the proton-pumping photocycle.

Photobleaching experiments showed that the photo-stability of cR3 decreased dramatically when the trimeric structure of cR3 was disassembled into monomers by mixing cR3-rich claret membranes with an excess amount of detergent. This result suggests that formation of the trimeric structure contributes to enhancement of the photo-stability of cR3.

A carotenoid, bacterioruberin, was observed at the crevice between subunits of cR3 trimer; It is surrounded by helices AB and DE of neighboring subunits. The similar binding of carotenoid molecules between subunits within the trimeric structure was also observed in different archaeal rhodopsins. It is possible that such binding strengthens the structural stability of the trimeric structure of microbial rhodopisn.

A homologous protein of archaerhodopsin-2, archaerhodopsin-3 (aR3), was also crystallized by the membrane fusion method. Although the size of aR3 crystals were too tiny for crystallography study at the moment when this thesis was finished, the spectroscopic properties and the absorption kinetics of aR3 were determined.

Acknowledgement

First of all I would like to thank my supervisor Prof. T. Kouyama, not only for giving me a chance to make my doctor research under his supervision, but also for his support on both my research and live in Japan in this four years.

Understanding my situation and giving me suitable advices, for this I would like to acknowledge all the staffs of Education Center for International Students (ECIS), Graduate school of Sciences and Department of Physics.

Continuous encouragement from the group members from the Supramolecular Biophysics Laboratory (J-lab.) is important for me to keep going, here I would like to express my acknowledgment to all of them: Prof. M. Midori, Mr. T. Nakanishi, Mr. S. Hayashi, Mr. K Mizuno, Mr. H. Kubo, Ms. N.F. Juber, Ms. Y. Itoh, Mr. S. Uyama, Mr. H. Kawaguchi, Mr. N. Kuga.

Kindly supports from the technical staffs contribute greatly to my understanding of the experiments, I would like to thank Mr. T. Gotoh, Ms T. Kitajima-Ihara and Ms. Y. Nakamura.

Yielding a thesis is not easy, working until mid-night or even overnight is quit common for researcher. I would like to thank my girl friend for her understanding.

On the path of becoming a doctor is very tough, I am glad to have "teammates" to share all the joyful and difficulty we had gone through. I would like to express my acknowledgment to all of them: Mr. Ray Chung, Dr. Toby Lai, Dr. Joey Leung, Dr. Lok Chan, Dr. HK Cheung, Mr. WH Chan, Mr. KK Kwan and Mr. Keith Yu.

Unlike other typical families in Hong Kong that parents always want their children to get a well paid job, my parents allow and support me to choose my own career path. I would like to express my thank to them.

Besides, I would also like to acknowledge all of my friends for the encouragements and supports.

On the other hand, I would also like to acknowledge Prof. S.K. Foong for sharing his teaching experience with me when I was working

as his TA.

Sincere thanks are due to the financial support from the MEXT scholarship.

Special thanks are addressed to my thesis committee : Prof. T. Kouyama, Prof. T. Noguchi, Prof. N.K. Sato, Prof. T. Yamato and Prof. K. Maki, not only for their advise and comments, but also for they patiently waiting for my belated final version of thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Sunlight is one of the most important energy sources for organisms on the earth. Organisms have developed different strategies to convert light energy for metabolism, orient themselves to receive more light or avoid being damaged from light (e.g. UV). For these purposes, organisms make use of light sensing materials. Retinylidene proteins are membrane proteins containing retinal (vitamin A aldehyde) as light receiver. Retinylidene proteins can be divided into two categories; namely, retinylidene proteins found in Eukaryotes, Bacteria and Archaea are structurally similar to one another and specified as microbial rhodopsin (type 1), while retinylidene proteins found in animals are specified as animal rhodopsin (type 2).

Retinylidene proteins were assumed to be observed in animals only before the discovery of bacteriorhodopsin in early 1970s. Besides bacteriorhodopsin, halorhodopsin, sensory rhodopsin-I and -II were soon discovered in *Halobacterium halobium* and their homologs were also found in other archaea. The family of archaeal rhodopsin was then further expanded to microbial rhodopsin after the discovery of proteorhodopsin in marine planktonic bacteria and xanthorhodopsin in eubacterium [2, 3]. Animal rhodopsins, on the other hand, belong to a larger family of protein - the G protein-coupled receptors (GPCRs), with a wide range of physiological functions from light vision to circadian rhythms regulation [4].

1.1 Archaeal rhodopsin

Among all microbial rhodopsins, archaeal rhodopsin is the best understood in terms of working mechanism and structural information. At least four types of archaeal rhodopsins have been identified in halophilic archaea (archaea which live in high salt concentration); physiologically they function as light-driven ion pumps for proton and chloride, and photo-sensors. Light-driven ion pumps harvest light energy to generate a proton gradient across membrane, which energizes ATP synthesis processes by ATP synthase. On the other hand, light sensors connected to their transducers transmit information of light signal to regulate the rotation of flagella, i.e. sensory rhodopsin regulates the photo-tactic properties of the cell (Figure 1.1).

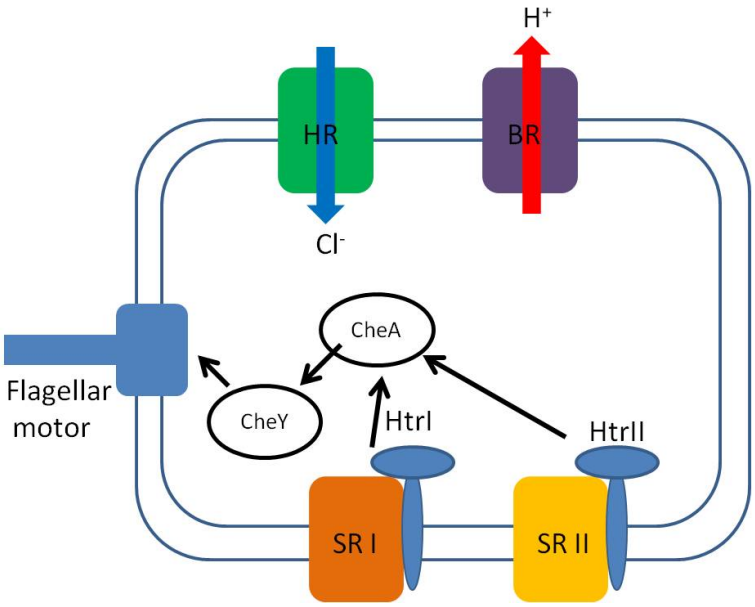

Figure 1.1: Four archaeal rhodopsins in *Halobacterium salinarum*. Bacteriorhodopsin (bR) functions as a light-driven outward proton pump; halorhodopsin (hR) functions as a light-driven inward chloride pump; sensory rhodopsin-I and -II with their transducers provide light sensing ability to the cell.

Despite the variety of functions and sequences, all members of the rhodopsin superfamily share the same structural motif; namely all of them consist of seven trans-membrane alpha helices, with a retinal covalently bound to a lysine residue at the middle of the helix G (Figure 1.3). A major difference between animal rhodopsins and microbial rhodopsins is seen in the retinal configuration in the resting state; namely, in microbial rhodopsins, retinal with the all-trans configuration is employed, and it isomerizes to 13-cis configuration when activated by light (Figure 1.2); on the other hand, animal rhodopsins employ 11-cis retinal, which isomerizes to the all-trans configuration when illuminated. Because of the structural similarity between animal and microbial rhodopsins, some scientists suggested that two protein families may have evolved from the same progenitor, though the hypothesis is still under debate.

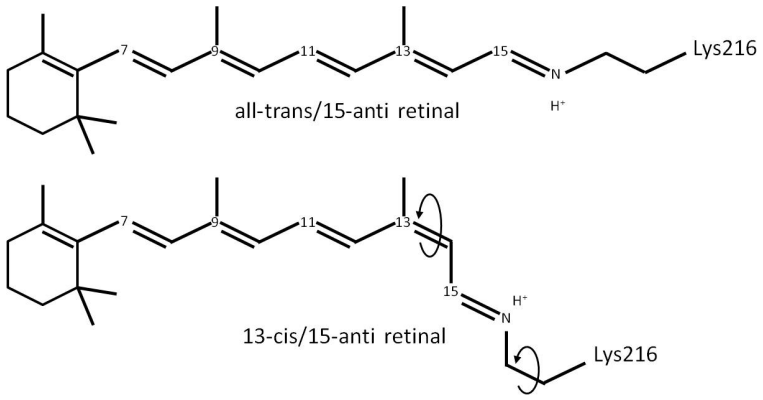

Figure 1.2: The all-trans retinal in light-adapted bR isomerizes to 13-cis / 15-anti configuration upon absorption of a photon.

1.2 Protein structure

Proteins are linear polymers built from monomeric amino acids. Twenty types of amino acids have been found in nature. Each amino acid consists of a central carbon (C_α), a carboxyl group, an amino group and a side chain (Figure 1.4 a). According to the chemical nature of the side chains, amino acids could be classified into three different

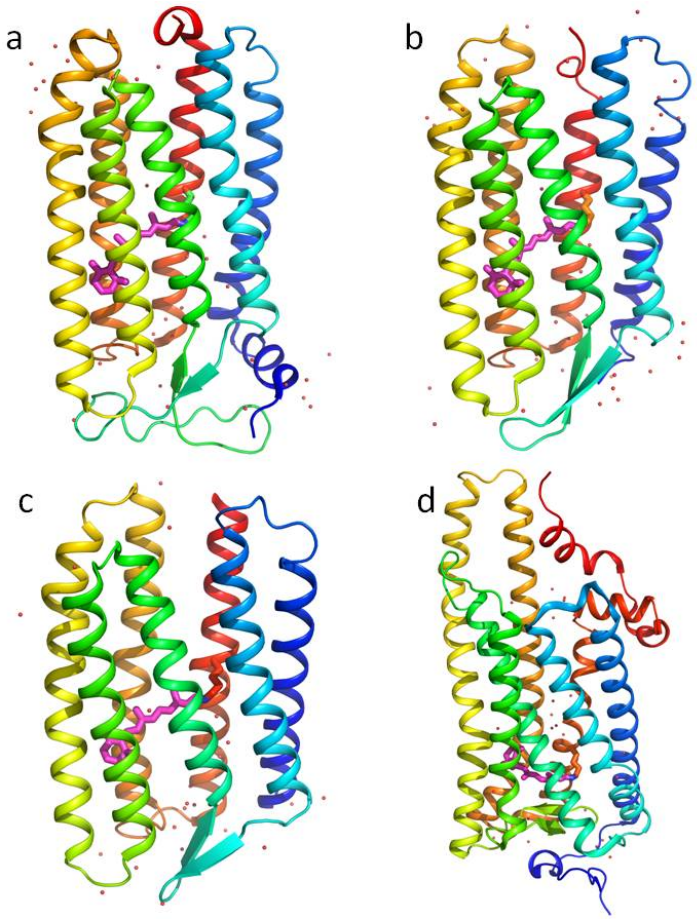

Figure 1.3: Structures of rhodopsins. a: *N. Pharaonis* halorhodopsin bromine ion bound form (pHR, pdb id: 3QBK). b: bacteriorhodopsin (bR, pdb id: 1IW6). c: sensory rhodopsin-II (SR2, pdb id: 1H68). d: squid rhodopsin (*squid*Rh, pdb id: 2Z73).

classes: hydrophobic (alanine, valine, pheylalanine, proline, methionine, isoleucine and leucine), charged (aspartic acid, glutamic acid, lysine and arginine), and polar (serine, threonine, tyrosine, histidine, cysteine, asparagine, glutamine and tryptophan). Glycine is an exceptional case because it has only a hydrogen atom as its side chain.

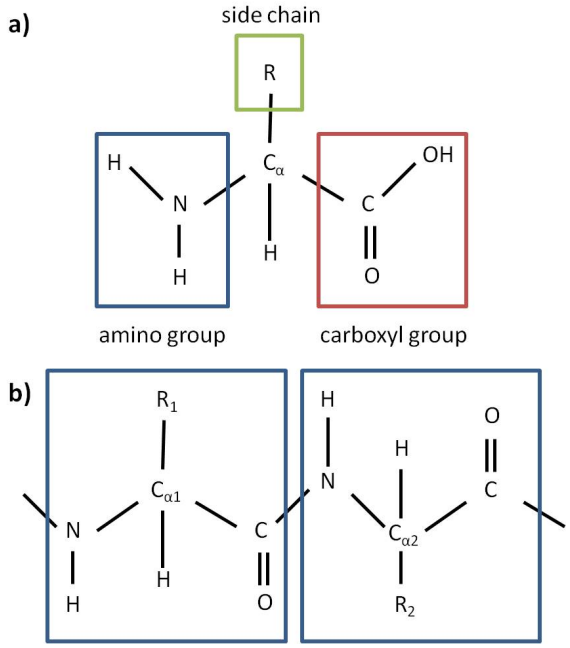

Figure 1.4: a) General structure of an amino acid in a polypeptide chain. It consists of a carboxyl group, a amino group, a central carbon and a side chain. b) Amino acids polymerized by forming a peptide bond between the amino group of one residue and the carboxyl group of a neighboring residue.

During protein synthesis, amino acids are connected via peptide bond, in which the carboxyl group of one amino acid reacts with the amino group of the neighboring amino acid to eliminate water, forming the protein main chain or backbone (Figure 1.4 b).

There are three levels of protein structure classification; namely primary structure, secondary structure and tertiary structure. Primary structure refers to the amino acid sequence of the protein polypeptide chain, in which the end with a carboxyl group is called C-terminus and the end with an amino group is called N-terminus. Each amino acid in the protein polypeptide chain is numbered according to their

position along the chain. The primary structure / amino acid sequence of a protein is always presented from N-terminus (amino group end) to C-terminus (carboxyl group end).

Secondary structure refers to the local structure formed by hydrogen bonding between amine hydrogens and carboxyl oxygen atoms contained in the main chains. There are two major types of secondary structures: alpha helix and beta sheet. In alpha helix, carbonyl group (C=O) of the n^{th} residue forms a hydrogen bond with the amino group (NH) of the $(n+4)^{th}$ residue (Figure 1.5). Each turn consists of 3.6 residues. In contrast to alpha helix which is built up from continuous residues, beta sheet is built up from different regions of a polypeptide chain. Several beta strands are aligned to form a planer sheet, in which the carboxyl group of one beta strand forms a hydrogen bond with an amino group in an adjacent beta strand.

1.3 Bacteriorhodopsin

In early 1970s, Oesterhelt and Stoeckenius reported their discovery of photo-sensitive purple membrane in *Halobacterium salinarum*, an archaeon that lives hypersaline lakes with extremely high concentration of salt. According to their report, the purple membrane consists of small patches of 0.5 μm diameter, occupying a half of the cell membrane of *Halobacterium salinarum*. The result of SDS PAGE showed that purple membrane of *Halobacterium salinarum* contains only one type of membrane protein as major component [5, 6]. This protein was named bacteriorhodopsin (bR), after the retinylidene protein found in higher animals retinae – rhodopsin. The proton pumping ability of bR and the role of converting light energy to an electrochemical energy for ATP synthesis were also determined.

Since its discovery, bacteriorhodopsin (bR) has been extensively studied for more than four decays. BR have attracted many researchers, not only because bR has high stability (in both temperature and pH) and can be easily produced in large quantities, but also because it is the simplest known biological light-driven proton pump. More than twenty molecular structures of bR at different states can be found in the pro-

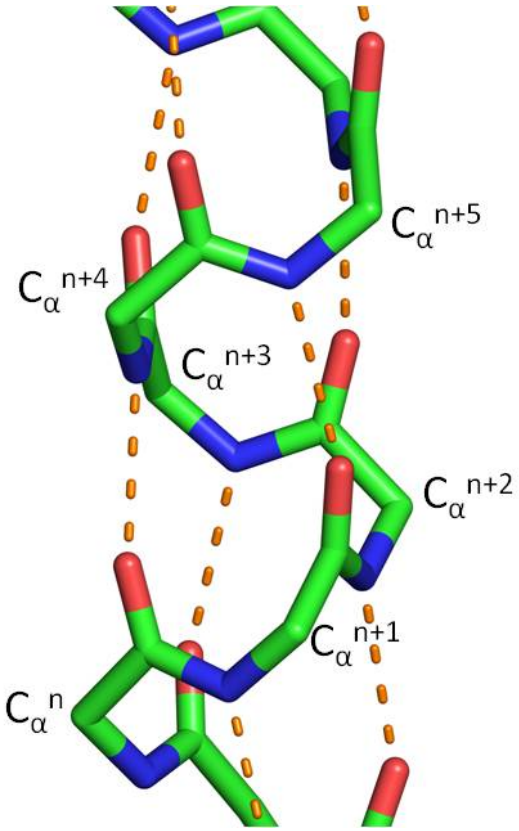

Figure 1.5: Structure of an alpha helix (the side chains are not shown). The central alpha carbon atoms are labeled, nitrogen atoms are colored in blue and oxygen atoms are colored in red. Hydrogen bonds are indicated with orange dashed lines.

tein data bank, including structures of the resting state and different photo-intermediates, structures at various pH levels, and structures of many bR mutants. Recently it was reported that a single residue modification (from Asp85 to Thr) can convert bacteriorhodopsin from a light-driven proton pump into a light-driven chloride pump [7]; bacteriorhodopsin can also be converted into a sensory receptor by modifying a few key residues [8].

1.3.1 General properties of bR

Bacteriorhodopsin is a membrane protein of 26 kDa, in the light-adapted state of which the all-trans retinal is bound via a Schiff base linkage to the ϵ-amino group of a lysine residue (Lys216) on helix G. The absorption maximum of the light-adapted state of bacteriorhodopsin is at 569 nm (green light, i.e. bR is purple in color and responsible for the characteristic purple color of *H. salinarum*).

The overall structure of bacteriorhodopsin at the light-adapted state is shown in Figure 1.6, with the functionally important residues labeled. Upon illumination, the retinal chromophore isomerizes from the all-trans to 13-cis / 15-anti configuration initiating a cyclic reaction. In a sub-millisecond time region after excitation, a proton attached to the Schiff base is transfered to Asp85. In the next step, a proton trapped at the proton-releasing group, which consists of a pair of glutamates (Glu194 and Glu204) pointing towards each others, is released to the extracellular surface. Then the cytoplasmic end of helix F and G tilts and the channel connecting Asp96 and the Schiff base is widened to accommodate several water molecules. The Schiff base regains a proton from Asp96 through a water chain, and Asp96 receives a proton from the cytoplasmic side afterwards. Finally a proton at Asp85 is transfered to the proton-releasing group and the protein conformation relaxes into the resting state.

In the dark, thermal isomerization of retinal from the all-trans to the 13-cis / 15-syn configuration takes place, yielding a mixture of the all-trans and 13-cis isomers (the dark-adapted state) (Figure 1.7). Although both the all-trans and 13-cis isomers undergo their own photocycles, there is a branching reaction from the photocycle of the 13-

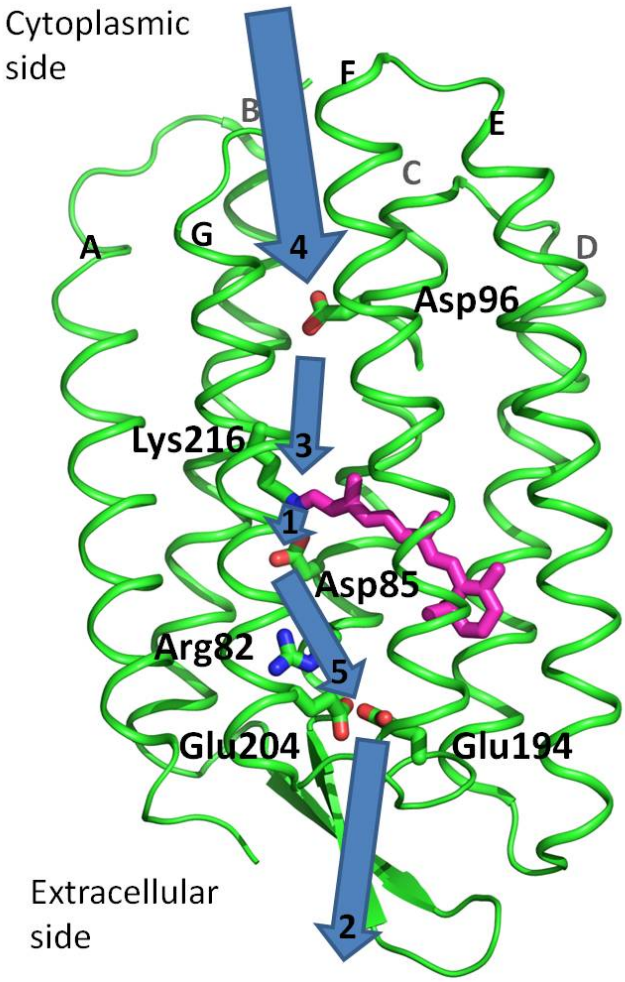

Figure 1.6: Structure of bacteriorhodopsin (pdb id: 1IW6), with the residues important for proton pumping activity labeled. Upon illumination, retinal isomerizes and subsequently (1) a proton bound to the retinal Schiff base is transfered to Asp85; (2) a proton at the proton releasing group is released to the extracellular surface; (3) the Schiff base regains a proton from Asp96; (4) Asp96 receives a proton from the cytoplasmic side; and (5) a proton at Asp85 is transfered to the proton-releasing group and (6) the protein conformation returns to the resting state.

cis isomer to the all-trans isomer. As a consequence, the dark-adapted state is rapidly converted to the light-adapted state under illumination. On the other hand, thermal equilibrium between these two isomers is slowly established in the dark, and bacteriorhodopsin in the fully dark-adapted state contains a mixture of all-trans and 13-cis isomers in a ratio of 1:1 – 1:2 [9].

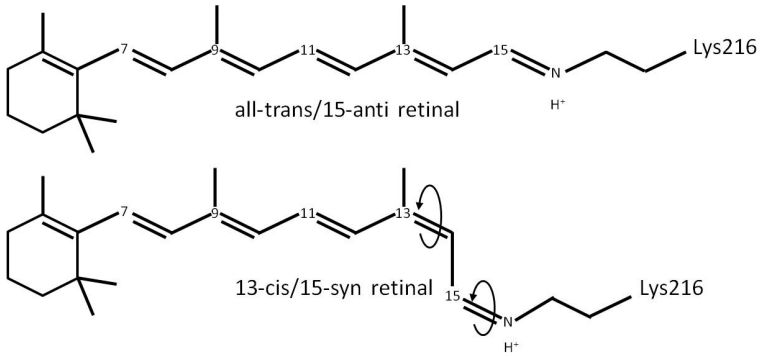

Figure 1.7: Two retinal isomers found in the dark-adapted state of bR; one of them takes on the all-trans/15-anti configuration, while the other takes on the 13-cis / 15-syn configuration. Interconversion between the two isomers is accompanied by the double isomerization around the C13=C14 and C15=N bonds.

The structure of bacteriorhodopsin is also influenced by the pH level. When the medium pH is decreased below pH 3, the absorption peak is red shifted to 604 nm (i.e. it exhibits blue in color) [10]. In the acidic blue form, the extracellular half of helix C moves towards helix G; the extracellular halves of helices D, E and F are tilted away from the center of the bR trimer; and the glutamate pair (Glu194 and Glu204) in the proton-releasing complex points away from each other. These structural changes were suggested to resemble the structural changes observed in the formation of the O intermediate. At alkaline pH (> 10), the overall structure of bacteriorhodopsin is basically conserved, but significant changes can be seen at the proton-releasing complex. The glutamate pair (Glu194 and Glu204) in the proton-releasing complex is broken in such a manner that Glu194 turns upwards pointing at Tyr83, and two additional water molecules are inserted between Glu194 and Glu204.

1.3.2 Molecular structure of ground state bR

Henderson and Unwin reported the first structural analysis on bacteriorhodopsin purple membrane by electron microscopy at 7 Å in 1975 [11]. Periodic structure with *P3* symmetry was reported, with three bacteriorhodopsin subunits sitting around a three-fold symmetric axis, forming a trimeric structure. The seven transmembrane alpha-helical segments in bacteriorhodopsin monomer were first revealed by analyzing electron micrographs of purple membrane sheets tilted from the electron beam. Meanwhile, high-resolution X-ray crystallographic studies of bacteriorhodopsin were reported in 1996 when the liqid cubic phase crystallization method was introduced [12]. Resolution at 2.4 Å was reported in 1997 [13], the resolution was then improved to 2.3 Å in 1998 [14], 1.9 Å [12] and 1.55 Å [15] in 1999, and eventually to 1.43 Å in 2002 [16].

The structure of the unphotolyzed state of bacteriorhodopsin shows that the cytoplasmic side of the proton-pumping channel is closed tightly; there is only one water molecule (Wat601) in the cytoplasmic channel. In the extracellular side, on the other hand, the space inside the channel is large enough to hold multiple water molecules. In the extracellular vicinity of the Schiff base, water molecules form a complex hydrogen-bonding network to stabilize the protonated Schiff base. Asp85, Wat602, Asp212, Wat603 and Wat604 form a pentagonal hydrogen bonding network, where the Schiff base is connected to Wat602. The hydrogen-bonding network further extends towards the extracellular surface, being mediated by Wat604, Arg82, Wat606, Glu204 and Glu194 of the proton-releasing complex (Figure 1.8). The side chains of two glutamates, Glu194 and Glu204, in the proton releasing complex are pointing towards each other. Our recent study on archaerhodopsin-2, a homologue of bR, has suggested that they are connected by a low-barrier hydrogen bond [17].

1.3.3 Proton pumping mechanism

An overview of bacteriorhodopsin structure is shown in Figure 1.6 with the key residues for proton pumping function (Asp96, Lys216, Asp85,

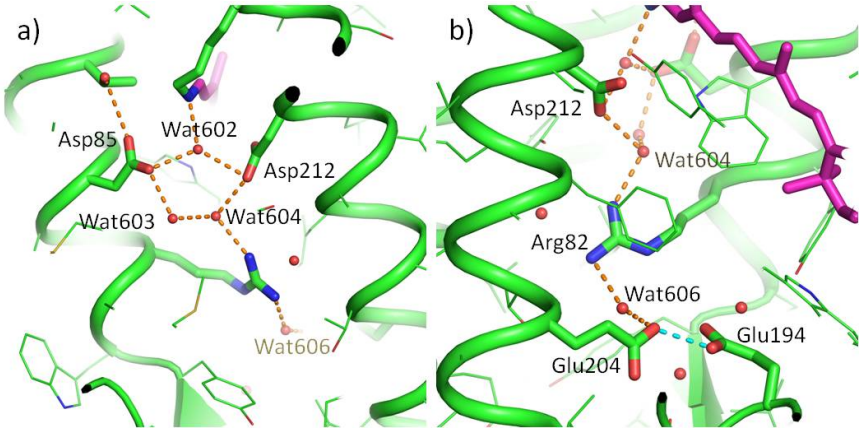

Figure 1.8: Hydrogen-bonding networks at the extracellular side of bacteriorhodopsin. a) Wat602, Asp212, Wat604, Wat603 and Asp85 form a pentagonal hydrogen-bonding network; b) The hydrogen-bonding networks connecting the Schiff base to the proton releasing group (Glu194 and Glu204).

Arg82, Glu204 and Glu194) labeled. When illuminated by light, retinal in the light-adapted state of bR isomerizes from all-trans configuration to 13-cis / 15-anti configuration (Figure 1.2), initiating a series of structural changes (photocycle); the photo-reaction intermediates can be distinguished by their absorption maximum, namely K_{590}, L_{550}, M_{410}, N_{560} and O_{640} (Figure 1.9). During the photoreaction cycle of bR, one proton is transferred from the cytoplasmic side to the extracellular side (i.e. a proton is pumped out from the cell).

A generally accepted scheme of the proton-pumping cycle of bR is described as followed: Retinal isomerizes from all-trans to 13-cis / 15-anti configuration (rest $\rightarrow$ K $\rightarrow$ L), such that the N-H dipole of the Schiff base flips from the orientation pointing towards the extracellular side to the orientation pointing towards the cytoplasmic side. This orientation change disturbs the water molecule Wat602, which is structurally important to stabilize the hydrogen network at the extracellular side. The disruption of the hydrogen bonding network accompanies significant changes in the pK_a values of the Schiff base and the nearby aspartate (Asp85), leading to reorientation of the side chain of Arg82 towards the extracellular surface. The reorientation of Arg82 also causes the proton release from the proton-releasing

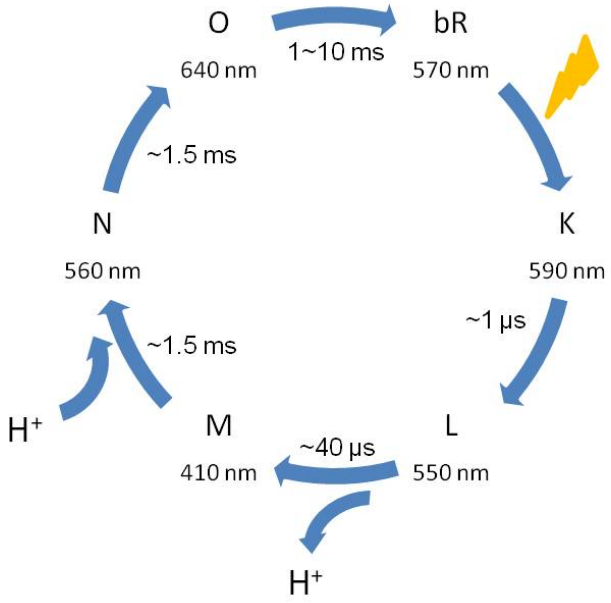

Figure 1.9: A simplified scheme of the proton-pumping cycle of bacteriorhodopsin. After illumination, bR undergoes a series of structural changes which can be identified by flash-induced absorption changes. The decay time constant of each reaction state at neutral pH is shown on the arrow.

complex (Glu194 and Glu204) to the extracellular space.

A large conformation change (outwards tilt of cytoplasmic half of helix F, away from helix G) opens the cytoplasmic channel and allows water molecules to form a proton transfer pathway between the Schiff base and Asp96, by which the Schiff base regains a proton from Asp96 (M $\rightarrow$ N). Asp96 then regains a proton from the cytoplasmic side, and the retinal re-isomerizes to all-trans configuration (N $\rightarrow$ O). The deprotonation at Asp85 relaxes the twisted retinal, and the proton releasing group regains a proton to complete the cycle. As a result, one proton is pumped from cytoplasmic side towards the extracellular side during the photo-cycle.

1.4 Homologous proteins of bacteriorhodopsin in archaea

Since the discovery of the bacteriorhodopsin, a blast of homologs of bR were discovered in different species of archaea. Up to date, more than 25 archaeal retinal proteins have been identified and their sequences are available. Archaerhodopsin-2, -3 and cruxrhodopsin-3 are discovered recently as the homologs of the well known bacteriorhodopsin. Although the sequence identity of aR2 to bR is not high (55 %)(Figure 1.10), the sequence identity between aR2 and aR3 is high (86 %). All of them function as light-driven proton pumps.

Archaerhodopsin-2, which was found in *Halorubrum* sp. *aus-2*, was crystallized into a crystal belonging to space group $C222_1$ by Enami et al. in 2006 [18]. Later two different crystal forms (space group *P321* and $P6_322$) of aR2 were prepared by Yoshimura and Kouyama in 2006 [19]. More recently, we have obtained another crystal form (space group *H32*) of aR2 that diffracts X-rays to 1.8 Å [17]. Interestingly, although these crystal forms were grown under similar conditions, there is a significant difference in the protein packing; namely, whereas the trimeric assembly of aR2 is retained in the *P321*, $P6_322$ and *H32* crystals, it is broken in the $C222_1$ crystal.

Archaerhodopsin-3, a homolog of aR2 with high sequence identity (86 %), was found in *Halorubrum sodomense*. Archaerhodopsin-3 is a

powerful tool commonly used as neuron silencer in optogenetics, yet very few structural nor absorption kinetics data about aR3 was reported. The sequence identity of aR3 to bR is 61 %. Cruxrhodopsin-3, a new member of the archaeal rhodopsin family, was found in *Haloarcula vallismortis* [20]. The sequence identity of cR3 to bR is low (54 %).

Structural analyses of homologous proteins of bR would allow us to clarify the important structural motifs for the proton pumping activity.

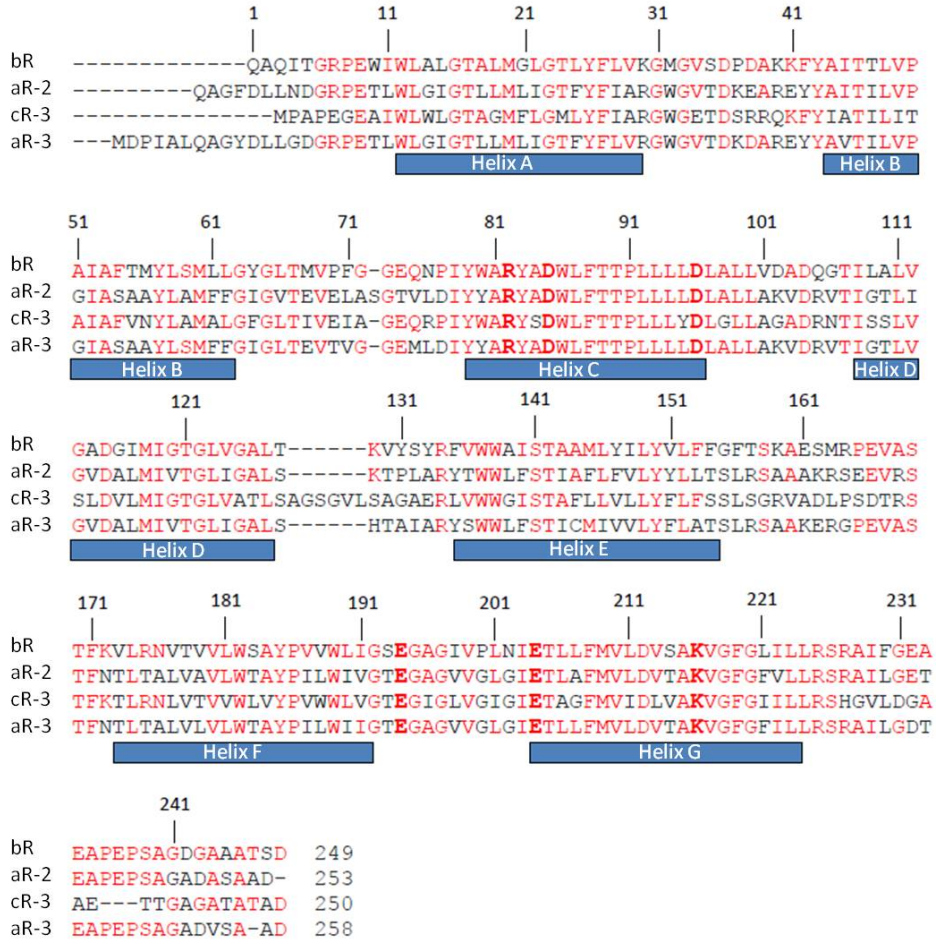

Figure 1.10: Amino acid sequences of bR, aR2, cR3, and aR3. The concerved residues are shown in red, while key residues for proton pumping activity are shown in bold style. The rectangles below the sequences indicate the regions forming α-helices in bR.

Chapter 2

X-ray crystallographic study

The first protein crystal reported in history was a hemoglobin crystal by Hunefeld in 1840 [21]. At that moment, crystallization of protein was just a procedure to extract and purify protein. In 1912, X-ray diffraction technique was first reported by Max Von Laue, who won the 1914 Nobel Prize in physics. The first protein structure was published in 1950s, and afterwards scientists started to put efforts on getting structural information of biological macromolecules through X-ray crystallography. The work on structure of myoglobin led in 1962 to the award to Max Perutz and John Kendrew of a Nobel Prize in chemistry. Up to date, roughly 95,000 protein structures determined by X-ray crystallography have been stored in Protein Data Bank (PDB) [22], and more than 10 Nobel Prizes in chemistry or medicine have been awarded to protein crystallographic studies including the recent 2012 Nobel Prize in chemistry awarded to Lefkowitz and Kobilka on the study of G-protein-coupled receptors (GPCRs). Protein crystallographic technique is also utilized for drug design.

Membrane proteins play important roles in maintaining various biological activities, including membrane trafficking and environmental sensing. It is estimated that about 30 percent of the gene codes correspond to membrane proteins. More than half of the pharmaceutical products target on membrane proteins. However, the importance of membrane proteins did not make their molecular structures very well studied. There are roughly 1,000 published reports of membrane pro-

tein structures in the PDB up to 2014 [1].

Generally, protein crystallographic studies on membrane proteins are carried out according to the following processes: Expression of protein, extraction and purification of protein, crystallization, obtaining diffraction data, building molecular model and refinement (Figure 2.1).

2.1 Expression, extraction and purification

Since crystallization requires a large amount of a target membrane protein (in milligram quantities), proteins are sometime expressed in another host cell (e.g. E. coli) to shorten the culturing time and increase the yield of protein production. To produce a sufficient amount of membrane protein, usually larger than 20 liter of culture solution is prepared; after the cells are harvested, the target protein has to be extracted and purified.

One of the main difficulties in studying the membrane proteins is that membrane proteins are usually unstable once leaving the membrane environment, so that the proteins become difficult to handle. This problem is overcame by using mild detergents, which consist of a polar head group and a hydrocarbon tail, to mimic the environment of membrane.

2.2 Crystallization

Three major crystallization methods have been developed for crystallization of membrane proteins; namely the lipid cubic phase method, the lipid bicelle method and the membrane fusion method. In this study, the membrane fusion method was used to crystallize cR3 and aR3.

2.2.1 Lipid cubic phase method

Crystallization by the lipid cubic phase (LCP) method was first introduced by Landau and Rosenbusch in 1996 [12]. This method is

[1] Stephen White Lab at UC Irvine, http://blanco.biomol.uci.edu/mpstruc/

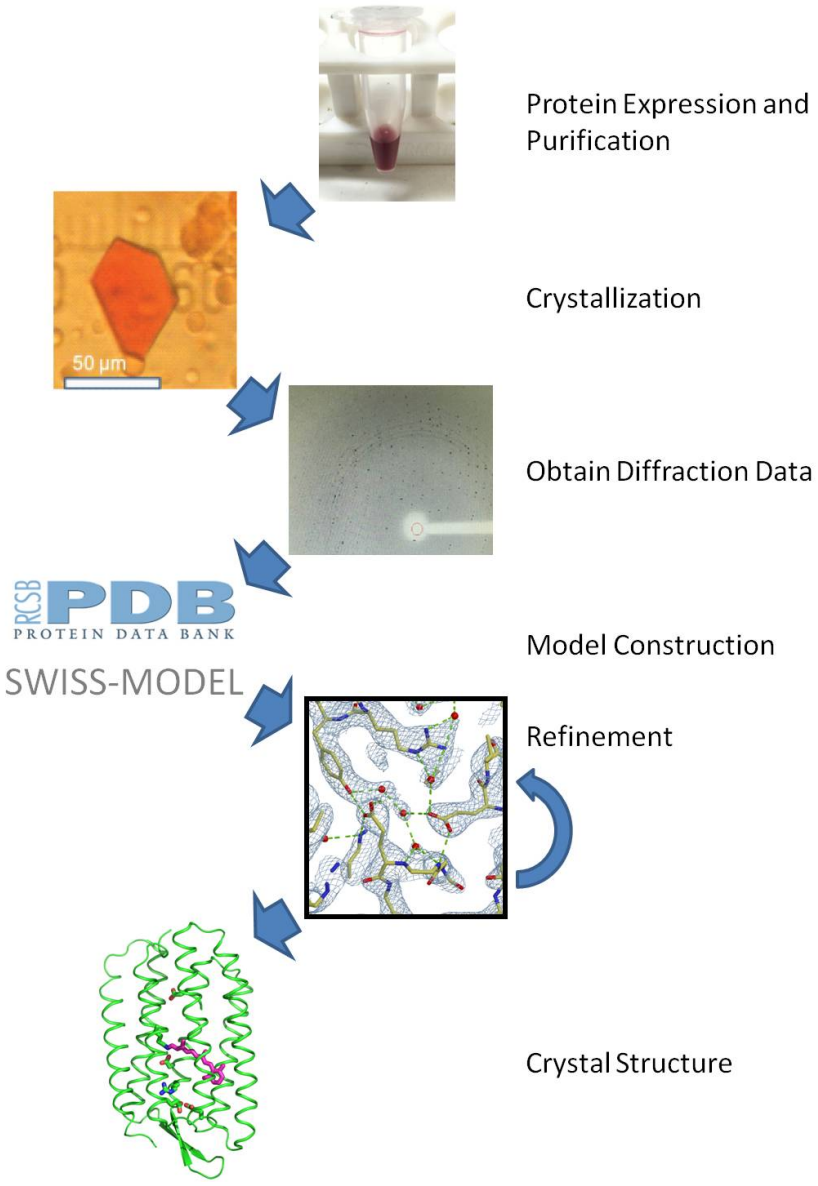

Figure 2.1: Flow chart of experimental procedures that are necessary for protein crystallography.

specifically designed for crystallizing membrane proteins. Under appropriate proportion of water, protein and lipid, three-dimensional lipid array can be produced. Such matrices provide sites with membrane-like environment for membrane proteins to nucleate and diffuse. Bacteriorhodopsin crystals with the size about 0.1 mm x 0.1 mm x 0.03 mm were observed, and diffracted to 3.7 Å resolution, with a space group $P6_3$ ($a = b = 62$ Å, $c = 108$ Å; $\alpha = \beta = 90°$ and $\gamma = 120°$). Diffraction data at 2.4 Å resolution was reported in 1997 [23].

2.2.2 Lipid bicelle method

The lipid bicelle method was invented by Faham and Bowie in 2002 [24]. Unlike the LCP method, in the lipid bicelle method membrane proteins were incorporated into bicelles (small bilayer disks of lipid / amphiphile mixtures), the membrane proteins embedded bicelles suspension was then mixed with precipitant solution in the well of sitting drop apparatus and incubated at 37 °C. Bacteriorhodopsion crystals belong to space group $P2_1$ ($a = 45.0$ Å, $b = 108.9$ Å, $c = 55.9$ Å, $\beta = 113.58°$) were observed, and diffraction data at 2.0 Å resolution was reported. The outline of bicelle method is shown in Figure 2.2.

2.2.3 Membrane fusion method

The membrane fusion method was reported by Takeda et al. in 1998. The main concept is shown in Figure 2.3; protein-rich membranes were converted into vesicles by incubating at high temperature (32 °C) with a small amount of detergent, after mixing with precipitant, membrane vesicles condense and fuse to yield a three-dimensional crystal. The introduction of the membrane fusion method has successfully yielded high quality protein crystals including bacteriorhodopsin [25], archaearhodopsin-1 and 2 [18, 26], cruxrhodopsin-3 [1], deltarhodopsin-3 [27], halorhodopsin [28], and squid rhodopsin [29]. One advantage of the membrane fusion method over the other two is that not much detergent is used throughout the crystallization process, so that the structures of membrane proteins to be crystallized are not as much altered. In both LCP method and lipid bicelle method, a step involving protein

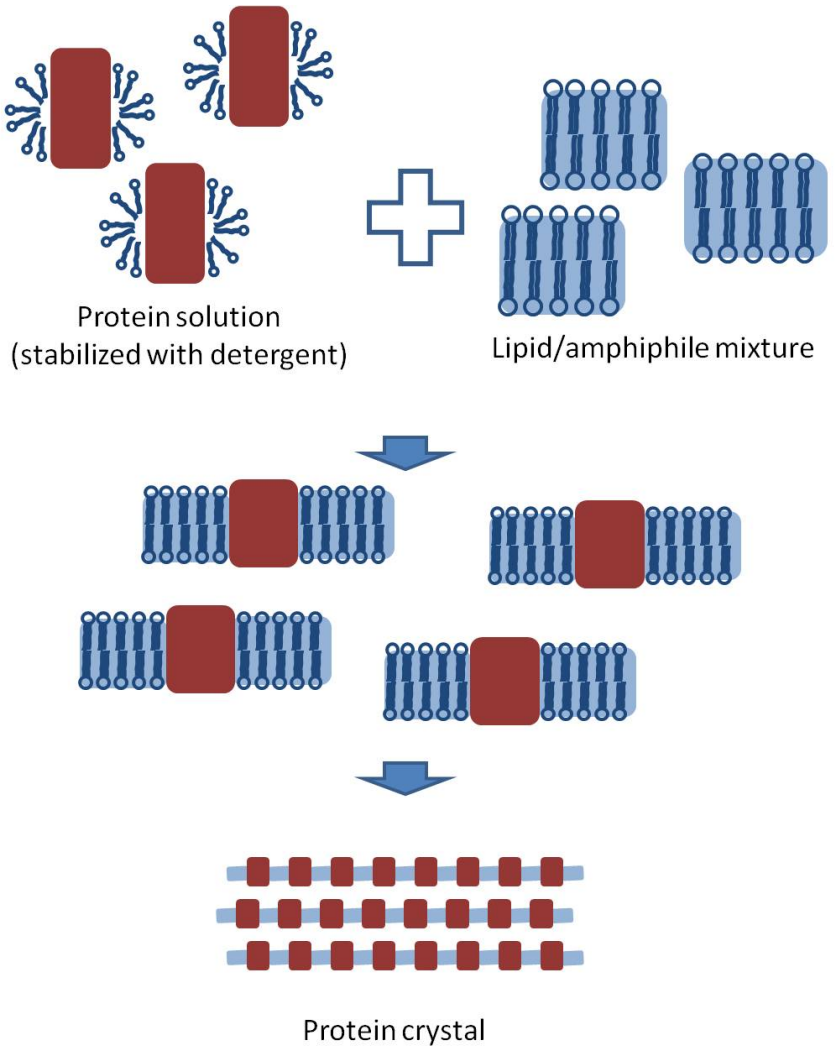

Figure 2.2: Schematic diagram of crystallization by the bicelle method. Protein-rich membrane is re-constructed by mixing purified protein with lipid / amphiphile mixture; and protein crystals are grown after mixing the reconstructed bicelle with precipitant.

solubilization is included. This step might alter the lipid distribution around the protein. It has been reported that lipids play an important role in stabilizing and supporting bacteriorhodopsins function [10, 30].

The importance of lipid composition to protein structure can be illustrated by comparing the crystal structures of bacteriorhodopsin yielded by these methods. The bR crystals grown by the bicelle lipid method formed bR dimers as asymmetric unit; while the native trimeric asymmetric unit with honeycomb-like lattice can be reporduced by the bR crystals yielded by the membrane fusion method and LCP method. On the other hand, the proton-releasing group, a paired structure of glutamic acids (Glu204 and Glu194) that is maintained by a low barrier hydrogen bond, was broken in some crystals of bR prepared by LCP method; furthermore, the lipid components observed in the central opening of the bR trimer in the crystals prepared by the membrane fusion method were missing in the crystals prepared by the LCP method, suggesting that native lipid components are important to stabilize the native structure of bR.

2.3 Collection of diffraction data

2.3.1 Synchrotron radiation

X-ray is an electromagnetic wave with a frequency range from 3×10^{16} Hz to 3×10^{19} Hz. Synchrotron is a circular device, in which charged particles (usually negatively charged electrons or positively charged positrons) circulate in a speed near to the speed of light. It is originally designed as a particle collider in high energy physics research. The charged particles are accelerated and stored in the storage ring; the circulating charged particles accelerate towards the center of the ring and therefore emit electromagnetic radiation. This byproduct of high energy physics research was found to be very useful for biologists and chemists in determining structures of macromolecules.

The main advantage of synchrotron radiation over the traditional in-house X-ray tubes is that the synchrotron radiation has much higher intensity (several orders of magnitude). The high intensity of X-ray allows shorter illumination period and data collection from weakly

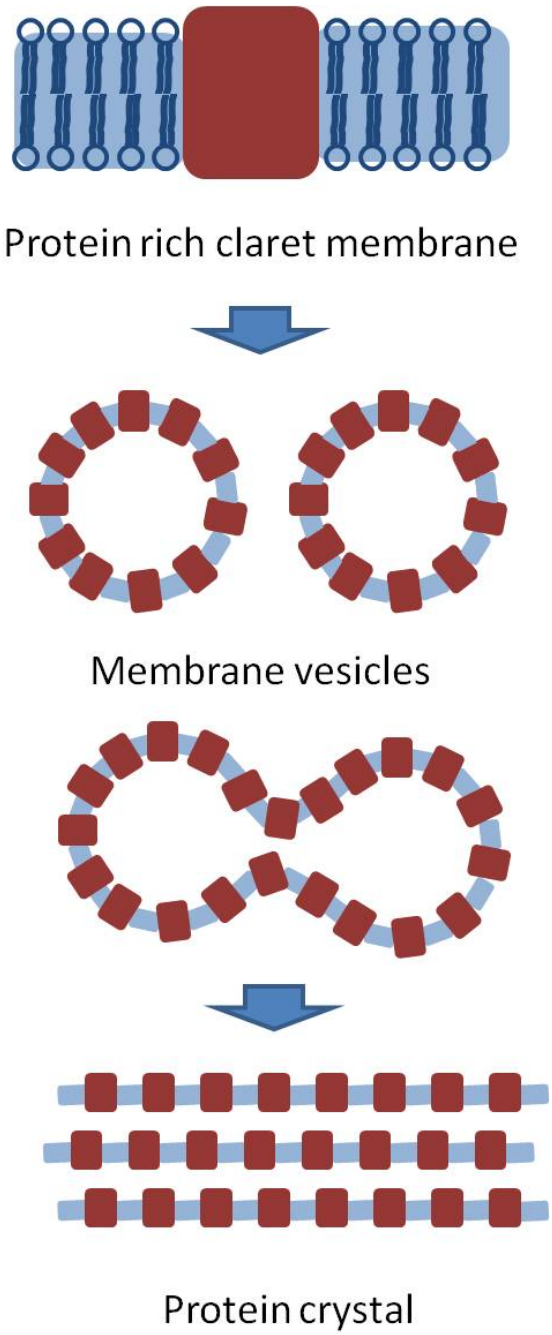

Figure 2.3: Schematic diagram of protein crystallization by the membrane fusion method. Purple membrane sheets are converted to vesicles by incubating the membranes with a small amount of detergent at a high temperature. Then vesicles are fused with one another, yielding a protein crystal in which membranous structures are piled up.

diffracting crystals or crystals with very large unit cells can be benefited. However high intensity X-ray also increases the chance of X-ray induced structural change (damage) of the proteins. Previous studies have shown that the X-ray damage can be suppressed by cooling the protein crystal to cryogenic temperatures[31], though specific damages in local regions remain unavoidable even at a cryogenic temperature [32].

Another advantage of synchrotron radiation over X-ray tubes is that the wavelength of synchrotron radiation can be adjusted by using a monochromator. In a X-ray tube, the electrons are emitted from the cathode, accelerated by high voltage and reach the target anode at a very high speed, and the high speed electrons bombard and eject the electrons from the low-energy orbitals of the atoms in the anode. X-rays are emitted when electrons transit from high-energy orbitals to low-energy orbitals to fill the vacancies. The wavelength of X-ray emitted is specific according to the anode material and cannot be tuned (e.g. x-ray wavelength emitted by an X-ray tube which use copper as anode is 1.54 Å). On the other hand, free electrons (charged particles) traveling in the synchrotron storage ring are not quantized and therefore, the wavelength of emitted X-ray is variable and determined according to the energy of the circulating charged particles and the applied magnetic field (to accelerate the charged particles). Usually X-rays at 1 Å or even shorter wavelength is used in protein X-ray diffraction experiments.

2.3.2 Data collection strategy

Crystal soaking and mounting

Protein crystals are soaked into the soaking solution with the environment crystallographers are interested in (e.g. pH value, metal ion or halide ion). A single crystal soaked in soaking solution is picked up with a cryoloop, which is a nylon fiber loop attached on the tip of a metal stick. The solution droplet stays in the cryoloop due to the surface tension. During data collection of X-ray diffraction from a protein crystal, cryogenic temperature is applied to reduce the X-ray damage. It is, however, possible that ices form during the cooling

process. Diffraction from ices interferes with the diffraction pattern of protein crystal. The formation of large ice crystals can be prohibited by adding cyroprotectant such as trehalose.

Rotation method

The diffraction pattern of a crystal is determined by the Bragg's law. The diffraction of X-rays is considered to be equivalent to X-ray reflection by the lattice planes with the crystal. The Bragg's law is given by

$$\frac{1}{d} = \frac{2\sin\theta}{\lambda} \tag{2.1}$$

in which d is the distance between reflection planes, λ is the wavelength of the incident X-ray, and θ is the angle of reflection. By using Ewald's sphere, we can illustrate the Bragg's law in a three-dimensional graphical space. The reciprocal space of a crystal is represented by a reciprocal lattice, while the incident X-ray is represented by a wave vector **S**. The Ewald's sphere is a sphere with a radius equal to the length of the incident ray wave vector and centered at the origin of the crystal reciprocal lattice [circle in 2-D case (Figure 2.4 a)]. The Bragg's law is fulfilled when the reciprocal lattice point lies on the surface of the Ewald sphere. As a result, diffraction beams pointing from the center of the Ewald's sphere towards the reciprocal lattice points that lie on the surface of Ewald sphere can be observed (Figure 2.4 b).

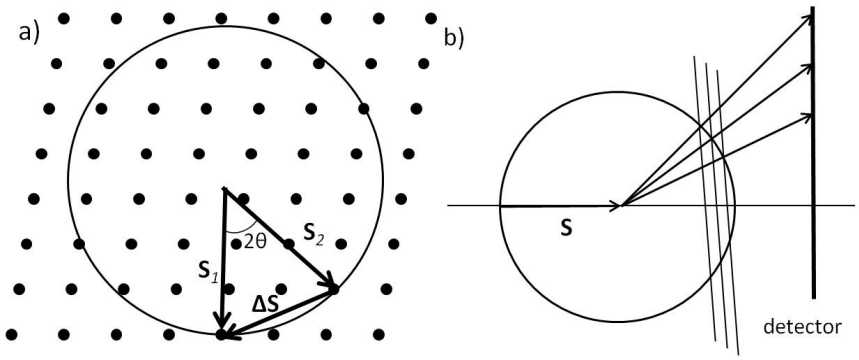

Figure 2.4: a) Ewald construction. b) Schematic diagram for prediction of X-ray diffraction pattern.

The number of the diffraction points that can be detected depends on the density of reciprocal lattice point (i.e. the size of the unit cell) and the radius of the Ewald sphere (the length of wave vector $1/\lambda$). The size of unit cell of protein crystal is usually much larger than the wavelength of the applied X-ray, only several reciprocal lattice points lie on the surface of the Ewald sphere, i.e. a very few diffraction points can be detected for a crystal mounted at a random orientation. In precession photography, the crystal and the camera are fixed when a photograph is taken and therefore, an alignment process is needed so that maximum number of the diffraction spots can be observed.

In the rotation photography, on the other hand, a single crystal is rotating around the spindle axis while a photograph is taken. Because the reciprocal lattice follows the rotation of the crystal in real space, the surface of Ewald sphere sweeps through a reciprocal lattice plane, generating a series of diffraction spots. A typical example of photographs taken in rotation photography is as shown in Figure 2.5; the nearly circular concentric pattern is called lune, and each lune arises from a single plane of reciprocal lattice. As diffraction spots in a single rotation photograph are partially recorded, their full intensities are recorded by rotating the crystal over a large rotational angle.

2.4 X-ray diffraction

2.4.1 X-ray diffraction by a crystal

When electromagnetic waves collide with an atom, the electromagnetic wave exerts a force on a charged particle (electrons and protons), causing it to oscillate with the same frequency as the incident ray. The oscillating charged particle acts as a secondary source of electromagnetic wave (However, protons are tightly bound in nucleus with neutrons and the mass to charge ratio of proton is much larger than that of electrons, the contribution by proton is negligible). X-ray diffraction experiments provide structure information by measuring the interference pattern of the scattered X-ray.

The intensity of X-rays scattered to a certain direction is determined by the interference of the X-rays scattered by many electrons within

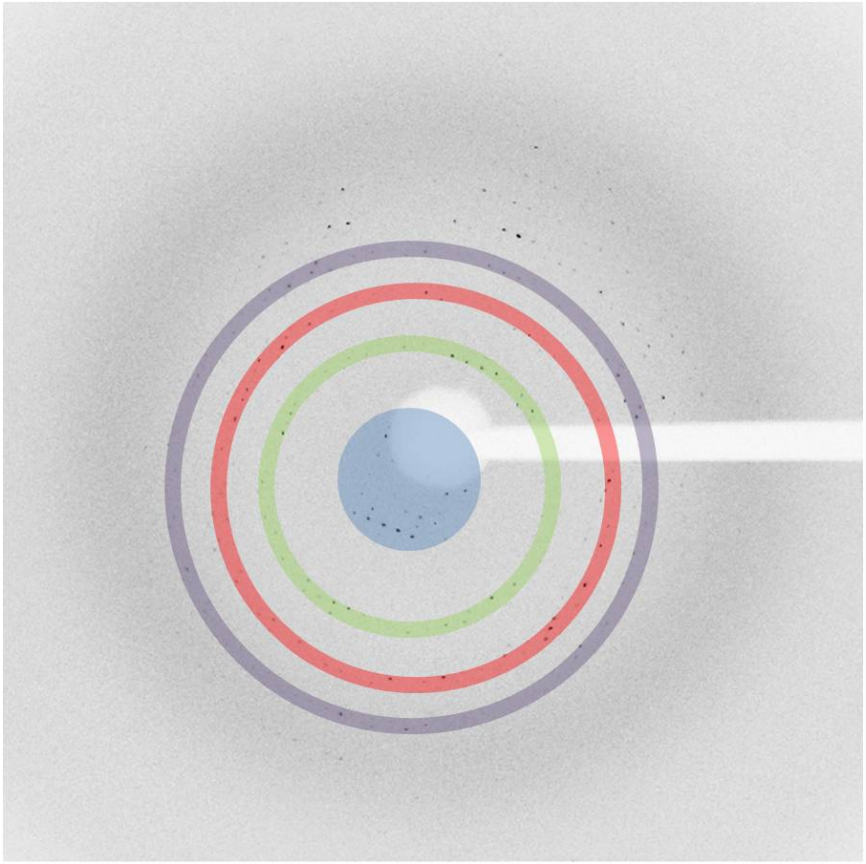

Figure 2.5: A 1° rotation photo of cR-3 taken at beamline BL38-B1 in SPring-8. Diffraction spots arising from different reciprocal lattice planes are observed in different lunes marked with different colors.

the crystal. Figure 2.6 shows a two-electrons system, where the path difference $p + q$ depends on the relative position of two electrons ($\mathbf{r}$) and the direction of diffraction (2θ). The path difference is $p + q = \lambda \cdot \mathbf{r} \cdot (\mathbf{s_0} - \mathbf{s})$, where $\mathbf{s_0}$ and $\mathbf{s}$ are the wave vectors of incident and scattered waves respectively. The phase difference is

$$\frac{2\pi \mathbf{r} \cdot (\mathbf{s_0} - \mathbf{s}) \cdot \lambda}{\lambda} = 2\pi \mathbf{r} \cdot \mathbf{\Delta S} \tag{2.2}$$

with $|\mathbf{\Delta S}| = 2 \sin\theta/\lambda$

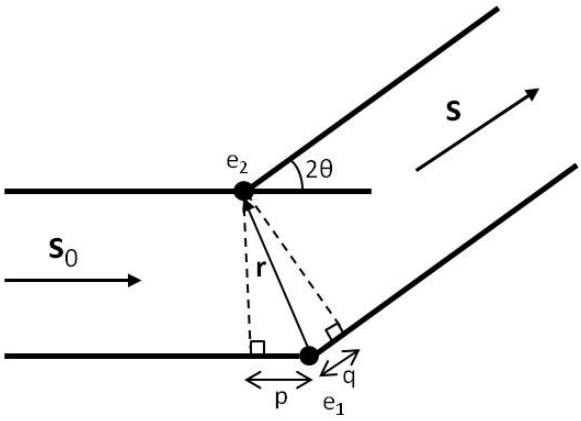

Figure 2.6: Scattering of X-ray by a two-electron system. The path difference between the two scattered waves is $p + q$.

The distribution of electrons of an atom are described by an electron density $\rho(\mathbf{r})$. The atomic scattering factor f is the Fourier transform of the electron density. i.e.

$$f = \int_{\mathbf{r}^3} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}) \mathbf{d^3 r} \tag{2.3}$$

The atomic scattering factors of different atoms can be looked up in tables, in which f is expressed as a function of $|\mathbf{S}| = 2 \sin\theta/\lambda$. The wave scattered by an atom is given by

$$\mathbf{f_i} = f_i \exp(2\pi i \mathbf{r}_i \cdot \mathbf{S}) \tag{2.4}$$

For a unit cell with n atoms, the total scattering is then

$$\mathbf{F}(\mathbf{S}) = \sum_{i=1}^{n} f_i \exp(2\pi i \mathbf{r}_i \cdot \mathbf{S}) \tag{2.5}$$

where $\mathbf{F}(\mathbf{S})$ is called the structure factor as it reflects the configuration of the atoms in the unit cell.

The scattering by a crystal is the sum of the scattering of X-rays from all unit cells within it. Consider translation vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ and assume that the position vector of all unit cells can be expressed as $t \cdot \mathbf{a} + u \cdot \mathbf{b} + v \cdot \mathbf{c}$, where t, u, and v are integers. The total scattering is

$$\mathbf{K}(\mathbf{S}) = \mathbf{F}(\mathbf{S}) \times \sum_{t=0}^{n_1} \exp(2\pi i t \mathbf{a} \cdot \mathbf{S}) \times \sum_{u=0}^{n_2} \exp(2\pi i u \mathbf{b} \cdot \mathbf{S}) \times \sum_{v=0}^{n_3} \exp(2\pi i v \mathbf{c} \cdot \mathbf{S}) \tag{2.6}$$

with n_1, n_2 and n_3 are the numbers of unit cells along $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ directions, respectively. Since n_1, n_2 and n_3 are very large, the summation $\sum_{t=0}^{n_j} \exp(2\pi i t \mathbf{a} \cdot \mathbf{S})$ is zero unless $\mathbf{a} \cdot \mathbf{S}$ is an integer (usually labeled as h, k and l), i.e. when $\exp(2\pi i t \mathbf{a} \cdot \mathbf{S})$ is equal to 1. This condition is known as Laue diffraction conditions,

$$\mathbf{a} \cdot \mathbf{S} = h \tag{2.7}$$

$$\mathbf{b} \cdot \mathbf{S} = k \tag{2.8}$$

$$\mathbf{c} \cdot \mathbf{S} = l \tag{2.9}$$

in which h, k, and l are integers (either positive, negative or zero).

2.4.2 The temperature factor (B-factor)

The atoms in a crystal vibrate around its equilibrium position at non-zero temperature. As a result, there may be a slight difference in the relative positions of the equivalent atoms in successive unit cells. The atomic scattering factor of the atom is multiplied by a temperature factor to descript this phenomenon. When the vibration is isotropic (i.e. same magnitude in all directions), the temperature factor is given

by

$$T(iso) = \exp\left[-B\frac{\sin^2\theta}{\lambda^2}\right] \tag{2.10}$$

with B as the thermal parameter that is proportional to the mean square displacement $\bar{u^2}$ of the atom vibration.

$$B = 8 \times \pi^2 \times \bar{u^2} \tag{2.11}$$

In protein crystallography the temperature factor (B-factor) can also be increased by the static disorder of the protein structure; i.e. when the protein consist of more than one configuration. For example, B-factor of the C_α atoms of cR3 crystal structure is shown in figure 2.7. The B-factors of C_α atoms of the helices regions is significantly lower than that of the flexible loop regions.

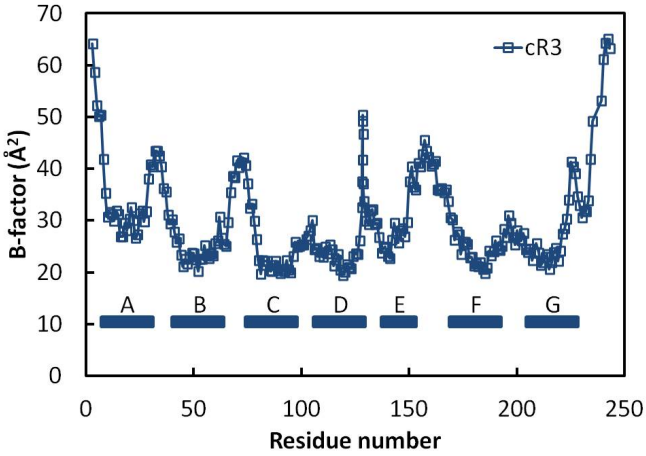

Figure 2.7: The B-factor of C_α atoms of cR3. Residues forming α helices are more stable and thus their C_α atoms have low B-factors.

The static disorder effect cannot be distinguished from the dynamic disorder caused by thermal vibration unless diffraction data at different temperatures are collected. The diffraction pattern fades away at certain diffraction angle θ_{max} due to the disorder of the crystal, the resolution of the diffraction pattern is defined as the lattice distance d_{min}

$$d_{min} = \frac{\lambda}{2\sin\theta_{max}} \tag{2.12}$$

The accuracy of the structure increases as the value of resolution decreases.

From diffraction pattern to electron density $\rho(\mathbf{x}, \mathbf{y}, \mathbf{z})$

The intensity of the diffracted X-ray is proportional to the magnitude of the square of the structure factor $|\mathbf{F}(\mathbf{S})|^2$, in which the structure factor $\mathbf{F}(\mathbf{S})$ is a function of the electron density distribution in the unit cell.

$$\mathbf{F}(\mathbf{S}) = \int_V \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}) dV \tag{2.13}$$

where V is the volume and $\rho(\mathbf{r})$ is the electron density distribution of the unit cell. By rewriting the coordinate as fraction coordinate of the unit cell (x, y, and z, such that $dV = V dx dy dz$), and by the Laue diffraction condition,

$$\mathbf{r} \cdot \mathbf{S} = hx + ky + lz \tag{2.14}$$

we have,

$$\mathbf{F}(hkl) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x, y, z) \exp(2\pi i(hx + ky + lz)) dx dy dz \tag{2.15}$$

The structure factor $F(hkl)$ is the Fourier transform of the electron density. By reverse Fourier transformation, we can obtain the electron density as:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}(hkl) \exp\left[-2\pi i(hx + ky + lz)\right] \tag{2.16}$$

The integration is replaced by summation over all Miller indexes (h, k and l) because the diffraction pattern is discrete according to Laue diffraction condition. Although the value of $\mathbf{F}$ can be determined from the intensity $I(hkl) = |\mathbf{F}(\mathbf{S})|^2$, the phase angle of the structure factor ($\exp(i\alpha)$) cannot be observed directly from the diffraction pattern. To determine the phase information, indirect methods were developed, namely the molecular replacement method and isomorphous replacement and so on. In this study the molecular replacement is used, in

which priori information of a homologous structure is used to determine the phase information. A brief introduction on the molecular replacement method is as followed.

2.4.3 Molecular replacement method

For the proteins with any homologous protein in which its structure has already been determined, the simplest way to determine the phase information is by the molecular replacement method. The molecular replacement method 'borrows' the phase information of the homologous protein as the initial phase of the structure factor $F(hkl)$ for further refinement. The molecular replacement involve of two steps: rotation and translation. In the rotation step the orientation of the unknown structure is rotated to match the orientation of the reference structure. In the translation step two structures are superimposed.

Patterson function

The Patterson function is a Fourier transform of intensities of the diffraction pattern. It can also be shown that Patterson function is equal to the convolution of electron density with itself.

$$P(\mathbf{u}) = \frac{1}{V}\sum_{\mathbf{S}} |\mathbf{F}(\mathbf{S})|^2 \exp\left[-2\pi i \mathbf{u}\cdot\mathbf{S}\right] = \int_{\mathbf{r}} \rho(\mathbf{r}) \times \rho(\mathbf{r}+\mathbf{u}) \qquad (2.17)$$

in which V is the volume of unit cell and the integration is for $\mathbf{r}$ over all positions in the real unit cell.

From equation 2.17 it is expected that $\rho(\mathbf{r}) \times \rho(\mathbf{r}+\mathbf{u})$ is significant only when $\mathbf{u}$ is a vector connecting two atoms (both from atom 1 to atom 2 and vice versa). Patterson function provides information of the relative position of atoms. Depending on the atoms to be considered, Patterson map can be divided into two regions: if the atoms connected by $\mathbf{u}$ is within the same molecule, they are called self-Patterson vectors; the vectors connecting intermolecular atoms are called cross-Patterson vectors. For the same molecule in different crystal forms, the inner region of Paterson map (i.e. the self-Patterson vectors region) should be the same apart from a rotation difference (that of homologous pro-

teins is not equal but similar); from the cross-Patterson vectors we can derive the translation required to move the molecules to their correct positions.

Rotation

For an atom at position $\mathbf{u}$, if we rotate it while keeping the same origin, the new position is related to the original position by a rotation matrix $[C]$, i.e. $\mathbf{u}_r = [C]\,\mathbf{u}$. Applying the rotation matrix $[C]$ to the Patterson function $P(\mathbf{r})$ gives the rotated Patterson function $P_r(\mathbf{r}_r)$.

$$P_r(\mathbf{u}_r) = \frac{1}{V}\sum_{\mathbf{h}} |\mathbf{F}(\mathbf{h})|^2 \exp\left[-2\pi i \mathbf{h}\mathbf{u}_\mathbf{r}\right] \tag{2.18}$$

Since $\mathbf{u}_r = [C]\,\mathbf{u}$,

$$P_r(\mathbf{u}_r) = \frac{1}{V}\sum_{\mathbf{h}} |\mathbf{F}(\mathbf{h})|^2 \exp\left[-2\pi i \mathbf{h}\,[C]\,\mathbf{u}\right] \tag{2.19}$$

rewrite $\mathbf{h}\,[C]$ as $[C^{-}1]\,\mathbf{h}$ and $\mathbf{h}$ as $[C]\,\mathbf{h}'$, we have

$$P_r(\mathbf{u}_r) = \frac{1}{V}\sum_{[C]\mathbf{h}} |\mathbf{F}([C]\,\mathbf{h})|^2 \exp\left[-2\pi i \mathbf{h}\mathbf{u}\right] \tag{2.20}$$

The overlap function R is defined as followed:

$$R = \int_U P(\mathbf{u_1}) P_r(\mathbf{u_{1r}}) d\mathbf{u_1} \tag{2.21}$$

where U is the volume of self-Patterson map. The overlap function gives maximum value when the rotation is optimized.

Now

$$R = \frac{1}{V^2}\sum_{\mathbf{h}}\sum_{\mathbf{h}'} |\mathbf{F}(\mathbf{h})|^2 |\mathbf{F}([C]\,\mathbf{h})|^2 \int_U \exp\left[-2\pi i(\mathbf{h}' + \mathbf{h})\mathbf{u}\right] d\mathbf{u} \tag{2.22}$$

The integral part can be written as $U/V|\mathbf{G}\,[-(\mathbf{h}+\mathbf{h}')]$, if we further assume that the protein molecule U has a spherical volume, then

$$G = \frac{3(\sin 2\pi x - 2\pi x \cos 2\pi x)}{(2\pi x)^3} \tag{2.23}$$

in which x is $(\mathbf{h}+\mathbf{h}')\cdot\mathbf{r}$. G has a maximal value when $-\mathbf{h}=\mathbf{h}'$ and it drops rapidly when $\mathbf{h}'$ moves away from $-\mathbf{h}$. i.e. Only the terms when $\mathbf{h}'$ is close to $-\mathbf{h}$ contribute to the sum significantly. This property reduces the number of terms in the summation as well as computation time.

Translation

In the rotation step the orientation of the model molecule is rotated around axises through the origin. To give the final solution, the molecule must be translated in the real space to explain the observed diffraction data. In the translation step the translation vector $\mathbf{t}$ is to be determined by considering the cross-Patterson. The translation vector is defined according to the symmetry axis of the crystal, i.e. it is not necessary to calculate the translation function for a crystal belonging to space group $P1$, in which no crystallographic symmetry exists and we can choose any position in a unit cell as the origin. The translation function is defined as equation 2.24.

$$T(\mathbf{t}) = \int_V P_{1,2}(\mathbf{u},\mathbf{t}) \times P(\mathbf{u}) d\mathbf{u} \tag{2.24}$$

in which $P_{1,2}(\mathbf{u},\mathbf{t})$ and $P(\mathbf{u})$ are the computed Patterson map and observed Patterson map respectively. The translation function has a maximal value when the intermolecular vector $\mathbf{t}$ is equal to the true intermolecular vector $\mathbf{t_0}$. It can be shown that the translation function can be expressed as

$$T(\mathbf{t}) = \sum_{\mathbf{h}} |F_{obs}(\mathbf{h})|^2 \cdot \mathbf{F_M}(\mathbf{h}) \cdot \mathbf{F}^*_{\mathbf{M}}(\mathbf{h}\cdot[C]) \exp\left[-2\pi i \mathbf{h}\cdot\mathbf{t}\right] \tag{2.25}$$

with $\mathbf{F}_M(\mathbf{h})$ and $\mathbf{F}_{obs}(\mathbf{h})$ is the structure factor of the model molecule and that of observed data respectively.

2.5 Refinement

After the molecular replacement process an approximate model of crystal structure can be obtained. Although some biological insight could be obtained by this approximate model, the crystal structure at this stage is usually too rough to be trusted. The agreement of the constructed model and observed data is defined by the R-factor.

$$R = \frac{\sum_{hkl} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|} \times 100\% \tag{2.26}$$

where k is a scale factor for the intensity. The R-factor is 59 % for a model with randomly positioned molecules [33], while a typical well refined crystal structure of membrane protein crystal has R-factor of around 20 %.

2.5.1 Least squares method

The goal of least squares method is to minimize the weighted sum of the difference between observed and calculated structural factors:

$$Q = \sum_{hkl} w(hkl)(|F_{obs}(hkl)| - |F_{calc}(hklu)|)^2 \tag{2.27}$$

where $w(hkl)$ is the inverse square of the standard deviation; u are atomic parameters that determines $|F_{calc}|$. This can be done by setting $\partial Q/\partial u_i = 0$, i.e.

$$\sum_{hkl} \frac{(|F_{obs}| - |F_{calc}|)}{\sigma^2} \frac{\partial |F_{calc}|}{\partial u_i} = 0 \tag{2.28}$$

Expand F_{calc} by Taylor series with a small change of one of the atomic parameters $\Delta u_j = u_{j1} - u_{j0}$, we have

$$\sum_{hkl} \frac{(|F_{obs}| - |F_{calc}(u_0)|)}{\sigma^2} \times \left[\frac{\partial |F_{calc}(u)|}{\partial u_j}\right]_{u_0} - \sum_i \Delta u_i \sum_{hkl} \frac{1}{\sigma^2} \left[\frac{\partial |F_{calc}(u)|}{\partial u_i}\right]_{u_0} \times \left[\frac{\partial |F_{calc}(u)|}{\partial u_j}\right]_{u_0} = 0 \quad (2.29)$$

in which we take the first two terms of the Taylor expansion. Now there are n equations (in which n is the number of the parameters to be refined), with each equation containing n variables (u_1 to u_n). The solution (final set of $[u_i]$ such that Q is minimum) can be generated by iteration.

2.5.2 Rigid body refinement

In the least squares method we assume that every atomic parameters are independent and thus we refine each of them individually. In the rigid body refinement, instead of refining the atomic parameters of each atom in the unit cell, constraints are applied to keep some domains in the molecule to be refined as a rigid body. The parameters of the constraints are now to be refined and the entire structure of each domain is considered as a rigid body.

Rigid body refinement is usually done as the first step in the refinement procedures, after the initial structure is generated by molecular replacement method.

2.5.3 Simulated annealing

In the least square method the refinement can be easily trapped into a local minimum if the initial model is too far from the real structure. To avoid getting a pseudo-structure calculated from the local minimum, a refining method that allows searching both uphills and downhills is required. Simulated annealing (or Molecular Dynamics) method meets

this requirement by providing an initial velocity to each atom at the beginning, and simulating the dynamic behavior of the system.

At certain temperature, multiple numbers of structure orientations are allowed. The structure ensemble follows the Boltzmann distribution $N = \exp[-\epsilon_p/kT]$, in which k is the Boltzmann constant and T is the absolute temperature. The potential energy of the structure is approximated by

$$\epsilon_p = E_{bond}+E_{bondangle}+E_{torsion}+E_{dihedral}+E_{vanderWaals}+E_{electro} \quad (2.30)$$

with E_{bond} is the potential energy of a bond, $E_{torsion}$ is the potential energy of bond angle, and $E_{torsion}$ is the torsion energy around a bond. These potential energies are approximated by harmonic oscillators in the form of $E = \frac{1}{2}K(b-b_0)^2$, in which K and b_0 can be determined by vibration spectra of small molecules. The energy of a dihedral angle is given by:

$$E_{dihedral} = K_\theta(1+\cos(m\theta+\delta)) \quad (2.31)$$

in which θ is the rotation angle, δ is a phase angle and m is an integer reflecting the rotation frequency. The van der Waals interaction is given by:

$$E_{vanderWaals} = \frac{A}{r^{12}} + \frac{B}{r^6} \quad (2.32)$$

$E_{electro}$ is the electrostatic potential which can be calculated as:

$$E_{electro} = \sum \frac{q_i \times q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \quad (2.33)$$

In simulated annealing method each atom is assigned with an initial velocity derived from Maxwell-Boltzmann distribution with a preferred temperature. For each atom i, a force $\partial\epsilon_p/\partial\mathbf{r}_i$ is acting on the atom such that the atom will have an acceleration $m_i^{-1}\partial\epsilon_p/\partial\mathbf{r}_i$; so atoms move according to their initial velocities and accelerations for a very short period (e.g. 1 femtosecond) and the new potential energy distribution is calculated by the new position of the atoms. The process is repeated until a minimum ϵ_p is reached.

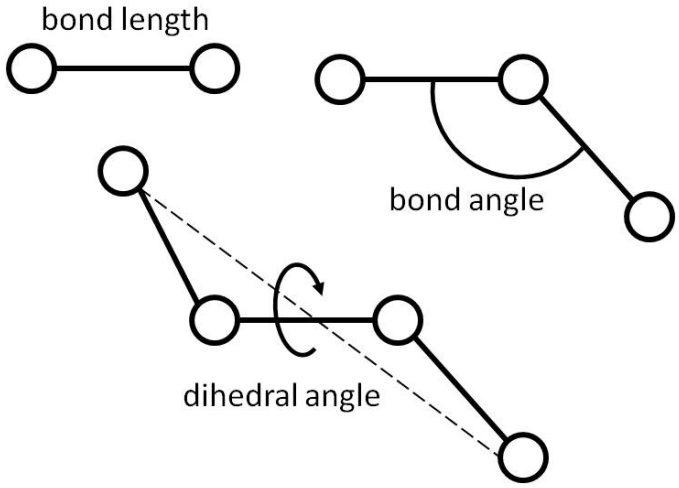

Figure 2.8: The potential energy of a molecular structure depends on the bond lengths, bond angles and dihedral angles between atoms.

Chapter 3

Method and Materials

3.1 Protein crystallization

To conduct a crystallographic study, one must obtain a high quality crystal of the target protein. The crystallization process is briefly separated to several steps: plasmid construction, culturing, protein extraction and purification, as well as crystallization. Plasmid construction produces a strain that produces the target protein at an efficient rate. Then a large amount of cells could be produced through culturing step. After harvesting the cells, the targeted protein was extracted and purified (e.g. by chromolography or washing with detergent), and protein crystals were produced afterwards. In this section, crystallization processes of aR3 as well as cR3 will be introduced.

3.1.1 Plasmid construction

Cruxrhodopsin-3 was expressed in a bR-deficient strain of *Halobacterium salinarum* (MPK 409). Briefly, a *bop* gene fragment in the vector pMPK85 [34] was cloned into the vector pUC18ΔNdeI and, after introducing a *Nde*I site at the start codon of the *bop* gene, the 700 bp *Nde*I-*Not*I fragment was substituted with the synthetic adaptor 5'-CATATGCTCGAGGAGATCTGAGCGGCCGC. The *Bam*HI fragment in the modified vector was cloned into the *Bam*HI site of pMPK85, producing a vector pKI72. Using *Haloarcula vallismortis* genome as a template, the cruxopsin-3 (*cop3*) gene fragment was amplified using

the polymerase chain reaction with the two primers CATATGGCCG-CAACAGTTGGCCCA and CTCGAGTCAGGTCGGGGCAGCCGT-CGG, and cloned into pKI72 at *Nde*I and *Xho*I sites. The vector with the *cop3* gene was transformed into the strain MPK409 [35]. Complete substitution of the *ura3* gene of the strain MPK409 with the *cop3* gene was achieved by positive selection with simvastatine and negative selection with 5-FOA [36].

Archaerhodopsin-3 was extracted from the cell membranes of the wild type *Halorubrum Sodomense*.

3.1.2 Mass Culture

In the mass culture of *Halorubrum Sodomense*, each liter of the culture medium contained 125 g of NaCl, 160 g of $MgCl_2 \bullet 6H_2O$, 0.13 g of $CaCl_2 \bullet 2H_2O$, 5 g of K_2SO_4, 1 g of polypepton (Nihon Pharm. Co.) and 1 g of yeast extract (BD).

In the mass culture of *Halobacterium salinarum* (MPK 409), each liter of the culture medium contained 250 g of NaCl, 2 g of KCl, 20 g of $MgSO_4 \bullet 7H_2O$, 10 g of peptone (Oxoid), 10 g of Na-DL-Malate and 50 mg of Uracil.

The culture medium was adjusted to pH 7 by sodium hydroxide and sterilized by autoclave (121°C, 20 minutes); after the autaclave, 1 ml of micro-nutrient (per 100 ml: $ZnSO_4 \bullet 7H_2O$ 0.1 g, $MnCl_2 \bullet 4H_2O$ 0.03 g, H_3BO_3 0.3 g, $CoCl_2 \bullet 6H_2O$ 0.2 g, $CuCl_2 \bullet 2H_2O$ 0.01 g, $NiCl_2 \bullet 6H_2O$ 0.02 g, $Na_2MoO_4 \bullet H_2O$ 0.03 g) was added to each liter of the culture solution. In the case of *Halorubrum Sodomense*, 2 g of starch dissolved in hot water was also added after the autoclave.

Halobacteria grown in a pre-cultural solution were added to the corresponding mass culture solution (10 – 20 L) and were then incubated at 38 °C with adequate aeration and under constant illumination with fluorescent lamp; 1 ml of antifoam A (Sigma) was added to prevent foam formation. The cultural solution turned milky after incubated for 2 weeks. Cells were harvested 1-2 days after the absorbance at 600 nm reached a saturation level.

For culture of halobacteria on an agar plate, 2 g of agar was dissolved per 100 ml of the culture solution and solidated at room temperature.

3.1.3 Preparation of cR3-rich claret membrane

The protein-rich claret membrane was isolated by the method according to the procedure reported by Oesterhelt and Stoekenius [5]. The cultured cells were harvested by centrifugation at 8,000 rpm for 20 minutes (Rotor: R9A; Centrifuge: Hitachi himac CR20GII), and the pellets were then suspended in 4 M sodium chloride (or 3 M potassium chloride) solution with 1 mg of Deoxyribonuclease I (DNase I, Sigma-Aldrich, D5025) per 10 g of cells collected. After stirring at room temperature for about 30 minutes, the mixture was then frozen in a deep freezer at -80 °C. After that, the sample was thawed and stirred at room temperature for 1 hour. This freeze-thaw process was repeated 3 times. The freeze-thawed sample was then collected by centrifuging at 27,000 rpm for 30 minutes (Rotor: Hitachi Koki P45AT; Centrifuge: Hitachi Koki CP65β) and resuspended in 3 M KCl, and this washing process was repeated 3 times. The washed claret membrane was then further washed with deionized water (38,000 rpm, 30 minutes) and this washing was repeated until the supernatant become clear.

3.1.4 Concentration estimation

The concentration of the target membrane protein (cR3 or aR3) in a purified sample solution was estimated from the absorption spectrum using the following equation:

$$C = OD_{\lambda} \frac{MW}{\epsilon_{\lambda}} \tag{3.1}$$

With OD_{λ} is the absorbance at the absorption maximum of the visible band of retinal, MW is the molecular weight of the opsin (aR3: 28 kDa; cR3: 26 kDa) and ϵ_{λ} is the extinction coefficient of the retinal chromophore. The concentrations of aR3 or cR3 were estimated on the assumption that the extinction coefficients of the retinal chromophores in aR3 or cR3 are the same as that determined for bR (64,000 $M^{-1}cm^{-1}$).

3.1.5 Tween 20 treatment

Partial delipidation of crude claret membranes was performed by Tween20 treatment [37]. The claret membrane isolated from the cell membrane was mixed with a solution containing 0.1 M NaCl, 0.1 M HEPES (pH 8) and 0.3 % Tween 20, and the final concentration of claret membrane was adjusted to 1 mg/ml. The mixture was then incubated for 1 hour at room temperature. After that, the claret membrane was washed with deionized water for 2 times (Rotor: P45AT; 30,000 rpm, 30 minutes.)

3.1.6 Crystallization

A high concentration of claret membrane ($\geq$ 3 mg/ml) was mixed with a solution containing detergent, salt (ammonium sulfate), pH buffer and NaN_3. The salt concentration was then gradually increased by the sitting drop vapor diffusion method (Figure 3.1) using 2.5 – 3.2 M ammonium sulfate as a reservoir solution. The crystallization condition in which high quality crystals were grown is listed in Table 3.1.

	aR3	cR3
Protein mixture		
Nonylglucoside (NG)		5 mg/ml
Octylthioglucoside (OG)	1 mg/ml	
Sodium citrate (pH 4)		0.04 M
HEPES (pH 7)	0.04 M	
Ammonium sulfate	1 M	1 M
Sodium chloride	0.08 M	0.08 M
Sodium azide	0.04 M	0.04 M
Reservoir Solution		
Ammonium sulfate	1.8 - 2.2 M	2.2 - 2.8 M
Sodium citrate (pH 4)		0.1 M
HEPES (pH 7)	0.1 M	

Table 3.1: Crystallization condition yielding high quality crystals. CR3 crystals were grown during incubation at 15 °C for ∼1 month, and aR3 crystals were grown during incubation at 20 °C for 2 months.

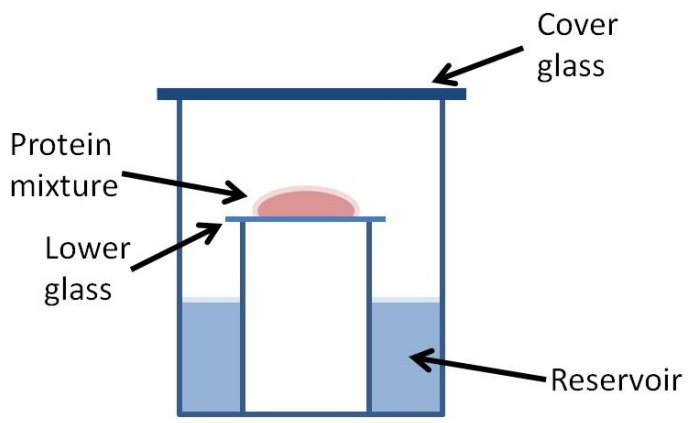

Figure 3.1: Schematic diagram of sitting drop diffusion method. Composition of protein mixture and reservoir solution are shown in Table 3.1. The cover glass was sealed with paraffin oil to prevent leakage of vapor. Concentration of protein and precipitant in the drop increase gradually as the drop and reservoir equilibrate.

3.2 Electron microscopy

The claret membrane structure can be studied by cyo-electron microscope as described by Kouyama et al. [38]. The claret membrane suspended in deionized water was mounted on a carbon-coated grid. The sample was flash-cooled with liquid propane at its melting temperature after removing the excess water. Cyo-electron micrographs were recorded with a CCD camera (Gatan SC200D) installed in a JEM2010 (Jeol) electron microscope.

3.3 Measurement of absorption kinetics by flash photolysis

In order to determine the reaction kinetics of the retinylidene proteins, absorption changes induced by green laser pulses were tracked at various wavelengths.

3.3.1 Experimental setup

The flash-photolysis experiment was set up as shown in Figure 3.2; a sample container was placed at the intercept between the laser beam and the measuring beam. An Nd-YAG laser (Quantel Ultra) producing light pulses at 532 nm and with pulse duration of 3 ns was connected to a function generator by which the timing of laser pulse is controlled. A xenon lamp (Hamaphotonics E7536, Hamamatsu photonic, Japan) was used as the light source for measuring beam and the wavelength of measuring beam was selected by a monochromator (Jobin Yuvon H29V) placed between the light source and the sample corvette. The transmitted light through the sample solution was detected by a photomultiplier tube and its output signal was amplified and sent to a digital oscilloscope (LeCroy Wavesurfer 422). Another monochromator, the wavelength scanning of which synchronizes with that of the other monochromator placed between the xenon lamp and the sample, was placed before the photomultiplier tube to keep the record signal clean from scattering components of actinic light pulses (by the quartz container or sample). Digital signal from the function generator was also fed to the oscilloscope to trigger the recording of light-induced absorption changes.

3.3.2 Data collection and analysis

Flash-induced absorption changes in the wavelength region ranging from 390 nm to 680 nm were measured with 10 nm interval (the measurements at 530 nm and 540 nm were omitted as the measuring light was attenuated by a notch filter at 534 nm that was inserted to block scattered actinic light). At each wavelength, about 1000 signal recordings were performed and their averaged signal was used for analysis.

Absorbance change after stimulation was calculated by subtracting the average absorbance before the triggering from the absorbance at time t. The data set was then reduced by singular value decomposition (SVD), and the apparent rate constants (k_r) were determined by fitting the truncated data with a sum of multiple exponential decay

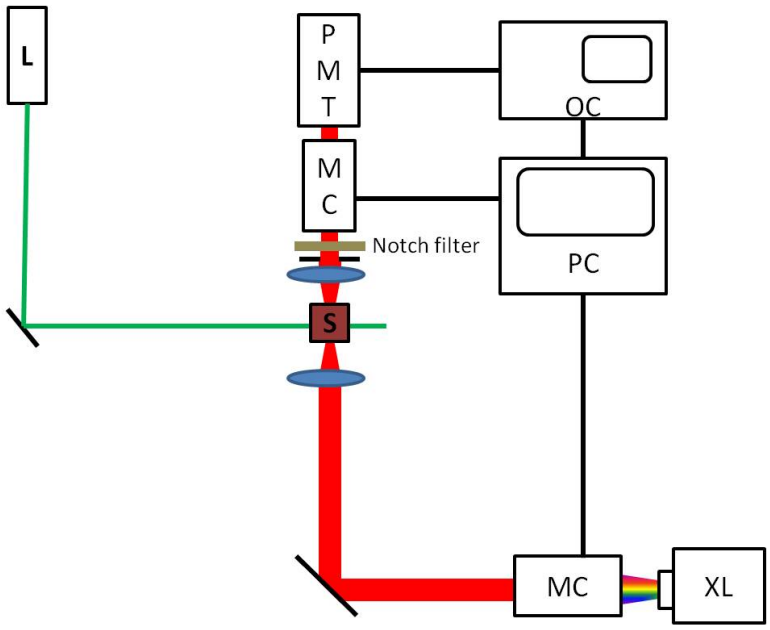

Figure 3.2: Schematic diagram of flash photolysis experiment. L: Pulse laser; PMT: Photomultiplier tube; OC: Oscilloscope; S: Sample; MC monochromator; XL: Xenon lamp.

components.

$$Ae^{\frac{-t}{\tau_1}} + Be^{\frac{-t}{\tau_2}} + Ce^{\frac{-t}{\tau_3}} + ... \tag{3.2}$$

The original data matrix was then reconstructed by minimizing the difference f between the actual data and the sum of weighted exponential with these time constants

$$f = \sum_{i}^{n_\lambda}\sum_{j}^{n_t} \Delta A_{measured}(\lambda_i, t_j) - \sum_{i}^{n_\lambda}\left[\sum_{j}^{n_t} A_r(\lambda_i)e^{-k_r t_j} - A_0(\lambda_i)\right] \tag{3.3}$$

Where $A_r(\lambda_i)$ are the absorption coefficients of the reaction state r at wavelength λ_i. The absorption change due to the reaction transition corresponding to each time constant can be determined by plotting the absorption coefficient against wavelength.

3.4 Blue native page

In this study, we performed modified native polyacrylamide gel electrophoresis (BN-PAGE, [39]) to study the stability of protein-lipid complex. The BN-PAGE consist of a stacking gel and separating gel, the composition of which is shown in Table 3.2. The separating gel was first poured inside a narrow space between two glass plates, and isopropanol was added to flatten the top surface of the separating gel. The gel was allowed to sit for around 30 minutes at room temperature, and then the top of the separating gel was rinsed with deionized water. The stacking gel was then poured on the top of the separating gel, and a comb was inserted to produce wells for sample loading. The gel was then placed at room temperature for around 1 hour before use.

The electrophoresis was run at 4 °C, with applied voltage of 180 V for 100 minutes. The cathode buffer consisted of 15 mM of Bis-tris, 50 mM of tricine and 0.02 % of coomassie blue G250. The anode buffer consisted of 50 mM of Bis-tris, and was adjusted to pH 7 by HCl.

4 % Stacking Gel (600 μl)	
Gel buffer	200 μl
30 % Acrylamide	80 μl
10 % Nonylglucoside	10 μl
10 % Ammonium persulfate (APS)	8 μl
Tetramethylethylenediamine (TEMED)	0.8 μl
Deionized water	300 μl

15 % Separating Gel (3600 μl)	
Gel buffer	1200 μl
30 % Acrylamide	1800 μl
10 % Nonylglucoside	72 μl
10 % Ammonium persulfate (APS)	9 μl
Tetramethylethylenediamine (TEMED)	0.9 μl
Deionized water	520 μl

Table 3.2: Salt composition in the stacking and separating gels. The running buffer consists of 150 mM Bis-tris and 200 mM ϵ-aminocaproic acid.

3.5 Diffraction data collection, scaling and refinement

3.5.1 Diffraction data collection

X-ray diffraction data were collected at beamline BL38B1 in SPring-8 (Figure 3.3). Protein crystals to be studied were kept at 100 K (-173°C) to reduce X-ray damage. The crystal was then exposed to monochromic X-ray beam of wavelength 1.0 Å, at the X-ray flux rate of about 2×10^{12} photons /mm^2/sec. Each diffraction image was recorded with an oscillation range of 1 degree.

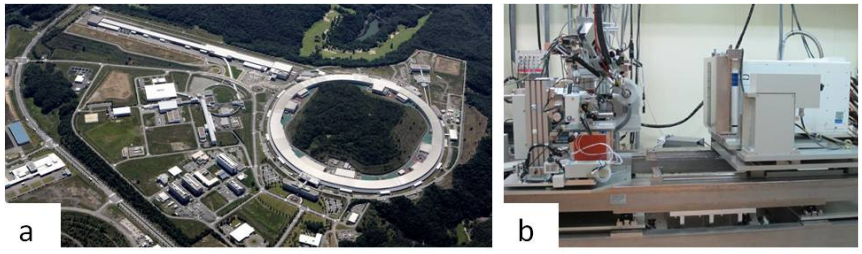

Figure 3.3: a) SPring-8 and b) Beamline BL38B1

3.5.2 Index and scaling

Indexing and integration of diffraction data were carried out by MOSFILM 6.1 [40]. The scaling of data was accomplished using SCALA in the CCP4 program suite [41]. Structure analysis was performed with CNS [42] and XtalView [43].

3.5.3 Refinement

To build a structural model of cR3, we chose the previously determined structure of bR (pdb entry: 1IW6 [44]) as the initial model. The structural model of bR was modified by Swiss-Model [45], by which non-conserved residues between bR and cR3 were automatically replaced, and the whole part of the modified polypeptide was used as an input model for the molecular replacement analysis [46] (The root mean square deviation between the input model and the final model of cR3 is 0.64 Å for 193 C_α atoms). After a rotational and translational searching with MOLREP, retinal, water and lipid molecules were added and the protein conformation was modified manually on the ground of the $2F_o - F_c$ map. Subsequent refinements including simulated annealing and individual B-factor refinement resulted in significant decreases in R_{cryst} and R_{free} values as shown in Table 4.1.

Chapter 4

Result

4.1 Spectroscopic properties of archaeal proton-pumping rhodopsins

After aR3-rich or cR3-rich claret membranes were purified by repeatedly washing the cell membranes with deionized water, their spectroscopic properties were investigated at various pH levels. The absorption spectra of aR3 and cR3 in claret membranes are shown in Figure 4.2 and Figure 4.3, respectively. The absorption spectra of these claret membrane are characterized by three fine peaks at 475, 505 and 541 nm, which are due to the vibrionic bands of the carotenoid bacterioruberin (Figure 4.1) as reported for aR2-rich claret membrane [18]. The visible absorption band of retinal can be recognized as a shoulder in the red region, yet it is difficult to analyze its exact profile due to a high content of bacterioruberin in claret membrane.

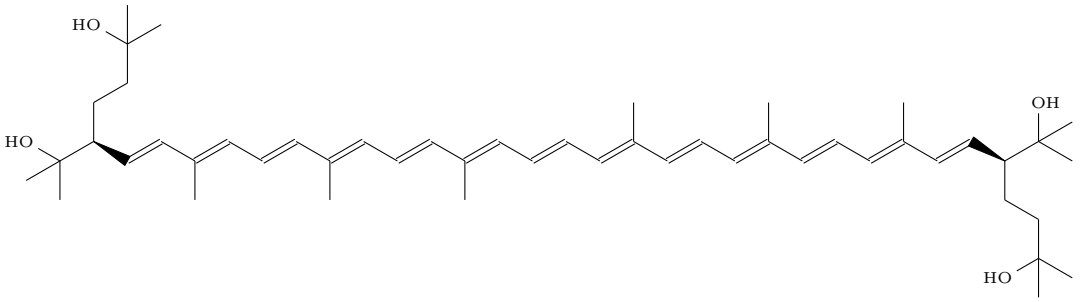

Figure 4.1: Chemical formula of a bacterioruberin.

It has been reported that when bacteriorhodopsin (bR) is placed in the dark for a long period, a thermally stable isomer state of bacteriorhodopsin containing 13-cis / 15-syn retinal is generated [9]. The 13-cis isomer contained bR dark-adapted state exhibits a blue shifted absorption maximum ($\lambda_{max} \sim 550$ nm) [47]. In cR3 and aR3, similar absorption changes were observed during the incubation in the dark (Solid-line of Figure 4.3 and red line of Figure 4.2). The profile of the difference absorption spectra associated with the light-/dark-adaption of cR3 and aR3 are very similar to that observed for bR.

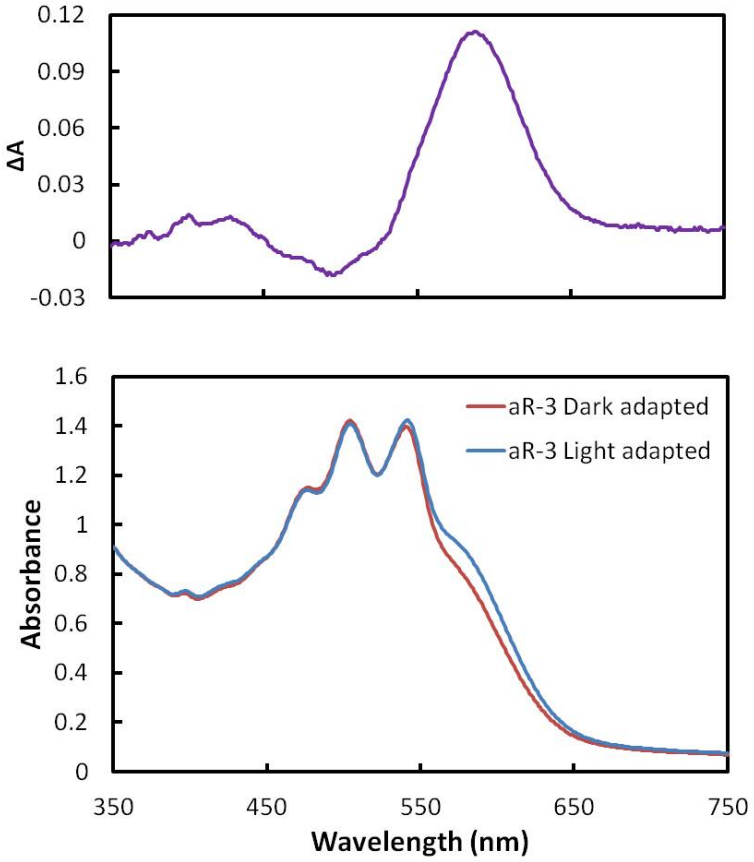

Figure 4.2: Absorption spectra of aR3 in claret membrane suspended in an aqueous solution. Upper panel: Difference spectrum associated with the light-adaption of aR3. Lower panel: Absorption spectra of the light-adapted state (blue line) and dark-adapted state (red line) of aR3.

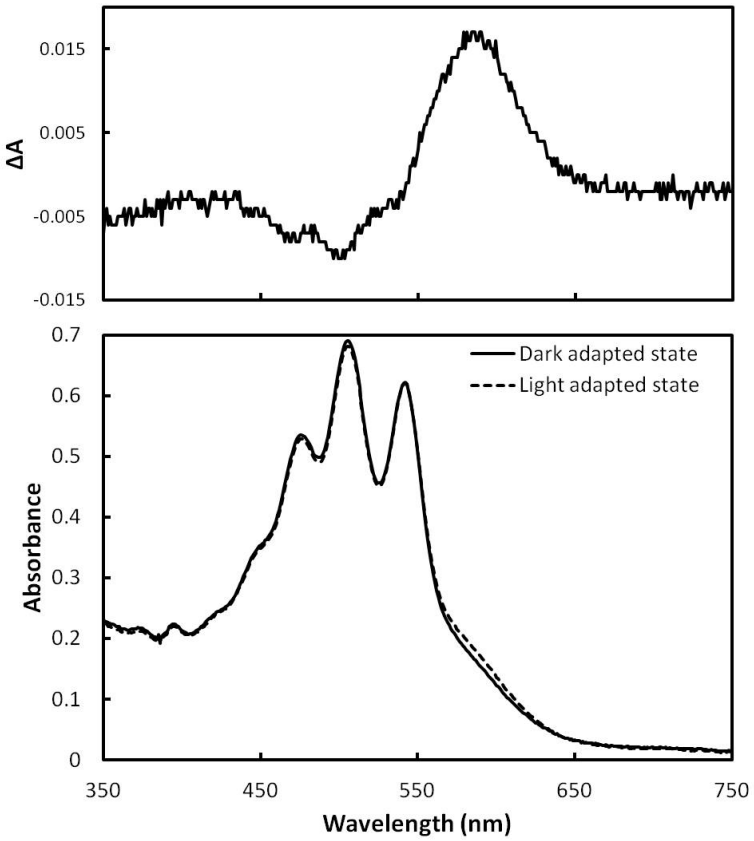

Figure 4.3: Absorption spectra of cR3 in claret membrane suspended in an aqueous solution. Upper panel: Difference spectrum associated with the light-adaption of cR3. Lower panel: Absorption spectra of the light-adapted state (dash line) and dark-adapted state (solid line) of cR3.

4.2 Flash-induced absorption kinetics

4.2.1 Photoreaction of cR3

Figure 4.4 shows the flash-induced absorption changes observed for cR3-rich claret membranes, which were suspended at pH 8.0 and excited with light pulses of 532 nm. In the investigated time region (0.001 – 500 ms), the absorption kinetics changes were fitted with four exponential components (P_1, P_2, P_3 and P_4). The difference absorption spectrum associated with the P_1 component ($\tau_1 = 0.031$ ms) exhibits a positive peak at 590 nm and a negative peak at 410 nm. This component was suggested to correspond to the transition from the K state to the M state. The difference spectrum associated with the P_2 component ($\tau_2 = 0.341$ ms) exhibits a positive peak at 410 nm, which corresponds to the decay of the M state, and a negative peak at 560 nm which is attributable to the rise of N state, i.e. the P_2 component represents the M-to-N transition; At first glance, the difference spectrum associated with the P_3 component ($\tau_3 = 0.854$ ms) looks similar to the P_2 component, with a positive peak of 410 nm and a negative peak of 590 nm. Meanwhile, another negative peak was also seen at 650 nm, this peak can be explained by supposing that the lifetime of N becomes much longer than that of O state, and the negative peak reflects the influence of the decay process of O state. The difference spectrum associated with the P_4 component ($\tau_4 = 10.5$ ms) exhibits a negative peak of 580 nm, which is the same as the difference spectrum between the long N state and the resting state of bR [48].

The absorption kinetics for cR3 at different pH levels (6.0 – 9.0) are shown in Figure 4.5. The absorption kinetics at pH 9 is similar to that at pH 8. Noticeable differences can be observed in the early stages of the photo-cycle at or below pH 7. At pH 7, the absorption kinetics changes were fitted with four exponential components (P_1, P_2, P_3 and P_4). The difference absorption spectrum associated with the P_1 component ($\tau_1 = 0.018$ ms) exhibits a positive peak at 590 nm and two negative peaks at 410 nm and 490 nm. This component was suggested to correspond to the transition from the K state to a mixture state of the L state and the M state (the L/M state). The difference

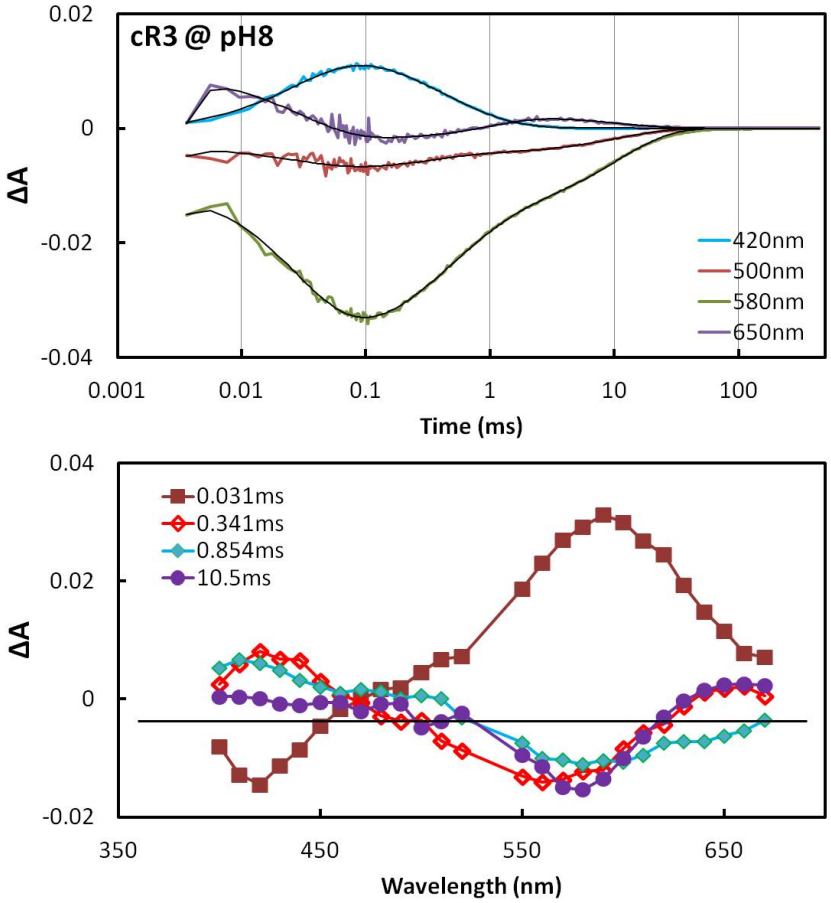

Figure 4.4: Flash-induced absorption changes of cR3 in claret membrane at pH 8. Lower panel: the amplitudes of absorption changes associated with different reaction transitions are plotted against the wavelength of measuring beam. Upper panel: Flash-induced absorption changes observed at different wavelengths.

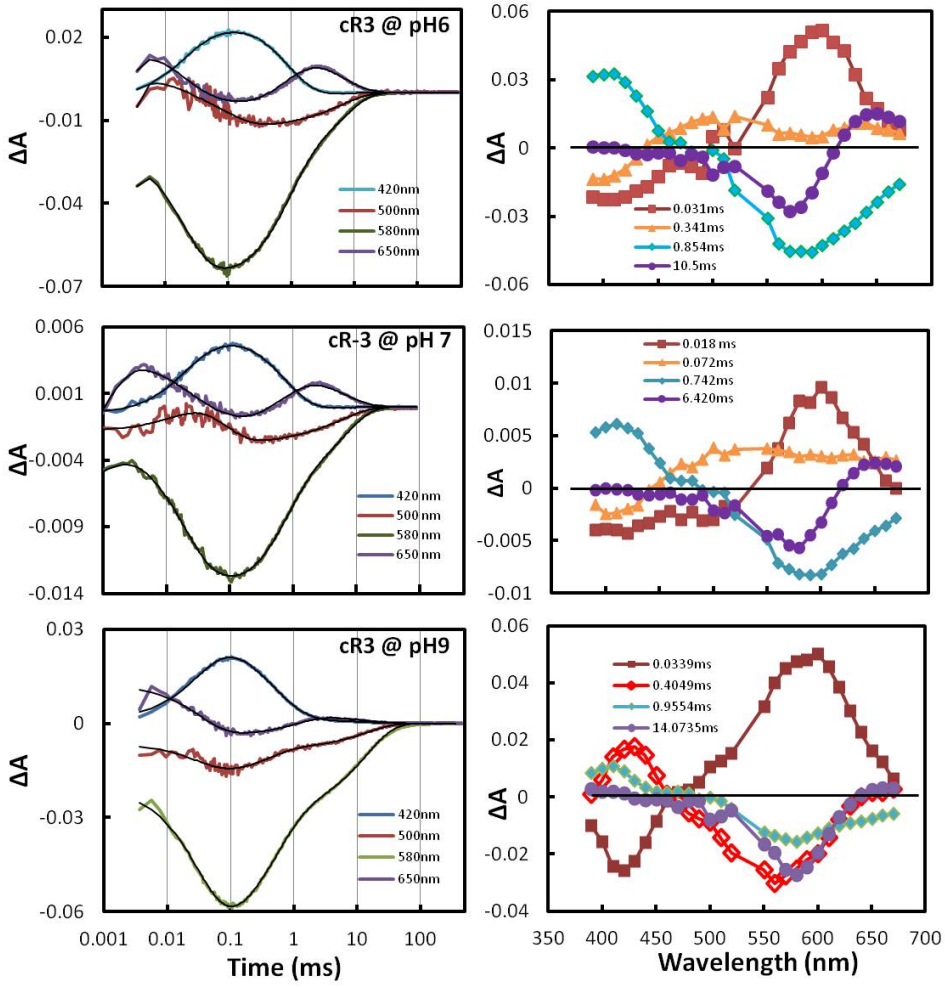

Figure 4.5: Flash-induced absorption changes of cR3-rich claret membranes suspended at pH 6 (top panels), pH 7 (middle panels) and pH 9 (bottom panels). In right panels: the amplitudes of absorption changes associated with different reaction transitions are plotted against the wavelength of the measuring light.

spectrum associated with the P_2 component ($\tau_2 = 0.072$ ms) exhibits a negative peak at 410 nm, which corresponds to the rise of the M state, and a broad positive peak around 500 nm which is attributable to the decay of L/M state (or possibly mixture of K and L/M state). i.e. the L intermediate which existed in the form of a mixture of L and M states (L/M state) became detectable when the medium pH was decreased below pH 7. A possible explanation is that the deprotonation of proton-releasing group was prohibited at pH levels lower than pH 7, so that conformation change of the proton releasing channel, which affects the pKa of Asp85, was delayed and therefore the equilibrium between L and M was extended. As a result, the L/M-to-M transition was delayed and became detectable.

The photo-cycle of cR3 can be described by K $\rightarrow$ (L/M) $\rightarrow$ M $\rightarrow$ N $\rightarrow$ O $\rightarrow$ cR3, with the L/M state become undetectable at high pH levels, where its decay rate is possibly faster than the decay rate of K state.

4.2.2 Photoreaction of aR3

Figure 4.6 shows the light-induced absorption changes observed for aR3-rich claret membranes suspended at pH 7. The absorption kinetics changes in the time investigated (0.001 – 500 ms) were fitted by five exponential components (P_1 to P_5). The difference spectrum associated with the P_1 component ($\tau_1 = 3.9$ μs) has a positive peak at 600 nm and a negative peak at 510 nm. This spectrum is similar in profile to the difference spectrum associated with the K-to-L transition of bR, indicating that in aR3 the K state decays to the L state with a time constant of 3.9 μs (i.e. at a similar rate observed for bR). The difference spectrum associated with the P_2 component ($\tau_2 = 0.054$ ms) exhibits a broad positive peak around 570 nm (a positive peak at 570 nm and a shoulder at 510 nm) and a negative peak at 410 nm, corresponding to the transition of the L state (possibly mixture of K, L and L/M state) to the M state. The third component P_3 ($\tau_3 = 0.43$ ms) with a positive peak at 410 nm, and a negative peak at around 520 nm is attributed to the M-to-N transition. The difference spectrum associated with the forth component P_4 ($\tau_4 = 3.63$ ms) consists of two positive peaks at

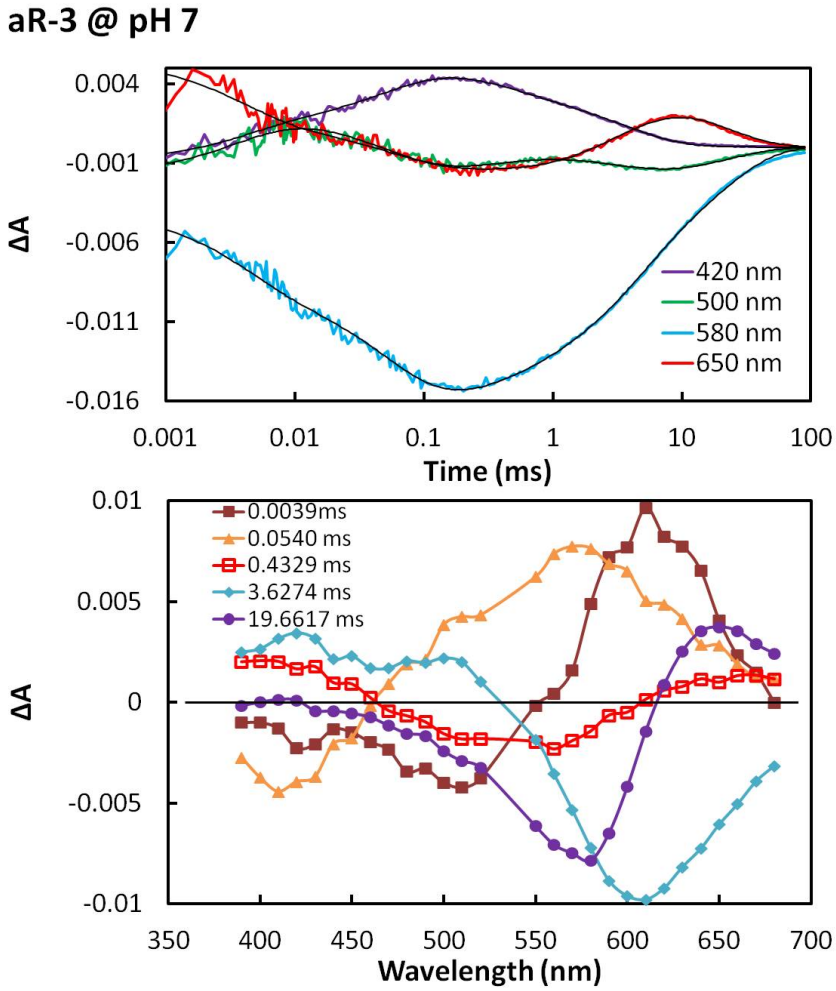

Figure 4.6: Flash-induced absorption changes of aR3 in claret membrane at pH 7. Lower panel: the amplitudes of absorption changes associated with different reaction transitions are plotted against the wavelength of measuring beam. Upper panel: Flash-induced absorption changes observed at different wavelengths.

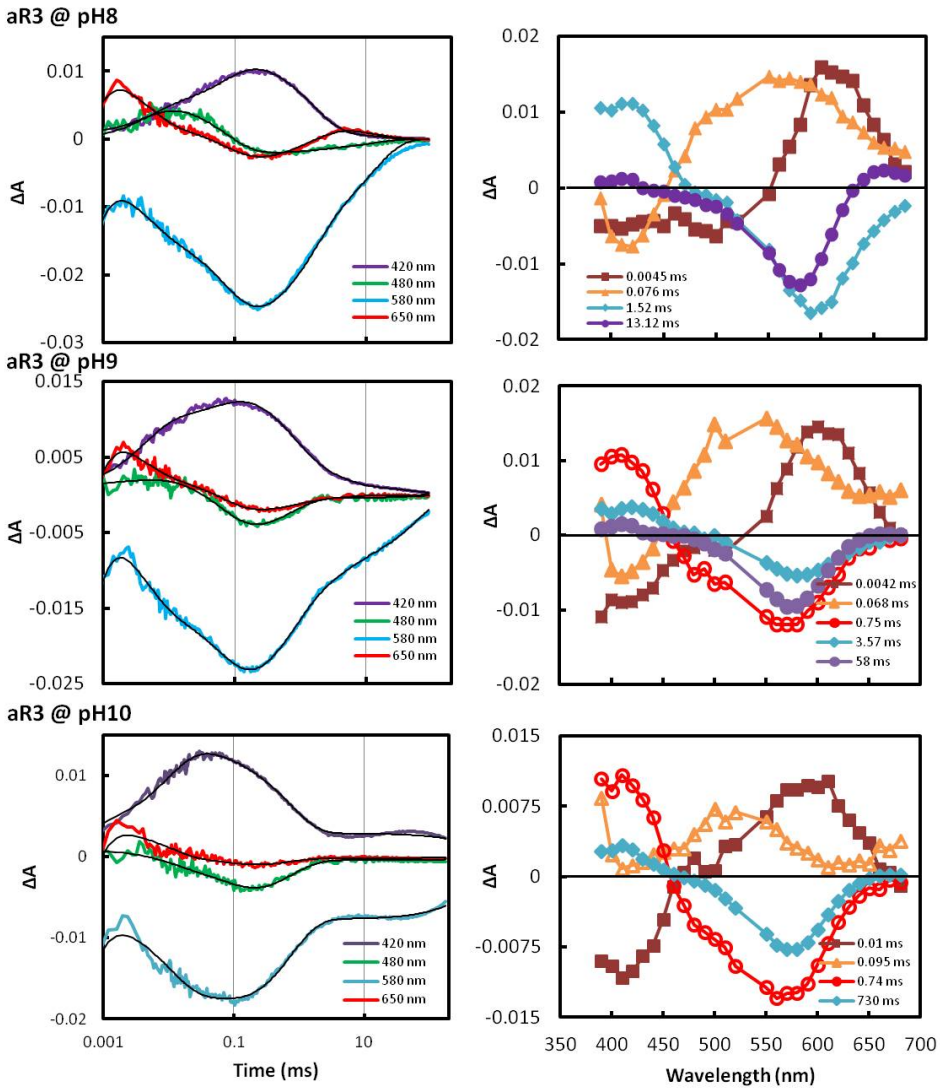

Figure 4.7: Flash-induced absorption changes of aR3-rich claret membranes suspended at pH 8 (top panels), pH 9 (middle panels) and pH 10 (bottom panels). In right panels: the amplitudes of absorption changes associated with different reaction transitions are plotted against the wavelength of the measuring light.

410 nm and 500 nm and negative peak at 600 nm, suggesting that this component corresponds to the decay of a mixture of M and N (or L) to the O state. The N state is not clearly detected because its decay rate is comparable to that of the M state. The difference spectrum associated with the last component P_5 (τ_5 = 19.67 ms) exhibits a positive peak at 640 nm and a negative peak at 580 nm, indicating that the P_5 component corresponds to the last step of the photo-cycle, i.e. the transition from the O state to the resting state. The photo-cycle of aR3 can be described by the following scheme: K $\rightarrow$ L $\rightarrow$ L/M $\rightarrow$ M $\rightarrow$ N $\rightarrow$ O $\rightarrow$ aR3.

The absorption kinetics of aR3 at alkaline pH levels (8.0 – 10.0) are shown in Figure 4.7. At pH 8, the K-to-L transition and the L-to-M transition are represented by the P_1 component and the P_2 component, respectively. The difference spectrum associated with the P_3 component (τ_3 = 1.52 ms) has a positive peak at 410 nm and a negative peak at 590 nm; this component is attributed to the transition from M state to N state (possibly mixture of the N and O state). The difference spectrum associated with P_4 component has a large negative peak at 590 nm and a small positive peak at 640 nm; this component corresponds to the transition from a mixture of N and O to the resting state. The photo-cycle of aR3 at pH 8 is the same as that at pH7.

At pH 9 and pH 10, however, the difference spectra associated with the P_1 component, which has a positive peak at 590 nm and a negative at 410 nm, showed that the K state of aR3 directly decay to the M state (or the decay rate of L state is too fast to be detected). Meanwhile, the difference spectra associated with P_2 components, which consist of a positive peak at 520 nm and a negative peak at 410 nm. This component can be explained by assuming that at pH 9 and pH 10, the aR3 claret membranes contain a mixture of neutral purple form and alkaline pink form of aR3, in which they undergo independent photo-cycles. To examine the assumption, absorption spectra of cR3 at alkaline pH were measured. Figure 4.8 shows the absorption spectra of aR3 at pH 7.31, pH 8.45 and pH 9.78, the difference spectra show a negative peaks at 600 nm and the positive peaks at 530 nm, indicating the transition from neutral purple form to alkaline pink form occurred

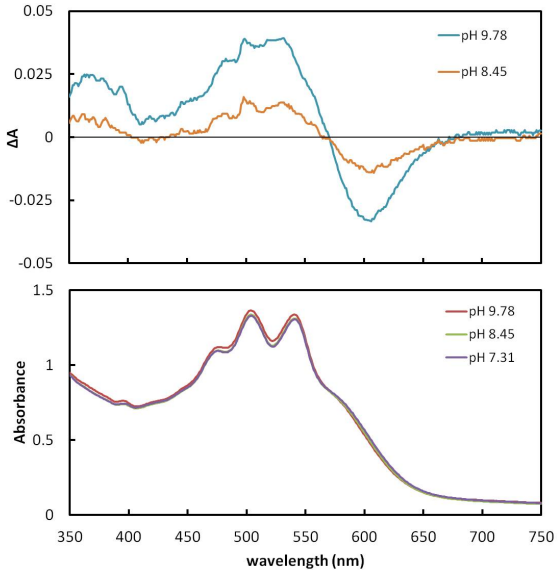

Figure 4.8: Upper penal: Difference spectra obtained by subtracting the absorption spectra of pH 8.45 and pH 9.78 by that of pH 7.31. Lower panel: Absorption spectra of aR3 at pH 7.31, pH 8.45 and pH 9.78.

above pH 8.45.

4.3 Electron microscopy

The morphology of cR3-rich claret membrane was investigated by cryo-electron microscopy. After the tween 20 treatment, vesicles of diameter from 100 nm to 200 nm were observed.

4.4 Crystal structure

4.4.1 3 D crystals of aR3 and cR3

Three dimensional crystals of aR3 and cR3 were prepared by the membrane fusion method as described in section 3.1.6. As shown in Figure 4.10, trigonal crystal of cR3 grew within a protein-rich phase in the crystallization solution. On the other hand, aR3 crystallized into a rectangular crystal. X-ray diffraction experiment showed that cR3 crystal can diffract up to 2.1 Å resolution. Unfortunately, the aR3 crystals

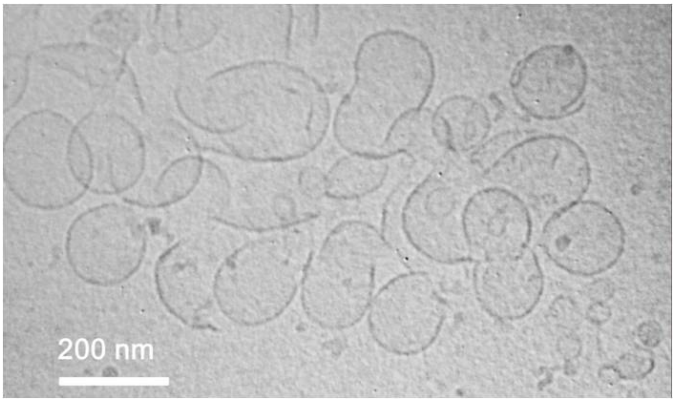

Figure 4.9: Cyro-electron micrograph of cR3-rich claret membranes.

obtained were too tiny to perform diffraction experiment before this thesis was finished.

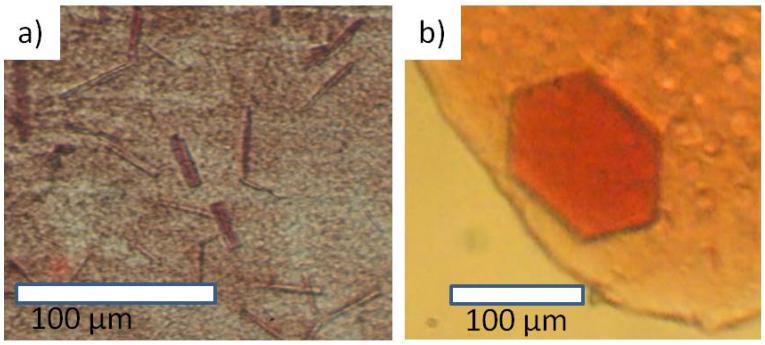

Figure 4.10: Protein crystals of: a) aR3 ,and b) cR3.

4.4.2 Spectroscopic property of the trigonal crystal of cR3

The absorption spectra of cR3 *P321* crystal under polarized measuring beam at different angles are shown in Figure 4.11. The spectrum with the measuring beam being polarized in paralleled to the c-axis is marked in red, and that perpendicular to the c-axis (along the normal direction of the trigonal surface in Figure 4.10 b) is marked in blue. The characteristic peaks of bacterioruberin showed only in the absorption spectrum by the polar angle parallel to the c-axis, indicating that

the bacterioruberin is oriented in parallel to the c-axis. Meanwhile, only a single peak at 560 nm was observed when the polarization of measuring beam was perpendicular to the c-axis of cR3 crystal. The dipole moment of retinal is expected to be tilted largely from the c-axis of the crystal, and the absorption peak of cR3 is estimated to be 560 nm.

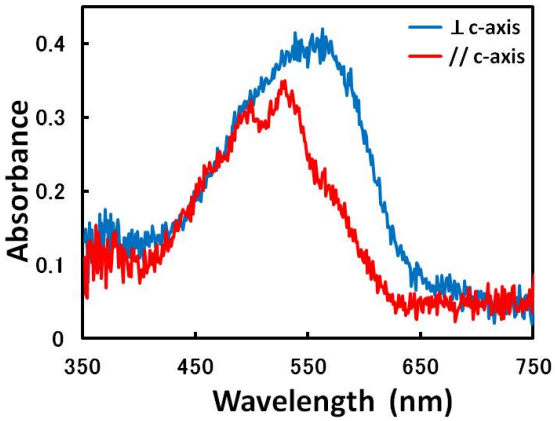

Figure 4.11: Spectroscopic property of the trigonal crystal of cR3, recorded with the measuring beam being polarized in paralleled (red) and perpendicular (blue) to the c-axis of cR3 *P321* crystal.

4.4.3 Crystal packing

Diffraction data shows that the trigonal crystal of cR3 belongs to space group *P321*. The top and side views of the crystal packing are shown in Figure 4.12. Stacking of membranous layers along the c-axis is observed. In each layer cR3 trimers are arranged on a honeycomb lattice such that the nearest neighboring trimers are orientated in opposite direction. The cR3 trimers are piled up along the c-axis. Previously, the trimeric crystal structures of aR2 in the crystal belonging to space group *P321* and $P6_322$ were reported by Yoshimura and Kouyama (2007) [26]. Both *P321* crystals show similar crystal packing, but the cell dimension of cR3 crystal is considerably larger than that of aR2 crystal ($a = b = 106.2$ Å, $c = 55.2$ Å in cR3 crystal vs. $a = b = 98.2$ Å, $c = 56.2$ Å in aR2 crystal). In their report, a serious problem in

the crystal twinning was observed. The problem was not serious in the *P321* crystal of cR3, which was grown at an acidic pH (i.e. at pH 4).

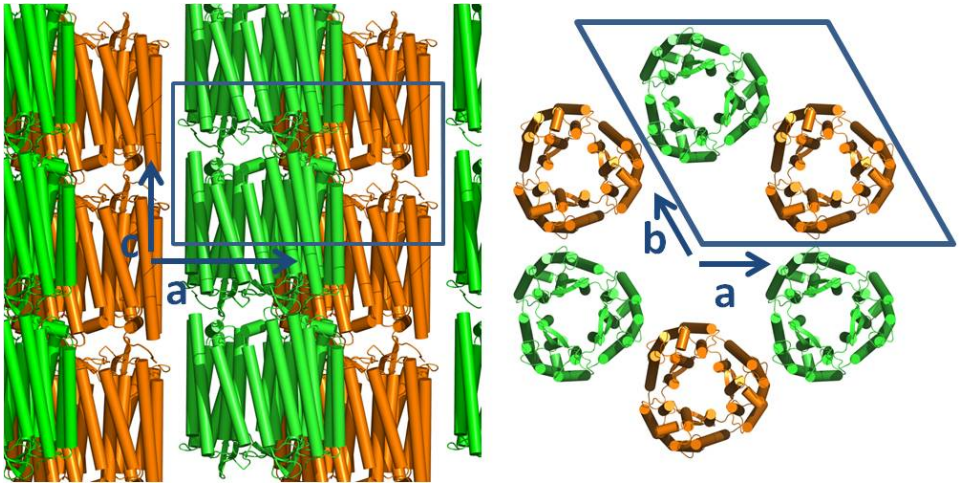

Figure 4.12: Protein packing in the *P321* crystal of cR3.

Diffraction data at the highest resolution were obtained when the *P321* crystal of cR3 was soaked at pH 5. Structural analyses show that the crystal structure remain unchanged in acidic pH region (i.e. between pH 4 and pH 5). It is interesting to note that the cell dimension along c-axis increased significantly (about 5 Å, i.e. 10 %) when the *P321* crystal of cR3 was soaked in a post-crystallization solution at or above pH 6 (Table 4.1). The expansion of the cell dimension along the c-axis seems to be caused by deprotonation of several residues at the contact region between different membranous layers. Crystal structures at low pH values show that along the c-axis, Glu5 and Glu72 at the extracellular surface of one subunit of cR3 are interacting with Glu239 and Asp166, respectively, on the cytoplasmic surface of another subunit in a neighboring membranous layer. Increase in the medium pH will cause deprotonation of some of these residues so that the electrostatic repulsive force among them increases, eventually pushing each other away along the c-axis. The expansion of the cell dimension along the c-axis was also accompanied by disorder of BC loop and the C-terminal polypeptide region, yet the orientation of cR3 trimers in the membranous layer was not affected (Figure 4.13).

Data collection				
Soaking solution pH	pH 4	pH5	pH 6	pH 7
Resolution ($\AA$)	47.6 - 2.3 (2.42 - 2.3)	47.5 - 2.1 (2.21 - 2.1)	45.9 - 2.3 (2.42 - 2.3)	50.0 - 2.8 (2.95 - 2.8)
Space group	$P321$	$P321$	$P321$	$P321$
Unit cell a, b ($\AA$)	106.12	106.01	106.10	106.21
Unit cell c ($\AA$)	55.62	55.44	60.12	60.01
No. unique reflections	16232 (2312)	19657 (2982)	15673 (2299)	9163 (1420)
Multiplicity	7.5 (7.4)	7.6 (7.4)	6.4 (6.4)	5.6 (6.3)
Completeness (%)	99.6 (98.9)	92.9 (96.9)	90.0 (92.4)	92.5 (100)
R^1_{sym} (%)	9.2 (48.8)	7.3 (49.4)	8.9 (52.0)	7.9 (49.6)
$I/\sigma(I)$	17.5 (4.8)	20.8 (4.6)	18.3 (3.9)	21.4 (4.4)
mosaicity	0.36	0.40	0.90	0.75
Refinement				
Resolution limit ($\AA$)	15.0 - 2.30	15.0 - 2.1	15.0 - 2.30	15.0 - 2.8
Protein residues	4 - 237	1 - 244	4 - 237	4 - 237
Number of lipids	1	1	1	1
Number of water	14	26	14	14
R^2_{crysr} (%)	25.0	23.1	24.7	25.7
R_{free} (%)	27.0	25.4	26.8	27.2
Rmsd				
Bond Length ($\AA$)	0.0066	0.0056	0.0060	0.077
Bond angle (°)	1.17	1.15	1.07	1.24
B factor ($\AA$)				
Protein	39.2	29.6	32.7	50.2
Water	42.7	34.0	31.8	47.7
lipids	67.4	74.5	37.9	64.8

Table 4.1: Data collection and final refinement statistics of cR3. $^1R_{sym} = \sum_{hkl} |I_i - < I >| / \sum_{hkl} \sum_i I_i$ where I_i is the intensity of an individual reflection and $< I >$ is the mean intensity obtained from multiple observation of symmetry related reflections. $^2R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{clac}|| / \sum_{hkl} |F_{obs}|$ 5 % randomly selected reflection were used for R_{free}.

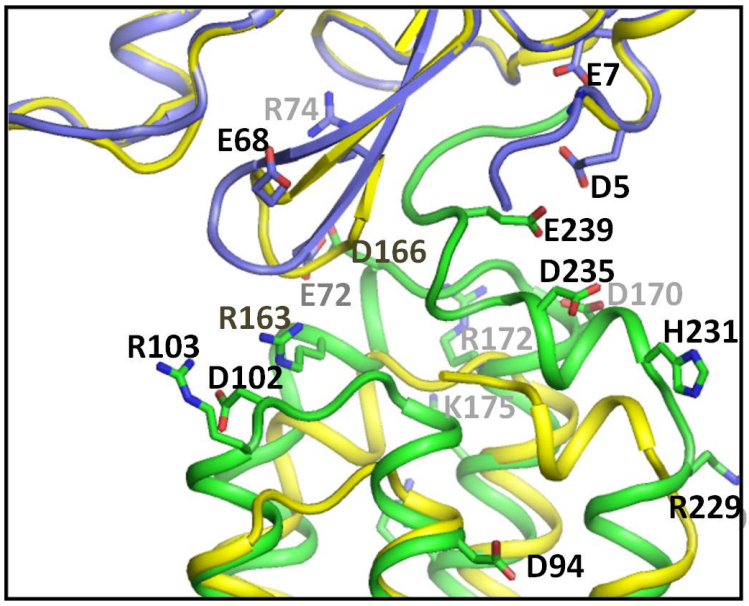

Figure 4.13: Contact region between a cR3 subunit and its nearest neighboring subunit along c-axis direction.

4.4.4 Structural properties

The structural model of cR3 is shown in Figure 4.14. Similar to other arhcaeal light-driven proton pumps, the polypeptide of cruxrhodopsin-3 in folded into seven trans-membrane helices (named from helix A to helix G), a β-sheet at the BC loop, and a short helix at the C-terminal region (Figure 4.15a). Retinal is bound to the ϵ amino group of Lys220, and it takes on the *all-trans* configuration.

At low pH (pH 5 and 4), the β-sheet at BC loop of cR3 is oriented towards its DE loop, and the main chain at Ala70 in the BC loop is hydrogen-bonded with the main chain at Ala127 in the DE loop (Figure 4.16a). The short helix at C-terminus, which covers the central area of the cytoplasmic surface, is fixed by interactions with residues (Arg36, Arg38, Tyr41, Asp166, and Thr171) from helices B, E and F. Upon the crystal soaking above pH 6, the BC loop and the C-terminal polypeptide underwent large structural alterations and became disordered (Figure 4.13), whereas the structure of the protein inside scarcely changed.

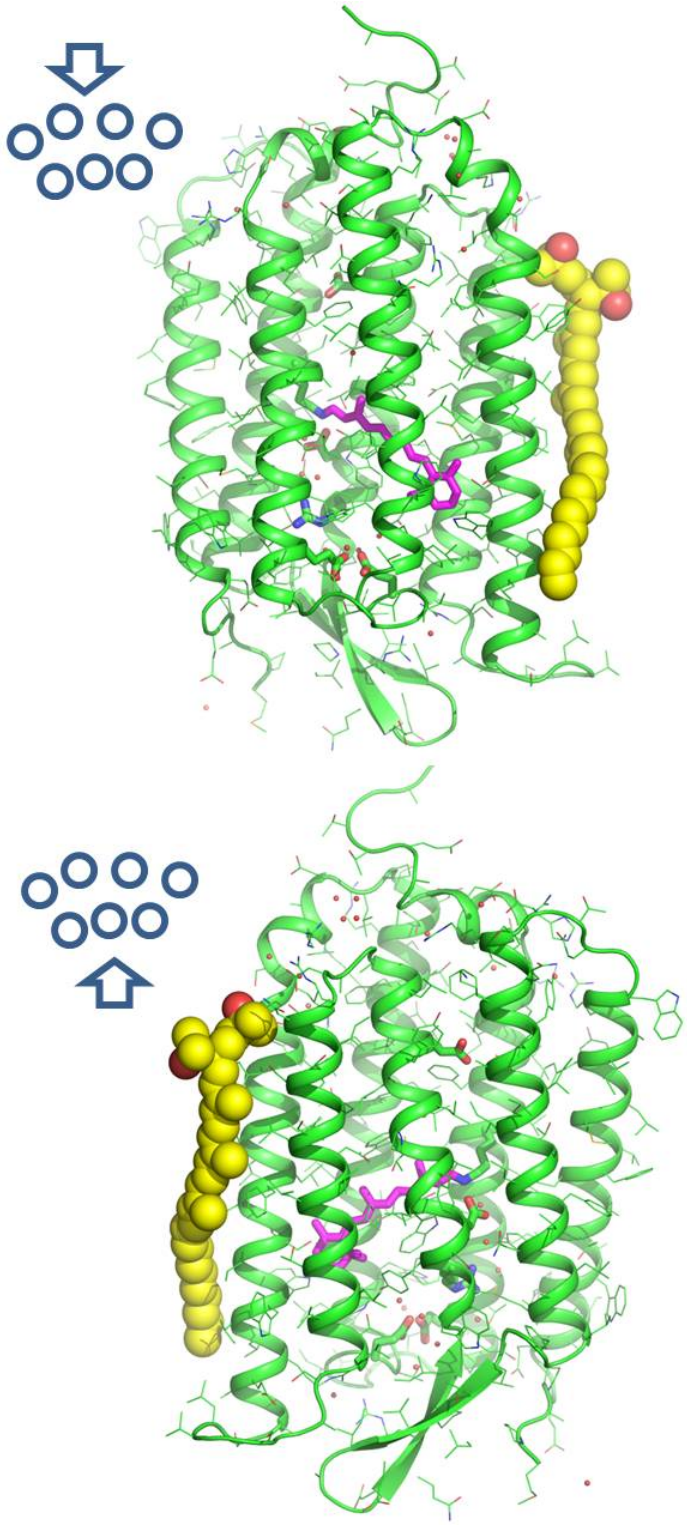

Figure 4.14: Side view of the cruxrhodopsin-3 monomer.

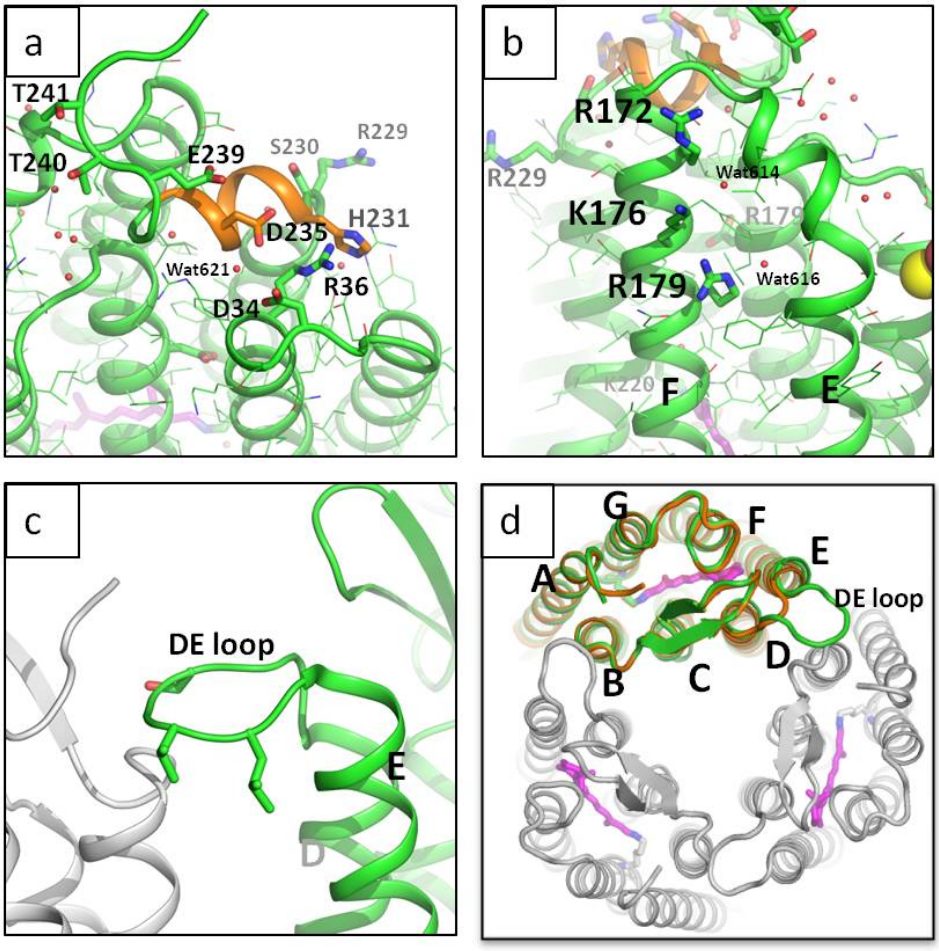

Figure 4.15: Specific structural properties of cR3. a) a short α-helix at the C-terminus (orange) covers the central region of the cytoplasmic surface; b) the cytoplasmic region of helix E is largely bent to create an opening; c) and d) the side view and top view (from the extracellular side) of the long extend DE loop.

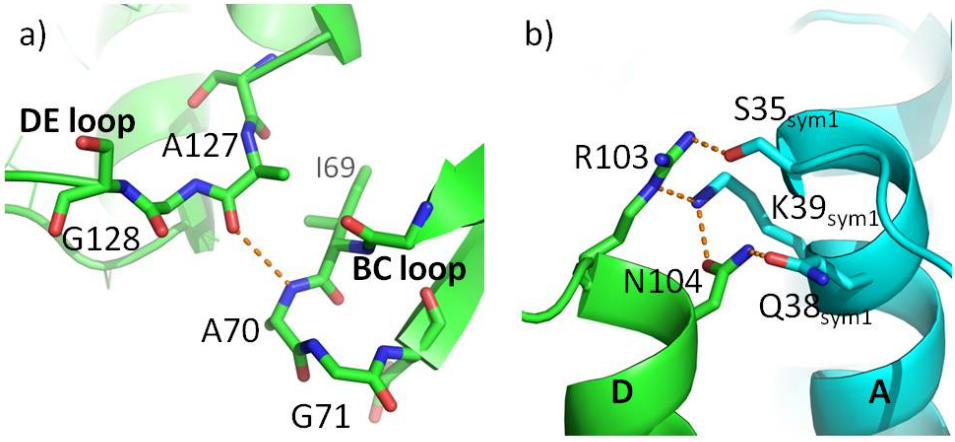

Figure 4.16: a) The BC loop and DE loop of cR3 at pH 5. Hydrogen bond was formed between the main chains of Ala127 and Ala70; b) The inter-subunit interaction between helix D (green) of one subunit and helix A (cyan) of a neighboring subunit.

The structures of the retinal-binding site and the proton release pathway in cR3 are shown in Figure 4.17. The retinal chromophore is linked to a lysine residue (Lys220) at the middle of the helix G, the remarkable similarity of the retinal-binding pocket for aR2 and cR3 explains the spectroscopic similarity (both absorption peak at around 580 nm) of aR2 and cR3. Note that at the proton-release channel, two glutamates (Glu198^{cR3} and Glu208^{cR3}, which correspond to E194 and E204 of bR, respectively) formed a paired structure. The close distance between these residues (2.4 Å) can be explained by a low-barrier hydrogen bond [49]. This paired structure in cR3 remained unchanged within the investigated pH range (pH 4 – pH 7).

The indole ring of Trp186^{cR3} (and Trp187^{aR2} of aR2) near the retinal chromophore is in contact with a residue on helix E (Leu149^{cR3} and Phe150^{aR2}), so that the motional freedom of Trp186^{cR3} (Trp187^{aR2}) is reduced (Figure 4.18).

The cytoplasmic side of helix E is largely bent in such a manner that a cavity was created between helices E and F near the cytoplasmic surface (Figure 4.15b). The DE loop of cR3 is long enough to reach

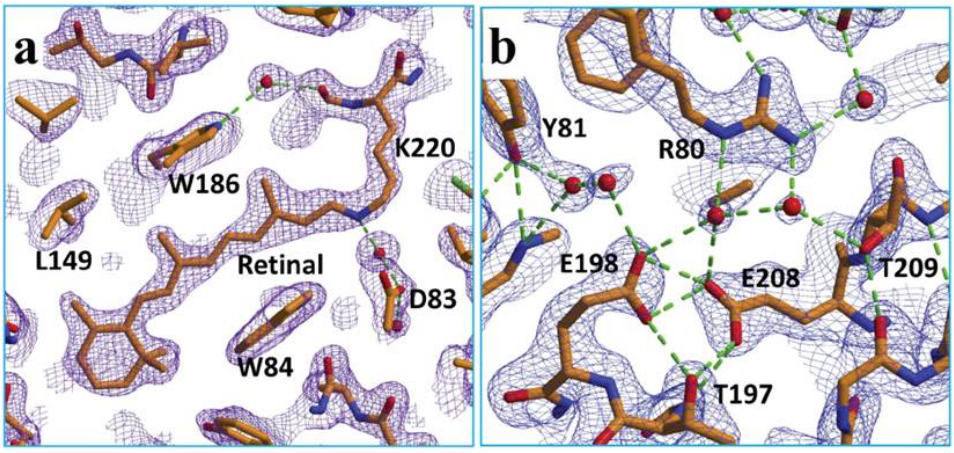

Figure 4.17: The internal structure of cR3. a) The retinal-binding pocket; b) the proton-releasing group.

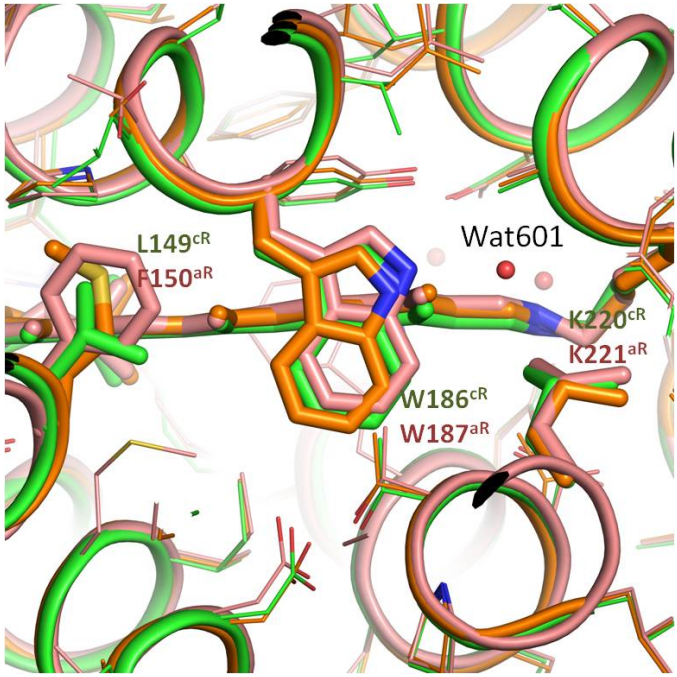

Figure 4.18: The retinal-binding pocket of aR2 (pink), bR (orange) and cR3 (green).

a neighboring subunit and interact with the N-terminal region (main chain at Pro2) and helix B (Leu60 and Gly61)(Figure 4.15c and d). As a result, the trimeric structure of cR3 is strengthened. On the other hand, inter-subunit interactions are observed on the cytoplasmic side of helix D (Asn104 and Arg105) and helix A (Ser35, Gln38 and Lys39) of the neighboring subunit (Figure 4.16b). Three positively-charged residues (Arg172, Lys176 and Arg179) are located at the cytoplasmic end of helix F. It is likely that this charge distribution affects the higher order structure of cR3 in the cell membrane (Figure 4.15b).

The trimeric structure of cR3 is further strengthened by the carotenoid bacterioruberin, which binds to the crevice between neighboring subunits (Figure 4.19). One terminal end of the bacterioruberin is surrounded by helices A and B of one subunit and helices E and F of the neighboring subunit, while the other end is extruded to the lipid bilayer. Similar structural feature has been observed in the trimeric aR2-bacterioruberin complex.

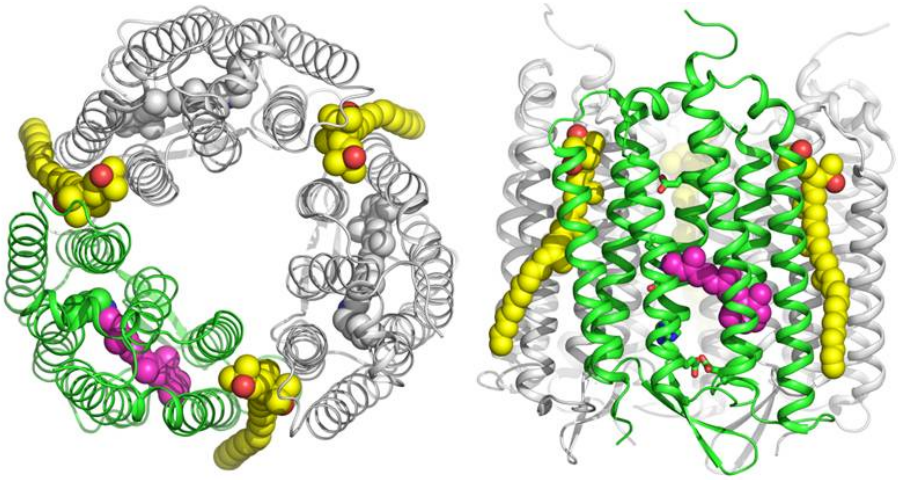

Figure 4.19: The trimeric structure of cR3.

4.5 Higher-order structure of cR3

The peculiar structure of cR3 is seen in that an extended DE loop interacts with a neighboring subunit. This observation leds to an hy-

pothesis that the trimeric assembly of cR3 is more stable than that of bR. To investigate this possibility, we investigated the photo-stability of cR3 at various concentrations of detergent.

4.5.1 Photobreaching

In this experiment, an aqueous suspension of cR3-rich claret membranes was exposed to strong orange light (wavelength $\geq$ 580 nm) from a 150 W xenon lamp, and the absorption changes at 560 nm under continuous illumination were recorded (Figure 4.20).

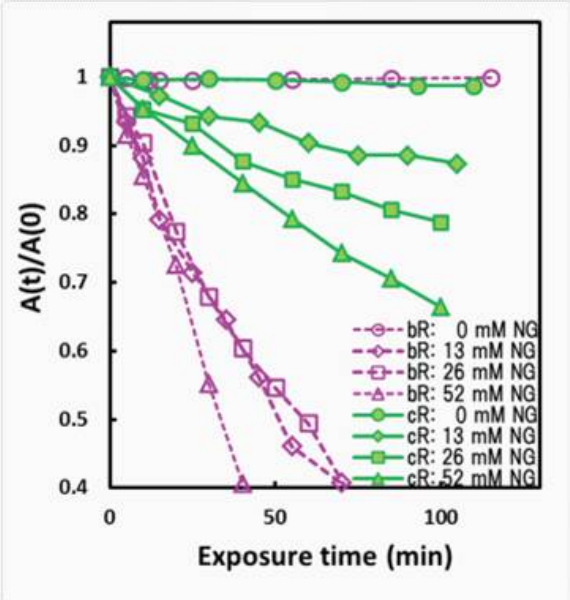

Figure 4.20: Photobleaching experiment of cR3.

No obvious absorption change was induced under illumination when cR3 was embedded in claret membrane; i.e. the retinal chromophore is protected from photohleaching. However, when a large amount of detergent (nonylglucoside) was added to the claret membrane suspension, a noticeable photobleaching of the retinal chromophore in cR3 was observed. The rate of photobleaching increased with the concentration of detergent in the suspension. The photobleaching data indicate that the photostability of cR3 is reduced when the trimeric structure and/or the native lipid-protein / protein-protein interactions are destroyed by

an excess amount of detergent. The photobleaching of bR was also shown in Figure 4.20. Note that the photobleaching of bR became significant at a much lower detergent concentration than observed for cR3 (13 mM); for example, the photobleaching of rate cR3 in 13 mM nonylglucoside is five times slower than that observed for bR at the same detergent concentration. This result suggests that even after addition of detergent, a considerable fraction of cR3 remains to exist in the trimeric state or the lipid-protein interaction is still maintained.

4.5.2 Blue native PAGE

The stability of the trimeric assembly of cR3 at various detergent concentrations was investigated using BN-PAGE. The result is shown in Figure 4.21. The result shows that cR3 remains to keep the trimeric structure in 13 mM nonylglucoside. When the detergent concentration was increased above 0.8 %, a large fraction of cR3 were dissociated into the monomeric state.

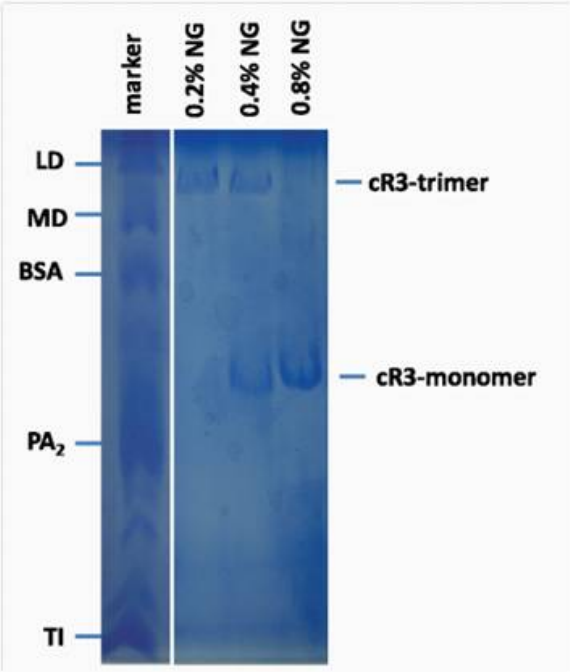

Figure 4.21: Blue native PAGE of cruxrhodopsin-3

Chapter 5

Discussion

In the present study, cruxrhodopsin-3 was crystallized into trigonal crystals belonging to space group *P321*, and the crystal structures at various pH levels were determined. Structural analysis shows that cR3 monomer consists of seven transmembrane α-helices (A to G) and a β sheet at the BC loop. The retinal chromophore is covalently linked to a lysine ($K220^{cR}$) at the middle of helix G. As compared with other archaeal proton-pumping rhodopsins (bR, aR1, aR2, dR3), cR3 has the following peculiar structural properties: i) a short helix at the C-terminal, covering the central region of the cytoplasmic surface; ii) the cytoplasmic side of helix E is largely bent to create a cavity between helix E and F near the cytoplasmic surface; iii) a long DE loop extends towards a neighboring subunit, strengthening the trimeric structure; iv) the side chain of Leu141 in helix E is in contact with the indole ring of Trp180, restricting the motional freedom of Trp180 and thereby making the retinal-binding pocket more rigid than that observed for bR. To discuss how these structure properties affect the proton pumping kinetics, we compare the absorption kinetics and structures of cR3 with those observed for other microbial rhodopsins.

5.1 Trimeric assemblies of archaeal rhodopsins

The crystal structure of cR3 crystal shows that the trimeric assembly is maintained under the crystallization condition. In fact, cR3 is not the unique microbial rhodopsin that forms trimeric structure. Crystal-

lographic studies reveal that the ability to assemble into trimers is possessed by several proton-pumping rhodopsins (bacteriorhodopsin (bR), archaerhodopsin-1 (aR1), archaerhodopsin-2 (aR2), deltarhodopsin-3 (dR3)) and also anion-pumping rhodopsins (halorhodopsin from *N. Pharaonis* (pHR) and halorhodopisn from *H. salinarum* (sHR)). Structural comparison shows that these trimeric assemblies are similar to one another (Figure 5.1). Comparison of the trimer of cR3 with those of aR2, bR and dR3 shows RMSD of 0.93 Å(over 618 residues), 0.98 Å(over 591 residues) and 0.56 Å(over 558 residues), respectively. Especially their transmembrane regions are superimposed very well; helices B, C and D are oriented in parallel to the 3-fold axis; while helices A, E, F, G are largely tilted from the membrane normal. Helices A and B, as well as helices D and E mediate direct contacts between neighboring subunits.

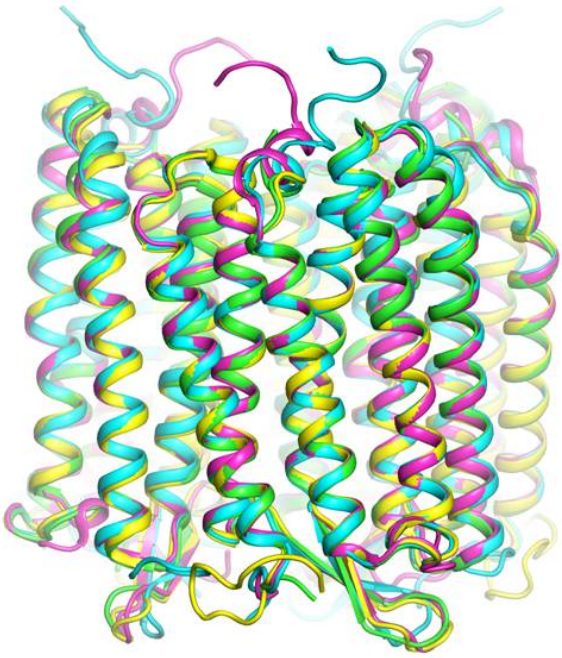

Figure 5.1: Trimeric assemblies of microbiol rhodopsins. Green: bacteriorhodopsin (pdb id: 1IW6); Cyan: cruxrhodopsin-3 (pdb id: 4L35); Yellow: archaerhodopsin-2 *H32* (pdb id: 3WQJ); Magentas: deltarhodopisn-3 (pdb id: 4FBZ).

Recently it was reported that a proton-pumping rhodopsin from Gloeobacter (a cyanobacterium) has the ability to form a trimeric assembly in the presence of dodecylmaltoside [50]. Boichenko et al. (2006) reported that the xanthorhodopsin (xR), a microbial rhodopsin found in eubacterium, may form a trimeric assembly in the cell membrane of *Salinibacter ruber* [51]. It seems reasonable to postulate that most proton-pumping microbial rhodopsins are able to form trimeric assemblies in their native membrane environments. It is interesting to ask: what advantages can these rhodopsins get by forming a trimeric assembly? Note that the proton pumping activity of bacteriorhodopsin monomer was reported to be similar to that of bR trimer [52], indicating that the trimeric assembly is not an essential structural motif for the proton-ion pumping activity. Meanwhile, it was reported that the thermal stability of bR monomer under illumination is much lower than that of bR trimer [53]. It is possible that formation of the trimeric assembly contributes to prohibition of unfavorable photo-reactions that may lead to inactivation of protein. Cruxrhodopsin-3 possesses a long DE loop that interacts with a neighboring subunit (Figure 4.15). This structural property enhances the ability of trimeric assembly formation and hence the photostability.

5.1.1 Role of bacterioruberin

Bacterioruberin is a carotenoid pigment abundant in the membrane of *H. salinarum* (Figure 4.1). It is responsible for the claret color of the *H. salinarum*. The primary role of this carotenoid pigment is to protect their host from ionizing radiation and UV by acting as an antioxidant [54]. In the crystal structure of cR3, a terminal end of bacterioruberin binds to the crevice at the cytoplasmic side between two subunits within a trimer (between helices A, B of one subunit and helices E, F of the neighboring subunit), while the other terminal is mobile in the lipid bilayer. In contrast to cR3, bacterhodopsin trimers contain another lipid component at the same position.

Recent studies have shown that xanthorhodopsin (xR), a new member of microbial rhodopsin found in eubacterium *Salinibacter ruber*, employs a different carotenoid (salinixanthin) as its second chromophore

to harvest sun light [2]. Crystal structure of xR shows that salinixanthin is within 5 Å from the β-ionone ring of retinal [55], which is in the range of the Van der Waals interaction distance (3 – 6 Å) such that excitation energy can be transfered effectively from salinixanthin to the retinal of xR. In cR3, on the other hand, bacteriorhodopsin is separated from the retinal; the closest distant between these pigments is 10.9 Å (Figure 5.2). It is unlikely that efficient excitation energy transfer can occur between bacteriorhodopsin and retinal.

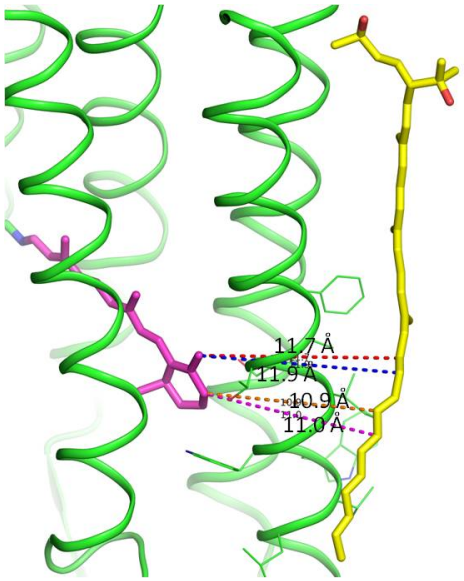

Figure 5.2: The closest distance between retinal and bacterioruberin in the trimeric assembly of cR3 is longer than 10 Å.

Structural comparison between cR3, aR2 and dR3 shows that the residues surrounding bacterioruberin are not necessarily conserved (Figure 5.3). In the *P321* crystal of aR2, the N terminal polypeptide chain forms an omega loop that provides a binding platform for bacterioruberin. As a consequence, the electron density map of the whole part of bacterioruberin can be identified. However, bacterioruberin in cR3 and dR3 is fixed only at one terminal end, and the other end is allowed to move freely in the lipid bilayer. The electron density of this mobile terminal of bacterioruberin is too weak to determine the exact configuration.

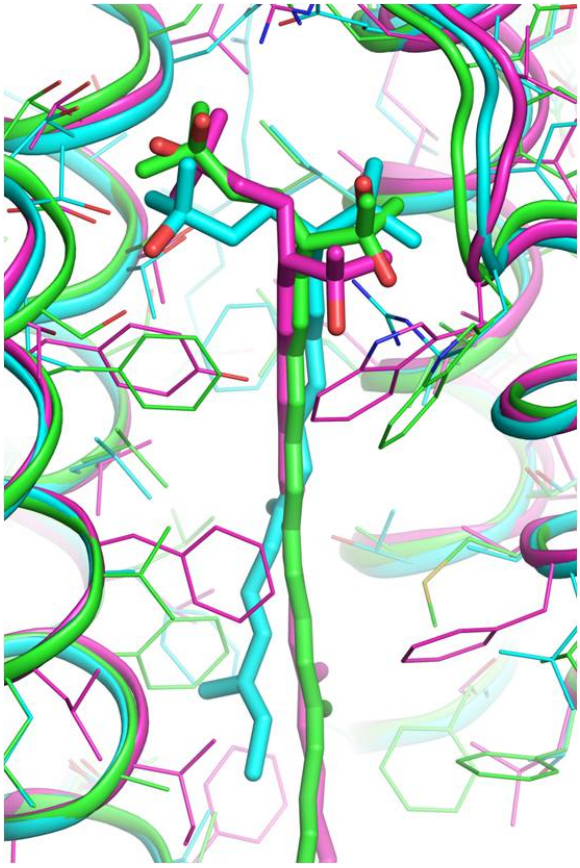

Figure 5.3: Structures of the bacterioruberin-binding site observed in the *P321* crystal of aR2 (magenta), the *P321* crystal of cR3 (green) and the *R32* crystal of dR3 (cyan)

A previous crystallographic study of the chloride ion pump halorhodopsin from *N. Pharaonis* (pHR) has shown that bacterioruberin is inserted in the crevices between neighboring subunits within the tirmeric assembly [28]. A recent study has shown that although the trimeric assembly of pHR is strengthened in the presence of bacterioruberin, pHR can form a trimeric assembly even in the absence of bacterioruberin [56]. Meanwhile, structural analysis shows that the bacterioruberins in pHR trimeric assembly are partially replaced by detergent when pHR crystals were grown at high detergent concentration. Also it is known that bacterioruberin is missing in the purple membrane of *H. salinarum*. We cannot exclude the possibility that native membrane of *Haloar-*

cula vallismortis contains a different type of lipid, which is replaced by bacterioruberin in the host cells in which cR3 is expressed.

Although the absorption spectrum of aR3-rich claret membrane shows characteristic peaks of bacterioruberin at 475 nm, 501 nm and 545 nm, the bacterioruberin content in aR3 is relatively lower than that observed for cR3. It is possible that the trimer of aR3 does not efficiently bind bacterioruberin or the binding affinity is low.

5.2 Structural conservation of proton releasing group

In bacteriorhodopsin, the proton-releasing pathway is composed of the Schiff base, Lys216^{bR}, Arg82^{bR}, Asp85^{bR}, Glu194^{bR} and Glu204^{bR}. Comparison between cR3 and bR shows that configurations of these functionally important residues are well conserved (Lys220^{cR}, Arg86^{cR}, Asp89^{cR}, Glu198^{cR} and Glu208^{cR} respectively). A minor difference is found in the water distribution within the proton-releasing channel: In cR3, two extra water molecules are observed in the proton-releasing pathway; Tyr83^{cR} (Tyr81^{bR}) is hydrogen-bound to Thr124^{cR}; an extra water Wat608^{cR} is placed between Glu198^{cR} and Tyr83^{cR}; and Wat607^{cR} is placed between Arg82^{cR} and Thr209^{cR}.

In the neutral purple form of bR in the *P622* crystal, the side chains of two glutamates in the proton-release channel (Glu194^{bR} and Glu204^{bR}) point towards each other. This glutamates pair is suggested to be maintained by a low barrier hydrogen bond [49]. Such a paired structure is also observed in the crystals of aR2, dR3 and cR3 that were prepared by the membrane fusion method (Figure 5.4 a). Recently we resolved the crystal structure of aR2 in the *H32* crystal at 1.8 Åresolution. This high resolution structural analysis allowed us to determine the detailed structure of the proton-release complex. Structural refinement shows that the distance between Glu199^{aR2} OE2 and Glu209^{aR2} OE1 is 2.3 Å, indicating that the glutamate pair is maintained by a low barrier hydrogen bond. Meanwhile, in the alkaline pink form (pH $>$ 10) and the acidic blue form (pH $<$ 3.5) of bR, such paired structure is broken [10]. In the acidic blue form of bR, the side chain of Glu194^{bR} points towards Arg134^{bR} and that of Glu204^{bR} points towards

helix F, creating a large cavity to accommodate a sulfate ion between Arg82^{bR}, Tyr83^{bR}, Glu194^{bR} and Glu204^{bR}. In the alkaline pink form, Glu194^{bR} turns upward to point at Tyr83^{bR}, and three water molecules fills an open space created between Glu194^{bR} and Glu204^{bR}.

It is also reported that in some crystals of bR (*C2* crystal [57]), the paired structure of Glu194^{bR} and Glu204^{bR} is broken. Interestingly, the orientations of side chains of these glutamates observed in the *C2* crystal are similar to those found in the alkaline pink form of bR in the *P622* crystal (Figure 5.4 b). On the other hand, although Glu194^{bR} and Glu204^{bR} can form a paired structure in another crystal of bR (*P6*$_3$ crystal), the configuration of the proton release complex is different from that seen in the neutral purple form of bR in the *P622* crystal. Furthermore the B-factor of the Glu194^{bR} OE2 was found to has an unusually high value (62 Å^2), which is much higher than the corresponding value observed in the *P321* crystal of cR3 (39 Å^2) and of *H32* crystal of aR2 (31 Å^2). A possible explanation is that the pKa of the alkaline transition of bR from the purple form into the pink form is dependent on the lipid-protein interactions. When the *P321* crystal of cR3 that was soaked in a wide range of pH (4 – 7), Glu198^{cR3} and Glu208^{cR3} remained to form a paired structure as seen in the *H32* crystal of aR2 and in the *P622* crystal of bR (the neutral purple form).

5.3 Structure factors affecting the reaction kinetics

The absorption kinetics of the bR in purple membrane at pH 7.5 and the aR2 claret membrane at pH 7 are shown in Figure 5.5 [1]. The flash-induced absorption changes of bR purple membranes at pH 7.5 were fitted by five exponential components (P_1 to P_5), which correspond to the K-to-L transition, the L-to-L/M transition, the L/M-to-M transition, the M-to-N (possibly a mixture of N and O state) transition and the decay of the long lifetime N state respectively, indicating a photoreaction scheme of K $\rightarrow$ L $\rightarrow$ L/M $\rightarrow$ M $\rightarrow$ N $\rightarrow$ O $\rightarrow$ bR. On the other hand, the flash-induced absorption changes of aR2 claret membrane at pH 7 were fitted by four exponential components (P_1, P_3, P_4 and P_5).

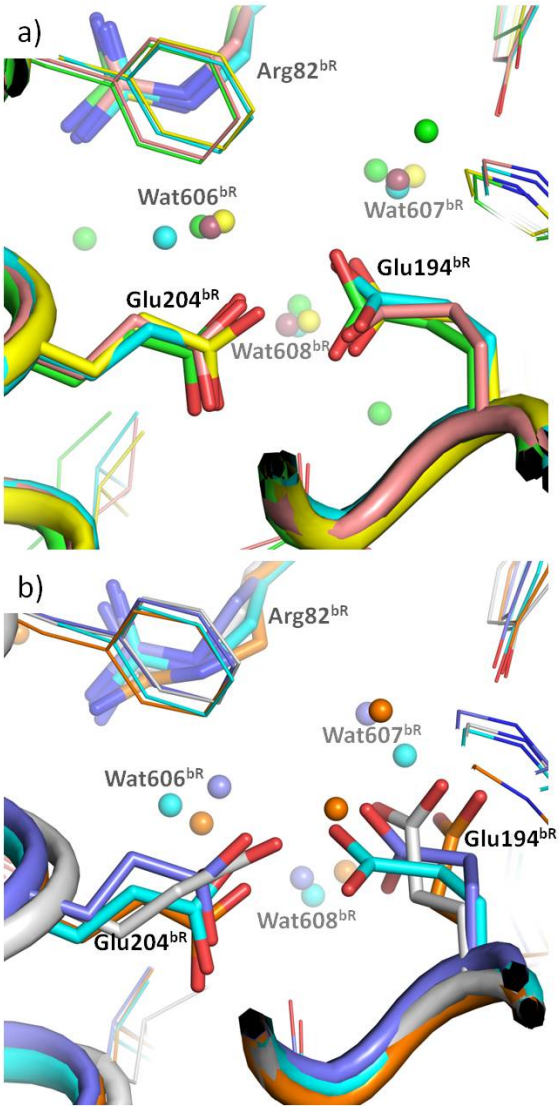

Figure 5.4: Structural conservation of the proton-releasing pathway among different proton-pumping rhodopsins. a) green: cR3 in the *P321* crystal (pdb id: 1C3W); pink: aR2 in the *H32* crystal (pdb id: 3WQJ); cyan: bR in the *P622* crystal (pdb id: 1IW6); yellow: dR3 in the *R32* crystal (pdb id: 4FBC). b) cyan: bR in the *P622* crystal (pdb id: 1IW6); purple: bR in the *P63* crystal (pdb id: 1C3W); white: bR in the *C2* crystal (pdb id: 1BRR); orange: the alkaline pink form of bR in the *P622* crystal (pdb id: 1X0K).

The four components correspond to the K-to-M transition, the M-to-N transition, the M/N-to-O transition and O-to-aR2 respectively. The photoreaction scheme of aR2 was described by K $\rightarrow$ M $\rightarrow$ N $\rightarrow$ O $\rightarrow$ aR2.

Comparison of absorption kinetics data shows that although more than half of the amino sequence of cR3 conserve with bR, the photo-cycle of cR3 is different from bR's photo-cycle in several aspects. Namely, the decay rate of the K state of cR3 is much slower than that of bR (time constant of the K-to-L transition of bR is $\tau_1 = 3.1$ μs at pH 7.5 versus that of cR3 at pH 8 is $\tau_1 = 0.031$ ms); the L state of cR3 could not be clearly detected; also the decay rate of M state of cR3 is much faster than that of bR (time constant of the M-to-N transition of bR is $\tau_4 = 2.67$ ms at pH 7.5 versus that of cR3 at pH 8 is $\tau_2 = 0.341$ ms). In the following section we tried to discuss the structural factors corresponding to the differences in reaction kinetics.

The slow decaying K state was not only observed in cR3, but also in aR2. Flash-induced absorption change of aR2 claret membrane showed that the K state decayed slowly to M state in a wide range of pH levels (pH 6 – pH 9, $\tau_1 = 0.061$ ms at pH 7) [17]. The decay rate of K state is expected to correlate with the structure around the retinal chromophore. Detailed comparison between the structure of the retinal binding pockets of aR2, bR and cR3 shows that most residues are conserved among microbial rhodopsins, the methionine in bR (Met145) is replaced by leucine (Leu149^{cR3}) in cR3 and by phenylalanine (Phe150^{aR2}) in aR2 (Figure 5.6 b). Such replacements is accompanied by a shift of a neighboring tryptophan residue (Trp182^{bR}, Trp186^{cR3} and Trp187^{aR2}) that makes contact with the C13 methyl of retinal, so that the structure around the C13 methyl of retinal is made more rigid and the horizontal movement of tryptophan (Trp186^{cR3}) is inhibited (Figure 5.6 b). It is reported that in bR, the K-to-L transition is accompanied with a horizontal movement of Trp182^{bR} towards Met145^{bR} and rotation of the side chain of Leu93^{bR} [58, 32], such configuration changes allow relaxation of the retinal polyene chain from a twisted 13-cis configuration to a planar 13-cis configuration (Figure 5.6 a). It is possible that the relaxation of retinal (twisted 13-cis to pla-

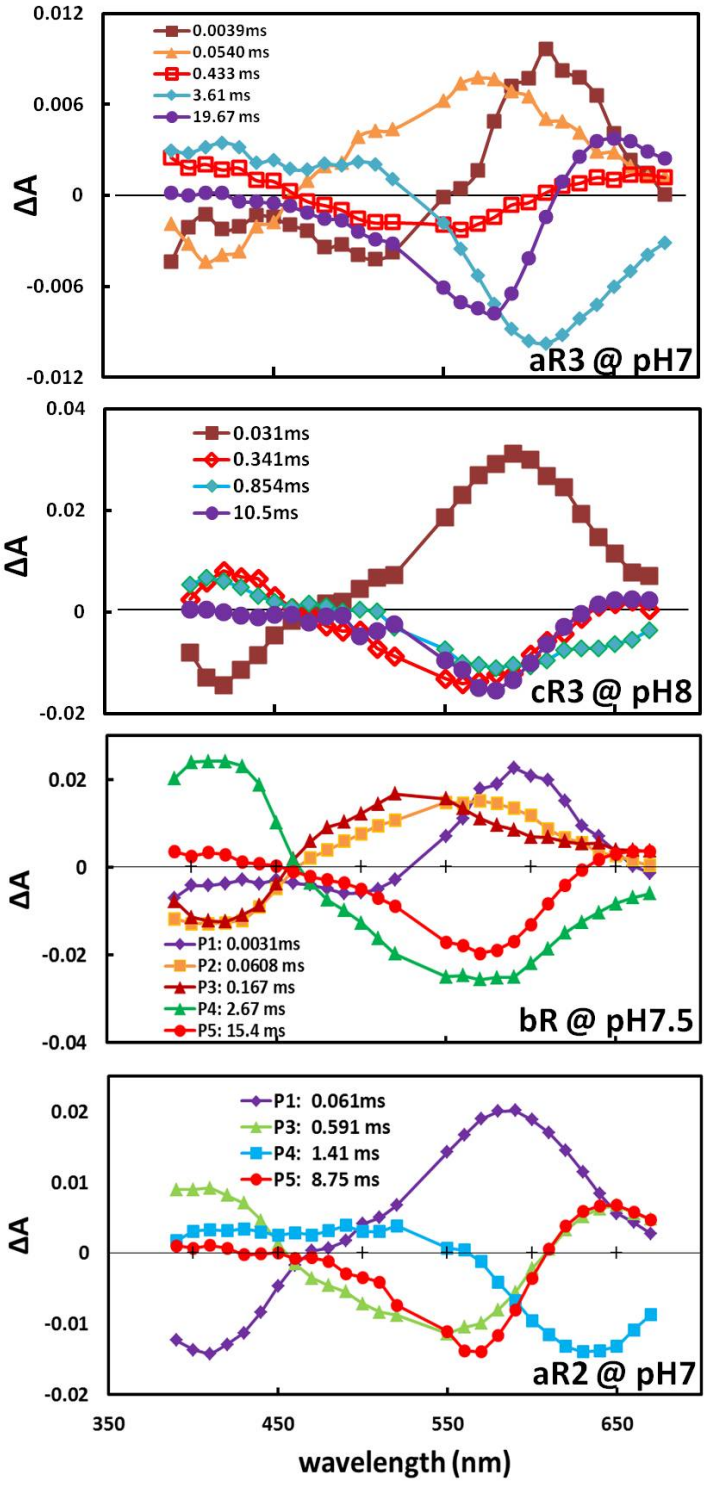

Figure 5.5: Flash-induced absorption changes of aR3 claret membrane at pH 7 (top panel); cR3 at pH 8 (second panel); bR purple membrane at pH 7.5 (third panel) and aR2 claret membrane at pH 7 (bottom panel). The flash-induced absorption changes of bR and aR2 are edited from [1].

nar 13-cis configuration) of cR3 and aR2 takes place in the formation process of the L/M state, during which the main chain of Lys220^{cR3} (Lys221^{cR3}) undergoes a large structural change.

Although the amino acid sequence of aR3 is highly conserved with that of aR2 (86 %), the decay rate of K state is much faster in aR3 then that in aR2 (time constant of K decay of aR3 at pH 7 is 0.0039 ms, similar to that of bR), structural comparison shows that the methionine sitting next to the tryptophan (Met155^{aR3}) is conserved with bR.

The difference in the rate of the M-to-N transition can be explained by the structural difference in the proton-uptake pathway. The cytoplasmic side of helix E of cR3 is largely bent to create a cavity between helix E and F. This cavity is large enough to accommodate a few water molecules. Previous structural analyses of bR have suggested that the cytoplasmic half of helix F is tilted outwards upon formation of the N state. A similar structural change has been shown to take place during the anion-pumping cycle of pHR in complex with azide. When its long-living N/M state of the pHR-azide complex is generated at high pH, the cytoplasmic half of helix F is largely deformed and a linear water cluster is formed between the retinal Schiff base and Lys215^{pHR} located near the cytoplasmic surface of pHR. It has been postulated that a similar linear water cluster is transiently generated during the proton pumping cycle of any proton-pumping rhodopsin. Because the unphotolyzed state of cR3 already contains a water molecule in the vicinity of Arg179 (the counterpart of Lys215^{pHR}), it would be expected that the formation of a linear water cluster in the cytoplasmic inter-helical space (i.e., the formation of the N state) takes place more rapidly in cR3 than in bR. It should be pointed out that in the *P321* crystal of cR3, the M state decays very slowly (t, 100 ms at 24 °C). This elongated lifetime of M is explained by supposing that the opening of the cytoplasmic half is inhibited by the protein-protein interactions, as previously reported for the M-to-N transition of bR.

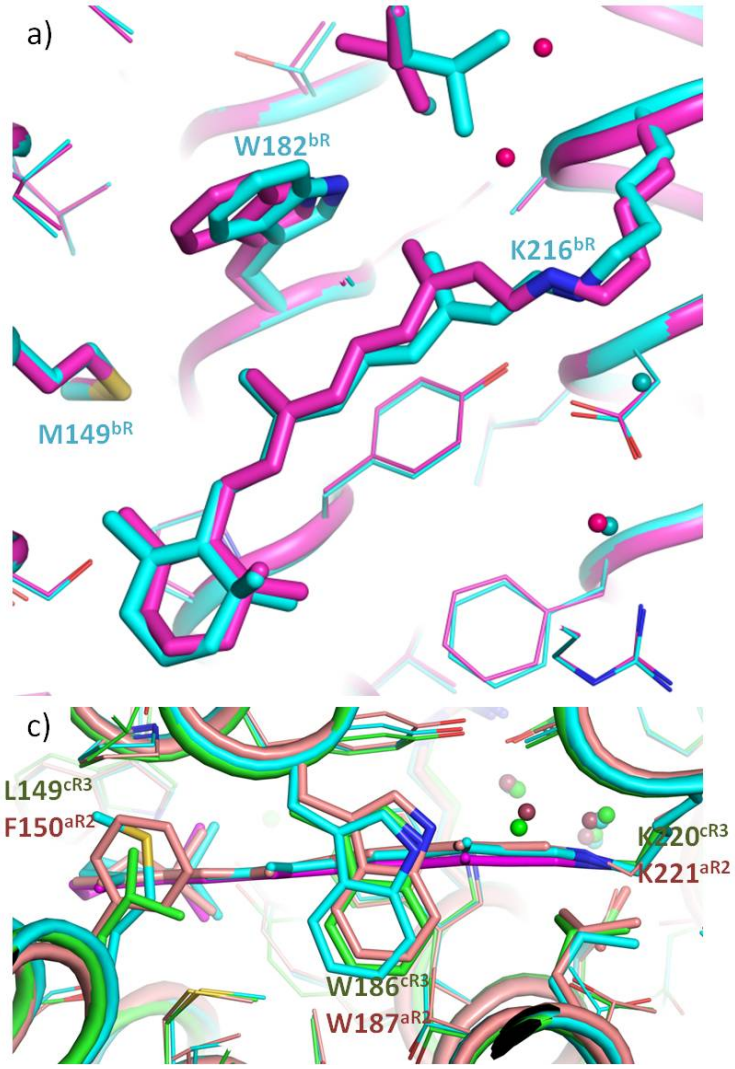

Figure 5.6: Comparison of the retinal-binding pocket. a) cyan: bR in the *P622* crystal (pdb id: 1IW6); purple: the L-intermediate of bR in the *P622* crystal (pdb id: 1UCQ). b) cyan: the unphotolyzed state of bR in the *P622* crystal (pdb id: 1IW6); green: cR3 in the *P321* crystal; pink: aR2 in the *H32* crystal (pdb id: 3WQJ).

Bibliography

[1] Siu Kit Chan, Tomomi Kitajima-Ihara, Ryudoh Fujii, Toshiaki Gotoh, Midori Murakami, Kunio Ihara, and Tsutomu Kouyama. Crystal structure of cruxrhodopsin-3 from haloarcula vallismortis. *PLOS ONE*, 9, sep 2014.

[2] Sergei P. Balashov, Eleonora S. Imasheva, Vladmir A. Boichenko, Josefa Anton, Jennifer M. Wang, and Janos K. Lanyi. Xanthorhodopsin: A proton pump with a light-harvesting carotenoid antenna. *Science*, 309(5743), 2005.

[3] Oded Beja, L Aravind, Eugene V. Koonin, Marcelino T. Suzuki, Andrew Hadd, Linh P. Nguyen, Stevan B. Jovanovich, Christian M. Gates, Robert A. Feldman, John L. Spudich, Elena N. Spudich, and Edward F. DeLong. Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science*, 289, 2000.

[4] Oliver P. Ernst, David T. Lodowsk, Marcus Elstner, Peter Hegemann, Leonid S. Brown, and Hideki Kandori. Microbial and animal rhodopsins: Structures, functions, and molecular mechanisms. *Chem Rev.*, 114(1), 2014.

[5] D Oesterhelt and W Stoeckenius. Rhodopsin-like protein from the pruple membrane of halobacterium halobium. *Nat New Biol*, 39(223):149–52, sep 1971.

[6] D Oesterhelt and W Stoeckenius. Functions of new photoreceptor membrane. *Proc Natl Acad Sci USA*, 70(10):2853–7, oct 1973.

[7] J. Sasaki, L. S. Brown, Y. S. Chon, H. Kandori, A. Maeda, R. Needleman, and J. K. Lanyi. Conversion of bacteriorhodopsin into a chloride ion pump. *Science*, 269, 1995.

[8] Y. Sudo and J. L. Spudich. Three strategically placed hydrogen-bonding residues convert a proton pump into a sensory receptor. *Proc. Natl. Acad. Sci. USA*, 103, 2006.

[9] Taichi Nishikawa, Midori Murakami, and Tsutomu Kouyama. Crystal structure of the 13-*cis* isomer of bacteriorhodopsin in the dark-adapted state. *J. Mol. Biol.*, 352, 2005.

[10] Hideo Okumura, Midori Murakami, and Tsutomu Kouyama. Crystal structures of acid blue and alkaline purple forms of bacteriorhodopsin. *J. Mol. Biol.*, 351, 2005.

[11] R. Henderson and P. N. T. Unwin. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature*, 257, sep 1975.

[12] E M Landau and J P Rosenbusch. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci USA*, 93:14532–5, 1996.

[13] Eva Pebay-Peyroula, Gabriele Rummel, Jurg P. Rosenbusch, and Ehud M. Landau. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science*, 277, 1997.

[14] Hartmut Luecke, Hans-Thomas Richter, and Janos K. Lanyi. Proton transfer pathways in bacteriorhodopsin at 2.3 angstrom resolution. *Science*, 280, may 1998.

[15] Hartmut Luecke, Brigitte Schobert, Hans-Thomas Richter, Jean-Philippe Cartailler, and Janos K. Lanyi. Structure of bacteriorhodopsin at 1.55 $\AA$ resolution. *J. Mol. Biol.*, 291, 1999.

[16] Janos K Lanyi and B Schobert. Crystallographic structure of the retinal and the protein after deprotonation of the schiff base: the

switch in the bacteriorhodopsin photocycle. *J. Mol. Biol.*, 321, 2002.

[17] Tsutomu Kouyama, Ryudo Fujii, Soun Kanada, Taichi Nakanishi, Siu Kit Chan, and Midori Murakami. Structure of archaerhodopsin-2 at 1.8 å resolution. *Acta Cryst.*, D70, 2014.

[18] Nobou Enami, Keiko Yashimura, Midori Murakami, Hideo Okumura, Kunio Ihara, and Tsutomu Kouyama. Crystal structures of archaerhodopsin-1 and -2: Common structural motif in archaeal light-driven proton pumps. *J. Mol. Biol.*, 358:675–685, 2006.

[19] Keiko Yoshimura and Tsutomu Kouyama. Structural role of bacterioruberin in the trimeric structure of archaerhodopsin-2. *J. Mol. Biol.*, 375, 2008.

[20] Tomomi Kitajima, Jun ichi Hirayama, Kunio Ihara, Yasuo Sugiyama, Naoki Kamo, and Yasuo Mukohata. Novel bacterial rhodopsins from *Haloarcula vallismortis*. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS*, 220, 1996.

[21] Hunefeld. Der chemismus in der thierischen organisation. *Leipzig: F. A. Brockhaus*, pages 158–63, 1840.

[22] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28:235–242, 1999.

[23] Eva Pebay-Peyroula, Gabriele Rummel, Jurg P. Rosenbusch, and Ehud M. Landau. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science*, 277, sep 1997.

[24] Salem Faham and James U. Bowie. Bicelle crystallization: A new method for crystallixing membarne proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.*, 316, 2002.

[25] Kazuki Takeda, Hidenori Sato, Tomoya Hino, Masahiro Kono, Kazuya Fukuda, Ikuko Sakurai, Tetsuji Okada, and Tsutomu

Kouyama. A novel three-dimensional crystal of bacteriorhodopsin obtained by successive fusion of the vesicular assemblies. *J. Mol. Biol.*, 283, 1998.

[26] Keiko Yoshimura and Tsutomu Kouyama. Structural role of bacterioruberin in the trimeric structure of archaerhodopsin-2. *J. Mol. Biol.*, 375:1267–1281, 2008.

[27] Jin Zhang, Katsuhide Mizuno, Yuki Murata, Hideaki Koide, Midori Murakami, Kunio Ihara, and Tsutomu Kouyama. Crystal structure of deltarhodopsin-3 from haloterrigena thermotolerans. *Proteins*, 81:1585–1592, 2013.

[28] Tsutomu Kouyama, Soun Kanada, Yuu Takeguchi, Akihiro Narusawa, Midori Murakami, and Kunio Ihara. Crystal structure of the light driven chloride pump halorhodopsin from natronomonas pharaonis. *J. Mol. Biol.*, 396:564–579, 2010.

[29] Midori Murakami and Tsutomu Kouyama. Crystal structure of squid rhodopsin. *Nature*, 453, may 2008.

[30] Akemi Nasuda-Kouyama, Kazuya Fukuda, Takayoshi Iio, and Tsutomu Kouyama. Effect of a light-induced ph gradient on purple-to-blue and purple-to-red transitions of bacteriorhodopsin. *Biochemistry*, 29, 1990.

[31] H. A. Hope. Cryocrystallography of iological macromolecules, a generally applicable method. *Acta Crystlalog. sect. B*, 44, 1988.

[32] Y Matsui, K Sakai, M Murakami, Y Shiro, S Adachi, H Okumura, and T Kouyama. Specific damaga induced by x-ray radation and structural changes in the primary photoreachtion of bacteriorhodopsin. *J Mol Biol*, 324, 2002.

[33] Jan Drenth. *Principles of Protein X-ray Crystallography*. Springer-Verlag New York, Inc., 1995.

[34] RF Peck, S DasSarma, and MP Krebs. Homologous gene knockout in the archaeon halobacterium salinarum with ura3 as a counter selectable marker. *Mol Microbiol*, 35(3), feb 2000.

[35] S W Cline, L C Schalkwyk, and W F Doolittle. Transformation of the archaebacterium halobacterium volcanii with genomic dna. *J Bacteriol*, 171(9), sep 1989.

[36] J D Boeke, F LaCroute, and G R Fink. A positiv selection for mutants lacking orotidine-5'-phoshate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet*, 197(2):345–346, 1984.

[37] S. G. Taneva, A. Dobrikova, I. B. Petkanchin, and F. M. Goni. Partial delipidation of purple membranes by tween 20: changes in surface electric properties. *Bioelectrochemistry and Bioenergetics*, 38(1):111–15, aug 1995.

[38] H Hino, E Kanamori, J R Shen, and T Kouyama. An icosahedral assembly of the light-harvesting chlorophyll a/b protien complex from pea chloroplast thylakoid membranes. *Acta Crystallgr D*, 60, 2004.

[39] Gina J. Fiala, Wolfgang W. A. Schamel, and Britta Blumenthal. Blue native polyacrylamide gel electrophoresis (bn-page) for analysis of multiprotein complexes from cellular lysates. *Journal of Visualized Experiments*, 48, feb 2011.

[40] I Steller, R Bolotovsky, and M G Rossmann. An algorithm for automatic indexing of oscillation images using fourier analysis. *J Appl Crystallogr*, 30:1036–1040, 1997.

[41] E J Dodson, M Winn, and A Ralph. Collaborative computational project, number 4: providing programs for protein crystallography. *Methods Enzymol*, 277:620–633, 1997.

[42] A T Brunger, P D Adams, G M Clore, W L DeLano, P Gros, R W Grosse-Kunstleve, J S Jiang, J Kuszewski, M Nilges, N S Pannu, R J Read, L M Rice, T Simonson, and G L Warren. Crystallography & nmr system: a new software suite for macromolecular structure determination. *Acta Crystallogr*, D54:905–921, 1998.

[43] D E McRee. *Practical protein crystallography*. San Diego: Academic Press, 1993.

[44] K Takeda, Y Matsui, N Kamiya, S Adachi, H Okumura, and Tsutomu Kouyama. Crystal structure of the m intermediate of bacteriorhodopsin: allosteric structural changes mediated by sliding movement of a transmembrane helix. *J. Mol. Biol.*, 341, 2004.

[45] K Arnold, L Bordoli, J Kopp, and T Schwede. The swiss-model workspace: A web-based enviroment for protein structure homology modelling. *Bioinformatics*, 22, 2006.

[46] A. Vagin and A. Teplyakov. Molrep: an automated program for molecular replacement. *J. Appl. Cryst.*, 30, 1997.

[47] T Kouyama, R A Bogomolni, and W Stoeckenius. Photoconversion from the light-adapted to the dark-adapted state of bacteriorhodopsin. *Biophys J.*, 48(2).

[48] Tsutomu Kouyama, A Nasuda-Kouyama, A Ikegami, MK Mathew, and W Stoeckenius. Bacteriorhodopsin photoreaction: indentification of a long-lived intermediate n (p,r350) at high ph and its m-like photoproduct. *Biochemistry*, 25, 1988.

[49] P Phatak, N Ghosh, H Yu, Q Cui, and M Elstner. Amino acids with an intermolecular proton bond as proton storage site in bacteriorhodopsin. *Proc. Natl Acad Sci U S A*, 105, 2008.

[50] T Tsukamoto, T kikukawa, T kurata, K H Kamo, and M Demura. Salt bridge in the conserved his-asp cluster in cloeobacter rhodopsin contributes to trimer formation. *FEBS Lett*, 587(4).

[51] Vladimir A Boichenko, Jennifer M Wang, Joesfa Anton, Janos K Lanyi, and Sergei P Balashov. Functions of carotenoids in xanthorhodopsin and archaerhodopsin, from action spectra of photoinhibition of cell respiration. *BBA Bioenergetics*, 1757(12), 2006.

[52] N A Dancher and M P Heyn. Bacteriorhodopsin monomers pump protons. *FEBS Lett*, 108, 1979.

[53] Yuri Mukai, Naoki Kamo, and Shigeki Mitaku. Light-induced denaturation of bacteriorhodopsin solubilized by octyl-β-glucoside. *Protein Engineering*, 12(9), 1999.

[54] Hamid Reza Shahmohammadi, Ezat Asgarani, Hiroaki Terato, Takeshi Saito, Yoshihiko Ohyama, Kunihiko Gekko, Osamu Yamamoto, and Hiroshi Ide. Protective roles of bacterioruberin and intracellular kcl in the resistance of textslHalobacterium salinarium against dna-damaging agents. *J. Radiat. Res.*, 39, 1998.

[55] Hartmut Luecke, Brigitte Schobert, Jason Stagno, Eleonora S Imasheva, Jennifer M Wang, Sergei P Balashov, and Janos K Lanyi. Crystallographic structure of xanthorhodopsin, the light-driven proton pump with a dual chromophore. *Proc natl Acad Sci USA*, 105(43).

[56] Takashi Tsukamoto, Takanori Sasaki, kazuhiro J Fujimoto, Takashi Kikukawa, Masakatsu Kamiya, Tomoyasu Aizawa, Keiichi Kawano, Naoki Kamo, and Makoto Demura. Homotrimer formation and dissociation of pharaonis halorhodopsin in detergent system. *Biophysical Journal*, 102, 2012.

[57] L. O. Esssen, R Siegert, W D Lehmann, and D Oesterhelt. Lipid patches in membrane protein oligomers: Crystal structure of the bacteriorhodopsin-lipid complex. *Proc Natl Acad Sci USA*, 95(20).

[58] T Kouyama, T Nishikawa, T Tokuhisa, and H Okumura. Crystal structure of the l intermediate of bacteriorhodopsin: Evidence for vertical translocation of a water molecule during the proton pumping cycle. *J Mol Biol*, 396, 2004.