

A Thesis for the Degree of Ph.D. in Engineering

**Interactions between water molecules and
biological materials: a molecular dynamics study**

February 2018

Graduate School of Science and Technology
Keio University

Katsufumi Tomobe

©2018
Katsufumi Tomobe
All Rights Reserved

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Professor Kenji Yasuoka, for his academic support and encouragement. I would like to thank him for enthusiastic comments and advices to this research, and I could receive any supports by him whenever I want. Moreover, he also provided his guidance in the Program for Leading Graduate Schools “Global Environmental System Leaders (GESL) Program”. With Professor Yasuoka’s education philosophy, I have done this research and would like to contribute to the society in future, thanks to Professor Yasuoka.

I would also like to thank Professor Masato Yasui in the department of pharmacology for valuable discussions and encouragement on my research. He taught me passions for researches of water molecules and provided many opportunities to learn various topics.

I also express my deepest gratitude to Dr. Eiji Yamamoto. We have collaborated in many research topics, and he spent a lot of his time for teaching me. I learned the professional spirit for academic activities.

I would also like to thank Associate Professor Takuma Akimoto in Tokyo University of Science. In the first study, he kindly help me to publish the paper, and he also supported my activities in the GESL program.

Furthermore, I extend my gratitude to all the past and present laboratory members. I especially would like to thank Professor Yohei Sarto in the department of system design engineering, Professor Kholmurzo Kholmurodov in the Joint Institute for Nuclear Research, Associate Professor Dusan Kojic, Mr. Takashi Iijima, Mr. Shinjiro Nakamura, Mr. Katsuhiko Endo, and Mr. Satoshi Yamaguchi for research collaborations. Moreover, I spent long time in the laboratory with Mr. Kentaro Nomura, Mr. Takuma Nozawa, and Mr. Daisuke Yuhara. They made my laboratory life enjoyable. I also express my gratitude for Dr. Donguk Suh, Dr. Taizo Shibuya, Dr. Masaki Hiratsuka. When I was down and discouraged, they brightened

my day. We chatted many times.

I wish to thank my thesis examining committee members: Professor Koji Fukagata, Assistant Professor Ayori Mitsutake, and Assistant Professor Keita Ando, for their thorough comments.

I acknowledge the support by MEXT Grant-in-Aid for the Program for Leading Graduate Schools “Global Environmental System Leaders (GESL) Program”, Keio University Program for the Advancement of Next Generation Research Projects.

And finally, I would like to express my great thanks to my parents, Fuki and Kenichi, for their encouragement and unending support.

Abstract

Water molecules play important roles for behavior, stability, and functions of biological materials. In this study, all-atom MD simulations were performed to investigate the interactions between water molecules and biological materials, e.g., proteins and monosaccharide.

First, we showed typical behavior of water molecules around two proteins: prion protein and rhodopsin. In the case of prion protein, we found that the pathogenic point mutation T188R makes the buried water molecules unstable and fluctuated. This fluctuation causes the instability of subdomains in the prion protein. In the case of rhodopsin, we found the putative solvent pore on cytoplasmic side in the intermediate state of rhodopsin. Moreover, we provided evidence that this pore is important for the function of rhodopsin.

Then, hydration structure and water dynamics around monosaccharide in the aqueous solutions were investigated using both the classical and *ab initio* MD simulations. The results of classical MD simulations revealed the effect of isomers, monosaccharide clustering, and conversion of conformers on its hydration shell. Using the *ab initio* MD simulations, we reported the molecular mechanism of blue shift, and this blue shift was exhibited by the water molecules in the first hydration shell of monosaccharide.

These results provide new insights on the interactions between water molecules and various biological materials.

Contents

| | | |
|----------------|---|-----|
| Acknowledgment | | i |
| Abstract | | iii |
| Chapter 1 | Introduction | 1 |
| 1.1 | Water in biology | 1 |
| 1.2 | Proteins | 2 |
| 1.2.1 | Prion protein | 3 |
| 1.2.2 | Rhodopsin | 4 |
| 1.3 | Monosaccharide | 6 |
| 1.4 | Thesis objective and outline | 8 |
| Chapter 2 | Methodology of molecular dynamics simulation | 9 |
| 2.1 | Numerical integration method for Newton's equations of motion | 9 |
| 2.2 | Control of temperature and pressure | 12 |
| 2.2.1 | Nosé-Hoover method | 12 |
| 2.2.2 | Langevin Piston method | 13 |
| 2.3 | Potential energy between molecules | 14 |
| 2.3.1 | Intramolecular potential energy | 14 |
| 2.3.2 | Intermolecular potential energy | 15 |
| 2.4 | Periodic boundary condition and Ewald method | 16 |
| 2.5 | Ab initio molecular dynamics simulation | 17 |
| 2.5.1 | Born-Oppenheimer approximation | 18 |
| 2.5.2 | Hohenberg-Kohn theorems | 19 |
| 2.5.3 | Kohn-Sham equation | 19 |
| 2.5.4 | Car-Parrinello MD simulation | 20 |
| Chapter 3 | Interactions between water molecules and prion protein | 23 |

| | | |
|-----------|---|----|
| 3.1 | Materials and methods | 23 |
| 3.2 | Results | 24 |
| 3.2.1 | Equilibrium of the systems | 24 |
| 3.2.2 | Fluctuations of amino acids in the protein | 25 |
| 3.2.3 | Time series of secondary structure | 25 |
| 3.2.4 | Buried water molecules | 25 |
| 3.2.5 | Reorientational time correlation function | 30 |
| 3.2.6 | Relationship between the fluctuations and stability of the buried hydration sites | 31 |
| 3.2.7 | Molecular configuration around the mutation | 32 |
| 3.2.8 | Summary | 33 |
| Chapter 4 | Interactions between water molecules and rhodopsin | 35 |
| 4.1 | Materials and methods | 35 |
| 4.2 | Results | 36 |
| 4.2.1 | Water molecule accessibility | 36 |
| 4.2.2 | Configuration of the solvent pore in the Meta II state . | 38 |
| 4.2.3 | Water displacements within the solvent pore in the Meta II state | 40 |
| 4.2.4 | All water pathways between the bulk and the retinal- binding pocket in the Meta II state | 43 |
| 4.2.5 | Summary | 47 |
| Chapter 5 | Interactions between water molecules and monosaccharide | 49 |
| 5.1 | Materials and methods | 49 |
| 5.2 | Results | 50 |
| 5.2.1 | Hydration structure of monosaccharides | 50 |
| 5.2.2 | Water arrangement in the hydration shell | 53 |
| 5.2.3 | Clustering of monosaccharides | 57 |
| 5.2.4 | Boat and chair conformers | 58 |
| 5.2.5 | Summary | 60 |
| Chapter 6 | Blue shift of water molecules in the first hydration shell around monosaccharide | 63 |
| 6.1 | Blue shift and dangling water molecules | 63 |

| | | |
|-----------|---|----|
| 6.2 | Materials and methods | 64 |
| 6.2.1 | Raman Data Collection and Multivariate Curve Resolu- tion Methods | 64 |
| 6.2.2 | <i>Ab Initio</i> Molecular Dynamics Simulations | 65 |
| 6.3 | Methods for analysis | 67 |
| 6.3.1 | Frequency Calculations: Fourier and Wavelet Transforms | 67 |
| 6.3.2 | Tetrahedral Order Parameter | 67 |
| 6.4 | Results | 68 |
| 6.4.1 | Hydration-shell spectra of monosaccharides | 68 |
| 6.4.2 | The relationship between water arrangement and the observed blue shift | 71 |
| 6.4.3 | Distribution of the dangling water molecules | 73 |
| 6.4.4 | Summary | 76 |
| Chapter 7 | Conclusions | 79 |
| | Abbreviations and symbols | 83 |
| | Appendix | 85 |
| | References | 87 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Schematic view of the hydration | 2 |
| 1.2 | Crystal structure of three states of rhodopsin and stability of our systems | 5 |
| 2.1 | Schematic flow of <i>ab initio</i> MD simulations | 18 |
| 3.1 | Stability of prion proteins | 24 |
| 3.2 | Secondary structure of WT-PrPc and T188R | 26 |
| 3.3 | The mean residence times were mapped on atoms of prion protein for each calculation | 28 |
| 3.4 | Position of the water molecules in WT-PrPc 2 | 29 |
| 3.5 | The reorientational correlation function of the buried water molecules in WT-PrPc 2 | 31 |
| 3.6 | Normalized CV and normalized residence time of each site. | 32 |
| 3.7 | Distance between the closest atoms of R188 and G200 | 33 |
| 4.1 | System for MD simulations of rhodopsin | 37 |
| 4.2 | Accessibility of water molecules | 38 |
| 4.3 | Accessibility of water molecules | 39 |
| 4.4 | Hourglass-shaped entrance of the solvent pore at the cytoplasmic site | 40 |
| 4.5 | Configuration of the solvent pore on the cytoplasmic side with ribbon representation of the Meta II structure | 41 |
| 4.6 | Characteristic regions within the solvent pore | 42 |
| 4.7 | Water displacements in <i>z</i> -coordinate within the solvent pore when one water molecule passes through the solvent pore | 43 |
| 4.8 | Trajectories of all water molecules which passed through the solvent pore in <i>z</i> coordinate | 44 |
| 4.9 | Times series of the number of internal water molecules in each state | 45 |

| | | |
|------|---|----|
| 4.10 | Other water pathways between the bulk and the retinal-binding pocket | 46 |
| 4.11 | Stability of the simulation systems | 47 |
| 5.1 | Monosaccharide conformation and simulation systems | 50 |
| 5.2 | RDFs of water oxygen-oxygen | 51 |
| 5.3 | SDFs of water molecules around each monosaccharide | 51 |
| 5.4 | Typical characteristics of the hydration shell around a monosaccharide | 53 |
| 5.5 | SDFs of the oxygen atoms of water molecules around each monosac- charide at 298 and 350 K | 53 |
| 5.6 | SDFs of water molecules around each monosaccharide | 54 |
| 5.7 | Water structure in the hydration shells of monosaccharides | 55 |
| 5.8 | Clustering of monosaccharides | 56 |
| 5.9 | Conversion of monosaccharide conformers | 57 |
| 5.10 | Definition of the boat conformer | 59 |
| 5.11 | Location of the flagpole in the boat conformer | 61 |
| 6.1 | I-shaped channel used for Raman experiments | 64 |
| 6.2 | Raman spectra of D-glucose solutions at concentrations between 0 and 0.5 M | 68 |
| 6.3 | Spectra of water molecules in D-glucose solution | 69 |
| 6.4 | OH vibrational spectra for all water molecules in the system ob- tained by CPMD simulations | 70 |
| 6.5 | OH vibrational spectra of water molecules within 3.5 Å of all isomers | 71 |
| 6.6 | Relationship between water arrangement and vibration in α -D- glucose solution | 72 |
| 6.7 | Density of water molecule oxygen atoms around all the monosac- charides | 74 |
| 6.8 | Formation process of the observed dangling water molecules around the monosaccharide | 75 |
| 6.9 | Average number of H-bonds between water molecules and RDF as a function of the distance between oxygen atoms of monosaccharide and water molecules. | 75 |
| 6.10 | Spectra and water arrangement of water molecules in the NaBr solution | 77 |

List of Tables

| | | |
|-----|---|----|
| 3.1 | Proportion of secondary structures in WT-PrPc and T188R | 27 |
| 3.2 | Mean residence times, total residence times of (A) WT-PrPc and (B) T188R | 27 |
| 3.3 | Mean distance and coefficient of variance between subdomains S1- H1-S2 and H2, 3 | 31 |
| 5.1 | Average lifetime, content, and free energy change of the conformers. T , C , L are temperature, concentration and lifetime, respectively . | 58 |

Chapter 1

Introduction

1.1 Water in biology

The human body contains 70% water, and most biological phenomena occur with water molecules. Thus, interactions between water molecules and biological materials play crucial roles in biological process, and there are various contributions of water molecules to biological materials [1]. A key interaction of water molecules is hydrogen bond (H-bond). Water molecules form stable H-bonds with other water molecules as H-bond network in the bulk, and they also form H-bonds with biological materials. In the biological solutions, these H-bonded water molecules are distributed around the biological materials, called first hydration shell. These water molecules contribute to the stability of the biological materials (see Fig. 1.1A) [2]. In the hydration shell, there are two types of H-bonds interacting with biological material: a donor and acceptor. Donor H-bond is defined when the water molecule provides a hydrogen atom; on the other hand, acceptor H-bond is defined when the biological material provides a hydrogen atom (see Fig. 1.1B).

In such a non-bulk environment, however, the interactions are perturbed by other materials, and the dynamics of the water molecules becomes anomalous. For example, in the hydration shell of biological and chemical materials, the rotational dynamics of water molecules becomes slower than that in the bulk, and the translational dynamics is similarly retarded [3–9]. Retardation of the rotational and translational dynamics has also been observed on the lipid-membrane surfaces as well as other biological materials [10–12]. Although these phenomena are crucial important for many reactions and dynamics in biological and chemical fields, molecular mechanism and process of many phenomena are still unclear.

In this study, we investigated the interactions between water molecules and bi-

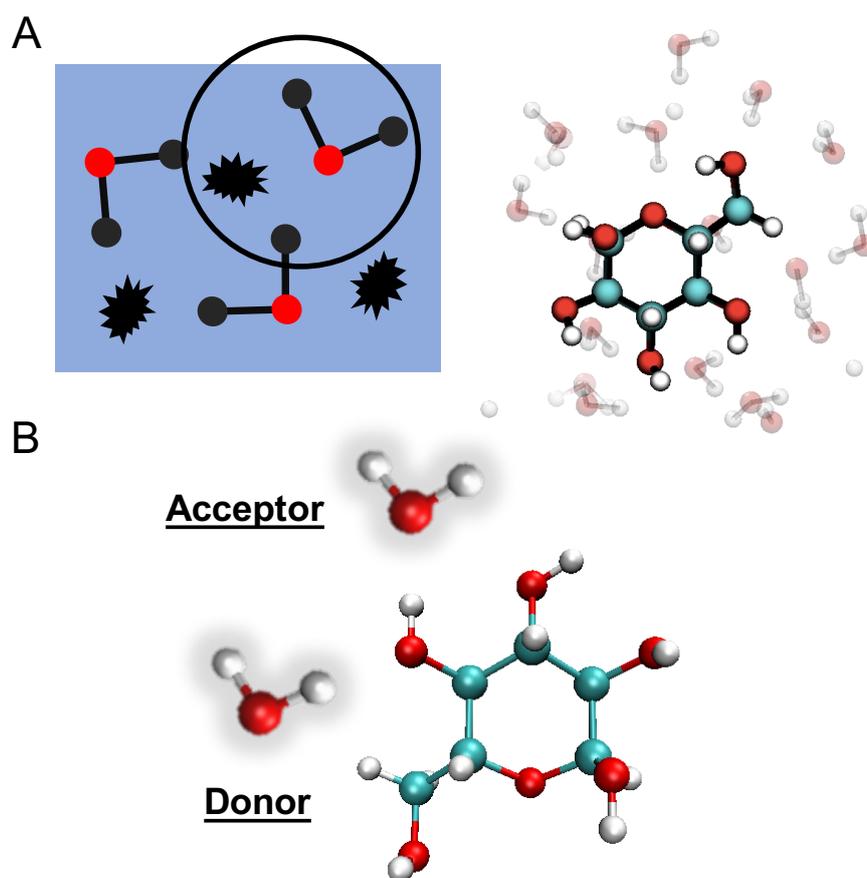


Fig. 1.1 **Schematic view of the hydration.** (A) In an aqueous solution, solutes are distributed in the solution. The first hydration shell is located around the solutes. (B) Two types of H-bonds: a donor and acceptor.

ological materials through two representative biological materials: proteins and monosaccharide.

1.2 Proteins

Protein is a chain of 20 kinds of amino acids. Water molecules play an important role for biochemical properties of proteins [13], e.g. protein conformation [14], aggregation of proteins, docking of proteins [15,16], interaction with other biological materials [17], and the stability of biological materials [18,19]. In this study, we investigated the interactions between water molecules and proteins through two important proteins: prion protein and rhodopsin.

1.2.1 Prion protein

The first target protein is a prion protein which is a causal protein for prion diseases. Prion diseases, including Creutzfeldt-Jakob disease and mad cow disease, are one of the neurodegenerative diseases caused by aggregation of misfolding proteins [20]. Dynamics of the prion protein has attracted researchers for elucidating prion diseases. The prion diseases are composed of mainly three processes: (i) conformational change of secondary structure, (ii) aggregation of transformed prion proteins, and (iii) accumulation of protein amyloids on nerve cells. Conformational change and aggregation of prion proteins are quite important because they may be related to the mechanism of Alzheimer's disease and Parkinson disease [21, 22]. However, these processes remain an unresolved biological phenomena.

The prion protein has two small β -sheets ($\beta 1$, $\beta 2$) located near C-terminal and three α -helices (H1, H2, H3) with a disulfide bond between C179 and C214 [23]. The secondary structural change from the α -helix-rich cellular prion protein (PrPc) to a protease-K resistant β -sheet-rich scrapie prion protein (PrPsc) is also a key process of prion diseases [24], and is caused by low pH [25–28], high temperature [29], or pathogenic mutations [30–34].

Water molecules in hydration shell are usually located around a protein; however, several water molecules are buried in the protein, called buried water molecules [35–37]. Buried water molecules have hydrogen bonds with amino acids, and contribute to the formation of secondary structure and the stability of protein [38]. 2 Å crystal structure has showed that prion protein also has buried water molecules [39]. Using all-atom MD simulations, De Simone *et al.* clarified the location of the buried water molecules and showed that these buried water molecules are quite important for the stability of secondary structure [40]. They have also investigated an influence on the buried hydration sites by a pathogenic point mutation (Q217R) located nearby the buried hydration sites. The results showed that the pathogenic point mutation disrupts the buried hydration site. However, some recognized pathogenic mutations are far from the sites. Therefore, it is important to investigate an influence on the buried hydration sites because there are no studies so far clarifying the effect.

In this study, we perform MD simulations of both the human prion protein wild type (WT)-PrPc and the pathogenic mutant prion protein (T188R) to clarify an influence of the pathogenic point mutation T188R on the buried hydration sites.

The T188R mutation, which is not adjacent to the buried hydration sites, has been investigated by both experiments [41, 42] and MD simulation [43]. Our results provide a first evidence that the pathogenic point mutation affects the stability of buried water molecules far from the mutation point.

1.2.2 Rhodopsin

The second target protein is rhodopsin. Rhodopsin is a GPCR which is transmembrane (TM) proteins that transmit a signal from the extracellular to the cytoplasmic side of cell membranes via G-proteins. GPCRs have become the targets of 30 % of marketed drugs and are still attractive materials for pharmaceutical and biophysical studies [44–46]. Rhodopsin is composed of three parts: seven TM helices, a small helix parallel to the membrane that works as an anchor for the interaction with the G-protein, and a light-sensitive chromophore (11-cis retinal) [47]. When the dark-adapted rhodopsin absorbs light, 11-cis-retinal is isomerized to 11-trans-retinal within 200 fs, which is one of the fastest chemical reactions in the human body [48–50]. This reaction triggers cascade reactions in the rhodopsin. Through several intermediate states including the Meta II state [47, 51, 52], a Schiff base linkage between the 11-trans-retinal and K296 is disrupted by hydrolysis (see Fig. 1.2A-D). Finally, the rhodopsin decays into the active state of the rhodopsin, called Opsin, and the 11-trans-retinal [53].

Rhodopsin has two pores in the cell membrane, called ligand pores. One is located between TM5 and TM6, and the other is located between TM1 and TM7, and two ligand pores function in release and uptake of the retinal [54–56]. In addition to the two pores, it is thought that the Meta II state and Opsin have one narrow pore on the cytoplasmic side [57, 58]. This narrow pore, called a ‘solvent pore’, was suggested by Angel *et al.* using radiolytic hydroxyl labeling and liquid chromatography coupled to mass spectroscopy [59]. The solvent pore is thought to provide water molecules from the cytoplasmic bulk to the retinal-binding pocket. Because it is narrow, it only allows permeations of water molecules and some small compounds [60, 61]. The solvent pore is regarded as functionally important pathway for the cascade reaction because hydrolysis of the Schiff base linkage requires a water molecule [56]. Moreover, the internal water molecules around the retinal stabilize local electric fields by hydrating near polar and charged amino acids [62]. Other GPCRs also have a continuous solvent pore, and its function depends on the

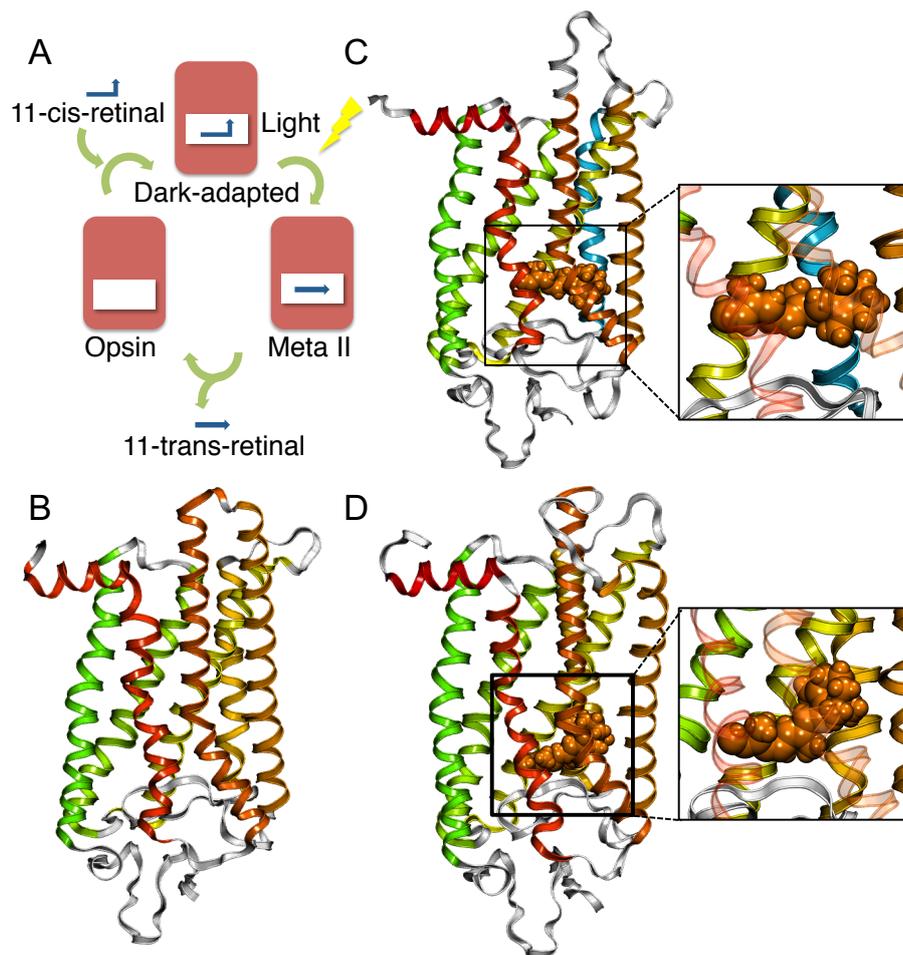


Fig. 1.2 **Crystal structure of three states of rhodopsin and stability of our systems.** (a) Activation cycle of rhodopsin. Through several intermediate states including the Meta II state, the rhodopsin is decayed into the Opsin and the 11-trans-retinal. Crystal structures of the rhodopsin in (b) the Opsin, (c) the Meta II state, and (d) the dark-adapted rhodopsin are shown. The retinal is shown in orange VDW format.

state of the GPCRs [63]. Additionally, water molecules passing through the solvent pore might affect important internal water molecules within the rhodopsin. In general, internal water molecules play important roles in biological materials, e.g., transferring oxygens and protons [64,65], regulating other material permeations [66, 67], and stabilizing biological materials, such as proteins and DNA [1]. In case of the rhodopsin, internal water molecules contribute to the thermodynamic and conformational stabilities of the protein [68]. Furthermore, these water molecules contribute to protein plasticity, and they mediate the activation of protein in other GPCRs [14]. Fourier-transform infrared spectroscopy has shown that configuration

of the internal water molecules changes during photoactivation process [62], which indicates that water molecules exchange between the bulk and the internal hydration sites. Although the importance of internal water molecules is well known, the precise pathways of water molecules between the bulk and protein inside remain poorly understood. Additionally, it remains difficult to investigate the dynamics of internal water molecules within the solvent pore directly by spectroscopy.

Here, using all-atom MD simulations, we reveal details of water flux between the bulk and rhodopsin inside. Due to the different secondary and hydration structures for each intermediate in rhodopsin, we separately prepare and investigate three states of rhodopsin: the dark-adapted rhodopsin, the intermediate Meta II state, and the Opsin. Our simulations and computational analyses show that a putative solvent pore exists on the cytoplasmic side in the Meta II state and the Opsin while the solvent pore does not exist in the dark-adapted rhodopsin. We clearly provide a detailed configuration of the solvent pore, showing that there are two narrow regions that distinguish the bulk and internal hydration sites.

1.3 Monosaccharide

Saccharides are abundant biological materials and used for basic energy storage, including starch in plants, glycogen in animal cells, and dextran in bacteria. In addition to energy storage, saccharides play various important roles in living organisms; e.g., cellulose in cell walls, protein recognition, blood clotting [69], mediation of cell–cell interactions [70], and bioprotection in tardigrades [71–73]. Because monosaccharides are the basic unit of all glycans, knowledge about monosaccharides provides fundamental information about all glycans.

Most biological saccharides are hydrated to form pyranose. Pyranose has α and β anomers, the ratio of which depends on the monosaccharide isomer [74, 75]. In addition to isomers, all monosaccharides have conformers: boat and chair. The most stable conformer is the chair. A previous study revealed the difference of free energy change between the two conformers in α/β -D-glucose solution [76]. Over the last two decades, the molecular behavior of water molecules in saccharide solutions has been investigated experimentally. In the hydration shell, water dynamics is retarded by the monosaccharide [77], and the hydration structure depends on the saccharide [78]. The reorientational dynamics of water molecules is also markedly slowed down in glucose solution [79]. However, the dynamics of molecules in saccharide solution

remains poorly understood.

Theoretical approaches are powerful to investigate molecular behavior and conformations of monosaccharide and hydrated water molecules [80–84]. Many previous studies focused on glucose solution due to its impact on chemical and biological fields, and they revealed characteristics of glucose and water molecules [85–88]. Kräutler *et al.* performed the MD simulations of β -D-glucose, mannose, and galactose in vacuum and aqueous solutions, and they reported the differences of its structure in detail [81]. The hydration properties around monosaccharide have also been investigated. Previous studies using MD simulations revealed the typical behavior of water molecules around saccharides; e.g., retardation of translational and rotational motion [85] and increased dipole moment [88]. In β -D-glucose, trehalose, and sucrose solutions, these saccharides disrupt the tetrahedral structure of water molecules [89]. However, no simulation study investigating the properties of monosaccharide solutions considering multiple factors, including temperature, concentration, and monosaccharide isomer has been reported. In addition, most previous studies have used one or a few monosaccharides in the solution systems; thus they failed to capture monosaccharide clustering and determine its effect on hydration dynamics.

Here, we prepared two types of simulations to investigate the interactions between water molecules and monosaccharide. Firstly, we performed MD simulations of six monosaccharide isomers with water molecules to investigate the effects of temperature and concentration differences on the structure and hydration of monosaccharide isomers. Using large simulation systems that enable simulations of monosaccharide–monosaccharide interactions, we examine the hydration structure of monosaccharides under more physically realistic conditions than previously reported. Secondly, we investigate the mechanism of the blue shift around various monosaccharides using *ab initio* and classical molecular dynamics simulations, and Raman spectroscopy. Notably, our results show that the blue shift arises from weak H-bonded water molecules located in the first hydration shell. The vibrational state of the OH bond, which is oriented towards the bulk solvent, contributes significantly to the observed blue shift. Importantly, these observations provide an enhanced depiction of the current experimental blue shift at aqueous interfaces.

1.4 Thesis objective and outline

In this thesis, we investigate the interactions between water molecules and biological materials. In chapter 2, the computational methodology of MD simulations is introduced. In chapter 3 and 4, interactions between water molecules and proteins are revealed. In chapter 5, interactions between water molecules and monosaccharide are investigated. In chapter 6, relationship between the interactions and spectral changes is discussed.

Chapter 2

Methodology of molecular dynamics simulation

2.1 Numerical integration method for Newton's equations of motion

In a many-body system which is consist of N particles, the Newton's equations of motion is written as

$$m_i \ddot{\mathbf{x}}_i(t) = \mathbf{F}_i(t), \quad (2.1)$$

where $\mathbf{F}_i(t)$, m_i , and $\mathbf{x}_i(t)$ are a force, mass, and coordinate of the i th particle, respectively. In molecular dynamics simulations, the Newton's equations of motion are solved using a numerical time integration algorithm. In this section, the most common time integration algorithms, Verlet and Velocity Verlet algorithms are described. Coordinates $\mathbf{r}_i(t + \Delta t)$ and $\mathbf{r}_i(t - \Delta t)$ are expanded in Taylor series,

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \dot{\mathbf{r}}_i(t) + \frac{\Delta t^2}{2!} \ddot{\mathbf{r}}_i(t) + \frac{\Delta t^3}{3!} \dddot{\mathbf{r}}_i(t) + O(\Delta t^4), \quad (2.2)$$

$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \dot{\mathbf{r}}_i(t) + \frac{\Delta t^2}{2!} \ddot{\mathbf{r}}_i(t) - \frac{\Delta t^3}{3!} \dddot{\mathbf{r}}_i(t) + O(\Delta t^4). \quad (2.3)$$

Adding Eqs. (2.2) and (2.3), we have

$$\mathbf{r}_i(t + \Delta t) + \mathbf{r}_i(t - \Delta t) = 2\mathbf{r}_i(t) + \Delta t^2 \ddot{\mathbf{r}}_i(t) + O(\Delta t^4), \quad (2.4)$$

and with Eq. (2.2), the Eq. (2.3) becomes

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i(t) + O(\Delta t^4). \quad (2.5)$$

Then, we substitute Eq. (2.1) in the Eq. (2.2), and we obtained

$$\dot{\mathbf{r}}_i(t) = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t} + O(\Delta t^2). \quad (2.6)$$

Equation (2.5) denotes time evolution of coordinates and Eq. (2.6) denotes time evolution of velocities. Totally, the two equations are called as Verlet method.

In this study, we mainly used Velocity Verlet method; thus, this section shows the derivation and characteristic of the Velocity Verlet method. The Velocity Verlet method is the following equations:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\Delta t}{m_i} \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)}{2}, \quad (2.7)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i} \mathbf{F}_i(t), \quad (2.8)$$

where Δt shows a time step and $\mathbf{v}_i(t)$ is the velocity of the i th particle at time t .

Derivation of the equation is as follows. Using time evolution operator, time evolution of generalized coordinates and generalized conjugate momenta is described

$$q_i(\Delta t) = \exp(i\Phi\Delta t)q_i(0), \quad (2.9)$$

$$p_i(\Delta t) = \exp(i\Phi\Delta t)p_i(0), \quad (2.10)$$

where Φ is the Liouville operator and this operator is defined using the Hamiltonian of the system

$$\Phi = i \left(\frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} \right) \quad (2.11)$$

$$= i \left(-F_i \frac{\partial}{\partial p_i} - \dot{q}_i \frac{\partial}{\partial q_i} \right). \quad (2.12)$$

H is the Hamiltonian of the system, q_i and p_i show generalized coordinate and conjugate momentum, respectively, and kinetic and potential energies contains only the generalized coordinates and momenta as independent variables. Then, the Liouville operator is decomposed into two operators

$$\Phi = \Phi_1 + \Phi_2, \quad (2.13)$$

$$\Phi_1 = -iF_i \frac{\partial}{\partial p_i}, \Phi_2 = -i\dot{q}_i \frac{\partial}{\partial q_i}. \quad (2.14)$$

According to the Lie-Trotter product formula, since the Liouville operator is based on exponential functions, this formula can be decomposed into

$$\exp(i\Phi\Delta t) = \exp(i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_2\Delta t) \exp(i\Phi_1 \frac{\Delta t}{2}) + o(\Delta t^3). \quad (2.15)$$

Therefore, after we ignore Δt^3 error, the time evolution of generalized coordinates and generalized conjugate momenta of Eqs. (2.9) and (2.10) deforms

$$q(\Delta t) = \exp(i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_2 \Delta t) \exp(i\Phi_1 \frac{\Delta t}{2}) q(0), \quad (2.16)$$

$$p(\Delta t) = \exp(i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_2 \Delta t) \exp(i\Phi_1 \frac{\Delta t}{2}) p(0), \quad (2.17)$$

where $q(\Delta t)$ and $p(\Delta t)$ are the generalized coordinates and momenta at time t . Using the Taylor series expansion, we obtain the polynomial representation of time evolution equations. As a result, the equations are described as

$$q_i(\Delta t) = \exp(i\Phi_2 \Delta t) q_i(0) = \exp(\dot{q}_i \frac{\partial}{\partial q_i} \Delta t) q_i(0) = q_i(0) + \frac{\Delta t}{m_i} p_i(0), \quad (2.18)$$

$$p_i(\Delta t) = \exp(i\Phi_1 \frac{\Delta t}{2}) p_i(0) = \exp(F_i \frac{\partial}{\partial p_i} \Delta t) p_i(0) = p_i(0) + \frac{\Delta t}{2m_i} F_i(0), \quad (2.19)$$

and

$$q_i(\Delta t) = q_i(0) + \Delta t v_i(0) + \frac{\Delta t^2}{2m_i} F_i(0), \quad (2.20)$$

$$v_i(\Delta t) = v_i(0) + \frac{\Delta t}{2m_i} \{F_i(0) + F_i(\Delta t)\}. \quad (2.21)$$

Equations (2.20) and (2.21) are called as the Velocity Verlet method. In the process of the derivation, we ignored Δt^3 error and the truncation error is on the order of 3, this method is second-order accurate. Moreover, since this method was derived by using the time evolution operator, the Hamiltonian H is precisely conserved during the time evolution. The methods with this characteristic are called the symplectic integrator. The symplectic characteristic shows the physical validity of this integrator, and this Hamiltonian is useful for validation of the simulations. In addition to the symplectic characteristic, the Velocity Verlet method has reversibility of time (reversible reference system propagator algorithm; RESPA). Using the RESPA method, we reversibly evolves the time with conserved precision. For example, we applied the RESPA method as $-\Delta t$ to the system at Δt , these operators are canceled by each other and the system come back at time 0 as follows:

$$\begin{aligned} & \exp(-i\Phi_1 \frac{\Delta t}{2}) \exp(-i\Phi_2 \Delta t) \exp(-i\Phi_1 \frac{\Delta t}{2}) q(\Delta t) \\ &= \exp(-i\Phi_1 \frac{\Delta t}{2}) \exp(-i\Phi_2 \Delta t) \exp(-i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_2 \Delta t) \exp(i\Phi_1 \frac{\Delta t}{2}) q(0) \\ &= \exp(-i\Phi_1 \frac{\Delta t}{2}) \exp(-i\Phi_2 \Delta t) \exp(i\Phi_2 \Delta t) \exp(i\Phi_1 \frac{\Delta t}{2}) q(0) \\ &= \exp(-i\Phi_1 \frac{\Delta t}{2}) \exp(i\Phi_1 \frac{\Delta t}{2}) q(0) = q(0). \end{aligned}$$

2.2 Control of temperature and pressure

In most real or experimental systems, temperature and pressure are constant; therefore, methods for controlling temperature and pressure are important. In this chapter, we described the algorithm of controlling method for temperature and pressure.

2.2.1 Nosé-Hoover method

The MD method simulates statistical ensembles of the system, and we can obtain macroscopic physical properties and thermodynamic properties as the expected value using the statistical ensembles. The representative ensembles are (1) micro-canonical ensemble (the total number of particles N , the volume V , and the total energy E of the system are constant), (2) canonical ensemble (N , V , and the temperature T are constant), and (3) isothermal-isobaric ensemble (N , the pressure P , and T are constant). In order to simulate these ensembles, previous studies have introduced some methods for controlling the temperature and pressure. In the derivation process, these methods basically considered the NVT or NPT ensembles as the NVE ensemble with a virtual heat or pressure reservoir. Nosé-Hoover method is one of the most famous methods for controlling the temperature of the system. Firstly, Nosé considered a virtual system which is s times faster than the real system,

$$\Delta t = \frac{\Delta t'}{s} \quad (2.22)$$

$$\mathbf{p}_i = \frac{\mathbf{p}'_i}{s} \quad (2.23)$$

Then, the Hamiltonian of the whole system is

$$H = \sum_{i=1}^N \frac{\mathbf{p}'_i{}^2}{2m_i s^2} + U(\mathbf{r}) + \frac{p_s^2}{2Q} + gk_b T_0 \log s, \quad (2.24)$$

where \mathbf{r} is the coordinates of all the particles. The first and second terms of the equation show the kinetic and potential energies, and p_s is the conjugate momentum of the s faster coordinates in the virtual system. g is the degree of freedom of the virtual system, and k_b and T_0 are the Boltzmann's constant and the target

temperature. From this Hamiltonian, equation of motions for the whole system is

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \quad (2.25)$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \frac{\dot{s}}{s} \mathbf{p}_i, \quad (2.26)$$

$$\dot{s} = s \frac{p_s}{Q}, \quad (2.27)$$

$$p_s = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - g k_b T_0. \quad (2.28)$$

These equations are Nosé thermostat method. Hoover introduced the following transformation of the variable:

$$\zeta \equiv \frac{1}{s} \frac{\partial s}{\partial t} = \frac{p_s}{Q}, \quad (2.29)$$

and Eq. (2.25) becomes

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \quad (2.30)$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \zeta \mathbf{p}_i, \quad (2.31)$$

$$\dot{\zeta} = \frac{1}{Q} \left(\sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - g k_b T_0 \right). \quad (2.32)$$

Compared with Eq. (2.25), Eqs. (2.30), (2.31), and (2.32) are consist of three equations. Equations from (2.30) to (2.32) are called Nosé-Hoover method.

2.2.2 Langevin Piston method

Anderson found the following equations to simulate NPT ensemble with a isothermal integrator

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} + \frac{1}{3} \frac{\dot{V}}{V} \mathbf{r}_i, \quad (2.33)$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i + \frac{1}{3} \frac{\dot{V}}{V} \mathbf{p}_i, \quad (2.34)$$

$$\ddot{V} = \frac{1}{W} [P(t) - P_0], \quad (2.35)$$

$$(2.36)$$

where V is the volume, $P(t)$ is the pressure at time t , P_0 is the target pressure. W is the piston mass and related to the frequency of pressure control. The W does not affect any values of thermodynamic properties but behavior depends on

the W . Thus, variance of the simulated system is not agreement with that of the real system. In order to remove the bias, a white noise produce by the Langevin equation was added to the equations

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} + \frac{1}{3} \frac{\dot{V}}{V} \mathbf{r}_i, \quad (2.37)$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i + \frac{1}{3} \frac{\dot{V}}{V} \mathbf{p}_i, \quad (2.38)$$

$$\ddot{V} = \frac{1}{W} [P(t) - P_0] - \gamma \dot{V} + R(t). \quad (2.39)$$

Equations from (2.37) to (2.39) is called the Langevin Piston method. Where γ is frequency of collisions, and $R(t)$ is a random force, and mean of this force is 0 and variance is

$$\langle R(0)R(t) \rangle = \frac{2\gamma k_b T \delta(t)}{W}. \quad (2.40)$$

2.3 Potential energy between molecules

In the MD simulations, the force is calculated by the potential energy between particles. However, it is difficult to define the multi-particle interactions strictly. In general, the potential energy of a multi-particle system can be considered as a function of coordinates of particles

$$U_N(\{\mathbf{r}\}) = u_0 + \sum_i u_1(\mathbf{r}_i) + \sum_{i,j} u_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i,j,k} u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k). \quad (2.41)$$

u_0 is the constant term. The second term corresponds to external fields, such as electric or magnetic fields. When we perform the MD simulations, we do not consider these fields; thus, we omit this and constant terms. The third term shows the two-body interactions which are defined by coordinates of two particles. As well as the third term, the forth and fifth terms are defined by coordinates of three and four particles, respectively (three and four-body interactions). Theoretically, as the number of considered interactions increases, the precision and computational cost of the simulations also increase.

2.3.1 Intramolecular potential energy

Interaction can be divided into two types: intramolecular and intermolecular interactions. Intramolecular interactions are defined as the interactions which are

considered only for atoms within the same molecule. A bond potential energy is one of the intramolecular potential energy, and it represents the energy between covalently bonded atoms

$$U_{\text{Distance}} = K (r_{ij} - r_{ij}^0)^2, \quad (2.42)$$

where r_{ij} is a distance between particles i and j , and r_{ij}^0 is the equilibrium bond length. This potential makes the system being equilibrium bond length. An angle potential energy is also the intramolecular and three-body potential energy. This potential energy is defined by

$$U_{\text{Angle}} = A (\theta_{ijk} - \theta_{ijk}^0)^2, \quad (2.43)$$

where θ_{ijk} is an angle of particles i , j , and k , and θ_{ijk}^0 is the equilibrium angle. A dihedral potential is the intramolecular and four-body potential energy and described as

$$U_{\text{Dihedral}} = \frac{V}{2} [1 + \cos(n\varphi - \gamma)], \quad (2.44)$$

where φ is a dihedral angle between the plane with the particles i , j , and k , and the plane with the particles j , k , and l . n is frequency, and represents the number of the optimal points in the conformation. γ is the phase.

2.3.2 Intermolecular potential energy

Intermolecular potential energy is the interaction between the molecules. The van der Waals interaction is one of the intra- and intermolecular interaction. Intensity of the attraction by the van der Waals is proportional to r^{-6} which is shown by quantum mechanics. On the other hand, repulsion is often described empirically. Lennard-Jones (LJ) potential is one of the most famous potential energy to simulate the van der Waals interactions:

$$U_{\text{LJ}} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right]. \quad (2.45)$$

This interaction is proportional to r^{-6} and r^{-12} ; thus, near molecules highly contribute to this interaction. Therefore, we usually use cut-off method and omit interactions with far molecules.

Coulomb interaction is one of the inter- and intramolecular interactions. In the classical MD simulations, change of molecules is defined by using point charges.

Coulomb interaction is defined by interactions between these point charges:

$$U_{\text{Coulomb}} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \quad (2.46)$$

where q_i is the charge, r_{ij} is a distance between point charges i and j , and ϵ_0 is the vacuum permittivity. In case of the LJ potential energy, we can omit interactions with far molecules because this function rapidly converges to 0. However, in case of the Coulomb interaction, we can not omit interactions with far molecules since the coulomb interaction is proportional to r^{-1} . In this study, we mainly used Ewald method with periodic boundary condition for calculation of the Coulomb interactions.

2.4 Periodic boundary condition and Ewald method

Real events are mainly observed as macroscopic phenomena which are composed of Avogadro's number order of particles. The amount of particles is much larger than that of particles in the MD simulations. In order to remove the artificial interactions with the interface and simulate the bulk situation like the real events, the periodic boundary condition is applied to the simulation system. The periodic boundary conditions are conditions that a particle is moved to opposite side when the particle passed through the interface. In the periodic boundary conditions, we used a space-filling system and calculate interactions and time evolution as all the space is filled by the system. With the periodic boundary conditions, we can use a useful method for calculating the Coulomb interactions: the Ewald method. The Coulomb potential energy under the periodic boundary conditions can be calculated as

$$U = \frac{1}{2} \sum_{\mathbf{n}} \sum_i \sum_{j(j \neq i, \mathbf{n}=\mathbf{0})} \frac{q_i q_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|}, \quad (2.47)$$

where \mathbf{n} is a vector of the cell position in the periodic boundary conditions, and $\mathbf{n} = (0, 0, 0)$ is the standard cell. The Ewald method calculates a fast-convergence part in the real space and a slow-convergence part in the reciprocal space, which results in the efficient calculations. The Ewald method divides Eq. (2.47) into two

equations:

$$U = U_1 + U'_2, \quad (2.48)$$

$$U_1 = \frac{1}{2} \sum_{\mathbf{n}} \sum_i \sum_{j(j \neq i)} \frac{q_i q_j}{4\pi\epsilon_0} \left\{ \frac{1}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|} - \frac{\text{erf}(\alpha|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|)}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|} \right\}, \quad (2.49)$$

$$U'_2 = \frac{1}{2} \sum_{\mathbf{n}} \sum_i \sum_{j(i \neq j, \mathbf{n}=\mathbf{0})} \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erf}(\alpha|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|)}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|}, \quad (2.50)$$

$$(2.51)$$

where erf is the error function

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \quad (2.52)$$

U_1 is the fast-convergence function in the real space; thus, we can apply cut-off with high precision. U'_2 is the slow-convergence function in the real space; however, becomes fast-convergence in the reciprocal space. Let $r = |\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|$, we obtained the following function

$$U'_2 = U_2 + U_3, \quad (2.53)$$

$$U_2 = \frac{1}{2} \sum_{\mathbf{n}} \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erf}(\alpha r)}{r}, \quad (2.54)$$

$$U_3 = -\frac{1}{2} \sum_i \frac{q_i^2}{4\pi\epsilon_0} \lim_{r \rightarrow 0} \frac{\text{erf}(\alpha r)}{r}. \quad (2.55)$$

Then, U_2 is calculated by the Fourier transform,

$$U_2 = \frac{1}{2L^3} \sum_{\mathbf{G}(\mathbf{G} \neq \mathbf{0})} \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon_0} \exp[i\mathbf{G} \cdot (\mathbf{r}_i - \mathbf{r}_j)] \frac{4\pi}{\mathbf{G}^2} \exp\left(-\frac{\mathbf{G}^2}{4\alpha^2}\right), \quad (2.56)$$

where \mathbf{G} is the reciprocal lattice vector, and we omit $\mathbf{G} = \mathbf{0}$ when the total charge is 0. This function is the fast-convergence function; thus we can apply cut-off.

2.5 Ab initio molecular dynamics simulation

To simulate the vibrational spectra of water molecules, we utilized the *ab initio* MD simulations. The *ab initio* MD is a MD simulation with the quantum mechanics for calculating the molecular interactions (see Fig. 2.1). For the calculation based on the quantum mechanics, we have to solve the many-body Schrödinger equation; however, the equation is basically intractable. Density function theory (DFT) is a computational quantum mechanical modeling method, and this method provides the properties of a many-electron system using functionals of electron density.

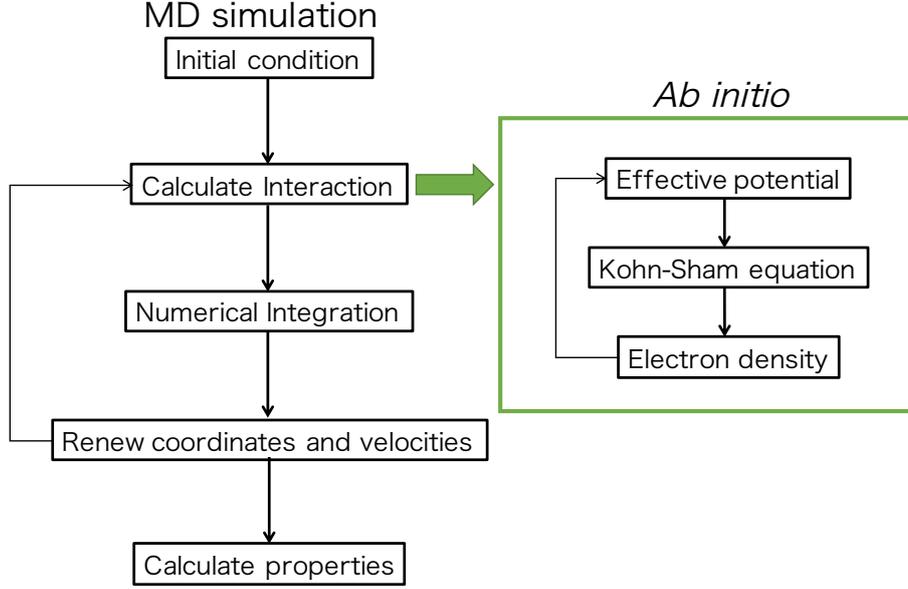


Fig. 2.1 Schematic flow of *ab initio* MD simulations.

2.5.1 Born-Oppenheimer approximation

Let a system be composed of N electrons and N_a nuclei, the Schrödinger equation is given by

$$H_{tot}\Psi_{tot}(\mathbf{r}, \mathbf{R}) = E_{tot}\Psi_{tot}(\mathbf{r}, \mathbf{R}), \quad (2.57)$$

where $\Psi_{tot}(\mathbf{r}, \mathbf{R})$ is the wave function of the system, E_{tot} is the total energy, and \mathbf{r} and \mathbf{R} are coordinates of electrons and nuclei. H_{tot} is Hamiltonian operator of the total energy, and is composed of five terms

$$\begin{aligned} H_{tot} &= -\sum_i \frac{1}{2} \nabla_i^2 - \sum_I \frac{1}{2} \nabla_I^2 + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= T_e + T_n + V_{ee} + V_{ne} + V_{vv}. \end{aligned} \quad (2.58)$$

The first and second terms represent the kinetic energy of the electrons and nuclei, and the other terms represent potential energies of electron-electron, nucleus-electron, and nucleus-nucleus interactions. In real systems, the mass of nucleus is much greater than that of electron, e.g., the mass of the lightest atom (hydrogen atom) is 1836 times greater than that of the electron. Therefore, the kinetic energy of nuclei is much smaller than that of electrons, and we can ignore the second term in Eq. (2.58). Then, since the motions of electrons and nuclei are separated, and

the eigenfunction of the Eq. (2.57) also can be described in the electron and nucleus components

$$\Psi_{tot}(\mathbf{r}, \mathbf{R}) = \Psi_{el}(\mathbf{r}, \mathbf{R})\Psi_{nub}(\mathbf{R}), \quad (2.59)$$

$\Psi_{el}(\mathbf{r}, \mathbf{R})$ is the eigenfunction of the Hamiltonian of electrons.

2.5.2 Hohenberg-Kohn theorems

The Hamiltonian of electrons is

$$H = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_i V_{\text{ext}}(\mathbf{r}_i) + \sum_{i \neq j} U(\mathbf{r}_i, \mathbf{r}_j), \quad (2.60)$$

where $U(\mathbf{r}_i, \mathbf{r}_j)$ is the potential energy of electron-electron interactions, and V_{ext} is the potential energy from the external field by the nuclei. Hohenberg and Kohn introduced the electron density function $\rho(\mathbf{r})$ as a basic variable. Then, they made the two theorems,

- The external potential energy is uniquely determined by the corresponding ground-state electronic density function $\rho_0(\mathbf{r})$ within a additive constant.
- The energy of the electrons as a functional of $\rho(\mathbf{r})$ is the minimum when $\rho(\mathbf{r}) = \rho_0(\mathbf{r})$. The energy function $E_{HK}[\rho]$ is

$$E_{HK}[\rho] = T[\rho] + E_{\text{int}}[\rho] + \int d\mathbf{r} V_{\text{ext}}\rho(\mathbf{r}) + E_{II} \quad (2.61)$$

$$\equiv F_{HK} + \int d\mathbf{r} V_{\text{ext}}\rho(\mathbf{r}) + E_{II}, \quad (2.62)$$

where E_{II} is nuclei-nuclei interactions.

2.5.3 Kohn-Sham equation

The Kohn-Sham equation is the one electron Schrödinger equation of a virtual system, called Kohn-Sham system. The energy of ground-state in Kohn-Sham method is

$$E_{KS}[\rho] = T_s[\rho] + \int d\mathbf{r} V_{\text{ext}}\rho(\mathbf{r}) + E_H[\mathbf{r}] + E_{xc}[\mathbf{r}], \quad (2.63)$$

where T_s is the Kohn-Sham kinetic energy

$$T_s = -\frac{1}{2} \sum_{\sigma} \sum_i \langle \psi_i^{\sigma} | \nabla^2 | \psi_i^{\sigma} \rangle \quad (2.64)$$

$$= \frac{1}{2} \sum_{\sigma} \sum_i \int d\mathbf{r} |\nabla \psi_i^{\sigma}(\mathbf{r})|^2, \quad (2.65)$$

where σ is spins of electrons. $E_H[\mathbf{r}]$ is the Hartree energy which is same as classical Coulomb interactions of electron density

$$E_H[\mathbf{r}] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho'(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.66)$$

E_{xc} is the exchange-correlation energy, and any other external potentials are represented by the energy including many-body interactions of electrons and interactions from the nuclei. Then, Kohn-Sham equation is

$$(H_{HK}^\sigma - \epsilon_i^\sigma)\psi_i^\sigma(\mathbf{r}) = 0, \quad (2.67)$$

where $H_{KS}^\sigma(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + V_{\text{ex}}^\sigma(\mathbf{r})$. It provides self-consistent solutions.

2.5.4 Car-Parrinello MD simulation

The Car-Parrinello MD (CPMD) method is an approximation of the Born-Oppenheimer MD method which separately develops electron and nuclei. The CPMD method develops electrons close to the ground-state without the calculation of self-consistent process. Thus, the computational costs of the CPMD method is much lower than those of the Born-Oppenheimer MD method. In the CPMD method, fictitious dynamics of the electrons was introduced and the Lagrangian of the system is

$$L_{CP} = \frac{1}{2} \sum_I M_I \dot{\mathbf{R}}_I^2 + \sum_i^N \int \mu |\dot{\psi}_i(\mathbf{r})|^2 d\mathbf{r} \quad (2.68)$$

$$-E_{KS}[\psi, \mathbf{R}] - \sum_{i,j} \Lambda_{ij} \left[\int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) d\mathbf{r} - \delta_{ij} \right], \quad (2.69)$$

where μ is the fictitious mass of the orbital degree of freedom and Λ is the Lagrangian multiplier matrix. Newton's equations of motions are derived by the Euler-Lagrange equations

$$\frac{d}{dt} \left(\frac{\delta L_{CP}}{\delta \dot{\psi}_i^*(\mathbf{r})} \right) = \frac{\delta L_{CP}}{\delta \psi_i^*(\mathbf{r})}, \quad (2.70)$$

$$\frac{d}{dt} \left(\frac{\delta L_{CP}}{\delta \dot{\mathbf{R}}} \right) = \frac{\delta L_{CP}}{\delta \mathbf{R}}. \quad (2.71)$$

Then, equations of motions of Kohn-Sham orbital and nuclei are

$$\mu\ddot{\psi}_i(\mathbf{r}) = -\frac{\delta E_{KS}[\psi, \mathbf{R}]}{\delta \psi_i^*(\mathbf{r})} + \sum_j \Lambda_{ij} \psi_j(\mathbf{r}), \quad (2.72)$$

$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial E_{KS}[\psi, \mathbf{R}]}{\partial \mathbf{R}_I} + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}} \int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r}. \quad (2.73)$$

Chapter 3

Interactions between water molecules and prion protein

3.1 Materials and methods

We performed two independent all-atom MD simulations for two systems: the human WT-PrP^c and the pathogenic mutant T188R. All simulations were performed totally for 1.2 μ s. The initial structure of WT-PrP^c was obtained from NMR structural data (PDB code:1HJM) which contains residues 125-228 with a disulfide bond between C179 and C214 [90]. We made the initial structure of T188R using the Swiss-PdbViewer [91] by replacing residue 188 from threonine to arginine. Each prion protein was set on the center of a cuboid box and solvated with TIP3P water molecules. The systems contained 12,761 water molecules and three and two sodium ions in WT-PrP^c and T188R as counterions, respectively. All the MD simulations were carried out with NAMD 2.9 simulation package [92]. The AMBER ff99SBildn force field was used [93]. The SHAKE algorithm was used for restraint of the covalent bonds of both prion proteins and water molecules. Langevin dynamics thermostat and Nosé-Hoover Langevin piston barostat were applied for the temperature and pressure coupling. Switching cutoffs starting from 0.8 nm to 1.4 nm was applied to both electrostatic and van der Waals interactions. The Particle Mesh Ewald method was used to calculate the electrostatic interactions. The periodic boundary condition was applied for each system. First, we performed 1000 steps of minimization using the conjugate gradient algorithm only for water molecules. After isothermal 1 ns *NVT* simulation (number of particles, volume, and temperature) at a temperature of 310 K, we performed two 600 ns isothermal isobaric *NPT* simulations (number of particles, pressure and temperature) for each system at 310

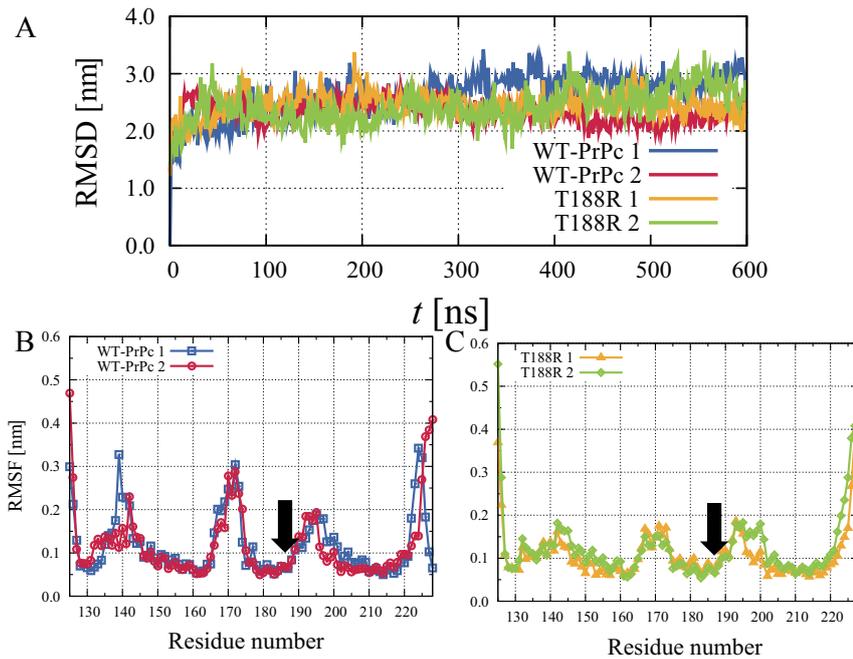


Fig. 3.1 (A) The $C\alpha$ -RMSD from 0 ns to 600 ns. We used the notations 1 and 2 to express two different simulations for WT-PrPc and T188R. The reference structure is a crystal structure obtained from Protein Data Bank. We have calculated the $C\alpha$ -RMSDs after fitting to the reference. (B) $C\alpha$ -RMSF of WT-PrPc and (C) $C\alpha$ -RMSF of T188R averaged from 200 ns to 600 ns. The vertical axis shows amino acid number. Black arrow indicates mutation point residue 188.

K and a pressure of 0.1 MPa with a time step of 2 fs.

3.2 Results

3.2.1 Equilibrium of the systems

$C\alpha$ -RMSD is an indicator of stability and conformational change of proteins. Since we are only interested in the region related to secondary structure (from residue 125 to 225), three N and C terminal residues are omitted from following analysis.

Figure 3.1A shows the $C\alpha$ -RMSDs of WT-PrPc and T188R for all simulations. All $C\alpha$ -RMSDs reached equilibrium states at 200 ns; therefore, we determined that every system reached an equilibrium state at 200 ns. Later analysis was conducted using the data after equilibrium, which accounts for totally 800 ns simulation data for WT-PrPc and T188R.

3.2.2 Fluctuations of amino acids in the protein

For analysis of structural flexibility, we calculated the C α -RMSF for all simulations. The C α -RMSF indicates the intensity of fluctuations, which is defined as

$$\text{RMSF}_i = \sqrt{\frac{1}{T} \sum_{t=1}^T (\mathbf{x}_i(t) - \bar{\mathbf{x}}_i)^2}, \quad (3.1)$$

where $\mathbf{x}_i(t)$ is the coordinates of C α of the i th amino acid, $\bar{\mathbf{x}}_i$ is the averaged coordinate and T is the measurement time after reaching equilibrium state. Before the RMSF analysis, we rotate the structure to minimize the C α -RMSD. Figures 3.1B and 3.1C show the C α -RMSFs of WT-PrPc and T188R. The distributions of the C α -RMSF values are not dramatically different. There are five peaks in the RMSFs of WT-PrPc and T188R. These results show that secondary structure does not dramatically change by the point mutation in this time scale.

3.2.3 Time series of secondary structure

The secondary structure is defined by the Dictionary Secondary Structure of Protein (DSSP) method [94]. Figures 3.2A and 3.2B show the time series of secondary structure and snapshots of the WT-PrPc and the T188R, respectively. We confirm that there are two stable sheets and three helical structures in all simulations. In T188R, the secondary structure is stable after around 200 ns. Secondary structure does not dramatically change by the mutation; however, several regions change into different secondary structures. By the T188R mutation, a part of helix 2 (residue 174 to 178) disrupts and changes into a coil structure, which could induce the transition into the scrapie form [95–97], and β -sheet 2 also increases. β -sheets are putative elements for aggregation [98]. Table 3.1 shows the proportion of secondary structures. This also indicates that T188R causes a decrease of α -helix content and an increase of β -sheet content.

3.2.4 Buried water molecules

In this section, we focused on the interactions between water molecules and the prion protein. Hydration structure is directly related to conformation, thermodynamic properties, and stability of protein [2]. Here, we focus on buried hydration sites. To identify the buried hydration sites, we used the mean residence time analysis. The

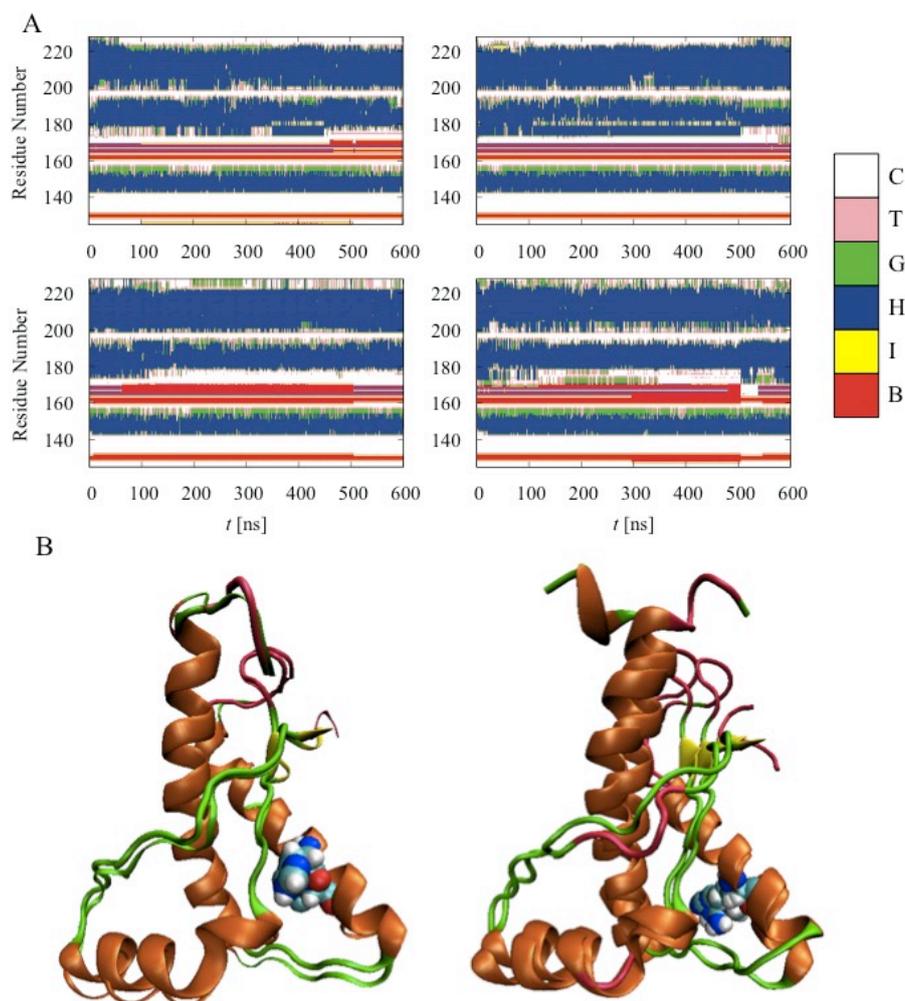


Fig. 3.2 Secondary structure of WT-PrPc and T188R. (A) Evolution of secondary structure of WT-PrPc (upper) and T188R (bottom). Secondary structure is calculated by DSSP method. C, T, G, H, I and B indicate coil, turn, 3_{10} -helix, α -helix, π -helix and β -sheet or anti-parallel β -sheet, respectively. (B) The secondary structures of WT-PrPc and T188R. The secondary structure is the mean coordinates. The α -helix, β -sheet, random coil, and turn regions are colored orange, yellow, green, and gold. Residue 188 is shown as beads form.

residence time is defined as the time in which a water molecule resides in a sphere centered at an atom in the protein. We used amino-acid-based residence time, while previous study used grid-based residence time. Because secondary structure is not stable during the simulations, the amino-acid-based analysis is better than the grid-based analysis especially for the case of the mutant [40]. The radius of the sphere was set to 0.4 nm which is larger than ordinal hydrogen bond length. However,

Table 3.1 Proportion of secondary structures in WT-PrPc and T188R

| | α -helix | 3_{10} -helix | π -helix | β -sheet | anti β -sheet |
|-----------|-----------------|-----------------|--------------|----------------|---------------------|
| WT-PrPc 1 | 49.17 | 6.14 | 0.003 | 2.72 | 8.10 |
| WT-PrPc 2 | 50.47 | 5.29 | 0.000 | 3.85 | 5.77 |
| T188R 1 | 46.44 | 7.69 | 0.003 | 1.92 | 11.54 |
| T188R 2 | 46.35 | 8.60 | 0.000 | 4.62 | 9.26 |

Table 3.2 Mean residence times, total residence times of (A) WT-PrPc and (B) T188R

| (A) WT-PrPc | | | | | |
|-------------|-------------------|------------|---------|-------------------|------------|
| PrPc 1 | Mean [ns] | Total [ns] | PrPc 2 | Mean [ns] | Total [ns] |
| Site 1 | 199.8 ± 37.81 | 399.6 | Site 1 | 3.577 ± 2.509 | 114.5 |
| Site 2 | 57.34 ± 52.87 | 401.4 | Site 2 | 7.413 ± 4.383 | 244.6 |
| Site 3 | 62.94 ± 33.93 | 377.9 | Site 3 | 10.94 ± 5.469 | 404.7 |
| Site 4 | 7.277 ± 2.890 | 356.7 | Site 4 | 53.22 ± 28.57 | 266.1 |
| (B) T188R | | | | | |
| T188R 1 | Mean [ns] | Total [ns] | T188R 2 | Mean [ns] | Total [ns] |
| Site 1 | 0.202 ± 0.002 | 31.28 | Site 1 | 13.08 ± 8.207 | 366.2 |
| Site 2 | 3.182 ± 0.736 | 308.6 | Site 2 | 9.159 ± 5.139 | 338.9 |
| Site 3 | 2.955 ± 0.526 | 428.5 | Site 3 | 15.13 ± 8.914 | 408.0 |
| Site 4 | 6.187 ± 1.043 | 377.4 | Site 4 | 42.78 ± 22.24 | 299.4 |

this definition cannot distinguish whether a water molecule completely leaves from the site or not. Therefore, we determine that a water molecule leaves from the site when the distance between the protein atom and the water molecule is greater than 0.8 nm.

In Figure 3.3, the mean residence times on each atom are mapped on protein atoms. Most of the mean residence times are less than 100 ps; however, there are quite high mean residence times at some atoms, i.e., the buried hydration sites (see table 3.2). From the high residence time areas, we identified four buried hydration sites, Site 1 (G131, V161, and Q217), Site 2 (Q217, V220, and R228), Site 3 (F175, R163, and R164), and Site 4 (Y162, F175, and T183) in both WT-PrPc and T188R. Hence, the mutation does not disrupt the buried hydration sites.

Figure 3.4A shows the locations of the four buried hydration sites superposed on the secondary structure of WT-PrPc 1. All buried water molecules are basically fixed by three amino acids. Sites 1 and 2 are composed of G131, V161, Q217, R220, and R228 [see Fig. 3.4B]. In Site 1, the buried water molecule has hydrogen bonds

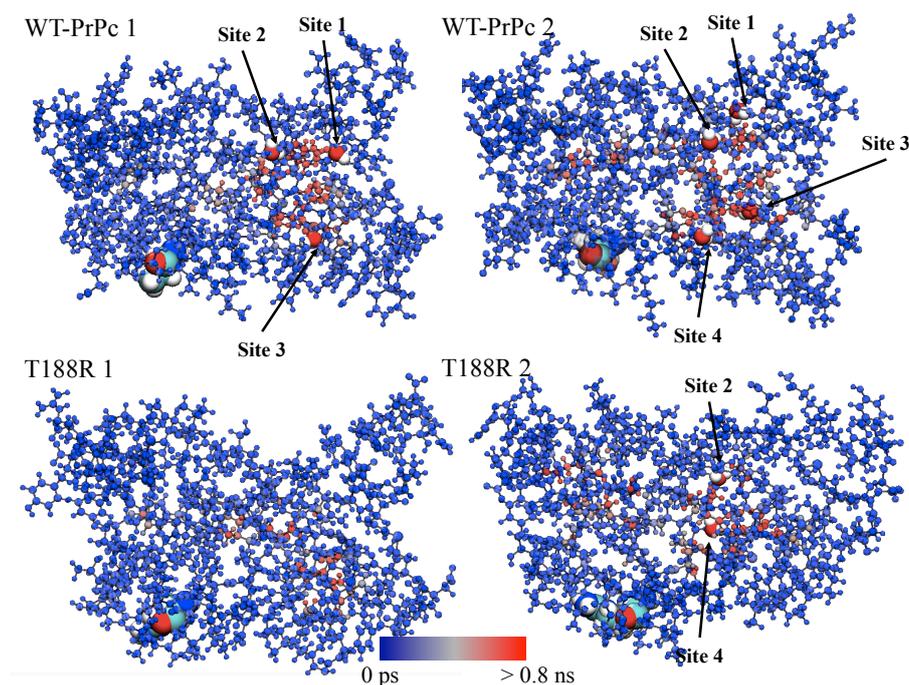


Fig. 3.3 The mean residence times were mapped on atoms of prion protein for each calculation. The buried water molecules are explicitly visualized in the sites with residue 188 shown as beads form.

with G131 and V161 as a acceptor and Q217 as a donor. The buried water molecule in Site 2 also makes hydrogen bonds with Q217 as a donor and with R220 and R228 as a acceptor. In other words, Q217 is structurally stabilized by two buried water molecules.

The buried water molecules in Sites 3 and 4 are wrapped by Y162, Y163, R164, F175, C179, and T183 (see Fig. 3.4C). In Site 3, the buried water molecule is fixed by hydrogen bond with F175 as a donor and with R164 or Y163 as a acceptor. Site 4 is composed of Y162 and F175, and T183 as an acceptor and a donor, respectively. These buried hydration sites are stable during the MD simulations. Table 3.2 shows that the mean residence times of the buried hydration sites are much longer than those of non buried hydration sites. We confirmed that Sites 1 and 3 are identical to two of the buried hydration sites reported in De Simone *et al.* [40]. Site 4 was not reported in previous study, but Sites 4 is located quite near Site 3. In other words, grid-based analysis cannot distinguish these sites. Site 2 was also not reported because R228 was truncated in previous study.

We also identify the four buried hydration sites in T188R. However, in T188R

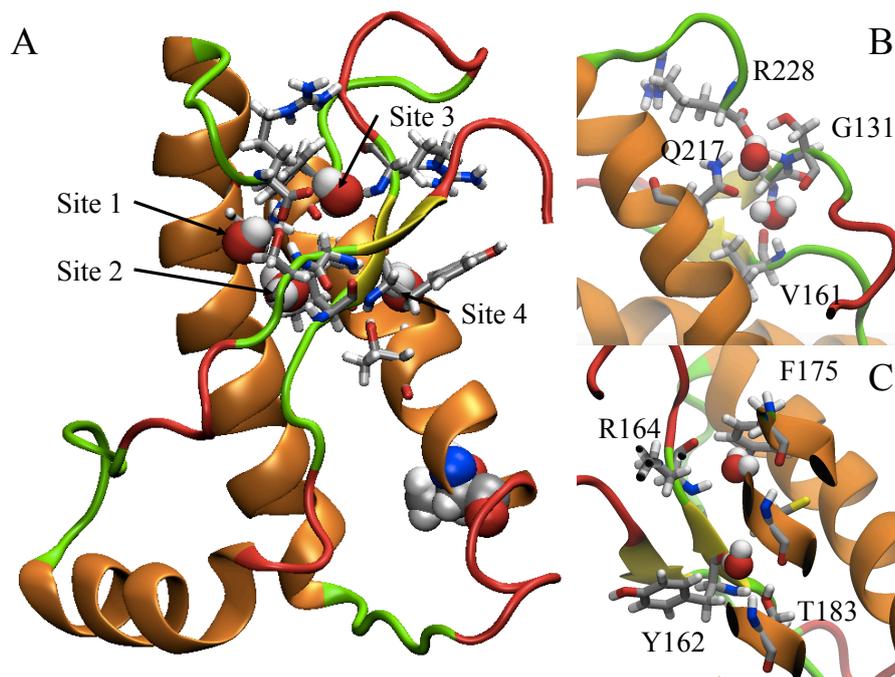


Fig. 3.4 Position of the water molecules in WT-PrPc 2. (A) The location of the buried water molecules superposed onto the secondary structure of WT-PrPc. Amino acids forming hydrogen bonds with the buried water molecules are explicitly visualized. The pathological mutation is shown as beads shape. (B) Hydrogen bond networks with amino acids around the buried water molecules. Top and bottom buried hydration sites are site 1 and 2. (C) Hydrogen bond networks around site 3 (top) and 4 (bottom).

1, the mean residence times become smaller, and there are no areas which have long mean residence times (more than 10 ns). This result implies that the buried hydration sites become unstable, and the sites cannot keep hydration configuration. On the other hand, the several buried hydration sites are still stable in T188R 2. These results show that the pathogenic point mutation T188R sometimes makes the buried hydration sites unstable.

To investigate a molecular mechanism of the decrease of the mean residence times, we consider the total residence time defined as a sum of residence times of all water molecules residing at the site. There are two mechanisms of the decrease of mean residence times. First, the buried hydration site is more exposed to other bulk water molecules. In this case, the mean residence time decreases but total residence time increases or does not change. Second, the buried hydration site is more wrapped by amino acids and less exposed to other bulk water molecules. Water molecules trying to access to the site; however, they are interrupted by other amino acids,

which leads to the decrease of both the total residence time and the mean residence time. In T188R 1, the mean residence times of Sites 2, 3 and 4 decrease while the total residence times do not change, which is caused by the first mechanism. However, both the total and the mean residence times of Site 1 decrease because of the second mechanism.

3.2.5 Reorientational time correlation function

Because these buried water molecules are stabilized by hydrogen bonds, the rotational motions of the water molecules are also restricted; therefore, reorientational relaxation time becomes slow [99]. We calculate the reorientational time correlation function according to the following equation:

$$C(t) = \left\langle \frac{1}{2} \left[3 \{ \mathbf{u}(t_0) \cdot \mathbf{u}(t_0 + t) \}^2 - 1 \right] \right\rangle, \quad (3.2)$$

where t_0 is the entry time of a water molecule into the sphere, $\mathbf{u}(t_0 + t)$ is the dipole vector of the water molecule at time $t + t_0$. Figure 3.5 shows the reorientational time correlation function of buried water molecules within 1 ns. The results clearly show that the rotational relaxation of all buried water molecules is dramatically slower than that of bulk water molecules. All buried water molecules are fixed by hydrogen bonds with amino acids in all systems.

Basically, the rotational relaxation of water molecules consists of two processes. First, hydrogen of water molecules librates around the hydrogen bond, called wobbling-in-the-cone motion. This motion gives rise to a fast rotational relaxation less than 1 ps [100]. The longer relaxation motion represents breaking and re-forming of hydrogen bonds with other water molecules, called a jumping and frame tumbling. Rotational relaxation by jumping and frame tumbling has a different relaxation time depending on the condition. Figure 3.5 shows that all water molecules including bulk water molecules have fast relaxation, which results in initial decrease of reorientational time correlation function. For a longer relaxation, all rotational relaxations of buried water molecules are found to be much slower than that of bulk. Because intensity is different for each other, the availability of switching new hydrogen bond differs [101].

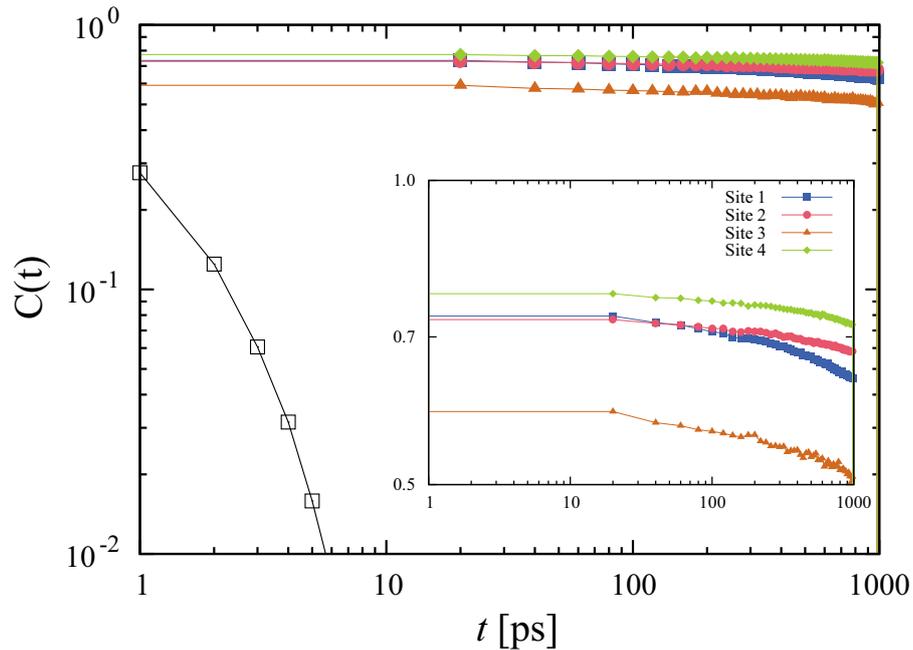


Fig. 3.5 The reorientational correlation function of the buried water molecules in WT-PrPc 2. Blue, orange, red, green, and black lines show water molecules in Site 1, 2, 3, 4, and bulk, respectively. The inset is zoomed range from 0.5 to 1.0.

Table 3.3 Mean distance and coefficient of variance between subdomains S1-H1-S2 and H2, 3

| | Mean distance [nm] | CV (10^{-4}) |
|-----------|--------------------|------------------|
| WT-PrPc 1 | 0.399 | 26.8 |
| WT-PrPc 2 | 0.429 | 44.2 |
| T188R 1 | 0.347 | 59.2 |
| T188R 2 | 0.461 | 26.9 |

3.2.6 Relationship between the fluctuations and stability of the buried hydration sites

Performing MD simulations of WT-PrPc and T188R, we found the following two key points: (i) T188R mutation affects secondary structure leading to the increase of β -sheet content and the decrease of helix 2. (ii) Although the mutation point is located far from the buried hydration sites, the sites become unstable by the mutation. Previous studies have shown that T188 is an interface between two subdomains [102, 103]. We divide the prion protein into two subdomains: S1, H1 and S2 (from residue 125 to residue 171), and H2 and H3 (from residue 172 to 228).

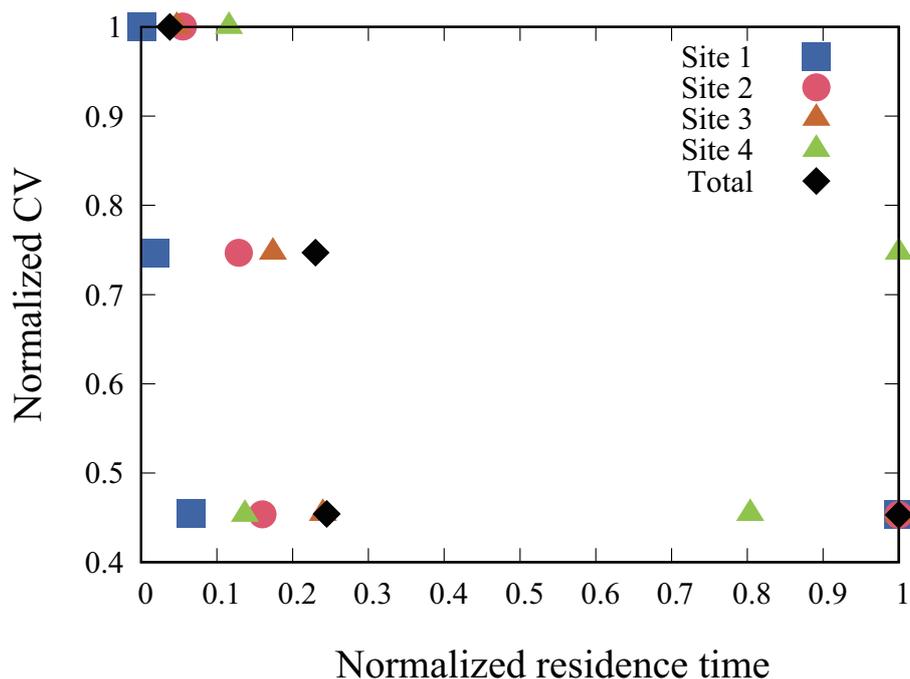


Fig. 3.6 Normalized CV and normalized residence time of each site.

We calculate the mean distance between the centers of mass of subdomains, and the coefficient of variance (CV) (see table 3.3). In T188R 1, the CV is higher than those in other simulations. Because all buried water molecules connect two subdomains, the stability of the distance between two subdomains is critical for survival of the buried hydration sites. We observe a clear correlation between the CV and the mean residence time (see Fig. 3.6). It suggests that these large fluctuations of the distance make the buried hydration sites fragile.

3.2.7 Molecular configuration around the mutation

We also focus on a configuration around the mutation. In WT-PrPc, side chain of T188 was embedded in the protein including G200. In T188R, we replaced the residue 188 from threonine to arginine. Because the side chain of arginine is larger than that of threonine, the side chain of R188 cannot be embedded in protein and exposed to the bulk in T188R 1 (see Fig. 3.7A). However, R188 in T188R 2 was mainly connected to the side chain of G200, which may contribute to the stability of T188R 2 (see Fig. 3.7B). In the former case, distance between the closest atoms of R188 and G200 is around 0.3 nm while that in latter case is around 0.2 nm (see Fig. 3.7) Figure 3.7 shows the distance between the closest atoms of the two residues.

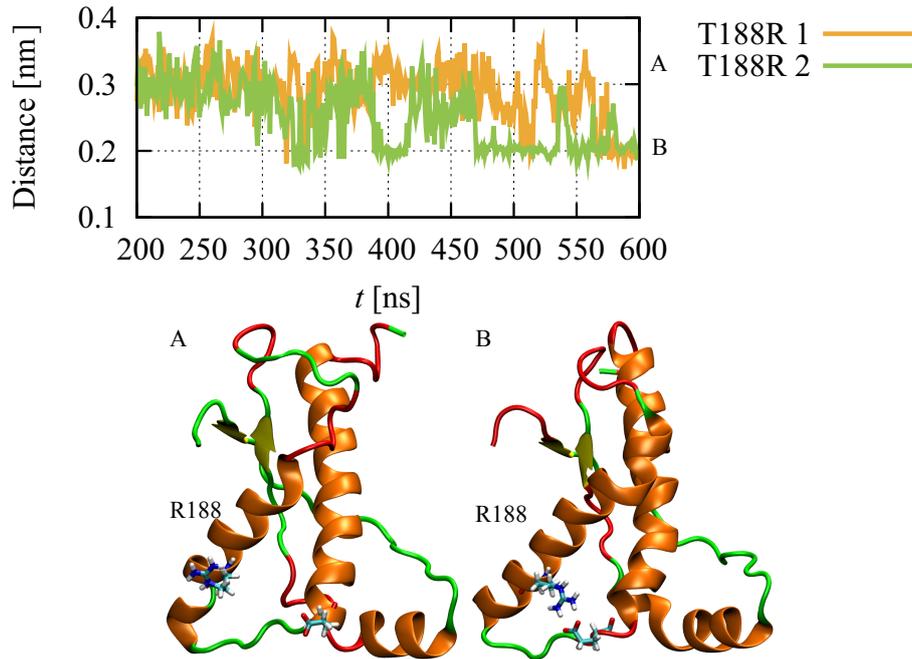


Fig. 3.7 Distance between the closest atoms of R188 and G200. The mean distances for T188R1 and T188R2 are 0.288 nm with the standard deviation 0.04 nm and 0.249 nm with the standard deviation 0.05 nm, respectively. (A) and (B) are snapshots of T188R1 at $t=450$ ns and T188R2 at $t=500$ ns, respectively.

R188 is sometimes bent and thus the interaction between the two residues are weak. This configuration is dominant in T188R 1, while it is not in T188R 2. Moreover, the statistical test implies that the mean distance in T188R1 is significantly larger than that in T188R 2. Overall, these data support following three things: 1) T188R which is not adjacent to the buried hydration sites sometimes makes the sites fragile, 2) subdomains fluctuations are tightly correlated with the stability of the buried hydration sites, and 3) because of the size of side chain, R188 cannot be embedded in protein, and this instability may trigger the subdomains fluctuations. Our findings provide a possible explanation for the instability by mutation which is not directly adjacent to the buried hydration sites.

3.2.8 Summary

we have performed all-atom MD simulations of WT-PrPc and a pathogenic point mutant T188R to investigate conformational changes and the buried hydration sites. By mutation T188R, a part of helix 2 (residue 174 to 178) is disrupted and β -sheet 2 increases in both T188R. In WT-PrPc, we have identified the four buried hydration

sites, and we have found that the rotational relaxations of all buried water molecules are much slower than that of bulk water. This means that all buried water molecules are fixed in stable position by hydrogen bonds with amino acids stably. In one of T188R simulations, on the other hand, there are no stable buried hydration sites. These results show that T188R sometimes makes the buried hydration sites fragile. Subdomain analysis has revealed that the instability of the buried hydration sites is related to the subdomains fluctuations. While there is the connection between R188 and G200 in T188R 2, R188 is mainly not connected to G200 in T188R 1. These data suggest that R188 cannot be embedded due to the size of side chain and this instability may trigger subdomains fluctuations. Our results provide a possible explanation for the instability by mutation T188R.

Chapter 4

Interactions between water molecules and rhodopsin

4.1 Materials and methods

We used the crystal structures of bovine rhodopsin because only bovine rhodopsin can be obtained as crystal structure in the three states. The structure of dark-adapted rhodopsin (PDB ID: 1U19) [104] was used as the initial structure. The dark-adapted rhodopsin was embedded in a lipid bilayer composed of 170 POPC lipids, and these systems were hydrated in 6,499 water molecules. In the case of the Opsin, we used the Opsin structure (PDB ID: 3CAP) [105] in 166 POPC lipids and 6627 water molecules. The crystal structure of the Meta II state was obtained from PDB ID: 3PXO [106]. The system was hydrated with 7,400 water molecules with 169 POPC lipids. The Asp83, Glu122 and Glu113 were protonated in Meta II state, and Glu113 was protonated in the Opsin. Co-crystallized water molecules were utilized as initial internal water molecules while palmitic acids were removed. In all simulations, we added 5 chloride, 2 chloride, and 1 sodium ions to neutralize the system of the dark-adapted rhodopsin, the Meta II state, and the Opsin, respectively. These ions were randomly placed in bulk of the systems. A disulfide bond was added between Cys110 and Cys187. All MD simulations were conducted using NAMD 2.9 [107] software. The CHARMM 36 forcefield was used for the lipids [108] and the protein [109] and TIP3P model modified for the CHARMM forcefield [110]. Force field parameters for the ligand molecules were generated with the CHARMM General Force Field [111]. Using the VMD [112], we built the initial structure of our systems. First, we embedded the rhodopsin on the center of membrane bilayer, and removed overlapping lipid molecules and overlapping water

molecules. Snapshot of the system is shown in Fig. 4.1A. After the 1 μ s simulations, box sizes of the simulation systems were $80 \times 80 \times 116$, $81 \times 81 \times 113$, and $81 \times 81 \times 116$ \AA^3 for the systems of the dark-adapted rhodopsin, the Meta II state, and the Opsin, respectively.

After the initial system setup, the system was subject to 2000 steps minimization using conjugate gradient and line search method, and *NPT* simulations of 5 ns were performed with the constraint on the protein to initial positions by a harmonic potential. Langevin dynamics thermostat and Nosé-Hoover Langevin piston barostat were applied for temperature and pressure coupling at 310 K and 1 bar. Van der Waals interactions were smoothly truncated by force switching from 10 to 12 \AA . A time step of 2 fs was used with the SHAKE algorithm applied to constrain the bond lengths involving hydrogen atoms. All systems were subjected to periodic boundary conditions. Electrostatic interactions were computed using the particle-mesh Ewald method. Each simulation was performed for 1 μ s and the analysis was conducted over the last 0.8 μ s trajectories. The data for the first 0.2 μ s were discarded for the equilibration of the simulation, which was judged by backbone RMSD (see Fig. 4.1B).

To validate the reproducibility of our results, we performed the same systems using AMBERff99SB-ildn [113] for protein, Slipid for POPC [114], and TIP3P water models [115]. Additional calculations were performed with GROMACS 5.1.2 [116]. The systems were subject to pressure scaling to 1 bar using Parrinello-Rahman barostat [117], temperature scaling to 310 K using velocity-rescaling method [118]. Force field parameters for the protonated 11-cis-retinal in the rhodopsin and the deprotonated all-trans-retinal in Meta II state were prepared using antechamber with the General Amber Force Field [119] and the AMI-BCC point charges. 1 ns *NVT* and *NPT* simulations were performed with the restrained protein for equilibration. *NPT* simulations were performed for 1 μ s with 2.0 fs time-step.

4.2 Results

4.2.1 Water molecule accessibility

To investigate interactions between water molecules and internal amino acids of the rhodopsin, we focused on internal water molecules. Rhodopsins have many internal water molecules and a flux of water molecules in its inside. Figure 4.2 shows regions

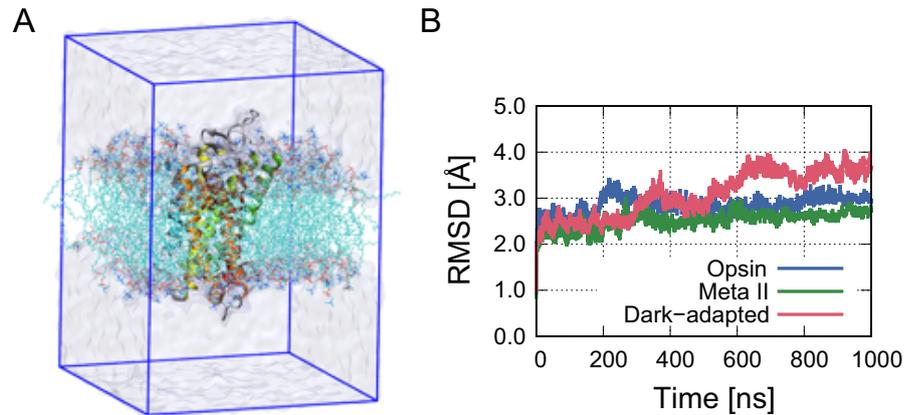


Fig. 4.1 **System for MD simulations of rhodopsin.** (A) The system used for MD simulations of the dark-adapted rhodopsin in a POPC lipid bilayer. The rhodopsin and lipids are shown in cartoon and cyan lines, respectively. Explicit water molecules correspond to the upper and lower transparent coatings. The blue line of the box is a periodic boundary. (B) Backbone RMSDs of the dark-adapted rhodopsin, the Meta II state and the Opsin. We calculated the RMSDs only using helical parts of rhodopsin.

where water molecules have accessed during the simulation. A putative solvent pore was identified on the cytoplasmic side in the Meta II state and the Opsin (see Figs. 4.2B and 4.2C). This solvent pore flows from the cytoplasmic bulk to the retinal-binding pocket. However, in the dark-adapted rhodopsin the solvent pore is clearly separated at the entrance of the solvent pore (see Fig. 4.2A). These results show that the solvent pore emerges after Meta II state due to the conformational change of the secondary structure. To validate the reproducibility of these results, we also performed MD simulations for three states of the rhodopsin using a different force field. Notably, the solvent pore was also observed in the Meta II state and the Opsin, while the solvent pore does not exist in the dark-adapted rhodopsin (see Fig. 4.3).

In all states, the solvent pore has hourglass-shaped entrance at the cytoplasmic site (see Fig. 4.4). When the water molecules enter into the rhodopsin from the cytoplasmic side, they pass through the hourglass-shaped entrance, which mediates optimal water permeation in the case of aquaporin water channels [120]. These results show that amino acids around the entrance become closed in the dark-adapted rhodopsin. Although the radius of the solvent pore is wider than the entrance of aquaporin [121], water molecules are not able to easily pass through

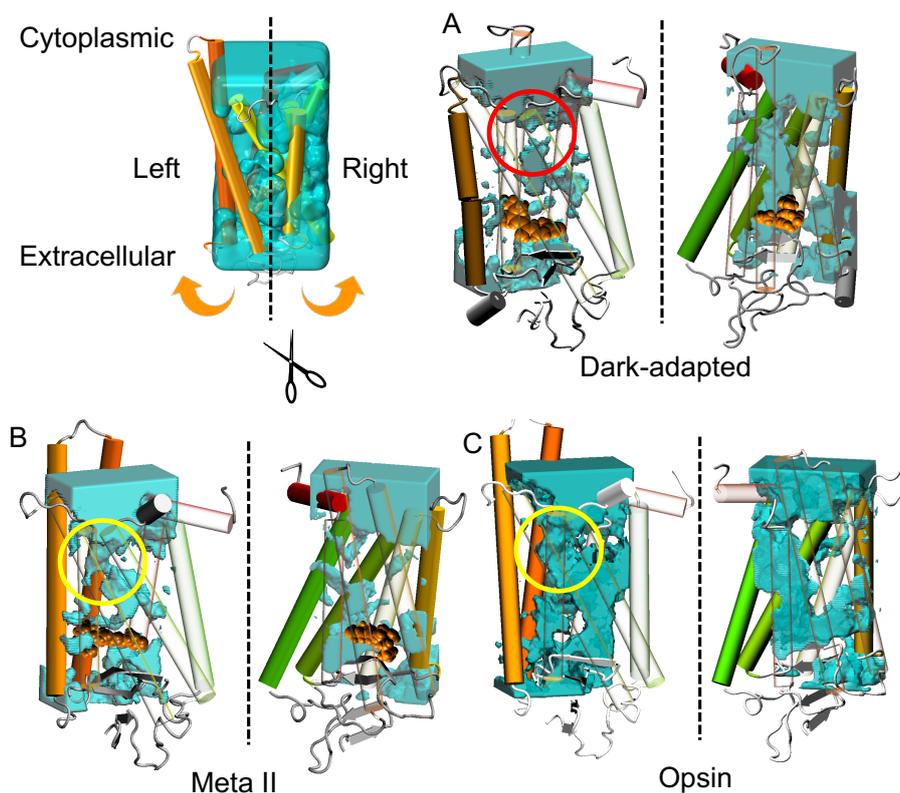


Fig. 4.2 **Accessibility of water molecules.** Cross sectional diagrams of water accessibility with the cross section taken at the middle of the rhodopsin in (A) the dark-adapted rhodopsin, (B) the Meta II state, and (C) the Opsin. Blue surfaces represent places where water molecules have reached during last $0.8 \mu\text{s}$ in the equilibrium state (see Fig. 4.1B). The protein molecules are depicted with each helix colored from red (the N terminus) to green (the C terminus). The retinal is shown in orange VDW format. A solvent pore can be identified in the Meta II state and the Opsin (yellow circle). However, the solvent pore does not exist in the dark-adapted rhodopsin (red circle).

the solvent pore. This is attributed to two narrow regions shown in the following analyses.

4.2.2 Configuration of the solvent pore in the Meta II state

Water molecules pass bi-directionally through the solvent pore. From the $1 \mu\text{s}$ MD simulation of Meta II state, we could observe 16 times permeations of water molecules through the solvent pore between the bulk and the retinal-binding pocket. All the water molecules reached the K296, which means that the solvent pore can provide water molecules from the bulk to the Schiff base linkage. Figure 4.5 shows extracted trajectories of 16 water molecules passing through the solvent pore. The

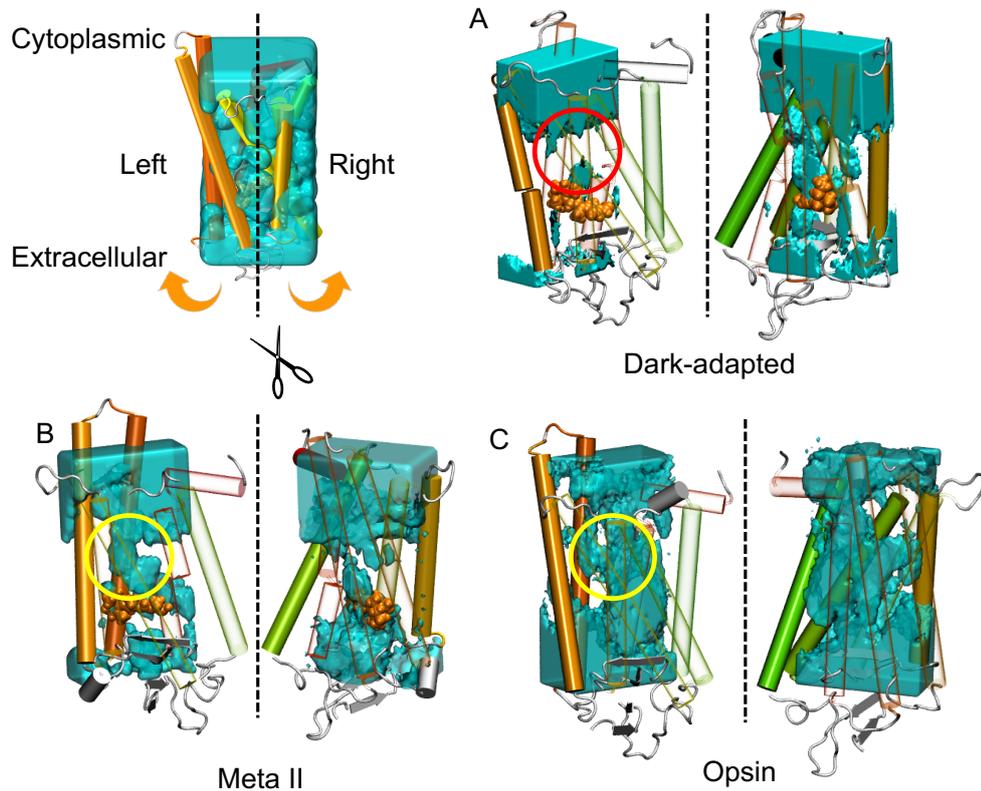


Fig. 4.3 **Accessibility of water molecules.** Cross sectional diagrams of water accessibility with the cross section taken at the middle of the rhodopsin in (A) the dark-adapted rhodopsin, (B) the Meta II state, and (c) the Opsin using Amber force field. Blue surfaces represent places where water molecules have reached during the last $0.2 \mu\text{s}$. The protein molecules are depicted with each helix colored from red (the N terminus) to green (the C terminus). The retinal is shown in orange VDW format. A solvent pore can be identified in the Meta II state and the Opsin (yellow circle). However, the solvent pore does not exist in the dark-adapted rhodopsin (red circle).

trajectories reveal two narrow regions that distinguish the bulk and the internal hydration sites. The first narrow region, which is bent into the L shape (see Fig. 4.5A), comprises three amino acids, L128, M257, and Y306. The water pathway is narrow owing to the L128 and M257 amino acids, which are parallel to z axis. Below this region, the water pathway becomes bent due to Y306, and it is connected to internal hydration sites. The second narrow region comprises two amino acids, F261 and Y306. The two narrow regions are hydrophobic due to three amino acids, leucine, tyrosine, and phenylalanine. We calculated the distance distributions of the amino acids, and the presence of single peak indicates that the solvent pore have only one state (see Fig. 4.6).

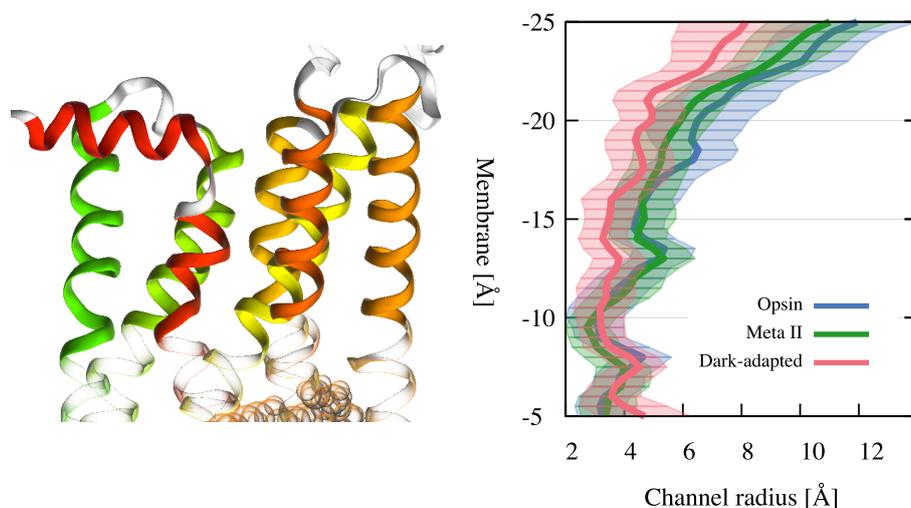


Fig. 4.4 **Hourglass-shaped entrance of the solvent pore at the cytoplasmic site.** Right figure shows the channel radius versus position along the pore axis. The center of the membrane is at 0 Å. The scale of the membrane corresponds with the left figure. The error bars drawn as transparent are given by standard deviation.

After passing through the two regions, water molecules flow into important hydration sites, including the sites adjacent to the conserved structural motif, which is called “NPxxY”. This NPxxY motif, which is highly conserved in the GPCR family, is crucial for activation of GPCRs, and comprises asparagine (N; 75% , D; 21%), proline (P; 96%), two hydrophobic residues X and tyrosine (Y; 92%, F; 3%) [122]. The internal water molecules adjacent to the NPxxY motif mediate GPCR activation and stabilize the active state of GPCRs [68, 123].

These results mean that the solvent pore provides water molecules not only for the retinal-binding pocket, but also for internal hydration sites adjacent to the key conserved motif NPxxY.

4.2.3 Water displacements within the solvent pore in the Meta II state

The analysis of water trajectories in the solvent pore provides more details about the permeation process of water molecules [124]. Figure 4.7 shows water displacements in z -coordinate within the solvent pore when one water molecule passed through the solvent pore (other 15 trajectories of water permeation observed during 1 μ s MD simulation are shown in Fig. 4.8). There are three characteristic positions

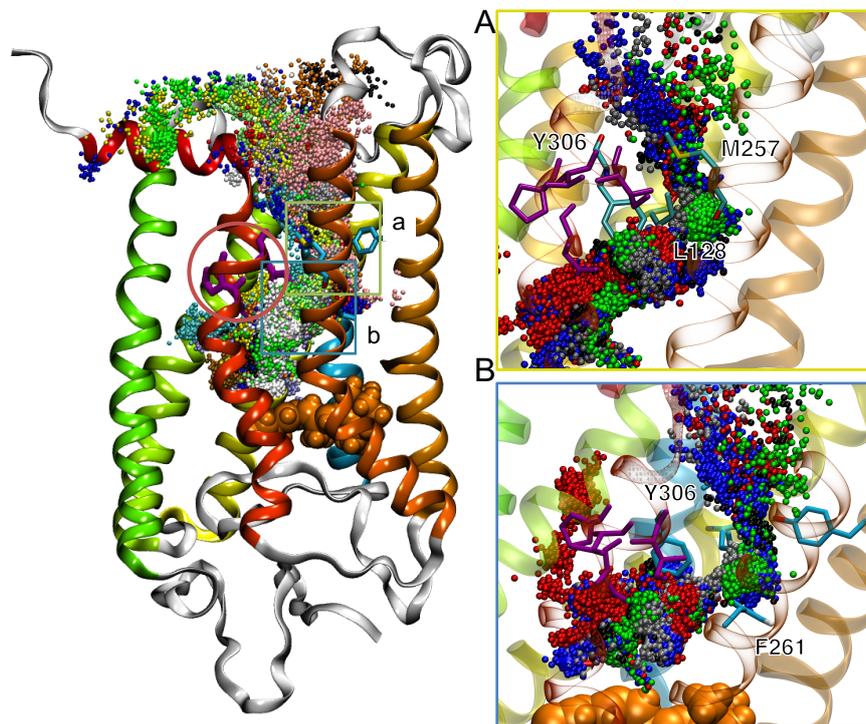


Fig. 4.5 **Configuration of the solvent pore on the cytoplasmic side with ribbon representation of the Meta II structure.** Extracted 16 trajectories of water molecules that passed through the solvent pore are shown in different colors. Residues of the NPxxY are shown as purple and in the red circle. The retinal is shown in orange VDW format. (A) Close-up view of the first narrow region and (B) the second region are shown. Only five trajectories are shown. The first narrow region comprises L128, M257, and Y306, and the second narrow region comprises F261 and Y306.

in the trajectories distinguished by the two narrow regions. The entrance of the solvent pore (less than -10 Å) is filled with bulk water, and most of the water molecules are blocked by the first narrow region which is composed of L128, M257 and Y306. The first narrow region is located between the entrance and the nearest internal hydration site (position 1). After this position 1, there is the second narrow region which is composed of F261 and Y306. The second narrow region is located between the positions 1 and 2. Internal hydration sites between the position 2 and sites around the retinal (position 3) are separated by N302 and W265. We found stable hydration sites (around -5 and -2.5 Å) in positions 1 and 2 (see Fig. 4.6B). The former water molecule has a hydrogen bond with Y306 and M257 as a donor, while the latter water molecule has a hydrogen bond with Y306 as an acceptor. Other water molecules that pass through the solvent pore are also trapped

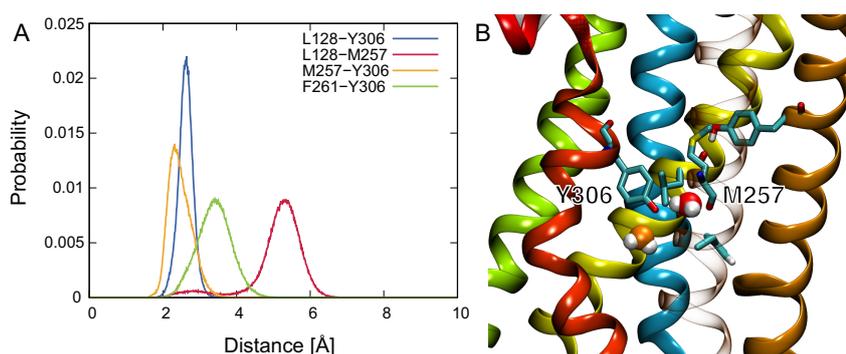


Fig. 4.6 **Characteristic regions within the solvent pore.** (A) Distance distribution of amino acids related to the narrow regions. The distance was calculated between the nearest two atoms. The first narrow region is composed of L128, M257 and Y306, and the second narrow region is composed of F261 and Y306. The distance distributions of amino acids related to the narrow regions are unimodal. (B) In hydrophobic layer, two stable hydration sites are stable during the simulation. Red water molecule has hydrogen bond with Y306 and M257 as donor (z coordinate is -5 \AA). Orange water molecule has hydrogen bond with Y306 as acceptor (z coordinate is -2.5 \AA).

in these positions (see Fig. 4.8). Because the two hydration sites are stable, we could always identify water molecules within the two hydration sites during the simulation. These water molecules have been already reported as the extended hydrogen bond network using the crystal structure [57]. Time series of the number of internal water molecules show that these internal water molecules were stable during our simulations, in particular in the Meta II state (see Fig. 4.9). The definition of internal water molecules was that water molecules within 7 \AA of some amino acids located on the center of helices, where G51, V87, F88, I123, I217, F261, V300, Y301, and N302 are chosen for the dark-adapted rhodopsin; and G51, I123, A124, L125, W126, S127, I217, and K296 are chosen for the Meta II state and Opsin. We note that the number of internal water molecules in MD simulations is generally

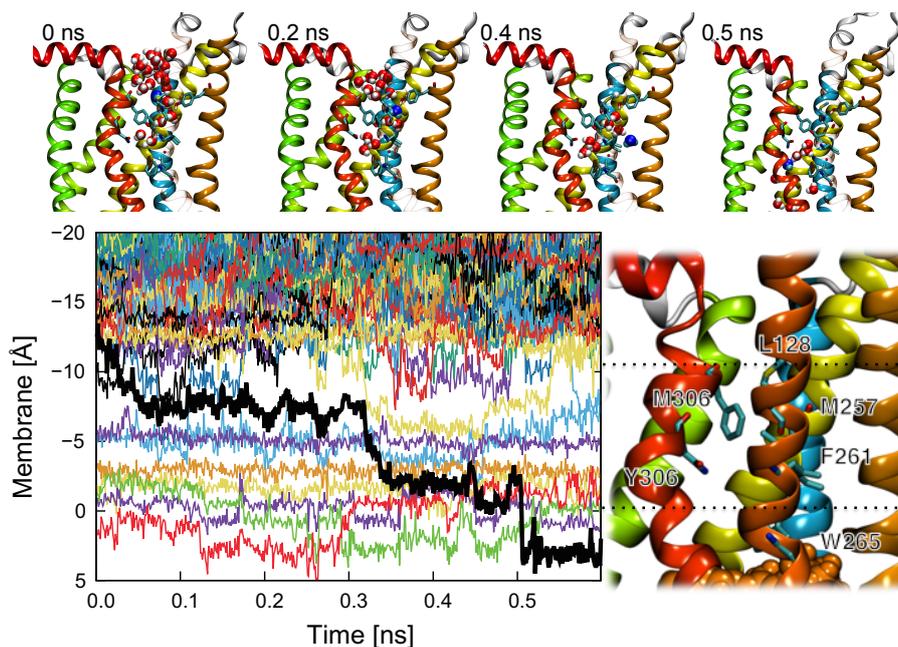


Fig. 4.7 **Water displacements in z -coordinate within the solvent pore when one water molecule passes through the solvent pore.** The trajectory of the target water molecule is shown by a thick black line. The hydrophobic region and some internal hydration sites (from -10 to 0 Å) correspond to the area between the first and second regions. The trajectory has four characteristic states and upper snapshots express each characteristic state with the target water colored as blue. Residues of two narrow gates are shown as cyan. Positions 1, 2 and 3 are located between -10 and -5 Å, -5 and 0 Å, and 0 and 5 Å, respectively.

more than that in the crystal structure, because MD simulations can consider all the internal water molecules which include dynamical water molecules. When a water molecule passes through the narrow regions, the water molecule jumps or replaces a preceding water molecule. These results show that unlike in aquaporin, where water molecules diffuse in single file, water molecules in the solvent pore diffuse by pushing or sometimes jumping a preceding water molecule due to the geometry of the solvent pore.

4.2.4 All water pathways between the bulk and the retinal-binding pocket in the Meta II state

As we mentioned above, the direction of water molecules through the solvent pore was not unidirectional. Therefore, one or more other pathways are needed for the equilibration of internal water molecules. The accessibility of water molecules (see

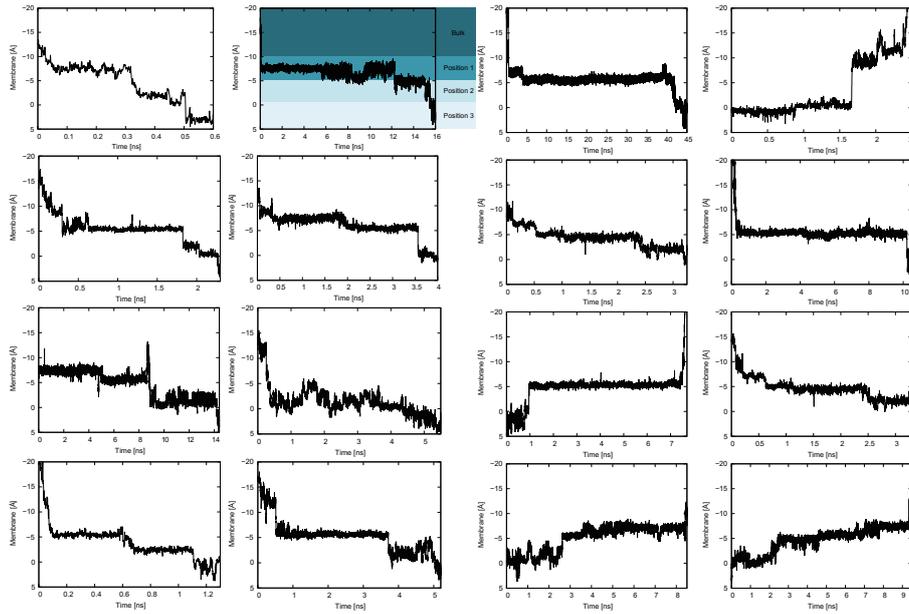


Fig. 4.8 **Trajectories of all water molecules which passed through the solvent pore in z coordinate.** The positions 1 and 2 are separated by the first narrow region which is composed of L128, M257 and Y306. The positions 2 and 3 are separated by the second narrow region which is composed of F261 and Y306.

Fig. 4.1) also indicated the presence of other pathways. Extracted trajectories of water molecules passing between the bulk and the retinal-binding pocket allowed us to observe other pathways. First, we found that a cleft between TM4 and TM5 also becomes functional as a water pathway (see Fig. 4.10). Results from a previous study using random acceleration MD simulation suggested that the biggest cleft exists between TM4 and TM5 [54]. We also observed that water molecules passed through the two ligand pores. The ligand pore between TM1 and TM7 is composed of M39, L40, Y43, M44, F91, F94, T289, and F293, and the ligand pore between TM5 and TM6 is composed of V204, I205, M207, F208, A272, F273, and F276 [56]. Two ligand pores also become functional as water pathways in the present study. A water molecule penetrated into the membrane from the cytoplasmic side after diffusing on the membrane. Since the insides of the two ligand pores are hydrophobic, the number of water molecules which passed through the ligand pores is few. There are two mechanisms for water permeation across the membrane. (i) Although membranes in our systems are stable (see Fig. 4.11), the water molecule penetrates into the membrane. Then, since the inside of the membrane is high free energy for water molecules [125], the water molecule evacuates the inside of

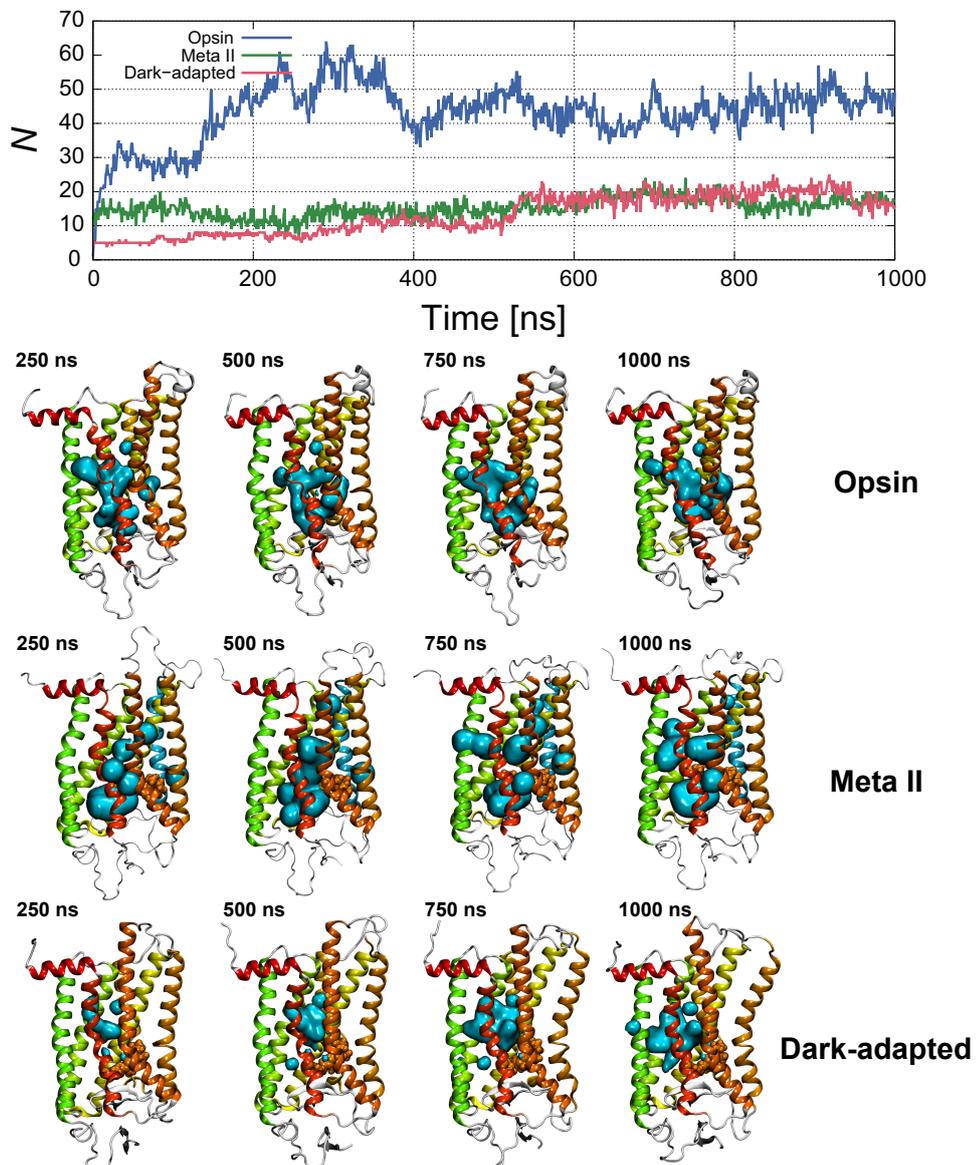


Fig. 4.9 **Times series of the number of internal water molecules in each state.** The distributions of the internal water molecules (transparent cyan surface) are shown in the lower figures. The definition of internal water molecules is that water molecules within 7 Å of some amino acids located on the center of helices.

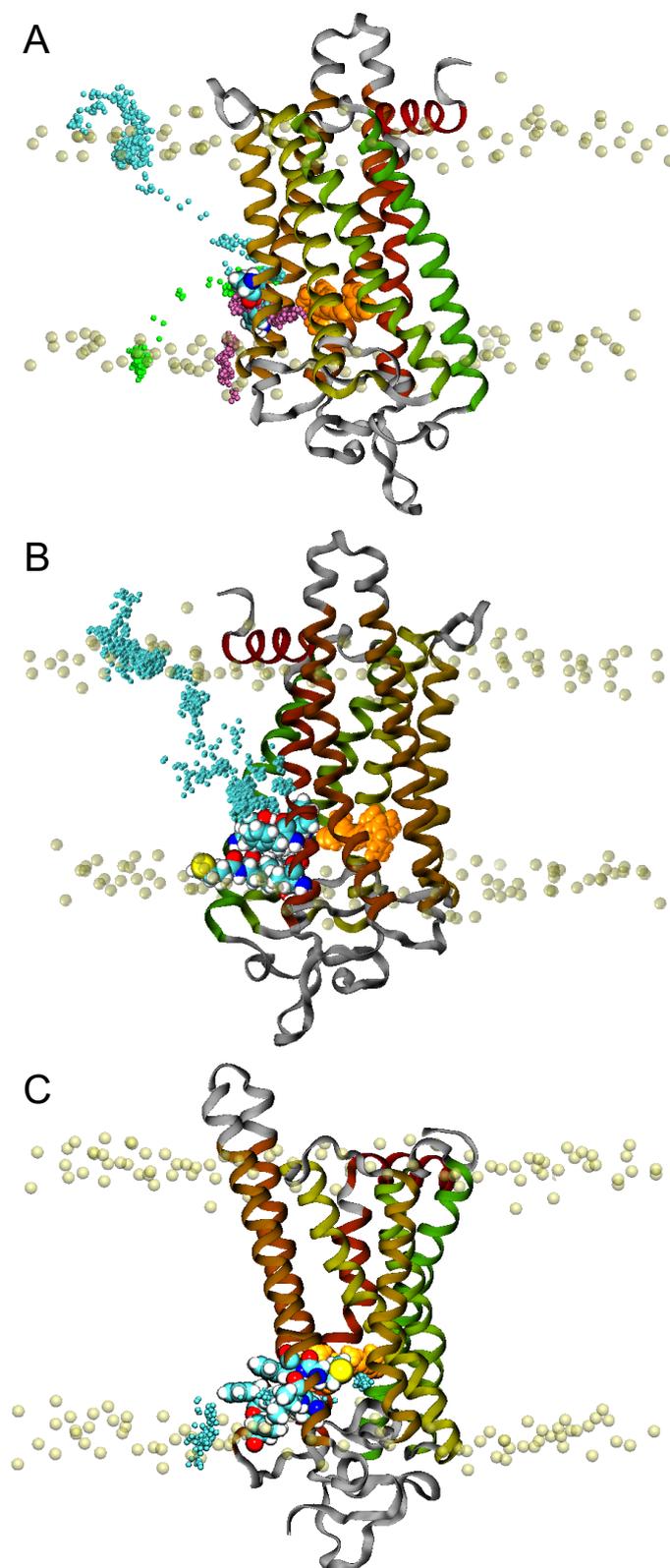


Fig. 4.10 **Other water pathways between the bulk and the retinal-binding pocket.** Blue spheres show trajectories of oxygen atoms in a water molecule every 1 ps. (A) Water pathway through the cleft between TM4 and TM5. Each color shows trajectories of different water molecules. (B) and (C) show trajectories of water molecules through the ligand pore A and ligand pore B, respectively.

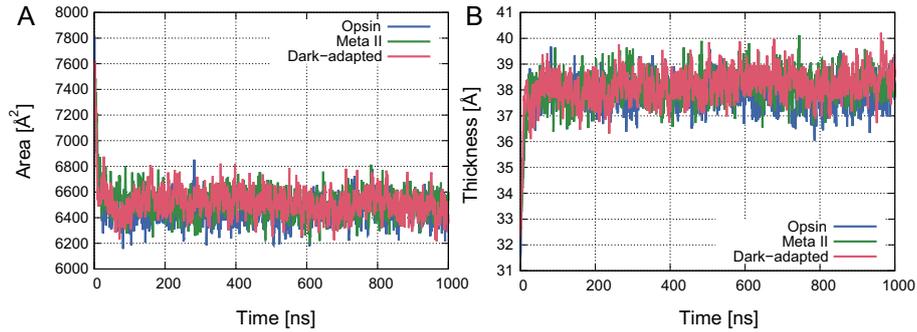


Fig. 4.11 **Stability of the simulation systems.** (A) Area (box size $L_x \times L_y$) and (B) thickness (difference between averaged z coordinates of phosphorus atoms in each leaflet) of membrane in each simulation. Membranes are stable during our simulations.

the membrane, which consequently leads to entrance into the cleft. (ii) The water molecule reaches the cleft along the interface between the protein surface and the membrane.

4.2.5 Summary

In summary, using all-atom MD simulations, we have examined the identification of the putative solvent pore on the cytoplasmic side. The solvent pore was identified in the Meta II state and Opsin, which is consistent with the experimental results. In the solvent pore, there are two narrow regions that distinguish the bulk and the internal hydration sites. The solvent pore also passes through the important hydration sites adjacent to the conserved structural motif “NPxxY”. We also revealed that unlike in aquaporin, where water molecules diffuse in single file, water molecules in the solvent pore diffuse by pushing or sometimes jumping a preceding water molecule due to the geometry of the solvent pore.

In association with conformational changes, the number of internal water molecules also changes through the intermediate states [126,127]. A previous study showed that internal water molecules contribute to the thermodynamic stability of the entire rhodopsin [68]. In particular, the internal water molecules around the retinal stabilize local electric fields by hydrating near polar and charged amino acids [62]. In this study, we also identified the stable internal water molecules around the retinal, which has been reported in crystal structure [106]. These water molecules have hydrogen bonds with E113, E181, S186, K296 and the retinal. It

was also suggested that internal water molecules contribute to functional plasticity and mediate the structural transitions from the dark-adapted rhodopsin to Opsin in all family A GPCRs [59].

Ronny *et al.* suggested that the solvent pore mediates water access [56]. We were able to show that the flow in solvent pore connects to the retinal-binding pocket, and the solvent pore also passes across the important hydration sites, such as the sites adjacent to the NPxxY motif and the extended hydrogen bond network. These results suggest two significance of the solvent pore: (1) the solvent pore provides water molecules to the retinal-binding pocket, and (2) the configuration and amount of important internal water molecules are controlled through the solvent pore.

Chapter 5

Interactions between water molecules and monosaccharide

5.1 Materials and methods

We prepared only the chair conformer as the initial structure of all the monosaccharides, and all the monosaccharides were arranged so that the molecules were more than 6.5 Å apart. We performed the MD simulations using TIP3P [115] and the modified GLYCAM06 force field [128] for water and all the monosaccharides, respectively. The modified GLYCAM06 force field improves modeling of monosaccharide aggregation compared with that of the original GLYCAM06 force field [129]. The 20, 100, and 200 monosaccharides were added to 11,070 water molecules to give monosaccharide solutions with concentrations of 0.1, 0.5, and 1 M, respectively (see Fig. 5.1).

Each system was first minimized for 5000 steps. After 1.2 ns of *NVT* constant simulation, 60 ns *NPT* constant simulations were carried out for production runs. The data for the first 12 ns were discarded because the relaxation of monosaccharide conformers occurred during this time (see Fig. 5.9A). The temperature was maintained using the Langevin thermostat with a Langevin damping coefficient of 1.0 ps⁻¹ and the Langevin piston barostat at 1 bar. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald method with a real-space cut-off of 10 Å. Van der Waals interactions were cut off at 10 Å. A time step of 2 fs was used with the SHAKE algorithm applied to constraint all hydrogen bonds (H-bonds). Periodic boundary conditions were applied to all systems.

We used six types of monosaccharides to investigate the differences between the isomers: α/β -D-glucose, α/β -D-mannose, and α/β -D-galactose. All MD simula-

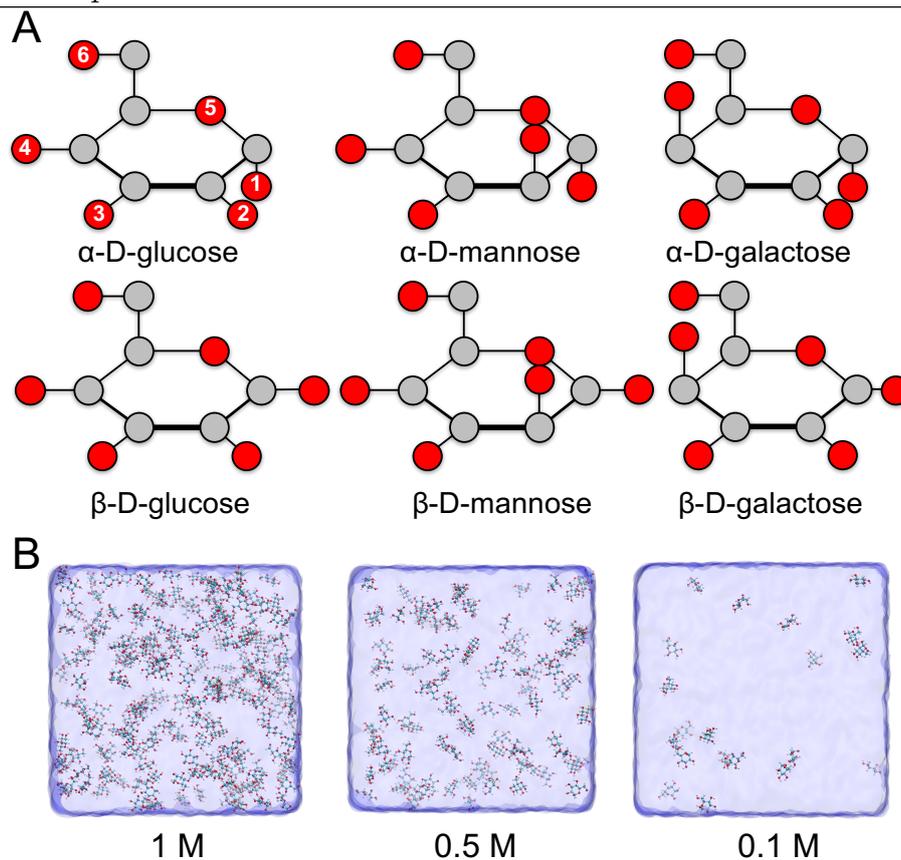


Fig. 5.1 **Monosaccharide conformation and simulation systems.** (A) Schematic of the conformations of all the monosaccharides. Oxygens are labeled with the same number as that of the bonded carbon. (B) Snapshots of 0.1, 0.5, and 1 M α -D-glucose solutions after equilibration.

tions were conducted using NAMD 2.9 software [130]. Various temperatures (298, 310, 320, 330, 340, and 350 K) and concentrations (0.1, 0.5, and 1 M) were used to explore a wide range of conditions.

5.2 Results

5.2.1 Hydration structure of monosaccharides

We analyzed the hydration structure around each monosaccharide. The RDF of water oxygen-oxygen is in good agreement with a previous study which used MD simulations and neutron diffraction with isotopic substitution experiments [87] (see Fig. 5.2). As concentration increases, the first peak of the RDF becomes stronger and broader. The SDF [131] shows the probability that water molecules exist around

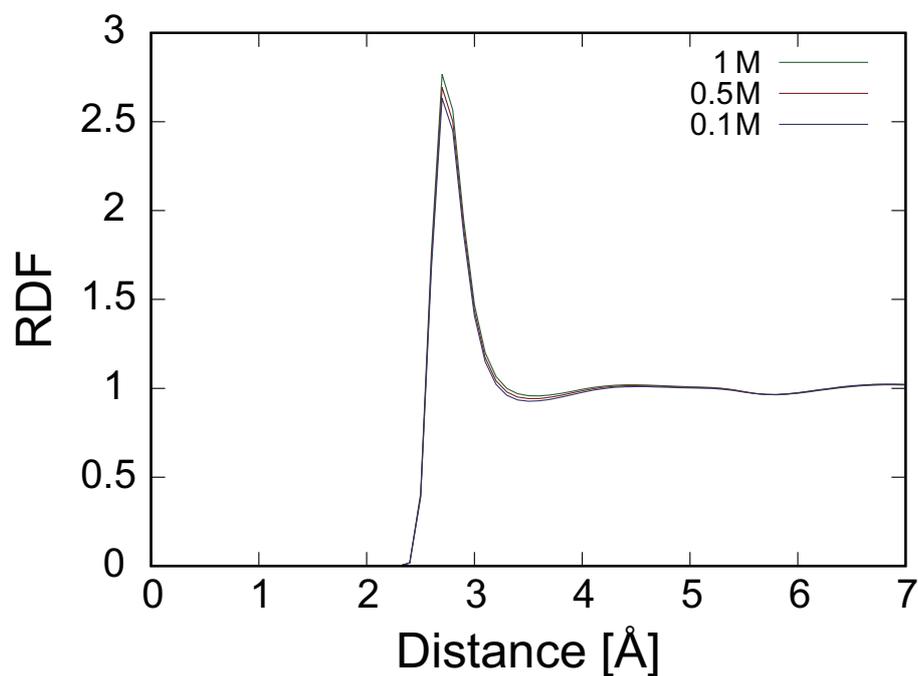


Fig. 5.2 **RDFs of water oxygen-oxygen.** Green, red, and blue lines show α -D-glucose solution at 0.1, 0.5, and 1 M, respectively.

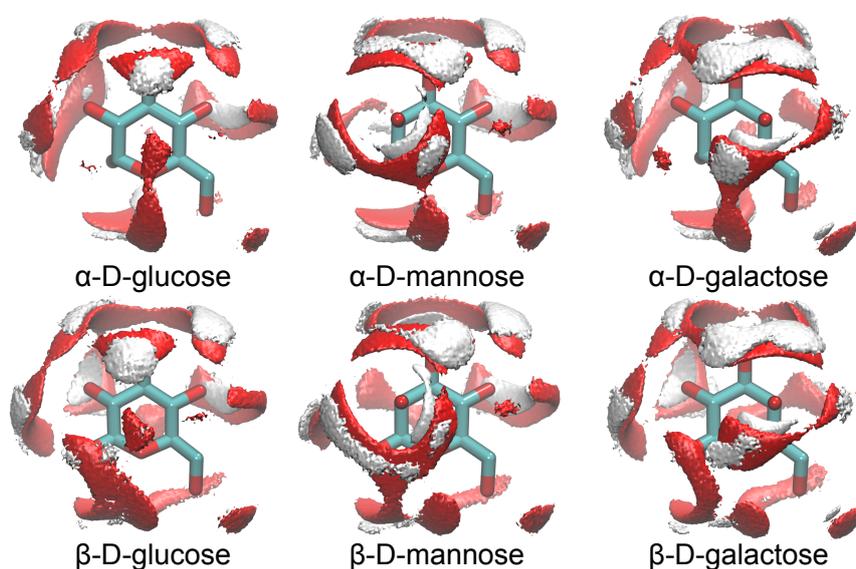


Fig. 5.3 **SDFs of water molecules around each monosaccharide.** Red and white regions represent oxygen and hydrogen atoms of water molecules, respectively. Only higher-value regions (more than 3.5 and 2.5 for oxygen and hydrogen atoms, respectively) are shown. The field is approximated on a square mesh with 0.1 Å mesh resolution.

a monosaccharide, which is calculated by

$$g(x, y, z) = \frac{1}{N} \sum_{i=1}^N \frac{\langle n_i(x, y, z) \rangle}{\rho dv}, \quad (5.1)$$

where ρ is the density of target particles, dv is the unit volume, and $n_i(x, y, z)$ is the number of target particles at the relative coordinate x , y , and z fixed to the i th monosaccharide, $\langle \rangle$ is time average, and N is the number of monosaccharides. Figure 5.3 shows the SDFs of water molecules around each monosaccharide. Using the SDFs of the oxygen and hydrogen atoms of water molecules, we clarified not only the locations but also the orientations of water molecules. The SDFs clearly showed that the orientation of monosaccharide OH groups (equatorial or axial) causes dramatic changes of the hydration shell around each monosaccharide. Around all the OH groups, there are two hydration sites where water molecules are acceptors, and one hydration site where a water molecule is a donor (see Fig. 5.4A). Most of the hydration shells overlap with those of neighboring OH groups. Because the degree of freedom of a hydroxymethyl group is higher than that of an OH group, the SDF around hydroxymethyl groups is bigger than that around OH groups (see Fig. 5.4B). Because of the bigger hydration shell around a hydroxymethyl group compared with that around each OH group, hydration shells of OH groups (O4, see Fig. 5.4A) next to a hydroxymethyl group are skewed (see Fig. 5.4C). We found that a stable hydration shell formed between axial OH groups (O2 or O4) and pyranose oxygen (O5), like a bridge. This bridged H-bond only existed in α/β -D-mannose and α/β -D-galactose because α/β -D-glucose do not have an axial OH group. Among all donor H-bonds, the bridged H-bond is the most stable (see Fig. 5.6). As temperature increases, the high-value regions of SDFs shrink uniformly in all the monosaccharide solutions (see Fig. 5.5). This result indicates that thermodynamic fluctuation makes the first hydration shell of the monosaccharides unstable. Notably, the magnitude of hydration shell shrinkage depends on the location of the hydration shell and type of monosaccharide isomer. For example, as temperature increases, the hydration shell of the hydroxymethyl group in α -D-glucose disappears. In contrast, the hydration shells of the hydroxymethyl groups in β -D-glucose and α/β -D-galactose are still stable.

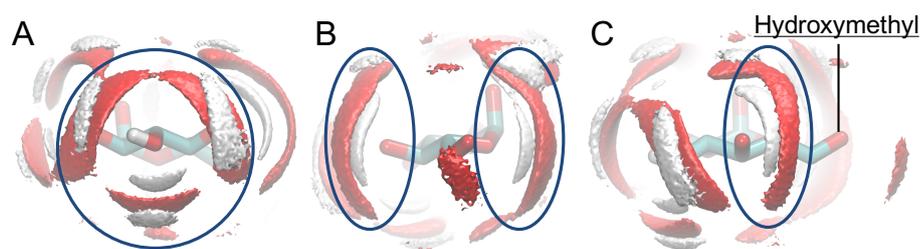


Fig. 5.4 **Typical characteristics of the hydration shell around a monosaccharide.** (A) Around all the hydroxyl (OH) groups of the monosaccharides, there are two hydration sites where water molecules are acceptors, and one hydration site where a water molecule is a donor. (B) The SDF around a hydroxymethyl group is bigger than that around a OH group. (C) Hydration shells of OH groups next to a hydroxymethyl group are skewed.

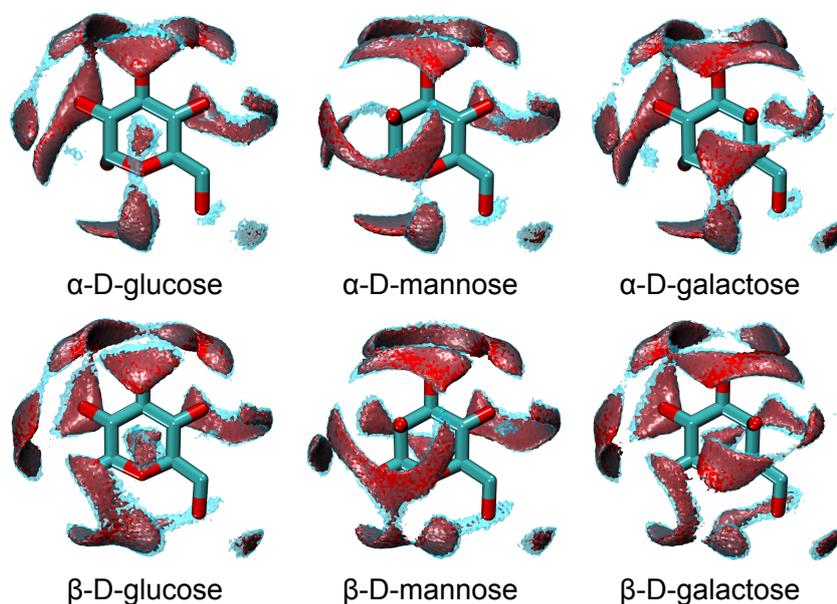


Fig. 5.5 **SDFs of the oxygen atoms of water molecules around each monosaccharide at 298 and 350 K.** Transparent cyan and glass red regions show SDFs at 298 and 350 K, respectively. Only high-value regions (more than 3.5) are shown. The field is approximated on a square mesh with 0.1 Å mesh resolution.

5.2.2 Water arrangement in the hydration shell

The stability of the hydration shells is strongly related to the status of H-bonds. In this section, we investigate the status of H-bonds using the TOP [132]. In the bulk water, each water molecule can form four H-bonds, two donors and two acceptors, and the water molecules are tetrahedrally arranged [133]. The TOP measures the

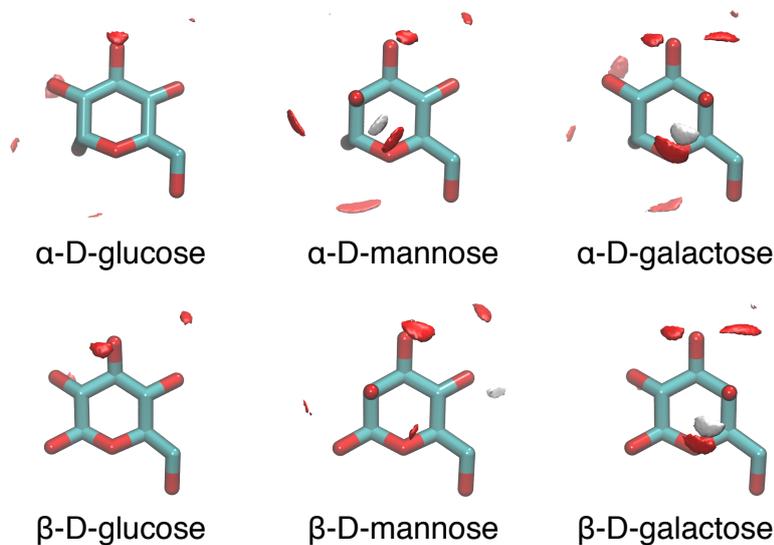


Fig. 5.6 **SDFs of water molecules around each monosaccharide.** Only higher-value regions (more than 8.5 and 7.5 for oxygen and hydrogen atoms, respectively) are shown. The field is approximated on a square mesh with 0.1 Å mesh resolution.

degree of the tetrahedral arrangement of water molecules. The TOP for a water molecule i is calculated by

$$q_i = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 (\cos(\theta_{j,k}) + \frac{1}{3})^2, \quad (5.2)$$

where $\theta_{j,k}$ is the angle between the vectors $r_{i,j}$ and $r_{i,k}$. j and k are the four nearest-neighbor oxygen atoms of the water molecule i . We calculated two types of TOPs [134], one using the nearest oxygen atoms of only water molecules (without monosaccharide), and the other using the nearest oxygen atoms of water and monosaccharide molecules (with monosaccharide). Figure 5.7A shows the TOPs of water molecules in the first hydration shell with and without monosaccharide oxygen atoms. The TOP without monosaccharide decreases near the monosaccharide because the monosaccharide perturbs the tetrahedral arrangement of water molecules. Conversely, the TOP with monosaccharide is equal to or larger than that of the bulk. These results reveal that the oxygen atoms of some water molecules are replaced with monosaccharide oxygen atoms as a component of the tetrahedral arrangement of water molecules. Figure 5.7B shows effects of temperature and concentration

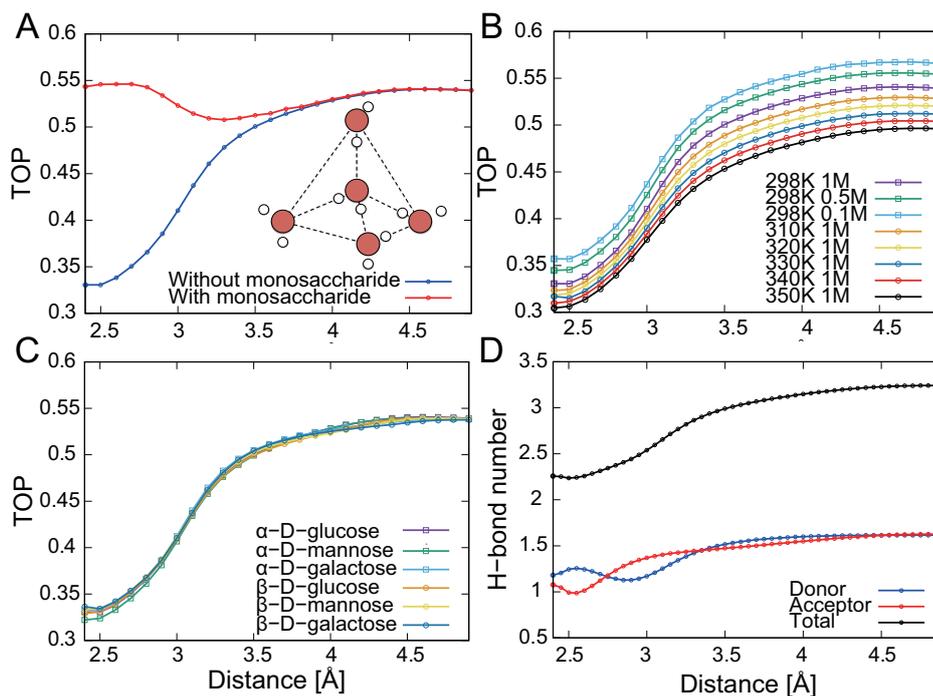


Fig. 5.7 **Water structure in the hydration shells of monosaccharides.** (A) Average TOPs as a function of the distance between monosaccharide oxygen atoms and water molecule oxygen atoms in 1 M α -D-glucose at 298 K. Red and blue show the calculation result without and with monosaccharide oxygen atoms, respectively. (B) Effects of temperature and concentration on the TOPs without monosaccharide in α -D-glucose. Squares and circles show the results for different concentrations and temperatures, respectively. (C) Effects of monosaccharide isomer type at 298 K and 1 M. (D) Average number of H-bonds between water molecules in 1 M α -D-glucose at 298 K. Blue, red, and black show the target water molecules that form H-bonds as a donor, acceptor, and either donor or acceptor, respectively.

on the TOPs without monosaccharide. In general, as temperature increases, water molecule dynamics is enhanced, which leads to a decrease of the TOPs. In the first hydration shell of monosaccharide, the TOPs also lower with rising temperature. In contrast, the TOPs increase with decreasing concentration because of the decrease in perturbation caused by the monosaccharide. The SDFs revealed the difference of hydration distribution in the various monosaccharide solutions; however, average TOPs were not effected by the type of monosaccharide isomer (see Fig. 5.7C).

When water molecules are arranged tetrahedrally, a water molecule has two donor H-bonds and two acceptor H-bonds. We calculated the distribution of the number of H-bonds, and separated the H-bonds into donor and acceptor types. H-bonds were defined geometrically as $R_{oo} \leq 3.5 \text{ \AA}$, and $\theta \leq 30^\circ$, where R_{oo} is the distance

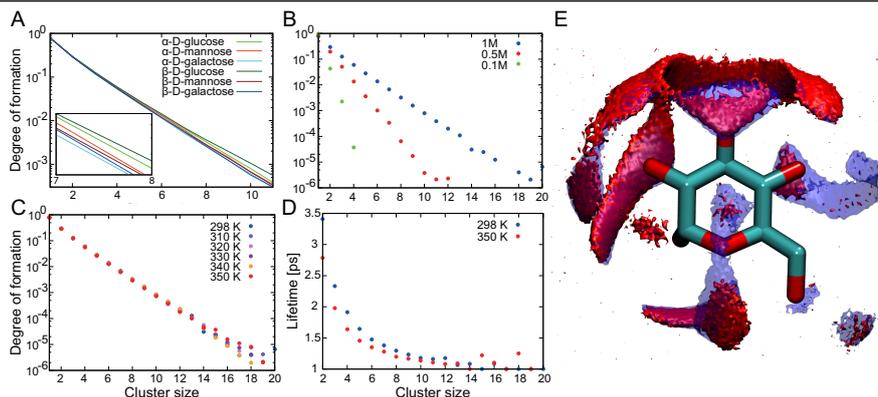


Fig. 5.8 Clustering of monosaccharides. Cluster formation is defined as when the distance between monosaccharide oxygen atoms is less than 3.5 \AA . (A) Distribution of monosaccharide clusters as a function of cluster size for each 1 M monosaccharide solution at 298 K. The inset shows a magnified view of the distribution between the cluster sizes of 7 and 8. (B) Distribution of α -D-glucose clusters as a function of cluster size for each solution concentration at 298 K. (C) Distribution of monosaccharide clusters in 1 M α -D-glucose solution at different temperatures. (D) Lifetimes of the α -D-glucose clusters as a function of cluster size at 298 and 350 K. (E) SDF of monosaccharides around a monosaccharide. Red and blue regions represent monosaccharides and oxygen atoms of water molecules, respectively. The SDF was calculated using monosaccharide oxygen atoms.

between two oxygen atoms, and θ is the angle of the H-bond [135]. Figure 5.7D shows the average number of H-bonds between water molecules as a function of the distance between monosaccharide oxygen atoms and water molecule oxygen atoms in α -D-glucose. We note that the total number of H-bonds around sugar is similar to that in the first hydration shell of various proteins [7]. As we expected, the decrease of the total number of H-bonds at less than 3.5 \AA corresponds to the decrease of the TOP without monosaccharide. In the $2.4\text{--}2.8 \text{ \AA}$ region, donor H-bonds with other water molecules are more abundant than acceptor ones. In this region, water molecules act as the acceptor in the H-bonds forming with the monosaccharide. In the $2.8\text{--}3.3 \text{ \AA}$ region, acceptor H-bonds with other water molecules are more abundant than donor ones. In this region, water molecules act as the donor in the H-bonds they form with the monosaccharide.

In the $2.4\text{--}2.8 \text{ \AA}$ region, the TOP with monosaccharide oxygen atoms is larger than or equal to that in the bulk around 0.55; conversely, the TOP with monosaccharide is smaller than that in the bulk in the $2.8\text{--}3.3 \text{ \AA}$ region (see Fig. 5.7A). These results suggest that two interesting types of H-bonds form between monosac-

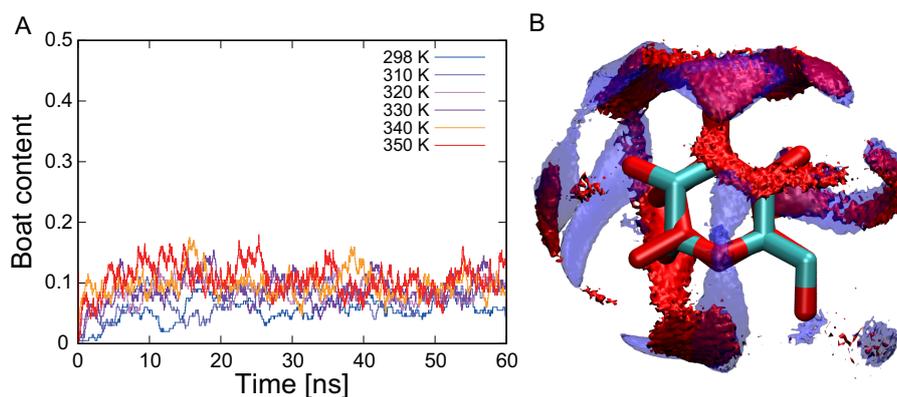


Fig. 5.9 **Conversion of monosaccharide conformers.** (A) Content of the boat conformer as a function of time. (B) SDFs of the oxygen atoms of water molecules around each conformer. Red and blue regions show the SDFs around the boat and chair conformers, respectively.

charide and water molecules: (i) when a water molecule behaves as an acceptor, the oxygen atom of the water molecule is located at 2.4–2.8 Å from monosaccharide oxygen atom, and these oxygen atoms are arranged tetrahedrally; and (ii) when a water molecule acts as a donor, the oxygen atom of the water molecule is located 2.8–3.3 Å from the monosaccharide oxygen atom, and the tetrahedral arrangement of these oxygen atoms is perturbed.

5.2.3 Clustering of monosaccharides

Clustering of monosaccharides strongly affects their hydration shell. When solutes form a cluster, the hydration shells are overlapped or depleted [9]. A cluster is defined as any array of monosaccharides such that (i) every monosaccharide is connected to at least one other monosaccharide, and (ii) two monosaccharides are said to be connected if the nearest distance between two heavy atoms is less than 3.5 Å. Existence probabilities decrease exponentially as the cluster size increases for all the monosaccharide isomers (see Fig 5.8A-C). Figure 5.8A reveals that the degree of clustering in all the solutions decreases exponentially as the cluster size increases. We found that α/β -D-glucose have the highest clustering propensity and α/β -D-galactose have the lowest of the monosaccharides. As monosaccharide concentration increases, the degree of clustering increases uniformly for all cluster sizes (see Fig. 5.8B). Figure 5.8C shows the degree of clustering at various temperatures. The results indicate that temperature does not have a large effect on clustering propen-

Table 5.1 Average lifetime, content, and free energy change of the conformers. T , C , L are temperature, concentration and lifetime, respectively

| | T [K] | C [M] | L [ns] | Content [%] | ΔG [kcal/mol] |
|-----------------------|---------|---------|----------|-------------|-----------------------|
| α -D-glucose | 350 | 1.0 | 0.96 | 10.9 | 0.633 |
| α -D-glucose | 298 | 1.0 | 3.69 | 5.99 | 0.708 |
| α -D-glucose | 298 | 0.5 | 3.46 | 6.29 | 0.694 |
| α -D-glucose | 298 | 0.1 | 3.31 | 6.61 | 0.681 |
| α -D-mannose | 298 | 1.0 | 0.63 | 4.06 | 0.813 |
| α -D-galactose | 298 | 1.0 | 6.60 | 4.18 | 0.805 |
| β -D-glucose | 298 | 1.0 | 0.24 | 0.283 | 1.51 |
| β -D-mannose | 298 | 1.0 | 0.01 | <0.01 | 2.40 |
| β -D-galactose | 298 | 1.0 | 0.19 | 0.03 | 2.08 |

sity. There are two main attributes that increase cluster probability: (i) increase of cluster stability, and (ii) increase of clustering frequency. As the monosaccharide concentration increases, increasing cluster frequency mainly contributes to the increase of cluster probability. Figure 5.8D shows the lifetime of the clusters as a function of cluster size. For all cluster sizes, the lifetime decreases as temperature increases. These results mean that increasing clustering frequency cancels out the effect of decreasing lifetime with rising temperature.

5.2.4 Boat and chair conformers

During the MD simulations, monosaccharides converted between boat and chair conformers. The chair conformer, which is the most stable, has an average dihedral angle of 60° because of the gauche interactions. The average dihedral angle of the boat conformer is much smaller than that of the chair. We defined the boat state such that: (i) the average dihedral angle is smaller than 45° ; and (ii) state (i) is stable for more than 5 ps. Our definition clearly separated two monosaccharide conformers (see Fig. 5.10). In this study, the twist-boat conformer was also considered as a boat conformer. Figure 5.9A shows the content of the boat conformer as a function of time. In the initial setup, we prepared only the chair conformer, and the content of the boat conformer increased over time. From this result, we considered that the systems were relaxed after around 10 ns. The content of conformers influences hydration properties because hydration shell is dramatically different between the boat and chair conformers (see Fig. 5.9B). Table 5.1 lists the average lifetime and average content of the boat conformer for the different monosaccharide solutions. The lifetime of the boat conformer is the order of 1-10 ns. Experimental

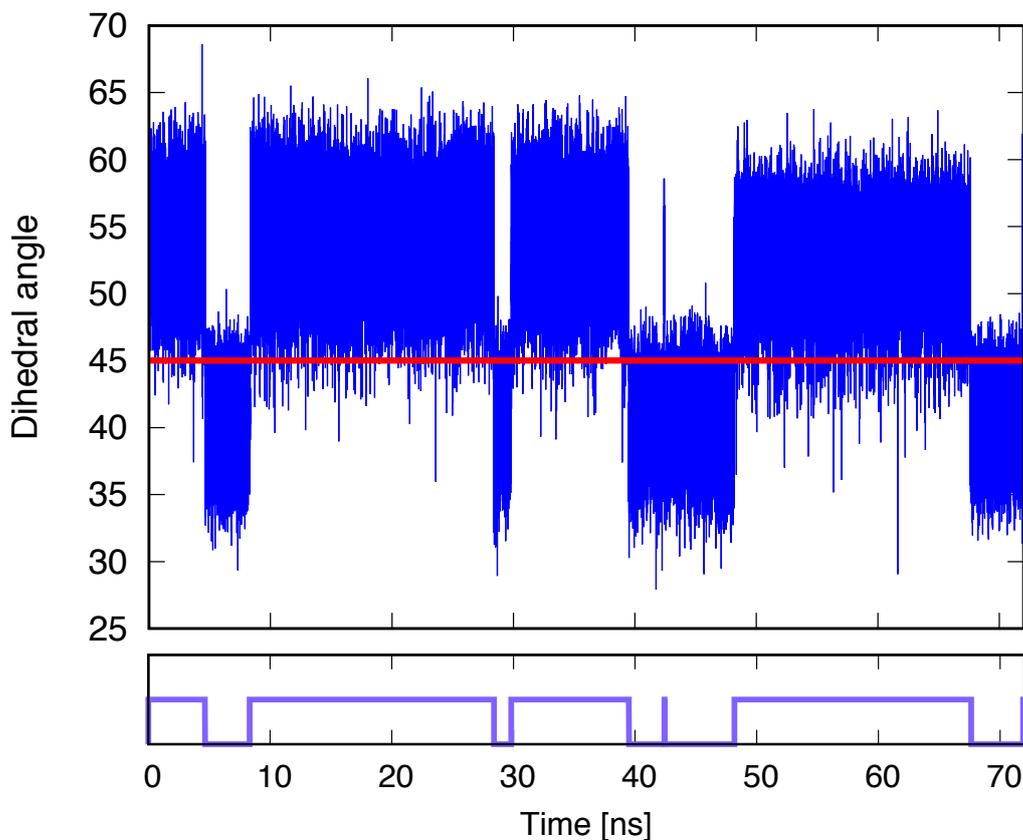


Fig. 5.10 **Definition of the boat conformer.** Upper shows dihedral angle of a monosaccharide as a function of time with the threshold 45° . Lower shows the boat states during the simulation. The results show that our definition clearly separated two monosaccharide conformers.

studies using the ultrasonic absorption spectroscopy provided valuable insights on relaxation time of sugar solutions [136,137], and relaxation time of pseudorotation is on the order of 100 ns [138]. The boat conformer is much unstable than the chair conformer; thus, we infer the lifetime of the boat conformer is much less than 100 ns. Our results show that the scale of the lifetime is in good agreement with the experimental suggestions. Moreover, a previous study using MD simulations also showed similar lifetime of the boat conformer in α -D-glucose [76].

Using the average content of the boat conformer, we calculated the free energy change between the boat and chair conformers as follows:

$$\Delta G = -RT \ln K, \quad (5.3)$$

where K is an equilibrium constant, T is temperature, and R is the gas constant. As temperature increases, the free energy change between the boat and chair con-

formers decreases, which means that the number of molecules in the boat conformer increases. These results suggest that the increased content of the boat conformer contributes to the change of the average hydration shell of the monosaccharides with rising temperature (see Fig. 5.5). The content of the boat conformer increased with decreasing monosaccharide concentration. The α -D-mannose and α -D-galactose show higher free energy changes than α -D-glucose. It is immediately apparent that the free energy change of β -D-monosaccharides is much higher than that of α -D-monosaccharides.

The difference between each monosaccharide is only the direction of OH groups; thus the differences of conformer stability between the monosaccharides are also attributed to the direction of OH groups. Figure 5.11 illustrates the location of flagpole of the boat conformer. Atoms are rendered as the average rate of being the flagpole when a monosaccharide adopts the boat conformer. The results indicated that the pair consisting of C2 and C6 has the highest probability of becoming the flagpole of the boat conformer in α -D-glucose, α -D-galactose, and β -D-galactose, while C1 and O5 pair has the highest probability of being the flagpole in α -D-mannose, β -D-glucose, and β -D-mannose. Because we considered that the twist-boat conformer was also a boat conformer, the ratio of flagpole atoms became asymmetric on the pyranose ring.

5.2.5 Summary

We performed MD simulations of monosaccharide solutions to investigate the effects of temperature, concentration, and monosaccharide isomer on monosaccharide hydration structure and water dynamics in the hydration shell. Around the monosaccharide OH groups, there are two hydration sites where water molecules act as acceptors, and one hydration site where a water molecule is a donor. We found that a stable hydration shell formed between an equatorial OH group and pyranose oxygen like a bridge only in α/β -D-mannose and α/β -D-galactose.

In the hydration shell, the water molecule network is perturbed by the monosaccharide. Our analysis revealed that some water molecules are replaced with monosaccharide oxygen atoms as a component of the tetrahedral arrangement of water molecules. According to donor/acceptor H-bond analysis, we separated this perturbation into two regions. In the 2.4–2.8 Å region, water molecules mainly form acceptor H-bonds with the monosaccharide, and the number of donor H-bonds

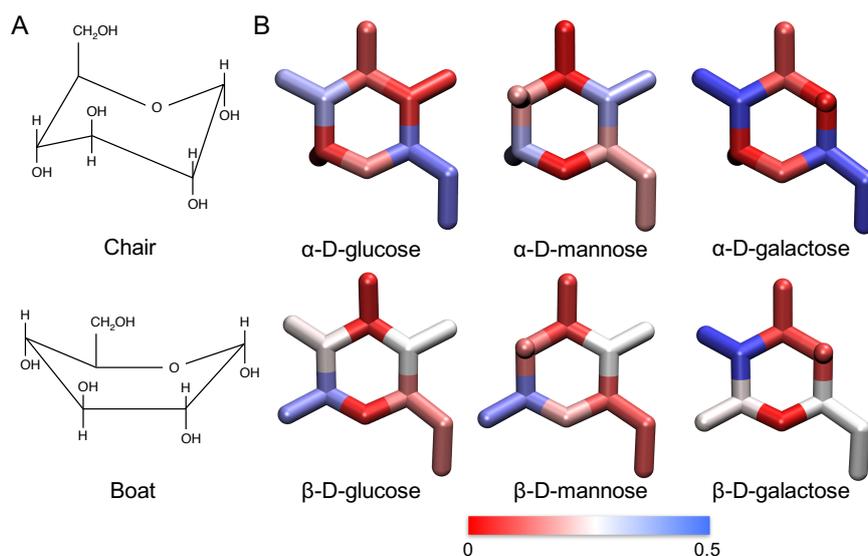


Fig. 5.11 **Location of the flagpole in the boat conformer.** (A) Chemical structure of the chair and boat conformers. (B) Atoms are rendered as the average ratio of being the flagpole when a monosaccharide adopts the boat conformer at 298 K and 1 M.

is greater than that of acceptor H-bonds. In contrast, in the 2.8–3.3 Å region, water molecules mainly form donor H-bonds with the monosaccharide. These results reveal that two types of H-bonds form between the monosaccharides and water molecules: (i) when a water molecule acts as an acceptor, the oxygen atoms of the water molecule is located 2.4–2.8 Å from the monosaccharide oxygen atoms, and these oxygen atoms are arranged tetrahedrally; and (ii) when a water molecule acts as a donor, the oxygen atom of the water molecule is located 2.8–3.3 Å from monosaccharide oxygens, and the tetrahedral arrangement of these oxygen atoms is perturbed. Our results for the clustering of monosaccharides demonstrated that this factor affects the hydration structure of the monosaccharides. In clusters, monosaccharides are distributed around the OH groups of other monosaccharides as well as water molecules. This means that water molecules are replaced with monosaccharides as their concentration increases. In all isomer solutions, we observed the same characteristics about the tetrahedral arrangement of oxygen atoms and the clustering of monosaccharide; thus, these results could be generalized to other monosaccharide isomers.

Conversion of monosaccharide conformers also strongly affects the hydration structure of the monosaccharides. We revealed that the hydration shell is dramatically different between the boat and chair conformers of the monosaccharides, and

the increased content of the boat conformer contributes to the changes of the average hydration shell with rising temperature. This study provided comprehensive information about monosaccharide solutions.

Chapter 6

Blue shift of water molecules in the first hydration shell around monosaccharide

6.1 Blue shift and dangling water molecules

The molecule behavior of water molecules at interfaces of materials plays an important role in all chemical, thermodynamic, and biological processes [1, 139, 140]. Materials are classified as either being hydrophobic or hydrophilic, according to their affinity to water. At hydrophobic interfaces, the hydrogen-bond (H-bond) network between water molecules is perturbed by the material, and consequently the H-bonds of some water molecules are broken forming a “dangling water molecule” [141]. Using either Raman spectroscopy or infrared-visible sum frequency generation spectroscopy, these dangling water molecules have been observed to exhibit a blue shift at many interfaces, such as alcohol [142, 143], tetraalkylammonium [144], fullerene derivatives [145], polyethylene glycol [146], and neopentane [147].

Hydrophilic materials, on the other hand, form stable H-bonds with water molecules. This interaction causes the spectral peak around 3200 cm^{-1} to become red shifted, as a response to change in the strength of the H-bond [148, 149]. Nevertheless, some experimental studies have reported that a blue shift was also observed in the aqueous solutions of hydrophilic materials, e.g., alkali halide ions [150] and acetonitrile [151]. These controversial phenomena demonstrated by hydrophilic materials have been attributed to a combination of both inter- and intra-molecular interactions. Although the observed blue shift has been linked to

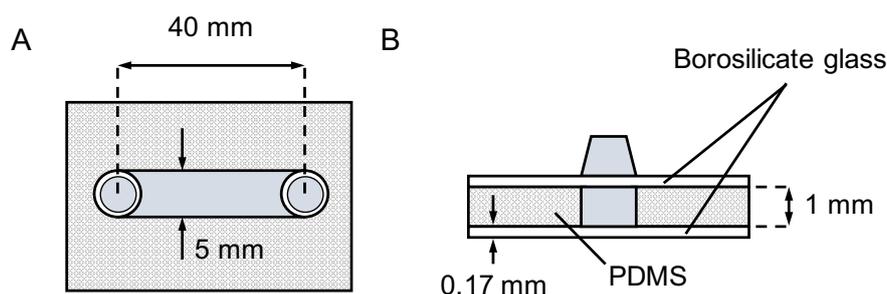


Fig. 6.1 **I-shaped channel used for Raman experiments.** (A) top view and (B) cross sectional view of the I-shaped channel are illustrated.

the presence of dangling water molecules at the interfaces, until now the exact molecular mechanism governing the blue shift has remained unclear for both hydrophobic and hydrophilic interfaces.

Here, we investigate the mechanism of the blue shift around various monosaccharides using *ab initio* and classical molecular dynamics simulations. Notably, our results show that the blue shift arises from dangling water molecules located in the first hydration shell. The vibrational state of the OH bond, which is oriented towards the bulk solvent, contributes significantly to the observed blue shift. Importantly, these observations provide an enhanced depiction of the current experimental blue shift at aqueous interfaces.

6.2 Materials and methods

6.2.1 Raman Data Collection and Multivariate Curve Resolution Methods

In this study, Raman spectra of the D-glucose solutions at 0, 0.05, 0.1, 0.3, and 0.5 M were collected. These experiments were conducted using an I-shaped channel (the detail is shown in Figure 6.1). The top and bottom walls of the channel were made of borosilicate glass, and the side wall was made of polydimethylsiloxane (PDMS). The depth and width of the channel was set to 1 mm and 5 mm, respectively. The measurement point was midpoint of the channel.

All the Raman spectra were collected using the Raman spectrometer Invia Raman Microscope (RENISHAW) equipped with the Nd:YAG laser (532 nm) and DM2500 (LEICA) model microscope. The experiments were coupled with a CCD camera (operating with 1024×256 pixel) mounted to a 250 mm focal length imaging spec-

trograph with a 1800 l/mm grating. The laser beam was focused on the samples with a 50 \times objective lens and a numerical aperture (NA) of 0.75 to the spot of 0.9 μm . The average laser excitation power was 150 mW, and exposure time of the camera was set to be 5 s. 10 images were accumulated as one spectrum and this procedure was repeated three times at each concentrations.

The self-modeling curve resolution (SMCR) algorithm was performed as a multivariate curve resolution method. In the pre-processing phase, almost zero signals in the spectral regions ($\sim 2800\text{ cm}^{-1}$ and $\sim 3900\text{ cm}^{-1}$) were subtracted, and the input Raman spectra were normalized by their area. Then, we concatenated all the spectra into a data matrix ($n \times m$) containing n spectra of the length m , and the SMCR algorithm was applied to the data matrix.

6.2.2 *Ab Initio* Molecular Dynamics Simulations

Ab initio molecular dynamics simulations were performed by the CPMD code [152] using the Car-Parrinello method [153]. Before the CPMD simulations, we performed classical MD simulations to obtain initial structure of the systems. The MD simulations and energy minimizations were conducted using the NAMD 2.9 software [107] with the GLYCAM 06 [154] and TIP3P water model [115]. Periodic boundary conditions were applied in all three directions of the cubic box. The last length of the *NPT* MD simulation was used to determine the side length of cubic box in the CPMD simulation. Furthermore, the initial structures in the CPMD simulation were taken from the last structures of 10 ns MD simulations.

In the CPMD simulations, our systems were treated according to the density function theory using the Perdew-Burke-Ernzerhof functional for the approximation of the exchange-correlation terms [155]. The core electrons were treated using norm-conserving pseudo potentials with the Kleinman-Bylander separation scheme for carbon, hydrogen, and oxygen [156]. We described the valence-core interaction using the Martins-Troullier pseudo potentials [157]. Plane wave functions were used for the basis of the calculations and set with an energy cutoff of 80 Ry. The systems were composed of a monosaccharide and a set number of water molecules. Specifically, the monosaccharide α and β -D-glucose/D-mannose/D-galactose were paired with 108, 108, 101, 107, 103, and 103 water molecules, respectively. The simulation time step was 0.097 fs in the *NVT* ensemble using Nosé-Hoover thermostat, and the fictitious electron mass was set as 0.22 atomic mass unit. Each simulation was performed for

60 ps, and the analyses were conducted over the last 52 ps trajectories. We also performed the CPMD simulation of NaBr solution with the same parameters and conditions as the monosaccharide solutions. The system contains a Na^+ , Br^- , and 100 water molecules. The simulation was performed for 96.6 ps, and the analyses were conducted over the last 92.6 ps trajectories.

Classical MD simulations of the α -D-glucose solution with TIP3P [115] and TIP5P [158] were performed using the NAMD-2.9 [130] and GROMACS-5.1.2 [116] softwares, respectively. We performed the MD simulations using the modified GLYCAM06 force field [128] for α -D-glucose. The modified GLYCAM06 force field improves the modeling of monosaccharide aggregation compared with that of the original GLYCAM06 force field [129]. For the simulation we chose a 0.1 M α -D-glucose system composed of 20 α -D-glucose and 11,070 water molecules. *NPT* constant simulations at 298 K and 1 bar were carried out for 200 ns, with the data from the first 10 ns being discarded due to the relaxation. Periodic boundary conditions were applied to all systems. In the simulation using the NAMD software, the temperature was maintained using the Langevin thermostat with a Langevin damping coefficient of 1.0 ps^{-1} and the Langevin piston barostat. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald method with a real-space cut-off of 10 Å. Van der Waals interactions were cutoff at 10 Å. A time step of 2 fs was used for the SHAKE algorithm that was applied to constraint all H-bonds. In the simulation using the GROMACS software, the pressure and temperature of the system were controlled using Parrinello-Rahman barostat [117] and velocity-rescaling method [118], respectively. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald method with a real-space cutoff of 12 Å. Van der Waals interactions were cut off at 12 Å. A time step of 2 fs was used for the linear constraint solver (LINCS) algorithm applied to constrain all H-bonds [159].

6.3 Methods for analysis

6.3.1 Frequency Calculations: Fourier and Wavelet Transforms

Vibrational spectra were obtained from the vibrational density of states (VDOS) [160],

$$P(\omega) = \sum_{k=1}^N \int_{-\infty}^{\infty} \langle \dot{r}_k(0) \dot{r}_k(t) \rangle e^{i\omega t} dt, \quad (6.1)$$

where $\langle \dot{r}_i(0) \dot{r}_i(t) \rangle$ is the time-series velocity auto-correlation function. The velocity auto-correlation function was calculated for 2.8 ps and averaged over all sugar-hydroxyl groups and simulation times. The Hann window function was employed for filtering before the Fourier transform [161] and the Loess method was used for smoothing in the MATLAB package with a span of 0.015 %.

The time dependent vibrational spectra were calculated by wavelet analysis [162],

$$L_{\psi} f(a, b) = a^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (6.2)$$

where a and b are the parameters, and the time series of OH relative velocities were used as $f(t)$. The parameter b shifts a window at $t = b$, and the parameter a determines the window scale and is directly related to frequency. The wavelet function provides a frequency at time $t = b$. In previous studies, wavelet analysis was successfully applied for spectral dynamics of water molecules in ion solutions [163–165]. In this study, the Morlet-Grossman function [166] was chosen as the mother wavelet $\psi(t)$,

$$\psi(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{2\pi\lambda it} e^{-\frac{t^2}{2\sigma^2}}, \quad (6.3)$$

where λ and σ are parameters. Following a previous study [162], the values 1 and 2 was set for λ and σ , respectively.

6.3.2 Tetrahedral Order Parameter

The TOP represents the degree of tetrahedral arrangement of water molecules. The TOP for a water molecule i is calculated by

$$q = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 \left(\cos(\theta_{j,k}) + \frac{1}{3} \right)^2, \quad (6.4)$$

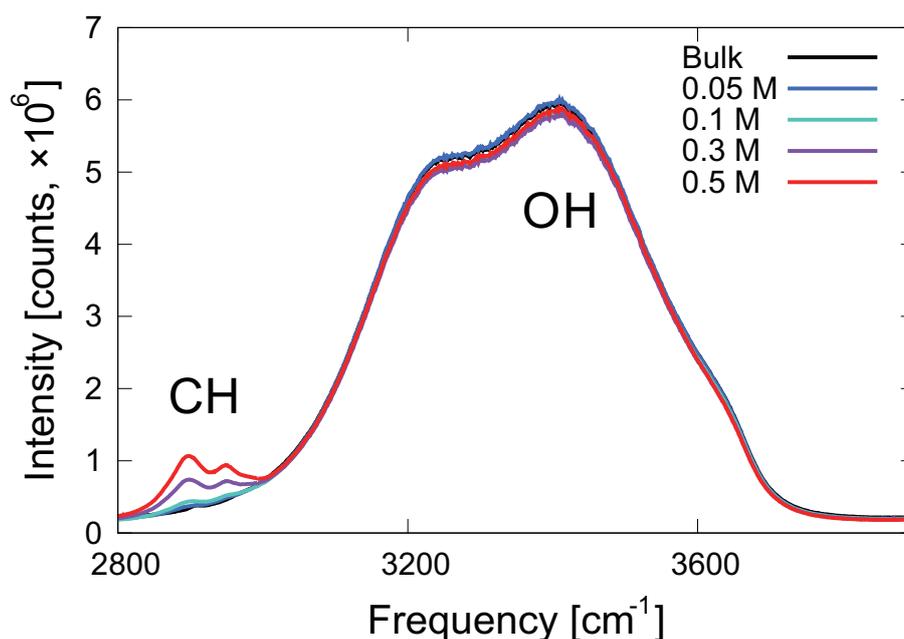


Fig. 6.2 Raman spectra of D-glucose solutions at concentrations between 0 and 0.5 M.

where $\theta_{j,k}$ is the angle between the vectors $r_{i,j}$ and $r_{i,k}$. The terms j and k represent the four nearest-neighbor water molecule oxygen atoms of the original water molecule i .

6.4 Results

6.4.1 Hydration-shell spectra of monosaccharides

We applied multivariate curve resolution (MCR) to Raman spectra of D-glucose solutions at concentrations between 0 and 0.5 M (see Fig. 6.2). Our Raman-MCR results were obtained using the self-modeling curve resolution (SMCR) algorithm (see Materials and Methods for more details). The SMCR algorithm is used to decompose spectra into bulk-water-correlated and solute-correlated (SC) components. Figure 6.3A shows the SC and bulk-water spectra. The peak at 2900 cm^{-1} denotes the monosaccharide CH groups, and the two overlapping peaks at 3200 cm^{-1} and 3400 cm^{-1} correspond to H-bonded OH groups [143]. The OH peaks in the SC spectra arise from the monosaccharide itself and water molecules in the hydration shell. However, a previous study showed that the SC spectra mostly contains information arising from water molecules in the hydration shell, by using various aqueous

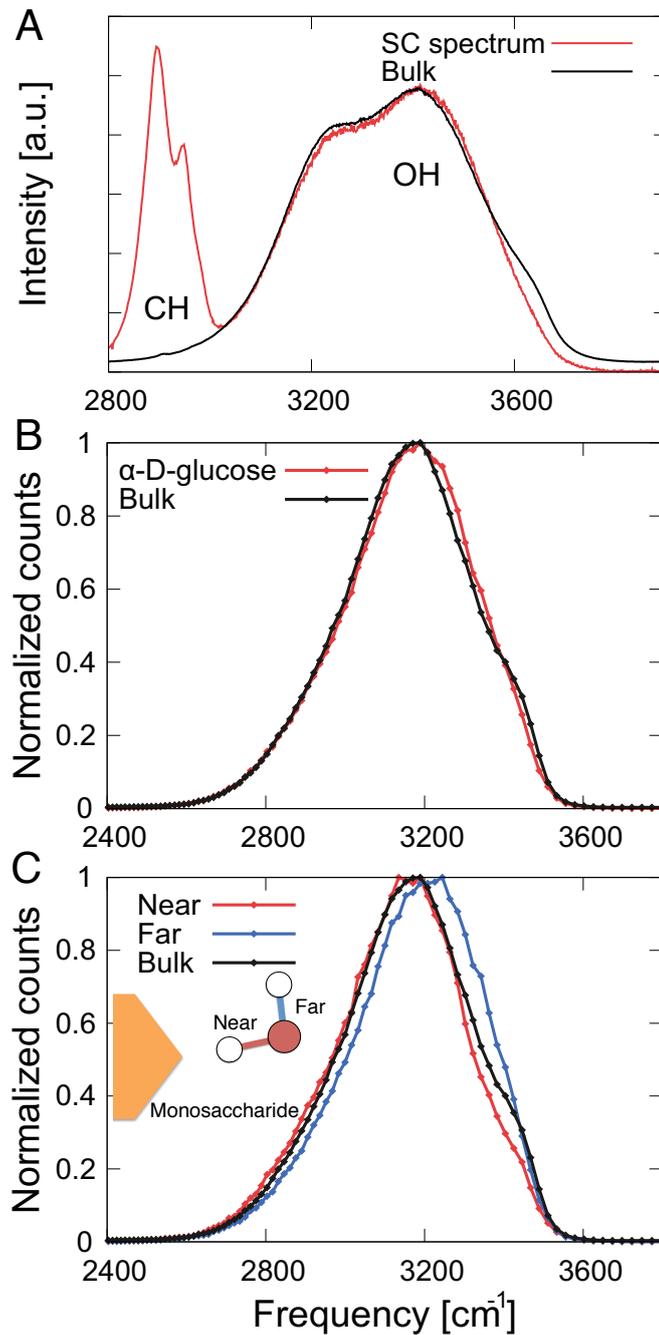


Fig. 6.3 **Spectra of water molecules in D-glucose solution.** (A) Experimental Raman spectrum of the bulk (black line) and SC spectrum (red line). The SC spectrum was obtained using the spectra of D-glucose solutions in 0, 0.05, 0.1, 0.3, and 0.5 M. (B)(C) The OH vibrational spectra of water molecules obtained by the CPMD simulations using the wavelet analysis; (B) within 3.5 Å of α-D-glucose (red line), and (C) “near” (red line) and “far” (blue line) OH bonds. The “near” and “far” represent OH bonds from the monosaccharide oxygen atoms.

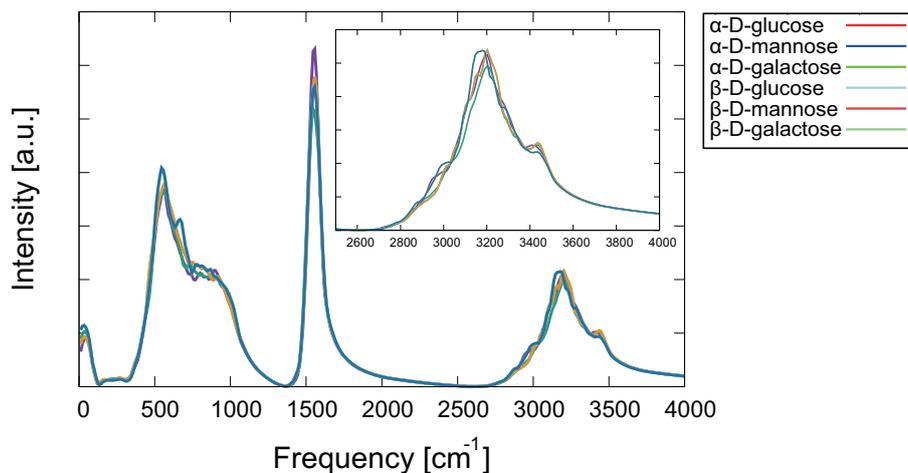


Fig. 6.4 **OH vibrational spectra for all water molecules in the system obtained by CPMD simulations.** The spectra were obtained by the Fourier transform of the OH relative velocity correlation function. The Loess method was used for smoothing. The results are in good agreement with the water spectra obtained experimentally

alcohols [143]. Moreover, the OH peaks of the SC spectrum were as broad as those of the bulk-water spectrum, which provides evidence that the peaks were correlated to water molecules in the hydration shell of monosaccharide. Compared with the bulk-water spectrum, blue shift of the OH peaks was observed in the SC spectrum.

To reveal molecular mechanism of the blue-shift, we also calculated OH vibrational spectra around monosaccharide using the *ab initio* MD simulations (see Fig 6.4). Most hydration shells in aqueous monosaccharide solutions are known to be distributed around monosaccharide oxygen atoms [167, 168]. Therefore, in this study, we focused on the hydration shells around all oxygen atoms present in the monosaccharide. Figure 6.3B shows the OH vibrational spectra obtained by the wavelet analysis (see Materials and Methods for more details) of the water molecules that are within 3.5 Å of the monosaccharide oxygen atoms. Notably, we also observed blue-shift of OH peak in hydration shell of monosaccharide, which is in agreement with the result of SC Raman spectrum. In a water molecule, the OH bond that is H-bonded with the monosaccharide is classified as the “near” OH bond, and the opposing OH bond known as the “far” OH bond. We observed a large blue shift in the “far” OH vibrational spectra, however, no blue shift was observed in the spectra of the “near” OH vibrational spectra (see Fig. 6.3C). Presence of this blue shift around all six isomers of the monosaccharide suggests that blue shift is

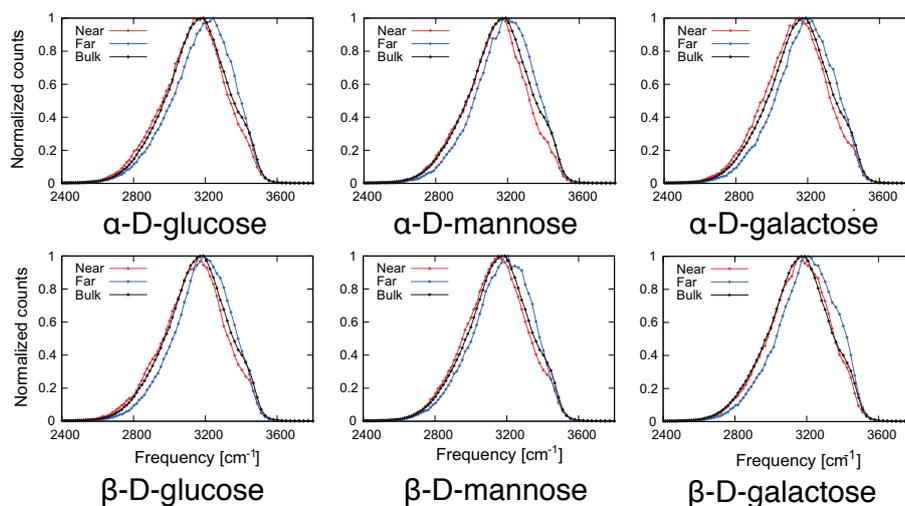


Fig. 6.5 **OH vibrational spectra of water molecules within 3.5 Å of all isomers.** The spectra were obtained by the wavelet analysis. Red and blue lines show “Near” and “far” OH bonds from monosaccharide oxygen atoms.

a fundamental phenomenon found at the interfaces of many cyclic compounds (see Fig. 6.5).

6.4.2 The relationship between water arrangement and the observed blue shift

In liquid water, the water molecules are tetrahedrally arranged and each water molecule has four H-bonds: two donor and two acceptor H-bonds [133]. To understand the correlation between the observed blue shift and the water molecule network, first we calculated the TOP [134], which represents the degree of tetrahedral arrangement of water molecules. Two forms of the TOP were calculated: (i) the TOP between water and sugar oxygen atoms, and (ii) the TOP between only water oxygen atoms (see Materials and Methods for more details). Figure 6.6A shows the TOPs of water molecules as a function of the distance between oxygen atoms of monosaccharide and water molecules. In the first hydration shell, the TOP between only water oxygen atoms is lower than that of the bulk solution because the tetrahedral arrangement of these water molecules is perturbed by the monosaccharide. On the other hand, the TOP between water and sugar oxygen atoms is equal to or greater than that of the bulk solution, which means that the sugar oxygen atoms are involved in the tetrahedral arrangement of water molecules. At a distance of 4 Å, both the TOPs converge to 0.8 corresponding to the value for the

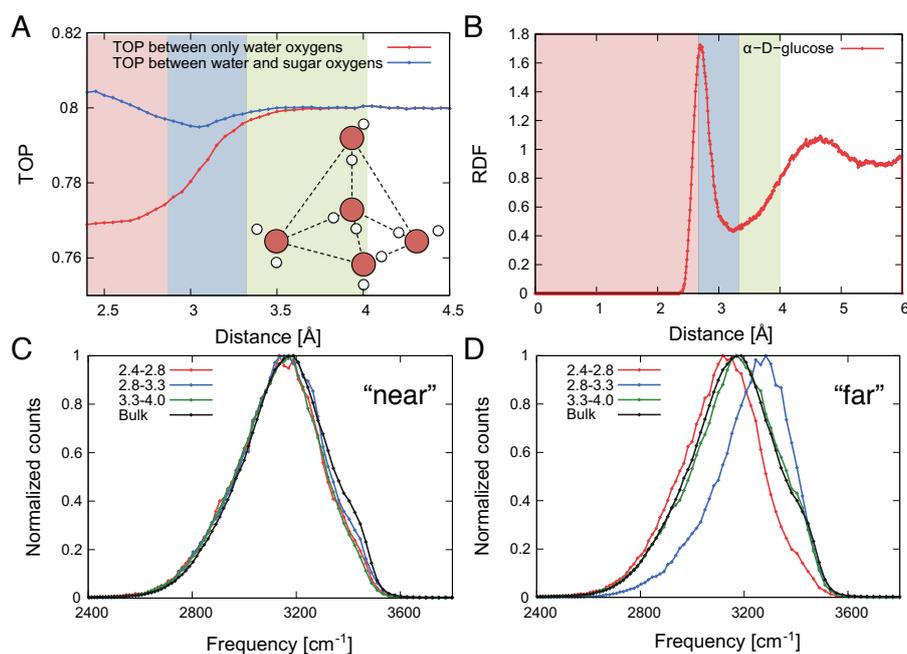


Fig. 6.6 **Relationship between water arrangement and vibration in α -D-glucose solution.** (A) Average tetrahedral order parameters as a function of the distance between oxygen atoms of monosaccharide and water molecules. Red and blue show the calculation without and with monosaccharide oxygen atoms, respectively. (B) Radial distribution function as a function of the distance between oxygen atoms of monosaccharide and water molecules. (C and D) “near” and “far” OH vibrational spectra of water molecules separated by the characteristic regions in the TOP between water and sugar oxygen atoms. Red, blue, and green colored lines are the spectra using water molecules located in 2.4–2.8, 2.8–3.3, and 3.3–4.0 Å, respectively.

bulk solvent.

Next, we separated the hydration shell into three regions according to the TOP values between water and sugar oxygen atoms: (i) the TOP is equal to or greater than that in the bulk (near_{TOP} region; the water molecule oxygen is located between the 2.4–2.8 Å), (ii) the TOP is lower than that in the bulk ($\text{middle}_{\text{TOP}}$ region; the water molecule oxygen is located between 2.8–3.3 Å), and (iii) the TOP is equal to that in the bulk (far_{TOP} region; the water molecule oxygen is located between 3.3–4.0 Å). To clarify the relationship between these regions and the hydration shell depth, we calculated the RDF between the water and monosaccharide oxygen atoms (see Figure 6.6B). The boundary between the near_{TOP} and $\text{middle}_{\text{TOP}}$ regions is represented by the first peak of the RDF at a distance of 2.8 Å. The boundary between the $\text{middle}_{\text{TOP}}$ and far_{TOP} regions is depicted by the first minimum of the

RDF, at the distance of 3.3 Å. We demonstrated that the first hydration shell is the most perturbed by the monosaccharide at a distance of 2.8–3.3 Å, i.e. in the middle region.

Figures 6.6C and 6.6D show the “near” and “far” OH vibrational spectra of water molecules in each of the above three regions. In the near_{TOP} region, water molecules form a cooperative H-bond network with other water molecules and monosaccharide oxygen atoms. Due to this stable network, the “far” OH vibrational spectra are shifted to lower frequencies (red-shift). Notably, in the $\text{middle}_{\text{TOP}}$ region, the “far” OH vibrational spectra are clearly shifted to higher frequencies (blue shift). As we showed above, the weak H-bonded water molecules in the $\text{middle}_{\text{TOP}}$ region contribute to this blue shift, which is similar to the observation at hydrophobic interfaces [144].

6.4.3 Distribution of the dangling water molecules

We performed simulations for six isomers of a monosaccharide, which differ only in the orientation of their hydroxyl groups. Using the SDF, we visualize the probability density of the water molecules around the monosaccharides (see Fig. 6.7). Due to the different orientation of the hydroxyl groups, distribution of the hydration shell is unique to each isomer. The weak H-bonded water molecules are distributed throughout this hydration shell, with all the hydroxyl groups encompassing the water molecules. Thus, the blue shift was observed in all the monosaccharide solutions. Furthermore, these results suggest that weak H-bonded water molecules can be found at the interfaces of many other cyclic compounds.

To further investigate the mechanism of the blue shift, we analyzed the H-bond status (donor or acceptor) in the different regions of the first hydration shells. To achieve a large ensemble for this analysis, we additionally performed classical MD simulations. Figure 6.8A shows the average number of H-bonds between water molecules, and the formation process of the observed weak H-bonded water molecules around the monosaccharides is schematically illustrated in Fig. 6.8B. Importantly, the regions defined by the TOP using the CPMD are in good agreement with results from the classical MD simulations. In the near_{TOP} region (2.4–2.8 Å), donor H-bonds between water molecules are more abundant than acceptor H-bonds, meaning that the two hydrogen atoms are largely oriented towards the bulk solvent (observed as the red region in Fig. 6.8B). The H-bonds of these water molecules are

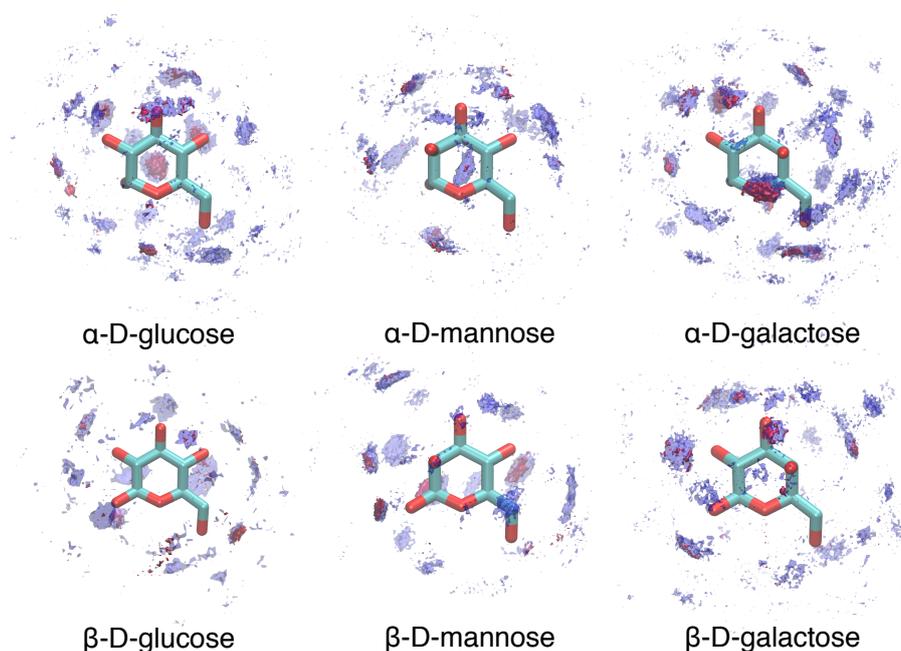


Fig. 6.7 **Density of water molecule oxygen atoms around all the monosaccharides.** Blue colored regions represent the first hydration shells, and red colored regions represent the location of dangling water molecules, which are defined as water molecules located in the middle_{TOP} region with low value of the TOP between water and sugar oxygen atoms.

very stable, and essentially, the water molecules are compatible with the structure of both the monosaccharide and the bulk solvent. Conversely, in the middle_{TOP} region (2.8–3.3 Å), acceptor H-bonds between water molecules are more abundant than donor H-bonds, meaning that the water molecules predominantly act as a donor to the monosaccharide (observed as the blue region in Fig. 6.8B). These results are in agreement with the CPMD result that the “far” OH exhibits a blue shift, whereas the “near” OH exhibits a red shift (see Figs. 6.6C and 2D). We confirmed these results had no dependence on the water model of classical MD simulation and H-bond definitions (see Fig. 6.9).

From these results, we propose a mechanism of the formation process of the first hydration shell around the monosaccharide. First, the closest water molecules (in the region of 2.4–2.8 Å) primarily form acceptor H-bonds with the monosaccharide, and these water molecules are as stable as the bulk solvent. Secondly, empty spaces in the first hydration shell are filled with other water molecules (in the region of 2.8–3.3 Å). In this region, water molecules mainly form donor H-bonds with the monosaccharide. These inner H-bonds of the water molecules are strongly attracted

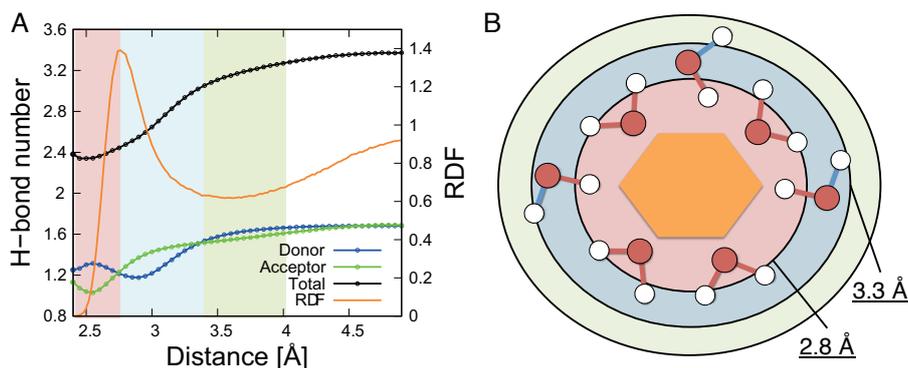


Fig. 6.8 **Formation process of the observed dangling water molecules around the monosaccharide.** (A) Average number of H-bonds between water molecules and RDF as a function of the distance between oxygen atoms of monosaccharide and water molecules. These results were obtained by the classical MD simulations of α -D-glucose solution at 298 K and 1 M using TIP3P water model. Blue, red, and black represent the number of donor, acceptor, and total H-bonds, respectively. (B) Schematic view of the formation process of the dangling water molecules. Red region shows 2.4–2.8 Å where water molecules mainly form acceptor H-bonds with monosaccharide. These water molecules are compatible with structure of both the monosaccharide and the bulk. Blue region shows 2.8–3.3 Å where water molecules mainly acts as a donor with monosaccharide. In this regions, H-bonds with outer water molecules are unstable.

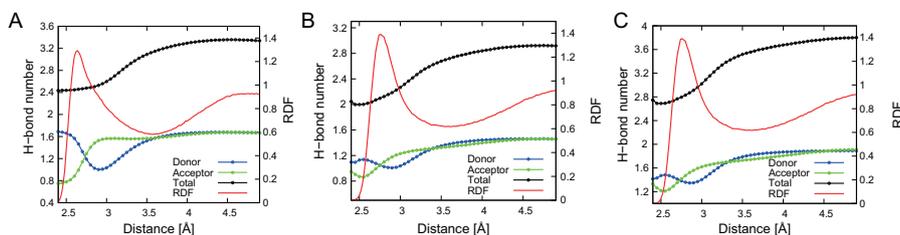


Fig. 6.9 **Average number of H-bonds between water molecules and RDF as a function of the distance between oxygen atoms of monosaccharide and water molecules.** (A) These results were obtained by the classical MD simulations of α -D-glucose solution at 298 K and 0.1 M using TIP5P water model. (B)(C) Different H-bond criteria using TIP3P water model with (B) $R_{oo} \leq 3.3 \text{ \AA}$, and $\theta \leq 25^\circ$ and (C) $R_{oo} \leq 3.7 \text{ \AA}$, and $\theta \leq 35^\circ$.

and perturbed by the monosaccharide; thus, the outer H-bonds between other water molecules become relatively unstable. The unstable H-bond causes the dangling motions, and these water molecules exhibit the observed blue shift.

From these results, we propose a mechanism of the formation process of the first hydration shell around the monosaccharide. First, the closest water molecules (in

the region of 2.4–2.8 Å) primarily form acceptor H-bonds with the monosaccharide, and these water molecules are as stable as the bulk solvent. Secondly, empty spaces in the first hydration shell are filled with other water molecules (in the region of 2.8–3.3 Å). In this region, water molecules mainly form donor H-bonds with the monosaccharide. These inner H-bonds of the water molecules are strongly attracted and perturbed by the monosaccharide; thus, the outer H-bonds between other water molecules become relatively unstable. The unstable H-bond causes the dangling motions, and these water molecules exhibit the observed blue shift.

6.4.4 Summary

In summary, we have used Raman spectroscopy and *ab initio* MD simulations to investigate the molecular mechanism of the blueshift of water molecules around six isomers of a monosaccharide. We observed the blueshift of water molecules in hydration shells of monosaccharides in both experiments and simulations. Using the simulations, it was demonstrated that in the first hydration shell, the closest water molecules (2.4 to 2.8 Å) mainly form acceptor H-bonds with the monosaccharide, and the water molecules are stable. Conversely, in the region of 2.8 to 3.3 Å, water molecules mainly form donor H-bonds with the monosaccharide, and these water molecules exhibit a blueshift. The vibrational state of the OH bond, which is oriented toward the bulk solvent, highly contributes to the observed blueshift. These results indicate that the weak H-bonded water molecules are arranged to be compatible not with the bulk water molecule but with the monosaccharide.

This molecular mechanism of the blue shift may be general for many other solutions. A previous study also observed the blue shift in an aqueous NaBr solution by using SC-Raman spectroscopy [150]. To validate the molecular mechanism, we performed a CPMD simulation of a NaBr solution. We applied the same analysis as we used for the monosaccharide solutions to the hydration shell around Na^+ and Br^- . By the “near” and “far” analysis, we found the blue shift arising only from “far” OH bond of water molecules around Br^- (see Fig. 6.10). This result supports the validity and generality of this molecular mechanism of the blue shift.

Furthermore, we demonstrated the relationship between the blue shift and the perturbation of the tetrahedral structure in water molecules at the hydrophilic interface. This perturbation of water molecules in the first hydration shell has previously been observed for many biological materials, e.g., DNA [169], small pep-

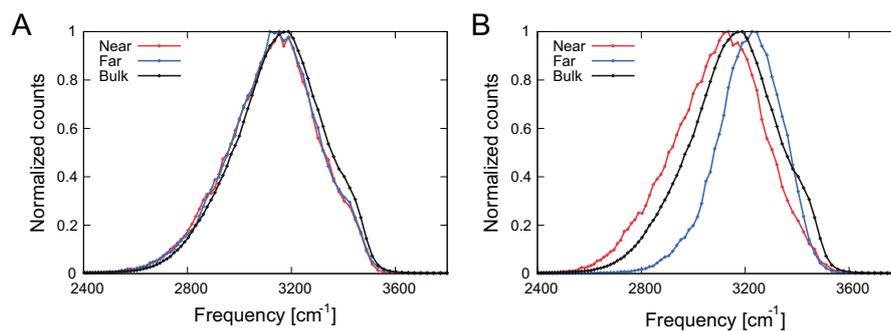


Fig. 6.10 **Spectra and water arrangement of water molecules in the NaBr solution.** (A and B) “near” and “far” OH vibrational spectra of water molecules in the first hydration shell of Na⁺ and Br⁻, respectively.

tides [170–172] and for various proteins [134]. Therefore, it can be extrapolated that the observed blue shift occurs not only at monosaccharide interfaces, but also at the other materials, and that these properties are essential to understanding the complex nature of many water-material interfaces.

Chapter 7

Conclusions

In this study, we have investigated the interactions between water molecules and biological materials using the MD simulations of the prion protein, rhodopsin, and monosaccharide.

First, we focused on the interactions between water molecules and the prion protein. The conformational change from the PrP^c to the PrP^{sc} is a key process in prion diseases. The prion protein has buried water molecules which significantly contribute to the stability of the protein; however, there has been no report investigating the influence on the buried hydration sites by a pathogenic mutation not adjacent to the buried hydration sites. We performed molecular dynamics simulations of wild type (WT) PrP^c and pathogenic point mutant T188R to investigate conformational changes and the buried hydration sites. In WT-PrP^c, four buried hydration sites were identified by residence time and rotational relaxation analysis. However, there are no stable buried hydration sites in one of T188R simulations, which indicates that T188R sometimes makes the buried hydration sites fragile. We also found that fluctuations of subdomains S1-H1-S2 and H1-H2 increase in T188R when the buried hydration sites become unstable. Since the side chain of arginine which is replaced from threonine in T188R is larger than of threonine, the side chain cannot be embedded in the protein, which is one of the causes of the instability of subdomains. These results showed correlations between the buried hydration sites and the mutation which is far from them, and provide a possible explanation for the instability by mutation.

Second, we investigated the interactions between water molecules and rhodopsin. Rhodopsin is a light-driven G-protein-coupled receptor that mediates signal transduction in eyes. Internal water molecules mediate activation of the receptor in a rhodopsin cascade reaction and contribute to conformational stability of the recep-

tor. However, it remains unclear how internal water molecules exchange between the bulk and protein inside, in particular through a putative solvent pore on the cytoplasmic side. Using all-atom molecular dynamics simulations, we identified the solvent pore on cytoplasmic side in both the Meta II state and the Opsin. On the other hand, the solvent pore does not exist in the dark-adapted rhodopsin. We revealed two characteristic narrow regions located within the solvent pore in the Meta II state. The narrow regions distinguish bulk and the internal hydration sites, one of which is adjacent to the conserved structural motif “NPxxY”. Water molecules in the solvent pore diffuse by pushing or sometimes jumping a preceding water molecule due to the geometry of the solvent pore. These findings revealed a total water flux between the bulk and the protein inside in the Meta II state, and suggested that these pathways provide water molecules to the crucial sites of the activated rhodopsin.

Finally, we focused on the interactions between water molecules and monosaccharide. Water-monosaccharide coupled interactions are essential for the function, stability, and dynamics of all glycans. Using molecular dynamics simulations, we investigated the effects of temperature, concentration, and monosaccharide isomer on the hydration structure and water dynamics in the hydration shell of monosaccharides in solution. We found that perturbations of the hydrogen-bond (H-bond) network in the first hydration shell around each monosaccharide molecule can be separated into two regions: one rich in water molecules with donor H-bonds (in the 2.4-2.8 Å region) and the other rich in water molecules with abundant acceptor H-bonds (in the 2.8-3.3 Å region). Moreover, we investigated the dependencies of clustering and conversion of the conformers of the monosaccharides on temperature and concentration. Increasing the concentration enhances monosaccharide clustering in all the monosaccharide solutions, while cluster formation does not depend on temperature. In the clusters, some water molecules in the hydration shell are replaced with monosaccharide oxygen atoms, which contributes to the shrinkage of the hydration shell with increasing monosaccharide concentration. The monosaccharides basically adopt one of two conformers, the stable chair or the unstable boat conformer. We revealed that the hydration structures of the boat and chair conformers were dramatically different. As the temperature increases, the content of the chair conformer decreases. Thus, the conversion of conformers strongly affects the hydration structure around the monosaccharide. These results are critical to

understand the important roles of the hydration structure in glycan solutions.

We also found that the interactions between water molecules and monosaccharide caused the high-frequency motions of water molecules. A variety of spectroscopic studies have observed a high-frequency motion in these water molecules, represented by a blueshift, at both hydrophobic and hydrophilic interfaces. However, the molecular mechanism behind this blueshift has remained unclear. Using Raman spectroscopy and *ab initio* molecular dynamics simulations, we revealed the molecular mechanism of the blueshift of water molecules around six monosaccharide isomers. In the first hydration shell, we found weak hydrogen-bonded water molecules that cannot have a stable tetrahedral water network. In the water molecules, the vibrational state of the OH bond oriented toward the bulk solvent strongly contributes to the observed blueshift. Our work suggests that the blueshift in various solutions originates from the vibrational motions of these observed water molecules.

From these results, we concluded that water molecules play fatal roles for the stability, fluctuations and functions of many biological materials, and this study provides essential information for understanding of the interactions between water molecules and biological materials.

Abbreviations and symbols

| | |
|-------------------|---|
| MD | Molecular dynamics |
| CPMD | Car-Parrinello molecular dynamics |
| DFT | Density function theory |
| PrP ^c | α -helix-rich cellular prion protein |
| PrP ^{sc} | β -sheet-rich scrapie prion protein |
| GPCR | G protein coupled receptor |
| POPC | Palmitoyl-oleoyl-phosphatidylcholine |
| RMSD | Root mean square deviation |
| RMSF | Root mean square fluctuation |
| RDF | Radial distribution function |
| SDF | Spatial distribution function |
| TOP | Tetrahedral order parameter |
| SC | Solute correlated |

Appendix

Publication list

Articles (related to this thesis)

- [1] Katsufumi Tomobe, Eiji Yamamoto, Takuma Akimoto, Masato Yasui, and Kenji Yasuoka, “Instability of buried hydration sites increases protein subdomains fluctuations in the human prion protein by the pathogenic mutation T188R”, *AIP Adv.* **6**, 055324 (2016).
- [2] Katsufumi Tomobe, Eiji Yamamoto, Kholmirzo Khomurodov, and Kenji Yasuoka, “Water permeation through the internal water pathway in activated GPCR rhodopsin”, *PLOS ONE* **12**, e0176876 (2017).
- [3] Katsufumi Tomobe, Eiji Yamamoto, Masato Yasui, and Kenji Yasuoka, “Effects of temperature, concentration, and isomer on the hydration structure in monosaccharide solutions”, *Phys. Chem. Chem. Phys.* **12**, e0176876 (2017).
- [4] Katsufumi Tomobe, Eiji Yamamoto, Dušan Kojić, Masato Yasui, and Kenji Yasuoka, “Origin of the blueshift of water molecules at interfaces of hydrophilic cyclic compounds”, *Science Adv.* **3**, e1701400, 2017.

Other articles

- [1] Katsufumi Tomobe, Eiji Yamamoto, Dušan Kojić, Masato Yasui, and Kenji Yasuoka, “Velocity auto-correlation function of ions and water molecules in different concentrations, anions and ion clusters”, *Mol. Sim.* **41**, 840 (2015).
- [2] Dušan Kojić, Roumiana Tsenkova, Katsufumi Tomobe, Kenji Yasuoka, and Masato Yasui, “Water Confined in the Local Field of Ions”, *ChemPhysChem* **15**, 4077 (2014).

References

- [1] Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **108**, 74–108 (2008).
- [2] Lu, Y., Yang, C.-Y. & Wang, S. Binding free energy contributions of interfacial waters in HIV-1 protease/inhibitor complexes. *J. Am. Chem. Soc.* **128**, 11830–11839 (2006).
- [3] Bizzarri, A. R. & Cannistraro, S. Molecular dynamics of water at the protein-solvent interface. *J. Phys. Chem. B* **106**, 6617–6633 (2002).
- [4] Dastidar, S. G. & Mukhopadhyay, C. Structure, dynamics, and energetics of water at the surface of a small globular protein: a molecular dynamics simulation. *Phys. Rev. E* **68**, 021921 (2003).
- [5] Fogarty, A. C. & Laage, D. Water dynamics in protein hydration shells: the molecular origins of the dynamical perturbation. *J. Phys. Chem. B* **118**, 7715–7729 (2014).
- [6] Duboué-Dijon, E. & Laage, D. Comparative study of hydration shell dynamics around a hyperactive antifreeze protein and around ubiquitin. *J. Chem. Phys.* **141**, 12B633.1 (2014).
- [7] Rani, P. & Biswas, P. Diffusion of hydration water around intrinsically disordered proteins. *J. Phys. Chem. B* **119**, 13262–13270 (2015).
- [8] Duboué-Dijon, E., Fogarty, A. C., Hynes, J. T. & Laage, D. Dynamical disorder in the DNA hydration shell. *J. Am. Chem. Soc.* **138**, 7610–7620 (2016).
- [9] Kojić, D., Tsenkova, R., Tomobe, K., Yasuoka, K. & Yasui, M. Water confined in the local field of ions. *ChemPhysChem* **15**, 4077–4086 (2014).
- [10] Róg, T., Murzyn, K. & Pasenkiewicz-Gierula, M. The dynamics of water at the phospholipid bilayer surface: a molecular dynamics simulation study. *Chem. Phys. Lett.* **352**, 323–327 (2002).
- [11] Murzyn, K., Zhao, W., Karttunen, M., Kurdziel, M. & Róg, T. Dynamics of

- water at membrane surfaces: Effect of headgroup structure. *Biointerphases* **1**, 98–105 (2006).
- [12] Yamamoto, E., Akimoto, T., Hirano, Y., Yasui, M. & Yasuoka, K. Power-law trapping of water molecules on the lipid-membrane surface induces water retardation. *Phys. Rev. E* **87**, 052715 (2013).
- [13] Schröder, C., Rudas, T., Boresch, S. & Steinhauser, O. Simulation studies of the protein-water interface. I. Properties at the molecular resolution. *J. Chem. Phys.* **124**, 234907 (2006).
- [14] Angel, T., Chance, M. & Palczewski, K. Conserved waters mediate structural and functional activation of family a (rhodopsin-like) G protein-coupled receptors. *Proc. Nat. Acad. Sci. U.S.A.* **106**, 8555 (2009).
- [15] Takano, N. *et al.* Solvent site-dipole field accompanying protein-ligand approach process. *Chem. Biol. Info. J.* **8**, 14–24 (2008).
- [16] Chong, S. & Ham, S. Impact of chemical heterogeneity on protein self-assembly in water. *Proc. Nat. Acad. Sci. U.S.A* **109**, 7636–7641 (2012).
- [17] Jana, M. & Bandyopadhyay, S. Restricted dynamics of water around a protein-carbohydrate complex: Computer simulation studies. *J. Chem. Phys.* **137**, 055102 (2012).
- [18] Yamamoto, E., Akimoto, T., Hirano, Y., Yasui, M. & Yasuoka, K. Power-law trapping of water molecules on the lipid-membrane surface induces water retardation. *Phys. Rev. E* **87**, 052715 (2013).
- [19] Yamamoto, E., Akimoto, T., Yasui, M. & Yasuoka, K. Origin of subdiffusion of water molecules on cell membrane surfaces. *Sci. Rep.* **4**, 4720 (2014).
- [20] Soto, C. Transmissible proteins: Expanding the prion heresy. *Cell* **149**, 968–977 (2012).
- [21] Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
- [22] Apostol, M. I., Perry, K. & Surewicz, W. K. Crystal structure of human prion protein fragment reveals a motif for oligomer formation. *J. Am. Chem. Soc.* **135**, 10202–10205 (2013).
- [23] Zahn, R. *et al.* NMR solution structure of the human prion protein. *Proc. Nat. Acad. Sci. U.S.A.* **97**, 145–150 (2000).
- [24] Pan, K. *et al.* Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc. Nat. Acad. Sci. U.S.A.* **90**,

- 10962–10966 (1993).
- [25] Swietnicki, W., Petersen, R., Gambetti, P. & Surewicz, W. pH-dependent stability and conformation of the recombinant human prion protein PrP (90–231). *J. Biol. Chem.* **272**, 27517–27520 (1997).
- [26] Hornemann, S. & Glockshuber, R. A scrapie-like unfolding intermediate of the prion protein domain PrP (121–231) induced by acidic pH. *Proc. Nat. Acad. Sci. U.S.A.* **95**, 6010–6014 (1998).
- [27] Matsunaga, Y. *et al.* Cryptic epitopes in N-terminally truncated prion protein are exposed in the full-length molecule: Dependence of conformation on pH. *Proteins* **44**, 110–118 (2001).
- [28] Gerber, R., Tahiri-Alaoui, A., Hore, P. & James, W. Conformational pH dependence of intermediate states during oligomerization of the human prion protein. *Protein Sci.* **17**, 537–544 (2008).
- [29] El-Bastawissy, E., Knaggs, M. & Gilbert, I. Molecular dynamics simulations of wild-type and point mutation human prion protein at normal and elevated temperature. *J. Mol. Graph. Model.* **20**, 145–154 (2001).
- [30] Riek, R. *et al.* Prion protein NMR structure and familial human spongiform encephalopathies. *Proc. Nat. Acad. Sci. U.S.A.* **95**, 11667–11672 (1998).
- [31] Van Der Kamp, M. & Daggett, V. The consequences of pathogenic mutations to the human prion protein. *Protein. Eng. Des. Sel.* **22**, 461–468 (2009).
- [32] Chen, X., Zhu, S., Wang, S., Yang, D. & Zhang, J. Molecular dynamics study on the stability of wild-type and the R220K mutant of human prion protein. *Molec. Simul.* 1–10 (2013).
- [33] Apetri, A., Surewicz, K. & Surewicz, W. The effect of disease-associated mutations on the folding pathway of human prion protein. *J Biol. Chem.* **279**, 18008–18014 (2004).
- [34] Chakroun, N. *et al.* Decrypting prion protein conversion into a β -rich conformer by molecular dynamics. *J. Chem. Theory Comput.* **9**, 2455–2465 (2013).
- [35] Malin, R., Zielenkiewicz, P. & Saenger, W. Structurally conserved water molecules in ribonuclease T1. *J. Biol. Chem.* **266**, 4848–4852 (1991).
- [36] Sreenivasan, U. & Axelsen, P. H. Buried water in homologous serine proteases. *Biochemistry* **31**, 12785–12791 (1992).
- [37] Shaltiel, S., Cox, S. & Taylor, S. S. Conserved water molecules contribute to

- the extensive network of interactions at the active site of protein kinase A. *Proc. Nat. Acad. Sci. U.S.A* **95**, 484–491 (1998).
- [38] Knight, J., Hamelberg, D., McCammon, J. & Kothary, R. The role of conserved water molecules in the catalytic domain of protein kinases. *Proteins* **76**, 527–535 (2009).
- [39] Haire, L. *et al.* The crystal structure of the globular domain of sheep prion protein. *J. Mol. Biol.* **336**, 1175–1183 (2004).
- [40] De Simone, A., Dodson, G., Verma, C., Zagari, A. & Fraternali, F. Prion and water: tight and dynamical hydration sites have a key role in structural stability. *Proc. Nat. Acad. Sci. U.S.A* **102**, 7535–7540 (2005).
- [41] Lorenz, H., Windl, O. & Kretzschmar, H. A. Cellular phenotyping of secretory and nuclear prion proteins associated with inherited prion diseases. *J. Biol. Chem.* **277**, 8508–8516 (2002).
- [42] Roeber, S. *et al.* Evidence for a pathogenic role of different mutations at codon 188 of PRNP. *PLoS One* **3**, e2147 (2008).
- [43] Guo, J., Ning, L., Ren, H., Liu, H. & Yao, X. Influence of the pathogenic mutations T188K/R/a on the structural stability and misfolding of human prion protein: Insight from molecular dynamics simulations. *Biochim. Biophys. Acta.* **1820**, 116–123 (2011).
- [44] Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
- [45] Garland, S. L. Are GPCRs still a source of new targets? *J. Biomol. Screen.* **18**, 947–966 (2013).
- [46] Lim, H. *et al.* Large-scale off-target identification using fast and accurate dual regularized one-class collaborative filtering and its application to drug repurposing. *PLoS Comput. Biol.* **12**, e1005135 (2016).
- [47] Smith, S. O. Structure and activation of the visual pigment rhodopsin. *Annu. Rev. Biophys.* **39**, 309–328 (2010).
- [48] Kukura, P., McCamant, D. W., Yoon, S., Wandschneider, D. B. & Mathies, R. A. Structural observation of the primary isomerization in vision with femtosecond-stimulated raman. *Science* **310**, 1006–1009 (2005).
- [49] Frutos, L. M., Andruniów, T., Santoro, F., Ferré, N. & Olivucci, M. Tracking the excited-state time evolution of the visual pigment with multiconfigurational quantum chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7764–7769

- (2007).
- [50] Röhrig, U. F., Guidoni, L., Laio, A., Frank, I. & Rothlisberger, U. A molecular spring for vision. *J. Am. Chem. Soc.* **126**, 15328–15329 (2004).
- [51] Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12729–12734 (2006).
- [52] Nakamichi, H. & Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew. Chem.* **118**, 4376–4379 (2006).
- [53] Blazynski, C. & Ostroy, S. Pathways in the hydrolysis of vertebrate rhodopsin. *Vision Res.* **24**, 459–470 (1984).
- [54] Wang, T. & Duan, Y. Retinal release from opsin in molecular dynamics simulations. *J. Mol. Recognit.* **24**, 350–358 (2011).
- [55] Hildebrand, P. W. *et al.* A ligand channel through the G protein coupled receptor opsin. *PLoS One* **4**, e4382 (2009).
- [56] Piechnick, R. *et al.* Effect of channel mutations on the uptake and release of the retinal ligand in opsin. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5247–5252 (2012).
- [57] Standfuss, J. *et al.* The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656–660 (2011).
- [58] Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).
- [59] Angel, T. E., Gupta, S., Jastrzebska, B., Palczewski, K. & Chance, M. R. Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14367–14372 (2009).
- [60] Jastrzebska, B., Palczewski, K. & Golczak, M. Role of bulk water in hydrolysis of the rhodopsin chromophore. *J. Biol. Chem.* **286**, 18930–18937 (2011).
- [61] Piechnick, R., Heck, M. & Sommer, M. E. Alkylated hydroxylamine derivatives eliminate peripheral retinylidene schiff bases but cannot enter the retinal binding pocket of light-activated rhodopsin. *Biochemistry* **50**, 7168–7176 (2011).
- [62] Furutani, Y., Shichida, Y. & Kandori, H. Structural changes of water molecules during the photoactivation processes in bovine rhodopsin. *Biochemistry* **42**, 9619–9625 (2003).
- [63] Yuan, S., Filipek, S., Palczewski, K. & Vogel, H. Activation of G-protein-

- coupled receptors correlates with the formation of a continuous internal water pathway. *Nat. Commun.* **5**, 4733–4742 (2014).
- [64] Ogata, K., Yuki, T., Hatakeyama, M., Uchida, W. & Nakamura, S. All-atom molecular dynamics simulation of photosystem ii embedded in thylakoid membrane. *J. Am. Chem. Soc.* **135**, 15670–15673 (2013).
- [65] Espinoza-Fonseca, L. M. & Ramírez-Salinas, G. L. Microsecond molecular simulations reveal a transient proton pathway in the calcium pump. *J. Am. Chem. Soc.* **137**, 7055–7058 (2015).
- [66] Aryal, P., Abd-Wahab, F., Bucci, G., Sansom, M. S. & Tucker, S. J. A hydrophobic barrier deep within the inner pore of the TWIK-1 K2P potassium channel. *Nat. Commun.* **5**, 4377 (2014).
- [67] Dong, H., Fiorin, G., Carnevale, V., Treptow, W. & Klein, M. L. Pore waters regulate ion permeation in a calcium release-activated calcium channel. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17332–17337 (2013).
- [68] Sun, X., Ågren, H. & Tu, Y. Functional water molecules in rhodopsin activation. *J. Phys. Chem. B* **118**, 10863–10873 (2014).
- [69] Esko, J. D. & Selleck, S. B. Order out of chaos: Assembly of ligand binding sites in heparan sulfate 1. *Annu. Rev. Biochem.* **71**, 435–471 (2002).
- [70] Ghazarian, H., Idoni, B. & Oppenheimer, S. B. A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta. Histochemica.* **113**, 236–247 (2011).
- [71] Crowe, J. H., Crowe, L. M. & Chapman, D. Preservation of membranes in anhydrobiotic organisms: the role of trehalose. *Science* **223**, 701–703 (1984).
- [72] Crowe, J. H. & Crowe, L. M. Preservation of mammalian cells learning nature’s tricks. *Nat. Biotechnol.* **18**, 145–146 (2000).
- [73] Magno, A. & Gallo, P. Understanding the mechanisms of bioprotection: a comparative study of aqueous solutions of trehalose and maltose upon supercooling. *J. Phys. Chem. Lett.* **2**, 977–982 (2011).
- [74] Zhu, Y., Zajicek, J. & Serianni, A. S. Acyclic forms of [1-13c] aldohexoses in aqueous solution: quantitation by 13C NMR and deuterium isotope effects on tautomeric equilibria. *J. Org. Chem.* **66**, 6244–6251 (2001).
- [75] Schnupf, U., Willett, J. & Momany, F. DFTMD studies of glucose and epimers: anomeric ratios, rotamer populations, and hydration energies. *Carbohydr. Res.* **345**, 503–511 (2010).

-
- [76] Plazinski, W. & Drach, M. Kinetic characteristics of conformational changes in the hexopyranose rings. *Carbohydr. Res.* **416**, 41–50 (2015).
- [77] Paolantoni, M. *et al.* Light scattering spectra of water in trehalose aqueous solutions: Evidence for two different solvent relaxation processes. *J. Phys. Chem. B* **113**, 7874–7878 (2009).
- [78] Branca, C., Magazu, S., Maisano, G., Bennington, S. & Fåk, B. Vibrational studies on disaccharide/H₂O systems by inelastic neutron scattering, raman, and IR spectroscopy. *J. Phys. Chem. B* **107**, 1444–1451 (2003).
- [79] Groot, C. & Bakker, H. Hydration dynamics of aqueous glucose probed with polarization-resolved fs-IR spectroscopy. *J. Chem. Phys.* **140**, 234503 (2014).
- [80] Bordat, P., Lerbret, A., Demaret, J.-P., Affouard, F. & Descamps, M. Comparative study of trehalose, sucrose and maltose in water solutions by molecular modelling. *Europhys. Lett.* **65**, 41 (2004).
- [81] Kräutler, V., Müller, M. & Hünenberger, P. H. Conformation, dynamics, solvation and relative stabilities of selected β -hexopyranoses in water: a molecular dynamics study with the GROMOS 45A4 force field. *Carbohydr. Res.* **342**, 2097–2124 (2007).
- [82] Zielinski, F., Mutter, S. T., Johannessen, C., Blanch, E. W. & Popelier, P. L. The raman optical activity of β -D-xylose: where experiment and computation meet. *Phys. Chem. Chem. Phys.* **17**, 21799–21809 (2015).
- [83] Mutter, S. T. *et al.* Conformational dynamics of carbohydrates: Raman optical activity of D-glucuronic acid and N-acetyl-D-glucosamine using a combined molecular dynamics and quantum chemical approach. *Phys. Chem. Chem. Phys.* **17**, 6016–6027 (2015).
- [84] Melcrová, A., Kessler, J., Bouř, P. & Kaminský, J. Simulation of raman optical activity of multi-component monosaccharide samples. *Phys. Chem. Chem. Phys.* **18**, 2130–2142 (2016).
- [85] Brady, J. W. Molecular dynamics simulations of alpha-D-glucose in aqueous solution. *J. Am. Chem. Soc.* **111**, 5155–5165 (1989).
- [86] Driesner, T., Seward, T. & Tironi, I. Molecular dynamics simulation study of ionic hydration and ion association in dilute and 1 molal aqueous sodium chloride solutions from ambient to supercritical conditions. *Geochim. Cosmochim. Acta.* **62**, 3095–3107 (1998).
- [87] Mason, P., Neilson, G., Enderby, J., Saboungi, M.-L. & Brady, J. Structure

- of aqueous glucose solutions as determined by neutron diffraction with isotopic substitution experiments and molecular dynamics calculations. *J. Phys. Chem. B* **109**, 13104–13111 (2005).
- [88] Suzuki, T. The hydration of glucose: the local configurations in sugar–water hydrogen bonds. *Phys. Chem. Chem. Phys.* **10**, 96–105 (2008).
- [89] Lee, S. L., Debenedetti, P. G. & Errington, J. R. A computational study of hydration, solution structure, and dynamics in dilute carbohydrate solutions. *J. Chem. Phys.* **122**, 204511 (2005).
- [90] Calzolari, L. & Zahn, R. Influence of pH on NMR structure and stability of the human prion protein globular domain. *J. Biol. Chem.* **278**, 35592–35596 (2003).
- [91] Guex, N. & Peitsch, M. C. Swiss-model and the swiss-pdb viewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
- [92] Phillips, J. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
- [93] Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).
- [94] Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
- [95] Dima, R. & Thirumalai, D. Exploring the propensities of helices in PrPc to form β sheet using NMR structures and sequence alignments. *Biophys. J.* **83**, 1268–1280 (2002).
- [96] Dima, R. I. & Thirumalai, D. Probing the instabilities in the dynamics of helical fragments from mouse PrPc. *Proc. Nat. Acad. Sci. U.S.A.* **101**, 15335–15340 (2004).
- [97] Lu, X., Wintrode, P. L. & Surewicz, W. K. β -sheet core of human prion protein amyloid fibrils as determined by hydrogen/deuterium exchange. *Proc. Nat. Acad. Sci. U.S.A.* **104**, 1510–1515 (2007).
- [98] De Simone, A., Dodson, G., Fraternali, F. & Zagari, A. Water molecules as structural determinants among prions of low sequence identity. *FEBS Lett.* **580**, 2488–2494 (2006).
- [99] Laage, D. & Hynes, J. A molecular jump mechanism of water reorientation.

-
- Science* **311**, 832–835 (2006).
- [100] Laage, D. & Hynes, J. On the molecular mechanism of water reorientation. *J. Phys. Chem. B* **112**, 14230–14242 (2008).
- [101] Laage, D. & Hynes, J. Do more strongly hydrogen-bonded water molecules reorient more slowly? *Chem. Phys. Lett.* **433**, 80–85 (2006).
- [102] De Simone, A., Zagari, A. & Derreumaux, P. Structural and hydration properties of the partially unfolded states of the prion protein. *Biophys. J.* **93**, 1284–1292 (2007).
- [103] Schwarzinger, S., Horn, A. H., Ziegler, J. & Sticht, H. Rare large scale sub-domain motions in prion protein can initiate aggregation. *J. Biomol. Struct. Dyn.* **23**, 581–590 (2006).
- [104] Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 crystal structure. *J. Mol. Biol.* **342**, 571–583 (2004).
- [105] Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).
- [106] Choe, H.-W. *et al.* Crystal structure of metarhodopsin ii. *Nature* **471**, 651–655 (2011).
- [107] Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
- [108] Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
- [109] Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
- [110] MacKerell Jr, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- [111] Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
- [112] Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
- [113] Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the amber

- ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
- [114] Jambeck, J. P. & Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **8**, 2938–2948 (2012).
- [115] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
- [116] Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
- [117] Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
- [118] Bussi, G., Zykova-Timan, T. & Parrinello, M. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *J. Chem. Phys.* **130**, 074101 (2009).
- [119] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
- [120] Gravelle, S. *et al.* Optimizing water permeability through the hourglass shape of aquaporins. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16367–16372 (2013).
- [121] Fischer, G. *et al.* Crystal structure of a yeast aquaporin at 1.15 reveals a novel gating mechanism. *PLoS Biol.* **7**, e1000130 (2009).
- [122] Mirzadegan, T., Benkő, G., Filipek, S. & Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **42**, 2759–2767 (2003).
- [123] Pardo, L., Deupi, X., Dölker, N., López-Rodríguez, M. L. & Campillo, M. The role of internal water molecules in the structure and function of the rhodopsin family of G protein-coupled receptors. *Chem. Bio. Chem.* **8**, 19–24 (2007).
- [124] Yamamoto, E., Akimoto, T., Hirano, Y., Yasui, M. & Yasuoka, K. 1/f fluctuations of amino acids regulate water transportation in aquaporin 1. *Phys. Rev. E* **89**, 022718 (2014).
- [125] Bemporad, D., Essex, J. W. & Luttmann, C. Permeation of small molecules through a lipid bilayer: a computer simulation study. *J. Phys. Chem. B* **108**, 4875–4884 (2004).

-
- [126] Grossfield, A., Pitman, M. C., Feller, S. E., Soubias, O. & Gawrisch, K. Internal hydration increases during activation of the G-protein-coupled receptor rhodopsin. *J. Mol. Biol.* **381**, 478–486 (2008).
- [127] Mitchell, D. C. & Litman, B. J. Effect of ethanol and osmotic stress on receptor conformation reduced water activity amplifies the effect of ethanol on metarhodopsin ii formation. *J. Biol. Chem.* **275**, 5355–5360 (2000).
- [128] Lay, W. K., Miller, M. S. & Elcock, A. H. Optimizing solute–solute interactions in the GLYCAM06 and CHARMM36 carbohydrate force fields using osmotic pressure measurements. *J. Chem. Theory Comput.* **12**, 1401–1407 (2016).
- [129] Kirschner, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field carbohydrates. *J Comput. Chem.* **29**, 622–655 (2008).
- [130] Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
- [131] Dabkowska, A. P., Foglia, F., Lawrence, M. J., Lorenz, C. D. & McLain, S. E. On the solvation structure of dimethylsulfoxide/water around the phosphatidylcholine head group in solution. *J. Chem. Phys.* **135**, 12B611 (2011).
- [132] Bhattacharjee, N. & Biswas, P. Structure of hydration water in proteins: a comparison of molecular dynamics simulations and database analysis. *Biophys. Chem.* **158**, 73–80 (2011).
- [133] Errington, J. R. & Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature* **409**, 318–321 (2001).
- [134] Rani, P. & Biswas, P. Local structure and dynamics of hydration water in intrinsically disordered proteins. *J. Phys. Chem. B* **119**, 10858–10867 (2015).
- [135] Tamai, Y., Tanaka, H. & Nakanishi, K. Molecular dynamics study of polymer-water interaction in hydrogels. 2. hydrogen-bond dynamics. *Macromolecules* **29**, 6761–6769 (1996).
- [136] Stenger, J. *et al.* Molecular dynamics and kinetics of monosaccharides in solution. a broadband ultrasonic relaxation study. *J. Phys. Chem. B* **104**, 4782–4790 (2000).
- [137] Behrends, R. & Kaatze, U. Molecular dynamics and conformational kinetics of mono- and disaccharides in aqueous solution. *ChemPhysChem* **6**, 1133–1145 (2005).
- [138] Polacek, R., Stenger, J. & Kaatze, U. Chair–chair conformational flexibil-

- ity, pseudorotation, and exocyclic group isomerization of monosaccharides in water. *J. Chem. Phys.* **116**, 2973–2982 (2002).
- [139] Diebold, U., Li, S.-C. & Schmid, M. Oxide surface science. *Annu. Rev. Phys. Chem.* **61**, 129–148 (2010).
- [140] Verdaguer, A., Sacha, G., Bluhm, H. & Salmeron, M. Molecular structure of water at interfaces: Wetting at the nanometer scale. *Chem. Rev.* **106**, 1478–1510 (2006).
- [141] Scatena, L., Brown, M. & Richmond, G. Water at hydrophobic surfaces: weak hydrogen bonding and strong orientation effects. *Science* **292**, 908–912 (2001).
- [142] Perera, P. *et al.* Observation of water dangling OH bonds around dissolved nonpolar groups. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12230–12234 (2009).
- [143] Davis, J. G., Gierszal, K. P., Wang, P. & Ben-Amotz, D. Water structural transformation at molecular hydrophobic interfaces. *Nature* **491**, 582–585 (2012).
- [144] Davis, J. G., Rankin, B. M., Gierszal, K. P. & Ben-Amotz, D. On the cooperative formation of non-hydrogen-bonded water at molecular hydrophobic interfaces. *Nat. Chem.* **5**, 796–802 (2013).
- [145] Varanasi, S., Guskova, O., John, A. & Sommer, J.-U. Water around fullerene shape amphiphiles: A molecular dynamics simulation study of hydrophobic hydration. *J. Chem. Phys.* **142**, 224308 (2015).
- [146] Verma, P. K., Kundu, A., Ha, J.-H. & Cho, M. Water dynamics in cytoplasm-like crowded environment correlates with the conformational transition of the macromolecular crowder. *J. Am. Chem. Soc.* **138**, 16081–16088 (2016).
- [147] Tomlinson-Phillips, J. *et al.* Structure and dynamics of water dangling OH bonds in hydrophobic hydration shells. comparison of simulation and experiment. *J. Phys. Chem. A* **115**, 6177–6183 (2011).
- [148] Ratajczak, H., Orville-Thomas, W. & Rao, C. Charge transfer theory of hydrogen bonds: relations between vibrational spectra and energy of hydrogen bonds. *Chem. Phys.* **17**, 197–216 (1976).
- [149] Rozenberg, M., Loewenschuss, A. & Marcus, Y. An empirical correlation between stretching vibration redshift and hydrogen bond length. *Phys. Chem. Chem. Phys.* **2**, 2699–2702 (2000).
- [150] Perera, P. N., Browder, B. & Ben-Amotz, D. Perturbations of water by alkali

- halide ions measured using multivariate raman curve resolution. *J. Phys. Chem. B* **113**, 1805–1809 (2009).
- [151] Perera, P., Wyche, M., Loethen, Y. & Ben-Amotz, D. Solute-induced perturbations of solvent-shell molecules observed using multivariate raman curve resolution. *J. Am. Chem. Soc.* **130**, 4576–4577 (2008).
- [152] Hutter, J. & Iannuzzi, M. CPMD: Car-Parrinello molecular dynamics. *Zeitschrift für Kristallographie-Crystalline Materials* **220**, 549–551 (2005).
- [153] Car, R. & Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, 2471 (1985).
- [154] Kirschner, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field. carbohydrates. *J. Comput. Chem.* **29**, 622–655 (2008).
- [155] Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
- [156] Kleinman, L. & Bylander, D. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* **48**, 1425 (1982).
- [157] Troullier, N. & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, 1993 (1991).
- [158] Mahoney, M. W. & Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **112**, 8910–8922 (2000).
- [159] Hess, B., Bekker, H., Berendsen, H. J., Fraaije, J. G. *et al.* LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
- [160] Tanzi, L., Ramondo, F. & Guidoni, L. Vibrational spectra of water solutions of azoles from QM/MM calculations: Effects of solvation. *J. Phys. Chem. A* **116**, 10160–10171 (2012).
- [161] Hiratsuka, M., Ohmura, R., Sum, A. K., Alavi, S. & Yasuoka, K. A molecular dynamics study of guest–host hydrogen bonding in alcohol clathrate hydrates. *Phys. Chem. Chem. Phys.* **17**, 12639–12647 (2015).
- [162] Vela-Arevalo, L. V. & Wiggins, S. Time-frequency analysis of classical trajectories of polyatomic molecules. *Int. J. Bifurcat. Chaos.* **11**, 1359–1380 (2001).
- [163] Mallik, B. S., Semparathi, A. & Chandra, A. Vibrational spectral diffusion and hydrogen bond dynamics in heavy water from first principles. *J. Phys. Chem. A* **112**, 5104–5112 (2008).

- [164] Mallik, B. S. & Chandra, A. An ab initio molecular dynamics study of supercritical aqueous ionic solutions: Hydrogen bonding, rotational dynamics and vibrational spectral diffusion. *Chem. Phys.* **387**, 48–55 (2011).
- [165] Karmakar, A. & Chandra, A. Water in hydration shell of an iodide ion: Structure and dynamics of solute-water hydrogen bonds and vibrational spectral diffusion from first-principles simulations. *J. Phys. Chem. B* **119**, 8561–8572 (2015).
- [166] Carmona, R., Hwang, W.-L. & Torr sani, B. *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms, with an Implementation in S*, vol. 9 (Academic Press, 1998).
- [167] Molteni, C. & Parrinello, M. Glucose in aqueous solution by first principles molecular dynamics. *J. Am. Chem. Soc.* **120**, 2168–2171 (1998).
- [168] Suzuki, T. & Sota, T. Circular hydrogen bond networks on the surface of β -ribofuranose in aqueous solution. *J. Phys. Chem. B* **109**, 12603–12611 (2005).
- [169] Chakraborty, K., Mantha, S. & Bandyopadhyay, S. Molecular dynamics simulation of a single-stranded DNA with heterogeneous distribution of nucleobases in aqueous medium. *J. Chem. Phys.* **139**, 08B618_1 (2013).
- [170] Agarwal, M., Kushwaha, H. R. & Chakravarty, C. Local order, energy, and mobility of water molecules in the hydration shell of small peptides. *J. Phys. Chem. B* **114**, 651–659 (2009).
- [171] Nayar, D., Agarwal, M. & Chakravarty, C. Comparison of tetrahedral order, liquid state anomalies, and hydration behavior of mTIP3P and TIP4P water models. *J. Chem. Theory Comput.* **7**, 3354–3367 (2011).
- [172] Nayar, D. & Chakravarty, C. Sensitivity of local hydration behaviour and conformational preferences of peptides to choice of water model. *Phys. Chem. Chem. Phys.* **16**, 10199–10213 (2014).