Hydration Water Dynamics Near Biological Interfaces

Margaret E. Johnson[1,2,3], Cecile Malardier-Jugroot[4], Rajesh K. Murarka[5], Teresa Head-Gordon[1,2,3*]
[1]*UCSF/UCB Joint Graduate Group in Bioengineering*
[2]*Department of Bioengineering, University of California, Berkeley*
[3]*Physical Biosciences Division, Lawrence Berkeley National Laboratory*
Berkeley, California 94720 USA
[4]*Department of Chemistry and Chemical Engineering, Royal Military College of Canada, Kingston, K7K 7B4, Canada*
[5]*Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, 14853 USA*

We perform classical molecular dynamics simulations using both fixed-charge and polarizable water and protein force fields to contrast the hydration dynamics near hydrophilic and amphiphilic peptides as a function of temperature. The high peptide concentrations we use serve as a model for the surface of folded proteins where hydration layers around each residue overlap significantly. Through simulation we determine that there are notable differences in the water dynamics analyzed from the outer and inner hydration layer regions of the amphiphilic peptide solution that explains the experimentally observed presence of two translational relaxations, while the hydrophilic peptide solution shows only a single non-Arrhenius translational process with no distinction between hydration layers. Given that water dynamics for the amphiphilic peptide system reproduces all known rotational and translational hydration dynamical anomalies exhibited by hydration water near protein surfaces, our analysis provides strong evidence that dynamical signatures near biological interfaces arises because of frustration in the hydration dynamics induced by chemical heterogeneity, as opposed to just topological roughness, of the protein surface.

**Corresponding author*

INTRODUCTION

Folded proteins in aqueous environments have roughly equal percentages of surface exposed hydrophobic and hydrophilic residues[1]. The question that we have been exploring in recent work[2-8] is whether the peculiarities of protein hydration water dynamics, manifest in a number of spectroscopies[9-27], as well as computer simulations[28-36], are due to the amphiphilic character of the protein surface. We have recently reported quasi-elastic neutron scattering (QENS) experiments at two resolutions that probe the hydration dynamics at relatively high concentrations of an amphiphilic[4,7,8] and a hydrophilic peptide[2,6] in water as a function of temperature. The high peptide concentrations we use should provide a reasonable model of the hydration dynamics for the folded state of a protein since the protein surface will also have hydration shells around surface amino acids that overlap, whereas the dilute concentrations of the same peptides explored by recent NMR experiments are more representative of the hydration dynamics of the unfolded state[37].

At room temperature, our QENS investigation of the hydration dynamics around the amphiphilic peptide, N-acetyl-leucine-methyl-amide (NALMA), revealed evidence of dynamical heterogeneity through stretched exponential fits to the intermediate scattering function, while the corresponding anomalous dynamical signatures were found to be absent for the hydrophilic peptide[6]. As the temperature is lowered, the hydration dynamics near the NALMA peptide splits into two translational components; the slower time scale corresponds to a localized relaxation process of what we believe is bound surface water with an Arrhenius temperature dependence, while the second faster relaxation process is non-Arrhenius, but still suppressed with respect to bulk water[2,4]. By contrast the water dynamics near the hydrophilic peptide, N-acetyl-glycine-methyl-amide (NAGMA), collapses to a single translational relaxation with little evidence of dynamical heterogeneity at any temperature[2]. Furthermore, our recent simulation work suggests that models used to analyze QENS experiments in order to extract macroscopic time constants introduce approximations that oversimplify the underlying microscopic dynamics[5]. Together this body of work suggests that a more complete molecular analysis is essential to give insight into the origin of the unusual dynamical properties of competing peptide chemistries as models for understanding water dynamics near heterogeneous biological interfaces.

In this study we use molecular dynamics simulations with fixed charge and polarizable potentials to test their abilities to reproduce the correct temperature trends in solution structure

and the quasi-elastic neutron scattering dynamical data. We then analyze different populations of water molecules in both peptide solutions in order to understand the effects of different side chain chemistries on the water dynamics. We find that the simulated solution structure with fixed-charge force fields (AMBERff03[38] with TIP4P-Ew[39]) predicts too much aggregation of both the hydrophobic and hydrophilic peptide solutes, freeing up too much bulk-like water so to yield water diffusion constants that are notably faster than experiment, displaying an Arrhenius temperature dependence, contradicting the experiments. When we fix the center of mass of the solutes such that they remain separate throughout the solution as determined from liquid diffraction experiments[40-42], we find that the simulated hydration dynamics are very accurate with respect to the experimental trends with temperature. Due to the somewhat unphysical perturbation introduced by fixing the solutes, we also performed the same simulations with the AMOEBA polarizable force field[43-45]. In contrast to the fixed-charge simulations, the polarizable force field nicely reproduces a non-aggregated, uniform distribution of solutes throughout the volume, and provides reasonable agreement with the experimental temperature trends, although the dynamics are far too slow at the lowest temperatures. Overall we find that the simulation models are instructive for isolating the origin of dynamical anomalies seen in our peptides systems that have also been measured near protein surfaces by various spectroscopies[6,8,19-27,32,46,47] which are similar to those seen in glassy liquids[33].

METHODS

Simulation Models and Protocols.

Fixed-Charge Model. Classical molecular dynamics simulations using a fixed charge force field were performed using an in-house simulation program of aqueous solutions of NALMA at two concentrations of 2M and 1M, and the hydrophilic model peptide NAGMA at 1.5M. Although the hydrophilic peptides are smaller in size, the ratio of peptide to water molecules is increased in this solution such that the volume available to the water molecules is comparable to that available in the amphiphilic NALMA solution. This keeps the water at the same level of spatial confinement. Separate MD simulations were carried out at temperatures coinciding with those used in experiment[2,4] which for the 1M NALMA solution had 220 water molecules and 4 solutes, and for the 1.5M NAGMA had 220 water molecules and 6 solutes. We use the AMBERff03[38] all-atom protein force field and potential parameters to model the

NALMA and NAGMA solutes, and the rigid, non-polarizable TIP4P-Ew model[39] for the water. All the electrostatic interactions were calculated using standard Ewald summation with tin-foil boundary conditions. The intra-molecular geometry of the water molecule (r_{OH} and θ_{HOH}) was constrained by applying the M_RATTLE[39] algorithms using an absolute geometric tolerance of 10^{-10}Å. The velocity Verlet algorithm[48] with a time step of 1 fs was used to integrate the equations of motion. For each temperature, all systems were equilibrated in the NPT ensemble at a pressure of 1 atm using the Nose-Hoover Andersen technique to determine the density[49-51]. These simulations ran from anywhere between 0.5-1.6ns, depending on temperature, and the average equilibrated volume was then used for the subsequent NVT and/or NVE simulations. The NPT simulations were equilibrated for 0.5-2.0ns using separate Nose-Hoover thermostats attached to solutes, solvent, and the barostat, with relaxation time constants of 1.0ps and 0.5ps for barostat and thermostat, respectively. For production runs we performed simulations in the NVT ensemble using lightly coupled Nose-Hoover chain thermostats[52] with a relaxation time constant of 5ps to minimally perturb system dynamics. Statistics were collected for 4.0-6.0ns, with longer runs used for lower temperatures. In addition, five independent trajectories were run for an additional 1.5ns for each temperature to collect better statistics on the diffusion constants. During this stage, coordinates were saved both linearly in time (every 0.06 ps), as well non-linearly (in powers of two in time) over multiple 5ps lengths.

We noted that these fixed-charge simulations displayed a large degree of aggregation of the solutes, even when the system size was increased eight fold to 32 NALMA molecules or 48 NAGMA molecules. Unfortunately, the aggregation is not in agreement with the x-ray structural study of NALMA in solution[40]. For instance, the inter-solute (carbon-carbon) radial distribution function (RDF) from the 1M NALMA simulation has extremely high density at short distances that drops off continuously at longer distances, indicating aggregation. A similar degree of aggregation was observed in the NAGMA simulations. In Figure 1 we contrast this fixed-charge RDF with that measured from the AMOEBA polarizable model, which shows a much more uniform distribution of solutes throughout the volume. In addition, the diffusion constants of the fixed-charge 2M NALMA simulations displayed an Arrhenius temperature dependence, which contradicted previous experimental QENS studies on aqueous 2M NALMA[7,8] (results not shown). Hence we switched to running fixed-charge simulations in which the centers-of-mass of the NALMA or NAGMA solutes were held fixed at a configuration showing no van der Waals

contact of the solutes. All the results presented later for the fixed-charge simulations therefore have the centers-of-mass of the solutes fixed.

AMOEBA Polarizable Model. For the polarizable AMOEBA force field field[43-45], we used the parallel version with the particle mesh Ewald (PME) method included in the TINKER software package. These simulations were considerably larger, with the NALMA solution consisting of 32 peptides and 1760 water molecules to produce a 1M solution, while the NAGMA solution consisted of 48 peptides with 1760 water molecules, producing the same 1.5M solution used in the experiments. The Beeman algorithm with a timestep of 1fs was used to numerically integrate the equations of motion. For each of the temperatures, simulations were initially run in the NPT ensemble for 0.5-2ns to equilibrate the box size to 1atm. We then performed NVT simulations at the determined density, using the Berendsen thermostat[36] and equilibration times between 0.5-2.0ns, with longer times used for lower temperatures. The NVE ensemble was used to collect all of the dynamics data, with initial equilibrations from 0.2-1.0ns, followed by production runs of 4.0-6.0ns. In order to best conserve energy, the NVE runs were run with the induced dipole iterative tolerance value set at 1E-5, instead of the default 1E-2. Additional NVE simulations were run for 600-900ps at each temperature where configurations were saved every 10fs for the population analysis. As mentioned above and illustrated in Figure 1, these simulations did not show aggregation of the solutes and therefore no restrictions were imposed on the simulations.

Dynamical observables

The water translational diffusion coefficient was obtained from the MD simulations by calculating the mean square displacement (MSD) of the particles, and using the Einstein relation

$$D_t = \frac{1}{2M} \lim_{t \to \infty} \frac{d}{dt} \left\langle \left| \mathbf{r}(t) - \mathbf{r}(0) \right|^2 \right\rangle \qquad (1)$$

where ***r(t)*** is the position vector of each atomic center at time *t* and M is the dimensionality. For this and subsequent equations, angular brackets denote an ensemble average at a given temperature. The convergence of the diffusion constant was verified by calculating the slope of the MSD over multiple time intervals, over the range 0.2 to 1.5 ns depending on temperature, to ensure the linear regime was reached. The diffusion constants as a function of temperature are fit with the Vogel-Fulcher-Tamman (VFT) equation,

$$D = D_0 \exp\left(\frac{K}{(T - T_0)}\right) \tag{2}$$

where K is the fragility parameter, and T_0 the temperature at which the dynamics diverge at the glass transition temperature. We also evaluate the self-intermediate scattering function (ISF) for water hydrogens:

$$F_T^H(Q,t) = \left\langle exp\{i\boldsymbol{Q}\cdot[\boldsymbol{r}_H(t) - \boldsymbol{r}_H(0)]\}\right\rangle \tag{3}$$

where Q is the momentum transfer. This allows for direct comparison with the QENS observables before any models are necessary.

We find that fits to the ISF required the use of a stretched exponential function

$$A\,exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \qquad . \tag{4}$$

Because the timescale for the relaxation of the stretched exponential decay is determined by both τ and β, we used an average relaxation time to recover a single parameter for comparison with the single exponential decay[53]

$$\overline{\tau} = \int_0^{\infty} dt \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] = \frac{\tau}{\beta}\Gamma\left(\frac{1}{\beta}\right) \quad . \tag{5}$$

The rotational relaxation is captured by the rotational ISF

$$F_R^H(\boldsymbol{Q},t) = \left\langle exp(-i\boldsymbol{Q}\cdot\boldsymbol{b}(0)\,exp(i\boldsymbol{Q}\cdot\boldsymbol{b}(t)\right\rangle = \sum_0^{\infty}(2l+1)\,j_l^2(Qb)C_l(t) \tag{6}$$

where the second term on the right is the exact expansion due to Sears[54]. The vector $\boldsymbol{b}$ is the vector connecting the center of mass to a hydrogen atom, and the lth order rotational correlation functions are given as

$$C_l(t) = \left\langle P_l(\boldsymbol{b}(0)\cdot\boldsymbol{b}(t))\right\rangle \tag{7}$$

where P_l is the lth order legendre polynomial, and in which only terms up to l=2 make significant contributions for the Q<2Å^{-1} range studied[55]. For the experimental analysis,[4] we used a rotational diffusion model in which the correlations functions decay exponentially as

$$C_l(t) = \exp(-l(l+1)D_R t) \tag{8}$$

where D_R is the rotational diffusion constant, and the rotational timescale is defined as $\tau_R = 1/(6D_R)$.

Population analysis.

In order to evaluate the dynamics of different populations of water, we consider two regions: water solvating the hydrophilic backbone atoms (O and N atoms), and hydration water that solvates the hydrophobic side chain carbons, shown in Figure 2. The water molecules of the first hydration shell are defined to be within 4.25Å of the peptide heavy atoms as obtained from the first minimum of the radial distribution function[3]. For the hydrophobic side chains the first nearest-neighbor peak of the radial distribution function of water oxygen is around 3.9Å and the corresponding minimum is around 4.25Å. For the backbone atoms, the radial distribution function has two maxima within 4.25Å, the first one is from the waters associated with the backbone via hydrogen bonds around 2.73Å, and the second one belonging to waters which are in the vicinity of the backbone but without a hydrogen bond at 3.9Å[3]. Therefore in both cases the reasonable cutoff of 4.0Å is chosen for the hydrophobic regions and backbone nitrogens, and 3.0Å for the backbone oxygens. Based on these different domains, configurations were saved every 10 steps, thereby allowing small motions in and out of these regions that smoothes the various metrics for analysis.

In the analysis of the regions, we calculate the time-dependent evolution of each water population within and between regions such that they can be used to regenerate the full water average given in Eq. (1).

$$\left\langle \left| \mathbf{r}(t+\Delta t) - \mathbf{r}(t) \right|^2 \right\rangle = \sum_{i=1}^{\text{Re}\,gions} P_i(\Delta t) * \left\langle \left| \mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t) \right|^2 \right\rangle \tag{9}$$

The separate regions are defined as: water molecules that have remained within the surface hydrophilic regions for the entirety of Δt, those which remained in the surface hydrophobic region for the entirety of Δt, those that remained in the "bulk" (outside of these regions) for the entirety of Δt, and those remaining waters that at some point in Δt transferred from one region to another. For the NAGMA peptide there is no defined hydrophobic region. Additionally, the surface and transfer region data could be combined into a single non-bulk contribution. The percentages, $P_i(\Delta t)$, reflect both the propensity to stay within a region i, as well as the average size of surface vs. bulk populations. The probability of a water molecule remaining isolated within either the surface or bulk regions inevitably drops as time evolves, illustrating an exchange rate between these two regional water populations. We also generalize Eq. (3) and (6) using the same regional analysis for calculation of distinct contributions to the ISF.

RESULTS

Translational Motion

Figures 3 and 4 report the translational diffusion coefficients of water obtained from the AMBERff03/TIP4P-Ew and the AMOEBA polarizable simulations for the 1.5M NAGMA (Figure 3) and the 1M NALMA (Figure 4) solutions compared to the QENS measurements[2,4]. The simulated diffusion constants are from the full water population, and were calculated using Eq. (1), and the QENS values were obtained from experimental data using the jump diffusion model[2,4]. The simulated values of the translational diffusion coefficients for the fixed-charge simulations of 1.5M NAGMA are reasonably close to the experimental value for temperatures ranging from 271K to 288K, and show a non-Arrhenius trend over the full temperature range as per experiment. In addition, the simulated water dynamics for the fixed-charge simulations of 1M NALMA, obtained as an average over all hydration shells of water present in the system, falls between the slow and fast dynamics determined from the separately resolved dynamics of the two QENS experiments for the 1M NALMA solutions from 263K to 248K.

For the AMOEBA force field, we found excellent agreement with experiment at higher temperatures for both peptides. For lower temperatures, the translational diffusion values for 1.5M NAGMA is in reasonable agreement with experiment (Figure 3), while the simulation values decrease far too rapidly with respect to experiment for the 1M NALMA peptide (Figure 4). Although not perfect, both simulation models give reasonable qualitative correlations with experiment, allowing us to investigate the molecular origins of the observed differences between the hydration dynamics for the two peptide chemistries at the same level of confinement.

To investigate the possibility of two separate translational motions in the MD simulations for 1M NALMA solutions, we evaluate and analyze the intermediate scattering function (ISF) as we did in previous work[2,4] for the experimental data for temperatures over the range 298K-248K. Figure 5 compares the experimental and simulated ISF spectra for the 1M NALMA solution at 288K for the AMBERff03/TIP4P-Ew and AMOEBA models, in which there is overall good agreement. In both cases, the short time decay of the correlations is due to the initial ballistic motion as well as decorrelation from the rotational motion of the waters, while the longer time decay is due to translational diffusion. The stretched exponential form (Eq. (4)) is fit beyond the

initial short time region in order to capture the majority of the translational decay region, which for higher temperatures starts at 1ps, while for T<271K they start at 5ps, consistent with the fitting protocol used for the analysis of the experimental ISF[2,4]. The stretched exponential form, when it measures deviations from β=1, is a signature of non-exponential relaxation with a characteristic relaxation time τ that is believed to be related to spatial heterogeneity in the dynamics[56].

As shown in Figure 6a (and Table 1), the 1M NALMA data at 288K using the polarizable AMOEBA model shows a good fit with Eq. (4). However, when we use the same stretched exponential fit to the simulated ISF data taken at 248K (Figure 6b), we see that the fit is far less good. The same trends in fits with temperature were also found using the AMBERff03/TIP4P-Ew with fixed solutes model. Given that a separation of the two translational motions was observed experimentally at lower temperatures, we use a combination of a stretched exponential form and an exponential function (Table 1) to describe the simulated ISF data:

$$F_H(Q,t) = A_1 \exp\left[\left(-\frac{t}{\tau_1}\right)^{\beta}\right] + A_2 \exp\left(-\frac{t}{\tau_2}\right) \qquad (10)$$

Figure 7a shows that the fit using Eq. (10) to the ISF is in good agreement for the AMOEBA data at the low temperature for NALMA. The stretched exponential is the slower translational process for all *Q* values, while the faster component coincides with the single exponential. Unlike the low temperature NALMA data, the fit of the ISF for the hydrophilic NAGMA peptide at 248K is consistent with a single stretched exponential decay (Figure 7b), with no separation of translational diffusion time scales observed, consistent with experiment.

We therefore analyze the dynamics in different regions of the solution in order to determine the molecular origin of the multiple timescales observed in the NALMA solution vs the single timescale observed for NAGMA. Figures 8 and 9 display the time evolution of the inner (surface water) and outer hydration layers ("bulk") for the two peptides using the polarizable model. Figure 8 shows the percentage of water molecules found in each region after a time Δt, as defined in Eq. (9), as a function of temperature. The probabilities in Figure 8 are particularly informative since it reflects the rapidity of exchange between the outer bulk layer and inner hydration surface regions; the more readily the bulk mixes and exchanges with the surface areas, the faster its curve crosses below the transfer water curve. These results indicate

that the bulk and surface hydration populations of water in the NALMA solution remain separated for long time scales which becomes more exaggerated as temperature is lowered, while for NAGMA the transfer and bulk regions lose their timescale distinction, with a much weaker dependence with temperature.

This is amplified in Figure 9 in which we display the mean square displacement with time of water in a given region. For the NALMA solution, there is a noticeable separation in the MSD for the bulk and transfer regions since the outer bulk water is faster than average, compared to the water diffusing through the surface area, which is slower than average. For NAGMA however, the bulk and the surface transfer water curves are nearly indistinguishable. Together, these two curves, bulk and transfer, provide nearly all the contribution to the observed average signal (note Figure 7), indicating that these two populations are good prospects for the sources of the two time scales observed in NALMA, and consistent with the lack of observation of two timescales for NAGMA. We also calculated diffusion constants for these bulk and transfer MSD curves, where we found the diffusion values were largely converged, giving an approximate idea of the different time scales in the regions. Although the simulated diffusion values differ in the two regions for 248K NALMA, (results not shown) they do not separate as significantly as those seen in experiment. This is perhaps a result of needing to specifically define fixed hydration regions in order to evaluate these regional MSDs, when in reality the hydration regions may be less rigidly defined.

From these MSD curves it is also clear that the hydration water confined to either the hydrophobic or hydrophilic region over its lifetime diffuses less for the NALMA solution when compared to the NAGMA solution. In fact there is sub-linear diffusion (a plateau in the MSD) for the isolated hydration regions (the hydrophobic and hydrophilic curves) in both solutions, and most manifest in the hydrophobic region of the amphiphilic peptide. This is one of the known anomalies of protein hydration water that is also evident for these peptide systems.[33] However, it must be noted that the populations of water that are entirely restricted to the hydrophobic or backbone regions contribute very little to the observed total water signal. Very similar trends are observed in the fixed charge model for all the regional data, although the populations for all solutions do not stay separated for as long, likely due to the faster diffusion in this model, and possibly the fixing of the solute centers of mass.

In Figure 10, we continue characterizing the different regions for the two peptides at 248K using the polarizable model by calculating the regional ISFs to compare to the actual experimental observable. It is evident for the NALMA peptide that the outer hydration dynamics is faster than the average, while the hydration surface dynamics is slower than the average, providing further evidence that these two regions are responsible for the observed dual time scales from the experiments. By contrast, both the outer and inner hydration populations for NAGMA give rise to very similar relaxation time scales, indicating that the interaction with the hydration surface of the hydrophilic peptide does not cause a distinctive change in the relaxation of the average water molecule.

By fitting these regional ISF curves to a stretched exponential function, we can measure the average decay times for the water molecules from these bulk and surface/transfer populations for 248K NALMA. In Figure 11 we compare these regional time scales with the two timescales derived from the Eq. (10) fit to the full water ISF curves. We find that the time scales in the two regional populations, the outer bulk-like hydration layer, and the surface transfer water population, agree very well with the separate time scales measured from the full water population. This further emphasizes that splitting the water molecules into bulk-like and hydration layer populations can account for the observed separation in time scales throughout the total water solution.

Rotational Motion

To further compare our simulation results with those from experiment, we calculate the relaxation of the water rotational correlations. The results for the simulations however, did not conform to the exponential decay of the rotational diffusion model used to model the experiment. Instead, the functions $C_l(t)$ given in Eq. (7) were much better described by a stretched exponential, as has been observed in other simulation studies of water[56,57]. Because the experimental results were fit to the l=1 correlation function, we report results in Table 2 for fits to those decay functions of the form

$$C_1(t) = A\exp\left(-\left(\frac{t}{3\tau_R}\right)^{\beta}\right). \qquad (11)$$

Clearly for both models, there is a much more dramatic slow down in the rotational relaxation time for the simulated water compared to the experimental values. These simulated values are

more commensurate with the recent results from molecular dynamics simulations of dielectric relaxation experiments and analysis of QENS models[3,5] as well as NMR experiments on similar peptide systems measured here[37]. The discrepancies are likely due to the assumption of the rotational diffusion model for describing the relaxation within the QENS experiment, which is an approximation introduced in order to extract a single macroscopic timescale, and which is not strictly comparable with the timescales reported with the above stretched exponential fit.

We additionally performed the regional analysis for the rotational relaxation and found the same trends as those observed for the translational motion (Figure 12). In the NALMA solution we see a faster decorrelation in the bulk water population, compared to the slower transfer/surface water. Again we see long-lived states in the hydrophilic and hydrophobic regions of the NALMA peptide that show virtually no rotational decay over the time scale analyzed. Again, for the aqueous NAGMA solution, both populations decay on the same timescale (data not shown). These results further indicate that the hydrophobic side chain of the NALMA solute introduces a greater perturbation into the ability of the water to translate and rotate uniformly throughout the solution. This is because of the amphiphilic peptide chemistry since water translation and rotation relaxation are uniform for the hydrophilic peptide.

Structural Trends

Because we measure differences in dynamics arising from the surface hydration layer and the remaining bulk waters, we finally compare the structural correlations of water molecules in these distinct regions of the solution for NALMA and NAGMA at 248K. We have measured the angular pair correlation function

$$g_{OM}(r,\theta,\varphi) = \left\langle \frac{N_O(r,\theta,\varphi)}{\int_{r1}^{r2} r^2 dr \int_{\varphi 1}^{\varphi 2} \sin(\varphi) d\varphi \int_{\theta 1}^{\theta 2} d\theta} * \frac{V_{tot}}{N_M N_{Wat}} \right\rangle \tag{12}$$

where r, ϕ, and θ are the centers of the volume elements defined by dr, dϕ, and dθ, $N_O(r,\phi,\theta)$ counts the number of water oxygens measured in that element, and the histogram is normalized by the number of reference atoms (denoted M) to water oxygen pairs in the total volume. Eq. (12) yields a value of one when the local density is the same as the average water density of the solution and one is the limit at large r. We considered two reference points in our analysis. Because both peptides have the same backbone structure, we first measured the effects of the solute chemistry on the structure of the water by calculating the distribution of the water oxygen

atoms around the peptide backbone group, where the nitrogen atom was chosen as the origin for the local coordinate frame, the N-H bond determined the z-axis, and the N-C bond then fixed the x-plane. We also considered a second reference point of a central water molecule in the surface or bulk regions, and the corresponding angular correlations with other water molecules. These distributions are calculated by using the $O\text{-}H_1$ bond to define the z-axis and the $O\text{-}H_2$ bond to define the x-plane.

One can see from Figure 13 that the angular correlations of water molecules around the peptide backbone are more uniformly distributed about NAGMA, while greater directionality in the hydrogen-bonding region is found around the NALMA solute, such that waters molecules in the amphiphilic solution less freely sample the volume around the peptide backbone. An analysis of the water distribution around water molecules confined to the hydrophobic side chain of NALMA displays sharply peaked angular correlations, which are nearly as pronounced as they are in the bulk region of the NALMA solution (Figure 14). Because these hydration layer water molecules cannot form hydrogen bonds with the hydrophobic side chain, they instead prefer to orient with respect to any adjacent water molecules. This necessary reorienting of molecules to maximize favorable contacts around the side chain could frustrate the dynamics with respect to the bulk waters. When averaged over all angles for water molecules in the bulk regions, we see enhanced structure for NALMA relative to NAGMA (Figure 15), indicating that the hydrogen-bonded network is more rigidly constrained for the amphiphilic peptide solution. This likely raises the barriers to rotational and translational diffusion between the inner and outer layer regions of water, so that the hydration dynamics are more strongly suppressed for NALMA relative to NAGMA.

DISCUSSION AND CONCLUSION

In a large body of work on the observed transport properties of hydration water near different peptide chemistries[2-8], we have determined that the amphiphilic chemistry is a primary source of anomalous hydration dynamics that is also observed for hydration water near protein surfaces. Experimental quasi-elastic neutron scattering investigation of the hydration dynamics around the amphiphilic NALMA reveals stronger dynamical heterogeneity through stretched exponential fits to the intermediate scattering function, while the corresponding anomalous dynamical signatures were found to be largely absent for the hydrophilic NAGMA peptide.

When investigated as a function of temperature, the dynamical heterogeneity for hydration water dynamics for NALMA further resolves cleanly into two experimentally observable translational components. In this work, through analysis of molecular dynamics simulations, we have confirmed that the origin of the two experimentally observed timescales is the separation of the water into a surface hydration layer that undergoes slow exchange with the bulk water in the amphiphilic NALMA peptide. In the hydrophilic NAGMA peptide, however, the surface hydration waters exchange more readily with the bulk, and both these populations of water in the solution produce nearly the same relaxation rates. One important clarification is that the hydrogen-bonded network near the hydrophobic regions are in fact the slowest, likely due to limited hydrogen-bonding arrangements available, frustrating the remaining hydration layer completely. Furthermore we have shown that sub-linear diffusion is most pronounced near the hydrophobic side chain, and provides a prediction for the molecular origin for the same observed sub-linear diffusion near protein surfaces.

Our results on the simulated single-particle rotational dynamics offer complementary information to our recent room temperature molecular dynamics simulations of the frequency-dependent dielectric relaxation (DR) spectra[3], which probes collective rotational relaxations, of these same peptide systems. We found that our high concentration NALMA solution showed two rotational relaxation processes with times on the order of ~20-60ps and ~1ns, timescales that are consistent with the two weak δ–relaxation processes present in the dielectric dispersion profiles measured at room temperature for aqueous protein systems. While there is consensus that the faster δ-dispersion corresponds to water dynamics near the protein surface, the slower δ-dispersion assignment has been more controversial, with molecular origins thought to be due to either protein-water coupling[58-60] or hydration and bulk water exchange[22,61]. Our single-particle rotational analysis remains consistent with our analysis of the DR simulations, which is that the slower δ–relaxation is attributable to the small populations of water that remain strongly coupled to the peptide in the hydrophobic or hydrophilic region (Figure 12).

These concentrated peptide solution systems provide a challenge to the effectiveness of fixed-charge force fields, which should break down when studying the dynamics of chemically heterogeneous systems away from ambient conditions. While the TIP4P-Ew model is a significant improvement over the SPC model[62] when combined with the AMBERff03 force field, the non-polarizable force field overemphasizes aggregated solution structure, that in turn frees up

too much bulk water that results in poor agreement with the experimentally observed non-Arrhenius temperature trends in the water translational diffusion. However, when the centers of mass of the peptide solutes are fixed so that they are not in van der Waals contact, then we find excellent agreement of the simulation with all experimental dynamical signatures, including the non-Arrhenius behavior. Because the aggregation of the solutes prevents more exposure of the solute surface to the water molecules, we conclude that this contact of solute with the solvent is crucial to observe the known non-Arrhenius behavior of water dynamics in these solutions. The need to fix the center of mass however suggests that the interaction potential of the protein-water components needs to be reoptimized, a worthwhile effort for the TIP4P-Ew non-polarizable water model that performs remarkably well for bulk properties over large temperature and pressure ranges[39,63]. The AMOEBA polarizable force field[44,45] by contrast has a much more accurate solution structure with little aggregation, and reproduces the temperature trends of the experimentally measured hydration dynamics, although quantitatively its hydration dynamics are too slow at the lower temperatures studied, especially for the NALMA peptide. While this highlights the importance of including polarizability for simulations of heterogeneous protein solutions, there is still further improvement needed in their accuracy.

We have shown that both AMBERff03/TIP4P-Ew and AMOEBA models yield a simulated ISF that is in good agreement with the experimental ISF at 288K, while direct comparison to experiment is not possible for NALMA as temperature is lowered because of the separation of the two translational dynamical processes, requiring more detailed characterization and analysis. We find that the presence of the hydrophobic leucine side chain in the amphiphilic peptide solution provokes a larger disruption to the solution water network, which gives rise to distinguishable populations of water molecules whose translational and rotational correlations relax at two different time scales. By contrast, these distinctions disappear entirely for NAGMA, which shows no splitting of translational timescales at all temperatures no matter how cold.

This strong heterogeneity and homogeneity in dynamics depending on peptide chemistry would appear to be important for biological interfaces that are almost always amphiphilic. In fact, the water dynamics for the high concentration amphiphilic peptide system, appropriate as a model for water near protein surfaces in which there are overlapping hydration shells due to close proximity of different amino acid chemistries, reproduces all known rotational and translational hydration dynamical anomalies exhibited by hydration water near protein surfaces.

This work provides proof that hydration dynamical signatures near biological interfaces arises from, in part, chemical heterogeneity (or energy disorder[28]) as opposed to mere topological roughness of the protein surface[17].

ACKNOWLEDGMENTS. We gratefully acknowledge the support of the Department of Energy, Condensed Phase and Interfacial Molecular Science Program, DE-AC02-05CH11231 and NERSC for computational resources. We would like to thank the NIST Center for Neutron Research (NCNR) for the beam time used to collect the experimental results.

REFERENCES

(1) Janin, J. *Structure with Folding & Design* **1999**, *7*, R277.
(2) Malardier-Jugroot, C.; Johnson, M. E.; Murarka, R. K.; Head-Gordon, T. *Physical Chemistry Chemical Physics* **2008**, *10*, 4903.
(3) Murarka, R. K.; Head-Gordon, T. *J Phys Chem B* **2008**, *112*, 179.
(4) Malardier-Jugroot, C.; Head-Gordon, T. *Phys Chem Chem Phys* **2007**, *9*, 1962.
(5) Murarka, R. K.; Head-Gordon, T. *J Chem Phys* **2007**, *126*, 215101.
(6) Russo, D.; Murarka, R. K.; Copley, J. R.; Head-Gordon, T. *J Phys Chem B* **2005**, *109*, 12966.
(7) Russo, D.; Murarka, R. K.; Hura, G.; Verschell, E. R.; Copley, J. R. *J Phys Chem B* **2005**, *108*, 19885.
(8) Russo, D.; Hura, G.; Head-Gordon, T. *Biophysical Journal* **2004**, *86*, 1852.
(9) Bhattacharyya, K. *Chemical Communications* **2008**, *25*, 2848.
(10) Helms, V. *Chemical Physics Physical Chemistry* **2007**, *8*, 23.
(11) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. *Proc Natl Acad Sci U S A* **2007**, *104*, 20749.
(12) Li, T.; Hassanali, A. A.; Kao, Y. T.; Zhong, D.; Singer, S. J. *Journal of the American Chemical Society* **2007**, *129*, 3376.
(13) Chen, S. H.; Liu, L.; Fratini, E.; Baglioni, P.; Faraone, A.; Mamontov, E. *Proc Natl Acad Sci U S A* **2006**, *103*, 9012.
(14) Doster, W.; Settles, M. *Biochimica et. Biophysica Acta* **2005**, *1749*, 173.
(15) Bagchi, B. *Chemical Reviews* **2005**, *105*, 3197.
(16) Modig, K.; Liepinsh, E.; Otting, G.; Halle, B. *J Am Chem Soc* **2004**, *126*, 102.
(17) Halle, B. *Philos Trans R Soc Lond B Biol Sci* **2004**, *359*, 1207.
(18) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. *Journal of Physical Chemistry B* **2002**, *106*, 12376.
(19) Dellerue, S.; Petrescu, A. J.; Smith, J. C.; Bellissent-Funel, M. C. *Biophysical Journal* **2001**, *81*, 1666.
(20) Dellerue, S.; Bellissent-Funel, M. C. *Chemical Physics* **2000**, *258*, 315.
(21) Bellissent-Funel, M. C. *Journal of Molecular Liquids* **2000**, *84*, 39.
(22) Nandi, N.; Bagchi, B. *Journal of Physical Chemistry* **1998**, *102*, 8217.
(23) Diehl, M.; Doster, W.; Petry, W.; Schober, H. *Biophysical Journal* **1997**, *73*, 2726.

(24) Bellissentfunel, M. C.; Zanotti, J. M.; Chen, S. H. *Faraday Discussions* **1996**, 281.
(25) Pethig, R. *Annu Rev Phys Chem* **1992**, *43*, 177.
(26) Bellissentfunel, M. C.; Teixeira, J.; Bradley, K. F.; Chen, S. H. *Journal de Physique I* **1992**, *2*, 995.
(27) Grant, E. H.; Mitton, B. G.; South, G. P.; Sheppard, R. J. *Biochem J* **1974**, *139*, 375.
(28) Pizzitutti, F.; Marchi, M.; Sterpone, F.; Rossky, P. J. *J Phys Chem B* **2007**, *111*, 7584.
(29) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Bagchi, B. *J Am Chem Soc* **2005**, *127*, 4071.
(30) Pal, S.; Balasubramanian, S.; Bagchi, B. *Journal of Chemical Physics* **2004**, *120*, 1912.
(31) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Pal, S.; Bagchi, B. *Journal of Physical Chemistry B* **2004**, *108*, 12608.
(32) Orecchini, A.; Paciaroni, A.; Bizzarri, A. R.; Cannistraro, S. *Journal of Physical Chemistry B* **2001**, *105*, 12150.
(33) Bizzarri, A. R.; Cannistraro, S. *Journal of Physical Chemistry B* **2002**, *106*, 6617.
(34) Tarek, M.; Tobias, D. J. *Biophysical Journal* **2000**, *79*, 3244.
(35) Merzel, F.; Smith, J. C. *Proc Natl Acad Sci U S A* **2002**, *99*, 5378.
(36) Golosov, A. A.; Karplus, M. *J Phys Chem B* **2007**, *111*, 1482.
(37) Qvist, J.; Halle, B. *J Am Chem Soc* **2008**, *130*, 10345.
(38) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J Comput Chem* **2003**, *24*, 1999.
(39) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *Journal of Chemical Physics* **2004**.
(40) Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. *Perspectives in Drug Discovery & Design* **1999**, *17*, 97.
(41) Sorenson, J. M.; Hura, G.; Soper, A. K.; Pertsemlidis, A.; Head-Gordon, T. *Journal of Physical Chemistry B* **1999**, *103*, 5413.
(42) Pertsemlidis, A.; Soper, A. K.; Sorenson, J. M.; Head-Gordon, T. *Proc Natl Acad Sci U S A* **1999**, *96*, 481.
(43) Ponder, J. W.; Case, D. A. *Adv Protein Chem* **2003**, *66*, 27.
(44) Ren, P.; Ponder, J. *J Phys Chem B* **2003**, *107*, 5933.
(45) Ren, P.; Ponder, J. W. *J Comput Chem* **2002**, *23*, 1497.
(46) Bellissent-Funel, M. C. *Journal of Molecular Liquids* **2002**, *96-7*, 287.
(47) Wanderlingh, U.; Giordano, R.; Teixeira, J. *Journal of Molecular Structure* **1993**, *296*, 271.
(48) Anderson, H. C. *J. Comp. Phys.* **1983**, *52*, 24.
(49) Hoover, W. G. *Phys Rev A* **1985**, *31*, 1695.
(50) Nose, S. *Journal of Chemical Physics* **1984**, *81*, 511.
(51) Andersen, H. C. *Journal of Chemical Physics* **1980**, *72*, 2384.
(52) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *Journal of Chemical Physics* **1992**, *97*, 2635.
(53) Zanotti, J. M.; Bellissent-Funel, M. C.; Chen, S. H. *Physical Review E* **1999**, *59*, 3084.
(54) Sears, V. F. *Canadian Journal of Physics* **1967**, *45*, 237.
(55) Chen, S. H.; Gallo, P.; Sciortino, F.; Tartaglia, P. *Physical Review E* **1997**, *56*, 4231.
(56) Sciortino, F.; Gallo, P.; Tartaglia, P.; Chen, S. H. *Physical Review E* **1996**, *54*, 6331.
(57) Kammerer, S.; Kob, W.; Schilling, R. *Phys Rev E* **1997**, *56*, 5450.

(58) Oleinikova, A.; Sasisanker, P.; Weingartner, H. *Journal of Physical Chemistry B* **2004**, *108*, 8467.
(59) Boresch, S.; Willensdorfer, M.; Steinhauser, O. *Journal of Chemical Physics* **2004**, *120*, 3333.
(60) Boresch, S.; Hochtl, P.; Steinhauser, O. *Journal of Physical Chemistry B* **2000**, *104*, 8743.
(61) Nandi, N.; Bhattacharyya, K.; Bagchi, B. *Chemical Reviews* **2000**, *100*, 2013.
(62) Berendsen, H. J.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(63) Horn, H. W.; Swope, W. C.; Pitera, J. W. *Journal of Chemical Physics* **2005**, *123*.

FIGURES

Figure 1. *Representative radial distribution functions (RDF) for the 1M NALMA solution at 298K in the fixed charge (black) vs the polarizable (red) force fields.* Carbon-carbon RDF for the fixed-charge simulations shows an excess of solute density at short range, indicating aggregation.

Figure 2. Definition of the hydrophilic (green) and hydrophobic (blue striped) parts around NALMA.

Figure 3. *Arrhenius representation of the theoretically and experimentally determined D_t for the 1.5M NAGMA solution.* VFT fit (solid line) is to the simulation data (black circles). (a) Fixed-charge model with fixed solutes and (b) polarizable model.

Figure 4. *Arrhenius representation of the theoretically and experimentally determined D_t for the 1M NALMA solution.* VFT fit (solid line) is to the simulation data (black circles). (a) Fixed-charge model with fixed solutes and (b) polarizable model. The Disc Chopper Spectrometer (DCS) and High Flux Backscattering Spectrometer (HFBS) were the two neutron scattering instruments used at the NIST Center for Neutron Research (NCNR).

Figure 5: *Experimental and simulated intermediate scattering function for 1M NALMA at 288K for three different Q values, 0.694 (black), 0.947 (blue), and 1.253 (magenta).* Results are shown for both the fixed-Charge model (lines with symbols), and the AMOEBA polarizable model (dashed lines).

Figure 6: *Intermediate scattering function obtained from the MD simulations using the AMOEBA polarizable model of 1M NALMA (symbols) and fit obtained with a stretched exponential function (black line).* (a) 288K and (b) 248K. The 248K emphasize that the fits using Eq. (4) are inadequate.

Figure 7: *Intermediate scattering function obtained from the AMOEBA simulations and fits to one- or two-translational processes at low temperature.* (a) For 1M NALMA at 248K, the fit was obtained with a stretched exponential function plus a single exponential as per Eq. (10). (b) For 1.5M NAGMA at 248K, the fit is based on Eq. (4), so that no second process is present for the hydrophilic peptide.

Figure 8. *Populations of water molecules found in the surface and bulk regions of the two peptides as a function of time for different temperatures using the AMOEBA model.* (a) Bulk water in the NALMA solution more slowly mixes with the hydration water as the temperature is lowered, which causes more perceptible differences in the time scales from these regions. (b) Bulk water in the NAGMA solution mixes with the hydration water more rapidly, giving rise much more quickly to a population of indistinguishable water molecules, and is only weakly dependent on temperature. Solid lines are the bulk water, dashed lines are the hydration layer transfer water.

Figure 9. *Mean square displacement (MSD) as a function of time for populations of water molecules found in the hydration layer and bulk regions at 248K for the AMOEBA model (a) NALMA and (b) NAGMA.* For the NALMA solution, there is a clear separation in diffusivity between the different bulk and hydration layer populations. For NAGMA, all populations diffuse on the same time scale. For both models, the water molecules restricted to just the hydrophilic or hydrophobic region move much slower, but contribute very little to the average observed RMSD, as shown in Figure 8.

Figure 10: *Intermediate scattering function (ISF) for populations of water molecules found in the surface and bulk regions for the AMOEBA model at 248K for Q=0.9538 (a) NALMA and (b) NAGMA.* The origin of the two translational motions arise from the well-separated slow

hydration water dynamics and transfer region relative to the fast bulk dynamics for NALMA, while it is evident that no second process is exhibited for the hydrophilic NAGMA peptide.

Figure 11: *Average ISF relaxation times,* $\bar{\tau}$ *(Eq. 5), for the AMOEBA model at 248K for NALMA as a function of Q.* The separation of the full water ISF into two distinct timescales, τ_2 and τ_1 (Eq. 10), is well reproduced by the individual time scales of water molecules in the bulk region, and water molecules transferring through the surface regions, respectively.

Figure 12: *Rotational correlation functions,* $C_1(t)$*, for populations of water for the AMOEBA model at 248K for NALMA* Just as in the translation motion, the rotational correlations decay at two different rates for the different populations for the NALMA solution, but at the same rate for the NAGMA. The noise in the hydrophilic/hydrophobic curves is due to the small number of water molecules that persist for long times in those regions.

Figure 13: *Density of water oxygens around the peptide bond for the AMOEBA model for (a) NALMA and (b) NAGMA*. The density is shown to a radius of 4.1Å from the backbone nitrogen atom (gray), and around the amide hydrogen (gold) and the backbone carbon atom (purple). A value of 1 (green) indicates the average water density. For the NALMA solution, the density is more sparsely distributed around the peptide, but more concentrated directly above the N-H bond. For NAGMA, the distribution is more uniform, suggesting it is less disruptive to the structure of the water network.

Figure 14: *Density of water oxygens around a water molecule from (a) the surface of the hydrophobic region and (b) the bulk region for the NALMA solution.* The density is shown at a radius from 2.7Å to 4.1Å from the reference water oxygen atom (gray). (hydrogen in red determines the z-axis, and the hydrogen in purple determines the x-plane). The angular regions of high density are very similar between the hydrophobic surface water and the bulk water.

Figure 15: *Radial distribution function of water in peptide solutions.* By performing angular averages in the surface and bulk regions, we see that the water is more sharply distributed around the NALMA relative to NAGMA.

TABLES

Table 1: ISF fit parameters for Q=0.95 Å^{-1} using equation $F_H(Q,t) = A_1 exp\left(-\frac{t}{\tau_1}\right)^{\beta} + A_2 exp\left(-\frac{t}{\tau_2}\right)$ for both solutions with both simulation models at 288K and 248K.

Parameters	A_1	τ_1/ps	A_2	τ_2/ps	β
NALMA 1M 288K-Fixed-Charge	0.95	10.43	-	-	0.80
NALMA 1M 248K-Fixed Charge	0.72	69.25	0.23	70.94	0.73
NAGMA 1.5M 298K-Fixed Charge	0.93	7.27	-	-	0.86
NAGMA 1.5M 248K-Fixed Charge	0.91	52.90	-	-	0.87
NALMA 1M 288K-Polarizable	0.95	11.70	-	-	0.79
NALMA 1M 248K-Polarizable	0.68	334.00	0.27	302	0.64
NAGMA 1.5M 298K-Polarizable	0.92	7.89	-	-	0.86
NAGMA 1.5M 248K-Polarizable	0.92	202.5	-	-	0.80

Table 2: Fits to rotational correlation, $C_1(t)$ using Eq. (8), compared to experimental fits using the rotational diffusion approximation.

Parameters	*A*	β	τ_R / ps	$\bar{\tau}_R / ps$
NALMA (Amoeba) 248K	0.961	0.709	85.9	107.4
NAGMA (Amoeba) 248K	0.946	0.765	50.56	59.233
NALMA 248K (Exp- DCS)	-	-	3.238	-
NALMA (Fixed charge) 248K	0.936	0.832	18.76	20.69
NAGMA (Fixed charge) 248K	0.94	0.83	14.81	16.37

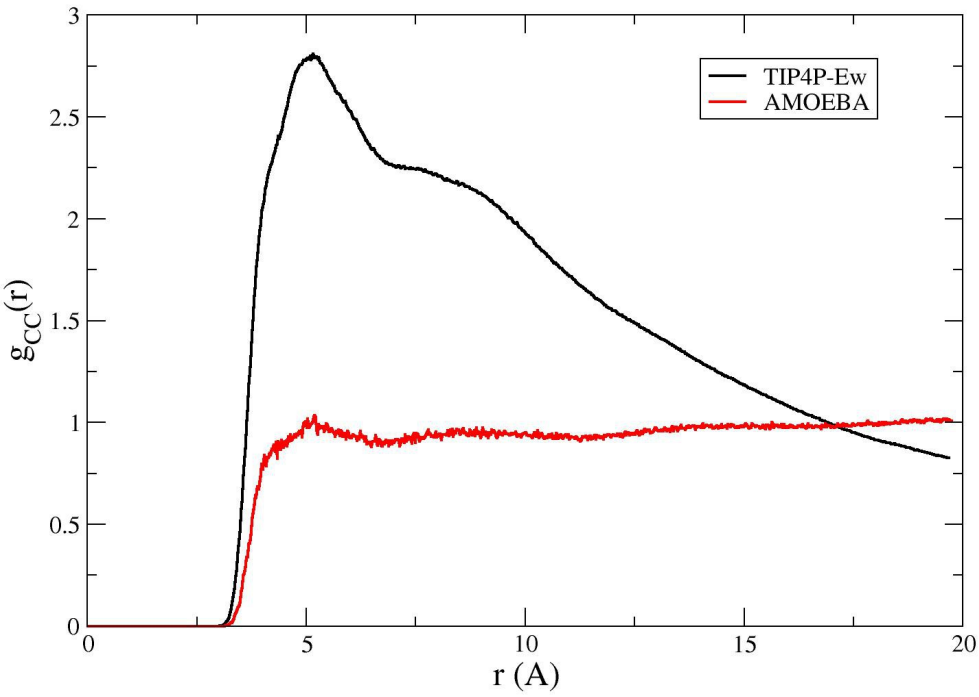


Figure 1. Johnson and co-workers

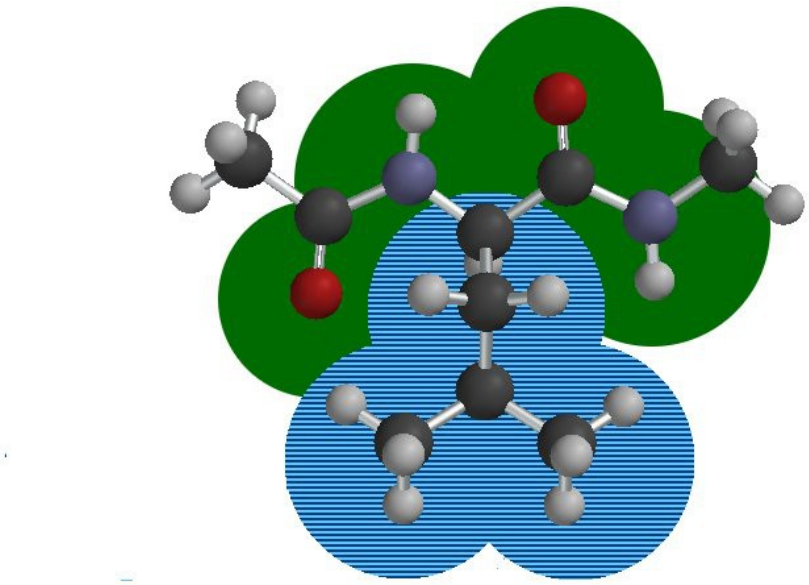


Figure 2. Johnson and co-workers

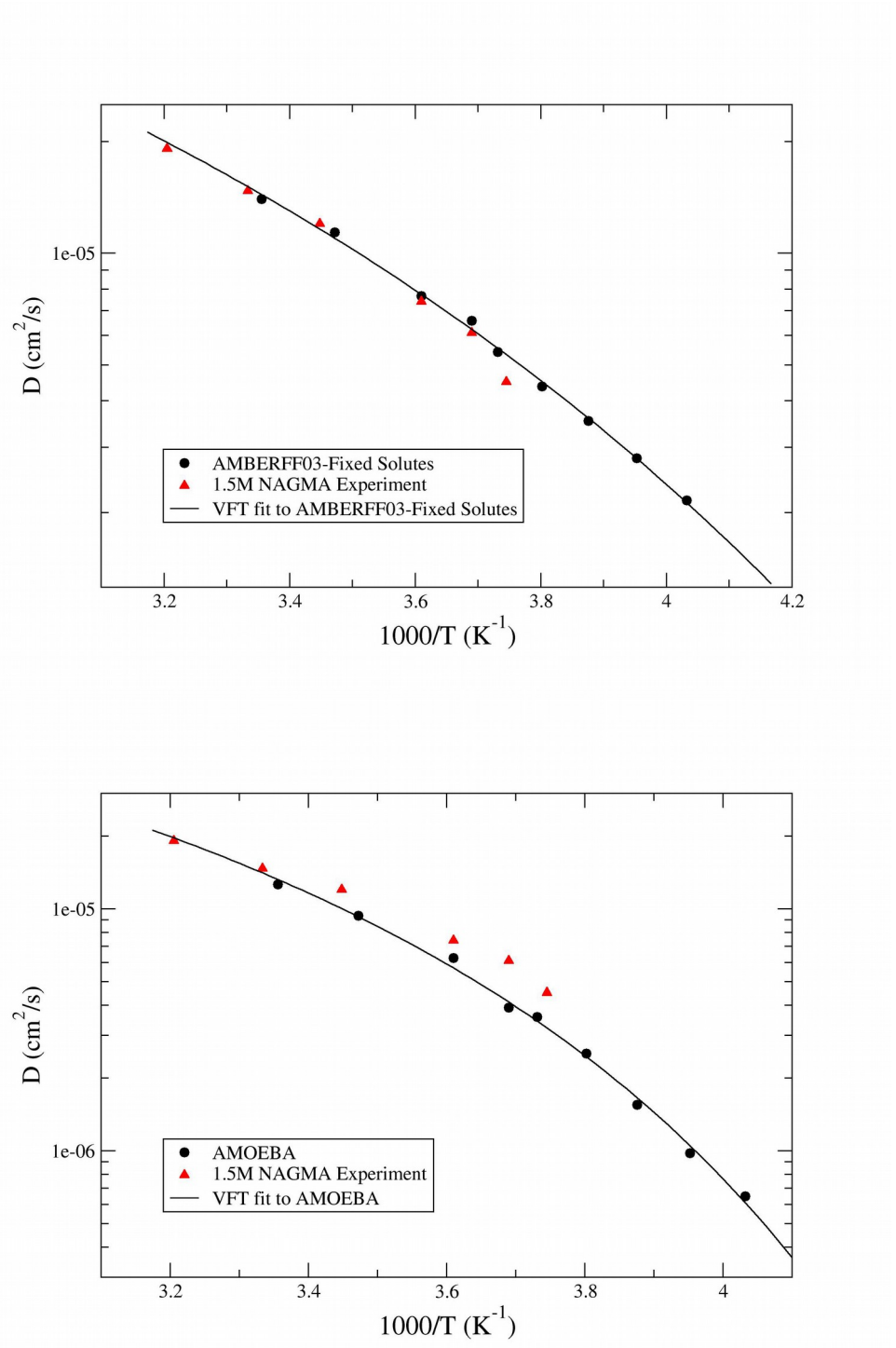


Figure 3. Johnson and co-workers

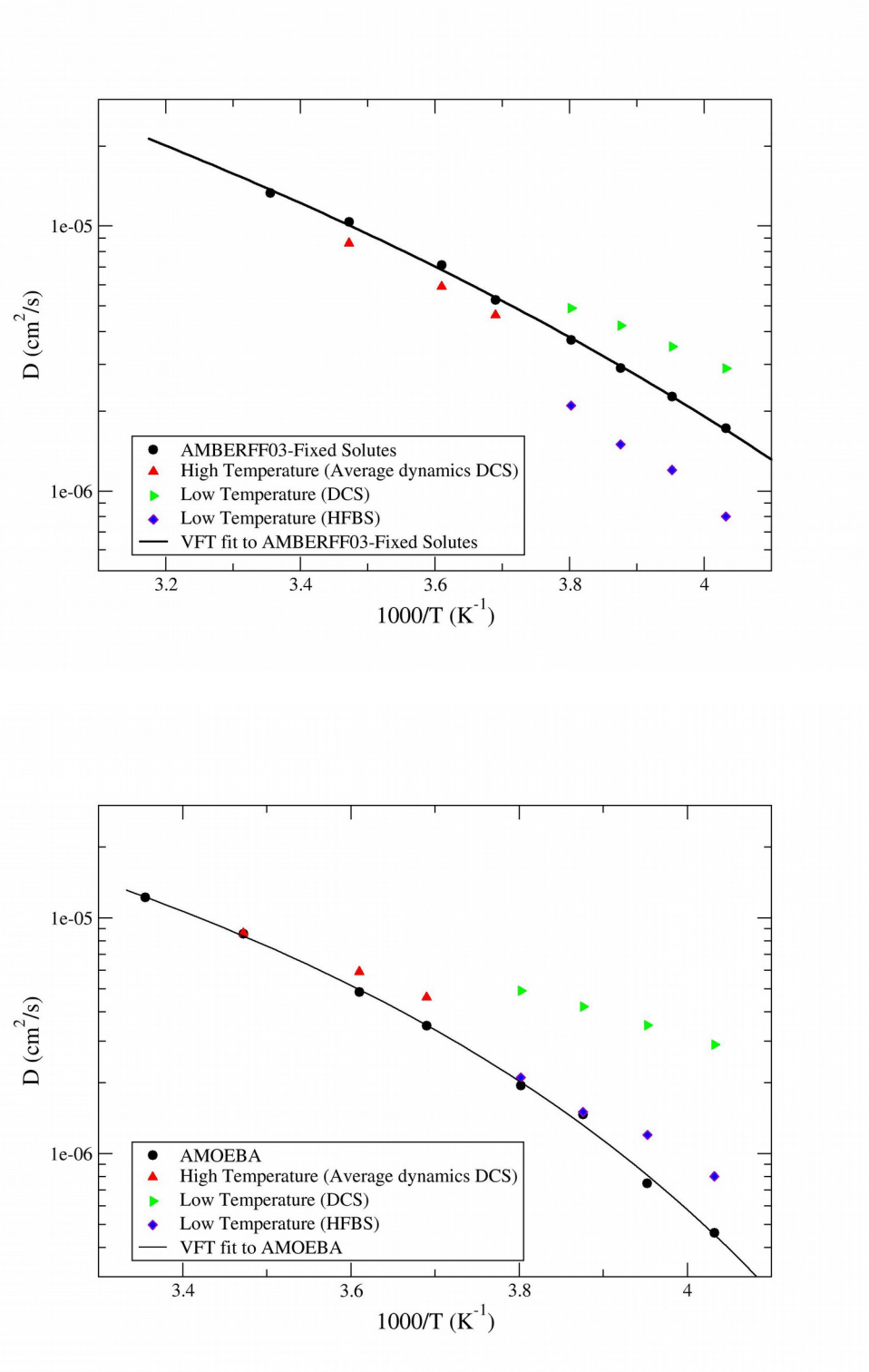


Figure 4. Johnson and co-workers

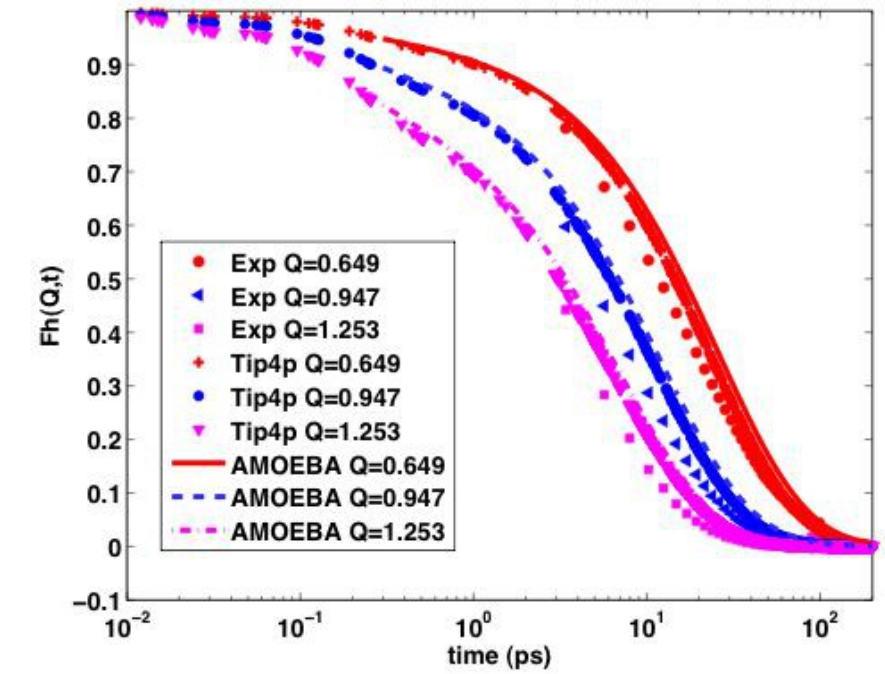


Figure 5. Johnson and co-workers

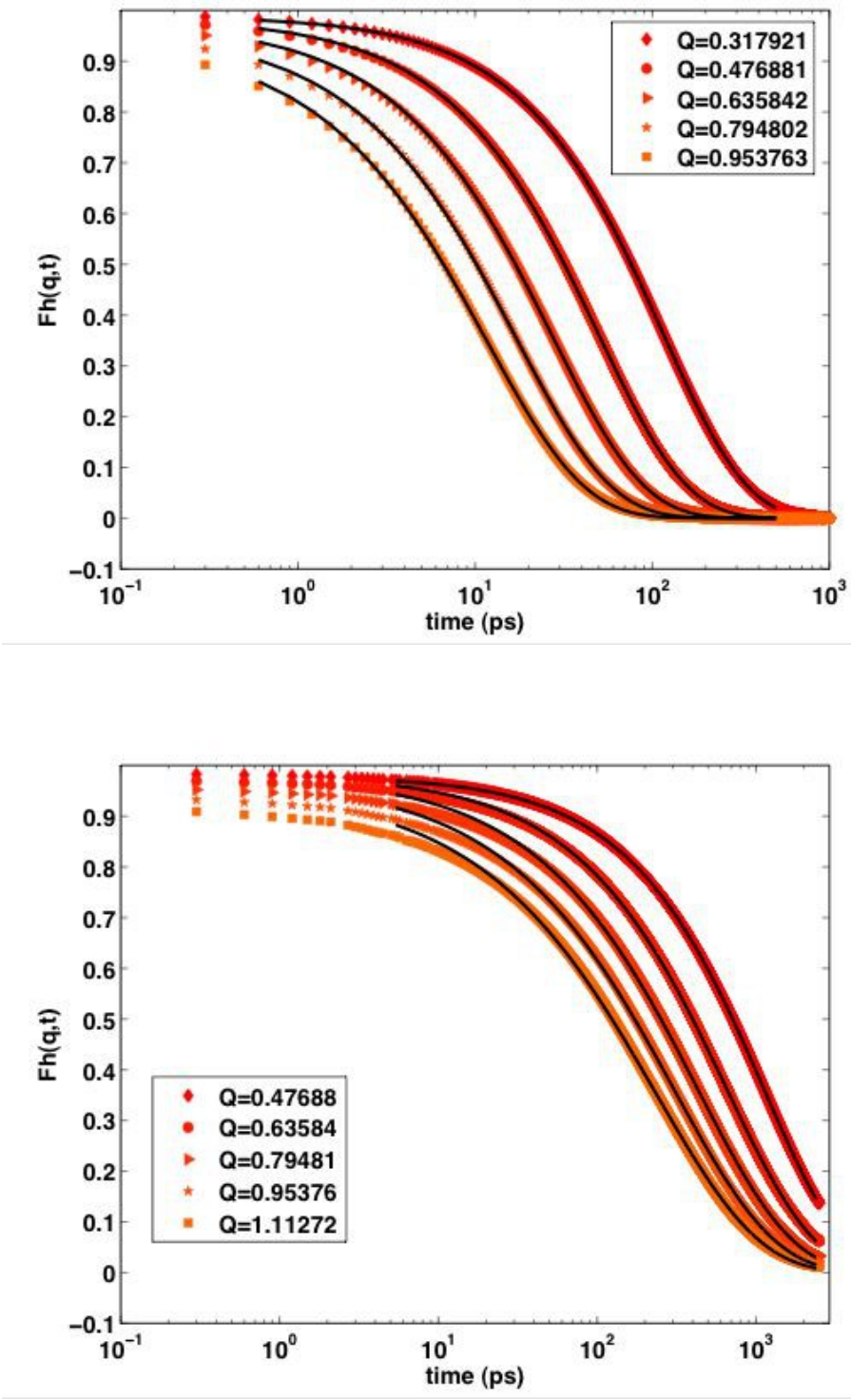


Figure 6. Johnson and co-workers

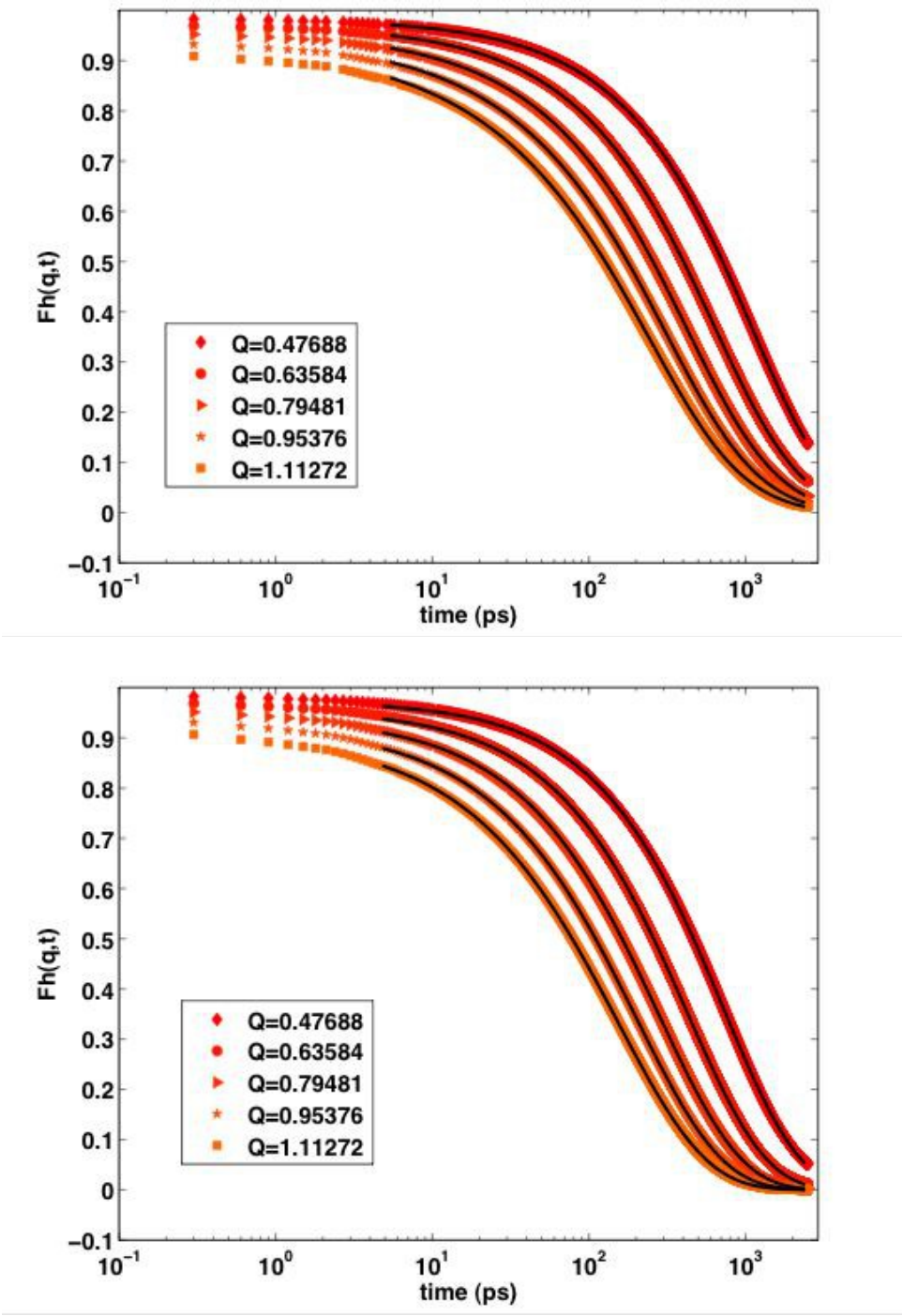
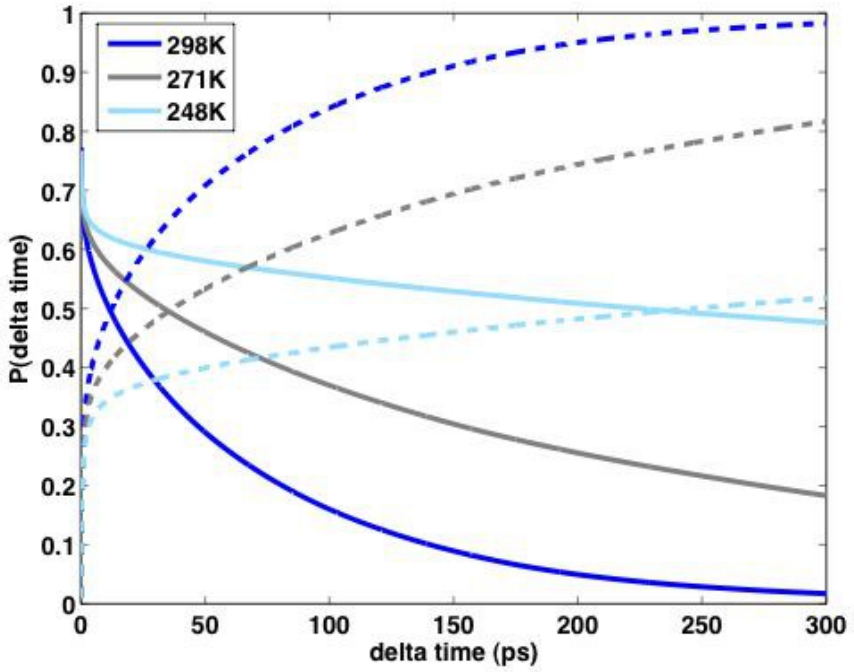


Figure 7. Johnson and co-workers



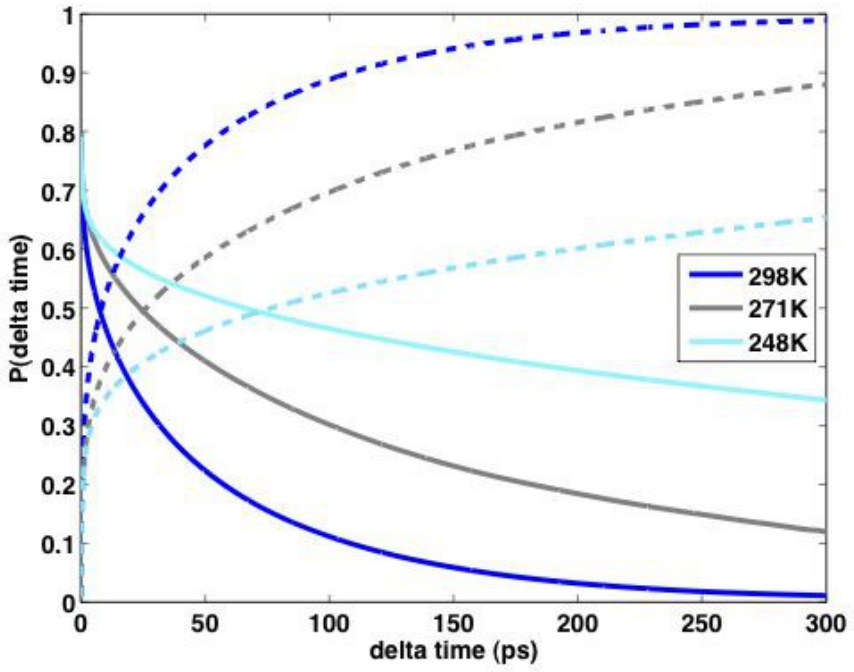
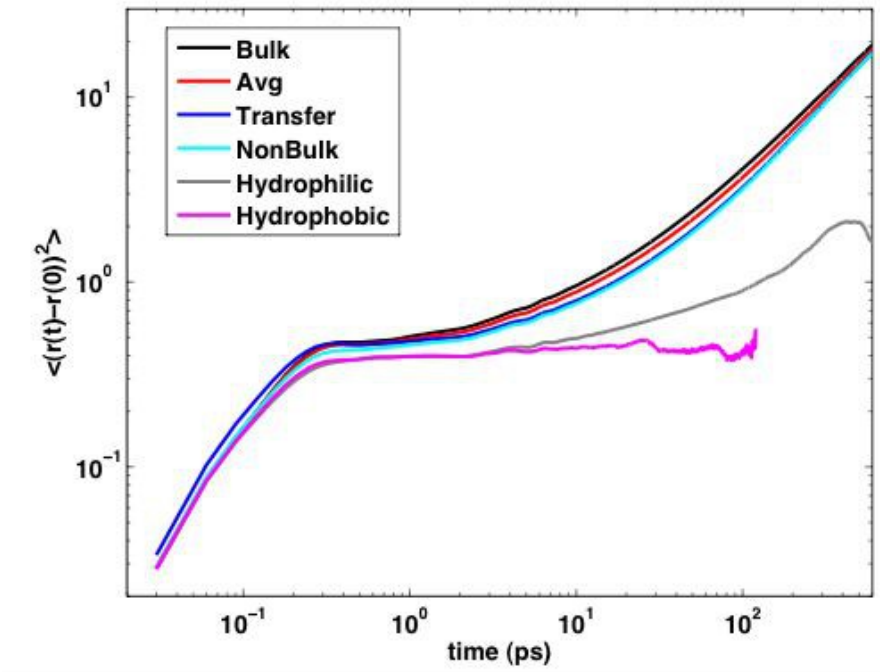


Figure 8. Johnson and co-workers



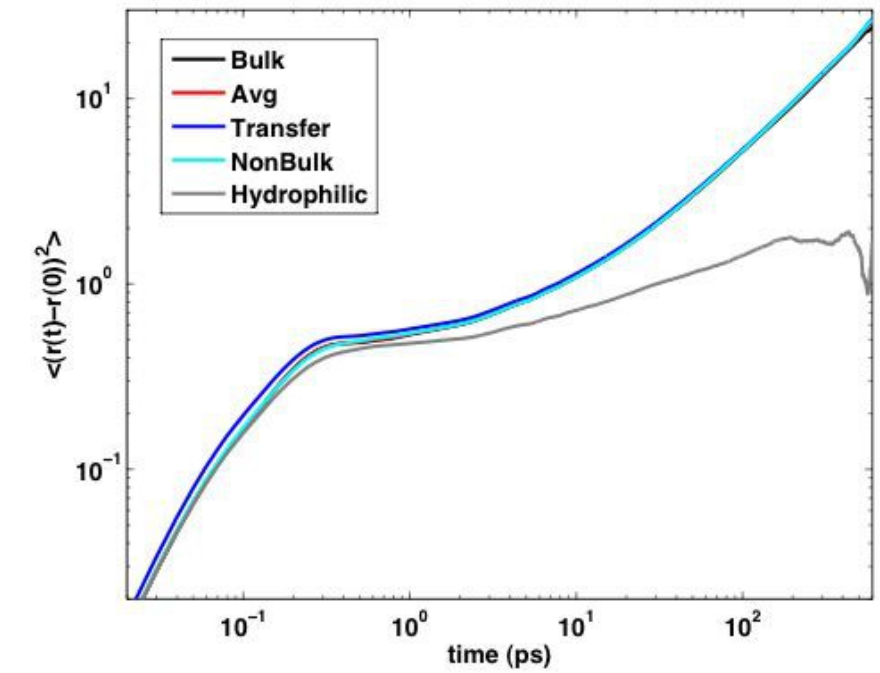


Figure 9. Johnson and co-workers

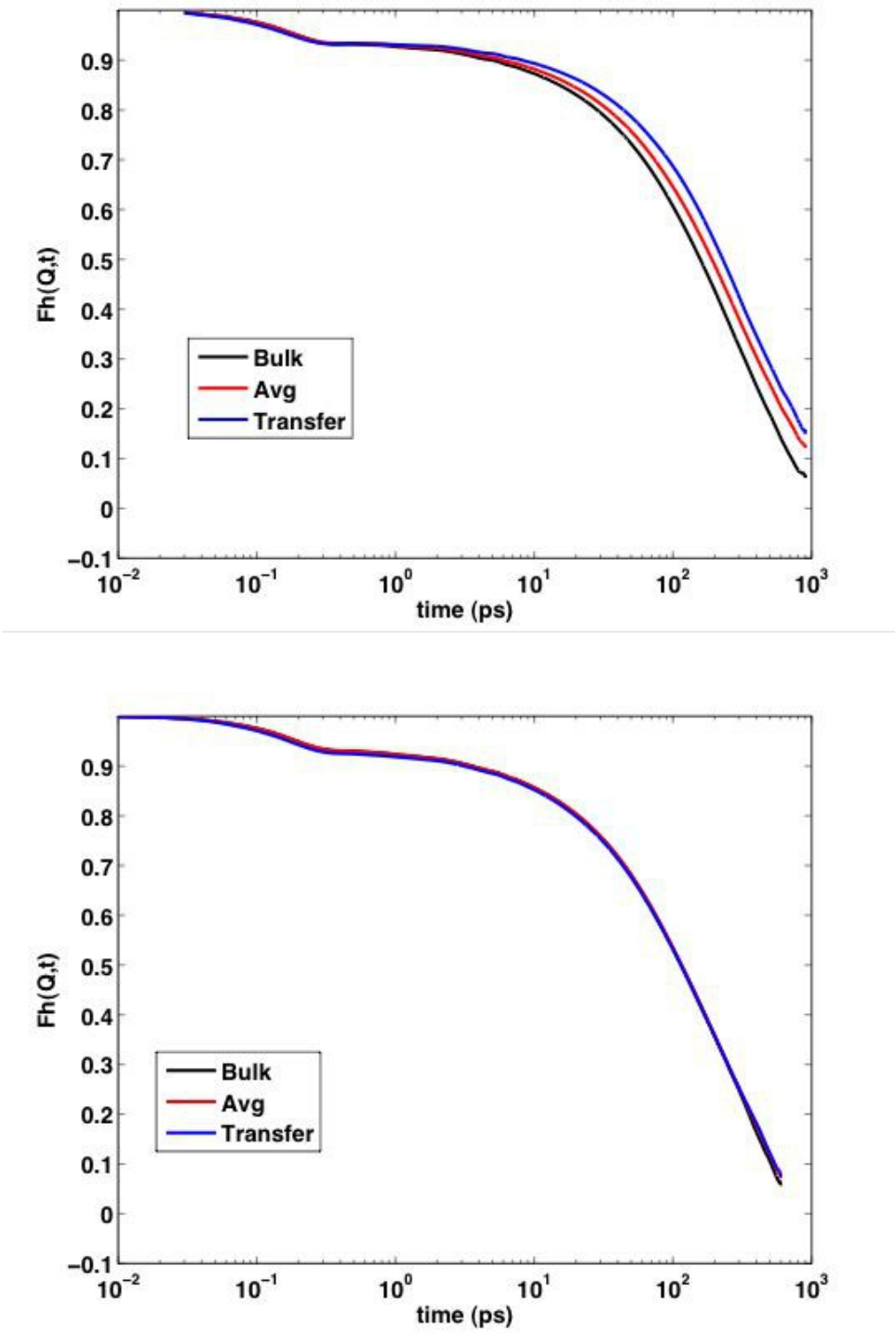


Figure 10. Johnson and co-workers

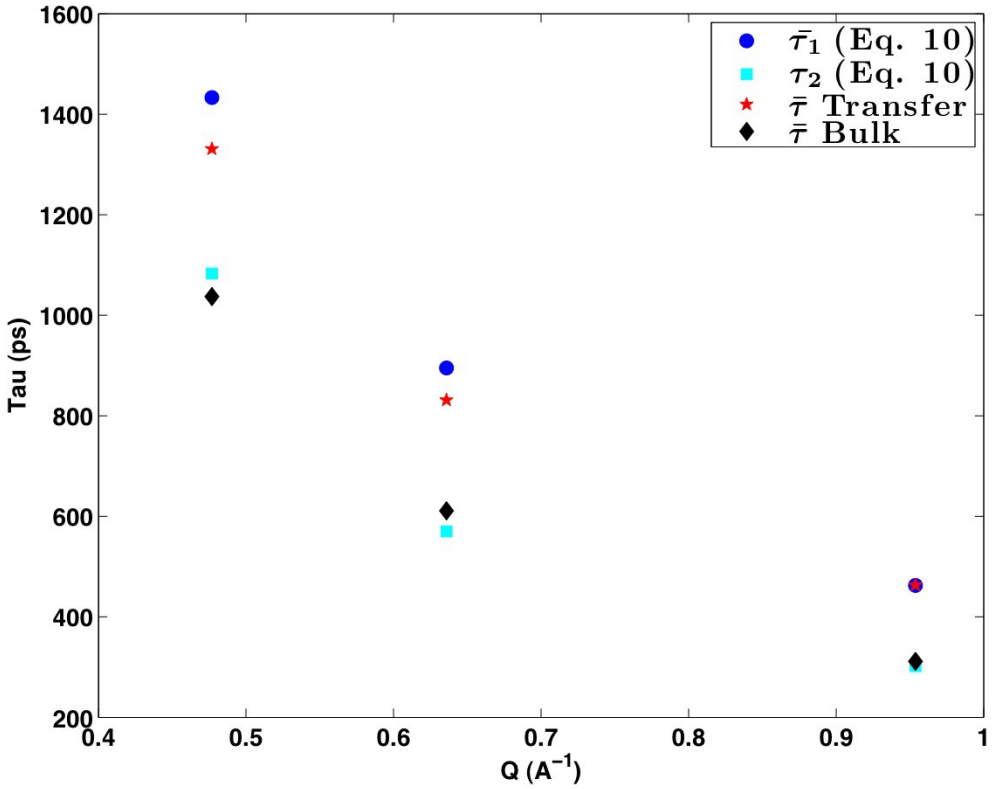


Figure 11. Johnson and co-workers

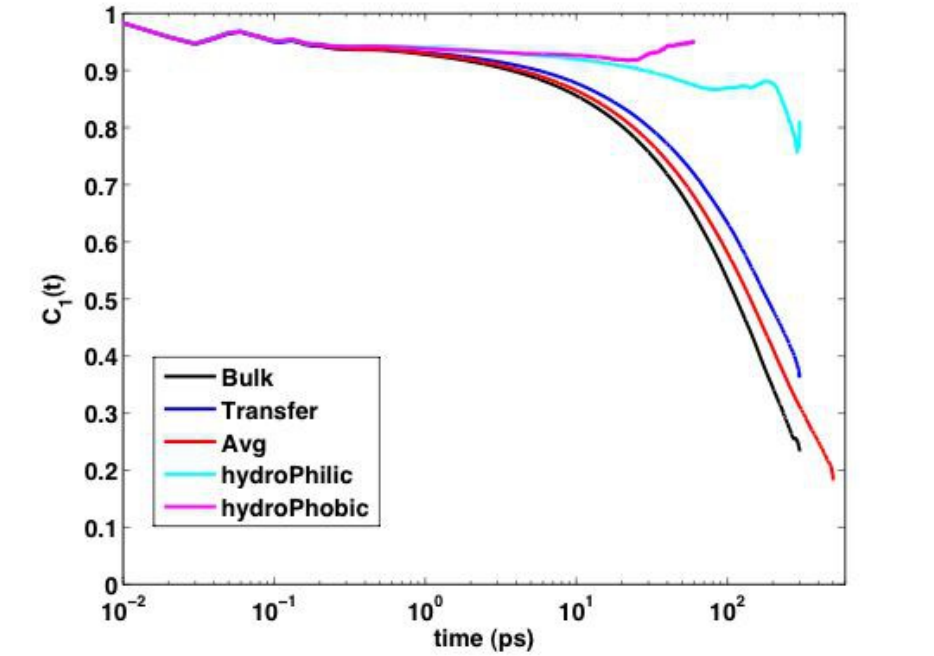


Figure 12. Johnson and co-workers

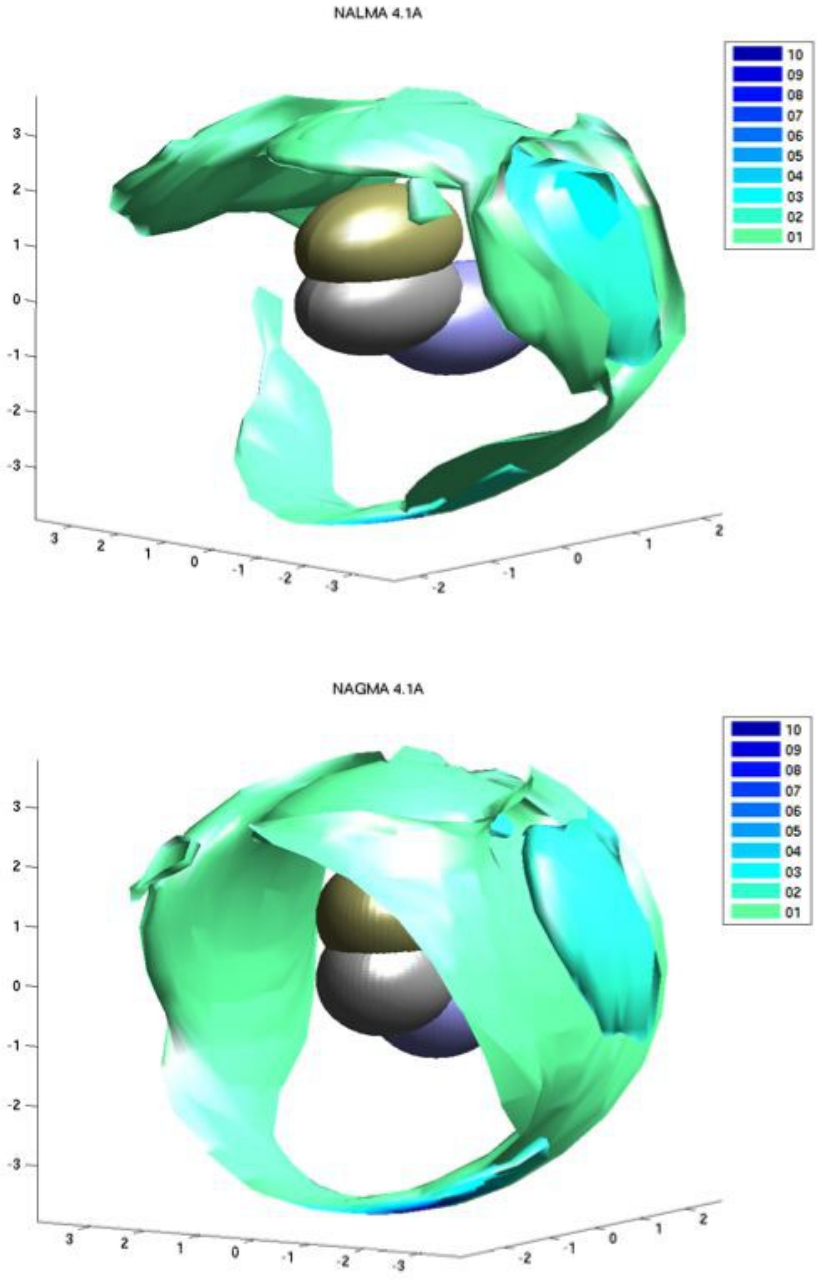


Figure 13. Johnson and co-workers

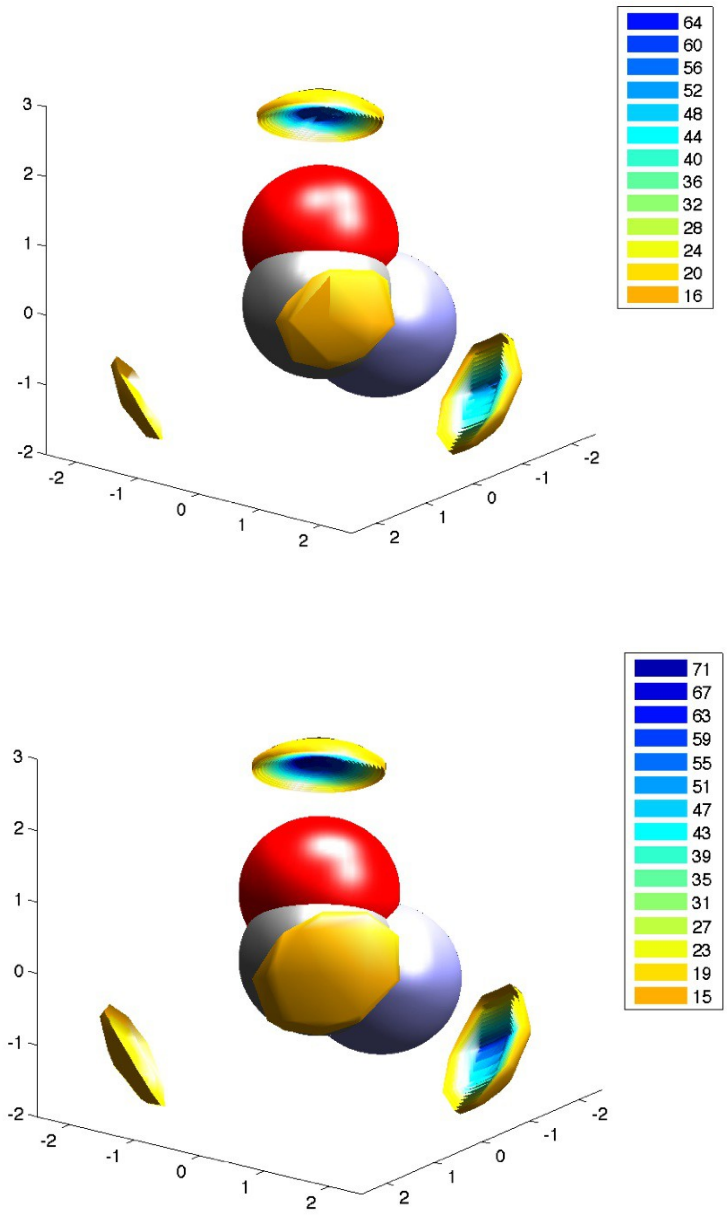


Figure 14. Johnson and co-workers

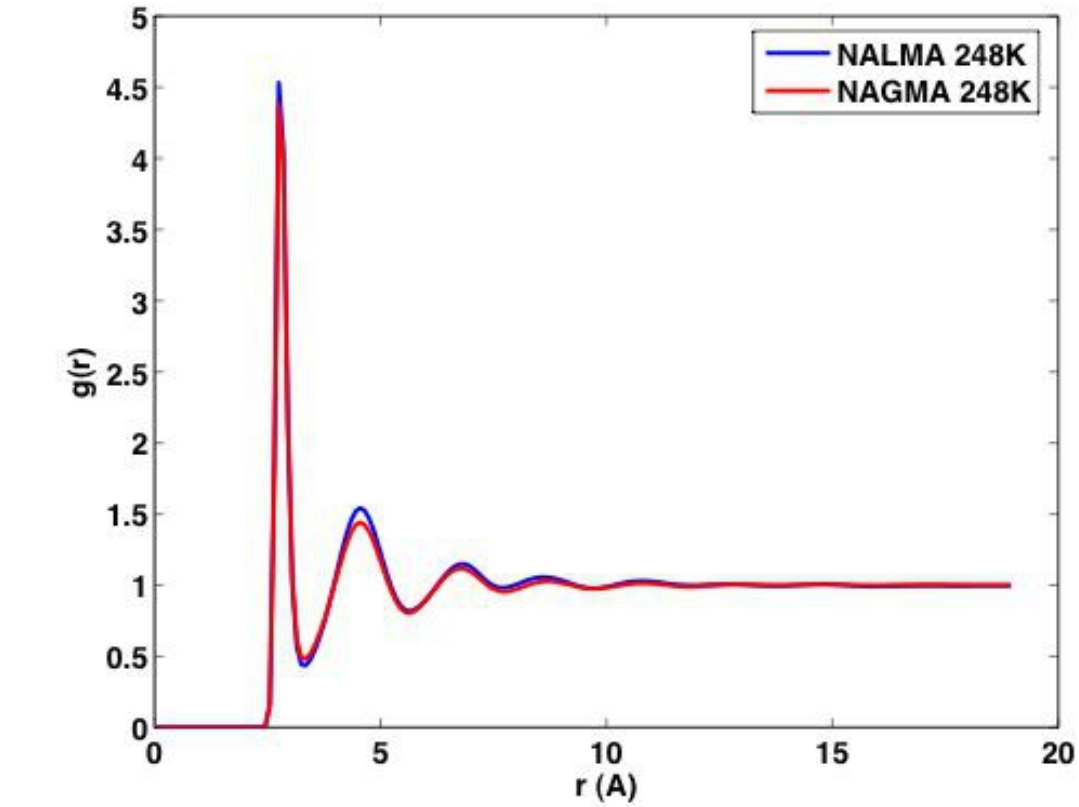


Figure 15. Johnson and co-workers