Structural studies of cyanobacterial anion pump

Roman Astashkin

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique pour les Sciences du Vivant

Arrêté ministériel : 25 mai 2016

Présentée par

Roman ASTASHKIN

Thèse dirigée par **Valentin GORDELIY**

préparée au sein du **Laboratoire Institut de Biologie Structurale**
dans **l'École Doctorale Physique**

Etudes structurales de pompe anionique cyanobactérienne

Structural studies of cyanobacterial anion pump

Thèse soutenue publiquement le **24 janvier 2022**,devant le jury composé de :

Monsieur Martin ENGELHARD
PROFESSEUR, Max Planck Institute, Dortmund, Rapporteur
Madame Ekaterina ROUND
DOCTEUR EN SCIENCES, European XFEL, Examinatrice
Monsieur Marc JAMIN
PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes, Président
Monsieur Victor SOURJIK
PROFESSEUR, MPI Marburg, Rapporteur
Monsieur Johann KLARE
DOCTEUR EN SCIENCES, University of Osnabrück, Examinateur

Acknowledgements

First of all, I would like to sincerely thank my scientific advisor, Valentin Gordeliy, for the wonderful supervising of my scientific project. He has helped me incredibly in many aspects of my scientific work and has taught me a lot of the necessary skills.

I would also like to thank the reviewers of my doctoral dissertation: Martin Engelhard (MPI Dortmund) and Victor Surjik (MPI, Marburg) for the fruitful comments and suggested corrections for my dissertation. Their comments helped me a lot to make this job better.

I would especially like to thank my colleagues from Grenoble, who helped me to carry out this and other work. Kirill Kovalev, a former postdoc in our laboratory and now a postdoc at EMBL Hamburg, provided me with invaluable help and support, helping me to carry out this work, as well as teaching me the necessary skills for working with microbial rhodopsins. Sergei Bukhdruker became a Ph.D. student in our laboratory not very long ago, but he helped me a lot in the last stages of this project. I'd also like to thank Maxim Rulev for helping out with some of the experiments. Thanks to Vitaly Polovinkin who taught me many experimental procedures when I first arrived in Grenoble. Many thanks to Ivan Guschin for fruitful discussions on this project and valuable advices.

I would like to express special thanks to my colleagues from Forshungtsentrum Julich, namely Svetlana Vaganova, Taras Balandin, Christian Baeken and Dmitry Bratanov. Their advice and assistance in protein production helped me a lot during my business trips to Julich. In addition, they were always ready to help with experiments when I could not come to them due to a pandemic.

Thanks to the researchers whose equipment was used in this project. Many thanks to Antoine Royant (ESRF, icOS), Gleb Bourenkov (EMBL Hamburg), Alexander Popov (ESRF) and Montserrat Soler-Lopez (ESRF).

Sincere compassion to the other members of the jury of my doctoral defense, namely Marc Jamin (UGA, IBS, Grenoble), Johan Klare (University of Osnabrück) and Ekaterina Round (European XFEL, Hamburg).

Finally, of course, I want to thank my family and closed people who supported me during my Ph.D. studies.

List of abbreviation

CCCP - carbonyl cyanide m-chlorophenyl hydrazone

DDM – n-Dodecyl-B-D-maltoside

EDTA - Ethylenediaminetetraacetic acid

ESRF - European Synchrotron Radiation Facility

GPCR – G protein – coupled receptor

IPTG - isopropyl β-D-1-thiogalactopyranoside

LB – Lycogeny broth

LCP – lipidic cubic phase

MO - monoolein

MR – microbial rhodopsin

MP – membrane protein

Ni-NTA – Nickel-Nitrilotriacetic Acid

OD_{600} – Optical density at 600 nm wavelength

PCR – Polymerase Chain Reaction

PDB – Protein Data Bank

PMSF – Phenylmethane sulfonyl fluoride

RPM – rotation per minute

RSB – retinal Schiff base

SDS – Sodium dodecyl sulphate

SEC – size-exclusion chromatography

TB – Terrific broth

TM – transmembrane

TRIS – 2-amino-2-(hydroxymethyl)propane-1,3-diol

WT – wild type

Table of content

List of figures

Abstract

Ion transport across the cell membrane is an essential process for living cells. Active transport of ions is carried out by various ion transporters, including microbial rhodopsins (MRs).

Microbial rhodopsins are light-driven membrane proteins that can perform various functions such as proton or ion pumping, photo-sensing, ion channeling, and others. Ion pumps MRs are known to be able to pump various monovalent ions such as Na+, K+, Cl-, I-. However, until recently, no microbial rhodopsin was known capable of pumping divalent ions.

In 2017 first microbial rhodopsin able to pump divalent ions was described. It is rhodopsin from cyanobacteria Synechocystis sp. PCC 7509 called Synechocystis halorhodopsin or SyHR. This protein belongs to the family of cyanobacterial anion pumps. It is a chloride pump with a unique ability to pump sulfate in the absence of chloride.

In this thesis, the author obtained high-resolution three-dimensional structures of different states of this protein for the first time. There are chloride-bound form, sulfate-bound form, and two structures of active states: O and presumably K. All four structures have resolution better than 2.0 Å resolution. Obtained results give insights into the unique ability of this protein to pump divalent ions. In addition, the structures of active states explain the molecular mechanism of chloride transport in cyanobacterial anion pumps. This mechanism is incredibly similar to archaeal halorhodopsins, despite the fact that these two groups of MRs are not the closest relatives.

The results obtained make it possible to understand better the molecular principles of ion transport by cyanobacterial anion pumps and assume that proteins from this family can become promising tools for optogenetics.

Résumé

Le transport des ions à travers la membrane cellulaire est un processus essentiel pour les cellules vivantes. Le transport actif des ions est assuré par divers transporteurs d'ions, dont les rhodopsines microbiennes (MR).

Les rhodopsines microbiennes sont des protéines membranaires lumineuses qui peuvent remplir diverses fonctions telles que le pompage de protons ou d'ions, la photodétection, la canalisation des ions et autres. Les pompes à ions Les MR sont connues pour être capables de pomper divers ions monovalents tels que Na+, K+, Cl-, I-. Cependant, jusqu'à récemment, aucune rhodopsine microbienne n'était connue capable de pomper des ions divalents.

En 2017, la première rhodopsine microbienne capable de pomper des ions divalents a été décrite. Il s'agit de la rhodopsine de la cyanobactérie Synechocystis sp. PCC 7509 appelé Synechocystis halorhodopsin ou SyHR. Cette protéine appartient à la famille des pompes à anions cyanobactériennes. C'est une pompe à chlorure avec une capacité unique à pomper le sulfate en l'absence de chlorure.

Dans cette thèse, l'auteur a obtenu pour la première fois des structures tridimensionnelles à haute résolution de différents états de cette protéine. Il existe une forme liée au chlorure, une forme liée au sulfate et deux structures d'états actifs : O et vraisemblablement K. Les quatre structures ont une résolution meilleure que la résolution 2,0 Å. Les résultats obtenus donnent un aperçu de la capacité unique

de cette protéine à pomper des ions divalents. De plus, les structures des états actifs expliquent le mécanisme moléculaire du transport du chlorure dans les pompes à anions cyanobactériennes. Ce mécanisme est incroyablement similaire aux halorhodopsines archéennes, malgré le fait que ces deux groupes de MR ne sont pas les plus proches parents.

Les résultats obtenus permettent de mieux comprendre les principes moléculaires du transport d'ions par les pompes à anions cyanobactériennes et supposent que les protéines de cette famille peuvent devenir des outils prometteurs pour l'optogénétique.

1 Introduction

1.1. Microbial rhodopsins

Microbial rhodopsins (MRs) are a large family of photoactive membrane proteins. They consist of seven transmembrane (TM) helices and contain cofactor retinal which is covalently bound to lysine residue of the 7th TM helix via the retinal Schiff base (RSB)[1]. In the dark or ground state, the retinal is usually in the all-trans state, however, upon absorption of a photon, the retinal is isomerized to the 13-cis state[2]. This process triggers a chain of spectroscopic and structural changes in a protein called a photocycle. During the photocycle, the protein passes through a number of metastable states, called intermediate states, and returns back to the ground state. Each intermediate of the photocycle is characterized by its own absorption wavelength and is denoted by capital letters from I to O. During the photocycle, changes occurring in the protein make it possible to fulfill its function (Fig. 1.1.1). It is known that in nature, microbial rhodopsins are capable of performing various functions, such as proton or ion pumping or channeling, sensory and enzymatic activities[3].

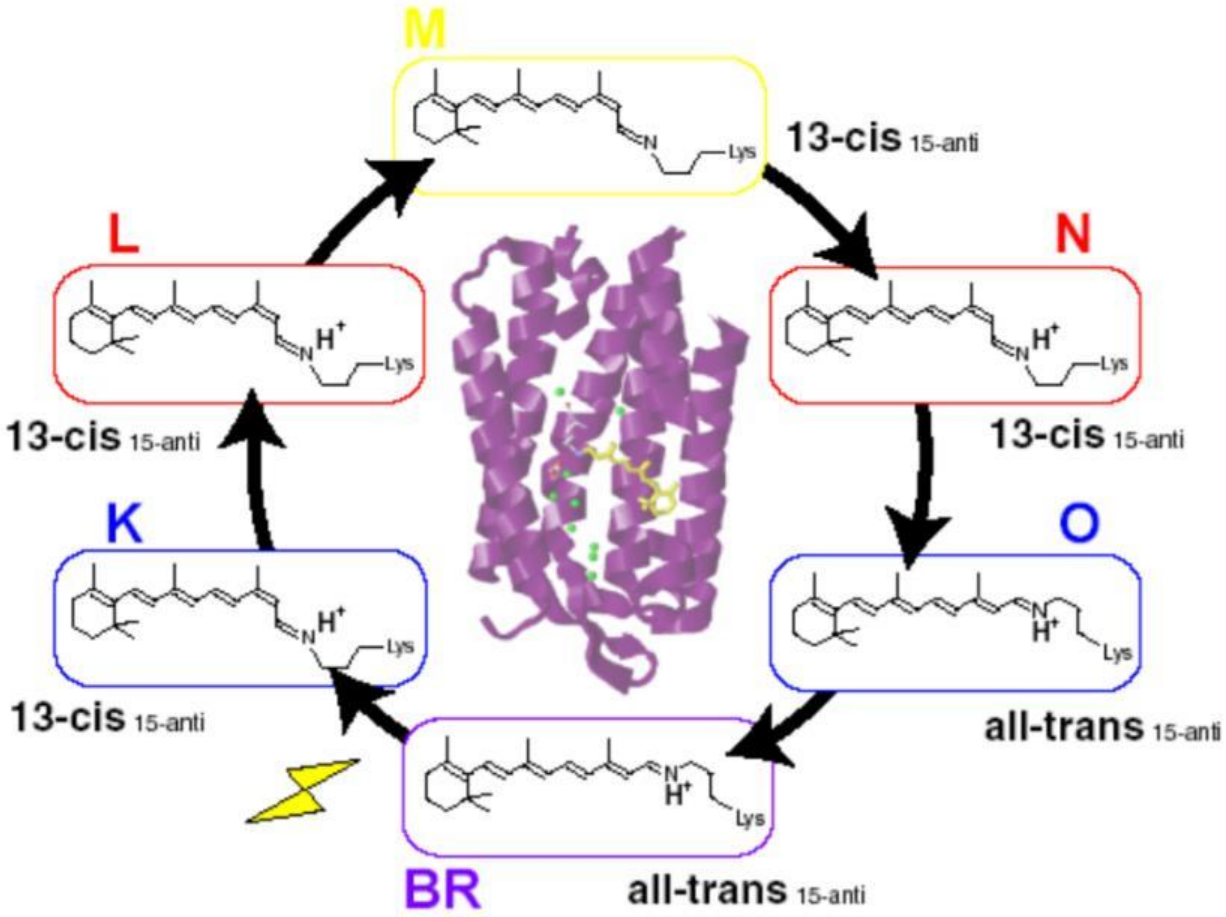

Figure 1.1.1. Photocycle of bacteriorhodopsin. Picture is taken from [4].

Microbial rhodopsins have been found in all domains of life as well as in viruses[5,6]. These proteins play an essential role in the bioenergetics of microorganisms. It has been shown that microbial rhodopsins are one of the main consumers of sunlight in the oceans[7]. Microbial rhodopsins are also called type-1 rhodopsins, as opposed to animal visual rhodopsins called type-2 rhodopsins[8].

The first discovered microbial rhodopsin was bacteriorhodopsin from archaea *Halobacterium salinarum*. This protein was discovered in 1971 by D.

Oesterhelt and W. Stoeckenius[9]. This protein was abundant (over 75%) in the purple membranes of these archaea[10]. Later it was shown that bacteriorhodopsin in the composition of purple membranes is the basis for the bioenergetics of the *Halobacterium salinarum*[11]. Also, this protein became the first among microbial rhodopsins, for which crystals and a three-dimensional structure were obtained[12,13].

The second discovered microbial rhodopsin was found in the same archaea *Halobacterium salinarum* in 1977[14]. Later, it was shown that this protein is capable of pumping chloride ions under the action of light, and it was named *Halobacterium salinarum* halorhodopsin or *Hs*HR[15].

The next discovered microbial rhodopsins were sensory rhodopsins I and II[16,17]. These proteins were unusual because they function in a complex with another protein, a transducer, and allow archaea to respond to light[18].

A revolution in the discovery of new microbial rhodopsins was the use of metagenomics to search for them. O. Beja et al. discovered in 2000 a new proton pump made from marine bacteria, which they named proteorodopsin (pR)[19]. This protein was the first microbial rhodopsin not found in archaea. It is proteorodopsins that are the main users of sunlight in the world's oceans.

Then, using metagenomics methods, a large number of new groups of microbial rhodopsins were discovered that perform very different functions. Channel rhodopsins 1 and 2 from green algae *Chlamydomonas reinhardtii* were discovered next. Channelrhodopsin-1 was discovered in 2002 by G. Nagel et al. and the channelrhodopsin-2 was discovered in 2003 by the same team[20,21]. It has been shown that channelrhodopsin-1 is a proton channel that opens under the action of light. Channelrhodopsin-2 turned out to be a similar light-sensitive channel, but not only for protons, but for cations in general. These proteins helped

the algae sense light and control its movement. Today, channelrhodopsin-2 has become the main instrument in optogenetics for activating neurons[22].

Within 20 years since the beginning of the use of metagenomics to search for new microbial rhodopsins, many new groups have been discovered and characterized (Fig. 1.1.2). Among these groups are reverse proton pumps - xenorhodopsins and eubacterial ion pumps (anionic ClRs[23] and cationic NaRs[24]). Group such as anion chanelrhodopsins (ACRs) was discovered in 2015[25]. Enzymorhodopsins have been found to function as enzymes in the cell[26]. A group of microbial rhodopsins was also found in giant viruses[6], for which it was shown that they can demonstrate channel activity[27,28]. However, the reason why viruses need microbial rhodopsins is unknown to date. Schizorodopsins have recently been found in Asgard archaea. These proteins have been shown to be inward proton pumps[29].

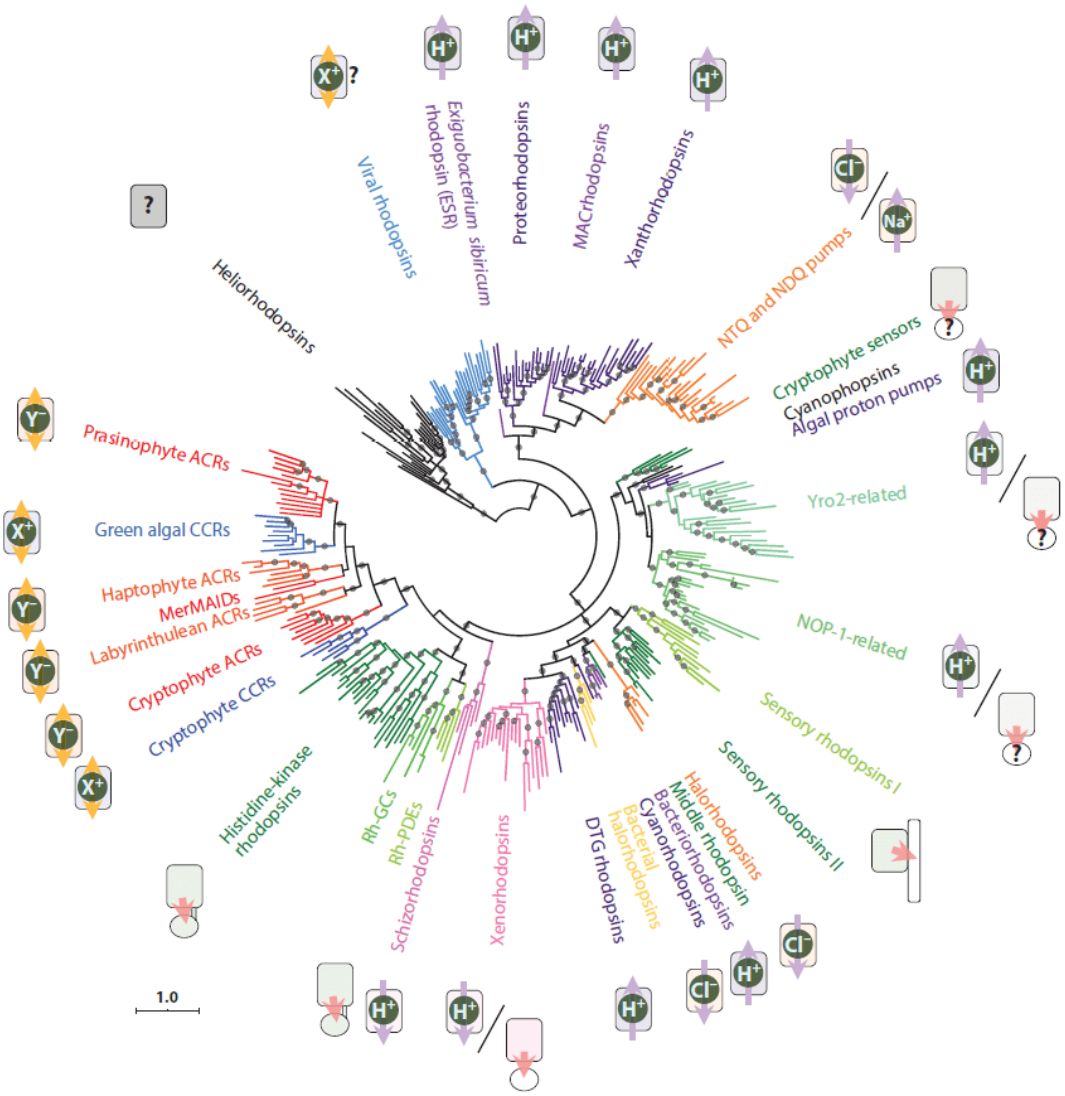

Figure 1.1.2. A phylogenetic tree of known for today groups of the microbial rhodopsins. The image taken from[30].

Heliorhodopsins, discovered in 2018, should be singled out separately[31]. These proteins have a topology unique to membrane rhodopsins (Fig. 1.1.3). They have an inverted position in the membrane with respect to all other known rhodopsins, that is, their N-end is facing the cytoplasm, and not into the

extracellular space. Common in all domains of life, they are divided into over 10 different groups. The variety of heliorhodopsin groups suggests that they can perform different functions. Nevertheless, the function of any of them is not yet precisely known, despite the fact that high resolution crystal structures have been published for two rhodopsins of this type (48C12 and TaHeR)[32–34]. Due to the large differences from other microbial rhodopsins of type 1, heliorhodopsins are often isolated into a separate group, different from type 1 rhodopsins and type 2 rhodopsins.

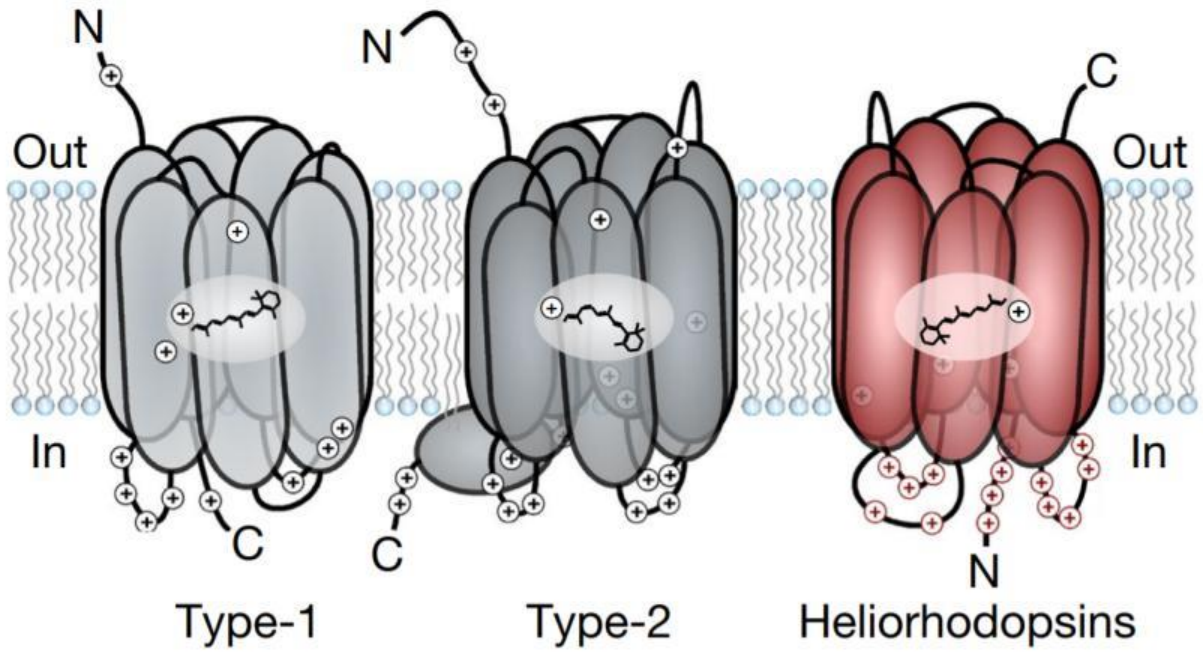

Figure 1.1.3. Topology of different rhodopsins in a cell membrane. Picture is taken from [31].

1.2 Functions of microbial rhodopsin

For many microbial rhodopsins, their biological role in cells is clear. For example, archaeal bacteriorhodopsins or bacterial proteorhodopsins pump a proton out of the cell. The high concentration of these proteins in the membrane of the microorganism allows maintaining the proton gradient required for the work of ATP synthase. In this way, these proteins allow microorganisms to receive energy from sunlight. According to research, bacteriorhodopsins or proteorhodopsins can be found in up to 80% of marine bacteria or archaea[35].

Sensory rhodopsins control the movement of the cell's flagellar motors and thus determine the parameters of its movement, depending on the light. This mechanism is called phototaxis. Sensory rhodopsin works in tandem with another protein, a transducer. The functional unit is a pair of heterodimers, that is, two molecules of sensory rhodopsin and two molecules of a transducer. The transducer has a long cytoplasmic portion with multiple methylation sites. Upon absorption of a photon, sensory rhodopsin undergoes structural changes and, due to structural rearrangements, transmits a mechanical signal to the transmembrane part of the transducer. Further in the cytoplasm, the signal passes through the HAMP domain to the kinase control module, which in turn regulates the activity of the CheA or CheW kinases, phosphorylating the CheY secondary messenger (Fig. 1.2.1)[36].

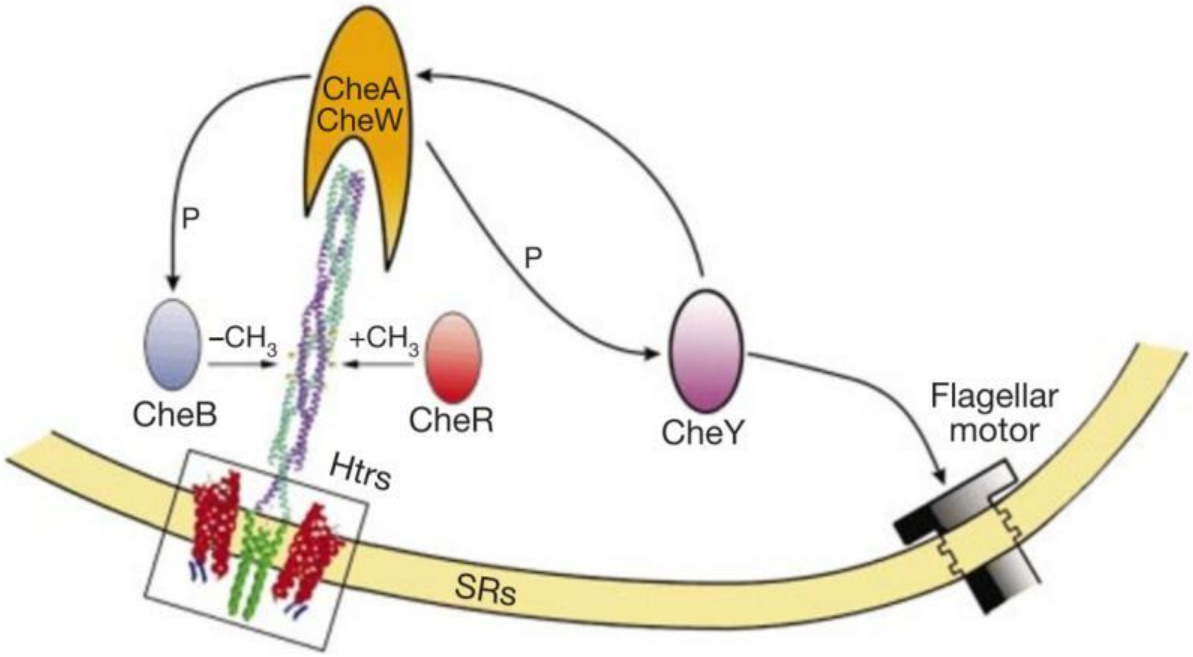

Figure 1.2.1. Two-component signalling cascade of *Np*SRII.
Activation of the HtrII transducer by its SRII receptor results in a conformational change in TM2 that extends to the tip of the cytoplasmic domain of the helix. The next steps in the signaling cascade include - by analogy with the bacterial sensory system - the homodimeric histidine kinase CheA, the binding protein CheW and the response regulators / aspartate kinase CheY and CheB. Phosphorylated (P) CheY acts as the switching factor of the flagellar engine. CheB (methylesterase) together with CheR (methyltransferase) are involved in the adaptation processes of archaea. Image is taken from [37].

Sensory rhodopsin I from *Natronomonas pharaonis* (*Np*SRI) induces positive phototaxis (movement towards light) upon exposure to light at 565 nm, and negative phototaxis (movement from light) upon exposure to light at 370 nm. Sensory rhodopsin II from *Natronomonas pharaonis* (*Np*SRII) induces only negative phototaxis on illumination at 480 nm[38]. It has been shown that bacteriorhodopsin HsBR can be converted into a light sensor by three mutations[39].

Despite the fact that a large number of anionic and cationic transporters are currently known, their exact biological role is still unknown.

There is an assumption that cationic, namely sodium, pumps arose in the era of ancient imperfect membranes, which were permeable to protons. It is believed that then ATP synthase worked on gradient of sodium ions, not protons[40]. Correspondingly, sodium pumps played the role of today's bacteriorhodopsins and proteorhodopsins[41]. However, what are sodium pumps for in the modern world is unknown. It has been hypothesized that these proteins may help the cell maintain osmotic balance, but these assumptions are not supported by experimental studies.

For chloride pumps, there is also no reliable information about what kind of biological function they perform. For archaeal halorhodopsins, they are believed to maintain an osmotic balance within the cell during cell growth[42]. Despite the prevailing consensus in the literature that the function of halorhodopsins is precisely this, all publications refer to one single work in which the idea of maintaining the osmotic balance is put forward only as a theoretical assumption. To date, no experiments have been conducted to confirm or refute this hypothesis.

For NTQ - rhodopsins it is also unclear. What function do they perform. As with archaeal halorhodopsins, it has been suggested that they may help maintain osmotic balance, but this has not been supported by any experimental tests either[30]. In the case of cyanobacterial anion pumps, it is also not known what their biological role is.

However, for some microbial rhodopsins, their biological role is completely incomprehensible. For example, it is completely unknown what biological role is played by representatives of the heliorodopsin family[34,43].

1.3 Identification of chloride pumping microbial rhodopsins

1.3.1 Archaeal halorhodopsins

At the moment, only three groups of microbial rhodopsins capable of active transport of anions are reliably known[44]. These are the well-known archaeal halorhodpsins, eubacterial chlorine transporters ClRs, and the recently discovered cyanobacterial anion pumps (Figure 1.2.1.1).

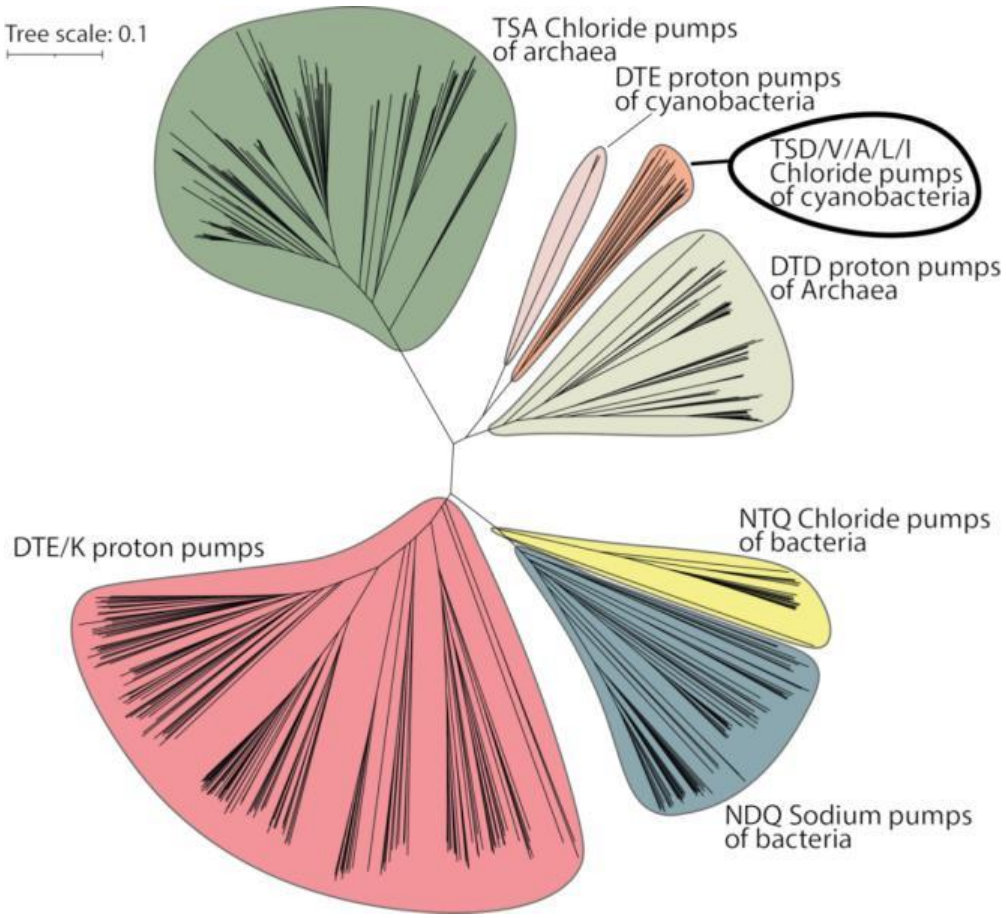

Figure 1.3.1.1. Phylogenetic tree of some microbial rhodopsins.

The other microbial rhodopsin was discovered in 1977 in the same archaea *Halobacterium salinarum*, in which bacteriorhodopsin was previously found[14]. Later it was shown that this protein is capable to perform active transport of chloride ion inside the cell and this protein was called *Halobacterium salinarum* halorhodopsin of *Hs*HR[45]. In 1986, another halorhodopsin was discovered in an extremely haloalkaliphilic archaea *Natronomonas pharaonis* and was called *Np*HR[46]. To this day, these two proteins remain the most studied archaeal halorodpsins. Despite the fact that the use of metagenomics has allowed the discovery of a large number of new archaeal halorodpsins, virtually all experiments were carried out with these two proteins.

A distinctive feature of archaeal halorhodopsins is the TSA motif, which means that Thr, Ser, and Ala, respectively, are located at the positions of the functionally important amino acid residues of *Halobacterium salinarum* bacteriorhodopsin (D85, T89, and D96)[47]. Like bacteriorhodopsin, these proteins are organized in the lipid membrane in the cell into trimers[48]. In general, halorodpsins, along with Middle rhodopsins, are the closest relatives of archaeal bacteriorhodopsins among all microbial rhodopsins. Halorodopsins have 25-40% sequence identity with bacterirodopsins[49,50]. It is known that replacing aspartate 85 in bacteriorhodopsin with threonine, which is in the corresponding position in halorhodopsins, it was possible to convert bacteriorhodopsin into a chloride pump[51]. Interestingly, even 10 bacteriorhodopsin-mimicking mutations failed to make halorodpsin a proton pump[52]. Halorhodopsins bind chloride in the ground state next to RSB; in addition, some structures show that another chlorine atom is bound to a protein in the cytoplasmic region[53,54]. Affinity constants for chloride

ions in the active center are 10 mM for halorhodopsin from *Halobacterium salinarum* and 2 mM for halorhodopsin from *Natronomonas pharaonis*[55,56].

Initially, it was believed that chalorodopsins can only pump chloride. It is known that in the absence of chloride archaeal halorhodopsins are capable of pumping various anions such as bromide[57], iodide[58] or nintrate[59]. An interesting observation is that all these anions are monovalent and at the moment there are no known archaeal halorhodopsins capable of pumping bivalent ions.

Photocycles of *Hs*HR and *Np*HR were experimentally characterized for different conditions in which the proteins were placed. In general terms, their photocycle is similar to the photocycle of the bacteriorhodopsin and has similar photocycle states. For example, NpHR under physiological conditions has the following photocycle[60]:

(NpHR $\rightarrow$ K $\leftrightarrow$ L^1 $\leftrightarrow$ L^2 $\leftrightarrow$ N $\leftrightarrow$ O $\leftrightarrow$ *Np*HR' $\rightarrow$ *Np*HR)

The main difference between the archaeal halorhodopsin photocycle and bacterirodopsin photocycle is the absence of the M state in the photocycle under normal conditions (Fig. 1.2.1.1.). However, the M state can be obtained by placing the protein in special conditions, for example, in a buffer containing azide instead of chloride[61].

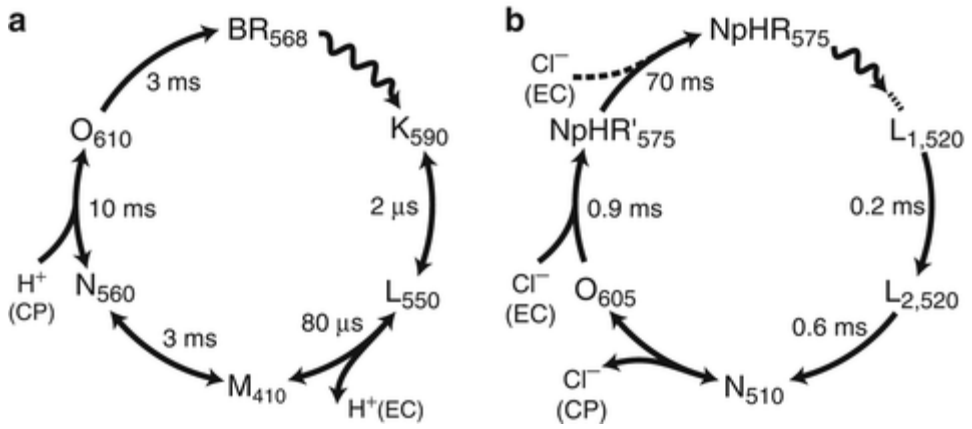

Figure 1.3.1.1 Comparison of photocycles of *Hs*BR (a) and *Np*HR (b). Image is taken from[62].

Archaeal halorhodopsins are widely used in optogenetics[63]. If you place proteins in the cytoplasmic membrane of a neuron that are capable of light-driven pumping of chlorine into the cell, then illumination of the cell with light will polarize the membrane and stop the incoming or emerging wave of depolarization[64]. This will lead to silencing of the neuron. Due to this fact, archaeal halorodpsins have become the main silencing tools in the optogenetics. The most widespread silencing instruments are NpHR and its modification eNpHR3.0[65].

1.3.2 Eubacterial chloride pumps

In 2014, a new family of microbial rhodopsins capable of pumping chloride was found in the marine flavobacteria *Nonlabens marinus*[23]. These rhodopsins have an NTQ motif (Asn, Thr and Gln at positions corresponding to D85, T89 and D96 in *Halobacterium salinarum* bacteriorhodopsin). These rhodopsins are called NTQ - rhodopsins or ClRs (formed from Chloride pumping rhodopsin)

These rhodopsins are closely related to sodium pumps NaRs with the NDQ motif. It is assumed that these proteins, like NaRs, function in the cell in a pentameric state[66]. Like archaeal halorhodopsins, these proteins contain a chloride ion next to RSB in the ground state[67]. In 2014, the first photocycle of a protein from the NTQ - rhodopsin family was measured[68]. It was a photocycle of a protein from the marine bacterium *Fulvimarina pelagi* called FR (Fulvimarina Rhodopsin). This protein shares 25% identity of amino acid sequence with halorhodopsin from *Halobacterium salinarum* (*Hs*HR) and 23% with halorhodopsin from *Natronomonas pharaonis* (*Np*HR). At the same time, it has 33% identity with sodium pump KR2 from *Krokinobacter eikastus*.

Photocycle measurements of FR showed that this photocycle is quite similar to the photocycle of archaeal halorhodopsins, since it also lacks the M - state. In general, the photocycle looks like this:

(FR → K ↔ L ↔ O ↔ FR’ → FR)

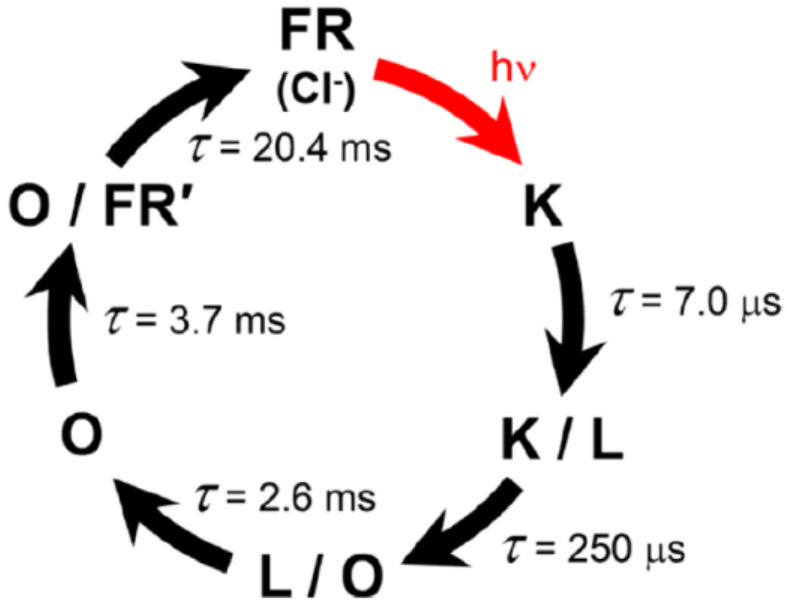

Figure 1.3.2.1. Photocycle of FR. Image is taken from[69].

Despite the fact that NTQ - rhodopsins are fairly well studied chloride pumps, there is currently no data on the successful use of bacterial anion pumps in optogenetics.

1.3.3 Cyanobacterial anion pumps

In 2015, a new family of chloride pumping rhodopsins was discovered in cyanobacteria[52]. These proteins do not have a strictly conservative three-letter motif. They have conserved Thr and Ser at positions D85 and T89 in bacteriorhodopsin, respectively. However, at the D96 position, these proteins have

different variants[70]. Asp is most common in this position, but Val, Leu or Ile are also found.

So far, only two proteins from this family have been experimentally characterized. The first is a protein discovered in 2015 from the cyanobacterium *Mastigocladopsis repens*. In the literature, it has 2 variations of the name: MrHR for *Mastigocladopsis Repens* HaloRhodopsin or MastR (*Mastigocladopsis* Rhodopsin).

The second protein from this family was the protein from the cyanobacterium *Synechocystis sp. PCC 7509*, called *Sy*HR[71]. This protein was found to have a unique feature among all microbial rhodopsins. Like *Mr*HR, *Sy*HR is a chloride pump. However, in the absence of chloride, this protein is capable of pumping sulfate. Before that, not a single microbial rhodopsin was known capable of pumping sulfate. In addition, this is the first known case of the pumping of a divalent ion by microbial rhodopsin.

Both of these proteins MrHR and SyHR have a TSD motif. The two proteins have a fairly similar amino acid sequence with 68% identity and 91% similarity. The introduction of SyHR-mimicking amino acid mutations made it possible to obtain MrHR mutants capable of pumping sulfate as well[72].

Both of these proteins have a chloride located next to the RSB in the ground state. In addition, the affinity of these proteins for chloride is, on average, higher than that of archaeal halorhodopsins. For *Mr*HR, this value is 2 mM, which is equal to the affinity of *Np*HR and 5 times better than that of *Hs*HR. In *Sy*HR, this value is 0.112 mM, which is the strongest affinity for chloride among all known microbial rhodopsins[71].

Photocycles for both of these proteins were measured. The photocycle of these proteins also does not have an M-state and under physiological conditions looks like this using *Mr*HR as an example:

(*Mr*HR → K → L^1 → L^2 ↔ N ↔ O ↔ *Mr*HR' → *Mr*HR)

As can be seen, the states of this photocycle are very similar to the states of archaeal halorhodopsins and, in particular, *Np*HR.

Attempts have also been made to understand, using FTIR spectroscopy, at which moments of the photocycle important stages of ion transfer occur, such as ion uptake and ion release. However, different researchers received different results, so at the moment there was no definite opinion on this matter. Harris et al. obtained data that the chloride release occurs after L^2 and before the N state[70]. And the chloride uptake is taken place between N and O state. But Takatoshi et al. got different data. They suggest that the chloride release occurs similarly after L state, but uptake happens after the O state[73]. A similar conclusion was made for the SyHR protein, which has a photocycle that is almost identical to the MrHR photocycle[71].

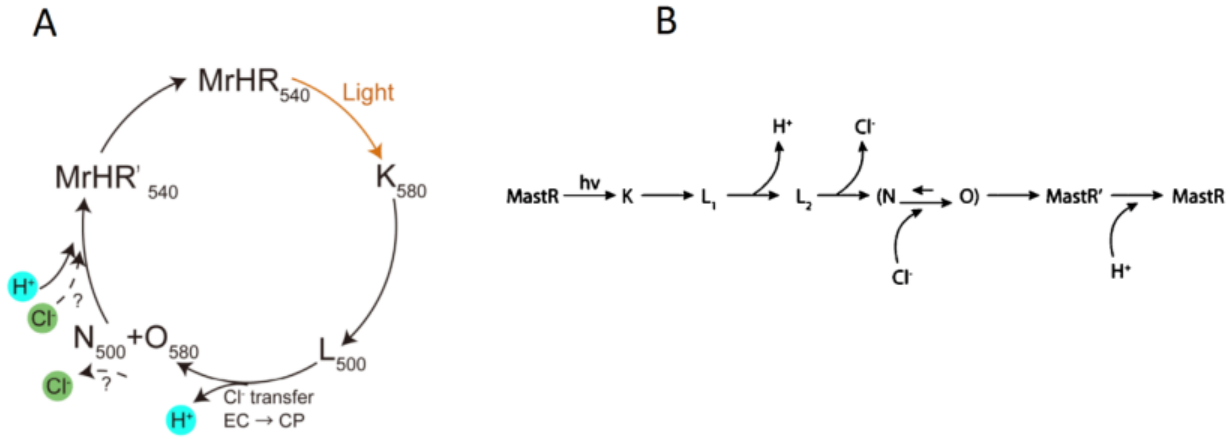

Figure 1.3.3.1. Photocycles of *Mr*HR. Different ideas about at what point in the photo cycle chloride release and uptake occurs. (A) Assumption from Takatoshi et al.[74] (B) Assumption from Harris et al.[70]. Image is adapted from [74] and [70].

1.4 Structural studies of chloride pumping rhodopsins

1.4.1 Introduction to structural studies of microbial rhodopsins

Attempts to obtain the structure of bacteriorhodopsin by X-ray structural analysis have led to the development of a large number of crystallization methods based on various lipid phases. For the first time, successful crystallization of microbial rhodopsins and obtaining three-dimensional structures became possible with the advent of crystallization technology in lipid cubic phases (Fig. 1.4.1.1) in 1997[13]. However, it was possible to obtain crystals of bacteriorhodopsin with the

help of other lipid phases. In 1998, crystals of bacteriorhodopsin were obtained using lipid vesicles[75]. In 2002, bacteriorhodopsin was successfully crystallized by crystallization from the bicelles[76].

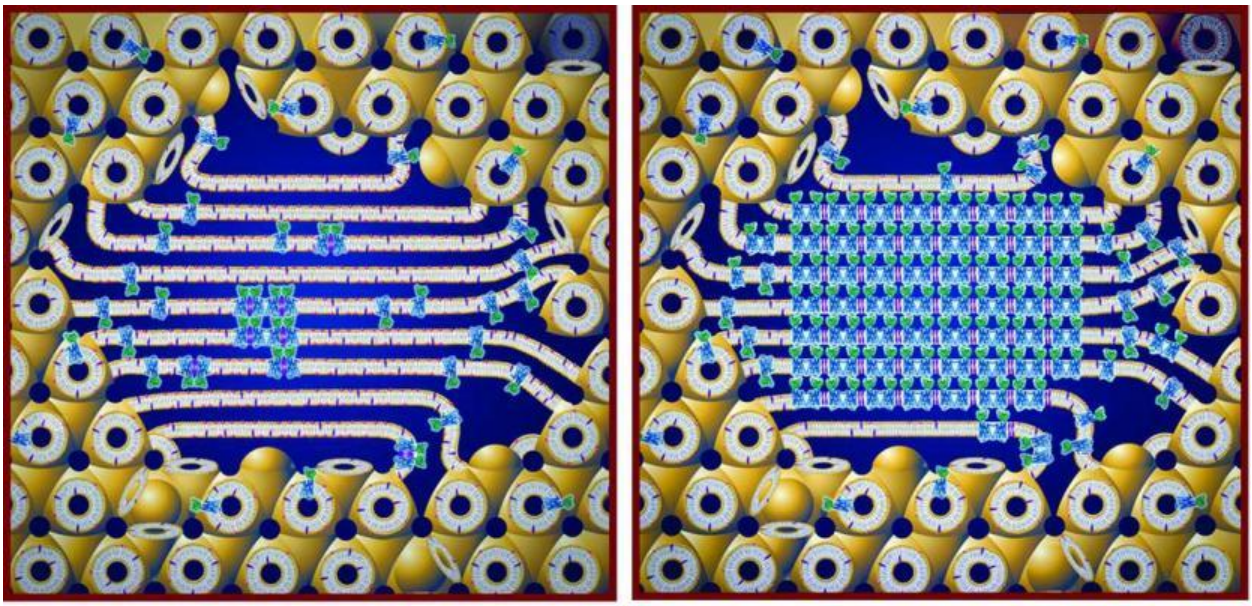

Figure 1.4.1.1. Schematic representation of presumptive protein crystallization in the lipid cubic phase. (A) Proteins freely migrating in the lipid cubic phase can spontaneously transform the lipid cubic phase into a sheet-like or lamellar domain at some site. (B) Protein molecules are concentrated in this region and form a growing crystal. Image is adapted from[77].

Initially, these techniques seemed too complex for routine crystallization, but automation and the creation of robots for high-throughput crystallization have led to widespread crystallization using lipidic mosephases also called *in meso*[78–81].

Later, more new methods were developed. Using bacteriorhodopsin as an example, the possibility of using amphipoles for the stabilization and

crystallization of membrane proteins was shown[82]. HsBRs in amphipoles or lipid nanodiscs can be successfully transferred to the lipid mesophase for conventional crystallization[83,84].

In meso crystallization methods, originally developed for the crystallization of bacteriorhodopsin, have become widespread for the crystallization of proteins from another family of 7-helical transmembrane proteins, the G-protein-coupled receptors[85–88].

1.4.2 Structures of archael halorhodopsins.

The structure of halorhodopsin from Halobacterium salinarum was obtained in 2000 by the lipid cubic phase method[42]. This protein has become the second microbial rhodopsin with a resolved three-dimensional structure, after bacteriorhodopsin. The resolution of the structure is 1.80 Å. The structure showed that HsHR, like bacteriorhodopsin, is organized into trimers, with three palmitic acid molecules in the center of the trimer (Fig. 1.4.2.1, A). Next to the RSB is a chloride ion bound to two water molecules, Ser115, Thr111, and nitrogen from RSB (Fig. 1.4.2.1, B).

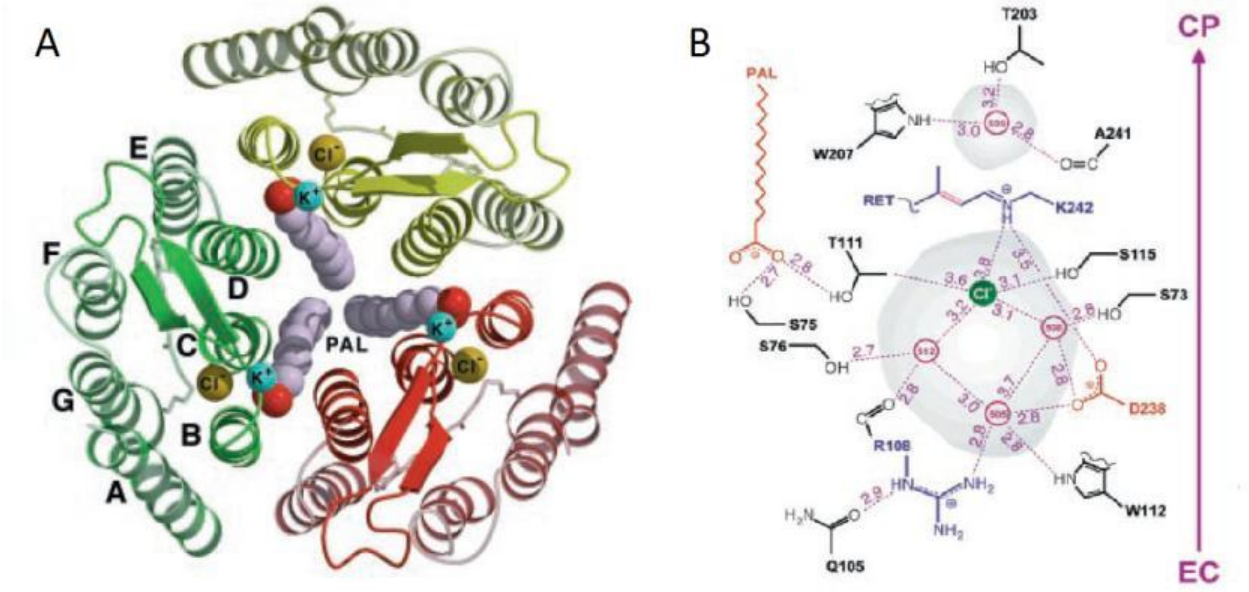

Figure 1.4.2.1 Structure of *Hs*HR. (A) Structural organization of trimer of *Hs*HR. View from the extracellular side. Cl^- is located near the retinal Schiff base and K^+ is located in the extracellular part of the protein. (B) Schematic representation of hydrogen bond network around RSC and chloride ion. Image is adapted from [42].

The next structure of archaeal halorhodopsin was obtained in 2007 with the T203V mutant of *Hs*HR[89]. This mutant has 10 times slower decay of the L state in comparison with the wild type protein. As a result of this work, the structure of the ground state of this mutant was obtained with a resolution of 1.7 Å, as well as the structure of the L1 state with a resolution of 1.9 Å. The structure of the ground state was obtained at a temperature of 100 K, and the structure of the L1 state was obtained using cryotrepping. Cryostream was blocked for 10 seconds and crystal was illuminated with 633 nm laser. The delayed decay of the L - state in this mutant simplified the preparation of this intermediate using trapping.

In contrast to the 2000 structure, two chloride binding sites were identified in this structure. A second site has been identified in the protein's cellular region

(Fig. 1.4.2.2, A). Comparison of the structures of the ground state and the L1 state showed very small structural changes between the two states. Minor changes occurred in the positions of amino acids Y109 and E219. Negative differential densities were observed on the retinal, which is evidence of isomerization from all-trans to 13-cis configuration of the retinal (Fig. 1.4.2.2, B).

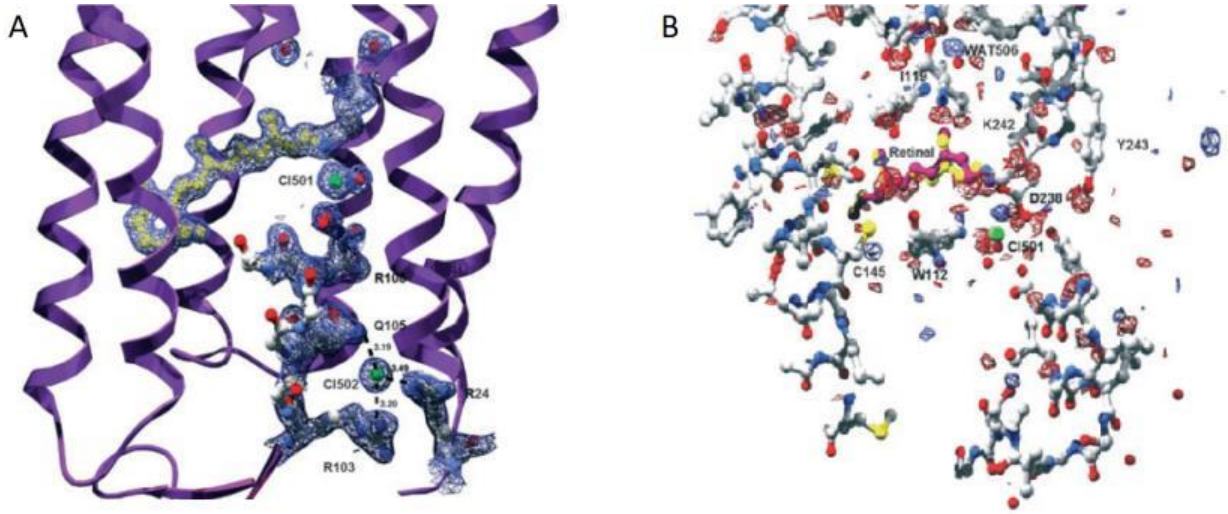

Figure 1.4.2.2. Structures of L1 state of T203V mutant of *Hs*HR. (A) Structure of the inner space of the protein. Electronic densities are shown for some amino acids. Two chloride binding sites are visible. (B) Difference maps of L1 densities and the ground state of the protein. Image is adapted from [89].

No corresponding positive density was found for this change; which can be explained by the presence of the intermediate 13-cis retinal in different places, possibly with different degrees of bending. In addition, small changes have occurred in the positions of amino acids such as S115, T116 and I119.

Two more papers have been published on the structures of halorhodopsin from *Halobacterium salinarum* in 2015 and 2016 by the same team[90,91]. In the first

work, *Hs*HR crystals were obtained in a new crystalline form with a resolution of 2.15 Å. The structure of the protein practically does not differ from those published earlier, however, the authors argue that the new crystalline is more free and may be better suited for studying active states by trapping intermediate or time resolved studies, including using XFEL.

In the second study, similar crystals were soaked in a strongly alkaline buffer to cause deprotonation of the Retinal Schiff base and loss of chloride in the active site. This procedure made it possible to get rid of chloride by about 50%; however, this greatly deteriorated the quality of the crystals and the resulting structure has only 2.6 Å resolution. The structure is no different from that described in the previous work, except for occupancy for chloride and less detail due to poorer resolution.

In 2010, the structure of archaeal rhodopsin was obtained for the first time from another organism, *Natronomonas pharaonis*[54]. The resolution of the resulting structure was 2 A. In the resulting structure, the protein was also packed into trimers and, in general, the structure was very similar to the *Hs*HR structure, which is not surprising given the relatedness of these proteins.

In 2011, the same team received the O - like blue form and anion - free yellow form of NpHR structures[92]. In this work, the researchers obtained anion-free blue form of the protein, which resembled an O-state with a resolution of 1.8 A. They also obtained anion-free yellow form, which was obtained from the blue form when the pH changed to more acidic. In the O-like blue form, a T126 flip was observed, which rotated and took the place of chloride in the active center (Fig.1.4.2.3). In addition, other amino acids such as Y124, E234, W127 have changed their position.

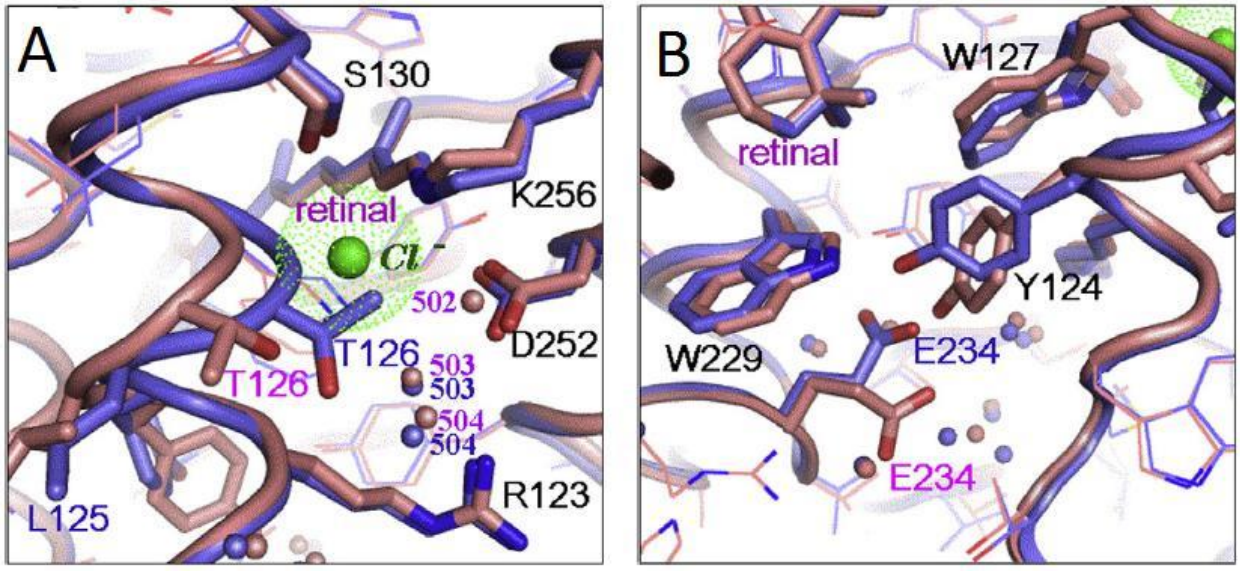

Figure 1.4.2.3. Comparison of O-like anion free form of *Np*HR (blue) and chloride – bound form (purple). (A) Structure of active center of the protein. (B) Structure of chloride uptake region. Image is adapted from [92].

In 2013, the same researchers determined the structure of NpHR in a complex with an azide. This structure is of interest, since the azide strongly influences the functions of halorhodopsins[61]. In *Hs*HR, it inhibits activity, and turns *Np*HR into a proton pump[93,94].

In 2015, the same group published an article on the structures of active states L1, L2, N and O[95]. The studies were carried out with bromide-bound protein. A series of obtained structures showed the following changes during the photocycle. In the L1-to-L2 transition, the bromide ion that initially exists in the extracellular vicinity of retinal moves across the retinal Schiff base. Upon the formation of the N state with a bromide ion bound to the cytoplasmic vicinity of the retinal Schiff base, the cytoplasmic half of helix F moves outward to create a water channel in the cytoplasmic interhelical space, whereas the extracellular half of helix C moves inward (Fig. 1.4.2.4).

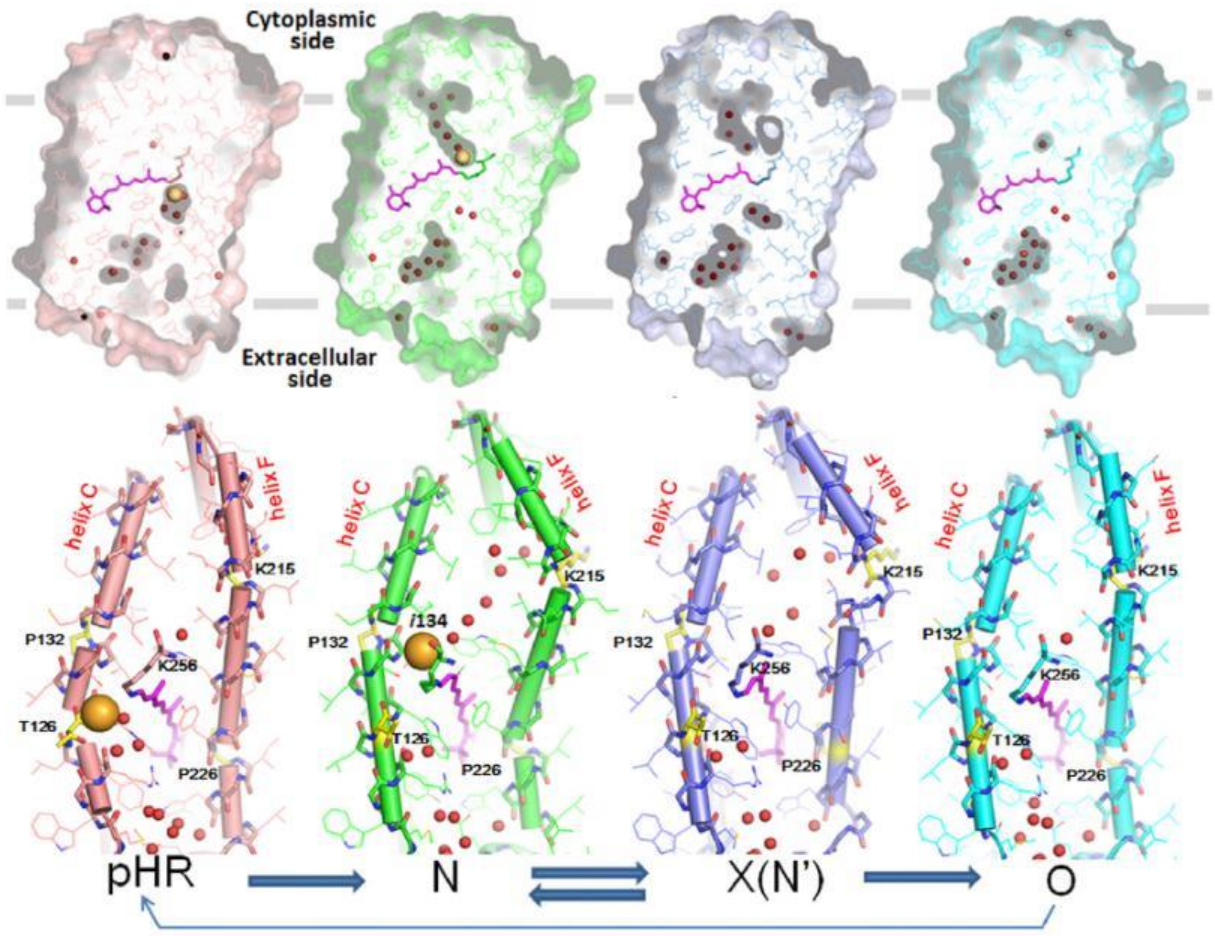

Figure 1.4.2.4. Structural conformational changes during the anion-pumping cycle of *Np*HR. Image is taken from [95].

Finally, in 2016, the unusual structure of the 11-cis isomer of HsHR was obtained[96]. It is shown in the work that when HsHR crystals are illuminated with a red laser in the anion-free form, a red-shifted state appears, which is characterized by 11-cis isomerization of retinal. The authors suggest that this isomerization may occur at later stages of the HsHR photocycle.

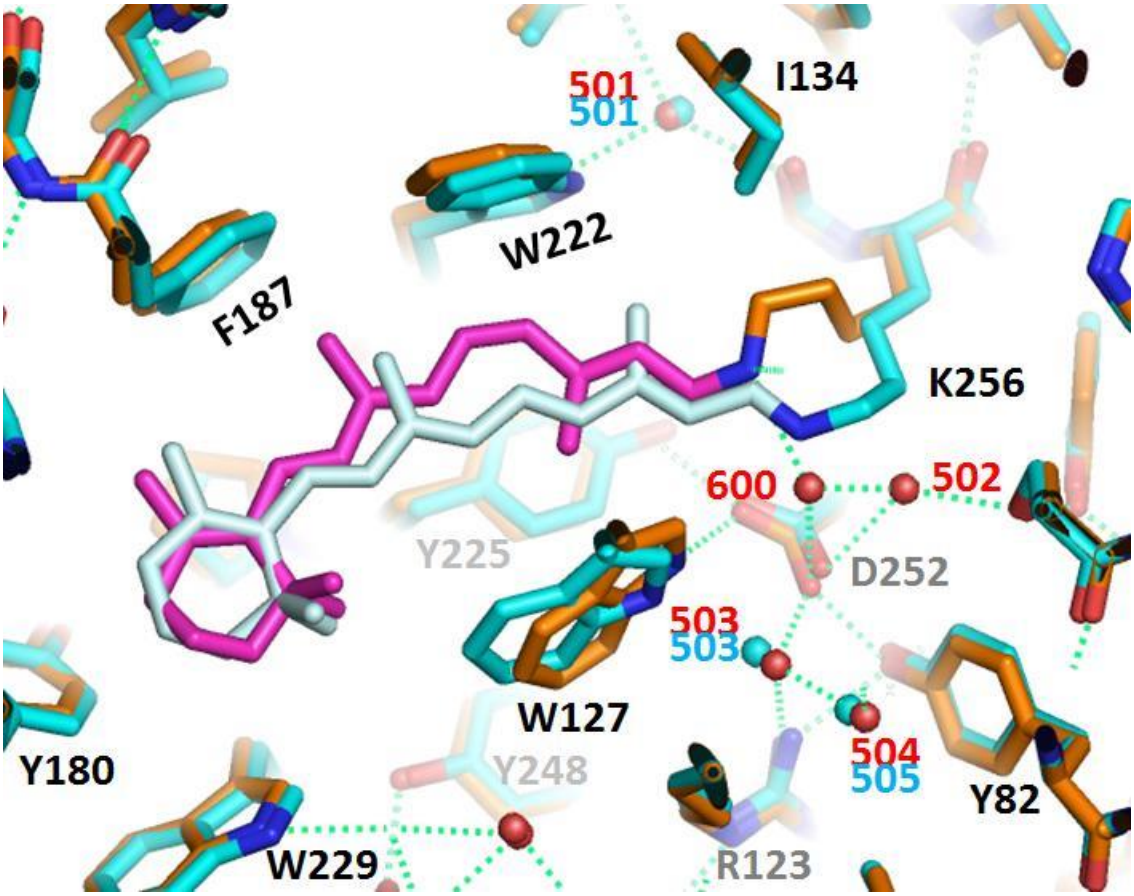

Figure 1.4.2.5. Comparison of all-trans and 11-cis retinal in anion free form of *Np*HR. Image is taken from [96].

1.4.3 Structures of NTQ – rhodopsins

Today, a large number of three-dimensional structures of eubacterial chloride pumps are known. There are 26 published three-dimensional structures of NTQ - rhodopsins, but all these structures belong to the same protein - rhodopsin 3 from *Nonlabens marinus* (*NM*-R3, or *Nm*ClR).

In 2016, the first NmClR structure was obtained[97]. The three-dimensional structure was solved with a resolution of 1.58 Å and contained one protomer in an assymetrical unit. This protein also has a chlorine ion next to RSB; however, in general, the structure of the active site is somewhat different from archaeal halorhodopsins (Fig. 1.4.3.1).

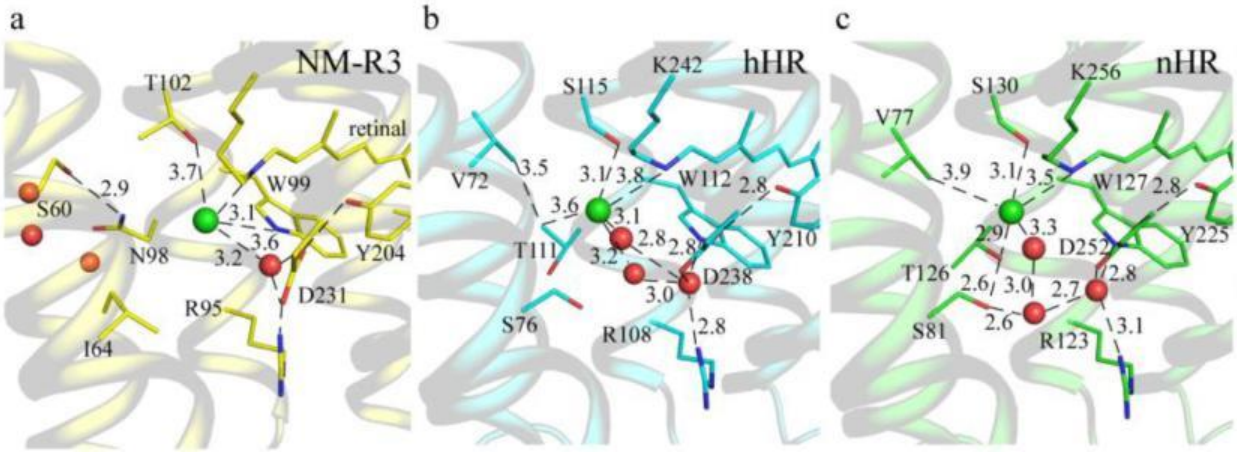

Figure 1.4.3.1. Comparison of structures of RSB region in NM-R3 and archaeal rhodopsins. Structures of (a) NM-R3, *Hs*HR (b) and *Np*HR (c). Image is adapted from [97].

Unlike archaeal halorhodopsins, the active site of *Nm*ClR contains 1 water molecule, not 3. Because of this, chloride forms fewer bonds with surrounding amino acids and water molecules.

At almost the same time, in 2016, another structural work on *Nm*ClR was published[67]. In this work, 5 different structures were obtained with a better resolution of 1.57 Å. Among the structures obtained: the structure of *Nm*ClR at pH 6.0 and 4.5, the structure in a complex with bromide instead of chloride at pH 6.0,

as well as the structure of the two mutants at pH 4.5. These mutants were T102N and T102D. Thr 102 is the second amino acid of the three-letter NTQ motif. In the active center of the protein, it forms a bond with chloride, therefore these mutations strongly affect the pumping of chloride. Also in this work, a second chloride binding site was identified, located in the cytoplasmic part of the protein in the loop between helices A and B (Fig. 1.4.3.2).

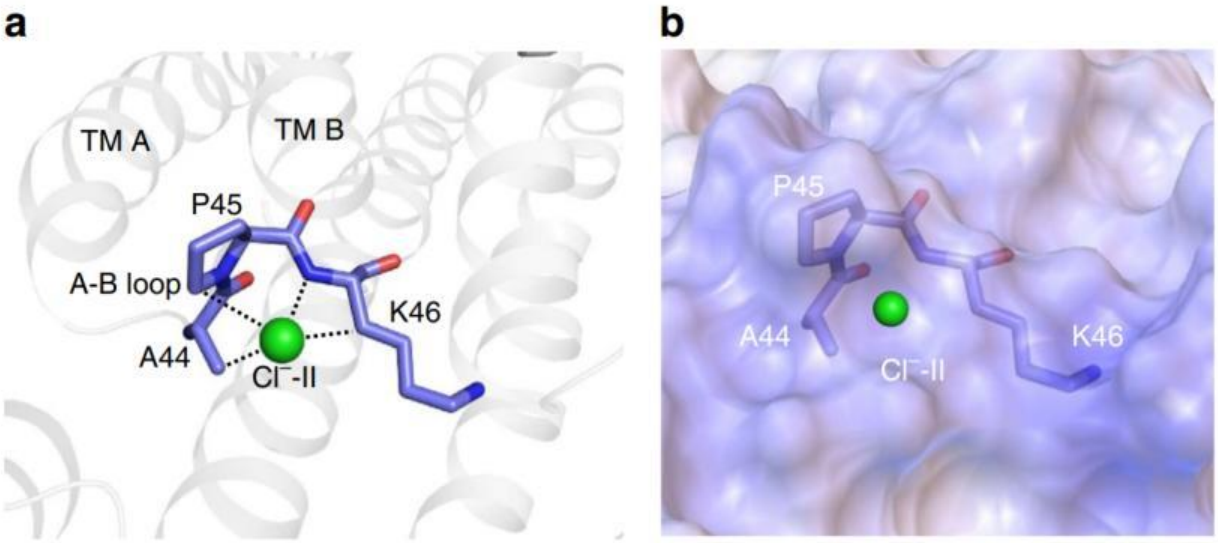

Figure 1.4.3.2. Coordination of the chloride ion by the AB loop. (a) Hydrogen – mediated bonds between chloride and aminoacids of AB loop. (b) Surface electrostatic potential around chloride binding site. Image is adapted from [67].

In 2018, 2 more *NM*-ClR structures were obtained[98]. The authors of this work obtained the structure of the protein at non-cryogenic temperature using XFEL, and for comparison, the structure at cryogenic temperature obtained at the synchrotron. The obtained resolution for the first structure is 1.85 Å, and for the second - 1.75 Å. The obtained structures had slight differences, the main of which

is an additional water molecule bound to the chloride in the active center (Fig. 1.4.3.3).

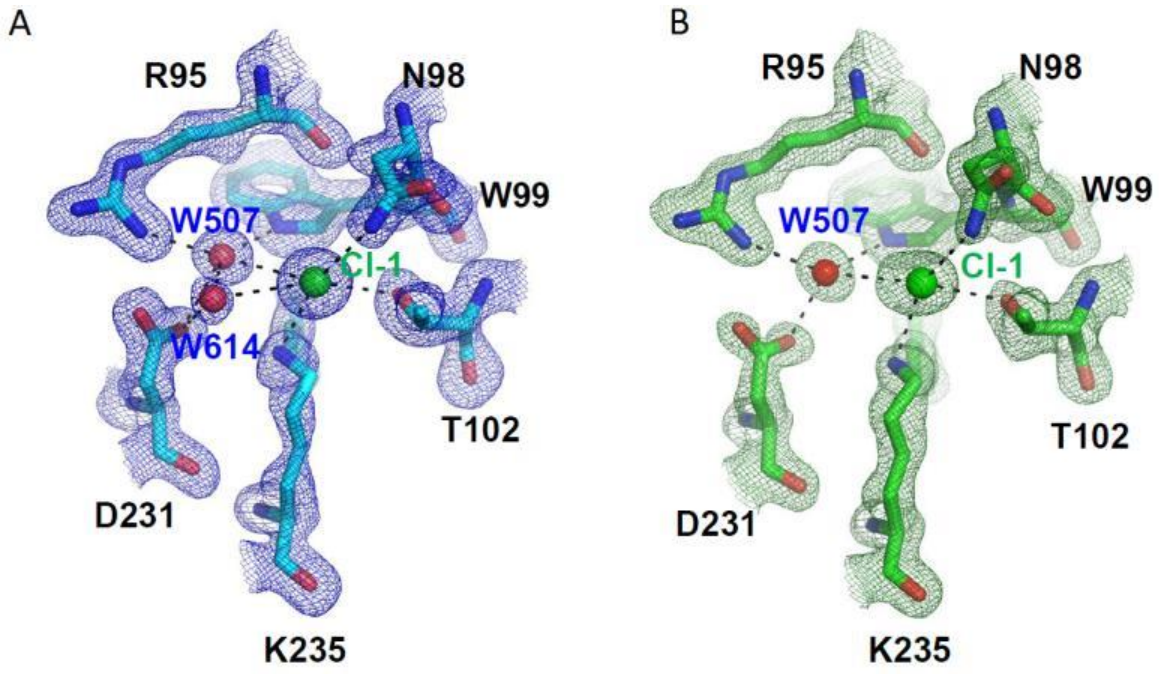

Figure 1.4.3.3 Comparison of active centers of *NM*-ClR. (A) XFEL non-cryogenic structure. (B) Synchrotron cryogenic structure. Image is taken from[98].

In 2020, the same team of researchers published a large-scale structural study containing 10 structures of this protein[99]. The authors obtained structures of early intermediates of *NM*-ClR using cryotrapping using laser pulses of different intensities at a temperature of 95 K or 140 K. The authors obtained a large number of difference maps and suggested that those amino acids whose position changed under different laser illumination conditions play an important role in chloride pumping. Unfortunately, it has not been established to which states in the photocycle the resulting structures correspond.

The same team published another structural paper on NM-ClR in 2021. This work contained 8 high-resolution structures (1.65 Å - 1.85 Å) obtained with the XFEL. In the study, structures were obtained in the time interval from 1 ps to 100 ps from laser excitation of the protein. A series of structures will allow to trace the movement of amino acids and chloride in the active site, depending on time[100].

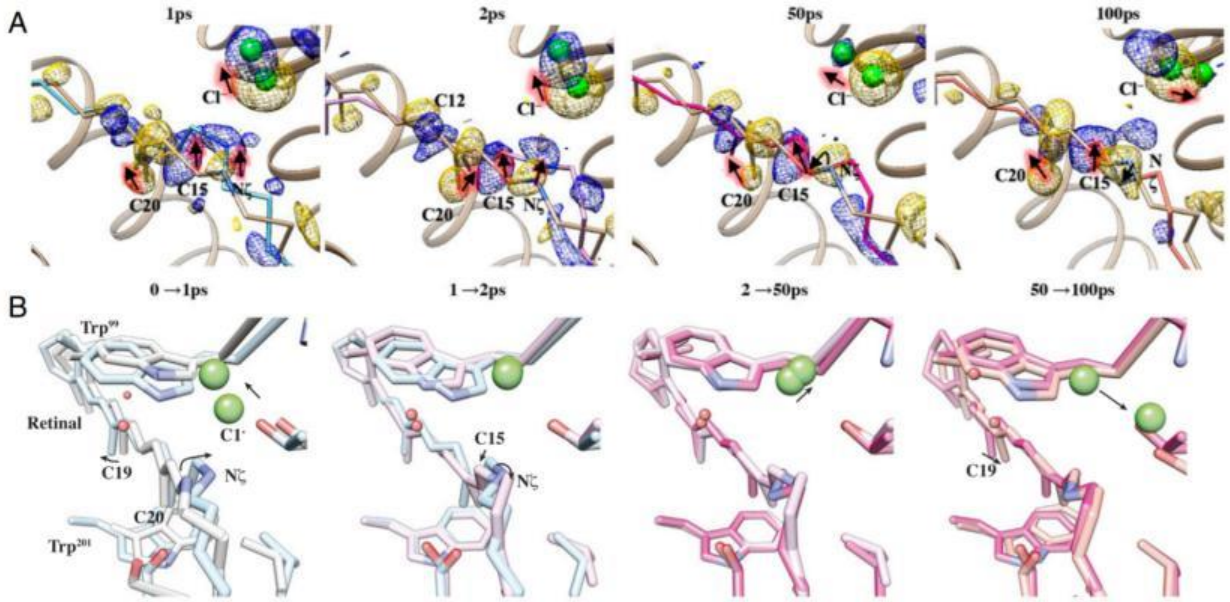

Figure 1.4.3.4 Change in the structure of the active center of the protein depending on time. (A) Difference maps in the vicinity of the retinal. (B) View of conformational changes from the RSB side. Image is taken from [100].

1.4.4 Structures of cyanobacterial anion pumps.

Compared to archaeal halorhodopsins and NTQ-rhodopsins, much less structural information is known about cyanobacterial anion pumps. Until recently,

there was not a single three-dimensional structure of a protein from this family. However, 2 structural studies of the *Mr*HR protein from *Mastigocladopsis repens* were published in 2020.

In the first work, 2 structures of this protein were obtained[101]. The structure of the protein ground state with a resolution of 2.33 Å, as well as the structure of the T74D mutant with a resolution of 2.5 Å. The structure of the wild-type protein showed that *Mr*HR is quite similar to archaeal halorhodopsins. In the active site, chloride is bound to two amino acids of the three-letter motif, just like in archaeal halorhodopsins (Fig. 1.4.4.1, A).

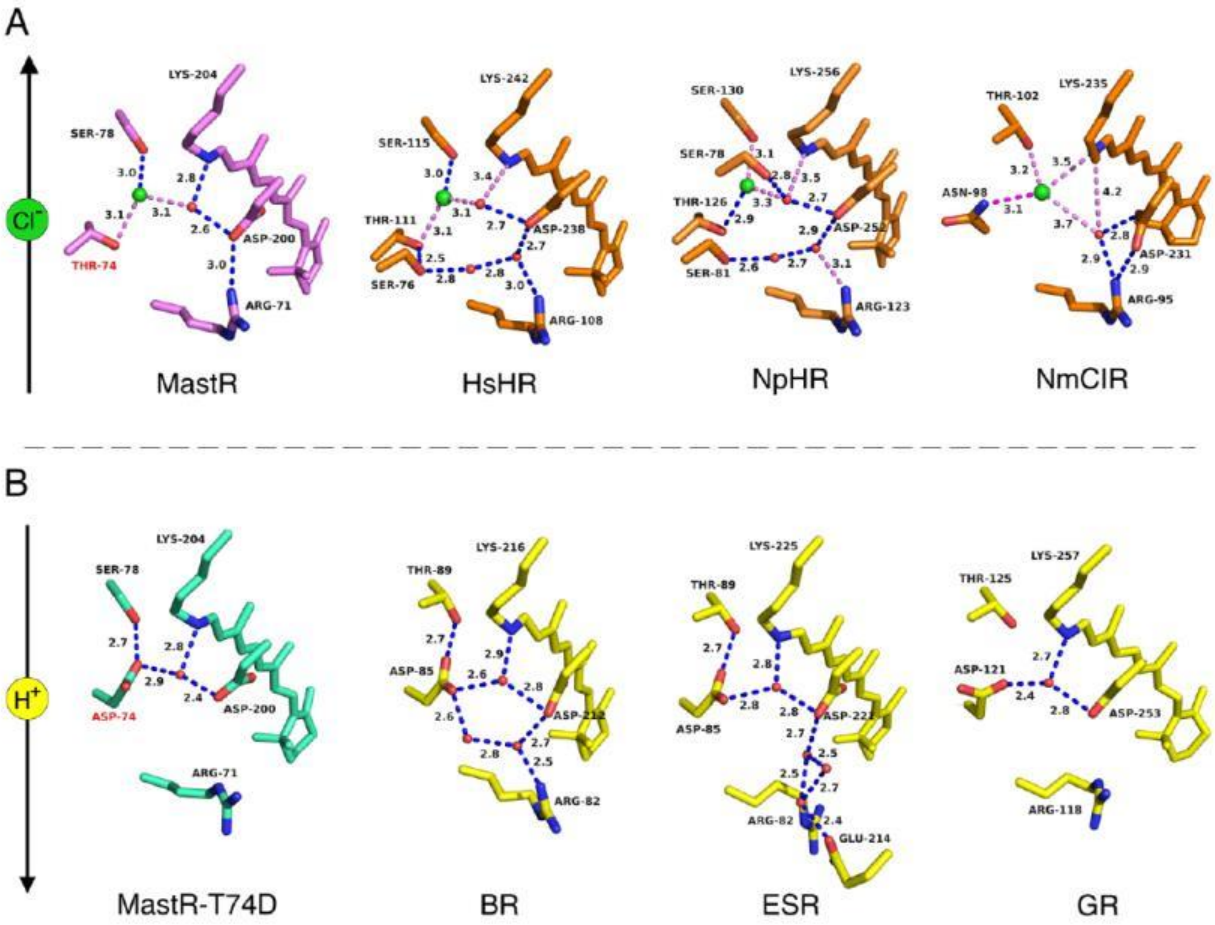

Figure 1.4.4.1 Comparison of structures of active centers of chloride and proton pumping microbial rhodopsin. (A) Comparison of chloride pumping MrHR, HsHR, NpHR and NmClR. (B) Comparison of proton pumping rhodopsins: MrHR T74D, bR, ESR and GR. Image is taken from [101].

The structure of the active site of the T74D mutant resembles a proton pump; in particular, it is very similar to the active site of bacteriorhodopsin. There is no chloride in the active site of this mutant, and there is a water molecule bound to RSB and two aspartates (Fig. 1.4.4.1 B).

A little later in 2020, a second structural paper was published on the MrHR study. This work contained three MrHR structures. The structure of the protein

with chloride was solved with a resolution of 1.9 Å, and the structure of the protein in a complex with bromide with a resolution of 2.5 Å was also obtained. In addition, the authors determined the structure of the double mutant N63A / P118A. These mutations duplicate amino acids in corresponding positions in SyHR, a close relative of the protein under study. A unique feature of the SyHR is its sulphate pumping capability. Making the corresponding double mutation to MrHR also gives him sulfate boost. It is likely that less bulky amino acids open up to form a larger hole in the protein in the area of BC loop and sulfate can enter the protein (Fig. 1.4.4.2).

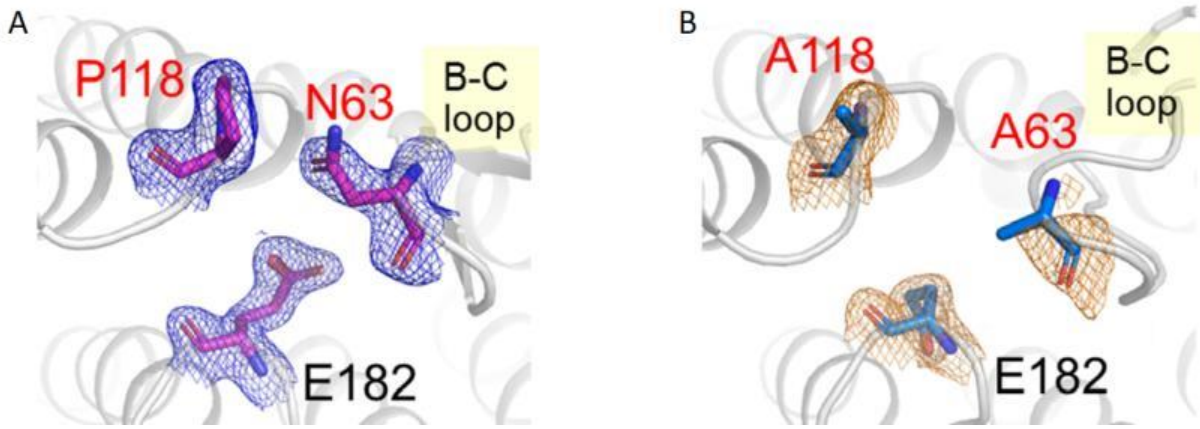

Figure 1.4.4.2 Structure of sulphate entrance region in the wild type and mutant *Mr*HR. (A) Structure of the wild type protein. (B) Structure of N63A/E182A mutant. Image is taken from [72].

2 Materials and Methods

2.1.1 Cloning

Nucleotide sequence of gene of *Sy*HR (Protein accession no. WP_009632765) was optimized for codon composition for *E.coli* protein expression using Thermofisher GeneArt online tool (https://www.thermofisher.com/fr/fr/home/global/forms/geneart-genes-synthesis-requests.html). Gene was ordered as DNA fragment from Eurofins Genomics (Luxembourg).

Amplification of DNA fragments was performed using PCR technique, following the guidelines of the manufacturer. The reaction mixture contained 100 ng of DNA vector, 0.2 mM of each of the four dNTPs, 1 μM of each of the primers, 1 unit of Phusion polymerase, 1/10 of the volume of 10-fold reaction buffer, 3% DMSO, and 10 μM magnesium chloride.

The restriction reaction was carried out using restriction endonucleases FD XbaI, FD XhoI (Thermo, USA) for an hour at 37 ° C in FD buffer. For the reaction with a volume of 20 μL, 10U restriction enzymes and 200 ng of PCR fragment and 1 μg of plasmid. Then a dye for DNA electrophoresis containing SDS was added, the mixture was heated at 65 ° C for 5 minutes and cooled to room temperature. The results were analyzed and purified on an agarose gel.

DNA electrophoresis was performed in agarose gel (1%). Ethidium bromide was added to a concentration of 0.5 μg / ml. Tris-acetate buffer (TAE) was used as a buffer. The dye 6X DNA Gel Loading Dye (Thermo, USA) was added to the studied DNA, the prepared solution was transferred into the wells. GeneRuler 1 kb Plus DNA Ladder (Thermo, USA) was used as a marker. The voltage was 120 V

for 30-40 minutes at room temperature. Extraction of DNA from agarose gel was carried out as follows. Using an ultraviolet lamp, DNA with ethidium bromide intercalated into it was visualized. A piece of gel containing a fragment of the required length was excised, DNA from the agarose gel was purified according to the instructions supplied with the Gel Extraction Kit (QIAGEN, USA).

The DNA insert was inserted into the plasmid after restriction by sticky end ligation. To the ligase mixture with a volume of 20 μl, 1 μl of 100 mM ATP, 1 unit of ligase and fragments in a molar ratio vector: insert = from 3: 1 to 5: 1 were added, the total amount of DNA was about 500 ng. Ligation was performed overnight at 21 ° C.

An aliquot of competent cells, frozen at –80 ° C, prepared according to the high efficiency transformation protocol[102] was thawed on ice, and 50–100 ng of plasmid DNA was added under sterile conditions. The mixture was incubated on ice for 40-50 minutes, optionally shaking was carried out every 10 minutes. Then the mixture was placed in a thermostat at 42 ° C for 1 minute, after which it was incubated on ice for 10 minutes. Under sterile conditions, 700 μL of SOC medium was added to the cells. The mixture was incubated in a shaker at 37 ° C in a horizontal position at 180 rpm within 1 hour. After that, the tube with cells was spun in a centrifuge for 3000g for two minutes. 750 μl of supernatant was removed and the pellet was resuspended in the remaining medium. Then the suspension was placed on an agar petri dish containing the selective antibiotic and incubated overnight at 37 °C.

Isolation of plasmid DNA was performed using one of the kits according to the instructions recommended by the manufacturers: MiniPrep Kit (QiaGen, USA). Colonies were taken from the plates with a sterile loop and placed in 10-15 ml of

LB medium containing the antibiotic. The culture was incubated at 37 ° C for 14-18 hours.

2.1.2 Protein expression and purification

For the expression of *Sy*HR rhodopsin, cells of *E. coli* strain C41 (DE3) (Lucigen, USA) were transformed with the *Sy*HR expression vector. Transformed cells were grown at 37°C in the shaking incubator (Infors HT, Switzerland) at 180 rpm in autoinducing medium called ZYP-5052[103] containing 100 mg/ml of kanamycin. Induction of the expression was carried out upon reaching the optical density OD_{600} of 0.8-1.0 with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and 10 μM all-trans-retinal. 3 hours after the induction cells were collected using centrifugation at 4500g for 10 minutes.

The centrifuged cells were disrupted with an M-110P Lab Homogenizer (Microfluidics, USA) at a pressure of 25000 psi in a buffer containing 20 mM Tris-HCl pH 8.0, 200 mM NaCl (0 mM NaCl fo *Sy*HR^{v}), 5% glycerol, 0.25% Triton X-100 (Sigma-Aldrich, USA), 0.6 mM PMSF (phenylmethane sulfonyl fluoride, Sigma-Aldrich, USA), and 50 mg/L DNase I (Sigma-Aldrich, USA). The membrane fraction was obtained by ultracentrifugation at 100000g for 1 hour at 4°C and collecting the pellet. From the obtained membranes, the protein was solubilized as follows. The membranes were resuspended in a buffer containing 50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.2 M NaCl (50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0 mM NaCl fo *Sy*HR^{v}), and 1% n-dodecyl β-D-maltoside (DDM) (Anatrace, Affymetrix, USA) and stirred at 4 °C overnight. The non-solubilized fraction was

separated by ultracentrifugation at 100000g for 1 hour at 4 ° C, and the supernatant was applied to a Ni-NTA column (Qiagen, Germany).

Protein was eluted with a buffer containing 50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.2 M NaCl (50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.0 M NaCl fo *Sy*HR^{v}), 0.3 M imidazole, and 0.1% DDM. Then the size-exclusion chromatography on 24 ml Superdex 200i column (GE Healthcare Life Sciences, USA) were performed with the eluate. Buffer for SEC contained 50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.2 M NaCl (50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.0 M NaCl fo *Sy*HR^{v}) and 0.05% DDM. Protein fractions with the minimum absorption ratio A280/A525 were selected for further analysis and crystallization, concentrated to 40 mg/ml and were frozen for long-term storage.

2.1.3 Protein crystallization and crystal preparation

Protein was crystallized using *in meso*[12] method according to protocol described in ref.[37] with some modifications. Frozen protein samples were at defrosted at ice bath. Later frozen samples were diluted to 20 mg/ml with buffer containing 50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.2 M NaCl (50 mM NaH_2PO_4 / Na_2HPO_4 pH 7.5, 0.0 M NaCl fo *Sy*HR^{v}). The cubic phase was obtained as follows. Monoolein lipid (Nu-Chek Prep, Elysian, USA) was heated to 42 °C and mixed with diluted protein in the ratio 3:2. Mixing of protein and lipid to obtain a cubic phase was carried out using two syringes connected by a coupler (Hamilton, USA)[80]. Then, using an NT8 crystallization robot (Formulatrix, USA), 100 nL of mesophase was dispensed into each well of a 96-well glass from the LCP sandwich plate set (Marienfeld, Germany) and the resulting drops were coated with 500 nL of precipitant. The obtained plate was covered with a coverslip on top. Then the

plates were incubated at 22 °C. Crystals grow to their maximum size within 2-3 months. Best crystals of red form of *Sy*HR were grown at 1.6 M Ammonium Phosphate pH 4.6 precipitant solution and the best crystals of *Sy*HR^{v} were obtained in precipitant solution of 2.0 M Ammonium Sulfate, 0.1 M HEPES pH 7.0.

Crystals with the maximum size and the smoothest shape were harvested as described at ref.[104]. Crystallization drop was opened and covered with the following precipitant solution to avoid dehydratation: 1.6 M Ammonium Phosphate pH 4.6, 10% v/v glycerol for *Sy*HR and 2.0 M Ammonium Sulfate, 0.1 M HEPES pH 7.0, 15% v/v glycerol for *Sy*HR^{v}. Then the crystals were harvested using micromounts (MiTeGen, USA) and were flash-cooled in liquid nitrogen for further X-ray or spectroscopic experiments.

2.1.4 Cryotrapping of active states and its spectroscopic characterization

Accumulation of active states in crystals and its spectroscopic characterization was performed at icOS Lab located at ESRF[105]. Crystals in micromounts were placed under the cryostream with 100K temperature. First, the spectrum of crystals in the ground state was measured using a QE65 Pro spectrometer (Ocean Optics, USA). Active states have been accumulated with 532-nm green laser (CNI Laser, Changchun, P.R. China). Power density of 7.5 mW/cm^2 was used.

The following protocol was used to obtain the O-state. A crystal installed in a goniometer under a cryo stream at 100 K temperature was illuminated by a green laser. After that, the cryo stream was blocked for 2 seconds to thaw the crystal. After two seconds, the cryo stream returned, and then the laser was turned off.

From the crystals obtained with the accumulated O-state, the spectrum was recorded.

The K-state was obtained as follows. Crystal which was located at cryo stream with 100 K temperature was illuminated with green laser for 1 minute. Such illumination resulted in a stable accumulation of red-shifted intermediates in the crystal. After illumination the spectrum was recorded from these crystals.

2.1.5 X-ray crystallography data collection

X-ray diffraction data from *Sy*HR crystals was obtained at the P14 beamline of PETRAIII synchrotron, DESY, Hamburg, Germany. Data collection was performed at 100 K using EIGER 16 M detector (Dectris, Baden-Daettwil, Switzerland). To obtain active states in the crystal, the protocol described in the previous paragraph was used. The same green laser and power settings were used.

2.1.6 Data refinement and structure determintaion

The data were integrated using the XDS[106] software package, scaled and combined with AIMLESS from the CCP4 suite[107]. Initial phases was obtained using molecular replacement method using MOLREP suite[108]. Reference model (ESR, PDB 4HYJ) for the molecular replacement was chosen with the RaptorX Structure Prediction Server[109]. The initial model was iteratively refined using REFMAC5[110], PHENIX and Coot[111,112]. The cavities were calculated using HOLLOW[113]. Hydrophobic-hydrophilic boundaries of the membrane were calculated using PPM server[114].

3 Results

3.1 Expression and purification of *Sy*HR protein

3.1.1 SyHR gene optimization.

Protein accession number of SyHR sequence (Protein accession no. WP_009632765) was taken from the publication[71]. Then the sequence was *E.coli* optimized and was supplemented with C-terminal 8-His tag. After this modified sequence was commercially ordered and inserted into a pBKT7. This plasmid is often used in our lab for microbial rhodopsin expression[115,116].

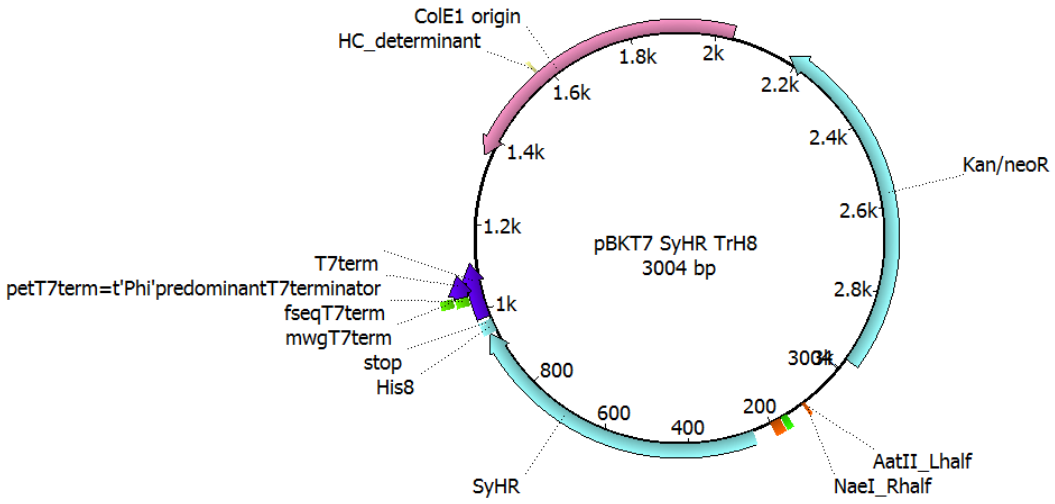

Figure 3.1.1.1 Plasmid map of *Sy*HR expression plasmid.

3.1.2 Optimization of expression of *Sy*HR

All tests of expression were performed using *E.coli* C41 strain. This strain is often used in our laboratory to work with microbial rhodopsins[117,118]. In order to increase the protein yield, various expression induction options were tested. In all variants, induction was carried out at OD_{600} of the cells equal to 0.9. Two main approaches were tested: induction for 4 hours at 37 ° C and overnight expression at 20 ° C. Within these approaches, various expression medium formulations and various expression induction additives were tested.

Number on sample (In the same order were applied to the blot)	Composition of the expression medium	Expression Induction Supplement
Expression 4 hours at 37 ° C		
1	Culture medium (1% yeast extract + 1% tryptone + 5% NPS + 0.1% Mg_2SO_4 + 0.1% Kan) + 0.05% glucose+0,5% glycerol	-
2	Culture medium + 0.05% glucose+0,5% glycerol	0.5 mM IPTG
3	Culture medium + 0.05% glucose+0,5% glycerol	1 mM IPTG
4	Culture medium + 0.05% glucose + 0,5% glycerol +	0.2 mM IPTG

	0.2% α-lactose	
5	Culture medium + 0.05% glucose + 0,5% glycerol + 0.2% α-lactose	0.5 mM IPTG
6	Culture medium + 0.05% glucose + 0,5% glycerol + 0.2% α-lactose	1 mM IPTG
11	Culture medium + 0,5% glycerol + 0.1% α-lactose	
Expression overnight at 20 ° C		
7	Culture medium + 0.05% glucose + 0,5% glycerol + 0.1% α-lactose	
8	Culture medium + 0.05% glucose + 0,5% glycerol + 0.2% α-lactose	
9	Culture medium + 0.05% glucose + 0,5% glycerol + 0.02% α-lactose	
10	Culture medium + 0.05% glucose + 0,5% glycerol	1 mM IPTG

Table 3.1.2.1 Optimization of the expression protocol of *Sy*HR

Expression efficiency was tested using a Western-blot. The amount of expressed protein was estimated using Anti-HIS antibodies. The amount of

expressed protein was determined in relative amount, that is, in terms of 1 gram of cells in the test sample. The results obtained show that condition number 3 shows the highest protein yield per 1 gram of cells (Fig. 3.1.2.1). Therefore, later, precisely such expression conditions were used, namely: expression with the following Culture medium: 1% yeast extract + 1% tryptone + 5% NPS + 0.1% Mg_2SO_4 + 0.1% kanamicyn + 0.05% glucose + 0,5% glycerol and induction at OD_{600} = 0.9. Induction is performed using 1 mM IPTG and and expression is carried out 4 hours after induction at 37 ° C.

Figure 3.1.2.1 Western-blot image of *Sy*HR expression test. Condition with number 3 showed the highest yield of the protein per 1 g of cells.

For the expression of *Sy*HR in chloride-free form the similar protocol of the expression was used. To obtain protein in the chloride-free form, changes were made at the protein purification stage and in subsequent stages.

3.1.3 Purification of *Sy*HR

After expression, purification was carried out according to the standard purification protocol for our laboratory for membrane proteins[119]. The cells containing the protein were lysed using a french press, the resulting membranes were solubilized in DDM. The protein was then collected using Ni-NTA chromatography and then purified using size-exclusion chromatography. For chloride bound protein, the buffers described in the Methods section were used. In addition to other components, these buffers contained 200 mM NaCl.

In order to purify the protein in chloride-free form, the exact same purification protocol was used, however, NaCl was removed from all buffers. We also made sure that HCl was not used to adjust the pH of other solutions.

As a result of purification with chloride, a pink protein was obtained, and when purification without chloride - violet, therefore, in what follows we will call the chloride-free form of the protein - *Sy*HR^{v}.

In addition to the fact that different forms of protein had different colors, they had a different profile of the column passage with Size exclusion chromatography. The SEC tested standard protein isolation with 200 mM NaCl as well as isolation without NaCl. In addition, we tested the protein without NaCl, to

which 200 mM Na_2SO_4 was added (Figure 3.1.3.1).

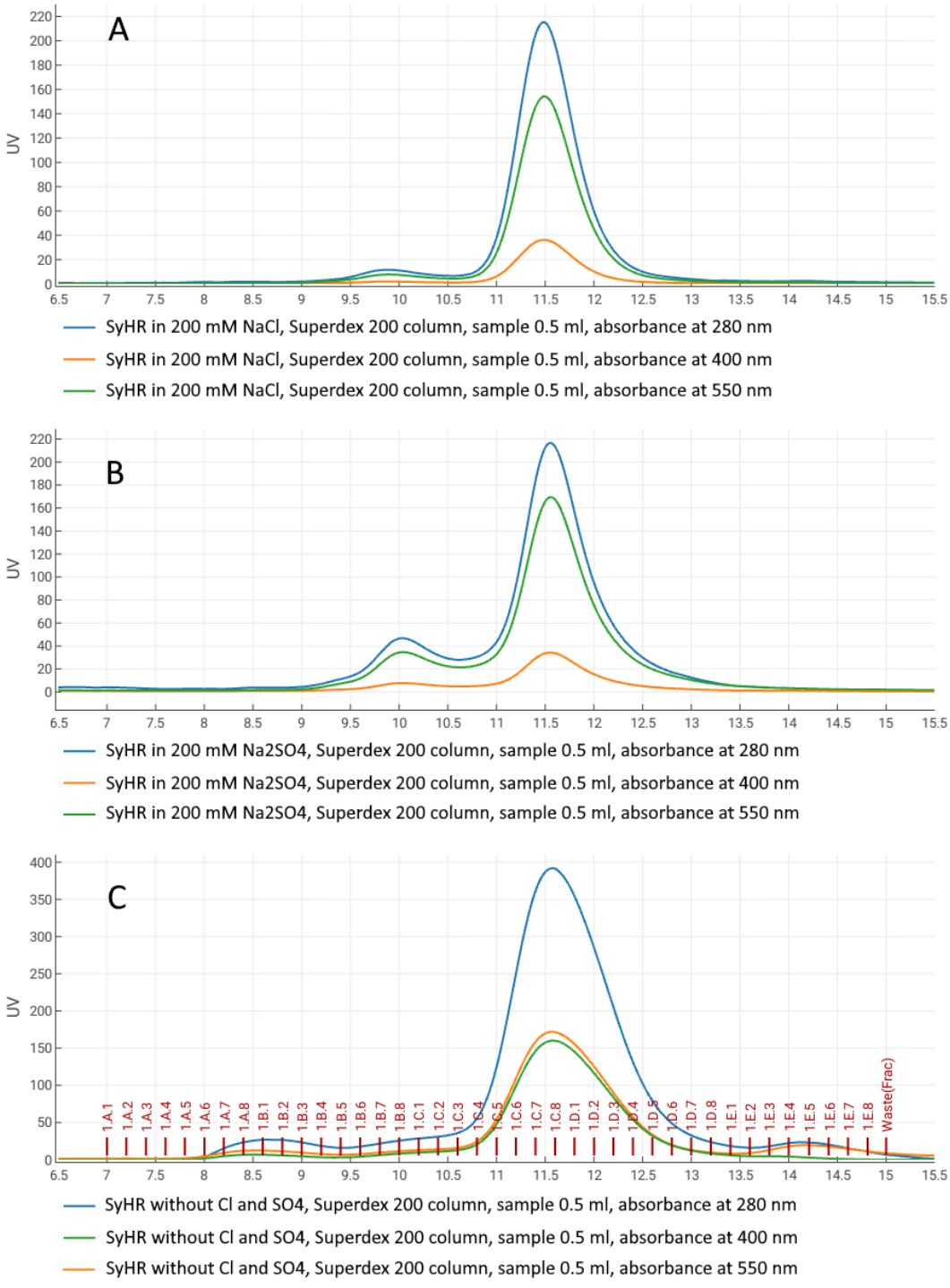

Figure 3.1.3.1. Different SEC profiles of *Sy*HR in presence of diffenrent ions. (A) SEC plot of *Sy*HR + 200 mM NaCl at 280, 400, and 550 nm. (B) SEC plot of *Sy*HR + 200 mM Na_2SO_4 at 280, 400, and 550 nm. (C) SEC plot of *Sy*HR without chloride and sulphate at 280, 400, and 550 nm. Milliliters are plotted on the X-axis

For all three cases, the maximum protein yield from the column is observed at 11.5 ml, which is approximately what, according to the experience of our laboratory, approximately corresponds to the trimers of microbial rhodopsins on this column.

For a protein with NaCl, the observed peak is the narrowest, which may indicate that the protein is the most stable in NaCl out of these three conditions.

In the protein with Na_2SO_4, a peak at 10 ml is observed, which indicates the aggregation of the protein in Na_2SO_4.

Finally, for protein without chloride and without sulfate, the peak appears to be the widest. This may indicate a lower stability of the protein or that this protein forms fewer trimers and more other forms, such as, for example, monomers or multimers.

3.1.4 Spectral characterization of *Sy*HR

To measure the effect of changing the absorption maximum depending on chloride or sulfate, the absorption spectrum of the protein was measured using a spectrophotometer in the presence of 200 mM NaCl or 200 mM Na_2SO_4 (Fig. 1.3.4.1).

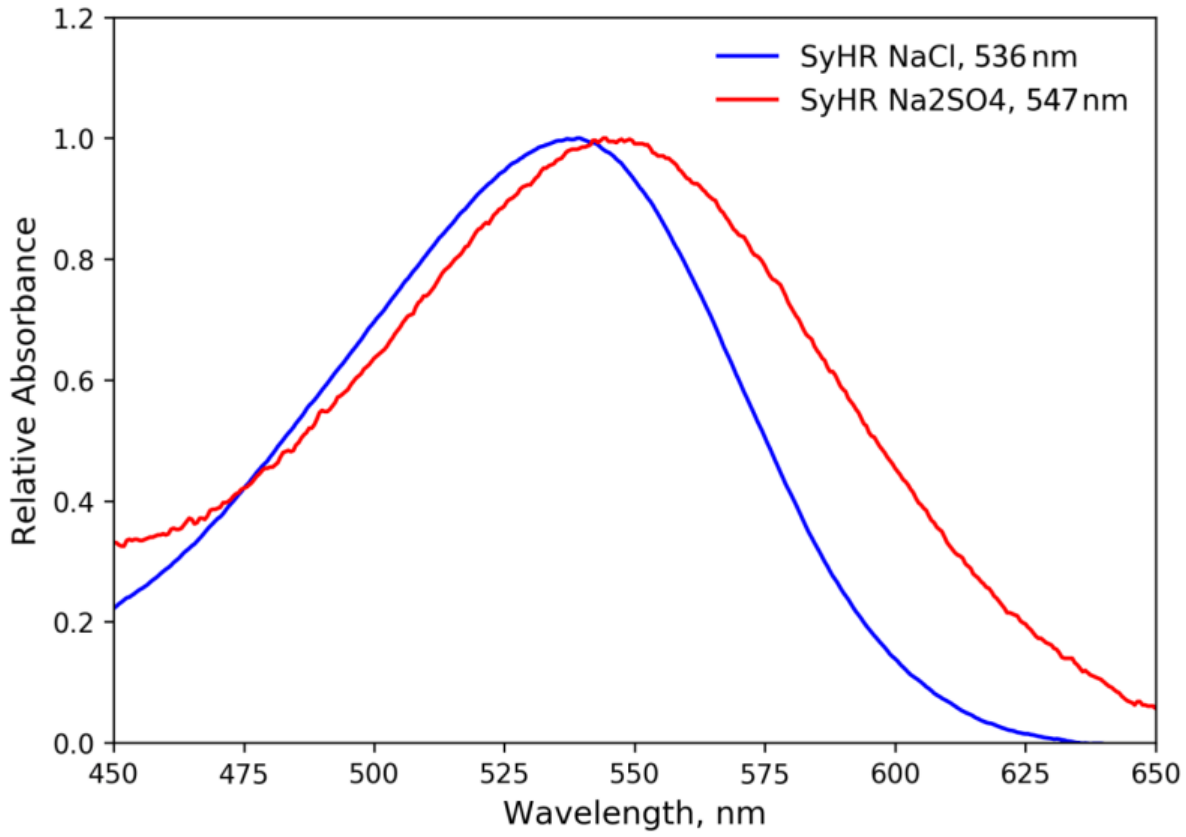

Figure 1.3.4.1. Absorbance maximum of SyHR in presence of different ions. Measurements were performed at pH 6.0. Maximum absorbance for NaCl-bound *Sy*HR was detected at 536 nm and for Na_2SO_4-bound was detected at 547 nm.

It is known that the UV-Vis spectrum of *Sy*HR is highly dependent on the pH of the buffer[71]. At alcaline pH, one more shoulder appears in the spectrum near 400 nm wavelength (Fig. 3.1.4.1). It is supposed that this shift is caused by the deprotonation of the Retinal Schiff base.

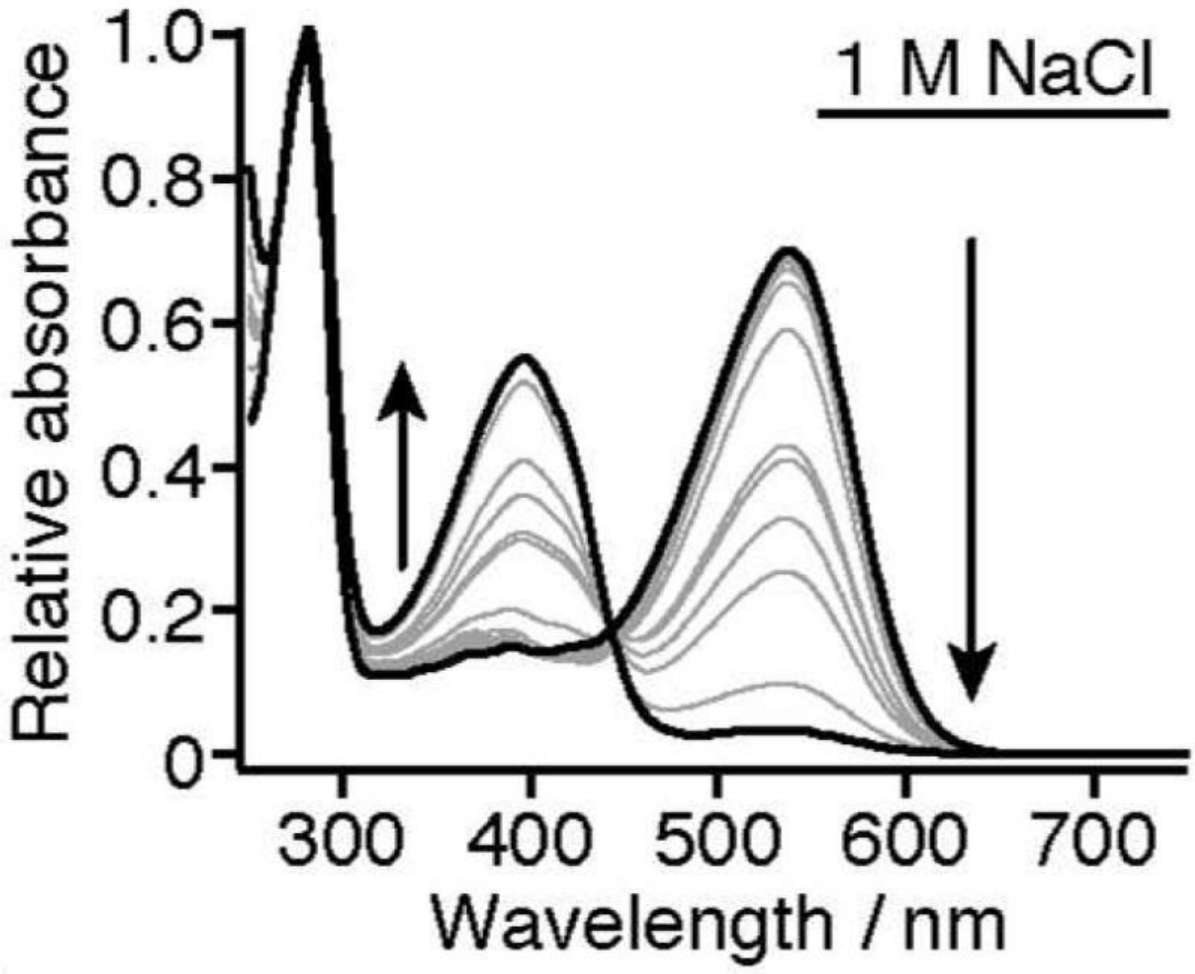

Figure 1.3.4.2. Schematic representation of the change in the spectrum of chloride-bound *Sy*HR depending on the pH. With more alkaline pH, the peak increases in the region of 400 nm, while with more acidic pH, the peak increases at 536 nm. Figure is adapted from[71].

Our goal was to establish the dependence of the spectrum on pH and sulfate concentration for sulphate-bound form of the protein. For this, the spectra of the protein were studied for different pH values at 100 mM Na_2SO_4, as well as spectra at different concentrations of Na_2SO_4 and pH 6.8 (Figure 1.3.4.3).

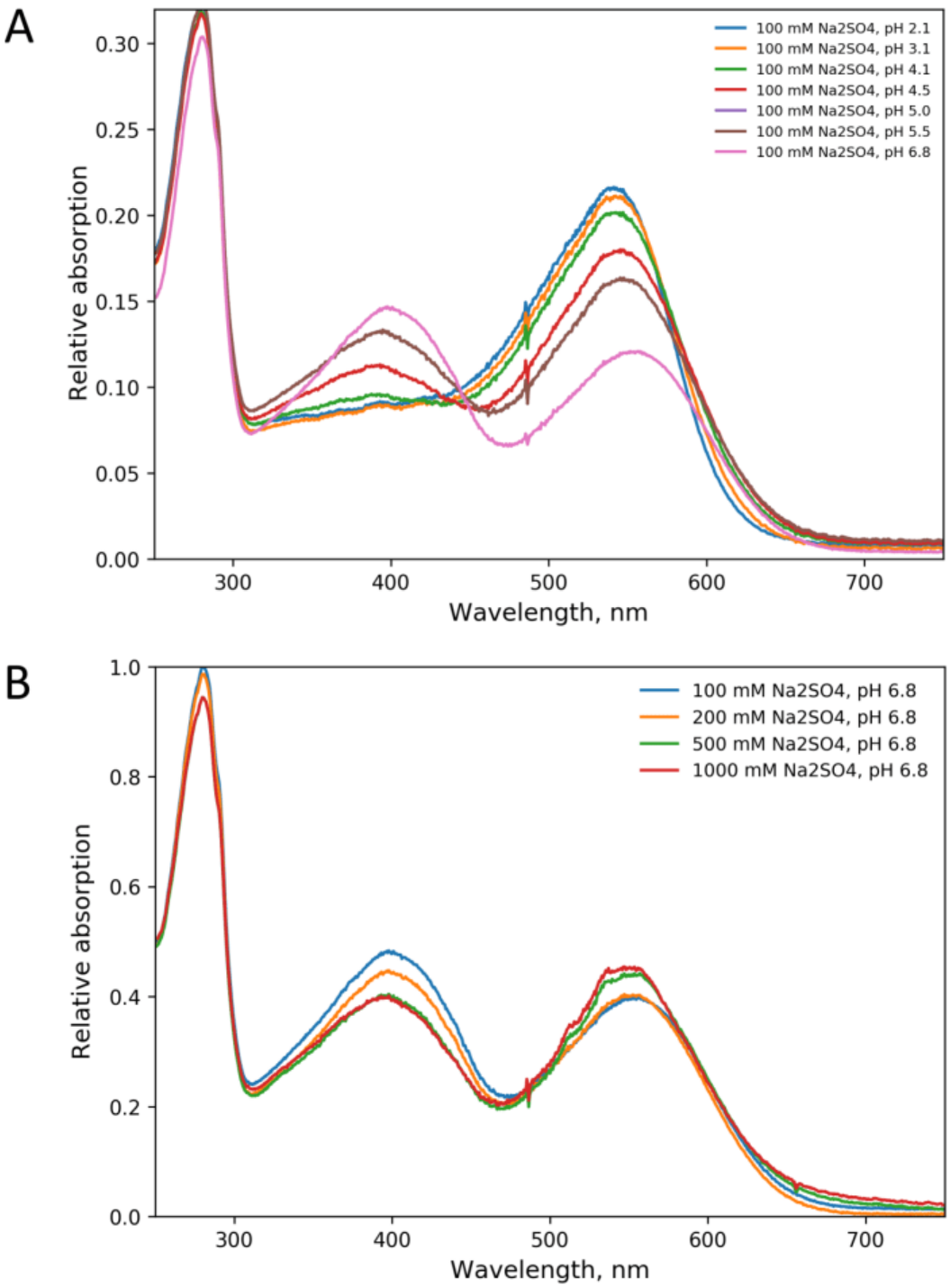

Figure 1.3.4.3. Absorbance spectra of *Sy*HR in presence of Na_2SO_4 in different pH and different concentrations of sulphate. (A) Series of spectra at titration pH from 2.1 to 6.8. (B) Series of spectra at titration of concentration of Na_2SO_4 from 100 mM to 1000 mM.

These spectra showed that the spectra of the protein bound to sulfate are highly pH dependent. As in the protein with chloride, at alkaline pH a peak appears around 400 nm. In addition, the ratio of the peaks also depends on the concentration of sulfate. The more sulfate, the more the ratio will move towards the peak around 550 nm.

3.1.5 Protein crystallization

Obtained protein was crystallized using *in meso* crystallization method according to protocol described in §2.2.4. For the initial crystallization of Cl-bound *Sy*HR commercial screens Cubic Phase I and Cubic Phase II (Quigen, Germany were used). As a result, the best crystals grew in a condition containing 1.6 M Ammonium Phosphate pH 4.6. For crystallization of Cl-free *Sy*HR, custom screens were prepared based on Quigen Cubic Phase I and Cubic Phase II, with the following change. All conditions in which chloride was contained were prepared without chloride. Best crystals for the Cl-free protein were grown at 2.0 M Ammonium Sulfate, 0.1 M HEPES pH 7.0 precipitant solution. For both Cl-bond and Cl-free *Sy*HR crystals appeared after 1-2 weeks after crystallization and reached their maximal size within 2-3 months. In accordance with the fact that the absorption spectrum for protein Cl-bound and Cl-free is different, the crystals also turned out to be of different colors. Crystals of Cl-bound protein had pink color, whereas crystals of Cl-free protein were violet.

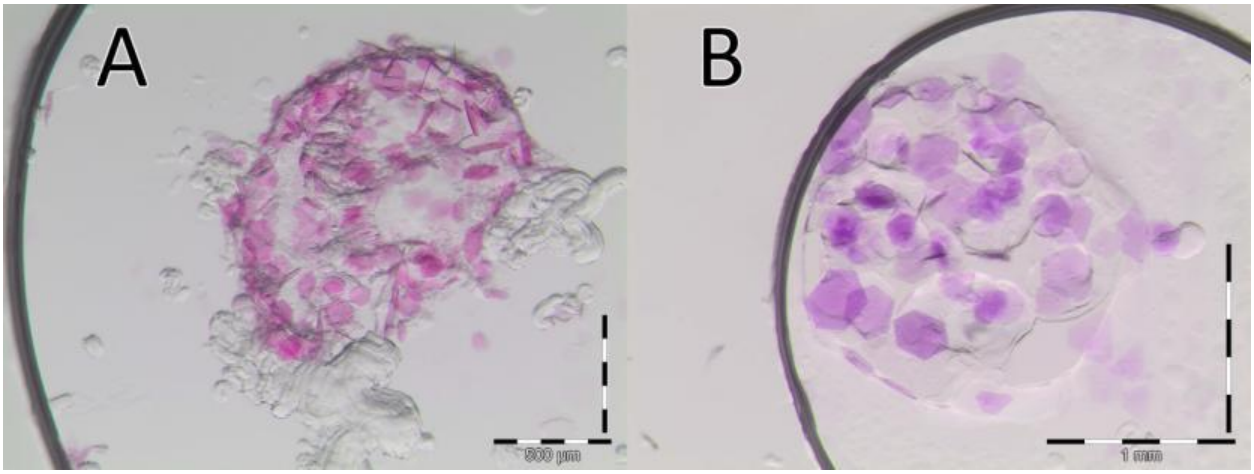

Figure 3.1.5. Microscope image of *Sy*HR crystals. (A) Red crystals of Cl-bound *Sy*HR. (B) Violet crystals of sulphate-bound *Sy*HR.

An important point is that despite all attempts at crystallization, crystals of Cl-free protein failed to obtain crystals under conditions not containing sulfate or chloride. This may be due to the fact that the protein is less stable in the absence of these ions and cannot form a crystal. This assumption is supported by SEC-profile data showing the broadest protein profile without the presence of chloride or sulfate.

3.1.6 Spectroscopy studies of crystals

The obtained protein crystals in Cl-bound form were pink, and in SO4-bound they were violet. Obviously, the absorption spectrum of the protein in these crystals was different. In order to understand how it differs, we measured the absorption spectra of these crystals on the icOS protein crystal spectroscopy unit at

ESRF (Grenoble). The data obtained confirmed the difference in the spectra of these crystals (Fig. 3.1.6.1).

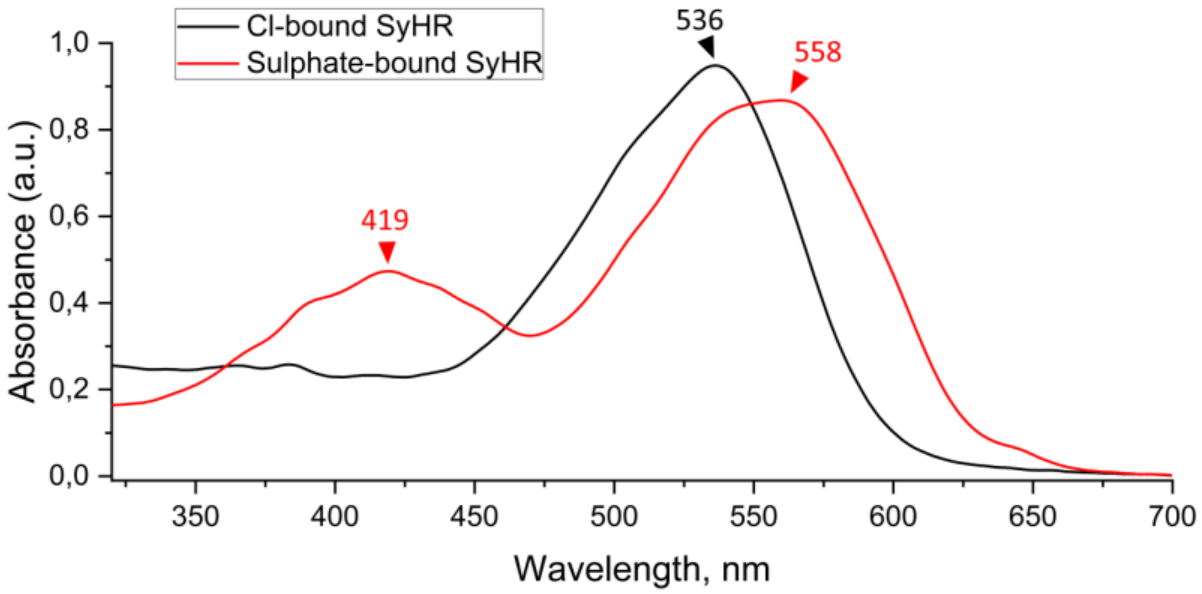

Figure 3.1.6.1. Spectra of red and violet *Sy*HR crystals. Chloride-bound protein has absorbtion maximum at 536 nm (black) and sulphate-bound *Sy*HR has maximum at 558 nm.

The spectra taken from the crystals show that the absorption maximum of the crystals of sulfate-bound protein is red-shifted. The maximum is observed at 558 nm. In addition, since these crystals were obtained from pH 7.0, a small peak at 420 nm is visible. The spectrum of Cl-bound protein crystals has a maximum at 536 nm, which is standard for Cl-bound *Sy*HR. This spectrum has peak at 420 as the crystals are obtained at pH 4.6.

3.1.7 Cryotrapping of crystals and spectroscopic validation

High-quality red crystals of *Sy*HR were used for the cryotrapping of the intermediate states of the Cl^--pumping mode. For that, we used an approach, similar to our previous works[115]. With violet crystals, due to their worse diffraction quality, we were not successful in solving structure of active states.

We validated the cryotrapped intermediates using microspectrophotometry of the *Sy*HR in crystals at the *ic*OS station of the European Synchrotron Radiation Facility (ESRF, Grenoble, France)[105].

The analysis of this structural data together with the previously reported data on *Sy*HR and other anion-pumping rhodopsins allowed us to provide a model of the observed structural rearrangements to the O and K intermediates of the Cl^--pumping mode of the *Sy*HR photocycle in solution.

First, we worked on cryotrapping of the late O intermediate state of the Cl^--pumping mode of the *Sy*HR photocycle. Since the O state is the dominant intermediate of the photocycle, it could be accumulated upon continuous green light illumination at room temperature. For the accumulation and trapping of the O state, we applied the following procedure. When the crystal was in the cryostream, we blocked the cryostream for 2 seconds while illuminating the crystals with a 532-nm laser. Then the cryobeam was returned. Microspectrophotometry indicated appearance of the red-shifted state fraction corresponding to presence of the O-state in the *Sy*HR crystals after the procedure (Fig. 3.1.7.1, B).

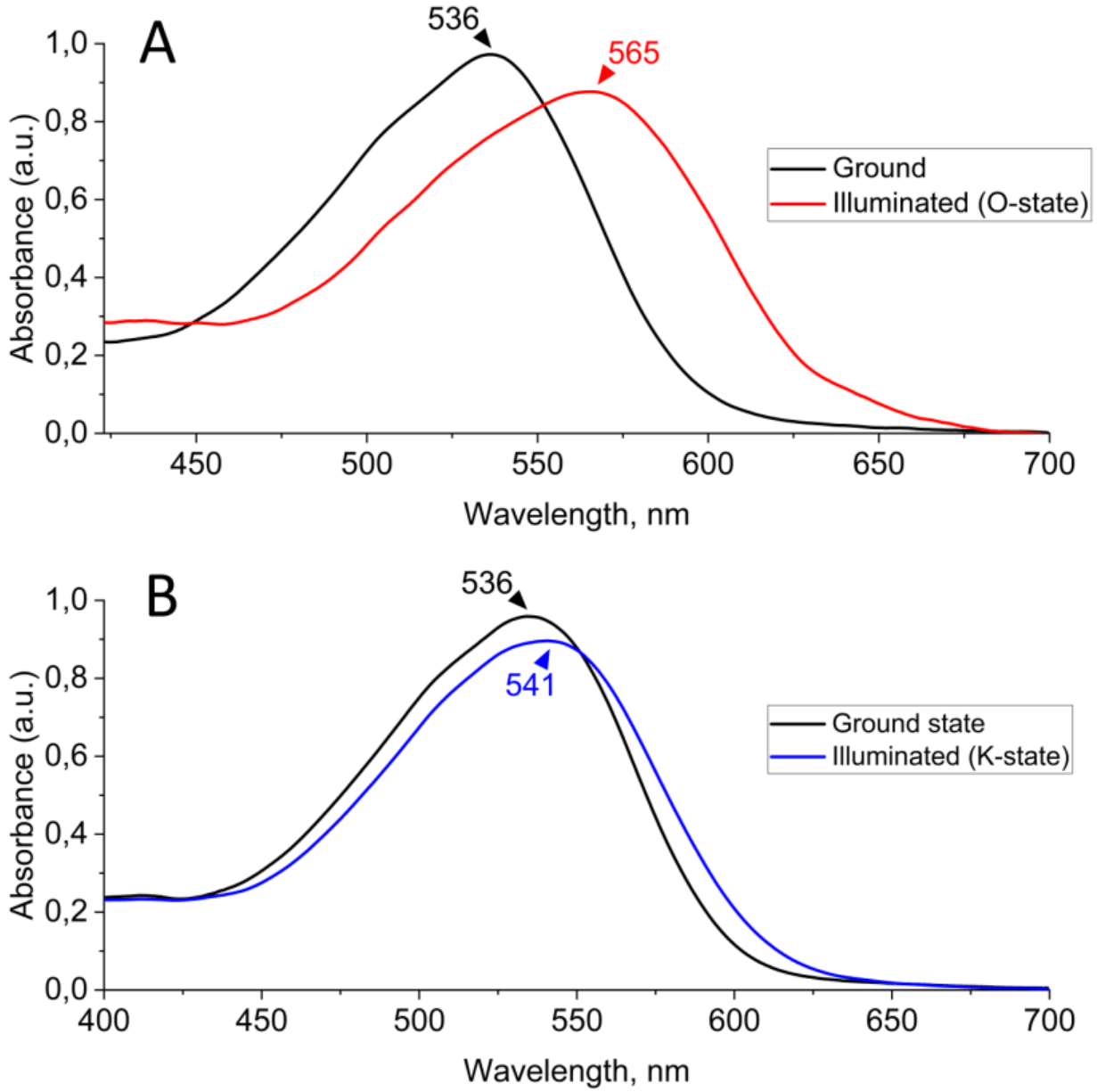

Figure 3.1.7.1. Changes in *Sy*HR crystals spectra after cryotrapping. (A) Spectral changes of 2 second illumination of crystals in room temperature (O - state). (B) Spectral changes after illumination of SyHR crystals for 1 min at 100 K with green laser (potential K-state).

The following structural analysis indicated the absence of the Cl^- ions inside the protomer in the corresponding fraction, which is expected for the O state of *Sy*HR. Therefore, the accumulated state was tentatively assigned to the O state of the rhodopsin

Second, we aimed at accumulation and cryotrapping of the K state of the *Sy*HR photocycle in red crystals. It is known that in the K state the structural rearrangements in all studied MRs are small; typically, only the retinal and a few surrounding residues are slightly affected. Such rearrangements may occur at cryogenic temperatures. For instance, in BR, the K state could be formed at 100 K[120]. In KR2, the K state was obtained at an even lower temperature of 77 K[121]. In this work we accumulated trapped red-shifted state by illuminating *Sy*HR crystals with a laser at 532 nm for 1 minute at 100 K (Fig. 3.1.7.1, B). As it will be described later in the manuscript, the structural analysis showed that the rearrangements associated with the formation of the accumulated state only occur in the retinal cofactor and two surrounding residues. Thus, taking into account the low temperature of trapping, red-shift of the absorption maximum, and only local corresponding structural changes in the *Sy*HR molecule, we tentatively assigned the accumulated state to the K state of the protein photocycle.

Importantly, in a test experiment, after the cryotrapping of both intermediates, the crystals were heated up to room temperature without laser illumination by blocking the cryostream for 2 seconds. This was done for the additional validation of the reversibility of *Sy*HR in crystals. Indeed, immediately after defreezing the crystal contained a fraction of the proteins in the cryotrapped intermediate, the latter relaxed fast to the ground state, as demonstrated by the spectrophotometry (Fig. 3.1.7.2). This fact indicates that the observed red shifts of the maximum absorption wavelengths in the cases of the cryotrapping of the K and

O states are not associated with non-reversible changes in *Sy*HR, such as, for instance, protein crystals dehydration.

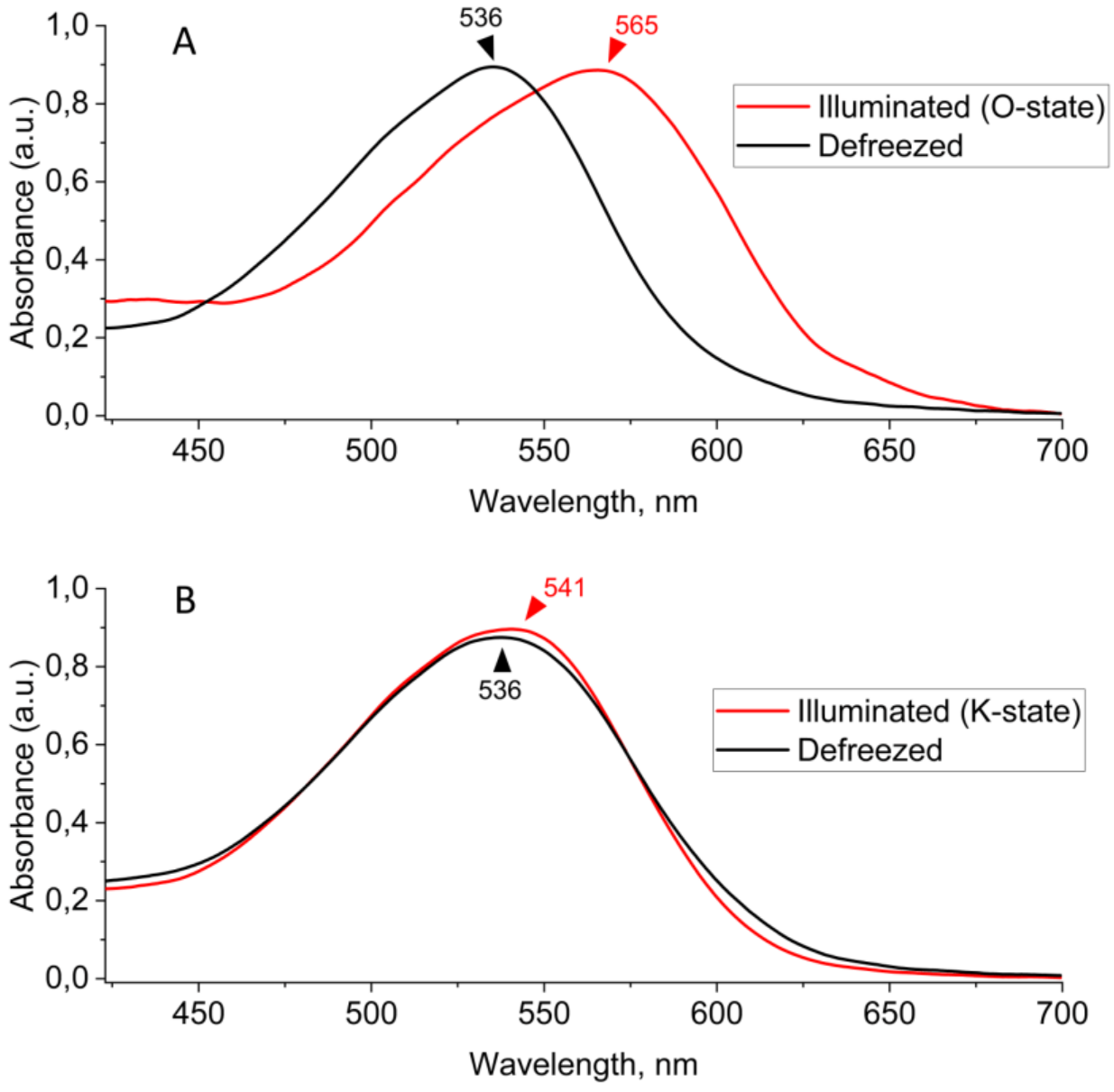

Figure 3.1.7.2. Relaxation of active states of *Sy*HR after annealing. (A) Relaxation of the O-state of the protein. (B) Relaxation of the potential K-state of *Sy*HR. Both spectra returned to the 536 nm absorbance maximum as it was before the illumination.

3.1.8 Structure determination

Using X-ray crystallography, 4 different structures of *Sy*HR were obtained. Red crystals without illumination were used to obtain the structure of Cl-bound *Sy*HR. Sulphate-bound or *Sy*HR^{v} structure was obtained using the violet crystals.

Using the same cryotrapping scheme that was used on icOS, we managed to obtain two structures of the active state Cl-bound *Sy*HR. These are the structures of the O-state and the red-shifted early intermediate, presumably the K-state.

All crystals were of P 3 2 1 symmetry and contained one molecule of *Sy*HR in the asymmetric unit. Analysis of the crystal packing indicated that *Sy*HR is in its trimeric form in the crystals (Figure 3.1.8.1)

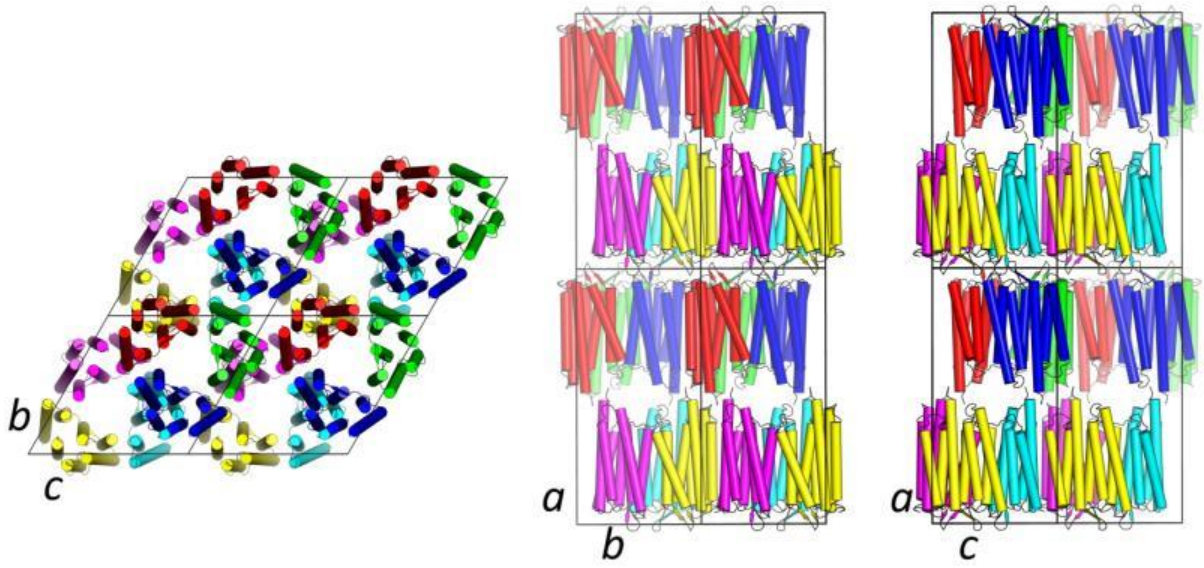

Figure 3.1.8.1. Crystal packing of *Sy*HR in three projections.

Structure of Cl-bound ground state of SyHR was solved at 1.57 Å resolution, structure of sulphate-bound form of SyHR was solved 2.0 Å resolution. O-state and potential K-state were solved at 1.6 and 1.7 resolution respectively (Table 3.1.8.1).

Data collection	Ground Cl	Ground SO_4	K state	O state
Space group	P 3 2 1	P 3 2 1	P 3 2 1	P 3 2 1
Cell dimensions				
a, b, c (Å)	61.98, 61.98, 109.83	62.30, 62.30, 110.14	61.99, 61.99, 109.82	62.41, 62.41, 110.08
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	53.68 - 1.58 (1.61-1.58)	48.45 - 1.91 (1.94 - 1.91)	48.23 - 1.70 (1.73 - 1.70)	48.52 - 1.60 (1.64-1.60)
Anisotropic resolution limits (Å)	1.54, 1.54, 1,62	1.87, 1.87, 1.89	1.67, 1.67, 1.70	1.56, 1.56, 1.71
No. of observations	352260 (7541)	211230 (8597)	294113 (14721)	349218 (17878)
No. of unique reflections	33368 (1667)	19823 (992)	27378 (1379)	31399 (1572)
Rpim	0.059 (0.993)	0.081 (0.816)	0.101 (0.975)	0.065 (0.820)
I/σI	8.0 (0.7)	7.5 (0.9)	5.5 (0.9)	7.0 (1.1)
CC1/2	99.8 (22.5)	99.7 (39.9)	99.6 (30.6)	99.9 (41.0)
Completeness spherical (%)	96.9 (72.8)	99.6 (92.3)	99.2 (91.7)	93.8 (62.3)

Completeness anisotropic (%)	98.2 (90.3)	99.6 (92.3)	99.2 (91.7)	97.5 (93.8)
Multiplicity	10.6 (4.5)	10.7 (8.7)	10.7 (10.7)	11.1 (11.4)
Wilson B-factor	17.56	22.36	18.76	22.21
Refinement statistics				
Resolution (Å)	19.95 - 1.58	19.27 - 1.91	19.95 - 1.70	19.28 - 1.60
No. unique reflections	33331	19758	27263	31363
Rwork/Rfree	0.1716/0.2021	0.1943/0.2250	0.1849/0.2217	0.1946/0.2322
No. of atoms				
Protein	1900	1922	1930	1888
Retinal	21	21	42	21
Cl	1	1	1	0
SO_4	0	10	0	0
Water	126	75	126	121
Lipids	238	165	238	233
No. of TLS groups	4	0	0	4
B-factors ($Å^2$)				
Protein	19.92	25.42	19.84	20.09
Retinal	13.99	20.52	15.00	15.72
Cl	14.61	18.68	15.86	-
SO_4	-	72.22	-	-

Water	34.15	34.43	33.42	33.22
Lipids	41.65	43.66	41.83	40.84
R.m.s.d.				
Bong lengths (Å)	0.012	0.015	0.006	0.007
Bond angles (°)	1.16	1.68	0.82	0.89
Ramachandran statistics				
Favoured (%)	99.09	99.09	99.09	98.63
Allowed (%)	0.91	0.91	0.91	1.37

Table 3.1.8.1. Data collection and refinement statistics for the 4 obtained structures of SyHR

Each of these four structures will be described in details in the following parts §3.2 – §3.5.

3.2 Structure of Cl-bound form in ground state

3.2.1 Overal structucture and retinal binding pocket

Using red crystals of *Sy*HR the Cl^--bound structure of the protein were solved at 1.57 Å resolution. The high-resolution structure reveals 126 protein-associated water molecules, including 27 inside *Sy*HR, and 20 lipid fragments surrounding each protomer (Fig. 3.2.1.1). Residues 1-225 from the total 234 of the *Sy*HR molecule are resolved in the model. The *Sy*HR protomer consists of 7 transmembrane α-helices (A-G) connected by three intracellular and three extracellular loops like in other type I rhodopsins. The BC extracellular loop is 13 residues long and contains a β-sheet formed by two β-strands.

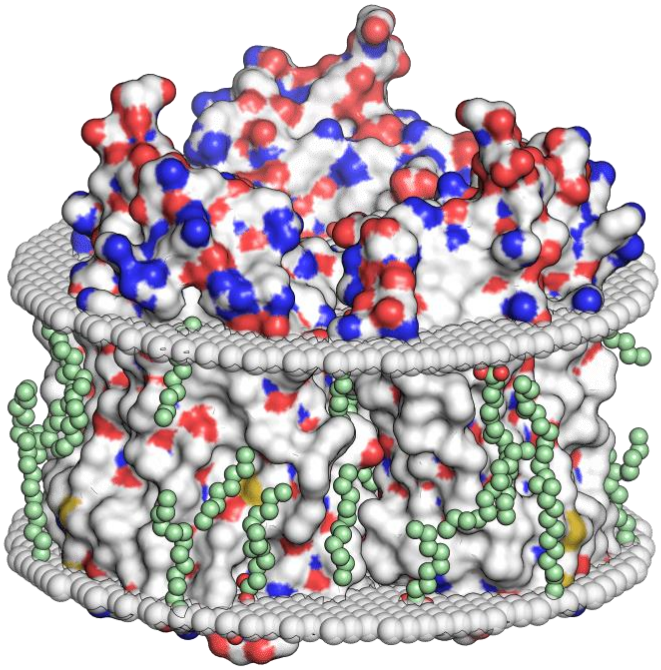

Figure 3.2.1.1. Surface representation of SyHR trimer in membrane.

As it was already mentioned, *Sy*HR is organized into trimers in the crystals. Similar trimers are known for many MRs, including archaeal H^+ and Cl^- pumps, but also *Mr*HR from the same clade of cyanobacterial anion-pumping rhodopsins. Interprotomer contacts are formed via pairs of residues A46 - Y106 and Q57 - W126 of the helices B with D and BC loop with helix E respectively (Fig. 3.2.1.2).

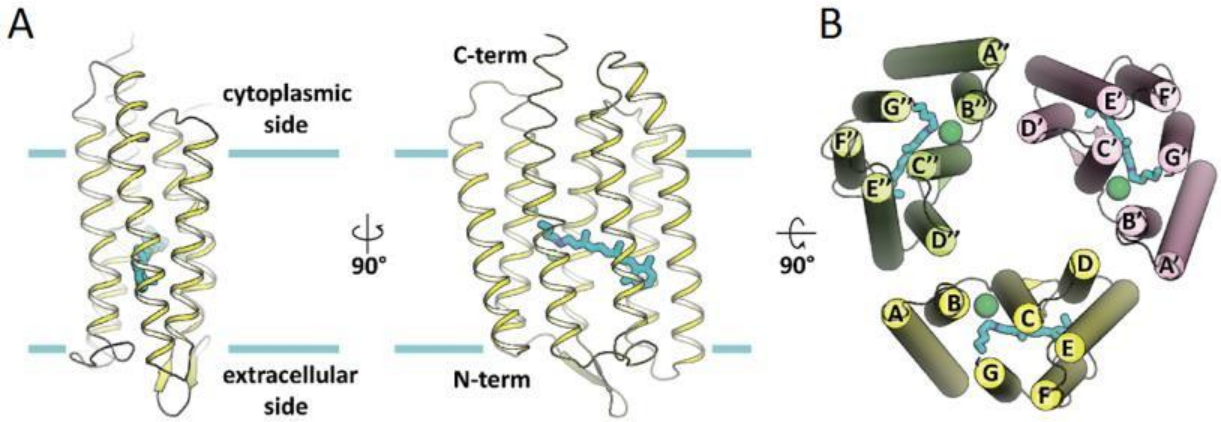

Figure 3.2.1.2. General view of *Sy*HR. (A) Two projections of *Sy*HR monomer in membrane. (B) Top view on *Sy*HR trimer.

Opposite to the structures of all other Cl^- pumps, we have not identified any Cl^- ions on the surface of *Sy*HR.

A retinal chromophore is covalently bound to K205 and is in the all-*trans* configuration in the ground state of *Sy*HR. As in archaeal HRs, Cl^- ion is found in the RSB region of *Sy*HR in its ground state (Fig. 3.2.1.3).

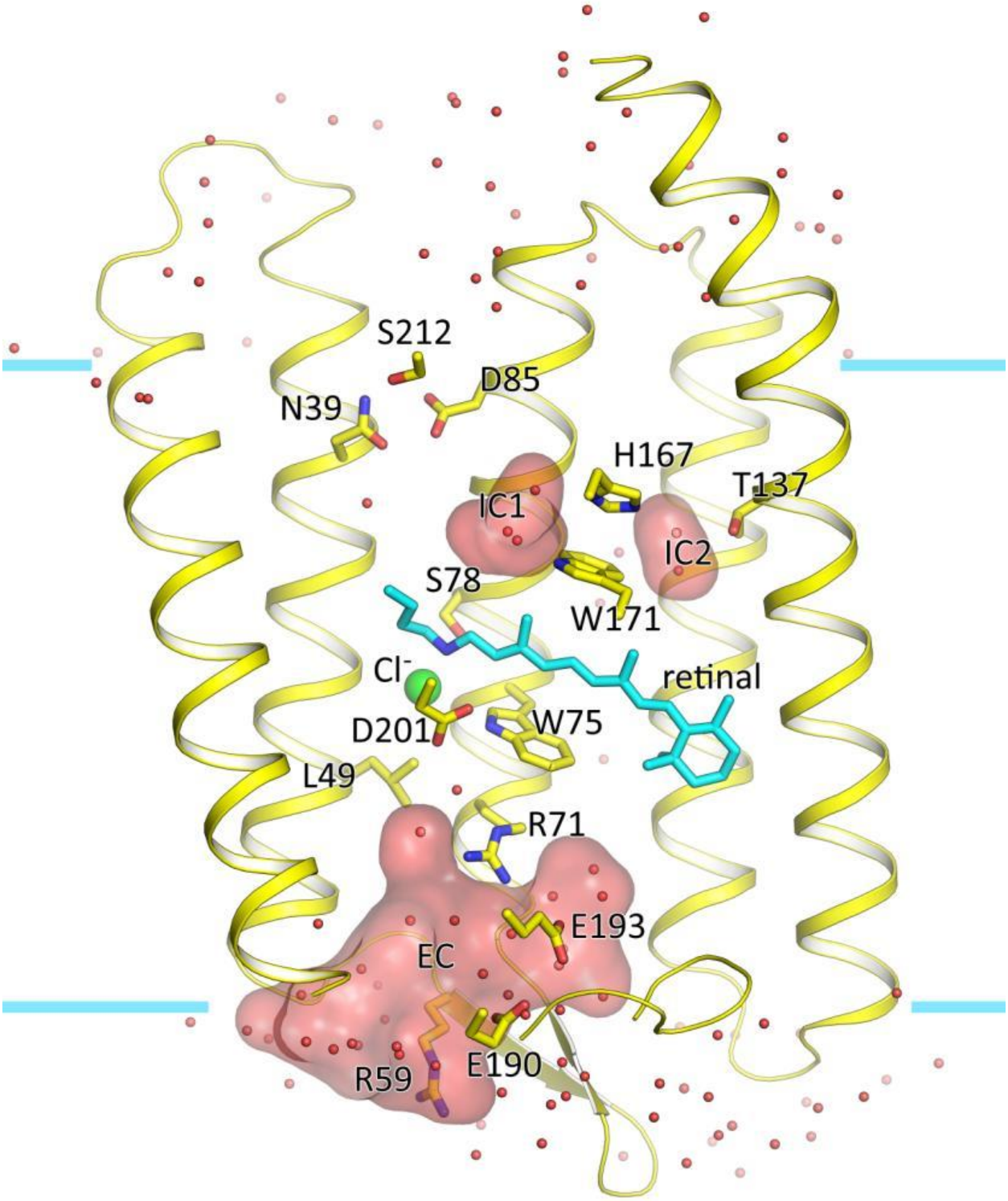

Figure 3.2.1.3. Structure of monomer of *Sy*HR

The chloride ion is coordinated by the T74 and S78 side chains of the characteristic TSD motif and also one water molecule (w401). The water molecule is further H-bonded to the RSB and D201. The distance between the RSB nitrogen and the Cl^- ion is 3.8 Å (Fig.3.2.1.4).

It should be noted that the retinal binding pocket of *Sy*HR is very similar to that of *Mr*HR. The similar retinal environment is in line with the fact of the similar maximum absorption wavelengths of the two proteins in their Cl^--bound state (536 and 537 nm for *Sy*HR and *Mr*HR, respectively).

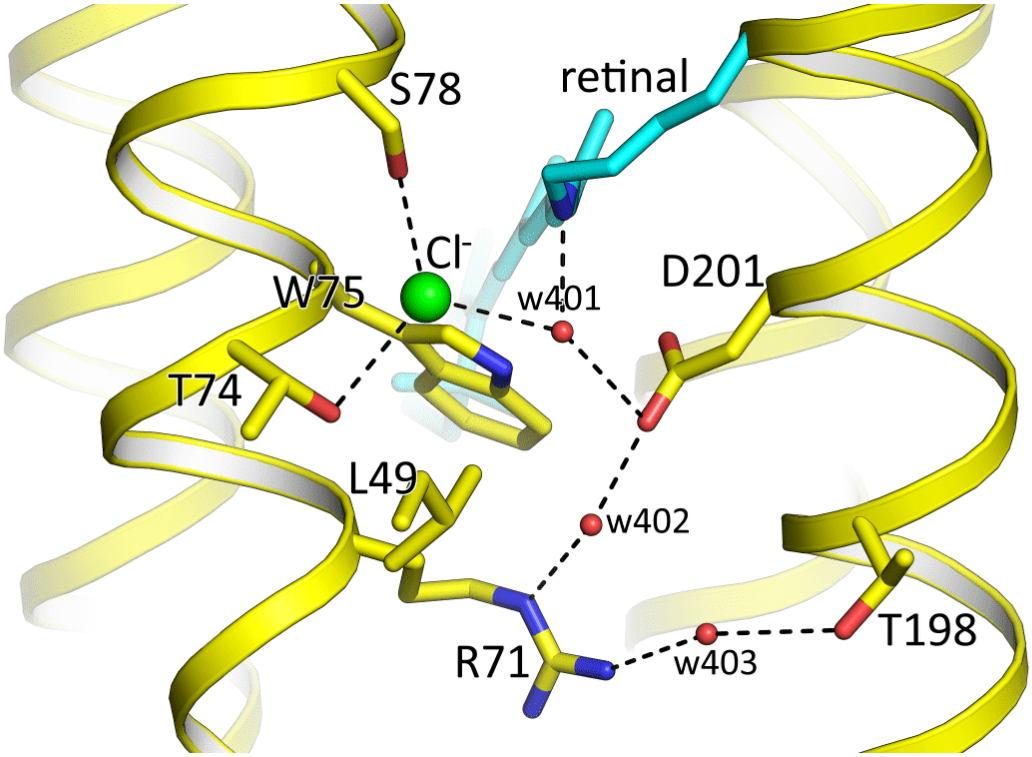

Figure 3.2.1.4. Structure of RSB region of *Sy*HR.

3.2.2 Structure of the ion release region of *Sy*HR

The structure of *Sy*HR in the ground state reveals two internal cavities at the cytoplasmic half of the protein, which we named as intracellular cavity 1 (IC1) and intracellular cavity 2 (IC2). The cavities are located in the region of H167 and are separated by the side chain of this residue. The presence of H167 is characteristic for only the clade of cyanobacterial anion-pumping rhodopsins. It is absent in both archaeal HRs and eubacterial Cl^- pumps. It is located in helix F and it is stabilized in *Sy*HR by the H-bonds with two water molecules (Fig.3.2.2.1).

The size of IC1 in *Sy*HR is dictated by the conformation of the L170 side chain, which adopts two alternative orientations. In one of them, L170 is pointed towards the lipid bilayer, which results in the enlargement of IC1. In this case, IC1 is filled with three water molecules. In the second conformation, the CD1 atom of the L170 side chain is moved closer to the inside of the *Sy*HR protomer. In this configuration, there are only two water molecules in IC1. Water molecules are H-bonded to each other and are stabilized by the side chains of W171 and H167 and also by the carbonyl oxygen of S204. Interestingly, the analogous cavity to IC1 is found also in *Mr*HR. However, in *Mr*HR it is smaller and is filled with only one water molecule, which does not interact with the characteristic histidine residue of the helix F (H166 in *Mr*HR).

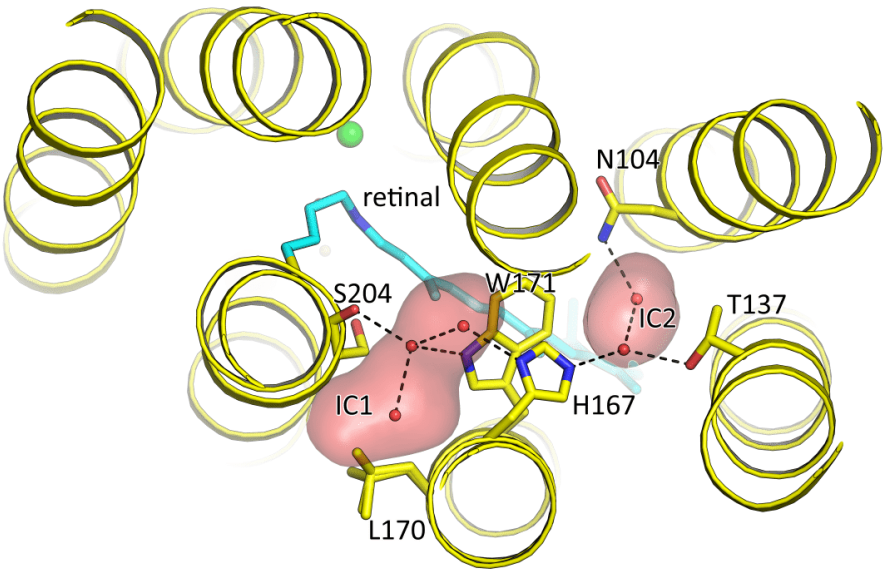

Figure 3.2.2.1 Top view of H167 region.

IC2 is characteristic only for *Sy*HR, but not for *Mr*HR. This is due to the presence of polar T137 in *Sy*HR at the position of hydrophobic I136 in *Mr*HR. Consequently, the IC2 is a relatively compact hydrophilic cavity filled by two water molecules, coordinated by N104, T137, and H167 residues. In total, there are five water molecules in the region of H167 in *Sy*HR, compared to only one molecule near H166 in *Mr*HR.

Another feature of the cyanobacterial anion-pumping rhodopsins is the presence of D85 in the helix C at the cytoplasmic side of the protein. This residue is part of the characteristic TSD motif. As mentioned above, it is an analog of the proton donor D96 residue of BR. It should be noted that a corresponding aspartate is absent in the archaeal HRs as well as in the eubacterial Cl^- pumps. In *Sy*HR, D85 is stabilized by the H-bonds with N39. N39 is further stabilized by the H-bonds with the water molecule w501 and S212. Water w501 is also connected with the

carbonyl oxygen of K205, to which the retinal is attached. The area is surrounded by a hydrophobic region formed by L35, L89, F208, I42, and L82 residues (Fig. 3.2.2.2.). A similar organization is found in *Mr*HR; the only minor difference is that in *Sy*HR the S212 side chain occupies two alternative conformations.

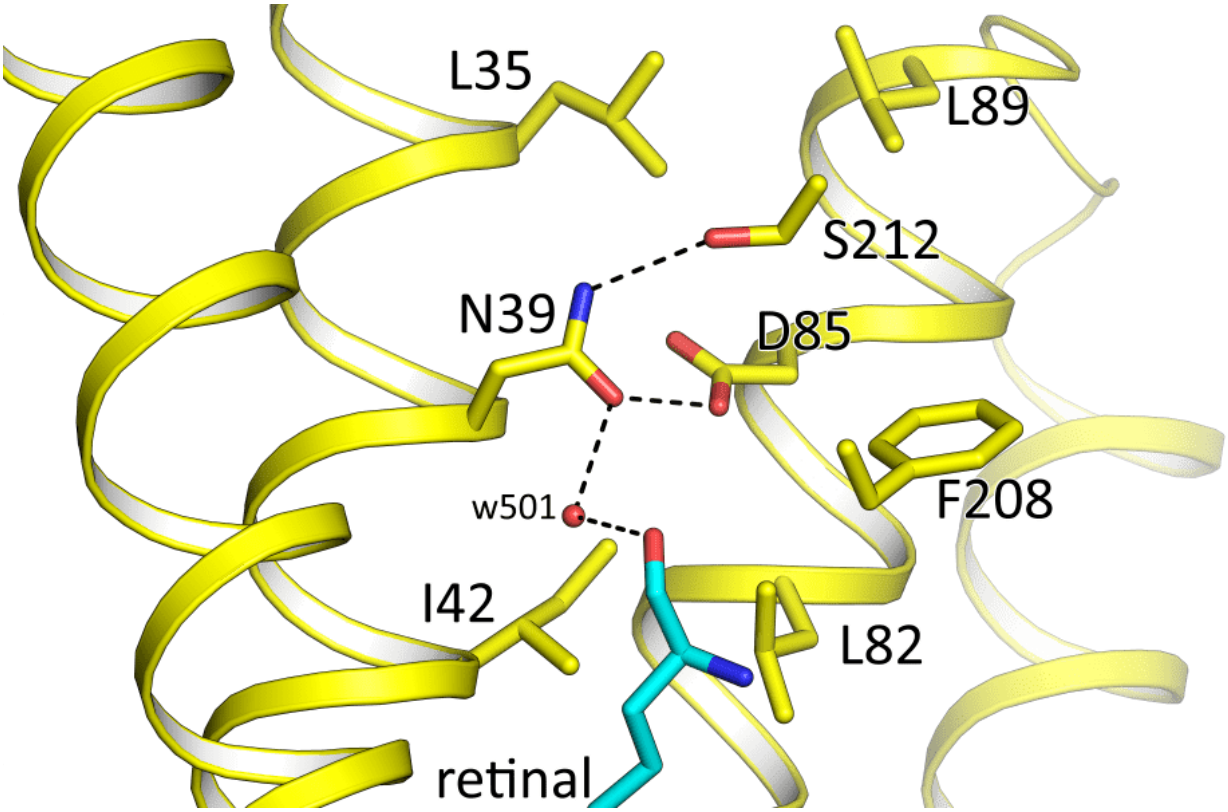

Figure 3.2.2.2. Structure of D85 region.

3.3 Structure of SO_4-bound form in ground state

As it was mentioned earlier, we obtained not only red but also violet crystals of *Sy*HR. The precipitant solution in this case contained 2.0 M of SO_4^{2-} and lacked Cl^- ions. Moreover, the protein used for crystallization was purified without Cl^-. These crystals allowed us to solve the structure of the protein at 2.0 Å resolution. Since the concentration of SO_4^{2-} in precipitant solution was much higher than its K_d value of *Sy*HR for this type of ions[71] (5.81 mM) and the maximum absorption wavelength of the protein in violet crystals is similar to that reported for the SO_4^{2-}-bound form of *Sy*HR in solution[71] (556 nm), we assume that the determined structure represents *Sy*HR in the SO_4^{2-}-bound form. However, a shoulder at 536 nm indicated the presence of chloride.

Since it is known that *Sy*HR has low K_d to Cl^- (0.112 mM), even after using Cl^--free solutions during protein purification and crystallization, we found that approx. 25.5% of protein molecules within the violet crystal contain Cl^- ion bound near the RSB. We suggest that it is due to the trace amounts of Cl^- in the high-molarity salts used as precipitants for crystallization or co-purification of chloride ions with SyHR. Nevertheless, the major fraction (74.5%) is Cl^--free. This high concentration of Cl^--free molecules enabled to determine the structure of violet *Sy*HR (*Sy*HR^v) in the presence of sulfate.

*Sy*HR^v contains two SO_4^{2-} ions on its surface. Both ions were found at the cytoplasmic side of the protein (Fig. 3.3.1). The first ion is placed between the helices D and E and interacts with the side chain of Q148 and the main chain oxygen of Q144 of the helix E. It is also within 3.2 Å from the K33 of the neighbouring protomer. The second ion is 13 Å away from the latter one and strongly interacts with R146 and also with E150 of the helix E. Both ions are

distant from the RSB (28 and 30 Å, respectively). The positions of the SO_4^{2-}-coordinating residues are similar in the Cl^-- and SO_4^{2-}-bound structures of *Sy*HR.

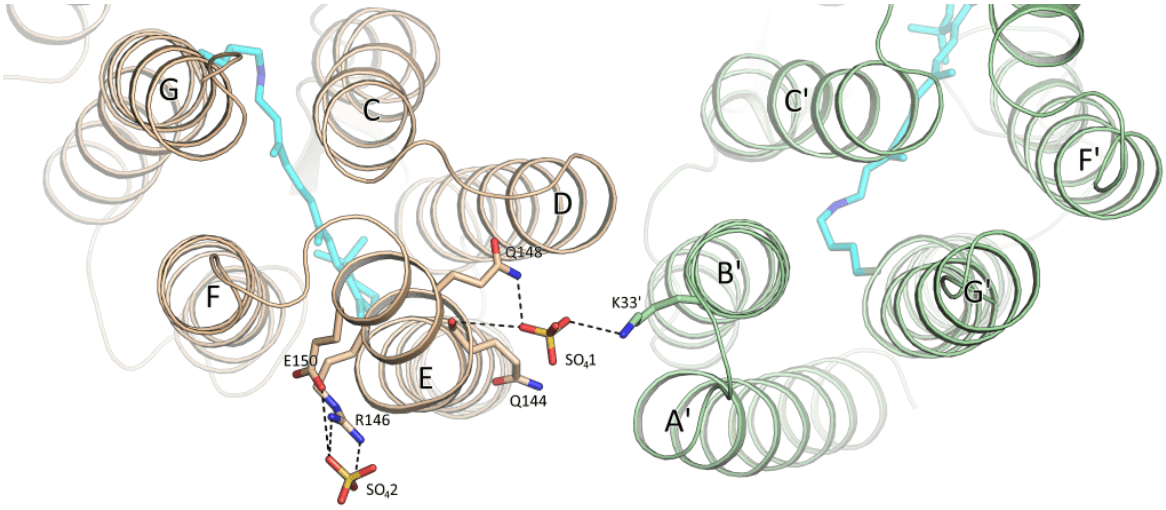

Figure 3.3.1. Coordination of sultphate ions between two SyHRv protomers.

Surprisingly, the structural data showed that there is no SO_4^{2-} ion close to the RSB in *Sy*HR^v. This stands in contrast with the earlier spectroscopy study, where it was proposed that SO_4^{2-} occupies the same site as Cl^- near T74 and S78 residues[71]. The hypothesis was supported by the notable red shift (542 nm to 556 nm) of the absorption peak with the addition of SO_4^{2-} to the detergent-solubilized *Sy*HR in the absence of Cl^-. On the contrary, in the presence of Cl^-, SO_4^{2-} does not affect the *Sy*HR spectrum. Thus, our structure, together with the spectroscopy of *Sy*HR^v in crystals leads to the conclusion that such red-shifted maximum absorption wavelength is observed without SO_4^{2-} binding near the RSB. We suggest that there are long-distance interactions of the surface-bound SO_4^{2-} ions with the RSB, resulting in the red shift of the spectrum of *Sy*HR. For instance, such interactions

were recently demonstrated for a light-driven Na^+ pump KR2[122]. In KR2, similar to the SO_4^{2-}-bound form of *Sy*HR, the transported substrate is not bound in the core of the protomer in the ground state. Instead, in both rhodopsins, ions are located on the protein surface, near the putative ion-release regions. Since there is no structure of *Sy*HR in the absence of both Cl^- and SO_4^{2-}, the exact influence of the surface-bound ions on this protein is not completely clear at the moment.

The RSB region is altered in the SO_4^{2-}-bound form of *Sy*HR compared to that of the Cl^--bound form. In general, the organization of RSB pocket is similar to that of the anion-free blue form of *Np*HR[92]. T74 is reoriented notably and moved closer to the RSB, while the side chain of L49 is flipped outside of the protomer to allow enough space for T74 (Fig. 3.3.2). As a result, the positions of residues 71-78 of the helix C are also altered in *Sy*HR^v. Consequently, the side chain of R71 is moved towards the extracellular space by 1 Å. However, the latter has only a minor effect on the positions of the surrounding water molecules and residues in the extracellular basin.

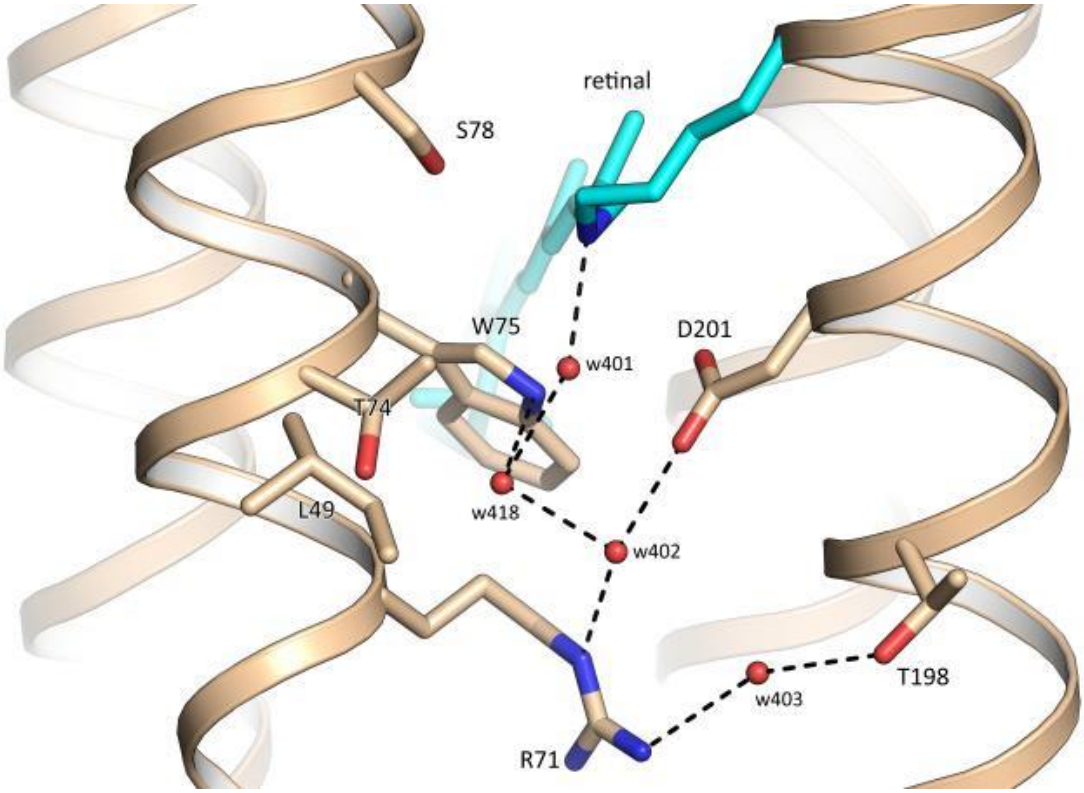

Figure 3.3.2. Organization of RSB region of *Sy*HR[v].

Notably, the synchronous rearrangements of T74 and L49 (Fig. 3.3.2) are similar to those observed in the light-driven sodium pump KR2. Indeed, the L74 and N112 residues of KR2, which are analogs of L49 and T74 of *Sy*HR, are reoriented in the same manner in course of photocycle upon Na^+ binding in the core of the KR2 protomer[115].

3.4 Structure of O-state of Cl-bound SyHR

By applying mentioned protocol, we solved 1.6 Å-resolution structure of the O state of the *Sy*HR photocycle. Crystallographic data analysis identified that the major fraction of the protein molecules in the crystal are in the O state. F_o-F_c difference electron density maps showed the trace amount of the ground state, insufficiently low for the proper fitting of this residual conformation (Fig. 3.4.1). The state composition of the *Sy*HR crystals after the cryotrapping procedure is also supported by the microspectrophotometry experiments. Therefore, we fitted the crystallographic data exclusively with the O state structure in the final model.

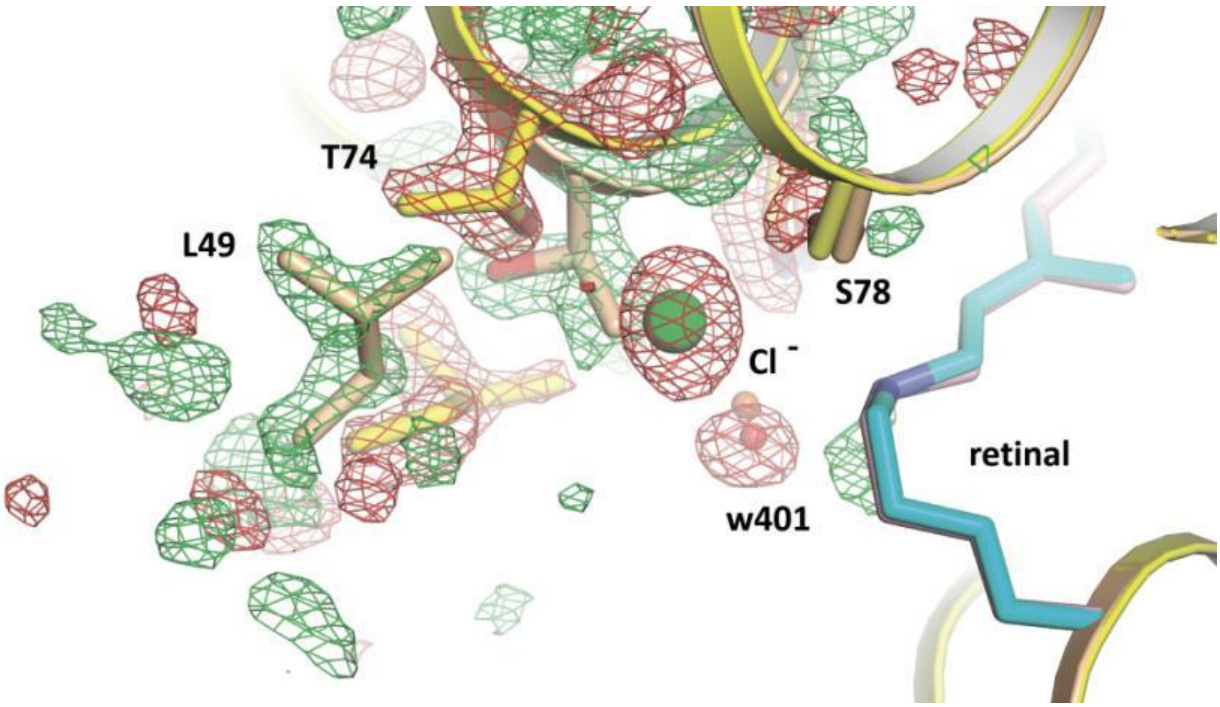

Figure 3.4.1. $2F_o$-F_c densities for the O-state. Structure of ground state is shown yellow. O-state is shown wheat.

Major conformational changes in the structure of the O state, in comparison to the ground state, are observed exclusively in the area around the RSB. Interestingly, the O state structure of the RSB region is very similar to the SO_4^{2-}-bound form in the ground state. Namely, there is no chloride ion near the RSB. Also, we observed the same flip of L49 and rearrangement of T74.

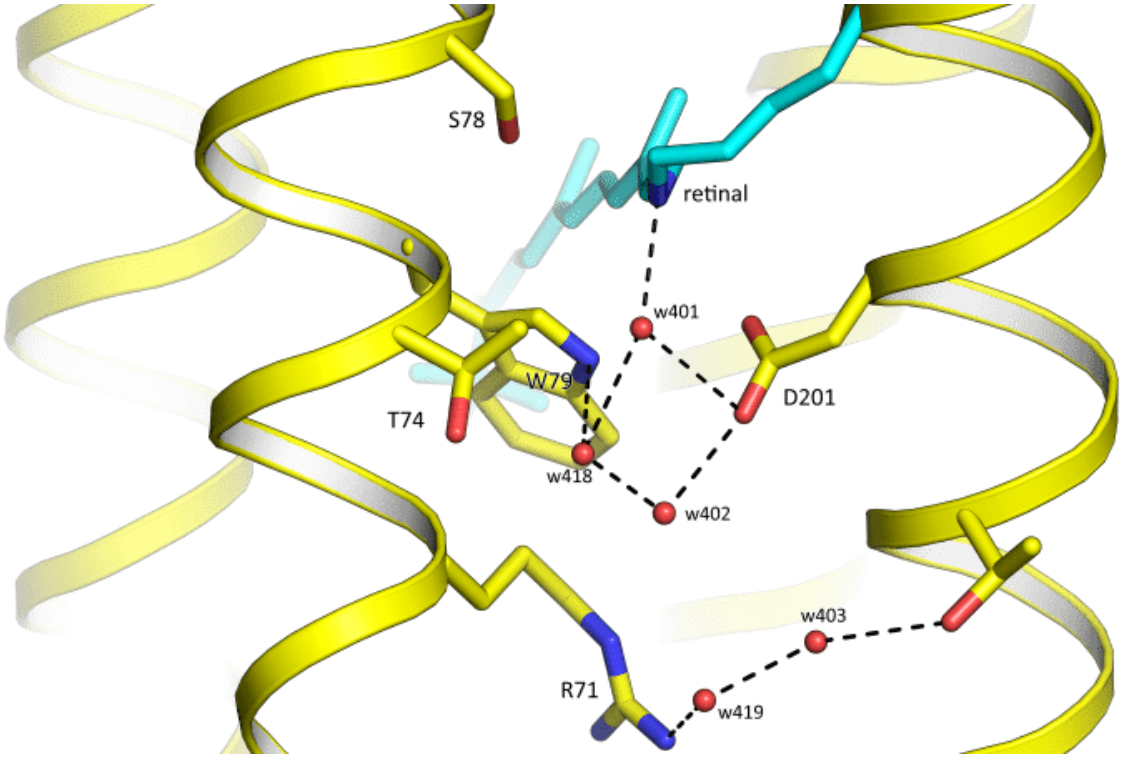

Figure 3.4.2. Organisation of the RSB region in the O-state.

The only major difference between the O state and the SO_4^{2-}-bound form is the position of R71, which is slightly shifted towards the extracellular side in the O state structure. The change in the position of this residue led to rearrangement of H-bonds and appearance of a new water molecule w419, not found in both the ground-state and SO_4^{2}-bound form. Thus, in the O state, R71 forms an H-bond

with this water w419 instead of w403. In its turn, w419 is H-bonded to w403, which is coordinated by T198 (Fig. 3.4.1).

3.5 Structure of the potential K-state.

Using the crystals of *Sy*HR with the cryotrapped K state we solved the structure of the protein at the 1.7 Å resolution. $2F_o$-F_c electron density maps showed that the major fraction of the proteins in the crystal remains in the ground state of the Cl^--bound form of *Sy*HR. Thus, the small difference in the maximum of the absorption spectrum between the ground and these active states is also explained by the low population of this state during cryotrapping (Fig. 1.3.7.1). The difference F_{oK}-F_{oGr} electron density maps indicated the presence of the strong signals above the level of 5σ near the C12, C13, C19, and C20 atoms of the retinal cofactor and also around its β-ionone ring as well as near the W75 and L82 side chains. The analysis of the structural data indicated that the fraction of the active state is around 35% in the crystal and the rest 65% is ground state. To build the model more accurately, extrapolated electron density maps was used.

The data showed structural rearrangements in the retinal region of *Sy*HR in this active state. Surprisingly, the retinal cofactor is in very the unusual configuration for microbial rhodopsins. The polyene chain of retinal is turned upside down by 12-s-cis isomerization. In this case, the β-ionone ring is not reversed due to 6-s-cis isomerization.

The C20 atom of the retinal is thus moved towards the L82 residue by 1.7 Å, which, in its turn, is reoriented to allow enough space for the cofactor isomerization. At the same time, the C19 atom of the retinal is moved closer to the

W75 residue. Consequently, the side chain of W75 is rotated by 10 degrees (Fig. 3.5.1). The β-ionone ring of the retinal is also shifted by almost 2Å in this state. The described structural rearrangements occur synchronously as indicated by the simultaneous appearance of the strong signal at the difference electron density maps near the retinal and W75 and L82 residues.

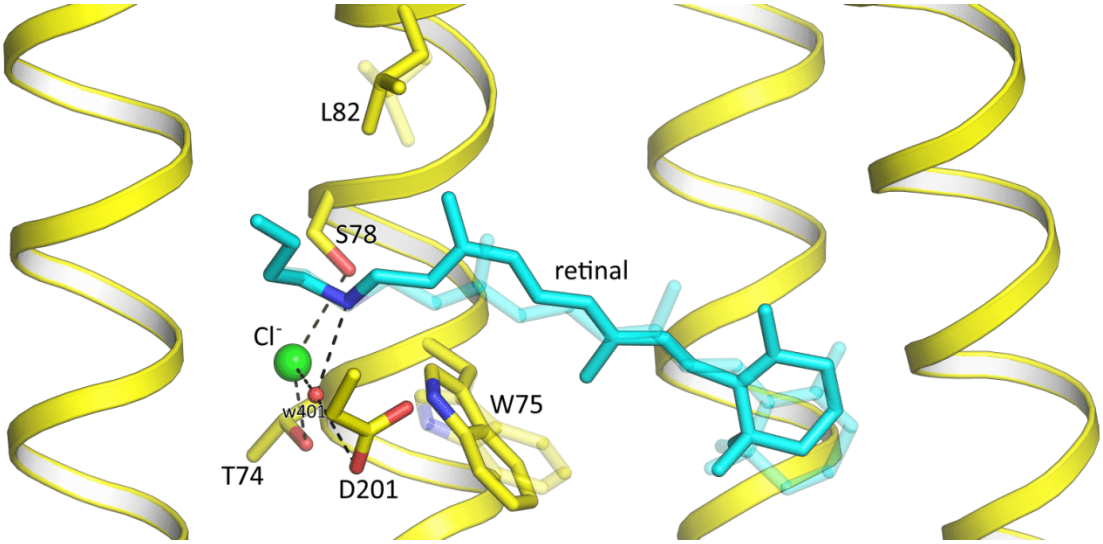

Figure 3.5.1. Changes in RSB region in a potential K-state. Ground-state is shown as transparent.

The RSB is not displaced considerably in this state. The surrounding of the RSB, such as water molecule w401, residues T74, and S78, as well as the Cl^- ion, occupy the same positions as in the ground state of the *Sy*HR photocycle.

Thus, the main change was the flip of the retinal polyene chain and the change in the amino acid position caused by this flip. This conformation of retinal has never been observed in microbial rhodopsins before. However, unusual conformations of retinal, such as, for example, 11-cis, were also observed in

halorhodopsins[96]. In addition, isomerization of 6-s-cis was also observed in microbial rhodopsins[123]. Also, the retinal polyene chain twist was observed in schizorodopsins[124].

Summing up, in this state there is a red shift in the spectrum, small structural changes affecting mainly the retinal, as well as the unchanged position of the chloride, its binding amino acids and the Schiff base.

Thus, the resulting state is an early state with a red shift. Based on this information, it can be assumed that the resulting state is a K - state.

Under this assumption, we observe a disturbance of retinal occurring prior to the 13-cis isomerization, since chloride and RSB are still in the same places as in the ground in this state. Probably, further in the course of the photocycle, 13-cis isomerization and changes in the position of the chloride will occur.

However, this information is only an assumption. It is possible that this state is not native, but only artifact associated with the low mobility of the protein in the crystal at cryo-temperatures.

Further investigation of the position of the retinal in the K - state using, for example, FTIR - spectroscopy can confirm or refute our hypothesis. Such research would be a useful follow-up to research on this protein.

3.6. Potential mechanism of the chloride pumping.

The structural data on the ground, K, and O functional states of the *Sy*HR photocycle presented here allowed us to propose the potential molecular mechanism of chloride pumping by the protein. In general, the suggestion is that the mechanism is similar to that proposed for archaeal HRs, such as *Np*HR[95].

The retinal binding pocket and the RSB region are organized similarly in *Sy*HR and *Np*HR[54]. In both cases, chloride ion is coordinated by two residues, serine and threonine, and one water molecule (Fig. 3.6.1).

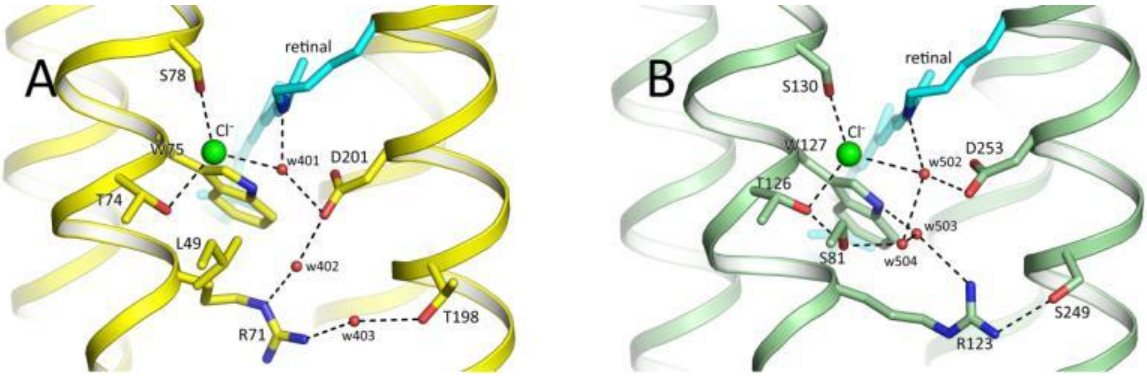

Figure 3.6.1. Comparison of active center of *Sy*HR and *Np*HR. (A) Retinal binding area of *Sy*HR. (B) Retinal binding area of *Np*HR.

Notably, the coordination of the chloride ion and water molecule w401 in *Sy*HR is tighter compared to *Np*HR. Indeed, the distance between chloride and the closest water molecule is 3.0 and 3.3 Å in *Sy*HR and *Np*HR, respectively. Also, in *Np*HR, this water molecule forms a weak 3.5 Å H-bond with the RSB, whereas in *Sy*HR the corresponding distance is 2.6 Å. These tighter hydrogen bonds explain well why *Sy*HR has a much higher affinity to chloride than *Np*HR and *Hs*HR.

Indeed, *Sy*HR has a Kd of 0.112 mM, while *Np*HR only 2 mM and *Hs*HR even 10 mM[71,56,125]. Notably, the presence of L49 in *Sy*HR at the position of S81 in *Np*HR also makes the region between the RSB and R71 less hydrated in cyanobacterial rhodopsin.

Comparing the O structures of *Sy*HR and *Np*HR, similarities are evident. Structure of the O state of *Np*HR was obtained by (Koyama et al.)[95]; unfortunately, the coordinates corresponding to this state are not available in the Protein Data Bank (PDB). However, there is a structure of the anion-free blue form of *Np*HR, which resembles the structure of the O state of *Np*HR[92]. Therefore, the structure of the blue form of *Np*HR were used for the comparison of the structures of the O states of *Sy*HR and *Np*HR.

The main structural difference between the ground and anion-free form of *Np*HR is the flip of T126. Thus, in the O state of *Np*HR, the sidechain of T126 occupies the original space of the chloride binding site. Also, a similar reorientation of S81 is observed which corresponds to the L49 flip in *Sy*HR (Fig. 3.6.2).

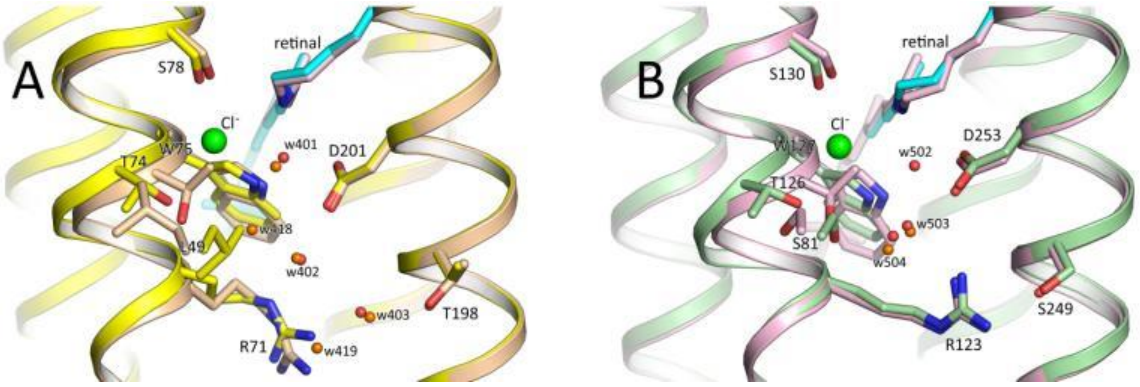

Figure 3.6.2. Ground and O states of *Sy*HR and *Np*HR. (A) Comparison of ground form of SyHR (yellow) and O-state (wheat). (B) Comparison of ground state of NpHE (green) and anion-free O-like state (pink).

Otherwise, the structure of the cytoplasmic and intracellular parts of the proteins are mostly the same for the ion-bound and ion-free forms. The exceptions are amino acids Y124 and G234 in *Np*HR. These residues undergo significant rearrangements during the switch between the ion-bound and ion-free forms.

Interestingly, *Sy*HR has a corresponing tyrosine (Y72), which remains in the same place, exactly as T183, which corresponds to G234 in *Np*HR.

The remarkable similarity of the anion-free forms and the O states suggests that these proteins have similar chloride pumping mechanisms, although they are phylogenetically quite distant. The high affinity for chloride and the ability to pump chloride against strong concentration gradients makes *Sy*HR potentially a good candidate for being a next-generation optogenetic tool.

According to the available experimental data and presented here high-resolution structures of *Sy*HR in different functional states we propose the following key events to occur during the chloride transport by the protein. First, upon absorption of the light photon, retinal become excited and passes from the

all-*trans* to 6-s-cis and 12-s-cis configuration in the K state. This differs from other MRs and likely means a unique primary photochemical reaction in *Sy*HR. However, the ability of MRs, including HRs, to maintain the retinal in unusual (11-cis) configuration as a result of the red-light adaptation was described several times in literature[126,96]. Also 6-s-cis configuration of retinal was already observed in the microbial rhodopsins[123]. It is important that for the further validation of our hypothesis additional experiments are required. Nevertheless, the high resolution in the K-state allows to position the retinal only in this way. Thus, the structural data provide initial insights into the early-stage dynamics of *Sy*HR.

This initial reconfiguration of the retinal does not affect the organization of the RSB region and Cl- ion bound near the RSB. However, residues W75 and L82 of the retinal binding pocket are rearranged to avoid steric conflict with the cofactor. We suppose that retinal transition occurs from 11-*cis* to 13-*cis* occurs later in photocycle, because at this stage we do not observe any relocation of the RSB.

Next, during the transitions of *Sy*HR through the L-like and N-like states, the Cl^- ion, originally bound in the RSB region, is released to the cytoplasm with the formation of the O state. Our structural data together with the previous spectroscopy studies of the mutant forms of *Sy*HR allow us to suggest that the release of the Cl^- occurs through the region of the H167 residue. A large cavity near H167 might serve as part of the ion-release pathway. The mutation of the corresponding histidine to alanine in homologous *Mr*HR results in the dramatic deceleration of the L and N states decay, associated with the ion release. At the same time, the only alternative way for Cl^- release passes near the region of D85, which is not conserved within the family. Moreover, the substitution of the

aspartate D85 in *Mr*HR corresponding to D85 in *Sy*HR does not strongly affect the photocycle of the protein[70].

We hypothesize that Cl^- is translocated within the protein upon the K-to-L and the L-to-N state transitions in the same manner as it was shown for the *Np*HR. We cannot exclude that Cl- ion could also be transiently bound near H167. It should be noted, that despite the absence of analogs of H167 in archaeal HRs, the Cl^- release pathway is formed transiently during their photocycle in the same region of the protein as we suggest for *Sy*HR. Indeed, according to the structure of the N state of *Np*HR, chloride releasing way is formed near the cytoplasmic part of helix F[95]. In *Np*HR, the appearance of the kink of helix F at the K215 residue creates a water channel leading outside of the protein. K215 of *Np*HR corresponds to L164 in *Sy*HR, which is located close to H167. Therefore, the release channel in *Np*HR is located in nearly the same area as the proposed anion release pathway in *Sy*HR. This fact highlights again the similarity of the anion pumping mechanisms of archaeal and cyanobacterial HRs.

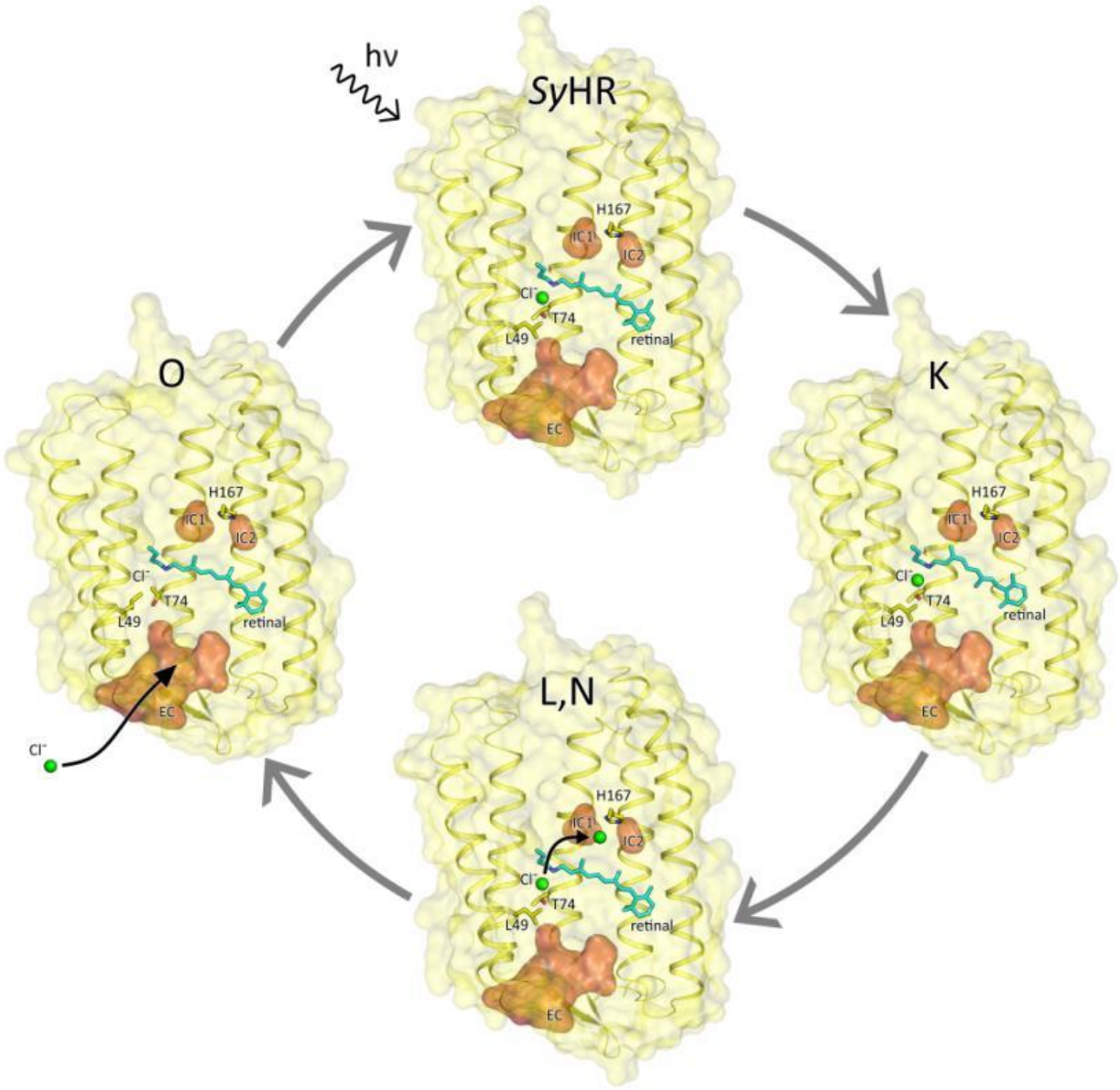

Figure 3.6.3. Schematic representation of chloride pumping process by *Sy*HR during photocycle.

Regarding the question of what steps of the photocycle correspond to the ion release and uptake in cyanobacterial anion pumps, there is no definitive opinion in the literature. Some researchers hypothesized that Cl^- release occurs upon the L-to-N state transition, while the ion is uptaken with the rise of the O state. Others

assume that ion release and uptake occur with the formation and the decay of the O state, respectively[70,71]. The structure of the O state of the *Sy*HR photocycle presented here clearly supports the second idea. Our structural data show that in the O state the retinal is isomerized back to the all-*trans* configuration, but the Cl^- ion is not yet bound in the region of the RSB.

The last step of the photocycle is therefore the uptake of the Cl^- through the extracellular side of the protein to its original binding site near the RSB. We suggest that the uptake occurs through the large polar concavity protruding from the extracellular bulk in the region between helices A, B, G to the R71 residue as identified in the structure of the protein. It is somewhat more difficult to specify what amino acids are involved in chloride uptake. Information from site-specific mutagenesis in *Mr*HR suggests that E193 (E192 in *Mr*HR) can play a role in chloride uptake because its mutation to alanine significantly slows down the turnover of the photocycle[70]. Finally, the binding of the Cl^- to its original site in the central part of the protein results in the *Sy*HR returning to its initial ground state (Fig. 3.6.1).

3.7 Sulphate pumping and selectivity mechanism.

3.7.1. Structural insights into sulphate pumping mechanism

It is known that *Sy*HR is capable of pumping sulfate ions in absence of chloride. But the molecular mechanism of the sulfate pumping was completely unknown. The obtained structure of the SO_4^{2-}-bound form of *Sy*HR reveals interesting details about sulfate transport.

According to the presented structure of $SO4^{2-}$ bound protein, there is no sulfate inside the rhodopsin in its ground state, despite the crystals being grown using $(NH_4)_2SO_4$ as the precipitant at a concentration of 2 M. Moreover, given the size of the active site of *Sy*HR, it is unlikely that sulfate can fit in it in the ground state. However, previous studies have identified a red shift in the absorption spectrum of a protein upon the addition of sulfate to an ion-free protein solution[71]. We suggest that the spectral shift is a result of long-range interactions between the central part of the protein near the retinal cofactor and sulfate ions bound at the two sites at the surface of *Sy*HR. Indeed, under sulfate bound conditions the absorption peak of *Sy*HR in violet crystals in the absence of Cl^- is 556 nm, which corresponds to the sulfate-bound, but not to the anion-free form of the protein in solution[71].

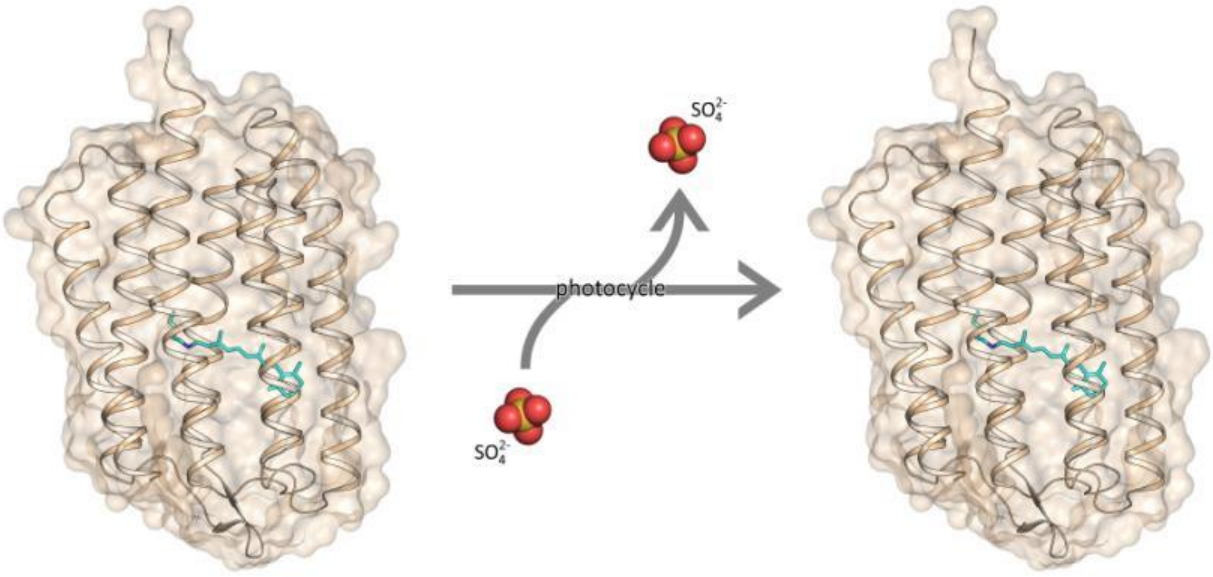

Figure 3.7.1.1. Schematic representation of sulphate pumping by *Sy*HR in course of photocycle. Sulphate ion uptakes after the beginning of photocycle and releases before the photocycle ends.

Thus, given our data, we assume that *Sy*HR does not bind sulfate inside the protein in the ground state. The ion enters the protein during the photocycle after the photon absorption by the retinal cofactor (Fig. 3.7.1.1). Unfortunately, our data is not sufficient to say with certainty at what stage of the photocycle the binding and translocation of sulfate occur. As soon as the ground state of $SO4^{2-}$-bound form is very similar to the O state of Cl-bound *Sy*HR, structures of the protein in the intermediate states may differ between the $SO4^{2-}$ and Cl^- pumping modes of *Sy*HR.

3.7.2 Sulphate selectivity mechanism

The high-resolution structure of *Sy*HR and reported earlier structures of the *Mr*HR and its SO_4^{2-}-pumping mutant allows us to suggest the key determinants of anion selectivity in cyanobacterial HRs.

We suggest that the unique ability of cyanobacterial HRs to pump sulfate is due to the presence of a large cavity at the extracellular side of the protein, serving as an ion uptake vestibule. Indeed, in *Sy*HR, the cavity protrudes from the extracellular space through the large pore in the protein to the R71 residue. In the wild-type *Mr*HR, which is incapable of pumping sulfate, the entrance of the pore is dramatically smaller. In *Sy*HR, we suppose that sulfate enters through a large extracellular cavity and then passes through a channel going inside the protein past the FG loop and more specifically T183. Through this channel, sulfate enters the central part of the protein, where it reaches R71 and T198 and further binds near T74 and S78 in the end of the photocycle. In *Mr*HR, this entrance to the cavity is much tighter and the transition from the cavity to the internal part of the channel is hindered by the presence of bulky E182 instead of T183 in *Sy*HR. It is known that both E182T and E182A variants of *Mr*HR are capable of pumping sulfate. Interestingly, E182A mutant pumps sulfate even more efficiently than *Sy*HR-like E182T mutant[72]. This suggests that the size of this amino acid is a determinant of sulfate pumping in cyanobacterial HRs. It should be noted that E182 is conservative and found in *Mr*HR, *Np*HR, and *Hs*HR, but only in *Sy*HR it is substituted by threonine (Fig. 3.7.2.1).

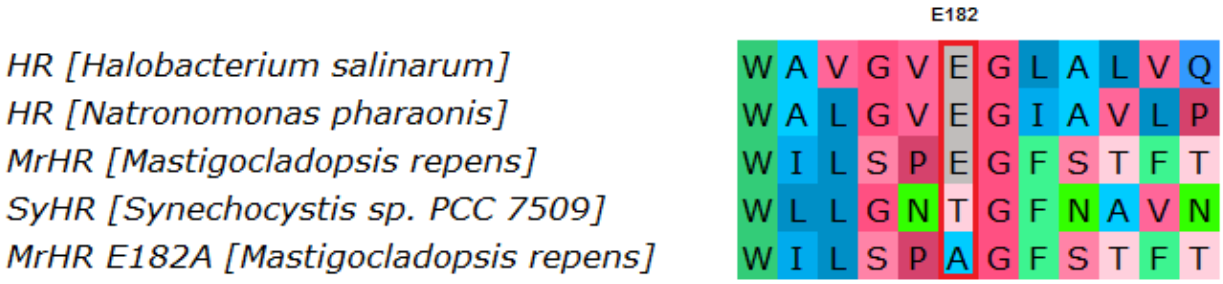

Fig. 3.7.2.1. Sequence aligment of the FG loop of different halorhodsins.

Nevertheless, unlike *Sy*HR, mutant forms of *Mr*HR possess another possible way for the sulfate to reach the internal part of the channel. Thus, double mutation N63A/P118A opens a new way for sulfate located between the BC and DE loops. This double mutant of *Mr*HR pumps sulfate although comprising a bulky negatively charged E183 residue. Interestingly, *Sy*HR also has alanines in positions corresponding to N63 and P118, but we have not found any additional channel in *Sy*HR in the region between BC and DE loops. In addition, the entire FG loop of *Sy*HR is very different from the corresponding loop of archaeal HRs and also from that of *Mr*HR.

Finally, *Sy*HR and *Mr*HR demonstrate some essential differences between the cytoplasmic internal regions, close to H167, and the putative ion release pathway. Namely, the cavities in *S*yHR are notably larger than those in *Mr*HR, which might also be a determinant of the effective translocation of a bulky divalent anion through the cytoplasmic side specific for *Sy*HR.

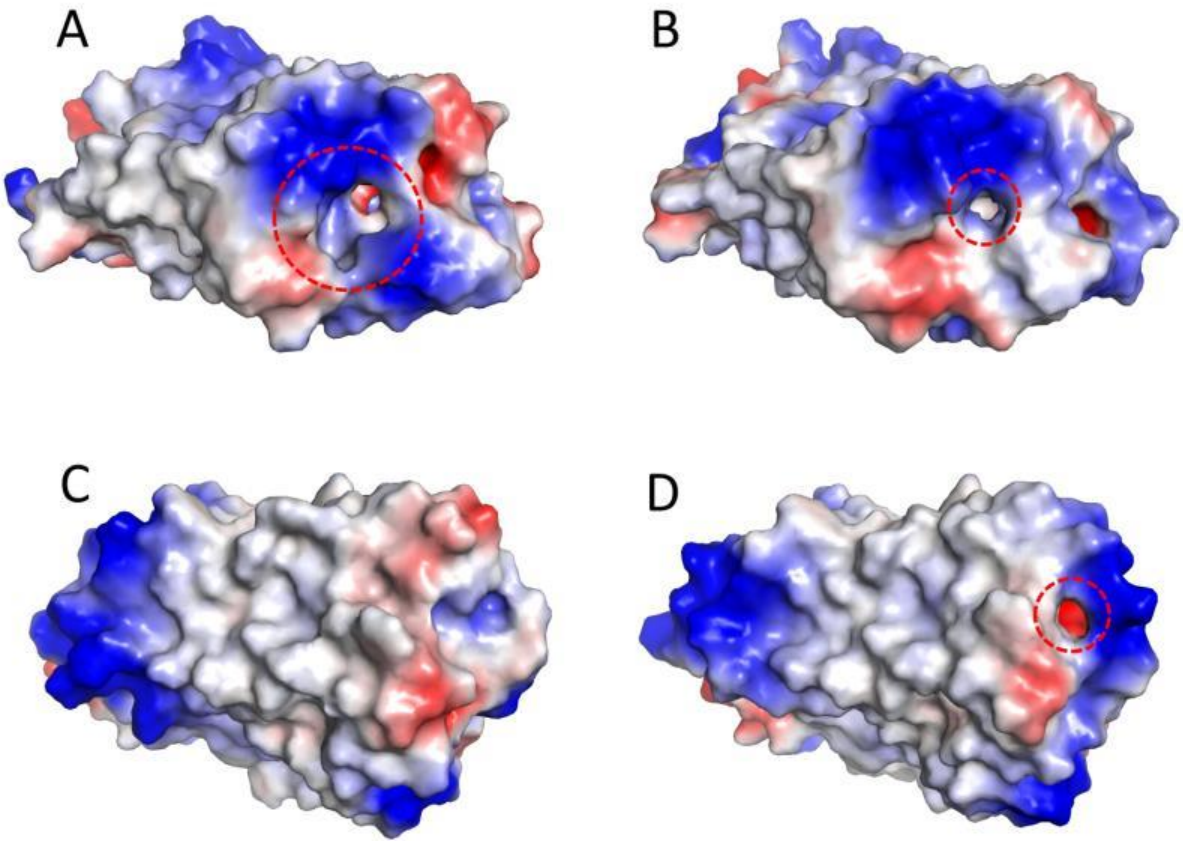

Fig. 3.7.2.2. Possible entry for sulphate in *Sy*HR and *Mr*HR. The main putative entrance to the cavity in SyHR (A) and MrHR (B).The second way for sulfate to enter the protein, which is obstructed in *Sy*HR (C) and opened in *Mr*HR-N63A/P118A (D).

4. Conclusion

In this work, a comprehensive structural study of the *Sy*HR cyanobacterial anion pump was carried out. As part of the work, the following new scientific results were obtained

1. Development of the protocol for large-scale expression of the *Sy*HR protein in chloride-bound and anion-free forms in E. coli cells in an amount sufficient for crystallization.

2. Obtaining crystals of *Sy*HR in chloride-bound and sulfate-bound forms.

3. Cryotrapping of the obtained crystals and spectral analysis of the resulting intermediates.

4. Determination of the first three-dimensional structure of the *Sy*HR protein with a resolution of 1.57 Å, which is a record value among microbial rhodopsins pumping anion.

5. Determination of the three-dimensional structure of *Sy*HR in sulfate-bound form with a resolution of 2.0 Å.

6. Determination of the three-dimensional structure of the O-state of the chloride-bound form of *Sy*HR with a resolution of 1.6 Å. This is the first structure of the active state among cyanobacterial anion pumps.

7. Determination of the three-dimensional structure of the potential K - state with a resolution of 1.7 Å.

The data obtained suggested a mechanism for pumping chloride by the *Sy*HR protein, and also gave structural insigths into unique sulphate pumping. Since the data obtained indicate that the mechanism of chloride pumping is very similar to that of archaeal rhodopsins. This fact suggests that *Sy*HR may become a more promising optogenetic tool than archaeal halorhodopsins, since it has a much stronger affinity for chlorine in the active site and therefore can pump chloride through a stronger negative chloride gradient. Studying *Sy*HR as a potential tool for optogenetics is a promising topic for future research.

5. References

1. Sharma AK, Spudich JL, Doolittle WF. Microbial rhodopsins: functional versatility and genetic mobility. *Trends Microbiol*. 2006. doi:10.1016/j.tim.2006.09.006

2. Haupts U, Tittor J, Bamberg E, Oesterhelt D. General concept for ion translocation by halobacterial retinal proteins: The isomerization/switch/transfer (IST) model. *Biochemistry*. 1997. doi:10.1021/bi962014g

3. Govorunova EG, Sineshchekov OA, Li H, Spudich JL. Microbial rhodopsins: Diversity, mechanisms, and optogenetic applications. *Annu Rev Biochem*. 2017. doi:10.1146/annurev-biochem-101910-144233

4. Kawanabe A, Kandori H. Photoreactions and structural changes of Anabaena sensory rhodopsin. *Sensors*. 2009. doi:10.3390/s91209741

5. Ugalde JA, Podell S, Narasingarao P, Allen EE. Xenorhodopsins, an enigmatic new class of microbial rhodopsins horizontally transferred between archaea and bacteria. *Biol Direct*. 2011. doi:10.1186/1745-6150-6-52

6. Yutin N, Koonin E V. Proteorhodopsin genes in giant viruses. *Biol Direct*. 2012. doi:10.1186/1745-6150-7-34

7. Gómez-Consarnau L, Raven JA, Levine NM, et al. Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Sci Adv*. 2019. doi:10.1126/sciadv.aaw8855

8. Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H. Microbial and animal rhodopsins: Structures, functions, and molecular

mechanisms. *Chem Rev*. 2014. doi:10.1021/cr4003769

9. Oesterhelt D, Stoeckenius W. Rhodopsin-like protein from the purple membrane of Halobacterium halobium. *Nat New Biol*. 1971. doi:10.1038/newbio233149a0

10. Henderson R. The purple membrane from Halobacterium halobium. *Annu Rev Biophys Bioeng*. 1977. doi:10.1146/annurev.bb.06.060177.000511

11. Schreckenbach T. LIGHT-ENERGY CONVERSION BY THE PUBLIC MEMBRANE FROM HALOBACTERIUM HALOBIUM. In: ; 1977. doi:10.1016/b978-0-08-021791-8.50027-5

12. Landau EM, Rosenbusch JP. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A*. 1996. doi:10.1073/pnas.93.25.14532

13. Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science (80-)*. 1997. doi:10.1126/science.277.5332.1676

14. Matsuno-Yagi A, Mukohata Y. Two possible roles of bacteriorhodopsin; a comparative study of strains of Halobacterium halobium differing in pigmentation. *Biochem Biophys Res Commun*. 1977. doi:10.1016/0006-291X(77)91245-1

15. Schobert B, Lanyi JK. Halorhodopsin is a light-driven chloride pump. *J Biol Chem*. 1982. doi:10.1016/s0021-9258(18)34020-1

16. Meyer TE, Yakali E, Cusanovich MA, Tollin G. Properties of a Water-Soluble, Yellow Protein Isolated from a Halophilic Phototrophic Bacterium

That Has Photochemical Activity Analogous to Sensory Rhodopsin. *Biochemistry*. 1987. doi:10.1021/bi00376a012

17. Takahashi T, Yan B, Spudich JL, et al. Color Regulation in the Archaebacterial Phototaxis Receptor Phoborhodopsin (Sensory Rhodopsin II). *Biochemistry*. 1990. doi:10.1021/bi00488a038

18. Yao VJ, Spudich JL. Primary structure of an archaebacterial transducer, a methyl-accepting protein associated with sensory rhodopsin I. *Proc Natl Acad Sci U S A*. 1992. doi:10.1073/pnas.89.24.11915

19. Beja O, Aravind L, Koonin E V., et al. Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science (80-).* 2000. doi:10.1126/science.289.5486.1902

20. Nagel G, Ollig D, Fuhrmann M, et al. Channelrhodopsin-1: A light-gated proton channel in green algae. *Science (80-).* 2002. doi:10.1126/science.1072068

21. Nagel G, Szellas T, Huhn W, et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc Natl Acad Sci U S A*. 2003. doi:10.1073/pnas.1936192100

22. Hegemann P, Nagel G. From channelrhodopsins to optogenetics. *EMBO Mol Med*. 2013. doi:10.1002/emmm.201202387

23. Yoshizawa S, Kumagai Y, Kim H, et al. Functional characterization of flavobacteria rhodopsins reveals a unique class of light-driven chloride pump in bacteria. *Proc Natl Acad Sci U S A*. 2014. doi:10.1073/pnas.1403051111

24. Inoue K, Ono H, Abe-Yoshizumi R, et al. A light-driven sodium ion pump in

marine bacteria. *Nat Commun*. 2013. doi:10.1038/ncomms2689

25. Govorunova EG, Sineshchekov OA, Janz R, Liu X, Spudich JL. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science (80-)*. 2015. doi:10.1126/science.aaa7484

26. Avelar GM, Schumacher RI, Zaini PA, Leonard G, Richards TA, Gomes SL. A Rhodopsin-Guanylyl cyclase gene fusion functions in visual perception in a fungus. *Curr Biol*. 2014. doi:10.1016/j.cub.2014.04.009

27. Bratanov D, Kovalev K, Machtens JP, et al. Unique structure and function of viral rhodopsins. *Nat Commun*. 2019. doi:10.1038/s41467-019-12718-0

28. Zabelskii D, Alekseev A, Kovalev K, et al. Viral rhodopsins 1 are an unique family of light-gated cation channels. *Nat Commun*. 2020. doi:10.1038/s41467-020-19457-7

29. Inoue K, Tsunoda SP, Singh M, et al. Schizorhodopsins: A family of rhodopsins from Asgard archaea that function as light-driven inward H+ pumps. *Sci Adv*. 2020. doi:10.1126/sciadv.aaz2441

30. Rozenberg A, Inoue K, Kandori H, Béjà O. Microbial Rhodopsins: The Last Two Decades. *Annu Rev Microbiol*. 2021. doi:10.1146/annurev-micro-031721-020452

31. Pushkarev A, Inoue K, Larom S, et al. A distinct abundant group of microbial rhodopsins discovered. *Nature*. 2018. doi:10.1038/s41586-018-0225-9

32. Shihoya W, Inoue K, Singh M, et al. Crystal structure of heliorhodopsin. *Nature*. 2019. doi:10.1038/s41586-019-1604-6

33. Lu Y, Zhou XE, Gao X, et al. Crystal structure of heliorhodopsin 48C12.

Cell Res. 2020. doi:10.1038/s41422-019-0266-0

34. Kovalev K, Volkov D, Astashkin R, et al. High-resolution structural insights into the heliorhodopsin family. *Proc Natl Acad Sci U S A*. 2020. doi:10.1073/pnas.1915888117

35. DeLong EF, Béjà O. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol*. 2010. doi:10.1371/journal.pbio.1000359

36. Klare JP, Bordignon E, Engelhard M, Steinhoff HJ. Transmembrane signal transduction in archaeal phototaxis: The sensory rhodopsin II-transducer complex studied by electron paramagnetic resonance spectroscopy. *Eur J Cell Biol*. 2011. doi:10.1016/j.ejcb.2011.04.013

37. Gordeliy VI, Labahn J, Moukhametzianov R, et al. Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. *Nature*. 2002. doi:10.1038/nature01109

38. Sasaki J, Spudich JL. Signal transfer in haloarchaeal sensory rhodopsin-transducer complexes. In: *Photochemistry and Photobiology*. ; 2008. doi:10.1111/j.1751-1097.2008.00314.x

39. Sudo Y, Spudich JL. Three strategically placed hydrogen-bonding residues convert a proton pump into a sensory receptor. *Proc Natl Acad Sci U S A*. 2006. doi:10.1073/pnas.0607467103

40. Mulkidjanian AY, Galperin MY, Makarova KS, Wolf YI, Koonin E V. Evolutionary primacy of sodium bioenergetics. *Biol Direct*. 2008. doi:10.1186/1745-6150-3-13

41. Shalaeva DN, Galperin MY, Mulkidjanian AY. Eukaryotic G protein-coupled receptors as descendants of prokaryotic sodium-translocating rhodopsins. *Biol Direct*. 2015. doi:10.1186/s13062-015-0091-4

42. Kolbe M, Besir H, Essen LO, Oesterhelt D. Structure of the light-driven chloride pump halorhodopsin at 1.8 Å Resolution. *Science (80-)*. 2000. doi:10.1126/science.288.5470.1390

43. Flores-Uribe J, Hevroni G, Ghai R, et al. Heliorhodopsins are absent in diderm (Gram-negative) bacteria: Some thoughts and possible implications for activity. *Environmental Microbiology Reports*. 2019.

44. Engelhard C, Chizhov I, Siebert F, Engelhard M. Microbial Halorhodopsins: Light-Driven Chloride Pumps. *Chem Rev*. 2018. doi:10.1021/acs.chemrev.7b00715

45. Mukohata Y, Kaji Y. Light-induced membrane-potential increase, ATP synthesis, and proton uptake in Halobacterium halobium R1mR catalyzed by halorhodopsin: Effects of N,N′-dicyclohexylcarbodiimide, triphenyltin chloride, and 3,5-di-tert-butyl-4-hydroxybenzylidenemalononitrile. *Arch Biochem Biophys*. 1981. doi:10.1016/0003-9861(81)90067-9

46. Bivin DB, Stoeckenius W. Photoactive retinal pigments in haloalkaliphilic bacteria. *J Gen Microbiol*. 1986. doi:10.1099/00221287-132-8-2167

47. Kandori H. Ion-pumping microbial rhodopsins. *Front Mol Biosci*. 2015. doi:10.3389/fmolb.2015.00052

48. Sasaki T, Aizawa T, Kamiya M, et al. Effect of chloride binding on the thermal trimer-monomer conversion of halorhodopsin in the solubilized system. *Biochemistry*. 2009. doi:10.1021/bi901380c

49. Mukohata Y, Ihara K, Tamura T, Sugiyama Y. Halobacterial rhodopsins. *J Biochem*. 1999. doi:10.1093/oxfordjournals.jbchem.a022332

50. Essen LO. Halorhodopsin: Light-driven ion pumping made simple? *Curr Opin Struct Biol*. 2002. doi:10.1016/S0959-440X(02)00356-1

51. Sasaki J, Brown LS, Chon YS, et al. Conversion of bacteriorhodopsin into a chloride ion pump. *Science (80-)*. 1995. doi:10.1126/science.7604281

52. Hasemi T, Kikukawa T, Kamo N, Demura M. Characterization of a cyanobacterial chloride-pumping rhodopsin and its conversion into a proton pump. *J Biol Chem*. 2016. doi:10.1074/jbc.M115.688614

53. Rüdiger M, Oesterhelt D. Specific arginine and threonine residues control anion binding and transport in the light-driven chloride pump halorhodopsin. *EMBO J*. 1997. doi:10.1093/emboj/16.13.3813

54. Kouyama T, Kanada S, Takeguchi Y, Narusawa A, Murakami M, Ihara K. Crystal Structure of the Light-Driven Chloride Pump Halorhodopsin from Natronomonas pharaonis. *J Mol Biol*. 2010. doi:10.1016/j.jmb.2009.11.061

55. Steiner M, Oesterhelt D, Ariki M, Lanyi JK. Halide binding by the purified halorhodopsin chromoprotein. I. Effects on the chromophore. *J Biol Chem*. 1984. doi:10.1016/s0021-9258(17)43334-5

56. Scharf B, Engelhard M. Blue Halorhodopsin from Natronobacterium pharaonis: Wavelength Regulation by Anions. *Biochemistry*. 1994. doi:10.1021/bi00187a002

57. Schobert B, Lanyi JK, Cragoe EJ. Evidence for a halide-binding site in halorhodopsin. *J Biol Chem*. 1983. doi:10.1016/s0021-9258(17)43787-2

58. Lanyi JK. Halorhodopsin: a light-driven chloride ion pump. *Annu Rev Biophys Biophys Chem*. 1986. doi:10.1146/annurev.bb.15.060186.000303

59. Bálint Z, Lakatos M, Ganea C, Lanyi JK, Váró G. The Nitrate Transporting Photochemical Reaction Cycle of the Pharaonis Halorhodopsin. *Biophys J*. 2004. doi:10.1016/S0006-3495(04)74234-4

60. Ludmann K, Ibron G, Lanyi JK, Váró G. Charge motions during the photocycle of pharaonis halorhodopsin. *Biophys J*. 2000. doi:10.1016/S0006-3495(00)76653-7

61. Hegemann P, Oesterbelt D, Steiner M. The photocycle of the chloride pump halorhodopsin. I: Azide-catalyzed deprotonation of the chromophore is a side reaction of photocycle intermediates inactivating the pump. *EMBO J*. 1985. doi:10.1002/j.1460-2075.1985.tb03937.x

62. Kikukawa T, Kamo N, Demura M. Photochemistry of halorhodopsin. In: *Optogenetics: Light-Sensing Proteins and Their Applications*. ; 2015. doi:10.1007/978-4-431-55516-2_4

63. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nat Neurosci*. 2015. doi:10.1038/nn.4091

64. Raimondo J V., Kay L, Ellender TJ, Akerman CJ. Optogenetic silencing strategies differ in their effects on inhibitory synaptic transmission. *Nat Neurosci*. 2012. doi:10.1038/nn.3143

65. Liu X, Tonegawa S. Optogenetics 3.0. *Cell*. 2010. doi:10.1016/j.cell.2010.03.019

66. Shibata M, Inoue K, Ikeda K, et al. Oligomeric states of microbial rhodopsins

determined by high-speed atomic force microscopy and circular dichroic spectroscopy. *Sci Rep*. 2018. doi:10.1038/s41598-018-26606-y

67. Kim K, Kwon SK, Jun SH, et al. Crystal structure and functional characterization of a light-driven chloride pump having an NTQ motif. *Nat Commun*. 2016. doi:10.1038/ncomms12677

68. Inoue K, Koua FHM, Kato Y, Abe-Yoshizumi R, Kandori H. Spectroscopic study of a light-driven chloride ion pump from marine bacteria. *J Phys Chem B*. 2014. doi:10.1021/jp507219q

69. Koua FHM, Kandori H. Light-induced structural changes during early photo-intermediates of the eubacterial Cl- pump Fulvimarina rhodopsin observed by FTIR difference spectroscopy. *RSC Adv*. 2015. doi:10.1039/c5ra19363j

70. Harris A, Saita M, Resler T, et al. Molecular details of the unique mechanism of chloride transport by a cyanobacterial rhodopsin. *Phys Chem Chem Phys*. 2018. doi:10.1039/c7cp06068h

71. Niho A, Yoshizawa S, Tsukamoto T, et al. Demonstration of a Light-Driven SO42- Transporter and Its Spectroscopic Characteristics. *J Am Chem Soc*. 2017. doi:10.1021/jacs.6b12139

72. Yun JH, Park JH, Jin Z, et al. Structure-Based Functional Modification Study of a Cyanobacterial Chloride Pump for Transporting Multiple Anions. *J Mol Biol*. 2020. doi:10.1016/j.jmb.2020.07.016

73. Hasemi T, Kikukawa T, Watanabe Y, et al. Photochemical study of a cyanobacterial chloride-ion pumping rhodopsin. *Biochim Biophys Acta - Bioenerg*. 2019. doi:10.1016/j.bbabio.2018.12.001

74. Hasemi T. Functional analysis of a novel cyanobacterial rhodopsin.

75. Takeda K, Sato H, Hino T, et al. A novel three-dimensional crystal of bacteriorhodopsin obtained by successive fusion of the vesicular assemblies. *J Mol Biol*. 1998. doi:10.1006/jmbi.1998.2103

76. Faham S, Bowie JU. Bicelle crystallization: A new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J Mol Biol*. 2002. doi:10.1006/jmbi.2001.5295

77. Caffrey M. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallogr Sect FStructural Biol Commun*. 2015. doi:10.1107/S2053230X14026843

78. Cherezov V, Peddi A, Muthusubramaniam L, Zheng YF, Caffrey M. A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases. *Acta Crystallogr Sect D Biol Crystallogr*. 2004. doi:10.1107/S0907444904019109

79. Cherezov V, Caffrey M. Membrane protein crystallization in lipidic mesophases. A mechanism study using X-ray microdiffraction. *Faraday Discuss*. 2007. doi:10.1039/b618173b

80. Cherezov V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc*. 2009. doi:10.1038/nprot.2009.31

81. Liu W, Cherezov V. Crystallization of membrane proteins in lipidic mesophases. *J Vis Exp*. 2011. doi:10.3791/2501

82. Lee SC, Bennett BC, Hong WX, et al. Steroid-based facial amphiphiles for

stabilization and crystallization of membrane proteins. *Proc Natl Acad Sci U S A*. 2013. doi:10.1073/pnas.1221442110

83. Nikolaev M, Round E, Gushchin I, et al. Integral Membrane Proteins Can Be Crystallized Directly from Nanodiscs. *Cryst Growth Des*. 2017. doi:10.1021/acs.cgd.6b01631

84. Polovinkin V, Gushchin I, Sintsov M, et al. High-Resolution Structure of a Membrane Protein Transferred from Amphipol to a Lipidic Mesophase. *J Membr Biol*. 2014. doi:10.1007/s00232-014-9700-x

85. Cherezov V, Rosenbaum DM, Hanson MA, et al. High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. *Science (80-)*. 2007. doi:10.1126/science.1150577

86. Jaakola VP, Griffith MT, Hanson MA, et al. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science (80-)*. 2008. doi:10.1126/science.1164772

87. Wu B, Chien EYT, Mol CD, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science (80-)*. 2010. doi:10.1126/science.1194396

88. Chien EYT, Liu W, Zhao Q, et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science (80-)*. 2010. doi:10.1126/science.1197410

89. Gmelin W, Zeth K, Efremov R, Heberle J, Tittor J, Oesterhelt D. The Crystal Structure of the L1 Intermediate of Halorhodopsin at 1.9 Å Resolution†. *Photochem Photobiol*. 2007. doi:10.1562/2006-06-23-ra-947

90. Schreiner M, Schlesinger R, Heberle J, Niemann HH. Structure of Halorhodopsin from Halobacterium salinarum in a new crystal form that imposes little restraint on the E-F loop. *J Struct Biol*. 2015. doi:10.1016/j.jsb.2015.04.010

91. Schreiner M, Schlesinger R, Heberle J, Niemann HH. Crystal structure of Halobacterium salinarum halorhodopsin with a partially depopulated primary chloride-binding site. *Acta Crystallogr Sect Struct Biol Commun*. 2016. doi:10.1107/S2053230X16012796

92. Kanada S, Takeguchi Y, Murakami M, Ihara K, Kouyama T. Crystal structures of an o-like blue form and an anion-free yellow form of pharaonis halorhodopsin. *J Mol Biol*. 2011. doi:10.1016/j.jmb.2011.08.021

93. Váró G, Brown LS, Needleman R, Lanyi JK. Proton transport by halorhodopsin. *Biochemistry*. 1996. doi:10.1021/bi9601159

94. Kulcsár Á, Groma GI, Lanyi JK, Váró G. Characterization of the proton-transporting photocycle of pharaonis halorhodopsin. *Biophys J*. 2000. doi:10.1016/S0006-3495(00)76508-8

95. Kouyama T, Kawaguchi H, Nakanishi T, Kubo H, Murakami M. Crystal Structures of the L1, L2, N, and O States of pharaonis Halorhodopsin. *Biophys J*. 2015. doi:10.1016/j.bpj.2015.04.027

96. Chan SK, Kawaguchi H, Kubo H, et al. Crystal Structure of the 11-cis Isomer of Pharaonis Halorhodopsin: Structural Constraints on Interconversions among Different Isomeric States. *Biochemistry*. 2016. doi:10.1021/acs.biochem.6b00277

97. Hosaka T, Yoshizawa S, Nakajima Y, et al. Structural mechanism for light-

driven transport by a new type of chloride ion pump, nonlabens marinus rhodopsin-3. *J Biol Chem*. 2016. doi:10.1074/jbc.M116.728220

98. Yun JH, Li X, Park JH, et al. Non-cryogenic structure of a chloride pump provides crucial clues to temperature-dependent channel transport efficiency. *J Biol Chem*. 2019. doi:10.1074/jbc.RA118.004038

99. Yun JH, Ohki M, Park JH, et al. Pumping mechanism of NM-R3, a light-driven bacterial chloride importer in the rhodopsin family. *Sci Adv*. 2020. doi:10.1126/sciadv.aay2042

100. Yun JH, Li X, Yue J, et al. Early-stage dynamics of chloride ion–pumping rhodopsin revealed by a femtosecond X-ray laser. *Proc Natl Acad Sci U S A*. 2021. doi:10.1073/pnas.2020486118

101. Besaw JE, Ou WL, Morizumi T, et al. The crystal structures of a chloride-pumping microbial rhodopsin and its proton-pumping mutant illuminate proton transfer determinants. *J Biol Chem*. 2020. doi:10.1074/jbc.RA120.014118

102. Inoue H, Nojima H, Okayama H. High efficiency transformation of Escherichia coli with plasmids. *Gene*. 1990. doi:10.1016/0378-1119(90)90336-P

103. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*. 2005. doi:10.1016/j.pep.2005.01.016

104. Li D, Boland C, Aragao D, Walsh K, Caffrey M. Harvesting and cryo-cooling crystals of membrane proteins grown in lipidic mesophases for structure determination by macromolecular crystallography. *J Vis Exp*. 2012. doi:10.3791/4001

105. Von Stetten D, Giraud T, Carpentier P, et al. In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr Sect D Biol Crystallogr*. 2015. doi:10.1107/S139900471401517X

106. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010. doi:10.1107/S0907444909047337

107. Winn MD, Ballard CC, Cowtan KD, et al. Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr*. 2011. doi:10.1107/S0907444910045749

108. Vagin A, Teplyakov A. MOLREP: An Automated Program for Molecular Replacement. *J Appl Crystallogr*. 1997. doi:10.1107/S0021889897006766

109. Källberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc*. 2012. doi:10.1038/nprot.2012.085

110. Murshudov GN, Skubák P, Lebedev AA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*. 2011;67(Pt 4):355-367. doi:10.1107/S0907444911001314

111. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr*. 2004;60(12):2126-2132. doi:10.1107/S0907444904019158

112. Adams PD, Afonine P V., Bunkóczi G, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr*. 2010. doi:10.1107/S0907444909052925

113. Ho BK, Gruswitz F. HOLLOW: Generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct Biol*. 2008. doi:10.1186/1472-6807-8-49

114. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res*. 2012;40(D1). doi:10.1093/nar/gkr703

115. Kovalev K, Astashkin R, Gushchin I, et al. Molecular mechanism of light-driven sodium pumping. *Nat Commun*. 2020:1-11. doi:10.1038/s41467-020-16032-y

116. Kovalev K, Polovinkin V, Gushchin I, et al. Structure and mechanisms of sodium-pumping KR2 rhodopsin. *Sci Adv*. 2019. doi:10.1126/sciadv.aav2671

117. Gushchin I, Shevchenko V, Polovinkin V, et al. Crystal structure of a light-driven sodium pump. *Nat Struct Mol Biol*. 2015;22(5):390-395. doi:10.1038/nsmb.3002

118. Shevchenko V, Mager T, Kovalev K, et al. Inward H+ pump xenorhodopsin: Mechanism and alternative optogenetic approach. *Sci Adv*. 2017. doi:10.1126/sciadv.1603187

119. Gushchin I, Melnikov I, Polovinkin V, et al. Mechanism of transmembrane signaling by sensor histidine kinases. *Science (80-)*. 2017. doi:10.1126/science.aah6345

120. Schobert B, Cupp-Vickery J, Hornak V, Smith SO, Lanyi JK. Crystallographic structure of the K intermediate of bacteriorhodopsin: Conservation of free energy after photoisomerization of the retinal. *J Mol Biol*. 2002. doi:10.1016/S0022-2836(02)00681-2

121. Ono H, Inoue K, Abe-Yoshizumi R, Kandori H. FTIR spectroscopy of a light-driven compatible sodium ion-proton pumping rhodopsin at 77 K. *J Phys Chem B*. 2014. doi:10.1021/jp500756f

122. Otomo A, Mizuno M, Inoue K, Kandori H, Mizutani Y. Allosteric Communication with the Retinal Chromophore upon Ion Binding in a Light-Driven Sodium Ion-Pumping Rhodopsin. *Biochemistry*. 2020. doi:10.1021/acs.biochem.9b01062

123. Kato HE, Kamiya M, Sugo S, et al. Atomistic design of microbial opsin-based blue-shifted optogenetics tools. *Nat Commun*. 2015. doi:10.1038/ncomms8177

124. Shionoya T, Singh M, Mizuno M, Kandori H, Mizutani Y. Strongly Hydrogen-Bonded Schiff Base and Adjoining Polyene Twisting in the Retinal Chromophore of Schizorhodopsins. *Biochemistry*. 2021. doi:10.1021/acs.biochem.1c00529

125. Steiner M, Oesterhelt D. Isolation and properties of the native chromoprotein halorhodopsin. *EMBO J*. 1983. doi:10.1002/j.1460-2075.1983.tb01595.x

126. Sudo Y, Ihara K, Kobayashi S, et al. A Microbial rhodopsin with a unique retinal composition shows both sensory rhodopsin II and bacteriorhodopsin-like properties. *J Biol Chem*. 2011. doi:10.1074/jbc.M110.190058

5. Publications

During the doctoral studies, 11 scientific papers were published:

Kovalev, K.*, D. Volkov*, **R. Astashkin***, A. Alekseev, I. Gushchin, J. M. Haro-Moreno, I. Chizhov, S. Siletsky, M. Mamedov, A. Rogachev, T. Balandin, V. Borshchevskiy, A. Popov, G. Bourenkov, E. Bamberg, F. Rodriguez-Valera, G. Büldt, and V. Gordeliy. 2020. "High-Resolution Structural Insights into the Heliorhodopsin Family." ***Proceedings of the National Academy of Sciences of the United States of America***. doi:10.1073/pnas.1915888117.

*authors contributed equally

Kovalev, Kirill, **Roman Astashkin**, Ivan Gushchin, Philipp Orekhov, Dmytro Volkov, Egor Zinovev, Egor Marin, Maksim Rulev, Alexey Alekseev, Antoine Royant, Philippe Carpentier, Svetlana Vaganova, Dmitrii Zabelskii, Christian Baeken, Ilya Sergeev, Taras Balandin, Gleb Bourenkov, Xavier Carpena, Roeland Boer, Nina Maliar, Valentin Borshchevskiy, Georg Büldt, Ernst Bamberg, and Valentin Gordeliy. 2020. "Molecular Mechanism of Light-Driven Sodium Pumping." ***Nature Communications***. Springer US, 1–11. doi:10.1038/s41467-020-16032-y.

Alleva, C., K. Kovalev, **R. Astashkin**, M. I. Bernd, C. Baeken, T. Balandin, V. Gordeliy, Ch Fahlke, and J. P. MacHtens. 2020. "Na+-Dependent Gate Dynamics and Electrostatic Attraction Ensure Substrate Coupling in Glutamate Transporters." *Science Advances*. doi:10.1126/sciadv.aba9854.

Volkov, Oleksandr, Kirill Kovalev, Vitaly Polovinkin, Valentin Borshchevskiy, Christian Bamann, **Roman Astashkin**, Egor Marin, Alexander Popov, Taras

Balandin, Dieter Willbold, Georg Büldt, Ernst Bamberg, and Valentin Gordeliy. 2017. “Structural Insights into Ion Conduction by Channelrhodopsin 2.” ***Science***. doi:10.1126/science.aan8862.

Bratanov, Dmitry, Kirill Kovalev, Jan Philipp Machtens, **Roman Astashkin**, Igor Chizhov, Dmytro Soloviov, Dmytro Volkov, Vitaly Polovinkin, Dmitrii Zabelskii, Thomas Mager, Ivan Gushchin, Tatyana Rokitskaya, Yuri Antonenko, Alexey Alekseev, Vitaly Shevchenko, Natalya Yutin, Riccardo Rosselli, Christian Baeken, Valentin Borshchevskiy, Gleb Bourenkov, Alexander Popov, Taras Balandin, Georg Büldt, Dietmar J. Manstein, Francisco Rodriguez-Valera, Christoph Fahlke, Ernst Bamberg, Eugene Koonin, and Valentin Gordeliy. 2019. “Unique Structure and Function of Viral Rhodopsins.” ***Nature Communications***. doi:10.1038/s41467-019-12718-0.

Kovalev, Kirill, Vitaly Polovinkin, Ivan Gushchin, Alexey Alekseev, Vitaly Shevchenko, Valentin Borshchevskiy, **Roman Astashkin**, Taras Balandin, Dmitry Bratanov, Svetlana Vaganova, Alexander Popov, Vladimir Chupin, Georg Büldt, Ernst Bamberg, and Valentin Gordeliy. 2019. “Structure and Mechanisms of Sodium-Pumping KR2 Rhodopsin.” ***Science Advances***. doi:10.1126/sciadv.aav2671.

Zabelskii, Dmitrii, Alexey Alekseev, Kirill Kovalev, Vladan Rankovic, Taras Balandin, Dmytro Soloviov, Dmitry Bratanov, Ekaterina Savelyeva, Elizaveta Podolyak, Dmytro Volkov, Svetlana Vaganova, **Roman Astashkin**, Igor Chizhov, Natalia Yutin, Maksim Rulev, Alexander Popov, Ana Sofia Eria-Oliveira, Tatiana Rokitskaya, Thomas Mager, Yuri Antonenko, Riccardo Rosselli, Grigoriy Armeev, Konstantin Shaitan, Michel Vivaudou, Georg Büldt, Andrey Rogachev, Francisco Rodriguez-Valera, Mikhail Kirpichnikov, Tobias Moser, Andreas Offenhäusser,

Dieter Willbold, Eugene Koonin, Ernst Bamberg, and Valentin Gordeliy. 2020. “Viral Rhodopsins 1 Are an Unique Family of Light-Gated Cation Channels.” ***Nature Communications***. doi:10.1038/s41467-020-19457-7.

Zabelskii, Dmitrii, Natalia Dmitrieva, Oleksandr Volkov, Vitaly Shevchenko, Kirill Kovalev, Taras Balandin, Dmytro Soloviov, **Roman Astashkin**, Egor Zinovev, Alexey Alekseev, Ekaterina Round, Vitaly Polovinkin, Igor Chizhov, Andrey Rogachev, Ivan Okhrimenko, Valentin Borshchevskiy, Vladimir Chupin, Georg Büldt, Natalia Yutin, Ernst Bamberg, Eugene Koonin, and Valentin Gordeliy. 2021. “Structure-Based Insights into Evolution of Rhodopsins.” ***Communications Biology*** 4 (1). Springer US. doi:10.1038/s42003-021-02326-4.

Maliar, Nina, Kirill Kovalev, Christian Baeken, Taras Balandin, **Roman Astashkin**, Maksim Rulev, Alexey Alekseev, Nikolay Ilyinsky, Andrey Rogachev, Vladimir Chupin, Dmitry Dolgikh, Mikhail Kirpichnikov, and Valentin Gordeliy. 2020. “Crystal Structure of the N112A Mutant of the Light-Driven Sodium Pump KR2.” ***Crystals***. doi:10.3390/cryst10060496.

Remeeva, Alina, Vera V. Nazarenko, Kirill Kovalev, Ivan M. Goncharov, Anna Yudenko, **Roman Astashkin**, Valentin Gordeliy, and Ivan Gushchin. 2021. “Insights into the Mechanisms of Light-Oxygen-Voltage Domain Color Tuning from a Set of High-Resolution X-Ray Structures.” ***Proteins: Structure, Function and Bioinformatics***. doi:10.1002/prot.26078.

Varaksa, Tatsiana, Sergey Bukhdruker, Irina Grabovec, Egor Marin, Anton Kavaleuski, Anastasiia Gusach, Kirill Kovalev, Ivan Maslov, Aleksandra Luginina, Dmitrii Zabelskii, **Roman Astashkin**, Mikhail Shevtsov, Sviatlana Smolskaya, Anna Kavaleuskaya, Polina Shabunya, Alexander Baranovsky, Vladimir Dolgopalets, Yury Charnou, Aleh Savachka, Raisa Litvinovskaya, Alaksiej Hurski,

Evgeny Shevchenko, Andrey Rogachev, Alexey Mishin, Valentin Gordeliy, Andrei Gabrielian, Darrell E. Hurt, Boris Nikonenko, Konstantin Majorov, Alexander Apt, Alex Rosenthal, Andrei Gilep, Valentin Borshchevskiy, and Natallia Strushkevich. 2021. “Metabolic Fate of Human Immunoactive Sterols in Mycobacterium Tuberculosis.” ***Journal of Molecular Biology***. doi:10.1016/j.jmb.2020.166763.