博士論文

Molecular dynamics studies of lysozyme solution and bio-protective sugar solutions for controlling water rotational relaxation time

(水分子の回転緩和時間の制御を目指したリゾチーム水溶液と生体保護物質糖類水溶液の分子動力学的研究)

胡 康

Doctoral Dissertation

Molecular dynamics studies of lysozyme solution and bio-protective sugar solutions for controlling water rotational relaxation time

By

Kang HU

Presented to

GRADUATE SCHOOL OF ENGINEERING, THE UNIVERSITY OF TOKYO

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the Field of Mechanical Engineering

February 2023

Abstract

Along with the great advancements in therapeutic proteins, the stabilization methods are in urgent demand to maintain the high quality of these biopharmaceuticals, given the necessity of protein stability for the successful implication of its therapy. Protein pharmaceuticals are usually in a liquid formulation, and the surrounding water layers (hydration layer, HL) are known to have a significant influence on the stability and dynamics of protein conformation. During the stabilization processes, protective agents (PAs) are usually added to help the preserved biological materials overcome the induced physiochemical stresses, such as ice crystallization, low temperature, dehydration, etc. Therefore, the capability of predicting the effect of additives and designing protein stabilizers could facilitate the preservation of biopharmaceuticals. Among PAs, sugars, the membrane impermeable molecules, have been widely employed to protect protein drugs and food during the freezing and drying processes, due to the unique sugar-water interactions. The investigations of water dynamics surrounding proteins and the effect of sugars on water dynamics may advance our knowledge of controlling water dynamics, thus the design of protein stabilizers.

In this dissertation, after a systematical review of recent processes in the experimental and computational studies on water dynamics in protein/sugar solutions, the molecular origin of the gigahertz domain of the dielectric spectra is first revealed in the following Chapter 2, by combining molecular dynamics (MD) simulation with broadband dielectric spectroscopy (BDS) experiments. The dielectric spectra for protein solutions and sugar solutions are obtained by the probability distribution of water relaxation time from MD simulation. The dielectric relaxation time obtained from this distribution is in good agreement with the measured γ relaxation time, which suggests that the rotational self-correlation of water molecules underlies the gigahertz domain of the dielectric spectra. This original finding associates the molecular water behavior with the macroscopic dielectric measurements which are widely employed to characterize different food products.

Upon validating the present MD simulation, two questions of the ongoing debate regarding the fundamental water-protein interaction are addressed (Chapter 3): 1) how far proteins can affect water rotational dynamics, i.e., the thickness of protein HL; and 2) to what extent the characteristics of hydration water differ from those of the bulk

water. To disentangle the molecular scale distribution of water dynamics around a solute biomolecule, the rotational dynamics of water around lysozyme are investigated here by MD simulation. Water rotational relaxation time is found to range from several picoseconds to tens of nanoseconds in protein solutions. Depending on the overall time it stays within the HL (total residence time), water can be classified into 3 kinds: highly mobile water (HMW), partially restricted water (PRW), and completely restricted water (CRW). In dilute solutions, water relaxation time exhibits a Gaussian probability distribution. In concentrated solutions, the distribution is distorted to a longer time constant due to the increasing proportion of PRW and CRW. A stochastic analysis combining the relaxation times and trajectories of every single water molecule is proposed and the two-dimensional probability distribution of water relaxation time and its distance from the protein surface is obtained. This analysis suggests that water rotational dynamics accelerate with the farther distance from the protein surface. Regardless of protein concentration, this spatial distribution of relaxation time reveals that the water rotation is severely retarded within 3 Å from the lysozyme surface and is nearly comparable to pure water when farther than 10 Å. The HL thickness is subsequently identified as this 3 Å thick water layer from the protein surface, in view of the remarkable slowdown wherein.

In the next chapter (Chapter 4), the molecular mechanism of water rotational dynamics is investigated from the perspective of electrostatic interactions, by MD simulation. By monitoring the variation of the exerted electric field on the individual water molecule, how the exerted intrinsic electric field determines water rotational dynamics is presented. It is found for the first time that the relaxation time of water rotation is equivalent to that of the reorientation of the exerted electric field for each water molecule, regardless of its translational motion. Namely, water rotation synchronizes with the experienced electric field reorientation. Furthermore, bringing the MD's superiority into full play, the reorientation process of the local field near the protein surface is obtained, which provides insight into the local hydration dynamics. Within the HL, the relaxation time of local field reorientation ranges from a few hundred femtoseconds to a little surpassing one hundred picoseconds. Comparing the spatial distribution of local field reorientation relaxation time with that of water rotational relaxation time, it is further suggested that water rotation dynamics are subject to the reorientation of the local overall field within the HL. While outside the

HL, the relaxation time of the local electric field reorientation is short enough (sub-picosecond) to assume the delta function, showing the electric force with randomly changing orientation is applied to each water molecule. This observation suggests the importance of the total residence time, and it is consequently found that water rotation relaxation time is approximately exponentially dependent on its total residence time.

Based on findings in Chapters 2-4, water dynamics in aqueous PAs solutions (sugars and sugar alcohol here) are discussed in Chapter 5 by analyzing the probability distributions of both water residence time and water relaxation time obtained by MD simulations. In bio-protective sugar and sugar alcohol solutions, water is just mildly retarded with the relaxation time spanning from several picoseconds to merely a few tens of picoseconds. Water residence time exhibits a gaussian distribution under various concentrations. These two observations are discrepant from the case of protein solutions. Moreover, I find that in all solutions, water relaxation time also increases exponentially with its residence time. Based on this, a decuple-retarding period which specifies the prerequisite residence time to obtain tenfold retardation of the relaxation time is defined. The larger the solute dipole moment, the smaller the decuple-retarding period, therefore the faster the increment rate of relaxation time versus residence time. The discovered random walk of water within the solution further suggests that the larger ratio of the hydration volume (total volume of HLs) to the overall water accessible space results in a longer residence time, thus the longer relaxation time distribution. This result indicates that the specific surface area of the additive is another key factor affecting water dynamics. When water residence time distribution is similar, the solute dipole begins to manifest. The above findings suggest the effect of balance between PA dipole and size on water rotational dynamics. Acquiring such kinds of basic laws for controlling water dynamics will have a meaningful impact on the development of protein stabilizers, which is of great significance in the long-term preservation of pharmaceutical proteins.

To summarize, this dissertation mainly focuses on the water rotational dynamics (relaxation time) of individual water molecules in biomolecule solutions and bio-protective sugar solutions. The proposed stochastic analysis is an original method that circumvents the balance of spatial resolution and calculation accuracy of water relaxation time, compared with the widely accepted multiple HLs model. The macroscopic dielectric measurements are associated with molecular water behavior.

The perspective of electric field reorientation reveals the driving mechanic of water rotational relaxation. The finding that water relaxation time is exponentially proportional to residence time paves the way to controlling water rotational relaxation time, thus to the design of protein protective agents.

Contents

Nomenclature and Abbreviations

x	Distance between water molecule and the solute surface, Å
t	Time, ps
τ	Relaxation time, ps
t_R	Normalized residence time
$t_{R,Max}$	Normalized maximum continuous residence time
$t_{R,Total}$	Normalized total residence time
D	Dipole moment
E	Electric field
θ	Angle between electric field and water dipole, °
ζ	Rotational retardation factor (RRF)
η	Excess RRF
$\boldsymbol{\mu}$	Water dipole vector, D
ξ	Decuple-retarding period
α	Overall percentage of HLs overlapping
r	Intermolecular distance in the solvent-sharing model, Å
a	External radius of the HL in the solvent-sharing model, Å
h	Height of two spherical caps in the solvent-sharing model, Å
β, δ, γ	Dielectric relaxation time, ps
A	Relaxation amplitude for a water molecule
k	Order of relaxation mode
ψ	Correlation function
P	Probability function
c	Protein concentration, mg/mL
$\boldsymbol{M}$	Total dipole moment, D
ε	Dielectric permittivity

ε''	Imaginary part of Dielectric permittivity
ε'	Real part of Dielectric permittivity
$\Delta\varepsilon$	Dielectric relaxation strength
ω	Circular frequency, rad/sec
ε_∞	High-frequency limit of ε
ε_0	Dielectric permittivity of vacuum
$\boldsymbol{E}$	Electric field vector
$\boldsymbol{f}$	Force exerted on atom
N_{solute}	Number of solute molecules
N_{hyd}	Number of hydration water
r_{solute}	Radius of a solute molecule
V_{access}	Accessible (free) space for water, $Å^3$
V_{HLs}	Volume of the hydration layers
$n_{\text{HL,s}}$	Sharing (overlapped) number of the hydration layer
PDF	Probability density function
CDF	Cumulative density function
MD	Molecular dynamics
DS	Dielectric spectroscopy
BDS	Broadband dielectric spectroscopy
PAs	Protective agents
SASA	Solvent-accessible surface area
HL	Hydration layer
HB	Hydrogen bond
SAS	Small-angle scattering
NS	Neutron scattering
ODNP	Overhauser dynamic nuclear polarization
NMR	Nuclear magnetic resonance

TDFSS	Time-dependent fluorescence Stokes shift
RRF	Rotational retardation factor
MRD	Magnetic relaxation dispersion
EDLS	Extended depolarized light scattering
ACF	Auto-correlation function
EJM	Extended jump model
EMOR	Exchange-mediated orientational randomization
NOE	Nuclear Overhauser effect
NQR	Nuclear quadrupole relaxation
MSD	Mean square displacement

Chapter 1 Introduction

1.1 Stabilization of therapeutic proteins

In recent decades, the major process in liposome technology and development in molecular biology have brought up the great advancement in therapeutic proteins[1]. Since protein stability is one of the key factors that can hinder the successful implication of its therapy, stabilization methods are in urgent demand to maintain the high quality of these biopharmaceuticals[1-2]. Currently, the four prevailing protein preservation and stabilization methods are classic slow freezing, ice-free vitrification, freeze-drying (lyophilization), and drying. During the above stabilization processes, ice formation and/or glass transition occur(s), probably inducing an array of detrimental stresses on the preserved biomolecules, such as ice crystallization, low temperature, dehydration, etc[2].

In this regard, protective agents (PAs), also named cryoprotectants are usually added to help the preserved biological materials survive the aforementioned physiochemical stresses. PAs have been reported to suppress ice crystallization, facilitate glass transition, and structure the hydration water around protein both directly and indirectly[1-5].

1.1.1 Molecular preservation mechanism of additives

Weng et.al summarized the four main hypotheses used to explain the biomolecule stabilization mechanism, by the example of protein in a sugar–water system in **Figure 1-1**[2]. 1) First, the vitrification hypothesis (Green and Angell) suggests that carbohydrates form a glassy matrix with high viscosity and that this matrix rigidity inhibits the protein molecular motion, thereby preventing the conformation change of the embedded protein[6]. This hypothesis is also known as mechanical entrapment. 2) Second, the water replacement hypothesis (Crowe *et al.*) suggests that, in a dehydration state, sugar molecules replace the hydration water molecules by forming direct hydrogen bonds (HBs) with the biomolecule, to preserve the native structure of the embedded biomolecule[7-9]. 3) Third, the water entrapment hypothesis (Belton and Gil) presumes that the existence of sugar molecules helps trap the water molecules between protein and sugar molecules upon cooling or/and dehydration, thereby maintaining

sufficient hydration water to prevent the protein denaturation[10-11]. 4) Fourth, the preferential exclusion/hydration theory (Xie and Timasheff) proposes that sugar molecules are preferentially excluded from the protein hydration layer (HL), thus increasing the solvent-accessible surface area (SASA) of proteins and resulting in preferential hydration of proteins. Owning to this reason, protein denaturation becomes thermodynamically unfavorable[12-13]. This exclusion could originate from the excluded volume effects and an increment in the surface tension of water caused by solutes (i.e., the free energy cost to remove water from the surface)[14-15].

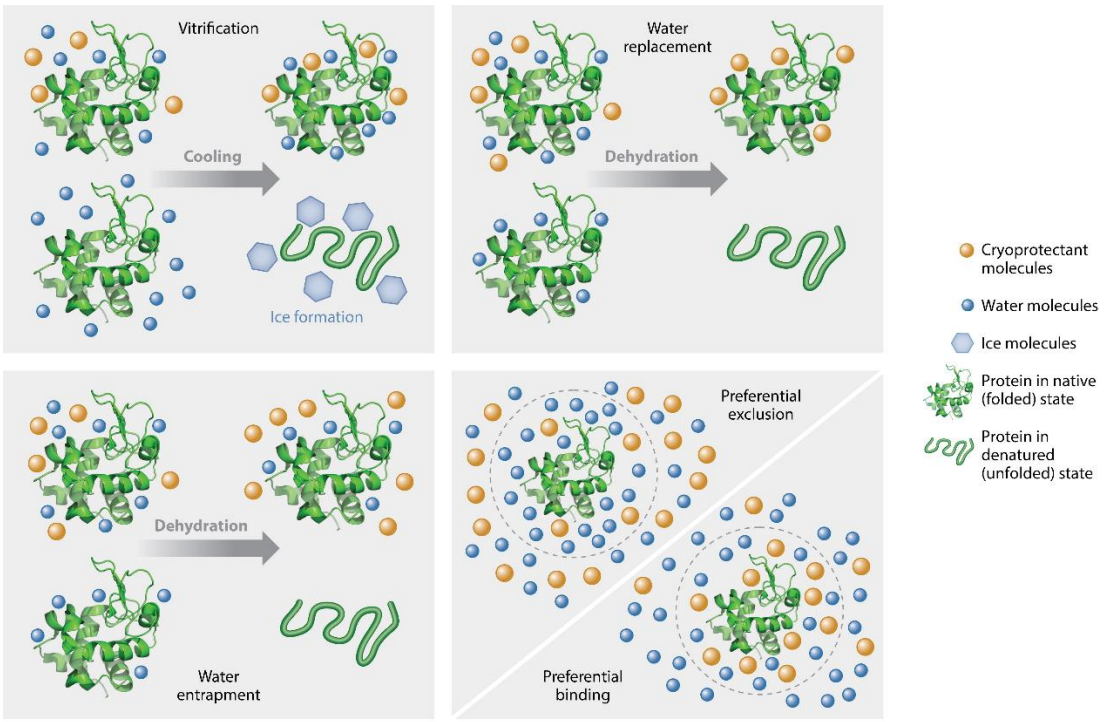

Figure 1-1. A summary of current hypotheses regarding the cryoprotectants protection mechanism in preventing protein damage induced by freezing or dehydration. Gold spheres indicate cryoprotectant molecules (for instance, trehalose), and blue ones indicate water molecules[2].

Corradin et.al visualized this preferential exclusion by MD simulation[16]. As shown in **Figure 1-2**, there is a region in the vicinity of the protein (from ~12 to ~17 Å, T = 300 K for instance) where the trehalose density is evidently lower than the average and an outer area (from ~17 to ~30 Å, T = 300 K) where the trehalose density is higher[16]. The latter part signals the formation of the transient cage of trehalose around the protein. The constrained, thin water layer between the protein and the sugar could lock the protein surface and hinder the large-scale, solvent-coupled protein motions[15]. It also works as the bridge between the protein and the externally excluded sugar-water

matrix[17-20] and this "*double-shell*" structure could significantly reduce the motional freedom of the loop and helix structures of the protein conformation, with respect to the water-only solvated system[17]. On the other hand, this trapped water layer was also believed to possess reduced electrostatic solvation properties toward the protein, since the trehalose may compete with the target protein for forming HBs with the trapped water molecules and finally reduce the number of protein–solvent HBs by a factor of four[21]. Such a reduction of the solvation might take responsibility for a compensating enhancement of intra-protein interactions and, thus, a stabilization of the protein in its native structure[21].

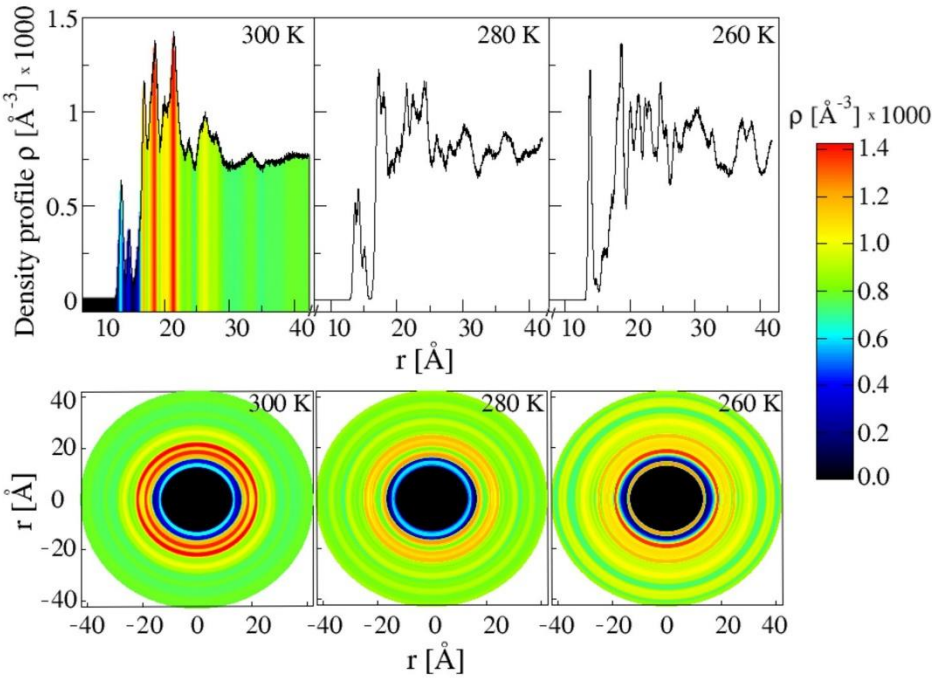

Figure 1-2. (**Top**) Density profile of trehalose molecules around the lysozyme. This density profile is derived by dividing the number of trehalose molecules found between an interval (r and $r + \Delta r$) by the corresponding spherical shell volume $4\pi[(r + \Delta r)3 - r^3]/3$, with Δr = 0.02 Å. (**Bottom**) Color heat maps of the 2D projections of the density profile of trehalose molecules around the lysozyme. The variable r indicates the distance between the center of mass (COM) of trehalose and that of lysozyme[16].

The water entrapment scenario is somewhat the extension of the preferential exclusion presumption, while the water replacement theory seems to discord with them. However, it is worth noting that the preferential exclusion and water entrapment mechanisms are applicative when the protein is hydrated which means that it is solvated in an aqueous sugar-water solution with a sugar mass percent less than 90%[2]. As the solution becomes dehydrated (for example, sugar is highly concentrated), the water replacement theory begins to take effect. Besides, the vitrification hypothesis does not

conflict with the other theories either, given the observation by Corradin et.al that the presence of trehalose introduces a glass-like behavior of water in the cage of trehalose molecules near proteins, of which the rotational diffusion is significantly slower[16, 22-24]. In general, slowing down the water dynamics around protein seems to be the key to achieving high-quality stabilization.

1.2 Water dynamics around protein molecules

1.2.1 Role of water dynamics in protein stability

Water layers surrounding proteins (hydration shells/layers) are known to have a significant influence on the stability and dynamics of protein conformation[25-26]. The conformational fluctuations in proteins are slaved to the dynamics of water both in the hydration shell and in the bulk solvent (the so-called "*unified model of protein dynamics*")[27-29]. And the protein folding rate, which has been reported directly related to the degradation rate[30-31], is inversely proportional to the power of the solvent viscosity[28]. The "master-slave" model shedding light on the water-protein interactions has been recently revisited by the Molecular dynamics (MD) simulations[32]. Even relatively distant water molecules can influence the dipolar functional groups of protein due to the low static dielectric constant of both protein core and hydration water[32]. Furthermore, by the Pearson correlation coefficient, the correlation between the energy of any two separate parts of the system (E_A and E_B) is defined as

$$\rho_{AB} = \frac{\langle \delta E_A \delta E_B \rangle}{\sqrt{\langle (\delta E_A)^2 \rangle \langle (\delta E_B)^2 \rangle}}, \tag{1-1}$$

where δE denotes the energy fluctuation and ρ_{AB} denotes the correlated contributions. For five different protein-water solutions, the values of ρ_{AB} between the protein self-energy ($E_{P\text{-}P}$) and protein-water interaction energy ($E_{P\text{-}W}$) are all around -0.9, which signifies the strong anticorrelation of the energy fluctuations between protein and water contributions [32]. The trajectories of $\delta E_{P\text{-}P}$ and $\delta E_{P\text{-}W}$ for the myoglobin-water system are shown in **Figure 1-3** which shows this anticorrelated fluctuation clearly[32]. Energy flows from protein to water (and vice versa) aided by the side-chain conformational fluctuations like a coupled-oscillator scenario and a single transition in the protein energy landscape corresponds to multiple transitions in the case of water[32].

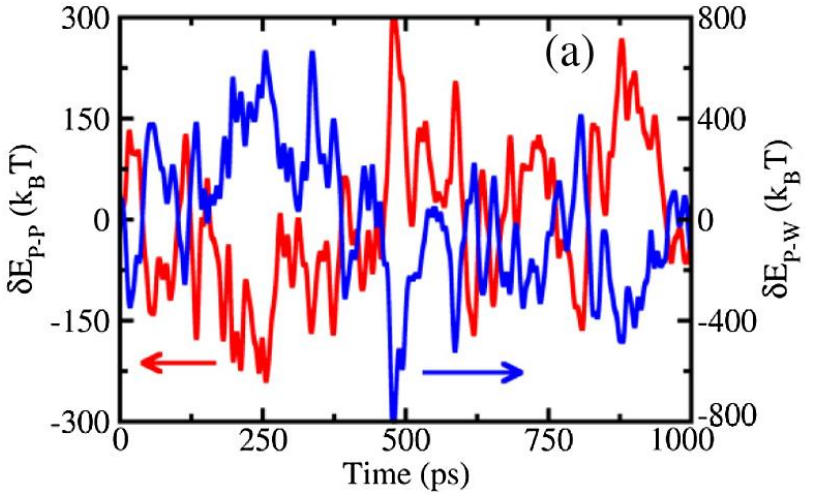

Figure 1-3. Protein (myoglobin) self-energy fluctuations ($\delta E_{P\text{-}P}$) and protein-water cross-energy fluctuations ($\delta E_{P\text{-}W}$)[32].

Recently, as shown in **Figure 1-4**, the deterioration rate of lactose dehydrogenase (LDH) in solution was found to have the power-law relationships with the dielectric relaxation times of water in the solution, particularly sensitive to the γ relaxation time attributed to the rotation of bulk water which was viewed as the distant water unperturbed by biomolecule[33]. With Debye's theory where the relaxation time is linear to the solvent viscosity, this work supports the high-viscosity hypothesis[34] where the molecular mobility surrounding the target protein dominates its deterioration from molecular-scale phenomena and the unified model of protein dynamics[29]. The extremely lower T_1 relaxation times from the ^{1}H NMR experiments (viz. the longer rotational correlation time) have been ascribed to a more "ice-like" behavior of water[35]. The simulation study together with free energy calculations by Saladino et.al showed that the driving force behind the native state protection is a slowdown of the solvent rotational dynamics which corresponds to a higher unfolding transition temperature of protein HP35[36]. The retarded solvent behaves like the "colder" aqueous solvent which works less effectively in interfering with the intra-protein interactions thus sustaining the native fold[36]. In the same spirit, a state-of-the-art THz spectroscopy study on aqueous ribonuclease A solutions with the addition of various osmolytes by Hishida et.al reported that the inhabited water collective rotational dynamics induce an increment in protein denaturation temperature in the presence of protein stabilizers and vice versa[37]. These findings suggest that controlling and evaluating water dynamics (e.g, the β relaxation time in **Figure 1-4**) leads to managing the deterioration of biomolecules,

which might be a potentially useful method for screening the bioprotective agents in the long-term preservation of pharmaceutical proteins.

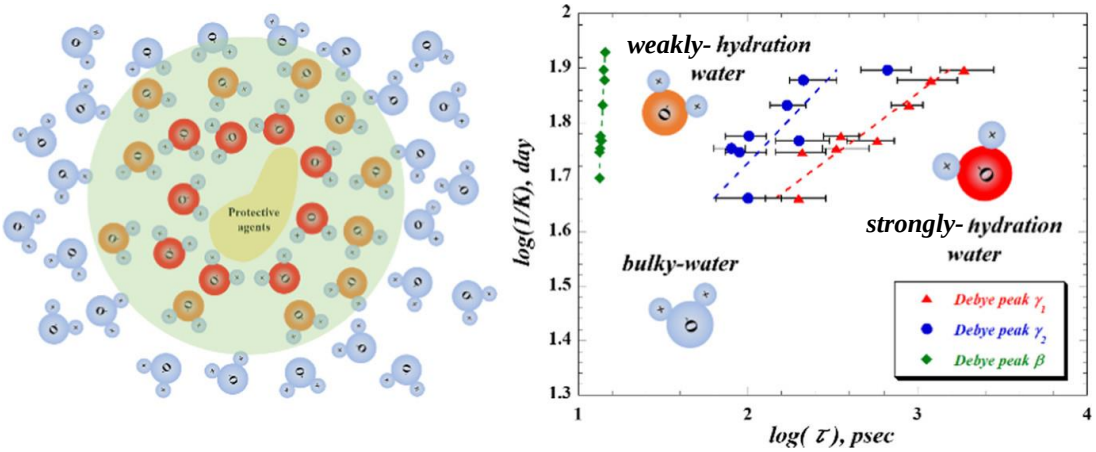

Figure 1-4. (**Left**) Schematic description of the distribution of water molecules in the preservative solution with protective agents. Bulk water (τ_β) is shown in blue spheres, which is the farthest water from the protective agents. Water molecules in stronger interaction with protective agent molecules (τ_γ) are shown in red (strongly) and orange (weakly) spheres, which interacted with the protective agents (Trehalose or ε-poly-l-lysine). (**Right**) Power law between the dielectric relaxation times of corresponding water molecules in the preservative solution and the deterioration time of LDH in the preservative solution at 40 °C. Error bars signify the standard deviation[33].

1.2.2 Hydration layer thickness

With regard to the understanding of water dynamics surrounding proteins, one of the relevant questions is how thick such an HL is, i.e., how distant the bulk water is. The first HL was first determined as the area within 3 Å from the lysozyme surface in terms of density and structural properties by both small-angle scattering (SAS, **Figure 1-5**) and MD simulation[38-39]. In a successive simulation performed by Ebbinghaus et. al, water dynamics were characterized by the HB correlation function, $C(t)$, which quantifies the probability that an HB which is formed between two water molecules at a given time, $t = 0$, still exists at a later time, t, regardless whether the bond has once been broken between 0 and t[40]. They found that this HB correlation function of water within ~ 10 Å around protein λ^*_{6-85} was distinct from that of bulk water[40]. A following simulation by Heyden et.al suggested that the vibrational motion between atoms of the protein surface and water extended up to 10 Å from the λ-repressor protein[41]. In the recent decade, except that the Neutron scattering (NS) spectroscopy has demonstrated the slowdown of water relaxation time within ~5.5 Å from a green fluorescent protein[42], most studies have reported the protein-induced perturbation reaching over 10 Å, such

as the Overhauser dynamic nuclear polarization (ODNP)-enhanced nuclear magnetic resonance (NMR) measurement on a peripheral membrane protein[43] and the time-dependent fluorescence Stokes shift (TDFSS) studies on rat liver fatty acid-binding protein (rLFABP)[44] and fish eye lens protein γM7-crystallin (γM7)[45]. In terms of the water relaxation time obtained from MD simulation, with the advancement of the Voronoi tessellation method, several groups[46-49] have dissected the solvent into 6~9 HLs based on the initial position of water and calculated the averaged relaxation time thereof. Regardless of the various proteins simulated and differing HL thicknesses adopted, both Steinhauser's group[46] and Halle's group[48] have observed that water rotational dynamics are severally retarded within their self-defined first HL (5~6 Å for globular protein). Beyond the first layer, water rotational relaxation time approximates that of bulk water (pure water). **Figure 1-6** shows the four-protein average excess rotational retardation factor (RRF) versus the layer index n[48]. The RRF is the water relaxation time normalized by that of bulk water, often expressed as $\zeta = \tau_R/\tau_{R,bulk}$ where τ indicates relaxation time and the subscript '$_R$' indicates water rotation. The excess RRF is $\eta \equiv \zeta - 1$.

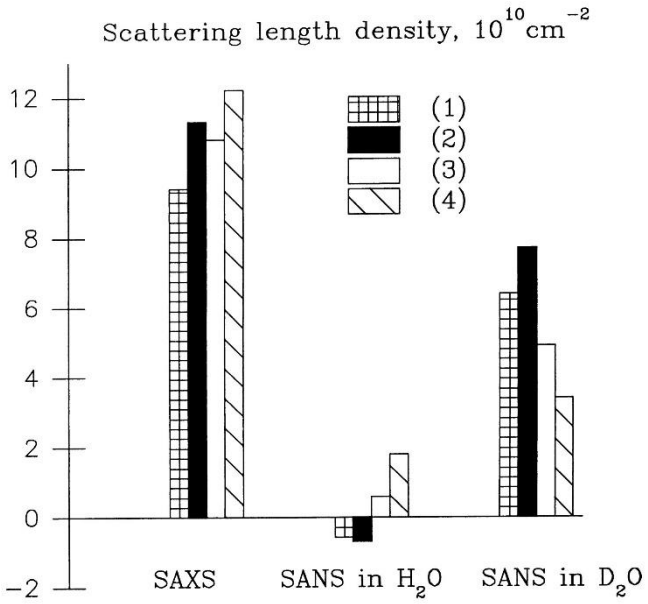

Figure 1-5. Relationship among the scattering densities of the bulk solvent, protein, and protein–solvent interface for x-rays and neutrons in H_2O and D_2O. (1) Bulk solvent, (2) a shell with a density 20% above that of the bulk solvent, (3) mobility of the side chains on the protein surface; scattering density in the interface is drawn in the middle between those of protein and of bulk; (4) protein. The scattering density of protein in D_2O is larger than that in H_2O because of H/D exchange[38].

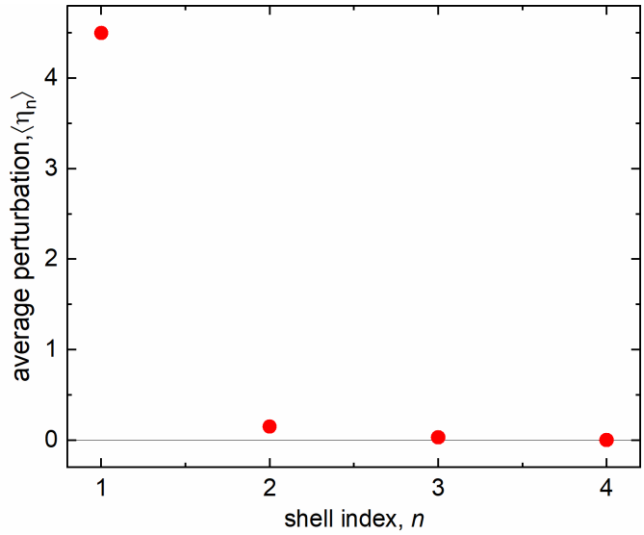

Figure 1-6. Excess rotational retardation factor for water molecules in the *n*-th hydration shell averaged over the four proteins (a triple mutant of the immunoglobulin-binding domain B1 of protein G (GB1), bovine pancreatic trypsin inhibitor (BPTI), mammalian ubiquitin without the three C-terminal residues (UBQ), and isoform 2-14 of the *β*-helical antifreeze protein from *Tenebrio molitor* (AFP))[48].

1.2.3 Heterogeneity of hydration water dynamics

Another essential question is to what extent the characteristic of hydration water differs from those of bulk water. As the time scales and processes for bulk (pure) water summarized in **Figure 1-7**, water librational motion occurs on a time range between several tens and several hundreds of femtoseconds[26]. For the molecular rotation/reorientation, it occurs on a time scale of around several picoseconds[25-26, 50]. In the present work, the focus is on how water rotational relaxation is affected by a biomolecule, particularly the hydration water molecules.

Early dielectric spectroscopy (DS) measurements on protein solutions implied different water dynamics in the hydration shell[51]. The subsequent analyses of such measured spectra categorized 4 kinds of dielectric relaxations for whale myoglobin solution (viz, τ_β = 74 ns, $\tau_{\delta 1}$ = 10.3 ns, $\tau_{\delta 2}$ = 40 ps, and τ_γ = 8.29 ps) and concluded that the δ-dispersions originate from the bound water (hydration water)[52-54]. It should be noted that the nomenclature of dielectric relaxation time here is different from that used by Wei et.al (**Figure 1-4**)[33]. A DS study on lysozyme solution found a bimodal relaxation composed of a low-frequency process (1.22 ns) assigned to the relaxation of bound water and a high-frequency process (10.0 ps) to free water[55]. In a similar study on equine myoglobin solutions, 5 relaxations were detected, namely $\tau_1 \approx$ 29 ns

ascribable to the tumbling of the whole myoglobin molecule, τ_4 = 10 ps and $\tau_5 \approx 5$ ps to free water, and $\tau_3 \approx 10$ ps and $\tau_4 \approx 200$ ps to hydration water together with the internal motions of the solute[56]. In recent decades, quite extensive DS studies[57-58] have measured the dielectric spectra of diverse protein solutions where several discrete relaxation components were detected consisting of a β (~ 10 ns), one or two δ (dozens of ps ~ several hundred ps), and a γ relaxation times (approximately 10 ps). A following NMR experiment discerned the presence of such a slow relaxation component subject to bound water[59]. However, NMR[60-61] and magnetic relaxation dispersion (MRD)[62-66] techniques provided an average relaxation time of all water molecules in the solution, which is 2-6-fold slower than the average relaxation time of bulk water. Extended depolarized light scattering (EDLS)[67-68] and NS[42] spectroscopy experiments also reported a comparable retardation range (2-10) for hydration water. TDFSS[44-45, 50, 69-70] of protein solutions even suggested the dramatic slowdown of water relaxation by several orders of magnitude. To summarize, although the results of both experiments[33, 42-45, 50, 57-75] and *in silico* simulations[46, 48-49, 76-96] suggested that water dynamics are retarded by the protein molecules, quantitative descriptions of the slowdown in hydration water dynamics remain controversial.

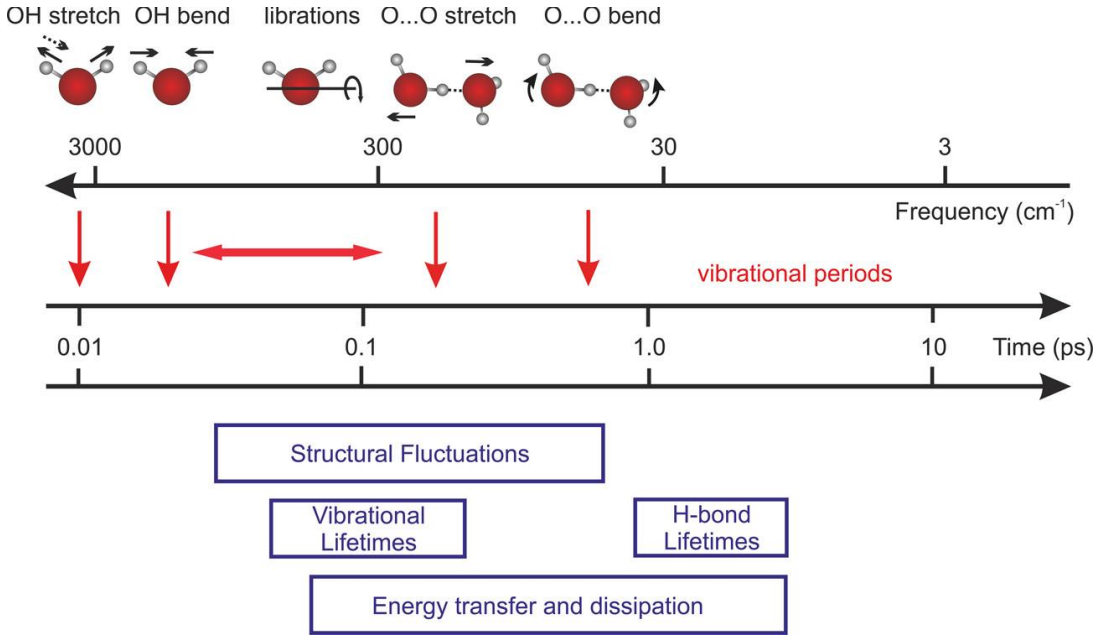

Figure 1-7. Time scales of different processes in bulk (pure) water. Red arrows above the logarithmic time axis denote the periods of vibrational and librational degrees of freedom schematically illustrated at the top of the figure. Horizontal boxes below the time axis depict the time span covered by the particular processes of bulk water dynamics[26].

Capable of providing insight into molecular-scale water dynamics, MD simulations have been intensively employed to study the rotational movement of single or collective motions of water in HLs to probe the corresponding relaxation times. Some pioneering works[88-89, 94] on protein solutions calculated the rotational relaxation time of each single water molecule within 4.4 Å from the atom belonging to a protein molecule, which they defined as hydration water. However, it is worth noting that although Laage and his coworkers calculated the time auto-correlation function (ACF) with trajectories of 5 or even 20 ns, they merely used the 2–10 ps interval of the ACF to derive the relaxation time for the sake of removing the contributions from water which have moved out of the HL[88-89]. Later simulation studies by Bagchi's group[95-96] defined water within 1 nm of its nearest protein atom as water of the first HL and water within 1 and 2 nm of the protein surface as water of the second layer. They also reported a rotational relaxation time distribution of water molecules that stay in the first protein HL longer than 100 ps (**Figure 1-8**).

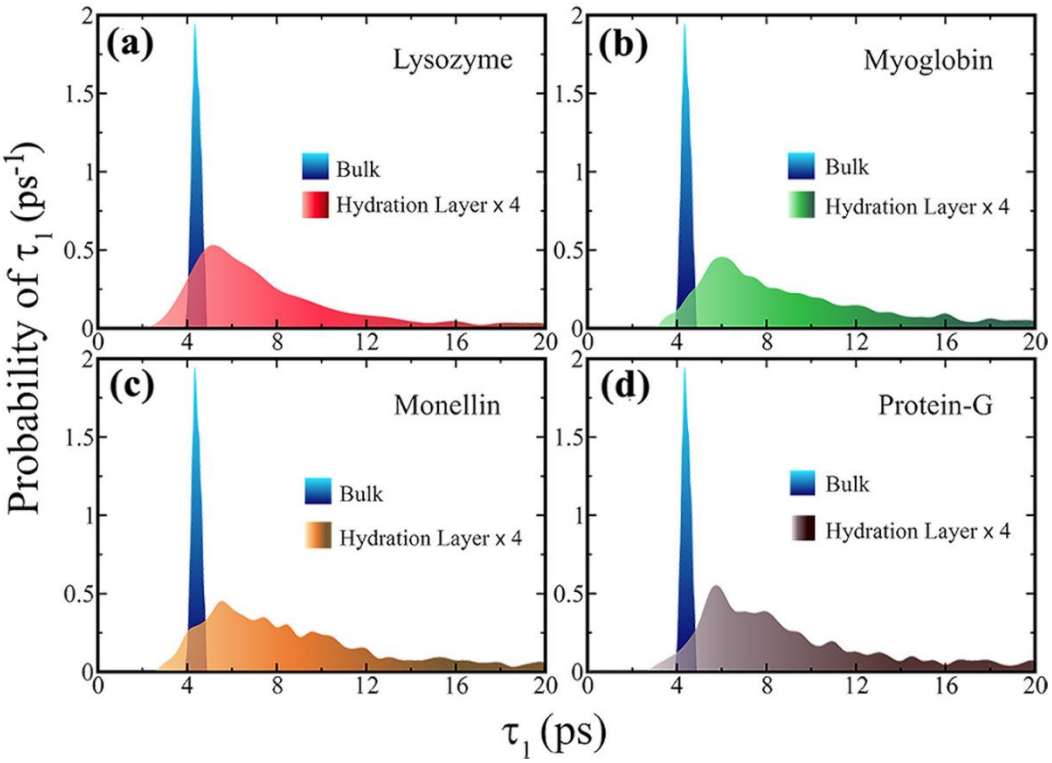

Figure 1-8. Probability distribution of rotational relaxation time of water molecules inside the protein HL for four different proteins and bulk water (blue). In the insets, the normalized and averaged rotational time auto-correlation function is shown with the same color codes. (**a**) Lysozyme, (**b**) myoglobin, (**c**) monellin, and (**d**) protein-G[95-96].

All the above studies claimed that the rotational relaxation time of the hydration water was distributed much more broadly than that of the bulk water. Additionally, the

relaxation times of hydration water derived from MD are on average 2~3 times longer than that of bulk water. In a recent report[97] of MD simulation study coupled with DS, water molecules within 3.5, 5.5, and 9.0 Å from bovine serum albumin (BSA) protein surfaces were categorized as tightly bound, loosely bound, and bulk water, respectively. The relaxation times of these water were found to correspond well to the measured β, δ, and γ dielectric relaxation time. Recently, dividing the solvent into even more consecutive HLs has been attracting more attention with the advancement of Voronoi tessellation analysis[46, 48-49]. Apparently, the statistical populations of water molecules used to calculate the local water relaxation time in these MD simulations are confined to empirical or adjusted "*hydration layers*". Moreover, the hydration water which directly interacts with the proteins usually possesses a time constant of at least on the scale of hundreds of picoseconds. In this regard, a 10 ps or 100 ps long time period may be insufficient to characterize its rotational dynamics. For instance, Fogarty and Laage[89] have assessed the simulation convergence by computing the reorientation time-correlation function and the distribution of reorientation slowdown factors respectively by two independent 20 ns trajectories and a shorter 4 ns trajectory, and found that the 4 ns trajectory could provide a reliable determination of these quantities. Therefore, defining the HLs based on the initial position of water molecules may call the computation accuracy of relaxation time into question, especially for water near the protein surface.

1.3 Water rotational dynamics mechanism

In addition to the quantity of water rotational relaxation, the knowledge of the predominant factors of water rotational dynamics is also necessary to manipulate the water dynamics around proteins. Although water dynamics dominating these characteristics have been extensively studied by means of various computational[83, 95, 98] and experimental methods[42, 44, 66, 69], the underlying mechanism characterizing the water rotational dynamics is still a field of active research.

1.3.1 Extended jump model

For example, the extended jump model (EJM) introduced by Laage and Hynes[99] provided a quantitative description of the HB reorientation relaxation time in liquid water. In this model, the water molecule reorients predominantly via the sudden, large-

amplitude jumps when any of its water hydroxyl (OH) groups trades HB with the surrounding water molecules. The HB exchange process is regarded as a chemical reaction, as illustrated in **Figure 1-9 (a)**. First, a new water oxygen acceptor approaches from outer shells while the initial HB elongates. Once the initial and the approaching oxygen acceptors are equidistant from the target water oxygen, the water OH can suddenly jump orienting from one acceptor to another. Finally, the new HB stabilizes, and the previous partner leaves. Additionally, a minor contribution results from the slower tumbling of the frame of the intact H-bonded complex[98]. They further applied this HB jump model to water on a biomolecule interface (**Figure 1-9 (b)**).

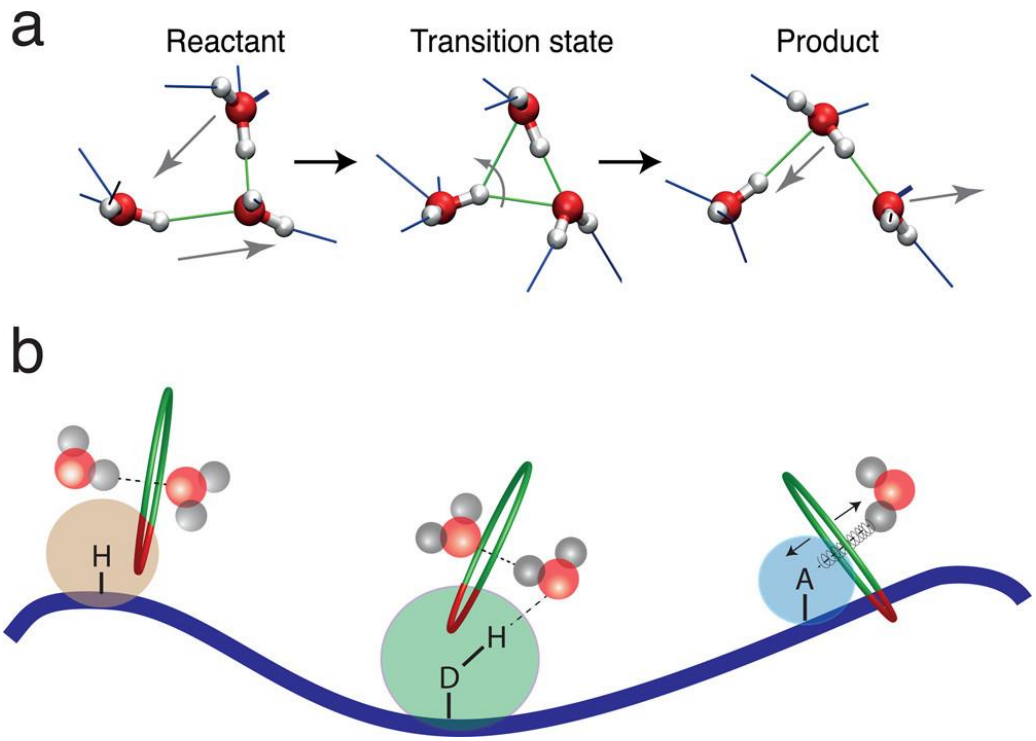

Figure 1-9. (**a**) Molecular H-bond jump mechanism for water reorientation process. (**b**) Schematic diagram with a protein interface and the three types of sites, respectively hydrophobic, H-bond donor, and H-bond acceptor, together with a pictorial representation of the types of perturbation they induce on water dynamics (excluded volume and H-bond strength factors)[26].

Hydrophobic groups affect HB jump dynamics merely by impeding the approach of a new water HB acceptor through an excluded volume effect, while the hydrophilic groups by forming strong HBs. HB donors and acceptors also act differently: although the donors can form bonds of different strengths, these bonds basically act only on the oxygen atom of the rotating water molecule. Thus, the resulting torque's influence on the OH reorientation is negligible, and these HB donors impact water dynamics mainly via their excluded-volume effect like hydrophobic groups. On the other hand, the H-

bong acceptors can disturb water reorientation dynamics via both the excluded-volume effect and the formed HB strength[88-89]. However, in their work, water reorientation relaxation time was derived from ab exponential fit within the interval 2–10 ps of the ACF to avoid contributions from water molecules "which have moved to a different environment"[88-89], as introduced above. Such concerns may be warranted for water near hydrophobic groups, but not for water near hydrophilic residues or in confined sites where water features essentially zero diffusion in such a short time span[48, 100]. Therefore, the relaxation time of the water was underestimated by an order of magnitude[48] and the rotational retardation factor (RRF, water relaxation time normalized by that of bulk water) distribution over the protein surface was "quite uniform"[88].

1.3.2 Exchange-mediated orientational randomization mechanism

Persson and Halle developed an exchange-mediated orientational randomization (EMOR) picture to account for the effect of protein confinement on the periphery water rotation motion[101], as illustrated in **Figure 1-10**[48]. For unconfined water molecules, water rotation is governed by the cooperative motion of neighboring several water molecules which "*lowers the transition-state barrier by enabling water-water H-bonds to be broken and formed in concert*"[48], just as bulk water. This is consistent with the EJM in some respects. For water in low-confinement sites (**top panel**), the mechanism is also essentially the same as that collective mode for bulk water. Nevertheless, due to the presence of the protein on one side, the distribution of water neighbors is less symmetric than bulk water. Besides, in the polar sites, the O and N atoms on proteins cannot act as efficiently in this cooperative process as the more mobile water molecules in the bulk region. For the highly confined circumstance, owning to the steric exclusion of most water neighbors and the essentially fixed HB polarity and position of the most protein polar atoms, water molecules are therefore orientationally restricted. The exchange event is thus required to rescue the target water molecule, whereby a new water molecule enters the confined cavity and replaces its position and the original one now can rotate with little retardation. This rotation mechanism is termed EMOR. They thus concluded that "*water rotation interpolates between a perturbed but bulk-like collective mechanism at low confinement and an exchange-mediated orientational randomization (EMOR) mechanism at high confinement*."[48] It should be noted that this EMOR picture does not provide a quantitative description of water rotational relaxation,

despite that it affords an understanding of the protein surface structure effect on water rotational dynamics in the view of the HBs exchange scenario.

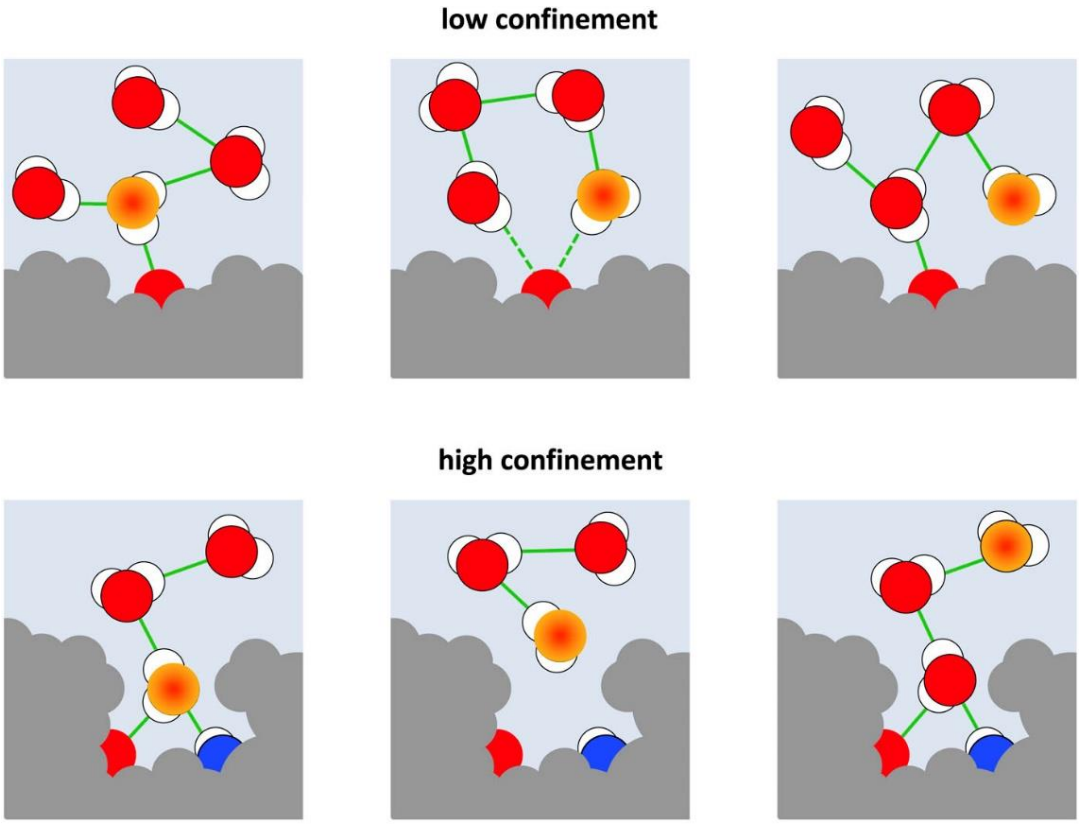

Figure 1-10. Schematic illustration of water rotation in low and high confinement sites, showing H-bonds (green), protein-O (red), and protein-N (blue) atoms[48].

1.3.3 The perspective of electrostatic interactions

In consideration of the essence of intermolecular interactions, the effect of electrostatic force has aroused wide interest in the research field of the water rotation mechanism. Earlier Time-Dependent Fluorescence Stokes Shift (TDFSS) experiments performed by Zhong and his coworkers suggested that the fundamental water-protein electrostatic interaction was crucial to HL formation[50, 102-103]. It was presumed that the exerted electro-static force from the protein immobilized the water molecules in its rotational motion, and hence to reorient, the exchange with water of the outer layer was required. As shown in **Figure 1-11**, their recent TDFSS determination further demonstrated that the hydration dynamics were highly related to the local residue charge[44-45, 104]. Combing with the recently developed fluctuation theory, the EJM also revealed that the electrostatic interactions dominated the activation energies for HBs jump and water reorientation (**Figure 1-12**), independent of the water model adopted[105-106].

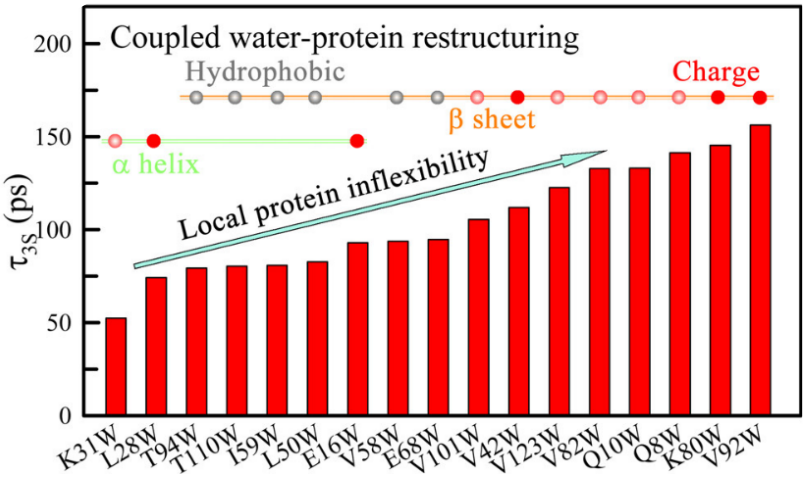

Figure 1-11. Time scales of the water-network relaxations τ_{3S}. The mutants are classified based on their local chemical identities (solid red, heavy charge; pink, partial charge; gray, hydrophobic) and secondary structures (orange bar, β-sheet; green bar, α-helix)[44].

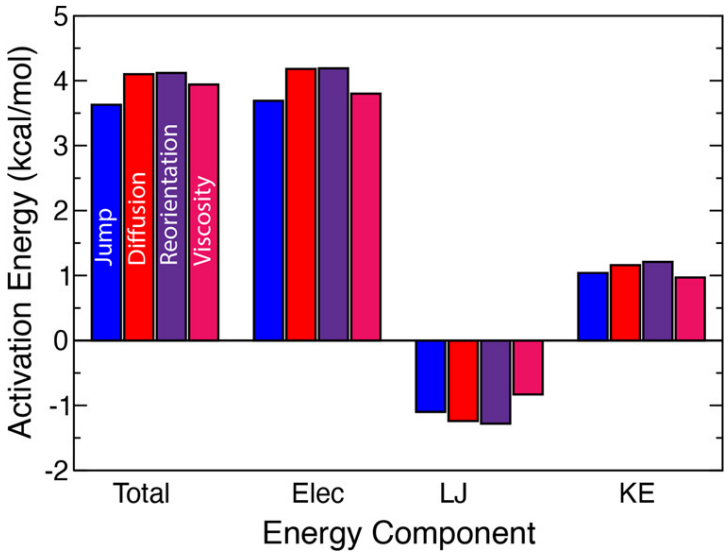

Figure 1-12. Total activation energy and its electrostatic, Lennard-Jones, and kinetic energy components for the H-bond jump time (blue), diffusion coefficient (red), reorientation time (purple), and shear viscosity (magenta) of TIP4P/2005 water[106].

As a kind of dipolar molecule with a permanent dipolar moment, a water molecule can be oriented by the local electric field[107-109]. In a protein solution, the effective dipole moment of the protein molecule is usually 2 orders of magnitude larger than the water dipole[95], which renders the alignment of the water dipole with the protein dipole plausible. Even though neither water orientation nor the exerted electric field is currently accessible by experiments, the molecular dynamics (MD) simulation can calculate these properties. To date, several simulation works[39, 46-49, 110-112] have addressed the affinity of a single water dipole toward the protein's electric field although the studies are fairly limited and have reached only few consensuses.

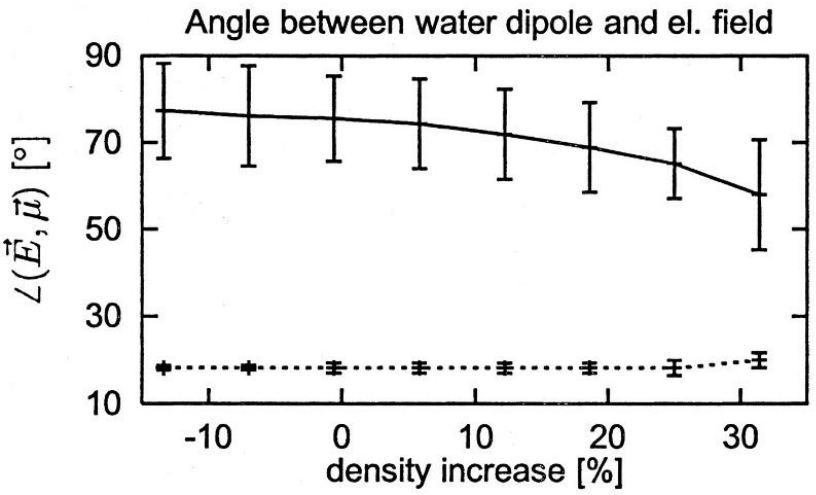

Figure 1-13. Correlation of angle between water dipoles and electrostatic field with the density. Solid line: electrostatic field calculated only from protein atoms. Dashed line: electrostatic field calculated from the entire system (protein and explicit water atoms)[39].

A seminal simulation work reported that water dipole moments within the shell of 3 Å thickness exhibited a more perpendicular alignment ordering with regard to the normal vectors from the lysozyme surface and that the water dipole was electrostatically driven to orient towards the line of electric field originating from the protein atoms (**Figure 1-13**)[39]. The long-range electrostatic interactions also lead to the heterogeneous orientation of the hydration water depending on the electrostatic potential of the local protein surface[110]. With the recent advancement of the Voronoi analysis of water shells, the long-range ordering influence of the protein on the structure of solvent distant from the protein was further demonstrated[47, 49, 111]. The parallelization of water dipole to the electric field outside the first HL was implied to be achieved by water-water interactions from shell to shell, instead of by the dielectric response of the solvent to the decaying electric field of the protein[111]. Similarly, for water outside the first HL, the tendency of water dipoles to be aligned with the electric field is concluded as an intrinsic property of bulk water[47, 49]. In the state-of-the-art experimental and computational studies of chiral sum frequency generation (chiral SFG) spectroscopy[113], Konstantinovsky *et al*. determined that the ordering of water dipoles extends 30−40 Å away from the protein and gradually fades into random arrangements beyond 40 Å. Besides, the much less symmetrical ordering within the first HL relative to that beyond provides the chiral SFG sensitivity to the water in the first hydration shell.

To summarize, all the reports introduced here suggest a strong relationship, albeit not quantitatively clear enough, between water dipole and the electric field (especially

from protein). In addition, the above investigations mainly paid attention to the static alignment (the angle) of the water dipole to the electric field generated from protein, but not to its effect on the dynamics of the water dipole. The very few works on investigating the effect of electric field alignment on these dynamics were made by probing the ^{17}O nuclear quadrupole relaxation (NQR), a superposition of the second-rank Legendre polynomials which describe the rotation of the electric field gradient[46, 112]. However, the authors themselves reported that this technique was incapable of calculating the strong retardation of the water rotation caused by an electric field exerted by the protein interface[46, 112]. Using a similar method, the later study further concluded that the short-range mechanisms in the first hydration shell of protein dwarfed the alignment effect of the electric field on the rank-1 time correlation functions (TCF)[48].

1.4 Water dynamics in cryoprotectant solutions

Among the PAs, sugars, the nonpermeable additives, have been widely employed to protect protein drugs and food during the freezing and drying processes. The sugar-water interactions are also deemed to be responsible for its unique protection mechanism[1-5]. Therefore, not only the inspection of the sugar effect on water dynamics helps the understanding of its protection mechanism in protein stabilization, but the water dynamics in aqueous sugar solutions can also be extrapolated to the study of that surrounding a large biomolecule by a reductionist approach[4].

1.4.1 Retardation of water relaxation by sugars

It has reached a consensus that water dynamics close to sugars are retarded relative to bulk water, by a variety of experiments[114-119] and MD simulations[120-122]. However, current studies disagree on to what extent the sugar solute can retard the surrounding water molecules and whether the disaccharide trehalose perturbs the water dynamics to a larger extent than other monosaccharides or disaccharides.

Ultrafast fluorescence spectroscopy experiments[114] suggested that trehalose slowed down the solvation dynamics more effectively than sucrose. The computation studies by MD simulation[120-122] showed that trehalose formed more HBs with water molecules than both sucrose and maltose did. A femtosecond infrared spectroscopy measurement[115] found water rotation motion was more strongly retarded in trehalose

solutions than that in glucose or sorbitol solutions. Recent broadband dielectric spectroscopy (BDS) experiments[116-117] also illustrated that trehalose induced a stronger slowdown of water dynamics than glucose and sucrose. The time-dependent Stokes shift (TDSS) measurements[118] suggested that sucrose slowed down the translational diffusion in greater magnitude with respect to trehalose and maltose. However, the depolarized light scattering[123] did not detect any obvious difference between trehalose and other monosaccharides or disaccharides. A later MD simulation[122] calculated the intermolecular nuclear Overhauser effect (NOE), nuclear quadrupole relaxation (NQR), and dielectric relaxation spectroscopy (DS), and concluded that trehalose slowed down the water dynamics a little bit stronger than other sugars. The very recent wide-angle X-ray scattering measurement[119] showed that trehalose preferred a more disordered arrangement, the bulky arrangement, which could disturb the water dynamics more effectively compared to sucrose. Therefore, whether trehalose is more effective in slowing down water dynamics is still highly debatable.

The BDS experiments[116-117] on the aqueous sucrose and trehalose solutions performed by Shiraga et.al showed that the HL extended up to ~ 3.5Å (corresponding to the average hydration number, $N_{hyd} \approx 25$) from the sugar surface and that the average retardation factor ranged from 2.4 to 5.5 depending on the disaccharide concentration. The NMR measurements[124] found a modest retardation factor of merely 1.67, by assuming the HL was comprised of 47 water molecules. The EDLS experiments[68, 123] suggested a higher retardation factor of 5-6, which is quite independent of the sugar species (fructose, glucose, trehalose, and sucrose) and the sugar concentrations. The HL is around 3.1 Å from the solute surface, which corresponds to the hydration number of 16 for glucose and 25 for trehalose. A femtosecond mid-infrared study[115] suggested that the hydration number for glucose, trehalose, and sorbitol was 24 ± 3, 46 ± 5, and 21 ± 3, respectively. The retardation factor was found to be 1.65 ± 0.15 for all three sugars, similar to the NMR measurements. It should be pointed out that the deduction of such results was aided by a probability model where water molecules were assigned to zero, one, two, or three HLs according to their increasing relaxation times, since they believed that water within the HLs of multiple sugar molecules rotated 2.7 times slower.

1.4.2 Random close-to-contact condition

As a kind of small molecule compared with proteins, sugar molecule has a higher specific surface area and a higher solubility. Although different experimental techniques have given the various number of hydration water (N_{hyd}), they all have suggested that N_{hyd} decreases with the increasing sugar biomolecule solution[68, 116-117, 123], probably due to the HL overlapping and solute aggregation both occurring in concentrated solutions and releasing the hydration water into the bulk. It is straightforward to understand that the latter phenomenon happens in the higher concentrated case.

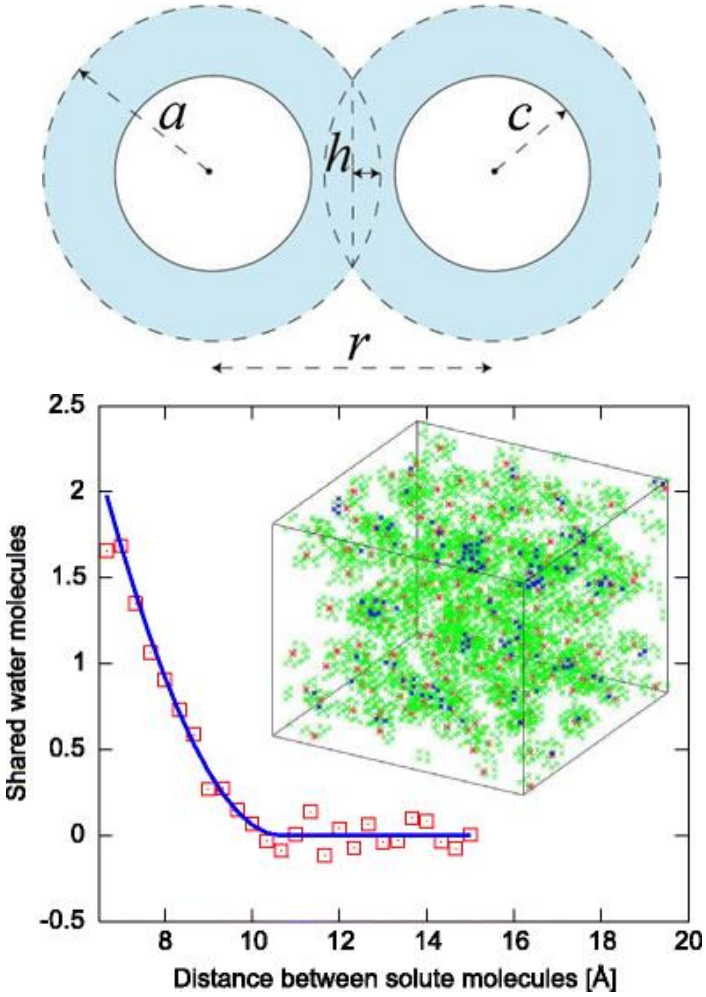

Figure 1-14. (**Top**) Schematical diagram of the solvent-sharing model. The radius of the spherical solute molecule is *c*, the external radius of the HL is *a*, the intermolecular distance is *r*, and the height of two spherical caps is $h = a\text{-}r/2$[125]. (**Bottom**) The number of shared hydration water molecules of glucose molecules was obtained from MD simulations (squares) and the geometric model (solid line). The inset shows the glucose molecules (red), hydration water (green), and shared hydration water (blue) in the simulation box at the solute mass percent of 9.1 wt%[125].

To discriminate these two factors, Fioretto *et al*.[125] propose a solvent-sharing model for small molecules, where hydration water is merely subject to the *finite probability of close-to-contact condition among different solute molecules*. This simple geometric model (**Figure 1-14 (top)**) which assumes the small solute molecule to be spherical estimates the hydration number through the predefined hydration thickness, sugar concentration, and number density of water. Once the intermolecular distance among sugars is obtained from the concentration or molar ratio, the number of hydration water shared by these two solute molecules can be deduced from the overlapping volume and water number density. They also calculated the number of the shared hydration water molecules from the MD simulation, by assuming the HL ~ 2Å and counting the water molecules belonging to different solute molecules[125]. The corresponding results versus the distance between the solute molecules are shown in **Figure 1-14** (**bottom**). Here, the glucose molecule has been modeled as a sphere of radius c=3.41 Å. The results from both MD and the analytic model show that a hydration water molecule may be shared among different solute HLs even in this pretty dilute case (9.1 wt%).

In addition to sugars, the slowdown effect of several zwitterionic molecules has also been studied, such as triethylamine N-oxide (TEAO), TMAO, and 4-methylmorpholino N-oxide (MMO)[4, 37]. Despite that all these works explored how different additives influence water dynamics through water-structure-making or water-structure-breaking effect[37], the practical principles for retarding water rotational dynamics by protective agents are not sufficiently discussed. For developing effective protective agents, a systematic investigation to extract the essential and specific parameters is of great significance.

1.5 Predictive role of MD simulation

As is known to all, the experiment techniques can provide valuable insights into certain aspects of water dynamics, such as HBs[115], water dielectric relaxation time[116-117], and water self-diffusion coefficients[126]. However, due to the general limitation in their temporal and spatial resolution, the output is usually the collective response or ensemble average properties, instead of the individual water dynamics or the site-specific water-biomolecule interactions. In contrast, the MD simulation can record the behavior of water, protective agents, and protein molecules at the atom level. In consequence, this computation tool has been widely implemented to unravel the

mystery of water motions and the water-biomolecule interplay with sufficient temporal and spatial resolution. With the recent rapid advancement of computer hardware, parallel algorithm, and force field development, the MD simulation has evolved to be more precise and the longer simulation span on the scale of milliseconds is also realizable[2]. Therefore, MD simulation should work as a powerful tool in predicting the effect of additives and designing the protein stabilizer. **Figure 1-15** summarizes the spatial and temporal scales of MD simulation for biopreservation and other relative fields.

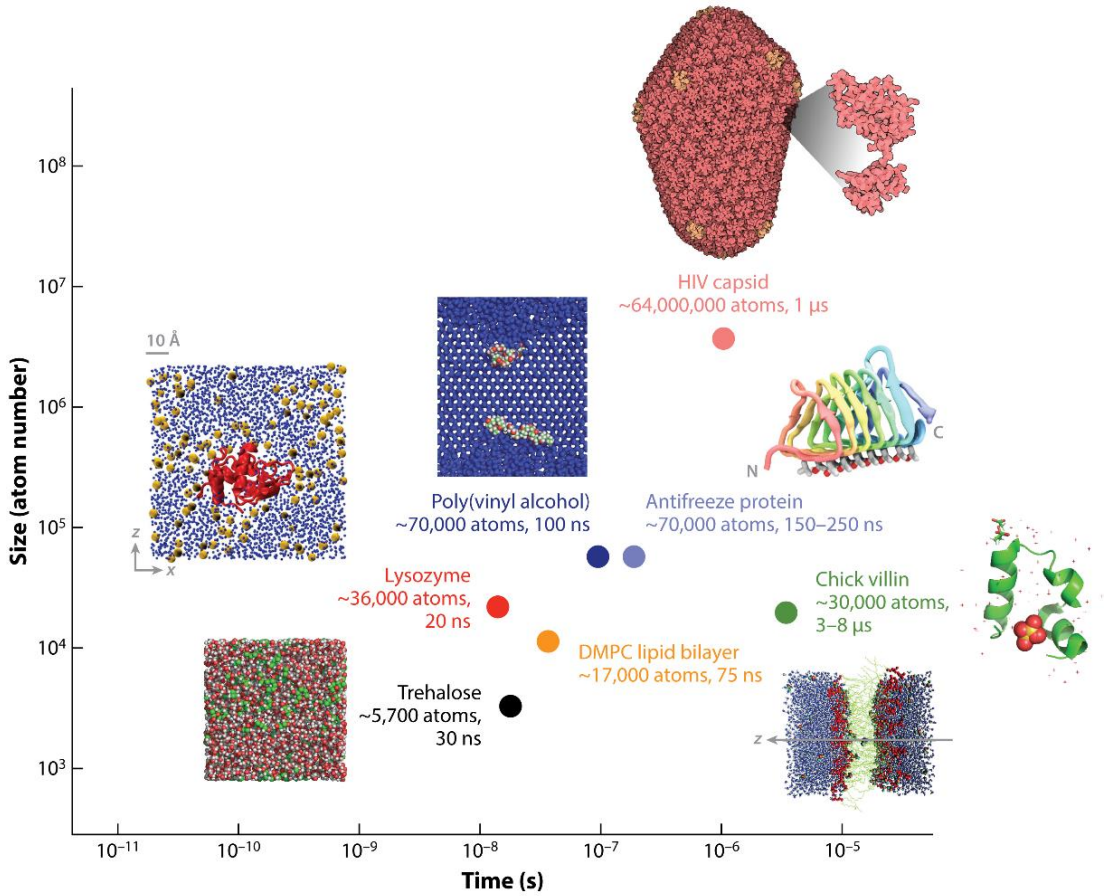

Figure 1-15. Solute size and timescales covered by representative MD simulations for biopreservation, along with state-of-the-art biomolecular simulations beyond the field. Abbreviation: DPMC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine[2].

1.6 Motivation and objective of this Dissertation

In order to better predict the experiment outcome and facilitate the design of the protein stabilizer, the motivation of this dissertation is to have a better understanding of the water rotational dynamics surrounding proteins and how the protective additives slow down the water dynamics, from a molecular level.

Specifically, the main objectives of this dissertation are highlighted as follows:

1) To reveal the molecular nature behind the dominant dielectric response measured by broadband dielectric spectroscopy (the γ dielectric dispersion).

2) To depict the probability distribution of water molecular rotational relaxation time (heterogeneity of individual water dynamics) and illustrate how far proteins can perturb the water dynamics by mapping the spatial dependence of water relaxation time.

3) To probe the molecular mechanism of water rotational dynamics from the perspective of electrostatic interactions, whereby the role of protein in water rotational relaxation and the correlation between water rotational and translational dynamics can be analyzed, and to disentangle the myth over the local hydration dynamics.

4) To shed some light on the dependence of water relaxation time on its residence time within the HL and on the solute dipole in aqueous sugar solutions, in order to provide guidance for controlling water rotational relxation.

1.7 Organization of this Dissertation

In chapter 1, the background of this work, including the dominant role of hydration water in protein conformation, water rotational dynamics in aqueous protein/sugar solutions, and the four prevailing hypotheses about the protection mechanism of protective agents (sugars here) in various preservation methods, is introduced. Meanwhile, chapter 1 also summarizes the recent progress in probing the underlying mechanism of water rotational dynamics and the slowdown effect of a variety of sugar molecules, by differing experimental techniques and MD simulations. The practicability of MD simulation in designing the protective additives for therapeutic proteins is also reviewed briefly wherein.

In chapter 2, the molecular water behavior behind the macroscopic dielectric measurement in the high-frequency domain is investigated. The dielectric spectra of various concentrations of protein solutions are derived from the relaxation times of all water molecules by MD simulation and measured by broadband dielectric spectroscopy (BDS). Comparison of the dielectric relaxation time between the calculated value from MD and the measured one by BDS reveals the molecular origin of the dielectric spectra in the gigahertz domain (γ relaxation in **Figure 1-16 (a)**).

In chapter 3, the preliminary understanding of individual water rotational dynamics in protein solutions is explored by MD simulation of lysozyme solutions. The probability density function of the rotational relaxation time of all water molecules in lysozyme solutions is calculated without predefining the "HLs", whereby the empirical or adjusted hydration shells are excluded. Based on that, the spatial distribution of water relaxation time is illustrated. Practically, MD simulations of the pure water system and lysozyme-water systems of various concentrations in the equilibrium state are carried out. From the trajectories and relaxation times of all water molecules, the two-dimensional probability distribution of water at a distance from the lysozyme surface with a rotational relaxation time can be derived. Based on this 2-D probability distribution, the expected value of water rotational relaxation time as a function of the distance from the lysozyme surface (the distance, *x*, in **Figure 1-16 (b)**) is obtained, from which the average retardation factor relative to bulk water and the thickness of the first HL are discussed ultimately.

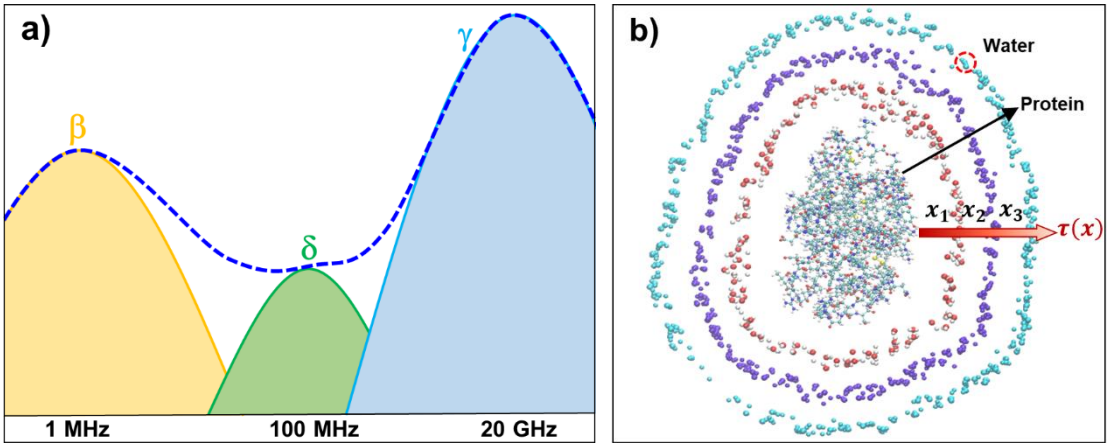

Figure 1-16. **(a)** Schematic diagram of the spectrum of the imaginary part of the complex permittivity (dielectric loss spectrum) of the solution (break line); β, δ and γ are Debye peaks decomposed from the dielectric spectrum[57-58, 71, 73, 127]. **(b)** Schematic representation of partial water molecules surrounding lysozyme.

In chapter 4, the mechanism of water molecular rotational dynamics from the perspective of electrostatic interactions is probed by MD simulation. The reorientation of the exerted electric field on the water molecule is analyzed and compared with its rotational dynamics. The relaxation time distribution of the local electric field (electric field at interior fixed positions) reorientation is also computed meanwhile. This picture provides insight into the local hydration dynamics. The discrepant water rotation mechanisms within the HL and outside of the HL are illustrated. This observation leads

to further discussion about the dependence of water relaxation time on its residence time, which provides a basic understanding of controlling water rotational dynamics surrounding proteins.

In chapter 5, water dynamics in aqueous PAs solutions (sugars and sugar alcohol here) are studied based on the findings in the above chapters, by MD simulation and BDS measurement. The probability distributions of both water residence time and water relaxation time are depicted in an array of sugar and sugar alcohol solutions (glucose, trehalose, sucrose, maltose, and sorbitol). Water molecular dynamics in these PAs solutions are compared with that in protein solutions. Moreover, through a systematic study of water residence time distribution with a geometric model, this chapter demonstrates how the solute size (specific surface area) governs water residence time. The effect of the HL overlapping on water residence is also discussed. Meanwhile, the summary of the relationship between water relaxation time and its total residence time indicates the importance of solute dipole moment in slowing down water dynamics. Finally, the effect of balance between PA dipole moment and its size on water dynamics is discussed. The basic laws acquired in this chapter for controlling water dynamics might have a meaningful impact on the development of protein stabilizers.

The main conclusions of the present work as well as the future perspectives are summarized in the last session of this dissertation.

Chapter 2 Dielectric Relaxation

The dielectric spectroscopy technique has been widely employed to characterize different food products. As introduced in **Sec. 1.2.1**, a previous work in the current lab found that the deterioration rate of lactose dehydrogenase (LDH) in solution has the power-law relationships with the dielectric relaxation times of water in the solution, particularly sensitive to the γ relaxation time (20 GHz)[33]. This chapter reveals the molecular water behavior behind the macroscopic dielectric time constant.

2.1 MD simulation and post-analysis theory

2.1.1 MD simulation setup

Molecular dynamics (MD) simulations were carried out for lysozyme aqueous solutions and pure water employing an extended point charge water model (SPC/E[128]) with periodic boundary conditions by the GROMACS[129] package (Version: 2019.4). For protein-water systems, the initial configuration of lysozyme (molecular weight: 14 kDa, total charge: +8e) was obtained from the crystallographic structures with PDB code 1AKI, and the OPLS-AA force field[130] was used in the simulation. The protein was solvated in a cubic box of length ~9 nm containing several lysozymes and ~20000-24000 water molecules, corresponding to effective mass concentrations ranging from 33 to 235 mg/mL. The protonation state of the amino acids corresponds to a pH of 7. Next, chloride ions of 8 times lysozyme molecules were added to neutralize the system. First, the potential energy of the whole system was minimized using the steepest descent algorithm. Then, the system was equilibrated in the NVT ensemble for 10 ns at constant temperature (298.15 K), followed by equilibration in the NPT ensemble at constant temperature (298.15 K) and pressure (1 bar) for another 10 ns. During the equilibration processes, the positions of protein-heavy atoms were restrained with the Lincs algorithm[131]. Finally, the production runs were performed in the NPT ensemble (298.15 K and 1 bar) for 100 ns with a 2-fs time step, and the last 5 ns - continuous running molecular trajectories were taken as sample populations for further analysis. A cutoff radius of 1 nm was used for the nonbonded Lennard−Jones interactions and neighbor searching, which was updated every 20 fs. The modified Berendsen thermostat[132] (coupling time constant of 0.1 ps) and the Parrinello-Rahman barostat[133] (coupling time constant of 2.0 ps) were used to keep the temperature and pressure constant,

respectively. The electrostatic interactions were calculated by the Particle Mesh Ewald (PME) method[134-135] with a Fourier grid spacing of 0.16 nm and a 1 nm cutoff for the short-range part in real space. For pure water, 11225 water molecules were included in an ~7 nm cubic box, and the simulation protocols were completely the same as the protocols implemented in the protein solution case.

2.1.2 Water rotational relaxation time

The rotational dynamics of a single water molecule was investigated by measuring the time evolutionary reorientation of its dipole vector (symmetrical axis vector of a water molecule) ***u***, i.e., the dipole moment time autocorrelation function (ACF), which is defined as

$$\psi(t) = \frac{\langle \boldsymbol{u}(0) \cdot \boldsymbol{u}(t) \rangle}{\langle \boldsymbol{u}(0) \cdot \boldsymbol{u}(0) \rangle}, \tag{2-1}$$

where dipole moment vector ***u(t)*** is the summation of single water OH-bond vector ***e*** at time *t*. Here, dipole moment ***u*** was calculated for all the water molecules in the lysozyme aqueous solution within a 5 ns continuous dynamic trajectory. Since water rotation should depend on the position of water relative to the lysozyme surface, the present work took the water position data relative to the lysozyme surface to the end of the calculation. The dipole moment ACF obtained was then approximated with a multi-decaying exponential by the least squares method to extract the rotational relaxation time constants, as expressed by **Eq. (2-2)**. The optimized fitting was derived based on a combination of a correlated function (R^2) and a two-tailed significance test (Student's t-test; $|t| > 2$) to guarantee the reliability of the calculated relaxation times. The number of the exponential curve was increased one by one starting from a single exponential curve until the *t* value of any of the fitting parameters became smaller than 2.

$$\psi(t) = \sum_{k=1} A_{D,k} \exp(-\Delta\tau / \tau_{D,k}) \tag{2-2}$$

Here, A_k and τ_k are the amplitude and time constant of the k^{th} relaxation mode, respectively. The subscript '$_D$' represents the water molecular dipole which corresponds to the symmetric axis of the water molecule, and this subscript is used to distinguish from the reorientation process of the electric field discussed in the following chapter. Note that the summation of A_k is equal to 1. The average rotational relaxation time of each water molecule was calculated as[95]

$$\tau_{D,\text{int}} = \sum_k A_{D,k} \cdot \tau_{D,k}. \tag{2-3}$$

2.1.3 Theory of calculation of dielectric loss from MD

Statistically, the dielectric spectrum is calculated by Fourier transformation of the ACFs of the total dipole moments in the measured material[107, 136]. Thus, the correlation function of the collective dipole moment, which is composed of all water and lysozyme molecules (**Eq**. **(2-4)**), should be calculated to obtain the dielectric spectrum of the lysozyme solution.

$$\boldsymbol{M}_t(t) = \sum_j \boldsymbol{M}_j(t) + \sum_l \boldsymbol{M}_{lys,l}(t) \tag{2-4}$$

Here, $\boldsymbol{M_t}(t)$ is the total dipole moment of the solution at time *t*, obtained from the vector summation of all molecular dipoles present in the solution, i.e. those of water, yielding $M_j(t) = \sum \mu_j(t)$, and those of the protein molecules, $M_{lys,l}(t) = \sum \mu_{lys,l}(t)$. Subscript *j* is a certain single water molecule, and *lys,l* is a certain single lysozyme molecule. This correlation function $\langle \boldsymbol{M}_t(0)\boldsymbol{M}_t(t)\rangle$ has five terms, namely, ACF of water, ACF of lysozyme, cross correlation function (CCF) between two water molecules, CCF between two lysozyme molecules, and CCF between lysozyme and water molecules.

$$\begin{aligned} <\boldsymbol{M}_t(0)\boldsymbol{M}_t(t)> = & \sum_j <\boldsymbol{M}_j(0)\boldsymbol{M}_j(t)> + \sum_{j\neq k} <\boldsymbol{M}_j(0)\boldsymbol{M}_k(0)> \\ & + \sum_{j,l} <\boldsymbol{M}_j(0)\boldsymbol{M}_{lys,l}(t)> + \sum_l <\boldsymbol{M}_{lys,l}(0)\boldsymbol{M}_{lys,l}(t)> \\ & + \sum_{l\neq m} <\boldsymbol{M}_{lys,l}(0)\boldsymbol{M}_{lys,m}(t)> \end{aligned} \tag{2-5}$$

Because the focus here is paid on the dynamics of all solvents in the system, only the ACF of water (the first term on the right-hand side) was calculated to investigate its contribution to the dielectric spectrum (neglecting other terms). Therefore, the imaginary part of the dielectric spectrum is

$$\varepsilon'' = \frac{1}{N}\sum_j \sum_k \frac{A_{j,k}\cdot(\omega\tau_{j,k})}{1+(\omega\tau_{j,k})^2}, \tag{2-6}$$

where ω is the circular frequency. Note that the imaginary part of dielectric relaxation strength $\Delta\varepsilon$ was not presented here.

2.2 Experimental method: Broadband dielectric spectroscopy

2.2.1 Experiment apparatus

The dielectric spectra of lysozyme solutions in the frequency range from 5 kHz to 43.5 GHz were measured using two impedance analyzers (4294A, Agilent Technologies, available frequency range of 40 Hz to 110 MHz and E4991B, Keysight Technologies, 1 MHz to 3 GHz) and a vector network analyzer (N5224A, Agilent Technologies, 10 MHz to 43.5 GHz). These three instruments are connected to an open-ended coaxial probe with a diameter of 1.2 mm via a multiport coaxial switch (8767M, Keysight Technologies, DC to 50 GHz) and coaxial flexible cable (FC-182, Flexco Microwave, DC to 40 GHz). For the calibration of measurement, several standard materials with known dielectric properties, air, 2-propanol[137], ethylene glycol[137], and water[136], were used. The sample temperature was controlled at 25.0 ± 0.1 °C by circulating water in the sample cell holder from a thermostatic bath (TRL-108H, THOMAS KAGAKU).

2.2.2 Sample preparation

Solute lysozyme powder from chicken egg white was purchased from Sigma–Aldrich (62970, ~100000 U/mg) and used without further purification. Lysozyme solutions of 34-284 mg/mL were prepared by dissolving the solute in ultrapure water (Simplicity UV, Merck).

2.2.3 Calculation of dielectric loss ε'' from measured real part ε' using the Kramers-Kronig relations

The imaginary part $\varepsilon^{"}$ (dielectric loss) of the complex permittivity, $\varepsilon^* = \varepsilon' - i\varepsilon''$ (i is the imaginary unit), was calculated from the measured real part ε' using the Kramers–Kronig transform[107, 136-141]. This transformation from the real part ε' allows the elimination of the effect of static electrical conductivity, which often overlaps with the dielectric loss peaks of water[107, 136]. Concretely, Kramers-Kronig (KK) relations provide a physical connection between the real and imaginary parts of an analytic signal that describes causal processes, and they are commonly used in the study of optical materials[142], where the refractive phenomena are derived by the measurements of absorptive phenomena or vice versa. For the complex permittivity $\varepsilon^* = \varepsilon' - i\varepsilon''$ in the

presence of static electrical conductivity σ, the real and imaginary parts are connected by KK relations as expressed in **Eqs**. **(2-7)** and **(2-8)**[138].

$$\varepsilon'(\omega) = \varepsilon_\infty + \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\varepsilon''(\omega')}{\omega' - \omega} \mathrm{d}\omega' \tag{2-7}$$

$$\varepsilon''(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\varepsilon'(\omega') - \varepsilon_\infty}{\omega' - \omega} \mathrm{d}\omega' + \frac{\sigma}{\varepsilon_0 \omega} \tag{2-8}$$

where ω is the angular frequency, ε_∞ is the high-frequency limit of the complex dielectric permittivity, and ε_0 is the dielectric permittivity of vacuum. The symbol *P* in **Eqs**. **(2-7)** and **(2-8)** denotes the Cauchy principal value of the integral. **Eqs**. **(2-7)** and **(2-8)** show that $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ are the Hilbert transform pairs. The Hilbert transform of $x(t)$ as a function of time t is defined as **Eq. (2-9)**[141].

$$\hat{x}(t) = \frac{1}{\pi t} * x(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\lambda)}{t - \lambda} \mathrm{d}\lambda \tag{2-9}$$

Here, symbol * indicates the convolution operation and ^ indicates a Hilbert-transformed function. From **Eqs**.**(2-8)** and **(2-9)**, dielectric loss without the contribution of electrical conductivity, $\varepsilon''(\omega) - \sigma/(\varepsilon_0 \omega)$, can be obtained by the Hilbert transform of $\varepsilon'(\omega)$[143].

2.2.4 Decomposition of the measured dielectric loss spectra

The obtained dielectric loss spectra were decomposed into Debye relaxation functions of several relaxation dispersions by the least square approximation.

$$\varepsilon'' = \sum_n \frac{\Delta\varepsilon_n \cdot (\omega\tau_n)}{1 + (\omega\tau_n)^2} \tag{2-10}$$

Each Debye relaxation has a relaxation time, τ_n, and relaxation strength, $\Delta\varepsilon_n$. In the present study, the Bayesian information criterion (BIC) was used to determine the number of relaxation terms on the right-hand side in **Eq**. **(2-10)** where the value of BIC in the approximation is preferred to be the minimum in the model selection[144-145].

2.3 Dielectric loss spectra

2.3.1 Probability distribution of water relaxation time

As expressed by **Eq. (2-6)**, the dielectric loss spectrum is obtained from the rotational ACF of all water molecules, and the rotational relaxation times of all individual water molecules are first calculated. **Figure 2-1** shows the probability density of the rotational relaxation time, $P(\tau)$, of water molecules in a pure water system and in lysozyme solutions with various protein mass concentrations. In the inset of (**Figure 2-1** (**a**)), the PDF of pure water exhibits a sharp narrow distribution and is perfectly fitted to a single Gaussian distribution, where the averaged rotational relaxati-

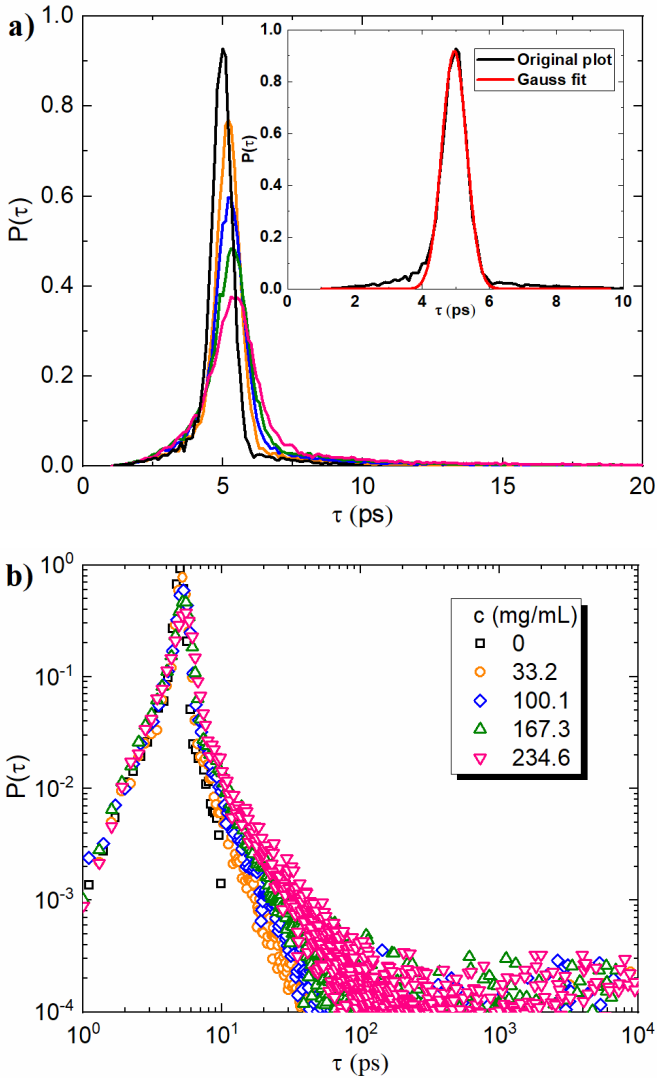

Figure 2-1. Probability density of water relaxation time in pure water and lysozyme solutions with various concentrations with abscissa in (**a**) linear and (**b**) logarithmic scale. The inset is the Gaussian fit (red) result of the probability distribution of relaxation time (black) for pure water. Panel (a) and (b) share the same legend in color.

-on time is 4.96 ps with a small deviation of 0.42 ps, indicating the homogeneous relaxation behavior of pure water. However, water molecules influenced by lysozyme have a broad spectrum of rotational relaxation times, and the relaxation time distribution is distorted to the longer relaxation time side from a symmetric Gaussian, principally because of the increasing ratio of hydration water with longer rotational relaxation time. However, at a relatively lower concentration (<100.1 mg/mL), the probability density of relaxation time can still be well fitted by Gaussian (results not shown here), possibly because only a small portion of water in the solution is hydration water at lower lysozyme concentrations. Hence, the PDF is not distorted too much from the Gaussian distribution. Upon increasing the solute concentration again, this relationship between PDF and Gaussian distribution never holds.

2.3.2 Dielectric loss spectra calculated from MD

Dielectric loss spectra calculated by the solvent component from MD simulation (**Eq. (2-6)**) for pure water (inset) and aqueous lysozyme solutions are shown with black solid lines in **Figure 2-2**, where the red (blue for pure water) dashed-dotted lines indicate the corresponding fitting curves by a single Debye function. The dielectric loss spectrum of pure water overlaps perfectly with the single Debye curve with a dielectric relaxation time of 4.96 ps, which is identical to the mean value of the relaxation time distribution of pure water. Thus, the dielectric relaxation property is proven to be exactly the macroscopic consequence of the molecular rotational relaxation, and the cross correlation among water molecules is negligible for pure water. For dilute lysozyme solutions (< 100.1 mg/mL), as shown in **Figure 2-2** (**a**) and (**b**), the calculated dielectric loss is well fitted by the single Debye function (**Eq. (2-10)**) with a negligibly small residual error, probably because the relaxation time distribution used for calculating dielectric loss is not a complete normal distribution. When the solute density increases, this residual error derived from the fitting with the Debye function becomes larger, especially in the lower frequency region. However, the measured dielectric loss spectrum has a single peak without any shoulder-like shape, even in the most concentrated case.

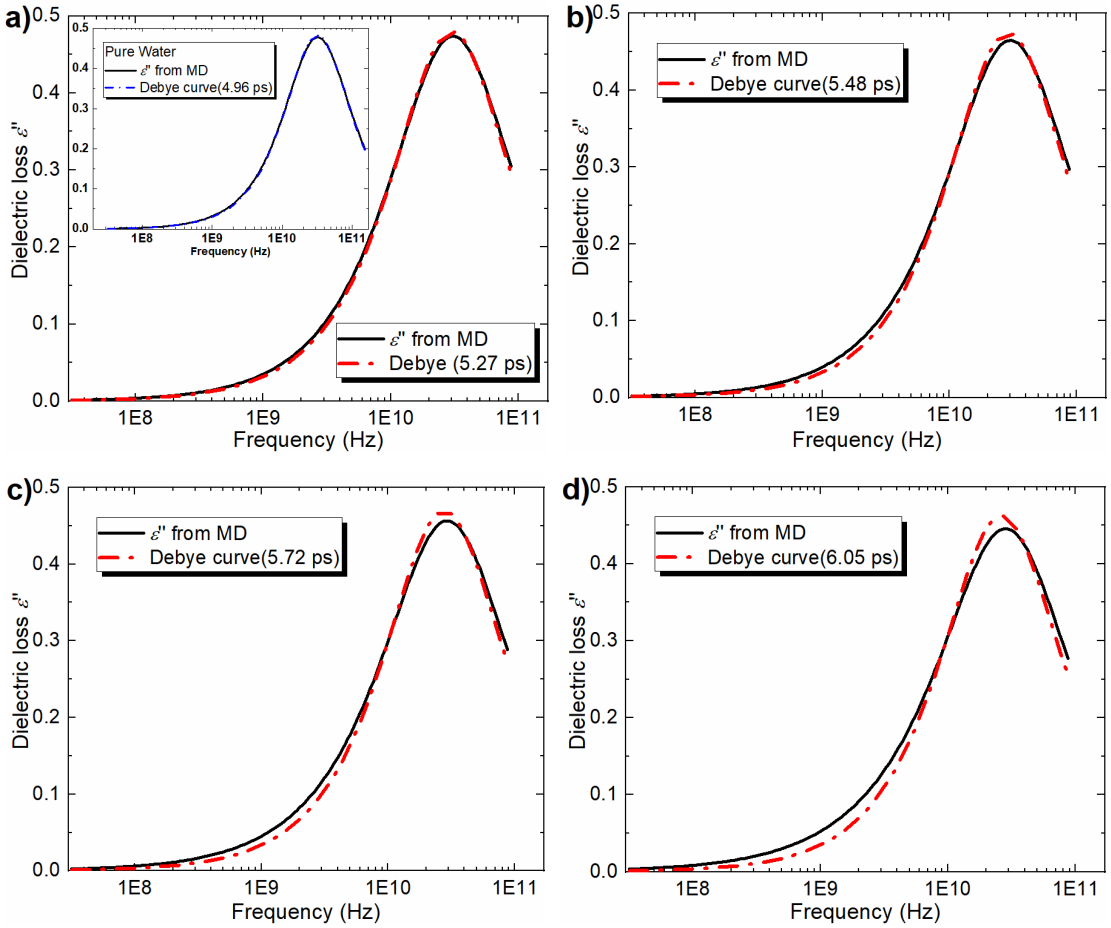

Figure 2-2. Dielectric loss spectra of water in various concentrations of protein solutions from MD simulation using **Eq. (2-6)**. Debye fit (red or blue dashed-dotted line) is also shown for each case with its relaxation time. **(a)** 33.2 mg/mL; inset indicates pure water, **(b)** 100.1 mg/mL, **(c)** 167.3 mg/mL, and **(d)** 234.6 mg/mL.

2.3.3 Dielectric loss spectra measured by BDS

In comparison, **Figure 2-3** and **Figure 2-4** show the dielectric spectra, measured by the BDS experiments. **Figure 2-3** is the real part of the permittivity. In the low-frequency range, the effect of electrode polarization (EP), which is induced by the ions accumulating on the electrode/sample interface[107, 136, 146], is dominant. Since EP often exists around 10 MHz, the decomposition of so-called β relaxation related to rotational dynamics of the protein[57-58, 73, 127] in this frequency range is hindered. In the present study, however, the β relaxation is not the target relaxation.

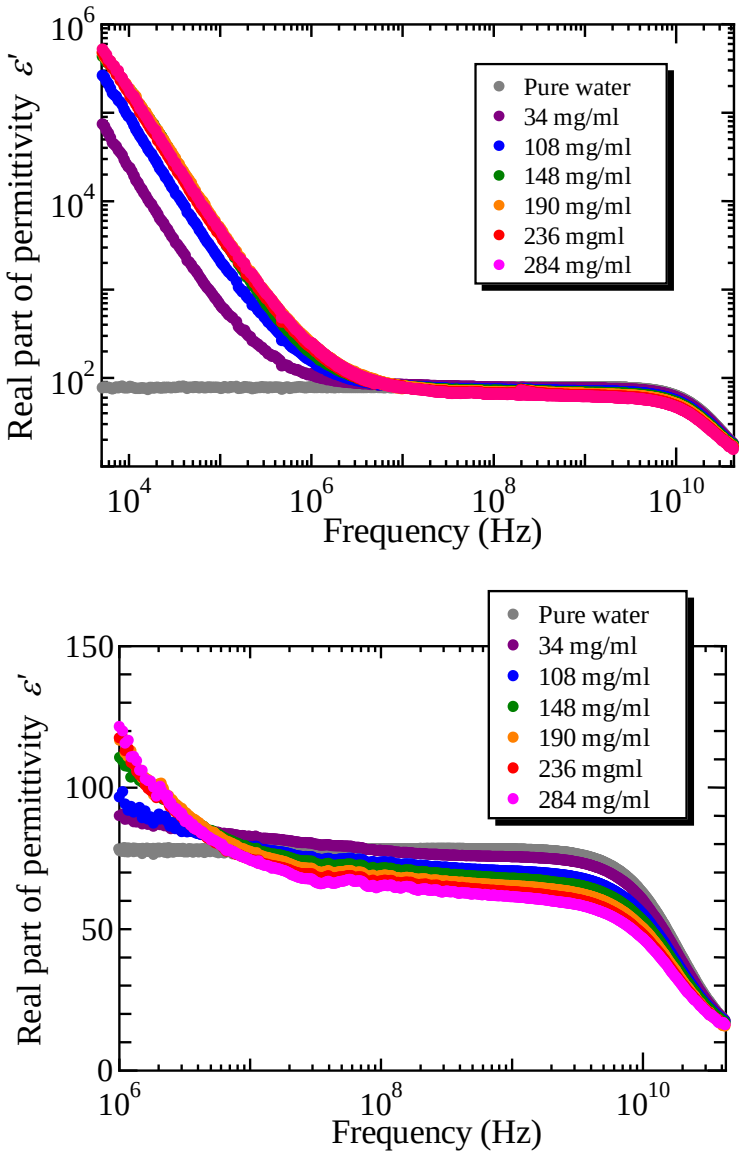

Figure 2-3. Dielectric dispersion obtained in the experiment. (**Top** panel) Raw data measured by the impedance analyzers and the vector network analyzer (25.0 °C). (**Bottom** panel) Dielectric dispersion after removing EP contribution $\varepsilon_{EP} = A\omega^{-\lambda}$ (ω is the angular frequency; A and λ are the constants of the system.[107, 136, 146]).

Figure 2-4 compares the dielectric loss spectrum of lysozyme aqueous solution (34 mg/ml) obtained by the measurement (raw data) with that by the transformation from experimentally obtained real part ε' using KK relations. While they agree well in the high frequency region, there is a significant difference in the lower frequency region. This is because the contribution of the static electrical conductivity σ overlaps the dielectric relaxation of water for the raw data. On the other hand, the dielectric loss spectrum without the contribution of σ was obtained by the transformation using KK relations, which is convenient for the analysis of dielectric relaxation of water.

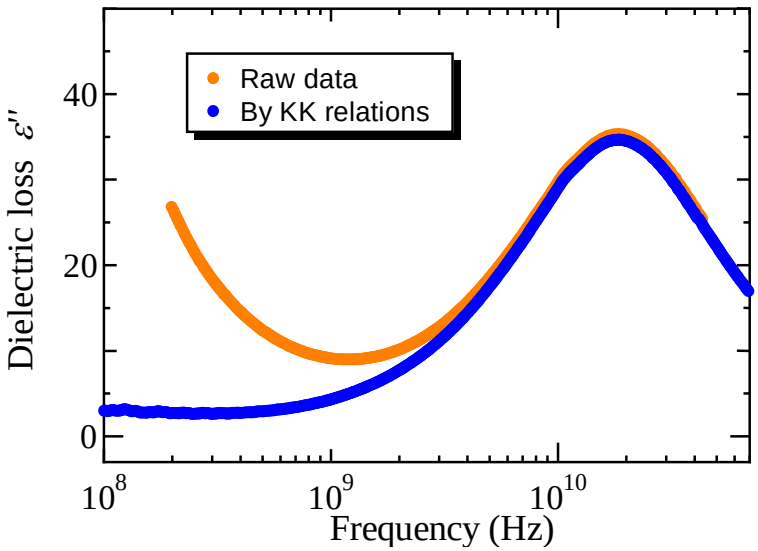

Figure 2-4. Comparison of the dielectric loss spectrum obtained by the measurement with the vector network analyzer (raw data) with that by the transformation from experimentally obtained real part ε' using KK relations for lysozyme aqueous solution with the concentration of 34 mg/ml at 25 °C.

Figure 2-5 summarizes the dielectric loss spectra of lysozyme aqueous solutions obtained by the transformation from experimentally obtained real parts ε' using KK transformation. At approximately 20 GHz, the fastest relaxation, γ process, attributed to the collective orientational motion of the dipole moments of water molecules[57-58, 73, 136] was observed.

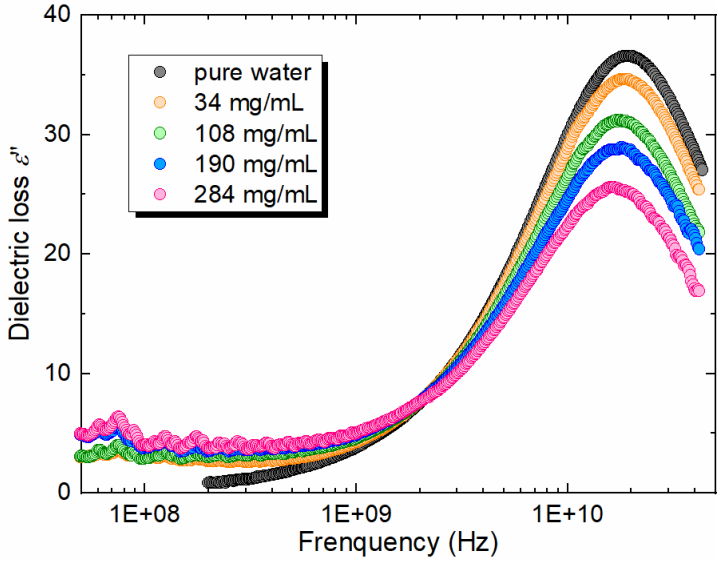

Figure 2-5. Dielectric loss spectra for lysozyme solutions of various concentrations at 25.0 ℃ by BDS.

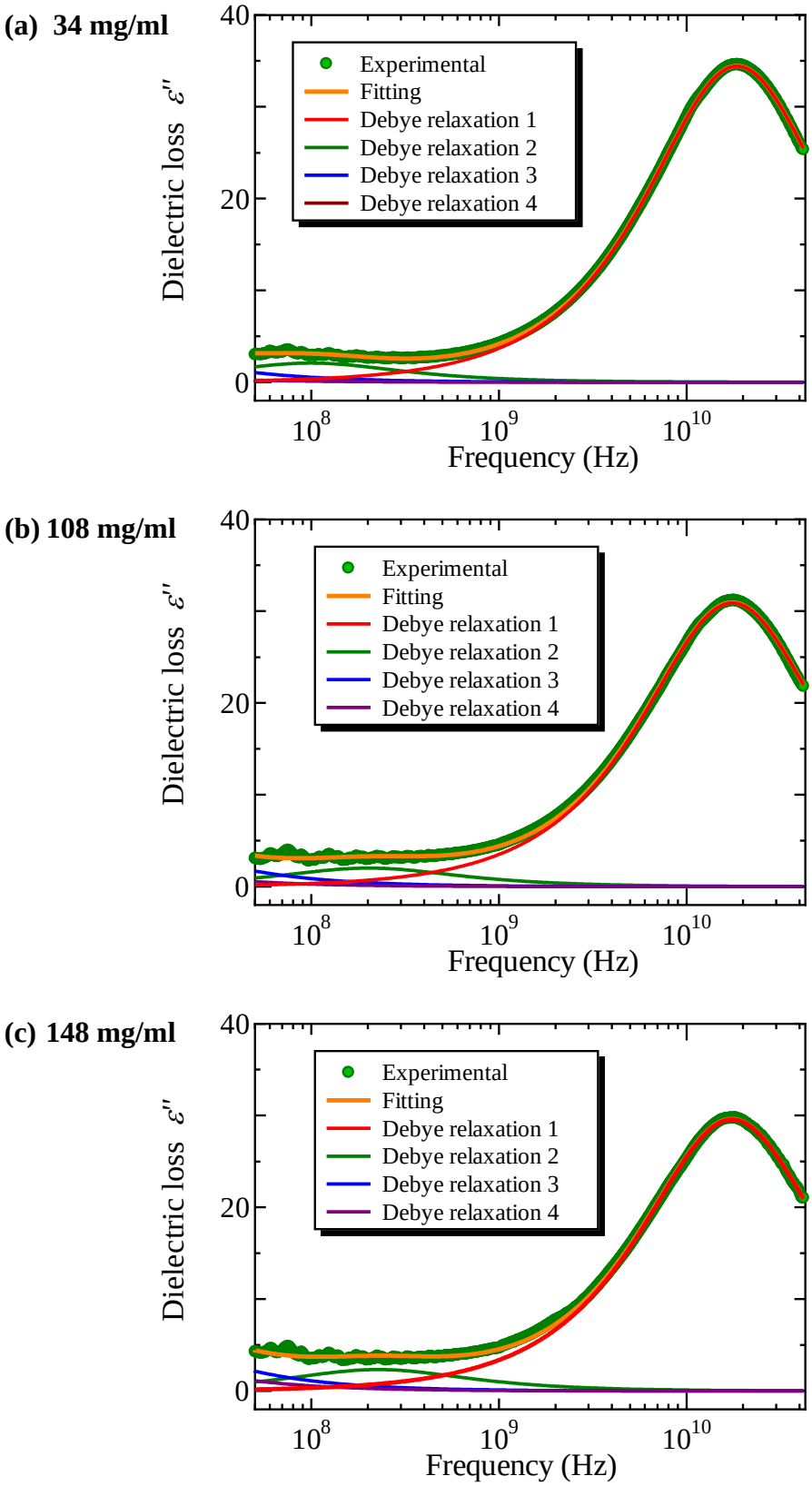

Figure 2-6. Decomposition of the dielectric loss spectra of lysozyme solutions to the Debye relaxation functions (by **Eq. (2-6)**; to be continued).

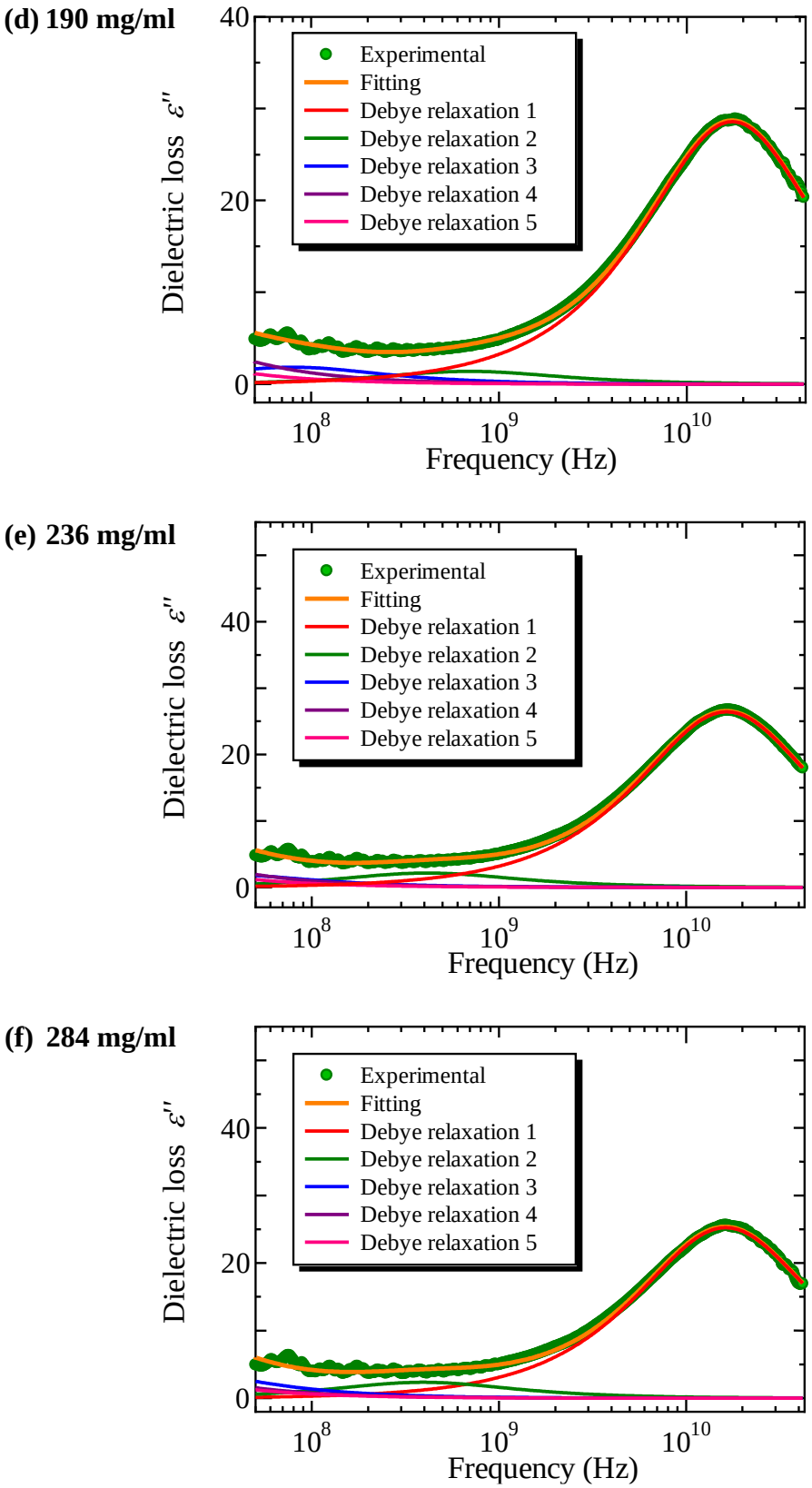

Figure 2-6. Decomposition of the dielectric loss spectra of lysozyme solutions to the Debye relaxation functions (by **Eq. (2-6)**). The spectrum summarizing all decomposed Debye relaxation curves is shown as an orange line (Fitting), and the γ relaxation is shown as a red line (Debye relaxation 1).

Figure 2-6 shows the decomposition of the dielectric loss spectrum for the various concentrations of lysozyme solutions. The dielectric loss spectra are identical to the sum of decomposed Debye-type relaxation curves. At all concentrations, γ relaxation at approximately 20 GHz is perfectly approximated by single Debye relaxation with the highest peak frequency (red curve (Debye relaxation 1)). A few small relaxations known as the δ process[57-58, 73], the molecular origin of which is still under discussion[71, 136, 147], are in the frequency range below γ relaxation. These relaxations have often been attributed to hydration water or bound water[57-58, 73, 136], but it is more accepted today to consider that the δ process is related to not only slowed water molecules but also other factors, including local peptide chains on the surface of protein molecules[136] and cross correlation between water and protein[71, 147-148]. According to this consideration, only the γ relaxation was focused on and compared with the MD simulation results in this study.

2.3.4 Comparison of the γ relaxation between MD and BDS

To investigate the correspondence relationship between the dielectric spectrum (measured by BDS) and the molecular rotational movement of water, the dielectric relaxation times derived from MD simulation (**Figure 2-2**) were compared with the γ relaxation time measured by BDS (red line in **Figure 2-6**) for various concentrations of lysozyme solution, as shown in **Figure 2-7**. In **Figure 2-7**, all the relaxation times are normalized by the relaxation time of pure water, respectively, i.e., 4.96 ps for the calculated value and 8.27 ps for the measured value. Both normalized relaxation times increase monotonically with the protein concentration, and they agree well within the maximum relative difference of 6%. This agreement unambiguously suggests that the distribution of rotational relaxation times of water molecules in **Figure 2-1** composes γ dielectric relaxation in lysozyme solutions. Since the rotational relaxation times are originally derived from the ACF of the water molecular dipole moment, it is further concluded that the collective effect of the self-correlation of water rotational movement is responsible for the dielectric loss peak in the gigahertz region (γ relaxation part) in protein solutions.

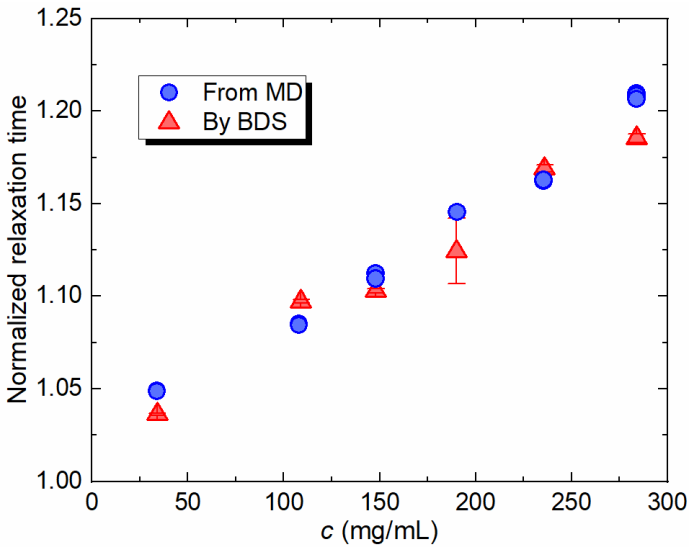

Figure 2-7. The normalized relaxation time of various concentrations of lysozyme solutions, in mg/mL. Blue circles depict the normalized dielectric relaxation time of water rotation calculated by **Eq. (2-6)**using the results from the MD simulation. The red triangle depicts the normalized g dielectric relaxation time measured by BDS with the error bar. Each relaxation time was normalized by the calculated or measured relaxation time of pure water.

Given the significant practicability of the dielectric spectrum in the bio-preservation industry, the successful link between molecular-level water rotation and dielectric relaxation may foster computational protein modeling to develop new proteins for biomedical and biotechnological applications. Moreover, it is noteworthy that the γ relaxation times presented here were derived from the dielectric loss spectra composed by the rotational autocorrelation of the single water molecule, rather than from the exponential fit of the collective ACF of defined HLs. Though, reproducing the δ dielectric relaxation will be intractable with this method. Besides, the approximation of the classical forcefield used in our MD simulation (SPC/E) has the disadvantages of the fixed charge points and the fixed distance between the charged points. The intramolecular vibration and bending that may affect the rotational relaxation time cannot be calculated. Nevertheless, the fixed-charge force fields are equally capable of describing the change in an average electric field associated with solvent polarity[149]. On account of this, M. Sega and C. Schröder successfully reproduced the main peak of the experimental dielectric loss spectra of water using three-charged site (SPC, SPC/E) and further complex six-charged site model (SWM6)[150], but failed to quantitatively obtain the peak around 5 THz which has been suggested to originate from the coupling of the electronic cloud and intermolecular motions.

Including the electronic polarization and charge transfer effects, the ab initio quantum mechanical calculation has become a progressively more important approach that allows deriving the molecular dipoles through decomposing the electronic structure in terms of maximally-localized Wannier orbitals[151]. In a state-of-art ab initio and force field MD simulation work, the dielectric loss spectra in the GHz range obtained by the classical forcefield still agreed well with that obtained by ab initio MD and dielectric spectroscopy experiments[152]. Thus, the classical forcefield is capable of analyzing the γ peak of the spectral response of water in interest here. For higher frequency domains (around or over 200 GHz), the ab initio MD is more suitable.

2.4 Summary

In this chapter, BDS experiments and MD simulations were carried out for lysozyme aqueous solutions with various concentrations, to reveal the molecular origin of the gigahertz domain of the dielectric spectra.

Dielectric loss spectra calculated from the probability density function of water relaxation time could be approximated by the Debye function. Upon increasing the protein solutions, the dielectric loss spectrum was gradually distorted from the Debye curve due to the broader non-Gaussian distribution of the probability density function of water relaxation time. The dielectric relaxation time calculated from this probability function is nearly identical to the γ relaxation time measured by broadband dielectric spectroscopy. The excellent agreement unambiguously proves that the dielectric relaxation in this gigahertz domain (γ relaxation) originates from the collective effect of the rotational self-correlation of water molecules only. This finding provides a good connection between the molecular water behavior with the macroscopic dielectric measurements.

Chapter 3 Water Rotational Dynamics around Protein

As summarized in Chapter 1, the current few MD simulations which reported on the water dynamics first assigned water molecules to multiple predefined HLs based on their initial positions relative to the protein surface. This kind of analysis may compromise the calculation accuracy of water relaxation time. To circumvent the trade-off between spatial solution and calculation accuracy, a stochastic analysis model is presented in this chapter and the spatial distribution of water relaxation time around the protein is obtained. Focused on the dynamics of individual water dynamics, two fundamental questions about the water dynamics introduced in **Sec. 1.2** in protein solutions have been addressed.

3.1 Theory of stochastic analysis

The definition of molecular rotational relaxation time requires a longer calculation period than the expected relaxation time, during which the molecule is in translational motion. Thus, to define the probability that water at a certain distance *x* (which is the distance from the nearest atom in lysozyme) has the relaxation time τ, $P(\tau|x)$, stochastic calculations regarding water position, *x*, and molecular rotational relaxation time, τ, were performed, assuming an ergodic system. By using Bayes' theorem, the conditional probability density function (PDF), $P(\tau|x)$, of the system is expressed as

$$P\left(\tau|x\right)=\frac{P\left(\tau\cap x\right)}{P(x)}, \tag{3-1}$$

where $P(x)$ is the PDF of the water molecule at a certain position *x*, which could be treated as the water distribution, and $P(\tau\cap x)$ is the joint PDF of relaxation time and distance. $P(\tau\cap x)$ was calculated by taking the statistics of the joint PDF, $P_j(\tau\cap x)$, of all the water molecules as:

$$P\left(\tau\cap x\right)=\frac{1}{N}\sum_{j=1}^{N}P_j\left(\tau\cap x\right). \tag{3-2}$$

Here, *N* is the number of water molecules in the simulation unit and the joint PDF, $P_j(\tau\cap x)$, of a single water molecule *j* is expressed as:

$$P_j\left(\tau\cap x\right)=A_{j,k}P_j\left(x\right),\quad \tau=\tau_{j,k}, \tag{3-3}$$

where A_{jk} corresponds to the probability that molecule j has relaxation time $\tau_{j,k}$. $P(x)$ was also calculated by taking the ensemble average of the PDFs of the surface distance of all single water molecules $P_j(x)$ as

$$P(x) = \frac{1}{N}\sum_{j=1}^{N} P_j(x). \tag{3-4}$$

$P_j(x)$ was derived by taking the statistics of x in the calculation period. That is, the probability of water j at x in the calculation period, $P_j(x)$, was calculated using the kernel density estimation method. Finally, by substituting **Eqs**. **(3-2)**, **(3-3)**, and **(3-4)** into **Eq**. **(3-1)**, the conditional PDF $P(\tau|x)$ was obtained as

$$P(\tau|x) = \frac{\sum_{j=1}^{N}\sum_{k} A_{j,k} \cdot P_j(x)}{\sum_{j=1}^{N} P_j(x)}, \quad \tau = \tau_{j,k}. \tag{3-5}$$

Finally, to denote the relaxation time at a specific position $\tau(x)$, the expected value was calculated by integrating the conditional PDF at x:

$$\tau(x) = \int P(\tau|x) \cdot \tau d\tau. \tag{3-6}$$

3.2 The rotational relaxation time of individual water molecules

A few reports have suggested that water rotation dynamics are highly correlated with its translation movement[83, 153]. In light of this, the water molecules are categorized into 3 types based on the time spent within the HL (the residence time), namely, highly mobile water (HMW), partially restricted water (PRW), and completely restricted water (CRW). It is first observed that the nearest distance between water and the protein surface is around 2 Å. Meanwhile, given that both SAS experiments (Svergun et. al)[38] and MD simulation (Merzel et. al)[39] have concluded that the thickness of lysozyme was 3 Å due to the found markedly high water density wherein, this study first adopted the definition of the HL as a 3 Å thick layer from the lysozyme surface. In this categorization, during the entire calculation period (5 ns), HMW always stays outside the HL, CRW is always in this HL, and PRW moves in and out of this HL from time to time (residence time is longer than 0 and shorter than 5 ns). It should be marked that the present classification is completely different from those defined in Refs[46-49]. Firstly,

the 3 classes of water are defined by the length of residence time, rather than based on the initial water position of a water molecule. Secondly, there are no bi or multi HLs defined throughout this dissertation. Thirdly, the 5-ns long trajectories are used to obtain the ACF and thus the relaxation time for every water molecule, whatever kind of water class a water molecule belongs to.

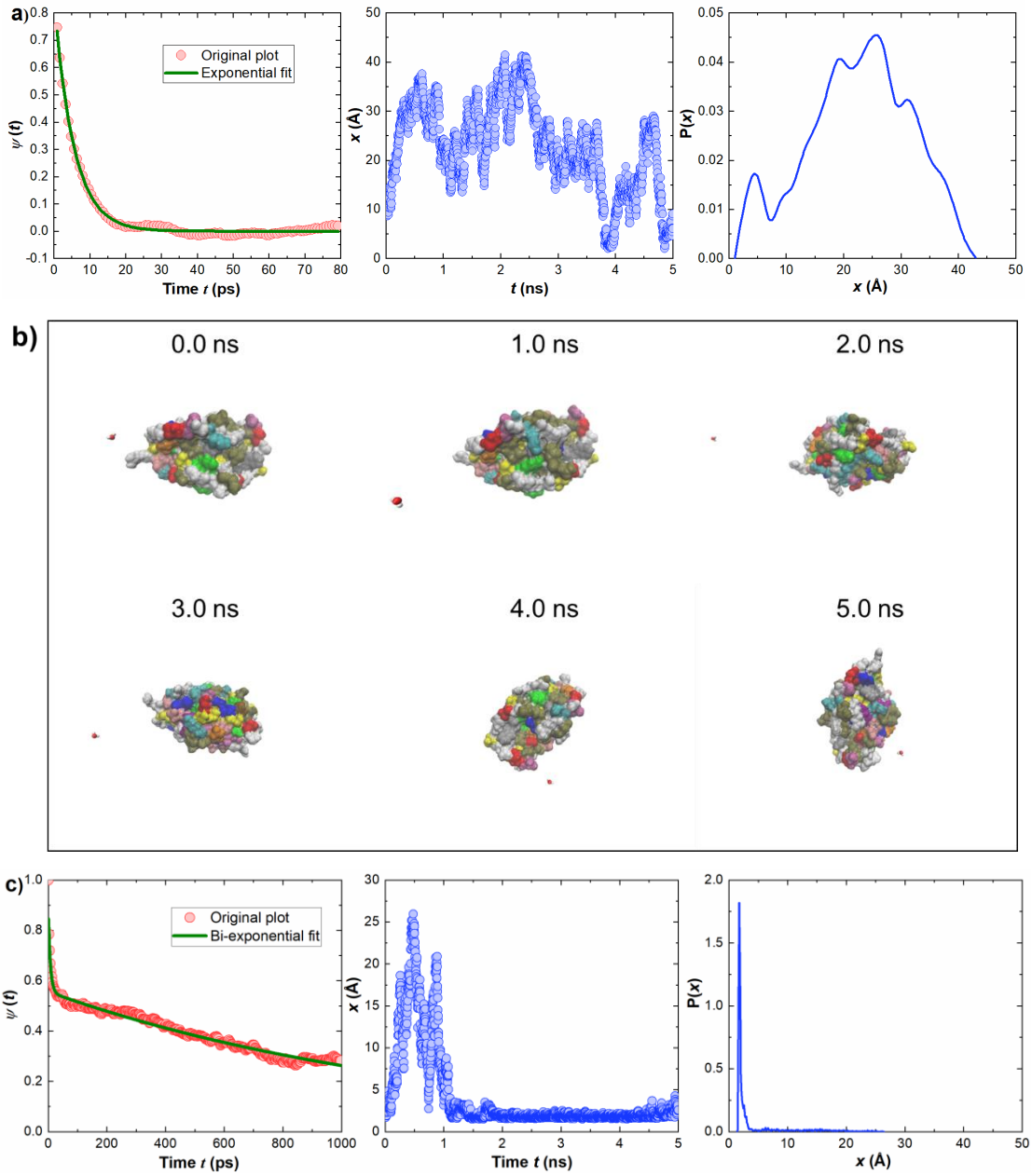

Figure 3-1. ACF of dipole moment (left), trajectory for 5 ns (middle), and probability distribution of distance to protein surface (right) for 3 typical types of water molecules in the solution of 33.2 mg/mL. **(a)** Highly mobile water (HMW) and **(c)** partially restricted water (PRW). The snapshots of the trajectories for HMW are shown in panel **(b)**. To be continued.

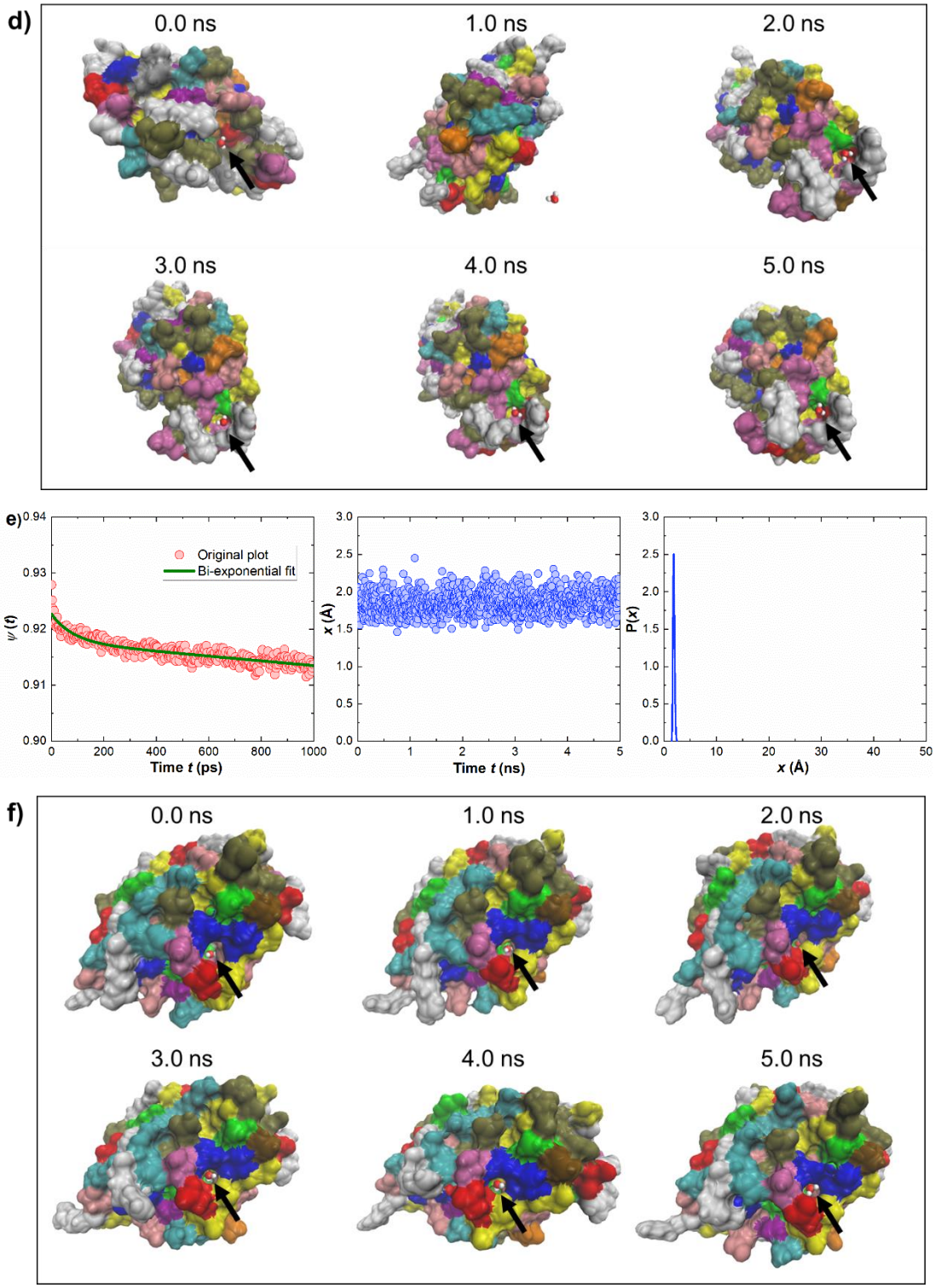

Figure 3-1. ACF of dipole moment (left), trajectory for 5 ns (middle), and probability distribution of distance to protein surface (right) for 3 typical types of water molecules in the solution of 33.2 mg/mL. **(a)** Highly mobile water (HMW), **(b)** partially restricted water (PRW), **(c)** completely restricted water (CRW). The snapshots of the trajectories for HMW, PRW, and CRW are shown in panels **(b)**, **(d)**, and **(f)**, respectively. The black arrows indicate the positions of corresponding water molecules.

Figure 3-1 shows the rotation ACF, the momentary distance between the lysozyme surface and water, and the probability density of this distance during the sampling period of 5 ns for 3 kinds of water. As shown in **Figure 3-1 (a)(center)** and **(b)**, a typical HMW molecule traveled all over the simulation box but never drifted into the protein HLs even though it may be close to them. HMW exhibits a single relaxation time of around 5.0 ps comparable to pure water (mean value: 4.96 ps) (**Figure 3-1 (a) (left)**). The active translation movement resulted in a very broad distribution of $P(x)$ (**Figure 3-1 (a) (right)**).

In contrast, a PRW molecule in **Figure 3-1** (**c**) and **(d)** moved actively as HMW when it was outside the HL for a long period, but then stopped its translational motion upon being captured by the cavities on the solute lysozyme (snapshots of 2.0 - 4.0 ns in **Figure 3-1 (d)**). Water in this category had two relaxation times of ACF, a short relaxation time (7.1 ps) slightly slower than pure water and an extremely long relaxation time of 3 orders of magnitude slower (1332.3 ps) than pure water. This water was moving around in the simulation box at the beginning (< 1 ns) and then trapped onto the protein surface (probably surface depressions, cavities, hydrophilic groups) without release. Thus, its probability density function of position had a peak close to the surface of the lysozyme with variance depending on the trapping time during the sampling period. Some water molecules in this category rotate like HMW when it stays on the protein surface for only tens of ps at most. It is also worth noting that at the initial moment (0.0 ns), the water molecule moves into the depressions on the lysozyme surface rather than into the cavities like in its later life. In consequence, this molecule is more viable in the translational motion and escapes from the protein surface tens of picoseconds later.

CRW, as exhibited in **Figure 3-1** (**e**) and **(f)**, stayed within the cavity on the lysozyme surface (distance from lysozyme surface ~ 2.0 Å) during the entire sampling time and exhibited a significantly slow rotational relaxation time of ACF (over 10 ns) with a sharp peak in its probability density function of the position. Although all the CRW maintains similar positions (2.0 Å from the lysozyme surface), the rotational relaxation time distributes in an extremely wide range. Besides, water with extremely slow relaxation time longer than 10 ns was usually attributed to protein tumbling in the studies of Dielectric Spectrum[57, 73]. MD simulation by Persson et. al also found such kind of long-lived buried water molecules, and they pointed out that protein tumbling

motion played the role in their rotation which therefore might reach the microsecond range[48]. As introduced in **Sec. 1.3.2**, the proposed EMOR mechanism suggested that it was the confined site declining the H-bonding opportunities that accounted for such a severe retardation[48].

A similar variety of water dynamics in the HL was reported by other groups. In the simulation work by Heyden and coworkers[100], three distinct hydration water species were defined in addition to "bulk water" based on their thermodynamic and vibrational signatures, which were named "bound", "weakly bound", and "unbound" water. Like the CRW classified, the "bound" water molecules were also reported to feature essentially zero diffusion and exhibit significant heterogeneity in rotational dynamics.

In the following analysis within this chapter, water with extremely slow relaxation time longer than 10 ns, which was usually attributed to protein tumbling in the studies of Dielectric Spectrum[57, 73], would be ignored, since the number ratio of such kind of slow water was less than 0.4% even for the most concentrated case.

3.3 Distribution of water position

The probability distribution of water at a certain distance *x* between water and the closest protein surface, i.e., the position distribution of water within the simulation box, is shown in **Figure 3-2**. An increasing number of water molecules were shown to approach the solute molecule of protein with increasing protein concentration. Two distinct peaks appear at 1.9 and 2.7 Å, which are in excellent agreement with the work of Ghattyvenkatakrishna *et al.*[78] and Gu *et al.*[154]. They stated that the first peak at 1.9 Å is the best hydrogen bonding distance formed between water oxygen and protein hydrogen, while the second at 2.7 Å is associated with a disparate configuration where a water molecule makes two HBs with the HB acceptor groups on the protein surface. In this context, water within 2.7 Å from the protein surface can form strong hydrogen-bond interactions with the protein surface and thus be categorized as hydration water, as defined by Laage *et al.*[88-89]. At a higher solute concentration, the distance between proteins is shorter. Interestingly, while the overall water distribution changes as the protein concentration increases, the first (1.9 Å) and second peaks (2.7 Å) of the water distribution remain at their positions.

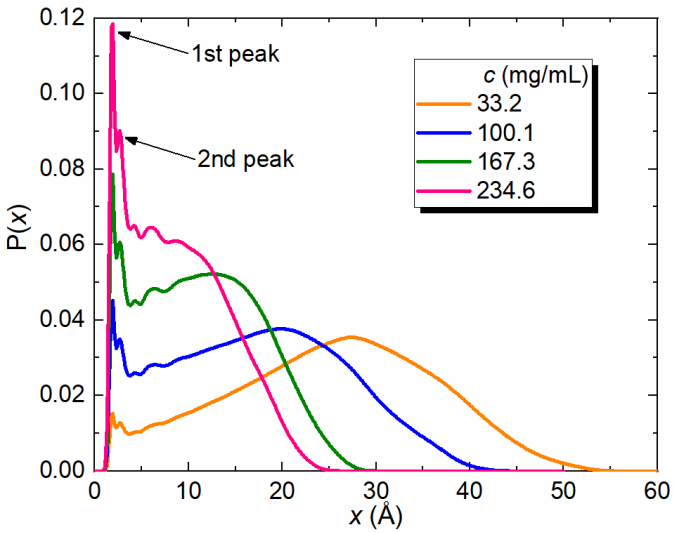

Figure 3-2. Distribution of water molecules regarding their distance to the nearest protein in various concentrations of lysozyme solutions.

Moreover, the result that the PDF height of the first and second peaks increase linearly with the protein concentration (see **Figure 3-3**) is due to the increase in the ratio of the volume close to the protein molecular surface in the whole system volume. Based on these results and the fact that augmentation of solute concentration reduces the water content relative to solute, it is straightforward to deduce that the amount of hydration water (<2.7 Å) per single protein declines with the protein concentration, as estimated by the dielectric measurement[127] where this phenomenon was interpreted by the aggregation process forming protein clusters.

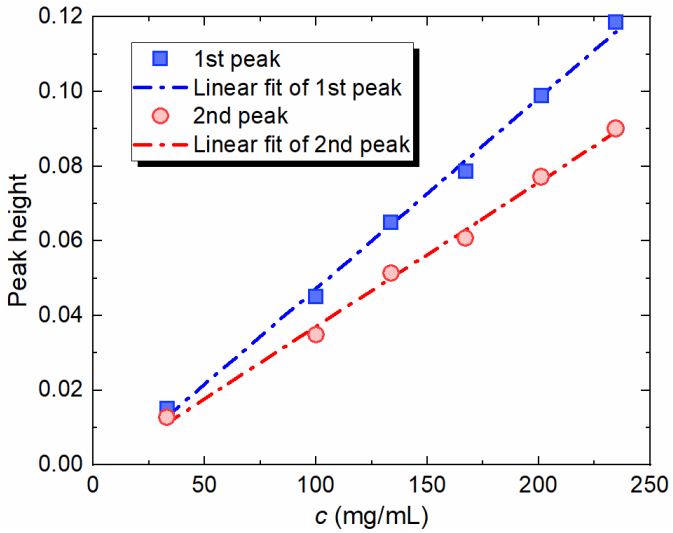

Figure 3-3. Height of the first and second peaks in the water distribution (**Figure 3-2**) for various lysozyme concentrations.

As shown in **Figure 2-1 (b)**, **Sec. 2.3.1**, water relaxation time distribution is distorted to the longer side from a symmetric Gaussian, principally because of the increasing

ratio of hydration water (x < 2.7 Å; see **Figure 3-2**), which tends to have longer rotational relaxation times (see **Figure 3-1** (**c**)). The relaxation time distribution exhibits a long tail extending from dozens of picoseconds up to the magnitude of nanoseconds owing to the restricted water. With the increment in lysozyme concentration, the proportion of this restricted water, corresponding to the progressively improving 1st and 2nd peak in **Figure 3-2**, accumulates and broadens the longer region of relaxation time distribution. It should be pointed out that this dissertation is focusing on every single water molecule within the solutions without defining any hydration shells based on the water's initial positions. Upon defining the HLs in advance and merely paying attention to the hydration water, the relaxation time probability function would be readily expected to exhibit more broadly since this customized (predefined) hydration water resembles somewhat the restrained water (PRW and CRW) categorized here. For instance, Bagchi's group obtained such distribution of water molecules that have residence times more than 100 ps inside of the protein HL defined within 10 Å from the protein atoms, which is found to be log-normal and comparatively broader[95].

3.4 Relaxation time distribution

3.4.1 Joint probability distribution

Compared with the global distributions of water distribution (P(x)) and water relaxation time (P(τ)), the following 2D stochastic analysis regarding water position and relaxation time is of more interest in the present work. **Figure 3-4** shows the joint PDFs of the rotational relaxation time τ of water at a distance x from the protein surface **(Eq. (3-5)** in **Sec. 3.1**), where the population of the statistic was taken over the calculation region and the calculation time. First, for all the simulated lysozyme solutions, the closer the distance from the protein surface, the broader the distribution of relaxation time, $P(\tau \cap x)$, although its peaks are nearly constant at 5 ps. For a solution concentration of 33.2 mg/mL, most water existed approximately 27.5 Å from the protein surface (peak in **Figure 3-4** (**a**)), behaving like pure water (5 ps), and a small portion of water molecules could stroll up to 55 Å away. With more lysozymes in the solution (100.1 mg/mL), almost all water molecules were within 40 Å from the protein surface and crowded at approximately 21 Å (peak in **Figure 3-4** (**b**)), although they still had relaxation time similar to pure water (~5 ps). However, at this concentration, a small but noticeable peak appeared in the region next to the protein (the peak at the

black arrow in **Figure 3-4** (**b**)). When the concentration of lysozyme was higher (**Figure 3-4** (**c**) and (**d**)), this small peak near the protein surface became the highest peak in $P(\tau \cap x)$, although the relaxation time of the peak (5.6 ps) at this position only changed slightly. However, as can be observed in the insets of **Figure 3-4** (**a-d**), the deviation of the relaxation time distribution at the position of this peak becomes broader with increasing protein concentration. The longest relaxation time in the relaxation time distribution thereof reaches the order of nanoseconds. In contrast, the farther the position, *x*, is, the smaller the deviation of the relaxation time becomes. In addition, the relaxation time at these main peak positions seems to be slightly longer than that of pure water in the most concentrated solution. A series of these results on $P(\tau \cap x)$ with various lysozyme concentrations suggests that the water distribution around the solute lysozyme depends on the concentration of lysozyme, which might be one of the reasons for the change in the probability density of water relaxation time (**Figure 2-1**).

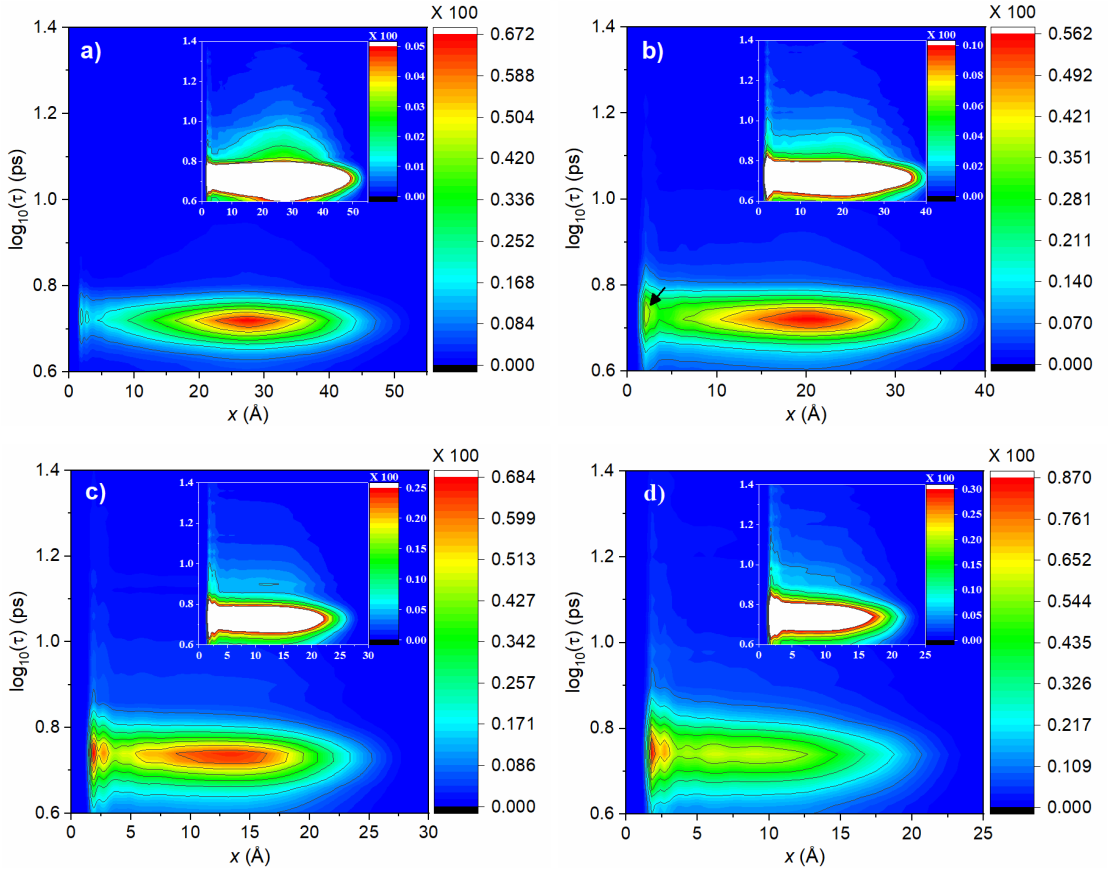

Figure 3-4. The joint probability density function of the distance to the protein surface and the relaxation time, $P(\tau \cap x)$, for water in various concentrations of lysozyme solutions. The inset highlights the percentage of longer relaxation time by showing the color bar on a limited scale. For visual convenience, only relaxation times shorter than 30 ps are shown here. **(a)** 33.2 mg/mL, **(b)** 100.1 mg/mL, **(c)** 167.3 mg/mL, and **(d)** 234.6 mg/mL.

3.4.2 Spatial distribution of the expected relaxation time

To elucidate the relationship between the expected number of water relaxation time and its position (**Eq.(3-6)**), the statistical distribution of the rotational relaxation time along the radial distance from the protein surface is obtained and illustrated in **Figure 3-5 (a**). In this figure, the rotational dynamics of water accelerate as water moves farther from the protein surface. And it is evident that the extent of protein perturbation is beyond 10 Å.

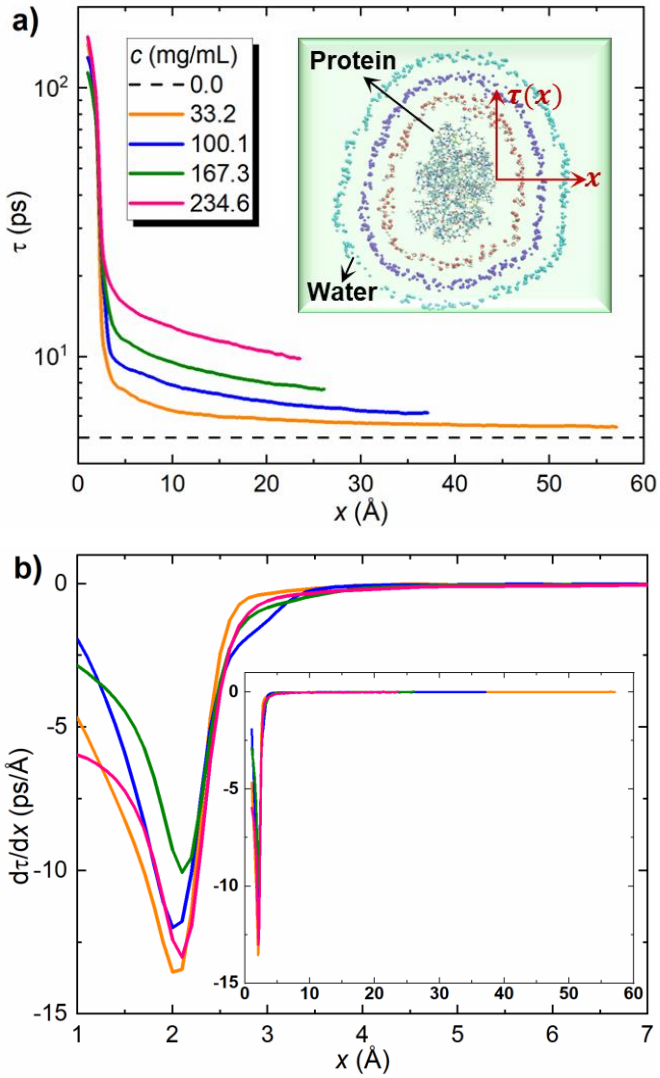

Figure 3-5. (a) Relaxation time of water rotation, τ, and **(b)** its first derivative versus distance, $d\tau/dx$, along the distance from lysozyme surface, x, in various concentrations of lysozyme solutions. Panel (a) and (b) share the same legend. The black dashed line in Panel (a) indicates the calculated mean value of pure water. The inset in Panel (b) shows the same plot with a wider x range.

Evidently, the rotational relaxation of water molecules prominently tends to be strongly retarded in the region adjacent to the protein surface for all the conditions

studied here. The relaxation time in **Figure 3-5** (**a**) declines drastically within 3-4 Å from the protein surface and gradually converges to the relaxation time of pure water farther than 3-4 Å. To express this trend more clearly, the first derivative of relaxation time versus the distance from the protein surface is plotted in **Figure 3-5** (**b**). Interestingly, independent of the solution concentration, the rotational relaxation time most drastically decreases at 2 Å from the protein surface and scarcely changes at a position beyond 3 Å.

This consistency suggests that the first HL, where the water rotational relaxation process is restricted by protein, lies within a thickness of 3 Å from the protein surface, which is identical to several reports. Svergun *et al.*[38] claimed that the average density of the first hydration shell (0–3 Å from the protein surface) is markedly higher than the average density of bulk water for lysozyme by combining small-angle scattering (SAS) of X-rays in H_2O with SAS of neutrons in H_2O and D_2O. Merzel *et al.*[39] consolidated this measurement by MD simulation, and they also found that the dipole moments within this 3-Å-thick first shell align more parallel to each other and more perpendicular to the surface normal than that outside this shell does. The significant retardation of the expected number of water rotational relaxation time within 3Å obtained here is in perfect agreement with these experimental and theoretical reports.

Besides, it is worth noting that the relaxation time of every single water molecule was obtained from the 5-ns long trajectories (see **Sec. 3.2**) and that all water molecules in the protein solutions were used in this stochastic analysis (see **Sec. 3.1**) not only those CRW molecules or those of which the initial positions were in the first HL. Thus, the excellent agreement has nothing to do with the water classification in **Sec.3.2** and is undoubtedly not a coincidence. It should also be noted that the trend of the relaxation time as a function of the distance from the protein surface is nearly the same regardless of lysozyme concentration, but shifts to a longer relaxation time as the concentration becomes higher. This shift happens because the probability density of relaxation time at a given position (see $P(\tau \cap x)$ at certain x in **Figure 3-4**) moved to a longer relaxation time as the lysozyme concentration increased.

Using the expected relaxation time as a function of distance from the protein surface (termed as the spatial distribution of water relaxation time, **Figure 3-5** (**a**)) together with the water distribution ($P(x)$; **Figure 3-2**), the PRF was derived from

$$\xi(x) = \frac{\int_0^x \tau(x')P(x')dx'}{\tau_{\mathrm{R,Bulk}}}. \qquad (3\text{-}7)$$

The PRF was compared with other studies that were also focused on lysozyme solutions. Mukherjee *et al*.[95] reported that the relative retardation factors within 10 Å from the protein surfaces for four different protein-water systems, including lysozyme, are 2.5-3.0, which is identical to our relative retardation factor of 2.72. Marchi *et al*.[83] calculated the retardation factor within 3.3 Å from the lysozyme surface to be 5.3, in good agreement with our result, 5.19. They also indicated that how much the rotational dynamics of water in the vicinity of lysozyme is slower than that in the bulk depends on how the hydration shell is defined, which is also consistent with our finding that water relaxation time decreases with the farther distance from the protein surface. Braun *et al*.[46] reported the retardation factor of the first HL for the globular protein ubiquitin as 5.7, while our result also shows an identical value of 5.68. All this consistency with various studies using hydration shell models with differing thicknesses proved the feasibility of the proposed stochastic method of evaluating the average retardation of water rotational dynamics within the HL.

Among the current studies endeavoring to establish the correlation between water relaxation time and the location relative to the solute molecule, two approaches have been reported available. Given that the dielectric spectroscopy experiments always detect several dielectric relaxation times for protein solutions and that these discrete time constants are assigned to some distinct water species (for instance, Charkhesht et.al[97]), one method divides the solvent into several HLs and optimizes the shell thicknesses to make the time constants derived from the collective ACF of the defined HLs match the experimentally measured dielectric relaxation times; then the protein HLs can be identified, as mentioned in the introduction part. The second method divides the HLs and assigns the thicknesses manually, without involving experiment techniques. Besides the studies where only one or two layer[95] were defined, some research groups sliced the solvent into 9 consecutive layers with the help of Voronoi tessellation analysis[46, 48-49] as reviewed in **Sec.1.2.2**. Afterwards they calculated the relaxation time of individual water molecules residing in each hydration shell, and finally obtained the average retardation relaxation time which can be associated with the position of the respective defined HLs.

In short, defining the HLs seems to be inevitable in the two presently available methods. However, predefining the HLs has the disadvantage of balancing spatial resolution and calculation accuracy of water relaxation time. Nevertheless, as the trajectories shown in **Figure 3-1**, it seems to be difficult for a water molecule to stand still or diffuse slowly at some place beyond 3 Å from the protein surface. For instance, Laage and his coworkers obtained the reorientation time from an exponential fit only on the 2–10 ps interval in order to avoid the contributions from waters which have moved to a different environment[88-89]. Hence, the active translation movement of water molecules will somewhat limit the sampling time within the HLs, thus reducing the time range to compute the relaxation time of the target water molecule, especially for those within the shells distant from the protein. Fewer HLs may improve the accuracy; however, it will sacrifice the spatial resolution instead. The stochastic approach proposed here circumvents the process to define any non-a priori HL and thus provides a continuous relaxation time distribution regarding the distance from the protein surface with a higher spatial resolution, although quite a computation effort is demanded.

3.5 Summary

In this chapter, MD simulations were performed for pure water and lysozyme aqueous solutions with various concentrations, to have a better understating of the fundamental water rotational dynamics around protein.

Water rotational relaxation time is found to range from several picoseconds to tens of nanoseconds in protein solutions. Depending on the time it stays within the hydration layer (HL) (the residence time), water can be classified into 3 kinds: highly mobile water (HMW), partially restricted water (PRW), and completely restricted water (CRW). In dilute solutions, water relaxation time exhibits a Gaussian probability distribution. In concentrated solutions, the distribution is distorted to a longer time constant due to the increasing proportion of restricted water with longer relaxation time (PRW and CRW).

A stochastic analysis combining the relaxation times and trajectories of every single water molecule is proposed and the two-dimensional probability distribution of water relaxation time and its distance from the protein surface is obtained. This analysis suggests that water rotational dynamics accelerate with the farther distance from the protein surface. This spatial distribution surprisingly reproduces the average slowdown

factors of hydration water reported by other studies[83,95,46]. Moreover, regardless of protein concentration, this spatial distribution of relaxation time reveals that the water rotation is severely retarded within 3 Å from the lysozyme surface and is nearly comparable to pure water when farther than 10 Å. In view of the remarkable slowdown, the HL thickness is subsequently identified as this 3 Å thick water layer from the protein surface, in excellent agreement with the determination by X-ray and neutron SAS of solution[38-39]. The presented stochastic analysis is an original method that circumvents the balance of spatial resolution and calculation accuracy of relaxation time, compared with widely accepted multiple HLs. This work may contribute to a better understanding of the picture of hydration water and its coupling with protein.

Chapter 4 Molecular Mechanism of Water Rotational Dynamics

To control and slow down the water rotational dynamics around proteins, a better understanding of water rotational mechanism is necessary. This chapter explores the behind mechanism from the perspective of electrostatic interactions. For the rotational relaxation of a single water molecule, the reorientation of the exerted electric field is studied; for local hydration dynamics within the HL, the reorientation of the electric field at fixed positions (local electric field) is illustrated.

4.1 Methodology

4.1.1 Calculation of the electric field by Coulomb's law

The electric field, $\boldsymbol{E}$, at the target water's position is defined at the midpoint of its 2 hydrogen atoms which corresponds to the center point of the water dipole, at its mass center, or at the position of its oxygen atom. To avoid any confusion over the terminology in this paper, it should be pointed out that this electric field exerted on the target water molecule is termed the "*exerted electric field* ". Therefore, it signifies the electric field that is always targeted at the diffusing water, while the "*local electric field* " discussed in the subsequent **Sec. 4.5** indicates the electric field at the fixed positions within the simulation boxes. Since the direct output from MD simulation provides only the force applied onto each atom, $\boldsymbol{E}$ is calculated with the Coulomb's law using the charge amount of each atom and its coordinates. Therefore, the electric field at a given position $\boldsymbol{r}$, $\boldsymbol{E}(\boldsymbol{r})$, is derived by **Eq. (4-1)**, where the Columbic interactions generated by all the charged atoms (atoms of protein and water molecules, ions) in the simulation unit and its neighboring periodic images are summarized. L indicates the side length of the simulation box. Here, the first summation is for all periodic images *n*, the second is for all charged atoms *j* in the image *n*, and the nearest 26 images ((±L, 0, 0), (0, ±L, 0), (0, 0, ±L), (±L, ±L, 0), (0, ±L, ±L), (±L, 0, ±L), and (±L, ±L, ±L)) are considered in the calculation. Expanding the area for the summation to 125 images (-2L ~ 2L, -2L ~ 2L, -2L ~ 2L) gives an identical $\boldsymbol{E}(\boldsymbol{r})$ to that using the nearest 26 images and the unit box.

$$\boldsymbol{E}(\boldsymbol{r}) = \sum_{\boldsymbol{n}} \sum_{j} \frac{1}{4\pi\varepsilon_0} \frac{q_j \cdot (\boldsymbol{r}_j + \boldsymbol{n}L - \boldsymbol{r})}{\left\| \boldsymbol{r}_j + \boldsymbol{n}L - \boldsymbol{r} \right\|^3} \tag{4-1}$$

4.1.2 Extraction of the electric field from GROMACS package

In MD simulation, the electrostatic interactions were dealt with the Particle Mesh Ewald (PME) method which is slightly different from the direct summation of the Columbic interaction in the above section. To guarantee the validity of the present method, the electrostatic field applied onto the atoms during the sampling time was extracted from the docking package according to the approach proposed by Fried et. al,[155] and compared with the direct summation method.

In Fried's method, first, 1) the total force on a target atom (hydrogen atom was chosen here), $\boldsymbol{f}^i_{tot}$, at each time step was extracted and saved. Then, 2) to remove the long-range electrostatic interactions, a charge-free topology file was prepared for the simulation system in which the partial charges for all the atoms were set to zero (charges on all the water molecules, all the protein molecules, and all the ions). Meanwhile, all other charge-free parameters remained unchanged. Next, 3) the trajectory from the production run was post-processed with the generated charge-neutralized topology using GROMACS' *rerun* utility. In the resulting trajectory, only the electrostatic interactions were absent. Likewise, the modified non-electrostatic total force on the target atom, $\boldsymbol{f}^i_{nonelectro}$, was extracted for each snapshot. Consequently, 4) the total electric force, $\boldsymbol{f}^i_{electro}$, was obtained by subtracting all the non-electrostatic forces extracted by rerunning the trajectory with the charge-free topology, $\boldsymbol{f}^i_{nonelectro}$, from the total force, $\boldsymbol{f}^i_{tot}$, (**Eq. (4-2)**). Finally, 5) the electric field on the target atom exerted by all the charged particles, $\boldsymbol{E}^i_{electro}$, can be converted by simply dividing by the charge of the atom with **Eq. (4-3)**.

$$\boldsymbol{f}^{\,i}_{electro} = \boldsymbol{f}^{\,i}_{tot} - \boldsymbol{f}^{\,i}_{nonelectro} \tag{4-2}$$

$$\boldsymbol{E}^{\,i} = \boldsymbol{f}^{\,i}_{electro} / q_i \tag{4-3}$$

Here, *i* is the index of time, $\boldsymbol{f}$ indicates the force, and $\boldsymbol{E}$ denotes the electric field.

4.1.3 Relaxation times of rotational water dipole and exerted electric field

To find out the effect of the reorientation of the exerted electric field, $\boldsymbol{E}$, on the water molecular rotational movement, the momentary experienced electric field for each target water molecule was monitored. Similarly, the time auto-correlation function (ACF) of the experienced electric field $\boldsymbol{E}$, EACF, is defined as

$$\psi_E(t) = \frac{\langle \boldsymbol{E}(0) \cdot \boldsymbol{E}(t) \rangle}{\langle \boldsymbol{E}(0) \cdot \boldsymbol{E}(0) \rangle}, \tag{4-4}$$

where $\boldsymbol{E}(t)$ is the orientation of the electric field exerted on the target water molecule at time t and subscript '$_E$' depicts the electric field. The relaxation time of electric field reorientation, $\tau_{E,k}$, was calculated by the same processes as that for calculating the relaxation times of dipole rotation (**Eqs. (2-2)** and **(2-3)**),

$$\psi_E(t) = \sum_{k=1} A_{E,k} \exp(-t/\tau_{E,k}) \tag{4-5}$$

and

$$\tau_{E,\text{int}} = \sum_k A_{E,k} \cdot \tau_{E,k}, \tag{4-6}$$

where $\tau_{E,int}$ is the averaged relaxation time of the electric field reorientation for each water molecule.

4.2 Validity of the direct summation method

The angle and strength of the electric field, $\boldsymbol{E}$, derived from these two approaches for a target H atom during the sampling time (6251 data points corresponding to 5 ns) are shown in **Figure 4-1**, for a water molecule in 33.2 mg/mL lysozyme solution. It clearly shows that the direct summation of Coulomb force (**Sec. 4.1.1**) gives the identical value of electric field strength with that extracted from the GROMACS package (**Sec. 4.1.2**) which was experienced in practice through the simulation run. Since the direction of $\boldsymbol{E}$ is the very quantity of interest in the present work, the angular difference between these $\boldsymbol{E}$ vectors obtained from these 2 different methods was also compared and found to be less than 4 degrees (see **Figure 4-1** (**Right**)). To clarify the difference between the two methods, the two-dimensional probability density distribution of the angular difference and the relative difference between the magnitude, ΔE, is plotted in **Figure 4-2**. The highest peak in the contour is at around ΔE = 0.5%, angle = 0.46°, and over 99% of the data set has a strength difference within 2.5% and an orientation difference less than 1.5°. The excellent agreement validates that the method of the direct summation of the Coulomb interaction is consistent with that used in the production MD simulation. The direct summation method was used to compute the electric field at the midpoint of the 2 hydrogen atoms and at the water mass center,

while that at the oxygen atom's position was extracted from GROMACS package due to the much less computation load.

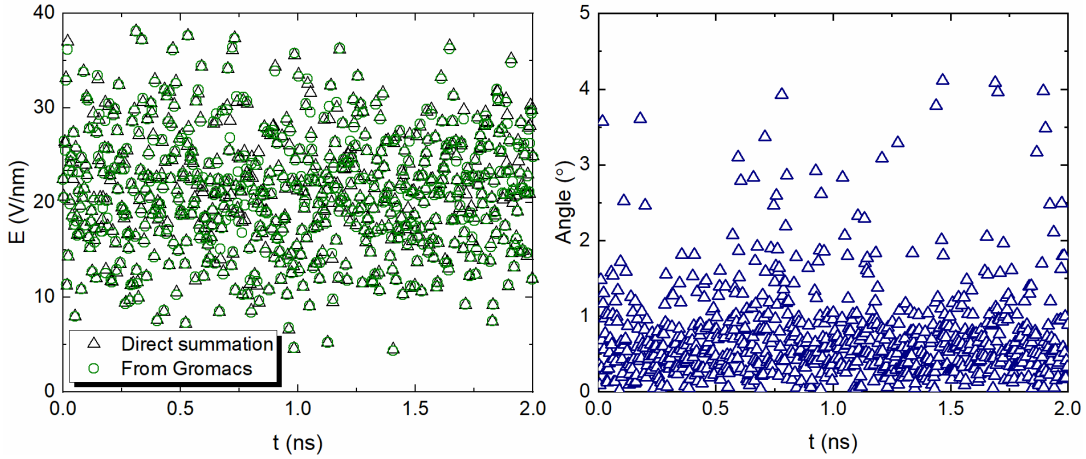

Figure 4-1. (**Left**) Strength of the electric field, *E*, experienced by one hydrogen atom of a water molecule during the sampling time. The black hollow triangle indicates the result obtained by directly summing up the Columbic terms, and the green hollow circle indicated that extracted from GROMACS package. (**Right**) The angle between the two electric field vectors was calculated from these two methods during the sampling time.

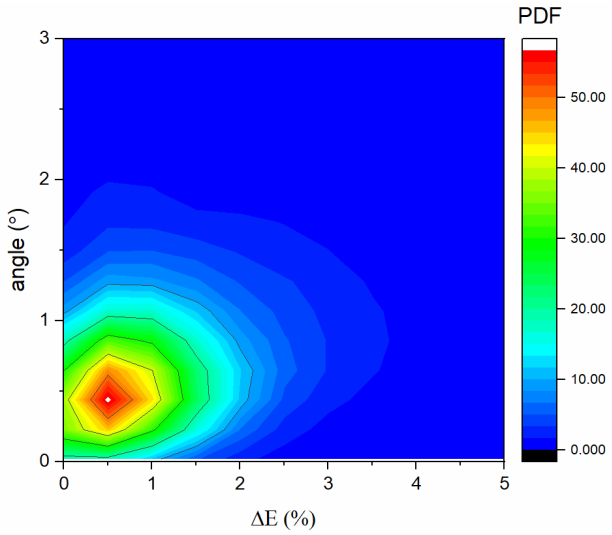

Figure 4-2. Two-dimensional probability density distribution of the angular difference and the relative difference between magnitude, for the two electric field vectors calculated from two different methods.

4.3 Electric field-water dipole alignment effect

To evaluate the tendency of the water dipole alignment to the exerted electric field, the angle between water dipole (***D***) and the exerted electric field (***E***) for every single water, θ, was investigated at first. For every moment of a single water molecule, a data set (x, θ) could be obtained. By setting a series of continuous small intervals for x ($x \pm$

dx), the whole data sets among all water molecules at all snapshots were categorized into an interval based on the first parameter (x). Then, the distribution of the angle (θ) was obtained within each interval. The result on the time and ensemble average with standard deviation is shown in **Figure 4-3 (a)** with a solid line as a function of the distance from the lysozyme surface, x. Wherever a water molecule is (x > 2.2Å), the averaged angle between $\boldsymbol{E}$ and $\boldsymbol{D}$ keeps a constant value of around 16° with a deviation of ~8.9°. Meanwhile, the angle between the water dipole and the overall field in the state of pure water, θ, is also obtained and θ on average is 16° ± 8.7°. Due to this global strong alignment of water dipole along the exerted electric field in the lysozyme solution, a strong correlation between water rotation and electric field reorientation can be expected.

In addition, to dissect the respective contribution from solute and solvent to this tendency of water dipole ($\boldsymbol{D}$) – electric field ($\boldsymbol{E}$) alignment, the overall exerted electric field, $\boldsymbol{E}$, was decomposed into the electric field originating in water solvent and that originating in protein molecules and ions, $\boldsymbol{E}_{\mathrm{PI}}$. It is worth noting that, for the ion concentrations investigated here, the effect of chloride ions on water rotational relaxation time is negligible compared with that of protein molecules[156]. Clearly, θ_{PI} (**Figure 4-3 (a)**), the angle between $\boldsymbol{D}$ and $\boldsymbol{E}_{\mathrm{PI}}$, is nearly perpendicular (~86°) especially when water is far from the protein. This observation indicates that the exerted electric field generated from protein ($\boldsymbol{E}_{\mathrm{PI}}$) has a negligible effect on the alignment of the water dipole especially when water is distant from the protein surface. However, this angle reduces drastically as water approaches within circa 5 Å from the protein surface, which suggests that the alignment of the water dipole is being gradually affected by the local electric field from the protein ($\boldsymbol{E}_{\mathrm{PI}}$) when close to the protein. As claimed by Merzel et. al[39], water dipole is being electrostatically driven to orient towards the progressively stronger electrostatic field lines originating from the protein atoms when it approaches the protein surface. Soper and Weckstrom[157] also observed a similar trend for water in the HLs of ions. On the other hand, **Figure 4-3 (a)** also illustrates the angle between the water dipole and the field from water solvent, θ_{W} (the olive long-dashed line), which exhibits the opposite trend from that of θ_{PI}. Away from the protein surface (x > 5 Å), θ_{W} shows excellent agreement with θ (red solid line), which suggests that the alignment of water dipole orientation with the electric field therein is due to the water-water interactions like pure water. In the vicinity of protein, θ_{W} begins to increase markedly,

which together with the progressively improving alignment with the protein's field (the decreasing θ_{PI}), contributes to the global strong alignment of water dipole along the exerted electric field (red line). It is also intriguing that θ_W and θ_{PI} are identical at the nearest position from the protein surface.

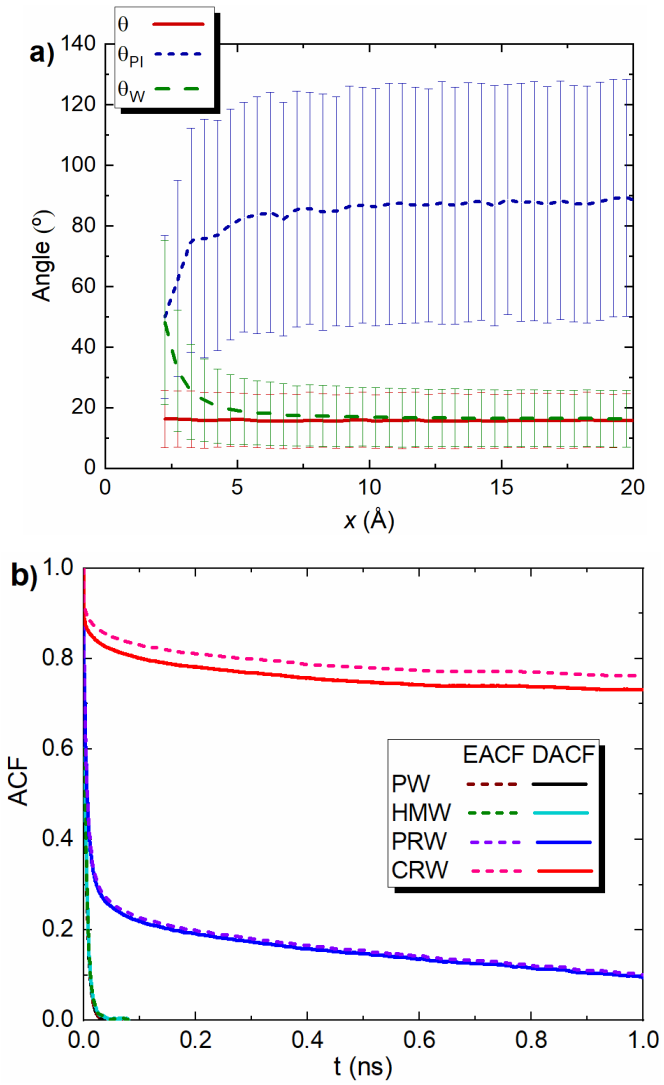

Figure 4-3. Illustration of electric field-water dipole alignment effect. **(a)** Distribution of the angle between water dipole and the exerted electric field in lysozyme solutions of 33.2 mg/mL. The red solid line represents the angle between the water dipole (***D***) and the exerted overall electric field (***E***), θ; the navy short-dashed line represents the angle between the water dipole (***D***) and the exerted electric field from protein ($\boldsymbol{E}_{PI}$), θ_{PI}; and the olive long-dashed line represents the angle between water dipole (***D***) and the exerted electric field from water solvent ($\boldsymbol{E}_W$), θ_W. The same results hold for concentrations up to 234 mg/mL. **(b)** Average (time) autocorrelation function (ACF) of electric field reorientation (EACF, dashed line) and dipole moment rotation (DACF, solid line) for highly mobile water (HMW), partially restricted water (PRW), and tightly bound water (CRW) in lysozyme solution of 234 mg/mL. To be continued.

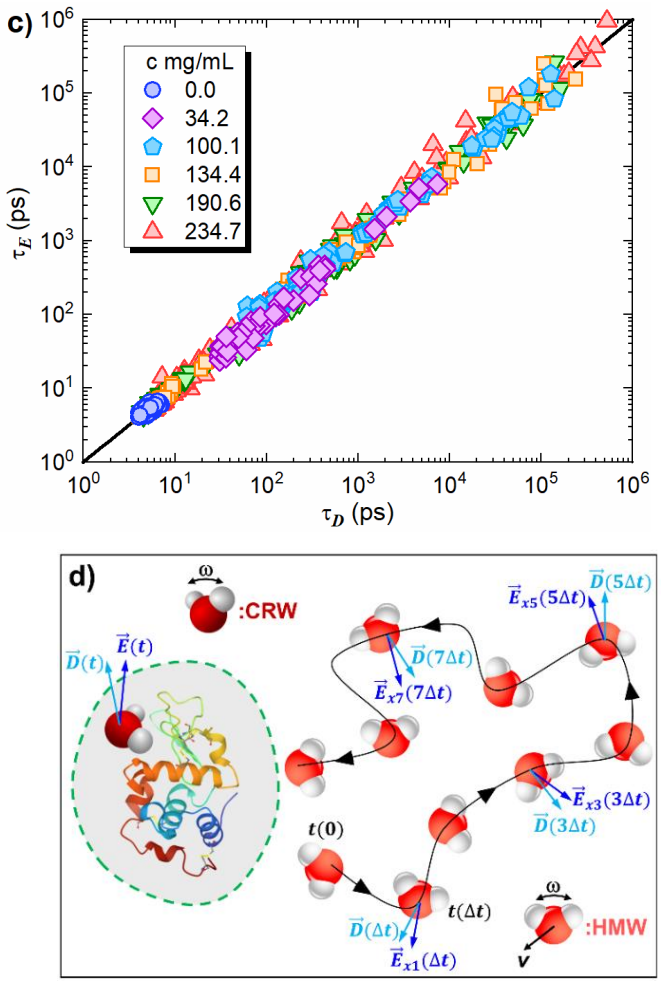

Figure 4-3. Illustration of electric field-water dipole alignment effect. **(c)** Comparison of relaxation time of dipole rotation ($\tau_{D,int}$) and that of the exerted electric field reorientation ($\tau_{E,int}$) for every single water molecule in various concentrations of lysozyme solutions and in pure water. The black solid line represents the function $\tau_{D,int} = \tau_{E,int}$. **(d)** Illustration of synchronization between water dipole (***D***) rotation and the exerted electric field (***E***) reorientation for individual CRW and HMW in lysozyme solutions. The results plotted here are based on the electric field at the midpoint between the two hydrogen atoms and hold for the field at the water mass center and the oxygen atom, even though the local electric field is non-uniform at the molecular scale.

As elaborated in the above chapter (**Sec. 3.2**), the HMW is the most viable water, showing a fast rotational relaxation time similar to pure water. In contrast, CRW is severally suppressed in diffusion, possessing an extremely long relaxation time. Behaving like the combination of HMW and CRW, PRW exhibits a relaxation time ranging from a few picoseconds to tens of nanoseconds depending on its residence time. To clearly discriminate PRW from HMW, this section only focuses on the PRW with a rotational relaxation time longer than 60 ps here (**Figure 4-3 (b)** and **Figure 4-4**). Again,

multiple "HLs" are not divided here, and the focus is on the dynamics of each individual water molecule wherever it translates.

The aforementioned results from **Figure 4-3 (a)** clearly show that a water dipole is aligned with the exerted electric field on average. With this trend of alignment, it is expected to deduce the mechanism of water rotational dynamics related to the electrostatic force applied to a water dipole. In **Figure 4-3 (b)**, the averaged autocorrelation functions of electric field reorientation (EACF) and dipole moment rotation (DACF) of HMW (500 molecules), PRW (374 molecules), and CRW (47 molecules) are plotted for lysozyme solution of 234 mg/mL, respectively. The standard deviation of each ACF is shown in **Figure A1 (a)** (see the **Appendix**). The DACF and EACF of pure water (PW, 3000 molecules) are also plotted here, showing an excellent agreement with that of the highly mobile water (HMW). More interestingly, the DACF of the relatively more mobile waters (HMW and PRW), which occupy over 99.7% of the water population in total, is overlapped with their respective EACF. For CRW, DACF and EACF still exhibit similar relaxation except for the initial drastic drop which results from the strength difference of a sub-picosecond relaxation time. The more precise similarity of relaxation times between DACF and EACF of CRW, PRW, and HMW can be seen in **Table S1**, which are derived by fitting each ACF to the multiple decaying exponential functions. Thus, in general, water dipole and the exerted electric field lose the memory of orientation at the same speed.

The identical decaying curves of DACF and EACF are found not only in the lysozyme solution of high concentration shown above, but in the wide range of concentrations of lysozyme solutions. In **Figure 4-3 (c)**, the averaged relaxation times (**Eq. (4-6)**) of individual water dipole rotation, $\tau_{D,int}$, vis-a-vis that of the exerted electric field reorientation, $\tau_{E,int}$, are plotted for pure water and lysozyme solutions of various concentrations. Although the averaged value may have confounded the corresponding motions of a single water molecule, the identical value of $\tau_{D,int}$ and $\tau_{E,int}$ with nearly the same relative deviation shows a near-perfect correlation between the electric field reorientational ACF and the rotational ACF time-decay for every single water molecule, which clearly illustrates that the water molecular rotation movement synchronously follows the reorientation of the exerted electric field (**Figure 4-3 (d)**), as implied in **Figure 4-3 (a)**. Furthermore, the data points in **Figure 4-3 (c)** cover all the categories of water molecules with rotational relaxation times ranging from

4 picoseconds to tens of nanoseconds, including the CRW which stays within the HL during the entire sampling period. Thus, this result unambiguously demonstrates that the electrostatic interactions play a dominating role in the water rotational relaxation wherever water is, even in the first HL.

This agreement of $\tau_{D,int}$ and $\tau_{E,int}$ here shows clear contrasts to the reported illustrative short-range EMOR mechanism which underestimates the alignment effect of electric field on the rank-1 time correlation functions (TCF) in the first hydration shell (2.8 Å thick in their case).[48] This discrepancy might be attributed to two factors: 1) in the present study, the exerted electric field from all charged particles (viz., protein, water, and ions) is taken into account, not merely that from protein; 2) both the DACF and EACF are built on individual water molecule here, not on the "*hydration shells or the subsets of the first shell based on the initial location of water oxygen*".[48]

Much as I agree that water rotational perturbation is proportionally related to the confinement degree of the hydration sites[48, 158], it is believed that this confinement effect is probably achieved by modifying the local electric field rather than by the short-range perturbation, as manifested by the quantitative correlation established in **Figure 4-3**. The femtosecond-resolved measurements also reported that the electrostatic interactions with the protein's residues affected the water-network rearrangements[44, 50]. Link and Heng even inferred that the stronger interplay between water and protein surfaces in combination with ions in solution led to an increase in the hydrogen bond strength between hydration water and the insulin residues.[159] Among the studies associating the water reorientation process with HBs, the EJM introduced by Laage and Hynes[99] provided a quantitative description of the H-bond reorientation in liquid water. However, combing with the recently developed fluctuation theory, they also pointed out that the electrostatic interactions dominated the activation energies for HBs jump and water reorientation[105-106]. From this viewpoint, the electrostatic interactions are the driving force for water rotational dynamics and the dynamic dipole-filed alignment perspective presented here has complemented this understanding.

Moreover, in the very latest MD simulation study on a dilute artificial dumbbell water system, Matyushov et.al also reported a good agreement between the relaxation time of solute rotation and the relaxation time of the electric field from water, in case that the partial charge of the investigated dumbbell is larger than 0.6 e.[160] Despite that the target particle of our research is different or even opposite (solute or solvent), the

intriguing agreement of both relaxation times in this report demonstrates how the molecular rotation movement synchronously follows the exerted electric field reorientation process. It should also be noted that the intrinsic field investigated in the present dissertation is around 2 V Å^{-1} in magnitude (see **Figure A1** in the **Appendix**), while several studies reported the asymmetric response of interfacial water near a monolayer graphene electrode[161] or a Janus interface[162] to the applied external electric fields of 0.03 V Å^{-1} or weaker.

4.4 Contribution of the electric field from protein/solvent to water rotational dynamics

With the dynamic field alignment picture proposed in the former section, the respective contribution of the electric field from protein and solvent to water rotational dynamics (relaxation time) is unpicked. The typical EACFs using the overall exerted electric field, $\boldsymbol{E}$, and using that originating solely from protein and ions, $\boldsymbol{E}_{\mathrm{PI}}$, of water in the three categories are shown in **Figure 4-4**, respectively. The dashed-dotted lines indicate the results using $\boldsymbol{E}_{\mathrm{PI}}$, while the other dashed lines indicate those using $\boldsymbol{E}$ as shown in **Figure 4-3 (b)**. The standard deviation is plotted in **Figure A1 (b)**. Only the EACFs of CRW using $\boldsymbol{E}$ and $\boldsymbol{E}_{\mathrm{PI}}$ are nearly overlapped, leading to the corresponding relaxation time ratio between both EACFs being 1.008 (**Table S1)**.

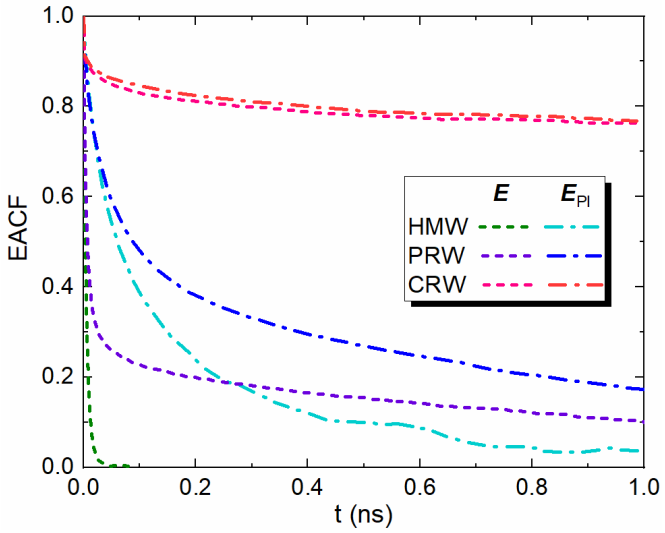

Figure 4-4. Average EACF obtained by using the overall electric field ($\boldsymbol{E}$) and the electric field originating solely from proteins and ions ($\boldsymbol{E}_{\mathrm{PI}}$), for the above 3 classes of water in a lysozyme solution of 234 mg/mL. The dashed lines indicate the EACF using the overall exerted electric field while the dashed-dotted ones indicate that using the electric field only from proteins and ions (viz., without the contribution from water).

The agreement of both EACFs of CRW elucidates that it is the electric field from the protein, $\boldsymbol{E}_{\mathrm{PI}}$, that dominates the rotation of CRW which stays inside the HL. Since both the protein and the CRW hardly ever diffuse, it is further concluded that the extremely slow protein tumbling gives rise to the slow orientating electric field, thus resulting in the severely retarded water rotation dynamics. In consequence, the reorientation of $\boldsymbol{E}_{\mathrm{PI}}$ might reflect the protein tumbling which has been deemed to be responsible for the γ relaxation time in the dielectric spectroscopy studies. Interestingly, the reorientation time of protein tumbling speculated to be around 22 ns (= the relaxation time of EACFs of CRW using $\boldsymbol{E}_{\mathrm{PI}}$, see **Table S1**) is commensurate with the detected γ relaxation time which was reported to be in the order of 10-30 ns in lysozyme solutions[57, 73, 75, 147]. The time-dependent fluorescence Stokes shift (TDFSS) experiments by Zhong and his coworkers also suggested a protein tumbling motion in ~20 ns[104].

On the other hand, for PRW, the difference between the initial drop of EACF using $\boldsymbol{E}$ and that using $\boldsymbol{E}_{\mathrm{PI}}$ is observed. EACF using $\boldsymbol{E}_{\mathrm{PI}}$ shows a relaxation time of 1.5 times as long as that using $\boldsymbol{E}$ (**Table S1**). Therefore, the rotation of these comparatively mobile waters that occasionally move outside the HL is accelerated by the electric field from the surrounding water. In this sense, the intrinsic solvent-solvent electrostatic interactions also play a significant role in addition to the effect of the protein's electric field.

For HMW, the most mobile water which is always outside of the HL, the difference in the initial drop of EACFs is more remarkable. Consequently, $\tau_{E,int}$ of EACF using $\boldsymbol{E}_{\mathrm{PI}}$ is even 2 orders of magnitude slower than that using $\boldsymbol{E}$ (**Table S1**). This wide discrepancy indicates that the influence of protein is already inappreciable for water outside the HL. Thus, the fact that EACF (and DACF) of HMW and EACF (and DACF) of pure water are identical leads to the conclusion that the dynamic alignment of water dipole and electric field outside the HL is just an intrinsic property like pure water, as suggested in Refs[47, 49, 111]. However, the objective here is the effect of protein's electric field on water rotational dynamics, rather than the shell-averaged dipolar orientational order parameters[47, 49, 111].

4.5 Reorientation of the local electric field

The electric field reorientation analysis of CRW, PRW, and HMW shown in **Sec. 4.4** concludes that the rotation dynamics of water near the protein surface are subject to the interaction with protein. It is also illustrated in **Figure 4-4** that the rotational dynamics of the relatively more mobile waters (PRW moving inside and outside the HL and HMW moving only outside) are strongly affected by the electric field from the surrounding water molecules. These observations imply that the active translation movement of water might facilitate the electric field of the surrounding solvent overwhelming the effect of a slow protein electric field.

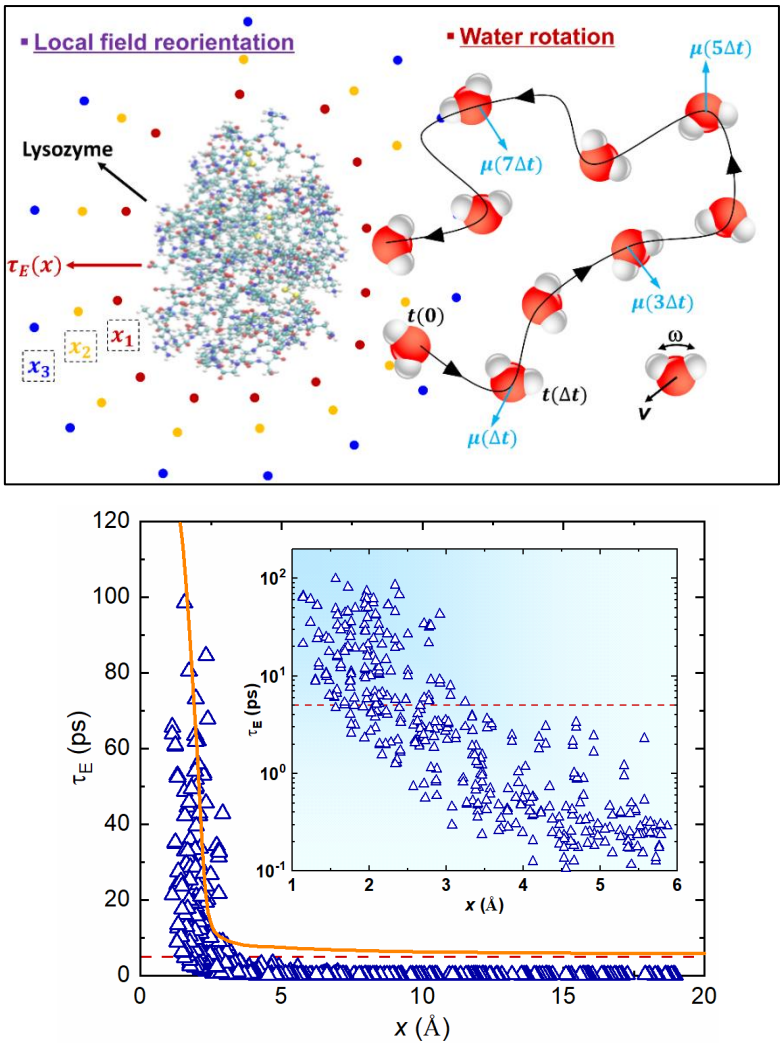

Figure 4-5. (**Top**) Schematic diagram of local field reorientation and water rotational relaxation. (**Bottom**) Relaxation time of electric field reorientation at fixed positions, $\tau_E(x)$, in lysozyme solution of 33.2 mg/mL. The red dotted line represents the average value of water rotational relaxation time, $\langle\tau_{D,int}\rangle$(4.96 ps), of pure water or HMW in lysozyme solutions. The solid orange line is the spatial distribution of water rotational relaxation time obtained from the stochastic analysis.[158]

To address the obvious slowdown of water near proteins, the local overall electric field dynamics at multitudes of "*fixed* " positions within the simulation boxes are first monitored. Then the local electric field reorientation relaxation time, $\tau_E(x)$, was obtained. Its distribution versus the distance from the protein surface is depicted in **Figure 4-5** where the average rotational relaxation time of a single pure water molecule (4.96 ps) is also plotted by a dotted line for comparison. As a whole, the reorientation of the local electric field drastically accelerates with the farther distance from the protein surface. In addition, it is evident that the local electric field reorients much slower within a layer 3 Å from the protein surface that corresponds to the HL.

Inside the HL, the spatial distribution of the local electric field reorientation relaxation time, $\tau_E(x)$, also shows a decent consistency with that of the water rotational relaxation time derived from the stochastic analysis (the orange line in **Figure 4-5** and **Figure 3-5 (a)**). This agreement strongly consolidates that the local water rotation in the HL is ruled by the slow variation of the electric field in the HL. One striking example is the CRW, which merely stays within the HL, has almost no diffusion, and rotates synchronously with the exerted electric field. This is consistent with the MD simulation from S. Balasubramanian and B. Bagchi[163]. In their report, the extremely slow mean-square-displacement of water molecules near the micellar surface indicated that the transient partial localization was reflected in their DACF. In contrast, the proposed electrostatic perspective in this dissertation demonstrates a more quantitative agreement.

Moreover, using tryptophan (Trp) as a local optical probe, the TDSFF experiment has enabled the measurement of the dynamical properties of the protein hydration water near various amino acid side chains. Through this technique, Qin et.al found the heterogeneous dynamics spanning from a few hundred femtoseconds to a few hundred picoseconds[104], which is consistent with the wide range of $\tau_E(x)$ (**Figure 4-5**) in the HL, *i.e.* the relaxation time of water rotation in the HL. As a non-invasive method, the local filed reorientation by MD simulation presented here provides insight into the local hydration dynamics, which is pretty interesting.

In contrast to the observation inside of the HL, the reorientation relaxation time of the local electric field, $\tau_E(x)$, outside the HL, is almost 20 times faster (0.25 ps) than the water rotational relaxation time (a little over 5 ps).[158] Interestingly, the averaged $\tau_E(x)$ in pure water is also merely 0.25 ps (standard deviation: 0.07 ps), which is much

smaller than the relaxation time of pure water or HMW (~ 4.96 ps). The extremely fast relaxation time of the local electric field outside the HL as well as in pure water suggests that the reorientation of the electric field outside the HL and in pure water is practically a random process. Clearly, the orientation of any dipole in such a randomly changing electric field should show the relaxation process intrinsic to the dipolar molecule.

As introduced in **Sec. 1.3.1**, Laage et.al have found that even water molecules near the hydrophobic residues of the protein exhibit slower reorientation dynamics than bulk water, due to the excluded volume effect[98]. Moreover, hydrophilic groups induce a marked slowdown of the water dynamics by accepting strong HBs from water[98]. Water rotational dynamics was also reported to most positively correlate with the tetrahedral structure, although the anti-correlation is observed for the hydrogen-bonded water molecules. Persson et.al also performed a systematical study on the correlation between relaxation time and confinement, polarity, and charge properties of the proximal residues[48]. The inspection of the local structural effects will foster our understanding of the rotational dynamics of the hydration water. The relation between these surface properties and the rotational relaxation time of nearby water merits further study and might be elucidated from the perspective of local electric field reorientation (e.g., EACF) provided here. For instance, the local electric field reorientation relaxation times, $\tau_E(x)$,within the HL plotted in **Figure 4-5** also cover those sites near the solvent-exposed hydrophobic residues on the studied lysozyme (Ala11, Tyr 23, Trp123, Ile124, and His15). Nevertheless, the local field reorientation therein is also visibly slower than that outside the HL which reorientates randomly with a universal time constant of merely 0.25 ps, and this observation might shed some light on why water rotation is also slower even approximate to the hydrophobic groups compared with the bulk water.

4.6 Water relaxation time versus residence time

Considering the prominent discrepancies in the local electric field reorientation inside and outside the HL, the rotational relaxation time of PRW should be related to its translational movement during which it moves into and out from the HL. To quantify the relationship here, the rotational relaxation time of PRW and the normalized total residence time ($t_{R,\,Total}$) were compared. In **Figure 4-6**, the rotation relaxation times of all water in lysozyme solutions are plotted as a function of the normalized residence time. If not specified, residence time (t_R) indicates the normalized total residence time.

The rotational relaxation time of water nearly exponentially depends on the length of its residence time, and the rotation relaxation time of water in protein solutions converges to that of pure water as $t_{R,Total}$ becomes shorter. Since all water molecules within each solution are plotted here, several points even around 10 ps are observed when $t_{R,\,Total} = 0$. However, it should be noted that **Figure 4-6** does not reflect the density of data points. The rotational relaxation times of HMW molecules ($t_{R,\,Total} = 0$) are de facto centered at around 5 ps identical to that of pure water with an extremely small deviation (see **Figure 2-1**).

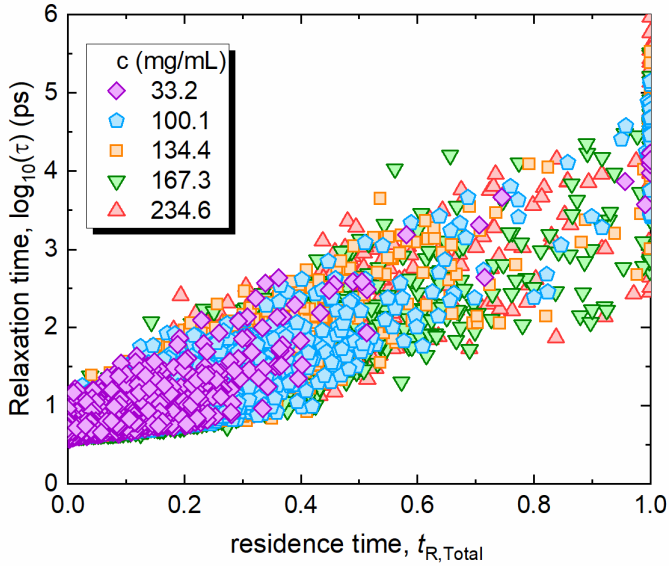

Figure 4-6. Correlation between rotation relaxation time and residence time for individual water molecules in lysozyme solutions. The abscissa, residence time t_R, is normalized by the sampling time (5 ns).

While at the other extreme ($t_{R,Total} = 1$: CRW), the rotation relaxation times prominently range from several hundred picoseconds to dozens of nanoseconds. As introduced above, this heterogeneous water rotational dynamics within the HL has been suggested to pertain to the discrepant confinement degree, hydropathy, polarity, and charge neutrality among the local surface residues of the protein[48, 94, 100]. At all events, in general, depending on the residence time, the averaged water rotation time, $\tau_{D,int}$, is interpolated between the relaxation time controlled by the bulk-like intrinsic solvent electric field outside the HL and that controlled by the extremely slow protein's electric field inside the HL. I also hope the endogenous electrostatic field reorientation perspective provided here could contribute to a better understanding of the rotational relaxation of dipolar molecules and could also help with the increasing scrutiny of the effect of extraneous electric fields on water hydrations surrounding biomolecules.[164-165]

4.7 Summary

The present MD simulation has probed the driving mechanics behind the water rotation dynamics.

By analyzing the reorientation of the exerted electric field on every single water molecule, a global excellent static alignment (~ 16°) between the water dipole and the exerted electric field is revealed. I find for the first time that the relaxation time of water rotation is equivalent to that of the reorientation of the exerted electric field for each water molecule, regardless of its trajectory. Namely, water rotation synchronizes with the experienced electric field reorientation. By separating the exerted electric field generated by protein and ions ($\boldsymbol{E}_{\mathrm{PI}}$) from the overall electric field ($\boldsymbol{E}$), it is found that the electric field from protein governs the rotational relaxation of water which stays within the HL invariably (CRW). When water frequently drifts outside of the HL (PRW), the effect of protein continues to decline. For the other extreme water never traversing the HL, the effect is already inconsiderable.

Bringing the MD's superiority into full play, the reorientation process of the local field (electric field at fixed positions) is depicted. The corresponding relaxation time is shown to range from a few hundred femtoseconds to a little longer time than one hundred picoseconds. The observed results show that the severe retardation of water rotation within the HL is probably attributed to the slow reorientation of the local electric field. For the very few CRW molecules which always stay in the HL, it is almost merely contributed from the protein's field. On the other hand, outside the HL, the rotational relaxation is predominantly determined by the electric field of the surrounding water which randomly reorientates akin to the case in pure water. This observation underlies the discovery that water rotational relaxation time is approximately exponentially dependent on its total residence time (time water spends within the HL). The relationship between water relaxation time and residence time will have a meaningful impact on controlling water dynamics in protein solutions.

Chapter 5 Water Dynamics around Sugar

Sugar additives are effective in preserving protein drugs by slowing down the surrounding water dynamics, thus reducing protein conformation fluctuations. In this chapter, how sugars affect the rotational dynamics of individual water molecules is studied. First, the main of the above findings on protein solutions are tested on sugars solutions. More importantly, the slowdown effect of sugar on water relaxation time is illustrated by its effect on water residence time distribution, based on the findings in **Sec. 4.6.** The importance of sugar dipole is also discussed in this chapter.

5.1 Method: MD simulations

5.1.1 MD simulation setup

GROMACS is a molecular dynamics package mainly designed for simulations of proteins, lipids, and nucleic acids. It is straightforward to import the lysozyme structure (PDB file), implement the empirical force field, and solvate it. However, it is not easy for some small molecules (other than water and ions) such as sugars and sugar alcohols. The necessary topology file for MD simulation contains not only its structure information but the bonded and unbonded interactions information. Thus, the topology file of sugar for MD simulation in GRMACS was prepared by the CHARMM Graphical User Interface (CHARMM-GUI) (http://www.charmm-gui.org)[166-168]. The embedded Ligand Reader & Modeler module can build the desired structure and generate various ligand structures and carbohydrate topology files for various MD simulation packages, such as NAMD, GROMACS, Amber, OpenMM, and so on. Initial structures of all simulated sugars (glucose, trehalose, sucrose, maltose) and sugar alcohol (sorbitol) were downloaded from the GlycoSciences Database (http://www.glycosciences.de/).

The sugars and sugar alcohol were described by the CHARMM36m force field[169-170] and the simple point charge (SPC/E) model was used for water[128]. For these carbohydrate-water systems, tens of solute molecules were solvated in a cubic box of length around 3 nm with 6~10 hundreds of water molecules, corresponding to a mass percent from 10 to 40 wt%. After two energy minimization processes with the steepest descent algorithm, the system was equilibrated 10 ns in the NVT ensemble (298.15 K) and another 10 ns in the NPT ensemble (298.15 K, 1 bar). During the equilibration stage,

all heavy atoms were restrained with the Lincs algorithm[131]. Finally, I performed a long 100 ns NPT simulation with a 2-fs time step and took the last 5 ns track as the sample populations for post analysis. A cutoff radius of 1 nm was used for the nonbonded Lennard−Jones interactions and neighbor searching and the updating frequency is 20 fs. The temperature was controlled using the modified Berendsen thermostat[132] (coupling time constant of 0.1 ps) and the pressure using the Parrinello-Rahman barostat[133] (coupling time constant of 2.0 ps). The electrostatic interactions were calculated by the Particle Mesh Ewald (PME) method[134-135] with a Fourier grid spacing of 0.16 nm and a 1 nm cutoff for the short-range part in real space.

5.1.2 Self-diffusion coefficient

For any particle, the diffusion coefficient, *D*, can be calculated through the Einstein equation[171] below:

$$D = \lim_{t\to\infty} \frac{1}{6Nt} \sum_{i=1}^{N} \left\langle \left| r_i(t) - r_i(0) \right|^2 \right\rangle \tag{5-1}$$

where t is time, r_i is the position of the particle center of mass, and N is the number of particles of a given type. The summation on the right hand of **Eq. (5-1)** is the mean square displacement (MSD) of particles from a set of initial positions.

$$MSD(t) = \frac{1}{N} \sum_{i=1}^{N} \left\langle \left| r_i(t) - r_i(0) \right|^2 \right\rangle \tag{5-2}$$

5.2 Self-diffusion coefficient of water and sugars

Figure 5-1 shows the MSD of trehalose and water in a 10 wt % solution, respectively. Except for the very initial stage (0~100 ps), the MSDs of both trehalose and water increase linearly with time. Fitting the MSD to a linear relationship with the range [20, 60 ns], the self-diffusion coefficient of trehalose and water is 0.37±0.08 and 2.21±0.08, respectively.

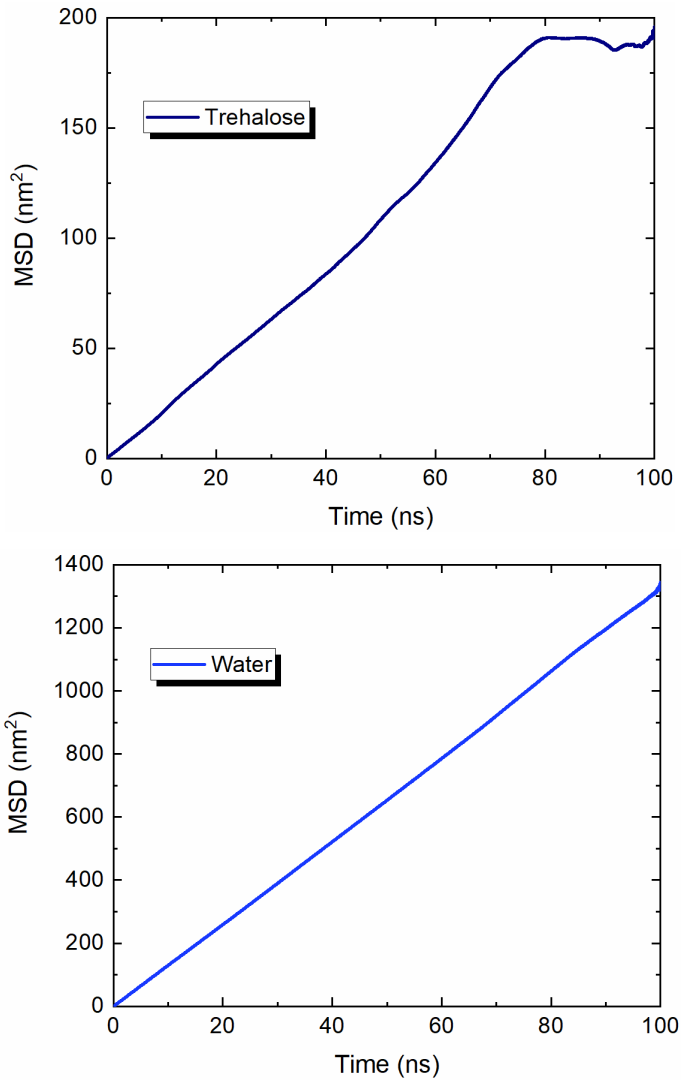

Figure 5-1. MSD of trehalose (**Top**) and water (**Bottom**) in 10 wt% trehalose solution.

Figure 5-2 summarizes the self-diffusion coefficient of both water and sugar in trehalose solutions and sucrose solutions with mass percent ranging from 10 to 40 wt%, calculated from the Einstein equation by MD. Results measured by NMR experiments[126, 172] are also plotted for comparison. It should be noted that Rampp *et al.*[172] measured the self-diffusion coefficient at 303 K and Ekdawi-Sever *et al.*[126] at 298 K, while the present MD was performed at 298 K. **Figure 5-2** shows that the MD simulations agree well with the measurement of Ekdawi-Sever *et al.* (olive symbols) in molecular translational motion at the same physical situation (10 wt% and 30wt% trehalose solutions, and 10 wt% and 30 wt% sucrose solutions, 298 K), which validates the MD protocols employed in this section.

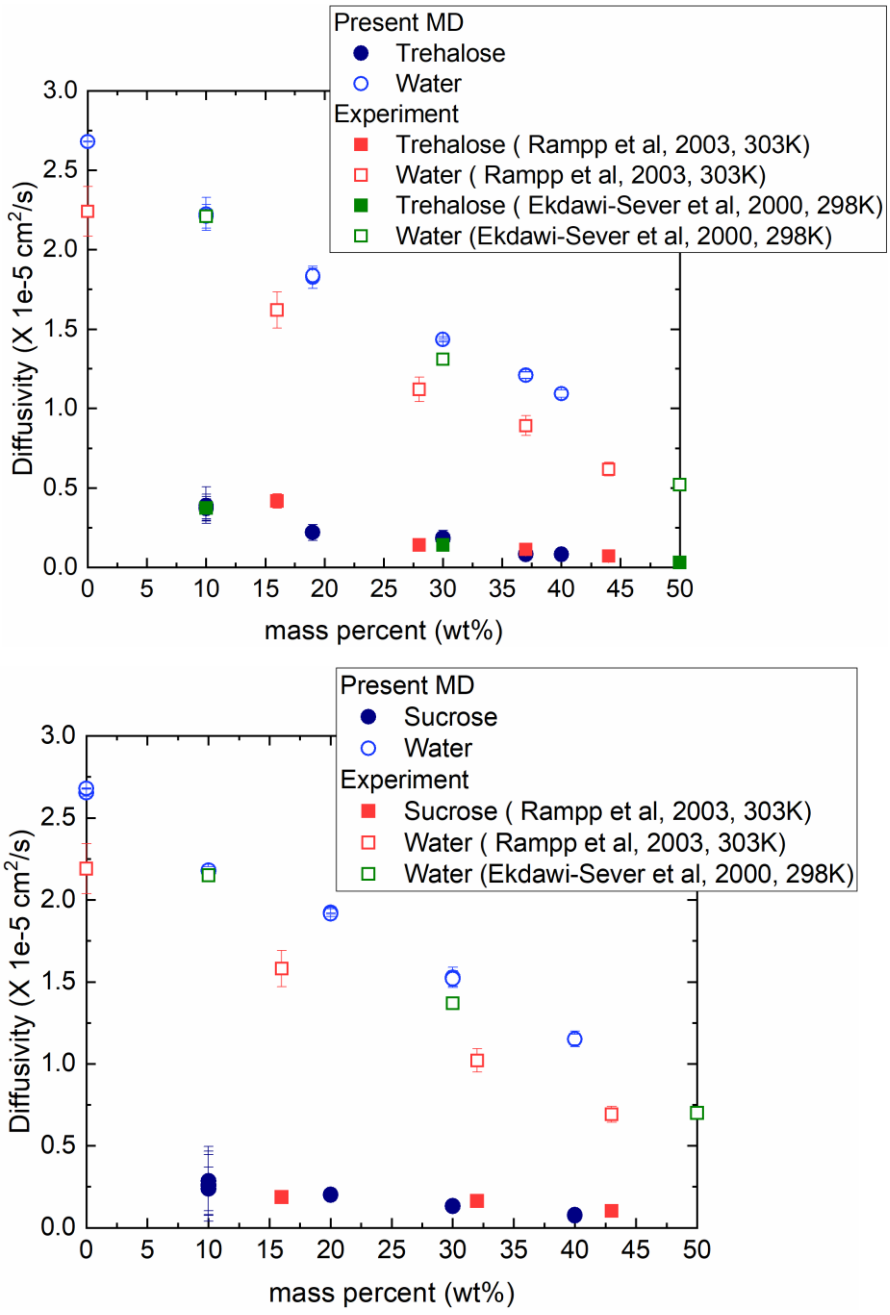

Figure 5-2. Comparison of self-diffusion coefficient between MD and NMR measurements (by Rampp *et al*.[172] and Ekdawi-Sever *et al*.[126]). The solid symbols indicate the sugar solute, and the hollow symbols indicate the water solvent. The circles indicate results from MD simulation, and the squares indicate experimental results. (**Top**) Trehalose; (**Bottom**) Sucrose.

5.3 Water relaxation time

5.3.1 Probability distribution

Figure 5-3 shows the PDF of water relaxation time in various concentrations of glucose solutions and trehalose solutions. In dilute sugar solutions (30 wt% or below), water relaxation time exhibits a normal distribution. When it is more concentrated (40

wt%), relaxation time distribution is slightly distorted to a longer time scale like the case in protein solutions. In all concentrations of monosaccharide and disaccharide solutions, water relaxation time ranges from around 3 to about 40 ps. In addition, in dilute solutions (10 and 20 wt%), water in glucose solutions and trehalose show an almost identical distribution. With the increasing concentration, the percent of the water with relaxation longer than 10 ps in trehalose solution is found to be higher than that in glucose solutions, i.e., water in trehalose solutions exhibits broader distribution.

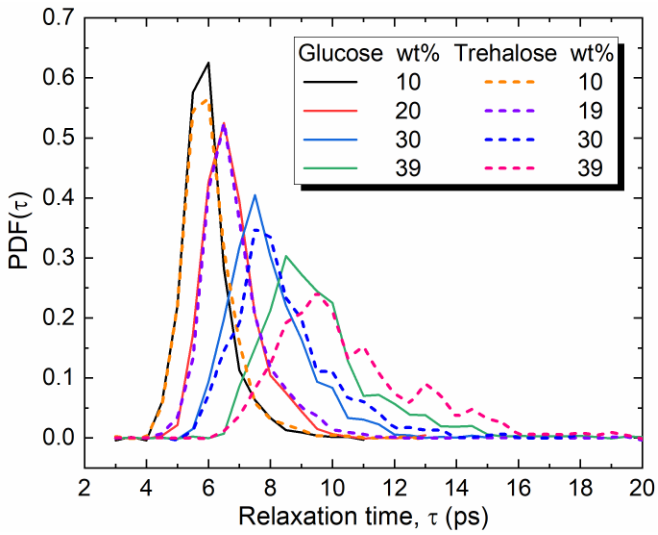

Figure 5-3. PDF of water relaxation time in various concentrations of glucose solutions and trehalose solutions. Only the probability distribution of relaxation time less than 20 ps is shown here.

It is also worth noting that in all concentrations of monosaccharide and disaccharide solutions, water relaxation time ranges from around 3 to about 40 ps. This range is extremely narrow compared with the case of water in lysozyme solutions where some water molecules with an extremely slow relaxation time of several hundreds of picoseconds or tens of nanoseconds have been observed[48, 88, 158]. In addition, at 10 wt%, the main peak of water relaxation time PDF of lysozyme solution (**Figure 2-1**) is around 5.2 ps[158], while that of sugar solutions (glucose and trehalose) is around 6.0 ps. At ~20 wt%, the main peak of lysozyme solution is about 5.3 ps[158], while that of sugars solutions is around 6.6 ps. This finding demonstrates that although a few water molecules in lysozyme solution are severely retarded by hundreds or even thousands of times, sugars slow down the water rotational dynamics more effectively on the whole.

5.3.2 Comparison of dielectric spectra between MD and BDS

With the same method explained in **Sec. 2.1.3**, dielectric loss spectra calculated by the solvent component from MD simulation for aqueous trehalose solutions are shown with a black solid line in **Figure 5-4** , where the red dashed-dotted lines indicate the corresponding fitting curves by a single Debye function. For dilute sugar solutions (< 20 wt%), as shown in **Figure 5-4** (**a**) and (**b**), the calculated dielectric loss is well fitted by the single Debye function (**Eq. (2-10)**) with merely a small residual error. When the solute density increases, this residual error by fitting with the Debye function becomes larger, especially in the lower frequency region. However, the dielectric loss spectrum has a single peak without any shoulder-like shape, even in the most concentrated case.

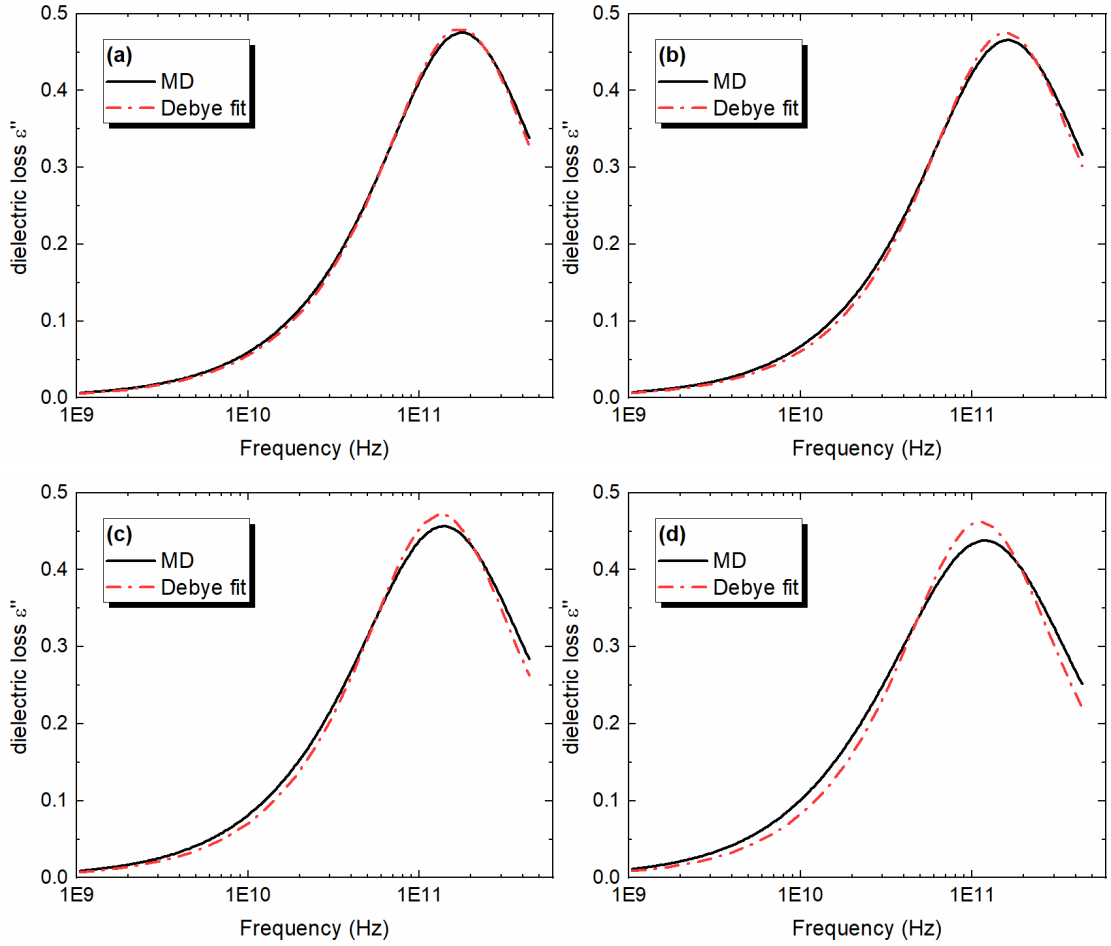

Figure 5-4. Dielectric loss spectra of water in various concentrations of trehalose solutions from MD simulation using **Eq. (2-6)**. Debye fit (red dashed-dotted line) is also shown for each case with its relaxation time. (a) 10 wt%, (b) 20 wt%, (c) 30 wt%, and (d) 40 wt%.

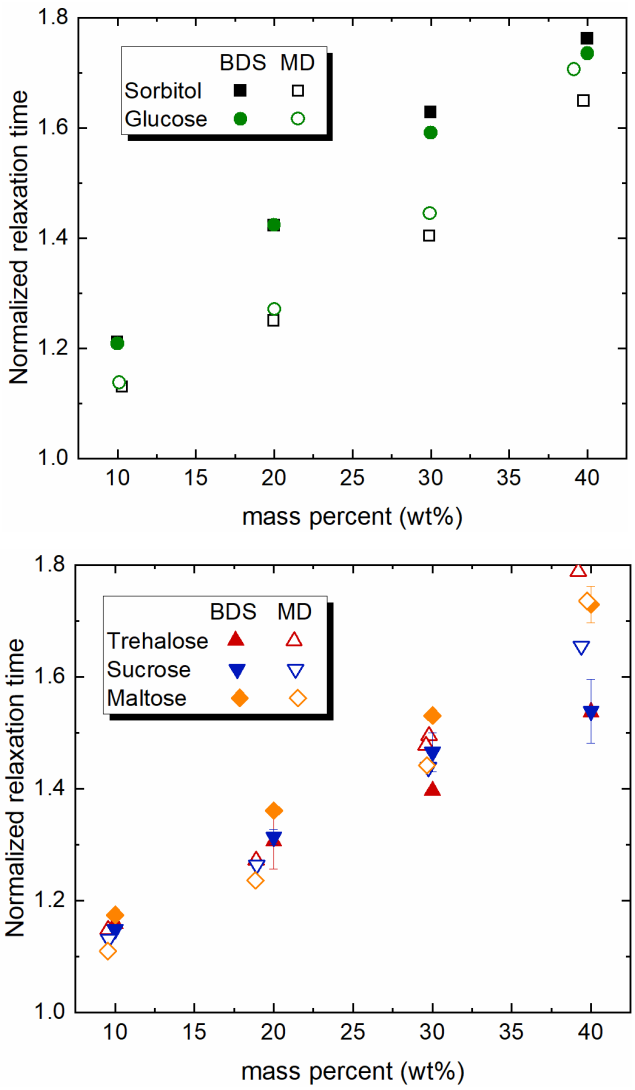

Figure 5-5. The normalized relaxation time of various concentrations of sugar/sugar alcohol solutions. (**Top**) sorbitol and glucose; (**Bottom**) disaccharides. Solid symbols depict the normalized dielectric relaxation time of water rotation calculated by **Eq. (2-6)** using the results from the MD simulation. The swallow symbols depict the normalized γ dielectric relaxation time measured by BDS with the error bar[173]. Each relaxation time was normalized by the calculated or measured relaxation time of pure water.

Results from the MD simulation (**Figure 5-4**) are compared with the γ relaxation time measured by BDS[173] for various concentrations of sugar solutions, as shown in **Figure 5-5**. As in the case of protein solutions, all the relaxation times are normalized by the relaxation time of pure water, respectively, i.e., 4.96 ps for the calculated value and 8.27 ps for the measured value. Both normalized relaxation times increase monotonically with the sugar concentration, and they agree well within the maximum relative difference of 16% resulting from the most concentrated trehalose solution (40 wt%). Although the difference between MD simulation and BDS experiments is acceptable for sorbitol and glucose solutions, γ relaxation time obtained from MD is

always faster than that from BDS and this small difference is worth further discussion. This observation on sugar and sugar alcohol solutions cements the finding that the collective effect of the self-correlation of water rotational movement is responsible for the dielectric loss peak in the gigahertz region (γ relaxation part) in protein solutions.

5.3.3 Electric field-water dipole alignment

Figure 5-6 shows the averaged relaxation times (**Eq. (4-6)**) of individual water rotation, $\tau_{D,int}$, vis-a-vis that of the exerted electric field reorientation, $\tau_{E,int}$, for sugar solutions and sugar alcohol solutions with the mass percent of 10 wt%. For each water molecule,$\tau_{D,int}$ is nearly equal to $\tau_{E,int}$ within around 1psec. For several very few water molecules, the discordance appears but the proportion of such water molecules is merely 1.1% which is negligible. This indicates that for water in sugar and sugar alcohol solutions, the water molecular rotation movement also synchronously follows the reorientation of the exerted electric field (**Figure 4-3 (d)**). Whether the solute is charged (for instance, lysozyme) or electrically neutral (such as sorbitol, glucose, trehalose, sucrose, maltose), this result demonstrates that the electrostatic interactions play a dominating role in the water rotational relaxation.

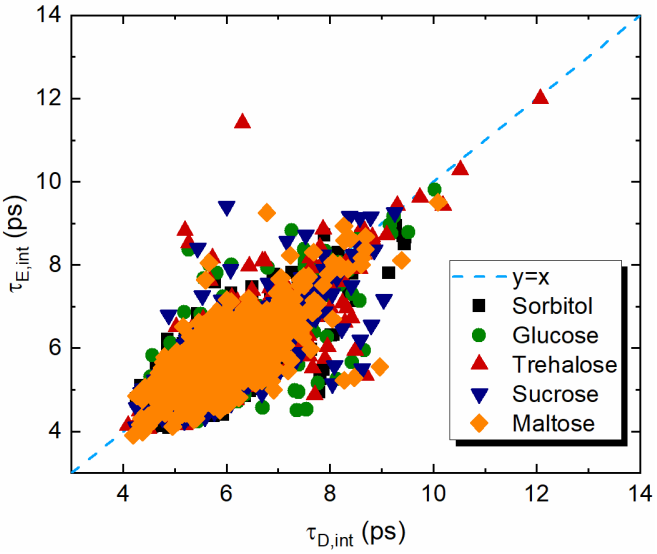

Figure 5-6. Comparison of relaxation time of dipole rotation ($\tau_{D,int}$) and that of the exerted electric field reorientation ($\tau_{E,int}$) for every single water molecule in various concentrations of sugar solutions and sugar alcohol solutions (10 wt%). The black solid line represents the function $\tau_{D,int} = \tau_{E,int}$.

5.3.4 Water relaxation time versus residence time

By terahertz absorption measurements, Heyden *et al*. found that the THz absorption of the hydration water is increased in a range of 3.7 ± 0.9 Å and defined the HL thickness of 3-4 Å, for monosaccharide solutions[174]. EDLS experiments by Lupi et.al suggested the evident retardation of water within 3.1Å from any solute atom relative to the bulk water, for monosaccharide and disaccharide solutions[123]. Therefore, the sugar HL in this dissertation is also adopted as a 3 Å water layer from the sugar surface. Consequently, the residence time ($t_{R,Total}$) of each water molecule in all the sugar and sugar alcohol solutions studied here is obtained.

The relationship between water rotational relaxation time and residence time is depicted in **Figure 5-7** for various sugar solutions and sugar alcohol solutions. Water relaxation time shows a decent exponential correlation with its residence time, for all the solutions studied here. The correlation between the relaxation time and the residence time is approximated by an exponential equation as follows,

$$\tau/\tau_{PW} = 10^{\frac{t_{R,Total}}{\xi}}. \quad (5\text{-}3)$$

Here, τ_{PW} is the relaxation time of pure water (5.0 ps). ξ is the decuple-retarding period which represents the prerequisite residence time for water relaxation time to become 10 times that of pure water. In short, it indicates how effectively the solute molecule retards the rotational relaxation of the water molecule. It is straightforward that the smaller the ξ is, the more effectively the water relaxation time increases with its residence time. The respective fitting result is plotted in **Figure 5-7** with a solid black line, and the fitting expression is also denoted for each sugar solution. The intercept of all solutions corresponds to the relaxation time of HMW, which is also ~5.0 ps. The decuple-retarding period is around 2.5 for sorbitol and monosaccharide (glucose), while it is about 1.8 for disaccharide solutions. Therefore, the water rotational relaxation is retarded more quickly in the HL of disaccharide than that in the HL of monosaccharide, i.e., the increasing rate of the water relaxation time with the residence time is faster in disaccharide solutions than that in monosaccharide solutions. In this regard, the present result coincides with the BDS experiments[116-117], the femtosecond infrared spectroscopy measurement[115], and the calculations of the intermolecular NOE and NQR by MD simulation[122]. Particularly, this observation implies that slowing down water

rotation dynamics might be achieved by shortening the increment rate and elongating the water residence time.

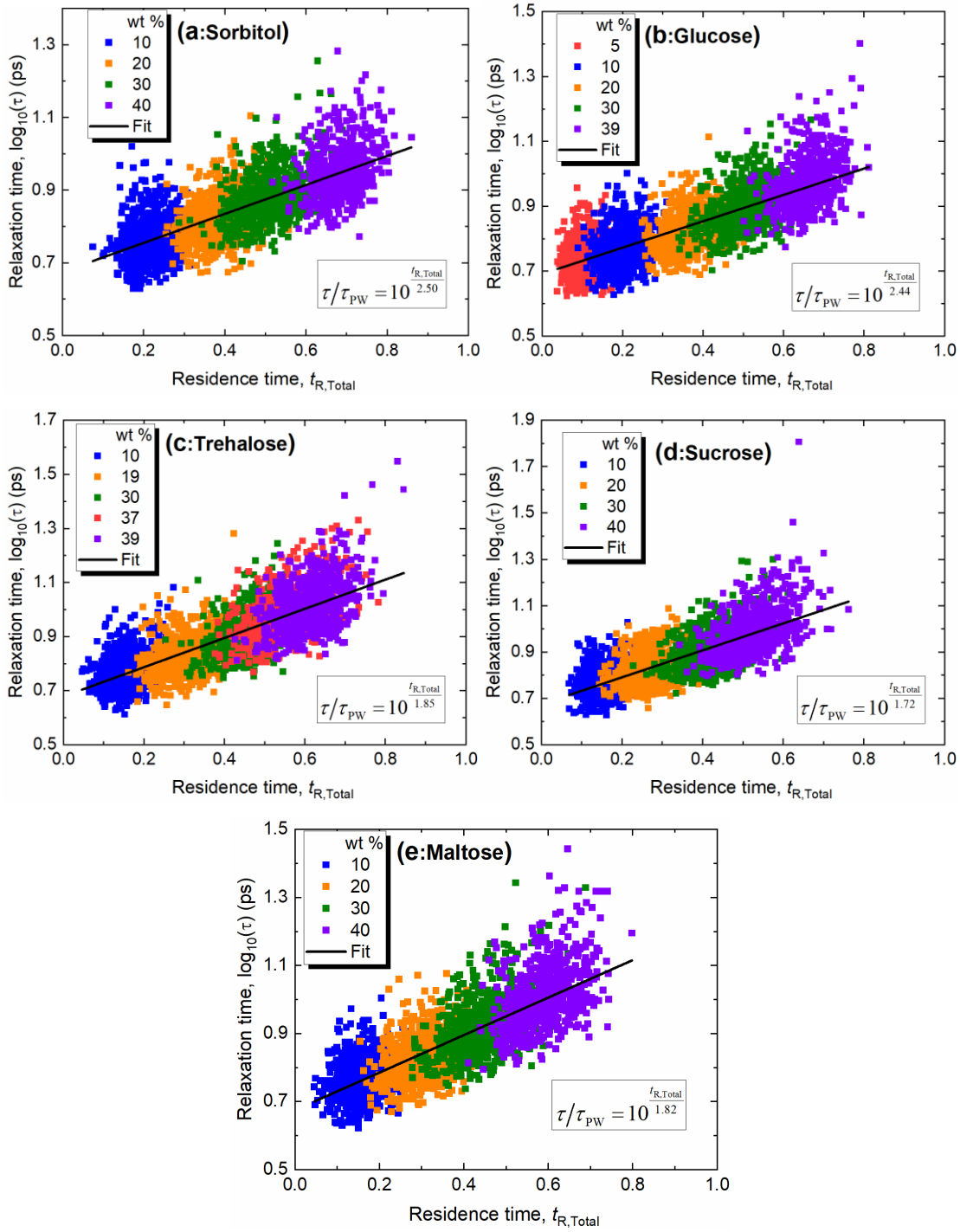

Figure 5-7. Correlation between rotational relaxation time and residence time for individual water molecules in sugar alcohol and sugar solutions. Each symbol represents one single water molecule. (**a**) sorbitol, (**b**) glucose, (**c**) trehalose, (**d**) sucrose, and (**e**) maltose.

5.4 Increment in water relaxation time with residence time (Effect of the dipole moment of protective agents)

Figure 5-7 clearly shows that the water relaxation time exponentially increases with its residence time and the corresponding decuple-retarding period (ξ) rests with the solute species. In consequence, the water relaxation time is different among species even though the residence time is identical. As a kind of dipolar molecule with a permanent dipolar moment, a water molecule can be oriented by the local electric field[107-109]. Water in the vicinity of a solute molecule may react to the solute tumbling, i.e., the solute dipole reorientation. The aforementioned **Sec.4.4** elaborates on the predominant role of protein's electric field on the rotational relaxation of hydration water(CRW). This fundamental physics leads to the idea that the solute dipole moment may be a simplified key factor in characterizing the effect of protective agents on this decuple-retarding period. **Figure 5-8** shows the decuple-retarding period versus the solute dipole moment. It is found that the larger the solute dipole, the shorter the decuple-retarding period, therefore the faster the water relaxation time increases with the residence time. Given that water relaxation time is exponentially proportional to its residence time (**Eq. (5-3)**), upon the increment in solute dipole, water relaxation time increases much faster with the residence time.

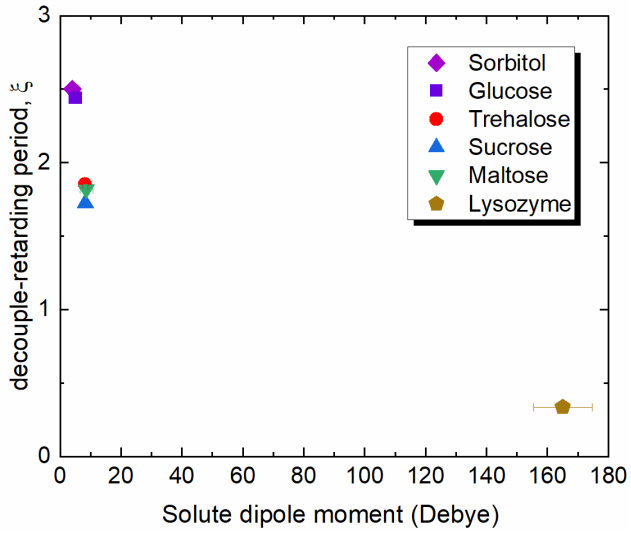

Figure 5-8. Correlation between the decuple-retarding period obtained from the exponential fit in **Figure 5-7** and the dipole moment of various solute molecules. The molecular dipole of glucose is from the calculation of Fulton[175]and the BDS experiment of Shiraga et.al[116]; for maltose, it is also from Fulton[175]; for trehalose and sucrose, they are from Shiraga et.al[1]; for sorbitol, it is from the quantum repository of University of Pittsburgh[176].

5.5 Water residence time distribution (Effect of the size of protective agents)

Not only the water rotational relaxation time in a certain concentration shows normal distribution (**Figure 5-3**), but water residence time for the same concentration also disperses according to **Figure 5-7**. To quantify this dispersion, the PDFs of water residence time in various concentrations of glucose solutions and trehalose solutions are plotted in **Figure 5-9** and **Figure 5-10** with the black lines, respectively. For each condition of these two saccharide solutions, the PDF of water residence time is perfectly fitted to a normal distribution (red line). Interestingly, unlike the deviation of the PDF of water rotational relaxation time that increases with sugar concentration, the deviation of the PDF of water residence time is nearly identical at all concentrations. This observation suggests that the residence time distribution of water in sugar solutions is uniform and that water molecule has an equivalent probability of entering the sugar HLs.

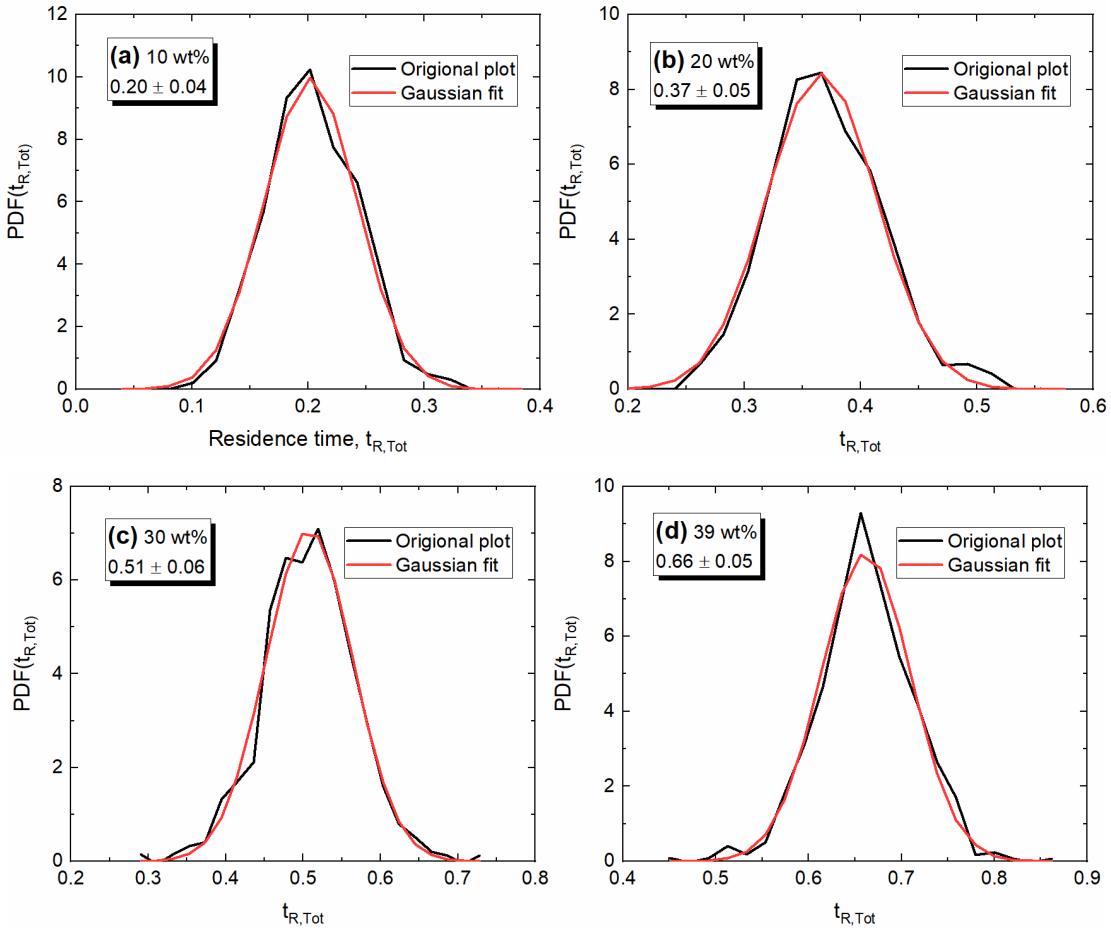

Figure 5-9. PDF of water residence time in various concentrations of glucose solutions. The abscissa shown here is the normalized residence time. The mean value and standard deviation of the Gaussian distribution are shown in the upper left legend of each panel.

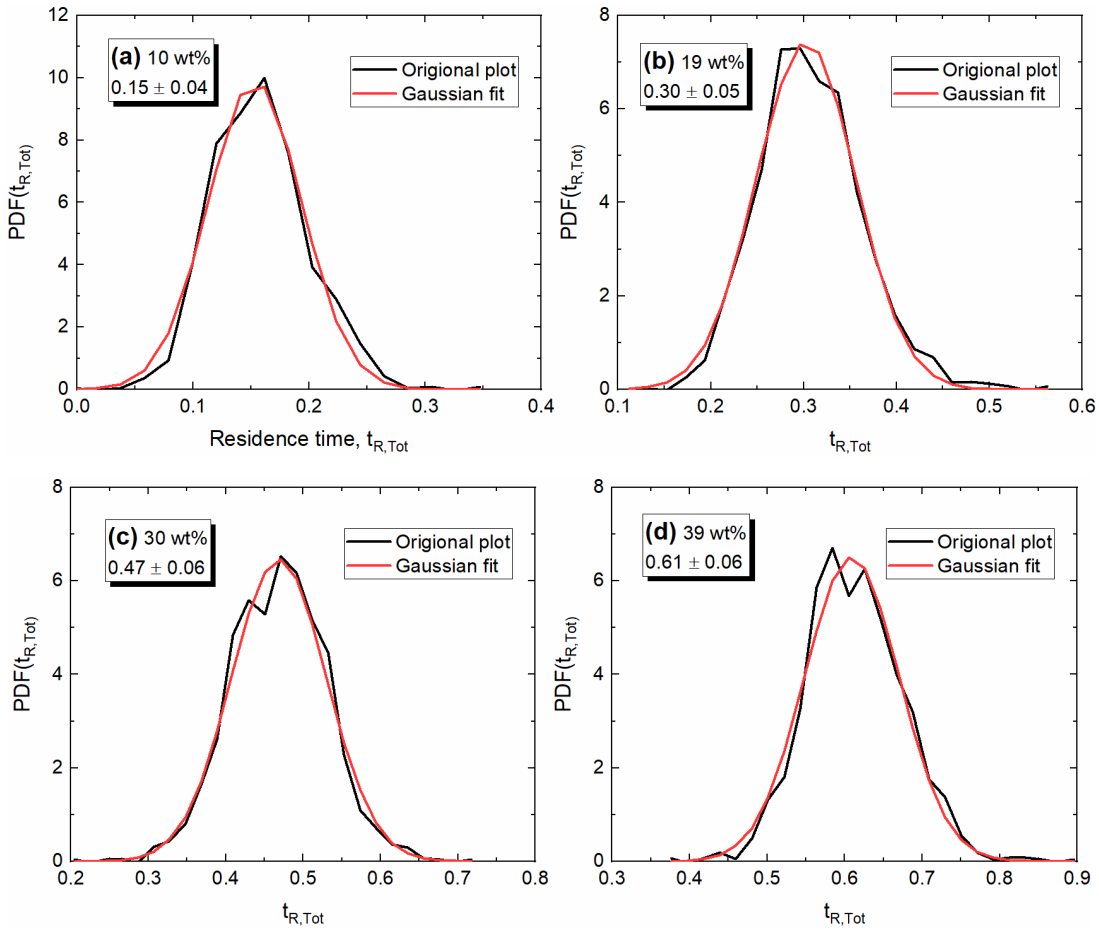

Figure 5-10. PDF of water residence time in various concentrations of trehalose solutions. The abscissa shown here is the normalized residence time. The mean value and standard deviation of the Gaussian distribution are shown in the upper left legend of each panel.

Due to this random walk of water in sugar HL which may be caused by the weak restriction of the water translational movement in sugar HL, no HMW or CRW is observed in sugar solutions within the concentration investigated here. Even in the most concentrated 39 wt% sugar solutions, CRW is not found. All water molecules move in and out of the sugar HLs (free translation), which is markedly different from that in protein solutions where all 3 kinds of water (HMW, PRW, and CRW) coexist in each studied condition. These results are also observed in sorbitol, sucrose, and maltose solutions.

The HL thickness in this study is defined as 3 Å, and the water residence time is defined as the total time when a water molecule stays within the HLs. Now that the normal distribution of water residence time in sugar solutions (**Figure 5-9** and **Figure 5-10**) has suggested that water molecules have the same probability of moving into and out from the sugar HLs (random walk of water), the ratio of sugar HLs volume to the

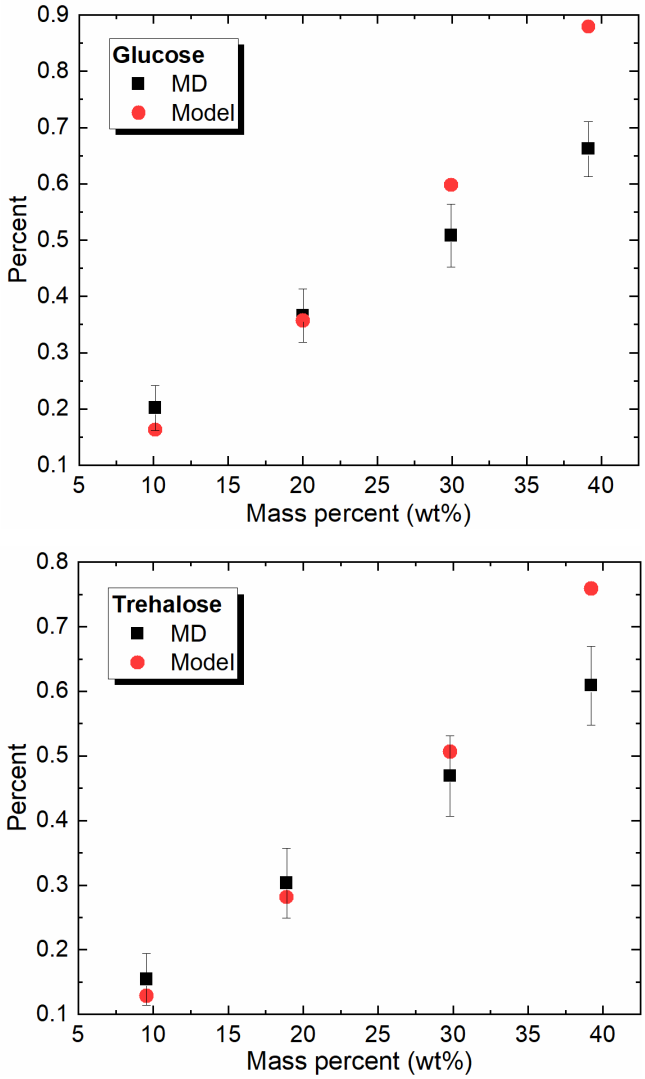

Figure 5-11. The percent of water within the HLs among total water in sugar solutions, obtained from MD (black squares) and a simple geometric model (red circles. The percent of water within the HLs among total water could be regarded as the average residence time of water molecules.

total solvent-accessible space should be equivalent to the normalized water residence time for a given sugar solution. In case that HLs never overlap each other, the volume of sugar HLs is the geometric summation of a single HL space. Consequently, the ratio of sugar HLs volume (V_{HLs}) to the total solvent-accessible space (V_{access}) is

$$\frac{V_{\mathrm{HLs}}}{V_{\mathrm{access}}} = \frac{N_{\mathrm{solute}} \cdot \mathrm{SASA} \cdot \Delta L}{V_{\mathrm{solution}} - N_{\mathrm{solute}} \cdot V_{\mathrm{solute}}}, \tag{5-4}$$

where N_{solute} indicates the number of solute molecules in the solution, V_{solute} the solute molecular volume, SASA the solvent-accessible surface area, ΔL the HL thickness, and V_{solution} the solution volume. This equation could further be expressed as

$$\frac{V_{\mathrm{HLs}}}{V_{\mathrm{access}}} = \frac{\mathrm{SSA} \cdot \Delta L}{1/\left(\rho_{\mathrm{solute}} \cdot V_{\mathrm{solute}}\right) - 1}, \tag{5-5}$$

where SSA is the specific surface area of the solute (=SASA /V_{solute}) and ρ_{solute} is the solute number density (=$N_{\mathrm{solute}}/V_{\mathrm{solution}}$). **Eq. (5-5)** indicates that a larger solute SSA is necessary for a longer residence time. Given that the extended depolarized light scattering experiments together with the MD simulations have shown that glucose could be modeled as a sphere of the radius of 3.41 Å[177] and trehalose of 4.30 Å[178], this ratio in sugar solutions is

$$\frac{V_{\mathrm{HLs}}}{V_{\mathrm{access}}} = \frac{1/r_{\mathrm{solute}} \cdot \Delta L}{1/\left(\rho_{\mathrm{solute}} \cdot 4\pi r_{\mathrm{solute}}^3/3\right) - 1}, \tag{5-6}$$

where r_{solute} is the radius of a single solute molecule.

Figure 5-11 compares the $V_{\mathrm{HLs}}/V_{\mathrm{access}}$ obtained from the geometric model (red circles) and the averaged water residence time obtained from MD simulations (gray squares) for various concentrations of glucose solutions and trehalose solutions. $V_{\mathrm{HLs}}/V_{\mathrm{access}}$ shows good agreement with the mean t_{R} in the dilute sugar solution (20 wt% or below), but evidently overestimates the water residence time when glucose is over 30 wt% and trehalose is over 40 wt%. This overestimation is probably due to the overlapping of HLs at higher sugar concentrations (**Figure 5-12 (Top)**), that is a hydration water molecule may stay within the HLs of two or even more different solute molecules, as reported by Fioretto *et al.*[125].

MD simulation is able to evaluate the area of overlapped HLs in a sugar solution. For instance, for a randomly chosen water molecule in the 10 wt% glucose solution, during the total sampling time (5000.0 ps, 6251 snapshots), it stays in the HLs for 1343 snapshots (1074.4 ps). Concretely, it stays within only 1 HL for 1123 snapshots, within 2 HLs at the same time for 193 snapshots, and within 3 HLs at the same time for 27 snapshots. Divided by the total snapshots it stays within the HLs (1343) respectively, the probability that this water molecule is shared by the HLs of different glucose molecules when it is in the HLs is obtained, as shown in **Figure 5-12 (Bottom)**. Taking the average over all water molecules, the probability that a hydration water molecule is shared by the HLs of several sugar molecules is derived and shown in **Figure 5-13 (a)**.

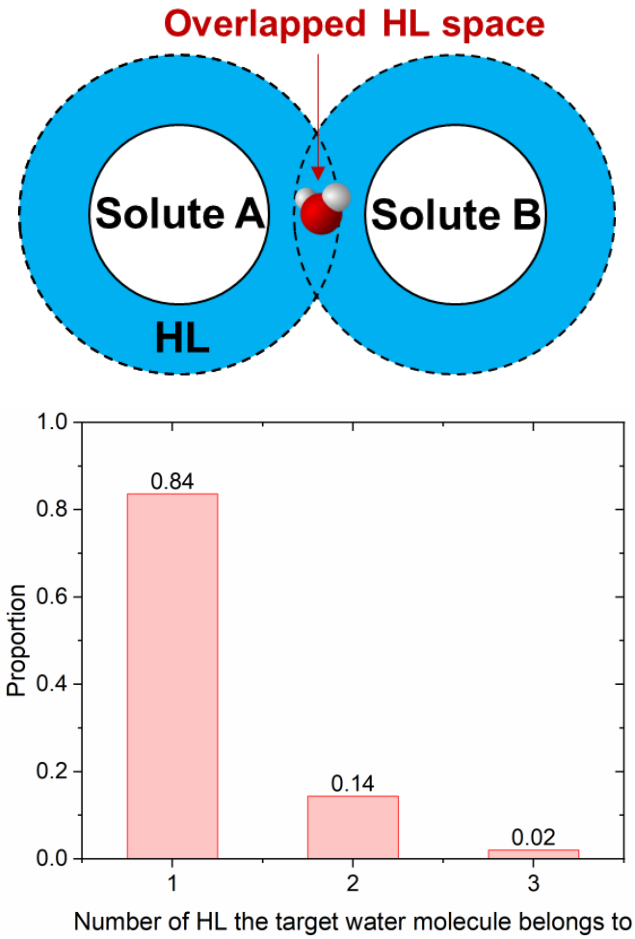

Figure 5-12. (**Top**) Schematical illustration of the HL overlapping. (**Bottom**)The probability that a water molecule (#8) in the 10 wt% glucose solution is shared by the HLs of different glucose molecules when it is in the HLs.

Multiplied by the average residence time, $t_{R,Total}$, which corresponds to the probability for water to be in HL (see **Figure 5-9** and **Figure 5-10**), the probability that a water molecule is shared by the HLs of different sugar molecules is shown in **Figure 5-13 (b)**. It should be emphasized again that there is no HMW (water always staying outside the HLs) in the sugar solutions studied here. From **Figure 5-13**, it is observed that 19.93% of water is shared by 2 HLs and that 5.62% of water is even shared by 3 HLs at the same time in the concentrated 39 wt% glucose solution. For 39 wt% trehalose solution, the corresponding sharing rates are 16.53% and 3.28%, respectively. In the 39 wt% glucose and trehalose solutions, the overall overlapping percent reaches 23.78% and 17.53%, respectively. Therefore, in the concentrated sugar solutions, the overlapping of HLs is non-negligible and the estimation of such an overlapping rate merits further study to predict the water residence time. Nonetheless, the normal distribution of water residence time (viz. the random walk of water within the sugar

solutions) suggests that the ratio of the hydration volume (total volume of HL) to the overall water accessible space equals the water residence time. This indicates the significant role of solute size in affecting water residence time.

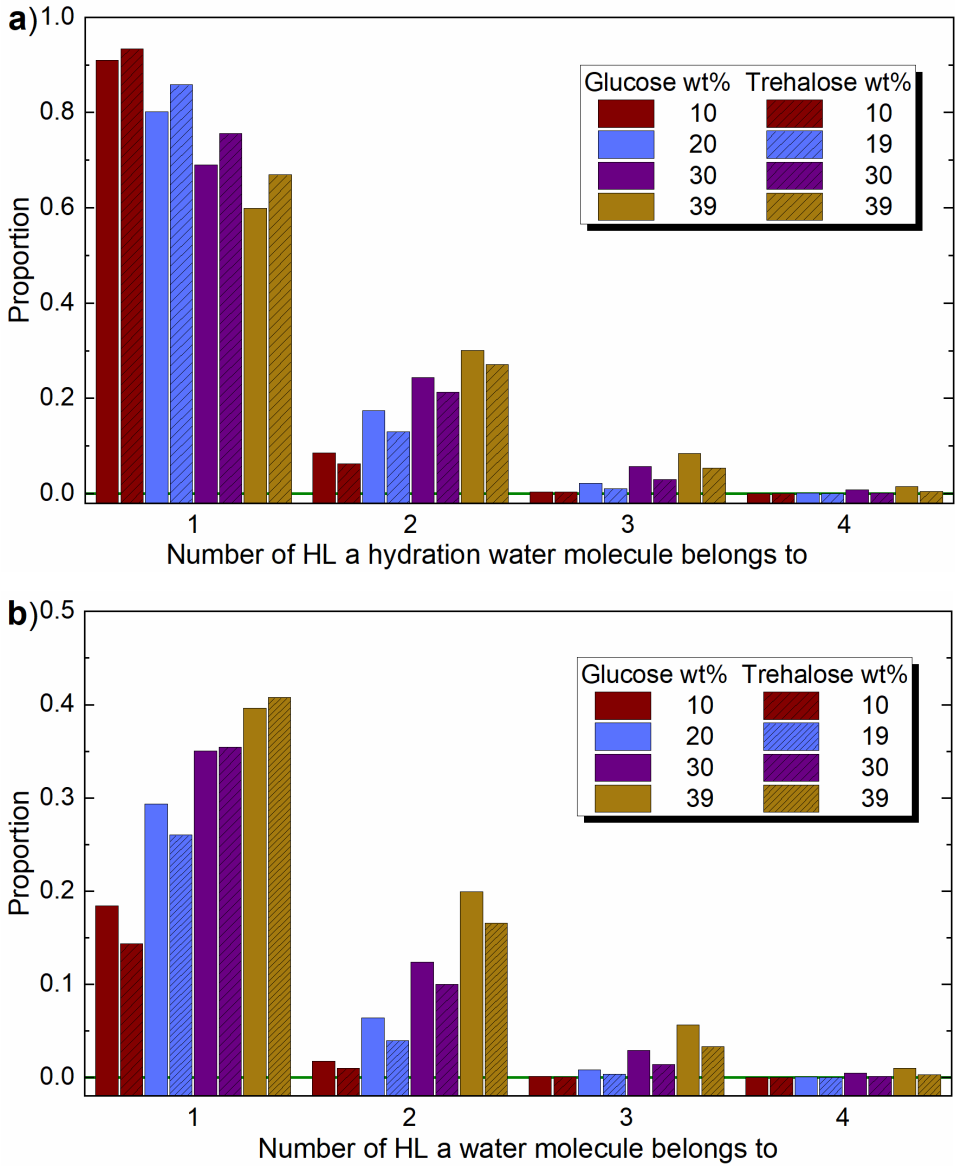

Figure 5-13. (**a**) The probability that a hydration water molecule is shared by the HLs of different glucose/trehalose molecules. (**b**) The probability that a water molecule is shared by the HLs of different glucose/trehalose molecules.

5.6 Trade-off between the molecular size and dipole moment of protective agents

In previous work, It was shown that the protein deterioration rate is most sensitive to the γ dielectric relaxation time[33] and that this γ relaxation originates from the rotational ACF of all water molecules[158]. In light of these discoveries, the water relaxation time distribution shall work as the evaluation criterion for protective agent

effectiveness. In the above **Secs. 5.4** and **5.5**, it has been established that the electric dipole (**Figure 5-8**) and molecular size (**Eq. (5-6)**)) of the solute molecule have great impacts on the decuple-retarding period and residence time, thus on the water rotational relaxation time. According to **Eqs. (5-3)** and **(5-6)**, water relaxation time could be simply expressed as

$$\log_{10}(\tau / \tau_{PW}) = \frac{1/r_{\text{solute}} \cdot \Delta L}{1/(\rho_{\text{solute}} \cdot 4\pi r_{\text{solute}}^3 / 3) - 1}(1-\alpha) \cdot \frac{1}{\xi} \tag{5-7}$$

, where α is the overall percentage of HLs overlapping within the solutions. For a unique solute molecule, ξ determines the increment rate of water relaxation time vis-à-vis its residence time whilst *r* and ρ dominate water residence time. Although the HLs overlapping rate, α, is also a notable parameter, it is probably related to the solute size (*r*) and density (ρ) based on the geometric definition of HL in the present work. Besides, since molarity is a variable of extensive property not any inherent physical or chemical attribute of the solute, only how solute dipole and size exert their influences simultaneously is discussed here for better guidance in the development of protein stabilizers.

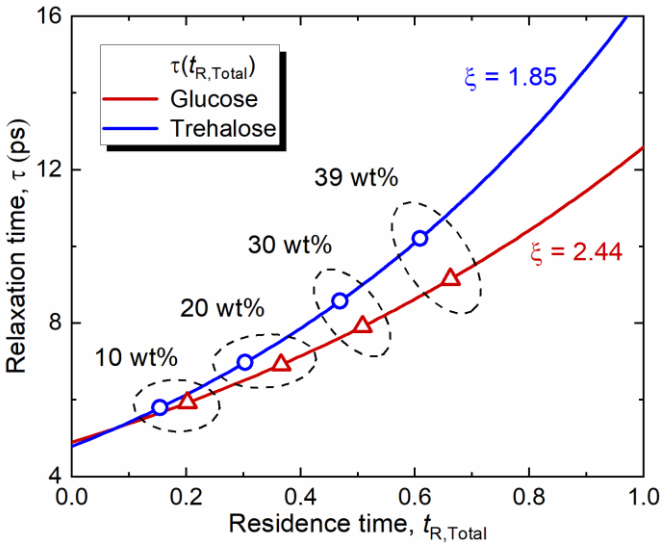

Figure 5-14. Correlation between water relaxation time and residence time (**Eq. (5-3)**) for glucose solutions and trehalose solutions. The abscissa of the symbols is the average water residence time obtained from **Figure 5-9** and **Figure 5-10**.

Figure 5-14 summarizes the relationships between water relaxation time and residence time (**Eq. (5-3)**, plotted with solid line) for glucose solutions and trehalose solutions. A clear comparison between these two solutions with the same mass percent is shown with the symbols. Under each condition, the water residence time in glucose

solutions is ~0.05 higher than that in trehalose concentration, due to the larger specific surface area of the smaller glucose molecule. Although water relaxation time increases exponentially with the residence time and the decouple-retarding period (ξ) for trehalose is 1.3 times smaller than that for glucose, when sugar mass percent is 10 wt and 20 wt%, water relaxation time distributions in these two sugar solutions are almost identical because the slighter longer water residence time distribution in glucose solutions can counteract the slower relaxation time rise rate while the residence time is 0.2~0.4. This observation suggests that a larger specific surface area to extend the hydration layers and thus elongate water residence time is of great significance. When the residence time reaches 0.5 and even larger (concentration: 30 wt% and 39 wt%), the exponential growth rate (viz. the decouple-retarding period) of relaxation time manifests itself. As a result, at these two dense concentrations, the relaxation time PDF for trehalose solutions broadens to the slightly longer side compared with that for glucose solutions (**Figure 5-3**). In light of this, the solute dipole moment which measures the decouple-retarding period seems to be dominant.

Therefore, in the design of a protective agent, the trade-off between its molecular size and dipole moment is key to controlling water rotational dynamics. Ideally, the protective agent should also be designed to possess a larger specific surface area and a stronger electric dipole. It is our hope that the definition of decuple-retarding period (**Eq. (5-3)** and the estimation of water residence time (**Eq. (5-6)**) proposed in the present work could help to better understand the slowdown effect of some protein stabilizers, such as polyol, zwitterionic molecules (TMAO), amid acids, and other carbohydrates. A whole picture of water relaxation time and residence time in solutions of additives from small to large is also expected to shed more light on the design of bioprotective agents.

5.7 Difference in water residence time between protein solution and sugar solution

The above findings provide a principal law for evaluating water rotational relaxation time, thus for developing the bio-protective agents. The period water spends within the HLs of solute molecules (water residence time) is of great significance. Therefore, the understanding of how water behaves on the surface (i.e., in the HLs) of different additives is also noteworthy.

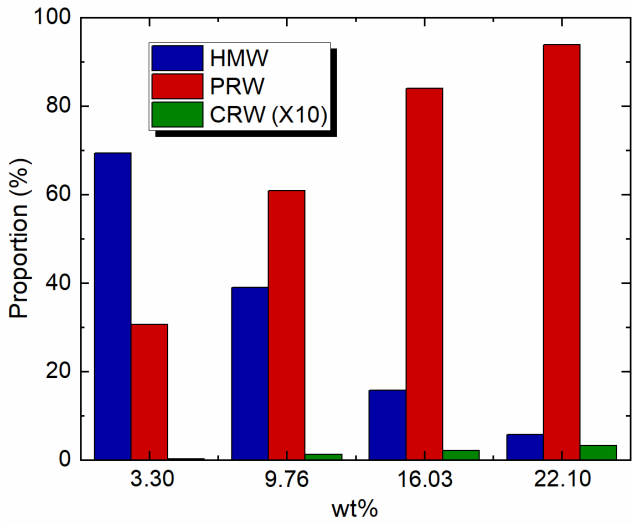

Figure 5-15. Percentage of the 3 typical types of water molecules in the lysozyme solutions (proportion of CRW is magnified 10 times). HMW represents the highly mobile water, PRW the partially restricted water, and CRW the completely restricted water.

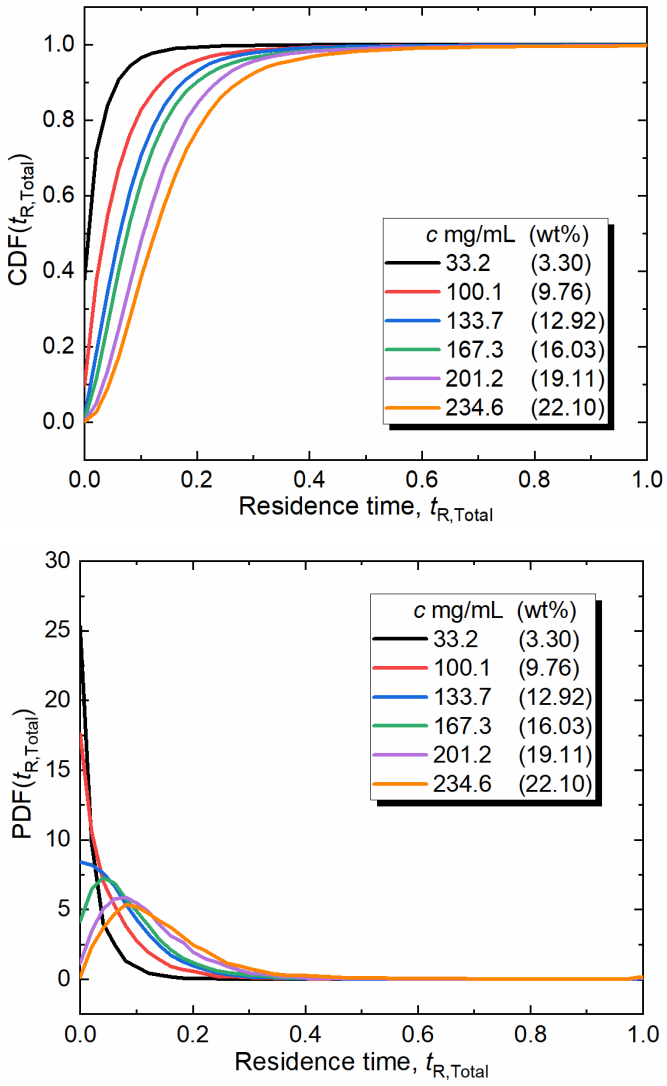

Figure 5-16. CDF and PDF of water residence time in protein solutions. The legend indicates lysozyme mass concentration in unit mg/mL and mass percent.

For example, the aforementioned 3 types of water coexist in lysozyme solutions of from 3.3 to 22.2 wt%. In the most dilute 3.3 wt% lysozyme solution, there are still several CRW molecules that are attracted on the protein surface; in the most concentrated 22.2 wt% protein solution, the HMW still occupies a percentage of 7 % (**Figure 5-15**). The coexistence of the above 3 classes of water indicates that water residence time in protein solutions always ranges from 0 to 1, as shown in **Figure 5-16**. Unlike sugar solutions, the residence time in lysozyme solution has a more distorted distribution (**Figure 5-16**). And due to larger molecular size and smaller SSA, the relative averaged value is even less than half of that in sugar solution at the same mass percent as shown in **Figure 5-14**.

It is also observed that when the concentration is less than 100.1 mg/mL (black and red lines), HMW takes quite a proportion and the PDF at $t_{\mathrm{R,Total}} = 0$ is remarkably high. For instance, around 40% water nerve comes into the protein HLs for 33.2 mg/mL lysozyme solutions. With the increment in protein concentration, the main peak of PDF shifts to a longer residence time. Clearly, the residence time PDFs for all the concentrations studied here exhibit pretty large deviations and are far from the normal distribution, due to the existence of both the HMW and CRW. The following **Table 5-1** summarizes the ratio of the HLs volume to the total solvent space for various concentrations of lysozyme solutions. The molecular SASA and volume of lysozyme used for the calculation are 69±1 nm^2 and 22±2 nm^3, respectively[95]. When comparing the V_{HLs}/V_{access} in **Table 5-1** with the mono peak of the PDF in **Figure 5-16** (c >= 167.3 mg/mL), it is found that V_{HLs}/V_{access} is almost twice the residence time corresponding to the main peak. Therefore, it is evident that the water molecule does not have an equal chance to come into the HL.

Table 5-1. V_{HLs}/V_{access} estimated for various concentrations of lysozyme solutions

wt %	V_{HLs}/V_{access}
3.30	0.030
9.76	0.096
12.92	0.133
16.03	0.173
19.11	0.216
22.10	0.261

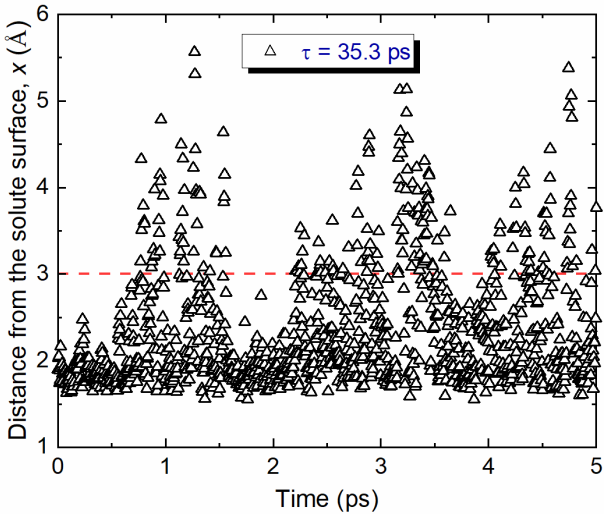

Figure 5-17. Momentary distance from the solute surface *x*, of the slowest water molecule in trehalose solution (39 wt%).

Water residence time in protein solutions is in contrast with the sugar solutions (**Figure 5-9** and **Figure 5-10**) where only PRW exists. Furthermore, even for the PRW, water movement in sugar solutions is also discrepant from that in protein solutions. **Figure 5-17** shows the time-history of the distance from the sugar (trehalose) surface for the slowest water molecule in the studied trehalose solutions here, which has a rotational relaxation time of 35.3 ps and a normalized total residence time of 0.85 (4.23 ns). Although trapped on the solute surface for a period as long as in protein solutions, it drifts in and out of the HLs many times, which means that the total residence time is segmented into plenty of short time spans. The longest continuous period of this water is around 0.8 ns starting from around 1.5 ns.

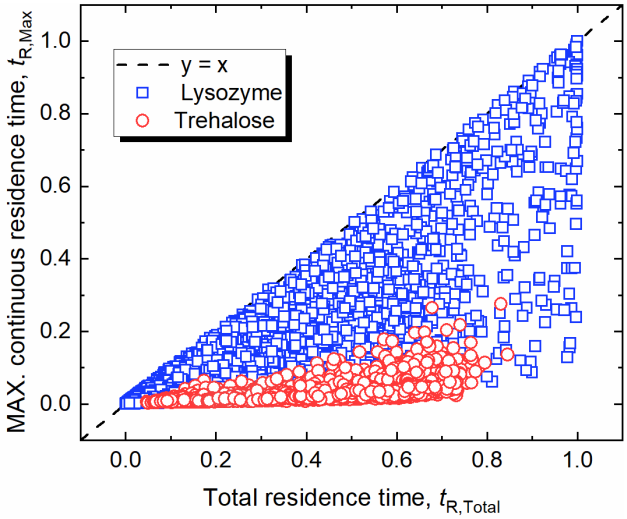

Figure 5-18. Comparison of the total residence time ($t_{R,Total}$) and the maximum continuous residence time ($t_{R,Max}$) of each water molecule in lysozyme solutions and trehalose solutions, including all the concentrations studied here.

By counting the number of the continuous snapshots where a target water molecule resides in the HL, the maximum continuous residence time ($t_{R,Max}$) can be obtained. These two kinds of residence times ($t_{R,Total}$ and $t_{R,Max}$) of all water molecules in lysozyme and trehalose solutions are summarized and compared in **Figure 5-18**. It is found that they could be equivalent for some water molecules in protein solutions. But $t_{R,Max}$ is always much shorter than $t_{R,Total}$ for water in sugar solutions. For water in sugar solutions, the total residence time consists of plenty of short periods, while it is usually contributed by a long commensurate continuous period for water in protein solutions.

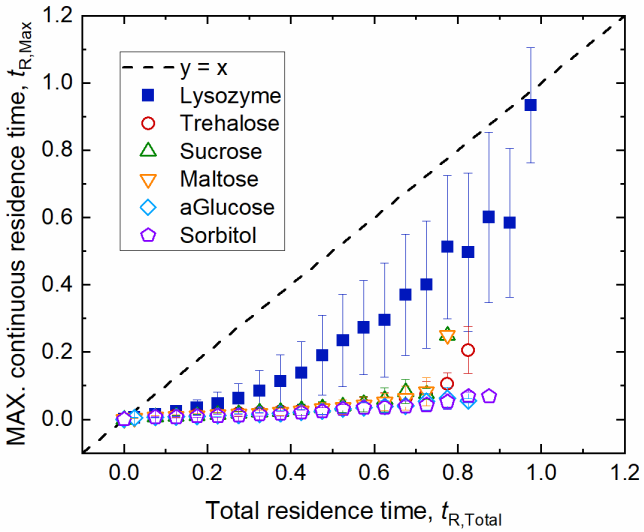

Figure 5-19. Comparison of the total residence time ($t_{R,Total}$) and the maximum continuous residence time ($t_{R,Max}$) on ensemble average, for water in different solutions.

For every single water molecule, a data set ($t_{R,Total}$, $t_{R,Max}$) could be obtained. By setting a series of continuous small intervals for t ($t \pm dt$), the whole data sets among all water molecules (data shown in **Figure 5-18**) were categorized into an interval based on the first parameter ($t_{R,Total}$). Then, the distribution of the maximum continuous residence time ($t_{R,Max}$) was obtained within each interval. The result on the ensemble average with standard deviation is shown in **Figure 5-19**. The maximum continuous residence time ($t_{R,Max}$) is closer to the total residence time ($t_{R,Total}$) for lysozyme solution, compared with sugar solutions or sugar alcohol solutions. For sugar solutions or sugar alcohol solutions, $t_{R,Max}$ increases slowly with $t_{R,Total}$; but for disaccharide solutions, the increment rate becomes higher when $t_{R,Total}$ is larger than 0.7. For lysozyme solutions, $t_{R,Max}$ also increases slowly at first then fast with the increment in $t_{R,Total}$; when $t_{R,Total}$ is larger than 0.3, the slope of the linear increase is around 1. Due to the heterogeneous protein surface properties, $t_{R,Max}$ for lysozyme solution exhibits a larger deviation

compared with other solutions. Nevertheless, it is concluded that sugar does not trap water molecules as strongly as protein does. Some water molecules tend to be trapped on the protein surface for a long time, reducing the possibility of other water coming into the HLs. This is in part due to the depressions or cavities on the protein surface[48, 88].

5.8 Summary

This chapter mainly studies the water rotational dynamics in sugar solutions by MD simulation. A good agreement of the dielectric relaxation time in the gigahertz region measured by BDS experiments and the relaxation time of the self-correlation of water rotation calculated by MD simulation strongly suggests that the origin of the dielectric loss peak in the gigahertz region (γ relaxation part) for protein solutions and sugars solutions is a collected rotational relaxation of the single water molecule. The relaxation time of water rotation is also equivalent to that of the reorientation of the exerted electric field for a water molecule in sugars solutions, which also consolidates the water dipole alignment along exerted electric field proposed in **Sec. 4.3**. In addition to the extension of the findings on protein solutions to sugars solutions, the following conclusions are obtained:

(1) In sugar solutions, only PRW is observed which indicates that all water molecules move in and out of the sugar HLs. Unlike water in protein solutions, the total water residence time in sugar solutions shows a Gaussian norm distribution, which suggests the random walk of water with the sugar HL, i.e., within the solution.

(2) Owing to the conclusion (1), the total water residence time in sugar solutions is identical to the ratio of sugar hydration volume to the total solvent accessible area. Due to the smaller specific surface area, water in monosaccharide solutions shows a longer residence time compared with that in disaccharide solutions, on average.

(3) The water relaxation time increases exponentially with the total residence time. Water relaxation time in monosaccharide solutions increases slower than that in disaccharide solutions. These decuple-retarding periods of relaxation time in all sugars solutions are much slower than that in protein solutions.

(4) A larger dipole moment of solute enhances the increasing rate of water relaxation time to the water residence time.

(5) The relaxation time in sugar solutions is about 1~1.5 ps longer than that in protein solutions due to the longer residence time, although the increasing rate of the relaxation time versus its residence time in protein solutions is much faster than that in sugar solutions at the same solute mass percent. Increasing the ratio of the HLs volume (i.e., the specific surface area of solute) is key to increasing the residence time, thus slowing down the rotational relaxation time. Comparison between disaccharide and monosaccharide also suggests the increment in water residence time could counteract the effect of a slower relaxation time rise rate.

(6) An effective solute for a longer rotational relaxation time requires a larger specific surface area of solute with a larger dipole moment.

Chapter 6 Conclusions and Outlook

6.1 Main conclusions

The dynamics of surrounding water are essential to the stability of protein pharmaceuticals. The research conducted in this dissertation mainly by molecular dynamics (MD) simulation contributes to the understanding of water rotational dynamics in protein solutions, as well as the designing of the protective agents in terms of their specific surface area and dipole moment.

Firstly, the dielectric spectra for protein solutions and sugar solutions are obtained by the probability of water relaxation time from MD simulation. Together with the dielectric spectroscopy experiments, it is discovered that the dielectric relaxation time obtained from this distribution is in good agreement with the measured γ relaxation time, which suggests that the rotational self-correlation of water molecules underlies the gigahertz domain of the measured dielectric spectra. This original finding associates the molecular water behavior with the macroscopic dielectric measurements.

Secondly, two basic questions of water rotational dynamics in protein solutions are addressed by MD simulation of lysozyme solutions: 1) how far proteins can affect water rotational dynamics, i.e., the thickness of protein HL; and 2) to what extent the characteristics of hydration water differ from those of the bulk water. Water rotational relaxation time ranges from several picoseconds to tens of nanoseconds in protein solutions. Depending on the time it stays within the HL(the residence time), water can be classified into 3 kinds: highly mobile water (HMW), partially restricted water (PRW), and completely restricted water (CRW). In dilute solutions, water relaxation time exhibits a Gaussian probability distribution. In concentrated solutions, the distribution is distorted to a longer time constant due to the increasing proportion of PRW and CRW. A stochastic analysis using the relaxation times and trajectories of every single water molecule is proposed and the two-dimensional probability distribution of water at a distance from the lysozyme surface with its rotational relaxation time is obtained. It is found that water rotational dynamics accelerate with the farther distance from the protein surface. Regardless of protein concentration, water rotational relaxation time versus the distance from the lysozyme surface reveals that the water rotational relaxation is severely retarded within 3 Å from the lysozyme surface and is nearly

comparable to pure water when farther than 10 Å. The thickness of the first HL is subsequently identified as this 3 Å thick water layer from the protein surface, in view of the remarkable slowdown wherein. The presented stochastic analysis of water rotational relaxation time is an original viewpoint compared with the widely accepted multiple HLs model through the Voronoi tessellation method. This method circumvents the balance of spatial resolution and calculation accuracy of relaxation time, and provides a precise HL thickness in great agreement with the SASX experiments and other MD simulations.

Thirdly, the driving mechanism behind the single water molecular rotational dynamics is investigated from the perspective of electrostatic interactions, by MD simulation. By analyzing the reorientation of the exerted electric field on a water molecule, I first find that the relaxation time of water rotation is equivalent to that of the reorientation of the exerted electric field for each water molecule, regardless of its trajectory. Namely, water rotation synchronizes with the experienced electric field reorientation. Furthermore, bringing the MD's superiority into full play, the reorientation process of the local field (electric field at fixed positions) is explored. The corresponding relaxation time is shown to range from a few hundred femtoseconds to a little longer time than one hundred picoseconds. The observed results show that the severe retardation of water rotation within the HL is probably attributed to the slow reorientation of the local electric field. For the very few CRW molecules which always stay in the HL, it is almost merely contributed from the protein's field. On the other hand, outside the HL, the rotational relaxation is predominantly determined by the electric field of the surrounding water which randomly reorientates akin to the case in pure water. This observation suggests the importance of the time length of water residing in the HL, and it is found that water rotation relaxation time is approximately exponentially dependent on its total residence time.

Finally, water dynamics in aqueous PAs solutions (sugars and sugar alcohol here) are discussed, by showing the probability distributions of both water residence time and water relaxation time through MD simulations. In PAs solutions studied here, water is just mildly retarded with the relaxation time spanning from several picoseconds to merely a few tens of picoseconds. Water residence time exhibits a gaussian distribution for various concentrations of sugar solutions and sugar alcohol solutions. This finding suggests the random walk of water within the sugar HL, i.e., within the solution. These

two observations are discrepant from the case of protein solutions. Furthermore, it is found that in all these sugar solutions (monosaccharides and disaccharides) and sugar alcohol solutions, water rotational relaxation time also increases exponentially with the increase of its residence time. Based on this, a decuple-retarding period which specifies the prerequisite residence time to obtain tenfold retardation of the relaxation time is defined. The larger the solute dipole moment, the smaller the decuple-retarding period, therefore the faster the increment rate of relaxation time versus residence time. It is also found that the larger ratio of the hydration volume (total volume of HL) to the overall water accessible space results in a longer water residence time, thus a longer relaxation time distribution. This result indicates that the solute size is one of the key factors affecting water dynamics. When water residence time distribution is similar, the solute dipole begins to manifest. The above findings suggest the effect of balance between PA dipole and size on water dynamics, which paves the way to controlling water rotational dynamics, thus to the design of bioprotective agents.

6.2 Prospect

This dissertation mostly focuses on the rotational dynamics of the individual water molecule in proteins solutions and PAs solutions, and successfully obtains several basic but useful findings about the factors dominating water rotational dynamics. Therefore, this research provides the fundamental stone for controlling water dynamics and designing PAs. Some continuing studies could pay attention to:

(1) The reorientation process of the local electric field and the electrostatic energy of water near various protein side chains. These 2 factors will provide a better understanding of the site-specific water dynamics surrounding proteins, which may account for the existence of CRW, therefore the residence time span (0~1), in protein solutions. The heterogeneous hydration dynamics may also be explained with such a viewpoint.

(2) The relationship between the relaxation time of local electric field reorientation and the commensurate relaxation time measured by TDFSS experiments.

(3) The relationship between water rotational relaxation time and residence time, for aqueous solutions of various additives with size and dipole from small to large, such as polyol, zwitterionic molecules (TMAO), amino acids, and peptides. The sizes of such solutes are between that of carbohydrates and

lysozyme studied in this dissertation. A whole picture of water relaxation time and residence time in solutions of additives from small to large is expected to facilitate the design of bioprotective agents.

Appendix

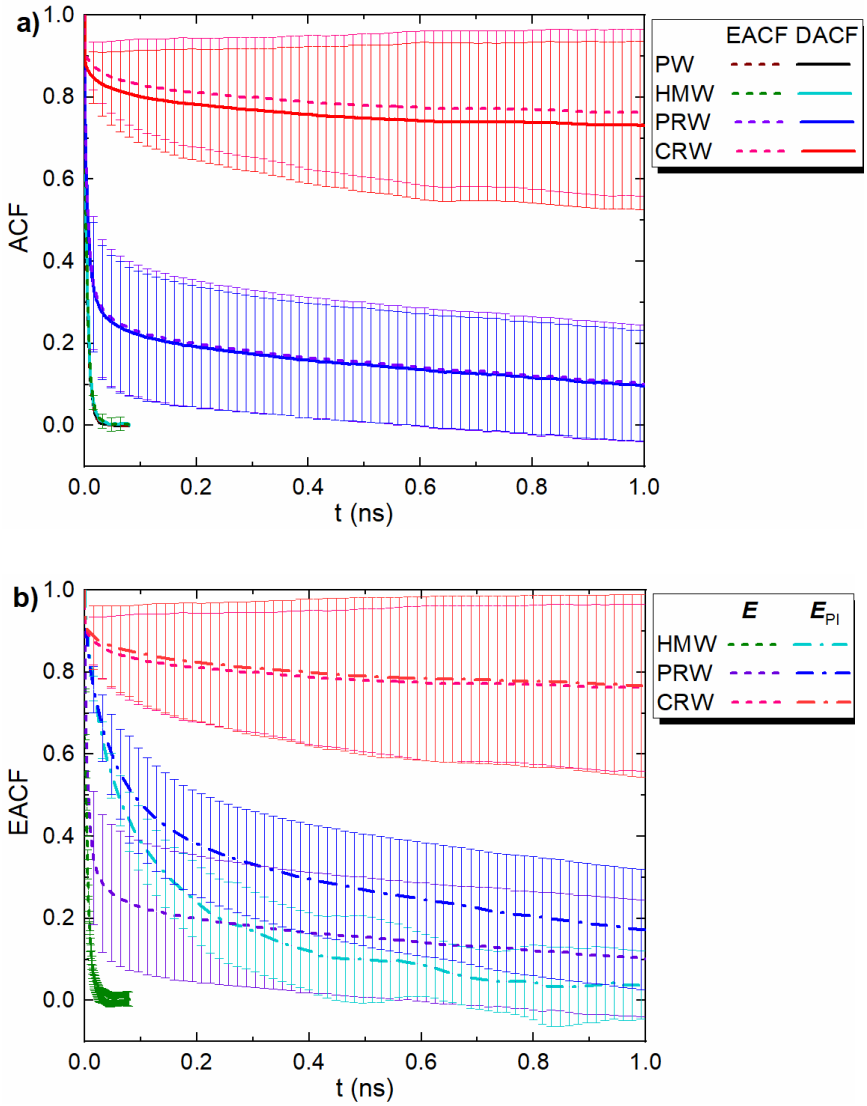

Figure A1. **(a)** Average (time) autocorrelation function (ACF) of the exerted electric field reorientation (EACF) and dipole moment rotation (DACF) for highly mobile water (HMW), partially restricted water (PRW), and tightly bound water (CRW) in lysozyme solution of 234 mg/mL. The dashed line indicates the EACF, and the solid line indicates the DACF. Results of pure water (PW) were also plotted here for comparison, which are overlapped with the results of HMW. Results shown here are the same as those in **Figure 4-3 (b) in the main text**, except that the stand deviation of each ACF is shown here. Since EACF and DACF have identical deviations for PW and HMW, only the standard deviation of EACF for HMW is plotted here (the green one). Due to the heterogeneous dynamics of the restricted water (PRW and CRW), the corresponding DACF and EACF show a larger deviation compared with that of PW of HMW. However, for each of PRW and CRW, the deviation of DACF and EACF is also identical at the same *t*.

(b) Average EACF obtained by using the overall electric field (***E***) and the electric field originating solely from proteins and ions ($\boldsymbol{E}_{\mathrm{PI}}$), for the above 3 classes of water in a lysozyme solution of 234 mg/mL. The dashed lines indicate the EACF using the overall electric field (with the part from water) while the dashed-dotted ones indicate that only uses the electric field from proteins and ions (without the part from water). Results shown here are the same as those in **Figure 4-4 in the main text**, except that the stand deviation of each EACF is shown here.

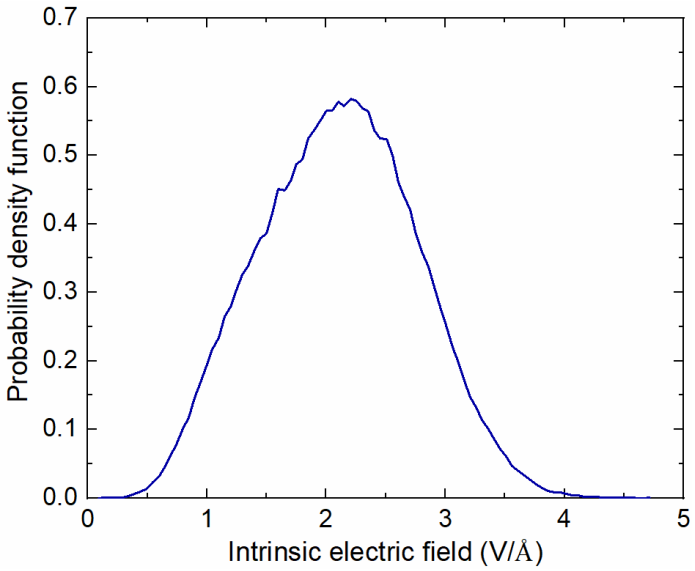

Figure A2. Intrinsic electric-field probability distribution in the 33.2 mg/mL lysozyme solutions. The internal electric field within our simulated protein-water system is around 2 V $Å^{-1}$.

TABLE S1. Multi-exponential fitting parameters of the averaged DACF and FACF for 3 classes of water (**Figure 4-3 (b)** and **Figure 4-4**), in ps unit. τ_{int} depicts the averaged relaxation from **(4-6)**. The superscript '^{PI}' indicates the EACF using the electric field only from proteins and ions. τ_i is the fitted relaxation time, and A_i the corresponding strength. In dielectric spectroscopy studies, the dielectric relaxation time of pure water is 8.27 ps,[73, 158] while the rotational relaxation time of pure water from MD is 4.87 ps. When comparing the relaxation time of protein tumbling with the γ relaxation time which is ascribable to protein tumbling, the results shown here should be normalized by the value of pure water. In this sense, the relaxation time of DACF and FACF for CRW is around 22 ns.

Subset	ACF	A_1	τ_1	A_2	τ_2	A_3	τ_3	τ_{int}	R^2
CRW	DACF	0.78	16179.87	0.09	103.67	--	--	12536.40	0.97
	EACF	0.81	16729.12	0.09	98.02	--	--	13495.02	0.98
	$EACF^{PI}$	0.83	16402.05	0.08	124.84	--	--	13598.53	0.98
PRW	DACF	0.22	1198.98	0.14	38.49	0.57	4.65	277.04	1.00
	EACF	0.23	1224.25	0.140	40.08	0.57	4.77	292.94	1.00
	$EACF^{PI}$	0.43	1100.65	0.34	76.94	0.18	14.58	497.11	1.00
HMW	DACF	0.80	5.92	0.20	0.68	--	--	4.87	1.00
	EACF	0.80	5.99	0.20	0.74	--	--	4.92	1.00
	$EACF^{PI}$	0.30	431.32	0.45	89.68	0.22	17.87	174.58	1.00

References

[1] Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nature Reviews Drug Discovery* **2008,** *7*, 21-39.

[2] Weng, L.; Stott, S. L.; Toner, M., Exploring Dynamics and Structure of Biomolecules, Cryoprotectants, and Water Using Molecular Dynamics Simulations: Implications for Biostabilization and Biopreservation. *Annu. Rev. Biomed. Eng.* **2019,** *21*, 1-31.

[3] Cordone, L.; Cupane, A.; Emanuele, A.; Giuffrida, S.; Cottone, G.; Levantino, M. J. C. O. C., Proteins in saccharides matrices and the trehalose peculiarity: Biochemical and biophysical properties. *Curr. Org. Chem.* **2015,** *19*, 1684-1706.

[4] Comez, L.; Paolantoni, M.; Sassi, P.; Corezzi, S.; Morresi, A.; Fioretto, D., Molecular properties of aqueous solutions: a focus on the collective dynamics of hydration water. *Soft Matter* **2016,** *12*, 5501-5514.

[5] Gallo, P.; Bachler, J.; Bove, L. E.; Böhmer, R.; Camisasca, G.; Coronas, L. E.; Corti, H. R.; de Almeida Ribeiro, I.; de Koning, M.; Franzese, G.; Fuentes-Landete, V.; Gainaru, C.; Loerting, T.; de Oca, J. M. M.; Poole, P. H.; Rovere, M.; Sciortino, F.; Tonauer, C. M.; Appignanesi, G. A., Advances in the study of supercooled water. *Eur. Phys. J. E* **2021,** *44*, 143.

[6] Green, J. L.; Angell, C. A., Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly. *J. Phys. Chem.* **1989,** *93*, 2880-2882.

[7] Crowe, J. H.; Crowe, L. M.; Chapman, D., Preservation of Membranes in Anhydrobiotic Organisms: The Role of Trehalose. *Science* **1984,** *223*, 701-703.

[8] Crowe, J. H.; Carpenter, J. F.; Crowe, L. M., THE ROLE OF VITRIFICATION IN ANHYDROBIOSIS. *Annu. Rev. Physiol.* **1998,** *60*, 73-103.

[9] Carpenter, J. F.; Crowe, J. H., An infrared spectroscopic study of the interactions of carbohydrates with dried proteins. *Biochemistry* **1989,** *28*, 3916-3922.

[10] Belton, P. S.; Gil, A. M., IR and Raman spectroscopic studies of the interaction of trehalose with hen egg white lysozyme. *Biopolymers* **1994,** *34*, 957-961.

[11] Arsiccio, A.; Pisano, R., Water entrapment and structure ordering as protection mechanisms for protein structural preservation. *J. Chem. Phys.* **2018,** *148*, 055102.

[12] Xie, G.; Timasheff, S. N., The thermodynamic mechanism of protein stabilization by trehalose. *Biophys. Chem.* **1997,** *64*, 25-43.

[13] Timasheff, S. N., Protein Hydration, Thermodynamic Binding, and Preferential Hydration. *Biochemistry* **2002,** *41*, 13473-13482.

[14] Lerbret, A.; Bordat, P.; Affouard, F.; Hédoux, A.; Guinet, Y.; Descamps, M., How Do Trehalose, Maltose, and Sucrose Influence Some Structural and Dynamical Properties of Lysozyme? Insight from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2007,** *111*, 9410-9420.

[15] Cottone, G., A Comparative Study of Carboxy Myoglobin in Saccharide−Water Systems by Molecular Dynamics Simulation. *J. Phys. Chem. B* **2007,** *111*, 3563-3569.

[16] Corradini, D.; Strekalova, E. G.; Stanley, H. E.; Gallo, P., Microscopic mechanism of protein cryopreservation in an aqueous solution with trehalose. *Scientific Reports* **2013,** *3*, 1218.

[17] Cottone, G.; Cordone, L.; Ciccotti, G., Molecular Dynamics Simulation of Carboxy-Myoglobin Embedded in a Trehalose-Water Matrix. *Biophys. J.* **2001,** *80*, 931-938.

[18] Cottone, G.; Ciccotti, G.; Cordone, L., Protein–trehalose–water structures in trehalose coated carboxy-myoglobin. *J. Chem. Phys.* **2002,** *117*, 9862-9866.

[19] Cottone, G.; Giuffrida, S.; Ciccotti, G.; Cordone, L., Molecular dynamics simulation of sucrose- and trehalose-coated carboxy-myoglobin. *Proteins: Structure, Function, and Bioinformatics* **2005,** *59*, 291-302.

[20] Cordone, L.; Cottone, G.; Giuffrida, S., Role of residual water hydrogen bonding in sugar/water/biomolecule systems: a possible explanation fortrehalose peculiarity. *J. Phys.: Condens. Matter* **2007,** *19*, 205110.

[21] Lins, R. D.; Pereira, C. S.; Hünenberger, P. H., Trehalose–protein interaction in aqueous solution. *Proteins: Structure, Function, and Bioinformatics* **2004,** *55*, 177-186.

[22] Paolantoni, M.; Comez, L.; Gallina, M. E.; Sassi, P.; Scarponi, F.; Fioretto, D.; Morresi, A., Light Scattering Spectra of Water in Trehalose Aqueous Solutions: Evidence for Two Different Solvent Relaxation Processes. *J. Phys. Chem. B* **2009,** *113*, 7874-7878.

[23] Lupi, L.; Comez, L.; Paolantoni, M.; Fioretto, D.; Ladanyi, B. M., Dynamics of Biological Water: Insights from Molecular Modeling of Light Scattering in Aqueous Trehalose Solutions. *J. Phys. Chem. B* **2012,** *116*, 7499-7508.

[24] Camisasca, G.; Marzio, M. D.; Corradini, D.; Gallo, P., Two structural relaxations in protein hydration water and their dynamic crossovers. *J. Chem. Phys.* **2016,** *145*, 044503.

[25] Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E., Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016,** *116*, 7673-7697.

[26] Laage, D.; Elsaesser, T.; Hynes, J. T., Water Dynamics in the Hydration Shells of Biomolecules. *Chem. Rev.* **2017,** *117*, 10694-10725.

[27] Fenimore, P. W.; Frauenfelder, H.; McMahon, B. H.; Young, R. D., Bulk-solvent and hydration-shell fluctuations, similar to α- and β-fluctuations in glasses, control protein motions and functions. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 14408-14413.

[28] Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H., Protein folding is slaved to solvent motions. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103*, 15469-15472.

[29] Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D., A unified model of protein dynamics. *Proceedings of the National Academy of Sciences* **2009,** *106*, 5129-5134.

[30] Horwich Arthur, L.; Weber-Ban Eilika, U.; Finley, D., Chaperone rings in protein folding and degradation. *Proceedings of the National Academy of Sciences* **1999,** *96*, 11033-11040.

[31] Yang, Y.; Guo, R.; Gaffney, K.; Kim, M.; Muhammednazaar, S.; Tian, W.; Wang, B.; Liang, J.; Hong, H., Folding-Degradation Relationship of a Membrane Protein

Mediated by the Universally Conserved ATP-Dependent Protease FtsH. *J. Am. Chem. Soc.* **2018,** *140*, 4656-4665.

[32] Mukherjee, S.; Mondal, S.; Bagchi, B., Mechanism of Solvent Control of Protein Dynamics. *Phys. Rev. Lett.* **2019,** *122*, 058101.

[33] Wei, L.; Shirakashi, R., Effect of Relaxation Times in a Preservative Solution on Protein Deterioration Rate. *J. Phys. Chem. B* **2020,** *124*, 8741-8749.

[34] Caliskan, G.; Kisliuk, A.; Tsai, A. M.; Soles, C. L.; Sokolov, A. P., Protein dynamics in viscous solvents. *J. Chem. Phys.* **2003,** *118*, 4230-4236.

[35] Lever, M.; Blunt, J. W.; Maclagan, R. G. A. R., Some ways of looking at compensatory kosmotropes and different water environments. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **2001,** *130*, 471-486.

[36] Saladino, G.; Marenchino, M.; Pieraccini, S.; Campos-Olivas, R.; Sironi, M.; Gervasio, F. L., A Simple Mechanism Underlying the Effect of Protecting Osmolytes on Protein Folding. *Journal of Chemical Theory and Computation* **2011,** *7*, 3846-3852.

[37] Hishida, M.; Anjum, R.; Anada, T.; Murakami, D.; Tanaka, M., Effect of Osmolytes on Water Mobility Correlates with Their Stabilizing Effect on Proteins. *J. Phys. Chem. B* **2022,** *126*, 2466-2475.

[38] Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, G., Protein hydration in solution: Experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95*, 2267-2272.

[39] Merzel, F.; Smith, J. C., Is the first hydration shell of lysozyme of higher density than bulk water? *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 5378-5383.

[40] Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M., An extended dynamical hydration shell around proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 20749.

[41] Heyden, M.; Tobias, D. J., Spatial Dependence of Protein-Water Collective Hydrogen-Bond Dynamics. *Phys. Rev. Lett.* **2013,** *111*, 218101.

[42] Perticaroli, S.; Ehlers, G.; Stanley, C. B.; Mamontov, E.; O'Neill, H.; Zhang, Q.; Cheng, X.; Myles, D. A. A.; Katsaras, J.; Nickels, J. D., Description of Hydration Water in Protein (Green Fluorescent Protein) Solution. *J. Am. Chem. Soc.* **2017,** *139*, 1098-1105.

[43] Fisette, O.; Päslack, C.; Barnes, R.; Isas, J. M.; Langen, R.; Heyden, M.; Han, S.; Schäfer, L. V., Hydration Dynamics of a Peripheral Membrane Protein. *J. Am. Chem. Soc.* **2016,** *138*, 11526-11535.

[44] Yang, J.; Wang, Y.; Wang, L.; Zhong, D., Mapping Hydration Dynamics around a β-Barrel Protein. *J. Am. Chem. Soc.* **2017,** *139*, 4399-4408.

[45] Houston, P.; Macro, N.; Kang, M.; Chen, L.; Yang, J.; Wang, L.; Wu, Z.; Zhong, D., Ultrafast Dynamics of Water–Protein Coupled Motions around the Surface of Eye Crystallin. *J. Am. Chem. Soc.* **2020,** *142*, 3997-4007.

[46] Braun, D.; Schmollngruber, M.; Steinhauser, O., Rotational dynamics of water molecules near biological surfaces with implications for nuclear quadrupole relaxation. *PCCP* **2016,** *18*, 24620-24630.

[47] Persson, F.; Söderhjelm, P.; Halle, B., The geometry of protein hydration. *J. Chem. Phys.* **2018,** *148*, 215101.

[48] Persson, F.; Söderhjelm, P.; Halle, B., How proteins modify water dynamics. *J. Chem. Phys.* **2018,** *148*, 215103.

[49] Persson, F.; Söderhjelm, P.; Halle, B., The spatial range of protein hydration. *J. Chem. Phys.* **2018,** *148*, 215104.

[50] Zhong, D.; Pal, S. K.; Zewail, A. H., Biological water: A critique. *Chem. Phys. Lett.* **2011,** *503*, 1-11.

[51] Grant, E. H., The dielectric method of investigating bound water in biological material: An appraisal of the technique. *Bioelectromagnetics* **1982,** *3*, 17-24.

[52] Grant, E. H.; McClean, V. E. R.; Nightingale, N. R. V.; Sheppard, R. J.; Chapman, M. J., Dielectric behavior of water in biological solutions: Studies on myoglobin, human low-density lipoprotein, and polyvinylpyrrolidone. *Bioelectromagnetics* **1986,** *7*, 151-162.

[53] Pethig, R., Protein-Water Interactions Determined by Dielectric Methods. *Annu. Rev. Phys. Chem.* **1992,** *43*, 177-205.

[54] Nandi, N.; Bhattacharyya, K.; Bagchi, B., Dielectric Relaxation and Solvation Dynamics of Water in Complex Chemical and Biological Systems. *Chem. Rev.* **2000,** *100*, 2013-2046.

[55] Mashimo, S.; Kuwabara, S.; Yagihara, S.; Higasi, K., Dielectric relaxation time and structure of bound water in biological materials. *J. Phys. Chem.* **1987,** *91*, 6337-6338.

[56] Dachwitz, E.; Parak, F.; Stockhausen, M., On the Dielectric Relaxation of Aqueous Myoglobin Solutions. *Berichte der Bunsengesellschaft für physikalische Chemie* **1989,** *93*, 1454-1458.

[57] Wolf, M.; Gulich, R.; Lunkenheimer, P.; Loidl, A., Relaxation dynamics of a protein solution investigated by dielectric spectroscopy. *Biochim. Biophys. Acta* **2012,** *1824*, 723-730.

[58] Wolf, M.; Emmert, S.; Gulich, R.; Lunkenheimer, P.; Loidl, A., Dynamics of protein hydration water. *Physical Review E* **2015,** *92*, 032727.

[59] Fukuzaki, M.; Miura, N.; Shinyashiki, N.; Kurita, D.; Shioya, S.; Haida, M.; Mashimo, S., Comparison of Water Relaxation Time in Serum Albumin Solution Using Nuclear Magnetic Resonance and Time Domain Reflectometry. *J. Phys. Chem.* **1995,** *99*, 431-435.

[60] Mattea, C.; Qvist, J.; Halle, B., Dynamics at the Protein-Water Interface from 17O Spin Relaxation in Deeply Supercooled Solutions. *Biophys. J.* **2008,** *95*, 2951-2963.

[61] Bryant, R. G., Dynamics of water in and around proteins characterized by 1H-spin-lattice relaxometry. *Comptes Rendus Physique* **2010,** *11*, 128-135.

[62] Denisov, V. P.; Halle, B., Protein hydration dynamics in aqueous solution. *Faraday Discuss.* **1996,** *103*, 227-244.

[63] Modig, K.; Qvist, J.; Marshall, C. B.; Davies, P. L.; Halle, B., High water mobility on the ice-binding surface of a hyperactive antifreeze protein. *PCCP* **2010,** *12*, 10189-10197.

[64] Daniel, R. M.; Finney, J. L.; Stoneham, M.; Halle, B., Protein hydration dynamics in solution: a critical survey. *Philosophical Transactions of the Royal Society B* **2004,** *359*, 1207-1224.

[65] Grebenkov, D. S.; Goddard, Y. A.; Diakova, G.; Korb, J.-P.; Bryant, R. G., Dimensionality of Diffusive Exploration at the Protein Interface in Solution. *J. Phys. Chem. B* **2009,** *113*, 13347-13356.

[66] Korb, J.-P.; Goddard, Y.; Pajski, J.; Diakova, G.; Bryant, R. G., Extreme-Values Statistics and Dynamics of Water at Protein Interfaces. *J. Phys. Chem. B* **2011,** *115*, 12845-12858.

[67] Perticaroli, S.; Comez, L.; Paolantoni, M.; Sassi, P.; Lupi, L.; Fioretto, D.; Paciaroni, A.; Morresi, A., Broadband Depolarized Light Scattering Study of Diluted Protein Aqueous Solutions. *J. Phys. Chem. B* **2010,** *114*, 8262-8269.

[68] Perticaroli, S.; Comez, L.; Sassi, P.; Paolantoni, M.; Corezzi, S.; Caponi, S.; Morresi, A.; Fioretto, D., Hydration and aggregation of lysozyme by extended frequency range depolarized light scattering. *J. Non-Cryst. Solids* **2015,** *407*, 472-477.

[69] Halle, B.; Nilsson, L., Does the Dynamic Stokes Shift Report on Slow Protein Hydration Dynamics? *J. Phys. Chem. B* **2009,** *113*, 8210-8213.

[70] Pal, S. K.; Peon, J.; Zewail, A. H., Biological water at the protein surface: Dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 1763-1768.

[71] Oleinikova, A.; Sasisanker, P.; Weingärtner, H., What Can Really Be Learned from Dielectric Spectroscopy of Protein Solutions? A Case Study of Ribonuclease A. *J. Phys. Chem. B* **2004,** *108*, 8467-8474.

[72] Mijović, J.; Bian, Y.; Gross, R. A.; Chen, B., Dynamics of Proteins in Hydrated State and in Solution As Studied by Dielectric Relaxation Spectroscopy. *Macromolecules* **2005,** *38*, 10812-10819.

[73] Cametti, C.; Marchetti, S.; Gambi, C. M. C.; Onori, G., Dielectric Relaxation Spectroscopy of Lysozyme Aqueous Solutions: Analysis of the δ-Dispersion and the Contribution of the Hydration Water. *J. Phys. Chem. B* **2011,** *115*, 7144-7153.

[74] Yamamoto, N.; Ohta, K.; Tamura, A.; Tominaga, K., Broadband Dielectric Spectroscopy on Lysozyme in the Sub-Gigahertz to Terahertz Frequency Regions: Effects of Hydration and Thermal Excitation. *J. Phys. Chem. B* **2016,** *120*, 4743-4755.

[75] Nakanishi, M.; Sokolov, A. P., Protein dynamics in a broad frequency range: Dielectric spectroscopy studies. *J. Non-Cryst. Solids* **2015,** *407*, 478-485.

[76] Lerbret, A.; Affouard, F.; Bordat, P.; Hédoux, A.; Guinet, Y.; Descamps, M., Molecular dynamics simulations of lysozyme in water/sugar solutions. *Chem. Phys.* **2008,** *345*, 267-274.

[77] Lerbret, A.; Affouard, F.; Hédoux, A.; Krenzlin, S.; Siepmann, J.; Bellissent-Funel, M.-C.; Descamps, M., How Strongly Does Trehalose Interact with Lysozyme in the Solid State? Insights from Molecular Dynamics Simulation and Inelastic Neutron Scattering. *J. Phys. Chem. B* **2012,** *116*, 11103-11116.

[78] GhattyVenkataKrishna, P. K.; Carri, G. A., The effect of complex solvents on the structure and dynamics of protein solutions: The case of Lysozyme in trehalose/water mixtures. *Eur. Phys. J. E* **2013,** *36*, 14.

[79] Camisasca, G.; Marzio, M. D.; Corradini, D.; Gallo, P., Two structural relaxations in protein hydration water and their dynamic crossovers. **2016,** *145*, 044503.

[80] Camisasca, G.; Iorio, A.; De Marzio, M.; Gallo, P., Structure and slow dynamics of protein hydration water. *J. Mol. Liq.* **2018,** *268*, 903-910.

[81] Martin, D. R.; Forsmo, J. E.; Matyushov, D. V., Complex Dynamics of Water in Protein Confinement. *J. Phys. Chem. B* **2018,** *122*, 3418-3425.

[82] Sterpone, F.; Ceccarelli, M.; Marchi, M., Dynamics of hydration in hen egg white lysozyme. *J. Mol. Biol.* **2001,** *311*, 409-419.

[83] Marchi, M.; Sterpone, F.; Ceccarelli, M., Water Rotational Relaxation and Diffusion in Hydrated Lysozyme. *J. Am. Chem. Soc.* **2002,** *124*, 6787-6791.

[84] Bernini, A.; Spiga, O.; Ciutti, A.; Chiellini, S.; Menciassi, N.; Venditti, V.; Niccolai, N., On the dynamics of water molecules at the protein solute interfaces. *Homeopathy* **2004,** *93*, 199-202.

[85] Kyritsis, A.; Panagopoulou, A.; Pissis, P.; Serra, R. S. I.; Ribelles, J. L. G.; Shinyashiki, N., Water and protein dynamics in protein-water mixtures over wide range of composition. *IEEE Trans. Dielectr. Electr. Insul.* **2012,** *19*, 1239-1246.

[86] Panuszko, A.; Wojciechowski, M.; Bruździak, P.; Rakowska, P. W.; Stangret, J., Characteristics of hydration water around hen egg lysozyme as the protein model in aqueous solution. FTIR spectroscopy and molecular dynamics simulation. *PCCP* **2012,** *14*, 15765-15773.

[87] Ngai, K. L.; Capaccioli, S.; Paciaroni, A., Dynamics of hydrated proteins and bio-protectants: Caged dynamics, β-relaxation, and α-relaxation. *Biochim. Biophys. Acta, Gen. Subj* **2017,** *1861*, 3553-3563.

[88] Sterpone, F.; Stirnemann, G.; Laage, D., Magnitude and Molecular Origin of Water Slowdown Next to a Protein. *J. Am. Chem. Soc.* **2012,** *134*, 4116-4119.

[89] Fogarty, A. C.; Laage, D., Water Dynamics in Protein Hydration Shells: The Molecular Origins of the Dynamical Perturbation. *J. Phys. Chem. B* **2014,** *118*, 7715-7729.

[90] Servantie, J.; Atilgan, C.; Atilgan, A. R., Depth dependent dynamics in the hydration shell of a protein. *J. Chem. Phys.* **2010,** *133*, 085101.

[91] Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M., An extended dynamical hydration shell around proteins. **2007,** *104*, 20749-20752.

[92] Oleinikova, A.; Smolin, N.; Brovchenko, I., Influence of Water Clustering on the Dynamics of Hydration Water at the Surface of a Lysozyme. *Biophys. J.* **2007,** *93*, 2986-3000.

[93] Tenuzzo, L.; Camisasca, G.; Gallo, P., Protein-Water and Water-Water Long-Time Relaxations in Protein Hydration Water upon Cooling—A Close Look through Density Correlation Functions. **2020,** *25*, 4570.

[94] Brotzakis, Z. F.; Groot, C. C. M.; Brandeburgo, W. H.; Bakker, H. J.; Bolhuis, P. G., Dynamics of Hydration Water around Native and Misfolded α-Lactalbumin. *J. Phys. Chem. B* **2016,** *120*, 4756-4766.

[95] Mukherjee, S.; Mondal, S.; Bagchi, B., Distinguishing dynamical features of water inside protein hydration layer: Distribution reveals what is hidden behind the average. *J. Chem. Phys.* **2017,** *147*, 024901.

[96] Mondal, S.; Mukherjee, S.; Bagchi, B., Protein Hydration Dynamics: Much Ado about Nothing? *The Journal of Physical Chemistry Letters* **2017,** *8*, 4878-4882.

[97] Charkhesht, A.; Regmi, C. K.; Mitchell-Koch, K. R.; Cheng, S.; Vinh, N. Q., High-Precision Megahertz-to-Terahertz Dielectric Spectroscopy of Protein Collective Motions and Hydration Dynamics. *J. Phys. Chem. B* **2018,** *122*, 6341-6350.

[98] Laage, D.; Stirnemann, G.; Sterpone, F.; Rey, R.; Hynes, J. T., Reorientation and Allied Dynamics in Water and Aqueous Solutions. *Annu. Rev. Phys. Chem.* **2011,** *62*, 395-416.

[99] Laage, D.; Hynes James, T., A Molecular Jump Mechanism of Water Reorientation. *Science* **2006,** *311*, 832-835.

[100] Pattni, V.; Vasilevskaya, T.; Thiel, W.; Heyden, M., Distinct Protein Hydration Water Species Defined by Spatially Resolved Spectra of Intermolecular Vibrations. *J. Phys. Chem. B* **2017,** *121*, 7431-7442.

[101] Persson, E.; Halle, B., Nanosecond to Microsecond Protein Dynamics Probed by Magnetic Relaxation Dispersion of Buried Water Molecules. *J. Am. Chem. Soc.* **2008,** *130*, 1774-1787.

[102] Zhang, L.; Wang, L.; Kao, Y.-T.; Qiu, W.; Yang, Y.; Okobiah, O.; Zhong, D., Mapping hydration dynamics around a protein surface. *Proceedings of the National Academy of Sciences* **2007,** *104*, 18461-18466.

[103] Zhang, L.; Yang, Y.; Kao, Y.-T.; Wang, L.; Zhong, D., Protein Hydration Dynamics and Molecular Mechanism of Coupled Water−Protein Fluctuations. *J. Am. Chem. Soc.* **2009,** *131*, 10677-10691.

[104] Qin, Y.; Wang, L.; Zhong, D., Dynamics and mechanism of ultrafast water–protein interactions. *Proceedings of the National Academy of Sciences* **2016,** *113*, 8424-8429.

[105] Piskulich, Z. A.; Laage, D.; Thompson, W. H., Activation energies and the extended jump model: How temperature affects reorientation and hydrogen-bond exchange dynamics in water. *J. Chem. Phys.* **2020,** *153*, 074110.

[106] Piskulich, Z. A.; Laage, D.; Thompson, W. H., Using Activation Energies to Elucidate Mechanisms of Water Dynamics. *J. Phys. Chem. A* **2021**.

[107] Kremer, F.; Schönhals, A., *Broadband Dielectric Spectroscopy*. Springer Berlin Heidelberg: 2012.

[108] Evans, M. W.; Lie, G. C.; Clementi, E., Molecular dynamics of water in a strong electric field: Electric-field-induced translational anisotropy. *Physical Review A* **1987,** *36*, 226-231.

[109] Baer, A.; Miličević, Z.; Smith, D. M.; Smith, A.-S., Water in an electric field does not dance alone: The relation between equilibrium structure, time dependent viscosity and molecular motions. *J. Mol. Liq.* **2019,** *282*, 303-315.

[110] Chong, S.-H.; Ham, S., Interaction with the Surrounding Water Plays a Key Role in Determining the Aggregation Propensity of Proteins. *Angew. Chem. Int. Ed.* **2014,** *53*, 3961-3964.

[111] Neumayr, G.; Rudas, T.; Steinhauser, O., Global and local Voronoi analysis of solvation shells of proteins. *J. Chem. Phys.* **2010,** *133*, 084108.

[112] Braun, D.; Schmollngruber, M.; Steinhauser, O., Revival of the Intermolecular Nuclear Overhauser Effect for Mapping Local Protein Hydration Dynamics. *The Journal of Physical Chemistry Letters* **2017,** *8*, 3421-3426.

[113] Konstantinovsky, D.; Perets, E. A.; Santiago, T.; Velarde, L.; Hammes-Schiffer, S.; Yan, E. C. Y., Detecting the First Hydration Shell Structure around Biomolecules at Interfaces. *ACS Central Science* **2022,** *8*, 1404-1414.

[114] Shukla, N.; Pomarico, E.; Chen, L.; Chergui, M.; Othon, C. M., Retardation of Bulk Water Dynamics by Disaccharide Osmolytes. *J. Phys. Chem. B* **2016,** *120*, 9477-9483.

[115] Groot, C. C. M.; Bakker, H. J., A femtosecond mid-infrared study of the dynamics of water in aqueous sugar solutions. *PCCP* **2015,** *17*, 8449-8458.

[116] Shiraga, K.; Suzuki, T.; Kondo, N.; Tajima, T.; Nakamura, M.; Togo, H.; Hirata, A.; Ajito, K.; Ogawa, Y., Broadband dielectric spectroscopy of glucose aqueous solution: Analysis of the hydration state and the hydrogen bond network. *J. Chem. Phys.* **2015,** *142*, 234504.

[117] Shiraga, K.; Adachi, A.; Nakamura, M.; Tajima, T.; Ajito, K.; Ogawa, Y., Characterization of the hydrogen-bond network of water around sucrose and trehalose: Microwave and terahertz spectroscopic study. *J. Chem. Phys.* **2017,** *146*, 105102.

[118] Roy, A.; Dutta, R.; Kundu, N.; Banik, D.; Sarkar, N., A Comparative Study of the Influence of Sugars Sucrose, Trehalose, and Maltose on the Hydration and Diffusion of DMPC Lipid Bilayer at Complete Hydration: Investigation of Structural and Spectroscopic Aspect of Lipid–Sugar Interaction. *Langmuir* **2016,** *32*, 5124-5134.

[119] Hirai, M.; Ajito, S.; Iwasa, T.; Wen, D.; Igarashi, N.; Shimizu, N., Short-Distance Intermolecular Correlations of Mono- and Disaccharides in Condensed Solutions: Bulky Character of Trehalose. *ACS Omega* **2020,** *5*, 10815-10825.

[120] Indra, S.; Biswas, R., How Heterogeneous Are Trehalose/Glycerol Cryoprotectant Mixtures? A Combined Time-Resolved Fluorescence and Computer Simulation Investigation. *J. Phys. Chem. B* **2016,** *120*, 11214-11228.

[121] Magno, A.; Gallo, P., Understanding the Mechanisms of Bioprotection: A Comparative Study of Aqueous Solutions of Trehalose and Maltose upon Supercooling. *The Journal of Physical Chemistry Letters* **2011,** *2*, 977-982.

[122] Heid, E.; Honegger, P.; Braun, D.; Szabadi, A.; Stankovic, T.; Steinhauser, O.; Schröder, C., Computational spectroscopy of trehalose, sucrose, maltose, and glucose: A comprehensive study of TDSS, NQR, NOE, and DRS. *J. Chem. Phys.* **2019,** *150*, 175102.

[123] Lupi, L.; Comez, L.; Paolantoni, M.; Perticaroli, S.; Sassi, P.; Morresi, A.; Ladanyi, B. M.; Fioretto, D., Hydration and Aggregation in Mono- and Disaccharide Aqueous Solutions by Gigahertz-to-Terahertz Light Scattering and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2012,** *116*, 14760-14767.

[124] Winther, L. R.; Qvist, J.; Halle, B., Hydration and Mobility of Trehalose in Aqueous Solution. *J. Phys. Chem. B* **2012,** *116*, 9196-9207.

[125] Fioretto, D.; Comez, L.; Corezzi, S.; Paolantoni, M.; Sassi, P.; Morresi, A., Solvent Sharing Models for Non-Interacting Solute Molecules: The Case of Glucose and Trehalose Water Solutions. *Food Biophysics* **2013,** *8*, 177-182.

[126] Ekdawi-Sever, N.; de Pablo, J. J.; Feick, E.; von Meerwall, E., Diffusion of Sucrose and α,α-Trehalose in Aqueous Solutions. *J. Phys. Chem. A* **2003,** *107*, 936-943.

[127] Cametti, C.; Marchetti, S.; Onori, G., Lysozyme Hydration in Concentrated Aqueous Solutions. Effect of an Equilibrium Cluster Phase. *J. Phys. Chem. B* **2013,** *117*, 104-110.

[128] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The missing term in effective pair potentials. *J. Phys. Chem.* **1987,** *91*, 6269-6271.

[129] Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **2008,** *4*, 435-447.

[130] Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988,** *110*, 1657-1666.

[131] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18*, 1463-1472.

[132] Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126*, 014101.

[133] Parrinello, M.; Rahman, A., Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980,** *45*, 1196-1199.

[134] Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993,** *98*, 10089-10092.

[135] Pollock, E. L.; Glosli, J., Comments on P3M, FMM, and the Ewald method for large periodic Coulombic systems. *Comput. Phys. Commun.* **1996,** *95*, 93-110.

[136] Raicu, V.; Feldman, Y., *Dielectric Relaxation in Biological Systems: Physical Principles, Methods, and Applications*. Oxford University Press: 2015.

[137] Gregory, A.; Clarke, R., *Tables of the complex permittivity of dielectric reference liquids at frequencies up to 5 GHz*. National Physical Laboratory Report, 2012.

[138] Axelrod, N.; Axelrod, E.; Gutina, A.; Puzenko, A.; Ishai, P. B.; Feldman, Y., Dielectric spectroscopy data treatment: I. Frequency domain. *Meas. Sci. Technol.* **2004,** *15*, 755-764.

[139] Lucas, J.; Géron, E.; Ditchi, T.; Holé, S., A fast Fourier transform implementation of the Kramers-Kronig relations: Application to anomalous and left handed propagation. *AIP Advances* **2012,** *2*, 032144.

[140] Shirakashi, R.; Amano, Y.; Yamada, J., Measurement of the Water Relaxation Time of $${\upvarepsilon }$$-Polylysine Aqueous Solutions. *Int. J. Thermophys.* **2017,** *38*, 75.

[141] Castro, F.; Nabet, B., Numerical computation of the complex dielectric permittivity using Hilbert transform and FFT techniques. *Journal of the Franklin Institute* **1999,** *336*, 53-64.

[142] Lucarini, V.; Saarinen, J.; Peiponen, K.; Vartiainen, E., Kramers-Kronig Relations in Optical Materials Research. *Kramers-Kronig Relations in Optical Materials Research, by V. Lucarini, J. Saarinen, K. Peiponen, and E. Vartiainen. X, 162 p. 37 illus. 3-540-23673-2. Berlin: Springer, 2005.* **2005,** *110*.

[143] Matsuura, H.; Shirakashi, R., Exclusion of DC conductivity effect from dielectric loss spectrum using Kramers-Kronig relations for evaluation of slow dynamics of water molecules. *Japanese Journal of Applied Physics* **2022**.

[144] Schwarz, G., Estimating the Dimension of a Model. *The Annals of Statistics* **1978,** *6*, 461-464.

[145] Shinotsuka, H.; Yoshikawa, H.; Murakami, R.; Nakamura, K.; Tanaka, H.; Yoshihara, K., Automated information compression of XPS spectrum using information criteria. *J. Electron. Spectrosc. Relat. Phenom.* **2019,** *239*, 146903.

[146] Ishai, P. B.; Talary, M. S.; Caduff, A.; Levy, E.; Feldman, Y., Electrode polarization in dielectric measurements: a review. *Meas. Sci. Technol.* **2013,** *24*, 102001.

[147] Braun, D.; Schmollngruber, M.; Steinhauser, O., Towards a complete characterization of the δ-dispersion in dielectric spectroscopy of protein–water systems. *PCCP* **2017,** *19*, 26980-26985.

[148] Floros, S.; Liakopoulou-Kyriakides, M.; Karatasos, K.; Papadopoulos, G., Detailed study of the dielectric function of a lysozyme solution studied with molecular dynamics simulations. *European biophysics journal : EBJ* **2015,** *44*.

[149] Fried, S. D.; Wang, L.-P.; Boxer, S. G.; Ren, P.; Pande, V. S., Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013,** *117*, 16236-16248.

[150] Sega, M.; Schröder, C., Dielectric and Terahertz Spectroscopy of Polarizable and Nonpolarizable Water Models: A Comparative Study. *J. Phys. Chem. A* **2015,** *119*, 1539-1547.

[151] Marx, D.; Hutter, J., *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press: Cambridge, 2009.

[152] Hölzl, C.; Forbert, H.; Marx, D., Dielectric relaxation of water: assessing the impact of localized modes, translational diffusion, and collective dynamics. *PCCP* **2021,** *23*, 20875-20882.

[153] Tan, P.; Huang, J.; Mamontov, E.; García Sakai, V.; Merzel, F.; Liu, Z.; Ye, Y.; Hong, L., Decoupling between the translation and rotation of water in the proximity of a protein molecule. *PCCP* **2020,** *22*, 18132-18140.

[154] Gu, W.; Schoenborn, B. P., Molecular dynamics simulation of hydration in myoglobin. **1995,** *22*, 20-26.

[155] Fried, S. D.; Bagchi, S.; Boxer, S. G., Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013,** *135*, 11181-11192.

[156] Stirnemann, G.; Wernersson, E.; Jungwirth, P.; Laage, D., Mechanisms of Acceleration and Retardation of Water Dynamics by Ions. *J. Am. Chem. Soc.* **2013,** *135*, 11824-11831.

[157] Soper, A. K.; Weckström, K., Ion solvation and water structure in potassium halide aqueous solutions. *Biophys. Chem.* **2006,** *124*, 180-191.

[158] Hu, K.; Matsuura, H.; Shirakashi, R., Stochastic Analysis of Molecular Dynamics Reveals the Rotation Dynamics Distribution of Water around Lysozyme. *J. Phys. Chem. B* **2022,** *126*, 4520-4530.

[159] Link, F. J.; Heng, J. Y. Y., Unraveling the Impact of pH on the Crystallization of Pharmaceutical Proteins: A Case Study of Human Insulin. *Crystal Growth & Design* **2022,** *22*, 3024-3033.

[160] Matyushov, D. V., War and Peace between Electrostatic and van der Walls Forces Regulate Translational and Rotational Diffusion. *J. Chem. Phys.* **2022**.

[161] Montenegro, A.; Dutta, C.; Mammetkuliev, M.; Shi, H.; Hou, B.; Bhattacharyya, D.; Zhao, B.; Cronin, S. B.; Benderskii, A. V., Asymmetric response of interfacial water to applied electric fields. *Nature* **2021,** *594*, 62-65.

[162] von Domaros, M.; Bratko, D.; Kirchner, B.; Luzar, A., Dynamics at a Janus Interface. *The Journal of Physical Chemistry C* **2013,** *117*, 4561-4567.

[163] Balasubramanian, S.; Bagchi, B., Slow Orientational Dynamics of Water Molecules at a Micellar Surface. *J. Phys. Chem. B* **2002,** *106*, 3668-3672.

[164] Wu, H.; Ghaani, M. R.; Nandi, P. K.; English, N. J., Investigation of Dipolar Response of the Hydrated Hen-Egg White Lysozyme Complex under Externally Applied Electric Fields: Insights from Non-equilibrium Molecular Dynamics. *J. Phys. Chem. B* **2022,** *126*, 858-868.

[165] Nandi, P. K.; Futera, Z.; English, N. J., Perturbation of hydration layer in solvated proteins by external electric and electromagnetic fields: Insights from non-equilibrium molecular dynamics. *J. Chem. Phys.* **2016,** *145*, 205101.

[166] Jo, S.; Kim, T.; Iyer, V. G.; Im, W., CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008,** *29*, 1859-1865.

[167] Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W., CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009,** *97*, 50-58.

[168] Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L., III; MacKerell, A. D., Jr.; Klauda, J. B.; Im, W., CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **2016,** *12*, 405-413.

[169] Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W., Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010,** *114*, 7830-7843.

[170] Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017,** *14*, 71-73.

[171] Frenkel, D.; Smit, B., Chapter 4 - Molecular Dynamics Simulations. In *Understanding Molecular Simulation (Second Edition)*, Frenkel, D.; Smit, B., Eds. Academic Press: San Diego, 2002; pp 63-107.

[172] Rampp, M.; Buttersack, C.; Lüdemann, H.-D., c,T-Dependence of the viscosity and the self-diffusion coefficients in some aqueous carbohydrate solutions. *Carbohydr. Res.* **2000,** *328*, 561-572.

[173] Zhang, J.; Matsuura, H.; Shirakashi, R., Prediction of water relaxation time using near infrared spectroscopy. *J. Food Process Eng.* **2022,** *n/a*, e14095.

[174] Heyden, M.; Bründermann, E.; Heugen, U.; Niehues, G.; Leitner, D. M.; Havenith, M., Long-Range Influence of Carbohydrates on the Solvation Dynamics of Water—Answers from Terahertz Absorption Measurements and Molecular Modeling Simulations. *J. Am. Chem. Soc.* **2008,** *130*, 5773-5779.

[175] Fulton, R. L., The effect of the spatial nonlocality of the Kirkwood g-factor on the determination of the long wavelength dielectric functions in dipolar fluids. *J. Chem. Phys.* **2012,** *136*, 084502.

[176] University of Pittsburgh, D. o. C. Sorbitol. http://pqr.pitt.edu/mol/FBPFZTCFMRRESA-JGWLITMVSA-N.

[177] Fioretto, D.; Comez, L.; Gallina, M. E.; Morresi, A.; Palmieri, L.; Paolantoni, M.; Sassi, P.; Scarponi, F., Separate dynamics of solute and solvent in water–glucose solutions by depolarized light scattering. *Chem. Phys. Lett.* **2007,** *441*, 232-236.

[178] Gallina, M. E.; Comez, L.; Morresi, A.; Paolantoni, M.; Perticaroli, S.; Sassi, P.; Fioretto, D., Rotational dynamics of trehalose in aqueous solutions studied by depolarized light scattering. *J. Chem. Phys.* **2010,** *132*, 214508.

Publication List

Peer-reviewed Journal Paper

1. **Hu, K.**; Shirakashi, R., Dynamic Electric Field Alignment Determines the Water Rotation movement. *The Journal of Physical Chemistry B*. (Accepted, DOI: 10.1021/acs.jpcb.2c07405)

2. **Hu, K.**; Matsuura, H.; Shirakashi, R., Stochastic Analysis of Molecular Dynamics Reveals the Rotation Dynamics Distribution of Water around Lysozyme. *The Journal of Physical Chemistry B*, 126(24), 4520-4530, 2022.

Conference Contributions

1. **Hu, K**.; Shirakashi, R., Molecular dynamics studies of sugar solutions for controlling water rotational relaxation time. *The 2023 Summer Biomechanics, Bioengineering and Biotransport Conference*, Colorado, USA, June, 2023. (Submitted)

2. **Hu, K**.; Shirakashi, R., Coupling between the rotational and translational dynamics of water in lysozyme solutions. *The 19th International Conference on Flow Dynamics*, Tohoku, Japan, Novermber, 2022. (Best Presentation Award for Young Researcher, Student Best Presentation Award (OS8))

3. **Hu, K.**; Shirakashi, R., Electric field alignment dominates the water rotation dynamics in protein solutions. *The 13th Asian Thermophysical Properties Conference*, Tohoku, Japan, September, 2022.

4. **Hu, K.**; Shirakashi, R., Effect of local Electric field on the rotational dynamics of water dipole in protein solutions. *The 60th Annual Meeting of the Biophysical Society of Japan*, Hakodate, Japan, September, 2022.

5 **Hu, K.**; Shirakashi, R., Relationship between the internal electric field and the water rotational relaxation time in lysozyme solution. *The 59th National Heat transfer Symposium of Japan*, Gifu, Japan, May, 2022.

6. **Hu, K.**; Matsuura, H.; Shirakashi, R., Characterizing water molecular rotational relaxation in lysozyme aqueous solution by MD Simulation. *The 58th National Heat transfer Symposium of Japan*, Online, Japan, May, 2021.

Acknowledgement

Upon the completion of this thesis, I would like to take this opportunity to give my heartfelt thanks to all those who have supported me through my academic studies, guided me in my academic career, and motivated me to pursue my goals in life.

First and foremost, I would like to express my deep gratitude to my supervisor, Professor Ryo Shirakashi, for his continued advice, support, and encouragement. Prof. Shirakashi has offered constructive guidance and invaluable advice on the completion and improvement of my doctoral dissertation. He is always patient with failures and mistakes and encourages me to analyze the causes and improve the models over and over again. His profound vision and knowledge are inspiring me during the whole research period.

Also, many thanks to my defence committee: Professor Taku Ohara, Professor Fumitoshi Sato, Professor Hirofumi Daiguji, Professor Junichiro Shiomi, and Professor Ikuya Kinefuchi, for their insightful comments and constructive suggestions.

I sincerely appreciate Dr. Hiroaki Matsuura, the research assistant for his continuous support of the broadband dielectric spectroscopy experiments and all his useful suggestions on writing academic papers and making presentations. Collaboration with Dr. Matsuura is always a pleasure. I am also grateful to Dr. Lin Wei and Junkai Zhang who help me enter the territory of the study of hydration water dynamics. I would like to thank all other labmates, Dr. Kiyoshi Takano, Mr. Mitsuhiro Uemura, Ms. Natsumi, Adachi, Mr. Daichi Tokuda, Mr. Daichi Tokuda, Mr. Nobuhiro Kaneoya, Mr. Kousei Kawai, Mr. Akio Nomura, Mr. Yuuma Nishikawa, Mr. Kishi Serizawa, and Dr. Yuanheng Zhao, for their help in my daily life in The University of Tokyo. My Ph.D. life wouldn't be this pleasant without their support and encouragement. As my tutor, Mr. Tokuda picks me up at the Narita Airport and helps me adapt to the life in Japan. Dr. Yuanheng Zhao is my old school chum when we were in college. What a coincidence! He joined the lab in the last half year of my Ph.D. course. He also offers me some great suggestions on academic publications and academic career decision-making.

I also would like to thank all my friends for always encouraging and supporting me. I sincerely appreciate Mr. Yangyang Ge who is always willing to chat about anything and cheers me up. My thanks also go to my friends Prof. Zhihao Chen, Dr.Hua Li, Dr.

Peihua Du, Dr.Shuaijie Zhao, Dr.Kunming Li, Dr.Yuanzhi Qi, Dr.Jifeng Xue, Dr.Weiyan Chen, Dr. Yushi Yi, Mr. Junjun Liu, Mr. Haibao Ma, and Ms. Yuwen Kong. Many of you are not mentioned here.

Finally, I would like to express my cordial gratitude to my parents for their endless understanding and support, and to my grandmother and grandfather for their enduring love. Words cannot express how grateful I am. This dissertation is dedicated to them and in the memory of my grandfather, who always encouraged me to pursue an academic career, whom I miss very much!