博　士　論　文

(Doctoral Thesis)

論　文　題　名(The Thesis Title)

Solid-State NMR Methodologies for Membrane Proteins and Application to a Halorhodopsin

膜蛋白質の固体 NMR 解析の方法論とハロロドプシンへの展開

Date: November, 2024

専 攻 名 (Department)	Biological Science
氏　　名(Name)	Xin ZHANG

大阪大学大学院理学研究科

(Graduate School of Science OSAKA UNIVERSITY)

Abstract

Membrane proteins, which account for 20–30% of genomically encoded proteomes, are fundamental to a variety of important life activities. Despite their important roles in life science and extensive research interest, understanding of membrane proteins' structure-function relationship remains insufficient, as evidenced by their underrepresentation in the structural database. For example, less than 2% of the protein structures deposited in the Protein Data Bank (PDB) are membrane proteins. The difficulty of structural study for membrane proteins stems from their physicochemical properties, particularly their integration within lipid bilayers. Solid-state NMR with magic-angle spinning (MAS ssNMR) emerges as a promising alternative technique for the structural analysis of membrane proteins assembled in lipid bilayers at atomic resolution under near-physiological conditions. Although MAS ssNMR offers unprecedented access to structural and dynamic information, it is still a new method and faces challenges related to sensitivity and resolution, necessitating advanced sensitivity enhancement strategies and new sample preparation protocols.

In this study, I worked on the structural analysis of *Natronomonas pharaonis* halorhodopsin (*Np*HR) an archetypical Cl^- ion inward pump derived from thermophilic archaea. *Np*HR inwardly pumps Cl^- into cells for regulating osmotic pressure with light irradiation and has been widely used as an optogenetic tool. Natively, *Np*HR forms a homotrimer in cell membrane and each monomer contains seven transmembrane helices (namely helix A-G) with a retinal chromophore binding to the sidechain of the Lys256 residue through Schiff base.

The light driven Cl^- transport function of *Np*HR is coupled with a cyclic conformational change which is called photocycle. The photocycle is generated by the light-induced isomerization of the retinal, leading to a series of structural changes in the whole protein that facilitate Cl^- transport and ended by the restoration of the ground state. Despite extensive X-ray crystallographic studies having been performed and revealed the molecular details of the ion release step in the early photocycle, the Cl^- uptake mechanism involved in the late photocycle remains unclear.

The mechanistic uncertainty becomes apparent when examining the crystal structures of *Np*HR both in the ground Cl^--bound state and the Cl^--free state, the latter representing an established mimic of the photocycle's O intermediate. Both structures display a similar closed conformation, contradicting the anticipated open conformation preceding the Cl^- uptake. Moreover, the extracellular side of this O-mimic state is entirely obstructed by the large extracellular loop region (B-C loop), making it difficult to identify any Cl^- uptake pathway. Crystal packing, which restricts large molecular motion, may contribute to the observed limited structural differences. This situation highlights the uniqueness of ssNMR analysis under near physiological lipid bilayer conditions for elucidating the Cl^- uptake process in the late photocycle in debate.

In the first part of this study (Section 2). I first developed a refined unfolding and refolding protocol for preparing the perdeuterated *Np*HR with enhanced amide proton (^1H_N) back exchange in its native trimeric state for ^1H detection ssNMR. While this perdeuteration and ^1H_N back exchange step proved essential for obtaining high resolution ^1H detection ssNMR data, achieving sufficient ^1H_N back exchange represents a challenge in membrane proteins due

to limited solvent accessibility, resulting in severe signal losses. My protocol, although retaining some signal losses, significantly improved the $^{1}H_{N}$ back exchange efficiency. This enabled me to obtain 3D CANH spectra with enhanced information content using approximately 1 mg sample. I observeed numerous peak shifts in the spectra between the Cl^--bound and -free states, which indicated significant structural modifications between the two states widely distributing in the molecule.

In the second part of this study (Section 3), I performed signal assignments for *Np*HR using ^{13}C detection ssNMR and successfully assigned residues in major part of helix D, the C-D and E-F loop and some residues in other helices. Through chemical shift perturbation (CSP) analysis, which reflects local environment changes, I have revealed that one of the significant structural changes was located at the extracellular (EC) part of helix D between the Cl^--bound and -free states.

Based on the above data, I examined in the final part of this study (Section 4) the functional implications on the identified structural change to the Cl^- uptake process through targeted mutations. I prepared point mutants of *Np*HR for residues in the EC part of Helix D and analyzed their molecular properties. Notably, the Ala to Val mutant at position 165 (A165V) exhibited distinct Cl^- uptake characteristics, including lower Cl^- binding affinity and a significantly accelerated late photocycle kinetics compared to the wild type. This mutation induced trimer dissociation both in the Cl^--free state and following photocycle activation. These results suggest that an open conformation, which pushes adjacent monomers apart, exists in the Cl^--free state or during the O intermediate. DNP-enhanced ^{15}N detection ssNMR measurement confirmed that the A165V mutant retains protein folding through preserved Schiff base ^{15}N chemical shift, supporting a model where the EC part of helix D undergoes outward movement during O intermediate formation. This movement may be critical for the Cl^- uptake process, facilitating the opening of the Cl^- entrance on the extracellular side, which was not observed in X-ray structures.

This study significantly advances understanding of the Cl^- uptake mechanism of *Np*HR in the late photocycle, while demonstrating the utility of ssNMR for examining structural changes of membrane proteins under near-physiological lipid bilayer conditions.

Abbreviations

Abbreviations	Definition
PDB	Protein Data Bank
MAS	magic-angle spinning
ssNMR	solid-state NMR
*Np*HR	*Natronomonas pharaonis* halorhodopsin
pSB	protonated Schiff Base
bR	bacteriorhodopsin
CP	cytoplasmic
EC	extracellular
ASR	*Anabaena* Sensory Rhodopsin
DNP	dynamic nuclear polarization
CD	circular dichroism
P/L ratio	protein to lipid molar ratio
SEC	size-exclusion chromatography

Contents

1 General introduction

Membrane proteins account for 20-30% of the total proteins encoded by the genome, play critical roles in various functions essential for life, including molecule transport, signal transduction, and enzymatic reactions. The molecular basis of these functions is often associated with the conformational changes of membrane proteins, which modify their molecular properties. Thus, understanding the structural dynamics of membrane proteins is crucial for elucidating their functions and represents an important area of life sciences. For example, ~50 % of modern pharmaceutical drugs target on membrane proteins[1]. Despite their critical roles in life activities and substantial study interest of them, the understanding for membrane proteins in their structures and functions remains insufficient.

The most widely employed X-ray crystallography contributed a lot in membrane protein research, over 80 % of the membrane protein structures deposited in Protein Data Bank (PDB)[2] were solved by X-ray crystallography, including many key proteins such as aquaporin[3], the water channel and ATP-synthase[4]. However, membrane proteins only account for less than 2 % of the total protein structures deposited in PDB[2], highlighting the challenges for membrane protein structural study using X-ray crystallography. The unique molecular properties of membrane proteins contribute to this insufficient understanding of their structural information. They contain the highly hydrophobic lipid bilayer inserted regions as well as hydrophilic regions exposed to extracellular fluid or cytoplasm. Such structural complexity makes membrane proteins hard to be purified and crystallized for X-ray crystallography structural analysis. Besides the difficulty in membrane protein sample preparation, crystallization itself can introduce structural artifacts to due to crystal packing effect, which can pose a challenge in detailed structural analysis of membrane proteins[5,6].

Cryo-EM, the repaid developing protein structural analysis method, achieved significant breakthroughs in structural study of membrane proteins[7]. However, the resolution of cryo-EM is still limited to sub-atomic level[8], which restricts its ability to analyze subtle conformational changes of membrane proteins, while such kind of fine structural changes are important to understand the molecular properties of membrane protein. In addition, although some protein with small molecular weight has been analyzed by cryo-EM, the lower limitation remains around 40 kDa[9], constraining its applicability within larger membrane proteins.

Solution NMR spectroscopy, on the other hand, is effective for analyzing small membrane proteins solubilized in detergent micelle with high resolution[10]. Solution NMR also provides important molecular dynamic information[11] which is critical for understanding the movement of membrane protein in relation to their function. However, solution NMR cannot be applied for large proteins or insoluble membrane protein samples. It also needs discussions on structural information obtained in the non-native micelle environment, rather than the physiological lipid bilayer.

Solid-state NMR with magic-angle spinning (MAS ssNMR) is a promising alternative for structural analysis of membrane proteins[12–14]. It allows direct measurement for membrane proteins assembled in lipid bilayer at atomic level under near physiological conditions[15–22] or even direct analysis native cell membrane[23] (in-cell ssNMR) and provides molecular dynamic information together with structural information. In addition, ssNMR is particularly suited for investigating tiny structural behaviors in specific regions of membrane proteins. For example,

ssNMR has been widely used to analyze the Schiff base conformation in microbial rhodopsins[24–29]. ssNMR allows to reveal sub-angstrom conformational difference of the Schiff base which is critical to their function[28]. However, ssNMR is still a developing method in the field of protein structural analysis and facing technique challenges related to sensitivity and resolution.

In this study, I worked on the structural analysis of the halorhodopsin form *Natronomonas pharaonis* (*Np*HR), a Cl^- pump with ssNMR method to elucidate the uncleared structural mechanistic details for the Cl^- uptake process.

1.1 *Natronomonas pharaonis* halorhodopsin

*Np*HR is a purple-colored membrane protein, belongs to microbial rhodopsin (type I rhodopsin) family, found in *Natronomonas pharaonis*, a thermophilic archaeon in 1986[30]. Microbial rhodopsin comprises a large protein family with similar structural properties.[31] They have seven-transmembrane helices (namely Helix A to G) with a retinal binding inside the protein through a Schiff base.[31,32] Microbial rhodopsins are widely distributed across all kinds of microbes and serve diverse functions in signaling, ion transport, enzyme activity.[31,33,34] Although they are involved in completely different life activities, all microbial rhodopsins share the same unique light triggered activity, which means they generate light energy to drive their function.

As for *Np*HR, it is a light-induced Cl^- inward pump.[35] This light induced transport function is critical for regulating the osmotic pressure of *Natronomonas pharaonis* which naturally lives in salt lakes. This light-activated transport function has led to the widespread use of *Np*HR as an optogenetic tool for silencing neurons.[36,37] The retinal chromophore of *Np*HR binds to the side chain amino group of Lys256 in the seventh helix (helix G) via a protonated Schiff base (pSB) (Fig. 1) and serves as the active center. In its native state, *Np*HR forms a homotrimer associated with bacterioruberin, a C50 carotenoid red-colored pigment synthesized by archaea on the cell membrane[38], while adopts a twisted trimer without bacterioruberin when heterologous expressed in other species[39,40].

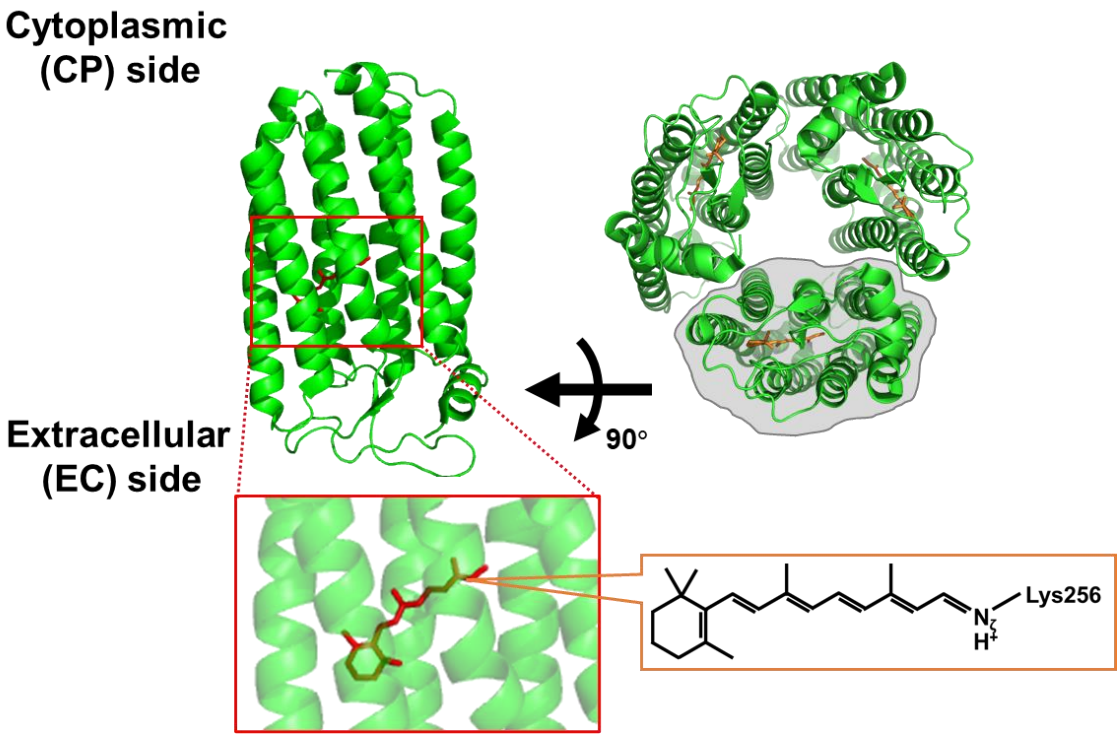

Fig. 1 The crystal structure of *Np*HR in the Cl^- bound state (PDB entry 3A7K[38]). Structural visualization is performed using PyMOL.

The Cl^- transport mechanism of *Np*HR is characterized by a photocycle, which is a common structural and photochemical feature shared for all microbial rhodopsins. The photocycle of microbial rhodopsins was first described for bacteriorhodopsin (bR), the first identified[41] and most extensively studied microbial rhodopsin[42]. The photocycle begins with the photoexcitation of the retinal by the light induced *all-trans* to 13-*cis* isomerization leading to several photochemically distinguishable intermediates and ends by the protein returning to its ground state. In the photocycle of bR, it contains several photo intermediates namely from K to O which distinguished by the retinal absorption changes after photoexcitation (Fig. 2a)[43]. For *Np*HR, the photocycle is described following this manner with several modifications (Fig. 2b)[43–46]. In the photocycle of *Np*HR, it leaks the M intermediate, which contains a deprotonated Schiff base. Additionally, two spectrally silent transients, L_2 and *Np*HR′ states are included, these two intermediates show similar absorption with L_1 and ground states, respectively. However, they are crucial for achieving accurate kinetic analyses for the photocycle of *Np*HR. In this photocycle, the isomerization of retinal and the corresponding structural changes in the whole protein allow the Cl^- ion, which initially binding at the primary binding site near the Schiff base, diffusing from the extracellular (EC) side of the protein to the cytoplasmic (CP) side. Then, the retinal molecule relaxes back into all-tarns form and the Cl- ion diffuses into cytoplasm. Finally, a new Cl^- ion is recaptured into the primary binding site and the whole protein structure returns to the ground state.

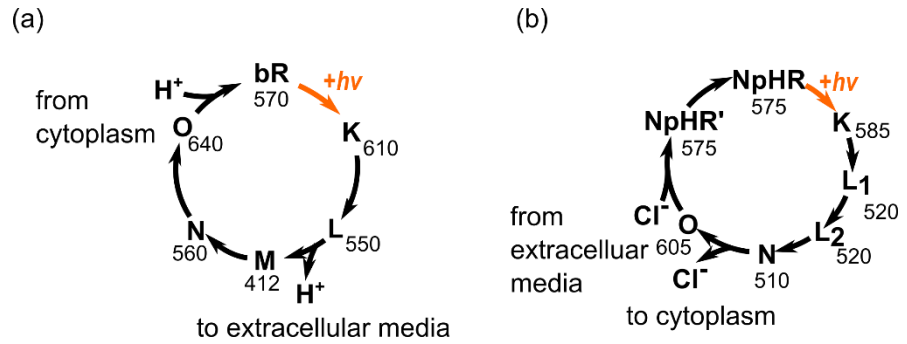

Fig. 2 The photocycle scheme of bR and *Np*HR, the number with each state is the absorption maxima (λ_{max}) of the corresponding state.

In this photocycle, the O intermediate is a Cl^--free intermediate characterized by a red-shifted absorption, while *Np*HR′ is a Cl^--bound state with similar absorption to the ground state. Consequently, the Cl^- uptake process is supposed to occur during the O to *Np*HR′ transition.[46] The N and O intermediates appear and decay simultaneously, that is under a quasi-equilibrium; and the equilibrium of these two states is known Cl^- dependent which means the Cl^- release process happens during the N to O transition.[44,45]

Numerous studies on wild-type (WT) *Np*HR and its mutants have been published and identified key residues for the Cl^- transport process in the photocycle.[44,47–58] Time-resolved X-ray crystallography has provided valuable insights into the early stage of photocycle[56], that is related to the Cl^- release process. The detailed structural changes in the whole protein and key residues for Cl^- release process with the Cl^- ion movement from the primary binding site around pSB to the cytoplasmatic side were traced in crystal[56] and agreed well with the photochemical studies[48–51], showed that the Cl- ion position is different between L_1 and L_2 intermediate[56]. However, the structural details remain still unclear in the late photocycle.

Crystal structures of *Np*HR are available for both the ground state[38] (referred to as the Cl^--bound state) and an O-like chloride-free state[59] (referred to as the Cl^--free state). Although the latter one is known to represent the O intermediate due to their structural similarities[60,61], it appears to be functionally impaired in the crystal from following observations (Fig. 3). First, the large B-C loop region on the extracellular side (EC side) obscures the Cl^- entrance in both the Cl^--bound and Cl^--free states, complicating the identification of the Cl^- uptake pathway. Second, the processes of Cl^- removal and recapture between the Cl^--free state and the ground state were not observed in the crystal. An M-like yellow form, which serves as an intermediate for Cl^- removal and recapture, is necessary for this transition[59]. Third, the transition from the O intermediate back to *Np*HR was not traced in the crystal due to lattice packing forces[56], resulting in a lack of direct structural insights into the Cl^- recapture process.

Additionally, photochemical studies on *Np*HR showed that the photocycle is slowed down under a high-pressure condition[46,49], suggesting large molecular volume changes between the O-intermediate and ground state. The crystal structures between Cl^--bound and -free states only showed structural change in the extracellular part of the helix C. These results suggest that the physiological conformational change between the O-intermediate and ground state is difficult to observe under crystal packing conditions.

For a better understanding of the Cl^- recapture process of *Np*HR and the associated structural changes between the Cl^--bound and Cl^--free state, an analysis at atomic level under

a near-physiological condition is necessary. To this end, I used ssNMR to analyze the structural differences of *Np*HR in its Cl^--bound and Cl^--free state in a near physiological lipid bilayer condition.

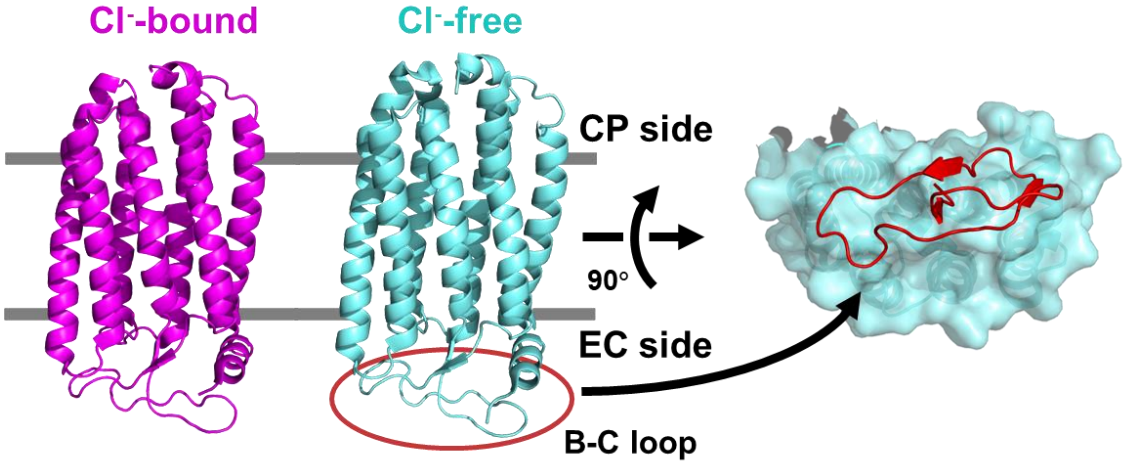

Fig. 3 The crystal structures of *Np*HR in the Cl^--bound (PDB entry: 3A7K[38], purple) and -free (PDB entry 3QBG, cyan[59]) state. The structure visualization is performed using PyMOL.

1.2 Solid-state NMR spectroscopy and its application in membrane protein structural analysis

Nuclear magnetic resonance (NMR) is the phenomenon of nuclei with non-zero nuclear spin quantum number under an applied magnetic field absorb electromagnetic wave at a certain radio frequency (RF) and produce a resonance signal. In most cases, nuclei with 1/2 nuclear spin quantum number are utilized for biomolecular NMR spectroscopy because nuclei with larger nuclear spin quantum number provide much more complex NMR spectra which are difficult to analysis.

Fig. 4 shows a scheme for the generation of NMR signal. The absorption of RF power is generated by the nuclear energy level splits under magnetic field which is called Zeeman split, and the frequency of the resonance signal corresponding to the energy difference between the energy levels. The Zeeman split for the nuclei with 1/2 nuclear spin quantum number is given by $\Delta E=\hbar\gamma B_0$. Therefore, it is sensitive to the local magnetic field. The local magnetic field is not only determined with the applied magnetic field but also affected by the local chemical and structural environment through the magnetic effect of surrounding nuclei and electrons. Therefore, the resonance signal of nuclei under different chemical environment shows in different frequency and makes NMR spectroscopy a powerful tool for molecular structure analysis. However, the frequency of NMR signal is dependent on magnetic field strength, which makes it difficult to direct analysis NMR data obtained under different magnetic fields. Chemical shift, δ, therefore, is introduced, which is described as $\delta = \frac{\nu_{sample}-\nu_{standard}}{\nu_{spectrometer}} \times 10^6$, a standard chemical is used as the zero point of the NMR signal and erases the influence of

magnetic field strength.

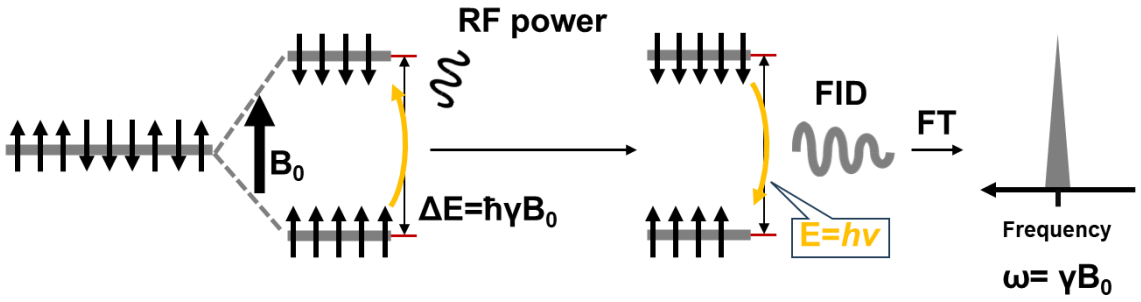

Fig. 4 A scheme for NMR signal generation.

As mentioned above, NMR signal is sensitive to the local magnetic field, while the main nuclear magnetic interactions have orientational dependence as showed in eq. 1[62]:

$$\hat{H}_{CS} = \left(\omega_0\sigma_{iso} + \frac{\omega_0\delta_\sigma}{2}[(3cos^2\theta - 1) - \eta_\sigma sin^2\theta cos2\phi]\right)\cdot \hat{I}_z$$

$$\hat{H}_Q^{(1)} = \frac{eV_{zz}Q}{4I(2I-1)\hbar}[(3cos^2\theta - 1) - \eta_\sigma sin^2\theta cos2\phi]\cdot\left(3\hat{I}_z^2 - \hat{I}^2\right)$$

$$\hat{H}_D^{(1)} = -\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi r_{IS}^3}\frac{(3cos^2\theta - 1)}{2}\cdot\begin{cases}\left[3\hat{I}_{IZ}\hat{I}_{SZ} - \left(\hat{I}_I\cdot\hat{I}_S\right)\right] \text{ if } I = S\\ 2\hat{I}_{IZ}\hat{I}_{SZ} \text{ if } I \neq S\end{cases} \quad \text{eq. 1}$$

where θ and ϕ define characteristic orientation of each spin interaction with respect to the external field direction.

In solution, such anisotropy is largely averaged out by the rapid Brownian motion (Fig. 5, left) allowing high resolution NMR measurement, however it remains in solid state because Brownian motion in solid condition is limited. Therefore, magic-angle spinning (MAS) technique is introduced (Fig. 5, right). MAS is rotating the ssNMR tube against the applied magnetic field at the magic angle (54.7°, the diagonal of cube where $3\cos^2\theta$-1 = 0) with high frequency (usually > 10 kHz). This rotation averages the anisotropic terms, irrespective of their alignment along the x, y, or z axes, thereby enabling high-resolution solid-state NMR measurements. Since the MAS ssNMR method does not rely on rapid Brownian motion to achieve signal narrowing, it significantly broadens the application range of NMR measurements beyond solubilized samples to include various systems, such as protein microcrystals[19], protein fibers[63], proteoliposome[13,16,17], and whole cells[23].

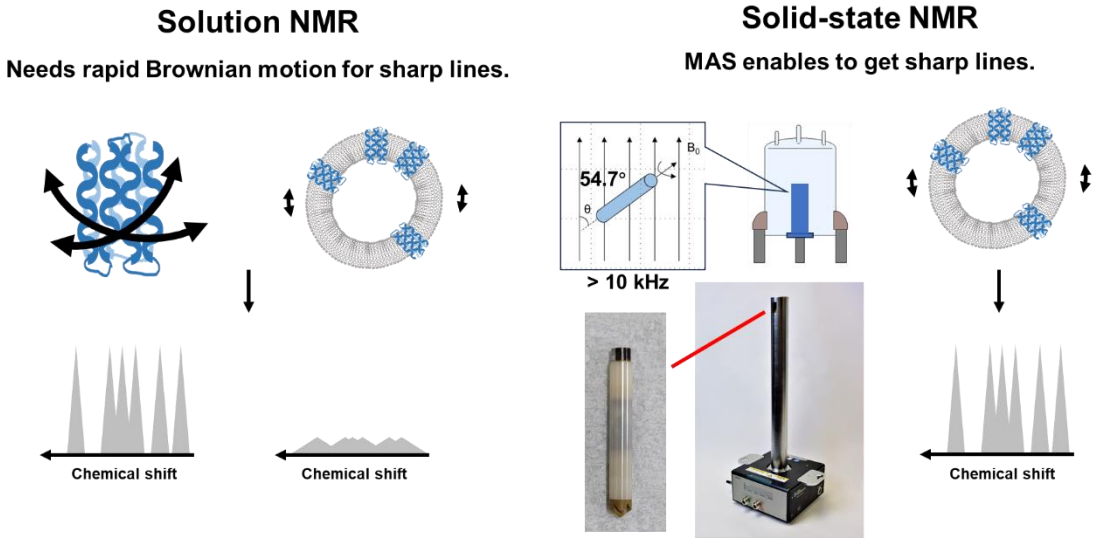

Fig. 5 the scheme of solution NMR and MAS ssNMR difference in averaging orientational dependent anisotropy

Comparing with the other two widely used protein structural analysis method, X-ray crystallography and cryo-EM. ssNMR is unique in its ability to directly measure reconstituted membrane proteins in lipid bilayers after purification or even directly on cell membrane without purification (in-cell NMR). In addition, ssNMR in most cases can be performed under physiological temperature that further minimizes the artificial influences on membrane protein structure. ssNMR also provides molecular dynamic information[64–66], which is valuable for membrane proteins because membrane proteins often perform their function through molecular movements.

These advantages make ssNMR, a powerful candidate, as a tool for membrane protein structural analysis. ssNMR now has been used for various structural studies of membrane proteins and revealed plenty of important features for membrane proteins. For example, the *de novo* structural determination of *Anabaena* Sensory Rhodopsin (ASR) was achieved in lipid bilayer based on accumulatio of sufficient amount of inter-nuclear distance information.[16] It showed that ASR forms a homotrimer in lipid bilayers[16], while in crystal ASR forms a dimer[67]. ssNMR also showed its irreplaceability in many other membrane protein structural studies on the near physiological measurement condition which provides additional structural information and more molecular flexibility compared with X-ray crystallography. Therefore, in this study I used ssNMR method to reveal the uncleared ion uptake mechanism of *Np*HR.

However, ssNMR still has its challenge. The Zeeman split $\Delta E=\hbar\gamma B_0$ is relatively small, while under thermal equilibrium, the population of each energy level follows Boltzmann distribution. Therefore, the population difference between the two energy levels, $P= (N_\alpha-N_\beta)/(N_\alpha+N_\beta) = \tanh(\hbar\gamma B_0/2kT)$ is also very small. Generally, it is about one of ten thousand nuclei under room temperature under > 10 T magnetic field. Therefore, NMR signals are very weak and need a large amount of samples with long measurement time to obtain high quality data. In addition, for proteins the commonly exist nuclei are H, N and C and only ^{1}H is the major isotope for H (99.985%) with 1/2 nuclear spin quantum number and can be detected by commonly used NMR spectroscopy. While ^{12}C (98.89% of C) and ^{14}N (99.64% of N) have

nuclear spin quantum number of 0 and 1, respectively, which are undetectable or not the commonly used nuclei for biomolecular NMR spectroscopy. These further reduce the sensitivity of NMR spectroscopy and made it necessary to use ^{13}C (1.11% of C) and ^{15}N (0.36% of N), which have 1/2 nuclear spin quantum number, enriched protein sample for NMR measurement. Advanced techniques have been developed to cope with this low sensitivity. The Zeeman split is proportional to the strength of applied magnetic field B_0 and gyromagnetic ratio γ, while the larger energy level split means the stronger NMR signal. Therefore, much stronger superconductive magnets are designed to provide stable magnetic field for NMR measurement, nowadays 28.2 T magnetic field was achieved and applied to protein structural studies of ssNMR.[68] The other method is to utilize nucleus with larger γ, however, ^{1}H is the nucleus with largest γ. Then, hyperpolarization by electrons was introduced. Electron has a much large γ, which is about 660 times larger than ^{1}H, therefore by transfer the polarization of electron to nuclei, much higher polarization can be generated on nuclei which means more nuclei can be excited in NMR measurement thus give higher sensitivity. Such technique is called dynamic nuclear polarization (DNP)[69], for protein studies DNP is usually performed under ultra-low temperature to obtain the best sensitivity enhance ratio.[70–75]

1.3 Scope of this study

In this study I aimed to use ssNMR method to reveal the missed structural details of the ion uptake mechanism of *Np*HR in its late photocycle by analyzing the Cl^--bound and Cl^--free state *Np*HR directly in lipid bilayer.

Two conclusions related to *Np*HR Cl^- uptake mechanism are obtained: (1) Structural changes between the Cl^--bound and Cl^--free state *Np*HR were confirmed, which was unclear before. And a major structural change, the outward movement of helix D in the extracellular (EC) part was revealed. (2) This newly identified movement contributed to the Cl^- uptake process of *Np*HR. It helped to form the Cl^- uptake pathway in the EC region, while this pathway was not observed in *Np*HR crystal structures due to crystal packing.

In addition to these biological insights regarding *Np*HR, I also overcame a technical challenge related to sample preparation for ^{1}H detection ssNMR measurement. I developed a novel sample preparation procedure that employs folding and refolding techniques. This method was crucial for achieving a higher ^{1}H back-exchange ratio for perdeuterated *Np*HR samples. Only through this approach were high-quality ^{1}H detection ssNMR spectra obtained, which are essential for structural analysis.

2 ^{1}H detection MAS ssNMR method development and application on *Np*HR structural analysis

2.1 Introduction

^{1}H detection MAS ssNMR is a highly sensitive NMR method because of the large γ ratio

of ^{1}H. Therefore ^{1}H detection MAS ssNMR measurement requires small sample amount (sub milligram scale), while the commonly used ^{13}C detection MAS ssNMR measurement requires ~10 mg sample.

^{1}H detection MAS ssNMR measurement has been used for membrane protein structural analysis from different sides. For example, the T_1 relaxation time analysis, which represents the flexibility for residues, is widely used to find the structural hotspot because the highly flexible region undergoes more conformational changes and may be important to protein function[65]. Titration measurement, which measures a serial of spectra under different conditions and analyzes the chemical shift changes and the peak intensity changes provides valuable information for protein-protein/ligand interaction[76]. ^{1}H detection MAS ssNMR measurement also was applied for membrane protein structural determination. For example, the structure of CXCR1, a GPCR involves immunologic process was solved by ^{1}H detection ssNMR directly in liposome[17].

However, ^{1}H detection MAS ssNMR has a significant limitation: the MAS frequency is often insufficient. MAS can completely average the anisotropic interactions when MAS frequency exceeds the magnitude of the interaction. While for ^{1}H, typical ^{1}H-^{1}H dipolar interaction in proteins which defined as $\frac{\mu_0\gamma^2\hbar}{4\pi r^3}$ (c.a. 20 kHz[77]) is too strong to be completely removed and resulting in severe line broadening in ^{1}H detection MAS ssNMR spectra, make the spectral analysis difficult.

Further increasing the MAS frequency is a potential solution for this problem. However, as the MAS frequency increases, the size of the ssNMR sample rotor must decrease. For example, a 1 mm rotor can be used at 70 kHz MAS, but to achieve 111 kHz MAS, a 0.7 mm rotor is required, which only has half the volume of the 1 mm rotor. Even with such ultra-fast MAS frequencies, line broadening due to ^{1}H-^{1}H dipolar interactions still observed[78]. Therefore, alternative methods need to be explored to address this challenge.

Reducing ^{1}H content is a good idea, as the dipolar interaction decreases with the cube of the inter-nuclear distance. Perdeuteration and $^{1}H_N$ back-exchange at amide site is a commonly employed method to remove all ^{1}H from proteins, while regenerating ^{1}H at the amide sites for ^{1}H detection[79]. However, protein folded structure often prevents the $^{2}H_N$ to be exchanged with solvent ^{1}H, resulting in heavy signal losses. This incomplete $^{1}H_N$ back-exchange becomes a severe problem for membrane proteins because of their hydrophobic transmembrane core.

In this study, I first developed a novel *Np*HR sample preparation method that introduced an unfolding and refolding step to prepare perdeuterated *Np*HR with improved $^{1}H_N$ back-exchange. The unfolding step enhanced water accessibility facilitating a higher ^{1}H back-exchange ratio. Subsequently, the refolding step allowed the ssNMR measurements for that $^{1}H_N$ back-exchanged *Np*HR in its native state. I successfully obtained native-state *Np*HR samples using this method and measured ^{1}H detection ssNMR spectra for both the Cl^--bound and Cl^--free states. The results indicated that conformational changes were widely dispersed throughout the entire protein.

2.2 Materials and methods

2.2.1 *Np*HR expression method

In this study, a truncated wild type *Np*HR at position from 18 to 277 in residue number was used. The truncated *Np*HR gene was amplified from a *Np*HR gene coded in pET-21c(+) vector[53] with primers that have NdeI or SalI recognition sequences through the standard PCR method. Then the fragments were ligated into the NdeI/XhoI site of pET-22b(+) vector. Consequently, a truncated wild type *Np*HR with 6-His-tag at the C terminus was constructed, the final sequence was ME^{18}VTQR…VVSGS^{277}VEHHHHHH.

All *Np*HRs were overexpressed in *E. coli* BL21(DE3) cells. The isotope labeled *Np*HRs for NMR measurements were expressed using M9 minimal media with proper isotope enriched chemicals, and those for unfolding and refolding method development were expressed using LB media without isotope labeling.

For ^{1}H detection ssNMR, all the proteins were uniformly labeled by ^{13}C and ^{15}N. As for hydrogen atoms, ^{1}H labeling with light water-based media (referred to as [U-^{1}H]-*Np*HR) and perdeuterated with H_N back-exchanged *Np*HR expressed in heavy water-based media (referred to as [^{2}H with $^{1}H_N$]-*Np*HR) were used. The uniform ^{15}N labeling was performed by using 1 g $^{15}NH_4Cl$ in 1 L media (or 1 kg heavy water-based media) as the sole nitrogen source for protein expression. While for ^{13}C labeling, when preparing [^{2}H with $^{1}H_N$ back-exchanged]-*Np*HR, 2 g [U-$^{13}C_6$, D_7]-glucose in 1 kg heavy water-based media was used as the sole carbon source; and 2 g [U-$^{13}C_6$]-glucose in 1 L media was used as sole carbon source for preparing [U-^{1}H]-*Np*HR

The *E. coli* cells were grown at 37 ℃ by monitoring of the growth by measuring the OD_{660}. The *Np*HR expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when OD_{660} reached 1.4 (for *E. coli* grew in M9 minimal media) or 1.6 (for *E. coli* grew in LB media). Simultaneously, 10 μM *all-trans* retinal was added exogenously to regenerate the expressed opsins. After 4 h (6 h when using heavy water-based M9 minimal media) of induction, the cells were harvested by centrifugation (4,000 ×*g*, 10 min, 4 ℃), then washed by 50 mM Tris-HCl (pH 8.0) with 5 mM $MgCl_2$ once and pelleted via centrifugation (4,000 ×*g*, 10 min, 4 ℃), stored at -20 ℃.

2.2.2 Native state *Np*HR purification method

The purification procedure was essentially the same as the previously reported one[80] with modifications. In this study, *Np*HR was purified in its native state or in an unfolded state.

To solubilize and purify *Np*HR in its native state, n-dodecyl-β-D-maltoside (DDM) was used. The cells were resuspended using 50 mM Tris-HCl (pH 8.0) with 5 mM $MgCl_2$ and disrupted by ultrasonication. Then, the crude membranes were collected by ultracentrifugation (120,000 ×*g*, 60 min, 4 ℃) and suspended using 50 mM Tris-HCl (pH 8.0) with 100 mM NaCl and 5 mM imidazole. The suspended membrane fraction was solubilized by addition of 1.5 % DDM with stirring overnight at 4 ℃. After solubilization, the insoluble part was removed by ultracentrifugation (120,000 ×*g*, 30 min, 4 ℃). The supernatant containing *Np*HR was incubated with Ni-NTA resin (Qiagen) for 1 h at 4 ℃, and then the resin was transferred into an empty chromatography column. Then the resin was washed by 50 mM Tris-HCl (pH 8.0)

with 100 mM NaCl, 50 mM imidazole and 0.1 % DDM of 5-fold the resin volume to eliminate nonspecifically bound components. Finally, the purple fraction containing *Np*HR was collected by elution using 50 mM Tris-HCl (pH 8.0) with 100 mM NaCl, 500 mM imidazole and 0.1 % DDM.

2.2.3 Unfolded *Np*HR purification method

To obtain the purified *Np*HR in unfolded state, sodium dodecyl sulfate (SDS) was used. The cells were resuspended using 50 mM Tris-HCl (pH 8.0) with 5 mM $MgCl_2$ and disrupted by ultrasonication. Then, the crude membranes were collected by ultracentrifugation (120,000 ×*g*, 60 min, 4 ℃) and suspended using 50 mM Tris-HCl (pH 8.0) with 100 mM NaCl and 5 mM imidazole. The suspended membrane fraction was solubilized and unfolded by addition of 0.5 % SDS with gentle shake at room temperature. After solubilization and unfolding, which can be seen as the solution became clean with a yellow color in several minutes, the insoluble part was removed by ultracentrifugation (120,000 ×*g*, 30 min, 20 ℃). The supernatant containing unfolded *Np*HR was applied to in a chromatography column per-packed with Ni-NTA resin. Then the resin was washed by 50 mM Tris-HCl (pH 8.0) with 100 mM NaCl, 5 mM imidazole and 0.5 % SDS of 5-fold the resin volume to eliminate nonspecifically bound components. Finally, the fraction with unfolded *Np*HR collected by elution using 50 mM Tris-HCl (pH 8.0) with 100 mM NaCl, 50 mM imidazole and 0.5 % SDS.

2.2.4 *Np*HR refolding method

The *Np*HR refolding protocol was described as follows: First, a PD-10 column was used to remove imidazole from the purified SDS unfolded *Np*HR sample through exchanging against 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl and 0.5 % SDS. Then, the SDS unfolded *Np*HR was 1:1 mixed with a refolding buffer which contains 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl, 20 µl *all-trans* retinal, 2 % DMPC and 2 % CHAPS, a mild detergent. Thus, this sample contains 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl, 10 µl *all-trans* retinal, 1 % DMPC, 0.25 % SDS and 1 % CHAPS. Then, the sample was incubated in dark at 30 ℃ for 1 h under gentle shake to mix them thoroughly. After incubation, the sample was packed into a dialysis bag with a 3.5 kDa cut-off membrane and dialysis against 10-fold of 50 mM Tris-HCl (pH 7.0) with 500 mM KCl, 10 µl *all-trans* retinal under room temperature overnight in dark. Then, a second dialysis step using 10 kDa cut-off membrane against 100-fold of 50 mM Tris-HCl (pH 7.0) with 500 mM KCl in dark at 4 ℃ overnight was applied. This proteoliposome was collected by ultracentrifugation (120,000 ×*g*, 30 min, 4 ℃), then annealed by a heat-chill cycle (7 min 30 ℃ water bath and 8 min on ice, 10 cycles) for *Np*HR reconstitution into trimeric form in DMPC liposome. Finally, the reconstituted *Np*HR was solubilized and purified in the native state by DDM in the same manner as described previously in section 2.2.2. The redundant DMPC was removed by this solubilization step and the purified trimeric native state *Np*HR was obtained. The trimeric assembly was examined by visible Circular dichroism (CD) spectroscopy.

The UV-visible spectra were measured with a UV-vis spectrophotometer (SHIMADZU UV1280) at room temperature. The CD spectra were measured using a JASCO J-1500 circular

dichroism spectrophotometer at 25 ℃. All spectra were recorded in the region of 400-700 nm with a scanning speed of 20 nm/min 4 times.

2.2.5 Solution NMR method for unfolding and refolding method evaluation

A 400 MHz (B_0 = 9.4 T) Bruker Avance III NMR spectrometer was used to examine whether SDS can be selectively removed by the first dialysis step in the refolding protocol. Four samples were prepared: (1) 0.25 % SDS solved in 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl; (2) 1 % CHAPS solved in 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl; (3) 0.25 % SDS and 1 % CHAPS solved in 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl; (4) the dialysis product of (3) against 10-fold of 50 mM Tris-HCl (pH 7.0) with 500 mM KCl using 3.5 kDa cut-off membrane overnight. (1) and (2) were used as the reference to obtain the SDS and CHAPS 1D ^{1}H spectra pattern, while (3) represented the buffer condition before dialysis and (4) represented the buffer condition after dialysis. The samples were loaded into standard 5 mm solution NMR sample tubes (~0.5 mL) for measurement.1D ^{1}H spectra of them were measured at 298 K to analyze the detergent concentration before and after dialysis. The NMR data was processed and visualized by JEOL Delta software. The detailed parameters are described in the appendix.

2.2.6 *Np*HR sample preparation for NMR analysis

A PD-10 column was used to remove the imidazole from the purified samples by exchanging against 50 mM Tris-HCl (pH 7.0) with 100 mM NaCl and 0.1 % DDM. After the buffer exchange, *Np*HR protein amount was determined by the absorbance of opsin-bound retinal at 580 nm using the extinction coefficient, ε_{max} of 54,000 $M^{-1}\cdot cm^{-1}$ [81]using an UV-vis spectrophotometer (SHIMADZU UV1280).

To prepare denatured *Np*HR sample for solution NMR analysis. The buffer exchanged samples were dehydrated by freeze-drying and resolubilized using the 1:1 mixture of CD_3OH and $CDCl_3$ (v/v) with 100 mM $LiClO_4$ to the final protein concentration of 100 μM.

To prepare *Np*HR-DMPC proteoliposome for ^{1}H detection ssNMR measurement. 20-fold of 14:0 PC (DMPC) compared to *Np*HR amount (P/L ratio, in molar ratio n_{NpHR}:n_{DMPC} = 1:20), was solved with chloroform in a glass bottle then formed lipid film under dried N_2 or Ar gas flow by evaporating chloroform. The DMPC film was rehydrated with the *Np*HR solution and incubated for 1 h at 30℃ under gentle shake. Then, DDM was removed by incubating the samples with bio-beads (Bio-Rad, 50 mg bio-beads for 1 mg DDM) overnight at 4 ℃ to form *Np*HR-DMPC proteoliposome. The proteoliposome was collected by centrifugation (18,000 ×*g*, 10 min, 4 ℃) and annealed by a heat-chill cycle (7 min 30 °C water bath and 8 min ice bath) for 10 times then washed by 1 mL of 50 mM sodium acetate (pH 5.0) with 100 mM NaCl or 67 mM Na_2SO_4 for 5 times and resuspended in the same buffer to prepare Cl^--bound or Cl^--free *Np*HR samples, respectively. The Cl^- depletion of the Cl^--free *Np*HR ssNMR samples was checked with the opsin bound retinal absorbance wavelength (λ_{max}) change by visible absorption spectra. The proteoliposome suspension was directly measured with the UV-vis

spectrometer and the λ_{max} changes were analyzed.

2.2.7 Solution NMR method for evaluation $^{1}H_{N}$ back-exchange ratio

A 600 MHz (B_0 = 14.1 T) Bruker Avance III HD NMR spectrometer was used to analyze the $^{1}H_{N}$ back-exchange ratio of *Np*HRs in denatured condition. Three samples were prepared for this measurement: (1) [U-^{1}H]-*Np*HR (2) [^{2}H with $^{1}H_{N}$ back-exchanged in native state]-*Np*HR and (3) [^{2}H with $^{1}H_{N}$ back-exchanged through unfolding and refolding]-*Np*HR. The samples were loaded into standard 5 mm solution NMR sample tubes (~0.5 mL) for measurement. 2D ^{1}H-^{15}N TROSY[82]-HSQC spectra were measured for all the samples at 298 K and the number of peaks were calculated. The NMR data was processed with NMRPipe and visualized by NMRFAM-Sparky software. The detailed parameters are described in the appendix.

2.2.8 ^{1}H detection MAS ssNMR measurement method

All ^{1}H detection MAS ssNMR measurements were carried out on a 700 MHz (B_0 = 16.4 T) JEOL ECAII spectrometer with a triple resonance probe. Samples were packed into 1 mm ZrO_2 ssNMR sample rotors for the measurements, which requires approximately 1 mg of *Np*HR sample for each measurement with a homemade rotor packing tool via ultracentrifugation[83] (135,000 ×*g*, 20 min, 4 °C). The MAS frequency was set to 70 kHz and a cooling gas was used to maintain the sample temperature to ~20 °C under MAS, which was estimated from the temperature-dependent K^{79}Br longitudinal relaxation time[84]. CANH[85] experiment was performed for [U-^{1}H]-*Np*HR and [^{2}H with $^{1}H_{N}$]-*Np*HR back-exchanged in native state in the Cl^--bound state. For [^{2}H with $^{1}H_{N}$]-*Np*HR back-exchanged via unfolding and refolding sample, the CANH spectra were collected both in the Cl^--bound and -free state for structural comparation. The ^{13}C chemical shift was referenced using the carbonyl carbon chemical shift of L-alanine to 179.778 ppm in the sodium 2,2-dimethyl- 2-silapentane-5-sulfonate (DSS) scale. ^{1}H and ^{15}N chemical shift was indirectly referenced to DSS and liquid ammonia, respectively[86]. The NMR data was processed with NMRPipe[87] and visualized by NMRFAM-Sparky[88] software. The detailed parameters of the pulse sequences and data processing are described in the appendix.

2.3 Results

2.3.1 Cl^- depletion of *Np*HR in DMPC liposome

To make sure the DMPC liposome provided a near physiological condition for *Np*HR. I first checked the Cl^- depletion-related *Np*HR absorption change for *Np*HR reconstituted in DMPC liposome. It has been reported that, in the crystal, Cl^- removal of *Np*HR is hard to achieve (and *vice versa*); an alkaline condition with a yellow form M-like *Np*HR intermediate is needed to remove Cl- from the purple form Cl^--bound state *Np*HR and form the O-like Cl^--free blue form *Np*HR.[59] While such conformational changes can be easily achieved in solutions or in native cell membranes under neutral or acidic condition by a simple buffer exchange.[46]

Therefore, I measured the retinal absorption of *Np*HR in DMPC liposome by visible spectroscopy using the NaOAc buffer (50 mM, pH 5.0) with 100 mM NaCl or 67 mM Na_2SO_4 (Fig. 6). The data showed the Cl^- removal was achieved in DMPC liposome and the retinal absorption maxima (λ_{max}) for the two states were identical with the previous reported values[50]. This result suggested the DMPC liposome provided a good mimic for the native lipid membrane condition for *Np*HR and the ssNMR analysis based this proteoliposome can provide structural information in a near physiological condition.

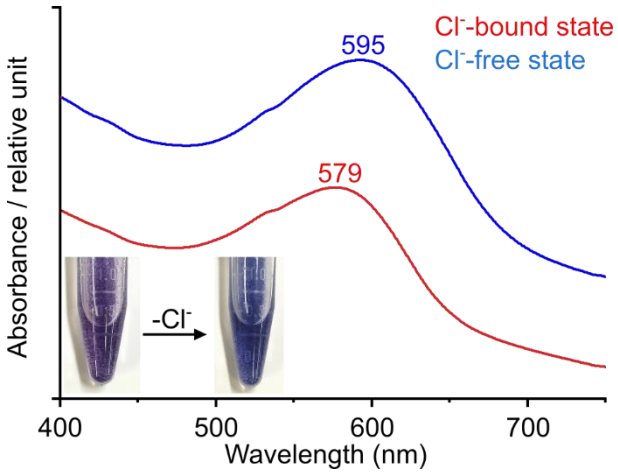

Fig. 6 The UV-vis spectra of reconstituted native state *Np*HR in DMPC liposome in Cl^--bound sate (red) and Cl^--free state (bule). The λ_{max} of the two states are shown in the spectra and the color change of *Np*HR proteoliposome after the Cl^- deletion is shown in the lower-left.

2.3.2 *Np*HR refolding by two-step dialysis

To improve ^1H_N back-exchange. I developed an unfolding and refolding method to prepare [^2H with ^1H_N]-*Np*HR (section 2.2.4). The unfolded *Np*HR preparation was described in section 2.2.3. Unfolded *Np*HR was successfully obtained with 0.5 % SDS which can be observed by the complete disappearance of the *Np*HR bound retinal absorption peak (Fig. 7).

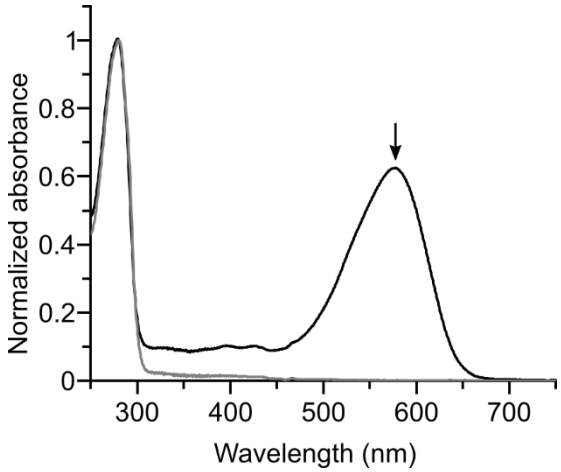

Fig. 7 The UV-vis spectra of purified native state *Np*HR in 0.1 % DDM (black) and

unfolded state *Np*HR in 0.5 % SDS (gray). The arrow indicates the *Np*HR bound retinal absorption peak.

Then, the *Np*HR refolding was performed by the two-step dialysis method described in section 2.2.4. The key step of this method was the 1st dialysis which allows SDS selective removal and *Np*HR refolding. The concept of the dialysis was: The micelle molecular weight of CHAPS (6.15 kDa) was larger than the cut-off range (3.5 kDa) therefore cannot be removed by this dialysis. However, the solubility of potassium dodecyl sulfate (KDS, 415 mg/L) was much lower than SDS (200 g/L) and the room temperature was lower than the Krafft temperature of KDS (34 ℃), dodecyl sulfate can be selectively precipitated as KDS and removed from the sample by this dialysis (Fig. 8). CHAPS, as a mild detergent, allows *Np*HR to keep its native fold, therefore, refolded *Np*HR can be generated after SDS removal in CHAPS-DMPC micelle (Fig. 9).

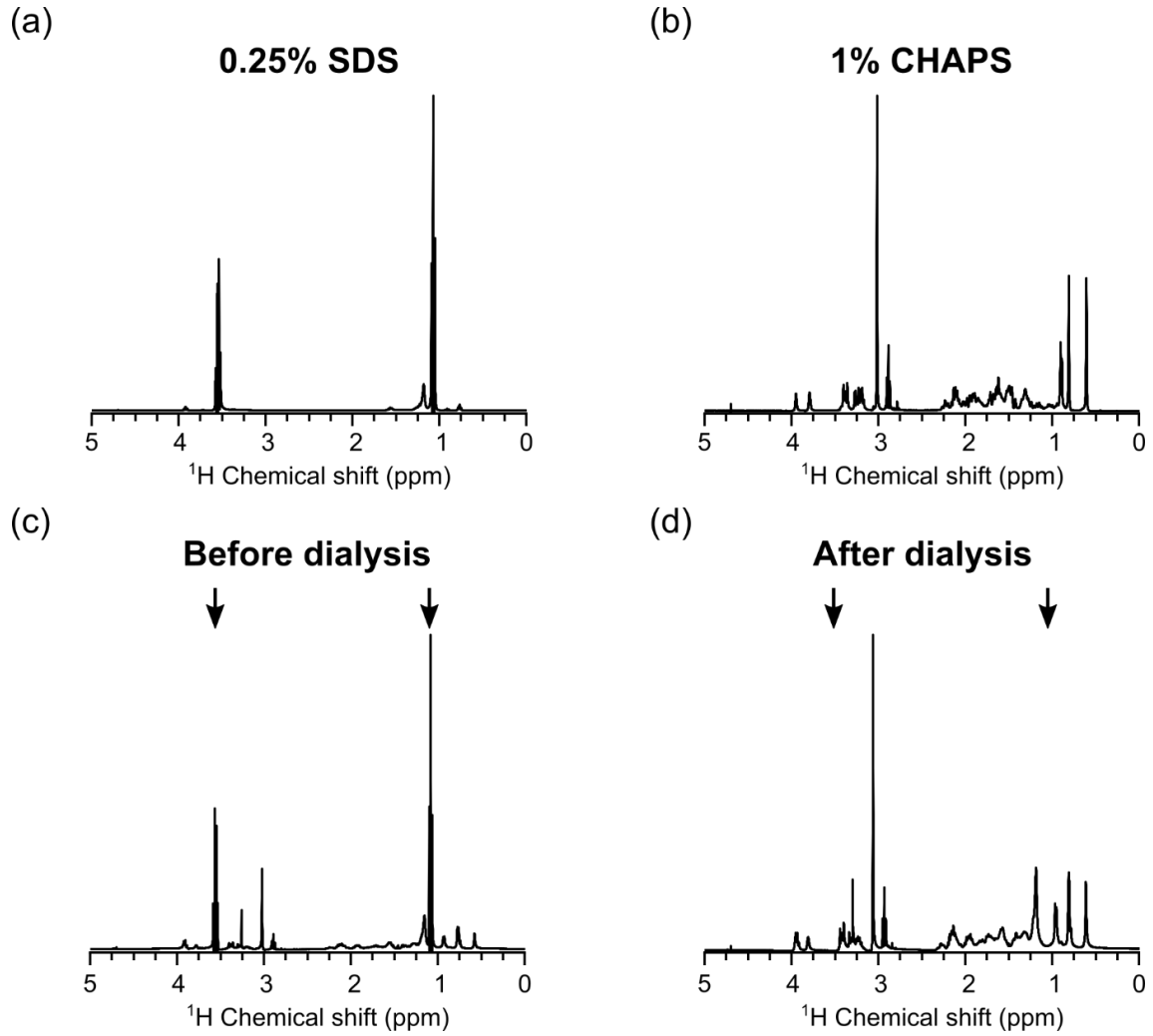

Fig. 8 ^{1}H detection solution NMR spectra for monitoring SDS removal efficiency with the dialysis method. (a) and (b) the sample buffer [50 mM Tris-HCl (pH 7.0) with 100 mM NaCl] only with 0.25% SDS or 1% CHAPS, respectively. (c) the sample buffer with 0.25% SDS and 1% CHAPS, which represented the detergent concentration before dialysis before dialysis. (d) (c) after 1 day dialysis against the sample buffer with 0.5 M KCl with 3.5 kDa cut-off

membrane. Arrows in (c) and (d) indicate the typical ^{1}H peaks of SDS around 3.6 ppm and 1 ppm in (a). In each spectrum, the peak height was normalized with the highest peak.

Fig. 8 a and b show the solution NMR spectra of the sample buffer [50 mM Tris-HCl (pH 7.0) with 100 mM NaCl] with 0.25 % SDS or 1 % CHAPS as reference. In Fig. 8c, which condition respect the sample condition before dialysis, peaks form SDS can be clearly observed, suggests the existing of SDS. While in Fig. 8d, those peaks disappeared and the peak shape pattern was similar with that in Fig. 8b, suggesting CHAPS was still remained in the sample. These results clearly mean the dialysis works as designed and SDS can be selectively removed

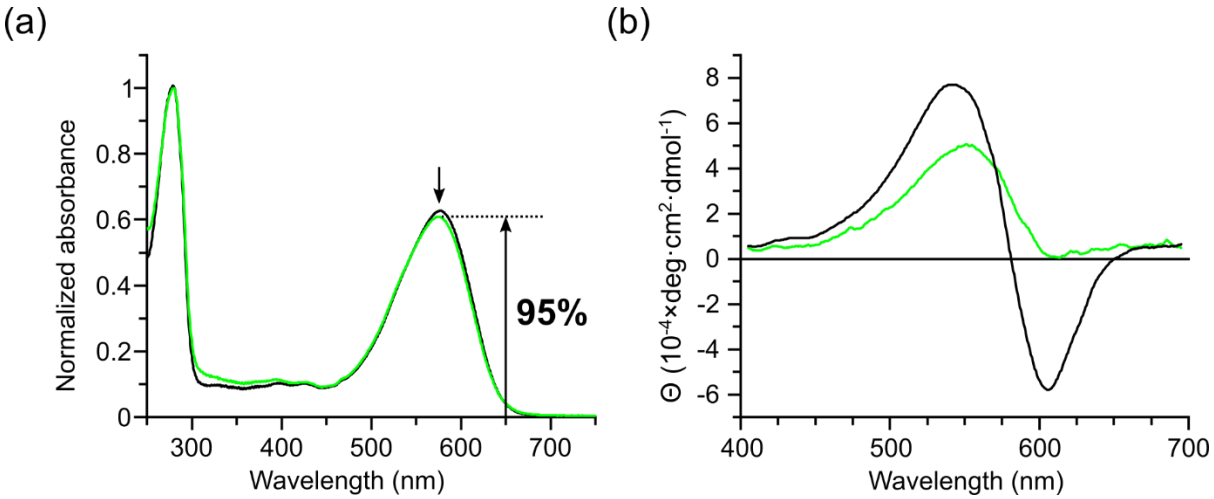

Fig. 9 The UV-vis spectra (a) and visible CD spectra (b) for refolded *Np*HR after 1st dialysis. *Np*HR solubilized in 0.1% DDM (the native state, black) data was shown as reference. (a) After dialysis, the retinal absorption peak was recovered (green) with an intensity of 95% comparing with the native state, suggesting *Np*HR can be refolded into native fold with high efficiency. (b) the refolded *Np*HR (green) showed the only positive peak at λ_{max} in visible CD spectra, suggesting the refolded *Np*HR was in the monomer state.

The *Np*HR folding condition was monitored by UV-vis spectroscopy using the peak intensity ratio between NpHR bound retinal absorption (A_{580}) and the protein absorption (A280) (Fig. 9a). The peak intensity of the refolded *Np*HR was ~95 % compared with the native state, suggesting the native fold of *Np*HR can be generated after SDS removal. However, in visible CD spectra, the refolded sample only showed a positive absorption at λ_{max} (Fig. 9b, green). For the native trimeric *Np*HR, the retinal-retinal interaction in *Np*HR trimer resulting in a typical bi-polar CD couplet[39], (Fig. 9b. black). These results suggested this refolded *Np*HR had a correct tertiary structure but non-native quaternary structure (monomeric native form) in the CHAPS-DMPC micelle.

Therefore, the 2nd dialysis step was applied. In this dialysis step, a large cut-off (10 kDa) membrane and much more dialysis buffer was used to diluent the concentration of CHAPS below its CMC value. Therefore, CHAPS was in monomer form and DMPC formed liposome solely with *Np*HR assembled in the liposome. Then, a heat-chill cycle was applied for *Np*HR reconstitution. Finally, the reconstituted *Np*HR was solubilized and purified by DDM again to remove the redundant DMPC lipid. The visible CD spectrum of the reconstituted *Np*HR (Fig.

10, red) showed a similar bi-polar pattern as the native state, suggesting the reconstituted *Np*HR was in trimeric form.

The above results clearly showed that *Np*HR can be unfolded by 0.5 % SDS and then refolded back into its native state through the refolding method described in this study with high efficiency. This protocol should be useful for preparing [^{2}H with $^{1}H_N$]-*Np*HR samples for ^{1}H detection ssNMR because the SDS unfolded *Np*HR offers better $^{1}H_N$ back exchange ratio with higher water accessibility to the protein hydrophobic core.

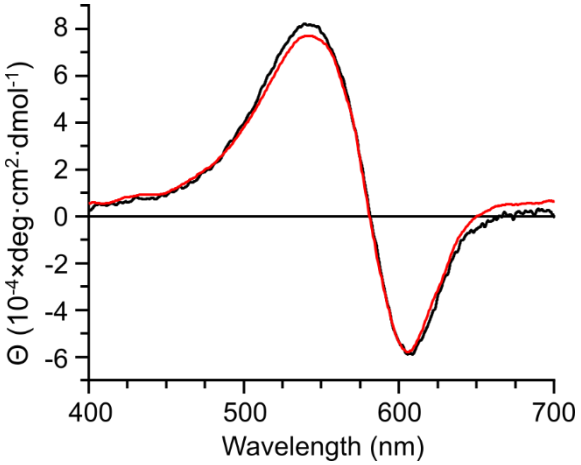

Fig. 10 The visible CD spectra of re-purified *Np*HR form DMPC liposome obtained by 2nd dialysis after heat-chill cycle (red). *Np*HR solubilized in 0.1% DDM (the native state, black) data was shown as reference.

2.3.3 $^{1}H_N$ back-exchange ratio analysis by solution NMR

To examine whether the unfolded and refolded sample provided better $^{1}H_N$ back-exchange ratio, I counted the number of amide peaks using solution NMR, which is suited to obtain very sharp signals through rapid Brownian motion although in a non-native solubilized condition. The proteins were solubilized in organic solvent (CD_3OD and $CDCl_3$ mixture) under a denatured state for solution NMR measurement. 2D TROSY-HSQC spectra were collected, which provide one ^{1}H-^{15}N signal form each amide site with $^{1}H_N$ labeled. The spectra of [U-^{1}H]-*Np*HR (blue); [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged in native state (red) and [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding (green) are shown in Fig. 11. The results clearly showed incomplete $^{1}H_N$ back-exchange happened for [^{2}H with $^{1}H_N$]-*Np*HR which performed $^{1}H_N$ back-exchange directly in native state Fig. 11b. Approximately 40 % peaks disappeared comparing with the spectra form [U-^{1}H]-*Np*HR. While, by applying the unfolding and refolding method, significant improvement in peak numbers can be observed (Fig. 11c). 50 % of missed peaks were recovered, showing the benefit of this novel *Np*HR sample preparation method.

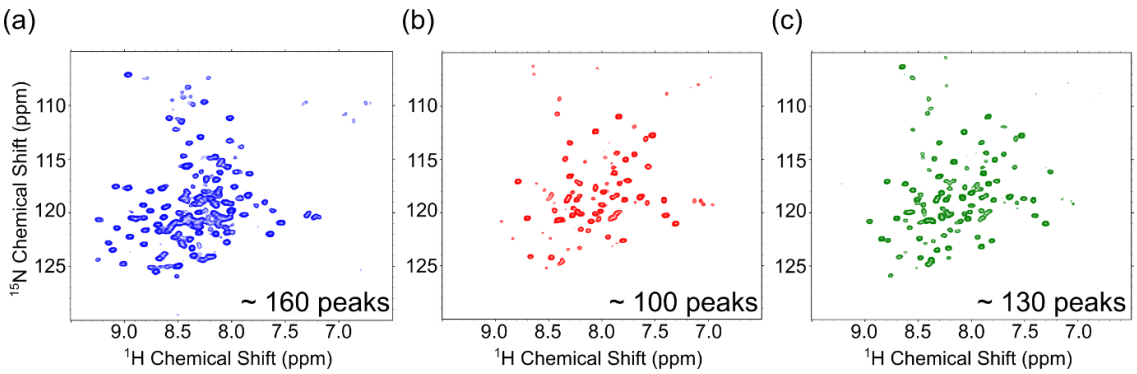

Fig. 11 2D ^{1}H-^{15}N TROSY-HSQC solution NMR spectra of non-native state *Np*HRs solubilized in 1:1 CD_3OD:$CDCl_3$ (v/v) for counting the total peak numbers. (a-c) [U-^{1}H]-*Np*HR (blue), [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged in native state (red) and [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding (green).

2.3.4 ^{1}H detection MAS ssNMR spectra of *Np*HR

The previous solution NMR data suggested the advantage of the novel *Np*HR unfolding and refolding protocol for maximizing the ^{1}H exchange rate in the [^{2}H with $^{1}H_N$]-*Np*HR sample preparation with denatured *Np*HR samples. To examine its benefit in ^{1}H detection ssNMR measurement, the three kinds of samples were reconstituted in DMPC liposome and 3D CANH ^{1}H detection ssNMR spectra of them were measured (Fig. 12). The spectra for perdeuterated *Np*HR showed significantly increased the spectra resolution: ^{1}H linewidth is much smaller which provided well sparse signals compared with the spectrum from [U-^{1}H]-*Np*HR. The ^{1}H linewidth in spectrum from [U-^{1}H]-*Np*HR was ~ 220 Hz and was ~ 170 Hz in spectra from [^{2}H with $^{1}H_N$]-*Np*HR. While the spectrum form [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged in native state (red) showed severe signal losses due to incomplete $^{1}H_N$ back-exchange, only small number of peaks can be observed and made this sample not ideal for subsequent analysis. In the CANH spectra, ~ 100 peaks was observed in the spectrum from [U-^{1}H]-*Np*HR, and the number of peaks was ~60 and ~100 for the spectrum form [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged in native state and [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding, respectively, indicated that by using unfolding and refolding, such signal losses are significantly reduced. Still, the number of peaks was less than that counted in the solution NMR data, which can be accounted for by the peak overlapping due to relatively broad linewidth in ssNMR spectra. Notably, in the glycine region which marked by square in (Fig. 12a), the [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding step even provided additional peaks which was missed in in the fully protonated *Np*HR. Such improvement is related to the longer transverse relaxion time (T_2) of perdeuterated sample than fully protonated sample. Glycine, which contains two hydrogen atoms bound with Cα, suffers stronger effect on T_2 form H content than other amino acids. Therefore, perdeuteration allowed the observation of such fast-relaxed glycine signals.

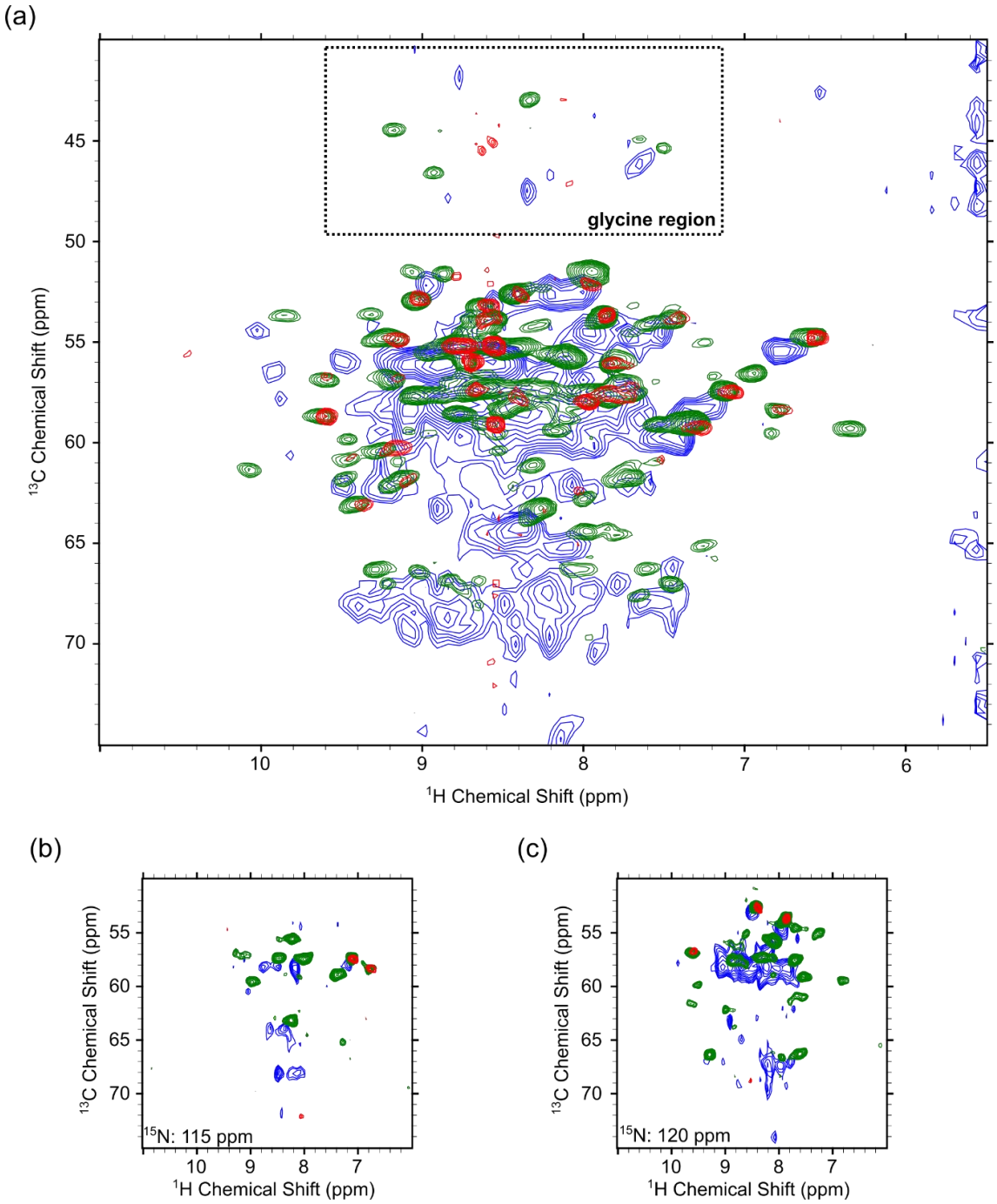

Fig. 12 The 3D ssNMR CANH spectra of [U-^{1}H]-*Np*HR (blue), [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged in native state (red) and [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding (green). (a) 2D projection of 3D CANH ssNMR spectra in H-C_α plane. (b and c) 2D slices of 3D CANH ssNMR spectra in H-C_α plane. The ^{15}N chemical shift of each plane is indicated in the top left corner.

The above results clearly showed better ^{1}H detection ssNMR data can be obtained from the [^{2}H with $^{1}H_N$]-*Np*HR back-exchanged via unfolding and refolding. Therefore, this sample was used for the structural comparation of *Np*HR between its Cl^--bound and -free state. Fig. 13 shows the CANH spectra of Cl^--bound state (red) and -free state (blue). CANH spectrum provides the correlations between Cα with the amide site N and H, and the peak position reflects

the local chemical environment of protein backbone. Peak shifts are widely dispersed among the spectra and indicated by rectangles (peaks still can be seen in the same ^{15}N plane) and arrows (peak shifted to other ^{15}N plane) in the spectra, which suggesting structural changes between Cl^--bound and -free state *Np*HR widely exist among the whole protein in liposome.

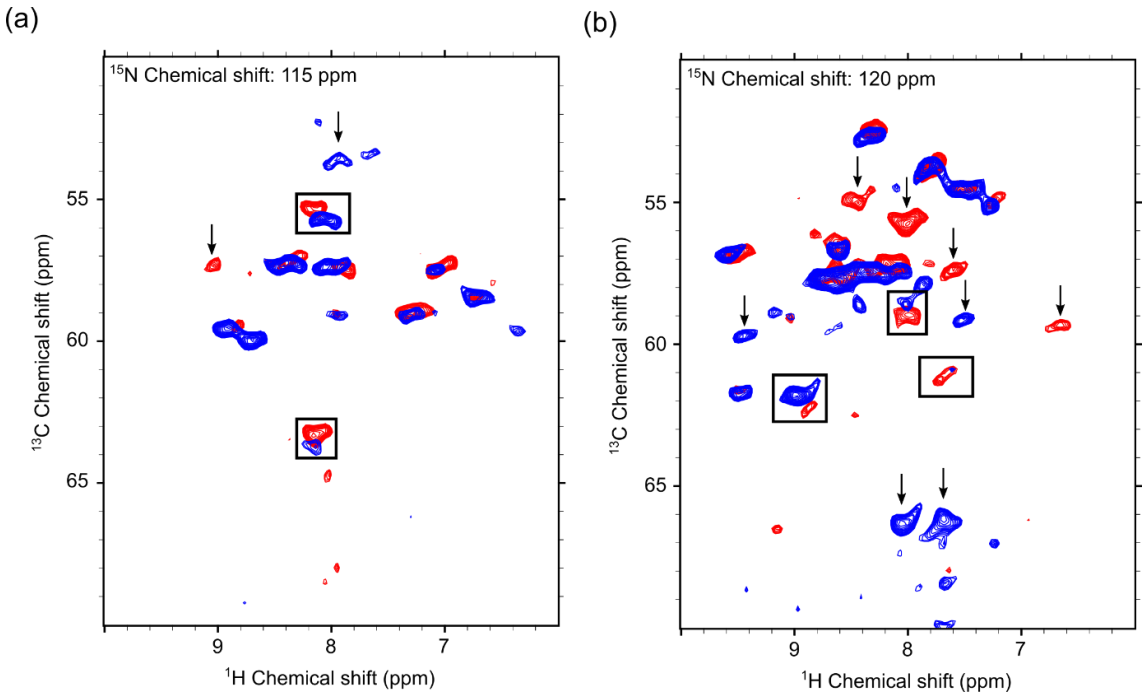

Fig. 13 2D slices of 3D CANH ssNMR spectra in H-C_α plane of *Np*HR in Cl^--bound sate (red) and Cl^--free state (bule). The spectra were obtained from perdeuterated with ^1H_N back-exchanged via unfolding and refolding *Np*HR samples. The ^{15}N chemical shift of each plane is indicated in the top left corner.

2.4 Discussion

The above results indicated the novel *Np*HR unfolding and refolding protocol significantly improved the ^{1}H back-exchange ratio for perdeuterated *Np*HR, which is essential for ^{1}H detection ssNMR measurements. This allowed me to conclude there are relatively major conformational changes of *Np*HR between its Cl^--bound -free state from observed peak shifting in the CANH spectra. However, even with *Np*HR unfolding step, ^1H_N back-exchange was still not perfect, ~20 % residues remained deuterated in amide site and resulting in signal losses in Fig. 11 and showed signal losses in ssNMR spectrum (Fig. 12). The total signal number of the 3D CANH spectra of [^{2}H with ^1H_N]-*Np*HR back-exchanged via unfolding and refolding was ~100 while the residue number used in this study was 269. This disparity prevented further signal assignment using ^{1}H detection ssNMR spectra. Consequently, although structural changes were evident through ^{1}H detection ssNMR, it was not possible to identify which specific regions of *Np*HR were undergoing these movements.

To address this challenge, ^{13}C detection ssNMR was considered, as it does not require perdeuteration or ^1H_N back-exchange, thus completely avoiding signal losses associated with incomplete ^1H_N back-exchange.

3 ^{13}C detection MAS ssNMR based *Np*HR structural analysis

3.1 Introduction

^{13}C detection MAS ssNMR is a more mature method than ^{1}H detection MAS ssNMR. Interactions of ^{13}C are much smaller than those for ^{1}H because of the low γ ratio of ^{13}C (~ 1/4 of ^{1}H). Therefore, a moderate MAS frequency (~10 kHz) is enough to average the anisotropic interactions. However, the low γ ratio means low sensitivity and much more sample (~ 10 mg protein) than ^{1}H detection method is required to obtain ^{13}C detection MAS ssNMR spectra within reasonable time.

Essentially, ^{13}C detection MAS ssNMR spectroscopy offers similar structural and dynamic information as ^{1}H detection, including distance restraints and insights into molecular dynamics through analysis of peak positions and intensity changes. For instance, the *de novo* structural determination of Anabaena Sensory Rhodopsin (ASR) was achieved using uniformly ^{13}C and ^{15}N labeled samples, as well as ^{13}C sparse-labeled samples using 1,3-$^{13}C_2$ glycerol and 2-^{13}C glycerol.[16,21,64] These studies on ASR serve as exemplary cases for employing ^{13}C detection MAS ssNMR in membrane protein structural investigations.

In this study, I applied a similar approach for signal assignment of *Np*HR using ^{13}C detection MAS ssNMR. Both sparse ^{13}C labeled and uniformly ^{13}C labeled *Np*HR samples were prepared for this purpose. I successfully assigned several residues of *Np*HR and conducted chemical shift perturbation analysis on the spectra from the Cl^--bound and Cl^--free states. The results indicated that the extracellular portion of helix D undergoes significant conformational changes upon Cl^- depletion, suggesting that this region is involved in the Cl^- uptake process.

3.2 Materials and methods

3.2.1 *Np*HR sample preparation method

The same *Np*HR protein was used for ^{13}C detection ssNMR measurement. For ^{13}C detection ssNMR, all the proteins were uniformly labeled by ^{15}N and ^{1}H using light water based M9 media. 1 g $^{15}NH_4Cl$ in 1 L media was still used for ^{15}N labeling. While for ^{13}C labeling, uniformly ^{13}C labeling (referred to as [U-^{13}C, ^{15}N]-*Np*HR) and sparse ^{13}C labeling with 1,3-^{13}C glycerol or 2-^{13}C glycerol (referred to as [1,3-^{13}C, U-^{15}N]-*Np*HR and [2-^{13}C, U-^{15}N]-*Np*HR, respectively) were used. 3 g [1,3-$^{13}C_2$]-glycerol or [2-^{13}C]-glycerol in 1 L media was used as sole carbon source for preparing [1,3-^{13}C, U-^{15}N]-*Np*HR and [2-^{13}C, U-^{15}N]-*Np*HR, respectively and 2 g [U-$^{13}C_6$]-glucose in 1 L media was used as sole carbon source for preparing [U-^{13}C, ^{15}N]-*Np*HR.

The culture condition was the same as previously described in section 2.2.1 and the expressed protein was purified in the same manner as described in section 2.2.2 in the native state. The *Np*HR-DMPC proteoliposome preparation method for ^{13}C detection ssNMR

measurement was essentially the same as it for ^{1}H detection ssNMR (section 2.2.6). The only modification for ^{13}C detection ssNMR measurement was the protein to lipid molar ratio (P/L ratio) was increased from 1:20 to 1:5. The prepared proteoliposomes were packed into 3.2 mm Varian-style rotors with a home-made rotor packing tool via ultracentrifugation[83] (100,000 ×*g*, 20 min, 4 °C).

3.2.2 ^{13}C detection ssNMR measurement method

All ^{13}C detection ssNMR experiments were carried out on a 600 MHz (B_0 = 14.1 T) JEOL ECAII spectrometer with a triple resonance probe. The MAS frequency was set to 12.5 kHz or 14 kHz and a cooling gas was used to maintain the sample temperature to ~20 °C under MAS using the temperature-dependent K^{79}Br longitudinal relaxation time [84]. ~ 10 mg *Np*HR samples were packed into 3.2 mm ZrO_2 ssNMR sample rotors for the measurements. NCA[89], NCACX[90], NCOCX[90], CANCO[91], CANcoCA[92] and NcoCACX[93] spectra were recorded for [U-^{13}C, ^{15}N]-*Np*HR in the Cl^--bound state, while NCACX and CANCO spectra were also performed for [U-^{13}C, ^{15}N]-*Np*HR in the Cl^--free state. For [1,3-^{13}C, ^{15}N] and [2-^{13}C, ^{15}N]-*Np*HR, NCACX, NCOCX and CANCO experiments were recorded in the Cl^--bound state. The ^{13}C chemical shift was directly referenced using the adamantane CH_2 peak to 40.48 ppm in DSS scale.[94] ^{1}H and ^{15}N chemical shift was indirectly referenced to DSS and liquid ammonia, respectively.[86]

The NMR data was processed with NMRPipe[87]. NMRFAM-Sparky[88] and CcpNmr[95] were used for spectra visualization and analysis. The FLYA algorithm in CYANA software was used for the automated resonance assignment to aid manual assignment procedure.[96,97] The CYANA library file was modified for the automated assignment using data for the sparsely ^{13}C-labeled samples. The detailed parameters of the pulse sequences and data processing are described in the appendix.

3.3 Results

3.3.1 ^{13}C detection ssNMR spectra of *Np*HR with different P/L ratio

Because ^{13}C detection ssNMR has relatively low sensitivity due to the small γ of ^{13}C, it is necessary to pack as many as possible *Np*HR sample into ssNMR rotor for higher signal to noise ratio (S/N ratio). However, the total volume of the 3.2 mm rotor is limited, the only possible way to increase *Np*HR amount in the sample was to use a higher protein to lipid (P/L) ratio. In this study, I tried to prepare the *Np*HR-DMPC liposome with the P/L ratio of 1:15, 1:10 and 1:5, which corresponding with the weight ratio of *Np*HR in the total proteoliposome increased from ~67 % (1:20 P/L ratio used for ^{1}H detection ssNMR) to ~90 % (1:5 P/L ratio). 2D HCA spectra of these samples were measured (Fig. 14 a-c) to evaluate the influence of P/L ratio to the sample condition.

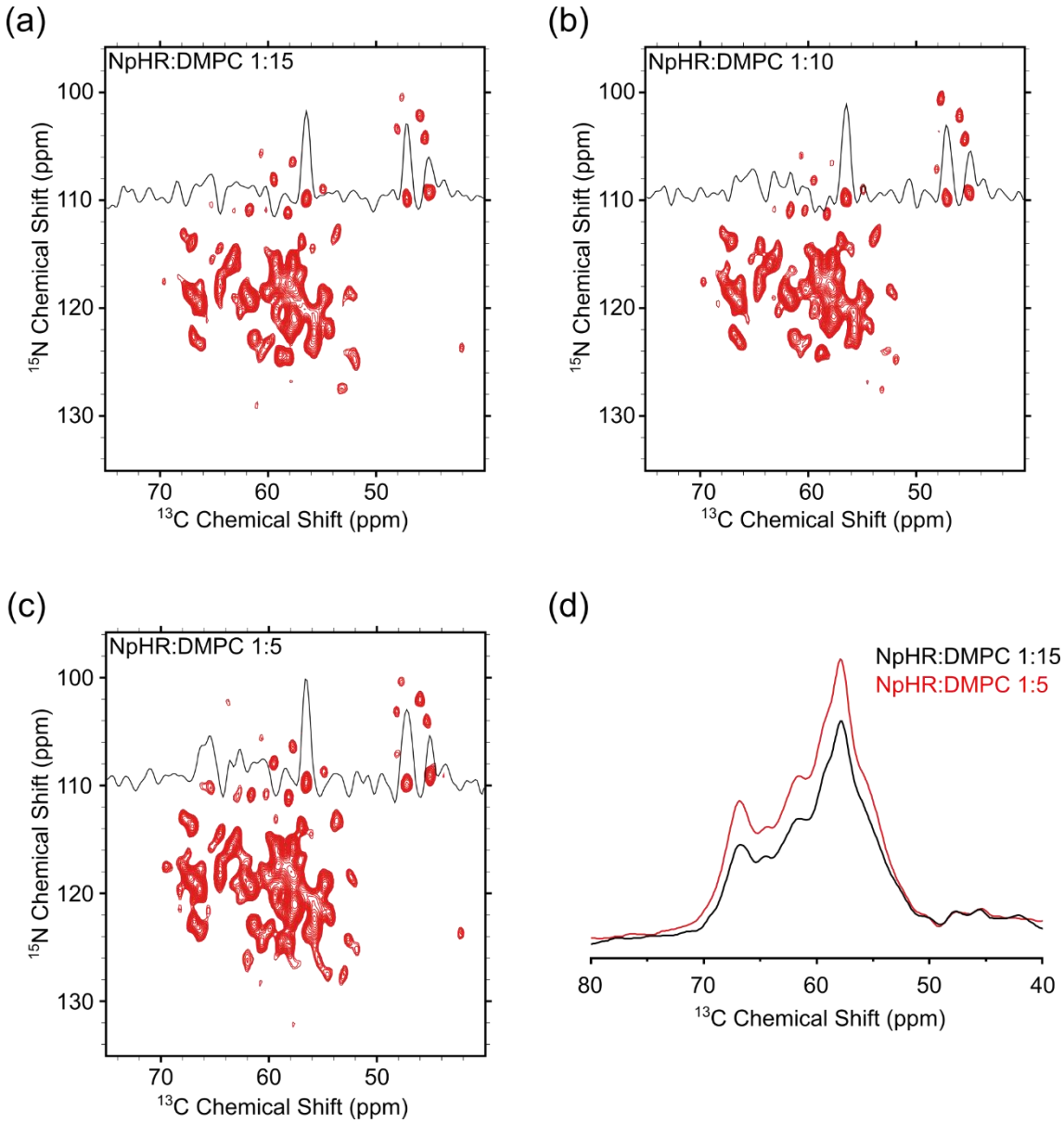

Fig. 14 2D double-CP based NCA spectra of [U-^{13}C,^{15}N]-*Np*HR for protein-to-lipid (P/L) ratio optimization. *Np*HR to DMPC mole ratio is 1:15, 1:10 and 1:5 for (a)-(c), respectively. The same contour level is used for all panels. Slices of ^{15}N chemical shift at 110 ppm are shown. (d) The 1D projection generated from the NCA spectra of [U-^{13}C,^{15}N]-*Np*HR. The bulk signal intensity was ~1.3 times higher for the sample with the P/L ratio 1:5 relative to that with 1:15.

The NCA spectra from different P/L ratio showed almost same patten with similar linewidths. These results suggested that in these samples *Np*HR were in similar condition and homogeneous *Np*HR sample can be obtained in a very high P/L ratio (1:5). While, in the spectrum from the highest P/L ratio (Fig. 14c, P/L ratio = 1:5), more peaks were observed and the bulk signal intensity was ~1.2 times of that in the spectrum of P/L ratio = 1:15 (Fig. 14d). The bulk signal intensity improvement was consistent with the corresponding *Np*HR weight ratio in the proteoliposome of 90 % (P/L ratio = 1:5) and 75 % (P/L ratio = 1:15), assuming the same density for protein and lipid bilayer portions. Based on the above results, the P/L ratio of 1:5 was used for further ^{13}C detection ssNMR measurement.

3.3.2 Signal assignment based on backbone walk strategy

The backbone walk strategy was used for signal assignment of *Np*HR. The concept of backbone walk is demonstrated in Fig. 15. This strategy utilizes a set of 3D spectra containing intra-residue correlations (NCACX) and inter-residue correlations (NCOCX and CANCO). By matching the peak pattern of those spectra, signal assignment can be performed along the protein backbone, therefore called backbone walk. However, spectra signal overlapping results in ambiguous peak matching and additional spectra with different correlations are usually required for unambiguous signal assignment. In this study, chemical shift assignment was performed by combining the CANCO, NCACX, NCOCX, CANcoCA and NcoCACX data, while ^{13}C sparse labeling with 1,3-$^{13}C_2$ glycerol or 2-^{13}C glycerol were also utilized for signal assignment.

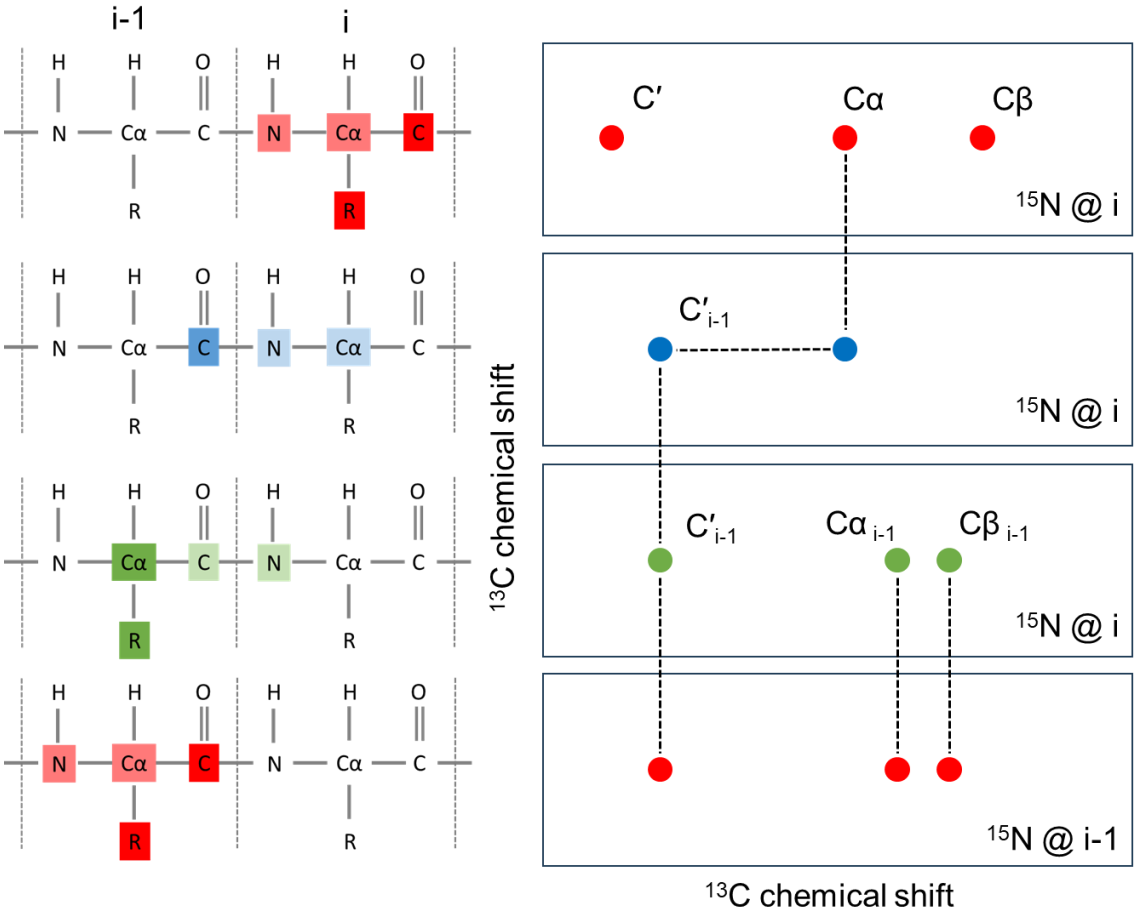

Fig. 15 A schematic figure for backbone walking using 3D NCACX (red), CANCO (blue) and NCOCX (green) spectra.

This sparse ^{13}C labeling reduced the total peak number while the unique labeling pattern helped to recognize residue types[98], as shown in Fig. 16. In addition, the sparse ^{13}C labeling reduced the signal linewidths by removing the homonuclear one-bond ^{13}C scalar couplings and mitigated the signal overlap.[22,98] Such advantages made the sparse labeled *Np*HR samples

benefit a lot in signal assignment.

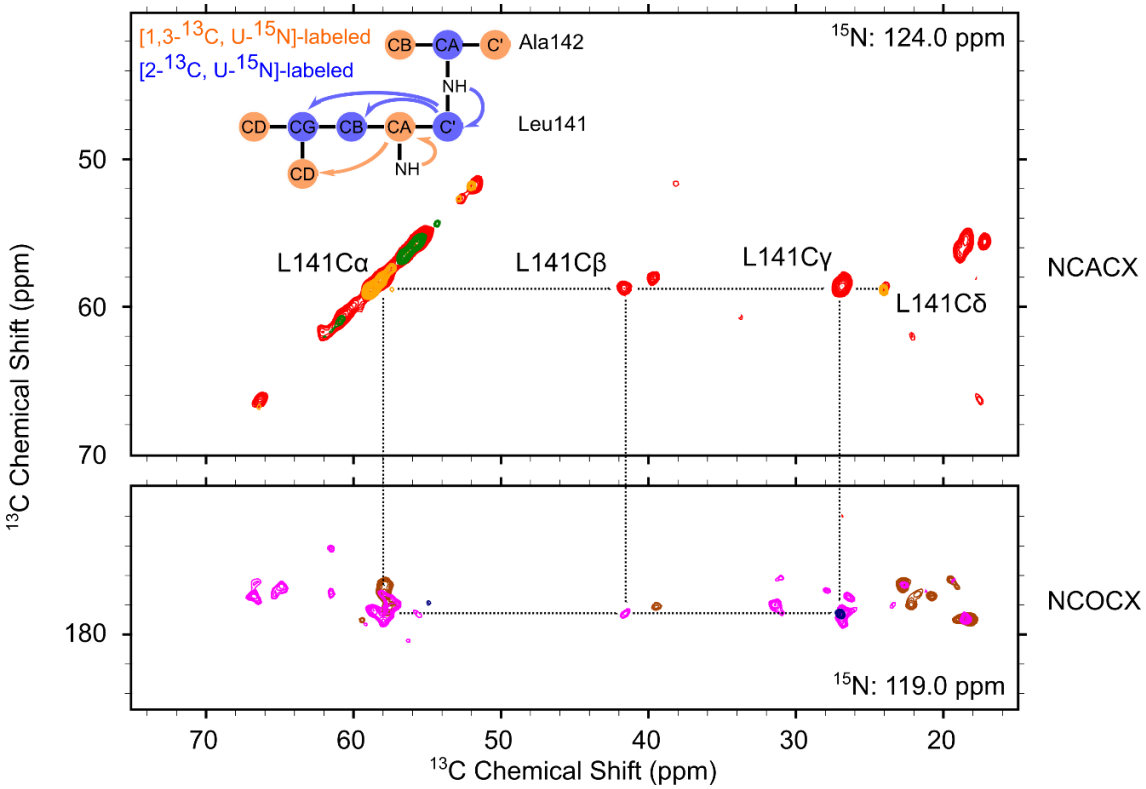

Fig. 16 Leu141 assignment assisted by spectra from sparsely ^{13}C-labeled *Np*HR. NCACX spectra from [1,3-^{13}C, U-^{15}N]-*Np*HR (orange) and [2-^{13}C, U-^{15}N]-*Np*HR (green); NCOCX spectra from [1,3-^{13}C, U-^{15}N]-*Np*HR (brown) and [2-^{13}C, U-^{15}N]-*Np*HR (blue) as well as these two spectra from [U -^{13}C, -^{15}N]-*Np*HR (red and magenta, respectively) are showed. ^{15}N chemical shift of NCOCX spectra is at 119.0 ppm where is the N_H chemical shift of Ala142. The labeling scheme of Leu141-Ala142 is showed in the top left corner, this labeling scheme is adopted from Castellani et al. The sparse ^{13}C labeling allowed unambiguous assignments in the crowd NCOCX plane of Ala142 that enabled the extending of "backbone walk".

Examples of the sequential assignment stretches with spectra from [U-^{13}C, ^{15}N]-*Np*HR are shown in Fig. 17. Total 27 backbone nitrogen, 32 Cα and 30 C′ signals (25 N- Cα-C′ intra-residue pairs) were assigned to a major part of helix D, C-D loop and E-F loop on the cytoplasmic side and the nearby helical residues, together with a few residues in the other helices, the position of these assigned residues is demonstrated in *Np*HR X-ray structure in Fig. 18 and assignment results are tabulated in Table S7. Other nuclei were not assigned without ambiguity because of the amino acid type degeneration and signal losses in the flexible regions.

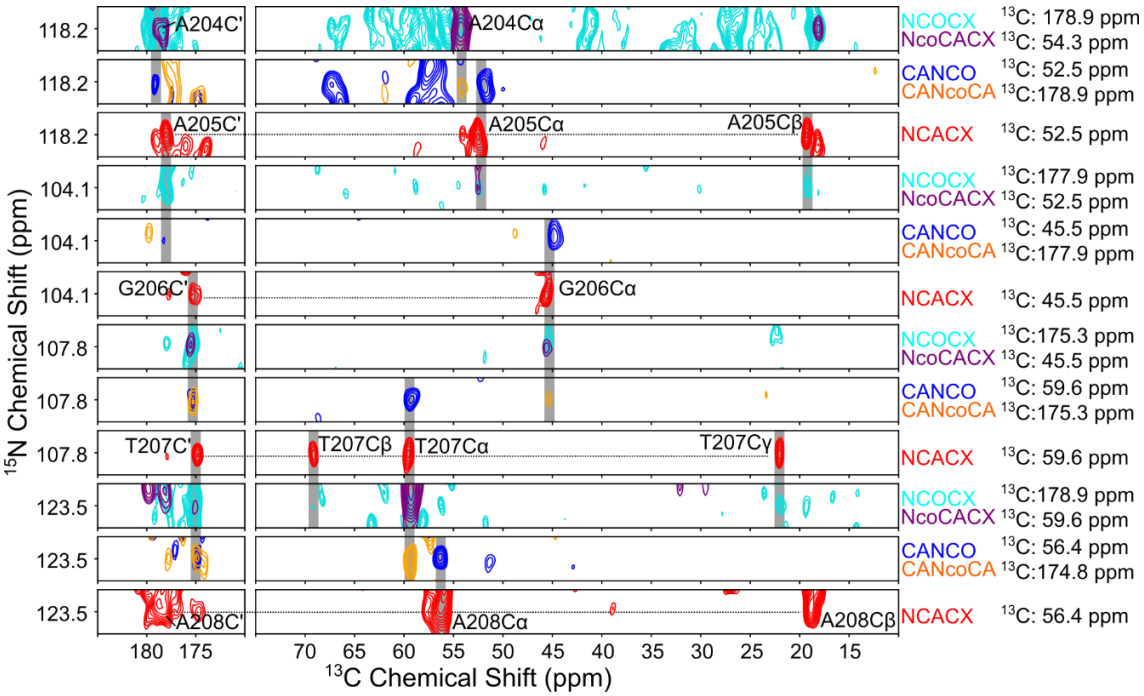

Fig. 17 An example of the "backbone walk" from A205 to A208 with the set of 3D spectra, shown for the Cl^--bound state [U-^{13}C, ^{15}N]-*Np*HR. The experiment names and chemical shifts for the 3rd axis are shown on the right of each plane.

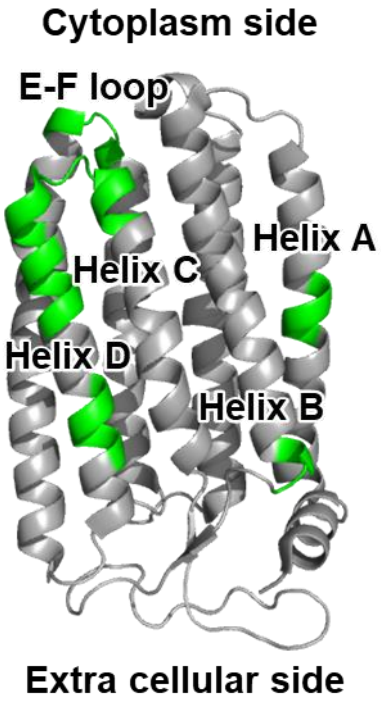

Fig. 18 The location of the assigned residues. Residues colored in green are the residues with assignment in backbone, the other residues colored in gray are unassigned ones. This structure is the crystal structure of Cl^--bound state *Np*HR (PDB entry: 3A7K[38]), the structure visualization is performed by PyMOL software.

3.3.3 Chemical shift perturbation based structural analysis of Cl^--bound and -free state *Np*HR

Next, NCACX, NCOCX and CANCO spectra were recorded for [U-^{13}C, ^{15}N]-*Np*HR in the O-like Cl^--free state to evaluate the structural changes because of the Cl^- depletion. A similar assignment strategy was applied for the signal assignment of the Cl^--free state *Np*HR with referenced by the previous results from the Cl^--bound state. The assignment results of the Cl^--free state *Np*HR are tabulated in Table S8. Compared with the spectra from Cl^--bound state, Cl^--free state *Np*HR provided a smaller number of signals. In Cl^--free state CANCO spectrum, ~120 peaks were observed. The decreased number of signals was related to the increased molecular mobility of *Np*HR in Cl^--free state. ~ 100 peaks in CANCO spectra were found both in the Cl^--bound and -free state, and numerous peak shifting were observed. This result meant the conformational changes still observed in ^{13}C detection ssNMR spectra and I have obtained assignments for several peaks therefore further structural analysis based on chemical shift perturbation (CSP) was performed. The CSP value for the peak pairs in CANCO spectra were calculated with:

$$\mathrm{CSP} = \sqrt{(\delta C_{\alpha,\ with\ Cl^-} - \delta C_{\alpha,\ without\ Cl^-})^2 + 0.4(\delta N_{H,\ with\ Cl^-} - \delta N_{H,\ without\ Cl^-})^2}.$$

where δC_α and δN_H is the chemical shift value of backbone ^{13}Ca and ^{15}N, respectively. CSP basically reflects how much the peak moved which corresponds to the local chemical/structural change for the atom. Therefore, CSP is a very nice tool to sense and quantity structural changes. Fig. 20 shows the Histogram of the number of peak pairs against their CSP value. The average CSP for all the peak pairs detected with a finite peak shift (assigned or not assigned) was 0.22 ppm; among them, only 30% of peak pairs gave a CSP value larger than 0.275 ppm (1.25 times of the mean value). This result suggested the conformational changes were widely distributed among the whole protein and some regions showed relatively major conformational change corresponding to the large CSP value. This result was consistent with previous ^{1}H detection ssNMR data.

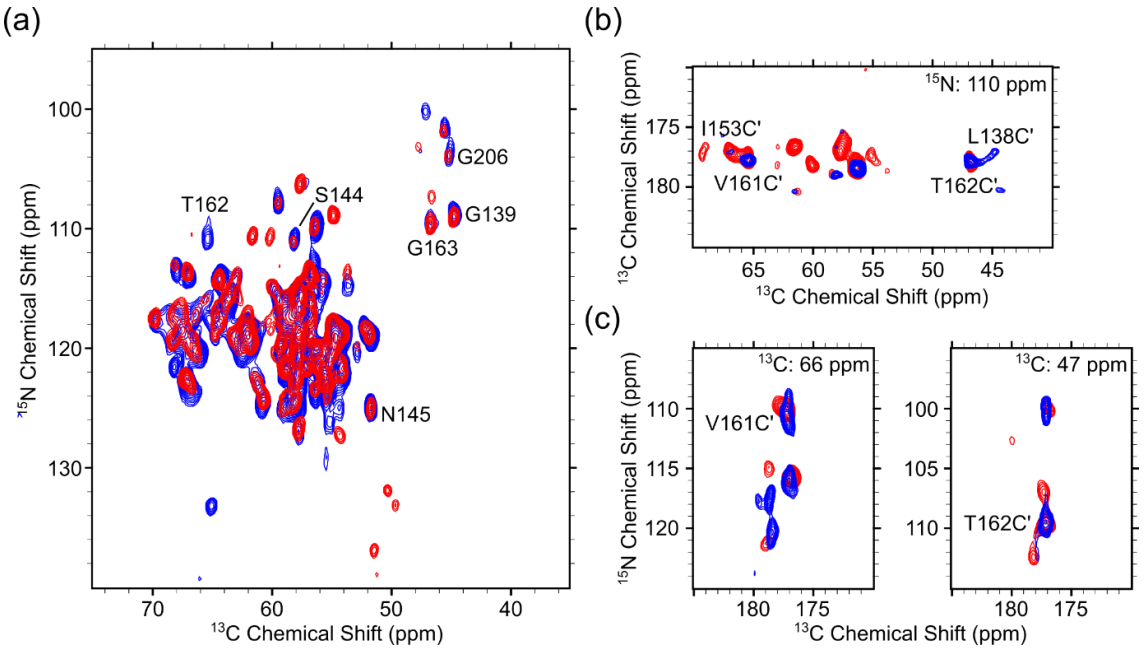

Fig. 19 The spectra comparation of [U-^{13}C,^{15}N]-*Np*HR in the Cl^--bound state (red) and -free state (blue) (a) The 2D projection of 3D CANCO ssNMR spectra recorded with [U-^{13}C,^{15}N]-*Np*HR in the Cl^--bound state (red) and -free state (blue). (b) A 2D C-C plane taken at δN_H =110 ppm. (c) 2D C′-N planes taken at δC_α indicated in the upper conner. Assigned peaks are labeled.

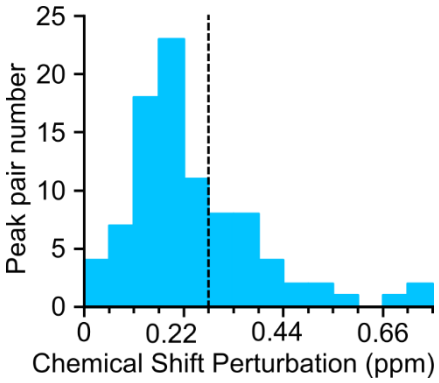

Fig. 20 Histogram of the number of peak pairs in CANCO spectra against their CSP value. Total 91 peak pairs were distinguished regardless of whether they were assigned or not and are shown in the histogram. The average CSP value of the peak pairs was 0.22 ppm, and 30% of them showed CSP value > 0.275 ppm ($\Delta\delta_{30th}$), which is indicated by the dash line.

Then, I focused on the CSPs of assigned residues, which are illustrated in Fig. 21. Those assigned residues in Helix A and B were not included for this CSP analysis because the assigned residue chain was too short to discuss conformational changes.

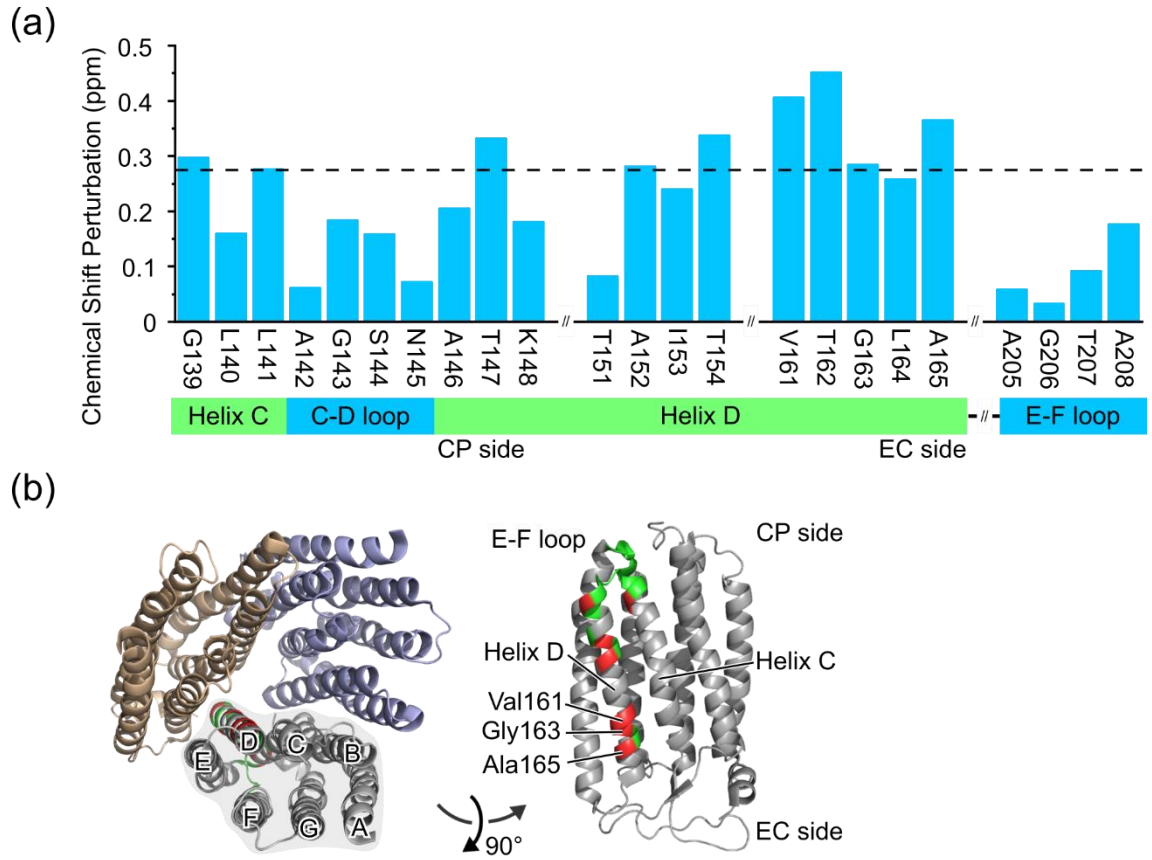

Fig. 21 The CSPs for the assigned residues in helix C, C-D loop, helix D and E-F loop region. (a) CSPs plotting for those assigned residues, the dash line is at 0.275 ppm ($\Delta\delta_{30th}$). (b) the location for those residues. The trimer structure in left is a homologous model from *Hs*HR (PDB entry: 1E12[40]) because the protein expressed in *E. coli* lacks bacterioruberin, which is necessary for the trimeric assembling in the 3A7K trimeric form; and the monomer structure is the Cl^--bound state *Np*HR (PDB entry: 3A7K[38]). The assigned residues in helix A and B were not included for this CSP analysis. The structure visualization is performed using PyMOL.

A few resides in helix C and D (Gly139-Thr154) on the cytoplasmic side showed significant CSPs, but the average CSP among this region (~0.19 ppm) was lower than $\Delta\delta_{30th}$. These residues are close to the Cl^- releasing channel in the cytoplasmic side[56]. Photochemical studies of the intermediates previously suggested structural changes arising for O formation remained until *Np*HR′ formation[44], and the N and O intermediates are under a structural quasi-equilibrium[45,57,58]. Therefore, the observed CSPs may be related to Cl^- releasing process. On the other hand, a stretch of five residues in the extracellular (EC) part of helix D: from Val161 to Ala165 showed large CSP values close to or greater than the threshold of 0.275 ppm, which suggested relatively major conformation changes in the EC part of helix D. Importantly, in the crystal structures, conformational differences around the EC side of helix D are not observed, while only a part of helix C and B-C loop on the EC side showed finite changes[59]. Considering the location of these residues on the EC side, and the fact that EC part of helix D is adjacent to several residues important for Cl^- uptake such as His100, Arg176 and Glu234[38], we presume the large CSPs observed were related to the Cl^- uptake process.

3.4 Discussion

Through ^{13}C detection ssNMR measurements, I successfully assigned several residues for *Np*HR and the CSP analysis indicated at least one of the major conformational changes occurred at the EC part of helix D. The assigned number of assigned residues is still limited because of the difficulty due to the nature of membrane proteins whose amino acid types are strongly biased toward the hydrophobic residues. When counting the total peak numbers in CANCO spectra, which provides one signal from each residue in most cases, ~140 peaks were observed, and this number was larger than it from ^{1}H detection CANH spectrum. The discrepancy with respect to the total residue number 269 can be related to signal overlapping due to relative wide signal linewidth and small distribution range of carbonyl carbon chemical shift, while the unique structural characteristic of *Np*HR also contributed. *Np*HR contains the large B-C loop region (~ 30 residues) in the extracellular side, which is highly flexible and was not observed in ssNMR measurements.

Some advanced ssNMR techniques may be beneficial for more assignments. For example, reverse isotope labeling[99,100], which selectively made certain amino acids unlabeled, simplifies the spectra with less overlapping should be helpful for the assignment. This method was already utilized in the assignment of two different microbial rhodopsins: proteorhodopsin[99] and sodium ion pump rhodopsin, KR2[100].

4 Functional analysis with point mutants for important residues

The previous ^{13}C detection ssNMR data concluded that the EC part of helix D undergoes some major conformational changes between *Np*HR Cl^--bound and -free state. To further investigate what kind of conformational change occurred and the relationship of these conformational changes with the Cl^- transport function, I tried to prepare point mutants for the residues at EC part of helix D. I successfully obtained the following three mutants: the Val161 to Ala (V161A), Gly163 to Ala (G163A) and Ala165 to Val (A165V). The molecular properties of these mutants were measured and compared with WT to reveal the details of this conformational change with the Cl^- transport function of *Np*HR.

4.1 Materials and methods

4.1.1 *Np*HR sample preparation method

The truncated *Np*HR at position from 18 to 277 in residue number was used as WT *Np*HR. All the mutants were prepared based on it through the QuikChange method (Agilent). A $^{15}N\zeta$-Lysine-labeled *Np*HR (referred to as [$^{15}N\zeta$-Lys]-*Np*HR) in wild type and mutants for DNP enhanced ^{15}N detection NMR measurements were expressed using M9 minimal media with $^{15}N\zeta$-Lysine, and those for molecular property analyses were expressed using LB media without isotope labeling.

To prepare [$^{15}N_{\zeta}$-Lys]-*Np*HR, 100 mg/L of $^{15}N_{\zeta}$-Lysine was added into the *E. coli* culture using natural abundant NH_4Cl (1 g/L) and glucose (3 g/L) as the nitrogen and carbon source when the optical density of culture at 660 nm (OD_{660}) reached 0.8. Then, the protein expression was still induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when OD_{660} reached 1.4 with 10 μM *all-trans* retinal and harvested after 4 h, which was described in section 2.2.1. The expression method for *Np*HR mutants was the same as pervious described in section 2.2.1

All *Np*HRs were purified in the native state with the same method as 2.2.2 and applied for the following measurements.

4.1.2 Cl^- affinity analysis method

The purified *Np*HR was passed through the PD-10 column and equilibrated with 50 mM PIPES (pH 7.0) and 0.1 % DDM in Cl^--free state. Then the UV-vis spectra were measured under various concentrations of NaCl from 0 to 1 M by adding 5 M NaCl. The Cl^- dependent absorption changes were analyzed by Hill equation:

$$\Delta A_{\mathrm{rel.}} = \frac{[Cl^-]^n}{K_{\mathrm{d}}^n + [Cl^-]^n}.$$

where $\Delta A_{\mathrm{rel.}}$ was the normalized reduced absorbance at the wavelength where the difference spectra between Cl^--bound and -free state reached maximum value, and n is the Hill-coefficient.

4.1.3 Transient absorption measurement method

For transient absorption measurement, the measuring medium exchanged to 50 mM PIPES (pH 7.0), 1 M NaCl and 0.1 % DDM by PD-10 column, while the *Np*HR concentration was adjusted to $Abs_{max} \approx 0.5$ by concentration using ultrafiltration devices (Amicon® Ultra centrifugal filter 10kDa cutoff, Sigma). The transient absorption measurement apparatus was kindly provided by Demura group (Laboratory of Biological Information Analysis Science, Hokkaido University) and described previously[101]. The transient absorption measurement apparatus is essentially a time-resolved UV visible spectrometer equipped with an additional Nd-YAG laser light source. This Nd-TAG laser (532 nm, 7 ns) was used for exciting *Np*HRs and the and the absorbance changes (transient absorption) after this excitation was recorded. For each measured wavelength (400-700 nm, at 10 nm intervals) 60 laser pulses were used to improve the S/N ratio. The data then analyzed with an irreversible sequential model[52,58] by global fitting:

$$P_0 \xrightarrow{h\nu} P_1 \xrightarrow{\tau_1} P_2 \xrightarrow{\tau_2} P_3 \xrightarrow{\tau_3} \cdots P_i \xrightarrow{\tau_i} P_0.$$

In this model, the P_i (i =1,2 …) were the photo-excited kinetic distinguishable intermediates and P_0 was the ground state. τ_i (i =1,2 …) are the corresponding time delay constant for each intermediate transition. Therefore, the contents of them after light excitation follows the following equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}P_1\\P_2\\\vdots\\P_i\\\vdots\end{bmatrix}=\begin{bmatrix}-1/\tau_1 & 0 & \cdots & & 0\\ 1/\tau_1 & -1/\tau_2 & \ddots & & \\ 0 & \ddots & \ddots & & \\ \vdots & \ddots & 1/\tau_{i-1} & -1/\tau_i & \\ 0 & & & \ddots & \ddots\end{bmatrix}\begin{bmatrix}P_1\\P_2\\\vdots\\P_i\\\vdots\end{bmatrix}_t \qquad \text{eq. 2}$$

At t_0, just after photo excitation, the photo excited *Np*HR was all in P_1 state, therefore P_1 =1 and other P_i = 0. Then eq. 2 was solved as:

$$\begin{bmatrix}P_1\\P_2\\\vdots\\P_i\\\vdots\end{bmatrix}_t=\begin{bmatrix}a_{1,1} & 0 & \cdots & & 0\\ a_{2,1} & a_{2,2} & \ddots & & \\ \vdots & & \ddots & & \\ a_{i,1} & \dots & a_{i,i-1} & a_{i,i} & \\ \vdots & & & & \ddots\end{bmatrix}\begin{bmatrix}\exp(-t/\tau_1)\\\exp(-t/\tau_2)\\\vdots\\\exp(-t/\tau_i)\\\vdots\end{bmatrix} \qquad \text{eq. 3}$$

Where $a_{i,j}=\left\{\prod_{m=1}^{i-1}\tau_m\prod_{m=1,m\neq j}^{i}\left(\frac{1}{\tau_m}-\frac{1}{\tau_j}\right)\right\}^{-1}$, defined the content for the intermediates as the function of time.

While the measured transient absorption change, $\Delta A(\lambda,t)$ can be described as:

$$\Delta A(\lambda,t)=F_c\sum_{i=1}^{n}B_i(t)\exp(-t/\kappa_i) \qquad \text{eq. 4}$$

where F_c was the fraction of *Np*HR excited by the laser, B_i (i=1,2 …) was the absorption difference of the i^{th} photochemically defined intermediate with the ground state and κ_i (i =1,2 …) was the corresponding delay time constant. The intermediates in eq. 4 were permitted to be a mixture of the intermediate defined in eq. 3. Therefore, $B_i(t)$ (i=1,2 …) can be described by $\Delta\epsilon_i(\lambda)$ (i = 1,2 ...), which was the absorption difference for the i^{th} intermediate with ground state at the wavelength of λ and the content of the i^{th} intermediate, which described in eq. 3 as:

$$B_i(\lambda,t)=F_c\sum_{j=i}^{n}a_{j,i}\Delta\epsilon_j(\lambda) \qquad \text{eq. 5}$$

Which corresponding to a linear system:

$$F_c\begin{bmatrix}a_{1,1} & a_{2,1} & a_{3,1} & \cdots & a_{n,1}\\ 0 & a_{2,2} & a_{3,2} & \cdots & a_{n,2}\\ \vdots & \ddots & a_{3,3} & \cdots & a_{n.3}\\ & & & \ddots & \vdots\\ 0 & \cdots & & 0 & a_{n,n}\end{bmatrix}\begin{bmatrix}\Delta\epsilon_1(\lambda)\\\Delta\epsilon_2(\lambda)\\\Delta\epsilon_3(\lambda)\\\vdots\\\Delta\epsilon_n(\lambda)\end{bmatrix}=\begin{bmatrix}B_1(\lambda)\\B_2(\lambda)\\B_3(\lambda)\\\vdots\\B_n(\lambda)\end{bmatrix}$$

Therefore $\Delta\epsilon_i(\lambda)\equiv\epsilon_i(\lambda)-\epsilon_0(\lambda)$ (i = 1,2 …) can be determined. And the absolute absorption for i^{th} intermediate, ϵ_i (i = 1,2 …) can be obtained by solving this equation under different wavelength.

The analyses were performed as the following procedure:

The absorption spectra of the ground state $\epsilon_0(\lambda)$ was per-measured with the UV-vis spectrometer with the Rayleigh scattering component given by form of c_0+c_1/λ^4 was subtracted., then the laser was applied to excite *Np*HR and the absorption changes $\Delta A(\lambda,t)$ were recorded at each measured wavelength.

The data was fitted by eq. 4 and eq. 5 to obtain κ_i and $B_i(\lambda)$, and the fitting results was selected by brute-forcing $F_c\epsilon_i(\lambda) \geq 0$ because the negative part only originated from the depletion of P_0 by the flash excitation. The fitting result yielded the absolute absorption for i^{th} intermediate $\epsilon_i(\lambda)$ and the corresponding time delay constant τ_i. The number of intermediates, n was determined was determined by χ^2 test and the well fitted result with smallest n was adopted as the model to describe *Np*HR photocycle kinetics.

The global fitting was performed using homemade scripts written in Python3 with NumPy and SciPy packages.

4.1.4 Quaternary structure analysis method

Visible CD spectroscopy and size-exclusion chromatography (SEC) were used for quaternary structure analysis of *Np*HR mutants.

The CD measurement condition was the same as described previously in section 2.2.4. The samples after light illumination were prepared with a homemade LED array (12 × LEDs, 590 nm, total output power ~10W, OptoSupply). The samples were illuminated for 15 h under 4°C with cooling by circulating water to analysis the effect of photocycle in the quaternary structure.

SEC measurement was conducted using a Superdex^{TM} 200 Increase 10/300 GL column with AKTA pure™ 25 M1 system (Cytiva). The *Np*HR samples for SEC were in 10 mM PIPES buffer (pH 7.0) with 0.1 % DDM and 100 mM NaCl or 67 mM Na_2SO_4 for Cl^--bound or -free samples, respectively. 200 µL samples containing ~ 27 µM *Np*HR (Abs_{580} = 1.5) were applied to the column pre-equilibrated with the same buffer. The flow rate was 0.75 mL/min and the eluted *Np*HRs were detected by the absorption at 580 nm.

4.1.5 Ultra-low temperature DNP enhanced ^{15}N detection MAS ssNMR sample preparation and measurement method

The *Np*HR-DMPC proteoliposome preparation was the same as that used for ^{13}C detection ssNMR measurement (section 3.2.1). Preparation, while [$^{15}N_\zeta$-Lys]-*Np*HR was used for this DNP ^{15}N detection ssNMR measurement. The proteoliposome that contain [$^{15}N_\zeta$-Lys]-*Np*HR in Cl^--bound or -free state were washed once with 100 µL deuterated "DNP juice", which contains 1:3:6 of the buffer solution: D_2O: glycerol-d_8 (v/v/v). The buffer solution contained 200 mM AMUPoL[102], 500 mM sodium acetate with 1M NaCl or 670 mM Na_2SO_4 (pH 5.0), so that the final concentration becomes 20 mM AMUPoL, 50 mM sodium acetate with 100 mM NaCl or 67 mM Na_2SO_4 (pH 5.0) to ensure the ion concentration in the sample. AMUPoL, which chemical structure shows in Fig. 22 is a stable biradical. It is used as the polarizing agent to generate nuclei hyper polarization from the electron spins of its unpaired electrons for DNP sensitivity enhancement. The suspended sample in the "DNP juice" was well vortexed and centrifuged to remove extra "DNP juice", then packed into a JEOL Si_3N_4 3.2 mm rotor using a home-made packing tool under centrifugation (2,000 ×g, 15 min). One of the rotor spacers was hollowed to contain $K^{79}Br$ powder for the temperature measurement[84].

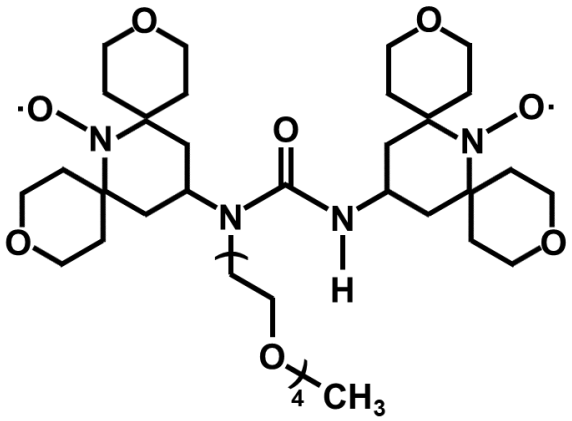

Fig. 22 The structure of AMUPoL.

The DNP MAS ssNMR experiments were carried out on a 700 MHz (B_0 = 16.4 T) JEOL ECAII spectrometer, equipped with a dedicated triple-resonance DNP MAS NMR probe, the custom-built closed-cycle helium gas sample spinning system and a 460 GHz gyrotron (output power ~ 10 W) for generating microwaves.[103] The probe was cooled with helium gas to 30 K and MAS frequency was set to ~ 7.5 kHz, while the sample temperature under MW irradiation was estimated to be 35 K using $K^{79}Br^{84}$. The data were processed and analyzed with the JEOL Delta software. The detailed parameters of the pulse sequences and data processing are described in the appendix.

4.2 Results

I tried to prepare the mutants following a concept of: (1) alanine into valine (2) other 19 amino acids into alanine. Because the EC part of helix D is a transmembrane region which favors hydrophobic residues and alanine is the smallest hydrophobic amino acid, while valine is the second smallest hydrophobic amino acid. Such mutation minimized the influence to the protein folding by the sidechain of amino acids therefore suitable to analysis if those residues involve in Cl- uptake process with conformational changes. I successfully obtained the following three mutants: Val161 to Ala, Gly163 to Ala and Ala165 to Val, which are referenced to as V161A, G163A and A165V, respectively. T To elucidate the functional significance of these residues in the ion uptake process, I measured the molecular properties of these mutants and compared them with WT *Np*HR. This comparative analysis aimed to clarify the relationship of these residues with the Cl^- uptake process.

4.2.1 Cl^- dependent absorption changes

I first analyzed Cl- binding affinity of these *Np*HRs by dissociation constant (K_d) through Cl^- dependent absorption changes. Fig. 23 shows the retinal absorption in Cl^--bound and -free state and the Cl^- dependent absorption changes are showed in Fig. 24. K_d was calculated with a Hill equation. The values of λ_{max} and the fitting results of the Hill equation is summarized in Table. 1.

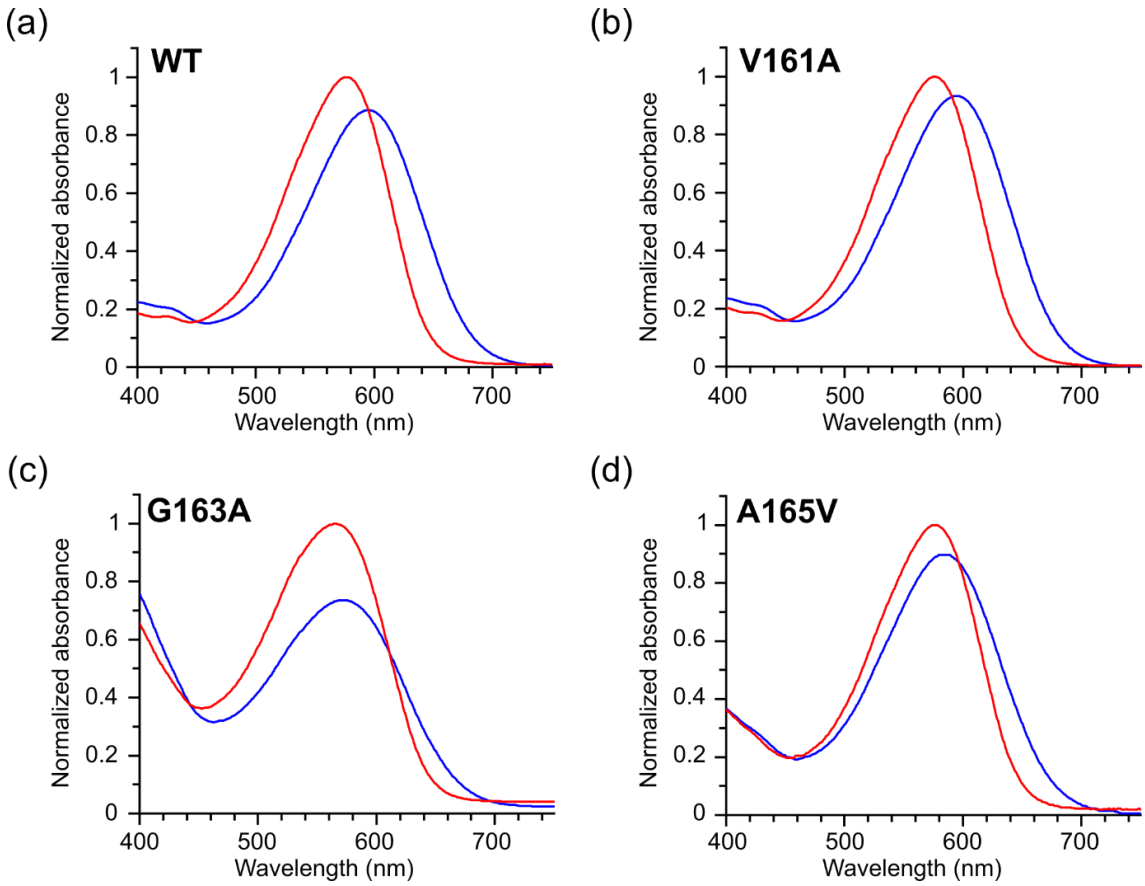

Fig. 23 Cl^- concentration-dependent absorption change observed for *Np*HR in wild type and mutants. (a-d) Visible spectra recorded with Cl^--bound (100 mM Cl^-, red) and -free (blue) state WT *Np*HR, V161A, G163A and A165V mutants, respectively.

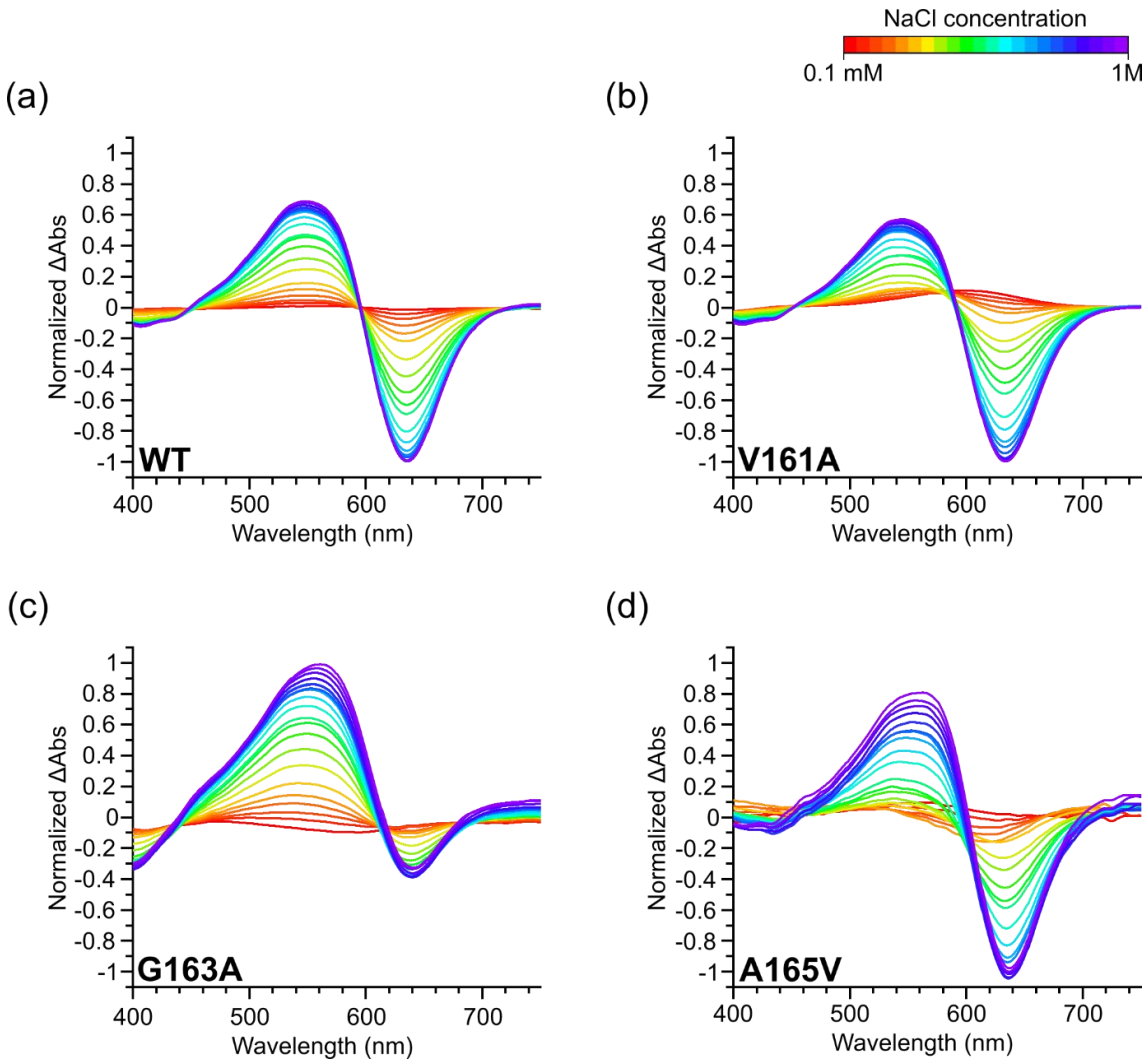

Fig. 24 Difference visible spectra of *Np*HRs taken under different Cl^- concentration relative to the Cl^--free condition. (a-d) WT, V161A, G163A and A165V mutant, respectively. The positive absorbance maxima changes are found at 552 nm, 547 nm, 559 nm and 568 nm for WT, V161A, G163A and A165V, respectively.

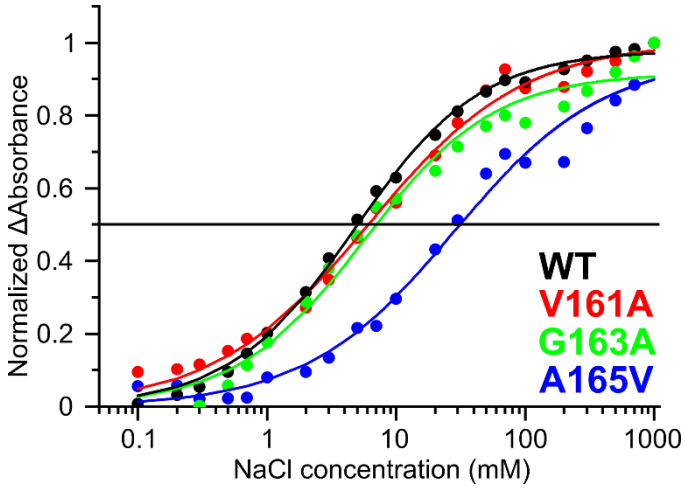

Fig. 25 Change of the absorption maximum *versus* Cl^- concentration. Each curve is

normalized by the maximum positive change data in Fig. 24. The solid curves show the best-fit results using the hill equation $\Delta Abs_{\mathrm{rel.}} = \frac{[Cl^-]^n}{K_\mathrm{d}^n + [Cl^-]^n}$.

Table 1. Absorption maxima (λ_{max}), Cl^- dissociation constants (K_d), Hill coefficients (n) for solubilized WT *Np*HR and the mutants in 0.1 % DDM at pH 7.0.

Opsin Type	λ_{max} / nm		K_d / mM	n
	Cl^- free	100 mM NaCl		
WT	595	577	4.8	0.89
V161A	595	576	6.1	0.73
G163A	575	560	5.7	0.87
A165V	584	576	28	0.75

V161A mutant showed almost identical Cl- dependent retinal absorption and K_d. For G163A mutant, we observed ~17 nm and ~20 nm blue-shift relative to WT for the Cl^--bound and -free state, respectively. For A165V mutant, no shift was observed for the Cl^--bound state, while ~10 nm blue-shift was observed for the Cl^--free state. The only altered absorption in Cl^--free state was reported previously for F150 mutants of *Np*HR with unstable trimeric assembly[39]. This mutant also has a less Cl^- binding affinity than WT through the increased K_d value.

4.2.2 Transient absorption spectroscopy analysis of photocycle

Next, to compare the photocycle properties, flash laser-induced transient absorption changes were measured for WT and the mutants with 1 M NaCl (Fig. 26). The photocycle was analyzed by n-components sequential irreversible model in the method section (section 4.1.3) by global fitting. The fitting results yielded the number of intermediates was 4 and the fraction of *Np*HR excited by laser, $F_c \approx 0.12$. Then a 4-components sequential irreversible model: $P_0 \rightarrow P_1 \overset{\tau_1}{\rightarrow} P_2 \overset{\tau_2}{\rightarrow} P_3 \overset{\tau_3}{\rightarrow} P_4 \overset{\tau_4}{\rightarrow} P_0$ was adopted.

The time constants τ_i $(i = 1, \cdots, 4)$ of corresponding component intermediate P_i are tabulated in Table 2. The absorption spectrum of each component is shown in Fig. 27 and the population change of each intermediate during photocycle is shown in Fig. 28. P_1 and P_2 showed blue-shifted absorption with identical λ_{max} and generated in the early photocycle, therefore been identified as L_1 and L_2 intermediates (Fig. 27 orange and green lines). P_3 showed the typical double-peak absorption and been identified as the N-O quasi-equilibrium (Fig. 27 red lines). P_4 shows the almost identical absorption as P_0 and generated in the last, therefore was assigned as the *Np*HR′ state (Fig. 27 purple lines). The photocycle of V161A was almost identical to WT, while those for G163A and A165V were different. The smaller peak of the differential absorption (Δ_{abs}) at 650 nm found in the late photocycle (~1 ms) for G163A and A165V (Fig. 26 c and d) suggests the suppressed formation of the O intermediate; this was consistent with the only minor peak seen around 650 nm in P_3 absorption spectra (Fig. 27 c and d). Moreover, G163A and A165V mutant showed, respectively, much slower (Fig. 26 c and Fig. 28 c) and faster recovery of Δ_{abs} at 580 nm (Fig. 26 d and Fig. 28 d), i.e., significantly

longer/shorter τ_3 and τ_4 than WT, respectively. Altogether, these observations indicated the mutations in the EC side of helix D perturb the late photocycle indeed, including the Cl^- release/uptake steps. Interestingly, τ_4 of A165V was ~5 times shorter than that for WT and close to that reported for the F150A mutant. The F150A mutant had unstable trimeric assembly and resulted into the faster photocycle while inefficient ion pumping

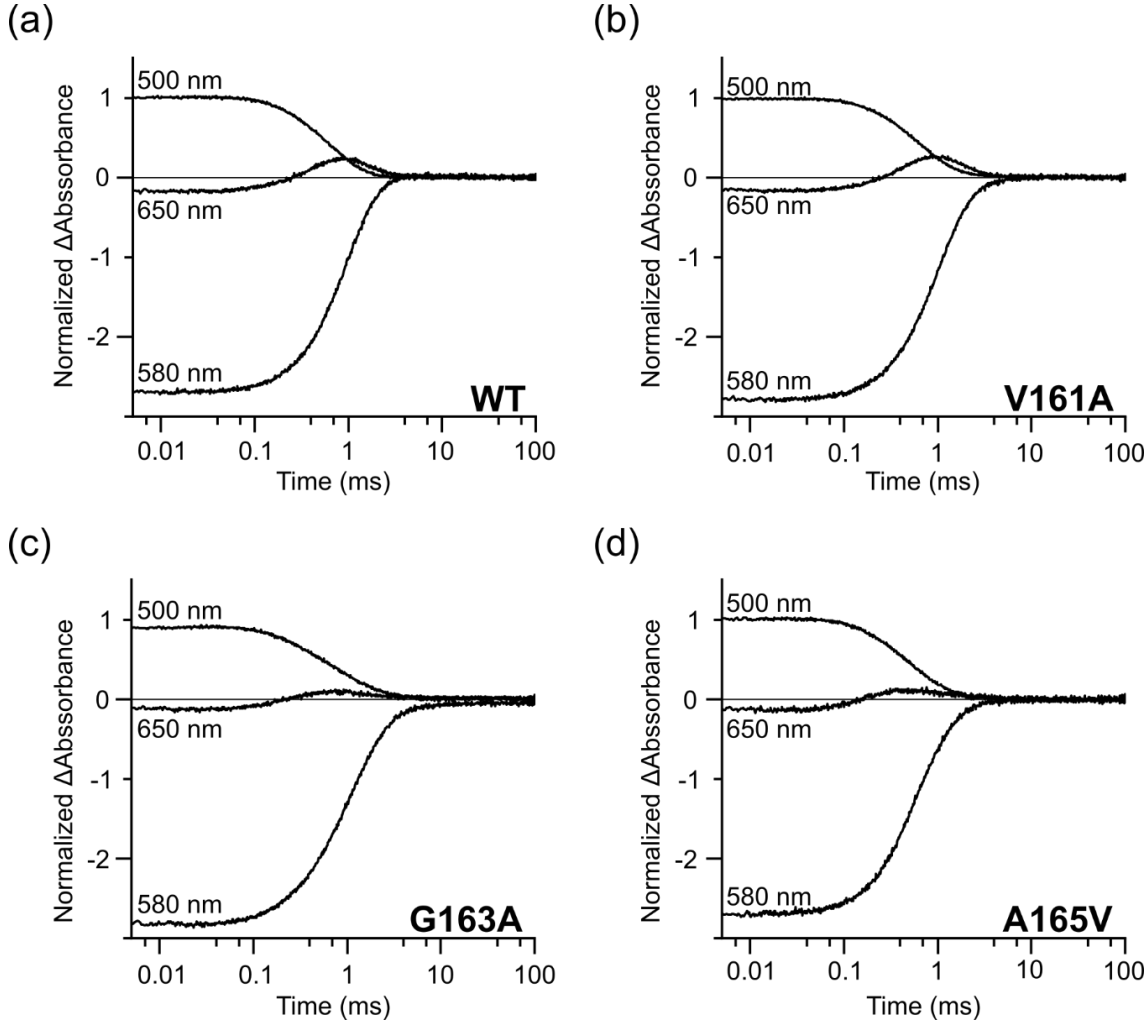

Fig. 26 Photocycle of *Np*HR in wild type and mutants. (a-d) Flash-induced transient absorption measured at indicated wavelengths is plotted as a function of time for WT *Np*HR, V161A, G163A and A165V mutants, respectively.

Table 2. The time constants of the components in sequential irreversible model calculated by global fitting. The corresponding intermediate formations are marked in the table.

Opsin	τ_1 / ms	τ_2 / ms	τ_3 / ms	τ_4 / ms
Type	$L_1 \rightarrow L_2$	$L_2 \rightarrow$ N-O	N-O$\rightarrow$*Np*HR′	*Np*HR′$\rightarrow$*Np*HR
WT	0.228	0.307	0.792	34.283
V161A	0.167	0.395	0.866	30.169
G163A	0.043	0.324	1.085	64.500
A165V	0.101	0.111	0.556	6.500

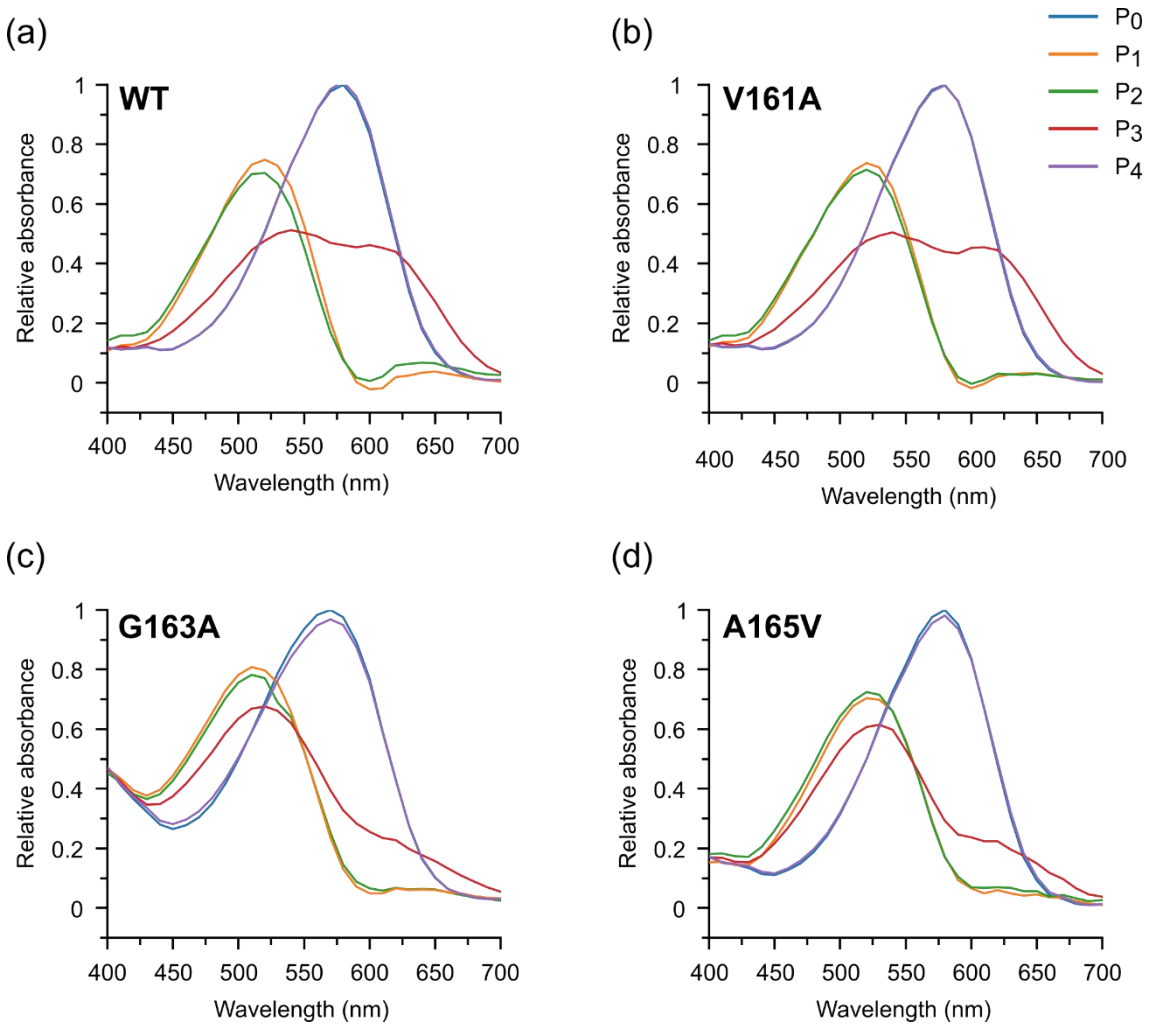

Fig. 27 The simulated absorbance spectra of the photochemically defined states P_0-P_4 obtained from the transient absorption data by global fitting analysis using the irreversible sequential model. P_0 represents the ground state. P_1 and P_2 exhibit a blue shift of λ_{max} to ~520 nm, therefore were identified as L_1 and L_2 intermediate, respectively. P_3 shows the typical doublet absorption at λ_{max} ~540 nm and ~600 nm and was identified as the N-O quasi-equilibrium. P_4 shows the almost identical absorption as P_0, therefore was assigned as the *Np*HR′ state. The P_3 (N-O quasi-equilibrium) spectra of G163A and A165V are significantly different to those for WT and V161A mutant, exhibiting much less red-shifted O component.

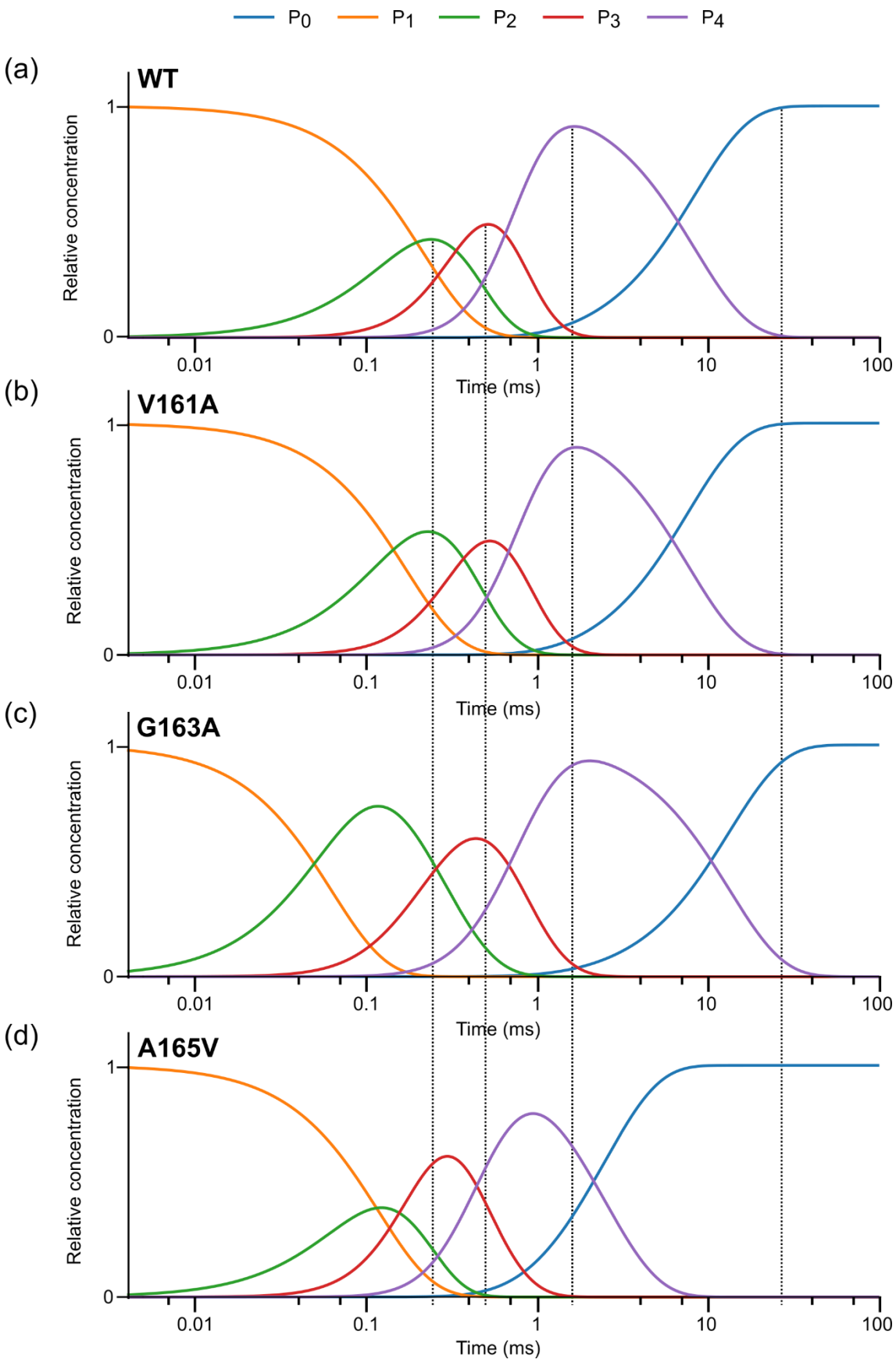

Fig. 28 The concentration profiles for the photochemically defined states obtained from the transient absorption data by global fitting analysis using the irreversible sequential model.

Dash lines indicate the times when the P_2-P_0 reached the maximum concentration for WT *Np*HR in the photocycle.

4.2.3 Quaternary structure analysis

Results in the previous section showed the A165V mutant shared similar properties with F150 mutants, which were reported have unstable trimeric assemble. Therefore, I measured the quaternary structures of the mutants prepared in this work by visible CD spectroscopy both in the Cl^--bound and -free state (Fig. 29). It is interesting that the typical bi-polar CD pattern[39] seen for A165V mutant in the Cl- bound state was perturbed in the Cl^--free state. This result suggested the A165Vmutant preserved a trimeric structure in the Cl^--bound state, while it partially dissociates in the Cl^--free state.

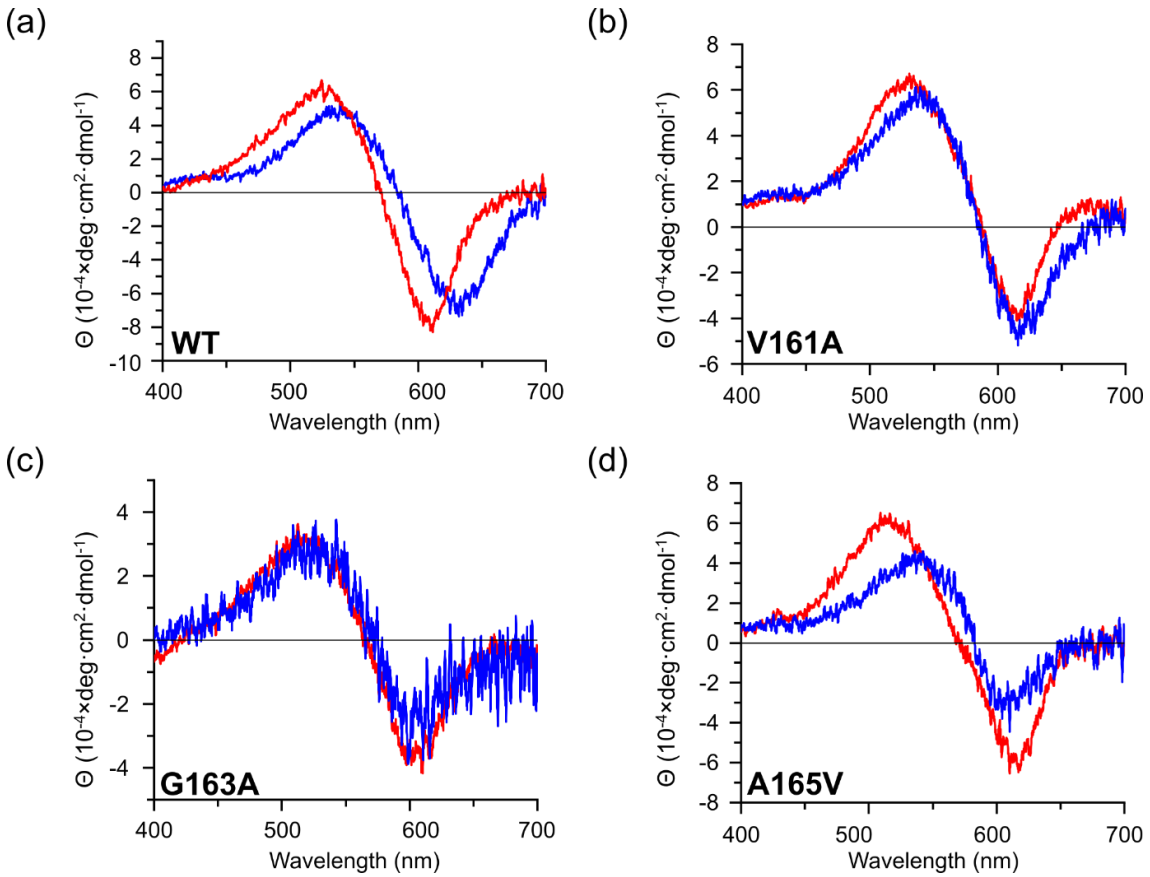

Fig. 29 Visible CD spectra of *Np*HRs in Cl^--bound (red) and -free (blue) state. (a-d) WT *Np*HR, V161A, G163A and A165V mutant, respectively. The bi-polar CD couplet did not change for WT *Np*HR, V161A and G163A mutants by the Cl^- depletion. The CD spectra of A165V in Cl^--free state showed a smaller negative peak with the red-shifted positive peak that indicates the trimer dissociated into monomer. While some trimer components remained in the sample and resulted in the negative peak still observable.

To confirm the trimer dissociation, I analyzed the protein samples with size-exclusion chromatography (SEC) measurement was applied (Fig. 30). Although the solubilized A165V mutant in 100 mM NaCl with 0.1% DDM was in trimeric form from its molecular size, we

observed the dissociation of the A165V trimer under low Cl^- concentration (Fig. 30b blue line). Upon removal of Cl^- ion by buffer exchange, the trimer dissociated into monomer form. To evaluate whether the dissociation was reversible, the whole *Np*HR fractions of the Cl^--free A165V sample (elution volume of 12–15 mL) were re-analyzed after those fractions were gathered, buffer exchanged to Cl^--containing buffer and concentrated ((Fig. 30b, black line). The result showed the trimeric component ratio was not recovered, which means the dissociation was irreversible.

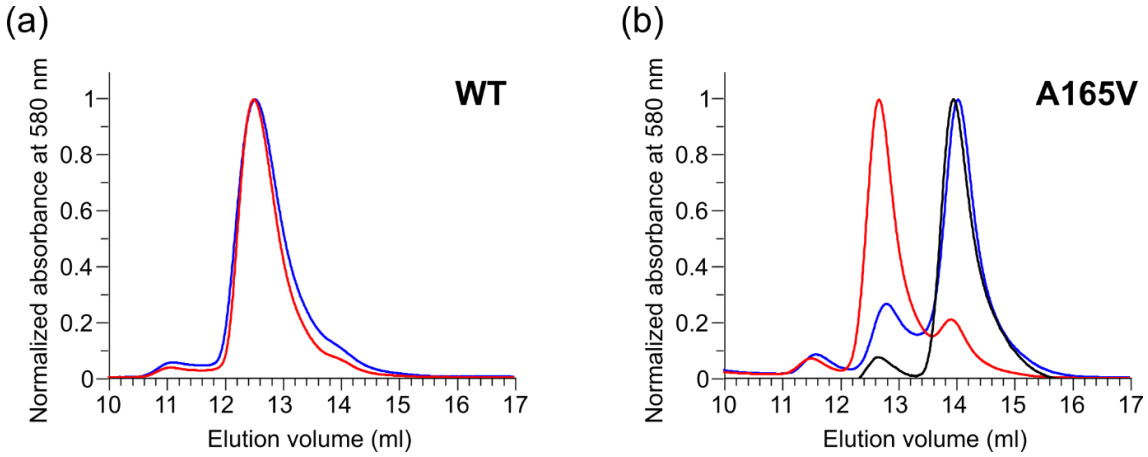

Fig. 30 Size-exclusion chromatograms recorded for WT *Np*HR (a) and A165V mutant (b) in Cl^--bound (red) and -free (blue) state. The black line shows the data taken the A165V monomer fraction (elution volume of 12–15 mL), back-exchanged to the Cl^--containing (100 mM NaCl) buffer and applied to SEC again after concentration. The *Np*HR elution peaks at 12.6 mL and 14.2 mL corresponds to the trimeric and monomeric form, respectively.[39] The small lead peak around 11-12 mL is from *Np*HR multimers.

The above results showed the A165V mutant undergoes some major conformational change in the Cl^--free state that irreversibly disrupts its trimeric assembly upon Cl^- depletion. To confirm whether this conformational change also exists for the O intermediate during photocycle, I measured visible CD spectra of the A165V mutant before and after light illumination with 100 mM NaCl. After 15 h of illumination under 590 nm, the trimeric A165V typical bi-polar CD couplet[39] was altered: the intensity of red-shifted negative peak was diminished, and the blue-shifted positive peak shifted toward λ_{max} (Fig. 31). This indicated the A165V trimer dissociates into monomers over the course of the photocycle in the O intermediate.

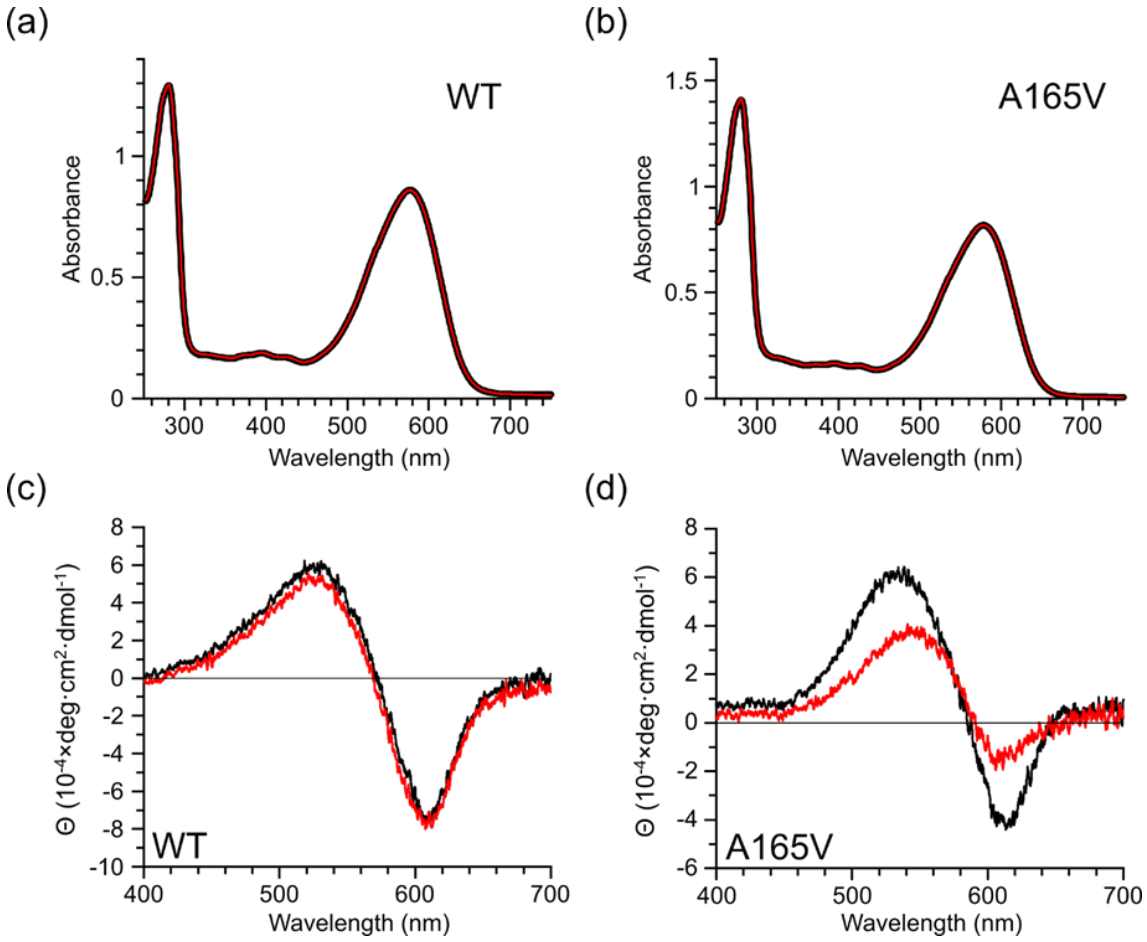

Fig. 31 The light-induced trimer dissociation of A165V mutant. (a and b) UV-vis spectra of WT and A165V *Np*HR before (black) and after (red) illumination under 590nm. (c and d) Visible CD spectra of them before (black) and after (red) illumination. The UV-vis spectra showed the illumination did not cause bleaching, while A165V showed an altered CD spectrum after illumination, which indicated trimer dissociation.

4.2.4 DNP enhanced ^{15}N detection MAS ssNMR measurement for pSB under ultra-low temperature

The above spectroscopic data showed the chromophore absorption wavelength was affected by the mutation in helix D and A165V mutant was suggested to disrupt its trimeric assembly in Cl^--free state with structural changes. To gain more insight into this, I measured the ^{15}N chemical shift of *Np*HR pSB (N_ζ of Lys256) by DNP-enhanced ssNMR measurements. Because the chromophore absorption wavelength of microbial rhodopsins is sensitive to the environment around the protonated Schiff base (pSB), which can be used to reveal the folding condition of *Np*HR.

One-dimensional DNP-enhanced ^{15}N spectra of [$^{15}N_\zeta$-Lys]-WT and G163A and A165V mutants are shown in Fig. 32. The DNP enhancement factor ε was ~30 at T ~30 K for all samples. This corresponds to roughly 300-fold sensitivity gain together with the effect from the sample cooling relative to the conventional room-temperature ssNMR. This sensitivity

allowed us to detect weak ^{15}N signals from minor conformation within hours. The ultra-low temperature DNP condition freeze-traps dynamical conformational ensembles thermally sampled by *Np*HR. Therefore, metastable conformational substates around pSB can be detected as ^{15}N peaks with distinct chemical shifts, where the linewidth and integral intensity of each peak reflects the degree of conformation disorder and population of each substate, respectively. The pSB of Cl^--bound WT and A165V exhibited a single ^{15}N peak at ~162 ppm, however that of G163A was split into three peaks at ~161, ~168 and ~180 ppm (indicated by dash lines in Fig. 32b), corresponding to three major conformational substates around the chromophore. The major peak was now at ~168 ppm while the peak at ~162 ppm seen also with WT and A165V mutant became much smaller. An additional broad peak was visible at ~180 ppm, for which an extensive conformational disorder was suggested from the broadened linewidth. This showed that the pSB environment for G163A was significantly different from that for WT and A165V mutant in the Cl^--bound state. Then, the same measurement was carried out for *Np*HRs in the Cl--free state (Fig. 32c). The similar three substates were observed for all three samples. Again, spectra of WT and A165A were similar: the peak at ~162 ppm remained major with additional minor peaks seen at ~168 ppm and ~ 180 ppm. For the G163A mutant, the broad peak at ~180 ppm, corresponding to highly disordered substate, became highest in intensity while that at ~160 ppm was much smaller. Interestingly, the spectra of WT and A165V in the Cl^--free state looks similar to that of G163A mutant in Cl^--bound state, suggesting similar pSB structure.

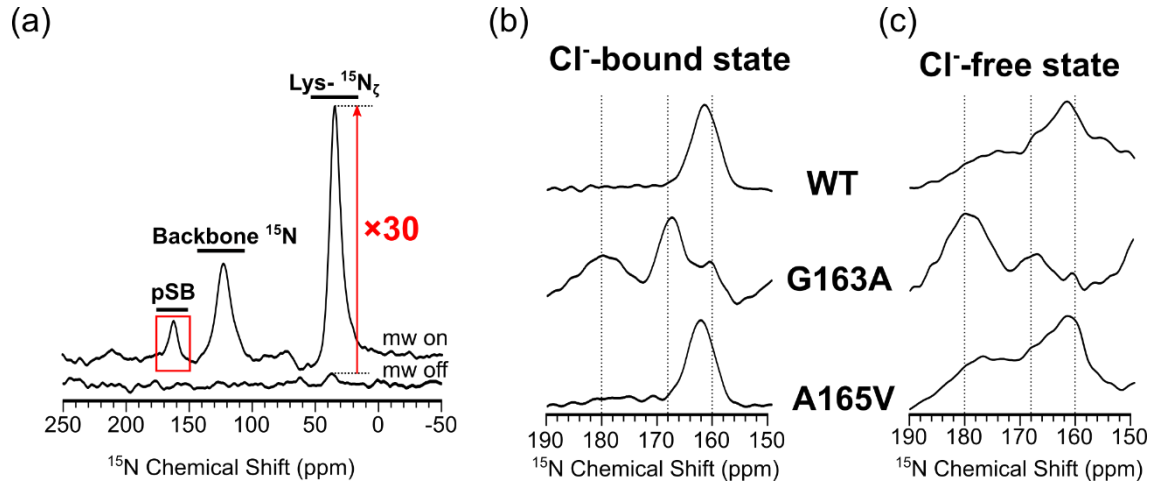

Fig. 32 ^{15}N NMR analysis for [$^{15}N_\zeta$-Lys]-*Np*HR WT and mutants. (a) Full NMR spectrum recorded for WT *Np*HR in the Cl^--bound state with and without the microwave irradiation for DNP enhancement. The pSB region blown up in (b,c) is indicated with the red rectangle. (b, c) Blowups of the pSB region of the CP-based DNP-enhanced ^{15}N spectra recorded for *Np*HRs in Cl^--bound (b) and -free state (c).

The ^{15}N chemical shift of pSB in many microbial rhodopsins follow an empirical linear relationship between $1/\lambda_{max}$ and $1/d^2$ where d is the distance between the pSB ^{15}N atom and the counterion.[26,29,104,105] In case of *Np*HR, the Cl^- bound to pSB acts as the counterion. Fig. 33 shows the pSB ^{15}N chemical shifts we observed for *Np*HRs *versus* $1/\lambda_{max}$. The Cl^--bound WT and A165V followed this relationship. The chemical shift of the major ^{15}N peak observed for G163A mutant in the Cl^--bound state (~168 ppm) also matched this relationship well. However,

those for the two other peaks deviated from this relationship, suggested a lot looser coordination of counter ion. To summarize, I found for the first time that the O-like Cl^--free intermediate generally comprises three substates, and that the conformational setting for WT and A165V mutant are similar to each other, indicating similar protein folding. While G163A mutant is distinct with significantly more disordered conformation, perhaps a destabilized opsin fold. In particular, in G163A mutant the surrounding environments of pSB are very different and disordered compared with WT in terms of the geometry of the counterions. And the A165V mutant preserved a similar protein folding as the WT both in the Cl^--bound and -free state.

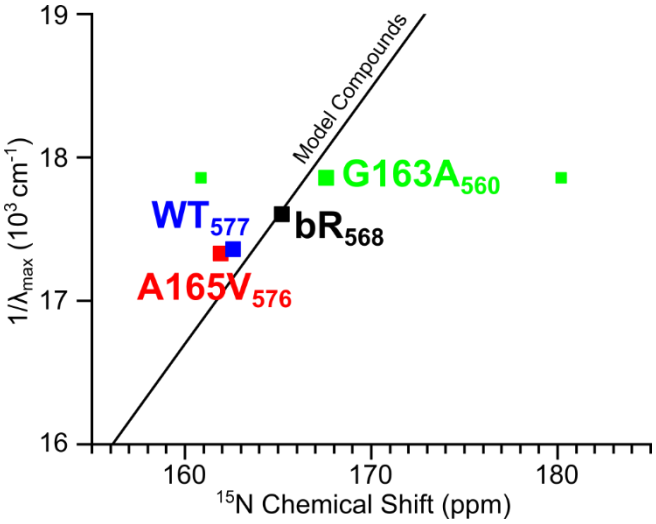

Fig. 33 The typical linear relationship between $1/\lambda_{max}$ and the pSB ^{15}N chemical shift observed for the counterion-bound retinal derivatives. Data taken with *Np*HR WT, A165V and G163A mutants are shown in blue, red and green, respectively. The two smaller green squares correspond to the two minor ^{15}N peaks observed for G163A. Data of the model compounds and bR_{568} are adopted from Hu et al[105].

4.3 Discussion

The G163A and A165V mutant showed altered retinal absorption and photocycle kinetics in the late photocycle, clearly showed the EC part of Helix D involved in the Cl^- uptake process.

The 163rd residue, which faces the retinal β-ionone ring, is highly conserved among rhodopsin family and is known as a color tuner.[106] Mutation in this residue was predicted to result in a blue shift of λ_{max} because the bulky sidechain pushes against the retinal β-ionone ring, distorting the polyene chain.[106,107] The G163A mutant indeed exhibited a blue-shifted λ_{max} both in the Cl^--bound and -free state. Additionally, this mutant displayed a loosened structure, as indicated by the peak-splitting in the pSB ^{15}N ssNMR spectra. Such loosened structure therefore may be related to the steric crush between the bulky sidechain and the retinal β-ionone ring.

Importantly, in the Cl^--free state of G163A mutant, ion depletion significantly increased the intensity of the broad ^{15}N peak at ~180 ppm, suggesting a more disordered conformation. Such peak-splitting also occurred for the A165V mutant and WT in the Cl^--free state, indicating

that *Np*HR undergoes looser packing around the retinal in this state. These results were consistent with the previous ssNMR data showing conformational changes between the Cl^--bound and -free state.

It is worth noting that A165V disrupted its trimeric structure in the Cl^--free state or during the photocycle involving the O intermediate formation through some major structural event in the late photocycle. ^{15}N chemical shift of pSB further indicated that the protein folding of A165V was similar with that in WT, suggesting the trimer dissociation is not due to overall protein denaturing but related with some specific conformational event occurred during the Cl^- uptake process.

The 165th position is in the monomer-monomer interface, however, the Ala to Val mutation seems not introduce severe steric crush there (Fig. 34). Also, the above presented CD data confirms the A165V mutant retains the trimeric formation in the Cl^--bound state (Fig. 29). This indicates that the trimer dissociation is not simply due to the bulky valine sidechain and needs some additional conformational event. According to the ssNMR CSP data, the EC part of helix D implements large conformational change in the late photocycle; this may be thought as an outward helix movement considering the disruption of the trimer in the late photocycle. This is supported by previous X-ray data[38,59]. The X-ray structural studies showed the Tyr124 sidechain rotates toward the helix D by Cl^- depletion, which is close to the 165th position (Fig. 35a). This rotation may push the EC part of helix D outward toward the adjacent monomer. A165V mutant, which contains a bulky sidechain on the interface at 165th position, disturbs its trimer integrity because of such structural changes (Fig. 35b). From these results and the previous X-ray crystallographic studies[38,59], we hypothesize that an outward movement of the EC part of helix D was involved in the late photocycle including the O-intermediate. More specifically, since its mutation to valine with a bulky sidechain disturbed the trimer integrity upon the D-helix movement during the photocycle.

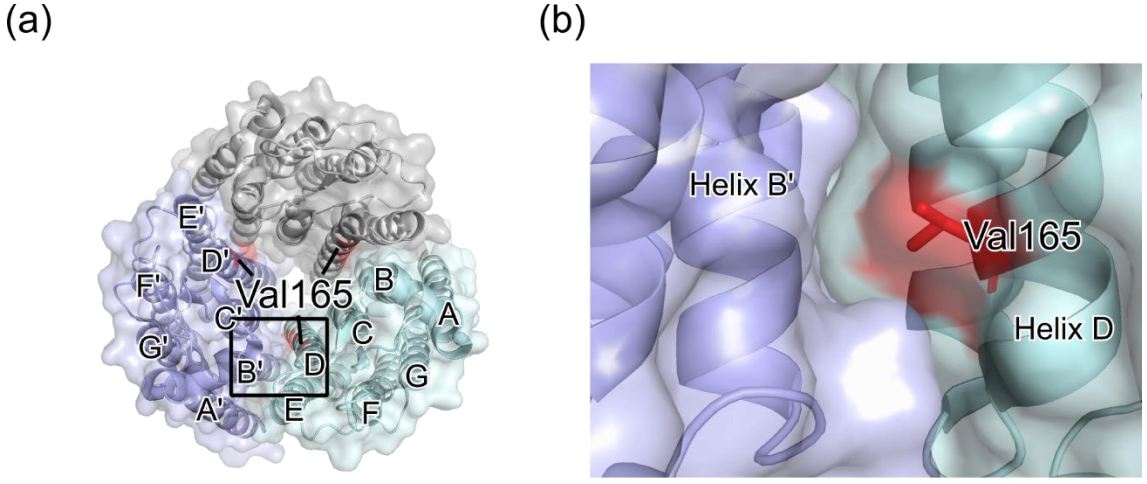

Fig. 34 (a) *Np*HR trimeric structure with the 165th position highlighted in red. (b) A lateral blow-up view at the trimer interface, in which the 165th residue is mutated to valine. The monomer structure is based on the Cl^--bound state *Np*HR (PDB entry 3A7K[38]). For the trimeric assembly, we used the homologous model from *Hs*HR (PDB entry 1E12[40]) because the protein expressed in *E. coli* lacks bacterioruberin, which is necessary for the trimeric assembling in the 3A7K trimeric form. The structure visualization and the amino acid replacement for A165V

were performed using PyMOL.

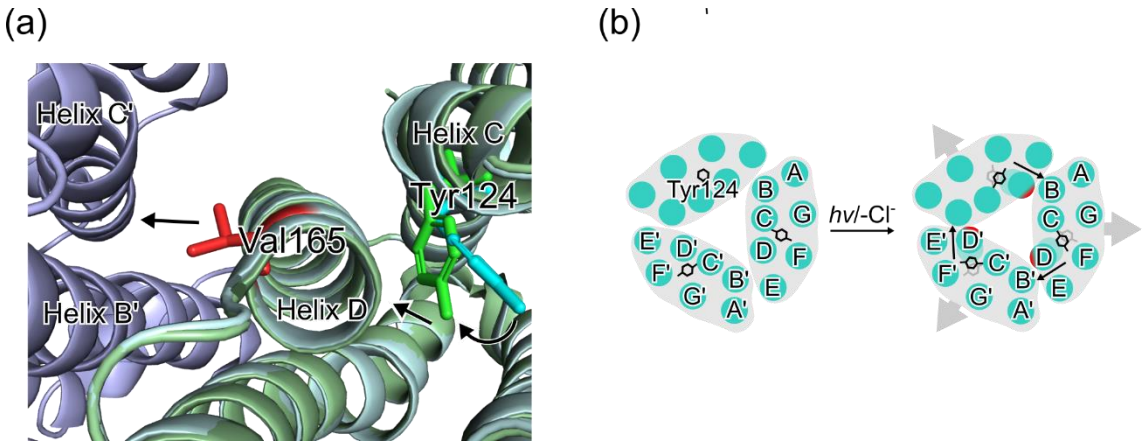

Fig. 35 The proposed trimer dissociation mechanism for the A165V mutant. The Cl^--free O-like structure (green, PDB entry 3QBG[59]) and the Cl^--bound state *Np*HR (cyan, PDB entry 3A7K[38]) is presented. The sidechain of Tyr124 rotates toward the helix D by Cl^- depletion or during the O intermediate formation in the photocycle, which may push the helix D outward and lead to the trimer dissociation. The structure visualization and the amino acid replacement for A165V were performed using PyMOL.

Based on the above discussion, I propose a model for the Cl^- uptake process comprising a movement of the EC side of the helix D as a crucial step for the ion uptake. It has been reported that His100 in B-C loop, Arg123 in helix C, Arg176 in D-E loop and Glu234 in F-G loop together with the nearby water molecules forms a putative Cl^- uptake pathway[38]. Indeed, some open space between helices (colored cyan in Fig. 36) can be found around these residues in the crystal structure of the O-like Cl^--free blue form[59]. However, these residues are totally covered by the B-C loop and there is no obvious ion entrance path toward this space in crystal structures. The His100, Arg176 and Glu234 are clustered near the EC part of helix D, meanwhile helix D is not tightly covered by B-C loop. Therefore, an outward movement of helix D may be helpful to create a path for the Cl^- entrance. The Glu234[47] in F-G loop and Tyr124[38] in helix C are known as key residues for Cl^- transport and showed sidechain rotation between Cl^--bound and -free state. These movements together with the outward movement of the EC part of the helix D can allow Cl^- ion to enter the inner protein space and captured, which is further transported to the pSB region (Fig. 37).

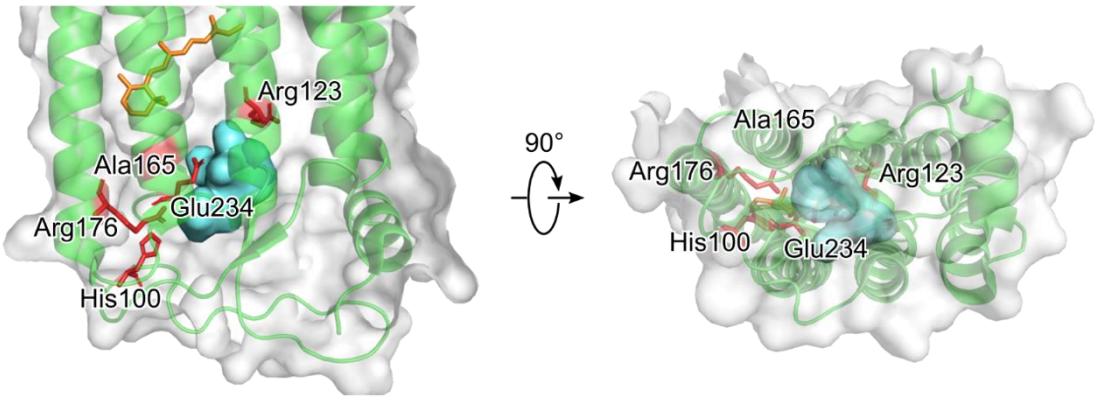

Fig. 36 Side view from the trimer interface side and the top view from the EC side of the Cl^--free blue form *Np*HR (PDB entry 3QBG[59]). The putative key residues for Cl^- uptake and Ala165 are shown in red. The inter-helix space mentioned in the text is highlighted in cyan. The retinal is colored orange. The protein structures were visualized using PyMOL.

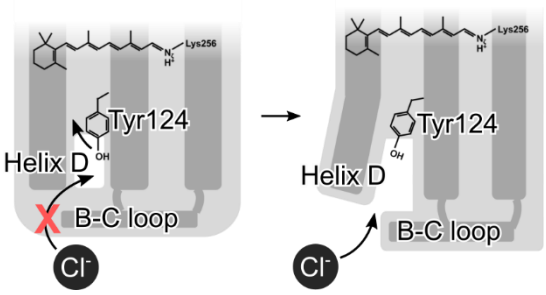

Fig. 37 Schematics for the Cl^- uptake process with the observed structural changes. The hydrophobic B-C loop covers the Cl^- entrance path in the Cl^--bound state and prevents the Cl^- uptake. The outward movement of helix D permits expose the inter-helix space to solvent which allows the Cl^- to enter.

5 Conclusion

In this study, I combined ssNMR spectroscopy with photochemical property analyses of wild-type (WT) and mutant forms of *Np*HR within a lipid bilayer environment, focusing on the late photocycle, particularly the Cl^- ion uptake process.

In section 2, I first developed an unfolding and refolding protocol for the sample preparation of ^1H detection ssNMR. This novel *Np*HR preparation method significantly improved the ^1H_N back-exchange ratio in perdeuterated *Np*HR samples, which is critical for achieving high resolution ^1H detection ssNMR spectra with reduced signal losses. By using this sample preparation method, I obtained high quality 3D CANH detection ssNMR spectra of *Np*HR reconstituted in DMPC liposome both in the Cl^--bound and -free state with ~1 mg [^2H with ^1H_N]-*Np*HR sample. Peak shifting between the two states clearly indicated conformational differences between the Cl^--bound and -free state *Np*HR.

Then in section 3, I performed signal assignments with ^{13}C detection ssNMR measurements. The ssNMR results identified one of the major structural changes that occurred in the EC part of the helix D between the Cl^--bound and -free states.

Finally in section 4, mutants for residues in the EC part of helix D were prepared and their molecular properties were measured. Functional analyses revealed that residues in the EC part of Helix D, particularly Gly163 and Ala165, were involved in the Cl^- uptake process. DNP-enhanced ^{15}N NMR spectroscopy elucidated the geometry changes around the protonated Schiff Bases, highlighting the existence of three principal substates in the O-intermediate. Most importantly, a Cl^- uptake scheme was proposed based on the existing X-ray crystal structures with the newly revealed helix D outward movement in the extracellular region.

Overall, the results filled the gap between the limited structural changes reported for *Np*HR by X-ray crystallography and the relatively large molecular volume changes observed during the photocycle in spectroscopic studies. This work highlighted the utility of ssNMR methods in the structural studies of membrane proteins, offering site-specific structural information under near-physiological conditions.

References

1. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? *Nature Reviews Drug Discovery 2006 5:12*. 2006;5(12):993-996. doi:10.1038/nrd2199
2. Burley SK, Bhikadiya C, Bi C, et al. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res*. 2023;51(D1):D488-D508. doi:10.1093/NAR/GKAC1077
3. Sui H, Han BG, Lee JK, Walian P, Jap BK. Structural basis of water-specific transport through the AQP1 water channel. *Nature 2001 414:6866*. 2001;414(6866):872-878. doi:10.1038/414872a
4. Stock D, Leslie AGW, Walker JE. Molecular architecture of the rotary motor in ATP synthase. *Science (1979)*. 1999;286(5445):1700-1705. doi:10.1126/SCIENCE.286.5445.1700
5. Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The Limit of Accuracy of Protein Modeling: Influence of Crystal Packing on Protein Structure. *J Mol Biol*. 2005;351(2):431-442. doi:10.1016/J.JMB.2005.05.066
6. Guo Y. Be Cautious with Crystal Structures of Membrane Proteins or Complexes Prepared in Detergents. *Crystals (Basel)*. 2020;10(2):86. doi:10.3390/cryst10020086
7. Piper SJ, Johnson RM, Wootten D, Sexton PM. Membranes under the Magnetic Lens: A Dive into the Diverse World of Membrane Protein Structures Using Cryo-EM. *Chem Rev*. Published online 2021. doi:10.1021/ACS.CHEMREV.1C00837
8. Ashmore J, Carragher B, Rosenthal PB, Weis W. A resolution record for cryoEM. *Fac Rev*. 2021;10:64. doi:10.12703/R-01-000002
9. Castells-Graells R, Meador K, Arbing MA, et al. Cryo-EM structure determination of small therapeutic protein targets at 3 Å-resolution using a rigid imaging scaffold. *Proc Natl Acad Sci* 2023;120(37):e2305494120. doi:10.1073/PNAS.2305494120
10. Danmaliki GI, Hwang PM. Solution NMR spectroscopy of membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2020;1862(9):183356. doi:10.1016/J.BBAMEM.2020.183356
11. Osawa M, Takeuchi K, Ueda T, Nishida N, Shimada I. Functional dynamics of proteins revealed by solution NMR. *Curr Opin Struct Biol*. 2012;22(5):660-669. doi:10.1016/J.SBI.2012.08.007
12. Brown LS, Ladizhansky V. Membrane proteins in their native habitat as seen by solid-state NMR spectroscopy. *Protein Science*. 2015;24(9):1333-1346. doi:10.1002/PRO.2700
13. Ladizhansky V. Applications of solid-state NMR to membrane proteins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 2017;1865(11):1577-1586. doi:10.1016/J.BBAPAP.2017.07.004
14. Kriebel CN, Becker-Baldus J, Glaubitz C. Solid-State NMR Spectroscopy on Microbial Rhodopsins. In: *Methods in Molecular Biology*. Vol 2501. Humana Press Inc.; 2022:181-206. doi:10.1007/978-1-0716-2329-9_9
15. Xie H, Zhao Y, Zhao W, Chen Y, Liu M, Yang J. Solid-state NMR structure determination of a membrane protein in E. coli cellular inner membrane. *Sci Adv*. 2023;9(44). doi:10.1126/SCIADV.ADH4168

16. Wang S, Munro RA, Shi L, et al. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nature Methods 2013 10:10*. 2013;10(10):1007-1012. doi:10.1038/nmeth.2635
17. Park SH, Das BB, Casagrande F, et al. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature*. 2012;491. doi:10.1038/nature11580
18. Verardi R, Shi L, Traaseth NJ, Walsh N, Veglia G. Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proc Natl Acad Sci* 2011;108(22):9101-9106. doi:10.1073/PNAS.1016535108
19. Shahid SA, Bardiaux B, Franks T, et al. membrane-protein structure determination by solid-state nmr spectroscopy of microcrystals. Published online 2012. doi:10.1038/nmeth.2248
20. Shi L, Kawamura I, Jung KH, Brown LS, Ladizhansky V. Conformation of a Seven-Helical Transmembrane Photosensor in the Lipid Environment. *Angewandte Chemie International Edition*. 2011;50(6):1302-1305. doi:10.1002/ANIE.201004422
21. Wang S, Shi L, Okitsu T, Wada A, Brown LS, Ladizhansky V. Solid-state NMR ^{13}C and ^{15}N resonance assignments of a seven-transmembrane helical protein Anabaena Sensory Rhodopsin. *Biomol NMR Assign*. 2013;7(2):253-256. doi:10.1007/S12104-012-9421-Y
22. Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*. 2002;420(6911):99-102. doi:10.1038/nature01070
23. Narasimhan S, Folkers GE, Baldus M. When Small becomes Too Big: Expanding the Use of In-Cell Solid-State NMR Spectroscopy. *Chempluschem*. 2020;85(4):760-768. doi:10.1002/CPLU.202000167
24. Pfleger N, Wörner AC, Yang J, et al. Solid-state NMR and functional studies on proteorhodopsin. *Biochim Biophys Acta*. 2009;1787(6):697-705. doi:10.1016/j.bbabio.2009.02.022
25. Pfleger N, Lorch M, Woerner AC, Shastri S, Glaubitz C. Characterisation of Schiff base and chromophore in green proteorhodopsin by solid-state NMR. *J Biomol NMR*. 2008;40(1):15-21. doi:10.1007/s10858-007-9203-5
26. Kumagai S, Kawamura I. Solid-state NMR of the retinal protonated Schiff base in microbial rhodopsins. *Magnetic Resonance Letters*. Published online April 25, 2024:200132. doi:10.1016/j.mrl.2024.200132
27. Makino Y, Kawamura I, Okitsu T, et al. Retinal Configuration of ppR Intermediates Revealed by Photoirradiation Solid-State NMR and DFT. *Biophys J*. 2018;115(1):72-83. doi:10.1016/j.bpj.2018.05.030
28. Bajaj VS, Mak-Jurkauskas ML, Belenky M, Herzfeld J, Griffin RG. Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization–enhanced solid-state NMR. *Proceedings of the National Academy of Sciences*. 2009;106(23):9244-9249. doi:10.1073/pnas.0900908106
29. Creemers AFL, Klaassen CHW, Bovee-Geurts PHM, et al. Solid State ^{15}N NMR Evidence for a Complex Schiff Base Counterion in the Visual G-Protein-Coupled Receptor Rhodopsin. *Biochemistry*. 1999;38(22):7195-7199. doi:10.1021/bi9830157
30. Bivin DB, Stoeckenius W. Photoactive retinal pigments in haloalkaliphilic bacteria. *J Gen Microbiol*. 1986;132(8):2167-2177. doi:10.1099/00221287-132-8-2167

31. Rozenberg A, Inoue K, Kandori H, Béjà O. Microbial Rhodopsins: The Last Two Decades. *Annu Rev Microbiol*. 2021;75(Volume 75, 2021):427-447. doi:10.1146/ANNUREV-MICRO-031721-020452
32. Henderson R, Nature PU. Three-dimensional model of purple membrane obtained by electron microscopy. *nature*. 1975;257. doi:10.1038/257028a0
33. Govorunova EG, Sineshchekov OA, Li H, Spudich JL. Microbial rhodopsins: Diversity, mechanisms, and optogenetic applications. *Annu Rev Biochem*. 2017;86:845-872. doi:10.1146/ANNUREV-BIOCHEM-101910-144233/1
34. Kurihara M, Sudo Y. Microbial rhodopsins: wide distribution, rich diversity and great potential. *Biophys Physicobiol*. 2015;12:121. doi:10.2142/BIOPHYSICO.12.0_121
35. Lanyi JK, Duschl A, Hatfield GW, May K, Oesterhelt D. The primary structure of a halorhodopsin from Natronobacterium pharaonis. Structural, functional and evolutionary implications for bacterial rhodopsins and halorhodopsins. *Journal of Biological Chemistry*. 1990;265(3):1253-1260. doi:10.1016/S0021-9258(19)40006-9
36. Gradinaru V, Thompson KR, Deisseroth K. eNpHR: a Natronomonas halorhodopsin enhanced for optogenetic applications. *Brain Cell Biol*. 2008;36(1-4):129-139. doi:10.1007/s11068-008-9027-6
37. Zhang F, Wang LP, Brauner M, et al. Multimodal fast optical interrogation of neural circuitry. *Nature*. 2007;446(7136):633-639. doi:10.1038/nature05744
38. 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;396(3):564-579. doi:10.1016/j.jmb.2009.11.061
39. Tsukamoto T, Sasaki T, Fujimoto KJ, et al. Homotrimer Formation and Dissociation of pharaonis Halorhodopsin in Detergent System. *Biophys J*. 2012;102(12):2906-2915. doi:10.1016/j.bpj.2012.05.008
40. Kolbe M, Besir H, Essen LO, Oesterhelt D. Structure of the Light-Driven Chloride Pump Halorhodopsin at 1.8 Å Resolution. *Science*. 2000;288(5470):1390-1396. doi:10.1126/science.288.5470.1390
41. Oesterhelt D, Stoeckenius W. Rhodopsin-like Protein from the Purple Membrane of Halobacterium halobium. *Nature New Biology.* 1971;233(39):149-152. doi:10.1038/newbio233149a0
42. Lanyi JK. Bacteriorhodopsin. *Annu Rev Physiol*. 2004;66:665-688. doi:10.1146/ANNUREV.PHYSIOL.66.032102.150049
43. Yawoo H, Kandori H, Koizumi A. *Optogenetics*. (Yawo H, Kandori H, Koizumi A, eds.). Springer Japan; 2015. doi:10.1007/978-4-431-55516-2
44. Kikukawa T, Kusakabe C, Kokubo A, et al. Probing the Cl−-pumping photocycle of pharaonis halorhodopsin: Examinations with bacterioruberin, an intrinsic dye, and membrane potential-induced modulation of the photocycle. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 2015;1847(8):748-758. doi:10.1016/j.bbabio.2015.05.002
45. Varo G, Brown LS, Sasaki J, et al. Light-Driven Chloride Ion Transport by Halorhodopsin from Natronobacterium pharaonis. I. The Photochemical Cycle. *Biochemistry*. 1995;34(44):14490-14499. doi:10.1021/bi00044a027
46. Varo G, Needleman R, Lanyi JK. Light-Driven Chloride Ion Transport by Halorhodopsin from Natronobacterium pharaonis. II. Chloride Release and Uptake, Protein Conformation

Change, and Thermodynamics. *Biochemistry*. 1995;34(44):14500-14507. doi:10.1021/bi00044a028
47. Shibata M, Saito Y, Demura M, Kandori H. Deprotonation of Glu234 during the photocycle of Natronomonas pharaonis halorhodopsin. *Chem Phys Lett*. 2006;432(4-6):545-547. doi:10.1016/j.cplett.2006.10.111
48. Sato M. Roles of Ser130 and Thr126 in Chloride Binding and Photocycle of pharaonis Halorhodopsin. *J Biochem*. 2003;134(1):151-158. doi:10.1093/jb/mvg123
49. Shibasaki K, Shigemura H, Kikukawa T, et al. Role of Thr218 in the Light-Driven Anion Pump Halorhodopsin from *Natronomonas pharaonis*. *Biochemistry*. 2013;52(51):9257-9268. doi:10.1021/bi401295e
50. Sato M, Kikukawa T, Araiso T, et al. Ser-130 of Natronobacterium pharaonis halorhodopsin is important for the chloride binding. *Biophys Chem*. 2003;104(1):209-216. doi:10.1016/S0301-4622(02)00368-X
51. Kubo M, Kikukawa T, Miyauchi S, et al. Role of Arg123 in Light-driven Anion Pump Mechanisms of *pharaonis* Halorhodopsin. *Photochem Photobiol*. 2009;85(2):547-555. doi:10.1111/j.1751-1097.2009.00538.x
52. Sato M, Kubo M, Aizawa T, et al. Role of Putative Anion-Binding Sites in Cytoplasmic and Extracellular Channels of *Natronomonas pharaonis* Halorhodopsin. *Biochemistry*. 2005;44(12):4775-4784. doi:10.1021/bi047500f
53. Sato M, Kanamori T, Kamo N, Demura M, Nitta K. Stopped-Flow Analysis on Anion Binding to Blue-Form Halorhodopsin from *Natronobacterium pharaonis* : Comparison with the Anion-Uptake Process during the Photocycle. *Biochemistry*. 2002;41(7):2452-2458. doi:10.1021/bi011788g
54. Inoue K, Kubo M, Demura M, Kamo N, Terazima M. Reaction Dynamics of Halorhodopsin Studied by Time-Resolved Diffusion. *Biophys J*. 2009;96(9):3724-3734. doi:10.1016/j.bpj.2008.12.3932
55. Ihara K, Narusawa A, Maruyama K, Takeguchi M, Kouyama T. A halorhodopsin-overproducing mutant isolated from an extremely haloalkaliphilic archaeon *Natronomonas pharaonis*. *FEBS Lett*. 2008;582(19):2931-2936. doi:10.1016/j.febslet.2008.07.030
56. Kouyama T, Kawaguchi H, Nakanishi T, Kubo H, Murakami M. Crystal Structures of the L_1, L_2, N, and O States of pharaonis Halorhodopsin. *Biophys J*. 2015;108(11):2680-2690. doi:10.1016/j.bpj.2015.04.027
57. Chizhov I, Engelhard M. Temperature and Halide Dependence of the Photocycle of Halorhodopsin from Natronobacterium pharaonis. *Biophys J*. 2001;81(3):1600-1612. doi:10.1016/S0006-3495(01)75814-6
58. Hasegawa C, Kikukawa T, Miyauchi S, et al. Interaction of the Halobacterial Transducer to a Halorhodopsin Mutant Engineered so as to Bind the Transducer: Cl^- Circulation Within the Extracellular Channel†. *Photochem Photobiol*. 2007;83(2):293-302. doi:10.1562/2006-06-09-RA-916
59. 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;413(1):162-176. doi:10.1016/j.jmb.2011.08.021
60. Guijarro J, Engelhard M, Siebert F. Anion Uptake in Halorhodopsin from *Natromonas*

pharaonis Studied by FTIR Spectroscopy: Consequences for the Anion Transport Mechanism. *Biochemistry*. 2006;45(38):11578-11588. doi:10.1021/bi060753j

61. Hackmann C, Guijarro J, Chizhov I, Engelhard M, Rödig C, Siebert F. Static and Time-Resolved Step-Scan Fourier Transform Infrared Investigations of the Photoreaction of Halorhodopsin from Natronobacterium Pharaonis: Consequences for Models of the Anion Translocation Mechanism. *Biophys J*. 2001;81(1):394-406. doi:10.1016/S0006-3495(01)75708-6
62. Polenova T, Gupta R, Goldbourt A. Magic angle spinning NMR spectroscopy: A versatile technique for structural and dynamic analysis of solid-phase systems. *Anal Chem*. 2015;87(11):5458-5469. doi:10.1021/AC504288U
63. Arnold AA, Marcotte I. Studying natural structural protein fibers by solid-state nuclear magnetic resonance. *Concepts in Magnetic Resonance Part A*. 2009;34A(1):24-47. doi:10.1002/CMR.A.20132
64. Ward ME, Brown LS, Ladizhansky V. Advanced solid-state NMR techniques for characterization of membrane protein structure and dynamics: Application to Anabaena Sensory Rhodopsin. *Journal of Magnetic Resonance*. 2015;253:119-128. doi:10.1016/j.jmr.2014.11.017
65. Medeiros-Silva J, Jekhmane S, Paioni AL, et al. High-resolution NMR studies of antibiotics in cellular membranes. *Nature Communications 2018 9:1*. 2018;9(1):1-10. doi:10.1038/s41467-018-06314-x
66. Delaforge E, Kragelj J, Tengo L, et al. Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *J Am Chem Soc*. 2018;140(3):1148-1158. doi:10.1021/JACS.7B12407
67. Vogeley L, Sineshchekov OA, Trivedi VD, Sasaki J, Spudich JL, Luecke H. Anabaena sensory rhodopsin: A photochromic color sensor at 2.0 Å. *Science (1979)*. 2004;306(5700):1390-1393. doi:10.1126/SCIENCE.1103943/SUPPL_FILE/VOGELEY.SOM.PDF
68. Callon M, Malär AA, Pfister S, et al. Biomolecular solid-state NMR spectroscopy at 1200 MHz: the gain in resolution. *J Biomol NMR*. 2021;75(6-7):255-272. doi:10.1007/S10858-021-00373-X/FIGURES/12
69. Overhauser AW. Polarization of Nuclei in Metals. *Physical Review*. 1953;92(2):411. doi:10.1103/PhysRev.92.411
70. Costello WN, Xiao Y, Frederick KK. DNP-Assisted NMR Investigation of Proteins at Endogenous Levels in Cellular Milieu. *Methods in Enzymology*. Vol 615. Academic Press Inc.; 2019:373-406. doi:10.1016/bs.mie.2018.08.023
71. Bajaj VS, Mak-Jurkauskas ML, Belenky M, Herzfeld J, Griffin RG. DNP enhanced frequency-selective TEDOR experiments in bacteriorhodopsin. *Journal of Magnetic Resonance*. 2010;202(1):9-13. doi:10.1016/j.jmr.2009.09.005
72. Barnes AB, Corzilius B, Mak-Jurkauskas ML, et al. Resolution and polarization distribution in cryogenic DNP/MAS experiments. *Physical Chemistry Chemical Physics*. 2010;12(22):5861. doi:10.1039/c003763j
73. Ni QZ, Daviso E, Can T V, et al. High frequency dynamic nuclear polarization. *Acc Chem Res*. 2013;46(9):1933-1941. doi:10.1021/ar300348n
74. Barnes AB, Mak-Jurkauskas ML, Matsuki Y, et al. Cryogenic sample exchange NMR

probe for magic angle spinning dynamic nuclear polarization. *Journal of Magnetic Resonance*. 2009;198(2):261-270. doi:10.1016/j.jmr.2009.03.003

75. Ni QZ, Markhasin E, Can T V., et al. Peptide and Protein Dynamics and Low-Temperature/DNP Magic Angle Spinning NMR. *J Phys Chem B*. 2017;121(19):4997-5006. doi:10.1021/acs.jpcb.7b02066
76. Klein A, Rovó P, Sakhrani V V., et al. Atomic-resolution chemical characterization of (2x)72-kDa tryptophan synthase via four- and five-dimensional ^{1}H-detected solid-state NMR. *Proceedings of the National Academy of Sciences*. 2022;119(4):e2114690119. doi:10.1073/pnas.2114690119
77. Thureau P, Mollica G, Ziarelli F, Viel S. Solid-State 1H NMR Studies of Homonuclear Dipolar Couplings. *Annu Rep NMR Spectrosc*. 2014;82:217-249. doi:10.1016/B978-0-12-800184-4.00005-9
78. Cala-De Paepe D, Stanek J, Jaudzems K, Tars K, Andreas LB, Pintacuda G. Is protein deuteration beneficial for proton detected solid-state NMR at and above 100 kHz magic-angle spinning? *Solid State Nucl Magn Reson*. 2017;87:126-136. doi:10.1016/J.SSNMR.2017.07.004
79. Zhou DH, Franks WT, Rienstra CM, Graesser DT, Franks WT, Rienstra CM. Sensitivity and resolution in proton solid-state NMR at intermediate deuteration levels: quantitative linewidth characterization and applications to correlation. *Elsevier*. Published online 2006. doi:10.1016/j.jmr.2005.10.008
80. Kajimoto K, Kikukawa T, Nakashima H, et al. Transient Resonance Raman Spectroscopy of a Light-Driven Sodium-Ion-Pump Rhodopsin from *Indibacter alkaliphilus*. *J Phys Chem B*. 2017;121(17):4431-4437. doi:10.1021/acs.jpcb.7b02421
81. Scharf B, Engelhard M. Blue Halorhodopsin from Natronobacterium pharaonis: Wavelength Regulation by Anions. *Biochemistry*. 1994;33(21):6387-6393. doi:10.1021/BI00187A002
82. Pervushin K, Riek R, Wider G, Wüthrich K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci*. 1997;94(23):12366-12371. doi:10.1073/PNAS.94.23.12366
83. Böckmann A, Gardiennet C, Verel R, et al. Characterization of different water pools in solid-state NMR protein samples. *J Biomol NMR*. 2009;45(3):319-327. doi:10.1007/s10858-009-9374-3
84. Thurber KR, Tycko R. Measurement of sample temperatures under magic-angle spinning from the chemical shift and spin-lattice relaxation rate of ^{79}Br in KBr powder. *Journal of Magnetic Resonance*. 2009;196(1):84-87. doi:10.1016/j.jmr.2008.09.019
85. Kay LE, Ikura M, Tschudin R, Bax A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*. 1990;89(3):496-514. doi:10.1016/0022-2364(90)90333-5
86. Markley JL, Bax A, Arata Y, et al. Recommendations for the presentation of NMR structures of proteins and nucleic acids. *Eur J Biochem*. 1998;256(1):1-15. doi:10.1046/j.1432-1327.1998.2560001.x
87. Delaglio F, Grzesiek S, Vuister GeertenW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*.

1995;6(3):277-293. doi:10.1007/BF00197809
88. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*. 2015;31(8):1325-1327. doi:10.1093/bioinformatics/btu830
89. Baldus M, Petkova AT, Herzfeld J, Griffin RG. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol Phys*. 1998;95(6):1197-1207. doi:10.1080/00268979809483251
90. Pauli J, Baldus M, van Rossum B, de Groot H, Oschkinat H. Backbone and Side-Chain ^{13}C and ^{15}N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem*. 2001;2(4):272-281. doi:10.1002/1439-7633(20010401)2:4<272::AID-CBIC272>3.0.CO;2-2
91. Li Y, Berthold DA, Frericks HL, Gennis RB, Rienstra CM. Partial ^{13}C and ^{15}N Chemical-Shift Assignments of the Disulfide-Bond-Forming Enzyme DsbB by 3D Magic-Angle Spinning NMR Spectroscopy. *ChemBioChem*. 2007;8(4):434-442. doi:10.1002/cbic.200600484
92. Franks WT, Kloepper KD, Wylie BJ, Rienstra CM. Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J Biomol NMR*. 2007;39(2):107-131. doi:10.1007/s10858-007-9179-1
93. Schuetz A, Wasmer C, Habenstein B, et al. Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s(1-227). *ChemBioChem*. 2010;11(11):1543-1551. doi:10.1002/cbic.201000124
94. Morcombe CR, Zilm KW. Chemical shift referencing in MAS solid state NMR. *Journal of Magnetic Resonance*. 2003;162(2):479-486. doi:10.1016/S1090-7807(03)00082-X
95. Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR*. 2016;66(2):111-124. doi:10.1007/s10858-016-0060-y
96. Schmidt E, Güntert P. A New Algorithm for Reliable and General NMR Resonance Assignment. *J Am Chem Soc*. 2012;134(30):12817-12829. doi:10.1021/ja305091n
97. Schmidt E, Gath J, Habenstein B, et al. Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. *J Biomol NMR*. 2013;56(3):243-254. doi:10.1007/s10858-013-9742-x
98. Higman VA, Flinders J, Hiller M, et al. Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively ^{13}C-labelled proteins. *J Biomol NMR*. 2009;44(4):245-260. doi:10.1007/s10858-009-9338-7
99. Shi L, Lake EMR, Ahmed MAM, Brown LS, Ladizhansky V. Solid-state NMR study of proteorhodopsin in the lipid environment: Secondary structure and dynamics. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2009;1788(12):2563-2574. doi:10.1016/j.bbamem.2009.09.011
100. Kaur J, Kriebel CN, Eberhardt P, et al. Solid-state NMR analysis of the sodium pump Krokinobacter rhodopsin 2 and its H30A mutant. *J Struct Biol*. 2019;206(1):55-65. doi:10.1016/j.jsb.2018.06.001
101. Kikukawa T, Saha CK, Balashov SP, et al. The Lifetimes of *Pharaonis* Phoborhodopsin Signaling States Depend on the Rates of Proton Transfers—Effects of Hydrostatic Pressure and Stopped Flow Experiments. *Photochem Photobiol*.

2008;84(4):880-888. doi:10.1111/j.1751-1097.2008.00318.x

102. Sauvée C, Rosay M, Casano G, et al. Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angewandte Chemie International Edition*. 2013;52(41):10858-10861. doi:10.1002/anie.201304657

103. Matsuki Y, Nakamura S, Fukui S, Suematsu H, Fujiwara T. Closed-cycle cold helium magic-angle spinning for sensitivity-enhanced multi-dimensional solid-state NMR. *Journal of Magnetic Resonance*. 2015;259:76-81. doi:10.1016/j.jmr.2015.08.003

104. Hu JG, Sun BQ, Petkova AT, Griffin RG, Herzfeld J. The Predischarge Chromophore in Bacteriorhodopsin: A ^{15} N Solid-State NMR Study of the L Photointermediate. *Biochemistry*. 1997;36(31):9316-9322. doi:10.1021/bi970416y

105. Hu J, Griffin RG, Herzfeld J. Synergy in the spectral tuning of retinal pigments: complete accounting of the opsin shift in bacteriorhodopsin. *Proceedings of the National Academy of Sciences*. 1994;91(19):8880-8884. doi:10.1073/pnas.91.19.8880

106. Karasuyama M, Inoue K, Nakamura R, Kandori H, Takeuchi I. Understanding Colour Tuning Rules and Predicting Absorption Wavelengths of Microbial Rhodopsins by Data-Driven Machine-Learning Approach. *Sci Rep*. 2018;8(1):15580. doi:10.1038/s41598-018-33984-w

107. Kato HE, Kamiya M, Sugo S, et al. Atomistic design of microbial opsin-based blue-shifted optogenetics tools. *Nat Commun*. 2015;6(1):7177. doi:10.1038/ncomms8177

Acknowledgements

This research was carried out under the supervision of Professor Toshimichi Fujiwara and Associate Professor Yoh Matsuki. I deeply appreciate the opportunity of conducting this research and the invaluable instruction.

I would like to express my sincere thanks to Professor Hironobu Hojo, Professor Genji Kurisu, and Associate Professor Yohei Miyanoiri, the reviewers of this thesis, for their insightful advice.

The photocycle kinetics measurements were performed in Hokkaido University with the help from Associate Professor Takashi Kikukawa. Solution NMR measurements were carried out with the support of Associate Professor Yohei Miyanoiri. I am truly grateful for their generous help.

I deeply appreciate the advice, discussions, and support from all the members of the Laboratory of Molecular Biophysics, the laboratory to which I belong for my study.

Finally, I thank my parents and Institute for Protein Research for the financial support during my study.

Appendix

Solution NMR experiment parameters

1D ^{1}H solution NMR measurements were carried out on a 400 MHz (B_0 = 9.4 T) Bruker Avance III NMR spectrometer. The ^{1}H carrier was set at 4.8 ppm. The spectral width was 16 ppm (16384 points). Each spectrum was obtained with 8 scans and processed with JEOL Delta by applying zero filling to double the number of data points.

2D ^{1}H-^{15}N TROSY-HSQC solution NMR measurements were carried out on a 600 MHz (B_0 = 14.1 T) Bruker Avance III NMR spectrometer. The ^{1}H carrier was set at 4.8 ppm and ^{15}N carrier was set at 117.1 ppm. The spectral width was 30 ppm for ^{1}H (1024 points) and 35 ppm for ^{15}N (128 points). Each spectrum was obtained with 216 scans and processed with NmrPipe. Eero filling was applied to double the number of data points and a shifted squared-sine-bell (from 0.3π to π) window function was applied in both dimensions.

ssNMR experiment parameters

^{1}H detection ssNMR experiment parameters

^{1}H detection ssNMR measurements with fast MAS were carried out on a 700 MHz (B_0 = 16.4 T) JEOL ECAII spectrometer with a HXY triple resonance probe. The MAS frequency was set to 70 kHz. The typical 90° pulse widths for ^{1}H, ^{13}C, and ^{15}N were 1.25 μs, 1.5 μs and 3.2 μs respectively. The carriers for ^{13}C were set to 55 ppm for Cα. The ^{15}N carrier was set at 118 ppm. The ^{1}H carrier was set at 5 ppm. Magnetization transfer was performed by RAMP-CP[1], and the ^{1}H effective RF field strength was ~ 105 kHz and 30 kHz for ^{13}C. The detailed parameters for CANH spectra are listed below.

Table S1 Parameters of CANH experiment.

H- Cα CP	Contact time 1.5 ms, linear gradient applied on ^{1}H channel.		
Cα-N CP	Contact time 11 ms, linear gradient applied on ^{15}N channel.		
N-H CP	Contact time 850 μs, linear gradient applied on ^{15}N channel.		
Experiment setup			
FID size	1024 (^{1}H)	32 (^{15}N)	25 (^{13}C)
Spectrum width (ppm)	100	40	55
Increment size (μs)	10	282	162
Acquisition time (ms)	10.24	9.02	4.59

^{13}C detection ssNMR experiment parameters

The conventional ^{13}C detection ssNMR measurements were carried out on a 600 MHz (B_0 = 14.1 T) JEOL ECAII spectrometer with a Varian T3-HXY triple resonance probe. The MAS frequency was set to 12.5 kHz or 14 kHz. The typical 90° pulse widths for ^{1}H, ^{13}C, and ^{15}N

were 3.2, 3.2, and 6.25 μs, respectively. The carriers for ^{13}C were set to 55, 178 and 100 ppm for Cα, C′ and non-selective C, respectively. The ^{15}N carrier was set at 118 ppm. The ^{1}H carrier was set at 5 ppm. The $^{13}C/^{15}N$ polarization was prepared by RAMP-CP[1] from ^{1}H, for which the ^{1}H effective RF field strength was ~ 37.5 kHz with the ^{13}C and ^{15}N effective RF field strength ~50 kHz. The RF amplitude gradient was applied to the ^{13}C or ^{15}N channel. The ^{15}N to ^{13}C (or ^{13}C to ^{15}N) magnetization transfer was performed by RAMP-CP[1], for which the ^{15}N effective RF field strength was ~ 17.5 kHz with the ^{13}C effective RF field strength ~5 kHz. The RF amplitude gradient was applied to the ^{13}C channel. The ^{1}H spins were decoupled by ~90 kHz of CW during the CP. The C′ to Cα magnetization transfer was achieved by RFDR[2] with XY-16 super-cycled 50 kHz of rotor-synchronized π-pulse trains under ~105 kHz of CW ^{1}H decoupling. The broadband ^{13}C homonuclear magnetization transfer was achieved by DARR[3]. The ^{1}H decoupling during acquisition was performed by ~70 kHz of SPINAL64[4]. The Hartmann-Hahn matching conditions and decoupling sequence parameters were further experimentally optimized. The detailed parameters for each experiment are listed below.

Table S2 Magnetization transfer parameters of ^{13}C detection experiments.

NCA	
H-N CP	Contact time 1 ms, linear gradient applied on ^{15}N channel.
N-Cα CP	Contact time 7 ms, linear gradient applied on ^{15}N channel.
NCACX	
H-N CP	Contact time 1.5 ms, linear gradient applied on ^{15}N channel.
N-Cα CP	Contact time 6 ms, linear gradient applied on ^{15}N channel.
Cα-CX DARR mixing	Mixing time 100 ms.
NCOCX	
H-N CP	Contact time 1.1 ms, linear gradient applied on ^{15}N channel.
N-C′ CP	Contact time 8 ms, linear gradient applied on ^{15}N channel.
C'-CX DARR mixing	Mixing time 50 ms.
CANCO	
H-Cα CP	Contact time 80 μs, linear gradient applied on ^{13}C channel.
Cα-N CP	Contact time 7 ms, linear gradient applied on ^{15}N channel.
N-C′ CP	Contact time 6 ms, linear gradient applied on ^{15}N channel.
NcoCACX	
H-N CP	Contact time 1.1 ms, linear gradient applied on ^{15}N channel.
N-C′ CP	Contact time 6.5 ms, linear gradient applied on ^{15}N channel.
C'-Cα RFDR mixing	XY-16 RFDR, mixing time 1.38 ms.
Cα-CX DARR mixing	Mixing time 20 ms.
CANcoCA	
H-Cα CP	Contact time 80 μs, linear gradient applied on ^{13}C channel.
Cα-N CP	Contact time 8 ms, linear gradient applied on ^{15}N channel.
N-C′ CP	Contact time 8 ms, linear gradient applied on ^{15}N channel.
C'-Cα RFDR mixing	XY-16 RFDR, mixing time1.28 ms.

Table S3 Experiment setup of ^{13}C detection experiments.

NCA

	F2 (Cα)	F1 (^{15}N)	
FID size	1024	32	
Spectrum width (ppm)	664.29	45	
Increment size (μs)	10	366.3	
Acquisition time (ms)	10.24	11.72	
NCACX			
	F3 (^{13}C)	F2 (^{13}Cα)	F1 (^{15}N)
FID size	1536	35	35
Spectrum width (ppm)	664.29	40	40
Increment size (μs)	10	166.04	412.08
Acquisition time (ms)	15.36	5.8114	14.4228
NCOCX			
	F3 (^{13}C)	F2 (^{13}C′)	F1 (^{15}N)
FID size	1024	35	40
Spectrum width (ppm)	664.29	25	50
Increment size (μs)	10	265.68	329.64
Acquisition time (ms)	10.24	9.2988	13.1856
CANCO			
	F3 (^{13}C′)	F2 (^{15}N)	F1 (^{13}Cα)
FID size	1024	40	35
Spectrum width (ppm)	664.29	50	40
Increment size (μs)	10	329.64	166.02
Acquisition time (ms)	10.24	13.1856	5.8107
NCAcoCX			
	F3 (^{13}C)	F2 (Cα)	F1 (^{15}N)
FID size	1024	30	30
Spectrum width (ppm)	664.29	45	45
Increment size (μs)	10	147.6	366.3
Acquisition time (ms)	10.24	4.428	10.989
CANcoCA			
	F3 (^{13}Cα)	F2 (^{15}N)	F1 (^{13}Cα)
FID size	1024	40	30
Spectrum width (ppm)	664.29	50	35
Increment size (μs)	10	329.64	189.78
Acquisition time (ms)	10.24	13.1856	5.6934

Table S4 FID process parameters for ^{1}H detection and ^{13}C detection ssNMR experiments

2D spectra	
Zero filling	Data point numbers were filled by 0 to 4096 and 256 in direct and indirect dimension.
Apodization function	A shifted squared-sine-bell (from 0.3π to π) applied in both dimensions.
3D spectra	
Zero filling	Data point numbers were filled by 0 to 4096 and 256 in direct and

	indirect dimensions.
Apodization function	A Lorentz-to-Gauss window applied in all dimensions with 50 Hz Lorentz width and 100 Hz Gauss width.

Ultra-low temperature DNP enhanced ^{15}N ssNMR experiment parameters

The DNP enhanced ^{15}N detection ssNMR measurements were carried out on a 700 MHz (B_0=16.4 T) JEOL ECAII spectrometer equipped with the closed-cycle helium gas sample spinning system.[5] The MAS frequency was set to ~7.5 kHz. The 90° pulse width of ^{1}H was 3.7 μs. The ^{1}H carrier was set at 4.7 ppm. The ^{15}N carrier was set at 145 ppm. Electron spins were saturated by a continuous microwave at ~ 460 GHz and transferred to ^{1}H. RAMP-CP[1] was used to transfer ^{1}H polarization to ^{15}N, the ^{1}H effective RF field strength was ~ 46 kHz with the ^{15}N effective RF field strength ~ 30 kHz. A ~70 kHz SPINAL64[4] was used for ^{1}H decoupling during acquisition period. The 90° pulse power, Hartmann-Hahn matching conditions and decoupling power were further experimentally optimized. The data sampling and processing parameters for each experiment are listed below.

Table S5 Experiment setup of DNP enhanced ^{15}N detection experiments.

FID size	Spectrum width (ppm)	Increment size (μs)	Acquisition time (ms)
256	706.33	20	5.12

The data point number was zero filled to 4096. No apodization was applied to the FIDs of DNP enhanced ^{15}N detection experiments. A backward linear prediction was applied to the FIDs to retrieve the first data point for flat baseline.

Signal assignment method and results

I adopted the FLYA algorithm in CYANA software for the automated resonance assignment.[6,7] The CYANA library file was modified to adapt the ^{13}C sparse labeling scheme and additional experiment type for FLYA automatic assignment. First, 12 carbon atom types were added as described below:

Table S6 Additional carbon atom definition.

Additional definition	^{13}C Labeling scheme
C_AL1, C_BY1, C_AR1 or C_VI1	Only ^{13}C labeled by 1,3-$^{13}C_2$-glycerol
C_AL2, C_BY2, C_AR2 or C_VI2	Only ^{13}C labeled by 2-^{13}C-glycerol
C_AL3, C_BY3, C_AR3 or C_VI3	Both ^{13}C labeled

Correspondingly, the amino acid atoms definitions were modified based on the ^{13}C sparse labeling scheme of 1,3-$^{13}C_2$-glycerol or 2-^{13}C-glycerol.[8] For example, the modified glutamine definition of carbon atoms is showed below:

```
RESIDUE   GLN    6    23  3    22
1 C C_BYL 0    0.0000  0.0000  0.0000  0.0000  2    3    0    0    0
```

```
5 CA     C_AL3  0    0.000   2.0925   0.0007   -1.2423  3    6    7    21   0
7 CB     C_AL3  0    0.0000  2.8764   -1.3051  -1.3846  5    8    9    11   0
11 CG    C_AL1  0    0.0000  4.3161   -1.1050  -1.8297  7    12   13   15   0
15 CD    C_BY2  0    0.0000  4.6207   -1.7918  -3.1464  11   16   17   0    0
21 C     C_BY3  0    0.0000  1.1744   0.1901   -2.4453  5    22   23   0    0
```

Finally, the formal magnetization pathways were modified as described below:

For experiments with uniformly ^{13}C labeled, the latest letter for all the carbon atoms in the magnetization pathway was instead by a "*". For example:

```
SPECTRUM CANCO C N CA
  0.980 C:C_BY* N:N_AMI CA:C_AL* C_BY*
```

And 6 new magnetization pathways were added:

```
SPECTRUM NCACX13gly N CA C
  0.980 N:N_AMI CA:C:C_AL1
  0.980 N:N_AMI CA:C:C_AL3
  0.980 N:N_AMI CA:C_AL1 ~1.5 C:C_AL3
  0.980 N:N_AMI CA:C_AL3 ~1.5 C:C_AL1
  0.980 N:N_AMI CA:C_AL3 ~1.5 C:C_AL3
  0.850 N:N_AMI CA:C_AL1 ~2.7 C:C_AL1
  0.850 N:N_AMI CA:C_AL1 ~2.7 C:C_AL3
  0.850 N:N_AMI CA:C_AL3 ~2.7 C:C_AL1
  0.850 N:N_AMI CA:C_AL3 ~2.7 C:C_AL3
  0.800 N:N_AMI CA:C_AL1 ~4.0 C:C_*1
  0.800 N:N_AMI CA:C_AL1 ~4.0 C:C_*3
  0.800 N:N_AMI CA:C_AL3 ~4.0 C:C_*1
  0.800 N:N_AMI CA:C_AL3 ~4.0 C:C_*3

SPECTRUM NCACX2gly N CA C
  0.980 N:N_AMI CA:C:C_AL2
  0.980 N:N_AMI CA:C:C_AL3
  0.980 N:N_AMI CA:C_AL2 ~1.5 C:C_AL2
  0.980 N:N_AMI CA:C_AL3 ~1.5 C:C_AL2
  0.850 N:N_AMI CA:C_AL2 ~2.7 C:C_AL2
  0.850 N:N_AMI CA:C_AL2 ~2.7 C:C_AL3
  0.850 N:N_AMI CA:C_AL3 ~2.7 C:C_AL2
  0.850 N:N_AMI CA:C_AL3 ~2.7 C:C_AL3
  0.800 N:N_AMI CA:C_AL2 ~4.0 C:C_*2
  0.800 N:N_AMI CA:C_AL2 ~4.0 C:C_*3
  0.800 N:N_AMI CA:C_AL3 ~4.0 C:C_*2
  0.800 N:N_AMI CA:C_AL3 ~4.0 C:C_*3
```

SPECTRUM NCOCX13gly N CO C
0.980 N:N_AMI CO:C:C_AL1
0.980 N:N_AMI CO:C:C_AL3
0.980 N:N_AMI CO:C_BY3 ~1.5 C:C_AL3
0.850 N:N_AMI CO:C_BY1 ~2.7 C:C_AL1
0.850 N:N_AMI CO:C_BY1 ~2.7 C:C_AL3
0.850 N:N_AMI CO:C_BY3 ~2.7 C:C_AL1
0.850 N:N_AMI CO:C_BY3 ~2.7 C:C_AL3
0.800 N:N_AMI CO:C_BY1 ~4.0 C:C_*1
0.800 N:N_AMI CO:C_BY3 ~4.0 C:C_*1
0.800 N:N_AMI CO:C_BY1 ~4.0 C:C_*3
0.800 N:N_AMI CO:C_BY3 ~4.0 C:C_*3

SPECTRUM NCOCX2gly N CO C
0.980 N:N_AMI CO:C:C_AL2
0.980 N:N_AMI CO:C:C_AL3
0.850 N:N_AMI CO:C_BY2 ~2.7 C:C_AL2
0.850 N:N_AMI CO:C_BY2 ~2.7 C:C_AL3
0.850 N:N_AMI CO:C_BY3 ~2.7 C:C_AL2
0.800 N:N_AMI CO:C_BY2 ~4.0 C:C_*2
0.800 N:N_AMI CO:C_BY2 ~4.0 C:C_*3
0.800 N:N_AMI CO:C_BY3 ~4.0 C:C_*2
0.800 N:N_AMI CO:C_BY3 ~4.0 C:C_*3

SPECTRUM CANCO13gly C N CA
0.980 C:C_BY1 N:N_AMI CA:C_AL1 C_BY*
0.980 C:C_BY1 N:N_AMI CA:C_AL3 C_BY*
0.980 C:C_BY3 N:N_AMI CA:C_AL1 C_BY*
0.980 C:C_BY3 N:N_AMI CA:C_AL3 C_BY*

SPECTRUM CANCO2gly C N CA
0.980 C:C_BY2 N:N_AMI CA:C_AL2 C_BY*
0.980 C:C_BY2 N:N_AMI CA:C_AL3 C_BY*
0.980 C:C_BY3 N:N_AMI CA:C_AL2 C_BY*
0.980 C:C_BY3 N:N_AMI CA:C_AL3 C_BY*

Where 13gly and 2gly stands for the spectrum obtained from ^{13}C sparsely labeled by 1,3-$^{13}C_2$-glycerol and 2-^{13}C-glycerol, respectively.

The modified library file was used for FLYA automatic assignment. The tolerances for chemical shift matching were set to 0.5 ppm for^{13}C and ^{15}N. The automatic assignment result was manually verified through backbone walking strategy. The assigned chemical shifts are listed below.

Table S7 The assigned chemical shifts of *Np*HR in Cl^- bound state.

Residue	Chemical shift (ppm)				
	N	Cα	C′	Cβ	Cγ
Ala45		55.618	179.933	17.179	-
Gly46	103.162	47.691	177.563	-	-
Leu47	121.302	57.416			
Ala86		54.724	179.106	17.349	-
Ser87	105.508	60.575	174.503	64.995	-
Gly88	109.560	44.570	171.825	-	-
Leu138		57.578	178.037	43.566	26.938
Gly139	109.018	45.091	176.250	-	-
Leu140	121.399	57.297	181.774	41.594	
Leu141	124.016	58.750	178.552	41.615	26.947
Ala142	118.992	51.820	173.655	18.015	-
Gly143	101.836	45.990	176.072	-	-
Ser144	111.028	58.236	175.032	63.202	-
Asn145	124.881	51.778	177.395	38.112	
Ala146	121.072	55.187	179.101	18.806	-
Thr147	115.469	57.553	179.658	66.244	18.342
Lys148	124.372	58.149	178.907		
Thr151	117.044	67.189	173.932		21.619
Ala152	120.521	55.629	178.583	18.712	
Ile153	113.696	67.006	178.039	38.114	31.721 (Cγ1) 18.463 (Cγ2)
Thr154	110.369	66.574	176.772		
Cys160		64.577	176.772		-
Val161	117.434	66.274	177.329	31.552	22.560
Thr162	110.287	65.399	177.052	68.863	22.618
Gly163	109.747	46.959	174.228	-	-
Leu164	120.913	57.551	177.565		
Ala165	120.365	54.985			-
Ala204		54.289	178.861	18.233	-
Ala205	118.249	52.514	177.949	19.212	-
Gly206	103.971	45.469	175.339	-	-
Thr207	107.781	59.565	174.829	69.130	22.017
Ala208	123.529	56.432	177.973	18.984	-

Table. S8 The assigned chemical shifts of *Np*HR in Cl^- free state.

Residue	Chemical shift (ppm)				
	N	Cα	C′	Cβ	Cγ
Leu47	121.170	57.694			
Ser87	105.371	60.349			-
Gly88	109.459	44.826	175.631	-	-
Leu138			178.029		
Gly139	108.801	44.948	176.243	-	-

Leu140	121.500	57.151	181.699		
Leu141	123.892	58.296	178.528	41.499	
Ala142	119.058	51.775	173.800	18.010	-
Gly143	101.620	45.867	176.008	-	-
Ser144	110.840	58.306	175.306	63.363	-
Asn145	124.780	51.811	177.485		
Ala146	120.939	55.022	179.485		
Thr147	115.766	57.815	179.516		
Lys148	124.587	58.030			
Thr151	116.915	67.931	174.200	68.367	21.432
Ala152	120.949	55.704	178.960	18.744	-
Ile153	114.043	67.104	178.124		
Thr154	110.832	66.741			
Val161	117.985	66.066	177.088		
Thr162	111.000	65.410	177.075	68.942	22.882
Gly163	109.759	47.144	174.168	-	-
Leu164	120.747	57.659		42.059	
Ala165	120.939	55.022			
Ala204		54.465	178.640		
Ala205	118.292	52.446	177.986	19.245	-
Gly206	104.023	45.470	175.375	-	-
Thr207	107.862	59.476	174.838		22.031
Ala208	123.667	56.278			-

Refences

1. Metz G, Wu X, Smith SO. Ramped-Amplitude Cross Polarization in Magic-Angle-Spinning NMR. *J Magn Reson A*. 1994;110(2):219-227. doi:10.1006/JMRA.1994.1208
2. Bennett AE, Ok JH, Griffin RG, Vega S. Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J Chem Phys*. 1992;96(11):8624-8627. doi:10.1063/1.462267
3. Takegoshi K, Nakamura S, Terao T. 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett*. 2001;344(5-6):631-637. doi:10.1016/S0009-2614(01)00791-6
4. Fung BM, Khitrin AK, Ermolaev K. An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids. *Journal of Magnetic Resonance*. 2000;142(1):97-101. doi:10.1006/JMRE.1999.1896
5. Matsuki Y, Nakamura S, Fukui S, Suematsu H, Fujiwara T. Closed-cycle cold helium magic-angle spinning for sensitivity-enhanced multi-dimensional solid-state NMR. *Journal of Magnetic Resonance*. 2015;259:76-81. doi:10.1016/j.jmr.2015.08.003
6. Schmidt E, Gath J, Habenstein B, et al. Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. *J Biomol NMR*. 2013;56(3):243-254. doi:10.1007/s10858-013-9742-x

7. Schmidt E, Güntert P. A New Algorithm for Reliable and General NMR Resonance Assignment. *J Am Chem Soc*. 2012;134(30):12817-12829. doi:10.1021/ja305091n
8. Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*. 2002;420(6911):99-102. doi:10.1038/nature01070