Spectral Graph Analyses of Water Hydrogen-Bonding Network and Osmolyte Aggregate Structures in Osmolyte–Water Solutions

Hochan Lee,‡,§ Jun-Ho Choi,‡,§ Pramod Kumar Verma,‡,§ and Minhaeng Cho∗,‡

†Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 136-701, Republic of Korea
‡Department of Chemistry, Korea University, Seoul 136-713, Republic of Korea
§Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 136-701, Republic of Korea

ABSTRACT: Recently, it was shown that the spectral graph theory is exceptionally useful for understanding not only morphological structural differences in ion aggregates but also similarities between an ion network and a water H-bonding network in highly concentrated salt solutions. Here, we present spectral graph analysis results on osmolyte aggregates and water H-bonding network structures in aqueous renal osmolyte solutions. The quantitative analyses of the adjacency matrices that are graph-theoretical representations of aggregates of osmolyte molecules and water H-bond structures provide the ensemble average eigenvalue spectra and degree distribution. We show that urea molecules form quite different morphological structures compared to other protecting renal osmolyte molecules in water, particularly sorbitol and trimethylglycine, which are well-known protecting osmolytes, and at high concentrations exhibit a strong propensity to form morphological structures that are graph-theoretically similar to that of the water H-bond network. Conversely, urea molecules, even at similarly high concentrations, form separated clusters instead of extended osmolyte–osmolyte networks. This difference in morphological structure of osmolyte–osmolyte aggregates between protecting and destabilizing osmolytes is considered to be an important observation that led us to propose a hypothesis on the osmolyte aggregate growth mechanism via either osmolyte network formation or segregated osmolyte cluster formation. We anticipate that the present spectral graph analyses of osmolyte aggregate structures and their interplay with the water H-bond network structure in highly concentrated renal osmolyte solutions could provide important information on the osmolyte effects of not only water structures but also protein stability in biologically relevant osmolyte solutions.

I. INTRODUCTION

Osmolytes are small organic molecules that are commonly found in all living beings.1 Their presence in the intracellular environment affects the solubility and stability of dissolved proteins. The effects depend on the type of osmolytes, namely, protecting and destabilizing osmolytes. The mammalian kidney accumulates five different protecting osmolytes (myo-inositol, sorbitol, taurine, trimethylglycine (TMG), and glycero phosphocholine) to counteract deleterious effects of destabilizing osmolyte urea.2 In addition, they also neutralize cellular stress, such as potentially harmful fluctuations in temperature and solution composition (salt concentrations, pH, etc.).3

The underlying mechanism of osmolyte operations is the key to understand how proteins in vivo maintain their stability and functionality. Despite considerable research efforts, experts differ in their opinion on the exact mechanism of osmolyte action.4–7 Some emphasize direct interaction of osmolyte molecules with protein backbone peptides or amino acid side chains,8–11 whereas others believe that osmolytes act indirectly by interacting with surrounding water molecules and subsequently modulating (weakening or strengthening) water H-bonding network structures and thermodynamic properties.14–18 Here, the latter hypothesis is the main subject of research in this paper.

As an experimental tool, a variety of vibrational spectroscopic methods have been used to investigate the structure and dynamics of water H-bonding structures in bulk water,19–31 salt solutions,32–38 water-miscible organic solvents,39–42 osmolyte solutions,43–45 and so forth.46 In particular, isotopically diluted water containing a small amount of HDO molecules has been studied quite extensively because the OD stretch vibration of HDO is an excellent IR probe.47 HDO is an isotopically diluted water solution that is isolated from the other OD groups such that there is a negligible chance of couplings between OD groups, making the interpretation of vibrational spectroscopic data comparatively straightforward.22,31,39,45,46,47 Bakker and co-workers performed IR pump–probe (IR PP) measurements of the OD bands of HDO in aqueous solutions of sorbitol and urea.33,46 They showed that the addition of sorbitol substantially slows the orientational mobility of water molecules, whereas urea does not strongly
affect water H-bond dynamics.\textsuperscript{53,46} Skinner and co-workers also studied urea–water mixtures and suggested that urea does not strongly perturb water structure.\textsuperscript{48}

The excellence of the OD stretch mode of HDO as an IR probe in various aqueous systems to monitor water H-bond dynamics is due to the capability of its parent HDO molecule to form multiple H-bonds with surrounding water molecules as well as other dissolved species (osmolyte, salt, protein, etc.). Consequently, its vibrational dynamics provides important information from HDO’s point of view on H-bond making and breaking processes and solvation dynamics. However, if only the OD stretch mode is used as an IR probe for linear and nonlinear IR spectroscopic studies, it is impossible to extract information on how a given osmolyte dictates water’s choice of making a H-bond with either another water or a third molecular component (e.g., protein or solute) in solution. Therefore, to have a stereoscopic and complementary view on osmolyte-induced changes in water structure, we recently employed two different IR probes that are the OD stretch mode of HDO and the azide asymmetric stretching vibration of hydrazoic acid (HN\(_3\)).\textsuperscript{49} In addition, we carried out MD simulations to calculate the O–O and O–H (between water molecules) radial distribution functions (RDFs) in different osmolyte solutions at various concentrations. There, we showed that the vibrational dynamics of the two different IR probes in aqueous urea solutions are significantly different from those in protecting osmolyte (e.g., sorbitol and TMG) solutions, which led us to suggest that the osmolyte aggregate structures could be important in understanding their effects on water structure.\textsuperscript{49−51}

Although the water–water RDFs and average H-bond number of water molecule in osmolyte solutions have provided useful information on how dissolved osmolyte molecules change water structure, they are local properties reflecting just the fleeting environment around individual water molecules. Therefore, to explore spatially large-scale structures of water H-bonding network and osmolyte aggregate, we have used the spectral graph theory method and examined various graph-theoretical properties, much like our previous works on spectral graph theory analyses of ion aggregates and specific ion effects on water structure in highly concentrated salt solutions.\textsuperscript{52}

In the present paper, we for the first time report the spectral graph analysis results on the water H-bonding network and osmolyte aggregate structures in highly concentrated renal osmolyte (Scheme 1) solutions, which in turn provides convincing evidence that protecting renal osmolytes with a strong propensity to form large-scale network-like aggregates significantly disrupts the water H-bonding network. Conversely, urea even at a fairly high concentration brings only a marginal effect on the water H-bonding network because they tend to form rather localized hydrated urea clusters instead of large-scale network structures. In section II, the results from MD simulations and statistical analyses are presented for five different osmolytes at various concentrations. The snapshot structures taken from MD trajectories are used to perform spectral graph analyses (section III). Finally, the main results are summarized in section IV with a few concluding remarks.

\textbf{II. MOLECULAR DYNAMICS SIMULATION RESULTS}

\textbf{A. MD Simulation Method.} For MD simulations, we considered the aqueous solutions of five different renal osmolytes (Scheme 1). The concentration along with the corresponding number of osmolyte molecules in each MD simulation are presented in Table 1. There are 1000 TIP\(_3\)P\textsuperscript{53} water molecules in a period box. The highest osmolyte concentrations studied are very close to their solubility limits, which are approximately 9, 5.5, 5, 0.86, and 0.80 M for urea, sorbitol, TMG, myo-inositol, and taurine, respectively. The GAFF\textsuperscript{54} (General Amber Force Field) parameters were used to describe osmolyte molecules, where the corresponding restrained electrostatic potential (RESP)\textsuperscript{55} charges were obtained from B3LYP/6-311++G(3df,2pd) calculations.\textsuperscript{56} The non-bonding interaction cutoff distance was assumed to be 10 Å, and the particle mesh Ewald method\textsuperscript{57,58} was used for long-range electrostatic interactions—note that we carried out an additional MD simulation of 8.5 M urea solution with 14 Å cutoff distance for nonbonding interactions and compared the calculated RDFs to those with a smaller cutoff distance (10 Å) but found no noticeable difference between them. The composite solution system was first energy-minimized with the steepest descent method and the conjugate gradient method prior to running equilibrium MD simulations. A constant \(N, p,\) and \(T\) ensemble simulation at \(p = 1\) atm and \(T = 298\) K was carried out for 10 ns to adjust the density of the composite system. An additional 10 ns constant \(N, V,\) and \(T\) simulation at 298 K was performed so that the osmolyte–water solution system was allowed to reach its thermal equilibrium state. We further ran a 20 ns equilibrium MD simulation to ensure that the solution with large osmolyte molecules having a number of internal degrees of freedom became fully equilibrated. Finally, the production run was performed for 10 ns at constant \(N, V,\) and \(T\) conditions, where the simulation time step was set to be 1 fs for MD trajectories, and atomic coordinates were saved for every 100 fs for subsequent statistical and graph theoretical analyses.

\textbf{B. Radial Distribution Functions between Osmolyte Molecules.} In Figure 1, the calculated site–site RDF between \(H_{\osm}\) (osmolyte H atom of OH or NH group) and \(O_{\osm}\) (osmolyte O atom) is shown for myo-inositol, sorbitol, taurine, and urea. Because of the absence of a H-bond donor group in TMG, RDF between \(O_{\osm}\) and \(N_{\osm}\) is instead shown in Figure 1D. Although the calculated RDFs in Figure 1 show notable concentration dependence, the first minimum position between the first and second solvation shells (peaks in RDFs) does not change much with respect to osmolyte concentration. This first minimum distance in RDFs is used as the criterion for determining whether a given pair of osmolyte molecules are connected or not via H-bonding (or electrostatic) interactions. In cases of myo-inositol and sorbitol solutions, the first minimum distance in the RDFs between \(H_{\osm}−(−OH)\) and \(O_{\osm}−(−OH)\) is 2.42 Å and that between \(H_{\osm}−(−NH_2)\) and \(O_{\osm}−(−CO)\) of urea is 2.65 Å. For taurine, the first minimum

\begin{scheme}
\caption{Molecular Structure of the Renal Osmolytes Studied in the Present Work}
\includegraphics[width=0.5\textwidth]{Scheme.png}
\end{scheme}
distance in the RDFs between $H_{\text{osm}}$ (−NH$_4^+$) and $O_{\text{osm}}$ (−SO$_3^-$) is 2.46 Å. For zwitterion TMG with no available H-bond donating sites, the first minimum distance between $O_{\text{osm}}$ (−COO$^-$) and $N_{\text{osm}}$ (−N(CH$_3$)$_3^+$) is found to be 5.62 Å. Finally, the criterion is set such that a given pair of osmolyte molecules form a dimer through electrostatic (or H-bonding) interactions if the distance between the two interacting sites is less than their first minimum distance in the corresponding RDF. In the spectral graph analyses of osmolyte aggregates, such electrostatic (or H-bonding) interaction between a pair of osmolyte molecules will be represented as an edge.52

C. Radial Distribution Functions between Water Molecules. The calculated water–water RDFs (between $O_w$−$H_w$ and $O_w$−$O_w$ where $O_w$ and $H_w$ mean water O and water H atoms, respectively) are shown in Figure 2. At a low concentration (~1M), the water–water ($O_w$−$H_w$, and $O_w$−$O_w$) RDFs of aqueous solutions of osmolytes are very similar to that of pure water, indicating that osmolytes at such a low concentration do not disrupt the water H-bond network structure, as expected. As the concentration rises, both the first and second peak heights of the $O_w$−$H_w$ (and the first peak of $O_w$−$O_w$) RDFs increase significantly in the case of sorbitol (see Figure 2A and D), whereas the increase is moderate in the case of TMG. Furthermore, the observation that the second water–shell peak (of $O_{w}$−$H_{w}$) increases greatly with respect to sorbitol concentration but marginal with TMG concentration indicates that sorbitol’s impact on water structures goes beyond one hydration layer. This long-range impact of sorbitol is due to its linear and flexible structure and its multiple H-bonding sites. Only a marginal increase in the second peak height of $O_{w}$−$H_{w}$ with increasing concentration of TMG suggests that TMG’s impact on water structure is mostly seen in the first hydration layer. Another observation unique to TMG is that the first peak position of the $O_{w}$−$H_{w}$ RDF shifts from 1.77 to 1.75 Å with the concentration increasing from 1.09 to 5.35 M. Such shortening of the water–water H-bond indicates that TMG causes a strong perturbation to local H-bond structure in its immediate vicinity compared to other renal osmolytes where no such shortening is observed. Here, the $O_w$−$H_w$ and $O_w$−$O_w$ RDFs of myo-inositol and taurine solutions are not presented because their solubility limits are less than 1 M and the global water structures are not strongly affected by the presence of these two osmolytes except for those water molecules in the first hydration layer.

Interestingly, urea that has very high solubility in water compared to other protecting osmolytes does not cause any considerable increase in the peak heights of the water–water RDF (Figure 2C and F). This suggests that urea perturbs the water structure weakly corroborating well with the work of Kokubo et al.59 They, based on their MD simulations and analysis of the water structures in urea solutions, showed that the water properties like average H-bond number and H-bond lifetime distribution both between urea and urea and between water and water are negligibly dependent on urea concentration.59

Table 1. Concentration and Number of Molecules of Osmolytes in MD Simulation

<table>
<thead>
<tr>
<th>Osmolyte</th>
<th>Concentration M (m)</th>
<th>Number of Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>myo-inositol</td>
<td>0.43 (0.44)</td>
<td>8</td>
</tr>
<tr>
<td>myo-inositol</td>
<td>0.88 (0.94)</td>
<td>17</td>
</tr>
<tr>
<td>taurine</td>
<td>0.39 (0.39)</td>
<td>7</td>
</tr>
<tr>
<td>taurine</td>
<td>0.86 (0.89)</td>
<td>16</td>
</tr>
<tr>
<td>TMG</td>
<td>0.54 (0.56)</td>
<td>10</td>
</tr>
<tr>
<td>TMG</td>
<td>1.09 (1.17)</td>
<td>21</td>
</tr>
<tr>
<td>TMG</td>
<td>2.10 (2.50)</td>
<td>45</td>
</tr>
<tr>
<td>TMG</td>
<td>3.28 (4.44)</td>
<td>80</td>
</tr>
<tr>
<td>TMG</td>
<td>5.04 (8.55)</td>
<td>154</td>
</tr>
<tr>
<td>TMG</td>
<td>5.35 (9.44)</td>
<td>170</td>
</tr>
<tr>
<td>sorbitol</td>
<td>0.53 (0.56)</td>
<td>10</td>
</tr>
<tr>
<td>sorbitol</td>
<td>1.00 (1.11)</td>
<td>20</td>
</tr>
<tr>
<td>sorbitol</td>
<td>2.04 (2.66)</td>
<td>48</td>
</tr>
<tr>
<td>sorbitol</td>
<td>3.07 (4.83)</td>
<td>87</td>
</tr>
<tr>
<td>sorbitol</td>
<td>4.08 (7.99)</td>
<td>144</td>
</tr>
<tr>
<td>sorbitol</td>
<td>4.98 (12.71)</td>
<td>229</td>
</tr>
<tr>
<td>sorbitol</td>
<td>0.50 (0.50)</td>
<td>9</td>
</tr>
<tr>
<td>urea</td>
<td>1.03 (1.05)</td>
<td>19</td>
</tr>
<tr>
<td>urea</td>
<td>2.09 (2.22)</td>
<td>40</td>
</tr>
<tr>
<td>urea</td>
<td>3.11 (3.44)</td>
<td>62</td>
</tr>
<tr>
<td>urea</td>
<td>4.19 (4.88)</td>
<td>88</td>
</tr>
<tr>
<td>urea</td>
<td>5.08 (6.16)</td>
<td>111</td>
</tr>
<tr>
<td>urea</td>
<td>6.07 (7.83)</td>
<td>141</td>
</tr>
<tr>
<td>urea</td>
<td>7.11 (9.66)</td>
<td>174</td>
</tr>
<tr>
<td>urea</td>
<td>7.82 (11.05)</td>
<td>199</td>
</tr>
<tr>
<td>urea</td>
<td>8.52 (12.49)</td>
<td>225</td>
</tr>
</tbody>
</table>

Figure 1. Osmolyte–osmolyte RDFs in taurine, myo-inositol, sorbitol, TMG, and urea solutions.

Figure 2. Water–water RDFs in sorbitol, TMG, and urea solutions. The upper panels correspond to the RDFs of $O_{w}$–$H_{w}$ and the lower ones represent those of $O_{w}$–$O_{w}$.
In Figure 3, the snapshot structures of osmolyte (upper panels) and water (lower panels) molecules in aqueous sorbitol (4.98 M), TMG (5.35 M), and urea (8.52 M) solutions are shown—note that these are aqueous solutions at concentrations close to their solubility limits. Although the number of sorbitol molecules in a given simulation box (229) is close to that of urea (225), the osmolyte aggregate structures in the two solutions (compare Figure 3A and C) are distinctly different. Note that sorbitol molecules form large-scale networks, whereas urea molecules prefer to form clusters separated from one another. Sorbitol being a linear and flexible molecule with six hydroxyl groups is well-capable of forming H-bonds with both water and other sorbitol molecules. Because of such a strong ability to form network-like aggregate structures by sorbitol molecules, water molecules in the highly concentrated sorbitol solutions form clusters within the cages of sorbitol networks (see Figure 3D). The water clusters are stabilized by not only water−water but also water−sorbitol H-bonds. In fact, our recent IR PP studies have shown that a considerable population of water−water H-bonding structures exists in highly concentrated sorbitol solution (5 M).\footnote{Moreover, our experimental and MD simulation results conroborate with the recent variable-temperature neutron scattering study on sorbitol−water mixtures, which revealed clustering behavior of water molecules that are surrounded by (and interacting with) sorbitol molecules.\cite{60}} Conversely, the aggregation of TMG molecules does not seem to make a profound impact on the water structure as compared to sorbitol. This again shows that the water H-bonding network is perturbed locally in the vicinity of TMG molecules.

As seen in Figure 3C, urea molecules form aggregates too, but most of the aggregates are dimers. These small aggregates or dimers are well-hydrated and remain separated from other urea aggregates. The water H-bonding network structure in urea−water binary mixture solutions remains similar to that in pure water, which is consistent with our recent time- and frequency-resolved IR PP studies with dual IR probes approach (both OD and azide stretch) and other previous works.\footnote{However, the water H-bond number distribution is affected by the presence of urea, as shown in Figure 4A.}

**D. Water H-bond Number Distribution.** To have a quantitative analysis of water H-bonding structures with respect to osmolyte concentration, we calculated the ensemble average H-bond numbers of a single water molecule in aqueous osmolyte solutions (see ref \footnote{Moreover, our experimental and MD simulation results conroborate with the recent variable-temperature neutron scattering study on sorbitol−water mixtures, which revealed clustering behavior of water molecules that are surrounded by (and interacting with) sorbitol molecules.\cite{60}} for the distance criterion of a water−water H-bond), which are plotted in Figure 4A. Note that the average H-bond number of a single water molecule in pure TIP3P water is 3.68.\footnote{Moreover, our experimental and MD simulation results conroborate with the recent variable-temperature neutron scattering study on sorbitol−water mixtures, which revealed clustering behavior of water molecules that are surrounded by (and interacting with) sorbitol molecules.\cite{60}} The osmolytes make the H-bond number decrease with increasing concentration, but the extent of decrease depends on the nature of the osmolyte. The protecting osmolytes studied here cause a significant decrease in the H-bond number (a considerable disruption of water H-bond structure), whereas urea does not. This again confirms that urea does not alter water’s intrinsic propensity to form a H-bond network.\footnote{Moreover, our experimental and MD simulation results conroborate with the recent variable-temperature neutron scattering study on sorbitol−water mixtures, which revealed clustering behavior of water molecules that are surrounded by (and interacting with) sorbitol molecules.\cite{60}} In Figure 4B, the average H-bond lengths (\(O_{w}−H_{w}\)) are plotted with respect to osmolyte concentration. Interestingly, the slope of the broken line (connecting the blue

---

**Figure 3.** Snapshots corresponding to aqueous solutions of sorbitol (4.98 M), TMG (5.35 M), and urea (8.52 M). In the upper panels, snapshot structures of osmolyte molecules within the periodic box are shown, whereas those of water molecules are shown in the lower panels. Each configuration is taken from MD trajectories, and the width of the slab is 5 Å. Here, H, O, C, and N atoms are represented by white, red, green, and blue colored spheres, respectively. The van der Waals volume ratio between osmolyte aggregates and the composite of osmolyte−water is inserted with osmolyte concentration in parentheses.

**Figure 4.** Ensemble average H-bond number and H-bond length plotted against osmolyte concentration (mol/L).
In graph theory, the object under consideration is represented by a set of vertices \( V \) and a set of edges \( E \) connecting vertices.\(^{67-69} \) Two vertices \( v_i \) and \( v_j \) of a given graph \( G = G(V,E) \) are adjacent if they are connected by the edge \( e_{ij} \). The degree \( d_j \) of the \( j \)th vertex corresponds to the number of adjacent vertices. One of the most successful ways to numerically analyze graphs is to represent a given graph in the form of a matrix. Because we are interested in connectivity patterns between osmolyte molecules and between water molecules, we consider the adjacency matrix

\[
A = A(G),
\]

where \( A \) is an \( N \times N \) symmetric matrix with \( N \) being the number of vertices (osmolyte or water molecules). Because the electrostatic (or H-bonding) interactions between osmolyte molecules as well as between water molecules do not have any directional property, the adjacency matrices considered in this work are all undirected, and the adjacency matrix elements \( A_{ij} \) (for \( i \neq j \)) are given as

\[
\begin{align*}
A_{ij} &= 1, \text{if } i \text{ and } j \text{ are connected by an edge}, \\
A_{ij} &= 0, \text{if } i \text{ and } j \text{ are not adjacent}, \\
A_{ii} &= 0, \text{for all } i \text{'s}.
\end{align*}
\]

Here, the H-bond between water molecules and the osmolyte–osmolyte connection via either intermolecular H-bonding or electrostatic interaction are represented by an edge of water H-bond graph or osmolyte aggregate graph. In this work, even though the strengths of water–water H-bonding interaction and osmolyte–osmolyte interaction could have distributions, we treat all the edges equally so that the graphs considered are nonweighted. The constructed adjacency matrix for a given undirected graph is diagonalized to obtain the characteristic values and eigenvectors.

**B. Spectral Graph Analysis of Osmolyte Aggregates.**

An application of spectral graph theory begins with characteristic value (eigenvalue) analysis of the adjacency matrix representing either osmolyte–osmolyte aggregate or the water–water H-bonding network. We considered 100,000 snapshot configurations taken from MD trajectories. In the cases of water H-bonding networks, each adjacency matrix is of size \( 1000 \times 1000 \) symmetric matrix. The size of the adjacency matrix graph-theoretically representing osmolyte aggregate is of course determined by the number of osmolyte molecules. The resulting eigenvalues and eigenvectors of the \( A \)-matrix therefore provide graph-theoretical properties of the relevant network structures.

The graph spectrum, which is the normalized and ensemble-averaged eigenvalue distribution, is plotted in Figure 5. If all the osmolyte molecules in solutions are not connected and isolated from all the others, the eigenvalues would all be zero. For a dimer with two vertices and one edge, the eigenvalues are 1 and −1. If osmolyte molecules begin to form large aggregates with extended network formation, the eigenvalue spectrum would become broad and continuous. First, consider the low concentration solutions of sorbitol (1 M), TMG (1.09 M), taurine (0.86 M), and myo-inositol (0.88 M), which are shown in Figure 5A, D, L, and K, respectively. Interestingly, the eigenvalue spectrum of sorbitol is particularly broad, and the populations of eigenvalues other than ±1 (dimers) are quite large. The other protecting osmolyte solutions like TMG, myo-inositol, and taurine exhibit similarly broad and rather continuous distributions of eigenvalues. However, the graph

---

**Figure 5.** Normalized eigenvalue spectra of the adjacency matrices of osmolyte aggregates in five different osmolyte solutions at various concentrations. The ratio between solute volume and whole volume constituted of water and osmolyte molecules is described in parentheses. For the volume ratio of 0.15 (sorbitol), 0.12 (TMG), and 0.14 (urea), the eigenvalue spectra for sorbitol exhibits higher density compared to other two cases of TMG and urea, representing the formation of sorbitol aggregates.

---

In Figure 4B) associated with TMG is relatively steep, but the H-bond lengths between water molecules in sorbitol and urea solutions do not change much compared to TMG. The reason is that the water molecules in the immediate vicinity (first hydration shell) of TMG are tightly packed due to stronger H-bonding interaction via assistance from TMG−water interaction. Although the MD and statistical analysis results presented and discussed throughout this section provide important information on local water H-bond structure, large-scale water structures as well as osmolyte aggregate structures can further be extracted from spectral graph analyses\(^{62-66} \) as presented in the following section.
spectrum of urea at similarly low concentration (1.03 M) (see Figure 5G) appears to differ from the graph spectra of protecting osmolytes at similar concentration. Specifically, despite the fact that the urea has multiple H-bonding sites, it exists preferentially as monomer or dimer, which is manifested by the strong peaks at $±1$ and 0 in the corresponding spectrum (Figure 5G). In fact, from the graph spectra of 3.11 and 5.08 M urea solutions, we believe that urea molecules prefer to be hydrated by water molecules instead of forming large-scale urea–urea aggregates and hence perturb the water H-bond structure weakly. Our observation of weak aggregation behaviors of urea not only corroborates well with the suggestion provided by Weerasinghe et al. (based on the excess coordination numbers and RDF) but also provide further insight into its network property and aggregation behavior. Another comparison (Figure 5A, D, and H, see further insight into its network property and aggregation excess coordination numbers and RDF) but also provide suggestion provided by Weerasinghe et al. (based on the structure weakly. Our observation of weak aggregation behavior of urea at a given concentration, whereas urea exhibits quite different distributions compared to those of protecting osmolytes.

As the concentration of the protecting osmolyte increases, the graph spectra becomes quite broad and continuous (Figure 5B, C, E, and F), indicating that the size of osmolyte aggregate molecules has a broad distribution. Such considerable change in eigenvalue distributions (graph spectra) means that protecting osmolytes (sorbitol, TMG) form osmolyte–osmolyte aggregates via network formation, and the spatial extent of the sorbitol–sorbitol network is found to be the largest among the studied renal osmolytes. The aggregation behavior of the osmolytes is not a completely new observation. For example, trehalose, a well-known osmolyte, shows a strong tendency to self-associate based on vapor pressure osmometry and MD simulation studies. Such an aggregation tendency is also found in sucrose to some extent. Sorbitol, much like trehalose, has multiple hydroxyl groups that are readily available for several H-bonding interactions with water as well as other neighboring sorbitol molecules. Therefore, similar to trehalose, our simulation results suggest that sorbitol molecules form spatially extended network structures at high concentrations, and for a denaturant like urea, only at an extremely high concentration (8.52 M); the graph spectrum (Figure 5J) appears to be similar to that (Figure 5E) of 3.28 M TMG solution and that (Figure 5A) of 1 M sorbitol solution. From the present graph spectral analyses, we have found that the intrinsic propensity of forming large osmolyte aggregates depends on (i) its size, (ii) number of H-bonding or electrostatic interaction sites, and (iii) detailed chemical structure (flexibility (linear versus cyclic form), compactness, and compatibility with the water molecule).

One of the easy to understand (intuitively) and important properties of a graph is its degree (number of adjacent vertices) distribution because it provides information on size distribution of subgraphs and connectivity pattern. In Figure 6, the degree distributions of five different osmolytes at various concentrations are plotted. At low concentrations ($≤1$ M), the population of zero degree vertex is always the highest, as expected, which just means that osmolyte molecules remain fully hydrated without making considerable osmolyte–osmolyte aggregates. As osmolyte concentration increases, the peak position of the degree distribution shifts toward higher values. The magnitude of shift is largest for sorbitol and TMG. Specifically, the peak position shifts to a value of 3 for approximately 5 M concentration of both sorbitol and TMG, meaning that each osmolyte molecule has about three adjacent osmolyte molecules. Sorbitol does not only induce the largest shift in the peak position of the degree but also makes the distribution broad (compare left triangle (symbol) lines in Figure 6C and D). Sorbitol is a linear (flexible due to multiple internal rotational degrees of freedom) molecule with six hydroxyl groups, whereas zwitterionic TMG has a bulky trimethyl group with only two oxygen atoms (Scheme 1). What is interesting is that the denaturing osmolyte urea shows little dependence of degree distribution on its concentration. Up to 5 M concentration urea follows a monotonically decaying pattern similar to those of protecting osmolytes at relatively much lower concentrations (~1 M) (compare the left triangle (symbol) line in Figure 6E with the circle (symbol) line in Figures 6A–D). Only at concentrations significantly higher than 5.0 M urea molecules begin to form larger aggregates beyond dimer states.

To make a comparison among different osmolytes at similar concentrations, we have plotted the degree distributions (osmolyte–osmolyte) of sorbitol, TMG, and urea all together in Figure 7A, which clearly shows how these three osmolytes have different connectivity patterns. From the degree distributions, we additionally calculated the average degrees and have plotted them in Figure 7B. Again, the increasing rate depends on the nature of the osmolyte. At a given concentration, the average degree for sorbitol is always the highest among all the studied osmolytes. For instance, the average degree values for osmolyte–water solutions of 4.98 M sorbitol, 5.35 M TMG, and 8.35 M urea solutions are 3.62, 3.01, and 1.47, respectively. The values of 3.62 and 3.01 for sorbitol and TMG mean that they can form ~4 H-bonds with neighboring sorbitol molecules or electrostatic interactions with ~3 TMG molecules, respectively. An individual water molecule in pure water also forms 3 to 4 H-bonds with other water molecules, and this H-bond number decreases upon the addition of osmolytes, including urea (Figure 7D). For direct comparisons, we have also plotted the H-bond degree distribution of water in pure water in both Figure 7A and C (open circle). Upon comparing the degree distributions of water in sorbitol (4.98 M), TMG (5.04 M), urea (5.08 M), and
pure water (Figure 7C), there appears minor difference in the degree distribution of water in urea and pure water (diamond and open circle in Figure 7C), but both sorbitol and TMG have distinctively low degree distribution compared to pure water’s degree distribution. This shift of degree distribution of water for protecting osmolytes (sorbitol and TMG) is mainly due to their ability to form multiple interactions with neighboring osmolyte molecules in an extended fashion (network type) that distinguishes them from urea.

In general, the MD simulation results are to be compared with experimentally available results to validate a given set of force field parameters. In a recent MD simulation study on urea–peptide interactions in aqueous solutions by Garcia and co-workers, not only were the physical properties associated with osmotic pressure based on two kinds of AMBER force fields calculated but the obtained results were also directly compared with the experimentally measured data to investigate water–water, osmolyte–osmolyte, osmolyte–peptide, and water–peptide interactions. Both of the employed AMBER force fields were reasonably good at reproducing the physical properties associated with osmotic pressure. In 2003, Weerasinghe and Smith developed a novel nonpolarizable force field and used it to study a urea–water binary mixture and to calculate the Kirkwood–Buff (KB) integrals for water–water, water–urea, and urea–urea, which tallied well with the experimental data. They also calculated excess coordination numbers representing an excess of species in the vicinity of a given species and showed no significant aggregation of urea molecules even at very high concentration (8M). The two works mentioned above did use different force fields to study urea–water binary mixtures, and their simulated data correlated well with the experimental observables. However, none of them found any strong evidence for the self-aggregation of urea. This is in line with the outcome of our present study. Nevertheless, it will be of great interest to further investigate the force field dependence of other osmolytes propensity for forming large-scale aggregates in water by using different molecular mechanical parameters for osmolyte and water molecules.

C. Spectral Graph Analysis of Water H-Bond Structure. In this subsection, we address the question about how the water H-bonding network structure changes in response to osmolytes aggregate formation, employing the spectral graph analysis method. From the spectral graph analyses of the water H-bonding network in pure water, we found that the eigenvalue distribution (graph spectrum) is broad (in a range from −4 to 4) and continuous. In Figure 8, the normalized eigenvalue spectra of water H-bonding networks in osmolyte solutions at various concentrations are plotted. The water graph spectra mostly appear to be featureless and continuous regardless of the nature and concentration of osmolyte. However, there are some noticeable differences, especially in the number of discrete peaks. Panels C, F, and I in Figure 8 represent three different osmolytes at a similar concentration (∼5 M), but the number of discrete peaks is larger in sorbitol and minimal in urea, meaning sorbitol
considerably disrupts water H-bonding structure but that urea causes only minor perturbation to the water H-bond network.

Now, let us compare the water graph spectra of sorbitol and TMG (Figure 8) with their osmolyte aggregate spectra as depicted in Figure 5C and F. The spectral similarity, called isospectral property, of these graph spectra is important evidence of the formation of network-type aggregates by the protecting osmolytes (sorbitol and TMG) that can be viewed as three-dimensional networks, similar to water H-bonding networks. Urea, however, does not show such a tendency to form large interconnected network structures at concentrations up to 5 M (compare Figure 5G−I with Figure 8G−I).

In Figure 9, the water degree distributions for all the studied osmolyte solutions are plotted. Overall, the degree distribution shifts to a lower value upon increasing osmolyte concentration, but the magnitude of the shift depends on the nature of the osmolyte. Urea causes only a marginal shift in the peak position of the degree distribution of water H-bond structure (see Figure 9E) upon increasing concentration from 1 to 5 M, whereas protecting osmolytes result in a considerable shift (see arrow in Figure 9). These shifts further confirm that protecting osmolytes like sorbitol and TMG significantly disrupt the water H-bond network due to their tendency to form large-scale network-like aggregates whereas urea does not.

D. Comparison between Theory and Experiment.

Recently, we carried out fs IR PP measurements of the OD stretch mode of HDO and the azide stretch mode of HN$_3$ in these renal osmolyte solutions to understand the osmolyte effects on water structure. The IR probe, OD stretch mode, provides crucial information on water structure from the water’s point of view, whereas the other IR probe, azide stretch mode, allows us to understand the change in water structure from dissolved solute’s point of view. All the protecting osmolytes (sorbitol, TMG, myo-inositol, and taurine) showed very strong concentration dependence of vibrational frequencies and rotational motions of HDO and HN$_3$. In contrast, urea did not cause any considerable changes in either the vibrational frequencies of OD and azide stretch modes or their vibrational lifetimes or even rotational motions. Our IR PP studies suggested that TMG is very effective at disrupting water H-bond structure but only in its immediate vicinity, whereas sorbitol has a long-range and diffusive impact on the water H-bonding network. In contrast, urea is neutral, neither weakening nor strengthening the H-bonding network-forming ability of water nor affecting the solute (HN$_3$)−water partnership. Although the focus of our previous article was to probe the detailed changes in the local water structure and dynamics due to dissolved osmolyte molecules, the focus of the present article is to quantify the global changes in these osmolyte−water solution structures with particular attention to morphological differences (Figure 3). In Figure 10, we now present representative structures where the direct H-bonds between osmolyte and water are shown in red and indirect H-bonds between osmolyte and water with a water molecule acting as a bridge between them is shown in black. Both sorbitol and TMG can be seen to hinder tetrahedral packing of surrounding water molecules, whereas urea positions itself in such a way as to preserve the tetrahedral arrangements of the surrounding water molecules.

Before we close this section, it is worth discussing the correlation between the orientational relaxation times (τ$_{\text{rot}}$) of OD in osmolyte solutions (data taken from ref 49) and the average H-bond numbers (shown in Figure 4A) in the same osmolyte solutions (see Figure 11A). The water reorientation involves two independent pathways according to the jump model that has been successfully used to describe the
orientational relaxation processes of water molecules in the bulk around halide ions and hydrophobic solutes, reproducing both the experimental and stimulated orientation relaxation times. The first major one is exchange of H-bond acceptors via a special bifurcated H-bonded configuration in which a water OH bond executes a large-amplitude angular jump from its former H-bond partner to a new H-bond acceptor. The second but minor contribution comes from the slower diffusive reorientation of the intact H-bond axis between successive jumps. Because the rate-limiting step is in the jump mechanism, it is expected that there is a correlation between the average H-bond number and τrot. This is indeed the case, as can be seen in Figure 11A. Protecting osmolytes (like sorbitol and TMG) reduce the average H-bond number considerably and, as such, the availability of the incoming H-bond acceptor is also decreased, which slows the water rotation. Interestingly, the τrot (OD) value is almost linearly correlated with H-bond number. Conversely, urea lies on a horizontal line in this correlation graph (Figure 11A) that again justifies its ideal compatibility with water.

We have also plotted τrot (HNO3) against H-bond number in Figure 11B. There seems to also be a linear dependence of τrot (HNO3) on the H-bond number for all three osmolytes irrespective of their protecting or denaturing nature. Because this linear correlation is sensitive only to the concentration but not the type of osmolytes, the rotational motion of HNO3 is independent of osmolyte type, unlike its vibrational peak frequency position or vibrational lifetime. The reason for concentration-dependent behavior of τrot (HNO3) irrespective of osmolyte type might be due to the larger contribution of the slower diffusive reorientation (viscosity) of HNO3.

IV. SUMMARY AND A FEW CONCLUDING REMARKS

Spectral graph theory, which has been used to describe properties of a graph by analyzing eigenvalues and eigenvectors of various matrices representing the graph, is shown to be useful for identifying differences in morphological structures of osmolyte aggregates as well as water H-bond structures in highly concentrated osmolyte solutions. The protecting renal osmolytes, e.g., sorbitol and TMG, form large-scale network-like aggregates at high concentrations, whereas the destabilizing osmolyte urea fits into water’s tetrahedral H-bonding structure and forms cluster-like aggregates in high urea solutions. Despite prolonged efforts to understand osmolyte effects on water structure, we believe that the present work is the first attempt showing an interesting interplay between morphological structure of osmolyte aggregate and water H-bonding network structure in high osmolyte solutions. A speculative but far-reaching implication of this result is that the intrinsic propensity of an osmolyte to form large-scale aggregate structures determines the thermodynamic stability of the protein in such high osmolyte solutions found in mammalian kidneys. Currently, this is under investigation by using the time-resolved vibrational spectroscopic measurement method.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b08029.

Snapshot structures based on a similar volume percentage of osmolytes (PDF)


(67) Deo, N. *Graph Theory with Applications to Engineering and Computer Science*; 2nd ed.; Prentice Hall of India Private Limited: New Delhi, 1984.


